На этапе распознавания фрагмент изображения, согласно принципу целостности, будет интерпретирован как некий объект (символ), только если на нём присутствуют ***все*** структурные части этого объекта, и эти части находятся в соответствующих отношениях. Поэтому ***ABBYY FineReader*** не пытается принимать решение, перебирая тысячи эталонов в поисках наиболее подходящего, вместо этого выдвигается ***ряд гипотез*** относительно того, на что похоже обнаруженное изображение, затем каждая гипотеза ***целенаправленно*** проверяется. Причём проверять, верна ли выдвинутая гипотеза, система будет, используя ***принцип адаптивности,*** опираясь на накопленные ранее сведения о возможных начертаниях символа в распознаваемом документе.

На начальном этапе OC7?-система производит ***бинаризацию,*** т. е преобразование цветного или полутонового образа в монохромный. Однако после типовой процедуры бинаризации любая текстура оставит большое количество «лишних» точек, расположенных вокруг символов и резко снижающих качество распознавания. После бинаризации объекта и создания его математической модели из имеющегося (ранее сформулированного) множества гипотез формируется список ***рациональных*** гипотез, каждой гипотезе приписывается определённый вес (степень уверенности).

Соответствующий каждой из гипотез набор графических объектов уровня «символ» поступает на вход механизма распознавания символов. Последний представляет собой комбинацию ряда элементарных распознавателей, называемых ***классификаторами.*** Отметим, что гипотезы в основном сформированы для символов. Наличие пар «гипотеза-символ» позволяет построить ***классификаторы,*** общий вид которых иллюстрирует рис. 1.

По окончании обработки классификатор порождает новый список гипотез относительно принадлежности очередного изображения к тому или иному классу Выходной список всегда ранжирован по весу (уверенности).

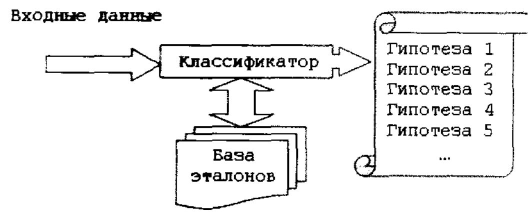


Рис. 1. Упрощенная схема работы классификатора1

В системах распознавания, построенных на технологиях ***ABBYY,*** применяются следующие типы классификаторов.

Растровый классификатор. Принцип действия основан на прямом сравнении изображения символа с эталоном.

Признаковый классификатор

Принцип действия: изображению ставится в соответствие У-мерный вектор признаков. Собственно классификация заключается в сравнении его с набором эталонных векторов той же размерности. Тип и количество признаков определяют качество распознавания. Формирование вектора (вычисление его координат в ***N*** -мерном пространстве) производится во время анализа предварительно подготовленного изображения. Данный процесс называют извлечением признаков. Эталон для каждого класса получают путём аналогичной обработки символов обучающей выборки. Сравнение каждой пары векторов заключается в вычислении оценки, характеризующей расстояние между точками в Л^-мерном пространстве (точка - геометрическое представление такого вектора). Основные достоинства признакового классификатора - простота реализации, хорошая обобщающая способность. Наиболее серьёзный его недостаток - ***неустойчивость к различным дефектам изображения.*** Кроме того, признаковые классификаторы обладают другим серьёзным недостатком - на этапе извлечения признаков происходит необратимая потеря части информации о символе. Точность работы признакового классификатора сильно зависит от качеботчиков оперируют различными наборами признаков.

Контурный классификатор. Обособленная разновидность признакового классификатора. Отличается от последнего тем, что для извлечения признаков использует *контуры (скелеты),* предварительно выделенные на изображении символа.

Признаковый дифференциальный классификатор. Предназначен для различения похожих друг на друга объектов, таких, например, как литера « *m* » и сочетание « *r n* ». Анализирует только те области изображения, где может находиться информация, позволяющая отдать предпочтение одному из вариантов. Так, в случае с « *m* » и « *r n* » ключом к ответу служит наличие и ширина разрыва в месте касания предполагаемых литер. Признаковый дифференциальный классификатор (ПДК) представляет собой набор признаковых классификаторов.

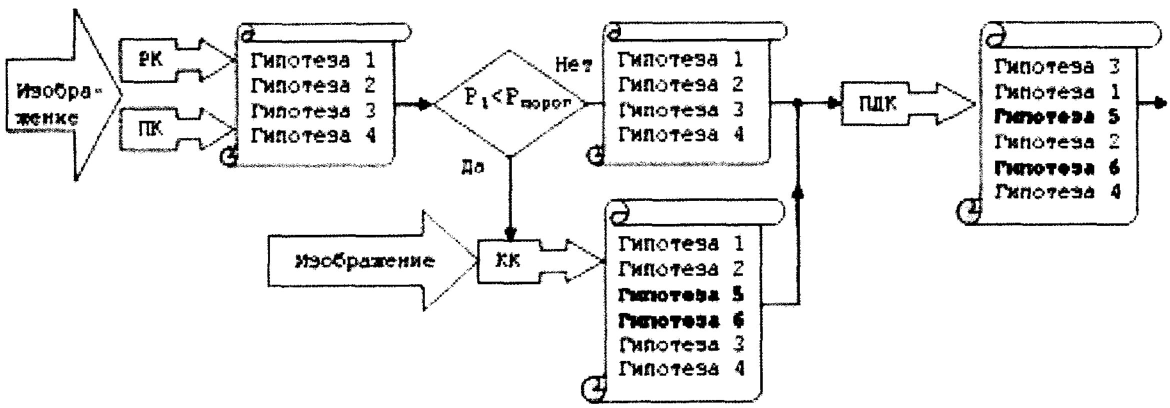


Рис. 2. Обобщенная блок-схема алгоритма распознавания (первый уровень) [1]

Эти последние оперируют эталонами, полученными для каждой пары схожих символов. Для всех пар используется один и тот же набор признаков, аналогичный имеющемуся у соответствующего признакового классификатора.

На рис. 2 представлена схема распознавателя, использующегося в системах *ABBYY*. Растровый (РК) и признаковый (ПК) классификаторы используются для быстрого порождения предварительного списка гипотез. Список поступает на вход ПДК, который производит сортировку.

Построение математических моделей. Анализ работ, посвященных построению математических моделей показал, что в настоящее время в этом направлении ведутся интенсивные исследования. Можно выделить следующие основные направления: методы теории нечётких множеств, стохастические методы [1], нейронные сети [2, 3], представления изображения с помощью гранул [4], с помощью графов [5], методами теории групп Ли [6], с помощью кластерного анализа. Математические модели должны позволять выделить характерные особенности изображения, инвариантные к изменению масштаба и малочувствительные к искажениям изображения в результате перемещений.

Одним из перспективных направлений, позволяющих решить многие из этих проблем, является тензорный анализ [7], [8], представление бинарного изображения тензором отвечает принципу целостности.

Изображение сканируется с достаточно высоким разрешением, что дает возможность получить матрицу высокого порядка, а затем её компаунд-тензор $m \times n$ и определить его численные характеристики (характеристическое уравнение, сингулярные числа, инварианты, собственные числа, магнитуду), некоторые из которых будут являться эталоном для сравнения.

Постановка задачи

Литера представляется тензором [9]. Наиболее простое действие с тензором A состоит в том, чтобы трактовать все его компоненты как тензоры т. е. тензор, у которого каждая компонента сама есть тензор той же валентности, а не обычная величина, называется «компаунд-тензором» [7]. На рис. 4 показано матричное представление литер. На литеру накладывается квадратная сетка. Сторона квадрата равняется высоте литеры. Если литера имеет прямоугольную форму, ее сдвигают в одну сторону, а свободную часть квадрата дополняют нулями.

Исходная матрица для тензора 9x9 (рис. 4 а) – матрица тензора 5-го ранга. Отметим, что тензор 4-го ранга наиболее близок к *скелетному* представлению символа, если любая его часть занимает в дискретном представлении не более одной клетки (рис. 3).

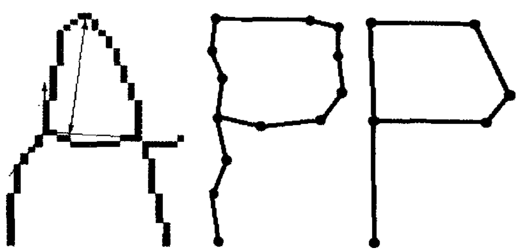
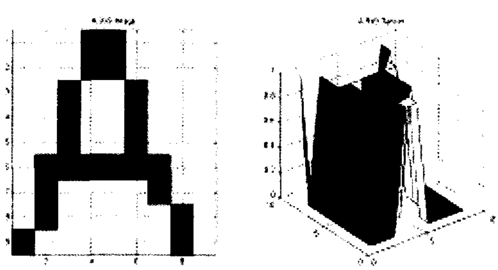
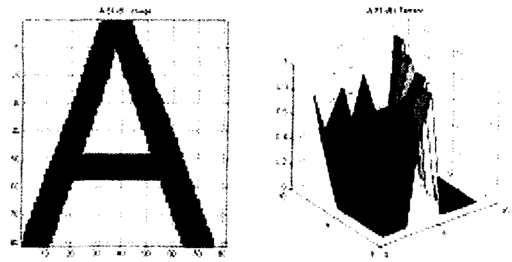


Рис. 3. Примеры скелетных представлений - любая часть символа укладывается в одну клетку

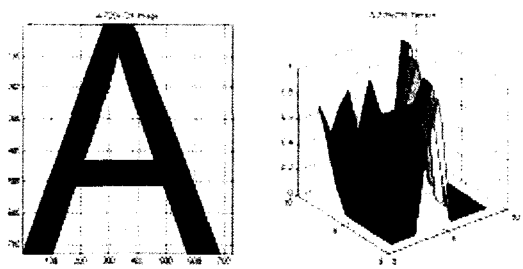


$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

а) (9x9)



б) (81x81)



в) (729x729)

Рис. 4. Представление литер тензорами разной размерности

Собственные числа и коэффициенты характеристического уравнения. Будем трактовать полученный компаунд-тензор m x n как линейное преобразование $y = Ax$ в ортонормированном базисе e_1, e_2. Вектор $x\ (x_1, x_2, \ldots, x_n) \neq 0$ называется собственным вектором этого линейного преобразования,если $Ax = \lambda$, где λ – некоторое действительное число, λ называется собственным значением вектора $\boldsymbol{x}$. Поскольку по предположению вектор $\boldsymbol{x}$ $(x_1, x_2, \ldots, x_n) \neq \boldsymbol{0}$, то определитель матрицы A должен быть равен нулю: $| \lambda E - A | = 0$. Выражение вида: $| \lambda E - A |$, где E – единичная матрица, λ – собственные числа, называется характеристическим (или вековым) уравнением матрицы A, а выражение вида:

$$det\, | \lambda E - A | = c_1\lambda^n + \ldots + c_n\lambda + c_{n+1}$$

называется многочленом характеристического уравнения матрицы A с коэффициентами c_i.

Если матрица имеет размерность $(m \times n)$, то у нее m - собственных векторов, n - собственных чисел, из которых можно составить матрицы собственных векторов $U = [x_1, x_2, \ldots, x_n]$ и собственных чисел

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Таким образом для матрицы A имеет место $AU=U\Lambda$, где U - матрица собственных векторов, Λ - диагональная матрица. Матрица U ортогональная. Согласно теории матриц это значит, что $U^1 = U^T, AUU^T = U\Lambda U^T,\ A = U\Lambda U^T$.

Сингулярные числа. Для матрицы A общего вида, т. е. неквадратной и не симметрической, существует понятие сингулярных чисел. Сингулярное разложение матрицы общего вида имеет вид:

$$A_{(mxn)} = V_{(mxm)}D_{(mxn)}U^T_{(nxn)}.$$

Матрицы V и U ортогональные квадратные матрицы. D - диагональная, но не квадратная матрица. Если m больше n, то m-n строки (случай а) содержат только нулевые элементы. Если m меньше n, n-m столбцы (случай б) содержат только нулевые элементы. Элементы матрицы D называются сингулярными числами. Таким образом, сингулярным разложением действительной матрицы А размеров m x n называется всякое ее разложение вида $A = UDV$, где U – ортогональная матрица размеров m x m, V - ортогональная матрица размеров n x n, D - диагональная матрица размеров m x n, элементы которой $d_{ij}= 0$, если i не равно j, и $d_{ii} = d_i$ в противном случае. Величины d_i называются сингулярными числами матрицы и равны арифметическим значениям квадратных корней из соответствующих собственных значений матрицы AA^T. В англоязычной литературе сингулярное разложение принято называть *SVD*-разложением [10]

$$D_{(n\times m)} = \begin{pmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & d_n \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

а)

$$D_{(n\times m)} = \begin{pmatrix} d & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & d & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & d & \dots & 0 & \dots & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & d & \dots & 0 \end{pmatrix}.$$

б)

Инварианты и магнитуда. Инварианты матрицы определяются по формулам представленным ниже.

$$I_1 = \lambda_1 + \lambda_2 + \dots + \lambda_8,$$
$$I_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \dots + \lambda_1\lambda_8,$$
$$I_3 = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_2\lambda_5 + \lambda_1\lambda_2\lambda_6 + \lambda_1\lambda_2\lambda_7 + \lambda_1\lambda_2\lambda_8 + \lambda_2\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_5 + \lambda_2\lambda_3\lambda_6 + \lambda_2\lambda_3\lambda_7 + \lambda_2\lambda_3\lambda_8 + \lambda_3\lambda_4\lambda_5 + \lambda_3\lambda_4\lambda_6 + \lambda_3\lambda_4\lambda_7 + \lambda_3\lambda_4\lambda_8 + \dots + \lambda_6\lambda_7\lambda_8,$$

где λ - собственные числа матрицы A (9 x 9) ($|\lambda E - A| = 0$);

$$I_1 = a_{11} + a_{22} + a_{33},,$$

$$I_2 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} + \det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}$$

$$I_3 = \det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Во втором случае инварианты считаются с помощью детерминантов составленных из элементов матрицы A. В обоих случаях представлены формулы для трех первых инвариант. В процессе отладки программы использовались формулы обоих вариантов. Результаты совпали и в дальнейшем использовались формулы для детерминантов. Инвариантной характеристикой тензора также является магнитуда, которая определяется по формуле

$$\|[\mathrm{A}]\| = \sqrt{0.5(\mathrm{AA^T})} = \sqrt{0.5\sum_i \sum_j (a_{ij})^2},$$

где a_{ij} – коэффициенты матрицы А.

Таким образом, система инвариантов символа (литеры) может служить основой формирования признаков для решения задачи распознавания.

Список литературы

1. *Анциферов С. С., Евтихеев Н. Н.* Адаптивные системы распознавания образов пространственно-временных полей // Искусственный интеллект. – 2004, – № 3. – С. 405 – 416.

2. *Al Dhouyani Saud, Бутенков С. А.* Синергетический подход к задаче генерализации при обучении нечетких нейронных сетей для обработки изображений // Искусственный интеллект. – 2005, – № 4. – С. 319 – 323.

3. *Бутенков С. А., Al Dhouyani Saud.* Применение нейронных сетей для семантической интеллектуальной сегментации // Искусственный интеллект. – 2006, – № 3. – С. 262 – 269.

4. *Бутенков С. А.* Грануляция и инкапсуляция в системах эффективной обработки многомерной информации // Искусственный интеллект. – 2005, – № 4. – С. 106 – 115.

5. *Агарков А. В.* Иерархическое представление с помощью графа // Искусственный интеллект. – 2007, – № 1. – С. 101 – 109.

6. *Макарычев В. П.* Распознавание плоских изображений методами теории групп Ли // Искусственный интеллект. – 2004, – № 3. – С. 493 – 498.

7. *Габриель Крон.* Тензорный анализ. Москва, «Советское радио», 1978. – 719 с.

8. *Акивис М. А., Гольдберг В. В.* Тензорное исчисление. – М.: 1969. – 351 с.

9. *Минаев Ю. Н., Филимонова О. Ю., Винник Д. Н.* Тензорные модели изображений // Матеріали YIII міжнародної науково-технічної конференції «АВІА-2007». – Т. 1. – С. 17 – 20.

10. *Brannon R. M.* Functional and Structured Tensor Analysis for Engineers. – New Mexico, 2003. – 300 p.