

Учреждение образования  
“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ”

УДК 681.327.12

МАЛЕНКО  
Олег Григорьевич

**РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ ДЛЯ СИСТЕМ  
РАСПОЗНАВАНИЯ РУКОПИСНЫХ СИМВОЛОВ НА ОСНОВЕ  
ОБУЧАЕМЫХ КЛАССИФИКАТОРОВ**

Специальность 05.13.15 – Вычислительные машины и системы

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
кандидата технических наук

Минск 2001

Работа выполнена в Учреждении образования “Белорусский государственный университет информатики и радиоэлектроники”

Научный руководитель –  
доктор технических наук, профессор Садыхов Р.Х.  
(БГУИР, каф. ЭВМ)

Официальные оппоненты:  
доктор технических наук, профессор Птичкин В.А.  
(БГУИР, каф. Информационных технологий автоматизированных систем)  
кандидат физико-математических наук, доцент Краснопрошин В.В.  
(БГУ, каф. Математического обеспечения АСУ)

Защита состоится 6 декабря 2001 года в 14 часов на заседании совета по защите диссертаций Д 02.15.04 при Учреждении образования “Белорусский государственный университет информатики и радиоэлектроники” по адресу: 220027, г.Минск, П.Бровки, 6, а. 232-1 корп., тел. 2398989.

С диссертацией можно ознакомиться в библиотеке Учреждения образования “Белорусский государственный университет информатики и радиоэлектроники”.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы диссертации.** В настоящее время планка развития средств вычислительной техники достигла уровня, позволяющего переложить на плечи ЭВМ часть задач, ранее решаемых только человеком. Возрастание объемов и сложности обрабатываемой графической информации привело к необходимости разработки методов ее компьютерного анализа и распознавания. Одной из классических задач анализа изображения, решаемых человеком, является восприятие текстовой информации. Для того, чтобы компьютер мог анализировать такую информацию она должна быть представлена в цифровом виде – цепочке кодов символов. Исторически сложилось так, что текстовая информация чаще всего изначально существует в виде изображения на бумажном носителе, поэтому его необходимо ввести в компьютер, то есть выделить, распознать и закодировать кодами цепочки символов на нем. При больших объемах текстовой информации использование человека для ее ввода крайне неэффективно из-за малой скорости набора, а также ошибок, обусловленных рутинной и усталостью. Поэтому как никогда остро стоит задача автоматического распознавания текста по его изображению, полученному с помощью устройств ввода изображения (сканер, видеокамера).

Существуют успешные решения этой задачи для ввода печатного текста. Однако наиболее сложной для восприятия машиной текстовой информации остается рукописная. Актуальность задачи машинного восприятия такой информации определяется тем, что рукописный текст очень удобен для кодирования информации человеком, поскольку требует минимальных средств (бумага и ручка) и производится с приемлемой для человека скоростью. Большие вариации рукописного написания символов, сложность их выделения, являются основной причиной того, что задача их распознавания с использованием ЭВМ решается с невысокими показателями по достоверности. Существуют лишь отдельные успешные попытки ее решения, как правило, ориентированные на конкретную задачу. Тем не менее, набор методов и алгоритмов распознавания рукописной информации постоянно расширяется и совершенствуется. Для практического решения конкретной задачи распознавания необходимо владеть всем этим набором и развивать его, поскольку универсального способа решить любую задачу в этой области пока не существует.

Практическое применение методов распознавания рукописных символов возможно в широком спектре задач:

- использование почтовыми службами при чтении адресов, рукописных или печатных почтовых кодов;
- использование в издательской деятельности для ввода рукописей;
- применение для непосредственной обработки документов: как многоцелевое считывающее средство в обработке широкого диапазона данных, от считывания микрофильмов - до электронных обработчиков больших объемов корреспонденции, бланков, анкет;
- для автоматического распознавания номеров транспортных средств в реальном масштабе времени по изображению, полученному с помощью видеокамеры.

Спектр методов и алгоритмов распознавания символов может применяться и в смежных задачах:

- Верификация подписей (сигнатур).
- Анализ сцен.
- Обработка и идентификация снимков земной поверхности.
- Контроль и идентификация изображений электронных схем и печатных плат.

**Связь работы с крупными научными программами, темами.**  
Диссертационная работа выполнена на кафедре ЭВМ Белорусского Государственного Университета Информатики и Радиозлектроники в сотрудничестве с лабораторией идентификации систем Института технической кибернетики НАН Беларуси в рамках тем: “Разработка концепции синтеза нового класса мультипроцессорных систем на принципах самоорганизации”, решение Совета ФФИ РБ от 12.01.96 пр. №1, разд. 1; “Разработка структурных моделей клеточных автоматов для задач синтаксического распознавания образов”, программа “Волна”, приказ Минобразования РБ от 16.02.96 №60; “Разработать методы, алгоритмы анализа и синтеза изображений и пространственных образов на основе спектральных моделей для создания оптимальных систем распознавания и компьютерного зрения”, письмо Минобразования РБ исх. 05-9/5 от 13.01.97; “Разработать систему реального времени для обработки видеоизображений”, программа “Информатика”, утверждено постановлением Совета Министров РБ от 17.01.97 №7, задание 01.02.10; “Разработка методов и алгоритмов статистического распознавания образов на основе спектральных инвариантов и моделей нейронных сетей”, задание Минобразования РБ исх. 05-8/291 от 15.12.98; “Разработка методов и алгоритмов синтеза клеточных автоматов на основе нейросетевых технологий”, решение Совета ФФИ РБ от 26.02.99 пр. №1; “Разработать компьютерную систему реального времени для идентификации образов в задачах экспертизы”,

программа “Информатика”, приказ ГКНТ от 27.07.99 №198, задание 01.05; “Intelligent neural system for autonomous control of a mobile robot”, фонд INTAS 97-2028.

**Цель диссертационной работы** - разработать методы и алгоритмы обработки изображений и распознавания образов применительно к идентификации рукописных символов.

Для достижения поставленной цели необходимо решить следующие задачи:

- выполнить классификацию существующих направлений методов распознавания изображений, выделив наиболее перспективные;
- разработать алгоритмы предварительной обработки изображений рукописных символов;
- исследовать и создать алгоритмы выделения информативных признаков для распознавания рукописных символов;
- разработать универсальные и надежные методы построения классификаторов для распознавания рукописных символов;
- исследовать и практически реализовать указанные выше алгоритмы в конкретных системах распознавания.

**Объект и предмет исследования.** В данной диссертационной работе объектами исследования являются изображения, причем особый акцент сделан на изображения рукописных символов. Предметом исследования являются методы распознавания графических образов.

**Методология и методы проведенного исследования.** В качестве аппарата исследования использовались методы статистического и структурного распознавания образов, нейронные сети. Для практической реализации разработанных алгоритмов использовались методы создания программных систем, методы оптимизации программных комплексов, объектно-ориентированный подход к технологии проектирования программного обеспечения.

**Научная новизна и значимость полученных результатов.**

- Разработан метод аппроксимации и устранения разрывов при структурном распознавании, позволяющий значительно упростить полученную из изображения символа структурную модель.
- Усовершенствован метод структурного распознавания, основанный на последовательностях примитивов, что позволило расширить область его применения.

- Разработан метод минимизации обучающей выборки для задачи распознавания рукописных символов, позволяющий увеличить скорость с сохранением качества.

- Создан метод сжатия, ориентированный на уменьшение размеров изображения рукописного символа с сохранением значимой информации о его форме.

- Разработана методика оценки качества алгоритмов выделения признаков с помощью конкурентной нейронной сети, позволяющая оптимально подобрать набор и количество информативных признаков, а также количество и состав классов в системе распознавания.

- Выполнен анализ наиболее перспективных методов и алгоритмов каждого этапа работы системы распознавания для структурного и статистического подхода и даны практические рекомендации по применению этих методов в конкретных системах.

**Практическая значимость полученных результатов.** Рассмотренные в диссертационной работе структурные и статистические методы распознавания рукописных символов могут использоваться для создания самых различных систем распознавания образов и обработки изображений, оптимизации скорости и качества их функционирования. Результаты диссертационной работы внедрены и используются в аппаратно-программном комплексе для создания электронных архивов автоматически вводимых больших массивов первичных документов “СИМВОЛ”, разработанным в УП “НИИЭВМ”, а также в учебном процессе на кафедре ЭВМ БГУИР в дисциплинах “Цифровая обработка сигналов и изображений”, “Машинная графика”.

#### **Основные положения диссертации, выносимые на защиту.**

- 1) Метод аппроксимации и устранения разрывов, позволяющий значительно упростить полученный по результатам скелетизации граф, представив его минимальным числом вершин и ребер, необходимых для передачи структуры символа, и устранить имеющиеся в нем ошибки. Применение метода приводит к значительному сокращению размерности описания, ускорению дальнейшего распознавания, уменьшению вариации входных символов и улучшению работы системы без усложнения алгоритмов других этапов и увеличения обучающей выборки.

- 2) Метод структурного распознавания, основанный на описании структуры рукописного символа с помощью последовательностей примитивов, устойчивых к геометрическим искажениям, свойственным распознаваемым образам. В комбинации с указанным методом описания используется классификатор, основанный на сравнении с эталоном.

3) Метод минимизации обучающей выборки для задачи распознавания рукописных символов, позволяющий значительно сократить размер базы идеальных описаний и ускорить процесс классификации, основанный на вычислении функции расстояния с эталоном. В основе метода лежит понятие покрытия одного идеального описания другим.

4) Метод сжатия изображения рукописного символа с сохранением сходства сжатого изображения с исходным, позволяющий сократить размерность задачи распознавания.

5) Методика для анализа и оптимизации выбора алгоритма выделения информативных признаков и его параметров, основанная на оценке качества кластеризации обучающей выборки с помощью самоорганизующихся нейронных сетей. Этот подход позволяет добиться качественного улучшения функционирования системы распознавания, недостижимого иными путями.

**Личный вклад соискателя.** Все предлагаемые методы были разработаны и программно реализованы лично автором. Научный руководитель принимал участие в постановке задач, определении возможных путей решения.

**Апробация результатов диссертации.** Результаты диссертационной работы докладывались и обсуждались на The Fourth International Conference "Pattern Recognition and Information Processing" (PRIP'97, Minsk, 1997), The Fifth International Conference on Advanced Computer Systems ACS-98 (Szczecin, Poland, 1998), Международной конференции "Цифровая обработка информации и управление в чрезвычайных ситуациях" ЦОИУЧС-98 (1998), Научно-технической конференции профессорско-преподавательского состава, аспирантов и студентов (Брест, 1998), Международной научно-технической конференции "Новые информационные технологии в науке и производстве", БГУИР (Минск, 1998), IEEE International Workshop of Intelligent Signal Processing (Budapest, Hungary, 1999), IAPR Workshop on Machine Vision and Applications (Tokyo, Japan, 2000).

**Опубликованность результатов.** По материалам выполненных дований опубликовано 12 научных работ, в том числе: 5 статей, 7 доводов на международных и республиканской конференциях. Общее количество страниц публикаций: 78.

**Структура и объем диссертации.** Диссертация изложена на 175 ицах машинописного текста, содержит 58 рисунков (30 страниц), 4 цы (2 страницы), и состоит из введения, общей характеристики работы, их глав, заключения, списка использованных источников на 7 страницах, ающего 101 названий, и 2 приложений на 26 страницах.

## ОСНОВНОЕ СОДЕРЖАНИЕ

**Во введении** обоснована актуальность темы диссертационной работы, сформулированы цель и задачи исследования, приведена краткая характеристика работы.

**В первой главе** исследуется задача распознавания рукописных символов. Приводится общая структура системы распознавания, состоящая из блока предварительной обработки, блока выделения признаков и классификатора, вводятся основные понятия: образ, класс, признак, пространство признаков.

Рассматриваются два основных подхода к распознаванию рукописных символов: структурный и статистический, их достоинства и недостатки. Приводится обзор известных методов и алгоритмов их реализации. Перспективным представляется гибридное направление, хотя для его развития необходимо значительно усовершенствовать оба подхода с целью повышения универсальности и уменьшения зависимости от человека.

Определяется, что наиболее важным этапом работы системы распознавания, определяющим скорость и качество ее работы и обучения, является выделение информативных признаков. В настоящее время практически отсутствуют методы, позволяющие оптимально выбрать алгоритм выделения признаков и оценить их количество, необходимое для разделения классов в пространстве признаков. Тем не менее, выбор информативных признаков представляется основным направлением оптимизации систем распознавания. Необходимо доказать, что оптимальный выбор информативных признаков позволяет достичь максимального эффекта, не достижимого другими средствами, такими как усложнение и оптимизация классификатора.

Наряду с этим определяются и другие перспективные направления развития методов распознавания рукописных символов: повышение их надежности, быстродействия, универсальности и независимости от человека.

**Во второй главе** рассматриваются структурные методы распознавания рукописных символов. Приводится структура системы распознавания на основе сравнения с эталоном, как наиболее универсальная. Структурные методы, позволяющие выделять и распознавать структуру, характеризуются устойчивостью к искажениям, свойственным рукописным символам: вариациям размера, пропорций, угла наклона, толщины линии.

Основным направлением выделения признаков в структурном распознавании рукописных символов является выделение структуры символа с помощью скелетизации, выделения графа и его упрощения (аппроксимации).

К важным проблемам структурного распознавания необходимо добавить проблему устранения разрывов. Актуальным представляется исследовать

разрывы, характерные для рукописных символов, и разработать быстрый алгоритм их устранения. Разрывы можно разделить на три типа (рис. 1).

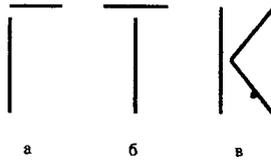


Рис. 1. Классификация разрывов

Назовем разрывом второго типа ситуацию, когда квадрат расстояния между вершиной кратности 1 графа и ближайшей точкой некоторого ребра, не инцидентного вершине и не связанного с ней ребром, меньше порога  $P_2$ .

Пусть задано ребро  $A_1A_2$  (рис. 2) и есть точка  $B$ , не лежащая на нем. Требуется найти координаты  $x$  и  $y$  точки  $C$ , принадлежащей ребру и самой близкой к  $B$ .

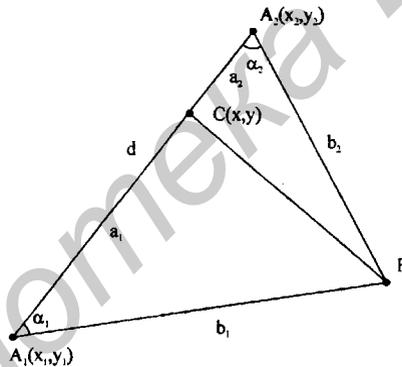


Рис. 2. Нахождение ближайшей точки

В диссертационной работе получено следующее соотношение:

$$C(x, y) = \begin{cases} A_1(x_1, y_1), & \text{if } d^2 - b_2^2 + b_1^2 \leq 0 \\ A_2(x_2, y_2), & \text{if } d^2 - b_1^2 + b_2^2 \leq 0 \\ \left( x = \frac{(z-1)x_2 + x_1}{z}, y = \frac{(z-1)y_2 + y_1}{z} \right), & \text{if } \begin{matrix} d^2 - b_2^2 + b_1^2 > 0 \\ d^2 - b_1^2 + b_2^2 > 0 \end{matrix} \end{cases} \quad (1)$$

Таким образом, определяется ближайшая точка и если квадрат расстояния  $BC$  меньше порога  $P_2$ , можно говорить о наличии разрыва, который легко устраняется введением ребра  $BC$ . Следует отметить, что в результате устранения разрыва в графе могут появиться ребра, которые не подвергались

аппроксимации. То есть алгоритмы аппроксимации и устранения разрывов целесообразно применять в комплексе.

Далее во второй главе предлагается метод упрощенного описания и классификации, основанный на последовательностях примитивов (рис. 3), представляющих собой изгиб кривой определенного направления.

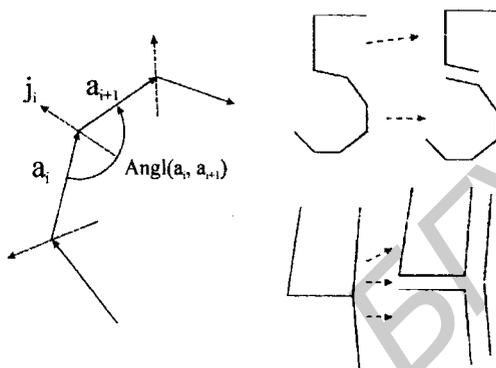


Рис. 3. Последовательности примитивов

Для каждой последовательности примитивов можно определить числовую характеристику, основанную на величине ее поворота. При этом поворот оценивается по разности углов биссектрис каждого угла кривой. Доказывается, что такой подход более устойчив к геометрическим искажениям, чем простая разность углов примитивов. Для упрощения расчетов угол биссектрис кодируется как  $j$  согласно его диапазону. Тогда, для последовательности примитивов:

$$a_1 \xrightarrow{j_1} a_2 \xrightarrow{j_2} a_3 \longrightarrow \dots \longrightarrow a_{n-1} \xrightarrow{j_{n-1}} a_n,$$

где  $j_i$  - характеристика соответствующей биссектрисы,  $i = \overline{1, n-1}$ , характеристика последовательности, определяющая поворот, определяется в виде соотношения:

$$P = 2 + \sum_{i=1}^{n-2} \{(j_{i+1} - j_i) \pmod{4}\}. \quad (2)$$

Соотношение (2) позволяет корректно оценить последовательность даже в случае, если угол ее искривления больше  $360^\circ$ . Если последовательность состоит только из двух примитивов, то  $P=2$ . Если из одного – считаем, что  $P=1$ . Число  $P$  является основной характеристикой последовательности, которая устойчива как к повороту, так и к наклону. Так, в ряде случаев для успешного распознавания очень важна информация о наклоне. Поэтому в описание

последовательности включается показатель наклона – характеристика первой биссектрисы  $j_i$ . Таким образом, пара  $(P, j_i)$  может быть принята в качестве описания последовательности примитивов.

Описание, основанное на этих характеристиках, используется при распознавании. Таким способом кодируется как распознаваемый объект, так и каждый класс в базе данных. Распознавание сводится к поиску наиболее близкого класса.

Пусть  $(P^{r^1}, j_i^{r^1}), (P^{r^2}, j_i^{r^2}), \dots, (P^{r^S}, j_i^{r^S})$  – список описаний последовательностей примитивов распознаваемого класса, а  $(P^{i^1}, j_i^{i^1}), (P^{i^2}, j_i^{i^2}), \dots, (P^{i^M}, j_i^{i^M})$  – список, соответствующий  $i$ -ому классу. Алгоритм распознавания заключается в следующем:

- 1) Распознаваемый символ сравнивается с  $i$ -ым классом, если он имеет то же число последовательностей примитивов ( $S=M$ ).
- 2) При каждом сравнении находится такая последовательность  $(P^{r^l}, j_i^{r^l}), (P^{r^2}, j_i^{r^2}), \dots, (P^{r^k}, j_i^{r^k})$ , чтобы число

$$R_i = \sum_{k=1}^S [A(P^{r^k} - P^{i^k}) + B\{(j_i^{r^k} - j_i^{i^k}) \pmod{4}\}] \quad (3)$$

было минимально для  $i$ -ого класса, где  $A$  и  $B$  – коэффициенты.

- 3) Если  $R_i$  меньше некоторого числа  $C$ , то  $i$ -ый класс помечается.
- 4) Если после поиска во всей базе помечен только один класс, то считается, что распознаваемый символ принадлежит этому классу.
- 5) Иначе, если классов несколько, то для выбора между ними используются информация о взаимном расположении последовательностей примитивов.

Числа  $A$ ,  $B$ ,  $C$  позволяют задавать степень влияния каждой характеристики на результат распознавания.  $A$  и  $B$  определяют влияние на  $R_i$  характеристик  $P$  и  $j_i$ , а  $C$  влияет на скорость работы системы и качество распознавания. При экспериментах было установлено, что задание  $A=6$ ,  $B=2$ ,  $C=12$  дает приемлемые результаты.

Обучение системы производится путем добавления в базу данных описаний тех символов, распознавание которых по текущей базе не приводило к успеху. Качество работы системы распознавания сильно зависит от обучения и выборки, которая при этом применяется. Для проверки качества распознавания применяется другая – тестовая выборка, образы из которой не входят в обучающую. Размер тестовой выборки обычно раз в 10 больше. При обучении на стилизованные рукописные символы успешно распознавалось до 95% символов из тестовой выборки. Попытка обучить систему на любые символы (стилизованные и раздельно написанные рукописные) приводила к 88% правильно распознанных символов.

Далее предложен метод минимизации обучающей выборки для структурного классификатора на основе сравнения с эталоном с помощью функции расстояния, примером которой является соотношение (3). Этот метод

специально ориентирован для работы в структурном классификаторе с функцией расстояния, где классические методы кластеризации неприменимы. В основе метода лежит понятие покрытия одним идеальным описанием другого, которое можно удалить без ущерба для качества распознавания, если данные описания относятся к одному классу. Применение метода позволило сократить объем базы идеальных описаний (ИО) на 25%, без существенной потери качества:

Тип образа	Кол-во классов	Базовое количество ИО	Кол-во ИО после минимизации	Размер тестовой выборки	Кол-во правильно распознанных	Время на символ, мс
Стилизованный символ	43	360	360	3500	3320 (94,8%)	19
Стилизованный символ	43	360	275 (76,4%)	3500	3309 (94,5%)	18
Произвольный символ	43	492	492	5000	4389 (87,8%)	23
Произвольный символ	43	492	370 (75,2%)	5000	4361 (87,2%)	20
Рукописная цифра	10	64	64	700	691 (98,7%)	10
Рукописная цифра	10	64	52 (81,2%)	700	690 (98,6%)	9

В третьей главе рассматриваются статистические методы распознавания рукописных символов. Рассматривается ряд алгоритмов выделения признаков, в том числе преобразование BIFORE (двоичное представление) – усеченное преобразование Адамара, перспективный метод сокращения размерности, применявшийся в экспериментах. Рассматриваются классические статистические классификаторы и их основной недостаток: необходимость линейной разделимости классов в пространстве признаков. Большое внимание уделено нейронным классификаторам, лишенным этого недостатка.

В рамках рассмотрения методов предварительной обработки и выделения информативных признаков особое внимание уделяется сокращению размерности задачи – сжатию. Предлагается метод сжатия, специально ориентированный на изображения рукописных символов, основной особенностью которого является сохранение визуального сходства между исходным и сжатым изображением. Метод является продолжением обычного масштабирования и позволяет добиться существенно большего сжатия без потери значимой информации о структуре и линиях символа.

Так как в основе сжатия лежит приведение окна большего размера к окну меньшего размера, то целесообразно привести его к преобразованию квадратной матрицы  $N*N$  (входная матрица) в матрицу  $M*M$  (выходная матрица), где  $N > M$  и  $N$  существенно меньше размеров исходного окна. Наилучшие результаты показали параметры  $N=4$  и  $M=2$ .

Теперь следует определить, каким образом покрывать входное окно матрицами  $4*4$ . Самое очевидное является покрытие (рис. 4а), не является

оптимальным, т.к. может привести к потере информации о линии, проходящей по границе матриц (рис. 4б).

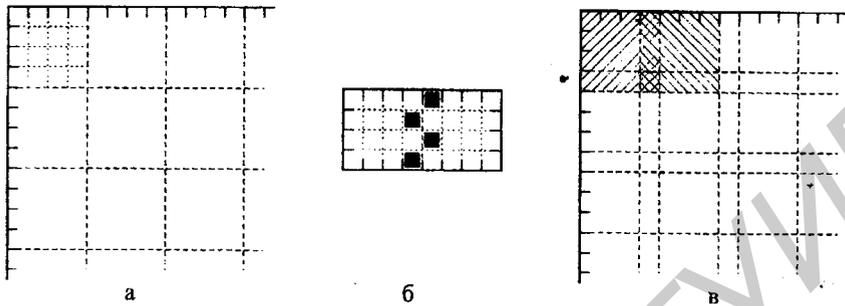


Рис. 4. Варианты разбиения входного окна

Перспективным является обеспечение перекрытия соседних как по горизонтали, так и по вертикали матриц на 1 строку или столбец (рис. 4в). В главе предложены четыре логические функции для начального кодирования преобразования матриц. Так как соседние входные матрицы перекрываются, то следует позаботиться о том, чтобы при этом дублировалось минимум информации. Это достигается путем присвоения каждому пикселю входной матрицы приоритета, с которым он влияет на выходную матрицу. Следует отметить, что вычислять логические функции при преобразовании каждой матрицы нет необходимости. Вместо этого функции вычисляются для каждой матрицы при инициализации системы распознавания, результат заносится в массив из 65536 элементов по индексу, кодирующему соответствующую входную матрицу. В каждом элементе массива хранится тетрада, кодирующая выходное окно. Таким образом, операция преобразования эквивалентна операции индексации в массиве и занимает минимальное время. Кроме того, массив можно редактировать, в случае необходимости изменений в функционировании алгоритма без трудоемкого изменения логических функций.

Далее показано, что применение нейросетевых классификаторов на базе многослойного персептрона позволяют осуществлять разделение классов в пространстве признаков произвольными границами. Экспериментально доказано, что наилучшие результаты при обучении и работе нейронного классификатора достигаются в условиях хорошей разделимости классов простыми границами. Поэтому в качестве главного направления для улучшения функционирования статистической системы распознавания предлагается поиск оптимального набора информативных признаков и алгоритма их выделения.

Отмечается, что хорошая разделимость достигается тогда, когда классы образуют кластеры в пространстве признаков. Предлагается методика для анализа и улучшения разделимости, основанная на применении

самоорганизующихся свойств конкурентной нейронной сети или сети Кохонена.

Допустим, что имеется несколько перспективных методов выделения признаков. В таком случае обработанную каждым из них обучающую выборку можно кластеризовать с помощью конкурентной нейронной сети, и по качеству разбиения выбрать наилучший. В том случае, если хорошее разбиение на кластеры не удастся получить сразу, конкурентную сеть можно использовать для подбора основных параметров системы, построенной на базе самого лучшего алгоритма выделения признаков. Предложены два способа улучшить разделимость классов, не меняя основу алгоритма выделения признаков:

1) Увеличить количество признаков. Это возможно за счет введения небольшого числа дополнительных признаков, полученных из исходного образа другими алгоритмами. Увеличенный набор можно снова проверить с помощью конкурентных методов.

2) Разбить обучающую выборку на классы по другому. Число классов можно искусственно увеличить и уменьшить. В первом случае несколько классов на выходе системы будут соответствовать одному решению. Для проверки возможности увеличения числа классов необходимо использовать сеть с числом нейронов большим, чем естественное число классов. В случае уменьшения числа классов, система строится на базе иерархического классификатора. Для проверки возможности следует использовать конкурентную сеть с числом нейронов меньшим, чем число классов. Признаком возможности объединения двух и более классов в один будет факт проецирования образов этих классов на один кластер и улучшение остальной кластеризации.

Можно использовать гибридный подход, разбивая одни классы и объединяя другие. Ключом к правильному выбору опять становится сопоставление результатов кластеризации с помощью конкурентной нейронной сети или сети Кохонена с информацией о разбиении обучающей выборки на классы. Кроме этого, конкурентную сеть можно использовать для проверки и фильтрации обучающей выборки.

**Четвертая глава** посвящена приложению разработанных методов и алгоритмов в различных системах распознавания.

Рассмотрена концепция построения программной оболочки экспериментальной системы распознавания и обработки изображения. Главная ее цель – объединить усилия разработчиков различных этапов системы распознавания для достижения общей цели, обеспечить повторное использование реализованных программных средств. Программная оболочка, как хранилище разработанных алгоритмов, особенно важна для систем распознавания, характеризующихся разнообразнейшим набором этапов и

методов их реализации. В рамках программной оболочки проводилась разработка систем, описанных в четвертой главе.

Разработана система распознавания номеров железнодорожных вагонов, функционирующая в реальном времени на базе статистических и нейронных методов (рис. 5).

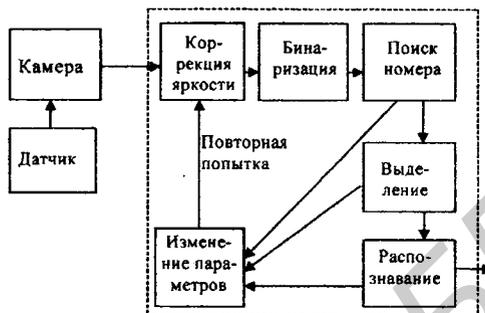


Рис. 5. Структура системы распознавания номеров вагонов.

Создана подсистема управления мобильным роботом основанная на анализе двух последовательных кадров с камеры, установленной на роботе. По двум изображениям, полученным в движении, имеется возможность получить представление о трехмерном рельефе местности перед роботом и избежать столкновения. В данной подсистеме используются структурный метод аппроксимации и устранения разрывов, предложенный во второй главе. Этот пример наглядно демонстрирует, что методы обработки и распознавания изображения рукописных символов могут успешно применяться и в других задачах, таких как задача анализа сцен (рис. 6).

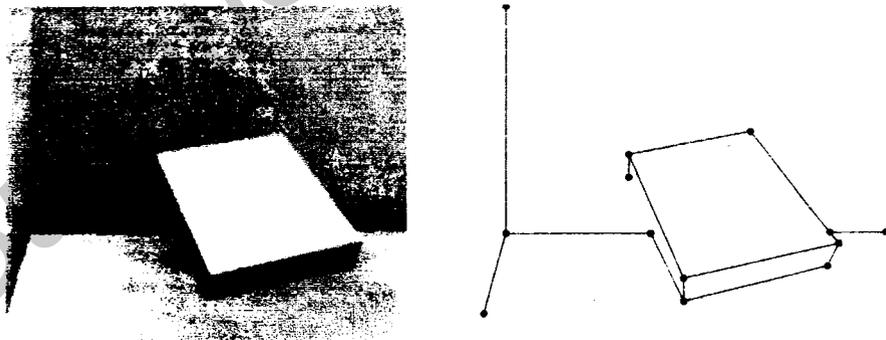


Рис. 6. Использование структурных методов для анализа сцены

Рассмотрена экспериментальная система распознавания рукописных цифр, разработанная для проверки всего комплекса структурных методов,

предложенных во второй главе, функционирующая с достоверностью около 99%. Примеры работы системы приведены ниже.

Исходное изображение	Упрощенная структура	Описание $(P, j)$	Класс	Расстояние
		$(3, 0)$	"1"	2
		$(3, 3), (3, 2)$	"2"	0
		$(3, 0), (2, 2), (3, 0)$	"3"	6
		$(5, 2), (2, 2), (2, 0), (2, 2)$	"4"	2
		$(3, 2), (3, 0)$	"5"	0
		$(7, 3), (4, 2)$	"6"	3
		$(2, 1), (3, 0)$	"7"	2
		$(0, 0), (0, 0), (3, 0), (3, 2)$	"8"	0
		$(11, 2), (3, 0)$	"9"	8
		$(0, 0)$	"0"	0

## ЗАКЛЮЧЕНИЕ

Результаты, полученные в ходе диссертационного исследования, можно сформулировать следующим образом:

1. Было выявлено, что каждое направление распознавания рукописных символов имеет свои достоинства и недостатки. В каждой конкретной задаче целесообразно применять один из подходов: структурный или статистический. Перспективным представляется гибридное направление. Определено, что наиболее важным этапом работы системы распознавания, определяющим скорость и качество ее работы и обучения, является выделение информативных признаков. Выбор информативных признаков представляется основным направлением оптимизации систем распознавания. Экспериментально доказано, что оптимальный выбор информативных признаков позволяет достичь эффекта, не достижимого другими средствами, такими как усложнение и оптимизация классификатора.

2. Разработан метод аппроксимации структуры символа, позволяющий значительно упростить полученный по результатам скелетизации граф, представив его минимальным числом вершин и ребер, необходимых для передачи структуры символа. Этот подход способствует ускорению дальнейшей обработки образа системой и сокращает размерность информации. Устранение разрывов способствует сокращению вариаций входных символов и позволяет улучшить функционирование системы без усложнения алгоритмов и увеличения обучающей выборки. Разработан эффективный метод поиска и устранения разрывов, предназначенный для применения вместе с методом аппроксимации [4, 9].

3. Предложен метод упрощения структуры и классификации, основанный на выделении последовательностей примитивов, устойчивых к искажениям [2, 3]. В диссертационной работе предложено использовать метод классификации, основанный на сравнении с эталоном. Это позволяет упростить формирование последовательностей примитивов, а качество сравнения обеспечивается кодированием отношений между ними. Кроме того, многообразие можно значительно сократить с помощью методов устранения разрывов. Применение классификатора, сравнивающего с эталоном, позволило упростить и ускорить обучение системы распознавания, повысить ее универсальность, а упрощение формирования окончательного описания – существенно повысить быстродействие. Разработан метод минимизации обучающей выборки, ориентированный на применение в комплексе с описанным выше классификатором и позволяющий сократить размер базы идеальных описаний и тем самым ускорить классификацию [5, 8]. Этот метод специально ориентирован для работы в структурном классификаторе с функцией расстояния, где классические методы кластеризации неприменимы. В

основе метода лежит понятие покрытия одним идеальным описанием другого, которое можно удалить без ущерба для качества распознавания, если данные описания относятся к одному классу. Применение метода позволило на 25% сократить объем базы идеальных описаний.

4. Предложен метод сжатия изображения рукописного символа, обеспечивающий визуальное сходство сжатого изображения с исходным [1, 6]. Указанный метод позволяет значительно сократить размерность задачи распознавания без потери значимой информации и, следовательно, повысить скорость без потери качества.

5. Применение нейросетевых классификаторов позволяют осуществлять разделение классов в пространстве признаков произвольными границами. Экспериментально доказано, что наилучшие результаты при обучении и работе нейронного классификатора достигаются в условиях хорошей разделимости классов простыми границами. Поэтому в качестве главного направления для улучшения функционирования статистической системы распознавания предложен поиск оптимального набора информативных признаков и алгоритма их выделения [7]. Отмечено, что хорошая разделимость достигается тогда, когда классы образуют кластеры в пространстве признаков. Предложена методика для анализа и улучшения разделимости, основанная на применении самоорганизующихся свойств конкурентной нейронной сети или сети Кохонена. Оценка строится на основе сравнения результатов самообучения указанных нейронных сетей с исходным разбиением обучающей выборки на классы. В случае плохого разбиения на кластеры в результате самообучения имеется возможность управлять процессом настройки параметров оптимизации, таких как число информативных признаков, алгоритм их выделения и количество классов.

6. Приведены описания некоторых экспериментальных и практических систем и подсистем, использующих методы, предложенные в диссертационной работе: системы распознавания номеров железнодорожных вагонов [7, 10, 12], подсистемы анализа изображения для управления мобильным роботом [11], системы структурного распознавания арабских [2, 3]. Все эти приложения создавались в рамках концепции единой программной оболочки, позволяющей эффективно экспериментировать и объединять труд нескольких разработчиков. Различные сферы практического применения указанных систем позволяют сделать вывод о возможности широкого использования полученных результатов в самых различных областях.

Полученные в диссертационной работе результаты предназначены не только для реализации в системах распознавания рукописных символов, но и многих других приложениях, связанных с обработкой изображения: распознавание снимков земной поверхности, анализ сцен, идентификация объектов на изображениях сверхбольших интегральных схем, идентификация дефектов печатных плат в системах технического зрения.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

## Статьи в сборниках и журналах

1. Садыхов Р.Х., Маленко О.Г. Метод предварительного сжатия бинарных изображений в системах распознавания рукописных символов / Институт технической кибернетики НАН Беларуси.– Минск, 1997.– 16 с.– Деп. В ВИНТИ 17.03.97, № 813-В1397 // Известия НАНБ. Сер. физ.-тех. наук.– 1997.– № 3.– С. 133.
2. Садыхов Р.Х., Маленко О.Г. Метод распознавания рукописных символов на основе графовых представлений // Автоматика и вычислительная техника.– 1998.– № 6.– С. 70-77.
3. Маленко О.Г., Садыхов Р.Х. Метод структурного описания графа в системе распознавания рукописных символов // Цифровая обработка изображений: Сб. ст. / Под ред. С.В.Абламейко.– Минск, 1999.– С. 51-60.
4. Садыхов Р.Х., Маленко О.Г. Метод устранения разрывов при структурном распознавании рукописных символов // Известия НАНБ. Сер. физ.-тех. наук.– 1999.– № 1.– С. 94-96.
5. Маленко О.Г., Садыхов Р.Х. Метод самоорганизации обучения и классификации при распознавании рукописных символов // Известия НАНБ. Сер. физ.-мат. наук.– 2000.– № 3.– С. 123-127.

## Тезисы докладов и материалы конференций

6. Rauf H. Sadykhov, Oleg G. Malenko. Preliminary compression of the binary images in systems of handwritten symbols recognition // Proceedings of fourth international conference “Pattern Recognition and Information Processing”, PRIP'97, Minsk, 20-22 May 1997. / Minsk–Szczecin, 1997. – Vol. 1.– P. 226-231.
7. Rauf Kh. Sadykhov, Oleg G. Malenko. Application of neural networks in system of numbers recognition for railway carriages // Proceedings of The Fifth International Conference “Advanced Computer Systems”, ACS-98, Szczecin, Poland, 19-20 November 1998. / Szczecin, 1998.– P. 99-105.
8. Маленко О.Г., Садыхов Р.Х. Метод минимизации обучающей выборки в задачах принятия решения // Первая международная конференция “Цифровая обработка информации и управление в чрезвычайных ситуациях”, Минск, 22-25 сентября 1998 г. Институт технической кибернетики НАН Беларуси.– Минск, 1998 – Тд2.– С. 114-119.

9. Маленко О.Г. Исследование и устранение разрывов в задачах структурного распознавания рукописных символов // Труды X научно-технической конференции профессорско-преподавательского состава, аспирантов и студентов, Брест, 31 марта–1 апреля 1998 г. / Брестский политехнический институт.– Брест, 1998.– Т.2.– С. 113-117.
10. Маленко О.Г., Садыхов Р.Х. Система автоматического поиска и распознавания номеров железнодорожных вагонов // Материалы международной научно-технической конференции “Новые информационные технологии в науке и производстве”, Минск, 25-27 ноября 1998 г. / БГУИР.– Минск, 1998.– С. 255-258.
11. Sadykhov R.Kh., Klimovich A.N., Malenko O.G. Method of spatial map forming for autonomous mobile robot control // Proceedings of IEEE International Workshop on Intelligent Signal Processing, Budapest, Hungary, 4-7 September 1999. / Budapest, 1999.– P. 242-245.
12. Sadykhov R.Kh., Klimovich A.N., Malenko O.G. and Podenok L.P. Automated TV-set Raster Tuning System // Proceedings of IARP Workshop on Machine Vision Applications, The University of Tokyo, Japan, 28-30 November 2000. / Tokyo, 2000.– P. 59-62.



МАЛЕНКА Алэг Рыгоравіч

РАСПРАЦОЎКА МЕТАДАЎ І АЛГАРЫТМАЎ ДЛЯ СІСТЭМ  
РАСПАЗНАВАННЯ РУКАПІСНЫХ СІМВАЛАЎ НА АСНОВЕ НАВУЧАЕМЫХ  
КЛАСІФІКАТАРАЎ

Ключавыя словы: распазнаванне вобразаў, рукапісныя сімвалы, структурныя спосабы распазнавання, нейронныя сеткі.

У дысертацыйнай рабоце даследавалася задача распазнавання рукапісных сімвалаў. Мэтай было распрацаваць і ўдасканаліць метады і алгарытмы апрацоўкі відарысу і распазнавання вобразаў адносна ідэнтыфікацыі рукапісных сімвалаў.

Распрацаван метада апраксімацыі і ліквідацыі разрываў, дазваляючы значна спрасціць апісанне сімвала і ліквідаваць наяўныя ў ім памылкі.

Удакладнён метада структурнага распазнавання, які заснаваны на апісанні структуры рукапіснага сімвала з дапамогай паслядоўнасцяў прымітываў, устойлівых да геаметрычных скажэнняў, уласцівых распазнаваемым вобразам. У камбінацыі з гэтым метадам апісання выкарыстоўваецца класіфікатар, заснаваны на параўнанні з эталонам.

Распрацаван метада мінімізацыі навучаючай выбаркі для задачы распазнавання рукапісных сімвалаў, дазваляючы значна скараціць размер базы ідэальных апісанняў і паскорыць працэс класіфікацыі, заснаваны на вылічэнні функцыі адлегласці з эталонам. У аснове метада пакладзена паняцце пакрыцця аднаго ідэальнага апісання другім.

Распрацаван метада сціскання відарысу рукапіснага сімвала з захаваннем падабенства сціснутага відарыса з зыходным, дазваляючы скараціць размернасць задачы распазнавання.

Распрацавана метадыка для аналізу і аптымізацыі выбара алгарытма выдзялення інфарматыўных прыкмет і яго параметраў, заснаваная на адзнацы якасці кластэрызацыі навучаючай выбаркі з дапамогай самаарганізуючыхся нейронных сетак. Гэты падыход дазваляе дасягнуць якаснага паляпшэння функцыянавання сістэмы распазнавання.

Тэарэтычна і эксперыментальна паказаны перавагі даных метадаў перад існуючымі.

Распрацаван шэраг сістэм распазнавання і апрацоўкі відарысу: сістэма распазнавання нумароў чыгуначных вагонаў, падсістэма аналізу відарыса для кіравання мабільнымі робатам, сістэма структурнага распазнавання арабскіх лічбаў.

МАЛЕНКО Олег Григорьевич

## РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ ДЛЯ СИСТЕМ РАСПОЗНАВАНИЯ РУКОПИСНЫХ СИМВОЛОВ НА ОСНОВЕ ОБУЧАЕМЫХ КЛАССИФИКАТОРОВ

Ключевые слова: распознавание образов, рукописные символы, структурные методы распознавания, нейронные сети.

В диссертационной работе исследовалась задача распознавания рукописных символов. Целью было разработать и усовершенствовать методы и алгоритмы обработки изображений и распознавания образов применительно к идентификации рукописных символов.

Разработан метод аппроксимации и устранения разрывов, позволяющий значительно упростить описание символа и устранить имеющиеся в нем ошибки.

Усовершенствован метод структурного распознавания, основанный на описании структуры рукописного символа с помощью последовательностей примитивов, устойчивых к геометрическим искажениям, свойственным распознаваемым образам. В комбинации с указанным методом описания используется классификатор, основанный на сравнении с эталоном.

Разработан метод минимизации обучающей выборки для задачи распознавания рукописных символов, позволяющий значительно сократить размер базы идеальных описаний и ускорить процесс классификации, основанный на вычислении функции расстояния с эталоном. В основе метода лежит понятие покрытия одного идеального описания другим.

Создан метод сжатия изображения рукописного символа с сохранением сходства сжатого изображения с исходным, позволяющий сократить размерность задачи распознавания.

Разработана методика для анализа и оптимизации выбора алгоритма выделения информативных признаков и его параметров, основанная на оценке качества кластеризации обучающей выборки с помощью самоорганизующихся нейронных сетей. Этот подход позволяет добиться качественного улучшения функционирования системы распознавания, недостижимого иными путями.

Теоретически и экспериментально показаны преимущества данных методов перед существующими.

Разработан ряд систем распознавания и обработки изображения: система распознавания номеров железнодорожных вагонов, подсистема анализа изображения для управления мобильным роботом, система структурного распознавания арабских цифр.

## RESUME

MALENKO Oleg Grigorevich

DEVELOPMENT OF METHODS AND ALGORITHMS FOR SYSTEMS OF HAND-WRITTEN SYMBOL RECOGNITION BASED ON TRAINED CLASSIFIERS

Key words: image recognition, hand-written symbols, structural recognition methods, neural networks.

The problem of hand-written symbols recognition has been investigated in PhD thesis. The main goal is development and improvement of methods and algorithms of image processing and recognition of hand-written symbols.

The method of approximation and breaks elimination has been developed. This approach allows to simplify the description of symbols and remove available errors.

The method of structural recognition based on the description of a structure for hand-written symbols with the help of primitive sequences, steady against geometrical distortions peculiar to images has been improved. In a combination with the specified method of the description the classifier based on comparison with the ideal image has been used.

The minimization method of training sample for a task of recognition of hand-written symbols allowing considerably to reduce the size of the base of the ideal descriptions and speed up process of classification based on distance function has been developed. The basis of a method is a concept of covering of one ideal description by an other.

The compression method of hand-written symbol image with similarity preservation of the compressed image with initial has been created, that allows to reduce dimension of recognition task.

The technique for the analysis and optimization of a choice of feature extraction algorithm and its parameters based on an estimation of clusterization quality of training sample with the help of self-organizing neural networks has been developed. This approach allows to achieve qualitative improvement of system functioning.

The advantages of the given methods in comparison with existing are shown both theoretically and experimentally.

A number of recognition and image processing systems has been developed: system of railway carriages number recognition, image analysis subsystem of mobile robot control, system of structural recognition of Arabian numerals.

МАЛЕНКО Олег Григорьевич

**РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ ДЛЯ СИСТЕМ  
РАСПОЗНАВАНИЯ РУКОПИСНЫХ СИМВОЛОВ НА ОСНОВЕ  
ОБУЧАЕМЫХ КЛАССИФИКАТОРОВ**

Специальность 05.13.15 – Вычислительные машины и системы

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
кандидата технических наук

---

Подписано в печать	31.10.2001.	Формат 60x84 1/16.
Бумага офсетная.	Печать ризографическая.	Усл.печ.л. 1,51.
Уч.-изд.л. 1,6.	Тираж 90 экз.	Заказ 517.

---

Издатель и полиграфическое исполнение:

Учреждение образования “Белорусский государственный университет  
информатики и радиоэлектроники”  
Лицензия ЛП № 156 от 05.02.2001  
Лицензия ЛВ № 509 от 03.08.2001  
220013, Минск, П. Бровки, 6.