

УДК 004.89:004.93

Николенко А.А., канд. техн. наук, доц.,
Котляр А.В., студент,
Тьен Т.К. Нгуен, магистр,
Одес. нац. политехн. ун-т

ОБНАРУЖЕНИЕ ТЕКСТОВЫХ ОБЛАСТЕЙ И ВЫДЕЛЕНИЕ СИМВОЛОВ НА ИЗОБРАЖЕНИЯХ С НЕОДНОРОДНЫМ ФОНОМ

А.О. Николенко, А.В. Котляр, Тьен Т.К. Нгуен. **Виявлення текстових областей та виділення символів на зображеннях з неоднорідним фоном.** Розглянута задача виявлення текстових областей та виділення символів на зображеннях з неоднорідним фоном. Запропоновано метод виявлення текстових областей та виділення символів на зображеннях з неоднорідним фоном на основі алгоритму перетворення по товщині штриха символу тексту з використанням перетворення по товщині штриха символу тексту. Запропонований метод дозволив виявити та виділити в середньому 85 % символів тексту на зображеннях з неоднорідним фоном незалежно від мови та розміру шрифту.

Ключові слова: обробка зображень, виділення символів, текстові області, перетворення по товщині штриха.

А.А. Николенко, А.В. Котляр, Тьен Т.К. Нгуен. **Обнаружение текстовых областей и выделение символов на изображениях с неоднородным фоном.** Рассмотрена задача обнаружения текстовых областей и выделения символов на изображениях с неоднородным фоном. Предложен метод обнаружения текстовых областей и выделения символов на изображениях с неоднородным фоном на основе алгоритма преобразования по толщине штриха символа текста. Предложенный метод позволил обнаружить и выделить в среднем 85 % символов текста на изображениях независимо от языка и размера шрифта.

Ключевые слова: обработка изображений, выделение символов, текстовые области, преобразование по толщине штриха.

A.A. Nikolenko, A.V. Kotlyar, Tien T.K. Nguyen. **Text regions detection and symbol extraction in images with non-uniform background.** The problem of text regions detection and symbol extraction in images with non-uniform background was considered. The method of text regions detection and symbol extraction in images with non-uniform background based on algorithm of the stroke width transform of text symbol was proposed. The proposed method can detect and extract on average 85% of text characters in images regardless of the language and font size.

Keywords: image processing, symbol extraction, text regions, stroke width transform.

С каждым днем возрастает количество воспринимаемой и анализируемой человеком текстовой информации, нацеленной на информирование его в различных жизненных ситуациях. Объявления, вывески, надписи, визитные карточки, номерные знаки транспортных средств — лишь краткий перечень форм представления такой информации. С ростом количества информации назревает и вопрос — как автоматизировать и улучшить процесс восприятия информации человеком с помощью информационных технологий? В таких условиях весьма актуальной является разработка методик и алгоритмов обработки визуальной информации с последующим использованием в информационных системах различного назначения. Одним из видов таких систем являются системы автоматического распознавания текста.

Базовой процедурой, применяемой на первоначальном этапе обработки в таких системах, является обнаружение текстовых областей. Существует множество алгоритмов и методов для

ее выполнения, однако до настоящего времени нет окончательного решения этой задачи.

Существующие методы обнаружения текстовых областей на изображениях можно условно разделить на три категории: текстурные, методы на основе областей и гибридные. Текстурные рассматривают текст как особый тип текстуры и используют для обнаружения такие его характеристики как локальная интенсивность или вейвлет-коэффициенты [1...3]. Эти методы в основном определяют текстовые области, строки и столбцы которых параллельны соответствующим координатным осям. Методы на основе областей [4, 5] сначала находят кандидатов в текстовые области, используя выделение контуров или кластеризацию, а затем проводят фильтрацию областей, используя эвристические правила для отсеивания лишних. Третья категория, гибридные методы, представляют собой смесь методов первых двух категорий [5, 6].

Перечисленные методы дают хорошие результаты лишь при экспериментальной оценке многих параметров (например, соотношение интенсивности текста и фона, размер шрифта и т.д.). Это отрицательно сказывается на быстродействии самих методов либо сужает область их применения. В то же время, нельзя утверждать, что на сегодняшний день уже исчерпаны все возможные средства для выделения отличительных особенностей текстовых областей, используемые для обнаружения последних на изображениях с последующим выделением символов.

В данной работе при анализе текстовых областей предлагается использовать такую характеристику как толщина штриха символа текста [3, 6]. Она основывается на априорных знаниях о том, что в пределах одной текстовой области (надписи, строки текста) толщина штриха символа текста и соотношение высоты/ширины символа остаются примерно неизменными.

Целью данной работы является разработка метода обнаружения текстовых областей и выделения символов на изображениях с неоднородным фоном, на основе алгоритма преобразования по толщине штриха символа текста.

Исходное изображение, как правило, может иметь ряд искажений из-за наличия шума, низкого разрешения, бликов, недостаточной или излишней экспозиции и т.д. Всё это снижает вероятность обнаружения текстовой области и последующего выделения символов на изображениях.

Для обнаружения текстовых областей и выделения символов на изображениях воспользуемся алгоритмом SWT (от английского Stroke Width Transform - преобразование по толщине штриха), первоначально предложенным в [6]. Данный алгоритм использует толщину штриха символа текста, которая практически не меняется в пределах текстовой области изображения и, следовательно, является подходящей характеристикой для обнаружения текстовых областей.

Общая схема метода обнаружения текстовых областей и выделения символов на изображениях с неоднородным фоном, основанного на методах и алгоритмах, предложенных в [3, 6], представлена на рис. 1.

Реализация метода требует выполнения следующих этапов: предварительная обработка изображения, SWT-преобразование, фильтрация и объединение компонент.

Этап 1. Предварительная обработка изображения.

Предварительная обработка изображения проводится по схеме (рис. 1): исходное изображение переводится из цветного в оттенки серого, затем полученное изображение сглаживается с использованием фильтра Гаусса апертурой 3×3 либо 5×5 .

После сглаживания изображения к нему применяется детектор границ Канни. Полученное с помощью детектора изображение инвертируется и сохраняется в памяти. В дальнейшем копия этого изображения сглаживается фильтром Гаусса апертурой 3×3 либо 5×5 с целью поиска нормали к контуру искомого символа в данной точке.

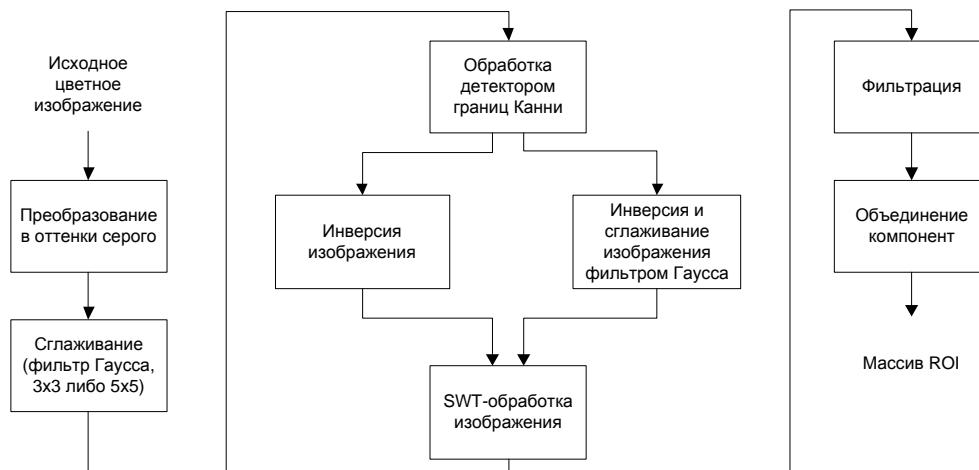


Рис 1. Общая схема метода обнаружения текстовых областей и выделение символов на изображениях с неоднородным фоном

Этап 2. Выполнение SWT-преобразования изображения.

SWT-преобразование изображения реализуется по следующему алгоритму (рис. 2).

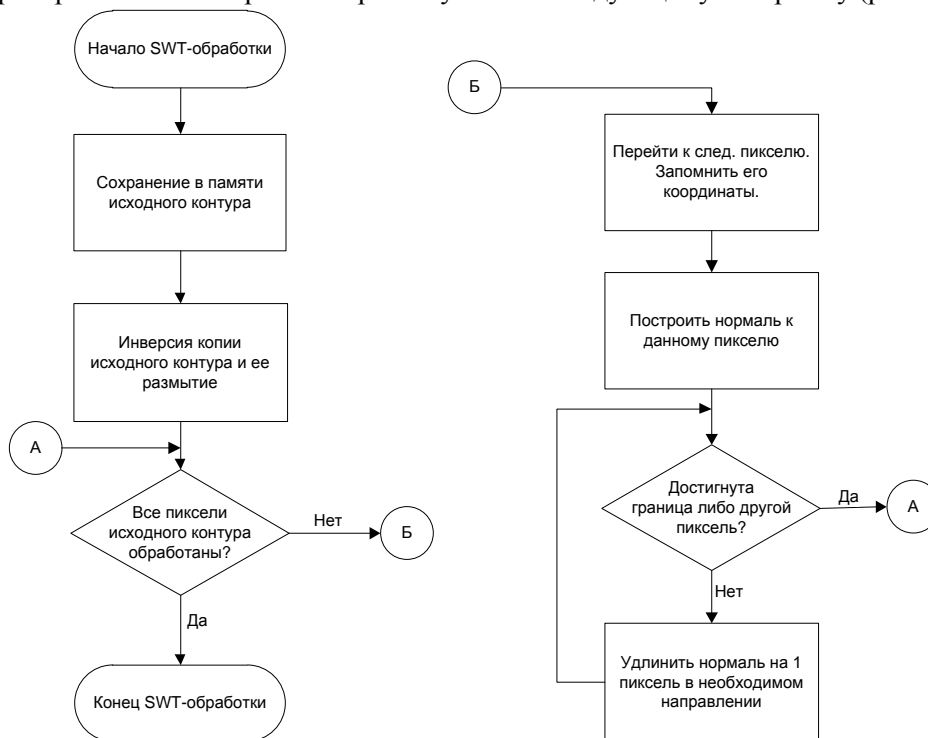


Рис. 2. Блок-схема алгоритма SWT-обработки изображения

Шаг 1. Выбирается пиксель контурного изображения. Его координаты сохраняются в памяти как P_1 . В зарезервированном изображении с размытым контуром находится пиксель, соответствующий этим координатам, его окрестности проверяются на наибольшую (если текст

светлее фона) либо наименьшую (в противном случае) интенсивность. Координаты этого пикселя также сохраняются в память (как P_2).

Шаг 2. Проводится луч из пикселя с координатами P_1 в направлении пикселя с координатами P_2 . Если построенный луч пересекает любой необработанный пиксель контура, либо достигает границы изображения – он заполняется пикселями определенной интенсивности, рассчитываемой исходя из длины получившегося отрезка.

Шаг 3. Повторяются шаги 1...2 до тех пор, пока каждый пиксель неразмытого контурного изображения не будет принадлежать какому-либо построенному лучу.

В результате данного этапа обработки получаем так называемое SWT-изображение. При этом текстовые области явно отличаются от любых других областей, что можно увидеть ниже, на рис. 3.

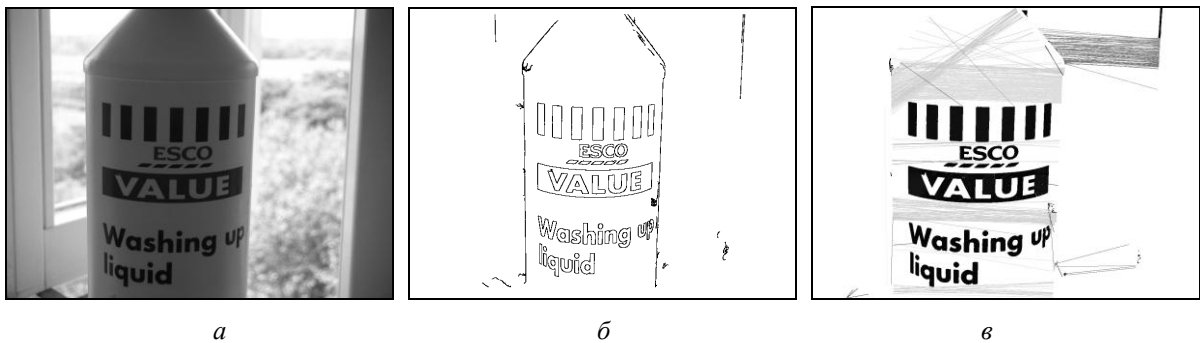


Рис. 3. Построение SWT-изображения: а — исходное изображение, б — инвертированное контурное изображение, в — SWT-изображение

Особенностью метода обнаружения текстовых областей на этапе построения SWT-изображения является учет зависимости соотношения средней интенсивности цвета шрифта (I_c) и фона (I_ϕ). При значении $I_c/I_\phi > 1$ цвет шрифта считается светлее фона, следовательно, лучи следует строить от контуров символа к фону (рис. 3, а). В противном случае направление градиента меняется в обратную сторону и лучи строятся по направлению от фона к контуру символа (рис. 3, б).

Использование двукратной SWT-обработки изображения для разного соотношения I_c/I_ϕ с последующим объединением полученных результатов позволяет улучшить точность обнаружения текстовых областей на неоднородном фоне.

Этап 3. Фильтрация полученного SWT-изображения.

В начале этапа любая пара контуров проверяется на наличие общих точек. Если два контура не имеют ни одной общей точки — они считаются отдельными. В противном случае они объединяются в один контур.

Для проверки принадлежности контура к символам рассчитывается среднее значение толщины штриха в этом контуре. Для каждого символа текста толщина штриха должна отличаться от среднего значения не более чем на 10%. При превышении данного предела символ

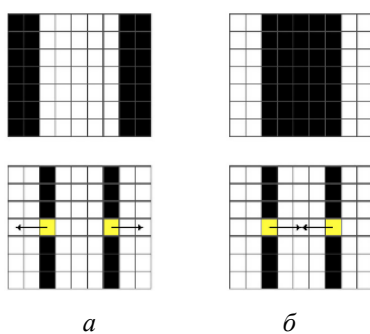


Рис. 4. Варианты выбора направления обработки контурного изображения: а — текст светлее фона; б — текст темнее фона

не будет считаться кандидатом на символ текста, соответствующая область изображения, ограниченная таким контуром, исключается из дальнейшей обработки.

Для выделенных символов проверяется соотношение пропорций по формуле

$$c = h/w,$$

где h — высота символа;

w — его ширина.

При значениях $c > 1,5$ символ исключается из дальнейшего процесса обработки.

Этап 4. Полученные после фильтрации компоненты объединяются в строки, согласно координатам полученных символов.

Если координаты центров каких-либо кандидатов на символ находятся на одной прямой либо прямоугольники, в которые вписаны выделенные символы, имеют общие стороны или отрезки, то рассматриваемые символы являются частью последовательности символов и должны быть объединены.

В конечном итоге получаем объединенные в компоненты символы, готовые к последующему процессу классификации либо распознавания. При этом фон удаляется с изображения, остается только массив текстовых областей (region of interest, ROI).

Для проверки эффективности предложенного метода, было разработано программное обеспечение (язык программирования C++, среда разработки Qt SDK, 4.8, библиотека OpenCV 2.4.2). Проведенное тестирование показало, что количество обнаруженных и впоследствии выделенных символов составляет 75...90 % от общего количества символов на изображениях. Снижение результатов правильного обнаружения и выделения символов связано с зависимостью метода в целом от параметров детектора границ Канни и сложностью получения качественного исходного изображения. Улучшить результат обнаружения и сегментации можно применением предварительной обработки изображения, но вычислительная сложность при этом увеличивается.

Ниже приведена таблица с процентным соотношением обнаруженных символов для шрифтов, чаще всего используемых в документообороте и наружной рекламе.

Зависимость количества обнаруженных символов от шрифта, используемого в надписи

Название шрифта	Символов обнаружено, %
Times New Roman	74,6
Arial	85,0
Calibri Regular	90,0
Free Sans	86,2
MS Sans Serif	85,1
Courier Regular	85,0
Helvetica Regular	84,1

Количество обнаруженных и выделенных символов зависит от особенностей алгоритма и шрифтов. В частности, символы шрифтов с засечками и переменной толщиной штриха (большие кегли Times New Roman), обнаруживались намного хуже. Этот недостаток могут частично устранить морфологические операции. Применение операции дилатации для светлых, либо эрозии - для темных шрифтов увеличивало количество обнаруженных символов на 10...15 %.

Предложенный метод обнаружения текстовых областей и выделения символов с использованием SWT-преобразования позволяет обнаружить и выделить в среднем 85 % символов на изображениях независимо от языка и размера шрифта, что доказывает его работоспособность в системах автоматического распознавания текста.

Использование двукратной SWT-обработки для разного соотношения средней интенсивности цвета шрифта и фона позволяет улучшить точность обнаружения текстовых областей на неоднородном фоне.

Литература

1. Шапиро, Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман. — Спб.: Бинум. Лаборатория знаний. — 2006. — 752 с.
2. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. — М.: Техносфера, 2006. — 1072 с.
3. Cong, Y. Detecting Texts of Arbitrary Orientations in Natural Images [Electronic resource]. / [Y. Cong, B. Xiang, L. Wenyu and others]. — Lab of Neuro Imaging and Department of Computer Science, UCLA — http://www.loni.ucla.edu/~ztu/publication/cvpr12_textdetection.pdf — 10.05.2012.
4. Николенко, А.А. Локализация характерных фрагментов изображения на основе двумерных вейвлет-фильтров. / А.А. Николенко, О.Ю. Бабилунга, В.Н. Зайковский // Вісник нац. техн. ун-ту «ХПІ». Зб. наук. праць. — 2011. — № 36. — С. 122 — 127.
5. Николенко, А.А. Обнаружение текстовых областей в видеопоследовательностях / А.А. Николенко, Тьен Т.К.Нгуен // Искусственный интеллект. — 2012. — № 4. — С. 227 — 234.
6. Epshtein, B. Detecting Text in Natural Scenes with Stroke Width Transform [Electronic resource] / B. Epshtein, E. Ofek, Y. Wexler. — Microsoft Corporation, 2009. — <http://research.microsoft.com/pubs/149305/1509.pdf>. — 2010. — 23.12.12.

References

1. Shapiro, L. Komp'yuternoe zrenie [Computer vision] / L. Shapiro, Dzh. Stokman. — St.-Petersburg. — 2006. — 752 p.
2. Gonsales, R. Cifrovaja obrabotka izobrazhenij [Digital image processing] / R. Gonsales, R. Vuds. — Moscow. — 2006. — 1072 p.
3. Cong, Y. Detecting Texts of Arbitrary Orientations in Natural Images [Electronic resource] / [Y. Cong, B. Xiang, L. Wenyu and other]. — Lab of Neuro Imaging and Department of Computer Science, UCLA — http://www.loni.ucla.edu/~ztu/publication/cvpr12_textdetection.pdf — 10.05.2012.
4. Nikolenko, A.A. Lokalizacija harakternyh fragmentov izobrazhenija na osnove dvumernyh wavelet-fil'trov [Localization of specific image fragments based on two-dimensional wavelet filters] // A.A. Nikolenko, O.Yu. Babilungaa, V.N. Zajkovskij // Visnik nac. tehn. un-tu «KhPI». Sb. nauk. prac. — 2011. — # 36. — pp. 122 — 127.
5. Nikolenko, A.A. Obnaruzhenie tekstovih oblastej v videoposledovatelnostyah [Text regions detection in video frames] / A.A. Nikolenko, Tien T.K. Nguyen // Iskusstvennij intellect [Artificial intelligence]. — 2012. — # 4. — pp. 227 — 234.
6. Epshtein, B. Detecting Text in Natural Scenes with Stroke Width Transform [Electronic resource] / B. Epshtein, E. Ofek, Y. Wexler. — Microsoft Corporation, 2009. — <http://research.microsoft.com/pubs/149305/1509.pdf>. — 2010. — 23.12.12.

Рецензент канд. техн. наук, доц. Одес. нац. политехн. ун-та Болтєнков В.А.

Поступила в редакцию 24 декабря 2012 г.