

В. П. Фраленко

Локализация текстовых фрагментов на смешанном фоне: краткий научный обзор

Аннотация. В работе предлагается краткий научный обзор исследований, посвященных задаче выделения текстовых фрагментов на изображениях со смешанным фоном. Представлены применяемые схемы обработки данных, выделены достоинства и недостатки используемых методов и алгоритмов. Определены факторы, влияющие на качество получаемых результатов.

Ключевые слова и фразы: текстовый фрагмент, сегментация, локализация, оптическое распознавание, алгоритм.

Введение

Оптическое распознавание текстов (Optical Character Recognition, OCR) стало наиболее востребованным применением методов искусственного интеллекта в жизни человека. Оно используется в видео-корректорской и охранной деятельности, мультимедийных библиотеках, в бухгалтерском документообороте, картографии и пр. сферах деятельности. Искусственно наложенный текст (например, субтитры, подписи, пояснения) или естественный текст на видео/изображениях, как и текст на отсканированных бумажных документах (например, на картографических материалах), является важным источником информации. При хорошем качестве исходных данных (изображения с высоким разрешением, контрастные и с однородным фоном) с обнаружением и распознаванием текстовых фрагментов справляются практически все современные OCR-системы. Однако при наличии

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 13-07-00025 «Исследование методов анализа интегрированной текстовой, графической и речевой информации в системах интеллектуального управления динамическими объектами») и Отделения нанотехнологий и информационных технологий Российской академии наук (программа фундаментальных исследований «Интеллектуальные информационные технологии, системный анализ и автоматизация»).

© В. П. Фраленко, 2014

© ИНСТИТУТ ПРОГРАММНЫХ СИСТЕМ ИМЕНИ А. К. АЙЛАМАЗЯНА РАН, 2014

© ПРОГРАММНЫЕ СИСТЕМЫ: ТЕОРИЯ И ПРИЛОЖЕНИЯ, 2014

сложной фоновой структуры доля корректно обработанных текстовых областей существенно снижается. На результат также влияют следующие факторы: количество цветовых оттенков или градаций текста, пространственная ориентация, наличие на фоновом изображении фрагментов, похожих на текст. Цель настоящей работы — провести научный обзор методов и алгоритмов обнаружения текстовых фрагментов на изображениях со сложной фоновой структурой.

1. Основные этапы обнаружения текстовых блоков на изображении

Во всех системах оптического распознавания символов вначале выполняются то или иное улучшение качества и анализ изображения, поданного на обработку. В том числе применяются специальные фильтры восстановления поврежденных изображений, например, с помощью гипоеллиптической диффузии [1], фильтры, устраняющие смазы [2], и пр. В ходе этого этапа определяются области, рекомендуемые для распознавания, делаются оценки ориентации текста, выделяются отдельные символы и строки. На втором этапе выполняется работа по непосредственному распознаванию текста. Обобщающее сравнение методов распознавания, применяемых в исследуемой области, дано в работе [3]. В частности, выделены следующие методы: признаковые, эталонные (растровые), структурные и искусственные нейронные сети (ИНС); методы сравниваются по способу описания изображения, вычислительным затратам, инвариантности к искажениям и универсальности (по мнению автора, лишь эталонные (растровые) и структурные методы пригодны для анализа сложных классов изображений и сцен).

На первом этапе обработки изображения осуществляется выбор цветовой модели, предпочтение отдается тем, в которых под яркость выделяется отдельный цветовой канал, например, моделям YCbCr, HSV, HSV, YUV, LAB и др. За счет игнорирования яркости получается устойчивость к разным условиям освещения и снижается вычислительная сложность. Однако, во многих исследованиях авторы до сих пор работают с RGB-моделью. Получив необходимые данные, следует осуществить сегментацию изображения. Выделяют пороговые методы с фиксированным и адаптивным порогом; методы с наращиванием областей: центроидное связывание, слияние/расщепление и метод водоразделов; методы на основе использования контурной информации: градиентные методы, методы с вычислением вторых

производных и др.; текстурные методы: статистические и структурные. Достаточно полное описание основных известных классификаций методов сегментации представлено в работе [4]. Далее обычно осуществляется непосредственно распознавание текста в виде отдельных символов.

Известны следующие подходы к обнаружению текстов на изображениях:

- (1) На основе использования контурной информации (каждый символ обладает четко выраженной контурной структурой); для локализации текста тут используют такие приемы как скелетизация (skeletonization) [5], выделение краев (edge detection) и выделение углов (corner detection) [6], методы на основе инвариантных моментов (invariant methods) [7] и пр. В случае изображений со сложным фоном быстрая обработка данных, полученных на этапе предобработки, может представлять из себя нетривиальную задачу.
- (2) На основе цветовой информации (целевые текстовые области обладают однородными цветами/интенсивностью и удовлетворяют ограничениям по размеру и форме); популярными инструментами являются метод гистограмм (histogram method) [8], анализ главных компонент (connected component analysis) [9] и различные алгоритмы адаптивной бинаризации (adaptive binarization) — в том числе алгоритмы Niblack, Sauvola, Chistian, Bernsan, Otsu и пр. [6]. Методы позволяют работать с произвольными размерами шрифта и произвольной направленностью текста, однако плохо работают на изображениях со сложным фоном, шумом и нерезкостями, используют большое количество эвристик.
- (3) На основе анализа текстурной информации (текстовые зоны могут заметно отличаться от фона, что позволяет использовать различные частотные фильтры для «пирамиды» изображений); для выявления нужных зон могут использоваться как классические методы распознавания образов — метод опорных векторов (support vector machines), искусственные нейронные сети (neural networks) [10], экспертные системы (expert systems) и др., так и специальные, например, метод спектрографических текстур [11]. Подходы позволяют работать с изображениями со сложным фоном, обладают высокой вычислительной сложностью из-за необходимости масштабирования изображений. В основополагающих работах [12–14] среди прочего рассмотрены вопросы выделения

текстурных признаков.

Далее приводится информация о наиболее интересных решениях и методах выделения текстовых фрагментов на изображениях.

2. Известные решения, упоминаемые в открытых источниках

Программа Abbyy FineReader [15] предназначена для конвертации изображений документов в редактируемые документы. Однако, в случае смешанного фона у обрабатываемого изображения качество результатов обработки значительно снижается, кроме того, структура результата обработки может значительно отличаться от структуры оригинального документа. Распознавание ухудшается при больших углах поворота текстовых строк (более 15°). Программа плохо справляется с текстами, где разные блоки имеют разную ориентацию.

Программа Datacube MaxVision Toolkit [16] позволяет осуществлять поиск целевого образца, измененного с помощью аффинных преобразований, имеется поддержка нелинейной коррекции освещенности. В качестве недостатка можно назвать существенный уровень вычислительной сложности: в основе программы лежит метод нормализованной корреляции оттенков серого; время подготовки системы к работе составляет $O(n^9)$, где n — число пикселей в образце. Также необходимо отметить отсутствие гибкости настройки преобразований, вызванное дискретностью шкалы масштабов и углов; ряд ограничений, вызванных применением метода пирамиды с малым числом уровней. Программа не поддерживает работу с полноцветными изображениями [17].

В исследовании [18] представлены: а) метод, применяющий реконструирование контурного представления и яркостное и цветное распределения для выделения текстовых символов; б) метод, использующий карты яркостной разницы для подбора фильтра, размера и формы окна фильтрации; в) алгоритм формирования ограничивающего прямоугольника с расчетом плотности распределения элементов; г) алгоритмы морфологической обработки. Метод обнаружения текстовых зон при поиске символов размером менее 10 пикселей показал степень обнаружения 92,1%, а для символов размером более 10 пикселей — 97,3%. Распознавание образов текстовых символов реализовано на основе ассоциативной ИНС Хемминга. При этом процент распознавания зашумленных изображений букв со смещениями составил 80–83%, что было недостаточным. Предложенный алгоритм

псевдотрехмерного анализа повысил качество распознавания на 9% при тех же условиях [19]. При отсутствии смещений и искажений образов точность распознавания достигает 100%.

В работе [20] проведено сравнение ряда алгоритмов сегментации. Сравнение алгоритмов выполнялось на базе изображений университета Беркли (Berkeley Segmentation Dataset). Лучшие результаты были получены у алгоритмов JSEG, Canny и Roberts. Анализ результатов сегментации, полученных с помощью данных алгоритмов, показал, что, в отличие от JSEG, области, выделенные с помощью алгоритмов Canny и Roberts, содержат большое количество мелких отверстий, для удаления которых необходима дополнительная обработка. Поэтому для сегментации в работе в дальнейшем использовался алгоритм JSEG.

В исследовании [21] предложен метод локализации текстовых областей на изображениях со сложным фоном, основывающийся на применении сверточной нейронной сети (СНС) и многомасштабного представления изображения на основе дискретного вейвлет-преобразования. По обрабатываемому изображению движется окно с размерами 36×64 пикселя. Для выделенных фрагментов формируется пирамида изображений: для этого производится преобразование значений интенсивности фрагмента к диапазону $[-1, 1]$ и выполняется дискретное вейвлет-преобразование с разложением изображения на два уровня по каждой составляющей каналов R, G и B. Полученные уровни пирамиды обрабатываются заранее обученной СНС, классифицирующей отдельные участки изображения как содержащие текст или не содержащие. Точность классификации обучающей выборки составила 99,3%, а контрольной — 77,7%. СНС, как известно, весьма хорошо работает с такими данными. Низкую точность классификации контрольной выборки можно объяснить несовершенством применяемых алгоритмов предобработки и малым объемом обучающей выборки.

В работе [10] процесс выделения и распознавания фрагментов текста разбит на несколько этапов: подготовка эталонных изображений и обучение ИНС Хемминга; удаление фона волновым алгоритмом; маркировка целевых объектов и их распознавание. Полученные результаты сравниваются с результатами работы программы FineReader, которая ошибается при распознавании некоторых букв, например, заглавной буквы «И» и набранных курсивом «и», «е» и «ь». Программа уступает по времени обработки решению от АBBYY.

Автор исследования [22] предлагает воспользоваться нейронной сетью типа неокогнитрон. Модуль распознавания символов (библиотека TRecS_net), наряду с модулем разделения текста на знаки (библиотека TRecS_t2s) и модулем группировки знаков в строки и слова (библиотека TRecS_s21), входит в систему распознавания TRecS. Распознавание текстов с известными шрифтами осуществляется почти со 100%-ой точностью. Однако, на тесте с неизвестными шрифтами система распознала корректно лишь 73% символов. Корректное распознавание разных форм символов требует значительного расширения разнообразия обучающей последовательности, чтобы она покрывала все множество входных образов.

В исследовании [23] для выделения текстовых линий на сложном фоне используется штриховой фильтр: штриховые карты вычисляются по горизонтали, вертикали, левой и правой диагоналям. Далее извлекается 24-мерная функция для каждого раздвижного окна и используется метод опорных векторов для приблизительного обнаружения текстовых регионов. Используются статистические признаки, в том числе математическое ожидание, дисперсия, весовая энергия. На последнем шаге используется другой классификатор на основе метода опорных векторов — для проверки кандидатов в текстовые линии. Точность метода составила 95,8%, что значительно выше по сравнению с результатами, полученными в работах [24, 25].

В исследовании [26] представлен алгоритм отделения графики от текста, отличающийся от известных тем, что корректность результата отделения не зависит от формы текстовых/графических блоков. Сегментация на текст и графику выполняется на основе анализа геометрических и пространственных соотношений между текстовыми и графическими компонентами. Показано, что эти соотношения характеризуются двумя свойствами, названными свойством площади и свойством компактности текста. Первое свойство выражает различие между площадями связанных компонент текста и графики, а второе — различие в расстояниях между соседними текстовыми компонентами в пределах одного и того же блока и между текстовыми и графическими компонентами. Разработаны две программные системы. Первая нашла практическое применение при оцифровке карт и лесоустроительных планшетов. Вторая система предназначена для создания электронного архива таких документов, как журналы и газеты.

В статье [27] представлен следующий алгоритм детектирования

и локализации текста на изображениях: строится гауссова пирамида изображений (на каждом уровне пирамиды далее ищется текст определенного размера); применяется оператор Собеля для выделения вертикальных границ, слишком короткие отрезки удаляются; к вертикальным границам применяется морфологическая операция закрытия. После объединения результатов, полученных на разных уровнях пирамиды, делается адаптивная бинаризация. Время обработки изображения размером 640x480 пикселей на „iPhone 5“ порядка 0,3 сек. (использовалась пирамида с тремя уровнями).

В работе [28] представлено решение задачи сегментации цветных изображений печатных документов, состоящих, в том числе, из сплошных и градиентных фонов, символов, заголовков, букв, многострочных таблиц и графических элементов: на основе методов многомасштабной сегментации линейных срезов изображения; определения внутренних областей объектов малого размера и модифицированного метода водоразделов. Реализованная модификация преобразования водоразделов позволяет выполнять сегментацию изображения только на основе информации о компонентах градиента. Предлагаемый метод включает подход, позволяющий исключить излишнюю сегментацию изображения, характерную для метода водоразделов, и критерий проверки ошибочного слияния сегментируемых областей. Метод многомасштабного вычисления компонент градиента изображения позволяет получать комбинированные аппроксимирующие производные линейных срезов, корректно отображающие информацию как о резких, так и плавных границах объектов изображения. Метод аппроксимации производной линейных срезов изображения повышает точность сегментации объектов. Метод определения внутренних областей объектов малого размера обобщен до возможности обработки изображений, имеющих эффект неточного совмещения.

Представленный в [29] программный комплекс для распознавания рукописных и старопечатных текстов включает в себя инструмент, позволяющий выделять символы из цветных и черно-белых растровых изображений кириллических рукописных и старопечатных исторических памятников и группировать схожие по начертанию символы и надстрочные знаки в отдельные образы-кластеры для дальнейшей процедуры распознавания. Алгоритм осуществляет адаптивную бинаризацию, сегментацию на основе метода прорисовки контура [30] и непосредственно кластеризацию. На этапе кластеризации каждому из сегментов присваивается числовая метка и вычис-

ляется вектор характерных признаков; сегменты, степень совпадения векторов которых превышает заданное пороговое значение, объединяются в кластеры, далее распознаваемые ИНС путем анализа векторов признаков. Работа интересна тем, что архитектура нейронной сети автоматически выбирается/синтезируется исходя из извлеченных характеристик обрабатываемых данных.

В работе [17] представлен ряд решений задачи поиска целевого фрагмента в изображении. При этом фрагмент может быть повернут, растянут, отличаться по контрасту или цвету, встречаться в разных масштабах и т.д. Однако доработанный алгоритм, устойчивый к аффинным преобразованиям, на практике не реализовывался. Применяется метод, основанный на использовании опорных точек и лексикографического рекурсивного поиска. Предварительная обработка включает в себя низкочастотную и медианную фильтрацию и выделение контуров. Наиболее сложным этапом алгоритма является построение поисковой структуры, ассоциируемой с обрабатываемым изображением. После ее создания появляется возможность находить опорные точки за логарифмическое время. Также в работе было предложено несколько оригинальных модификаций классических алгоритмов (например, модификация алгоритма Брезенхейма в подзадаче быстрого поворота фрагмента изображения). Программное обеспечение работает в режиме, близком к реальному времени.

В работе [31] представлен метод особо контрастных пикселей для полутоновых изображений, основывающийся на подсчете количества особо контрастных соседей для каждого пикселя. По своей сути метод является эвристическим. Он содержит четыре этапа:

- конвертирование цветного изображения в полутоновое (на этом же этапе вычисляется пороговое значение для выделения наиболее контрастных пикселей);
- обход восьми соседей (для каждого пикселя подсчитывается число особо контрастных соседей, где оттенок серого превышает заданный порог; если это количество больше или равно трем или меньше восьми, то такой пиксель маркируется единицей, в противном случае нулем);
- сегментация изображения (осуществляется построчный (60 пикселей в строке-маске) обход из верхнего левого угла с единичным шагом; если на шаге обнаруживается более 24 единиц, то предполагается наличие текста, к маркеру пикселя прибавляется две единицы; в случае наличия 40 единиц подряд область считается

линией; на каждом шаге проверяется 15 впереди стоящих пикселей, если они все нули, то это конец блока строки);

- поиск блоков строк (по различным характеристикам происходит отсев блоков с шумом, например, по площади, занимаемой текстом, по соотношению сторон блока, по расстоянию до соседних блоков и пр.).

Возможности метода ограничены высотой, длиной и углом наклона текста, что не позволяет говорить об универсальности решения.

В работе [32] представлен алгоритм для сегментации изображения страницы с текстом. Решена задача соотнесения каждого элемента страницы к одному из двух классов — текст или рисунок. На первом шаге бинаризованное изображение разбивается на небольшие области. Для классификации отдельных областей используются следующие критерии: доля черных пикселей во всей области, величина разброса толщины отдельных элементов в области (нашел применение алгоритм, предложенный в [33]) и наличие чередующихся строк/междустрочий. Вероятность ошибки первого рода при классификации областей составила 0,04; вероятность ошибки второго рода — 0,11. Основной причиной ошибок является относительно высокая чувствительность метода к ориентации строк текста. Возможно улучшение алгоритма за счет добавления инвариантности относительно наклона строк текста в пределах исследуемого фрагмента, а также за счет использования дополнительных инструментов для уточнения границ.

Заключение

Для современных приложений, в том числе дополненной виртуальной реальности, крайне важно повышение показателей точности выделения текста в ситуации сложных графических сцен, характеризующихся неоднородностью фона, отсутствием четких критериев отличия фона от текста и большой вероятностью разнообразных искажений. Развитие интеллектуальных технологий поиска текста на фотографиях и видеокадрах позволит в дальнейшем усилить степень внедрения этих технологий в различные сферы деятельности.

Выполненный научный обзор затрагивает лишь часть проблем настоящей области исследований. При решении задачи локализации текстовых фрагментов на смешанном фоне возникает ряд подзадач,

универсальных решений для которых не существует. На результат обработки влияют многие факторы: например, особенности текста, его пространственная ориентация и пр. До сих пор остается популярной и востребованной цветовой модель RGB, хотя она и не обеспечивает устойчивость к разным условиям освещения.

Известные коммерческие решения не позволяют получить необходимое качество обработки целевых данных. Исследованные работы объединяет общая схема обработки: улучшение изображения, сегментация для обнаружения текстовых областей, кластеризация и непосредственно распознавание. Представленные в открытой печати алгоритмы с применением искусственных нейронных сетей и адаптивной бинаризации обеспечивают высокую долю обнаружения текстовых зон. Недостатком подходов с применением нейронных сетей является высокая вычислительная сложность, необходимость работы с большими объемами обучающих выборок и неоднозначности при выборе архитектуры ИНС. Также интересны и методы, использующие штриховые фильтры, показывающие практически такую же точность работы за счет применения метода опорных векторов для обработки полученных данных. Можно констатировать, что и другие подходы достойны внимания.

Список литературы

- [1] U. Boscaïn, J.-P. Gauthier, A. O. Ремизов. *Восстановление изображений с помощью гипоеллиптической диффузии*, 2013 (russian), URL <http://gct.math.nsc.ru/wordpress/wp-content/uploads/2013/10/Remizov.pdf>. ↑34
- [2] X. Li, J. Jiaya. *Two-phase kernel estimation for robust motion deblurring* // ECCV 1. Lecture Notes in Computer Science, 2010. Vol. **6311**, p.157–170. (english) ↑34
- [3] А. Г. Зотин. *Методы и алгоритмы обнаружения наложенных текстовых символов в системах распознавания изображений со сложной фоновой структурой*, Дис... кан. тех. наук, ГОУ ВПО Сибирский государственный аэрокосмический университет имени академика М.Ф. Решетнева, Красноярск, (2007). ↑34
- [4] С. В. Поршнева, А. О. Левашкина. *Универсальная классификация алгоритмов сегментации изображений* // Журнал научных публикаций аспирантов и докторантов, 2008. Т. **3**, с.163–172, URL <http://www.jurnal.org/articles/2008/inf23.html>. ↑35
- [5] С. В. Погодин. *Выделение и анализ скелетов объектов на цветных снимках* // Программные продукты и системы, 2009. Т. **2**, с.42–45. ↑35
- [6] Л. Шапиро, Дж. Стокман. *Компьютерное зрение*. М.: Бином. Лаборатория знаний, 2006. — 752 с. ↑35

- [7] Н. С. Абрамов, В. М. Хачумов. *Распознавание на основе инвариантных моментов* // Вестник РУДН. Серия Математика. Информатика. Физика, 2014. Т. 2, с. 142–149. ↑35
- [8] А. Н. Виноградов, Ф. В. Калугин, М. Д. Недев. *Выделение и распознавание локальных объектов на аэрокосмических снимках* // Авиакосмическое приборостроение, 2007. Т. 9, с. 39–45. ↑35
- [9] Н. В. Завалишин, И. Б. Мучник. *Модели зрительного восприятия и алгоритмы анализа изображений*. М.: Наука, 1974. — 344 с. ↑35
- [10] А. А. Талалаев, И. П. Тищенко, М. В. Хачумов. *Выделение и кластеризация текстовых и графических элементов на полутоновых снимках* // Искусственный интеллект и принятие решений, 2008. Т. 3, с. 72–84. ↑35, 37
- [11] В. П. Фраленко. *Анализ спектрографических текстур данных дистанционного зондирования Земли* // Искусственный интеллект и принятие решений, 2010. Т. 2, с. 11–15. ↑35
- [12] M. Sharma, M. Markou, S. Singh. *Evaluation of texture methods for image analysis* // Pattern Recognition Letters, 1980 (english), URL <http://www.eng.iastate.edu/ee528/Papers/texturecomparison.pdf>. ↑35
- [13] R. M. Haralick, K. Shan-mugam, I. Dinstein. *Textural features for image classification* // IEEE Trans. on Systems, Man and Cybernetics, 1973. Vol. 3, p. 610–621. (english) ↑35
- [14] K. I. Laws. *Rapid Texture Identification* // Image Processing for Missile Guidance. SPIE, 1980. Vol. 238, p. 376–380. (english) ↑35
- [15] *Официальный сайт программы ABBYY FineReader*, URL <http://www.abbyy.ru/finerreader/>. ↑36
- [16] *Программа MaxVision Toolkit фирмы Datacube*, URL <http://datacube.industrialpartner.com/products-m/maxvision-toolkit.htm>. (english) ↑36
- [17] И. В. Лобив. *Программные системы для идентификации и локализации объектов в изображениях*, Дис... кан. ф.-м. наук, Институт систем информатики СО РАН, Новосибирск, (2004). ↑36, 40
- [18] А. Г. Зотин. *Обнаружения текстовых зон в изображениях и кадрах видеопоследовательностей со сложной структурой фона*, 2013, URL <http://professorjournal.ru/PJGrantsPrograms/GrantMaterialsServlet?grantmId=2839760&grantmType=presentation>. ↑36
- [19] М. Н. Фаворская, А. Г. Зотин, Н. Д. Торгашин, М. В. Дамов. *Идентификация ограниченного набора образов на основе ассоциативных нейронных сетей* // Вестник Сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева, 2004. Т. 5, с. 109–116. ↑37
- [20] А. О. Левашкина. *Разработка методов поиска изображений на основе вычислительных моделей визуального внимания*, Дис... кан. тех. наук, ГОУ ВПО Уральский государственный технический университет, ГОУ ВПО Нижнетагильский институт (филиал) Уральского государственного технического университета, Новосибирск, (2009). ↑37
- [21] А. А. Николенко, О. Ю. Бабилунга, Т. К. Нгуен Тьен. *Локализация текстовых областей на изображениях с использованием сверточной нейронной сети* // Вестник НТУ «ХПИ», 2013. Т. 19, с. 121–127. ↑37

- [22] Р. Ижбулатов. *Модульная система распознавания текста*, 2009., URL <http://lrn.no-ip.info/other/Diplom/ППЗ.doc>. †38
- [23] X. Li, W. Wang, S. Jiang, Q. Huang, W. Gao. *Fast and effective text detection* // 15th IEEE International Conference on Image Processing (ICIP), 2008, p. 969–972. (english) †38
- [24] Q. Ye, Q. Huang, W. Gao, D. Zhao. *Fast and robust text detection in images and video frames* // Image and Vision Computing, 2005. Vol. **23**, no. 6, p. 565–576 (english). †38
- [25] Q. Liu, C. Jung, S. Kim, Y. Moon, J. Kim. *Stroke filter for text localization in video images* // Int. Conf. Image Process.— Atlanta, GA, USA, 2006, p. 1473–1476. (english) †38
- [26] О. Г. Окунь. *Сегментация изображений графических документов на основе метрических преобразований*, Дис... кан. тех. наук, Институт технической кибернетики НАН Беларуси, Минск, (1996). †38
- [27] А. Ермолаев. *Детектирование и локализация произвольного текста на изображениях, полученных с помощью фотокамер мобильных телефонов*, 2013, URL <http://habrahabr.ru/post/180609/>. †38
- [28] М. Д. Казанов. *Сегментация цветных изображений печатных документов*, Дис... кан. тех. наук, Институт системного анализа РАН, Москва, (2008). †39
- [29] С. Корниенко, Ю. Айдаров, Д. Гагарина, Ф. Черепанов, Л. Ясницкий. *Программный комплекс для распознавания рукописных и старопечатных текстов* // Информационные Ресурсы России, 2011. Т. **1**, URL http://www.aselibrary.ru/digital_resources/journal/irr/irr2725/number_1/number_12754/number_127542755/. †39
- [30] F. Chang, C.-J. Chen, C.-J. Lu. *A linear-time component-labeling algorithm using contour tracing technique* // Computer Vision and Image Understanding, 2004. Vol. **93**, no. 2, p. 206–220 (english). †39
- [31] А. Макаров. *Обнаружение текста. Метод «Особо контрастные пиксели»*, 2010, URL http://macarov.net/news/obnaruzhenie_teksta/2010-03-03-13. †40
- [32] С. В. Южиков. *Сегментация изображений страниц древних рукописей* // Электронные библиотеки: Перспективные Методы и Технологии, Электронные коллекции (RCDL).— Переславль-Залесский, 2007. Т. **1**, с. 236–240. †41
- [33] Е. В. Щепин, Г. М. Непомнящий. *К топологическому подходу в анализе изображений* // Геометрия, топология и приложения.— М.: Мин. высшего и средн. спец. образ. РСФСР, Московский институт приборостроения, 1990, с. 13–25. †41

Рекомендовал к публикации

к.т.н. Е. П. Куршев

Об авторе:



Виталий Петрович Фраленко

К.т.н., старший научный сотрудник ИПС им. А.К. Айла-мазяна РАН, автор более 60 публикаций. Область научных интересов: интеллектуальный анализ данных и распознавание образов, искусственный интеллект и принятие решений, параллельные алгоритмы, сетевая безопасность.

e-mail:

alarmod@pereslavl.ru

Образец ссылки на эту публикацию:

В. П. Фраленко. *Локализация текстовых фрагментов на смешанном фоне: краткий научный обзор* // Программные системы: теория и приложения: электрон. научн. журн. 2014. Т. 5, № 2(20), с. 33–45.

URL: http://psta.psiras.ru/read/psta2014_2_33-45.pdf

Vitaly Fralenko. *Localization of text fragments on mixed background: short scientific review.*

ABSTRACT. This paper proposes the short scientific review of research in the text fragments localization for images with mixed background. Presented applicable processing circuits, highlighted the strengths and weaknesses for used methods and algorithms. Detected factors that affecting on results quality. (in Russian).

Key Words and Phrases: text fragment, segmentation, localization, optical character recognition, algorithm.