

Федеральное государственное бюджетное образовательное учреждение
высшего образования «Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»

На правах рукописи



Лоханов Александр Васильевич

АЛГОРИТМЫ РАСПОЗНАВАНИЯ И МОДЕЛИ ЦИФРОВОЙ ОБРАБОТКИ
ДИНАМИЧЕСКИХ ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЙ

Специальность 05.12.04

Радиотехника, в том числе системы и устройства телевидения

ДИССЕРТАЦИЯ

на соискание ученой степени
кандидата технических наук

Научный руководитель:
Давыдов Николай Николаевич,
доктор технических наук, доцент

Владимир – 2021

Содержание

Введение	4
Глава 1 Современные методы и средства распознавания, контроля и обработки телевизионных изображений	14
1.1 Цифровые системы, устройства и технологии контроля и распознавания полутоновых телевизионных изображений	14
1.2 Методы распознавания, модели и особенности построения алгоритмов контроля и обработки телевизионных изображений	21
1.3 Постановка задачи	35
1.4 Выводы по 1-ой главе	37
Глава 2 Исследование и разработка методов и средств информационно-технологического обеспечения процессов идентификации, контроля и обработки телевизионных изображений	38
2.1 Обоснование технологии выбора процесса распознавания и обработки телевизионных изображений	38
2.2 Исследование и разработка информационно-технологической модели и алгоритмов распознавания и обработки телевизионных изображений	40
2.3 Исследование и разработка средств программного обеспечения удаленного контроля, диагностики и обработки телевизионных изображений	45
2.4 Выводы по 2-й главе	58
Глава 3 Исследование и разработка средств технического и методического обеспечения процессов распознавания, контроля, диагностики, испытания и управления видеосистемами.....	59
3.1 Разработка структуры информационно-технических средств визуализации процессов распознавания, контроля, диагностики и испытания видеосистем ...	59

3.2 Разработка средств информационно-программного кодирования для взаимодействия элементов визуализации с элементами интерактивного управления видеосистемами	66
3.3 Разработка методики управления средствами видеоконтроля процессов распознавания	83
3.4 Выводы по 3-ей главе	113
Глава 4 Экспериментальные исследования и внедрение в промышленности алгоритмов и моделей процессов распознавания динамических телевизионных изображений	114
4.1 Результаты экспериментальных исследований средств распознавания, контроля и обработки телевизионных изображений	114
4.2 Внедрение результатов исследования в промышленности и в учебный процесс	133
4.3 Выводы по 4-й главе	135
Заключение	137
Список литературы.....	140
Приложение А. Свидетельства о государственной регистрации программ для ЭВМ (справочное)	161
Приложение Б. Акт о внедрении результатов кандидатской диссертации в учебный процесс (справочное)	165
Приложение В. Акт о практическом использовании результатов диссертационных исследований в инновационной научной деятельности (справочное).....	166
Приложение Г. Дипломы (справочное)	168
Приложение Д. Акты внедрения (справочное)	169

Введение

Актуальность темы исследования, степень её разработанности

Методы и средства цифровой обработки и распознавания изображений относятся к числу одних из наиболее изучаемых и перспективных направлений развития телевизионных систем, используемых в различных отраслях народного хозяйства, и активно модернизируемых с учетом необходимости повышения качества изображения и помехоустойчивости работы видео устройств специального назначения [1].

На современном этапе развития телевизионных систем наблюдается увеличение объема видеoinформации, подлежащей оперативной передаче и аналитическому исследованию, что предопределяет необходимость повышения как скорости обработки видеoinформации, так и достоверности распознавания динамических телевизионных изображений. Например, при передаче цифрового сигнала цветного изображения скорость видеопотока составляет порядка 216 Мбит/с, а при трансляции видеоизображения высокой четкости скорость передачи сигнала достигает 1 Гбит/с [2].

Примером востребованности скоростных методов и средств обработки и достоверного распознавания видеоизображений (образов) являются, как системы ретрансляции и приема программ цифрового телевидения, так и сигнальные и охранные системы, системы распознавания лиц, системы кодирования биометрических данных, системы слежения за подвижными объектами и субъектами, системы дорожно-патрульной службы для контроля, регистрации и распознавания дорожно-транспортной обстановки, системы мониторинга запрещенного видео контента информационных платформ Интернета, медицинские диагностические системы, системы видео регистрации процессов электронной микроскопии и быстропротекающих в реальном масштабе времени научных экспериментов и др. При этом под образами понимаются не только регистрируемые динамические телевизионные видеоизображения, но и

структурированное описание изучаемого объекта, субъекта или явления, представленное вектором признаков, каждый элемент которого характеризуется числовым значением параметров, отображающих конкретные свойства образа.

Наряду с этим и в связи с распространением угрозы террористической опасности и продвижением в преступной среде противоправных действий с использованием биометрических данных, возрастает актуальность исследования и разработки методов и средств оперативной и достоверной цифровой обработки динамических изображений и распознавания образов в процессе телевизионного мониторинга городской среды. В контексте данной проблемы ведутся теоретические и прикладные научные исследования, основанные на использовании биометрических технологий и направленные на автоматизацию процессов обнаружения и опознания опасных субъектов в местах массового скопления и хаотичного перемещения людей, и, в первую очередь, в таких местах, как вокзалы, аэропорты, метрополитен и торговые центры [3-8].

Под обработкой телевизионных изображений понимается не только стремление к улучшению их зрительского восприятия, но и процедура классификации видео субъектов, выполняемая в ходе анализа динамических изображений. При этом особое внимание уделяется вопросу обеспечения достоверности автоматизированного распознавания личности человека по биометрическим признакам относительно эталонных снимков за счет снижения ошибок 1 рода "False Rejection Rate" (FRR – вероятность отказа «своему») и ошибок 2 рода "False Acceptance Rate" (FAR – вероятность пропуска «чужого»).

Создание подобных систем распознавания отличается определенной сложностью их воспроизведения, так как проблемы, в основном, группируются в двух принципиально важных направлениях совершенствования данных систем:

- во-первых, в инженерно-технической области, - в связи с тем, что в обязательном порядке необходимо обеспечить надлежащую вычислительную мощность и высокую производительность ключевых функциональных узлов и блоков системы, а также следует обеспечить высокую разрешающую способность и чувствительность оптических модулей при регистрации изображений;

- во-вторых, в алгоритмической области, - что обусловлено необходимостью повышения эффективности функционирования системы за счет модернизации алгоритмов детектирования и распознавания образов, в том числе, при использовании узкоспециализированных алгоритмов (типа «только лица»), основанных на методе сравнения с изменяющейся мерой в виде набора (класса) как эталонных, так и случайных динамических видеоизображений или снимков субъекта.

Кроме того, подлежащие решению научные задачи характеризуются, в определенной степени, свойством двойственности исходных данных, а именно:

- при наличии исчерпывающей совокупности эталонных наборов необходимо оценить степень схожести некоторого множества регистрируемых тестовых классов с имеющимися эталонами;

- при наличии ограниченной совокупности эталонов требуется достоверно констатировать факт принципиального отличия регистрируемых тестовых классов от доступных эталонных наборов.

Особый вклад в развитие методов и средств цифровой обработки телевизионных изображений и распознавания образов внесли советские и российские ученые и их ученики Бакут П.А., Галушкин А.И., Гуляев Ю.В., Журавлев Ю.И., Загоруйко Н.Г., Потапов А.А., Снетков В.А., Сойфер В.А., Фурман Я.А. [9-42], а также такие отечественные и зарубежные авторы расчетных методов, моделей и алгоритмов, - в том числе, сопровождения объектов и субъектов на динамичных телевизионных изображениях, - как Алпатов Б.А., Бабаян П.В., Визильтер Ю.В., Васильев В.И., Гренандер У., Дуда Р., Кононюк А.Е., Приоров А.Л., Fisher Y., Khart P., Lucas B.D., Marr D., Kanade T., Pratt W.K., Roberts L.G., Sobel I.E. и др. [43-87]. Например, пакетный способ сопоставления динамичных телевизионных изображений и регистрируемых образов имеет свою предысторию [6, 7, 14, 15, 21, 25] и в определенном смысле основан на результатах ряда научных работ по исследованию текстур. Так при изучении текстуры изображения был применен метод, базирующийся на анализе семейства изображений, полученных при комбинации искажений анализируемого изображения с помощью

разнообразных фильтров, при этом комбинация фильтров определялась интуитивно при зрительном анализе образов (полосы, пятна и т.п.). Задачи, которые решались в этих работах, не были связаны с распознаванием лиц, но позволили сделать практические значимые выводы: во-первых, не следует усложнять фильтры, достаточно в ряде случаев применять линейные комбинации гауссовых фильтров; во-вторых, по возможности необходимо использовать такие методы фильтрации, которые генерируют зрительно понятные образы, то есть не следует отказываться от интуиции.

Учитывая высокий уровень востребованности и степени разработанности вышеуказанного направления научных исследований, в российских университетах актуализированы учебно-методические комплексы дисциплин для подготовки квалифицированных специалистов, владеющих методами и средствами цифровой обработки телевизионных изображений и распознавания образов [2, 16-19, 28, 29, 35-44, 55-70, 74-78].

Основу современного подхода к решению альтернативных задач составляет утверждение, состоящее в том, что если тестовый набор в полном объеме соответствует эталонам, то существенное число известных методик анализа и оценки изображений обеспечивает высокий уровень достоверности распознавания образов. Однако в этом случае методики должны обладать принципиальными структурно-функциональными различиями, носить универсальный характер и не быть привязаны к конкретному виду изображений, - в результате отпадает необходимость применения процедуры обучения системы распознавания изображений. В этом случае наиболее существенными обстоятельствами при организации натурального эксперимента является качество снимков при недостаточно устойчивом освещении субъектов и определенная сложность извлечения теста из непрерывно регистрируемого видео движущегося потока изображений [9-15; 36-38; 57-60; 78-86].

Объектом диссертационного исследования являются цифровые системы, устройства и технологии [1, п.5, п.7, п.10] контроля и распознавания полутонных, в том числе, цветных телевизионных изображений.

Предметом исследования являются методы, алгоритмы и модели [1, п.4, п.7] распознавания, контроля и обработки телевизионных изображений и распознавания лиц.

Цель и задачи исследования – это повышение эффективности и достоверности обработки и распознавания динамических телевизионных изображений [1] в условиях хаотичности пространственного перемещения, визуального однообразия и многочисленности одновременно контролируемых субъектов съемки.

Для достижения поставленной цели необходимо решить следующие **научные задачи**:

1) выполнить аналитические исследования современных методов и аппаратно-программных средств распознавания телевизионных изображений и лиц в условиях сложной помеховой обстановки;

2) разработать информационно-технологические модели, алгоритмы и методики распознавания лиц в условиях пониженной освещенности, при использовании одиночных архивных эталонов (фото) низкого качества и полутонных динамических телевизионных изображений в виде сплошного потока кадров видеосъемки;

3) выполнить программную реализацию алгоритмов, обеспечивающих распознавание динамических телевизионных изображений, образов и лиц;

4) подготовить опытный образец комплекса и провести лабораторные и натурные испытания разработанных аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц;

5) реализовать результаты диссертационных исследований в инновационной научной и образовательной деятельности высших учебных заведений, а также в научно-производственной деятельности предприятий реального сектора промышленности.

Отличительная **научная новизна** результатов решения поставленных задач состоит в следующем:

- данные аналитического исследования современных методов и аппаратно-программных средств распознавания телевизионных изображений и лиц в условиях сложной помеховой обстановки определяют возможность повышения вероятности распознавания в условиях изменения освещения, оптических искажений и перекрытия объектов за счет применения пакетного способа сравнения и принципа модульного объединения независимых методик и искусственных трансформаций регистрируемых изображений;

- результаты разработки информационно-технологических моделей, алгоритмов и методик распознавания лиц в условиях пониженной освещенности отличаются возможностью использования одиночных архивных эталонов (фото) низкого качества и полутоновых динамических телевизионных изображений в виде сплошного потока кадров видеосъемки за счет структурной минимизации входных данных и совместного использования нескольких методов распознавания образов различного типа;

- результаты программной реализации алгоритмов, обеспечивающих распознавание динамических телевизионных изображений, образов и лиц, отличаются универсализацией аппаратно-программных средств, а именно, алгоритмов, реализующих принципы ближайшего соседа и попарного сравнения регистрируемых изображений с эталонами;

- результаты лабораторных и натурных испытаний разработанных аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц, вошедших в состав изготовленного образца комплекса «Стенд», отличаются более высокой достоверностью и результативностью по сравнению с широко распространенным узкоспециализированным алгоритмическим комплексом "COGNITEC";

- результаты разработки комплекса «Стенд» отличаются высокой востребованностью и практической значимостью аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц для использования в инновационной научной, образовательной и производственной

деятельности учебных заведений и промышленных предприятий, что подтверждено актами практического использования и реализации.

Теоретическая и практическая значимость работы

Научно-техническая новизна и практическая значимость достигнутых результатов диссертационного исследования подтверждены публикациями в журналах, рекомендованных ВАК, и в сборниках научных трудов, докладами на научных конференциях международного и российского уровня, а также государственными Свидетельствами РФ о регистрации программ для ЭВМ.

Разработанные в диссертации алгоритмы, математические и информационно-технологические модели и аппаратно-программные средства составили основу перспективной системы распознавания телевизионных изображений и лиц, созданной в рамках выполнения государственных заданий и госбюджетных опытно-конструкторских и научно исследовательских работ «Разработка технических средств распознавания образов по анализу видеоизображения для автоматизации работы оператора» по Государственному контракту с Минобрнауки РФ от 17 ноября 2010г. №02.524.11.4010 в рамках реализации Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России». Система реализована и успешно функционирует на опытном образце аппаратно-программного комплекса в ВлГУ, г. Владимир. Данное обстоятельство предопределило целесообразность и возможность освоения промышленного производства разработанных аппаратно-программных средств на предприятиях региона.

Кроме того, результаты диссертации были положены в основу работ по государственному заданию ВлГУ (от 17.03.2020 соглашение №075-03-2020-046/1).

Результаты диссертационного исследования востребованы и внедрены в инновационную научную и образовательную деятельность ВлГУ в г. Владимир, а также одобрены к промышленной реализации в рамках НИОКР Федерального казённого предприятия «Государственный лазерный полигон «Радуга» в г. Радужный Владимирской области, предприятия ООО «Техника и технология» в г.

Владимире и предприятия ООО «РУСОКСИД» в г. Москве на территории инновационного центра «Сколково», что подтверждено актами практического использования и реализации.

Методы исследования

При решении поставленных научных задач использовались современные методы цифровой обработки изображений, компьютерного зрения, распознавания образов, математического анализа, теории вероятностей и математической статистики. Для практической реализации алгоритмов применялись современные численные методы и методы объектно-ориентированного программирования на языках Matlab, C++, C#.

Положения, выносимые на защиту

По результатам выполненных исследований при решении поставленных научных задач необходимо отметить следующие достигнутые преимущества:

- результаты аналитического исследования и разработки информационно-технологических моделей, алгоритмов и методик распознавания лиц при использовании одиночных архивных эталонов (фото) низкого качества в условиях пониженной освещенности и сложной помеховой обстановки обеспечивают повышение вероятности детектирования образов до 0,98 и, в ряде случаев, до 1,00 за счет применения пакетного способа сравнения классов и исключения ложных срабатываний;

- программная реализация алгоритмов цифровой обработки полутоновых динамических телевизионных изображений в виде сплошного потока регистрируемых кадров видеосъемки, обеспечивает повышение помехоустойчивости системы и вероятности среднего уровня распознавания лиц с 0,65 до 0,79, в том числе, при деградации описания классов;

- опытный образец комплекса аппаратно-программных средств «Стенд» при проведении лабораторных и натурных испытаний, а также при его применении в образовательной и производственной деятельности обеспечивает совместное использование несколько методов распознавания различного типа и повышение быстродействия процедуры распознавания за счет модульного принципа

построения комплекса, допускающего возможность добавления, исключения или замены подсистем верхнего уровня.

Степень достоверности и апробация результатов

Достоверность достигнутых научных результатов обусловлена применением адекватного математического аппарата, что подтверждается согласованностью с данными компьютерного моделирования и экспериментального исследования разработанных аппаратно-программных средств, а также непротиворечивостью при сопоставлении полученных научных результатов с ранее опубликованными данными известных российских и зарубежных ученых передовых научных школ.

Результаты диссертационного исследования прошли научную экспертизу в Оргкомитетах, докладывались, обсуждались и опубликованы в сборниках научных трудов, следующих международных, российских и региональных научных и научно-прикладных конференций:

— XX международная научно-практическая конференция «Теоретические и методологические проблемы современных наук», г. Новосибирск, 2017 г.;

— Международная конференция по математической теории управления и механике, г. Суздаль Владимирской области, 2015 г.;

— XVIII всероссийская научно-методическая конференции «Телематика'2011», г. Санкт-Петербург, 2011 г.;

— Региональная научно-практическая конференция «Многоядерные процессоры и параллельное программирование», г. Барнаул, 2011 г.;

— XVII Всероссийская научно-методическая конференции «Телематика'2010», г. Санкт-Петербург, 2010 г.;

— IX международная научно-практическая конференция «Исследование, разработка и применение высоких технологий в промышленности», г. Санкт-Петербург, 2010 г.

Результаты диссертационной работы прошли научную экспертизу в государственной комиссии по назначению стипендии Президента РФ для "аспирантов очной формы обучения по имеющим государственную аккредитацию

образовательным программам, соответствующим приоритетным направлениям модернизации и технологического развития экономики России на 2011-2012 годы".

Достигнутые научно-технические результаты прошли научно-общественную экспертизу, отмечены в качестве Победителя программы и явились основанием для заключения Договора от 21.03.2014г. №1769ГУ1/2014 «Разработка алгоритмов биометрической идентификации человека в условиях единичного низкокачественного эталона» с Федеральным государственным бюджетным учреждением «Фонд содействия развитию малых форм предприятий в научно-технической сфере» ("Фонд содействия инновациям") при выполнении научно-практических работ по Программе «Участник молодежного научно-инновационного конкурса «УМНИК-2014».

Научно-техническая новизна и практическая значимость результатов диссертационного исследования подтверждены государственными Свидетельствами РФ о регистрации программ для ЭВМ:

- Полуавтоматическая аппроксимация экспериментальных данных базисным набором гауссиан. - Свидетельство о государственной регистрации программ для ЭВМ № 2017611477; заявлено 05.12.2016 г.; зарегистрировано 03.02.2017 г.;

- Программное обеспечение автоматизированного тестирования алгоритмов распознавания лиц. - Свидетельство о государственной регистрации программ для ЭВМ № 2014612123; заявлено 22.10.2013 г.; зарегистрировано 19.02.2014 г.

- Модульная программа идентификации человека по растровому двумерному изображению лица. - Свидетельство о государственной регистрации программ для ЭВМ № 2011611917; заявлено 22.11.2010 г.; зарегистрировано 28.02.2011 г.

- База данных биометрических признаков. - Свидетельство о государственной регистрации программ для ЭВМ № 2011620172; заявлено 22.11.2010 г.; зарегистрировано 28.02.2011 г.

Глава 1 Современные методы и средства распознавания, контроля и обработки телевизионных изображений

1.1 Цифровые системы, устройства и технологии контроля и распознавания полутоновых телевизионных изображений

Решаемые задачи и исходные данные

В настоящее время проблема реализации процедуры контроля и обработки телевизионных изображений предопределяет необходимость эффективного и достоверного решения разноплановых задач [1], таких, например, как распознавание на экранах цифровых мониторов динамических образов и лиц в условиях сложной помеховой обстановки и недостаточной освещенности объектов и субъектов видеосъемки [17, 30, 33, 35, 43-45, 49-52, 54, 75, 79-95]. При этом требуется, в первую очередь, оценить структуру и объем информации, поступающей для выполнения процедуры распознавания конкретного лица, а также структуру и объем информации, имеющейся в системе базы данных о лице [16, 18, 20, 26, 31, 38, 46, 55, 61, 65, 68, 76].

Современные методы и средства обработки видеоизображений относятся к числу динамично развивающегося научно-технического направления, востребованного, как уже отмечалось выше, при построении систем цифрового телевидения, охранных систем, систем мониторинга состояния городской среды, системы распознавания государственных регистрационных номерных знаков, систем распознавания лиц, систем слежения за объектами и субъектами, медицинских диагностических системы, систем телеметрии, игровых систем и систем, дополняющих функции перспективных гаджетов. Кроме того, при создании и совершенствовании перспективных интеллектуальных систем мониторинга, контроля и управления как производственными (в том числе, научно-техническими и технологическими), так и социальными процессами необходимо отслеживать положение подвижных объектов и субъектов в реальном времени на

основе достоверной зрительной информации, поступающей от видеокамеры [88-91].

Основополагающей методологической частью в системах обработки видеоизображений являются процедуры распознавания образов, определяющие возможности эффективного решения задачи по определению принадлежности регистрируемого изображения к числу сохраненных эталонных образов. Располагая рядом последовательных и динамичных во времени цифровых видеоизображений, появляется возможность выявления специфической информации, обладающей свойством прогнозирования и предсказания перспектив видеотрансформации объекта (и субъекта), что способствует повышению эффективности обнаружения контролируемых, но частично внешне изменяющихся объектов при их хаотическом перемещении в пространстве [92-96].

По существу, результатом решения задачи распознавания является ответ на вопрос, обладают ли исследуемые объекты фиксированным конечным набором признаков, позволяющим отнести данные объекты к определенному классу, при этом к числу наиболее актуальных задач целесообразно отнести следующие типы процедуры распознавания [2, 4, 7, 23, 25]:

- выбор информативного набора признаков описания для распознавания объектов (образов);
- приведение исходных данных о признаках описания исследуемого объекта к виду, приемлемому для распознавания последовательности объектов;
- разделение множества объектов по их признакам описания на непересекающиеся подмножества классов в режиме автоматической классификации;
- отнесение предъявленного для распознавания объекта по признакам его описанию к одному из востребованных классов – задача обучения аппаратно-программного комплекса;
- реализация процесса динамического распознавания и динамической классификации исследуемых подвижных объектов;

- выработка и совершенствование универсальных признаков описания подвижных объектов – задача прогнозирования.

Определено [2, 4, 7], что задачи распознавания подлежат решению в два этапа: формирование исходных данных в наиболее приемлемом для распознавания виде и выполнение непосредственно процедуры распознавания, то есть установление с высокой степенью достоверности факта принадлежности объекта к конкретному искомому классу. При этом в качестве основания для зачисления объектов в один и тот же класс или в различные классы вводятся понятия аналогии (или подобия объектов) и близости объектов. Кроме того, при решении подобных задач рекомендуется оперировать набором примеров-прецедентов, классификация которых заранее известна и которые в виде формализованных описаний могут быть предъявлены алгоритму распознавания для настройки аппаратно-программного комплекса в процессе обучения.

Особенность решения вышеуказанных задач в условиях обработки исходной видеоинформации, изобилующей структурными пробелами и пропусками отдельных фрагментов регистрируемой видеопоследовательности, а также характеризующейся признаками разнородности видеофайлов, их косвенности, нечеткости и неоднозначности, состоит в том, что возможность формализации теоретической процедуры распознавания образов для применения классических математических методов зачастую весьма затруднительна из-за недоступности полного объема информации, необходимой для построения достоверной математической модели объекта. В ряде случаев возможность применения классических математических методов нецелесообразна из-за финансовой неприемлемости данного подхода, то есть в силу значительной несоизмеримости достижимых результатов и преимуществ с фактически запредельными ресурсными затратами [7].

Системы обработки телевизионных изображений широко используются в технологических операциях видеоконтроля изделий микроэлектронного производства при изготовлении интегральных коммуникационных оснований и многослойных печатных плат [80, 96-98]. В данном случае системы обработки

видеоизображений решают задачи ориентации и совмещения, например, рабочих и эталонных фотошаблонов. Позиционирование выполняется путем обнаружения и распознавания реперных знаков на телевизионном изображении многослойной печатной платы (МПП). Реперные знаки представляют собою объекты известной формы, наносимые на поверхность МПП. Для обеспечения точности позиционирования применяют одиночные или многочисленные взаимосвязанные метки прямоугольной формы либо характерные по форме элементы топологического рисунка.

Здесь необходимо отметить [99, 100], что подобные системы должны достоверно функционировать при минимальном участии оператора. При этом достигнутые результаты современной качественной видеосъемки наряду с высоким уровнем развития аппаратно-программных средств информационно-вычислительной техники определяют актуальность построения систем обработки изображений, обеспечивающих автоматизацию процессов сбора и классификации регистрируемых зрительных данных (образов) при снижении числа ошибок субъективного характера [8-55, 71-78].

В общем виде структурная схема системы [7], реализующей процесс распознавания образов и являющаяся по сути объектом диссертационного исследования, представлена на рисунке 1.1.

При решении задачи распознавания образов особое значение имеет состояние источников информации, отображающих различные варианты исходных данных о распознаваемом объекте. К их числу следует отнести: многочисленные или ограниченное число качественных или искаженных фотографий, выполненных из стационарных или сменяемых точек наблюдения (точек фотосъемки объекта с различным ракурсом); подвижные или статичные телевизионные изображения из строго фиксированных или сменяемых точек видеонаблюдения; фрагменты видеопоследовательности кадров съемки объекта в условиях сложной помеховой обстановки, включая атмосферные искажения; наличие субъективных отрывочных словесных описаний и/или полноценных либо фрагментарных графических моделей поверхности лица распознаваемого субъекта; результаты

математического двух- или трехмерного компьютерного моделирования лица по доступным антропологическим данным [101].



Рисунок 1.1 - Структурная схема системы распознавания образов.

Особый интерес при выполнении исследований представляет возможность решения задачи распознавания образов при наличии следующих ограничений по исходным данным:

- для распознавания используется только одно изображение (фотография) объекта или субъекта;
- база образов (БО), имеющихся в "памяти" разрабатываемой системы, содержит только по одному изображению для каждого распознаваемого класса, при этом изображения относятся к категории эталонных.

Обе поставленные задачи являются одними из числа наиболее проблематичных в части обеспечения достоверности результатов распознавания образов, так как в основу исходных данных положен малый объем информации, обеспечивающий идентификацию объекта. В зарубежных источниках информации [81-87, 103-105] данное направление идентифицируется как «One Sample Problem»

(OSP), то есть "проблема одного эталона" (ПОЭ). В опубликованных отечественных научных трудах [91-95, 106] данный тип задач определяется как "распознавание при малом числе обучающих прецедентов". Здесь следует подчеркнуть, что подобная постановка задачи распознавания образов весьма актуальна, так как зачастую фотография потенциального нарушителя и/или террориста, хранящаяся в государственных архивах регистрационно-паспортных подразделений МВД РФ, является единственным имеющимся источником информации о его внешности.

В данном направлении ведутся фундаментальные и научно-прикладные изыскания, в том числе, и во Владимирском государственном университете имени А.Г и Н.Г. Столетовых (ВлГУ) [107-115, 128-133, A116-A127]. Из результатов выполненных экспериментальных работ следует, что качество и достоверность алгоритмов распознавания субъектов снижается при уменьшении числа изображений лица человека, имеющихся в БО. Так, например, при наличии в БО лишь одного единственного ракурса изображения лица человека при использовании наиболее распространенных известных алгоритмов [107, 109, 111, 113, 115, A116, A118, A121, A122, A124, A125, A127] вероятность распознавания субъекта опускается ниже 0,65, то есть менее 65%.

Востребованная результативность решения подобных задач предопределила актуальность поиска различных вариантов оптимизации алгоритмов, связанных с "проблемой одного изображения". Данные сопоставительного анализа известных алгоритмов приведены в таблице 1.1 [131], причем анализируемые алгоритмы тестировались на различных БО, таких, как "FERET" [134], "ORL" [135], "AR" [136].

Наиболее высокой результативностью характеризуется метод эластичных графов ("Elastic Bunch Graph Matching"). Однако известно [131], что данный метод проблематично применять на практике, так как для его реализации, во-первых, требуются значительные вычислительные ресурсы, и, во-вторых, при распознавании лица расчет ведется по ключевым точкам, которые являются крайне неустойчивыми в пространстве даже при незначительных изменениях в

изображении лица (мимика, очки, ниспадающие на лицо волосы, надвинутые на лицо головные уборы и т.п.).

Таблица 1.1 – Сопоставительный анализ известных методов и алгоритмов распознавания лиц.

Наименование методов и алгоритмов распознавания лиц	Наименование БО	Количество классов в БО	Количество изображений лиц в БО	Вероятность распознавания лиц
Эластичных графов	FERET	1196	1196	0,95
Локальных бинарных образов	FERET	1196	5980	0,97
Шаблонного распознавания	FERET	200	200	0,835
Шаблонного распознавания (улучшенный)	FERET	200	200	0,855
Линейного дискриминантного анализа	FERET	70	350	0,786
Дискриминантный метод главных компонент	FERET	256	914	0,72
Аналитически-глобальный подход	ORL	40	160	0,84
Одномерные скрытые марковские модели	AR	120	1440	0,898

Также отличные результаты получены при использовании метода локальных бинарных образов ("Local Binary Pattern"). В данном методе изображение объекта анализируется на трех уровнях сегментации [131], а именно: уровень точечной сегментации, уровень частичной сегментации и уровень целостного изображения. Частичная сегментация производится путем деления изображения на небольшие прямоугольные блоки. Для каждого уровня сегментации строятся гистограммы, с помощью которых происходит распознавание образа. Метод характеризуется высокими результатами распознавания, однако "при обучении средств распознавания" используются пять фотографий для каждого лица.

К числу специальных исходных данных при решении задачи распознавания образов необходимо отнести энергетические характеристики регистрируемого видеосигнала. Так, например, по результатам ранее выполненных исследований

известно [88-95], что эффективность работы средств оптимального обнаружения объекта, находящегося в пределах поля видео обзора, подлежащего регистрации, определяется энергией видеосигнала, исходящего от объекта, и не зависит от внутренних структурных особенностей в построении самого сигнала. Поэтому при решении задачи распознавания субъекта по регистрируемому изображению необходимо учитывать, что достоверность распознавания в значительной степени предопределяется физическими (а именно, энергетическими) параметрами и характеристиками данного сигнала.

1.2 Методы распознавания, модели и особенности построения алгоритмов контроля и обработки телевизионных изображений

Известные методы распознавания образов подразделяются на общие и специализированные. Общие методы характеризуются универсальностью применения и могут использоваться не только для распознавания лиц, но и при решении других задач. Специализированные методы оперируют с признаками, характерными для описания лиц людей, и разработаны для решения данной конкретной задачи и, как правило, мало применимы для распознавания других образов.

Общие методы

Метод главных компонент основан на одном из самых распространенных алгоритмов распознавания образов [36, 97, 137]. Здесь входные векторы представляют собой отцентрированные и приведённые к единому масштабу изображения лиц. Векторы, вычисленные для всего набора изображений лиц, определяются как собственные лица и имеют полезное свойство, заключающееся в том, что изображение, соответствующее каждому такому вектору, имеет лицеподобную форму (см. рисунок 1.2).

С помощью вычисляемых матриц входное изображение разлагается на набор линейных коэффициентов, являющихся главными компонентами. Сумма главных

компонент, умноженная на соответствующие собственные векторы, является реконструкцией изображения.

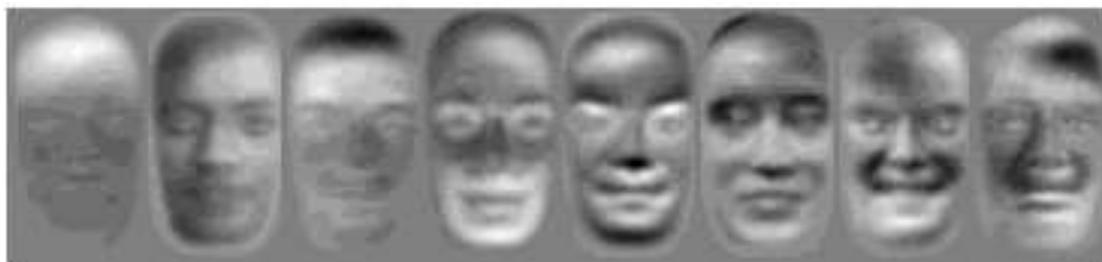


Рисунок 1.2 - Лицеподобные формы собственных векторов.

Для каждого изображения лица вычисляются его главные компоненты. В большинстве случаев используют от 5 до 200 главных компонент. Остальные компоненты кодируют мелкие различия между лицами, а также шум. Процесс распознавания заключается в сравнении с помощью метрик главных компонент неизвестного изображения с главными компонентами всех остальных изображений.

В **линейном дискриминантном анализе** [108] выбирают проекцию пространства изображений на пространство признаков таким образом, чтобы минимизировать внутриклассовое и максимизировать межклассовое расстояние в пространстве признаков. В этих методах предполагается, что классы линейно разделимы. В данном случае проецирование на собственное пространство смешивает классы, что предопределяет распознавание как невозможное, а линейный дискриминант определяет проекцию на пространство признаков таким образом, чтобы разделить разные классы. Также, как и собственные векторы, изображения базисных дискриминантных векторов имеют лицеподобную форму.

Алгоритмы **метода оптического потока** применяют в основном для анализа движения. Используя два или более последовательных кадра изображения, выполняется расчет двумерного векторного поля, являющегося оптическим

поток, которое отражает актуальное или наиболее вероятное смещение точек изображения от кадра к кадру [138].

В этом методе оптический поток рассчитывается для двух произвольных изображений лица с целью получения меры соответствия изображений. Эти два изображения считаются последовательными кадрами. Затем вычисляется векторное поле, наилучшим образом отображающее одно изображение в другом по условию минимизации расстояния между изображениями и с учётом геометрических ограничений, таких как относительное взаимное расположение точек изображения. Алгоритм определяет блоки (участки) изображений, наиболее взаимно подобные. Поиск осуществляется иерархически, начиная с больших блоков, и затем разбивая их на меньшие блоки. Таким образом, строится пирамида соответствия изображений. Используя векторное поле оптического потока, определяются различные меры соответствия, с помощью которых в БО находится изображение, ближайшее к неизвестному.

По своей сути **метод Марковских моделей** непосредственно учитывает пространственно-временные характеристики сигналов, поэтому данный метод получил наиболее широкое применение в задачах распознавания изображений.

Каждая модель $\lambda = (A, B, \pi)$ представляет собой набор N состояний $S = \{S_1, S_2, \dots, S_N\}$, между которыми возможны переходы [139, с.31; 140, с.745; 141, с.111; 142]. В каждый момент времени система находится в строго определённом состоянии. В наиболее распространённых Марковских моделях первого порядка полагается, что следующее состояние зависит только от текущего состояния. При распознавании образов скрытые Марковские модели (СММ) применяются следующим образом.

Каждому классу i соответствует своя модель λ_i . Распознаваемый образ представляется в виде последовательности наблюдений O . Далее для каждой модели λ_i вычисляется вероятность того, что эта последовательность могла быть сгенерирована именно данной моделью. Модель λ_j , получившая наибольшую вероятность, считается наиболее подходящей, и образ относят к классу j .

Сущность двумерных Марковских моделей заключается в том, что в отличие от одномерных линейных СММ, эти модели отражают искажения изображений и взаимного расположения участков лица не раздельно по горизонтали или вертикали, но и в обоих направлениях одновременно. В данном случае для снижения вычислительной сложности процесса распознавания применяют псевдодвумерные СММ. Такие модели состоят из нескольких линейных вертикальных моделей нижнего уровня и одной линейной горизонтальной модели верхнего уровня, на вход которой поступают выходы моделей нижнего уровня. Каждое состояние модели верхнего уровня включает в себя последовательность состояний соответствующей модели нижнего уровня. Модели нижнего уровня не связаны между собой.

Таким образом, псевдодвумерная модель позволяет учесть локальные деформации и взаимное расположение участков изображений. Но в отличие от метода оптических потоков и других методов сопоставления изображений, псевдодвумерная модель учитывает, как форму деформации, так и возможные её пределы изменения, то есть участок лица, соответствующий глазу, не будет сопоставлен, с участком по месту расположения рта.

Нейросетевые методы распознавания человека по изображению лица

Искусственные нейронные сети частично моделируют работу мозга человека, поэтому с их помощью решаются разнообразные задачи, которые способен решать человек. При помощи нейронных сетей осуществляются попытки производить распознавание образов так, как это осуществляет мозг человека. Нейросетевые технологии отличаются от многих других технологий распознавания тем, что они не анализируют содержимое входного изображения как такового, а строят некоторую реакцию на него, причем эти реакции могут быть проанализированы в контексте решения задачи.

Нейросетевые методы, основанные на применении различных типов искусственных нейронных сетей, получили широкое распространение [143-145]. Ниже проанализированы основные преимущества, которыми обладают нейронные сети.

Настройка нейронной сети для решения определённой задачи производится в процессе обучения на наборе тренировочных примеров. Таким образом, не требуется вручную определять параметры модели – нейросеть извлекает параметры модели автоматически наилучшим образом в процессе обучения. Остаётся только создать тренировочную выборку. В задачах классификации при этом происходит неявное выделение ключевых признаков внутри сети, определение их значимости и системы взаимоотношений между признаками (см. рисунок 1.3).

Нейронные сети обладают высокой обобщающей способностью, а именно, "опыт", полученный в процессе обучения на конечном наборе образов, нейронная сеть "может успешно переносить" на всё множество образов. Кроме интерполяционных обобщающих способностей, нейронные сети обеспечивают возможность экстраполяции результатов анализа, так как "допускают применение" достигнутого опыта к качественно иным образам, отличным от входящих в обобщающую выборку.

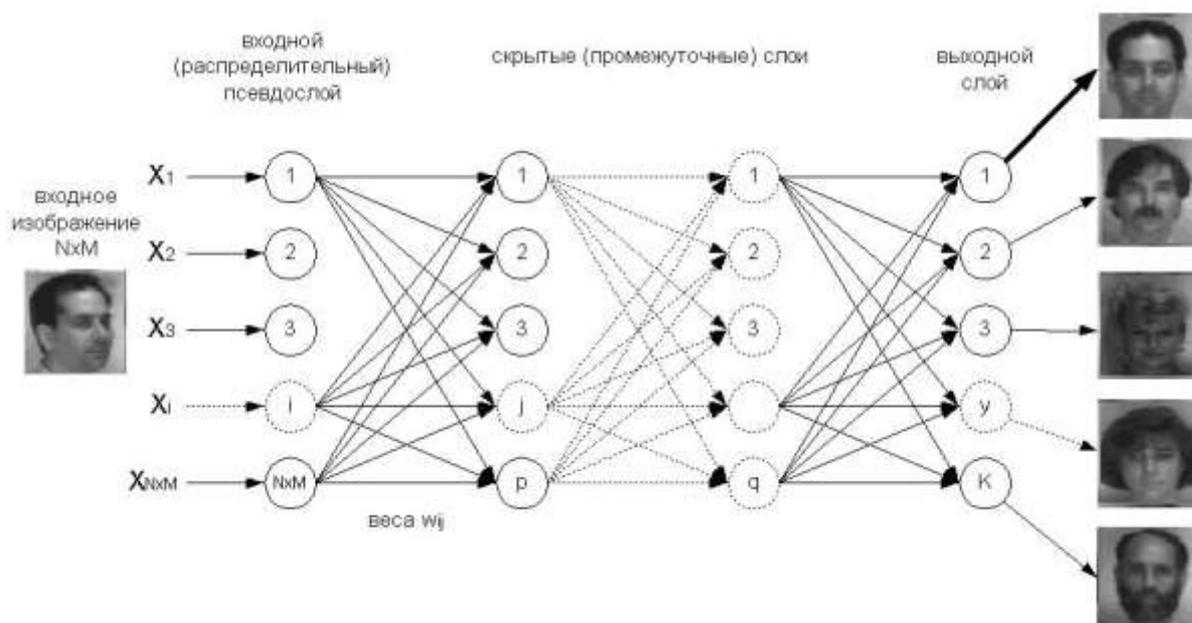


Рисунок 1.3 - Архитектура многослойной нейронной сети и её применение для распознавания изображений. Нейрон с максимальной активностью указывает принадлежность к распознанному классу.

Нейронная сеть характеризуется нечётким и распределённым хранением информации. То есть нет отдельного нейрона, отвечающего за конкретное понятие или признак, и удаление или искажение работы данного нейрона не приведёт к фатальным последствиям.

Основные разновидности нейронных сетей:

- многослойные нейронные сети [146], персептрон Розенблата [147]. Нейронная сеть с двумя решающими слоями позволяет формировать любые выпуклые области в пространстве решений, а с тремя решающими слоями – области любой сложности, в том числе и невыпуклые [148];

- нейронные сети высокого порядка [143]. Такие сети также могут формировать сложные разделяющие поверхности;

- радиально-базисные нейронные сети [148] состоят из двух слоёв. Скрытый слой представляет собой набор кластеров в пространстве образов и выполняет первый этап кластеризации входного образа;

- когнитрон [149] был разработан, основываясь на анатомии и физиологии мозга, и своей архитектурой похож на строение зрительной коры. Каждый слой мозга реализует различные уровни обобщения. Входной слой чувствителен к простым образам таким, как линии и их ориентация в определенных областях визуальной области. В то время как реакция других слоев является более сложной, абстрактной и независимой от позиции образа. Аналогичные функции реализованы в когнитроне путем моделирования организации зрительной коры;

- неокогнитрон является дальнейшим развитием идеи когнитрона и более точно отражает строение зрительной системы, позволяет распознавать образы независимо от их преобразований таких, как смещение, вращение, изменение масштаба и другие искажения [150].

Специализированные методы

Один из первых методов, известный еще со времен Бертильона, – это **метод анализа геометрических характеристик лица** [58, 148, 151, 152]. Суть метода состоит в выделении набора ключевых точек лица с последующим выделением

набора признаков. Каждый признак является либо расстоянием между ключевыми точками, либо соотношением таких расстояний. В отличие от метода сравнения эластичных графов, здесь расстояния выбираются не как дуги графов. Наборы наиболее информативных признаков выделяются экспериментально [153].

Ключевыми точками могут быть уголки глаз, губ, кончик носа и центр глаз (см. рисунок 1.4). В качестве ключевых областей могут быть прямоугольные области, включающие в себя глаза, нос и рот. В процессе распознавания сравниваются признаки неизвестного лица, с признаками, хранящимися в БО.

Правильное нахождение ключевых точек на изображении во многом определяет успех распознавания. Поэтому изображение лица человека должно быть без помех, мешающих процессу поиска ключевых точек. К таким помехам относят очки, бороды, украшения, элементы причёски и макияжа. Кроме того, изображение лица должно иметь фронтальный ракурс, возможно с небольшими отклонениями. Выражение лица должно быть нейтральным. Это связано с тем, что в большинстве методов нет модели учёта таких изменений.

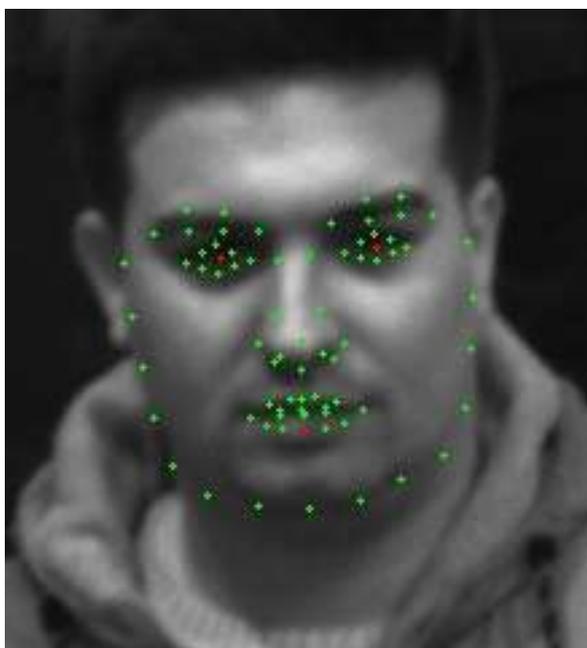


Рисунок 1.4 - Идентификационные точки и расстояния, наиболее часто применяемые при построении автоматизированных систем идентификации.

Метод сравнения эталонов [112] состоит в выделении областей лица на изображении (см. рисунок 1.5) и последующем сравнении этих областей для двух различных изображений. Каждая совпавшая область увеличивает меру сходства изображений. Данный метод является одним из исторически первых методов распознавания человека по изображению лица. Для сравнения областей используются простейшие алгоритмы, вроде попиксельного сравнения.

Недостаток данного метода заключается в том, что для его использования необходимы значительные информационно-технические ресурсы как для хранения данных о геометрии участков лица, так и для их сравнения. В виду того, что используется простейший алгоритм сравнения, изображения должны быть сняты в строго установленных условиях - недопустимы заметные изменения ракурса, освещения, эмоционального выражения лица и пр.



Рисунок 1.5 - Сравнимые области-эталонны лица.

Метод сравнения трехмерных моделей основан на сканировании лица человека и построении трехмерной модели его поверхности (см. рисунок 1.6) [36, 61, 101, 111, 154-156]. В данном методе 3D-распознавания физические и, в том числе, геометрические параметры лица известны и могут использоваться для распознавания в явном виде.

Среди различных вариантов применения метода 3D-распознавания необходимо отметить три основных направления их совершенствования, а именно, развитие средств анализа формы 3D-поверхности лица, применение законов статистического анализа и использование параметрической модели лица.

Методы, основанные на анализе формы, оперируют с геометрическими параметрами поверхности, которые численно характеризуют лицо. При этом методы могут быть классифицированы по группам, использующим либо локальные, либо глобальные свойства поверхности (например, кривизну), либо основанные на профиле линий, либо использующие метрику расстояний между двумя поверхностями.

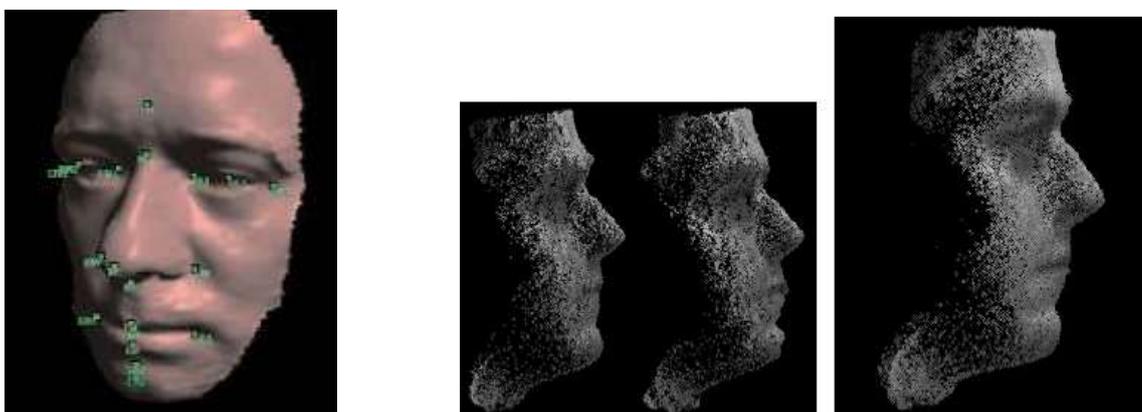


Рисунок 1.6 - Трехмерные модели поверхности лица.

Известны также **гибридные методы**, основанные на объединении информации о локальных параметрах поверхности лица с глобальной трехмерной сеткой, охватывающей все лицо [157].

Глобальными являются также методы, которые в качестве исходных данных используют все 3D-параметры лица как информацию на входе системы распознавания. Существенным ограничением по применению подобных методов является так называемое условие жесткости поверхности лица, то есть искомое лицо не может деформироваться.

Для сравнения трехмерных моделей применяются статистические методы, например, **метод главных компонент** и альтернативный ему **метод линейного дискриминантного анализа**.

Известны также методы распознавания, основанные на параметрических моделях лица. Ключевая идея распознавания по моделям основана на так называемых параметрических 3D-моделях. В данном подходе форма лица контролируется набором параметров модели. Эти коэффициенты описывают 3D-форму лица и также могут задавать цвет (текстуру) на ее поверхности. Подобная модель затем проецируется на двумерные изображения, из которых определяются параметры коэффициентов модели для данного изображения [101]. Недостатком метода является высокая вычислительная сложность и чувствительность к начальной инициализации параметров модели.

Общим недостатком, характерным для 3D-распознавания, является ресурсоемкость подобных методов, предопределяющая необходимость выполнения значительного объема вычислений и применения видео устройств с высоким разрешением.

Метод сравнения эластичных графов

В данном методе лицо представляется в виде графа, вершины которого расположены на ключевых точках лица, таких как контуры головы, губ, носа и их крайних точках [158] (см. рисунок 1.7). Каждая грань помечена расстояниями между её вершинами. В каждой такой точке вычисляются коэффициенты Габоровых функций для пяти различных частот и восьми ориентаций. Набор вычисленных коэффициентов $\{J_j\}$ является джетом. Джеты характеризуют локальные области изображений и служат для двух целей. Во-первых, для нахождения точек соответствия в заданной области на двух различных изображениях. Во-вторых – для сравнения двух соответствующих областей различных изображений. Каждый коэффициент $J_j = a_j \exp(i\phi_j)$ для точек из одной области различных изображений, характеризуется амплитудой a_j , которая медленно меняется с изменением положения точки, и фазой ϕ_j . Поэтому в

простейшем случае для поиска на новом изображении точки с аналогичными характеристиками в функции подобия фазу не учитывают:

$$S_a(J, J') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'_j{}^2}}.$$

Функция подобия с одним джетом в фиксированной позиции и с другим джетом с переменной позицией является достаточно гладкой и обеспечивает возможность получения быстрой и надёжной сходимости при поиске с применением простейших методов, таких как диффузия или градиентный спуск.

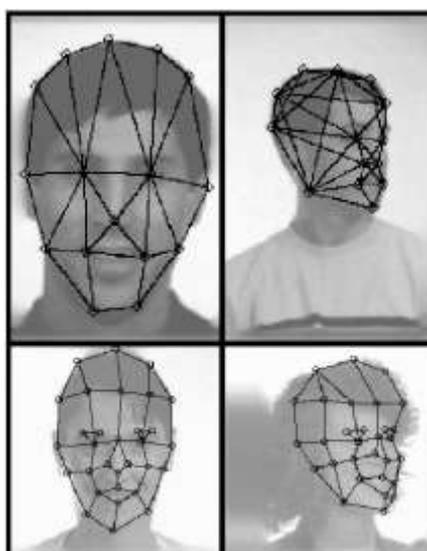


Рисунок 1.7 - Эластичный граф, покрывающий изображение лица.

Для различных ракурсов соответствующие ключевые точки отмечаются вручную на тренировочном наборе. Кроме того, чтобы для одного и того же лица представить различные вариации его изображения в одном и том же графе, для каждой точки используются несколько джетов, каждый из которых может соответствовать различным локальным характеристикам данной точки, например открытому и закрытому глазу.

Процесс распознавания неизвестного лица состоит в сравнении графа изображения лица G^I со всеми остальными графами из набора B при помощи функции подобия:

$$S_B(G^I, B) = \frac{1}{N} \sum_n \max_m (S_\phi(J_n^I, J_n^{Bm})) - \frac{\lambda}{E} \sum_e \frac{(\Delta \vec{x}_e^I - \Delta \vec{x}_e^B)^2}{(\Delta \vec{x}_e^B)^2},$$

где N – количество вершин, E – количество граней, λ – коэффициент относительной важности топографической информации.

При оценке рассмотренных методов распознавания образов с точки зрения решения проблемы одного эталона необходимо отметить следующее.

Метод главных компонент допускает возможность использования его для решения задачи распознавания субъекта по одному изображению, но обладает следующими недостатками:

- умеренные показатели корректности распознавания [137] в рамках ПОЭ;
- необходимость пересчета собственных векторов при добавлении образов в эталонную БО;
- необходимость точного центрирования и масштабирования распознаваемых изображений, что усложняет алгоритм.

Линейный дискриминантный анализ является развитием метода главных компонент и отличается более высокими показателями корректности распознавания (см. таблицу 1.1). Однако имеет те же недостатки, характерные для метода главных компонент.

Распознавание образов при помощи оптического потока вполне применимо к решению задачи ПОЭ. Однако значительным недостатком данного метода является его алгоритмическая сложность. Возникает необходимость попарного сравнения распознаваемого образа с каждым образом эталонной базы. Также для каждой пары строится векторное поле оптического потока, и только затем получают меры схожести образов. Таким образом, весьма затруднительно использовать оптический поток в системах распознавания реального времени с большим количеством образов-эталонов.

Скрытые Марковские модели и искусственные нейронные сети вполне успешно могут применяться для распознавания образов. Однако все они требуют обучения. В СММ для каждого распознаваемого класса строится модель

наблюдений. Увеличение для каждого класса массива изображений, предъявляемых для обучения, улучшает работу системы распознавания. Для достижения корректности распознавания порядка 89%, система на базе СММ требует для своего обучения до 12 фотографий каждого класса (см. таблицу 1.1). Тем самым затрудняется использование СММ для решения ПОЭ.

Искусственные нейронные сети наряду с необходимостью обучения обладают еще одним недостатком. Проектирование искусственных нейронных сетей ведется, исходя из количества распознаваемых классов. От этого зависит количество нейронов на последнем слое. Таким образом, весьма затруднено масштабирование сети при добавлении в эталонную базу дополнительных образов. Искусственные нейронные сети могут эффективно использоваться для предварительной классификации образов.

Методы, основанные на анализе геометрических характеристиках регистрируемого образа, являются специализированными. Они предназначены для распознавания исключительно конкретных объектов – лиц. Недостатком алгоритмов являются низкие показатели распознавания из-за необходимость нахождения на лице ключевых точек, что представляет собою достаточно сложную для решения задачу из-за существенных погрешностей вычисления. Особенно значительные погрешности возникают при наличии одного эталона, а также в случае некоторого отклонения распознаваемого лица от фронтального ракурса.

Метод сравнения эталонов является весьма чувствительным к ракурсу и освещенности. Метод сравнения эластичных графов характеризуется более устойчивым алгоритмом. Однако для построения графа требуется несколько изображений распознаваемого объекта. Следовательно, данный метод неприменим при решении ПОЭ. Кроме того, данные алгоритмы являются исключительно специализированными.

Метод сравнения 3D-моделей лица также является специализированным методом. Однако его применение для распознавания других объектов теоретически возможно. Недостатком метода является фактическая его неприменимость для решения ПОЭ, так как построение трехмерной модели требует или большого

количества фотографий в различных ракурсах, или быстродействующего трехмерного сканера-дальномера. Также метод весьма требователен к вычислительным ресурсам, так как обрабатывает большое количество данных.

Отдельно следует изучить методы восполнения недостатка информации, обусловленного наличием в базе только одного эталона для каждого класса. Эта информация восполняется различными, иногда искусственными методами.

В работе [159] предлагается модифицировать метод линейного дискриминанта Фишера. Этот метод неприменим в чистом виде для решения ПОЭ, потому что ему требуется более одного обучающего изображения для каждого человека. Авторы предлагают создать некоторый обобщенный обучающий набор, состоящий из нескольких изображений для каждого человека. Обучить на нем систему распознавания. А затем, используя прогнозирование, применить результаты обучения для распознавания людей из базы уже с единственным изображением для одного человека.

В исследованиях, проводимых ранее [160] предлагается использовать механизмы регрессии (то есть вероятностные механизмы) в стандартных алгоритмах распознавания, а именно в методе главных компонент и методе наименьших частных квадратов. Кроме того, предлагается [109] устанавливать дополнительные ракурсы лица, представленного одной обучающей фотографией, и затем использовать имеющиеся алгоритмы для распознавания субъекта.

В общем виде [161] структурная схема средств программно-алгоритмического обеспечения оптико-электронной системы обнаружения и распознавания образов представлена на рисунке 1.8.

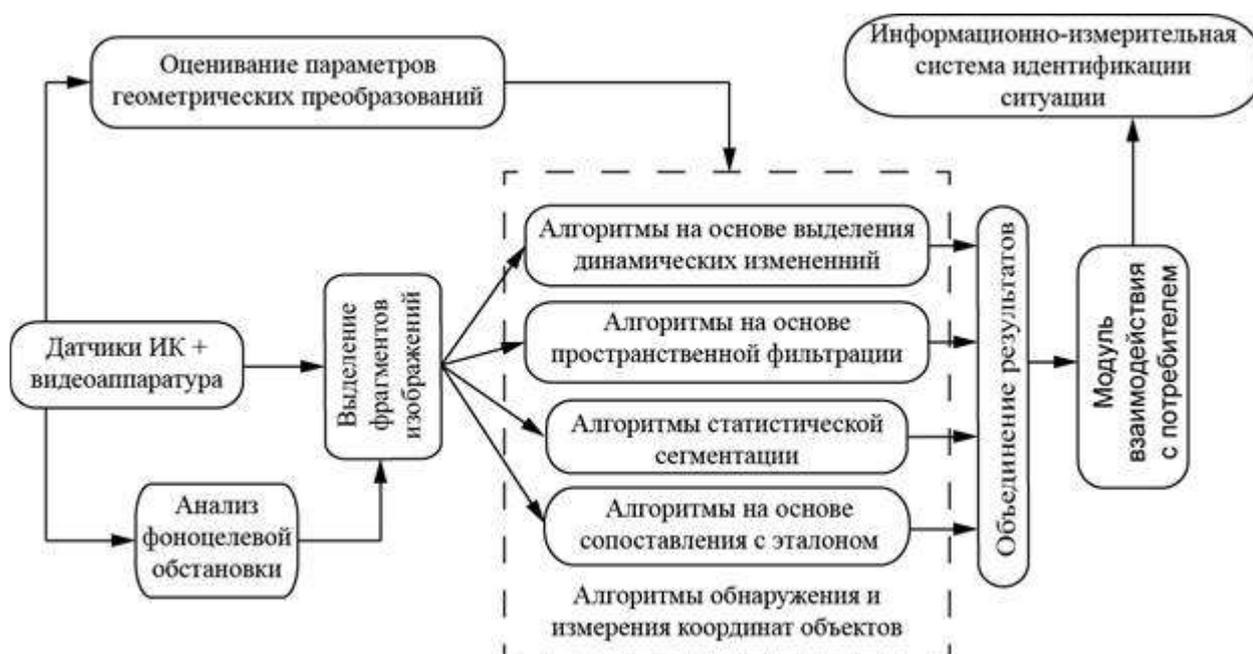


Рисунок 1.8 - Структура программно-алгоритмического обеспечения оптико-электронной системы обнаружения и распознавания объектов.

1.3 Постановка задачи

Современные методы и средства распознавания, контроля и обработки телевизионных изображений относятся к числу наиболее востребованных, перспективных и изучаемых направлений развития техники связи, передачи и приема информации, систем и устройств телевидения для различных отраслей народного хозяйства. В связи с распространением угрозы террористической опасности и продвижением в преступной среде противоправных действий с использованием биометрических данных, возрастает актуальность исследования и разработки методов и средств оперативной и достоверной цифровой обработки динамических изображений и распознавания образов в процессе телевизионного мониторинга городской среды в условиях сложной помеховой обстановки при отсутствии надежных источников первичной информации о внешнем облике субъектов противоправных действий.

В контексте данной проблемы необходимо выполнить теоретические и прикладные исследования, основанные на использовании биометрических технологий и направленные на автоматизацию процессов обнаружения и опознания опасных субъектов в местах массового скопления и хаотичного перемещения людей, и, в первую очередь, в таких местах, как вокзалы, аэропорты, метрополитен и торговые центры. Для этого необходимо решить следующие научно-прикладные задачи:

- выполнить аналитические исследования современных методов и аппаратно-программных средств распознавания телевизионных изображений и лиц в условиях сложной помеховой обстановки;

- разработать информационно-технологические модели, алгоритмы и методики распознавания лиц в условиях пониженной освещенности, при использовании одиночных архивных эталонов (фото) низкого качества и полутоновых динамических телевизионных изображений в виде сплошного потока кадров видеосъемки;

- выполнить программную реализацию алгоритмов, обеспечивающих распознавание динамических телевизионных изображений, образов и лиц;

- подготовить опытный образец комплекса и провести лабораторные и натурные испытания разработанных аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц;

- реализовать результаты диссертационных исследований в инновационной научной и образовательной деятельности высших учебных заведений, а также в научно-производственной деятельности предприятий реального сектора промышленности.

При этом основополагающими сферами научных исследований являются методы и средства информационно-технологического обеспечения процессов идентификации, контроля и обработки телевизионных изображений, а также средства технического и методического обеспечения процессов распознавания, контроля, диагностики, испытания и управления видеосистемами.

1.4 Выводы по 1-ой главе

1) Результаты аналитических исследований состояния современных цифровых систем, устройств и технологий контроля и распознавания полутоновых телевизионных изображений, а также методов, алгоритмов и информационно-технологических моделей обработки изображений и распознавания лиц подтвердили актуальность повышения эффективности и достоверности обработки и распознавания динамических телевизионных изображений в условиях хаотичности пространственного перемещения, визуального однообразия и многочисленности одновременно контролируемых субъектов съемки.

2) Сформулированы основные научные задачи, подлежащие решению и определяющие необходимость выполнения теоретических и прикладных исследований в части:

- разработки информационно-технологических моделей, алгоритмов и методик распознавания лиц в условиях пониженной освещенности, при использовании одиночных архивных эталонов (фото) низкого качества и полутоновых динамических телевизионных изображений в виде сплошного потока кадров видеосъемки;

- программной реализации алгоритмов, обеспечивающих распознавание динамических телевизионных изображений, образов и лиц;

- подготовки опытного образца исследовательского комплекса для лабораторных и натурных испытаний аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц;

- реализации данных научно-практических исследований в научной и образовательной деятельности высших учебных заведений, а также в производственной деятельности предприятий реального сектора промышленности.

Глава 2 Исследование и разработка методов и средств информационно-технологического обеспечения процессов идентификации, контроля и обработки телевизионных изображений

2.1 Обоснование технологии выбора процесса распознавания и обработки телевизионных изображений

Этап нормализации входного изображения. Алгоритм анализа изображения с возможностью маршрутизации

Входными данными для алгоритма анализа и маршрутизации являются:

- а) не менее двух видеокадров, поступающих из передающей части системы видеорегистрации;
- б) информация о положении объектов относительно кадра (силуэтная маска, ограничивающий прямоугольник).

Алгоритм анализа изображения с возможностью маршрутизации основан на процедуре слежения за точечными особенностями изображения (ТОИ). Под точечной особенностью понимается такая точка изображения, окрестность которой отличается от окрестности любой другой точки в некоторой другой окрестности этой особой точки.

Процедура слежения за ТОИ, используемая в системе распознавания, обеспечивает возможность поиска ТОИ на кадре из видеопотока и сопоставление точечных особенностей двух последующих кадров с целью определения смещения точечных особенностей сцены за промежутки времени между последовательными кадрами. Данная процедура слежения за точечными особенностями подлежит описанию следующим образом.

В качестве процедуры слежения за точечными особенностями изображения в системе применена "Open-source" реализация алгоритма KLT (Kanade-Lucas-Tomasi) - слежения является адаптивной к изменению освещённости. Данный алгоритм слежения за точечными особенностями сцены назван в честь

разработчиков: Т. Kanade, В. D. Lucas, С. Tomasi, - метод Лукаса-Канаде-Томаси [162-167]. Реализация использует вычисления на GPU, что позволяет разгрузить центральный процессор для других задач и существенно повысить производительность алгоритма. Описание данной реализации процедуры слежения за точечными особенностями изображения приведено в [110], при этом возникает необходимость анализа особенностей данной реализации.

Адаптация к изменению параметров экспозиции

Подстраиваясь к изменению освещённости сцены, современная цифровая камера меняет параметры экспозиции, что затрудняет сопоставление изображений последовательных кадров видеопотока, поступающего с выхода камеры. Основы метода адаптации алгоритма к изменению параметров экспозиции, применённого в данной реализации, приведены в авторской научной статье [A101]. С целью компенсации отличий в параметрах экспозиции и для оценки соотношения параметров экспозиции для различных кадров применяется линейная модель отклика камеры.

В дополнении к версии алгоритма KLT-слежения, адаптирующейся к параметрам экспозиции, описываемая реализация содержит, в том числе, классический вариант KLT-слежения, предполагающий постоянные параметры экспозиции и требующей для каждой отслеживаемой точечной особенности на каждой итерации решения системы размерности 2×2 . Таким образом, если камера функционирует с постоянными параметрами экспозиции или, если освещённость в зоне съёмки сравнительно постоянна, то возникает возможность использования данной более простой версии алгоритма без снижения эффективности вычислений, что обеспечивает сохранность вычислительных ресурсов.

Выбор технологии GPGPU

На данном этапе развития технологий неграфических вычислений с помощью графических процессоров принято использовать разработанную фирмой "NVidia" технологию "CUDA" ("Computer Unified Device Architecture"), то есть "унифицированная архитектура вычислительных устройств". Но по результатам выполненных измерений параметров и оценки характеристик производительности

[A116-A127] следует вывод, что реализации алгоритмов, основанных на шейдерах во многих случаях, обеспечивает результате с более высокой производительностью. В частности, оценка производительности версий алгоритма, реализованных с использованием "Cg/OpenGL" показала, что вариант, использующий шейдеры, является более быстродействующим, чем тот, который использует технологию "CUDA". Здесь необходимо отметить, что обычно KLT-сопровождение точечных особенностей является высокопараллельным алгоритмом и не требует взаимодействия между потоками, обрабатывающими индивидуальные треки. Главное преимущество технологии "CUDA" состоит в том, что перенос данных между потоками с использованием быстрой общей памяти не используется. Кроме того, применение "Cg/OpenGL" позволяет запускать данную реализацию KLT-слежения на менее дорогостоящем IT-оборудовании, что значительно повышает её ценность.

2.2 Исследование и разработка информационно-технологической модели и алгоритмов распознавания и обработки телевизионных изображений

Этап обнаружения объекта

Разработанный алгоритм обнаружения и детектирования лиц [A119, A125, A127] основан на методе Виола-Джонса [135, 168-171].

Особенности алгоритма

Характерные черты (feature) – представляют собой комбинацию двух, трех или четырех прямоугольников. Вычисление характерных черт заключается в вычислении разности значений суммарных интенсивностей светлого и темного прямоугольников (см. рисунок 2.1).

Изображение, представленное ниже поясняет принцип выявления характерных черт (см. рисунок 2.2).

Интегральное изображение – это особое представление исходного изображения, позволяющее производить быстрое вычисление характерных черт (особенностей), используемых при классификации изображения. Каждый пиксель

интегрального изображения представляет собой сумму значений всех пикселей исходного изображения, расположенных выше и левее текущего пикселя (см. рисунок 2.3).

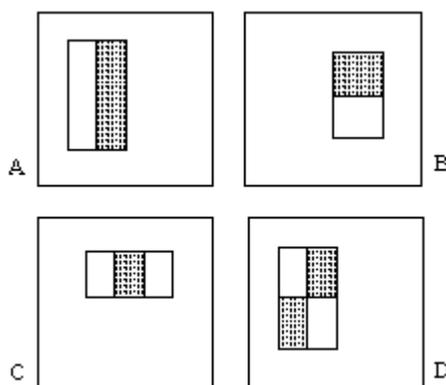


Рисунок 2.1 – Варианты комбинаций для характерных черт

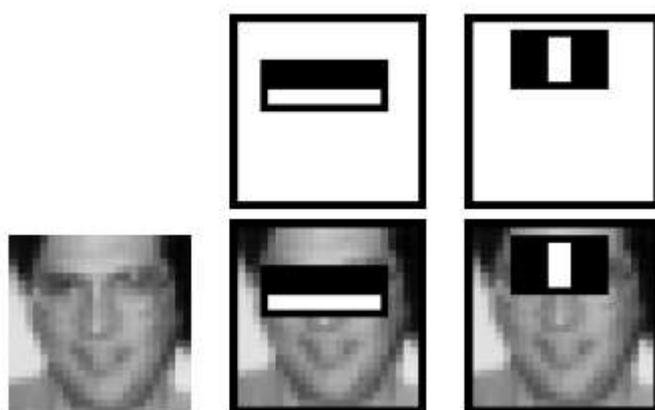


Рисунок 2.2 - Принцип выявления характерных черт.

Это позволяет вычислять сумму пикселей в любом прямоугольнике за четыре операции: $1+4-2-3$, то есть за постоянное время (см. рисунок 2.4).

Слабый классификатор – представляет собой дерево решений, узлами которого являются характерные черты.

Сильный классификатор – это комбинация простых классификаторов, построенная по алгоритму "AdaBoost" [172, 173].

Стандартный сценарий "AdaBoost" состоит в следующем:

- на каждой итерации выбирается «наилучший» слабый классификатор;
- строится линейная комбинация классификаторов;
- на каждой итерации формируется «наилучший» сильный классификатор из фиксированного числа слабых классификаторов;
- ошибка при обучении сводится к нулю с экспоненциальной скоростью.

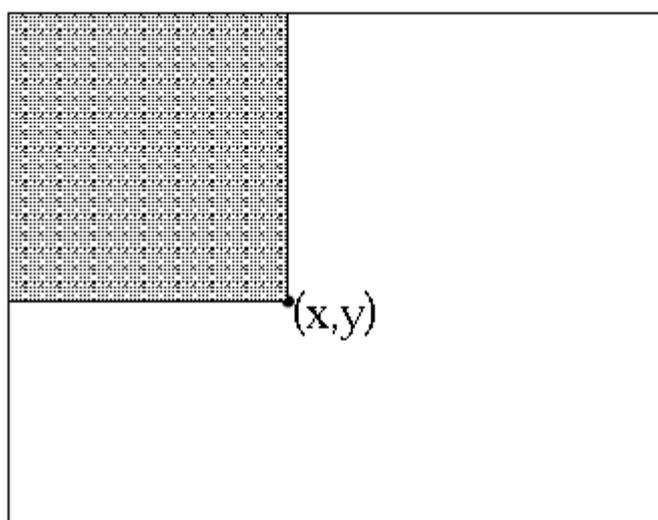


Рисунок 2.3 - Интегральное изображение.

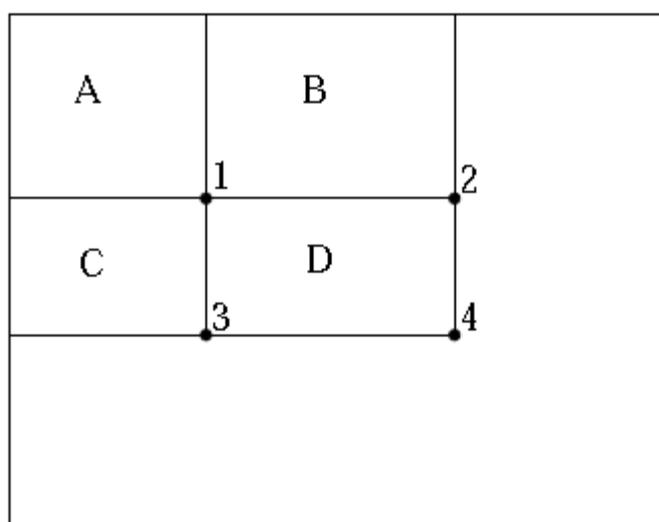


Рисунок 2.4 - Вычисление суммы пикселей в прямоугольнике.

Каскад классификаторов – представляет собой объединение в цепочку сильных классификаторов (выход одного является входом для другого). Если участок изображения не проходит хотя бы один классификатор, он отбрасывается, то есть не идет по всей цепочке. Данное действие значительно ускоряет процесс классификации, так как позволяет сконцентрировать внимание на участках изображения, на которых вероятность обнаружения лица выше. На рисунке 2.5 представлен пример каскада классификаторов.

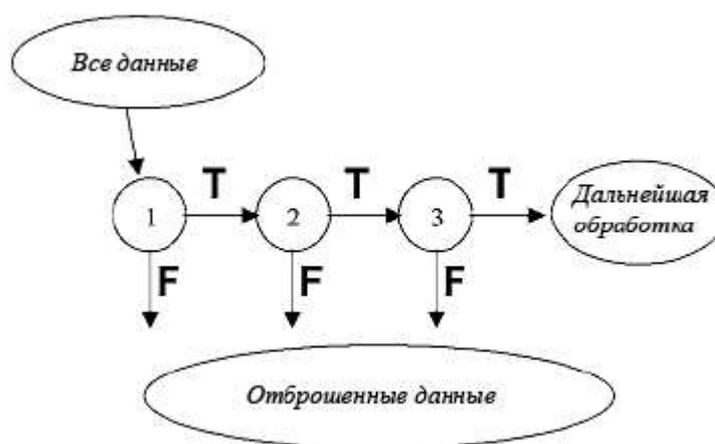


Рисунок 2.5 - Пример каскада классификаторов.

При обучении каскад классификаторов настраивается на определенный размер окна, в котором ведется поиск лица субъекта. Лица, размер которых менее размера окна, каскад «найти» не сможет.

Анализ исходного изображения происходит с использованием скользящего окна, то есть всё изображение сканируется окном, на которое настроен каскад. При этом сканирование осуществляется с шагом в один или два пикселя (в зависимости от масштаба) как по горизонтали, так и по вертикали.

Для обеспечения возможности поиска человеческих лиц различных масштабов производится масштабирование входного изображения с шагом масштабирования до тех пор, пока размер окна поиска не станет больше максимально возможного лица человека или превысит размеры исходного изображения. Если размер этого окна изначально больше входного изображения, то алгоритм прекращает работу. Если окно поиска меньше минимально

возможного (детектируемого) лица, то изображение масштабируется, и снова проверяются перечисленные выше условия.

Далее происходит основной процесс анализа изображения. Сам процесс представляет собой тройной цикл по всем сильным классификаторам каскада, а также по всем слабым классификаторам каждого сильного классификатора и по всем узлам деревьев решений слабых классификаторов. Для каждого слабого классификатора (проверяемого сильного классификатора) вычисляется взвешенная сумма характерных черт. Для сильного классификатора вычисляется сумма значений его слабых классификаторов. Если эта сумма больше некоторого порога (определяется во время обучения), то проверяется следующий сильный классификатор. Если меньше, то делается вывод, что на данном участке изображения нет человеческого лица. Если прошли все сильные классификаторы каскада, то делается вывод о наличии человеческого лица на данном участке (см. рисунок 2.6). Затем происходит усреднение результатов поиска. В итоге обнаруживаются участки изображения, содержащие человеческие лица.

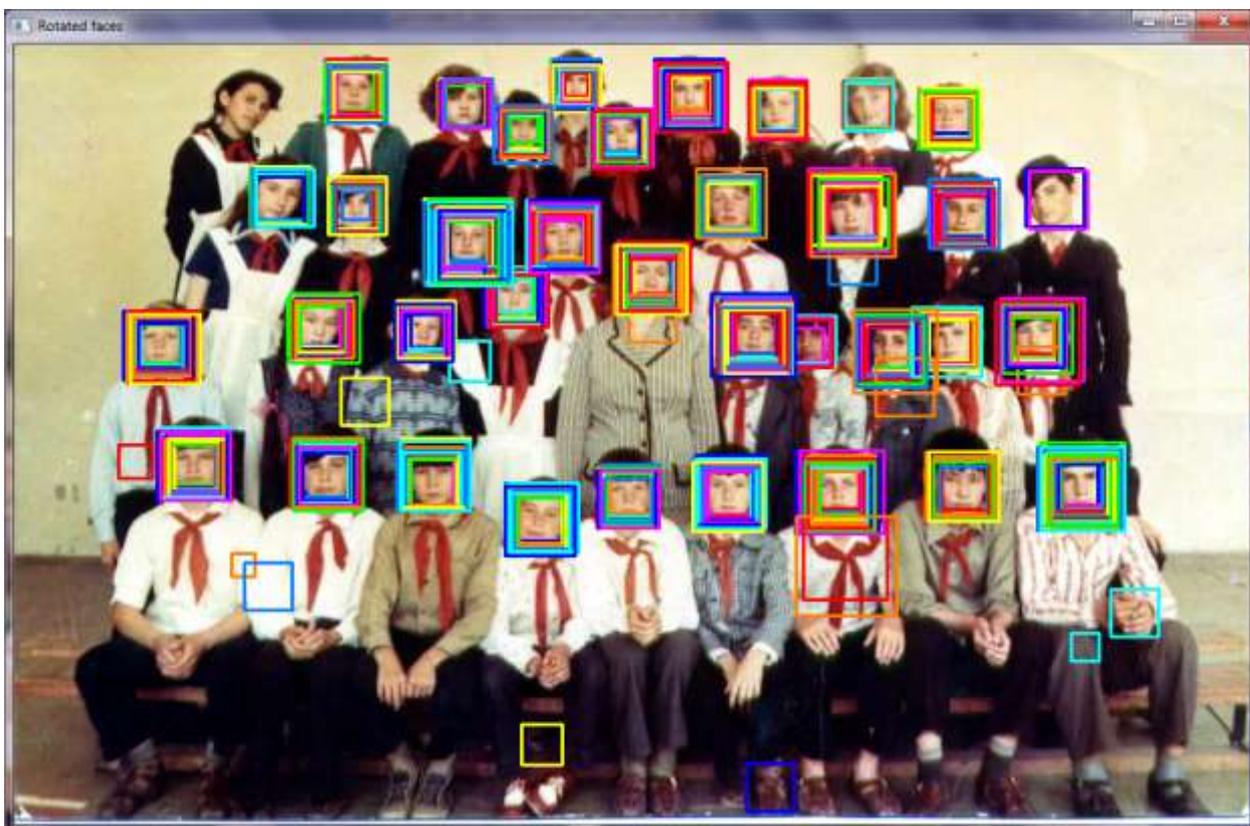


Рисунок 2.6 - Сравнение результатов поиска лица с усреднением и без усреднения.

При работе алгоритма одно лицо может быть обнаружено в разных масштабах. Подобная избыточность позволяет отсечь ложные срабатывания, при этом задается минимальная кратность, при которой лицо должно быть обнаружено. Если число обнаружений меньше заданного значения, то делается вывод о ложном срабатывании. Алгоритма позиционирования приведен на рисунке 2.7.

2.3 Исследование и разработка средств программного обеспечения удаленного контроля, диагностики и обработки телевизионных изображений

Этап применения трансформаций к входному изображению

В литературе [174] изложен способ сравнения изображений, в основе которого лежит принцип построения "групп размытий" исходных изображений и их производных карт фильтрами Гаусса. Для обозначения этих групп введено понятие "пакеты трансформаций". Их сравнение ведется по векторам атрибутов, полученным с помощью инструментов сканирования-сжатия, основная задача которых заключается в корректном уменьшении пиксельной размерности изображений. Исходное изображение может быть подвергнуто некоторому преобразованию, результат которого в дальнейшем является производной картой яркостей (или производной картой). Существенно то, что размерность полученной карты равна размерности исходного изображения. Комбинация такого преобразования и инструмента сжатия определяет процедуру обработки изображения, в результате которой формируются подпространства векторов атрибутов для пакетов трансформаций. Расстояния между получаемыми подпространствами рассчитываются по метрике " $d_{m,n}$ " на базе определителя Грамма. Особенностью метрики является нахождение минимального расстояния между всеми возможными парами векторов атрибутов, образующих подпространства. В зависимости от числа сравниваемых векторов атрибутов возможны частные варианты метрики, которые будут рассмотрены ниже, а именно: $d_{1,1}$ – для сравнения единичных векторов и $d_{3,3}$ – для сравнения подпространств, образованных тремя векторами. Применение соответствующего варианта метрики

к полученным после процедуры «обработки изображения» подпространствам обеспечивает формирование «механизма сравнения». Для сравнения изображений по их небольшим участкам используется изложенный выше пакетный способ сравнения. Однако для формирования пакетов сравнения необходимо применять не размытия исходного изображения, а его небольшие соседние участки. Благодаря особенностям метрики, используемой при пакетном сравнении, перемещение сравниваемых участков в рамках пакета на разных изображениях не повлияет на результат сравнения, то есть позволит компенсировать возможные неточности позиционирования объекта на изображении после юстировки. Более формально преобразования, изложенные выше, могут быть записаны в виде

$$M \rightarrow t(M) \rightarrow P_i t(M) \rightarrow a_i \equiv a(P_i t(M)), i = 1, \dots, m,$$

где M – карта яркости изображения (исходное изображение, хранимое в памяти компьютера в виде матрицы значений интенсивности отдельных его пикселей); $t(M)$ – необязательное преобразование исходной карты, не изменяющее ее размерности, например, свертка фильтрами Собеля или преобразование с использованием текстурных карт Лавса. Задачей данного этапа является извлечение максимального объема информации об объекте, представленном на изображении, например, о его текстуре или форме. Этап " $P_i t(M)$ " формирования пакетов для сравнения из участков изображения необходимо изложить подробнее.

Анализируемый участок изображения следует обозначить "окном интереса". Параметрами, определяющими форму окна, являются его высота и ширина. Предлагается выбрать вокруг исходного окна интереса (номер 1) четыре соседних окна (см. рисунок 2.8). Взаимное расположение выбранных окон определяются значением переменных «шаг 1» и «шаг 2». При таком сравнении участок изображения, представленный окном 1, будет компенсировать возможное его смещение в пределах небольшой окрестности (окна 2–5). Полученный таким образом пакет является «пакетом сравнения». Расстояние между подпространствами векторов атрибутов i а, полученных в результате применения инструментов сканирования-сжатия к сформированным пакетами i а, $P_i t(M)$, определяется метрикой dm, n .



Рисунок 2.7 - Алгоритм позиционирования образа.



Рисунок 2.8 - Схема анализируемого участка изображения.

Необходимо оценить эффективность предложенного способа сравнения изображений при решении задачи классификации фотографий людей с разными выражениями лиц.

Для реализации процедуры выбраны 865 пар изображений FA и FB из базы «ColorFERET», для которых имеются координаты глаз и рта [175]. Данные наборы различаются выражением лиц, изображенных людей. Проведена юстировка выбранных изображений (см. рисунок 2.9). Из этих изображений были составлены два множества следующим образом: первое множество содержит пары изображений FA и FB одного класса (человека), а второе – фотографии разных классов (разных людей). Количество пар в обоих множествах одинаковое и равно 865. Выбор комбинации изображений для пар второго множества при проведении эксперимента несущественен.

В качестве критерия оценки эффективности предлагаемого подхода используется отношение межклассового расстояния к внутриклассовому (K) [148]. Чем выше значение K, тем более эффективно возможно разделение на классы. В

ходе исследования оценивается влияние на значение коэффициента K таких параметров формирования «пакета сравнения», как размер окон и их взаимное расположение.

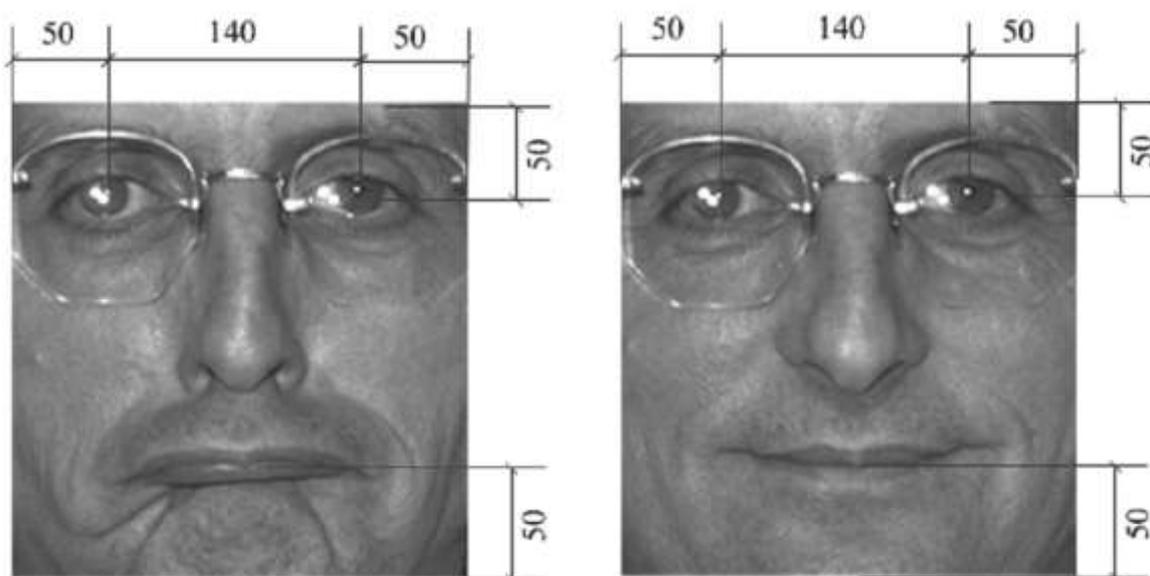


Рисунок 2.9 - Пример изображений после их юстировки.

На процесс формирования пакетов из участков изображения оказывают влияние несколько параметров, таких, как: ширина и высота окна, «шаг 1» и «шаг 2». В ходе исследования процедуры используется квадратное окно, то есть ширина окна равна его высоте, и данный параметр определен как «размер окна». Также для упрощения процедуры исследований зафиксирован размер «шага 2» относительно «шага 1», а именно «шаг 2» равен «шагу 1», умноженному на 2. Таким образом, для вариации процесса формирования пакета определены два параметра, а именно, размер окна и шаг между окнами.

При выполнении процедуры исследования использовались окна с размерами 24×24 , 36×36 и 48×48 пикселей, при этом шаг между ними был равен трети, половине или размеру окна (см. рисунок 2.10). Все исследуемые комбинации этих параметров приведены в таблице 2.1.

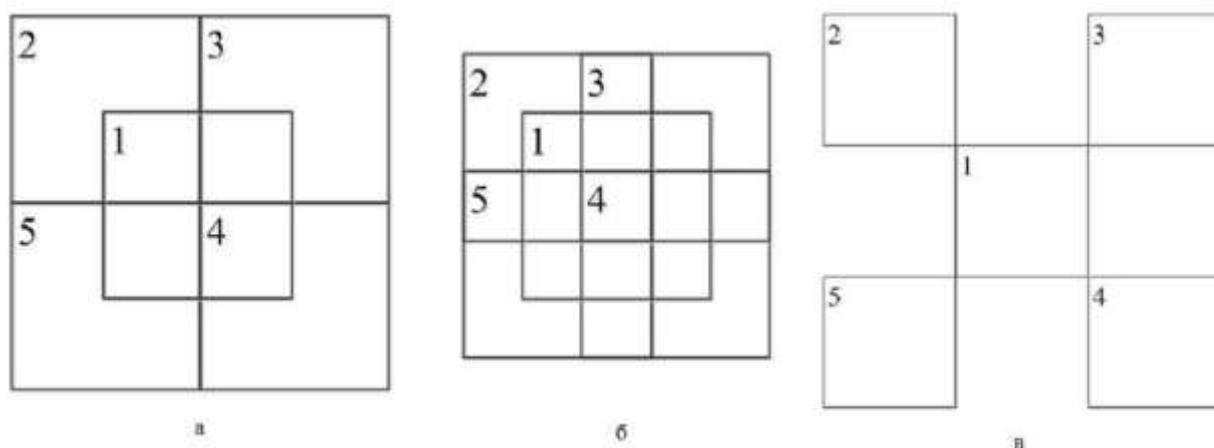


Рисунок 2.10 - Взаимное расположение окон в зависимости от выбранного шага: а – шаг равен половине размера окна; б – шаг равен трети размера окна; в – шаг равен размеру окна

Таблица 2.1 - Исследуемые комбинации параметров для построения «пакетов сравнений».

Идентификатор комбинации	a	b	c	d	e	f	g	h	i
Размер окна	24	24	24	36	36	36	48	48	48
Шаг	8	12	24	12	18	36	16	24	48

Средства преобразования изображений при выполнении процедуры обработки t (М) приведены в таблице 2.2. При использовании инструментов сканирования было установлено ограниченное число полученных в результате их работы коэффициентов. Учитывая незначительные размеры окон, определено число коэффициентов в количестве 50.

Для каждой процедуры обработки изображения (см. таблица 2.2) определены следующие действия:

1) Формирование пакетов сравнения, состоящих лишь из окна 1. Выбор окна осуществляется путем его перемещения по всему изображению с интервалом, равным половине размера окна. Выполнены измерения расстояния для каждой пары изображений по метрике $d_{1,1}$ и расчет критерия оценки эффективности $K_{1,1}$. В результате для каждого из рассматриваемых размеров окон отобраны по 10

участков изображения, на которых было достигнуто максимальное значение величины $K_{1,1}$.

Таблица 2.2 - Состав процедуры обработки изображения.

Идентификатор	Название процедуры обработки изображения	Преобразование	Инструмент сканирования
1	Дискретное косинусное преобразование (ДКП)	Тождественно	Коэффициенты ДКП-преобразования
2	Полиномы Лежандра	Тождественно	Коэффициенты разложения по полиномам Лежандра
3	Карты Лавса	Преобразование с использованием текстурных карт Лавса	Коэффициенты ДКП
4	Маски Собеля	Свертка фильтрами Собеля	Коэффициенты ДКП
5	Карты Лавса + Полиномы Лежандра	Преобразование с использованием текстурных карт Лавса	Коэффициенты разложения по полиномам Лежандра
6	Маски Собеля + Полиномы Лежандра	Свертка фильтрами Собеля	Коэффициенты разложения по полиномам Лежандра

2) Для каждого из 10 выбранных участков построены пакеты сравнения из трех соседних участков изображения следующим образом.

Окно 1 считается основным, соответствующим одному из 10 рассматриваемых участков. Для формирования пакета выбраны еще два соседних к нему участка. Таким образом, возможны следующие комбинации соседних участков: 1 – 2 – 3, 1 – 2 – 4, 1 – 2 – 5, 1 – 3 – 4, 1 – 3 – 5 и 1 – 4 – 5.

Для всех возможных комбинаций параметров формирования пакетов (см. таблица 2.2) выполнены измерения расстояния для каждой пары изображений по метрике $d_{3,3}$ и расчет величины $K_{3,3}$.

3) Определены соотношения $K_m = K_{3,3} \times 3 / K_{1,1} \times 1$ для каждого из анализируемых участков и сформированных на его основе пакетов. Данный критерий характеризует степень увеличения межклассового расстояния относительно внутриклассового при сравнении изображений разработанным способом по отношению к простому анализу независимых участков изображения. Увеличение численного значения данного критерия определяет возможность более качественного различения двух сформированных тестовых множеств.

На рисунке 2.11 приведены средние значения коэффициента K_m , полученные для каждой процедуры обработки изображения.

Выполненные исследования подтвердили возможность применения «пакетного сравнения» с формированием пакетов из пересекающихся участков карты изображения для сравнения изображений. Все комбинации карт изображений и инструментов сканирования показали различные степени уменьшения внутриклассового расстояния относительно межклассового. Наиболее эффективно пакетное сравнение реализуется при анализе изображения относительно крупных окон (порядка 48 пикселей) с достаточной степенью перекрытия соседних окон (шаг равен трети размера окна). Полученные данные подтвердили эффективность использования пакетов сравнения из соседних участков изображения по отношению к анализу отдельных его участков. Интерес представляет оценка эффективности описанного выше пакетного способа сравнения для извлечения признаков из изображения в процессе построения классификатора людей по изображениям их лиц.

Построение классификатора людей по изображениям их лиц

При выполнении исследования отобраны 865 пар изображений из базы «ColorFERET».

Цель исследования - сравнение эффективности различных признаков для построения классификатора лиц людей.

Исследуемые признаки:

1. «Глобальные признаки». Вектор атрибутов, полученный с использованием любой процедуры обработки изображений из таблицы 2.2, примененной ко всему изображению в целом. Количество значащих коэффициентов инструментов сжатия равно 200. Вычисление расстояния между векторами атрибутов производится по метрике $d_{1,1}$.

2. «Локальные признаки 1×1 ». Вектор атрибутов, полученный с использованием процедуры обработки изображений из таблицы 2.2, примененной к участку изображения. Перебор участков изображения осуществляется окном размером 48×48 пикселей, скользящим по всему изображению с шагом в 24

пикселя. Поскольку размер окна в 5 раз меньше размеров полного изображения, то и значащих коэффициентов инструментов сжатия можно использовать пропорционально меньше. Использованы 50 старших коэффициентов. Вычисление расстояния между векторами атрибутов производится по метрике $d1,1$.

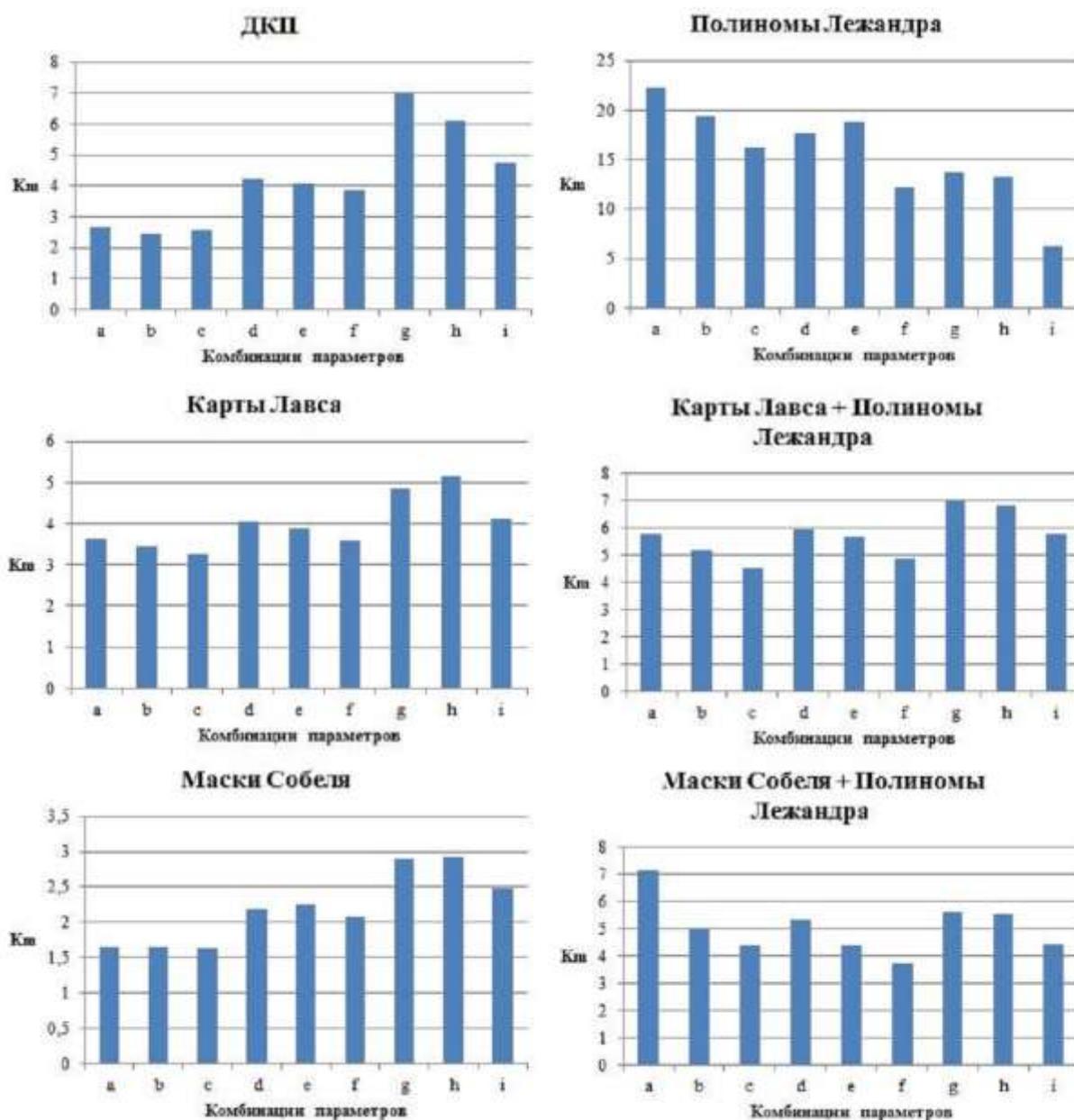


Рисунок 2.11 - Средние значения коэффициента Km.

3. «Локальные признаки 3×3 ». Подпространство векторов атрибутов, полученных с использованием одной из процедур обработки изображений, описанной в таблице 2.2, примененной к участку изображения и его окрестности.

Для выбора участка изображения используется окно размером 48×48 пикселей, скользящее по всему изображению с интервалом в 24 пикселя. Шаг, задающий расположение соседних окон, равен 16. Количество значащих коэффициентов инструментов сжатия равно 50. Вычисление расстояния между подпространствами векторов атрибутов производится по метрике $d_{3,3}$.

Этап принятия решения

Последовательные процедуры проверки гипотез, вследствие их известных достоинств широко используются в различных приложениях [101, 175-178]. Одним из важных достоинств последовательных процедур принятия решений является то, что они минимизируют количество необходимых данных и, как следствие, среднее время принятия решения. Однако известны и трудности их синтеза и анализа, а именно, отсутствие (за исключением отдельных частных случаев) простых аналитических выражений для оценки показателей качества в зависимости от параметров процедур. Синтез и анализ оптимальных последовательных процедур необходимо осуществлять численными методами, что достаточно трудоемко и допускает риск получения лишь частных закономерностей. Поэтому представляют интерес пусть приближенные, но простые аналитические выражения для оптимальных параметров последовательных процедур и показателей их качества в зависимости от существенных факторов, а именно, от числа проверяемых гипотез, от характеристик используемых для проверки гипотез результатов наблюдений (признаков), от распределения результатов наблюдений (признаков) по этапам процедуры и т.п. [101]. Данные факторы могут быть учтены для нахождения первого приближения при поиске оптимальных параметров процедур численными методами [176].

Модель и закономерности принятия решений

При распознавании образов по их видео изображениям (или по синтезируемым трехмерным изображениям) решение принимается по реализации доступных наблюдению данных $X = (x_1, x_2, \dots, x_n)$. Это могут быть первичные или преобразованные данные [179]. Эти данные содержат случайные компоненты той

или иной природы. В работе эти случайные компоненты обусловлены влиянием условий съемки, наличием фактора «маскировки», «шумами» при получении и обработке изображений и другими причинами. Как следствие, возникает типовая ситуация статистической теории распознавания образов. Особенности в данном приложении являются: размеры базы образов (число гипотез); наличие неаддитивных помех; отсутствие статистических моделей влияния на доступные наблюдению данные вышеуказанных случайных компонентов. Вместе с тем, изложенные далее закономерности принятия решений на этапах и в результате последовательной процедуры проверки многих гипотез в целом имеют достаточно общий характер и применимы в рассматриваемом приложении.

Обоснование алгоритмов решения задачи распознавания образа по биометрическим инвариантным признакам

Исходя из требований к разрабатываемой системе, должно обеспечиваться автоматическое обнаружение изображений людей, попавших в зону наблюдения, с последующим их распознаванием. При этом изображение является движущимся со скоростью до 10 км/час люди и необходимо выполнить видео сопровождение и маршрутизацию. Здесь следует отметить, что функция сопровождения может быть использована не только при реагировании на срабатывание системы распознавания, но и для повышения качества синтеза трехмерных изображений субъектов по их двумерным проекциям.

Селекция на изображениях движущихся объектов, их классификация (человек или нет) и сопровождение представляется относительно простой задачей.

При скорости движения человека от 1 км/час (медленное движение) до 10 км/час (бег) его изображение будет смещаться в поле зрения камер со скоростью от 0.3 до 3 м/сек. Как следствие, разностное изображение кадров, регистрируемых, например, через 1 секунду, будет содержать необходимую и достаточную информацию для обнаружения и сопровождения движущихся объектов.

Среди выделенных таким образом движущихся объектов могут быть не только люди. Это могут быть движущиеся блики и тени окружающей среды или предметы (например, газета или полиэтиленовый пакет, движимые ветром).

Классификация этих объектов (человек это или нет) может быть выполнена разными методами.

Во-первых, могут быть использованы методы «траекторной» селекции (человек или, напротив, газета или полиэтиленовый пакет, движимые ветром, движутся по-разному).

Во-вторых, может быть использована селекция «по шаблону» фигуры человека.

В-третьих, может быть использована селекция на основе выделения (распознавания) в изображении движущегося объекта фрагмента, который может принадлежать только человеку (например, изображение лица). Именно такой способ используется в выбранном варианте реализации алгоритма осуществления маршрутизации. Данный алгоритм отслеживает только те объекты, которые предварительно обнаружил алгоритм позиционирования образа объекта, то есть классифицированные как человек.

Выделение лица как фрагмента изображения

Положение человека и его лица на изображениях может достаточно существенно изменяться. Поэтому в конечном виде (после экспериментальных исследований на различных этапах ОКР) выделение лица может потребовать использования совокупности частных алгоритмов, объединенных логическими функциями «И» и «ИЛИ».

Одним из таких частных алгоритмов может быть алгоритм, осуществляющий поиск на изображении фрагментов, удовлетворяющих инвариантным для любого человеческого лица признакам. Описание этих инвариантных признаков приведено в приложении Б.

Логическими «И» и «ИЛИ» можно объединить несколько частных алгоритмов, что позволит получить робастный к разным лицам и их ориентации составной алгоритм выделения лица. Но с целью снижения вычислительной сложности подобного составного алгоритма, следует предусмотреть его адаптацию к текущей ситуации с использованием оператора «ЕСЛИ...,ТО...».

Распознавание объектов, классифицированных как «человек»

Распознавание объектов, классифицированных как «человек», может выполняться как по отдельным изображениям с камер (изображениям, на которых обнаружено лицо), так и по совокупности изображений, позволяющим сформировать трехмерную модель лица. Первый вариант принимается как основной, второй вариант выполняется по запросу оператора.

В обоих случаях используются биометрические инвариантные признаки лица людей.

К настоящему времени в габитоскопии (науке о внешнем облике человека, методах собирания и использования данных о внешнем облике человека [180, 181]) накоплено достаточно много знаний о биометрических инвариантных признаках лиц людей. В этих работах предлагается использовать десятки и даже сотни признаков, позволяющих более или менее точно описать и распознать лицо конкретного человека.

Теория распознавания образов рекомендует относиться к проблеме выбора признаков с особым вниманием [141, 143, 149, 179, 182]. От выбора признаков существенно зависит вычислительная сложность алгоритмов распознавания объектов. От их выбора зависит также и эффективность этих алгоритмов.

В частности, компьютер в процессе обработки и распознавания изображений объектов «не обязан видеть» 2D образ объекта или 3D образ объекта точно так же, как его видит человек (это может потребоваться лишь на этапе представления результата распознавания человеку-оператору комплекса). Однако для распознавания объектов часто достаточно их 2D или 3D моделей, учитывающих лишь контуры объектов и их элементов.

Использование подобных моделей необходимо и потому, что изображения из базы данных зачастую получены при совсем других условиях освещения и качестве съемки. Приведение изображений к единым условиям освещения является задачей, которую потенциально (при достаточном вычислительном ресурсе) решить можно, но ее решение представляется в настоящее время актуальным лишь для этапа

представления результатов автоматического распознавания объекта человеку-оператору комплекса для их утверждения.

2.4 Выводы по 2-й главе

1) Обоснованы методы и средства процесса распознавания и обработки телевизионных изображений. Разработана структура процедуры распознавания лица человека на цифровом изображении.

2) Исследованы и разработаны информационно-технологические модели и алгоритмы распознавания и обработки телевизионных изображений. Обоснованы этапы предварительной подготовки изображения. Определена эффективность применения методов признакового описания.

3) Исследованы и разработаны средства программного обеспечения удаленного контроля, диагностики и обработки телевизионных изображений. Проработан способ пакетного сравнения эталонных и тестовых изображений, предопределяющий процедуру предварительного применения вполне определенных трансформаций.

4) Впервые разработана программа полуавтоматической аппроксимации экспериментальных данных базисным набором гауссиан, вошедшая в состав аппаратно-программного комплекса (см. Свидетельство № 2017611477).

5) Впервые разработано программное обеспечение автоматизированного тестирования алгоритмов распознавания лиц, вошедшее в состав аппаратно-программного комплекса (см. Свидетельство №2014612123) (см. Приложение А).

Глава 3 Исследование и разработка средств технического и методического обеспечения процессов распознавания, контроля, диагностики, испытания и управления видеосистемами

3.1 Разработка структуры информационно-технических средств визуализации процессов распознавания, контроля, диагностики и испытания видеосистем

Общая структура

Варианты применения разрабатываемой системы могут быть разделены на две принципиально различные группы:

- однонаправленный поток людей высокой интенсивности, сконцентрированный в небольшом пространстве, то есть пересекающий ограниченный участок некоторой плоскости (вход в здание, на эскалатор, к стойке регистрации пассажиров и т.п.) – схема 1;

- одно- или двунаправленный поток людей средней интенсивности (два чел/сек) на достаточно протяженной (порядка 50×15 м) области пространства (коридор, подземный переход, тротуар и т.п.) – схема 2.

Для реализации каждого из этих вариантов применяется специализированная последовательность действий, однако в обеих схемах используется унифицированный набор операций алгоритмического обеспечения.

Решение задачи в первом варианте по схеме 1

Блоки одного модуля передающей части размещаются стационарно с таким расчетом, чтобы обеспечить стереосъемку зоны контроля с разрешением, достаточным для надежной идентификации человека (около 1 пикселя на 1 мм согласно ГОСТ Р ИСО/МЭК 19794-5-2006). Пример размещения видеокамер приведен на рисунке 3.1.

При установке видеокамер проводится их калибровка в плоскости зоны контроля. В связи с этим каждому объекту, находящемуся в данной плоскости и

видимому на кадре одной из камер, может быть сопоставлен тот участок кадра любой другой камеры, где он должен находиться в данный момент времени.

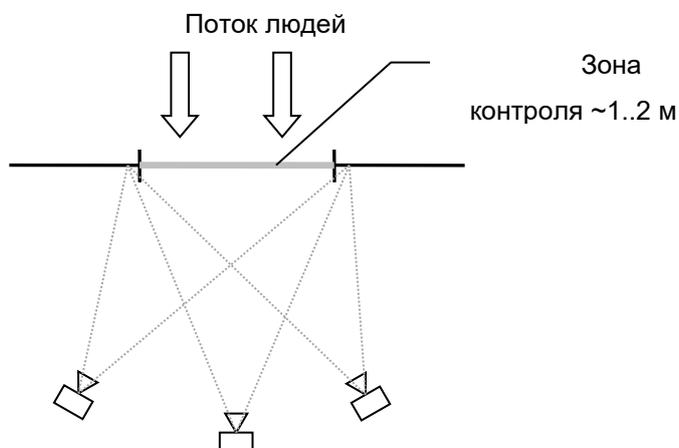


Рисунок 3.1 - Размещение видеокамер передающей части в схеме 1.

Принцип (алгоритм) функционирования системы на примере взаимодействия двух видеокамер из трёх изложен на рисунке 3.2.

Здесь необходимо отметить, что в схеме 1 потенциально возможно ошибочное сопоставление с одним человеком изображений лиц различных людей. Пример подобной ситуации приведен на рисунке 3.3. Здесь изображения объектов O_1 , O_2 , O_3 проецируются в одну и ту же область плоскости зоны контроля и могут быть ошибочно сопоставлены с несуществующим объектом O_4 .

Для предотвращения подобных ошибок в алгоритме сопоставления необходимо предусмотреть дополнительные средства анализа «сходства» изображений. Такими средствами могут быть некоторые статистические показатели (например, среднее значение, дисперсия и др.), вычисленные для определенного цветового диапазона.

Решение задачи во втором варианте по схеме 2

Блоки передающей части видеосистемы размещаются на значительном удалении от зоны контроля (до 50 м), которая представляет собою прямоугольную площадку до 15м×15м. В составе блоков задействуются как стационарные камеры

(СК) с фиксированными фокусами в составе ИК-датчиков (выполняющие обзорную функцию), так и камеры на поворотных устройствах (ВК) с управляемыми вариообъективами (выполняющие следящую функцию). Вариант размещения камер для однонаправленного потока приведен на рисунке 3.4. Если поток движения людей двунаправленный, то модули передающей части располагаются друг другу навстречу.

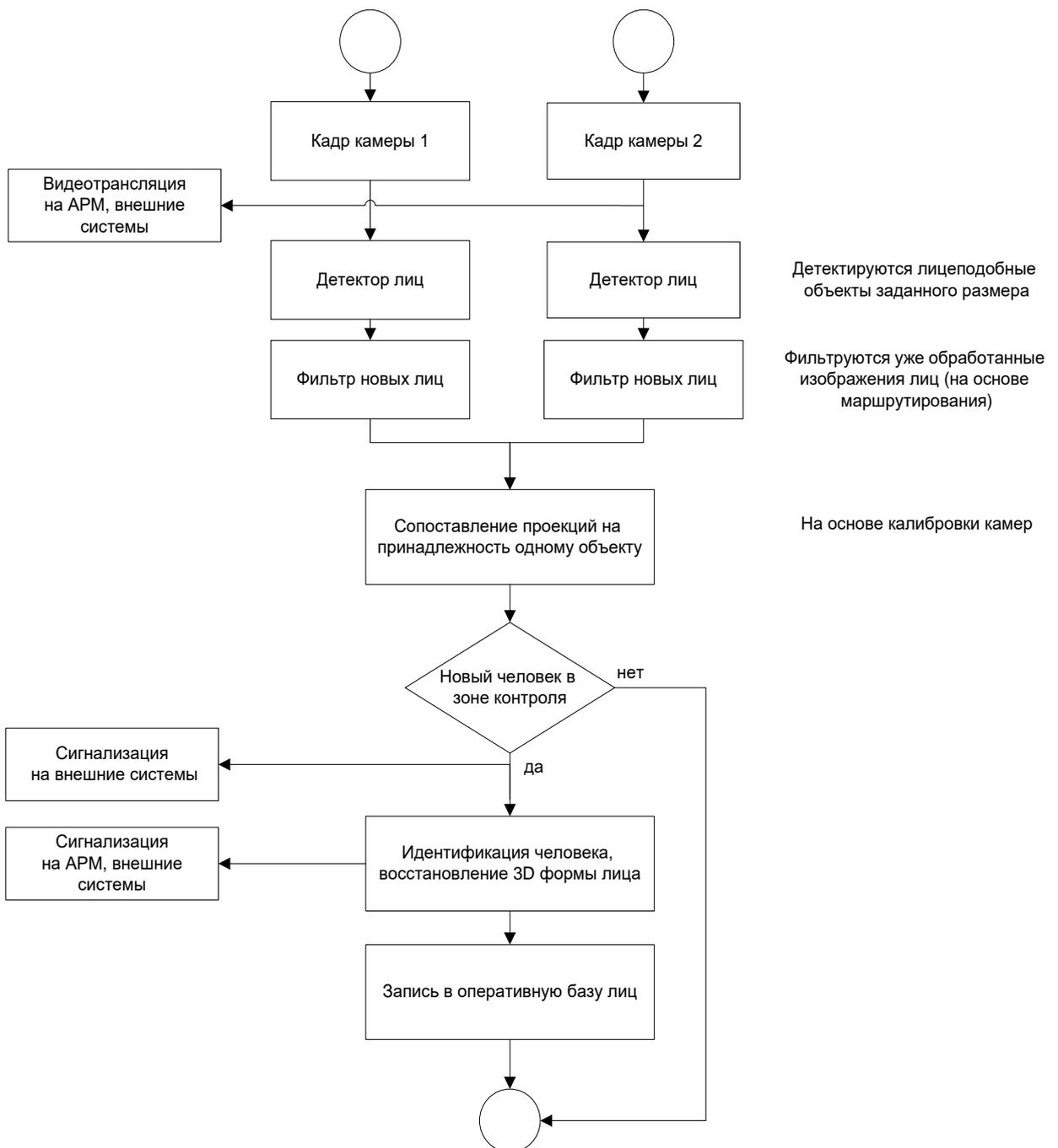


Рисунок 3.2 - Алгоритм взаимодействия двух видеокамер из трёх.

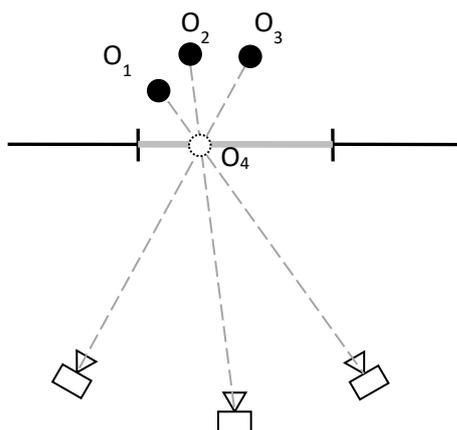


Рисунок 3.3 – Пример ошибочной классификации в схеме 1.

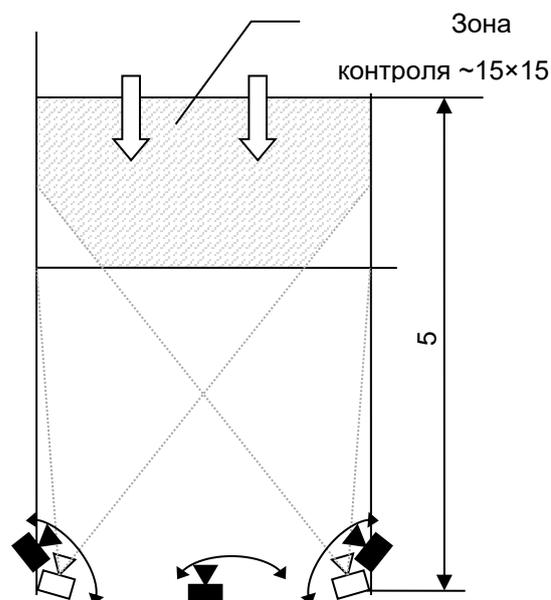


Рисунок 3.4 - Размещение камер передающей части в схеме 2.

Ширина зоны контроля (до 15м) обусловлена тем, что на максимальном удалении 50 м обзорная камера должна фиксировать лицо человека с разрешением, достаточным для его детектирования (не менее 20 пикселей в ширину).

При установке комплекса стационарные камеры калибруются так, чтобы для любого объекта в зоне контроля, видимого на кадре СК1, можно было вычислить вероятную область его нахождения на кадре СК2. Данная информация

используется как для ускорения процедуры сопоставления двух проекций одного человека, так и в дальнейшем для вычисления расстояния до объекта. Процедура сопоставления для схемы 2 отличается от используемой процедуры в схеме 1 и является более сложной. В схеме 2 предполагается, что проекции объекта на кадрах СК1 и СК2 отличаются незначительно вследствие большой удаленности. Это дает возможность использовать корреляционный анализ для определения идентичности фрагментов двух кадров. Каждый фрагмент кадра СК1, идентифицированный как «лицо», сравнивается попарно со всеми аналогичными фрагментами кадра СК2, попадающими в область, вычисленную на основании данных калибровки. Максимум корреляционной функции дает искомый фрагмент.

Принцип (алгоритм) работы комплекса по схеме 2 приведен на рисунке 3.5.

Структурная организация комплекса

Стенд реализуется в составе следующих подсистем:

- подсистема регистрации;
- подсистема распознавания;
- подсистема управления;
- подсистема интеграции;
- подсистема пользовательского интерфейса.

Схема размещения указанных подсистем в составе вычислительных узлов приемной части комплекса приведена на рисунке 3.6.

Конструктивно, каждая подсистема состоит из набора ряда компонентов – сборок “.NET”, исполняемых файлов, файлов конфигурации, баз данных, динамических библиотек и т.п. Размещение компонент в узлах комплекса определяется функциями, выполняемыми каждой из подсистем в составе того или иного узла. При этом некоторые компоненты связаны с внешними службами, дополнительно устанавливаемыми в соответствующих узлах (см. рисунок 3.7).

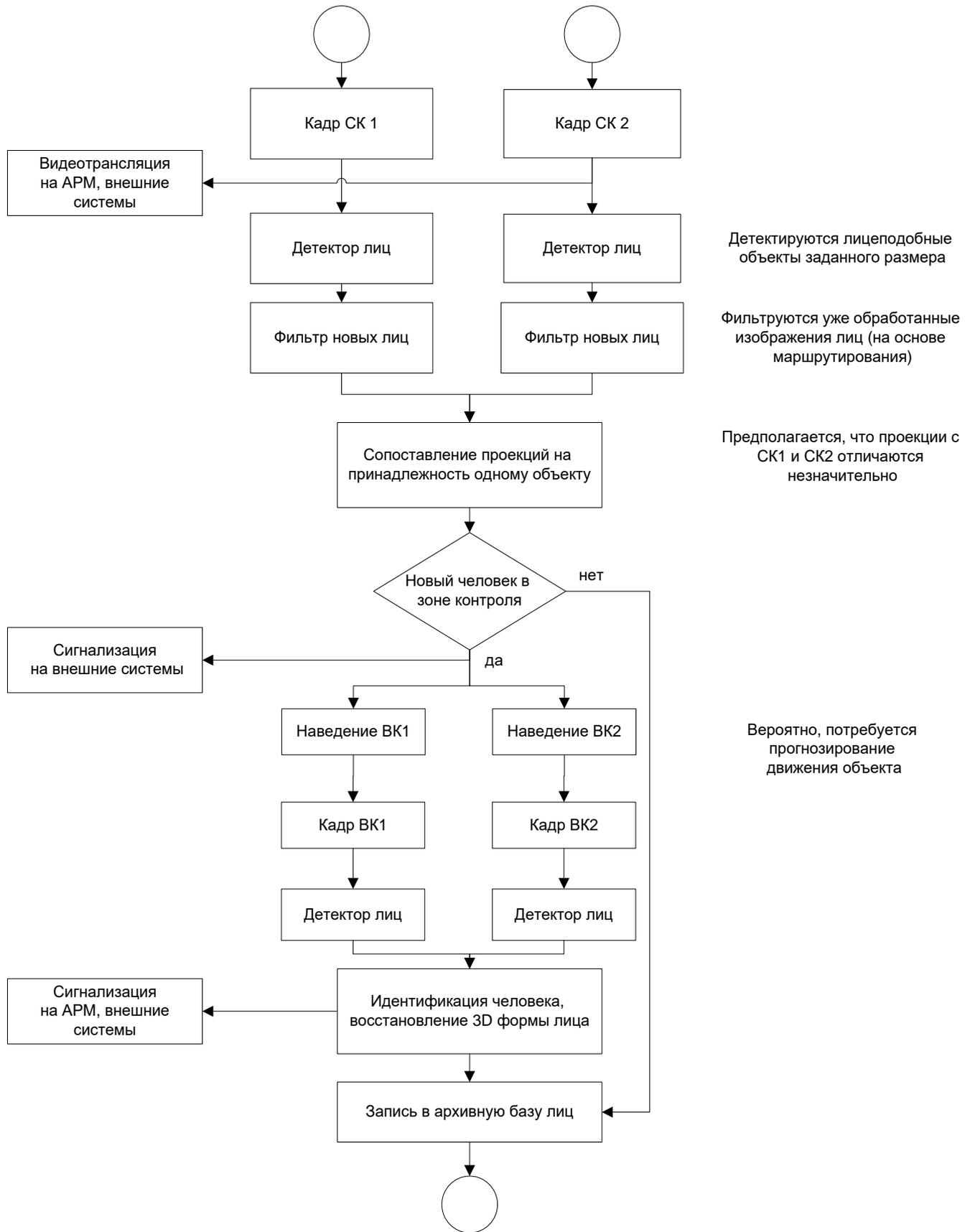


Рисунок 3.5 - Алгоритм работы комплекса по схеме 2.

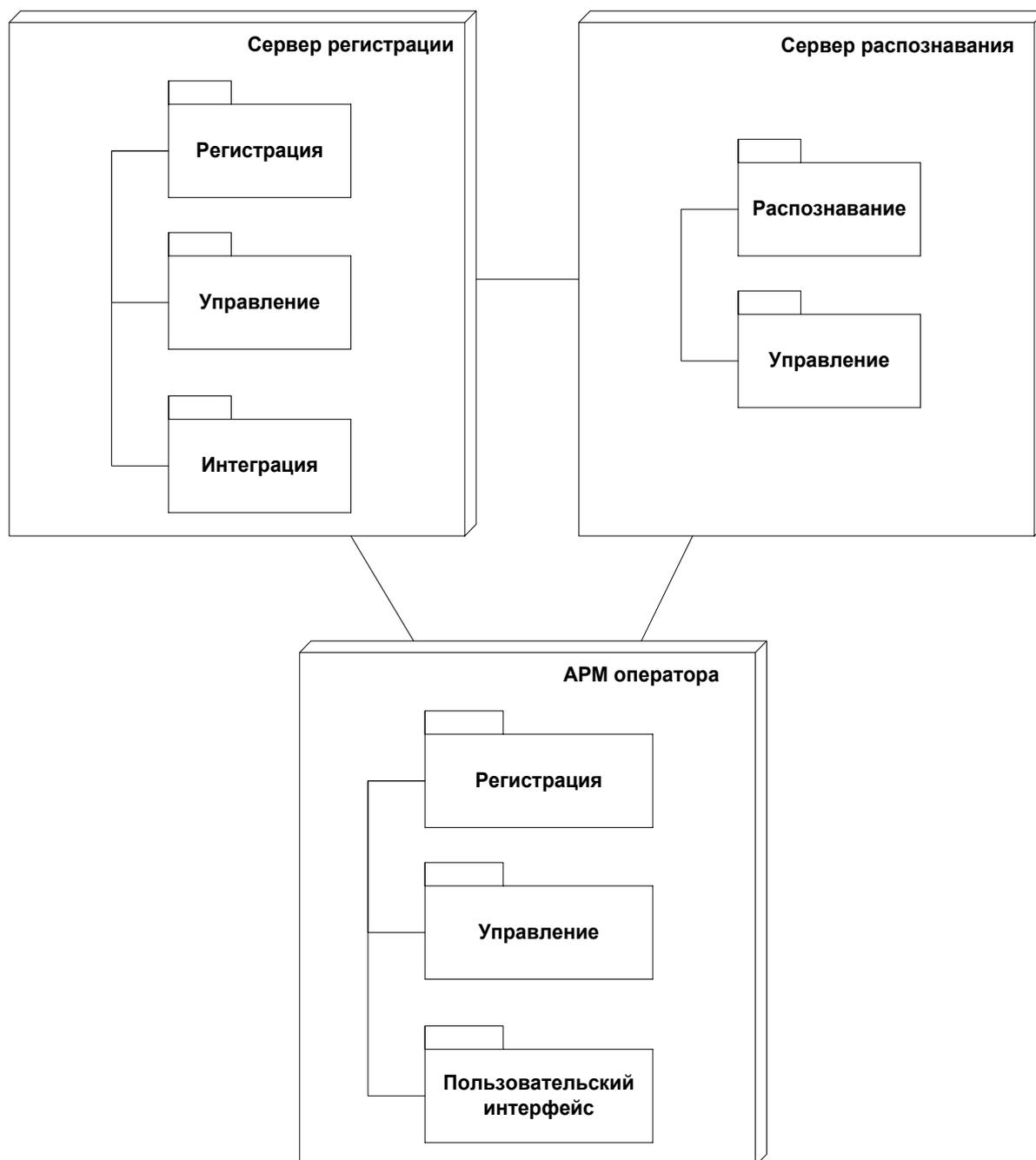


Рисунок 3.6 - Размещение подсистем в узлах приемной части комплекса.

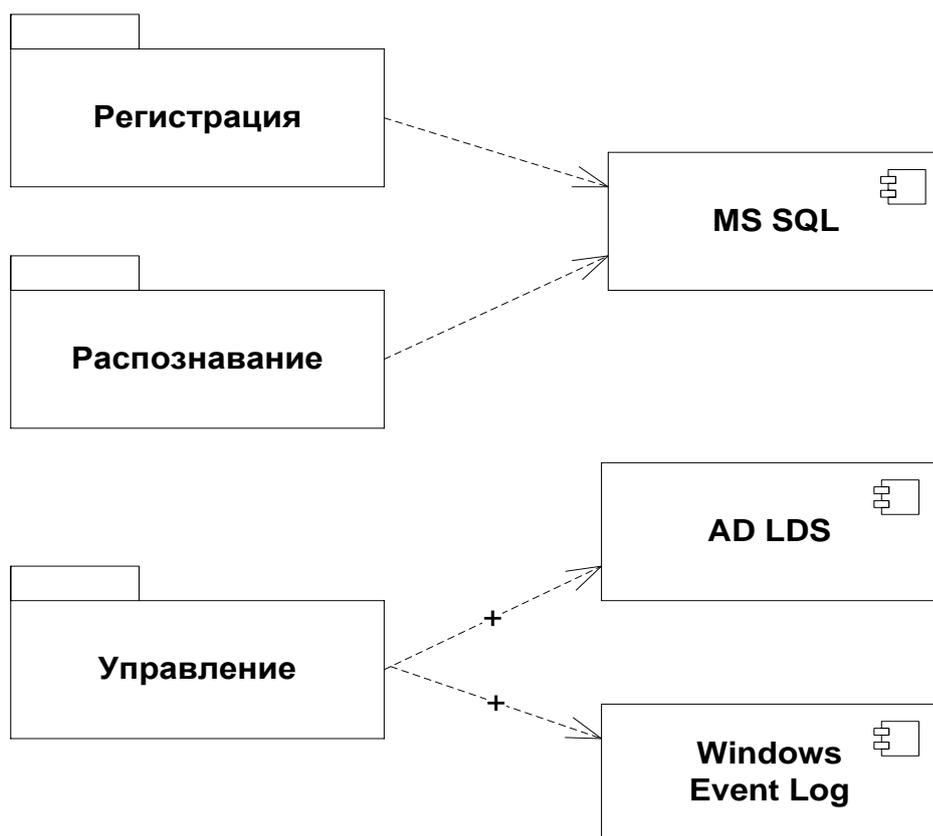


Рисунок 3.7 - Внешние взаимосвязи подсистем комплекса.

3.2 Разработка средств информационно-программного кодирования для взаимодействия элементов визуализации с элементами интерактивного управления видеосистемами

Описание подсистем

Подсистема регистрации. Назначение и состав

Подсистема регистрации предназначена для получения видеоданных с устройств передающей части, их декодирования и последующего анализа с целью обнаружения объектов (лиц людей) в зоне наблюдения. В задачи подсистемы регистрации также входит ведение архивной базы данных и видеозапись.

Подсистема регистрации является источником данных для ключевых функций комплекса, поэтому она тесно интегрирована со всеми остальными

подсистемами. В состав подсистемы регистрации входят компоненты, изображенные на рисунке 3.8.

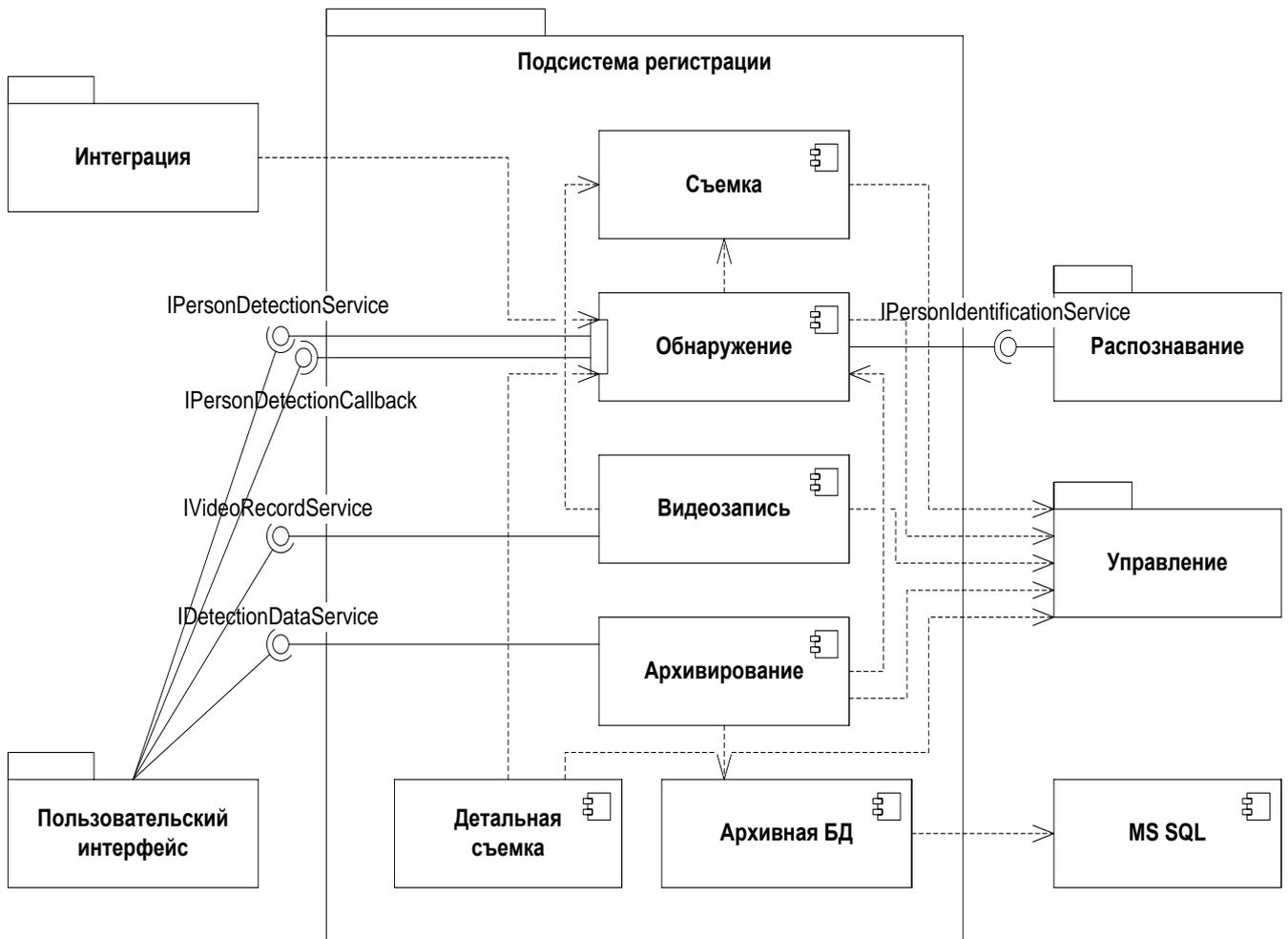


Рисунок 3.8 – Состав подсистемы регистрации.

Все компоненты непосредственно участвуют в конвейере видеообработки.

Компонент «Съемка» предназначен для получения и декодирования видеоданных с цифровых камер. Декодирование представляет собой распаковку видеоданных из потокового формата (например, MPEG) в формат RGB24, пригодный для применения алгоритмов обработки изображений.

Компонент включает два блока конвейера («DirectShow»-фильтра):

- источник видеопотока, осуществляющий подключение к камере и получение видеоданных по сетевому интерфейсу. Данный блок разрабатывается

отдельно для каждого типа устройств регистрации передающей части (мегапиксельная камера, ИК-камера, электронно-оптический преобразователь), и реализуется на основе комплектов программных библиотек, поставляемых производителями устройств.

- декодер видеопотока.

В качестве основного формата передачи видеоданных от передающей к приемной части предполагается использовать MPEG-4 Part 10 AVC (H.264). Для его декодирования запланировано применить аппаратные декодеры графических адаптеров «Nvidia», входящие в состав узлов. Их возможности позволяют одновременно декодировать до 6 видеопотоков H.264 (на двух видеоадаптерах), практически полностью освобождая от данной задачи центральный процессор.

Таким образом, данный блок реализуется в виде DirectShow-фильтра на основе «Nvidia SDK». Необходимо отметить, что при использовании другого (менее ресурсоемкого) формата кодирования видеопотока в качестве данного блока может быть задействован один из стандартных DirectShow-фильтров декодирования.

Компонент «Обнаружение» предназначен для выявления видимых на видеокдрах лиц людей и их последующего сопровождения с целью исключения повторного распознавания. Компонент обеспечивает выполнение действий на следующих этапах в конвейере видеообработки:

- обнаружение лиц людей - реализует алгоритм Виола-Джонса;
- трассировка движущихся объектов;
- фильтрация новых объектов и анализ ракурса лица для предварительной оценки возможности распознавания.

Конструктивно на 1-м этапе выполняется один блок конвейера, а на 2-м и 3-м этапе – другой. Эти два блока функционируют параллельно с целью минимизации совокупной временной задержки. При этом второй блок обеспечивает реализацию методов интерфейса IPersonDetectionService, сообщая другим подсистемам комплекса о появлении и передвижении объектов в зоне наблюдения.

Компонент «Детальная съемка» является вспомогательным и задействуется только в режиме работы комплекса по «Схеме-2», то есть при наличии «следящих» камер. Данный компонент реагирует на обнаружение новых объектов в зоне контроля и обеспечивает наведение и фокусировку камер для укрупненной съемки распознаваемого лица на большом расстоянии. На рисунке 3.9 приведена упрощенная схема взаимодействия компонентов подсистемы регистрации в данном режиме. Для наглядности изображены только два конвейера – для одной обзорной камеры и одной следящей.

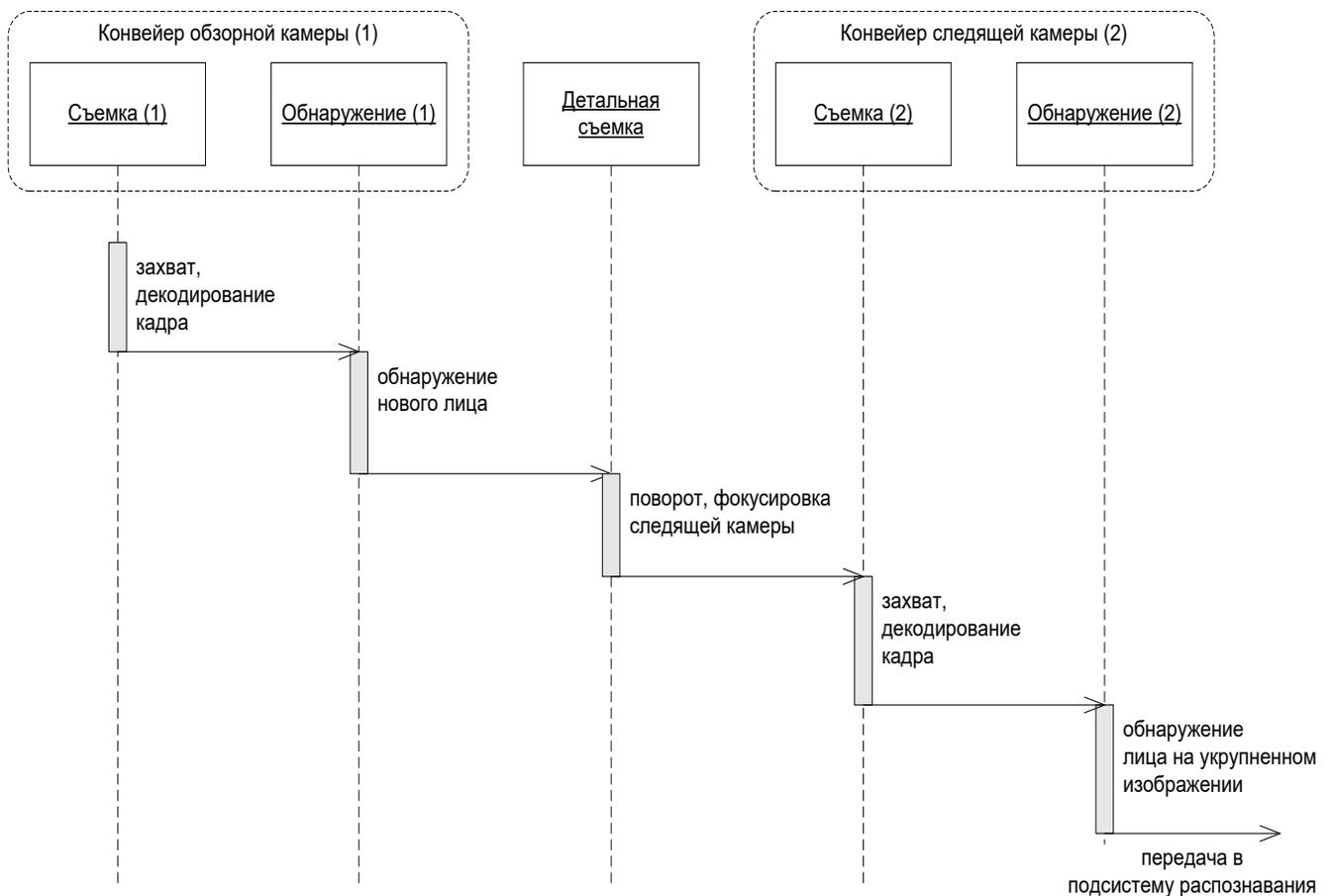


Рисунок 3.9 - Взаимодействие конвейеров обзорной и следящей камер по «Схеме-2».

Компонент «Архивирование» (см. рисунок 3.10) предназначен для ведения архивной базы данных, при этом в базу вносятся изображения всех обнаруженных

объектов, время их обнаружения и результаты распознавания. Компонент реализует интерфейс IDetectionDataService для просмотра и поиска данных в архивной базе.

Компонент «Видеозапись» (см. рисунок 3.10) предназначен для ведения потоковой записи данных, получаемых с камер наблюдения. Запись осуществляется в виде видеофайла на жесткий диск узла системы в одном из стандартных форматов. Компонент реализуется на основе существующих DirectShow-фильтров видеозаписи.

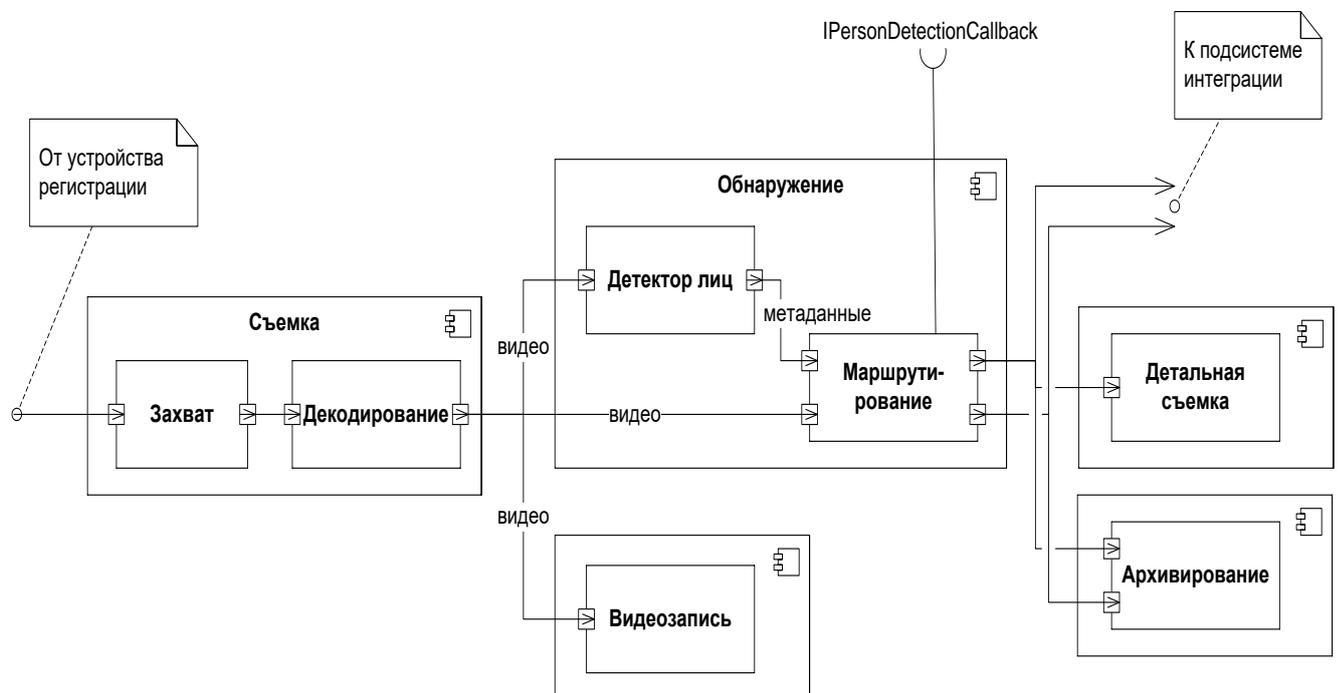


Рисунок 3.10 - Схема взаимосвязи основных компонентов комплекса.

Подсистема распознавания. Назначение и состав

Подсистема распознавания предназначена для идентификации личности человека по изображению лица, фиксируемого подсистемой регистрации. Также в задачи подсистемы входит обслуживание оперативной базы идентифицируемых личностей.

В состав подсистемы распознавания (см. рисунок 3.11) входит четыре компонента. Подсистема локализована на выделенном сервере распознавания, и её

взаимодействие с другими подсистемами ограничено внешними интерфейсами WCF-сервисов, то есть её компоненты не участвуют в конвейере видеообработки.

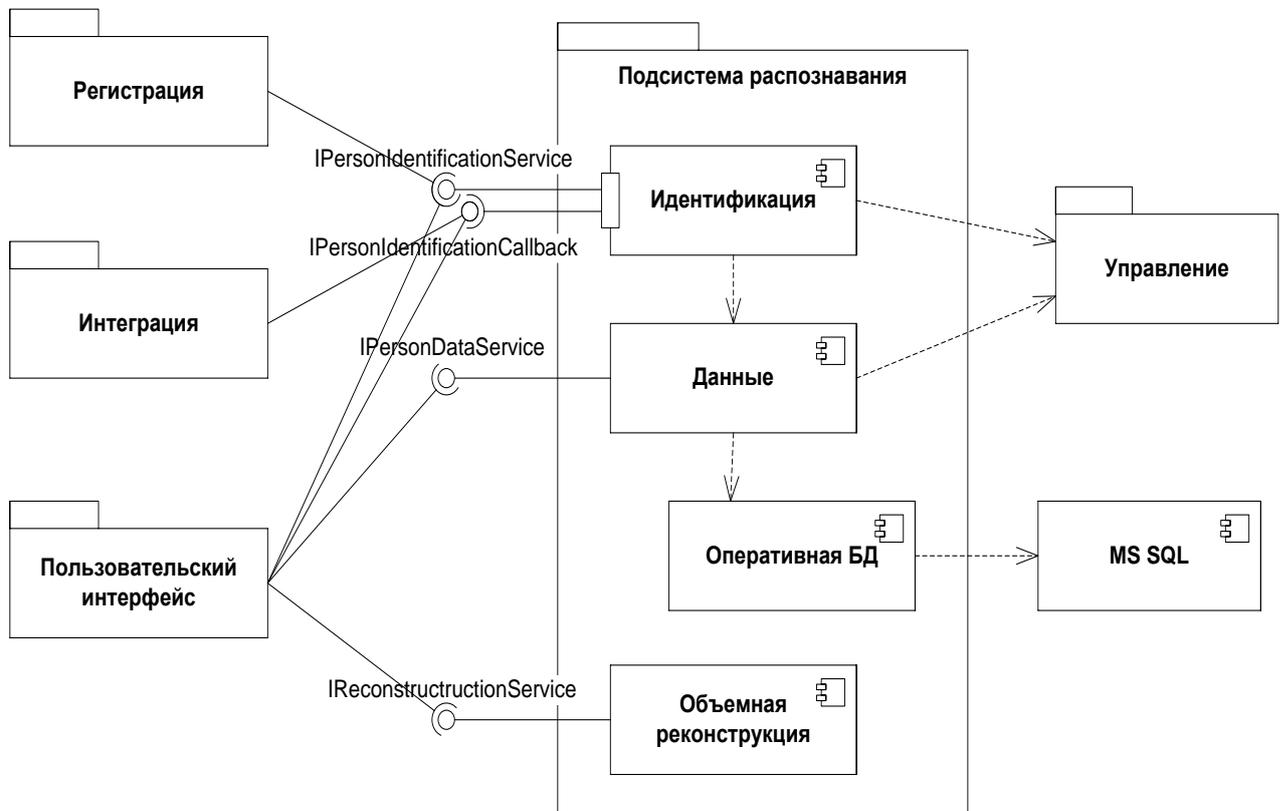


Рисунок 3.11 - Подсистема распознавания.

Компонент «Идентификация» выполняет комплекс операций по обработке изображений с целью их сопоставления с изображениями в оперативной базе данных. Компонент является центральным звеном всей системы и реализует набор специально разработанных алгоритмов. Согласно запроектированной конструкции, механизм работы компонента «Идентификация» заключается в следующем.

Подсистема регистрации в автоматическом режиме (через интерфейс IPersonIdentificationService) запускает процесс распознавания для каждого вновь обнаруженного объекта. При этом в качестве входного параметра передается откадрированное фронтальное изображение лица человека. По завершении процесса распознавания подсистема оповещает о его результатах всех

«подписавшихся» на уведомление клиентов, для чего предназначен интерфейс обратного вызова `IPersonIdentificationCallback`. В качестве клиентов выступают подсистема интеграции и подсистема пользовательского интерфейса.

Компонент «Объемная реконструкция» предназначен для воссоздания трехмерной формы распознаваемого лица по зарегистрированным планарным изображениям (проекциям). Компонент задействуется по запросу оператора и позволяет выполнить дополнительный анализ указанного человека, попавшего в зону наблюдения видеокамер передающей части. При наличии трехмерного эталона данный компонент является дополнительным средством повышения надежности распознавания. Принцип действия и предполагаемые алгоритмы функционирования компонента изложены в главе 2.

Компонент «Данные» обслуживает все операции с оперативной базой данных. Посредством интерфейса `IPersonDataService` данный компонент предоставляет возможность для пополнения, редактирования, поиска и резервного копирования содержащейся в базе данных информации. Данный интерфейс задействуется подсистемой пользовательского интерфейса, в частности, для отображения детальной информации о распознанных объектах на экране монитора оператора.

Подсистема управления. Назначение и состав

Подсистема управления объединяет несколько разнообразных компонент, выполняющих разнообразные служебные функции, необходимые для продуктивной работы остальных подсистем. По сути, подсистема управления выступает в роли промежуточной среды, обеспечивающей инфраструктурную поддержку всем функциональным компонентам специализированного прикладного программного обеспечения (СППО). Ключевые задачи, решаемые подсистемой, это: организация взаимодействия как между компонентами СППО, так и со службами операционной системы; обработка конфигурационных данных; разграничение доступа; "журналирование". Помимо этого, подсистема управления включает загрузочный модуль, содержащий основную точку входа в программу (служба, запускаемая при старте операционной системы).

В связи со спецификой подсистемы, все её компоненты разворачиваются в каждом узле приемной части аппаратно-программного комплекса (АПК).

Подсистема управления

Подсистема содержит следующие компоненты.

Компонент «Сервисы» объединяет все внешние определения интерфейсов программного взаимодействия между компонентами (то есть запускает определение интерфейсов сервисов, согласно проектной модели). К данному компоненту относятся как сами интерфейсы, так и типы операндов и возвращаемых значений. Поскольку подобные определения необходимы всем компонентам, участвующим во взаимодействии (как предоставляющим сервис, так и использующим его), их обособление в отдельный модуль позволяет с одной стороны избежать циклических ссылок в зависимостях, а с другой – исключить дублирование этих определений и, как следствие, необходимость их синхронизации при изменении (см. рисунок 3.12).

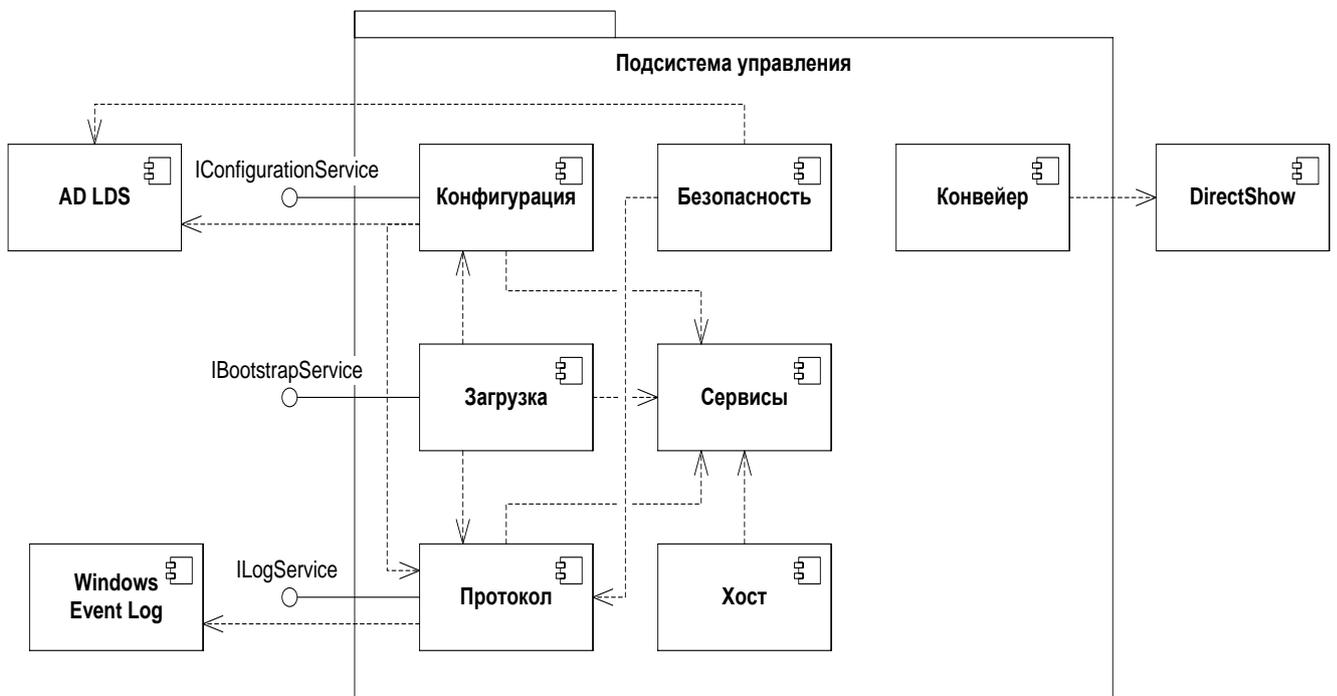


Рисунок 3.12 - Подсистема управления.

Компонент «Хост» является исполняемым файлом службы операционной системы и содержит код для автоматического запуска всей программной системы. Данный компонент создает в каждом узле процесс, в адресном пространстве которого функционируют остальные программные модули, за исключением подсистемы пользовательского интерфейса.

Компонент «Загрузка» предназначен для начальной инициализации и управления запуском всех подсистем, в зависимости от текущего узла. Данный компонент получает управление сразу же после загрузки компонента «Хост». Конфигурационные данные предусматривают перечень сервисов для каждого узла, которые необходимо запустить. Обработывая эти данные, компонент «Загрузка» в указанной последовательности создает экземпляры соответствующих классов, а затем запускает их, позволяя установить межкомпонентные связи и начать работу. Таким образом, распределение сервисов по узлам может достаточно гибко настраиваться, что является весьма важным, как с эксплуатационной точки зрения, так и с точки зрения дальнейшего развития системы.

Предоставляемый компонентом «Загрузка» интерфейс `IBootstrapService` обеспечивает перезапуск всех сервисов в текущем узле, а также получение детальной информации о запущенных сервисах. Этот же интерфейс дает возможность выполнить «останов» или перезапуск всей операционной системы на текущем узле – благодаря чему оператор имеет возможность выключить весь комплекс напрямую, средствами только клиентского приложения АРМ.

Компонент «Конфигурация» предназначен для считывания и записи параметров конфигурации всех подсистем СППО. В качестве конфигурационного хранилища используется объектный каталог службы ADLDS. Компонент содержит отдельный модуль с определениями всех конфигурационных структур данных (конфигурируемых объектов), а также модуль реализации механизмов считывания и записи значений параметров этих структур в виде объектов LDAP-каталога. По умолчанию все операции производятся с копией каталога, хранящейся в текущем узле. Интерфейс `IConfigurationService`, предоставляемый компонентом

«Конфигурация», включает метод форсирования репликации каталога между узлами.

Компонент «Безопасность» содержит средства аутентификации и авторизации пользователей СППО. Компонент используется подсистемой пользовательского интерфейса для проверки введенных учетных данных пользователя и разграничения доступа к функциям подсистем в процессе работы.

Учетные данные хранятся в том же LDAP-каталоге, что и конфигурационные параметры. Для аутентификации пользователя компонент «пытается» установить подключение к службе ADLDS, используя введенные пользователем учетные данные. Если они не верны (несуществующее имя пользователя или неверный пароль), подключение установлено не будет. При успешной аутентификации подключение используется для определения роли пользователя: оператор, системный администратор или администратор баз данных. В зависимости от роли, пользователю предоставляется либо запрещается доступ к тем или иным функциям клиентского приложения АРМ.

Компонент «Конвейер» представляет собой набор вспомогательных классов, предназначенных для создания и управления совокупностью конвейеров видеобработки. Каждый конвейер – это последовательность блоков, выполняющих операции с видеокадрами, поступающими от устройств регистрации передающей части. В терминах DirectShow конвейер видеобработки является графом фильтров.

Компонент Конвейер является надстройкой над интерфейсом “DirectShow”, призванной облегчить взаимодействие «неуправляемого» кода (фильтров, реализованных на языке C++) и «управляемого» кода “.NET” (компонентов и WCF-сервисов, написанных на C#).

Компонент «Протокол» реализует "журналирование" всех событий, возникающих в процессе работы программного обеспечения СППО. В качестве хранилища записей протокола используется стандартная служба операционной системы Windows – WindowsEventLog. При первом запуске компонент создает отдельный файл журнала, в который записываются все события. Каждая запись

содержит развернутую информацию, включающую тип события (ошибка, предупреждение, сообщение), источник (имя компонента и подсистемы), сообщение с описанием. Размер журнала автоматически отслеживается операционной системой: при достижении максимального размера файла самые старые записи перекрываются новыми. Преднамеренно удалить записи из журнала возможно, только если полностью очистить журнал – но для этого требуются полномочия локального администратора.

Просмотр записей протокола следует осуществлять как встроенными средствами операционной системы, так и средствами клиентского приложения АРМ подсистемы пользовательского интерфейса. Для второго варианта компонент «Протокол» предоставляет интерфейс ILogService, который позволяет агрегировать сообщения со всех узлов системы.

Подсистема интеграции. Назначение и состав

Подсистема интеграции предназначена для сопряжения системы с внешними системами безопасности и контроля доступа. Сопряжение обеспечивается путем трансляции видеоданных, полученных от подсистемы регистрации, сопутствующих данных об обнаруженных объектах, а также информации о событиях распознавания и решениях, принимаемых оператором. Подсистема интеграции осуществляет трансляцию этих данных посредством разработанных протоколов.

Подсистема интеграции включает два компонента (см. рисунок 3.13): «Видеотрансляция» и «Трансляция событий».

Компонент «Трансляция событий» реализует ТСР-сервер, ожидающий подключения клиентов к заданному порту. Протокол трансляции событий не предусматривает каких-либо диалоговых средств взаимодействия, поэтому сразу после открытия сокета начинается трансляция текущих событий. Источниками событий для данного компонента являются:

- интерфейс IPersonIdentificationService подсистемы распознавания;
- компонент «Приложение оператора» подсистемы пользовательского интерфейса (через интерфейс IEventSourceService).

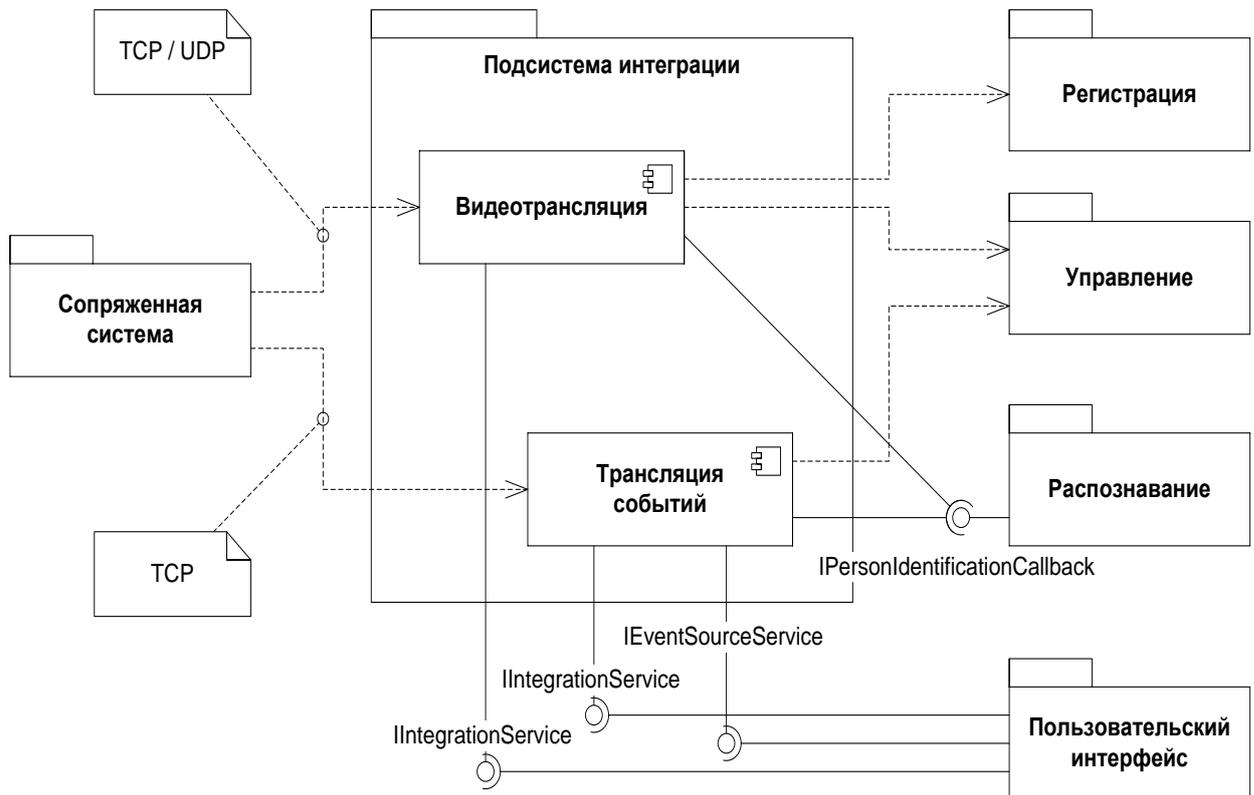


Рисунок 3.13 - Подсистема интеграции.

Подключение к этим двум источникам происходит при создании экземпляра компонента «Трансляция событий» и при запуске приложения оператора соответственно.

Протокол трансляции событий реализуется на базе стандартных средств объектной библиотеки “.NET Framework”. В частности, основные алгоритмы строятся при помощи классов `System.Net.Sockets.TcpListener` и `System.Net.Sockets.TcpClient`, обеспечивающих абстракцию серверного и клиентского сокетов.

Механизм работы компонента представляет собой асинхронную событийную модель. Подключаясь к интерфейсам-источникам, компонент «подписывается» на их события, регистрируя интерфейсы обратного вызова. При наступлении событий объекты-источники осуществляют вызовы методов этих интерфейсов, передавая управление компоненту «Трансляция событий».

Обработывая параметры методов, компонент формирует требуемые протоколом пакеты и пересылает их подключенным клиентам.

Конфигурационными параметрами для компонента «Трансляция событий» являются:

- IP-адрес сетевого интерфейса серверного сокета (по умолчанию – 0.0.0.0, то есть все доступные интерфейсы);

- номер порта серверного сокета (число в диапазоне от 1024 до 65535, по умолчанию 5010);

- максимальное число одновременно обслуживаемых клиентских подключений (неотрицательное целое число, по умолчанию 0 – не ограничено).

Компонент «Трансляция событий» предоставляет интерфейс `IntegrationService` для мониторинга и управления подключенными клиентами. Данный интерфейс используется подсистемой пользовательского интерфейса, давая оператору возможность просмотра текущих подключений и отключения любого из них при необходимости.

Вторая составляющая подсистемы интеграции – это компонент «Видеотрансляция». Данный компонент реализует RTSP-сервер, предоставляющий возможность получения потоковых видеоданных по запросу. Протокол RTSP (RFC 2326) регламентирует набор команд для выбора конкретного потока, начала/остановки трансляции, согласования параметров кодирования и т.д. Непосредственно передача видеоданных реализуется протоколами RTSP (RTPcontrolprotocol, RFC 3550) и модифицированным протоколом RTP, согласно описанию протокола видеотрансляции системы.

Для получения видеоданных от подсистемы регистрации компонент «Видеотрансляция» встраивается в конвейер видеообработки. Для этой цели в его состав входит несколько экземпляров блоков конвейера (DirectShow-фильтров), которые совместно используют один экземпляр RTSP-сервера (см. рисунок 3.14). Число блоков конвейера соответствует числу устройств регистрации передающей части АПК.

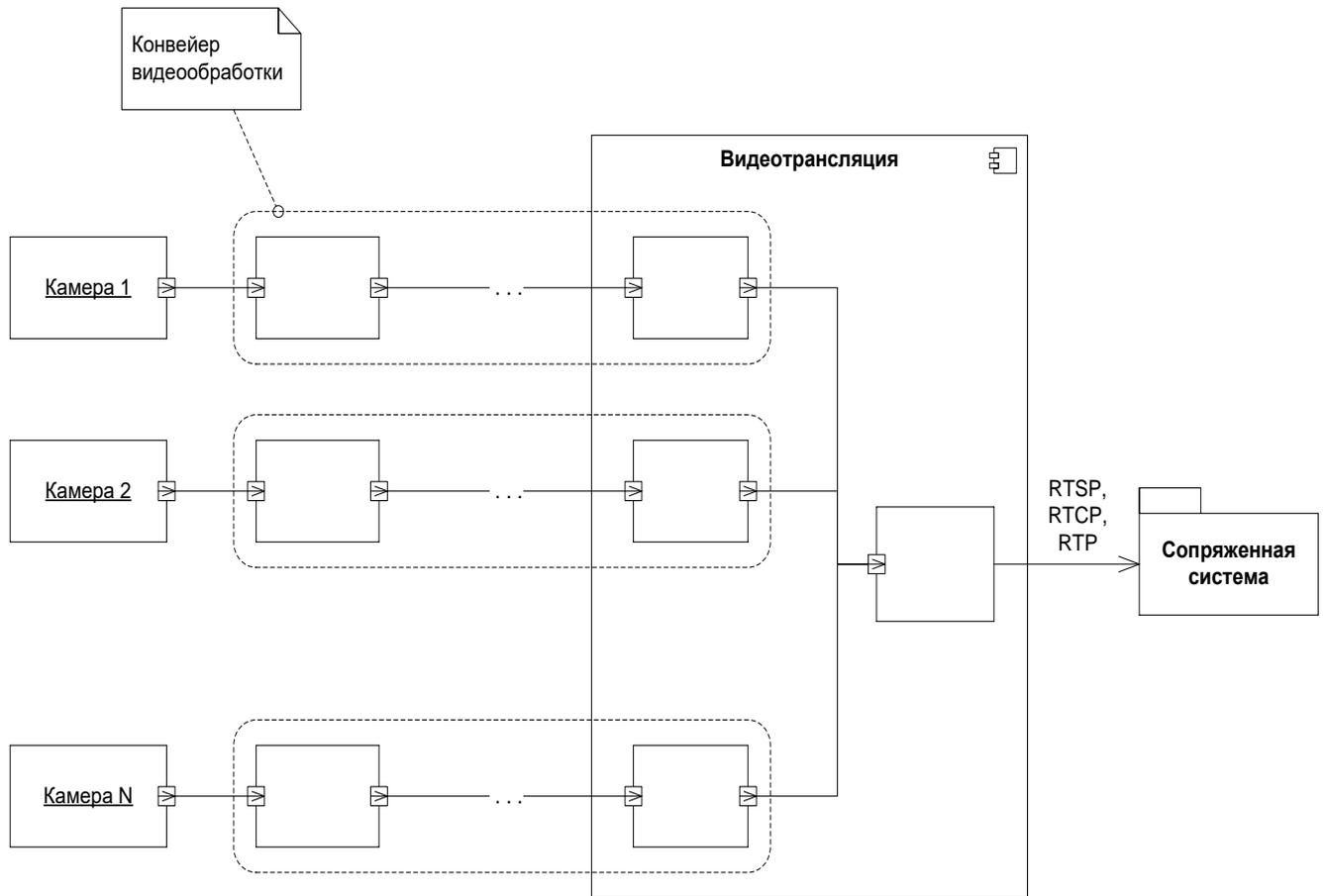


Рисунок 3.14 - Интеграция компонента «Видеотрансляция» в конвейер видеобработки.

Наряду с видеоданными, протокол трансляции обеспечивает передачу сопутствующих данных об обнаруженных и распознанных объектах. Эти данные включаются в зарезервированные поля заголовка RTP-пакета. Источниками данных об объектах являются:

- компонент «Обнаружение» подсистемы регистрации. Поскольку этот компонент также встраивается в конвейер видеобработки, данные от него передаются непосредственно, минуя интерфейс `IPersonDetectionService`;
- компонент «Идентификация» системы распознавания. посредством интерфейса `IPersonIdentificationService` компонент «Видеотрансляция» подписывается на события распознавания объекта и с помощью интерфейса обратного вызова `IPersonIdentificationCallback` получает информацию о них.

Таким образом, компонент «Видеотрансляция» осуществляет мультиплексирование трех потоков данных (см. рисунок 3.15). При этом опционально к видеопотоку может быть применено сжатие одним из стандартных системных кодеков.

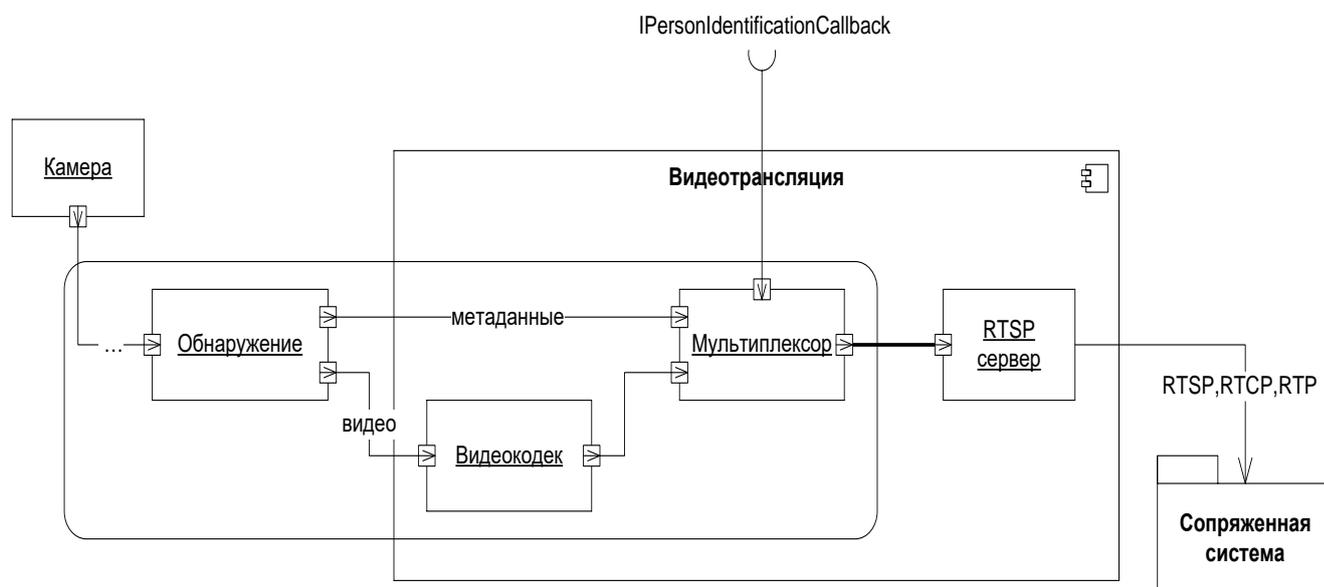


Рисунок 3.15 - Потоки данных компонента «Видеотрансляция».

Мультиплексор реализуется в виде «DirectShow» фильтра. RTSP-сервер реализуется на основе общедоступной библиотеки «LiveMedia» и конструктивно объединяется с мультиплексором в одном DLL-модуле. При этом экземпляр RTSP-сервера функционирует согласно шаблону «singleton». Согласование всех составляющих компонента «Видеотрансляция» и их интеграция с другими подсистемами реализуется средствами «.NET Framework».

Конфигурационные параметры компонента «Видеотрансляция»:

- IP-адрес RTSP-сервера (по умолчанию 0.0.0.0);
- порт RTSP-сервера (по умолчанию 554, стандартный порт протокола RTSP);
- тип транспортного протокола для RTP: TCP либо UDP (по умолчанию TCP, согласно описанию протокола видеотрансляции системы).

Для мониторинга и управления клиентскими подключениями компонент «Видеотрансляция», так же, как и компонент «Трансляция событий», предоставляет интерфейс IntegrationService.

Подсистема пользовательского интерфейса. Назначение и состав

Подсистема пользовательского интерфейса представляет собой средство взаимодействия обслуживающего персонала со всеми подсистемами СППО для использования комплекса по назначению (см. рисунок 3.16). Задачей подсистемы является наиболее полное отображение информации о функционировании всех модулей АПК – как программных, так и аппаратных, а также обеспечение удобства выполнения функций оператором.

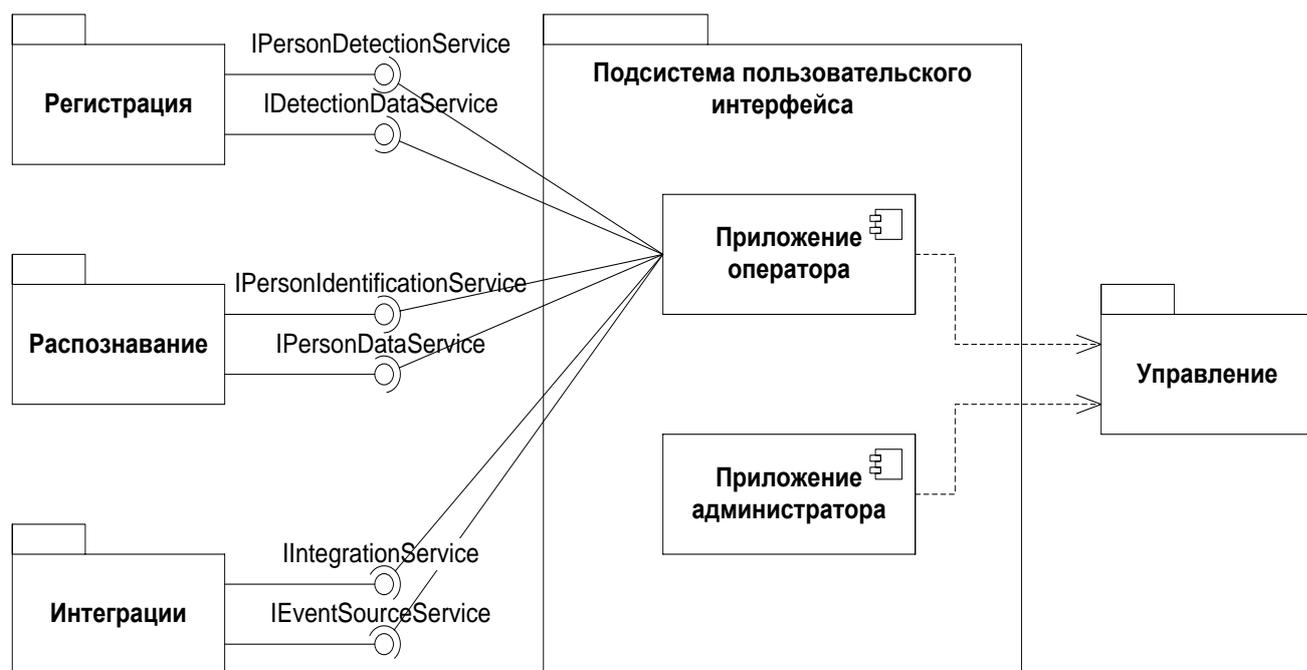


Рисунок 3.16 - Подсистема пользовательского интерфейса.

Два компонента подсистемы пользовательского интерфейса содержат в своем составе по одному исполняемому файлу. В целом, каждый компонент представляет собой стандартное оконное приложение WindowsForms. При запуске приложение отображает диалоговое окно для ввода учетных данных – имени и пароля пользователя. Интерфейс состоит из набора экранных форм,

скомпонованных из стандартных элементов управления: меню, кнопок, текстовых полей, списков, переключателей и т.д. Взаимодействие с экранными формами осуществляется посредством клавиатуры и манипулятора «мышь».

Запуск приложений не является обязательным для функционирования остальных подсистем СППО. Подсистема управления может быть сконфигурирована таким образом, чтобы начинать работу ключевых функций (обнаружение, распознавание, видеозапись) при включении АПК, не дожидаясь команды оператора. В этом случае результаты работы фиксируются в архивной базе данных, в протоколе, в файлах видеозаписи, а также транслируются сопряженным системами безопасности/контроля доступа.

Подсистема не предоставляет никаких внешних интерфейсов, поскольку выполняет исключительно клиентские функции.

Компонент «Приложение оператора» предназначен для выполнения оператором основных служебных обязанностей: ведение наблюдения, контроль функции распознавания, принятие решения о реакции на обнаружение/опознание объекта, просмотр и поиск информации в архивной и оперативной базах данных. Для отображения видеоданных «Приложение оператора» интегрируется в конвейер видеообработки. Начальные блоки конвейера (получение потока данных с камер и их декодирование) заимствуются из подсистемы регистрации, компонент «Съемка». Затем подключается блок визуализации, входящий в «Приложение оператора». Данный блок осуществляет прорисовку видеокадров на экране монитора.

Компонент «Приложение администратора» предназначен для начального конфигурирования и переконфигурирования подсистем СППО, а также для проведения обслуживающих работ с базами данных в составе СППО: пакетного импорта/экспорта данных, резервного копирования.

Подсистема не содержит интерфейсов и типов данных, используемых другими подсистемами.

3.3 Разработка методики управления средствами видеоконтроля процессов распознавания

Этап формирования признакового описания (методики)

Понятие методики определяется как последовательность выполняемых действий, которые характеризуются следующими наименованиями:

1. **ТРАНСФОРМАЦИЯ** - фото субъекта обрабатывается фильтром (комбинацией фильтров) общего назначения;
2. **ПРИЗНАКОВОЕ ОПИСАНИЕ** – формируется набор чисел, характеризующих фото (вектор, гистограмма, пакет);
3. **СРАВНЕНИЕ** – правило использования метрического способа сравнения числовых характеристик.

Процедура наполнения «Стенда» содержанием, то есть выбор методики и параметров, является весьма кропотливой и трудоемкой процедурой. Далее приводится каталог отобранных методик, состоящий из двух частей I и II. В части I содержатся данные, которые сформированы на теоретических этапах работы над стендом [A116, A117, A122-A127]. В части II изложены методики, построенные и апробированные на практике и в лабораторных условиях на последующих этапах разработки стенда. Все разработанные методики обеспечены оригинальной программной реализацией, защищенной авторскими правами [A118-A121].

Методика – дискретные косинусные преобразования (обозначение D)

Последовательность действий:

- трансформация исходного фото не проводится;
- применение разновидности Фурье –преобразования, а именно дискретное косинусное преобразование (сканирование-сжатие);
- применение синус–метрики.

Дискретное косинусное преобразование (ДКП) представляет собой разновидность преобразования Фурье [107, 183, 184,] и имеет обратное преобразование. В ДКП изображение рассматривается как совокупность

пространственных волн, для которых оси X и Y проводятся параллельно продольной и поперечной осям, а по оси Z откладывается значение яркости соответствующего пикселя изображения, при этом порядок нумерации элементов представлен на рисунке 3.17.

Для изображения $M \times N$ дискретное косинусное преобразование имеет следующий вид [185]:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I[x, y] \cos \left[\frac{(2x+1)u\pi}{2M} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right],$$

где $I[x, y]$ - яркость пикселя (x, y) ,

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{M}}, u = 0 \\ \sqrt{\frac{2}{M}}, u = 1, 2, \dots, M-1 \end{cases}, \quad \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}}, v = 0 \\ \sqrt{\frac{2}{N}}, v = 1, 2, \dots, N-1 \end{cases}.$$

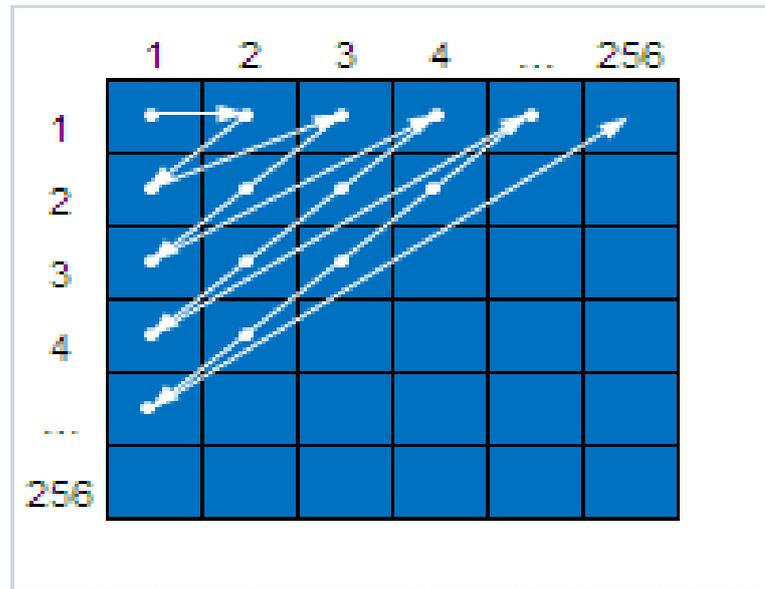


Рисунок 3.17 - Порядок нумерации элементов.

Величины $C(u, v)$ образуют матрицу коэффициентов преобразования ДКП. Преобразование матрицы в вектор с упорядоченными коэффициентами матрицы имеет вид:

$$\{C_{1,1}, C_{2,1}, C_{1,2}, C_{3,1}, \dots, C_{256,256}\}.$$

Процедура формирования вектора признаков в подсистеме нижнего уровня, использующей ДКП, схематично представлена на рисунке 3.18.

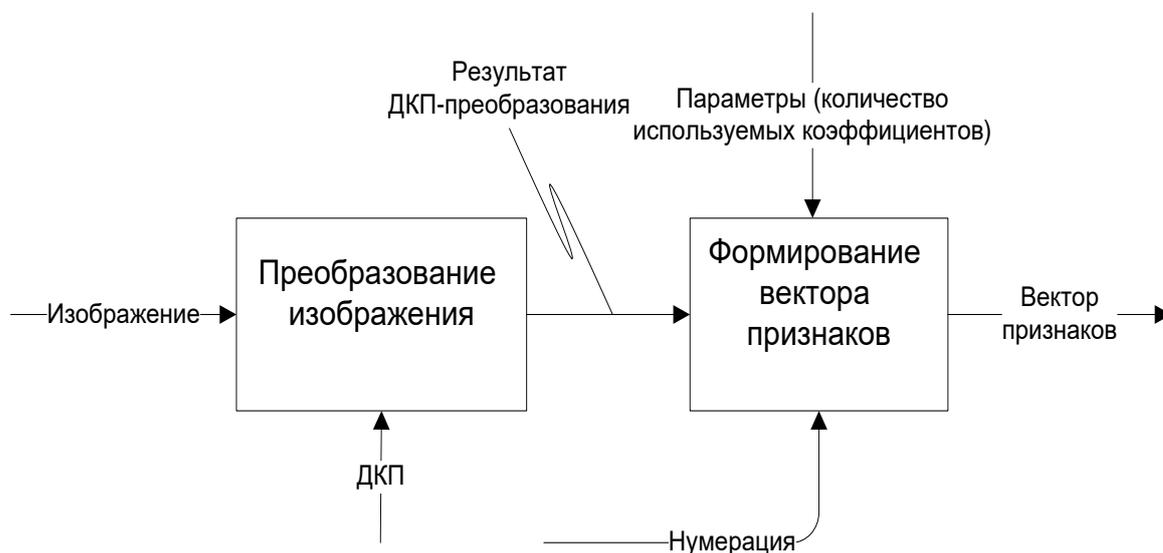


Рисунок 3.18 - Схема формирования вектора признаков при использовании ДКП.

Для определения параметров алгоритма и процедуры формирования вектора признаков используется не весь результат ДКП-преобразования изображения лица человека, а лишь его часть. Низкочастотные коэффициенты в наибольшей степени определяют основные черты изображения. Высокочастотные коэффициенты соответствуют мелким деталям изображения. Так как априори неизвестно, какое количество коэффициентов преобразования обеспечивает лучшие результаты распознавания, необходимо определить это опытным путем.

При рассмотрении процедуры построения вектора признаков необходимо обозначить элементы в матрице коэффициентов преобразования ДКП через i, j – индексы. Построение вектора признаков необходимо проводить из коэффициентов, индексы которых удовлетворяют условию $i + j \leq n$, где параметр n – натуральное число. Нумерацию коэффициентов необходимо выполнять по схеме, представленной на рисунке 3.19.

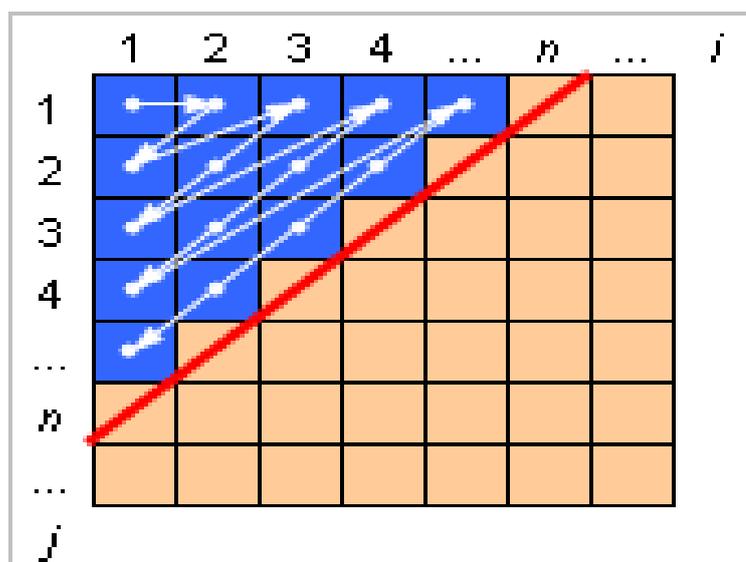


Рисунок 3.19 – Схема нумерации элементов, удовлетворяющих условию $i+j \leq n$

В результате формируется вектор меньшего размера, состоящий из значений, полученных при ДКП-преобразовании, со следующими номерами:

$$\{C_{1,1}, C_{2,1}, C_{1,2}, C_{3,1}, \dots, C_{1,n-1}\} .$$

В качестве метрики для вычисления расстояния между векторами признаков используется синус угла между векторами:

$$D(h^1, h^2) = \sin(\widehat{h^1, h^2}) = \sqrt{1 - \cos^2(\widehat{h^1, h^2})} = \sqrt{1 - \left[\frac{\langle h^1, h^2 \rangle}{\|h^1\| \cdot \|h^2\|} \right]^2} .$$

Методика - моменты Лежандра (обозначение LG)

Последовательность действий:

- трансформация исходного фото не проводится;
- вычисление моментов при помощи полиномов Лежандра;
- применение синус-метрики.

Отличительной особенностью данной методики является нечувствительность либо слабая чувствительность к аффинным преобразованиям и следующие преимущества:

- допускается различие изображения лица человека по масштабу (например, различная высота и ширина лица) и по условиям съемки;

- на фотографиях могут быть лица людей, имеющие поворот относительно центра. Например, в базе лиц «ColorFERET» имеются фотографии, где поворот лица составляет до 18 градусов (см. рисунок 3.20);

- лица, изображенные на фотографиях, могут быть неудовлетворительно отцентрированы.

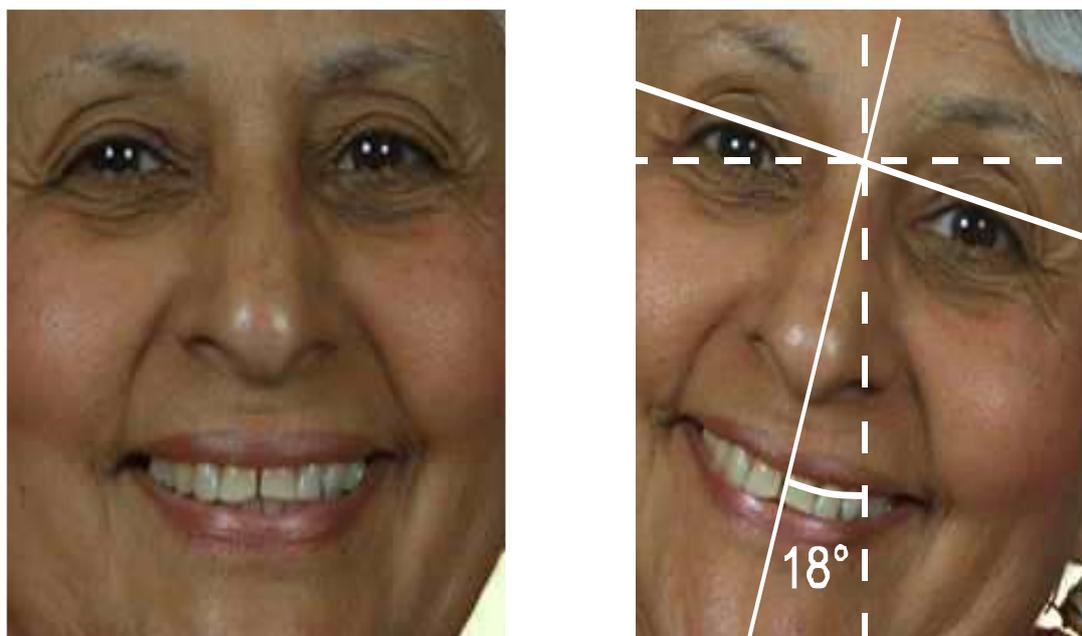


Рисунок 3.20 - Пример повернутого лица из базы ColorFERET.

Процедура формирования вектора признаков состоит в следующем.

Моменты обладают инвариантностью к аффинным преобразованиям и устойчивостью к шуму, а моменты первых порядков позволяют с достаточно высокой точностью реконструировать и распознавать изображения. В отличие от ДКП, в моментах используются не синусные, а полиномиальные базисные функции [150].

Для изображения $M \times N$ моменты Лежандра вычисляются следующим образом:

$$L(k, l) = \frac{(2k+1)(2l+1)}{4} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} P_k(x)P_l(y) \cdot f(x, y) ,$$

где $f(x, y)$ – яркость пикселя (x, y) , $P_0(x) = 1$; $P_1(x) = x$;

$P_k(x) = \frac{(2k-1)xP_{k-1}(x) - (k-1)P_{k-2}(x)}{k}$ – полином Лежандра степени k , для $k \in \mathbb{N}, k > 1$.

Величины $k \geq 1, l \geq 1$ определяют порядок моментов.

Рассчитав некоторое количество моментов Лежандра для ракурса, формируется следующий вектор:

$$\{L_{1,1}, L_{2,1}, L_{1,2}, L_{3,1}, \dots, L_{k,l}\}.$$

Данный вектор является вектором признаков, построенным при помощи моментов Лежандра. Моменты Лежандра вычисляются для некоторых натуральных чисел $k \geq 1, l \geq 1$. В силу симметричности принципа исчисления по паре (k, l) , вычисление моментов выполняется при условии $k + l \leq n$. Нумерация и выбор значений, из которых формируется вектор признаков, аналогичен процедуре, используемой в ДКП (см. рисунок 3.9). Величина n является параметром построения образа. Значение n , при котором достигается высокая корректность ПНУ, определяется экспериментально.

Методика – энергетические карты Лавса (обозначение LW)

Последовательность действий:

- трансформация исходного фото проводится при помощи карт Лавса;
- построение вектора признаков при помощи сканирования-сжатия выполняется с применением ДКП-преобразований, если исследуется карта. Если исследуется достаточно малое окно интереса, то сама карта яркостей служит таким вектором. В первом случае используется синус-метрика, во втором – метрика L1.

В методику вводятся пять векторов:

$$L5 (level - \text{уровень}) = [1 \ 4 \ 6 \ 4 \ 1],$$

$$E5 (edge - \text{край}) = [-1 \ -2 \ 0 \ 2 \ 1],$$

$$S5 (spot - \text{пятно}) = [-1 \ 0 \ 2 \ 0 \ -1],$$

$$R5 (ripple - \text{рябь}) = [1 \ -4 \ 6 \ -4 \ 1].$$

Вектора $L5, E5, S5, R5$ при использовании их в качестве маски, позволяют выделить такие характерные особенности изображения вдоль выбранной оси (x или

у) как рябь, пятнышко, некоторую границу, общий уровень яркости. Все вектора, за исключением вектора $L5$, нормированы, сумма их элементов равна 0.

При помощи комбинации этих векторов получаются маски, являющиеся масками Лавса [179]. К примеру, маска $E5L5$ получается умножением векторов $(E5)^t \cdot L5$. Индекс «t» означает транспонированный вектор. Подробнее:

$$\begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \times [1 \ 4 \ 6 \ 4 \ 1] = \begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}.$$

Таким же образом, получаются остальные 15 масок (всего их 16).

Далее вводится параметр N_l и определяется окно с размерами $N_l \times N_l$ с центром в точке $[r, c]$. Данное окно помещается в левый нижний угол изображения. Затем это окно интереса последовательно смещается вправо по изображению на шаг, равный ширине окна, до достижения правого края изображения. Далее происходит возврат к левому краю и окно смещается вверх на шаг, равный высоте окна. Затем – снова вправо и т.д. Процесс повторяется, пока окно последовательно не накроет все изображение.

На каждом шаге внутри окна вычисляется среднее значение яркости по окну и в каждом пикселе всего окна вычитается это среднее от значения яркости. В итоге формируется матрица (той же размерности, что и изображение), значение элементов которой близко к нулю. Возможны и отрицательные элементы.

Полученную таким образом усредненную матрицу необходимо использовать для формирования всех 16 карт Лавса. Для этого к точкам усредненной матрицы необходимо применить все 16 масок Лавса 5×5 . То есть каждая точка рассматриваем в окрестности 5×5 .

Далее необходимо: обозначить через $O(x, y)$ окрестность 5×5 точки (x, y) , а элементы этой окрестности $o^i(x, y) \in O(x, y)$; обозначить через w^i элементы некоторой маски 5×5 . Тогда результат применения некоторой маски к точке выглядит следующим образом: $\sum_i o^i(x, y) w^i$. На рисунке 3.21 приведен пример карт Лавса.

Если обозначить через $F_k[i, j]$ результат применения к точке $[i, j]$ усредненной матрицы некоторой k -ой маски, то энергетическая карта пикселя $[r, c]$ определяется выражением

$$E_k[r, c] = \sum_{i=r-\frac{N_L}{2}}^{r+\frac{N_L}{2}} \sum_{j=c-\frac{N_L}{2}}^{c+\frac{N_L}{2}} |F_k[i, j]|.$$

Далее, берется среднее от некоторых комбинаций масок. Например, $\frac{L5E5}{E5L5}$ есть среднее арифметическое от применения двух масок $L5E5$ и $E5L5$. Таким образом из 16 промежуточных карт Лавса формируются девять окончательных (см. рисунок 3.11):

$$\frac{L5E5}{E5L5}, \frac{L5S5}{S5L5}, \frac{L5R5}{R5L5}, E5E5, \frac{E5S5}{S5E5}, \frac{E5R5}{R5E5}, S5S5, \frac{S5R5}{R5S5}, R5R5.$$

Если выбрать одну из карт Лавса, то она будет являться вектором признаков, построенным при помощи энергетических карт Лавса. Энергетическая карта Лавса имеет такое же количество точек, как и исходное изображение. Следовательно, для исходного изображения $M * N$ точек вектор признаков будет состоять из $M * N$ чисел.

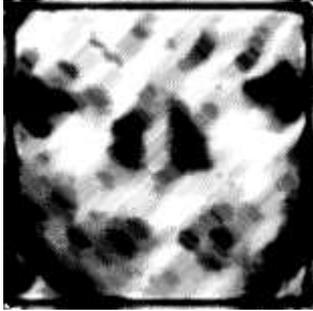
Возможны два варианта сравнения фото.

Вариант 1. Строится карта Лавса по всему изображению. В этом случае используется стандартный способ построения вектора признаков. Проводится сканирование–сжатие, которое осуществляется при помощи ДКП-преобразования. В данном случае используется синус–метрика.

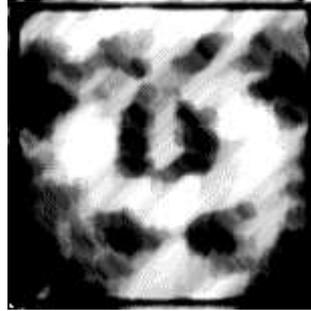
Вариант 2. Фото разбивается на фрагменты. В этом случае вектором признаков каждого фрагмента является само изображение фрагмента (после трансформации). Фрагменты сравниваются в метрике $L1$. На основании всех сравнений строится консолидированное расстояние между исходными фото.



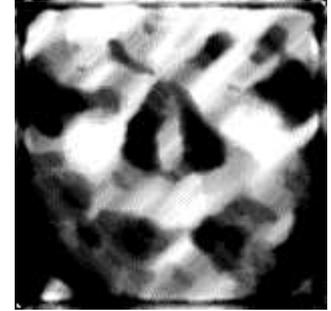
а)



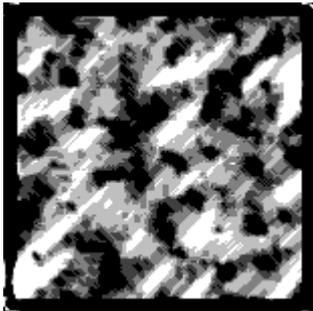
б)



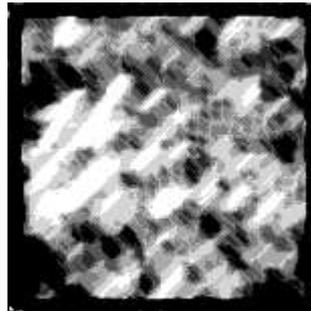
в)



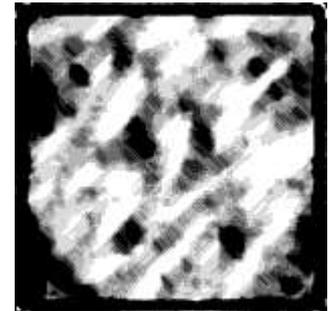
г)



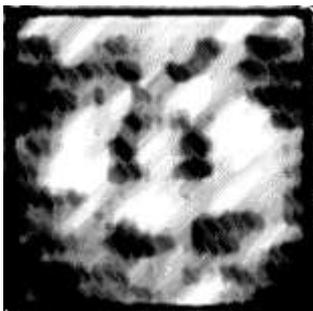
д)



е)



ж)



з)



и)



к)

Рисунок 3.21 - Исходное изображение (а) и построенные карты Лавса: E5E5 (б), E5R5/R5E5 (в), E5S5/S5E5 (г), L5E5/E5L5 (д), L5R5/R5L5 (е), L5S5/S5L5 (ж), R5R5 (з), S5R5/R5S5 (и), S5S5 (к).

Методика - порядковые гистограммы (обозначение FC)

Последовательность действий:

- использование базиса Фрея–Чена для построения откликов в 3×3 окрестности пикселя;
- формирование содержания порядковых признаков по всем используемым пикселям, строится порядковая гистограмма;
- применение метрики L1.

Вводятся девять матриц размерами 3×3 следующего вида:

Градиент	$W_1 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix};$	$W_2 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix};$
Пульсации	$W_3 = \frac{1}{\sqrt{8}} \begin{bmatrix} 0 & -1 & \sqrt{2} \\ 1 & 0 & -1 \\ -\sqrt{2} & 1 & 0 \end{bmatrix};$	$W_4 = \frac{1}{\sqrt{8}} \begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -\sqrt{2} \end{bmatrix};$
Линия	$W_5 = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix};$	$W_6 = \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix};$
Лапласиан	$W_7 = \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix};$	$W_8 = \frac{1}{6} \begin{bmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{bmatrix};$
Константа	$W_9 = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$	

Указанные девять матриц образуют ортонормированный базис в пространстве матриц 3×3 и являются базисом Фрея-Чена [179]. Матрицы базиса Фрея-Чена используются в качестве масок для изображения, то есть каждый пиксель рассматривается в окрестности 3×3 .

Если обозначить: через $O(x, y)$ окрестность 3×3 пикселя (x, y) , элементы этой окрестности - $o^i(x, y) \in O(x, y)$. Далее обозначить через w^i элементы некоторой маски 3×3 , то результат применения некоторой маски к пикселю выглядит следующим образом: $\sum_i o^i(x, y) w^i$.

Реакция окрестности пикселя (x, y) на некоторую маску $W_i, i = 1, \dots, 9$ есть фиксация наличия примитива (градиент, пульсация, линия...) в этой окрестности.

Величина $\sum_i o^i(x, y) w^i$ характеризует выраженность соответствующего примитива в окрестности пикселя (x, y) . Чем она выше, тем ярче выраженность того или иного примитива.

Если ввести понятие и обозначение энергии градиента как величину

$$E_{\text{градиента}} = W_1^2 + W_2^2,$$

и аналогично:

$$E_{\text{пульсации}} = W_3^2 + W_4^2,$$

$$E_{\text{линии}} = W_5^2 + W_6^2,$$

$$E_{\text{лапласиана}} = W_7^2 + W_8^2,$$

$$E_{\text{константы}} = W_9^2,$$

то полная энергия $E_{\text{общ}}$:

$$E_{\text{общ}} = E_{\text{градиента}} + E_{\text{пульсации}} + E_{\text{линии}} + E_{\text{лапласиана}} + E_{\text{константы}}.$$

В дальнейшем для построения гистограммы будет использоваться только

$$E_{\text{градиента}}, E_{\text{пульсации}}, E_{\text{линии}} \text{ и } E_{\text{лапласиана}}.$$

Величина $E_{\text{константы}}$ не используется, так как данная величина соответствует значению средней яркости в окрестности точки и не содержит текстурной информации.

Если обозначить градиент, пульсацию, линии и лапласиан через a , b , c , d , соответственно, то:

$$a = \frac{E_{\text{градиента}}}{E_{\text{общ}}}; b = \frac{E_{\text{пульсации}}}{E_{\text{общ}}}; c = \frac{E_{\text{линии}}}{E_{\text{общ}}}; d = \frac{E_{\text{лапласиана}}}{E_{\text{общ}}}.$$

В каждом используемом пикселе, определяются порядок. Например, порядок $abcd$ означает, что $a \leq b \leq c \leq d$; в случае равенства приоритет отдается букве, стоящей ранее по алфавиту. То есть, если $a < b = c < d$, то порядок $abcd$, а не $acbd$ и т.д. Порядок $bcad$ означает, что $b \leq c \leq a \leq d$. В силу того, что используются порядковые характеристики и отношение рассматриваемых величин не зависит от яркости текущего пикселя, получаются инвариантные к яркости отношения. То есть точки равноправны и в темной, и в светлой областях.

Если построить гистограмму, состоящую из $4!=24$ счетчиков, в зависимости от порядка, то в каждом пикселе определяется конкретный порядок, который там получился, и единица добавляется к соответствующему счетчику.

Пусть $H = \{h_i\}$; $H' = \{h'_i\}$ – две гистограммы, тогда расстояние

$$D(H, H') = \frac{1}{K} \sum_i |h_k - h'_k|^L,$$

где K – количество счетчиков гистограммы. L – показатель степени. Значение L , дающее лучшие результаты корректности ПНУ, определяется экспериментально.

Методика – маски Собеля (обозначение S)

Последовательность действий:

- трансформация исходного изображения при помощи масок Собеля;
- формирование вектора признаков при помощи ДКП-преобразования (сканирования-сжатия);
- применение синус-метрики.

Алгоритм вычисления вектора признаков использует процедуру каскадного применения алгоритмов обработки изображения и состоит из нескольких этапов. Схема формирования вектора признаков представлена на рисунке 3.22.

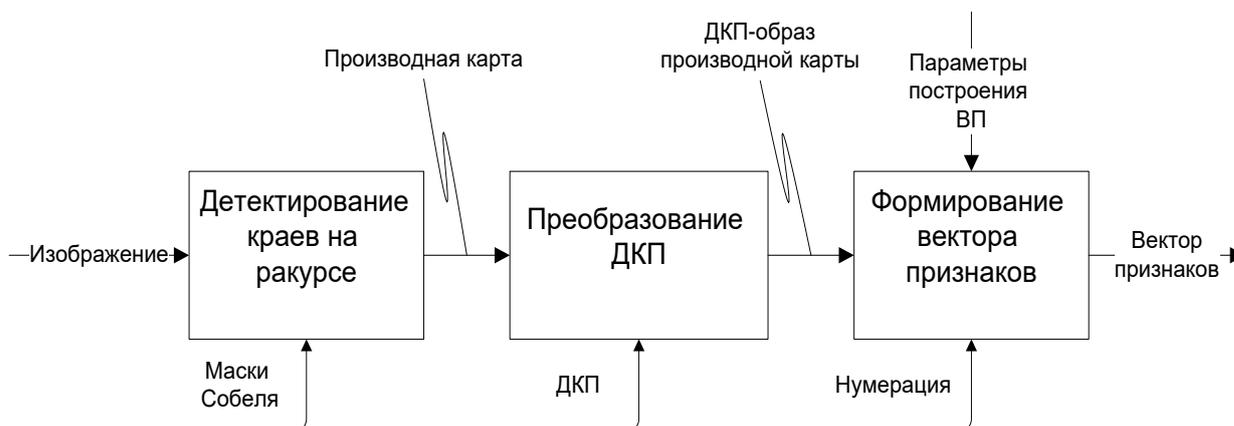


Рисунок 3.22 - Схема формирования вектора признаков по методике маски Собеля.

Алгоритм формирования вектора признаков использует маски Собеля для детектирования краев на фотографии. Исходное изображение преобразуется при

помощи этих масок в производную карту. Применяются маски Собеля следующего вида:

$$G_X = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}; \quad G_Y = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}.$$

Детектирование краев осуществляется с применением обеих масок. При этом маска G_X осуществляет поиск краев по горизонтали, а маска G_Y – по вертикали. Каждый пиксель фотографии рассматривается в окрестности 3×3 .

Если обозначить через $O(x, y)$ окрестность 3×3 пикселя (x, y) , элементы этой окрестности $o^i(x, y) \in O(x, y)$ и через w^i – элементы некоторой маски 3×3 , то результат применения этой маски к пикселю выглядит следующим образом:

$$\sum_i o^i(x, y) w^i.$$

Для каждого пикселя исходной фотографии применяются обе маски при вычислении следующих величин:

$$G_X(x, y) = \sum_i o^i(x, y) g_X^i,$$

$$G_Y(x, y) = \sum_i o^i(x, y) g_Y^i,$$

где $g_X \in G_X$, $g_Y \in G_Y$. Для учета детектированных краев, как по вертикали, так и по горизонтали рассчитывается следующая величина:

$$I(x, y) = \sqrt{G_X(x, y)^2 + G_Y(x, y)^2}.$$

Данное число $I(x, y)$ будет являться значением яркости для пикселя (x, y) в формируемой производной карте.

Далее к полученному изображению применяется ДКП и осуществляется выбор коэффициентов, которые и составят вектор признаков. В качестве расстояния используется синус-метрика.

Методика - пакетный принцип сравнения

В данном способе сравнения при построении набора гауссовых фильтров предполагается подбор двух групп параметров. Одна группа – это параметры самих гауссовых фильтров; другая группа параметров – это весовые коэффициенты, с которыми данные параметры включаются в линейную комбинацию.

В отличие от известного способа [6, 7, 14, 15, 21, 25] предлагается исключить настройку по второй группе параметров при рассмотрении всех возможных линейных комбинаций фото, к которым применяются гауссовы фильтры из заранее фиксированного набора. В данном случае используются для сравнения не сами фото, а результаты их сканирования-сжатия, а именно, векторы, состоящие из части коэффициентов ДКП-преобразования. Таким образом возникает объект, который является специализированным пакетом. Подобный пакет представляет собой линейное пространство, образованное набором векторов незначительной размерности. Например, если в качестве способа сканирования используется ДКП-преобразование на картах размером 256x256, то необходимо ограничиться примерно двумястами коэффициентами и т.д. При этом возникают большие возможности построения необходимого пакета, так как появляется возможность комбинировать последовательность применения преобразующих фильтров, применять различные приемы сканирования-считывания и т.д. Важным является тот факт, что на линейном пространстве вектор признаков заменяется пакетом. В данном случае необходимо более подробно определить математическое описание процедуры построения и сравнения пакетов, отображающих фото.

Краткое описание последовательности математических действий

Под понятием «пакет» подразумевается простое линейное пространство, образованное системой векторов. Для сравнения «пакетов признаков» двух классов предлагается следующая последовательность специализированных действий.

1) ОРТОГОНАЛИЗАЦИЯ

Пусть

a_1, \dots, a_k - векторы признаков;

π - пространство, образованное этими векторами, тогда

- ШАГ 1(первый):

Вход: $\{a_1, \dots, a_k\}$.

$$a'_1 = a_1; e_1 = \frac{a'_1}{\|a'_1\|}.$$

Выход: e_1, \dots, a_K .

- ШАГ k:

Вход: $e_1, \dots, e_{k-1}, \dots, a_K$.

$$a'_k = a_k - (\langle e_1, a_k \rangle e_1 + \dots + \langle e_{k-1}, a_k \rangle e_{k-1}); e_k = \frac{a'_k}{\|a'_k\|}$$

Выход: $e_1, \dots, e_k, \dots, a_K$.

- ШАГ K (последний):

Вход: $e_1, \dots, e_{K-1}, \dots, a_K$.

$$a'_K = a_K - (\langle e_1, a_K \rangle e_1 + \dots + \langle e_{K-1}, a_K \rangle e_{K-1}); e_K = \frac{a'_K}{\|a'_K\|}$$

Выход: e_1, \dots, e_K .

e_1, \dots, e_K - ортонормированный базис пространства π .

2) ПРИМЕНЕНИЕ

Для тестирования поступает вектор b

$$b' = b - (\langle e_1, b \rangle e_1 + \dots + \langle e_K, b \rangle e_K); \sin(\widehat{b, \pi}) = \frac{\|b'\|}{\|b\|}.$$

ЗАМЕЧАНИЕ. На этапе тестирования могут использоваться индикаторы δ_k . Они принимают значения 0 или 1; $b' = b - (\delta_1 \cdot \langle e_1, b \rangle e_1 + \dots + \delta_K \cdot \langle e_K, b \rangle e_K)$.

Так, например, комбинация $\delta_1 = 1, \delta_2 = \dots = \delta_K = 0$ означает обычный синус между направлениями.

Сравнение пакетов (метрика $d_{m,n}$):

1) ОПРЕДЕЛИТЕЛЬ ГРАМА

Пусть a_1, \dots, a_K - набор векторов в пространстве большой размерности. Тогда определитель Грама задается формулой:

$\Gamma(a_1, \dots, a_K) = \det(\langle a_i, a_j \rangle)$ - определитель матрицы, составленной из скалярных произведений.

Определитель равен квадрату K -мерного объема параллелепипеда, построенного на векторах a_1, \dots, a_K . Определитель $\Gamma \geq 0$; $\Gamma = 0$ тогда и только тогда, когда векторы a_1, \dots, a_K линейно зависимы.

При рассмотрении двух пространств, построенных на наборах векторов a_1, \dots, a_m и a_{m+1}, \dots, a_{m+n} , справедливо неравенство

$$\Gamma(a_1, \dots, a_{m+n}) \leq \Gamma(a_1, \dots, a_m) \cdot \Gamma(a_{m+1}, \dots, a_{m+n}).$$

Равенство возможно в следующих случаях:

- во-первых, когда пространства $L(a_1, \dots, a_m)$ и $L(a_{m+1}, \dots, a_{m+n})$ ортогональны;

- во-вторых, когда один из определителей равен нулю.

В приложениях данный вариант не реализуется, поэтому в качестве расстояния между пространствами $L(a_1, \dots, a_m), L(a_{m+1}, \dots, a_{m+n})$ следует взять

$$d_{m,n} = \frac{\Gamma(a_1, \dots, a_{m+n})}{\Gamma(a_1, \dots, a_m) \cdot \Gamma(a_{m+1}, \dots, a_{m+n})}.$$

Конкретные вычислительные формулы, которые используются в методиках «Стенд», имеют следующий вид.

Случай 1x1.

Сравниваются одномерные пакеты. Система состоит из 2-х векторов a_1, a_2 .

Тогда

$$\begin{aligned} \Gamma(a_1) &= \langle a_1, a_1 \rangle, \Gamma(a_2) = \langle a_2, a_2 \rangle; \Gamma(a_1, a_2) = \begin{vmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle \end{vmatrix} = \\ &= \langle a_1, a_1 \rangle \langle a_2, a_2 \rangle - \langle a_1, a_2 \rangle^2. \end{aligned}$$

Расстояние между векторами равно

$$d_{1,1} = \frac{\langle a_1, a_1 \rangle \langle a_2, a_2 \rangle - \langle a_1, a_2 \rangle^2}{\langle a_1, a_1 \rangle \langle a_2, a_2 \rangle} = 1 - \frac{\langle a_1, a_2 \rangle^2}{\langle a_1, a_1 \rangle \langle a_2, a_2 \rangle} = \sin^2(\widehat{a_1, a_2}).$$

Случай 2x1.

Сравниваются 2-мерный и 1-мерный пакеты.

Пусть e_1, e_2 ортонормированный базис 2-мерного пространства; вектор e_3 просто вектор длины 1. Необходимо определить синус угла между плоскостью и одномерным пространством. Пусть $e'_3 = e_3 - \langle e_1, e_3 \rangle \cdot e_1 - \langle e_2, e_3 \rangle \cdot e_2$.

Тогда квадрат синуса равен квадрату длины e'_3

$$\langle e'_3, e'_3 \rangle = 1 - \langle e_1, e_3 \rangle^2 - \langle e_2, e_3 \rangle^2.$$

Вычисляется скалярное произведение вектора на себя с учетом условий, наложенных на векторы. Тогда определитель Грама равен:

$$\Gamma(e_1, e_2, e_3) = \begin{vmatrix} 1 & 0 & \langle e_1, e_3 \rangle \\ 0 & 1 & \langle e_2, e_3 \rangle \\ \langle e_3, e_1 \rangle & \langle e_3, e_2 \rangle & 1 \end{vmatrix} = 1 - \langle e_1, e_3 \rangle^2 - \langle e_2, e_3 \rangle^2;$$

$$\Gamma(e_1, e_2) = 1, \Gamma(e_3) = 1.$$

В результате $d_{1,2} = \Gamma(e_1, e_2, e_3)$. Тогда квадрат синуса угла между одномерным направлением и плоскостью и есть расстояние, которое вычисляется по формуле, приведенной выше.

Случай 3x3.

Сравниваются 3-мерные пакеты.

Пусть $e_1, e_2, e_3; e_1', e_2', e_3'$ ортонормированные базисы, представляющие пакеты. Матрица Грама строится и выглядит следующим образом.

Вводятся сокращенные обозначения:

$$\begin{aligned} \langle e_1, e_1' \rangle &= \alpha_1; \langle e_1, e_2' \rangle = \alpha_2; \langle e_1, e_3' \rangle = \alpha_3, \\ \langle e_2, e_1' \rangle &= \beta_1; \langle e_2, e_2' \rangle = \beta_2; \langle e_2, e_3' \rangle = \beta_3, \\ \langle e_3, e_1' \rangle &= \gamma_1; \langle e_3, e_2' \rangle = \gamma_2; \langle e_3, e_3' \rangle = \gamma_3. \end{aligned}$$

Матрица Грама принимает вид:

$$\begin{pmatrix} 1 & 0 & 0 & \alpha_1 & \alpha_2 & \alpha_3 \\ 0 & 1 & 0 & \beta_1 & \beta_2 & \beta_3 \\ 0 & 0 & 1 & \gamma_1 & \gamma_2 & \gamma_3 \\ \alpha_1 & \beta_1 & \gamma_1 & 1 & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 & 0 & 1 & 0 \\ \alpha_3 & \beta_3 & \gamma_3 & 0 & 0 & 1 \end{pmatrix}.$$

Выполняются следующие операции со строками:

$$\begin{aligned} (4) - (\alpha_1(1) + \beta_1(2) + \gamma_1(3)) &\rightarrow (4), \\ (5) - (\alpha_2(1) + \beta_2(2) + \gamma_2(3)) &\rightarrow (5), \\ (6) - (\alpha_3(1) + \beta_3(2) + \gamma_3(3)) &\rightarrow (6). \end{aligned}$$

Это означает, например, что из строки 4 вычитается линейная комбинация строк 1,2,3, умноженных на соответствующие числа. Результат заносится снова в 4-ю строку.

В итоге получается матрица вида:

$$\begin{pmatrix} 1 & 0 & 0 & \alpha_1 & \alpha_2 & \alpha_3 \\ 0 & 1 & 0 & \beta_1 & \beta_2 & \beta_3 \\ 0 & 0 & 1 & \gamma_1 & \gamma_2 & \gamma_3 \\ 0 & 0 & 0 & h_{11} & h_{12} & h_{13} \\ 0 & 0 & 0 & h_{21} & h_{22} & h_{23} \\ 0 & 0 & 0 & h_{31} & h_{32} & h_{33} \end{pmatrix}.$$

Определитель матрицы 6x6 совпадает с определителем правого нижнего угла матрицы 3x3. Элементы этой матрицы вычисляются при помощи формул:

$$h_{ij} = -(\alpha_i \cdot \alpha_j + \beta_i \cdot \beta_j + \gamma_i \cdot \gamma_j); i \neq j,$$

$$h_{ii} = 1 - (\alpha_i^2 + \beta_i^2 + \gamma_i^2),$$

$$i, j = 1, 2, 3.$$

Матрица 3x3 получается симметричной $h_{ij} = h_{ji}$, что приводит к некоторым упрощениям. Итак, определитель равен:

$$d_{3,3} \equiv \Gamma = (h_{11}h_{22}h_{33} + 2h_{12}h_{13}h_{23}) - (h_{11}h_{23}^2 + h_{22}h_{13}^2 + h_{33}h_{12}^2).$$

Случай 2x2.

Итоговые формулы имеют вид:

$$\begin{pmatrix} 1 & 0 & \alpha_1 & \alpha_2 \\ 0 & 1 & \beta_1 & \beta_2 \\ \alpha_1 & \beta_1 & 1 & 0 \\ \alpha_2 & \beta_2 & 0 & 1 \end{pmatrix}.$$

После аналогичных преобразований формула приводится к виду:

$$\begin{pmatrix} 1 & 0 & \alpha_1 & \alpha_2 \\ 0 & 1 & \beta_1 & \beta_2 \\ 0 & 0 & h_{11} & h_{12} \\ 0 & 0 & h_{21} & h_{22} \end{pmatrix},$$

$$h_{ij} = -(\alpha_i \cdot \alpha_j + \beta_i \cdot \beta_j); i \neq j,$$

$$h_{ii} = 1 - (\alpha_i^2 + \beta_i^2),$$

$$i, j = 1, 2,$$

$$d_{2,2} \equiv \Gamma = h_{11}h_{22} - h_{12}^2.$$

Выше выделены лишь те частные случаи, которые были использованы при выборе методик, использующих пакетный способ сравнения.

Лабораторные исследования

Методики, использующие пакетный способ сравнения фото, проверены в лабораторных условиях при проектировании «Стенда», входящего в состав разработанной системы. Основная цель исследований заключалась в оценке перспективы применения подобных методик в рамках «Стенда». Результаты, полученные в процессе лабораторных испытаний исследований, привели к выводу о целесообразности их применения. Исследования носили многоплановый характер. Ниже приведены некоторые наиболее значимые фрагменты испытаний.

Первый фрагмент (ФИЛЬТРЫ) относится к решению задачи формирования пакетов с использованием набора различных гауссовых фильтров.

Второй фрагмент (ОКНА ИНТЕРЕСА) включает результаты исследования задачи, связанной с локализацией изображения путем выделением некоторых специальных окон.

Часто встречающаяся на практике задача распознавания изображения на основе использования только одного хранимого эталона или только одной предъявляемой фотографии, относящаяся к «проблеме одного эталонного изображения» OSPP или SSPP [150], удовлетворительного решения с помощью указанных выше типов алгоритмов не имеет, поскольку точное построение элементов структуры лица по одной фотографии невозможно. Практически исключается использование алгоритмов, относящихся к первому классу. С применением алгоритмов второго класса также возникают затруднения. Так, в случае применения нейросетевых методов исключается возможность обучения, а метод главных компонент хоть формально и может применяться, но дает слишком низкие результаты [175]. Таким образом, следует или искать новые подходы, или совершенствовать уже известные методы, например, метрические. Однако проблемой, которая снижает эффективность применения метрического подхода, является слишком большой внутриклассовый разброс по отношению к межклассовому разбросу [147].

Предлагаемый метод исследования

Карта – есть результат установления соответствия каждому пикселю некоторого числа. Если рассматривать карту как вектор, то данный вектор характеризуется значительной размерностью, например, до 256×256 , и любому изображению может быть сопоставлена его карта. Карты позиционируются как исходные или производные.

Исходная карта — это карта яркостей. Производная карта — это любая трансформация исходной карты, например, экваризация (выравнивание гистограммы), или экваризация с последующим гауссовым размытием и т.д.

Существенным необходимо признать тот факт, что при подобных трансформациях размерность карты, рассматриваемой как вектор, не меняется, то есть остается неприемлемо большой для решения задачи распознавания. Для уменьшения размерности применяются инструменты сканирования-сжатия, например, такие, как коэффициенты ДКП1 разложения и коэффициенты разложения по полиномам Лежандра. При этом возможно применение только части всех коэффициентов.

Результат сканирования является атрибутом. Атрибут представляет собой вектор существенно меньшей размерности, например, с размерностью порядка 200. Сравнение изображений осуществляется через сравнение их атрибутов.

Пусть M — это карта, например, карта яркостей или производная от нее карта. Тогда изложенные выше преобразования имеют следующий вид:

$$M \rightarrow G_i M \rightarrow t(G_i M) \rightarrow a_i \equiv a(t(G_i M)), i = 1, \dots, m.$$

Здесь через $G_i M$ обозначены размытия. Они получены из карты M при помощи набора Гауссовых фильтров, при этом обусловлено, что $G_i M \equiv M$.

Через $t(G_i M)$ обозначены любые преобразования карты, не изменяющие ее размерности, например, свертка фильтрами Собеля или преобразование с использованием текстурных карт Лавса.

Здесь необходимо отметить, что промежуточное преобразование $M \rightarrow t(M)$ может быть тождественным. Векторы a_i , полученные в результате сканирования, образуют в пространстве атрибутов линейное подпространство $L = L(a_1, \dots, a_m)$.

Если обозначить через e_1, \dots, e_m ортонормированный базис этого подпространства, то его можно получить, например, из a_1, \dots, a_m методом ортогонализации Грама-Шмидта [58]. Особо необходимо выделить одномерный случай $M \rightarrow t(M) \rightarrow a_i \equiv a(t(M))$, когда вектор a выделяет в пространстве атрибутов направление.

Для сравнения линейных подпространств атрибутов вводится метрика $d_{m,n}$. Тогда определитель Грама $\Gamma(a_1, \dots, a_k) = \det(\langle a_i, a_j \rangle)$ является определителем матрицы, составленной из скалярных произведений. Определитель равен квадрату K -мерного объема параллелепипеда, построенного на векторах a_1, \dots, a_k . При этом $\Gamma \geq 0$; $\Gamma = 0$ тогда и только тогда, когда векторы a_1, \dots, a_k линейно зависимы. Если имеются два пространства, построенных на наборах векторов a_1, \dots, a_m и a_{m+1}, \dots, a_{m+n} , то справедливо неравенство

$$\Gamma(a_1, \dots, a_{m+n}) \leq \Gamma(a_1, \dots, a_m) \cdot \Gamma(a_{m+1}, \dots, a_{m+n}).$$

Равенство возможно, когда:

- 1) пространства $\Gamma(a_1, \dots, a_m)$ и $\Gamma(a_{m+1}, \dots, a_{m+n})$ ортогональны;
- 2) один из определителей равен нулю [58].

В приложениях данный случай маловероятен, поэтому в качестве расстояния между пространствами $\Gamma(a_1, \dots, a_m), \Gamma(a_{m+1}, \dots, a_{m+n})$ необходимо взять

$$d_{m,n} = \frac{\Gamma(a_1, \dots, a_{m+n})}{\Gamma(a_1, \dots, a_m) \cdot \Gamma(a_{m+1}, \dots, a_{m+n})}. \quad (1)$$

В дальнейшем используются только метрики $d_{1,1}$ и $d_{3,3}$.

Для метрики $d_{1,1}$ - пусть a, a' два направления, тогда расстояние между ними в вышеуказанной метрике задается формулой

$$d_{1,1}(a, a') = \sin(a, a') = \sqrt{1 - \frac{\langle a, a' \rangle^2}{|a|^2 \cdot |a'|^2}}. \quad (2)$$

Для метрики $d_{3,3}$ - пусть два трехмерных под-пространства L и L' пространства атрибутов представлены своими ортонормированными базисами e_1, e_2, e_3 и e'_1, e'_2, e'_3 , соответственно (см. выше).

Введены обозначения:

$$\begin{aligned} \langle e_1, e'_1 \rangle &= \alpha_1, \langle e_2, e'_2 \rangle = \alpha_2, \langle e_3, e'_3 \rangle = \alpha_3, \\ \langle e_2, e'_1 \rangle &= \beta_1, \langle e_2, e'_2 \rangle = \beta_2, \langle e_2, e'_3 \rangle = \beta_3, \\ \langle e_3, e'_1 \rangle &= \gamma_1, \langle e_3, e'_2 \rangle = \gamma_2, \langle e_2, e'_3 \rangle = \gamma_3. \end{aligned}$$

Для реализации возможности эффективного вычисления определителя матрицы 6×6 , стоящего в числителе, введены обозначения:

$$\begin{aligned} h_{ij} &= -(\alpha_i \alpha_j + \beta_i \beta_j + \gamma_i \gamma_j); i \neq j, \\ h_{ji} &= 1 - (\alpha_i^2 + \beta_i^2 + \gamma_i^2), \\ i, j &= 1, 2, 3. \end{aligned}$$

Непосредственно из формул следует, что $h_{ij} = h_{ji}$. С учетом данного обстоятельства получается выражение для вычисления определителя, стоящего в числителе базовой формулы, а, следовательно, и расстояния между соответствующими подпространствами в метрике d_{33}

$$d_{3,3}(L, L') = (h_{00}h_{11}h_{22} + 2h_{01}h_{02}h_{12}) - (h_{00}h_{01}^2 + h_{11}h_{02}^2 + h_{22}h_{01}^2) \quad (3)$$

Смысл метрики $d_{m,n}$ состоит в следующем.

Пусть a, a' два направления, взятые соответственно из подпространств L и L' . Вычисляется расстояние между ними в метрике d , (1). Минимум расстояний по всем возможным парам подобных направлений и есть расстояние между подпространствами атрибутов в метрике d_{mn} . Из этого следует, что добавление искажений в принципе не приводит к увеличению расстояния. Таким образом, необходимо оценить относительную динамику уменьшения внутриклассового и межклассового расстояния при различных конкретных вариантах применяемых трансформаций.

Апробация метода

Используя базу изображений человеческих лиц The Yale Face Database В [174], был составлен набор из 100 изображений по следующему принципу.

Для каждого из 10 представленных в базе людей (классов) отобраны 10 фронтальных изображений (экземпляров) таким образом, чтобы вариация

положения источника освещения достигала 20 градусов. Основа всех дальнейших построений исходная карта, с которой выполняются следующие действия.

Производится автоматизированный поиск центров глаз и рта. Затем, используя найденные характерные точки, выполняется поворот, масштабирование и кадрирование изображения. В результате формируются строго вертикальные изображения лица фиксированной ширины в 240 пикселей (140 пикселей между центрами глаз с отступами по 50 пикселей в каждую сторону) и различной для разных классов высоты (отступы по 50 пикселей вверх от центров глаз и вниз от центра рта). Разница в высоте получаемых изображений является средством, усиливающим различия между классами. Полученные в результате исходные карты приведены на рисунках 3.23 и 3.24.



Рисунок 3.23 - Исходные карты разных классов с одинаковым освещением.

Алгоритм в зависимости от используемого промежуточного преобразования $M \rightarrow t(M)$ и инструментов сканирования обеспечивает различные способы реализации, которые приведены в таблице 3.1.

Для каждого способа преобразования устанавливается следующая последовательность шагов:

1. Выбирается карта яркости изображения. Расстояние между векторами атрибутов рассчитывается по формуле (2). Данный алгоритм является технологией 1x1.

2. Вводятся исследуемые трансформации (см. таблица 3.2). Объединение трех трансформаций является пакетом трансформаций. Процесс составления

пакетов трансформаций выполняется интуитивно на основе простого перебора. Строятся искажения путем применения пакета трансформаций к карте яркости изображения. Далее применяется способ, аналогичный технологии 1×1 . Расстояние между пространствами векторов атрибутов рассчитывается по формуле (3). Данный алгоритм является технологией 3×3 .

3. Для каждого пакета трансформаций оценивается относительное уменьшение внутриклассового и межклассового расстояния, полученного по технологиям 1×1 и 3×3 .

Выполненная процедура позволяет оценить относительную эффективность технологии 3×3 по сравнению с технологией 1×1 , а заодно определить рекомендации по комбинации исследуемых способов и пакетов трансформаций.



Рисунок 3.24 - Исходные карты одного класса с разным освещением.

При выборе преобразований, которые используются в исследуемых трансформациях, предложено руководствоваться зрительным восприятием. Так, если взять разность двух изображений, полученных размытием фильтрами Гаусса с соседними радиусами, то результат будет содержать контуры исходного изображения (см. рисунок 3.25). А именно в контурах изображения содержится значительная часть характеризующей его информации, которая к тому же относительно устойчива к изменениям условий освещения. При вычислении

метрики dm возникает сходная ситуация с учетом контуров изображений при расчете расстояния между пространствами векторов их атрибутов.

Таблица 3.1 - Исследуемые способы промежуточного преобразования изображений.

Идентификатор способа	Наименование способа	Преобразование $M \rightarrow t(M)$	Инструмент сканирования
1	ДКП	Тождественное	Коэффициенты ДКП
2	Полиномы Лежандра	Тождественное	Коэффициенты полиномов Лежандра
3	Карты Лавса	Преобразование с использованием текстурных карт Лавса	Коэффициенты ДКП
4	Маски Собеля	Свертка фильтрами Собеля	Коэффициенты ДКП

Таблица 3.2 - Исследуемые трансформации изображений.

Идентификатор трансформации	Используемая карта	Фильтр Гаусса
1	Карта яркости	Отсутствует
2	Экваризация 64	Отсутствует
3	Экваризация 128	Отсутствует

4	Карта яркости	С радиусом 1
5	Карта яркости	С радиусом 2
6	Карта яркости	С радиусом 3
7	Экваризация 64	С радиусом 3
8	Карта яркости	С радиусом 4
9	Экваризация 64	С радиусом 4
10	Карта яркости	С радиусом 5
11	Экваризация 64	С радиусом 5
12	Карта яркости	С радиусом 6



Рисунок 3.25 - Контуры изображения, полученные с использованием фильтров Гаусса.

Свойство экваризация приводит к тому, что большие однородные области, например, лицо человека, будут содержать больше различных значений интенсивности (см. рисунок 3.26). Следовательно, они будут иметь более выраженную текстуру, что немаловажно для процесса анализа изображения. Кроме того, применение данного фильтра обеспечит нормализацию (стандартизацию) представления анализируемых изображений.



Рисунок 3.26 - Результат экваризации изображений.

При оценке относительного уменьшения внутриклассового и межклассового расстояний для разных пакетов трансформаций после применения технологии 1x1 ко всему набору изображений формируется линейный набор их векторов атрибутов $\{a_1^1, a_2^1, \dots, a_n^1\}$, где n — число классов, $n = 1, 2, \dots, N$, l — число экземпляров класса, $l = 1, 2, \dots, M$. Среднее расстояние между векторами атрибутов изображений одного класса равно $d_{1x1}^{IN} = \frac{\sum d_{1x1}(a_{n1}, a'_{n2})}{P}$, где P — число возможных комбинаций двух векторов атрибутов внутри класса, $n_1=n_2$. Далее необходимо усреднить полученные значения для всех классов $d_{1x1}^{IN} = \frac{\sum d_{1x1}^{IN}}{N}$. Среднее расстояние между векторами атрибутов изображений разных классов вычисляется по формуле $d_{1x1}^{OUT} = \frac{\sum d_{1x1}(a_{n1}, a'_{n2})}{Q}$, где Q — число возможных комбинаций двух векторов атрибутов

между классами $n_1 \neq n_2$. Отношение $K_{1x1} = \frac{d_{1x1}^{OUT}}{d_{1x1}^{IN}}$ показывает, во сколько раз измеренное по технологии 1x1 межклассовое расстояние превосходит внутриклассовое расстояние. Аналогично для каждого пакета трансформаций m по технологии 3x3 вычисляется $K_{3x3}^m = \frac{d_{3x3}^{OUT}}{d_{3x3}^{IN}}$.

Таким образом, отношение $K_m = \frac{K_{3x3}^m}{K_{1x1}}$ позволяет сопоставить технологии 1x1 и 3x3 по параметру относительного межклассового расстояния.

Полученные результаты

По итогам выполненных исследований удается достичь увеличения относительного межклассового расстояния, измеренного по технологии 3x3, по сравнению с технологией 1x1 более чем в 7 раз (см. рисунок 3.27). Полученный результат подтверждает практическую применимость предложенного метода. Например, в задаче распознавания полученные результаты означают уменьшение вероятности пересечения векторов атрибутов разных классов, то есть неправильного распознавания. В таблице 3.3 приведены пакеты трансформаций, показавшие лучшие результаты по параметру K_m для каждого исследуемого способа.

Следует отметить, что наибольшего увеличения данного показателя удастся достичь при использовании способа «Полиномы Лежандра» со следующим пакетом трансформаций:

- экваризация 64;
- экваризация 64 с последующим размытием фильтром Гаусса с радиусом 1;
- размытие фильтром Гаусса с радиусом 6.

Анализируя данные, приведенные в таблицах 3.1-3.3, следует заметить, что для разных способов перечень наиболее применимых трансформаций не совпадает.

Для анализа данного факта построена гистограмма процентного распределения трансформаций в 10 наиболее эффективных по параметру K_m пакетах (см. рисунок 3.28). Из рисунка следует, что способы «ДКП» и «Полиномы

Лежандра» используют практически одинаковые трансформации такие, как экваризация 64 и экваризация 64 с последующим размытием фильтром Гаусса с радиусом 3.

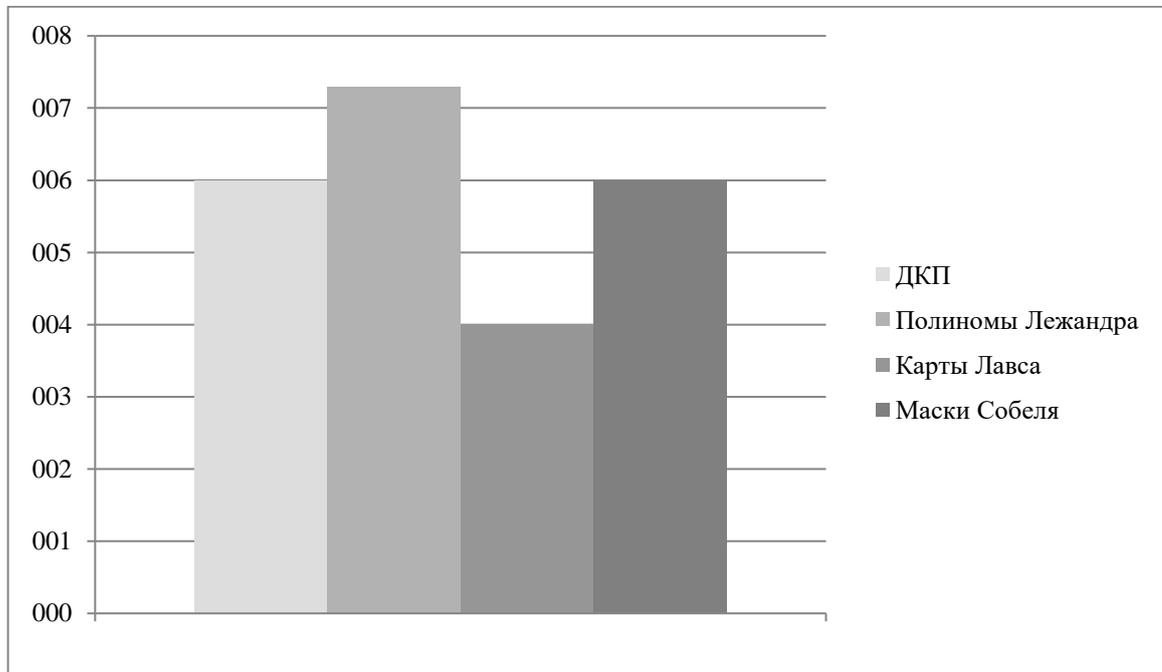


Рисунок 3.27 - Наибольшее увеличение межклассового расстояния относительно внутриклассового.

Способ «Карты Лавса» оказался единственным, в которой эффективно используется исходная карта яркостей без модификаций и в совокупности с экваризацией 128. Следует отметить, что только способ «Маски Собеля» из всех исследованных способов может эффективно применяться без использования экваризации исходной карты яркостей.

Таким образом, в целях достижения наилучших результатов в процессе сравнения изображений перед применением любого нового способа, сформированного на основе разработанного алгоритма преобразований, следует тщательно подбирать используемый пакет трансформаций.

Таблица 3.3 - Экваризация 64 с последующим размытием фильтром Гаусса с радиусом 1.

Идентификатор способа	Идентификатор трансформации 1	Идентификатор трансформации 2	Идентификатор трансформации 3	K_m
1	2	7	12	6,121
2	2	7	12	7,38833
3	2	5	8	6,02564
4	1	3	10	4,07163

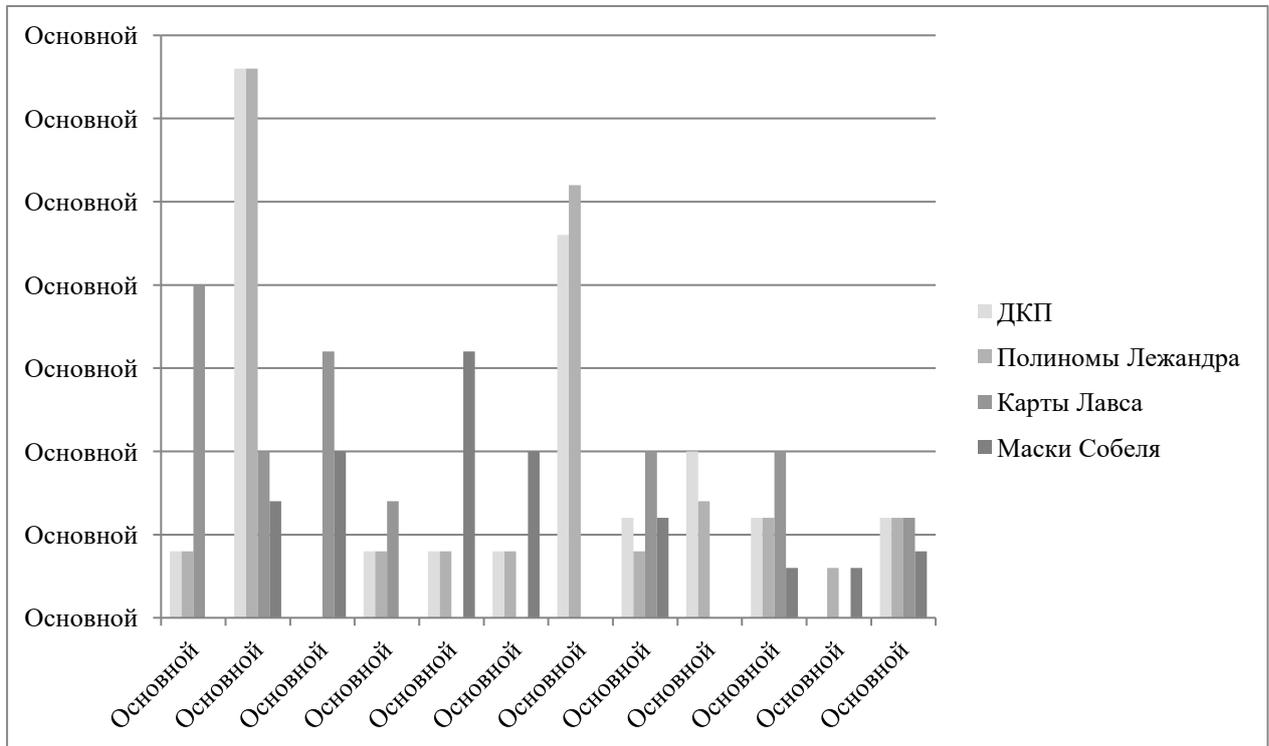


Рисунок 3.28 - Процент применения исследуемых трансформаций в пакетах показывающих лучшие результаты по параметру K_m

Разработан метод сравнения двух изображений, основанный на анализе расстояний между пространствами векторов атрибутов сравниваемых изображений. Пространства строятся по предложенному алгоритму, использующему размытие анализируемых изображений фильтрами Гаусса.

Задача нахождения расстояния между пространствами векторов не тривиальна, поэтому для ее решения разработан алгоритм на основе вычисления определителя Грама.

Результаты проверки алгоритма подтвердили его работоспособность и эффективность применения для практического решения задачи распознавания в условиях, когда количество хранимых в базе эталонных изображений ограничено одним единственным.

3.4 Выводы по 3-ей главе

1) Разработана структуры информационно-технических средств визуализации процессов распознавания, контроля, диагностики и испытания видеосистем. Выполнено проектирование программной части системы, проведена декомпозиция на подсистемы и модули.

2) Разработаны средства информационно-программного кодирования для взаимодействия элементов визуализации с элементами интерактивного управления видеосистемами. Определен состав средств системного программного обеспечения.

3) Разработана методика управления средствами видеоконтроля процессов распознавания. Выполнена программная реализация алгоритмов и структур данных, входящих в состав разработанных модулей.

4) Впервые разработана модульная программа идентификации человека по растровому двумерному изображению лица, вошедшая в состав аппаратно-программного комплекса (см. Свидетельство №2011611917).

5) Впервые разработана база данных биометрических признаков, вошедшая в состав аппаратно-программного комплекса (см. Свидетельство №2010620768).

Глава 4 Экспериментальные исследования и внедрение в промышленности алгоритмов и моделей процессов распознавания динамических телевизионных изображений

4.1 Результаты экспериментальных исследований средств распознавания, контроля и обработки телевизионных изображений

Аттестационные экспериментальные исследования разработанных АПК и системы в целом выполнены в ходе испытаний, проведенных на станции метро «Ладожская» в г. Санкт-Петербург по завершении специализированных научно-исследовательских работ в лабораториях ВлГУ.

Состав передающей части, представленной на испытаниях

На испытаниях передающая часть системы была представлена тремя однотипными цифровыми камерами Sanyo VCC-2500. Каждая из камер транслировала цветное видеоизображение с разрешением 1920×1080 пикселей в формате H.264.

На камерах были установлены следующие объективы:

- в турникете №10 – объектив с переменным фокусным расстоянием $4 \div 10$ мм, выставленный в положение с фокусным расстоянием около 6 мм;
- в турникете №9 – объектив с фиксированным фокусным расстоянием 4 мм;
- на колонне – трансфокатор с переменным фокусным расстоянием $10 \div 120$ мм, выставленный в положение с фокусным расстоянием около 75 мм.

Видеокамеры находились в режиме автоматической регулировки экспозиции, усиления и шумоподавления.

Взаимодействие средств АПК в ходе испытаний

Обобщенная схема взаимодействия средств АПК разработанной системы распознавания лиц представлена на рисунке 4.1. Основные моменты взаимодействия средств состоят в следующем:

1) Кадры с каждой из двух камер, установленных в турникете и на колонне, обрабатываются независимо детектором лиц. Результаты обработки отображаются оператору в окне просмотра видеоизображения (а именно, «рамки», ограничивающие лицо);

2) Обнаруженные изображения лиц поступают в общую очередь регистрации. Из этой очереди изображения последовательно записываются в базу данных АПК и отображаются в интерфейсе оператора как детектированные лица;

3) После записи в базу изображение поступает в один из 10 параллельных потоков распознавания. Каждый из них выполняет обработку изображения с извлечением биометрических атрибутов по 5 методикам, а также поиск ближайших эталонов по базе данных.

4) По завершении распознавания выполняется сохранение результатов поиска в базе АПК и отображение их оператору.

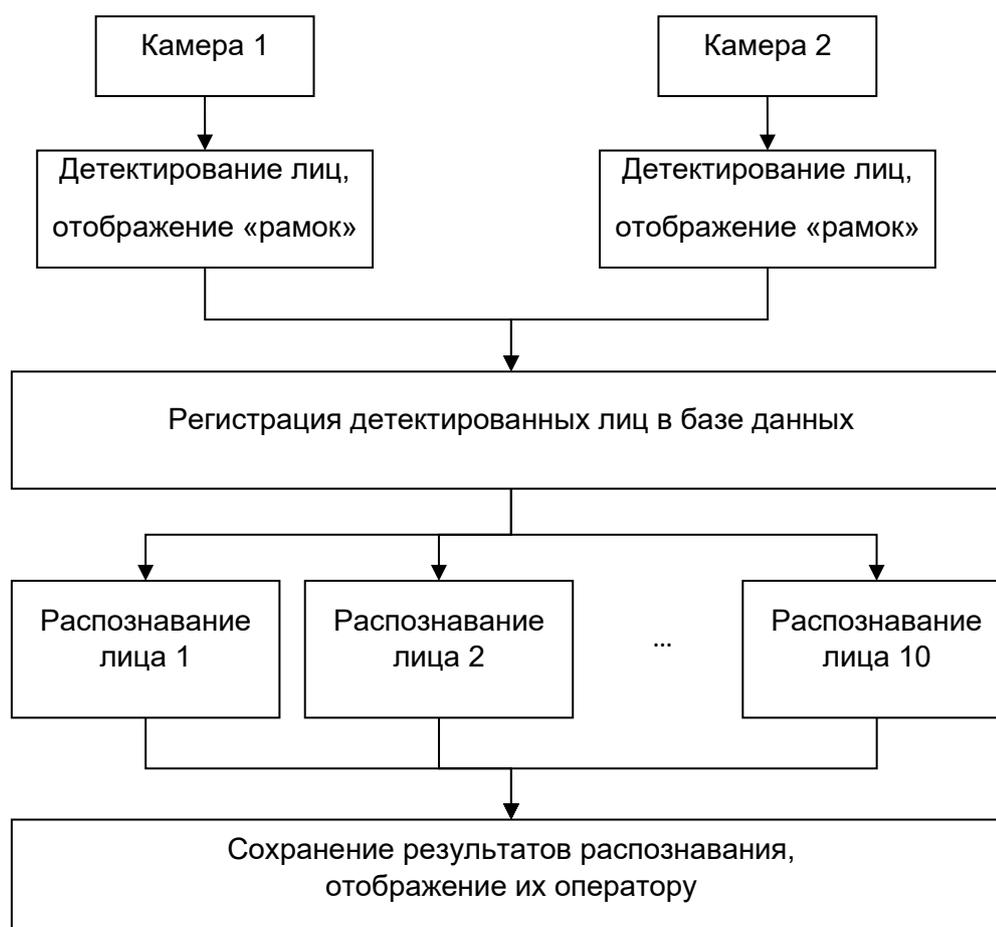


Рисунок 4.1 - Схема взаимодействия средств АПК в ходе испытания системы.

Результаты, достигнутые в ходе предварительных испытаний

Вероятность корректного детектирования лица на экране монитора оператора составила:

Турникет 10	1,0 (100%)
Турникет 9	1,0 (100%)

Вероятность регистрации (сохранения) детектированных лиц в базе данных АПК составила:

Турникет 10	0,95 (95%)
Турникет 9	0,76 (76%)

Вероятность корректного распознавания зарегистрированных лиц составила:

Турникет 10	0,463 (46,3%)
Турникет 9	0,408 (40,8%)

Основные выводы по результатам предварительных испытаний

1) Камера, размещенная на колонне напротив турникетов, малоприспособлена для использования в целях распознавания лиц. Это обусловлено её запредельным месторасположением по отношению к направлению движения людей, проходящих зону контроля и находящихся между турникетами. В подавляющем большинстве случаев распознаваемое лицо находилось под неприемлемым вертикальным углом наклона к оси камеры (см. рисунок 4.2). Такой ракурс превышает максимально допустимое значение угла для корректного сопоставления регистрируемых изображений с фронтальным эталоном.



Рисунок 4.2 - Типичное изображение лица с камеры на колонне.

2) Распознавание лиц, находящихся на значительном удалении от камеры, требует применения более высококачественной оптики и хорошего освещения в зоне контроля. Использованный объектив и условия испытаний не позволили получить с камеры изображения с достаточной для распознавания детализацией.

3) На камерах, установленных в турникетах, использование 6-мм объектива по отношению к 4-мм имеет преимущество. Более длиннофокусный объектив позволяет получать более крупные и менее искаженные изображения, при этом поле зрения 6-мм объектива оказалось достаточным для захвата всех лиц людей, участвовавших в испытаниях (хотя люди существенно различались по росту).

4) На эффективность распознавания в целом (с учетом регистрации) существенное влияние оказывает интенсивность потока людей через зону контроля.

Недостатки, выявленные в ходе предварительных испытаний системы

По результатам испытаний АПК имеются расхождения между количеством детектируемых людей (100%) и числом регистрируемых лиц (от 76% до 95%). Это означает, что в среднем около 15% лиц вообще «не обрабатываются алгоритмом распознавания», так как в данном случае не было ни одного факта регистрации

изображения этих лиц. Кроме того, в ходе обработки данных было установлено, что для 10-15% из числа зарегистрированных людей количество изображений лица составляет не более двух по причине низкого качества изображения при интенсивном потоке людей через зону контроля. Поэтому эти люди также «выпадали» из процесса распознавания. Таким образом, было установлено, что наиболее узким местом при построении современных систем распознавания лиц является очередь регистрации.

Меры, принятые для устранения недостатков

Для устранения выявленных недостатков в работе АПК в систему внесены следующие изменения:

- 1) решено отказаться от единой очереди регистрации и распределить её между потоками, подлежащими распознаванию;
- 2) при детектировании усовершенствована функция формирования качественного входного изображения за счет ужесточения требований к результатам фильтрации изображений, поступающих на регистрацию. Тем самым снижена нагрузка на АПК при регистрации и распознавании лиц;
- 3) выполнена оптимизация потоков распознавания, обеспечившая более эффективную обработку большего числа лиц;
- 4) введено ограничение на максимальное число изображений лица одного человека и разделение приоритетов при обработке лиц других людей или лиц уже отслеживаемых субъектов.

В совокупности принятые меры позволили практически полностью исключить проблему отсутствия/недостатка входных изображений для алгоритма распознавания.

Результаты, полученные на основе видеозаписей

После устранения выявленных недостатков было проведено моделирование тех же условий испытания в режиме «отложенного задания», то есть с использованием видеозаписи прохода людей через турникеты. Также было проведено исследование возможности повышения вероятности корректного

распознавания путем изменения параметров, характеризующих разработанный алгоритм распознавания.

В итоге вероятность регистрации (сохранения) детектированных лиц в базе данных АПК повысилась:

Турникет 10	0,96 (96%)
Турникет 9	0,91 (91%)

Для оценки вероятности корректного распознавания зарегистрированных лиц средствами модернизированного АПК выполнены дополнительные экспериментальные исследования.

Эксперимент 1 (повторение условий предварительных испытаний)

Последовательность действий:

- использование тех же пяти методик распознавания;
- использование того же правила для принятия решения о распознавании, а именно, формирование списка длиной не более двух элементов.

Результаты эксперимента 1:

Точки видеонаблюдения	Вероятность корректного распознавания	Вероятность ложноположительного распознавания
Турникет 10	0,583 (58,3%)	0,063 (6,3%)
Турникет 9	0,637 (63,7%)	0,033 (3,3%)

Эксперимент 2 (модификация правила принятия решения)

Последовательность действий:

- использование тех же пяти методик распознавания;
- правило принятия решения о распознавании - формирование списка длиной не более трех элементов.

Результаты эксперимента 2 (три методики дали взаимосогласованный результат):

Точки видеонаблюдения	Вероятность корректного распознавания	Вероятность ложноположительного распознавания
-----------------------	---------------------------------------	---

Турникет 10	0,677 (67,7%)	0,104 (10,4%)
Турникет 9	0,725 (72,5%)	0,066 (6,6%)

Эксперимент 3 (более полное использование методик)

Последовательность действий:

- использование тех же пяти методик распознавания после предварительного применения других 21-ой методики, а именно, 21 методика формируют «длинный список», из которого затем 5 методик формируют «короткий список»;

- правило принятия решения о распознавании - формирование списка длиной не более трех элементов.

Результаты эксперимента 3 (три методики дали взаимосогласованный результат):

Точки видеонаблюдения	Вероятность корректного распознавания	Вероятность ложноположительного распознавания
Турникет 10	0,708 (70,8%)	0,188 (18,8%)
Турникет 9	0,747 (74,7%)	0,121 (12,1%)

Эксперимент 4 (адаптация методик)

Последовательность действий:

- использование других пяти методик распознавания, имеющих наилучший результат;

- правило принятия решения о распознавании - формирование списка длиной не более трех элементов.

Результаты эксперимента 4 (три методики дали взаимосогласованный результат):

Точки видеонаблюдения	Вероятность корректного распознавания	Вероятность ложноположительного распознавания
Турникет 10	0,77 (77,0%)	0,094 (9,4%)
Турникет 9	0,791 (79,1%)	0,088 (8,8%)

Результаты экспериментальных исследований для каждой из использованных методик приведены в таблице 4.1.

Таблица 4.1 – Данные экспериментальных исследований методик.

Обозначение методики (* - использовалась в СПб)	Распознано людей (из 100)		Распознано изображений, %	
	Турникет 10	Турникет 9	Турникет 10	Турникет 9
LW:	55	38	36,8	25,1
LW:	52	47	31,0	27,8
LW:	49	42	29,5	22,7
LW:	43	22	26,9	14,0
LW:	38	27	18,1	13,7
LW2:	71	65	53,5	44,1
LW2:	65	60	43,3	44,5
LW2:	70	62	51,8	37,1
LW2:	69	53	49,7	33,1
LW2:	52	49	30,4	26,4
LW3:	71	68	54,4	46,5
LW3:*	64	62	43,6	47,8
LW3:*	68	62	52,6	38,5
LW3:	67	51	49,4	32,1
LW3:*	44	49	26,3	27,1
LG:	54	50	39,2	35,5
DC:	58	52	39,5	35,8
LG2:	55	48	40,1	35,1
DC2:	65	58	48,8	43,5
LG3:	53	47	37,4	32,1
DC3:*	62	60	47,7	45,2
FC:	11	4	3,8	1,3
FCw:	9	5	3,2	2,0
FCw:	3	6	1,2	2,7
FC2:	0	0	0,0	0,0
FCw2:	8	4	2,9	1,3

FCw2:	2	3	0,9	1,0
FC3:	6	12	1,8	5,7
FCw3:	8	4	2,9	1,3
FCw3:	2	3	0,9	1,0
SB:	61	53	40,4	33,8
SB+:	65	55	48,8	38,1
SB2:	80	73	62,9	54,5
SB+2:	4	6	1,2	2,0
SB3:*	77	78	62,3	58,5
SB+3:	4	8	1,2	2,7

Сравнение с узкоспециализированным алгоритмическим комплексом "COGNITEC".

При изучении изображений целесообразно использовать многоуровневые базисы (например, вейвлет-базисы), составляющие специфический аппарат отклика масок в виде базиса Фрея-Чена для выделения именно локальных текстурных особенностей изображения (см. главу 2). В данном случае предлагается рассматривать изображения на двух уровнях детализации.

Первый уровень основной - обычно пиксельный. При этом фото дополнительно разбивается на клетки 3x3 (на мегапиксели). Каждая клетка рассматривается как отдельная самостоятельная структура, которая отображается при помощи откликов специальных масок. Отклики или их производные в дальнейшем формируют микротекстуру. Наличие микротекстуры позволяет идентифицировать мегапиксели и отслеживать динамику их преобразования.

Второй уровень – мегапиксельный. Фото также разбивается на клетки 3x3, но при этом яркости внутри клеток выравниваются, то есть в каждом пикселе определяется средняя яркость по каждой клетке. Таким образом мегапиксели лишаются микротекстуры.

Искажения изображений, возникающие в условиях турбуленции, имеют специфические особенности. Характерны следующие базисные типы

преобразований, из которых строятся искажения. Условное наименование и формальное описание искажений состоит в следующем.

1) «Туман» - его моделирует следующее преобразование яркости мегапикселя, а именно, $(x_1, \dots, x_9) \rightarrow \alpha(1, \dots, 1) + \beta(x_1, \dots, x_9)$, являющееся «наложением нейтрального тумана». В частности, допускается предельный случай $\beta = 0$, при этом рассматривается только простейший случай добавления константы. Коэффициенты α, β разные для каждого мегапикселя; текстурное содержание мегапикселей при этом не меняется, кроме предельного случая.

2) «Преломление» - происходят локальные изменения геометрии. Для их моделирования предлагается использовать перестановку мегапикселей. Подобные преобразования также не меняют текстуру мегапикселей.

3) «Смазывание» — это локальные размытия фото. Их предлагается осуществлять, применяя гауссовы размытия отдельно по мегапикселям, то есть для каждого формирующего его пикселя применяется гауссов фильтр. Текстура при этом претерпевает изменения, поскольку используется информация, распложенная в соседних мегапикселях.

Таким образом, предлагается реконструировать изменение фото сцены, которое возникает при помощи последовательного применения набора элементарных трансформаций, которые описаны выше.

Текстурные характеристики

В качестве оценки меры отличия искажений предлагается использовать следующий инструмент – «порядковые гистограммы». Окно интереса имеет размеры $3m \times 3n$ и представляет собой объединение непересекающихся мегапикселей. Для построения «порядка», являющегося текстурной характеристикой мегапикселя, используются буквы алфавита, например, a, b, c, d . В случае равенства энергий приоритет отдается букве, стоящей ранее по алфавиту. Например, если $a < b = c < d$, то порядок именно $abcd$, а не $acbd$ и т.д. В этом случае порядок $abcd$ обозначает выполнение неравенства $a \leq b \leq c \leq d$. В частности, такой же «порядок» присваивается при полном равенстве энергий

откликов. Таким образом, если все яркости пикселей одинаковые (например, если a, b, c, d равны нулю), то присваивается именно порядок $abcd$.

Далее, каждому «порядку» присваивается некоторый идентификатор, например, порядковый числовой номер (всего их $4!=24$), а номеру соответствует накопитель. Так как каждому мегапикселю соответствует «порядок», то по нему определяется номер накопителя, в который помещается единица. Таким образом, $m \times n$ единиц распределяется по 24 накопителям. Если нормировать содержание накопителей, то в результате получается объект G , который является «порядковой гистограммой». Таким образом, совокупная текстурная характеристика представляет собой гистограмму G .

Если G^* аналогичный объект, который характеризует содержание другого окна, то степень похожести текстурного содержания окон может быть измерена величиной $1 - 0.5|G - G^*|$. Здесь необходимо отметить, что данная мера является весьма грубой мерой сравнения. Обычно для сравнения фото выбирают сразу несколько окон интереса. Вектор, образованный степенями похожести сравниваемых окон, предоставляет более совершенный инструмент оценки.

Инструмент «порядковые гистограммы» изначально включен в набор методик, входящих в «Стенд» и никак не связан с таким предметом, как «влияние турбулентности атмосферы на изображение сцены».

Эксперимент

В качестве исследуемого объекта предлагается рассматривать набор фото, полученных при помощи следующей процедуры.

В основу процедуры положен эталон (см. рисунок 4.3). Затем при помощи трансформаций строится набор искажений. Далее предлагается оценить степень похожести пар фото, которые выбираются из множества различными способами. Предполагается оценивать похожесть по текстурному содержанию, являющемуся индикатором «текстурный анализатор» (ТА), с применением алгоритма «COGNITEC». Последний представляет собой узкоспециализированный алгоритм высокой эффективности, используемый для сравнения лиц. Узкая специализация характеризует его «не универсальность». Алгоритм сравнивает только

изображения лиц и, если повреждены критические зоны лица, то не выполняет построение атрибутов сравнения. Так как текстурный способ сравнения универсален, то вводятся ограничения для обеспечения возможности работы «COGNITEC». Для этого в качестве исходного эталона используется фото лица, например, из базы FERET. Существенным обстоятельством для анализа результатов сравнения является факт «не отказа» «COGNITEC» от детектирования построенного искажения эталона.

Применение базисных преобразований, описанных выше, в принципе, позволяет существенно изменить и текстурное содержание фото. Однако важным является то обстоятельство, чтобы искажение воспринималось алгоритмом «COGNITEC» именно как лицо.

ВАРИАНТ 1 (кривое зеркало)

Построены искажения, в основном сохраняющие микротекстуру эталона (см. рисунок 4.4)



Рисунок 4.3 – Эталон.



A1



A2



A3



A4



A5



A6



A7



A8



A9



A10



A11



A12

Рисунок 4.4 - Искажения A1-A12

Результаты эксперимента представлены в таблице 4.2.

Таблица 4.2 – Результаты эксперимента «кривое зеркало».

Обозначение фото	ТА (Э)	«COGNITEC» (Э)
A1	0.994075	0,999723
A2	0.994075	0,999886
A3	0.998926	0,999904
A4	0.998784	0,999918
A5	0.998334	0,999984
A6	0.996057	0,999886
A7	0.997990	0,999894
A8	0.998557	0,999808
A9	0.998168	0,999868
A10	0.997467	0
A11	0.998622	0,999868
A12	0.998634	0,999851

Выводы: практически все искажения признаны алгоритмом «COGNITEC» похожими на эталон со степенью близкой к единице; кроме одного (A10), который алгоритмом не детектировался. Высокая степень схожести предопределена тем, что при построении искажений сохранялась именно микротекстура. Близость текстурного содержания сравниваемых фото является одной из основных причин высокой степени схожести.

В принципе, в качестве модели «нейтрального тумана» может служить произвольное фото с отсутствующей микротекстурой. Однако, руководствуясь спецификой эксперимента, необходимо использовать фото лица, отличное от эталона. Внутри каждого мегапикселя, во-первых, вычисляется средняя яркость. Во-вторых, внутри каждого мегапикселя выравниваются яркости пикселей, заменяя их на среднее значение. В итоге получается более размытое по сути фото низкого разрешения. Данное фото квалифицируется, с одной стороны, как «нейтральный туман». С другой стороны фото воспринимается алгоритмом

«COGNITEC» как лицо. Данная модель нейтрального тумана является «маской». Затем по исходному фото эталона может быть построено искажение при помощи следующей процедуры.

Пусть $I_k^{\text{эталон}}$ - карта яркостей k -го мегапикселя эталона; $I_k^{\text{маска}}$ - карта яркостей k -го мегапикселя маски ($k = 1, \dots, K$), $\mathcal{E}_i, \mathcal{E}'_i (i = 1, 2)$ фото эталона и его маски. Тогда для каждого мегапикселя необходимо выполнить следующую операцию $I_k^{\text{эталон}} \rightarrow \alpha_k \cdot I_k^{\text{маска}} + \beta_k \cdot I_k^{\text{эталон}} (\alpha_k \beta_k > 0)$.

ВАРИАНТ 2 (наложение маски)

Если $\alpha_k = \alpha$, $\beta_k = \beta$ и не зависят от мегапикселя, то операция, по сути, означает наложение друг на друга двух фото, взятых в разных пропорциях. В дальнейшем для удобства следует полагать, что $\alpha + \beta = 1$. Это означает, что за основу берется одно фото, а другое используется для построения нейтрального тумана. Возможны два варианта, они определяются распределением ролей, а именно, $\{\mathcal{E}_1, \mathcal{E}'_2\}$ или $\{\mathcal{E}_2, \mathcal{E}'_1\}$. Рассматриваемый вариант схемы $\{\mathcal{E}_1, \mathcal{E}'_2\}$ приведен на рисунке 4.5.



а) Эталон \mathcal{E}_1



б) Маска \mathcal{E}'_2

Рисунок 4.5 - Наложение маски.

Со смесями $\alpha\mathcal{E}_1 + (1 - \alpha)\mathcal{E}'_2$ (см. рисунок 4.6) сравнивается сначала эталон \mathcal{E}_1 , затем его маска \mathcal{E}'_1 , при этом алгоритм «COGNITEC» выдает следующие результаты:

α	COGNITEC(\mathcal{E}_1)
0.9	0.9999642
0.8	0.9999011
0.7	0.999809
0.6	0.9997096
0.5	0.9995313
0.4	0.9573702
0.3	0.5024528
0.2	0.4000888
0.1	0.3571693

 $\alpha = 0.9$  $\alpha = 0.8$  $\alpha = 0.7$  $\alpha = 0.6$  $\alpha = 0.5$  $\alpha = 0.4$

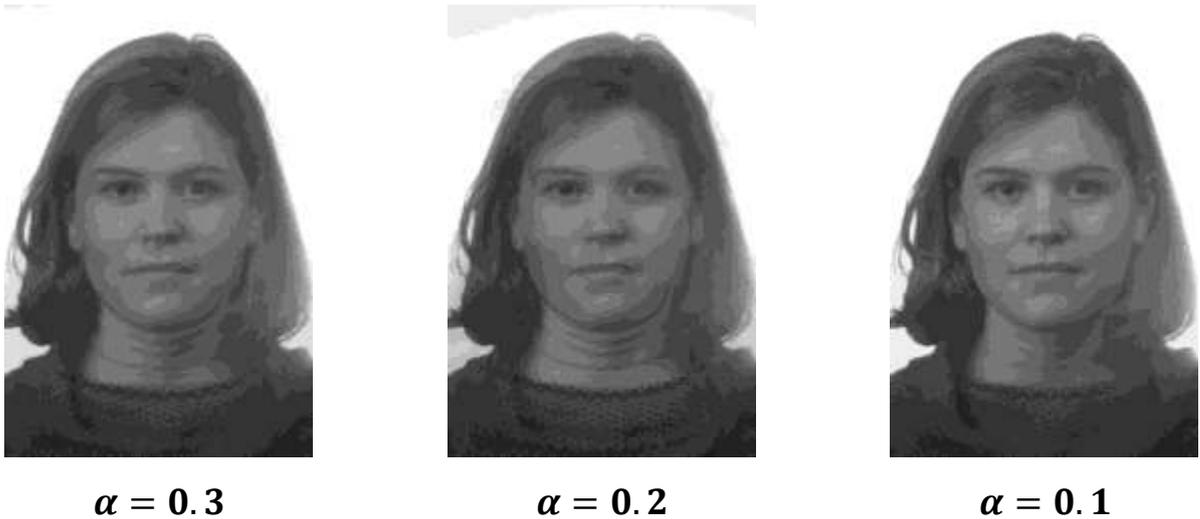


Рисунок 4.6 - Смеси $\alpha\mathcal{E}_1 + (1 - \alpha)\mathcal{E}'_2$.

Вывод. Индикатор «ТА» указывает на полную похожесть «смеси» (эталон + туман) эталону, так как при построении смеси накладывалась маска, которая изначально была лишена микротекстуры. Данное обстоятельство противоречит визуальному восприятию. По-видимому, именно это объясняет и высокую устойчивость показателя похожести при сравнении смеси с эталоном \mathcal{E}_1 , предоставляемую алгоритмом «COGNITEC». Таким образом, алгоритм «COGNITEC» дает результаты, противоречащие визуальному восприятию человека (см. $\alpha = 0.4$; степень похожести эталону 0.9573702). Далее в качестве эталона выбирается \mathcal{E}'_1 и вновь сравнивается с теми же смесями. С визуальной точки зрения оценка выглядит более реалистично. Подтверждается высокий уровень зависимости оценок алгоритма «COGNITEC» от компонентов микротекстуры.

ВАРИАНТ 3 (фотороботы)

Известно, что узкоспециализированные алгоритмы распознавания лиц теряют устойчивость, если сталкиваются с вариантами маскировки. Например, уже наличие очков вызывают у них затруднение. Для проверки реакции алгоритмов на различные виды маскировки построены несколько фотороботов. Искажению подвергалось не всё фото эталона, а только его часть - некоторая зона интереса. В

этом случае в мегапикселях, представляющих зону интереса, α_k берется близким к единице, в остальных мегапикселях $\alpha_k = 0$. При этом возможны две схемы $\mathcal{E}_1, \mathcal{E}'_2$ или $\mathcal{E}_2, \mathcal{E}'_1$. В первом случае фоторобот несет микротекстуру \mathcal{E}_1 во втором - \mathcal{E}_2 . Здесь эталон и его маска являются антиподами в текстурном смысле. Последнее означает, что их сравнение при помощи индикатора «ТА» фиксирует малую степень похожести. Сравниваются эталон \mathcal{E}_1 , его маска \mathcal{E}'_1 и фотороботы, полученные по схеме $\mathcal{E}_1, \mathcal{E}'_2$ (см. рисунок 4.7):



R1



R2



R3



R4



R5



R6



R7

Рисунок 4.7 - Фотороботы R1-R7.

Результаты сравнения представлены таблицей

R	COGNITEC(\mathcal{E}_1)
R1	0.9998242
R2	0.9998744
R3	0.999836
R4	0.9997308
R5	0.99968
R6	0.9997503
R7	0.9996186

Вывод

Варианты искажений «Фотороботы», с точки зрения индикатора «ТА», эквиваленты эталону \mathcal{E}_1 (по построению), что является ошибочным результатом. Используемые программным обеспечением «COGNITEC» алгоритмы при сравнении «фотороботов» с эталоном \mathcal{E}_1 также дают неоднозначные результаты. Во-первых, алгоритм «COGNITEC» определил «фотороботы» как лицо; это свидетельствует о его неустойчивости к маскировке. Во-вторых, была определена степень похожести фоторобота эталону \mathcal{E}_1 , как близкую к единице; в то время как

визуально эти изображения крайне несхожи. Двойная ошибка, совершенная алгоритмами «COGNITEC» может быть объяснена только тем, что «фотороботы» имеют одинаковую по построению микротекстуру с эталоном.

Индикатор «ГА» верно фиксирует непохожесть «фотороботов» изображению \mathcal{E}'_1 , в то время как алгоритм «COGNITEC» ошибочно даёт высокую степень схожести.

В итоге проблемы, выявленные в результате данных экспериментов, подтвердили следующее. Алгоритм «COGNITEC», действуя в идеальных или близких к ним условиях, очень эффективен. Однако, если изображения в анализируемом потоке имеют ряд искажений или потерь, то его эффективность сводится практически к нулю. При этом зачастую правильный ответ даёт наиболее примитивный индикатор, например, оценивающий только текстурное содержание фото. Анализ результатов эксперимента подтверждает выводы о целесообразности использования универсальной платформы, что предполагает одновременное применение большого числа разноплановых инструментов. В данном случае использование подобной платформы не является альтернативой узкоспециализированным алгоритмам. Использование такой платформы целесообразно для решения задач, когда отсутствуют алгоритмы узкой специализации либо требуется существенное улучшение достижимых показателей качества ожидаемых решений.

4.2 Внедрение результатов исследования в промышленности и в учебный процесс

Результаты диссертационных исследований в части изучения и применения разработанных алгоритмов и методик цифровой обработки, а также распознавания образов, внедрены в Федеральном государственном бюджетном образовательном учреждении высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (ВлГУ) и использованы в учебном процессе при подготовке бакалавров и магистров по

направлениям «Лазерная техника и лазерные технологии», «Прикладная математика и информатика», «Нанотехнологии и микросистемная техника» в рамках проведения учебных занятий по дисциплинам «Оптоэлектроника», «Проектирование электронных средств в наноэлектронике», «Прикладное программирование» и «Инструментальные средства разработки программного обеспечения».

Результаты работы использованы при проведении лекционных, лабораторных, практических и семинарских занятий, а также в научно-исследовательской работе студентов, в курсовом и дипломном проектировании (см. Приложение Б).

Научно-технические результаты диссертационных исследований включены в итоговый отчет ВлГУ по проблеме: «Разработка технических средств распознавания образов по анализу видеоизображения для автоматизации работы оператора» по Государственному контракту с Минобрнауки РФ от 17 ноября 2009 г. № ГК №02.524.11.4010 в рамках реализации Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России» на 2007 – 2013 годы, а также используются в научно-исследовательской работе в рамках Государственного задания ВлГУ по теме «Моделирование и прогнозирование многопараметрических задач и многофакторных процессов на основе новых физических принципов и математических методов нелинейной динамики с использованием систем искусственного интеллекта, параллельных вычислений и обработки видеоинформации» и по теме «Новые физические методы лазерного синтеза микро- и наноструктурированных углеродосодержащих материалов: экспериментальное и теоретическое изучение процессов их формирования и распознавания наноструктур» (см. Приложение В).

Достигнутые научно-технические результаты явились основанием для заключения от 21.03.2014 договора 1769ГУ1/2014 «Разработка алгоритмов биометрической идентификации человека в условиях единичного низкокачественного эталона» с Федеральным государственным бюджетным

учреждением «Фонд содействия развитию малых форм предприятий в научно-технической сфере» (Фонд содействия инновациям) и успешного выполнения работ по Программе «УМНИК-2014» (см. Приложение Г).

Разработанная и научно обоснованная система распознавания образов в условиях единичного низкокачественного эталона, предложенные алгоритмы, методики, а также их программная реализация подтвердили свою работоспособность и высокие качественные показатели в ходе проведения сравнительных испытаний и конкурсов проводимых для тестирования различных систем распознавания образов.

Результаты работы используются в научно-практической деятельности магистрантов и аспирантов базовых кафедр ВлГУ, функционирующих на промышленных предприятиях Владимирской области.

Результаты работы используются или планируются к внедрению на предприятиях Владимирской области, таких как: в рамках НИР и ОКР проводимых Федеральным казённым предприятием «Государственный лазерный полигон «Радуга», г.Радужный Владимирской области, ООО «Техника и технология», г.Владимир, и ООО «РУСОКСИД», г.Москва, территория Сколково, что подтверждается актами практического использования и реализации (см. Приложение Г).

4.3 Выводы по 4-й главе

1) Разработан и изготовлен опытный образец аппаратно-программного комплекса «Стенд» для проведения лабораторных и натурных испытаний в реальных условиях эксплуатации.

2) Выполнены экспериментальные лабораторные и натурные исследования средств распознавания, контроля и обработки телевизионных изображений на объекте транспортной инфраструктуры метрополитена в г. Санкт-Петербург. Дан анализ недостатков в работе программного обеспечения и выполнена доработка аппаратно-программных средств. Достигнута вероятность корректного

распознавания лиц при использовании единичных низкокачественных эталонов до 0,7 ... 0,79 (70...79%).

3) Проведены экспериментальные исследования аппаратно-программных средств в сравнении с известными системами на основе «SDK Cognitec Systems GmbH» в условиях неравномерного естественного освещения, перекрытия лиц субъектов в сплошном потоке пассажиров в «пиковые» часы работы метрополитена. Показаны преимущества разработанных аппаратно-программных средств.

4) Подтверждена востребованность и практическая значимость аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц для использования в инновационной научной, образовательной и производственной деятельности учебных заведений и промышленных предприятий.

Заключение

Основные результаты диссертационного исследования и следующие из них теоретические и практические выводы сформулированы следующим образом:

- обоснованы методы и средства процесса распознавания и обработки телевизионных изображений, разработана структура процедуры распознавания лица человека на цифровом изображении;

- исследованы и разработаны информационно-технологические модели и алгоритмы распознавания и обработки телевизионных изображений, обоснованы этапы предварительной подготовки изображения, определена эффективность применения методов признакового описания;

- исследованы и разработаны средства программного обеспечения удаленного контроля, диагностики и обработки телевизионных изображений, проработан способ пакетного сравнения эталонных и тестовых изображений, предопределяющий процедуру предварительного применения вполне определенных трансформаций;

- разработана структуры информационно-технических средств визуализации процессов распознавания, контроля, диагностики и испытания видеосистем, выполнено проектирование программной части системы, проведена декомпозиция на подсистемы и модули;

- разработаны средства информационно-программного кодирования для взаимодействия элементов визуализации с элементами интерактивного управления видеосистемами, определен состав средств системного программного обеспечения;

- разработана методика управления средствами видеоконтроля процессов распознавания, выполнена программная реализация алгоритмов и структур данных, входящих в состав разработанных модулей;

- разработан и изготовлен опытный образец аппаратно-программного комплекса «Стенд» для проведения лабораторных и натурных испытаний в реальных условиях эксплуатации;

- выполнены экспериментальные лабораторные и натурные исследования средств распознавания, контроля и обработки телевизионных изображений на объекте транспортной инфраструктуры метрополитена в г. Санкт-Петербург, дан анализ недостатков в работе программного обеспечения и выполнена доработка аппаратно-программных средств, достигнута вероятность корректного распознавания лиц при использовании единичных низкокачественных эталонов до 0,7 ... 0,79 (70...79%);

- проведены экспериментальные исследования аппаратно-программных средств в сравнении с известными системами на основе «SDK Cognitec Systems GmbH» в условиях неравномерного естественного освещения, перекрытия лиц субъектов в сплошном потоке пассажиров в «пиковые» часы работы метрополитена, показаны преимущества разработанных аппаратно-программных средств;

- подтверждена востребованность и практическая значимость аппаратно-программных средств обработки динамических телевизионных изображений и распознавания лиц для использования в инновационной научной, образовательной и производственной деятельности учебных заведений и промышленных предприятий.

- впервые разработаны: программа полуавтоматической аппроксимации экспериментальных данных базисным набором гауссиан (см. Свидетельство № 2017611477); программное обеспечение автоматизированного тестирования алгоритмов распознавания лиц (см. Свидетельство №2014612123); модульная программа идентификации человека по растровому двумерному изображению лица (см. Свидетельство №2011611917); база данных биометрических признаков (см. Свидетельство №2010620768), - вошедшие в состав аппаратно-программного комплекса.

Научные и научно-прикладные результаты, достигнутые в диссертационной работе, могут быть использованы при создании новых и модернизации существующих методов биометрической идентификации людей для систем

безопасности и контроля доступа, в том числе, в части поиска лиц, находящихся в розыске.

Список литературы

1. Паспорт специальности 05.12.04 – Радиотехника, в том числе, системы и устройства телевидения [Электронный ресурс]. - Режим доступа: <http://arhvak.minobrnauki.gov.ru/316>. Дата обращения: 15.03.2021г.
2. Фисенко, В.Т., Фисенко, Т.Ю. Компьютерная обработка и распознавание изображений [Текст]. - СПб: Изд-во СПбГУ ИТМО, 2008. – 192 с.
3. Можейко, В.И., Фисенко, В.Т., Фисенко, Т.Ю. Автоматическое сопровождение объектов в компьютерных системах обработки изображений [Текст] // Оптический журнал. – 2007. - № 11. - Том 74. - С. 39-46.
4. Ганин, А.Н. Сопровождение и распознавание объектов на телевизионных изображениях [Текст]: автореф. дисс. ... канд. техн. наук: 05.12.04 / Ганин Александр Николаевич. – Ярославль, 2013. – 20 с.
5. Можейко, В.И., Фисенко, В.Т., Фисенко, Т.Ю. Адаптивный метод автоматического сопровождения объектов по последовательности цифровых сигналов телевизионных изображений [Текст] // Материалы 4-й международной конференции «Телевидение: передача и обработка изображений», СПб., 2005г. - С. 91-92.
6. Охотников, С.А. Распознавание видеоизображений объектов заданной формы на основе анализа их контуров [Текст]: дисс. ... канд. техн. наук: 05.12.04 / Охотников Сергей Аркадьевич. - Йошкар-Ола, 2014. – 181 с.
7. Афонин, В., Макушкин, В. Распознавание изображений [Электронный ресурс]. - Режим доступа: <https://intuit.ru/studies/courses/46/46/lecture/1374>. Дата обращения: 14. 03.2021г.
8. Пат. 2189121 Российская Федерация, МПК Н04N 13/00. Способ получения стереоскопического цветного телевизионного изображения [Текст] / Дрязгов В.Д., Колобков В.С., Петров В.С., Фисенко В.Т. - № 2000125634/09; заявл. 11.10.2000; опубл. 10.09.2002, Бюл. № 25. – 3 с.: ил.

9. Ануфриев, А.В., Бакут, П.А., Зимин, Ю.А., Толмачев, А.И. Использование функций резкости для компенсации фазовых искажений [Текст] // Квантовая электроника. - 1985. - №2. – Том 12. - С. 441–443.
10. Бакут, П.А., Польских, С.Д., Ряхин, А.Д., Свиридов, К.Н., Устинов, Н.Д. Статистический синтез алгоритмов оптимальной обработки серии искаженных атмосферой “пятенных” изображений астрономического объекта [Текст] // Проблемы передачи информации. – 1985. - №2. – Том 21. – С.32–41.
11. Бакут, П.А., Польских, С.Д., Ряхин, А.Д., Свиридов, К.Н., Устинов, Н.Д. Статистический синтез алгоритмов оптимальной обработки искаженного атмосферой пуассоновского изображения [Текст] // Квантовая электроника. - 1984. - №12. – Том 11. - С. 2405–2412.
12. Бакут, П.А., Свиридов, К.Н., Устинов, Н.Д. Методика оценки возможностей оптимального приема при обнаружении астрофизических объектов через турбулентную атмосферу [Текст] // Квантовая электроника. - 1981. - №2. – Том 8. - С. 341–346.
13. Александров, А.Б., Бакут, П.А., Логинов, В.А. Оптимальная обработка оптического сигнала, искаженного случайно-неоднородной средой [Текст] // Проблемы передачи информации. – 1979. - №4. – Том 15. – С.106–110.
14. Бакут, П.А., Баринов, В.В., Димов, Н.А., Мандросов, В.И., Печенов, А.С., Троицкий, И.Н. Об изображении диффузного объекта в случайно-неоднородной среде с амплитудными искажениями [Текст] // Квантовая электроника. - 1979. - №10. – Том 6. - С. 2093–2102.
15. Бакут, П.А., Троицкий, И.Н., Устинов, Н.Д. Анализ точности измерения координат центра тяжести оптического изображения [Текст] // Проблемы передачи информации. – 1978. - №1. – Том 14. – С.68–76.
16. Теория когерентных изображений [Текст] / П.А. Бакут, В.И. Мандросов, И.Н. Матвеев, Н.Д. Устинов; Под ред. Н.Д. Устинова. - М.: Радио и связь, 1987. – 263 с.: ил.
17. Обнаружение движущихся объектов [Текст] / П.А. Бакут, Ю.В. Жулина, Н.А. Иванчук; Под общ. ред. П.А. Бакута. - М.: Сов. радио, 1980. - 287 с.

18. Галушкин, А.И. Синтез многослойных систем распознавания образов [Текст]. — М.: Энергия, 1974. — 234 с.
19. Галушкин, А.И. Многослойные системы распознавания образов [Текст]. — М.: Изд-во МИЭМ, 1970. — 270 с.
20. Гуляев, Ю.В., Казарян, М.А., Мокрушин, Ю.М., Шакин, О.В. Акустооптические лазерные проекционные системы отображения телевизионной информации [Текст] // Квантовая электроника. - 2015. - №4. — Том 45. - С. 283–300.
21. Гуляев, Ю.В. Функциональные изображения биологических объектов [Текст] // Доклады АН СССР. - 1986. - №5. — Том 287. - С. 1088–1092.
22. Журавлев, Ю.И. Экстремальные алгоритмы в математических моделях для задач распознавания и классификации [Текст] // Доклады АН СССР. Математика. — 1976. - №3. — Том 231. — С. 96-104.
23. Журавлев, Ю.И. Об алгебраическом подходе к решению задач распознавания или классификации [Текст] // Проблемы кибернетики. — 1978. — Том 33. — С. 5-68.
24. Журавлев, Ю.И. Об алгоритмах распознавания с представительными наборами (о логических алгоритмах) [Текст] // Журнал ВМиМФ. — 2002. — №9, - Том 42. — С. 1425—1435.
25. Журавлев, Ю.И. Об алгебраических методах в задачах распознавания и классификации [Текст] // Распознавание, классификация, прогноз. — 1988. — Т. 1. — С. 9-16.
26. Журавлев, Ю.И., Гуревич, И.Б. Распознавание образов и распознавание изображений [Текст] // Распознавание, классификация, прогноз. — 1989. — Т. 2. — С. 5-73.
27. Журавлев, Ю.И., Никифоров, В.В. Алгоритмы распознавания, основанные на вычислении оценок [Текст] // Кибернетика. - 1971. - №3. — С. 83–91.
28. Загоруйко, Н.Г. Методы распознавания и их применение [Текст]. — М.: Сов. радио, 1972. — 286 с.

29. Загоруйко, Н.Г., Заславская, Т.И. Распознавания образов в социальных исследованиях [Текст]. – Новосибирск: Наука, 1968. – 326 с.
30. Загоруйко, Н.Г., Дюбанов, В.В. Методы ускорения процесса поиска ближайшего аналога при распознавании большого числа образов. Алгоритм «Локатор» [Текст] // Автометрия. – 2004. - Том 6. – С. 34-45.
31. Загоруйко, Н.Г. Распознавание образов методом попарного сравнения эталонов [Текст] // Доклады РАН. – 2002. - №1. – Том 382. - С. 1-3.
32. Загоруйко, Н.Г. Распознавание с использованием знаний о пространстве образов [Текст] // Труды IX Международной конференции «Знание-Диалог-Решение» (KDS-2001). Санкт-Петербург, июнь 2001 года. - С. 246-251.
33. Загоруйко, Н. Г., Самохвалов, К.Ф. Распознавание образов по динамическим признакам (алгоритм ДИП) [Текст] // Вычислительные системы, вып. 36, Новосибирск, 1969.
34. Загоруйко, Н.Г. Классификация задач распознавания образов [Текст] // Вычислительные системы, вып. 22, Новосибирск, 1966. - С. 146-153.
35. Потапов, А.А., Гуляев, Ю.В., Никитов, С.А., Пахомов, А.А., Герман, В.А. Новейшие методы обработки изображений [Текст] / Под ред. А.А. Потапова. - М.: Физматлит, 2008. - 496 с.
36. Зинин, А.М., Виниченко, И.Ф., Житников, В.С, Овсянникова, М.Н. Криминалистическое описание внешности человека (функциональные и сопутствующие элементы и признаки): Справочное пособие [Текст] / Под ред. проф. В. А. Снеткова. — М.: Изд-во ВНИИ МВД СССР, 1988. — 242 с.: ил.
37. Сойфер, В.А. Обработка изображений в автоматизированных системах научных исследований [Текст]. - М.: Наука, 1982. – 316 с.
38. Методы компьютерной обработки изображений [Текст] / Под ред. В.А. Сойфера. — 2-е изд., испр. — М.: ФИЗМАТЛИТ, 2003. — 784 с
39. Комплексные и гиперкомплексные системы в задачах обработки многомерных сигналов [Текст] / Я.А. Фурман, А.В. Кревецкий, А.А. Роженцов и др. – М.: Физматлит, 2004. – 456 с.

40. Фурман, Я.А., Яншин, В.В., Юрьев, А.Н. Цифровые методы обработки и распознавания бинарных изображений [Текст]. - Красноярск: Изд-во Красноярского ун-та, 1992. – 248 с.
41. Введение в контурный анализ и его приложения к обработке изображений и сигналов [Текст] / Я.А. Фурман, А.В. Кревецкий, А.К. Передреев и др.; под ред. Я.А. Фурмана. – М.: ФИЗМАТЛИТ, 2002. - 592 с.
42. Введение в контурный анализ: Приложение к обработке изображений и сигналов [Текст] / Фурман Я.А. [и др.]; под ред. Я.А. Фурмана. - 2. изд., испр. - М.: ФИЗМАТЛИТ, 2003. - 588 с.
43. Алпатов, Б.А., Бабаян, П.В., Балашов, О.Е., Степашкин, А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление [Текст]. - М.: Радиотехника, 2008. - 176 с.
44. Обработка изображений и управление в системах автоматического сопровождения объектов [Текст] / Алпатов, Б.А. [и др.]. - Рязань: Изд-во РГРТУ. 2011. - 236с.
45. Алпатов, Б.А., Бабаян, П.В., Ершов, М.Д. Подходы к обнаружению и оценке параметров движущихся объектов на видеопоследовательности применительно к транспортной аналитике [Текст] // Компьютерная оптика. – 2020. - №5. - Том 44. – С. 746–756.
46. Визильтер, Ю.В., Горбачевич, В.С., Моисеенко, А.С. Одноэтапный детектор лиц и особых точек на цифровых изображениях [Текст] // Компьютерная оптика. – 2020. - №4. - Том 44. – С. 589–593.
47. Рубис, А.Ю., Лебедев, М.А., Визильтер, Ю.В., Выголов, О.В., Желтов, С.Ю. Компаративная фильтрация изображений с использованием монотонных морфологических операторов [Текст] // Компьютерная оптика. – 2018. - №2. - Том 42. – С. 306–311.
48. Лебедев, М.А., Рубис, А.Ю., Визильтер, Ю.В., Выголов, О.В. Выделение отличий на изображениях с помощью референтных EMD-фильтров [Текст] // Компьютерная оптика. – 2018. - №2. - Том 42. – С. 291–296.

49. Визильтер, Ю.В., Горбацевич, В.С., Вишняков, Б.В., Сидякин, С.В. Поиск объектов на изображении с использованием морфлетных описаний [Текст] // Компьютерная оптика. – 2017. - №3. - Том 41. – С. 406–411.
50. Визильтер, Ю.В., Горбацевич, В.С., Воротников, А.В., Костромов, Н.А. Идентификация лиц в реальном времени с использованием свёрточной нейронной сети и хэширующего леса [Текст] // Компьютерная оптика. – 2017. - №2. - Том 41. – С. 254–265.
51. Рубис, А.Ю., Лебедев, М.А., Визильтер, Ю.В., Выголов, О.В. Морфологическая фильтрация изображений на основе взаимного контрастирования [Текст] // Компьютерная оптика. – 2016. - №1. - Том 40. – С. 73–79.
52. Визильтер, Ю.В., Горбацевич, В.С., Рубис, А.Ю., Выголов, О.В. Сравнение изображений по форме с использованием диффузной морфологии и диффузной корреляции [Текст] // Компьютерная оптика. – 2015. - №2. - Том 39. – С. 265–274.
53. Визильтер, Ю.В., Горбацевич, В.С., Каратеев, С.Л., Костромов, Н.А. Обучение алгоритмов выделения кожи на цветных изображениях лиц [Текст] // Информатика и её применение. – 2012. - №1. - Том 6. – С. 108–113.
54. Каратеев, С.Л., Бекетова, И.В., Визильтер, Ю.В. и др. Автоматизированный контроль качества цифровых изображений для персональных документов [Текст] // Информатика и её применение. – 2010. - №1. - Том 4. – С. 65–73.
55. Васильев, В.И. Распознающие системы. Справочник [Текст]. – Киев: Наукова думка, 1983. – 422 с.
56. Приоров, А.Л., Ганин, А.Н., Хрящёв, В.В. Цифровая обработка изображений: Учебное пособие [Текст]. - Ярославль: Изд-во ЯрГУ, 2001. - 218 с.
57. Приоров, А.Л., Апальков, И.В., Хрящев, В.В. Цифровая обработка изображений: Учебное пособие [Текст]. - Ярославль: Изд-во ЯрГУ, 2007. - 235 с.

58. Самаль, Д.И. Алгоритмы идентификации человека по фотопортрету на основе геометрических преобразований [Текст]: автореф. дисс. ... канд. техн. наук: 05.13.01. / Самаль Дмитрий Иванович. — Минск, 2002. — 22 с.
59. Анисимов, Б.В. Распознавание и цифровая обработка изображений [Текст] / Анисимов Б.В., Курганов В.Д., Злобин В.К. — М.: Высш. шк., 1983. — 295 с.
60. Головкин, В.А. Нейроинтеллект: Теория и применения. Книга 1. Организация и обучение нейронных сетей с прямыми и обратными связями [Текст]. — Брест: Изд-во БПИ, 1999. — 260 с.
61. Дуда, Р., Харт, П. Распознавание образов и анализ сцен [Текст] / Пер. с англ.; под ред. В.Л. Стефанюка. — М.: Мир, 1976. — 512 с.
62. Винклер, Г. Анализ изображений, случайные поля и динамические методы Монте-Карло. Математические основы [Текст]. — Новосибирск: Изд-во СО РАН, 2002. - 343 с.
63. Гренандер, У. Лекции по теории образов. Том 1. Синтез образов [Текст] / В 3-х томах. Пер. с англ. — М.: Мир, 1979. — 383 с.
64. Журавлев, Ю.И., Рязанов, В.В., Сенько, О.В. Распознавание. Математические методы. Программная система. Практические применения [Текст]. — М: Фазис, 2005. - 159 с.
65. Катус, Г.П. Обработка визуальной информации [Текст]. — М.: Машиностроение, 1990. - 320 с.
66. Кононюк, А.Е. Общая теория распознавания. В 2 книгах. Книга 1: Начала [Текст]. — Киев: Освіта України, 2012. — 584 с.
67. Кононюк, А.Е. Общая теория распознавания. В 2 книгах. Книга 2: Математические средства описания распознаваемых объектов и распознающих процессов [Текст]. — Киев: Освіта України, 2012. — 588 с.
68. Кручинин, А.Ю. Оптимальный подход к распознаванию протяженных объектов в реальном времени [Текст]. — М.: Научное обозрение, 2016. — 305 с.

69. Лбов, Г.С., Бериков, В.Б. Устойчивость решающих функций в задачах распознавания образов и анализа разнотипной информации [Текст]. – Новосибирск: Изд-во Института математики, 2005. — 218 с.
70. Лепский, А.Е., Броневиц, А.Г. Математические методы распознавания образов [Текст]. – Таганрог: Изд-во ТТИ ЮФУ, 2009. - 155 с.
71. Марр, Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов [Текст] / Пер. с англ. — М.: Радио и связь, 1987. — 400 с.: ил.
72. Математические методы распознавания образов: 15-я Всероссийская конференция, г. Петрозаводск, 11-17 сентября 2011 г. [Текст] / Сборник докладов. Научное издание — М.: МАКС Пресс, 2011. — 618 с.
73. Ньейн, Э. Разработка и исследование программного обеспечения для распознавания и анализа объектов на изображении с помощью нейронной сети [Текст]: дисс. канд. техн. наук: 05.13.11 / Ньейн Эй. - Москва, 2007. – 104 с.
74. Панин, С.Д. Теория принятия решений и распознавание образов [Текст]. – М.: Изд-во МГТУ им. Н. Э. Баумана, 2017. — 240 с.
75. Веселова, С.В. Цифровая обработка изображений [Текст]. - СПб.: Изд-во СПбГИКиТ, 2018. — 349 с.
76. Васильев, К.К., Крашенинников, В.Р. Статистический анализ изображений [Текст] / Научное издание; Под. науч. ред. докт. техн. наук, проф. К.К. Васильева. - Ульяновск: Изд-во УлГТУ, 2014. — 214 с.: ил.
77. Визильтер, Ю.В., Желтов, С.Ю., Бондаренко, А.В. и др. Обработка и анализ изображений в задачах машинного зрения [Текст]. – М.: Физматкнига, 2010. - 672 с.
78. Гришенцев, А.Ю., Коробейников, А.Г. Методы и модели цифровой обработки изображений [Текст]. – СПб.: Изд-во Политехн. ун-та, 2014. — 191 с.
79. Verschae, R., Ruiz-del-Solar, J., Correa, M. Face Recognition in Unconstrained Environments: A Comparative Study // Workshop on Faces in Real-Life Images: Detection, Alignment, and Recognition [Текст]. — Marseille, 2008.

80. Image Analysis, Sediments and Paleoenvironments / Edited by P. Francus [Текст]. — N.Y.: Springer, 2005. — ISBN 1-4020-2061-9.
81. Beymer, D., Poggio, T. Face recognition from one example view [Текст] // Proceedings of the IEEE International Conference on Computer Vision, 1995. — P. 500—507.
82. Bhat, G., Achary, K.K. Face Recognition using FSS-DSOP for Small Sample Size Problem with Illumination Variations [Текст] // Soft Comput. Appl, 2009. — Vol. 1. — No. 2. — P. 105—118. — ISSN 2074-8523.
83. Blanz, V., Romdhani, S., Vetter, T. Face identification across different poses and illuminations with a 3D morphable model [Текст] // International Conference on Automatic Face and Gesture Recognition, 2002. — P. 202—207.
84. Choi, S., Kim, C., Choi, C. Shadow compensation in 2D images for face recognition [Текст] // Pattern Recognition, 2007. — No. 40. — P. 2118—2125.
85. Face recognition as a search tool "Foto-Fahndung". Final Report [Текст]. — Wiesbaden, 2007. — 30 p.
86. Fractal Image Compression. Theory and Application [Текст] / Edited by Y. Fisher — N.Y.: Springer, 1995. — ISBN 0-387-29211-4.
87. Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision [Текст]. Proc. of Imaging Understanding Workshop, 1981, pp. 121-130.
88. Алфимцев, А.Н., Лычков, И.И. Метод обнаружения объекта в видеопотоке в реальном времени [Текст] // Вестник ТГТУ. — 2011. — Т. 17. - №1. — С. 44-54.
89. Алфимцев, А.Н., Лычков, И.И., Девятков, В.В. Метод отслеживания транспортных средств в видеопотоке [Текст] // Вестник Иркутского ГТУ. - 2012. - Т. 60. - № 1. - С. 79 – 85.
90. Локтев, Д.А., Алфимцев, А.Н. Измерение расстояния до движущегося объекта с помощью комплексной системы видеомониторинга [Электронный ресурс] // Инженерный журнал: наука и инновации. — 2013. - № 11(23). - Режим

доступа: <http://engjournal.ru/cata...g/it/hidden/996.html>. Дата обращения: 15.03.2021г.

91. Алфимцев, А.Н., Демин, А.И. Захват и отслеживание удаленных объектов в видеопотоке [Электронный ресурс] // Инженерный журнал: наука и инновации. – 2013. - № 11(23). - Режим доступа: <http://engjournal.ru/authors/1545.html>. Дата обращения: 15.03.2021г.

92. Девятков, В.В., Алфимцев, А.Н., Таранян, А.Р. Селективно-ковариационный метод локализации, классификации и отслеживания людей в видеопотоках от множества видеокамер [Текст] // Вестник МГТУ им.Н.Э.Баумана. Серия "Приборостроение". – 2016. - № 6. - С. 54 – 70.

93. Локтев, А.А., Алфимцев, А.Н., Локтев, Д.А. Алгоритм размещения видеокамер, и его программная реализация [Текст] // Вестник МГСУ. – 2012. - № 5. - С. 167 – 176.

94. Локтев, А.А., Алфимцев, А.Н., Локтев, Д.А. Алгоритм распознавания объектов [Текст] // Вестник МГСУ. – 2012. - № 5. - С. 194 – 202.

95. Алфимцев, А.Н., Кукин, П.А. Автоматизация визуальной оценки городского пейзажа методами компьютерного зрения [Текст] // Автоматизация. Современные технологии. – 2016. - № 7. - С. 28 – 33.

96. Сакулин, С.А., Алфимцев, А.Н., Локтев, Д.А. и др. Защита изображения человека от распознавания нейросетевой системой на основе состязательных примеров [Текст] // Вестник компьютерных и информационных технологий. – 2020. - Т. 17. - № 2 (188). - С. 32 – 38.

97. Оптико-электронные устройства обработки и распознавания изображений [Текст] / В.С. Титов, М.И. Труфанов [и др.]. – Тула: Изд-во ТГУ, 2008. – 121 с.

98. Файн, В.С. Опознавание изображений (основы непрерывно-групповой теории и её приложения) [Текст]. – М.: Наука, 1970. - 299 с.

99. Коледов, Л.А. Технология и конструкции микросхем, микропроцессоров и микросборок [Текст]. – СПб.: Лань, 2007. – 400 с.: ил.

100. Крылов, В.Н. Совмещение изображений в системах оптического контроля печатных плат [Текст] / В.Н. Крылов, Г.Ю. Щербакова // Технология и конструирование в электронной аппаратуре. – 2000. – №1. – С. 23-25.
101. Сидоров, А. Особенности реперных знаков для систем автоматического совмещения [Текст] / А. Сидоров // Электронные компоненты. – 2011. – №9. – С. 18-23.
102. Татарченко, Н.В., Тимошенко, С.В. Биометрическая идентификация в интегрированных системах безопасности [Электронный ресурс]. – Режим доступа: http://www.vashdom.ru/articles/st_15.htm. Дата обращения: 01.05.2021г.
103. Звягин, М.Ю. Распознавание людей по изображению лица с использованием текстурных характеристик [Текст] / М.Ю. Звягин, В.Г. Прокошев, О.А. Новикова, П.Ю. Шамин // Материалы XVII Всероссийской научно—методической конференции "Телематика 2010". — СПб., 2010. — С. 257—259.
104. Bronstein, A. M. Three-dimensional face recognition [Текст] / Bronstein A. M., Bronstein M. M., Kimmel R. // International Journal of Computer Vision (IJCV), 2005. — № 1 (64). — P. 5—30.
105. Brunelli, R., Poggio, T. Face recognition: features versus templates [Текст] // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993. — Vol. 15. — No 10. — P. 235—241.
106. Chen, S. Making FLDA applicable to face recognition with one sample per person [Текст] / S. Chen, J. Liu, Z. Zhou // Pattern Recognition, 2004. — Vol. 37. — No. 7. — P. 1553—1555.
107. Каменская, Е.И. Алгоритмы обработки и представления изображений лиц в пространстве канонических переменных [Текст]: автореф. дисс. ... канд. техн. наук: 05.13.18 / Каменская Екатерина Ивановна. — СПб., 2010. — 18 с.
108. Разработка технических средств распознавания образов по анализу видеоизображения для автоматизации работы оператора: пояснительная записка [Текст] / Владимирский государственный университет; рук. Аракелян С.М.; исп. Голубев А.С. [и др.] — Владимир, 2010. — 262 с. — ВлГУ 425718.001 ПЗ.

109. Рожков, М.М. Актуальность 2D алгоритмов в определенных задачах автоматического распознавания человека [Текст] / М.М. Рожков // Материалы научной конференции "Технические науки: проблемы и перспективы" — СПб, 2011. — С. 145—146. — ISSN 2072-0297.

110. Рожков, М.М. Использование текстурных карт Лавса и дискретного косинусного преобразования в задаче распознавания лиц [Текст] / М.М. Рожков // Прикладная информатика. — 2011. — №1 (31). — С. 98—103. — ISSN 1993-8314.

111. Рожков, М.М. К вопросу использования текстурных карт Лавса для распознавания лиц в рамках проблемы одного эталонного изображения [Текст] / М.Ю. Звягин, М.М. Рожков, П.Ю. Шамин // Труды Владимирского государственного университета. — Владимир, 2010. — Выпуск 7, Физико-математические основы индустрии наносистем и материалов. — с. 90—92. — ISBN 978-5-9984-0115-2.

112. Рожков, М.М. Некоторые вопросы распознавания лиц в ситуации одного эталонного изображения [Текст] / Рожков М.М., Звягин М.Ю., Шамин П.Ю. // Материалы II-ой международной научно-практической конференции студентов, аспирантов и молодых ученых "Прогрессивные технологии и перспективы развития". — Тамбов, 2010. — С. 53—55. — ISBN 978-5-905039-03-4.

113. Рожков, М.М. Построение подпространств атрибутов на базе одного эталона для обеспечения устойчивости работы в перспективных системах автоматического распознавания лиц [Текст] / Прокошев В.Г., Рожков М.М., Шамин П.Ю., Голубев А.С. // Прикладная информатика. — М, 2011. — №2 (32). — С. 100—107. — ISSN 1993-8314.

114. Рожков, М.М. Проблема автоматического распознавания лиц с одним эталонным изображением [Текст] / В.Г. Прокошев, М.М. Рожков, П.Ю. Шамин // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Серия "Информатика. Телекоммуникации. Управление". — СПб., 2010. — №5. — С. 13—18.

115. Рожков, М.М. Программа тестирования и обработки алгоритмов распознавания [Текст]: свидетельство о государственной регистрации программ

для ЭВМ № 2011611918 / С.М. Аракелян, М.Ю. Звягин, М.М. Рожков и др. - № 2010617338; заявл. 22.11.2010; зарегистр. 28.02.2011. Специальное прикладное программное обеспечение: пояснительная записка [Текст] / Владимирский государственный университет; рук. Аракелян С.М.; исп. Шамин П.Ю. [и др.]. — Владимир, 2010. — 241 с. — № 643.02068048 00628-01 81 01.

116. Лоханов, А.В. Исследование искаженных изображений при помощи инструментов текстурного анализа [Текст] / Д.Г. Васильченкова, А.С. Голубев, М.Ю. Звягин, А.В. Лоханов // Проектирование и технология электронных средств. — 2017. — №3. — С. 41-47.

117. Лоханов, А.В. Усовершенствованное распознавание образов на основе квази-эталонных образов [Текст] / А.С. Голубев, М.Ю. Звягин, А.В. Лоханов, Л.А. Семин // Известия высших учебных заведений. Технология текстильной промышленности. - 2015. - № 4 (358). - С.236-240.

118. Лоханов, А.В. Полуавтоматическая аппроксимация экспериментальных данных базисным набором гауссиан [Текст]: свидетельство о государственной регистрации программ для ЭВМ № 2017611477 / А.А. Куприянов, А.С. Голубев, М.Ю. Звягин, А.В. Лоханов. - № 2016663261; заявл. 05.12.2016; зарегистр. 03.02.2017.

119. Лоханов, А.В. Программное обеспечение автоматизированного тестирования алгоритмов распознавания лиц [Текст]: свидетельство о государственной регистрации программ для ЭВМ № 2014612123 / А.С. Голубев, И.В. Курочкин, А.В. Лоханов, П.Ю. Шамин. - № 2013619560; заявл. 22.10.2013; зарегистр. 19.02.2014.

120. Лоханов, А.В. Модульная программа идентификации человека по растровому двумерному изображению лица [Текст]: свидетельство о государственной регистрации программ для ЭВМ № 2011611917 / А.С. Голубев, А.В. Лоханов, В.Г. Прокошев и др. - № 2010617337; заявл. 22.11.2010; зарегистр. 28.02.2011.

121. Лоханов, А.В. База данных биометрических признаков [Текст]: свидетельство о государственной регистрации программ для ЭВМ № 2011620172 /

А.С. Голубев, А.В. Духанов, Д.С. Квасов, А.В. Лоханов - № 2010620768; заявл. 22.11.2010; зарегистр. 28.02.2011.

122. Лоханов, А. В. Текстуальный анализ искаженных изображений [Текст] / А.В Лоханов // XX международная научно-практическая конференция «Теоретические и методологические проблемы современных наук». Сб. тез. докладов. – Новосибирск. – 2017. – С.78-84.

123. Лоханов, А.В. Численно-аналитические исследования в фазовом пространстве динамической системы [Текст] / М.Ю. Звягин, А.В Лоханов, В.А. Складенко // Сборник трудов Международной конференции по математической теории управления и механике. Сб. тез докладов. - Суздаль, 03-07.07.2015г. – С.74-81.

124. Лоханов, А.В. Сравнительный анализ перспектив использования источников видеоизображения с компрессией и без компрессии в технических системах с функциями распознавания людей [Текст] / А.В. Лоханов, П.Ю. Шамин // Труды XVIII Всероссийской научно-методической конференции «Телематика'2011». 20-23.06.2011, Санкт-Петербург, Россия – СПб.: Изд-во Санкт-Петербургского государственного университета информационных технологий, механики и оптики. - 2011. - Т. 2. - С 264-266.

125. Лоханов, А.В. Поиск лиц людей на изображениях высокого разрешения в реальном времени [Текст] / И.И. Зиновьев, А.В. Лоханов // Региональная научно-практическая конференции «Многоядерные процессоры и параллельное программирование». - 25.02 2011г.; Барнаул. Сб. докладов - Барнаул: Изд-во БГУ, 2011. – С. 71-77.

126. Лоханов, А.В. Компьютерное моделирование Mesh сети в системах GPSS и MATLAB [Текст] / А.В. Лоханов, А.В. Осин, М.В. Руфицкий // Девятая международная научно практическая конференция "Исследование, разработка и применение высоких технологий в промышленности". - 22-23.04.2010г. - СПб.: Изд-во Политехн. университета. - 2010. - Т. 1. – С.165-166.

127. Лоханов, А.В. Аппаратно-программный комплекс автоматизированного распознавания и биометрической идентификации людей

(АПК АР-БИЛ) [Текст] / С.М. Аракелян, А.В. Лоханов, В.Г. Прокошев и др. // Труды XVII Всероссийской научно-методической Конференции "Телематика'2010". 21-24.06.2010, Санкт-Петербург, Россия – СПб.: Изд-во Санкт-Петербургского государственного университета информационных технологий, механики и оптики. - 2010. - Т. 2. - С 255-256.

128. Пат. по заявке 2004123248 РФ, МПК G06K 9/80, G06T 7/60, A61B 5/117. Система и способ отслеживания объекта [Текст] / Мурынин А.Б., Базанов П.В, Кузнецов В.Д. - № 2004123248/09; заявл. 29.07.2004; опубл. 20.06.2009, Бюл. № 17. – 3 с.: ил.

129. Степин, В.С., Савушкин, А.В., Зотов, А.Б. Криминалистическое отождествление человека по разноракурсным портретам [Текст]. - М.: Изд-во ЭКЦ МВД РФ, 1992. – 23 с.: ил.

130. Стокман, Д. Компьютерное зрение [Текст] / Д. Стокман, Л. Шапиро. Пер. с англ. — М: БИНОМ. Лаборатория знаний, 2010. — 752 с. — ISBN 5-94774-384-1. — ISSN 0-13-030796-3.

131. Тартаковский, А.Г. Последовательные методы в теории информационных систем [Текст]. – М.: Радио и связь, 1991. — 280 с.

132. Форсайт, Д.А., Понс, Ж. Компьютерное зрение. Современный подход [Текст] / Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 928 с.: ил. – ISBN 5-8459-0542-7 (рус.), ISBN 0-13-085198-1 (англ.).

133. Шеридан, Т.Б., Феррелл, У.Р. Системы человек-машина [Текст]. – М.: Машиностроение, 1980. – 398 с.

134. Пат. по заявке: 2005132003 РФ, МПК H04B 7/00, G08C 17/02, G06K 9/00. Терминал для передачи трехмерных изображений в биометрических системах (варианты) [Текст] // Мосиенко С.А. - № 2005132003/22, заявл. 06.06.2005, опубл. 10.11.2005, Бюл. № 31. – 10 с.: ил.

135. Ту, Дж., Гонсалес, Р. Принципы распознавания образов [Текст] / Пер. с англ.; под ред. Ю.И. Журавлева. – М.: Мир, 1978. – 414 с.

136. Сосулин, Ю.Г., Фишман, М.М. Теория последовательных решений и её применение [Текст]. – М.: Радио и связь, 1985. — 272 с.

137. Пат. по заявке 2001121986 РФ, МПК G06K 9/78, G06T 7/00. Система распознавания объектов и слежения за ними [Текст] / Коуэн П. Д., - №2001121986/09; заявл. 01.02.2000, опубл. 20.07.2003, Бюл. №. 20. – 12 с.: ил.

138. Сосулин Ю.Г. Последовательное обнаружение сигналов: проблемы и перспективы. [Текст] – М.: Радиотехника, 1998 г., № 10. - С. 1-6.

139. Брилюк, Д.В. Распознавание человека по изображению лица нейросетевыми методами. [Текст] / Д.В. Брилюк, В.В. Старовойтов. — Препринт. — Минск: Изд-во Ин-та техн. кибернетики НАН Беларуси, 2002. — 54 с.

140. Патрик, Э. Основы теории распознавания образов [Текст] / Пер. с англ. Под ред. Б.Р. Левина. – М.: Сов. радио, 1980. – 408 с., ил. / Пер. изд.: США, 1972.

141. Плотникова, И.Н. Анализ погрешностей приближений для вероятностей решений на этапе последовательной процедуры проверки многих гипотез [Текст] // Теория и техника радиосвязи. – Воронеж, 2009. – Вып.3. - С.26-29.

142. Пат. по заявке 2001122361 РФ, МПК G06T 11/00. Способ представления человека [Текст] / Свириденко А.В. - №2001122361/09, заявл. 10.08.2001, опубл. 27.01.2005, Бюл. №. 3. – 1 с.: ил.

143. Поляков, А.Ю. Методы и алгоритмы компьютерной графики в примерах на Visual C++ [Текст] / 2-е изд., перераб. и доп. Поляков А.Ю., Брусенцев В.А. — СПб.: БХВ, 2003. — 560 с. — ISBN 5-94157-377-4.

144. Асаи, К. Прикладные нечёткие системы [Текст] / Пер. с япон.; К. Асаи., Д. Ватада, С. Иваи и др.: под ред. Т. Тэрано, К. Асаи, М. Сугэно. – М.: Мир, 1993. – 368 с., ил.

145. Рогозин, А.А., Поветко, В.Н. Восстановление оптических и геометрических свойств объектов в трехмерном пространстве по их двумерным проекциям. [Текст] // Телекоммуникации. – 2002. - №11. – С.28-37.

146. Головкин, В.А. Нейроинтеллект: Теория и применения. Книга 1. Организация и обучение нейронных сетей с прямыми и обратными связями [Текст]. — Брест: Изд-во БПИ, 1999. — 260с.

147. Биометрия. Изображение лица. [Электронный ресурс]. - Режим доступа: http://wiki.oszone.net/index.php/%D0%91%D0%B8%D0%BE%D0%BC%D0%B5%D1%82%D1%80%D0%B8%D1%8F._%D0%98%D0%B7%D0%BE%D0%B1%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5_%D0%BB%D0%B8%D1%86%D0%B0 Дата обращения: 14. 03.2021г.

148. Введение в контурный анализ. Приложения к обработке изображений и сигналов [Текст] / Под ред. Я.А. Фурмана. — 2-е изд., испр. — М.: ФИЗМАТЛИТ, 2003. — 592 с. — ISBN 5-9221-0374-1.

149. Морзеев, Ю. Зачем компьютеру зрение / Журнал «Компьютер-Пресс». - 2002. - №8. - [Электронный ресурс]. - Режим доступа: <http://www.compress.ru/article.aspx?id=11585&iid=453>. Дата обращения: 14. 03.2021г.

150. Анисимов, Б.В. Распознавание и цифровая обработка изображений: Учеб. пособие для студентов вузов. [Текст] / Анисимов Б.В., Курганов В.Д., Злобин В.К. — М.: Высш. шк., 1983. — 295 с.

151. 3D технология распознавания лиц [Электронный ресурс]. - Режим доступа: http://www.sec-group.ru/index.php?show_aux_page=23. Дата обращения: 14. 03.2021г.

152. Манолов, А.И. Некооперативная биометрическая идентификация по 3d-моделям лица с использованием видеокамер высокого разрешения [Текст] / А.И. Манолов, А.Ю. Соколов, О.В. Степаненко и др. // Труды 19 Международной конференции по компьютерной графике и зрению ГрафиКон'2009. — М., 2009. - С. 12-18.

153. Люгер, Джордж Ф. Искусственный интеллект: стратегии и методы решения сложных проблем [Текст] — М.: Вильямс, 2005. — 864 с. ISBN 5-8459-0437-4, 0-201-64866-0.

154. Пат. по заявке 2005100267 РФ, МПК G06T 7/00, G06K 9/78, G10L 15/00, G10L 17/00. Способ и система автоматической проверки присутствия лица живого человека в биометрических системах безопасности [Текст] / Мун В.Д. (KR),

Мурынин А.Б. (RU), Базанов П.В. (RU) и др.- № 2005100267/09; заявл. 12.01.2005, опубл. 20.06.2006, Бюл. №. 3. – 8 с.: ил.

155. Пат. по заявке 95115528 РФ, МПК G06K 9/00, G07C 9/00. Способ распознавания личности и система для его осуществления [Текст] / Серебрянников О.А., Мурынин А.Б., Кулаков В.В., Большаков С.А. - № 2005100267/09; заявл. 08.09.1995, опубл. 20.10.1997. – 8 с.: ил.

156. Уоссермен, Ф. Нейрокомпьютерная техника: Теория и практика [Текст] / Уоссермен Ф. — М.: Мир, 1992. — 184 с.

157. Фурман, Я. А., Юрьев, А.Н., Яншин, В.В. Цифровые методы обработки и распознавания бинарных изображений [Текст]. – Красноярск: Изд-во Краснояр. ун-та, 1992. – 248 с.

158. Фу, К. Структурные методы в распознавании образов [Текст] / Пер. с англ.; под ред. М.А. Айзермана. – М.: Мир, 1977. – 320 с.

159. Стокман, Д. Компьютерное зрение [Текст] / Пер. с англ.; Стокман, Д. Шапиро Л. — М.: БИНОМ. Лаборатория знаний, 2006. – 761 с. — ISBN 0-13-030796-3, 5-94774-384-1.

160. Степин, В.С., Савушкин, А.В., Зотов, А.Б. Криминалистическое отождествление человека по разнокурсным портретам [Текст]. – М.: ИНФРА-М -НОРМА, 2009. – 261 с.

161. Бойко, И.А. Распознавание объектов на основе видеосигнала, полученного с камеры, установленной на подвижной платформе [Текст] / И.А. Бойко, Р.А. Гурьянов // Молодой ученый. – 2013. – №6. – С. 34-36.

162. Шувалова, И.В., Литвинская, О.С. Выбор признаков изображения для слежения за объектом [Текст] // Современные наукоемкие технологии. – 2014. – № 11. – С. 67-68. – Режим доступа: URL: <http://top-technologies.ru/ru/article/view?id=34796>. Дата обращения: 06.05.2021.

163. Shi, J., Tomasi, C. Good features to track [Текст] // IEEE Conference on Computer Vision and Pattern Recognition, 1994. – P. 593–600.

164. Bouguet, J.-Y. Pyramidal implementation of the Lucas-Kanade feature tracker [Текст] // Open CV documentation, Intel Corporation, Microprocessor Research Labs, 1999.
165. Horn Berthold K.P., Schunck Brian G. Determining Optical Flow [Текст] // Artificial Intelligence, 1981. – Vol. 17. – P. 185–203.
166. Lucas, B.D., Kanade, T. An iterative image registration technique with an application to stereo vision [Текст] // 7th International Joint Conference on Artificial Intelligence, 1981. – P. 674–679.
167. Bruhn, A, Weickert, J. Lucas. Kanade Meets Horn/Schunck: Combining Local Global Optic Flow Methods [Текст] // International Journal of Computer Vision, 2005. – Vol. 61. – №. 3. – P. 211–231.
168. Viola, P., Jones, M.J. Rapid Object Detection using a Boosted Cascade of Simple Features [Текст] / Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001.
169. Viola, P., Jones, M.J. Robust real-time face detection [Текст] // International Journal of Computer Vision, vol. 57, no. 2, 2004, pp.137–154.
170. Гонсалес, Р., Вудс, Р. Цифровая обработка изображений [Текст]. – М.: Техносфера, 2005. – 1072 с. - ISBN 5-94836-028-8.
171. Местецкий, Л.М. Математические методы распознавания образов [Текст]. – М.: Изд-во МГУ - ВМиК, 2002–2004. - 96 с.
172. Jan Šochman, Jiří Matas, «AdaBoost», Center for Machine Perception [Текст] / Czech Technical University, Prague, 2010.
173. Yoav Freund, Robert E. Schapire, «A Short Introduction to Boosting» [Текст] / Shannon Laboratory, USA, 1999. - pp. 771-780.
174. Вапник, В.Н. Теория распознавания образов. Статистические проблемы обучения [Текст] / В.Н. Вапник, А.Я. Червоненкис. — М.: Наука, 1974. — 416 с.
175. Василенко, О.Н., Теоретико-числовые алгоритмы в криптографии [Текст] / О.Н. Василенко. — М.: МЦНМО, 2003. — 328 с.

176. Зинин, А.М. Габитоскопия и портретная экспертиза [Текст]. - М.: Изд-во Московск. Акад. МВД России, 2002. - 157 с.
177. Зинин, А.М., Овсянникова, М.Н. Криминалистическая габитоскопия [Текст]. - М., 2000. - 257 с.
178. Красильников, Н.Н. Теория передачи и восприятия изображений [Текст]. – М.: Радио и связь. 1986. - 283 с.
179. Моржаков, В., Мальцев, А. Современные биометрические методы идентификации / Архив журнала БДИ №2, 2009. [Электронный ресурс]. – Режим доступа: URL:<http://www.bdi.spb.ru/arch/?id=37&a=916> Дата обращения: 06.04.2021.
180. Вежневек, В.П. Алгоритмы анализа изображения лица человека для построения интерфейса человек—компьютер: автореф. дис. ... канд. физ.-мат. наук: 05.13.11 / Вежневек Владимир Петрович. — М., 2004. — 24 с.
181. Журавлев, Ю.И. Об алгебраических методах в задачах распознавания и классификации [Текст] / Ю.И. Журавлев // Ежегодник "Распознавание. Классификация. Прогноз". Математические методы и их применение. Выпуск 1. — М.: Наука, 1989. — С. 9—16.
182. Прэтт, У. Цифровая обработка изображений [Текст] / Пер. с англ. — М.: Мир, 1982. — Кн. 1. — 312 с.
183. Житенев, С.А., Плотникова, И.Н., Поветко, В.Н. Оценка зависимостей вероятностей решений на этапе последовательной процедуры проверки многих гипотез от ее параметров [Текст]. – Девятая междунар. научно-технич. конфер.: Кибернетика и высокие технологии XXI века. 13 – 15 мая 2008 г., Воронеж, Россия. Труды конференции, Том 1, с.267 – 276.
184. Нгуен, В.Х. Методы биометрической идентификации личности по внешним статическим образам - Общая структура систем [Электронный ресурс]. – МФТИ, Май 2006 – Режим доступа: URL: http://www.re.mipt.ru/infsec/2006/essay/2006_Biometric_authentication_Nguyen_Viet_Hung.pdf Дата обращения: 06.05.2021.

185. Житенев, С.А., Змаева, С.А. Оптимизация и анализ последовательной процедуры проверки многих гипотез при нелинейных затратах ресурса системы на получение данных. Теория и техника радиосвязи [Текст]. – Воронеж, 2009. – Вып.4. с. 67 – 73.

Приложение А
Свидетельства о государственной регистрации программ для ЭВМ
(справочное)

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ
№ 2017611477

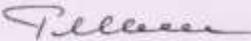
**Полуавтоматическая аппроксимация экспериментальных
данных базисным набором гауссиан**

Правообладатель: *Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Владимирский государственный университет имени Александра
Григорьевича и Николая Григорьевича Столетовых» (RU)*

Авторы: *Голубев Андрей Сергеевич (RU), Лоханов Александр
Васильевич (RU), Звягин Михаил Юрьевич (RU), Куприянов Артем
Андреевич (RU)*

Заявка № **2016663261**
Дата поступления **05 декабря 2016 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **03 февраля 2017 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

 **Г.П. Исхов**



РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2014612123

"Программное обеспечение автоматизированного тестирования алгоритмов распознавания лиц"

Правообладатель: *Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования "Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых" (ВлГУ) (RU)*

Авторы: *Курочкин Иван Владимирович (RU), Шамин Павел Юрьевич (RU), Голубев Андрей Сергеевич (RU), Лоханов Александр Васильевич (RU)*

Заявка № 2013619560

Дата поступления 22 октября 2013 г.

Дата государственной регистрации
в Реестре программ для ЭВМ 19 февраля 2014 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Б.П. Симонов

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2011611917

Модульная программа идентификации человека
по растровому двумерному изображению лица

Правообладатель(ли): *Государственное образовательное
учреждение высшего профессионального образования
«Владимирский государственный университет» (RU)*

Автор(ы): *Прокошев Валерий Григорьевич,
Звягин Михаил Юрьевич, Лоханов Александр Васильевич,
Шамин Павел Юрьевич, Голубев Андрей Сергеевич,
Рожков Максим Михайлович, Боченина Клавдия Олеговна (RU)*

Заявка № 2010617337

Дата поступления 22 ноября 2010 г.

Зарегистрировано в Реестре программ для ЭВМ
28 февраля 2011 г.



Руководитель Федеральной службы по интеллектуальной
собственности, патентам и товарным знакам

Б.П. Симонов

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации базы данных
№ 2011620172

База данных биометрических признаков

Правообладатель(ли): *Государственное образовательное учреждение высшего профессионального образования «Владимирский государственный университет» (RU)*

Автор(ы): *Прокошев Валерий Григорьевич, Квасов Дмитрий Сергеевич, Духанов Алексей Валентинович, Шамин Павел Юрьевич, Голубев Андрей Сергеевич, Лоханов Александр Васильевич (RU)*

Заявка № 2010620678

Дата поступления 22 ноября 2010 г.

Зарегистрировано в Реестре баз данных
28 февраля 2011 г.



Руководитель Федеральной службы по интеллектуальной собственности, патентам и товарным знакам

Б.П. Симонов

Приложение Б

Акт о внедрении результатов кандидатской диссертации в учебный процесс (справочное)

«УТВЕРЖДАЮ»

Проректор по образовательной деятельности
Владимирского государственного университета
имени Александра Григорьевича и Николая
Григорьевича Столетовых, к.т.н., доцент



А.А. Панфилов

« 21 » апреля 2021 г.

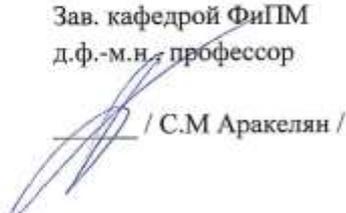
АКТ

о внедрении результатов кандидатской диссертации Лоханова А.В. в
учебный процесс Института прикладной математики, физики и
информатики ВлГУ

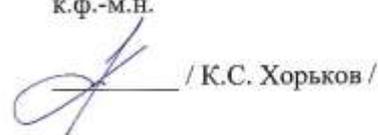
Результаты диссертационных исследований Лоханова А.В., выполненных по теме «Алгоритмы распознавания и модели цифровой обработки динамических телевизионных изображений», в части изучения и применения разработанных алгоритмов и методик цифровой обработки а также распознавания образов, внедрены в университете и использованы в учебном процессе при подготовке бакалавров и магистров по направлениям «Лазерная техника и лазерные технологии», «Прикладная математика и информатика», «Нанотехнологии и микросистемная техника» в рамках проведения учебных занятий по дисциплинам «Оптоэлектроника», «Проектирование электронных средств в нанoeлектронике», «Прикладное программирование» и «Инструментальные средства разработки программного обеспечения».

Результаты работы использованы при проведении лекционных, лабораторных, практических и семинарских занятий, а также в научно-исследовательской работе студентов, в курсовом и дипломном проектировании.

Зав. кафедрой ФиГМ
д.ф.-м.н., профессор

 / С.М. Аракелян /

Директор ИПМФИ ВлГУ
к.ф.-м.н.

 / К.С. Хорьков /

Приложение В

Акт о практическом использовании результатов диссертационных исследований в инновационной научной деятельности (справочное)

«УТВЕРЖДАЮ»

Проректор по научной и инновационной
работе Владимирского государственного
университета имени Александра
Григорьевича и Николая Григорьевича
Столетовых, д.т.н., профессор



А.В. Федин
« 21 » 2021 г.

АКТ

о практическом использовании результатов диссертационных исследований
Лоханова А.В. в инновационной научной деятельности Владимирского
государственного университета имени Александра Григорьевича и Николая
Григорьевича Столетовых (ВлГУ).

Научно-технические результаты диссертационных исследований Лоханова А.В., выполненных по теме «Алгоритмы распознавания и модели цифровой обработки динамических телевизионных изображений», включены в итоговый отчет ВлГУ по проблеме: «Разработка технических средств распознавания образов по анализу видеозображения для автоматизации работы оператора» по Государственному контракту с Минобрнауки РФ от 17 ноября 2009 г. № ГК №02.524.11.4010 в рамках реализации Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России» на 2007 – 2013 годы, а также используются в научно-исследовательской работе в рамках Государственного задания ВлГУ по темам «Моделирование и прогнозирование многопараметрических задач и многофакторных процессов на основе новых физических принципов и математических методов нелинейной динамики с использованием систем искусственного интеллекта, параллельных вычислений и обработки видеoinформации» и «Новые физические методы лазерного синтеза микро- и наноструктурированных углеродосодержащих материалов: экспериментальное и теоретическое изучение процессов их формирования и распознавания наноструктур».

Достигнутые Лохановым А.В. научно-технические результаты явились основанием для заключения Договора 1769ГУ1/2014 от 21.03.2014 «Разработка алгоритмов биометрической идентификации человека в условиях единичного низкокачественного эталона» с Федеральным государственным бюджетным

учреждением «Фонд содействия развитию малых форм предприятий в научно-технической сфере» (Фонд содействия инновациям) и успешного выполнения работ по Программе «УМНИК-2014».

Разработанная Лохановым А.В. и научно обоснованная система распознавания образов в условиях единичного низкокачественного эталона, предложенные алгоритмы, методики, а также их программная реализация подтвердили свою работоспособность и высокие качественные показатели в ходе проведения сравнительных испытаний и конкурсов проводимых для тестирования различных систем распознавания образов.

Результаты работы Лоханова А.В. используются в научно-практической деятельности магистрантов и аспирантов базовых кафедр ВлГУ, функционирующих на промышленных предприятиях Владимирской области.

Начальник управления научно-исследовательской
деятельности и развития информационных технологий,
д.ф.-м.н., доцент



А.О. Кучерик

Заведующий кафедрой
Физики и прикладной математики,
д.ф.-м.н., профессор



С.М. Аракелян

Приложение Г
Дипломы (справочное)



 **ФОНД СОДЕЙСТВИЯ РАЗВИТИЮ**
малых форм предприятий в научно-технической сфере

ДИПЛОМ

Победитель программы “Участник молодежного
научно-инновационного конкурса” (“УМНИК”)
ЛОХАНОВ

АЛЕКСАНДР

ВАСИЛЬЕВИЧ

*Председатель
Наблюдательного совета*

И.М. Бортник

*Генеральный директор
Фонда содействия развитию
малых форм предприятий
в научно-технической сфере*

С.Г. Поляков

Приложение Д
Акты внедрения (справочное)

АКТ

об использовании результатов кандидатской диссертационной
работы Лоханова Александра Васильевича
«Алгоритмы распознавания и модели цифровой обработки
динамических телевизионных изображений»

Алгоритмы и модели распознавания образов, предложенные Лохановым А. В., используются в научно-исследовательских и опытно-конструкторских работах, проводимых ФКП «Государственный лазерный полигон «Радуга» при коррекции регистрируемых изображений, искаженных турбулентной атмосферой.

Разработанные алгоритмы и модели являются результатом диссертационных исследований Лоханова А. В. и обеспечивают существенное повышение достоверности распознавания объектов, находящихся в турбулентном атмосферном потоке, и регистрируемых изображений, характеризующихся турбулентными искажениями.

Врио заместителя генерального директора
по НИИР ФКП «ГЛП «Радуга»
кандидат физико-математических наук



С.Л. Лысенко
25.12.2020

УТВЕРЖДАЮ


 Генеральный директор
 ООО «РУСОКСИД»
 _____ Е. Е. Алясова
 11 марта 2021 г.

АКТ

о практическом использовании результатов диссертационной научной работы Лоханова Александра Васильевича «Алгоритмы распознавания и модели цифровой обработки динамических телевизионных изображений»

Комиссия в составе ведущего технолога В.В. Кривовой и инженера по силовой электронике А. А. Гаджиева подтверждает, что разработанные Лохановым А. В. алгоритмы и модели цифровой обработки динамических изображений рекомендованы к использованию в научно-производственных работах, проводимых в ООО «РУСОКСИД».

Предложенные в диссертации Лоханова А. В. математические модели и программные средства планируются к использованию в системах технического зрения при контроле качества применения электрохимической алюмооксидной технологии.

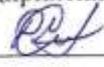
Предложенные Лохановым А.В. алгоритмы и модели распознавания образов отличаются от известных аналогов повышенным быстродействием, достоверностью результатов оценки качества и состояния изделий, способствуют автоматизации контрольных и измерительных методик на производстве.

Ведущий технолог
 Инженер по силовой электронике


 В. В. Кривова

 А. А. Гаджиев

УТВЕРЖДАЮ

Генеральный директор
предприятия ООО «Техника и технология»


 Синавский Д.В.
 “ 17 ” мая 2021 г.

Заключение

о практической реализации результатов диссертационной работы Лоханова Александра Васильевича «Алгоритмы распознавания и модели цифровой обработки динамических телевизионных изображений».

Разработанные в диссертации Лоханова А.В. средства информационно-технологического и технического обеспечения процедур визуального контроля продукции промышленных предприятий характеризуются высоким уровнем готовности к использованию в производственно-технологических процессах при построении комплексов технического зрения и цифровой обработки регистрируемых изображений.

Предложенные Лохановым А.В. алгоритмы и модели распознавания образов выгодно отличаются от известных аналогов, обеспечивают повышение достоверности результатов объективной оценки состояния изделий по результатам цифровой обработки динамических изображений и способствуют оперативному внедрению в производственный процесс автоматизированных контрольно-измерительных комплексов.

Результаты диссертационных исследований Лоханова А.В. представляют практический интерес для промышленных предприятий реального сектора экономики.

Главный инженер
кандидат технических наук


 А.Н. Золотов