

Анализ особенностей использования стационарных и мобильных малоразмерных цифровых видео камер для распознавания документов

В.В. Арлазаров, А.Е. Жуковский, В.Е. Кривцов, Д.П. Николаев, Д.В. Полевой

Аннотация. Статья посвящена актуальной проблеме распознавания текстовых документов с использованием стационарных и мобильных малоразмерных цифровых камер в качестве устройств получения изображения. Авторами показаны различия в постановках задач, связанных, во-первых, с природой устройств, особенностями процесса формирования изображения и, во-вторых, с возможностью использования для распознавания текста видеопотока взамен одиночных изображений. На примере документов удостоверяющих личность рассмотрены основные проблемы и новые задачи, возникающие в процессе распознавания документа.

Ключевые слова: распознавание текста, видеопоток, мобильные видеокамеры.

Введение

За последнее десятилетие технологии распознавания текстовых документов из актуальной научной задачи превратились в стандартные, надёжные функции информационных систем [1-3]. Более того, свое разрешение получили многие задачи обработки структурированных документов, и хотя эта тематика все еще насчитывает ряд нерешенных научных и технических проблем, технологии распознавания структурированных документов активно применяются для решения различных задач ввода.

Все эти технологии базируются на обработке полученного путем сканирования документа, в тоже время в современном мире наблюдается серьёзная тенденция к решению различных прикладных задач с использованием стационарных и мобильных малоразмерных цифровых видео камер, к которым можно в первую очередь отнести web-камеры и камеры мобильных устройств. Кроме того, за послед-

ние несколько лет мобильные устройства показали серьёзный прогресс как в области роста вычислительных ресурсов, так и в области повышения качества и разрешения фото- и видеосъемки, последнее можно отнести и к web-камерам. Все это дает возможность рассматривать такие устройства в качестве платформы для решения задач получения изображения и дальнейшего распознавания текстовых структурированных и не структурированных документов. В данной статье мы рассмотрим вопросы и проблемы возникающие в процессе использования фото- и видео-съемки в задачах ввода документов.

Обзор технологий получения изображений документов

Для начала рассмотрим физические отличия процессов получения изображений документов методами сканирования, фото- и видео-съемки и их влияние на конечный результат. Для чего

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 13-07-12171 офи_м и № 13-07-12172 офи_м

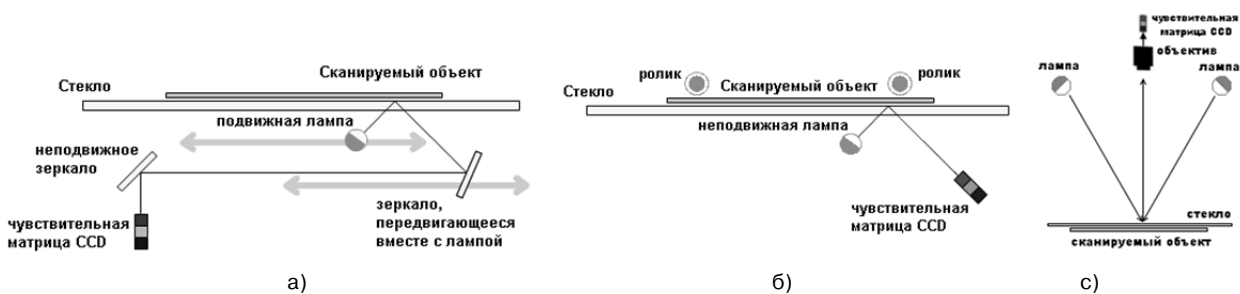


Рис. 1. Принципиальная схема формирования изображения для разных типов сканеров

а – планшетные, б - протяжные, с – планетарные

детально рассмотрим способы построения изображения в каждом случае, начав с процесса сканирования. В рамках рассмотрения мы будем изучать не только физические свойства регистрирующих устройств, но и особенности самих документов, сцены и процесса регистрации.

Особенности формирования изображения для сканирующих устройств

Без ограничения общности можно считать, что при сканировании документ расположен в перпендикулярной оптической оси плоскости на фиксированном расстоянии от регистрирующей матрицы [4]. Этим достигается гомотетичность исходного документа и его изображения, а незначительные искажения при небольших отклонениях от такого расположения (будут рассмотрены далее) легко детектируются и корректируются.

Для всех типов сканеров (Рис. 1) регистрируемый участок документа может считаться неподвижным во время экспозиции, поэтому исключены связанные со смещением исходного документа дефекты (размытие) изображения.

В планшетных устройствах сканируемый объект прижимается к стеклу и сверху накрывается специальной крышкой, которая обеспечивает правильное фоновое заполнение изображения вне зоны документа. При использовании протяжного механизма регистрирующая подсистема неподвижна, сканируемый объект (бумага) протягивается системой специальных роликов, при этом фоновое заполнение изображения вне зоны документа формируется специальным роликом белого или черного цвета. В планетарных сканерах используют специ-

альные столы, на фоне которых производится съемка документов. Таким образом, во всех типах сканеров использование собственной подложки для документа обеспечивает известные характеристики заполнения изображения за пределами зоны документа.

Освещение во всех типах сканеров формируется специальными мощными лампами подсветки, которые гарантируют стабильные характеристики и отсутствие теней. Известный материал фона и стабильность освещения позволяет использовать в сканерах функции автоматической самокалибровки для улучшения качества изображения.

Рассмотрим основные причины нарушения гомотетичности исходного документа и его изображения при сканировании. Для протяжных сканеров несогласованность протягивающего и регистрирующего механизма приводит к «растяжению/сжатию» изображения. Для планшетных сканеров наиболее типичным источником геометрических искажений и дефокусировки изображения является неплотное прилегание бумаги к стеклу, что часто происходит при сканировании сшитых в тома материалов и/или в области сгибов (Рис. 2).

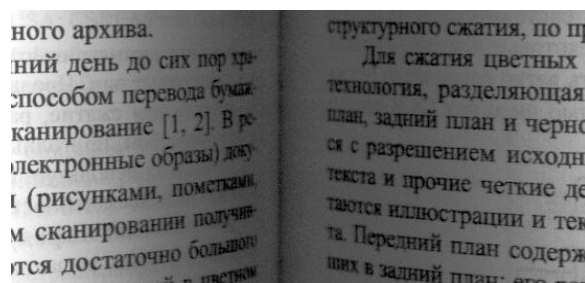


Рис. 2. Искажения (дефокус и изгиб строк) в области переплета книги

Дополнительным источником помех при сканировании на планшетном сканере является краевая засветка при открытой или неплотно закрытой крышке сканера.

Таким образом, сканер является специализированным устройством для получения высококачественных изображений. Образ документа при сканировании формируется в известной «лабораторной» сцене, все физические параметры которой зафиксированы и определены. Известны модели искажений и помех, поэтому сами дефекты изображений хорошо детектируются и в большинстве случаев устраняются программными методами.

Особенности формирования изображения для цифровых фото и видео устройств

Рассмотрим использование фото и видео камер для получения изображений документов, при этом будем использоваться общий термин «съемка камерой» [5]. При такой съемке с помощью объектива на светочувствительной матрице лучами отраженного от объекта съемки света формируется изображение документа (Рис. 3).

По сравнению со сканерами, оптическая схема камеры является более сложной и сама по себе вносит больше искажений вследствие аберраций, бликов и отражений внутри оптической системы.

Использование фотосенсоров (матриц) и аналоговой электроники устройствами для регистрации изображений неизбежно приводит к появлению искажений изображений, которые называют цифровой шум [6]. Источниками цифрового шума являются сам процесс оцифровки аналогового сигнала (ошибки квантования сигнала, тепловой шум и переноса заряда на матрице) и его дальнейшее усиление. Дополнительными источниками искажений явля-

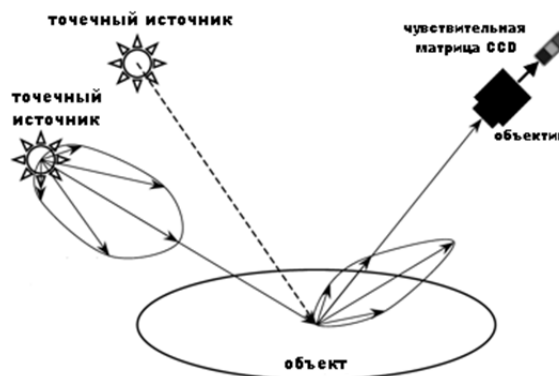


Рис.3. Принципиальная схем формирования изображения в фото и видео камерах

ются загрязнения матрицы и дефектные элементы сенсора. Еще один источник искажений - алгоритмы сжатия изображений, что особенно заметно для кадров видеопотока [7, 8]. Цифровой шум заметен на изображении в виде наложенной маски из пикселей случайного цвета и яркости. Шум более заметен на однотонных участках изображения, в особенности – на темных. В сканерах гарантировано качественное освещение, в отличие от камер, для которых часто возникает ситуация недостаточной освещенности и влияние цифрового шума (Рис. 4) многократно усиливается.

Рассмотрим разрешающую способность различных устройств формирования изображений документа. Типичным разрешением для систем оптического ввода документов является 300 точек на дюйм (dpi). Такое разрешение обеспечивается всеми современными сканерами и позволяет качественно распознавать печатный текст с обычными для печатных документов размерами шрифтов. При таком разрешении изображение страницы формата А4 имеет размеры 3394x2400 пикселей. Характерное разрешение для камеры мобильного устройства в фоторежиме (5 мегапикселей) дает размер кадра 2592x1936 пикселей. Таким образом при фотосъемке макси-

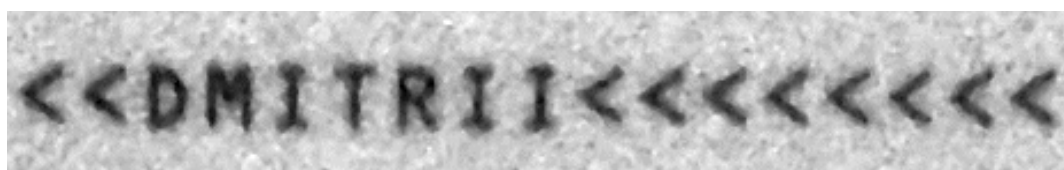


Рис.4. Цифровой шум в условиях низкой освещенности



Рис.5. Примеры а - слабого и б - сильного проективного искажения

мальное рабочее разрешение не превышает 200dpi (для формата А4 - 2262x1600 пикселей). Характерное разрешение web-камеры среднего уровня определяется режимом съемки FullHD с размером кадра 1920x1080 пикселей, что дает еще более низкое рабочее разрешение, которое можно оценить как примерно 120dpi (для формата А4 - 1358x960 пикселей).

В отличие от сканеров, при съемке камерой сам документ расположен в произвольной плоскости относительно плоскости сфокусированного изображения. Отклонение от перпендикулярной оптической оси плоскости приводит к проективному искажению (Рис. 5) изображения документа [9].

Предположения о выровненности, параллельности и перпендикулярности различных частей документа, например, изображений линий разграфки или строк текста часто используется в алгоритмах оптического распознавания текста [10, 11]. Такого рода свойства справедливы для исходных документов и выполняются для гомотетичных исходному документу изображений, получаемых при сканировании. Очевидно, что при съемке камерой углы и их отношения, а также пропорции объектов изменяются в зависимости от ракурса съемки. Это приводит к тому, что классические алгоритмы не могут применяться напрямую, а требуют предварительной проективной нормализации изображения.

Проективная нормализация основана на том, что съемка камерой хорошо моделируется

преобразованием центральной проекции. Матрицу преобразования плоскости оригинального документа на плоскость изображения можно легко восстановить исходя из знания точного соответствия между документом и его изображением всего 4-х точек. А из гомологии, при условии что найденные точки составляют прямоугольник, становится возможным определить отношение сторон оригинального документа. В большинстве случаев решается задача определения положения документа на изображении, а в качестве особых точек выбираются углы документа - таким образом одновременно находятся параметры проективной нормализации и границы документа, в которых ее нужно проводить.

В сканере фиксированы взаимное расположение документа и регистрирующей подсистемы, поэтому известен коэффициент подобия, а разрешение выходного изображения дает возможность вычислить точные размеры для различных элементов изображения. При съемке камерой коэффициент подобия с исходным документом неизвестен, поскольку определяется оптической схемой и взаимным расположением камеры и документа.

В зависимости от характеристик объектива и положения документа относительно плоскости наводки на резкость часть или все изображение документа может быть «размыто» (Рис. 6 а).

Подвижность самого документа или камеры, особенно в условиях недостаточной освещенности, приводит к ухудшению качества изображе-

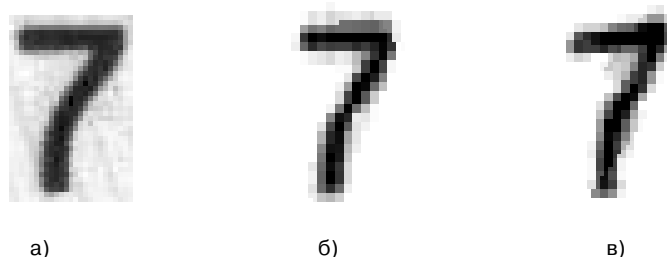


Рис. 8. Пример искажения изображения отдельного символа при механической деформации исходного документа

а - отсканированный символ, б – символ при слабом изгибе, в – символ при сильном изгибе



Рис. 9. Различные варианты организации сцены с точки зрения фона

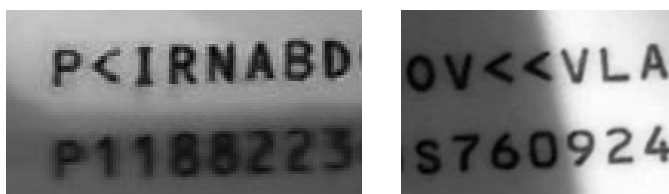


Рис. 10. Изображения документов при наличии затенения и перепада цветов

угольную форму - он может иметь скругленные, загнутые или обрезанные углы, дефекты прямых краев и деформации в плоскости документа. Сцена, в которой происходит съемка, может содержать в себе множество примитивов, не отличимых от границ документа и быть схожей по цвету или текстуре на заполнение документа (Рис. 9).

В таких условиях нахождение зоны документа и выделение текстовых строк само по себе становится довольно сложной задачей.

Схема освещения документа в сканерах минимизирует появление теней и бликов даже для «глянцевых» страниц документов. При съемке камерой в естественных сценах на изображениях часто возникают перепады яркости (тени, отражения, рефлексы и т.д.) и цветовые иска-

жения, которые усложняют задачи анализа изображений и распознавания, например, за счет потери существующих или появления фальшивых границ объектов (Рис. 10).

Дополнительно элементы защиты [12] документа часто содержат области с «голографическими» элементами (Рис. 11 б, в), которые тоже искажают изображение.

Особенности использования видеопотока в качестве источника изображений документа

Физический процесс формирования изображения документа для малоразмерной цифровой видеокамеры аналогичен формированию фотоизображения. Необходимое для оптиче-

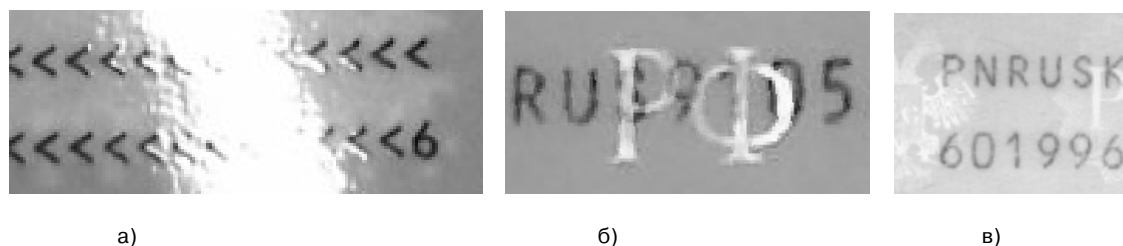


Рис. 11. Фрагмент зоны а – блик от протяженного источника света (лампа дневного света), б, в – «голографическая защита»

ского распознавания разрешение изображения можно считать высоким с точки зрения технологий видеосъемки, что приводит к необходимости передачи большого потока информации. Последнее решается использованием аппаратных и программных средств сжатия потока кадров [7]. Таким образом, при использовании в задаче оптического распознавания изображения документа в качестве источника самого изображения видеокамеры необходимо учитывать еще более низкое, по сравнению с использованием сканера или фотоустройства, разрешение отдельного кадра. При этом каждый отдельный кадр изображения подвергается упаковке/распаковке и может содержать артефакты использования алгоритмов сжатия с потерями. Одновременно с этим использование видеопотока позволяет использовать для распознавания последовательность изображений, на которых документ может располагаться под разными углами, в разных условиях освещения и при разных условиях фокусировки.

Новые задачи

Рассмотрим основные задачи, которые должны быть решены для получения высокого качества результатов распознавания при создании систем оптического ввода документов с использованием стационарных и мобильных малоразмерных цифровых видео камер в качестве источника изображения на примере документов, удостоверяющих личность.

В первую очередь должно быть определено множество логических типов документов, которые обрабатываются [13]. Логический тип документа определяется формально с точки зрения законодательства, что находит свое отражение в названии и составе реквизитов документа. Примерами логических типов являются

«паспорт гражданина Российской Федерации» или «водительское удостоверение». После этого необходимо исследовать вариативность экземпляров документов внутри одного логического типа и определить типы физические, которые отражают основные визуальные отличия. Необходимость выделения физических типов связана с тем, что часто одному документу соответствуют разные визуальные представления, например «старого» и «нового» образца. При анализе физических типов особое внимание необходимо обратить внимание на:

- различные варианты размеров и формы (например, прямые и скругленные края);
- различные варианты исполнения фона;
- материал самих страниц и их покрытий;
- расположение, ориентация, точность и качество печати реквизитов;
- наличие и расположение фотографии;
- наличие и расположение штриховых кодов;
- специальные методы (металлографический, орловский, ирисовый и т.д.) и материалы (метамерные краски) печати;
- использование специальных элементов защиты (голографические, гильоширные).

Все эти особенности могут приводить, во-первых к существенным различиям между изображениями документа между разными физическими классами, во-вторых, именно для документов с высокой степенью защиты характерны существенные изменения внешнего вида (наличия, формы и окраски элементов) на изображении при изменениях условий освещения и съемки. Основными результатами такого исследования должны стать выводы о границах вариативности для тех или иных элементов изображения документа, возможности появления бликов или засветок, а также изменения степени контрастности текста в различных условиях.

Для выбранных типов документов необходимо построить модель механических деформации исходных документов, провести анализ влияния таких деформаций на распознавание и выбирать некоторые уровни «хороших», «допустимых» и «неприемлемых» деформаций. Если документы исполняются на плотном пластике, то страницы практически не деформируются, поэтому возможными малыми изгибами можно просто пренебречь. Выполненные на бумаге документы подвержены «изгибам» и «скручиванию» (чаще всего вдоль или поперек основного направления чтения), причем иногда возникают «волны», когда изгибы разнонаправленны в разных местах страницы. Отдельный важный и довольно сложный случай возникает при распознавании разворота документа, когда между двумя страницами есть естественный сгиб, а сами страницы могут быть по-разному ориентированы в пространстве.

Различные варианты организации сцены и процесса съемки очень существенно влияют на качество изображений, поэтому обязательным этапом является анализ процесса видеосъемки. По результатам такого анализа определяются и формализуются границы применимости разрабатываемой технологии. Критическими пунктами являются:

- наличие и степень влияния на изображение особенностей оптической системы и регистрирующей подсистем (например, дисторсия, различные шумы, сжатие изображений и т.д.);
- расположение и крепление документа, камеры и источников освещения (в том числе допустимые диапазоны углов отклонений документа от плоскости фокусировки, диапазон расстояний от камеры до документа, скорость изменения различных характеристик);
- степень контролируемости фона и освещения.

Помимо общих вопросов желательно исследовать особенности конкретных моделей камер, например качество и скорость автофокусировки, цветопередача и т.д.

Распознавание документа в видеопотоке можно условно разделить на следующие этапы:

- детектирование документа в кадре;
- нахождения зоны и параметров проективного восстановления;

- проективное исправление изображения документа;
- позиционирование зон реквизитов на исправленном изображении;
- распознавание реквизитов;
- постобработка.

Эти задачи могут решаться как на отдельном кадре, так и на последовательности. В последнем случае добавляется возможность использовать межкадровую информацию для повышения качества и скорости работы.

Многие существующие методы (детектирования, распознавания и т.д.) в качестве входных данных используют полутоновые или даже монохромные изображения, поэтому важной задачей является правильная постановка и решение различных вариантов цветоредукции [14, 15] в условиях неизвестной сцены и ее освещенности, частным случаем которой является настройка методов бинаризации. В то же время вместе с задачей уменьшения размерности цветового пространства стоит задача восстановления истинных цветовых характеристик изображения документа, что требуется, например, для реализации эффективного сжатия с помощью технологий смешанного растрового представления [16] (Mixed Raster Content, MRC) или векторизации. С точки зрения использования малогабаритных камер в неконтролируемых условиях важной является оценка «четкости» и контрастности зон изображения, которые помогают решать различные задачи в условиях дефокусированных (частично или полностью) изображений и смазывания при движении

Детектирование наличия документа в кадре, идентификация его типа, а так же грубая локализация на изображении возможна без распознавания реквизитов [17], например, при помощи метода Виола-Джонс [18]. Этот же подход можно использовать для выделения реперных элементов. Примерами искомым объектов могут служить лица или логотипы, встречающиеся на документах. При этом часто необходимо детектировать объекты на изображении, как до проективной нормализации, так и после. Это позволяет учитывать их положение, как инварианта, для оценки найденного положения документа. Для получения детекто-

ров зачастую используются механизмы обучения машин. Одним из ключевых аспектов достижения должного качества распознавания является качественно подготовленная обучающая выборка. В идеальном случае она должна быть репрезентативна относительно генеральной совокупности рассматриваемых изображений. Описанные ранее особенности формирования изображения должны найти отражения и при построении обучающей выборки. Например, в задаче поиска объектов на проективно искаженных документах выборка должна содержать примеры одного и того же документа, снятого с разных ракурсов.

Практически всегда сбор и разметка данных для построения выборки требуют огромных человеческих усилий, но, тем не менее, редко покрывают пространство всевозможных параметров. А так как большинство искажений возникающих при съемке камерой (включая перспективное), могут быть смоделированы по изображению отсканированного документа, для решения этих проблем можно использовать искусственно синтезированные примеры [19]. Это позволит сократить затраты на разметку, точнее учесть особенности задачи и, при необходимости, быстро перестроить выборку. Подобный подход можно использовать для моделирования произвольных искажений: разной освещенностью, бликами, размытием объектов и пр.

Для точного позиционирования полей необходимо найти границы документа и определить проективный базис на изображении. Для этого требуется в условиях шума выделить линейные границы, углы, скругления и т.д., провести анализ точек схода, сгенерировать варианты и выбрать варианты границ документа, наиболее соответствующие модели. После определения проективного базиса необходимо проективно исправить зону изображения, уточнить положения реперных элементов на проективно исправленном изображении документа, произвести позиционирование и распознавание полей [20, 21].

Для распознавания текстовых реквизитов возможно требуется разработка специализированных методов оптического распознавания, как отдельных символов, так и текстовых фрагментов. Возможны постановки задачи рас-

познавания как для проективно искаженных изображений документов, так и для изображений после проективного исправления. Особенности является исходно достаточно низкое разрешение (не превышающее 150-200 dpi) при наличии помех и искажений, в частности бликов и засветки, дефокусировки изображения и смазывания. Ошибки определения проективного базиса могут существенно исказить геометрические характеристики символов. Все сказанное выше о возможности моделирования искажений и генерации изображений крайне актуально для изображений строк и символов, поскольку имеются объективные сложности получения достаточного количества образцов самих документов, например, удостоверяющих личность.

Заключение

Неблагоприятные для распознавания условия и высокая значимость информации в выбранном классе документов заставляют обратить особое внимание на разработку методов контекстного улучшения результатов распознавания, а также получения оценок точности и надежности [22-24]. Существенным вкладом в решение могут стать накопительные контексты распознавания, которые используют информацию о характеристиках изображений и результаты распознавания на разных кадрах.

Обработка каждого кадра видеопотока требует настолько больших вычислительных ресурсов, что в большинстве случаев либо невозможно, либо нецелесообразно. Это приводит к необходимости правильной постановки и решения задачи выбора оптимального подмножества кадров из последовательности, например, для получения заданных точности и надежности результатов распознавания за требуемое время.

Ввиду общей сложности задачи ввода документов и невозможности ее решения одним разработчиком особое значение приобретает правильный инструментарий [25, 26], который позволяет организовать протоколирование, наблюдение и анализ работы алгоритмов на отдельных кадрах и для последовательности. При этом правильная организация стендирования

и разметка эталонных данных может потребовать нетривиальных технических и организационных решений.

Литература

1. Полевой Д.В. Актуальные задачи создания систем массового ввода с использованием оптического распознавания для преобразования сложно структурированных бумажных документов в гибридных информационных системах // Системный анализ и информационные технологии: тр. Четвертой междунар. конф. (Абзаково, Россия, 17-23 августа 2011 г.): в 2 т. Т.2. Челябинск: Изд-во Челяб. гос. ун-та, 2011. С. 192-195.
2. Арлазаров В.В., Постников В.В., Шоломов Д.Л. Cognitive Forms - система массового ввода структурированных документов // Труды ИСА РАН. 2002 С.35-46.
3. Postnikov V.V., Sholomov D.L., Marchenko A.E. Flexi-Docs: The Template Driven Document Recognition Technology. Proceedings of the 6th German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW-6), 2003.
4. Image scanner. – URL: http://en.wikipedia.org/wiki/Image_scanner
5. Кошушин А. Геометрия камеры и структура движения // Компьютерная графика и мультимедиа. – URL: <http://cgm.computergraphics.ru/content/view/141>
6. Image noise. – URL: http://en.wikipedia.org/wiki/Image_noise
7. Тропченко А.А., Тропченко А.Ю. Методы сжатия изображений, аудиосигналов и видео. – СПб: СПбГУ ИТМО, 2009.
8. Compression artifact. – URL: https://en.wikipedia.org/wiki/Compression_artifact
9. Hartley R., Zisserman A. Multiple view geometry in computer vision // New York: Cambridge University Press, 2003.
10. Nagy G. Twenty years of document image analysis in PAMI // Pattern Analysis and Machine Intelligence, IEEE Transactions on In Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 22, No. 1. (2000), pp. 38-62
11. Арлазаров В.Л., Куратов П.А., Логинов А.С., Славин О.А. Алгоритмы поиска границ печатных символов, используемые при оптическом распознавании символов // Информационные технологии и вычислительные системы № 4, 2004, С. 59-70.
12. Способы защиты документов. – URL: <http://www.bnti.ru/showart.asp?aid=940&lvl=02.09.05>.
13. Арлазаров В.В., Кривцов В.Е., Полевой Д.В., Слугин Д.Г., Янишевский И.М. От графического образа к универсальному представлению Формы документа // Труды ИСА РАН, том 63, выпуск 3, 2013. С. 95-05.
14. Kim S.K., Nikolayev D.P. Method and apparatus for sectioning image into plurality of regions // US patent. Application No. 09/983032, Publication No. US 2002/0102017 A1, 2002, Issued 11.07.2006 No. 7076095, 25 pp.
15. Николаев Д.П., Карпенко С.М., Постников В.В. Преобразование цветного изображение в монохромное с сохранением градиентной структуры // Искусственные интеллектуальные системы и интеллектуальные САПР. Труды международной конференции IEEE AIS'04 и CAD-2004. М.: Изд-во Физматлит, 2004. Т. 2, стр. 320-323.
16. Queiroz R. De, Buckley R., Xu M. Mixed Raster Content (MRC) Model for Compound Image Compression. – 2000. – P. 1–12.
17. Usilin S., Nikolaev D., Sholomov D. Guilloche Elements Recognition Applied to Passport Page Processing // Proceedings of the 8th Open German-Russian Workshop "Pattern Recognition and Image Understanding" OGRW-8-2011, 2011. Pp. 303-306.
18. Усилин С.А., Николаев Д.П., Постников В.В. Идентификация изображений документов методами, предназначенными для поиска лиц // Информационные технологии и системы (ИТиС'10): сборник трудов конференции. М., 2010. С. 132–137.
19. Жуковский А.Е., Тарасова Н.А., Усилин С.А., Николаев Д.П. Синтез обучающей выборки на основе реальных данных в задачах распознавания изображений// Информационные технологии и системы (ИТиС'12): сборник трудов конференции. М., 2012. С. 377-382.
20. Усилин С.А., Николаев Д.П., Постников В.В. Быстрый алгоритм совмещения изображений документов в произвольной геометрической модели // Труды конференции "Информационные технологии и системы". - 2008. - С. 471 - 477.
21. Усилин С.А., Николаев Д.П., Постников В.В. Локализация, ориентация и идентификация документов с фиксированной геометрией на изображении // Труды Института системного анализа РАН. Обработка информационных и графических ресурсов / под ред. В.Л. Арлазаров. М.: КРАСАНД, 2010. С. 248–261.
22. Слугин Д.Г. Особенности контекстного распознавания российского заграничного паспорта в системе Cognitive Passport. // Труды Института системного анализа РАН, 2009. С. 174–183.
23. Арлазаров В.В., Булатов К.Б., Карпенко С.М. Метод определения надежности распознавания в задаче распознавания тисненых символов. // Труды ИСА РАН, том 63, выпуск 3, 2013. С. 117-122.
24. Арлазаров В.В., Малых В.А., Шоломов Д.Л. Распознавание изображений документов с использованием алгоритма «рулетки» // Труды ИСА РАН, том 63, выпуск 4, 2013. С. 35-38.
25. Полевой Д.В., Самойлов О.С. Задача контроля качества при создании и развитии систем оптического распознавания печатного текста. // Технологии программирования и хранения данных / Под ред. В.Л. Арлазарова, Н. Е. Емельянова, Труды Института системного анализа РАН; Т. 45. — М.: ЛЕНАНД, 2009. С. 251-259
26. Полевой Д.В., Самойлов О.С. Оценка качества сегментации изображений печатных документов на примере системы оптического распознавания текстов OpenOCR / Обработка информационных и графических ресурсов / Под ред. В.Л. Арлазарова. - М.: КРАСАНД, 2010. (Труды Института системного анализа РАН; т.58.) С. 164-171.

Арлазаров Владимир Викторович. Старший научный сотрудник ИСА РАН. Окончил МИСиС в 1999 году. Кандидат технических наук. Автор 10 печатных работ. Область научных интересов: распознавание образов, обработка изображений, системы массового обслуживания. E-mail: vva777@gmail.com.

Жуковский Александр Евгеньевич. Аспирант МФТИ. Окончил МФТИ в 2013 году. Область научных интересов: распознавание образов, анализ изображений, машинное обучение. E-mail: zhukovsky@phystech.edu

Кривцов Валерий Евгеньевич Старший научный сотрудник, декан МФТИ. Окончил МФТИ в 1971 году. Кандидат физико-математических наук. Автор 94 печатных работ и 3 монографий. Область научных интересов: информационные технологии и компьютерные науки, математическое программирование, математические модели и методы в экономике. E-mail: krivtsov@phystech.edu

Николаев Дмитрий Петрович. Заведующий сектором ИППИ РАН. Окончил МГУ им. М.В. Ломоносова в 2000 году. Кандидат физико-математических наук. Автор 116 печатных работ. Область научных интересов: быстрые алгоритмы обработки изображений. E-mail: dimonstr@iitp.ru

Полевой Дмитрий Валерьевич. Старший научный сотрудник ИСА РАН. Окончил МФТИ в 2004 году. Кандидат технических наук. Автор 10 печатных работ. Область научных интересов: методы искусственного интеллекта, оптическое распознавание, системы обработки документов. E-mail: dvpsun@gmail.com