

Мучник М.М., Прохоров В.Г.

ОБ ОДНОМ МЕТОДЕ ПРЕДСТАВЛЕНИЯ ГРАФИЧЕСКИХ ОБРАЗОВ ТЕКСТОВЫХ СИМВОЛОВ В ВИДЕ ХАРАКТЕРИСТИЧЕСКИХ ВЕКТОРОВ

Широкое внедрение компьютеров во все сферы народного хозяйства приводит к необходимости перехода от бумажных к электронным носителям информации. Для решения этой задачи необходимо использовать системы автоматического распознавания текстов.

***Ключевые слова:** система автоматического распознавания фрактальное сжатие изображений, эталоны характеристических векторов коэффициентов.*

Введение

Одной из актуальных проблем современных информационных технологий является разработка методов создания систем автоматического распознавания печатных и/или «рукопечатных» бумажных документов и формирования их электронных версий. В настоящее время в таких системах наиболее часто используются следующие методы:

- распознавание с помощью нейронных сетей;
- распознавание с помощью шаблонов.

Метод нейронных сетей основан на принципах искусственного интеллекта, требует значительных вычислительных ресурсов и трудно предсказуем (заранее сложно оценить сложность сети).

Метод шаблонов сложен для использования для решения данной задачи в силу своей негибкости (привязанности к шрифту), однако с определенными модификациями используется достаточно часто.

Эти методы требуют много времени на распознавание и не обеспечивают однозначного распознавания текстов, хотя, при соответствующей настройке, позволяют получить практически приемлемые результаты. Параллельно развивались методы сжатия и восстановления изображений [1], в частности, представления сжатых изображений в виде характеристических векторов [2,3].

Можно выдвинуть гипотезу о том, что, если после сжатия изображение может быть восстановлено по характеристическому вектору (т.е. значения элементов характеристического вектора сохраняют существенные особенности графического образа), то можно говорить о том, что по полученному в результате сжатия графического образа характеристическому вектору некоторого символа (печатного или «рукопечатного») можно определить, какой именно это был символ, т.е. распознать символ.

Настоящая статья посвящена описанию разработанных авторами методов автоматического распознавания печатных и рукопечатных символов при использовании предложенного подхода для построения системы автоматического распознавания символов «Cunning Eye».

1. Постановка задачи распознавания при использовании метода представления в виде характеристических векторов

Суть распознавания заключается в анализе символа, расчёте его характеристик и установлении соответствия между обрабатываемым символом и одним из эталонных (с заранее вычисленными характеристиками) за счёт сравнения характеристик. Следует отметить, что речь не идёт об однозначном сопоставлении символов, а лишь о нахождении эталона, характеристики которого наиболее точно соответствуют характеристикам данного символа (сто-процентное совпадение характеристик невозможно в связи с различием шрифтов и почерков, а также некоторыми другими факторами).

Рассмотрим математическую модель алгоритма распознавания. Символ представляется в виде вектора, состоящего из характеристик этого символа.

$X = (x_1, x_2, \dots, x_n)$ Здесь, x_1 – первая характеристика вектора X , x_2 – вторая характеристика вектора X , и т.д.

Рассмотрим задачу сравнения двух символов, представленных в форме векторов $X = (x_1, x_2, \dots, x_n)$ и $Y = (y_1, y_2, \dots, y_n)$. В качестве критерия близости используем евклидово расстояние.

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1.1)$$

С учетом того, что одни характеристики при сравнении имеют большее значение, нежели другие, формула расстояния приобретает вид

$$D(x, y) = \sqrt{\sum_{i=1}^n w_i (x_i - y_i)^2}, \quad (1.2)$$

где w_i - весовой коэффициент важности i -той характеристики. Поскольку при распознавании нас интересует вычисление минимального расстояния между векторами, большее значение весового коэффициента w_i соответствует более важному признаку и наоборот.

В данной статье будут рассмотрены некоторые характеристики, используемые для распознавания символов.

2. Коэффициенты фрактального сжатия.

Метод фрактального сжатия изображений основан на представлении изображения в более компактной форме – с помощью коэффициентов системы итерируемых функций (Iterated Function System – далее по тексту как IFS). Строго говоря, IFS представляет собой набор трехмерных аффинных преобразований, переводящих одно изображение в другое. Преобразованию подвергаются точки в трехмерном пространстве (x - координата, y - координата, яркость).

Каждая итерация алгоритма фрактального сжатия приводит к тому, что по исходному изображению строится новое изображение, которое и берется в качестве исходного для следующей итерации. В процессе итераций мы получим изображение, которое не будет изменяться при очередной итерации. Оно будет зависеть только от коэффициентов, используемых при данном преобразовании, и не будет зависеть от исходной картинке. Это изображение называется “неподвижной точкой” или *аттрактором* данной IFS. Соответствующая теория гарантирует наличие ровно одной неподвижной точки для каждой IFS. Этот факт и позволяет однозначно поставить в соответствие каждому символу набор фрактальных коэффициентов.

В контексте данной задачи при сжатии графический образ символа, разбивается на 4 квадранта, а затем к каждому из них применяется алгоритм фрактального сжатия. Найденные при этом коэффициенты IFS и являются искомыми характеристиками символа.

Применение фрактальных коэффициентов в качестве характеристик символа является новшеством системы “Cunning Eye”: в других OCR и ICR системах они не применяются. В контексте формулы (1.2), коэффициентам фрактального сжатия назначается наибольший весовой коэффициент.

3. Инвариантные моменты распределения точек.

Моменты различных порядков используются в таких задачах обработки изображений, как зрение роботов, обнаружение и распознавание летательных аппаратов и судов по снимкам,

анализ сцен и распознавание символов. В последнем случае в качестве признаков используют значения статистических моментов совокупности "черных" точек относительно некоторого выбранного центра.

Наиболее общеупотребительными в приложениях такого рода являются построчные, центральные и нормированные моменты.

Для цифрового изображения, хранящегося в двумерном массиве, построчные моменты являются функциями координат каждой точки изображения следующего вида:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y), \quad (3.1)$$

где $p, q=0, 1, \dots, \infty$; M и N являются размерами изображения по горизонтали и вертикали и $f(x, y)$ является яркостью пикселя в точке (x, y) на изображении.

Центральные моменты являются функцией расстояния точки от центра тяжести символа:

$$\mu = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y), \quad (3.2)$$

где (\bar{x}, \bar{y}) - координаты центра тяжести.

Наконец, нормированные центральные моменты получаются в результате деления центральных моментов на моменты нулевого порядка.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \quad \text{где } \gamma = \frac{p+q}{2} \quad (3.3)$$

Отдельный вид моментов является инвариантным по отношению к операциям масштабирования, параллельного переноса, поворота. Такие моменты называют *инвариантными*. Они вычисляются через нормированные центральные моменты второго и третьего порядков. Использование инвариантных моментов позволяет распознавать слегка повернутые символы, а также символы разных размеров. Инвариантные моменты вычисляются по формулам

$$\phi_1 = \eta_{20} + \eta_{02} \quad (3.4)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (3.5)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (3.6)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (3.7)$$

и т.д. Всего в расчетах используется семь инвариантных моментов, значения которых затем логарифмируются.

На практике применение инвариантных моментов было сопряжено с небольшими различиями значений (которых по идее не должно было быть) при небольших углах поворота. Данный эффект связан с дискретностью рассматриваемого изображения.

4. Центр тяжести рисунка

Вычисления центра тяжести рисунка является одной из базовых операций применяемых при анализе статистического распределения точек рисунка.

Координаты (\bar{x}, \bar{y}) центра тяжести рисунка определяется по формуле:

$$\bar{x} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} xf(x, y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)}, \quad \bar{y} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} yf(x, y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)}, \quad (4.1)$$

где M и N являются размерами изображения по горизонтали и вертикали, $f(x, y)$ является яркостью пикселя в точке (x, y) на изображении.

С практической точки зрения, вычисление центра тяжести является «улучшенным вариантом» вычисления средней яркости изображения, которая подробно описана в [5]: центр тяжести характеризует распределение черных точек по всему рисунку.

5. Структурные характеристики символа

Структурные признаки обычно используются для выделения общей структуры образа. Они описывают геометрические и топологические свойства символа.

Одними из наиболее используемых признаков являются штрихи и пробелы, применяемые для определения следующих характерных особенностей изображения: концевых точек, пересечения отрезков, замкнутых циклов, а также их положения относительно рамки, объемлющей символ.

Пусть матрица, содержащая скелетизированный символ, разделена на ряд областей, каждой из которых присвоены буквы A, B, C и т.д. Символ рассматривается как набор штрихов. При этом штрих является прямой линией (ℓ) или кривой (c), и соединяет некоторые две точки в начертании символа. Штрих является кривой, если его точки удовлетворяют следующему выражению:

$$ABS \left| \frac{\sum_{i=1}^n ax_i + by_i + c / \sqrt{a^2 + b^2}}{n} \right| > 0.69,$$

в противном случае - это прямая. В данной формуле (x_i, y_i) является точкой, принадлежащей штриху; $ax+by+c=0$ - уравнение прямой, проходящей через концы штриха, коэффициент 0.69 получен опытным путем.

При введенных обозначениях признаки символа могут быть записаны, например, в виде 'A ℓ C' и 'AcD', что означает наличие прямой, проходящей из области 'A' в область 'C', и кривой, проходящей из области 'A' в область 'D' соответственно. Структурные признаки, по сравнению с другими методами, обладают устойчивостью к сдвигу, масштабированию и повороту символа на небольшой угол, возможным дисторсиям и различным стилевым вариациям шрифтов [4-6]. Структурные характеристики выделяются на этапе скелетизации.

Недостатком структурных признаков является повышенное требование к алгоритмам скелетизации, а также ресурсоемкость. Поскольку данные характеристики невозможно записать в строгой аналитической форме, применение таких признаков при распознавании требует использования дополнительных алгоритмов, которые находят соответствия между эталонами символов и распознаваемыми символами. Одним из таких алгоритмов является метод маски.

6. Метод маски (пересечений)

На изображение накладывается маска, которая состоит из прямых, которые выходят из начала координат (верхний левый угол рисунка) под углами 0, 45, 90, 135 градусов. Затем подсчитывается количество пересечений скелета символа с прямыми маски. Этот метод инвариантен к дисторсии и небольшим стилистическим вариациям символов, а также обладает

достаточно высокой скоростью и не требует значительных вычислительных затрат.

В качестве начала координат используются все углы изображения с целью избегания возможных ошибок, вызванных, например, скелетизацией, или нестандартным написанием символов.

7. Обучение системы

Перед началом использования системы автоматического распознавания необходимо обеспечить формирование набора эталонов характеристических векторов для каждого символа (это задача разработчика системы). На Рис.1 в качестве примера, иллюстрирующего выделение характеристик символов, приведена распечатка формы обучающего модуля системы “Cunning Eye”

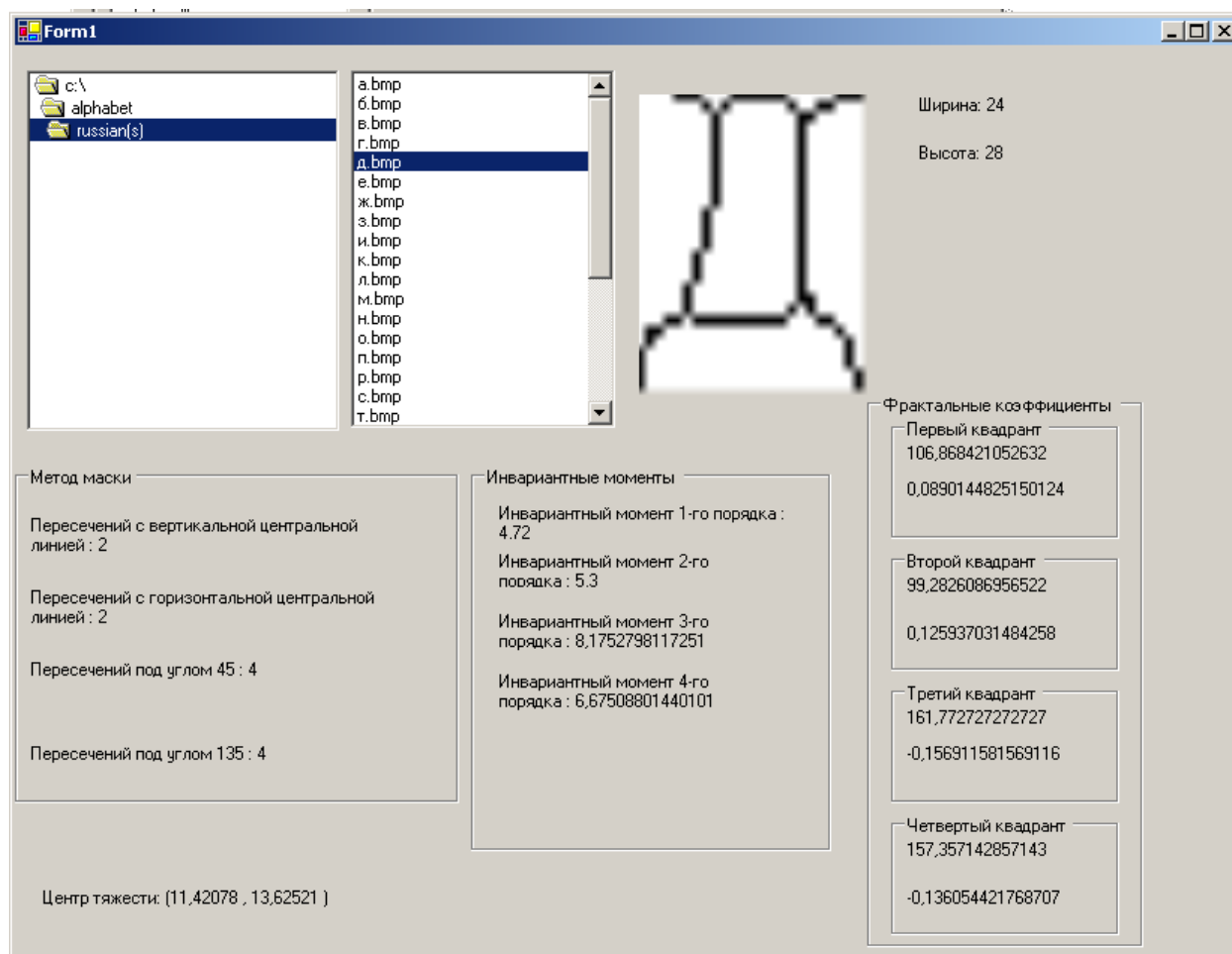


Рис.1. Формирования эталона характеристического вектора для символа

В процессе обучения, система вычисляет характеристики различных букв, формирует алфавит и записывает его в файл на жестком диске.

Собственно распознавание документа, представленного в виде графического образа, предлагается проводить по таким этапам:

- предпроцессорная обработка - образ обрабатывается программой фильтрации, убирающей шум и выделяющей отдельные символы;
- скелетизация – выделенные из текста символы скелетизируются;
- формирование векторов коэффициентов – для каждого символа формируется вектор коэффициентов;
- распознавание – сравнение полученного вектора с векторами символов-эталонов и принятие решения о типе символа.

Выводы

1. В статье описан новый подход к построению систем автоматического распознавания, основанный на использовании метода представления графических образов текстовых символов в виде характеристических векторов;
2. Разработана математическая модель процесса распознавания;
3. Подтверждена возможность представления текстовых символов в виде векторов с характеристическими коэффициентами и их последующего распознавания путём сравнения с эталонными векторами;
4. Проведенные исследования по выбору характеристик для определения вида символа показали, что ни одна из них в большинстве случаев не способна «в одиночку» установить однозначное соответствие между символом и его образом, т.е. для распознавания необходимо использовать совокупность взаимодополняющих характеристик;
5. Для апробации разработанных методов разработана система автоматического распознавания печатных и «рукопечатных» текстов “Cunning Eye”, обеспечивающая быстрое и качественное распознавание текстов.

ЛИТЕРАТУРА

1. Ватолин Д.С. Алгоритмы сжатия изображений – М.: МГУ, 1999.- 231с.
2. Fisher Y. “Fractal Image Compression: Theory and Application to Digital Images”, Springer Verlag, 1995- 452 p..
3. Уэлстид С. Фракталы и вейвлеты для сжатия изображений – М.: Триумф , 2003 – 320 с.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений - М.: Техносфера, 2005. - 1072 с
5. Форсайт Д., Понс Д. Компьютерное зрение. Современный подход.- Вильямс, Москва, 2004.- 928 с.
6. Бондаренко А.В. и др. Исследование подходов к построению систем автоматического считывания символьной информации - Изд-е ИПМ им. М.В. Келдыша РАН, Москва, 2003, с.32-43.

УДК 621.396

Богом'я В.І.

ОСОБЛИВОСТІ ІДЕНТИФІКАЦІЇ ГЕОСТАЦІОНАРНИХ КОСМІЧНИХ ОБ'ЄКТІВ

У даній статті розглядаються особливості ідентифікації геостаціонарних космічних об'єктів. Наведена методика ведення зонального каталогу геостаціонарних космічних об'єктів за умови відсутності даних про космічні апарати у міжнародних каталогах.

Ключові слова: *методика ведення зонального каталогу, геостаціонарний космічний об'єкт*

У даний час приділяється значна увага спостереженню за космічними апаратами, які знаходяться на геостаціонарній орбіті та уточненню даних про їх місцезнаходження [1-5].

Відомо, що на космічний апарат впливають сили гравітації, тому він постійно змінює своє місцезнаходження. Для підтримки космічного апарата у визначених границях застосовуються коректувальні двигуни, які встановлені на самому космічному апараті та за допомогою яких здійснюються маневри щодо стабілізації космічного апарата. У випадку закінчення ре-