

УДК 004.932.75'1

Н. С. Исупов, аспирант, Ижевский государственный технический университет имени М. Т. Калашникова
А. В. Кучуганов, кандидат технических наук, Ижевский государственный технический университет имени М. Т. Калашникова

РАСПОЗНАВАНИЕ СЛИТНЫХ РУКОПИСНЫХ ТЕКСТОВ С ИСПОЛЬЗОВАНИЕМ АППАРАТА НЕЧЕТКОЙ ЛОГИКИ*

Представлен краткий обзор основных методик распознавания. Предложен новый подход в распознавании слитных рукописных текстов, основанный на представлении векторизованного текста в виде нечетких нагруженных графов.

Ключевые слова: рукописный текст, аппарат нечеткой логики.

Существует множество работ, связанных с автоматизацией процесса распознавания графических изображений, начиная с печатных символов и до обработки трехмерных изображений [1, 2]. Используются различные методики по улучшению качества исходного изображения и выделению признаков, по которым ведется распознавание. Тем не менее возможности интеллектуального анализа изображений с помощью компьютеров оставляют желать лучшего. Так, например, существует острая необходимость в создании эффективных систем распознавания рукописных текстов.

Известны несколько крупных проблем в задаче распознавания рукописных текстов:

- большое количество разновидностей почерка;
- индивидуальные особенности почерка, такие как раздельное написание некоторых элементов слов или наличие декоративных элементов;
- сложность выявления отдельных символов в слитном рукописном слове;
- зависимость написания символа от его положения в слове.

На данный момент существует ряд систем распознавания текстов, например ABBYY Finereader, OmniPage, OCR CUNEIFORM, Readiris, Microsoft Office Document Imaging. Все они довольно успешно справляются с задачей распознавания печатных текстов, в то время как задача распознавания слитных рукописных текстов является нерешенной.

В теории распознавания образов выделяются различные группы признаков, по которым ведется сравнение распознаваемой области и эталона. По типу выделяемых признаков методы распознавания делятся [3]:

- на статистические;
- детерминированные;
- логические;
- структурные или лингвистические;
- нейросистемные.

Нередко системы распознавания создаются на основе комбинации этих методов.

Статистические методы основаны на вычислении различных статистических величин для последующего их сравнения (коэффициент корреляции, мо-

менты различных порядков, закон распределения, матрицы смежности). Недостатки:

- значительные вычислительные затраты;
- неизвестность закона распределения;
- неоднозначность описания области изображения.

Детерминированные методы основаны на выделении из исходного изображения признаков, принимающих конкретные числовые значения, например геометрических величин, таких как периметр, площадь фигуры, средняя длина хорды, эксцентриситет, периметр наименьшей выпуклой формы, описывающей объект и т. д. Недостатком этих методов является необходимость качественной сегментации изображения. В противном случае достоверность признаков в значительной степени снижается.

Логические методы основаны на выделении таких признаков, которые можно рассматривать как элементарные высказывания, принимающие два значения истинности («да», «нет» или «истина», «ложь») с полной определенностью, что затруднительно при построении реальных систем.

Структурные методы основаны на представлении как всего изображения, так и изображения отдельного субъекта в виде совокупности некоторых примитивных геометрических элементов (непроизводные элементы) и их отношений между собой [4]. Главным недостатком структурных методов является необходимость четкого разделения изображения на объект и фон, что затруднительно в условиях реального изображения.

Нейросистемные методы. В основу этого подхода положен принцип действия нейронных клеток головного мозга человека. Нейрон представляет собой устройство, состоящее из нескольких входов, каждый из которых имеет свою весовую функцию, а также из блока суммирования и одного выхода [5]. Значение признака, образующееся на выходе нейрона, зависит от входного сигнала, весовых коэффициентов входов и функции возбуждения. Недостатком нейросети является необходимость качественной предобработки изображения. Другой недостаток заключается в отсутствии инвариантности ко многим видам топологических преобразований.

Проанализировав достоинства и недостатки вышеперечисленных методов, было решено создать систему распознавания на основе структурно-лингвистического подхода с использованием лингвистических переменных и аппарата нечеткой логики.

Основы нечеткой логики были заложены в конце 60-х годов в работах известного американского математика Лотфи Заде [6]. Лингвистическая переменная может принимать значения понятий (фраз) естественного языка и используется при описании объектов и явлений с помощью нечетких множеств. Например, лингвистическая переменная «скорость» может принимать значения «очень быстро», «быстро», «медленно», «очень медленно»; переменная «направление» – «север», «юг», «запад», «восток». Они, в свою очередь, являются нечеткими переменными и изменяются в некотором диапазоне числовых значений.

Сущность предлагаемого подхода заключается в следующем.

1. Изображение символа можно представить в виде графа, состоящего из узлов и ветвей (рис. 1).

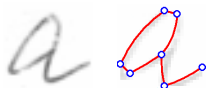


Рис. 1. Пример представления рукописной буквы «а» в виде графа

2. Выделяется ограниченное множество вариантов узлов (рис. 2), называемых типовыми опорными узлами (ТОУ).

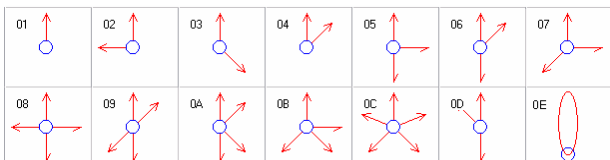


Рис. 2. Типы узлов

3. Параметры ветвей графа преобразуются в нечеткие переменные. Ветви имеют форму, а также показатели направления и ориентации. Отличие ориентации от направления заключается в том, что направление высчитывается относительно предыдущего отрезка, а ориентация не зависит от других элементов. Направление, ориентация, форма и длина ветви являются лингвистическими переменными, принимающими ограниченное количество значений. В табл. 1 показаны различные значения лингвистических переменных направления и ориентации ветви.

4. Эталоны символов хранятся в базе данных в виде аналогичных графов.

Проводить распознавание предлагается путем сравнения ТОУ, а также направлений и количеств исходящих из них ветвей в исходном изображении и в эталонах.

Внешний вид системы представлен на рис. 3.

Имеются следующие режимы работы с программой:

- 1) работа с базой типовых опорных узлов;
- 2) работа с библиотекой эталонов;
- 3) формирование скелетного изображения;
- 4) предобработка скелетного изображения;
- 5) распознавание.

Таблица 1. Сводная таблица значений направления и ориентации ветвей в зависимости от угла

Количественный диапазон	Качественное значение направления	Качественное значение ориентации
-11,25...11,25°	ForWard	North
11,25...33,75°	FFR	NNE
33,75...56,25°	FR	NE
56,25...78,75°	FRR	EEN
78,75...101,25°	Right	East
101,25...123,75°	RRB	EES
123,75...146,25°	RB	SE
146,25...168,75°	RBB	SSE
168,75...191,25°	Back	South
191,25...213,75°	BBL	SSW
213,75...236,25°	BL	SW
236,25...258,75°	LLB	WWS
258,75...281,25°	Left	West
281,25...303,75°	FLL	WWN
303,75...326,25°	FL	NW
326,25...348,75°	FFL	NNW

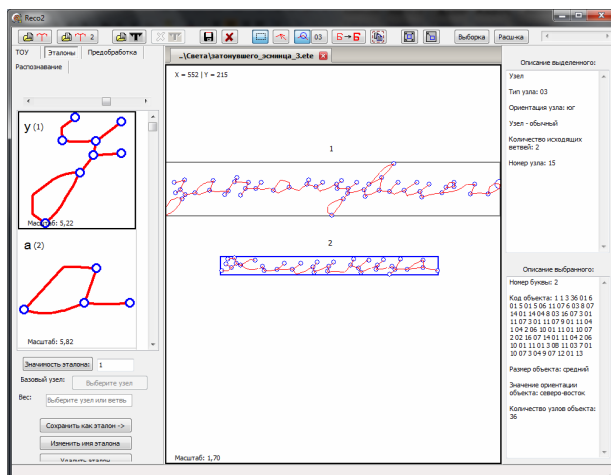


Рис. 3. Внешний вид системы Reco2

В режиме работы с базой ТОУ имеется возможность просматривать, добавлять и удалять типовые опорные узлы, выбирая их из контрольного изображения.

Перед предобработкой входное изображение обрабатывается векторизатором ArtPhoto [7] для получения скелетного изображения (рис. 4).

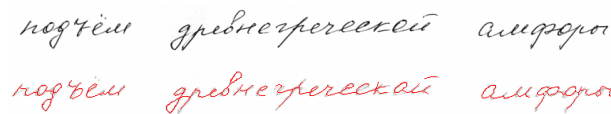


Рис. 4. Исходное и скелетное изображение текста

Процесс предобработки состоит из нескольких этапов.

1. Фрагментация изображения – разбиение всего текста на слова и символы (рис. 5). В программе реализован статистический метод разбиения исходного текста на фрагменты.



Рис. 5. Пример фрагментации текста

2. Устранение мелких разрывов между концами отрезков.

3. Исключение коротких отрезков (рис. 6).



Рис. 6. Текст после исключения коротких отрезков и устранения мелких разрывов

4. Корректировка направления отрезков.

5. Аппроксимация дугами.

6. Распознавание ТОУ типа разветвление, конец, угол (рис. 7).

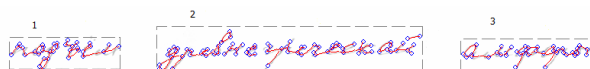


Рис. 7. Текст после распознавания узлов

7. Формирование нечеткого графа.

Можно выполнять каждый этап по отдельности или задать автоматическое выполнение всех этапов предобработки.

В режиме распознавания фрагменты изображения, прошедшие обработку, сравниваются с эталонами, и выдается результат. Любой из обработанных фрагментов можно задать в качестве эталона в режиме работы с библиотекой эталонов.

В процессе экспериментальных исследований было выявлено, что при представлении рукописных слов в виде нечетких нагруженных графов некоторые буквы включают в себя более простые (рис. 8).

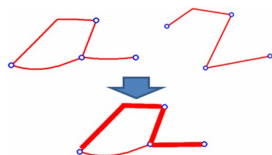


Рис. 8. Символ «а» включает в себя символ «г»

Соответственно, можно выделить две группы символов – составные и простые (табл. 2).

В результате проведенных экспериментов было решено реализовать алгоритм на основе метода сборки.

Сущность предлагаемого метода заключается в следующем.

1. Поиск простых символов.

2. Поиск фрагментов составных символов по предполагаемым ключевым элементам.

Таблица 2. Зависимости между различными символами отдельно взятого рукописного шрифта

Составной символ	Простые символы
а	г
в	е, а, г, о
е	о
з	н, о
и	а, г
м	н, а, ш
н	а, и
у	о, з, а, г, и
ш	г, о, н, м, и, с

Таким образом, любой сложный символ собирается по частям (рис. 9).

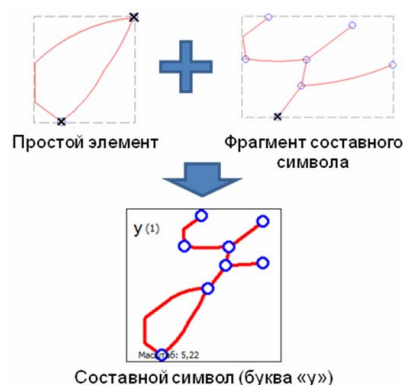


Рис. 9. Сущность метода сборки

Экспериментальные исследования с рукописными текстами показали, что на данный момент надежность распознавания составляет 70 % при среднем качестве носителя информации.

В настоящее время ведется доработка экспериментальной системы распознавания с целью включения морфемного словаря в состав системы и детальной проработки алгоритма распознавания на основе метода сборки. В дальнейшем планируется проведение экспериментальных исследований на архивных документах (скорости 17–19 вв).

Библиографические ссылки

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М. : Техносфера, 2006. – 1072 с.
2. Горский Н., Анисимов В., Горская Л. Распознавание рукописного текста. – СПб. : Политехника, 1997.
3. Горелик А. Л., Скрипкин В. А. Методы распознавания. – М. : Высш. шк., 1989.
4. Фу К. Структурные методы в распознавании образов. – М. : Мир, 1977.
5. Казаков Я. В. Повышение качества киноизображения средствами вычислительной техники. – Ч. 2. Методы признакового описания статичного изображения. – URL : <http://www.sciteclibrary.ru/rus/catalog/pages/5026.html>
6. Заде Л. А. Роль мягких вычислений и нечеткой логики в понимании, конструировании и развитии информационных/интеллектуальных систем // Новости Искусственно Интеллекта. – 2001. – № 2-3. – С. 7–11.
7. Kuchuganov A. V. Recursions in Image Analysis Problems // Pattern Recognition and Image Analysis. – 2009. – Vol. 19. – No. 3. – P. 501–507.

N. S. Isupov, Post-graduate, Kalashnikov Izhevsk State Technical University

A. V. Kuchuganov, PhD in Engineering, Kalashnikov Izhevsk State Technical University

Joined-up Writing Recognition with Fuzzy Logic Application

A brief survey of main recognition methods is represented in this paper. A new approach of joined-up writing recognition based on fuzzy graph presentation of vectorized text is proposed.

Key words: handwriting, fuzzy logic.

УДК 004.93

С. И. Зыкин, студент, Ижевский государственный технический университет имени М. Т. Калашникова

М. А. Сенилов, доктор технических наук, профессор, Ижевский государственный технический университет имени М. Т. Калашникова

РЕАЛИЗАЦИЯ АЛГОРИТМА ВИОЛЫ – ДЖОНСА НА ОСНОВЕ ТЕХНОЛОГИИ *CUDA* ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Реализация алгоритма Виолы – Джонса для архитектуры *NVIDIA CUDA* позволяет в несколько раз ускорить его работу по сравнению с реализацией на центральном процессоре при распознавании объектов на изображениях.

Ключевые слова: алгоритм Виолы – Джонса, технология *CUDA*, обнаружение объектов, анализ изображений.

Задача распознавания объектов на изображениях является в настоящее время весьма актуальной. Она возникает при создании естественного человеко-машинного интерфейса, построении охранных систем и пр.

Существует множество методов обнаружения объектов на изображениях, таких, например, как методы с использованием контурных моделей, световой сегментации, нейронных сетей и др. Одним из наиболее известных алгоритмов нахождения объектов является детектор Виолы – Джонса.

Алгоритм Виолы – Джонса был представлен в 2001 г. Полом Виолой и Майклом Джонсом [1]. Этот детектор является одним из лучших по соотношению показателей «эффективность распознавания/скорость работы».

Для быстрого вычисления суммы пикселей вычисляется промежуточное представление изображения, которое называется интегральным изображением [2]. Значение элемента интегрального изображения с координатами x, y равняется сумме пикселей исходного изображения выше и левее координат x, y включительно:

$$IntImage(x, y) = \sum_{x' \leq x, y' \leq y} Image(x', y'), \quad (1)$$

где $IntImage$ – интегральное изображение; $Image$ – исходное изображение.

Детектор сканирует входное изображение на многих масштабах, начиная с базового значения. Процесс сканирования представляет собой процесс просмотра вырожденного дерева решений, которое называется каскадом (рис. 1) [3].

Отрицательный результат в любой точке каскада приводит к немедленному отказу от субокна. После

нескольких этапов обработки число субокон сокращается радикально, и оставшиеся окна являются решением поставленной задачи.

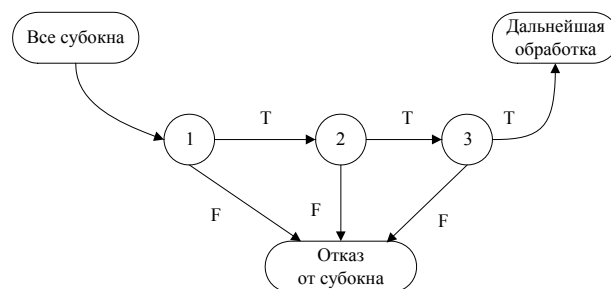


Рис. 1. Схематичное представление каскада обнаружения

Технология *CUDA* использует параллельную модель вычислений, когда каждый из *SIMD*-процессоров выполняет одну и ту же инструкцию над разными элементами данных параллельно [4].

Из-за параллельных вычислений трудно сразу вычислить интегральное изображение, так как нужно контролировать порядок вычислений. Поэтому вычисление разбито на две части.

1. Нахождение суммы по столбцам с использованием следующих формул:

$$\begin{aligned} b(x, -1) &= 0; \\ b(x, y) &= b(x, y-1) + Image(x, y), \end{aligned} \quad (2)$$

где $b(x, y)$ – сумма значений по столбцам.

2. Вычисление интегрального изображения по полученным ранее значениям:

$$\begin{aligned} IntImage(-1, y) &= 0; \\ IntImage(x, y) &= IntImage(x-1, y) + b(x, y). \end{aligned} \quad (3)$$