

УДК 621.865.8

МЕТОДЫ БЫСТРОГО РАСПОЗНАВАНИЯ СИМВОЛОВ, ПРИГОДНЫЕ ДЛЯ АППАРАТНОЙ РЕАЛИЗАЦИИ

И. Л. Ерош,

доктор техн. наук, профессор

М. Б. Сергеев,

доктор техн. наук, профессор

Н. В. Соловьев,

старший преподаватель

Санкт-Петербургский государственный университет
аэрокосмического приборостроения

Рассматривается метод распознавания печатных символов, ориентированный на аппаратную реализацию в микропроцессорных системах учета и сортировки документов по их идентификационным номерам. Устанавливаются ограничения на число распознаваемых символов и виды шрифта. Формулируются требования к скорости и точности распознавания.

The method of recognition of printed symbols for microprocessor systems of the account and sorting of documents under their identification numbers is considered. Restrictions of symbols number and kinds of a font are defined. Requirements for speed and accuracy of recognition are formulated

Введение

Программные системы для персональных компьютеров (ПК), распознающие печатные символы по изображениям, известны уже давно. Примером может служить Fine Reader, позволяющий распознавать печатный текст с использованием ПК с достаточными скоростью и степенью точности. Он представляет собой сложный процесс, основанный на выявлении топологических признаков отдельных символов и построении семантических деревьев [1], требует большого объема памяти, высокого разрешения сканирующей линейки (не менее 300 dpi) для получения удовлетворительных (около 1 % ошибок) результатов распознавания и большого времени работы ПК. Попытка реализовать аппаратно указанный алгоритм и подобные ему в реальном масштабе времени, т. е. в темпе поступления сканируемого изображения, сталкивается с серьезными трудностями.

Следует отметить, что подобная задача вполне успешно была решена еще в 70-е годы прошлого века при автоматизации сортировки почтовых конвертов по цифровому коду почтового отделения адресата. Основные отличия данной системы от рассматриваемой далее заключаются в том, что, во-

первых, цифры кода на конверте заполнялись от руки по трафарету и имели стилизованный вид, специально приспособленный для автоматического распознавания, и, во-вторых, код на конверте состоял только из цифр, в то время как документ может иметь в номере как цифры, так и буквы.

Постановка задачи

Каждый документ, подлежащий учету, имеет цифровой или цифробуквенный идентификационный номер, нанесенный в фиксированном месте документа печатным образом известным шрифтом.

Документы поступают на устройство сканирования с заранее известной ориентацией со скоростью не менее 20 м/с. По различным причинам документы могут поступать на устройство считывания со смещением, достигающим 10 % от ширины документа, и перекосом в пределах $\pm 5^\circ$.

Требуется в темпе сканирования документа распознавать символы идентификационного номера документа с вероятностью ошибки не более 0,01 % и отказом от распознавания не более 0,1 %.

В качестве документа, подлежащего учету по распознаваемым цифробуквенным номерам, будем



■ Рис. 1. Исходное изображение купюры

рассматривать денежные купюры. На рис. 1 приведено изображение десятирублевой купюры с наличием как яркостных, так и пространственных искажений.

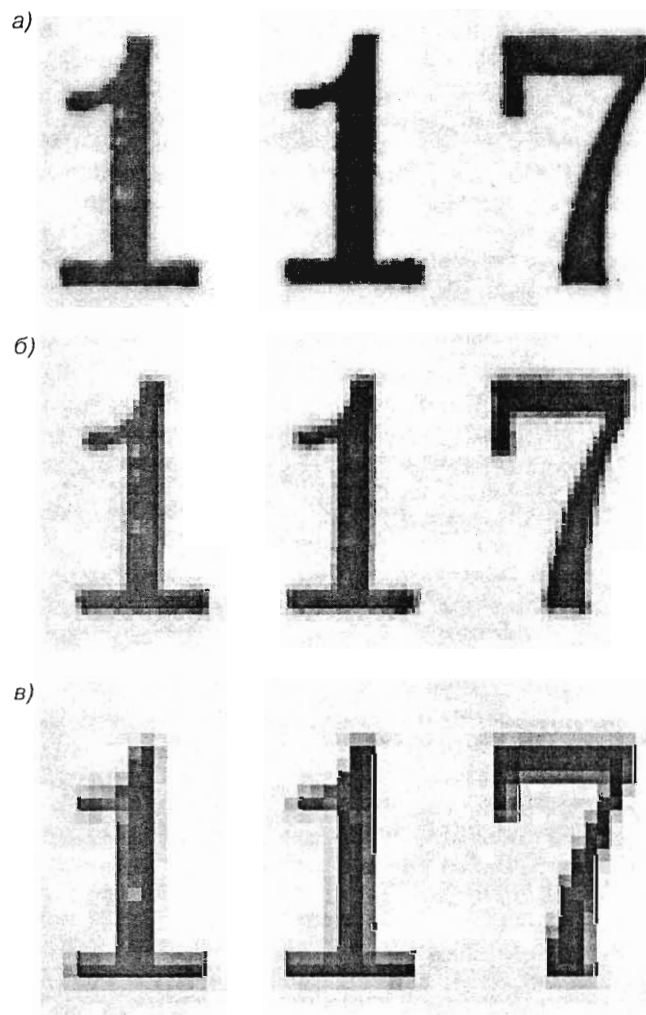
Процедура распознавания символов

Анализ возможного быстродействия устройства сканирования купюр при заданной скорости их поступления показал, что размер распознаваемого символа не может превышать 25x11 пикселей, что несколько меньше 150 dpi. Естественно, чем выше разрешающая способность устройства получения изображения, тем больше (в пикселях) размер распознаваемых символов, что позволяет применить более широкий спектр методов как предварительной обработки, так и собственно распознавания. На рис. 2, а-в показаны фрагменты номеров купюр, полученных при разрешении 600, 300 и 150 dpi соответственно.

Даже визуальное сравнение изображений с различным разрешением позволяет сказать, что, например, методы, основанные на выделении контурной линии, могут быть вполне успешно применены при разрешении 600 dpi, но вряд ли дадут положительный эффект на изображениях с разрешением 150 dpi. Однако увеличение разрешения изображения ограничено быстродействием устройства ввода, скоростью поступления купюр и требованием распознавания номеров в реальном масштабе времени.

Как известно [2], процедуру распознавания образов принято разделять на два этапа. На первом этапе в диалоговом режиме происходит обучение системы распознавания по репрезентативной выборке образов, относящихся к распознаваемым классам, на втором этапе обученная система в автоматическом режиме производит собственно распознавание предъявленного ей образа, т. е. отношение его к одному из известных ей классов или отказ от распознавания. Последнее тоже можно считать положительным результатом распознавания в отличие от ошибочной классификации распознаваемого образа.

Существует большое количество разнообразных методов классификации, причем большинство из



■ Рис. 2. Фрагменты номеров при разном разрешении: а — 600 dpi; б — 300 dpi; в — 150 dpi

них можно представить как поиск минимального расстояния в пространстве признаков между распознаваемым образом и известными классами. В рассматриваемом случае основная проблема связана с выбором метрики пространства признаков и собственно признаков, которые должны, с одной стороны, возможно дальше разнести классы в признаковом пространстве и иметь минимально возможный разброс значений, а с другой стороны, достаточно быстро вычисляться.

Система распознавания объектов по их изображениям, в том числе и символов идентификационного номера купюры (документа), предполагает наличие следующих операций:

- получение монохромного полутонового изображения в цифровой форме;
- подавление яркостных помех и компенсация пространственных искажений;
- бинаризация изображения, т. е. приведение его к черно-белому виду;
- сегментация, т. е. выделение на изображении номера отдельных символов;

— вычисление значений признаков и распознавание по ним каждого символа.

Для каждой из перечисленных операций известно множество алгоритмов [3], из которых необходимо отобрать подходящие по скорости выполнения. Следует отметить, что и при таком ограничении число возможных алгоритмов достаточно велико и дальнейший отбор следует проводить по результатам экспериментов с тестовыми изображениями.

Выбор методов классификации

При отборе методов классификации образов в рассматриваемом случае распознавания символов основными критериями наряду с величиной ошибки распознавания, несомненно, являются скорость выполнения и возможность аппаратной реализации. Подобным ограничениям в первую очередь удовлетворяют различные модификации метода «маски». Действительно, аппаратная реализация метода «маски» имеет достаточно простое решение, а при отсутствии яркостных и пространственных помех метод дает исключительно высокую надежность распознавания.

Анализ изображений символов в идентификационных номерах купюр не позволяет говорить об отсутствии помех, т. е. различные изображения одного и того же символа существенно отличаются друг от друга по яркости и расположению относительно центра изображения, причем *даже на одной купюре*. Причины этих отличий следующие:

- наличие систематических и случайных погрешностей аппаратуры сканирования;
- изменение освещенности сканируемого документа;
- возможность смещения документа относительно эталонного положения при сканировании;
- неточность при печати символа и непостоянство расстояния между символами в номере;
- изменчивость отражательной способности бумажной основы документа.

Полностью компенсировать имеющиеся помехи не представляется возможным, поэтому наиболее перспективным представляется выбрать в качестве признаков те, которые в большей степени инвариантны к неустраняемым яркостным и пространственным искажениям.

К последним можно отнести:

- метод гарантированной маски, при котором с пикселями бинарного изображения распознаваемого символа сравниваются на предмет их совпадения только те пиксели эталона, значения которых оставались постоянными на этапе обучения у всех изображений данного символа из выборки;
- метод статистической маски, при котором расстояние между распознаваемым образом и эталоном в пространстве признаков определяется по всем пикселям изображения с учетом для каждого пикселя распознаваемого символа вероятности появления данного значения у соответствующего пикселя сравниваемого эталона;

— метод построчной топологии, при котором пространство признаков определяется числом строк изображения, причем для каждой строки вычисляется топологический код, описывающий размер и расположение по строке участков символа;

— метод признакового пространства с Евклидовой метрикой, при котором признаком является число пикселей символа по строкам и столбцам.

После окончания этапа обучения, на котором определяются статистические характеристики классов распознаваемых образов для выбранных признаков пространств, желательно провести отбор наиболее информативных признаков в каждом пространстве по известным критериям, например, критерию Фишера [2]:

$$d_{ij}^n = \frac{(m_i - m_j)^2}{D_i + D_j}, \quad (1)$$

где d_{ij}^n — значение критерия по n -му признаку для пары классов i и j ; m_i, D_i, m_j, D_j — математические ожидания и дисперсии по n -му признаку для классов i и j .

Критерий (1) позволяет выбрать наиболее информативный признак для пары классов с учетом статистических характеристик разброса значений каждого признака для сравниваемых классов.

Выбор методов предварительной обработки изображения

Анализ качества монохромных изображений символов номеров купюр, получаемых в цифровой форме, показал, что коррекция яркостных помех непосредственно после получения изображения не приводит к существенному повышению надежности распознавания. Временные затраты на обработку монохромного изображения медианным или усредняющим фильтром практически не окупаются повышением надежности распознавания.

Пространственные искажения, вызванные смещением купюры относительно эталонного положения при сканировании, необходимо компенсировать, так как они не позволяют надежно выделить на изображении область, содержащую распознаваемые символы, и эффективно применять выбранные методы распознавания. Известно [4], что подобные искажения описываются аффинной группой преобразования, включающей в себя подгруппы поворота и смещения:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad (2)$$

где (x', y') и (x, y) — соответственно координаты точек преобразованного и исходного изображений; α, b_1, b_2 — соответственно угол поворота и смещение по осям X и Y изображения относительно эталонного положения.

Особенностью группового преобразования является то, что все соответствующие точки исходного и преобразованного изображений связаны уравнением (2), что позволяет найти параметры норма-

лизации α , b_1 , b_2 , если известны координаты двух характерных точек на искаженном изображении и их координаты при эталонном положении купюры. В качестве характерных точек можно выбрать вершины купюры, поступающие на сканирование первыми, что позволит проводить вычисление параметров нормализации параллельно со сканированием остальной части купюры. Для определения координат вершин можно использовать известные алгоритмы нахождения вершин многоугольника [5].

После нормализации пространственных искажений необходимо произвести бинаризацию изображения. Известно большое количество различных методов выбора порога бинаризации [3], отличающихся по сложности алгоритма и, соответственно, по времени выполнения. Следует отметить, что выбор конкретного метода бинаризации зависит от статистических характеристик функции яркости как конкретного изображения, так и всей совокупности обучаемой выборки.

Если яркость пикселей символов и фона практически постоянна на всех изображениях выборки и существенно отличается друг от друга, то вполне удовлетворительные результаты можно получить установкой постоянного порога, что позволяет проводить бинаризацию практически одновременно со сканированием. Данный метод следует признать наиболее эффективным с точки зрения быстродействия.

Больших временных затрат требует порог, определяемый как $p = 0,5(f_{\min} + f_{\max})$, где p — порог бинаризации; f_{\min} , f_{\max} — соответственно минимальная и максимальная яркость изображения (поскольку значение p можно вычислить только после окончания сканирования всего изображения). Данный метод следует применять, если яркость пикселей символов и фона существенно отличаются на одном изображении, но сами значения яркости имеют значительный разброс на изображениях выборки.

Наибольших временных затрат требует метод, в котором порог бинаризации определяется для каждого пиксела изображения по гистограмме яркости некоторой окрестности данного пиксела. Этот метод при правильном подборе размера окрестности дает удовлетворительные результаты при плавном изменении яркости пикселей фона на изображении, причем яркость пикселей символов в одной области изображения может совпадать с яркостью пикселей фона в другой области этого изображения.

Для подавления случайных помех на бинаризованном изображении желательно провести логическую фильтрацию одиночных пикселей и сгладить контуры символов одним из известных методов [4]. Специфическая форма цифробуквенных символов, характеризующаяся в большинстве случаев высоким значением отношения длины контура к площади, требует обратить внимание на сохранение связности при проведении логической фильтрации.

Сегментация бинаризованного изображения (выделение на изображении фрагмента, содержа-

щего один символ) не вызывает в данном случае каких-либо затруднений. Регулярное расположение символов в строке и отсутствие перекрытия проекций символов на линию, параллельную строке, позволяет легко сегментировать символы обычным сканированием бинарного изображения с шириной, равной известной ширине символа.

После сегментации для каждого символа производится вычисление признаков и его классификация одновременно всеми изложенными выше методами с целью повышения вероятности правильного распознавания. Окончательное решение об отнесении символа к одному из известных классов производится методом конкурентного распознавания.

Реализация конкурентного распознавания символов

Приведенные выше методы классификации изображений символов в цифробуквенном номере купюры были выбраны на основании обработки результатов распознавания на большой выборке сканированных номеров реальных купюр различного достоинства.

Оценка сложности методов позволяет считать их пригодными для аппаратной реализации в программируемой логике, универсальных или DSP-процессорах. Однако отсутствие гарантии правильной классификации символов в цифробуквенном коде купюры методами статической маски, построчной топологии, а также специфика методов гарантированной маски и признакового пространства с Евклидовой метрикой не позволяют однозначно выбрать существенно лучший по вероятности правильного распознавания метод.

Возможным решением, незначительно усложняющим реализацию устройства учета и сортировки документов, является применение для распознавания одновременно нескольких методов, перекрывающих по своим возможностям весь спектр основных помех.

Данная реализация может быть представлена в виде совокупности конкурирующих процессоров [6, 7], начинающих одновременно реализовывать каждый свой алгоритм распознавания изображения, полученного со сканера.

Конкурентная классификация изображений отдельных символов завершается решением арбитра по результатам классификации каждым из используемых методов с учетом весомости статистических характеристик, полученных для них в процессе обучения системы распознавания на тестовой выборке.

Заключение

Анализ результатов применения перечисленных методов к тестовым изображениям цифровых символов показал, что ни один из них не дает 100%-ную гарантию распознавания всех цифр. В большинстве случаев наилучшие результаты дают методы

статистической маски и метод построчной топологии, но примерно для 5 % изображений символов более эффективными оказываются методы гарантированной маски и признакового пространства.

Экспериментального исследования требует и вопрос определения порога достоверности, т. е. максимально допустимого расстояния между распознаваемым образом и известными классами в пространстве признаков, при котором образ можно отнести к одному из имеющихся классов. Если

для какого-либо распознаваемого образа расстояние до ближайшего класса превышает порог достоверности, то данный образ следует отнести к нераспознанному.

Эксперименты на реальных изображениях денежных купюр показали приемлемые результаты распознавания цифровых символов конкурентным методом при правильном подборе объема обучающей выборки, порога бинаризации и порога достоверности.

Литература

1. Харченко Н. В., Чезганцов Я. В. Современные технологии оптического распознавания текста в системе ABBY FINEREADER // Экология, мониторинг и рациональное природопользование: Научные труды. Вып. 318. Российская академия естественных наук. — М., МГУЛ, 2002. — С. 162–172.
2. Ту Дж., Гонсалес Р. Принципы распознавания образов. — М.: Мир, 1978. — 411 с.
3. Катус Г. П., Катус П. Г. Системы распознавания визуальной информации // Приборы и системы. Управление, контроль, диагностика. — 2001. — № 11. — С. 54–60.
4. Ерш И. Л., Игнатьев М. Б., Москалев Э. С. Адаптивные робототехнические системы: Методы анализа и системы обработки изображений. — Л.: ЛИАП, 1985. — 144 с.
5. Техническое зрение роботов / Под ред. А. П. Ю; Пер. с англ. Д. Ф. Миронова. — М.: Машиностроение, 1987. — 320 с.
6. Байков В. Д., Сергеев М. Б. Конкурентные вычислительные системы // Тез. докл. X Всес. симпозиума по проблемам избыточности в информационных системах. — Л.: ЛИАП, 1989. — С. 96–97.
7. Байков В. Д., Сергеев М. Б., Мохаммед М. А. Архитектура вычислительной системы для решения систем линейных алгебраических уравнений // Управляющие системы и машины. — 1990. — № 4. — С. 33–35.

Ю. П. Иванов, В. Г. Никитин, В. Ю. Чернов

Контроль и диагностика измерительно-вычислительных комплексов: Учеб. пособие / СПб.: СПбГУАП, 2004. — 98 с.: ил. ISBN 5-8088-0114-1

Изложены основные понятия, задачи, методы и способы контроля и диагностики технического состояния измерительно-вычислительных комплексов летательных аппаратов. Описаны обобщенная структура и основные характеристики систем контроля, основные показатели достоверности контроля и ее составляющие, а также выбор допусков на параметры контроля и методы принятия решений в процессе контроля. Рассмотрены методы, аппаратные и программные средства цифровых измерительно-вычислительных комплексов с помощью систем встроенного контроля и диагностики.

Предназначено для студентов, обучающихся по специальностям 190300 «Авиационные приборы и измерительно-вычислительные комплексы», 131000 «Техническая эксплуатация авиационных электросистем и пилотажно-навигационных комплексов», а также магистров по направлению 5515 «Приборостроение» и бакалавров по направлению 5520 «Эксплуатация авиационной и космической техники».

Ю. П. Иванов
В. Г. Никитин
В. Ю. Чернов

КОНТРОЛЬ И ДИАГНОСТИКА
ИЗМЕРИТЕЛЬНО-ВЫЧИСЛИТЕЛЬНЫХ
КОМПЛЕКСОВ

