



Учреждение Российской академии наук
Институт космических исследований РАН
(ИКИ РАН)

ТЕХНИЧЕСКОЕ ЗРЕНИЕ В СИСТЕМАХ УПРАВЛЕНИЯ МОБИЛЬНЫМИ ОБЪЕКТАМИ-2010

Выпуск 4

Под редакцией Р. Р. Назирова

Таруса

16–18 марта 2010 года

УНИВЕРСИТЕТ
КНИЖНЫЙ ДОМ

Москва
2011

УДК [004.896:681.5](063)
ББК 32.816я431(2Рос)
Т38

Т38 **Техническое зрение в системах управления мобильными объектами-2010: Труды научно-технической конференции-семинара.** Вып. 4 / Под ред. Р. Р. Назирова.— М. : КДУ, 2011.— 328 с. : табл., ил., цв. ил.

ISBN 978-5-98227-794-7

Настоящий сборник содержит материалы научно-технической конференции-семинара «Техническое зрение в системах управления мобильными объектами», организатором которой является Учреждение Российской академии наук Институт космических исследований РАН. В её подготовке и проведении участвовали специалисты ведущих организаций, отраслевых и академических научно-исследовательских институтов, а также высшей школы. На конференции-семинаре были представлены доклады по основным областям применения технического зрения в системах управления мобильными объектами, а также вопросам разработки таких систем: авиационно-космические приложения систем технического зрения; системы технического зрения в наземных и подводных роботах; программно-аппаратное обеспечение систем технического зрения. В сборник включены пленарные и секционные доклады конференции-семинара.

Конференция проходила 16–18 марта 2010 г. в г. Таруса на базе гостиницы «Интеркосмос» ИКИ РАН. Предыдущие выпуски трудов семинаров в Тарусе: Труды семинара по вычислительным технологиям в естественных науках : Вычислительная физика. Вып. 1. / Под ред. Р. Р. Назирова. М.: КДУ, 2009. 288 с.; Методы нелинейного анализа в кардиологии и онкологии : Физические подходы и клиническая практика. Вып. 2 / Под ред. Р. Р. Назирова. М.: КДУ, 2010. 206 с.; Труды семинара по вычислительным технологиям в естественных науках : Вычислительная химия и биология. Вып. 3 / Под ред. Р. Р. Назирова, Л. Н. Шура. М.: КДУ, 2010. 118 с.

Редакционная коллегия: *Желтов С. Ю., Визильтер Ю. В., Платонов А. К., Соколов С. М., Кротонов А. Н., Носков В. П., Мещеряков А. Ю., Гришин В. А.*

Редактор: *Корниленко В. С.*

Компьютерная верстка: *Комарова Н. Ю.*

COMPUTER VISION IN CONTROL SYSTEMS OF THE MOBILE OBJECTS PROCEEDINGS OF THE CONFERENCE-WORKSHOP-2010

Ed. R. R. Nazirov, 16–18 March, 2010

These are the Proceedings of the scientific and technological conference-workshop “Computer vision in control systems of the mobile objects”. The conference was organized by the Space Research Institute of the Russian Academy of Sciences. Experts from the leading organizations both the industrial and academic research institutes as well as universities take part in preparation and participate in the conference-workshop. The reports, which are presented in the conference-workshop, reflect the major applications of computer vision in control systems of mobile objects as well as the development of such systems: aerospace applications of computer vision systems; computer vision systems in terrestrial and underwater robots; hardware and software of the computer vision systems. The proceedings include plenary and panel reports at the conference-workshop.

УДК [004.896:681.5](063)
ББК 32.816я431(2Рос)

Отдельные материалы даны в авторской редакции.

ISBN 978-5-98227-794-7

© Учреждение Российской академии наук
Институт космических исследований (ИКИ) РАН, 2011
© Оформление. Обложка. Издательство «КДУ», 2011

Содержание

Предисловие	5
-------------------	---

АВИАЦИОННО-КОСМИЧЕСКИЕ ПРИЛОЖЕНИЯ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ

<i>Визильтер Ю. В., Желтов С. Ю.</i> Проблемы технического зрения в современных авиационных системах	11
<i>Костяшкин Л. Н., Бабаев С. И., Логинов А. А., Павлов О. В.</i> Технологии систем улучшенного / синтезированного зрения для управления летательными аппаратами	45
<i>Алпатов Б. А., Бабаян П. В., Костяшкин Л. Н., Романов Ю. Н.</i> Обработка и анализ изображений в бортовых оптико-электронных системах ..	57
<i>Белоглазов И. Н.</i> Теоретические аспекты и приложения стереоскопических систем навигации, наведения и дистанционного зондирования местности	63
<i>Блохинов Ю. Б., Гнилицкий В. В., Инсаров В. В., Чернявский А. С.</i> Алгоритм анализа и принятия решения в задаче селекции объектов на изображениях наземных сцен	92
<i>Васильев Д. В.</i> Фрагменты прикладной теории систем технического зрения для беспилотных летательных аппаратов	109
<i>Блажевич С. В., Винтаев В. Н., Ушакова Н. Н.</i> Автоматическое выделение и сопровождение псевдоточечных мерцающих объектов в реальном времени — модель фоноцелевой обстановки, алгоритм работы бортового процессора и технические решения по процессору	132
<i>Соколов С. М., Богуславский А. А.</i> Система информационного обеспечения задач сближения, стыковки, посадки космического аппарата на основе компьютерного видения	140
<i>Аванесов Г. А., Жуков Б. С., Краснопевцева Е. Б., Форш А. А.</i> Задачи бортовой обработки информации телевизионной системы навигации и наблюдения при посадке на Фобос	157
<i>Гришин В. А.</i> Анализ влияния динамики космического аппарата на характеристики алгоритмов обработки изображений системы технического зрения проекта ФОБОС-ГРУНТ	166
<i>Веселов Ю. Г., Карпиков И. В.</i> К вопросу оценки корреляционных характеристик шума цифровых оптико- электронных систем видимого диапазона	179

СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ В НАЗЕМНЫХ И ПОДВОДНЫХ РОБОТАХ

<i>Михайлов Б. Б.</i> Техническое зрение мобильных роботов	191
<i>Выголов О. В., Желтов С. Ю., Визильтер Ю. В.</i> Обнаружение препятствий перед наземным мобильным объектом в бортовой системе технического стереозрения реального времени	202
<i>Овчинников А. М., Платонов А. К.</i> Гиперспектральная система зрения для мобильных роботов	216
<i>Артюхов М. Ю., Кропотов А. Н., Макашов А. А., Сахарова Е. И.</i> Опыт создания системы локальной видеонавигации для подводных аппаратов	230

ПРОГРАММНО-АППАРАТНОЕ ОБЕСПЕЧЕНИЕ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ

<i>Жданов А. А.</i> Биологически инспирированное техническое зрение в системах автономного искусственного интеллекта	245
<i>Кий К. И.</i> Автоматическая система реального времени для обнаружения объектов и ориентиров на изображении, основанная на обработке цветных изображений	268
<i>Стрельников К. Н.</i> Использование особенностей современного представления цифровых видеоданных для оптимизации алгоритмов обработки	277
<i>Пытьев Ю. П., Чуличков А. И.</i> Морфологические методы анализа сцен по их изображениям	282
<i>Ланге М. М., Новиков Н. А.</i> Сегментация изображений на основе описаний кластеров структурированными гауссовыми смесями	291
<i>Корнилов Ф. А., Костоусов В. Б., Первалов Д. С.</i> Быстрые алгоритмы анализа изображений со структурными искажениями	303
<i>Домунян А. А.</i> Программно-аппаратная реализация кортексного сканера	310
<i>Бахшиян Б. Ц., Назиров Р. Р., Федяев К. С.</i> Точность оценивания координат реперных знаков на неподвижной цели по измерениям бортового оптического датчика	317

Предисловие

Системы технического зрения (СТЗ) в настоящее время являются одним из главных средств развития автоматических систем управления движением в условиях, когда объём априорной информации не достаточен и для решения задач управления необходим анализ внешней обстановки в режиме реального времени. СТЗ находят свое применение в современных космических, авиационных, наземных, надводных и подводных мобильных объектах. Благодаря дальнодействию и достаточно высоким уровням пространственного и цветового разрешения современных линейных и матричных приемников оптического излучения СТЗ могут служить незаменимыми источниками информации при автоматическом решении задач распознавания, навигации или наведения.

Вместе с тем следует признать наличие технических трудностей внедрения зрительного информационного канала в контуры систем управления мобильными объектами. Указанные трудности связаны с проблемами преобразования зрительной информации в данные результатов распознавания, навигационной привязки или определения параметров движения этих объектов в режиме реального времени.

Требование сокращения времени преобразований зрительных сигналов при высоких скоростях движения мобильных объектов достаточно сложно в реализации, но оно — не единственное препятствие для расширения использования СТЗ. Не менее сложные проблемы связаны с синтезом алгоритмов требуемых преобразований. Отсутствуют универсальные алгоритмы решения задач зрительного слежения, навигации, распознавания и наведения в общем случае движения аппарата или объекта зрительного слежения. Для каждой задачи СТЗ и в каждой конкретной ситуации фотометрических и траекторных условий оптимален лишь некий конкретный алгоритм, причём даже слабое изменение обрабатываемой зрительной сцены может потребовать смены используемого алгоритма преобразования фотометрического сигнала. Это вызывает необходимость обеспечения алгоритмической полноты и структурной устойчивости решения задач технического зрения на множестве возможных фотометрических ситуаций и траекторий движения.

Не менее сложная проблема алгоритмического обеспечения СТЗ связана с конструктивными недостатками их технической реализации — ограничениями динамических диапазонов свето- и цветопередачи, вопросами дискретности фотоприёмников, астигматизмом оптического канала, ошибками калибровочных характеристик и временных привязок потока видеоданных. Образно говоря, «красота алгоритма» гибнет под ударами шумов и искажений. Поэтому борьба с помехами в видеоданных является главной алгоритмической задачей при внедрении зрительной обратной связи в контур системы управления. Это требует серьёзного развития математических методов и алгоритмов зрительных преобразований в процессе решения конкретных задач управления мобильными объектами.

В свете сказанного большое значение имела организованная Институтом космических исследований (ИКИ) РАН научно-техническая конференция-семинар, посвященная обсуждению вопросов использования СТЗ в задачах управления мобильными объектами. Важная роль СТЗ и сложность перечисленных проблем их совершенствования собрали на конференции ведущих специалистов промышленности и научных организаций, известных своими результатами в создании

и внедрении отечественных СТЗ, их математического, алгоритмического и программного обеспечения. Доложенные на конференции результаты этих работ публикуются в настоящем сборнике. Большое значение для развития СТЗ имеет и то, что на конференции в плодотворном обсуждении докладов смогли принять активное участие не только крупные учёные и инженеры, но и работающие в этой области молодые специалисты и аспиранты.

В программный комитет конференции-семинара входили: **Назирев Р. Р.** — (председатель) заместитель директора ИКИ РАН, доктор технических наук, профессор; **Желтов С. Ю.** — член-корреспондент РАН, генеральный директор Федерального государственного унитарного предприятия «ГосНИИАС», доктор технических наук, профессор; **Гришин В. А.** — старший научный сотрудник Института космических исследований РАН, кандидат технических наук, доцент; **Платонов А. К.** — заведующий сектором Института прикладной математики им. М. В. Келдыша, доктор физико-математических наук, профессор; **Соколов С. М.** — ведущий научный сотрудник Института прикладной математики им. М. В. Келдыша, доктор физико-математических наук, профессор; **Визильтер Ю. В.** — начальник лаборатории Федерального государственного унитарного предприятия «ГосНИИАС», доктор физико-математических наук, старший научный сотрудник; **Кропотов А. Н.** — заведующий лабораторией Московского государственного технического университета им. Н. Э. Баумана, кандидат технических наук, доцент; **Носков В. П.** — заведующий сектором Московского государственного технического университета им. Н. Э. Баумана, кандидат технических наук; **Мещеряков А. Ю.** — ведущий научный сотрудник Института проблем управления РАН, кандидат технических наук, доцент.

Были заслушаны следующие доклады (соответствующие им статьи сборника могут иметь другие названия, состав авторов и разбивку по темам).

Пленарные доклады:

- Проблемы технического зрения в современных авиационных системах (*Желтов С. Ю., Визильтер Ю. В.*).
- Теоретические аспекты и приложения стереоскопических систем навигации, наведения и дистанционного зондирования местности (*Белоглазов И. Н.*).
- Фрагменты корреляционной теории в разработках прецизионных быстродействующих систем технического зрения (*Васильев Д. В.*).
- Результаты создания комплексированных систем технического зрения для обеспечения управления движением (*Носков В. П., Рубцов И. В.*).
- Биологически инспирированное техническое зрение в системах автономного искусственного интеллекта (*Жданов А. А.*).

Секция 1. Авиационно-космические приложения систем технического зрения:

- Технологии систем улучшенного / синтезированного зрения для управления летательными аппаратами (*Костяшкин Л. Н., Логинов А. А., Павлов О. В.*).
- Видеоинформационные технологии в аппаратуре бортовых оптико-электронных систем (*Блохин А. Н., Костяшкин Л. Н., Романов Ю. Н., Шанка С. В.*).
- Обработка и анализ изображений в бортовых оптико-электронных системах (*Алпатов Б. А., Бабаян П. В., Костяшкин Л. Н., Романов Ю. Н.*).

- Алгоритм анализа и принятия решения в задаче селекции объектов на изображениях наземных сцен (*Блохинов Ю. Б., Гнилицкий В. В., Инсаров В. В., Чернявский А. С.*).
- Автоматическое выделение и сопровождение псевдоточечных мерцающих объектов в реальном времени — модель фоноцелевой обстановки, алгоритм работы бортового процессора и технические решения по процессору (*Блажевич С. В., Винтаев В. Н., Ушакова Н. Н.*).
- Система информационного обеспечения задач сближения, стыковки и посадки космического аппарата на основе компьютерного видения (*Соколов С. М., Богуславский А. А.*).
- Задачи бортовой обработки информации телевизионной системы навигации и наблюдения при посадке на Фобос (*Аванесов Г. А., Жуков Б. С., Краснопецева Е. Б., Форш А. А.*).
- Анализ влияния динамики космического аппарата на характеристики алгоритмов обработки изображений системы технического зрения проекта ФОБОС-ГРУНТ (*Гришин В. А.*).
- К вопросу оценки корреляционных характеристик шума цифровых оптико-электронных систем видимого диапазона (*Веселов Ю. Г., Карпиков И. В.*).
- Система видеосклейки для беспилотных летательных аппаратов (*Макашов А. А., Кропотов А. Н., Артюхов М. Ю., Сахарова Е. И.*).
- Построение геометрических моделей объекта и снимков с заданными параметрами (*Говоров А. В.*).

Секция 2. Системы технического зрения в наземных и подводных роботах:

- Повышение эффективности применения мобильных робототехнических комплексов на базе использования средств наблюдения беспилотных летательных аппаратов (БПЛА) (*Аникин В. А., Ким Н. В., Носков В. П., Рубцов И. В.*).
- Техническое зрение мобильных роботов (*Михайлов Б. Б.*).
- Обнаружение препятствий перед наземным мобильным объектом в бортовой системе технического стереозрения реального времени (*Выголов О. В., Желтов С. Ю., Визильтер Ю. В.*).
- Гиперспектральная система зрения для мобильных роботов (*Овчинников А. М., Платонов А. К.*).
- Опыт создания системы локальной видеонавигации для подводных аппаратов (*Макашов А. А., Кропотов А. Н., Артюхов М. Ю., Сахарова Е. И.*).

Секция 3. Программно-аппаратное обеспечение систем технического зрения:

- Автоматическая система реального времени для обнаружения объектов и ориентиров на изображении, основанная на обработке цветных изображений (*Кий К. И.*).
- Морфологические методы поиска и распознавания объектов на изображении (*Пытьев Ю. П., Чуличков А. И.*).
- Сегментация изображений с использованием древовидных описаний кластеров в многомерном пространстве признаков (*Ланге М. М., Новиков Н. А.*).
- Точность оценивания координат реперных знаков неподвижной мишени по измерениям с помощью бортового оптического датчика (*Бахшиян Б. Ц., Назиров Р. Р., Федяев К. С.*).

- Обнаружение прямых и окружностей на изображениях с использованием многомасштабной оценки локальной кривизны границ (*Семейкина Е. В., Юрин Д. В.*).
- Обнаружение объектов с помощью поворотной IP-камеры (*Стрельников К. Н.*).
- Быстрые алгоритмы анализа изображений со структурными искажениями (*Корнилов Ф. А., Костоусов В. Б., Перевалов Д. С.*).
- Программно-аппаратная реализация кортексного сканера (*Домулян А. А.*).

В обсуждениях большой интерес присутствующих вызвали примеры успешно-го практического применения СТЗ в авиационной и космической областях. Особое внимание в этих работах было привлечено к средствам анализа последовательности видеоданных в режиме реального времени для решения сложной проблемы малого числа дискретных элементов зашумлённой сцены, приходящихся на объект интереса. В опытных разработках зрения мобильных роботов и подводных аппаратов внимание привлекли проблемы расширения состава средств получения зрительных данных (лазерные дальномеры, гиперспектрометры, системы подводного зрения), связанные с расширением областей применения робототехнических систем.

Алгоритмическая сторона процессов преобразований зрительных данных в публикуемых докладах связана с анализом математических свойств преобразований двумерных растровых изображений. Была убедительно продемонстрирована эффективность использования классических корреляционных методов в стереосистемах, в системах дистанционного зондирования поверхности Земли, при идентификации и сопровождении плохо видимых объектов.

В последнее время всё большее внимание уделяется методам сегментации зрительной сцены. В связи с этим на конференции оживлённо дискутировалась проблема «контур или область?» в задачах распознавания зрительных образов, а в докладах была показана эффективность использования новых математических методов описания областей изображения в алфавите специфических характеристик яркости или цвета. К этому направлению относятся и поиски формальных методов морфологического или структурного анализа сцен. Следует заметить, что предложенные на конференции методы сегментации изображения вполне пригодны для их параллельной реализации. Это связано с развитием современной элементной базы — программируемых логических интегральных схем (ПЛИС), цифровых сигнальных процессоров (ЦСП), многопроцессорных и многоядерных вычислительных средств, обеспечивающих возможность многопоточной и параллельно-конвейерной обработки данных видеопотока.

Безусловно, предлагаемый сборник принесёт заметную пользу специалистам в области создания систем технического зрения.

**АВИАЦИОННО-КОСМИЧЕСКИЕ ПРИЛОЖЕНИЯ
СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ**

ПРОБЛЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ В СОВРЕМЕННЫХ АВИАЦИОННЫХ СИСТЕМАХ

Ю. В. Визильтер, С. Ю. Желтов

*Федеральное государственное унитарное предприятие «Государственный научно-исследовательский институт авиационных систем» (ФГУП «ГосНИИАС»),
125319, Москва, Викторенко ул., 7*

Статья посвящена основным проблемам построения систем технического зрения в современных и перспективных авиационных системах. Рассмотрены основные типы датчиков, используемых в таких системах, и основные требования, предъявляемые к алгоритмам технического зрения. Описаны задачи и методы обнаружения объектов в сцене наблюдения; анализа изображений на этапе планирования авиационных операций; получения, обработки и комлексирования многоспектральных видеоданных. Отдельно рассмотрен класс систем улучшенного видения для самолётов гражданской авиации. Перечислен ряд теоретических и практических результатов, полученных в ГосНИИАС в области разработки перспективных методов технического зрения.

Работа выполнена при поддержке РФФИ (проект № 09-07-13551-офи_ц).

Ключевые слова: авиация, техническое зрение, обработка изображений, анализ изображений.

Введение

Системы управления современными летательными аппаратами (ЛА) предназначены для управления сложными многофункциональными объектами, действующими в сложной окружающей обстановке. При этом канал зрительного восприятия является одним из наиболее важных источников информации как в автоматических, так и автоматизированных (человеко-машинных) системах управления. Вследствие этого в последние годы на передний план всё в большей степени выходят задачи создания систем технического зрения (СТЗ) для различных типов ЛА двойного назначения.

Как показывают исследования Всемирного фонда безопасности, почти 75 % аварий самолётов при заходе на посадку и посадке происходят в аэропортах, где недоступны или отсутствуют приборы точного захода на посадку, в условиях плохой видимости. В связи с этим одним из важных направлений совершенствования бортовой авионики является разработка аппаратно-программных комплексов «улучшенного видения» (Enhanced Vision Systems, EVS). В качестве источников информации в таких системах могут использоваться телевизионные видеодатчики, инфракрасные (ИК) датчики различных диапазонов, миллиметровые радары (МР), лазерные локаторы (ЛЛ), базы данных рельефа местности вдоль маршрутов полёта, базы данных аэропортов и объектов взлётно-посадочной полосы (ВПП), навигационные параметры и ряд других. Сформированная системой EVS оперативная графическая информация далее представляется пилоту в реальном времени на соответствующем устройстве отображения — индикаторе на лобовом стекле

Визильтер Юрий Валентинович — начальник лаборатории компьютерного машинного зрения, доктор физико-математических наук, старший научный сотрудник, e-mail: viz@gosniias.ru.

Желтов Сергей Юрьевич — член-корреспондент РАН, генеральный директор, доктор технических наук, профессор.

(ИЛС) или многофункциональном индикаторе (МФИ), представляющем собой компьютерный дисплей, входящий в состав бортовой системы отображения информации. По функциональным характеристикам системы улучшенного видения разделяются на три типовых класса: системы улучшенной визуализации, системы синтезированного видения и системы автоматизированного видения. При этом в составе EVS выделяются две основные подсистемы: система технического зрения (СТЗ), выполняющая операции ввода и обработки видеоинформации, и система компьютерной визуализации (СКВ), непосредственно формирующая и представляющая пилоту графические образы закабинной обстановки.

В то же время во многих областях применения авиационной техники речь сегодня идёт уже не об улучшении каких-то отдельных характеристик единичных ЛА, а о наиболее эффективных способах построения и использования целостной системы управления, связи, обработки и сбора информации (так называемые системы С4I). Эта система включает множество ЛА, а также целый ряд подсистем, служб и сетей, в том числе связанных с изображениями и географическими данными. В неё входят космические и авиационные платформы получения данных, службы географической поддержки, наземные системы сбора информации, системы планирования операций, моделирования, навигации, управления движением, целеуказания и ряд других. Неотъемлемой частью построения системы С4I является сбор и комплексирование информации, поступающей от датчиков различной физической природы, таких как космоснимки, авиаснимки, видео- и ИК-данные, данные от радаров, лазерных систем и т. д. С учётом этого самолёт начала XXI в. должен обеспечивать быструю и помехоустойчивую передачу больших объёмов данных, высокую точность навигационных систем, сверхбольшую дальность применения, всепогодность, высокую точность и гибкость при выполнении поставленных задач. Создания и применения различных технологий анализа изображений требуют сегодня системы сбора информации на базе космических, авиационных и БЛА носителей, системы подготовки полётных заданий, системы высокоточной геопространственной привязки, системы автономного управления высокоточных беспилотных ЛА и летающих робототехнических комплексов.

В данной статье мы постараемся дать обзор основных задач, методов и технологий технического зрения, связанных с авиационными приложениями. При этом в качестве примеров будут приведены некоторые теоретические и практические результаты, полученные в данной области в ГосНИИАС.

1. ОБНАРУЖЕНИЕ ОБЪЕКТОВ И ИДЕНТИФИКАЦИЯ ЦЕЛЕЙ

1.1. Автоматическое и автоматизированное обнаружение целей

Назначение автономной системы навигации и целеуказания сводится к максимально эффективному обнаружению определённых объектов на местности, их классификации (идентификации) в пределах установленных классов и выдаче соответствующих директив исполнительной системе управления. В автоматизированной системе те же задачи решаются человеком-оператором на основании информации, предоставленной ему системой обработки и комплексирования информации. При этом основные функции систем машинного зрения и состав необходимых алгоритмов анализа изображений оказываются в обоих случаях достаточно сходными.

Для зарубежных исследований в области технологии автоматического обнаружения целей (ATR — Automatic Target Recognition) было характерно чередование относительных успехов и неудач примерно до 1987 г., когда научно-технический комитет МО США окончательно признал её перспективной в рамках программы LANTIRN (Low Altitude Navigation and Targeting Infrared for Night). В ходе реализации этой программы были преодолены трудности, связанные с распознаванием воздушных и крупных стационарных целей в условиях отсутствия или наличия незначительных местных помех. В настоящее время исследования технологии ATR ведутся главным образом с точки зрения её применения в оптико-электронных системах, РЛС и системах с комбинацией датчиков. Именно в этих направлениях сосредоточены усилия научных лабораторий и промышленных фирм, причём основное внимание уделяется усовершенствованию алгоритмов, датчиков и процессоров.

Анализ разработок зарубежных фирм в области создания систем ATR показывает, что практическое применение находят четыре основные группы алгоритмов распознавания цели:

- алгоритмы согласованной фильтрации с проверкой совпадения, использующие шаблоны для целей, которые необходимо обнаружить;
- алгоритмы статистического распознавания образов, предусматривающие выборку изображений целей с учётом их конкретных характерных признаков;
- алгоритмы технического зрения на основе моделей, с помощью которых сравниваются характерные признаки наблюдаемой цели с хранящимися в памяти ЭВМ моделями данного изображения;
- нейросетевые алгоритмы, предполагающие обучение на примерах искусственной нейронной сети, имитирующей структуру мозга человека.

Выбор того или иного способа распознавания целей так же, как и разработка системы ATR в целом, требует чётко определённой основы, предусматривающей ответы на следующие вопросы:

- каковы требования к выполнению задачи данным ЛА?
- насколько сложна классификация целей для разрабатываемой системы?
- потребуется ли для выполнения поставленной задачи дедуктивный или индуктивный метод рассуждений?
- если в процессе принятия решений предусматривается участие человека, какова должна быть степень подробности «подсказок» со стороны машины?
- какие характеристики датчиков являются в данном случае ключевыми?
- как совмещать разнородную информацию от датчиков, установленных на различных платформах, имеющих разную внутреннюю геометрию и представленных в различных форматах?

Фирмы Martin Marietta и Texas Instruments ведут разработку подсистем обнаружения и алгоритма распознавания целей для системы самонаведения крылатых ракет (КР), предназначенных для автономного поиска и уничтожения особо важных объектов. Подсистема фирмы Martin Marietta включает РЛС миллиметрового диапазона и систему FLIR. Фирма Texas Instruments в качестве средства обнаружения избрала лазерный локатор LADAR и систему FLIR. Алгоритмы распознавания целей обеих конкурирующих фирм используют сопоставление их трёхмерного

изображения с моделью, введенной в память системы обработки данных, или же их сигнатур, полученных под различными углами обзора. Весь процесс распознавания занимает доли секунды. В случае недостатка данных КР может сойти с маршрута, сблизиться с объектом и совершить его облёт. По заявлению представителей фирмы Texas Instruments, разработанный фирмой алгоритм АТР позволяет различать пусковые установки ЗРК и ракет класса «поверхность — поверхность» или же транспортные автомашины и подвижные пункты связи. По утверждению представителей фирмы Martin Marietta, разработанный ею алгоритм с вероятностью 85 % обеспечивает распознавание трёх типов армейских транспортных средств с примерно равными размерами и сходной конструкцией.

Алгоритмы распознавания целей с использованием нейронных сетей разрабатываются в США управлением DARPA. В центре исследований — отработка алгоритмов на основе абстрактного представления целей вместо использования детальных изображений. Это особенно важно в тех случаях, когда возникает необходимость в усовершенствовании систем АТР с включением изображений новых целей, так как значительно сокращает время на сбор информации и обработку алгоритма.

В 1999 г. на вооружение палубной авиации ВМС США поступила тактическая крылатая ракета Boeing AGM-84H SLAM-ER — первая американская система оружия, обладающая способностью автоматического распознавания целей — режим АТР. По сравнению с режимом автоматического захвата цели (АТА — Automatic Target Acquisition), реализованным ранее в ряде авиационных средств поражения, в режиме АТР изображение потенциальной цели, получаемое бортовыми датчиками, в реальном масштабе времени сравнивается с её цифровым образом, заложенным в память БЦВМ, что позволяет осуществлять автономный поиск объекта, его идентификацию и нацеливание ракеты при наличии лишь приблизительных данных о местонахождении цели. Автономные ЛА типа КР JASSM также снабжаются комбинированной системой наведения — инерциально-спутниковой на маршрутовом участке полёта и тепловизионной, с режимом автоматического распознавания цели — на конечном. Такая комбинированная система позволяет, в частности, реализовать режим перенацеливания в полёте.

Однако, несмотря на явный прогресс в области разработки систем АТР, за рубежом по-прежнему наблюдается большой интерес к созданию автоматизированных человеко-машинных систем управления ЛА с элементами технического зрения. Авиационные специалисты не торопятся переходить на полностью автоматические системы целеуказания. На наш взгляд, такая осторожность во многом оправдана, и это связано с двумя следующими основными факторами. Во-первых, решение о нанесении удара по цели слишком ответственно, и потому необходима гораздо более высокая степень уверенности обнаружения и распознавания, чем та, которая обеспечивается в настоящее время лучшими системами АТР. Во-вторых, в условиях постоянно обновляющейся номенклатуры изделий, принимаемых на вооружение (как в нашей стране, так и за рубежом), для поддержания боеспособности систем АТР в каждой из них необходимо создать и постоянно обновлять банк моделей объектов, подлежащих точному распознаванию, что является достаточно тяжелой задачей в техническом, экономическом и организационном плане. Поэтому на настоящем этапе представляется более целесообразным ориентироваться на включение элементов АТР в системы информационной поддержки автоматизированных систем управления ЛА. При этом требования к надежности распознавания немного смягчаются, поскольку окончательное решение прини-

мает человек-оператор, который может проигнорировать или уточнить информацию, если она покажется ему недостаточно достоверной. В этом случае оператор в обязательном порядке должна быть доступна и «сырая» информация от датчиков, которую он мог бы интерпретировать самостоятельно, «в обход» системы АТР. В автоматизированных системах снижаются и требования к «разрешению» распознающих алгоритмов. Системе информационной поддержки достаточно привлечь внимание оператора к определённому участку сцены, после чего распознавание точного типа объектов и принятие решения о необходимости тех или иных действий осуществит сам оператор. При такой постановке задачи нет необходимости поддерживать сверхподробную базу моделей возможных целей, как это имеет место в случае «чистого» АТР. База моделей может включать лишь общее описание крупных классов целей. В то же время уменьшение подробности выдаваемых оператору «подсказок» позволяет резко увеличить скорость обработки информации, что ведёт к высвобождению вычислительных ресурсов для решения других задач управления ЛА.

В заключение данного раздела необходимо ещё раз отметить, что задача автоматического или автоматизированного обнаружения целей является безусловно базовой, но всё же частной технологической задачей по отношению ко всему комплексу задач машинного зрения в перспективных ЛА, которые в общих чертах могут быть сформулированы следующим образом:

- обнаружение объектов и изменений в сцене наблюдения;
- высокоточные измерения элементов сцены;
- слежение за объектами;
- самоориентация и самопозиционирование ЛА;
- реконструкция наблюдаемых поверхностей и обнаружение трёхмерных структур;
- описание сцены и идентификация объектов.

1.2. Требования к алгоритмам обнаружения и идентификации в СТЗ ЛА

В последние годы использование цифровых видеодатчиков с высоким разрешением позволило существенно приблизить информативность искусственных зрительных устройств к соответствующим характеристикам человеческого глаза. При этом аппаратные возможности, предоставляемые последними достижениями электроники и вычислительной техники, настолько высоки, что вычислительные мощности перспективных бортовых ЭВМ уже сегодня вплотную приближаются к характеристикам «вычислительных мощностей», используемых для обработки изображений в мозгу. Таким образом, на пути к осуществлению некогда фантастической мечты будущего — созданию полностью автоматизированных и автоматических ЛА, способных выполнять любые самые сложные боевые или хозяйственные задания, — сегодня в значительной степени стоит необходимость решения лишь одной технической задачи — разработки методов и алгоритмов «понимания» изображений, получаемых по различным каналам информационного обеспечения современного ЛА. Однако именно эта задача во многих случаях оказывается и наиболее трудной.

Специфика разработки конкретных систем технического зрения ЛА определяется следующими основными особенностями. Во-первых, при разработке подсистем информационного обеспечения ЛА требуется решать не общую проблему автоматического понимания изображения произвольной сцены, а гораздо более

определённую и узкую задачу *проблемно-ориентированной интерпретации изображений*. Наиболее часто эта задача сводится к *обнаружению и идентификации* на изображении некоторого набора объектов заранее известных типов, присутствие которых может повлиять на формирование управления. Во-вторых, к алгоритмам обработки изображений в подсистемах информационного обеспечения предъявляются *специальные требования*, связанные с назначением и конкретными характеристиками разрабатываемой или уже существующей системы управления. В зависимости от специфики конкретной задачи, система автоматического управления предъявляет к подсистеме анализа изображений такие требования как *робастность, локализация и вычислительная реализуемость* в рамках заданной аппаратной архитектуры. При этом одна из центральных проблем, отличающих область обработки изображений, например, от теории обработки сигналов, заключается в разработке методов обнаружения объектов, слабо чувствительных к разнообразным видам изменчивости, характерным только для изображений, таким как различные типы искажений оптических сенсоров, блики, затенения, загораживания, искажения формы, ракурсные искажения, шумовые компоненты. В то же время необходимо отметить, что существенное расширение возможностей подсистем технического зрения ЛА может быть достигнуто при совместном использовании датчиков различной физической природы (ДРФП). При этом финальная эффективность решения задачи управления будет зависеть как от используемого набора и параметров датчиков, так и от характера алгоритмов обработки и комплексирования информации.

Рассмотрим требования к методам и алгоритмам технического зрения в системах управления ЛА на примере наиболее специфической группы алгоритмов — алгоритмов обнаружения и идентификации объектов на изображениях.

Робастность. Алгоритмы технического зрения часто отрабатываются на математических моделях характерных целей, которые могут быть получены искусственным путём с помощью ЭВМ. Этот метод считается более быстрым и даёт возможность отрабатывать алгоритмы в отношении объектов, реальные изображения которых отсутствуют. Однако создаваемые алгоритмы должны работать не только на модельных, но и на реальных изображениях. Между тем, каждый, кто конструировал алгоритмы обработки реальных изображений, хорошо знает, насколько изменчивы и неформализуемы могут быть факторы, влияющие на качество реальных изображений от датчиков и соответственно — на вероятность обнаружения и распознавания объектов на этих изображениях.

Перечислим эти факторы более подробно:

- шумовые эффекты — имеют десятки видов источников возникновения, к числу которых можно отнести несовершенство сенсоров приёмо-передающей аппаратуры, аппаратуры оцифровки изображений, трудные условия съёмки, недостаток освещения и ряд других;
- сложный текстурированный фон, на котором должно происходить обнаружение объектов, например, обнаружение штриховой наклейки на газетной странице и т. п.;
- эффекты загораживания (заслонения) одних объектов другими, как правило, не определённой заранее формы, например — облако на космофото-снимке и т. п., загораживающие помехи;
- искажающие оптические эффекты в виде различных расфокусировок и дисторсий, ракурсные искажения и др.;

- эффекты резкой смены освещения, блики, тени, особенно в динамически меняющихся сценах;
- разнообразие или изменчивость самих объектов обнаружения — переменная структура (как у штриховых кодов), дефекты, временные изменения формы, вегетационные циклы для растительности и т. п.;
- эффекты изменения среды между сенсорами и объектами наблюдения — задымления, атмосферные осадки, пыль, искусственные помехи и многие другие;
- несинхронная запись и обработка данных в динамических задачах обнаружения, связанная с ограничениями компьютерных средств хранения и анализа изображений, особенно критическими для приложений с требуемыми высокими временами реакции системы обнаружения объектов; сюда можно отнести также сбои в компьютерных программах обработки.

Даже беглый анализ приведенных факторов демонстрирует практическую невозможность их полного формального математического описания — вероятностного, радиометрического или геометрического. Отсутствие формализованного описания ключевых факторов, вносящих неопределённость в процесс обработки, приводит к тому, что говорить о существовании единственного оптимального алгоритма для решения той или иной задачи обработки изображений в подобных случаях будет невозможно ещё многие годы. Представим себе, что существует несколько алгоритмов, достигающих примерно одинаковых результатов на «идеальных» (неискажённых) изображениях. Тогда возникает естественный вопрос, как сравнить эти алгоритмы по качеству их работы. При разработке реальных алгоритмов в настоящее время стандарт де-факто состоит в проверке эффективности работы сконструированных алгоритмов на огромных выборках реальных данных или изображениях, содержащих все неприятные ситуации. Такие алгоритмы, которые обладают устойчивостью к значительным искажениям и меняющимся факторам, принято называть *робастными*. Робастность следует отнести к основным практическим требованиям, предъявляемым при разработке алгоритмов обнаружения объектов и других алгоритмов машинного зрения.

Локализация. Важное отличие, присущее собственно проблеме обнаружения объектов на изображениях по сравнению с задачами распознавания или интерпретации заранее сегментированного образа, заключается в том, что обнаружение в практических задачах всегда связано с процедурой *поиска* объекта. Именно реализация процедуры поиска объекта связана с угрозой взрывообразного роста необходимого объёма вычислений. Проиллюстрируем это на примере простой задачи поиска объекта путём сравнения текущего изображения сцены с растровым эталоном или шаблоном формы объекта. Если построить какой-либо функционал соответствия между объектом размером $M \times M$ и фрагментом $M \times M$ из изображения размера $N \times N$, то простой перебор фрагментов требует количества вычислений не менее чем $M^2 N^2$ операций, что, например, при размере объекта 50×50 , а изображения — 2000×2000 элементов составляет 10 миллиардов операций. Даже с учётом значительного увеличения мощности современных БЦВМ, такие объёмы вычислений по-прежнему далеко выходят за пределы возможностей реализации бортовых систем реального времени, предназначенных для таких задач как навигация и наведение ЛА. Более того, реальные задачи обработки визуальной информации, как правило, изобилуют дополнительными степенями свободы, когда искомая

яркостно-геометрическая структура на изображении может иметь не только произвольные положение, угловую ориентацию и масштаб, но и подвергаться разным преобразованиям, не только аффинным или проективным, но и гораздо более сложным, таким как различные «резиновые» модели или «коробление» (warping). Всё это катастрофически увеличивает потребное для корреляционного перебора время расчётов и требует применения качественно иных идей по организации процесса обнаружения.

В связи с этим второе важнейшее свойство, которым должны, как правило, обладать алгоритмы обнаружения объектов на изображениях, можно определить как *точную локализацию*. Это понятие означает, что необходимо не только обнаружить объект, но и точно указать в системе координат изображения (или сцены) его положение в каком-либо смысле. Несколько неясное толкование понятия «локализации», приведенное выше, связано с тем, что по сравнению со своей эталонной моделью объект на реальном изображении может быть заметно искажён геометрически, причём аналитическая модель искажения может отсутствовать. В этих случаях локализация объекта является нетривиальной задачей. В более простой ситуации, при аналитически известной с точностью до параметров геометрии искажений, под точной локализацией можно понимать знание о положении какой-либо характерной точки объекта и параметрах геометрии искажения (углы поворота, элементы проективного преобразования и т. п.). При этом встречающиеся случаи ошибок локализации целесообразно разделить на две группы — нормальные и аномальные ошибки. *Нормальная ошибка* — это правильная локализация объекта с некоторой позиционной или параметрической неточностью, характеризующейся количественными оценками. Для объектов, характеризующихся габаритными размерами, большими $3 \times 3 \dots 5 \times 5$ элементов изображения, позиционные нормальные ошибки могут быть значительно меньше размера элемента изображения, уменьшаясь с величиной объекта. В этом случае принято говорить о возможности *субпиксельной локализации*. Это особенно важно для задач стереонаблюдения, так как при малых параллелях 3D-объектов субпиксельная локализация самым существенным образом определяет точность их пространственного положения. К *аномальным ошибкам* следует отнести ситуацию перепутывания объектов или возникновения артефактов (ложных объектов) на фоне, что связано с фатальными количественными ошибками позиционирования или просто ложным обнаружением. Требования по исключению или ограничению уровня аномальных ошибок составляют очень важную часть требований к алгоритмам обнаружения, так как ошибочное целеуказание непосредственно приводит к формированию неэффективного управления.

Вычислительная реализуемость. Несмотря на отмеченный ранее колоссальный прогресс вычислительной техники и создание обширной специализированной процессорной базы для обработки изображений, для основной массы бортовых приложений реального времени характеристики вычислителей и их свойства всё ещё далеки от желаемых. Даже в случае реализации простейших алгоритмов оконной фильтрации изображения с минимальной апертурой 3×3 элемента объём вычислений составляет десятки операций на точку изображения. При обработке более высокого уровня необходимый объём вычислений колеблется в пределах от сотен до тысяч операций на пиксел. Если размер анализируемого изображения составляет порядка 1000×1000 элементов, что не является чем-либо необычным для современных видеодатчиков (можно вспомнить, что бытовые цифровые

фотоаппараты давно превзошли отметку 5 Мпикс. в ПЗС-матрице), мы получим оценку количества потребных вычислений порядка нескольких гигабайтов операций на кадр. Между тем, для приложений реального времени необходимо выполнять эти вычисления в темпе кадровой развертки (не менее 25 кадров в секунду), что приводит к оценке быстродействия около 50 Гфлопс. Сами по себе эти оценки сегодня не являются запредельными для ЭВМ последнего поколения, однако следует учесть, что в случае создания систем управления перспективных ЛА массогабаритные характеристики конструируемых вычислительных устройств должны быть весьма ограничены. Таким образом, вычислительная реализуемость алгоритмов по-прежнему относится к числу наиболее важных факторов, учитываемых при их разработке.

1.3. Современный подход к обнаружению объектов в сцене наблюдения

Вплоть до 80-х гг. прошлого века основным классом алгоритмов обнаружения и идентификации объектов, использовавшихся в бортовых системах технического зрения ЛА, являлся класс корреляционно-экстремальных алгоритмов [Красовский и др., 1979]. Однако по мере возрастания требований к точности и надежности алгоритмов обнаружения всё более сложных объектов во всё более сложной реальной обстановке недостатки данной группы методов стали проявляться всё более явно. Это прежде всего высокая вероятность аномальных ошибок, необходимость иметь большое число эталонов для описания разноракурсных образов трёхмерных объектов, неустойчивость по отношению к яркостно-геометрической изменчивости изображений, имеющей место в реальных условиях регистрации. Таким образом, намечился переход от корреляционных детекторов заданных образов к методам и алгоритмам структурного анализа изображений.

В настоящее время последовательность процедур обработки изображений принято рассматривать в соответствии с так называемой парадигмой Марра [Марр, 1987]. Эта парадигма, предложенная Д. Марром на основе длительного изучения механизмов зрительного восприятия человека, утверждает, что обработка изображений опирается на несколько последовательных уровней восходящей информационной линии «иконическое представление объектов (растровое изображение, неструктурированная информация) — символическое представление (векторные и атрибутивные данные в структурированной форме, реляционные структуры)» и должна осуществляться по модульному принципу посредством следующих этапов обработки:

- предобработка изображения;
- первичная сегментация изображения;
- выделение геометрической структуры видимого поля;
- определение относительной структуры и семантики видимой сцены.

Связанные с этими этапами уровни обработки обычно называются соответственно обработка нижнего уровня, среднего уровня, высокого уровня. В то время как алгоритмы обработки нижнего уровня (фильтрация простых шумов, гистограммная обработка) могут рассматриваться как хорошо проработанные и детально изученные, алгоритмы среднего уровня (сегментация) продолжают сегодня оставаться центральным полем приложения исследовательских усилий. За последние годы значительный прогресс достигнут по отношению к проблемам сопоставления точек и фрагментов изображений (*matching*) [Gruen, Baltsavias, 1988; Schenk,



Рис. 1. Основные типы характерных черт

1996], выделения признаков внутри малых фрагментов [Grimson, 1990; Marr, Hildreth, 1980; Mumford, 1987], высокой точности 3D-позиционирования точек [Gruen, Baltsavias, 1985; Haralick, Chu, 1984], что подразумевает соответствующее моделирование и калибровку датчиков и их комбинаций, выделение простых яркостно-геометрических структур типа «точка», «край», «пятно», «прямая линия», «угол» [Burns et al., 1986; Canny, 1986; Kitchen, Rosenfeld, 1982; Marr, Hildreth, 1980; Seeger, Seeger, 1994]. Эти «первичные» особенности изображения, также называемые *характерными чертами* (ХЧ) (рис. 1), играют базовую роль и при составлении яркостно-геометрических моделей объектов и разработке робастных алгоритмов их выделения. Ввиду высокой значимости характерных черт для разработки алгоритмического обеспечения задач управления ЛА, в ГосНИИАС были проведены исследования свойств различных типов ХЧ на типичных сюжетах — авиационных снимках местности.

Приведём классификацию характерных черт, которые могут присутствовать на изображениях и иметь следующие виды атрибутов:

- 1) *положение*: концы отрезка, центр отрезка, центр тяжести области, вершины многоугольников;
- 2) *геометрические* атрибуты: ориентация, длина, кривизна, площадь, периметр, ширина линии, минимальный и максимальный диаметр области, оси симметрии, число и положение особых точек, показатель компактности и др.;
- 3) *радиометрические* атрибуты: контраст, статистика распределения яркости, знак и величина края, автокорреляция;
- 4) *текстурные* атрибуты: матрица смежности, показатель однородности, энергия, энтропия, статистика градиентов текстуры, результаты применения текстурных фильтров, моменты;
- 5) *топологические* атрибуты: связность, соседство, общие точки, пересечение, параллельность, перекрытие, включение;
- 6) *цветовые / многозональные* атрибуты: вектор атрибутов для каждого канала;
- 7) *динамические* атрибуты: атрибуты статических и движущихся объектов;
- 8) *временные* атрибуты: функции изменения атрибутов со временем.

Выбор конкретных ХЧ и их атрибутов для построения алгоритмов обнаружения должен основываться на следующих основных критериях:

- 1) *присутствие / плотность*: наличие данных ХЧ на всех используемых изображениях, достаточная плотность ХЧ для покрытия интересующего района;
- 2) *редкость / уникальность*: редкость конкретной ХЧ на изображении, уникальность ХЧ в окрестности;
- 3) *инвариантность / устойчивость*: робастность по отношению к геометрическим и радиометрическим искажениям, нечувствительность к шуму;
- 4) *локализация*: возможность точной локализации;
- 5) *интерпретация*: возможность быстрого распознавания и интерпретации;
- 6) *скорость*: время выделения данного класса ХЧ из исходного изображения.

При работе с реальными изображениями перечисленные критерии являются противоречивыми. Поэтому конкретный выбор ХЧ и их атрибутов зависит от доступной вычислительной мощности и от минимальной требуемой робастности описания модели объекта в терминах ХЧ.

Методы обработки высокого уровня, относящиеся собственно к «пониманию изображений», находятся ещё в начальной фазе развития и по-прежнему представляют собой «вызов» для сообщества исследователей в области компьютерного зрения и искусственного интеллекта. Безусловно, перспектива создания будущих поколений интеллектуальных систем управления в основном зависит от дальнейшей разработки именно этого круга алгоритмов.

Если говорить о методах и алгоритмах обнаружения сложных объектов, то одна из центральных проблем, отличающих методы обработки изображений от хорошо изученной теории обработки сигналов, заключается в разработке методов обнаружения объектов, слабо чувствительных к разнообразным видам изменчивости, характерным лишь для изображений. Такими специфическими видами изменчивости являются ракурсные и радиометрические искажения, а также различные виды искажений, не сводимые к вероятностным моделям (шумы формы). На пути борьбы с ними были предложены как огромное множество эвристических алгоритмов обнаружения конкретных типов объектов, так и ряд подходов, обладающих большей общностью: методы корреляционного обнаружения [Василенко, Цибулькин, 1985; Левшин, 1987; Ярославский, 1987], преобразование Хафа [Davies, 1987; Duda, Hart, 1972; Hough, 1962; Illingworth, Kittler, 1987], морфологические подходы Пытьева [Пытьев, 1983, 1984] и Серра [Serra, 1982, 1986]. Значительный вклад в разработку методов и алгоритмов обработки изображений и машинного зрения применительно к обсуждаемым задачам обнаружения внесли работы Л. П. Ярославского, В. К. Злобина, В. Л. Лёвшина, Р. Харалика, Е. Дэвиса, Р. Неватиа, Е. Дикманнса, В. Фёрстнера и многих других. Однако, несмотря на достигнутые результаты, общее состояние проблемы выделения и идентификации сложноструктурированных объектов на моноскопических изображениях можно охарактеризовать как неудовлетворительное. Ещё сложнее выглядит задача обнаружения трёхмерных структур на стереоскопических изображениях. Здесь только намечаются подходы к более общим постановкам.

Рассмотрим, например, часто встречающуюся задачу обнаружения зданий (см., например, известные работы [Forstner, 1995; Huertas, Nevatia, 1988]). На рис. 2 (см. с. 22) представлены этапы работы реализованного в ГосНИИАС модульного алгоритма автоматического обнаружения зданий с борта летательного аппарата [Zheltov et al., 1997]. Данный алгоритм включает следующие основные этапы: выделение контурного препарата, выделение первичных линеаментов, выделение крупных и /или параллельных линеаментов, отбор линеаментов по размеру, поиск

П-образных структур, формирование окончательной 2D-модели сцены, установление межкадрового соответствия между 2D-моделями, формирование окончательной 3D-модели наблюдаемой городской сцены (рис. 3), определение целевых зданий в составе наблюдаемой 3D-модели городской сцены.

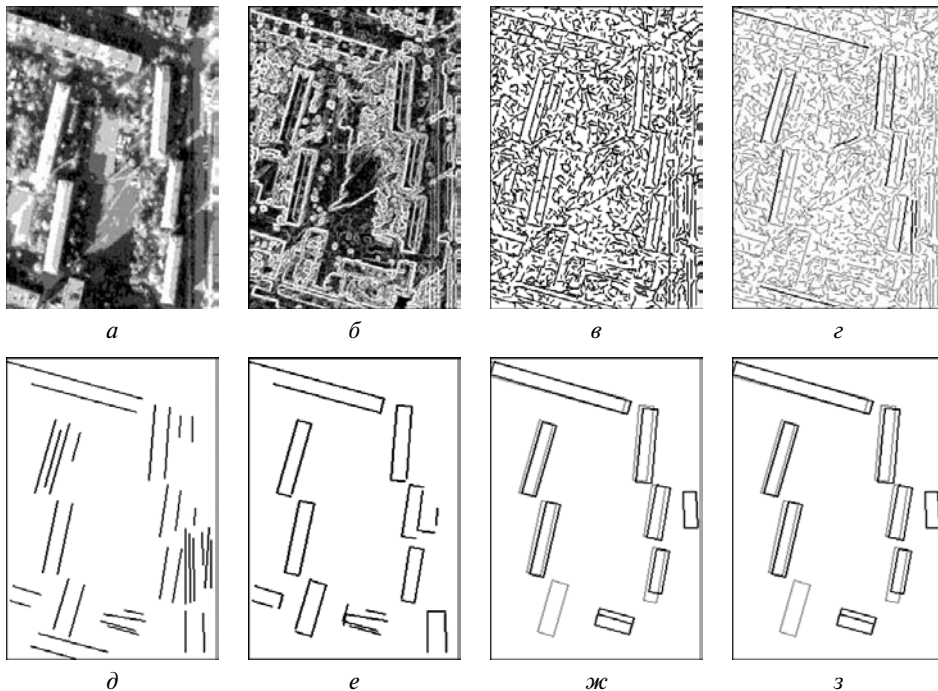


Рис. 2. Пример модульного алгоритма автоматического обнаружения зданий с борта летательного аппарата: *a* — исходное авиационное изображение; *б* — контурный препарат; *в* — первичные линеаменты; *г* — крупные и/или параллельные линеаменты; *д* — окончательный список линеаментов; *е* — П-образные структуры; *ж* — 2D-модель сцены; *з* — стереотождествление 2D-моделей

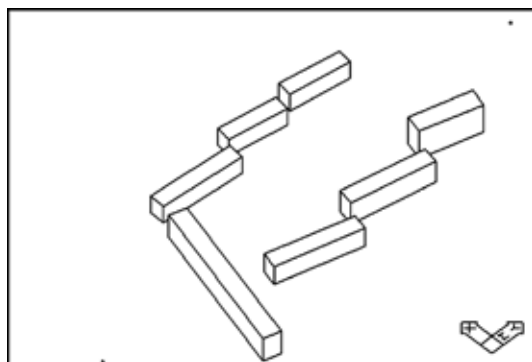


Рис. 3. Сформированная 3D-модель наблюдаемой городской сцены

1.4. Методы и подходы, развиваемые в ГосНИИАС

На протяжении последних десятилетий в ГосНИИАС ведутся регулярные исследования и разработки, связанные с созданием перспективных систем технического зрения ЛА. За это время, помимо ряда прикладных систем технического зрения, был создан и ряд новых теоретических подходов к построению алгоритмов анализа изображений. Ниже приводится краткое описание основных направлений работ, ведущихся в данной области, и полученных при этом результатов.

Разработка методов анализа свидетельств на изображениях. Для непосредственного описания и разработки алгоритмов обнаружения сложных структурированных объектов был предложен обобщённый метод *совместного анализа свидетельств на изображениях* [Visilter, 1996a, b]. Данный метод изначально основан на теоретико-вероятностном подходе, причём обнаружение сводится к проверке гипотезы о нахождении изображения объекта на реальном изображении. При этом любой *яркостно-геометрической модели* объекта ставится в соответствие некоторая экспертная *вероятностная модель*, описывающая статистические зависимости между различного рода особенностями изображения и гипотезой о принадлежности наблюдаемого объекта данной яркостно-геометрической модели. Полученная вероятностная модель используется уже на этапе обнаружения объекта, непосредственно *в ходе низкоуровневого анализа* конкретного предъявляемого изображения. При этом каждая обнаруженная особенность данного изображения (характерная черта) рассматривается как *событие, свидетельствующее* в пользу гипотезы (ряда гипотез) о наличии и характеристиках искомого объекта. *Инвариантность* к условиям регистрации и *робастность* обнаружения обеспечиваются за счёт соответствующего выбора типа используемых свидетельств (ХЧ), таких, что их связь с оцениваемыми параметрами регистрации не зависит или слабо зависит от значений параметров, которые на этапе обнаружения не оцениваются.

Было показано, что обобщённый метод анализа свидетельств определяет следующие основные возможности повышения вычислительной эффективности алгоритмов детектирования: независимое аккумулирование свидетельств, декомпозиция вектора параметров, редукция вектора параметров, закругление модели объекта. Соответствующая *модульная схема обобщённого алгоритма обнаружения* объектов содержит по крайней мере три основные процедуры, применяемые последовательно: *обработка изображения* по схеме голосования с целью выделения объектов или их составляющих; *анализ аккумулятора* с целью определения положения и/или ориентации объектов; *повторный анализ изображения* с целью проверки природы обнаруженных объектов и уточнения их параметров. Для реализации процедур анализа свидетельств в виде модульных алгоритмов обработки изображений предложена *формальная последовательность шагов разработки* алгоритмов обнаружения, позволяющая разрабатывать эффективные процедуры инвариантного детектирования конкретных классов объектов на изображениях.

Необходимо отметить, что, в отличие от общих вероятностных, информационных, алгебраических, нейросетевых и т. п. подходов, предложенная в ГосНИИАС технология анализа свидетельств имеет чётко очерченную область применения в задачах обнаружения и локализации сложных структурных объектов на цифровых изображениях. Данный метод предполагает наличие *яркостно-геометрической модели двумерного образа* искомого объекта, исходя из которой он позволяет стандартным образом порождать необходимые *яркостно-геометрические* модели

подлежащих обнаружению характерных черт («событий») и устанавливая их *геометрическую* связь с параметрами модели объекта для последующей локализации и идентификации объекта на изображении. При этом о содержательной яркостно-геометрической модели объекта можно говорить только в том случае, если существует определённая неоднородность внутри самого изображения объекта, которая позволяет выделить некоторые характерные черты и затем выявить их пространственные отношения, которые послужат основой для локализации объекта.

Разработка методики автоматизированного конструирования процедур анализа изображений. Как было отмечено, условием применимости метода анализа свидетельств на изображениях является наличие известной яркостно-геометрической модели объекта. В литературе описан широкий спектр таких моделей — от простейших признаковых описаний до высоко специализированных и изощренных структурных моделей. Однако общий метод составления работоспособных моделей для новых классов объектов до сих пор не известен. Отсюда следует, что разработка и использование яркостно-геометрических моделей, пригодных для эффективного решения задачи обнаружения соответствующих объектов, в значительной степени остается на грани науки и искусства, т. е. требует особого *know-how* или, другими словами, знания предметной области, отражающего многолетний опыт исследований по решению частных задач. В связи с этим в ГосНИИАС была поставлена задача разработки методики полностью автоматического конструирования модульных алгоритмов анализа изображений на основе моделей, включая и автоматическую генерацию самих исходных моделей на основе базы эталонных изображений объекта.

С этой целью была разработана методика автоматизированного конструирования процедур обнаружения и идентификации объектов на основе их структурно-вероятностного описания [Буряк, Визильтер, 2002б]. В рамках данной методики модели искомых объектов описываются разработчиком при помощи специально ориентированного на работу с изображениями декларативного языка логического программирования (типа ПРОЛОГ). После того как модель описана, осуществляется автоматическое преобразование исходного описания модели к одной из известных системе форм описания, соответствующих различным классам типовых алгоритмов вероятно-логического вывода на изображениях (голосование, прослеживание, привязка и т. п.), включая и класс алгоритмов, основанных на анализе свидетельств.

Также была разработана методика автоматизированного конструирования близких к оптимальным модульных процедур обнаружения и идентификации объектов на изображении с использованием *генетических алгоритмов* [Буряк, Визильтер, 2002а]. Метод основан на «генетическом отборе» информативных элементов эталонных изображений в процессе обучения. Критерием отбора является функция качества, вычисляемая по значениям точности, надежности и времени работы процедуры на изображениях обучающей выборки.

Разрабатываемые в рамках данного направления методики позволят в будущем существенно упростить и ускорить цикл разработки алгоритмического обеспечения при создании различных систем компьютерного и машинного зрения.

Разработка методов оценки движения видимых объектов в движущемся видимом поле. Системы технического зрения, использующие технологии анализа движения, как правило, предполагают выделение подвижного объекта на неподвижном

(относительно датчика) фоне. Между тем, для мобильных технических средств (ЛА, автомобили, поезда, мобильные роботы и т. п.) фон изображения сцены также изменяется со временем за счёт относительного движения приёмника изображения. Разработанный в ГосНИИАС подход к выделению движущихся объектов по признаку их движения [Визильтер и др., 2006] характеризуется следующими основными особенностями: быстрая корреляционная привязка фона по пирамиде изображений либо по набору динамически выделяемых особых точек; пытьевское морфологическое проецирование яркости; сравнение текущей межкадровой разности с накопленной разностной матрицей; опора на несколько динамически формируемых базовых кадров. Разработанные алгоритмы обеспечивают автоматическую оценку и компенсацию движения камеры относительно сцены наблюдения.

Для адаптивной настройки алгоритмов на крупно-, средне- и мелкоразмерные объекты, в зависимости от задачи наблюдения, поля зрения и текущих настроек камеры предложен так называемый модифицированный метод оптических потоков. В рамках данного метода на этапе предобработки используются временные разностно-накопительные схемы, оптический поток вычисляется по разности накопленных изображений, движущиеся элементы изображения определяются на основе оценки накопленного оптического потока. Для задачи оценивания параметров видимого движения и размеров объектов на изображениях разработана процедура повторной сглаживающей авторегрессионной фильтрации. Для вычислительно эффективной реализации процедуры анализа оптических потоков предложен способ исследования пиксельного временного сигнала на базе вычисления так называемых кратнорегрессионных псевдоспектров [Визильтер, Вишняков, 2008].

Разработка методов стереобнаружения объектов в сценах наблюдения. Для решения данного класса задач был предложен так называемый *метод дифференциального ортофото* [Желтов и др., 2002], основная идея которого заключается в следующем. По известной аналитической модели поверхности строятся ортоскопические проекции данной поверхности на некоторую удобную плоскость с использованием двух изображений, синхронно регистрируемых откалиброванной стереоскопической системой. Вычисление дифференциальной разности полученных ортоскопических изображений приводит к появлению характерных яркостно-геометрических структур в окрестности 3D-объектов, не лежащих на исследуемой поверхности. Поскольку ортоскопические проекции устраняют искажения, вызванные криволинейностью поверхности, существенная дисперсия яркости на разностном изображении проявляется только в окрестности 3D-объектов. В случае простых объектов с прямолинейными краями на разностном изображении образуются чёткие структуры с формой угла. При этом величина и форма угла являются известными функциями от положения объекта и не зависят от формы объекта. В случае объектов более сложной формы (с большим количеством краев) на разностном изображении возникает комбинация угловых структур, зависящих только от положения объекта. Поэтому алгоритм обнаружения, основанный на согласованной фильтрации угловых структур, близок в данной задаче к оптимальному. Таким образом, метод дифференциального ортотрансформирования позволяет свести задачу обнаружения 3D-объектов к задаче обнаружения на синтетическом изображении 2D-структур с предсказанными свойствами.

Разработка современных морфологических методов анализа цифровых изображений. Морфологический подход к задачам анализа данных отличается от других

схем анализа данных тем, что в качестве обязательного этапа предполагает *обоснованное* (т. е. в некотором смысле оптимальное) построение *модельного описания гипотетического (скрытого) прообраза наблюдаемых данных*. Иными словами, обязательным этапом решения любого типа задач здесь является *модельная сегментация* данных, допускающая их последующую полную или частичную *реконструкцию*. Легко заметить, что морфологический подход не является ни единственно возможным, ни наиболее объемлющим по отношению к остальным методам анализа данных, однако на сегодня он представляется единственным *универсальным*, т. е. *функционально полным* подходом, обеспечивающим возможность единообразного решения всех типов задач обработки и анализа данных, возникающих в реальных технических приложениях, а именно: *фильтрации* (преобразования) данных, *сжатия / декомпрессии* (сегментации / реконструкции) данных, *классификации* данных (тестирования гипотез, распознавания образов), *обнаружения* объектов (локализации гипотез в пределах одного наблюдения), *обучения* алгоритмов анализа данных (автоматического конструирования моделей).

В последние годы в области математической морфологии и морфологического анализа изображений вновь наметился существенный прогресс, и во многом он связан с теоретическими и практическими разработками, ведущимися в ГосНИИАС.

В рамках математической морфологии Серра был предложен ряд принципиально новых направлений [Visilter, 2000, 2002; Визильтер, 2005]. Предложен формализм «монотонной математической морфологии» (МММ), представляющий собой схему построения новых операторов математической морфологии Серра на базе произвольных неморфологических операторов при помощи специального оператора «монотонизации». Предложена математическая морфология в смысле Серра, построенная на базе преобразования Хафа и его модификаций (Хаф-морфология).

Разработан *метод анализа морфологических свидетельств* [Визильтер, 2007], объединяющий в рамках единого математического формализма процедуры анализа свидетельств и процедуры морфологического анализа изображений и таким образом позволяющий в едином ключе конструктивно и содержательно описывать логико-вероятностные и яркостно-геометрические составляющие таких подходов к анализу изображений как преобразование Хафа, обобщенное преобразование Хафа, математическая морфология Серра, морфологическое сравнение изображений Ю. П. Пытьева. Разработана теория *проективных морфологических разложений*, опирающаяся на однородные структурные модели, представляющие изображения в виде объединения независимых линейных проекций на образующие (структурные примитивы) из некоторого набора, причём коэффициенты линейной связи образующих и анализируемого образа составляют регулярный вектор морфологического разложения, характеризующий данный образ. Определены морфологические корреляционные меры и соответствующие схемы сравнения и яркостно-геометрической привязки изображений на основе их морфологических разложений. Предложен единый подход к построению алгоритмов структурной фильтрации, сжатия и сегментации цифровых изображений, а также обнаружения и идентификации объектов на изображениях, основанный на проективных морфологических разложениях [Визильтер, Желтов, 2008а, б; Визильтер, 2008в]. Предложен и обоснован новый широкий класс *проективных морфологий на базе структурной интерполяции* [Визильтер, 2008б].

В качестве наиболее общего на сегодняшний день морфологического формализма был предложен математический аппарат *критериальной проективной морфологии*

[Визильтер, 2008а], позволяющий использовать как однородные, так и неоднородные модели данных. Предложена и исследована схема построения критериальных морфологических операторов анализа изображений на основе *критерия максимальной достоверности*, включающего *критерий соответствия* проекции и проецируемого образа, *модельный критерий* проекции, характеризующий её принадлежность модели, и *структурирующий параметр*, обеспечивающий компромисс между требованиями соответствия и качества. Определены достаточные условия проективности критериальных морфологических операторов. Показано, что структурирующий параметр морфологического критерия определяет морфологическую сложность по Пытьеву соответствующей модели. Разработаны эффективные реализации критериальных морфологических фильтров на основе метода динамического программирования [Визильтер и др., 2009].

Результаты данного направления находят практическое применение в таких задачах как автоматическое и полуавтоматическое обнаружение искусственных объектов при дешифрировании аэро- и космофотоснимков, информационное обеспечение автономных ЛА и ряде других.

2. АНАЛИЗ ИЗОБРАЖЕНИЙ ЗЕМНОЙ ПОВЕРХНОСТИ НА ЭТАПЕ ПЛАНИРОВАНИЯ АВИАЦИОННЫХ ОПЕРАЦИЙ

В последние годы трёхмерное изображение местности с точной координатной привязкой характерных объектов и целей становится необходимой основой планирования авиационной операции (рис. 4). Принципиальный выигрыш от использования таких «виртуальных» 3D-представлений заключается в максимальной наглядности геопространственных данных и совершенно новых интерактивных



Рис. 4. Изображение местности с точной координатной привязкой объектов

информационных возможностях, которые обеспечивают компьютерные технологии машинной графики по сравнению с обычными бумажными картами и схемами.

Помимо этого, рассматриваемые в данном разделе технологии реконструкции поверхностей и обнаружения трёхмерных структур применяются для создания высокоточных метрических геопространственных данных при навигации и наведении ЛА, для точной координатной привязки целей, для использования стереоскопических демаскирующих факторов с целью повышения вероятности обнаружения объектов, а также при создании реалистических авиационных тренажеров нового поколения.

Разработанный в ГосНИИАС набор методов, алгоритмов и программных средств построения 3D-описания сцены и методов обнаружения объектов позволил создать автоматизированные средства обработки данных отечественной аэрокосмической съёмки для построения «виртуальных» моделей реальных территорий. При этом был спроектирован и реализован программно-алгоритмический пакет Z-Space, включающий:

- быструю генерацию цифровых моделей рельефа (ЦМР) в виде регулярной матрицы высот по стереопарам космических снимков или аэроснимков;
- создание цифровых ортофотопланов;
- съёмку векторных контуров по стереопарам и ортофото;
- визуализацию элементов рельефа с использованием разнообразных возможностей компьютерной графики.

В составе пакета реализованы как стандартные фотограмметрические процедуры ориентирования, геометрической и радиометрической коррекции снимков, так и новые, оригинальные процедуры автоматического выделения объектного состава. В частности, для решения задач реконструкции поверхностей и обнаружения трёхмерных структур предложен вычислительно эффективный метод поиска соответствующих (сопряженных) точек стереопары изображений трёхмерной сцены с учётом априорной информации об их приблизительном положении [Zhel'tov, Sibiryakov, 1997]. Особенность разработанного метода стереотождествления заключается в том, что соответственные точки находятся с максимальной (*субпиксельной*) точностью и снабжаются коэффициентом надёжности стереотождествления. Алгоритм стереотождествления состоит из следующих этапов:

- 1) выбор наиболее информативных участков изображений;
- 2) стереотождествление с пиксельной точностью;
- 3) стереотождествление с субпиксельной точностью (субпиксельное уточнение).

При этом большие области поиска ограничиваются за счёт применения пирамиды изображений, использования эпиполярной геометрии и априорного оценивания диапазона высоты рельефа видимой сцены. Выбор уникальных эталонов осуществляется на основе анализа информативности фрагментов изображений, тем самым преодолевается проблема сходства близлежащих эталонов. Трудности в выборе начального приближения положения соответствующих образов преодолеваются применением иерархической стратегии построения карты диспаратностей. Геометрические искажения образов устраняются геометрическими преобразованиями изображений. Яркостные искажения образов устраняются явным введением параметров искажений и их оценкой. Ложное отождествление (аномальные ошибки) выявляется с помощью анализа формы корреляционной функции. Для

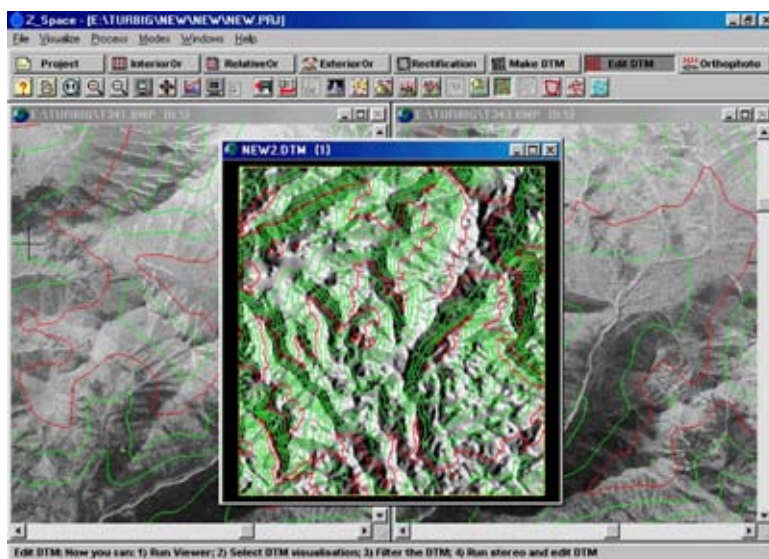


Рис. 5. Интерфейс пользователя программного пакета Z-Space

решения задачи выбора информативных участков изображений разработан новый статистический метод определения информативности на основе анализа статистических свойств шума по оптическому клину, который представляет собой изображение с плавным нелинейным изменением яркости.

На рис. 5 представлен вид рабочего окна разработчика ЦМР в пакете Z-Space при работе с космическими фотоснимками.

На рис. 6 (см. с. 30) показана типовая последовательность обработки данных авиационной съёмки для получения 3D-модели местности с объектовым составом с помощью пакета Z-Space. На рис. 7 (см. с. 30) приведён пример синтеза модели «Каир» (Египет) по данным отечественной космической съёмки камерами ТК350 (стереоданные) и КВР1000 (панорамная съёмка), а на рис. 8 (см. с. 31) — комбинация моделей, построенных по данным космических (общий план) и авиационных снимков (детали аэропорта).

Апробация разработанных технологий проводилась на комплексах полунатурного моделирования авиационных приложений в ГосНИИАС. Анализ показал, что, например, для демонстратора системы подготовки полётного задания перспективных самолётов могут быть значительно сокращены сроки подготовки полётного задания.

Необходимо также отметить роль компьютерного моделирования изображений земной поверхности в процессе разработки авиационных систем технического зрения различного назначения. В настоящее время средствами компьютерного моделирования могут быть получены достаточно реалистичные изображения любой конкретной местности, включая все необходимые объекты инфраструктуры, имитирующие результаты регистрации в видимом диапазоне длин волн. К сожалению, программное обеспечение, позволяющее осуществлять высококачественное и натуралистическое моделирование изображений других диапазонов, на рынке практически отсутствует или крайне труднодоступно. Тем не менее, даже возможность моделирования изображений видимого диапазона (или, точнее, трёхмерной

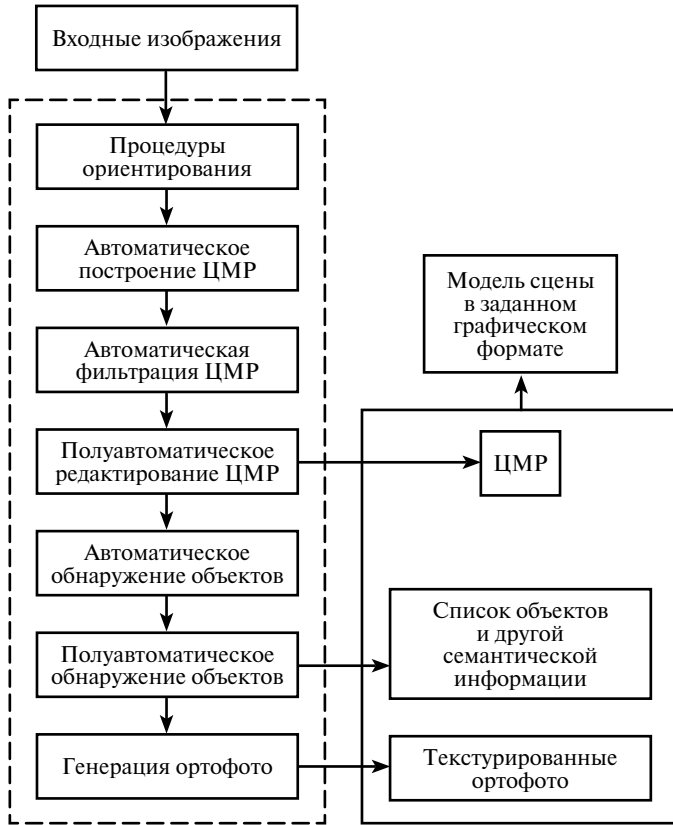


Рис. 6. Последовательность обработки данных в пакете Z-Space



Рис. 7. Результат моделирования. Каир, Гиза, район пирамид

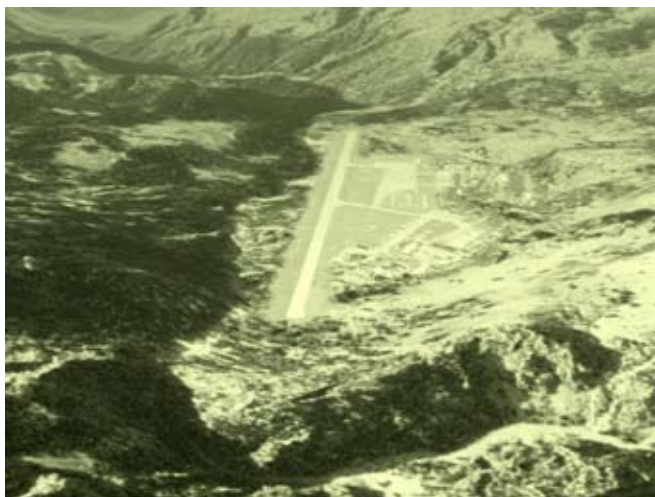


Рис. 8. Результат моделирования. Аэропорт в горах

яркостно-геометрической информации на базе таких моделей) является весьма существенным подспорьем при разработке СТЗ, которые должны функционировать в чрезвычайно широком диапазоне условий, определяемых множеством факторов: все сочетания этих факторов крайне трудно и экономически нецелесообразно пытаться зарегистрировать в реальном лётном эксперименте. Отметим хотя бы такие факторы изменчивости обстановки как различные траектории полёта, различный характер окружающего рельефа, различные погодные и другие условия видимости, наконец, различные ситуации в связи с наличием и положением различного рода подвижных объектов. Все эти факторы целесообразно моделировать, и программа моделирования должна быть достаточно обширной, чтобы гарантировать надёжность работы СТЗ в реальных условиях. Естественно, качество моделирования должно постоянно проверяться на соответствие реальным регистрациям, получаемым от датчиков в натуральных условиях. Создание практической СТЗ путём отработки на одних лишь модельных данных пока остается принципиально невозможным.

3. ПОЛУЧЕНИЕ, ОБРАБОТКА И КОМПЛЕКСИРОВАНИЕ МНОГОСПЕКТРАЛЬНЫХ ВИДЕОДАННЫХ

Как в процессе функционирования автономной системы машинного зрения, так и в работе человека-оператора автоматизированной системы управления, многие проблемы возникают вследствие естественной низкой различимости целей на оптическом изображении (дождь, туман и т. д.), либо возможного присутствия активных видов противодействия (маскировка, пиротехнические средства, прожекторы и т. п.). Значительную роль в снижении характеристик обнаружения играют также шумовые компоненты сигналов, образующиеся из-за особенностей применяемых датчиков и неидеальности каналов передачи данных. Большинство разработчиков перспективных систем управления ЛА видят основной путь борьбы с этой группой проблем в совместном использовании датчиков различной физической природы.

Изображения, получаемые от различных ДРФП, имеют свои характерные яркостно-геометрические особенности, которые обусловлены как физикой формирования самого изображения, так и характеристиками оптико-электронных трактов. Данные особенности в значительной степени определяют выбор алгоритмов предварительной обработки и распознавания. Так, существенные сложности при обработке изображений в оптическом диапазоне представляют затемнённые участки, на которых теряют эффективность локальные и градиентные алгоритмы. В то же время изображение в ИК-диапазоне обладает свойством отсутствия теней, что дает возможность выделить тени на оптическом изображении в случае совместной обработки изображений, получаемых от ТВ- и ИК-датчиков. Изображения, получаемые в миллиметровом диапазоне, характеризуются хорошим качеством в условиях дождя, тумана, облачности, однако геометрические размеры объектов сцены на таких изображениях значительно отличаются от действительных. Таким образом, комплексная обработка изображений в различных спектральных диапазонах предоставляет дополнительные возможности при автоматизации обработки и распознавания.

В табл. 1 приведены основные характеристики и наиболее известные признаки восьми двумерных датчиков, обладающих существенной разделительной способностью по отношению к типовым объектам интереса ЛА. Все перечисленные датчики могут практически использоваться при решении задач управления и выпускаются серийно по существующим технологиям.

Таблица 1

Тип датчика	Формат представления данных	Полезные признаки
Тепловизионный датчик	2D-тепловое изображение	Форма, max / min эмиссия, количество и расположение горячих пятен, окружение (среда)
Радар миллиметрового диапазона	<ul style="list-style-type: none"> • 1D-отражённый профиль • 1D или 2D-поляризационные изображения • 2D-поле скоростей 	<ul style="list-style-type: none"> • Распределение и степень протяженности рассеивателей • Чётный и нечётный номер и расположение упругих рассеивателей • Частоты пульсации и ширина пучка рассеивания
Лазерный локатор	<ul style="list-style-type: none"> • 3D-изображение • доплеровская модуляция (вибрация) • 2D-поле скоростей 	<ul style="list-style-type: none"> • Размер, 3D-форма, расположение объектов • Пульсация, структурная и поверхностная частоты • Пространственное распределение движущихся участков сцены
Локатор с синтезированным раскрывом	2D-изображение	Размер и расположение объектов
Телевизионный датчик	2D-полутонное видео-изображение	Форма, размеры, текстура, внутренняя структура объектов, окружение
Микроволновый радар	<ul style="list-style-type: none"> • Доплеровская модуляция • 2D-изображение 	<ul style="list-style-type: none"> • Скорость, частота пульсации и ширина пучка • Размер, формат, количество и расположение объектов
Акустический датчик	Отражённый звуковой сигнал	Частоты пульсации, гармоники отношения частот, источники специфического шума
Интерферометр	Спектральная и временная зависимость микроволнового излучения	Частота, частотная модуляция, амплитудная модуляция, продолжительность пульсации, интервалы пульсации

Для экспериментальной съёмки ДРФП в ГосНИИАС был создан целый ряд специальных аппаратно-программных стендов. Ниже приводятся примеры различных многоспектральных регистраций (телевизионные, лазерно-локаторные, инфракрасные). На рис. 9 представлены ТВ-, ЛЛ- и ИК-регистрации движущегося объекта на малой дальности (150 м). Пример демонстрирует надежную селекцию объекта на ЛЛ-изображении, а также яркое пятно от выхлопных газов двигателя в ИК-диапазоне. На рис. 10 показаны ТВ-, ЛЛ- и ИК-изображения, полученные при наблюдении объекта на средней дальности (400 м). На ЛЛ-изображении селектируется только подвижный объект. На рис. 11 (см. с. 34) представлены ТВ-, ЛЛ- и ИК-регистрации изображений объекта на большой дальности (2000 м). Объект наблюдается на изображениях всех трёх датчиков, но форма объекта визуально не различима. На рис. 12 (см. с. 34) представлены изображения городской сцены в ТВ- и двух ИК-диапазонах (3...5 и 8...14 мкм).

Как видно из приведенных примеров, изображения, получаемые от датчиков различных диапазонов, имеют существенную корреляцию из-за естественных взаимосвязей между физическими величинами, которые измеряются этими датчиками. Это означает, что со статистической точки зрения данные, формируемые различными каналами многоканальных систем дистанционного наблюдения, не могут рассматриваться как независимые источники свидетельств об объектах сцены наблюдения. Тем не менее, каждый диапазон несет и определённую специфическую информацию об объекте наблюдения, что позволяет системам машинного

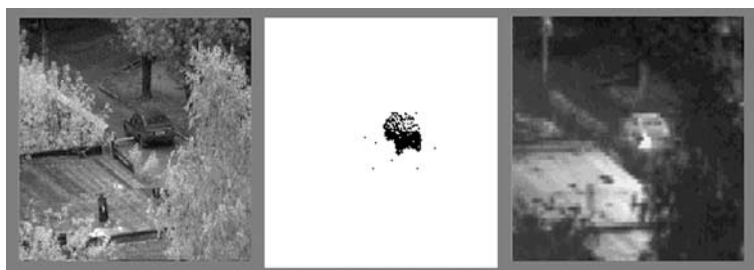


Рис. 9. ТВ-, ЛЛ- и ИК-изображения движущегося объекта на малой дальности (150 м)



Рис. 10. ТВ-, ЛЛ- и ИК-изображения объекта на средней дальности (400 м)

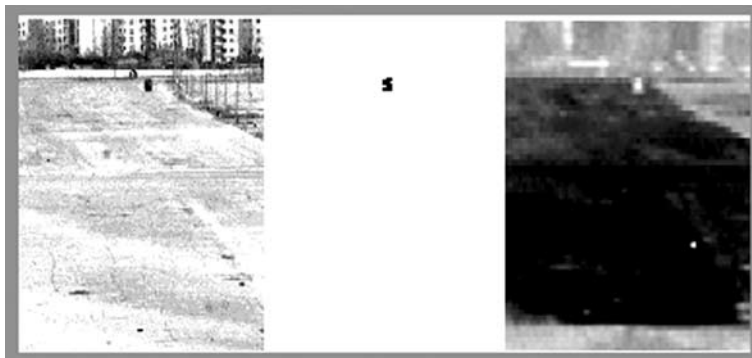


Рис. 11. ТВ-, ЛЛ- и ИК-изображения объекта на большой дальности (2000 м)



Рис. 12. Изображения городской сцены в ТВ- и двух ИК-диапазонах (3...5 и 8...14 мкм)

зрения, использующим несколько различных датчиков, демонстрировать лучшие результаты по сравнению с системами, основанными на одном спектральном диапазоне.

Комплексирование данных различной физической природы может осуществляться на основе логического или вероятностного объединения поканальных решений, полученных в ходе отдельного анализа информации, поступающей от каждого датчика. Известны также схемы комплексирования, основанные на классификации объединенных наборов признаков, формируемых по совокупности признаков, выделяемых в ходе анализа изображений разных каналов. Однако большинство современных подходов к построению многоканальных систем машинного зрения предполагают уже не «комплексирование» многозональной информации в прежнем, узком, смысле, а сложный многоэтапный процесс *совместного анализа данных*, на различных этапах которого в центре рассмотрения оказываются изображения от различных датчиков, происходит их перекрестный опрос, запрашиваются различного рода подтверждения и уточнения и т. д. Разработанный в ГосНИИАС оригинальный подход к классификации алгоритмов и структур для работы с данными различной физической природы получил название ПОМИ (*платформа обработки многоспектральной информации*) [Степанов и др., 1995, 1996]. Он позволяет в едином ключе описывать, анализировать и разрабатывать наиболее широкий спектр схем обработки и комплексирования многозональной информации. Для того чтобы описать общую структуру ПОМИ, необходимо сначала определить набор типов данных, а затем, опираясь на них, — набор типов процедур обработки.

В последние годы в области комплексирования данных существует принятая концепция *семантических уровней* представления данных, включающая *уровень измерений, признаковый уровень, уровень оценок достоверности, символичный уровень*. Упомянутые *оценки достоверности* представляют собой векторы вероятностного или нечёткого описания принадлежности объекта некоторому классу или типу объектов. В то же время любой тип данных в вычислительных системах есть совокупность *элементов*, организованная в некоторую *структуру*. Соответствующая схема *уровней организации данных* включает: растровые данные, 2D-структурированные данные, 3D-структурированные данные, а также данные, структурированные во времени.

Таким образом, любой практический тип данных может быть описан путём использования двух «координат» — *семантического уровня и организационного уровня*. Комплексирование также может выполняться на любом уровне первой схемы и одновременно на любом уровне второй схемы. Перечислим вкратце, какие виды процедур обработки и комплексирования можно определить для представленного таким образом набора типов данных. *Семантические межуровневые процедуры* сохраняют структуру, изменяя тип содержащихся в ней элементов. *Структурные межуровневые процедуры* сохраняют типы элементов при изменении структурной организации данных. *Процедуры попарного комплексирования* объединяют две входные структуры в одну выходную структуру того же типа. Необходимо также выделить два различных типа комплексирования: *комплексирование элементов и комплексирование структур*. Эти процедуры должны применяться для каждой из структур данных, поддерживаемых системой обработки данных различной физической природы, с использованием следующих процедур комплексирования структур: объединение, пересечение и ранговое комплексирование структур. *Процедуры фильтрации* являются внутриуровневыми процедурами с одним входом и одним выходом. Они всегда сохраняют размерность элементов данных. В их число входят: *преобразование элементов* — преобразование каждого элемента единообразным способом, часто используются на измерительном и признаковом уровнях, например, гистограммные преобразования полутоновых изображений или отображение признаков пространств; *линейные и нелинейные преобразования* — обработка растровых данных в пространственной или частотной областях; *геометрические преобразования* — преобразуют любые пространственные данные, используя некоторую геометрическую модель. Может быть использован широкий набор таких процедур, в частности, они могут применяться для обеспечения условий взаимной привязки и синхронизации данных при регистрации.

В частности, на основании проводившихся в ГосНИИАС численных экспериментов по комплексированию многоспектральных данных на уровне векторов достоверности оказалось возможным сделать следующие выводы.

1. Комплексирование всегда обеспечивает более высокую достоверность распознавания, чем каждый из каналов в отдельности. При отказе одного из каналов решение практически принимается по одному оставшемуся каналу.
2. При комплексировании по методу Демпстера–Шафера [Dempster, 1967; Shafer, 1976] больший вес придается показаниям канала с меньшим коэффициентом неуверенности.
3. В присутствии помех высокой интенсивности комплексирование на уровне векторов достоверности позволяет принять решение даже тогда, когда ни по одному из каналов в отдельности его принять невозможно.

Последний вывод можно проиллюстрировать следующим числовым примером (табл. 2).

Таблица 2

Диапазон	k_1	k_2	$S_f(1)$	$S_f(2)$	$S_f(\theta)$
ТВ	0,303	0,345	0,161	0,184	0,655
ИК	0,469	0,549	0,253	0,296	0,451
ТВ+ИК	0,568	0,674	0,308	0,366	0,326

Здесь k_1 и k_2 — исходные значения оценок принадлежности наблюдаемого объекта одному из двух близких классов, формируемых независимо по каждому спектральному каналу алгоритмами обнаружения и идентификации объектов. После этого на их основе рассчитывается нормированный вектор оценок достоверности S_f . Как видно, в данном примере коэффициент неуверенности распознавания типа цели $S_f(\theta)$ по каждому из каналов в отдельности много больше оценки достоверности принадлежности цели обоим классам ($S_f(1)$ и $S_f(2)$ соответственно). В то же время по результатам комплексирования мы вправе принять решение об обнаружении объекта класса 2, так как достоверность распознавания объекта этого класса оказывается больше, чем достоверность неопределённого решения: $S_f(2) = 0,366 > 0,326 = S_f(\theta)$.

4. УЛУЧШЕННОЕ ВИДЕНИЕ ДЛЯ САМОЛЁТОВ ГРАЖДАНСКОЙ АВИАЦИИ

В настоящее время на рынке присутствует ряд известных систем улучшенного видения (EVS, EFVS).

Компания Kollsman, Inc. (США) предлагает два EVS-решения — Kollsman All Weather Window (рис. 13, см. с. 37) и Kollsman Night Window. Первая система предназначена для работы во всепогодных условиях. В состав системы входят: охлаждаемый ИК-датчик (cooled FLIR), спецвычислитель, дисплей с подогревом, ПО обработки и визуализации изображений. Система Kollsman Night Window представляет собой более компактное и дешёвое решение, предназначенное для работы в тёмное время суток, но при хороших погодных условиях. В состав системы входит неохлаждаемый микроболометр, работающий в диапазоне 8...14 мкм.

Компанией Gulfstream Aerospace Corporation разработана система EVS (рис. 14, см. с. 37) в составе: ИК-датчика Kollsman FLIR с сапфировым экраном; ИЛС компании Honeywell, оригинального ПО обработки и визуализации изображений.

Компания CMC Electronics (Канада) предлагает два вида систем: CMA-2600 I-Series™ и более компактный и дешёвый вариант CMA-2610 M-Series™. В I-Series используется охлаждаемый, двухдиапазонный (1...3 и 3...5 мкм) ИК-датчик, в случае M-Series — неохлаждаемый, работающий в диапазоне 8...14 мкм. В стадии разработки находится система на базе миллиметрового радара. Индикация изображений выполняется на ИЛС и/или на многофункциональном индикаторе (дисплее). Также может устанавливаться подсистема синтезированного зрения (SVS), предоставляющая данные о рельефе местности.

Компания Max-Viz, Inc. (США) предлагает систему EVS 2500 на базе двух ИК-датчиков: длинноволнового — для формирования изображений рельефа



Рис. 13. Примеры работы системы Kollsman All Weather Window: EVS View — синтезированное изображение; Window View — вид из кабины пилота



Рис. 14. Примеры работы системы Gulfstream EVS. Показаны изображения, формируемые EVS (with EVS), и вид из кабины пилота без EVS (without EVS)



Рис. 15. Пример работы системы Rockwell Collins EVS

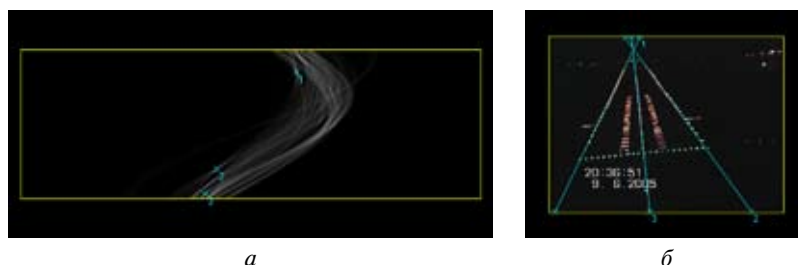


Рис. 16. Пример обнаружения элементов ВПП на основе преобразования Хафа: *a* — изображение аккумулятора; *б* — выделенные прямые линии

и потенциальных препятствий и коротковолнового — для обнаружения ярких сигнальных огней ВПП. Информация с обоих датчиков поступает на спецвычислитель, где комплексировается и выдается на МФИ лётчику. Благодаря использованию неохлаждаемых ИК-датчиков EVS 2500 имеет компактные размеры, небольшую массу и достаточно проста в установке. Кроме того, отсутствие криогенной системы охлаждения значительно уменьшает её стоимость. Компания также предлагает более дешёвое решение на базе одного неохлаждаемого ИК-датчика EVS 1000 для вертолётов и бизнес-самолётов.

Система Rockwell Collins EVS (рис. 15, см. с. 37) формирует ИК-изображения внешней среды, которую видит пилот, и выводит их на дисплей навигационной системы на лобовом стекле HGS[®]-4000, разработки этой же компании. Интерфейс EVS-системы обеспечивает взаимодействие ИК-датчика и навигационной системы HGS с целью предоставления пилоту оптимального по качеству ИК-изображения.

Следует отметить, что все перечисленные сертифицированные коммерческие системы EVS первого поколения представляют собой простые неинтеллектуальные системы датчик — дисплей, обеспечивающие передачу на индикатор в кабине изображения, полученного от датчиков. Только некоторые системы предоставляют возможность визуализации синтезированных географических данных с использованием GPS. Лишь в одной из систем используется комплексирование изображений от различных датчиков, причём оба датчика ИК (различных диапазонов). Задача комплексирования и отображения данных, одновременно поступающих от ТВ, ИК, МР, ЛЛ, не рассматривается и не решается ни в одной из этих систем. Также ни в одной из коммерческих систем EVS не рассматриваются задачи автоматического выделения ВПП, препятствий и распознавания объектов из базы геопространственных данных. Тем более не рассматриваются задачи стыковки EVS с системой автоматического управления посадкой. В то же время известно, что работы по всем указанным направлениям активно ведутся. Таким образом, на повестке дня сейчас стоит создание EVS следующего поколения, представляющих собой авиационные системы многоспектрального технического зрения и улучшенной компьютерной визуализации закабинной обстановки, существенно превосходящие ныне известные EVS гражданского назначения по набору функций комплексирования, и особенно — в части интеллектуальной обработки видеоданных.

Важнейшую роль в таких EVS нового поколения должна играть СТЗ, выполняющая следующие основные функции:

- получение и оцифровка многоспектральной видеоинформации;
- комплексирование многоспектральной видеоинформации;
- автоматическая привязка оперативной видеоинформации к априорной информации о закабинной обстановке с учётом имеющихся навигационных данных;
- автоматическое обнаружение ВПП и других типовых объектов интереса в отсутствие навигационных данных и априорной информации о закабинной обстановке;
- автоматическое обнаружение препятствий на ВПП.

В состав аппаратного обеспечения СТЗ могут входить: телевизионные видеодатчики (ТВ) различного разрешения; инфракрасные (ИК) видеодатчики различных диапазонов; миллиметровые радары (МР); лазерные локаторы (ЛЛ); специализированные бортовые вычислители.

В состав программно-алгоритмического обеспечения перспективной бортовой многоспектральной системы технического зрения должны входить следующие основные модули: управления видеодатчиками различных диапазонов; обмена информацией с другими бортовыми подсистемами; организации информационного обмена, синхронизации захвата и обработки многоспектральных видеоданных; раздельной поканальной обработки видеоданных; комплексирования многоспектральной видеоинформации; автоматической привязки оперативной двумерной видеоинформации к априорной двумерной или трёхмерной информации с учётом текущих навигационных данных; автоматического обнаружения ВПП и других типовых объектов интереса в отсутствие навигационных данных и априорной информации; автоматического обнаружения препятствий на ВПП. На протяжении последних лет необходимые для решения перечисленных задач методы, подходы и алгоритмы были реализованы, опробованы и протестированы сотрудниками ГосНИИАС. В частности, на рис. 16 (см. с. 37) приведён пример автоматического обнаружения ВПП с использованием модифицированного преобразования Хафа.

Накопленный к настоящему моменту практический задел позволяет с уверенностью говорить о том, что реализация описанных модулей перспективной СТЗ для EVS действительно возможна и должна обеспечить необходимые характеристики функционирования такой системы.

5. ПРОГРАММНАЯ АРХИТЕКТУРА ДЛЯ ОБРАБОТКИ И АНАЛИЗА ИЗОБРАЖЕНИЙ

Опыт проводившихся в ГосНИИАС разработок СТЗ для различных ЛА показывает, что в программно-аппаратных системах, включающих большое количество датчиков и других источников информации, а также модулей обработки информации, которые должны работать согласованно в синхронном или асинхронном режиме, важную роль играет использование специальной программной платформы, которая позволяет рассматривать все эти модули как своего рода «конструктор». В таком случае разработчики в состоянии легко менять конфигурацию системы, подключать и отключать различные датчики, «на лету» изменять и пробовать различные схемы обработки данных, быстро формировать необходимые отчёты о проводимых исследованиях. При этом в случае одновременной работы с несколькими изображениями необходимо особенно наглядным и удобным образом представлять в системе разработки как сами исходные изображения, так и различные этапы их обработки, предоставляя все необходимые инструменты для их просмотра и исследования. Для выполнения этих функций в ГосНИИАС была разработана специальная технология модульного визуального программирования, которая используется уже более полутора десятилетий и многократно доказала свою эффективность.

Разработанная программная технология была названа *фрейм-ориентированным программированием* (ФОП) [Visilter, 2002; Визильтер, 2005]. ФОП отличается следующими основными особенностями. Базовым элементом программной архитектуры являются программные фреймы — объекты, объединяющие данные, процедуры и линки (связи с другими фреймами). Любая схема обработки в этой архитектуре представляется в виде сети фреймов. Фреймы принимают и порождают сообщения, которые распространяются по линкам (связям) и определяют порядок «запуска» соответствующих процедур обработки данных. Такая структура позволяет организовывать автоматическую обработку данных по мере её поступления

независимо от числа источников информации и степени их временного рас- согласования. Обеспечивается динамический механизм защиты данных. Любые внешние для фрейма данные доступны для чтения и записи только посредством «входных» и «выходных» информационных линков и только тогда, когда фреймы — «хозяева данных» — разрешают такой доступ. Три принципиальных типа фреймов необходимы для полноценного функционирования сети фреймов: фреймы данных, фреймы обработки и фреймы управления. Конкретный набор типов фреймов данной предметной области полностью определяет функциональные возможности проектируемых программных систем. После создания базового на- бора фреймов любые схемы обработки данных могут быть сформированы непо- средственно путём графического визуального программирования.

Программные системы, разработанные с использованием фреймовой техно- логии, обладают оригинальным фреймовым интерфейсом пользователя; возмож- ностью интерактивно формировать из имеющихся готовых блоков любые схемы обработки без использования какого-либо специального командного языка (пу- тём фреймового визуального программирования); осуществлять автоматическую обработку данных в заданной схеме по мере их обновления.

Фреймовый интерфейс пользователя обеспечивает: адаптивную настройку меню системы в зависимости от её текущего состояния, текущего активного окна и характера выполняемых действий; адаптивную настройку режима ввода и выво- да интерактивной векторной графической информации (маркеры, линии, кривые, прямоугольники и т. п.) во всех необходимых окнах одновременно; удобную про- блемно-ориентированную навигацию в многооконном интерфейсе; адаптивную настройку системы помощи в зависимости от текущего состояния системы, теку- щего активного окна и характера выполняемых действий.

На рис. 17 показан интерфейс фреймовой системы обработки информации ДРФП версии 1997 г. На рис. 18 (см. с. 41) приведён пример пользовательского интерфейса среды обработки видеoinформации версии 2009 г.

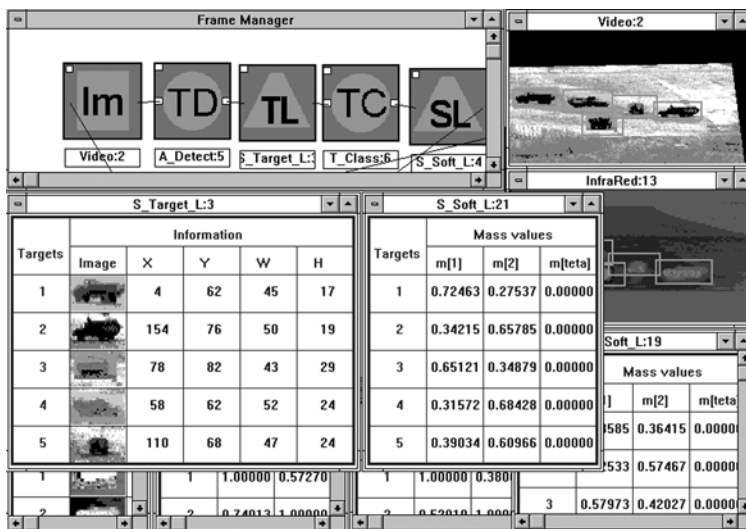


Рис. 17. Пример интерфейса фреймовой системы обработки информации ДРФП (1997)

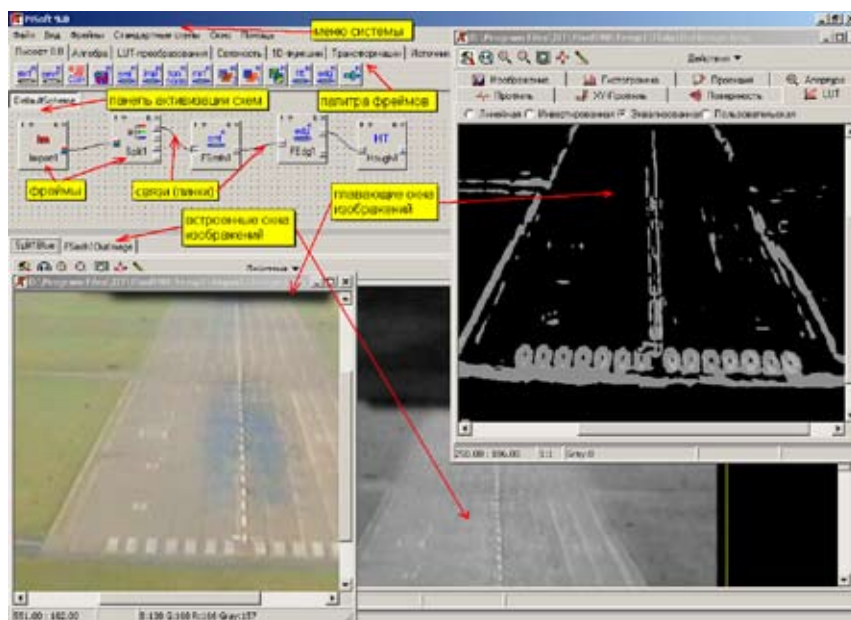


Рис. 18. Пример интерфейса фреймовой системы обработки изображений (2009)

Среда визуального программирования была реализована на языке высокого уровня Borland Delphi для работы в среде Windows. Реализованные на сегодняшний день типы фреймов позволяют осуществлять комплексный анализ и обработку изображений различной физической природы и в полной мере реализовать концепцию визуального программирования.

Заключение

Данная статья посвящена проблематике технического зрения в современных авиационных системах. При этом рассматривались в основном наиболее общие методические и теоретические аспекты. Между тем, большинство упомянутых в данной работе методов и технологий могут применяться не только при создании систем управления и информационного обеспечения ЛА, но и при решении множества других задач.

За последние годы в ГосНИИАС было создано значительное число систем компьютерного и машинного зрения в таких областях как бесконтактные измерения и технический контроль, биометрические системы персональной идентификации, автоматизированное управление наземными транспортными средствами, автоматизация учёта и документооборота с использованием цифробуквенной и штрихкодовой машиночитаемой информации, медицинские системы анализа цифровых изображений (радиологических, микроскопических), системы интеллектуального видеонаблюдения и ряд других.

Таким образом, описанные в работе методы и технологии машинного зрения являются достаточно универсальными и помимо авиационных приложений находят практическое применение в самых различных областях.

Литература

- Буряк Д. Ю., Визильтер Ю. В. (2002а) Автоматизированное конструирование близких к оптимальным процедур идентификации и обнаружения объектов на изображениях с использованием генетических алгоритмов // *Материалы 12-й Международ. конф. по компьютерной графике и машинному зрению (Графикон'2002)*. Нижний Новгород, 16–21 сент. 2002. С. 292–298.
- Буряк Д. Ю., Визильтер Ю. В. (2002б) Возможности применения генетических алгоритмов для автоматизированного конструирования процедур анализа изображений // *6-я Международ. конф. «Распознавание образов и анализ изображений: новые информационные технологии» (РОАИ-6-2002)*. Великий Новгород, РФ. 21–26 окт. 2002: Тр. конф. Великий Новгород: НовГУ им. Ярослава Мудрого, 2002. С. 87–92.
- Василенко Г. И., Цибулькин Л. М. (1985) Голографические распознающие устройства. М.: Радио и связь, 1985. 312 с.
- Визильтер Ю. В. (2005) Метод компьютерного анализа изображений, основанный на объединении морфологических свидетельств // Тр. ГосНИИАС. Сер. «ИТ». 2005. № 2.
- Визильтер Ю. В. (2007) Применение метода анализа морфологических свидетельств в задачах машинного зрения // *Вест. компьютерных и информац. технологий*. 2007. № 9. С. 11–18.
- Визильтер Ю. В. (2008а) Обобщенная проективная морфология // *Компьютерная оптика*. 2008. Т. 32. № 4.
- Визильтер Ю. В. (2008б) Проективные морфологии на базе интерполяции // *Вестн. компьютерных и информац. технологий*. 2008. № 4. С. 11–18.
- Визильтер Ю. В. (2008в) Структурная фильтрация цифровых изображений с использованием проективных морфологий // *Вестн. компьютерных и информац. технологий*. 2008. № 3. С. 18–22.
- Визильтер Ю. В., Вишняков Б. В. (2008) Исследование поведения авторегрессионных фильтров в задаче выделения и анализа движения на цифровых видеопоследовательностях // *Вестн. компьютерных и информац. технологий*. 2008. № 8. С. 2–8.
- Визильтер Ю. В., Желтов С. Ю. (2008а) Проективные морфологии и их применение в структурном анализе цифровых изображений // *Изв. РАН. ТиСУ*. 2008. № 6. С. 113–128.
- Визильтер Ю. В., Желтов С. Ю. (2008б) Сравнение и локализация фрагментов изображений с использованием проективных морфологий // *Вестн. компьютерных и информац. технологий*. 2008. № 2. С. 14–22.
- Визильтер Ю. В., Желтов С. Ю., Ларетина Н. А. (2009) Проективные морфологии на базе операторов фильтрации и сегментации изображений, вычисляемых методом динамического программирования // *Вестн. компьютерных и информац. технологий*. 2009. № 6.
- Визильтер Ю. В., Лагутенков А. В., Ососков М. В., Выголов О. В., Блохинов Ю. Б. (2006) Выделение и межкадровое прослеживание движущихся объектов при регистрации изображений сложных пространственных сцен произвольно движущимися двумерными сенсорами // *Вестн. компьютерных и информац. технологий*. 2006. № 3. С. 34–39.
- Желтов С. Ю., Визильтер Ю. В., Князь В. А., Выголов О. В. (2002) Обнаружение препятствий в системе управления транспортным средством // *Докл. 3-й Международ. конф. «Цифровая обработка информации и управление в чрезвычайных ситуациях»*. Минск, 2002. Т. 2. С. 121–126.
- Красовский А. А., Белоглазов И. Н., Чигин Г. П. (1979) Теория корреляционно-экстремальных навигационных систем. М.: Наука, 1979. 448 с.
- Левшин В. Л. (1987) Биокрибернетические оптико-электронные устройства автоматического распознавания изображений. М.: Машиностроение, 1987. 176 с.
- Март Д. (1987) Зрение. Информационный подход к изучению представления и обработки зрительных образов. М.: Радио и связь, 1987. 400 с.
- Пытёв Ю. П. (1983) Морфологический анализ изображений // *Докл. АН СССР*. 1983. Т. 269. № 5. С. 1061–1064.

- Пытывев Ю. П. (1984) Задачи морфологического анализа изображений // Математич. методы исслед. природных ресурсов Земли из космоса. М.: Наука, 1984. С. 41–83.
- Ярославский Л. П. (1987) Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. М.: Радио и связь, 1987. 296 с.
- Burns J., Hanson A., Riseman E. (1986) Extracting Straight Lines // IEEE Trans. Pattern Analysis and Machine Intelligence. 1986. V. 8. N. 4.
- Canny J. (1986) A Computational Approach to Edge Detection // IEEE Trans. Pattern Analysis and Machine Intelligence. 1986. N. 8. P. 679–698.
- Davies E. R. (1987) Machine Vision: Theory, Algorithms, Practicalities. 2nd ed. San Diego: Academic Press, 1997. 750 p.
- Dempster A. P. (1967) Upper and Lower Probabilities Induced by a Multivalued Mapping // Annals of Mathematical Statistics. 1967. V. 38. P. 325–339.
- Duda R. O., Hart P. E. (1972) Use of the Hough Transformation to Detect Lines and Curves in Pictures // Communications of the ACM. 1972. V. 15. P. 11–15.
- Forstner W. (1995) Mid-level vision processes for automatic building extraction, Automatic Extraction of Man-Made Objects from Aerial and Space Images. Basel: Birkhauser Verlag, 1995.
- Grimson W. E. L. (1990) Object Recognition by Computer: The Role of Geometric Constraints. Cambridge, USA: MIT Press Bookstore, 1990.
- Gruen A., Baltsavias E. (1985) Adaptive Least Squares Correlation with Geometrical Constraints // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. 1985. V. 595.
- Gruen A., Baltsavias E. (1988) Geometrically Constrained Multiphoto Matching // Photogrammetric Engineering and Remote Sensing. 1988. V. 54. N. 5. P. 633–641.
- Haralick R. M., Chu Y. H. (1984) Solving Camera Parameters from the Perspective Projection of a Parameterized Curve // Pattern Recognition Letters. 1984. N. 17(6). P. 637–695.
- Hough P. V. C. (1962) Methods and Means for Recognizing Complex Patterns. U. S. Patent 3069654, 1962.
- Huertas A., Nevatia R. (1988) Detecting Buildings in Aerial Images // Computer Vision, Graphics and Image Processing. 1988.
- Illingworth J., Kittler J. (1987) The Adaptive Hough Transform // IEEE Trans. Pattern Analysis and Machine Intelligence. 1987. V. 9. P. 690–698.
- Kitchen L., Rosenfeld A. (1982) Gray-Level Corner Detection // Pattern Recognition Letters. 1982. N. 1. P. 95–102.
- Marr D., Hildreth E. (1980) Theory of Edge Detection // Proc. Royal Soc. B: Biological Sciences (London). 1980. V. 207. N. 1167. P. 187–217.
- Mumford D. (1987) The Problem of Robust Shape Descriptors // Proc. 1st Intern. Conf. Computer Vision. 1987. P. 602–606.
- Schenk A. F. (1996), Automatic Generation of DEM's, Digital Photogrammetry: An Addendum to the Manual of Photogrammetry // American Soc. for Photogrammetry and Remote Sensing, 1996.
- Seeger U., Seeger R. (1994) Fast Corner Detection in Grey-Level Images // Pattern Recognition Letters. 1994. N. 15(7). P. 669–675.
- Serra J. (1982) Image Analysis and Mathematical Morphology. L.: Academic Press. 1982.
- Serra J. (1986) Introduction to mathematical morphology // Computer Vision, Graphics and Image Processing. 1986. V. 35. N. 3.
- Shafer G. (1976) A Mathematical Theory of Evidence. Princeton University Press, 1976.
- Stepanov A. A., Visilter Yu. V., Morzeev Yu. V., Zheltov S. Yu. (1995) The Object-Oriented Frame Approach to Image Processing and Data Management in the Multisensory Remote Sensing // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. 1995. V. 2587.
- Stepanov A. A., Visilter Yu. V., Morzeev Yu. V., Zheltov S. Yu. (1996) Frames Based Analysis of Multisensor Image Sequences // ISPRS Proc. Intern. Archives of Photogrammetry and Remote Sensing. 1996. V. 31. Pt. B2.

- Vizilter Yu.* (2000) Image Analysis Using Select-Only Morphological Operators // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. Boston. 2000. V. 4197. P. 107–118.
- Vizilter Yu.* (2002) Design of Morphological Operators Based on Selective Morphology // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. Sun Jose. 2002. V. 4667. P. 215–226.
- Vizilter Yu., Zheltov S., Stepanov A.* (1996a) Events-based Image Analysis for Machine Vision and Digital Photogrammetry // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. 1996. V. 31. Pt. 5.
- Vizilter Yu., Zheltov S., Stepanov A.* (1996b) Object Detection and Recognition using Events-based Image Analysis // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. 1996. V. 2823. P. 184–195.
- Zheltov S. Yu., Blochinov Yu. B., Stepanov A. A., Sibiriyakov A. V.* (1997) Computer 3D Site Model Generation Based on Aerial Images // SPIE Proc. Intern. Archives of Photogrammetry and Remote Sensing. 1997. V. 3084.
- Zheltov S. Y., Sibiriyakov A. V.* (1997) Adaptive Subpixel Cross-correlation in a Point Correspondence Problem // Optical 3-D Measurement Techniques IV. Heidelberg: Wichmann Verlag, 1997. P. 86–95.

PROBLEMS OF TECHNICAL VISION IN AVIATION SYSTEMS

Yu. V. Vizilter, S. Yu. Zheltov

*Federal State Unitary Enterprise State Scientific Research Institute for Aviation Systems
(GosNIIAS), 125319, Moscow, 7 Victorenko st.*

The paper describes common problems of design of technical vision systems for current and prospective aviation systems. Types of image sensors are considered. Basic requirements for vision algorithms are listed and characterized. Problems and methods of image analysis are discussed in context of automatic target recognition, aviation operations' planning, multi-spectral data acquisition and fusion. The class of enhanced flight vision systems for civil aviation is addressed. The brief summary of some theoretical and practical results obtained in GosNIIAS in the field of technical vision is included.

Keywords: aviation, technical vision, image processing, image analysis.

Vizilter Yury Valentinovich — Head of the Laboratory of Computer and Machine vision, Doctor of Physical and Mathematical Sciences, Senior Scientist, e-mail: viz@gosniias.ru.

Zheltov Sergey Jurevich — Corresponding Member of the Russian Academy of Sciences, Doctor of Technical Science, General Director, Professor.

**ТЕХНОЛОГИИ СИСТЕМ
УЛУЧШЕННОГО / СИНТЕЗИРОВАННОГО ЗРЕНИЯ
ДЛЯ УПРАВЛЕНИЯ ЛЕТАТЕЛЬНЫМИ АППАРАТАМИ
Л. Н. Костяшкин¹, С. И. Бабаев², А. А. Логинов¹, О. В. Павлов¹**

¹ Федеральное государственное унитарное предприятие «Государственный рязанский приборный завод» (ФГУП «ГРПЗ»), 390000, Рязань, Семинарская, 32

² Государственное образовательное учреждение высшего профессионального образования «Рязанский государственный радиотехнический университет» (ГОУ ВПО «РГРУ»), 390005, Рязань, Гагарина, 59/1

В работе приводится обзор способов интеграции информации от бортовых систем технического зрения летательного аппарата. Рассматриваются вопросы возникновения и компенсации геометрических искажений изображений. Приводятся технологии улучшения и комплексирования теле- и тепловизионных изображений, а также способы совмещения радиолокационных, теле-тепловизионных изображений и цифровой карты местности.

Ключевые слова: системы технического зрения, цифровая карта местности, геометрические искажения, улучшенное изображение, совмещение изображений, корреляционно-экстремальные навигационные системы.

Введение

Необходимость повышения безопасности полётов летательных аппаратов (ЛА), обеспечения безопасного захода на посадку и приземления требует реализации новых способов интеграции и интерпретации информации от бортовых систем технического зрения (СТЗ) различных спектральных диапазонов, навигационных приборов и цифровых карт местности (ЦКМ).

С одной стороны, данные от СТЗ могут быть использованы для формирования управляющих воздействий в автоматизированном навигационном комплексе для коррекции текущих координат местоположения ЛА, что позволяет разгрузить экипаж от ряда традиционных функций по обработке информации и управлению. С другой стороны, видеoinформация от СТЗ совместно с синтезированным изображением рельефа местности, картографической и навигационной информацией может быть представлена экипажу в реальном времени для осуществления визуального пилотирования и посадки в условиях плохой видимости, что реализуется в системах улучшенного и синтезированного зрения (Enhanced synthetic vision systems — ESVS).

В составе бортовых СТЗ на ЛА могут присутствовать телевизионные (ТВ) и тепловизионные (ТПВ) сенсоры, бортовая радиолокационная станция (БРЛС), лидар. Каждый из сенсоров СТЗ в отдельности представляет специфические характеристики окружающей обстановки (яркостные, тепловые либо радиолокационные

Костяшкин Леонид Николаевич — директор — главный конструктор Научно-конструкторского центра видеокомпьютерных технологий (НКЦ ВКТ) ФГУП «ГРПЗ», кандидат технических наук, доцент, e-mail: hunter-rzn@yandex.ru.

Бабаев Сергей Игоревич — старший преподаватель кафедры электронных вычислительных машин РГРУ, e-mail: evm@rgta.ryazan.ru.

Логинов Александр Анатольевич — главный конструктор по направлению НКЦ ВКТ ФГУП «ГРПЗ», кандидат технических наук, доцент, e-mail: hunter-rzn@yandex.ru.

Павлов Олег Вячеславович — начальник лаборатории НКЦ ВКТ ФГУП «ГРПЗ», e-mail: hunter-rzn@yandex.ru.

контрасты объектов), не обеспечивающие достаточного количества информации о подстилающей поверхности. Кроме того, изображения от этих сенсоров формируются в различных системах координат: азимут – дальность для БРЛС и азимут – угол места для ТВ и ТПВ.

Система ESVS состоит из пяти компонентов: системы улучшенного зрения, средства формирования интуитивного синтезированного изображения рельефа местности, средства обнаружения и отображения опасных препятствий, средства контроля целостности и предупреждения о её нарушении, системы точной навигации.

Концепция построения ESVS зависит от типа и назначения ЛА, для которого она строится. Можно выделить следующие основные этапы реализации данной технологии.

- *Создание виртуальной модели местности (ВММ)*. ВММ формируется при предполётной подготовке на основе данных ЦКМ с учётом возможных навигационных параметров ЛА и может использоваться в качестве эталонного изображения (ЭИ) в корреляционно-экстремальной навигационной системе (КЭНС).
- *Предварительная обработка в реальном времени текущих изображений (ТИ) от СТЗ* (фильтрация, контрастирование, сегментация, выделение типовых объектов). Приведение ТИ различных сенсоров СТЗ к единому масштабу и формату. Комплексирование (слияние) ТИ различных сенсоров (например, ТВ и ТПВ или ТПВ и РЛС мм-диапазона волн) в одно улучшенное ТИ, обладающее повышенной информационной насыщенностью.
- *Совмещение (привязка) ТИ* с применением корреляционно-экстремальных методов, где в качестве эталонного изображения используется соответствующая ВММ. Здесь также возможен индивидуальный учёт объектового состава ВММ и ТИ. Результаты этого этапа имеют самостоятельное значение для коррекции координат ЛА.
- *Распознавание объектов на изображениях*, полученных от разных подсистем СТЗ, определение их характеристик и унификация с базой знаний для получения дополнительной информации об объектах такого рода.
- *Построение синтезированного изображения подстилающей поверхности* на основе совмещения ТИ и ВММ. Результатом этой задачи должно стать синтезированное изображение, на котором максимально эффективно и полно отображена вся доступная информация об окружающей обстановке ЛА, такая как рельеф местности, статические и динамические объекты с определёнными для них характеристиками.
- *Визуализация улучшенного либо синтезированного изображений закабинного пространства* совместно с навигационной информацией на бортовых индикаторах, в качестве которых могут выступать многофункциональные индикаторы на приборной панели, индикаторы на лобовом стекле и нашлемные индикаторы. В качестве конкретного примера реализации первых трёх этапов технологии рассматривается КЭНС, базирующаяся на совместной обработке радиолокационной и картографической информации с использованием поля радиолокационного контраста [Костяшкин и др., 2007]. Его использование в КЭНС эффективно как на больших высотах над относительно ровной местностью, когда рельефометрические системы не обеспечивают заданную точность, так и на малых высотах, где рельефометрические системы создают большие затенения точек профиля высот.

Основная трудность при реализации данной технологии КЭНС — обеспечение максимальной «похожести» ТИ с соответствующей ВММ и минимизация зависимости нестабильности поля радиолокационного контраста от сезонных и погодных условий, от угловой ориентации антенны. В качестве информативных участков радиолокационного изображения (РЛИ) для корреляции ТИ и ВММ используют переходы из областей с высокими значениями удельных эффективных площадей рассеяния (ЭПР) в соседние области с пониженными значениями. Форма таких участков достаточно стабильна во времени и может быть получена по ЦКМ. В этом случае необязательно знание абсолютных величин удельных ЭПР, так как все данные об информативном участке будут сосредоточены в его контурном изображении. К информационным элементам ЦКМ относятся площади, занятые населенными пунктами, лугами, лесом, дорожная и речная сеть, перепады высот рельефа.

Общая схема корреляционной привязки включает предварительную обработку ТИ для исключения избыточной информации, устранение которой не приводит к снижению вероятности и точности корреляционной привязки, а также использование критериальной функции, инвариантной к определённому геометрическому несоответствию ТИ и ВММ, что позволяет существенно сократить требования к бортовой БД для хранения ВММ.

ТЕХНОЛОГИЯ СОЗДАНИЯ УЛУЧШЕННОГО / СИНТЕЗИРОВАННОГО ИЗОБРАЖЕНИЯ

Для совместной обработки РЛИ, телевизионного и тепловизионного изображений (ТВИ и ТПВИ) при создании улучшенного/синтезированного изображения необходимо провести ряд предварительных действий над ТИ, включающих:

- коррекцию геометрических искажений;
- приведение ТИ к общей проекции;
- корреляционную привязку ТИ между собой и ЭИ.

Геометрические искажения ТИ могут быть вызваны несоответствием пространственной ориентации СТЗ ЛА при подготовке ЭИ. Пространственная ориентация СТЗ ЛА определяется его географическими координатами, высотой, а также углами тангажа, крена, курса. Схематично возникновение геометрических искажений представлено на рис. 1 (см. с. 48), плоскостям P и P' соответствуют плоскости текущего и эталонного изображений, точкам x и x' — положение одного и того же ориентира на них.

Для компенсации геометрических искажений ТВИ и ТПВИ в данном случае применяется перспективное преобразование. Координаты любой точки изображения $m(x, y)$ задаются уравнением:

$$m = PRM,$$

где P — матрица проекции; R — матрица преобразования системы координат между видами; $M(X, Y, Z)$ — координаты ориентира на местности.

Перспективное преобразование данного вида называется гомографией, а матрица $H = PR$ — матрицей гомографии.

Если географические координаты ЛА при формировании ТИ и ЭИ совпадают, то геометрические искажения вызваны различной пространственной ориентацией СТЗ ЛА по углам тангажа, крена и по курсу. На рис. 2 (см. с. 48) показан результат разанной (т. е. несовпадающей) ориентации по углам тангажа, крена и по курсу.

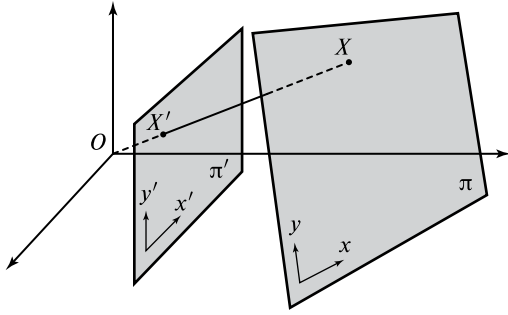


Рис. 1. Геометрические искажения изображений

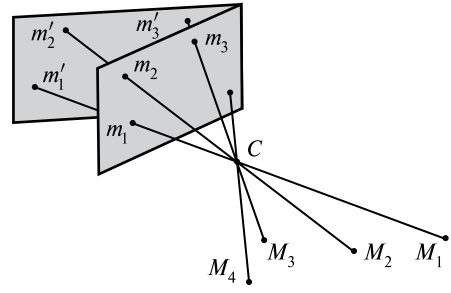


Рис. 2. Гомография плоскостей

При этом ТИ и ЭИ переводятся друг в друга перспективным преобразованием плоскости. В этом случае преобразование точек между изображениями не зависит от формы рельефа, а обусловлено только взаимным положением плоскостей изображений.

Пусть π — наблюдаемая плоскость, $H_{1\pi}$ — преобразование гомографии между плоскостью π и изображением I_1 , $H_{2\pi}$ — преобразование гомографии между плоскостью π и изображением I_2 . Тогда преобразование гомографии H_{12} — между изображениями I_1 и I_2 можно вывести следующим образом:

$$m_1 = H_{1\pi} m_\pi;$$

$$m_2 = H_{2\pi} m_\pi;$$

$$m_2 = H_{2\pi} H_{1\pi}^{-1} m_1;$$

$$H_{12} = H_{2\pi} H_{1\pi}^{-1}.$$

Преобразование H_{12} не зависит от параметризации плоскости π , а обусловлено только взаимной ориентацией в пространстве плоскостей изображений I_1 и I_2 .

Рассмотрим случай, когда положения ЛА при формировании ТИ и ЭИ не совпадают, что соответствует рис. 3.

Пусть C_1 и C_2 — положения ЛА, m_1 и m_2 — проекции точки M на первое и второе изображение соответственно. Пусть π — плоскость, проходящая через точку M и точки C_1 и C_2 . Плоскость π пересекает плоскости изображений по прямым l_1 и l_2 . Проекция любой точки M' , лежащей в плоскости π , лежат на прямых l_1 и l_2 . Эти прямые называются эпиполярными линиями [Бэр, 2000], а плоскость π — эпиполярной плоскостью (см. рис. 3).

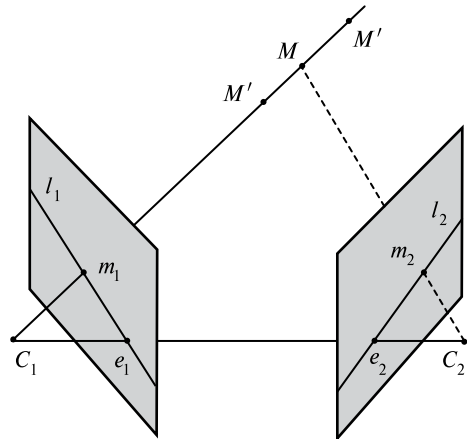


Рис. 3. Эпиполярная геометрия двух изображений

Пусть $C_1 C_2$ пересекает первое и второе изображение в точках e_1 и e_2 соответственно. Точки m_1 и m_2 называются соответствующими, если они являются проекциями одной и той же точки рельефа M . Если плоскость π проходит через точку m_1 , то лежащие в ней эпиполярные линии l_1 и l_2 называются соответствующими точке m_1 .

Ограничение на положение соответствующих точек m_1 и m_2 , вытекающее из эпиполярной геометрии, можно сформулировать следующим образом: точка m_2 , соответствующая m_1 , должна лежать на эпиполярной линии l_2 , соответствующей точке m_1 . Для соответствующих эпиполярных линий l_1 и l_2 верно:

$$l_2 \sim \left(P_2^T\right)^{-1} P_1^T l_1 = H^{-1T} l_1;$$

$$l_2 \sim H^{-1T} [e_1] m_1 = F m_1.$$

Матрица F представляет собой линейный оператор, сопоставляющий каждую точку m_1 с соответствующей ей эпиполярной линией l_2 . Для каждой пары соответствующих точек m_1 и m_2 верно:

$$m_2^T F m_1 = 0.$$

Применение данного вида преобразований для коррекции РЛИ напрямую невозможно из-за несоответствия принципов формирования изображений. Для применения рассмотренных методов геометрической коррекции к РЛИ преобразуем изображение в полярной системе координат азимут–дальность в изображение в местной прямоугольной системе координат ТИ с началом в центре масс ЛА [Кузьмин, 1986].

Величина дискретизации по оси X вычисляется по формуле:

$$\Delta x = \frac{r_{\min} \gamma}{a},$$

где γ — ширина диаграммы направленности антенны (ДНА) БРЛС в горизонтальной плоскости, a — количество элементов дискретизации на элемент разрешения БРЛС.

Величина дискретизации по оси Y вычисляется по формуле:

$$\Delta y = \frac{\Delta r}{a}.$$

Формулы пересчёта из полярной системы координат в местную прямоугольную:

$$\begin{aligned} dx &= \frac{r_{\max} \cos(\Delta\alpha)}{\Delta x}, \\ x &= \frac{[(r_{\min} + r\Delta r)\cos(\alpha\Delta\alpha)]}{\Delta x} + dx, \\ y &= \frac{[(r_{\min} + r\Delta r)\sin(\alpha\Delta\alpha)]}{\Delta y}. \end{aligned}$$

Результаты применения описанного алгоритма представлены на рис. 4 (см. с. 50).

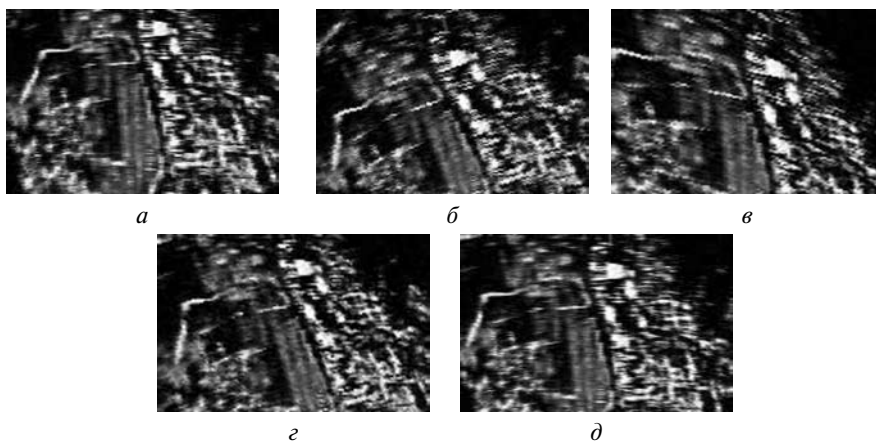


Рис. 4. Результаты экспериментов: *a* — РЛИ при $\theta = -7^\circ$; *б* — РЛИ при $\psi = 15^\circ$; *в* — РЛИ при $\varphi = 10^\circ$; *г* — РЛИ при комплексном отклонении ЛА; *д* — результаты коррекции

Улучшение видения на основе комплексирования ТВ- и ТПВ-каналов

Изображения пилотажной картины от ТВ- или ТПВ-каналов перед выводом на средства отображения проходят предварительную обработку с целью улучшения их восприятия при визуализации. Для этой цели предназначена система улучшения видения (УВ).

Эффективность «интеллектуальной» обработки видеоданных в системе УВ определяется её математическим обеспечением и способом его технической реализации.

Улучшение видения обеспечивается цифровыми преобразованиями исходных ТВ- и ТПВ-изображений, что приводит к расширению возможностей пилота по предельным дальностям видимости, обнаружения и идентификации объектов, особенно при слабой контрастности изображений, низкой заметности целей и объектов в условиях плохой видимости (туман, пыль, дым, дождь, снег или наличие искусственных помех).

Главной целью системы УВ является усиление «разборчивости» визуализируемых изображений, достигаемое алгоритмическим решением задач:

- усиления контрастности при оптимальном яркостном распределении;
- усиления контурного препарата изображения, вносящего наибольший вклад в «разборчивость» элементов сцены;
- устранения и ограничения помех тракта формирования и передачи изображений на устройство отображения.

Для этого в состав математического обеспечения включены следующие операторы:

- адаптивное линейное яркостное преобразование, которое расширяет диапазон изменения входного сигнала, увеличивает глобальный контраст и поддерживает максимально необходимое или реально имеющееся количество информационных уровней в выходном изображении;
- нелинейные яркостные преобразования, основанные на гистограммных преобразованиях (локальных и глобальных) с использованием различных

законов, в том числе по специально разработанным шаблонам, позволяющим эффективно контрастировать изображения и нормализовывать их среднюю яркость;

- адаптивная временная фильтрация (рекурсивная фильтрация, или накопление кадров с попиксельной межкадровой оценкой), обеспечивающая подавление шумов, особенно проявляющихся на уровне порогов чувствительности ТВ- или ТПВ-датчиков;
- пространственная фильтрация (ранговая или линейная), подавляющая структурные и импульсные шумы видеодатчиков;
- градиентная обработка, позволяющая усилить разборчивость изображения.

Адаптивная комбинация различных преобразований, подбор их параметров под текущие условия наблюдения позволяют осуществить в конечном счёте улучшение видения изображений от ТВ- и ТПВ-каналов за порогом чувствительности человеческого глаза. На рис. 5 и 6 показаны примеры преобразований исходных видеоизображений, полученных от ТВ- и ТПВ-каналов. Алгоритмическое обеспечение предусматривает ручной и автоматический способы настройки аппаратных средств задачи УВ для достижения максимально высоких показателей качества выходного изображения. Автоматическая настройка производится по показателям контраста, средней яркости и отношению сигнал / шум.

Другим аспектом повышения информативности изображения для восприятия является комплексирование разноспектральных видеоизображений. Алгоритмическое обеспечение задачи комплексирования включает нормирование изображений



Рис. 5. Улучшение ТВИ: *а* — ТВИ; *б* — улучшенное ТВИ



Рис. 6. Улучшение ТПВИ: *а* — ТПВИ; *б* — улучшенное ТПВИ

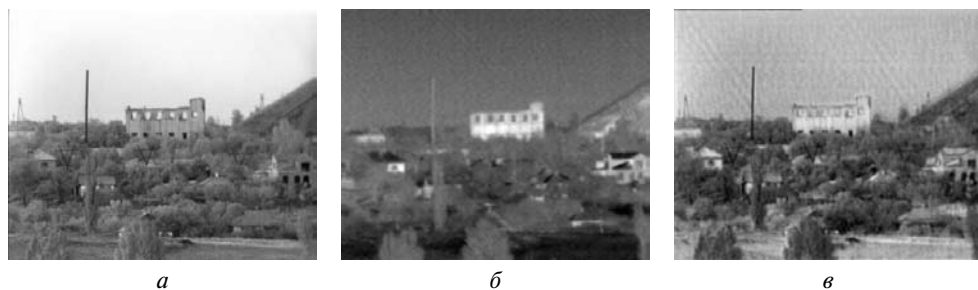


Рис. 7. Комплексирование ТВИ и ТПВИ:
 а — ТВИ; б — ТПВИ; в — результат комплексирования

по динамическому диапазону, пространственному разрешению, геометрическим характеристикам и их объединение (микширование). Выполняется точное совмещение полей зрения и линий визирования, синхронизация и выравнивание видеопотоков ТВ- и ТПВ-каналов, а затем их попиксельное слияние.

Один из вариантов комплексирования основан на приоритете телевизионного изображения и описывается выражением:

$$F_{xy} = I_{xy} + \left| Q_{xy} - \bar{A} \right| - \bar{A}_{\Delta},$$

где F_{xy} — комплексированное изображение; I_{xy} — телевизионное изображение; Q_{xy} — тепловизионное изображение, \bar{A} — среднее значение яркости элементов ТПВИ; \bar{A}_{Δ} — среднее значение абсолютных отклонений яркости элементов ТПВИ.

В результате формируется комплексированное изображение, телевизионное по характеру яркостного распределения (объёмное, с наличием теней и линией горизонта), но дополненное информативными элементами, взятыми из тепловизионного изображения. ТПВ-канал используется как обнаружитель потенциальных целей, препятствий, характерных элементов картины окружающего пространства. Результирующее изображение оказывается более эффективным при восприятии, чем каждое из исходных в отдельности. На рис. 7 показан пример выполнения комплексирования разноспектральных изображений.

Комплексирование ТВИ (ТПВИ) и РЛИ

Изображения земной поверхности, полученные с помощью ТВ, ТПВ и БРЛС, отличаются друг от друга различным спектральным диапазоном и системами координат представления изображений. ТВИ (ТПВИ) и РЛИ могут быть приведены к общей системе координат с помощью геометрических преобразований [Кузьмин, 1967].

Совмещение изображений производится в декартовой системе координат $\{x, y, z\}$. Таким образом, возникают задачи преобразования координат ТИ и получения проекции РЛИ на плоскость, перпендикулярную линии визирования, с координатами $\{x_n, y_n, z_n\}$ (рис. 8, см. с. 53).

Первый этап выполняется за счёт вращения в обратном направлении относительно углов α и β , второй — поворотом системы координат $\{x, y, z\}$ на угол α_n вокруг оси z и последующим вращением на угол β_n вокруг оси y .

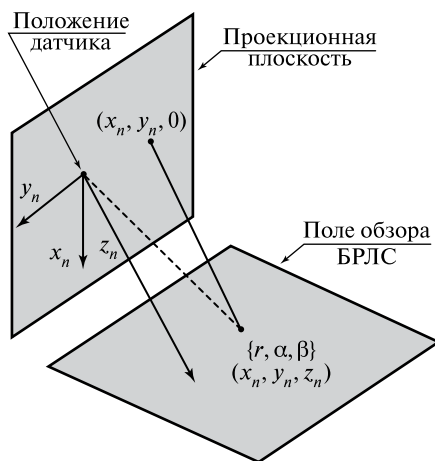


Рис. 8. Преобразование координат РЛИ

С использованием матриц эти операции можно записать в виде:

$$\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = B_n A_n \begin{pmatrix} x \\ y \\ z \end{pmatrix} = B_n A_n A^{-1} B^{-1} \begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix},$$

где A_n и B_n — матрицы, полученные заменой α и β на α_n и β_n в матрицах A и B ; A^{-1} и B^{-1} — матрицы, полученные заменой α и β на $-\alpha$ и $-\beta$ в матрицах A и B .

В результате выполнения данных операций изображение $I(r, \alpha, \beta)$ в полярных координатах будет преобразовано в изображение $I(x_n, y_n, z_n)$ в декартовых координатах, где I — интенсивность отражённого сигнала. В проекционной плоскости РЛИ будет представлено в виде $I(x_n, y_n)$.

Для данных геометрических преобразований необходимо, чтобы РЛИ в моменты времени t_0, t_1, \dots, t_n содержали информацию по всем трем полярным координатам, геометрическим преобразованиям подвергаются трёхмерные РЛИ в виде массивов данных $S_n(r_i, \beta_j, \alpha_k)$, $n = 0, 1, 2, \dots$, представляющих амплитуду отражённого сигнала в каждом i -м, j -м, k -м элементе диаграммы направленности антенны (ДНА). В каждый момент времени РЛИ представляет собой матрицу $S_{\text{РЛИ}}(i, j, k)$ с размерностями $i = 1, N$; $j = 1, M$; $k = 1, K$.

При использовании в качестве датчика РЛИ БРЛС с фазо-кодовой модуляцией (ФКМ), работающей в режиме обзора земной поверхности, результаты измерений интенсивностей отражённого сигнала выдаются в полярных координатах: наклонная дальность r , азимут α , в то время как ТВИ (ТПВИ) содержат отсчёты интенсивности в координатах азимут — угол места β . Таким образом, возникает задача формирования трёхмерного РЛИ. Данную задачу предлагается решать с использованием ЦКМ.



Рис. 9. Привязка РЛИ к ЦКМ: *а* — РЛИ; *б* — ЦКМ; *в* — результат привязки

Для привязки текущего РЛИ к ЦКМ предлагается использовать эталонные РЛИ, подготовленные заранее для заданного района полёта. Эталонные РЛИ предполагается хранить в бортовой базе ЭИ в специальном формате, позволяющем ассоциировать РЛИ с параметрами ориентации ЛА в пространстве и его географическими координатами. ТИ и ЭИ совмещаются с использованием корреляционно-экстремальных алгоритмов. В качестве критериальной функции используется нормированная взаимная корреляционная функция [Баклицкий и др., 1986]:

$$K(m, n) = \frac{\sum_i^M \sum_j^N (S_{\mathcal{E}}(i, j) S_{\mathcal{T}}(i-m, j-n))^2}{\sqrt{\sum_i^M \sum_j^N S_{\mathcal{E}}^2(i, j) \sum_i^M \sum_j^N S_{\mathcal{T}}^2(i-m, j-n)}}$$

где $S_{\mathcal{E}}(i, j)$, $S_{\mathcal{T}}(i, j)$ — яркость эталонного и текущего изображений в точке (i, j) ; M, N — размеры эталонного изображения.

В результате данного этапа производится привязка текущего РЛИ к ЭИ, ассоциированному с определённым участком местности (районом полёта) (рис. 9).

Для получения трёхмерных РЛИ в виде массивов данных $S_n(r_i, \beta_j, \alpha_k)$ производится расчёт координат соответствующих точек рельефа $M'\{x_m, y_m, z_m\}$ в системе координат, связанной с центром масс ЛА по формулам:

$$x_m = x_{\text{ЛА}} + \min R \sin \left(\arccos \left(\frac{H_0}{\min R} \right) \right) + u \, dU;$$

$$y_m = y_{\text{ЛА}} + v \, dV;$$

$$z_m = H_A - H_{\text{РЛИ}}(x_m; y_m),$$

где x_m, y_m, z_m — координаты точки рельефа в локальной системе координат; $x_{\text{ЛА}}, y_{\text{ЛА}}$ — координаты ЛА на ЦКМ; $\min R$ — минимальный отсчёт по дальности для данного РЛИ; H_0 — относительная высота рельефа; H_A — абсолютная высота ЛА; u и v — декартовы координаты элемента РЛИ; dU, dV — дискретность РЛИ по соответствующим координатам.

Отображение точки объекта $M'\{x_m, y_m, z_m\}$ на радиолокационном изображении БРЛС может быть связано с системой координат $\{x_n, y_n, h\}$ оптической системы следующими уравнениями:

$$x_m = f \frac{r_{11}x_n + r_{12}y_n + r_{13}h}{r_{31}x_n + r_{32}y_n + r_{33}h},$$

$$y_m = f \frac{r_{21}x_n + r_{22}y_n + r_{23}h}{r_{31}x_n + r_{32}y_n + r_{33}h},$$

где h — высота полёта носителя БРЛС; f — фокусное расстояние оптической системы тепловизора; $R = \{r_{ij}\}$, $i = 1, 3$, $j = 1, 3$ — матрица вращения.

Для вычисления элементов матрицы R применим следующий порядок разворота системы координат: угол азимута, угол крена и угол тангажа. Тогда, учитывая, что $R = R_x R_y R_z$ [Грузман и др., 2000], получим следующие выражения для r_{ij} :

$$r_{11} = \cos \varphi \cos \psi;$$

$$r_{12} = \sin \theta \cos \psi + \sin \varphi \sin \psi \cos \theta;$$

$$r_{13} = \sin \varphi \cos \psi \cos \theta + \sin \psi \sin \theta;$$

$$r_{21} = \cos \varphi \sin \psi;$$

$$r_{22} = \sin \varphi \sin \psi \sin \theta + \cos \psi \cos \theta;$$

$$r_{23} = \sin \varphi \sin \psi \sin \theta - \cos \psi \cos \theta;$$

$$r_{31} = -\sin \varphi;$$

$$r_{32} = \cos \varphi \sin \theta;$$

$$r_{33} = \cos \varphi \cos \theta.$$

Выводы

Рассмотренные методы получения улучшенного / синтезированного изображения могут применяться как ещё одно средство автономной коррекции счисленных инерциальными навигационными системами текущих координат ЛА в дополнение к спутниковым навигационным системам, а также для создания интуитивно понятного пилоту улучшенного / синтезированного изображения подстилающей поверхности в условиях недостаточной видимости и быстро меняющейся обстановки.

Использование ЦКМ дает более оперативный и дешевый инструмент формирования ЭИ по широким зонам полёта с возможностью моделирования радиолокационного поля и позволяет совмещать РЛИ с ТВИ и ТПВИ, представленными в различных системах координат.

Литература

- Баклицкий В. К., Бочкарев А. М., Мусьяков М. П. (1986) Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. М.: Радио и связь, 1986. 216 с.
- Бэр Р. (2000) Линейная алгебра и проективная геометрия. М.: Физматлит, 2000.
- Грузман И. С., Киричук В. С., Косых В. П. и др. (2000) Цифровая обработка изображений в информационных системах. Новосибирск: Изд-во НГТУ, 2000.

- Костяшкин Л. Н., Логинов А. А., Конкин Ю. В.* (2007) Технология определения координат летательного аппарата на основе совмещения радиолокационной и картографической информации // Мир авионики. 2007. № 6. С. 32–38.
- Кузьмин С. З.* (1967) Цифровая обработка радиолокационной информации. М.: Сов. радио, 1967. 400 с.
- Кузьмин С. З.* (1986) Основы проектирования систем цифровой обработки радиолокационной информации. М.: Радио и связь, 1986. 352 с.

AIRCRAFT CONTROL ENHANCED VISION SYSTEMS TECHNOLOGIES

L. N. Kostyashkin¹, S. I. Babaev², A. A. Loginov¹, O. V. Pavlov¹

¹ Ryazan State Instruments-Making Enterprise (GRPZ),
32 Seminarskaya st., Ryazan, 390000

² Ryazan State Radio Engineering University (RSREU),
390005, Ryazan, 59/1 Gagarin st.

This paper contains overview of aircraft technical vision system information integration methods. Questions geometrical distortions appearing and compensation are discussed. Methods of enchanting and merging TV and FLIR images are given. Methods of RADAR, TV, FLIR images and digital map merging are suggested.

Keywords: technical vision systems, digital map, geometrical distortions, enhanced vision, image combination, extreme correlation navigation systems.

Kostyashkin Leonid Nikolaevich — Director of the Scientific-Design Center of Video Computer Technology at Ryazan State Instrument-making Enterprise, Candidate of Technical Science, Associate Professor, e-mail: hunter-rzn@yandex.ru.

Babaev Sergey Igorevich — Senior Lecturer in the Chair of Computers Ryazan State Radio Engineering University, e-mail: evm@rgta.ryazan.ru.

Loginov Alexander Anatolevich — Chief Designer of Direction of the Scientific-Design Center of Video Computer Technology at Ryazan State Instrument-making Enterprise, Candidate of Technical Science, Associate Professor, e-mail: hunter-rzn@yandex.ru.

Pavlov Oleg Vjacheslavovich — Head of the Laboratory of the Scientific-Design Center of Video Computer Technology at Ryazan State Instrument-making Enterprise, e-mail: hunter-rzn@yandex.ru.

ОБРАБОТКА И АНАЛИЗ ИЗОБРАЖЕНИЙ В БОРТОВЫХ ОПТИКО-ЭЛЕКТРОННЫХ СИСТЕМАХ

Б. А. Алпатов¹, П. В. Бабаян¹, Л. Н. Костяшкин², Ю. Н. Романов²

¹ Государственное образовательное учреждение высшего профессионального образования «Рязанский государственный радиотехнический университет» (ГОУ ВПО «РГРТУ»), 390005, Рязань, Гагарина ул., 59/1

² Федеральное государственное унитарное предприятие «Государственный рязанский приборный завод» (ФГУП «ГРПЗ»), 390000, Рязань, Семинарская ул., 32

Рассмотрены основные задачи обработки и анализа изображений, решаемые в оптико-электронных системах, предназначенных для использования в составе бортового радиоэлектронного оборудования самолётов, вертолётов, судов, наземных транспортных средств. Описаны базовые подходы к решению задач обнаружения и высокоточного сопровождения движущихся и неподвижных объектов, перспективные направления развития информационных технологий в рассматриваемой области. Особое внимание уделено методике принятия решений о выборе алгоритма измерения координат объектов.

Работа проводилась при финансовой поддержке Федерального агентства по науке и инновациям (госконтракт № 02.740.11.0002).

Ключевые слова: изображение, обнаружение, измерение координат, сопровождение.

Разработка и совершенствование бортовых оптико-электронных систем обработки и анализа изображений является важным направлением повышения тактико-технических характеристик современного вооружения и специальной техники. Одними из наиболее актуальных задач, решаемых рассматриваемыми оптико-электронными системами, являются [Алпатов и др., 2008]:

- 1) улучшение видения в сложных условиях наблюдения, в частности, при наличии шумов датчика, вибраций, в условиях слабой освещённости;
- 2) обнаружение и оценка параметров движущихся и неподвижных объектов (самолётов, вертолётов, автомобилей, судов, людей, зданий), наблюдаемых на фонах различных типов: от ясного неба до пересечённой местности;
- 3) высокоточное сопровождение объектов, связанное с организацией непрерывного управления перемещением датчика изображений в целях поддержания постоянного визуального контакта с объектом наблюдения.

При решении этих задач наблюдение может вестись в любое время суток с помощью датчиков изображений видимого или инфракрасного диапазонов, а также посредством мультиспектральных датчиков.

Алпатов Борис Алексеевич — заведующий кафедрой автоматики и информационных технологий в управлении, доктор технических наук, профессор, e-mail: aitu@rgta.ryazan.ru.

Бабаян Павел Варганович — начальник научной лаборатории кафедры автоматики и информационных технологий в управлении, кандидат технических наук, e-mail: aitu@rgta.ryazan.ru.

Костяшкин Леонид Николаевич — директор — главный конструктор Научно-конструкторского центра видеокомпьютерных технологий ФГУП «ГРПЗ», кандидат технических наук, доцент, e-mail: hunter-rzn@yandex.ru.

Романов Юрий Николаевич — заместитель директора-главного конструктора Научно-конструкторского центра видеокомпьютерных технологий ФГУП «ГРПЗ», e-mail: hunter-rzn@yandex.ru.

При разработке методов и алгоритмов обработки и анализа изображений для бортовых оптико-электронных систем следует учитывать следующие основные особенности и ограничения, связанные с характером решаемых задач:

- 1) априорная информация о характеристиках наблюдаемых объектов и фона чаще всего отсутствует или включает в себя лишь приблизительные размеры объектов;
- 2) обработка и анализ изображений должны выполняться в реальном масштабе времени, что обусловлено необходимостью оперативного принятия решений и наличием замкнутого контура управления движением;
- 3) работа бортовой оптико-электронной системы должна обеспечиваться в автономном режиме или при минимальном участии оператора.

Базовые подходы к улучшению видения в сложных условиях наблюдения основаны на применении различных методов линейной и нелинейной пространственно-временной фильтрации, алгоритмов оценивания параметров геометрических преобразований изображений.

Среди зарекомендовавших себя подходов к обнаружению и измерению координат объектов можно выделить четыре основных класса методов.

1. *Методы на основе сопоставления с эталоном* могут использоваться для измерения координат движущихся и неподвижных объектов, наблюдаемых на однородном и неоднородном фоне, при малых отношениях сигнал/шум. Разработаны и используются в различных вариантах аппаратуры эффективные алгоритмы, в основе которых лежит разностное совмещение эталонного и текущего изображений и межкадровая фильтрация эталонного изображения [Алпатов и др., 1985].
2. *Методы статистической сегментации* предназначены для выделения движущихся и неподвижных объектов, наблюдаемых на относительно однородном фоне. Они основаны на использовании априорной информации о различии статистических свойств объекта и фона. Алгоритмы на основе многомерной статистической сегментации использованы в аппаратуре и показали высокую эффективность.
3. *Методы выделения объектов с помощью пространственной фильтрации* основаны на использовании операций линейной и нелинейной пространственной фильтрации изображений. Данный класс методов наиболее эффективен при выделении движущихся и неподвижных объектов на фоне ясного или облачного неба.
4. *Методы на основе выделения динамических изменений* ориентированы на решение задачи выделения движущихся объектов, наблюдаемых на однородном или неоднородном фоне [Алпатов, 1994]. Принцип работы подобных алгоритмов основан на выявлении изменений, происходящих с течением времени в наблюдаемой последовательности изображений.

Теоретическое исследование базовых вопросов, связанных с разработкой пространственно-временных методов обнаружения и сопровождения объектов, проведено, в частности, в работах Б. А. Алпатова и др. [2008, 1994].

Среди перспективных концепций развития информационных технологий обработки и анализа изображений в бортовых оптико-электронных системах можно выделить:

- 1) применение комплексирования информации, полученной различными каналами регистрации изображений: телевизионными, тепловизионными, радиолокационными; это позволяет повысить надёжность обнаружения, выделения и оценки параметров объектов при работе на больших дальностях, в условиях слабой видимости и при использовании различных средств маскировки;
- 2) использование структурных методов обнаружения и оценки параметров объектов, основанных на обнаружении и анализе визуальных примитивов (узловых точек, отрезков, дуг), соответствующих наблюдаемому объекту; данный класс методов наиболее эффективен для обнаружения и измерения координат протяженных объектов искусственного происхождения, наблюдаемых в условиях аффинных и проективных деформаций;
- 3) анализ фоноцелевой обстановки и автоматический выбор алгоритма обнаружения и измерения координат объектов, наиболее эффективного в текущих условиях наблюдения; это дает возможность повысить степень автономности оптико-электронной системы, во многих случаях исключая необходимость вмешательства человека-оператора.

На рис. 1 (см. с. 60) представлена структура программно-алгоритмического обеспечения оптико-электронной системы обнаружения и сопровождения объектов, в которой отражены перечисленные выше концепции.

Существенной особенностью представленной структуры является наличие этапа анализа фоноцелевой обстановки, по результатам работы которого принимается решение об использовании того или иного алгоритма измерения координат объектов для различных фрагментов наблюдаемого изображения. При изменении фоноцелевой обстановки возможен переход от одного метода измерения координат к другому.

Ишем методику принятия решений о работоспособности алгоритмов обнаружения и измерения координат объектов [Алпатов и др., 2009]. Предлагаемая методика основана на применении двух типов количественных метрик, характеризующих работоспособность соответствующих алгоритмов. Первый тип метрик основан на сопоставлении результатов измерения координат с эталонными параметрами наблюдаемого объекта. Эти метрики называются количественными характеристиками работоспособности (КХР). Второй тип метрик основан на анализе пространственно-временных характеристик фона и наблюдаемого объекта. Эти метрики называются признаками работоспособности (ПР). Каждому алгоритму измерения координат соответствует отдельный ПР.

Предварительно, с использованием обширной базы данных видеосюжетов, оценивается характер статистической зависимости КХР от ПР. При наблюдении конкретного видеосюжета вычисляются ПР, и на основе найденной статистической зависимости определяются наиболее вероятные значения КХР для каждого алгоритма измерения координат. Путём сравнения КХР выбирается алгоритм, наиболее эффективный в текущих условиях наблюдения. Результаты исследований показывают, что вероятность принятия ошибочного решения при использовании предлагаемой методики не превышает 6 %. Основные этапы методики отражены на рис. 2 (см. с. 61).

Описанные методы и алгоритмы широко используются в настоящее время при создании программно-аппаратного обеспечения для разрабатываемых и промышленно выпускаемых на ФГУП «Государственный рязанский приборный завод»

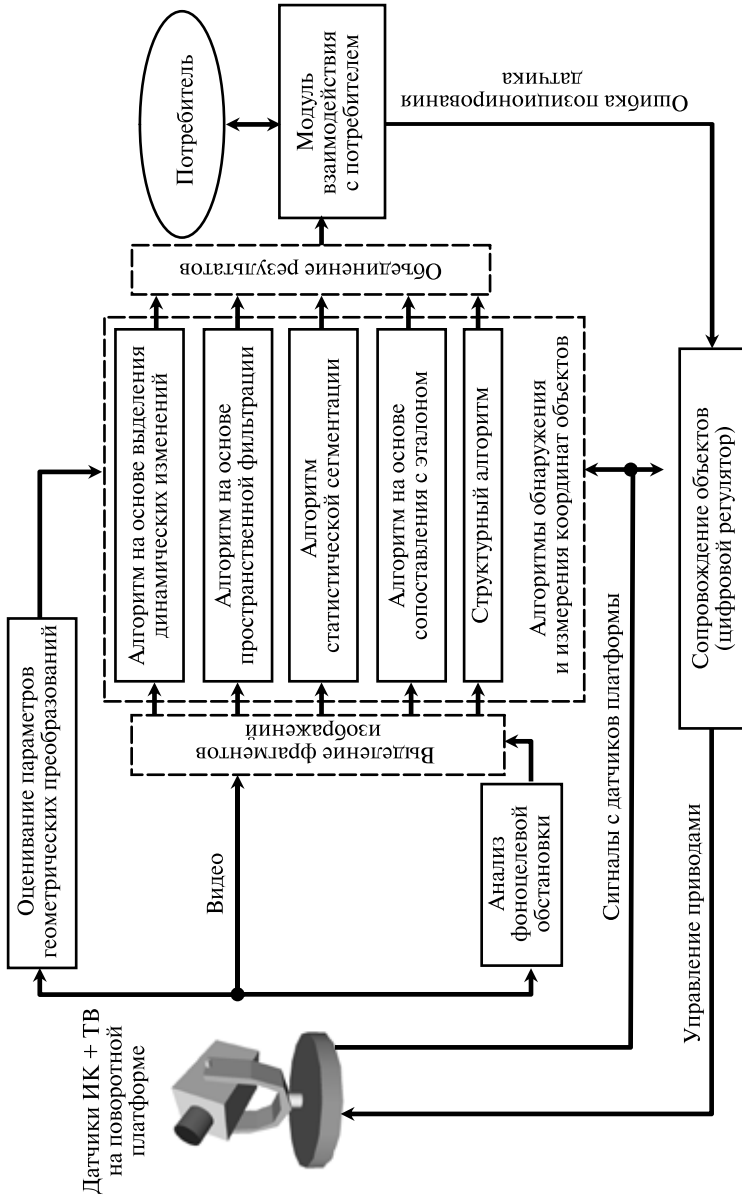


Рис. 1. Структура программно-алгоритмического обеспечения оптико-электронной системы обнаружения и сопровождения объектов

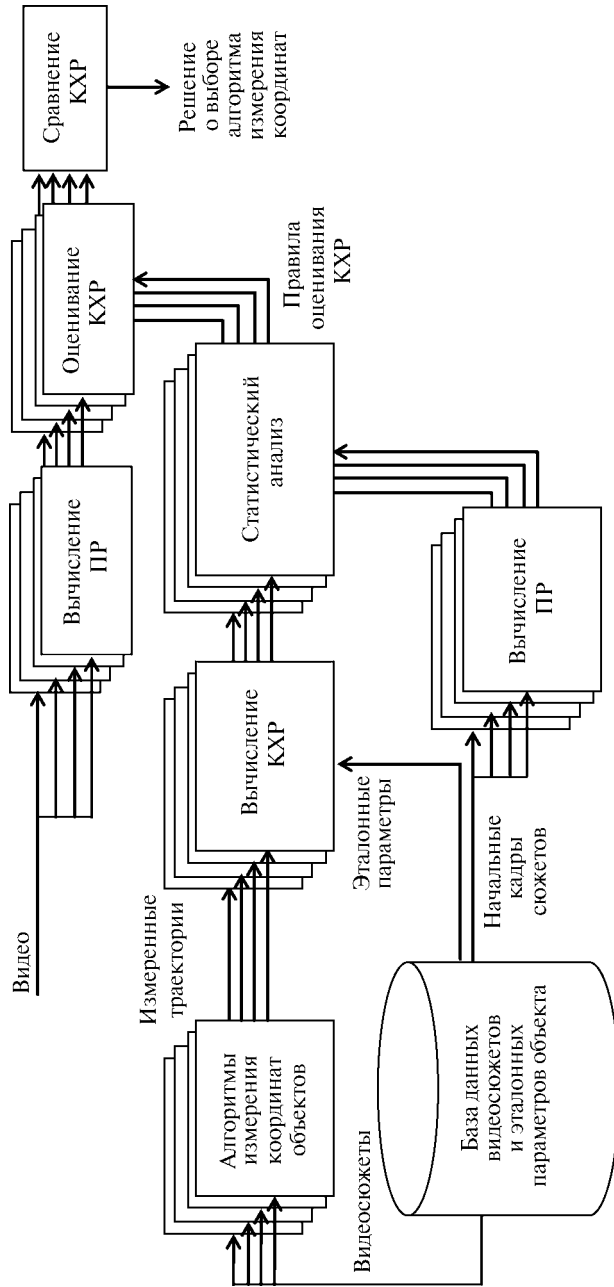


Рис. 2. Методика принятия решений о работоспособности алгоритмов обнаружения и измерения координат объектов

различных модификаций системы обработки и анализа изображений реального времени «Охотник», нашедшей применение в различных образцах специальной техники [Алпатов и др., 2002].

Литература

- Алпатов Б. А. (1994) Оптимальное оценивание параметров движущегося объекта в последовательности изображений // Автометрия. 1994. № 2. С. 32–37.
- Алпатов Б. А., Бабаян П. В., Балашов О. Е., Степашкин А. И. (2008) Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. М.: Радиотехника, 2008. 176 с.
- Алпатов Б. А., Бабаян П. В., Смирнов С. А. (2009) Автоматическое слежение за объектами при отсутствии априорных сведений о фоновой обстановке // Цифровая обработка сигналов. 2009. № 3. С. 52–56.
- Алпатов Б. А., Зеленюк Ю. И., Костяшкин Л. Н., Романов Ю. Н., Семенов В. П. (2002) Видеоинформационные технологии в бортовом радиоэлектронном оборудовании перспективных авиационных систем: концепция, алгоритмы и решения // Радиотехника. 2002. № 8. С. 102–112.
- Алпатов Б. А., Селяев А. А., Степашкин А. И. (1985) Цифровая обработка изображений в задаче отслеживания движущегося объекта // Изв. вузов. Сер. «Приборостроение». 1985. № 2. С. 39–43.

IMAGE PROCESSING AND ANALYSIS IN VEHICLE-BORNE OPTICAL-ELECTRONIC SYSTEMS

B. A. Alpatov¹, P. V. Babayan¹, L. N. Kostyashkin², Yu. N. Romanov²

¹ Ryazan State Radio Engineering University (GRPZ),
390005, Ryazan, 59/1 Gagarin st.

² Ryazan State Instrument-making Enterprise (RSREU),
390000, Ryazan, 32 Seminar'skaya st.

This paper is dedicated to the basic tasks of image processing and analysis, that need to be solved during the development of optical-electronic systems for helicopters, planes, ships and ground vehicles. The most popular methods of object detection, precise tracking, and the conceptions of the perspective approaches are explained. The main attention is given to the task of automatic selection of the algorithm of the object position estimation.

Keywords: image, detection, parameter estimation, tracking.

Alpatov Boris Alekseevich — Head of the Department of Automation and Information Technology in Control at Ryazan State Radio Engineering University, Doctor of Technical Science, Professor, e-mail: aitu@rgta.ryazan.ru.

Babayan Pavel Vartanovich — Head of the Scientific Laboratory of the Department of Automation and Information Technology in Control at Ryazan State Radio Engineering University, Candidate of Technical Science, e-mail: aitu@rgta.ryazan.ru.

Kostyashkin Leonid Nikolaevich — Director of the Scientific-Design Center of Video Computer Technology at Ryazan State Instrument-making Enterprise, Candidate of Technical Science, Senior Lecturer, e-mail: hunter-rzn@yandex.ru.

Romanov Yury Nikolaevich — Deputy Director of the Scientific-Design Center of Video Computer Technology at Ryazan State Instrument-making Enterprise, e-mail: hunter-rzn@yandex.ru.

ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ И ПРИЛОЖЕНИЯ СТЕРЕОСКОПИЧЕСКИХ СИСТЕМ НАВИГАЦИИ, НАВЕДЕНИЯ И ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ МЕСТНОСТИ

И. Н. Белоглазов

*Военный учебный научный центр (ВУНЦ) ВВС «Военно-воздушная академия
им. проф. Н. Е. Жуковского и Ю. А. Гагарина»,
125167, Москва, Планетная, 3*

В статье исследуются стереоскопические навигационные системы (СтНС), предназначенные для решения задач высокоточной навигации, наведения и дистанционного зондирования местности. Приводится краткое изложение теории совместной оптимальной фильтрации, идентификации и проверки гипотез в дискретной динамической системе специального типа. Предложен комбинированный алгоритм СтНС, основанный на этой теории. Произведена линеаризация фотограмметрических уравнений СтНС в режиме стабилизированного горизонтального полёта летательного аппарата. Синтезирован алгоритм обработки изображений, обеспечивающий максимум апостериорной совместной плотности-вероятности вектора состояния динамической системы, вектора неизвестных параметров и выбранной гипотезы. Получены уравнения, характеризующие точность СтНС при полётах над чистыми ландшафтами и урбанизированной местностью и рассмотрена процедура выбора участков коррекции.

Работа выполнена при финансовой поддержке РФФИ (проект № 08-08-00157).

Ключевые слова: стереоскопические системы; навигация; дистанционное зондирование; оптимальная фильтрация; местность.

Введение

Стереоскопические системы коррекции базовой навигационной системы могут применяться для решения задач высокоточной навигации, наведения, дистанционного зондирования местности (при определении координат наблюдаемых объектов).

Достоинствами стереоскопической навигационной системы (СтНС) являются: использование в качестве эталона топографической карты рельефа местности и отсутствие необходимости в оптическом эталоне, высокая точность коррекции базовой навигационной системы.

В области распознавания образов и обработки изображений при решении задачи стереоотожествления, привязки изображений часто используется эвристический подход, связанный с робастной оценкой параметров модели [Чернявский, 2008; Блохинов и др., 2008]. В отличие от такого подхода синтез алгоритмов обработки изображений в СтНС базируется на двух доказанных результатах: прямых и обратных уравнениях фотограмметрии, примененных к двум изображениям стереопары [Лобанов, 1984], а также теории совместной оптимальной нелинейной фильтрации, идентификации и проверки гипотез в дискретной динамической системе специального типа [Белоглазов, 1977, 1983, 2007; Белоглазов, Казарин,

Белоглазов Иннокентий Николаевич — профессор кафедры электронной автоматики (и авиационных тренажёров), доктор технических наук, профессор.

1997; *Beloglazov, Kazarin, 2001*], когда критерием оптимизации является максимум апостериорной совместной плотности-вероятности вектора состояния динамической системы X_N , вектора неизвестных параметров C и выбранной гипотезы D_j $p\left(X_N, C, D_j \mid z^N\right)$.

Первой публикацией, в которой началась разработка теории, стала статья [*Белоглазов, 1977*]. В работе [*Белоглазов, Казарин, 1998*] содержатся все необходимые доказательства. Более подробное изложение и развитие теории можно проследить по публикациям [*Белоглазов, Казарин, 1997; Красовский и др., 1979; Белоглазов, 1983, 2003а, б, 2009, 2010; Белоглазов и др., 1985; Красовский, Белоглазов, 2000; Beloglazov, Kazarin, 1999, 2001; Лобанов, 1984*].

1. КРАТКОЕ ИЗЛОЖЕНИЕ ТЕОРИИ СОВМЕСТНОЙ ОПТИМАЛЬНОЙ ФИЛЬТРАЦИИ, ИДЕНТИФИКАЦИИ И ПРОВЕРКИ ГИПОТЕЗ В ДИСКРЕТНОЙ ДИНАМИЧЕСКОЙ СИСТЕМЕ СПЕЦИАЛЬНОГО ТИПА

1.1. Исходная постановка задачи

Задано:

– $\{D_1, D_2, \dots, D_j, \dots, D_l\}$ — полное множество гипотез и вероятности гипотез — $p_j = P(D_j)$, где $j = \overline{1, l}$, а $\sum_{j=1}^l p_j = 1$; в общем случае гипотеза не является метрической величиной;

– случайный вектор неизвестных параметров C и множество условных плотностей его распределения $\{p(c \mid D_j)\}$, $j = \overline{1, l}$;

– дискретная динамическая система — $X_{k+1} = A_k(C, D)X_k + B_k(C, D)U_k + \xi_k$;

– канал наблюдения — $Z_k = H_k(C, D)X_k + F_k(C, D) + \eta_k$, где $k = 0, 1, \dots, N$.

Прописными буквами будем обозначать случайные величины:

$X_k = (X_{k1}, X_{k2}, \dots, X_{kn})^T$ — случайный вектор состояния динамической системы в дискретный момент времени k ;

$C = (C_1, C_2, \dots, C_r)^T$ — случайный вектор неизвестных параметров;

$Z_k = (Z_{k1}, Z_{k2}, \dots, Z_{km})^T$ — случайный вектор наблюдения, зафиксированный в момент k , а строчными буквами:

$x_k = (x_{k1}, x_{k2}, \dots, x_{kn})^T$;

$c = (c_1, c_2, \dots, c_r)^T$;

$z_k = (z_{k1}, z_{k2}, \dots, z_{km})^T$ — возможные значения этих величин;

$\xi_k = (\xi_{k1}, \xi_{k2}, \dots, \xi_{kn})^T$ — вектор случайных возмущений, действующих на объект в момент времени k ;

$\eta_k = (\eta_{k1}, \dots, \eta_{km})^T$ — вектор шумов наблюдения в момент k ;

$U_k = (U_{k1}, U_{k2}, \dots, U_{ks})^T$ — вектор управления в момент k ;

$A_k(C, D)$ — матрица объекта; $B_k(C, D)$ — матрица объекта по управлению; $H_k(C, D)$ — матрица канала наблюдения; $F_k(C, D) = [F_{k1}(C, D), \dots, F_{km}(C, D)]^T$ — вектор сигнала в момент k ;

ξ_k, η_k — нормальные, центрированные, взаимно не коррелированные, дискретные белые шумы для $\forall k, c, D_j$:

$$M\langle \xi_k | c, D_j \rangle = 0, \quad M\langle \eta_k | c, D_j \rangle = 0, \quad M\langle \xi_k \eta_q^T | c, D_j \rangle = 0,$$

где M — знак условного математического ожидания.

Заданы ковариационные матрицы $Q_k(c, D_j), R_k(c, D_j)$ шумов ξ_k, η_k :

$$M\langle \xi_k \xi_q^T | c, D_j \rangle = Q_k(c, D_j) \delta_{kq}, \quad M\langle \eta_k \eta_q^T | c, D_j \rangle = R_k(c, D_j) \delta_{kq},$$

где δ_{kq} — символ Кронекера, $\delta_{kq} = \begin{cases} 1, & k = q, \\ 0, & k \neq q; \end{cases}$

$A_k(C, D), B_k(C, D), H_k(C, D), F_k(C, D)$, а также характеристики шумов $Q_k(C, D), R_k(C, D)$ зависят от вектора неизвестных параметров C и гипотез D .

Задано множество априорных возможных условных нормальных распределений: $p\{X_0 | C, D\} \in N[m(c, D_j), G(c, D_j)]$, где $m(c, D_j), G(c, D_j)$ — условное математическое ожидание и ковариационная матрица X_0 и $p\{c, D_j\} = p\{c | D_j\} p_j, j = \overline{1, l}$.

Рассматривается вариант, когда для различных гипотез D_j вектор C может принадлежать различным областям Ω_j .

Требуется по всем проведённым наблюдениям $z^N = (z_0^T, \dots, z_k^T, \dots, z_N^T)^T$ оптимальным образом оценить вектор состояния \hat{x}_N , вектор неизвестных параметров \hat{c} и выбрать гипотезу $D = D_v$, имевшую место при проведении наблюдений z^N .

Критерий оптимальности — максимум условной апостериорной плотности-вероятности $p\{x_N, c, D_j | z^N\}$, которая является плотностью вероятности по отношению к случайным векторам X_N, C и вероятностью — по отношению к гипотезам D_j . Величина $p\{x_N, c, D_j | z^N\} dx_N dc$ равна условной вероятности события $x_N < X_N \leq x_N + dx_N, c < C \leq c + dc, D = D_j$ при условии, что случайный вектор всех проведённых наблюдений $Z^N = (Z_0^T, \dots, Z_k^T, \dots, Z_N^T)^T$ принял значение $z^N = (z_0^T, \dots, z_k^T, \dots, z_N^T)^T$.

1.2. Условная апостериорная плотность-вероятность

Формула этой характеристики получена в работе [Белоглазов, Казарин, 1998]. Она имеет вид

$$p\{x_N, c, D_j | z^N\} = \frac{1}{p\{z^N\} \sqrt{(2\pi)^{n+m(N+1)}}} \times \\ \times \exp\left[-\frac{1}{2} \left\| x_N - \hat{x}_{cjN} \right\|_{\Gamma_{cjN}^{-1}}^2 + I_{cjN}\right],$$

где $p\{z^N\}$ — плотность-вероятность вектора всех проведённых наблюдений z^N ;

$\hat{x}_{cjN} = M\langle X_N | C = c, D = D_j, Z^N = z^N \rangle$ — условное математическое ожидание вектора X_N ;

$\Gamma_{cjN} = M \left\langle \left(X_N - \hat{x}_{cjN} \right) \left(X_N - \hat{x}_{cjN} \right)^T \middle| C = c, D = D_j, Z^N = z^N \right\rangle$ — условная ковариационная матрица ошибок оценивания вектора X_N по N наблюдениям;

I_{cjN} — функционал, соответствующий значениям $C = c, D = D_j$ и полученным измерениям z^N .

Обозначим $A_{cjk} = A_k(c, D_j); B_{cjk} = B_k(c, D_j); R_{cjk} = R_k(c, D_j); F_{cjk} = F_k(c, D_j); G_{cj} = G(c, D_j)$.

Вычисление $p \left(x_N, c, D_j \middle| z^N \right)$ связано с решением трёх групп уравнений для $k = 0, 1, \dots, N$.

Уравнения оценок

$$\hat{x}_{cjk} = \hat{x}_{cjk/k-1} + K_{cjk} \Delta_{cjk}, \hat{x}_{cj0/-1} = m_{cj};$$

$$\hat{x}_{cjk+1/k} = A_{cjk} \hat{x}_{cjk} + B_{cjk} U_k;$$

$$K_{cjk} = \Gamma_{cjk} H_{cjk}^T R_{cjk}^{-1};$$

K_{cjk} — коэффициент усиления фильтра Калмана;

$$\Delta_{cjk} = z_k - H_{cjk} \hat{x}_{cjk/k-1} - F_{cjk};$$

$$\hat{x}_{cjk} = M \left\langle X_k \middle| C = c, D = D_j, Z^k = z^k \right\rangle;$$

$$\hat{x}_{cjk/k-1} = M \left\langle X_k \middle| C = c, D = D_j, Z^{k-1} = z^{k-1} \right\rangle;$$

Δ_{cjk} — невязка наблюдений, расхождение между наблюдением z_k и прогнозом $\hat{z}_k = H_{cjk} \hat{x}_{cjk/k-1} + F_{cjk}$.

Уравнения ковариаций

$$\Gamma_{cjk} = M \left\langle \left(X_k - \hat{x}_{cjk} \right) \left(X_k - \hat{x}_{cjk} \right)^T \middle| C = c, D = D_j, Z^k = z^k \right\rangle$$

— условная ковариационная матрица ошибок оценивания вектора X_k по k наблюдениям; условная ковариационная матрица ошибок прогноза —

$$\Gamma_{cjk+1/k} = M \left\langle \left(X_{k+1} - \hat{x}_{cjk+1/k} \right) \left(X_{k+1} - \hat{x}_{cjk+1/k} \right)^T \middle| C = c, D = D_j, Z^k = z^k \right\rangle;$$

$$\left\{ \begin{array}{l} \Gamma_{cjk} = \Gamma_{cjk/k-1} - \Gamma_{cjk/k-1} H_{cjk}^T \left(H_{cjk} \Gamma_{cjk/k-1} H_{cjk}^T + R \right)^{-1} H_{cjk} \Gamma_{cjk/k-1}, \\ \Gamma_{cj0/-1} = G_{cj}; \end{array} \right.$$

или эквивалентная формула

$$\Gamma_{cjk}^{-1} = \Gamma_{cjk/k-1}^{-1} + H_{cjk}^T R_{cjk}^{-1} H_{cjk};$$

$$\Gamma_{cjk+1/k} = A_{cjk} \Gamma_{cjk} A_{cjk}^T + Q_{cjk}.$$

Уравнения ковариации не зависят от измерений, их можно рассчитать заранее по априорным данным.

Уравнения функционалов

$$I_{cjk} = I_{cjk-1} + \left\| \Delta_{cjk} \right\|_{\Sigma_{cjk}^{-1}}^2 + \varepsilon_{cjk}, \quad k = 0, 1, \dots;$$

$$\Sigma_{cjk}^{-1} = R_{cjk}^{-1} - R_{cjk}^{-1} H_{cjk} \Gamma_{cjk} H_{cjk}^T R_{cjk}^{-1};$$

$\varepsilon_{cjk} = \ln \frac{|\Gamma_{cjk/k-1}| |R_{cjk}|}{|\Gamma_{cjk-1}|}$, $|\Gamma_{cjk/k-1}|$, $|R_{cjk}|$, $|\Gamma_{cjk-1}|$ — определители соответствующих матриц;

$$I_{c_{j,-1}} = -2 \ln p(c, D_j), \quad |\tilde{A}_{c_{j,-1}}| = 1.$$

Решением этой системы является:

$$I_{cjk} = -2 \ln p(c, D_j) + \sum_{k=0}^N \left\| \Delta_{cjk} \right\|_{\Sigma_{cjk}^{-1}}^2 + \sum_{k=0}^N \varepsilon_{cjk}.$$

Функционал должен быть наименьшим, чтобы достигался максимум апостериорной плотности-вероятности.

1.3. Алгоритм, оптимальный по критерию максимума апостериорной плотности-вероятности

Надо найти такие оценки \hat{x}_N , \hat{c} , $\hat{D} = D_v$, которые обращают в максимум плотность-вероятность $p(x_N, c, D_j | z^N)$.

Эти оптимальные оценки обращают в минимум функционал

$$J_{c_{jN}} = \left\| x_N - \hat{x}_{c_{jN}} \right\|_{\Gamma_{c_{jN}}^{-1}}^2 + I_{c_{jN}},$$

т. е. надо найти минимум $\left[\left\| x_N - \hat{x}_{c_{jN}} \right\|_{\Gamma_{c_{jN}}^{-1}}^2 + I_{c_{jN}} \right]$ по всевозможным x_N , c , D_j ; при этом предполагается, что $\Gamma_{c_{jN}}^{-1}$ — положительно определённая матрица и квадратичная форма $\left\| x_N - \hat{x}_{c_{jN}} \right\|_{\Gamma_{c_{jN}}^{-1}}^2 \geq 0$.

Рассмотрим оптимальный алгоритм последовательности операций.

1. Для любой D_j находятся значения вектора неизвестных параметров c_j , которые минимизируют $I_{c_{jN}}$,

$$c_j = \arg \inf_{c \in \Omega_j} I_{c_{jN}}, \quad j = \overline{1, l}, \Leftrightarrow I_{c_{jN}} = \inf_{c \in \Omega_j} I_{c_{jN}}, \quad \text{где } \inf \text{ — нижняя грань.}$$

Возможны два варианта:

$$\arg \inf_{c \in \Omega_j} I_{c_{jN}} \text{ — либо точка минимума, либо точка на границе области } \Omega_j.$$

В случае минимума c_j находятся так:

$$\frac{\partial I_{cjN}}{\partial c_1} = 0, \frac{\partial I_{cjN}}{\partial c_2} = 0, \dots, \frac{\partial I_{cjN}}{\partial c_r} = 0$$

и матрица вторых производных

$$\left\| \frac{\partial^2 I_{cjN}}{\partial c_p \partial c_q} \right\|_{c=c_j} \text{ — должна быть положительно определённой.}$$

2. Определяется гипотеза

$$\hat{D} = D_v(z^N) = \operatorname{argmin}_{D_j} I_{c_j N}.$$

Общей оптимальной оценкой \hat{c} является $c_v(z^N)$.

3. Оптимальной оценкой вектора состояния будет

$$\hat{x}_N = \hat{x}_{c_v, vN}(z^N).$$

Тройка $D_v(z^N)$, $\hat{c}_v(z^N)$, $\hat{x}_N(z^N)$ является решением задачи оптимизации.

Структурная схема оптимального алгоритма приведена на рис. 1 (см. с. 69), где τ обозначает элемент единичного запаздывания.

1.4. Частные случаи

В исходной постановке рассмотренной задачи в критерии оптимизации присутствовали: вектор состояния динамической системы X_k , вектор неизвестных параметров C и проверяемые гипотезы D_j . Если в критерии оптимизации или в исходной постановке отсутствуют один или два из этих объектов, то возникает целое семейство важных частных случаев, имеющих самостоятельное значение. Будем обозначать рассматриваемые варианты дробной записью: в числителе перечислены объекты, участвующие в постановке задачи, а в знаменателе — объекты, входящие в критерии оптимизации. Все возможные варианты иллюстрирует таблица.

Таблица частных случаев

Объекты, входящие в критерий оптимизации	Объекты, входящие в постановку задачи						
	X, C, D	X, C	X, D	C, D	X	C	D
X, C, D	$X, C, D/X, C, D$						
X, C	$X, C, D/X, C$	$X, C/X, C$					
X, D	$X, C, D/X, D$		$X, D/X, D$				
C, D	$X, C, D/C, D$			$C, D/C, D$			
X	$X, C, D/X$	$X, C/X$	$X, D/X$		X/X		
C	$X, C, D/C$	$X, C/C$		$C, D/C$		C/C	
D	$X, C, D/D$		$X, D/D$	$C, D/D$			D/D

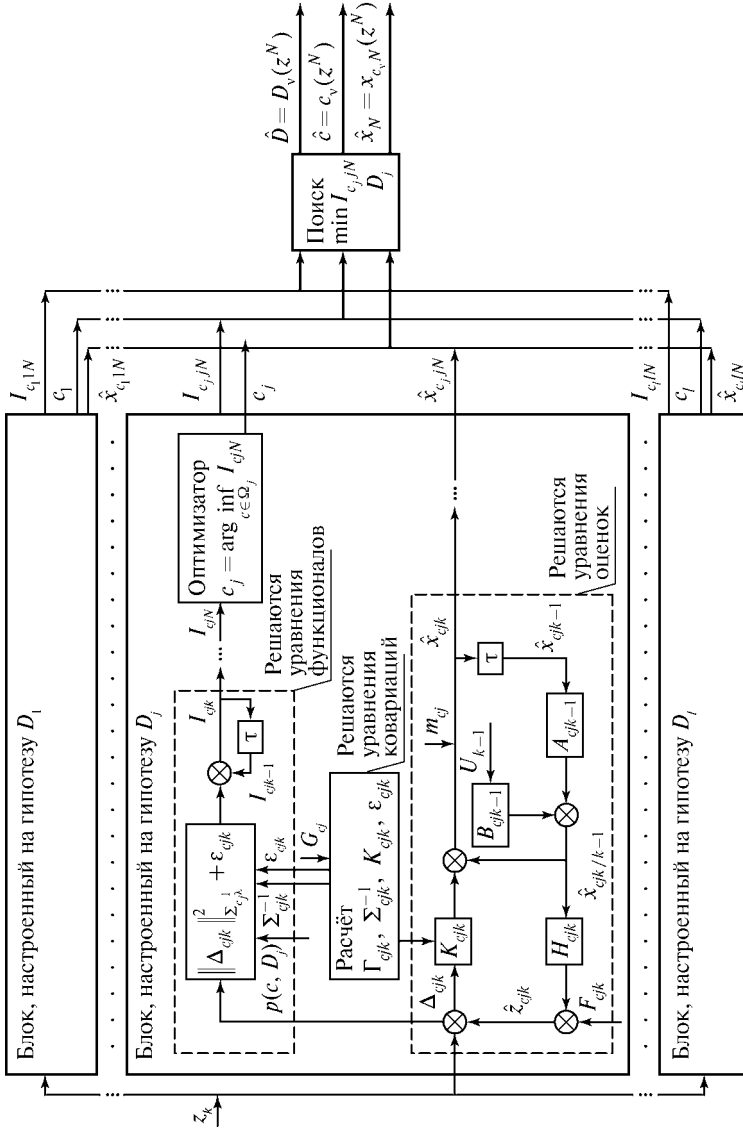


Рис. 1. Структурная схема оптимального алгоритма

2. УРАВНЕНИЯ СТЕРЕОНАБЛЮДЕНИЯ

2.1. Уравнения наблюдения для стереопары

Рассмотрим стереопару (рис. 2), т. е. два изображения, имеющие общую область наблюдения.

В плоскости первого изображения вводится равномерная сетка с шагом δ , на пересечении i -го столбца и j -й строки которой находится точка с координатами $x_i^I = i\delta$, $y_j^I = j\delta$, δ — линейный размер элементарного приёмника ПЗС-матрицы.

Уравнения фотограмметрии [Лобанов, 1984] позволяют найти точку второго изображения x_{ij}^{II} , y_{ij}^{II} , соответствующую точке x_i^I , y_j^I .

Рассмотрим стабилизированный горизонтальный полёт [Белоглазов, Казарин, 1997] на высоте $H^{*SI} = H^{*SII} = H^*$ с постоянной скоростью V_x^* , с нулевым курсом $\psi^{*I} = \psi^{*II} = 0$ и с вертикальным расположением оптической оси датчика изображений (рис. 3, см. с. 71)

$$\theta^{*I} = \theta^{*II} = 0, \quad \gamma^{*I} = \gamma^{*II} = 0.$$

В этом случае

$$Y^{*SI} = Y^{*SII} = 0, \quad X^{*SI} = X^{*SII} - B,$$

$B = V_x \tau$ — стереобаза наблюдения, τ — время перемещения между точками получения первого и второго изображений стереопары.

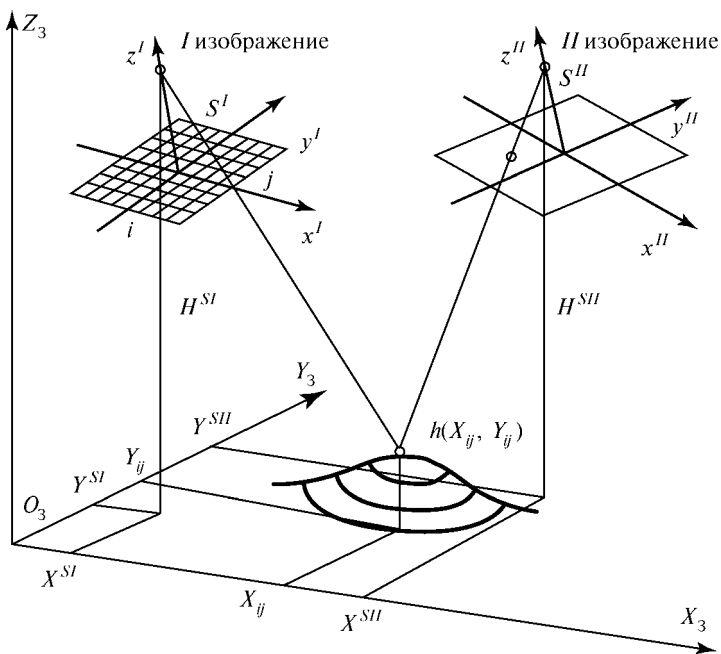


Рис. 2. Определение соответственных точек на I и II изображениях стереопары

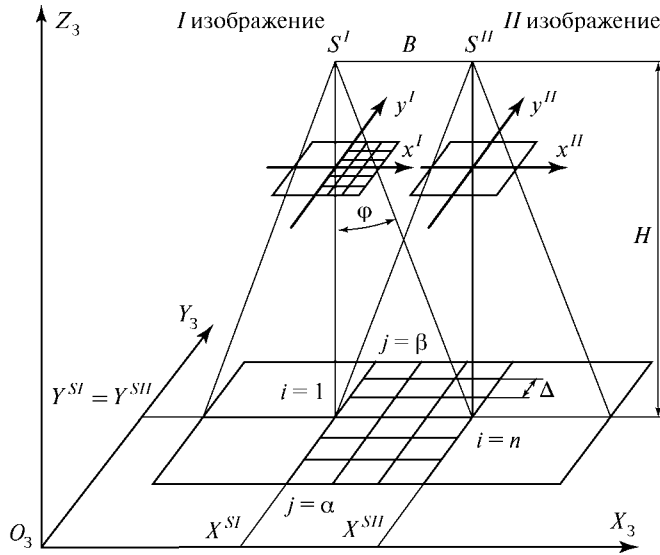


Рис. 3. Определение соответственных точек на I и II изображениях стереопары при стабилизированном горизонтальном полёте

Навигационной системой на борту летательного аппарата (ЛА) измеряются значения координат, скорости и угловой ориентации видеокамеры:

$$\left. \begin{aligned} X^{*SII} &= X^{SII} + \Delta X^{II}, & X^{*SI} &= X^{SI} + \Delta X^I, & V_X^* &= V_X + \Delta V_X, \\ Y^{*SII} &= Y^{SII} + \Delta Y^{II}, & Y^{*SI} &= Y^{SI} + \Delta Y^I, & V_Y^* &= V_Y + \Delta V_Y, \\ H^{*SII} &= H^{SII} + \Delta H, & H^{*SI} &= H^{SI} + \Delta H, & \gamma^{*II} &= \gamma^{II} + \Delta \gamma, \\ \psi^{*II} &= \psi^{II} + \Delta \psi, & \theta^{*II} &= \theta^{II} + \Delta \theta, & \gamma^{*I} &= \gamma^I + \Delta \gamma, \\ \psi^{*I} &= \psi^I + \Delta \psi, & \theta^{*I} &= \theta^I + \Delta \theta, & & \end{aligned} \right\} \quad (2.1)$$

где

$X^{*SII}, Y^{*SII}, H^{*SII}, \psi^{*II}, \theta^{*II}, \gamma^{*II}, X^{*SI}, Y^{*SI}, H^{*SI}, \psi^{*I}, \theta^{*I}, \gamma^{*I}$ — измеренные навигационной системой значения элементов внешнего ориентирования стереопары;

$X^{SI}, Y^{SI}, H^{SI}, \psi^I, \theta^I, \gamma^I; X^{SII}, Y^{SII}, H^{SII}, \psi^{II}, \theta^{II}, \gamma^{II}$ — действительные значения элементов внешнего ориентирования I и II изображений стереопары;

$\Delta X^{II} = \Delta X^I + \Delta V_X \tau, \Delta Y^{II} = \Delta Y^I + \Delta V_Y \tau$; в этих записях ошибки измерения высоты, скорости и углов курса ψ , тангажа θ и крена γ считаются постоянными в предположении их несущественного изменения за время τ .

Если яркость наблюдаемой точки на местности $X_{ij}, Y_{ij}, Z_{ij} = h(X_{ij}, Y_{ij})$ обозначить f_{ij} , то сигналы приёмников в соответственных точках x_i^I, y_j^I и x_{ij}^{II}, y_{ij}^{II} первого и второго изображений стереопары будут равны:

$$u_{ij}^I = k_{ij}^I f_{ij}^I + \delta f_{ij}^I, u_{ij}^{II} = k_{ij}^{II} f_{ij}^{II} + \delta f_{ij}^{II}, \quad (2.2)$$

где u_{ij}^I — сигнал элементарного приёмника излучения I изображения, расположенного в точке $x_i^I = i\delta, y_j^I = j\delta$;

u_{ij}^{II} — сигнал элементарного приёмника излучения II изображения, расположенного в точке x_{ij}^{II} , y_{ij}^{II} , сопряженной с точкой x_{ij}^I , y_{ij}^I ;

f_{ij}^I , f_{ij}^{II} — определяемые яркостью f_{ij} облучённости элементарных приёмников, расположенных в соответственных точках x_{ij}^{II} , y_{ij}^{II} и x_{ij}^I , y_{ij}^I I и II изображений стереопары;

k_{ij}^I , k_{ij}^{II} — коэффициенты усиления элементарных приёмников I и II изображений, δf_{ij}^I , δf_{ij}^{II} — инструментальные погрешности (шумы) этих приёмников.

Облучённости f_{ij}^I , f_{ij}^{II} , а также коэффициенты k_{ij}^I , k_{ij}^{II} в общем случае могут отличаться друг от друга из-за неламбертовости отражения (излучения) поверхности, влияния атмосферы и других причин, но эти зависимости известны с достаточной точностью. В первом приближении можно считать $f_{ij}^I = f_{ij}^{II}$, отнеся допускаемую при этом неточность к ошибкам δf_{ij}^I , δf_{ij}^{II} . В дальнейшем для сокращения записей принимаем $k_{ij}^I = k_{ij}^{II} = 1$. Чтобы в дальнейшем можно было воспользоваться теорией фильтрации, вектор используемых сигналов элементарных приёмников I изображения стереопары представим в виде вектора-столбца

$$u^I = \left(f_{1\alpha}^I, f_{1\alpha+1}^I, \dots, f_{1\beta}^I; f_{2\alpha}^I, \dots, f_{2\beta}^I; \dots; f_{i\alpha}^I, \dots, f_{ij}^I, \dots, f_{i\beta}^I; \dots; f_{n\alpha}^I, \dots, f_{n\beta}^I \right)^T + \left(\delta f_{1\alpha}^I, \dots, \delta f_{ij}^I, \dots, \delta f_{n\beta}^I \right)^T;$$

$i = 1, \dots, n; j = \alpha, \dots, \beta; n, \alpha, \beta$ определяются размерами общей области наблюдения стереопары.

При сокращенных обозначениях

$$f^I = \left(f_{1\alpha}^I, \dots, f_{ij}^I, \dots, f_{n\beta}^I \right)^T; \quad \delta f^I = \left(\delta f_{1\alpha}^I, \dots, \delta f_{ij}^I, \dots, \delta f_{n\beta}^I \right)^T;$$

$$u^I = f^I + \delta f^I. \quad (2.3)$$

Далее обозначим:

— вектор ошибок бортовых измерений в момент записи двух изображений стереопары

$$X = \left(\Delta X^{II}, \Delta V_X, \Delta H, \Delta \theta, \Delta Y^{II}, \Delta V_Y, \Delta \psi, \Delta \gamma \right)^T; \quad (2.4)$$

— вектор измеренных навигационной системой значений элементов внешнего ориентирования изображений

$$X = \left(X^{*SII}, Y^{*SII}, H^{*SII}, \psi^{*II}, \theta^{*II}, \gamma^{*II}, X^{*SI}, Y^{*SI}, H^{*SI}, \psi^{*I}, \theta^{*I}, \gamma^{*I} \right)^T; \quad (2.5)$$

— вектор значений высот рельефа в визируемых точках стереопары

$$h = \left(h_{1\alpha}, h_{1\alpha+1}, \dots, h_{1\beta}; h_{2\alpha}, \dots, h_{2\beta}; \dots; h_{i\alpha}, \dots, h_{ij}, \dots, h_{i\beta}; \dots; h_{n\alpha}, \dots, h_{n\beta} \right)^T;$$

— вектор значений высот, извлекаемых из имеющейся цифровой карты рельефа (либо топографической карты, либо карты 3D-поверхности, наблюдаемой оптико-электронной системой)

$$h^k = \left(h_{1\alpha}^k, h_{1\alpha+1}^k, \dots, h_{1\beta}^k, h_{2\alpha}^k, \dots, h_{2\beta}^k, \dots; h_{i\alpha}^k, \dots, h_{ij}^k, \dots, h_{i\beta}^k, \dots; h_{n\alpha}^k, \dots, h_{n\beta}^k \right)^T;$$

— вектор ошибок картографирования рельефа

$$\delta h = \left(\delta h_{1\alpha}, \dots, \delta h_{ij}, \dots, \delta h_{n\beta} \right)^T.$$

Вектор используемых сигналов элементарных приёмников второго изображения стереопары также представим в виде вектора-столбца:

$$u^{II} = \left(f_{1\alpha}^{II}, \dots, f_{ij}^{II}, \dots, f_{n\beta}^{II} \right)^T + \left(\delta f_{1\alpha}^{II}, \dots, \delta f_{ij}^{II}, \dots, \delta f_{n\beta}^{II} \right)^T, \text{ при сокращённом обозначении}$$

$$u^{II} = f^{II} + \delta f^{II}, \quad (2.6)$$

$$\text{вектор шумов — } \delta f^{II} = \left(\delta f_{1\alpha}^{II}, \dots, \delta f_{ij}^{II}, \dots, \delta f_{n\beta}^{II} \right)^T.$$

При допущении $f^I = f^{II}$ с учётом равенств (2.3), (2.6) получим

$$u^I = f^{II} + \delta f^I. \quad (2.7)$$

Первое слагаемое правой части уравнения (2.6) определяет вектор облучённостей f^{II} , соответственных точек второго изображения стереопары, которые зависят от X^* , X , h . Для дальнейшего рассмотрения его удобно представить в виде

$$f^{II} = F(X^*, X, h), \quad (2.8)$$

$F(X^*, X, h)$ — вектор-столбец, составленный из столбцов матрицы $[F_{ij}(X^*, X, h)]$;

$$\begin{aligned} F_{ij}(X^*, X, h) = f^{II}(x_{ij}^{II}, y_{ij}^{II}) = F_{ij}(X^{*SII} - \Delta X^{II}, Y^{*SII} - \Delta Y^{II}, \\ H^{*SII} - \Delta H, \psi^{*II} - \Delta \psi, \theta^{*II} - \Delta \theta, \gamma^{*II} - \Delta \gamma, \\ X^{*SI} - \Delta X^{II} - \Delta V_X \tau, Y^{*SI} - \Delta Y^{II} - \Delta V_Y \tau, \\ H^{*SI} - \Delta H, \psi^{*I} - \Delta \psi, \theta^{*I} - \Delta \theta, \gamma^{*I} - \Delta \gamma, i\delta, h(X_{ij}, Y_{ij})), \end{aligned}$$

$$i = 1, \dots, n; j = \alpha, \dots, \beta,$$

где $X_{ij} = X_{ij}(X^*, X)$, ..., $Y_{ij} = Y_{ij}(X^*, X)$ — неявно заданные зависимости, определяемые фотограмметрическими уравнениями, в которых аргументы Y^{*SII} , H^{*SII} , ψ^{*II} , θ^{*II} , γ^{*II} ; X^{*SI} , Y^{*SI} , H^{*SI} , ψ^I , θ^I , γ^I выражены через X^* и X .

Функция $(F_{ij}(X^*, X, h) = F_{ij}(X^{*SII} - \Delta X^{II}, Y^{*SII} - \Delta Y^{II}, \dots, h(X_{ij}, Y_{ij})) = f^{II}(x_{ij}^{II}, y_{ij}^{II})$ задаёт связь координат $x_i^I = i\delta$, $y_j^I = j\delta$ точки I изображения стереопары с координатами x_{ij}^{II} , y_{ij}^{II} соответственной точки II изображения. Сначала

с использованием прямых фотограмметрических уравнений и измеренных корректируемой навигационной системой элементов внешнего ориентирования X^{*SI} , Y^{*SI} , H^{*SI} , ψ^{*I} , θ^{*I} , γ^{*I} изображения стереопары с использованием карты рельефа местности определяются координаты X_{ij} , Y_{ij} , $h(X_{ij}, Y_{ij})$ наблюдаемой точки местности (см. рис. 2). Затем по этим координатам с использованием обратных фотограмметрических уравнений, измеренных элементов внешнего ориентирования

II изображения $X^{*SII}, Y^{*SII}, H^{*SII}, \psi^{*II}, \theta^{*II}, \gamma^{*II}$ находятся координаты x_{ij}^{II}, y_{ij}^{II} «псевдосоответственной» i, j -й точки на II изображении.

После определения всех «псевдосоответственных» точек II изображения на основе теории совместной оптимальной фильтрации, идентификации и проверки гипотез в дискретной динамической системе [Белоглазов, Казарин, 1998; Beloglazov, Kazarin, 2001; Белоглазов, 2007] и закономерностей авиационной светотехники [Савельев, 1967; Белоглазов и др., 1987] по критерию максимума совместной апостериорной плотности-вероятности $p(X, C, D_v | z^N)$ с помощью комбинированного алгоритма [Белоглазов, Казарин, 1997, 1998; Beloglazov, Kazarin, 1999], включающего процедуры фильтрации, идентификации и проверки гипотез, связанного с минимизацией функционалов I_{cvN} , содержащих суммы квадратов разностей $\left\| u^I(x_i^I, y_j^I) - f^{II}(x_{ij}^{II}, y_{ij}^{II}) \right\|^2$ всех обрабатываемых точек I и II изображений, находится вектор X ошибок корректируемой навигационной системы.

Таким оптимальным образом уточняются все элементы внешнего ориентирования I и II изображений и находятся скорректированные значения позиционных и угловых координат ЛА на момент получения II изображения стереопары. В процессе выполнения этой процедуры определяются соответственные точки на II изображении стереопары для всех анализируемых точек I изображения, хотя это не является целью используемого алгоритма, а лишь средством обеспечения оптимальности найденных ошибок позиционных и угловых координат ЛА.

При рассмотрении вопросов точности стереоскопической навигационной системы (СтНС), поскольку предполагается малость оцениваемых ошибок, вместо сложного рабочего комбинированного алгоритма можно ограничиться исследованием калмановского фильтра, являющегося частью комбинированного алгоритма.

Векторы $(f_{1\alpha}^I, \dots, f_{ij}^I, \dots, f_{n\beta}^I)^T, (f_{1\alpha}^{II}, \dots, f_{ij}^{II}, \dots, f_{n\beta}^{II})^T$ облучённостей соответственных точек первого и второго изображения стереопары совпадут, только если элементы внешнего ориентирования I и II изображений $X^{SI}, \dots, \gamma^I; X^{SII}, \dots, \gamma^{II}$ точно известны и отсутствуют шумы приёмников. Иначе «соответственные» точки x_{ij}^{II}, y_{ij}^{II} второго изображения будут определяться с ошибками, и совпадения облучённостей f_{ij}^I, f_{ij}^{II} не произойдет.

В работах [Белоглазов, Казарин, 1997; Белоглазов, 2009] проведена линеаризация функции $F(\dots)$ в окрестности точки $X = 0, h = h^k$, с учётом того, что $h = h^k - \delta h$, получена запись уравнения (2.6) в виде:

$$u^{II} = F(X^*, 0, h^k) + H(X^*, h^k)X + \delta f^{II} - \delta Z^{II}, \quad (2.9)$$

где

$$H(X^*, h^k) = \left. \frac{\partial F(X^*, X, h)}{\partial X} \right|_{\substack{h=h^k \\ X=0}},$$

$$\delta Z^{II} = \left. \frac{\partial F(X^*, X, h)}{\partial h} \right|_{\substack{h=h^k \\ X=0}} \delta n,$$

$$F(X^*, 0, h^k) = \left(F_{1\alpha}(X^*, 0, h^k), \dots, F_{ij}(X^*, 0, h^k), \dots, F_{n\beta}(X^*, 0, h^k) \right)^T$$

— вектор облучённостей, имеющий недостаточную точность соответственных точек второго изображения, рассчитанный на основе измеренных навигационной системой значений элементов внешнего ориентирования изображений и карты рельефа местности; назовем его условно вектором облучённостей «псевдосоответственных» точек второго изображения стереопары;

$H^T(X^*, h^k)X$ — слагаемое для корректировки «псевдосоответственных» точек и поиска оптимально возможного, наиболее точного положения соответственных точек второго изображения стереопары;

$$H^T(X^*, h^k) = \begin{bmatrix} H_{1\alpha}^1 & \dots & H_{ij}^1 & \dots & H_{n\beta}^1 \\ H_{1\alpha}^2 & \dots & H_{ij}^2 & \dots & H_{n\beta}^2 \\ \dots & \dots & \dots & \dots & \dots \\ H_{1\alpha}^8 & \dots & H_{ij}^8 & \dots & H_{n\beta}^8 \end{bmatrix}. \quad (2.10)$$

В работах [Белоглазов, Казарин, 1997, 1998] для случая стабилизированного горизонтального полёта получены выражения

$$\left. \begin{aligned} H_{ij}^1 &= -f'_{X_{ij}} \frac{B}{H} h'_{X_{ij}}, \\ H_{ij}^2 &= -f'_{X_{ij}} \left[1 - h'_{X_{ij}} \frac{B}{H} \right], \\ H_{ij}^3 &= -f'_{X_{ij}} \frac{B}{H} \left[1 - h'_{X_{ij}} \frac{\Delta X_{ij} + B}{H} - h'_{Y_{ij}} \frac{\Delta Y_{ij}}{H} \right], \\ H_{ij}^4 &= f'_{X_{ij}} \frac{B}{H} \left[2\Delta X_{ij} + B - h'_{X_{ij}} H \right] + f'_{Y_{ij}} \frac{B}{H} \Delta Y_{ij}, \\ H_{ij}^5 &= -f'_{X_{ij}} \frac{B}{H} h'_{Y_{ij}}, \\ H_{ij}^6 &= -f'_{X_{ij}} \frac{B}{H} h'_{X_{ij}} - f'_{Y_{ij}}, \\ H_{ij}^7 &= f'_{X_{ij}} \frac{B}{H} \left[h'_{Y_{ij}} (\Delta X_{ij} + B) - h'_{X_{ij}} \Delta Y_{ij} \right] - f'_{Y_{ij}} B, \\ H_{ij}^8 &= f'_{X_{ij}} \frac{B}{H} \left[\Delta Y_{ij} - h'_{Y_{ij}} H \right]; \end{aligned} \right\} \quad (2.11)$$

$\Delta X_j = X_{ij} - X^{SH}$, $\Delta Y_j = Y_{ij} - Y^{SH}$, а соответствующие ошибки модели стереоскопического наблюдения равны

$$\delta Z^H = \left(-H_{1\alpha}^3 \delta h_{1\alpha}, \dots, -H_{ij}^3 \delta h_{ij}, \dots, -H_{n\beta}^3 \delta h_{n\beta} \right)^T,$$

где $f'_{X_{ij}}$, $f'_{Y_{ij}}$ — градиенты яркости оптического поля в точке X_{ij} , Y_{ij} ; $h'_{X_{ij}}$, $h'_{Y_{ij}}$ — градиенты рельефа по земным осям X_3 , Y_3 .

Ковариационная матрица R ошибки $\delta Z = \delta f^I - \delta f^{II} - \delta Z^{II}$ равна, согласно (2.11) (подробное доказательство приведено в работе [Белоглазов, Казарин, 1997]),

$$R = \sigma_{\sim}^2 E, \quad \sigma_{\sim}^2 = \frac{B^2}{H^2} \sigma_{f'}^2 \sigma_k^2 + 2\sigma_{II}^2, \quad (2.12)$$

где E — единичная матрица; σ_{\sim}^2 , $\sigma_{f'}^2$, σ_k^2 — дисперсии суммарной случайной помехи, поля оптического градиента и ошибки карты рельефа; σ_{II}^2 — дисперсия ошибки элементарных приёмников ПЗС-матрицы; шумы используемых приёмников предполагаются независимыми.

2.2. Уравнения наблюдения для стереомножества изображений

Использование стереомножества изображений предполагает последовательную обработку ряда стереопар. Удачное использование стереопар получается при $B = H \operatorname{tg} \varphi$, φ — половина угла зрения иконического средства. В этом случае общая область занимает половину изображения, в формулах (2.3), (2.6), $\alpha = -n$, $\beta = n$.

Обозначим через k дискретный момент времени получения k -го изображения (и, соответственно, k -й стереопары). В этот момент знание об элементах ориентирования I и II изображений k -й стереопары определяется некорректированными показаниями навигационной системы X_k^* и прогнозируемым значением ошибки навигационной системы $\hat{X}_{k|k-1}$, рассчитанным по всем наблюдениям, проведённым до момента k .

Поэтому линеаризация уравнения (2.8) должна производиться относительно точки $X = \hat{X}_{k|k-1}$, и соотношение (2.7) для момента времени k примет вид

$$\begin{aligned} u_k^I &= F\left(X_k^*, \hat{X}_{k|k-1}, h^k\right) + \left. \frac{\partial F\left(X_k^*, X, h^k\right)}{\partial X} \right|_{X=\hat{X}_{k|k-1}} \left(X - \hat{X}_{k|k-1}\right) + \delta Z_k = \\ &= \tilde{F}\left(X_k^*, \hat{X}_{k|k-1}, h^k\right) + H_k X_k + \delta Z_k, \end{aligned}$$

где приняты обозначения:

$$\begin{aligned} \tilde{F}\left(X_k^*, \hat{X}_{k|k-1}, h^k\right) &= F\left(X_k^*, \hat{X}_{k|k-1}, h^k\right) - H_k \hat{X}_{k|k-1}; \\ H_k &= \left. \frac{\partial F\left(X_k^*, X, h^k\right)}{\partial X} \right|_{X=\hat{X}_{k|k-1}}; \end{aligned}$$

u_k^I , δZ_k , $X_k = (\Delta X_k, \Delta V_{Xk}, \Delta H_k, \Delta \theta_k, \Delta Y_k, \Delta V_{Yk}, \Delta \psi_k, \Delta \gamma)^T$ — значения векторов наблюдения, ошибок модели наблюдения с ковариационной матрицей $R_k = M\langle \delta Z_k \delta Z_k^T \rangle$ и ошибок навигационной системы в момент времени k .

Элементы H_{ijk}^I , $l = 1, \dots, 8$, матрицы H_k определяются выражениями (2.11), в которых градиенты $f'_{X_{ij}}$, $f'_{Y_{pq}}$, $h'_{X_{ij}}$, h'_{pq} рассчитываются в текущих точках местности X_{ij} , Y_{ij} , эти точки связаны с текущим местоположением ЛА, определяемым измеренными значениями X_k^* и прогнозируемым значением вектора ошибок $\hat{X}_{k|k-1}$.

3. ТОЧНОСТЬ СТЕРЕОСКОПИЧЕСКОЙ НАВИГАЦИОННОЙ СИСТЕМЫ ПРИ ПОЛЁТАХ НАД ЧИСТЫМИ ЛАНДШАФТАМИ И УРБАНИЗИРОВАННОЙ МЕСТНОСТЬЮ

3.1. Точность стереоскопической навигационной системы при полётах над чистыми ландшафтами

3.1.1. Беспосредственное оценивание координат по одной стереопаре

Выше было получено уравнение наблюдения стереоскопической навигационной системы

$$u^l = F(X^*, 0, h^k) + H(X^*, h^k)X + \delta Z.$$

Если обозначить $Z = u^l - F(X^*, 0, h^k)$, то получим уравнения наблюдения в стандартной, принятой в теории калмановской фильтрации форме

$$Z = H(X^*)X + \delta Z,$$

где $X = (\Delta X, \Delta V_x, \Delta H, \Delta \theta, \Delta Y, \Delta V_y, \Delta \psi, \Delta \gamma)^T$, а сам вектор наблюдений Z представляет собой множество разностей облучённости соответствующих точек первого и второго изображений стереопары, причём положение «псевдосоответственных» точек на втором изображении стереопары определяется по показаниям X^* корректируемой навигационной системы.

Матрица $H(X^*)$ может быть разбита на столбцы

$$H(X^*) = [H^1 \ H^2 \ \dots \ H^l \ \dots \ H^8];$$

каждый столбец H^l , $l = 1, \dots, 8$, содержит n столбцов H_i^l , $i = 1, \dots, n$ (по числу столбцов в совместной части двух изображений обрабатываемой стереопары):

$$H^l = [H_1^{lT} \ H_2^{lT} \ \dots \ H_i^{lT} \ \dots \ H_n^{lT}]^T, \text{ а } H_i^l = [H_{i-n}^l \ H_{i-n+1}^l \ \dots \ H_{ij}^l \ \dots \ H_{in}^l]^T,$$

причём H_{ij}^l для $l = 1, \dots, 8$ определяются соотношениями (2.11). Для обработки информации в беспосредственных СтНС используется метод калмановской фильтрации, состоящий в следующем [Аюки, 1979].

Если задана дискретная динамическая линейная система

$$X_{k+1} = A_k X_k + U_k + \xi_k, \quad k = 0, 1, \dots,$$

априорное распределение вектора X_0 нормальное: $p(X_0) \in N(m, G)$ с математическим ожиданием m и ковариационной матрицей G , а уравнение канала наблюдения имеет вид

$$Z_k = H_k X_k + \eta_k,$$

где ξ_k, η_k — нормальные, центрированные, не коррелированные с X_0 и взаимно не коррелированные дискретные белые шумы с ковариационными матрицами Q_k, R_k , то оптимальный в смысле минимума дисперсий ошибки оценивания алгоритм обработки информации содержит две группы уравнений:

уравнения оценок

$$\hat{X}_k = \hat{X}_{k|k-1} + K_k \Delta_k, \quad \Delta_k = Z_k - H_k \hat{X}_{k|k-1},$$

$$\hat{X}_{k+1|k} = A_k \hat{X}_k + U_k, \quad K_k = \Gamma_k H_k^T R, \quad \hat{X}_{0|-1} = m$$

и уравнения ковариаций

$$\Gamma_k = \Gamma_{k|k-1} - \Gamma_{k|k-1} H_k^T \left(H_k \Gamma_{k|k-1} H_k^T + R_k \right)^{-1} H_k \Gamma_{k|k-1},$$

$$\Gamma_k = \Gamma_{k|k-1} - \Gamma_{k|k-1} H_k^T \left(H_k \Gamma_{k|k-1} H_k^T + R_k \right)^{-1} H_k \Gamma_{k|k-1}.$$

В рассматриваемом случае $U = 0$, $Q = 0$, а в предположении несмещённости априорного распределения $p(X_0)$ и $m = 0$. Так как речь идет об одной-единственной стереопаре, полученной в начальный момент времени $k = 0$, то индекс дискретного времени $k = 0$ можно опустить, и уравнения фильтра Калмана примут вид

$$\begin{aligned} \hat{X} &= KZ, \quad K = \Gamma H^T R^{-1}; \\ \Gamma &= G - GH^T \left(HGH^T + R \right)^{-1} HG. \end{aligned} \quad (3.1)$$

Вектор \hat{X} содержит оценки ошибок позиционных координат $\Delta \hat{X}^H$, $\Delta \hat{Y}^H$, $\Delta \hat{H}$, горизонтальных составляющих скорости движения $\Delta \hat{V}_X$, $\Delta \hat{V}_Y$, курса $\Delta \hat{\psi}$ и вертикали $\Delta \hat{\theta}$, $\Delta \hat{\gamma}$.

Диагональные элементы матрицы Γ определяют апостериорные (после обработки информации стереопары) дисперсии ошибок навигационных координат σ_X^2 , $\sigma_{V_X}^2$, σ_H^2 , σ_θ^2 , σ_Y^2 , $\sigma_{V_Y}^2$, σ_ψ^2 , σ_γ^2 , а априорные значения этих дисперсий $\sigma_{X_0}^2$, $\sigma_{V_{X_0}}^2$, $\sigma_{H_0}^2$, $\sigma_{\theta_0}^2$, $\sigma_{Y_0}^2$, $\sigma_{V_{Y_0}}^2$, $\sigma_{\psi_0}^2$, $\sigma_{\gamma_0}^2$ задаются матрицей G , предполагаемой диагональной.

Соотношение (3.1) можно преобразовать к виду $\Gamma^{-1} = G^{-1} + H^T R^{-1} H$ [Аоки, 1979].

В работе [Белоглазов, Казарин, 1997] определены элементы матрицы $C = H^T R^{-1} H$; показано, что матрица $\Gamma^{-1} = G^{-1} + C$ будет блочно-диагональной. При больших дисперсиях начальных ошибок диагональные элементы матрицы Γ^{-1} незначительно отличаются от диагональных элементов матрицы C . На этом основании примем $\Gamma^{-1} \approx C$, увеличив, тем самым, апостериорные значения дисперсий ошибок оценок. Обратив матрицу C , найдем ковариационную матрицу Γ , диагональными элементами которой являются дисперсии σ_X^2 , $\sigma_{V_X}^2$, σ_H^2 , σ_θ^2 , σ_Y^2 , $\sigma_{V_Y}^2$, σ_ψ^2 , σ_γ^2 соответствующих величин, полученных после коррекции навигационных параметров по одной стереопаре.

На величину дискретности расстояния Δ в предыдущих рассуждениях было наложено лишь одно ограничение: Δ должно быть не меньше радиуса корреляции суммарной помехи $\delta Z = \delta f^I - \delta f^{II} - \delta Z^{II}$.

Рассмотрим вариант полёта над чистыми ландшафтами, когда используются не все элементарные приёмники, а только некоторая часть из них. Причём расстояние Δ между ними равно $N\delta$, и они по-прежнему образуют равномерную сетку в плоскости изображения xoy с числом строк $\frac{n}{N}$, числом столбцов $\frac{2n+1}{N}$ и количеством используемых элементарных приёмников $m = \frac{m_{\max}}{N^2}$, где $m_{\max} = (2n+1)n \approx 2n^2$. Такая ситуация может возникнуть, по крайней мере, в двух случаях:

- во-первых, когда для достижения нужной точности навигации нет необходимости использовать все элементарные приёмники ПЗС-матрицы;
 - во-вторых, чтобы не нарушались условия (2.12), принятые при выводе приводимых ниже уравнений, когда радиус корреляции шумов δZ больше δ ;
- при выводе этих формул полагалось, если $\Delta > \Delta_{\min} = \frac{H}{F} \delta$, то $m = 2 \frac{H^2 \text{tg}^2 \varphi}{\Delta^2}$.

Проведённые преобразования [Белоглазов, Казарин, 1997; Белоглазов, 2010] дали

$$\begin{aligned} \sigma_X &= \frac{2}{\sqrt{5}} \frac{\Delta \sigma_{\sim}}{\sigma_{f'} H \text{tg}^2 \varphi \sigma_{h'}}, & \sigma_Y &= \frac{2}{\sqrt{5}} \frac{\Delta \sigma_{\sim}}{\sigma_{f'} H \text{tg}^2 \varphi \sigma_{h'}}, \\ \sigma_{V_X} &= \sqrt{\frac{6}{5}} \frac{\Delta \sigma_{\sim} V}{\sigma_{f'} H^2 \text{tg}^3 \varphi \sigma_{h'}}, & \sigma_{V_Y} &= \sqrt{\frac{6}{5}} \frac{\Delta \sigma_{\sim} V}{\sigma_{f'} H^2 \text{tg}^3 \varphi \sigma_{h'}}, \\ \sigma_H &= \sqrt{\frac{6}{5}} \frac{\Delta \sigma_{\sim}}{\sigma_{f'} H \text{tg}^3 \varphi \sigma_{h'}}, & \sigma_{\psi} &= \sqrt{\frac{6}{5}} \frac{\Delta \sigma_{\sim}}{\sigma_{f'} H \text{tg}^3 \varphi \sigma_{h'}}, \\ \sigma_{\theta} &= \frac{\sqrt{3}}{2} \frac{\Delta \sigma_{\sim}}{\sigma_{f'} H^2 \text{tg}^3 \varphi}, & \sigma_{\gamma} &= \frac{\sqrt{3}}{2} \frac{\Delta \sigma_{\sim}}{\sigma_{f'} H^2 \text{tg}^3 \varphi}. \end{aligned} \quad (3.2)$$

На рис. 4 приведены графики, иллюстрирующие относительную точность оценивания координат $\sigma_X, \sigma_H, \sigma_{\psi}, \sigma_{\theta}$ стереоскопической навигационной системой при наблюдении одной стереопары в зависимости от высоты наблюдения H и градиента поля рельефа $\sigma_{h'}$; $a = \frac{\Delta \sigma_{\sim}}{\sigma_{f'}}$, $\text{tg} \varphi = 0,5$.

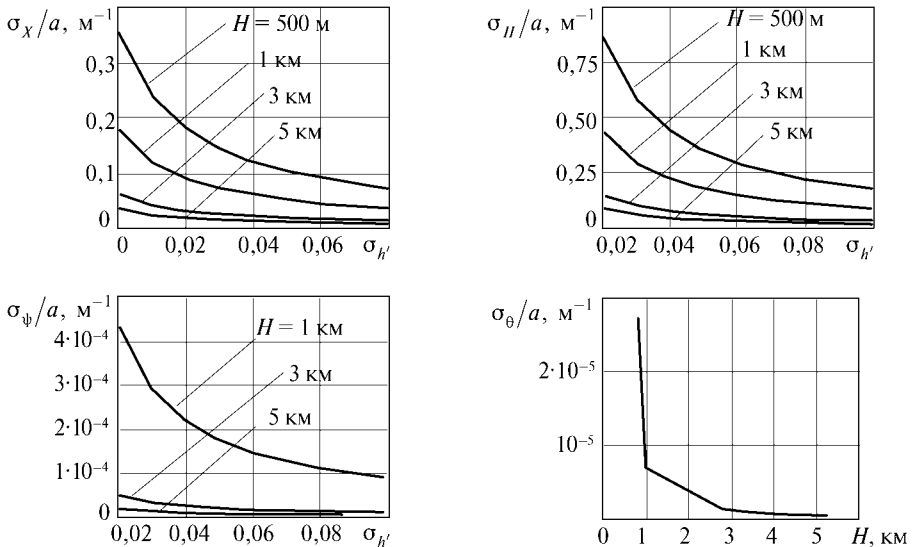


Рис. 4. Точность определения навигационных параметров по одной стереопаре изображений

3.1.2. Бесписковое оценивание координат по стереомножеству изображений

Необходимо оценить вектор

$$X_k = \left(\Delta X_k, \Delta V_{X_k}, \Delta H_k, \Delta \theta_k, \Delta Y_k, \Delta V_{Y_k}, \Delta \psi_k, \Delta \gamma_k \right)^T$$

в результате наблюдения ряда стереопар. В этой процедуре, кроме уравнений канала наблюдения, используются и уравнения контролируемого процесса:

$$\begin{aligned} \Delta X_{k+1} &= \Delta X_k + \Delta V_{X_k} \tau, \quad \Delta Y_{k+1} = \Delta Y_k + \Delta V_{Y_k} \tau, \quad \Delta V_{xk+1} = \Delta V_{xk}, \\ \Delta V_{yk+1} &= \Delta V_{yk}, \quad \Delta H_{k+1} = \Delta H_k, \quad \Delta \psi_{k+1} = \Delta \psi_k, \quad \Delta \theta_{k+1} = \Delta \theta_k, \quad \Delta \gamma_{k+1} = \Delta \gamma_k \end{aligned}$$

или в матричном виде:

$$X_{k+1} = A_k X_k.$$

Матрица A_k является блочно-диагональной и может быть определена блоками A_k^n и A_k^6 , соответствующими продольному и боковому каналам оценивания:

$$A_k = \begin{bmatrix} A_k^n & 0 \\ 0 & A_k^6 \end{bmatrix}, \quad \text{где } A_k^n = A_k^6 = \begin{bmatrix} 1 & \tau & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Такой вид матрицы A_k позволяет говорить о некоторой независимости продольного и бокового каналов оценивания. Поведение продольного канала определяется подвектором

$$X_k^n = \left(\Delta X_k, \Delta V_{xk}, \Delta H_k, \Delta \theta_k \right)^T,$$

а бокового — подвектором

$$X_k^6 = \left(\Delta Y_k, \Delta V_{yk}, \Delta \psi_k, \Delta \gamma_k \right)^T;$$

при этом $X_k = \begin{bmatrix} X_k^n \\ X_k^6 \end{bmatrix}$.

Процессы оценивания носят связанный характер

$$\hat{X}_k = \hat{X}_{k|k+1} + K_k \Delta_k,$$

$$\Delta_k = Z_k - \tilde{F} \left(X_k^*, \hat{X}_{k|k-1}, h^k \right) - H_k \hat{X}_{k|k-1} = Z_k - F \left(X_k^*, \hat{X}_{k|k-1}, h^k \right),$$

$$\hat{X}_{k+1|k} = A_k \hat{X}_k, \quad K_k = \sigma_{\sim}^{-2} \Gamma_k H_k^T, \quad \hat{X}_{0|-1} = 0,$$

так как вектор невязок наблюдений Δ_k для продольного и бокового каналов один и тот же.

Уравнения же ковариаций для продольного и бокового каналов в силу блочно-го строения матриц $H_k^T R H_k$ и G оказываются независимыми и определяются следующими рекуррентными зависимостями:

в продольном канале:

$$\left(\Gamma_k^n \right)^{-1} = \left(\Gamma_{k|k-1}^n \right)^{-1} + \sigma_{\sim}^{-2} \left(H_k^n \right)^T H_k^n, \quad \Gamma_{k+1|k}^n = A_k^n \Gamma_k^n \left(A_k^n \right)^T,$$

$$\Gamma_{0|1}^n = \text{diag}\left(\sigma_{X_0}^2, \sigma_{V_{X_0}}^2, \sigma_{H_0}^2, \sigma_{\theta_0}^2\right), k = 0, 1, \dots;$$

в боковом канале

$$\left(\Gamma_k^6\right)^{-1} = \left(\Gamma_{k|k-1}^6\right)^{-1} + \sigma_{\sim}^{-2} \left(H_k^6\right)^T H_k^6, \quad \Gamma_{k+1|k}^6 = A_k^6 \Gamma_k^6 \left(A_k^6\right)^T,$$

$$\Gamma_{0|1}^6 = \text{diag}\left(\Delta Y_k, \Delta V_{y_k}, \Delta \psi_k, \Delta \gamma_k\right), k = 0, 1, \dots;$$

где Γ_k^n, Γ_k^6 — ковариационные матрицы ошибок оценивания координат продольного и бокового каналов соответственно, причём

$$\Gamma_k = \begin{bmatrix} \Gamma_k^n & 0 \\ 0 & \Gamma_k^6 \end{bmatrix}.$$

На рис. 5 отражена зависимость относительных среднеквадратических ошибок оценивания навигационных параметров от числа использованных стереопар k для различных градиентов рельефа $\sigma_{h'}$ и разных высот полёта H .

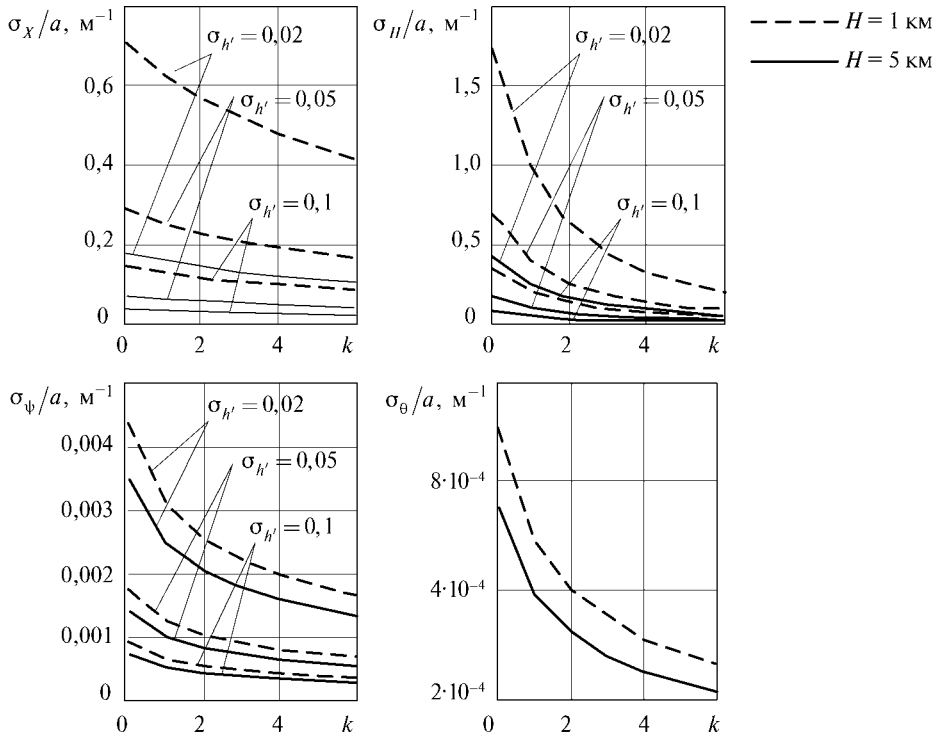


Рис. 5. Зависимости относительной точности определения навигационных параметров от числа используемых стереопар

3.2. Точность стереоскопической навигационной системы при полётах над урбанизированной местностью

3.2.1. Беспосредственное оценивание координат по одной стереопаре

Перейдём к рассмотрению точности стереоскопической навигационной системы при полёте над урбанизированной местностью, т. е. в ситуации, когда высоты наблюдаемой стереопарой 3D-поверхности существенно отличаются от топографической карты рельефа из-за наличия растительности и строений. В статье рассматриваются стереоскопические навигационные системы, использующие в качестве эталона топографические карты рельефа местности. Чтобы иметь такую возможность, необходимо разработать методики, позволяющие изолировать на топографической карте участки, заполненные растительностью, и работать только с теми приёмниками стереопары, которые принимают сигналы, отражённые от топографического рельефа. Участки коррекции должны выбираться с учётом характеристик поля рельефа и оптического поля местности. Обрабатываемые точки на первом изображении стереопары выбираются из системы точек, рассмотренных в

подразд. 3.1.1. Подставляя $\Delta = \sqrt{2} \frac{H \operatorname{tg} \varphi}{\sqrt{m}}$ в (3.2), получим новые зависимости

$$\left. \begin{aligned} \sigma_X = \sigma_Y &= 2\sqrt{\frac{2}{5}} \frac{\sigma_{\sim}}{\sqrt{m} \cdot \sigma_{f'} \sigma_{h'} \operatorname{tg} \varphi}, \\ \sigma_{V_X} = \sigma_{V_Y} &= 2\sqrt{\frac{3}{5}} \frac{\sigma_{\sim} V}{\sqrt{m} \cdot \sigma_{f'} \sigma_{h'} H \operatorname{tg}^2 \varphi}, \\ \sigma_H &= 2\sqrt{\frac{3}{5}} \frac{\sigma_{\sim}}{\sqrt{m} \cdot \sigma_{f'} \sigma_{h'} \operatorname{tg}^2 \varphi}, \\ \sigma_{\psi} &= 2\sqrt{\frac{3}{5}} \frac{\sigma_{\sim}}{\sqrt{m} \cdot \sigma_{f'} \sigma_{h'} H \operatorname{tg}^2 \varphi}, \\ \sigma_{\theta} &= \frac{\sigma_{\gamma}}{\sqrt{2}} = \sqrt{\frac{3}{2}} \frac{\sigma_{\sim}}{\sqrt{m} \cdot \sigma_{f'} H \operatorname{tg}^2 \varphi}, \\ \sigma_{\sim}^2 &= \frac{B^2}{H^2} \sigma_{f'}^2 \sigma_k^2 + 2\sigma_{\text{И}}^2. \end{aligned} \right\} \quad (3.3)$$

Формулы (3.3) характеризуют точность стереоскопической навигационной системы при полёте над урбанизированной местностью. В них m — число оставленных для обработки с использованием карты рельефа местности точек первого изображения. Оставленные для обработки точки не образуют, в отличие от варианта подразд. 3.1.1, равномерной сетки, но расстояние между ними Δ больше или равно

$$\sqrt{2} \frac{H \operatorname{tg} \varphi}{\sqrt{m}}.$$

На рис. 6 (см. с. 83) приведены графики точности стереоскопической навигационной системы при полёте над урбанизированной местностью на высоте

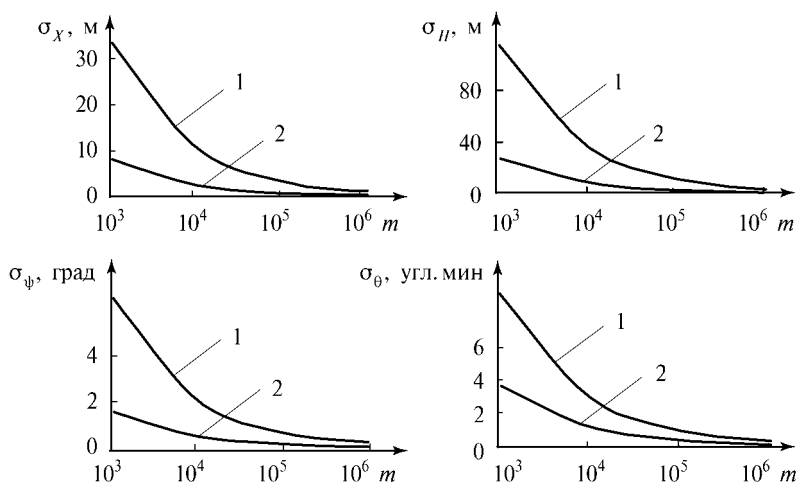


Рис. 6. Точность определения навигационных параметров при полёте над урбанизированным участком местности

$H = 1$ км. Рассмотрено два варианта участков коррекции, различающихся характеристиками рельефа и ошибками карты рельефа: вариант 1 — $\sigma_k = 25$ м, $\sigma_{h'} = 0,03$; вариант 2 — $\sigma_k = 10$ м, $\sigma_{h'} = 0,05$; $\frac{B}{H} = \text{tg}\theta = 0,5$.

3.2.2. Выбор участков коррекции

Следующим шагом развития предлагаемого подхода является разработка методики выбора участков коррекции, позволяющая изолировать на топографической карте участки, заполненные строениями и растительностью, и работать только с теми приёмниками, которые принимают сигналы, отражённые от топографического рельефа. Коррекция может осуществляться как при полёте над чистыми ландшафтами, когда отсутствуют строения и высота растительности не превышает ошибки карты рельефа, так и над урбанизированной местностью, если эти условия не выполняются.

Эффективность стереоскопической коррекции зависит от многих факторов: градиентов поля рельефа и оптического поля участка коррекции, высоты полёта и угла зрения видеокамеры, ошибки карты рельефа, дисперсии и радиуса корреляции ошибок элементарных приёмников матрицы ПЗС, количества используемых соответственных точек на двух изображениях стереопары, точности корректируемой базовой навигационной системы в определении позиционных, скоростных и угловых координат ЛА.

Последовательность действий при выборе участка коррекции следующая:

1) исходя из цели, условий и траектории полёта определяются высота H , скорость V и необходимая точность знания позиционных и угловых координат ЛА в результате проведённой коррекции;

2) выбирается базовая (корректируемая) навигационная система ЛА, что определяет СКО в начале коррекции $\left\{ \sigma_{X_0}, \sigma_{Y_0}, \sigma_{H_0}, \sigma_{\psi_0}, \sigma_{\theta_0}, \sigma_{V_0} \right\}$ централизованных позиционных и угловых ошибок этой системы;

3) выбирается видеокамера и ПЗС-матрица, это определяет угол зрения видеокамеры 2φ , СКО σ_{Π} и радиус корреляции ошибок элементарных приёмников ПЗС-матрицы;

4) рассматривается предполагаемый участок коррекции, это определяет степень урбанизации участка, СКО и радиусы корреляции градиентов поля рельефа $\sigma_{H'}$, $\rho_{H'}$ и оптического поля $\sigma_{f'}$, $\rho_{f'}$, а также максимально возможное число обрабатываемых соответственных точек стереопары $m_{\max} = \frac{2F^2 \operatorname{tg}^2 \varphi}{\Delta^2}$ (F — фокусное расстояние видеокамеры) или стереомножества изображений;

5) выбирается метод составления эталона поля рельефа, что определяет СКО ошибки карты рельефа σ_k ;

6) выбирается алгоритм стереоскопической коррекции базовой навигационной системы и по формулам точности для варианта стабилизированного горизонтального полёта над урбанизированным участком коррекции рассчитываются СКО ошибок скорректированной навигационной системы σ_x , σ_y , σ_{V_x} , σ_{V_y} , σ_H , σ_{ψ} , σ_{θ} , σ_{γ} в зависимости от числа обрабатываемых соответственных точек m ; если такое число допустимо ($m \leq m_{\max}$), то рассматриваемый участок пригоден для проведения коррекции;

7) из множества пригодных участков выбирается наиболее выгодный по дополнительным показателям, например, требованиям к вычислительной системе.

4. КОМБИНИРОВАННЫЙ АЛГОРИТМ СТЕРЕОСКОПИЧЕСКОЙ НАВИГАЦИОННОЙ СИСТЕМЫ, ОСНОВАННЫЙ НА ТЕОРИИ СОВМЕСТНОЙ ОПТИМАЛЬНОЙ ФИЛЬТРАЦИИ, ИДЕНТИФИКАЦИИ И ПРОВЕРКИ ГИПОТЕЗ (вариант X, C, D / X, C, D таблицы)

Из теории корреляционно-экстремальных навигационных систем [Красовский и др., 1979] известно, что беспойсковыми алгоритмами могут быть ликвидированы начальные отклонения, не превышающие первого экстремума — минимума корреляционной функции. Аналогичным свойством обладают и беспойсковые СтНС.

Рассмотрим такую систему, когда начальные ошибки навигационной системы могут быть достаточно велики, и возможна потеря работоспособности беспойсковой СтНС, основанной на линеаризации уравнений наблюдения и использующей калмановскую фильтрацию. В этом случае необходимо применять иные алгоритмы обработки, не имеющие ограничений на величину начальной ошибки навигационной системы и связанные с поиском правильного местоположения ЛА.

В момент времени, когда получено второе изображение стереопары, назначается доверительный квадрат $2\Delta_{\max} \times 2\Delta_{\max}$ с центром в точке X^{*SII} , Y^{*SII} , где Δ_{\max} — возможная максимальная ошибка корректируемой навигационной системы. Задается шаг дискретизации по координатам L , и в большом доверительном квадрате на плоскости $O_3 X_3 Y_3$ рассматриваются меньшие квадраты $L \times L$ и СтНС проверяет гипотезы о наличии ЛА в одном из этих малых квадратов. Произвольная гипотеза D_{pq} предполагает, что ЛА находится в малом квадрате \square_{pq} с центром в точке $X^{*SII} = pL$, $Y^{*SII} = qL$ (рис. 7, см. с. 85).

Величина L выбирается из условия, что на расстоянии L справедливо предположение о возможности линеаризации уравнений (2.6).

Синтезируем комбинированный алгоритм СтНС, использующий одну стереопару.

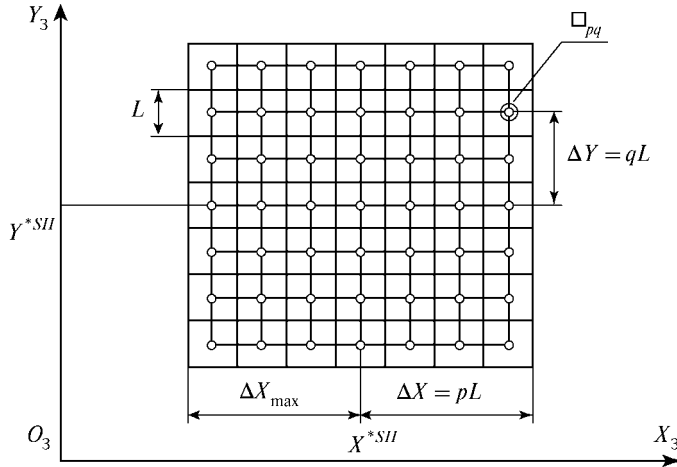


Рис. 7. Доверительный квадрат в плоскости $O_3X_3Y_3$

4.1. Гипотетическое уравнение наблюдения для стереопары при проверке гипотезы D_{pq}

Введем в рассмотрение гипотезу D_{pq} и вектор неизвестных параметров C . Гипотеза D_{pq} предполагает позиционные ошибки навигационной системы ограниченными

$$pL - \frac{L}{2} \leq \Delta X_{pq} < pL + \frac{L}{2}, \quad qL - \frac{L}{2} \leq \Delta Y_{pq} < qL + \frac{L}{2}.$$

Вектор неизвестных параметров C содержит ошибки знания высоты, скорости полёта, угловой ориентации и ошибки положения позиционных координат ЛА $\Delta X_{pq}, \Delta Y_{pq}$ внутри малого квадрата \square_{pq} :

$$C = (\Delta X_{pq}, \Delta V_X, \Delta H, \Delta \theta, \Delta Y_{pq}, \Delta V_Y, \Delta \psi, \Delta \gamma)^T.$$

Поэтому при рассмотрении гипотезы D_{pq} полный вектор ошибок бортовых измерений X_{pq} предполагается равным

$$\begin{aligned} X_{pq} &= X(C, D_{pq}) = \\ &= (pL + \Delta X_{pq}, \Delta V_X, \Delta H, \Delta \theta, qL + \Delta Y_{pq}, \Delta V_Y, \Delta \psi, \Delta \gamma)^T. \end{aligned} \quad (4.1)$$

Введем в рассмотрение $F[X^*, X(C, D_{pq}), h]$ — вектор сигналов элементарных приёмников u^{II} в соответствующих гипотетических точках второго изображения стереопары в случае, если верна гипотеза D_{pq} , а также вектор $\tilde{X} = C$ — часть вектора ошибок бортовых измерений, совпадающая с вектором неизвестных параметров; X^* — это вектор измеренных значений элементов внешнего ориентирования изображений (2.5), полученных навигационной системой ещё до использования информации второго изображения стереопары.

Линеаризуя $F[X^*, X(C, D_{pq}), h]$ подобно (2.9) в окрестности точки $X(0, D_{pq}) = (pL, 0, 0, 0, 0, qL, 0, 0, 0)^T$ и $h = h^k$, получим уравнение для рассматриваемого гипотетического случая

$$\begin{aligned}
 u^I = F[X^*, X(0, D_{pq}), h^k] + \frac{\partial F[X^*, X(C, D_{pq}), h]}{\partial \tilde{X}} \Bigg|_{\substack{h=h^k \\ \tilde{X}=0}} \tilde{X} - \\
 - \frac{\partial F[X^*, X(C, D_{pq}), h]}{\partial h} \Bigg|_{\substack{h=h^k \\ \tilde{X}=0}} \delta h + \delta Z^I.
 \end{aligned} \tag{4.2}$$

Примем, сохраняя пока обозначения разд. 1, $Z = u^I$ — вектор сигналов приёмников первого изображения стереопары,

$$F[X^*, D_{pq}] = F[X^*, X(0, D_{pq}), h^k] \tag{4.3}$$

— вектор гипотетических (соответствующих D_{pq}) сигналов приёмников второго изображения стереопары, рассчитанный на основе вектора измеренных значений элементов внешнего ориентирования изображений X^* .

$$\begin{aligned}
 H(X^*, D_{pq}) = \frac{\partial F[X^*, X(C, D_{pq}), h]}{\partial \tilde{X}} \Bigg|_{\substack{h=h^k \\ \tilde{X}=0}}, \\
 \frac{\partial F[X^*, X(C, D_{pq}), h]}{\partial h} \Bigg|_{\substack{h=h^k \\ \tilde{X}=0}} \delta h = \delta Z^II.
 \end{aligned} \tag{4.4}$$

В рассматриваемых условиях (до начала идентификации и использования второго изображения стереопары)

$$F(C, D_{pq}) = F(X^*, D_{pq}), \quad H(C, D_{pq}) = H(X^*, D_{pq})$$

и гипотетическое уравнение наблюдения примет вид

$$\begin{aligned}
 Z = H(X^*, D_{pq}) \tilde{X} + F(X^*, D_{pq}) + \delta Z, \\
 \delta Z = \delta Z^I + \delta Z^II.
 \end{aligned} \tag{4.5}$$

4.2. Частный случай теории совместной оптимальной фильтрации, идентификации и проверки гипотез при проведении единственного измерения, используемый при обработке стереопар

Именно этот вариант используется далее при синтезе оптимальных алгоритмов обработки стереопары. Запишем уравнения разд. 1 для единственного измерения, проведённого в момент $k = 0$. Индекс дискретного времени $k = 0$ опускаем, так как в этом случае он не несет никакой информации. Вместо обозначения гипотез D_j примем обозначение D_{pq} , что соответствует исследуемой ситуации. Также предполагается $U_k = 0, \xi_k = 0$. В результате приходим к следующей постановке задачи.

Задано:

— $\{D_{pq}\}$ — полное множество гипотез и вероятности гипотез

$$p_{pq} = P(D_{pq}), \quad \sum P(D_{pq}) = 1;$$

- случайный вектор неизвестных параметров C и множество условных плотностей распределения $\{p(c/D_{pq})\}$;
- канал наблюдения

$$Z = H(X^*, D_{pq})\tilde{X} + F(X^*, D_{pq}) + \eta;$$

- множество априорных условных нормальных распределений

$$p(X_{pq}/C, D_{pq}) \in N[m(c, D_{pq}), G(c, D_{pq})],$$

где m, G — условное математическое ожидание и ковариационная матрица; также задана ковариационная матрица $R(c, D_{pq})$ шума η .

Требуется по наблюдению Z оптимальным образом оценить вектор состояния \hat{X} , вектор неизвестных параметров \hat{c} , выбрать гипотезу $D = D_{v,u}$, имевшую место при проведении наблюдения Z . Критерий оптимальности — максимум апостериорной плотности-вероятности $p(x, c, D_{pq}/z)$.

Составляющие вектора искоемых ошибок бортовых измерений

$$X = (\Delta X, \Delta V_X, \Delta H, \Delta \theta, \Delta Y, \Delta V_Y, \Delta \psi, \Delta \gamma)^T$$

в комбинированном алгоритме разбиваются на две части:

$$\Delta X = pL + \Delta X_{pq}, \Delta Y = qL + \Delta Y_{pq} \text{ и оставшая часть вектора (4.1)}$$

$$(\Delta V_X, \Delta H, \Delta \theta, \Delta V_Y, \Delta \psi, \Delta \gamma)^T.$$

Значения $\Delta X, \Delta Y$ могут быть и больше, превышая зону линейности L , допускающую линеаризацию уравнения (2.9). Они определяются путём проверки гипотез D_{pq} и одновременного уточнения координат $\Delta X_{pq}, \Delta Y_{pq}$ внутри малого квадрата \square_{pq} .

Решение содержит три группы уравнений:

уравнения оценок

$$\hat{x}_{crq} = m_{crq} + K_{crq} \Delta_{crq},$$

где $K_{crq} = \Gamma_{crq} H_{crq} R_{crq}$, $\Delta_{crq} = z - \hat{z}_{crq}$, $\hat{z}_{crq} = H_{crq} m_{crq} + F_{crq}$, и приняты сокращенные обозначения:

$$H_{pq} = H(X^*, D_{pq}), F_{pq} = F(X^*, D_{pq}), R_{pq} = R(X^*, D_{pq}),$$

$$\hat{x}_{crq} = M \langle X | C = c, D = D_{pq}, Z = z \rangle, m_{crq} = m(c, D_{pq}), G_{crq} = G(c, D_{pq}),$$

$$\Gamma_{crq} = M \left\langle (X - \hat{x}_{crq})(X - \hat{x}_{crq})^T | C = c, D = D_{pq}, Z = z \right\rangle;$$

уравнения ковариаций

$$\Gamma_{crq} = G_{crq} - G_{crq} H_{crq}^T (H_{crq} G_{crq} H_{crq}^T + R_{crq})^{-1} H_{crq} G_{crq},$$

или эквивалентная формула

$$\Gamma_{crq}^{-1} = G_{crq}^{-1} + H_{crq}^T R_{crq}^{-1} H_{crq};$$

уравнения функционалов

$$I(C, D_{pq}) = I_{crq} = -2 \ln p(c, D_{pq}) + \|\Delta_{crq}\|_{\Sigma_{crq}^{-1}}^2 + \varepsilon_{crq},$$

$$\Sigma_{crq}^{-1} = R_{crq}^{-1} - R_{crq}^{-1} H_{crq} \Gamma_{crq} H_{crq}^T R_{crq}^{-1}, \quad \varepsilon_{crq} = \ln |\tilde{A}_{crq}| + \ln |R_{crq}|.$$

Оптимальный по критерию максимума условной апостериорной плотности вероятности алгоритм сводится к приведённой ниже последовательности операций.

1. Для любой гипотезы D_{pq} находится вектор неизвестных параметров c_{pq} , который минимизирует I_{cpq} :

$$c_{pq} = \arg \inf_{\Omega_{pq}} I_{cpq}, \text{ т. е. } I_{c_{pq}, pq} = \inf_{c \in \Omega_{pq}} I_{cpq}.$$

Методы идентификации вектора неизвестных параметров могут быть различными [Льюнг, 1991; Красовский, 1987]. Они зависят от конкретной задачи (навигации, наведения, определения координат наблюдаемых объектов), траектории движения ЛА, составляющих вектора неизвестных параметров. Могут использоваться метод градиента или метод наискорейшего спуска. Для метода наискорейшего спуска характерно относительно малое время выхода в окрестность экстремума при крупных шагах движения на начальном этапе поиска. Также возможно применение различных комбинаций методов поиска [Расстригин, 1968, 1980; Иоффе, Тихомиров, 1974; Красовский, 1987].

2. Определяется гипотеза

$$\hat{D} = D_{\mu\nu} = \arg \min_{D_{pq}} I_{c_{pq}, pq},$$

а общей оптимальной оценкой \hat{c} вектора неизвестных параметров (ошибок бортовых измерений) является $\hat{c}_{\mu\nu}(z)$.

3. Формируется оптимальная оценка вектора состояния

$$\hat{x} = \hat{x}_{c_{\mu\nu}, \mu\nu}(z).$$

Тройка $D_{\mu\nu}, \hat{c}_{\mu\nu}, \hat{x}_{c_{\mu\nu}, \mu\nu}(z)$ является решением задачи оптимизации.

Структурная схема оптимального алгоритма в момент, предшествующий началу работы оптимизатора, представлена на рис. 8 (см. с. 89).

Исходные значения величин, входящих в структурную схему рис. 8, определяются следующим образом.

Априорные характеристики $p(c, D_{pq})$ и $G(c, D_{pq})$ задаются из физических соображений с учётом точностных характеристик навигационной системы, R_{cpq} определяются равенством (2.12), $m_{cpq} = X(0, D_{pq})$. F_{pq} для каждой из проверяемых гипотез D_{pq} рассчитывается на основе измеренных навигационной системой значений элементов внешнего ориентирования изображений и карты рельефа местности. Матрица канала наблюдения H_{pq} , согласно (2.11), определяется для каждой гипотезы характеристиками рельефа и оптического поля в малом квадрате \square_{pq} .

Оптимизатор для каждой из проверяемых гипотез генерирует последовательность векторов неизвестных параметров c , принадлежащих области их возможных значений Ω_{pq} , зависящей от точности навигационной системы. В процессе идентификации и поиска экстремума минимизируемого функционала I_{cpq} для каждого проверяемого значения c находятся соответствующие ему новые матрица H_{cpq} и вектор F_{cpq} (вместо начальных H_{pq} и F_{pq}). После определения элементов матрицы H_{cpq} в соответствии с приведенными ранее формулами рассчитываются Γ_{cpq} , Σ_{cpq}^{-1} , K_{cpq} , ε_{cpq} и методами идентификации или поиска определяется нижняя грань функционала I_{cpq} .

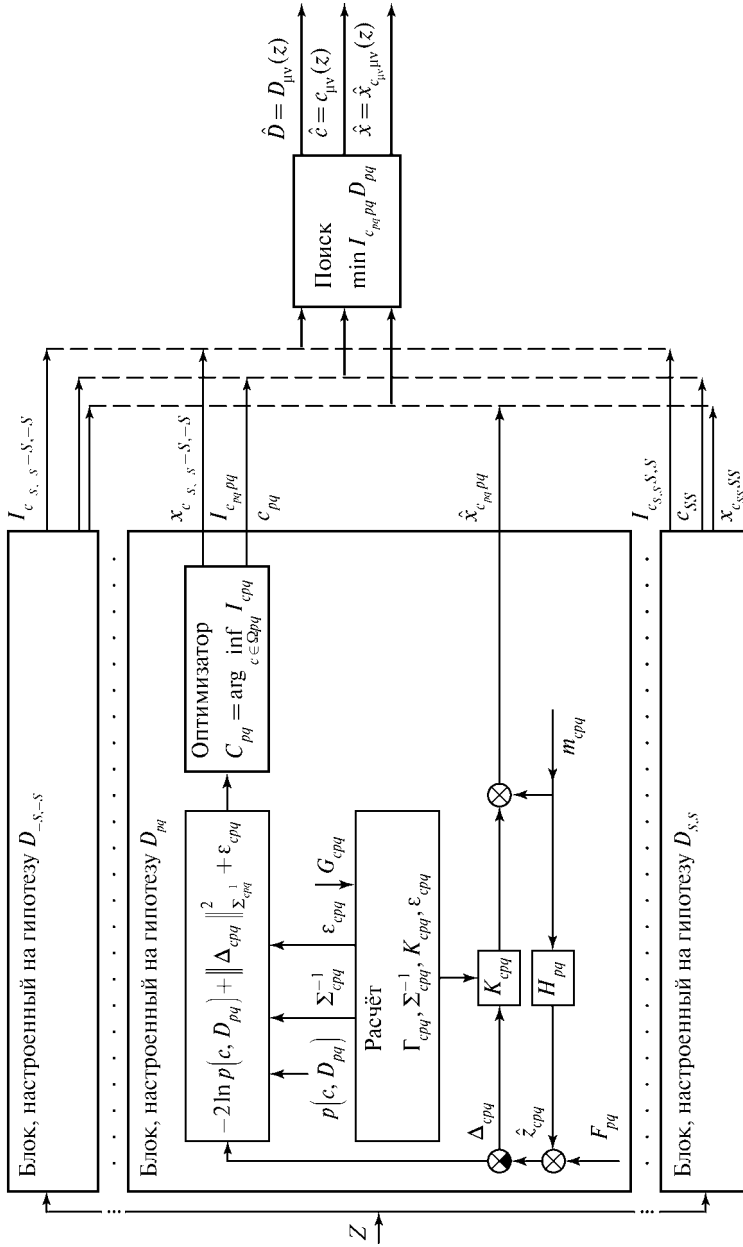


Рис. 8. Структурная схема оптимального алгоритма в момент, предшествующий началу работы оптимизатора

Заключение

В результате работы всего алгоритма, целью которого является оценка ошибок стереоскопической навигационной системы, каждой из рассматриваемых точек первого изображения стереопары противопоставляется соответствующая точка на втором изображении стереопары. В силу оптимальности используемого алгоритма обработки стереопар и физического смысла минимизируемого функционала средний квадрат разности координат найденных соответствующих точек является минимальным.

В статье оценена точность СтНС, предназначенной для решения задач высокоточной навигации, наведения и дистанционного зондирования местности. Произведена линеаризация фотограмметрических уравнений СтНС в режиме стабилизированного горизонтального полёта ЛА. Определены уравнения наблюдения СтНС с учётом единственной стереопары и стереомножества изображений.

Литература

- Аоки М. (1979) Оптимизация стохастических систем. М.: Наука, 1979.
- Белоглазов И. Н. (1983) Оптимальные совместные оценивания и идентификация в дискретных линейных системах // ДАН СССР. 1983. Т. 273. № 4.
- Белоглазов И. Н. (1977) Рекуррентно-поисковые алгоритмы оценивания // ДАН СССР. 1977. Т. 236. № 2.
- Белоглазов И. Н., Казарин С. Н. (1998) Совместное оптимальное оценивание, идентификация и проверка гипотез в дискретных динамических системах // Изв. РАН. Теория и системы управления. 1998. № 4.
- Белоглазов И. Н. (2003а) Синтез оптимального поискового алгоритма и оценка точности иконической навигационной системы, использующей изображения поверхностных полей Земли // Изв. РАН. Теория и системы управления. 2003. № 5.
- Белоглазов И. Н. (2003б) Автоматизированные системы наземных комплексов сбора и обработки данных воздушной разведки. М.: ВВИА им. проф. Н. Е. Жуковского, 2003.
- Белоглазов И. Н. (2007) Теория совместной оптимальной фильтрации, идентификации и проверки гипотез в дискретных динамических системах, её частные случаи и использование в задачах навигации и дистанционного зондирования // Информационные технологии в разработках сложных систем: Тр. ГосНИИАС. 2007. Вып. 1. № 15.
- Белоглазов И. Н. (2009) Синтез алгоритмов стереоскопической системы навигации, наведения и дистанционного зондирования местности // Изв. РАН. Теория и системы управления. 2009. № 5.
- Белоглазов И. Н. (2010) Точность стереоскопической навигационной системы при полётах над чистыми ландшафтами и урбанизированной местностью // Изв. РАН. Теория и системы управления. 2010. № 5.
- Белоглазов И. Н., Джанджгава Т. Н., Чигин Г. П. (1985) Основы навигации по геофизическим полям. М.: Наука, 1985.
- Белоглазов И. Н., Казарин С. Н. (1997) Стереоскопическая навигационная система // Изв. РАН. Теория и системы управления. 1997. № 6.
- Белоглазов И. Н., Казарин С. Н. (1999) Автоматическое стереоскопическое распознавание наблюдаемых сцен // Изв. РАН. Теория и системы управления. 1999. № 6.
- Белоглазов И. Н., Коваленко В. П., Халтобин В. М. (1987) Авиационная светотехника. М.: ВВИА им. проф. Н. Е. Жуковского, 1987.
- Блохинов Ю. Б., Грибов Д. А., Чернявский А. С. (2008) Задача привязки изображений для некоторых случаев ракурсной фотосъёмки // Изв. РАН. Теория и системы управления. 2008. № 6.

- Иоффе А. Д., Тихомиров В. М.* (1974) Теория экстремальных задач. М.: Наука, 1974.
- Красовский А. А.* (1987) Справочник по теории автоматического управления. М.: Наука, 1987.
- Красовский А. А., Белоглазов И. Н., Чигин Г. П.* (1979) Теория корреляционно-экстремальных навигационных систем. М.: Наука, 1979.
- Красовский А. А., Белоглазов И. Н.* (2000) Методы оптимальной и субоптимальной фильтрации в САУ // Машиностроение: Энцикл. Т. 1–4. М.: Машиностроение, 2000.
- Лобанов А. Н.* (1984) Фотограмметрия. М.: Недра, 1984.
- Льюнг Л.* (1991) Идентификация систем. М.: Наука, 1991.
- Расстригин Л. А.* (1968) Статистические методы поиска. М.: Наука, 1968.
- Расстригин Л. А.* (1980) Современные принципы управления сложными объектами. М.: Сов. радио, 1980.
- Савельев И. В.* (1967) Курс общей физики. Т. 3. М.: Наука, 1967.
- Чернявский А. С.* (2008) Робастная схема оценки параметров модели на основе метода роя частиц в задаче стереоотожествления // Изв. РАН. Теория и системы управления. 2008. № 5.
- Beloglazov I. N., Kazarin S. N.* (1999) Stereoscopic Navigation and Observation Systems // 6th Saint Petersburg Intern. Conf. Integrated Navigation Systems, Russia. 1999.
- Beloglazov I. N., Kazarin S. N.* (2001) Optimal Algorithm of the Nonlinear filtering, Identification and Hypothesis Checking for Dynamical Systems // 5th IFAC Symp. "Nonlinear Control Systems". Saint Petersburg, Russia. 2001.

THEORETICAL ASPECTS AND APPENDICES OF STEREOSCOPIC NAVIGATION SYSTEMS, POINTING AND AREA REMOTE SENSING

I. N. Beloglazov

MECS of the Air Force Academy n. a. Professors N. E. Zhukovsky and Ju. A. Gagarin, 125167, Moscow, 3 Planetary st.

In article stereoscopic navigation systems (SNS), intended for the decision of problems of high-precision navigation, pointing and area remote sensing are investigated. The summary of the theory of a joint optimum filtration, identification and check of hypotheses in discrete dynamic system of special type is resulted. The combined algorithm of SNS, based on this theory is offered. The linearization of SNS photogrammetric equations in a mode of the stabilized horizontal flight of a flying machine is made. The algorithm of image processing, providing maximum a posteriori joint density-probability of a vector of a condition dynamic system, a vector of unknown parameters and the chosen hypothesis is synthesized. The equations characterizing accuracy of SNS at flights over pure landscapes and the urbanized area are received and the procedure of a choice of correction sites is considered.

Keywords: stereoscopic systems; navigation; remote sensing; optimum filtration; area.

АЛГОРИТМ АНАЛИЗА И ПРИНЯТИЯ РЕШЕНИЯ В ЗАДАЧЕ СЕЛЕКЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ НАЗЕМНЫХ СЦЕН

Ю. Б. Блохинов, В. В. Гнилицкий, В. В. Инсаров, А. С. Чернявский

*Федеральное государственное унитарное предприятие «Государственный научно-исследовательский институт авиационных систем» (ФГУП «ГосНИИАС»),
125319, Москва, Викторенко, 7*

Одним из подходов к задаче обнаружения объектов на изображениях является сравнение текущего изображения (ТИ) с заранее заданным эталонным изображением объекта (ЭИ). Рассматривается задача сравнения ТИ, полученного с борта летательного аппарата, и ЭИ наземной сцены, причём эталонное изображение задается в виде набора контуров, без учёта яркостей объектов. Предложен алгоритм сопоставления ТИ и ЭИ, основанный на анализе поля направлений градиентов яркости ТИ и нормалей к контурам ЭИ. Также рассматривается иерархическая процедура учёта априорной информации о взаимном расположении объектов на сцене, позволяющая решить задачи локализации объектов сцены и сборки сцены. Алгоритм применяется к набору фотоснимков наземных сцен, содержащих от двух до четырёх объектов. Исследуется точность обнаружения объектов в зависимости от уровней шума.

Ключевые слова: селекция объектов, поиск по эталону, сопоставление изображений, контурный эталон, поле градиентов.

Введение

Одним из перспективных направлений в совершенствовании систем навигации летательных аппаратов является использование технологии машинного зрения. При этом центральная проблема — разработка и реализация алгоритмов обработки изображений и распознавания образов, обеспечивающих надежную автоматическую селекцию выбранного объекта наземной сцены и определение координат этого объекта при ограничении времени решения в бортовой вычислительной системе. В достаточно общей постановке эта проблема может быть сформулирована следующим образом. Во-первых, априорная информация о сцене (типы, количество, размеры объектов и их взаимное расположение, включая возможные значения ошибок этих характеристик) используется для формирования эталонного изображения сцены или её эталонного описания в виде некоторой структуры данных. Во-вторых, апостериорная информация о наблюдаемой в процессе полёта летательного аппарата сцене, включающая характеристики пассивного бортового сенсора различного диапазона электромагнитного спектра, используется для формирования текущего изображения сцены или текущего описания, которое отличается пространственно-временной трансформацией относительно эталонного изображения. Сопоставление полученного таким образом текущего и эталонного изображения позволяет решить задачу локализации объектов сцены, сборки сцены, принятия решения и, в конечном счёте, селекции объекта и оценки значений

Блохинов Юрий Борисович — начальник лаборатории, кандидат физико-математических наук, e-mail: achern@gosniias.ru.

Гнилицкий Владимир Васильевич — кандидат технических наук.

Инсаров Вильям Викторович — академик РАЕН, заместитель начальника отделения, доктор технических наук, профессор.

Чернявский Алексей Станиславович — инженер.

его координат. В данной статье подробно рассматриваются постановка задачи и особенности практической реализации подхода, базирующегося на указанных выше положениях, главным образом в части алгоритмизации процедуры анализа и принятия решения в задаче селекции заданного объекта наземной сцены.

Сцена состоит из N объектов, из которых M наиболее характерных (функционально и геометрически) используется для построения её эталонного описания на основе указанной выше априорной информации. В качестве основных, наиболее стабильных, признаков изображений объектов рассматриваются их геометрические характеристики — форма и размеры объектов, параметры их взаимного расположения. Эталонное описание сцены содержит только рёберные (скелетные) изображения объектов. Яркость объектов не входит в эталонное описание объекта (трёхмерную модель), поскольку при его создании достоверная информация о яркости и текстуре, как правило, не доступна. Кроме того, форма и расположение граней объектов меньше всего варьируются при изменении условий съёмки, в то время как яркостная составляющая ТИ может сильно меняться в зависимости от освещения, времени суток и времен года. Для каждого объекта, включённого в эталонное описание, область поиска на текущем изображении считается заданной. Предлагается иерархический алгоритм сопоставления ЭИ и ТИ, реализующий процедуры автоматического обнаружения, селекцию заданного объекта и определение его координат. Этот алгоритм включает в себя следующие этапы, описанные в соответствующих разделах данной статьи.

1. Предварительная обработка ТИ, позволяющая устранить или снизить интенсивность шумов различного происхождения.
2. Построение поля градиентов яркости ТИ и нормалей к ЭИ. Нахождение набора наиболее вероятных положений каждого объекта сцены на основе анализа поля градиентов.
3. Сборка сцены, принятие решения, оценка значений координат объекта.

1. ОПИСАНИЕ АЛГОРИТМОВ ПРЕДОБРАБОТКИ ТЕКУЩЕГО ИЗОБРАЖЕНИЯ

Пример текущего изображения и соответствующего ему контурного эталона приводится на рис. 1. Для точной локализации контурного эталона на текущем изображении сцены далее вводится специальная функция сходства. Отметим, что шумы, присутствующие на реальных изображениях, усложняют и без того нелегкую

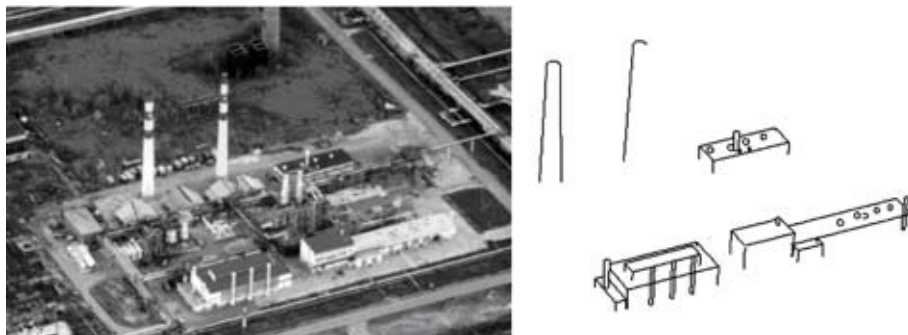


Рис. 1. Пример текущего изображения и контурного эталона

проблему сравнения изображений столь разной природы (бинарное и «текстурированное»). В данном исследовании предполагается, что текущие изображения могут быть зашумлены двумя видами шумов: нормально распределённым шумом и импульсным шумом вида «соль-перец».

При наличии нормально распределённого шума к каждому пикселу исходного (незашумленного) изображения $I(x, y)$ добавляется случайная величина $N(0, \sigma)$ с математическим ожиданием 0 и дисперсией σ , характеризующей амплитуду шума. В данной работе шум применялся по следующему правилу:

$$\tilde{I}(x, y) = \max\left(0, \min\left(I(x, y) + N(0, \sigma), 255\right)\right),$$

т. е. после добавления шума все значения яркости зашумленного изображения, выходящие за рамки $[0, 255]$, обрезались. Классическим методом устранения шума считается свёртка изображения с фильтром, ядро которого представляет собой двумерную функцию Гаусса. Такой метод, обладая бесспорным преимуществом в скорости, имеет, тем не менее, важный недостаток. Поскольку метод не учитывает информации об изображении, лежащем под ядром фильтра, в результате его применения происходит размытие границ. Это особенно нежелательно, поскольку алгоритм обнаружения контурного эталона на изображении построен именно на совпадении границ. В связи с этим, а также по итогам проведённых экспериментов, которые показали неэффективность гауссовской фильтрации зашумленных изображений, было решено применить другой метод фильтрации — двусторонний (билатеральный) фильтр [Tomasi, Manduchi, 1998]. В отличие от гауссовской фильтрации, в которой вклад каждого пиксела в итоговую яркость в центральной точке пропорционален его расстоянию от неё, ядро двустороннего фильтра зависит не только от расстояния, но и от разности яркостей между центральной точкой и пикселями, принадлежащими её окрестности. Яркость в т. (x_0, y_0) считается по формуле

$$I_f(x_0, y_0) = \frac{1}{W} \sum_{(x, y) \in S} G_{\sigma_s}(r(x, y)) G_{\sigma_I}(|I(x_0, y_0) - I(x, y)|) I(x, y), \quad (1)$$

где

$$r(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2};$$

$$G_{\sigma_s} = \frac{1}{\sqrt{2\pi\sigma_s}} \exp\left\{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma_s^2}\right\};$$

$$G_{\sigma_I} = \frac{1}{\sqrt{2\pi\sigma_I}} \exp\left\{-\frac{|I(x_0, y_0) - I(x, y)|^2}{2\sigma_I^2}\right\};$$

W — коэффициент нормализации; сумма считается по круглому окну S заданного радиуса; параметры σ_s и σ_I задают степень сглаживания в пространственной области и в области значений.

Использован двусторонний фильтр, входящий в стандартную библиотеку алгоритмов компьютерного зрения OpenCV. На рис. 2 (см. с. 95) приводятся результаты применения этого фильтра для устранения нормально распределённого шума. На рис. 2а, в изображён текущий снимок с шумом $\sigma = 25$ и $\sigma = 50$ соответственно.

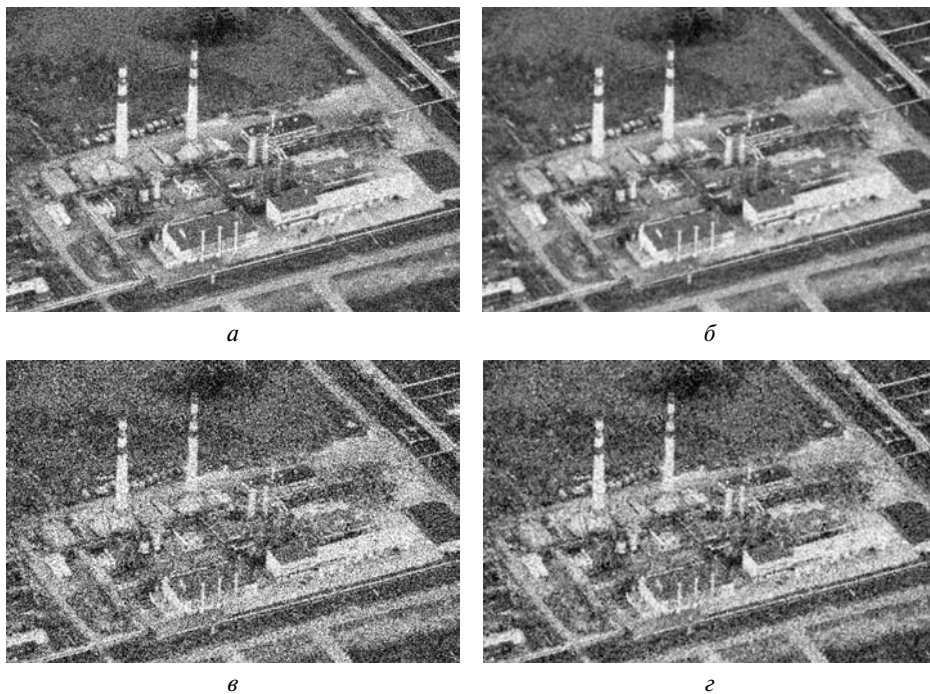


Рис. 2. Применение двустороннего фильтра для удаления нормально распределённого шума

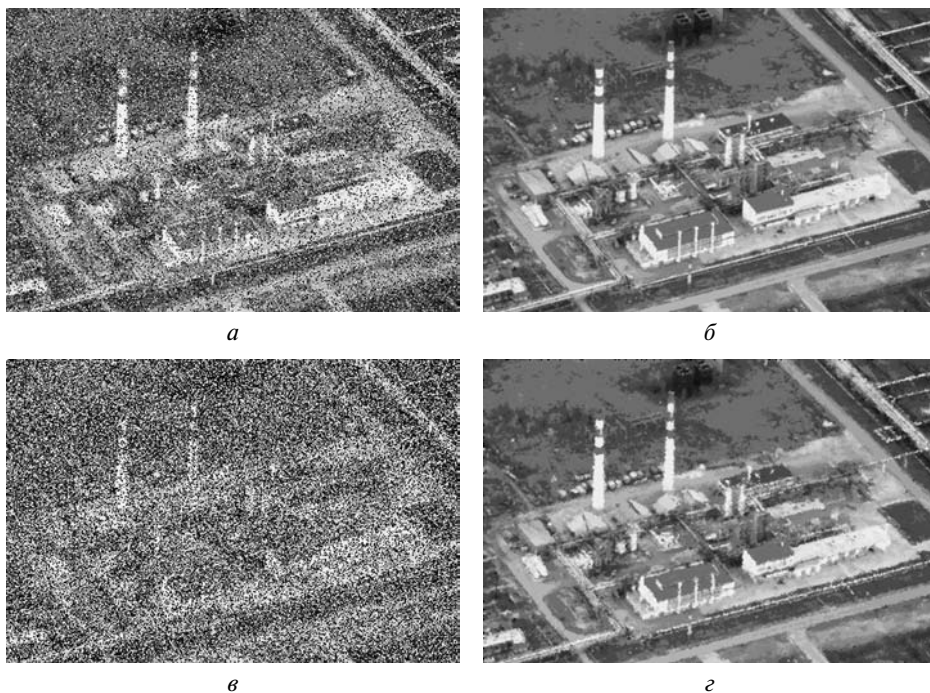


Рис. 3. Результат применения адаптивной медианной фильтрации

На рис. 2б, г показаны результаты применения двустороннего фильтра с параметрами $\sigma_s = 5$, $\sigma_f = 80$. Из данных рис. 2 следует, что двусторонний фильтр хорошо справляется с гауссовским шумом со значением дисперсии вплоть до $\sigma = 25$, что составляет 10 % динамического диапазона изображения. При более высоком уровне шума на фильтрованном изображении возникают пятна.

Импульсный шум, именуемый в англоязычной литературе как шум типа «соль-перец», возникает при некорректном функционировании (отказе) тех или иных элементов матрицы фото- или видеокамеры. Уровень шума регулируется вероятностью, т. е. числом, принадлежащим отрезку $[0, 1]$. Процесс добавления импульсного шума к изображению состоит в том, что значение яркости каждой точки изображения с вероятностью $P_S = P_a + P_b \leq 1$ заменяется случайным значением шума. Яркость изображения с вероятностью P_a заменяется значением a , с вероятностью P_b — значением b , а с вероятностью $(1 - P_S)$ остается неизменной. Как правило, $a = 0$, $b = 255$, т. е. величина яркости заменяется максимальным или минимальным значением динамического диапазона изображения.

На рис. 3 (см. с. 95) приводятся результаты применения адаптивной медианной фильтрации для удаления шума типа «соль-перец» высокой интенсивности. На рис. 3а, в изображён текущий снимок с импульсным шумом при вероятностях 0,2 и 0,5 соответственно; на рис. 3б, г — после применения медианного фильтра с адаптивным окном.

2. ФУНКЦИЯ СХОДСТВА ИЗОБРАЖЕНИЙ, ОСНОВАННАЯ НА АНАЛИЗЕ НАПРАВЛЕНИЙ ГРАДИЕНТОВ ЯРКОСТИ

В данном разделе описывается метод, с помощью которого ведётся подсчёт функции сходства контурных эталонов и областей текущего изображения. Для нахождения наиболее вероятного положения объекта на текущем снимке вычисляется функция сходства эталона и ТИ.

Алгоритм нахождения функции сходства основан на поиске минимума разности направлений градиентов яркости на текущем изображении и эталоне. Метод хорошо зарекомендовал себя в задаче регистрации медицинских изображений [Haber, Modersitzki, 2006]. В настоящей работе используется модификация, в которой в качестве расстояния между эталоном и исходным изображением в каждом пикселе считается величина угла α между направлением градиента на исходном снимке и нормалью к контуру на контурном эталоне (рис. 4). Поскольку нормаль к контуру можно рассчитать двумя способами (и получить внутреннюю или внешнюю нормаль), то из двух возможных значений угла (α и $\pi - \alpha$) берётся наименьшее.

Процедура поиска функции сходства методом разности направлений контуров выглядит следующим образом. В начале к изображению контурного эталона применяется фильтр Гаусса для того чтобы, «размазав»

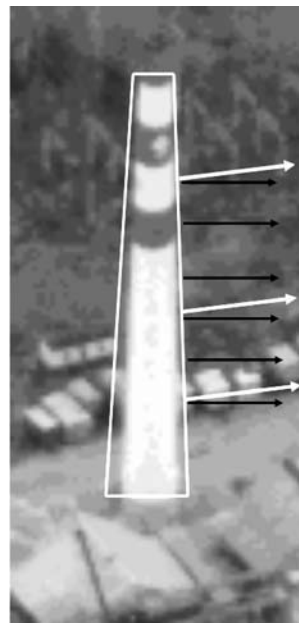


Рис. 4. Подсчёт угла между нормалью к контурному эталону и градиентом яркости текущего изображения

эталон, снизить эффект малой площади эталона (предполагается, что контуры, из которых состоит эталон, имеют ширину в один пиксел) и вычислять меру сходства эталона и фрагмента ТИ по большей части текущего изображения. Затем в каждой ненулевой точке сглаженного эталона находится направление градиента яркости:

$$R_E(i, j) = \arctg\left(E_y/E_x\right),$$

где $E(i, j)$ — изображение контурного эталона.

Та же процедура проводится и для ТИ, в результате чего получаем двумерный массив значений $R_I(i, j)$. Строго говоря, для вычисления направления градиентов на текущем изображении и нормалей к контурам на эталоне вместо арктангенса используется функция $\text{atan2}\left(\left|E_y\right|, \left|E_x\right|\right)$:

$$\text{atan2}(y, x) = \begin{cases} \varphi \text{sgn}(y), & x > 0, \\ \frac{\pi}{2} \text{sgn}(y), & x = 0, \\ (\pi - \varphi) \text{sgn}(y), & x < 0, \end{cases} \quad (2)$$

где $\varphi \in [0, \pi/2)$ — такой угол, что $\text{tg}(\varphi) = |y/x|$. Поскольку арктангенс берётся от модулей значений производных, то углы принадлежат отрезку $[0, \pi/2)$. Для облегчения вычислений значения углов квантуются в N_A ячеек. В данной работе используется шесть ячеек: $[0, \pi/6)$, $[\pi/6, \pi/3)$, $[\pi/3, \pi/2)$, $[\pi/2, \pi/3)$, $[2\pi/3, 5\pi/6)$ и $[5\pi/6, \pi/2)$.

Для нахождения функции сходства между эталоном и ТИ верхний левый угол двумерного массива $R_E(i, j)$, содержащего поле направлений градиента эталона, переносится в позицию (m, n) в системе координат ТИ и накладывается на массив $R_I(i, j)$, содержащий поле направлений градиентов ТИ. В каждой точке, принадлежащей контурному эталону, вычисляется величина

$$p(i, j) = \begin{cases} 1, & \text{если } \left| \text{bin}(R_I(i, j)) - \text{bin}(R_E(i, j)) \right| < T, \\ 0, & \text{в противном случае,} \end{cases} \quad (3)$$

где $\text{bin}(R_I(i, j))$ — номер ячейки, в которой лежит направление градиента текущего изображения в точке (i, j) , а $T > 1$ — значение порога на разность ячеек. Итоговая функция сходства равна сумме функций сходства, рассчитанных во всех пикселах эталона:

$$D(m, n) = \sum_{\substack{i \in [m, m+H_E] \\ j \in [n, n+W_E]}} p(i, j), \quad (4)$$

где H_E и W_E — высота и ширина эталонного изображения. Отметим, что мера сходства (4) называется также расстоянием Хэмминга. Здесь используется значение порога $T = 2$, так как оно обеспечивает наилучшее обнаружение объектов. При пороге, равном 1, также достигаются хорошие результаты. Использование порога, превышающего 2, снижает разрешающую способность поиска, и поиск осуществляется нестабильно.

При использовании метода направлений градиентов яркости функция сходства считается во всех точках, где сглаженное изображение эталонного контура не равно нулю. Из рассмотрения исключаются лишь те точки, в которых амплитуда

градиента мала, что соответствует абсолютно «плоским» участкам изображения. Следует отметить, что порог для формирования маски пикселей, входящих в выражение (4), зависит от диапазона яркости. Так, для ИК-изображений порог должен быть ниже в силу их более низкой контрастности по сравнению с телевизионными изображениями.

Заканчивая описание алгоритма поиска объектов, основанного на сравнении текущего изображения и контуров на эталоне, обращаем внимание на следующий важный факт. Поскольку функция сходства состоит из суммы расстояний, рассчитанных в каждой точке контурного эталона, и пространственное расположение пикселей внутри эталона не учитывается, функция сходства иногда имеет ложные локальные максимумы. Эталон дает хороший отклик на тех участках ТИ, в которых присутствуют контуры схожего с эталоном направления. В частности, для различных прямых, длинных перекладин, краев панелей функция сходства имеет ряд протяженных линейно «гребней», где сосредоточены локальные максимумы.

3. ПОСТАНОВКА ЗАДАЧИ СБОРКИ СЦЕНЫ И СЕЛЕКЦИИ ОБЪЕКТА

Поиск основного объекта проводится следующим образом. Контурный эталон (КЭ) переносится в положение $k(i, j)$ и накладывается на текущее изображение. Пусть сглаженный контурный эталон состоит из N ненулевых пикселей. Рассмотрим те пиксели, в которых направление нормали к КЭ и градиента яркости ТИ совпадают в смысле (3) (попадают в одну ячейку). Иными словами, рассмотрим множество $\Omega = \{(i, j) | \rho(i, j) = 1\}$. Пусть мощность множества Ω равна N^+ : $|\Omega| = N^+$, $N^+ \in [0, M]$. Функция сходства КЭ и ТИ вычисляется путём анализа каждого пикселя, принадлежащего множеству Ω . Будем рассматривать процесс вычисления функции сходства между КЭ и ТИ как проведение N независимых испытаний (по числу пикселей сглаженного контурного эталона) по схеме Бернулли. Введем следующие обозначения: W — изображение объекта, $P_{об} = P\{(i, j) \in (\Omega \cap W)\}$, т. е. $P_{об}$ — вероятность совпадения направлений градиентов в месте действительного расположения на ТИ изображения основного объекта; $P_{фон} = P\{(i, j) \in (\Omega \cap \neg W)\}$, т. е. $P_{фон}$ — вероятность совпадения направлений градиентов во всех других пикселях зоны поиска основного объекта, т. е. на «фоне» (шум, посторонние объекты).

Задачу локализации основного объекта на ТИ можно рассматривать как задачу его селекции. Согласно работе [Горелик, Скрипкин, 1989], решение в задаче селекции, в отличие от задач распознавания, принимается на основе всей совокупности информации по критерию максимума отношения правдоподобия для всех возможных гипотез. В контексте исследуемой в данной статье задачи для всех пикселей области возможного расположения/поиска изображения основного объекта на ТИ вычисляется отношение правдоподобия и решение принимается в пользу того пикселя, в котором обеспечивается максимум отношения правдоподобия. В каждом положении КЭ выдвигаются две гипотезы. Основная гипотеза $H_0: \theta = \theta_0$ состоит в том, что в данном положении находится основной объект. Альтернативная гипотеза $H_1: \theta = \theta_1$ гласит, что в данном положении основной объект отсутствует, а ненулевые значения $\rho(i, j)$ вызваны наличием шума или посторонних объектов. В каждом k -м положении КЭ составляется отношение правдоподобия, позволяющее оценить степень предпочтительности той или иной гипотезы. Функция плотности вероятности для указанных выше гипотез равна

$$f(x_k | \theta_0) = P_{об}^{N_k^+} (1 - P_{об})^{N - N_k^+},$$

$$f(x_k | \theta_1) = P_{фон}^{N_k^+} (1 - P_{фон})^{N - N_k^+}.$$

Отношение правдоподобия записывается как:

$$L_k = \frac{f(x_k | \theta_0)}{f(x_k | \theta_1)} = \left[\frac{P_{об}}{P_{фон}} \right]^{N_k^+} \left(\frac{1 - P_{об}}{1 - P_{фон}} \right)^{N - N_k^+}. \quad (5)$$

Прологарифмировав формулу (5), получаем

$$\log L_k = N \left[\frac{N_k^+}{N} \log \frac{P_{об}}{P_{фон}} + \left(1 - \frac{N_k^+}{N} \right) \log \left(\frac{1 - P_{об}}{1 - P_{фон}} \right) \right]. \quad (6)$$

В рамках метода отношения правдоподобия решение в пользу той или иной гипотезы принимается в результате сравнения величины L_k или $\log L_k$ с некоторым заданным наперед порогом. В данной задаче величину порога определить заранее невозможно, поэтому для всех пикселей области возможного расположения изображения основного объекта на ТИ вычисляется отношение правдоподобия и решение принимается в пользу того пикселя (расположения центральной точки), в котором обеспечивается максимум отношения правдоподобия.

Для практического использования соотношения (6) необходимо знание значений $P_{об}$ и $P_{фон}$, которые, вообще говоря, не известны. Покажем, что при любом фиксированном (неизвестном) значении $P_{фон}$ максимум логарифма отношения

правдоподобия (6) достигается при $\hat{P}_{об} = \frac{N_k^+}{N}$:

$$\frac{\partial \log L_k}{\partial P_{об}} = N \left[\frac{N_k^+}{N} \frac{1}{P_{об}} - \left(1 - \frac{N_k^+}{N} \right) \frac{1}{1 - P_{об}} \right] = N \left[\frac{N_k^+ / N - P_{об}}{P_{об}(1 - P_{об})} \right], \quad \frac{\partial \log L_k}{\partial P_{об}} = 0$$

$$\text{при } \hat{P}_{об} = \frac{N_k^+}{N}.$$

При такой оценке формула (6) принимает вид

$$\log L_k = N \left[\frac{N_k^+}{N} \log \frac{N_k^+}{P_{фон} N} + \left(1 - \frac{N_k^+}{N} \right) \log \left(\frac{1 - N_k^+ / N}{1 - P_{фон}} \right) \right]. \quad (7)$$

Величину $P_{фон}$ можно оценить следующим образом:

$$\hat{P}_{фон} = \frac{\bar{N}_k^+}{N},$$

где \bar{N}_k^+ — среднее число пикселей, в которых направления нормали к КЭ и градиентов яркости ТИ совпадают, для всех «фоновых» пикселей зоны поиска основного объекта. Величину \bar{N}_k^+ с достаточной для практических приложений точностью можно оценить, к примеру, методом скользящего среднего по мере подсчёта N_k^+ во всех точках зоны поиска изображения основного объекта. Применив разложение

Маклорена для функции $\log(1+x)$ и взяв в обоих слагаемых формулы (7) первый (линейный) член разложения, получим квадратичную аппроксимацию формулы (7):

$$\log L_k \approx \frac{N}{\hat{P}_{\text{фон}}(1 - \hat{P}_{\text{фон}})} \left(\frac{N_k^+}{N} - \hat{P}_{\text{фон}} \right)^2. \quad (8)$$

В формулах (7) и (8) полагаем, что $\log L_k = 0$ при $N_k^+ / N \leq \hat{P}_{\text{фон}}$.

Функция сходства иногда может достигать максимума в тех точках области поиска, которые не соответствуют реальному положению искомого объекта. Это может быть связано с наличием протяженных периодических структур либо с сильным зашумлением, в том числе с неточной информацией о положении камеры и об углах обзора. В таких случаях необходимо каким-либо образом использовать дополнительную информацию для повышения вероятности обнаружения. Одним из способов повышения устойчивости процедуры поиска является использование информации о взаимном расположении объектов сцены. Эта идея положена в основу структурно-лингвистического алгоритма, описанного в [Инсаров, 2004]. В данной статье этот алгоритм используется для того, чтобы для одного из объектов, названного основным, агрегировать значения функции сходства текущего изображения с каждым заданным контурным эталоном в одну *совокупную функцию сходства*, учитывая взаимные расстояния между объектами. Предполагается, что после выполнения процедуры сборки совокупная функция сходства основного объекта, составленная из функций сходства соседних с ним объектов, будет достигать своего максимума именно в той точке, в которой основной объект находится в действительности.

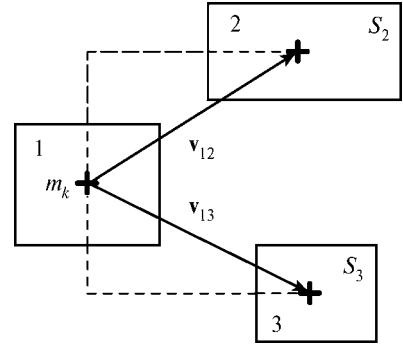
Для каждого объекта сцены, включённого в её эталонное описание, рассчитываются зоны поиска его изображения на ТИ. Выбор указанных зон зависит от координат расчётной точки визирования сцены, статистических характеристик начальных ошибок системы навигации, характеристик датчика изображений и априорных данных о самой сцене. В качестве расстояний между изображениями объектов используется набор векторов $\{\mathbf{v}_{ij}\}_{i,j=1}^N$ между центрами областей поиска всех N объектов сцены. Помимо взаимных расстояний между объектами, также предполагается заданной область поиска каждого объекта относительно основного. Здесь под «относительной областью поиска» понимается тот факт, что для расчёта совокупной функции сходства искомого объекта в процессе сборки учитываются значения функций сходства соседних с ним объектов сцены, причём их функция сходства анализируется в заданной области. В данной статье размеры области поиска каждого объекта относительно основного равны ± 10 пикс. по обеим координатам.

В статье [Инсаров, 2004] на первом этапе процесса сборки велся поиск K наибольших локальных максимумов функции сходства $D_{\text{очн}}(m_k)$ для основного объекта. Затем для каждого максимума $m_k = (i_k, j_k)_{k=1}^K$ определялась область $S_n(m_k)$, в которой анализировалась функция сходства остальных объектов (рис. 5, см. с. 101):

$$i_n \in [i_k + v_{kn}^x - \Delta x, i_k + v_{kn}^x + \Delta x],$$

$$j_n \in [j_k + v_{kn}^y - \Delta y, j_k + v_{kn}^y + \Delta y],$$

Рис. 5. Процесс сборки. Объект № 1 считается основным. Анализируются подобласти поиска соседних объектов ($n = 2, 3$)



где v_{kn}^x и v_{kn}^y — x - и y -компоненты вектора, соединяющего центры ограничивающих прямоугольников k -го и n -го объектов (см. рис. 5). Далее в этой области находился максимум функции сходства для n -го объекта D_n . Его значение сравнивалось со значением функции сходства основного объекта $D_{\text{осн}}(m_k)$, и наибольшее из этих двух значений сохранялось в $D_{\text{осн}}(m_k)$. В результате вычислялась обновленная функция сходства основного объекта:

$$D_{\text{осн}}(m_k) := \max(D_{\text{осн}}(m_k), \max_{\mathbf{r} \in S_n(m_k)} D_n(\mathbf{r})).$$

Описанная выше процедура проводится для всех объектов, кроме основного. В итоге среди всех значений $D_{\text{осн}}(m_k)$ выбиралось максимальное, и делался вывод о том, что основной объект находится в точке, в которой агрегированная функция сходства достигает глобального максимума. Тем самым, учитывалась не только функция сходства основного объекта, но и принимались во внимание «дополнительные голоса» в пользу основного объекта, предоставленные соседними объектами в виде их собственных функций сходства.

В данной работе были сохранены многие аспекты метода сборки, изложенного в [Исаров, 2004], с некоторыми существенными дополнениями. Одним из наиболее «уязвимых» этапов метода сборки является самый первый — отбор локальных максимумов функции сходства основного объекта для дальнейшего анализа. Дело в том, что в случае загороживания или при наличии схожих (но посторонних) объектов локальные максимумы функции сходства могут находиться не в тех местах, где находится объект. В свою очередь, функция сходства на самом объекте может быть меньше, чем K наибольших отобранных значений. В связи с этим в данной статье в процессе сборки рассматриваются все значения функции отклика основного объекта, а не только локальные максимумы. С учётом вышесказанного, процесс сборки состоит из трёх шагов.

Шаг 1. Для каждого объекта эталонного описания сцены в каждом пикселе его зоны поиска определяется N_k^+ и \bar{N}_k^+ .

Шаг 2. Сборка всей сцены идет от основного объекта (см. рис. 5). Обозначим N_1, N_2, \dots, N_m — число контурных элементов в эталонах первого (основного), второго и M -го объектов соответственно. Оценка вероятности совпадения направлений градиентов в месте действительного расположения на ТИ изображения основного объекта для рассматриваемой гипотезы определяется формулой

$$\hat{P}_{\text{об}} = \frac{N_1^+ + \sum_{i=2}^M \max N_i^+}{\sum_{i=1}^M N_i}, \quad (9)$$

где N_m^+ – число пикселей ТИ $m = \overline{1, M}$, в которых направление градиента яркости совпадает с направлением нормали к эталону m -го объекта для данной гипотезы, а максимум берётся по областям допуска $(\Delta x_i, \Delta y_i)$ на взаимное расположение всех объектов относительно основного объекта. Логарифм отношения правдоподобия для рассматриваемой гипотезы имеет вид

$$\log l_k = \sum_{i=1}^M N_i \left[\hat{P}_{об} \log \frac{\hat{P}_{об}}{\hat{P}_{фон}} + (1 - \hat{P}_{об}) \log \left(\frac{1 - \hat{P}_{об}}{1 - \hat{P}_{фон}} \right) \right], \quad (10)$$

где оценка $\hat{P}_{об}$ находится по формуле (9), а $\hat{P}_{фон} = \sum_{i=1}^M \bar{N}_i^+ / \sum_{i=1}^M N_i$ – оценка $P_{фон}$.

На этом этапе выполнения алгоритма логарифм отношения правдоподобия вычисляется для каждого пикселя области поиска основного объекта.

Шаг 3. Принятие решения, оценка координат точки прицеливания. Находится максимум логарифма отношения правдоподобия. Точка прицеливания находится в тех координатах, в которых достигается указанный максимум.

После того, как найден максимум совокупной функции сходства, необходимо оценить, соответствует ли его положение реальному положению искомого объекта на ТИ. Самым лучшим способом сделать это является визуальное наблюдение. В данной статье реальное положение объектов на сцене примерно известно, поэтому для оценки точности обнаружения используется расстояние от найденного центра объекта (вернее, от центра ограничивающего его прямоугольника) до реального центра объекта. Расстояние делится на диаметр (длину диагонали) ограничивающего прямоугольника объекта. Все величины считаются в пикселях.

В реальной ситуации, когда нет возможности привлечь человека для оценки качества обнаружения и когда точное расположение объекта не известно, предлагается использовать численный показатель, равный отношению амплитуд глобального максимума функции сходства к следующему по величине локальному максимуму:

$$Q_{приц} = \frac{\max_{S_{приц}} D_{приц}}{\max_{S_{приц} \setminus R} (D_{приц})}, \quad (11)$$

где $S_{приц}$ – область поиска основного объекта; $R = \{(i, j) | (i - i_{\max})^2 + (j - j_{\max})^2 \leq \varepsilon^2\}$ – точки, принадлежащие ε – окрестности глобального максимума, находящегося в точке с координатами (i_{\max}, j_{\max}) и $\varepsilon > 0$ – радиус окна. В данной работе используется значение радиуса, равное 10 пикс. Если объект найден уверенно, то все локальные максимумы функции сходства будут существенно меньше, чем глобальный максимум, и величина $Q_{приц}$ будет велика. Напротив, эта величина будет мала (снизу она ограничена единицей), если глобальный максимум мало отличается от локальных. Такая ситуация возникает при наличии шума, или если на ТИ присутствует несколько структур, похожих на искомый объект.

4. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

В качестве исходного материала для исследования алгоритма поиска контурных эталонов использовались три изображения трёх групп объектов (моделей). Информация об изображениях и сценах приводится в табл. 1 (см. с. 104). Фрагменты

текущих изображений, соответствующие областям поиска, указанным в табл. 1, эталонные изображения объектов и изображения сцен, собранные из ЭИ объектов, представлены на рис. 6–8 (масштаб ЭИ и ТИ отличается).

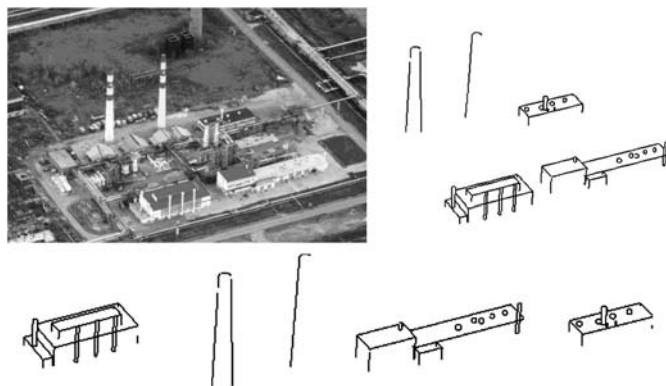


Рис. 6. Изображение № 1, модель А, состоящая из четырёх объектов

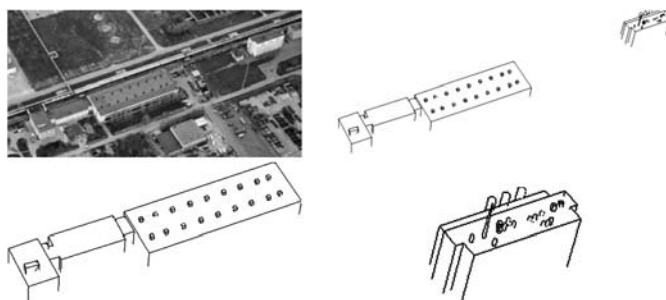


Рис. 7. Изображение № 2, модель В, состоящая из двух объектов

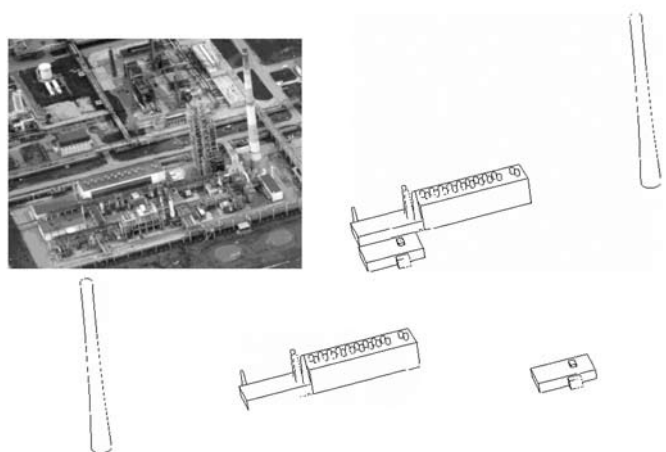


Рис. 8. Изображение № 3, модель С, состоящая из трёх объектов

Таблица 1

Описание использованных ТИ и эталонов

№ ТИ	Размер области поиска [пикс.]	Количество объектов в эталоне
1	507×333	4
2	1067×333	2
3	533×417	3

Таблица 2

Результаты поиска КЭ на ТИ № 1

№ п/п	dx	dy	100*dx/W	100*dy/H	Q
Без шума					
1	2,0	4,0	1,85	6,45	3,34
2	3,0	4,0	3,23	3,08	3,09
3	1,0	3,0	0,63	4,62	4,64
4	1,0	1,0	1,27	2,44	3,53
Импульсный P = 0,5					
1	2,0	4,0	1,85	6,45	2,29
2	3,0	4,0	3,23	3,08	2,47
3	1,0	3,0	0,63	4,62	3,28
4	1,0	1,0	1,27	2,44	3,13
Импульсный P = 0,7					
1	1,0	4,0	0,93	6,45	2,03
2	1,0	6,0	1,08	4,62	1,99
3	1,0	3,0	0,63	4,62	2,57
4	1,0	1,0	1,27	2,44	2,35
Нормальный σ = 10					
1	1,0	4,0	0,93	6,45	3,01
2	3,0	6,0	3,23	4,62	2,87
3	1,0	3,0	0,63	4,62	4,30
4	1,0	1,0	1,27	2,44	3,24
Нормальный σ = 25					
1	1,0	4,0	0,93	6,45	2,98
2	3,0	6,0	3,23	4,62	2,61
3	1,0	2,0	0,63	3,08	3,49
4	1,0	1,0	1,27	2,44	2,74

Таблица 3

Результаты поиска КЭ на ТИ № 2

№ п/п	dx	dy	100*dx/W	100*dy/H	Q
Без шума					
1	5,0	8,0	0,93	3,14	2,00
2	0,0	11,0	0,00	10,48	2,44
Импульсный P = 0,5					
1	5,0	8,0	0,93	3,14	1,57
2	0,0	15,0	0,00	14,29	1,99
Импульсный P = 0,7					
1	6,0	6,0	1,12	2,35	1,17
2	0,0	12,0	0,00	11,43	1,72
Нормальный σ = 10					
1	5,0	9,0	0,93	3,53	1,94
2	0,0	15,0	0,00	14,29	2,24
Нормальный σ = 25					
1	6,0	7,0	1,12	2,75	2,42
2	0,0	12,0	0,00	11,43	2,50

Таблица 4

Результаты поиска КЭ на ТИ № 3

№ п/п	dx	dy	100*dx/W	100*dy/H	Q
Без шума					
1	1,0	2,0	2,27	0,83	1,29
2	168,0	33,0	67,20	28,95	1,38
3	2,0	1,0	2,15	1,82	1,28
Импульсный P = 0,5					
1	1,0	2,0	2,27	0,83	1,34
2	168,0	25,0	67,20	21,93	1,32
3	2,0	0,0	2,15	0,00	1,21
Импульсный P = 0,7					
1	2,0	3,0	4,55	1,24	1,49
2	167,0	24,0	66,80	21,05	1,08
3	226,0	271,0	243,01	492,73	1,04
Нормальный σ = 10					
1	1,0	1,0	2,27	0,41	1,11
2	168,0	33,0	67,20	28,95	1,19
3	2,0	1,0	2,15	1,82	1,24
Нормальный σ = 25					
1	2,0	1,0	4,55	0,41	1,14
2	168,0	33,0	67,20	28,95	1,30
3	0,0	1,0	0,00	1,82	1,50

Алгоритм поиска контурных эталонов был протестирован на трёх изображениях сцен. На каждом изображении было проведено семь испытаний — без шума, с импульсным шумом с $P = 0,5$ и $0,7$, а также с нормально распределённым шумом с $\sigma = 10$ и 25 . Результаты работы алгоритма поиска контурных эталонов представлены в табл. 2–4 (см. с. 104). Столбцы таблиц содержат следующие данные: номер объекта, считающегося основным, абсолютное значение сдвига найденного центра объекта относительно истинного центра по осям абсцисс и ординат (в пикселах), относительное значение сдвига в процентах от ширины и высоты ограничивающего прямоугольника объекта и, наконец, величина Q , являющаяся отношением амплитуд глобального максимума к следующему по величине локальному максимуму. Полу жирным шрифтом в табл. 4 выделены значения, соответствующие неудовлетворительному обнаружению объектов (относительное значение сдвига центра объекта превосходит 25 % от высоты или ширины ограничивающего прямоугольника). Заметим, что величина Q для каждого объекта падает по мере увеличения зашумления. Это связано с тем, что границы объектов размываются, и изображения объектов становятся менее похожими на их эталонные изображения.

На рис. 9 изображены графики функции сходства для некоторых интересных случаев. На рис. 9а, б приведена функция сходства, рассчитанная по формуле

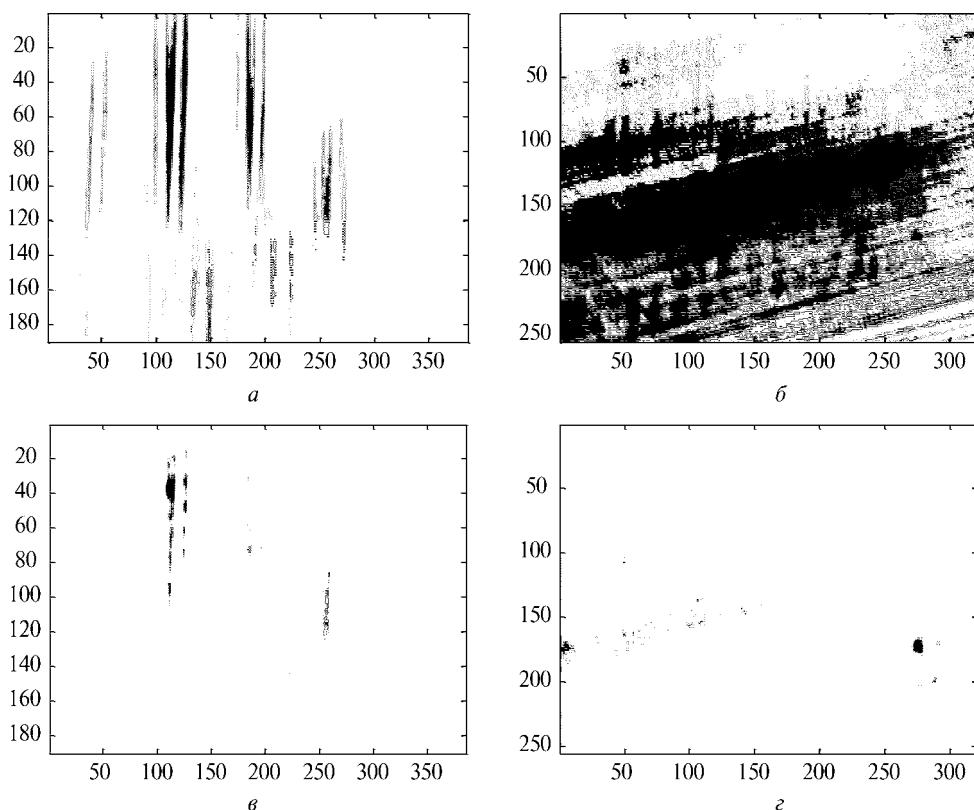


Рис. 9. Функция сходства для объектов № 2 и 3 на ТИ № 1: а, б — функция сходства по формуле (4); в, г — агрегированная функция сходства с учётом расположения основного объекта относительно остальных по формуле (10)

(4) для объектов № 2 и № 3 (трубы и длинное здание — рис. 10) на ТИ № 1. На рис. 9в, г показана агрегированная функция сходства для указанных объектов, вычисленная по формуле (10) во время сборки сцены. Из рисунков видно, что функция сходства для объекта «трубы» имеет ряд вертикальных областей, в которых сосредоточены локальные экстремумы сходства. Это связано с наличием на ТИ № 1 ряда вертикальных элементов — стен зданий, других труб меньшего размера. Обилие ложных максимумов функции сходства для объекта «трубы» коррелирует с величиной Q из табл. 2 — в большинстве случаев величина Q для объекта «трубы» является минимальной среди всех объектов, изображенных ТИ № 1. Линейчатый характер графика функции сходства для объекта «длинное здание» из рис. 9в отражает наличие на ТИ № 1 большого числа линий, сонаправленных с контурами здания (подъездные пути, трубопровод).

На рис. 10–12 приводятся изображения трёх сцен с наложенными на них контурными эталонами объектов. Контурные эталоны наложены в тех местах, которые соответствуют найденным координатам пиков агрегированной функции сходства для каждого объекта. При сборке ряд ложных максимумов функции сходства исчезает, поскольку области, в которых объект действительно располагается, получают «больше голосов». В то же время сборка сцены не всегда позволяет точно локализовать объекты. К примеру, на рис. 12 представлены результаты поиска объектов по ТИ № 3, из которых видно, что объект № 2 найден неверно.



Рис. 10. Контурные эталоны объектов, найденных на ТИ № 1



Рис. 11. Контурные эталоны объектов, найденных на ТИ № 2

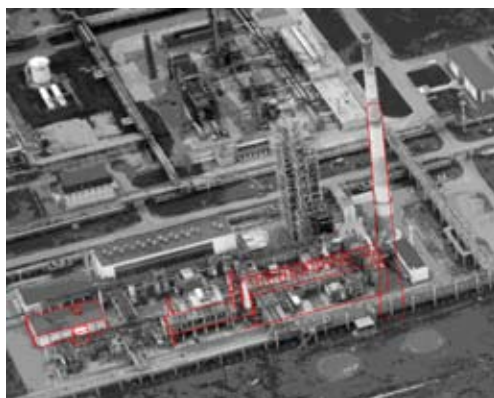


Рис. 12. Контурные эталоны объектов, найденных на ТИ № 3

Заключение

В данной работе был представлен метод сравнения контурных эталонных изображений трёхмерных объектов с текущими изображениями сцен. Для нахождения наиболее вероятного месторасположения того или иного объекта был предложен метод подсчёта функции сходства, основанный на разности направлений градиентов яркости текущего изображения и нормалей к контурам эталонного изображения. Использование контурных эталонов повышает устойчивость поиска объектов. Метод поиска может быть быстро реализован в бортовых вычислителях. Предложенный алгоритм реализован в процедуре селекции объекта (сборке), в которой учитывается априорная информация о взаимном расположении объектов сцены.

Представленный алгоритм позволяет проводить быстрый поиск по эталону, который может формироваться прямо во время облёта сцены, при условии, что трёхмерная модель известна заранее, наряду с приблизительными навигационными данными. Такая концепция привносит большую гибкость по сравнению с использованием статического дискретного набора эталонных изображений. В дальнейшем планируется изучить поведение алгоритма на снимках, полученных в инфракрасном диапазоне, а также более полно изучить проблему атмосферного зашумления и загромождения сцены.

Литература

- Гонсалес Р., Вудс Р. (2005) Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.
- Горелик А. Л., Скрипкин В. А. (1989) Методы распознавания. 3-е изд., перераб. и доп. М.: Высш. шк., 1989. 232 с.
- Инсаров В. В. (2004) Структурно-лингвистический алгоритм обработки изображений и распознавания образов наземных сцен в системе наведения летательного аппарата // Изв. РАН. ТИСУ. 2004. № 1. С. 145–154.
- Haber E., Modersitzki J. (2006) Intensity gradient based registration and fusion of multi-modal images // Intern. Conf. Medical Image Computing and Computer-Assisted Intervention. 2006. V. 9. P. 726–733.
- Tomasi C., Manduchi R. (1998) Bilateral filtering for gray and color images // Proc. IEEE Intern. Conf. Computer Vision. 1998. P. 839–846.

AN ALGORITHM OF ANALYSIS AND DECISION MAKING FOR OBJECT SELECTION ON IMAGES OF GROUND SCENES

Yu. B. Blokhinov, V. V. Gnilitskiy, V. V. Insarov, A. S. Chernyavskiy

State Research Institute of Aviation Systems (GosNIAS)

E-mail: achern@gosniias.ru

One approach to the problem of object detection in images is to compare the current image (CI), with a predetermined reference image of the object (RI). This article discusses the problem of comparing the CI obtained from an aircraft and a RI of a ground scene. The reference image is defined as a set of contours, without considering the brightness of objects. We propose an algorithm for comparing CI and RI which is based on the analysis of the brightness gradient direction field of CI and

the normals to the contours of RI. We also consider a hierarchical image registration procedure which takes into account prior information about the relative positions of objects in a scene. This procedure allows one to solve the problem of localization of objects in the scene and assemble the scene as a whole. The algorithm is applied to a set of photographs of ground scenes containing from two to four objects. We also study the accuracy of object detection depending on the noise level.

Keywords: object selection, template-based search, image matching, contour model, gradient direction field.

Blokhinov Yury Borisovich — Head of the Laboratory, Candidate of Physical and Mathematical Sciences, e-mail: achern@gosniias.ru.

Gnilitskiy Vladimir Vasilevich — Candidate of Technical Science.

Insarov William Viktorovich — Academician of the Russian Academy of Natural Sciences, Deputy Head of Division, Doctor of Technical Science, Professor.

Chernyavskiy Alexey Stanislavovich — Engineer.

ФРАГМЕНТЫ ПРИКЛАДНОЙ ТЕОРИИ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Д. В. Васильев

*Открытое акционерное общество Научно-производственная корпорация
«Системы прецизионного приборостроения» (ОАО «НПК «СПП»), 111250,
Москва, Авиамоторная, 53*

Дан обзор итогов исследований 1972–2009 гг., выполнявшихся под руководством автора в целях разработки и серийного производства следящих высокоточных систем технического зрения с беспойсковым корреляционным сравнением текущих и эталонных видеосигналов. Выделены центральные результаты и методологические приёмы, облегчающие проектирование систем для прецизионного оперативного ориентирования. В необходимой теории задачи корреляционного сравнения сигналов разделены на прямые и обратные. В первых из него извлекают необратимое «пороговое» решение в пользу одного из дискретного множества сигналов, используя корреляции как *меры сходства* претендента с образцом. Решения задач второго вида обеспечивают непрерывное измерение сдвигов между сигналами близкой формы. Поэтому в дополнение к корреляциям с чётной симметрией относительно нулевого сдвига определён класс «ортокорреляционных» функций (ОКФ) с нечётной симметрией, пригодных на роль *меры различия* сигналов из-за одномерного сдвига, имеющего недискретизованное множество значений. Рассмотрена линейная процедура оценивания вектора сдвигов финитных сигналов, порождающая беспойсковые алгоритмы, которые опираются на «взвешивающую» фильтрацию кадровых изображений. Содержание увязано со структурой систем технического зрения (СТЗ), где корреляционные дискриминаторы сдвигов измеряют ошибки слежения.

Ключевые слова: поле ориентиров, проективные сдвиги, обобщенная корреляция, взвешенная ортокорреляция, беспойсковое корреляционное слежение, автозахват, субпиксельная точность ориентирования.

Введение

Представлены сжато материалы обзорного доклада, сделанного автором на семинаре ИКИ в Тарусе 17 марта 2010 г. Вместо включённых в текст работы рисунков иллюстрациями к нему были слайды, входившие в презентацию, которая сопровождала обзор и выложена на сайте <http://tz2010.cosmos.ru/>.

Материал рассчитан на специалистов, знакомых с природой, техникой получения и составом видеосигналов, характерных для современных систем технического зрения, обеспечивающих с помощью собственной видеокамеры ориентирование и управление мобильным объектом через его автопилот. В такой системе изображение наблюдаемого в 3D-пространстве участка яркостного поля преобразуется в электрическую форму с помощью входной оптики и фоточувствительного 2D-прибора с устройствами «развертки», аналого-цифровым приёмным трактом и векторным процессором площадной обработки сигнального потока, извлекающим из него информацию, нужную для управления аппаратом, который

обслуживается этой СТЗ. В дальнейшем изложении подразумевается пропорциональная связь между удельным потоком входной энергии, приходящей в очередном кадре ТВ-типа на фотомишень приёмника изображений (ПИ), и его электрическим откликом с соответствующей площадки.

Опыт реализации беспоскоковых площадных корреляционных алгоритмов в следящих СТЗ для летательных аппаратов (ЛА) и модельные результаты ряда работ позволили установить, что для прецизионного автоматического ориентирования системы по ТВ-изображениям далеко не всегда нужно расходовать вычислительные ресурсы на поддержание качества изображения на высшем уровне, поскольку при этом возрастают динамические ошибки слежения из-за запаздываний в процессорах обработки сигналов. Поэтому при выборе параметров ПИ, анализе и проектных расчётах СТЗ повышенной точности целесообразно:

- не усложнять систему, стремясь к «сверхразрешению» регистрируемых изображений, если их субпиксельное качество не требуется в тракте визуализации ориентиров для оператора или автомата, производящего идентификацию наблюдаемых объектов на стадии «целеуказания» либо «перенацеливания»;
- не применять в роли основной затратную процедуру перебора альтернатив и принятия пороговых решений при покадровом корреляционном измерении сдвигов между изображениями.

Уместно заметить, что при анализе корреляционных СТЗ вместо классической «непрерывной» математики нередко опираются только на упрощённую целочисленную. Однако она затрудняет адекватную оценку достижимой субпиксельной точности корреляционного ориентирования и чревата алгоритмическими ошибками. Это обстоятельство учтено в дальнейшем изложении.

Представленный материал акцентирует внимание специалистов, обеспечивающих «зрительным интеллектом» непилотируемые аппараты, на вопросах, которые обусловлены назначением СТЗ и их функциональным сопряжением с комплексами из подсистем с высоким уровнем автоматизма.

1. СТРУКТУРНЫЙ ОБЛИК И ВЕДУЩИЕ ПРИНЦИПЫ ДЕЙСТВИЯ СТЗ ДЛЯ ЛА

1.1. Исходные предпосылки

Система точного корреляционного ориентирования ТВ-типа строится как контур авторегулирования, который обеспечивает в следящем режиме квазинепрерывное векторное измерение многомерных сдвигов «текущего» (наблюдаемого в очередном кадре) видеоизображения (ТИ) относительно «эталонного» (ЭИ). Роль последнего при целеуказании может играть видеоизображение, вводимое в память СТЗ извне, а на последующих стадиях — ЭИ, уже подвергнутое корректировке по результатам измерений в предшествующих кадрах. Таким образом, ведущим принципом действия системы является «распознавание» ЭИ в потоке сменяемых ТИ и выделение путём сравнения сигналов тех различий между ними, которые нужны для точного ориентирования ЛА в наблюдаемом пространстве при движении ЛА. Темп извлечения требуемой для этого информации из входного потока определяется частотой смены кадров в СТЗ и наличием эффективных алгоритмов решения обратных корреляционных задач.

Вид ЭИ, вводимого в оперативную память СТЗ на стадии «целеуказания», должен быть таким, чтобы СТЗ, включённая в «рабочий режим», могла за приемлемое время выполнить автоматический «захват» нужного направления оси видеокамеры, заданного на ЭИ её «точкой прицеливания» (ТП). Способ подготовки ЭИ зависит от назначения ЛА с СТЗ, состава доступной для применения «авионики» на борту ЛА, наличия и емкости каналов связи между ЛА и командным пунктом (КП) управляющего комплекса, от присутствия в нем средств типа GPS, а также от уровня кооперации между разработчиками «интеллектуальных» СТЗ и адекватных средств целеуказания.

Функциональный состав точной бортовой системы технического зрения (БСТЗ) (рис. 1) определяется её ролью в системе автоматического управления (САУ) для ЛА, где при его собственном движении относительно объектов окружающего пространства она служит аналогом зрительного аппарата человека, являясь основным источником информации для его оперативного ориентирования в процессе слежения за элементами наблюдаемого поля.

Проективные сдвиги — главные причины различий между ТИ и ЭИ из-за перемещения ЛА в 3D-пространстве относительно объектов, которые формируют в совокупности яркий сигнал, наблюдаемый в поле зрения ПИ (обрабатываемом «окне» из ТВ-кадра). Если СТЗ служит для оперативного ориентирования ЛА по некоторой точке прицеливания (ТП) на изображении подстилающей поверхности, то угловые размеры рабочего «окна» в системе не превышают единиц градусов. Чаще всего расстояние от ПИ до объектов намного превосходит их высоту, и тогда, пренебрегая трёхмерностью этих объектов, допустимо считать, что число компонент вектора таких сдвигов $m = 6$, которое равно количеству механических степеней свободы ПИ в 3D-пространстве (рис. 2, см. с. 112). Помимо параметров, обусловленных геометрией, к обобщенным сдвигам изображений можно отнести и показатели освещённости, цвета, характеристики мерцаний, наблюдаемости объектов и т. д. Таким образом, математическое содержание совокупности факторов, отслеживаемых в СТЗ «с интеллектом», определяется вектором параметров с числом компонент $m > 6$.

Наиболее динамичные возмущения в следящей СТЗ для ЛА при его движении вызываются «аддитивными» сдвигами s_1 и s_2 , которые смещают изображения в собственных (обычно декартовых) координатах $\{z = (z_1 z_2)^T\}$ фотомишени ПИ.

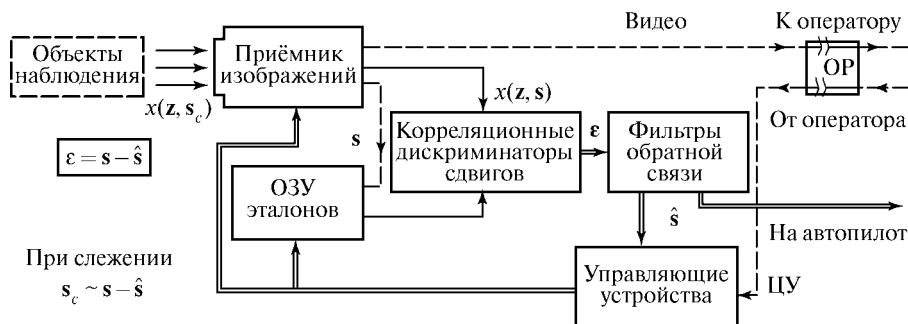


Рис. 1. Система технического зрения как комплекс слежения за многомерными сдвигами изображения (\hat{s} — оценка вектора сдвига, ЦУ — целеуказание, ОР — органы регулирования, ОЗУ — оперативное запоминающее устройство)

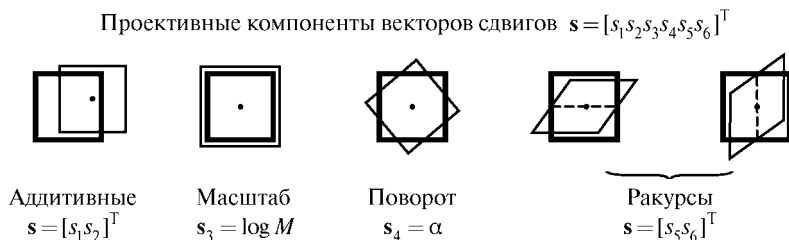


Рис. 2. Виды проективных сдвигов изображений, наблюдаемых СТЗ при движении (α — угол поворота изображения)

Такие сдвиги надо отслеживать в первую очередь и наиболее точно, что удастся лучше всего, если приёмник гиросtabilизирован. Это не освобождает от заботы о парировании также сдвигов изображений по масштабу, «крену» и двум «ракурсным» составляющим s_3, s_4, s_5, s_6 .

1.2. Обобщенное уравнение дискриминатора векторного сдвига изображений в СТЗ

Измерительным звеном в векторном контуре слежения вида рис. 1 служит многомерный корреляционный дискриминатор (МКД), представляющий блок из m одномерных корреляционных дискриминаторов (КД) сдвига. Их алгоритмы, сравнивая в каждом кадре текущее и эталонное изображения (ТИ и ЭИ), выявляют векторную ошибку слежения, т. е. различия между ними по m скалярным параметрам. Здесь и далее принято, что в ТИ и ЭИ исключены средние значения.

Сигналы ТИ и ЭИ также представимы векторами y и x , чья размерность $M \ll m$ и определяется количеством разрешаемых фотоприёмником элементов изображения в обрабатываемом поле зрения. Число скалярных выходов МКД совпадает с числом компонент сдвига в составе s , учитываемых в следящем режиме (например, поворотами ПИ, преобразованиями ЭИ и т. д.) Каждый выход МКД обслуживает свой парциальный контур авторегулирования. Для достижения оптимального качества совместного слежения желательно, чтобы алгоритм измерения соответствующей компоненты ε_i ошибки слежения $\varepsilon = s_r - s_s$ за вектором сдвига между ТИ и ЭИ обеспечивал её нормированную линейную оценку, хотя бы в некоторой окрестности нуля. В связи с этим для синтеза следящей СТЗ удобна процедура наилучшего линейного приближения функций, представляющих ТИ и ЭИ, которые различаются параметрами обобщенного сдвига.

Корреляционный состав этой процедуры обусловлен содержанием обобщенного решения этой задачи в метрическом пространстве. Результат выражает вектор наилучших парциальных оценок m -компонентного сдвига между y и x после прихода очередного кадра ТИ в форме матричного уравнения (1). В нем введена фильтрующая симметрическая $[M \times M]$ весовая (метризирующая) матрица Q , а также выражены отдельными сомножителями ненормированный «производящий» вектор оценок μ и обратная нормирующая матрица B^{-1} :

$$\varepsilon = B^{-1} \mu, \tag{1}$$

где

$$\mu = (dx/ds)^T Q (y - x); \tag{2}$$

dx/ds — производная вектора сигнала по вектору сдвигов — матрица Якоби размерностью $[M \times m]$; m — размерность пространства обобщённых сдвигов; M — число компонент вектора, представляющего изображение или число операций.

Нормирующую матрицу можно представлять в виде, не зависящем от сдвига ГИ:

$$\mathbf{B} = (dx/ds)^T \mathbf{Q} (dx/ds). \quad (3)$$

Легко видеть, что итогом уравнения (2) является выражение вектора, в котором каждый компонент находится скалярным умножением M -мерных векторов с весом, т. е. линейной процедурой «взвешенной» корреляции сомножителей. Содержимое первых скобок в (2) представляет производную вектора сигнала по вектору сдвигов и после транспонирования образует «ленточную» матрицу размера $[m \times M]$. Наиболее важен для практики выбор состава матрицы \mathbf{Q} , которая играет роль фильтрующего линейного оператора для сравниваемых изображений. Он допускает множество способов факторизации, что позволяет находить и сравнивать ряд равноправных версий алгоритмов по аппаратной и вычислительной реализуемости. Возможности выбора представлены формулами в сводке на рис. 3,

1. Постановка и решения вариационной задачи наилучшего приближения функций

$$\varepsilon = \mathbf{B}_x^{-1} \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} \mathbf{n} + r_s^{(k)}.$$

Фундаментальное решение: $\varepsilon = \mathbf{B}_x^{-1} \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} [y - x] = \mathbf{B}_x^{-1} \mu$ при $\mathbf{B}_x = \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} \left(\frac{dx(s)}{ds} \right)$.

Класс ортогональных МКД такой, где $\mu = \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} [y - x] = \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} y$.

2. Коммутативность сомножителей в алгоритмах ненормированных оценок

$$\mu = \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} y = dx^T \mathbf{Q} y = dy^T \mathbf{Q} x = d \left[(L^T x)^T (L^T y) \right] = [d(Ly)^T] Lx = [d(L_1 y)^T] L_2 x = d(Qx)^T y \dots,$$

где d — оператор дифференцирования: $\left(\frac{dx(s)}{ds} \right)^T = d^T x(s) = \left[\frac{d}{ds_1} \frac{d}{ds_2} \dots \frac{d}{ds_m} \right]^T x(s)$.

3. Многозначность решений задачи наилучшего приближения функции x к y .
 Эквивалентность решений по оптимальности. Разбросы множества решений $\xi \in \{\Gamma\}$.

$$\varepsilon^{(1)} = \mathbf{B}_x^{-1} \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} [y - x] = \mathbf{B}_x^{-1} \mu \neq \varepsilon^{(2)} = \mathbf{B}_y^{-1} \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} [y - x] = \mathbf{B}_y^{-1} \mu, \text{ так как } \mathbf{B}_x \neq \mathbf{B}_y,$$

$$\mathbf{B} = \mathbf{A} \mathbf{B}_x + (\mathbf{I} - \mathbf{A}) \mathbf{B}_y, \quad \mathbf{B} = \left(\frac{dx(s)}{ds} \right)^T \mathbf{Q} \left(\frac{dy(s)}{ds} \right).$$

4. Пути структурно-аппаратного синтеза каналов МКД по обобщенному алгоритму
 Парциальная ДХ в МКД по i -му выходу — зависимость i -й компоненты оценки при $s = s_j$.

Рис. 3. Математические средства анализа и синтеза ортокорреляционных алгоритмов многомерных измерительных звеньев для СТЗ в режиме слежения за изображениями: n — векторное представление аддитивной шумовой помехи; $r_s^{(k)}$ — остаточные члены разложения; L, L_1, L_2 — линейные матрицы; ξ — случайная величина; $\{\Gamma\}$ — множество гауссовых распределений; A — весовая матрица, задающая вклады оценок сдвига из дуальных алгоритмов; ДХ — дискриминационная характеристика

где показаны также варианты нормировок, обеспечивающих получение оценок, «эквивалентных по оптимальности».

Нормированный вектор оценок (1) представляет обобщенное решение «обратной» задачи на отыскание измеримых различий по m параметрам между коррелируемыми сигналами-изображениями. Анализ уравнений (1)–(3) с учётом осевых направлений в разложении ТВ-кадра согласно используемому в СТЗ стандарту позволяет факторизовать оператор Q так, чтобы его фильтрующие функции были выполнимы путём раздельной обработки элементов изображения в соответствии с его «разверткой» по строкам, по столбцам или иначе. Другими словами, это облегчает выбор фильтрующих звеньев в парциальных КД.

Состав уравнения (3) напоминает о наличии взаимной корреляции разноименных элементов реального изображения и возможном из-за неё перекрестном влиянии парциальных оценок сдвигов друг на друга. Упрощенный анализ таких связей дан в монографии Ю. М. Астапова, Д. В. Васильева, Ю. И. Заложнева [Астапов и др., 1988], где в силу методической преемственности беспоисковых методов этот вопрос увязан с рассмотрением свойств двумерных автокорреляционных функций (АКФ) с чётной центральной симметрией.

1.3. Свойства чётной симметрии автокорреляционных функций изображений

Эти свойства характерны для двумерных АКФ аддитивного сдвига $\mathbf{K}(\mathbf{s})$ относительно точки $\mathbf{s} = 0$. Они имеют место, если изображение является *финитным* [Астапов и др., 1988; Vassiliev, 2000]. Финитность означает, что $\mathbf{K}(\mathbf{s}) = 0$ за пределами области сдвигов s_1 и s_2 , в которой исходное и сдвинутое изображения «перекрываются». Возможный вид таких АКФ в плановых проекциях показан на рис. 4 (см. с. 115). Асимптотические свойства на примере двух изображений при малых сдвигах видны на рис. 5 (см. с. 116).

Вертикальное сечение тела функции $\mathbf{K}(\mathbf{s})$ плоскостью, совпадающей с осью s_1 , представляет усреднённую по рабочему «окну» АКФ строчных полезных сигналов. Если названное условие финитности не выполняется, возникают отклонения от симметрии, растущие по мере удаления аргумента $\mathbf{K}(\mathbf{s})$ от точки $\mathbf{s} = 0$. Это приводит к «краевым эффектам», которые при больших сдвигах могут влиять на качество работы КД. Их можно учитывать, следуя работе Д. В. Васильева, К. Ф. Григорьева, В. А. Никонова [Васильев и др., 2007б], но потребность в этом редка и здесь не обсуждается.

1.4. Пути обеспечения инвариантности следящей СТЗ для ориентирования

Способность СТЗ осуществлять основную функцию — ориентирование не должна существенно зависеть от структуры наблюдаемых сюжетов. Это качество часто является важнейшим требованием из предъявляемых к «зрительному интеллекту» таких комплексов. Его выполнение достигается выделением инвариантных качеств сигналов-изображений и их корреляционных характеристик при обработке последних фильтрами и процессорами в составе МКД.

Установлено, что корреляционные процедуры в БСТЗ наиболее продуктивны, если они опираются на линейные алгоритмы сигнальных преобразований, обеспечивающие извлечение точных покадровых скалярных оценок, пропорциональных аддитивным компонентам сдвига между ТИ и ЭИ. Пользуясь итогами проделанных работ, можно получать в каналах МКД стабильные по крутизне выходные

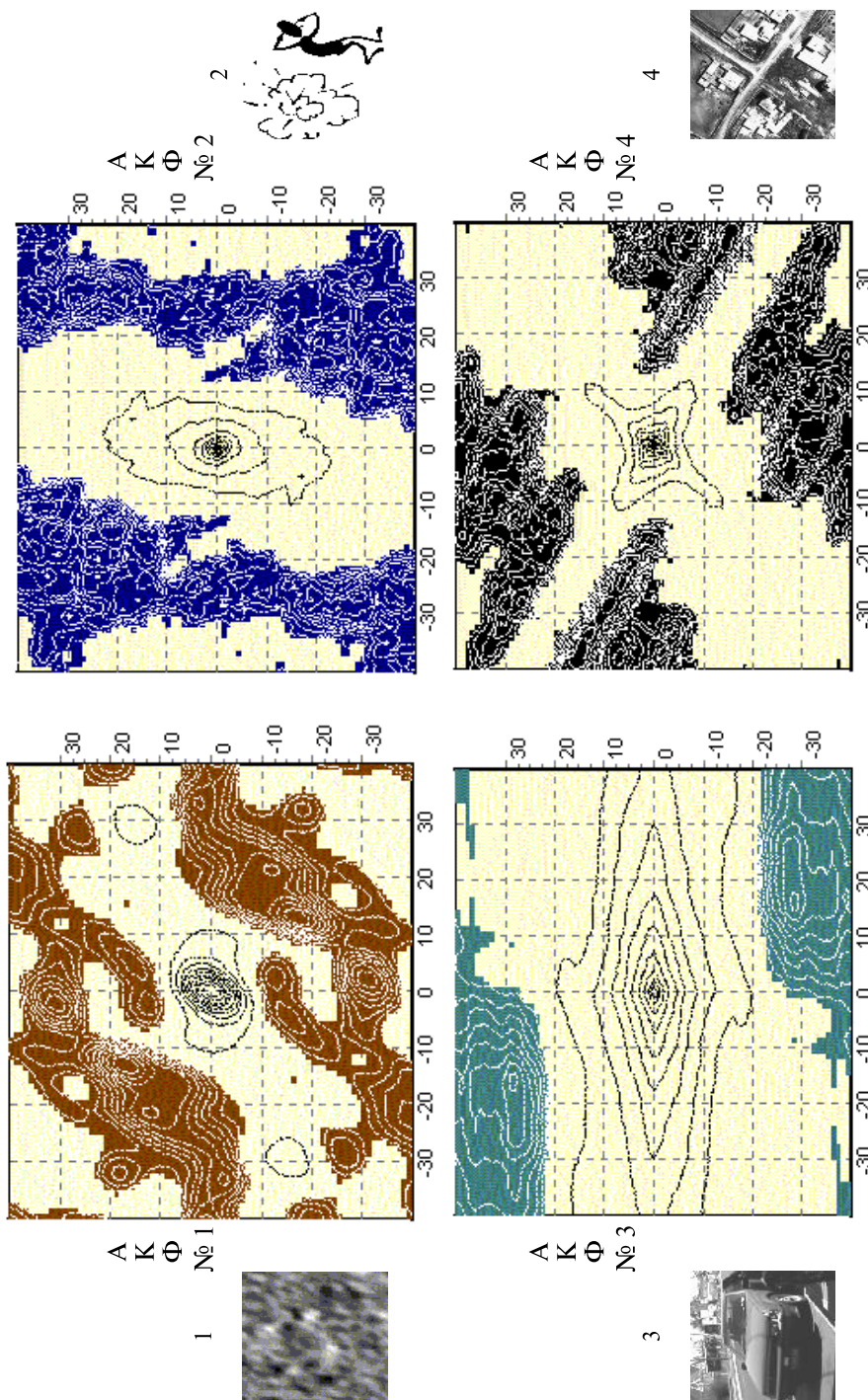


Рис. 4. Плановые проекции АКФ двумерных сдвигов $K(s_1, s_2)$ для финитных изображений формата 100×100 пикс. Области отрицательных значений функции затемнены, её аргументы даны в пикселах

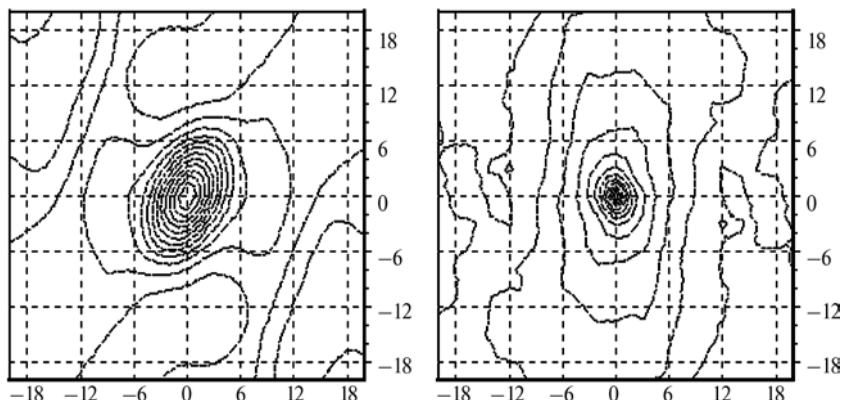


Рис. 5. Характерная эллиптичность сечений двумерных АКФ в области малых сдвигов

дискриминационные характеристики (ДХ), обладающие *нечётной симметрией* относительно нулевого сдвига и апертурой, достаточной для большого запаса устойчивости контура слежения за обслуживаемой им компонентой сдвига ТВ-изображений. Самыми важными качествами беспойскового ортокорреляционного алгоритма каждого из выходов МКД являются быстроедействие, хорошая стабильность нуля ДХ, малая зависимость её формы от структуры изображений и простота управления этой формой. Последние свойства обеспечиваются при проектировании системы подбором операторов фильтрации, участвующих в процессе наилучшего приближения функций путём авторегулирования изображений по итогам покадрового корреляционного сравнения согласно представленной выше процедуре оценивания вектора сдвигов между ними. Исследования и применение алгоритмов парциальных КД сдвига ТВ-изображений стали предметом одного из разделов общей теории [Астапов и др., 1988], расширенной в вузовских курсах лекций автора по СТЗ, работах [Гапон, 2009а и др.].

1.5. Алгоритмы взвешенной ортокорреляции сигналов в парциальном КД

Формальной базой, гарантирующей для КД реализуемость ДХ со «строгим нулем» и нечётной симметрией, могут служить решения обратной корреляционной задачи об оценке сдвига любого одномерного финитного сигнала, который всегда обладает собственной АКФ с чётной симметрией $K(s) = a_0 + a_2s^2 + a_4s^4 + \dots$, а её форму характеризует коэффициент корреляции $\kappa(s) = K(s)/K(0) = K(s)/a_0$ (a_i, b_i, c_i — коэффициенты разложения функций по чётным, нечётным или произвольным составляющим).

Принципы обратных решений корреляционных задач развиваются практически десятки лет. Они апробированы в серийных СТЗ, изучались в диссертациях автора, А. В. Русакова, В. П. Сергиенко и др., не раз излагались в вузовских курсах автора по СТЗ, а недавно получили более общую интерпретацию [Васильев, Гапон, 2009]. Достаточной основой теории обратных решений таких задач служит обращение одномерной АКФ, которое сводится к отысканию сходящегося разложения ОКФ (обратной или ортогональной корреляционной функции) $R(s)$ в ряд по нечётным степеням сдвига $R(s) = b_1s + b_3s^3 + b_5s^5 + \dots$, чью форму выражает коэффициент ортокорреляции $\rho(s) = R(s)/b_1$.

Для окрестности нулевого сдвига последнее выражение описывает идеальные свойства алгоритма беспойскового КД, достижимые в отсутствие мешающих факторов, если взаимная импульсная характеристика входных каналов дискриминатора является нечётной. Эквивалентная схема такого ненормированного КД со стационарными фильтрами (СКД) дана на рис. 6, где x_1, x_2, y_1, y_2 — входные и выходные сигналы фильтров в каналах КД; g_1, g_2 — импульсные отклики фильтров; γ_{12} — взаимная импульсная характеристика (ВИХ) каналов, нечётная симметрия которой и обеспечивает необходимую ортогонализацию в преобразовании взаимной корреляционной функции K_x в K_y . Двусторонние стрелки \leftrightarrow означают обратимое фурье-преобразование, символ \otimes — коммутативную свёртку функций, точка над функцией отмечает её комплексные свойства, а «птичка» над функцией указывает на «реверсирование» её аргумента, т. е. $\bar{g}(t) = g(-t)$.

Необходимое условие физической осуществимости алгоритма беспойскового КД определяется требованием, устанавливающим конечные границы рабочей области ДХ такого дискриминатора (например, апертуру, т. е. «раскрыв» монотонной

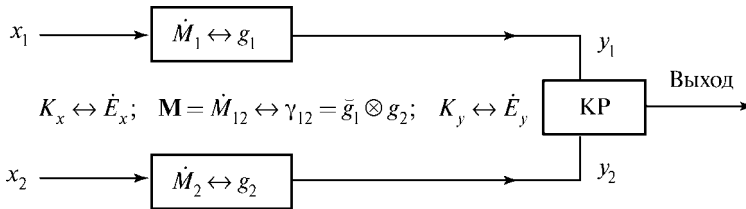


Рис. 6. Эквивалентная схема одномерного КД со стационарными фильтрами: E_x, E_y — спектральные плотности энергий сигналов; M — комплексная частотная характеристика; КР — коррелятор

Примеры 1D-процедур беспойсковой оценки сдвига по значению ОКФ

№ п/п	Ортогонализирующий оператор (отклик) $\gamma_{12}(s) = -\gamma_{12}(-s)$	Ортогонализирующая ЧХ $jM(f)$	Фурье-выход ортокоррелятора $R(s) \leftrightarrow iM(f)E_0(f)$	Наименование процедуры
1	$\otimes \frac{1}{s}$	$\frac{jf}{ f }$	$jE_0(f)$	Преобразование Гильберта
2	$\otimes \delta'(s) \equiv \frac{d}{ds}$	jf	$jfE_0(f)$	Первая производная $K(s)$
3	$\otimes \delta^{(2n-1)}(s) \equiv \left(\frac{d}{ds}\right)^{(2n-1)}$	$(jf)^{(2n-1)}$	$(jf)^{(2n-1)} E_0(f)$	Производные порядка $(2n - 1)$
4	$\otimes \frac{t}{ t } \exp(- t)$	$\frac{jaf}{1+a^2 f^2}$	$\frac{jafE_0(f)}{1+a^2 f^2}$	Аналоговый RC-алгоритм СКД
5	$\otimes C(t)$ и $\otimes S(t)$	$F_c(f)$ и $F_s(f)$	$\sim R(s) \leftrightarrow jM(f)$	Цифровой НКД с ДХ $\sim C \otimes S$

Рис. 7. Алгоритмы ортокорреляционных дискриминаторов (ОКД) сдвига одномерных сигналов: f — частота; δ — дельта-функция; $C(t)$ — базисная функция с чётной симметрией; $S(t)$ — базисная функция с нечётной симметрией; $F_c(f), F_s(f)$ — фурье-образы чётной и нечётной базисных функций

рабочей зоны). Для правильного подбора фильтрующих свойств каналов КД полезно доказанное в работе [Астанов и др., 1988] заключение, что ДХ любого одномерного линейного КД обладает максимальной крутизной при нулевом сдвиге между сигналами.

Алгоритмы ортокорреляционных дискриминаторов сдвига сигналов характеризует таблица примеров на рис. 7 (см. с. 117) с перечнем решений, подсказывающих разные пути реализации КД аддитивных сдвигов в аналоговой и цифровой технике. В него включён пример алгоритма КД с аналоговой фильтрацией [Васильев и др., 1981], апробированного в серийных СТЗ, а также назван класс «нестационарных» корреляционных дискриминаторов (НКД), впервые описанных в диссертации А. В. Русакова в 1975 г. и позже представленных у Д. В. Васильева, С. А. Серебрякова [Vassiliev, Serebjakov, 1997]. Принцип действия НКД иллюстрируют формулы и блок-схемы на рис. 8, 9.

$$x_i(t) = x[z_i(t)]; \quad a_{ik} = (x_i, C_k); \quad b_{ik} = (x_i, S_k);$$

$$x_{cmp}(t) = \sum_{i,k} q_{ik} x_{ik}(t); \quad x_{ik}(t) = a_{ik} C_k(t) + b_{ik} S_k(t); \quad \{C(t)\} = \{C(-t)\}; \quad \{S(t)\} = \{-S(-t)\}.$$

Все БФ нормированы и взаимно ортогональны.

$$\mathbf{x}_{ik} = [a_{ik} \ b_{ik}]^T; \quad \mathbf{x}_{ik}^o = \mathbf{A} \mathbf{x}_{ik} = [b_{ik} \ -a_{ik}]^T; \quad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix};$$

$$x_{cmp}^s(t) = \sum_{i,k} q_{ik} x_{ik}^s(t); \quad x_{ik}^s(t) = a_{ik} C_k(t) + b_{ik}^s S_k(t); \quad \mathbf{x}_{ik}^s = [a_{ik}^s \ b_{ik}^s]^T.$$

Парциальная ВКФ ортогонализированных сигналов:

$$k_{ik}(s) = \langle x_{ik}^{os}, x_{ik}^s \rangle = a_{ik} b_{ik}^s - a_{ik}^s b_{ik}.$$

Усреднённая по кадру ВКФ сдвига сигналов в направлении строк (ДХ КД):

$$K_{12}^q(s) = \sum_{i,k} q_{ik} (a_{ik} b_{ik}^s - a_{ik}^s b_{ik}).$$

Рис. 8. Сводка формул, образующих алгоритмы ортокорреляционных дискриминаторов одномерного сдвига сигналов на основе их нестационарной фильтрации (НКД). Показана ненормированная оценка сдвига кадрового ТВ-сигнала вдоль строк; индексы: i — номера строк; k — номера базисных функций (БФ); ВКФ — взаимнокорреляционная функция

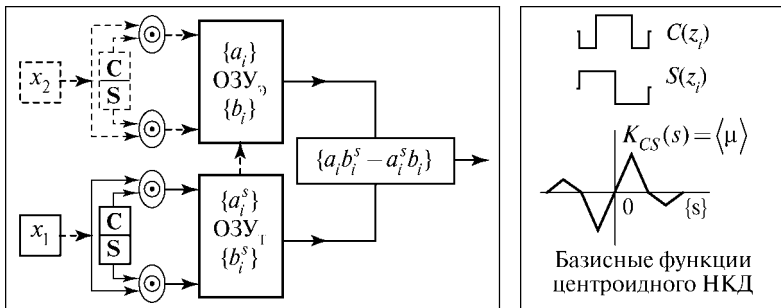


Рис. 9. Алгоритмическая схема НКД в СТЗ с ОЗУ для текущего и эталонного сигналов и пример выбора БФ: ОЗУ_э — ОЗУ эталона; ОЗУ_т — ОЗУ текущего изображения; $C(z_i)$ — чётная часть сигнала на i -й строке; $S(z_i)$ — нечётная часть сигнала на i -й строке

1. **Обобщение теоремы Карсона:** $Lim K_y(s) = K_x \otimes \gamma_{12}$ при $K_x(s) \Rightarrow \delta(x)$.
2. **ВК** \rightarrow индикатор сходства $K(s) = c_0 + c_2s^2 + c_4s^4 + \dots$ — чётная функция сдвига;
 - при: а) $\Pi \gg \Delta f_c \rightarrow \gamma_{12} \cong \delta(s)$ — идеальный коррелятор;
 - б) $\gamma_{12}(-s) = \gamma_{12}(s) \rightarrow \text{Im } \mathbf{M}_{12}(f) = 0$, где $\mathbf{M}_{12} = \mathbf{M}_1 * \mathbf{M}_2$ — реальный коррелятор.
 Если ВК \rightarrow ДС = КР + нелинейный э/п алгоритм оценки сдвига, в/затраты чрезмерны, ДХ неинвариантны к спектру сигнала.

- ВК** \rightarrow беспойсковый КД $\rightarrow R(s) = c_1s + c_3s^3 + c_5s^5 + \dots$ — нечётная функция сдвига;
 - при а) $\gamma_{12}(-s) = -\gamma_{12}(s) \rightarrow \text{Re } \mathbf{M}_{12}(f) = 0$ — реальный коррелятор;
 - б) $\Pi \ll \Delta f_c \rightarrow$ ДС с ДХ \sim инвариантной к спектру сигнала;
 - в) в/затраты малы;
 - г) увеличена устойчивость ОЭСС + область захвата.

Рис. 10. Сводка факторов, поддерживающих инвариантность выходных характеристик ортокорреляционных дискриминаторов сдвига изображений к их структуре: ВК — взаимная корреляция; П — эквивалентная шумовая полоса корреляционного дискриминатора; ДС — дискриминатор сдвигов; э/п — экстремально-поисковый; ОЭСС — оптикоэлектронная следающая система

1.6. Асимптотические характеристики КД

Проектируя измеритель сдвига сигналов, удобно использовать понятие «асимптотической ДХ» (АДХ). Оно опирается на те корреляционные свойства, которыми согласно теореме Карсона [Букингем, 1986] начинает обладать широкополосный случайный сигнал, будучи пропущенным через узкополосную фильтрующую линейную систему. Моделью такого сигнала служит пуассоновский поток независимых коротких импульсов. Очевидно сходство «многопиксельного» кадрового видеосигнала ТВ-типа с длинным отрезком из такого потока. Поэтому, обобщая теорему, можно полагать, что усреднённая на множестве строчных (или столбцовых) сигналов ДХ, которой обладает относительно узкополосный линейный КД аддитивного одномерного сдвига случайных видеоизображений, имеет форму, почти совпадающую с ВИХ его фильтрующих цепей. Это легко увидеть, связав согласно символам на рис. 6 входные и выходные корреляционные характеристики K_x и K_y с ВИХ фильтрующей системы и устремив K_x к δ -функции (рис. 10).

Справедливость достигаемого при этом обобщения теоремы Карсона подтверждается видом ДХ реальных КД в составе изделий, либо их моделей на случайных множествах сюжетов. Пример статистической проверки дискриминатора на модели с почти идеальной формой ДХ дан в статье и диссертации А. В. Гапона (2009а, б).

1.7. Отображения ортокорреляции над полем двумерных сдвигов

Нетрудно увидеть, сравнивая рис. 4 и 11 (см. с. 120), как меняется вид двумерной АКФ, если при выполнении двумерной свёртки изображений одно из них подвергнуть преобразованию Гильберта по направлению оси s_1 . Процедура равносильна пропуску всех фурье-компонент строчных сигналов через фазовращатель $\pi/2$.

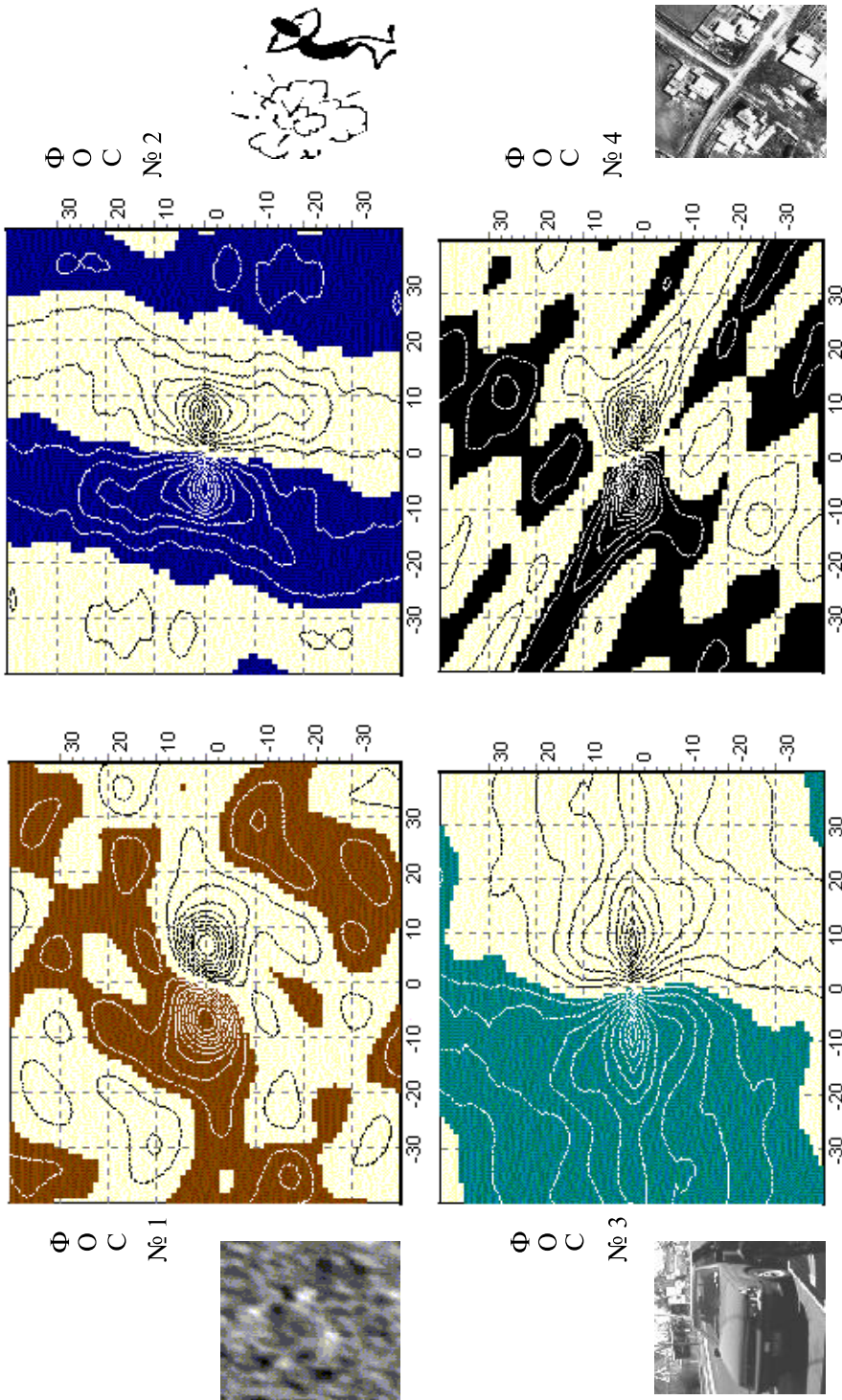


Рис. 11. Двумерные функции оценок сдвига в направлении $\{s_1\}$ для изображений с рис. 4

Результаты вычисления четырёх «функций оценки сдвига» (ФОС), сохраняя центральную симметрию, отличаются тем, что при $s = 0$ их значение АКФ обращается в нуль. Вертикальное сечение АКФ плоскостью, совпадающей с осью s_1 , представляет ОКФ этого сдвига, усреднённую по всем строкам сигнала x . Сместив секущую плоскость параллельно себе, можем увидеть в полученном сечении влияние «чужого» сдвига s_2 на ход ОКФ от s_1 . Представленные на рис. 12 и 13 графики

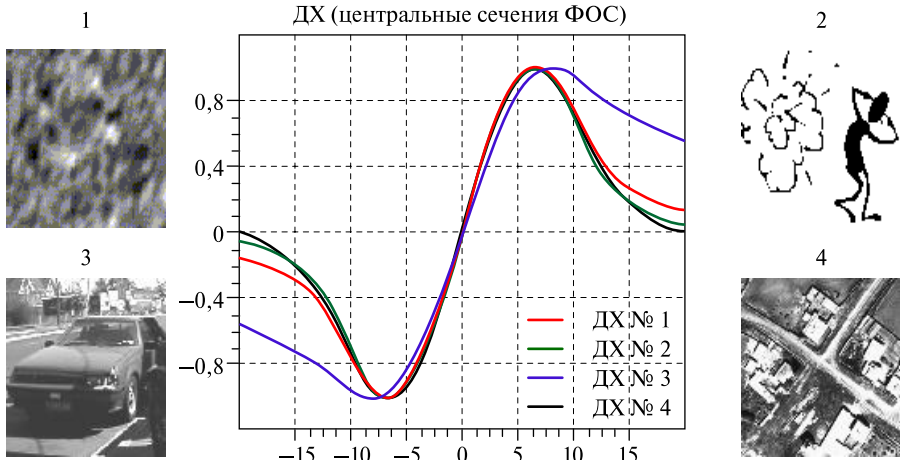


Рис. 12. Семейство ДХ в НКД усреднённого сдвига s_1 при $s_2 = 0$ для сюжетов рис. 4. В алгоритм оценивания вошла нормировка по средней за кадр крутизне ДХ в точке $s_1 = 0$. Инвариантность формы и апертюра ДХ обеспечены выбором фильтрующего оператора γ_{12}

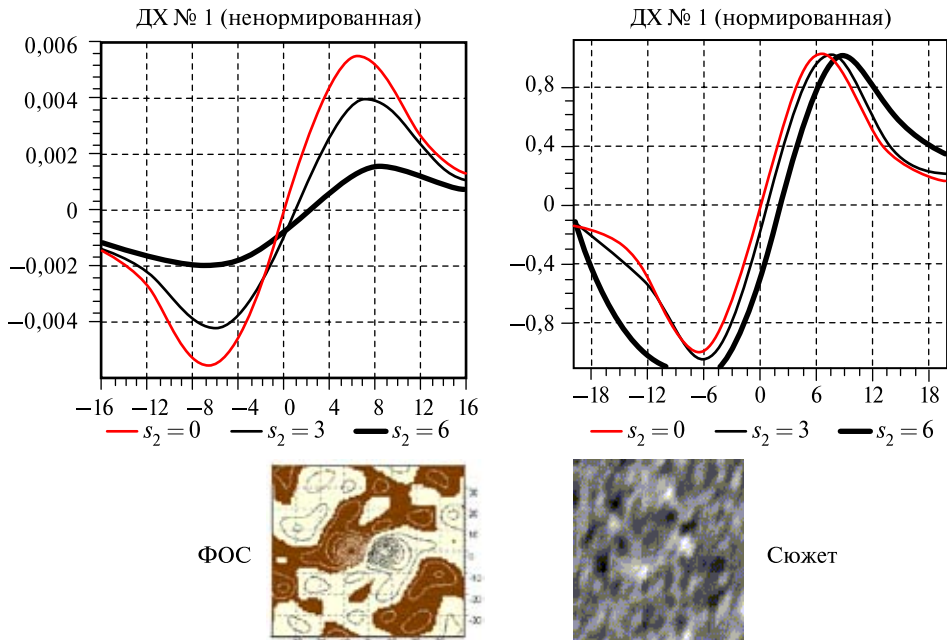


Рис. 13. Примеры ДХ в НКД усреднённого сдвига s_1 при $s_2 \neq 0$ для сюжетов рис. 4 и тех же условиях, что и для рис. 12

отображают часть результатов из статей [Васильев, 2002; Мишин, 2007; Васильев и др., 2007б], где также исследовались пути получения инвариантных ДХ в парциальных КД, способы ослабления перекрестных связей между ними и впервые использовались совмещённые «диаграммы нулей» ДХ по s_1 и s_2 для описания свойств векторных дискриминаторов аддитивного сдвига.

1.8. Обеспечение «автозахвата» направления по эталонному изображению

Прогнозируемость формы ДХ парциальных КД линейного двумерного контура слежения за текущими аддитивными сдвигами изображений любого состава можно считать одним из практических результатов развития теории корреляционных СТЗ для ориентирования ЛА. Именно он позволяет, опираясь на известный арсенал средств и методов точного авторегулирования, строить такие системы с учётом традиционной автоматики носителей и назначения комплекса в целом. Иллюстрирующими приложениями этой теории служат представленные в статьях Д. В. Васильева и др. (2007а, б) и в работах Е. А. Фирсова (2007б) и А. В. Гапона (2009б) итоги проверок алгоритмов, которые обеспечивают автозахват заданного направления (или ТП) на изображении подстилающей поверхности для СТЗ ЛА, либо на участке звёздного поля для СТЗ ориентации космических аппаратов (КА). Рисунки 14 и 15 содержат графические и текстовые комментарии к названным ре-

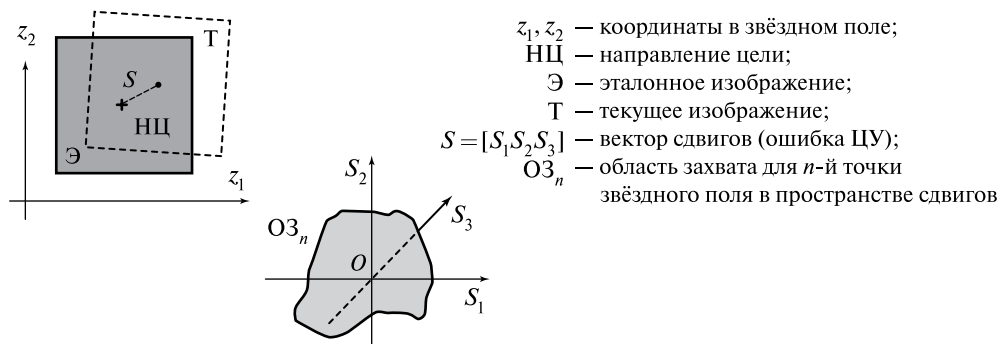


Рис. 14. Понятие области автозахвата направления в звёздном поле

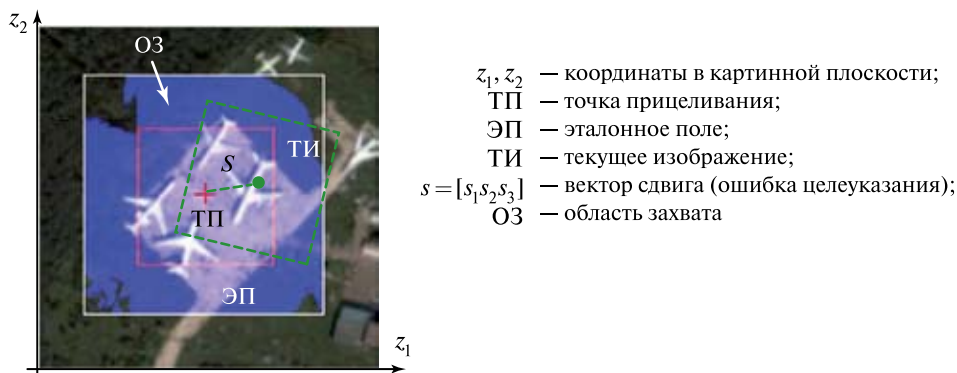


Рис. 15. Область автозахвата направления в точку прицеливания на земной поверхности

зультатам. Основными приёмами, определяющими их достижение, является априорная подготовка ЭИ, чей формат превосходит размеры ТИ, а также подбор таких фильтрующих операторов для алгоритмов КД, которые расширяют области захвата в парциальных контурах слежения.

1.9. Шумоустойчивость КД

Линейная фильтрация строк (или столбцов) ТВ-кадра до передачи на коррелятор, кроме ортогонализации сигналов, подчеркивает в их сплошных спектрах участки, обеспечивающие получение ДХ той формы, которая желательна для автослежения за сдвигами ТИ относительно ЭИ. Поскольку входные спектры таких сигналов широки, а при извлечении из них информации о сдвигах происходит её многоэлементное усреднение по площади рабочего «окна», помехоустойчивость таких алгоритмов весьма высока. Это подтверждено анализом [Астапов и др., 1988], параметрами серийных СТЗ и результатами работ А. В. Гапона (2009а, б). Иллюстрацией к ним служит рис. 16.

Там же представлен способ и формулы расчёта дисперсии выхода КД в окрестности нулевого сдвига из-за аддитивного входного шума при известной амплитудно-частотной характеристике (АЧХ) этого звена. Приведенные в статье А. В. Гапона (2009а) выводы подтверждают выявленную и ранее [Астапов и др., 1988; Русаков, Васильев, 1975] высокую шумоустойчивость КД сдвига ТВ-сигнала, а кроме того, устанавливают связь между параметрами ПИ и достижимыми показателями точности слежения за сдвигами изображения при известном уровне шумов в ТВ-тракте. Опыт разработок СТЗ подтвердил, что расчёты позволяют верно

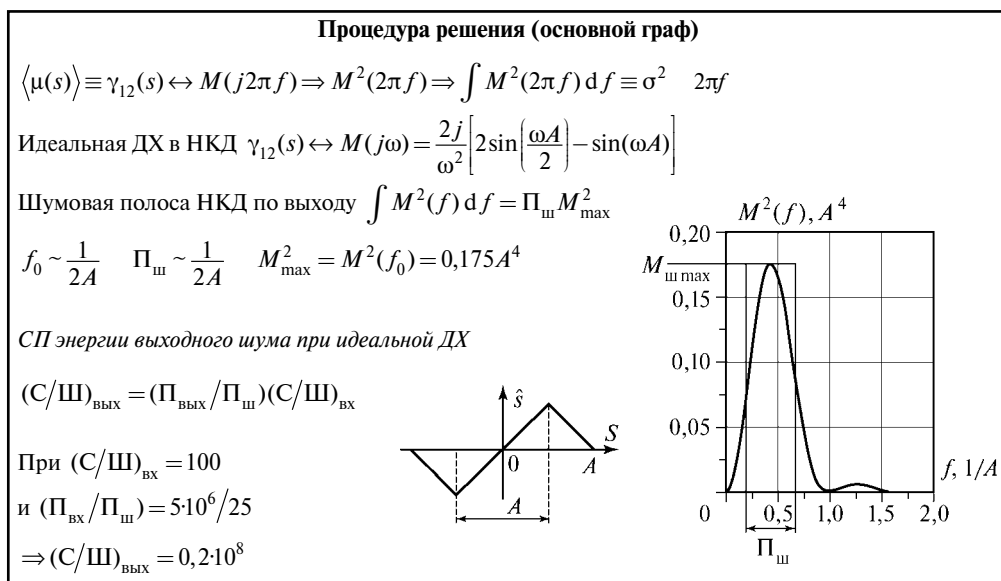


Рис. 16. Решение задачи о шуме в НКД аддитивного сдвига: σ^2 — дисперсия случайной величины; $\Pi_{ш}$ — эквивалентная шумовая полоса корреляционного дискриминатора; СП — спектральная плотность; С/Ш — отношение сигнал/шум; А — апертура дискриминационной характеристики

оценивать реализуемость прецизионного ориентирования ЛА с помощью проектируемой следящей системы из элементов доступного качества.

2. УЛУЧШЕНИЕ СТЗ КАК СЛЕДЯЩЕГО ИЗМЕРИТЕЛЯ ДЛЯ ТОЧНОГО ОРИЕНТИРОВАНИЯ

Стабилизация крутизны ДХ в парциальном КД сдвигов изображения поддерживает постоянство петлевого усиления парциальных следящих контуров в СТЗ, необходимое для достижения их предельного быстродействия при гарантированной устойчивости. Это обеспечивается тем, что алгоритмы КД выполняют, получая ДХ согласно подразд. 1.7 как коэффициент ортокорреляции $\rho(s) = R(s)/b_1$, где s — парциальный аддитивный сдвиг ТИ относительно ЭИ по одному из направлений, определяемых стандартом разложения в ПИ; b_1 — крутизна функции $R(s)$ в точке нулевого сдвига.

Для практики пригодны разные способы отыскания нормирующего множителя $1/b_1$, один из которых опирается на «пробные сдвиги», совершаемые, например, при записи ЭИ в оперативную память. Другие возможности многообразны и частично освещались в работе Ю. М. Астапова, Д. В. Васильева, Ю. И. Заложнева (1988) и вузовских курсах автора. Различия между ними не слишком велики, а выбор зависит от назначения СТЗ и требований к ней.

Электронное слежение за направлением на точку прицеливания в СТЗ позволяет достичь максимального быстродействия, наилучшей точности и устойчивости автосопровождения (АС) при совместном использовании следующих приёмов и средств:

- разделение функций «точного» и «быстрого» слежения соответственно между внешним «электромеханическим» и внутренним «электронным» контурами (ЭМК и ЭК);
- минимизация запаздываний в ПИ с АЦП и в процессорах обработки изображений по алгоритмам парциальных КД для ЭК при жесткой синхронизации с видеокамерой во избежание неконтролируемых фазовых сдвигов в петле регулирования;
- применение «субпиксельной» интерполяционной коррекции передаваемых из ЭК в ЭМК покадровых целочисленных оценок сдвига между ТИ и ЭИ в пикселах.

Следящим элементом в ЭК служит рабочее «окно», которое может «попиксельно» перемещаться в ТВ-растре видеокамеры по окончании кадра к приходу следующего так, чтобы центр «окна» возвращался обратно с точностью до пиксела согласно округленным до целого сигналам ошибок, которые измерил МКД из-за векторного аддитивного сдвига ТИ и ТП в нем. Поскольку ЭК смещает «окно» через программу разложения кадра без обращения к инерционному механическому приводу, такая организация процесса АС позволяет, располагая парциальными КД с конечным раскрытием (апертурой) дискриминационных характеристик (ДХ), расширять диапазон измеряемых сдвигов за пределы апертуры до границ допустимых смещений «окна» внутри кадра. Благодаря этому принципу и высокому быстродействию замкнутых ЭК удастся использовать их как корреляционные измерительные звенья с улучшенными свойствами для АС по каждой из компонент векторного сдвига ТИ относительно ЭИ, парируемых внешним ЭМК, например, за счёт бесступенчатого механического поворота видеокамеры вокруг рабочих осей.

Применение цифровых ЭК в СТЗ ЛА всегда оправдано расширением динамических диапазонов и повышением качества электромеханического управления видеокамерой в составе системы, так как ЭК служит для каждой из парциальных ветвей ЭМК почти идеальным измерителем ошибки слежения. Это звено имеет увеличенный динамический диапазон и может поддерживать субпиксельную точность АС. Она достижима, если каждый из КД в составе «целочисленного» ЭК имеет бесступенчатую линейную ДХ в окрестности истинно нулевого сдвига, а при передаче данных из ЭК в ЭМК к округленному до целого пиксела сдвигу добавляется с нужным знаком дробный субпиксельный «остаток», измеримый посредством подходящего интерполяционного алгоритма в интервале $\pm 0,5$ пикс.

Стабилизация собственных координат СТЗ для ЛА (рис. 17) является специфическим требованием, от выполнения которого может существенно зависеть успех применения системы на борту аппарата при его маневрировании. С этим связана нужда в электромеханических и /или «виртуальных» гиropлатформах и средствах быстрой астроориентации СТЗ. Удовлетворение этих требований требует её сопряжения с устройствами управления и навигации ЛА в едином комплексе. Следящее управление оптической осью СТЗ упрощается при наличии средств гиросtabilизации (ГС), которые сопряжены с навигационными приборами ЛА. Если ГС является трёхосной, а ЛА не должен быстро маневрировать, то слежение за ТП по ТИ удастся поддерживать, обходясь без компенсации сдвигов «по крену», оперативно парируя лишь аддитивные сдвиги и используя специально создаваемый контур для многоступенчатого согласования масштабов ТИ и ЭИ по сигналам, которые накапливаются на выходах парциальных КД масштабных сдвигов. Действующая система такого состава представлена в патенте В. В. Бабушкина и др. (1993) и применяется в технике ЛА.

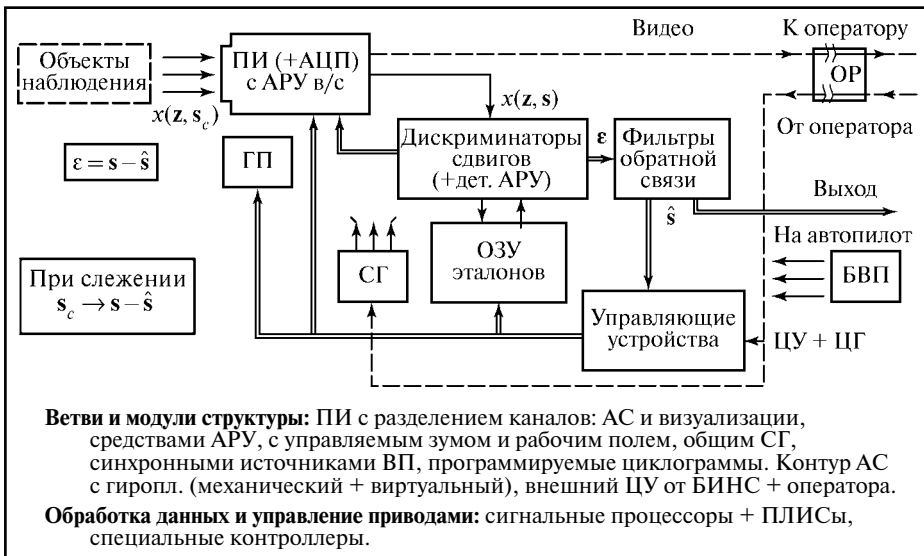


Рис. 17. Функциональная схема СТЗ в комплексе для самонаведения БПЛА: АРУ — автоматическая регулировка усиления; в/с — видеосигналы; ГП — гиropлатформа; СГ — синхрогенератор; БВП — блоки вторичного питания; ЦГ — гиросцентраль; БИНС — бесплатформенная инерциальная навигационная система; ПЛИС — программируемая логическая интегральная схема

Когда трёхосная ГС затруднена, для ориентирования СТЗ «по крену» строится парциальный контур слежения, где измерительным звеном является дискриминатор неаддитивного сдвига-поворота рабочего «окна» относительно центра вращения, которым служит ТП в наблюдаемом ТИ. При этом векторный дискриминатор сдвига, как и процесс корреляционного слежения в СТЗ, становится трёхкомпонентным. При компенсации «кренового сдвига» можно использовать компьютерные преобразования, оптико-механические устройства либо органы управления ориентацией ЛА.

3. ПУТИ СОЗДАНИЯ ДИСКРИМИНАТОРОВ НЕАДДИТИВНОГО СДВИГА ИЗОБРАЖЕНИЙ

Обобщенные АКФ изображений и преобразования сдвигов служат математическими понятиями и приёмами, необходимыми для приложения фундаментального описания многомерной процедуры 1.4. к выбору алгоритмов КД, которые могут служить измерителями сигналов ошибки в парциальных контурах слежения за сдвигами s_3, s_4, s_5, s_6 (см. подразд. 1.3). В левой части рис. 18 представлена геометрия 2D-изображения, подвергнутого «проективному» сдвигу T^S по шести составляющим одновременно. Результат действия такого оператора сводится к групповому преобразованию координат $\{z\}$. В правой части рисунка выписаны преобразования, совершаемые над полем $\{z\}$ для перехода от декартовых координат к полярным, а затем показано, как представлять операцией умножения «масштабно-креновые» смещения, сохраняя симметрию множества сдвигов $\{s_3\}$. Она является условием, которое необходимо, если надо сохранить свойство ортогональности сигнала и его производной по сдвигу. В случае выполнения этого условия упрощается принцип выбора алгоритма КД, измеряющего масштабный сдвиг изображения.

Возможность такого упрощения видна из рассмотрения рис. 19 (см. с. 127), на котором слева показано конформное отображение логарифмической полярной сетки из декартовых координат $\{z\}$ в развернутые цилиндрические координаты $\{\xi\}$, где масштабные и креновые сдвиги s_3 и s_4 представляются как аддитивные Δ_1 и Δ_2 .

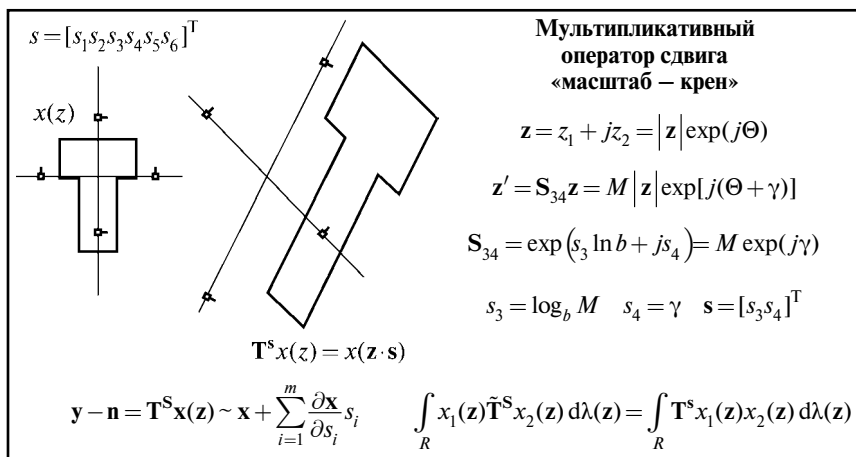


Рис. 18. Неаддитивные сдвиги, их преобразования и обобщенные АКФ изображений: M – масштаб; γ – угол поворота изображения

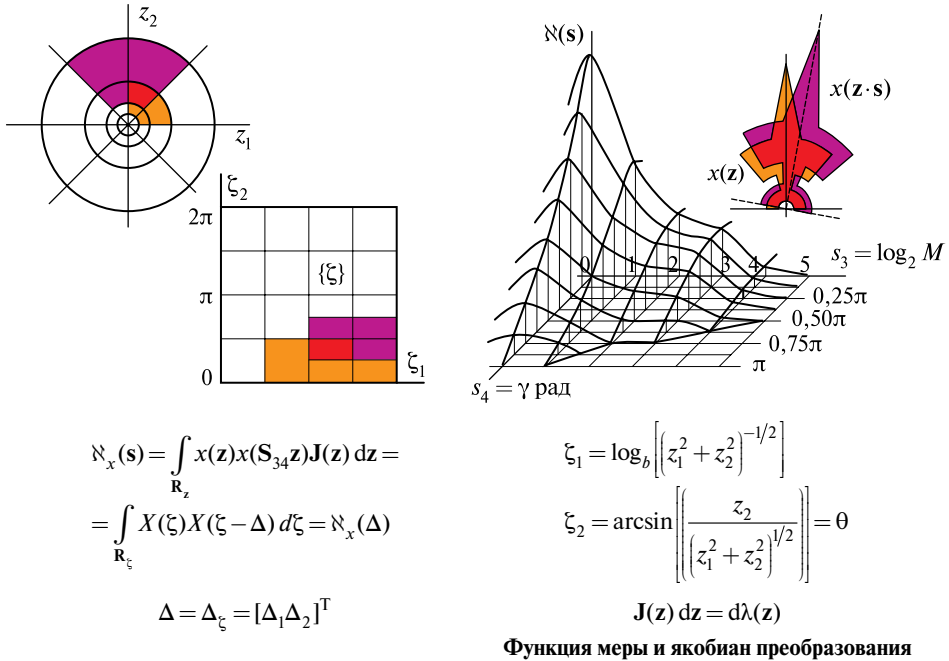


Рис. 19. Приведение сдвигов к аддитивной форме, пример обобщенной АКФ объекта: $\xi = [\xi_1, \xi_2]^T$ — преобразованные (логополярные) координаты сигналов-изображений; N — обобщенная АКФ изображения, зависящая от сдвигов по масштабу и крену; b — основание логарифма для преобразования; $J(z)$ — якобиан преобразования

В правой части того же рисунка построена часть обобщенной АКФ «масштабно-кренового» сдвига для стилизованных ЭИ и ТИ. Нижняя часть рисунка демонстрирует правила записи обобщенной АКФ и её преобразования при смене координат.

Рациональные средства парирования неаддитивного сдвига сигналов в СТЗ для ориентирования ЛА стали в последних десятилетиях XX в. целью ряда изобретений эвристического характера. Среди них более продуктивными явились принципы слежения за масштабными изменениями площадного изображения, происходящими по мере сближения ЛА с некоторой точкой прицеливания в поле зрения СТЗ. Один из удачных принципов представлен в упомянутом патенте [Бабушкин и др., 1996], где для компенсации малых масштабных сдвигов этого типа в оперативном запоминающем устройстве системы производят замену действующего ЭИ очередным ТИ в каждый момент, когда этот сдвиг достигает программно установленной величины в несколько процентов. Описанный метод «перезаписи эталонов» по существу характеризует эффективную систему ступенчатого слежения за масштабом ТИ относительно ТП, указанной на первом эталоне.

Примечательно, что гиросtabilизация видеоизображений в реальных СТЗ для ЛА делает малозаметным эффект неизбежного «сползания» ТП из-за шумов при многократной последовательной замене ЭИ на ТИ, которая должна выполняться при изменениях масштаба в десятки раз. Желаемый результат достижим в серийных СТЗ ценой тщательной отработки алгоритмов, средств автокалибровки, конструктивных решений и качественного контроля изделий при их производстве и в эксплуатации.

Среди альтернативных вариантов выбора алгоритмов дискриминатора масштабных сдвигов отметим такой, который не требует обращения к преобразованию координат поля в ЭИ и ТИ. Он может быть реализован путём выделения из состава изображений тех компонент сигнала, которые обеспечивают получение асимптотической ДХ нужного вида на множестве масштабных сдвигов, представленных в аддитивной форме $\{\Delta\}$, как на рис. 19. Удобством этого решения является возможность построения экономного алгоритма дискриминатора по принципу НКД (см. рис. 8), работающего с коэффициентами разложения видеосигналов по подходящим базисным функциям (БФ). Особенность решения в том, что БФ для масштабного НКД надо подбирать с учётом обратного перехода из пространства $\{\xi\}$ в $\{z\}$ и якобиана преобразования.

4. ОСОБЕННОСТИ СОПРЯЖЕНИЯ БОРТОВЫХ СТЗ И «АВИОНИКИ» ЛА

Эти особенности обусловлены многообразием предусматриваемых режимов применения и технологических требований по пригодности системы для её контроля, в том числе — по показателям точности, быстродействию, применимости в экстремальных условиях и т. д. Названная специфика отражается не только на сложности сопутствующих изделию измерительных средств, но и на составе входящих в СТЗ субблоков и связей между ними. Среди характерных черт такой СТЗ на рис. 17 отмечены роль внешних связей системы и место гироплатформы (ГП), а также единого синхрогенератора (СГ). Качество последнего, например, нередко определяется возможностью переключений режима его работы как в «ведущем», так и в «ведомом» состояниях, поскольку второе бывает нужным при сопряжении двух и более беспилотных СТЗ с авионикой одного ЛА для совместного целеуказания и/или управления ими.

Взаимная синхронизация внутри отдельной СТЗ характерна для высокоточных систем ТВ-типа, где шумы внешнего и теплового происхождения подавляются настолько, что в роли основных помех могут оставаться наводки и пульсации от источников вторичного питания с асинхронными импульсными преобразователями. Поэтому радикальные меры, поддерживающие электромагнитную совместимость блоков малогабаритных прецизионных СТЗ, по необходимости опираются на хорошее взаимное фазирование всех потенциальных источников технических помех.

5. АДАПТИВНЫЙ КОРРЕЛЯЦИОННЫЙ ИЗМЕРИТЕЛЬ СКОРОСТИ (АКИС)

Измеритель скорости для контроля параметров собственного движения локомотива, исследованный автором вместе с С. А. Денисовым, Д. Г. Конотопом, С. А. Серебряковым [Васильев и др., 1995, 1998], представляет простейшую СТЗ с электронным контуром слежения за временным сдвигом s между потоками идентичных сигналов случайной формы. Они поступают от двух фотодетекторов, которые направлены на движущуюся подстилающую поверхность и разделены базой D , расположенной вдоль вектора скорости. Её значение D/s и пройденный путь определяются тем точнее, чем выше качество синхронизации между «таймером» и процессором, где выполняется алгоритм взвешенной ортокорреляции, а сглаженный выход дискриминатора управляет темпом выборок в АЦП входных сигналов. Технология подобных корреляционных измерений может быть полезной при создании автоматических устройств посадки ЛА и КА.

Ходовые испытания макетных образцов АКИС подтвердили эффективность их принципа действия в диапазоне скоростей более двух порядков при погрешности измерения скорости менее 1 %. Они показали также, что для достоверной сертификации таких приборов в промышленном исполнении необходимы метрологические средства сравнительного контроля и организация трассовых испытаний.

Заключение

Общие выводы из представленных работ таковы.

1. Точное ориентирование подвижного летательного аппарата в 3D-пространстве требует применения системы технического зрения непрерывного действия, где методами системы автоматического управления поддерживается многомерное слежение за меняющимся 2D-отображением наблюдаемого поля ориентиров. Математическим объектом слежения в такой системе технического зрения является вектор сдвига между текущим и эталонным изображениями, а в роли последнего могут использоваться сжатые изображения такого состава, который нужен для вхождения в режим АС и его поддержания с приемлемой погрешностью ориентирования по совокупности требуемых навигационных показателей. Физическим результатом служит стабильное удержание оптического центра видеокамеры на точке прицеливания в поле ориентиров. Устойчивость этого состояния должны поддерживать и средства гиросtabilизации.

2. Проверенная рациональная структура системы технического зрения для ориентирования летательного аппарата реализует парциальные контуры быстрого слежения за компонентами вектора сдвигов. По крайней мере, два из них парируют наиболее динамичные аддитивные сдвиги, выделяя сигналы ошибок беспоисковыми дискриминаторами по алгоритмам взвешенной ортокорреляции изображений и управляя точкой прицеливания через электронный контур.

3. Принципы получения бесступенчатых дискриминационных характеристик, подбор и при необходимости — коммутация базисных функций, обеспечивающих инвариантность их формы к составу сюжетов, позволяет не только осуществлять режимы автозахвата, но и достигать субпиксельной точности ориентирования системы технического зрения на точку прицеливания.

4. Совместное применение алгоритмов следящего контроля за несколькими типами сдвига изображений в системы технического зрения для летательного аппарата (в том числе за «креном», масштабом, ракурсами и уровнем сигналов) выполнимо, целесообразно и перспективно.

Автор надеется, что его обзор привлечёт внимание разнопрофильных специалистов, создающих элементы СТЗ «с интеллектом», к тем сторонам теории сигналов, которые важно увязывать между собой в опытно-конструкторской разработке прецизионных комплексов ориентирования, где методы математики, достижения автоматике, оптики, фотоники, радиоэлектроники и навигации приходится объединять практикой.

Литература

Астатов Ю. М., Васильев Д. В., Заложнев Ю. И. (1988) Теория оптико-электронных следящих систем. М.: Физматлит, 1988. 328 с.

- Бабушкин В. В., Буадзе В. Ш., Васильев Д. В.* и др. (1996) Бортовой комплекс корректируемого летательного аппарата с ТВ-головкой самонаведения. Патент РФ № 2058011 с приор. 16.09.93. Опубл. в Бюл. изобретений, 1996. № 10.
- Букингом М.* (1986) Шумы в электронных приборах и системах / Пер. с англ. М.: Мир, 1986. 399 с.
- Васильев Д. В., Муранов Л. Н., Тимонин В. А.* (1981) Дискриминатор. А. С. СССР № 885918. Опубл. в Бюл. изобретений, 1981. № 44.
- Васильев Д. В., Гапон А. В.* (2009) Элементы теории решений обратных корреляционных задач // Электромагнитные волны и электронные системы. М.: Радиотехника, 2009. № 7. С. 30–39.
- Васильев Д. В., Гапон А. В., Сумерин В. В., Фирсов Е. А.* (2007а) Автозахват направления по изображению с использованием скользящей фильтрации // Электромагнитные волны и электронные системы. 2007. Т. 12. № 7. С. 49–52.
- Васильев Д. В., Григорьев К. А., Никонов В. А.* (2007б) Краевые эффекты в корреляционных дискриминаторах сдвига изображений // Электромагнитные волны и электронные системы. 2007. Т. 12. № 9. С. 61–71.
- Васильев Д. В., Денисов С. А., Серебряков С. А.* (1995) Исследование адаптивного корреляционного измерителя скорости с применением математического моделирования // Вестн. МЭИ. 1995. № 2. С. 9–18.
- Васильев Д. В., Конотоп Д. Г., Серебряков С. А.* (1998) Измерительные устройства на основе адаптивного корреляционного измерения скорости // Искусственный интеллект в технических системах: Сб науч. тр. Вып. 19. М.: Гос. ин-т физико-технич. проблем. 1998. С. 123–136.
- Васильев Д. В., Мишин В. В.* (2002) Исследование инвариантного дискриминатора сдвигов изображений // Радиотехнические тетради. М.: Изд. МЭИ, 2002. № 24. С. 61–68.
- Васильев Д. В., Мишин В. В.* (2007) Инвариантные дискриминаторы сдвига изображений // Электромагнитные волны и электронные системы. 2007. Т. 12. № 7. С. 35–43.
- Гапон А. В.* (2009а) Свойства ортокорреляционных дискриминаторов сдвига изображений // Электромагнитные волны и электронные системы. М.: Радиотехника, 2009. № 7. С. 23–29.
- Гапон А. В.* (2009б) Теория и применение инвариантных ортокорреляционных алгоритмов автозахвата направления по изображению в оптико-электронных следящих системах: Автореф. ... дис. канд. техн. наук. М.: МФТИ (ГУ), 2009.
- Русаков А. В., Васильев Д. В.* (1975) Ошибки дискретного вычисления функций автокорреляции случайного поля // Радиоприёмные устройства: Тр. М.: МИРЭА, 1975. Вып. 80.
- Фирсов Е. А.* (2007а) Корреляционный астроориентатор для космических аппаратов // Электромагнитные волны и электронные системы. 2007. Т. 12. № 7. С. 44–48.
- Фирсов Е. А.* (2007б) Развитие и применение корреляционных методов обработки изображений в быстродействующих оптико-электронных следящих системах: Автореф. ... дис. канд. техн. наук. М.: МФТИ (ГУ), 2007.
- Vassiliev D.* (2000) Some Invariant Features of Signals in the Correlation Image Tracking Systems // Proc. on SRIA Conf., BIT, Beijing, Sept. 2000. P. 104–109.
- Vassiliev D. V., Serebrjakov S. A.* (1997) Wavelet Shift Discriminators for Correlation Velocimetry and Other Applications // Radioelektronika'97. Conf. Proc. Slovak University of Technology, Bratislava, Slovakia. 1997. P. 477–480.

PARTIAL APPLIED THEORY OF UNMANNED AIR VEHICLES' VISION SYSTEMS

D. V. Vasilyev

*Open Joint Stock Company Research and Production Corporation
of Precision Scientific Instruments ("System of Precision Instrument Making"),
53 Aviamotornaya st., Moskva, 111024*

An overview of research conducted under the author's guidance over the period of 1972–2009 as a part of development and deployment of high-precision tracking vision systems with searchless correlation comparison of the current and reference videosignals. Core results and methodological approaches that can alleviate the design of vision systems for high-precision real-time orientation have been highlighted. The theory of correlation comparison gives two types of problems – the straight and the reverse. The straight problem makes an irreversible threshold-based choice out of a discrete set of signals, using correlation as a measure of similarity between the signal of choice and a reference signal. The reverse type of problem gives a continuous measure of shift between close signals. Therefore, in addition to correlation functions with an even parity near the zero-shift, a class of orthocorrelation functions with an odd parity, that can be used as a measure of difference between signals, produced by a one-dimensional continuous shift, is introduced. A linear procedure for evaluating the shift vector between finite signals, generating searchless algorithms using a weighted filtering operation on image frames, is considered. The contents correspond to the structure of a vision system, where correlation shift discriminators are used to measure tracking errors.

Keywords: cue field, projective shift, generalized correlation, weighted orthocorrelation, searchless correlation tracking, autolock, subpixel orientation accuracy.

Vasilyev Dmitry Valerievich — Leading Scientist, Doctor of Technical Science, Professor, e-mail: 02@niipp-moskva.ru.

АВТОМАТИЧЕСКОЕ ВЫДЕЛЕНИЕ И СОПРОВОЖДЕНИЕ ПСЕВДОТОЧЕЧНЫХ МЕРЦАЮЩИХ ОБЪЕКТОВ В РЕАЛЬНОМ ВРЕМЕНИ — МОДЕЛЬ ФОНОЦЕЛЕВОЙ ОБСТАНОВКИ, АЛГОРИТМ РАБОТЫ БОРТОВОГО ПРОЦЕССОРА И ТЕХНИЧЕСКИЕ РЕШЕНИЯ ПО ПРОЦЕССОРУ

С. В. Блажевич¹, В. Н. Винтаев², Н. Н. Ушакова²

¹ Белгородский государственный университет (БелГУ),
308015, Белгород, Россия, Победы, 85

² Белгородский университет потребительской кооперации (БУПК),
Белгород, Россия, Садовая, 116а

В работе представлены компьютерная модель фоноцелевой обстановки и программа, реализующая алгоритм непрерывного сопровождения движущегося мерцающего точечного объекта с вычислением картинных координат X и Y отслеживаемого объекта и выдачей прогноза поведения объекта на несколько шагов вперед. Модель демонстрирует высокую эффективность сопровождения цели на фоне интенсивной шумовой составляющей, имеющей статистические характеристики, практически совпадающие с характеристиками объекта, и возможность сопровождения до 1000 целей в секунду при использовании бортового процессора, построенного согласно этой модели.

Ключевые слова: автоматическое сопровождение, удаленные объекты, прогноз, бортовой процессор.

Введение

Один из немаловажных аспектов задач технического зрения — это обнаружение и сопровождение блестящих точек на фоне шумовой составляющей, статистические характеристики которой совпадают или мало отличаются от соответствующих характеристик мерцания сопровождаемого объекта. Такие задачи возникают в рамках проблем по слежению за астероидами и другими опасными объектами на звездном небе, по сопровождению в инфракрасном диапазоне наблюдения горячих точек (горячих выхлопных дюз реактивных аппаратов) летящих и взлетающих объектов, по слежению за слабоконтрастными треками малых частиц во взвесах и суспензиях, за распространением микросейсм — предвестников землетрясений, лавин и сходов покровов почв при наблюдении их с оптоэлектронных систем, размещаемых на подспутниковых и космических платформах.

В настоящей работе представлены компьютерная модель фоноцелевой обстановки и результаты работы программной реализации алгоритма непрерывного сопровождения движущегося объекта с вычислением картинных координат X и Y отслеживаемого объекта, выдающего прогноз поведения объекта на несколько шагов вперед.

Блажевич Сергей Владимирович — заведующий кафедрой, доктор физико-математических наук, доцент, e-mail: blazh@bsu.edu.ru.

Винтаев Виктор Николаевич — кандидат технических наук, доцент.

Ушакова Наталья Николаевна — кандидат технических наук, доцент.

МОДЕЛЬ ФОНОЦЕЛЕВОЙ ОБСТАНОВКИ

Представленный алгоритм предназначен для реализации в бортовом процессоре. Предполагается, что три координаты движущегося объекта и прогноз его движения по трём координатам вычисляются с помощью двух аналогичных процессоров, базирующихся на разнесенных в пространстве платформах. Эти координаты несут контролирующие исследуемое полупространство фокальные системы с сенсорами на ПЗС.

Траектория отслеживаемой цели представляется более высокой плотностью вычисляемых отсчётов, чем плотность отсчётов, порождаемая степенью дискретности распределения пикселей ПЗС на плоскости картинного транспаранта.

Оптические оси фокальных систем наблюдения пересекаются на существенном удалении от платформ — система из двух платформ предназначена для наблюдения за объектами на больших расстояниях. Модель фоноцелевой обстановки представляется в виде суммы регистрируемых излучений, соответствующих фоновой составляющей Земли (подстилающей поверхности) или неба, динамического изменяющегося от кадра к кадру шума (случайной составляющей) и точечных мерцающих подобно точкам шумовых составляющих движущихся объектов с параметрами движения, включающими случайные аддитивные слагаемые с дисперсией, ограничиваемой требованиями к динамическим характеристикам способности объекта к маневру.

На рис. 1 (см. с. 134) представлено рабочее окно моделирующей программы с выбранным изображением от подстилающей поверхности или неба (от фонового потока радиации) и динамическим (меняющимся от кадра к кадру) нормальным шумом с задаваемыми основными характеристиками.

На рис. 2 (см. с. 134) представлено рабочее окно программы с моделируемыми траекториями пролёта объектов. Объекты со случайной и регулярной составляющими в скоростях и ускорениях (показан их пролёт) являются выделенными пикселями, т. е. точками, мерцание которых соответствует мерцанию точек генерируемого шума. Для большей репрезентативности объекты на экране монитора представляются четырьмя связанными пикселями.

АЛГОРИТМ ОБНАРУЖЕНИЯ И СОПРОВОЖДЕНИЯ ЦЕЛИ

Компьютерная модель алгоритма слежения представляет следующий технологический процесс: на первом шаге выполняется расфокусировка цифрового кадра маской 7×7 и 13×13 с однородным заполнением ячеек маски и последующей нормализацией изображений на кадрах; далее выполняется попиксельное накопление с нормализацией кадров « 7×7 », а также накопление с нормализацией в другой вычислительной ветви кадров « 13×13 ». При накоплении более чем трёх кадров в каждой ветви реализуется подавление накапливающейся (паразитной) постоянной составляющей. Компьютерная модель предусматривает накопление до 40 кадров в каждой ветви. Таким образом формируются два суммокадра, из них попиксельным вычитанием синтезируется разностный кадр, на котором регулярные составляющие удаляются, так как являются в достаточной мере инвариантными к усредняющим маскам указанного размера, и остается разностная стохастическая составляющая кадра, имеющая в силу центральной предельной теоремы теории вероятности нормальное распределение с параметрами, идентичными параметрам исходных стохастических составляющих на кадрах — участниках операции

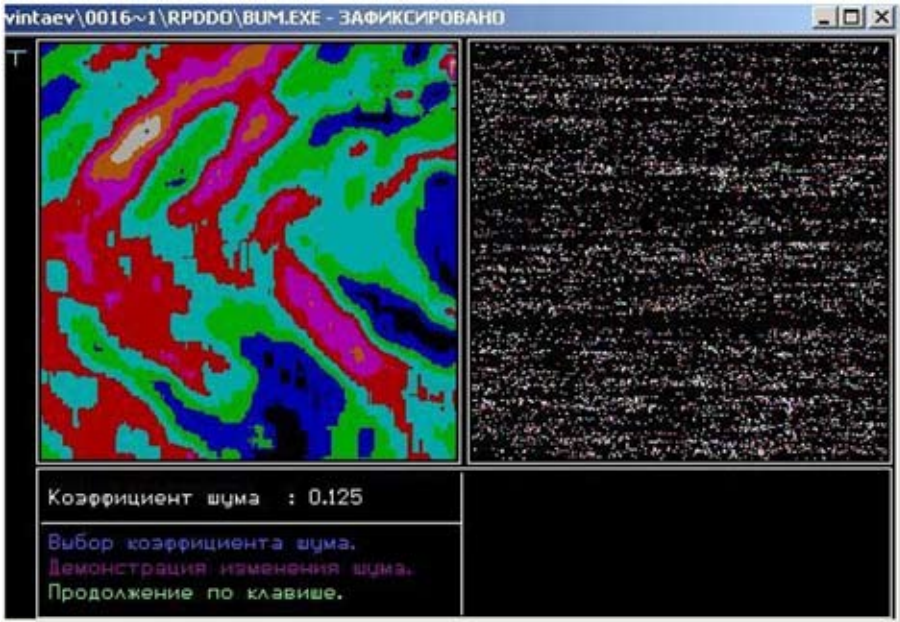


Рис. 1. Рабочее окно моделирующей программы с фоном и динамическим нормальным шумом

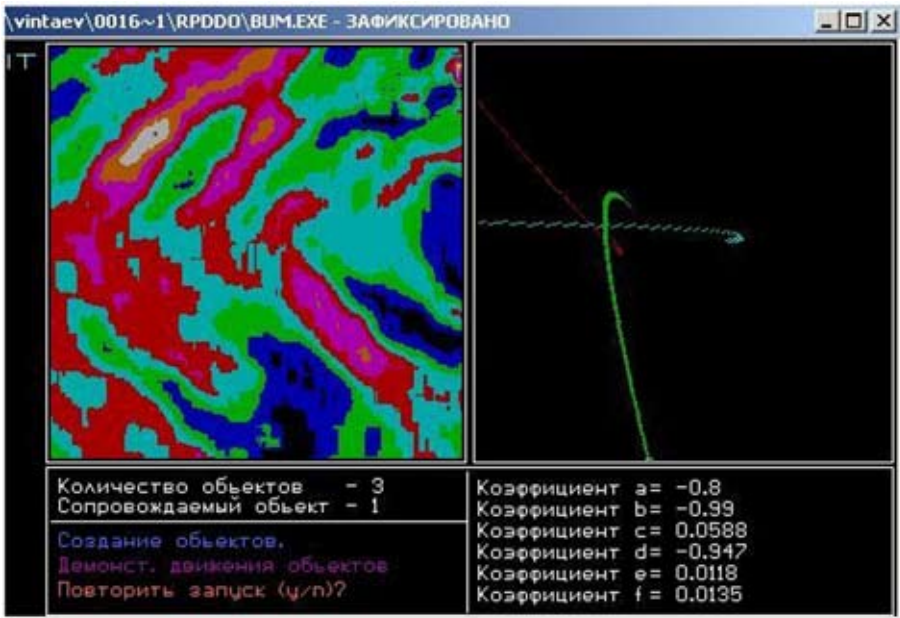


Рис. 2. Рабочее окно программы с моделируемыми траекториями пролёта объектов

вычитания. Далее простой процедурой восстанавливается выявленная регулярная составляющая, в состав которой могут входить фрагменты расфокусированных траекторий движения объектов и случайно образовавшиеся при выполнении операций накопления оси кратковременной корреляции расфокусированного шума.

Следующая пара вычислительных ветвей с разностным кадром на выходе начинает формироваться либо на втором первичном принятом с фокальной системы кадре, либо с задержкой на несколько кадров (до 39). Таким образом, на потоке первичных кадров формируется поток кадров (суммокадров), на котором при сопоставлении кадров по технологии «полупрозрачной кальки» пятна от расфокусированных точечных объектов будут частично наплзать друг на друга в соответствии с параметрами динамики маневра, на который способен объект (этим диктуется выбор апертур больших и малых масок усреднения и параметров накопления суммокадров). В модернизированном варианте моделирующей программы выполняется сначала накопление суммокадров с нормализацией, а потом свёртка их с расфокусирующей маской, так как процедура свёртки более затратная с вычислительной точки зрения.

На рис. 3 представлено рабочее окно с прослеженной траекторией одного из объектов, прогнозом его полёта на заданный период времени и эллипсом вероятности обнаружения, превышающей 0,7. В левой части рабочего окна виден незавершенный процесс свёртки смеси «фон + динамический шум + летящие мерцающие объекты» и малой расфокусирующей маски для вычисления следующей точки положения объекта.

В качестве критериев для выполнения процедур продолжения и сшивки траекторий с разрывами и построения эллипса с вероятностью обнаружения объекта, не меньшей 0,7, а также удаления остаточного шума на потоке суммокадров

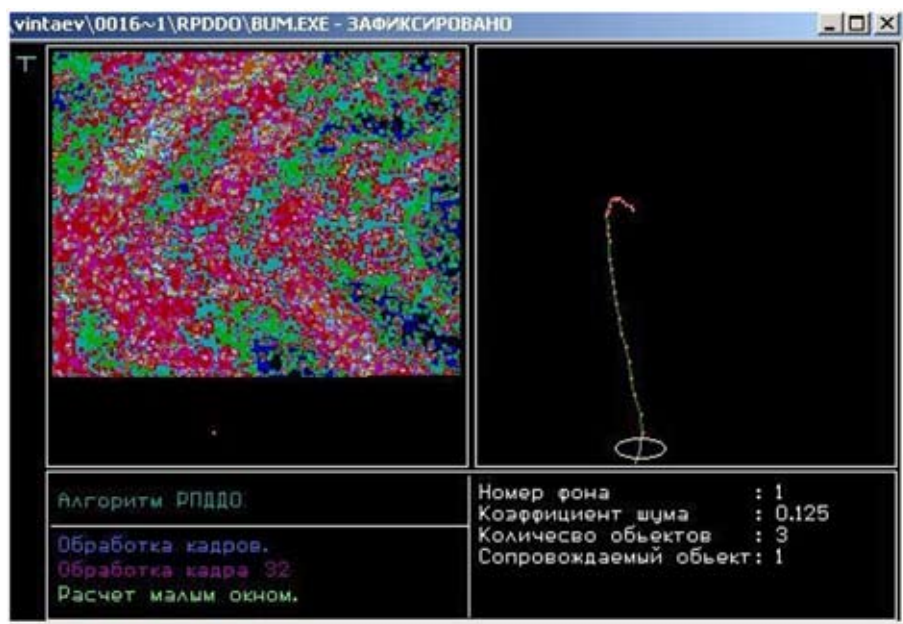


Рис. 3. Рабочее окно с прослеженной траекторией одного из объектов

используются параметрические критерии, вычисляемые на основе построения алгоритма-наблюдателя Неймана – Пирсона с модификацией решающих правил на нечёткое логическое исчисление с заданием разрешенной вероятности пропуска цели на ограниченном участке отслеживаемой траектории и с минимизацией вероятности ложной тревоги. При этом несколько кадров потока суммокадров накладываются друг на друга в режиме «полупрозрачной кальки», формируя изображение для анализа кинематических параметров выделяемой траектории и динамических параметров объекта, необходимых для продолжения траектории-прогноза (интегрирования уравнения движения на несколько шагов вперед по известным значениям скорости и ускорения объекта). Метод «полупрозрачной кальки» — это объединение нескольких нечётких множеств, где носитель множества — пиксели кадра, функция принадлежности каждого элемента множества — вероятность принадлежать к выявляемому классу регулярных объектов.

Процедура установки порогов, применяющаяся в процессе принятия решения (выявление регулярной составляющей в окне обработки данных и подавление остаточного шума) построена на использовании вычисленных в предварительных процедурах оценок интегральных параметров: σ (дисперсия исходного изображения), D_c (динамический диапазон контрастов выделяемых образов), Ω (полоса пространственных частот, занимаемая спектральными портретами регулярных составляющих), Δ_p (полоса пересечения пространственных частот спектральных портретов регулярных образов и шумов), τ (радиус корреляции шума), A (средняя апертура выделяемого образа) — и организуется в соответствии с методологией теории «Наблюдатель Неймана – Пирсона». При этом определяется ряд лингвистических переменных:

- переменная $\alpha_L = \alpha_L(\sigma, \Delta_p)$, задающая модальность атома: «фильтруемый или распознаваемый элемент есть шум». Лингвистические значения $\alpha_L = \alpha_L(\sigma, \Delta_p)$ соответствуют конкатенации лингвистических значений Δ и σ , определяемых функцией принадлежности высказываний «Нулевое», «Низкое», «Среднее», «Высокое» соответствующим гипотетическим нормированным интервальным значениям вида $[0, 0.1]$, $[0.1, 0.3]$, $[0.3, 0.7]$, $[0.7, 1]$, принадлежащим области значений вероятности ложной тревоги при заданных D_c и ε в методе «Наблюдатель Неймана–Пирсона»; в табл. 1 приведены лингвистические значения $\alpha_L = \alpha_L(\sigma, \Delta_p)$; в качестве значений элементов, являющихся результатом конкатенации, вновь использованы высказывания «Нулевое», «Низкое», «Среднее» и «Высокое» уже для переменной α_L ;

Таблица 1

Лингвистические значения $\alpha_L = \alpha_L(\sigma, \Delta_p)$

Δ_p	σ			
	Нулевое	Низкое	Среднее	Высокое
Нулевое	Нулевое	Нулевое	Нулевое	Нулевое
Низкое	Нулевое	Низкое	Низкое	Среднее
Среднее	Нулевое	Низкое	Среднее	Среднее
Высокое	Нулевое	Среднее	Среднее	Высокое

- переменная $\beta_L = \beta_L(\varepsilon, \Delta_p)$ — вероятность обнаружения в методе «Наблюдатель Неймана–Пирсона», определяющая модальность атома «фильтруемый или распознаваемый элемент есть регулярный образ»; пользуясь теми же терминами:

«Нулевое», «Низкое», «Среднее», «Высокое», — и для лингвистических значений параметров аргументов, и для значений оценки β_L можно получить соответствующую табл. 2;

Таблица 2

Лингвистические значения параметра β_L

ε	D_c			
	Нулевое	Низкое	Среднее	Высокое
Высокое	Нулевое	Нулевое	Нулевое	Нулевое
Среднее	Нулевое	Низкое	Низкое	Среднее
Низкое	Нулевое	Низкое	Среднее	Среднее
Нулевое	Нулевое	Среднее	Среднее	Высокое

• переменная $\gamma_L = \gamma_L(\alpha_L, \beta_L)$ — модальность атома «выявлен регулярный образ»; значение переменной γ_L определяется методом «взвешивания на рычажных весах», т. е. в какую сторону сместится значение γ_L , если на «чаши весов» класть различные значения α_L и β_L по обе стороны, соответственно, от стрелки весов (табл. 3).

Таблица 3

Зависимость значений лингвистической переменной γ_L от значений α_L и β_L

α_L	β_L			
	Нулевое	Низкое	Среднее	Высокое
Нулевое	Среднее	Высокое	Высокое	Высокое
Низкое	Нулевое	Среднее	Высокое	Высокое
Среднее	Нулевое	Низкое	Среднее	Высокое
Высокое	Нулевое	Низкое	Низкое	Среднее

В выбранном окне обработки для выделяемого элемента вычисляется соотношение $PG = \frac{\Delta_p \sigma \varepsilon}{D_c \Omega}$ ($D_c, \Omega \neq 0$), квазипропорциональное отношению «энергия шума/энергия сигнала», умноженному на верхнюю оценку вероятности ложной тревоги $\varepsilon = \tau/A$. При этом величина ε нормирует отношение $\Delta_p \sigma / D_c \Omega$ так, что при выборе для разрешения объектов, для которых $\varepsilon \ll 1$, она удерживает значения $PG \leq 0,3$. Это эквивалентно значению ошибки в распознавании $\Pi \leq 0,3$ и снимает возможные неопределённые ситуации в выводе лингвистических значений γ_L , так как определяемые по совокупности значений $\Delta_p, \sigma, D_c, \Omega$ при $PG \leq 0,3$ по методу «Наблюдатель Неймана – Пирсона» α_L и β_L будут $\alpha_L \leq 0,3, \beta_L > 0,3$ (что, в соответствии с данными табл. 3, определяет значение γ_L как «Высокое» ($\gamma_L \geq 0,7$)).

Процессор для реализации алгоритма построен по конвейерно-параллельной схеме с каскадным и тактируемым коммутатором блоков процессора друг с другом для реализации перепрограммирования его конфигурации.

Поле ПЗС-транспаранта 1024×1024 построено на основе блоков из ПЗС-матриц 32×32 с параллельным считыванием информации с блоков. Вычисления сверток выполняются погружением алгоритма свёртки и потоков данных в систему

остаточных классов [Ахо и др., 1979], выход в позиционную двоичную систему реализуется табличными преобразованиями.

Основное отличие предлагаемой программно-аппаратной системы от ранее реализованных вариантов [Алиева и др., 1987; Винтаев и др., 2001] заключается в следующем.

Предлагаемый вариант процессора предусматривает регулируемые дискретно расфокусировки при формировании кадров на ПЗС. При этом все расфокусировки и накопление изображений от кадра выполняются параллельно на оптической системе и подаются на ПЗС, одновременно суммируясь. Восстановление точечных объектов реализуется вычислением для каждого пятна его центра тяжести методом деления суммомомента пятна (суммы всех моментов — произведений координат зажженных пикселей в пятне на яркости пикселей) на сумму яркостей пикселей в пятне.

Разделение просуммированных кадров на ПЗС реализуется по следующей схеме.

Пусть S_0 — исходное изображение (кадр); F_0 — его фурье-пространственный портрет, а S_i — возмущенные на оптике изображения со спектральными портретами F_i . Тогда каждый портрет F_i есть F_0 , помноженный покомпонентно на R_i -спектральное представление оператора, моделирующего i -й возмущающий фактор. Все R_i известны, задаём и управляем ими в случае необходимости. Сложив все изображения на одной ПЗС: $S_0 + S_1 + S_2 + \dots = S_r$, на фурье-процессоре вычислим пространственный спектральный портрет $F_r = F_0(1 + R_1 + R_2 + \dots)$, откуда видно, что мы нашли и выделили все спектры F_i (перемножением найденного из предыдущего уравнения F_0 покомпонентно на R_i) и все изображения S_i и сохраняем их в памяти. Единица в скобке играет немаловажную роль, так как для некоторых R_i значения амплитуд высших гармоник будут нулевыми, а поскольку при определении F_0 эта скобка стоит в знаменателе, то эта единица не позволяет возникнуть сингулярности при вычислениях.

При сложении двух изображений на ПЗС следует учитывать, что разрядность цифрового преобразования сигналов с ПЗС на АЦП необходимо увеличить на единицу, при сложении четырёх изображений — на два и т. д. Кроме того, важно позаботиться о динамическом диапазоне чувствительности элементов фотоприёмной матрицы ПЗС, точнее: при сложении двух изображений, каждое из которых согласовано по динамическому диапазону яркости с ПЗС, необходимо обеспечить расширение динамического диапазона чувствительности ПЗС вдвое, при сложении четырёх изображений — вчетверо и т. д.

Принципиальные решения по процессору рассматривались в элементной базе исполнения на бескорпусных микросхемах ТТЛШ серии 700 (по 100 кристаллов на чип размером 10×10 см, но с предпринимаемыми особыми мерами по охлаждению чипа), а также на программируемых логических матрицах и характеризуются массогабаритными показателями: 0,5...1,0 кг и не более $20 \times 20 \times 10$ см.

Относительно характеристик быстродействия работы процессора следует отметить, что он в режиме поочередного сканирования объектов в потоке кадров может достаточно устойчиво сопровождать до 1000 объектов в секунду.

Заключение

В работе предложены модель фоноцелевой обстановки, алгоритм работы бортового процессора и технические решения по автоматическому выделению

и сопровождению цели, представляющей собой движущийся мерцающий псевдоточечный объект на фоне свечения поверхности Земли (или неба) и динамического изменяющегося от кадра к кадру светового шума. Для обеспечения решения задачи предприняты специальные меры по организации вычислительного процесса, позволяющие на шесть порядков снизить количество однократных операций по выполнению задачи. Разработанный алгоритм и технические решения по созданию процессора, его реализующего, могут быть эффективно использованы в широком спектре современных приложений.

Литература

- Алиева М. А., Винтаев В. Н., Исмаилов К. Х. (1987) Моделирование архитектуры бортового процессора с проблемной ориентацией // Исслед. Земли из космоса. 1987. № 2. С. 112–117.
- Ахо А., Хопкрофт Дж., Ульман Дж. (1979) Построение и анализ вычислительных алгоритмов / Пер. с англ. А. О. Слисенко; Под ред. Ю. В. Матиясевича. М.: Мир, 1979. 536 с.
- Винтаев В. Н., Константинов И. С., Ушакова Н. Н. (2001) Процессор целеуказания с матричным сенсорным полем // Технологич. конгресс «Современ. технологии при создании продукции военного и гражданского назначения»: Сб. докл. Омск, 2001. С. 330–333.

AUTOMATIC REAL-TIME SEPARATION AND TRACKING OF PSEUDO POINT BLINKING OBJECTS — MODEL OF THE TARGET-BACKGROUND SITUATION, ONBOARD PROCESSOR WORK ALGORITHM AND THE PROCESSOR TECHNICAL DECISIONS

S. V. Blazhevich¹, V. N. Vintaev², N. N. Ushakova²

¹ Belgorod State University, 85 Pobedy st., 308015, Belgorod Russia

² Belgorod University of Consumer Cooperatives, Belgorod, Russia

The computer model of the target-background situation as well as the program actualizing the continuous tracking of the moving blinking point object with the calculation of tangent coordinates X and Y of traceable object and the object behavior prediction for several steps ahead are presented. The model demonstrates the high efficiency of target tracking on the background of intensive noise component having the statistical characteristics practically coinciding with the characteristics of the object and an opportunity of tracking above the 1000 targets per second with the use of the on-board processor constructed under this model.

Keywords: automatic tracking, distant object, forecasting, onboard processor.

Blazhevich Sergey Vladimirovich — Head of Department, Doctor of Science, Associate Professor, e-mail: blazh@bsu.edu.ru.

Vintaev Viktor Nikolayevich — Candidate of Technical Science, Associate Professor.

Ushakova Nataliya Nikolayevna — Candidate of Technical Science, Associate Professor.

СИСТЕМА ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ЗАДАЧ СБЛИЖЕНИЯ, СТЫКОВКИ, ПОСАДКИ КОСМИЧЕСКОГО АППАРАТА НА ОСНОВЕ КОМПЬЮТЕРНОГО ВИДЕНИЯ

С. М. Соколов, А. А. Богуславский

*Учреждение Российской академии наук Институт прикладной математики
им. М. В. Келдыша РАН (ИПМ им. М. В. Келдыша РАН),
125047, Москва, Миусская пл., 4*

Описывается программный комплекс автоматического визуального мониторинга процесса сближения и стыковки космических аппаратов. Исходной информацией служит видеосигнал, идущий от телекамеры на одном из аппаратов. Получаемая последовательность кадров обрабатывается в реальном времени. В каждом кадре выделяются детали стыковочных узлов, размеры которых и геометрические параметры взаимного расположения используются как первичные измерения. По этим измерениям восстанавливается относительное движение космических аппаратов. Приводятся примеры обработки реальных видеоданных, полученных в процессе сближения и стыковки космических кораблей с международной космической станцией. Обсуждается возможность использования этого же комплекса для других задач информационного обеспечения целенаправленных перемещений космических аппаратов, в частности, посадки, на основе вычисления оптического потока.

Ключевые слова: компьютерное зрение, относительное движение, зрительная система реального времени, оптический поток, сближение, стыковка и посадка космических аппаратов.

Введение

Одним из эффективных средств информационного обеспечения решения навигационных задач в процессах сближения и стыковки, посадки космических аппаратов (КА) является визуальный канал.

До последнего времени этот канал использовался в «ручном» режиме, в виде визуальной обратной связи посредством зрительной системы человека в контуре управления КА (например, подсистема телеоператорного режима управления (ТОРУ) на транспортных кораблях и космических станциях). Человеческий фактор привносит в контур управления ряд особенностей, сдерживающих эффективность самого зрительного канала и всей системы управления (СУ) КА в целом.

Преимуществом человеческого восприятия является богатство его априорной модельной базы и возможностей комплексирования всего разнообразия данных о состоянии управляемой системы на текущий момент. Но в работе со зрительной системой человека есть ряд объективных ограничений, требующих использования дополнительных измерительных приспособлений. Как известно из психофизических исследований и опыта эксплуатации космической техники, человеку требуется довольно длительное время для аккомодации в условиях космоса (от 5 до

Соколов Сергей Михайлович — ведущий научный сотрудник, доктор физико-математических наук, профессор, e-mail: sokolism@keldysh.ru.

Богуславский Андрей Александрович — старший научный сотрудник, доктор физико-математических наук, доцент, e-mail: boguslav@keldysh.ru.

20 дней) и даже после этого периода скорость зрительных реакций остаётся существенно замедленной по отношению к земным показателям [Батулин, 2008; Грегори, 1970]. При этом человек может не ощущать каких-либо изменений в своём восприятии. Количественная оценка пространственно-временной информации в естественных зрительных системах также затруднена не только в космосе, а и в любых других условиях.

Современный уровень развития видеоаппаратуры и вычислительной техники позволяет формировать системы компьютерного зрения, которые дают возможность автоматизировать многие процессы сбора и обработки зрительных данных и включить зрительную обратную связь в СУ КА без/или с минимальным участием человека.

В статье описывается опыт автоматизации визуального мониторинга процессов сближения и стыковки космических кораблей (КК) с российским сегментом (РС) международной космической станции (МКС) [Boguslavsky et al., 2004, 2007]. Одной из особенностей зрительных данных, получаемых в космосе, является большая изменчивость образов объектов интереса из-за большого диапазона изменений освещения, от полной темноты или частичной освещённости с помощью прожекторов до предельной яркости при прямом освещении солнцем. В разработанных алгоритмах предусмотрены средства для парирования указанных трудностей.

Предлагаются способы настройки системы на обработку стыковок к различным стыковочным узлам и унифицированный интерфейс для работы с системой. В разработанном программном обеспечении поддерживается естественный переход от ручного режима работы к автоматическому и обратно с учётом текущей фоноцелевой обстановки. Приводятся результаты экспериментов по обработке реальных видеоматериалов.

Кроме того, анализируется возможность использования системы технического зрения (СТЗ) с той же аппаратной базой для определения движения КА относительно подстилающей поверхности при посадке [Богуславский и др., 2008, Sokolov et al., 2009]. Приводятся результаты экспериментов по использованию оптического потока для решения задач навигационного обеспечения наземных транспортных средств.

1. АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

В условиях массовых и габаритных ограничений на бортовые системы КА представляется весьма привлекательной возможность многофункционального использования имеющейся аппаратуры. Монокулярные СТЗ обеспечивают такую возможность с помощью замены алгоритмического обеспечения обработки зрительных данных.

С позиций обработки зрительных данных в алгоритмическом обеспечении описываемой системы выделяется два основных направления:

- обработка сцен, в которых присутствуют априори известные объекты (КА, стыковочные узлы, специальные мишени и др.), такая обработка необходима в информационном обеспечении задач сближения и стыковки космических аппаратов;
- обработка сцен, зрительные образы объектов которых априори не известны, такая обработка необходима в информационном обеспечении задач движения КА над поверхностью планет и посадки в необорудованных местах.

1.1. Автоматизация визуального мониторинга процессов сближения и стыковки КА

Конечная цель программного комплекса для автоматизации визуального мониторинга процессов сближения и стыковки КА — полная автоматизация визуального контроля стыковки КА с момента видимости в поле зрения ТВ-камеры КА (приблизительно 150...200 м) и до полной стыковки космических аппаратов.

На основных этапах сбора и обработки зрительных данных комплекс работает подобно человеку-оператору. Начинается обработка зрительных данных с попытки «узнать» в полученном изображении образы объектов, с которыми происходит сближение и планируется стыковка. Если объекты «узнаны», то выполняется оценка относительно расположения космических аппаратов.

В дополнение к действиям оператора комплекс вычисляет и отображает в виде, принятом для анализа, количественные параметры, характеризующие относительное движение КА: расстояние, скорость сближения, относительное угловое положение, угловые скорости.

Всё алгоритмическое обеспечение разбито на две основные части.

Измерительная часть. К этой части отнесены все алгоритмы обработки зрительных данных. Конечная цель: в цифровом изображении выделить характерные точки, по взаимному расположению которых может быть определено относительное положение/смещение видеокамеры и объекта наблюдения.

На входе измерительной части — видеопоследовательность с изображением области пространства, в котором может располагаться объект интереса (другой космический аппарат или участок поверхности для посадки).

На выходе — множество координат характерных точек и моментов времени, в которые они определены.

Вычислительная часть. Задача алгоритмов этой части: по заданному множеству характерных точек, известным геометрическим соотношениям между ними, временам их определения и параметрам видеокамеры рассчитать взаимное расположение видеокамеры и наблюдаемого объекта в зависимости от времени.

На входе вычислительной части — множество координат характерных точек и моментов времени, в которые они получены.

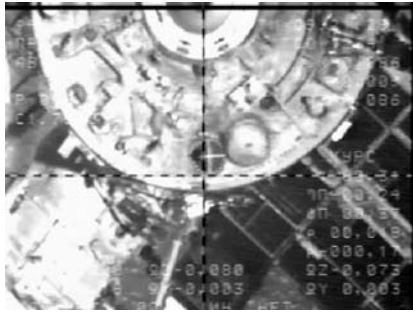
На выходе — закон относительного движения видеокамеры и объекта в поле зрения.

Рассмотрим алгоритмическое обеспечение этих частей более подробно.

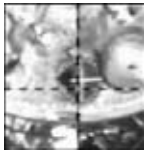
1.1.1. Измерительная часть

В алгоритмическом обеспечении этой части выделяется ряд этапов. В проведённых экспериментах использовались видеоданные, полученные на Земле по различным каналам. Перед предоставлением этих видеоданных к обработке выполняется их предобработка. Видеоданные приводятся к унифицированному формату (720×576 элементов изображения с 256 уровнями яркости), удаляются искусственные объекты, внесённые в видеоданные (перекрестие визира, данные радиолокационной системы).

Выделение областей интереса на текущем изображении (в англоязычной транскрипции ROI — region of interest). Задача этого этапа обработки зрительных данных — уменьшение объёма обрабатываемых данных, повышение надёжности результатов обработки. На входе — полное изображение поля зрения



a



б

Рис. 1. Пример выделения области интереса в поле зрения ТВ-камеры на борту КК: *a* — полное изображение поля зрения камеры на борту КК; *б* — изображение области интереса (ROI) (фрагмента, содержащего причальную мишень)

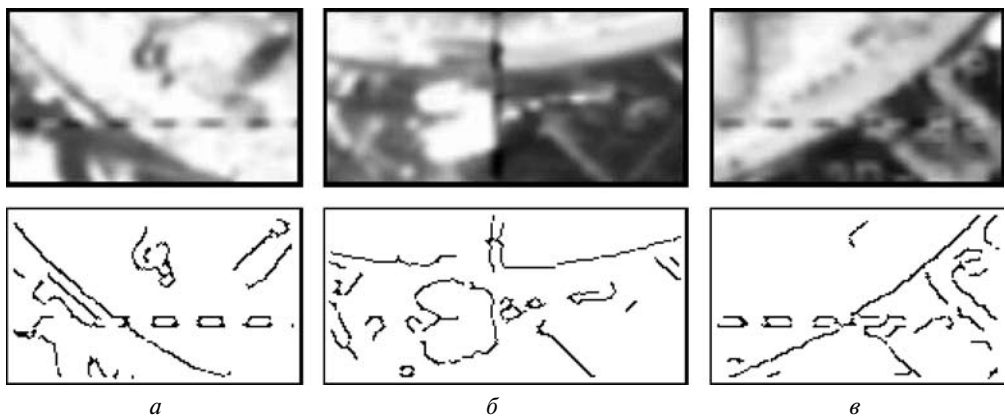
видеосистемы. На выходе — часть (фрагмент) изображения поля зрения видеосистемы, чаще всего прямоугольная область со сторонами, параллельными строкам и столбцам изображения, содержащая образ объекта интереса.

Метод решения — использование априорной информации, данных от других источников (например, указание оператора), результатов обработки предыдущих кадров видеопоследовательности.

На рис. 1 представлен пример выделения области интереса в задаче информационного обеспечения стыковки. В этом примере область интереса выделяется путём учёта результатов обработки предыдущих кадров видеопоследовательности.

Предобработка зрительных данных в области интереса. Как указывалось во введении, одной из особенностей зрительных данных, получаемых в космосе, является большая изменчивость образов объектов интереса из-за большого диапазона изменений освещения. Для частичного парирования этой изменчивости перед дальнейшей обработкой по гистограмме яркостей области интереса оценивается и нормируется яркость элементов этой области.

Выделение (распознавание) образов объектов интереса. Процесс распознавания начинается с выделения первичных признаков образов объектов на изображении: выделения областей с примерно равной яркостью (сегментация) и нахождения перепадов яркости — «краёв» образов объектов. Для сокращения



a

б

в

Рис. 2. Пример выделения края внешнего обвода станции: *a* — слева; *б* — по центру; *в* — справа

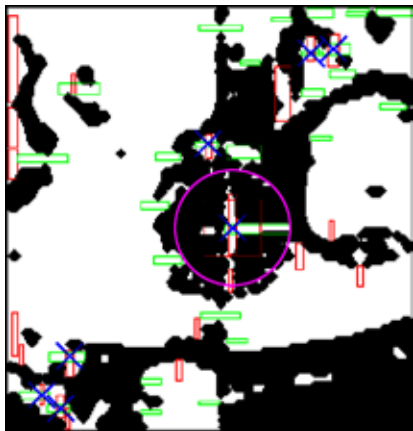


Рис. 3. Выделение креста мишени среди кандидатов на основе априорной информации (размеры и относительное расположение переключин креста)

времени обработки активно используется априорная информация о форме образов объектов интереса. Эта форма задаётся алгоритму в виде масок — областей интереса, грубо (с большим допуском по размеру) повторяющих очертания искомых объектов. Рисунок 2 (см. с. 143) иллюстрирует выделение внешней границы космической станции в отдельных фрагментах изображения.

На рис. 3 представлен пример выделения сегментов креста мишени на бинарном изображении с учётом априорной информации, а на рис. 4 — меток основания мишени на полутоновом изображении также с учётом априорной информации об их возможном расположении.

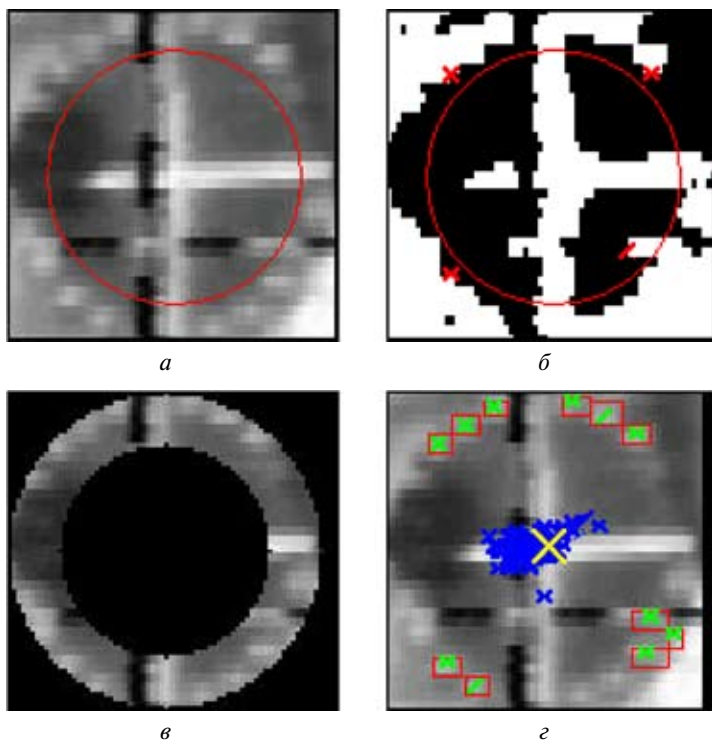


Рис. 4. Пример выделения меток мишени на основании априорной информации об их расположении: *a* — оценка расположения области интереса для выделения меток; *б* — грубая оценка радиуса мишени по результатам обработки диаметрально расположенных фрагментов/кандидатов меток; *в* — кольцообразная область интереса с метками мишени; *г* — результаты распознавания меток мишени и уточнения радиуса мишени

Определение размеров и координат характерных точек распознанных образов. После того как на предъявленном изображении выделены образы интересующих объектов, на них выбираются характерные точки. Проиллюстрируем формат данных о характерных точках на примере специальной причальной мишени, расположенной на стыковочном модуле (СМ) российского сегмента МКС. Измерительная информация, полученная в результате обработки одного телевизионного полукадра кадра, имеет вид:

N — порядковый номер измерения ($N = 1, 2, 3, \dots$);

$T, \Delta T$ — время привязки полукадра (в секундах от начала ввода видеоданных) и длительность выполнения измерений (мс);

X_C, Y_C — координаты центра креста (действительные числа);

$N_1, X_1, Y_1, X_2, Y_2, \dots, X_{N_1}, Y_{N_1}$ — количество точек на верхней горизонтальной стороне креста и их координаты (целые числа);

$N_2, X'_1, Y'_1, X'_2, Y'_2, \dots, X'_{N_2}, Y'_{N_2}$ — количество точек на нижней горизонтальной стороне креста и их координаты (целые числа);

$N_3, U_1, V_1, U_2, V_2, \dots, U_{N_3}, V_{N_3}$ — количество точек на левой вертикальной стороне креста и их координаты (целые числа);

$N_4, U'_1, V'_1, U'_2, V'_2, \dots, U'_{N_4}, V'_{N_4}$ — количество точек на правой вертикальной стороне креста и их координаты (целые числа);

X_O, Y_O, R — координаты центра окружности мишени и её радиус (действительные числа);

$N_5, A_1, B_1, A_2, B_2, \dots, A_{N_5}, B_{N_5}$ — количество точек на окружности мишени и их координаты (целые числа).

Здесь все координаты выражены в пикселах. Во всех вариантах обработки выполнены соотношения: $N_1 = N_2$ и, если $N_1 > 0$, то $X_i = X'_i$, $N_3 = N_4$ и, если $N_3 > 0$, то $V_i = V'_i$. Описываемые алгоритмы выполнения этих соотношений не учитывают. Несмотря на то, что результаты обработки кадра — величины X_C, Y_C и X_O, Y_O, R — в указанных данных имеются, в процедуре определения движения (вычислительная часть) эти величины находятся независимо по исходной информации.

1.1.2. Вычислительная часть

Результаты работы алгоритмов измерительной части — координаты характерных точек образов объектов и моменты времени их определения — передаются расчётным алгоритмам. Здесь методом наименьших квадратов с использованием формул пересчёта координат точек реальной станции в координаты тех же точек на её изображении в кадре находятся координаты фокуса камеры и её ориентация в строительной системе координат КА, который находится в поле зрения видеокамеры. Функционал метода наименьших квадратов формируется так, чтобы его минимизация по определяемым параметрам позволила наилучшим образом согласовать расчётные и измеренные координаты выбранных точек на видеоизображении. Найденные таким образом координаты фокуса камеры и параметры её ориентации пересчитываются в параметры положения строительной системы КА, на котором установлена камера, относительно строительной системы координат наблюдаемого КА [Богуславский и др., 2004].

В результате описанной обработки получается последовательность положений КА (точнее, последовательность параметров, задающих положение его строительной системы координат) во времени. Интерполяция этой последовательности позволяет рассчитывать положение КА для любого момента на представляющем интерес отрезке времени, а геометрические модели КА позволяют визуализировать это положение.

Алгоритм определения движения космического аппарата. Движение КА определяется в реальном времени в результате пошаговой обработки последовательности кадров ТВ-изображения креста и мишени. Данные обрабатываются отдельными порциями. Порции либо имеют фиксированный объём, либо состояются из данных, полученных на отрезках времени фиксированной длины. При обработке второй и последующих порций учитываются результаты обработки предыдущих порций.

Каждая порция обрабатывается в два этапа. На первом этапе определяется движение центра масс корабля, на втором этапе — его движение относительно центра масс. Математическая модель движения корабля принимается в виде

$$d_1 = z_1 + z_2 t, \quad d_2 = z_3 + z_4 t, \quad d_3 = z_5 + z_6 t + \frac{1}{2} z_7 t^2,$$

$$\varphi_1 = v_1 + v_2 t, \quad \varphi_2 = v_3 + v_4 t, \quad \varphi_3 = v_5 + v_6 t,$$

где t — время, отсчитываемое от начала обработки последовательности кадров; z_i и v_j — постоянные коэффициенты; $\varphi_1, \varphi_2, \varphi_3$ — активные углы тангажа, рысканья, крена. Подробно алгоритмы определения движения КА описаны в статье [Богуславский и др., 2004].

Для более удобного представления движения вычисляются величины:

$$\rho = \sqrt{d_1^2 + d_2^2 + d_3^2}, \quad \frac{d\rho}{dt}, \quad \alpha = \arctan \frac{d_2}{\sqrt{d_1^2 + d_3^2}}, \quad \beta = \arctan \frac{d_1}{d_3},$$

где ρ — расстояние, $d\rho/dt$ — скорость; α, β — пассивные углы тангажа и рысканья.

1.2. Определение движения КА относительно подстилающей поверхности в процессе посадки

В этой задаче нет априорной информации о наблюдаемых объектах. Определение смещений видеокамеры в подобных случаях связано с вычислением оптического потока.

За алгоритмическую основу был выбран метод Лукаса — Кэнэда (Bruce D. Lucas и Takeo Kanade) над пирамидой изображений [Bouguet, 2002]. Этот метод, с одной стороны, обладает возможностями автоматической настройки на свойства подстилающей поверхности, а с другой стороны, за счёт сокращения количества характерных точек (так называемый разреженный оптический поток) позволяет вести обработку в масштабе реального времени.

Метод разреженного оптического потока состоит из трёх этапов: выделение признаков областей, совмещение изображений, сегментация полученного векторного поля и его интерпретация. Опишем основные этапы алгоритма обработки зрительных данных.

Предобработка зрительных данных. Для устранения избыточности зрительных данных на дальнейшую обработку передаются только те точки исходного изобра-

жения, в которых присутствует изменчивость, определяемая по разности последовательных кадров. Тем самым исключаются из обработки точки однородных областей изображения, на которых оптический поток не может быть определён.

Первый этап обработки — вычисление местоположения характерных точек.

На этом этапе используется модифицированный алгоритм Харриса [Bouguet, 2002], который реагирует на углы. Под углом в данном случае понимается особенность локации изображения, где изменение функции интенсивности одновременно в двух направлениях максимально.

Второй этап обработки — определение векторов смещений характерных точек. Здесь сравниваются характерные точки на двух кадрах и вычисляются векторы смещений соответствующих точек.

Согласно методу Лукаса–Кэнада вектор, который оптимально соотносит точку первого изображения с точкой второго, может содержать погрешность. Чтобы уменьшить эту погрешность, данный метод применяется итеративно, т. е. найденный вектор является входным параметром алгоритма для нахождения нового, более точного. Операция повторяется до достижения определённой точности или количества итераций. Гауссова пирамида изображений — ряд изображений, разрешения которых последовательно уменьшаются в два раза. В основании пирамиды лежит исходное изображение.

Сначала вычисляются векторы на самом верхнем уровне пирамиды, результат уточняется до тех пор, пока не будет получена погрешность, достаточная для применения на следующем (более низком) уровне пирамиды. Этапы вычислений повторяются для всех уровней, на последнем уровне получаются векторы смещений — векторы оптического потока.

Третий этап — фильтрация полученных векторов. Эта задача требует учёта специфики конкретной фоноцелевой обстановки, позволяющей исключить из рассмотрения те векторы, которые являются «выбросами» в процессе измерений. Спроектирован алгоритм, выполняющий следующие действия:

- строится гистограмма направлений векторов смещений;
- устанавливается направление с максимальным количеством векторов;
- вычисляется средний вектор из всех векторов, которые попадают в промежуток с заданным отклонением от найденного направления; если средний вектор не найден, то используется значение из предыдущего цикла вычислений.

2. ПРОГРАММНЫЕ ТЕХНОЛОГИИ

Описанные выше алгоритмы решают поставленную задачу только при условии успешной реализации в реальном времени на доступных ресурсах универсальных ЭВМ. Обеспечить выполнение названных условий удалось на основе методов программирования систем технического зрения реального времени, разработанных в группе компьютерного видения ИПМ им. М. В. Келдыша РАН [Богуславский, Соколов, 2003, 2005, 2006а, б, 2007].

Компонентный каркас для разработки программного обеспечения СТЗ реального времени

Выполняется всё программное обеспечение (Про) на языке Си⁺⁺.

Каркас — заготовка прикладной программы СТЗ для доработки применительно к конкретной задаче. Каркас приложения представляет собой наращиваемую среду разработки ПрО СТЗ. Компоненты этой среды: классы и подсистемы.

Архитектура каркаса основана на применении трёх параллельных подсистем:

- ввода зрительных данных;
- интерфейса пользователя;
- обработки изображений.

Предлагаются два основных режима функционирования приложений: режим настройки (рис. 5) и автоматический режим (рис. 6, см. с. 149).

Для использования компонентного каркаса в прикладных задачах выполняются следующие действия (основные этапы):

- сбор представительных образцов зрительных данных;
- проектирование высокоуровневого алгоритма обработки изображений и оформление в виде объекта для подключения к каркасу;
- проектирование цикла реального времени и оформление в виде объекта для подключения к каркасу;
- тестирование ПрО СТЗ на представительных образцах зрительных данных;
- проектирование интерфейса пользователя.

В следующем разделе приведём примеры реализации вышеуказанных действий в нескольких прикладных задачах.

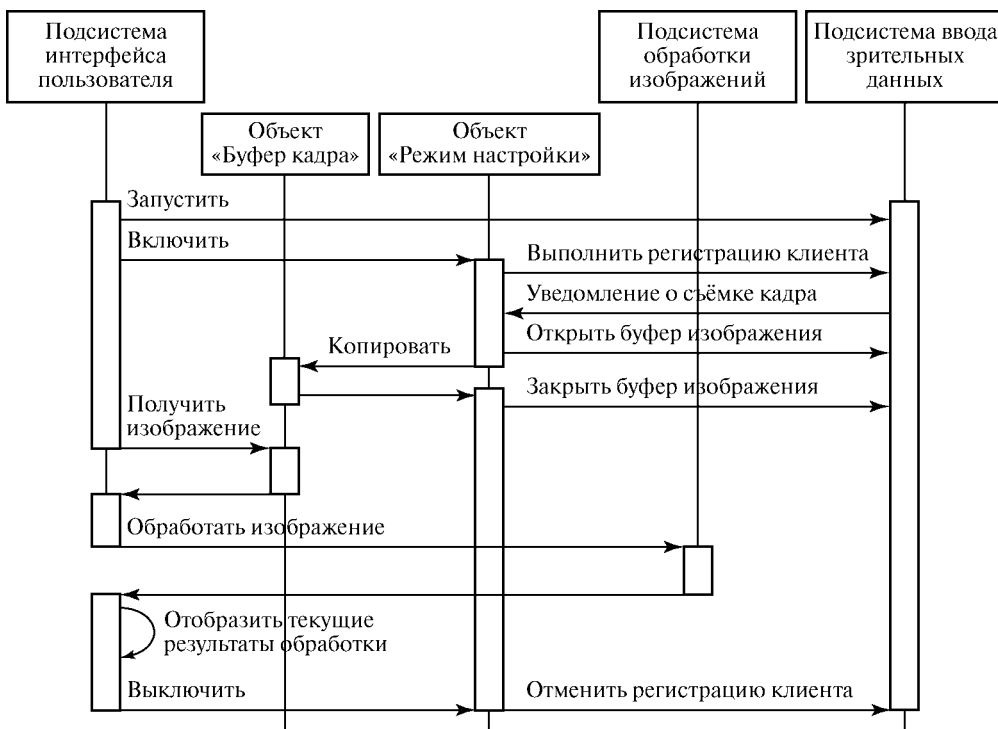


Рис. 5. Координация взаимодействия подсистем в режиме настройки

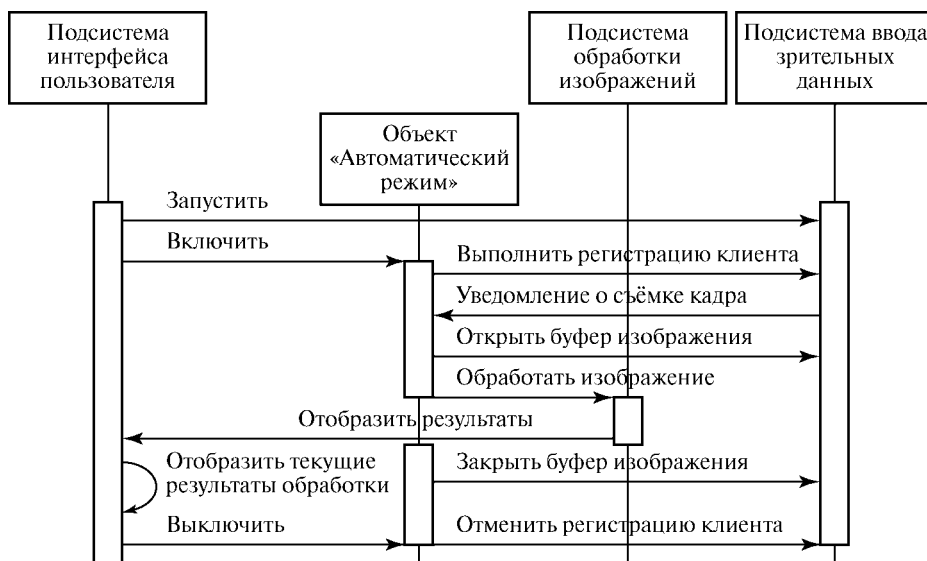


Рис. 6. Координация взаимодействия подсистем в автоматическом режиме

3. ПРИМЕРЫ РЕАЛИЗАЦИИ

3.1. Система автоматизированного мониторинга процессов сближения и стыковки КК «Прогресс» и МКС

На основе описанных алгоритмов была реализована система автоматизированного мониторинга процессов сближения и стыковки КК «Прогресс» и МКС [Boguslavsky, Sokolov et al., 2004]. Эта система опробовалась на видеоматериалах о сближении и стыковке КК «Прогресс» и МКС.

Видеокамера располагалась на КК, а специальная причальная мишень — на стыковочном узле МКС. На рис. 7 (см. с. 150) представлен основной экран для работы с системой.

Качество результатов определения параметров относительного движения оценивалось в рамках обычных допущений метода наименьших квадратов (расчёт стандартных отклонений величин, вычисленных по очередной порции данных) [Богуславский и др., 2008] и сравнением с данными традиционной радиолокационной системы (рис. 8, см. с. 150).

Точности, которые достигаются при расчёте перечисленных параметров, следующие: расстояние 0,2 %; углы пассива (все три): 0,1°; углы актива (все три): 0,2°.

3.2. Система автоматизированного мониторинга процессов сближения и стыковки КК ATV и МКС

В этой задаче видеокамера располагается на МКС, а специальная причальная мишень на КК. Это обстоятельство практически не изменило измерительную часть и внесло технические коррективы в расчётную часть программного обеспечения. Опробовалась система на видеоматериалах компьютерного моделирования

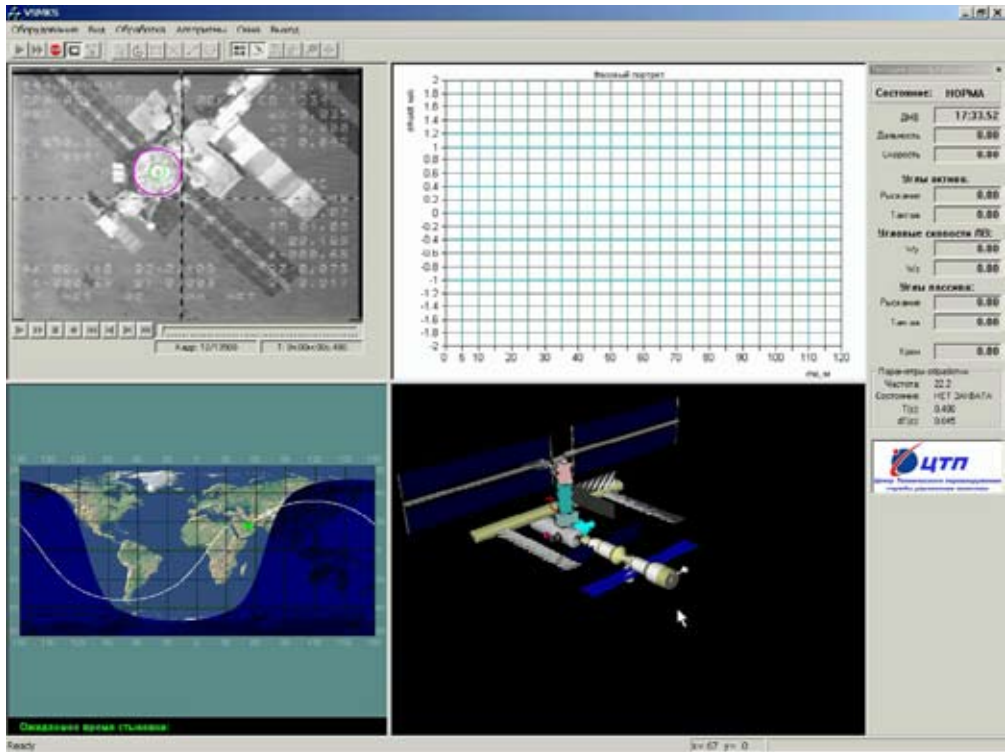


Рис. 7. Основной экран системы автоматизированного мониторинга процессов сближения и стыковки КК «Прогресс» и МКС

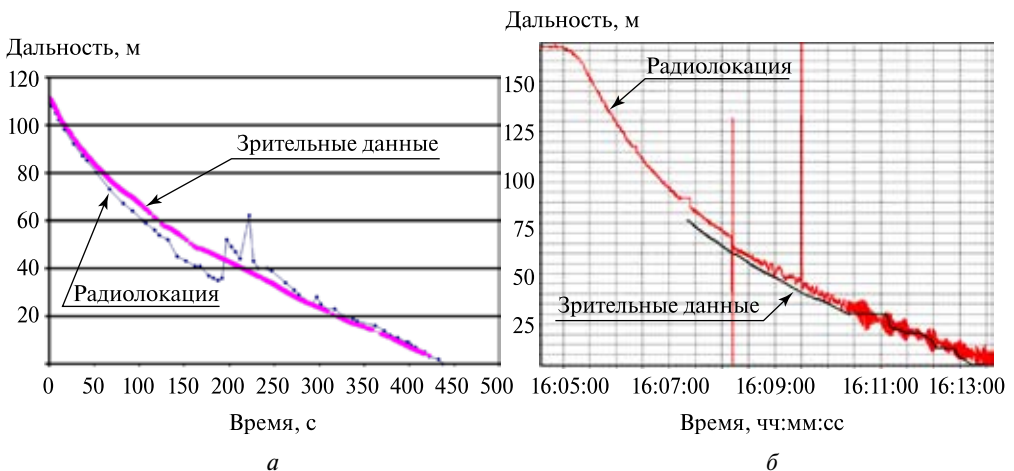


Рис. 8. Сравнение результатов определения расстояния между КК и МКС в зависимости от времени по зрительным данным и традиционным (радиолокация) методам: а — темным цветом обозначены результаты, полученные традиционными методами, розовым — с помощью СТЗ; б — красным цветом обозначены результаты, полученные традиционными методами, черным — с помощью СТЗ

названных процессов и видеоматериалах реального полёта КК ATV [Boguslavsky et al., 2007]. На рис. 9 представлен основной экран для работы с системой.

Этот вид панели — интерфейса СТЗ с оператором — принят в качестве унифицированного для всех подобных задач и содержит три основные зоны (присутствующие в окне программы в большинстве режимов работы с системой): левую, среднюю и правую.

В левой зоне представлены — окно с отображением поля зрения СТЗ (вверху) и окно поля зрения СТЗ с наложенной маской трёхмерной модели станции (внизу) (см. рис. 9).

В средней зоне представлены три закладки: «Боковые виды», «Графики» и «Таблица» (на рис. 9 выбрана/активизирована закладка «Боковые виды»). В окнах, появляющихся при выборе той или иной закладки, отображаются визуальные представления результатов обработки изображений различных видов.

В правой зоне располагается панель «Текущее состояние» для представления параметров текущего состояния системы в числовом виде.

Кроме того, в основном окне программы содержатся три группы интерфейсных элементов (см. рис. 9): «Верхнее меню» содержит команды для включения/выключения панелей инструментов и дополнительные панели — «Панели инструментов» для управления программой (выбор основных команд и режимов); «Группа элементов управления», связанных с полем зрения (управление видеоданными, подаваемыми на обработку АТИС), — под окном с отображением поля зрения СТЗ.

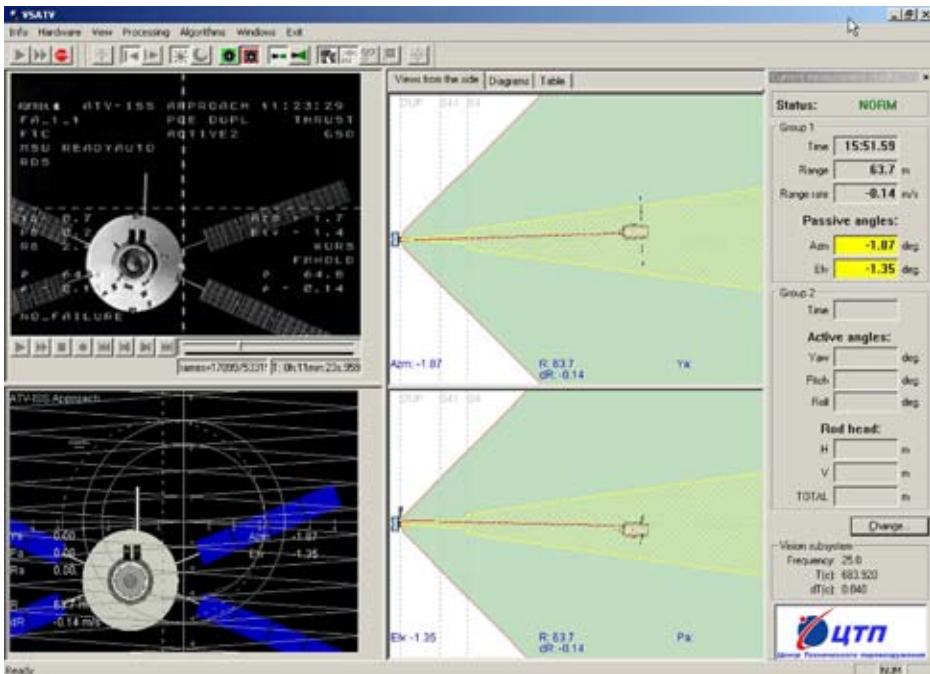


Рис. 9. Основной экран системы автоматизированного мониторинга процессов сближения и стыковки КК ATV и МКС

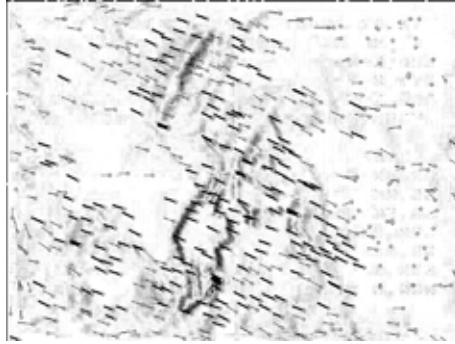


Рис. 12. Найденный оптический поток (тёмные векторы — векторы, прошедшие фильтрацию)

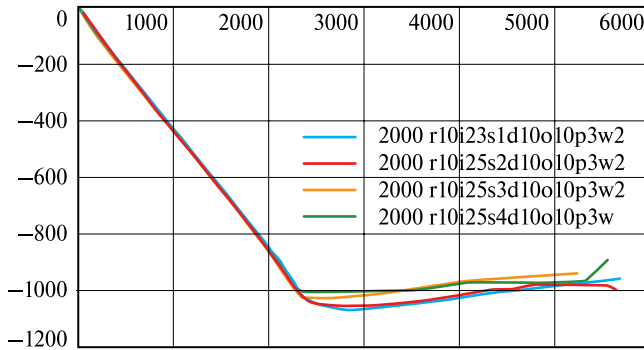


Рис. 13. Траектории движения космического аппарата, вычисленные при $r = 10$ и $p = 3$ (коэффициент уменьшения масштаба: $s = 1, s = 2, s = 3, s = 4$). Приняты обозначения вида: $2000\ r10i25s4d10o10p3w2$, где первое число — количество кадров в обрабатываемой последовательности, далее записываются параметры и их значения без знака «равно» для экономии места; i — количество итераций при субпиксельном уточнении точек и при поиске оптического потока; d — расстояние между признаковыми точками; o — размер окна поиска; w — параметр, задающий угловой диапазон разброса при фильтрации векторов

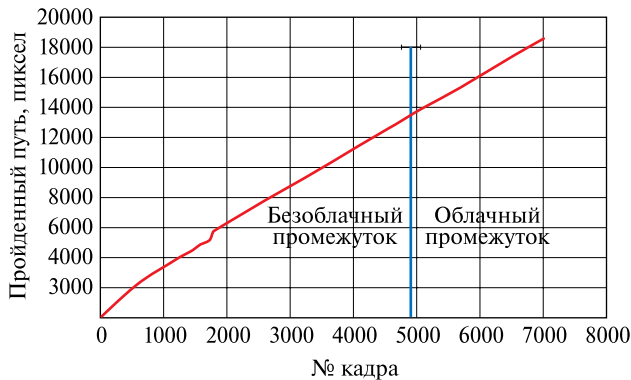


Рис. 14. Зависимость пройденного пути от времени. Синяя вертикальная линия обозначает примерное место перехода с безоблачной поверхности Земли на облачную поверхность

Вычисление оптического потока проводилось на одной и той же последовательности, состоящей из 2000 кадров, в каждой серии менялись разрешения видеоряда — уровень пирамид Гаусса $s = 1, 2, 3, 4$ ($768 \times 576, 384 \times 288, 256 \times 192, 192 \times 144$). Изменялись следующие параметры: r — дистанция между разностными кадрами для расчёта оптического потока; p — количество пирамид. Оценивались такие параметры как доля ошибочно найденных векторов, среднее время обработки одного кадра, общее время обработки последовательности. На рис. 13 представлен результат обработки одной из видеопоследовательностей.

В этой серии время обработки от 0,48 до 0,38 с. Доля ошибок при $s = 1$ и $s = 2$ равна 21 и 23 % соответственно. При $s = 3$ доля ошибок равна 38 %, однако время обработки кадра 0,20 с. При $s = 4$ доля ошибок составляет 66 %.

Значительное количество ошибок (более 30 %) обусловлено тем, что данная последовательность содержит ряд кадров без ярко выраженной структуры (около 25...30 % от всего ролика).

Важным моментом при нахождении траектории видимого движения является вопрос о работе алгоритма при переходе с безоблачной поверхности Земли на облачную и наоборот. За облачную принимаем такую поверхность, на которой количество областей с большой яркостью (области со значением яркости более 200 единиц в диапазоне 0...255) превышает 40 %. По результатам обработки (см. рис. 14) видно, что при данном переходе не происходит существенного изменения видимой скорости.

Заключение

Проведённые исследования и разработки показывают возможность успешного использования системы технического зрения для информационного обеспечения процессов сближения, стыковки и посадки космических аппаратов. Результаты экспериментов по определению параметров относительного движения КК и МКС на основе видеопоследовательностей, переданных с борта КК, подтверждают высокую надёжность и точность предложенных решений.

Открытая архитектура и компонентные технологии позволяют говорить о возможности формирования многофункционального информационного комплекса на основе системы компьютерного видения.

Все элементы описанной системы и её функциональные возможности продолжают совершенствоваться. Исследования развиваются в направлении разработки системы информационного обеспечения сближения и стыковки к различным стыковочным узлам (стыковочный отсек модуля «Пирс» (СО1), стыковочный узел функционально-грузового блока (ФГБ) и др.).

Литература

- Батурин Ю. М.* (2008) Восприятие пространства в космосе и живописи (Физика. Психология. Эксперименты) // Материалы Циолковских чтений. ГМИК им. К. Э. Циолковского, 2008.
- Богуславский А. А., Кирильченко А. А., Платонов А. К., Соколов С. М., Трифонов О. В., Ярошевский В. С.* (2008) Построение описания внешней среды в системах информационного обеспечения мобильных робототехнических комплексов // Мехатроника, автоматизация, управление. 2008. № 12. С. 50. Прилож. С. 15–24.
- Богуславский А. А., Сазонов В. В., Соколов С. М., Смирнов А. И., Сайгираев Х. У.* (2004) Автоматический мониторинг стыковки космического корабля с орбитальной станцией по ви-

- деоинформации: Препринт. № 74. М.: Ин-т прикладной математики им. М. В. Келдыша РАН, 2004. 23 с.
- Богуславский А. А., Соколов С. М.* (2003) Компонентные методы разработки программного обеспечения СТЗ // Материалы научной шк.-конф. «Мобильные роботы и мехатронные системы». МГУ им. М. В. Ломоносова, 17–18 нояб. 2003. М., 2003. С. 35–49.
- Богуславский А. А., Соколов С. М.* (2005) Компонентный каркас для разработки программного обеспечения систем технического зрения реального времени // Тр. 2-й Всерос. науч. конф. «Методы и средства обработки информации». МГУ им. М. В. Ломоносова. 5–7 окт. 2005. М., 2005. С. 337–343.
- Богуславский А. А., Соколов С. М.* (2006а) Обнаружение образов объектов на изображениях с использованием структурных геометрических моделей в системах технического зрения реального времени // Программирование. 2006. № 3. С. 66–80.
- Богуславский А. А., Соколов С. М.* (2006б) Обнаружение образов объектов с изменяющейся структурой в системах технического зрения реального времени // Информац. технологии и вычислит. системы. 2006. № 1. С. 61–77.
- Богуславский А. А., Соколов С. М.* (2007) Реализация высокоуровневых алгоритмов обработки изображений в системах технического зрения реального времени // Вестн. Коломенского гос. пед. ин-та (Мат. и естеств. науки). 2007. № 2. С. 211–221.
- Грегори Р. Л.* (1970) Глаз и мозг. М.: Прогресс, 1970. 271 с.
- Boguslavsky A. A., Sazonov V. V., Sokolov S. M., Smirnov A. I., Saigiraev K. U.* (2004) Automatic Vision-based Monitoring of the Spacecraft Docking Approach with the International Space Station // Proc. 1st Intern. Conf. Informatics in Control, Automation and Robotics (ICINCO 2004). Setúbal, Portugal. 2004. V. 2. P. 79–86.
- Boguslavsky A. A., Sazonov V. V., Sokolov S. M., Smirnov A. I., Saigiraev K. U.* (2007) Automatic vision-based monitoring of the spacecraft ATV rendezvous/ separations with the International Space Station // Proc. 4th Intern. Conf. Informatics in Control, Automation and Robotics (ICINCO 2007). Angers, France. 2007. V. 1. P. 284–291.
- Bouguet J.-Y.* (2002) Pyramidal Implementation of the Lucas Kanade Feature Tracker. Intel Corporation, Microprocessor Research Labs. OpenCV Documents, 2002.
- Shi J., Tomasi C.* (1994) Good features to track // Proc. IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition. 1994. P. 593–600.
- Sokolov S. M., Boguslavsky A. A., Kuftin F. A.* (2009) Vision System for Relative Motion Estimation from Optical Flow // Proc. 13th Intern. Conf. Systemics, Cybernetics and Informatics (WMSCI 2009). Orlando, USA. 2009. P. 321–327.

DATAWARE SYSTEM FOR PROBLEMS OF APPROACHING AND DOCKING, PLANTING SPACE VEHICLE ON THE BASIS OF COMPUTER VISION

S. M. Sokolov, A. A. Boguslavsky

Keldysh Institute of Applied Mathematics Russian Academy of Sciences (KIAM RAS)

The program complex is described which intended for the automatic visual monitoring of a space vehicle approaching and docking. Initial information for the software is a video signal received from the TV-camera, mounted on the space vehicle board. An obtained sequence of frames is processed in real time. In each frame, the algorithms find out the details which sizes and geometrical parameters are known and determine parameters of the space vehicle mutual arrangement. On these measurements relative movement of space vehicles is restored. Examples of processing of the real video data

obtained during the spacecraft Progress docking to International Space Station. Possibility of use of the same complex for other information supply problems of space vehicles purposeful moving, in particular, plantings, on the basis of an optical flow calculation is discussed.

Keywords: computer vision, relative movement, real time vision system, an optical flow, space vehicles rapprochement, docking and planting.

Sokolov Sergey Mikhailovich — Leading Scientist, Doctor of Physical and Mathematical Sciences, Professor, e-mail: sokolsm@keldysh.ru.

Boguslavsky Andrey Alexandrovich — Senior Scientist, Doctor of Physical and Mathematical Sciences, Associate Professor, e-mail: boguslav@keldysh.ru.

ЗАДАЧИ БОРТОВОЙ ОБРАБОТКИ ИНФОРМАЦИИ ТЕЛЕВИЗИОННОЙ СИСТЕМЫ НАВИГАЦИИ И НАБЛЮДЕНИЯ ПРИ ПОСАДКЕ НА ФОБОС

Г. А. Аванесов, Б. С. Жуков, Е. Б. Краснопецева, А. А. Форш

Учреждение Российской академии наук Институт космических исследований РАН (ИКИ РАН), 17997, Москва, Профсоюзная, 84/32

Телевизионная система навигации и наблюдения (ТСНН) разработана для орбитальных съёмок Фобоса и Марса в навигационных и научных целях, а также для информационной поддержки посадки на Фобос. В состав ТСНН входят две узкоугольные камеры с фокусным расстоянием 500 мм и две широкоугольные камеры с фокусным расстоянием 18 мм. Процессоры камер позволяют проводить обработку получаемых изображений в реальном времени. При посадке на Фобос с помощью ТСНН будут проводиться автономный выбор места посадки, а также резервные измерения высоты и скорости КА. Программно-алгоритмическое обеспечение ТСНН в режиме «Посадка» отлажено с использованием изображений Фобоса, Луны и шероховатых поверхностей со структурой, сходной со структурой поверхности Фобоса.

Ключевые слова: Фобос, миссия ФОБОС-ГРУНТ, телевизионная система навигации и наблюдения, бортовая обработка информации

Введение

Важная роль в проекте ФОБОС-ГРУНТ, предусматривающем посадку на Фобос и доставку на Землю образцов его грунта, отводится телевизионным средствам, которые будут использоваться для решения как научных, так и навигационных задач [Аванесов и др., 2009]. В данной работе после общего описания телевизионного эксперимента детально рассматриваются задачи бортовой обработки телевизионной информации при посадке на Фобос.

1. НАУЧНЫЕ И НАВИГАЦИОННЫЕ ЗАДАЧИ ТЕЛЕВИЗИОННОГО ЭКСПЕРИМЕНТА

Научные задачи телевизионного эксперимента включают комплекс исследований характеристик Фобоса и околомарсианского пространства:

- уточнение модели фигуры, карт рельефа и геологических карт Фобоса;
- детальные исследования характеристик кратеров и борозд на поверхности Фобоса;
- исследования пространственных вариаций отражательных характеристик поверхности Фобоса и оценки по ним характеристик грунта Фобоса;
- уточнение массы и положения центра масс Фобоса;
- исследования пылевых колец Марса.

Аванесов Генрих Аронович — главный научный сотрудник, доктор технических наук, профессор, e-mail: genrikh-avanesov@yandex.ru.

Жуков Борис Сергеевич — старший научный сотрудник, кандидат физико-математических наук, e-mail: bzhukov@nserv.iki.rssi.ru.

Краснопецева Елена Борисовна — младший научный сотрудник, e-mail: lkrasnop@ofo.iki.rssi.ru.

Форш Анатолий Анатольевич — кандидат физико-математических наук, e-mail: forsh@ofo.iki.rssi.ru

По данным съёмок с КА «Маринер-9», «Викинг», «Фобос-2» и «Марс-Экспресс», поверхность Фобоса картирована с разрешением от нескольких метров до нескольких десятков метров в зависимости от района. В проекте ФОБОС-ГРУНТ могут представлять интерес съёмки поверхности Фобоса с разрешением лучше 1 м с орбиты искусственного спутника Марса (ИСМ) и до 1 см в процессе посадки. Это позволит исследовать мелкомасштабную структуру кратеров и борозд, вариации их характеристик по глубине, толщину и характеристики слоя реголита.

Снимки, полученные в широком диапазоне значений фазового угла (угла между направлениями освещения и наблюдения), дадут возможность проанализировать характеристики реголита на поверхности Фобоса: нормальное альbedo, размер частиц реголита и размер неровностей поверхности субпиксельного масштаба путём обращения моделей рассеяния излучения на шероховатых поверхностях.

Анализ динамики КА при движении вблизи Фобоса позволит уточнить массу и положение центра масс в модели фигуры Фобоса и на этом основании сделать заключение о внутренней однородности или неоднородности Фобоса.

Значительный научный интерес представляет экспериментальное подтверждение существования теоретически предсказанных пылевых колец Марса. Поскольку КА «Фобос-Грунт» будет длительное время находиться вблизи орбиты Фобоса, данные телевизионных съёмок и измерений с помощью звёздных датчиков могут дать ценную информацию для подтверждения гипотезы о существовании пылевого тора Фобоса и оценки характеристик пыли.

Полученная информация будет представлять интерес для исследования механизмов, формирующих природу Фобоса, проверки гипотез о происхождении Фобоса и сравнительно-планетных исследований.

Телевизионные изображения будут использоваться также для решения навигационных задач:

- уточнения орбиты КА при маневрировании на орбитах ИСМ;
- уточнения орбиты Фобоса;
- выбора района посадки с орбиты ИСМ;
- навигационной поддержки посадки: автономного выбора места посадки при подлёте к поверхности, измерения расстояния до поверхности Фобоса и компонент относительной скорости КА «Фобос-Грунт» в процессе снижения.

2. ТЕЛЕВИЗИОННАЯ СИСТЕМА НАВИГАЦИИ И НАБЛЮДЕНИЯ

Для решения указанных задач разработана телевизионная система навигации и наблюдения (ТСНН) [Аванесов и др., 2007]. В состав ТСНН входят две узкоугольные телевизионные камеры (УТК) высокого разрешения с фокусным расстоянием $F = 500$ мм и две обзорные широкоугольные телевизионные камеры (ШТК) с фокусным расстоянием $F = 18$ мм. Общий вид УТК и ШТК показан на рис. 1 (см. с. 159), а их характеристики приведены в табл. 1 (см. с. 159). Размещение двух комплектов УТК и ШТК по диаметру платформы посадочного модуля с базисом около 2 м обеспечивает возможность проведения стереоизмерений при посадке на поверхность Фобоса. Все камеры оснащены процессорами ADSP-21060 с тактовой частотой 24 МГц, которые использовались в составе звёздных датчиков семейства БОКЗ на ряде космических аппаратов и подтвердили свою работоспособность в условиях открытого космоса. Связь между камерами осуществляется через бортовой вычислительный комплекс (БВК).



a

б

Рис. 1. Камеры ТСНН: *a* – УТК, *б* – ШТК

Таблица 1

Основные характеристики ТСНН

Параметры	ТСНН	
	УТК	ШТК
Фокусное расстояние, мм	500	18
Относительное отверстие	1:7	1:2
Спектральные зоны, мкм	0,4...1,0	0,4...1,0
Размер элемента ПЗС, мкм	7,4	7,4
Кол-во активных элементов	1000×1000	1000×1000
Угловое разрешение, угл. с	3,05	84,8
Поле зрения, град	0,85	23,3
Радиометрическое разрешение, бит	10	10
Объём флеш-памяти, Мбайт	16	16
Масса, кг	2,8	1,6
Энергопотребление, Вт	8	8
Количество приборов	2	2

Для привязки измерений ТСНН к инерциальной системе координат используются звёздные датчики БОКЗ-МФ. Приборы БОКЗ-МФ также дублированы и имеют неперекрывающиеся направления наблюдения для облегчения реализации необходимых углов их оптической оси к Солнцу и Марсу. Каждая пара камер УТК-ШТК жестко сопряжена с соответствующим прибором БОКЗ-МФ.

Камеры ТСНН прошли стандартный комплекс приёмо-сдаточных, конструктивно-доводочных и ресурсных испытаний. Проведённая калибровка распределения коэффициента чувствительности и темнового тока по полю кадра позволяет восстанавливать по полученным изображениям поле яркости снимаемых объектов. Геометрическая калибровка камер и привязка систем координат камер ТСНН к системам координат звёздных датчиков, которая будет проверяться и уточняться в полёте по синхронным снимкам звёздного неба, позволит с высокой точностью определять направление на снимаемые объекты в инерциальной системе координат.

3. ЗАДАЧИ, РЕШАЕМЫЕ ТСНН НА ОРБИТАХ ИСМ

Войдя в сферу притяжения Марса, КА сначала начинает движение по вытянутой эллиптической орбите, а затем переходит на круговую орбиту наблюдения, высота которой примерно на 500 км превышает высоту орбиты Фобоса. После навигационных съёмок и уточнения по их результатам орбит КА и Фобоса КА переводится на квазиспутниковую орбиту (КСО), имеющую период обращения вокруг Марса, равный периоду обращения Фобоса. При движении по КСО расстояние до поверхности Фобоса будет изменяться от 30 до 70 км. Переход на указанную «финальную» КСО может осуществляться через промежуточную (промежуточные) КСО с максимальным расстоянием до поверхности Фобоса ~100 км. Разрешение и поле зрения камер ТСНН на этих орбитах приведено в табл. 2.

Таблица 2

Пространственное разрешение и поле зрения на поверхности Фобоса, получаемое с помощью УТК и ШТК на КСО и орбите наблюдения

Параметр	Расстояние до поверхности Фобоса, км			
	КСО		Орбита наблюдения	
	30	100	500	20 000
УТК:				
разрешение, м	0,45	1,5	7,5	300
поле зрения, км	0,45	1,5	7,5	300
ШТК:				
разрешение, м	12	40	200	8000
поле зрения, км	12	40	200	8000

Съёмки на орбитах ИСМ будут использоваться для геометрической калибровки камер по звёздам, решения задач околопланетной навигации и картирования Фобоса. При этом камеры УТК позволят получать изображения поверхности Фобоса с субметровым разрешением, что на 1–2 порядка лучше разрешения существующих карт Фобоса. Эти данные дадут качественно новую информацию о формах рельефа и вариациях отражательных характеристик поверхности Фобоса.

Важнейшей задачей орбитальных съёмок Фобоса является уточнение района посадки. Требования к выбору района посадки определяются тем, что при относительно малых размерах посадочного аппарата и, возможно, не полной компенсации к моменту посадки боковых составляющих скорости «безопасной» считается

площадка, не имеющая выпуклостей и впадин с размером более 0,5 м. В настоящее время район посадки предварительно выбран на основе имеющихся снимков Фобоса с разрешением ~ 10 м, что не достаточно для выявления неровностей указанного размера.

Вместе с тем, учитывая, что эллипс рассеяния при посадке равен 800×400 м (соответственно по долготе и широте) и что ожидаемое среднее расстояние между кратерами, которые могут представлять опасность для посадки, составляет ~ 40 м, нельзя ожидать, что в пределах выбранного района посадки (размеры которого должны соответствовать эллипсу рассеяния) будут отсутствовать опасные неровности. Можно лишь надеяться по полученным с орбиты снимкам выбрать для посадки район с минимальным количеством таких неровностей, а задачу выбора конкретного места посадки решать автономно при спуске на поверхность Фобоса.

4. ЗАДАЧИ ТСНН ПРИ ПОСАДКЕ НА ФОБОС

При посадке предусматривается съёмка Фобоса всеми четырьмя камерами ТСНН.

С помощью двух камер ШТК будут решаться следующие задачи:

- формирование сжатых кадров репортажной съёмки, которые будут передаваться на Землю в реальном времени в процессе посадки;
- построение карты пригодности поверхности для посадки, которая будет использоваться для автоматического выбора места посадки;
- измерения расстояния до поверхности Фобоса;
- измерения смещения опорного объекта на последовательных изображениях, которые с учётом изменения масштаба изображения и ориентации КА позволят оценить боковую скорость КА.

Выполняемые с помощью ШТК измерения высоты и скорости КА будут резервировать аналогичные измерения, выполняемые штатными средствами — лазерным высотомером и доплеровским измерителем скорости и дальности.

Хотя измерения ШТК будут проводиться непрерывно начиная с момента схода с орбиты, наиболее важными они являются при приближении к поверхности на высотах ниже ~ 300 м для контроля высоты, скорости соприкосновения с поверхностью и автономного выбора места посадки.

Разрешение и поле зрения камер ТСНН при посадке приведены в табл. 3 (см. с. 162) с учётом эффекта дефокусировки. Смазом изображений из-за бокового смещения КА относительно поверхности Фобоса можно пренебречь. Для УТК ухудшение разрешения из-за дефокусировки в два раза происходит на высоте 5 км, а на высотах ниже 1 км разрешение УТК остается практически постоянным. Для ШТК ухудшение разрешения из-за дефокусировки более чем в два раза происходит на высотах ниже 20 м, а предельное разрешение составляет ~ 1 см.

Сжатый кадр репортажной съёмки будет передаваться из каждой камеры в БВК за 20 циклов. Всего каждой камерой за время спуска будет получено до 20 репортажных кадров.

После окончания торможения КА на высоте 300...200 м, когда разрешение ШТК достигнет 13...9 см, по снимкам ШТК должен производиться автономный выбор места посадки. С этой целью по данным ШТК будет строиться и передаваться в БВК карта пригодности поверхности для посадки, по которой можно определить пригодность прогнозируемого места посадки, а в случае его непригодности — выбрать наиболее ровный участок заданного размера, на который может

быть проведена посадка КА. Карта пригодности строится по признаку фотометрической однородности изображения с дополнительным учётом наличия теней и близости крупных неоднородностей.

Таблица 3

Изменение разрешения и поля зрения камер ТСНН в процессе посадки

Расстояние до поверхности Фобоса	ШТК ($f=18$ мм)		УТК ($f=500$ мм)	
	разрешение	поле зрения	разрешение, см	поле зрения, м
30 км	12 м	12 км	52	450
10 км	4,1 м	4,1 км	22	150
3 км	1,2 м	1,2 км	12	45
1 км	42 см	410 м	8,6	15
300 м	13 см	120 м	7,6	4,5
100 м	5,0 см	41 м	7,3	1,5
30 м	2,1 см	12 м	–	–
10 м	1,3 см	4,1 м	–	–
3 м	1,0 см	1,2 м	–	–
1 м	0,9 см	41 см	–	–

Расстояние до поверхности Фобоса будет определяться по изображениям, полученным камерами ШТК на стереобазе около 2 м. Ошибки измерений высоты камерами ШТК, оценённые с учётом разрешения камер и ошибок их геометрической калибровки, приведены в табл. 4. Относительные ошибки измерения высоты убывают от 7 до 0,7 % при уменьшении высоты от 300 до 10 м. Для экономии времени стереоизмерения первоначально проводятся по четырёхкратно заглубленному изображению, а затем уточняются по исходному изображению в соответствии с алгоритмом [Гришин, 2009]. Максимальное число используемых для стереоизмерений опорных точек равно 10. Измерения высоты, полученные по этим точкам, фильтруются, а затем определяется их среднее значение.

Таблица 4

Ошибки измерения высоты и горизонтальной скорости по данным ШТК

Расстояние до поверхности Фобоса	Ошибка измерения высоты, м	Ошибка оценки горизонтальной скорости, м/с
30 км	–	–
10 км	–	–
3 км	1900	0,82
1 км	210	0,28
300 м	20	0,084
200 м	9,1	0,056
100 м	2,5	0,029
30 м	0,32	0,0097
10 м	0,066	0,0042

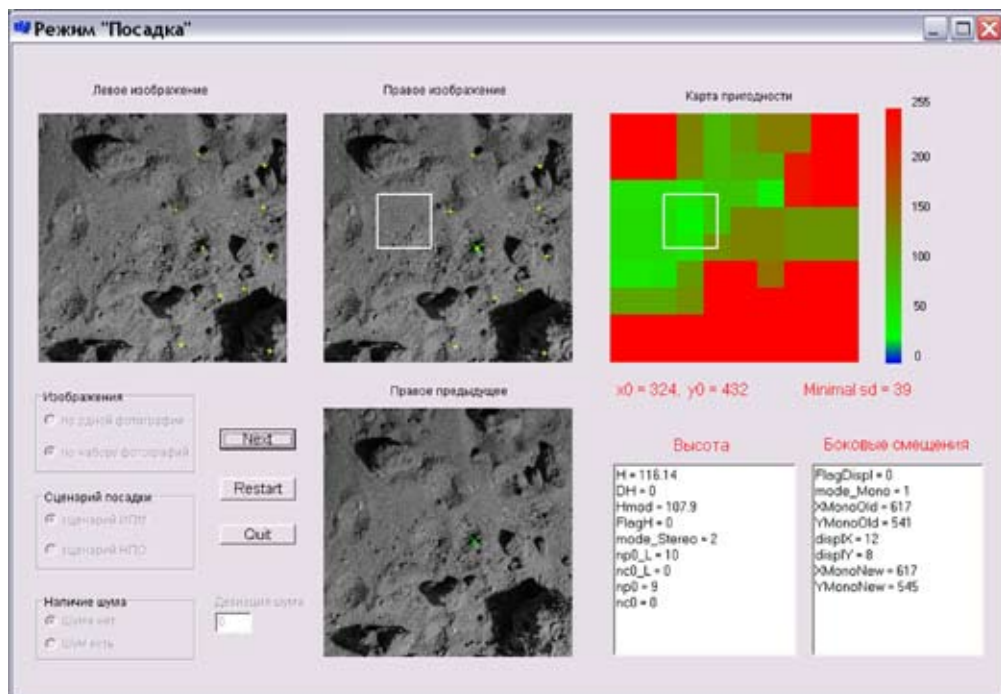


Рис. 2. Визуализация результатов работы ТСНН в режиме «Посадка»

Информацию для оценки горизонтальных компонент скорости КА дадут измерения смещения опорной точки на последовательных изображениях, полученных одной и той же камерой ШТК. Данные этих измерений будут обрабатываться в БВК с учётом изменения ориентации КА и высоты съёмки. Поскольку смещения объектов очень чувствительны к горизонтальной скорости, для экономии времени измерения бокового смещения первоначально проводятся по шестнадцатикратно загрубленному изображению, а затем уточняются по четырёхкратно загрубленному изображению в соответствии с алгоритмом [Гришин, 2009]. При этом обеспечивается достаточная точность измерения горизонтальной скорости (см. табл. 4). Ограничения по времени заставляют ограничиться определением бокового смещения лишь по одной опорной точке.

С накоплением результатов измерений ТСНН в БВК ошибки траекторных измерений будут существенно ниже, чем рассмотренные ошибки единичных измерений ТСНН.

Цикл съёмки и обработки информации в режиме «Посадка» составляет около 6 с.

Съёмки с помощью камер УТК будут использоваться для стереосъёмки поверхности Фобоса и записи полных изображений во флеш-память камер для передачи на Землю после посадки. Всего во флеш-память каждой камеры УТК будет записано по 12 полных изображений.

Отработка алгоритмических и программных решений информационной поддержки посадки на Фобос проводилась с использованием изображений поверхности Фобоса и Луны, а также стереоизображений модельных шероховатых

поверхностей [Жуков и др., 2009]. Для управления в наземных условиях реальной бортовой программой режима «Посадка» и визуализации получаемых результатов использовалась управляющая программа, интерфейс которой показан на рис. 2 (см. с. 163). Интерфейс отображает (слева направо и сверху вниз):

- текущий «левый» кадр стереопары с указанием положения выбранных опорных точек для стереоизмерения высоты;
- текущий «правый» кадр стереопары, с указанием положения: опорных точек для стереоизмерения высоты, для которых установлено соответствие с «левым» кадром; опорной точки для определения бокового смещения, для которой установлено соответствие с предыдущим «правым» кадром; оптимальной площадки для посадки;
- карту пригодности с указанием положения оптимальной площадки для посадки;
- управляющие элементы для выбора последовательности модельных изображений;
- предыдущий «правый» кадр (под текущим «правым» кадром) с указанием положения выбранной опорной точки для определения бокового смещения;
- окна с результатами измерения высоты и бокового смещения.

Устойчивость работы программы «Посадка» при различных сценариях посадки на Фобос обсуждается в статье настоящего сборника (см. с. 166): *Гришин В. А.* «Анализ влияния динамики космического аппарата на характеристики алгоритмов обработки изображений системы технического зрения проекта ФОБОС-ГРУНТ».

Заключение

Измерения высоты и скорости КА и автономный выбор места посадки КА «Фобос-Грунт» требуют бортовой обработки в реальном времени информации, получаемой телевизионной системой навигации и наблюдения ТСНН. Разработано и отлажено программно-алгоритмическое обеспечение ТСНН в режиме «Посадка», обеспечивающее эффективное решение этих задач в условиях ограничений на производительность бортовых процессоров и скорость каналов информационного обмена.

Литература

- Аванесов Г. А., Жуков Б. С., Краснопецева Е. Б.* (2007) Телевизионные средства навигации и наблюдения проекта ФОБОС-ГРУНТ // Космич. приборостроение. Координатно-временные системы с использованием космических технологий. Приборы для космических исследований Земли и планет: Сб. тр. по результатам выездного семинара / Под ред. Р. Р. Назирова. Россия, Таруса, 7–9 июня 2006. М.: ИКИ РАН, 2007. С. 226–231.
- Аванесов Г. А., Жуков Б. С., Краснопецева Е. Б.* (2009) Задачи, решаемые телевизионной системой навигации и наблюдения в проекте ФОБОС-ГРУНТ // Всерос. научно-технич. конф. «Современные проблемы определения ориентации и навигации космических аппаратов»: Сб. тр. Сер. «Механика, управление и информатика». Россия, Таруса, 22–25 сент. 2008. М.: ИКИ РАН, 2009. С. 239–250.
- Гришин В. А.* (2009) Алгоритмы измерения высоты и компонент скорости по телевизионным изображениям при посадке на Фобос // Всерос. научно-техн. конф. «Современные проблемы определения ориентации и навигации космических аппаратов»: Сб. тр. Сер. «Ме-

ханика, управление и информатика». Россия, Таруса, 22–25 сент. 2008. М.: ИКИ РАН, 2009. С. 279–293.

Жуков Б. С., Дунаев Б. С., Жуков С. Б. (2009) Моделирование детальных изображений поверхности Фобоса для отработки задач информационной поддержки посадки на Фобос // Всерос. научно-техн. конф. «Современные проблемы определения ориентации и навигации космических аппаратов»: Сб. тр. Сер. «Механика, управление и информатика». Россия, Таруса, 22–25 сент. 2008. М.: ИКИ РАН, 2009. С. 304–313.

TELEVISION SYSTEM OF NAVIGATION AND OBSERVATION: OBJECTIVES OF ON-BOARD DATA PROCESSING DURING LANDING ON PHOBOS

G. A. Avanesov, B. S. Zhukov, E. B. Krasnopevtseva, A. A. Forsh

*Space Research Institute (IKI RAN),
117997, 84/32 Profsoyuznaya st, Moscow, Russia*

Television System of Navigation and Observation (TSNN) was developed for scientific and navigational imaging of Phobos and Mars, as well as to support landing on the surface of Phobos. TSNN includes two narrow-angle cameras with a focal length of 500 mm and two wide-angle cameras with a focal length of 18 mm. The camera processors allow real-time image processing. TSNN will be used during the landing for autonomous landing site selection as well as for back-up measurements of s/c altitude and relative velocity. The TSNN landing software was tested using images of Phobos, Moon and of rough surfaces with a structure similar to Phobos surface.

Keywords: Phobos, Phobos-Grunt mission, television system of navigation and observation, on-board data processing.

Avanesov Henry Aronovich — Leading Scientist, Doctor of Technical Science, Professor, e-mail: genrikh-avanesov@yandex.ru.

Zhukov Boris Sergeevich — Senior Scientist, Candidate of Physical and Mathematical Sciences, e-mail: bzhukov@nserv.iki.rssi.ru.

Krasnopevtseva Elena Borisovna — Junior Researcher, e-mail: lkrasnop@of.iki.rssi.ru.

Forsh Anatoly Anatolevich — Candidate of Physical and Mathematical Sciences, e-mail: forsh@of.iki.rssi.ru.

АНАЛИЗ ВЛИЯНИЯ ДИНАМИКИ КОСМИЧЕСКОГО АППАРАТА НА ХАРАКТЕРИСТИКИ АЛГОРИТМОВ ОБРАБОТКИ ИЗОБРАЖЕНИЙ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ ПРОЕКТА ФОБОС-ГРУНТ

В. А. Гришин

Учреждение Российской академии наук Институт космических исследований РАН (ИКИ РАН), 117997, Москва, Профсоюзная 84/32

В статье рассматриваются методика и результаты отработки алгоритмов слежения и прогноза системы технического зрения для проекта ФОБОС-ГРУНТ. Отработка производилась с использованием расчётных баллистических данных процесса посадки и модели поверхности Фобоса.

Ключевые слова: посадка космических аппаратов, системы технического зрения, динамика космического аппарата, алгоритмы измерения и прогноза.

Введение

Проект ФОБОС-ГРУНТ предполагает полёт до окрестности планеты Марс, детальную съёмку поверхности его спутника — Фобоса (рис. 1), уточнение места посадки, выполнение посадки, взятие проб с поверхности Фобоса, проведение ряда исследований, старт и возвращение полученных проб на Землю.

В состав бортового комплекса аппаратуры входит телевизионная система навигации и наблюдения (ТСНН). В функции ТСНН входит, во-первых, проведение наблюдений и съёмка поверхности Марса и его спутника Фобоса с высоким разрешением, во-вторых, — ведение съёмки поверхности Фобоса в процессе выполнения посадки на его поверхность [Zhukov et al., 2007].

Помимо указанных задач ТСНН обеспечивает информационную поддержку процесса посадки. Производится измерение средней дальности (высоты) до поверхности в стереорежиме, измерение перемещений точек поверхности в поле зрения камер в монокулярном режиме. Измерения средней дальности в стереорежиме дублируют измерения лазерного высоотомера. Данные о перемещениях точек изображения поверхности в монокулярном режиме в сочетании с информацией, поступающей от гироскопов системы инерциальной навигации, используются для измерения компонент скорости космического аппарата (КА) и дублируют работу доплеровского измерителя скорости.

В состав ТСНН входят две узкоугольные камеры (УТК), предназначенные для наблюдения поверхности Марса и Фобоса с большого расстояния, и две широкоугольные камеры (ШТК), предназначенные для наблюдения поверхности со средних и малых расстояний.



Рис. 1. Изображение Фобоса

В состав каждой камеры входит цифровой процессор обработки сигналов и флеш-память для хранения изображений высокого разрешения. Для информационного обеспечения процесса управления посадкой КА используются широкоугольные камеры. Система технического зрения построена по симметричной схеме, когда обе камеры имеют одинаковое программное обеспечение. В монокулярном режиме каждая из камер работает независимо от другой камеры. В стереоскопическом режиме каждая камера реализует поиск особенностей (координаты и образы которых передаются во вторую камеру), установление соответствия и измерение средней дальности. Для уменьшения вычислительных затрат измерения в стереоскопическом и монокулярном режимах построены по двухканальной схеме. Первоначально установление соответствия происходит в канале низкого разрешения. После захвата на автосопровождение в канале низкого разрешения включается канал высокого разрешения.

Следует отметить, что большое разнообразие и ресурсоемкость параллельно решаемых задач вынуждает использовать весьма сложные циклограммы работы ТСНН, которые изменяются в процессе сближения с Фобосом. Ограниченная вычислительная мощность процессоров обуславливает необходимость тщательного анализа влияния динамики КА на функционирование системы технического зрения (СТЗ).

ВЛИЯНИЕ ДИНАМИКИ КА НА ПРОЦЕСС ИЗМЕРЕНИЙ

Рассматривая СТЗ как измерительную систему, естественно определить показатели качества работы такой системы: как динамические ошибки измерений, так и случайные ошибки. Случайные ошибки измерений порождаются, во-первых, нормальными ошибками установления соответствия, во-вторых — грубыми (аномальными) ошибками этого процесса. Важно также максимально снизить вероятность сбоев слежения за дальностью и скоростью, поскольку после возникновения такого сбоя инициируется процесс поиска по дальности (скорости). Только после перехода системы технического зрения в режим слежения будет выдаваться достоверная информация. Таким образом, каждый сбой вызывает потерю информации. Исключительно важное значение имеет вычислительная сложность алгоритмов. Факторы, влияющие на ошибки измерений и вычислительную сложность, показаны на рис. 2 (см. с. 168).

Динамика КА, а именно его угловое и линейное движение, вызывает соответствующее перемещение объектов и поверхностей, находящихся в поле зрения камер. Поэтому она оказывает непосредственное влияние на измеряемые параметры. Кроме того, работа алгоритмов оценивания, фильтрации, обнаружения сбоев слежения, переключения в режим поиска, обнаружения и захвата существенно учитывает законы распределения измеряемых случайных величин. А они также зависят от динамики КА. Например, закон распределения дальности (высоты) по полю кадра зависит от угла наклона оптической оси к визируемой поверхности. Для повышения устойчивости функционирования систем технического зрения целесообразно использовать (при наличии такой возможности) информацию от гироскопов и акселерометров.

В состав систем технического зрения входят различные алгоритмы слежения и прогноза [Гришин, 2009; Жуков и др., 2009]. В данном случае это алгоритмы слежения за средней дальностью в стереорежиме и алгоритмы слежения за средней скоростью перемещения точек поверхности в поле зрения монокулярного измерителя.

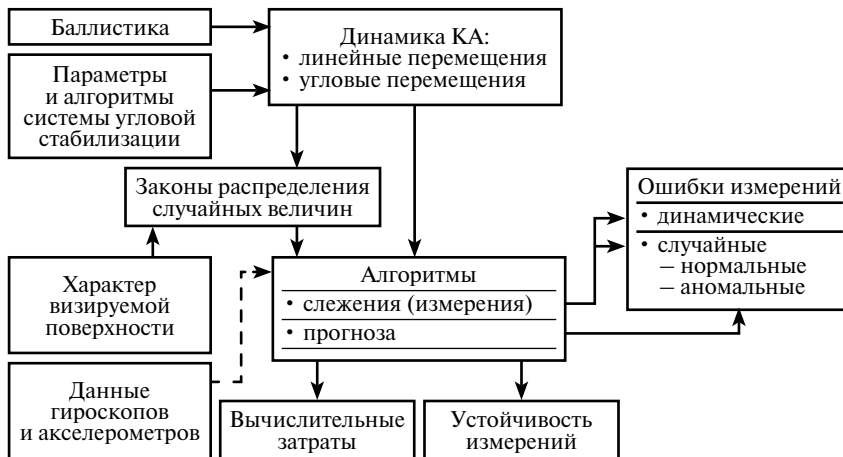


Рис. 2. Факторы, влияющие на ошибки измерений и вычислительную сложность

Указанные алгоритмы слежения (прогноза) обеспечивают, с одной стороны, уменьшение вероятности аномальных ошибок измерений за счёт уменьшения размеров области поиска соответствия, с другой стороны — уменьшение вычислительной сложности алгоритмов, которая прямо пропорциональна длине отрезка, вдоль которого производится поиск в процессе установления соответствия. Однако при этом растёт вероятность того, что резкое изменение дальности, обусловленное, например, быстрым разворотом КА, вызовет сбой слежения, инициирование процесса поиска по дальности, во время выполнения которого система технического зрения не будет выполнять измерений.

Для монокулярного режима используется слежение за средней скоростью перемещения точек в поле зрения и её прогнозирование на следующем такте измерения. Данный прогноз особенно важен в связи с тем, что в отличие от стереоскопического режима, где поиск и установление соответствия производится вдоль отрезков, в монокулярном режиме поиск и установление соответствия производится в некоторой области изображения, что требует существенно больших вычислительных затрат. Размер этой области целиком определяется степенью предсказуемости (прогнозируемости) перемещения КА за время цикла обработки информации.

Существует оптимальное значение прогноза размера области поиска с точки зрения минимизации вероятности аномальных ошибок. При превышении этого размера наблюдается медленный рост вероятности аномальных ошибок установления соответствия и рост вычислительных затрат. При размере области поиска меньше оптимального наблюдается быстрый рост вероятности аномальных ошибок установления соответствия, поскольку растёт вероятность того, что искомая точка окажется вне области поиска.

В статье представлены методика и результаты моделирования работы алгоритмов технического зрения в части выполнения алгоритмов прогноза положения и размера области поиска. Проверялось также взаимодействие между каналами низкого и высокого разрешения.

Моделирование производилось, в основном, по имеющейся баллистической информации, полученной от ИПМ им. М. В. Келдыша. Приведены также некото-

рые результаты моделирования, выполненные на основе баллистической информации о сценарии посадки, разработанном в НПО им. С. А. Лавочкина.

Первым рассмотрим сценарий посадки, разработанный в НПО им. С. А. Лавочкина («быстрый» вариант посадки). Сценарий предусматривает сход с квази-синхронной орбиты (КСО), снижение в течение 32,4 мин до высоты 691 м. При этом расчётное значение вертикальной скорости равно 13,8 м/с, а боковой — 2,2 м/с. После выполнения торможения скорость гасится практически до нуля, а сам КА оказывается на высоте 200 м над поверхностью Фобоса. Затем КА осуществляет свободное падение на поверхность Фобоса в течение 272 с. Расчётное значение вертикальной скорости в момент касания равно 1,44 м/с, а горизонтальной — 0,34 м/с.

Полёт осуществляется в режиме инерциальной стабилизации. Такой режим наиболее благоприятен для работы системы технического зрения. Здесь необходимо отметить два наиболее критичных участка. Во-первых, это высота 691 м, когда вертикальная скорость достигает максимума. Это обуславливает изменение масштаба изображения каждого последующего кадра и затрудняет установление соответствия в монокулярном режиме (коэффициент масштабирования 1,12). Второй критичный участок находится непосредственно перед касанием поверхности Фобоса. На высоте 50 м коэффициент масштабирования достигает 1,147. Ниже он ещё более растёт. Большой коэффициент масштабирования затрудняет процесс установления соответствия и повышает вероятность аномальных ошибок.

Вместе с тем следует отметить, что наличие участка свободного падения с высоты 200 м создает благоприятные условия для проведения измерений высоты и скорости перемещения КА относительно поверхности Фобоса, а также для автоматического выбора наиболее подходящего места посадки.

Рассмотрим сценарии посадки, разработанные в ИПМ им. М. В. Келдыша [Аким и др., 2009]. Сценарии предусматривают сход с КСО, перелёт в точку, расположенную над районом посадки (прицельную точку). Высота этой точки не превышает 10 км. Длительность перелёта — около 30 мин. Затем производится вертикальный спуск, который происходит на большой скорости для уменьшения времени выполнения и, соответственно, уменьшения накопления ошибки положения КА в горизонтальной плоскости. На высоте примерно 1000 м начинается прецизионное торможение. Вертикальная скорость гасится до величины, не превосходящей 1,5...2 м/с, а боковая скорость — до величины, не превосходящей 1 м/с.

В процессе посадки происходит совмещение средней нормали к поверхности с осью OX аппарата. Это сделано для того, чтобы избежать больших разворотов и сбоя в работе систем лазерного высотомера, доплеровского измерителя скорости и дальности, привода остронаправленной антенны. Необходимо отметить, что такой режим весьма неблагоприятен для работы системы технического зрения в монокулярном режиме, поскольку порождает интенсивные угловые колебания КА и как следствие — большие скорости перемещения точек визируемой поверхности в поле зрения камер. Колебания столь велики, что за время, равное циклу обработки информации в ТСНН, точки могут выходить за пределы поля зрения камеры, что неизбежно порождает аномальные ошибки измерений в монокулярном режиме.

Для стереоскопического режима интенсивные угловые колебания порождают ошибки измерения дальности из-за асинхронности моментов съёмки камер. Однако расчёты показывают, что эти ошибки невелики при величине асинхронности в 5 мс, и ими можно пренебречь. Для стереорежима критичными являются последние 10...20 м при подлёте к поверхности Фобоса. Проблема заключается во

всё возрастающей скорости роста диспаратности при подлёте к визируемой поверхности.

В ИПМ им. М. В. Келдыша разработано два сценария посадки:

- вариант схода с КСО с упреждением прохождения траверза точки посадки;
- вариант схода с КСО при прохождении траверза точки посадки.

В обоих случаях баллистическая информация, использованная для моделирования, включает в себя данные о времени (с шагом 0,1 с), текущие координаты КА в системе координат, связанной с точкой посадки, матрицу перехода из этой системы координат в связанную систему координат КА. Координаты точки посадки равны 5° широты и 235° долготы (рис. 3, 4). Центр системы координат, связанной с точкой посадки, помещён в номинальную точку посадки. Ось OZ направлена по нормали к поверхности, ось OX — по касательной к линии постоянной широты, ось OY дополняет систему координат до правой. Направления осей строятся по модели Фобоса — трёхосному эллипсоиду с осями, равными 13,0; 11,4 и 9,2 км (см. рис. 4). Для определения пространственных координат точек поверхности (в том числе координат точки посадки) используется модель Фобоса Института геохимии и аналитической химии им. В. И. Вернадского (ГЕОХИ). Модель задает профиль поверхности (длину радиус-вектора) для сетки углов широты и долготы с шагом 2° .

Для этих сценариев критичные участки, обусловленные масштабированием изображения, имеют непрерывный характер до самой поверхности Фобоса. Для варианта схода с КСО с упреждением прохождения траверза точки посадки коэффициент масштабирования изображения превышает 1,146 для высот ниже 654 м, а для варианта схода с КСО при прохождении траверза точки посадки коэффициент масштабирования превышает 1,148 для высот, меньших 624 м. Большой коэффициент масштабирования затрудняет процесс установления соответствия и повышает вероятность аномальных ошибок.

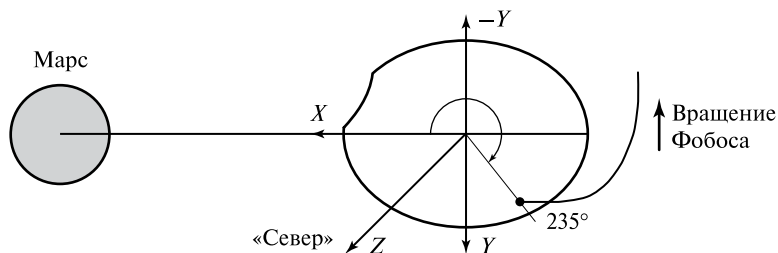


Рис. 3. Прямоугольная система координат

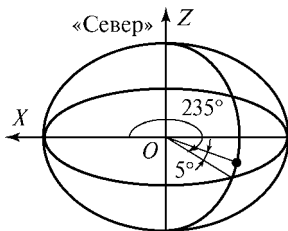


Рис. 4. Аппроксимирующей поверхность Фобоса эллипсоид

В процессе анализа динамики работы алгоритмов слежения и прогноза используется величина расстояния от камер до точки пересечения оптической оси камеры с поверхностью Фобоса. Для решения этой задачи используется аппроксимация поверхности Фобоса моделью ГЕОХИ. Дальность до поверхности и пространственные координаты точки пересечения оптической оси камеры с поверхностью получаются путём итерационного процесса. Процесс включает в себя поиск точек пересечения оптической оси со сферами: вписанной в поверхность Фобоса и описанной вокруг Фобоса. Для этих точек определяются соответствующие значения долготы и широты. По этим двум парам координат, согласно модели ГЕОХИ, определяются радиус-векторы, которые используются для построения вписанной и описанной сфер. Процесс сходится обычно за 2–4 итерации с точностью 0,1 м. Если расстояние от КА до центра модели Фобоса становится меньше радиуса описанной сферы, то радиус этой сферы устанавливается равным расстоянию от КА до центра модели Фобоса.

СТЕРОИЗМЕРЕНИЯ

Рассмотрим методику отработки алгоритмов стереорежима. Для стереорежима в качестве входной информации используется дальность по оптической оси с учётом текущего положения КА и его ориентации в пространстве. Данная информация применяется для моделирования процесса измерений, прогноза дальности на следующий цикл измерений, расчёта границ поиска соответствия, прогноза границ на следующий цикл и на случай сбоя слежения.

Следует отметить, что при большой длительности цикла обработки информации (6 с) в прогноз, особенно в прогноз на несколько циклов измерения в случае сбоя измерений, для повышения точности нужно закладывать величину ускорения, которая обусловлена работой двигателя и существенно меняется в зависимости от использованного сценария. Поскольку акселерометр в состав ТСНН не входит, то неопределённость ускорения вызывает отличие прогнозируемого движения от реального движения КА и ухудшает возможность восстановления слежения за дальностью после сбоя, особенно на последних десятках метров полёта до поверхности Фобоса.

По сценарию НПО им. С.А. Лавочкина торможение заканчивается на высоте 200 м. Этой высоте соответствует диспаратность в 6 пикс. (изображения 250×250 канала низкого разрешения). Можно устанавливать отрезок поиска максимум в 40 пикс., что соответствует дальности в 30 м. Таким образом, границу использования скорости для формирования прогноза нужно выбирать между этими двумя значениями. Была выбрана граница, равная 25 пикс. (высота равна 48,6 м) для использования оценки скорости. Следует также отметить, что оценка скорости, получаемая конечно-разностными методами, сильно флуктуирует на высотах, больших 50 м.

Рассмотрим сценарий посадки, разработанный в ИПМ им. М. В. Келдыша — вариант схода с КСО с упреждением прохождения траверза точки посадки. На графике (рис. 5, см. с. 172) показаны зависимости расстояния (высоты h) до поверхности Фобоса от времени, а также две оценки этого расстояния, полученные в каналах низкого и высокого разрешения. Заметен ступенчатый характер оценок, который обусловлен дискретным характером измеряемой диспаратности.

На рис. 6 (см. с. 172) представлена зависимость диспаратности от времени на изображении 250×250 пикс. (канал низкого разрешения).

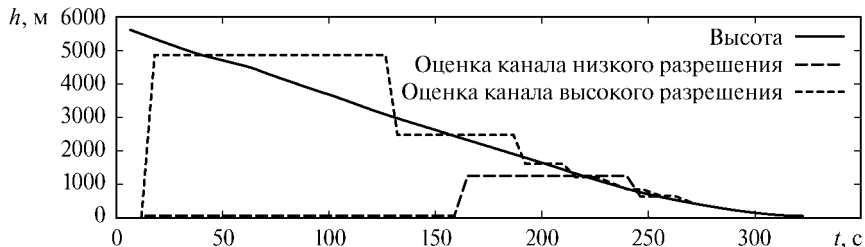


Рис. 5. Расстояние до поверхности Фобоса и его оценка в каналах низкого и высокого разрешения

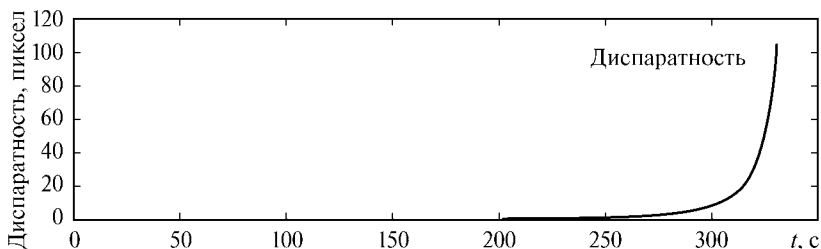


Рис. 6. Зависимость диспаратности от времени

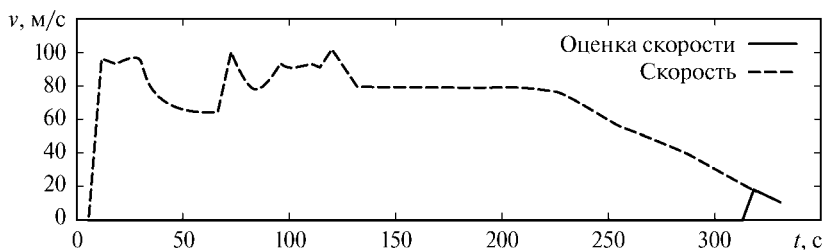


Рис. 7. Вертикальная скорость снижения и её оценка на последних 48,6 м

На рис. 7 представлена зависимость вертикальной скорости снижения КА от времени и её оценка на последних 48,6 м (что соответствует диспаратности 25 пикс. изображения 250×250).

На рис. 8 (см. с. 173) показаны прогнозируемые значения минимальной и максимальной диспаратности в режиме слежения и границы области поиска, рассчитанные на случай сбоя слежения. Указанные величины показаны относительно точного значения диспаратности (горизонтальная линия). Следует отметить, что на последних секундах полёта существенно возрастают ошибки прогноза, обусловленные заметным ускорением (торможением), которое порождает работающий двигатель.

Реально области поиска особенностей расширяются по сравнению с прогнозом для получения запаса устойчивости слежения при маневрировании КА (рис. 9, см. с. 173).

Переключение режимов (поиск, захват, слежение) происходит по количеству аномальных ошибок измерения дальности в группе из 10 измерений. Графики по сценарию — варианту схода с КСО при прохождении траверза точки посадки — не имеют существенных отличий и здесь не приводятся.

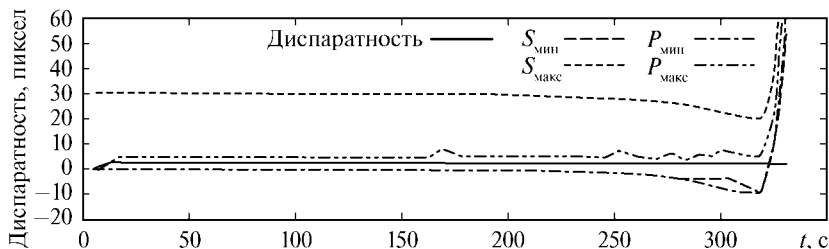


Рис. 8. Прогнозируемые границы диспаратностей P_{\min} и P_{\max} , а также границы поиска по диспаратности S_{\min} и S_{\max} в случае сбоя слежения

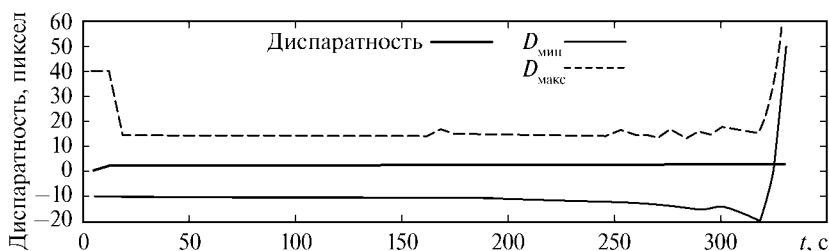


Рис. 9. Расширенные границы области поиска соответствия D_{\min} и D_{\max}

Анализировалось влияние разброса диспаратности в поле зрения (только в канале низкого разрешения). Для этого к каждому значению диспаратности на стереопаре добавлялась случайная величина с нормальным распределением и средним квадратическим отклонением, равным 0,1 от текущей диспаратности. Эта случайная величина моделирует разброс, обусловленный разбросом высот в пределах кадра в точках, которые используются для привязки. Следует отметить, что такой разброс является достаточно серьезным испытанием для алгоритмов обработки информации в стереорежиме, так как величина разброса не коррелирована на соседних кадрах, что может порождать заметные флуктуации. В реальных условиях влияние разброса дальностей будет менее существенным.

На рис. 10–12 (см. с. 174) приведены графики при разбросе высот h , СКО которого равно 0,1 от величины диспаратности для варианта схода с КСО с упреждением прохождения траверза точки посадки.

Из графиков следует, что при среднеквадратическом разбросе в 0,1 диспаратности происходит сглаживание «ступенек» оценок дальности на высотах 1000 м и ниже, т. е. уменьшаются ошибки измерения высоты. Это обусловлено эффектом стохастической линеаризации, описанным в литературе. При величине разброса в 0,1 диспаратности происходит искажение границ оценок области поиска соответствия. Однако работоспособность алгоритмов сохраняется. Графики по второму сценарию — варианту схода с КСО при прохождении траверза точки посадки — не имеют существенных отличий и здесь не приводятся.

Оценим, какому разбросу высот соответствуют эти разбросы диспаратности. Допустим, диспаратность равна 25 пиксел., при этом высота равна 48,6 м. Разброс по диспаратности 0,1 даёт разброс по дальности 44,2...54 м на площадке размером 20×10 м. Таким образом, это весьма значительная изрезанность поверхности,

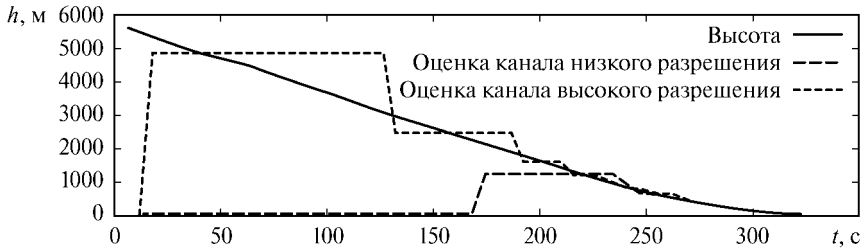


Рис. 10. Высота и её оценки в каналах низкого и высокого разрешения при разбросе измеренных высот

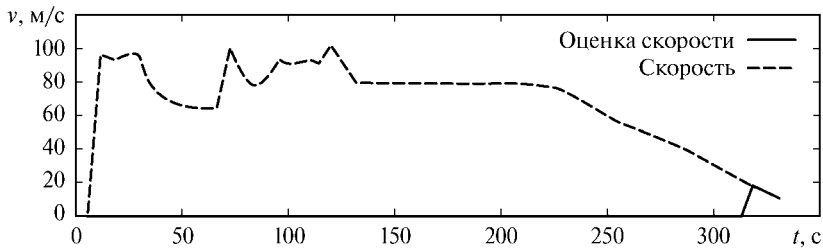


Рис. 11. Вертикальная скорость снижения и её оценка на последних 48,6 м при разбросе измеренных высот

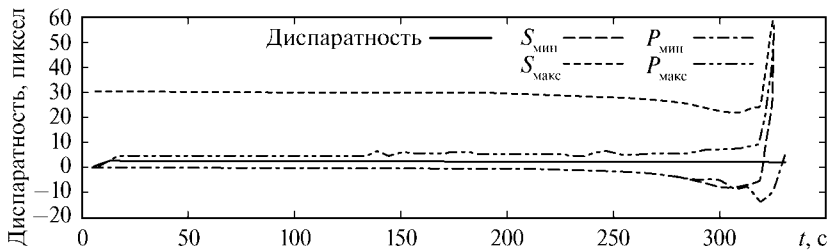


Рис. 12. Прогнозируемые границы диспаратностей P_{\min} и P_{\max} , а также границы поиска по диспаратности S_{\min} и S_{\max} в случае сбоя слежения при разбросе измеренных высот

которая маловероятна. Как уже отмечалось, в реальных условиях результаты измерений на последовательных циклах будут коррелированы, и величина флуктуаций будет меньше. По результатам моделирования была произведена оптимизация параметров алгоритма прогноза области поиска для установления соответствия в режиме слежения и прогноза области поиска соответствия в режиме сбоя (на несколько циклов измерения). Алгоритмы довольно сложны, поэтому они здесь не приводятся. Алгоритмы учитывают следующие величины:

- текущие оценки дальности;
- дисперсию измерений дальности;
- дискретность измерения диспаратности;
- флуктуации оценок скорости;
- необходимость коррекции приращений на малых дальностях для уменьшения динамических ошибок прогноза.

МОНОКУЛЯРНЫЙ РЕЖИМ

Рассмотрим методику отработки алгоритмов монокулярного режима. В этом режиме фиксируется точка на i -м кадре и определяется её положение на $i + 1$ кадре. При моделировании производился расчёт координат визируемой точки с учётом линейных и угловых перемещений камеры. Далее моделировался процесс измерений, прогнозирования, определения сбоев слежения и т. д.

Рассмотрим сценарий посадки, разработанный в ИПМ им. М. В. Келдыша — вариант схода с КСО с упреждением прохождения траверза точки посадки. Результаты приведены на графиках (рис. 13–17). На рис. 13 представлены высота над поверхностью и наклонная дальность до поверхности в метрах как функции времени; на рис. 14 — величины изменения координат и уменьшения наклонной дальности за время цикла обработки измерения в метрах; на рис. 15 — величины изменения углов ориентации за время цикла измерений в градусах, а также коэффициент масштабирования изображения; на рис. 16 (см. с. 176) — изменение

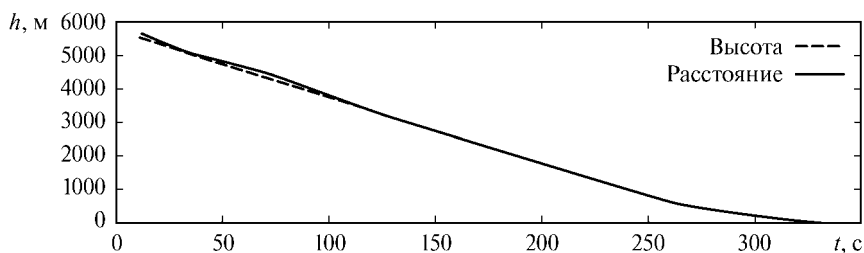


Рис. 13. Зависимость высоты и расстояния до поверхности от времени

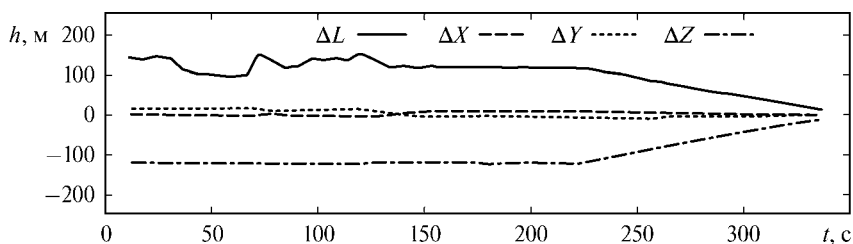


Рис. 14. Зависимость изменения координат и расстояния до поверхности за цикл обработки

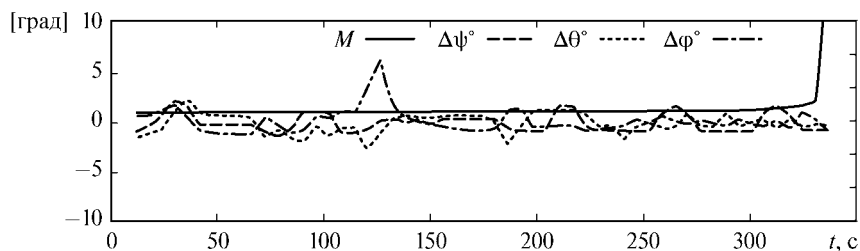


Рис. 15. Зависимость изменения углов ориентации за время цикла измерений (в град.), а также коэффициент масштабирования изображения M

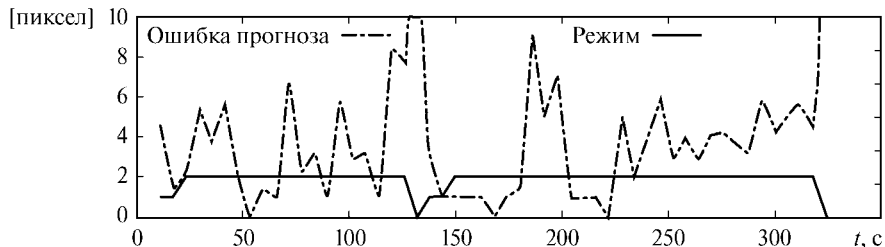


Рис. 16. Изменение величины ошибки прогноза вектора перемещения точки в поле зрения и режимов работы алгоритма

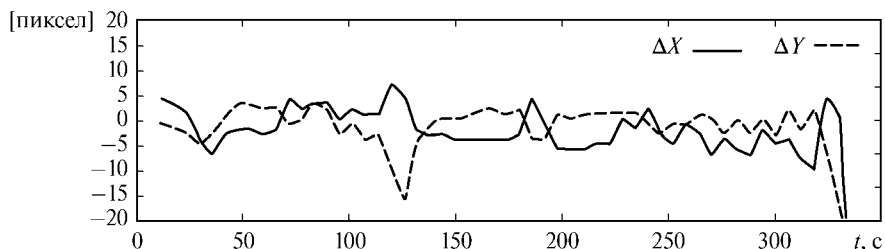


Рис. 17. Изменение компонент векторов перемещения точки в поле зрения на изображении 62×62 пикс. (канал низкого разрешения)

величины ошибки прогноза вектора перемещения точки в поле зрения и график режимов работы алгоритма (поиск, захват, слежение).

Из данных рис. 16 следует, что приблизительно на 125-й с происходил сбой слежения, ошибки прогноза имеют значительную величину. Переключение режимов (поиск, захват, слежение) происходило по результатам анализа величины ошибки прогноза. Аналогичные графики по сценарию посадки, разработанному в НПО им. С. А. Лавочкина, показывают гораздо более благоприятную картину.

Из графиков (рис. 17) следует, что в случае сценария ИПМ имеют место значительные выбросы вектора перемещения, что чревато сбоем процесса измерений. Графики по сценарию схода с КСО при прохождении траверса точки посадки не имеют существенных отличий и здесь не приводятся.

На важнейшие характеристики алгоритмов обработки информации в монокулярном режиме оказывает влияние характер прогноза положения области поиска соответствия. Была проведена оценка эффективности прогноза для двух сценариев посадки — ИПМ им. М. В. Келдыша и НПО им. С. А. Лавочкина. Для прогноза использовался вектор перемещения точки ΔX_i , измеренный по текущему i -му кадру. Указанный вектор подвергался низкочастотной фильтрации с целью подавления флуктуаций, обусловленных тем, что на каждом кадре точка выбиралась исходя из наилучших свойств привязки и может выбираться на различных участках области поиска, где вектор потока изображения существенно различен. Получался прогноз перемещения точки на следующем кадре:

$$S_{pi+1} = W_0 S_{pi} + W_1 \Delta X_i.$$

Использование фильтров высокого порядка не имело смысла, поскольку перемещения слабо коррелированы из-за большой длительности цикла и разброса

точек привязки по визируемой поверхности. Таким образом, с весом W_0 учитывался вектор потока, построенный по последовательности предыдущих кадров, и с весом W_1 — текущий вектор потока.

В качестве меры эффективности прогноза использована евклидова длина ошибки прогноза по отношению к реальному перемещению точки в поле зрения. Эта ошибка суммировалась по всей траектории с нарастающим итогом. Определялось оптимальное сочетание коэффициентов W_0 и W_1 для получения минимальной ошибки прогноза. Результаты для сценариев посадки, разработанных в ИПМ им. М. В. Келдыша, представлены в таблице.

Суммарная ошибка прогноза положения точки в поле зрения

Вариант схода с КСО с упреждением прохождения траверза точки посадки			Вариант схода с КСО при прохождении траверза точки посадки		
W_0	W_1	суммарная ошибка прогноза [пикс.]	W_0	W_1	суммарная ошибка прогноза [пикс.]
1	0	309	1	0	247
0,7	0,3	280	0,7	0,3	217
0,4	0,6	274	0,4	0,6	216
0,1	0,9	282	0,1	0,9	215
0	1	287	0	1	217

Было произведено 55 измерений для обоих сценариев. Как следует из таблицы, векторы перемещений точек в поле зрения слабо коррелированы и оптимальные значения весовых коэффициентов составляют $W_0 = 0,25$ и $W_1 = 0,75$.

При посадке по сценарию НПО им. С. А. Лавочкина суммарная ошибка прогноза оказывается меньше в 2...3 раза при большем числе измерений (87). Учёт этого обстоятельства позволил бы за счёт уменьшения числа операций сократить время вычислений примерно в 3...6 раз и уменьшить вероятность аномальных ошибок установления соответствия примерно в 4...9 раз. Однако это можно сделать только в том случае, если будет принят сценарий посадки, разработанный в НПО им. С. А. Лавочкина. К сожалению, нет никакой определённости по поводу используемого сценария посадки, поэтому приходится ориентироваться на сценарий посадки, разработанный в ИПМ им. М. В. Келдыша, хотя такое решение чревато увеличением числа аномальных ошибок измерений в случае, если бы выполнялась посадка по сценарию, разработанному в НПО им. С. А. Лавочкина.

По результатам моделирования работы монокулярного канала были оптимизированы значения параметров, определяющих переключение режимов (поиск, обнаружение, захват, слежение). В частности, уменьшено граничное значение счётчика циклов, обеспечивающее переход в режим захвата с 5 до 2. Значение ошибки прогноза, при которой происходит переключение из режима захвата в режим поиска, увеличено с 7 до 11. Это улучшает работу при посадке по сценарию, разработанному в ИПМ им. М. В. Келдыша (ускоряет переход в режим слежения и повышается устойчивость в этом режиме). Разумеется, это ухудшает точность измерений в режиме посадки по сценарию, разработанному в НПО им. С. А. Лавочкина.

Выводы

Моделирование показало работоспособность алгоритмов слежения и прогноза при посадке по сценариям, разработанным в НПО им. С. А. Лавочкина и ИПМ

им. М. В. Келдыша. Анализ результатов моделирования позволил оптимизировать параметры этих алгоритмов.

По результатам моделирования можно сделать вывод, что сценарий посадки, разработанный в НПО им. С. А. Лавочкина, существенно более благоприятен для работы системы технического зрения. В условиях неопределённости сценария посадки, ограниченности вычислительной мощности бортового процессора, а также особенностей архитектуры КА приходится оптимизировать параметры алгоритмов из компромиссных соображений.

Литература

- Аким Э. Л., Архангельский Р. Н., Зайко Ю. К., Лавренов С. М., Порошин А. Л., Рузский Е. Г., Степаньянц В. А., Тучин А. Г., Тучин Д. А., Федотов В. П., Ярошевский В. С. (2009) Баллистика, навигация и управление движением космического аппарата на этапе его посадки на поверхность Фобоса // Космич. исслед. 2009. Т. 47. № 4. С. 332–342.
- Гришин В. А. (2009) Алгоритмы измерения высоты и компонент скорости по телевизионным изображениям при посадке на Фобос // Всерос. научно-техн. конф. «Современные проблемы определения ориентации и навигации космических аппаратов»: Сб. тр. Сер. «Механика, управление и информатика». Россия, Таруса, 22–25 сент. 2008. М.: ИКИ РАН, 2009. С. 279–293.
- Жуков Б. С., Жуков С. Б., Гришин В. А., Гордеев Р. В. (2009) Бортовое программно-алгоритмическое обеспечение информационной поддержки посадки на Фобос // Всерос. научно-техн. конф. «Современные проблемы определения ориентации и навигации космических аппаратов»: Сб. тр. Сер. «Механика, управление и информатика». Россия, Таруса, 22–25 сент. 2008. М.: ИКИ РАН, 2009. С. 294–303.
- Zhukov B., Avanesov G., Grishin V., Krasnopevtseva E. (2007) On-board Real-Time Image Processing to Support Landing on Phobos // 7th Intern. Symp. Reducing the Costs of Spacecraft Ground Systems and Operations (RCSGSO). Russia, Moscow, 11–15 June, 2007.

ANALYSIS OF SPACECRAFT DYNAMIC'S EFFECTS ON THE CHARACTERISTICS OF THE IMAGE PROCESSING ALGORITHMS IN THE COMPUTER VISION SYSTEM OF THE MISSION PHOBOS-GRUNT

V. A. Grishin

Space Research Institute (IKI RAN), 117997 Moscow, 84/32 Profsoyuznaya st.

This paper describes the technique and analysis results of computer vision system measurement and prediction algorithms working through. This computer vision system is used in the spacecraft of PHOBOS-GRUNT mission. The working through was carried out with using of landing ballistics and Phobos model.

Keywords: spacecraft landing, computer vision systems, spacecraft dynamic, measurement and prediction algorithms.

К ВОПРОСУ ОЦЕНКИ КОРРЕЛЯЦИОННЫХ ХАРАКТЕРИСТИК ШУМА ЦИФРОВЫХ ОПТИКО- ЭЛЕКТРОННЫХ СИСТЕМ ВИДИМОГО ДИАПАЗОНА

Ю. Г. Веселов, И. В. Карпиков

*Военный учебный научный центр (ВУНЦ) ВВС «Военно-воздушная академия
им. проф. Н. Е. Жуковского и Ю. А. Гагарина», 125167, Москва, Планетная, 3*

Детально проанализированы корреляционные характеристики шума цифровых оптико-электронных систем видимого диапазона. Получены экспериментальные зависимости значений радиуса корреляции шума от условий съёмки и параметров аппаратуры наблюдения в лётных условиях.

Работа выполнена при финансовой поддержке РФФИ (проект № 08-08-00157).

Ключевые слова: цифровые фотоаппараты; статистические характеристики шума; автокорреляционная функция; площадь корреляции; радиус корреляции.

Введение

Требования к оперативности и достоверности, предъявляемые к воздушному и космическому мониторингу земной поверхности, вынуждают искать пути автоматизации процессов дешифрирования полученных изображений.

Применение стереосъёмки позволяет автоматизировать решение задач высокоточного определения координат объектов, обнаруженных любым из применяемых методов, путём привязки изображений к цифровой карте местности. Кроме того, стереосъёмка позволяет автоматизировать решение задачи обнаружения и распознавания объектов.

Точность цифровых стереоскопических систем зависит от статистических характеристик возмущений. Дисперсия $\sigma_{\delta_f}^2$ и радиус корреляции ρ_{δ_f} шума элементарных приёмников излучения ПЗС- или КМОП-матриц являются одними из факторов, влияющих на выбор участков коррекции для стереоскопической навигационной системы.

С целью исследования основных характеристик и оценки особенностей эксплуатации цифровых фотоаппаратов в полёте 1 октября 2007 г. в районе аэродрома Кубинка на самолёте Ан-30Б с использованием двух цифровых фотоаппаратов Canon EOS 350D (физический размер КМОП-матрицы $22,2 \times 14,8$ мм, размер кадра 3456×2304 пикс., линейный размер пиксела $a \approx 6$ мкм) с объективами, имеющими фиксированные фокусные расстояния, соответственно $f = 28$ и 60 мм, был проведён лётный эксперимент.

ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ОЦЕНКИ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ШУМА ИЗОБРАЖЕНИЯ

При неизменном входном сигнале и фиксированных условиях наблюдения на выходе приёмника излучения всегда имеется хаотически флуктуирующая составляющая сигнала, называемая шумом.

Веселов Юрий Геннадьевич — кандидат технических наук, доцент, e-mail: vesel_foto@mail.ru.
Карпиков Игорь Владимирович — адъюнкт.

Шумы, проявляющиеся в системах с приёмниками излучения, зависят не только от свойств фотоприёмников, но и от условий их работы. Перечислим основные составляющие шума приёмника излучения (на примере фоточувствительного прибора с зарядовой связью — ФПЗС-матрицы) [Пресс, 1991; Шипилов, 2002].

Шум темнового сигнала. При закрытом объективе цифрового фотоаппарата непрозрачной крышкой в осциллограмме выходного сигнала наряду с сигналом синхронизации присутствует темновой сигнал. Основной причиной темнового сигнала является термоэлектронная эмиссия, при которой темновой ток экспоненциально уменьшается при снижении температуры.

Фотонный шум. Этот шум является следствием дискретной природы света. Для максимальных сигналов среднеквадратическое отклонение (СКО) фотонного шума в элементе матрицы ФПЗС достигает сотен электронов.

Шум переноса. Во время переноса зарядового пакета по элементам ФПЗС некоторая часть электронов теряется, она захватывается на дефектах и примесях, существующих в кристалле. Эта неэффективность переноса является функцией количества переносимых зарядов, числа переносов и неэффективности отдельного акта переноса.

Шум считывания. Этот шум появляется в каждом элементе ФПЗС, когда сигнал, накопленный в этом элементе, выводится из матрицы, преобразуется в напряжение и усиливается. Шум считывания присутствует в изображении при отсутствии света на входе цифрового фотоаппарата и не зависит от шума темнового сигнала.

Шум сброса. Перед вводом в детектирующий узел сигнального заряда необходимо вывести предыдущий заряд, для чего используется транзистор сброса. Электрический уровень сброса зависит только от температуры, ёмкости узла считывания и его СКО с учётом заряда электрона и постоянной Больцмана.

Как уже отмечалось, в эксперименте использовались камеры с фоточувствительными КМОП-матрицами, основные достоинства которых в сравнении с ФПЗС — возможность интеграции на одном кристалле функций приёма и обработки изображения, т. е. реализации однокристалльной камеры с цифровым выходом, низкая потребляемая мощность и повышенная радиационная стойкость.

Основные недостатки — малая амплитуда выходных сигналов и высокий уровень шумов, низкая фоточувствительность, более высокий темновой ток, большие размеры активного элемента, меньшая, чем у ФПЗС, разрешающая способность [Стемковский, Шилин, 2003].

Рассмотренные компоненты шума ФПЗС- и КМОП-сенсоров являются первой причиной ограничения потенциально достижимого отношения сигнал/шум. Причём при наличии корреляции между значениями шума в моменты отсчёта выигрыш в отношении сигнал/шум всегда меньше, чем в случае независимых значений шума [Харкевич, 1963].

Предсказать величину шума в заданный момент времени невозможно, поэтому для оценки его вероятных значений пользуются статистическими методами.

По результатам проведённого эксперимента оценены шумовые характеристики цифрового фотоаппарата в лётных условиях.

Наиболее полной характеристикой шума является автокорреляционная функция, которая (в предположении эргодичности шума) может быть рассчитана по одной реализации:

$$K_{\text{ш}}(\xi, \eta) = \lim_{A \rightarrow \infty} \frac{1}{A} \iint_A [U(x, y) - \bar{U}] [U(x + \xi, y + \eta) - \bar{U}] dx dy,$$

где A — площадь изображения, по которой выполняется интегрирование (если A достаточно большая, но конечная величина, то соответствующую функцию $K_{\text{ш}}(\xi, \eta)$ назовем эмпирической корреляционной функцией, рассчитанной по одной реализации шума) [Белоглазов и др., 1985]; $U(x, y)$ — значение поля в точке (x, y) ; $\bar{U} = \lim_{A \rightarrow \infty} \frac{1}{A} \iint_A U(x, y) dx dy$ — эмпирическое математическое ожидание; ξ и η — сдвиги по осям x и y соответственно между точками изображения, для которых рассчитывается корреляционная связь.

Типовой вид автокорреляционной функции по осям ξ и η представлен на рис. 1. Вид автокорреляционной функции шума, рассчитанной по реальным изображениям, полученным в лётных условиях, представлен на рис. 2.

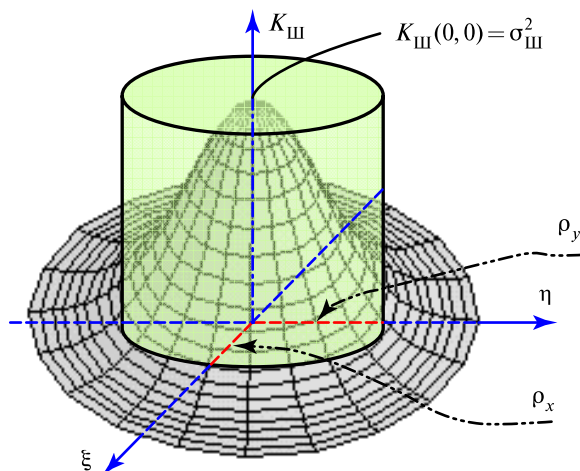


Рис. 1. Типовой вид автокорреляционной функции

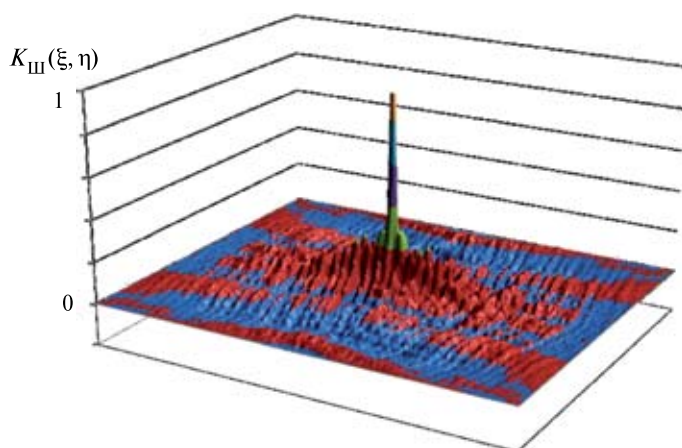


Рис. 2. Автокорреляционная функция шума для аэрофотоснимка, полученного цифровым фотоаппаратом Canon EOS 350D (высота фотографирования 500 м): $\rho = 9,028546$ пикс., $F = 256,085794$ пикс.×пикс., $\rho_x = 7,897064$ пикс., $\rho_y = 8,089217$ пикс.

По автокорреляционной функции вычисляется спектральная плотность (энергетический спектр) шумов изображения $S_{\text{ш}}(N_x, N_y)$ [Белоглазов и др., 1985; Кочетков, 1995]:

$$S_{\text{ш}}(N_x, N_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_{\text{ш}}(\xi, \eta) \exp\{-j2\pi(N_x \xi + N_y \eta)\} d\xi d\eta,$$

где N_x, N_y — пространственные частоты по осям x и y соответственно.

Можно также использовать понятие площади корреляции F двумерного случайного поля $U(x, y)$ [Белоглазов и др., 1985]:

$$F = \frac{\int_0^{\infty} \int_0^{\infty} K_{\text{ш}}(\xi, \eta) d\xi d\eta}{\sigma_{\text{ш}}^2},$$

где $K_{\text{ш}}(\xi, \eta)$ — корреляционная функция поля $U(x, y)$; $\sigma_{\text{ш}}^2 = K_{\text{ш}}(0, 0)$ — дисперсия этого поля. Площадь корреляции совпадает с площадью цилиндра высотой $\sigma_{\text{ш}}^2$, объём которого равен объёму фигуры, ограниченной поверхностью $K_{\text{ш}}(\xi, \eta)$ и горизонтальной плоскостью ξ и η (см. рис. 1), и характеризует спектральный состав поля. Если поле изотропно, т. е. его спектральные свойства вдоль осей x и y одинаковы, то корреляционная функция $K_{\text{ш}}$ является телом вращения, а площадь корреляции удобно представлять кругом радиуса ρ , так что

$$F = \pi\rho^2.$$

Здесь величина ρ называется радиусом корреляции изотропного поля. В случае, когда поле не изотропно, использование понятия площади корреляции остаётся правомерным. Однако наряду с F для более точного описания структуры поля вдоль направлений x и y применяются корреляции ρ_x и ρ_y , под которыми можно понимать радиусы корреляции одномерных сечений $U(x, 0)$ и $U(0, y)$:

$$\rho_x = \frac{\int_0^{\infty} K_{\text{ш}}(\xi, 0) d\xi}{K_{\text{ш}}(0, 0)}, \quad \rho_y = \frac{\int_0^{\infty} K_{\text{ш}}(0, \eta) d\eta}{K_{\text{ш}}(0, 0)}.$$

Необходимо также отметить, что введенные простые определения площади F и радиусов корреляции ρ , ρ_x и ρ_y , к сожалению, не универсальны. Их удобно применять только в случаях, когда корреляционная функция $K_{\text{ш}}(\xi, \eta)$ неотрицательна либо области её отрицательных значений незначительны [Белоглазов и др., 1985; Красовский и др., 1979].

Известно, что некоррелированными можно считать только значения случайного поля, отстоящие друг от друга не менее чем на интервал корреляции [Харкевич, 1963].

Интервал корреляции определяется [Харкевич, 1963; Тихонов, 1966] следующим образом:

$$\xi_0 = \frac{\int_{-\infty}^{\infty} K_{\text{ш}}(\xi) d\xi}{K_{\text{ш}}(0)}, \quad \xi_0 = 2\rho_x,$$

$$\eta_0 = \frac{\int_{-\infty}^{\infty} K_{\text{ш}}(\eta) d\eta}{K_{\text{ш}}(0)}, \quad \eta_0 = 2\rho_y.$$

При этом ξ_0, η_0 (аналогично определению радиусов корреляции) — основание прямоугольника, высота которого есть $K_{\text{ш}}(0)$, площадь равна площади под кривой (рис. 3). Одномерная автокорреляционная функция, полученная по реальным изображениям полей яркости, представлена на рис. 4.

Для изотропного поля интервал корреляции равен

$$\tau_0 = 2\sqrt{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_{\text{ш}}(\xi, \eta) d\xi d\eta}{\pi K_{\text{ш}}(0, 0)}}.$$

ОСОБЕННОСТИ ОЦЕНКИ РАДИУСА КОРРЕЛЯЦИИ ШУМА СЛУЧАЙНЫХ ПОЛЕЙ, ПРЕДСТАВЛЕННЫХ В ДИСКРЕТНОМ ВИДЕ

Для определения ρ случайных полей, представленных в дискретном виде, можно воспользоваться свойствами дискретного преобразования Фурье (ДПФ).

Автокорреляционная функция массива $f(r, l)$ задаётся формулой [Романюк, 1981]

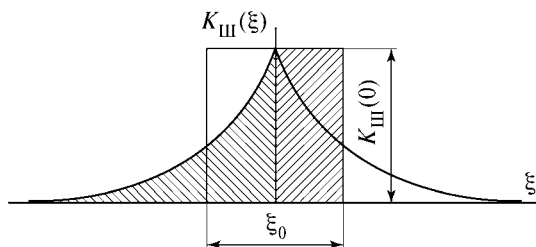


Рис. 3. Определение радиуса корреляции

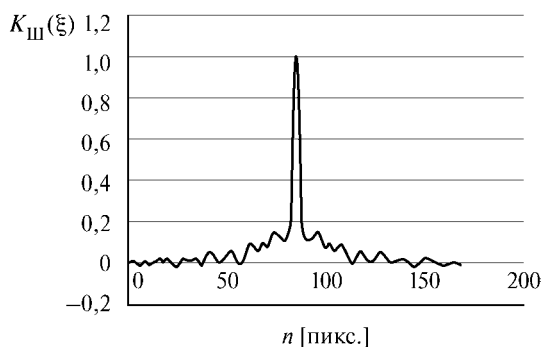


Рис. 4. Автокорреляционная функция, полученная по реальным изображениям полей яркости (лётные условия)

$$\begin{aligned}
 R(\xi, \eta) &= \frac{1}{L} \frac{1}{R} \sum_{l=0}^{L-1} \sum_{r=0}^{R-1} f(r, l) f(r + \xi, l + \eta) = \frac{1}{L} \frac{1}{R} \sum_{l=0}^{L-1} \sum_{r=0}^{R-1} \left[\sum_{\lambda=0}^{L-1} \sum_{q=0}^{R-1} F(\lambda, q) W^{(Lr+l)(R\lambda+q)} \right] \times \\
 &\times \left[\sum_{v=0}^{L-1} \sum_{\gamma=0}^{R-1} F(v, \gamma) W^{(Rv+\gamma)(Lr+l)} W^{(Rv+\gamma)(L\xi+\eta)} \right] = \\
 &= \frac{1}{N} \sum_{\lambda=0}^{L-1} \sum_{q=0}^{R-1} \sum_{v=0}^{L-1} \sum_{\gamma=0}^{R-1} F(\lambda, q) F(v, \gamma) W^{(Rv+\gamma)(L\xi+\eta)} \left[\sum_{l=0}^{L-1} \sum_{r=0}^{R-1} W^{((R\lambda+q)+(Rv+\gamma))(Lr+l)} \right],
 \end{aligned} \quad (1)$$

где $F(\lambda, q) = \sum_{l=0}^{L-1} \sum_{r=0}^{R-1} f(r, l) W^{(Lr+l)(R\lambda+q)}$; в этом выражении

$$W^{(Lr+l)(R\lambda+q)} = \exp \left\{ -\frac{2\pi}{N} (Lr+l)(R\lambda+q) \right\} \text{ — ядро преобразования ДПФ; } N = LR \text{ —}$$

количество элементов в двумерном массиве, состоящем из L столбцов с номерами $l = 0, 1, 2, \dots, L-1$ и R строк с номерами $r = 0, 1, 2, \dots, R-1$ [Белоглазов и др., 1985]; λ и q — переменные двумерного ДПФ.

Таким образом, $F(\lambda, q)$ ДПФ $f(r, l)$, а $F(v, \gamma)$ ДПФ $f(r + \xi, l + \eta)$.

Возвращаясь к выражению (1), отметим, что $\sum_{l=0}^{L-1} \sum_{r=0}^{R-1} W^{((R\lambda+q)+(Rv+\gamma))(Lr+l)} = N$, если $(Rv + \gamma) = -(R\lambda + \gamma)$.

Поэтому

$$R(\xi, \eta) = \sum_{\lambda=0}^{L-1} \sum_{q=0}^{R-1} F(\lambda, q) F(-\lambda, -q) W^{-(R\lambda+q)(L\xi+\eta)} = \sum_{\lambda=0}^{L-1} \sum_{q=0}^{R-1} G(\lambda, q) W^{-(R\lambda+q)(L\xi+\eta)}, \quad (2)$$

где $G(\lambda, q) = F(\lambda, q) F(-\lambda, -q) = F(\lambda, q) F^*(\lambda, q) = |F(\lambda, q)|^2$ — энергетический спектр массива $f(r, l)$. Таким образом, согласно (2) автокорреляционная функция массива $f(r, l)$ и энергетический спектр этого массива связаны парой ДПФ:

$$G(\lambda, q) = \frac{1}{N} \sum_{\xi=0}^{L-1} \sum_{\eta=0}^{R-1} R(\xi, \eta) W^{-(R\lambda+q)(L\xi+\eta)}, \quad (3)$$

$$R(\xi, \eta) = \sum_{\lambda=0}^{L-1} \sum_{q=0}^{R-1} G(\lambda, q) W^{(R\lambda+q)(L\xi+\eta)}.$$

На основе теории дискретного преобразования Фурье был разработан алгоритм и составлена программа для оценки радиуса корреляции пространственных полей, представленных в дискретном виде, с помощью которой выполнялись все расчёты в рассматриваемой работе.

ЛЁТНЫЙ ЭКСПЕРИМЕНТ И ЕГО РЕЗУЛЬТАТЫ

Съёмка на всех высотах, за исключением высоты 1200 м, проводилась с автоматической регулировкой экспозиции. На высоте 1200 м регулирование экспозиции осуществлялось вручную посредством изменения диафрагменного числа $f/4,5$; $f/5,6$; $f/8$.

Изображения записывались в двух форматах: jpg и tiff. Расчёты характеристик шума проводились отдельно по снимкам форматов jpg и tiff.

Для оценки статистических характеристик шума ОЭС использовались изображения полей яркости российской проверочной миры, используемой с оптическими фотоаппаратами (рис. 5).



Рис. 5. Серое поле яркости на традиционном тестовом полигоне, используемое для оценки характеристик шума

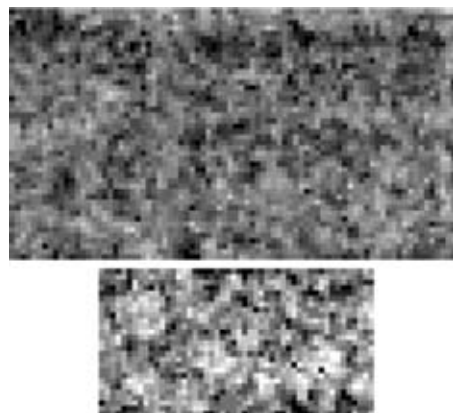


Рис. 6. Изображения шума, полученные при аэрофотосъемке с высот 1200 и 2000 м

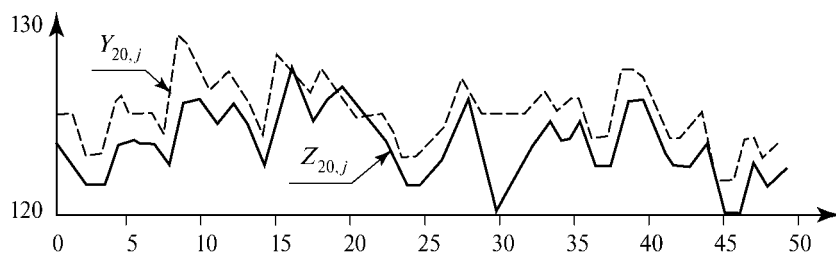


Рис. 7. Распределение яркости вдоль одной строки параллельной оси абсцисс, полученное с высот 1200 и 2000 м (Z — высота 1200 м, $\rho = 8,195$ пикс.; $\sigma = 1,875$; Y — высота 2000 м, $\rho = 6,125$ пикс., $\sigma = 2,065$)

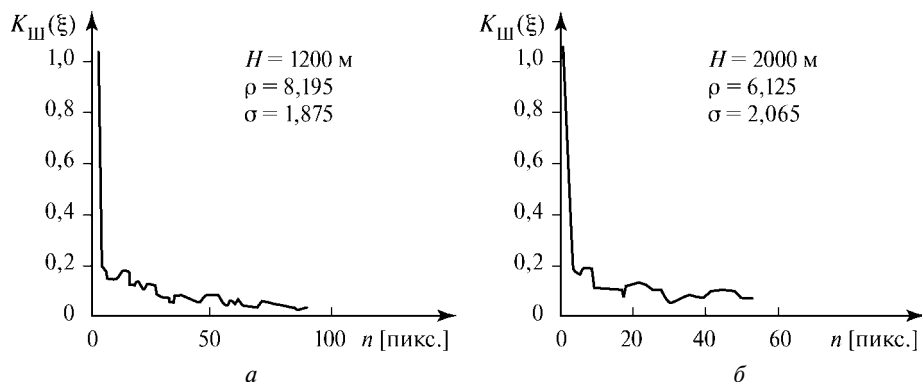


Рис. 8. Автокорреляционная функция изображения равномерного поля яркости

На аэрофотоснимке возможно найти равномерные по яркости объекты, которые можно использовать для расчётов характеристик шума (свежее асфальтовое покрытие, свежее рубероидное покрытие крыши и т. п.). В работе расчёты проводились в том числе и по этим объектам.

Зашумленные поля яркости с высот 1200 и 2000 м представлены на рис. 6 (см. с. 185).

Распределение яркости вдоль одной строки, параллельной оси абсцисс, представлено на рис. 7 (см. с. 185).

Автокорреляционная функция этих сечений представлена на рис. 8 (см. с. 185).

По результатам полученных расчётов были построены зависимости значений радиуса корреляции, СКО шума в уровнях серого тона (уст) от высоты полёта, фокусного расстояния и стандарта кодирования неподвижных изображений (рис. 9–15, см. также с. 187).

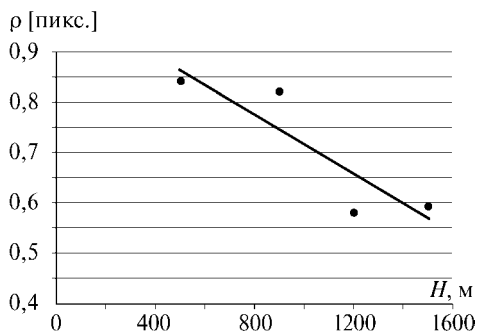


Рис. 9. Зависимость радиуса корреляции шума ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата jrg, $f = 28$ мм)

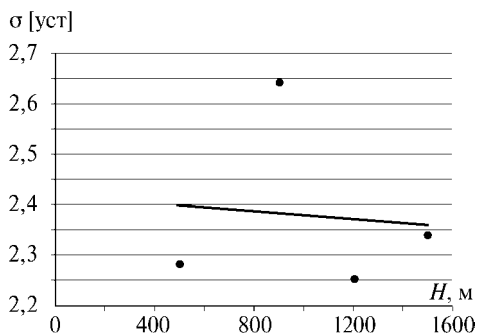


Рис. 10. Зависимость СКО шума ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата jrg, $f = 28$ мм)

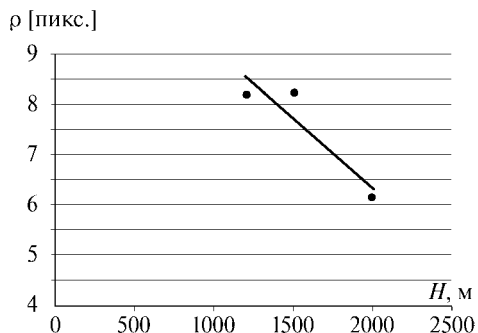


Рис. 11. Зависимость радиуса корреляции шума ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата tiff, $f = 28$ мм)

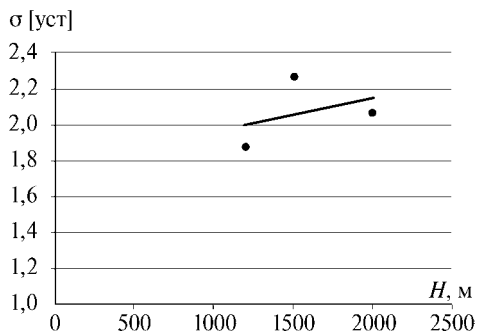


Рис. 12. Зависимость СКО шума ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата tiff, $f = 28$ мм)

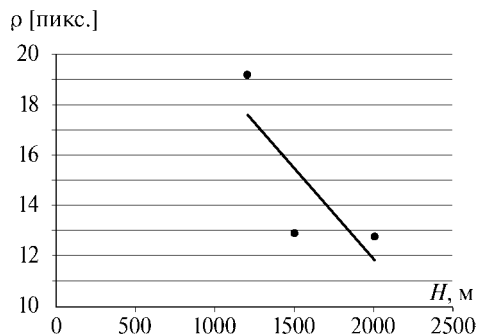


Рис. 13. Зависимость радиуса корреляции шума ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата tiff, $f = 60$ мм)

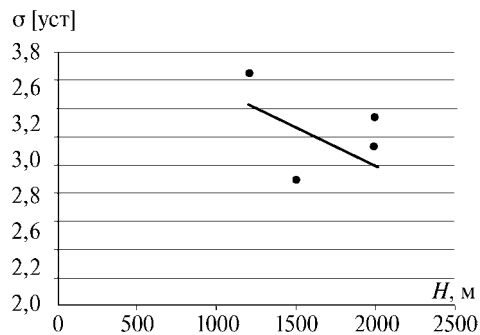


Рис. 14. Зависимость СКО шума ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата tiff, $f = 60$ мм)

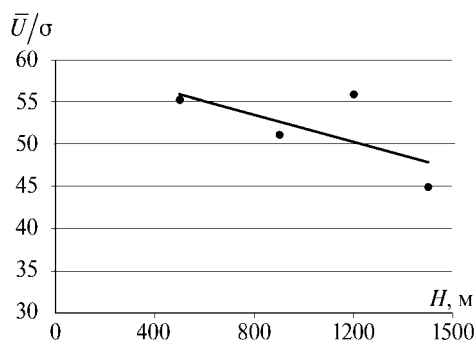


Рис. 15. Зависимость отношения сигнал/шум ЦФА Canon EOS 350D от высоты фотосъёмки (стандарт кодирования неподвижных изображений формата jpg, $f = 28$ мм)

Анализ полученных выше зависимостей показал, что стандарт кодирования неподвижных изображений не существенно влияет на значение радиуса корреляции шума и СКО шума изображения. С увеличением высоты фотографирования наблюдается небольшое уменьшение значения радиуса корреляции в плоскости изображения на 30 %, при пересчёте на местность — на 20 %. При увеличении фокусного расстояния с 28 до 60 мм значение радиуса корреляции шума в плоскости изображения увеличивается почти на 70 %, при пересчёте на местность — всего на 9 %.

Заключение

В работе подробно описаны алгоритмы оценки характеристик шума цифровых оптико-электронных систем. Продемонстрирована работоспособность предложенных методик.

Оценённые с учётом условий применения и для различных параметров съёмки значения СКО $\sigma_{\delta f}^2$ и радиуса корреляции $\rho_{\delta f}$ шума элементарных приёмников излучения КМОП-матриц в лётных условиях используются для разрежения сети

соответственных точек и оценки точности беспойсковой стереоскопической навигационной системы.

Литература

- Белоглазов И. Н., Джанджгава Г. И., Чигин Г. П.* (1985) Основы навигации по геофизическим полям. М.: Наука, 1985.
- Кочетков Ю. А.* (1995) Основы автоматики авиационного оборудования. М.: ВВИА им. проф. Н. Е. Жуковского, 1995.
- Красовский А. А., Белоглазов И. Н., Чигин Г. П.* (1979) Теория корреляционно-экстремальных навигационных систем. М.: Наука, 1979.
- Пресс Ф. П.* (1991) Фоточувствительные приборы с зарядовой связью. М.: Радио и связь, 1991.
- Романюк Ю. А.* (1981) Дискретные преобразования сигналов. М.: МФТИ, 1981.
- Стемковский А., Шилин В.* (2003) КМОП-фотодиодные СБИС. Перспективная элементная база однокристалльных систем приёма и обработки информации // Электроника, наука, технология, бизнес. 2003. № 2.
- Тихонов В. И.* (1966) Статистическая радиотехника. М.: Сов. радио, 1966.
- Харкевич А. А.* (1963) Борьба с помехами. М.: Гос. изд-во физ.-мат. лит., 1963.
- Шипилов А. Е.* (2002) Война с шумами // Е-фото. 2002. № 12.

TO A QUESTION OF AN ESTIMATION OF CORRELATION CHARACTERISTICS OF NOISE OF DIGITAL OPTICO-ELECTRONIC SYSTEMS OF A VISIBLE RANGE

Ju. G. Veselov, I. V. Karpikov

*Air Force Academy n. a. Professors N. E. Zhukovsky and Ju. A. Gagarin,
125167, Moscow, 3 Planetary st.*

Correlation characteristics of noise of digital optico-electronic systems of a visible range are in details analysed. Experimental dependences of values of radius of correlation of noise on conditions of shooting and parameters of equipment of supervision in flight conditions are received.

Keywords: digital cameras; statistical characteristics of noise; autocorrelation function; the correlation area; correlation radius.

Veselov Jury Gennadevich — Candidate of Technical Sciences, Senior Lecturer, e-mail: vesel_foto@mail.ru.

Karpikov Igor Vladimirovich — Post-Graduate Student.

**СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ
В НАЗЕМНЫХ И ПОДВОДНЫХ РОБОТАХ**

ТЕХНИЧЕСКОЕ ЗРЕНИЕ МОБИЛЬНЫХ РОБОТОВ

Б. Б. Михайлов

*Научно-учебный центр «Робототехника» Московского государственного
технического университета им. Н. Э. Баумана,
105037, Москва, Измайловская пл., 7*

Рассматриваются три системы технического зрения, которые обеспечивают управление движением мобильного робота в условиях недетерминированной среды.

Ключевые слова: мобильные роботы, техническое зрение, системы управления, осязание, видеокамеры, объёмное зрение, построение маршрута, навигация робота, корреляционная функция.

Введение

В последнее время экономическая ситуация в наукоемких отраслях народного хозяйства существенно улучшилась, в том числе и в робототехнике. Снова продолжились работы по созданию промышленных роботов и специальных мобильных робототехнических систем. К сожалению, системы управления этими роботами в большинстве своём остались на прежнем уровне: основной элемент системы управления — человек-оператор с пультом в руках, а на пульте — традиционные монитор, тумблеры, джойстики и т. д. Отсюда — все преимущества и недостатки этих систем. Такая ситуация связана с отсутствием средств, которые могли бы существенно облегчить работу оператору, а в некоторых случаях и совсем заменить его.

В этой статье рассматриваются системы технического зрения, с помощью которых собирается и анализируется до 90 % необходимой для управления информации. В частности, приводятся результаты работ, выполненных в научно-учебном центре «Робототехника» МГТУ им. Н. Э. Баумана в области систем зрения мобильных роботов (МР) за последние пять лет.

Рассмотрим простую задачу — движение мобильного робота из одной точки в другую в частично или полностью недетерминированной среде. Иными словами, роботу необходимо добраться из точки старта в точку финиша при условии, что ситуация в рабочем пространстве, которое необходимо преодолеть, известна не полностью и, более того, может меняться в процессе движения. Например, это мобильный транспортный робот для обслуживания станков в промышленном цехе, при условии, что там могут двигаться другие роботы, ходить рабочие и т. д. Другой пример — доставка лекарств в больничную палату. Здесь столы, шкафы, кровати стоят неподвижно, но в коридоре может появиться врач, а больному захочется посидеть за столом.

Для того чтобы робот мог работать автономно в таких условиях, необходимо решить три основные задачи:

- 1) проложить маршрут (траекторию) движения робота;
- 2) распознать и определить параметры препятствий, которые появляются перед роботом во время движения;

- 3) определить текущие координаты мобильного робота в рабочем пространстве.

Ниже приведено краткое описание систем технического зрения для решения этих задач.

1. ТЕЛЕВИЗИОННАЯ СИСТЕМА ДЛЯ ПРОКЛАДКИ МАРШРУТА

В этой системе зрения используются стационарные телекамеры, которые контролируют рабочее пространство робота. На основании поступающей видеоинформации составляется план сцены и выделяются свободные участки пространства, которые можно использовать для движения робота. Далее оператор программно или с помощью «мышки» указывает начальную (стартовую) и конечную (финишную) точки маршрута, и система находит траекторию, соединяющую эти точки, если решение существует. Причём из всех возможных вариантов система автоматически выбирает наиболее короткий маршрут движения.

Система работает с множеством различных объектов, расположенных на сложной подстилающей поверхности. Причём объекты, в общем случае, могут иметь произвольную форму и любое сочетание цветов. Выделение объектов в таких условиях является достаточно сложной задачей. Для её решения был использован ряд фильтров различной размерности: усредняющий, медианный и фильтр Гаусса. Кроме того, были использованы фильтрация объектов по площади и процедура слияния нескольких объектов в один. Результаты выделения объектов из исходного изображения (рис. 1а) показаны на рис. 1б.

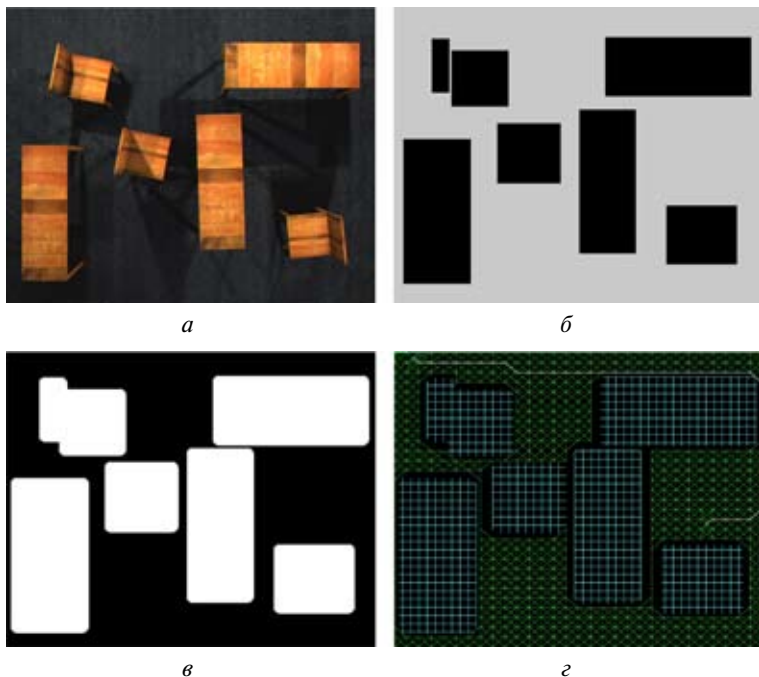


Рис. 1. Последовательность работы системы прокладки маршрута

Вторая задача, которая была решена, — это учёт геометрических размеров реального мобильного робота. Из известных способов решения в системе был применен следующий: геометрические размеры изображений объектов увеличивались (расширялись) на величину, равную радиусу окружности, описывающей робот. Пример решения показан на рис. 1в.

Непосредственно прокладка траектории основана на использовании процедуры Дейкстры. Для этого на изображение сцены накладывается матрица точек. Из матрицы исключаются точки, расположенные на изображениях объектов. Остальные точки соединяются между собой и строятся ломаные линии, соединяющие точки старта и финиша, если это возможно. Далее из всех линий выбирается самая короткая. Пример маршрута показан на рис. 1г.

В качестве примера рабочей зоны на рис. 2 (см. с. 194) приведено реальное изображение, полученное с видеокамеры. Оно имеет большое количество шумов и явно выраженную дисторсию из-за оптических искажений объектива камеры. В процессе первичной обработки в каждом цветовом канале — R, G и B — компенсируется эффект отрицательной дисторсии и выполняется медианная фильтрация шумов. Для выделения перепадов яркости каждый канал обрабатывается фильтром Собеля. Полученная информация суммируется, результирующий сигнал бинаризуется и фильтруется по площади.

При начальной настройке на конкретное изображение оператор использует панель управления этапами обработки (рис. 3, поз. 2, см. с. 194). Программа имеет возможность останова после каждого этапа обработки для оценки и корректирования работы фильтров. Интерфейс содержит цифровое поле ввода для коррекции уровня дисторсии (поз. 2, п. 1), который задаётся в диапазоне от 0 до 1 (поз. 7). Как видно на рис. 3 (поз. 9) фильтр Собеля (поз. 2, п. 2) подстраивается путём выбора размера матрицы фильтра (9×9 , 19×19 или 33×33) и количества повторов фильтрации (1, 3 или 6). Медианный фильтр (поз. 2, п. 3) имеет стандартную матрицу 3×3 .

Программно реализованы следующие процедуры (поз. 2, п. 4): бинаризация, выделение периметра объектов, введение поправки на габарит робота и фильтрация по площади. Бинаризация выполняется по фиксированному пороговому значению. Подстройка периметра (поз. 3) определяет степень точности описания объекта — чем больше введённое значение в цифровое поле, тем менее точный строится периметр. Поправка на габарит робота вводится ползунком (поз. 5) и может составлять от 0 до 50 единиц. Настройка фильтра по площади осуществляется ползунком (поз. 8) и может составлять от 0 до 500 единиц.

На последнем этапе обработки строится траектория движения робота с помощью алгоритма Дейкстры (рис. 3, поз. 2, п. 6). Для этого на полученное изображение накладывается сетка из $N \times M$ вершин графа. Далее вершины соединяются и формируется взвешенный граф. Результатом работы алгоритма является набор относительных координат точек, через которые необходимо двигаться роботу, чтобы добраться из точки старта в точку финиша. Время построения траектории зависит от требуемого качества — чем больше матрица NM вершин графа, тем выше точность. Например, если матрица имеет размер 40×40 , то время поиска траектории составляет около 5 с, требуемый объём памяти — 12 Мбайт.

Результат работы алгоритма показан на рис. 3 (поз. 10). Служебная информация — в правом нижнем углу экрана (поз. 4). Здесь сохраняются числовые значения координат точек построенной траектории либо сигнал ошибки, если траектории не существует. Далее значения этих координат преобразуются в абсолютные (в систему координат рабочей зоны робота) и передаются в систему управления робота.



Рис. 2. Пример реального изображения рабочей зоны

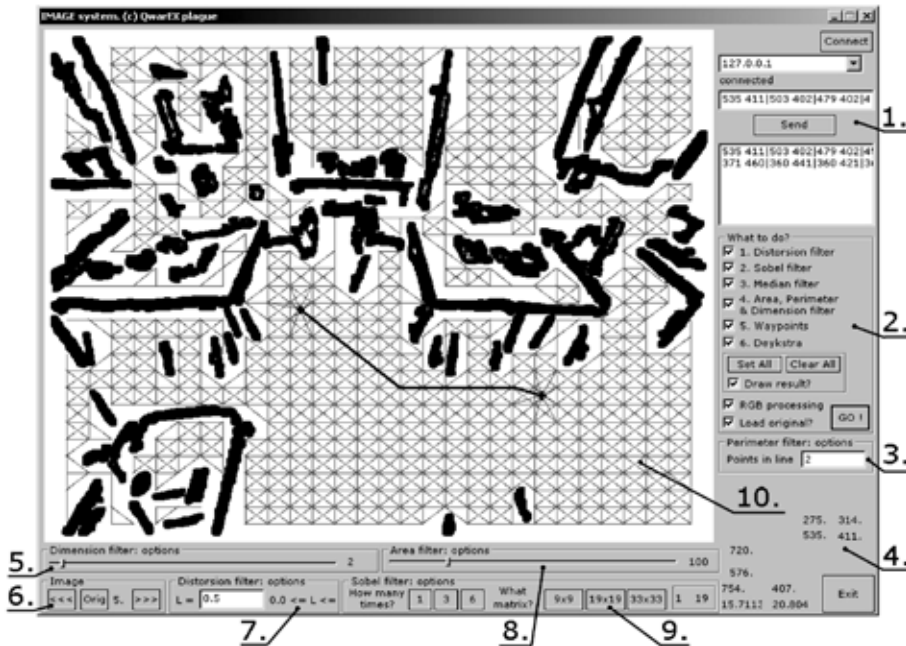


Рис. 3. Интерфейс оператора

2. ТЕЛЕВИЗИОННАЯ СИСТЕМА ОБЪЁМНОГО ЗРЕНИЯ

Телевизионная система объёмного зрения устанавливается на мобильном роботе. Она предназначена для распознавания препятствий, возникающих в процессе движения по траектории, определения их координат и геометрических размеров.

В телевизионной системе объёмного зрения используется метод структурной подсветки. Идея метода заключается в следующем. Объект подсвечивается структурированным светом, и спроецированная структура регистрируется телекамерой. Искажения структуры однозначно определяют форму объекта в поле зрения телекамеры. Генератор опорной подсветки работает синхронно с телекамерой. Основными параметрами системы являются базовое расстояние между генератором подсветки и камерой и параллаксный угол (угол между оптическими осями проекционного устройства и телекамеры). В общем случае в системе можно использовать две телекамеры — одну для измерения объёмного (3D) изображения, а другую — для передачи плоской картинки (2D) оператору.

Известно несколько типов структурной подсветки: линейная, точечная и т. д. Изображение структурной подсветки, которая была специально синтезирована для данной системы, приведено на рис. 4.

Для вычисления координат точки объекта применяется триангуляционный метод, в котором используются две геометрические модели: проекционного устройства и приёмной телекамеры. Первая модель связывает элемент подсветки с его положением в системе координат объекта; вторая модель содержит соотношения, связывающие координаты точки объекта и её изображения в системе координат телекамеры. Перед началом работы выполняется калибровка системы. В данной системе был использован метод калибровки при помощи плоских эталонных объектов.

Основные параметры системы объёмного зрения

Размеры рабочей зоны по дальности	50...2600 мм
Угловые размеры поля зрения	
по горизонтали	40°
по вертикали	30,7°
Угловое разрешение	0,004°
Допустимый диапазон внешней освещённости	0...1000 лк
Время сбора визуальной информации (одного кадра)	менее 1 мс
Частота кадров	7,7 Гц

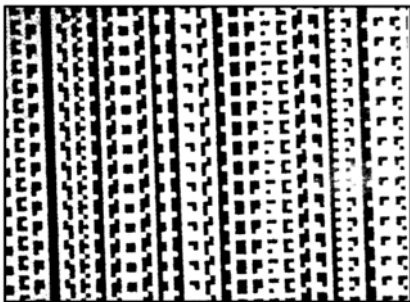


Рис. 4. Изображение структурной подсветки

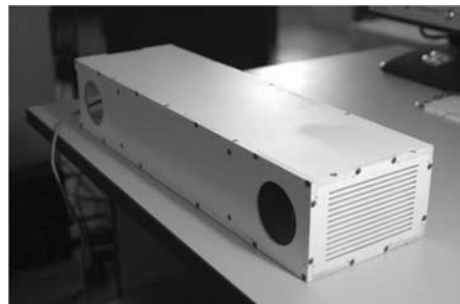


Рис. 5. Оптический блок системы объёмного зрения

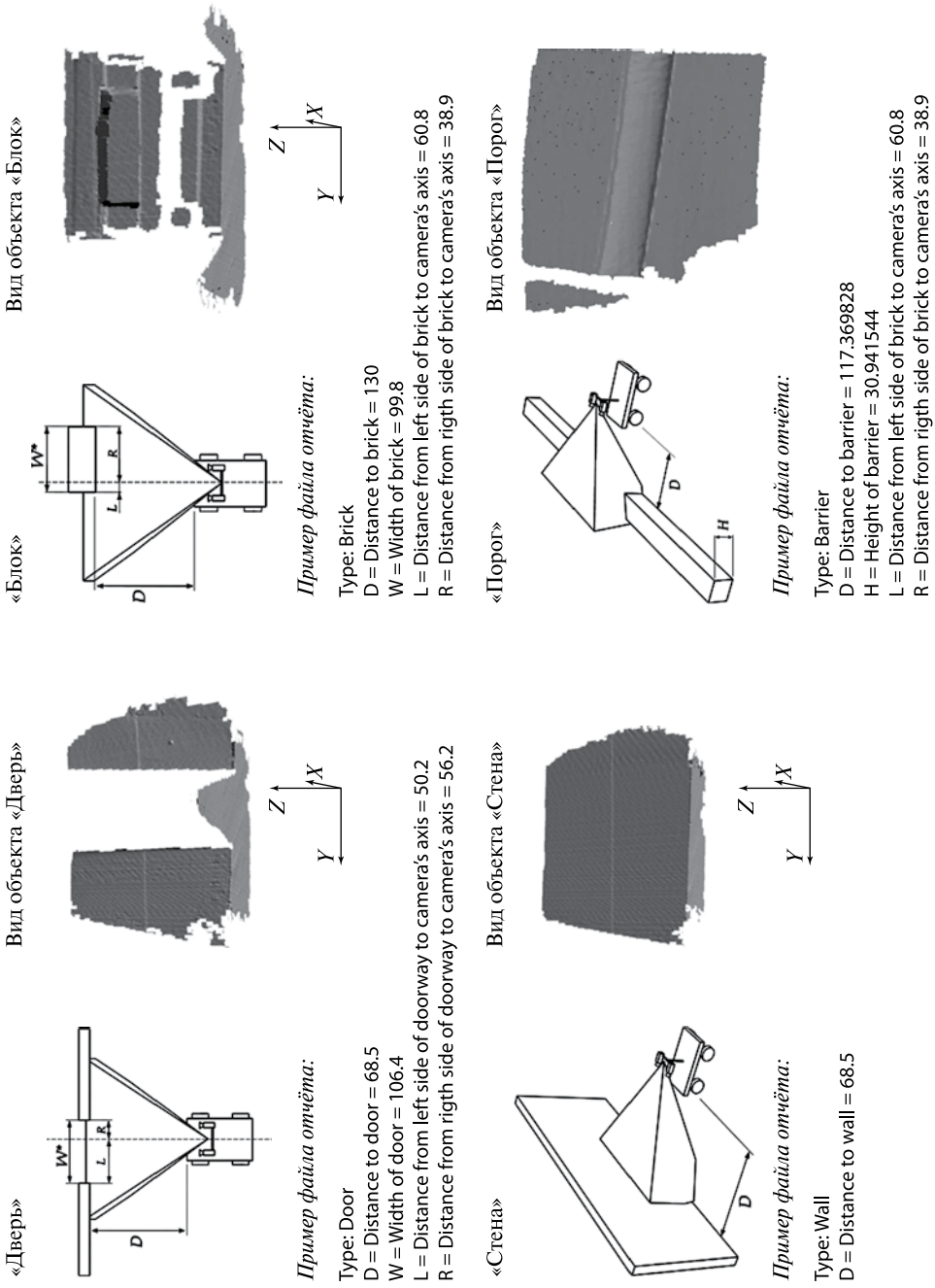


Рис. 6. Примеры препятствий и измеряемые параметры

На рис. 5 (см. с. 195) показан внешний вид оптического блока системы объёмного зрения. База системы составляет 360 мм, параллаксный угол — 14°. Блок построен на основе телевизионной камеры ВИДЕОСКАН-285, которая имеет разрешение 1392×1040 пикселей и удобный интерфейс подключения к компьютеру через порт USB 2.0.

При разработке системы объёмного зрения был выбран следующий набор типовых препятствий: «Дверь», «Блок», «Стена», «Порог». Авторы полагают, что с помощью данного набора препятствий можно описать большинство ситуаций, которые встречаются при движении робота внутри помещений. Список является открытым. Если на трассе встречаются другие препятствия, например, камень или дерево, то программа распознавания может быть легко дополнена.

На рис. 6 (см. с. 196) для каждого препятствия приведены внешний вид (слева), объёмное описание в системе зрения (справа) и список измеряемых параметров. Если в поле зрения системы встречается неизвестный объект, то он описывается параллелепипедом и его параметры измеряются так же, как у препятствия типа «Блок».

На рис. 7 и 8 приведены результаты экспериментальных исследований телевизионной системы объёмного зрения. Здесь по горизонтальной оси отложены реальные расстояния и размеры, а по вертикальной — результаты измерений; пунктиром показаны желаемые характеристики.

На рис. 7а–в показаны результаты измерения расстояния (D) до препятствий типа «Стена», «Дверь» и «Блок», соответственно. Как видно на рисунках, система

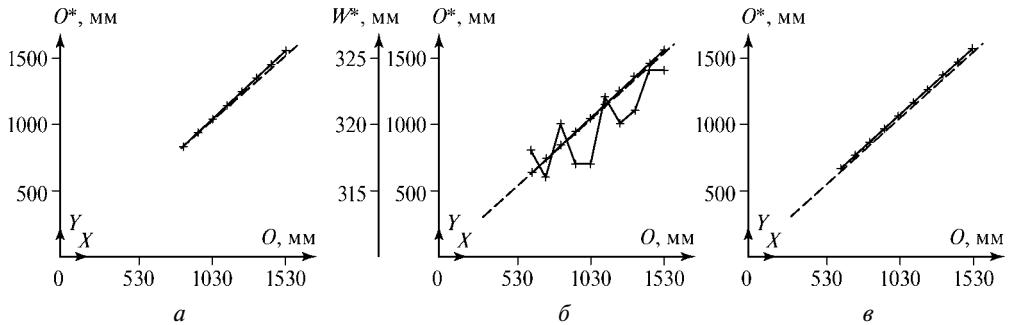


Рис. 7. Результаты измерений

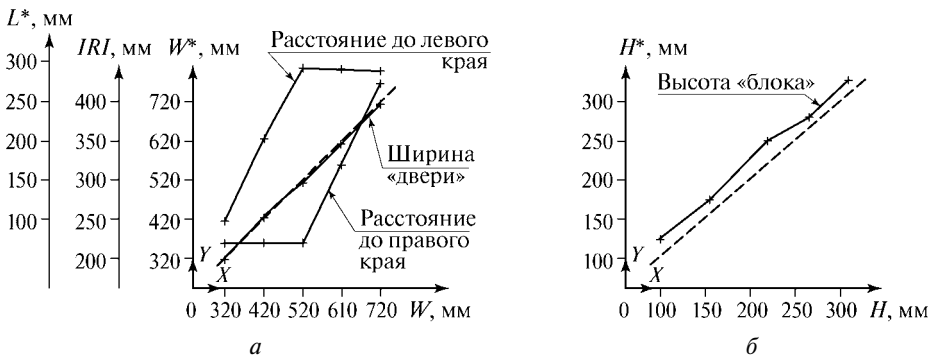


Рис. 8. Результаты измерений

обеспечивает достаточно высокую точность: ошибка измерения в рабочей зоне не превышает 5...25 мм. При этом разрешающая способность по дальности составляет 1...2 мм.

На рис. 8а показаны результаты измерения ширины двери (W), которая вычисляется как разность между расстояниями от оптической оси системы зрения до правого (R) и левого (L) краев дверного проема. Ошибка измерения ширины проема в рабочей зоне не превышает 5...10 мм. На рис. 8б показаны результаты измерения высоты (H) препятствия «Блок». Ошибка измерения достаточно велика и колеблется в пределах 15...20 мм. Её можно рассматривать как оценку высоты сверху, что достаточно для описания препятствия. В дальнейшем планируется компенсировать постоянную составляющую ошибки (рис. 8б), что существенно повысит точность измерения высоты препятствия.

Используя полученные результаты, можно принять решение о возможности движения робота в данном направлении. Например, если высота порога мала, то робот может двигаться, не снижая скорости, если высота порога больше определённого размера, то скорость движения необходимо снизить, а если высота порога превосходит допустимый размер, то препятствие становится непреодолимым и необходимо принимать решение об изменении траектории движения.

3. ТЕЛЕВИЗИОННАЯ СИСТЕМА ДЛЯ НАВИГАЦИИ МОБИЛЬНОГО РОБОТА

В этой системе зрения определение координат мобильного робота в рабочем пространстве осуществляется относительно нескольких реперных точек, координаты которых известны или могут быть измерены до начала движения робота. Минимально необходимое количество реперов — три. Зная координаты реперов и значение углов α и β (угол между направлениями на реперы), можно рассчитать положение самого робота (рис. 9). Это осуществляется путём решения типичной геодезической задачи — задачи трёх линейных засечек.

Для поиска реперных точек на изображении рабочей сцены используется математический аппарат автокорреляционных функций, что позволяет существенно увеличить скорость работы системы за счёт исключения этапов описания и распознавания реперов. Изображения реперов запоминаются в точке старта мобильного

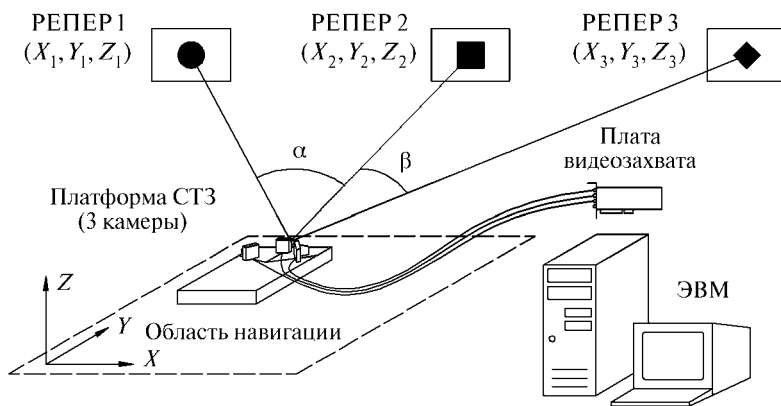


Рис. 9. Расположение реперных точек

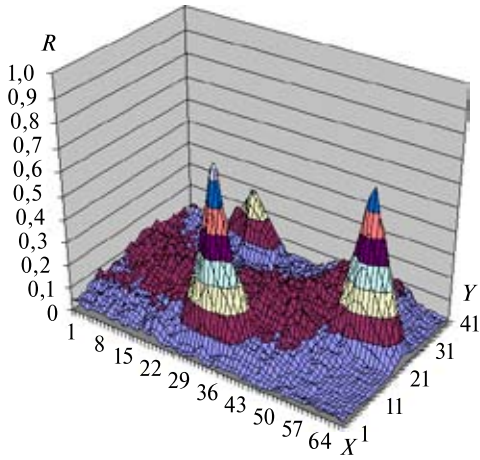


Рис. 10. Вид автокорреляционных функций



Рис. 11. Оптический блок системы навигации

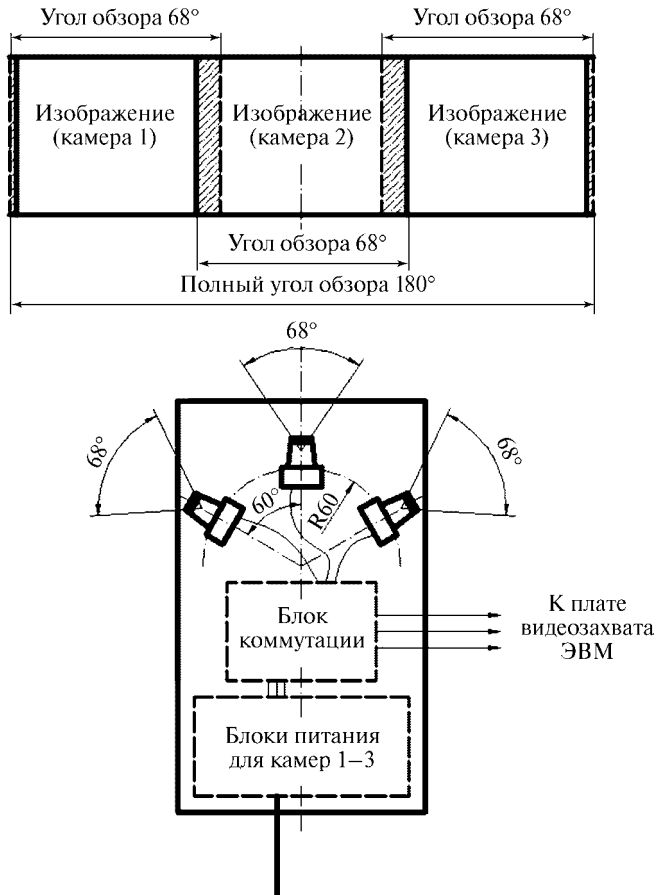


Рис. 12. Расположение телекамер системы навигации

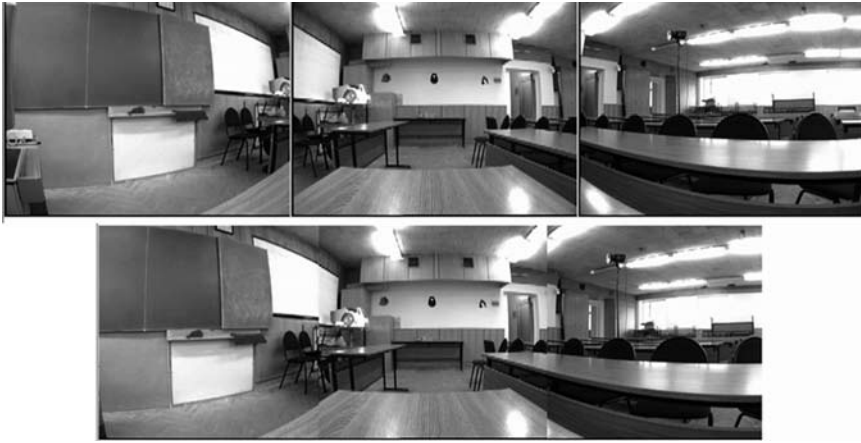


Рис. 13. Изображения рабочей сцены

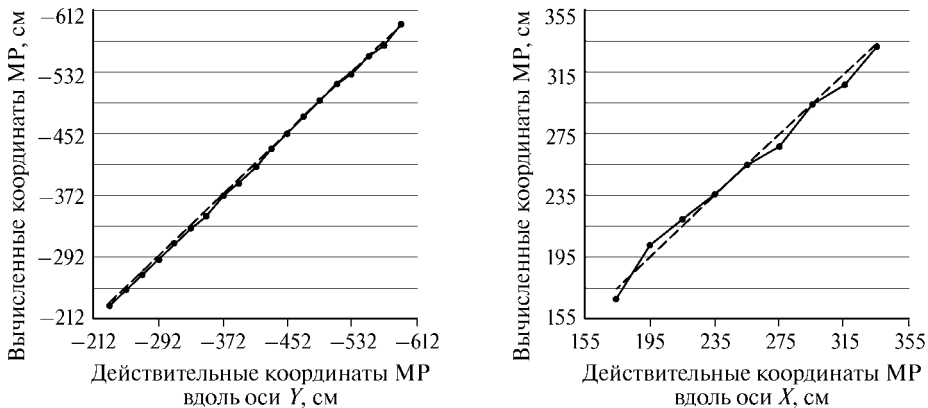


Рис. 14. Результаты измерений

робота и заносится в память как эталонные. Далее, в процессе движения выполняется поэлементное сравнение текущего изображения рабочего пространства с каждым эталонным изображением и вычисляются автокорреляционные функции. Для каждой функции находятся координаты глобального максимума, которые совпадают с координатами репера. Вид функций для трёх реперов приведён на рис. 10 (см. с. 199). Очевидно, что при движении робота изменяются угол наблюдения и линейные размеры репера. Поэтому в системе используется алгоритм, который после определения координат робота стирает текущее эталонное изображение репера и на его место заносит последнее (найденное) изображение этого репера.

На рис. 11 (см. с. 199) показан оптический блок системы навигации. В нем используются три телевизионные камеры, что позволяет анализировать переднюю полусферу перед мобильным роботом — угол обзора рабочего пространства составляет около 180° . Расположение телекамер и поле зрения системы показаны на рис. 12 (см. с. 199). На рис. 13 приведены изображения сцены, которые обрабатывались системой зрения: сверху — реальные изображения, полученные с телекамер; внизу

— совмещённое изображение, по которому проводится анализ сцены и определяются координаты реперов в системе координат, связанной с мобильным роботом.

Результаты измерения положения робота при движении в рабочем пространстве вдоль осей координат показаны на рис. 14 (см. с. 200). Здесь по горизонтальной оси отложены реальные расстояния (в сантиметрах), а по вертикальной — результаты измерений; пунктиром показаны желаемые характеристики. Как видно из данных рисунка, система имеет достаточно высокую точность измерения: ошибка составляет 6...10 см при удалении робота до 6 м от плоскости реперов.

Заключение

Как видно из вышеизложенного, все поставленные задачи могут быть решены с помощью нескольких систем технического зрения, что обеспечивает возможность управления мобильным роботом в частично детерминированной среде без участия человека-оператора. В результате действия оператора сводятся к постановке задачи и контролю за ходом её выполнения, т. е. к работе в супервизорном режиме. Это существенно упрощает процесс управления роботом и, как следствие, снижает требования к квалификации самого оператора.

В заключение необходимо отметить, что в создании описанных систем зрения, кроме автора, принимал участие коллектив исполнителей. Так, разработку алгоритмов системы для прокладки маршрута выполнил инженер А. Б. Гуров; оптический блок объёмного зрения был разработан и изготовлен в ООО НПО «Информация», а алгоритмы обработки информации составлены аспирантом Ю. С. Володиным и студентом А. В. Седовым; макет и алгоритмы системы навигации разработаны аспирантом А. С. Тищенко.

Литература

- Володин Ю. С. (2009) Метод кодирования структурной подсветки для телевизионной системы объёмного зрения мобильного робота // Мехатроника, автоматизация, управление. 2009. № 11.
- Гуров А. Б., Михайлов Б. Б. (2008) Система зрения для прокладки траектории мобильного робота и мониторинга рабочей зоны // Тр. 19-й Научно-техн. конф. «Экстремальная робототехника». СПб.: Изд-во НПО специальных материалов, 2008. Т. 5. С. 394–398.
- Тищенко А. С., Михайлов Б. Б. (2009) Навигация мобильного робота на основе бортовой системы технического зрения // Мехатроника, автоматизация, управление. 2009. № 12.

COMPUTER VISION OF MOBILE ROBOTS

B. B. Mikhailov

Scientific-Educational Center “Robotic Engineering” of Bauman Moscow State Technical University, 7 Izmailovskaya sq., Moscow 105037 Russia

Described three computer vision systems to provide motion control of mobile robot in undetermined environment.

Keywords: mobile robots, computer vision, control systems, sensor systems, camcorders, 3D-vision, route planning, robot navigation, correlation function.

ОБНАРУЖЕНИЕ ПРЕПЯТСТВИЙ ПЕРЕД НАЗЕМНЫМ МОБИЛЬНЫМ ОБЪЕКТОМ В БОРТОВОЙ СИСТЕМЕ ТЕХНИЧЕСКОГО СТЕРЕОЗРЕНИЯ РЕАЛЬНОГО ВРЕМЕНИ

О. В. Выголов, С. Ю. Желтов, Ю. В. Визильтер

*Федеральное государственное унитарное предприятие «Государственный научно-исследовательский институт авиационных систем» (ФГУП ГосНИИАС),
125319, Москва, Викторенко, 7*

Предложен вычислительно эффективный метод обнаружения трёхмерных объектов на относительно гладкой поверхности по стереоскопическим изображениям сцены, получаемым с двух видеокамер наземного мобильного объекта. Суть метода заключается в том, что с построением в полярной системе координат ортогональных проекций стереоизображений на подстилающую поверхность задача обнаружения трёхмерных объектов сводится к более простой задаче обнаружения на синтезированных изображениях яркостно-геометрической структуры известной формы — совокупности вертикальных границ контрастности. Вертикальность границ позволяет организовать поиск этой структуры на основе «быстрых», аппаратно поддерживаемых процедур обработки изображений, таких как свёртка изображения с прямоугольной маской и вычисление интегральных проекций яркости изображения в вертикальном и горизонтальном направлениях. Практическая применимость метода показана на примере создания бортовой системы технического стереозрения реального времени для обнаружения препятствий перед транспортным средством при его движении по скоростным автомагистралям.

Ключевые слова: обнаружение препятствий, стереозрение, реальное время, ортофото.

Введение

Обнаружение препятствий можно отнести к одной из частных, но, безусловно, ключевых задач автономного управления наземными мобильными объектами. С её решением тесно связаны перспективы автоматизации целого ряда важных функций, таких как самопозиционирование, анализ достижимости целей управления, оперативное планирование маршрутов перемещения, построение карты исследованного пространства. Стремительное развитие средств получения цифровых изображений и прогресс вычислительной техники открыли новые возможности для решения задачи обнаружения препятствий на базе интеллектуальных систем технического зрения.

На сегодняшний день ощутимые успехи достигнуты в создании систем обнаружения препятствий при движении мобильного объекта в пределах стационарного пространства известной геометрии (офиса, здания) [Valls et al., 2006; Wang et al., 2004]. Поскольку скорости движения в таких задачах невелики, то для без-

Выголов Олег Вячеславович — начальник сектора, кандидат технических наук, e-mail: o.vygolov@gosniias.ru.

Желтов Сергей Юрьевич — член-корреспондент РАН, генеральный директор, доктор технических наук, профессор.

Визильтер Юрий Валентинович — начальник лаборатории компьютерного машинного зрения, доктор физико-математических наук, старший научный сотрудник, e-mail: viz@gosniias.ru.

опасного перемещения обычно достаточно обнаруживать препятствия лишь в непосредственной близости от мобильного объекта. В качестве препятствий в основном рассматриваются объекты, отвечающие предопределённым моделям, которые могут представлять собой как совокупность простых признаков, так и сложные структурные описания.

Другим востребованным классом систем являются системы обнаружения препятствий при движении мобильного объекта по скоростным, протяженным маршрутам в условиях априори неизвестной, динамично меняющейся окружающей обстановки, например, при движении транспортного средства (ТС) по автомагистралям. Данная задача является неоспоримо более сложной и её текущее состояние можно охарактеризовать как совокупность перспективных подходов, на базе которых созданы первые, тестовые прототипы систем обнаружения [Sun et al., 2006]. Специфика задачи заключается в высокой изменчивости фоноцелевой обстановки, практически не поддающейся формальному математическому описанию. При этом задача усложняется ещё и тем, что из-за высоких скоростей движения время реакции системы управления на изменение в окружающей обстановке должно быть минимальным, поэтому необходимо обрабатывать информацию в реальном масштабе времени.

Одним из перспективных путей решения задачи обнаружения препятствий средствами машинного зрения является подход на базе стереозрения. Важное преимущество стереоскопических систем заключается в возможности из-за различия в положении камер различать трёхмерные (3D) объекты, отстоящие от поверхности, и объекты, принадлежащие этой поверхности (блики, тени, специальный рисунок), что потенциально снижает вероятность ложных обнаружений. На сегодня основным фактором, ограничивающим применение стереозрения для обнаружения препятствий, является вычислительная сложность методов анализа стереоинформации. Как следствие, актуальной остается проблема разработки новых методов и алгоритмов обнаружения 3D-объектов, отличающихся, с одной стороны, корректностью с точки зрения стереофотограмметрии, с другой — использованием «быстрых» процедур обработки, ориентированных на архитектуру современных вычислительных платформ. Решению этой проблемы и посвящена настоящая работа.

1. МЕТОД ОБНАРУЖЕНИЯ ТРЁХМЕРНЫХ ОБЪЕКТОВ, ОСНОВАННЫЙ НА АНАЛИЗЕ РАДИАЛЬНЫХ ОРТОФОТОИЗОБРАЖЕНИЙ

Одним из робастных методов обнаружения 3D-объектов на относительно гладкой поверхности является метод «дифференциального ортофото» [Zheltoy et al., 2002], основная идея которого заключается в следующем. Если известна аналитическая модель подстилающей поверхности, то с учётом этой модели строятся ортогональные проекции левого и правого изображений стереопары на вспомогательную (например, горизонтальную) плоскость. Будем называть такие проекции «ортофото». В отсутствие трёхмерного объекта, при условии, что модель идеально описывает наблюдаемую поверхность, ортофото левого и правого изображений совпадают. Наличие 3D-объекта приводит к возникновению отклонений от модели поверхности, в результате чего на ортофото объект оказывается «спроецированным» на загораживаемую им область (рис. 1, см. с. 204).

Из-за различия в положении камер стереосистемы на разности левого и правого ортофото (дифференциальном ортофото) в области границ 3D-объекта возникают

характерные яркостно-геометрические структуры угловой формы (рис. 2). Таким образом, задача обнаружения 3D-объектов может быть сведена к более простой — поиску на дифференциальном ортофото двумерной структуры угловой формы. Однако решение и этой задачи сопряжено с ощутимыми вычислительными затратами и труднореализуемо в системах реального времени. В данной работе метод дифференциального ортофото получил свое дальнейшее развитие, позволившее преодолеть указанное выше ограничение.

Пусть на подстилающей поверхности выбрана внешняя система координат ($OX_e Y_e Z_e$) такая, что ось OX_e направлена вперед по ходу движения мобильного объекта, ось OY_e перпендикулярна оси OX_e и ориентирована вправо, ось OZ_e направлена вверх по нормали к поверхности, плоскость $OX_e Y_e$ совпадает с плоскостью подстилающей поверхности. Пусть также имеется пространственная модель подстилающей поверхности $Z(X, Y)$.

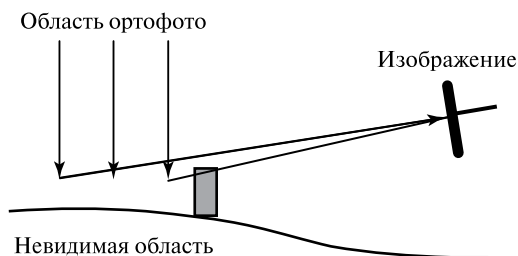


Рис. 1. На ортофото объект «проецируется» на загораживаемую им область

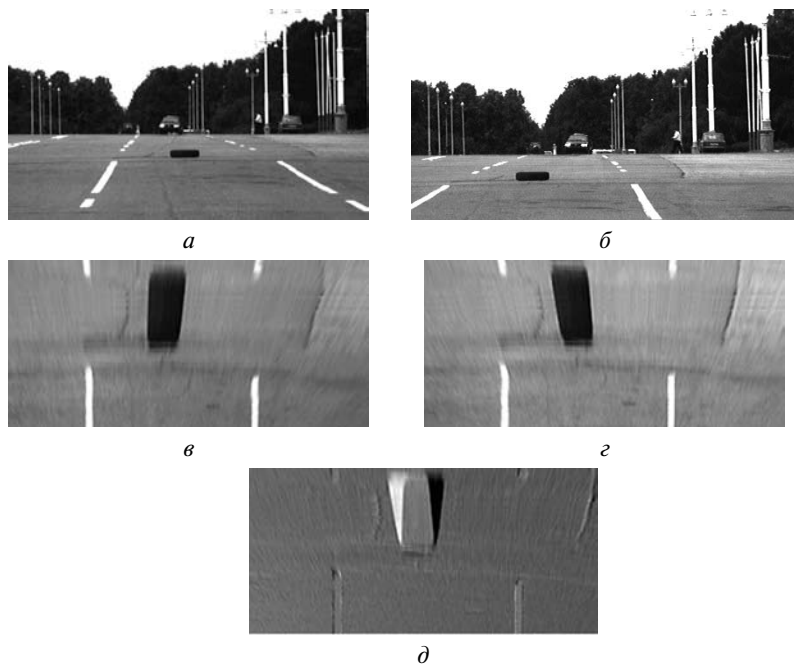


Рис. 2. Пример ортофото и их разности: *a, б* — левое и правое изображения; *в, г* — левое и правое ортофото; *д* — разность «*в*» минус «*г*» (дифференциальное ортофото)

Введем полярную систему координат (α, R) :

$$R = \sqrt{(X - X_s)^2 + (Y - Y_s)^2}, \quad \alpha = \arctan \frac{Y - Y_s}{X - X_s},$$

где R — расстояние от фокальной плоскости камеры до 3D-точки (X, Y, Z) ; X_s, Y_s — координаты оптического центра камеры; α — угол между проекциями на горизонтальную плоскость оптической оси камеры и линией, соединяющей оптический центр с точкой (X, Y, Z) .

Построим ортофотоизображение в полярной системе координат (α, R) . В дальнейшем будем называть такое изображение «радиальным ортофото». Пиксельные координаты (i, j) точки радиального ортофото в системе координат (α, R) соответствуют точке (α_i, R_i) , вычисляемой как

$$\alpha_i = iS_\alpha, \quad R_i = jS_R,$$

где S_α, S_R — масштабирующие коэффициенты вдоль осей α и R .

Для получения значения яркости пиксела (i, j) определяются координаты точки (α_i, R_i) в системе координат $(OX_e Y_e Z_e)$:

$$X_i = R_i \cos(\alpha_i) + X_s, \quad Y_i = R_i \sin(\alpha_i) + Y_s, \quad Z_i = Z(X_i, Y_i)$$

и полученная точка (X_i, Y_i, Z_i) проецируется на исходное изображение с помощью уравнений коллинеарности.

Важное преимущество радиального ортофото заключается в следующем. В отличие от ортофото, построенного в прямоугольной системе координат, на радиальном ортофото границы объекта не искажаются, т. е. невидимая область, загораживаемая объектом, имеет прямоугольную форму с вертикальными краями. Поскольку границы объекта есть области существенного перепада яркости, то в качестве характерного признака 3D-объекта на радиальном ортофото можно рассматривать яркостно-геометрическую структуру максимально простой формы — совокупность вертикальных границ контрастности (контуров). Поиск такой структуры может быть организован на базе вычислительно эффективных операций построения и анализа интегральных проекций яркости в вертикальном и горизонтальном направлении специального «признакового» изображения, т. е. уже на начальном этапе обработки выполняется переход от дорогостоящих в вычислительном плане операций анализа изображений к более быстрым операциям анализа одномерных массивов (проекций яркости).

Признаковое изображение получается путём свёртки радиального ортофото с маской модифицированного оператора Собела для выделения вертикального края (рис. 3, см. с. 206).

Далее строится вертикальная проекция яркости $V(x)$ признакового изображения:

$$V(x) = \sum_{y=0}^{DY-1} I(x, y), \quad x = 0, \dots, DX - 1,$$

где $I(x, y)$ — яркость пиксела (x, y) признакового изображения, DX, DY — ширина и высота признакового изображения.

Из-за вертикальности контуров в проекции $V(x)$ возникают области характерной формы — локальные пики, представленные комбинацией отрезков постоянного

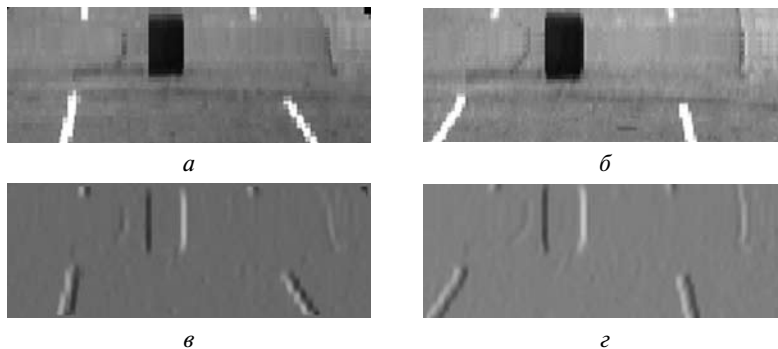


Рис. 3. Пример радиальных ортофото, полученных по стереопаре на рис. 2, и соответствующих им признаков изображений: *a, б* — левое и правое радиальные ортофото; *в, г* — левое и правое признаковые изображения

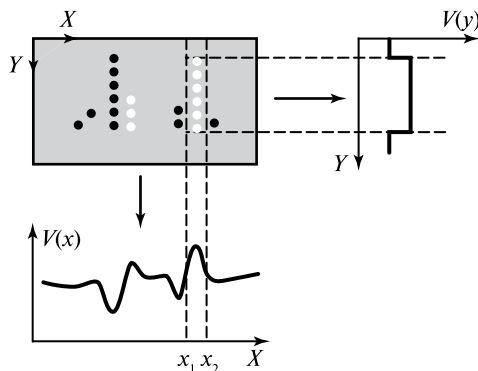


Рис. 4. Схема, показывающая соответствие вертикальных контуров областям характерной формы на вертикальной и горизонтальной проекциях яркости признакового изображения

значения и линейного возрастания-убывания (рис. 4). По положению этих пиков определяется горизонтальное положение контуров и оценивается их ширина.

Далее для каждого выделенного пика x_i в области, соответствующей ширине контура, вычисляется горизонтальная проекция яркости признакового изображения (см. рис. 4):

$$H_{x_i}(y) = \sum_{x=x_1}^{x_2-1} I(x, y), \quad y = 0, \dots, DY - 1,$$

где x_1 и x_2 — координаты левой и правой границы контура.

Концевые точки контуров в проекции $H_{x_i}(y)$ соответствуют перепадам, представленным комбинацией двух постоянных значений.

Таким образом, возникает задача поиска в одномерном массиве значимых перепадов известной формы. Приведем её решение на примере выделения локальных пиков на вертикальной проекции яркости.

Поскольку требуется не только выделить значимые перепады, но и оценить их форму, то подход на основе введения порогового значения в данном случае не

применим. Предложен метод, основанный на тестировании статистических гипотез о локальной форме сигнала. В окрестности каждой точки x близость формы сигнала к искомой модели оценивается как вероятностное отношение:

$$\rho(x) = \frac{S_{H_0}^2}{S_{H_1}^2},$$

где $S_{H_0}^2$, $S_{H_1}^2$ — несмещённые оценочные дисперсии наблюдаемого сигнала для моделей, описывающих гипотезы H_0 (сумма гауссовского шума и постоянного сигнала) и H_1 (сумма шума и полезного сигнала известной формы, представленной комбинацией отрезков постоянного значения и линейного возрастания-убывания).

Положение значимых пиков требуемой формы определяется путём поиска локальных максимумов в $\rho(x)$, превышающих пороговое значение.

Аналогичным образом определяется положение концевых точек отрезка на горизонтальной проекции яркости. Единственное отличие заключается в том, что в качестве искомой (гипотезы H_1) рассматривается форма, представленная комбинацией двух постоянных значений.

После обнаружения вертикальных контуров на левом и правом изображении для определения пространственного положения объекта необходимо решить задачу их стереотождествления. Задача решается на основе сопоставления вертикальных проекций яркости левого и правого признаковых изображений. Поскольку на проекциях яркости положение интервалов, соответствующих контурам, известно, отождествление проекций однозначно задает и отождествление контуров.

Основная идея метода заключается в установлении глобального стереосоответствия, максимизирующего суммарную меру сходства локальных интервалов вертикальных проекций яркости левого и правого признаковых изображений, т. е. в пользу каждого варианта сопоставления контура на левом изображении с контуром на правом изображении «голосуют» не только локальные признаки этой пары контуров, но и признаки других информативных областей изображения. Это позволяет снизить вероятность ложного стереотождествления в особо сложных случаях, например, когда границы объекта не имеют достаточных признаков для их локального отождествления, или при наличии на изображении областей периодической структуры (различного рода теней, бликов).

Проблема максимизации суммарной меры сходства интервалов левой и правой проекций яркости рассматривается как нахождение пути на графе. Строится двумерная сетка, узлами (вершинами) которой являются пересечения вертикальных прямых, задаваемых точками левой проекции, и горизонтальных прямых, задаваемых точками правой проекции (рис. 5, см. с. 208). Каждая вершина обладает двумя важными характеристиками: угловой диспаратностью и стоимостью, вычисляемой как

$$c(m) = \frac{\sum_{i=x_{lm}-s}^{x_{lm}+s} \sum_{j=x_{rm}-s}^{x_{rm}+s} V_l(i) V_r(j) - N \bar{V}_l \bar{V}_r}{\left[\sum_{i=x_{lm}-s}^{x_{lm}+s} V_l^2(i) - N \bar{V}_l^2 \right]^{1/2} \left[\sum_{j=x_{rm}-s}^{x_{rm}+s} V_r^2(j) - N \bar{V}_r^2 \right]^{1/2}},$$

где x_{lm} и x_{rm} — положение точки на левой и правой проекциях яркости, соответствующее вершине графа; \bar{V}_l и \bar{V}_r — среднее значение на интервалах $[x_{lm} - s, x_{lm} + s]$ и $[x_{rm} - s, x_{rm} + s]$ левой и правой проекций соответственно; $N = 2s + 1$, s — средняя ширина контуров, выделенных на этапе первичной сегментации.

Максимизация суммарной меры сходства интервалов проекций достигается нахождением на сетке глобального пути, такого, что суммарная стоимость входящих в него вершин максимальна. Для нахождения такого пути применяется метод динамического программирования. Выполняется обход вершин в направлениях слева направо, сверху вниз. Стоимость пути от начальной вершины до вершины m определяется как:

$$C(m) = c(m) + \max_{n_i < m} \{C(n_i)\}, \quad i = 1, \dots, 3,$$

где $c(m)$ — стоимость вершины m ; $C(m)$ — стоимость пути от начальной вершины до вершины n_i .

Таким образом, на каждом шаге решение принимается так, чтобы суммарная стоимость пути на данном шаге плюс суммарная стоимость на всех предыдущих шагах была максимальной. Максимум суммарной стоимости достигается на последней итерации, и оптимальный глобальный путь определяется посредством обратного «развертывания» пути от конечной к начальной вершине.

После установления стереосоответствия могут выполняться различные процедуры вторичной обработки для удаления контуров, заведомо не принадлежащих объекту, например, с высотой, меньшей порогового значения, или находящихся вне исследуемого диапазона дальностей. Также весьма эффективна независимая временная (межкадровая) фильтрация отдельных контуров с выявлением устойчивых кластеров, которые могут соответствовать реальным объектам сцены.

В следующем разделе будет показана возможность практического применения метода радиального ортофото и предложены конкретные пути его реализации в современных вычислительных средах на примере создания системы реального времени для обнаружения препятствий перед автомобилем.

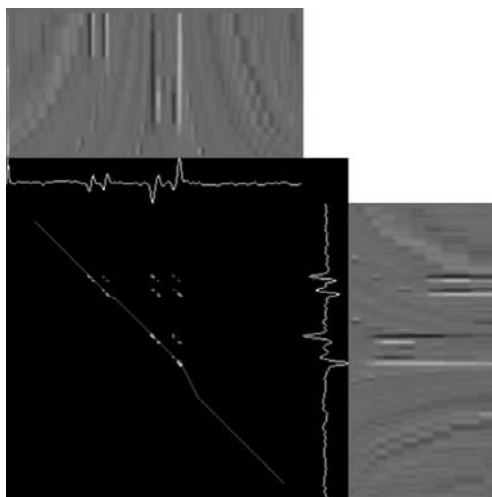


Рис. 5. Стереотождествление вертикальных проекций яркости признаков изображений как поиск пути на графе. На чёрном поле показаны вертикальные проекции яркости признаков изображений: вверху — левого, справа — правого. Яркость точки поля соответствует стоимости вершины графа. Показан путь, проходящий через вершины и задающий стереосоответствие

2. СИСТЕМА ТЕХНИЧЕСКОГО СТЕРЕОЗРЕНИЯ РЕАЛЬНОГО ВРЕМЕНИ ДЛЯ ОБНАРУЖЕНИЯ ПРЕПЯТСТВИЙ ПЕРЕД ТРАНСПОРТНЫМ СРЕДСТВОМ

2.1. Технические параметры задачи

К системе обнаружения препятствий (СОП) предъявляются следующие требования:

- 1) необходимо обнаруживать движущиеся и неподвижные непрозрачные препятствия различной формы с прямолинейными или закругленными краями, высотой над уровнем дороги не менее 10 см, шириной от 20 см до 1,5 м, находящиеся в диапазоне дальности от 5 до 100 м, в собственной полосе движения ТС;
- 2) необходимо отличать препятствия от контрастных объектов, лежащих в плоскости дороги (линии разметки, тени, блики);
- 3) система обнаружения должна быть круглосуточной и всепогодной;
- 4) обработка информации должна осуществляться в масштабе реального времени;
- 5) система обнаружения должна выдавать в систему управления ТС сигнал о наличии препятствия, а также его характеристики: дальность и положение в полосе движения ТС, ширину и высоту над уровнем дороги, относительную скорость сближения с ТС.

В состав аппаратного обеспечения СОП входит стереовидеосистема на базе двух CCD-видеокамер, устанавливаемых в области ветрового стекла ТС на высоте 106,5 см с базой 113,6 см, бортовой компьютер с платами оцифровки ТВ-изображений и спецвычислителем, поддерживающим набор функций обработки изображений на аппаратном уровне.

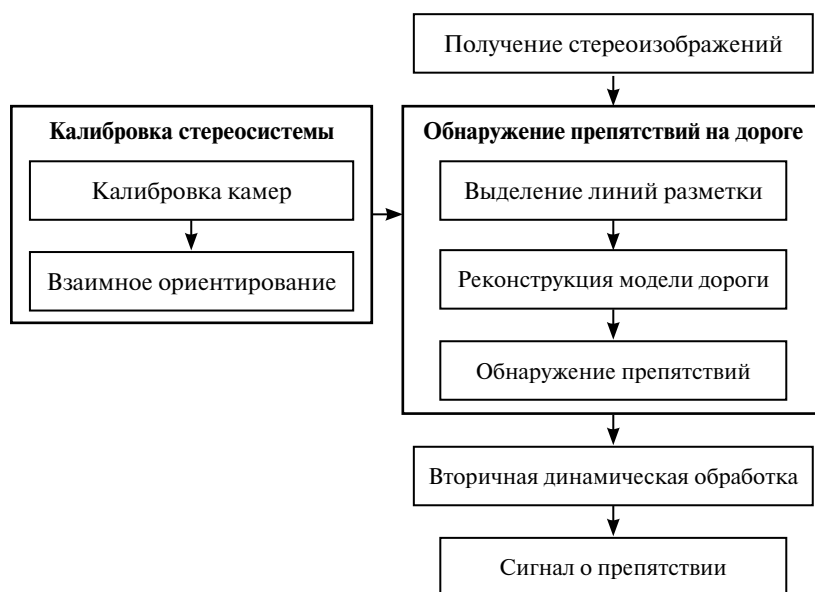


Рис. 6. Структура алгоритма обнаружения препятствий

2.2. Обобщенная структура алгоритма обнаружения препятствий

Общий подход к проблеме обнаружения препятствий на дороге, предлагаемый в настоящей работе, базируется на следующих основных принципах:

- 1) калибровка и ориентирование стереовидеосистемы осуществляется средствами цифровой фотограмметрии по оригинальной методике, предложенной в работе [Клуаз et al., 1999];
- 2) алгоритм обнаружения препятствий использует трёхмерную модель поверхности дороги;
- 3) трёхмерная модель поверхности дороги восстанавливается использованием продольных линий разметки, ограничивающих полосу движения автомобиля.

Структура алгоритма обнаружения препятствий представлена на рис. 6 (см. с. 209).

2.3. Выделение линий дорожной разметки

Алгоритм выделения линий дорожной разметки разработан в рамках так называемого подхода событийного анализа изображений, предложенного авторами ранее [Visilter et al., 1996]. Характерной особенностью процедур событийного анализа изображений является то, что яркостно-геометрическая модель объекта используется непосредственно в ходе низкоуровневого анализа изображения; при этом каждая значимая характеристика изображения рассматривается как событие, свидетельствующее в пользу гипотезы (ряда гипотез) о наличии и свойствах искомого объекта, а процесс проверки гипотез управляется событиями в том смысле, что каждое выявленное событие инициирует обработку лишь тех гипотез, на апостериорную вероятность которых данное событие может повлиять.

Ниже приводятся основные этапы процедуры выделения разметки.

1. *Первичное выделение сегментов разметки* — процедура основана на статистическом анализе профилей яркости в построчных горизонтальных сечениях изображения. Сегмент любой линии разметки на профиле соответствует существенному положительному яркостному контрасту относительно фона и возникновению резкого ступенчатого края. Сегменты с данными свойствами находятся методом сигма-фильтрации.

2. *Монокулярное обнаружение продольных линий разметки*. Обнаружение линий разметки основано на модификации алгоритма преобразования Хафа (Hough Transform). Суть предлагаемой процедуры заключается в голосовании пар сегментов разных строк, входящих в зону поиска, в пользу набора гипотез о положении и направлении линии разметки, представленных в виде массива-аккумулятора. Параметризация аккумулятора соответствует натуральной параметризации прямых линий, при которой линия определяется парой лежащих на ней точек. Однако вместо параметризации (x_{Top} , $x_{Horizontal}$), где x_{Top} — точка пересечения прямой с верхней границей зоны; $x_{Horizontal}$ — точка пересечения прямой с линией предполагаемого горизонта, используется параметризация (x_{Top} , x_{Bottom}), где x_{Top} — точка пересечения прямой с верхней границей зоны; x_{Bottom} — точка пересечения прямой с нижней границей (основанием) зоны.

Такая параметризация позволяет учитывать проективную геометрию. Пучку параллельных прямых в 3D-пространстве, лежащих в одной плоскости и на изображении сходящихся в одной точке, в пространстве (x_{Top} , x_{Bottom}) соответствует

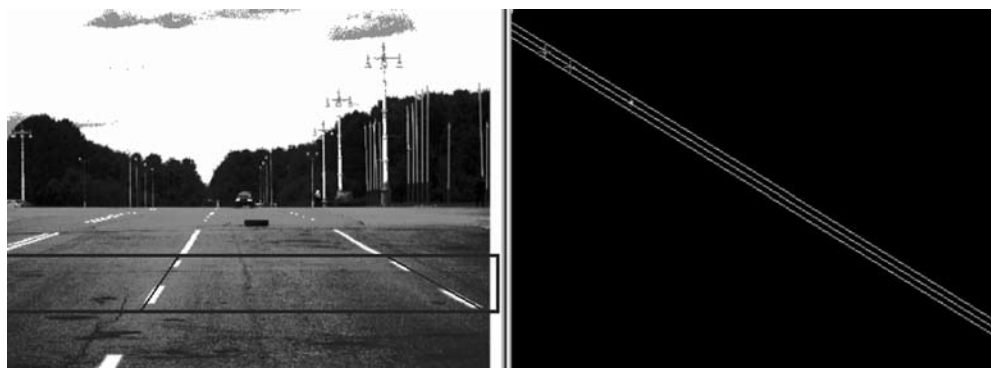


Рис. 7. Процедура монокулярного обнаружения линий разметки. Слева — исходное изображение. Показаны прямоугольная зона поиска и линии разметки, соответствующие локальным максимумам в аккумуляторе. Справа — аккумулятор модифицированного преобразования Хафа. Показана прямая, соответствующая предсказанию точки схода. Отмечены границы области точек, отстоящих от этой прямой не более чем на пороговое число пикселей

множество точек, лежащих на одной прямой (рис. 7). Поэтому при голосовании любая пара отрезков «голосует» в пользу множества прямых, таких что:

- а) прямые пересекают эти отрезки;
- б) их образы в пространстве (x_{Top}, x_{Bottom}) отстоят от прямой, соответствующей текущей гипотезе о нулевой плоскости, не более чем на заданное число пикселей.

3. *Стереотождествление линий разметки.* Для всех возможных пар линий, относящихся к разным изображениям, проверяются следующие условия: тангенсы углов наклона стереолинии относительно плоскостей $Z_e - Y_e$ и $X_e - Y_e$ и высота сдвига её стартовой точки от нулевой плоскости по оси Z_e не должны превышать пороговых значений. Пары линий, удовлетворяющие этим условиям, заносятся в список стереолиний. Линии разметки определяются на основе анализа взаимного расположения стереолиний из сформированного списка.

2.4. Обнаружение препятствий

В основе алгоритма обнаружения препятствий лежит метод радиального ортофото. На первом этапе выполняется пространственная реконструкция подстилающей поверхности (дороги). В предположении об относительной гладкости дороги для её описания была выбрана аналитическая модель:

$$Z(X, Y) = a_0 + a_1X + a_2Y + a_3Y^2,$$

где X, Y, Z — координаты точки поверхности в системе координат $(OX_e Y_e Z_e)$.

Параметры $a_i, i = 0, \dots, 3$ модели $Z(X, Y)$ определяются методом наименьших квадратов по набору трёхмерных точек дорожной разметки, которые принадлежат поверхности и заведомо не относятся к обнаруживаемому объекту.

Для построения радиального ортофото в масштабе реального времени авторами предложен вычислительно эффективный алгоритм, основанный на аппаратно поддерживаемой процедуре кусочно-билинейной трансформации изображений. Суть алгоритма состоит в следующем. Область дороги разбивается на участки

с фиксированным шагом по дальности. Каждый участок имеет четыре угловые точки, задаваемые положением дорожной разметки на текущей дальности. Для каждой точки указанного набора определяется соответствующая ей точка на трансформированном изображении.

Преобразование набора из четырёх точек исходного изображения в набор из четырёх точек трансформированного изображения, будучи записано в виде

$$x = x(u, v), \quad y = y(u, v),$$

где (x, y) — координата точки на исходном изображении; (u, v) — координата точки на трансформированном изображении, может быть реализовано в рамках билинейной формы вида:

$$x = Bu + Cuv + Ev + F, \quad y = Hu + Kuv + Mv + N,$$

где B, C, E, F, H, K, M, N — коэффициенты преобразования.

Алгоритм стереоотожествления вертикальных контуров использует условие расположения препятствия в собственной полосе движения ТС, что позволяет существенно уменьшить область поиска пути, задающего стереосоответствие. По аналогии с алгоритмом построения радиального ортофото область дороги разбивается на участки, ограниченные линиями дорожной разметки. На каждом участке область поиска пути, задающего стереосоответствие, ограничивается:

- прямыми, отвечающими значениям угловой диспаратности $A(D_{\min})$ и $A(D_{\max})$ на минимальной D_{\min} и максимальной D_{\max} дальности участка;
- положением собственной полосы движения ТС на минимальной дальности участка (рис. 8а).

В качестве результирующей области поиска рассматривается объединение областей, полученных на отдельных участках (рис. 8б). На форму искомого пути

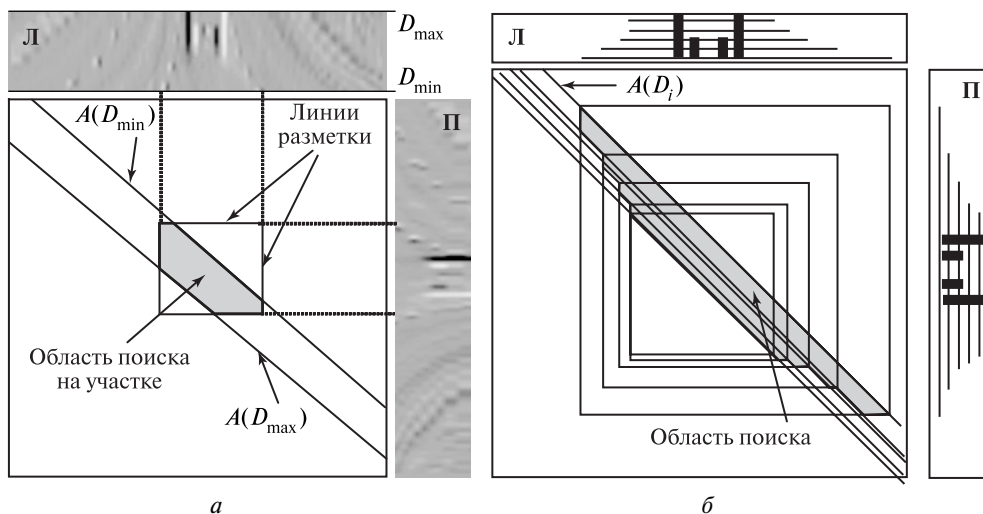


Рис. 8. Построение области поиска пути, задающего стереосоответствие: а — для интервала дальностей $[D_{\min}, D_{\max}]$ (показаны фрагменты левого (Л) и правого (П) признаковых изображений для $[D_{\min}, D_{\max}]$ и положение линий дорожной разметки); б — для всего диапазона дальности

дополнительно накладываются геометрические ограничения, вытекающие из условий однозначности отождествления и упорядоченности отрезков.

Полученные стереоконтуры передаются в блок многоканальной вторичной обработки информации, отличительной особенностью которой является сопровождение отдельных стереоконтуров, а не объектов в целом. Преимущество такой обработки проявляется в том, что в результирующей сборке гипотез участвуют только те контуры, информация о которых подтверждена последовательностью кадров и динамические атрибуты которых укладываются в рамки реально возможных перемещений исследуемых объектов.

2.5. Примеры работы системы

На рис. 9–13 (см. также с. 214) приведены примеры работы системы в различных условиях. Показаны левое и правое изображения. Обнаруженные линии разметки показаны чёрным цветом. Найденное препятствие — белый прямоугольник.



Рис. 9. Пример работы системы



Рис. 10. Пример работы системы при маневре препятствия

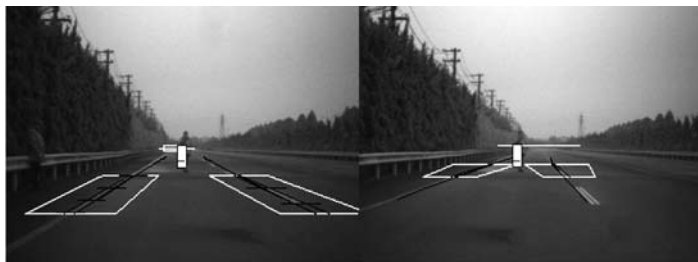


Рис. 11. Пример обнаружения пешехода

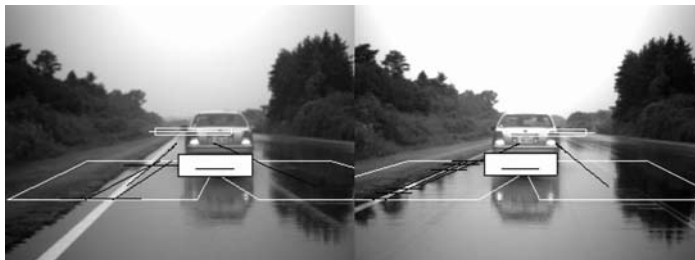


Рис. 12. Пример работы системы в сложных условиях мокрой дороги

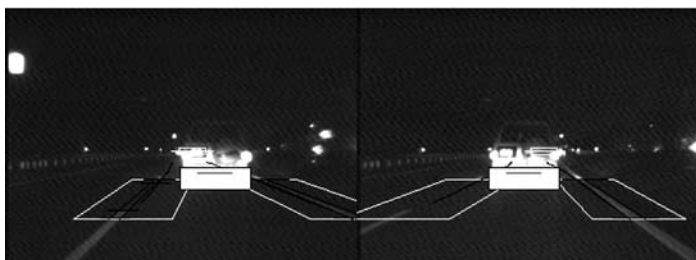


Рис. 13. Пример работы системы в ночных условиях

Выводы

Предложенный в работе метод обнаружения трёхмерных объектов на основе анализа радиальных ортофотоизображений был протестирован на обширной базе реальных дорожных сцен в системе обнаружения препятствий перед транспортным средством. Достоинства метода наиболее ярко проявились при работе системы в особо сложной информационной обстановке: при наличии различного рода засветок, бликов и теней на подстилающей поверхности, загораживании части препятствия, его резком маневрировании, ускорении или торможении. На имеющейся базе регистраций, включающей и указанные выше сложные случаи, вероятность обнаружения препятствий составила 0,97, вероятность ложного срабатывания — 0,02.

Алгоритмическая реализация метода полностью согласуется с архитектурой современных платформ обработки изображений, что позволяет выполнять наиболее затратные в вычислительном плане процедуры обработки изображений на аппаратном уровне, а также распараллеливать процесс вычислений и избегать хранения промежуточных результатов. Показатели быстродействия разработанных алгоритмов полностью отвечают требованиям систем реального времени.

Всё вышесказанное позволяет сделать положительный вывод о практической применимости предложенного метода и алгоритмов в перспективных системах управления сложными техническими объектами, имеющими видеодатчики.

Литература

Knyaz V. A., Zheltov S. Yu., Stepanyants D. G. (1999) Method for accurate camera orientation for automobile photogrammetric system // Proc. Intern. Workshop on Mobile Mapping Technology. Bangkok, Thailand. Apr. 21–23, 1999. P. 4-3-1–4-3-6.

- Valls M., Zhou W., Dissanayake G. (2006) Towards vision based navigation in large indoor environments // Proc. 2006 IEEE/RSJ Intern. Conf. Intelligent Robots and Systems (IROS 2006), 2006. P. 2096–2102.*
- Vizilter Yu., Zheltov S., Stepanov A. (1996) Events-based Image Analysis for Machine Vision and Digital Photogrammetry // SPIE Proc. International Archives of Photogrammetry and Remote Sensing. 1996. V. 31. Pt. 5.*
- Wang J., Kimura H., Sugisaka M. (2004) Intelligent control for the vision-based indoor navigation of an A-life mobile robot // Artificial Life and Robotics. 2004. V. 8. N. 1. P. 29–33.*
- Sun Z., Bebis G., Miller R. (2006) On-Road Vehicle Detection: A Review // IEEE Trans. Pattern Analysis and Machine Intelligence. May 2006. V. 28. N. 5. P. 694–711.*
- Zheltov S. Yu., Sybiryakov A. V., Vygolov O. V. (2002) Car collision avoidance system based on orthophoto transformation // Intern. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Greece, 2002. V. 34. Pt. 5. P. 125–130.*

MOBILE ROBOT STEREOVISION SYSTEM FOR REAL-TIME OBSTACLE DETECTION

O. V. Vygolov, S. Yu. Zheltov, Yu. V. Vizilter

*The Federal State Unitary Enterprise “State Research Institute of Aviation Systems”,
Russia, Moscow*

In this paper the computationally efficient method of 3D-objects detection is proposed. The object is located on a relatively smooth surface of known analytical model and registered stereoscopically by stereovision system of mobile robot. The key idea of proposed method is to reduce the problem of 3D-objects detection to the problem of simple 2D-structure detection on special orthogonal projections of stereo images to the underlying surface. Projections are generated in polar coordinate system (polar orthophoto). 3D-object raised above the surface corresponds to simple cluster of vertical straight-line edges on polar orthophotos. To find these edges the hardware supported image processing procedures are proposed such as image convolution and brightness projections. Implementation of the method is considered for real-time car collision avoidance system.

Keywords: obstacle detection, stereovision, real-time, orthophoto.

Vygolov Oleg Vyacheslavovich — Head of Group, Candidate of Technical Science, e-mail: o.vygolov@gosniias.ru.

Zheltov Sergey Jurevich — Corresponding Member of the Russian Academy of Sciences, General Director, Doctor of Technical Science, Professor.

Vizilter Yury Valentinovich — Head of the Laboratory of Computer Machine Vision, Doctor of Physical and Mathematical Sciences, Senior Scientist, e-mail: viz@gosniias.ru.

ГИПЕРСПЕКТРАЛЬНАЯ СИСТЕМА ЗРЕНИЯ ДЛЯ МОБИЛЬНЫХ РОБОТОВ

А. М. Овчинников, А. К. Платонов

*Учреждение Российской академии наук Институт прикладной математики
им. М. В. Келдыша РАН (ИПМ им. М. В. Келдыша РАН),
125047, Москва, Миусская пл., 4*

Описывается макетный образец гиперспектрометра, разработанный для исследования особенностей характеристик гиперспектральных сигналов, получаемых при наблюдении близко расположенных объектов. Такие исследования необходимы при создании сенсоров пространственно-спектральных характеристик окружения мобильных роботов или других распознающих мехатронных автоматов. Приводится модель гиперспектра, получаемого при наблюдении близких объектов, которая позволяет учесть влияние дисторсионных явлений в оптическом канале гиперспектрометра, заметно отличающихся «локальный гиперспектр» от гиперспектральных данных, получаемых при дистанционном зондировании объектов Земли с самолётов и спутников. Описываются результаты исследования достижимых характеристик локального гиперспектрометра и экспериментов спектрального анализа объектов в его ближайшем окружении.

Работа поддержана РФФИ (проект № 08-01-00908).

Ключевые слова: спектроскопия, гиперспектр, гиперспектрометр, спектральная сигнатура, спектральная база, спектральные эталоны, мехатроника, мобильные роботы.

Введение

Работа направлена на расширение состава информации, получаемой от зрительного сенсора мобильного робота при его движении в условиях бездорожья. Острота этой проблемы связана с необходимостью автоматического получения надёжной оценки свойств опорной проходимости видимых сенсором областей местности, по которой происходит движение. В развитых системах зрения мобильных роботов задача оценки профильной проходимости успешно решается по сигналам лазерного 3D-дальномера, а профильно-опорной проходимости — совместной обработкой сигналов дальномера и телевизионной видеокамеры [Носков, Рубцов, 2005]. Трудности зрительного распознавания в режиме реального времени именно опорных характеристик видимых объектов делают желательным обогащение получаемых от видеокамеры фотометрических измерений яркости сцены дополнительными измерениями спектрального диапазона длин волн получаемого панхроматического сигнала.

Принципиальной основой попыток использования спектрального анализа наблюдаемых объектов является предположение о соответствии спектра *отражённого* сигнала составу материала отражающей поверхности, т. е. о наличии у каждого материала его спектральной «сигнатуры». Это достаточно сильное предположение, поскольку характеристики отражённого сигнала зависят не только от материала

Овчинников Андрей Михайлович — аспирант, e-mail: akp31mail@gmail.com.

Платонов Александр Константинович — заведующий сектором, доктор физико-математических наук, профессор, e-mail: platonov@keldysh.ru.

наблюдаемого объекта. Связанные с этим трудности распознавания материала видимой поверхности при использовании цветных видеокамер с фотометрией всего трёх каналов спектра видеосигнала породили желательность более тонкого исследования свойств многоканального спектрального сигнала. Есть надежда, что при достаточно большом числе спектральных каналов измерение *изменений* интенсивности спектра отражённого сигнала позволит получить различные результаты для различных отражающих материалов. Для этого имеются следующие удачные примеры.

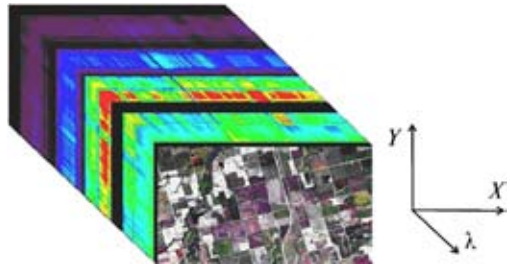


Рис. 1. Пример гиперкуба

Одним из хорошо известных и достаточно эффективных методов спектрального исследования веществ и материалов на поверхности Земли является её дистанционное *гиперспектральное* зондирование авиационными и космическими средствами — наблюдение земной поверхности в различных диапазонах электромагнитного спектра. Достоинством таких гиперспектральных измерений является одновременное описание пространственных и спектральных свойств наблюдаемой сцены в виде её «гиперкуба» этих данных. Техника получения такого «гиперкуба» (трёхмерной функции яркости спектра $\Phi(\lambda, x, y)$, для каждого пиксела $\{x, y\}$ фотометрического изображения видимой в поле зрения спектрометра части окружающего пространства — рис. 1) в настоящее время находит широкое применение для целей дистанционного зондирования Земли (<http://eo1.gsfc.nasa.gov/Technology/>)



Рис. 2. Макет гиперспектрометра

Hyperion.html) и Марса (<http://sci.esa.int/marsexpress>). Это свойство — одновременной привязки пространственных и спектральных данных — крайне ценно для сенсоров и систем распознавания операционной среды движущихся машин.

За последние два десятка лет в мире было создано довольно большое число удачно функционирующих гиперспектрометров авиационного и космического базирования. Однако эти дорогие* системы и, главное, их программное обеспечение с накопленными базами сигнатур (спектральных эталонов) не могут быть использованы в задачах локального зондирования, где измерения должны выполняться с более мелким пространственным разрешением на расстояниях порядка 1,5 м. Это потребовало специальной разработки недорогого «локального гиперспектрометра» со специальными настройками его оптического канала и требуемым набором программных средств. Описываемый ниже легко переносимый** макет такого локального гиперспектрометра (рис. 2, см. с. 217) предназначен и используется в настоящее время для исследования спектральных характеристик элементов среды движения робота и накопления нужных баз их сигнатур.

СХЕМА ГИПЕРСПЕКТРОМЕТРА

Стандартная схема гиперспектрометра (рис. 3) содержит два объектива (1 и 2), фокальные плоскости которых совпадают. Эти объективы (если отвлечься от потерь и искажений в стеклянных элементах оптического тракта) при соответствующей настройке фокусных расстояний реализуют преобразование входного оптического потока со сфокусированным изображением объекта, находящегося на конечном расстоянии от входного зрачка, в выходной оптический поток с мнимым изображением этого объекта на требуемом удалении.

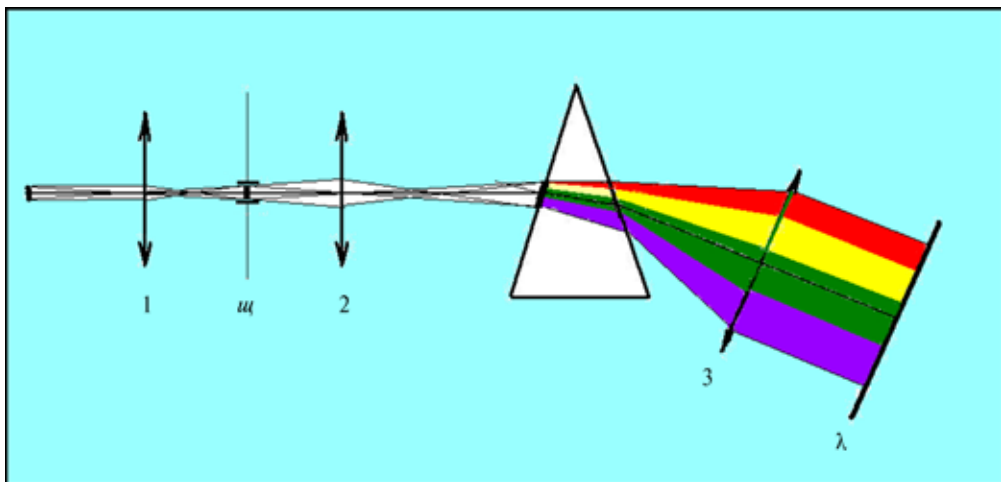


Рис. 3. Принципиальная схема гиперспектрометра

* Типовой проект гиперспектральной программы оценивается в несколько сотен миллионов долларов.

** Суммарная масса 2,5 кг, габариты 480×250×120 мм.

Если между объективами 1, 2 в их фокальной плоскости поместить щелевую диафрагму ω , а за объективом 2 расположить в нужной ориентации диспергирующий элемент (стеклянную призму или дифракционную решетку), то на входную плоскость последнего будет падать одномерная полоска изображения объекта. В результате на выходе диспергирующего элемента будет получен двумерный оптический поток со спектральными картинками каждой точки линии входной полоски (например, плоские горизонтальные сечения $\{x, \lambda\}$ гиперкуба на рис. 1). Эту двумерную сцену спектрального разложения точек щели можно увидеть с помощью объектива 3 и фотоприёмной матрицы за ним.

Если все три объектива вместе с диафрагмой и матрицей двигать или поворачивать в направлении, перпендикулярном оси щели диафрагмы, то на диспергирующем элементе будет выполняться щелевое сканирование изображения объекта с получением последовательности двумерных изображений спектральных сечений гиперкуба в нормальном для них направлении (т. е. — последовательности плоскостей $\{x, \lambda\}$ в направлении оси y гиперкуба на рис. 1). Использование цифровой видеокамеры на выходе диспергирующего элемента обеспечивает оцифровку и запоминание последовательности щелевых плоскостей $\{x, \lambda\}$ гиперкуба в виде последовательности кадров видеофильма с результатами сканирования изображения объекта.

Таким образом, полный гиперкуб в виде последовательности его горизонтальных сечений возникает либо при известном движении робота, либо при качании прибора. Это, с одной стороны, позволяет в данной реализации исключить вторую видеокамеру, обычно используемую в системах зондирования Земли для привязки спектральной информации к пространственным координатам наблюдаемой видеосцены. С другой стороны, этот способ требует создания программных средств преобразования структуры данных получаемого видеопотока в изображение видимого объекта (фронтальной плоскости на рис. 1) с возможностью доступа к спектру каждого пикселя такого изображения. Заметим, что в лабораторных условиях для исследования малых объектов удобным вариантом сканирования экспериментальных поверхностей является установка поворотного зеркала с программируемым управлением.

ОСОБЕННОСТИ ЛОКАЛЬНОГО ГИПЕРСПЕКТРОМЕТРА

Гиперспектральные измерения близких объектов обладают свойством повышенной дисторсии спектрального канала — зависимости спектрального отклонения лучей не только от длины волны, но и от угла наблюдения, а также от расстояния до близкого объекта (рис. 4, см. с. 220).

При фиксированном относительном расположении оптических элементов гиперспектрометра в общем положении близкого объекта его изображение оказывается не в фокусе, и в рамках модели геометрической оптики ход лучей отличается от идеальной схемы рис. 3. Борьба с этим явлением основана на оптимальном выборе параметров объективов и ширины щели, с применением модели хода лучей и с последующим использованием итерационного алгоритма коррекции спектральной дисторсии, основанного на модели связи пространственных и спектральных характеристик пиксельных картин. Это выполняется с помощью следующих соотношений.

Пусть точка A_1 на рис. 4 является изображением точки A_0 после собирающей линзы L_1 . Тогда координаты x_1 и y_1 точки A_1 в системе координат линзы L_1 , по

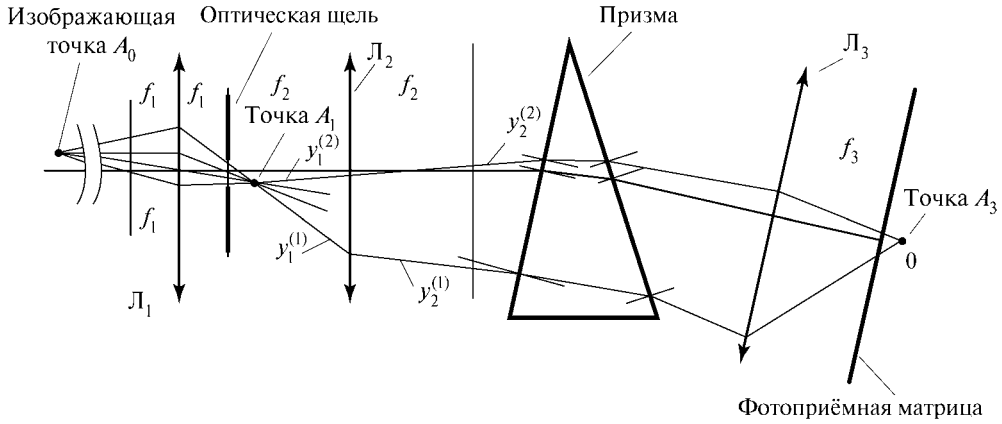


Рис. 4. Модель искажения хода лучей локального гиперспектрометра

правилам геометрической оптики, определяются через условия пересечения двух лучей (верхним индексом в скобках обозначается номер луча):

$$y_1^{(1)} = -\frac{h}{L}x, \quad y_1^{(2)} = -\frac{h}{f_1}x + h$$

по следующим общим формулам:

$$x_1 = f_1 \frac{1}{1 - f_1/L}, \quad y_1 = -h \frac{f_1/L}{1 - f_1/L}.$$

Здесь L — расстояние по оси Ox от точечного объекта до линзы L_1 ; h — отклонение наблюдаемой точки объекта от оси Ox ; f_1 — её фокусное расстояние (другие обозначения показаны на рис. 4). Запишем координаты точки A_1 относительно правого фокуса линзы L_1 (эти координаты обозначены верхним индексом «0» без скобок):

$$x_1^0 = f_1 \frac{f_1/L}{1 - f_1/L}, \quad y_1^0 = -h \frac{f_1/L}{1 - f_1/L}.$$

Введем безразмерный параметр $\delta = f_1/L$, характеризующий степень удаленности точечного объекта от линзы. Бесконечному удалению отвечает значение $\delta = 0$, при расположении объекта в фокальной плоскости $\delta = 1$. Для достаточно большого удаления параметр δ является малым параметром исследования погрешностей гиперспектрометра. Если точка A_1 лежит близко к фокальной плоскости, то $\delta \ll 1$, и тогда можно записать

$$x_1^0 = f_1 \delta + O(\delta^2), \quad y_1^0 = -h \delta + O(\delta^2).$$

Если $h \ll f_1$, то можно пренебречь смещением изображения точки A_1 из главной фокальной плоскости, т. е. считать расстояние $x_1^0 = 0$. В этом предположении точка A_1 лежит в главной фокальной плоскости, но она несколько смещена от главной оптической оси системы. Пусть апертура линзы L_2 равна величине d_2 , по определению она равна ширине плоскопараллельного пучка, порождаемого точ-

кой A_1 . В принятых предположениях пучок лучей, приходящих в точку A_1 , действительно стал плоскопараллельным — из-за пренебрежения смещением точки A_1 от главной фокальной плоскости первой линзы (его параллельность обеспечивается принадлежностью точки A_1 фокальной плоскости).

Построим такой пучок лучей, порождаемый точкой A_1 и лежащий между прямыми $y_2^{(1)}(x)$, $y_2^{(2)}(x)$ на рис. 4. Ось пучка задается прямой

$$y_2(x) = \frac{-y_1^0}{f_2} [x - (f_1 + f_2)].$$

Выражения для новых границ пучка лучей в точке A_1 имеют вид

$$\begin{aligned} y_2^{(1)} &= y_2 - d_2 = \frac{-y_1^0}{f_2} [x - (f_1 + f_2)] - d_2, \\ y_2^{(2)} &= y_2 - d_2 = \frac{-y_1^0}{f_2} [x - (f_1 + f_2)] + d_2. \end{aligned} \quad (1)$$

Следует отметить, что речь идет о монохромном пучке, т. е. точка A_0 порождает луч одной длины волны и линза не подвержена хроматической аберрации. Поэтому дисперсионным элементом рассматриваемой оптической системы является только оптическая призма.

Параллельный пучок, ограниченный прямыми (1), падает на собирающую линзу L_2 с фокусным расстоянием f_2 . Угол α_d , который составляет осевая линия этого пучка с оптической осью линзы L_2 , приближённо вычисляется по формуле

$$\alpha_d = \frac{y_1^0}{f_2},$$

в предположении, что образ A_1 точки A_0 также лежит в фокальной плоскости второй линзы L_2 и отстоит от её оси на расстоянии, значительно меньшем, чем её фокусное расстояние f_2 .

Искомое результирующее смещение образа точки A_2 на фотоматрице от точки «0» на рис. 4 обусловлено отклонениями и несимметричностью пучка лучей, порожденного точкой A_1 , проходящего через призму и собираемого линзой L_3 , от направлений главных оптических осей системы до и после призмы. Чтобы определить величину этого смещения, воспользуемся малостью отклонения точки A_1 от главной оптической оси. В этом предположении угол, под которым происходит падение луча на поверхность входной грани призмы, вычисляется так (согласно рис. 5, см. с. 222):

$$\alpha_{1d} = \alpha_1 - \alpha_d.$$

Пользуясь предположением о малости угла α_d , найдем отклонение луча, выходящего из выходной грани призмы, варьируя легко выводимую формулу (см. рис. 5) прохождения через призму луча вдоль её главной оптической оси (n_2 — коэффициент преломления в стеклянной призме, отнесённый к коэффициенту преломления в воздушной среде):

$$\sin \alpha_2 = n_2 \sin \beta_2 = n_2 \sin(\beta_1 + \gamma) = \sin \alpha_1 \cos \gamma + \sqrt{n_2^2 - \sin^2 \alpha_1} \cdot \sin \gamma. \quad (2)$$

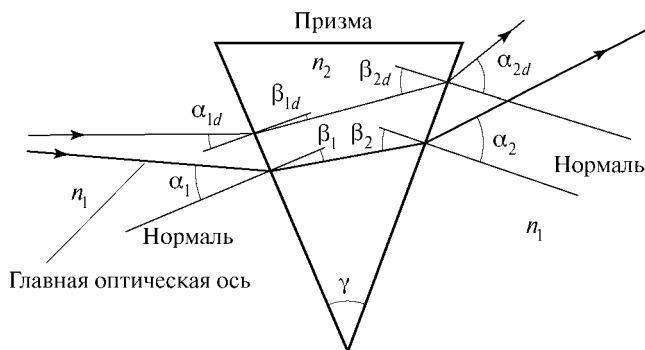


Рис. 5. Прохождение луча, порождаемого точкой A_0 (относящиеся к этому лучу углы имеют индекс d), на этом рисунке $\alpha_1 < 0$ и $\beta_1 < 0$, а $\alpha_2 > 0$ и $\beta_2 > 0$

Имеем:

$$\Delta\alpha_2 \cos\alpha_2 = \left(\cos\alpha_1 \cos\gamma - \frac{\sin\alpha_1 \cos\alpha_1}{\sqrt{n_2^2 - \sin^2\alpha_1}} \sin\gamma \right) \Delta\alpha_1.$$

Заменяя $\Delta\alpha_2$ на $-\alpha_d$ (минус перед α_d появляется из-за выбранного направления отклонения луча, порождаемого точкой A_1), получаем величину отклонения луча от направления главной оптической оси при выходе из призмы:

$$\Delta\alpha_2 = -\frac{1}{\cos\alpha_2} \left(\cos\alpha_1 \cos\gamma - \frac{\sin\alpha_1 \cos\alpha_1}{\sqrt{n_2^2 - \sin^2\alpha_1}} \sin\gamma \right) \Delta\alpha_1.$$

Знание этого угла позволяет вычислить положение точки A_2 относительно начала отсчёта на фотоприемной матрице. Действительно, этот луч проходит без искривления через середину линзы L_2 , пересекая фокальную плоскость линзы, в которой и расположена фотоприемная матрица. Параллельный ему пучок также фокусируется в той же точке A_2 .

Таким образом, искомая погрешность Δy_m отображения длины волны на фотоматрице в условиях локальных измерений гиперспектрометра при малых значениях угла $\Delta\alpha_2$ определяется с учётом предыдущего из (f_3 — фокусное расстояние линзы L_3):

$$\Delta y_m = f_3 \Delta\alpha_2.$$

ОТОБРАЖЕНИЕ НА ФОТОМАТРИЦЕ ДЛИНЫ ВОЛНЫ ИЗЛУЧЕНИЯ ТОЧКИ A_0

Рассмотрим изображение на фотоматрице точки A_0 , лежащей на главной оптической оси, с длиной волны её излучения, отличающейся на $\Delta\lambda$ от некоторой исходной длины волны λ , которой по определению соответствует главная оптическая ось системы и точка «0» на фотоматрице (см. рис. 4). Угол α_2 (рис. 6, см. с. 223) в этом случае будет меняться из-за разных коэффициентов преломления света при прохождении луча через призму лучей с разными длинами волн. Варьируя формулу (2) по длинам волн λ , получаем

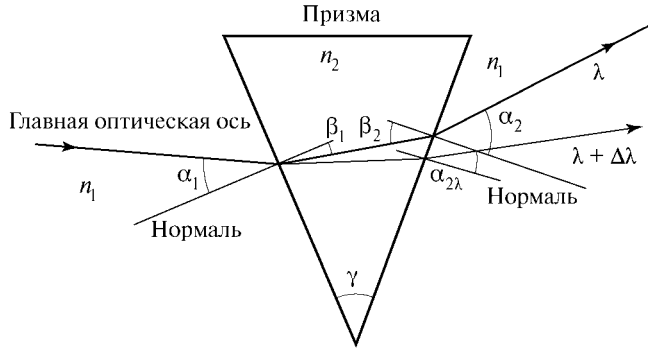


Рис. 6. Отклонение луча из-за отличия длины волны от исходной длины волны

$$\Delta\alpha_{2\lambda} = \frac{1}{\cos\alpha_2} \frac{n_2 \frac{\partial n_2}{\partial \lambda}}{\sqrt{n_2^2 - \sin^2\alpha_1}} \sin\gamma\Delta\lambda,$$

где $n_2 \frac{\partial n_2}{\partial \lambda}$ зависит от марки стекла призмы.

По аналогии с предыдущим имеем выражение для отклонения луча на фотоприемной матрице при изменении длины волны отражённого света: $\Delta y_{m\lambda} = f_3 \Delta\alpha_{2\lambda}$.

Суммарное отклонение луча Δy_m из-за изменения длины волны и бокового смещения точки A_0 в линейном приближении пропорционально фокусному расстоянию видеокамеры:

$$\Delta y_m = f_3 \left(\Delta\alpha_2 + \Delta\alpha_{2\lambda} \right).$$

Приведенная модель оптического тракта гиперспектрометра позволяет:

- оценить спектральное разрешение прибора;
- подобрать оптимальный параметр ширины щели d_2 ;
- определить размеры пятна размытия для любой длины волны;
- оценить наложение этих пятен друг на друга;
- задать в программном обеспечении размер биннинга* для коррекции шумовой погрешности сигнала, вызванной собственными шумами матрицы и наложением соседних каналов спектра.

КОРРЕКЦИЯ ДИСТОРСИИ ПРИЗМЕННОГО ГИПЕРСПЕКТРОМЕТРА

Дисторсия — аберрация оптической системы, возникающая из-за разницы длины пути лучей с их преломлением в толще стёкол, в результате чего интенсивность света различается в разных точках изображения. Этот эффект необходимо учитывать при анализе данных гиперспектрометра, поскольку дисторсия также приводит к нарушению соответствия места на фотоматрице наблюдаемой интенсивности света и его длины волны. Одним из источников дисторсии гиперспектрометра

* Биннинг — режим считывания соседних пикселей, при котором интенсивность сигнала на них интегрируется и присваивается каждому из этих пикселей так, что все они считаются «одним пикселем» изображения.

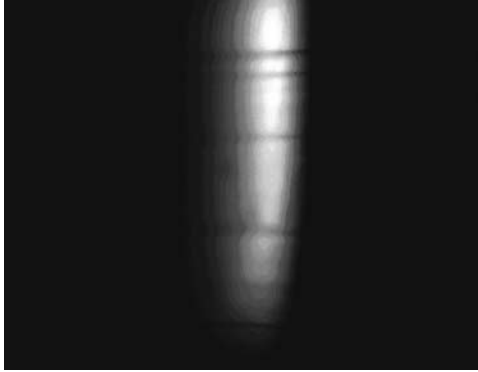


Рис. 7. Искажённое изображение. На этом рисунке хорошо видны тёмные линии

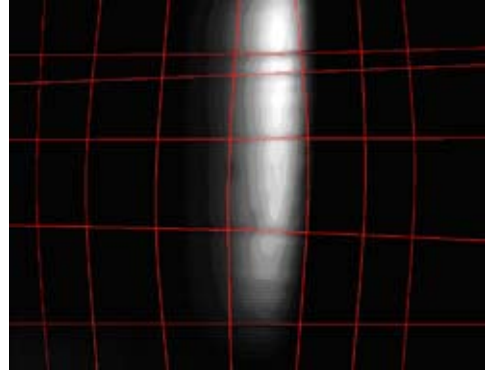


Рис. 8. Выделение реперных линий

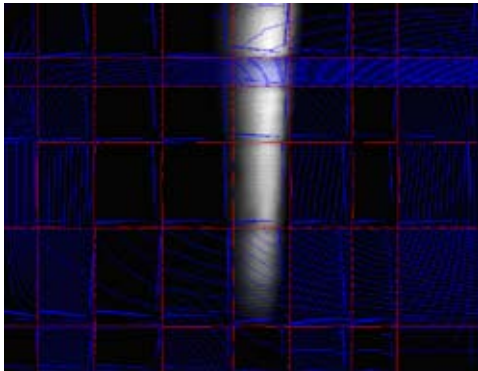


Рис. 9. Результат пиксельного округления исправленных координат (\hat{x}, \hat{y})

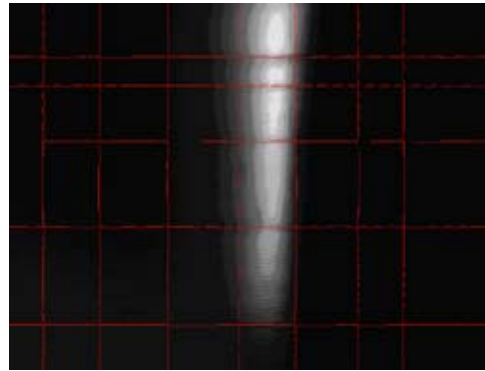


Рис. 10. Пример исправленного изображения

является использование призмы, что приводит к необходимости разработки алгоритма, позволяющего исправить изображение: перевести все точки полученного кадра фотоматрицы в точки, которые были бы найдены в случае использования идеальной оптической системы.

Для определения «правильного» положения спектральных точек можно использовать линии видеоплоскости гиперкуба, соответствующие видимым линейным протяженным предметам в поле видимости гиперспектрометра (рис. 7).

Поскольку известно, что тёмные линии, видимые на рис. 7, должны быть параллельны соответствующим краям изображения, можно создать набор точек, координаты которых известны как для искаженного, так и для исправленного изображения. На рис. 8 эти линии выделены красным цветом.

Для исправления изображения будем искать закон преобразования его координат

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = F \begin{pmatrix} x \\ y \end{pmatrix},$$

где (\hat{x}, \hat{y}) — координаты точки исправленного изображения, (x, y) — начального. Предполагая этот закон линейным, будем искать его отдельно для каждой области изображения, ограниченной красными линиями. Используя выражение линейной выпуклой комбинации вершин реперных четырёхугольников, имеем:

$$\begin{aligned}x &= x_1(1-a)(1-b) + x_2a(1-b) + x_3ab + x_4(1-a)b, \\y &= y_1(1-a)(1-b) + y_2a(1-b) + y_3ab + y_4(1-a)b.\end{aligned}$$

Здесь (x_i, y_i) — координаты вершин области, $a = \frac{\hat{x} - \hat{x}_1}{\hat{x}_2 - \hat{x}_1}$ и $b = \frac{\hat{y} - \hat{y}_1}{\hat{y}_3 - \hat{y}_1}$.

Приведенная зависимость позволила уточнять координаты (\hat{x}, \hat{y}) решением системы нелинейных уравнений методом Ньютона по выбранным значениям пиксельных координат (x, y) . Однако оказалось, что при округлении полученных координат не все точки исправленного изображения получают прообраз в старой системе. На рис. 9 (см. с. 224) синим цветом показаны точки, для которых пары (\hat{x}, \hat{y}) не было найдено.

Во избежание появления в исправленном изображении «пустых» точек яркость в них принималась равной осредненной яркости соседних пикселей. Результат работы представлен на рис. 10 (см. с. 224).

КОНСТРУКЦИЯ МАКЕТА ЛОКАЛЬНОГО ГИПЕРСПЕКТРОМЕТРА

В основе оптической части гиперспектрального прибора лежит описанная выше широко распространенная схема, изложенная в патенте США (United States Patent 7,199,876). Конструкция гиперспектрометра содержит два объектива, выбранных по соотношению цена / качество.

Регистрация информации происходит с помощью цифровой камеры с выводом на компьютер через интерфейс USB 2.0. Получаемые данные записываются на жесткий диск и одновременно выводятся на экран оператора.

В макете локального гиперспектрометра использовался одноплатный компьютер с установленной операционной системой Microsoft Windows XP и аккумуляторной батареей, дающей возможность автономной работы. В другом варианте использовался ноутбук.

Прибор обеспечивает исследования спектральных характеристик элементов среды движения робота в ближнем ИК-диапазоне 0,4...1 мкм и накопления нужных баз их сигнатур на жестком диске ноутбука. Разработанное программное обеспечение макета локального гиперспектрометра позволяет реализовать приём данных в одном из трёх вариантов рабочих режимов (табл. 1, см. с. 226).

1. В периодическом режиме запуск экспозиции происходит циклически с заданным периодом. Величина периода устанавливается программно. Минимальное значение периода зависит от установленных размеров зоны считывания. При установке минимального периода темп передачи данных камерой максимален и может достигать 20 Мбайт/с. Однако, ввиду высокой скорости съёмки, данный режим не может использоваться постоянно, но он идеально подходит для автоматических систем распознавания поверхности, когда робот сканирует пространство вокруг себя с заданной частотой.
2. При использовании режима «внешнего запуска» режим экспозиции происходит по положительному фронту внешнего сигнала, приходящего на

вывод разьема внешней синхронизации. Длительность экспозиции определяется программно установленным значением.

3. Работа в режиме программного запуска с фиксированной заданной экспозицией.

Таблица 1

Производительность обработки данных для различных вариантов

Операции	Вариант 1	Вариант 2	Вариант 3
Преобразование данных, с	26	25	48
Перемещение преобразованных данных на внешний жесткий диск, с	6	8	31
Тематическая обработка данных, с	17	16	16

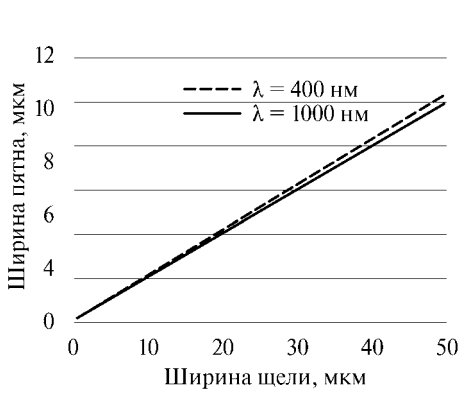


Рис. 11. Зависимость ширины размытия пятна d_n от ширины щели при увеличении фокусного расстояния коллиматорного объектива в 10 раз

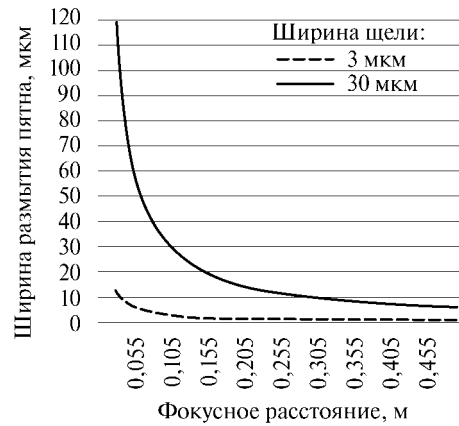


Рис. 12. Зависимость ширины размытия пятна d_n от фокусного расстояния коллиматорного объектива при двух значениях ширины щели

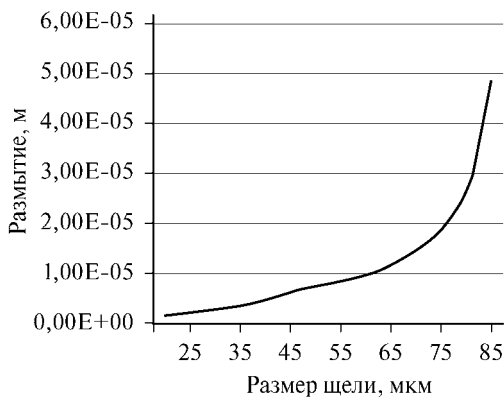


Рис. 13. Зависимость отклонения монохроматических лучей от ширины оптической щели

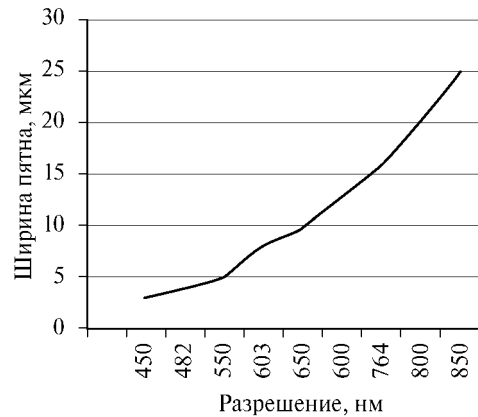


Рис. 14. Спектральное разрешение прибора при размере щели, равном 40 мкм

Одной из причин непараллельности лучей, падающих на призму, является неточность установки щели в фокусе коллиматорного объектива. При хроматической аберрации объектива точная установка щели невозможна, так как его фокусное расстояние зависит от длины волны. Кроме того, оптическая щель имеет конечный размер, и лучи света проходят сквозь призму под различными углами. Для наклонных лучей угол преломления призмы оказывается больше, и они преломляются сильнее. Вследствие этого спектральные линии смещаются в сторону длинных волн. С использованием описанных выше моделей оптического тракта гиперспектрометра была найдена требуемая зависимость размытия спектрального пятна от ширины оптической щели (рис. 11–13, см. с. 226).

Как видно из графика (рис. 14), размер щели 40 мкм приводит к погрешности менее одного пиксела фотоматрицы (размер пиксела составляет 6,45 мкм), что позволяет получить гиперспектральные данные с минимальными погрешностями и при этом обеспечить высокое светопропускание оптического тракта.

ЭКСПЕРИМЕНТЫ

Спектральный анализ основан на сравнении спектральных гистограмм, полученных в ходе съёмки, с гистограммами уже известных материалов. Необходимым инструментом анализа являются *спектральные библиотеки* — базы данных, содержащие информацию о зависимости коэффициента отражения материалов от длины волны. Использование существующих библиотек весьма затруднительно ввиду наличия в них эталонных сигнатур, полученных в неизвестных лабораторных условиях. Для работы с этими библиотеками необходимо провести процедуры калибровки используемого прибора и согласования его параметров с параметрами библиотечного гиперспектрометра. Эта процедура не всегда возможна. Особенным препятствием является то обстоятельство, что для многих материалов даются коэффициенты отражения для длин волн свыше 2,5 мкм, в то время как в описываемом макете был использован ограниченный диапазон 0,4...1 мкм. Всё это привело к необходимости создания собственной библиотеки эталонных сигнатур.

На первом этапе была проведена калибровка прибора по длине волны и получена оценка его разрешения по спектру (рис. 15). Измерение спектрального

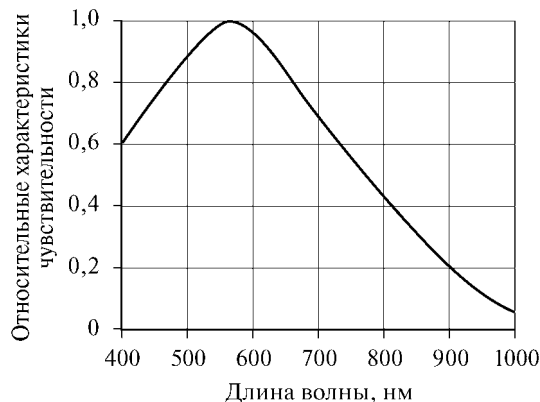


Рис. 15. Спектральная характеристика фотоматрицы

разрешения проводилось с помощью монохроматора МДР-6 для диапазона длин волн от 400 до 850 мкм.

Затем для оценки возможностей использования гиперспектрального сигнала в задачах мобильной робототехники была получена и обработана экспериментальная сцена, аналогичная той, которую движущийся робот может встретить на пути своего передвижения. Фотография сцены и её эталонное представление приведены на рис. 16.

Далее путём анализа фрагментов снимков были получены эталонные сигнатуры. Выбор участков для подготовки базы эталонных сигнатур подразумевает выбор фрагментов эталонного снимка, однородных по яркости и расположению. Согласно эмпирическому правилу каждый эталонный участок должен содержать в 10...100 раз больше пикселей, чем число спектральных каналов снимка. После подготовки полученных гиперспектральных данных (в частности, коррекция дисторсионных искажений) из первого полученного гиперкуба были выделены четыре эталонные сигнатуры, соответствующие бетонному бордюру, щебню, траве зелёной и немного высохшей зелёной траве. Построение сигнатур, как это следует из табл. 2, занимает время в пределах 5 мин.

Таблица 2

Время построения спектральных сигнатур

Количество эталонов	1	2	3	4
Время предварительной обработки, с	85	85	85	85
Время тематической обработки, с	90	140	200	265

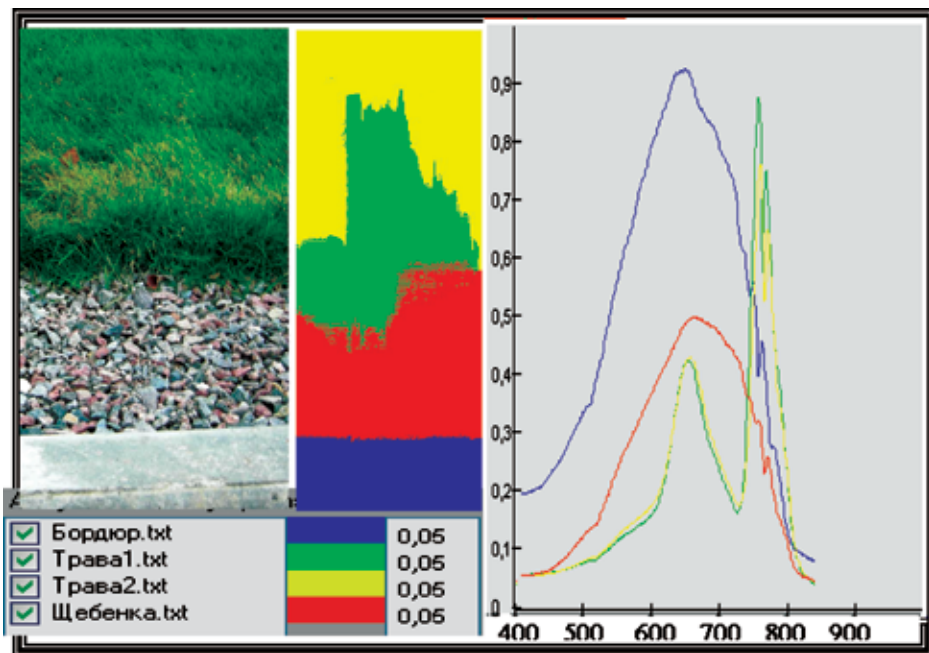


Рис. 16. Экспериментальная сцена в полевых условиях с результатами использования четырёх эталонов

Затем разработанным программным средством Analyzer [Овчинников, 2009] была проведена тематическая обработка последующих гиперспектральных данных с успешным распознаванием расположения на видеплоскости гиперкуба каждого из упомянутых элементов сцены.

Выводы

- У локальной гиперспектрологии есть безусловное будущее в качестве дистанционного средства информационного обеспечения мехатронных автоматов.
- Для внедрения локальной гиперспектрологии необходимо выполнить работы в следующих направлениях:
 - построение спектральных сигнатур материалов;
 - разработка способов комплексирования с сигналами дальномера;
 - создание средств интерактивной спектрологии.

Литература

- Носков В. П., Рубцов И. В. (2005) Опыт решения задачи автономного управления движением мобильных роботов // Мехатроника, автоматизация, управление. 2005. № 12. С. 21–24.
- Овчинников А. М., Ролдугин Д. С., Овчинников М. Ю. (2009) Аппаратно-программный комплекс для обработки спектральной информации // Датчики и системы. 2009. № 6. С. 41–46.

HYPERSPECTRAL VISION SYSTEM FOR MOBILE ROBOTS

A. M. Ovchinnikov, A. K. Platonov

Keldysh Institute of Applied Mathematics Russian Academy of Sciences (KIAM RAS)

A mock-up of the hyperspectrometer is described. The mock-up is constructed for the researches of the characteristic's peculiarities of the hyperspectral signals, that are registered from the closely set objects. These researches are necessary for the development of the sensors for the mobile robots or recognizing mechatronic automates. Hyperspectrometer's model, that is used for the recognizing of the closely set objects, is under consideration. The model allows reduce distortion's influence in the optical channel of the hyperspectrometer. These factors emphasize differences between the «local» hyperspectrometer's information and hyperspectral information, that is received during the remote sensing of the Earth surface. The paper described the results of the researches of the accessible characteristics of the «local» hyperspectrometer and the results of experiments with the closely set objects. Current work was supported by RFFI № 08-01-00908.

Keywords: spectroscopy, hyperspectr, hyperspectrometer, spectral signature, spectral base, spectral etalons, mechatronic, mobile robots.

ОПЫТ СОЗДАНИЯ СИСТЕМЫ ЛОКАЛЬНОЙ ВИДЕОНАВИГАЦИИ ДЛЯ ПОДВОДНЫХ АППАРАТОВ

М. Ю. Артюхов, А. Н. Кропотов, А. А. Макашов, Е. И. Сахарова

*Научно-исследовательский институт специального машиностроения
(НИИСМ) Московского государственного технического университета (МГТУ)
им. Н. Э. Баумана, 105005, Москва, Госпитальный пер., 10*

В статье рассмотрены вопросы реализации системы локальной видеонавигации и динамического позиционирования, описаны алгоритмы её построения с использованием стереопары. Приводится алгоритм выделения на изображении кольцевого маркера, определения его метрических координат. Рассматривается возможность динамического позиционирования аппарата над объектом работ и автоматизированная посадка подводного аппарата по этим показаниям.

Ключевые слова: локальная навигация, обработка видеоизображения, распознавание образов, системы технического зрения, динамическое позиционирование, подводный аппарат, склейка кадров, перспективное преобразование.

Расширение и усложнение подводно-технических работ у дна и/или у объекта работ сделало актуальной задачу повышения их точностных характеристик. Одним из способов решения этой задачи является разработка высокоточных бездрейфовых, малогабаритных и относительно дешевых систем определения линейных, угловых и скоростных параметров перемещения обитаемых и необитаемых подводных рабочих аппаратов.

В статье рассмотрены принципы реализации такой системы в рамках модернизации обитаемого спасательного глубоководного аппарата (СГА).

Был разработан опытно-поставочный образец системы обработки телевизионной информации (СОТИ), при проектировании которой поставлены и решены следующие задачи:

- реализация локальной навигации подводного аппарата в придонном режиме плавания;
- реализация режима динамического позиционирования подводного аппарата возле объекта работ;
- организация сплошной цифровой видеосклейки;
- распознавание комингс-площадки подводной лодки;
- определение координат комингс-площадки подводной лодки относительно СГА.

Основным условием применения СОТИ является наличие устойчивого телевизионного контакта видеодатчика СГА с донной поверхностью или объектом работ.

Артюхов Михаил Юрьевич — инженер, e-mail: michael.artuerhof@gmail.com.

Кропотов Александр Николаевич — заведующий лабораторией, кандидат технических наук, доцент, e-mail: tkropotova@rambler.ru.

Макашов Алексей Анатольевич — инженер, e-mail: amakashov@gmail.com.

Сахарова Екатерина Ивановна — инженер, e-mail: kunapucc@gmail.com.



Рис. 1.1. Общий вид забортной части СОТИ на СГА АС-28

1. УСТРОЙСТВО И РАБОТА СОТИ

Система представляет собой программно-аппаратный комплекс цифровой обработки изображений бортовых телекамер подводного аппарата.

Конструктивно СОТИ состоит из забортной (рис. 1.1) и бортовой частей, связанных друг с другом информационными связями.

Забортная часть включает в себя два одинаковых видеодатчика, внешний вид которых приведен на рис. 1.2.

Видеодатчик состоит из двух телебоксов, четырёх светодиодных светильников и коллектора — коммутатора электрических линий. Он обеспечивает получение изображений донной поверхности для бортовой части изделия.

На СГА АС-28 используются два одинаковых видеодатчика, расположенных, как показано на рис. 1.1. Кормовой видеодатчик предназначен для решения задач измерения линейных скоростей и линейных перемещений подводного аппарата, носовой видеодатчик — для решения задач распознавания и определения координат комингс-площадки ПЛ (подводной лодки) и видеонаблюдения.

Бортовая часть СОТИ состоит из блока компьютерной обработки, который представляет собой четырёхпроцессорную IBM-PC-подобную систему для обработки видеосигналов забортных видеобоксов видеодатчиков с целью определения параметров движения СГА и распознавания и определения координат комингс-площадки.

Использование четырёх IBM-PC-подобных процессорных плат типа Advantech PCM-3380, связанных между собой по интерфейсу Ethernet, обусловлено требованием



Рис. 1.2. Внешний вид видеодатчика

одновременного решения всех вышеперечисленных задач обработки и распознавания видеоизображений.

2. ВЫДЕЛЕНИЕ ХАРАКТЕРНЫХ ЭЛЕМЕНТОВ И ОПРЕДЕЛЕНИЕ ИХ СМЕЩЕНИЙ

Формально задачу определения параметров движения объектов по видеоизображениям можно разделить на две подзадачи:

- 1) нахождение соответствия между характерными элементами (объектами) пары последовательно предъявленных изображений и определение смещений этих характерных элементов на межкадровом интервале;
- 2) восстановление параметров движения на основании информации о смещениях всех характерных элементов (объектов).

Для анализа каждое изображение разбивается на малые элементарные области, обладающие достаточным контрастом, без анализа их принадлежности конкретным объектам. Часть выделенных областей будет по-прежнему принадлежать теням и другой помеховой информации, однако за счёт большого количества таких элементарных объектов информация от ложных элементов может быть отброшена. Кадровый массив размерностью $M \times N$ пикселей представляется в виде прямоугольной матрицы ячеек размерностью $(M/size) \times (N/size)$, где $size$ — размер ячейки в пикселях. Каждая ячейка определяет, таким образом, квадратную область размером $(size) \times (size)$ пикселей и представляет собой предложенный элементарный участок изображения.

Смещения объектов определяются по двум кадрам, один из которых является эталонным. Таким образом, задача может быть сформулирована следующим образом — определить смещение текущего изображения относительно эталона. Для определения экранных смещений изображения традиционно применяются методы, основанные на анализе функции взаимной корреляции исходного (эталона) и сдвинутого изображений. Координата главного максимума корреляционной функции определяет смещение изображения. Однако непосредственное вычисление функции взаимной корреляции является весьма емкой вычислительной задачей. Поэтому на практике часто используют приближения корреляционной функции, требующие меньшего числа вычислений, — функционал. Аналогичная ситуация имеет место и с самим алгоритмом поиска максимума корреляционной функции. Вместо расчёта корреляционной функции для всего множества интересных точек с последующим поиском среди них максимума часто используют итерационные алгоритмы с минимальным числом вычислений функционала схожести.

2.1. Определение смещений кадров

В качестве функционала схожести используется функционал вида:

$$F(a, b) = \sum_{x-n}^{x+n} \sum_{y-m}^{y+m} f(x, y, a, b), \quad (2.1)$$

$$f(x, y, a, b) = \left| \text{Video}(x + a, y + b) - \text{Sample}(x, y) \right|, \quad (2.2)$$

где $\text{Video}(x, y)$ — яркость изображения в точке (x, y) ; $\text{Sample}(x, y)$ — яркость эталонного изображения в точке (x, y) ; $m, n = size/2$ — половина размера ячейки по соответствующей координате, пиксел (рис. 2.1, см. с. 233).

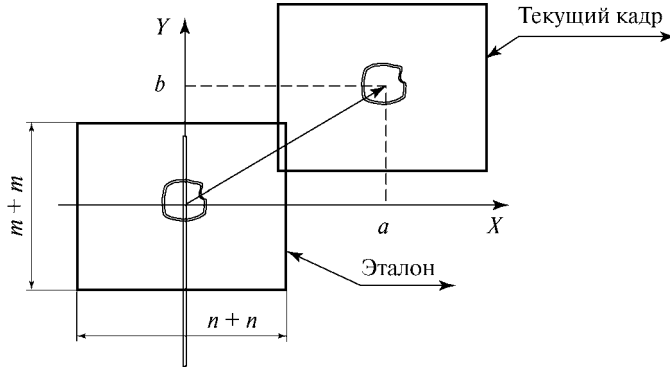


Рис. 2.1. Смещение элементарной ячейки

Данный функционал предложен как мера расстояния между эталоном и текущим изображением. Очевидно, что в случае отсутствия шумов и геометрических искажений функционал (2.1) будет равен нулю. Однако из-за наличия в реальном изображении различных шумов и искажений $F(a, b) = c > 0$ в точке минимума.

Для упрощения исследований представим тестовый кадр в одномерном виде:

$$Sample(x, y) = \begin{cases} c_1 > 0, & x < x_1, \\ c_2 > 0 \neq c_1, & x \geq x_1, \end{cases} \quad (2.3)$$

где c_1, c_2 — константы, задающие значение яркости в точке изображения.

Функция (2.3) описывает вертикальный переход в точке x_1 . Рассуждения, относящиеся к одномерному кадру (2.3), можно, не изменяя общности, перенести на двумерный случай. Функционал (2.1) в случае наличия перемещения имеет вид конуса.

На практике эталонное изображение искажено шумом той же амплитуды, что и текущее, поэтому в процессе моделирования на эталонное и текущее изображения накладывался аддитивный гауссовский шум различной амплитуды. Выбор гауссовского шума обусловлен известным положением о приближении шума с произвольным распределением к гауссовскому при прохождении им динамических звеньев. На рис. 2.2 и 2.3 (см. с. 234) представлена зависимость вида функционала схожести и взаимной корреляционной функции соответственно от уровня шумов.

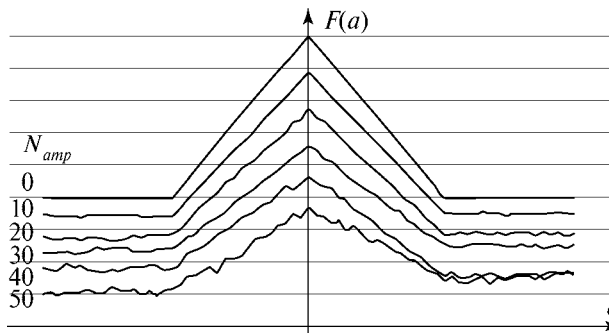


Рис. 2.2. Вид функционала при различных амплитудах шумов N_{amp}

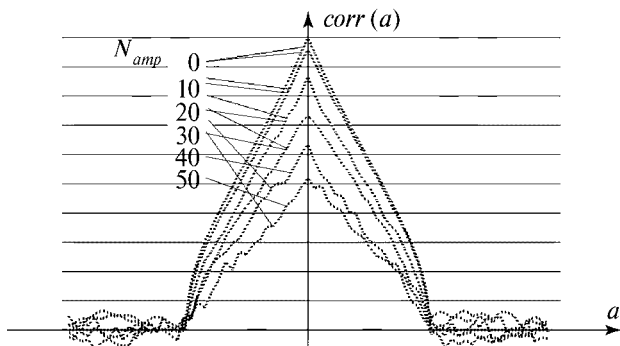


Рис. 2.3. Вид корреляционной функции при различных амплитудах шумов N_{amp}

Верхняя сплошная и прерывистая линии соответствуют функционалу схожести и взаимной корреляционной функции в отсутствие шума. Графики обеих функций монотонно смещаются вниз с ростом амплитуды шума (линии проведены для амплитуды N_{amp} шума 0, 10, 20, 30, 40 и 50 единиц при амплитуде полезного сигнала 100 единиц) от межкадрового смещения a . Самая нижняя линия соответствует максимальному шуму.

Из всех приведенных графиков видно, что форма функционала схожести вблизи экстремума может достаточно хорошо аппроксимироваться конусом, даже в случае нецелого x_1 и в условиях шумов. На основании этого полезного свойства функционала схожести удалось синтезировать простой итерационный алгоритм поиска координаты его экстремума.

В основе алгоритма лежит предположение о допустимости аппроксимации функционала схожести вблизи главного максимума эллиптическим (в двухмерном случае) конусом. Задача алгоритма — найти параметры аппроксимирующего конуса (координаты его вершины) по измерениям функционала схожести. Здесь будет рассмотрен одномерный случай, к которому приводят сигналы вида (2.3). Как видно на рисунках, аппроксимирующий конус должен менять свой наклон и ординату вершины в зависимости от амплитуды полезного сигнала и текущей шумовой обстановки. В одномерном случае конус с вершиной в т. x_0 задается следующей системой уравнений:

$$Z(x) = \begin{cases} k(x - x_0) + Z_0, & x < x_0, \\ -k(x - x_0) + Z_0, & x \geq x_0, \end{cases} \quad (2.4)$$

где Z_0 — ордината вершины; k — тангенс угла наклона образующей. В общем случае эта система уравнений имеет три неизвестных — Z_0, k, x_0 . Для нахождения параметров аппроксимирующего конуса необходимо измерение функционала схожести по крайней мере в трёх точках. Другой особенностью системы (2.4) является изменение вида функции $Z(x)$ в т. x_0 , что также затрудняет решение. Исходя из особенностей системы (2.4) был предложен следующий итерационный алгоритм:

1) определяется точка начального приближения координаты вершины конуса $x(i) = x_{00}$, $i = 1$ — номер итерации;

2) вычисляется функционал схожести — $Z(x)$ в точках $x(i)$, $x(i) + rd$, $x(i) - rd$, где положительное целое число rd — параметр алгоритма, зависящий от области определения ($m; n$) функционала схожести;

3) вычисляется вспомогательная величина — $SL(i) = [Z(x(i) + rd) - Z(x(i) - rd)]$;
 4) по знаку $SL(i)$ определяется направление на вершину по отношению к $x(i)$ и определяется наклон образующей конуса —

$$Den(i) = \begin{cases} Z(x(i) - rd) - Z(x(i)), & SL(i) > 0, \\ Z(x(i) + rd) - Z(x(i)), & SL(i) \leq 0; \end{cases} \quad (2.5)$$

5) корректируется координата вершины —

$$dx(i) = \frac{SL(i) rd}{2 Den(i)}, \quad (2.6)$$

модуль которой подвергается ограничению $|dx(i)| = \min(|dx(i)|, rd)$;

6) вычисляется новое приближение координаты вершины конуса — $x(i + 1) = x(i) + dx(i)$;

7) проверяется условие нахождения крайних точек на разных склонах конуса (условие завершения итераций) — $SL(i) < Den(i)$;

8) если условие п. 7 не выполняется, то осуществляется переход к п. 2;

9) конец итераций — $x(i + 1) \cong x_0$;

Описанный алгоритм реализует характеристику $dx(i) = F(x(i))$, линейную на участке $(x_0 - rd; x_0 + rd)$.

Аналитически характеристика $dx(i) = F(x(i))$ запишется следующим образом:

$$dx(i) = \begin{cases} rd, & x(i) < (x_0 - rd), \\ (x_0 - x(i)), & (x_0 - rd) < x(i) < (x_0 + rd), \\ -rd, & x(i) > (x_0 + rd). \end{cases} \quad (2.7)$$

В окрестности нуля смещение может быть также записано в аналитическом виде как

$$\begin{aligned} dx &= \frac{rd (Z(x(i) + rd) - Z(x(i) - rd))}{2(\max(Z(x(i) - rd), Z(x(i) + rd) - Z(x(i))))} = \\ &= \frac{rd (F(rd) - F(-rd))}{2(\max(F(rd), F(-rd)) - F(0))}. \end{aligned} \quad (2.8)$$

С учётом $\max(p, q) = 0,5(p + q + |p - q|)$, выражение (2.8) принимает вид

$$dx = \frac{rd (F(rd) - F(-rd))}{F(rd) + F(-rd) + |F(rd) - F(-rd)| - 2F(0)}. \quad (2.9)$$

Из (2.7) становится ясен практический смысл параметра rd — он задает область линейности алгоритма. При попадании x_0 в интервал $(-rd; rd)$ алгоритму потребуется только одна итерация для нахождения вершины конуса. С этой точки зрения rd должен быть как можно большим. С другой стороны, функционал схожести (2.9) вычисляется на конечной площади размером $2n \times 2m = csize \times csize$, где $csizе$ — размер области осреднения изображения при $m = n$, и при выборе большого rd есть вероятность попасть на участок функционала схожести, не являющийся конусом даже при незначительной ошибке в выборе начального приближения. На практике размер ячейки разбиения изображения $csizе$ выгодно выбирать

в пределах примерно от 4 до 64 пикселей. При этом функционал схожести может аппроксимироваться конусом на интервале $csize/2$ от ± 2 до ± 32 пикселей. Необходимо также потребовать, чтобы на склоне конуса могли разместиться все три точки, в которых вычисляется функционал схожести, в противном случае часть точек может не попасть на интервал $(-csize/2; csize/2)$ даже при малых отклонениях начального приближения, что в условиях шумов (см. рис. 2.2) не позволит реализовать горизонтальные участки характеристики $dx(i) = F(x(i))$ и приведёт к уменьшению скорости сходимости вплоть до потери устойчивости алгоритмом. Отсюда получается верхняя граница rd (или нижняя граница $csize$):

$$2rd \leq \frac{csize}{2}. \quad (2.10)$$

Из (2.10) вытекает минимальный размер ячейки $csize = 4$ при $rd = 1$.

На данном этапе работы алгоритма мы определили смещение кадров относительно друг друга, следующим шагом является восстановление на основании этих данных линейных и угловых смещений аппарата.

2.2. Восстановление параметров движения на основании смещений кадров

Для определения отстояния СГА от дна был реализован метод параллакса, использующий изображения, получаемые с двух видеокамер видеодатчика (образующих стереопару), а также предусмотрена возможность комплексирования с внешним эхолотом подводного аппарата. В основе алгоритма определения дальности лежит предположение, что плоское дно перпендикулярно оптической оси видеокамеры видеодатчика.

В идеальном случае «перемещение» на канале дальности есть расстояние между видеокамерами видеодатчика, т. е. база — величина, заранее известная. В нашем же случае имеет место неодновременность съёмки изображений с левой и правой камер, вследствие чего необходимо учитывать информацию о времени съёмки и движении телекамер.

Для учёта движения видеокамер используются два режима — «неподвижный» и «подвижный». Первый случай соответствует рассогласованию в пределах погрешности (точное значение рассогласования определяется экспериментально) — тогда базовый кадр остаётся неизменным, что позволяет избежать статической ошибки ухода. В неподвижном режиме время съёмки кадров на канале дальности можно не учитывать. В динамическом режиме движение камеры рассматривается как движение твёрдого тела по четырём координатам: Δx , Δy , ΔZ и θ , где $\theta = \Delta\varphi$ — поворот камеры за кадр (вычисляется относительно центра кадра). Тогда для каждой контрольной точки можно записать выражение:

$$\begin{cases} x'_i = \left(\Delta x + x_i \cos\theta + y_i \sin\theta \right) \frac{Z + \Delta Z}{Z}, \\ y'_i = \left(\Delta y + x_i \sin\theta - y_i \cos\theta \right) \frac{Z + \Delta Z}{Z}, \end{cases} \quad (2.11)$$

где Δx , Δy — смещение центра кадра по осям Ox , Oy в системе координат текущего кадра; x_i , y_i — координаты контрольной точки на предыдущем кадре; $\frac{Z + \Delta Z}{Z}$ — изменение масштабного коэффициента, связанное с изменением глубины;

x'_i, y'_i — измененные координаты характеристической точки (высчитываются посредством функционала (2.1)). Таким образом, получаем переопределённую систему уравнений с четырьмя неизвестными и числом уравнений, равным числу взятых на сопровождение ячеек с допустимым уровнем шума. Решая её методом наименьших квадратов, находим величину перемещений камеры.

Для нахождения Z в (2.11) необходимо два изображения: одно — с опорного кадра, другое — со стереоканала. Камеры находятся параллельно друг другу, на расстоянии B , тогда:

$$\begin{cases} x_i^1 + \Delta x \eta = \frac{\lambda X_i^1}{Z_i}, \\ x_i^2 = \frac{\lambda X_i^2}{Z_i}, \\ X_i^2 = X_i^1 + B, \\ y_i^1 = y_i^2. \end{cases} \Rightarrow Z_i = \frac{\lambda B}{x_i^2 - x_i^1 - \Delta x \eta} \Rightarrow Z = \frac{1}{N} \sum_{i=1}^N Z_i, \quad (2.12)$$

Точки в выражении (2.12) берутся вдоль эпиполярной линии; x_i^2 — координата точки на стереоканале; x_i^1 — координата той же точки на опорной камере; чтобы учесть факт неодновременности съёмки изображений на стерео- и опорном канале, необходимо ввести поправку $\eta = \frac{t_{\text{стерео}} - t_{\text{опорн}}}{t_{\text{текущ}} - t_{\text{опорн}}}$; $t_{\text{стерео}}, t_{\text{опорн}}, t_{\text{текущ}}$ — моменты времени, в которые были получены кадр со стереоканала, опорный и текущий кадры соответственно; Δx — перемещение на опорном кадре вдоль оси Ox ; усреднение Z осуществляется по всем N «хорошим» точкам эпиполярной линии.

Из (2.11) получаем $\Delta X = \frac{Z \Delta x}{\lambda}$, $\Delta Y = \frac{Z \Delta y}{\lambda}$, $\Delta \phi$, из (2.12) — последний параметр — отстояние.

Таким образом, можно определить параметры движения СГА в предположении о его движении над плоским дном.

Вышеописанный разработанный алгоритм локальной навигации позволяет определять три линейные координаты СГА и курсовой угол аппарата, две линейные скорости и отстояние от дна в режиме реального времени. Особенностью данного типа навигационных алгоритмов является отсутствие интегрирования по времени, что позволило реализовать на их основе высокоточную систему локальной видеонавигации и динамического позиционирования СГА. Разработанный алгоритм навигации-позиционирования работает в реальном времени с частотой не менее 7 Гц.

3. СИСТЕМА ВИДЕОСКЛЕЙКИ

Описанная выше методика может также применяться для составления фотопланов поверхности, над которой движется подводный аппарат. В этом случае используется одна видеокамера, а начальное отстояние аппарата от дна принимается равным единице. В таком случае осуществляется счисление не в метрических, а в пиксельных координатах, т. е. становится возможным определение расположения последовательности кадров относительно первого кадра.



Рис. 3.1. Пример работы алгоритма склейки

Нетрудно заметить, что, так как в данном случае используется только одна камера, задачу видеосклейки можно решать при постобработке заранее полученных изображений. При этом проблема сходимости алгоритма при значительных смещениях rd решается использованием исходного изображения в нескольких различных масштабах (multi resolution pyramid, описанная в работах [Koenderink, 1984; Zokai, Wolberg, 2005]).

Эта методика составления фотопланов применима при обработке результатов не только подводных видеосъёмок, но и данных сплошной видеосъёмки, полученной с произвольного носителя, совершающего плоское движение над поверхностью с незначительными перепадами высот. Так, изображение, приведённое на рис. 3.1, было получено при постобработке изображения с исходным разрешением 720×576 пикселей и при-

менении линейной межпиксельной интерполяции. Исходные изображения были получены при аэрофотосъёмке с летального аппарата.

Описанная методика позволяет осуществлять построение фотоплана не только по видеоизображению, но и, например, по изображениям, получаемым с многолучевого эхолота.

Особенностью предложенной системы видеосклейки является возможность:

- применения на других мобильных носителях (летательных или наземных), оснащённых видеокамерами;
- функционирования без дополнительных датчиков;
- видеосклейки в автоматическом режиме, без участия оператора;
- просмотра видеосклейки любыми графическими пакетами ОС Windows либо штатными программными средствами системы;
- наложений отдельных кадров при построении фотоплана с высокоточной детализацией снятых объектов, а при наличии данных отстояния от дна (о высоте полёта) точного определения расстояния на плане и размеров объектов;
- восстановления параметров движения подводного аппарата (летательного аппарата, наземного робота) — три составляющие линейной скорости, три координаты линейных перемещений либо в режиме реального времени, либо в режиме постобработки.

4. ВЫДЕЛЕНИЕ ЛЮКА КОМИНГС-ПЛОЩАДКИ

Задача выделения и сопровождения люка комингс-площадки возникает при решении задачи зависания над люком и/или автоматической посадки на него. Разработанная методика основана на выделении маркировочного кольца, нанесённого

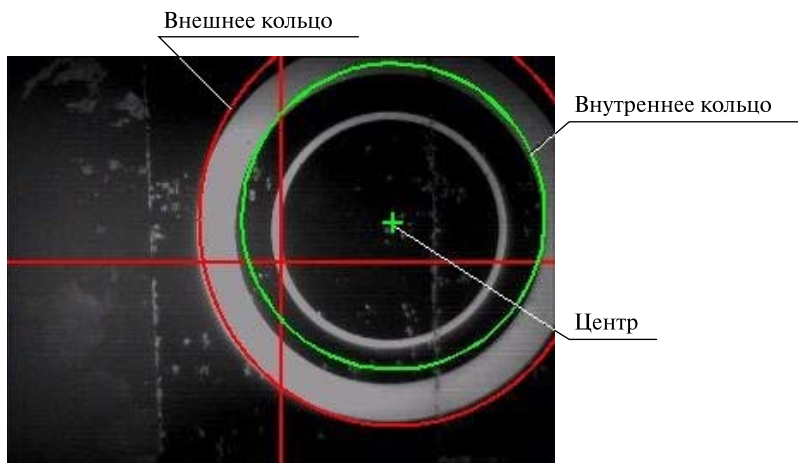


Рис. 4.1. Работа алгоритма по выделению люка

вокруг люка, и основывается на предположениях о существенном различии яркости изображения кольца и корпуса ПЛ, а также на отсутствии в кадре других кольцеобразных объектов аналогичного размера.

При обработке результатов реальных подводных съёмок возникает проблема неравномерной освещённости маркировочного кольца при подходе СГА к люку ПЛ. Для её устранения на этапе предварительной обработки из изображения вычитается фон, вычисляемый по окну постоянного размера. Затем проводится бинаризация полученного изображения с использованием метода Отсу [Otsu, 1979]. После этого осуществляется сегментация изображения методом разрастания регионов (region growing) и исключения из рассмотрения регионов, размер которых меньше порогового. Пороговая величина определяется на основании априорной информации о видимом размере кольца на максимальной рабочей дистанции. К выделенным регионам применяются морфологические операции открытия/закрытия для устранения артефактов [Яне, 2007].

Затем для каждого выделенного региона ищутся его границы с использованием оператора Робертса и определяется число связности множества граничных точек. Очевидно, что в идеальном случае изображение маркировочного кольца является двухсвязным и обе его границы аппроксимируются эллипсами.

В случае частичного перекрытия маркировочного кольца выделить такую область не удаётся. В реальных условиях такой эффект обусловлен, как правило, недостаточной освещённостью части маркировочного кольца или его частичным выходом за область видимости видеодатчика. В этом случае на выделенных границах выделяются точки перегиба, и проводится их дальнейшая сегментация. Предполагается, что граница выделенного региона состоит из элементов двух эллипсов, соединённых прямолинейными отрезками. Если выделенные эллиптические сегменты границы возможно разделить на два класса, каждый из которых описывает эллипс, то данный регион определяется как маркировочное кольцо.

На рис. 4.1 показана работа алгоритма на испытаниях системы в бассейне, используется имитатор комингс-площадки. Система отображает эллипсами различных цветов внутреннюю и внешнюю границы маркировочного кольца, его центр помечается крестом.

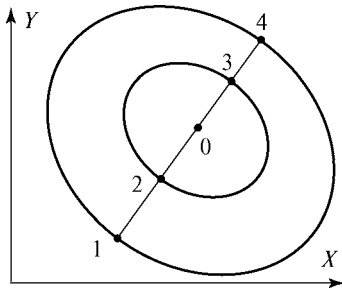


Рис. 4.2. Обработка выделенных контуров для определения углов наклона плоскости и координат центра люка

Для определения координат центра кольца комингс-площадки используются найденные внешняя и внутренняя окружности кольца. Выделенные эллипсы описываются параметрическими уравнениями. Используя матрицы перспективного проектирования [Роджерс, Адамс, 2001], можно получить уравнения для проекции окружности на изображение:

$$\begin{cases} x = \frac{f_{pix} (R \cos t \cos \psi - R \sin t \cos \psi \sin \varphi + x_0)}{f - R \sin t \sin \psi - z_0}, \\ y = \frac{f_{pix} (R \cos t \sin \varphi + R \sin t \cos \psi \cos \varphi + y_0)}{(f - R \sin t \sin \psi - z_0) \alpha}, \\ t = 0 \dots 2\pi, \end{cases} \quad (4.1)$$

где f_{pix} — фокусное расстояние камеры, отнесённое к ширине пиксела; R — радиус проецируемой окружности, м; t — полярный угол; x_0, y_0, z_0 — реальные координаты центра окружности относительно оптического центра камеры, м; f — фокусное расстояние камеры, м; ψ — угол наклона плоскости комингс-площадки; φ — угол между проекцией линии наискорейшего спуска плоскости комингс-площадки на изображение и осью ординат (рис. 4.2); α — коэффициент формы пиксела.

Можно доказать, что проекция линии наискорейшего спуска плоскости комингс-площадки проходит через точки проекционных эллипсов с координатами полярного угла $\pm 90^\circ$ или $[0^\circ, 180^\circ]$ в зависимости от выбора системы координат. Таким образом, мы выделяем четыре точки (см. рис. 4.2). Обозначив соотношение длин отрезков 12 и 34 через a , можно записать:

$$\frac{p_{1n} - p_{2n}}{p_{3n} - p_{4n}} = \frac{(f + R_1 \sin \psi - z_0)(f + R_2 \sin \psi - z_0)}{(f - R_1 \sin \psi - z_0)(f - R_2 \sin \psi - z_0)} = a,$$

где R_1 — радиус внешнего кольца, м; R_2 — радиус внутреннего кольца, м; $p_{1n}, p_{2n}, p_{3n}, p_{4n}$ — проекционные координаты точек 1, 2, 3, 4 на изображении.

Составив квадратное уравнение относительно синуса угла наклона плоскости комингс-площадки, получаем:

$$\sin \psi = s(f - z_0). \quad (4.2)$$

Подставив (4.2) в проекционные уравнения (4.1), составленные для $t = 1, 2, 3, 4$, получим выражение для z_0 :

$$z_0 = \frac{2R_1 f_{pix}}{\sqrt{\left(4f_{pix}^2 + (y_1 + y_4)^2 + (x_1 + x_4)^2\right) R_1^2 s^2 + \left(+2R_1 s(x_1^2 - x_4^2 + y_1^2 - y_4^2) + (x_1 - x_4)^2 + (y_1 - y_4)^2\right)}} + f.$$

Остальные параметры легко находятся из проекционных уравнений (4.1):

$$\sin \varphi = \frac{(z_0 - f)(x_1 R_1 s + x_4 R_1 s + x_1 - x_4)}{2 R_1 f_{pix} \sqrt{1 - s^2 (f - z_0)^2}},$$

$$x_0 = \frac{z_0 (x_4 R_1 s - x_4 - x_1 R_1 s - x_1)}{2 f_{pix}},$$

$$y_0 = \frac{z_0 (y_4 R_1 s - y_4 - y_1 R_1 s - y_1)}{2 f_{pix}}.$$

Предложенный алгоритм успешно прошел испытания в бассейне и был реализован в опытном образце. Алгоритм позволял устойчиво выделять маркировочное кольцо при его частичном перекрытии и определять три линейные координаты центра кольца относительно камеры, а также два угла наклона плоскости кольца. Разработанный алгоритм работает в реальном времени с частотой не менее 5 Гц.

Заключение

Все представленные в статье алгоритмы были реализованы в опытно-поставочном образце СОГИ и оптимизированы для работы в реальном времени. Склейка видеокадров производилась в режиме постобработки (после проведения подводных работ) на судне-носителе.

Разработанный опытно-поставочный образец системы обработки телевизионной информации успешно прошёл государственные испытания в составе спасательного глубоководного аппарата в реальных морских условиях в части реализации навигационного алгоритма и системы видеосклейки. Достигнутая погрешность измерения параметров движения спасательного глубоководного аппарата в режиме придонного плавания составила около 1,5 % от измеряемой величины. Однако достигнутая точность определения координат маркировочного кольца люка комингс-площадки не удовлетворяла точностным требованиям на режим автоматизированной посадки.

Неудача в реализации режимов распознавания и определения координат комингс-площадки системы обработки телевизионной информации, по мнению авторов, объясняется сложностью поставленной задачи в условиях изменяющихся параметров подводной среды, которая на сегодняшний момент не была решена в рамках реального промышленного образца.

Предложенные алгоритмы носят универсальный характер и могут успешно применяться на других типах обитаемых или необитаемых подводных аппаратов, а также на летательных и/или наземных мобильных аппаратах, оснащённых системами технического зрения.

Литература

- Роджерс Д., Адамс Дж. (2001) Математические основы машинной графики. М.: Мир, 2001. 604 с.
- Яне Б. (2007) Цифровая обработка изображений. М.: Техносфера, 2007. 584 с.

- Koenderink J.* (1984) The Structure of Images // *Biological Cybernetics*. 1984. N. 50. P. 363–370.
- Otsu N.* (1979) A Threshold Selection Method from Grey-Level Histograms // *IEEE Trans. Systems, Man, and Cybernetics*. 1979. V. 9. N. 1. P. 62–66.
- Zokai S., Wolberg G.* (2005) Image Registration Using Log-Polar Mappings for Recovery of Large-Scale Similarity and Projective Transformation // *IEEE Trans. on Image Processing*. Oct. 2005. V. 14. N. 10.

THE DEVELOPMENT OF LOCAL VIDEO NAVIGATION SYSTEM FOR UNDERWATER VEHICLES

M.Yu. Arturkhov, A.N. Kropotov, A.A. Makashov, E.I. Sakharova

*Moscow State Technical University named after N. E. Bauman Research Institute of
Special Machinery, Department of Underwater Systems, Moscow*

This paper describes a local video navigation systems development's problem. Stereo-pair based algorithm of navigation and dynamic positioning would be introduced. We will show the possibility to use same technique for post-processing mosaic's assembling. Also a ring-shaped marker's tracking algorithm would be described. The technique of real marker's coordinate calculation using perspective transformation would be shown.

Keywords: video navigation, underwater vehicle, image fitting, dynamic positioning, assembly of a mosaic, computer vision, perspective transformation.

Arturkhov Michael Yurievich — Engineer, e-mail: michael.artuerhof@gmail.com.

Kropotov Alexander Nikolaevich — Head of Laboratory, Candidate of Technical Science, Associate Professor, e-mail: tkropotova@rambler.ru.

Makashov Alexey Anatolevich — Engineer, e-mail: amakashov@gmail.com.

Sakharova Ekaterina Ivanovna — Engineer, e-mail: kunapucc@gmail.com.

**ПРОГРАММНО-АППАРАТНОЕ ОБЕСПЕЧЕНИЕ
СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ**

БИОЛОГИЧЕСКИ ИНСПИРИРОВАННОЕ ТЕХНИЧЕСКОЕ ЗРЕНИЕ В СИСТЕМАХ АВТОНОМНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

А. А. Жданов

*ОАО «Институт точной механики и вычислительной техники
им. С. А. Лебедева» (ИТМиВТ), 119991, Москва, Ленинский проспект, 51*

В статье представлены авторская биологически инспирированная концепция построения систем автономного адаптивного управления (ААУ), её основные положения и способы построения систем управления для прикладных объектов. Более подробно рассматриваются подсистемы «технического зрения», обеспечивающие восприятие входной информации и играющие роль самообучаемых систем распознавания образов. Особое внимание уделяется способам построения таких систем на базе специально для систем ААУ разработанных самообучаемых моделей нейронов и сетей из таких нейронов. Акцентируется внимание на том, что система восприятия входной информации у автономных роботов должна быть адекватной их системе принятия решений и совсем не обязательно должна копировать систему восприятия человека. Представлены некоторые примеры автономных роботов с адаптивными системами управления. В заключение говорится о возможных перспективах и направлениях развития данного подхода к построению систем автономного искусственного интеллекта.

Ключевые слова: автономные роботы, адаптивное управление, распознавание образов, нейроноподобные сети, метод автономного адаптивного управления.

Введение

Создавая технические системы и, в целом, техносферу как искусственную среду своего обитания, человек часто повторяет «изобретения» природы, сделанные ею в процессе формирования видов живых организмов. И это не случайно. У природы было достаточно много времени, чтобы найти и опробовать различные варианты решений, отбросить неудачные, отобрать, развить и оптимизировать эффективные решения для многочисленных управляемых систем, каковыми являются живые организмы и их популяции. В последние десятилетия, когда мощность интеллектуальных усилий в научно-технической области чрезвычайно высока, можно часто наблюдать, как специалисты в той или иной области технологий находят эффективные решения, аналогичные тем, которые давно найдены природой. Это закономерное явление, оно обусловлено наличием часто единственного оптимального решения в заданных условиях, поэтому и природа, и специалисты приходят к одному и тому же результату в своём поиске. Мы полагаем, что, в силу сказанного, бионический подход может играть роль путеводной нити при поиске технических решений. Он в состоянии существенно сузить перебор исследуемых вариантов и, следовательно, ускорить нахождение оптимальных решений при разработке новых технических систем.

Жданов Александр Аркадьевич — научный руководитель группы AAC-Lab, доктор физико-математических наук, профессор, главный научный сотрудник, e-mail: aazhdanov@ipmce.ru; a.zhdanov@mail.ru.

Настоящая работа не является обзорной и не содержит обзорной части, но представляет собой изложение одной, авторской, концепции, выработанной в результате логического анализа и моделирования принципов работы нервных систем, составивших теорию и метод «автономного адаптивного управления» (ААУ) [Жданов, 1999, 2009]. В статье более подробно рассматриваются основные аспекты, связанные с воспринимающей, «афферентной» частью систем ААУ, наиболее близкой к проблеме технического зрения. Затронуты вопросы анализа, синтеза и практических приложений метода ААУ.

Система зрения и распознавания живых организмов является важнейшей частью нервной системы, составляя, вместе с системой принятия решений, биологическую систему управления. В работе показано, что наиболее плодотворным для раскрытия механизмов распознавания является логический анализ нервной системы как адаптивной (самообучаемой) системы управления, вынужденной работать в определённых условиях, — автономности, недоопределённости и изменчивости свойств окружающей среды, необходимости её изучения с целью накопления знаний, пригодных для управления, а также дискретности её строения. Из анализа системы среда — объект управления — система управления следует, что автономная система управления вынужденно должна иметь две главные целевые функции: накопление знаний и выживание. Эти целевые функции, существенно отличающиеся от произвольных целевых функций в технических системах управления, вынужденно определяют состав функций и строение рассматриваемых систем управления. В состав функциональных подсистем должны входить: блок датчиков, самообучаемая система распознавания, аппарат качественных оценок (аппарат эмоций), база знаний, подсистема формирования знаний, в том числе — формального вывода новых знаний, подсистема принятия решений и некоторые другие. Каждая подсистема должна иметь также свою специальную память. Все другие системы управления являются частными случаями этой общей схемы. В статье основное внимание уделено входной части системы управления, обеспечивающей восприятие и распознавание информации.

Датчики понимаются как входной, нулевой уровень многоуровневой системы распознавания. Биологические датчики — рецепторные клетки, обладая небольшой собственной памятью, обеспечивают распознавание только 0-й и 1-й производных входной информации. Распознавание 2-й и более высоких производных, а также других закономерностей обеспечивается уже в прилегающих к рецепторам и последующих слоях нервных клеток. Самообучение системы распознавания в целом направлено на обнаружение повторяющихся пространственно-временных комбинаций входных сигналов. Нейрон — элементарная составляющая, из которой строится многоуровневая самообучаемая распознающая система, как и другие подсистемы, сам по себе является элементарной самообучаемой системой распознавания, но отнюдь не пороговым сумматором, как это принято в теории так называемых искусственных нейронных сетей (ИНС).

В настоящей работе представлена также соответствующая логически вынужденная модель самообучаемого нейрона, на базе которого можно реализовать все подсистемы системы автономного адаптивного управления. Подробнее описаны свойства распознающей сети, построенной из таких нейронов. Среди практически полезных свойств такой сети отмечены: а) способность формировать и запоминать образы неслучайных явлений, представленные повторяющимися пространственно-временными комбинациями сигналов; б) способность формировать и распознавать образы периодических, причинно-следственных и комбинаторных

явлений; в) способность распознавать образы в условиях помех и неполной информации; г) способность распознавать образы с упреждением (прогнозированием); д) эффект ассоциативности; е) языковые эффекты и некоторые другие свойства.

Рассматриваются также вопросы сочетания априорной информации и способности к самообучению. Показана необходимость многоуровневой технологии построения адаптивных систем в природе, где «правила» формирования образов для самообучаемой распознающей системы вырабатываются поисковыми механизмами более широкого уровня. В технических системах эта проблема может решаться несколькими способами, одним из которых является применение методов поисковой оптимизации, например, генетических алгоритмов. В статье описаны примеры оптимизации датчиков и нейроноподобной системы распознавания на прикладных примерах.

Практические прикладные системы управления на основе метода ААУ, включая и соответствующие системы технического зрения, представлены несколькими примерами: а) программной и физической моделями наноспутника с адаптивной системой управления угловым движением [Zhdanov et al., 2006]; б) физической и программной моделями мобильного робота с многоуровневой адаптивной системой управления [Zhdanov, Ryadovikov, 2000].

Таким образом, в статье представлены: логически вынужденное рациональное обоснование принципов построения систем зрения биологических организмов, формализация таких систем, способы их реализации, описаны их свойства, а также прикладные примеры, что в целом и составляет основные положения авторского метода «Автономного адаптивного управления». В заключение говорится о возможных перспективах и направлениях развития данного подхода к построению систем автономного искусственного интеллекта.

АКТУАЛЬНОСТЬ РАЗРАБОТКИ АДАПТИВНЫХ МАШИН

Если в самых общих чертах оценить прагматический вклад, который наука дала человечеству, то можно увидеть, что некогда наука позволила, прежде всего, увеличить количество вещей: сырья, пищи, оружия, жилищ и т. д. Затем наука дала возможность поднять качество, до сих пор нас поражают музейные образцы дорогих произведений искусства, оружия, дворцов, посуды и иных изделий. Ближе к XX в. пришло понимание необходимости оптимизации изделий. Оптимальные жилища, одежда, транспорт, пища и т. д. характеризуют достижения XX в. Более того, все высокие достижения науки и технологий — космическая и авиационная техника, транспорт, архитектура, приборостроение — демонстрируют достижения оптимизации во всех видах технологий. Что есть оптимизация? Это поиск экстремума некоторой «фитнес»-функции в многомерном признаковом пространстве параметров объекта и условий его функционирования, известных заранее. Объект оптимизируется на этапе его конструирования, затем найденные его параметры «замораживаются» в конструкторской документации, и этот объект далее многократно тиражируется и запускается в эксплуатацию. В течение своей «жизни» объект уже не изменяется, лишь старея и разрушаясь, пока не приходит в негодность. Каков же будет следующий этап, следующий вклад науки? В поисках ответа обратимся к природе. Мы увидим, что в природе все живые объекты не просто оптимально «спроектированы» изначально — но они адаптивны. Это означает, с нашей точки зрения, что все живые объекты в природе умеют в течение своей жизни

изменяться в зависимости от конкретных условий, в которые они попадают, а также при изменении этих условий. Нельзя найти организма, который не имел бы способностей к адаптации. Природа пришла к такому инженерному решению, и, по-видимому, человечеству тоже не избежать этого пути в своих технологиях.

Адаптивность (динамическая оптимизация) с неизбежностью должна стать следующим шагом современной науки и технологии. Должен появиться новый вид объектов — адаптивные машины, которые будут не однократно подвергаться оптимизации при их проектировании, но осуществлять оптимизацию постоянно в режиме реального времени в изменяющихся условиях.

Что адаптивность машин может дать людям? Сегодня мы стремимся научить машины делать хотя бы что-то из того, что умеют делать живые организмы. Но добиваемся мы этого совсем иными средствами — с помощью детерминированных систем, оптимизированных заранее. При этом каждая машина уникальна и приспособлена только под определённые условия. Шахматный компьютер не может пилотировать самолёт, а автопилот самолёта не может обучаться общению с себе-подобными. В природе же мы повсеместно видим чудеса управления, демонстрируемые животными, — это и высший пилотаж, и ходьба, и бег, и балансировка на одной ноге, и навигация, и выработка стратегий преследования, и общение в стае и т. д. и т. д., и всё это вместе часто демонстрируется одним устройством управления — мозгом какой-нибудь птицы, зверя, рыбы и т. п. Мы полагаем, что адаптивные машины, если мы научимся их делать, будут способны достигать аналогичных высоких результатов в своём управлении, а также снимут проблему дефицита внимания человека, возникшую с увеличением числа машин, требующих нашего внимания для управления ими.

Пока же в техносфере мы видим практически полное отсутствие адаптивных (самоприспосабливающихся) объектов. Наши машины рождаются сразу «взрослыми» и «умными», и в течение жизни ничему не учатся, лишь постепенно разрушаясь.

Поскольку сегодня не существует однозначной трактовки понятия «адаптивность», то необходимо договориться о терминах.

Адаптивностью будем называть способность системы управления автономного объекта добывать знания о свойствах системы «среда — объект управления — система управления», накапливать эти знания в своей памяти и использовать их для управления с целью дальнейшего пополнения и сохранения накопленных знаний путём обеспечения физического выживания объекта как носителя памяти и средств работы с нейю.

Отсутствие адаптивных машин сегодня имеет научные, технологические и психологические причины. Первые связаны с отсутствием полного понимания механизмов адаптивного управления, реализуемых в живой природе, и, соответственно, с отсутствием адекватных прагматических методов адаптивного управления. Сегодня существуют следующие основные методы адаптивного управления. Наибольшее развитие, главным образом благодаря работам Я. З. Цыпкина, получил подход, при котором закон управления представлен в аналитическом виде, а значения его коэффициентов автоматически уточняются в процессе управления. Адаптивными в определённой степени являются некоторые искусственные нейронные сети (ИНС), способные к автоматической кластеризации, но в них речь идёт в основном о распознавании образов, а не об управлении, либо об управлении, которое сводится только к проблеме распознавания. Однако в общем случае ИНС есть системы распознавания (либо аппроксимации), обучаемые с учителем,

и обучаемые априори — т. е. до того момента, когда они могут быть использованы в рабочем режиме. Как адаптивный метод часто рассматривают генетические алгоритмы (ГА). Однако ГА по определению являются методами эволюционной оптимизации, т. е. их следует использовать именно на этапе проектирования машины, а не в режиме её эксплуатации. Экспертные системы (ЭС) и системы на основе нечёткой логики (НЛ) также в своём каноническом виде требуют предварительной экспертной настройки — задания правил, определения функций принадлежности и т. д. Системы с подкрепляющим обучением (Reinforcement Learning — RL) требуют полного переобучения при изменении условий или целей управления, хотя эти системы достаточно близки к искомым. В каждом из названных методов существуют версии, в которых вводятся некоторые элементы самообучения. В целом все указанные подходы, которые сегодня рассматриваются как адаптивные методы, ещё достаточно далеки от тех свойств, которые мы видим у живых систем управления. Поэтому, хотя и существуют успехи в том, что такие методы позволяют оптимизировать системы управления или обучить их до начала применения в рабочем режиме, однако машин, демонстрирующих способность самообучаться в процессе работы, не встретишь на выставках робототехники. Самые эффективные антропоморфные роботы, ходящие, играющие в футбол и синхронно танцующие сложные танцы, демонстрируют не свои успехи в самообучении, а успехи детерминированного программирования. Глядя на хорошего танцора, мы говорим: «Какой молодец, как хорошо он научился танцевать», а глядя на танцующего робота, мы говорим: «Как хорошо программист составил программу движений робота!». Это разные вещи (рис. 1).

Технологические причины отсутствия адаптивных машин связаны с традиционной общей ориентацией технологии на одноразовую априорную оптимизацию изделий на этапе их проектирования, а также обусловлены понятной инерцией технологических традиций, тем более — принесших человечеству выдающиеся успехи. Все технологические процессы сегодня основаны на оптимизационных идеях. Проблемы адаптивности обходятся на этапе проектирования систем, которые приспособляются к усредненным условиям будущей эксплуатации. Инженеры сегодня не знают о возможности и необходимости создания адаптивных систем и машин и не умеют их делать. Но когда их внимание обращается на такую возможность, инженеры сразу начинают предлагать задачи для адаптивных машин, что говорит о «созревании» ситуации.



Рис. 1. Видел ли кто-нибудь, как роботы обучаются танцевать? Никто не видел, так как роботы не обучаются. Роботы работают по заранее отлаженным детерминированным программам

Психологические причины состоят в том, что соответствующий рынок ещё не сформирован и на адаптивные машины пока нет спроса. Никто не знает о возможности существования адаптивных машин и не спрашивает их в магазинах. Можно предположить также, что психологически будет непросто приспособиться к появлению адаптивных машин. Кроме того, человеку в некоторых отношениях приятнее приспосабливаться самому, чем видеть, как вещи приспособляются к нему. Однако появление адаптивных машин — систем автономного искусственного интеллекта — неизбежно, и специалисты уже понимают это.

Вернемся к научно-технологическим проблемам построения адаптивных машин. Прежде всего, обратим внимание на существенную разницу между современными техническими и биологическими системами управления. В технике сегодня господствуют детерминированные системы управления, основанные на априорной информации и математических моделях объектов. Типовая схема современной системы управления выглядит следующим образом (рис. 2).

В такой схеме заказчик системы управления и объект управления могут быть локализованы в разных местах. Уже отсюда следует, что задачи и целевые функции у такой системы управления могут быть любыми (например, объект управления — это ракета, которая должна долететь в нужную точку и там взорваться). Если целевая функция произвольна, то и устройство системы управления может быть каким угодно, лишь бы оно обеспечивало достижение цели. Совсем не то у биологических объектов, у которых объект управления и «заказчик»-мозг — это одно и то же, они находятся в одном теле, а потому там не может быть таких целевых функций, как «долететь и взорваться». У всех биологических объектов, по-видимому, одна целевая функция — обеспечить выживание объекта управления — тела в агрессивной, заранее малоизвестной среде, а для этого необходимо попутно решать задачу исследования свойств среды и накопления найденных знаний. При желании можно говорить, что более главная цель — накопление знаний, для сохранения которых надо обеспечить выживание тела. Так или иначе, налицо две определённые целевые функции — выживание и накопление знаний. Определённые целевые функции требуют и вполне определённого устройства системы управления, которая будет представлена ниже.

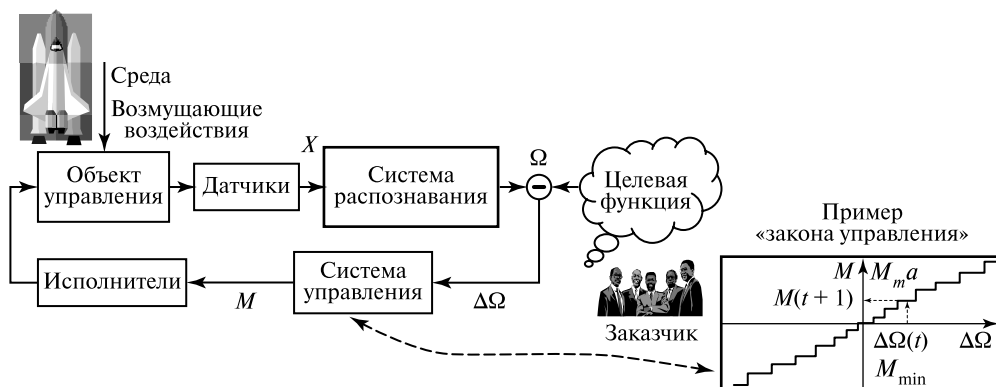


Рис. 2. Типовая схема системы управления: X — информация от датчиков; M — команда исполнительным устройствам; Ω — информация о текущем состоянии объекта управления; $\Delta\Omega$ — текущее рассогласование

Возвращаясь к техническим системам управления, обратим внимание, что все элементы системы, показанные на рис. 2, требуют использования априорных знаний и предварительной оптимизации. В результате априорных оптимизационных расчётов выбираются наиболее эффективные датчики (признаковое пространство), исполнительные устройства, система распознавания (алфавит классов, решающие правила и т. д.). Реализуемый системой управления «закон управления» во всех своих элементах опирается на априорную информацию и соответствующие расчёты. Такой подход к построению системы управления в живой природе был совершенно невозможен, так как у природы не было необходимой априорной информации.

Итак, основные отличия «технического» и «природного» подходов к построению систем управления состоят в следующем.

1. Биологический «заказчик» и объект управления (ОУ) в природе — это не разные сущности — это одно и то же. «Заказчик» сидит внутри ОУ и есть сам мозг. Отсюда следует, что важная **целевая функция мозга — это выживание** объекта управления.

2. Мозг не знает изначально и полностью оптимального закона управления, позволяющего выжить, он должен сам найти или уточнить «закон управления». Следовательно, **другая целевая функция — поиск закона управления — поиск знаний**. Отсюда следует, что мозг во многих отношениях есть поисковая система. Свойства и среды, и самого тела постоянно изменяются, поэтому поиск знаний должен происходить постоянно, т. е. мозг — система поисковой оптимизации.

3. Среда не есть только источник случайных возмущений, **среда закономерна, и мозг должен её понять, чтобы в ней выжить**.

4. Мозг и тело не могут быть неизменными. В результате накопления знаний содержимое памяти, морфология тела и морфология самого мозга **должны изменяться** в силу ограниченности памяти и поисковых алгоритмов.

Из сказанного следует вывод об основном принципе устройства мозга. Он состоит в том, что мозг, с одной стороны, должен представлять собой машину по поиску новых знаний — поисковую систему, направленную на исследование свойств малоизученной враждебной окружающей среды. В этом смысле мозг должен реализовать известный из гносеологии алгоритм познания, другого варианта у мозга нет (если только он не получает знания в готовом виде откуда-то извне). Напомним коротко этот алгоритм: а) *проанализируй имеющиеся знания и сделай обоснованный прогноз с доверительным интервалом*; б) *проведи практический эксперимент в случайном месте спрогнозированного доверительного интервала и получи новые достоверные знания*; в) *прибавь новые знания к уже имеющимся*; г) *возвращайся к шагу (а) и повторяй всё снова*. Этот алгоритм мозг и реализует постоянно, удовлетворяя своей целевой функции накопления знаний. С другой стороны, мозг, обязанный обеспечить выживание своему телу и своей «базе знаний», должен совершать такие действия, чтобы переводить тело в максимально благоприятные состояния из возможных. Для этого он должен всё время решать задачу оптимального управления, пользуясь теми знаниями, которые им накоплены к моменту принятия решения. Совместное решение обеих этих задач и составляет принцип функционирования мозга. При этом две указанные задачи противоречивы, например, поиск новых знаний требует совершения случайных действий, а принятие оптимального решения — действий максимально определённых. Тем не менее, именно компромиссное решение этих двух задач и составляет специфическое поведение живых объектов, существенно отличающееся от поведения детерминированных технических систем.

Мы полагаем, что в теории и на практике сегодня отсутствуют постановки задач построения систем управления, соответствующих биологическим нервным системам и мозгу. А именно, не ставится задача о разработке единого распознающе-управляющего комплекса, который в отсутствие полной априорной информации об окружающей среде и о самом объекте управления (которые, кроме того, могут меняться со временем) способен был бы в реальном времени управления согласованно решать следующие задачи:

- формировать и оптимизировать признаковое пространство;
- постоянно формировать образы существенных для управления явлений по комплексированным показаниям датчиков, а также строя образы над образами (понимая свои текущие состояния в терминах сформированных образов);
- исследовать закономерные переходы из известных состояний в результате совершения определённых действий в новые неслучайные результирующие состояния (искать закономерные обратные связи через исполнители, среду и датчики);
- искать новые действия;
- агрегировать и структурировать известные действия по отношению к вызываемым ими закономерным причинно-следственным связям в терминах образов состояний, сохраняя эти информационные структуры в памяти «базы знаний»;
- оценивать качественные характеристики различных состояний;
- постоянно экстраполировать «закон управления» на новые области его определения (новые образы) и новые области его значений (новые действия);
- принимать решения, оптимальные в смысле найденных качественных характеристик образов состояний, стремясь переводить объект управления в более благоприятные состояния;
- осуществлять не реактивное, но активное поведение, направленное на постоянный поиск лучших состояний и на поиск новых знаний;
- осуществлять формальный вывод новых знаний из уже имеющихся, прогнозируя их с последующей экспериментальной проверкой и обеспечивая направленный поиск и экономию ресурсов;
- автоматически развивать средства принятия решений за счёт развития средств многошагового принятия решений посредством формирования и использования языковых структур образов и действий;
- конечным проявлением свойств системы управления должна являться её способность при помещении в среду с малоизвестными и слабоформализуемыми условиями находить способ функционирования, приводящий к постепенному улучшению качества управления по заданным критериям, в том числе — при постепенном изменении свойств среды или самого объекта управления.

Все известные нам постановки задач по построению систем управления являются только частичными по отношению к данному решению.

МЕТОД «АВТОНОМНОГО АДАПТИВНОГО УПРАВЛЕНИЯ»

Попытка построения системы управления, соответствующей указанной постановке задачи, выполнена в исследовательской группе AAC-Lab и описана в ряде

публикаций, в том числе — в монографии [Жданов, 2009]. Не имея здесь возможности подробно описать систему автономного адаптивного управления (ААУ), отсылаем читателя к данной монографии или к соответствующим публикациям (например, [Жданов, 1999]). Здесь отметим только те аспекты, которые имеют отношение к теме. На рис. 3 представлена общая схема системы управления ААУ.

Принцип работы системы ААУ логически вынужденно выводится из анализа условий, в которых вынужден работать биологический мозг, эти условия состоят в следующем. Управляющая система (УС) помещена в тело объекта управления (ОУ), который погружен в среду. Априори свойства среды и ОУ известны УС мало (однако некоторая априорная информация обязательно должна быть). Входную информацию УС получает через сенсоры, её выходная реакция поступает вовне через исполнительные устройства (актуаторы). Среда закономерна, хотя и очень сложна. Неизвестные УС закономерные проявления среды воспринимаются УС как случайные. В определённом смысле задача УС — познать свойства среды, добиваясь превращения «случайных» воздействий среды в закономерные, чтобы научиться управлять ими. Структура и принцип действия УС дискретны. Множества её элементов конечны.

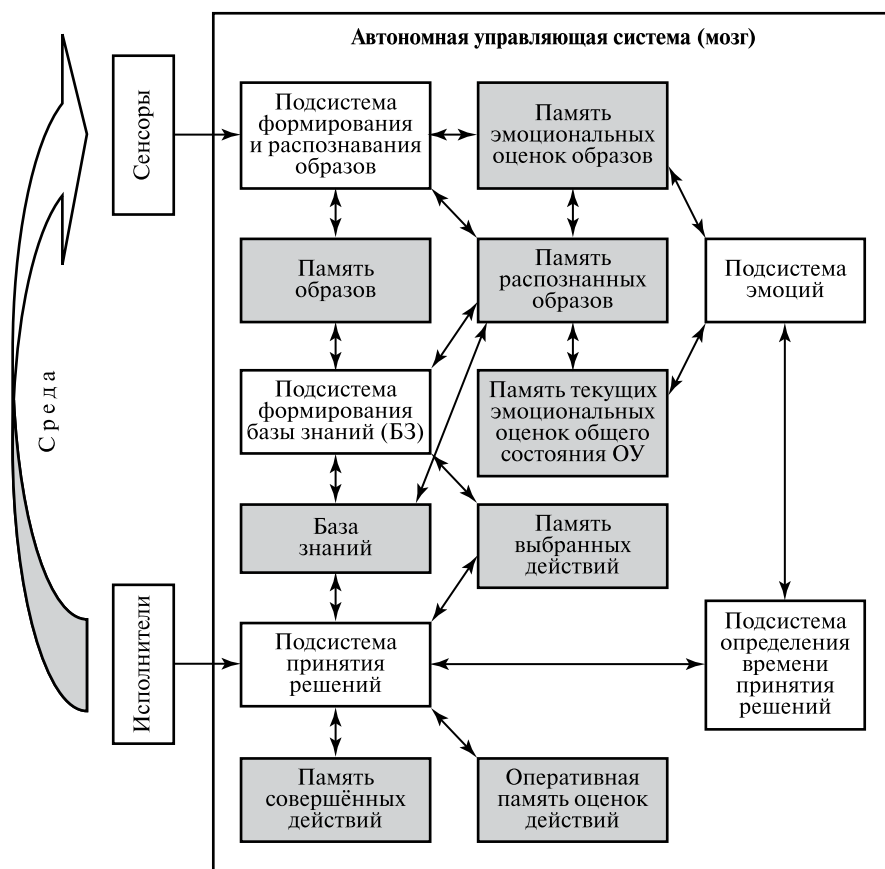


Рис. 3. Общая структура подсистем системы управления ААУ

Из этих условий вытекает следующий принцип действия УС, осуществляющей автономное адаптивное управление.

1. «Сенсоры» преобразуют некоторую часть воздействий, оказываемых средой на ОУ, в дискретные (в том числе — по времени) наборы бинарной информации (нервные импульсы).

2. «Подсистема формирования и распознавания образов» (ФРО) должна: а) научиться выделять во входной информации статистически неслучайные пространственно-временные комбинации сигналов, присваивая им идентификаторы — номера образов, и запоминая их в «Памяти образов», другими словами, это задача автоматической кластеризации входной информации в признаковом пространстве-времени, с формированием алфавита классов; б) распознавать в текущей входной информации образы, уже зафиксированные ранее в «Памяти образов», обозначая их в «Памяти распознанных образов».

3. «Подсистема формирования и распознавания образов» должна оказывать воздействие на среду, выбирая «действия» с помощью «Подсистемы принятия решений» и передавая номер выбранного действия исполнительным устройствам. «Подсистема принятия решений» использует информацию, записанную в «Базе знаний» (БЗ). Если в БЗ нет информации, адекватной текущему состоянию, то решения принимаются случайным (в определённых рамках) способом. Таким образом, решения принимаются всегда, случайно либо не случайно.

4. Анализируя «Память распознанных образов» и «Память выбранных действий» с учётом времени t , «Подсистема формирования БЗ» может отыскивать в памяти неслучайные цепочки причинно-следственных событий в виде повторяющихся последовательностей вида

$$\{O_i(t-2), D_j(t-1), O_z(t)\}, \quad (1)$$

где O_i — образы «условия»; D_j — образы действий; $O_z(t)$ — образы результатов этих действий. Такую тройку закономерных причинно-следственных событий вида (1) будем называть «элементарным знанием». Структуры из таких элементарных знаний сохраняются в памяти БЗ.

5. В каждый момент времени «Память распознанных образов» отделяет в БЗ ту её часть, которая содержит знания, адекватные текущей ситуации, — некоторую структуру из записей вида (1), где образы «условия» совпадают с тем, что содержится в текущем состоянии «Памяти распознанных образов». Отделённая часть БЗ содержит указания на то множество альтернативных действий, которые можно совершить в текущий момент.

6. Для обоснованного принятия решений — выбора действия — нужны качественные критерии, позволяющие отличать качество образов «результатов» совершения действий. Качественное оценивание образов, а тем самым и целеполагание при принятии решений осуществляет специальная «Подсистема эмоций», которая имеется у любого живого организма. «Подсистема эмоций» выполняет несколько функций: во-первых, оценивает качество («эмоциональные оценки») сформированных образов, во-вторых, даёт интегральную эмоциональную оценку текущего состояния, в-третьих, формирует универсальный «стимул» к активной деятельности УС — а именно, стремление к получению максимально больших эмоциональных оценок текущего состояния. «Подсистема эмоций» играет важнейшую, «системообразующую» роль в живых организмах, задавая им стимул к активному поведению, организуя и направляя это поведение. Если «объективная, фактологическая» информация о свойствах системы среда — ОУ накапливается в образах и

знаниях УС, то «Подсистема эмоций» определяет «субъективные» оценки и целевые функции управления, связанные с данным ОУ.

7. «Подсистема принятия решений» анализирует отобранные в БЗ возможные в текущий момент действия и выбирает то из них, которое обещает привести к распознаванию образов «результатов» с максимальной суммарной эмоциональной оценкой либо к вытеснению распознанных образов с максимально плохими эмоциональными оценками.

8. Выбор действия осуществляется с учётом интервала времени, отведенного на принятие решения, которое, в свою очередь, зависит от величины текущей общей оценки эмоционального состояния ОУ и его производной. Чем хуже текущее состояние и чем быстрее оно ухудшается, тем скорее надо принимать решение. Отведенное на принятие решения время используется для анализа только соответствующего числа наиболее важных образов результатов (с максимальными по модулю оценками). Остальные образы результатов УС учесть не успевают, тем самым (как один из механизмов) привнося случайность в процесс принятия решений, необходимую для поиска новых знаний.

9. По мере формирования всё более сложных, обобщенных и абстрактных образов УС, действуя по описанному алгоритму, начинает принимать всё более сложные и обоснованные решения. Адаптивности способствует также формирование новых действий, новых знаний, новых оценок образов.

Можно говорить, что другие способы построения систем управления являются частными случаями данной схемы, когда управление строится на использовании только некоторых из указанных на схеме (рис. 4) подсистем. Так, например, управление на основе «экспертной системы» (рис. 4а) — это заранее заполненная «База знаний» и «Подсистема принятия решений», когда все остальные подсистемы удалены и заменены «заглушками». А управление на основе ИНС возможно тогда, когда для управления достаточно только решить задачу распознавания (рис. 4б).

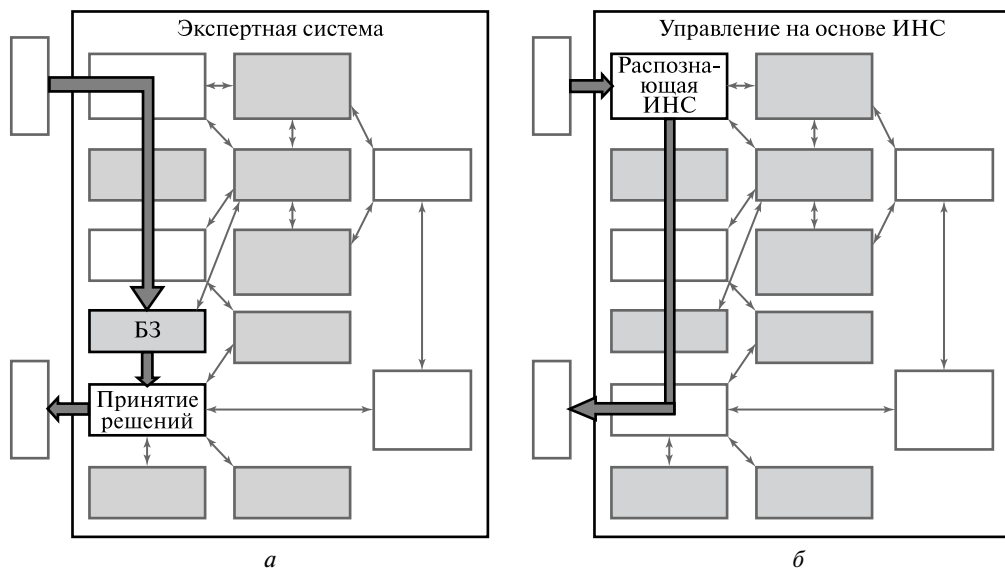


Рис. 4. Частные случаи систем управления, когда некоторые подсистемы системы ААУ не используются

ВОСПРИЯТИЕ ВХОДНОЙ ИНФОРМАЦИИ В СИСТЕМАХ ААУ

Остановимся на вопросах «зрения» системы ААУ и мозга. Из общих соображений и из анализа структуры задач, решаемых системой ААУ, можно заключить, что восприятием входной, афферентной информации занимается практически половина УС. Выскажем несколько заключений, следующих из нашего понимания рассмотренного принципа управления.

Датчики и система распознавания. Известно, что для управления динамическими объектами системе управления требуется информация о 0-й, 1-й, 2-й и других производных значений и комбинаций наблюдаемых параметров. Однако для вычисления уже 1-й и больших производных нужна память для хранения наблюдаемых значений на предыдущих тактах. В биологических сенсорах — рецепторных клетках — природа не может использовать много памяти. Поэтому природой выработаны «тонические» датчики, измеряющие 0-ю производную, практически без использования памяти, и «фазические» датчики, измеряющие 1-ю производную, с использованием небольшой памяти (условно — на один такт в прошлое). Для распознавания информации о 2-й и т. д. производных природа использует уже нейронные конструкции, которые дополняют рецепторные клетки, где нейроны располагают необходимой, хотя и небольшого объёма памятью. Поэтому, по сути, система распознавания мозга состоит из датчиков (0-й слой) и дополняющих их нескольких слоев нервных клеток, распознающих производные более высокого порядка, а также другие пространственно-временные образы, нужные для управления.

Семантика распознаваемых образов. Какие образы должна распознавать система зрения данного ОУ? Из общего понимания задачи об адаптивном управлении следует, что УС нуждается в тех образах, которые существенны для принятия решений. Это образы, на основании которых УС может принимать эффективные решения, позволяющие ОУ выживать в окружающей его среде. Подчеркнем, что физическая реальность, в которой находится данный ОУ, и система эффективных образов, распознаваемых его УС, — это «две большие разницы». Система образов, которыми оперирует данный ОУ, — это некоторое абстрактное отображение реальности, которое должно быть эффективно для управления (иначе организм бы не смог выживать в среде), но оно в общем случае может соотноситься с реальностью так же, как текст на странице книги, описывающий пейзаж, соотносится с этим пейзажем. Два разных животных видят одну и ту же сцену совершенно «разными глазами». Каждый из них видит в нем то, что важно для него, и может совсем не видеть того, что видит другое животное. Бессмысленно рассуждать о том, которое из животных видит более правильно. «Правильность» образов, действий и оценок проверяется только возможностью их использования для управления. Так, змея видит в инфракрасном диапазоне, а лягушка видит только движущиеся точки. Человек — не исключение: мы видим только то, что мы научились «видеть» и что оказалось полезным видеть именно нам, людям. Однако не следует забывать, что линейная перспектива, контуры, цвет и другие элементы — это всё не более чем «фокусы», «аппаратная функция», «образы» нашей системы восприятия, с помощью которых мы получаем информацию о взаимном расположении объектов в пространстве, их важных для нас свойствах и т. п. В реальной природе нет линейной перспективы, цвета и т. д., а есть только электромагнитные поля, ступки элементарных частиц...

Не следует забывать также о том, что мы видим только то, что мы понимаем, т. е. то, что обучена видеть наша система распознавания. Мы можем буквально не видеть находящийся перед нами объект, если мы не обучены его видеть.

Системы машинного зрения. Из сказанного следует, что для построения систем машинного зрения вовсе не обязательно копировать систему зрения человека. Мы это делаем по инерции, пытаясь перенести в машину свойства единственной хорошо знакомой нам системы восприятия — человеческой. Система машинного зрения должна снабжать УС машины образами, которые позволят ей принимать правильные решения. Однако, поскольку мы добиваемся от машины действий, «разумных» именно с нашей точки зрения, мы навязываем ей и нашу систему восприятия. Возможно, что видеоизображения, полученные видеокамерами, и их анализ не есть лучший подход к построению автономных роботов. Картинка с видеокамеры — это изображение, удобное только для человека. Боюсь, что, кроме человека, ни одно животное не воспринимает телевизионное изображение как отражение реальности и не может использовать его для своего управления. Возможно, что следует разрабатывать специальные датчики для роботов, ориентированные именно на их системы принятия решений. Элементы этого подхода заметны уже сейчас, когда в качестве датчиков начинают использоваться спектрометры, различные фильтры и другие способы выявления и представления информации, ориентированные на конкретное их использования в машине, а не копирующие человеческое восприятие.

САМООБУЧАЕМАЯ СИСТЕМА РАСПОЗНАВАНИЯ В МЕТОДЕ ААУ

Подсистема формирования и распознавания образов (ФРО) в системе ААУ может быть построена на базе любого подходящего самообучаемого метода распознавания образов, соответствующего требованиям прикладной системы ААУ. В частности, в общем случае требуется, чтобы подсистема ФРО допускала обучение непосредственно в процессе управления, в том числе дообучение и переобучение.

Однако, принимая во внимание, что метод ААУ имеет бионическое происхождение, наибольший интерес для нас представляет его нейроподобная реализация, подобная по форме и принципу действия нейронной системе мозга. Известные так называемые искусственные нейронные сети были сочтены нами принципиально не соответствующими нашему пониманию принципов работы биологических нервных систем и нейронов, поэтому были разработаны новые модели нейронов, соответствующие требованиям метода ААУ. Разработаны несколько моделей нейрона, отличающиеся уровнем сложности и функциональности, но основанные на одном принципе. Кратко опишем базовую модель нейрона ААУ.

Мы исходим из того, что отдельный биологический нейрон (рис. 5, см. с. 258) представляет собой простую самообучаемую систему распознавания образов. Нервная система и мозг представляют собой логическую машину. Поэтому в ней должен выполняться закон тождества, который гласит, что предмет суждения должен оставаться тождественным самому себе в этом суждении (*зд.* — в работе УС). Закон тождества был сформулирован Аристотелем в трактате «Метафизика» следующим образом: «...иметь не одно значение — значит не иметь ни одного значения...». Для нервной системы это означает, что выходной сигнал нейрона, если он появляется (при условии, что все сигналы одинаковы по форме), должен означать каждый раз одно и то же. Это «одно и то же» есть «одна и та же комбинация сигналов на входе нейрона». Если нейрон автономный и никаких

входных сигналов иного рода не получает, то некоторая комбинация сигналов на входе, которую он должен отличать от всех других комбинаций, может отличаться только своими пространственными и временными-частотными-вероятностными характеристиками.

Иными словами, нейрон должен сначала выявить некоторую неслучайную комбинацию, удовлетворяющую заданным пространственным и временным-частотным характеристикам, запомнить её и начать реагировать своим выходным сигналом только на появление именно этой комбинации на входе, но никакой иной. Для других нейронов в сети выходной сигнал данного нейрона будет означать всегда одно и то же — данный нейрон на предыдущем такте распознал закрепившуюся за ним всегда одну и ту же ситуацию на его входе. Если бы выходной сигнал нейрона означал каждый раз нечто иное, ни о каком управлении не могло бы быть и речи.

Работает нейрон, по-видимому, следующим образом. «Умножители» его входных сигналов — синапсы растут по мере их использования, т. е. при прохождении через них сигналов. Однако если несколько сигналов, пространственно недалеко отстоящих друг от друга (структурное правило), приходят одновременно, то соответствующие синапсы растут существенно быстрее, как бы взаимно стимулируют друг друга (посредством выделяемых ими в межклеточную среду химических медиаторов). Такой вид «пластичности» синапсов известен в биологии. Это и есть основной способ обнаружения коррелирующих событий нейроном. Входная комбинация сигналов, приходящая чаще других, приведёт к победному росту именно своих синапсов (кроме того — к угнетению других синапсов), которые однажды достигнут таких размеров, что именно эта комбинация на входе начнет вызывать срабатывание «триггерной» части нейрона и появление его выходного сигнала.

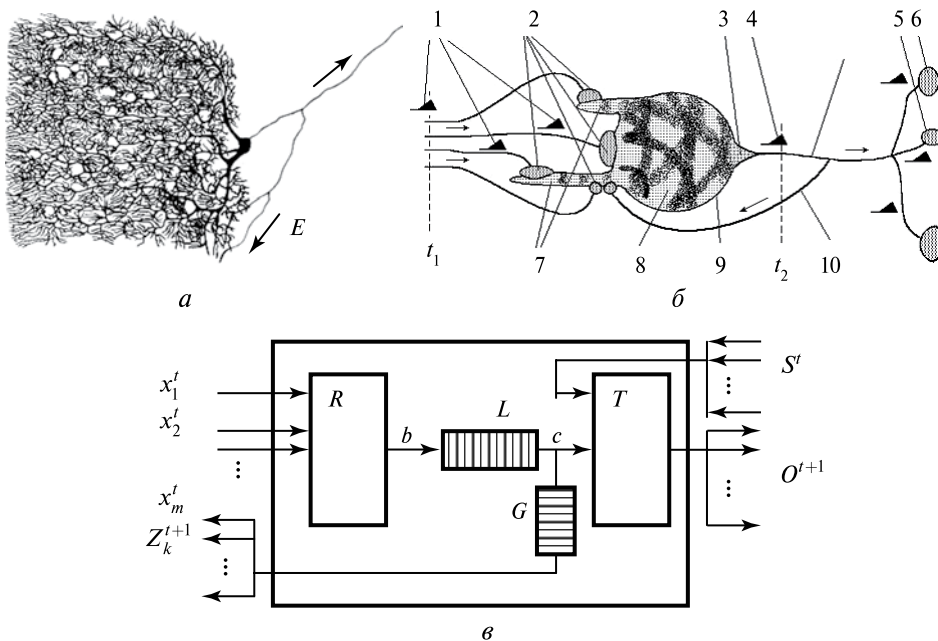


Рис. 5. Образец реального нейрона (а), его условное изображение (б) и модель (в)

Тем самым, нейрон обучается узнавать именно ту комбинацию входных сигналов, которая появлялась достаточно часто и чаще других, и начинает сигнализировать другим нейронам именно об её появлении, т. е. его выходной сигнал каждый раз будет означать одно и то же. Обучившийся нейрон будет игнорировать иные комбинации входных сигналов (некоторые нейроны могут распознавать несколько разных входных воздействий), реагируя на каждую из них определённым выходным сигналом, например, разной частотой, что тоже логично.

Простая модель нейрона, реализующая этот принцип, показана на рис. 5*в*. Здесь элемент R осуществляет проверку заданного структурного правила. В простейшем случае элемент R может быть представлен логическим элементом «И», либо пороговым «И». Проверка статистического правила осуществляется элементом L . В простейшем случае это может быть счётчик числа срабатываний элемента R , который начнет пропускать сигнал из точки b в точку c только после того, как единичный сигнал в точке b появится не менее чем M раз, где M — заданная константа. Тем самым, элемент L — это «коллективный» синапс кореллирующих входных сигналов, а M — пороговое значение, определяющее степень обученности нейрона, достаточную для вынесения суждения о неслучайном характере закономерности, выявляемой элементом R . Нейрон не начнет «узнавать» комбинацию на входе и реагировать на неё выходным сигналом, пока он не обучится — пока число N наблюдений данной комбинации не превысит заданного в элементе L порога M .

Если говорить об элементе R как о пороговом «И» (здесь речь идет уже о другом пороге ρ), то этот порог ρ может постепенно уменьшаться в зависимости от числа N таким образом, чтобы при накоплении статистики сохранять приемлемую вероятность ложной тревоги. Снижение порога ρ позволяет нейрону распознавать образ в условиях помех, а также обеспечивает несколько крайне важных свойств нейрона. К числу последних относятся: а) возможность распознавания образа с упреждением, что ведёт к возможности прогнозирования и принятия решения; б) возможность ассоциативной памяти и в) возможность возникновения языковых явлений [Zhdanov et al., 2006]. На выходе нейрона стоит триггерный элемент T , который поддерживает выходной сигнал O о распознавании образа дольше, чем длится предъявление прообраза на входе нейрона, это один из видов оперативной памяти нейрона. Отключать демонстрацию выходного сигнала можно, в простейшем случае, с помощью «тормозящего» сигнала S , поступающего на специальный вход нейрона (тормозящий синапс). Возвратную коллатераль нейрона можно организовать либо с помощью отвода выходного сигнала O , либо с помощью специального выхода Z . На срабатывание нейрона требуется один такт времени, что играет важную роль в сетевой структуре из нейронов. Формальное описание данной модели нейрона приведено, например, в работах [Жданов, 1999, 2009].

Покажем на примерах из работы [Жданов, 2009] некоторые свойства сети из описанных нейронов. Возьмём некоторый фрагмент сети из подсистемы ФРО (рис. 6, см. с. 260).

На рис. 7 (см. с. 260) покажем временную диаграмму, отражающую бинарные сигналы в этой сети. По оси ординат пронумерованы нейроны и датчики. По оси абсцисс идет дискретное время. Зачерненная клетка отражает наличие сигнала на выходе соответствующего нейрона в данный такт времени. Будем неоднократно подавать на датчики некоторую пространственно-временную комбинацию сигналов. Эта комбинация-прообраз обведена рамками в верхней части временной диаграммы.

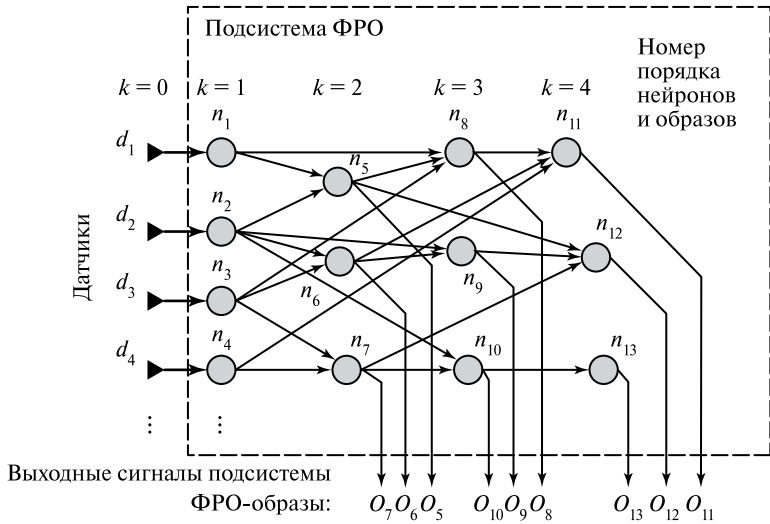


Рис. 6. Пример фрагмента сети ФРО, составленной из описанных нейронов. Датчики, вместе с прилегающей к ним сетью нейронов, составляют самообучаемую систему распознавания образов

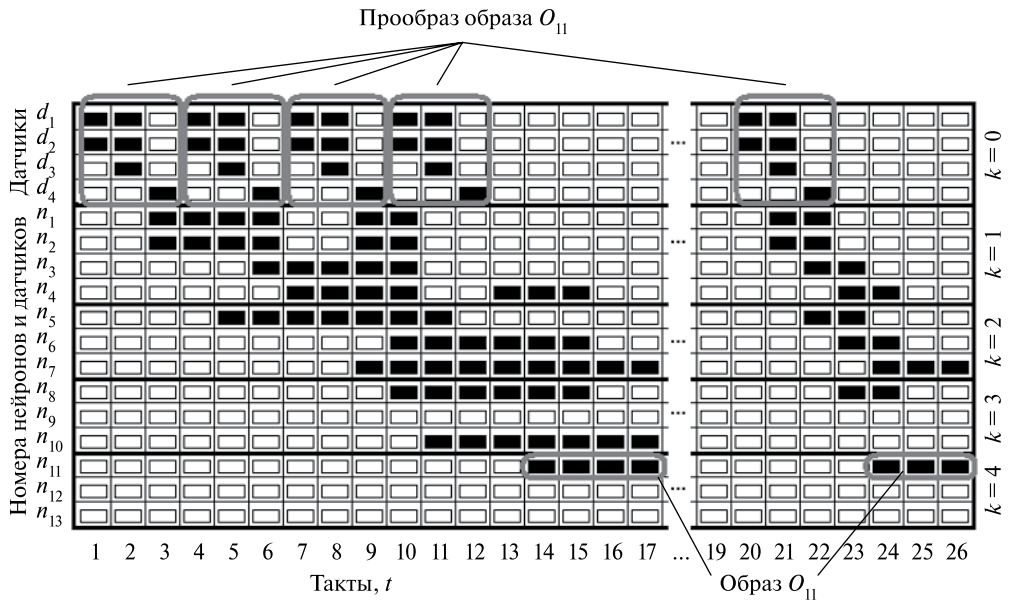


Рис. 7. Временная диаграмма, отражающая процессы в сети при неоднократной подаче некоторой комбинации сигналов

По мере обучения нейронов (здесь $M = 1$) на их выходах начинают появляться сигналы. Образ O_{11} , наиболее полно соответствующий всему данному прообразу, вырабатывается нейроном n_{11} , который обучается и приобретает способность распознавать такой образ на такте $t = 14$. Теперь этот нейрон будет в состоянии

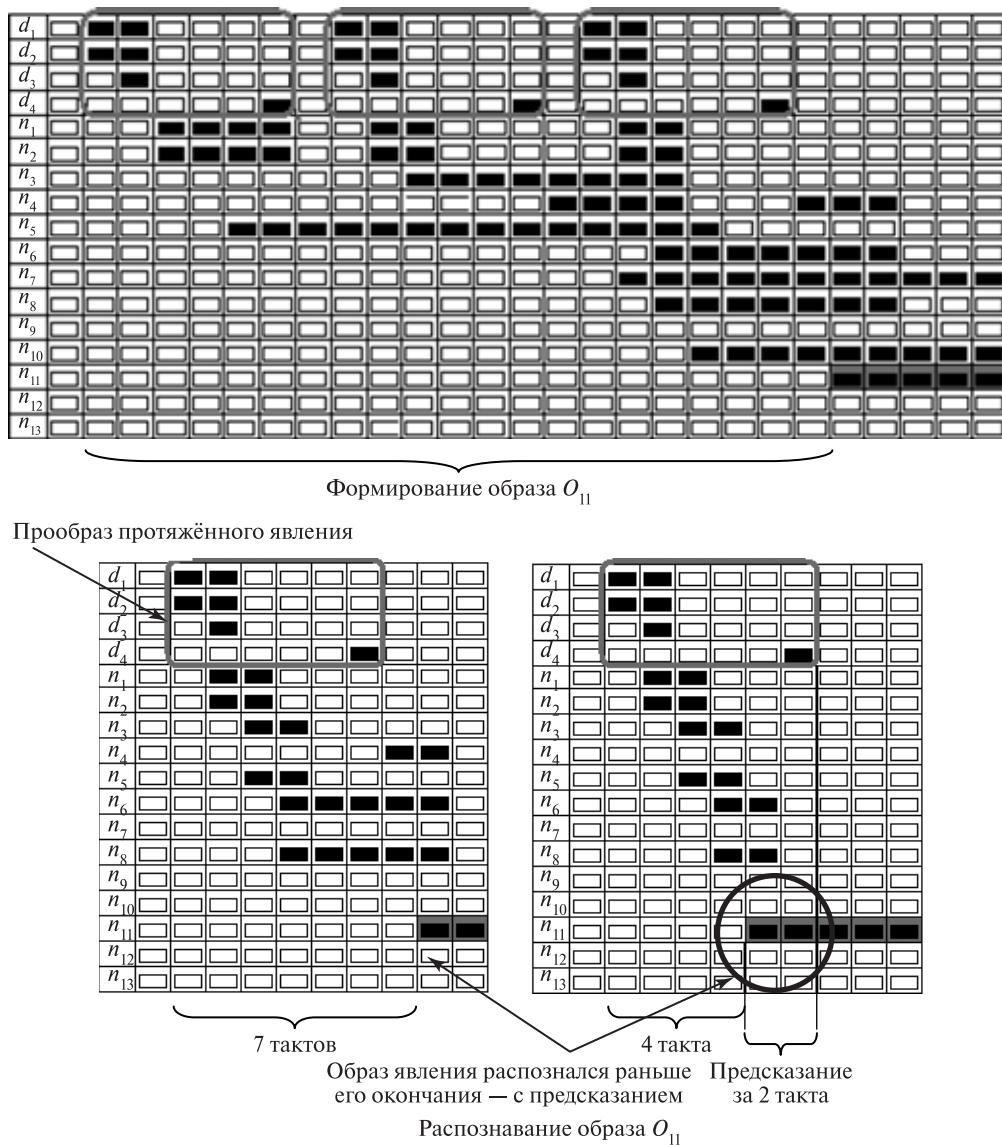


Рис. 8. Сеть из нейронов ААУ может распознавать явления раньше, чем закончится их демонстрация, т. е. с предсказанием

распознавать данный образ с первого предъявления — на фрагменте в правой части рис. 7 нейрон n_{11} реагирует на предъявление прообраза на такте $t = 24$. В данном случае нейрон n_{11} отражает конкретно данный образ, что и соответствует известной в биологии проблеме «нейрона моей бабушки» (известно, что в сетях ИНС такого нейрона нет).

Постепенное уменьшение порога ρ может обеспечивать возможность распознавания образов пространственно-временных явлений раньше, чем на датчиках закончится предъявление всего их прообраза (рис. 8), т. е. имеет место предсказание

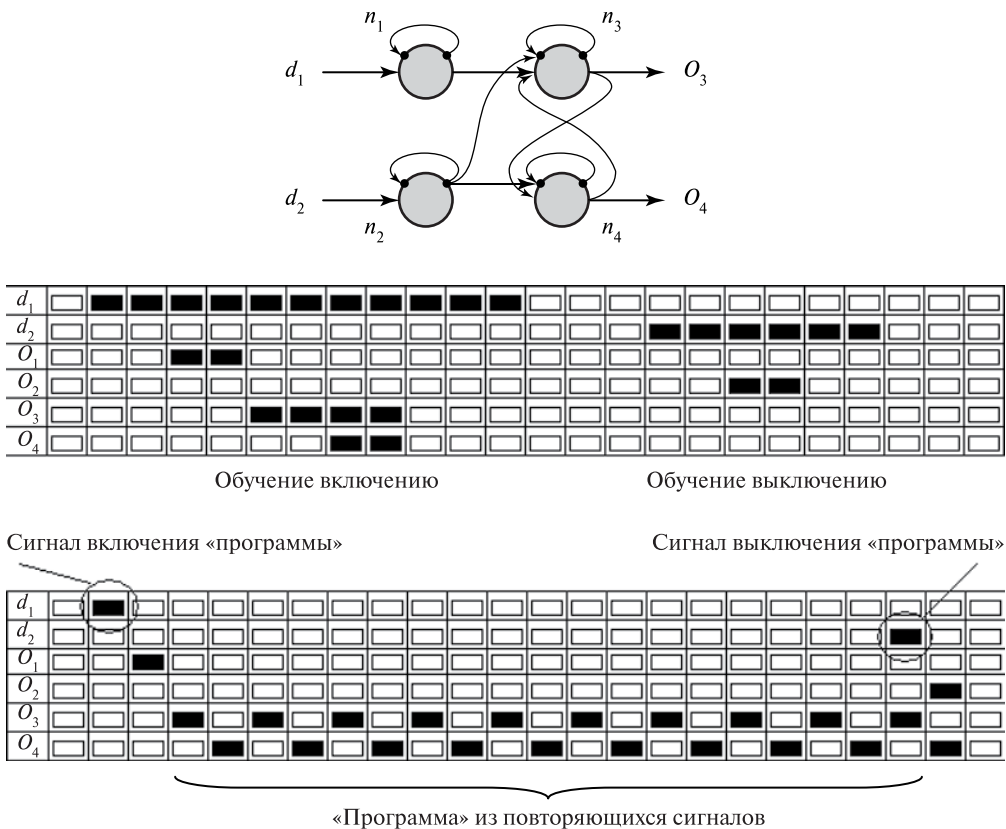


Рис. 9. Пример нейросетевой конструкции с обратными связями, управляющей периодическим процессом (например, ходьбой) её обучения, запуска и останова

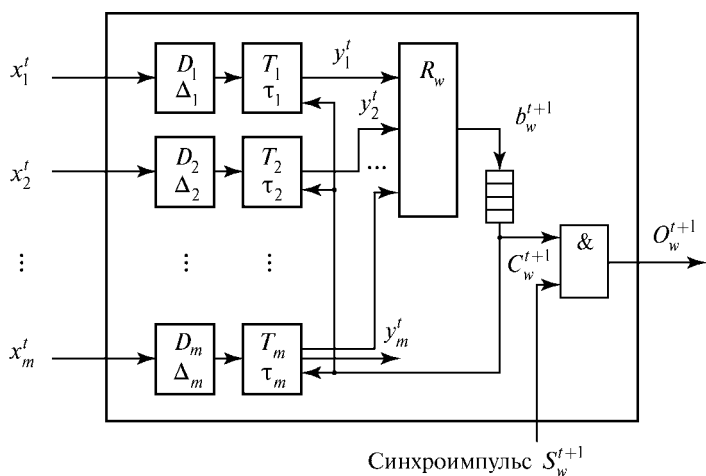


Рис. 10. Одна из модификаций модели нейрона, с более дифференцированным учётом синаптических временных задержек

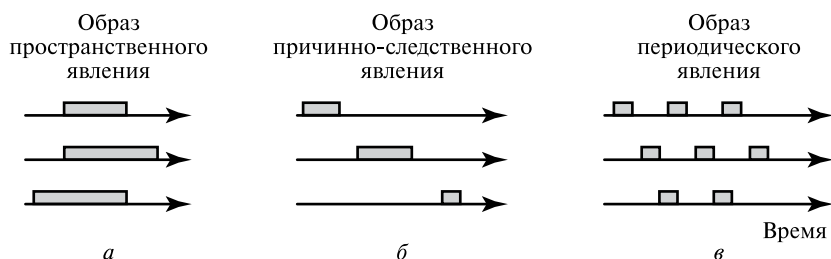


Рис. 11. Возможность формирования основных образов пространственно-временных явлений нейроном с временными синаптическими задержками

закономерностей, лежащих в основе возможности обоснованного принятия решений.

Используя возможности введения у нейрона обратных связей, моделирующих функции возвратной коллатерали, можно создавать обучаемые нейроноподобные «программы» управления периодическими движениями, например, исполнительных органов. Здесь имеет место не просто распознавание образов, но своего рода распознавание последовательных периодически повторяемых фаз некоторого длительного процесса, который можно запустить и в нужный момент остановить сигналами, подаваемыми на определённые нейроны (рис. 9, см. с. 262).

Нами разработаны несколько модификаций данной модели нейрона. Например, в одной из моделей учтены характеристики синаптических задержек, а также убрана внутрь нейрона отключающая обратная связь Z (рис. 10, см. с. 262). Формализация такого нейрона приведена, например, в работах [Жданов, 2009; Zhdanov et al., 2006].

Учитывая различные значения временных синаптических задержек, а также актуальные длительности возбуждений синапсов, можно заставить нейрон формировать образы пространственно-временных явлений всех основных существующих в природе типов, а именно (рис. 11): образы пространственных явлений, когда очередность распознаваний образующих не важна (*а*); образы причинно-следственных явлений, когда важна лишь последовательность событий (*б*) и образы строго регламентированных во времени (например, периодических) явлений (*в*).

Мы показали, что вместе с тоническими и фазическими сенсорами сети, состоящие из предлагаемых моделей нейрона, способны формировать образы очень сложных явлений, которые можно использовать для автономного адаптивного управления.

Надо упомянуть, что рассмотренные модели нейронов представляют собой, по сути, полный логический базис самообучаемых элементов «И», «ИЛИ» и «НЕ», из которых можно конструировать не только подсистемы ФРО, способные формировать образы любой сложности, но и все остальные подсистемы системы ААУ — «базу знаний», «аппарат эмоций» и др. Однако описание этих подсистем выходит за рамки темы технического зрения.

СИНТЕЗ СИСТЕМ ААУ

Синтез прикладных систем ААУ на основе описанных нейронов представляет собой отдельную проблему. В отличие от ИНС, где не слишком продвинутая

функциональность «формальных» нейронов оправдывается крайней простотой синтеза сетевых систем, в которых используются регулярные полнопереборные способы соединения нейронов, нерегулярные нейроноподобные сети ААУ подобны нерегулярным на вид сетям нейронов в нервных системах. Для синтеза подсистем системы ААУ из сетей описанных нейронов можно применять три подхода. Первый из них связан с максимальным использованием априорных знаний для конструирования сетей с определённым избытком нейронов, заведомо покрывающим те области нейронов, которые могут быть обучены. Второй подход [Жданов и др., 2001] связан с оптимизацией сетевых и других структур системы ААУ методами поисковой оптимизации, в частности генетическими алгоритмами. Здесь требуется моделирование многих поколений популяций индивидов. Этот подход соответствует природному способу эволюционной оптимизации видов в филогенезе. На рис. 12 показаны примеры нейроноподобной сети подсистемы ФРО для адаптивной системы управления мобильного робота, оптимизированные генетическими алгоритмами. Там же показаны примеры некоторых параметров системы технического зрения (радиусы и углы раствора полей зрения) для мобильного робота, оптимизированные генетическими алгоритмами в зависимости от предельной скорости движения робота и предельного радиуса поворота.

Третий подход связан с составлением таких программ автоматического синтеза нейросетевых структур, которые бы автоматически «выращивали» нейроноподобную сеть в том направлении, в котором она более всего востребована в данном приложении. Здесь речь идет о динамической оптимизации устройства отдельного индивида, что соответствует в биологии развитию в онтогенезе.

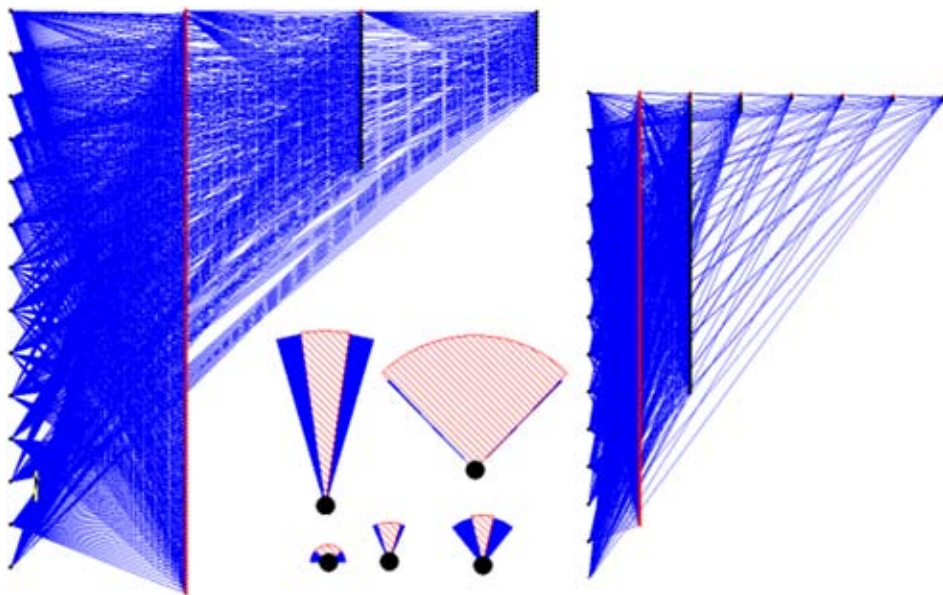


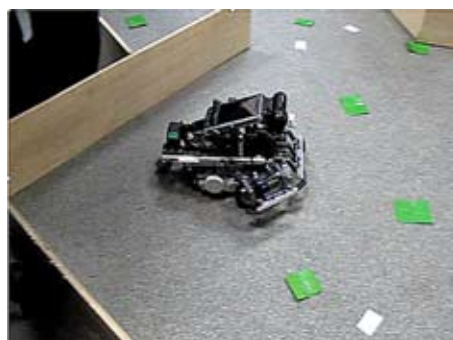
Рис. 12. Примеры нейроноподобной сети подсистемы ФРО для адаптивной системы управления мобильного робота, оптимизированные генетическими алгоритмами. Примеры характеристик системы технического зрения для мобильного робота, оптимизированные генетическими алгоритмами в зависимости от предельной скорости движения робота и предельного радиуса поворота

ОБ ИЕРАРХИЧНОСТИ АДАПТИВНЫХ СИСТЕМ УПРАВЛЕНИЯ

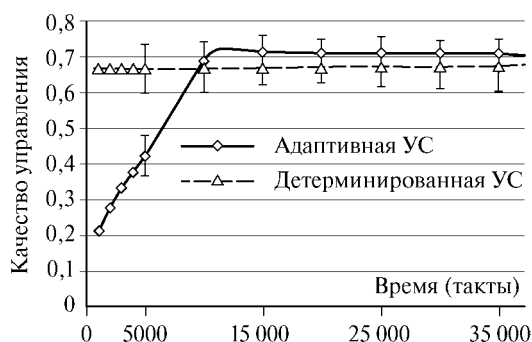
Необходимо отметить следующее. Всякий метод распознавания, обучающийся «без учителя», требует задания правил формирования образов. Это связано с тем, что любое множество объектов можно разбить на классы многочисленными способами, соответствующими разным принципам классификации. Другими словами, нельзя построить систему распознавания, которая научилась бы распознавать всё, что можно распознать в предъявляемых объектах. Надо обязательно определить правила, отвечающие на вопрос, «что именно мы хотим считать классами объектов». Поэтому для любой самообучаемой системы нужны априорно заданные правила формирования классов. Эта необходимая априорная информация отыскивается за пределами данной самообучаемой системы и некоторыми иными способами. Например, эффективные правила построения нейронной сети для мозга животного были найдены в процессе эволюционного развития вида его предками. Тем самым мы всегда должны иметь иерархическую систему адаптивного управления, в которой более широкие и медленндействующие механизмы адаптации находят правила для работы более локальных и быстродействующих уровней адаптивного управления.

НЕКОТОРЫЕ ПРИМЕРЫ ПРИКЛАДНЫХ СИСТЕМ

На основании описанной технологии построения систем ААУ было создано несколько прототипов таких машин и систем. На рис. 13 показан мобильный робот «Гном № 9» с адаптивной системой управления (данный робот является развитием его прототипной программной модели, описанной в работе [Жданов и др., 2004б]). Этот мобильный робот, подобно малолетнему ребенку, учится безаварийной езде на полигоне. Предварительно ему заданы только несколько качественных (эмоциональных) оценок для некоторых ситуаций: соударение со стенкой «больно» на «-3», а наезд на бумажку зелёного цвета «приятно» на «+5». Запущенный в полигон, робот вначале совершает беспорядочные движения, часто натываясь на препятствия. Постепенно он начинает «понимать», как связано то, что он видит, и то,



а



б

Рис. 13. Мобильный робот «Гном № 9» с адаптивной системой управления автоматически учится безаварийной езде и совершению полезной работы — сборке определённого мусора (а). Качество управления роботом постепенно растёт, догоняя и обгоняя качество управления, обеспечиваемое детерминированной системой управления, и выходя в насыщение, связанное с ограничением ресурсов бортового компьютера (б)

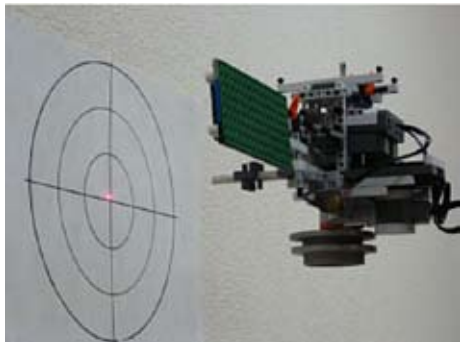


Рис. 14. Модель наноспутника с автономной системой стабилизации углового движения

что он делает, с тем, что из этого получается. Примерно через 10...15 мин робот уже успешно объезжает препятствия и начинает целенаправленно наезжать на зелёные бумажки. Робот также сам догадался, что взять зелёную бумажку, лежащую у стенки, выгодно, даже испытал соударение, так как « -3 » + « $+5$ » = « $+2$ ». Система машинного зрения робота состоит из лазерного луча, развернутого

в плоскость, параллельную полу, и видеокамеры, наблюдающей немного сверху линию пересечения луча с препятствиями. Передняя часть робота снабжена также тактильными датчиками. Подсистема ФРО робота построена по некоторой аналогии с глазом человека в том отношении, что также имеет трёхслойную нейроподобную сеть из описанных нейронов, которая формирует и затем распознает образы типичных препятствий — стенки, края стенки, ворот и некоторых других.

Ещё одним примером является модель наноспутника с автономной системой стабилизации углового движения (рис. 14). Программные прототипы этой системы управления были описаны в ряде работ например, в [Жданов и др., 2004a].

Модель спутника подвешена на нити и вращается в горизонтальной плоскости. «Эмоциональной подсистеме» спутника сказано, что её «ощущения» тем хуже, чем больше рассогласование реального углового положения спутника с заданным. Совершая вначале случайные воздействия, система ААУ постепенно находит способы управления данной машиной и через некоторое время самообучается устойчиво держать спутник в заданном положении. Подсистема ФРО спутника распознает образы углов рассогласования, угловой скорости, углового ускорения движения спутника и некоторых их комбинаций.

В двух рассмотренных случаях системы управления построены по одному и тому же методу ААУ и демонстрируют ожидаемые свойства — способность автоматически находить способ управления заданным объектом, не используя никаких «законов управления», априорно разработанных на основании аналитических моделей объекта управления.

Другие примеры прототипов прикладных систем на основе метода ААУ можно найти в работе [Новые подходы..., 2004].

Заключение

Нам представляется неизбежным, что следующий этап развития технологий будет состоять в создании автономных машин, системы управления которых по необходимости должны быть адаптивными. Примеры такого рода «техники» мы во множестве видим вокруг себя в живой природе. Эта технология должна прийти на смену нынешним технологиям одноразовой оптимизации систем на этапе их проектирования. Машины, способные самостоятельно накапливать информацию, обучаться и постепенно приспосабливаться к окружающим обстоятельствам и их изменениям, составят новое поколение машин, которое будет обладать *автономным искусственным интеллектом* и в общих чертах воспроизводить свойства всегда

и всюду приспособляющихся живых организмов. На пути создания такой техники стоит ещё много проблем. Надеемся, что данная работа внесёт свой вклад в их решение.

Литература

- Жданов А. А. (1999) Метод автономного адаптивного управления // Изв. РАН. Теория и системы управления. 1999. № 5. С. 127–134.
- Жданов А. А. (2009) Автономный искусственный интеллект: Монография. 2-е изд. М.: БИНОМ. Лаборатория знаний, 2009. 359 с.
- Жданов А. А., Земских Л. В., Беляев Б. Б. (2001) Применение генетических алгоритмов для оптимизации нейросетевой базы знаний адаптивной системы стабилизации углового движения космического аппарата // Сб. тез. докл. 15-х Академич. чтений по космонавтике. Москва, 24–26 янв. 2001. М.: Война и мир, 2001. С. 128–129.
- Жданов А. А., Земских Л. В., Беляев Б. Б. (2004а) Система стабилизации углового движения космического аппарата на основе нейроподобной системы автономного адаптивного управления // Космич. исслед. 2004. Т. 42. № 3. С. 1–15.
- Жданов А. А., Крыжановский М. В., Преображенский Н. Б. (2004б) Бионическая интеллектуальная автономная адаптивная система управления мобильным роботом // Мехатроника. Автоматизация. Управление. 2004. Ч. 1. № 1. С. 21–30; Ч. 2. № 2. С. 17–22.
- Новые подходы в нейроподобных и основанных на знаниях системах: Тр. Ин-та системного программирования РАН. Т. 7 / Под ред. А. А. Жданова. М.: ИСП РАН, 2004.
- Zhdanov A., Kondukov A., Naumkina T., Dmitrienko O. (2006) Automatic Origin of a Language in AAC Neuron-Like Systems // Proc. 11th Intern. Conf. “Speech and Computer SPECOM’2006”. 2006. P. 550–554.
- Zhdanov A. A., Ryadovikov A. V. (2000) Neuron Models in the Autonomous Adaptive Control Method // Optical Memory and Neural Network. Allerton Press, Inc. 2000. V. 9. N. 2. P. 115–132.

BIOLOGICALLY INSPIRED MACHINE VISION IN AUTONOMOUS ARTIFICIAL INTELLIGENCE SYSTEMS

A. A. Zhdanov

*JSC Lebedev Institute of Precision Mechanics and Computer Engineering
Russian Academy of Sciences*

The article presents the author’s biologically-inspired concept of the creation of autonomous artificial intelligence systems on basis of his method of autonomous adaptive control (AAC). Subsystem of «machine vision» as a self-learning recognition system is described in more details. The special attention is given to ways of construction of such systems on base of developed for AAC systems the self-learning models of neuron and neural networks. Author stresses that the machine vision system of robots should be adequate to their systems of decision-making and should not be a copy of human perception system. Some examples of robots with AAC control systems are presented. The AAC method can be used for creation of autonomous artificial intelligence systems.

Keywords: independent robots, adaptive management, recognition of images, neural networks, a method of independent adaptive management.

АВТОМАТИЧЕСКАЯ СИСТЕМА РЕАЛЬНОГО ВРЕМЕНИ ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ И ОРИЕНТИРОВ НА ИЗОБРАЖЕНИИ, ОСНОВАННАЯ НА ОБРАБОТКЕ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ

К. И. Кий

*Учреждение Российской академии наук Институт прикладной математики им.
М. В. Келдыша РАН (ИПМ им. М. В. Келдыша РАН),
125047, Москва, Миусская пл., 4*

В работе кратко излагается новый подход к качественному и количественному описанию цветных изображений и выделению объектов на изображениях, которые могут представлять интерес при решении различных практических задач. Проводится также анализ существующих методов для решения указанных задач и сравнение с предложенным методом. Приводятся примеры применения к задаче поиска ориентиров на изображениях и демонстрируются результаты обработки. Описываются параметры программной системы, разработанной на основе предложенного подхода.

Ключевые слова: компьютерное зрение, понимание изображений, сегментация изображений, зрение роботов.

Введение

Разработана система реального времени для обнаружения ориентиров и объектов на цветных изображениях, основанная на новом методе обобщенной сегментации и описания цветных изображений, предложенном автором. Несмотря на значительный прогресс методов сегментации цветных изображений, основанных на индексации изображений (разделении изображений на непересекающиеся области, которые однородны в том или ином смысле) [Forsyth, Ponce, 2003; Xie, Mirmehdi, 2007; Chang et al., 2007a, b; Deng, Manjunath, 2001; Mishra, Aloimonos, 2009] (также см. ссылки в указанных публикациях), практическая ценность этих методов желает много лучшего [Mishra, Aloimonos, 2009]. Анализ работы существующих методов может быть найден в [Mishra, Aloimonos, 2009; Kadir, Brady, 2001; Gribkov et al., 2009]. Особенно интересно обсуждение данной проблемы в [Mishra, Aloimonos, 2009]. Метод геометризованных гистограмм и метод, предложенный в [Mishra, Aloimonos, 2009], имеют некоторые общие предпосылки, но подходы к решению задач различны и, в некотором смысле, дополняют друг друга. В обоих методах вначале выделяются некоторые объекты внимания (хотя и разными способами), а затем уточняется их конфигурация. В методе геометризованных гистограмм это делается с упором на «тело» объекта, а в методе из [Mishra, Aloimonos, 2009] вначале осуществляется обнаружение характерных признаков по [Kadir, Brady, 2001], а затем находится контур, ограничивающий выделенный объект.

На реальных изображениях объекты, как правило, не являются однородными, а состоят из частей, которые могут быть однородными в разных смыслах (интенсивность, цвет, оттенок, текстура и т. д.). Кроме того, для решения задачи навигации в робототехнике не нужно решать задачу индексации всех точек изображения

Кий Константин Иванович — старший научный сотрудник, кандидат физико-математических наук, e-mail: kikip_46@mail.ru.

[Mishra, Aloimonos, 2009]. Необходимо лишь детектировать потенциально опасные (близкие, движущиеся быстро в некотором направлении) и интересные (ориентиры) объекты и предсказать их поведение во времени (разделить движущиеся и статические объекты). Задача обнаружения и распознавания опасных объектов и ориентиров требует нахождения в кадре малых объектов (если они движутся с большой скоростью — чтобы выработать управляющую реакцию заранее, или если они могут пропасть из кадра, заслонённые другими объектами, — чтобы не потерять их). Нахождение малых объектов также важно при обнаружении таких ориентиров как сигналы светофора, дорожные знаки, тормозные сигналы автомобилей и т. д. Заметим, что обнаружение малых объектов также полезно при создании робота-ассистента [Mishra, Aloimonos, 2009]. Работа с такими объектами наряду с обеспечением реального времени является слабым местом общих алгоритмов сегментации (см. примеры работы общих алгоритмов сегментации на изображении природной сцены в [Mishra, Aloimonos, 2009]). Поэтому в основу разработанной системы положен метод геометризованных гистограмм изображений, разработанный в [Kiy, 1998, 2008, 2010; Kiy, Dickmanns, 2004; Kuy, 2009]. Данный метод обеспечивает технику для обнаружения малых объектов и работу с объектами, составленными из нескольких однородных частей. С помощью этого метода также возможен поиск текстурных объектов. Кроме того, метод позволяет работать с цветовыми объектами с малым насыщением цветов и использует введение «наивных» белых, серых и чёрных цветов [Kuy, 2009] для приближения к методам, применяемым человеческим зрением при обработке изображений.

1. КРАТКОЕ ОПИСАНИЕ ПРЕДВАРИТЕЛЬНЫХ РЕЗУЛЬТАТОВ ОБРАБОТКИ ИЗОБРАЖЕНИЯ, ПОЛУЧАЕМЫХ СИСТЕМОЙ

В методе геометризованных гистограмм производится объединение информации о частотах значений функции, задающей изображение (скалярной или векторной, как в цветных изображениях), и некоторого способа геометрического описания её линий уровня. Во многих серьезных работах по компьютерному зрению [Chang et al., 2007a, b] было отмечено, что для разработки тонких инвариантов изображений необходимо произвести указанный синтез. Несмотря на то, что классические гистограммные методы на простых сценах позволяют получать результаты в реальном времени, замечание о том, что гистограммы сохраняются всеми взаимно-однозначными преобразованиями плоскости, показывает, что они применимы только для анализа очень ограниченного класса изображений. Всё это справедливо независимо от того, какие тонкие методы кластеризации применяются, чтобы найти закономерности распределения частот в этих гистограммах (независимо от того, какие гистограммы рассматриваются — векторные или скалярные), так как если объекты потеряны на уровне построения гистограммы (возможного её сглаживания), их уже не вернуть.

В методе геометризованных гистограмм анализ изображений основан на синтезе информации о частотах значений функции интенсивности и некотором геометрическом описании вариаций этих значений на плоскости изображения. В основу метода положена идея о том, что значения функции близки, не только когда они близки как целые числа, но и тогда, когда близки подмножества изображения, на которых эти значения принимаются. Близость между подмножествами меряется с помощью специальных мер близости [Kiy, 1998, 2008, 2010; Kiy, Dickmanns, 2004; Kuy, 2009]. Это предположение приводит к тому, что в математической задаче,

вместо классического n -мерного пространства, некоторое расслоение пространств отрезков на прямой является классифицирующим пространством. Чтобы формализовать задачу таким образом, при обработке исходное изображение разбивается на узкие полосы одинаковой ширины, параллельные горизонтальной или вертикальной оси системы координат плоскости изображения. В деталях процедура построения геометризованных гистограмм изображений описывается в [Kiy, 2010].

Приведем её краткое описание и мотивировку конструкции. Легко заметить, что для дискретного изображения одномерного мира, заданного одной строкой значений функции интенсивности, линии уровня, соответствующие фиксированным значениям функции, есть интервалы на оси одномерного мира (может быть, вырождающиеся в отдельные пиксели). Задача геометрического описания линий уровня в данном случае имеет тривиальное решение. Каждая линия уровня есть объединение непересекающихся отрезков. Если мы рассмотрим двумерное изображение, заданное на узкой полосе, то линии уровня функции, задающей изображения, также можно приблизительно описать с помощью проекции на ось полосы Ox . При проецировании множества точек, соответствующих фиксированному значению функции, мы получаем систему интервалов на оси Ox . Каждому значению функции будет соответствовать своя система интервалов. Ясно, что, в отличие от одномерного мира, мы получаем перекрывающиеся системы интервалов. Однако каждая из них приблизительно описывает место в полосе, где функция принимает фиксированные значения (точность представления зависит от ширины полосы). Кроме того, каждому интервалу, соответствующему некоторому фиксированному значению интенсивности, ставится в соответствие некоторое число — его мощность. Мощность интервала равна числу точек полосы, проецирующихся на данный интервал на оси Ox . По набору мощностей систем интервалов однозначно восстанавливается классическая гистограмма функции интенсивности изображения в данной полосе. Рисунок 1 иллюстрирует построение системы интервалов геометризованной гистограммы.

В верхней части рис. 1 показана функция интенсивности, задающая изображение в одномерном мире. Она принимает постоянные значения на интервалах, пересекающихся только в граничных точках. Нижняя часть рисунка показывает часть полосы, в которой функция интенсивности принимает два значения, определяемые разными способами штриховки. В самой нижней части рисунка изображены

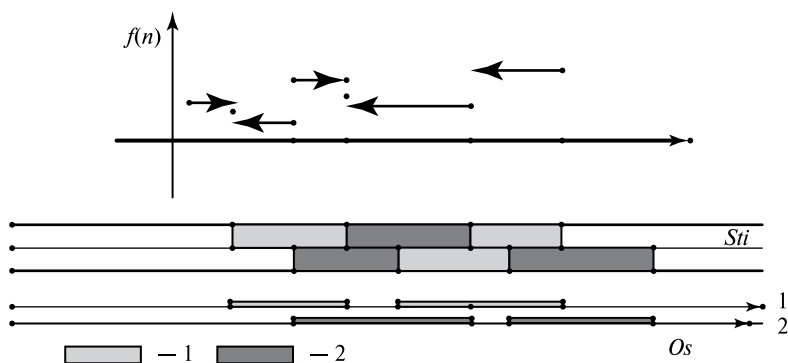


Рис. 1. Построение системы интервалов в геометризованной гистограмме

два экземпляра оси O_s с полученными двумя системами интервалов, соответствующих двум выбранным значениям.

Геометризованная гистограмма дает также некоторое формализованное описание текстур. Пусть выделена прямоугольная часть полосы, в которой функция интенсивности принимает значения z_1, \dots, z_k . Предположим, что точки, в которых принимаются эти значения, равномерно перемешаны в этой части полосы так, что в каждую точку оси O_s проецируются точки, соответствующие всем значениям z_j . Тогда при проекции на ось O_s этой части полосы отрезки, соответствующие всем значениям z_i , совпадут. Мощности отрезков зададут классическую гистограмму функции интенсивности, и совпадение всех отрезков всё равно даёт полезную дополнительную информацию, детектируя возможную текстурную структуру. Для более тонкого изучения текстуры необходимо рассматривать разбиения на более узкие полосы и одновременно разбиения на вертикальные и горизонтальные полосы. Вышесказанное дает некоторый способ выделения и характеристики текстур на изображениях.

Заметим, что геометризованную гистограмму сохраняют взаимно-однозначные автоморфизмы полосы, переставляющие точки, имеющие одинаковую проекцию на ось O_s . Это означает, что её группа автоморфизмов значительно меньше, чем у классической гистограммы. Если мы рассмотрим одновременно два разбиения изображения на вертикальные и горизонтальные узкие полосы, то у пары геометризованных гистограмм, соответствующих этим двум разбиениям, группа автоморфизмов редуцируется в общем случае ещё более значительно.

Пример построения геометризованной гистограммы и её графическая визуализация могут быть найдены в [Kiy, 1998, 2010; Kiy, Dickmanns, 2004]. В работе [Kiy, 2010] указано, как обобщить конструкцию геометризованной гистограммы на случай цветных изображений. Для цветного изображения в узкой полосе геометризованная гистограмма задается системой интервалов на оси O_s такой, что каждому интервалу поставлены в соответствие его мощность и набор чисел [$IntBeg$, $IntEnd$] — начало и конец интервала на оси O_s ; [H_{min} , H_{max}] и H_{mean} — диапазон и среднее значение цветового оттенка; [S_{min} , S_{max}] и S_{mean} — диапазон и среднее значение цветового насыщения; [I_{min} , I_{max}] и I_{mean} — диапазон и среднее значение полутоновой компоненты. Однако геометризованная гистограмма является всё ещё громоздким объектом. Чтобы редуцировать информацию, вводятся процедуры кластеризации на пространстве интервалов геометризованной гистограммы [Kiy, 2010]. В результате получается набор цветовых сгустков в полосе. Каждый цветовой сгусток в полосе задается тем же самым набором ($[IntBeg, IntEnd]$, [H_{min}, H_{max}], H_{mean} , [S_{min}, S_{max}], S_{mean} , [I_{min}, I_{max}], I_{mean}) и значением мощности, но уже объединяет несколько интервалов геометризованной гистограммы с подобными характеристиками.

Результатом работы алгоритмов обработки изображения является построение структурного графа цветовых сгустков данного цветного изображения. Вершинами данного графа являются цветовые сгустки, обнаруженные в полосах изображения (вертикальных или горизонтальных) [Kiy, 1998, 2008, 2010; Kiy, Dickmanns, 2004; Куй, 2009]. Цветовой сгусток может быть либо единственным в своей части полосы, либо одним из нескольких. Второй случай возможен в точках, где встречаются границы нескольких областей, окрашенных в разные цвета, либо в точках цветовой структуры. Каждый из цветовых сгустков является кластером окрашенных интервалов. Окрашенный интервал характеризуется координатами начала и конца на соответствующей оси и средними значениями цветового оттенка, насыщения,

интенсивности полутоновой компоненты и диапазонами изменения указанных характеристик. Рёбра графа соединяют вершины (цветовые сгустки) в одинаковых и соседних полосах, которые «непрерывно продолжают друг друга» [Kiy, 1998, 2008, 2010; Kiy, Dickmanns, 2004; Kuy, 2009]. Связные компоненты путей на этом графе, построенные по неким правилам, являются кандидатами на объекты в рассматриваемом изображении. Структурный граф позволяет отыскивать как части реальных объектов с однородными цветовыми характеристиками, объединяющие несколько цветовых сгустков (например, шкаф в помещении или входная дверь, окрашенные определённым образом), так и объекты, составленные из разнородных цветовых компонент (люди с предметами одежды и непокрытыми одеждой частями тела разного цвета, деревья (с вариацией цветов на стволах и темных местах), специальные ориентиры (дорожные знаки) и т. д.).

На структурном графе вводятся дополнительные отношения, которые позволяют получить с его помощью более точное описание исходного изображения. Для всех вершин, принадлежащих одному слою (полосе) и соседним слоям (полосам), определены отношения соседства в виде списков соседних вершин. Кроме того, цветовым сгусткам в каждой полосе присваиваются некоторые качественные характеристики типа «чистый сгусток, порождающий полосу», «доминирующий сгусток», «один из доминирующих сгустков», «существенный контрастный сгусток на фоне доминирующего», «заметный сгусток», «малозаметный сгусток». Для выполнения указанной классификации разработана система продукций, которая, основываясь на численных характеристиках цветовых сгустков, делает заключение о принадлежности сгустка к тому или иному классу. При написании продукций принимается во внимание субъективная экспертная информации, полученная в результате экспериментов с большим числом изображений.

Для решения задач робототехники, например, для построения робота-ассистента [Mishra, Aloimonos, 2009], необходимо научить робота различать цвета образом, подобным тому, который используется человеком при анализе сцены. Необходимо отметить, что цвета на телевизионном изображении существенно зависят от освещения и самой телевизионной камеры. Имеется ряд работ, посвященных задаче определения объективного цвета предмета в кадре (см. [Finlayson et al., 2006] и ссылки в этой работе). Однако, как правило, в работах данного направления изучается простая сцена (один предмет при разном освещении). Описания сцены, разработанные в методе геометризованных гистограмм, позволяют перейти в задаче определения реальных цветов предметов в сложных сценах. Об этом будет говориться в следующих работах.

2. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Предложенная система реализована в виде комплекса программ на C++, стыкованного с вводом изображений (с видеокамеры или видеофайла) на основе DirectX. Комплекс может работать через сеть (включая Internet) и обрабатывает цветные изображения разрешения 640×480 пикселей со скоростью 10 кадров/с на стандартном РС (3 ГГц). Скорость обработки пропорциональна объёму массива изображения (для разрешения 320×240 пикселей скорость обработки более 30 кадров/с).

Многочисленные компьютерные эксперименты по сегментированию изображений и видеопоследовательностей были выполнены с использованием данного комплекса программ. Они показали его способность находить ориентиры и реальные объекты при различных условиях освещённости. Изучаются возможности

разработанного комплекса для применения в анализе движения и создания стереосистемы на сложных многообъектных сценах. Примеры работы системы могут быть найдены в презентации, расположенной на сайте конференции (<http://tz2010.cosmos.ru>) и на сайте (Color vision). Так как они требуют существенного использования цвета, в настоящей работе число примеров сокращено.

3. ПРИМЕНЕНИЕ К ПОИСКУ ОРИЕНТИРОВ

Разработанный метод показал хорошие результаты при описании уличных сцен и выделении характерных объектов на них. На рис. 2 приведено цветное изображение уличной сцены и цветовые сгустки, найденные данным комплексом программ, наложенные на изображение, задаваемое полутоновой компонентой исходного изображения. Из рисунка видно, что большинство существенных объектов найдено на изображении. Цветовое описание исходного изображения разрешения 320×240 пикселей, которое разбито на 24 горизонтальные полосы одинаковой ширины, задается парой сотен цветовых сгустков со структурными связями. Цветовые сгустки наложены на средние линии каждой полосы. Структурный граф цветовых сгустков может быть найден в презентации доклада, размещенной на сайте (<http://tz2010.cosmos.ru>). Другие примеры результатов обработки для различных сцен размещены на сайте (Color vision). Проведены также эксперименты, которые показывают, что система способна отыскивать заданные предметы в помещении [Mishra, Aloimonos, 2009].

Для применения в навигационной задаче робота, для тестирования системы была изготовлена метка, состоящая из трёх последовательно соединенных прямоугольников зелёного, оранжевого и белого цветов, и эксперименты проводились в помещениях с различным освещением, которое могло быть естественным, искусственным и их сочетанием. Эксперименты показали, что реальный цвет на изображении указанных прямоугольников варьируется в очень широком диапазоне и сильно зависит от типа освещения и качества телевизионной камеры. Это означает, что если допускается изменение параметров освещённости в широком диапазоне, то поиск на изображении прямоугольных объектов с заданными цветами не может привести к результату. На ряде изображений белая метка может иметь цвет от синего до зелёного, при этом её насыщение может достигать до 50 %. Однако



Рис. 2. Результаты обработки уличной сцены (цветовые сгустки), наложенные на полутоновую компоненту изображения

её полутоновая компонента имеет наибольшие значения интенсивности. Поэтому поиск на структурном графе цветовых сгустков в виде тройки соседних окрашенных объектов с условием, что цвета оранжевой и зелёной части варьируются в некоторой достаточно большой окрестности этих цветов, а белая часть имеет большую интенсивность полутоновой компоненты, позволяет находить метку в условиях широкой вариации освещённости. Система показала возможность нахождения метки даже при условиях плохой освещённости на сложных загромождённых сценах и возможных фонах с близким цветом к одному из цветов метки. В свою очередь нахождение метки позволяет восстанавливать истинные цвета объектов в кадре. Пример структурного графа цветовых сгустков для сцены в помещении приведен в презентации (<http://tz2010.cosmos.ru>).

4. СРАВНЕНИЕ С ИЗВЕСТНЫМИ СИСТЕМАМИ

По сравнению с известными методами сегментации метод геометризованных гистограмм позволяет находить более тонкие детали, которые не доступны общим методам. Метод геометризованных гистограмм не требует гипотез о числе областей и не содержит задаваемых порогов. Также ввиду глобальности подхода он позволяет применять семантические понятия, например сегментацию по форме объекта. Необходимо также заметить, что общие методы сегментаций плохо работают в условиях, когда осуществляется постепенный переход между близкими цветами [Gribkov et al., 2009]. В методе геометризованных гистограмм этот недостаток в значительной степени устраняется с помощью специальной процедуры разделения цветовых диапазонов [Kiy, 2010]. Например, на основном изображении в работе [Mishra, Aloimonos, 2009], в отличие от результатов сегментации, полученных общими методами, которые приведены на рисунках в [Mishra, Aloimonos, 2009], метод геометризованных гистограмм позволяет наряду с большими объектами, такими как газон и деревья, находить более мелкие, но характерные объекты, такие как стадо овец вдали и изгородь на переднем плане.

Кроме упомянутых систем сегментации изображений, заслуживают отдельного рассмотрения так называемые детекторы характерных признаков (salient feature detectors) [Kadir, Brady, 2001]. Они выделяют характерные точки на основе локальных рассматриваний, может быть с использованием изменения шкалы рассматривания локальных признаков [Kadir, Brady, 2001]. Недостаток этих методов — невозможность использования семантики при поиске, а также время вычислений, которое является неприемлемо большим для обработки в реальном времени. Эти методы основаны на изучении локальных гистограмм (распределения вероятностей значений в окрестности точек) функций, задающих изображения. С другой стороны, метод геометризованных гистограмм, не снимая задачи локального изучения поведения дескрипторов изображения, позволяет решать задачу выделения характерных областей, используя глобальные аргументы, включая семантику. При сравнении работы на одних и тех же изображениях метода геометризованных гистограмм и детекторов характерных признаков мы получаем, что метод геометризованных гистограмм работает лучше во многих случаях, особенно для сложных сцен с объектами, окрашенными в разные цвета. Дополнительным его преимуществом является отсутствие скрытых настраиваемых порогов. Построению детектора характерных признаков на основе метода геометризованных гистограмм будет посвящена отдельная публикация.

Заключение

Разработан новый метод представления сцен, основанный на понятии геометризованной гистограммы, удобный для решения задач понимания изображений в реальном времени. Основу представления данных составляет структурный граф, поставленный в соответствие любому цветному изображению. Геометризованная гистограмма и структурный граф позволяют разделять и находить различные предметы на изображениях. Предложенная техника дает подход к интерпретации цветовых характеристик объектов на изображении, сходный с интерпретацией, выполненной человеком. Разрабатывается система поиска ориентиров с разнообразными описаниями на изображениях, основанная на развитой технике. Необходимо строить разные правила интерпретации, повысить точность используемых параметров и организовать процедуру обучения с целью выбора оптимальных правил для разных классов изображений.

Литература

- Кий К. И. (2009) Геометризованные гистограммы и понимание изображений // Тр. конф. по математическим методам распознавания образов ММРО-14. Суздаль, 2009. М.: Макс Пресс, 2009. С. 362–365.
- Техническое зрение в системах управления мобильными объектами-2010: Научно-технич. конф.-семинар. Москва, ИКИ РАН, 16–18 марта 2010. [Электронный ресурс]. Режим доступа: <http://tz2010.cosmos.ru/>.
- Chang Y. C., Archibald J. K., Wang Y.-G., Lee D. J. (2007a) Texture-based color image segmentation using local contrast information // Intern. J. Information Technology and Intelligent Computing. 2007. V. 2. N. 4.
- Chang Y. C., Lee D. J., Archibald J. K. (2007b) Using color variation to mimic human visual perception for color image quantization // Intern. J. Information Technology and Intelligent Computing. 2007. V. 2. N. 3.
- Color vision: [Электронный ресурс]. Режим доступа: <http://sites.google.com/site/colorvision-kikiy/>.
- Deng Y., Manjunath B. S. (2001) Unsupervised segmentation of color texture regions in images and video // IEEE Trans. Pattern Analysis and Machine Intelligence. 2001. V. 23. N. 8. P. 800–810.
- Finlayson G. D., Hordley S. D., Drew M. S., Lu C. (2006) On the removal of shadows from images // IEEE Trans. Pattern Analysis and Machine Intelligence. 2006. V. 28. N. 1. P. 59–68.
- Forsyth D. A., Ponce J. (2003) Computer Vision: a Modern Approach. N. Y.: Prentice Hall, 2003.
- Gribkov I. V., Koltsov P. P., Kotovich N. V. et al. (2009) Robustness of noisy and blurry images segmentation // Pattern Recognition and Image Analysis. 2009. V. 19. N. 3. P. 484–490.
- Kadir T., Brady M. (2001) Saliency, scale and image description // Intern. J. Computer Vision. 2001. V. 45. N. 2. P. 83–105.
- Kiy K. I. (1998) An unsupervised color vision system for driving unmanned vehicles // Proc. SPIE AeroSense'98 Symp. Enhanced and Synthetic Vision 1998. Orlando, USA / Ed. J. G. Verly. 1998. V. 3364. P. 371–382.
- Kiy K. I. (2008) A new method for description and generalized segmentation of color images in real time // Proc. Intern. Conf. Pattern Recognition and Image Analysis: New Information Technologies. Nizhni Novgorod, Russia. 2008. P. 297–300.
- Kiy K. I. (2010) A new real-time method for description and generalized segmentation of color images // Pattern Recognition and Image Analysis. 2010. N. 2. P. 169–178.
- Kiy K. I., Dickmanns E. D. (2004) A color vision system for real-time analysis of road scenes // Proc. IEEE Intelligent Vehicle'04 Symp. Parma, Italy. 2004. P. 54–59.

- Mishra A. K., Aloimonos Y.* (2009) Active segmentation // Intern. J. Humanoid Robotics. 2009. V. 6. N. 3. P. 361–386.
- Xie X., Mirmehdi M.* (2007) Colour image segmentation using texems // Annals of BMVA. 2007. N. 1. P. 1–10.

**AN AUTOMATIC REAL-TIME SYSTEM FOR DETECTING
OBJECTS AND LANDMARKS IN IMAGES BASED
ON PROCESSING COLOR IMAGES**

K. I. Kiy

*Keldysh Institute of Applied Mathematics of Russian Academy of Sciences,
4 Miusskaya sq., Moscow, 125047, Russia*

In this paper, a new approach to qualitative and quantitative description of color images and to detection of objects in images which can be of interest in solving various applied problems is described briefly. The analysis of existing methods for solving the specified problems is presented and the comparison of these methods with the proposed method is provided. Examples of application to the problem of finding landmarks in images are presented and the results of processing are demonstrated. The capability parameters of the software system, developed based on the proposed approach, are listed.

Keywords: computer vision, image understanding, image segmentation, robot vision.

Kiy Konstantin Ivanovich — Senior Scientist, Candidate of Physical and Mathematical Sciences, e-mail: kikip_46@mail.ru.

ИСПОЛЬЗОВАНИЕ ОСОБЕННОСТЕЙ СОВРЕМЕННОГО ПРЕДСТАВЛЕНИЯ ЦИФРОВЫХ ВИДЕОДАНЫХ ДЛЯ ОПТИМИЗАЦИИ АЛГОРИТМОВ ОБРАБОТКИ

К. Н. Стрельников

*Московский государственный университет им. М. В. Ломоносова (МГУ),
119991, Москва, ГСП-1, Ленинские горы*

В данной статье предлагается быстрый алгоритм обнаружения объектов с помощью поворотной IP-камеры. Высокая скорость достигается благодаря новому подходу к оптимизации обработки кодированного видеопотока с использованием служебной информации кодека.

Работа выполнена при поддержке РФФИ (проект № 10-01-00697-а).

Ключевые слова: оптимизация, видеокодеки, IP-камера, обнаружение объектов.

Введение

Все видеоданные можно разделить на два типа: цифровые и аналоговые. Обработка и тех и других, как правило, ведётся цифровым способом, поскольку такой подход предоставляет очень широкие возможности. При этом для обработки аналогового сигнала применяются аналогово-цифровые преобразования. Данный факт, а также удешевление чипов для обработки сигналов (ПЛИС, DSP, FPGA) приводит к повсеместному вытеснению аналоговых видеоданных цифровыми.

Важно отметить, что в цифровых видеоданных практически не встречаются последовательности независимых кадров. Такое представление требует больших объёмов памяти для хранения видеоданных и высокоскоростных каналов для их передачи. Значительное уменьшение размеров представления видеоданных достигается использованием высокой коррелированности соседних кадров. При этом хранятся не непосредственно кадры, а сжатые межкадровые разницы. Эта идея реализуется при использовании видеокодеков.

Для повышения коррелированности современные видеокодеки (H.264, MPEG-4) перед получением межкадровой разницы выполняют компенсацию движения между парой соседних кадров. Для этого используется аппарат векторов движения для блоков. Этот же аппарат может применяться во многих алгоритмах обработки видеоданных, таких как: адаптивное шумоподавление, стабилизация видео, изменение частоты кадров и др. При обработке кодированного видеосигнала информация о векторах движения доступна и её использование может значительно повысить скорость обработки видеоданных.

Для получения векторов движения кадр разбивается на блоки фиксированного размера и для каждого блока ищется наиболее похожий в некоторой области предыдущего кадра. Вектор движения — разница между положением в текущем и в предыдущем кадре. Обзор алгоритмов поиска векторов движения для блоков можно найти в статье [Гришин и др., 2008].

Далее описывается быстрый алгоритм обнаружения объектов с помощью поворотной IP-камеры, использующий информацию о векторах движения из кодированного потока.

ОБНАРУЖЕНИЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ

Задача автоматического обнаружения объектов в видеопотоке изучается длительное время. Большим импульсом к развитию данной тематики послужил переход на цифровые технологии видеонаблюдения в начале 1990-х гг. С этого момента технологии обнаружения объектов в видео стали востребованы повсеместно, а рост вычислительных мощностей позволил применять всё более сложные алгоритмы обработки. В настоящее время видеонаблюдение переживает очередную революцию. Повсеместное распространение получают IP-камеры. Их используют как профессионалы для решения задачи видеонаблюдения, так и обычные люди, например, транслируя вид из своего окна в Интернет.

Под IP-камерой понимают цифровую видеокамеру, особенностью которой является передача видеопотока в цифровом формате по сети Ethernet, использующей протокол IP. Являясь сетевым устройством, каждая IP-камера в сети имеет свой IP-адрес. В отличие от аналоговых камер при использовании IP-камер после получения видеокadra с ПЗС (CCD) или КМОП (CMOS)-матрицы камеры изображение остаётся цифровым вплоть до отображения на мониторе. Как правило, перед передачей полученное с матрицы изображение сжимается с помощью потоковых (MJPG) или потоковых (MPEG-4) методов сжатия видеоданных, благодаря чему в IP-камерах могут использоваться большие разрешения, включая мегапиксельные. Типичное разрешение для сетевых камер: 640×480 точек. Существуют камеры с мегапиксельными разрешениями: 1280×1024 , 1600×1200 и более высокими. Максимальные возможности для видеонаблюдения предоставляют поворотные (pan-tilt-zoom, PTZ) IP-камеры. Данная статья ориентирована на поворотные IP-камеры с потоковым методом сжатия видеоданных.

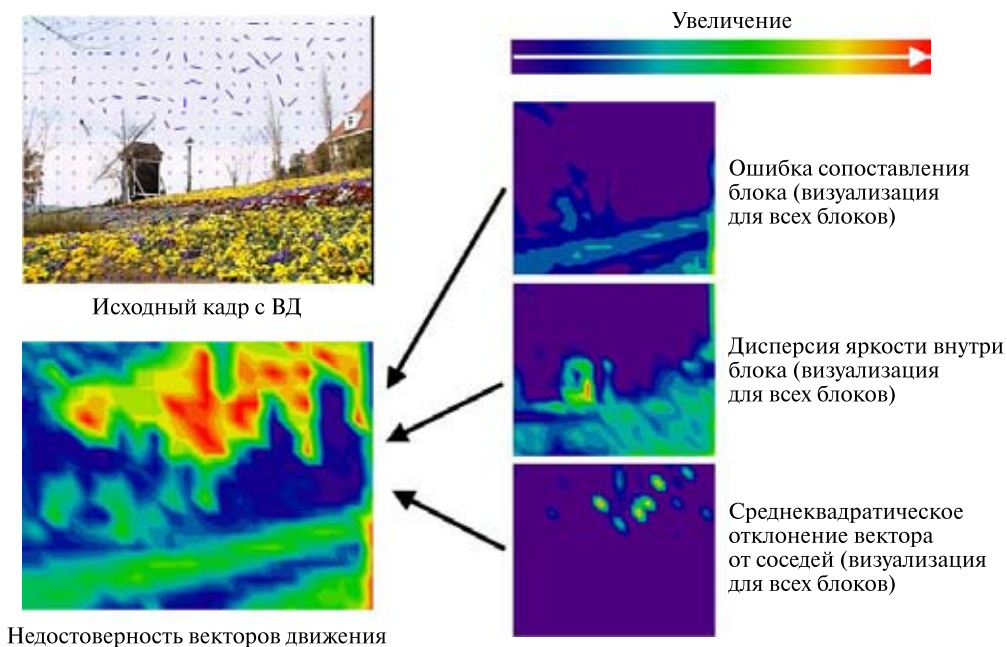


Рис. 1. Оценка достоверности векторов движения

Существует два подхода к обнаружению объектов: *моделирование фона* (когда строится статистическая модель фона и анализируется отличие кадра от построенной модели) и *анализ движения* (когда для каждого пиксела определяется вектор смещения относительно положения в предыдущем кадре, а выделение объектов осуществляется путём анализа полученного поля векторов). С последними разработками в области обнаружения объектов посредством моделирования фона можно ознакомиться в статьях [Staufer, Grimson, 2000; Пешков, 2004; Goldberger, Greenspan, 2006]. При использовании поворотных камер для моделирования фона используются сферические панорамы [Biswas et al., 2006]. Выделение объектов на видео, полученном от неподвижной камеры, при помощи анализа движения, как правило, сводится к поиску областей с движением, отличным от нулевого [Latecki et al., 2004]. В случае подвижных камер требуется отделить движение камеры от движения объектов в кадре, это актуально и для фиксированных камер, которые могут колебаться из-за ветра. Пример такого алгоритма можно найти в работе [Borshukov et al., 1997]. Скорость работы этих алгоритмов невысока, поскольку в процессе обработки строится и анализируется оптический поток.

В данной статье предлагается быстрый алгоритм обнаружения движущихся объектов в видео, полученном от движущейся камеры. Алгоритм базируется на информации о векторах движения, извлечённой из кодированного видеопотока от камеры. Таким образом, не требуется проводить трудоёмкие вычисления оптического потока.

БЫСТРЫЙ АЛГОРИТМ ОБНАРУЖЕНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ

Ключевым моментом предлагаемого метода обнаружения движущихся объектов в видео, полученном от движущейся камеры, является алгоритм выделения из векторного поля компоненты движения, связанной со смещением камеры. Эта компонента описывается моделью аффинного искажения:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_0 \\ a_5 \end{pmatrix}.$$

Для её определения в предлагаемом подходе также используются векторы движения (ВД).

Векторы движения не всегда соответствуют реальному движению в сцене, поэтому, прежде чем использовать их, необходимо избавиться от недостоверных векторов с помощью функции доверия векторам движения:

$$\Psi(B_{ij}, v_{ij}) = k_1 F(B_{ij}, v_{ij}) + k_2 G(v_{ij}, \{V\}) + k_3 H(B_{ij}),$$

где k_1, k_2, k_3 — весовые коэффициенты; $F(B_{ij}, v_{ij})$ — ошибка сопоставления блоков; $G(v_{ij}, \{V\})$ — отклонение вектора от соседей; $H(B_{ij})$ — оценка наличия деталей в блоке изображения.

На рис. 1 (см. с. 278) приведён пример оценки доверия векторам движения.

На рис. 2 (см. с. 280) приведена блок-схема алгоритма определения компоненты движения, связанной со смещением камеры. Алгоритм состоит из следующих шагов.

1. Выполнить фильтрацию поля векторов по значению функции доверия.

2. Повторить N раз:
 - а) выбрать случайным образом три вектора из поля векторов движения;
 - б) вычислить параметры аффинного преобразования по тройке векторов;
 - в) обновить гистограмму по каждому параметру аффинной модели глобального движения.
3. При помощи алгоритма водораздела выделить пики гистограмм, лежащие выше порога T_1 .
4. Из значений параметров, соответствующих пикам, построить набор параметров-кандидатов.
5. Выбрать параметры, на которых достигается минимум межкадровой разницы.

С помощью найденных параметров глобального движения вычисляются векторы глобального движения для блоков по формуле:

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ f \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix},$$

где (x, y) — координаты центра блока; (a, b, c, d, e, f) — параметры глобального движения; (v_x, v_y) — вектор глобального движения для блока.

По векторам вычисляется ошибка приближения блоков. Блоки, для которых ошибка приближения превышает порог T_2 , объединяются в связные области. Область помечается как объект, если она состоит более чем из четырёх блоков размером 16×16 пикселей.

Быстродействие алгоритма составляет 400 кадров (размером 640×480 пикселей) в секунду на PC Pentium-4 (2,8 ГГц), что позволяет в реальном времени обрабатывать информацию от 80 камер на одной машине.

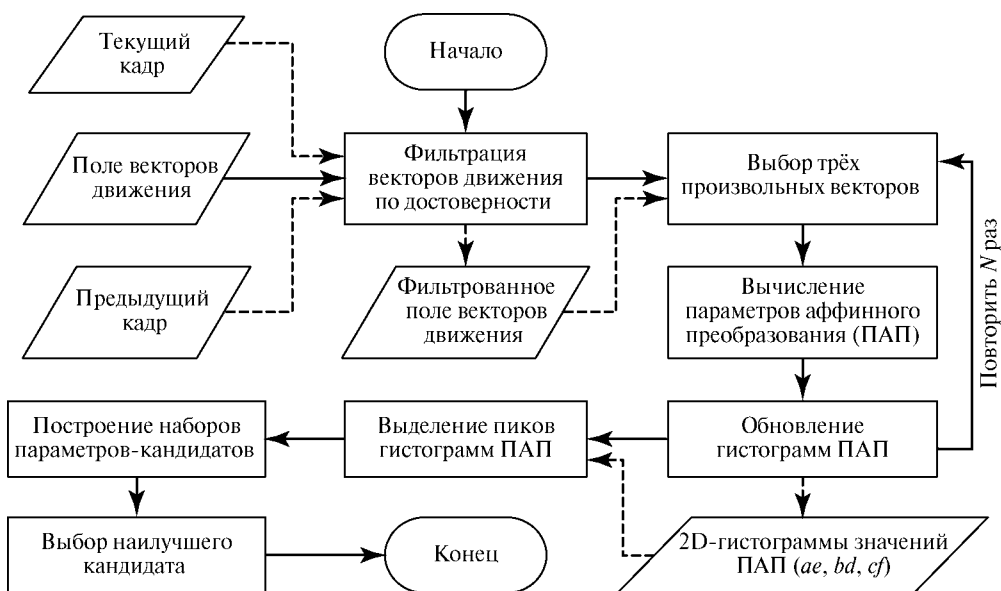


Рис. 2. Блок-схема алгоритма оценки глобального движения

Литература

- Гришин С. В., Ватолин Д. С., Стрельников К. Н. и др. (2008) Обзор блочных методов оценки движения в цифровых видеосигналах // Программные системы и инструменты: Тематич. сб. № 9. М.: Изд-во ф-та ВМиК МГУ, 2008. С. 50–62.
- Пешков Н. Н. (2004) Математическое моделирование и разработка алгоритмов обнаружения и измерения параметров сторонних объектов в системах наблюдения. Ставрополь: Ставропольский гос. ун-т, 2004. 146 с. РГБ ОД, 61:05-1/59.
- Biswas A., Guha P., Mukerjee A., Venkatesh K. S. (2006) Intrusion Detection and Tracking with Pan-Tilt Cameras // Visual Information Engineering. 2006. P. 565–571.
- Borshukov G. D., Bozdagi G., Altunbasak Y., Tekalp A. M. (1997) Motion Segmentation by Multistage Classification // IEEE Trans. Image Processing. 1997. V. 6. N. 11. P. 1591–1594.
- Goldberger J., Greenspan H. (2006) Context-based Segmentation of Image Sequences // IEEE Trans. Pattern Analysis and Machine Intelligence. 2006. V. 28. N. 3. P. 463–468.
- Latecki L. J., Miezianko R., Pokrajac D. (2004) Evaluating Reliability of Motion Features in Surveillance Videos // NIST Workshop on Performance Metrics for Intelligent Systems. Gaithersburg, MD, USA. 2004.
- Staufer C., Grimson W. E. L. (2000) Learning Patterns of Activity Using Realtime Tracking // IEEE Trans. Pattern Analysis and Machine Intelligence. 2000. V. 22. N. 8. P. 747–757.

USING THE FEATURES OF THE MODERN PRESENTATION OF DIGITAL VIDEO PROCESSING ALGORITHMS FOR OPTIMIZATION

К.Н. Стрельников

*Lomonosov Moscow State University,
119992, GSP-1, Leninskie Gory, Moscow, 119991*

In this article we propose a fast algorithm for detecting objects using pan-tilt-zoom IP-cameras. High speed is achieved thanks to a new approach to optimize the processing of encoded video using internal codec information.

Keywords: optimization, video codecs, IP-camera, object detection.

Strelnikov Konstantin Nikolaevich — Candidate of Physical and Mathematical Sciences, e-mail: kstrelnikov@gaphics.cs.msu.ru.

МОРФОЛОГИЧЕСКИЕ МЕТОДЫ АНАЛИЗА СЦЕН ПО ИХ ИЗОБРАЖЕНИЯМ Ю. П. Пытьев, А. И. Чуличков

*Московский государственный университет им. М. В. Ломоносова (МГУ),
119991, Москва, ГСП-1, Ленинские горы*

Морфологические методы анализа изображений предназначены для решения задач узнавания и классификации объектов, выделения отличий в сценах, оценки параметров сцен по их изображениям, полученным при неконтролируемых и неизвестных условиях. В статье вводится понятие сравнения изображения по форме, дается определение формы изображения как инварианта преобразований изображения, отвечающих изменениям условий их регистрации. Приводятся методы решения ряда задач.

Работа выполнена при финансовой поддержке РФФИ (проект № 08-07-00120).

Ключевые слова: анализ изображений, задачи наилучшего приближения, форма изображения, классификация, выделение отличий, наилучшая оценка параметров.

Введение

Изображения одной и той же сцены могут заметно различаться между собой при вариации условий их регистрации — таких как освещение, оптические свойства поверхности объектов, характеристики среды, влияющей на изображения объекта, и пр. Этот факт усложняет задачу анализа сцен по их изображениям, поскольку связь между расположением объектов и распределением яркости на поле зрения неоднозначна. Тем не менее, если всевозможные изображения сцены могут быть описаны определённым классом преобразований, выполняемых над некоторым изображением этой сцены, то характеристикой формы объектов на изображении естественно считать максимальный инвариант данного класса преобразований. Поскольку этот инвариант, как правило, не позволяет восстановить *форму объектов*, он назван *формой изображения*, а методы анализа изображений, основанные на этой идее, — *морфологическими* [Пытьев, 1984; Пытьев, Чуличков, 2010].

Рассматриваемые в статье морфологические методы ориентированы на решение задач, которые могут быть сгруппированы в два больших класса. Первый класс включает задачи выделения неизвестных объектов на изображениях одной сцены, полученных при неизвестных условиях. Пусть, например, имеется два изображения местности и объектов на ней, полученное зимой и летом (рис. 1, см. с. 283). Пусть при предьявлении этих изображений для анализа необходимо выделить объект, отсутствующий на одном из них, но различимый на другом. Трудность, с которой традиционно сталкиваются при решении этой задачи, состоит в том, что изменение условий регистрации приводит к изменениям в изображении, часто более существенным, чем появление/исчезновение объектов на местности.

Пытьев Юрий Петрович — заведующий кафедрой компьютерных методов физики, доктор физико-математических наук, профессор.

Чуличков Алексей Иванович — профессор кафедры компьютерных методов физики, доктор физико-математических наук, профессор, e-mail: achulichkov@gmail.com.



Рис. 1. Изображения одной и той же местности, полученные в разные сезоны года



Рис. 2. Результат выделения отличий в сценах. Слева — разность $f - g$ изображений, приведенных на рис. 1, справа — отличие изображения летнего пейзажа g от зимнего f по форме. Форму изображения f определяют конфигурации областей его одинаковой яркости. Отличие по форме получается вычитанием средней яркости $P_f g$ изображения g на этих областях из яркости изображения g в каждой точке поля зрения. Отличные от нуля разности $g - P_f g$ возникают вследствие изменения конфигураций областей поля зрения равной яркости изображения g по сравнению с конфигурациями соответствующих областей изображения f , которые вызваны изменением сцены

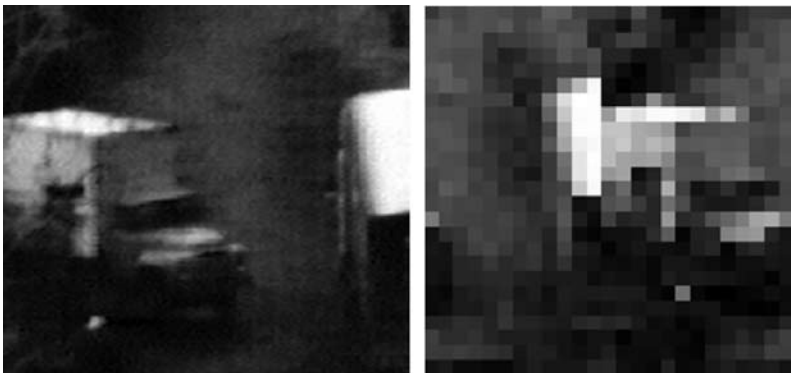


Рис. 3. Изображение сцены (слева) и её фрагмента, полученного видеосистемой с существенно меньшим пространственным разрешением (справа)

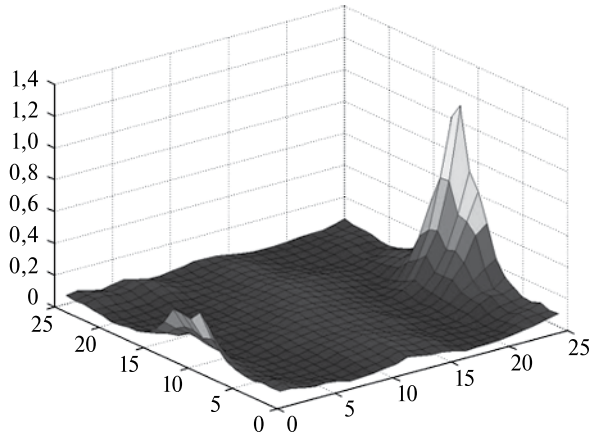


Рис. 4. Зависимость близости по форме сравниваемых участков изображений, представленных на рис. 3, от взаимного сдвига изображений. Близость по форме $1/t_{H,\gamma}(\xi)$, измеряемая в условных единицах, откладывается по вертикальной оси, а величина взаимного сдвига изображений в пикселах — по осям горизонтальной плоскости. Точка максимума соответствует наибольшему сходству изображений по форме и служит оценкой координат искомого фрагмента

Морфологические методы позволяют успешно справиться с этой задачей. На рис. 2 (см. с. 283) слева приведено изображение, яркость которого равна разности яркостей изображений сцены, полученных в разные сезоны года. Видно, что различие в условиях регистрации приводят к существенной разнице в яркостях. На этом же рисунке справа приведена *морфологическая разность* этих же изображений, позволившая выделить объект (яркое пятно в правом нижнем углу), появившийся на сцене и не связанный с изменениями условий наблюдения.

Второй класс задач связан с поиском известных объектов на неизвестной сцене. Например, пусть на изображении сцены, полученном в видимом диапазоне электромагнитного излучения (рис. 3, слева, см. с. 283), требуется найти фрагмент (автомобиль), видимый на изображении ИК-диапазона (рис. 3, справа).

Яркости участков поля зрения, соответствующие изображению автомобиля, существенно различаются, однако их структура («форма») сохраняется. Это позволило определить координаты искомого фрагмента ИК-изображения на видимом изображении. На рис. 4 приведен график зависимости близости по форме сравниваемых участков изображений от взаимного сдвига изображений (в условных единицах). Точка максимума соответствует наибольшему сходству изображений по форме и служит оценкой координат искомого фрагмента.

1. ФОРМА ИЗОБРАЖЕНИЯ КАК ИНВАРИАНТ ПРЕОБРАЗОВАНИЙ ИЗОБРАЖЕНИЙ, ОТВЕЧАЮЩИХ ВАРИАЦИЯМ УСЛОВИЙ РЕГИСТРАЦИИ

Поясним основную идею методов морфологического анализа изображений на следующем примере. Рассмотрим изображения кубика, сформированные при различных условиях наблюдения (рис. 5, см. с. 285). Они несут достаточно подробную информацию о геометрической форме объекта сцены (о кубике), так как области изображения, все точки которых имеют примерно одинаковые яркости, отобра-

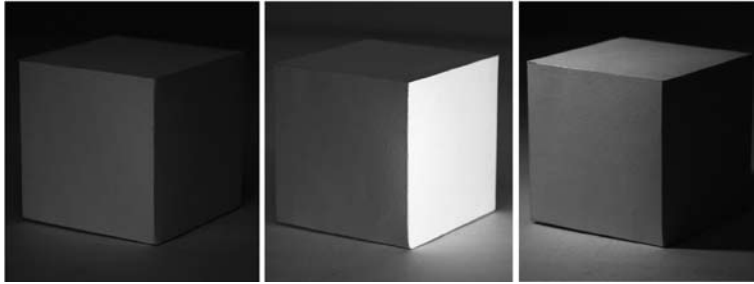


Рис. 5. Изображения кубика при различных условиях регистрации (освещения)

жают грани кубика, обладающие одинаковыми геометрическими и оптическими свойствами. Изменение условий освещения приводит к изменениям яркости точек поля зрения, однако области, соответствующие граням кубика и фону, останутся неизменными.

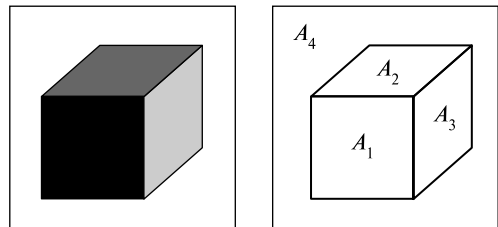
В этой ситуации математическую модель изображения кубика можно определить как кусочно-постоянную функцию $f(\cdot)$, заданную на поле зрения в виде прямоугольной области X на плоскости R^2 ; значение $f(x)$ функции в точке x поля зрения X интерпретируется как яркость изображения в точке $x \in X$, а области постоянной яркости $A_1, \dots, A_4 \subset X$ соответствуют фрагментам, изображающим три видимые грани кубика и фон (рис. 6). Вариация условий освещения приводит к изменению яркости исходного изображения точек, но не к произвольному, а такому, при котором области поля зрения постоянной яркости сохраняются либо объединяются, но не могут распадаться на более мелкие.

Таким образом, математическую модель множества изображений данной сцены, полученных при всевозможных условиях (однородного) освещения, можно определить как множество функций

$$V_f = \left\{ g(x) = \sum_{j=1}^4 c_j \chi_j(x), x \in X; c_j \in (-\infty, \infty), j = 1, \dots, 4 \right\},$$

где $\chi_j(\cdot)$ — индикаторная функция множества A_j , т. е. функция, равная единице, когда её аргумент принадлежит области A_j , и нулю — в противном случае. Изменение условий регистрации приводит к изменению яркостей c_1, \dots, c_4 . Разбиение поля зрения X на области A_1, \dots, A_4 одинаковой яркости является максимальным инвариантом этого класса преобразований яркости. Это разбиение определяет всё то, что относится к данной сцене, и не зависит от условий формирования изображений. Его можно назвать формой изображения сцены. Изменения в геометрической форме объектов сцены (в данном случае кубика и фона), в их размерах, в расположении на поле зрения, в ракурсе или появление новых объектов сцены приведут к тому, что области

Рис. 6. Математическая модель изображения кубика (слева) и разбиение поля зрения, определяющее его форму (справа)



равной яркости на изображении этой измененной сцены изменят свои конфигурации по сравнению с областями A_1, \dots, A_4 — форма изображения изменится.

Для того чтобы узнать сцену на предъявленном изображении $\xi(\cdot)$, следует ответить на вопрос, принадлежит ли изображение $\xi(\cdot)$ множеству V_f . С формальной точки зрения для этого следует вычислить расстояние от $\xi(\cdot)$ до V_f , т. е. решить задачу наилучшего приближения изображения $\xi(\cdot)$ изображениями из V_f :

$$\rho(V_f, \xi) = \inf \left\{ \left\| \xi - \sum_{j=1}^4 c_j \chi_j \right\|, c_j \in (-\infty, \infty), j = 1, \dots, 4 \right\};$$

для того, чтобы $\xi \in V_f$, необходимо и достаточно, чтобы $\rho(V_f, \xi) = 0$. Если изображения рассматривать как элементы евклидова пространства, то речь идет о расстоянии в пространстве R всех изображений от точки (вектора) ξ до четырехмерного линейного подпространства $V_f \subset R$, являющегося линейной оболочкой векторов χ_1, \dots, χ_4 . Как известно, это расстояние измеряется расстоянием между ξ и его ортогональной проекцией на V_f . Ортогональную проекцию ξ на V_f обозначим $P_f \xi$. Изображение $\xi(\cdot)$ будет изображением кубика (т. е. выполнено включение $\xi \in V_f$) тогда и только тогда, когда $P_f \xi = \xi$. Ортогональная проекция $P_f \xi$ изображения ξ на

$$V_f \text{ легко вычисляется и равна } P_f g = \sum_{i=1}^4 \frac{(g, \chi_i)}{\|\chi_i\|^2} \chi_i.$$

Итак, разбиение A_1, \dots, A_4 поля зрения X на непересекающиеся области A_1, \dots, A_4 взаимно однозначно связано с множеством всех изображений кубика V_f и с оператором ортогонального проецирования на V_f . В морфологических методах анализа изображений формой называют множество V_f или оператор проецирования на V_f .

2. СРАВНЕНИЕ ИЗОБРАЖЕНИЙ ПО ФОРМЕ

Следует отметить, что морфологические методы применимы не только к кусочно-постоянным изображениям. Действительно, для того чтобы задать форму изображения сцены как множество V_f всех возможных его изображений, следует задать какое-либо (достаточно подробное) изображение f сцены и определить, как оно преобразуется при изменении условий регистрации. Для этого следует указать класс \mathbf{F} преобразований, таких, что для любого $F \in \mathbf{F}$ результат $F * f$ преобразования изображения f тоже является изображением той же сцены. Тогда форму V_f можно определить соотношением

$$V_f = \{g = F * f, F \in \mathbf{F}\}.$$

Для рассмотренного выше примера формы изображения кубика множество V_f можно задать, указав в качестве изображения f сцены любое изображение

$$f(x) = \sum_{j=1}^4 c_j \chi_j(x), x \in X, \text{ где все яркости } c_1, \dots, c_4 \text{ попарно различны, а в качестве класса его преобразований — преобразования яркости } F * f = \sum_{j=1}^4 F(c_j) \chi_j(x), \text{ где}$$

$F \in \mathbf{F}$, а \mathbf{F} — класс всех (борелевских) функций, определённых на числовой оси и принимающих конечные числовые значения. Поскольку в результате таких преобразований может случиться так, что, несмотря на различие яркостей c_1, \dots, c_4 ,

некоторые из яркостей $F(c_1), \dots, F(c_4)$ могут совпасть, области одинаковой яркости изображения $F * f = \sum_{j=1}^4 F(c_j) \chi_j(x)$ будут объединением областей из разбиения A_1, \dots, A_4 , и форму такого изображения естественно рассматривать как более простую, чем форма исходного изображения f .

В общем случае будем считать, что g по форме не сложнее, чем f , если найдется такая функция $F \in \mathbf{F}$, что $g = F * f$. Про такие изображения говорят, что они сравнимы по форме с f . Форма изображения f , таким образом, состоит из изображений, сравнимых с f по форме.

Мерой близости формы изображения ξ к V_f в морфологическом анализе счита-

ется отношение $t(\xi) = \frac{\|\xi - P_f \xi\|^2}{\|\Pi \xi - P_f \xi\|^2}$, где $\Pi \xi$ — ортогональная проекция ξ на мно-

жество изображений, равных константе на X , т. е. $\Pi \xi$ — изображение, яркость которого в каждой точке поля зрения равна средней (по всему полю зрения X) яркости ξ . Дробь $t(\xi)$ тем меньше, чем меньше расстояние от ξ до V_f и чем больше отличие $P_f \xi$ от константы.

Заметим, что отношению $t(\xi)$ можно придать смысл отношения «шум / сигнал» в предположении, что изображение ξ принадлежит форме V_f . Действительно, числитель этой дроби характеризует отличие изображения ξ от изображений из V_f , которое при $\xi \in V_f$ можно объяснить только наличием погрешности, а знаменатель — величину той составляющей изображения ξ , которая сравнима по форме с изображением f и отлична от константы. В задачах узнавания сцены по изображению, искаженному шумовой погрешностью, следует указать значение порога от-

ношения шум/сигнал либо по заданному значению отношения $t(\xi) = \frac{\|\xi - P_f \xi\|^2}{\|\Pi \xi - P_f \xi\|^2}$

принять решение, можно ли считать полученное значение отношения «шум / сигнал» правдоподобным для данного изображения.

3. ВЫДЕЛЕНИЕ ОТЛИЧИЙ ИЗОБРАЖЕНИЙ ПО ФОРМЕ

Пусть форма изображения f задана в виде оператора ортогонального проецирования P_f на V_f и предьявляется некоторое изображение ξ . Требуется указать, в чем состоит отличие по форме изображения ξ от f . Так как изображение $P_f \xi$ является наилучшим приближением ξ изображениями, сравнимыми по форме с f , то изображение $\xi - P_f \xi$ представляет всё то, что отличает ξ от f по форме.

Иллюстрация эффективности такого способа выделения отличий по форме представлена на рис. 2. На рис. 7 (см. с. 288) приведен ещё один пример выделения отличий по форме. Предьявлены изображения двух сцен, приведенные на рис. 7 вверху, полученные при различных условиях регистрации, требуется найти различие в сценах. Изображение $\xi - f$ слева внизу показывает отличие в яркости предьявленных изображений и не дает возможности заметить отличие в сценах, морфологическая разность $\xi - P_f \xi$, приведенная на рис. 7 справа внизу, выделяет область поля зрения, на которой изображен объект, отсутствующий на исходной сцене.



Рис. 7. Выделение отличий по форме. Слева сверху — изображение f , задающее форму изображения сцены. Справа сверху — изображение ξ сцены, отличающейся от предыдущей наличием детали — бусинки, лежащей у ног статуи собачки. Слева внизу — изображение, яркость которого равна разности яркостей изображений f и ξ , по этому изображению заметить отличие в сценах невозможно. Справа внизу — разность $\xi - P\xi$, хорошо видна область поля зрения, на которой показана бусинка, изменившая форму изображения сцены.

4. ОБНАРУЖЕНИЕ ОБЪЕКТА ПО ЕГО ИЗОБРАЖЕНИЮ И ОЦЕНКА ЕГО КООРДИНАТ

На практике большой интерес представляют задачи совмещения и выделения фрагментов заданной формы на предъявленном изображении. Пример такой задачи приведен во введении. Как известно, если отличие яркостей сравниваемых по форме фрагментов не сводится к однородному изменению их яркости и контраста, то в таких задачах широко распространенные корреляционные методы работают недостаточно эффективно.

Пусть дано изображение f , причём на подмножестве H поля зрения X имеется фрагмент $f\chi_H$, форма которого рассматривается как эталон формы. Напомним, что функция χ_H равна нулю вне множества H и единице — на множестве H ; таким образом, яркость изображения $f\chi_H$ равна яркости изображения f на множестве H и равна нулю вне H . Рассмотрим группу G преобразований плоскости R^2 , и G_H — множество преобразований из G , таких, что γH полностью содержится в области X , если $\gamma \in G_H$. Следует найти такое преобразование $\gamma \in G_H$, чтобы фрагмент изображения ξ на подмножестве γH был близок по форме к фрагменту изображения f на подмножестве H . Опишем эту задачу как рассмотренную выше для сравнения формы двух изображений, заданных на поле зрения X .

Определим меру близости указанных фрагментов по форме. Назовем формой фрагмента f изображения на подмножестве H множество V_{f_H} изображений, форма которых не сложнее, чем форма любого изображения вида $\tilde{f}(\cdot) = f\chi_H(\cdot) + \varphi(\cdot)(1 - \chi_H(\cdot))$, где $\varphi(\cdot)$ — произвольное изображение. Иными словами, форма изображения на H определяется как множество изображений, яркость которых на H может быть получена преобразованиями изображения $f_H(\cdot)$ вида $F * f_H(\cdot)$, где F пробегает по всему множеству \mathbf{F} , а на дополнении к H (на фоне) яркость изображения произвольна. Проектор на это множество обозначим $P_{f,H}$.

Фрагмент заданной формы на предъявленном для анализа изображении ξ будет обнаружен, если $P_{f,H}\xi = \xi$, где $P_{f,H}$ — проектор на множество V_{f_H} . Для определения близости фрагментов по форме обозначим через $P_{0,H}$ — проектор на множество изображений, форма которых не сложнее, чем форма любого из изображений вида $\tilde{f} = f(\cdot)\chi_H(\cdot) + \varphi(\cdot)(1 - \chi_H(\cdot))$, где $\varphi(\cdot) \in L_2(X)$ — произвольное изображение. Иными словами, $P_{0,H}$ — проектор на множество изображений, яркость которых равна константе на подмножестве H и произвольна вне него. Тогда близость фрагмента предъявленного изображения ξ к форме фрагмента изображения f на под-

множестве H определим значением дроби $t_H(\xi) = \frac{\|\xi - P_{f,H}\xi\|^2}{\|P_{0,H}\xi - P_{f,H}\xi\|^2}$.

Если на изображении ξ фрагмент заданной формы расположен на подмножестве γH , то сравнивать по форме на подмножестве H следует фрагменты изображений $f_\gamma(\cdot)$ и $\xi(\cdot)$, где

$$f_\gamma(x) = \begin{cases} f(\gamma^{-1}x), & \text{если } \gamma^{-1}x \in X, \\ 0, & \text{если } \gamma^{-1}x \notin X. \end{cases}$$

Близость фрагмента изображения ξ на множестве γH к форме фрагмента изображения f на множестве H определим значением дроби $t_{H,\gamma}(\xi) = \frac{\|\xi - P_{f_\gamma,H}\xi\|^2}{\|P_{0,H}\xi - P_{f_\gamma,H}\xi\|^2}$.

Пусть G — группа сдвигов плоскости R^2 , так что при задании декартовых прямоугольных координат любой вектор $(a, b) \in R^2$ преобразование $\gamma \in G$ превращает в вектор $\gamma_{x,y}(a, b) + (a + x, b + y) \in R^2$. На рис. 4 приведены значения $1/t_{H,\gamma(\xi)}$, обратные близости фрагментов по форме, в зависимости от координат вектора сдвига. Максимальное значение величины $1/t_{H,\gamma(\xi)}$ определяет координаты вектора сдвига плоскости, совмещающие близкие по форме фрагменты изображений ξ и f .

Заключение

В работе определена операция сравнения изображений по форме, предназначенная для решения задач анализа сцен по их изображениям методами, инвариантными к преобразованиям условий регистрации изображений сцен, продемонстрирована их эффективность.

Литература

- Пытьев Ю. П.* (1984) Задачи морфологического анализа изображений // Мат. методы исслед. природных ресурсов Земли из космоса. М.: Наука, 1984. С. 41–83.
- Пытьев Ю. П., Чуличков А. И.* (2010) Морфологические методы анализа изображений. М.: Физматлит, 2010. 336 с.

MORPHOLOGICAL METHODS FOR THE ANALYSIS OF SCENES UNDER THEIR IMAGES

Yu. P. Pyt'ev, A. I. Chulichkov

*Lomonosov Moscow State University,
119992, GSP-1, Leninskie Gory, Moscow, 119991*

Morphological methods of the analysis of images are developed for the decision of problems of a recognition and classification of objects, revealing of differences in scenes, an estimation of parameters of scenes under their images received under uncontrollable and unknown conditions. The concept of comparison of the image under the shape is entered. The shape of an image is defined as the invariant of the transformations of an image which models its change corresponding to variations of conditions of registration. Methods for the decision of the problems of a recognition and classification of image shapes, revealing of differences in shapes, an estimation of parameters of shapes are described.

Keywords: image analysis, optimal decision making, image shape, classification, revealing of differences in scenes, best estimation.

Pyt'ev Yuri Petrovich — Head of the Chair of Computer Methods in Physics, Doctor of Physical and Mathematical Sciences, Professor.

Chulichkov Alexey Ivanovich — Professor of Chair of Computer Methods in Physics, Doctor of Physical and Mathematical Sciences, Professor, e-mail: achulichkov@gmail.com.

СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ОПИСАНИЙ КЛАСТЕРОВ СТРУКТУРИРОВАННЫМИ ГАУССОВЫМИ СМЕСЯМИ

М. М. Ланге¹, Н. А. Новиков²

¹ Учреждение Российской академии наук Вычислительный центр им. А. А. Дородницына РАН (ВЦ РАН), 119333, Москва, Вавилова, 40

² Государственное образовательное учреждение высшего профессионального образования «Московский государственный институт радиотехники, электроники и автоматики (технический университет)» (МИРЭА), 119454, Москва, Вернадского просп., 78

Предложен подход к решению задачи сегментации изображений с использованием древовидно-структурированных описаний кластеров гауссовыми смесями в многомерном пространстве признаков. Рассмотрены модификации таких описаний, строящиеся на основе оптимизации деревьев с помощью процедуры скользящего контроля. С использованием древовидных представлений кластеров сформулирован алгоритм сегментации в форме байесовской процедуры классификации по критерию максимума апостериорной вероятности. В экспериментах с аэрокосмическими изображениями разработанный алгоритм сегментации продемонстрировал существенно меньшее число отказов по сравнению с традиционным алгоритмом, использующим гистограмму. Сравнительные оценки затрат памяти демонстрируют выигрыш предложенного алгоритма относительно традиционного в 50 и более раз, в зависимости от модификаций применяемых древовидных структур.

Работа выполнена при частичной поддержке РФФИ, проект № 09-01-00573.

Ключевые слова: сегментация изображений, кластер, гауссова смесь, гистограмма, функция правдоподобия, апостериорная вероятность, байесовский классификатор, древовидное описание.

Введение

Сегментация является одной из важных задач анализа изображений, которая предполагает разбиение изображений на семантически однородные области, и может быть решена на основе методов распознавания образов [Дуда, Харт, 1976]. Процедуры сегментации используются для автоматического выделения информативных объектов на изображениях (лиц, жестов, графических символов, движущихся объектов и др.). Различные варианты процедур сегментации находят применение в робототехнике при проектировании систем технического зрения, в материаловедении для обнаружения дефектов деталей, в медицине для выявления пораженных тканей, в картографии и экологии для построения систем автоматической обработки аэрокосмических изображений и в других приложениях, имеющих практическое значение.

В известных подходах решение задачи сегментации основано либо на поиске однородных областей, либо на определении границ между областями [Гонсалес, Вудс, 2006; Malik et al., 2001]. Известные алгоритмы сегментации условно можно

Ланге Михаил Михайлович — ведущий научный сотрудник, кандидат технических наук, e-mail: lange_mm@ccas.ru.

Новиков Никита Александрович — аспирант, e-mail: nikknovikov@gmail.com.

разбить на две группы, которые объединяют последовательные и параллельные алгоритмы. Последовательные алгоритмы обрабатывают множество локальных областей изображения, например, методом выращивания [Mancas et al., 2005], а параллельные — рассматривают изображения как поля, описываемые заданными моделями, например, моделями марковского типа [Szirányi et al., 2000].

В настоящей работе исследуется подход к сегментации на основе выделения однородных областей путём последовательного отнесения элементов изображения к одному из классов, которые представлены древовидными описаниями в заданном многомерном пространстве признаков. В этом случае задача обучения классификатора сводится к обучению по прецедентам, которыми являются векторы признаков элементов изображения, характеризующие локальные свойства текстуры. Оно ориентировано на построение структурированных описаний функций правдоподобия, заданных кластерами в многомерном признаковом пространстве, и базируется на модификации подхода, развитого в работе [Ganebnykh, Lange, 2009] для построения древовидных представлений двумерных объектов. Это продиктовано стремлением ускорить процедуру обучения и сократить объём памяти, необходимый для хранения описаний функций правдоподобия. Экспериментальная часть работы выполнена для аэрокосмических снимков, в которых образцы изображений различных типов поверхности заранее известны. Поэтому для указанного источника изображений использована схема обучения с учителем [Дуда, Харп, 1976]. Приводятся результаты сегментации аэрокосмических изображений в трёхмерном пространстве признаков Харалика [Miyamoto, Merryman, 2005], полученные при различных модификациях древовидных описаний функций правдоподобия и при традиционном описании с помощью гистограммы. Получены оценки объёмов памяти, которые требуются для реализации алгоритмов, использующих как оптимизированные древовидные описания кластеров, так и оптимизированную по числу уровней гистограмму.

1. СХЕМА КЛАССИФИКАЦИИ И ЗАДАЧА ИССЛЕДОВАНИЯ

Процедура классификации включает два этапа: обучение и принятие решения. На этапе обучения объекты обучающей выборки с известными метками классов используются для построения описаний (представлений) классов и для оценивания параметров алгоритма принятия решения по заданному критерию. На этапе принятия решений каждый объект из тестовой выборки предъявляется алгоритму классификации, и полученная в результате его работы метка класса присваивается данному объекту.

В решаемой задаче объектами являются элементы изображения (пиксели) с заданными значениями яркости. Признаковым описанием пиксела служит набор значений различных функций яркости из заданной окрестности выбранного пиксела. Эти функции выражают различные локальные свойства текстуры.

Большинство известных классификаторов можно условно разделить на два типа: метрические и вероятностные. Метрические классификаторы основаны на вычислении расстояний между тестовым и эталонными объектами. Простейшим примером является метод «ближайшего соседа», в соответствии с которым тестовый объект присваивается тому же классу, что и ближайший к нему эталонный объект. Метрические классификаторы плохо применимы при обучающих выборках значительного объёма, так как перебор на этапе принятия решений требует в

этом случае слишком больших вычислительных затрат. Вероятностные классификаторы предполагают, что появление конкретного объекта является случайным событием, вероятность которого оценивается на основе обучающей выборки. Такие классификаторы предпочтительно использовать в условиях обучающих выборок большого объёма, поскольку при этом оценка неизвестных распределений оказывается более точной. Характерный объём используемых нами выборок составляет порядка 10^5 объектов, поэтому для исследования выбраны классификаторы вероятностного типа.

При использовании вероятностного (байесовского) подхода к классификации вероятность принадлежности объекта к классу определяется апостериорной вероятностью

$$P(\omega_i | \mathbf{x}) = \frac{P(\omega_i) P(\mathbf{x} | \omega_i)}{\sum_k P(\omega_k) P(\mathbf{x} | \omega_k)}.$$

Здесь \mathbf{x} — вектор признаков объекта; ω_i — метка i -го класса; $P(\omega_i)$ — априорная вероятность i -го класса; $P(\mathbf{x} | \omega_i)$ — значение функции правдоподобия для i -го класса (условного по i -му классу распределения) в точке, задаваемой вектором \mathbf{x} . Решение принимается в пользу класса с наибольшей апостериорной вероятностью. Обучение байесовских классификаторов сопряжено с построением оценок функций правдоподобия для классов. Большинство подходов опираются на принцип максимизации правдоподобия. Отметим три распространенных подхода к оцениванию функций правдоподобия.

Первый подход заключается в квантовании признакового пространства и построении гистограмм распределения объектов обучающей выборки для каждого класса. Преимущества подхода — сравнительно небольшие вычислительные затраты на этапах обучения и принятия решений и регулируемая точность аппроксимации условных распределений; недостатками являются экспоненциальный рост требуемого объёма памяти в зависимости от количества признаков, а также необходимость задания обучающей выборки, имеющей объём порядка числа квантов.

Другой подход основан на построении оценок функций правдоподобия в виде нормальных распределений, параметры которых совпадают с соответствующими выборочными величинами. Преимущества подхода — нетребовательность к вычислительным затратам и квадратичная зависимость объёма памяти от размерности признакового пространства; недостатком является достаточно грубая аппроксимация условных распределений.

Третий подход связан с построением оценок функций правдоподобия в виде смесей нормальных распределений. Число компонент смеси выбирается заранее, а их веса и параметры настраиваются при помощи EM-алгоритма, реализующего принцип максимизации правдоподобия [Дуда, Харт, 1976]. Преимущество подхода заключается в том, что он сочетает квадратичную зависимость объёма памяти от числа признаков с регулируемой точностью аппроксимации. Недостатками являются значительный объём вычислений на этапе обучения и трудности при выборе начальных значений параметров.

Метод, развиваемый в настоящей работе, сочетает особенности трёх вышеупомянутых подходов. Этот метод базируется на построении оценок функций правдоподобия в виде гауссовых смесей, но процедура настройки параметров смеси

существенно отличается от EM-алгоритма. Множество объектов обучающей выборки для каждого класса разбивается на непересекающиеся фрагменты, и каждому фрагменту ставится в соответствие отдельная компонента смеси. Параметры каждой компоненты вычисляются только один раз, поэтому оценки строятся значительно быстрее, чем при использовании EM-алгоритма. Такой выигрыш в быстродействии достигается с потерей оптимальности оценок в смысле принципа максимизации правдоподобия.

В рамках сформулированного подхода строятся древовидно структурированные аппроксимации функций правдоподобия по заданным кластерам в выбранном пространстве признаков. Выполнена оптимизация указанных аппроксимаций по параметрам древовидной структуры. Приведены примеры сегментации реальных изображений с использованием оптимизированных древовидных описаний функций правдоподобия и гистограммы, оптимизированной по числу уровней квантования. Получены сравнительные оценки объемов памяти, требуемой для реализации алгоритмов сегментации, использующих древовидные аппроксимации гауссовыми смесями и гистограмму.

2. СТРУКТУРНОЕ ПРЕДСТАВЛЕНИЕ КЛАСТЕРОВ В ПРОСТРАНСТВЕ ПРИЗНАКОВ

Дихотомическое разбиение кластеров на фрагменты

Объекты каждого класса, входящие в обучающую выборку, образуют в пространстве признаков кластеры. Оценивание функций правдоподобия эквивалентно формированию представлений кластеров. Предлагаемый подход основан на дихотомическом разбиении кластеров на фрагменты, каждый из которых описывается нормальным распределением, взятым с некоторым весом. Используемая процедура разбиения является многомерной модификацией алгоритма, предложенного в работе [Ganebnykh, Lange, 2009] для построения древовидных представлений образов на двумерных изображениях. Представление кластера формируется на множестве описаний его фрагментов и дает аппроксимацию функции правдоподобия в виде взвешенной суммы нормальных распределений.

Рассмотрим процедуру разбиения кластера, соответствующего некоторому классу ω_j . На первом шаге строится нормальное распределение, описывающее кластер наилучшим образом. Его вектор математического ожидания и ковариационная матрица равны соответствующим выборочным величинам, вычисленным для множества элементов кластера. Если полученное распределение не является удовлетворительным представлением кластера, то он разбивается на два непересекающихся фрагмента, и для каждого из них указанная процедура, по необходимости, повторяется.

Для каждого полученного в ходе разбиения фрагмента $B_{i,j}$ (i — номер кластера; j — номер фрагмента внутри кластера) определим его описание в виде:

$$\tilde{B}_{i,j} = (\alpha_{i,j}, N_{i,j}),$$

где $\alpha_{i,j}$ — вес фрагмента; $N_{i,j} = N(\mu_{i,j}, \Sigma_{i,j})$ — плотность соответствующего фрагменту нормального распределения с математическим ожиданием $\mu_{i,j}$ и ковариационной матрицей $\Sigma_{i,j}$. Вес фрагмента равен относительному числу элементов кластера, принадлежащих фрагменту:

$$\alpha_{i,j} = \frac{|\{x \in B_{i,j}\}|}{|\{x \in B_{i,1}\}|},$$

где исходный кластер соответствует фрагменту с номером $j = 1$. Нормальное распределение для фрагмента строится так же, как для исходного кластера: его параметры равны выборочным значениям, полученным на множестве принадлежащих фрагменту элементов:

$$\mu_{i,j} = E_{x \in B_{i,j}} [x], \quad \Sigma_{i,j} = E_{x \in B_{i,j}} \left[(x - \mu_{i,j})(x - \mu_{i,j})^T \right].$$

На каждом шаге перебираются фрагменты, для которых разбиение ещё не произошло. Относительно каждого фрагмента принимается решение: «разбивать» или «не разбивать». Если ни один фрагмент не подлежит разбиению, работа алгоритма заканчивается. Если для некоторого фрагмента $B_{i,j}$ требуется разбиение, то он разбивается на два непересекающихся фрагмента:

$$B_{i,j} \rightarrow \{B_{i,2j+1}, B_{i,2j+2}\}, \quad B_{i,2j+1} \cup B_{i,2j+2} = B_{i,j}, \quad B_{i,2j+1} \cap B_{i,2j+2} = \emptyset.$$

Для новых фрагментов строятся их описания.

В текущей реализации фрагмент разбивается гиперплоскостью, которая проходит через его центр перпендикулярно наибольшей оси эллипсоида рассеяния, порожденного соответствующим нормальным распределением. Такой способ разбиения продемонстрирован на рис. 1.

Из рекурсивного характера алгоритма разбиения следует, что полученное множество фрагментов имеет вид бинарного дерева, в корне которого расположен исходный кластер $B_{i,1}$. Каждый неконцевой узел $B_{i,j}$ имеет двух потомков: $B_{i,2j+1}$ и $B_{i,2j+2}$. В концевых узлах располагаются фрагменты, которые не подвергались разбиению.

Аппроксимация функций правдоподобия гауссовыми смесями

Сформированные деревья фрагментов и их описания используются для построения оценок функций правдоподобия. В текущей реализации оценка функции правдоподобия для класса строится в виде смеси нормальных плотностей, веса и параметры которых берутся из описаний фрагментов, расположенных в концевых узлах дерева, соответствующего данному классу:

$$P(x | \omega_i) = \sum_j \alpha_{i,j} N_{i,j}(x) [\tilde{B}_{i,j} - \text{концевой узел}].$$

Объединение используемых фрагментов представляет собой исходный кластер, поэтому сумма их весов равна единице. Следовательно, построенная оценка удовлетворяет условию нормировки, налагаемому на функции плотности вероятности.

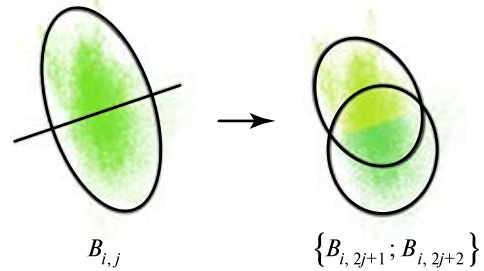


Рис. 1. Схема разбиения фрагмента кластера

3. КРИТЕРИЙ ПРИНЯТИЯ РЕШЕНИЙ

Оценки функций правдоподобия, построенные при обучении, используются на этапе принятия решений для вычисления апостериорных вероятностей (по формуле Байеса). Априорные вероятности классов в рамках решаемой задачи разумно положить равными друг другу.

Классификатор должен уметь выявлять объекты, не принадлежащие ни к одному из известных классов, и принимать решение об отказе от классификации. Будем считать, что в таких случаях классификатор относит объект к особому классу отказов, который обозначим ω_r . В текущей реализации отказ от классификации происходит в том случае, когда значение оценки функции правдоподобия на данном объекте меньше заданного порога ε для всех классов.

Если решение об отказе не принимается, объекту присваивается метка того класса, для которого апостериорная вероятность в условиях появления данного объекта имеет наибольшее значение. Суммируя сказанное, ответ классификатора на объекте \mathbf{x} можно записать следующим образом:

$$\omega(\mathbf{x}) = \begin{cases} \arg \max_{\omega_i} P(\omega_i | \mathbf{x}), & \exists \omega_i : P(\mathbf{x} | \omega_i) \geq \varepsilon, \\ \omega_r, & \forall \omega_i : P(\mathbf{x} | \omega_i) < \varepsilon. \end{cases}$$

4. ПОСТРОЕНИЕ ПОЛНЫХ И МИНИМАЛЬНЫХ ДЕРЕВЬЕВ

Рассмотренный выше алгоритм рекурсивного разбиения кластеров поддерживает различные стратегии последовательного выбора фрагментов для разбиения.

Простейшая стратегия заключается в построении полных деревьев, глубина которых заранее определена и одинакова для всех кластеров. При этом каждый раз разбиваются все фрагменты, полученные на предыдущем шаге, т. е. к деревьям добавляется по одному целиком заполненному уровню. Процедура повторяется до достижения заданной глубины.

Указанный подход часто приводит к формированию избыточных представлений и имеет ряд недостатков. Во-первых, число концевых узлов полного дерева находится в экспоненциальной зависимости от глубины. Поэтому при повышении точности аппроксимации резко возрастают требуемый объём памяти и вычислительные затраты, связанные с принятием решений. Во-вторых, каждая компонента смеси строится независимо от остальных, в результате чего между центрами компонент возникают области пониженной плотности. Таким образом, в рамках рассматриваемых древовидных представлений кластеров необходимо построить оптимальные, которые хорошо отражали бы структуру кластеров и содержали наименьшее число компонент. Ниже описан реализованный в работе способ построения минимальных представлений.

Пусть задана обучающая выборка:

$$X = \left\{ \left\{ \mathbf{x}_k, \hat{\omega}_k \right\} \right\}_{k=1}^m,$$

где \mathbf{x}_k — векторы признаков объектов обучающей выборки; $\hat{\omega}_k$ — соответствующие им метки классов; m — длина выборки.

Введем обозначение для множества, состоящего из деревьев фрагментов различных кластеров:

$$\mathbf{B} = \{B_i\},$$

где B_i — дерево фрагментов для i -го класса.

Качество набора представлений будем оценивать при помощи среднего числа ошибок, допускаемых соответствующим классификатором на обучающей выборке:

$$Q(\mathbf{B}, X) = \frac{1}{m} \sum_{k=1}^m [\omega(\mathbf{x}_k) \neq \hat{\omega}_k].$$

Стратегия разбиения кластеров основана на минимизации указанного критерия. На каждом шаге осуществляется перебор всех возможных разбиений, и реализуется то из них, которое приводит к набору представлений с наименьшим значением $Q(\mathbf{B}, X)$. Шаг алгоритма разбиения можно записать следующим образом:

$$\mathbf{B}^{t+1} = \arg \min_{\{\mathbf{B}_+^t(i', j') : B_{i', j'} \in S^t\}} Q(\mathbf{B}_+^t(i', j'), X),$$

где \mathbf{B}^t — множество деревьев фрагментов, сформированное к началу шага t ; $\mathbf{B}_+^t(i', j')$ — множество, полученное из него после разбиения некоторого фрагмента $B_{i', j'}$; S^t — множество подлежащих перебору фрагментов. В текущей реализации перебор ведётся по всем фрагментам всех кластеров, которые расположены в конечных узлах соответствующих деревьев.

При построении древовидных представлений используются параметры, значения которых с точки зрения приведенных алгоритмов могут быть произвольными. В случае полных деревьев таким параметром является максимальная глубина (число уровней), а в случае минимальных деревьев — общее число конечных узлов. Параметры такого рода принято настраивать с помощью процедуры скользящего контроля [Theodoridis, Koutroumbas, 2006]. Для этого обучающая выборка разбивается на две части: одна используется для обучения классификатора, а другая предъявляется полученному классификатору, и вычисляется доля допущенных на ней ошибок. Ошибкой скользящего контроля называется полученное указанным способом относительное значение ошибки, усредненное по нескольким разбиениям обучающей выборки. Параметры классификатора следует выбирать так, чтобы ошибка скользящего контроля принимала наименьшее значение.

5. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Схема эксперимента

Целью эксперимента является сравнение трёх типов байесовских классификаторов, использующих различные способы представления кластеров: на основе гистограмм, на основе полных деревьев фрагментов и их описаний, а также на основе минимальных деревьев.

Классификаторы исследуются применительно к задаче сегментации полутоновых изображений. Снимки взяты из базы данных Google Earth, на них изображена местность к югу от Берлина, высота камеры — 4,5 км. Для описания свойств текстуры использованы 2-й, 6-й и 9-й признаки Харалика (число уровней яркости — 32, размер окна — 9 пикселей).



Рис. 2. Области, формирующие обучающую выборку. Красным обозначены поля, зелёным — лесные массивы, синим — городские застройки

Рассматривается три класса объектов, соответствующих разным типам поверхности. Первый класс соответствует лугам и полям, второй — лесным массивам, третий — городским застройкам. Объекты, не относящиеся ни к одному из перечисленных типов, образуют особый класс отказов. Обучающая выборка для каждого класса задается набором прямоугольных областей (рис. 2). Область сегментации на изображении также имеет прямоугольную форму.

Представления кластеров гистограммами строились с использованием равномерного квантования с одинаковым числом уровней квантования по каждой координате. Представления на основе полных и минимальных деревьев строились в виде гауссовых смесей, компоненты которых соответствуют конечным узлам деревьев. Глубина (число уровней) полных деревьев полагалась одинаковой для всех кластеров.

Параметры классификаторов настраивались с использованием процедуры скользящего контроля [Theodoridis, Koutroumbas, 2006]. Обучающая выборка разбивалась на равные части, число разбиений равнялось 25.

Результаты сегментации аэрокосмических изображений

Зависимость ошибки скользящего контроля (cross-validation error e_{cv}) от числа уровней квантования n_q при представлении кластеров гистограммами приведена на рис. 3а. Наилучший результат достигается при 16 уровнях квантования. При

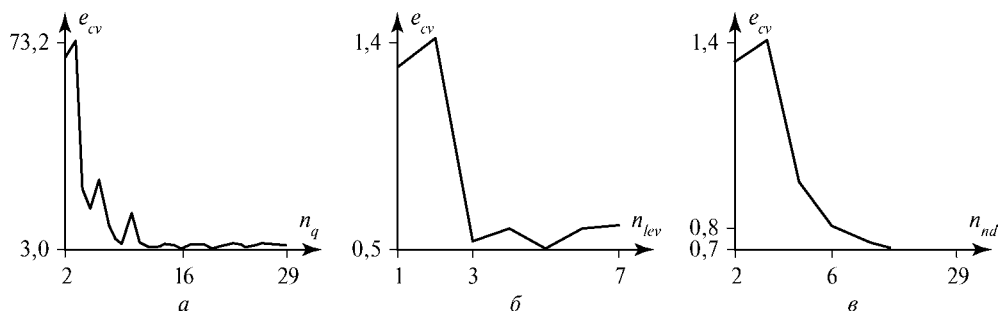
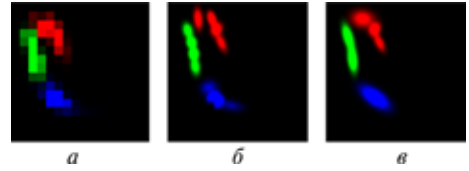


Рис. 3. Зависимость ошибки скользящего контроля e_{cv} от параметров классификаторов: *a* — от числа уровней квантования n_q по каждой координате при использовании гистограмм; *б* — от глубины деревьев n_{lev} при использовании полных деревьев; *в* — от числа конечных узлов n_{nd} при использовании минимальных деревьев

Рис. 4. Представления кластеров в двумерном пространстве признаков (цвета аналогичны рис. 2): *а* — гистограммы; *б* — представления полными деревьями; *в* — представления минимальными деревьями



более грубом квантовании представления кластеров слишком сильно перекрываются, и возрастает число ошибок классификации. При более точном квантовании объём обучающей выборки становится малым по сравнению с числом квантов, и возрастает число отказов от классификации. Зависимость ошибки скользящего контроля e_{cv} от глубины дерева (number of levels n_{lev}) при построении представлений на основе полных деревьев приведена на рис. 3б. Использование трёхуровневых деревьев приводит к заметному повышению качества по сравнению с двумя уровнями, а дальнейшее увеличение числа уровней влияет на качество достаточно слабо. Для случая представлений минимальными деревьями зависимость ошибки скользящего контроля e_{cv} от суммарного числа концевых узлов (number of end nodes n_{nd}) во всех деревьях приведена на рис. 3в. Видно, что при шести узлах значение ошибки заметно ниже начального уровня, а дальнейшее наращивание леса фрагментов уменьшает ошибку незначительно.

Описанные выше представления кластеров проиллюстрированы на рис. 4 для двумерного признакового пространства, образованного 6-м и 9-м признаками Харалика [Miyamoto, Merryman, 2005]. Видно, что оба представления на основе древовидных структур адекватно отражают структуру кластеров, но представление на основе минимальных деревьев лишено некоторых артефактов, связанных с избыточным дроблением фрагментов. Деревья, соответствующие представлениям, схематично изображены на рис. 5.

Результаты сегментации изображения с использованием трёх вышеперечисленных способов представления кластеров даны на рис. 6 (см. с. 300). Исходное изображение дано на рис. 6а, результат, полученный с использованием гистограмм с 16 уровнями квантования на признак, приведен на рис. 6б. В этом случае, даже при оптимальном квантовании, наблюдается значительное число отказов от классификации (чёрные пиксели). На рис. 6в показан результат, полученный на основе представлений полными трёхуровневыми деревьями. В этом случае качество сегментации существенно выше, чем при использовании гистограмм. На рис. 6г показано сегментированное изображение с использованием леса минимальных деревьев, в котором общее число концевых узлов равно шести. В этом случае качество сегментации сопоставимо с результатом, полученным для полных трёхуровневых деревьев.

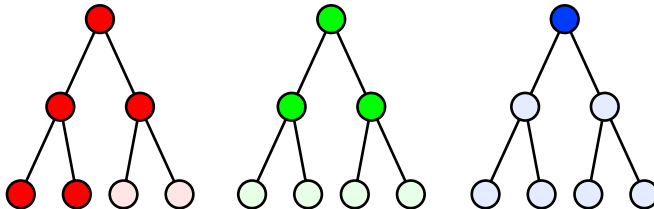


Рис. 5. Деревья, соответствующие оптимизированным представлениям кластеров (цвета аналогичны рис. 2): насыщенные цвета — узлы минимальных деревьев; ненасыщенные цвета — узлы полных деревьев, не входящие в минимальные деревья

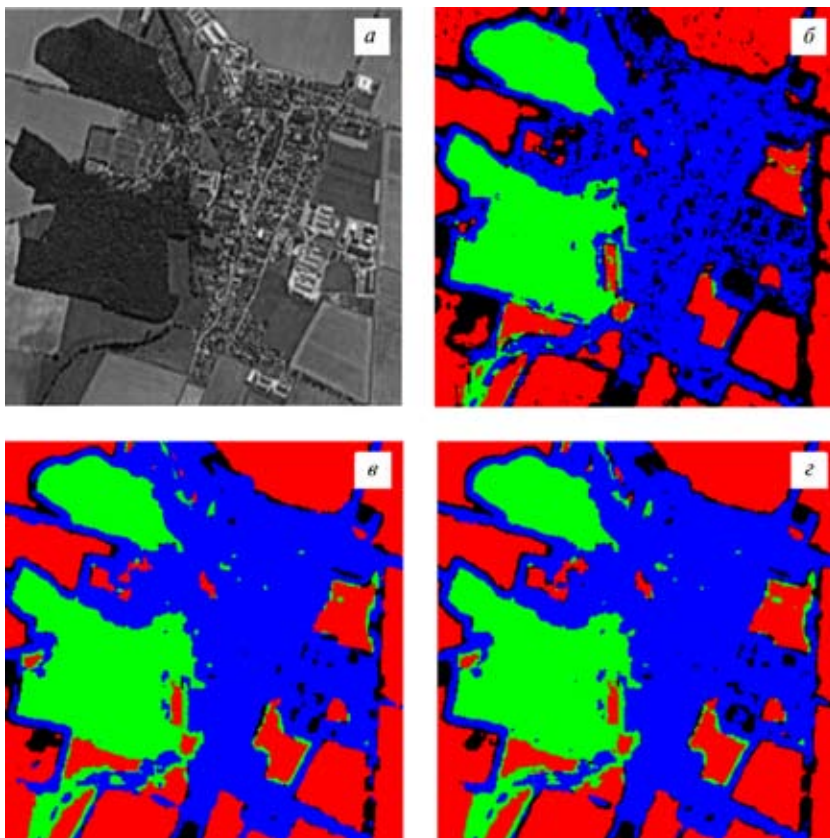


Рис. 6. Результаты сегментации аэрокосмического изображения (цвета аналогичны рис. 2): *a* — исходное изображение; *б–г* — результаты с использованием гистограмм, полных деревьев, минимальных деревьев соответственно

Объёмы памяти (числа ячеек), которые требуются для хранения рассмотренных представлений кластеров, приведены ниже. При расчётах полагалось, что необходимо хранить все узлы деревьев, а не только концевые. Для хранения одного узла в случае трёх признаков требуется 13 ячеек памяти (3 — на вектор средних значений, 9 — на ковариационную матрицу и 1 — на весовой коэффициент). Видно, что переход от гистограмм к древовидным представлениям позволяет сократить объём памяти почти на два порядка. Использование минимальных деревьев по сравнению с полными деревьями уменьшает объём памяти более чем в два раза.

Объём памяти для хранения различных представлений кластеров

Тип представления кластеров	Требуемый объём ячейки памяти
Гистограммы	$3 \times 16^3 = 12\,288$
Полные деревья	$3 \times 3(2^3 - 1)(3^2 + 3 + 1) = 273$
Минимальные деревья	$(3 + 2(6 - 3))(3^2 + 3 + 1) = 117$

Заключение

В статье рассмотрено построение байесовского классификатора, использующего оценки функций правдоподобия, основанные на древовидном представлении кластеров. Основными преимуществами классификатора являются небольшой объём памяти, требуемый для хранения представлений, регулируемая точность описания кластеров и ускоренное обучение (по сравнению с EM-алгоритмом).

Исследовано применение классификатора к задаче сегментации полутоновых изображений. Проведено сравнение предложенного подхода с представлением кластеров гистограммами. Показано, что в условиях проведённого эксперимента использование гистограмм не способно обеспечить приемлемое качество сегментации, а применение древовидных представлений даёт заметно лучший результат. Переход от гистограмм к древовидным представлениям позволил сократить требуемый объём памяти приблизительно в 50 раз. Применение минимальных деревьев привело к дополнительному уменьшению объёма памяти в 2 раза (по сравнению с полными деревьями) без ухудшения качества сегментации.

Рассмотренный классификатор позволяет работать в любом непрерывном признаковом пространстве, а также допускает различные модификации процедуры разбиения отдельных фрагментов, стратегии наращивания деревьев и правила построения оценок функций правдоподобия. Допускается также комплексирование с другими классификаторами путём взвешенного суммирования апостериорных вероятностей.

В будущем планируется реализовать некоторые модификации разработанного классификатора, например, использовать все узлы деревьев, а не только концевые при построении аппроксимаций функций правдоподобия гауссовыми смесями. Предполагается также исследовать возможности комплексирования и применимость классификатора к другим задачам распознавания образов.

Литература

- Гонсалес Р., Вудс Р. (2006) Цифровая обработка изображений. М.: Техносфера, 2006.
- Дуда Р., Харт П. (1976) Распознавание образов и анализ сцен. М.: Мир, 1976.
- Ganebnykh S. N., Lange M. M. (2009) Classification of 2D Grayscale Objects in a Space of the Multiresolution Representations // Pattern Recognition and Image Analysis, Nauka-Interperiodica. 2009. V. 19. N. 4. P. 591–602.
- Malik J., Belongie S., Leung Th., Shi J. (2001) Contour and Texture Analysis for Image Segmentation // Intern. J. Computer Vision 2001. V. 43. N. 1. P. 7–27.
- Mancas M., Gosselin B., Macq B. (2005) Segmentation Using a Region Growing Thresholding // Proc. Electronic Imaging Conf. of the Intern. Soc. for Optical Imaging (SPIE/EI 2005), San Jose (California, USA). 2005. V. 5672. P. 388–398.
- Miyamoto E., Merryman-Jr. T. (2005) Fast Calculation of Haralick Texture Features [Электронный ресурс]. Режим доступа: <http://www.ece.cmu.edu/~pueschel/teaching/18-799B-CMU-spring05/material/eizan-tad.pdf> (последний доступ 30.11.2008).
- Szirányi T., Zerubia J., Czúni L., Geldreich D., Kato Z. (2000) Image Segmentation Using Markov Random Field Model in Fully Parallel Cellular Network Architectures // Real-Time Imaging. 2000. V. 6. P. 195–211.
- Theodoridis S., Koutroumbas K. (2006) Pattern Recognition. 3rd ed. USA: Academy Press, 2006.

**IMAGE SEGMENTATION BASED ON DESCRIBING CLUSTERS
BY STRUCTURED GAUSSIAN MIXTURES**

M. M. Lange¹, N. A. Novikov²

¹ *Institution of Russian Academy of Sciences Dorodnicyn Computing Centre of RAS
(CC RAS), 40 Vavilov st., 119333, Moscow*

² *Moscow Institute of Radio Engineering and Automatics
78 Vernadsky pr., 119454, Moscow*

A new approach to image segmentation problem by using tree-structured descriptions of clusters in a multidimensional space of features is proposed. The description modifications based on optimization of the tree structures are developed by applying a cross validation procedure. Using the tree-structured cluster descriptions by the Gaussian mixtures, a segmentation algorithm is suggested as the Bayesian classification procedure by a maximum posterior probability criterion. For aerospace images, the proposed algorithm shows a significantly less number of the classification rejects as compared to the known histogram-based algorithm. The comparative estimations of memory costs provide more than fifty times profit of the suggested algorithm depending on the applied tree-structured cluster descriptions.

Keywords: image segmentation, cluster, Gaussian mixture, histogram, likelihood function, posterior probability, Bayesian classifier, tree-structured description.

Lange Mikhail Mikhailovich — Leading Researcher, Candidate of Technical Science, e-mail: lange_mm@ccas.ru.

Novikov Nikita Aleksandrovich — Post-Graduate Student, e-mail: nikknovikov@gmail.com.

БЫСТРЫЕ АЛГОРИТМЫ АНАЛИЗА ИЗОБРАЖЕНИЙ СО СТРУКТУРНЫМИ ИСКАЖЕНИЯМИ

Ф. А. Корнилов, В. Б. Костоусов, Д. С. Первалов

*Учреждение Российской академии наук Институт математики и механики
Уральского отделения РАН (ИММ УрО РАН),
620990, Екатеринбург, Софьи Ковалевской ул., 16*

В работе предлагаются и исследуются новые эффективные алгоритмы поиска по эталону и поиска изменений на изображениях, применимые для анализа изображений реальных сцен в случае наличия структурных искажений интересующих объектов.

Работа выполнена при финансовой поддержке программ фундаментальных исследований Президиума РАН (№ 09-П-1-1003, 09-П-1-1013) и РФФИ (проект № 09-01-00523).

Ключевые слова: структурные искажения, сопоставление с эталоном, матрицы сравнений, обнаружение изменений.

Введение

При решении задач автоматического анализа изображений реальных сцен возникают специфические проблемы, связанные с изменчивостью цвета и формы интересующих объектов. Имеется достаточно большое число работ, в которых предлагаются способы устранения таких проблем — например, удаление теней и бликов при условии наличия известной априорной информации о положении источника света или типа наблюдаемых объектов. Зачастую предлагаемые подходы эффективно решают одну конкретную задачу либо являются универсальными, но вычислительно неэффективными, так как в них используются методы глобальной оптимизации типа имитации отжига.

В работе будут рассмотрены два новых алгоритма анализа изображений реальных сцен, позволяющих эффективно работать со «структурными искажениями» интересующих объектов. Под структурными искажениями здесь понимаются такие изменения изображений, которые не моделируются с помощью аддитивного шума или линейных фильтров. Примерами структурных искажений являются изменения геометрии объекта, цвета типа бликов и теней или окраски некоторой части объекта.

Первый рассматриваемый алгоритм — сопоставление с эталоном на основе матриц сравнений, второй алгоритм — нахождение областей структурных изменений на двух геометрически выровненных изображениях. Для каждого алгоритма описаны модели соответствующих классов структурных искажений, при которых алгоритмы работают устойчиво.

Корнилов Федор Андреевич — программист.

Костоусов Виктор Борисович — заведующий отделом, кандидат физико-математических наук, доцент.

Первалов Денис Сергеевич — главный программист, e-mail: denis.perevalov@mail.ru.

1. ПОНЯТИЕ СТРУКТУРНЫХ ИСКАЖЕНИЙ

Компоненты формирования цифрового изображения — сцена, снимающая её камера и среда, через которую идут световые лучи от сцены к камере. Каждый из этих элементов является источником искажений результирующего цифрового изображения: снимаемые объекты в сцене меняют цвет и форму; среда вносит шумы, связанные с условиями освещения и факторами типа дождя, снега и теней от других объектов; камера дает дефокусировку и шумы, связанные с АЦП.

Будем называть *структурными* такие искажения изображения, которые нельзя смоделировать наложением аддитивного шума, применением линейной фильтрации и т. п. [Костоусов, Перевалов, 2008; Перевалов, 2009]. Иными словами, структурные искажения содержат «комбинаторный» элемент. Можно выделить следующие классы структурных искажений: цветовые — связанные с изменением цвета всего объекта или его части, а также бликами или тенями на объекте, и геометрические — изменение формы самого объекта либо загороживание его другими объектами.

Существуют различные сценарии преодоления проблем построения алгоритма анализа изображений при наличии структурных искажений:

- 1) описать условия применимости алгоритма с точки зрения структурных искажений, т. е. класс структурных искажений, с которыми гарантируется устойчивая работа алгоритма;
- 2) использовать для построения алгоритма специфику конкретной задачи (структурные свойства искомого объекта и окружающей сцены);
- 3) построить модель структурных искажений и на её основе разработать алгоритм, устойчивый к заданному классу искажений.

Наиболее предпочтительным является последний вариант и именно на него мы ориентировались в данной работе.

2. АЛГОРИТМ СОПОСТАВЛЕНИЯ С ЭТАЛОНОМ НА ОСНОВЕ АППАРАТА МАТРИЦ СРАВНЕНИЙ

Одним из простейших вариантов структурных искажений является изменение яркости пикселей изображения с помощью произвольной монотонно возрастающей функции. В этом случае отношение порядка между любой парой пикселей остается неизменным. Это свойство положено в основу алгоритма сопоставления с эталоном на основе порядковых статистик, с настроечным параметром, позволяющим регулировать соотношение качество/скорость [Перевалов, 2006б]. В основе алгоритма лежит использование аппарата матриц сравнений.

Для полутонового изображения $X = (X_1, \dots, X_n)$, состоящего из n пикселей, $X_i \in [0, 1]$, *матрицей сравнения* будем называть матрицу $M = M(X)$ размером $n \times n$, задаваемую по правилу:

$$M(X)_{ij} = \begin{cases} 1, & X_i \geq X_j, \\ 0, & X_i < X_j. \end{cases}$$

Матрица сравнений для полутонового изображения представляет собой матрицу бинарного отношения «больше», построенную для пикселей этого изображения. Она является порядковой характеристикой изображения, инвариантной к изменениям яркости и контраста, а также устойчива к изменениям яркости и контраста отдельных частей объекта.

Справедливо следующее утверждение [Перевалов, 2006б]. Пусть определён класс изображений, характеризуемый набором пар областей, для которых яркость любого пиксела из одной области не меньше яркости любого пиксела из другой области. В этом случае значение некоторого элемента (i, j) средней матрицы сравнений для этого класса имеет значение 1 тогда и только тогда, когда пиксела i и j пиксела лежат соответственно в некоторой паре областей.

Из утверждения следует, что если имеется набор эталонных изображений, характеризующих некоторый более широкий класс изображений, то элементы с максимальными значениями средней матрицы сравнений для набора эталонов приближенно характеризуют области различной яркости всего класса изображений.

Приведём описание алгоритма, построенного на основе данного следствия.

1. Расчёт по эталонным изображениям средней матрицы сравнения.
2. Выбор K пар индексов (i_k, j_k) , в которых соответствующие элементы матрицы принимают наибольшие значения. (Параметр позволяет в некоторых пределах регулировать соотношение качество/скорость получаемого алгоритма. Обычное значение этого параметра $K = n$.)
3. Для каждого положения сканирующего окна вычисляем значение следующего критерия сопоставления входного изображения X :

$$dist(X) = K - \sum_{k=1}^K M(X)_{i_k j_k}.$$

4. Окно с минимальным значением критерия является результатом работы алгоритма. Близость критерия к нулю в этом окне характеризует степень уверенности в правильности решения.

Работоспособность предложенного алгоритма была проверена путём сравнения с другими широко известными алгоритмами на примере одной задачи анализа реальных объектов, а именно, задачи обнаружения изображения автосцепки. На входном изображении требуется обнаружить автосцепку двух сцепленных железнодорожных вагонов (рис. 1) либо сообщить, что на изображении её нет.

Для тестирования использовались 685 кадров с изображением 137 автосцепок, при этом каждая автосцепка была представлена пятью кадрами с разных ракурсов; а также 23 211 кадров, на которых не было автосцепки. Использовались следующие критерии сопоставления:

- 1) $dist_{L_2}(X; T)$ — расстояние от X до ближайшего эталона в метрике L_2 ;
- 2) $r(X; T)$ — коэффициент корреляции изображения X с усредненным изображением эталонов;
- 3) $\rho(X; T)$ — коэффициент корреляции Спирмена рангового вектора изображения X с усредненным вектором рангов эталонов;
- 4) $dist(X; T)$ — предложенный критерий с $K = n$;

5) конкретный алгоритм Ad hoc, основанный на поиске темных прямоугольников в определённой конфигурации [Перевалов, 2006а].



Рис. 1. Пример изображения узла автосцепки

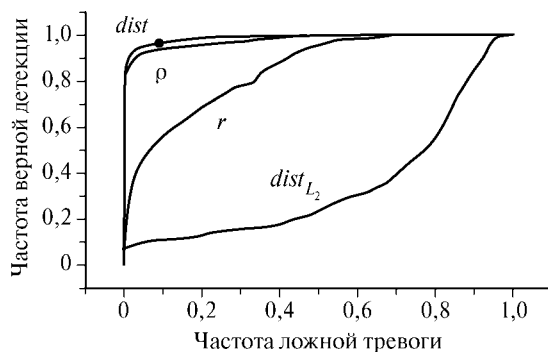


Рис. 2. ROC-кривые

На каждом тестовом изображении было проведено сканирование локальным окном, в каждой позиции окна было вычислено значение критерия. Экстремальное значение критерия для изображения (минимум для $dist_{L_2}(X; T)$ и $dist(X; T)$; максимум для $r(X; T)$ и $\rho(X; T)$) было записано в таблицы. На основании таблиц для всевозможных пороговых значений были вычислены частоты верной детекции и ложной тревоги. Результаты, отражающие эти данные, представлены на рис. 2 в виде ROC-кривых. (Алгоритму Ad hoc соответствует не график, а жирная точка рядом с графиком $dist$.)

Чем ближе график такой кривой к верхней горизонтали $y = 1$, тем алгоритм точнее, а площадь под кривой есть численная мера точности. В данном случае площади оказались равными 0,3423; 0,8491; 0,9793; 0,9894 для критериев $dist_{L_2}(X; T)$, $r(X; T)$, $\rho(X; T)$, $dist(X; T)$ соответственно. Таким образом, предложенный критерий показал лучший результат, сопоставимый по качеству с коэффициентом корреляции Спирмена. Число операций у всех методов асимптотически — $O(n)$, а фактически число операций наименьшее у предложенного метода, так как он требует только n операций сравнения яркостей при $K = n$.

3. АЛГОРИТМ НАХОЖДЕНИЯ ОБЛАСТЕЙ СТРУКТУРНЫХ ИЗМЕНЕНИЙ

Второй рассматриваемый алгоритм — нахождение областей структурных изменений на двух геометрически выровненных изображениях [Первалов, Корнилов, 2009а, б). Работа алгоритма основана на построении функции коррекции гистограммы одного изображения; для этого используется связь яркостей соответствующих пикселей между обоими изображениями.

Входными данными являются два цифровых полутоновых космических снимка одного и того же участка земной поверхности. Предполагается, что они имеют одинаковый размер в пикселах и геометрически выровнены. Геометрическая выровненность означает, что каждый пиксел (x, y) на обоих снимках соответствуют одному и тому же участку земной поверхности. Требуется найти на изображениях области структурных изменений, под которыми здесь понимаются области, в которых есть появившиеся, исчезнувшие или изменившиеся объекты. При этом не должны быть выбраны те объекты, которые изменили лишь свой цвет (например, здание с перекрашенной крышей).

Специфика поиска изменений объектов на снимках земной поверхности состоит в том, что области могут менять свой цвет сложным образом, и один и тот же объект на разных снимках может существенно различаться по набору яркостей соответствующих пикселей. В то же время необходима высокая точность определения факта изменения для весьма малых объектов и точность построения контура вокруг изменившихся областей.

При работе алгоритма исходная пара изображений сканируется окном фиксированного размера. Для устранения возможного несовершенства синхронизации исходных изображений проводится следующая геометрическая коррекция: при зафиксированном положении сканирующего окна на одном из изображений на другом изображении осуществляется сдвиг сканирующего окна в пределах некоторого радиуса. Для каждого вновь полученного окна вычисляется «степень биективности» (о которой речь пойдет ниже) и среди всех рассмотренных окон выбирается то, для которого значение «степени биективности» наименьшее. Это окно сдвигается на место исходного и дальнейшая работа ведётся с ним. А именно, сначала проводится локальная корректировка яркостей изображений с целью грубой оценки изменений. Затем осуществляется точная оценка изменений с помощью вычисления «степени биективности».

После этого с использованием пороговой обработки собираются все локальные окна с большой величиной изменения, из которых формируются связные области, являющиеся результатом работы алгоритма.

Опишем шаги грубого и точного оценивания подробней.

3.1. Грубая оценка изменений

Для построения грубой оценки проводится выравнивание яркостей двух сравниваемых изображений с сохранением их структуры. Для этого используется специальная функция преобразования яркости:

$$f_{AB}(c) = \overline{\{I_B(x, y) : I_A(x, y) = c\}},$$

где c — некоторая яркость; I_A, I_B — входные изображения; (x, y) — координаты текущей точки в сканирующем окне; $f_{AB}(c)$ — среднее значение набора яркостей второго изображения в точках, яркость которых на первом изображении равна c . Нужно заметить, что класс подобных функций преобразования появился впервые в теории морфологий Пытьева [Пытьев, 1983].

Структура грубой оценки изменений:

- 1) построение функций преобразования яркости $f_{AB}(c)$ и $f_{BA}(c)$;
- 2) создание преобразованных изображений: $I_{A'} = f_{AB}(I_A)$, $I_{B'} = f_{BA}(I_B)$;
- 3) вычисление числа точек, для которых модуль разницы между значениями на одном (преобразованном) изображении и на втором (исходном) превосходит некоторый порог T_1 :

$$S_{AB} = \left| \left\{ (x, y) : \left| I_{A'}(x, y) - I_B(x, y) \right| \geq T_1 \right\} \right|,$$

$$S_{BA} = \left| \left\{ (x, y) : \left| I_{B'}(x, y) - I_A(x, y) \right| \geq T_1 \right\} \right|;$$

- 4) из двух полученных значений выбирается наибольшее:

$$S = \max\{S_{AB}, S_{BA}\}.$$

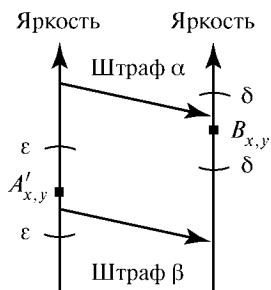


Рис. 3. Метод расчёта «степени биективности»

Использование такой схемы дает симметричный результат, не зависящий от порядка изображений.

3.2. Вычисление «степени биективности»

Точная оценка изменений проводится только для подозрительных фрагментов изображений, т.е., если величина грубой оценки изменений S в некоторой точке достаточно велика (превосходит значение некоторого порога), то для этой точки проводится вычисление «степени биективности», полученная величина полагается результатом обработки данного сканирующего окна и записывается в его центр. В противном случае это значение равно нулю.

В основе вычисления «степени биективности» лежит следующая модель: при отсутствии структурных изменений на изображениях малая окрестность яркости каждого пиксела должна перейти в малую окрестность яркости соответствующего пиксела, а точки вне этой окрестности не должны попасть в эту окрестность.

Формально это можно определить с помощью понятия штрафа. Для этого задаются два радиуса окрестностей — ϵ для первого и δ — для второго изображения. Пиксеты, яркость которых выходит за пределы ϵ -окрестности границы яркости центральной точки первого изображения, но на втором изображении попадает в δ -окрестность яркости соответствующей точки, «штрафуются» на величину α . Аналогично, пиксеты, яркость которых находится в ϵ -окрестности первого изображения, но на втором не попадает в δ -окрестность, «штрафуются» на величину β (рис. 3). Итоговая оценка в данной точке, которую будем называть «степенью биективности», есть сумма штрафов для всех пикселей локального окна.

Был проведён ряд экспериментов по проверке работы алгоритма на разновременных космических снимках. Результаты показали достаточно устойчивую работу алгоритма в случае использования снимков, снятых одним типом спутника. Алгоритм показал свою пригодность как инструмент для поиска изменений в городской застройке.

Заключение

Результаты настоящей работы показали эффективность использования моделей структурных искажений как средства построения и исследования эффективных алгоритмов анализа изображений. В дальнейшем предполагается продолжить исследования различных классов структурных искажений. Кроме того, планируется применить матрицы сравнений в качестве источника для построения признаков в алгоритмах распознавания типа Виолы – Джонса [Viola, Jones, 2001].

Литература

- Костоусов В. Б., Первалов Д. С. (2008) О задачах обнаружения и локализации объектов на изображении // Тез. докл. Международ. конф. «Алгоритмический анализ неустойчивых задач». Екатеринбург: Изд-во Урал. ун-та, 2008. С. 136–137.
- Первалов Д. С. (2006а) Вероятностная модель и эффективный алгоритм распознавания изображения автосцепки // Тр. 37-й Регион. молодёж. конф. «Проблемы теоретической

- и прикладной математики». ИММ УрО РАН. Екатеринбург: Изд-во УрО РАН, 2006. С. 468–472.
- Перевалов Д. С.* (2006б) Использование матриц сравнений в задаче поиска по эталону // Материалы 9-й Международ. конф. «Интеллектуальные системы и компьютерные науки». МГУ. М.: Изд-во мех.-мат. ф-та МГУ, 2006. С. 226–228.
- Перевалов Д. С.* (2009) Исследование алгоритмов обнаружения и локализации объекта на изображении в условиях структурных искажений // Вычисл. технологии. 2009. Т. 14. № 1. С. 94–104.
- Перевалов Д. С., Корнилов Ф. А.* (2009а) Алгоритм поиска структурных изменений на изображениях // СПИСОК-2009 [Системное программирование. Интеллектуальные системы. Обеспечение качества-2009]: Материалы межвуз. науч. конф. по проблемам информатики. 20–23 апр. 2009. Екатеринбург: Изд-во Урал. ун-та, 2009. С. 264–266.
- Перевалов Д. С., Корнилов Ф. А.* (2009б) Алгоритмы обнаружения структурных изменений на одновременных снимках // Тез. докл. Регион. научно-техн. конф. «Данные дистанц. зондирования Земли для решения практич. задач развития территорий». Екатеринбург: ФГУП «Уралгеоинформ», 2009.
- Пытнев Ю. П.* (1983) Морфологический анализ изображений // Докл. АН СССР. 1983. Т. 269. № 5. С. 1061–1064.
- Viola P., Jones M. J.* (2001) Rapid Object Detection Using a Boosted Cascade of Simple Features // CVPR'01: Proc. Conf. Computer Vision and Pattern Recognition. Los Alamitos, CA, USA. 2001. P. 511–518.

RAPID ALGORITHMS OF IMAGE ANALYSIS WITH STRUCTURAL NOISE

F. A. Kornilov, V. B. Kostousov, D. S. Perevalov

*Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy
of Sciences (IMM), 16 S. Kovalevskaja st., 620219, Ekaterinburg, Russia*

In this paper, it is proposed and studied the new efficient algorithms for pattern matching search and changes search on images. These algorithms are applicable to analysis of realistic scene images in the case of structural distortions of objects of interest.

Keywords: structural noise, pattern matching, comparison matrices, change detection.

Kornilov Fedor Andreevich — Programmer.

Kostousov Victor Borisovich — Department Leader, Doctor of Physical and Mathematical Sciences, Associate Professor.

Perevalov Denis Sergeevich — Chief Programmer, e-mail: denis.perevalov@mail.ru.

ПРОГРАММНО-АППАРАТНАЯ РЕАЛИЗАЦИЯ КОРТЕКСНОГО СКАНЕРА

А. А. Домунян

*Учреждение Российской академии наук Институт проблем
управления им. В.А. Трапезникова РАН (ИПУ РАН),
117997, ГСП-7, Москва, Профсоюзная, 65*

В статье рассматривается задача создания кортексного сканера, являющегося универсальным аппаратным средством выделения признаков на этапе предварительной обработки изображений для последующего использования признаков при распознавании изображений. Работа кортексного сканера моделирует процессы обработки информации, происходящие в первичной зрительной области коры головного мозга. Рассмотрены также аппаратные аспекты создания такого устройства, основанного на программируемых логических интегральных схемах, получающих информацию от микрокамер.

Ключевые слова: искусственный нейронный кортекс, кортексный сканер, распознавание образов, искусственный интеллект.

Введение

Машинное зрение находит применение во многих областях техники. Системы машинного зрения выполняют ввод изображений, их предобработку и распознавание. Если ввод изображений включает стандартизированные процедуры, то алгоритмы предобработки специфичны и зависят от решаемых задач, поскольку для различных задач, как правило, требуются те или иные специфические признаки, которые отыскиваются на изображениях на этапе предобработки. В данной статье рассматривается возможность создания относительно универсальных программно-аппаратных средств для этапа предобработки изображений, что позволит в некоторой степени стандартизировать признаки и процедуры их выделения. Такая универсализация может быть достигнута в результате механизмов работы первичной зрительной области VI коры головного мозга.

На первый взгляд нейроны расположены в коре головного мозга совершенно хаотично. На самом деле нейроны объединяются в группы интересов и сплетаются в колонки, содержащие примерно по 100 нейронов. При этом нейроны каждой колонки активизируются одновременно.

Из работы [Хьюбел, 1990] известно, что колонки в области V1 реагируют на линии, расположенные под определённым наклоном. Например, одна колонка будет активироваться, когда через неё проходит линия под углом 15° , а другая — когда через неё проходит линия под углом 20° .

На рис. 1 (см. с. 311) колонки, представленные секторами, активизируются, когда изображение линии проецируется на соответствующий сектор.

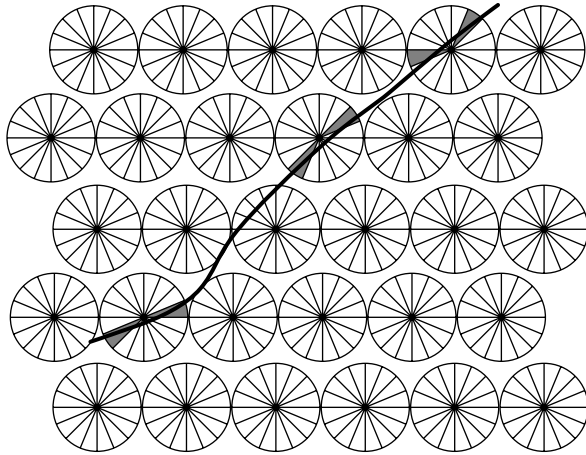


Рис. 1. Активация колонок первичной зрительной области в зависимости от углов проецируемых линий

ПОДХОД К ЦИФРОВОЙ РЕАЛИЗАЦИИ ФУНКЦИЙ ПЕРВИЧНОЙ ЗРИТЕЛЬНОЙ ОБЛАСТИ

Цифровым аналогом выходных данных этой области могут служить числа x , y и α , где x и y — пространственные координаты колонки, а α — угол, под которым проходит линия. В данной статье рассматривается задача создания устройства, называемого кортексным сканером, выходом которого служат (наряду с традиционными координатами x , y пикселей изображения) доминирующие градиенты α пикселей.

Если в области VI колонки, связанные с пикселями сетчатки, реагируют параллельно на интересующие их градиенты, то в кортексном сканере некоторое маленькое окно размером $M \times M$ пикселей, например, $M = 32$, последовательно сканирует все пиксели изображения. Над окном находятся 32×32 колонки памяти. Ячейки каждой колонки содержат имена градиентов, которые активизируются в зависимости от наклона линии (градиента изображения) в точке, соответствующей центру окна. Максимально активное, т. е. наиболее часто встречающееся имя, представляет собой величину градиента изображения в центральной точке. Если центр окна находится в равномерно освещённой области, то все имена градиентов будут активированы в равной степени, т. е. победителя нет, и, значит, в данной точке градиент отсутствует. Если центр окна находится на границе, где имеет место перепад яркости, то появляется доминирующий градиент.

Вышесказанное иллюстрируется рис. 2а–в (см. с. 312, 313). Изображенное на них окно может детектировать четыре направления градиента: 0, 4, 8 и 19, отмеченных соответствующими линиями. Так, например, в ячейках, через которые проходит линия 0, т. е. в ячейках с координатами (0, 0), (1, 1), (2, 2) и т. д. хранится величина 0 нулевого градиента, которому условно соответствует диагональное направление. В ячейках, через которые проходит линия 4, т. е. в ячейках с координатами (4, 0), (5, 1), (5, 2) и т. д., — хранится величина градиента 4, представляющего направление 45°. И так далее. В ячейке А, через которую проходят все четыре линии, хранятся значения всех четырёх градиентов, а в ячейках В и С — градиенты 4 и 8.

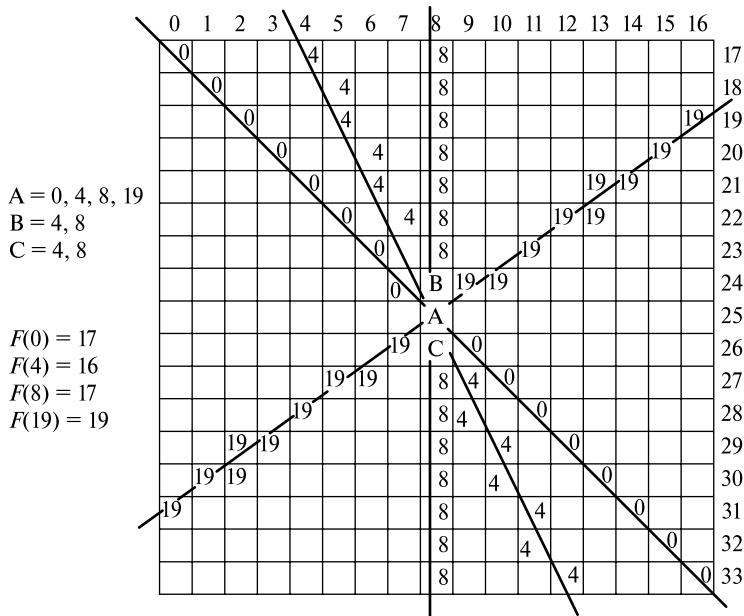


Рис. 2а. Случай равномерного освещения поля окна

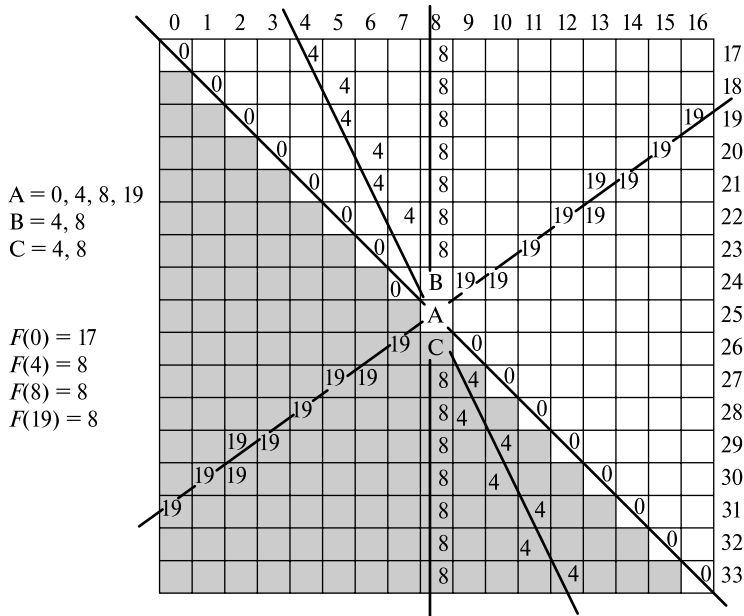


Рис. 2б. Тень покрывает левый нижний угол поля окна

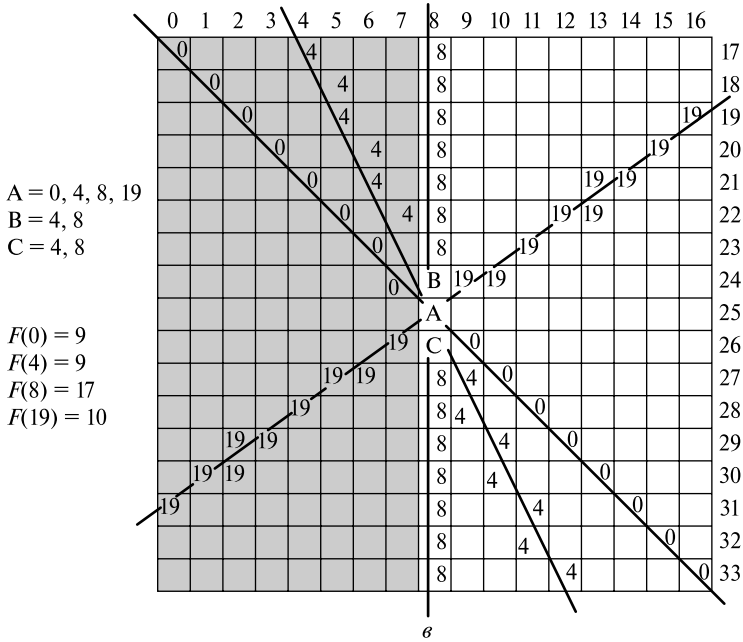


Рис. 2в. Тень покрывает левое поле окна

Можно показать, что максимальная точность аппроксимации контуров изображений достигается при минимальном размере окна. Однако минимальный размер окна ограничен эффектами дискретизации, в результате которых в одну ячейку попадают более одного значения градиента. Экспериментально установлено, что оптимальное качество достигается при размере окна 32×32 пиксела. На рис. 2а освещены все ячейки окна. Поэтому активность имен выглядит как $F(0) = 17$, $F(4) = 17$, $F(8) = 17$, $F(19) = 17$, что соответствует случаю отсутствия явного победителя и, тем самым, отсутствию градиента-победителя.

На рис. 2б освещены ячейки верхнего правого угла окна. Поэтому активность имен выглядит как $F(0) = 17$, $F(4) = 8$, $F(8) = 8$, $F(19) = 8$, что соответствует наличию нулевого градиента.

На рис. 2в освещены ячейки правой половины окна. Поэтому активность имен выглядит как $F(0) = 9$, $F(4) = 9$, $F(8) = 17$, $F(19) = 10$, что соответствует наличию вертикального градиента 8.

СХЕМОТЕХНИКА

На рис. 3 (см. с. 314) приведена блок-схема предполагаемого портативного устройства, предназначенного для предобработки изображения с целью последующего распознавания.

Здесь камера — это законченный модуль, состоящий из оптической линзы, матрицы ПЗС и внутреннего контроллера с последовательным интерфейсом ввода-вывода. КМОП-камера выглядит так, как показано на рис. 4 (см. с. 314).

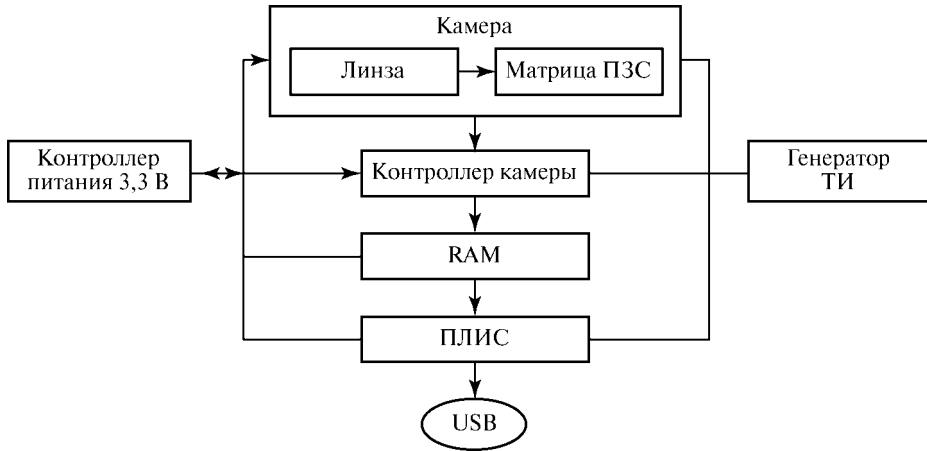


Рис. 3. Блок-схема кортексного сканера



Рис. 4. Внешний вид камеры

Контроллер камеры — это однокристальное устройство сопряжения КМОП-камеры с оперативной памятью (стек FIFO), служит для накопления данных и после получения законченного кадра передачи его в память, с которой работает ПЛИС. Этот контроллер может также выполнять и другие вычислительные функции, так как по своей сути является однокристальным микропроцессором, выполняющим функции DMA-контроллера.

Оперативная память (RAM) — это микросхема памяти, в которой умещается два изображения. В нашем случае объём памяти должен быть больше 2,6 Мбайт (так как мы используем 1,3-мегапиксельную камеру и 256 цветов каждого пиксела). Сначала данные текущего кадра записываются в первую половину памяти. Пока ПЛИС работает с первым кадром, контроллер камеры записывает следующий кадр во вторую часть (страницу) памяти. Потом, когда готов третий кадр, ПЛИС начинает обрабатывать второй кадр, а первый уже не нужен и контроллер камеры «затирает» первый кадр третьим. И так далее по очереди. Контроллер камеры записывает данные из камеры то в первую страницу памяти, то во вторую, а ПЛИС обращается за данными то во вторую страницу, то в первую.

Программируемая логическая интегральная схема ПЛИС состоит из программируемых ячеек «2И-НЕ», триггеров и системы ввода-вывода. ПЛИС позволяет при большой тактовой частоте за один такт выполнять логические битовые операции, например такие как побитовое «ИЛИ», побитовое «И». Именно эти операции присутствуют в алгоритме предобработки изображения. ПЛИС программируется на языке схемотехнических описаний Verilog или VHDL. Портом ввода-вывода для ПЛИС выбран интерфейс USB, который при помощи встроенного JTAG-интерфейса обменивается данными с управляющим устройством, например с ПК или другим аппаратным устройством, которому необходимо предобработанное изображение. ПЛИС имеет встроенный бит защиты, который не позволяет считывать записанный в неё код и тем самым защищает авторские права разработчиков. На этапе разработки имеется возможность трассировки программы, записанной в ПЛИС при помощи отладочного модуля и специализированной программы, которая по JTAG-интерфейсу считывает состояния внутренних триггеров и макроячеек ПЛИС.

Генератор тактовых импульсов — это микросхема с кварцевым генератором, задающая тактовую частоту для всех микросхем, кроме оперативной памяти, служит для синхронизации процессов.

Контроллер питания — это микросхема, стабилизирующая питание, которое необходимо для работы всех микросхем, входящих в аппаратный комплекс. Напряжение питания выбрано 3,3 В.

Технические требования к устройству: питание от аккумуляторов или батарей, оценочные параметры которых — 3,6 В; 2000 мА/ч; Li-ion или Li-pol. Устройство должно работать от одного источника не менее 12 ч без подзарядки. Оно может также подзарядать внутренний аккумулятор от порта USB, хотя полностью питание от USB исключается, поскольку по стандарту ток через порт USB не должен превышать 100 мА на устройство, а в режиме работы ток только через ПЛИС может достигать 400 мА. Качество обрабатываемого изображения не хуже 1,3 Мпиксела, 10 кадров/с, 256 цветов. Соответственно скорость шины обращения к памяти равна 13 Мбайт/с = 104 Мбит/с. Физические размеры должны быть достаточно малыми и сравнимы с размерами камеры и ПЛИС: 8,5×8,0×7,9 мм для камеры ADCM-1650-3011 и 15,2×15,2×1 мм для ПЛИС Xilinx XC95108.

Заключение

Рассмотрена архитектура аппаратно-программной реализации устройства выделения первичных признаков для задач распознавания изображений. Рассмотрены теоретические основы цифрового нейронного кортекса.

Для аппаратной реализации устройства предлагается использовать программируемые логические схемы.

Литература

- Максфилд К. (2007) Проектирование на ПЛИС. Архитектура, средства и методы / Пер. с англ. М.: Изд. дом «Додэка-XXI», 2007. 408 с.
- Хьюбел Д. (1990) Глаз, мозг, зрение. М.: Мир, 1990. 240 с.

- Mikhailov A.* (2007) Digital neural cortex // Proc. Artificial Neural Networks in Engineering Conf. (ANNIE 2007). Nov. 11–4, 2007. St. Louis, Missouri, USA.
- Felleman D.J., Van Essen D. C.* (1991) Distributed Hierarchical Proc. Primate Cerebral Cortex. Cerebral Cortex 1:1–47.

HARD AND SOFT REALIZATION CORTEX SCANER

A. A. Domunian

*Institute of Control Sciences Russian Academy of Sciences named after
V.A. Trapeznikov, 117997 Moscow, 65 Profsoyuznaya st.*

This paper considered to create a cortex scanner, that is a universal hard-devices to feature extraction in stage of image preprocessing for future uses to image recognition. Cortex scanner modeling process handling information in primary visual area cerebral cortex. Paper consist, against, hard aspect creation this devices, found in FPGA connective with cameras.

Keywords: machine neural cortex, cortex scanner, image recognition, machine intelligence.

Domunian Andrey Aleksandrovich — Post-Graduate Student, e-mail: inventor@pisem.net.

ТОЧНОСТЬ ОЦЕНИВАНИЯ КООРДИНАТ РЕПЕРНЫХ ЗНАКОВ НА НЕПОДВИЖНОЙ ЦЕЛИ ПО ИЗМЕРЕНИЯМ БОРТОВОГО ОПТИЧЕСКОГО ДАТЧИКА

Б. Ц. Бахшиян, Р. Р. Назиров, К. С. Федяев

Учреждение Российской академии наук Институт космических исследований РАН (ИКИ РАН), 117997, Москва, Профсоюзная, 84/32

Излагается способ оценивания точности определения координат реперных знаков, расположенных на неподвижной цели, по результатам их наблюдений с помощью бортового оптического датчика. Точность оценивается при различных предположениях о коррелированности измерений координат каждого из реперных знаков. Рассматриваются следующие случаи: а) измерения некоррелированы; б) информация о корреляционной матрице отсутствует; в) существуют несколько групп измерений, причём измерения в различных группах некоррелированы, а внутри каждой группы возможна произвольная корреляция. Рассматривается также задача нахождения наилучшего по точности плана измерений при условии произвольной коррелированности измерений. Приводятся результаты практических расчётов.

Работа выполнена при поддержке РФФИ (проект № 09-08-00202-а).

Ключевые слова: точность оценивания, гарантирующий подход, корреляционная матрица, метод наименьших квадратов, оцениваемый параметр.

Введение

Задача оценивания точности определения координат реперных знаков на неподвижной цели по результатам бортовых измерений является одним из этапов процесса управления стыковкой объектов. В частности, эта задача решалась при создании алгоритмов обеспечения стыковки глубоководного спасательного аппарата с потерпевшей аварию подводной лодкой. При решении этой задачи требовалось оценить координаты реперных знаков на так называемой комингс-площадке (от англ. *comings*) подводной лодки в подвижной системе координат, связанной со спасательным аппаратом. Координаты реперных знаков определялись по результатам измерений с помощью оптического датчика и априорным значениям неизвестных параметров в используемой модели оценивания. Совокупность всех этих исходных данных будем называть *расширенным вектором измерений*.

Характеристики оценок искомых координат реперных знаков существенно зависят от предположений относительно дисперсий и матрицы корреляций расширенного вектора измерений. Традиционно считается, что указанная корреляционная матрица равна единичной, т. е. постулируется отсутствие корреляции между различными измерениями. Но на практике такое допущение может быть слишком грубым и привести к неоправданно оптимистической оценке дисперсий определяемых параметров.

Бахшиян Борис Цолакович — ведущий научный сотрудник, доктор физико-математических наук, e-mail: bbakhshiyam@gmail.com.

Назиров Равиль Равильевич — заместитель директора, доктор технических наук, профессор, e-mail: rnazirov@cosmos.ru.

Федяев Константин Сергеевич — младший научный сотрудник, кандидат физико-математических наук, e-mail: ksfedyayev@gmail.com.

В данной работе для вычисления характеристик точности используется *гарантирующий подход*. В этом подходе полагается известным лишь некоторое множество, которому принадлежат недиагональные элементы корреляционной матрицы расширенного вектора измерений, и находятся максимальные на этом множестве дисперсии контролируемых параметров. Гарантирующий подход дает менее оптимистические гарантированные значения дисперсий оценок по сравнению с традиционными подходами, но позволяет получить более реальные характеристики точности оцениваемых параметров.

Согласно теореме Маркова – Гаусса [Pao, 1968], при допущении о некоррелированности всех измерений минимальная дисперсия оценки любого параметра в линейной модели оценивания достигается при определении всех неизвестных параметров методом наименьших квадратов (МНК). Однако определение всех параметров часто приводит к вычислительным проблемам, связанным с обращением больших плохо обусловленных матриц. Альтернативой является оценивание лишь нескольких неизвестных параметров, а остальные полагаются равными своим априорным значениям. Такая стратегия может быть предпочтительнее при условии, что измерения коррелированы. В частности, оптимальная стратегия выбора «уточняемых» параметров и используемого состава измерений достаточно эффективно решается для случая, когда допускается произвольная корреляция (см. подразд. 4.2). Эта стратегия может быть также использована при других, более гибких допущениях о множестве возможных матриц корреляций (см., например, подразд. 4.3).

1. МОДЕЛЬ ОЦЕНИВАНИЯ

В рассматриваемой модели используются две системы координат (рис. 1):

- 1) неподвижная $OXYZ$, связанная с посадочной плоскостью;
- 2) подвижная $O'xyz$, связанная с оптическим датчиком.

Неподвижная и подвижная системы координат связываются с помощью трёх углов: угла курса γ , угла дифферента ψ и угла крена θ (рис. 2). Пусть A — матрица

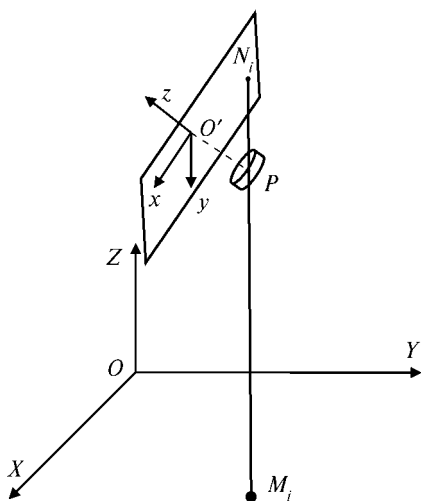


Рис. 1. Используемые системы координат

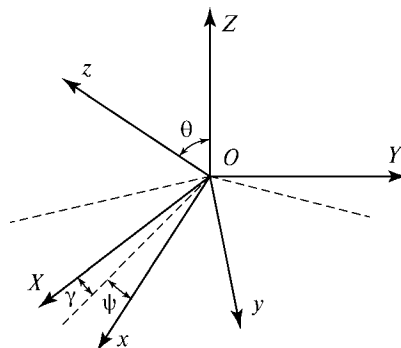


Рис. 2. Используемые угловые характеристики

поворота, переводящего неподвижную систему координат $OXYZ$ в систему координат Кенига с центром O и осями, параллельными осям $O'x, O'y, O'z$ подвижной системы координат. Матрица \mathbf{A} представляет собой произведение трёх матриц:

$$\mathbf{A} = \mathbf{A}_3(\gamma)\mathbf{A}_2(\psi)\mathbf{A}_1(\theta),$$

где

$$\mathbf{A}_3(\gamma) = \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{A}_2(\psi) = \begin{pmatrix} \cos \psi & 0 & -\sin \psi \\ 0 & 1 & 0 \\ \sin \psi & 0 & \cos \psi \end{pmatrix}, \mathbf{A}_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Обозначим через

$$\mathbf{r}(P) = \mathbf{f} = (f_x, f_y, f_z)' \quad (1)$$

радиус-вектор оптического центра P в подвижной системе координат, компоненты которого f_x, f_y, f_z могут считаться оцениваемыми параметрами.

Обозначим через M_i точку на плоскости мишени, соответствующую i -му реперному знаку, а через N_i — образ этой точки на матрице оптического датчика. Тогда точка M_i в неподвижной системе имеет координаты $\mathbf{R}(M_i) = (X(M_i), Y(M_i), 0)'$, а точка N_i в подвижной системе — координаты $\mathbf{r}(N_i) = (x(N_i), y(N_i), 0)'$. Модель измерений имеет следующий вид:

$$\begin{aligned} \tilde{x}(N_i) &= x(N_i) + \xi_{ix}, \\ \tilde{y}(N_i) &= y(N_i) + \xi_{iy}, \quad i = 1, \dots, s, \end{aligned} \quad (2)$$

где $\tilde{x}(N_i), \tilde{y}(N_i)$ — измеренные значения соответствующих координат точек N_i ; $i = 1, \dots, s$; ξ_{ix}, ξ_{iy} — ошибки измерений; s — число реперных знаков. По этим измерениям требуется найти координаты точек M_i в подвижной системе координат, т. е. величины $\mathbf{r}(M_i) = (x(M_i), y(M_i), z(M_i))'$.

Координаты точек M_i в подвижной и неподвижной системах связаны соотношением

$$\mathbf{r}(M_i) = \mathbf{r}(O) + \mathbf{A}'\mathbf{R}(M_i), \quad (3)$$

здесь $\mathbf{r}(M_i)$ — искомые величины; $\mathbf{R}(M_i)$ — известны; $\mathbf{A}' = \mathbf{A}'_1(\theta)\mathbf{A}'_2(\psi)\mathbf{A}'_3(\gamma)$; вектор $\mathbf{r}(O)$ находится из соотношения

$$\mathbf{f} = \mathbf{r}(P) = \mathbf{r}(O) + \mathbf{A}'\mathbf{R}(P),$$

где $\mathbf{R}(P)$ — координаты оптического центра в неподвижной системе.

Перепишем соотношение (3) в виде

$$\mathbf{r}(M_i) = \mathbf{f} - \mathbf{A}'(\mathbf{R}(P) - \mathbf{R}(M_i)). \quad (4)$$

Согласно (4), для нахождения оценок векторов $\mathbf{r}(M_i)$ достаточно по известным измерениям в модели (2) определить девятимерный вектор неизвестных параметров:

$$\mathbf{q} = (\mathbf{R}(P), \gamma, \psi, \theta, \mathbf{f})'. \quad (5)$$

Модель измерений (2) определяется $2s$ -координатами $x(N_i)$, $y(N_i)$ точек N_i . Преобразуем эту модель к виду, в котором измеряемые параметры будут функцией вектора \mathbf{q} . Для этого запишем очевидные соотношения, которые выражают условие принадлежности одной прямой точек N_i , P и M_i для каждого i :

$$\frac{\overline{PN_i}}{|PN_i|} = \frac{\overline{M_iP}}{|M_iP|}, i = 1, \dots, s. \quad (6)$$

Здесь $\overline{PN_i} = \mathbf{r}(N_i) - \mathbf{r}(P)$, $\overline{M_iP} = \mathbf{R}(P) - \mathbf{R}(M_i)$. Равенства (6) в подвижной системе координат переписуются в виде

$$\frac{\mathbf{r}(N_i) - \mathbf{r}(P)}{\|\mathbf{r}(N_i) - \mathbf{r}(P)\|} = \mathbf{g}_i = \begin{pmatrix} g_{ix} \\ g_{iy} \\ g_{iz} \end{pmatrix} = \mathbf{A}' \frac{\mathbf{R}(P) - \mathbf{R}(M_i)}{\|\mathbf{R}(P) - \mathbf{R}(M_i)\|}, \quad (7)$$

где вектор \mathbf{g}_i есть функция неизвестных параметров $\mathbf{R}(P)$, γ , ψ , θ . Учитывая, что $\|\mathbf{g}_i\| = 1$ и

$$\mathbf{r}(N_i) - \mathbf{r}(P) = \left(x(N_i) - f_x, y(N_i) - f_y, -f_z \right)',$$

нетрудно преобразовать модель (2) к виду

$$\begin{aligned} \tilde{x}(N_i) &= f_x - \frac{g_{ix}}{g_{iz}} f_z + \xi_{ix}, \\ \tilde{y}(N_i) &= f_y - \frac{g_{iy}}{g_{iz}} f_z + \xi_{iy}, \quad i = 1, \dots, s. \end{aligned} \quad (8)$$

Здесь i -я измеряемая функция зависит только от вектора \mathbf{q} неизвестных параметров и известного вектора $\mathbf{R}(M_i)$.

Кроме измерений (8) будем считать также измерениями известные заранее номинальные значения q_j^0 компонент вектора \mathbf{q} . Тогда модель (8) дополнится уравнениями

$$q_j^0 = q_j + \xi_{qj}, j = 1, \dots, 9. \quad (9)$$

2. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Пусть σ_{ix} , σ_{iy} , σ_{qj} — стандартные отклонения ошибок ξ_{ix} , ξ_{iy} , ξ_{qj} . Оценка $\hat{\mathbf{q}}$ вектора \mathbf{q} во взвешенном методе наименьших квадратов (МНК) находится из условия

$$\hat{\mathbf{q}} = \operatorname{argmin} \left\{ \sum_{i=1}^s \left[\left(\frac{\tilde{x}(N_i) - f_x + \frac{g_{ix}}{g_{iz}} f_z}{\sigma_{ix}} \right)^2 + \left(\frac{\tilde{y}(N_i) - f_y + \frac{g_{iy}}{g_{iz}} f_z}{\sigma_{iy}} \right)^2 \right] + \sum_{j=1}^9 \left(\frac{q_j^0 - q_j}{\sigma_{qj}} \right)^2 \right\}$$

и является решением системы нормальных уравнений

$$\sum_{i=1}^s \left[\left(\frac{\tilde{x}(N_i) - f_x + \frac{g_{ix}}{g_{iz}} f_z}{\sigma_{ix}} \right) \mathbf{H}'_{1i}(\mathbf{q}) + \left(\frac{\tilde{y}(N_i) - f_y + \frac{g_{iy}}{g_{iz}} f_z}{\sigma_{iy}} \right) \mathbf{H}'_{2i}(\mathbf{q}) \right] + \left(\frac{q_1^0 - q_1}{\sigma_{q_1}^2}, \dots, \frac{q_9^0 - q_9}{\sigma_{q_9}^2} \right) = 0, \quad (10)$$

где

$$\mathbf{H}'_{1i}(\mathbf{q}) = \frac{1}{\sigma_{ix}} \frac{\partial}{\partial \mathbf{q}} \left(f_x - \frac{g_{ix}}{g_{iz}} f_z \right), \quad \mathbf{H}'_{2i}(\mathbf{q}) = \frac{1}{\sigma_{iy}} \frac{\partial}{\partial \mathbf{q}} \left(f_y - \frac{g_{iy}}{g_{iz}} f_z \right)$$

— векторы-строки. Система (10) может быть решена методом последовательных приближений путём линеаризации модели измерений (8).

Пусть \mathbf{q}_0 — некоторое начальное значение вектора (5). Тогда линеаризованная модель измерений (8) будет иметь вид

$$\begin{aligned} \tilde{x}(N_i) - x(N_i) \Big|_{\mathbf{q}=\mathbf{q}^0} &= \sigma_{ix} \mathbf{H}'_{1i}(\mathbf{q}^0) (\mathbf{q} - \mathbf{q}^0) + \xi_{ix}, \\ \tilde{y}(N_i) - y(N_i) \Big|_{\mathbf{q}=\mathbf{q}^0} &= \sigma_{iy} \mathbf{H}'_{2i}(\mathbf{q}^0) (\mathbf{q} - \mathbf{q}^0) + \xi_{iy}, \quad i = 1, \dots, s, \\ 0 &= (q_j - q_j^0) + \xi_{q_j^0}, \quad j = 1, \dots, 9. \end{aligned} \quad (11)$$

Для модели (8), (9) оценка вектора \mathbf{q} определяется согласно известной формуле

$$\hat{\mathbf{q}} - \mathbf{q}^0 = \mathbf{Y}' \tilde{\mathbf{d}}, \quad (12)$$

где $\tilde{\mathbf{d}}$ — вектор размерности $(2s + 9)$, составленный из $2s$ нормированно-центрированных измерений и девяти нулевых значений центрированных априорных данных:

$$\frac{\tilde{x}(N_i) - x(N_i) \Big|_{q=q^0}}{\sigma_{ix}}, \quad \frac{\tilde{y}(N_i) - y(N_i) \Big|_{q=q^0}}{\sigma_{iy}}, \quad i = 1, \dots, s, \quad \frac{q_j^0 - q_j^0}{\sigma_j} = 0, \quad j = 1, \dots, 9;$$

матрица \mathbf{Y} определяется по формуле

$$\mathbf{Y} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-1}, \quad (13)$$

где \mathbf{H} — матрица размером $9 \times (2s + 9)$, составленная из столбцов $\mathbf{H}_{1i}(\mathbf{q}_0)$, $\mathbf{H}_{2i}(\mathbf{q}_0)$, $i = 1, \dots, s$ и столбцов $\mathbf{e}_j / \sigma_{q_j}$, $j = 1, \dots, 9$ (\mathbf{e}_j — столбец с единственной ненулевой компонентой 1 на j -м месте, который соответствует априорному измерению q_j^0). Тогда ковариационная матрица $\mathbf{K}_{\hat{\mathbf{q}}}$ вектора $\hat{\mathbf{q}}$ равна

$$\mathbf{K}_{\hat{\mathbf{q}}} = \mathbf{Y}' \mathbf{K}_{\tilde{\mathbf{d}}} \mathbf{Y}, \quad (14)$$

где $\mathbf{K}_{\tilde{\mathbf{d}}}$ — ковариационная матрица вектора нормированных измерений $\tilde{\mathbf{d}}$, которая является корреляционной матрицей исходного расширенного вектора измерений \mathbf{d} . В частности, если измерения некоррелированы, то $\mathbf{K}_{\tilde{\mathbf{d}}} = \mathbf{I}$ (\mathbf{I} — единичная матрица) и

$$\mathbf{K}_{\hat{\mathbf{q}}} = (\mathbf{H}\mathbf{H}')^{-1}. \quad (15)$$

3. ФОРМУЛЫ ДЛЯ ТОЧНОСТИ МНК-ОЦЕНОК В СЛУЧАЕ ИЗВЕСТНОЙ КОВАРИАЦИОННОЙ МАТРИЦЫ ОШИБОК

Будем рассматривать следующие контролируемые параметры $l_j, j = 1, \dots, 6$, дисперсии ошибок оценивания, которые мы будем вычислять для линеаризованной модели (11).

1–3. Вектор $\mathbf{r}(M_k)$ для некоторого k (параметры l_1, l_2, l_3).

4. Угол φ между плоскостью датчика и плоскостью комингс-площадки (параметр l_4).

5. Прицельный параметр $\rho_k = \sqrt{x(M_k)^2 + y(M_k)^2}$ (параметр l_5).

6. Расстояние от начала подвижной системы координат до k -го реперного знака $d_k = \sqrt{x(M_k)^2 + y(M_k)^2 + z(M_k)^2}$ (параметр l_6).

Ковариационная матрица вектора контролируемых параметров $\hat{\mathbf{I}}$ с компонентами l_1, \dots, l_6 определяется по формуле

$$\mathbf{K}_{\hat{\mathbf{I}}} = \frac{\partial \mathbf{l}}{\partial \mathbf{q}} \mathbf{K}_{\hat{\mathbf{q}}} \left(\frac{\partial \mathbf{l}}{\partial \mathbf{q}} \right)', \quad (16)$$

где

$$\frac{\partial \mathbf{l}}{\partial \mathbf{q}} = \left(\frac{\partial \mathbf{r}(M_k)}{\partial \mathbf{q}}, \frac{\partial \varphi}{\partial \mathbf{q}}, \frac{\partial \rho_k}{\partial \mathbf{q}}, \frac{\partial d_k}{\partial \mathbf{q}} \right)',$$

элементы матрицы $\frac{\partial \mathbf{r}(M_k)}{\partial \mathbf{q}}$ вычисляются по формулам, получаемым из (4) и (1):

$$\frac{\partial \mathbf{r}(M_k)}{\partial \mathbf{R}(P)} = -\mathbf{A}'; \quad \frac{\partial \mathbf{r}(M_k)}{\partial \gamma} = -\mathbf{A}'_1(\theta) \mathbf{A}'_2(\psi) \left(\frac{d\mathbf{A}_3(\gamma)}{d\gamma} \right)' (\mathbf{R}(P) - \mathbf{R}(M_k)), \quad \frac{\partial \mathbf{r}(M_k)}{\partial \mathbf{f}} = -\mathbf{I}.$$

Аналогично вычисляются производные от остальных компонент вектора \mathbf{I} .

4. ФОРМУЛЫ ДЛЯ ГАРАНТИРОВАННОЙ ТОЧНОСТИ ОЦЕНОК МНК В СЛУЧАЕ НЕПОЛНОГО ЗНАНИЯ КОРРЕЛЯЦИОННОЙ МАТРИЦЫ ОШИБОК

Пусть $l_j = \mathbf{b}_j' \mathbf{q}$ — один из контролируемых параметров, описанных в разд. 3. Тогда для линеаризованной модели оценка параметра l_j методом наименьших квадратов имеет вид

$$\hat{l}_j - l_j(\mathbf{q}^0) = \mathbf{X}_j' \tilde{\mathbf{d}},$$

где $\mathbf{X}_j' \doteq \mathbf{b}_j' \mathbf{Y}'$, $\mathbf{b}_j' \doteq \frac{\partial l_j}{\partial \mathbf{q}}$, а дисперсия ошибки оценивания —

$$D\hat{l}_j = \mathbf{X}_j' \mathbf{K}_{\hat{\mathbf{d}}} \mathbf{X}_j. \quad (17)$$

Величины $D\hat{l}_j, j = 1, \dots, 6$ представляют собой диагональные элементы матрицы $\mathbf{K}_{\hat{\mathbf{I}}}$ (см. формулу (16)). В зависимости от степени неопределённости знания матрицы $\mathbf{K}_{\hat{\mathbf{d}}}$ можно судить о точности оценок \hat{l}_j . Будем рассматривать следующие случаи.

4.1. Матрица корреляций — единичная (самый оптимистический случай)

Этот случай описывается формулами (13), (14). Исходя из (17), имеем

$$D\hat{l}_j \triangleq \left\{ D\hat{l}_j : \mathbf{K}_{\hat{a}} = \mathbf{I} \right\} = \sum_{i=1}^n X_{ji}^2. \quad (18)$$

4.2. Гарантированная точность при отсутствии сведений о матрице корреляций (самый пессимистический случай)

Найдем максимальное (гарантированное) значение дисперсии оценки контролируемого параметра l_j в (17) при условии, что коэффициенты корреляции могут принимать произвольные значения. Таким образом, допускается, что недиагональные элементы матрицы $\mathbf{K}_{\hat{a}}$ могут быть произвольными, т. е.

$$|K_{pq}| \leq 1 \quad \forall p \neq q. \quad (19)$$

При условии (19) гарантированное значение дисперсии равно

$$D\hat{l}_j \triangleq \max \left\{ D\hat{l}_j : (19) \right\} = \left(\sum_{i=1}^n |X_{ji}| \right)^2. \quad (20)$$

Это гарантированное значение для каждого контролируемого параметра l_j достигается при

$$K_{pq}^j = \text{sign}(X_{jp} X_{jq}), \quad p \neq q. \quad (21)$$

Замечание 1. Обозначим через $\mathbf{K}_{\hat{a}}^{1j}$ матрицу с элементами (21). Тогда дисперсия оценки любого параметра l_j , $i \neq j$, вычисленная для корреляционной матрицы $\mathbf{K}_{\hat{a}}^{1j}$ измерений, равна

$$D\hat{l}_i = \mathbf{X}_i' \mathbf{K}_{\hat{a}}^{1j} \mathbf{X}_i.$$

При этом легко показать, что коэффициент корреляции между любыми оценками \hat{l}_i и \hat{l}_k равен ± 1 .

4.3. Дисперсии оценок при отсутствии корреляции между измерениями и априорными данными и возможности произвольной корреляции между остальными исходными данными (умеренно пессимистический случай)

Такой случай менее пессимистичен, чем изложенный в подразд. 4.2. Разумно допустить, что априорные данные не коррелированы с измерениями, а внутри каждой группы возможна произвольная корреляция, т. е.

$$\begin{cases} K_{pq} = 0, & p = 1, \dots, 2s, \quad q = 2s + 1, \dots, 2s + 9, \\ |K_{pq}| \leq 1 & \text{в противном случае.} \end{cases} \quad (22)$$

При этих допущениях гарантированная дисперсия равна

$${}^{01}D\hat{l}_j \triangleq \max \left\{ D\hat{l}_j : (22) \right\} = \left(\sum_{i=1}^{2s} |X_{ji}| \right)^2 + \left(\sum_{i=2s+1}^{2s+9} |X_{ji}| \right)^2. \quad (23)$$

Указанное значение дисперсии достигается при следующих значениях элементов корреляционной матрицы (22):

$$\begin{cases} K_{pq} = 0, & p = 1, \dots, 2s, q = 2s + 1, \dots, 2s + 9, \\ K_{pq} = \text{sign}(X_{jp}X_{jq}) & \text{в противном случае.} \end{cases} \quad (24)$$

Обозначим через $\mathbf{K}_{\hat{\mathbf{a}}}^{01j}$ матрицу с элементами (24). Тогда ковариационную матрицу $\mathbf{K}_{\hat{\mathbf{a}}}^{01j}$ вектора $\hat{\mathbf{a}}$, соответствующую корреляционной матрице $\mathbf{K}_{\hat{\mathbf{a}}}^{01j}$, можно найти по формулам (14) и (16) при $\mathbf{K}_{\hat{\mathbf{a}}} = \mathbf{K}_{\hat{\mathbf{a}}}^{01j}$.

4.4. Об оптимизации стратегии оценивания

Рассмотрим задачу нахождения минимальной дисперсии оценки параметра l_j для случая 4.2. Минимум будем отыскивать на множестве всех линейных несмещенных алгоритмов оценивания, т. е. для всех векторов \mathbf{X}_j , удовлетворяющих условию несмещенности

$$\mathbf{H}\mathbf{X}_j = \mathbf{b}_j. \quad (25)$$

Взвешенный МНК с любой весовой матрицей принадлежит этому множеству алгоритмов. Оптимальная задача записывается в виде [Бахшиян и др., 1980].

$$\left({}^1D\hat{l}_j \right)^* \triangleq \min \left\{ {}^1D\hat{l}_j : (26) \right\}. \quad (26)$$

Решение задачи (26) имеет следующий вид [Бахшиян и др., 1980]. Оптимальный вектор \mathbf{X}_j^* содержит не более 9 ненулевых переменных. Пусть $X_{j,1}^*, \dots, X_{j,r}^*$ — компоненты, соответствующие измеренным значениям координат, а $X_{j,r+1}^*, \dots, X_{j,9}^*$ — компоненты, соответствующие априорным данным (нумерация условна). Тогда оптимальная стратегия оценивания для этого случая состоит в том, что $9 - r$ соответствующих априорных данных полагаются равными своим априорным значениям, а остальные s компонент вектора \mathbf{q} определяются однозначно по r измерениям.

Эта стратегия оптимальна для случая 4.2, но она может быть применена и для других моделей.

5. РЕЗУЛЬТАТЫ РАСЧЁТОВ

При проведении практических расчётов рассматривался случай, когда оптический центр P датчика находился на расстоянии 200 см от начала O неподвижной системы координат, плоскость матрицы оптического датчика была наклонена по дифференту к плоскости мишени под углом $\varphi = \psi = \pi/12$ рад, оптическая ось дат-

чика была направлена в начало координат неподвижной системы. Рассматривалось пять реперных знаков, их расположение показано на рис. 3. Координаты реперных знаков M_1, \dots, M_5 в неподвижной системе и координаты соответствующих точек N_1, \dots, N_5 на плоскости матрицы оптического датчика приведены в табл. 1 и 2 (здесь и далее все линейные размеры указаны в сантиметрах, все угловые — в радианах).

В качестве априорных выбирались следующие значения компонент вектора \mathbf{q} :

$$\mathbf{R}(P) = \begin{pmatrix} 52 \\ 0 \\ 193 \end{pmatrix}, \gamma = 0, \psi = \pi/12, \theta = 0,$$

$$f_x = f_y = 0, f_z = 0,5.$$

Стандартные отклонения этих априорных значений принимались следующими:

$$\sigma(\mathbf{R}(P)) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \sigma(\gamma) = \sigma(\psi) = \sigma(\theta) = \pi/1800, \sigma(f_x) = \sigma(f_y) = \sigma(f_z) = 10^{-4}.$$

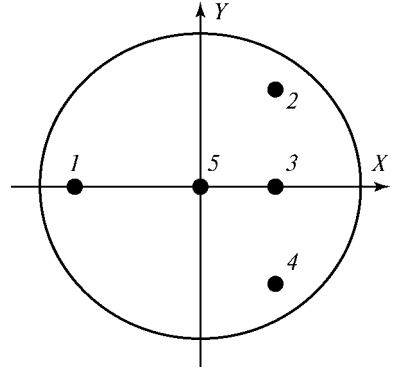


Рис. 3. Схема расположения реперных знаков на мишени

Таблица 1
Координаты реперных знаков в неподвижной системе координат

	M_1	M_2	M_3	M_4	M_5
X	-50	30	30	30	0
Y	0	40	0	-40	0
Z	0	0	0	0	0

Таблица 2
Измеренные координаты образов реперных знаков в подвижной системе координат

	N_1	N_2	N_3	N_4	N_5
x	0,113	-0,075	-0,075	-0,075	0
y	0	-0,104	0	0,104	0
z	0	0	0	0	0

Приведём результаты вычислений при различных допущениях, рассмотренных в разд. 4.

5.1. Характеристики точности оценок вектора \mathbf{q} и вектора \mathbf{l} контролируемых параметров в случае некоррелированных измерений

При допущении об отсутствии корреляции между измерениями получены следующие значения стандартных отклонений вектора $\hat{\mathbf{q}}$ и его корреляционной матрицы (см. формулу (15)):

$$\sigma(\hat{\mathbf{q}}) = (0,31; 0,33; 0,28; 0,0011; 0,0016; 0,0016; 0,0001; 0,0001; 0,0001).$$

$${}^0\mathbf{K}_{\hat{\mathbf{q}}} = \begin{pmatrix} 1 & 0 & -0,234 & 0 & 0,963 & 0 & 0,101 & 0 & 0,033 \\ 0 & 1 & 0 & 0,036 & 0 & -0,97 & 0 & 0,104 & 0 \\ -0,234 & 0 & 1 & 0 & -0,395 & 0 & -0,047 & 0 & 0,125 \\ 0 & 0,036 & 0 & 1 & 0 & -0,175 & 0 & 0,016 & 0 \\ 0,963 & 0 & -0,395 & 0 & 1 & 0 & -0,019 & 0 & 0,003 \\ 0 & -0,97 & 0 & -0,175 & 0 & 1 & 0 & 0,014 & 0 \\ 0,101 & 0 & -0,047 & 0 & -0,019 & 0 & 1 & 0 & 0 \\ 0 & 0,104 & 0 & 0,016 & 0 & 0,014 & 0 & 1 & 0 \\ 0,033 & 0 & 0,125 & 0 & 0,003 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Как видно, наибольшую корреляцию (более 0,96 по модулю) имеют оценки параметров $x(P)$ и ψ , $y(P)$ и θ .

Значения стандартных отклонений компонент вектора $\hat{\mathbf{I}}$ для данного случая приведены во втором столбце табл. 3. Корреляционная матрица этого вектора равна

$${}^0\mathbf{K}_{\hat{\mathbf{I}}} = \begin{pmatrix} 1 & 0 & -0,6 & -0,99 & -1 & 0,6 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -0,6 & 0 & 1 & 0,58 & 0,6 & -1 \\ -0,99 & 0 & 0,58 & 1 & 0,99 & -0,58 \\ -1 & 0 & 0,6 & 0,99 & 1 & -0,6 \\ 0,6 & 0 & -1 & -0,58 & -0,6 & 1 \end{pmatrix}.$$

Таблица 3

Стандартные отклонения компонент вектора оцениваемых параметров

Параметр	${}^0\sigma$	${}^1\sigma$	${}^{01}\sigma$	${}^1\sigma^*$
$x(M_5)$	0,59	1,26	0,93	0,81
$y(M_5)$	0,64	1,25	0,93	1,04
$z(M_5)$	0,32	0,93	0,68	0,72
$\cos(\varphi), \times 10^{-3}$	1,6	3,1	2,3	1,7
ρ	0,59	1,26	0,93	0,81
d	0,32	0,93	0,68	0,72

Так как корреляция между оценками параметров l_1 и l_5 , l_3 и l_6 равна -1 , то число контролируемых параметров в данном случае может быть уменьшено. Отметим также сильную корреляцию между параметрами l_1 и l_4 , l_4 и l_5 .

5.2. Гарантированные характеристики точности оценок вектора $\hat{\mathbf{I}}$ контролируемых параметров

В табл. 3 приводятся результаты, полученные при различных предположениях о коррелированности измерений (см. подразд. 4.2–4.4), в таблице также представлены

стандартные отклонения компонент вектора $\hat{\mathbf{1}}$ при различной информации о коррелированности измерений. При этом в случаях 4.2 и 4.3 для каждого параметра приведено гарантированное значение его стандартного отклонения. В последнем столбце приведены оптимальные значения стандартных отклонений в случае 4.2.

Заключение

Задача определения координат реперных знаков по измерениям бортового оптического датчика в рассматриваемой постановке сведена в работе к задаче оценивания вектора из 9 неизвестных параметров по 10 скалярным измерениям. Приведенные результаты расчётов характеристик точности для случая некоррелированных измерений и соответствующих гарантированных характеристик показывают возможность достаточно точного оценивания координат цели. Как показывают результаты, приведенные в табл. 3, отсутствие информации о коррелированности измерений не приводит к существенному ухудшению полученных характеристик точности.

Литература

- Бахшиян Б. Ц., Назиров Р. Р., Эльясберг П. Е. (1980) Определение и коррекция движения. М.: Наука, 1980. 360 с.
- Рао С. Р. (1968) Линейные статистические методы. М.: Наука, 1968. 548 с.

ACCURACY ESTIMATE OF BENCH MARKS COORDINATES DETERMINATION BY ONBOARD OPTICAL SENSOR OBSERVATIONS

B. Ts. Bakhshiyani, R. R. Nazirov, K. S. Fedyaev

Space Research Institute, 117997, Moscow, 84 / 32 Profsoyuznaya st.

Method of accuracy estimate of bench marks coordinates determination by onboard optical sensor observations is presented. We suppose that bench marks are situated at the stationary target. Accuracy is estimated under various assumptions about correlations of the observations. The following cases are discussed: a) the observations are uncorrelated; b) there is no information about the correlation matrix; c) there are several uncorrelated groups of observations, but inside each group observations may have an arbitrary correlation. We also discuss the problem of finding an observation plan that has the best accuracy in case of arbitrary observations correlation. Calculation results are given.

Keywords: accuracy estimate, guaranteed approach, correlation matrix, the least square method, estimated parameter.

Bakhshiyani Boris Tsolakovich — Leading Researcher, Doctor of Physical and Mathematical Sciences, e-mail: bbakhshiyani@gmail.com.

Nazirov Ravil Ravilyevich — Deputy Director, Doctor of Technical Science, Professor, e-mail: rnazirov@cosmos.ru.

Fedyaev Konstantin Sergeevich — Junior Researcher, Candidate of Physical and Mathematical Science, e-mail: ksfedyaev@gmail.com.

Научное издание

Техническое зрение в системах управления мобильными объектами-2010
Труды научно-технической конференции-семинара
Выпуск 4
Таруса, 16–18 марта 2010 года

Под редакцией Р. Р. Назирова

Заведующая редакцией *Игнатова Е. С.*
Ведущий редактор *Климкин М. С.*
Корректор *Федотова А. Л.*
Художник *Терехова Г. Д.*

Директор издательства *Чепыжов В. В.*

Подписано в печать 17.01.2011.
Формат 70×100/16. Бумага офсетная.
Гарнитура Newton. Печать цифровая.
Усл. печ. л. 26,65. Тираж 55 экз. Заказ № Т-009

ООО «Издательство «КДУ», 119234, а/я 587.
Тел./факс: (495) 939-57-32, 939-44-91.
E-mail: kdu@kdu.ru; <http://www.kdu.ru>

Отпечатано в типографии КДУ
Тел./факс: (495) 939-44-91.
E-mail: press@kdu.ru