

Х. АЛЬ-РАШАЙДЕХ, С. В. КУЛЕШОВ

Санкт-Петербургский институт информатики и автоматизации РАН

## МЕТОД РАСПОЗНАВАНИЯ АРАБСКО-ИНДИЙСКИХ ЦИФР

Представлена система распознавания арабско-индийских цифр с использованием метода классификации и комбинирования векторов признаков. Описаны методы классификации цифр, основанные на различных способах выделения характерных признаков.

**Введение.** Арабский язык — это рукописный язык с направлением письма справа налево, он используется как письменный в странах арабского мира и в некоторых языках ряда исламских государств. Арабско-индийские (арабские) цифры используются в настоящее время в большинстве восточных арабских стран (Иордании, Сирии, Саудовской Аравии и др.). На рис. 1, а представлены примеры печатных и рукописных арабско-индийских цифр от нуля до девяти, а на рис. 1, б — примеры иорданских автомобильных номерных знаков, для распознавания которых были разработаны представленные методы.

а)

0	٠							
1	١	١	١	١	١	١	١	١
2	٢	٢	٢	٢	٢	٢	٢	٢
3	٣	٣	٣	٣	٣	٣	٣	٣
4	٤	٤	٤	٤	٤	٤	٤	٤
5	٥	٥	٥	٥	٥	٥	٥	٥
6	٦	٦	٦	٦	٦	٦	٦	٦
7	٧	٧	٧	٧	٧	٧	٧	٧
8	٨	٨	٨	٨	٨	٨	٨	٨
9	٩	٩	٩	٩	٩	٩	٩	٩

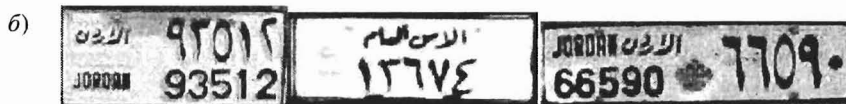


Рис. 1

Для решения задачи распознавания арабских символов было предложено много способов, например, сравнение с шаблонами, статистические и структурные методы, методы, использующие нейронные сети (НС) и др. Так, в работе [1] подробно рассмотрены этапы разработки систем распознавания символов, а в [2] — описаны методы выделения характерных признаков символов.

В [3] предложена система распознавания арабских цифр с применением вероятностных нейронных сетей. Система, использующая метод выделения вектора признаков символов, основана на вычислении моментов скелета цифры. Точность распознавания, достигнутая в этой системе, —  $\approx 99,72\%$  при испытаниях на множестве тестовых примеров различного начертания (наборе из 1080 символов).

В работе [4] описана система распознавания арабских цифр, в которой для выделения характерных признаков символов применялось вейвлет-преобразование. Точность распознавания для 1600 образцов символов составила 91,33 %. Система способна работать лишь с 8 цифрами.

В [5] была предложена схема сегментации изображения, основанная на 12-сегментном шаблоне, для выделения вектора признаков использовался теневой метод кодирования. Применялись две НС: для определения местонахождения контрольной точки шаблона и для распознавания. Была достигнута точность распознавания 97,60 % при испытаниях на 25 000 образцов.

В настоящей работе исследуется эффективность различных методов выделения векторов признаков для распознавания арабских цифр, а также рассматривается влияние на точность распознавания комбинирования векторов признаков и применения составных классификаторов.

**Система распознавания.** Системы распознавания символов обычно строятся по единой схеме: преобразование изображения символа в цифровую форму; предварительная обработка, выделение отдельных символов или областей, относящихся к символу; выделение характерных признаков; классификация; оценка результата.

*Предварительная обработка.* Эта стадия включает в себя нормализацию и приведение изображения к стандартному размеру, сглаживание и фильтрацию, коррекцию наклона символов. Приведение всех изображений, находящихся в базе данных, к стандартному размеру производится следующим образом: определяется (подбирается) прямоугольник минимальных размеров, ограничивающий цифру, полученный прямоугольник масштабируется до размеров 60×40 пикселей. На следующем этапе для повышения качества изображения (путем удаления шумов, улучшения границ и контуров изображения) применяется круговой сглаживающий фильтр радиуса 3.

Одной из существенных особенностей рукописных цифр является их наклон. Эксперименты показывают, что без коррекции наклона векторы признаков образцов внутри одного класса могут сильно варьироваться. Чтобы минимизировать внутриклассовую изменчивость признаков, необходимо фиксировать наклон символов. Коррекция наклона символов цифр играет важную роль в некоторых методах выделения характерных признаков символов. Существуют методы, например метод геометрических моментов, инвариантные к наклону.

Угол наклона арабских цифр задается самым длинным штрихом в цифре или углом между двумя такими штрихами. Из рис. 1, а видно, что цифры 1, 2, 3, 6, 9 содержат один вертикальный основной штрих, а 7 и 8 — два основных штриха, образующих острый угол.

Алгоритмы для коррекции наклона символов из других языков не могут быть напрямую применены к арабским цифрам, поэтому нами был разработан алгоритм, учитывающий особенности формы арабских цифр.

*Выделение признаков.* Этот этап является ядром любой системы распознавания, и для достижения высокой точности он должен быть спроектирован особенно тщательно. Значения признаков, вычисленные на основе входного изображения цифр, формируют векторы признаков. Для получения наилучшей точности распознавания вектор признаков должен быть различным для разных классов и иметь небольшую вариацию в пределах одного класса.

В работе [2] были рассмотрены разнообразные методы выделения характерных признаков: структурные, морфологические, преобразования Габора, Радона и Фурье и др. Каждый из этих методов в отдельности имеет свои преимущества, недостатки, особенности и свой уровень ошибок классификации. В работе были реализованы шесть методов для исследования точности распознавания каждого из них, а также при комбинировании векторов признаков и составном (комплексном) классификаторе.

*Классификация и распознавание.* Многослойный перцептрон (multilayer perceptron) представляет собой нейронную сеть прямого распространения, которая хорошо подходит для

задач классификации. В настоящей работе для обучения НС применяется градиентный метод поиска весов синапсов, минимизирующий среднеквадратичную ошибку.

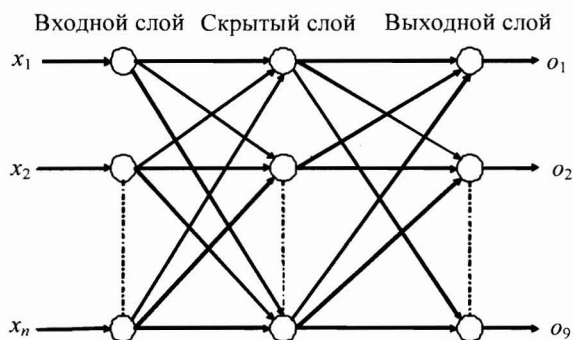


Рис. 2

Общая архитектура НС представлена на рис. 2, она включает входной нейронный слой, нескольких скрытых и выходной слой нейронов. Входной слой содержит количество нейронов, равное размерности вектора признаков ( $x_1—x_n$ ). Выходной слой содержит 9 нейронов, соответствующих классам распознаваемых цифр ( $o_1—o_9$ ). В скрытом слое содержится число нейронов, равное сумме нейронов во входном и выходном слоях. Шесть нейросетевых классификаторов, разработанных для оценки каждого из методов выделения признаков,

имеют одинаковую архитектуру и различаются лишь количеством нейронов во входном и скрытом слоях (см. таблицу).

Метод выделения признаков	Вид вектора признаков	Размерность слоя нейронов		Процент распознавания	Ранг
		входной слой	скрытый слой		
Средние расстояния зонирования	$\mathbf{X} = [x_1, x_2, \dots, x_{24}]^T$	24	35	<b>92,21</b>	<b>1</b>
Метод среднего значения зонирования	$\mathbf{X} = [x_1, x_2, \dots, x_{24}]^T$	24	35	94,50	3
Гистограммы проекции	$\mathbf{X} = [x_1, x_2, \dots, x_{30}]^T$	30	40	95,90	5
Дискретное преобразование Радона	$\mathbf{X} = [x_1, x_2, \dots, x_{40}]^T$	40	50	95,33	4
Кольцевое зонирование	$\mathbf{X} = [x_1, x_2, \dots, x_{30}]^T$	30	40	94,55	2
Геометрические моменты	$\mathbf{X} = [x_1, x_2, \dots, x_{18}]^T$	18	30	<b>96,88</b>	<b>6</b>

Аналогичная совокупность нейросетевых классификаторов была разработана для исследования эффективности комбинирования векторов признаков. Использование составных классификаторов (показывающих высокую точность распознавания во многих приложениях по распознаванию образов, включая символы) является основным направлением построения систем распознавания образов [6].

Существует множество подходов и соответствующих им архитектур для построения составных классификаторов [6]: простые методы, основанные на мажоритарном принципе (с учетом „ранга голоса“) [7], а также более сложные методы, как например, нечеткие модели и НС.

Процедура принятия решения в таких классификаторах имеет следующий вид. Пусть имеется  $N$  независимых классификаторов, каждый выдает значение того класса ( $1 \dots m$ ), которому принадлежит образец. Тогда составной классификатор выбирает подходящее выходное значение из множества выходов независимых классификаторов.

На рис. 3 представлена структура составного классификатора. Процедура „ранжирования голосов“, заключается в следующем. Для  $N$  независимых классификаторов ищутся совпадения выходных значений между  $(N + 1)/2$  классификаторами, если условие выполнено, то это значение становится выходным для составного классификатора, в противном случае ищутся совпадения для  $N/2$  классификаторов; если совпадение найдено, то выходным значением будет результат группы классификаторов, имеющих более высокий ранг. Следовательно, когда  $N \geq 4$ , существует возможность того, что более чем одна группа классификато-

ров сможет выдать различные результаты. Алгоритм продолжается рекурсивно, до тех пор пока остаются разногласия. В таком случае выходным значением составного классификатора станет результат классификатора с более высоким рангом.

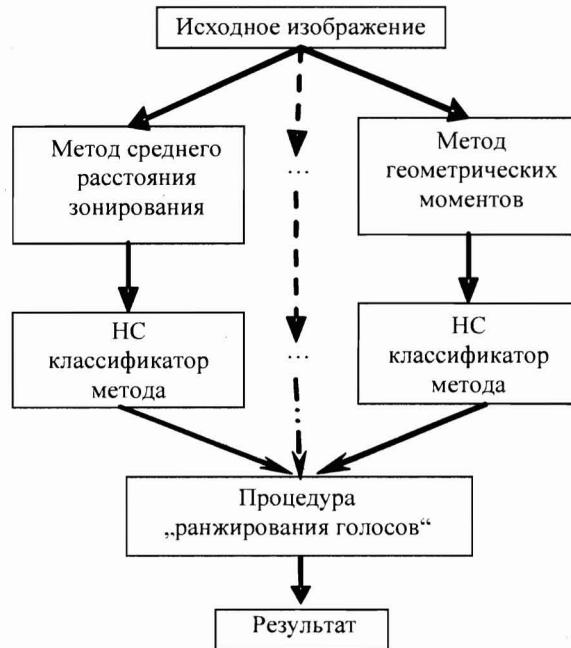


Рис. 3

**Оценка полученных результатов.** С целью проверки и оценки методов выделения признаков был проведен ряд испытаний. База данных образцов цифр содержала 4176 символов, которые использовались для обучения и тестирования системы [8]. Заметим, что в процессе обучения НС использовалось 55 % (2250 символов) содержимого базы данных, остальные 45 % (1926 символов) применялись для тестирования.

В первом эксперименте каждый вектор признаков подавался на вход НС. Из таблицы видно, что метод геометрических моментов дал наилучшую точность распознавания (96,88 %), а метод среднего расстояния зонирования имел худший результат (92,21 %).

В результате испытания независимый классификатор, показавший лучший результат распознавания, получил 6-й „ранг“, а худший — 1-й „ранг“.

Во втором эксперименте все комбинации векторов признаков подавались на вход соответствующего классификатора. В случае комбинирования векторов наблюдалось повышение точности распознавания. Так, например, комбинирование векторов признаков метода гистограмм проекций и геометрических моментов дало точность распознавания 99,27 %, а при комбинировании векторов методов геометрических моментов и среднего значения зонирования точность распознавания составила 99,53 %.

В третьем эксперименте исследовался составной классификатор. Результаты говорят об увеличении точности распознавания при использовании предложенного составного метода классификации. В случае комбинирования трех классификаторов лучший результат (99,43 %) был получен для методов гистограмм проекций, дискретного преобразования Радона и метода геометрических моментов. Как и ожидалось, при комбинировании четырех классификаторов дальнейшего повышения точности не происходило [8].

Также проводилось сравнение средней точности распознавания для составного классификатора, содержащего три классификатора, с методом комбинирования векторов признаков (соответственно 97,66 и 98,32 %).

**Заключение.** В работе исследовались и сравнивались шесть методов извлечения векторов признаков цифр. Были разработаны и испытаны различные классификаторы, основанные на этих методах. Индивидуальный классификатор, основанный на методе геометрических моментов, показал наилучшую точность распознавания.

Метод комбинирования векторов признаков и схема составного классификатора проявили свою эффективность в задаче распознавания арабско-индийских цифр.

Эксперименты показывают важность метода комбинирования векторов признаков. Даже если эффективность независимых классификаторов низка, она может увеличиться в комбинации с другими векторами признаков, например, при комбинировании векторов признаков метода среднего расстояния зонирования (92,21 %) и дискретного преобразования Радона (95,33 %) точность распознавания составила (98,39 %).

Результаты экспериментов показывают, что метод комбинирования векторов признаков дает намного лучший результат, чем независимые классификаторы, и незначительно лучший, чем составные классификаторы. Основное препятствие для применения такого подхода состоит в увеличении размерности векторов признаков, ввиду чего обосновано применение методов редукции факторного пространства признаков.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Plamondon R., Srihari S. N., Polytech E., Montreal Q.* Online and offline handwriting recognition: a comprehensive survey // *IEEE Transact. Pattern Analysis and Machine Intelligence*. 2000. Vol. 22. P. 63—84.
2. *Trier O. D., Jain A. K., Taxt T.* Feature extraction methods for character recognition a survey // *Pattern Recognition*. 1996. Vol. 29. P. 641—662.
3. *Al-Omari F. A., Al-Jarrah O.* Handwritten Indian numerals recognition system using probabilistic neural networks // *Advanced Engineering Informatics*. 2004. Vol. 18. P. 9—16.
4. *Mowlaei A., Faez K., Haghghat A. T.* Feature extraction with wavelet transform for recognition of isolated handwritten Farsi/Arabic characters and numerals // *14<sup>th</sup> Int. Conf. Digital Signal Proc.* 2002. Vol. 2. P. 923—926.
5. *Harifi A., Aghagolzadeh A.* A New Pattern for Handwritten Persian/Arabic Digit Recognition // *Int. J. of Information Technology*. 2004. Vol. 1. P. 293—296.
6. *Kittler J., Hatef M., Duin R. P. W., Matas J.* On combining classifiers // *IEEE Transact. Pattern Analysis and Machine Intelligence*. 1998. Vol. 20. P. 226—239.
7. *Lam L., Suen S. Y.* Application of majority voting to pattern recognition: an analysis of its behavior and performance // *IEEE Transact. Systems, Man and Cybernetics. Part A*. 1997. Vol. 27. P. 553—568.
8. *Kharm N., Ahmed M., Ward R.* A new comprehensive database of handwritten Arabic words, numbers, and signatures used for OCR testing // *IEEE Canadian Conf. Electrical and Computer Engineering*. 1999. P. 766—768.

Рекомендована институтом

Поступила в редакцию  
22.05.07 г.