

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ МЕТОДА РАДИАЛЬНЫХ ОКРЕСТНОСТЕЙ

Михайлов И.А.

Ярославский государственный университет им. П.Г. Демидова

Аннотация

Рассматривается задача распознавания чёрно-белых растровых изображений цифровых символов. Для решения данной задачи предлагаются три метода, основным из которых является метод радиальных окрестностей, двумя дополнительными – метод срезов и метод, основанный на модификации метрики Хаусдорфа. Эти методы сравниваются как между собой, так и с известными *CL*- и корреляционным подходами, для чего проводится серия экспериментов с использованием изображений цифровых символов малого размера. Результаты экспериментов показывают эффективность предложенных методов в целом и метода радиальных окрестностей в особенности.

Ключевые слова: оптическое распознавание символов, зашумлённые изображения, метод радиальных окрестностей.

Введение

Оптическое распознавание символов (Optical Character Recognition, OCR) является одним из самых ранних направлений в области распознавания образов (см., напр., [1]). В зависимости от условий и способов начертания символов, а также их последующего перевода в цифровую форму, данное направление разделяется на множество задач. Среди них можно выделить задачу распознавания номеров железнодорожных цистерн (см., напр., [2, 3]). Специфика данной задачи заключается в сравнительно низком разрешении изображений (в большей или меньшей степени), а также в их сильной зашумлённости. Шум может быть достаточно интенсивным, чтобы повредить не только очертания (границы) символа, но и его каркас (см. примеры в разделе 5). Поэтому использование структурного анализа (см. следующий раздел) в этом случае затруднительно. Цель настоящей работы – описать некоторые методы распознавания, применимые в данных условиях.

Статья устроена следующим образом. В разделе 1 кратко описаны существующие методы распознавания символов. Постановка задачи и общий подход к её решению приведены в разделе 2. В разделе 3 дано описание предложенных методов распознавания. В разделе 4 определены *CL*- и корреляционный подходы. Используемые в экспериментах наборы изображений-эталонов и тестовых изображений описаны в разделе 5, там же дана количественная оценка степени искажения тестовых изображений. Раздел 6 отведён для результатов экспериментов и заключения.

1. Краткий обзор методов распознавания символьных изображений

В данном разделе приведены некоторые подходы к задаче распознавания изображений. Выделяются и рассматриваются в первую очередь те методы, которые могут быть полезны для решения поставленной задачи.

В работе [4] приведена общая схема распознающих систем: восприятие → предобработка → выделение признаков → классификация. Для нас будут представлять интерес два последних этапа. Набор

(или система) признаков являются некоторым промежуточным представлением, позволяющим выразить наиболее существенное и исключить второстепенное в контексте классификации. Основными требованиями к такому представлению являются сходство описаний объектов одного и того же класса и различие описаний объектов разных классов.

Этапы выделения признаков и классификации являются взаимосвязанными: чем большая работа будет сделана на первом из них, тем легче выполнить второй.

Остановимся вначале на задаче классификации. В качестве критерия, позволяющего выделить различные семейства методов её решения, можно рассмотреть способ интерпретации совокупности признаков. Наиболее часто с признаками работают как с векторами в евклидовом пространстве. В этом случае для классификации чаще всего применяются методы, не требующие какого-либо предположения о распределении объектов в пространстве признаков. Среди них можно выделить такие распространённые подходы, как правило ближайших соседей (примеры работ см. ниже), машины поддерживающих векторов (Support Vector Machines, SVM; см., напр., [5 - 7]) и нейронные сети [7 - 11].

Интерпретация признаков как вектора (набора чисел) не всегда является удобной, например, если необходимо не только описать степень проявления характеристик объекта, но и выразить взаимосвязь между ними. Для символьных изображений существуют и другие представления. В качестве примера можно привести описание с помощью графа, рёбрами которого могут быть отрезки и дуги, а вершинами – точки их соединения. В случае использования нечисловой интерпретации системы признаков метод классификации является специфическим, он зависит от представления и смысла выделенных характеристик.

Заметим, что существуют методы классификации, применимые как в случае числового (векторного), так и нечислового описания объектов. Примером может служить правило ближайших соседей, для использо-

вания которого необходимо лишь определить меру близости на множестве допустимых объектов.

Перейдём теперь к задаче выделения признаков. Перечислим некоторые применяемые на практике системы признаков.

Видимо, самой простой из них является исходное представление – решётка пикселей. В этом случае промежуточное представление не порождается, эффективность распознавания целиком зависит от метода классификации. Если рассматривается задача распознавания символов с небольшим количеством возможных начертаний, но допускаются значительные шумовые искажения, то часто (и с успехом) применяется следующий подход. Классификация выполняется по правилу ближайших соседей (чаще одного), а в качестве меры близости могут выступать расстояние Хемминга, функции корреляции, взаимной корреляции и подобные. Приведём примеры задач, где применяется данный подход (возможно, с некоторыми модификациями): распознавание регистрационных номеров автомобилей [12-14], номеров железнодорожных цистерн [3], дверных номеров [15]. Среди описанных в настоящей работе методов решётку пикселей в качестве признаков используют корреляционный алгоритм и метод, основанный на модификации метрики Хаусдорфа.

В некоторых случаях используется описание изображений с помощью моментов. При этом применяются центральные моменты [16], моменты Зернике [11] и другие [9].

В рамках структурного анализа изображений используются следующие подходы и системы признаков: анализ каркаса [1, 17-20], анализ контура (границы символа) [1, 20-22], ориентационные признаки (в том числе с использованием градиентных характеристик) [5-8, 11, 23-25].

Кроме того, можно выделить семейство признаков, полученных в результате пересечений. Суть данного подхода заключается в наложении на изображение множества отрезков прямых, взаимное расположение которых может быть различным (например, они могут пересекать изображение по горизонтали, располагаясь равномерно вдоль его высоты, или пересекаться в его центре тяжести). Возможными характеристиками являются количество пересечений этих отрезков с каркасом символа, расстояние между пересечениями, расстояние от границы окаймляющего прямоугольника до пересечений и другие величины. Примеры использования можно найти в работах [9, 10, 26]. Данный подход к описанию изображений применяется в следующих методах, представленных в данной работе: метод срезов, *CL*-подход, метод радиальных окрестностей.

В заключение отметим, что для повышения эффективности различные методы распознавания могут объединяться друг с другом. Распространённой является такая схема работы, когда для анализа более простых изображений применяется быстрый грубый классификатор, а остальные изображения, с которы-

ми он не смог справиться, подаются на вход более сложному и медленному. Примеры объединения распознавателей можно найти в работах [8, 11, 20, 23, 27].

2. Постановка задачи и общий подход к её решению

Рассматривается задача распознавания чёрно-белых растровых изображений цифровых символов. Пусть имеется набор классов, для каждого из которых существует, по крайней мере, одно изображение-эталон. Пусть дано также тестовое изображение (тест). Размеры изображений: как эталонов, так и теста – могут быть различны. Задача состоит в том, чтобы определить, к какому классу относится тестовое изображение.

Опишем теперь общий подход к решению данной задачи для всех описанных в настоящей работе методов. Для классификации теста будем использовать правило ближайшего соседа: вычислим расстояние между тестом и каждым эталоном, сопоставим тесту класс того эталона, на котором реализуется минимум.

Таким образом, поставленная задача распознавания сводится к выбору меры близости на множестве символьных изображений. Поэтому описание каждого метода будет ограничено определением соответствующего расстояния.

3. Предложенные методы распознавания

3.1. Метод срезов

В этом разделе даётся описание метода срезов. Раздел содержит определения бесконечного вектора, линейного образа и определения используемых расстояний. В данной работе дано несколько улучшенное описание метода срезов по сравнению с исходным описанием в работе [28].

Под *бесконечным вектором* V будем понимать вектор, обладающий следующими свойствами:

$$1. |V| = \infty, \quad (1)$$

$$2. \exists n : (\forall i \leq n \ v_i \geq 0) \wedge (\forall i > n \ v_i = 0). \quad (2)$$

Пусть V', V'' – бесконечные векторы. *Расстоянием* на множестве бесконечных векторов будем называть

$$\rho(V', V'') = \sum_{p=1}^{\infty} |v'_p - v''_p|. \quad (3)$$

Пусть имеется матрица изображения $A = (a_{ij})$ размера $n \times m$ (n строк, m столбцов). Для каждого $i : 1 \leq i \leq n$ определим множество чисел $S(A, i)$:

$$S(A, i) = \{s_1, s_2, \dots, s_k\} = \\ = \{0\} \cup \{j : a_{i,j} = 0 \wedge a_{i,j+1} = 1, 1 \leq j \leq m-1\} \cup \{m\}. \quad (4)$$

Будем считать, что $s_1 < s_2 < \dots < s_k$.

Вектором разности для тройки (A, i, N) , где $1 \leq i \leq n$, N – натуральное, будем называть бесконечный вектор действительных чисел, построенный следующим образом:

$$V_p(A, i, N) = \begin{cases} (s_{p+1}(A, i) - s_p(A, i)) \cdot \frac{N}{m}, & 1 \leq p \leq |S(A, i)| - 1; \\ 0, & |S(A, i)| \leq p < \infty. \end{cases} \quad (5)$$

Таким образом, вектор разности $V_p(A, i, N)$ содержит в себе расстояния между соседними переходами с белого на чёрный на высоте i , а также расстояния от крайних переходов до границ изображения (первая и последняя компоненты). Параметр N играет роль коэффициента масштабирования.

Линейным образом матрицы изображения $A = (a_{ij})$ размера $n \times m$ с частотой сечения N , где N – натуральное, будем называть упорядоченное мультимножество векторов разности для этой матрицы

$$L(A, N) = \{l_p : l_p = V(A, \lceil n \cdot p / N \rceil, N), 1 \leq p \leq N\}. \quad (6)$$

Таким образом, линейный образ содержит векторы разности (*срезы*) для всех n строк матрицы изображения (a_{ij}) . При этом для масштабирования каждой строке сопоставляется несколько одинаковых срезов в зависимости от параметра N .

Пусть $L(A_1, N), L(A_2, N)$ – линейные образы матриц изображений A_1 и A_2 с частотой сечения, равной N .

Расстоянием на множестве линейных образов матриц изображений с частотой сечения N будем называть

$$\rho(L(A_1, N), L(A_2, N)) = \sum_{1 \leq i \leq N} \rho(l_i(A_1, N), l_i(A_2, N)). \quad (7)$$

Введённое расстояние между линейными образами будет использоваться в экспериментах в качестве расстояния между символьными изображениями.

3.2. Метод, основанный на модификации метрики Хаусдорфа

Метрика Хаусдорфа и её модификации находят своё применение в задаче распознавания изображений в качестве меры близости множеств [18, 22]. При использовании данной метрики изображения рассматриваются как конечные множества точек, причём точки одного изображения привязываются к точкам другого. В этом разделе описывается подход к решению задачи распознавания, использующий некоторую модификацию метрики Хаусдорфа для определения близости изображений. Исходное описание данного подхода дано в работе [29].

Для того чтобы показать отличие предлагаемого расстояния от оригинальной метрики Хаусдорфа, приведём вначале полное определение последней.

1. Расстоянием от множества из одной точки $\{z'\}$ до компактного множества G будем называть

$$\sigma(z', G) = \min_{z \in G} \rho(z', z), \quad (8)$$

где $\rho(z', z)$ – расстояние между точками z' и z .

2. Отклонением множества G_1 от G_2 называется $\sigma(G_1, G_2) = \max_{z' \in G_1} \sigma(z', G_2)$. (9)

3. Метрикой Хаусдорфа называется

$$\rho(G_1, G_2) = \max\{\sigma(G_1, G_2), \sigma(G_2, G_1)\}. \quad (10)$$

Отличия предлагаемого расстояния от метрики Хаусдорфа содержатся во втором и третьем пунктах.

4. Отклонением множества G_1 от G_2 будем называть

$$\tilde{\sigma}(G_1, G_2) = \frac{\sum_{z' \in G_1} \sigma(z', G_2)}{|G_1|}. \quad (11)$$

5. hm -расстоянием будем называть

$$\tilde{\rho}_{hm}(G_1, G_2) = \tilde{\sigma}(G_1, G_2) + \tilde{\sigma}(G_2, G_1). \quad (12)$$

Для того чтобы применить описанный подход к чёрно-белым изображениям, каждому чёрному (принадлежащему символу) пикселю с координатами (i, j) сопоставим одну точку с теми же координатами. Заметим, что до вычисления расстояния между изображениями тест приводится к размеру эталона с помощью алгоритма масштабирования, описанного в работе [28]. Введённое hm -расстояние вычисляется два раза: в первом случае совмещаются центры тяжести изображений, во втором – их верхние левые углы. Минимум из двух найденных величин используется в качестве меры близости между изображениями в экспериментах.

3.3. Метод радиальных окрестностей

В этом разделе даётся определение метода радиальных окрестностей. Этот метод имеет некоторые общие черты с *CL-подходом*, кратко описанным в работе [9]. Основная идея метода радиальных окрестностей была предложена в работе [29].

Дадим вначале определения используемых структур. Пусть имеется изображение A , заданное матрицей (a_{ij}) размера $n \times m$. Под *радиальной окрестностью* элемента матрицы с координатами (i, j) и коэффициентом нормирования s будем понимать упорядоченный набор из четырёх бинарных векторов (v^l, v^r, v^t, v^b) , построенных следующим образом:

$$\begin{aligned} \bullet v_k^l &= a_{i, j'}, j' = \left\lceil m \cdot \frac{k-1}{s} \right\rceil, \\ & 2 \leq k \leq \left\lfloor \frac{s \cdot (j-0,5)}{m} + 1 \right\rfloor; \end{aligned} \quad (13)$$

$$\begin{aligned} \bullet v_k^r &= a_{i, j'}, j' = \left\lfloor m \cdot \frac{s-k+1}{s} \right\rfloor, \\ & 2 \leq k \leq \left\lfloor \frac{s \cdot (m-j+0,5)}{m} + 1 \right\rfloor; \end{aligned} \quad (14)$$

$$\bullet v_k^t = a_{i',j}, i' = \left\lceil n \cdot \frac{k-1}{s} \right\rceil, \quad (15)$$

$$2 \leq k \leq \left\lfloor \frac{s \cdot (i-0,5)}{n} + 1 \right\rfloor;$$

$$\bullet v_k^b = a_{i',j}, i' = \left\lceil n \cdot \frac{(s-k+1)}{s} \right\rceil, \quad (16)$$

$$2 \leq k \leq \left\lfloor \frac{s \cdot (n-i+0,5)}{n} + 1 \right\rfloor.$$

При этом $v_1^l = v_1^r = v_1^t = v_1^b = 1$. Каждый из этих векторов соответствует одному из четырёх основных направлений *от границы окаймляющего прямоугольника к центру окрестности* (элементу с координатами (i, j)): v^l – слева направо, v^r – справа налево, v^t – сверху вниз, v^b – снизу вверх. Каждый вектор включает в себя масштабированную (с помощью параметра s) часть строки или столбца вдоль соответствующего направления от границы изображения до центра окрестности. Исключением являются первые компоненты векторов – они всегда равны 1 независимо от цвета пикселей.

Для иллюстрации определения окрестности рассмотрим пример изображения на рис. 1. Размер изображения – 8×14 пикселей. Каждый значащий (то есть, принадлежащий символу) пиксель изображения здесь выделен отдельным закрашенным квадратиком, а пиксель, для которого будет построена окрестность, – квадратиком с точкой.

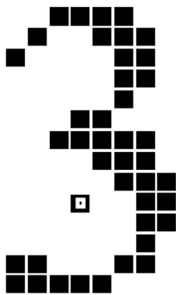


Рис. 1. Построение радиальной окрестности

Искомая окрестность включает в себя следующие векторы при $s = 20$:

- $v^l = 100000000$;
- $v^r = 111110000000$;
- $v^t = 11000000111000$;
- $v^b = 11000000$.

Под *радиальным образом* изображения будем понимать множество радиальных окрестностей, построенных для каждого элемента соответствующей матрицы (a_{ij}) (то есть, для каждого пикселя изображения).

Введём теперь расстояния на множестве бинарных векторов, радиальных окрестностей и радиальных образов изображений.

bt-расстоянием на множестве бинарных векторов будем называть

$$\tilde{\rho}_{bm}(a, b) = \sigma(a, b) + \sigma(b, a), \quad (17)$$

где $\sigma(a, b) = \sum_{a_i=1} \min_{b_j=1} |i - j|$.

Заметим, что при использовании данного расстояния для сравнения бинарных векторов радиальных окрестностей указанный минимум всегда может быть найден, так как каждый такой вектор имеет, по крайней мере, один ненулевой элемент по построению.

Введём расстояние на множестве радиальных окрестностей:

$$\begin{aligned} \tilde{\rho}(e_1, e_2) = & \tilde{\rho}_{bm}(v^l(e_1), v^l(e_2)) + \\ & + \tilde{\rho}_{bm}(v^r(e_1), v^r(e_2)) + \\ & + \tilde{\rho}_{bm}(v^t(e_1), v^t(e_2)) + \\ & + \tilde{\rho}_{bm}(v^b(e_1), v^b(e_2)). \end{aligned} \quad (18)$$

Определим *pp-расстояние* на множестве радиальных образов:

$$\tilde{\rho}_{pp}(R', R'') = \sigma(R', R'') + \sigma(R'', R'), \quad (19)$$

где $\sigma(R', R'') = \frac{1}{|R'|} \cdot \sum_{i=1}^{|R'|} \min_j \tilde{\rho}(R'_i, R''_j)$.

Здесь R'_i – i -ая окрестность образа R' , R''_j – j -ая окрестность образа R'' . Введённое *pp-расстояние* на множестве радиальных образов будет использоваться в экспериментах в качестве меры близости между изображениями.

4. Другие методы распознавания

4.1. Расширенный CL-подход

Усовершенствованная версия базового *CL*-подхода (Characteristic Loci, CL), первоначально предложенного в работе [30], была успешно применена к распознаванию рукописных цифр [9]. Как уже было сказано, этот метод имеет ряд сходных черт с методом радиальных окрестностей. Приведём теперь различия между ними и тем самым неформально опишем базовый *CL*-подход и его расширенный вариант.

При использовании этого метода в качестве системы признаков также выступает совокупность окрестностей пикселей изображения. Первое отличие состоит в том, что в *CL*-подходе окрестности вычисляются не для всех пикселей, но только для фоновых (не принадлежащих символу, белых). В качестве составляющих окрестности выступают количества переходов с белого на чёрный вдоль одного из четырёх основных направлений *от центра окрестности до границы окаймляющего прямоугольника*. Это второе отличие. Для примера приведём окрестность выделенного пикселя на рис. 1: $(0, 1, 2, 1)$ (порядок направлений следующий: налево, направо, вверх, вниз).

Заметим, что допустимое число переходов ограничено сверху: если количество пересечений сканирующего луча с символом больше определённого числа (назовём его T_M), то соответствующая компо-

нента окрестности заменяется этим числом. В базовом CL -подходе параметр T_M положен равным 2.

При использовании расширенного CL -подхода в процесс построения окрестностей вносятся два изменения. Во-первых, параметр T_M полагается равным 3. Во-вторых, после обрезки (см. раздел 5, первый абзац) каждое символьное изображение дополнительно окружается белой полосой шириной в 4 пикселя с каждой стороны. Во время экспериментов, описанных в следующем разделе, будет использоваться расширенная версия CL -подхода, как и в работе [9].

Легко видеть, что окрестности в CL -подходе представляют собой всевозможные упорядоченные четвёрки чисел, начиная от $(0, 0, 0, 0)$ и заканчивая (T_M, T_M, T_M, T_M) . При этом набор окрестностей для некоторого изображения может включать в себя несколько экземпляров одной и той же четвёрки (например, окрестность для всех пикселей, соседних выделенному на рис. 1, будет одной: $(0, 1, 2, 1)$). Таким образом, для каждой окрестности можно вычислить её *относительную частоту* (далее просто частоту) как отношение количества её экземпляров к суммарному количеству экземпляров всех окрестностей.

В работе [9] для классификации применяются различные методы, в том числе правило ближайшего соседа, однако строгое определение используемого расстояния отсутствует. В настоящей работе в качестве меры близости двух изображений будет использоваться евклидово расстояние между наборами частот окрестностей. Для каждой окрестности можно вычислить квадрат разности её частот в первом и втором изображениях. Квадратный корень из суммы таких величин для всевозможных окрестностей есть расстояние между изображениями.

4.2. Корреляционный алгоритм

Корреляционный алгоритм был использован для распознавания номеров железнодорожных цистерн на предприятии «Славнефть-ЯНОС» [3]. В качестве степени сходства изображений в нём использовалась следующая величина:

$$\lambda_{ij} = \max_{(i,j)} \left\{ \frac{k_{ij}}{n_{ij} + 1} \right\}. \quad (20)$$

Здесь (i, j) задаёт смещение теста относительно эталона (их взаимное расположение), k_{ij} – количество чёрных совпадающих пикселей в эталоне и тесте, n_{ij} – количество чёрных несовпадающих пикселей в тесте и эталоне. Заметим, что величина λ_{ij} возрастает по мере увеличения сходства между изображениями в противоположность различным расстояниям, описанным ранее. Поэтому в качестве меры близости изображений в экспериментах будет использоваться противоположная величина:

$$\overline{\lambda}_{ij} = -\lambda_{ij}. \quad (21)$$

5. Используемые наборы изображений

В этом разделе будут описаны используемые нами наборы эталонов и тестов и представлена количественная оценка искажения тестовых изображений.

5.1. Наборы изображений-эталонов

В данной работе используются два набора эталонов. Первый из них включает в себя 10 изображений, соответствующих 10 классам символов: «0» – «9». Все эти изображения были созданы с помощью шрифта Times New Roman и имеют одинаковую высоту, равную 14 пикселям. Данный набор далее будем называть классическим и обозначать CS.

Второй набор включает в себя изображения, вырезанные из бинаризованных фотографий цистерн. Этот набор использовался для тестирования в упомянутой выше работе [3]. Набор включает в себя экземпляры изображений для 9 классов символов: «0» и «2» – «9». Общее количество изображений в нём равно 55, их средний размер – 8×12 пикселей. На рис. 2 показаны примеры изображений из этого набора. Данный набор далее будем называть экспериментальным и обозначать ES.

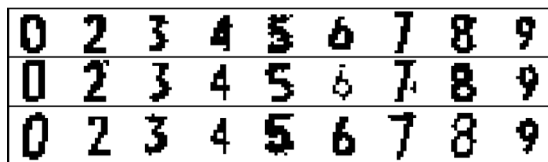


Рис. 2. Примеры изображений-эталонов

5.2. Наборы тестовых изображений

В настоящей работе используются 5 различных наборов тестовых изображений. Опишем каждый из них в отдельности.

Изображения из первых четырёх наборов были построены путём искажения изображений-эталонов из классического набора. Приведём теперь используемые модели искажения и тем самым дадим описание тестовых наборов. В конце описания каждой модели будем давать условное обозначение для соответствующего набора изображений.

- Нанесение на эталон одиночных чёрных пикселей. Расположение каждого пикселя выбиралось случайно. Их количество варьировалось от 1 до 50. Обозначение – NP.
- Нанесение на эталон прямых линий. Направление, длина и толщина линий были случайны; длина колебалась в пределах от 1 до 7 пикселей, толщина – от 1 до 3 пикселей. Цвет половины из них был чёрным, цвет остальных – белым. Обозначение – NL.

Алгоритмы искажения каркаса символа, описанные далее, на входе получают последовательность узлов (или цепь), построенную вручную. Узел представляет собой координаты некоторой точки каркаса. Цепь в среднем содержит 25 узлов, распределён-

ных по всему символу. В цепь символа-эталона вносились изменения, затем она отображалась в растр.

- Сдвиг каждого узла цепи на случайный вектор. При этом координаты узла изменялись не более чем на 20% от размеров изображения. Обозначение – SS.
- Разбиение каркаса символа на несколько сегментов, поворот каждого сегмента относительно узла, случайно выбранного в данном сегменте. Количество сегментов выбиралось случайно в диапазоне от 2 до 6. Угол поворота лежал в пределах от $-\pi/6$ до $\pi/6$. Обозначение – ST.

Каждый из четырёх приведённых тестовых наборов содержит 5000 изображений: по 500 для каждого из 10 эталонов. Размер каждого тестового изображения совпадает с размером соответствующего эталона.

В качестве пятого набора тестов были взяты изображения символов, вырезанные из бинаризованных фотографий цистерн. Набор включает в себя изображения, соответствующие 10 классам символов: «0» – «9». Общее количество изображений в нём равно 1614, их средний размер – 8×15 пикселей. На рис. 3 показаны некоторые изображения из этого набора. Данный набор будем так же называть экспериментальным (по аналогии с эталонами) и обозначать EXP.

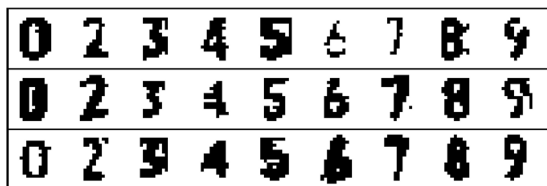


Рис. 3. Примеры тестовых изображений

5.3. Оценка степени искажения изображений

Для того чтобы облегчить интерпретацию, сравнение и практическое использование полученных в ходе экспериментов результатов, была выполнена оценка степени искажения всех изображений из тестовых наборов. Оценка была осуществлена следующим образом. Для каждого тестового изображения было найдено соответствующее ему изображение-эталон из классического набора, после чего вычислена метрика Хаусдорфа (оригинальная) между тестом и эталоном. Величина метрики в данном случае выражается в пикселях. Таким образом, в качестве оценки степени искажения тестового изображения выступает значение метрики Хаусдорфа между ним и соответствующим эталоном из классического набора.

Заметим, что, как и при использовании описанной выше модификации метрики Хаусдорфа, тестовое изображение вначале было приведено к размеру эталона, а в качестве результата возвращался минимум из двух вычислений: в первом случае совмещались левые верхние углы изображений, а во втором – их центры тяжести. В качестве расстояния между точками использовалась сумма модулей разностей координат.

В результате вычисления оценки искажения в зависимости от её величины было выделено 3 класса изображений. Первый класс включает в себя изображения, оценка искажения которых не превышает 2 единиц, второй класс – 3 единиц и третий класс – 7 единиц. В используемых тестовых наборах не найдено ни одного изображения, оценка искажения которого превышает 7 единиц. Приведём теперь в табличном виде количество изображений в каждом классе (верхняя подстрока) и их среднюю оценку искажения (нижняя подстрока) для всех используемых тестовых наборов.

Таблица 1. Средняя оценка искажения изображений по классам

Класс	NP	NL	SS	ST	EXP
1	1534	2353	2630	2268	714
	1,822	1,690	1,975	1,758	1,993
2	1718	1766	2112	1972	674
	3,000	3,000	3,000	3,000	3,000
3	1748	881	258	760	226
	4,356	4,243	4,213	4,222	4,168
Всего	5000	5000	5000	5000	1614
	3,113	2,602	2,523	2,623	2,718

6. Результаты распознавания изображений

Для оценки эффективности разработанных методов был проведён ряд экспериментов. Заметим, что каждое изображение до запуска алгоритмов распознавания подвергалось предобработке: из соответствующей матрицы удалялись крайние нижние и верхние строки, а также крайние левые и правые столбцы, не содержащие единичных (соответствующих символу) элементов.

Отметим также, что в алгоритме метода срезов частота сечений бралась равной 150, а в алгоритме метода радиальных окрестностей коэффициент нормирования s брался равным 100. В методе, основанном на модификации метрики Хаусдорфа, в качестве расстояния между точками использовалась следующая величина:

$$\rho(z', z) = \frac{|z_x - z'_x|}{w} + \frac{|z_y - z'_y|}{h}, \quad (22)$$

где z_x, z'_x – их абсциссы, z_y, z'_y – ординаты, w, h – ширина и высота изображения соответственно.

Заметим, наконец, что при проведении экспериментов с использованием набора эталонов ES и набора тестов EXP из тестового набора удалялись все изображения символа «1», так как набор ES не содержит эталонов для данного символа. Урезанный таким образом набор EXP содержит 1426 изображений.

При описании результатов будем использовать следующие обозначения:

- Corr – метод, использующий корреляционный алгоритм;
- ECL – расширенный CL-подход;
- Slice – метод срезов;

- MНaus – метод, основанный на модификации метрики Хаусдорфа;
- Radial – метод радиальных окрестностей.

Приведём теперь результаты экспериментов по распознаванию в табличном виде для каждого класса отдельно: в первом столбце через косую черту указаны используемые наборы эталонов и тестов соответственно, в остальных ячейках указано количество верно распознанных изображений (в процентах).

Таблица 2. Результаты распознавания для первого класса

Наборы	Согг	ECL	Slice	MНaus	Radial
CS/NP	100,00	59,26	97,46	99,61	99,93
CS/NL	99,79	76,75	98,68	96,94	95,79
CS/SS	69,35	45,70	71,14	88,25	89,73
CS/ST	88,71	89,86	88,10	91,58	84,57
CS/EXP	73,95	62,61	66,25	80,39	85,57
ES/EXP	94,27	89,07	96,24	98,03	99,64

Таблица 3. Результаты распознавания для второго класса

Наборы	Согг	ECL	Slice	MНaus	Radial
CS/NP	99,88	23,92	86,79	99,48	89,58
CS/NL	96,43	44,45	88,05	89,35	86,24
CS/SS	50,47	29,45	53,17	73,77	80,26
CS/ST	53,25	59,89	60,40	65,21	46,45
CS/EXP	38,13	48,07	38,28	56,38	74,93
ES/EXP	87,69	67,29	89,41	89,56	92,06

Таблица 4. Результаты распознавания для третьего класса

Наборы	Согг	ECL	Slice	MНaus	Radial
CS/NP	99,20	8,92	73,51	90,90	86,27
CS/NL	88,76	25,77	68,10	72,08	71,51
CS/SS	50,39	21,71	37,60	57,75	74,81
CS/ST	36,05	33,42	33,29	33,95	28,03
CS/EXP	40,71	19,03	32,30	48,23	66,37
ES/EXP	81,42	47,35	76,99	81,42	87,17

Приведём теперь общие результаты распознавания (без разбиения изображений на классы).

Таблица 5. Общие результаты распознавания

Наборы	Согг	ECL	Slice	MНaus	Radial
CS/NP	99,68	29,52	85,42	96,52	91,60
CS/NL	96,66	56,36	89,54	89,88	88,14
CS/SS	60,40	37,60	61,82	80,56	84,96
CS/ST	66,72	69,46	68,84	72,42	60,94
CS/EXP	54,34	50,43	49,81	65,86	78,44
ES/EXP	89,27	72,65	90,11	91,58	94,25

Теперь можно сделать некоторые выводы.

- Метод срезов оказался наиболее быстрым из предложенных методов.
- Метод радиальных окрестностей позволяет добиваться наибольшей степени распознавания реальных зашумлённых изображений сре-

ди описанных в данной работе подходов. Однако он же является и самым медленным из них.

- Расширение набора эталонов позволяет значительно повысить степень распознавания всех приведённых методов.

Предложенные подходы могут найти своё применение в системах распознавания бортовых номеров железнодорожных цистерн, регистрационных номеров автомобилей, печатного и рукописного текста и других.

Литература

1. **Mori, S.** Historical Review of OCR Research and Development / S. Mori, C.Y. Suen, K. Yamamoto // Proceedings of the IEEE. – 1992. – Vol. 80, N. 7. – P. 1029-1058.
2. **Волотовский, С.Г.** Система технического зрения для распознавания номеров железнодорожных цистерн с использованием модифицированного коррелятора в метрике Хаусдорфа / С.Г. Волотовский, Н.Л. Казанский, С.Б. Попов, Р.В. Хмельёв // Компьютерная оптика. – 2005. – № 27. – С. 177-184.
3. **Карлин, А.К.** Распознавание номеров железнодорожных цистерн с использованием корреляционного алгоритма / А.К. Карлин, А.Н. Малков, Е.А. Тимофеев, Г.П. Штерн // Математика, кибернетика, информатика. Труды международной научной конференции, посвящённой памяти профессора А.Ю. Левина (Ярославль, 25-26 июня, 2008), под ред. С.А. Кашенко, В.А. Соколова; – Ярославль: ЯрГУ, 2008. – С. 103-110.
4. **Дуда, Р.** Распознавание образов и анализ сцен / Р. Дуда, П. Харт; пер. с англ. – М.: Мир, 1976. – 511 с.
5. Chinese character recognition: history, status and prospects / R. Dai, C. Liu, B. Xiao // Frontiers of Computer Science in China. – 2007. – Vol. 1, N. 2. – P. 126-136.
6. **Malon, C.** Support Vector Machines for Mathematical Symbol Recognition / C. Malon, S. Uchida, M. Suzuki // Structural, Syntactic, and Statistical Pattern Recognition. – 2006. – Vol. 4109/2006. – P. 136-144.
7. Offline Chinese handwriting recognition: assessment of current technology / S.N. Srihari, X. Yang, G.R. Ball // Frontiers of Computer Science in China. – 2007. – Vol. 1, N. 2. – P. 137-155.
8. **AbdulKader, A.** A Two-Tier Arabic Offline Handwriting Recognition Based on Conditional Joining Rules / A. AbdulKader // Arabic and Chinese Handwriting Recognition, 2008. – Vol. 4768/2008. – P. 70-81.
9. **Abou-zeid, H.M.R.** Computer Recognition of Unconstrained Handwritten Numerals / H.M.R. Abou-zeid, A.S. El-ghazal, A.A. Al-khatib // Circuits and Systems. Proceedings of the 46th IEEE International Midwest Symposium on, 2004. – P. 969-973.
10. **Kacalak, W.** Selected Problems of Intelligent Handwriting Recognition / W. Kacalak, K.D. Stuart, M. Majewski // Analysis and Design of Intelligent Systems using Soft Computing Techniques. – 2007. – Vol. 41/2007. – P. 298-305.
11. **Song, L.** Method for Automatic Image Recognition based on Algorithm Fusion / L. Song, Y. Lin // Proceedings of the Third IEEE International Conference on Natural Computation. – 2007. – P. 671-675.
12. **Ahmed, M.J.** License plate recognition system / M.J. Ahmed, M. Sarfraz, A. Zidouri, W.G. Al-Khatib // Proc. IEEE Int. Electronics, Circuits and Systems Conf. – 2003. – Vol. 2. – P. 898-901.

13. **Comelli, P.** Optical Recognition of Motor Vehicle License Plates / P. Comelli, P. Ferragina, M.N. Granieri, F. Stabile // IEEE Transactions on Vehicular Technology. – 1995. – Vol. 44. – N. 4. – P. 790-799.
14. **Huang, R.** License Plate Character Recognition Using Artificial Immune Technique / R. Huang, H. Tawfik, A. Nagar // Computational Science, 2008. – Vol. 5101/2008. – P. 823-832.
15. Rapid Door Number Recognition by a Humanoid Mobile Robot / Q. Meng // Journal of Intelligent and Robotic Systems. – 2005. – Vol. 43, N 1. – P. 33-54.
16. **Einsele, F.** Recognition of Ultra Low Resolution Word Images Using HMMs / F. Einsele, R. Ingold, J. Hennebert // Computer Recognition Systems 2, 2007. – Vol. 45/2007. – P. 429-436.
17. **Assabie, Y.** Structural and Syntactic Techniques for Recognition of Ethiopic Characters / Y. Assabie, J. Bigun // Structural, Syntactic, and Statistical Pattern Recognition. – 2006. – Vol. 4109/2006. – P. 118-126.
18. **Kara, L.B.** Sketch Understanding for Engineering Software: Ph.D. Thesis Proposal / L.B. Kara. – Pittsburgh, 2003. – 68 p.
19. **Sternby, J.** Frame Deformation Energy Matching of On-Line Handwritten Characters / J. Sternby // Progress in Pattern Recognition, Image Analysis and Applications. – 2005. – Vol. 3773/2005. – P. 128-137.
20. **Suen, C.Y.** Computer Recognition of Unconstrained Handwritten Numerals / C.Y. Suen, C. Nadal, R. Legault, T.A. Mai, L. Lam // Proceedings of the IEEE. – 1992. – Vol. 80, N 7. – P. 1162-1180.
21. **Hu, J.** Structural Boundary Feature Extraction for Printed Character Recognition / J. Hu, D. Yu, H. Yan // Joint IAPR International Workshops SSPR'98 and SPR'98 Sydney, Australia, Proceedings. – 1998. – P. 500-507.
22. **Хмельёв, Р.В.** Совместное использование структурного анализа и метрики Хаусдорфа при сравнении объекта и эталона / Р.В. Хмельёв // Компьютерная оптика. – 2005. – № 27. – С. 174-176.
23. Dual Classifier System for Handprinted Alphanumeric Character Recognition / Y.-C. Chim, A.A. Kassim, Y. Ibrahim // Pattern Analysis & Applications. – 1998. – Vol. 1, N 3. – P. 155-162.
24. **Pal, U.** Handwritten Character Recognition of Popular South Indian Scripts / U. Pal, N. Sharma, T. Wakabayashi, F. Kimura // Arabic and Chinese Handwriting Recognition, 2008. – Vol. 4768/2008. – P. 251-264.
25. **Yagi, M.** A Human-Perception-Like Image Recognition System Based on PAP Vector Representation With Multi Resolution Concept / M. Yagi, T. Shibata // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. – 2002. – Vol. 1. – P. 1045-1048.
26. **Yu, K.** Skeleton-Based Recognition of Chinese Calligraphic Character Image / K. Yu, J. Wu, Y. Zhuang // Advances in Multimedia Information Processing. – 2008. – Vol. 5353/2008. – P. 228-237.
27. **Wachenfeld, S.** A Multiple Classifier Approach for the Recognition of Screen-Rendered Text / S. Wachenfeld, S. Fleischer, X. Jiang // Computer Analysis of Images and Patterns. – 2007. – Vol. 4673/2007. – P. 921-928.
28. **Михайлов, И.А.** Об одном методе распознавания изображений / И.А. Михайлов // Моделирование и анализ информационных систем. – 2007. – Т. 14, № 4. – С. 7-12.
29. **Михайлов, И.А.** Некоторые методы распознавания изображений / И.А. Михайлов // Моделирование и анализ информационных систем. – 2008. – Т. 15, № 4. – С. 56-64.
30. **Glucksman, H.** Classification of mixed-font alphabets by characteristic loci / H. Glucksman // Digest of 1st Annual IEEE Comp. Conf. – 1967. – P. 138-141.

References

1. **Mori, S.** Historical Review of OCR Research and Development / S. Mori, C.Y. Suen, K. Yamamoto // Proceedings of the IEEE. – 1992. – Vol. 80, N. 7. – P. 1029-1058.
2. **Volotovskiy, S.G.** A machine vision system for tank-wagon number recognition using a modified correlator in the Hausdorff distance / S.G. Volotovskiy, N.L. Kazanskiy, S.B. Popov, R.V. Khmelev // Computer Optics. – 2005. – N. 27. – P. 177-184. – (In Russian).
3. **Karlin, A.K.** Tank-wagon number recognition using a correlation algorithm / A.K. Karlin, A.N. Malkov, E.A. Timofeev, G.P. Shtern // Mathematics, cybernetics, informatics. Proceedings of the International Science Conference devoted to Prof. A.Yu. Levin (Yaroslavl, June 25-26, 2008) / Edited by S.A. Kaschenko, V.A. Sokolov; Yaroslavl State University. – Yaroslavl: YSU, 2008. – P. 103-110. – (In Russian).
4. **Duda, R.O.** Pattern classification and scene analysis / R.O. Duda, P.E. Hart. – Wiley-Interscience, John Wiley & Sons. N.Y.; London; Sydney; Toronto, 1973.
5. Chinese character recognition: history, status and prospects / R. Dai, C. Liu, B. Xiao // Frontiers of Computer Science in China. – 2007. – Vol. 1, N. 2. – P. 126-136.
6. **Malon, C.** Support Vector Machines for Mathematical Symbol Recognition / C. Malon, S. Uchida, M. Suzuki // Structural, Syntactic, and Statistical Pattern Recognition, 2006. – Vol. 4109/2006. – P. 136-144.
7. Offline Chinese handwriting recognition: assessment of current technology / S.N. Srihari, X. Yang, G.R. Ball // Frontiers of Computer Science in China. – 2007. – Vol. 1, N. 2. – P. 137-155.
8. **AbdulKader, A.** A Two-Tier Arabic Offline Handwriting Recognition Based on Conditional Joining Rules / A. AbdulKader // Arabic and Chinese Handwriting Recognition, 2008. – Vol. 4768/2008. – P. 70-81.
9. **Abou-zeid, H.M.R.** Computer Recognition of Unconstrained Handwritten Numerals / H.M.R. Abou-zeid, A.S. El-ghazal, A.A. Al-khatib // Circuits and Systems. Proceedings of the 46th IEEE International Midwest Symposium on, 2004. – P. 969-973.
10. **Kacalak, W.** Selected Problems of Intelligent Handwriting Recognition / W. Kacalak, K.D. Stuart, M. Majewski // Analysis and Design of Intelligent Systems using Soft Computing Techniques. – 2007. – Vol. 41/2007. – P. 298-305.
11. **Song, L.** Method for Automatic Image Recognition based on Algorithm Fusion / L. Song, Y. Lin // Proceedings of the Third IEEE International Conference on Natural Computation. – 2007. – P. 671-675.
12. **Ahmed, M.J.** License plate recognition system / M.J. Ahmed, M. Sarfraz, A. Zidouri, W.G. Al-Khatib // Proc. IEEE Int. Electronics, Circuits and Systems Conf. – 2003. – Vol. 2. – P. 898-901.
13. **Comelli, P.** Optical Recognition of Motor Vehicle License Plates / P. Comelli, P. Ferragina, M.N. Granieri, F. Stabile // IEEE Transactions on Vehicular Technology. – 1995. – Vol. 44, N 4. – P. 790-799.
14. **Huang, R.** License Plate Character Recognition Using Artificial Immune Technique / R. Huang, H. Tawfik, A. Nagar // Computational Science, 2008. – Vol. 5101/2008. – P. 823-832.

15. Rapid Door Number Recognition by a Humanoid Mobile Robot / Q. Meng // Journal of Intelligent and Robotic Systems. – 2005. – Vol. 43, N 1. – P. 33-54.
16. **Einsele, F.** Recognition of Ultra Low Resolution Word Images Using HMMs / F. Einsele, R. Ingold, J. Hennebert // Computer Recognition Systems 2. – 2007. – Vol. 45/2007. – P. 429-436.
17. **Assabie, Y.** Structural and Syntactic Techniques for Recognition of Ethiopic Characters / Y. Assabie, J. Bigun // Structural, Syntactic, and Statistical Pattern Recognition. – 2006. – Vol. 4109/2006. – P. 118-126.
18. **Kara, L.B.** Sketch Understanding for Engineering Software: Ph.D. Thesis Proposal / L.B. Kara. – Pittsburgh, 2003. – 68 p.
19. **Sternby, J.** Frame Deformation Energy Matching of On-Line Handwritten Characters / J. Sternby // Progress in Pattern Recognition, Image Analysis and Applications. – 2005. – Vol. 3773/2005. – P. 128-137.
20. **Suen, C.Y.** Computer Recognition of Unconstrained Handwritten Numerals / C.Y. Suen, C. Nadal, R. Legault, T.A. Mai, L. Lam // Proceedings of the IEEE. – 1992. – Vol. 80, N 7. – P. 1162-1180.
21. **Hu, J.** Structural Boundary Feature Extraction for Printed Character Recognition / J. Hu, D. Yu, H. Yan // Joint IAPR International Workshops SSPR'98 and SPR'98 Sydney, Australia, Proceedings. – 1998. – P. 500-507.
22. **Khmelev, R.V.** Combined application of structural analysis and the Hausdorff distance for comparison of object and template / R.V. Khmelev // Computer Optics. – 2005. – N 27. – P. 174-176. (In Russian).
23. Dual Classifier System for Handprinted Alphanumeric Character Recognition / Y.-C. Chim, A.A. Kassim, Y. Ibrahim // Pattern Analysis & Applications. – 1998. – Vol. 1, N. 3. – P. 155-162.
24. **Pal, U.** Handwritten Character Recognition of Popular South Indian Scripts / U. Pal, N. Sharma, T. Wakabayashi, F. Kimura // Arabic and Chinese Handwriting Recognition. – 2008. – Vol. 4768/2008. – P. 251-264.
25. **Yagi, M.** A Human-Perception-Like Image Recognition System Based on PAP Vector Representation With Multi Resolution Concept / M. Yagi, T. Shibata // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. – 2002. – Vol. 1. – P. 1045-1048.
26. **Yu, K.** Skeleton-Based Recognition of Chinese Calligraphic Character Image / K. Yu, J. Wu, Y. Zhuang // Advances in Multimedia Information Processing. – 2008. – Vol. 5353/2008. – P. 228-237.
27. **Wachenfeld, S.** A Multiple Classifier Approach for the Recognition of Screen-Rendered Text / S. Wachenfeld, S. Fleischer, X. Jiang // Computer Analysis of Images and Patterns. – 2007. – Vol. 4673/2007. – P. 921-928.
28. **Mikhaylov, I.A.** A certain image recognition method / I.A. Mikhaylov // Modeling and analysis of information systems. – 2007. – Vol. 14, N 4. – P. 7-12. – (In Russian).
29. **Mikhaylov I.A.** Some image recognition methods / I.A. Mikhaylov // Modeling and analysis of information systems. – 2008. – Vol. 15, N 4. – P. 56-64. – (in Russian).
30. **Glucksman, H.** Classification of mixed-font alphabets by characteristic loci / H. Glucksman // Digest of 1st Annual IEEE Comp. Conf., 1967. – P. 138-141.

IMAGE RECOGNITION USING A RADIAL NEIGHBORHOOD METHOD

I.A. Mikhaylov

P. G. Demidov Yaroslavl State University

Abstract

Numeral character recognition is considered in this paper. Three recognition methods are proposed. The radial neighborhood method is basic, whereas the slice method and the method, based on a modified Hausdorff distance, are additional. A series of experiments is performed to compare these methods with the *CL*-approach and the correlation algorithm. Artificial and real noised images are used as the test samples. Resolution of these images is low. Experimental results reveal an effectiveness of the proposed methods, most notably the radial neighborhood method.

Key words: optical character recognition, noised images, radial neighborhood method.

Сведения об авторе



Михайлов Иван Александрович, 1987 года рождения. В 2009 году окончил Ярославский государственный университет им. П.Г. Демидова (ЯрГУ) по специальности «Математическое обеспечение и администрирование информационных систем». В настоящее время обучается в аспирантуре факультета информатики и вычислительной техники ЯрГУ. **E-mail:** waskot@mail.ru. Область научных интересов: обработка и распознавание изображений.

Ivan Aleksandrovich Mikhaylov (b. 1987) graduated from the P.G. Demidov Yaroslavl State University (YSU) in 2009, majoring in Mathematical Support and Management of Information Systems. He is presently a post-graduate student at Computer Science Faculty of YSU. **E-mail:** waskot@mail.ru. His research interest is currently focused on image processing and recognition.

Поступила в редакцию 7 мая 2010 г.