

На правах рукописи



Хамухин Анатолий Владимирович

**Высокоэффективные алгоритмы семантической обработки
видеоизображений и управления приборными комплексами
технического зрения**

05.13.01 – «Системный анализ, управление и обработка информации»
(информационно-вычислительное обеспечение)

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
доктора технических наук

Москва — 2016

Работа выполнена в акционерном обществе «ЭЛВИС(Электронные вычислительно-информационные системы)-Неотек» в отделе видеоаналитики.

Научный консультант: ПЕТРИЧКОВИЧ Ярослав Ярославович
доктор технических наук, профессор

Официальные оппоненты: БАРСКИЙ Аркадий Бенционович,
доктор технических наук, профессор кафедры
«Вычислительные системы и сети» Московско-
го государственного университета путей сообще-
ния

МАТВЕЕВ Юрий Николаевич,
доктор технических наук, профессор кафедры
электронных вычислительных машин Тверско-
го государственного политехнического универ-
ситета

УМНЯШКИН Сергей Владимирович,
доктор физико-математических наук, профес-
сор кафедры высшей математики №1 на-
ционального исследовательского университета
«Московский институт электронной техники»

Ведущая организация: Московский научно-исследовательский телеви-
зионный институт

Зашита состоится «___» _____ 2017 г. в 11:00 на заседании диссертаци-
онного совета Д 002.073.04 при Федеральном государственном учреждении
«Федеральный исследовательский центр «Информатика и управление» Рос-
сийской Академии Наук» (ФИЦ ИУ РАН) по адресу: 117312, Москва, про-
спект 60-летия Октября, 9 (конференц-зал, 1-й этаж).

С диссертацией можно ознакомиться в библиотеке ФИЦ ИУ РАН по адресу:
119333, Москва, ул. Вавилова д.40.

Электронные версии диссертации и автореферата размещены на официаль-
ном сайте ФИЦ ИУ РАН <http://www.frccsc.ru>.

Электронная версия автореферата размещена на официальном сайте ВАК
Министерства образования и науки РФ по адресу: <http://vak.ed.gov.ru>.

Отзывы и замечания по автореферату в двух экземплярах, заверенные пе-
чатью, просьба высыпать по адресу: 117312, Москва, проспект 60-летия
Октября, 9, ФИЦ ИУ РАН, диссертационный совет Д 002.073.04.

Автореферат разослан «___» _____ 20___ года.

Телефон для справок: +7 (499) 135-51-64.

Учёный секретарь
диссертационного совета Д 002.073.04
д.т.н., профессор

В.Н. Крутько

Общая характеристика работы

Актуальность проблемы. По мере развития вычислительной техники системы технического зрения используются всё шире, охватывая многие области деятельности человека: от ставшей уже обыденной уборки помещений автоматами до автоматического контроля изменений ландшафтов при слежении из космоса. Системы технического зрения всё чаще применяют в автопилотах для автомобилей, для контроля качества конвейерного производства и даже в устройствах, предназначенных для игр и развлечений: при кажущейся несерьёзности последнего массовая реализация подобных устройств может дать очень значительный экономический эффект для предприятия-производителя в частности и для страны-производителя в целом. Тем не менее, наиболее значимым и актуальным применением систем технического зрения, на наш взгляд, остаются охранные системы, предназначенные для предотвращения неправомерных действий. Современные технические средства, предназначенные для охраны, улучшают такие важные показатели, как своевременная реакция на нарушение, количество задействованного в охранном комплексе персонала, трудоёмкость принятия решений при выполнении необходимых действий в ответ на нарушение. В последнее время проблема эффективной охраны порядка стала особенно актуальной в связи с повышением активности террористических организаций, поэтому задача совершенствования охранных приборов является насущной.

Современные системы технического зрения обладают двумя основными недостатками. Во-первых, на операторов, в обязанности которых входит следить за происходящим на изображении, налагается высокая нагрузка, и на практике её можно выдержать только в течение 2–3 часов, даже если камера наблюдения всего одна. Во-вторых, поиск события в видеоархиве, если время события известно не точно, может быть проведён лишь при помощи просмотра архива целиком. Часто видеорегистраторы снабжаются алгоритмами определения движения в кадре, что несколько повышает эффективность использования системы видеонаблюдения. Однако простой детектор изменения изображения бесполезен в условиях интенсивного движения, например, на транспортных магистралях или многолюдных улицах.

К текущему моменту создано множество разработок в области семантической обработки цифровых изображений, которые используются для решения проблем технического зрения. Работы Ярославского Л.П., Грузмана И.С., Киричука В.С., Косых В.П., Перетягина Г.И., Кузнецова А.Е, Спектора А.А, У. Претта, Ханта Б.Р. и многих других учёных позволяют улучшить качество изображений и сигналов с помощью первичной фильтрации для повышения эффективности основной семантической обработки и для увеличения информативности данных. В работах Алпатова Б.А, Блохина А.Н., Бабая-

на П.В., Степашкина А.И., а также в работах зарубежных учёных Т. Канаде, Р. Колинза, К. Стоуфера, В. Гrimсoнa, Д. Коминисью, Дж. Малика и других предложены различные алгоритмы выделения положения движущихся объектов в области изображений при обработке потока сигналов от видеоприборов. Подобные алгоритмы могут быть использованы для привлечения внимания оператора видеосистем к изображению видеокамеры, если на изображении происходит какое-то событие. Ещё большее влияние на развитие систем технического зрения оказали достижения в области машинного обучения и информационной семантической обработки, классификации и распознавании видеосигналов, которые основаны на теории, изложенной в работах Харкевича А.А., Ковалевского В.А., Ф. Розенблата, Бравермана Э.М., Айзermana M.A., Розоноэра Л.А. и получившей дальнейшее развитие в работах Галушкина А.И., Ивахненко А.Г., Вапника В.Н., Червоненкиса А.Я., Журавлёва Ю.И., Злобина В.К., Р. Дуда, Д. Сторка, Р. Харта и многих других. В теории распознавания образов существенно используется теория статистических решений, в развитие которой основополагающий вклад внесли Колмогоров А.Н., Дж. Нейман, К. Пирсен. Семантическая классификация изображений позволяет создавать приборы обработки, которые не только реагируют на произвольное изменение в области изображения видеокамер, но и анализируют тип изменений. В итоге внимание операторов систем технического зрения привлекается не к каждому изменению в области наблюдения, а только к тем, которые действительно представляют интерес с позиции решаемых аналитической видеосистемой задач. В связи с проблемой распознавания и классификации изображений следует отметить большие практические успехи, которые в последние несколько лет были достигнуты при применении аппарата так называемых свёрточных нейронных сетей в задачах машинного обучения: достоверность классификации изображений, полученная с помощью нейронных сетей в данных работах, сопоставима с качеством классификации, выполняемой человеком. Однако алгоритмическая сложность глубоких свёрточных нейронных сетей настолько высока, что на данный момент их невозможно полноценно применять в системах реального времени, поскольку для этого нет достаточно мощных вычислительных устройств.

Процесс развития систем технического зрения неотделим от процесса развития техники, и качество современных систем определяется характеристиками оптических приборов, эффективностью обработки сигналов и быстродействием вычислительных устройств. Большой вклад в развитие элементной базы обработки сигналов, вычислительной техники и в теорию построения и разработки алгоритмов внесли Котельников В.А., К. Шенон, Гуляев Ю.В., Глушков В.М., Валиев К.А., Гуськов Г.Я., Преснухин Л.Н., Стемпиковский А.Л., Филатов В.Н., Красников Г.Я., Бетелин В.Б., Немудров В.Г.,

Шахнов В.А., Корячко В.П., Бархоткин В.А, Петричкович Я.Я. и многие другие. Тем не менее, проблема быстрой обработки изображений до сих пор актуальна, поскольку многие теоретические разработки в области семантической обработки видеосигналов применимы лишь в режиме постобработки, а не в реальном времени, поскольку обработка одного кадра видеопоследовательности занимает слишком длительное время даже самыми современными вычислительными средствами.

В последнее время системы видеонаблюдения строятся по принципу объединения в единую сеть различных датчиков, регистрирующих обстановку в зоне наблюдения. Заметим, что датчики — это не только видеокамеры, но и, например, датчики объёма в помещениях, датчики открытия/закрытия дверей, датчики поворота турникетов и т.п. В том числе датчиками комплексов технического зрения являются видеоприборы: стационарные камеры и камеры с возможностью управления положением оптической оси и величиной оптической силы объектива. Цель объединения видеокамер в единую систему заключается в том, чтобы выдавать общую информацию о ситуации в зоне наблюдения в виде единого целого, а не как не связанные друг с другом изображения с видеокамер. Основное назначение подобной системы — максимально упростить анализ информации оператором за счёт первичной обработки и оптимизированного с точки зрения эргономики вывода данных о событиях, происходящих в зоне действия видеоприборов системы. Связь видеокамер между собой требует настройки ряда параметров, часть из которых нужно ввести вручную, а часть автоматически. Чем более прост и более автоматизирован процесс предварительной настройки, тем экономически более оправдано применение видеосистемы.

Таким образом, основная проблематика в области систем технического зрения заключается в трудоёмкости их настройки и эксплуатации. Несмотря на значительные усовершенствования, достигнутые благодаря развитию семантической обработки изображений, многие **проблемы** остаются не решёнными:

- операторы систем видеонаблюдения вынуждены постоянно следить за массивом изображений, поэтому из-за человеческого фактора велика вероятность пропуска событий в зоне наблюдения;

- для упрощения анализа информации, поступающей от приборов в информационных комплексах, используют связь видеокамер между собой, при этом возникает проблема оптимизации времени предварительной юстировки сопряжения перед вводом системы в эксплуатацию;

- необходимость увеличения разрешающей способности изображений, получаемых от неподвижных камер в составе видеосистем, требует значительного дополнительного количества неподвижных камер, поскольку зависи-

мость между требуемым линейным увеличением разрешения и количеством камер квадратичная, естественно, что повышение разрешающей способности путём установки дополнительных неподвижных камер приводит к повышению сложности комплексов и увеличению затрат;

— в ряде приборных комплексов технического зрения есть возможность управления поворотными видеокамерами, и с их помощью можно получить изображение с более высокой разрешающей способностью, чем на неподвижной камере, причём сектор обзора поворотных камер достигает 360° , однако для управления поворотными камерами в существующих системах требуется непосредственное участие оператора, что повышает вероятность пропуска событий в зоне наблюдения, кроме того, многие поворотные видеокамеры могут осуществлять сканирование области контроля в автоматическом режиме, но при этом эффективных алгоритмов семантической обработки панорамных изображений реального времени пока не существует;

— использование стереокамер потенциально может существенно увеличить достоверность идентификации объектов в приборных комплексах технического зрения, однако многие вопросы теории использования и обработки стереоизображений остаются также не решёнными.

Цель работы. Цель данной диссертации — повысить эффективность контроля зоны наблюдения с помощью комплексов технического зрения нового поколения за счёт разработки и практической реализации способов автоматического сопряжения и юстировки изображений видеокамер, создания и анализа алгоритмов синхронизации процессов семантической обработки информационных сигналов от сети видеоприборов и автоматического управления поворотными камерами при наведении на объекты в зоне наблюдения с последующим сопровождением или в режиме непрерывного сканирования.

Задачи работы. Для достижения поставленной цели необходимо решить следующие задачи:

1. Провести функциональный анализ процессов семантической обработки видеосигналов и разработать описание модели распределённой системы коллективной обработки изображений различными приборами технического зрения.

2. Вывести соотношения, определяющие параметры сопряжения приборов технического зрения, и создать на их основе алгоритмы сопоставления объектов в общей зоне контроля приборов и алгоритмы наведения поворотных камер на сопровождаемые объекты с адаптивным упреждением.

3. Разработать методику тестирования и контроля алгоритмов семантической обработки изображений.

4. Разработать алгоритмы семантического анализа стереоизображений и оценить эффективность их применения в системах технического зрения.

5. Разработать алгоритмы семантической обработки панорамных изображений реального времени для видеоприборов, установленных на поворотной платформе, которая движется в режиме непрерывного сканирования.

Методы исследований. Задачи диссертационной работы решены с применением теории сложности алгоритмов, методов обработки цифровых сигналов-изображений, теории вычислительных методов оптимизации, фактов проективной и сферической геометрии, теории вероятности и теории статистических решений.

Научная новизна состоит в разработке, анализе и применении в приборных комплексах и в системах технического зрения эффективных алгоритмов синхронизации процессов семантической обработки изображений и управления поворотными видеокамерами, а также в создании способов автоматизации юстировки параметров сопряжения приборов технического зрения. В работе развивается концепция систем видеонаблюдения, которая строится на совместной семантической обработке информации с различных видеокамер и на применении элементов теории распознавания образов. В процессе выполнения диссертационной работы впервые получены следующие новые научные результаты.

1. Впервые разработана теория параллельных вычислений для комплекса программ, осуществляющих коллективную обработку информации, поступающей от приборов технического зрения. Для данной модели создан новый сетевой протокол синхронизации данных, определены ограничения на количество сопряжённых приборов в сети комплексов технического зрения, а также сформулирована и решена проблема полуавтоматической юстировки начальных параметров сопряжения.

2. Выведены математические соотношения для сопряжения поворотных камер и приборов технического зрения, на основе которых созданы новые алгоритмы сопоставления изображений одного и того же объекта в поле зрения сопряжённых неподвижных видеокамер и алгоритмы автоматического наведения и сопровождения объектов поворотными камерами для получения изображений большего разрешения на основе адаптивного упреждения.

3. Разработана теория тестирования и контроля алгоритмов семантической обработки на основе формальной логики и математического аппарата теории множеств.

4. Предложен разработанный автором новый комбинированный алгоритм обработки моно- и стереоизображений, отслеживающий перемещение объектов с высокой достоверностью.

5. Создан новый быстродействующий параллельно-конвейерный алгоритм построения и анализа панорамных изображений в реальном времени для ви-

деоприборов, установленных на поворотном устройстве и непрерывно сканирующих область контроля.

Практическая значимость. Разработанные в диссертации алгоритмы, способы и модели используются в следующих серийно изготавливаемых системах и приборных комплексах технического зрения: системы видеонаблюдения с компьютерным зрением «Orwell2k» РАЯЖ 46652.001-ОС.ПЗ, ЛЦКБ.464412.002, ВАРШ.00001-01, ВАРШ.00003-01 (патенты РФ на полезные модели №36315 от 07.08.2003 и №36912 от 23.06.2003, патенты РФ №2265531 от 07.08.2003 и №2268497 от 23.06.2003); видеодетектор «Orwell2k-Barrier», определяющий объекты в состоянии свободного полёта; система подсчета зрителей в кинозалах «Orwell2k-Cinema» (патент РФ на полезную модель №47546 и на изобретение №2296434 от 14.05.2005); приборы подсчёта объектов «Statistics» ВАРШ.468196.001 и «Statistics-3d» ВАРШ.463135.002-01; системы контроля дорожного движения «Travio» ВАРШ.466452.001, ВАРШ.466452.002 (патенты РФ №2014104763, №2015102323, патенты РФ на полезную модель №2014104762, №2015102309, №2015102315). Вышеперечисленные комплексы разработаны при непосредственном участии автора, таким образом, тема диссертации полностью соответствует направлению научно-производственной деятельности организации, в которой выполнена работа.

Программное обеспечение семейства «Orwell2k», реализованное на основе разработанных в диссертации алгоритмов сопряжения, управления и юстировки поворотными видеокамерами, поставляется на внутренний и зарубежный рынок и является конкурентоспособным продуктом на мировом уровне. **Преимущества** «Orwell2k» заключаются: в работоспособности при тяжёлых климатических и механических условиях эксплуатации; в автоматическом наведении поворотных камер на движущиеся объекты с установкой высокой разрешающей способности, при этом погрешность наведения на подвижные объекты составляет 3° , что при типичном угле зрения неподвижных камер 30° – 60° позволяет улучшить разрешающую способность на порядок по сравнению с исходной разрешающей способностью изображения от неподвижных видеокамер; в выводе событий на топографический план с локализацией места события; в применении критерия идентичности изображений объекта в поле зрения пересекающихся камер с достоверностью сопоставления 99,0%; в классификации всех объектов по типу «человек/машина/группа людей»; в существенном сокращении времени реакции системы «Orwell2k» на новые объекты, появляющиеся в зоне наблюдения, до 0,5с, что более чем на порядок превышает показатели существующих систем; в повышении на порядок точности подсчёта числа объектов за счёт анализа стерео-изображений по сравнению с существующим принципом обработки моноизображений и

уменьшению погрешности подсчёта до 3%; в применении анализа панорамных изображений, что позволяет в 10 раз и более сократить количество камер, необходимых для контроля зон наблюдения.

Автор диссертации проводил исследования и разработки в рамках «Приоритетных направлений развития науки, технологий и техники РФ» и «Критических технологий РФ», утверждённых указом Президента РФ №899 от 7 июля 2011 года.

Достоверность результатов подтверждается хорошей сходимостью результатов моделирования основных параметров функционирования приборных комплексов технического зрения, разработанных с участием автора, со значениями параметров серийных образцов, полученными в ходе их промышленной эксплуатации на различных объектах.

Внедрение результатов. Результаты диссертационной работы внедрены в системах технического зрения «Orwell2k» (ВАРШ.00001-01, ВАРШ.00003-01, ЛЦКБ.464412.002), которые используются в целях обеспечения безопасности как в России, так и за рубежом, в системах контроля безопасности дорожного движения «Travio» (ВАРШ.466452.001, ВАРШ.466452.002), в приборах подсчёта объектов «Statistics» (ВАРШ.468196.001) и «Statistics-3d» (ВАРШ.463135.002-01). Данные приборные комплексы используются для охраны стратегически важных объектов, таких как аэропорты, железные дороги, ГЭС, АЭС. Некоторые из перечисленных систем применяют для мониторинга городских общественных территорий в целях обеспечения безопасности. Приборы подсчёта объектов широко используются для управления хозяйственной деятельностью предприятий, предоставляющих услуги населению. Применение систем подтверждено актами о внедрении и протоколами испытаний, проводимых на объектах внедрения.

Личный вклад. Автором выявлены проблемы повышения эффективности комплексов технического зрения и сформулированы задачи по их решению. Все материалы, изложенные в диссертации, а именно: научные положения, технические решения, математические модели, семантические алгоритмы и алгоритмы управления, методики и результаты испытаний опытных образцов — получены автором лично. Автор создал архитектуру программного обеспечения, осуществляющего семантическую обработку, и участвовал в разработке программ и комплексов технического зрения, в которых внедрены результаты диссертации. Работы, выполненные в соавторстве, посвящены конструкторскому моделированию, производственным и эксплуатационным вопросам.

На защиту выносится:

1. Впервые предложенная теория параллельного функционирования се-

мантических обработчиков в комплексах технического зрения и способы полуавтоматической юстировки параметров сопряжения видеоприборов, позволяющие посредством синхронизации информации об анализируемых изображениях объединять в сети сопряжённые приборы технического зрения.

2. Новые алгоритмы сопряжения поворотных и неподвижных видеокамер, обеспечивающие достоверность сопоставления изображений в общей зоне контроля приборов 99,0%, а также автоматическое наведение поворотных камер с малыми погрешностями $0,5^\circ$ на неподвижные объекты и 3° на подвижные объекты, позволяющие получать изображения на порядок большего разрешения, причём время автоматического наведения в 20 раз меньше времени, требуемого для ручного наведения.

3. Теория тестирования и контроля алгоритмов семантической обработки, на основе которой создана распределённая система контроля и тестирования «Супервизор», использованная в качестве программно-инструментального средства, позволившего в 20 раз сократить частоту ложных идентификаций во время анализа потока видеоизображений при неизменном уровне ошибок второго рода.

4. Новый комбинированный алгоритм семантического анализа стерео- и моно- изображений, позволяющий вести подсчёт объектов с погрешностью 3%.

5. Впервые созданный алгоритм построения и семантической обработки панорамных изображений с частотой 25 кадров в секунду в реальном времени.

6. Внедрение результатов диссертации при разработке приборных комплексов технического зрения нового поколения «Orwell2k», в которых реализованы разработанные в диссертации алгоритмы и способы (свидетельство о регистрации программы №2003612604 от 28.11.2003, патенты РФ на полезные модели №36315 от 07.08.2003, №36912 от 23.06.2003, патенты РФ №2265531 от 07.08.2003, №2268497 от 23.06.2003) и которые внедрены, как уже отмечалось, на ряде важнейших объектов.

Апробация работы. Результаты диссертации докладывались на XLVI, XLVII и XLVIII научной конференции Московского физико-технического института, а также на XV, XVI и XVII конференциях молодых ученых, аспирантов и студентов по современным проблемам машиноведения в институте машиноведения им. А.А. Благонравова РАН. Информационные приборы и видеосистемы семейства «Orwell2k», в которых внедрены результаты работы, демонстрировались на 20 выставках. Алгоритмы и информационные технологии, разработанные автором диссертации и реализованные в системе, получили признание специалистов и отмечены соответствующими дипломами. В

2005 году видеосистема с компьютерным зрением «Orwell2k» удостоена II Национальной премии по безопасности «За укрепление безопасности России».

Публикации. Основное содержание диссертации отражено в 70 опубликованных работах, в том числе в 29 статьях в журналах, входящих в перечень, утвержденный ВАК. Без соавторов опубликовано 36 статей. В соавторстве получены семь патентов на изобретения, шесть свидетельств на полезную модель и шесть свидетельств о регистрации программ.

Структура и объём диссертации. Диссертация состоит из введения, семи глав, заключения, списка литературы и приложений. Работа содержит 250 страниц основного текста, 50 страниц с рисунками и таблицами, список литературы из 215 наименований и приложений на 44 страницах.

Основное содержание работы

Во **введении** обосновывается актуальность темы диссертации, формулируются основные цели и задачи работы, научная новизна, практическое значение полученных результатов и их внедрение, личный вклад автора, результаты, выносимые на защиту, а также описываются апробация и публикации по содержанию работы, структура диссертации и краткое содержание глав.

В **первой главе** вводится понятие семантического сжатия последовательности изображений и проводится обзор алгоритмов семантического сжатия цифровых видеосигналов. Семантическое сжатие определяется как операция преобразования множества изображений в информационное описание значительно меньшего объёма, чем исходный объём, с обеспечением сохранения в сжатом описании элементов, достаточных для решения поставленной задачи перед приборными комплексами технического зрения. Практически, в комплексах технического зрения оператору требуется небольшой объём информации о происходящем в поле действия приборов. Например, для описания изображения сцены дорожно-транспортного происшествия, состоящего из 10^6 точек (несколько миллионов байт), во многих случаях достаточно текстовой информации «произошло дорожно-транспортное происшествие» (десяток байт по порядку величины), или просто информация о том, что есть событие, на которое нужно обратить внимание (один бит информации), что потенциально сокращает объём информации об изображении на 7 порядков. Подобное семантическое сжатие применяется в системах технического зрения с большим количеством приборов (видеокамер, датчиков объёма, радиолокационных станций, индукционных датчиков заграждений и т.д.) для того, чтобы упростить процесс контроля и обработки данных операторами системы, и уменьшить трудоёмкость процесса контроля.

Современные аналитические комплексы технического зрения состоят из камер видеонаблюдения, подключаемых к ЭВМ, которые обрабатывают и

анализируют поступающий видеосигнал с целью автоматического выделения событий, требующих внимания операторов, находящихся на автоматизированных рабочих местах, оборудованных терминалами видеосистемы. Обработка видеосигнала является многоэтапным процессом. Первичная обработка заключается в повышении таких характеристик изображения, как, например, контрастность или соотношение сигнал/шум. Улучшение качества изображения может быть реализовано на основе фильтра Винера-Хопфа, на основе применения статистических фильтров, использующих Марковские случайные поля для повышения качества искажённого изображения по аналогии с восстановлением ориентации диполей в модели Изинга, а также на основе фильтров, подробно рассмотренных в трудах Л.П. Ярославского, Грузмана И.С., Е. Претта, Н. Ахмеда, К.Р. Рао и других, посвящённых цифровой обработке сигнала. Затем из изображения могут быть автоматически выделены движущиеся элементы. В литературных источниках, посвящённых обработке изображений и опубликованных в России и за рубежом, наиболее часто упоминается так называемый алгоритм вычитания фона как наиболее эффективно решающий задачу выделения движущихся объектов на изображении неподвижных видеокамер с точки зрения соотношения требований к ресурсам ЭВМ и надёжности определения. Наконец, накопленный мировым сообществом учёных опыт в области распознавания образов, получивший отражение в работах Харкевича А.А., Ковалевского В.А., Ф. Розенблата, Бравермана Э.М., Галушкина А.И., Ивахненко А.Г., Вапника В.Н., Червоненкиса А.Я., Журавлёва Ю.И., Злобина В.К. и других, стал основой для автоматической классификации по определённым заранее категориям объектов и событий в области наблюдения видеокамер. Например, можно выделять только автомобили и не выделять людей при отражении событий на терминале операторов. В последнее время значительное распространение получил аппарат анализа изображений с помощью свёрточных нейронных сетей и теории статистических решений, достигнуты достаточно высокие показатели точности классификации, правда, для реализации нейросетевых подходов требуются значительные вычислительные ресурсы, что не позволяет использовать нейросети с большим, более миллиона, количеством весов нейронов в системах реального времени.

Анализ состояния современных систем видеонаблюдения выявил, что алгоритмы классификации изображений, сопряжение и синхронизация не применяются в полной мере, хотя они могут существенно повысить эргономические характеристики терминалов операторов, облегчить анализ событий в зоне наблюдения большого количества камер и повысить разрешение выводимых изображений. Применение ручного управления поворотными видеокамерами в существующих видеосистемах приводит к задержке реакции

операторов на появление новых объектов ($\approx 10\text{с}$), что повышает вероятность пропуска динамичных и важных событий в области наблюдения. На основе анализа недостатков известных видеосистем и методов обработки видеосигналов поставлена цель диссертационной работы: разработать, проанализировать и внедрить более эффективные способы и алгоритмы автоматического сопряжения, синхронизации и юстировки видеоизображений и управления поворотными камерами при создании интеллектуальных автоматических видеосистем нового поколения.

Во второй главе проводится разработка и функциональный анализ параллельных алгоритмов семантической обработки видеосигналов и создание теоретических основ синхронизации и сопряжения изображений.

Будем называть мастер-камерами или мастерами приборы технического зрения, на которых автоматически выделяются движущиеся объекты и которые, как правило, являются оптическими приборами либо радиолокационными датчиками. Поворотные видеокамеры, которые автоматически наводятся на эти объекты, будем именовать слейв-камерами или слейвами. Управление слейвов подчинено положению и скорости объектов, выделяемых на мастерах. Назовём кластером подмножество приборов системы технического зрения, связанное между собой сопряжением. Синхронизацию данных в кластере видеокамер обеспечивает специальный модуль-диспетчер, назначение которого — принять информацию о положении объектов от мастеров и назначить некоторые объекты для наведения на них тех слейвов, которые в данный момент свободны и сопряжены с теми областями видеокамеры, в которой находятся объекты. Если число мастер-камер M , а число слейв-камер N , тогда число мастеров и слейвов прежде всего подчиняется неравенству $M \geq N$ из экономических соображений. Ограничение пропускной способности локальной вычислительной сети (ЛВС) не позволяет использовать более

$$M_{max} = \frac{\eta w_0}{(1+k)n w_{obj} f + \frac{1}{T_R}(w_t + w_q + w_a + w_f)} \quad (1)$$

неподвижных видеокамер, где n — среднее количество объектов на одной обзорной камере, w_{obj} — объём данных в байтах, описывающих один объект при передаче от обработчика мастера диспетчеру или от мастера к мастеру, k — среднее количество мастеров, с которыми пересекается каждый мастер, и, следовательно, должен передавать объекты своим k соседям, w_t — размер сообщения-целеуказания от диспетчера, указывающее слейву, куда нужно повернуться, w_q — объём сообщения о положении объекта от поворотной камеры к обзорной, w_a — объём ответа от мастера слейву, w_f — объём сообщения о готовности принять следующее задание, передаваемое диспетчеру слейвом, f — частота обработки кадров с обзорных камер, T_R — среднее время исполнения поворота слейвами, $\eta = 0,5$ — коэффициент «полезного

действия» пересылки по сети: из-за издержек протоколов на физическом и транспортном уровне невозможно добиться максимальной пропускной способности w_0 . Оценка получена для предельного случая $N = M$, и поскольку установлено экспериментально и теоретически, что $n = 50$, $k = 3$ для предельного случая, $w_{obj} = 20$ байт, $w_t = w_q = w_a = w_f = 24$ байта, $f = 25\text{с}^{-1}$, $T_R = 500\text{мс}$, то при пропускной способности локальной вычислительной сети (ЛВС) $w_0 = 100\text{Мбит/с}$ получим $M_{max} = 65$. Данное ограничение действует только на один кластер. Поскольку различные кластеры не ведут сетевых обменов, количество кластеров ресурсами ЛВС не ограничивается.

Ограничение, обусловленное вычислительной мощностью ЭВМ, рассчитывается исходя из выведенной в диссертации формулы для доли ресурсов δ , занимаемых модулем диспетчера:

$$\delta = fM(C_1n + C_3Nn) + \frac{1}{T_R}N(C_2Mn + C_5N!Mn), \quad (2)$$

где $C_1 = 0,1\text{мкс}$, $C_2 = 2\text{мкс}$, $C_3 = C_5 = 1\text{мкс}$ — константы времени выполнения элементарных операций алгоритма синхронизации, измеренные с точностью 20% на ЭВМ с эталонной вычислительной мощностью. Понятие эталонной вычислительной мощности введено в работе для унификации расчётов и примерно соответствует ЭВМ с тактовой частотой процессора 3ГГц и частотой обмена данных с динамической памятью 400МГц. Обычно около 90% ресурсов ЭВМ резервируется для обработки сигналов от видеокамер, поэтому типично для диспетчера выделяется около 10% ресурсов ЭВМ, т.е. $\delta < 0,1$, и с помощью численного решения данного неравенства относительно неизвестных M и N получим

$$\begin{cases} \{N = 4, M = 4\}, \\ \{N = 3, 3 \leq M \leq 9\}, \\ \{N = 2, 2 \leq M \leq 14\}, \\ \{N = 1, M \leq 28\}, \end{cases} \quad (3)$$

то есть вычислительная мощность современных ЭВМ сильнее ограничивает количество камер в кластере, чем пропускная способность ЛВС.

Сопряжение приборов технического зрения с пересекающимися зонами контроля в случае плоской поверхности в области обзора проводится в работе на основе формул проективного преобразования

$$\begin{aligned} x' &= \frac{Ax + By + C}{Gx + Hy + I} \\ y' &= \frac{Dx + Ey + F}{Gx + Hy + I} \end{aligned}, \quad \text{где } \Delta = \begin{vmatrix} A & B & C \\ D & E & F \\ G & H & I \end{vmatrix} \neq 0. \quad (4)$$

Чтобы получить коэффициенты преобразования, с помощью специальной настроенной программы вводится сетка $\{(x_i, y_i)\}$ и значения искомого преобра-

зования на этой сетке $\{(x'_i, y'_i)\}$ за счёт указания соответствующих друг другу точек изображений. Затем методом градиентного спуска решается задача

$$\min \sum_{i=1}^k \left[\left(x'_i - \frac{Ax_i + By_i + C}{Gx_i + Hy_i + 1} \right)^2 + \left(y'_i - \frac{Dy_i + Ey_i + F}{Gx_i + Hy_i + 1} \right)^2 \right]. \quad (5)$$

Поиск начального приближения предложено провести с помощью решения линейной переопределённой системы, которая следует из системы уравнений, состоящих из соотношений вида (4)

$$x'_i = \frac{Ax_i + By_i + C}{Gx_i + Hy_i + I}, \quad y'_i = \frac{Dx_i + Ey_i + F}{Gx_i + Hy_i + I} \quad (6)$$

после домножения на знаменатели. Если форма наблюдаемой поверхности не является приближением плоскости, а видеокамеры не находятся в одной точке, то преобразование между изображениями не является проективным. Поэтому в качестве приближения предложено использовать остаточные суммы ряда Фурье:

$$\begin{aligned} x' &= \frac{Ax + By + C}{Gx + Hy + I} + \sum_{i=1}^4 \left(C_i \cos\left(\frac{x\pi}{w}i\right) + S_i \sin\left(\frac{x\pi}{w}i\right) \right), \\ y' &= \frac{Dx + Ey + F}{Gx + Hy + I} + \sum_{i=1}^4 \left(C'_i \cos\left(\frac{y\pi}{h}i\right) + S'_i \sin\left(\frac{y\pi}{h}i\right) \right), \end{aligned} \quad (7)$$

где $w \times h$ — разрешение исходного изображения, а неопределённые коэффициенты C_i, S_i, C'_i, S'_i требуется найти. В диссертации разработан алгоритм поиска данных коэффициентов с помощью минимизации квадрата невязки на сетке с известными значениями преобразования.

Следует отметить, что формулы сопряжения изображений видеокамер с картой применимы также для сопряжения пространства координат радиолокационных датчиков и видеокамер, что легло в основу разработанной в рамках диссертационной работы системы контроля за соблюдением скоростного режима и правил дорожного движения.

Сопряжение поворотных и неподвижных видеокамер проводится по аналогии с сопряжением неподвижных видеокамер между собой: вручную с помощью специальной программы задаётся сетка координат $\{(x_i, y_i)\}$ и соответствующие узлам сетки углы $\{(\alpha_i, \beta_i)\}$ оптической оси поворотной камеры в полярной системе координат с центром в точке закрепления камеры, а затем проводится интерполяция преобразования на всей области изображения. Разработан ряд алгоритмов для оптимизации процесса сопряжения. Например, для камер, которые поддерживают только команды управления, перемещающие оптическую ось с различными скоростями в течение заданного времени (без обратной связи по положению видеокамеры), разработан

алгоритм составного перемещения на основе измеренных векторов смещения (при заданном времени и скорости смещения), который не увеличивает неопределённость конечного положения за счёт сравнения изображений слайва и мастера. Предложены алгоритмы автоматического выбора необходимых параметров управления увеличением и фокусировкой.

На основании соотношений, выведенных для сопряжения видеокамер и других приборов технического зрения во второй главе, в **третьей главе** проведены исследование и разработка алгоритмов синхронизации, управления и семантической обработки информации в системах технического зрения, а также разработаны алгоритмы наведения поворотных видеокамер на движущиеся объекты и алгоритмы сопоставления изображений одного и того же объекта, регистрируемых различными приборами с общей зоной контроля. Также рассмотрены способы синтеза параметров изображений при применении классификации объектов по категориям.

Для организации наведения слайвов предложено несколько схем синхронизации данных о положении объектов между обработчиками слайвов и мастеров. Все схемы состоят из пяти типовых блоков: копирование данных об объектах при приёме сообщения от мастера (сложность C_1n), сортировка Mn объектов по приоритетности наведения (сложность $C_2Mn \log_2(Mn)$), проверка возможности наведения одного из слайвов для только что принятых объектов (сложность C_3Nn), добавление новых объектов в существующий отсортированный список (сложность $C_4n \log_2(Mn)$), распределение поворотных видеокамер для наведения на объекты в общем списке (сложность $C_5nMN!$). В результате сравнения вычислительной сложности выбрана схема синхронизации с обработкой в циклах приёма сообщений мастера и слайва. Вычислительная сложность алгоритма описывается формулой (2).

Для наведения поворотной камеры на движущийся объект в работе введено соотношение $\vec{\tau} - \vec{\tau}_1 = \vec{\omega}T(\vec{\tau}, \vec{\tau}_0, \gamma, \gamma_0, f, f_0)$, где $\vec{\tau} = (\alpha, \beta)$ — углы положения оптической оси, соответствующие положению объекта после перемещения поворотной камеры; $\vec{\tau}_1$ — углы начального положения объекта; $\vec{\tau}_0$ — углы начального положения поворотной камеры, $\vec{\omega}$ — угловая скорость объекта; γ, f, γ_0, f_0 — соответственно, координаты увеличения и фокусировки, конечные и начальные; T — функция, определяющая время изменения положения поворотной видеокамеры. Функцию T следует измерять экспериментально, тогда можно определить время упреждения, решая одномерную задачу оптимизации

$$\min_{t \leq 4c} |t - T(\vec{\tau}_1 + \vec{\omega}t, \vec{\tau}_0, \gamma, \gamma_0, f, f_0)| \quad (8)$$

(время поворота применяемых в работе камер менее 4с). Решив эту задачу, найдём конечные координаты поворотной камеры $\vec{\tau} = \vec{\tau}_1 + \vec{\omega}t$. Наведение, определяемое соотношением (8), назовём наведением с адаптивным

упреждением. В результате теоретической оценки погрешности наведения на неподвижные объекты, сделанной с помощью метода численного моделирования, получается значение $0,5^\circ$. Теоретическая погрешность наведения на подвижные объекты при максимальной погрешности определения скорости и времени синхронизации равна 20° .

Для сопоставления изображений одного и того объекта на различных видеокамерах с общей зоной обзора введён критерий

$$\begin{cases} \vec{u}_{ds}(\vec{v}^s) \in \{(x, y) | x_{min}^d \leq x \leq x_{max}^d, y_{min}^d \leq y \leq y_{max}^d\}, \\ \vec{u}_{sd}(\vec{v}^d) \in \{(x, y) | x_{min}^s \leq x \leq x_{max}^s, y_{min}^s \leq y \leq y_{max}^s\}, \end{cases} \quad (9)$$

где $[x_{min}^d, x_{max}^d] \times [y_{min}^d, y_{max}^d]$ — прямоугольник, ограничивающий изображение объекта на мастере d ; $[x_{min}^s, x_{max}^s] \times [y_{min}^s, y_{max}^s]$ — прямоугольник объекта на мастере s ; \vec{u}_{sd} — преобразование сопряжения с мастера d на мастер s ; \vec{u}_{ds} — преобразование сопряжения с мастера s на мастер d и $\vec{v}_d = (\frac{x_{min}^d + x_{max}^d}{2}, \frac{y_{min}^d + y_{max}^d}{2})$, $\vec{v}_s = (\frac{x_{min}^s + x_{max}^s}{2}, \frac{y_{min}^s + y_{max}^s}{2})$. Разработанный на основе предложенного критерия алгоритм позволяет не учитывать скорости движения объектов, поскольку время синхронизации данных о положении объектов значительно меньше, чем время между обработкой двух последовательных кадров, равное 40мс, и за это время объект сдвигается не более, чем на несколько пикселей. Использование экстраполяции траектории объекта может, наоборот, увеличить ошибку сопоставления. Поэтому предложенный критерий, определяемый неравенствами (9), оказывается наиболее эффективным. Рассмотрев модель перемещения объектов в поле зрения камер и сравнивая время перемещения и обработки, можно утверждать, что критерий (9) наиболее достоверен, если скорость объектов, для которых сопоставляются изображения, не превышает 14км/ч для объектов с размером 0,5м и 90км/ч для объектов размером 3м.

В четвёртой главе рассматриваются вопросы теории тестирования алгоритмов семантической обработки изображений, на основании которой проводится разработка, исследование и анализ методик сравнения алгоритмов семантического сжатия видеоинформации. Типичным подходом к оценке качества алгоритма сегментации является сравнение результата с шаблонной сегментацией, которая определяется либо вручную оператором, либо с помощью другого точного алгоритма. Будем обозначать множество точек шаблонной сегментации объекта как T . Обозначим $|\mathcal{M}|$ мощность (количество элементов) множества \mathcal{M} . Распространённым показателем качества сегментации изображения является отношение

$$\varrho(A, T) = \frac{|A \cap T|}{|A \cup T|}. \quad (10)$$

Очевидно, что $\forall A, T : \varrho(A, T) \leq 1$, и $\forall A, T : \varrho(A, T) = 1$ равносильно $A = T$,



Рисунок 1. Примеры некорректной оценки качества выделения шаблона с помощью меры $\varrho(A, T)$

то есть чем больше результат A повторяет форму шаблона T , тем ближе значение ϱ к единице.

В рамках проведённого обзора источников установлены ряд недостатков, присущих оценке (10) меры совпадения сегментации с эталонной. Существующие способы устранения данных недостатков были обобщены с привлечением теории нечётких множеств. Детерминированному множеству T мы поставим в соответствие нечёткое множество T' , являющееся парами из точек изображения p и числовой меры принадлежности точки множеству в диапазоне $[0, 1]$. В качестве меры принадлежности к нечёткому множеству шаблонных точек T' мы введём бинарную функцию

$$\mu^{T'}(p) = \begin{cases} 1, & \text{если } p \in T, \\ 0 & \text{в противном случае,} \end{cases}$$

т.е. принадлежность точек к шаблону определена точно. Для множества A , выдаваемого в результате алгоритма сегментации, мы определим соответствующее нечёткое множество A' , мера принадлежности каждой точки изображения которого будет определяться на основании функции $\mu^{T'}$:

$$\mu^{A'}(p) = \begin{cases} 1, & \text{если } \mu^{T'}(p) = 1 \text{ и } p \in A, \\ g(p), & \text{если } \mu^{T'}(p) = 1 \text{ и } p \notin A, \\ b(p), & \text{если } \mu^{T'}(p) = 0 \text{ и } p \in A, \\ 0, & \text{если } \mu^{T'}(p) = 0 \text{ и } p \notin A. \end{cases}$$

Функция $g(p)$ определяется как функция сглаживания с весами распределения Гаусса:

$$g(p) = \sum_{q \in A} \frac{\mu^{T'}(q)}{\sqrt{2\pi\sigma^2}} e^{-\frac{\|p-q\|^2}{2\sigma^2}},$$

где $\|p-q\|$ — евклидова норма вектора с началом в точке p и с концом в точке q . Благодаря сглаживанию в A' с некоторым весом входят даже те точки, которые не вошли в результат алгоритма A , причём вес определяется величиной окрестности суммирования σ^2 и количеством точек из пересечения $A \cap T$, которые находятся близко к точке p , при этом чем больше таких точек, тем выше вес. Принято использовать небольшие окрестности, например, $\sigma^2 = 5$.

Функция $b(p) = \frac{2-e^{\alpha d(T,p)}}{2}$, где $d(T,p) = \min_{q \in T} \|p - q\|$ — расстояние от точки p до множества T и α — отрицательная константа, определяет вес точек A' , находящихся вне T , в зависимости от того, насколько они удалены от множества T . Экспоненциальная функция $f(x) = \frac{2-e^{\alpha x}}{2}$ достаточно быстро возрастает до 1, при этом зону пониженного вклада в ошибку первого рода можно задать показателем α .

После того, как нечёткие множества T' и A' определены, меру совпадения исходных множеств T и A можно определить как

$$\varrho^*(A, T) = \frac{|A' \cap T'|}{|A' \cup T'|}, \quad (11)$$

где объединение, пересечение, и мощность множеств определяются в рамках теории нечётких множеств, т.е. $\mu^{A' \cap T'}(x) = \min(\mu^{A'}(x), \mu^{T'}(x))$, $\mu^{A' \cup T'}(x) = \max(\mu^{A'}(x), \mu^{T'}(x))$, $|M| = \sum_{x \in M} \mu^M(x)$ для $M = A' \cap T'$ и $M = A' \cup T'$.

Ввиду того, что принадлежность точек множеству A' задана более приближенно к шаблону T в тех случаях, когда это не мешает субъективному восприятию результата A , мера совпадения ϱ^* более устойчива к небольшим шумовым искажениям в результатах алгоритма сегментации. Например, для примеров на рис. 1 за счёт функции сглаживания $g(p)$ мера совпадения $\varrho^*(A_C, T)$ значительно меньше $\varrho^*(A_B, T)$, а за счёт уменьшенного веса $b(p)$ точек возле границ шаблона T мера совпадения $\varrho^*(A_E, T)$ значительно больше меры $\varrho^*(A_D)$, но при этом мера сравнения ϱ , рассчитанная по формуле (10), одинакова для соответствующих пар. Поэтому мера (11) на основе нечётких множеств более соответствует ожидаемому поведению меры сравнения.

В диссертационной работе на основе анализа открытых тестовых видеопоследовательностей и их семантической разметки показано, что одной только меры сравнения результата алгоритма семантической обработки одного изображения не достаточно для полноценной оценки аналитических систем технического зрения. Для проверки эффективности алгоритмов сопровождения объектов на наборе тестовых видеопоследовательностей введены следующие математические понятия и обозначения:

- $V^n = \{V_i^n\}$ для последовательности кадров видео, где i — номер кадра, n — номер видеопоследовательности в тестовой базе;
- $\{I_l^n\}$ для объектов идеальной разметки, где l — номер идеального объекта, n — номер видеопоследовательности, на которой данный объект отмечен;
- $\{A_m^{an}\}$ для объектов автоматической разметки, где m — номер объекта, n — номер видеопоследовательности, a — номер тестируемого алгоритма;
- $P_{il}^n = P(V_i^n, I_l^n)$ и $P_{im}^{an} = P(V_i^n, A_m^{an})$ — функции, возвращающие положение объектов I_l^n и A_m^{an} на кадре V_i^n , где положение P_{il}^n и P_{im}^{an} — это множество точек объектов в кадре, которое, в связи с трудоёмкостью поточечной раз-

метки изображений объектов, обычно представляется в виде прямоугольника наименьшего размера с параллельными границами кадра сторонами и содержащего все точки изображения объекта;

- $\rho(P_{il}^n, P_{im}^{an}) = \rho(P_{im}^{an}, P_{il}^n) = \frac{|P_{im}^{an} \cap P_{il}^n|}{|P_{im}^{an} \cup P_{il}^n|} = \rho_{lm}^{ian}$ — мера совпадения положений автоматического объекта A_m^{an} и идеального объекта I_l^n на кадре i последовательности V^n ;
- $\mu(\omega)$ — это число элементов в конечном множестве ω , т.е. мощность;
- $\mu(A_m^{an}) = \mu(\{i | P_{im}^{an} \neq \emptyset\})$ — число кадров, на которых есть объект A_m^{an} ;
- $\mu(I_l^n) = \mu(\{i | P_{il}^n \neq \emptyset\})$ — число кадров, на которых есть объект I_l^n .

На основе введённых обозначений можно рассчитать следующие метрики работы алгоритмов (для сокращения индексы алгоритма a и видеопоследовательности n опущены).

Число ошибок первого рода можно рассчитать по формуле

$$F_P^{\alpha\beta} = \mu\left(\left\{A_m \middle| \forall l : \frac{\mu(\{i | \rho_{lm}^i < \alpha\})}{\mu(A_m)} > \beta\right\}\right), \quad (12)$$

где $\alpha \in (0, 1]$ — порог меры совпадения автоматического и идеального объекта на одном кадре для того, чтобы иметь основания эти объекты сопоставлять на данном кадре, а $\beta \in [0, 1]$ — порог на долю кадров автоматических объектов, на которых идеальный объект должен быть сопоставлен автоматическому для того, чтобы автоматический объект не считался ложным срабатыванием. Чем ближе пороги α и β к единице, тем строже критерий для отбора объектов, не являющихся ложными, чем ближе эти пороги к 0, тем критерий мягче. Например, при $\beta = 0$ и достаточно малых α , например, $\alpha = \alpha_m = \frac{1}{2} \rho_{lm}^i$, автоматический объект будет считаться ложным только тогда, когда он ни разу не пересекался по положению ни с каким из идеальных объектов.

Число ошибок второго рода можно рассчитать по формуле, подобной (12):

$$F_N^{\alpha\beta} = \mu\left(\left\{I_l \middle| \forall m : \frac{\mu(\{i | \rho_{lm}^i < \alpha\})}{\mu(I_l)} > \beta\right\}\right), \quad (13)$$

т.е. идеальный объект считается пропущенным, если не нашлось ни одного автоматического объекта, который совпадал бы по положению лучше, чем на α , на более, чем на $(1 - \beta)\mu(I_l)$ кадрах. При $\beta = 0$ и $\alpha = \alpha_m$ идеальный объект считается обнаруженным, если хотя бы на одном кадре его положение пересекается с положением одного из автоматических объектов.

Качество сопровождения объекта можно оценить с помощью показателей

$$Q_T^\alpha = \sum_l Q_{Tl}^\alpha, \text{ где } Q_{Tl}^\alpha = \mu(\{i | \exists m : \rho_{lm}^i \geq \alpha\}), \quad (14)$$

$$E_T^\alpha = \sum_m E_{Tm}^\alpha, \text{ где } E_{Tm}^\alpha = \mu(\{I_l | \exists i : \rho_{lm}^i \geq \alpha\}), \quad (15)$$

где значение Q_{Tl}^α характеризует, насколько постоянно идеальный объект I_l сопровождается автоматическими объектами, а значение E_{Tm}^α показывает, сколько раз автоматический объект A_m ошибочно меняет сопровождаемые идеальные объекты. Для оценки качества выделения траекторий интересны интегральные соотношения $Q^\alpha = \frac{Q_T^\alpha}{\sum_l \mu(I_l)}$, $E^\alpha = \frac{E_T^\alpha}{\mu(\{A_m\})}$, которые, соответственно, показывают долю кадров, на которых идеальные объекты сопровождались автоматическими объектами, и число ошибок смены траектории по отношению к общему числу автоматических объектов.

Расчёт метрик (12) – (15) представляет собой вычислительно ёмкую задачу, если необходимо обработать большой объём видеоматериала. В рамках диссертационной работы создан программный комплекс тестирования «Супервизор», который работает с информационной системой, включающей в себя видеоматериалы, их разметку и результаты тестирования различных алгоритмов. «Супервизор» предоставляет набор программных инструментальных средств для анализа результатов тестов и для запуска распределённых расчётов характеристик на рамоченных видеоданных с помощью системы распределённых вычислений HTCondor, что позволяет в десятки раз сократить время тестирования за счёт масштабирования вычислительного кластера ЭВМ, подключенного к системе HTCondor.

В пятой главе алгоритмы синхронизации и сопряжения рассмотрены в рамках разработки алгоритмов анализа стереоизображений. Проведён обзор теории сопряжения двух видеокамер, образующих стереосистему, при котором автоматически определяется взаимное расположение видеокамер друг относительно друга (так называемая база и поворот оптических осей и чувствительных элементов камеры друг относительно друга) и теории автоматического построения так называемой карты диспаратностей, то есть смещений от точек изображения объектов, видимых на одной из камер (условно левой) к точкам изображения тех же объектов, видимых на второй камере (условно правой). Карта диспаратностей позволяет оценивать расстояние от стереокамер до изображённых объектов. Чтобы определить преимущества анализа стереоизображений перед алгоритмами семантической обработки изображений от одной камеры, сравним два этих подхода.

Пусть $\{F_i(x, y)\}$ — последовательность изображений, а значениями функций F_i являются интенсивности изображения в точках (x, y) . Условимся, что проведена предварительная калибровка, и в каждой точке изображения задана функция $H(x, y)$, значениями которой является характерный размер наблюдаемых объектов в точке. Последовательность изображений и функция $H(x, y)$ относятся к входным данным для алгоритма отслежива-

ния движения. К выходным данным алгоритма относится множество объектов $\mathbf{A} = \{A_m\}$, где для каждого объекта A_m определено его множество положений $\{P_{im}^a\}$ на i -ом кадре видеопоследовательности, индекс a показывает, что положение объекта определено автоматически. Само положение объекта определяется координатами минимального по площади прямоугольника со сторонами, параллельными координатным осям. Примером алгоритма обработки видеопоследовательности изображений, достаточно достоверно выделяющего движущиеся объекты в поле зрения одной камеры, при этом обладая невысокой вычислительной сложностью, является следующая последовательность шагов.

1. Осуществим подбор проективного преобразования координат

$$(x', y') = \mathbf{h}(x, y) = (h_x(x, y), h_y(x, y)),$$

такого, что на кадрах $F_i^p(x, y) = F_i(h_x(x, y), h_y(x, y))$ характерный размер изображения объектов в каждой точке примерно равен фиксированному значению $n_h = 16$. Преобразование $\mathbf{h}(x, y)$ представляет собой две дробно-линейные функции

$$\begin{aligned} x' = h_x(x, y) &= \frac{Ax + By + C}{Gx + Fy + 1}, \\ y' = h_y(x, y) &= \frac{Dx + Ey + F}{Gx + Fy + 1}, \end{aligned}$$

где коэффициенты A, B, C, D, E, F, G, H необходимо подобрать таким образом, чтобы в любой точке кадра выполнялось условие

$$h_y(x, y - H(x, y)) - h_y(x, y) \approx n_h.$$

При этом коэффициенты проективного преобразования могут быть заданы заранее с помощью графической утилиты подбора преобразования кадра, либо их можно вычислить, решив следующую задачу

$$\min_{A,B,C,D,E,F,G,H} \sum_{(x,y)} (h_y(x, y - H(x, y)) - h_y(x, y) - n_h)^2,$$

например, методом градиентного спуска. Начальное приближение можно получить с помощью решения несколько упрощённой задачи

$$\min_{A,B,C,D,E,F,G,H} \sum_{(x,y)} (h_y(x, y - H(x, y)) - h_y(x, y) - n_h)^2 \cdot (Gx + Fy + 1)^2,$$

например, методом наименьших квадратов, поскольку в данном случае целевая функция состоит из слагаемых, являющимися квадратами линейных функций относительно пространства, в котором ведётся поиск оптимальных значений.

2. Введём множество объектов T_i , отслеживаемых на кадре i . Элементами множеств T_i будем считать точки t , характеризующие положение объектов. Индекс начального кадра видеопоследовательности мы будем считать единственным, и положим $T_0 = T_1 = \emptyset$. Множество результирующих объектов в начале работы алгоритма положим $\mathbf{A} = \emptyset$.

3. Следующие шаги алгоритма повторяются по порядку для каждого кадра последовательности $\{F_i(x, y)\}$.

4. Кадр $F_i(x, y)$ преобразуется в кадр $F_i^p(x, y)$ с помощью преобразования, полученного на шаге 1.

5. Кадр $F_i^p(x, y)$ подвергается операции Гауссова размытия путём свёртки с функцией $g(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}}$, где $\sigma = \frac{n_h}{3}$. В результате получается кадр $F_i^{pg}(x, y) = g(x, y) * F_i^p(x, y)$.

6. Для кадра $F_i^{pg}(x, y)$ производится поиск точек локальных нестрогих экстремумов, положение которых запоминается с помощью характеристических функций-таблиц $\chi_{\min}(x, y)$ и $\chi_{\max}(x, y)$:

$$\chi_{\min}(x, y) = \begin{cases} 1, & \text{если} \\ \forall x' \in [x - \frac{n_h}{2}; x + \frac{n_h}{2}] \forall y' \in [y - \frac{n_h}{2}; y + \frac{n_h}{2}] : \\ & F_i^{pg}(x', y') \geq F_i^{pg}(x, y), \\ 0 & \text{в противном случае,} \end{cases}$$

$$\chi_{\max}(x, y) = \begin{cases} 1, & \text{если} \\ \forall x' \in [x - \frac{n_h}{2}; x + \frac{n_h}{2}] \forall y' \in [y - \frac{n_h}{2}; y + \frac{n_h}{2}] : \\ & F_i^{pg}(x', y') \leq F_i^{pg}(x, y), \\ 0 & \text{в противном случае.} \end{cases}$$

7. С помощью алгоритма волновой раскраски выделяются множества соседних точек, помеченные 1 в области значений функций $\chi_{\min}(x, y)$ и $\chi_{\max}(x, y)$. Множества центров масс выделенных областей обозначим как $R'_i = \{(c_x^r, c_y^r)\}_{r=1}^{n'_r} = \{c^r\}_{r=1}^{n'_r}$.

8. Из множества точек R'_i выделяются точки R_i с наиболее различимой текстурой. Это производится с помощью критерия Хариса: в каждой точке оцениваются градиенты $\nabla F_i^p = \left(\frac{\partial F_i^p}{\partial x}, \frac{\partial F_i^p}{\partial y} \right)$ функции $F_i^p(x, y)$, из них составляется матрица $(\nabla F_i^p)^T \cdot \nabla F_i^p$ размером 2×2 . Для устранения шума в градиентах данная матрица в каждой точке (x, y) усредняется по окрестности $[x - \frac{n_h}{2}; x + \frac{n_h}{2}] \times [y - \frac{n_h}{2}; y + \frac{n_h}{2}]$, а затем для усреднённой матрицы рассчитывается минимальное собственное значение. Если это собственное значение больше некоторой величины, значит найденный центр масс точек нестрогого экстремума на изображении $F_i^p(x, y)$ является достаточно существенным, и поэтому легко идентифицируемым на предыдущих и последующих кадрах. Именно эти точки из множества R'_i отбираются во множество R_i .

9. Из элементов множества $T_{i-1} = \{t^o\}$ и элементов множества $R_i = \{c^r\}$ строится соответствующий двудольный граф, каждому ребру (t^o, c^r) сопоставляется вес $w_o^r = \alpha \|t^o - c^r\| + \beta \varepsilon(t^o, c^r)$, где α и β — положительные весовые коэффициенты, $\|t^o - c^r\|$ — расстояние между соответствующими точками, $\varepsilon(t^o, c^r)$ — разность изображений по норме L_2 в окрестности точки (t_x^o, t_y^o) на предыдущем кадре F_{i-1}^p и в окрестности точки (c_x^r, c_y^r) на текущем кадре F_i^p . С помощью метода Эдмондса, решается задача об оптимальном паросочетании

$$\min \sum_{(r,o)} (C - w_o^r),$$

где C — константа, превышающая все веса w_o^r , а $E = \{(r, o)\}$ — множество пар индексов, однозначно определяющих рёбра построенного двудольного графа, входящие в оптимальные паросочетания, это подразумевает что каждый индекс вершин r и o может входить в найденное множество E только один раз. Используя найденное множество E , строится множество новых положений объектов на T_i текущем кадре: если $(r, o) \in E$, то включаем во множество T_i элемент t^o из множества T_{i-1} , при этом его положение обновляем новым положением c^r . Если существуют элементы R_i , которые не входят в рёбра, задаваемые множеством E , то в множество T_i включаются новые объекты, соответствующие положению областей R_i без пары. Все объекты $t \in T_{i-1}$, которые не обладают парой среди R_i , переносятся в T_i , при условии, что $\exists n < n_t : (\text{на кадре } T_{i-n-1} \text{ у объекта } t \text{ была пара в } R_{i-n})$. Значение n_t является постоянным параметром, определяющим время сохранения в памяти информации об объекте, измеряемое в количестве кадров видеопоследовательности.

10. Обновляются выходные данные $\mathbf{A} = \{A_m\}$. Информация о вновь появившихся объектах t множества T_i переносится в качестве элементов A_m при условии, что

$$\exists n_0 \forall n < n_b : (\text{на кадре } T_{i-n_0-n-1} \text{ у объекта } t \text{ была пара в } R_{i-n_0-n}),$$

где n_b — постоянный параметр, определяющий длину отрезка последовательности изображений, на каждом кадре которого рассматриваемый объект должен сопоставляться с какой-либо точкой экстремума, чтобы попасть в конечный результат работы алгоритма. Положения объектов отмечаются прямоугольными рамками, получить которые можно с помощью построенного преобразования $\mathbf{h}(x, y)$: для каждого объекта A_m на кадре i известно его текущее положение $t^m = (t_x^m, t_y^m)$. Окрестность положения объекта в координатах кадра $F_i^p(x, y)$ можно определить как $P''_m = [t_x^m - \frac{n_h}{2}; t_x^m + \frac{n_h}{2}] \times [t_y^m - \frac{n_h}{2}; t_y^m + \frac{n_h}{2}]$. Прообраз данного множества при преобразовании $\mathbf{h}(x, y)$ является оценочным положением объекта на исход-

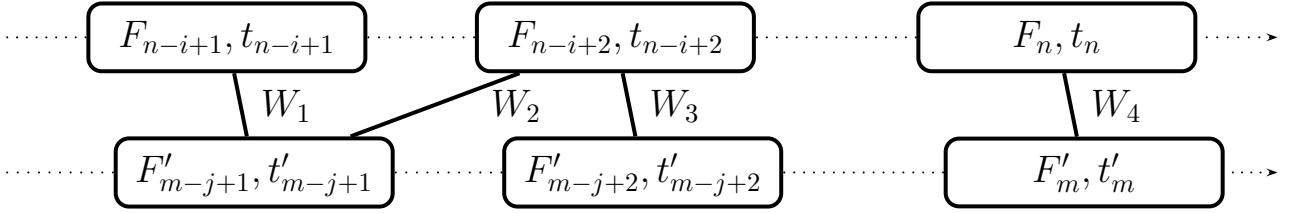


Рисунок 2. Двудольный граф из асинхронных кадров, веса рёбер равны $W_1 = T_{\max} - |t'_{n-i+1} - t_{m-j+1}|$, $W_2 = T_{\max} - |t'_{n-i+1} - t_{m-j+1}|$, $W_3 = T_{\max} - |t'_{n-i+2} - t_{m-j+2}|$, $W_4 = T_{\max} - |t'_n - t_m|$.

ном кадре $F_i(x, y)$, т.е. $P_{im}'^a = \mathbf{h}^{-1}(P_{im}''^a)$ определяет четырёхугольник, приближающий положение объекта. В качестве конечного результата, используя координаты вершин $P_{im}'^a$, строится P_{im}^a — минимальный по площади прямоугольник со сторонами, параллельными осям координат исходного кадра.

Таким образом, получается результат $\mathbf{A} = \{A_m\}$. Для краткости дальнейших ссылок на данный алгоритм, обозначим его \mathcal{A}_{mono} . Алгоритм \mathcal{A}_{mono} с высокой достоверностью сопровождает объекты при малой плотности движения не более одного объекта в течение 10 секунд в поле зрения камеры. В этом случае показатели $F_N^{\alpha\beta}$ и $F_P^{\alpha\beta}$ не выше 5% по отношению к общему числу объектов в кадре. Однако данные ошибки значительно возрастают до 20% при увеличении потока движения до нескольких объектов в секунду.

Алгоритм отслеживания движущихся объектов с помощью стереоизображений можно построить аналогично \mathcal{A}_{mono} . На входе алгоритма, кроме функции характерных размеров объектов $H(x, y)$, также задаётся высота подвеса L для пары камер, угол наклона оптических осей к наблюдаемой сцене γ , \vec{b} — калибровочные параметры стереопары, к которым относятся: смещение \vec{b} от одной камеры к другой, матрица поворота R , позволяющая привести положение оптической оси одной камеры к такому положению, при котором она параллельна расположению второй камеры, векторы коэффициентов устранения радиальной дисторсии D_1 и D_2 объективов стереопары, матрицы M_1 и M_2 , состоящие из внутренних параметров камер, характеризующие фокусное расстояние и разрешение изображений. Кроме последовательности кадров $\{F_i(x, y)\}$ с левой камеры на вход алгоритмов поступает последовательность соответствующих кадров с $\{F'_i(x, y)\}$ с правой камеры.

Отслеживание движущихся объектов с помощью стереоизображений состоит из следующих шагов.

1. Повторяется шаг 1 алгоритма \mathcal{A}_{mono} , и находится преобразование $\mathbf{h}(x, y)$, приводящее все изображения объектов к изображению с характерным размером $n_h = 16$.
2. Проводится предварительная инициализация множеств T_i и \mathbf{A} аналогично шагу 2 алгоритма \mathcal{A}_{mono} .
3. Дальнейшие шаги повторяются для каждой пары кадров последова-

тельностей $\{F_i(x, y)\}$ и $\{F'_i(x, y)\}$. В некоторых случаях требуется выбор синхронных пар кадров по алгоритму, описанному в работе: из очереди кадров левой и правой камеры с метками времени t_i и t'_j , соответственно, строится граф, изображённый на рис. 2, в котором с помощью алгоритма поиска максимальных паросочетаний выбирается такой набор пар кадров, который даёт наименьшее суммарное несовпадение меток времени в парах кадров. Асимптотика сложности данного алгоритма $O(|V|(|E| + \ln |V|))$, где $|V|$ — количество вершин графа, т.е. количество кадров в очереди обрабатываемой видеопоследовательности, а $|E|$ — количество рёбер графа, т.е. количество пар кадров с близкими значениями меток времени. Для данной задачи, как правило, $|E| \approx |V|$, т.е. сложность описываемого шага $O(|V|^2)$.

4. На основании параметров $\vec{b}, R, D_1, D_2, M_1, M_2$ проводим преобразование ректификации изображений, т.е. преобразуем $F_i(x, y)$ и $F'_i(x, y)$ в, соответственно, ректифицированные $F_i^r(x, y)$ и $F'^r_i(x, y)$. Далее либо с помощью метода BM, либо с помощью более качественного, но и более вычислительно сложного метода SGBM, рассмотренных в обзоре, на основе $F_i^r(x, y)$ и $F'^r_i(x, y)$ производится расчёт диспаритетов $F_i^d(x, y)$, т.е. расчёт смещений в каждой точке левого изображения к соответствующей точке правого изображения. Поскольку известны параметры сцены L и γ , то на основании $F_i^d(x, y)$ производится расчёт высот $F_i^h(x, y)$ наблюдаемых объектов на основе трёхмерной реконструкции сцены. Наконец, с помощью преобразования, полученного на шаге 1, получаем из изображения $F_i^h(x, y)$ изображение $F_i^p(x, y)$ с компенсацией проективных искажений размеров объектов.

5. Далее повторяются шаги 5–10 алгоритма \mathcal{A}_{mono} , за исключением шага 8, поскольку для карты высот фильтрацию максимумов проводить не требуется.

Приведённый алгоритм будем обозначать \mathcal{A}_{stereo} . На основе алгоритмов \mathcal{A}_{mono} и \mathcal{A}_{stereo} возможно создать комбинированный алгоритм \mathcal{A}_{comb} , приведённый ниже.

1. На последовательности кадров $\{F_i(x, y)\}$ запускается \mathcal{A}_{mono} . В результате получается набор объектов \mathbf{A}^{mono} .

2. На последовательности кадров $\{F_i(x, y)\}$ и $\{F'_i(x, y)\}$ запускается \mathcal{A}_{stereo} . В результате получается набор объектов \mathbf{A}^{stereo} .

3. Конечный результат \mathbf{A} формируется на основе \mathbf{A}^{mono} и \mathbf{A}^{stereo} по следующему принципу: в множество \mathbf{A} попадают те объекты из \mathbf{A}^{mono} и \mathbf{A}^{stereo} , которые достаточно долго, на протяжении n_I кадров подряд, были близки по положению. Если на текущем кадре существует два почти одинаковых по положению объекта из \mathbf{A}^{mono} и \mathbf{A}^{stereo} , то в конечный результат включается только положение одного объекта из \mathbf{A}^{stereo} . Приоритет отдан алгоритму \mathcal{A}_{stereo} , поскольку, как показано далее, он работает более достоверно.

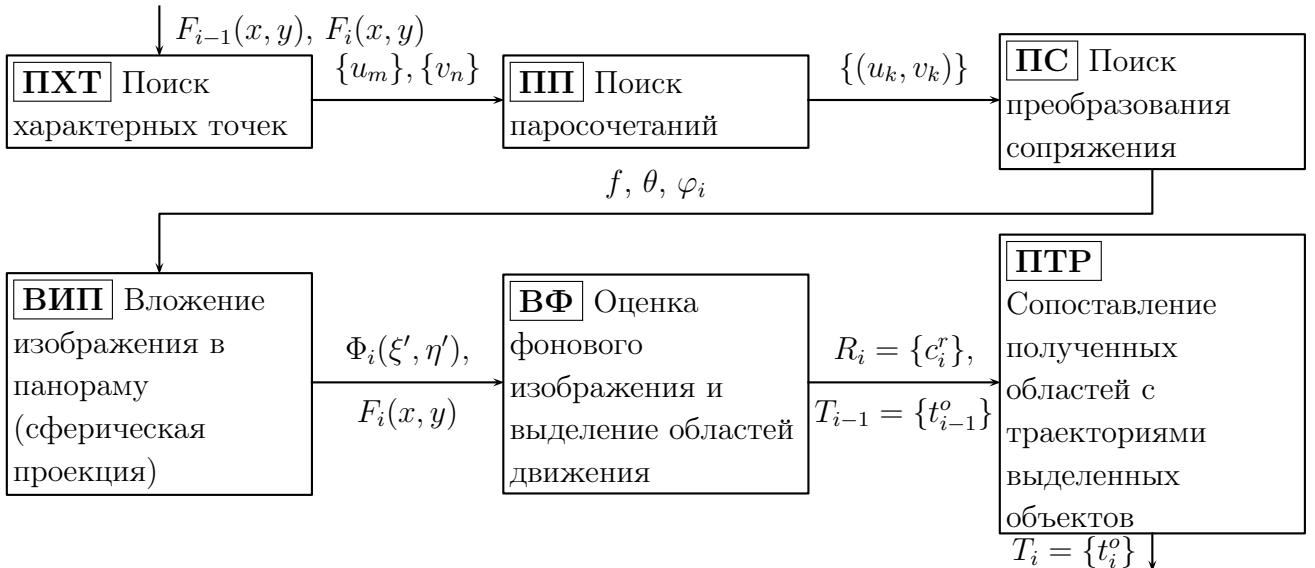


Рисунок 3. Граф-программа построения панорамного изображения и выделения на нём движущихся объектов (ПХТ — поиск характерных точек, ПП — поиск пар точек, ПС — поиск преобразования сопряжения, ВИП — вложение изображения в панораму, ВФ — выделение и вычитание фона, ПТР — поиск траекторий движущихся объектов).

Для сравнения эффективности алгоритмов \mathcal{A}_{mono} , \mathcal{A}_{stereo} и \mathcal{A}_{comb} необходимо провести анализ их работы на одинаковых последовательностях изображений, и сравнить результаты с помощью метрик (12), (13), (14), (15). Результаты тестирования, проведённого на основе теории контроля, изложенной в диссертации, для данных трёх алгоритмов показывают, что по качеству сопровождения и достоверности \mathcal{A}_{stereo} значительно опережает \mathcal{A}_{mono} , а \mathcal{A}_{comb} показывает несколько лучшие результаты по сравнению с \mathcal{A}_{stereo} , за исключением показателя E^α .

Таким образом, дополнительный признак изображений, а именно — глубина, получаемая с помощью стереокамер, позволяет существенно улучшить результаты сопровождения движущихся объектов. Тривиальная комбинация подходов к сопровождению объектов с помощью стерео- и обычных изображений позволяет дополнительно улучшить приведённые показатели.

Для эффективного вычисления глубины на стационарных стереокамерах предложено ускорить алгоритм SGBM путём вычисления диспаратности только в тех точках, в которых один из алгоритмов вычитания фона фиксирует движение, что позволяет существенно ускорить расчёт для сцен с небольшим количеством движения в поле зрения видеокамер.

Шестая глава посвящена исследованию алгоритмов семантической обработки панорамных изображений, получаемых автоматически управляемыми поворотными видеокамерами в режиме постоянного сканирования зоны обзора. Проведённый обзор показал, что описанные в литературе алгоритмы построения панорамных изображений обладают асимптотической вычислительной сложностью $O(n^2)$, где n — количество изображений, для которых

необходимо построить панораму, и это является основной причиной того, что построение панорамы для последовательности изображений, получаемой от поворотных камер в режиме панорамного сканирования, не может быть выполнено в режиме реального времени, т.к. уже для 10 изображений построение панорамы выполняется в среднем 1 секунду за кадр при частоте поступления кадров от камеры 25 кадров в секунду (один раз в 40 мс).

В работе впервые представлен алгоритм, при котором дополнение панорамы новым изображением последовательности осуществляется за $O(1)$, см. схему алгоритма в виде графа на рис. 3. Граф состоит из вершин с сокращёнными названиями процедур: ПХТ — поиск характерных точек $\{u_m\}$ и $\{v_n\}$ на последовательных изображениях $F_{i-1}(x, y)$ и $F_i(x, y)$, соответственно, ПП — поиск пар одинаковых точек $\{(u_k, v_k)\}$ среди найденных на предыдущем этапе на соседних кадрах видеопотока, ПС — поиск преобразования сопряжения вновь поступившего изображения с предыдущим, в результате которого устанавливаются параметры камеры: фокусное расстояние f , угол наклона θ и угол поворота φ_i для текущего кадра $F_i(x, y)$, ВИП — вложение изображения в панораму $\Phi_i(\xi, \eta')$, построенную к данному моменту и представляющую собой сферическую проекцию изображений видеокамеры на пространство угловых координат (ξ, η') , ВФ — выделение и вычитание фона, формирующее набор областей движения $R_i = \{c_i^r\}$, где c_i^r является связанным множеством точек движения на кадре, ПТР — поиск траекторий движущихся объектов, формирующий описание траекторий объектов $T_i = \{t_i^o\}$ для i -го кадра на основе анализа областей движения R_i .

При обработке кадров F_{i-1} и F_i процедуры ПХТ, ПП и ПС не требуют информации, полученной в ходе выполнения остальных процедур, что позволяет применить параллельно-конвейерный подход к ускорению обработки панорамных изображений. Время выполнения процедур построения панорамы на одном ядре современного компьютера архитектуры Intel Core i7 отображено в таблице 1. Таблица составлена с учётом оптимизации процедур ПХТ и ПС: в качестве характерных точек и их информационного описания используется быстрый подход ORB вместо более устойчивого к шуму и аффинным искажениям, но при этом более сложного метода SIFT, кроме того, при подборе преобразования сопряжения существенно используется то, что угол наклона θ , угол зрения камеры, определяемый фокусным расстоянием f и скорость изменения горизонтального угла $\dot{\varphi}$ при горизонтальном сканировании постоянны, и измерены в процедуре автоматической калибровки перед запуском обработки панорамного изображения. Без данного ускорения, предложенного в работе, процедуры ПХТ и ПС выполняются дольше более, чем в 2 раза. Анализ данных таблицы 1 показывает, что суммарно процедуры ПХТ, ПП и ПС, зависящие только от двух кадров F_i и F_{i-1} , вы-

Сокращённое название процедуры	ПХТ	ПП	ПС	ВИП	ВФ	ПТР
Время, мсек	20	20–30	40–60	5	5	5–10

Таблица 1. Времена обработки данных отдельными процедурами анализа панорамных изображений при выполнении с использованием одного ядра процессора Intel Core i7.

полняются на одном ядре 80–110мс, а остальные процедуры ВИП, ВФ и ПТР, результаты работы которых зависят от всей последовательности поступающих кадров, суммарно выполняются за 15–20мс, то есть в 3–5 раз быстрее, чем независимые процедуры.

Обозначим алгоритм выполняемых последовательно процедур ПХТ, ПП и ПС символом \mathcal{P} , а алгоритм выполняемых последовательно процедур ВИП, ВФ, ПТР символом \mathcal{T} . Если трактовать алгоритм как функцию, которую можно вычислить по входным данным за конечное число шагов, то имеет место

$$(\varphi_i, \{u_i\}, \{v_i\}) = \mathcal{P}(F_{i-1}, F_i, \{v_{i-1}\}), \quad \{u_i\} = \{v_{i-1}\}, \\ (\Phi_i, T_i, S_i) = \mathcal{T}(\varphi_i, F_i, \Phi_{i-1}, T_{i-1}, S_{i-1}),$$

где S_i — некоторое внутреннее состояние алгоритма \mathcal{P} , связанное, в частности, с данными о накопленном фоновом изображении. Стадией конвейера будем считать трёхкратное вычисление приведённых выше функций для данных с индексами $i = 3s + p$, где s — номер стадии, а p — целочисленный индекс операции внутри стадии от 0 до 2. Таким образом, в течении одной стадии три конвейера параллельно выполняют алгоритм \mathcal{P} , и один конвейер последовательно выполнит три операции \mathcal{T} .

Для того, чтобы все конвейеры, выполняющие алгоритмы \mathcal{P} и \mathcal{T} , были постоянно загружены, необходимо ввести отставание обработки данных конвейером, выполняющим \mathcal{T} , на одну стадию по сравнению с тремя конвейерами, выполняющими алгоритм \mathcal{P} . Тогда данные $\varphi_{3(s-1)+p}$ с предыдущей стадии конвейеров, выполняющих \mathcal{P} , поступают на вход текущей стадии конвейера, выполняющего операцию \mathcal{T} , и одновременно с этим три конвейера, выполняющие \mathcal{P} , начнут готовить данные φ_{3s+p} для работы следующей стадии. Графическая схема описанной работы конвейера представлена на рис. 4. Отметим, что результат с траекториями выделенных движущихся объектов T_i будет выдан конвейером с отставанием по сравнению с поступлением кадров с изображением объектов, так как для загрузки конвейера следует ввести очередь ожидания обработки кадров с длиной как минимум 3. В тот момент, когда новый кадр поступает в данную очередь, результат обработки будет готов только для кадра с меньшим на 3 индексом. При частоте поступления кадров 25 к/сек задержка составит около 160 мсек, что сопоставимо со средним временем самой быстрой моторной реакции у людей в ответ на какие-либо нервные раздражители, и, следовательно, такая задержка не должна

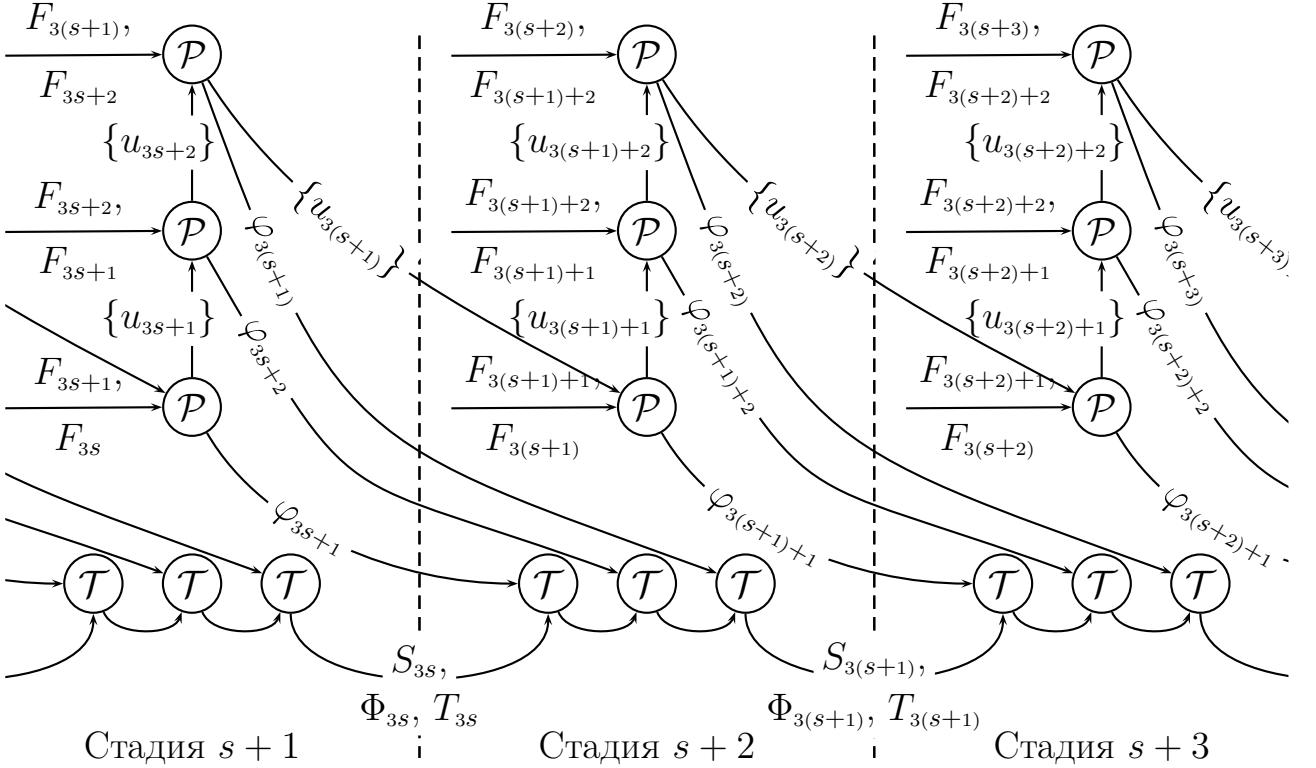


Рисунок 4. Схема параллельно-конвейерного построения и семантической обработки панорамных изображений.

создавать ощущимого дискомфорта для работы операторов систем видеонаблюдения. Конвейер, схема которого представлена на рис. 4, за одну стадию обрабатывает 3 кадра одновременно. При этом конвейер, последовательно выполняющий 3 операции \mathcal{T} , будет действовать 45–80 мсек в течение одной стадии, а каждый из трёх конвейеров, выполняющих операцию \mathcal{P} , будет действовать 80–110 мсек в течение одной стадии, то есть конвейер \mathcal{T} будет, как правило, недозагружен. При этом каждая стадия обрабатывает 3 кадра подряд, поэтому среднее время обработки одного кадра представленным конвейером будет 30–40 мсек, что позволяет обрабатывать изображения с частотой 25 кадров в секунду, если бы кроме семантической обработки не было необходимости выполнять другие операции, например, архивирование видеосигнала и декомпрессию изображений перед обработкой при получении потока видеофрагментов от сетевых камер.

Для дальнейшего увеличения быстродействия можно рассматривать конвейерный процесс обработки как набор асинхронных процессов \mathcal{P} и \mathcal{T} . Синхронизация между ними будет осуществляться на уровне очереди из входных данных: кадров F_i для процессов \mathcal{P} и углов φ_i для процессов \mathcal{T} . Если очередь из входных данных на одном из конвейеров сильно выросла, значит, пора пропустить кадры в очереди на обработку, чтобы сократить разность скорости выдачи выходных данных по сравнению со скоростью входных данных. Если конвейеры, выполняющие операцию \mathcal{P} , являются узким местом,

то вместо тяжёлой обработки изображения для оценки угла φ_i можно использовать формулу $\varphi_i = \varphi_{i-1} + \dot{\varphi}(t_i - t_{i-1})$, при условии, что для камеры проведена калибровка и время кадров t_i, t_{i+1} известно достаточно точно. Это позволит исключить кадр F_i из обработки довольно длительной операцией \mathcal{P} , при этом, тем не менее, получить оценку угла поворота на основе интерполяции. Совсем отказаться от операции \mathcal{P} нельзя, т.к. с её помощью оценка угла поворота более точная, чем с помощью интерполяции.

В седьмой главе приводится экспериментальное исследование эффективности разработанных алгоритмов сопряжения, синхронизации, управления и семантического сжатия изображений, а также описывается структура приборных комплексов технического зрения РАЯЖ 46652.001-ОС.ПЗ, ВАРШ.00001-01, ВАРШ.00003-01, ЛЦКБ.464412.002, ВАРШ.466452.002, разработанных с использованием полученных в диссертационной работе результатов и получивших широкое внедрение, в том числе при непосредственном участии автора, для контроля безопасности на таких стратегически важных объектах как аэропорты, гидро- и атомные электростанции, а также на объектах городской и транспортной инфраструктуры.

Для практической оценки эффективности систем видеонаблюдения проведено сравнительное испытание систем ВАРШ.00001-01 с неподвижными камерами и с панорамной камерой, а также системы ЛЦКБ.464412.002 с радиолокационным датчиком и системы ВАРШ.00003-01 с тепловизором, установленным на поворотном устройстве и сканирующим зону контроля в режиме шагового перемещения между тремя положениями поворотной платформы. Настройка всех перечисленных выше четырёх систем технического зрения была проведена таким образом, что зона контроля всех четырёх систем включала в себя одну и ту же область, в которой проводились испытания. Для сравнения использовалась специальная методика, разработанная в диссертационной работе, суть которой заключается в следующем. Пусть сравниваемые системы технического зрения пронумерованы от 1 до N . В процессе испытания для каждой из систем измеряются некие показатели качества её работы x_1, x_2, \dots, x_N . Рангом i -ой системы по показателю x мы будем обозначать величину $rg(i, x)$, равную порядковому номеру показателя системы i в последовательности, отсортированной по возрастанию показателя качества x . Например, если $x_1 = 1, x_2 = 3, x_3 = 1,5$ для трёх испытываемых систем, и увеличение показателя обозначает ухудшение качества, то $rg(1, x) = 3, rg(2, x) = 1, rg(3, x) = 2$. Если после упорядочивания показателей несколько значений равны друг другу, то ранг присваивается всем соседям одинаковый, равный среднему значению индексов позиций с одинаковыми значениями показателя. Т.е., если $x_1 = 2, x_2 = 2, x_3 = 2$, то $rg(1, x) = rg(2, x) = rg(3, x) = 2$, а если $x_1 = 1, x_2 = 2, x_3 = 2$, то

Показатель	Видеокамеры		Радар		Тепловизор		Панорама	
	кол-во	ранг	кол-во	ранг	кол-во	ранг	кол-во	ранг
Дневные ложные срабатывания	0	1,5	1	3	0	1,5	3	4
Ночные ложные срабатывания	1	4	0	2	0	2	0	2
Дневные пропуски	0	1	8	2	11	4	9	3
Ночные пропуски	1	1,5	1	1,5	11	3	30	4
Объекты вне поля зрения камер	0	1,5	1	3	16	4	0	1,5
Средний ранг	—	1,9	—	2,3	—	2,9	—	2,9

Таблица 2. Ранговая таблица показателей эффективности систем технического зрения.

$$rg(1, x) = 1, rg(2, x) = rg(3, x) = 2,5.$$

Для каждой системы были измерены 5 показателей: количество ложных срабатываний и количество пропущенных объектов за 20 минут в светлое время суток, те же показатели в тёмное время суток, и количество объектов, которые не попали в поле зрения видеокамер. По каждому из этих показателей был посчитан ранг, а затем для каждой системы был получен средний ранг. Результаты испытаний приведены в таблице 2. Анализ таблицы показывает, что панорамный детектор объектов хуже всего приспособлен для охраны объектов, поскольку вочных условиях он пропускает более половины событий, происходящих в поле зрения камеры. С другой стороны, плохая работа детектора обусловлена повышенными требованиями к чувствительности камеры: из-за того, что в ночное время сильно повышается экспозиция, изображения объектов становятся смазанными и не отличимыми от фона. Поэтому при соответствующем оборудовании панорамная система будет работать лучше тепловизионного шагового локатора, поскольку данный локатор пропускает очень много объектов из-за того, что они не попадают в область обзора или находятся в ней очень недолго. Анализ пропусков радарной системы показывает, что система из неподвижных видеокамер и радарная система сравнимы друг с другом по достоверности детектирования объектов, если видеосистема используется в приемлемых для неё условиях наблюдения (не далее 300м от видеокамеры). 8 пропусков, указанных в таблице для радарной системы, обусловлены либо несовершенством алгоритма генерации нового события в условиях плотного движения, когда несколько объектов формируют только одну радарную отметку на одном проходе луча, либо узким вертикальным углом диаграммы направленности радара, что приводит к пропускам на склоне непосредственно возле радара, поэтому в рамках паспортных ограничений у системы с видеокамерами нет больших преимуществ перед радарными детекторами, и наоборот, на расстояниях, при которых ат-

Алгоритмы	$F_P^{\alpha\beta}$	$F_N^{\alpha\beta}$	Q^α	Время обработки кадра, мс
2002	>10000	20	60%	80
2003	≈ 500	15	60%	40
2008	50	10	55%	50
2012	20	10	50%	50
2012 с фильтром оптического потока	10	12	47%	55

Таблица 3. Эволюционное улучшение показателей алгоритмов семантической обработки по мере совершенствования комплексов «Orwell2k» с помощью системы тестирования «Супервизор».

мосферные явления не вызывают помех в поле зрения видеокамер, системы детектирования объектов на основе семантической обработки видеоизображений являются надёжным средством автоматической фиксации событий в поле зрения видеокамер.

Разработанная в диссертационной работе теория тестирования широко и длительно использовалась для повышения эффективности алгоритмов семантической обработки. В таблице 3 приведены введённые ранее показатели (12), (13), (14), рассчитанные для тестового набора видеофайлов общей длительностью около 2 часов, содержащей примерно 150 объектов. Для иллюстрации эволюционного улучшения алгоритмов семантической обработки с помощью системы тестирования в таблице отображено постепенное совершенствование используемых в системах «Orwell2k» базовых алгоритмов семантической обработки (2002 год) с помощью последовательного внедрения новых элементов анализа изображений: нормализация изменений яркости во время перемены внешнего освещения сцены (2003), фильтры согласованного движения характерных точек изображения (2008), фильтры анализа подвижных границ и направления оптического потока (2012). Без системы автоматического тестирования было бы невозможно оптимальным образом подобрать параметры фильтрации, поэтому тестирование является неотъемлемой частью разработки алгоритмов семантической обработки. Например, результаты, полученные для счётчиков ВАРШ.468196.001 и ВАРШ.463135.002-01, измерены с помощью системы тестирования «Супервизор», в которой внедрена теория, разработанная в диссертации. В частности, с помощью системы тестирования было выявлено влияние синхронизации изображений от стереокамер на конечную точность результата подсчёта объектов в поле зрения стереосистемы: если точность детектирования движущихся объектов при синхронной обработке находится на уровне 95%, то при сдвиге на кадр (что соответствует 40мс при частоте 25к/с) результат подсчёта на 87% совпадает с действительным, а при сдвиге на 2 кадра (80мс) — на 72%.

Методика измерения погрешности сопряжения поворотных и неподвижных видеокамер, разработанная автором диссертации, основана на оценке достоверности сопоставления изображений: в процессе функционирования модулей обработки видеокамер накапливается статистика ошибок сопоставления изображений объектов либо на двух сопряжённых неподвижных видеокамерах, либо промахов при наведении слейвов на выделенные объекты обработчиками сопряжённых с ними мастеров. Если p — доля ошибок сопоставления объектов в ходе эксперимента, то тогда погрешность сопряжения можно оценить как $\sigma = \frac{3\Delta}{F^{-1}(1-p/2)}$, где Δ — предельное значение разброса координат, при котором фиксируется промах, то есть это либо ширина объекта в пикселях при измерении погрешности сопряжения мастеров, либо угол зрения слейва при наведении на объекты, а $F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx$, то есть F^{-1} — функция ошибок. Значения Δ должны быть примерно одинаковыми для всех сопоставлений в ходе эксперимента. В результате сбора статистики можно сравнить работу нескольких алгоритмов наведения: алгоритм наведения без учёта упреждения, алгоритм наведения с линейным упреждением и разработанный в диссертации алгоритм наведения с адаптивным временем наведения; а также алгоритм сопоставления изображений сопряжённых мастеров при синхронизации данных между обработчиками мастеров, запущенных на одной ЭВМ и при синхронизации посредством ЛВС между обработчиками мастеров, запущенных на разных ЭВМ. При применении алгоритма наведения без упреждения доля промахов в условиях эксперимента получилась 25% при погрешности наведения 7° , при применении наведения с линейным упреждением доля промахов 5% и погрешность наведения 4° , при применении разработанного в диссертации алгоритма наведения с адаптивным временем доля промахов 1% и погрешность наведения 3° . При сопоставлении изображений, выделенных обработчиками на одной ЭВМ, доля неверных сопоставлений 1% и погрешность равна 6 пикселям, при синхронизации положения объектов по ЛВС доля ошибочных сопоставлений 10% и погрешность преобразования сопряжения равна 10 пикселам.

Измерение времени выполнения программных функциональных блоков систем типа РАЯЖ 46652.001-ОС.ПЗ, ВАРШ.00001-01, ВАРШ.00003-01 показало, что время реакции системы на появление нового объекта в зоне наблюдения камер определяется временем автоматического наведения поворотной камеры на объект и равно 0,5с, что в 20 раз меньше времени реакции систем, использующих ручное наведение.

В заключении приведены основные выводы по работе и описана практическая значимость достигнутых результатов.

В приложения включены протоколы испытаний системы, в том числе проведённые на Курской АЭС, Бурейской ГЭС и в Международном аэро-

порту «Шереметьево», фрагменты кодов основных программных модулей на основе алгоритмов, разработанных в диссертации, свидетельства о регистрации программ и перечень патентов, включающих результаты диссертации, а также список выставок, на которых проводилась апробация результатов работы.

Основные результаты и выводы

При выполнении диссертационной работы в рамках решения поставленной научной проблемы по повышению эффективности приборных комплексов технического зрения на основе алгоритмов семантического сжатия, синхронизации и сопряжения семантических обработчиков и автоматического управления видеоприборами достигнуты следующие результаты.

1. Впервые разработана теория и выведены уравнения автоматизации начальной установки и последующего параллельного сопряжённого функционирования приборов семантической обработки изображений в комплексах технического зрения.

2. Выведены математические соотношения для сопряжения приборов технического зрения и разработаны новые алгоритмы сопоставления изображений в общей зоне контроля приборов и управления поворотными камерами с адаптивным упреждением, которые обеспечивают достоверность сопоставления изображений 99,0%, малые погрешности наведения для неподвижных ($0,5^\circ$) и движущихся (3°) объектов, а также повышение разрешения в 10 раз при значительном уменьшении количества приборов в комплексах технического зрения с обширными областями контроля, причём время автоматического наведения в 20 раз меньше времени ручного наведения.

3. Разработана теория тестирования и контроля алгоритмов семантического сжатия с применением формальной логики и теории множеств, и на основе этой теории построен прототип системы тестирования.

4. Создан и внедрён новый комбинированный алгоритм обработки моно- и стереоизображений для оценки плотности движения наблюдаемых объектов с погрешностью 3%.

5. Созданы новые алгоритмы семантического анализа панорамных изображений, позволяющие уменьшить в 3–10 раз количество видеоприборов в приборных комплексах технического зрения за счёт поворотных устройств.

6. Разработанные в диссертации алгоритмы и способы позволили создать приборные комплексы технического зрения нового поколения с параметрами выше мирового уровня, обеспечивающие автоматическое обнаружение, наведение и сопровождение объектов, значительное улучшение качества их визуализации, классификации и идентификации и автоматическое формирование

сигналов управления устройствами предупреждения и заграждения.

7. Результаты диссертационной работы применены в программно-аппаратных комплексах «Orwell2k», разработанных при непосредственном участии автора, что подтверждается рядом свидетельств о регистрации программ и патентами на полезные модели и изобретения, и используются для эффективного решения задач обеспечения безопасности, например, в аэропортах «Шереметьево», «Домодедово», «Пулково», «Череповец», на крупных промышленных предприятиях и объектах стратегического значения, например, Бурейской ГЭС, Курской АЭС, нефтеперерабатывающем заводе Оренбургской области, в городах Нижний Новгород, Саранск, Сочи при контроле общественных территорий, на железной дороге Сочи – Красная поляна (со времени проведения зимних Олимпийских игр 2014 года) и на других важных объектах государственной инфраструктуры.

Таким образом, представленная диссертация является законченной научно-квалификационной работой, в которой в рамках приоритетных направлений развития науки, технологий и техники, утверждённых указом Президента РФ от 7 июля 2011 г. №899, автором разработаны теоретические положения и научно обоснованные технические решения, а также созданы высокоэффективные алгоритмы, программные средства и приборные комплексы технического зрения, внедрение которых вносит значительный вклад в развитие экономики страны и в повышение её обороноспособности.

Список работ, опубликованных по теме диссертации

Публикации в журналах, входящих в перечень ВАК

1. А.В. Хамухин, Я.Я. Петричкович. Принцип семантического сжатия в алгоритмах обработки видеинформации. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2015, с. 82–85.
2. А.В. Хамухин. Разработка и анализ эффективных алгоритмов для достоверной идентификации номерных знаков транспортных средств. М.: «Оборонная техника», №8-9, 2012, с. 26–29.
3. А.В. Хамухин. Повышение быстродействия компьютерных видеосистем при обнаружении подвижных объектов. М.: «Оборонная техника», №8-9, 2012, с. 30–32.
4. А.В. Хамухин, Я.Я. Петричкович. Анализ современного состояния разработок детекторов изменения изображений видеокамер компьютерных систем видеонаблюдения и сопровождения. М.: «Оборонная техника», №7, 2015, с. 67–70.
5. А.С. Малистов, А.А. Солохин, А.В. Хамухин. Формальный подход к оценке качества алгоритмов обработки изображений в интеллектуальных си-

стемах видеонаблюдения. Журнал «Вопросы радиоэлектроники», серия «общетехническая(ОТ)», выпуск 2, ОАО ЦНИИ «Электроника», 2006.

6. А.Х. Ахриев, А.С. Малистов, А.В. Хамухин, П.А. Александров. Комплексный подход к созданию систем автоматического видеонаблюдения и видеоконтроля на объектах высокой сложности типа ITER и атомных станций. «Вопросы атомной науки и техники», серия «Термоядерный синтез», 2006, вып. 3., с. 69–81.

7. А.В. Хамухин, Я.Я. Петричкович. Анализ изображений с использованием стандарта OpenVX. М.: «Оборонная техника», №7, 2015, с. 71–77.

8. А.В. Хамухин, Я.Я. Петричкович. Принцип семантического сжатия в алгоритмах ускоренной обработки видеинформации. М.: «Оборонная техника», №7, 2015, с. 64–68.

9. А.В. Хамухин. Управление видеокамерами при сопровождении движущихся целей. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 233–234.

10. А.В. Хамухин. Алгоритм фокусировки для объективов с управляемым увеличением. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 216–217.

11. А.В. Хамухин. Калибровка охранных камер видеонаблюдения на поворотной платформе с использованием стереоэффекта. М.: «Оборонная техника», №7, 2015, с. 81–82.

12. А.В. Хамухин. Алгоритм сопоставления идентичных объектов на изображениях видеокамер с общей зоной обзора. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 214–215.

13. А.В. Хамухин. Алгоритмы калибровки видеокамер для фотометрической оценки размеров объектов. М.: «Компания Спутник+», «Естественные и технические науки», №6, 2013, с. 329–332.

14. А.В. Хамухин. Калибровка видеокамер по плану местности с помощью измерения дальности до топографических ориентиров. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 221–222.

15. А.В. Хамухин. Вычисление линии горизонта на изображении видеокамеры. М.: «Компания Спутник+», «Естественные и технические науки», №5, 2007, с. 197–198.

16. А.В. Хамухин. Преобразование координат с изображения камеры на топографический план зоны обзора. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 228–229.

17. А.В. Хамухин. Моделирование поверхности земли при преобразовании координат с изображения камеры на топографический план зоны обзора.

М.: «Компания Спутник+», «Естественные и технические науки», №5, 2007, с. 199–200.

18. А.В. Хамухин. Статистическая оценка погрешности сопряжения видеокамер на основе достоверности сопоставления изображений. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 230–232.

19. А.В. Хамухин. Определение линии горизонта на изображении видеокамеры на основе привязки к плану местности. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 226–227.

20. А.В. Хамухин. Классификация объектов с двумя признаками на основе триангуляции. М.: «Компания Спутник+», «Естественные и технические науки», №5, 2007, с. 81–82.

21. А.В. Хамухин, Я.Я. Петричкович. Теоретико-множественный анализ посещаемости на основе решающих правил в пространстве статических и динамических признаков изображений зрительных залов. М.: «Компания Спутник+», «Естественные и технические науки», №11-12, 2014, с. 290–293.

22. А.В. Хамухин. Анализ методов тестирования алгоритмов автоматической обработки изображений. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 218–220.

23. А.В. Хамухин. Архитектура информационных систем хранения видеоданных для разработки алгоритмов компьютерного зрения. М.: «Компания Спутник+», «Естественные и технические науки», №6, 2013, с. 336–337.

24. А.В. Хамухин. Алгоритм синхронизации видеопотоков в стереосистемах. М.: «Компания Спутник+», «Естественные и технические науки», №6, 2013, с. 333–335.

25. А.В. Хамухин. Калибровка камер на поворотной платформе с использованием стереоэффекта. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 223–225.

26. А.В. Хамухин, Я.Я. Петричкович. Теоретико-множественный анализ посещаемости занятий личным составом на основе решающих правил в пространстве статических и динамических признаков изображений учебно-тренировочной базы. М.: «Оборонная техника», №7, 2015, с. 78–80.

27. А.В. Хамухин. Управление положением поворотной видеокамеры посредством компьютерного анализа изображений. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2014, с. 235–236.

28. А.В. Хамухин, Я.Я. Петричкович. Особенности преобразования сопряжения координат панорамного изображения с координатами плана местности. М.: «Компания Спутник+», «Естественные и технические науки», №11-12, 2014, с. 287–289.

29. А.В. Хамухин. Выделение движущихся объектов на поворотных видеокамерах в режиме панорамного сканирования. М.: «Компания Спутник+», «Естественные и технические науки», №1, 2015, с. 96-98.

Свидетельства о регистрации программ, полезные модели и изобретения

30. С.Т. Иванченко, И.А. Кан, К.В. Лунин, А.С. Малистов, Я.Я. Петричкович, А.А. Солохин, В.П. Сомиков, А.В. Хамухин. Система обеспечения безопасности и мониторинга мобильных объектов. Патент РФ №2265 531, бюл. №34, 2005.

31. И.А. Кан, К.В. Лунин, А.С. Малистов, Я.Я. Петричкович, А.А. Солохин, В.П. Сомиков, А.В. Хамухин. Система и способ автоматизированного видеонаблюдения и распознавания объектов и ситуаций. Патент РФ №2268497, бюл. №02, 2006.

32. Я.Я. Петричкович, А.В. Хамухин, А.С. Малистов, А.А. Солохин и др. Устройство автоматизированного контроля обстановки в зрительных залах. Патент РФ №2296434, бюл. №9, 2007.

33. К.В. Лунин, А.С. Малистов, Я.Я. Петричкович, А.А. Солохин, В.П. Сомиков, А.В. Хамухин. Система автоматизированного видеонаблюдения и распознавания объектов и ситуаций «Orwell2k». Свидетельство об официальной регистрации программы для ЭВМ №2003612604, 2003г.

34. А.В. Хамухин, Я.Я. Петричкович, А.Н. Ракутин, С.Ю. Миллер, О.О. Колесников и др. Автоматическая система фотовидеофиксации нарушений правил дорожного движения «Перекрёсток». Свидетельство о государственной регистрации программы для ЭВМ №2014613905.

35. А.В. Хамухин, С.Ю. Миллер, Д.О. Соколов, А.М. Ерипилов, А.М. Глазков и др. Программный комплекс «Orwell 2k-base». Свидетельство о государственной регистрации программы для ЭВМ 2012615283.

36. А.В. Хамухин, А.С. Малистов. Программный комплекс «TraVio SDK». Свидетельство о государственной регистрации программы для ЭВМ 2012617248.

37. А.В. Хамухин, Д.Е. Кадочкин, Е.С. Иванченко, К.В. Лядвинский, А.М. Глазков, С.Ю. Миллер и др. Программный комплекс «Orwell 2k-touch». Свидетельство о государственной регистрации программы для ЭВМ 2012618571.

38. А.В. Хамухин, Анд.В. Хамухин, А.Е. Баскаков, В.В. Галкин, А.С. Малистов и др. Программный комплекс «Statistics». Свидетельство о государственной регистрации программы для ЭВМ 2012619211.

39. А.В. Хамухин, Я.Я. Петричкович, А.Н. Ракутин, О.О. Колесников, А.С. Малистов, А.В. Акимов. Система и способ определения государ-

ственного регистрационного номера транспортного средства. Патент РФ №2014104763.

40. А.В. Хамухин, Я.Я. Петричкович. Способ и система определения скорости транспортных средств. Патент РФ №2015102323.

41. А.В. Хамухин, Я.Я. Петричкович. Система определения скорости транспортных средств. Патент РФ на полезную модель №2015102309.

42. А.В. Хамухин, Я.Я. Петричкович. Способ и система калибровки комплекса измерения скорости транспортных средств. патент РФ №2015102320.

43. А.В. Хамухин, Я.Я. Петричкович. Система калибровки комплекса измерения скорости транспортных средств. Патент РФ на полезную модель №2015102315.

44. А.В. Хамухин, Я.Я. Петричкович, А.Н. Ракутин, О.О. Колесников, А.С. Малистов, А.В. Акимов. Система определения государственного регистрационного номера транспортного средства. Патент РФ на полезную модель №2014104762.

45. А.В. Хамухин, Я.Я. Петричкович, А.Н. Ракутин, О.О. Колесников, Н.Д. Бобков и др. Система и способ автоматической фиксации несанкционированного прохода через зону контроля. Патент РФ №2014152194.

46. С.Т. Иванченко, И.А. Кан, К.В. Лунин, А.С. Малистов, Я.Я. Петричкович, А.А. Солохин, В.П. Сомиков, А.В. Хамухин. Система обеспечения безопасности и мониторинга мобильных объектов. Патент РФ на полезную модель №36315, бюл. №7, 2004.

47. И.А. Кан, К.В. Лунин, А.С. Малистов, Я.Я. Петричкович, А.А. Солохин, В.П. Сомиков, А.В. Хамухин. Система и способ автоматизированного видеонаблюдения и распознавания объектов и ситуаций. Патент РФ на полезную модель №36912, бюл. №9, 2004.

48. Я.Я. Петричкович, А.В. Хамухин, А.С. Малистов, А.А. Солохин и др. Устройство автоматизированного контроля обстановки в зрительных залах. Патент РФ на полезную модель №47546, бюл. №24, 2005.

Публикации в других изданиях

49. А.В. Хамухин. Расчёт ресурсных затрат при применении автоматизированных систем видеонаблюдения. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №4, 2014, с. 283–285.

50. Я.Я. Петричкович, В.П. Сомиков, А.А. Солохин, А.С. Малистов, А.В. Хамухин. Интеллектуальная система телевизионного наблюдения с компьютерным зрением “Orwell 2K”. Журнал «Техника средств связи», серия «Техника телевидения», научно-технический сборник, ЗАО «МНИТИ», вып. 1, 2005, с. 54–57.

51. А.С. Малистов, А.А. Солохин, А.В. Хамухин, Алгоритмы обработки видеопотоков для отделения теней от объектов. XV Международная Интернет-конференция молодых ученых, аспирантов и студентов по современным проблемам машиноведения, тезисы докладов, с. 95, из-во ИМАШ РАН, Москва, 2003.
52. А.В. Хамухин. Системы мониторинга инфраструктуры мегаполисов и анализ состояния парковок. М.: «Компания Спутник+», «Техника и технология», №6, 2012.
53. А.В. Хамухин. Анализ ресурсных затрат при эксплуатации автоматизированных систем видеонаблюдения. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №6, 2013, с. 32–34.
54. А.С. Малистов, А.А. Солохин, А.В. Хамухин. Слежение за целями в мультисенсорных системах видеонаблюдения с компьютерным зрением. Труды XLVIII научной конференции МФТИ «Современные проблемы фундаментальных и прикладных наук», Москва, 2005.
55. А.В. Хамухин. Вычисление угла упреждения при управлении наведением поворотных видеокамер на движущуюся цель. М.: «Компания Спутник+», «Техника и технология», №5, 2007, с. 34–35.
56. А.В. Хамухин. Алгоритм управления синхронизацией распределённых вычислений в системах видеонаблюдения. М.: «Компания Спутник+», «Техника и технология», №5, 2007, с. 31–33.
57. А.С. Малистов, А.А. Солохин, А.В. Хамухин. Калибровка поворотных камер в системах видеонаблюдения с компьютерным зрением. Избранные труды XVII Международной интернет-конференции молодых ученых, аспирантов и студентов по современным проблемам машиноведения, с. 159–162, Москва, 2005.
58. А.В. Хамухин. Вычисление реперной точки юстировки оптической оси видеокамеры по нескольким измерениям дальности до топографических объектов на местности. М.: «Компания Спутник+», «Аспирант и соискатель», №5, 2007, с. 161–162.
59. А.В. Хамухин. Измерение зависимости координат увеличения и фокусировки от кратности увеличения для объективов с координатным управлением. М.: «Компания Спутник+», «Аспирант и соискатель», №5, 2007, с. 163–164.
60. А.В. Хамухин. Алгоритм управления положением поворотной видеокамеры при смещении центра изображения на заданный вектор. М.: «Компания Спутник+», «Аспирант и соискатель», №5, 2007, с. 159–160.

61. А.В. Хамухин. Критерий идентичности объектов на изображениях видеокамер с общей зоной обзора. М.: «Компания Спутник+», «Техника и технология», №5, 2007, с. 36–37.
62. А.С. Малистов, А.А. Солохин, А.В. Хамухин. Методы оценки эффективности алгоритмов в интеллектуальных системах видеонаблюдения. Труды XLVII научной конференции МФТИ «Современные проблемы фундаментальных и прикладных наук», ч. 3, с. 216–217, Москва, 2004.
63. А.В. Хамухин. Анализ современного состояния разработок детекторов изменений изображений видеокамеры. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №1, 2015, с. 131–134.
64. А.С. Малистов, А.А. Солохин, А.В. Хамухин. Оценка эффективности алгоритмов в системах видеонаблюдения. XVI Международная Интернет-конференция молодых ученых, аспирантов и студентов по современным проблемам машиноведения, тезисы докладов, с. 177, из-во ИМАШ РАН, Москва, 2004.
65. А.В. Хамухин. Оценка объёма тестовой выборки для вычисления ожидаемой доли ошибки при классификации объектов видеонаблюдения. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №6, 2007, с. 169–170.
66. А.В. Хамухин. Калибровка поворотных камер в системах компьютерного зрения. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №4, 2014, с. 279–282.
67. А.В. Хамухин. Система тестирования алгоритмов компьютерного зрения. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №1, 2014, с. 170–174.
68. А.В. Хамухин. Анализ применения стереосистем для решения задач бизнес-мониторинга. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №6, 2013, с. 342–347.
69. А.В. Хамухин. Анализ стерео-, видео- и панорамных изображений с использованием стандарта OpenVX. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №4, 2014, с. 269–278.
70. А.В. Хамухин. Обнаружение движения объектов компьютерными информационными видеосистемами с помощью двухмерных и стереоскопических характеристик изображения. М.: «Компания Спутник+», «Актуальные проблемы современной науки», №1, 2015, с. 135–141.

Отпечатано в АО «ЭЛВИС Неотек».
Формат бумаги 60 × 84 1/16.
Уч.-изд. л. 2.0. Тираж 100 экз.