

Метод восстановления следа пера с обучением при off-line распознавании рукописного текста

Нефедов Алексей
Факультет прикладной математики и кибернетики
Тверской государственной университет, Россия
solderer@mail.ru

Abstract

This article describes methods of pen trace reconstruction process in the handwriting documents, with the training possibility. The pen trace reconstruction method is one more approach for the off-line handwriting recognition problem solving.

Keywords: *Handwriting, segmentation, stroke, skeleton, container, classification, training.*

1. ВВЕДЕНИЕ

Задачи распознавания рукописного текста можно разделить в зависимости от типа входных данных на два класса off-line и on-line. К первому относятся задачи, входом для которых является статическое сканированное изображение рукописного текста, представленное в виде матрицы черных и белых пикселей. Так как в этом случае процесс письма предшествует процессу распознавания, задачи такого рода принято называть off-line распознаванием. Системы, решающие такие задачи, получают на вход сканированное изображение, производя анализ, выдают текст, представленный в виде кодов символов.

Ко второму классу относятся задачи, в которых входной информацией является последовательность пикселей, соответствующая траектории движения пера. Обладая такой информацией, можно осуществить распознавание непосредственно в момент письма, поэтому задачи такого рода называют on-line задачи распознавания. Они основываются, прежде всего, на знании о том, как движется перо при начертании того или иного символа.

Задачи распознавания остаются весьма актуальными. Off-line распознавание рукописного текста является наиболее сложной задачей. Существующие технологии в этой области применимы только для решения задач входные данные которых ограничены жесткими рамками. Например, идентификация подписи на банковском чеке, автоматическое распознавание почтового адреса, перевод анкетной информации в электронный вид. Однако, для таких задач, как индексация рукописных архивов при оцифровке, перевод произвольного рукописного текста в электронный вид и ряда других пока не существует технологий, способных решать их с необходимой степенью точности. Проблема заключается в том, что количество возможных вариантов начертания того или иного символа или их сочетаний чрезвычайно велико и даже человек не всегда способен правильно расшифровать рукописный текст при слитном письме.

Технологии, решающие задачи on-line распознавания, развиты гораздо лучше. Они находят применение во всевозможных электронных записных книжках, карманных

компьютерах и везде, где возможен непосредственный ввод информации при письме. Разработанные системы весьма эффективно решают задачи on-line распознавания [1,2].

Системы on-line распознавания незначительно отличаются способом построения. На вход подается последовательность пикселей, соответствующая траектории движения пера. В основном системы on-line распознавания отличаются от систем off-line распознавания блоком сегментации слов на штрихи. Информация, которую несут в себе on-line данные, позволяет намного эффективнее произвести сегментацию слова, что в свою очередь повышает эффективность распознавания.

В связи с этим, возникает идея преобразования off-line сигнала к on-line сигналу, т.е. восстановления динамической информации из статической. После этого становится возможным применение on-line методов распознавания, которые, как уже было сказано, значительно эффективней. Следовательно, можно провести редукцию задачи off-line распознавания к задаче on-line распознавания. Отсюда возникает задача восстановления следа пера (ВСП), которая составляет основную трудность такого перехода [3,4].

2. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим задачу ВСП, в следующей постановке. В качестве исходных данных принимается чёрно-белый растровый образ, в котором фон представлен белыми, а текст – чёрными пикселями. Вопрос получения такого бинарного представления из серого сканированного изображения лежит вне рамок нашего рассмотрения. Бинаризации сканированных текстов посвящено много работ и имеются вполне приемлемые решения этой задачи.

Требуется для исходного бинарного изображения построить описание «следа пера» в виде упорядоченной последовательности точек, которая соответствует реальной траектории движения пера.

3. ПРЕДЛАГАЕМЫЙ ПОДХОД

Основная идея предлагаемого подхода состоит в создании обучаемого алгоритма, осуществляющего решение задачи по образцам, созданным человеком-учителем. Человек в большинстве случаев может правильно восстановить последовательность и направление движения пера. По-видимому, каждый это и делает неявно при чтении рукописного текста. Поэтому можно рассчитывать, что создав представительную базу образцов, можно перейти к автоматическому решению задачи с помощью компьютера. Однако для такого «наивного» решения имеются следующие препятствия.

Во-первых, количество таких «образцов» чрезвычайно велико. Фактически, каждое слитное слово является отдельным образцом. К тому же разные люди имеют разные почерки, что ещё увеличивает объём базы. Ясно, что создать такую базу практически невозможно.

Во-вторых, сам механизм сравнения нового растрового изображения с базой образцов неочевиден. Полученный растровый образ не может быть непосредственно идентифицирован с растровыми образцами в базе.

Предлагаемое решение, позволяющее преодолеть указанные трудности, включает следующие элементы.

- Метод сегментации исходного растрового образа в векторном формате в виде так называемого штрихового графа, в котором рёбра соответствуют линиям, а вершины – точкам соединения или пересечения этих линий.
- Метод декомпозиции (разбиения) слитного рукописного текста на множество отдельных фрагментов, так называемых контейнеров, для каждого из которых восстановление следа пера может быть осуществлено независимо от остальных фрагментов. База образцов, разобранных учителем, включает в себя множество контейнеров.
- Интерактивный метод восстановления следа пера в отдельном контейнере человеком при обучении.
- Метод сравнения контейнеров, полученных при сегментации новых слов, с базой образцов.

4. СЕГМЕНТАЦИЯ

Предлагаемое решение задачи ВСП, построено на математической модели, которая основана на гранично-скелетном представлении растрового бинарного изображения рукописного слова.

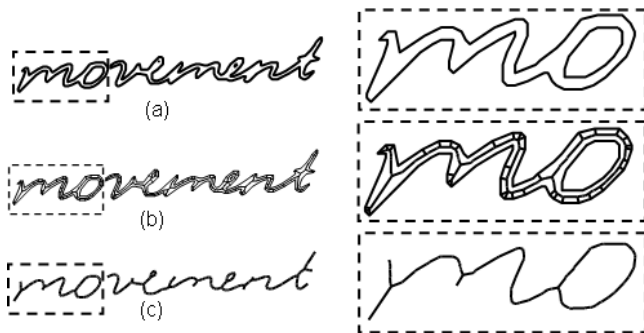


Рисунок 1: Математическая модель рукописного слова.

Модель включает многоугольную фигуру (рис.1а), аппроксимирующую растровое изображение с заданной точностью (обычно это размер одного пиксела), а также непрерывный скелет многоугольной фигуры (рис.1б), образованный центрами всех вписанных в нее окружностей. Из этого скелета получается так называемый базовый скелет (рис.1с)) путем отсечения тех ветвей полного скелета, которые оказываются несущественными при заданной точности аппроксимации. Базовый скелет имеет вид плоского графа с вершинами степени 1,2,3 и ребрами – серединными

осями многоугольной фигуры. Базовый скелет бинарного изображения будем называть *скелетным графом* [5,6].



Рисунок 2: Скелетный граф.

Ветвью скелетного графа называется его связный подграф, в котором нет циклов и всего две вершины имеют степень, отличную от второй.

Каждая вершина скелетного графа образована центром окружности.

Шириной ветви будем называть усредненный радиус окружностей, центры которых образуют вершины этой ветви.

Выровненной ветвью называется ветвь, полученная в результате операции выравнивания.

Короткой ветвью скелетного графа называется ветвь, соединяющая две вершины третьей степени, длина которой меньше некоторой величины, являющейся параметром регулирования системы (Рисунок 3).

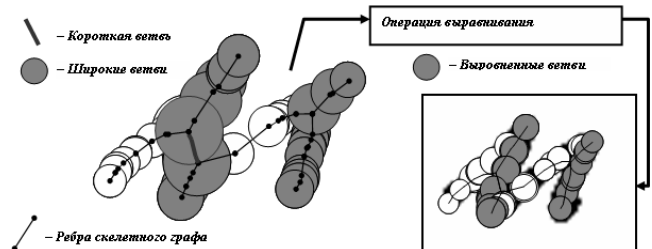


Рисунок 3: Операция выравнивания ветви скелетного графа.

Терминальными вершинами скелетного графа будем называть вершины первой степени (Рис. 2).

Каждая ветвь, соединяющая вершины третьей степени и не являющейся *короткой ветвью*, «разрезается» пополам. В результате образуются две терминальные вершины, не связанные между собой, но занимающие одно и то же положение на плоскости (Рис. 2).

Штрихом – будем называть цепь, состоящую из одной или нескольких ветвей, имеющих общие конечные точки, причем каждая ветвь имеет общую конечную точку не более чем с одной ветвью. Фактически штрих моделирует часть рукописного слова, которая была написана без отрыва пера и без резких изменений направления движения пера.

Выровненным штрихом – будем называть цепь, состоящую из соответствующих штриху выровненных ветвей.

На основе математической модели, может быть построено разбиение рукописного слова на *упорядоченные контейнеры* – независимые части, анализ которых возможен независимо друг от друга. На рисунке 4 показано разбиение рукописного слова на контейнеры.

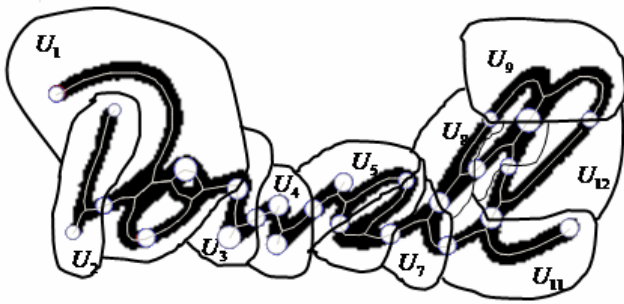


Рисунок 4: Разбиение слова на контейнеры.

Фактически контейнер представляет собой независимый подграф скелетного графа и построенное на нем множество штрихов. В свою очередь множество штрихов контейнера состоит из реальных штрихов и ложных (гипотетически возможных штрихов). Для примера рассмотрим рисунок 5, на котором изображен простой контейнер, множество штрихов которого состоит из трех элементов. Человек без труда определит траекторию движения пера при написании такого контейнера. В начале был написан штрих, проходящий через вершины 123, затем штрих 324, в свою очередь траектория представленная штрихом 124 является ложной, так как в реальности перо не шло по такой траектории.

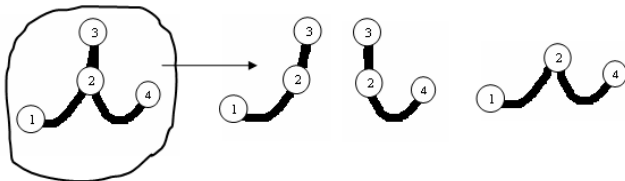


Рисунок 5: Реальные и ложные штрихи контейнера.

Для восстановления траектории движения пера необходимо для каждого контейнера осуществить следующие действия:

- из множества штрихов контейнера выбрать реальные штрихи и отсеять ложные
- установить последовательность написания реальных штрихов и направление движения по ним.

Предлагаемый механизм обучения описывается следующей схемой.



5. КОНТЕЙНЕРЫ ШТРИХОВ

Пусть G – скелетный граф. s_1, \dots, s_n – ветви скелетного графа. v_1, \dots, v_r – вершины скелетного графа.

Пусть g_1, \dots, g_m – множество штрихов (подграфов). Например, штрих может быть получен как $s_1 \cup s_2$ или $s_4 \cup s_6$ (Рисунок 2.).

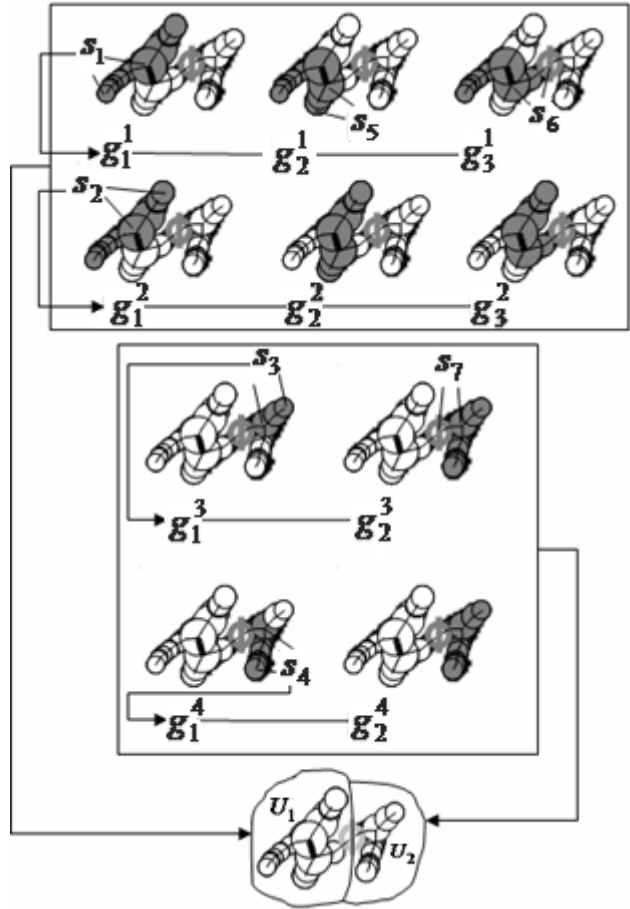


Рисунок 6: Пример построения контейнера.

Рассмотрим множество G'_i всех возможных штрихов g_1, \dots, g_l , $l < m$, включающих ветвь s_i , т.е. для которых выполняется условие $s_i \in g_j$, $j = 1, \dots, l$. G'_i – образует **пучок штрихов** для s_i .

На множестве пучков введем отношение смежности.

Два пучка штрихов называются смежными, если $\forall g' \in G', \exists g'' \in G'' : g' \cap g'' = s$.

Отношение смежности задается на множестве пучков графом.

Максимальное по включению множество попарно смежных пучков, называется **контейнером**. Множество контейнеров $U = \{U_1, \dots, U_h\}$, однозначно задает разбиение скелетного графа G на непересекающиеся подграфы. Каждому

контейнеру можно поставить в соответствие порядковый номер $1 \leq r \leq h$, причем $\bigcap_1^h U_r = 0$.

Рисунок 6 поясняет вышесказанное. Для ветви скелета s_1 , построен пучок штрихов $G_1 = \{g_1^1, g_2^1, g_3^1\}$, а для s_2 пучок $G_2 = \{g_1^2, g_2^2, g_3^2\}$. Пучки G_1 и G_2 являются смежными, так как $g_1^1 = g_1^2$, $g_2^1 \cap g_2^2 = s_5$ и $g_3^1 \cap g_3^2 = s_6$. Также, для s_3 построен пучок $G_3 = \{g_1^3, g_2^3\}$ и для s_4 $G_4 = \{g_1^4, g_2^4\}$. Пучки G_3 и G_4 , являются смежными, так как $g_2^3 = g_2^4$, $g_1^3 \cap g_1^4 = s_7$. Пучки для s_5, s_6, s_7, s_{short} не рассматриваются, так как они не образуют новых штрихов, отличных от множества штрихов образованного пучками G_1, G_2, G_3, G_4 . Отметим, что никакие другие пучки, кроме указанных, не являются попарно смежными. Таким образом, пучки G_1 и G_2 образуют **контейнер** $U_1 = \{g_1^1, g_2^1, g_3^1, g_2^2, g_3^2\}$, а пучки G_3 и G_4 **контейнер** $U_2 = \{g_1^3, g_2^3, g_1^4, g_2^4\}$.

6. ЭКСПЕРТНОЕ ВОССТАНОВЛЕНИЕ СЛЕДА ПЕРА НА КОНТЕЙНЕРЕ

Для экспертного восстановления следа пера было создано рабочее место оператора- эксперта, инструменты которого позволяют отобрать реальные штрихи, установить их порядок и направление движения по ним. Приведем пример экспертного восстановления следа пера на контейнере. Рассмотрим пример показанный на рисунке 7.

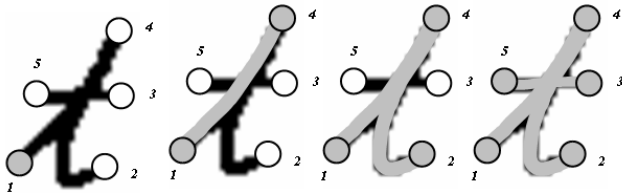


Рисунок 7: Восстановление следа пера контейнера оператором-экспертом.

Фактически оператор-эксперт должен указать, как двигалось перо при написании контейнера. Для этого на экране отображается контейнер и его терминальные узлы, посредством выбора которых эксперт восстанавливает траекторию движения. Для примера показанного на рисунке 6, с точки зрения экспертных знаний автора, сначала должна быть выбрана терминальная вершина номер 1, затем вершина 4, после чего штрих, соединяющий вершины 1 и 4, будет выбран как первый штрих траектории с заданным направлением движения от вершины 1 к вершине 4. Далее выбирается вершина 2 и штрих, соединяющий вершины 4 и 2 заносится в траекторию, как второй штрих с направлением движения от 4 к 2. После чего, перо перемещается в вершину 5, моделируя так называемый отрыв пера. Следующей выбирается вершина с номером 3 и штрих, соединяющий

вершины 5 и 3, заносится в траекторию, как третий штрих с направлением движения от 5 к 3.

Таким образом, на контейнере восстанавливается траектория движения пера, состоящая из следующих штрихов (1,4) (4,2) (5,3). Восстановленная траектория сохраняется в базе образов как шаблон для последующей классификации контейнеров.

7. АВТОМАТИЧЕСКАЯ КЛАССИФИКАЦИЯ КОНТЕЙНЕРОВ

Каждый контейнер можно отобразить в связанный граф без циклов с вершинами первой и третьей степени. Будем говорить, что два контейнера являются идентичными, если удалось установить взаимно однозначное метрико-топологическое соответствие между двумя графами.

Введём следующее правило нумерации вершин графа таких графов. В качестве примера рассмотрим контейнер U_1 (Рисунок 8(а)) из слова, изображенного на рисунке 4. Отобразим контейнер в граф с вершинами первой и третьей степени.

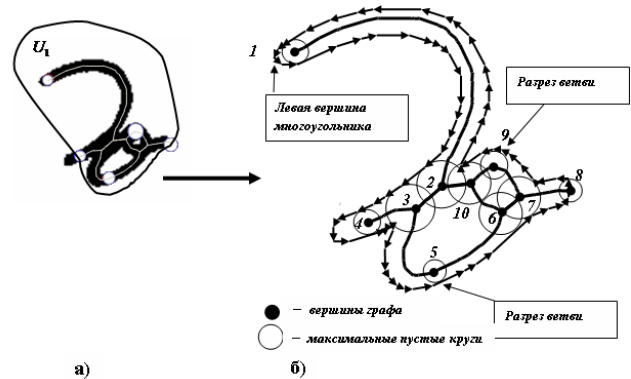


Рисунок 8: Упорядочивание вершин графа контейнера.

Гранично-скелетное представление включает в себя многоугольную фигуру, аппроксимирующую растровое бинарное изображение. С каждой вершиной графа связан максимальный пустой круг, для которого в многоугольной фигуре есть касательные ребра. Зададим отношение порядка для вершин графа следующим образом. Найдем в многоугольной фигуре самую левую вершину, затем осуществим обход многоугольника против часовой стрелки. В процессе обхода будем проверять каждое ребро многоугольника на касание с максимальным пустым кругом вершин графа. Если касание обнаружено, то вершине, чей максимальный пустой круг касается ребра, присваиваем очередной порядковый номер. В конце процесса каждой вершине графа будет присвоен порядковый номер (рис. 8(б)). Среди упорядоченных вершин найдем самую левую, правую, верхнюю и нижнюю, запомним их номера. В примере самая левая вершина 1, правая 8, нижняя 5, верхняя 1.

Найдем все пути на графе, соединяющие вершину 1 с другими вершинами графа первой степени. Вершины 5 и 9 являются по построению вершинами разреза ветви, поэтому при поиске пути на графе при попадании в одну из этих вершин очередной путь будет найден и добавлен в список путей. Таким образом, на графе изображенном на рисунке

8(б), будут найдены следующие пути (1,2,3,4) (1,2,3,5) (1,2,10,6,5) (1,2,10,9) (1,2,10,6,7,9) (1,2,10,6,7,8).

В базе данных хранятся контейнеры-шаблоны, для которых заранее проделана описанная процедура. Более того, для каждого шаблона восстановлен след пера, т.е. оператором-экспертом из множества штрихов контейнера отобраны реальные штрихи, установлен их порядок и направление движения по ним.

Таким образом, для автоматической классификации контейнера необходимо, сравнить список путей контейнера со списком путей каждого шаблона из базы данных. Если два списка однозначно отображаются друг в друга и самые левые, правые, нижние, верхние вершины, совпадают по номерам, то это означает, что найдено взаимнооднозначное метрико-топологическое соответствие между контейнером и шаблоном. Таким образом, становится возможным восстановить штрихи контейнера в соответствии с шаблоном.

8. ЗАКЛЮЧЕНИЕ

В текущий момент реализованы алгоритмы сегментации, алгоритмы декомпозиции скелетного графа на контейнеры. Реализовано рабочее место оператора эксперта для создания прецедентной базы шаблонов.

Предварительные результаты исследований показывают перспективность предлагаемых методов и алгоритмов. Основным преимуществом является возможность обучения. В дальнейшем планируется подключение системы *on-line* распознавания рукописного текста реализованной в библиотеке Microsoft Ink Analysis.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант 05-01-00542)

9. ЛИТЕРАТУРА

[1] Claus Bahlmann, Bernard Haasdonk, Hans Burkhardt. "On-line Handwriting Recognition with Support Vector Machines - A Kernel Approach", publ. in Proc. of the 8th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR), pp. 49–54, 2002.

[2] A.El-Nasan, M. Perrone. "On-line Handwriting Recognition Using Character Bigram Match Vectors", publ. in Proc. of the 8th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR), pp. 67–72, 2002.

[3] P.M. Lallican, C. Viard-Gaudin, S. Knerr, "From Off-line to On-line Handwriting Recognition", Seventh International Workshop on Frontiers in Handwriting Recognition September 11-13 2000, Amsterdam.

[4] V. Govindaraju, D. Wang, S. Srihari, "Using Temporal Information In Off-Line Word Recognition", Advanced Technology Conference, Vol. 1, pp. 529, 1992.

[5] Местецкий Л.М., Рейер И.А.. "Непрерывная гранично-скелетная модель дискретного изображения с контролируемой точностью аппроксимации", доклады XI Всероссийской конференции "Математические методы распознавания образов", Москва, 2003, С. 367-370.

[6] Местецкий Л.М. "Непрерывный скелет растрового изображения", графикон – 98, Москва, МГУ, 1998.

Об авторе

Нефедов Алексей аспирант Тверского государственного университета. Факультет Прикладной математики и кибернетики. Адрес электронной почты solderer@mail.ru.