

Использование полутонового изображения при поиске границ символов

А. В. БРУХТИЙ, П. А. КУРАТОВ

Аннотация. В работе рассмотрен метод сегментации символов печатного текста на основе информации о яркости пикселей изображения.

Ключевые слова: OCR, сегментация, полутоновое изображение.

Введение

Большинство алгоритмов распознавания символов требуют подачи на вход изображения только одного символа одновременно. Это требует предварительной сегментации строки на отдельные символы, которая производится исходя из геометрических свойств изображения. В простейшем случае выделение символа на изображении несложно. В случае бинарного изображения — это компонента связности. На полутоновом (сером) изображении символ можно выделить, предварительно определив цвет фона. На рис. 1 показан более сложный случай. Здесь два символа склеились, и для их распознавания требуется предварительная сегментация.



Рис. 1. Склеенные символы требуют сегментации перед распознаванием

1. Алгоритм сегментации

Сегментация символов производится в два этапа. Сначала из геометрических соображений определяются возможные границы символов. Затем из полученного множества выбираются различные подмножества, каждое из которых определяет один вариант сегментации. Полученные в результате сегментации символы распознаются. На основе их оценок формируется суммарная оценка варианта сегментации. В качестве решения признается вариант, обеспечивший наилучшую суммарную оценку [1].

В этой схеме множество возможных границ символов имеет ключевое значение. Для правильной

сегментации оно должно содержать все реальные границы символов. Если хотя бы одна граница пропущена, соответствующая пара символов будет сегментирована неправильно. Таким образом, при построении множества возможных границ критичным является пропуск границ. Наличие лишних кандидатов не является критичным, хотя большой объем множества возможных границ приводит к увеличению перебора и повышает возможность выбора ложных границ из-за несовершенства оценок распознавателя символов.

В случае бинарного (черно-белого) изображения возможная граница определяется по характерным точкам внешнего контура [2], по огибающей верхнего и нижнего контуров [3], вертикальной плотности (числу черных пикселей в вертикальном сечении) [4], точках пересечения линий скелетона [5] и другим признакам. Так, границу между склеенными буквами *X* и *A* можно определить по минимальной ширине вертикального сечения (рис. 1), а границу между *O* и *T* подсказывает изгиб верхнего контура, в то время как ширина вертикального сечения здесь имеет минимум в середине буквы *T* (рис. 2).



Рис. 2. Граница между символами определяется по изгибу верхнего контура

Использование изображения, содержащего промежуточные градации яркости (серого изображения), позволяет использовать дополнительную информацию для определения возможных границ. Так, на рис. 3 видно, что на границе между символами *A* и *B* плотность черного цвета имеет локальный минимум.



Рис. 3. Минимум плотности на границе между символами

В то же время на бинаризованном изображении нет никакого указания на границу символов (рис. 4).



Рис. 4. Черно-белое изображение не сохраняет информацию о границе

К сожалению, этот признак также не является универсальным. На рис. 5 плотность черного цвета не имеет минимума на границе символов.



Рис. 5. Склейка символов с равномерной плотностью

Таким образом, локальные минимумы плотности черного цвета не являются универсальным признаком, однако они служат хорошим дополнением к алгоритмам, использующим бинарное изображение.

Перейдем к рассмотрению критерия для поиска границ символов на сером изображении.

Рассмотрим вертикальные сечения изображения и функцию, вычисляющую сумму яркостей точек сечения

$$s(j) = \sum_{i=1}^h p_{i,j}, \quad (1)$$

где i — номер строки, j — номер столбца (сечения), h — высота раstra.

На рис. 6 показан график этой функции серым цветом. Как видно, на границе символов функция имеет минимум. Однако она также имеет минимумы

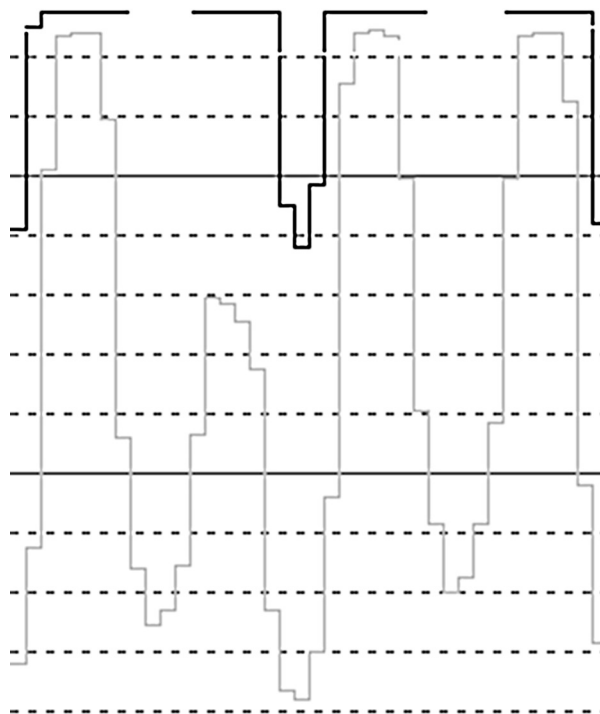


Рис. 6. Функции плотности сечения

в середине символов. Для отсеечения лишних минимумов в [6] используется набор критериев, основанных на сравнении средней и максимальной плотности рассматриваемого сечения и его соседей.

Поскольку набор критериев учитывает только локальные свойства изображения вблизи исследуемой точки, он оказывается недостаточно эффективным для широкого класса изображений. Рассмотрим другой способ. Модифицируем функцию плотности.

$$m(j) = \min_{i=1}^h (p_{i,j}). \quad (2)$$

Теперь мы имеем локальный минимум функции $m(j)$ точно на границе символов (на рис. 6 функция $m(j)$ показана черным цветом).

На практике часто встречаются более сложные случаи (рис. 7).

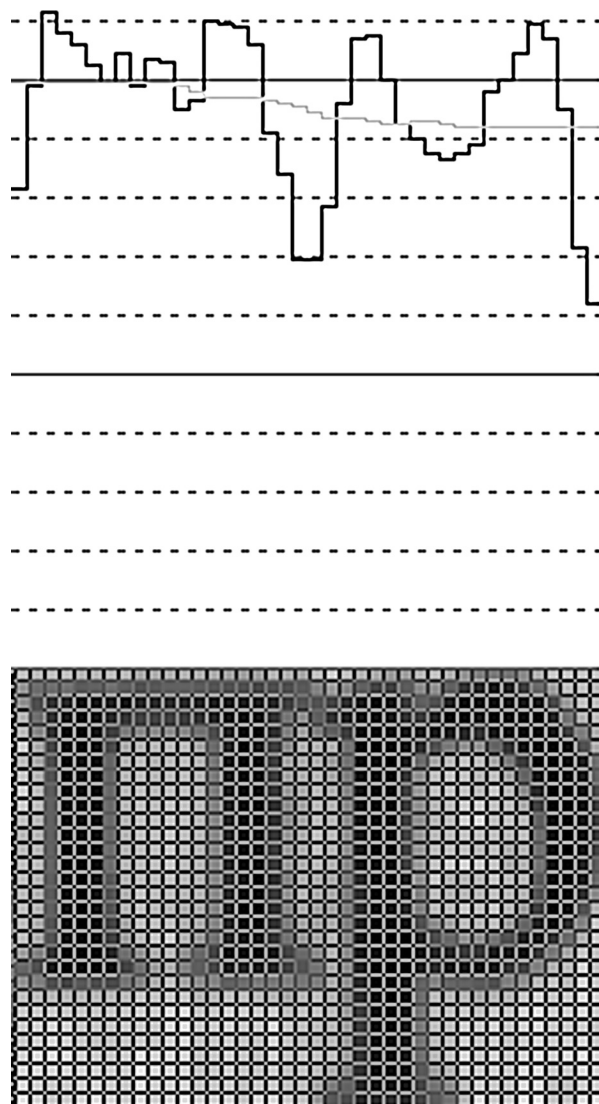


Рис. 7. Функция плотности со множеством экстремумов

Здесь функция $m(j)$ (показана черным цветом) имеет локальные минимумы не только на границах символов, но и внутри их. Многочисленные минимумы внутри символов обусловлены неравномерностью цвета символа. Их количество больше числа границ, уровни менее глубоки, чем у минимумов на границах. Поэтому для выделения минимумов, соответствующих границам символов, можно использовать пороговое значение на величину минимума. Однако использование статического порога невозможно. Так, на рис. 7 минимумы внутри символа глубже, чем минимумы на границе на рис. 6. Для надежного выделения минимумов, соответствующих границам символов, требуется адаптивный порог. Рассмотрим усредненное значение функции $m(j)$ на интервале длины $2l$, где $2l$ — средняя ширина на символа. Для ширины раstra n получим:

$$sm(j) = \frac{1}{2l} \sum_{d=-l}^l m(j+d), \text{ при } l \leq j < n-l;$$

$$sm(j) = \frac{1}{2l} \sum_{d=0}^{2l} m(j+d), \text{ при } 0 \leq j < l;$$

$$sm(j) = \frac{1}{2l} \sum_{d=0}^{2l} m(j-d), \text{ при } n-l \leq j < n.$$
(3)

График функции $sm(j)$ показан на рис. 7 серым цветом. Как видно из рисунка, функции $m(j)$ и $sm(j)$ сближаются на участках медленного изменения функции $m(j)$, а также на участках, где функция $m(j)$ имеет часто расположенные небольшие экстремумы. В то же время на границах букв значения функций $m(j)$ и $sm(j)$ существенно отличаются. Критерий отбора определим следующим образом: точка j_0 является кандидатом на границу символа, если

$$sm(j_0) - m(j_0) > r_o, \quad (4)$$

где r_o — экспериментально определенный порог.

2. Практические результаты

Для оценки качества сегментации использовался стенд из 98 изображений, представляющих собой анкеты, напечатанные на лазерном принтере. Одна анкета содержала в среднем 46 полей. Использование рассмотренного алгоритма увеличило количество правильно распознанных полей, то есть полей, у которых все символы распознаны правильно, на 46 единиц.

Заключение

Предложенный алгоритм поиска границ символов на основе анализа яркостной составляющей полутонового изображения является дополнением к традиционным методам поиска границ и обеспечивает повышение качества распознавания.

Литература

1. Распознавание строк печатных текстов // Сб. трудов ИСА РАН «Методы и средства работы с документами». М.: URSS, 2000. С. 31–51.
2. Dong-Yu Zhang, Xue-Dong Tian, and Xin-Fu Li. An Improved Method for Segmentation of Touching Symbols in Printed Mathematical Expressions // 2nd International Conference on Advanced Computer Control (ICACC). 2010. Vol. 2. P. 251–253.
3. Salman Amin Khan. Character Segmentation Heuristics for Check Amount verification. Дипломная работа на степень магистра Массачусетского Технологического Института, 1998.

4. *George Nagy, Thomas A. Nartker, Stephen V. Rice. Optical Character Recognition: "An illustrated guide to the frontier" // Proc. Document Recognition and Retrieval VII, SPIE. Vol. 3967. P. 58–69.*
5. *Jayarathna U. K. S., Bandara G. E. M. D. C. A Junction Based Segmentation Algorithm for Offline Handwritten Connected Character Segmentation // International Conference on Computational Intelligence for Modelling Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), IEEE, 2006.*
6. *Борисов Е. Сегментация изображений текста. СПбГУ, ПМ-ПУ, 2008.*

Брухтий Антон Владимирович. Н. с. ИСА РАН. Окончил в 1966 г. Московский государственный институт электроники и математики. Количество печатных работ: 3. Область научных интересов: распознавание образов. E-mail: tony@cs.isa.ru

Куратов Павел Александрович. Н. с. ИСА РАН. Окончил в 1978 г. МГУ. Количество печатных работ: 15. Область научных интересов: теория и методы распознавания образов. E-mail: paul@cognitive.ru