

На правах рукописи

Чернов Тимофей Сергеевич

Математические модели и алгоритмы оценки качества изображений в системах оптического распознавания

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

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Москва — 2018

Работа выполнена в Федеральном государственном учреждении «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», лаборатория № 9-1 «Системное программирование»

Научный руководитель: доктор технических наук
Славин Олег Анатольевич

Официальные оппоненты: **Николаев Петр Петрович**,
доктор физико-математических наук,
ФГБУН Институт проблем передачи информации им.
А.А. Харкевича Российской академии наук,
главный научный сотрудник лаборатории 11 «Зрительные системы»

Потапов Алексей Сергеевич,
доктор технических наук, доцент,
ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»,
профессор кафедры компьютерной фотоники и видеоинформатики

Ведущая организация: Публичное акционерное общество «Институт электронных управляемых машин им. И.С. Брука»

Защита состоится _____ 2018 г. в _____ часов на заседании диссертационного совета Д 002.073.04 на базе Федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН) по адресу: 117312, г. Москва, проспект 60-летия Октября, 9 (конференц-зал, 1-й этаж).

С диссертацией можно ознакомиться в библиотеке ФИЦ ИУ РАН по адресу: г. Москва, ул. Вавилова, д. 40 и на официальном сайте ФИЦ ИУ РАН: <http://www.frccsc.ru>.

Отзывы на автограферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 119333, г. Москва, ул. Вавилова, д. 44, кор. 2, ученому секретарю диссертационного совета Д 002.073.04.

Автограферат разослан _____ 2018 г.
Телефон для справок: +7 (499) 135-51-64.

Ученый секретарь
диссертационного совета
Д 002.073.04,
д-р. техн. наук, профессор

В.Н. Крутко

Общая характеристика работы

Актуальность темы. Развитие науки и технологий послужило причиной тому, что современное общество вступило в новый этап использования систем распознавания образов. Большой вклад в развитие методов распознавания внесли отечественные и зарубежные ученые М.А. Айзerman, В.Л. Арлазаров, Э.М. Браверман, В.Н. Вапник, Ю.И. Журавлев, Н.Г. Загоруйко, А.Г. Ивахненко, В.Д. Мазуров, В.А. Сойфер, А.Я. Червоненкис, Y. Bengio, G. Hinton, M. Jones, Y. LeCun, P. Viola и многие другие.

Как и несколько десятилетий назад, системы распознавания и машинного зрения широко задействуются в анализе документов и текстов, медицине, биометрии, обороне, промышленности и многих других сферах. Тем не менее, процесс их использования и требования к ним существенно изменились.

Точность, устойчивость и производительность методов распознавания образов повышаются, за счет чего расширяется область их применения в автоматизации процессов жизнедеятельности, где системам распознавания доверяются все более важные и критические задачи с высокой ценой ошибки. В существующих профессиональных сферах их применения снижается время и степень участия ответственного работника или оператора. В этих условиях система распознавания должна стабильно выдавать сигнал о недостаточной достоверности результата распознавания, также называемый отказом, при подозрении на некорректность результата или при возникновении нештатной ситуации, что является запросом на передачу контроля дальнейших действий оператору.

Помимо профессиональных сфер, системы распознавания получают свое распространение и среди обычных пользователей, все более плотно проникая в повседневную жизнь людей и становясь неотъемлемой частью быта. Важную роль в этом играют повсеместно использующиеся мобильные устройства, с каждым годом повышающие свою производительность, доступность и простоту в использовании. Вместе с тем, взаимодействие систем распознавания с непрофессиональной аудиторией открывает ряд новых задач.

Обычный пользователь не является оператором в привычном смысле этого слова: он не знает, как устроен тот или иной алгоритм распознавания и не понимает, как выглядит допустимое множество входных данных, на которых следует ожидать корректный результат. Как следствие, теряется контроль над изображениями, поступающими в систему распознавания, из-за чего алгоритмам необходимо подстраиваться под всевозможные искажения входных данных. Кроме того, пользователь не привык контролировать выход систем распознавания, что снова перекладывает ответственность за информирование пользователя об ошибке на разработчика системы.

Даже при текущем уровне прогресса построение систем распознавания с абсолютной точностью на практике невозможно. Тем не менее, разные задачи имеют разную цену ошибки или требуемый уровень надежности, который можно обеспечить с помощью правил определения достоверности результатов распознавания. Проблема заключается в том, что при искаженных входных данных поведение систем

распознавания не всегда предсказуемо. Следовательно, для построения систем распознавания заданной надежности необходима разработка методов контроля качества входных изображений.

На текущий момент вопросам оценки качества изображений уделяется достаточно большое внимание вследствие их высокой актуальности. Среди ученых, внесших значительный вклад в развитие области оценки качества изображений, можно отметить A. Bovik, Z. Wang, D. Chandler, H. Sheikh, L. Zhang, E. Simoncelli, P. Ye, D. Doermann, J. Kumar и других. Разработано множество объективных (автоматических) методов оценки и коррекции качества изображений. Несмотря на это, практическое применение имеющихся методов в промышленных системах обработки и распознавания изображений имеет существенные ограничения.

Во-первых, не существует универсального метода оценки качества изображений, применимого в любых приложениях, поскольку требования к качеству различаются как среди классов входных изображений, так и среди классов конкретных систем.

Во-вторых, отсутствуют модели, методологии или рекомендации к созданию систем распознавания, учитывающих качество входных изображений с целью повышения точности, надежности, производительности и других характеристик систем.

В-третьих, для многих прикладных систем распознавания изображений и видеопотока актуальна не только традиционная постановка задачи оценки и коррекции качества входных изображений, но и ее вариации, актуальные при выборе лучшего кадра в видеопотоке, отказе в обработке поврежденных областей, интеграции результатов распознавания на разных кадрах и т.д.

Таким образом, дальнейшее исследование вопросов контроля качества изображений в системах распознавания является актуальным.

В диссертации исследуются три важных различных промышленных приложения систем распознавания: в системах документооборота и долгосрочного хранения документов, обеспечения контроля доступа и безопасности, а также распознавания изображений в непосредственном технологическом процессе.

Основные результаты диссертации были получены в процессе выполнения работ по следующим научным грантам РФФИ:

- №17-29-03170 – «Исследование быстродействующих методов и алгоритмов обработки изображений и оптического распознавания для использования в мобильных устройствах с ограниченной вычислительной производительностью»;
- №16-07-01051 – «Создание методологии разработки систем массового ввода и распознавания изображений документов»;
- №16-07-00616 – «Исследование возможности создания новых методов поиска нечетких дубликатов в видеопотоке»;
- №15-07-06520 – «Методы контроля подлинности документов и их фрагментов в гибридных системах обработки, передачи и хранения документов»;
- №14-07-00730 – «Математическое моделирование шумовых помех при распознавании»;

- №13-07-12173 – «Исследование методов сегментации изображений документов на структурные блоки методами цветового и морфологического анализа»;
- №13-07-12172 – «Распознавание документов удостоверяющих личность с помощью веб камер и камер мобильных устройств».

Объектом исследования являются системы оптического распознавания.

Предметом исследования является оценка качества входных изображений систем оптического распознавания.

Целью данной работы является разработка математических моделей и алгоритмов улучшения характеристик систем оптического распознавания путем анализа и учета неравномерности качества входных данных.

Для достижения поставленной цели необходимо было решить следующие задачи:

1. Исследовать влияние искажений входных изображений на промышленные системы распознавания и провести обзор современных методов оценки качества изображений.
2. Построить математическую модель системы распознавания изображений и видеопотока с модулями оценки качества и обратной связью на каждом этапе обработки.
3. Разработать алгоритм автоматической оценки качества цветовой сегментации изображений печатных документов.
4. Исследовать влияние детектирования и фильтрации бликов на точность поиска и распознавания документа на изображении, а также влияние бликов на устойчивость алгоритма вычисления взвешенного проективного оптического потока.
5. Разработать алгоритм детектирования и локализации периодических фоновых элементов известной структуры на изображении документа.
6. Разработать алгоритм контроля качества входного сигнала двухканального рентгенографического сепаратора.
7. Реализовать разработанные в диссертации алгоритмы для обеспечения их внедрения в промышленные системы распознавания и провести их экспериментальный анализ.

Методология и методы исследования основаны на системном анализе, математическом моделировании, методах цифровой обработки изображений и распознавания образов, Фурье-анализе.

Основные положения, выносимые на защиту:

1. Построена математическая модель системы распознавания изображений и видеопотока с модулями оценки качества и обратной связью на каждом этапе обработки.
2. Разработан алгоритм автоматической оценки качества цветовой сегментации изображений печатных документов, учитывающий особенности текстовых и графических типов информационных слоев.
3. Разработан алгоритм детектирования и локализации периодических фоновых элементов известной структуры на изображении документа, основанный на анализе Фурье-спектра, учитывающий модель документа с целью увеличения точности и

устойчивости детектирования и локализации.

4. Показано положительное влияние подавления бликов на точность работы детектора документа методом Виолы-Джонса, на точность распознавания документов в видеопотоке, а также на устойчивость алгоритма вычисления взвешенного проективного оптического потока.

5. Разработан алгоритм контроля качества входного сигнала двухканального рентгенографического сепаратора на основе анализа входного рентгенографического изображения, позволяющий производить диагностику нештатных ситуаций и способный работать в реальном времени в процессе эксплуатации сепаратора.

Научная новизна:

1. Предложена новая математическая модель системы распознавания изображений и видеопотока с модулями оценки качества и обратной связью на каждом этапе обработки.

2. Разработан новый алгоритм автоматической оценки качества цветовой сегментации изображений печатных документов.

3. Разработан новый алгоритм детектирования и локализации регулярной пространственной помехи на изображении документа.

4. Выполнено оригинальное исследование воздействия подавления бликов на точность поиска и распознавания изображений документов, а также на устойчивость вычисления взвешенного проективного оптического потока.

5. Разработан новый алгоритм контроля качества и детектирования разладки регистрирующей системы оптического распознавания алмазов в рентгеновском диапазоне.

Практическая значимость. Реализации разработанных алгоритмов выбора наилучшего кадра для распознавания в видеопотоке, автоматической оценки качества цветовой сегментации изображений печатных документов, детектирования и фильтрации бликов на изображениях документов и детектирования и локализации периодических фоновых элементов известной структуры на изображении документа внедрены в программное обеспечение «Smart PDF/A», «Smart 3D OCR MRZ», «Smart PassportReader» и «Smart IDReader» компании ООО «Смарт Эндженс Сервис». Данные продукты интегрированы в информационные решения крупнейших российских банков, страховых компаний и организаций различных государственных структур, в том числе – международных. Реализация разработанного алгоритма контроля качества входного сигнала двухканального рентгенографического сепаратора внедрена в программный комплекс сепарации алмазов в рамках коммерческого проекта компании ООО «Визилект Сервис». Предложенная модель системы распознавания изображений и видеопотока с модулями оценки качества и обратной связью использовалась при проектировании данных программных продуктов.

Достоверность полученных результатов обеспечивается согласованностью разработанных моделей и алгоритмов с экспериментальными результатами, их успешной апробацией и внедрением в коммерческие программные системы распознавания.

Апробация работы. Основные результаты работы докладывались на следующих конференциях и семинарах:

1. Seventh International Conference on Machine Vision (Milan, Italy, 2014),
2. Школе-конференции «Информационные технологии и системы – 2014» (Нижний Новгород, Россия, 2014),
3. 29th European Conference on Modelling and Simulation (Varna, Bulgaria, 2015),
4. Eighth International Conference on Machine Vision (Barcelona, Spain, 2015),
5. Школе-конференции «Информационные технологии и системы – 2015» (Сочи, Россия, 2015),
6. Техническом семинаре лаборатории зрительных систем ИППИ РАН (Москва, Россия, 2015),
7. Международном научно-исследовательском семинаре «Анализ и понимание изображений (Математические, когнитивные и прикладные проблемы анализа изображений и сигналов)» (Москва, Россия, 2017),
8. Школе-конференции «Информационные технологии и системы – 2017» (Уфа, Россия, 2017),
9. 14th IAPR International Conference on Document Analysis and Recognition (Kyoto, Japan, 2017),
10. 7th International Workshop on Camera Based Document Analysis and Recognition (Kyoto, Japan, 2017),
11. Tenth International Conference on Machine Vision (Vienna, Austria, 2017).

Личный вклад. Результаты, изложенные в диссертации, принадлежат лично автору. В совместных работах автор принимал непосредственное участие в выборе направлений и задач исследований, разработке и обсуждении результатов.

Публикации. Основные результаты по теме диссертации изложены в 24 публикациях, в том числе: 8 изданы в журналах, рекомендованных ВАК, 10 – в сборниках трудов конференций (8 работ в сборниках, входящих в международные базы цитирования Scopus или Web of Science), 1 патент на изобретение, 1 патент на полезную модель и 4 свидетельства о государственной регистрации программы для ЭВМ.

Объем и структура работы. Диссертация состоит из введения, пяти глав и заключения. Полный объем диссертации составляет 157 страниц, включая 63 рисунка и 3 таблицы. Список литературы содержит 206 наименований.

Содержание работы

Во **введении** обосновывается актуальность работы и ее научная новизна, формулируются цели и задачи диссертационного исследования, приводится краткое содержание глав диссертации и положения, выносимые на защиту.

В **первой главе** описываются промышленные приложения систем распознавания изображений и анализируется их общее устройство. Рассматриваются типы искажений входных данных, приводящие к снижению точности и надежности систем распознавания. Приводится обзор существующих методов оценки качества изображений.

Распознавание изображений находит приложения в различных дисциплинах: анализе документов и текстов, медицине, биометрии, обороне, промышленности,

робототехнике и многих других. В структуре систем распознавания изображений можно выделить присущие большинству из них функциональные блоки, после захвата изображения включающие его предварительную обработку, сегментацию, извлечение признаков, классификацию, пост-обработку результатов и обратную связь. Поскольку построение системы распознавания с абсолютной точностью практически невозможно, важной задачей является определение достоверности результатов распознавания, что усложняется проблемой чрезмерной уверенности выходных оценок распознавателей, в частности, из-за искажений входных изображений.

Входные изображения систем оптического распознавания могут подвергаться большому количеству различных искажений, особенно в неконтролируемых или естественных условиях съемки. Основными причинами искажений, рассмотренными в разделе 1.2, являются: шум и аберрации оптических систем, сжатие и интерполяция изображений, дефокусировка и размытость, тени, блики, голограммические элементы, механическая деформация и артефакты объекта. На основании проведенного обзора можно сказать, что современные системы распознавания, в том числе, основанные на нейронных сетях, могут быть неустойчивы к подобным искажениям.

Методы оценки качества изображений, обзор которых приведен в разделе 1.3, можно разделить на субъективные, основанные на экспертном мнении, и объективные – автоматические, которые представляют наибольший интерес для встраивания в системы распознавания. Некоторые объективные методы оценки качества могут требовать наличие оригинального изображения, с которым сравнивается данное искаженное изображение, что актуально при мониторинге и настройке систем сжатия или обработки изображений. Другие методы оценки принимают на вход только искаженное изображение, что делает их применимыми при непосредственном распознавании изображений или видеопотока. Приводится обзор как универсальных методов оценки качества, так и специализированных, т.е. направленных на поиск определенных искажений и позволяющих их диагностику.

В конце главы ставятся основные задачи диссертации, затрагивающие вопросы оценки качества изображений в трех важных промышленных приложениях систем распознавания: в системах документооборота и долгосрочного хранения документов, обеспечения контроля доступа и безопасности, а также распознавания изображений в технологическом процессе.

Вторая глава посвящена разработке модели системы оптического распознавания с встроенными модулями оценки качества входных данных. В рамках модели рассматривается задача выбора лучших кадров при распознавании в видеопотоке.

Рассмотрим систему $G = \langle S, E \rangle$, представленную в виде ориентированного графа подсистем обработки $S = \{S_0, \dots, S_N\}$ и связей между ними $E = \{(S_i, S_j) | S_i, S_j \in S\}$. Выходом подсистемы S_i является упорядоченный набор объектов $r_i = (r_{i,0}, \dots, r_{i,N_i})$. Каждый отдельный объект $r_{i,j}$ содержит непосредственный набор выходных данных $v_{i,j}$, тип которых зависит от устройства S_i и может быть, например, изображением или вектором альтернатив распознавания. Кроме того, он содержит скалярную оценку уверенности подсистемы в корректности своего результата $c_{i,j} \in [0, 1]$ и опциональную информацию $l_{i,j}$ о местонахождении

результата в глобальной системе координат – например, аффинное или проективное преобразование. Итого, $r_{i,j} = \langle v_{i,j}, c_{i,j}, l_{i,j} \rangle$. Ограничимся подсистемами исходной системы S_i , и теми их выходами $r_{i,j}$, компоненты $v_{i,j}$ которых содержат данные из множества изображений, т.е. $v_{i,j} \in \mathcal{I}$.

Пусть $Q_{i,j}$ – модуль оценки качества изображений для j -го выхода подсистемы S_i , принимающий на вход $r_{i,j}$. Данный модуль реализует функцию оценки качества $Q(I)$ и возвращает оценки качества $q_{i,j} \in \mathcal{D}_{i,j}$ для поступившего изображения $v_{i,j}$. Оценки качества $q_{i,j}$ могут быть представлены как скалярным значением, соответствующим степени искажения всего изображения $v_{i,j}$, так и пространственным распределением качества изображения в системе координат, заданной $l_{i,j}$.

Введем модуль коррекции и принятия решений о дальнейшей обработке $\Psi_{i,j}$, принимающий на вход как оценки качества $q_{i,j}$, так и результат $r_{i,j}$. Данный модуль возвращает модифицированные результаты $r_{i,j}^*$, тип которых совпадает с типом $r_{i,j}$, но компонента уверенности в корректности результата $c_{i,j}^*$ может отличаться в соответствии с логикой учета пространственного распределения оценок качества $q_{i,j}$. Важной особенностью $\Psi_{i,j}$ является возможность выдачи отказа в дальнейшей обработке за счет присвоения $c_{i,j} = 0$, когда повреждение покрывает большую часть изображения $v_{i,j}$, и передачи сообщения об этом родительской подсистеме S_i с помощью обратной связи. Модуль возвращает восстановленное от дефектов изображение $v_{i,j}^*$ при таковой возможности.

Последним вводимым модулем в систему является optionalный модуль внимания или интереса A_i , соответствующий подсистеме S_i . Его задача – построить карту интереса $a_i(x, y)$ подсистемы S_i в участках обрабатываемого подсистемой объекта при условии его выделения на предыдущих этапах, за счет чего возможно контролировать приоритет обработки в видеопотоке. На вход данному модулю подаются оценки качества $q_{i,j}$ и модифицированные результаты $r_{i,j}^*$ работы подсистемы, причем в модуле хранится история о результатах на предыдущих кадрах, т.е. выходная карта является интегральной по видеопотоку на текущий момент. Применение карты интереса особенно актуально при распознавании объектов с фиксированными формами, например, документов – низкий интерес в регионе означает, что система уже имеет высокий уровень уверенности в результатах распознавания в этом регионе, или же что в нем отсутствуют какие-либо информативные участки. Модифицированные результаты $r_{i,j}^*$ передаются на следующий этап обработки в соответствии с существующими связями между подсистемами. На рисунке 1 проиллюстрирована модель графа подсистем и их связей после добавления модулей оценки качества изображений в процесс обработки.

В рамках разработанной модели рассматривается задача выбора кадров для распознавания в видеопотоке. Пусть Ψ_0 – модуль выбора кадров, имеющий устройство очереди с приоритетом максимального размера L , накапливающий входные кадры и при запросе выдающий кадр с наибольшим приоритетом. Введем дополнительный модуль оценки приоритета кадра P , зависящий от результатов оценки качества $q(x, y)$, времени t нахождения кадра в очереди, а также от карты интереса $a(x, y)$ системы распознавания в соответствующих регионах входного изображения.

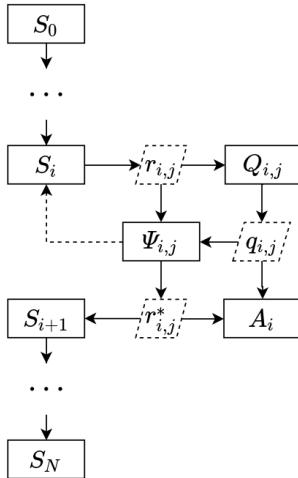


Рисунок 1 — Граф системы обработки с модулями оценки качества изображений

Вычислим коэффициент w , показывающий пригодность данного кадра с точки зрения качества интересующих систему регионов: $w = \sum_{x,y} q(x, y) a(x, y)$. В качестве способа вычисления итогового приоритета кадра p с учетом времени нахождения в очереди t и пригодности кадра w возьмем способ, использующий гауссово ядро: $p(w, t) = w^\alpha e^{-\frac{\beta t^2}{\sigma^2}}$, где коэффициент α регулирует важность оценки качества, а β и σ контролируют период релевантности кадра относительно времени захвата.

Эксперименты проводились с реализацией системы Smart IDReader, пред назначенной для распознавания удостоверяющих личность документов в видеопотоке. Были моделированы условия размытия входных кадров при фокусировке или движении камеры с помощью функции $\sigma(t) = a|\text{sinc}(b(t - t_0))| = \frac{a|\sin(b(t - t_0))|}{b|t - t_0|}$ зависимости радиуса фильтра Гаусса от времени захвата кадра $t > 0$. Степень размытия кадра оценивалась с помощью алгоритма, основанного на анализе модулей производной изображения в четырех направлениях; экспериментально подтверждено соответствие возвращаемых алгоритмом оценок настоящему размытию.

Для экспериментов с системой распознавания использовались тестовые выборки из 217 и 39 видеороликов удостоверяющих личность документов (ID-карты Германии, паспорт гражданина РФ) длительностью 5 секунд каждый, снятых с использованием различных мобильных устройств. Итоговой точностью работы системы на выборке считалась средняя точность распознавания всех полей документа. Каждый видеоролик разбивался на кадры с фиксированной частотой $F = 5$ кадров в секунду. Кадры передавались на распознавание с периодом $T = (1, \dots, 8)$ кадров (т.е. с итоговой частотой $\frac{F}{T}$ кадров в секунду), тем самым имитируя продолжительность распознавания одного кадра. Были введены искусственные ограничения на максимальную продолжительность видео $M = (5, 3, 1)$ секунд для проверки потенциала скорости распознавания, т.е. точности при фиксированном максимальном времени обработки

видеопотока. Эксперимент проводился в двух режимах: с простым выбором каждого T -го кадра («Fixed») и выбором наилучшего кадра в каждый T -й момент времени в соответствии с предложенным способом вычисления приоритета («Priority»).

На рисунке 2 показаны полученные результаты точности распознавания видеопотока для исходной (вверху) и размытой (снизу) выборки изображений с различными ограничениями M на максимальную продолжительность видео (для ID-карт Германии, аналогичные результаты были получены для паспортов гражданина РФ). По горизонтальной оси отложен период между кадрами T , а по вертикальной – средняя точность распознавания всех полей ACC .

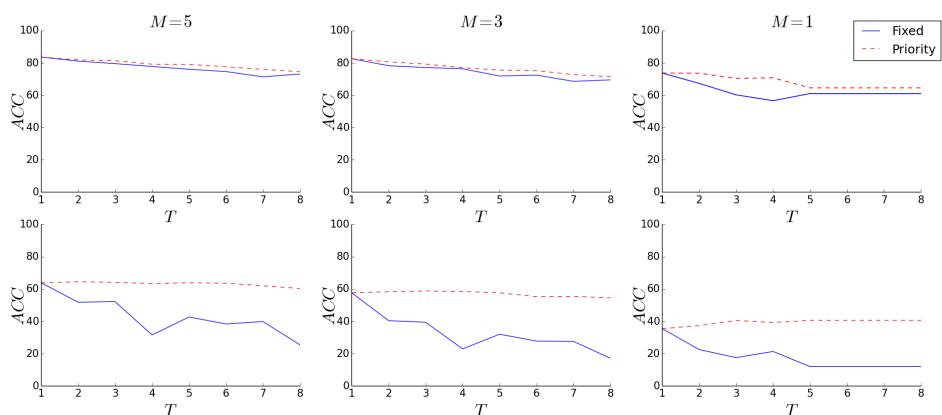


Рисунок 2 – Графики точности распознавания видеопотока для исходной (сверху) и размытой (снизу) выборки изображений

Таким образом, экспериментальные результаты демонстрируют значительное увеличение точности распознавания с различными ограничениями на продолжительность видеопотока при выборе кадров в соответствии с их приоритетом при искусственно моделируемых условиях автоматической фокусировки камеры, приводящих к размытию кадров.

В третьей главе представлен алгоритм автоматической оценки качества цветовой сегментации изображений печатных документов, актуальный при их сжатии для долговременного хранения. Цветовая сегментация (расслоение) изображения на информационные слои является важнейшим этапом работы систем, основанных на модели смешанного растрового содержимого (MRC), к которым принадлежит исследуемая в главе система сжатия изображений цветных печатных документов в формат PDF/A. Использование автоматических методов оценки качества необходимо для увеличения эффективности и снижения стоимости контроля качества и настройки параметров системы. Основной идеей разрабатываемых функций оценок информационных слоев (текстового и графического типов) является использование подготовленной вручную разметки, задающей модели качества информационных слоев с учетом специфики принадлежащих к ним объектов.

Оценка качества текстового слоя производится с помощью разметки слов в некоторых областях изображений, что снижает требуемые трудозатраты по сравнению с посимвольной разметкой (рисунок 3а). Для подсчета оценки используется формула:

$$Q_{OCR} = \frac{\sum_{w_i} |w_i| \cdot \text{recognized}(w_i)}{\sum_{w_i} |w_i|}, \quad (1)$$

где w_i – слово из разметки, $|w_i|$ – его длина (количество символов), $recognized(w_i)$ – показатель полностью успешного распознавания слова w_i . Распознавание производилось популярной кросс-платформенной системой Tesseract OCR.



Рисунок 3 – Разметка текстового и графического слоя

Для оценки качества графического слоя требуется оценить отклонение полученной бинарной маски слоя от некоторой ее заранее известной разметки. Однако, ручная разметка отдельных пикселей маски слоя крайне трудоемка и не может быть проведена за разумное время. Поэтому используется значительно более эффективный подход с точки зрения трудозатрат размечающего эксперта. Строится инвертированная карта цветовой насыщенности, где насыщенность вычисляется по формуле: $S = \max(R, G, B) - \min(R, G, B)$, а затем инвертируется: $\bar{S} = 255 - S$. Далее, экспертом выделяется набор непересекающихся прямоугольных областей, и для каждой области устанавливается оптимальный с точки зрения визуального восприятия порог бинаризации пикселей инвертированной карты насыщенности (рисунок 3б).

Классическим способом оценки отклонения двух бинарных векторов x и y длины N является среднеквадратичное отклонение: $MSE(x, y) = \frac{1}{N} \sum_{i=1}^N b(x_i, y_i)$, где $b(x, y) = 1 - [x_i = y_i]$. Тогда, для оценки качества построения маски графического слоя Y относительно размеченной инвертированной карты насыщенности X можно вычислить суммарное отклонение по размеченным областям:

$$Q_{BINMSE}(X, Y) = \sum_{i=1}^{N_A} MSE(X_i, Y_i), \quad (2)$$

где X_i , Y_i – соответствующие i размеченной области части разметки и полученного

графического слоя, а N_A – количество областей. Данная функция оценки имеет недостаток: в общем случае невозможно составить единственную идеальную разметку, соответствующую наилучшему визуальному восприятию, поскольку последнее может оставаться прежним от присутствия или отсутствия в маске слоя некоторых пикселей (чаще всего – на границе объектов). Поэтому, вводится еще одна функция оценки, основанная на плотности пикселей $\rho(A)$ в областях разметки:

$$Q_{BINCNT}(X, Y) = \frac{\sum_{i=1}^{N_A} (\rho(X_i) - \rho(Y_i))^2 \cdot (W_{A_i} \cdot H_{A_i})}{\sum_{i=1}^{N_A} (W_{A_i} \cdot H_{A_i})}, \quad \rho(A) = \frac{N_1(A)}{W_A \cdot H_A}, \quad (3)$$

где $N_1(A)$ – количество принадлежащих маске пикселей в области разметки A , имеющей размеры $W_A \times H_A$. Данный способ показал большую линейную корреляцию с экспертными оценками по сравнению с первым.

Итоговая функция оценки качества определяется как линейная комбинация независимых оценок слоев: $Q = w_1 \cdot Q_{BINCNT} + w_2 \cdot Q_{OCR}$, где веса w_1 и w_2 выставляются экспертом для регулировки значимости соответствующих слоев и в экспериментах принимались равными 0.5.

Работоспособность разработанного алгоритма оценки качества была экспериментально протестирована путем автоматической оптимизации реализации системы цветовой сегментации как системы типа «черный ящик». Входами являлись параметры цветовой сегментации, отвечающие за классификацию пикселей к текстовому или графическому слою, а выходами – вычисленные оценки качества на заранее размеченной тестовой выборке после сегментации с учетом входных параметров. Автоматическая оптимизация проводилась с помощью программного обеспечения NOMAD, реализующего алгоритм MADS. В результате оптимизации для всех изображений выборки качество проблемных (при различных вариантах исходных параметров) изображений также визуально улучшилось, а качество изначально удачных – осталось прежним, что подтверждает соответствие возвращаемых алгоритмом оценок визуальному качеству сегментации.

В четвертой главе рассмотрены приложения оценки качества изображений в системах распознавания при обеспечении контроля доступа и безопасности на предприятиях.

Раздел 4.1 содержит анализ негативного влияния бликов (рисунок 4а) на этапы работы системы поиска и распознавания документов. Показана актуальность детектирования бликов и интерполяции поврежденного ими фона (рисунок 4б) для систем распознавания в видеопотоке по причине неустойчивости методов поиска документа к выраженным бликам, а также из-за возможности комбинирования результатов частичного распознавания неповрежденных участков между кадрами.

Для детектирования бликов и восстановления фона были разработаны вычислительно простые алгоритмы, позволяющие работу в реальном времени на мобильных устройствах. Построение первоначальной бинарной маски принадлежности блику



а)

б)

Рисунок 4 — Пример выраженного блика на изображении паспорта РФ и результата восстановления фона

осуществляется пороговым отсечением по яркости и насыщенности. Следующим этапом из-за неравномерности поверхности документа применяется морфологическая операция закрытия с целью закрашивания небольших бассейнов внутри положительной маски, а затем – морфологическая операция дилатации.

Алгоритм восстановления фона основан на обходе компонент связности найденной бинарной маски в ширину. В очередь Q заносятся внутренние граничные пиксели компоненты, после чего на каждой итерации пиксель (x, y) достается из очереди и его значение для цветовых компонент пересчитывается, как усреднение по множеству $N(x, y, W)$ не принадлежащих блику или же уже закрашенных бликовых пикселей в окне с центром (x, y) и крылом W . Затем, еще не закрашенные непосредственные соседи текущего пикселя добавляются в очередь Q .

В экспериментах использовался набор из 579 изображений 3-й страницы (покрытой пленкой) паспортов РФ, полученных с камер мобильных устройств и содержащих выраженные блики от направленных источников света. Целью первого эксперимента был анализ устойчивости детектора документа Виолы-Джонса к бликам и проверка возможности повышения точности поиска документа за счет восстановления поврежденного бликами фона. Результаты показали, что на 75.3% изображений документ был изначально найден корректно, на 19.4% отказ детектора перешел в корректный результат после восстановления фона, а на оставшихся 5.3% изображениях не удалось добиться поиска документа даже после восстановления. Поскольку отрицательное срабатывание детектора приводит к невозможности дальнейшего распознавания, исправление детектирования одной пятой части выборки документов, что также является уменьшением числа ошибок в 4.6 раз, можно считать существенным. Результаты второго эксперимента показали, что средняя точность распознавания полей документа увеличилась с 45% до 57% (полное совпадение с идеальной разметкой) и с 60% до 74% (частичное совпадение) после рассмотрения восстановленных изображений.

Раздел 4.2 посвящен экспериментальному анализу проблемы бликов при вычислении проективного оптического потока, применяющейся для отслеживания движения плоских ригидных документов в видеопотоке.

Рассматривается модификация задачи поиска проективного оптического потока обобщением классического алгоритма Лукаса-Канаде, допускающая использование мультиплексивных весов в функционале с целью учета оценки достоверности значений пикселей (например, в случае потери информации из-за блика), которая формулируется следующим образом:

$$\sum_x \left[I_\tau(x) - I_\varphi(W(x, p + \Delta p)) \right]^2 \cdot \omega(x, p + \Delta p) \rightarrow \min_{\Delta p}, \quad (4)$$

где $I_\tau(x)$, $I_\varphi(x)$ – исходные изображения; $W(x, p)$ – некоторое параметрическое преобразование с вектором параметров p ; $W(x, p)$ задает проективное преобразование из СК изображения $I_\varphi(x)$ в СК изображения $I_\tau(x)$; значения $I_\varphi(W(x, p + \Delta p))$ вычисляются при помощи билинейной интерполяции; $\omega(x, W(x, p)) = \omega_\tau(x) \cdot \omega_\varphi(W(x, p))$ – итоговая функция весов, а $\omega_\tau(x)$ и $\omega_\varphi(x)$ – функции весов изображений $I_\varphi(x)$ и $I_\tau(x)$.

Для проведения экспериментов были получены наборы изображений синтетических и реальных данных с разметкой четырех не лежащих на одной прямой точек объекта для вычисления проективного преобразования между любыми двумя кадрами одного набора. Синтетические данные (SD) представляют собой последовательность из 200 изображений плоского прямоугольника с текстурой шахматной доски. В качестве реальных данных использовались последовательности изображений, содержащие карту «Тройка» (TD, 56 изображений) и банковскую карту (BD, 50 изображений). Эксперименты проводились для изображений в цветовых пространствах $c = \{RGB, \alpha\beta\}$.

Основной оценкой результата работы алгоритма вычисления проективного потока на данном множестве пар изображений O_s было выбрано отношение $\psi^c(O_s, t) = \frac{|O_s^c(t)|}{|O_s|}$, где O_s – множество пар индексов рассматриваемых кадров с шагом s , а $O_s^c(t)$ – множество успешно сопоставленных пар индексов, т.е. таких, для которых вычисленная ошибка сопоставления $Q_{ij} = Q(P_{est}(I_i^c, I_j^c), P_{markup}(I_i, I_j))$ была меньше порогового значения t , причем:

$$Q(P_{est}, P_{markup}) = \frac{1}{4} \sum_q |q - P_{markup}^{-1} \cdot P_{est} \cdot q| \quad (5)$$

для истинного P_{markup} и вычисленного P_{est} проективных преобразований и $q \in \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$.

В таблице 1 приведены значения показателей качества $\psi^{RGB}(O_s, t)$, $\psi^{\alpha\beta}(O_s, t)$ для каждого набора данных.

Таблица 1 – Показатели качества вычисления оптического потока

Данные	SD	TD	BD
Порог допустимой ошибки t	40	40	30
Показатель качества $\psi^{RGB}(O_s, t)$	0.396	0.348	0.096
Показатель качества $\psi^{\alpha\beta}(O_s, t)$	1.000	0.858	0.637

Таким образом, статистические характеристики алгоритма вычисления проективного оптического потока на рассматриваемых данных с бликами выше в цветовом подпространстве $\alpha\beta$ чем в RGB . Полученные результаты подтверждают негативное влияние бликов на алгоритм вычисления оптического потока, что также является дополнительным аргументом в пользу разработки и использования методов детектирования зашумляющих бликов.

Раздел 4.3 посвящен разработке алгоритмов детектирования и локализации периодических фоновых элементов известной структуры на изображении документа.

Детектирование, т.е. определение факта наличия особых периодических визуальных элементов, таких, как голограммы, делает возможным автоматическую проверку подлинности документа и классификацию его типа. Локализация периодических элементов (рисунок 5) позволяет системам распознавания регулировать параметры в областях нанесения таких элементов и задействовать методы их подавления с максимальным сохранением текстовой информации. Это обусловлено тем, что защитные элементы изменяют характеристики распознаваемых областей, тем самым являясь шумом для систем распознавания.



Рисунок 5 – Локализация голографического шаблона на паспорте РФ

Вводится следующая модель изображения документа, сначала рассматриваемая для одномерного случая. Исходный сигнал изображения $I(x)$ составлен из фонового изображения $h(x)$ и изображения периодического шаблона $g(x)$, комбинированными аддитивно. Сигнал $g(x)$ может быть представлен как свертка единичного элемента $f(x)$ и гребня Дирака $c(x)$, тогда:

$$I(x) = h(x) + g(x) = h(x) + f(x) * c(x) = h(x) + f(x) * \sum_{m=0}^{M-1} \delta(x - mT), \quad (6)$$

где M – заранее известное целое число периодических элементов, которое умещается в длину сигнала N , а $T = \frac{N}{M}$ – период одного элемента. При такой модели задача детектирования формулируется, как оценка мощности $f(x)$, а локализации – как оценка фазы (циклического сдвига) $c(x)$.

Дискретное преобразование Фурье импульсного сигнала $c(x)$ имеет особую форму другого импульсного сигнала с периодом M :

$$\mathcal{F}_k c(x) = \sum_{m=0}^{M-1} e^{-2\pi i k \frac{m}{M}} = M \cdot \sum_m \delta(k - mM), \quad (7)$$

из-за чего $\mathcal{F}_k g(x)$ равно $M \cdot \mathcal{F}_k f(x)$ при $k \in \{0, M, 2M, \dots\}$ и нулю в противном случае, порождая амплитудные пики с периодом M , что и является основой для дальнейшего анализа ДПФ спектра. Сдвиг $c(x)$ на S означает, что:

$$\mathcal{F}_k c(x - S) = \mathcal{F}_k c(x) \cdot e^{i\Phi_k}, \quad \Phi_k = \Phi \cdot k, \quad \Phi = -\frac{2\pi}{N}S, \quad (8)$$

а фазовый угол на ϕ_m на m 'ом амплитудном пике становится равен:

$$\phi_m = \phi \cdot m, \quad \phi = -\frac{2\pi}{T}s = \Phi \cdot M, \quad s = S \bmod T. \quad (9)$$

Первым этапом разработанного алгоритма является вырезание области изображения, содержащей целое количество периодов периодического шаблона по горизонтали и вертикали (рисунок 6а для случая решетки 2×2), которая затем подвергается существенному сжатию и сглаживанию для подавления фона и нивелированию различий между периодическими элементами (рисунок 6б), после чего вычисляется двумерное дискретное преобразование Фурье (рисунок 6в). С целью рассмотрения дополнительной полезной информации также используется маскирование и замещение областей изображения, позволяющее расширить область до периодической решетки 3×2 .

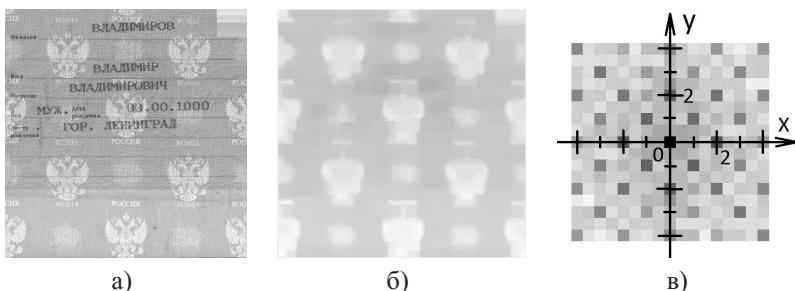


Рисунок 6 – Этапы обработки изображения и амплитуда ДПФ

Был разработан вычислительно простой алгоритм детектирования периодического шаблона, основанный на сопоставлении амплитуды ДПФ заранее известной структуре пиков, следующей из фиксированной структуры шаблона. Однако, второй разработанный алгоритм детектирования, а также алгоритм локализации, требуют рассмотрения фазы спектра.

Следующим этапом является предварительная обработка спектра $\mathcal{FI}(x)$, представленного суммой $\mathcal{F}h(x) + \mathcal{F}f(x) \cdot \mathcal{Fc}(x)$, для возможности получения информации из $c(x)$ о сдвиге $\phi = (\phi_x, \phi_y)$. Для экспериментальной оценки фазы периодического элемента в пике (i, j) используется среднее значение для угловых величин:

$$\arg \mathcal{F}f(i, j) = \arg \sum_k \exp(\arg \mathcal{F}f_k(i, j)), \quad (10)$$

где $\arg \mathcal{F}f_k(i, j)$ – разница между идеальным фазовым сдвигом в пике $\theta_k(i, j)$, вычисляемым по заранее подготовленной вручную разметке идеального сдвига всего шаблона, и актуальным значением $\phi_k(i, j)$ фазы в пике.

Подавление спектра фона $\mathcal{F}h(x)$ для пика в позиции (x, y) осуществляется линейной интерполяцией по ближайшим соседям пика: как простым усреднением по формуле $\mathcal{FI}(x, y) := \mathcal{FI}(x, y) - \overline{\mathcal{FI}(x', y')}$, так и более общим подходом с решением системы уравнений методом наименьших квадратов для нахождения коэффициентов линейной модели.

Нахождение итоговой информации о фазовом сдвиге производится путем решения системы из n уравнений $\{i \cdot \phi_x + j \cdot \phi_y = \phi_{i,j} \pmod{2\pi}\}$ или $A\phi = b \pmod{2\pi}$, в которой каждое уравнение соответствует пику (i, j) и правые части $\phi_{i,j}$ фазы спектра подвергаются описанной ранее предварительной обработке. Для решения данной переопределенной нелинейной системы уравнений сначала выбирается начальное решение ϕ^* (из множества решений, возникающих из рассмотрения уравнений пиков с одной нулевой координатой), минимизирующее невязку $\delta_k = \frac{1}{n} |A\phi_k^* - b|^2$. Большое значение δ_k может означать ложное обнаружение пика, на чем также основан второй алгоритм детектирования. Затем, производится итеративная процедура, на каждой итерации которой пространство решений сдвигается относительно текущего решения: $A\theta = b - A\phi^*$, а в систему добавляются уравнения для пиков по мере увеличения $(|i| + |j|)$, т.е. расстояния от центрального пика. С предположением о постепенном уменьшении ошибки, это обеспечивает нахождение каждого уравнения в пределах 2π , что позволяет решать систему методом наименьших квадратов на каждой итерации, при этом расширяя множество рассматриваемых пиков.

Финальным этапом алгоритма после нахождения итогового фазового сдвига $\phi = (\phi_x, \phi_y)$ является его преобразование в пиксельный сдвиг $s = (s_x, s_y)$ и экстраполяция периодического шаблона, что было показано на рисунке 5.

В экспериментах использовались отсканированные изображения российских паспортов из 497 позитивных изображений, т.е. содержащих периодический шаблон, и 714 – негативных. Каждое изображение позитивной выборки содержит вручную размеченный сдвиг периодического шаблона.

На рисунке 7а показано распределение множеств возвращаемых показателей алгоритма детектирования, основанного на анализе невязки решения системы после итерационного нахождения фазового сдвига. Алгоритм имеет абсолютную точность детектирования, а зазор между двумя множествами обеспечивает лучшую разделяющую способность по сравнению с алгоритмом детектирования, основанным на проверке соответствия амплитуды ДПФ пиковой структуре.

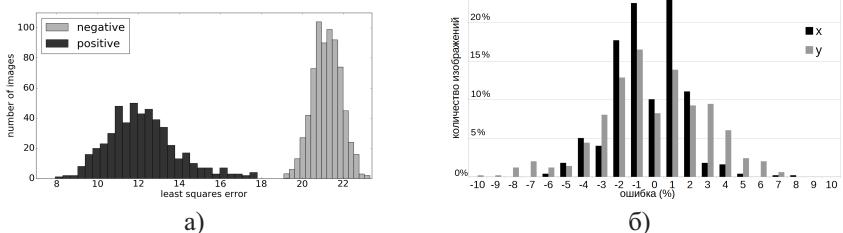


Рисунок 7 – Показатели алгоритма детектирования и ошибки локализации

Рисунок 7б содержит гистограмму распределения горизонтальных и вертикальных ошибок (в процентах от стороны окаймляющего прямоугольника периодического элемента) для лучшего варианта алгоритма локализации, использующего все предложенные этапы. Итоговая ошибка, т.е. корень из суммы квадратов x и y ошибок, равна 2.78 для региона изображения 64×64 и 162 уравнений. При сжатии региона до 256×256 и решая систему с 2698 уравнениями возможно достижение ошибки 2.64 ценой замедления алгоритма.

Таким образом, разработанные алгоритмы детектирования и локализации имеют высокие точностные характеристики, позволяющие их применение в промышленных системах распознавания. Были проведены дополнительные эксперименты для анализа устойчивости алгоритмов детектирования и локализации к моделируемым ошибкам нахождения границ документа, подтвердившие возможность использования разработанных алгоритмов при небольших ошибках нахождения границ.

В пятой главе рассматривается задача контроля качества входного сигнала двухканального рентгенографического сепаратора алмазосодержащей руды для трех возможных некорректируемых нештатных ситуаций: раскалибровки пары источник/детектор, заслонения детектора и загрязнения конвейерной ленты.

Метод рентгенографической (РГ) сепарации предложен в 2012 году и основан на учете зависимости ослабления проходящего рентгеновского излучения от элементного состава исследуемого материала. Принципы рентгенографической сепарации могут быть использованы в распознающем алгоритме только при условии калибровки излучения рентгеновских источников, чувствительности детекторов, оптической плотности ленты транспортера и скорости его движения. Однако даже в этом случае в процессе эксплуатации параметры могут постепенно отклоняться от калиброванных значений; кроме того, могут возникать другие нештатные ситуации, искажающие сигнал. Во всех таких случаях требуется либо остановка сепаратора, либо динамическая корректировка параметров, иначе качество сепарации может существенно упасть, что может привести к значительным убыткам. Следовательно, необходим способ контроля качества входного сигнала, позволяющий передавать сигнал о раскалибровке параметров или возникновении другой нештатной ситуации.

На основании анализа существующих методов контроля качества и специфики устройства РГ сепаратора было выяснено, что интерес представляет разработка алгоритмов контроля, обладающих следующими свойствами: высокая производительность

по причине высокой скорости движения конвейерной ленты в реальном времени, возможность диагностики нештатной ситуации вместе с ее детектированием для ускорения поиска неисправности и возможности ее автоматического исправления, работа только с входными РГ изображениями для избавления от необходимости в имитаторах алмаза и других внешних способах контроля.

Изменение чувствительности детектора (например, ее уменьшение за счет «выцветания» люминофора), а также интенсивности входного излучения (за счет изменения напряжения на трубке) может приводить к плавному изменению диапазона наблюдаемого сигнала (рисунок 8а). Задача контроля чувствительности канала может

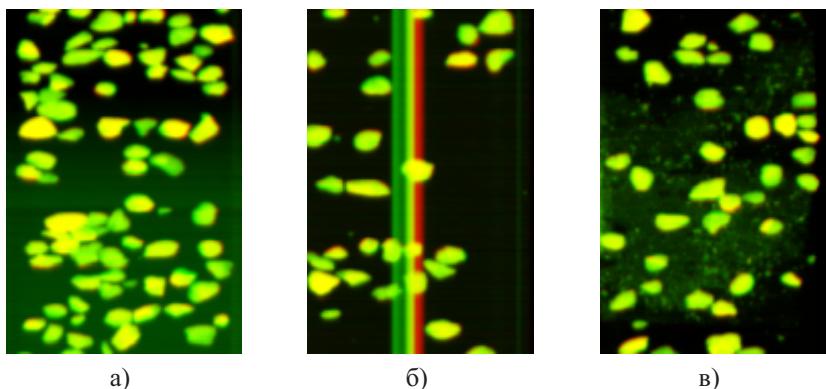


Рисунок 8 — Сигнал при раскалибровке детектора, заслонении зерном породы и загрязнении ленты

быть сведена к пороговому детектированию отклонения текущей средней яркости сигнала в области пустой ленты от начального расчетного значения, как правило, не превышающего 70%.

Для оценки яркости фона будем поддерживать гистограмму H яркости канала в экспоненциально затухающем окне, пересчитываемую для i -й позиции t -й строки по формуле: $H_i^t = \alpha \cdot H_i^{t-1} + (1 - \alpha) \cdot h_i^t$, $0 < \alpha < 1$, где α – коэффициент экспоненциального затухания, а h – гистограмма яркости новой строки. В получившейся гистограмме найдем наименьший номер e ячейки гистограммы, для которой сумма долей значений (квантиль) с нулевой ячейки до данной не меньше, чем параметр q_e , задающий предполагаемую долю фона, не занятого объектами. Для уменьшения числа ложных срабатываний вводится значение накопленного штрафа p (в начале работы системы равное нулю), которое наращивается в сторону единицы в случае, если отклонение e от исходного значения превышает порог. Пересчет накопленного штрафа производится по формуле: $p_t = \beta \cdot p_{t-1} + (1 - \beta) \cdot f$, $0 < \beta < 1$, где параметр β задает требуемую протяженность вертикальной серии превышающих порог пикселей, а f принимает значение единицы при превышении отклонения и нуля в противном случае. Сигнал о раскалибровке соответствующего канала передается при превышении значением накопленного штрафа порогового значения.

Алгоритм был протестирован на 100 тестовых пакетах данных, содержащих записанные в виде двумерных массивов сигналы каналов «высокая» и «низкая» энергия рентгенографического сепаратора (каждый массив имеет ширину 256 пикселей и длину 60 000 строк). Алгоритм обнаружил изменение интенсивности сигнала на 7 пакетах из 8 и не дал ни одного ложноположительного срабатывания на остальных 92 пакетах. Следует заметить, что на пакете с ложноотрицательным результатом изменение интенсивности оказалось подпороговым и для алгоритма классификации, что делает этот случай спорным.

Из-за конструкционных особенностей сепаратора теоретически возможно попадание зерен породы непосредственно на детектор, где они будут создавать постоянную помеху в нескольких ячейках детектирующей линейки (рисунок 8б). Для детектирования заслонения детектора требуется найти продолжительный по времени участок, яркость сигнала на котором выше порогового значения.

Разработанный алгоритм основан на модификации экспоненциального сглаживания (с целью уменьшения использования памяти) накапливаемого штрафа для каждой позиции строки, постепенно растущего вверх к известному значению яркости темнового тока d_i , перед этим ограничиваемым яркостью пикселя: $p_i^t = \alpha \cdot \min(p_i^{t-1}, s_i^t) + (1-\alpha) \cdot d_i$, $0 < \alpha < 1$, где α задает требуемую протяженность вертикальной серии превышающих порог пикселей. Алгоритм был протестирован на 92 тестовых пакетах, один из которых содержал сигнал с заслоненного детектора; все пакеты были обработаны без ошибок.

Последним рассматриваемым негативным эффектом является постепенное загрязнение ленты транспортера пылью и чешуйками породы, налипающими под действием статического электричества, затрудняющее детектирование (рисунок 8в).

Разработанный алгоритм состоит из нескольких этапов. Первым этапом является первичная классификация пикселей s_i входной строки, представленных парой компонент низкого и высокого каналов (l_i, h_i) на «пыль» ($L_{min}^d \leq l_i \leq L_{max}^d$ и $H_{min}^d \leq h_i \leq H_{max}^d$) и «камень» ($L_{min}^s \leq l_i \leq L_{max}^s$ и $H_{min}^s \leq h_i \leq H_{max}^s$). Бинарная маска «камней» расширяется морфологической операцией дилатации, после чего итоговая маска принадлежности к «пыли» получается в результате поэлементной конъюнкции с отрицанием расширенной маски «камней». Штраф p_i^t накапливается с помощью экспоненциального сглаживания ступенчатой функции: $p_i^t = \alpha \cdot p_i^{t-1} + (1 - \alpha) \cdot [s_i^t - \text{«пыль»}]$, $0 < \alpha < 1$, а затем усредняется; значение α регулирует требуемую продолжительность вертикальной загрязненной полосы. С целью сокращения ложных срабатываний важно учитывать периодичность появления загрязнений на ленте транспортера, для чего поддерживается список Q регистрации событий превышения порогом штрафа в виде пар (t', n_t) , где t' – время (позиция на ленте) первой регистрации события на данном участке, а n_t – количество срабатываний в 10%-й окрестности позиции. Если число событий для данного участка ленты стало кратным заданному порогу, то генерируется сигнал о наличии оптического загрязнения. Алгоритм был протестирован на 92 тестовых пакетах, в 3 из них на ленте присутствовало загрязнение; все пакеты были обработаны без ошибок.

В заключении приведены основные результаты работы, которые заключаются в следующем:

1. Построена модель системы распознавания изображений и видеопотока с модулями оценки качества и обратной связью на каждом этапе обработки. Данная модель позволяет увеличить точность, устойчивость и скорость работы систем распознавания в неконтролируемых условиях съемки за счет возможности контроля уверенности результата и отказа в дальнейшей обработке в случае низкого качества входного изображения. В рамках модели рассмотрен пример задачи выбора наилучших кадров для распознавания в видеопотоке в условиях автоматической фокусировки камеры, приводящих к размытию кадра. Экспериментально показано значительное увеличение точности работы системы распознавания удостоверяющих личность документов с различными ограничениями на продолжительность видеопотока.

2. Разработан алгоритм автоматической оценки качества цветовой сегментации изображений печатных документов, основанный на независимой оценке текстового слоя через точность распознавания и графического слоя с помощью вычисления разницы плотности принадлежащих слою пикселей на размеченной карте насыщенности изображения. Экспериментально подтверждена пригодность алгоритма для автоматической оптимизации параметров цветовой сегментации.

3. Выполнено исследование негативного влияния бликов на точность поиска и распознавания документа на изображении, а также на устойчивость алгоритма вычисления взвешенного проективного оптического потока. Разработаны алгоритмы детектирования бликов и интерполяции поврежденного ими фона, экспериментально показано увеличение точности поиска и распознавания документов после фильтрации бликов. Экспериментально подтверждена неустойчивость алгоритма вычисления взвешенного проективного оптического потока к бликам в RGB пространстве и возможность его заметного улучшения при переходе в $\alpha\beta$ пространство и исключении зашумляющих областей.

4. Разработаны алгоритмы детектирования и локализации периодических фоновых элементов известной структуры на изображении документа, основанные на введенной математической модели изображения документа, предварительной обработке изображения и спектра двумерного дискретного преобразования Фурье, анализе спектра и итерационном решении нелинейной системы уравнений для получения итогового результата локализации. Экспериментальные результаты демонстрируют высокую точность алгоритмов детектирования и локализации. Проведен анализ устойчивости разработанных алгоритмов к ошибкам нахождения границ документа с использованием искусственного моделирования, показавший устойчивость к небольшим ошибкам.

5. Разработаны алгоритмы детектирования нештатных ситуаций двухканального рентгенографического сепаратора для случаев раскалибровки пары источник/детектор, заслонения детектора и загрязнения конвейерной ленты, экспериментально показавшие высокую точность и позволяющие их применение в промышленных решениях для работы в реальном времени за счет высокой производительности и простоты реализации.

6. Практическая ценность полученных результатов подтверждается их использованием при проектировании и реализации промышленных систем распознавания. Реализации алгоритмов выбора наилучшего кадра для распознавания в видеопотоке, автоматической оценки качества цветовой сегментации изображений печатных документов, детектирования и фильтрации бликов на изображениях документов и детектирования и локализации периодических фоновых элементов известной структуры на изображении документа внедрены в программное обеспечение «Smart PDF/A», «Smart 3D OCR MRZ», «Smart PassportReader» и «Smart IDReader» компании ООО «Смарт Энджинс Сервис». Данные продукты интегрированы в информационные решения крупнейших российских банков, страховых компаний и организаций различных государственных структур, в том числе – международных. Реализация алгоритма контроля качества входного сигнала двухканального рентгенографического сепаратора внедрена в программный комплекс сепарации алмазов в рамках коммерческого проекта компании ООО «Визиллект Сервис». Построенная модель системы распознавания изображений и видеопотока с модулями оценки качества и обратной связью на каждом этапе обработки использовалась при проектировании данных коммерческих программных систем распознавания.

Публикации автора по теме диссертации

В изданиях из списка ВАК РФ

1. Chernov, T. S. An algorithm for detection and phase estimation of protective elements periodic lattice on document image / T. S. Chernov, S. I. Kolmakov, D. P. Nikolaev // Pattern Recognition and Image Analysis. — 2017. — Vol. 27, no. 1. — 53–65 (Scopus).
2. Оценка качества входных изображений в системах распознавания видеопотока / Т. Чернов, Н. Разумный, А. Кожаринов, Д. Николаев, В. Арлазаров // Информационные технологии и вычислительные системы. — 2017. — № 4. — С. 71–82.
3. Чернов, Т. Детектирование и фильтрация бликов в задачах распознавания документов с мобильных устройств / Т. Чернов // Труды Института системного анализа Российской академии наук. — 2017. — Т. 67, № 1. — С. 66–74.
4. Алгоритм взвешенного поиска проективного оптического потока, устойчивый к бликам / Д. Шепелев, Е. Ершов, А. Терешин, Т. Чернов, Д. Николаев // Сенсорные системы. — 2018. — в печати (RSCI).
5. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения / Т. Чернов, Д. Ильин, П. Безматерных, И. Фараджев, С. Карпенко // Вестник Российского фонда фундаментальных исследований. — 2016. — Т. 92, № 4. — С. 55–71.

6. Ключевые аспекты распознавания документов с использованием малоразмерных цифровых камер / Д. Полевой, К. Булатов, Н. Скорюкина, Т. Чернов, В. Арлазаров, А. Шешкус // Вестник Российского фонда фундаментальных исследований. — 2016. — Т. 92, № 4. — С. 97–108.
7. Арлазаров, В. Метод нечеткого поиска изображений в больших объемах видеоданных / В. Арлазаров, К. Булатов, Т. Чернов // Системы высокой доступности. — 2016. — Т. 12, № 1. — С. 53–58.
8. Николаев, Д. Метод автоматической оценки качества цветовой сегментации в задаче упаковки изображений печатных документов / Д. Николаев, Д. Полевой, Т. Чернов // Труды Института системного анализа Российской академии наук. — 2013. — Т. 63, № 3. — С. 78–84.

В сборниках трудов конференций

9. Image quality assessment for video stream recognition systems / T. Chernov, N. Razumny, A. Kozharinov, D. Nikolaev, V. Arlazarov // Tenth International Conference on Machine Vision. — International Society for Optics, Photonics. 2018. — in press (Scopus, WoS).
10. Smart IDReader: Document Recognition in Video Stream / K. Bulatov, V. Arlazarov, T. Chernov, O. Slavin, D. Nikolaev // 14th IAPR International Conference on Document Analysis and Recognition. Vol. 6. — IEEE, 2017. — 39–44 (Scopus, WoS).
11. Diamond recognition algorithm using two-channel x-ray radiographic separator / D. P. Nikolaev, A. Gladkov, T. Chernov, K. Bulatov // Seventh International Conference on Machine Vision. Vol. 9445. — 2015. — 944507 (Scopus, WoS).
12. Chernov, T. S. A method of periodic pattern localization on document images / T. S. Chernov, D. P. Nikolaev, V. M. Kliatskine // Eighth International Conference on Machine Vision. Vol. 9875. — International Society for Optics, Photonics. 2015. — 987504 (Scopus, WoS).
13. Chernov, T. S. A Method Of Periodic Pattern Detection On Document Images / T. S. Chernov, D. P. Nikolaev, V. M. Kliatskine // 29th European Conference on Modelling and Simulation. — 2015. — 506–510 (Scopus).
14. Snapscreen: TV-stream frame search with projectively distorted and noisy query / N. Skoryukina, T. Chernov, K. Bulatov, D. P. Nikolaev, V. Arlazarov // Ninth International Conference on Machine Vision. Vol. 10341. — International Society for Optics, Photonics. 2017. — 103410 (Scopus, WoS).
15. Контроль и коррекция входного сигнала двухканального рентгенографического сепаратора / Д. Николаев, Т. Чернов, А. Гладков, А. Борисов // Информационные технологии и системы, сборник трудов. — 2014. — С. 202–209.
16. Чернов, Т. Метод поиска периодических фоновых элементов на изображении документа / Т. Чернов, Д. Николаев, В. Кляцкин // Информационные технологии и системы, сборник трудов. — 2015. — С. 400–412.

17. Алгоритм взвешенного поиска проективного оптического потока, устойчивый к бликам / Д. Шепелев, Е. Ершов, А. Терещин, Т. Чернов, Д. Николаев // Информационные технологии и системы, сборник трудов. — 2017.
18. Segments Graph-Based Approach for Document Capture in a Smartphone Video Stream / A. Zhukovsky, D. Nikolaev, V. Arlazarov, V. Postnikov, D. Polevoy, N. Skoryukina, T. Chernov, J. Shemiakina, A. Mukovozov, I. Konovalenko, M. Povolotsky // 14th IAPR International Conference on Document Analysis and Recognition. Vol. 1. — IEEE, 2017. — 337—342 (Scopus, WoS).

Патенты и свидетельства о регистрации программ для ЭВМ

19. Система распознавания изображений символов на основе обучающей выборки: патент РФ на полезную модель №161580 / В.Л. Арлазаров, К.Б. Булатов, Д.А. Ильин, Д.П. Николаев, Т.С. Чернов, А.В. Шешкус, опубл. 27.04.2016 по заявке №2015148233/08 от 10.11.2015.
20. Способ детектирования голограмических элементов в видеопотоке: патент РФ на изобретение №2644513 / В.В. Арлазаров, Д.П. Николаев, Н.С. Скорюкина, Т.С. Чернов, опубл. 12.02.2018 по заявке №2017106048 от 27.02.2017.
21. Программа для распознавания идентификационных карт личности «Smart IDReader»: свидетельство о государственной регистрации программы для ЭВМ №2016616961 / В.В. Арлазаров, Д.П. Николаев, С.А. Усилин, К.Б. Булатов, Т.С. Чернов, Д.Г. Слугин, Д.А. Ильин, П.В. Безматерных, А.А. Муковозов, Е.Е. Лимонова, опубл. 22.06.2016 по заявке №2016612014 от 10.03.2016.
22. Библиотека для распознавания в видеопотоке паспорта гражданина Российской Федерации «Smart PassportReader»: свидетельство о государственной регистрации программы для ЭВМ №2015616071 / В.В. Арлазаров, К.Б. Булатов, Д.А. Ильин, А.В. Куроптев, Д.П. Николаев, Д.В. Полевой, С.А. Усилин, И.А. Фараджев, Т.С. Чернов, опубл. 29.05.2015 по заявке №2015612880 от 10.04.2015.
23. Библиотека для создания компактных цифровых копий документа «Smart PDF/A»: свидетельство о государственной регистрации программы для ЭВМ №2015615714 / В.В. Арлазаров, А.Е. Жуковский, А.В. Куроптев, Д.П. Николаев, С.А. Усилин, Т.С. Чернов, опубл. 22.05.2015 по заявке №2015612889 от 10.04.2015.
24. Библиотека для распознавания машиночитаемых строк в видеопотоке «Smart 3D OCR MRZ»: свидетельство о государственной регистрации программы для ЭВМ №2015615712 / В.В. Арлазаров, К.Б. Булатов, А.Г. Волков, Д.А. Ильин, А.В. Куроптев, А.Е. Марченко, Д.П. Николаев, Д.В. Полевой, Т.С. Чернов, Ю.С. Чернышова, опубл. 22.05.2015 по заявке №201561288 от 10.04.2015.