

УДК 621.398.727.12

Распознавание слитных рукописных слов с помощью модифицированных глирафов

А. В. Князев*

Рассмотрен метод распознавания слитного рукописного текста, основанный на понятии глирафа — конструкции, объединяющей свойства графа и линейчатого рисунка. Предложены процедуры построения модифицированного глирафа и распознавания слитных рукописных слов. Определена мера схожести глирафов.

Ключевые слова: распознавание, рукописный текст, глираф.

Введение

Распознавание рукописного текста является важной задачей. Однако заметные успехи достигнуты только при распознавании рукописного текста в режиме «on-line» при написании текста пером на специальном планшете, так как при этом доступна информация о направлении и скорости движения пера. Однако при распознавании текста в режиме «off-line», когда выполняется работа с предварительно отсканированным изображением текста и отсутствует дополнительная информация, таких результатов пока нет.

Решение задачи распознавания текста разбивается на ряд подзадач: отделение текста от графической информации, удаление шумов, выделение строк и отдельных слов, распознавание букв и слов. Наиболее сложной считается задача распознавания слитно написанных рукописных слов.

Существует ряд подходов к распознаванию рукописного текста, в частности, подходы, основанные на использовании нейронных сетей [1 — 3], на использовании скрытых моделей Маркова [4 — 6] и структурных особенностей букв и слов [7, 8].

В [9] описан метод распознавания рукописных слов, основанный на использовании глирафов — графоподобных конструкций, обладающих свойствами как графов, так и линейчатых рисунков. В данной работе предложен метод на основе модифицированных глирафов, т.е. глирафов, построенных на сетке с некоторым фиксированным шагом.

Понятие глирафа

Рассмотрим понятие глирафа, изложенное в [9].

После сканирования изображения выполняется предварительная обработка: проводятся фильтрация и скелетизация изображения с целью удаления шумов и получения линий толщиной в одну точку. Таким образом, слово представляется множеством точек на дискретной плоскости с шагом 1. Пример такого изображения слова приведен на рис. 1.

Точки, образующие линию, являются соседями. Для обозначения соседства используем нумерацию направлений, предложенную Г. Фрименом [10] и представленную на рис. 2.

Таким образом, точка, расположенная справа от рассматриваемой точки, является 0-соседом, слева — 4-соседом и т.д. Точки внутри линии (цепочки) имеют двух соседей.

* KniazevAV@mpei.ru

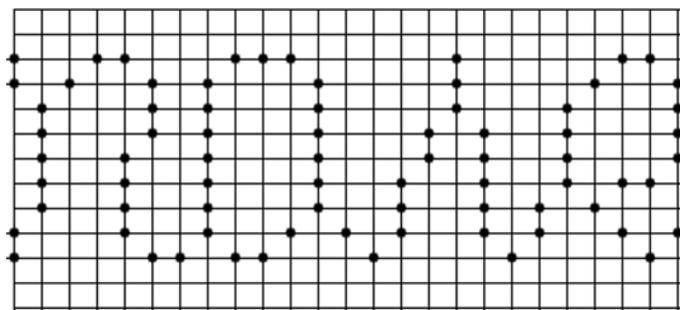


Рис. 1

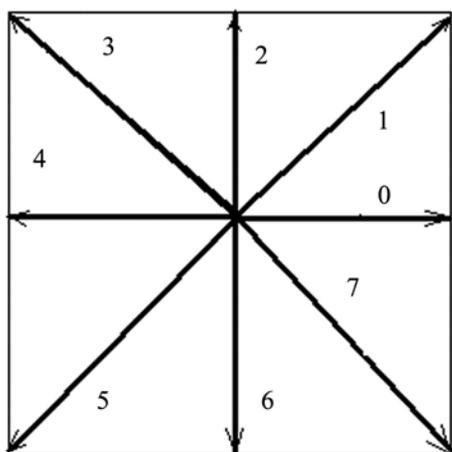


Рис. 2

Концевые точки (точки, имеющие только одного соседа) и точки ветвления (у которых более двух соседей) образуют множество узловых точек.

Последовательность точек (цепочку) между двумя узловыми точками, не содержащую других узловых точек, будем называть плетью. Следовательно, плети либо не пересекаются (не имеют общих точек), либо пересекаются только на концах.

Множество точек, составляющих плети, образует граф. Таким образом, граф (от слов graph — рельефно вырезанная фигура и graph — граф) — это объединение плетей.

Он представляет собой некоторую модель рукописного слова (рис. 3).

На рис. 3 граф слова «лес» содержит 8 плетей: $L1, L2, \dots, L8$, а узловые точки $x1, x2, \dots, x8$ обведены

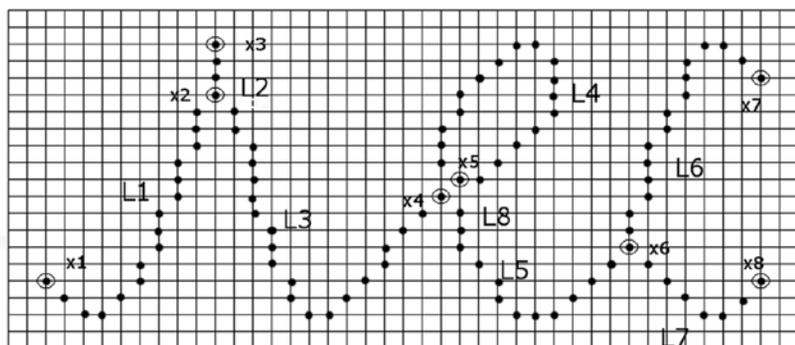


Рис. 3

кружочками. Плеть $L1$ — множество точек от $x1$ до $x2$; $L2$ — множество точек от $x2$ до $x3$ и т.д. Плеть $L8$ состоит из двух точек $x4$ и $x5$.

Таким образом, граф — это рисунок, образованный линиями, являющимися цепочками точек.

С другой стороны графу можно сопоставить взвешенный граф, вершинами которого будут точки графа, дуги связывать соседние точки, а веса дуг — номера направлений соседства. Такой граф назовем первичным графом, соответствующим данному графу или просто графом графа.

Графу можно сопоставить и другие графы, например граф, вершинами которого будут узловые точки, а ребра — соответствовать плетям. Такой граф определяет множество взвешенных графов, различающихся способом задания весов ребер, и называется опорным графом.

Вообще говоря, в качестве вершин графа, сопоставляемого графу, можно выбрать любое множество точек графа, включающее узловые точки.

Следовательно, граф определяет семейство графов, вложенных друг в друга.

Модифицированные графы

Для построения модифицированного графа используется отсканированное нескелетизированное изображение (после удаления шумов). При этом линии, образующие слово, имеют толщину более одной точки. Модифицированный граф строится следующим образом: на дискретную плоскость изображения накладывается сетка с шагом h ($h > 1$), узлы сетки, по которым (или рядом с которыми) проходят линии изображения слова, становятся вершинами модифицированного графа.

На рис. 4 приведена дискретная плоскость с наложенной сеткой (шаг равен 5) и изображением строчной буквы «л» (для удобства восприятия нарисован только внешний контур линий буквы). Жирные точки обозначают вершины модифицированного графа. Для построения модифицированного графа изображение сканируется сначала вдоль горизонтальных линий, затем вдоль вертикальных линий сетки. При этом определяются границы интервалов, задающих линии изображения. Узлы сетки, лежащие внутри интервала, включаются в граф. Если некоторый интервал не содержит узел сетки, то находится ближайший узел слева или справа от него (для горизонтальных интервалов) и включается в граф. Для вертикальных интервалов в граф входит ближайший узел сверху или снизу.

Рассмотрим алгоритм построения модифицированного графа. Обозначим матрицу изображения b ($b[i, j] = 1$, если в позиции $[i, j]$ есть точка изображения и $b[i, j] = 0$ в противном случае); N, M — чис-

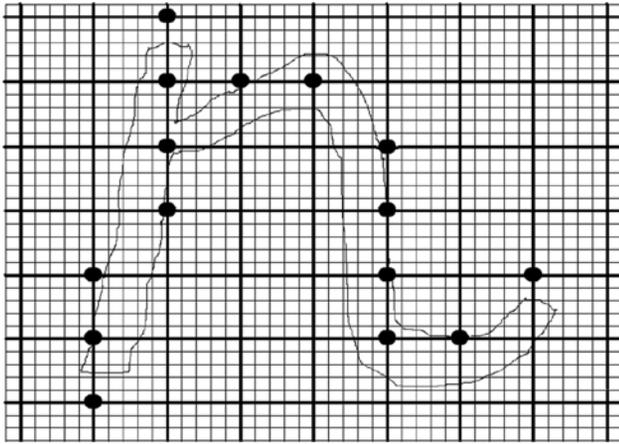


Рис. 4

ло строк и столбцов сетки; n , m — число строк и столбцов исходного изображения.

Алгоритм формирования модифицированного графа имеет следующий вид:

```

/* Цикл по горизонтальным линиям сетки */
ЦИКЛ  $i = 0, N - 1$ 
  Флаг Начала Интервала = 0
  ЦИКЛ  $j = 0, m - 1$ 
    ЕСЛИ  $b[i, j] = 1$  ТО
      Флаг Начала Интервала = 1
      Запомнить левую границу интервала
    КОНЕЦ_ЕСЛИ
    ЕСЛИ  $b[i, j] = 0$  & Флаг Начала Интервала = 1 ТО
      Флаг Начала Интервала = 0
      Запомнить правую границу интервала
      ЕСЛИ внутри интервала есть узлы сетки ТО
        Занести узлы в граф
      ИНАЧЕ
        Найти узел сетки, ближайший к интервалу
        Занести узел в граф
      КОНЕЦ_ЕСЛИ
    КОНЕЦ_ЕСЛИ
  КОНЕЦ_ЦИКЛ
КОНЕЦ_ЦИКЛ

/* Цикл по вертикальным линиям сетки */
ЦИКЛ  $j = 0, M - 1$ 
  Флаг Начала Интервала = 0
  ЦИКЛ  $i = 0, n - 1$ 
    ЕСЛИ  $b[i, j] = 1$  ТО
      Флаг Начала Интервала = 1
      Запомнить верхнюю границу интервала
    КОНЕЦ_ЕСЛИ
    ЕСЛИ  $b[i, j] = 0$  & Флаг Начала Интервала = 1 ТО
      Флаг Начала Интервала = 0
      Запомнить нижнюю границу интервала
      ЕСЛИ внутри интервала есть узлы сетки ТО
        Занести узлы в граф
      ИНАЧЕ
        Найти узел сетки, ближайший к интервалу
        Занести узел в граф
      КОНЕЦ_ЕСЛИ
    КОНЕЦ_ЕСЛИ
  КОНЕЦ_ЦИКЛ
КОНЕЦ_ЦИКЛ

```

На рис. 5 показан модифицированный граф для буквы «m», изображенной на рис. 4.

Обучение и распознавание

Рассмотрим процесс распознавания. Исходное изображение слова представляется в виде файла в определенном графическом формате, например в *.bmp.

Затем изображение предварительно обрабатывается. Предположим, что в нашем случае выполняется фильтрация изображения (устранение шумов), определение базовых линий, и выделяются слова.

Далее строим модифицированный граф (в дальнейшем просто граф). В итоге получаем его как модель рукописного слова. Граф слова по сути является объединением графов букв. При распознавании возникает обратная задача: выделение графов букв из графа слова. Это — непростая задача для слитно написанного слова.

В процессе распознавания предполагаемые графы букв будут выделяться и сравниваться с эталонами (заранее описанными графами). Библиотека эталонных графов составляется на стадии обучения. В общем случае для каждой буквы может быть несколько эталонных графов.

Таким образом, будем считать, что мы имеем граф распознаваемого слова.

Этап распознавания слов можно представить как взаимодействие двух процессов: объединения определенных множеств плетей в группу и идентификации сформированных групп как некоторых символов. Объединение плетей в группу — это выделение некоторого подграфа в исходном графе, а идентификация символа — это определение сходства выделенного подграфа и библиотечных графов, представляющих символы.

Разбиение слова на подграфы — задача неоднозначная, здесь возможны различные варианты, поэтому результатом данного этапа является граф символов слова $G^{(C)}$, т.е. ориентированный граф без циклов, ребра которого соответствуют символам. Каждый путь в графе $G^{(C)}$ от начальной до конечной вершины представляет собой один из возможных вариантов распознавания данного слова.

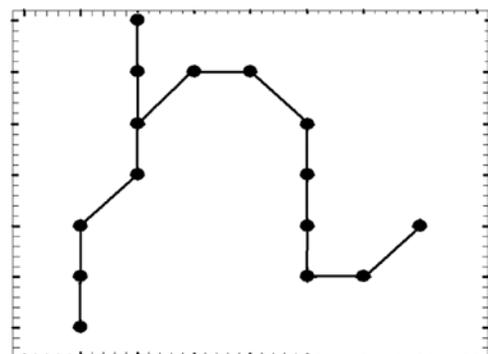


Рис. 5

При распознавании используем список упорядоченных плетей и два указателя на элементы этого списка. Здесь предполагается известной средняя ширина символа ρ .

Обобщенная процедура распознавания слова имеет следующий вид (за основу взята процедура из [9], которая была модифицирована):

1. Выполняется предварительная обработка изображения слова: устраняются шумы и строится модифицированный глираф.

2. Выделяются плети — последовательности точек между концевыми точками и/или точками ветвления.

3. Формируется упорядоченный список плетей исходного слова (плети упорядочиваются слева направо и сверху вниз). Формируется начальная вершина v графа $G^{(C)}$. Эта вершина становится текущей вершиной графа $G^{(C)}$. Указатель активной плети устанавливается на первую плеть в списке.

4. Сбрасывается признак наличия распознанных символов.

5. Указатель текущей плети устанавливается в позицию указателя на активную плеть.

6. Выполняется очистка множества плетей текущего символа.

7. Текущая плеть добавляется в текущее множество.

8. Для текущего множества плетей происходит распознавание символа путем определения схожести выделенного подглирафа с эталонными глирафами букв. Если символ распознан, то в графе $G^{(C)}$ формируется новая текущая вершина w и ребро e , соединяющее эту вершину с вершиной v . Новая вершина w включается в очередь текущих вершин. Ребро e помечается распознанным символом. Указатель текущей плети передвигается на следующую плеть. Устанавливается признак наличия распознанных символов.

9. Если число плетей, включаемых в текущее множество, не превышает заданный предел, а ширина формируемого кандидата в символы не превышает $3/2 \rho$, то переходим к этапу 7.

10. Если очередь активных вершин не пуста, то выбираем из очереди активную вершину v ; указатель текущей плети устанавливаем на плеть, следующую за плетью, включенной в распознанные символы последней. Переходим к этапу 6.

11. Если обработаны все плети, то работа заканчивается.

12. Если признак наличия распознанных символов установлен, то указатель активной плети устанавливаем на плеть, следующую за плетью, включенной в распознанные символы последней; иначе указатель активной плети сдвигается на одну позицию. Переходим к этапу 4.

В итоге формируется граф вариантов распознавания слова. Средняя ширина символа неизвестна для первого распознаваемого слова: в этом случае в качестве

первоначальной ширины используется расстояние между базовыми линиями. После того как первое слово распознано, вычисляется средняя ширина символа, и она далее используется в приведенной процедуре.

На рис. 7 приведен граф символов, построенный для слова «шина» [9], изображенного на рис. 6.

При распознавании возникают различные варианты группирования плетей, поэтому данное слово может быть распознано как «шина» или «ишна». Окончательное решение принимается с использованием словаря.

Перейдем к рассмотрению процесса обучения. На вход подаются образцы букв, для каждого образца формируется модифицированный глираф; если похожего глирафа в библиотеке эталонов нет, то новый глираф добавляется в библиотеку.

Сходство глирафов

Самым трудоемким этапом в приведенной выше процедуре распознавания слов является этап 8 — идентификация символа, т.е. сравнение выделенного глирафа с эталонными глирафами, хранящимися в библиотеке.

Идентификация выполняется в несколько шагов. На первом шаге выбираются эталонные глирафы, опорные графы которых изоморфны с опорным графом идентифицируемого глирафа, и определяется степень сходства опорных графов. На втором шаге сравниваются плети найденных эталонных глирафов с соответствующими плетями идентифицируемого глирафа. При этом вычисляются коэффициенты схожести соответствующих плетей и общий коэффициент схожести глирафов. На третьем шаге выбирается эталонный глираф, наиболее схожий с идентифицируемым глирафом, и идентифицируемому глирафу сопоставляется символ данного эталонного глирафа.

Рассмотрим, как определяются коэффициенты схожести. Коэффициент схожести опорных графов вычисляется по следующей формуле:

$$k_{\text{опор}} = 1 - \frac{\sum_i d_i}{2l},$$

где d_i — расстояние между i -ми вершинами опорных графов; l — расстояние между базовыми линиями.



Рис. 6

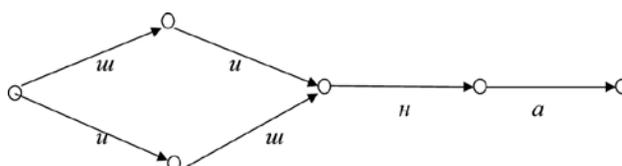


Рис. 7

Коэффициент схожести i -ых плетей определяется следующим образом:

$$k_i^{\text{плет}} = 1 - \frac{\sum_j d_j}{3l},$$

где d_j — расстояние между j -ми вершинами соответствующих плетей графов.

Коэффициент схожести графов определяется следующим образом:

$$k_{\text{схож}} = k^{\text{опор}} \prod_i k_i^{\text{плет}}.$$

Таким образом определяется коэффициент схожести идентифицируемого графа с каждым из отобранных эталонных графов, находится эталонный граф, наиболее схожий с идентифицируемым, и идентифицируемому графу приписывается символ данного эталонного графа.

В настоящей работе рассмотрен подход к распознаванию слитного рукописного текста, основанный на использовании модифицированных графов. Описана процедура построения модифицированного графа, приведена процедура распознавания слитного рукописного текста, изложен способ определения степени схожести графов.

Работа выполнена при поддержке РФФИ (проект № 14-01-00671а).

Литература

1. **Ghosh M., Ghosh R.** A fully automated offline handwriting recognition system incorporating rule based neural network validated segmentation and hybrid neural network classifier // Intern. J. pattern recognition and artificial intelligence. 2004. V. 18. N 7. P. 1267 — 1283.

2. **Астахова И.Ф., Мищенко В.А., Краснояров А.В.** Алгоритм обучения нечеткой нейронной сети Ванга-Менделя для распознавания рукопечатных символов в работе почтовой службы // Вестник ВГУ. Сер. Системный анализ и информационные технологии. 2011. № 2. С. 144 — 148.

3. **Куссуль Э.М., Касаткина Л.М., Байдык Т.Н., Лукович В.В.** Нейросетевые технологии распознавания рукописных текстов // Управляющие системы и машины. 2001. № 2. С. 64 — 83.

4. **Artieres Thierry, Marucatat Sanparith.** Online handwritten shape recognition using segmental hidden Markov models. // IEEE Trans. pattern analysis and machine intelligence. 2007. V. 29. N 2. P. 205 — 217.

5. **Xue H., Govindaraju V.** Hidden Markov models combining discrete symbols and continuous attributes in handwriting recognition // IEEE Trans. pattern analysis and machine intelligence. 2006. V. 28. N 3. P. 458 — 462.

6. **Gunter S., Bunke H.** HMM-based handwritten word recognition: on the optimization of the number of states, training iterations and Gaussian components. // Pattern recognition. 2004. V. 37. N 10. P. 2069 — 2079.

7. **Jou Chichang, Lee Hang-Chang.** Handwritten numeral recognition based on simplified structural classification and fuzzy memberships // Expert System Appl. 2009. V. 36. N 9. P. 11858 — 11863.

8. **Qiao Y., Nishiara M., Yasuhara M.** A framework toward restoration of writing order from single-stroked handwriting image // IEEE Trans. pattern analysis and machine intelligence. 2006. V. 28. N 11. P. 1724 — 1737.

9. **Князев А.В.** Методы распознавания рукописных слов // Вестник МЭИ. 2012. № 6. С. 192 — 197.

10. **Freeman H.** On the encoding of arbitrary geometric configuration // IEEE Trans. electronic computers. 1961. V. 10. P. 260 — 268.

Статья поступила в редакцию 23.10.2015