

РАСПОЗНАВАНИЕ СИМВОЛОВ НА БОЛЬШЕФОРМАТНЫХ РАСТРОВЫХ ИЗОБРАЖЕНИЯХ ДОКУМЕНТОВ СО СЛАБО ФОРМАЛИЗОВАННЫМ ОПИСАНИЕМ ОБЪЕКТОВ

Д. Ю. Васин¹, М.А. Редькин¹
dm04@list.ru|maxop151@mail.ru

¹ ИИТММ Нижегородского Государственного Университета им. Н.И. Лобачевского, Нижний Новгород, Россия

В докладе рассмотрены алгоритмы формирования эффективных, устойчивых к искажениям признаков и решающих правил распознавания объектов, представленных на большеформатных растровых изображениях со слабо формализованным описанием объектов. Рассмотрен вопрос влияния метода бинаризации исходного растрового описания на качество последующего автоматического распознавания объектов.

Ключевые слова: растровые изображения, моментные характеристики изображения, низкоуровневая структурированная модель описания бинарного растра, дискриминантные признаки, распознавание.

CHARACTER RECOGNITION ON LARGE-SCALE RASTER IMAGES OF DOCUMENTS WITH WEAKLY FORMALIZED DISCRIPTION OF OBJECTS

D. Y. Vasin¹, M. A. Redkin¹
dm04@list.ru|maxop151@mail.ru

¹ ITMM Nizhny Novgorod State University. N.I. Lobachevsky, Nizhny Novgorod, Russia

In the report the algorithms of building effective, resistant to distortion characteristics and decision rules of recognition of objects presented on large format raster images with weakly formalized description of objects. Considers influence of the method of binarization of the original raster descriptions on the quality of the subsequent automatic recognition of objects.

Keywords: bitmap, torque characteristics of the image, low-level structured model describing the binary image, discriminant features, recognition.

1. Введение

В настоящее время актуальной остается проблема создания электронных архивов большеформатных графических документов (топографических и морских навигационных карт, конструкторской документации, технических чертежей, электрических схем и др.). При этом источником, как правило, являются графические изображения на бумажных носителях. С целью автоматизации технологии ввода графических документов (ГД) в последние годы предложены различные информационные технологии на основе эвристических процедурных методов, а также на базе методов распознавания с учителем, которые эффективны для ограниченного набора объектов с жестким ограничением на размер и ориентацию. Однако, значительный класс большеформатных ГД выполняется ручным способом со слабо формализованным начертанием объектов, произвольной ориентацией, произвольными размерами символов, наличием большого количества межобъектных наложений, примыканий и пересечений (рис. 1).

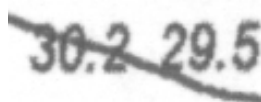


Рис. 1. Пример изображения ГД со слабо формализованным начертанием объектов.

Все это существенно снижает качество автоматического распознавания объектов, что, в конечном итоге, сказывается на эффективности всей технологической цепи создания электронных архивов большеформатных ГД.

Одной из основных проблем, возникающих при решении задачи распознавания образов и существенно влияющих как на эффективность распознавания, с одной стороны, так и на его качество, с другой, является выбор системы эффективных признаков классификации.

В работе [1] рассматривалось формирование решающих правил в задаче распознавания символов, находящихся в стандартной горизонтальной ориентации, но допускающих растяжение/сжатие по горизонтали/вертикали отдельных символов, на базе низкоуровневой структурированной модели описания бинарного растрового изображения (БРИ) [2]. Настоящая работа является продолжением исследований по созданию эффективных систем распознавания объектов по растровым данным. В частности, рассмотрен вопрос влияния различных методов бинаризации исходного полутонового растрового изображения на конечный результат распознавания объектов, расширен класс рассматриваемых документов.

Очевидно, что качество формируемых дискриминантных признаков, определяемых по низкоуровневой структурированной модели описания БРИ, напрямую зависит от качества выполнения операций предобработки и бинаризации исходного полутонового растрового изображения. Дело в том, что неудачный выбор порога бинаризации приводит к появлению различных шумов, которые вносят значительный вклад в снижение качества итогового результата распознавания. Среди основных помех можно выделить наличие на распознаваемых по БРИ объектах: неправомерных разрывов или слипаний, появление изолированных пустот внутри объектов, наличие «снега» — изолированных черных объектов малого размера, изрезанность краев объектов.

2. Моментные характеристики изображения

В работе [1] были введены моментные характеристики, нашедшие широкое применение во многих задачах обработки цифровых изображений [3].

Момент – характеристика контура изображения, объединённая (суммированная) со всеми пикселями контура. Момент $m_{p,q}$ для непрерывного изображения $f(x,y)$ определяется как:

$$m_{p,q} = \iint x^p y^q f(x,y) dx dy,$$

в случае дискретного растрового изображения:

$$m_{p,q} = \sum \sum x^p y^q f(x,y),$$

где p, q – мощность, на которой соответствующий компонент взят в сумме с другими отображенными. Если функция $f(x,y)$ рассматривается как бинарное изображение, тогда момент m_{00} вычисляет площадь чёрных пикселей этого изображения.

Для вычисления характеристик, инвариантных относительно сдвига, необходимо центрировать относительно средних значений $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$ –

центр тяжести фигуры.

Центральный момент определяется как:

$$\mu_{p,q} = \sum \sum (x - \bar{x})^p (y - \bar{y})^q f(x,y).$$

Доказано, что все центральные моменты выражаются через обычные, по формуле:

$$\mu_{p,q} = \sum_{i=0}^p \sum_{j=0}^q C_p^i C_q^j (-\bar{x})^{p-i} (-\bar{y})^{q-j} m_{ij},$$

где $C_n^k = \frac{n!}{k!(n-k)!}$.

С помощью центральных моментов определяются характеристики, инвариантные к повороту изображения (моментные инварианты). Моментные инварианты – это признаки изображения, построенные на основе степенных моментов и описывающие силуэт некоторого объекта. В соответствии со своим названием данные признаки являются инвариантными к аффинным преобразованиям изображения. Они стали важнейшим инструментом для распознавания образов независимо от их особого положения, ориентации, угла обзора и других изменений, также они нечувствительны к различным геометрическим преобразованиям (сдвиг, поворот, масштабирование) [4]. Это определяет их эффективность при использовании в качестве признаков в задаче обнаружения и распознавания на изображении объектов неизвестной ориентации.

Имея набор характеристик, можно определить следующие аффинные преобразования моментных инвариантов, которые инвариантны относительно сдвигов, поворотов, а также растяжений и сжатий:

$$I^1 = \frac{\mu_{20} + \mu_{02}}{m_{00}^2}, \quad I^2 = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{m_{00}^4},$$

$$I^3 = \frac{(\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{30})^2}{m_{00}^5},$$

$$I^4 = \frac{(\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2}{m_{00}^5}.$$

Значения $I^1 - I^4$ используются в качестве признаков классификации.

3. Низкоуровневая структурированная модель описания БРИ

Полное описание модели приведено в [2]. Пусть БРИ представлено в пиксельном формате, при этом функция яркости пикселей имеет вид:

$$f(x,y) = \begin{cases} 255, & \text{если пиксель фоновый,} \\ 0, & \text{если пиксель не фоновый.} \end{cases}$$

Непроизводным элементом модели является *штрих* – одномерный кластер связанных пикселей заданного цвета, представленный набором чисел $S=(t, N, K, C)$, где: t – номер строки раstra; N, K – координаты начала и конца связанной одноцветной пиксельной области вдоль линии раstra; C – код цвета связанных пикселей, образующих штрих. Для бинарных изображений штрих описывается тройкой типа $S=(t, N, K)$, где: t – номер строки раstra; N, K – координаты начала и конца штриха, определяемые соотношениями:

$$\begin{aligned} N &= x \wedge f(x-1, y) \wedge \bar{f}(x, y), \\ K &= x \wedge \bar{f}(x, y) \wedge f(x+1, y). \end{aligned} \quad (1)$$

Совокупность штрихов для заданной строки раstra t определяет строку штрихов $SH=\{S_i\}(i=1,2,...,m)$, где S_i – штрих. Описание раstra с помощью набора строк штрихов является однозначным. Отметим, что штриховой формат представления и хранения растровых данных является в 10–20 раз более компактным по сравнению с пиксельным.

Топологические свойства штриха $S=(t, N, K)$ определяются через его связность со штрихами смежных строк. Два штриха S_t и S_{t-1} смежных строк t и $t-1$ связаны, если $(N_{t-1} \geq N_t) \wedge (N_{t-1} \leq K_t) \vee (K_{t-1} \geq N_t) \wedge (K_{t-1} \leq K_t)$. Связный штрих – это набор $H = \{S, sw_p, sw_s\}$, sw_p, sw_s – количество связанных штрихов со штрихами смежных строк, а t – номер текущей строки штрихов раstra. Тогда $sw_p=0$, если в предыдущей строке $p=t-1$ нет ни одного штриха, для которого выполняется условие:

$$(N_p \geq N_t) \wedge (N_p \leq K_t) \vee (K_p \geq N_t) \wedge (K_p \leq K_t) \quad (2)$$

иначе, $sw_p = \mu_p, \mu_p$ – кратность выполнения условия (2), т.е. количество связанных штрихов предыдущей строки. Аналогично, $sw_s=0$, если в следующей строке $s=t+1$ нет ни одного штриха, для которого выполняется условие:

$$N_s \geq N_t \wedge N_s \leq K_t \vee K_s \geq N_t \wedge K_s \leq K_t \quad (3)$$

иначе, $sw_s = \mu_s, \mu_s$ – кратность выполнения условия (3), т.е. количество связанных штрихов последующей строки. Здесь N, K – координаты начала и конца исследуемого на связность штриха, N_t, K_t – координаты начала и конца текущего штриха предыдущей смежной строки штрихов, N_s, K_s – координаты начала и конца текущего штриха следующей смежной строки штрихов.

По суперпозиции значений sw_p и sw_s можно ввести следующую классификацию графических ситуаций. В текущей строке:

штрих изолированный (Ши): $sw_p = 0 \wedge sw_s = 0$; штрих начала растрового объекта (Шн): $sw_p = 0 \wedge sw_s = 1$; штрих конца растрового объекта (Шк): $sw_p = 1 \wedge sw_s = 0$; штрих слияния растровых объектов (Шс): $sw_p > 1$; штрих расщепления растровых объектов (Шр): $sw_s > 1$; штрих слияния и расщепления растровых объектов (Шср): $sw_p > 1 \wedge sw_s > 1$.

Алгоритм нахождения указанных особых штрихов приведен в [2].

Растровый простой объект (РПО) – кластер связанных штрихов, не содержащий графических ситуаций типа слияния и расщепления, то есть для любого штриха кластера выполняется условие $sw_p < 2 \wedge sw_s < 2$.

Растровый составной объект (РСО) – кластер связанных штрихов, для каждого штриха которого выполняется условие: $sw_p > 0 \vee sw_s > 0$, и для любых двух элементов (штрихов или пикселей) найдется хотя бы одна

соединяющая их пиксельная траектория, состоящая из связанных смежных элементов.

4. Постановка задачи

Особенностью ГД со слабо формализованным описанием объектов является то, что, несмотря на отклонение от нормативных требований к изображению объектов, они обладают некоторой стилизованной формой представления. Это означает, что возможно создание набора эталонов объектов, основанных на работе с низкоуровневыми моделями графических изображений, максимально адаптированных для данного класса ГД, с учетом сильной зависимости эффективности признаков классификации от искажений входных данных. В дальнейшем, по мере поступления новых документов, этот набор может соответствующим образом дополняться. Вместе с тем, требует внимания выбор наиболее эффективного метода бинаризации исходного растрового изображения, с целью минимизации возникающих в процессе бинаризации искажений.

В качестве исходных документов были взяты 24-битные растровые изображения текстов на английском языке, а также планшеты гидрографической съемки (рис. 2).

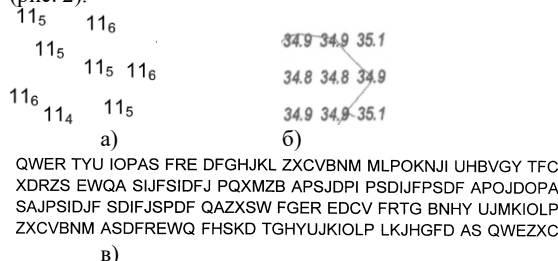


Рис. 2. Примеры исходных изображений: а) планшет гидрографической съемки хорошего качества; б) планшет гидрографической съемки среднего качества; в) отсканированный текст.

5. Признаки классификации

Для каждого объекта будем фиксировать следующие признаки:

1. Отношение сторон описанного вокруг символа прямоугольника со сторонами, параллельными осям координат (отношение аспекта).
2. Отношение площади символа к площади описанного вокруг символа прямоугольника со сторонами, параллельными осям координат (процент заполненности).
3. Значения моментных инвариантов $I^1 - I^4$.
4. Количество внутренних областей (ВО).
5. Количество особых штрихов: Шн, Шк, Шс, Шр.
6. Количество РПО.

Рассмотренный набор признаков не является окончательным и может быть расширен.

Число внутренних областей позволяет сразу разделить все объекты на 3 группы, т.е. объекты, у которых: ВО = 0 (К, X, S, 1, 2 и др.), ВО = 1 (R, O, P, 4, 6, 9 и др.) и ВО = 2 (B, 8). Учтем, что «В» – единственный среди символов английского алфавита, а цифра 8 – единственная цифра среди арабских, имеющие ВО=2.



Рис. 3. РПО, выделенные на символах.

Дальнейшее распознавание объекта можно выполнять в пределах этих групп.

Число ВО – достаточно эффективный признак, отношение аспекта, и процент заполненности, более

Количество РПО, а также их параметры, довольно хорошо характеризуют каждый символ (рис. 3) Символ «А» – 6 РПО, символ «К» – 5 РПО.

устойчивы к различным искажениям распознаваемых данных.

6. Преобразование исходного РИГД в оттенки серого

Учитывая, что источником данных являются различные ГД, то основными цветами их представления будут белый (цвет фона) и чёрный (цвет объектов). Однако, изначальное сканирование исходного документа в режиме монохромного, может привести к существенным искажениям, т.к. довольно малое количество информации представлено чистыми черным и белым цветами. Для минимизации подобных искажений требуется постепенное уменьшение глубины цвета путём преобразования цвета всех пикселей изображения к требуемому диапазону. Изначально сканируем ГД с глубиной цвета 24 бит/пиксель. Данный режим сканирования является избыточным, но позволяет выполнить последующую программную бинаризацию РИГД с минимизацией возможных искажений. На первом этапе осуществляем конвертацию полноцветного РИГД в оттенки серого. Это осуществляется путем применения к каждому пикселю исходного РИГД преобразования вида:

$$Y = 0.299 R + 0.587 G + 0.144 B, \quad (4)$$

в котором R, G, B – красная, зелёная и синяя компонента цвета исходного РИГД соответственно, а Y – полученная яркость в оттенках серого.

6.1. Бинаризация РИГД

Нами было исследовано влияние различных методов бинаризации РИГД на конечный результат распознавания объектов.

Бинаризация РИГД методом Оцу. В методе Оцу для расчета порогового значения используется нормализованная гистограмма изображения [5]. Гистограмма строится по значениям $p_i = n_i / N$ где N – общее число пикселей изображения, n_i – число пикселей с уровнем яркости i , ($0 \leq i \leq L$). Гистограмма позволяет разделить все пиксели изображения на два класса: «полезные» и фоновые, с помощью порогового значения уровня яркости t (t – целое значение $0 \leq t \leq L$). Каждому классу соответствуют относительные частоты ω_0 и ω_1 :

$$\omega_0(t) = \sum_{i=1}^t p_i, \quad \omega_1(t) = \sum_{i=t+1}^L p_i = 1 - \omega_0(t).$$

Средние уровни для каждого из двух классов изображения: $\mu_0(t) = \sum_{i=1}^t \frac{ip_i}{\omega_0}$, $\mu_1(t) = \sum_{i=t+1}^L \frac{ip_i}{\omega_1}$.

Суть метода Оцу заключается в том, чтобы выставить порог между классами так, чтобы каждый из них был наиболее «плотным». Метод Оцу определяет порог, уменьшающий дисперсию внутри класса, которая определяется как взвешенная сумма дисперсий двух классов: $\delta_{\omega}^2(t) = \omega_1(t) \delta_1^2(t) + \omega_2(t) \delta_2^2(t)$.

Далее вычисляется максимальное значение оценки качества разделения изображения на две части:

$$\eta(t) = \max \left(\frac{\sigma_{kl}^2(t)}{\sigma_{\omega}^2(t)} \right), \quad \sigma_{kl}^2(t) = \omega_0(t) \omega_1(t) (\mu_1(t) - \mu_0(t))^2$$

– межклассовая дисперсия.

На рис. 4 приведен фрагмент удачно бинаризованного изображения планшета гидрографической съемки по методу Оцу.

Результат обработки имеет и явный недостаток: при обработке объектов, имеющих различные размеры, но близкие цвета, в результате бинаризации Оцу получаются темные пятна (рис. 5).

117 117 116
117 116 116
117 117

Рис. 4. Хороший результат бинаризации Оцу.

35.5 35.5 35.6
35.4 35.6 35.6

Рис. 5. Неудачный выбор порога бинаризации Оцу.

Бинаризация с двойным ограничением. В этом случае пиксели, которые входят в диапазон $[t1, t2]$ – становятся черными, которые не входят – белыми [6].

$$F(x, y) = \begin{cases} 255, & f(x, y) < t1; \\ 0, & t1 \leq f(x, y) \leq t2; \\ 255, & f(x, y) > t2; \end{cases}$$

Пороговые значения яркости пикселей оценивались по гистограмме.

35.3 35.2 35.2
35.3 35.4 35.6
35.8 35.8 35.9

Рис. 6. Результат бинаризации с двойным ограничением.

На рис. 6 приведен фрагмент изображения планшета гидрографической съемки, обработанный по методу бинаризации с двойным ограничением. Сравнивая результаты бинаризации (рис. 5 и рис. 6)

видно, что на данном классе документов метод бинаризации с двойным ограничением устраняет недостатки бинаризации по методу Оцу.

7. Этап подготовки изображения

На данном этапе изображения разбиваются на связанные области чёрных пикселей (компоненты), каждая из которых является тем или иным объектом с присвоенной ему уникальной меткой и описанной вокруг нее прямоугольной областью. Отметим, что данная операция является наиболее затратной.

Во время выполнения данного этапа производится объектовая фильтрация БРИ, перевод БРИ в штриховой формат, а также выделение внутренних областей внутри компонент.

7.1. Алгоритм объектовой фильтрации исходного бинарного изображения

Заметим, что в целом, признаки, определяемые по низкоуровневой модели БРИ, не являются достаточно устойчивыми. Различного рода помехи на БРИ ведут к тем или иным изменениям штрихового описания, в результате чего могут быть выделены лишние штрихи расщепления, слияния, РПО а также внутренние области. Следовательно, эффективное использование штрихового описания для автоматического распознавания объектов требует выполнения предварительной фильтрации БРИ. Нами был разработан алгоритм объектовой фильтрации БРИ, суть которого в следующем. На вход алгоритма подается БРИ с распознаваемыми объектами. Входными параметрами являются код цвета удаляемой помехи и размеры описанного вокруг помехи прямоугольника со сторонами, параллельными осям координат, а также процент его заполненности пикселями цвета помехи. Работа алгоритма основана на выделении по исходному БРИ связанных компонент, состоящих из пикселей с цветом, равным заданному, с последующим удалением (перекрашиванием в альтернативный цвет) тех найденных связанных областей, которые по своим геометрическим характеристикам удовлетворяют параметрам помехи. Важным достоинством данного алгоритма является тот факт, что указанная фильтрация не вносит никаких искажений в те связанные области, которые не распознаются алгоритмом как шумовые.

7.2. Алгоритм конвертации БРИ в штриховое представление

На вход алгоритма подается БРИ, представленное в пиксельном формате. На выходе алгоритма формируется БРИ, представленное в

штриховом формате. Конвертация осуществляется построчно, используя соотношения (1).

7.3. Алгоритм подсчета числа внутренних областей

Пусть имеется изображение, представленное в штриховом формате, с объектом, представленным единственной связанной компонентой. В процессе работы на изображении находятся “особые” точки, хранящиеся в списке S .

1. Просматривая штриховой файл, находим штрихи расщепления.
2. Пусть штрих расщепления находится в столбце i . Просматриваем все связанные с ним штрихи в столбце $(i+1)$, для каждого $H = (t, N, K)$ добавляем в список S особые точки $(i+1, N-1)$, $(i+1, K+1)$.
3. Просматривая все особые точки (x, y) из списка:
 - 3.1. Если координаты точки выходят за рамки матрицы P , то удаляем её из списка как несуществующую.
 - 3.2. Запускаем поиск в ширину со стартом в (x, y) , при этом $цвет1$ – белый, $цвет2$ – любой дополнительный цвет (отличный от черного и белого).
 - 3.3. Проверяем все особые точки из S , следующие по списку за точкой (x, y) . Если они окрашены в $цвет2$, то удаляем их как лишние.

3.4. Просматриваем граничные точки всего изображения $[P_{ij}]$, $[P_{Nj}]$, $[P_{i1}]$, $[P_{iM}]$, $i = 1, \dots, N$; $j = 1, \dots, M$. Если хотя бы одна из них будет закрашена в $цвет2$, значит точка (x, y) не принадлежит внутренней области объекта (замкнутая область белого цвета). Удаляем её из списка S . Количество элементов в списке S равно количеству внутренних областей в объекте.

7.4. Алгоритм построения классификатора

На вход алгоритма поступают все эталоны для каждого класса, фиксируем: для каждого признака принимаемые значения, а также математическое ожидание и доверительный интервал их значений.

7.5. Алгоритм автоматического распознавания символов

На вход алгоритма поступает вектор признаков связанного объекта. На выходе распознанный символ алфавита базы эталонов или пустой символ. В алгоритме введено понятие штрафа – положительная величина, пропорциональная отношению расстояния до ближайшей границы интервала к длине этого интервала. Для признаков, сформированных по штриховому описанию: если значение признака меньше значения нижней границы интервала, то сразу начисляется штраф, больший порогового значения и класс убирается из дальнейшего рассмотрения.

Изначально формируем массив размером по количеству классов. В цикле по номеру признака:

1. Для каждого класса в соответствующую ему ячейку массива добавляется штраф, если значение признака не входит в допустимый интервал для данного класса.
2. Если штраф объекту для класса превосходит порог, то класс исключается из дальнейшего рассмотрения.
3. Если в рассмотрении остался один единственный класс, то соответствующий ему символ возвращается как результат распознавания.
4. Если были исключены все классы, то в качестве результата работы алгоритма возвращается пустой символ.

Если после завершения цикла осталось несколько классов, штраф которых не превосшел порог, то в качестве результата возвращается символ, соответствующий классу с минимальным штрафом.

7.6. Алгоритм автоматической сборки надписей

На вход алгоритма поступает набор описаний распознанных объектов. На выходе описание сформированной надписи в структуре выходной

тематической базы данных. Сборка осуществляется применением методов кластеризации. Из списка выбирается очередной непомеченный объект, который объявляется ядром нового кластера (надписи). Данный объект помечается уникальной меткой. Просматривается ε -окрестность относительно центра тяжести ядра кластера. Если в указанную ε -окрестность попадает центр тяжести очередного непомеченного объекта, то оценивается угол наклона вектора, соединяющего центр тяжести ядра кластера и объекта-кандидата на добавление в кластер. Если отклонение вектора от горизонтали не более, чем на заданную величину, то найденный объект присоединяется к кластеру, помечается меткой стартового объекта и для объектов, принадлежащих кластеру, высчитывается новый центр тяжести ядра, как среднее значение координат центров объектов, составляющих кластер. Поиск очередных объектов, принадлежащих формируемому кластеру, осуществляется относительно нового центра тяжести. Процесс повторяется до тех пор, пока на очередной итерации к кластеру не будет добавлено ни одного нового объекта. В результате, сформированный кластер может состоять либо из одиночного объекта, либо из нескольких объектов, расположенных на исходном документе вдоль строки. Процесс поиска новых кластеров завершается, если на входе нет непомеченных объектов.

8. Анализ сформированного признакового пространства

По обучающему тексту на английском языке, была сформирована выборка по 5 эталонов на символ. Анализируя значения мы пришли к выводу, что значений I^1 и I^2 достаточно чтобы выделить среди всех классов группы схожих, значения I^3 и I^4 необходимы для более тонкой классификации внутри группы. В целом, значения моментных инвариантов достаточно стабильны и потому распознавание начинается с них. Данные признаки с большой долей вероятности позволяют исключить множество классов, не относящихся к рассматриваемому объекту.

На практике, после использования моментных признаков остаётся 2 – 5 возможных классов, к которым может принадлежать распознаваемый объект. Для окончательного решения об отнесении объекта к определённой группе используются признаки, полученные из штрихового описания. У них есть одно замечательное свойство – чёткая нижняя граница возможных значений. Например, у символа “М” ни при каких условиях не может быть менее трех Шк, двух Шн и Шр, одного Шс и четырех РПО. Следовательно, по нижней границе штриховых признаков можно исключать классы, для которых объект нарушает указанные признаки. Данное утверждение справедливо для любого из распознаваемых нами объектов, при условии, что они представляются связными растровыми объектами.

9. Результаты распознавания

Шрифт написания обучающего и экзаменующего английских текстов – Consolas, размер шрифта – 14 пунктов. Обучающий текст содержал 1350, экзаменующий – 1850 символов без учета пробелов. Оба документа сканировались с глубиной цвета 24 бит/пиксель, с пространственным разрешением 300 dpi. размеры полученной растровой матрицы 2500x3500 пикселей. На изображениях присутствовали только символы английского алфавита в стандартной ориентации. Изображения были последовательно переведены из полноцветных в изображения в оттенках серого, которые были бинаризованы по методу Оцу. По полученному обучающему БРИ сформирована выборка из 150 эталонов. Использование обучающего документа в качестве экзаменующего обнаружило следующие результаты распознавания: неправильно распознано – 8

объектов, неопознанных объектов нет, качество распознавания более 99%. Высокий процент объясняется тем, что специально в качестве эталонных символов были выбраны наиболее различающиеся символы внутри своих классов. В результате были получены достаточно широкие интервалы допустимых значений, а сформированного набора признаков оказалось достаточно для распознавания. Эта же база эталонов была использована для проведения распознавания на экзаменующем документе. В результате: неправильно распознано 25 объектов, неопознано – 15, качество распознавания $\approx 93 - 96 \%$.

Для распознавания цифровых надписей на планшетах гидрографической съемки была сформирована база из 50 эталонов, по 5 эталонов на каждую цифру. При этом геометрический размер символов практически не играл роли в том смысле, что для разных размеров символов дискриминантные признаки принимают значения, хоть и слегка отличающиеся друг от друга, но гарантированно попадающие в границы рассчитанных доверительных интервалов. При этом качество распознавания было не ниже 85–90%. Основная неразрешенная к настоящему времени проблема, снижающая качество распознавания – «рассыпанные» объекты-цифры и наложение на объекты-цифры тонких линейных объектов.

10. Благодарности

Работа выполнена при финансовой поддержке РФФИ – проекты №16-07-01214А и № 16-07-01198А.

11. Литература

- [1] Д.Ю. Васин, А.В. Аратский. Распознавание символов на основе инвариантных моментов графических изображений. // Труды юбилейной 25-ой Международной конференции GraphiCon 2015, 22-25 сентября 2015 г, Протвино, Парк Дракино, Московская область, стр.259 – 264.
- [2] Д.Ю. Васин, Ю.Г. Васин, В.П. Громов. Структурное описание растровых данных. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 21–23.
- [3] Dan S. Bloomberg, Gary E. Kopec and Lakshmi Dasari. Measuring document image skew and orientation. // Xerox Palo Alto Research Center. – Режим доступа: <http://www.leptonica.com/papers/skew-measurement.pdf>, свободный.
- [4] S. X. Liao, «Image Analysis by Moments,» *The Department of Electrical and Computer Engineering The University of Manitoba*, 1993.
- [5] N. Otsu. A thresholding selection method from gray-scale histogram // IEEE Transactions on System, Man, and Cybernetics, 1979. No. 9. P.62-66
- [6] M.I. Sezan, A peak detection algorithm and its application to histogram-based image data reduction //Computer Vision, Graphics, and Image Processing, 1990. Vol. 49 (1) P.36-51.

Об авторах

Васин Дмитрий Юрьевич – к.т.н., ведущий научный сотрудник Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Государственного университета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского).
E-mail: dm04@list.ru

Редькин Михаил Александрович – магистр 1 года обучения ИИТММ ННГУ им. Н.И. Лобачевского.
E-mail: maxop151@mail.ru