

Сегментация штрихов и их соединений при распознавании рукописного текста

Иван Рейер

Вычислительный центр РАН

Москва, Россия

Аннотация

Предлагается метод сегментации рукописного текста, заключающийся в выделении отдельных штрихов и определении типов их соединений. Текст рассматривается как растровое бинарное изображение. Метод основан на совместном использовании сглаженной границы и непрерывного скелетного представления изображения.

Ключевые слова: сегментация, штрих, соединение, растровое бинарное изображение, сглаженная граница, непрерывное скелетное представление

1. ВВЕДЕНИЕ

Существует подход к распознаванию рукописного текста, при котором слово рассматривается целиком как “изображение” (черное пятно достаточно сложной формы на белом фоне) [1,2]. Данный подход предполагает структурный анализ бинарного изображения, который должен выявить топологические и метрические характеристики для распознавания.

Структурный анализ базируется на выделении в бинарном изображении относительно простых элементов. Такими элементами не обязательно являются буквы. Это могут быть, например, элементы письма, аналогичные используемым в школьных прописях. То есть, надпись рассматривается как выполненная в “предалфавите”, состоящем из штрихов (крючков, палочек, овалов, соединительных линий и т. п.). Выделение отдельных штрихов основывается на анализе топологической структуры изображения (обнаружение связанных компонентов, “отверстий” и т. д.). Кроме того, эффективно может быть использован и такой важный признак как характер соединения штрихов. Очевидно, что человек, читая рукописный текст, пользуется информацией о характере соединения при определении букв. Таким образом, для успешного распознавания текста целесообразно анализировать в совокупности и сегментированные штрихи, и типы их соединений. Можно выделить два типа соединения: пересечение (рис. 1) и сопряжение (рис. 2).



Рис. 1.



Рис. 2.

Традиционным инструментом, выявляющим топологию образа, является скелет - множество срединных осей. Однако рассматриваемый сам по себе скелет не позволяет выявить тип соединения. Определить тип соединения можно, анализируя степень гладкости границы изображения в месте соединения. В то же время, только по граничному представлению нельзя выделить штрихи. Поэтому для успешной сегментации штрихов и соединений желательно было бы совместное использование граничного и скелетного представлений.

Традиционно в системах распознавания используется растровое представление границы и скелета изображения. Но в этом случае оценить гладкость границы и установить соответствие между элементами границы и элементами скелета довольно сложно.

В данной статье предлагается метод, который позволяет обеспечить более полный учет информации

о видах соединения штрихов. Метод базируется на использовании двух следующих идей:

- 1) Вместо растрового скелета и границы мы будем использовать непрерывный скелет растрового бинарного изображения и непрерывную сглаженную границу.
- 2) Определение степени гладкости соединений штрихов будем проводить на основе выявления гладких участков границы исходного образа и идентификации участков скелета и участков границы.

Предлагаемый метод включает следующие этапы:

- Аппроксимация растрового бинарного изображения многоугольной фигурой.
- Построение скелетного представления многоугольной фигуры.
- Построение структурного графа по скелетному представлению.
- Выделение в структурном графе компонентов, соответствующих штрихам, и их соединений.

2. АППРОКСИМИРУЮЩАЯ МНОГОУГОЛЬНАЯ ФИГУРА (МФ)

Аппроксимирующая многоугольная фигура для бинарного растрового изображения строится как область, ограниченная конечным числом непересекающихся простых многоугольников. Каждый из них должен быть многоугольником минимального периметра, разделяющим черные и белые точки растрового изображения (рис. 3).



Рис. 3.

Такая фигура называется *многоугольной фигурой минимального периметра* [3].

3. СКЕЛЕТНОЕ ПРЕДСТАВЛЕНИЕ МНОГОУГОЛЬНОЙ ФИГУРЫ

Скелетное представление МФ строится с использованием множества вписанных пустых кругов. В соответствии с [3] вписанным пустым кругом (или просто пустым кругом) называется круг на плоскости, у которого все внутренние точки являются внутренними для МФ, а его граница содержит не менее двух граничных точек МФ.

Множество центров этих кругов образует *скелет* МФ (рис. 4).

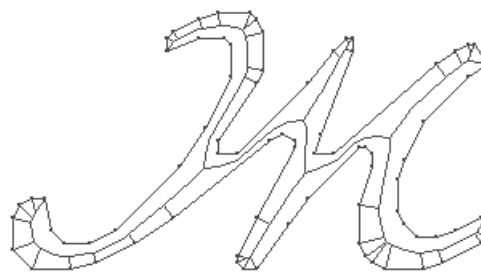


Рис. 4.

Под *скелетным представлением* растрового изображения будем понимать скелет аппроксимирующей МФ минимального периметра вместе с множеством всех вписанных пустых кругов.

Структуру скелета удобно описывать в терминах теории графов. Для этого надо совместно рассмотреть границу МФ и множество пустых кругов. Граница МФ состоит из множества отрезков, концами которых являются вершины МФ. Представим границу как объединение точек – вершин многоугольников и открытых сегментов – сторон многоугольников без концевых точек. Такие точки и открытые сегменты будем называть *сайтами*. Каждый пустой круг МФ касается двух или более сайтов. *Бисектором* пары сайтов называется максимальное связанное множество центров пустых вписанных кругов, касающихся этих сайтов. Очевидно, что бисектор состоит из внутренних точек МФ, равноудаленных от пары сайтов. При этом, если множество таких точек несвязно, то согласно данному определению бисекторами считаются все его максимальные связанные подмножества. Из геометрических соображений ясно, что бисектор является либо сегментом прямой линии (для пар сайт-сегмент и точка-точка), либо участком параболы (для пары точка-сегмент). Концевыми точками бисектора будут центры пустых кругов,

касающихся трех или более сайтов. Таким образом, скелет МФ можно рассматривать как плоский граф, у которого вершины – это центры вписанных пустых кругов, касающихся трех или более сайтов, а ребра – бисекторы. Особо следует отметить случай вырожденных пустых кругов нулевого радиуса с центрами в вершинах МФ, каждый из которых “касается” соответствующей вершины и двух инцидентных ей сайтов-сегментов. В графе-скелете вершины, определяемые этими вырожденными кругами, имеют по одному инцидентному ребру-бисектору.

4. СТРУКТУРНЫЙ ГРАФ

Скелетное представление хорошо описывает бинарное растровое изображение, однако не все элементы скелетного представления несут одинаковую информацию для распознавания рукописного текста. Часть из них действительно существенна и формирует основную, информативную часть скелета. Другие элементы представляют ответвления от основного скелета в сторону границы. Как правило их длина меньше, чем толщина следа пера, они лежат в пределах ширины контура. Поэтому их удаление не нанесет ущерба точности представления фигуры. За счет их удаления можно на основании исходного скелетного представления построить более простое без потери информации. Такое построение можно рассматривать как задачу аппроксимации представления G представлением G' таким, что мера различия G и G' не превосходит заданной величины. Скелетное представление является, по сути дела, объединением пустых кругов, поэтому в качестве меры различия удобно взять хаусдорфово отклонение множеств $h(G, G') = \max\{\max_{g \in G'} d(g, G), \max_{g \in G} d(g, G')\}$, где $d(g, G')$ и

$d(g, G)$ обозначает расстояние от точки g до множеств G' и G соответственно в евклидовой метрике. Задача состоит в том, чтобы найти представление G' , такое что $h(G, G') < \varepsilon$, где ε – заданная точность аппроксимации.

Рассмотрим ребро скелета P_1P_2 (рис. 5), у которого

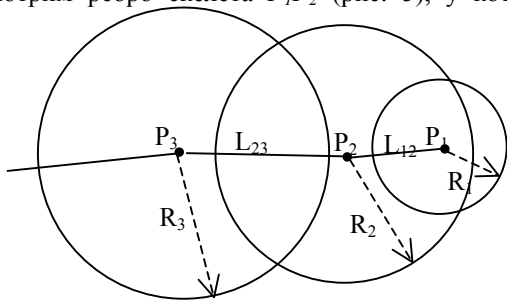


Рис. 5.

вершина P_1 имеет степень 1, а вершина P_2 – степень больше 1.

Пусть R_1 и R_2 – радиусы пустых кругов бисектора P_1P_2 , а L_{12} – расстояние между P_1 и P_2 . Из геометрических соображений очевидно, что хаусдорфово отклонение кругов P_1 и P_2 есть $\delta_{12} = L_{12} + |R_2 - R_1|$. Тогда если $\delta_{12} < \varepsilon$, то все внутренние точки пустых кругов, соответствующих этому ребру, удалены от круга P_2 не более, чем на ε . Следовательно, если задаться точностью аппроксимации ε , ребро P_1P_2 , вполне может быть исключено из скелета. В результате степень вершины P_2 понижается на 1 и может стать равной 1. Тогда с ребром P_2P_3 может быть проделана аналогичная проверка: если $\delta_{12} + \delta_{23} < \varepsilon$, где δ_{23} есть хаусдорфово отклонение кругов P_2 и P_3 , то ребро P_2P_3 также может быть исключено из скелета.

С каждым из терминальных ребер скелета может быть проделана аналогичная процедура. Оставшееся в результате представление может рассматриваться как аппроксимация исходного скелетного представления с заданной точностью ε (рис. 6).



Рис. 6.

5. СЕГМЕНТАЦИЯ ШТРИХОВ И ИХ СОЕДИНЕНИЙ

В структуре скелета можно выделить три типа вершин: вершины степени 3 или выше, терминальные вершины степени 1 и промежуточные вершины степени 2.

Будем считать каждую вершину степени 3 и выше соединением штрихов. Необходимо определить тип каждого такого соединения. Для этого воспользуемся следующими соображениями. Пусть V – произвольный сайт-точка границы, V' и V'' – его соседние сайты-точки (рис. 7).

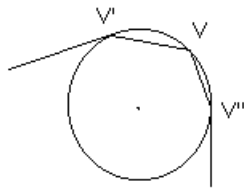


Рис. 7

Будем считать, что граница в точке V является гладкой, если отношение радиуса окружности, проведенной через точки V' , V , V'' к сумме длин отрезков $V'V$ и VV'' превышает некоторую пороговую величину (определяемую экспериментально). В противном случае считаем, что граница имеет в точке V излом. Для каждой вершины степени 3 и выше исследуем окрестность (соседние элементы границы) каждого из касательных сайтов пустого круга (рис. 8).

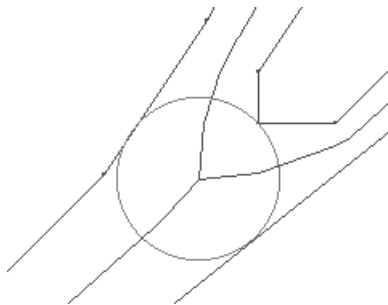


Рис. 8.

Тип соединения штрихов в этой вершине определяется количеством касательных сайтов, в окрестности которых граница имеет изломы, и их взаимным расположением.

В зависимости от типа соединения рассматриваемая вершина заменяется соответствующим числом терминальных и промежуточных вершин.



Рис. 9

В результате получаем граф, состоящий из отдельных связанных компонентов, соответствующих штрихам (рис. 9), и набор связей между ними (рис. 10).



Рис. 10

6. ЗАКЛЮЧЕНИЕ

Предложенный в данной работе метод позволяет удобным образом сегментировать фрагмент рукописного текста на совокупность штрихов и соединений. Метод наиболее полно использует информацию, содержащуюся в изображении слова, за счет совместного рассмотрения скелетного представления и характеристик гладкости границы. Приведенный алгоритм обеспечивает наиболее сложную часть сегментации. Получаемое с его помощью представление слова обеспечивает возможность дальнейшего анализа изображения с использованием разбиения компонентов на более элементарные, рассмотрения метрических характеристик компонентов, классификации штрихов и т. д.

7. ЛИТЕРАТУРА

- [1] Govindaraju, V., Krishnamurthy, R. K. Holistic handwritten word recognition using temporal features derived from off-line image. Pattern Recognition Letters 17 (1996), 537-540.
- [2] Rao P V S. Script recognition. Sadhana, vol. 19, part 2, April 1994, pp. 257-270.
- [3] Местецкий Л. М. Непрерывный скелет бинарного растрового изображения. Графикон-98. Труды конференции, с.71-78.

Segmentation of strokes and their
junctions at recognition of the hand-
written text

Ivan Reyer
CC RAS
Moscow, Russia

Abstract

We suggest the method of segmentation of a hand-written text consisting in allocation of separate strokes and definition of types of their junctions. The text is considered as the binary bitmap. The method is based on using the smoothed border and continuous skeleton of the image.

***Keywords:** segmentation, stroke, junction, binary bitmap, smoothed border, continuous skeleton*

Автор:

Рейер Иван Александрович –аспирант ВЦ РАН.
Адрес: Москва, Вавилова, 40, ВЦ РАН.
E-mail: john-reyer@usa.net