Модел. и анализ информ. систем. Т.**14**, №**4** (2007) 7–12

УДК 519.68:[681.5137+612.8.001.57+007.51/52]

Об одном методе распознавания изображений

Михайлов И. А.
Ярославский государственный университет
150 000, Ярославль, Советская, 14

получена 12 марта 2007

Аннотация

Рассматривается задача распознавания изображений алфавитно-цифровых символов. Предлагается метод распознавания, основанный на введенной в работе метрике. Приведены результаты экспериментов на случайных изображениях и сравнение с метрикой Хаусдорфа.

Введение

В этой работе рассматривается задача распознавания черно-белых растровых изображений алфавитно-цифровых символов, заданных матрицами из нулей и единиц соответствующего размера. Изображения: как образцы, так и тесты — могут быть различного размера; для приведения их к одному размеру используется масштабирование.

Работа содержит четыре раздела.

В первом разделе приводятся постановка задачи, обозначения, применяемые в статье, и понятие метрики на изображениях.

Во втором разделе описывается стандартный подход к решению задачи распознавания, использующий метрику Хаусдорфа для определения схожести изображений.

В третьем разделе дается описание другого метода распознавания, получившего название метода срезов.

В четвертом разделе описываются четыре модели искажения изображений: две из них накладывают шумовые помехи на изображение-образец, другие две искажают контур символа-образца. Для каждой модели строится 5000 тестовых изображений.

Далее на подготовленном наборе тестовых изображений проводится эксперимент с целью сравнения описанных выше методов. В результате сравнения делается вывод: количество верно распознанных тестовых изображений в среднем для моделей наложения шума у метода срезов на 35% больше, чем у метода, основанного на метрике Хаусдорфа, для моделей искажения контура — на 1% меньше.

1. Постановка задачи и обозначения

В этом разделе приводятся постановка задачи, обозначения, применяемые в статье, и понятие метрики на изображениях.

1.1. Постановка задачи

Рассматривается задача распознавания изображений цифровых символов. Будем работать только с растровыми черно-белыми изображениями. Пусть задан алфавит изображений-образцов и одно тестовое изображение. Размеры изображений: как образцов, так и теста — могут быть различны. Задача состоит в том, чтобы для тестового изображения найти наиболее похожее на него изображение-образец из заданного алфавита. Степень схожести определяется метрикой на множестве изображений.

1.2. Основная трудность

Главная проблема в задаче распознавания изображений — это наличие помех в тестах. Помехи могут быть разделены на две категории:

- создающие шум вокруг символа и удаляющие его части;
- изменяющие контур самого символа.

Это разделение условно, так как искажения, вносимые помехами второго вида, могут быть внесены помехами первого. Необходимость выделения категорий вызвана различием подходов к решению задачи распознавания. Методы, ориентированные на определенную форму символов и наличие шума, возможно, окажутся неэффективны или даже неприменимы в том случае, когда для каждого символа имеется множество различных вариантов начертания, но количество шума сравнительно невелико.

1.3. Обозначения

Как было сказано в постановке задачи, рассматриваемые нами изображения являются растровыми, то есть каждое из них представляет собой множество пикселей — узлов прямоугольной решетки.

Пусть I — некоторый черно-белый растр размеров $m \times n$. Будем считать, что наименьшие x- и y-координаты его пикселей равны 1. Сопоставим изображению матрицу $A = (a_{ij})$ размера $n \times m$ по правилу: $a_{ij} = 1$, если пиксель изображения I с координатами (j, i) отличен от фона, или $j = 1$. В противном случае $a_{ij} = 0$. Далее будем задавать изображения описанными матрицами.

1.4. Сведение задачи распознавания изображений к задаче сравнения двух изображений

Во всей работе решение поставленной задачи будем полагать наиболее очевидным: сравнение теста с каждым образцом из алфавита и выбор наиболее близкого из них. Поэтому при описании различных подходов к задаче распознавания сосредоточимся на алгоритме сравнения двух изображений. В случае использования метрики (ее определение будет дано в следующем пункте) выходом может быть значение метрики, тогда наиболее близкий образец тот, для которого это значение минимально.

1.5. Определение метрики

Один из стандартных подходов к задаче распознавания — это введение и использование метрики на множестве изображений. Метрика — это функция, отображающая множество пар изображений на множество действительных чисел. Она должна обладать следующими свойствами:

- расстояние (значение метрики) между двумя изображениями неотрицательно и равно нулю тогда и только тогда, когда изображения совпадают;
- свойство симметричности — расстояние от изображения A до изображения B равно расстоянию от B до A;
- неравенство треугольника — сумма расстояний от A до B и от B до C не меньше, чем расстояние от A до C.

Иными словами, метрика на изображениях — это мера различия (или разности) между ними.

2. Метод, основанный на метрике Хаусдорфа

В этом разделе описывается стандартный подход к решению задачи распознавания, использующий метрику Хаусдорфа для определения схожести изображений. Раздел включает в себя определение метрики Хаусдорфа и описание простейшего алгоритма сравнения изображений, использующего эту метрику.

2.1. Определение метрики Хаусдорфа

На практике достаточно часто применяется сравнение изображений с помощью метрики Хаусдорфа [1]. Метрика Хаусдорфа определяется на множестве всех непустых компактных подмножеств плоскости. Теперь дадим ее определение.

Отклонением множества из одной точки $\{x'_0\}$ от компактного множества G называется

$$\sigma(x'_0, G) = \min_{y' \in G} ||x'_0 - y'||,$$

где $\min ||x'_0 - y'||$ — евклидово расстояние между точками x'_0 и y'.

Отклонением множества G_1 от G_2 называется

$$\sigma(G_1, G_2) = \max_{x' \in G_1} \sigma(x', G_2).$$

Метрикой Хаусдорфа называется

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

2.2. Применение метрики Хаусдорфа в задаче распознавания изображений

Применение метрики Хаусдорфа в задаче распознавания изображений возможно, если научиться сопоставлять изображениям компактные множества. Будем сопоставлять, используя следующие правила:

- Каждому изображению соответствует матрица (см. пункт 1.3);
- Каждую такую матрицу очевидным образом можно соотнести с дискретным множеством точек на плоскости.

2.3. Описание алгоритма масштабирования изображений

Входные данные:

- матрица изображения (a_{ij}) размера $n_1 \times m_1$;
- размер, к которому нужно привести матрицу, — $n_2 \times m_2$.

Выходные данные:

- матрица изображения (b_{ij}) размера $n_2 \times m_2$.

Опишем теперь сам алгоритм.

1. На первом этапе строится вспомогательная матрица (c_{ij}) размера $n_2 \times m_1$ по следующему правилу: $c_{ij} = a_{kj}$, где $k =]i \cdot n_1/n_2[$, $1 \leq i \leq n_2$, $1 \leq j \leq m_1$.
2. На втором этапе строится матрица (b_{ij}) по правилу: $b_{ij} = c_{ik}$, где $k =]j \cdot m_1/m_2[$, $1 \leq i \leq n_2$, $1 \leq j \leq m_2$.

Матрица (b_{ij}) возвращается в качестве результата.

2.4. Описание простейшего алгоритма сравнения

Приведем теперь описание простейшего алгоритма сравнения на основе этой метрики.

Входные данные:

- Матрицы изображений теста и образца.

Выходные данные:

- Действительное число — значение метрики.

Так как изначально матрицы образца символа и теста обладают различными размерами, то на первом этапе тест приводится к размеру образца с помощью алгоритма масштабирования, описанного в предыдущем пункте.

На втором этапе производится поиск такого взаимного расположения изображений теста и образца, чтобы значение метрики Хаусдорфа между ними было минимальным. Поиск осуществляется следующим образом.

Вначале матрицы изображений сдвигаются так, чтобы совместились их центры тяжести. Вычисляется значение метрики. Если при уменьшении на 1 координаты x теста значение метрики между ним и образцом стало меньше, то продолжаем уменьшать координату x на 1, в противном случае увеличиваем ее на 1. Изменение (уменьшение или увеличение) x координаты продолжается до тех пор, пока значение метрики уменьшается. Теперь координата x фиксируется, и производится аналогичный поиск по столбцам, то есть изменяется координата y. После вычисления обеих координат вектора сдвига вычисляется значение метрики, которое возвращается в качестве результата.

Данный алгоритм будет использоваться в эксперименте, описанном в четвертом разделе.

3. Метод срезов

В этом разделе дается описание другого метода распознавания, названного методом срезов. Раздел содержит определения бесконечного вектора, линейного образа и определения используемых метрик.

3.1. Определение бесконечного вектора и метрики на множестве бесконечных векторов

Под **бесконечным вектором** будем понимать вектор, обладающий такими свойствами:

1. Размерность вектора бесконечна, то есть вектор содержит бесконечно много координат;
2. Для каждого бесконечного вектора V существует такое $n \geq 0$, что $v_i \geq 0$ для $i \leq n$ и $v_i = 0$ для $i > n$. Заметим, что значение n допускается равным 0.

Пусть V', V'' — бесконечные векторы.

Метрикой на множестве бесконечных векторов будем называть

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

Легко видеть, что выполняются соответствующие свойства метрики:

$$\rho(V_1, V_2) \geq 0, \quad \rho(V_1, V_2) = 0 \Leftrightarrow V_1 = V_2, \tag{1}$$

$$\rho(V_1, V_2) = \rho(V_2, V_1), \tag{2}$$

$$\rho(V_1, V_2) + \rho(V_2, V_3) \geq \rho(V_1, V_3). \tag{3}$$

3.2. Определение линейного образа и метрики на множестве линейных образов

Пусть имеется матрица изображения $A = (a_{ij})$ размера $n \times m$.

Для каждого $i : 1 \leq i \leq n$ определим множество индексов $S(A, i)$

$$S(A, i) = \{s_1, s_2, \ldots, s_k\} = \{j : a_{ij} = 1 \wedge (j = 1 \vee a_{i,j-1} = 0), 1 \leq j \leq m\}.$$

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

Вектором разности $V(A, i, N)$, где $A = (a_{ij})$ — матрица изображения размера $n \times m$, $1 \leq i \leq n$, N — натуральное, будем называть бесконечный вектор действительных чисел, построенный следующим образом:

$$v_p(A, i, N) = \begin{cases} (s_p(A, i) - s_{p-1}(A, i)) \cdot N/m, & 2 \leq p \leq |S(A, i)|, \\ 0, & |S(A, i)| < p < \infty, \end{cases}$$

где $v_p(A, i, N)$ — i -й элемент вектора $V(A, i, N)$.

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

$$L(A, N) = \{l_p : l_p = V(A,]n \cdot p \,/\, N[\,, N), 1 \leq p \leq N\}.$$

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

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

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

где $l_i(A_1, N)$ — i -й элемент мультимножества $L(A_1, N)$, $l_i(A_2, N)$ — i -й элемент мультимножества $L(A_2, N)$.

Легко проверить, что выполняются соответствующие свойства метрики:

1. (a) $\rho(L(A_1, N), L(A_2, N)) \geq 0$,
   (b) $\rho(L(A_1, N), L(A_2, N)) = 0 \Leftrightarrow L(A_1, N) = L(A_2, N)$.

   Доказательство следует из (1).

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

3. $$\rho(L(A_1, N), L(A_2, N)) + \rho(L(A_2, N), L(A_3, N)) \geq \rho(L(A_1, N), L(A_3, N)).$$

   Доказательство следует из (3).

3.3. Описание алгоритма сравнения

Входные данные:

- Матрицы изображений теста и образца.

Выходные данные:

- Действительное число - значение метрики.

На первом этапе этого алгоритма происходит построение линейных образов теста и образца, на втором вычисление значения метрики между ними. Полученное значение возвращается в качестве результата.

Данный алгоритм будет использоваться в эксперименте, описанном в четвертом разделе.

4. Описание моделей искажения изображений и результаты работы алгоритмов

Теперь приведем результаты работы алгоритмов сравнения на наборе тестовых изображений. Алфавит образцов включает в себя 10 изображений символов ’0’—’9’ высотой 14 пикселей, полученных с помощью шрифта Times New Roman. Тестовые изображения строились с использованием 4 алгоритмов искажения: в двух из них на изображение-образец из алфавита наносились шумовые помехи и еще в двух искажался контур символа-образца. Эксперимент с каждым видом искажения включал в себя 500 тестовых изображений для каждого из 10 образцов. В алгоритме метода срезов частота сечения бралась равной 100, а в алгоритме метода, основанного на метрике Хаусдорфа, в качестве расстояния между точками на плоскости использовалось евклидово расстояние.

Опишем все алгоритмы по порядку, используя при этом сокращенные обозначения: метод, основанный на метрике Хаусдорфа, — Hausdorf, метод срезов — Slices. В таблицах верхняя строка содержит названия методов, нижняя — соответствующее количество верно распознанных тестов (в процентах).

1. Нанесение на изображение-образец одиночных пикселей. Расположение каждого пикселя выбиралось случайно. Цвет всех пикселей совпадал с цветом контура. Их количество варьировалось от 1 до 50. Результаты показаны в таблице 1.

Таблица 1.

Hausdorf	Slices
34,92%	83,22%

2. Нанесение на изображение-образец 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 совпадал с цветом фона, поэтому они стирали части контура символа; цвет остальных совпадал с цветом символа. Результаты показаны в таблице 2.

Таблица 2.

Hausdorf	Slices
65,0%	85,96%

Алгоритмы искажения контура, описанные далее, на входе получают последовательность узлов контура (или цепь), построенную вручную. Узел представляет собой координаты некоторой точки контура. В среднем цепь содержит около 25 узлов, распределенных по всему контуру. В цепь символа-образца вносились изменения, затем она отображалась в растр.

3. Сдвиг каждого узла цепи на случайный вектор. При этом координаты изменялись не более, чем на 20% от размеров изображения. Результаты показаны в таблице 3.

Таблица 3.

Hausdorf	Slices
66,74%	67,8%

4. Разбиение контура символа на несколько сегментов, поворот каждого сегмента относительно узла, случайно выбранного в данном сегменте. Угол поворота лежал в пределах от $-\pi/6$ до $\pi/6$. Результаты показаны в таблице 4.

Таблица 4.

Hausdorf	Slices
66,32%	63,42%

Если измерять эффективность методов соответствующим количеством верно распознанных тестовых изображений, то можно сделать следующие выводы:

1. Для первой модели искажения метод срезов эффективнее на 48,3%.
2. Для второй модели искажения метод срезов эффективнее на 20,96%.
3. Для третьей модели искажения метод срезов эффективнее на 1,06%.
4. Для четвертой модели искажения метод срезов менее эффективен, чем метод, основанный на метрике Хаусдорфа, на 2,9%.

Таким образом, метод срезов более приспособлен для распознавания зашумленных черно-белых изображений, чем метод, основанный на метрике Хаусдорфа.

Список литературы

1. Скворцов, В.А. Примеры метрических пространств /В.А.Скворцов //Библиотека "Математическое просвещение". — 2001. — Вып. 9.

An Algorithm for the Recognition of Images

Mikhailov I.A.

A new algorithm for the recognition of digits is proposed. The result of computer experiments is described.