УДК 004.93'14

А.В. Гороховатский

Харьковский национальный экономический университет им. С. Кузнеца

ДЕТЕКТИРОВАНИЕ ТЕКСТОВЫХ ОБЛАСТЕЙ НА ИЗОБРАЖЕНИИ ДОКУМЕНТА МЕТОДОМ СЛИЯНИЯ

В данной статье представлен метод автоматического детектирования текстовых областей на изображении документа с помощью слияния областей. Предложенный метод комбинирует простоту реализации алгоритма слияния с анализом и отсечением графических областей, которые не несут текстовой информации. Проведено моделирование работы предложенного метода, выделены классы практических задач и условия наиболее эффективного его применения.

Ключевые слова: текстовая область, графическая область, слияние, изображение документа, пороговое значение, преобразование Хафа.

Введение

Задача распознавания документов в современном мире постоянного конвертирования информации, хранящейся в традиционной форме, в соответствующий электронный вариант приобрела широкое распространение. При распознавании документа логичным этапом видится создание общей структуры документа (page layout analysis), распознавание текстовых элементов (OCR – optical character recognition), формирование таблиц, перенос изображений и формул и т.п.

Основную информационную составляющую при конвертировании информации в электронную форму образует текстовое наполнение. Реализацией задачи конвертирования текста на изображении в соответствующий электронный текст занимается OCR, в то же время, существующие методы решения данной проблемы ориентируются на распознавание именно текста, а не графической информации внутри документа – такой как схемы, диаграммы, чертежи. Распознавание же этих элементов с помощью существующих механизмов OCR чаще всего приводит к генерации лишнего текста, символов, которых в реальности в документе не существует.

Данная статья посвящена физическому анализу [1] структуры документа, т.е. такому анализу, который позволяет выделить такие физические элементы документа, как текст, графические объекты, таблицы. Существующие методы сегментации часто базируются на анализе проекций [2, 11], связных компонент [3], фильтрации изображения, анализе границ и частот смены цвета [3, 11], слиянии-расщеплении элементов изображения и т.п.

Как логический, так и физический анализ документа может быть выполнен одним из двух наиболее распространенных подходов [4–6, 12]: «снизу вверх» («bottom-up» – анализ мелких элементов изображения с последующим их слиянием, образуя при этом элементы более высокого логического уровня: блоки, параграфы, абзацы и т.п.) либо «сверху вниз» («top-down» – разбиение изображения на абзацы, блоки с их последующим анализом и разбиением) [4, 5, 7, 8]. Предложенный подход подразумевает использование стратегии «снизу-вверх», используя слияние элементов нижнего уровня.

Статья посвящена разработке метода, который комбинирует простоту реализации стратегии слияния с эффективным детектированием текстовых областей, текстовой информации внутри таблиц и графических областей, и является адаптируемым к структуре изображения документа. Большинство нюансов реализации данного метода являются универсальными и позволяют обрабатывать изображения документов различного типа.

Предварительная обработка

Современные методы получения изображения документа, такие как сканирование или фотографирование, позволяют получить цветное или полутоновое изображение различного уровня четкости и качества. С практической точки зрения, целесообразно избегать обработки полноцветных и даже полутоновых изображений при решении таких проблем, в которых возможно пренебрежение подобной информацией. При обработке документа мы в дальнейшем будем оперировать понятиями «значимых» объектов (текста, таблиц, графических объектов) и «фона», разделить же сцену на изображении на фон и объект позволяет бинаризация. Следует отметить, что к проблеме выбора порога для бинаризации необходимо

подходить весьма серьезно, когда речь идет о практических задачах, требующих высокоточной обработки, таких, как собственно OCR, сегментация с целью измерения биологических объектов и т.п. При решении рассматриваемой проблемы определения текста мы не уделяем проблеме выбора порога особого значения в предположении, что изображение документа является высококонтрастным в том понимании, что текст от фона значительно отличается по яркости. В качестве значения порога бинаризации было выбрано толерантное значение в половину максимальной яркости, т.е. в порог был установлен в $\delta = 128$.

Кроме бинаризации часто также используют эквализацию гистограммы для улучшения контраста, фильтры шумоподавления (размытие, медианный фильтр), морфологические операции для устранения шумов и т.п.

В данной статье мы говорим об изображении документа, и о разбиении его на области. Под областью мы будем понимать атомарный объект, контент которого является однородным, т.е. содержит либо текст, либо изображение, либо таблицу и т.п. Область, содержащую какой-либо графический объект внутри изображения документа (схему, чертеж, изображение и.т.п) мы будем называть графической областью, область с контентом в виде текста – текстовой областью.

Выделение потенциальных областей интереса

Метод анализа изображения документа «снизу-вверх» предполагает слияние мелких элементов с образованием более крупных [9]. Первоначальным этапом является сканирование изображения с помощью сканирующего окна, которые дает возможность определить минимально возможные области, содержащие значимые объекты, для последующего слияния. В качестве таких областей на первом этапе обработки предлагается использовать классическую окрестность каждого пикселя, который потенциально может быть частью текстовой области. Чаще всего при выполнении подобных операций принято использовать сканирующее окно размером 3x3 пикселей, но в нашей задаче было использовано окно размера 5x5, поскольку большее окно позволяет ускорить процесс обработки.

Результатом обработки на этом этапе является множество областей, каждая из которых имеет площадь, равную сканирующему окну и расположена вокруг значимых пикселей, яркость которых удовлетворяет условию бинаризации (рис.1).

Переход на более высокий иерархический уровень выполняется путем слияния пересекающихся квадратных областей с образованием произвольных прямоугольных регионов (рис.2).

Рис. 1 Результат сканирования изображения плавающим окном размера 5x5

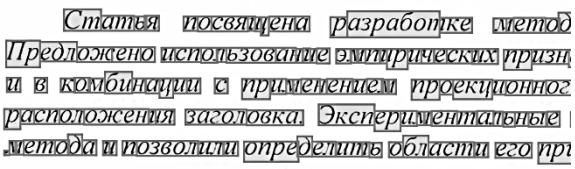


Рис. 2 Результат слияния пересекающихся областей

Результатом подобного технического слияния без анализа контента являются области, которые могут содержать контент типа, не являющегося потенциально интересным, например, линии разметки документа. Подобные области могут быть идентифицированы по нестандартному отношению высоты области к ширине либо с использованием других шумоподавляющих фильтров. Например, фильтрация областей по отношению $height_{r1} / width_{r1} > 0.02$, где r_1 – анализируемая область, $height, width$ – ее высота и ширина соответственно, позволяет получить результат, представленный на рис. 3.

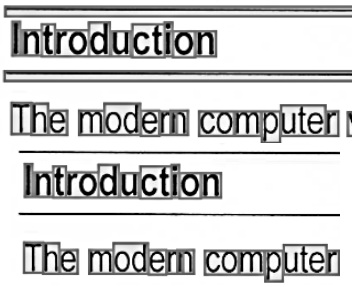


Рис. 3 Результат фильтрации потенциальных текстовых областей по отношению высоты к ширине: вверху – графическое отображение областей до фильтрации, внизу – графическое отображение областей после фильтрации.

Слияние квадратных пересекающихся областей позволяет сформировать совокупность прямоугольных областей, которые, в зависимости от выбранного ранее размера плавающего окна сканирования и особенностей шрифта могут содержать одну или несколько букв текста. Для получения области с контентом более высокого иерархического уровня (например, слова)

выполняется слияние тех областей, которые пересекаются в рамках допустимого радиуса δ_1. Для этого каждая из соседних областей увеличивается по горизонтали и вертикали на значение радиуса δ_1, после чего пересекающиеся области объединяются. Необходимость увеличения области на значение δ_1 по вертикали объясняется возможным соседством двух областей, которые содержат символы разного размера и расположены выше либо ниже базовой линии текста. Графическая интерпретация данного критерия приведена на рис. 4.

Рис.4 Графическая интерпретация критерия пересечения областей по радиусу δ_1

На рис.5 показан результат слияния набора текстовых областей с использованием различных пороговых значений.

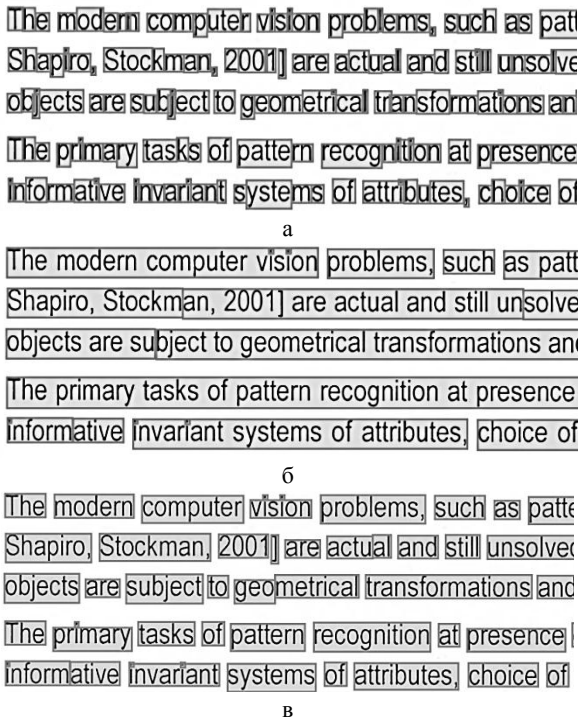


Рис.5 Результат слияния текстовых областей по горизонтали: а – графическое отображение областей перед слиянием, б – графическое отображение областей после слияния при $\delta_1 = 15$, в – графическое отображение областей после слияния при $\delta_1 = 5$.

Как видно, выбор значения δ_1 определяет строгость правила формирования единственной текстовой области в виде строки. При выборе значения порогового ограничения важную роль может иметь априорная информация о структуре изображения документа, например, δ_1 непосредственно указывает, могут ли быть объединены в единственную область такие области с текстом, которые находятся в одной горизонтальной физической строке, но в разных столбцах документа.

Выделение изображений

Разделение текстовых и графических областей при выполнении анализа документа является нетривиальной задачей, требующей индивидуального подхода к каждому типу изображения документа. Существует два основных класса сложности методов решения подобной проблемы, первый из которых оперирует количественными характеристиками области, такими, как плотность пикселей интереса, геометрические пропорции размеров [10, 15] и т.п., второй – использует более трудоемкие методы с большей точностью, которые опираются на глубокий анализ области и использование априорной информации о возможных различиях [13, 14].

В [15] предложено считать текстовой областью такую область, для которой выполнено условие:

$$0.045 < ratio_{r1} < 0.444,$$

где $ratio_{r1} = (Total\ black\ pixels) / (Total\ pixels)$ для области r_1, при этом попадание за пределы минимальной границы соответствует пустой области, попадание за пределы максимальной – графической области.

Тем не менее, подобное условие может быть выполнено для большинства изображений документов лишь в тех случаях, если графический элемент внутри области r_1 является достаточно контрастным и при этом его предварительная обработка (особенно бинаризация) сохраняет этот контраст.

В качестве обобщения предложенного метода может быть рассмотрено использование предварительной информации об области, содержащей изображение. Одной из самых легкодоступных в вычислительном плане характеристик является анализ физических размеров области, поскольку для множества изображений документов графические области значительно отличаются по размеру от текстовых областей.

Эмпирические исследования показали, что целесообразным может быть использование следующих условий для идентификации области r_1, содержащей изображение:

$$(ratio > 0.4 \vee ratio < 0.1) \wedge width_{r1} < 2.5 height_{r1},$$

где $width_{r1}$, $height_{r1}$ – ширина и высота прямоугольной области соответственно. Таким

образом, к графической области на изображении документа предъявляются требования повышенной или пониженной плотности точек, а также несоответствие размерам типичной текстовой области, для которой ширина значительно больше высоты.

Слияние областей по горизонтали

Получив множество текстовых областей после отсеивания графических, рассмотрим процедуру слияния тех текстовых областей, которые пересекаются в горизонтальном направлении. На рис. 6 приведена графическая интерпретация оценивания параметров d_1, d_2, по которым принимается решение об объединении. Значения определяются по двум прямоугольным областям r_1, r_2, каждая из которых имеет параметр top, определяющий значение ординаты верхней стороны, и параметр $bottom$, соответствующий ординате нижней стороны области: $d_1 = |top_{r1} - top_{r2}|$, $d_2 = |bottom_{r1} - bottom_{r2}|$. В качестве критерия принятия положительного решения объединения предложено использовать дизъюнкцию $(d_1 \le \delta_2) \vee (d_2 \le \delta_2)$, где δ_2 – некоторое пороговое значение, указывающее на возможное отклонение слов текста от горизонтальной оси.

Введем также в рассмотрение дополнительное пороговое ограничение δ_3, указывающее на максимально возможное горизонтальное расстояние между двумя текстовыми областями: $\delta_3 = left_{r1} + width_{r1} - left_{r2}$, где r_1, r_2 – анализируемые области, $left$ – абсцисса левой границы области, $width$ – ширина области. Параметр δ_3 дает возможность варьировать расстояние между текстовыми областями, определяющими слова.

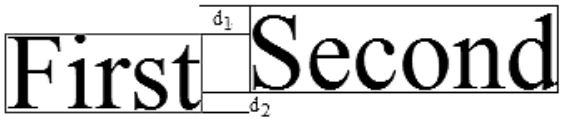


Рис.6 Графическая интерпретация значений d_1, d_2

Удаление таблиц

После выполнения всех предыдущих этапов выбор порога δ_3 с последующим слиянием может привести к тому, что таблицы на изображении документа будут определены как единый элемент (рис. 7, вверху). Для решения подобной проблемы целесообразно использовать метод для удаления линий, после чего выполнить повторное разбиение областей внутри исходной области с таблицей.

Известным методом детектирования линий на изображении является использование преобразования Хафа [16, 17]. Идея преобразования состоит в том, что каждый из значимых пикселей на изображении «голосует» за всевозможные прямые линии, частью которых он может являться. Соответственно, накопление «голосов» всех пикселей в массиве-накопителе с последующим поиском в нем глобальных и локальных максимумов позволяет определить прямые линии на изображении, за которые было отдано наибольшее число «голосов».

С практической точки зрения использование преобразования Хафа эффективно лишь для поиска линий на относительно небольших изображениях в силу его значительной трудоемкости (близка к полному перебору). Кроме того, классическая реализация преобразования Хафа позволяет определять идеальные линии, которые в результате оцифровки документа практически никогда не сохраняются.

Применение преобразования Хафа сопряжено с некоторыми нюансами рассматриваемой предметной области, которые связаны с обработкой текста. С практической точки зрения массив-накопитель содержит указатели на параметры линий, которые могут не являться частью таблицы, а проходить, например, сквозь текст. Отслеживание и отсечение подобных ситуаций можно реализовать с помощью дополнительных условий, выдвигаемых к найденной линии, подлежащей удалению, а именно: значение длины линии (его эквивалентом является значение текущего максимума в массиве-накопителе) должно превышать половину максимально возможной длины линии в этом направлении.

Практические исследования показали, что целесообразным может быть удаление всех линий в прямоугольной области, которые превышают значение в 60% от глобального максимума, но не более 20 таких линий. Поскольку алгоритмически преобразование Хафа определяет линии на изображении без анализа его контекста, то в каждом из прямоугольников также могут быть удалены линии, являющиеся частью текста, например, четко выраженные вертикальные линии в буквах “l”, “I”, “K”, “H” и т.п. латинского алфавита. Для устранения последствий подобного удаления предлагается использовать слияние областей по горизонтали (аналогично ранее рассмотренному этапу при поиске графических областей).

Результаты удаления вертикальных и горизонтальных линий с помощью преобразования Хафа с последующим повторным разбиением изменившейся области представлены на рис. 7 внизу.

В качестве последнего этапа обработки выполняется объединение текстовых областей, которые расположены друг над другом и имеют одинаковую ширину (в рамках погрешности δ_4), что соответствует объединению физических строк текста в вертикальном направлении в единственную область.

Type	Description	Example
object	The ultimate base type of all other types	object o = null;
string	String type; a string is a sequence of Unicode characters	string s = "hello";
sbyte	8-bit signed integral type	sbyte val = 12;
short	16-bit signed integral type	short val = 12;
int	32-bit signed integral type	int val = 12;
long	64-bit signed integral type	long val1 = 12; long val2 = 34L;
byte	8-bit unsigned integral type	byte val1 = 12;
ushort	16-bit unsigned integral type	ushort val1 = 12;
uint	32-bit unsigned integral type	uint val1 = 12; uint val2 = 34U;

ulong	64-bit unsigned integral type	ulong val1 = 12;
		ulong val2 = 34U;
		ulong val3 = 56L;
		ulong val4 = 78UL;

float	Single-precision floating point type	float val = 1.23F;
double	Double-precision floating point type	double val1 = 1.23; double val2 = 4.56D;
bool	Boolean type; a bool value is either true or false	bool val1 = true; bool val2 = false;
char	Character type; a char value is a Unicode character	char val = 'h';
decimal	Precise decimal type with 28 significant digits	decimal val = 1.23M;

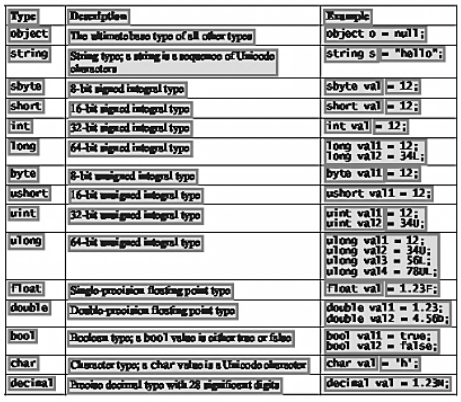


Рис. 7 Блок, содержащий таблицу: вверху – до удаления прямых линий, внизу – после удаления линий.

Эксперименты

Таким образом, все параметры предложенного метода обработки исходного изображения документа могут быть представлены в виде кортежа $M =< \delta_1, \delta_2, \delta_3, \delta_4 >$, где δ_1 – радиус касания областей для первичного объединения; δ_2 – максимально возможное отклонение текстовых областей вдоль вертикальной оси для объединения областей в строке текста; δ_3 – максимально возможный интервал между текстовыми областями вдоль горизонтальной оси; δ_4 – максимальное значение отклонения областей для конечного слияния текстовых областей по вертикали.

Выбор параметров $\delta_1 - \delta_4$ определяется многими факторами, такими как разрешение и размер изображения, параметры форматирования текста на документе и т.п. Решающее значение имеет размер используемого шрифта, на который могут оказывать влияние также все перечисленные факторы. Таким образом, предварительная оценка размеров текста на изображении позволяет выполнить грубую оценку пороговых ограничений $\delta_1 - \delta_4$.

На рис. 8 представлены результаты работы предложенного метода по детектированию текстовых блоков для двух изображений документов, первый из которых (рис. 8, вверху) имеет разрешение 96dpi и размер 1981x1389 пикселей, второй (рис. 8, внизу) – разрешение 300dpi при размере 1984x2928. Для обработки был использован кортеж с параметрами $M_1 =< 50, 50, 50, 25 >$. Как видно из рис.8, текстовые области, параметры которых не соответствуют параметрам кортежа, образуют отдельные области.

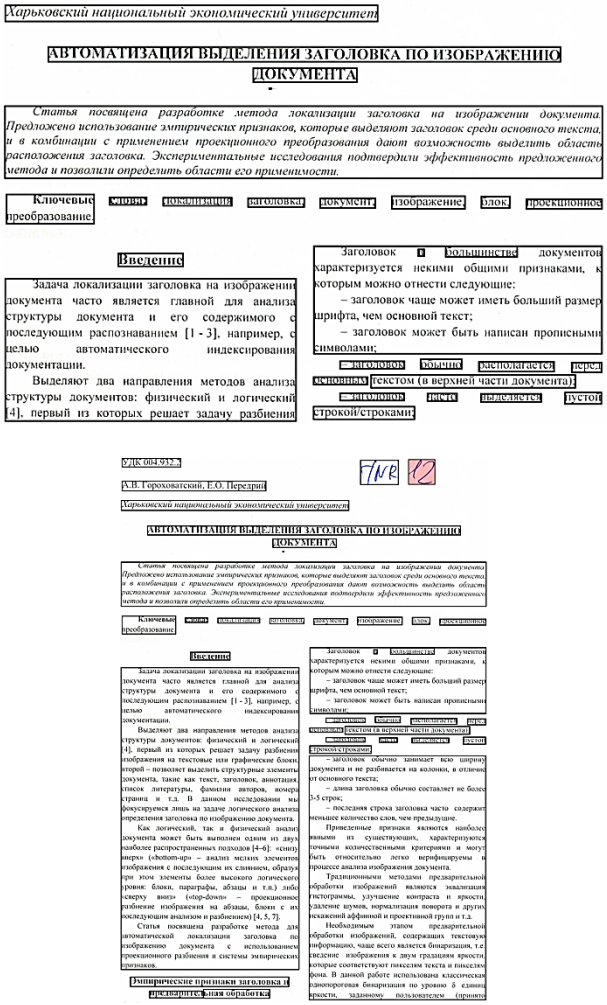
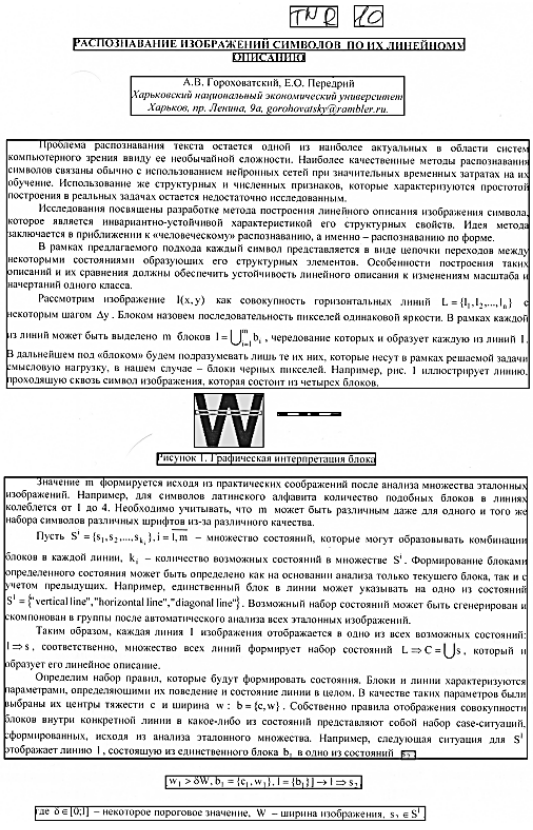


Рис. 8 Примеры обработки изображений текстовых документов.

На рис.9 показаны примеры обработки двух изображений документов, которые содержат графические области. Первое изображение (рис. 9, слева) имеет разрешение 600dpi, размер 4008x6167 пикселей и было обработано с помощью кортежа с параметрами $M_2 =< 70, 150, 57, 165 >$, второе (рис. 9, справа, [18]) – имеет разрешение 96dpi, размер 685x920 пикселей, для его обработки использовался

кортеж $M_3 = < 5,5,5,5 >$. Как видно, области, содержащие графические объекты (выделены на рис. 9 полупрозрачным заполнением), были успешно сегментированы и отделены от текстовых областей. Выбор пороговых значений для кортежа M_3 обоснован в первую очередь мелким размером шрифта на изображении документа.



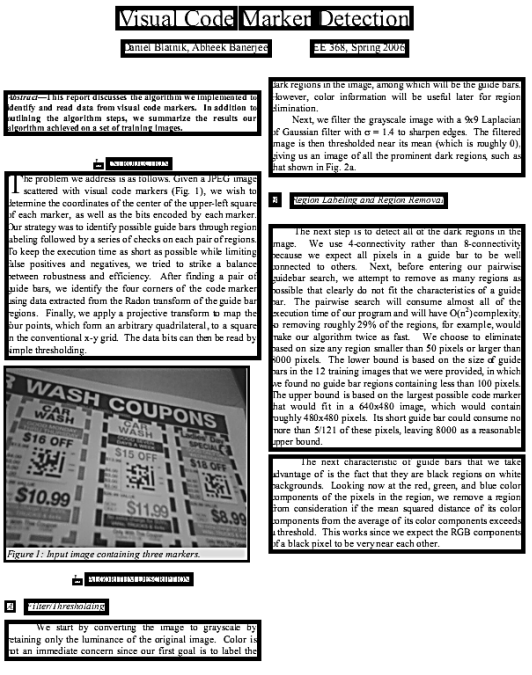
РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ СИМВОЛОВ ПО ИХ ЛИНЕЙНОМУ ОПИСАНИЮ

А.В. Гороховатский, Е.О. Передрий

Рисунок 1. Графическая интерпретация блока

$w_1 > \delta W, b_1 = \{c_1, w_1\}, l = \{b_1\} \to l \Rightarrow s_1$

где $\delta \in [0,1]$ – некоторое пороговое значение, W – ширина изображения, $s_1 \in S^1$



Visual Code Marker Detection

Daniel Bialnik, Abheek Banerjee EE 368, Spring 2006

Figure 1: Input image containing three markers.

Рис. 9 Примеры обработки изображений текстовых документов, которые содержат графические области.

Выводы

В данной статье предложен метод определения местоположения текстовых областей на изображении документа с использованием слияния областей минимального размера с последующим образованием прямоугольных областей вокруг текстовых объектов. Предложенный метод прост в реализации, а также поддерживает возможность сепарации изображений и удаления таблиц.

Наилучший результат применения рассмотренного метода может быть получен с учетом имеющейся априорной информации о структуре документа и размерах используемых элементов, для большинства классов практических задач можно выделить набор пороговых ограничений непосредственно в процессе обработки.

Научной новизной статьи является предложенный метод детектирования текстовых областей на изображении документа с помощью процедуры слияния «снизу вверх», который дает возможность отделить текстовые фрагменты на изображении документов от графических объектов, выделить текст внутри таблиц, сохранив при этом достаточный для решения практических задач уровень быстродействия.

Одним из недостатков предложенного метода является необходимость подбора каждого из пороговых значений $\delta_1 - \delta_4$. Тем не менее, часто начальные их значения можно генерировать в автоматическом режиме на основе анализа физических параметров областей в процессе обработки. Например, значение δ_1 может быть грубо оценено как средняя ширина области на соответствующем этапе. Автоматический выбор эффективных пороговых значений, как и улучшение стабильности метода отделения графической области от текстовой может быть предметом для дальнейшей работы в этой области.

Список литературы

1. Priti P. Rege. Text-Image Separation in Document Images Using Boundary/Perimeter Detection / Priti P. Rege, Chanchal A. Chandrakar // ACEEE International Journal on Signal & Image Processing. – Vol. 03, No. 01. – 2012. – Режим доступа: http://hal.archives-ouvertes.fr/docs /00/74/79/44/PDF/70.pdf.

2. Geometric Layout Analysis Techniques for Document Image Understanding: a Review [Электронный ресурс] / R. Cattoni, T. Coianiz, S. Messelodi, C. M. Modena // ITC-irst Technical Report TR#9703-09.– 1998.– Режим доступа: http://www.ee.bgu.ac.il/~dinstein/stip2002/Seminar_papers/David_Cahana_Geometric%20Layout%20Analysis%20Techniques %20-%20a%20Review.pdf. – Название с экрана.

3. Negi, A. Localization, Extraction and Recognition of Text in Telugu Document Images / Negi A., Shanker K.N.,

Chereddi C.K. // Proceedings of the Seventh International Conference on Document Analysis and Recognition.– 2003 .– P. 1193 – 1197.– Режим доступа: http://www.hserus.net/~cck/pubs/icdar.pdf

4. Smith, R. Hybrid Page Layout Analysis via Tab-Stop Detection [Электронный ресурс] / R. Smith // 10th International Conference on Document Analysis and Recognition (ICDAR), 2009, 26-29 July.– P. 214-245.– Режим доступа: http://dejanseo.com.au/research/google/35094.pdf. – Название с экрана.

5. Nagy, G. Hierarchical representation of optically scanned documents / G. Nagy, S. Seth // International Conference on Pattern Recognition (ICPR), Montreal, 1984.– P. 347-349.

6. Faure, C. Extracting the tables of contents from the images of documents [Электронный ресурс] / C. Faure // Режим доступа: http://perso.telecom-paristech.fr/~cfaure/articles/riao00.pdf. – Название с экрана.

7. Namboodiri, A. Document Structure and Layout Analysis [Электронный ресурс] / Anoop M. Namboodiri, Anil K. Jain // Digital Document Processing.– 2007.– P. 29-48.– Режим доступа: http://pdf.aminer.org/000/348/003/document_page_segmentation_and_layout_analysis_using_soft_ordering.pdf

8. Learning Logic Programs for Layout Analysis Correction [Электронный ресурс] / M. Berardi, M. Ceci, F. Esposito, D. Malerba // Proceedings of the 4th International Workshop on Multilingual OCR (MOCR '13).– 2013.– Режим доступа: http://www.di.uniba.it/~malerba/publications/icml03.pdf

9. Chang, F. Chinese Document Layout Analysis Using An Adaptive Regrouping Strategy [Электронный ресурс] / Fu Chang, Shih-Yu Chu, Chi-Yen Chen // Pattern Recognition, 38(2). – 2005.– P. 261-271.– Режим доступа: http://www.iis.sinica.edu.tw/papers/fchang/1566-F.pdf. – Название с экрана.

10. Text Detection and Translation from Natural Scenes [Электронный ресурс] / J. Gao, J. Yang, Y. Zhang, A. Waibel // School of Computer Science Carnegie Mellon University.– 2001. – Режим доступа: http://www.dtic.mil/dtic/tr/fulltext/u2/a455563.pdf. – Название с экрана.

11. Bukhari, S.S. High Performance Layout Analysis of Arabic and Urdu Document Images / S.S. Bukhari, F. Shafait, T.M. Breuel // Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR 2011).– 2011.– P. 1275 - 1279.

12. Historical Document Layout Analysis Competition / A. Antonacopoulos, C. Clausner, C. Papadopoulos, S. Pletschacher // Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR 2011).– 2011.– P. 1516 - 1520.

13. Bukhari, S.S. Improved Document Image Segmentation Algorithm using Multiresolution Morphology / S.S. Bukhari, F. Shafait, T.M. Breuel // Proceedings of the Document Recognition and Retrieval XVIII - DRR 2011, 18th Document Recognition and Retrieval Conference, San Jose, CA, USA, January 24-29.– 2011.– P. 1- 10.

14. Gupta, N. Image Segmentation for Text Extraction / N. Gupta, V.K. Bange // Proceedings of the 2nd International Conference on Electrical, Electronics and Civil Engineering (ICEECE'2012), Singapore, April 28-29.– 2012.

15. Le, D.X. Automated Borders Detection and Adaptive Segmentation for Binary Document Images / D. X. Le, G. R. Thoma, H. Wechsler // Proceedings of the International Conference on Pattern Recognition (ICPR '96).– Vol. 7276. – P. 737 - 741.

16. Hough P.V.C. Methods and Means for Recognizing Complex Patterns. U.S. Patent 069654, 1962.

17. Duda, R.D. Use of the Hough transform to detect lines and curves in pictures [Электронный ресурс] / R.D. Duda, P.E. Hart // Communications of the ACM. – 1972.– Vol. 15, No 1.– P. 11–15.– ISSN:0001-0782. – Режим доступа: www.ai.sri.com/pubs/files/tn036-duda71.pdf

18. Blatnik, D. Visual Code Marker Detection [Электронный ресурс] / D. Blatnik, A. Banerjee // Режим доступа: http://www.stanford.edu/class/ee368/Project_06/Project/ee368_reports/ ee368group12.pdf. – Название с экрана.

Рецензент: д-р техн. наук, проф. Путятин Е.П., Харьковский национальный университет радиоэлектроники

Авторы:

Гороховатский Алексей Владимирович

Харьковский национальный экономический университет им. С. Кузнеца,

Харьков, кандидат технических наук, доцент кафедры информатики и компьютерной техники.

Тел. 702-06-74 (4-38), E-mail: gorohovatsky@rambler.ru

Детектування текстових областей на зображенні документа методом злиття

О.В. Гороховатський

В статті запропоновано метод автоматичного детектування текстових областей на зображенні за допомогою злиття областей. Запропонований підхід комбінує простоту реалізації алгоритму злиття із аналізом та відокремленням графічних областей, котрі не несуть текстової інформації. Проведено моделювання роботи запропонованого методу, визначено класи практичних задач та умови найбільш ефективного його застосування.

Ключові слова: *текстова область, графічна область, злиття, зображення документа, порогове обмеження, перетворення Хафа.*

The detection of text regions on image of a document using merge method

O.V. Gorokhovatskyi

Paper is devoted to the construction of the method of automatical detection of text regions on image of a document using regions merging. Proposed approach combines the simplicity of merging algorithm implementation with an opportunity of removing image elements which do not contain text information. Experimental investigations were performed, classes of practical tasks to use proposed approach in the most effective way were distinguished.

Keywords: *text region, graphic region, merging, image of a document, threshold limitation, Hough transform.*