

## РАСПОЗНАВАНИЕ ИНДЕКСА НА КОНВЕРТАХ

### ЗАДАНИЕ:

**Разработать алгоритм и программу распознавания почтового индекса на изображении почтового конверта**

### ОСНОВНЫЕ ЭТАПЫ РЕШЕНИЯ ЗАДАНИЯ:

- 1). Преобразование цветного изображения в полутоновое.
- 2). Бинаризация с автоматическим определением порога.
- 3). Выделение связных областей бинарного изображения.
- 4). Расчет характеристик выделенных областей.
- 5). Удаление объектов не «похожих» на маркер по некоторым параметрам.
- 6). Определение маркеров индекса.
- 7). Локализация зоны индекса.
- 8). Вырезание и поворот индекса.
- 9). Выделение отдельных символов индекса.
- 10). Распознавание индекса.

## МИНИМАЛЬНЫЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

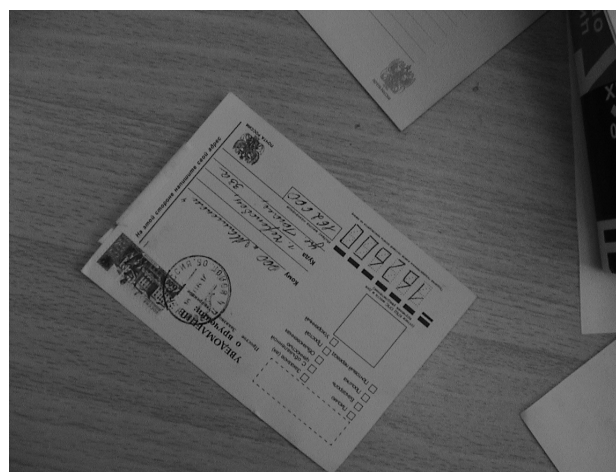
### 1). Преобразование цветного изображения в полутоновое.

Переводим изображение из цветовой системы RGB в цветовую систему YIQ [1, 3, 5, 6]:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.522 & 0.311 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Используем только яркостную составляющую Y, получаем полутоновое изображение. Т.е. преобразование можно записать в более простом виде:

$$Y=0.399 \cdot R+0.587 \cdot G+0.114 \cdot B$$



Исходное цветное изображение (DSC00375.JPG)

полутоновое изображение

Рис. 1. Преобразование цветного изображения в полутоновое

### 2). Бинаризация лучше с автоматическим определением порога.

Порог бинаризации можно определить при помощи дискриминантного метода [7, 11] (классический метод Otsu). Порог  $t^*$  выбирается, исходя из максимизации межклассовой дисперсии уровней яркости:

$$\sigma_B^2(t) = \frac{[\mu(L-1) \cdot F(t) - \mu(t)]^2}{F(t) \cdot (1 - F(t))}$$

$$\mu(t) = \sum_{i=0}^t i \cdot p(i), \quad F(t) = \sum_{i=0}^t p(i), \quad p(t) = \frac{n(t)}{h \cdot w}, \quad t = 0, 1, \dots, L-1,$$

$$\sigma_B^2(t^*) = \max_{0 \leq t \leq L-1} \sigma_B^2(t)$$

$n$  – гистограмма изображения;

$L$  – число уровней квантования яркости;

$h$  – высота изображения;

$w$  – ширина изображения;

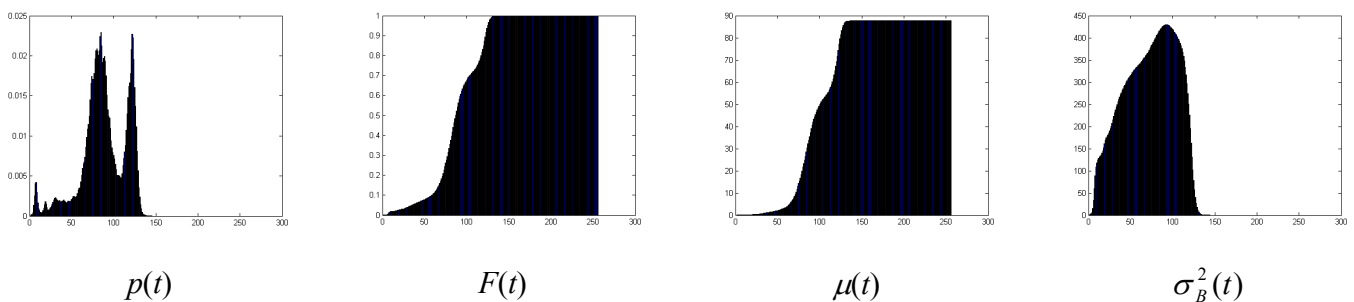
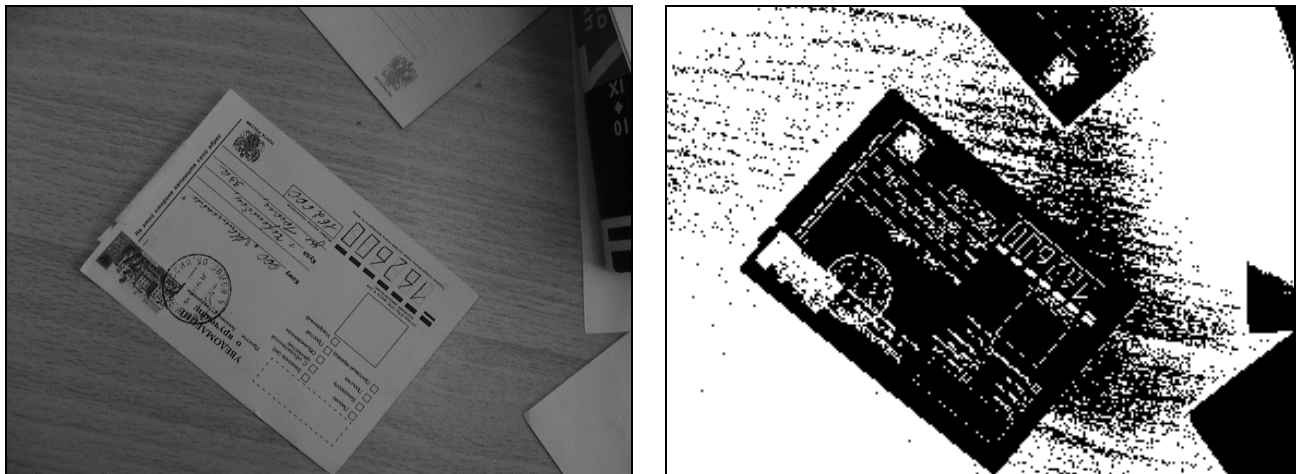


Рис. 2. Примеры графиков функций, используемых при расчете порога бинаризации. Функции получены для изображения на рис.1

Максимуму соответствует значение порога  $t^* = 91$ .



полутоновое изображение

бинарное изображение

Рис. 3. Бинаризация изображения с выбранным порогом

### 3). Выделение связных областей бинарного изображения.

На данном этапе необходимо выделить отдельные объекты на бинарном изображении, т.е. каждой ненулевой точке поставить в соответствие номер объекта.

Исходное изображение BW									Результат выделения серий									Результат разметки L								
1	1	1	0	0	0	0	0	→	1	3	5	0	0	0	0	0	→	1	1	1	0	0	0	0	0	
0	1	1	0	1	1	0	0		0	3	5	0	6	7	0	0		0	0	1	1	0	2	2	0	0
0	0	1	0	1	1	0	0		0	0	5	0	6	7	0	0		0	0	0	1	0	2	2	0	0
0	0	1	0	0	0	0	0		0	0	5	0	0	0	0	0		0	0	0	0	1	0	0	0	0
1	0	1	0	0	0	1	1		2	0	5	0	0	0	9	10			1	0	1	0	0	0	3	3
1	0	1	0	0	0	1	0		2	0	5	0	0	0	9	0			1	0	1	0	0	0	3	0
1	1	1	0	0	1	1	0		2	4	5	0	0	8	9	0			1	1	1	0	0	3	3	0

Рис. 4. Разметка бинарного изображения

Существует множество различных алгоритмов выделения связных областей изображения [8, 9 стр. 245-250 , 10]. Рассмотрим один из них [8].

а). В каждом столбце (или строке) изображения BW определяется начало и конец серии единиц. Результаты помещаем в таблицу вида (пример для приведенного выше изображения BW):

Таблица 2. Выделение серий единиц бинарного изображения

Номер серии	Начало серии (номер строки)	Конец серии (номер строки)	Номер столбца
1	1	1	1
2	5	7	1
3	1	2	2
4	7	7	2
5	1	7	3
6	2	3	5
7	2	3	6
8	7	7	6
9	5	7	7
10	5	5	8

б). Определяются серии, принадлежащие одному объекту с учетом выбранного определения связности. Результат можно представить в виде матрицы смежности графа, вершинам которого соответствуют выделенные серии. Две вершины соединяет ребро, если они принадлежат одному объекту.

Пересечение двух серий, имеющих координаты начала и конца 1:(a,b) 2:(c,d) можно определить следующим образом:

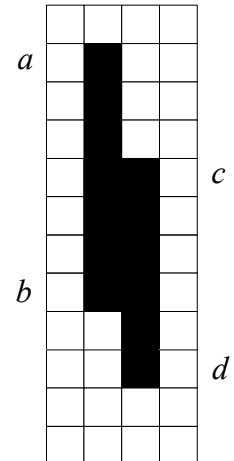
$$IntersectionPixelCount = \min(b, d) - \max(a, c) + 1$$

Серии считаются принадлежащими одному объекту, если

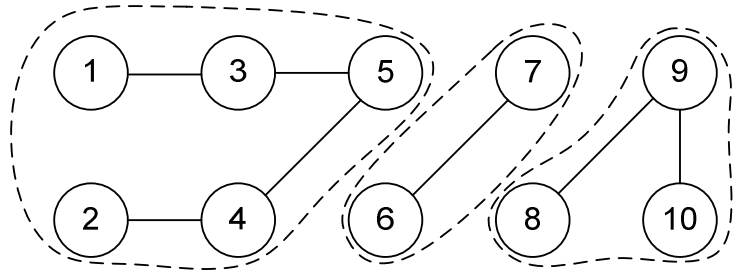
$IntersectionPixelCount > 0$  для четырехсвязности

$IntersectionPixelCount \geq 0$  для восьмисвязности

Матрица смежности и соответствующий ей граф для рассматриваемого примера имеют вид:



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



Поскольку данная матрица является разреженной (малое число ее элементов отлично от нуля), эффективнее данный граф представлять каким либо другим способом, например с помощью списков смежности. Связные компоненты данного графа соответствуют отдельным объектам. Выделение компонент связности может быть выполнено при помощи любого известного алгоритма, например [12, стр.227-228]. Для приведенного примера первая компонента связности содержит вершины 1, 2, 3, 4, 5, во вторая – 6, 7 и третья 8, 9 и 10. Это значит, что первому объекту соответствуют серии 1-5, второму 6,7, третьему 8-10. Результат разметки представлен на рис.4 справа.

#### 4). Расчет характеристик выделенных областей.

Для классификации объектов могут оказаться полезными следующие их характеристики:

- Площадь  $S = \sum_{(i,j) \in D} 1$

- Координаты, ограничивающего область, прямоугольника

$$x_{\min} = \min_{(i,j) \in D} \{j\}, \quad x_{\max} = \max_{(i,j) \in D} \{j\}$$

$$y_{\min} = \min_{(i,j) \in D} \{i\}, \quad y_{\max} = \max_{(i,j) \in D} \{i\}$$

- Максимальные размеры области

$$w = x_{\max} - x_{\min}, \quad h = y_{\max} - y_{\min}$$

- Координаты центра тяжести

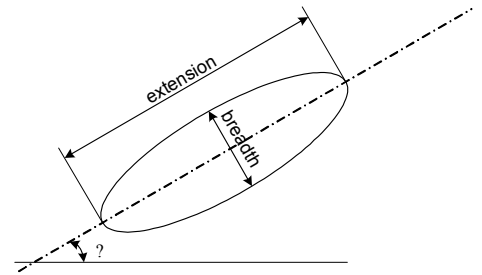
$$\bar{x} = \frac{1}{n} \sum_{(x,y) \in Reg} x; \quad \bar{y} = \frac{1}{n} \sum_{(x,y) \in Reg} y;$$

- Ориентация оси инерции

$$\theta = \frac{1}{2} \arctan \left( \frac{2m_{11}}{m_{20} - m_{02}} \right)$$

где  $m_{ij}$  дискретные центральные моменты области:

$$m_{ij} = \sum_{(x,y) \in Reg} (x - \bar{x})^i (y - \bar{y})^j$$



- Удлиненность, нецентрированность (эксцентриситет)

$$elongation = \frac{m_{20} + m_{02} + \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}}{m_{20} + m_{02} - \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}}$$

- Компактность

$$C = \frac{P^2}{A}$$

- Ширина и длина объекта (относительно оси инерции) extension, breadth.

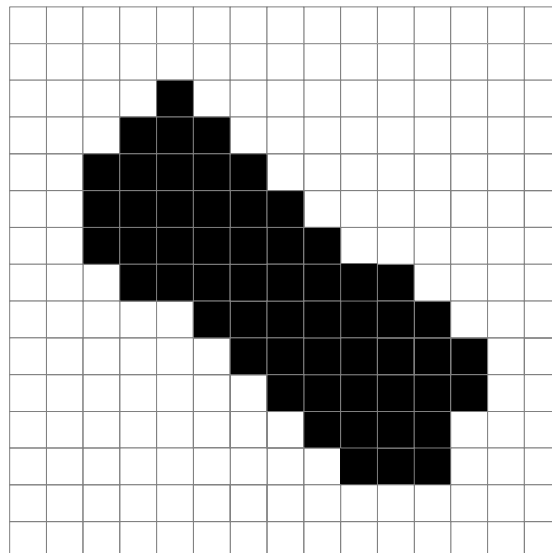


Рис. 5. Пример изображения (последний маркер) и некоторые его характеристики  
 $S = 57, x_{\min} = 3, x_{\max} = 13, y_{\min} = 3, y_{\max} = 13, w = 11, h = 11$ , Координаты центра тяжести:  
 $x_c = 7.9123, y_c = 8.3333, \theta = -41.1973, \text{extension} = 14.2731, \text{breadth} = 5.4798$

Часть характеристик объектов могут быть вычислены непосредственно во время разметки. Для этого необходимо расширить табл. 2.

Таблица 3. Оценка некоторых характеристик объектов в процессе разметки

Номер серии	Начало серии (номер строки)	Конец серии (номер строки)	Номер столбца	S	$x_{\min}$	$x_{\max}$	$y_{\min}$	$y_{\max}$	$\bar{x}$	$\bar{y}$
1	1	1	1	1	1	1	1	1	1	1
2	5	7	1	3	1	1	5	7	1	6
3	1	2	2	2	2	2	1	2	2	1.5
4	7	7	2	1	2	2	7	7	2	7
5	1	7	3	7	3	3	1	7	3	3.5
6	2	3	5	2	5	5	2	3	5	2.5
7	2	3	6	2	6	6	2	3	6	2.5
8	7	7	6	1	6	6	7	7	6	7
9	5	7	7	3	7	7	5	7	7	6
10	5	5	8	1	8	8	5	5	8	5

После выяснения принадлежности серий объектам их характеристики определяются на основе характеристик серий:

Таблица 5. Характеристики объектов, определенные в процессе разметки

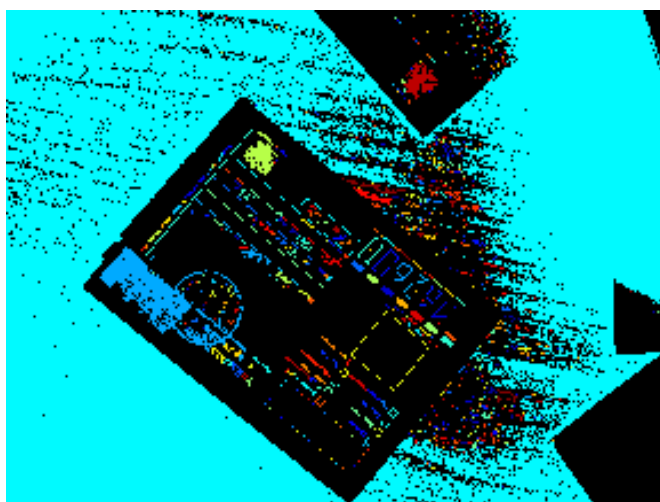
Номер объекта	Входящие серии	S	$x_{\min}$	$x_{\max}$	$y_{\min}$	$y_{\max}$	$\bar{x}$	$\bar{y}$
1	1, 2, 3, 4, 5	14	1	3	1	7	2.2143	4.0714
2	6, 7	4	5	6	2	3	5.5	2.5
3	8, 9, 10	5	6	8	5	7	7	6

### 5). Удаление объектов не «похожих» на маркер по некоторым параметрам.

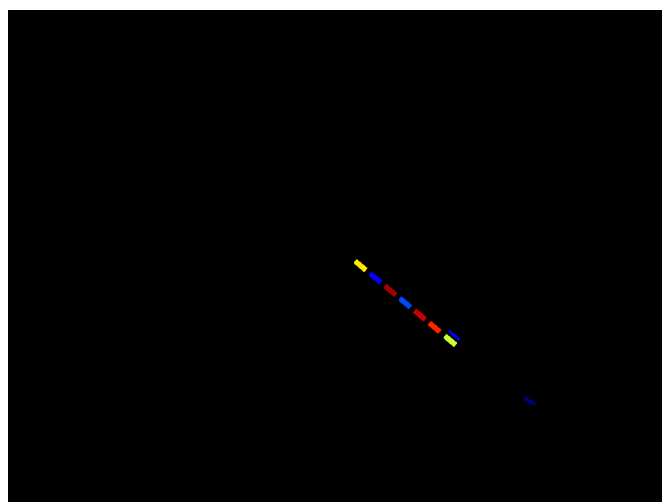
Многие параметры маркеров зоны индекса принадлежат некоторым интервалам доверия. Если какой-либо объект имеет значения характеристик вне этих интервалов его можно удалить. Возможно, различительными параметрами являются: площадь, отношение длины к ширине

объекта, линейные размеры, отношение площади объекта к площади прямоугольника минимальных размеров, содержащего объект целиком и т.д.

Рис.6 демонстрирует действие фильтрации. Исходное изображение содержало 4739 четырехсвязных объектов (2762 восьмисвязных), отфильтрованное изображение содержит 9 объектов (для обоих вариантов связности).



Результат разметки бинарного изображения, представленного на рис. 3 (различные объекты имеют разные цвета)



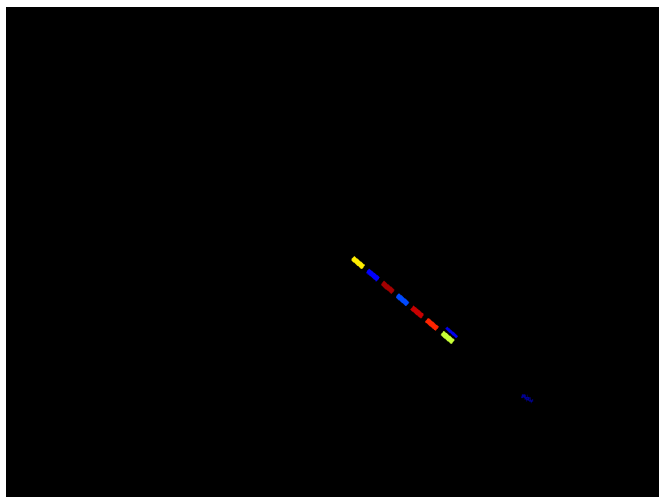
Результат фильтрации

Рис. 6. Фильтрация объектов

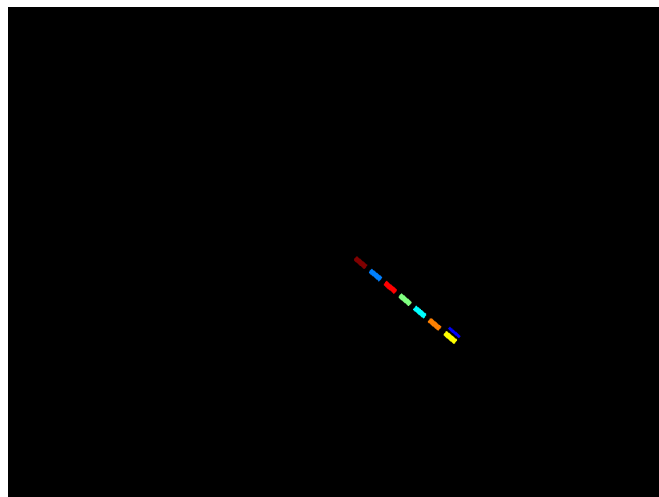


#### 6). Определение маркеров индекса.

На данном этапе среди оставшегося множества объектов необходимо выделить 8 маркеров индекса.



Объекты после фильтрации

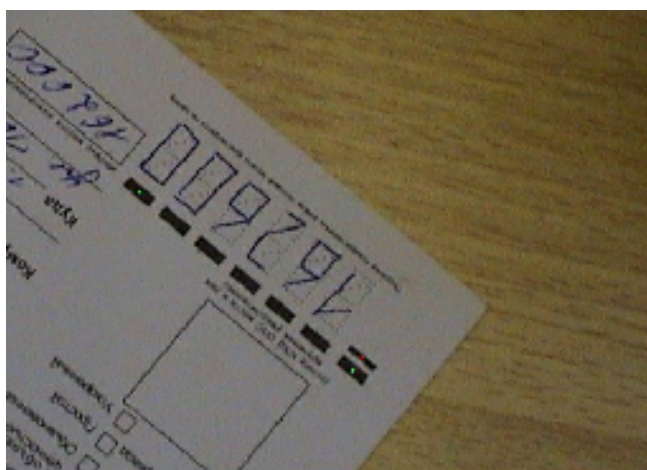


Выделенные маркеры индекса

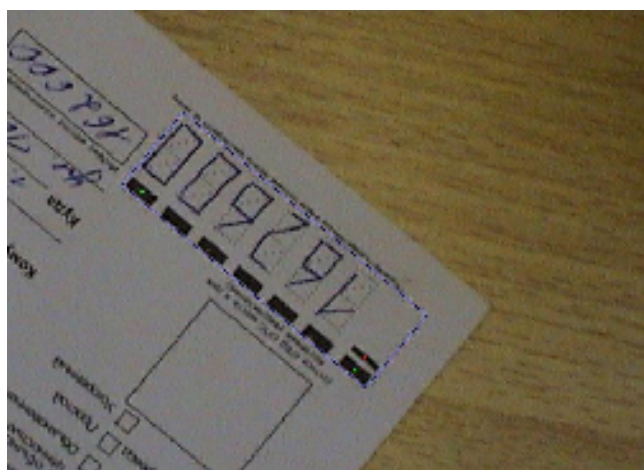
Рис. 7. Определение маркеров индекса

#### 7). Локализация зоны индекса.

На данном этапе на основе анализа характеристик выделенных маркеров и информации о размерах зоны индекса определяются координаты прямоугольника, содержащего индекс.



Опорные маркеры индекса



Выделенная зона индекса

Рис. 8. Локализация зоны индекса

Для определения угла наклона маркера необходимо рассчитать параметры прямой, проходящей через центры тяжести всех маркеров кроме маленького. Это можно сделать с помощью метода наименьших квадратов, или просто найти уравнение прямой проходящий через центры первого и последнего маркеров (на рис. 8 отмечены зелеными точками). Затем с учетом положения маленького маркера (на рис. 8 его центр отмечен красной точкой) определяется ориентация зоны индекса.

Результатом выделения зоны являются координаты четырех вершин прямоугольника, содержащего индекс и угол наклона маркера  $\theta$ .

### 8). Вырезание и поворот индекса.

Для вырезания зоны индекса создается матрица, размеры которой совпадают с размерами выделенной на предыдущем этапе зоны. Для каждой точки данной матрицы определяется ее положение на исходном изображении на основе обратного преобразования координат:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} t_x \\ t_y \end{pmatrix}, \quad (1)$$

где  $R^{-1}$  – матрица, обратная к матрице вращения на угол  $\theta$ .

вектором трансляции (сдвига)  $-\begin{pmatrix} t_x \\ t_y \end{pmatrix}$ ,

$t_x, t_y$  – координаты левого верхнего угла зоны индекса (угла в котором находится маленький маркер)

$$R^{-1} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

С помощью выражения (1) для каждой точки  $(x, y)^T$  формируемого изображения индекса определяется ее положение на исходном изображении  $(x', y')^T$ . Яркость точки  $(x, y)^T$  может быть определена при помощи метода ближайшего соседа или посредством билинейной интерполяции.

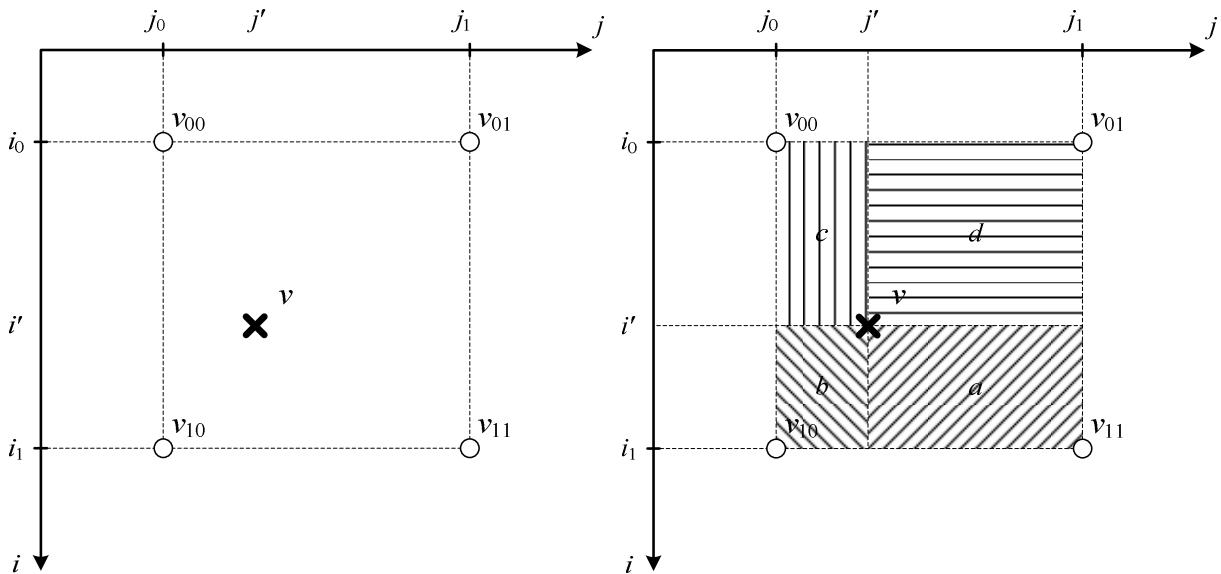
### Билинейная интерполяция

Данный метод интерполяции называется билинейным, потому что построенная с его помощью функция яркости (внутри области интерполирования) является линейной функцией относительно каждой из своих переменных при любом фиксированном значении другой. Суть метода состоит в том, чтобы получить искомое приближение в результате последовательного

применения процедур одномерной линейной интерполяции. Мы хотим построить оценку значения яркости  $v$  в точке  $(i', j')$ . Пусть  $(i_0, j_0)$  - ближайшая соседняя точка, значения координат которой являются целыми и не превосходят значений координат  $(i', j')$ , т.е.  $i_0 \leq i'$  и  $j_0 \leq j'$ . Если обозначить  $i_1 = i_0 + 1$ ,  $j_1 = j_0 + 1$ , то координаты четырех ближайших к точке  $(i', j')$  соседних точек примут вид:  $(i_0, j_0)$ ,  $(i_0, j_1)$ ,  $(i_1, j_0)$  и  $(i_1, j_1)$ . Обозначим известные значения яркости в этих точках  $v_{00}$ ,  $v_{01}$ ,  $v_{10}$  и  $v_{11}$  соответственно. Применим теперь процедуру одномерной линейной интерполяции по горизонтали, чтобы найти оценки для значений яркости в точках  $(i_0, j')$  и  $(i_1, j')$ , после чего, опять применим одномерную линейную интерполяцию, на этот раз по вертикали, найдем искомое значение яркости в точке  $(i', j')$ . Окончательный результат всей процедуры имеет вид:

$$v = (i_1 - i')(j_1 - j')v_{00} + (i_1 - i')(j' - j_0)v_{01} + (i' - i_0)(j_1 - j')v_{10} + (i' - i_0)(j' - j_0)v_{11}.$$

Изменение порядка применения процедур интерполяции (сначала по вертикали, затем по горизонтали) не меняет окончательного результата, что легко видеть из приведенной формулы. Кроме того, очевидно, что эта формула приводится к виду (5.11-7) и даст правильные значения для яркости четырех соседних точек.



Если обозначить площади областей как на рисунке, то формулу можно переписать в виде:

$$v = av_{00} + bv_{01} + cv_{11} + dv_{10}.$$

При использовании интерполяции по ближайшему соседу яркость  $v$  полагается равной яркости ближайшей к ней точки.

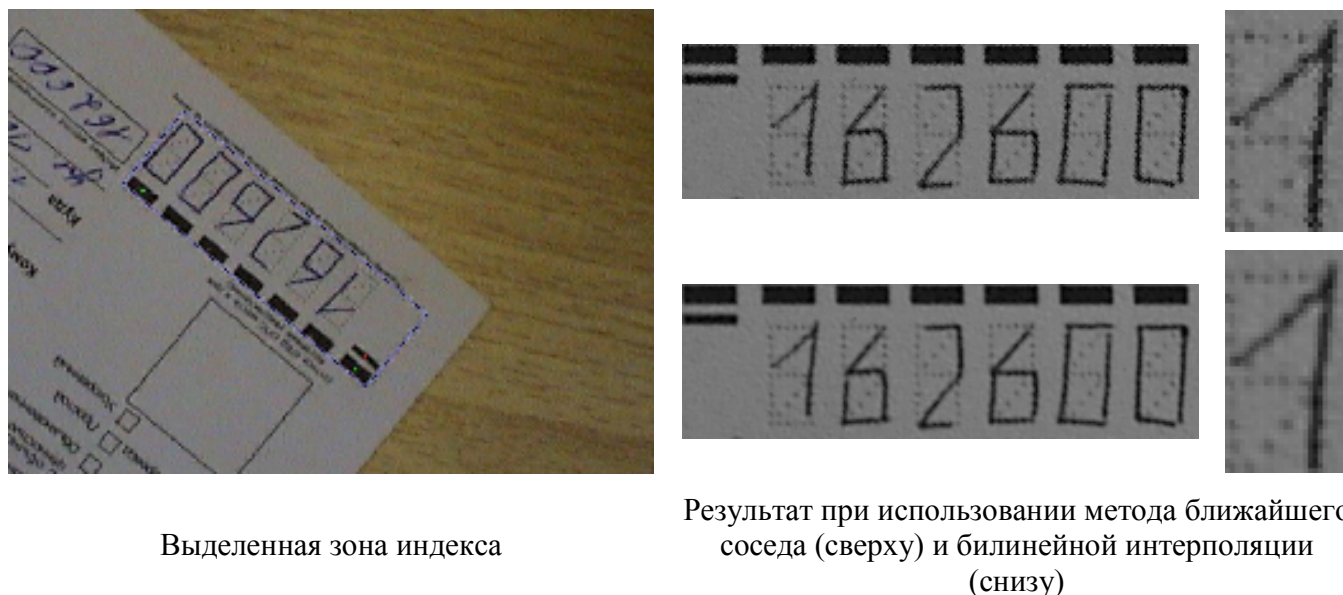


Рис. 9. Вырезание зоны индекса

#### 9). Выделение отдельных символов индекса.

После бинаризации изображения индекса отдельные символы могут быть обнаружены одним из способов:

- на основе относительных размеров и положения символов внутри зоны индекса;
- анализируя суммарную яркость столбцов/строк;
- на основе позиций и размеров маркеров;



Рис. 10. Выделение отдельных символов индекса

# 10). Распознавание индекса.

Распознавание символов может быть выполнено подсчетом количества точек в 9 дискриминантных зонах (рис. 11). Каждой зоне можно поставить в соответствие логическую переменную, указывающую на наличие (1) или отсутствие (0) в ней линии. Сравнивая данное описание с эталонными, можно принять решение об изображенном символе. Другой вариант распознавания заключается в подсчете количества точек  $N_i$  в зонах соответствующих различным символам  $i = 0 - 9$ . В качестве распознанного выбирается символ  $i$ , количество точек для которого оказалось максимальным.

Таблица 6. Эталонные описания символов

Символ	Номер зоны								
	1	2	3	4	5	6	7	8	9
0	1	1	1	0	0	0	1	1	1
1	0	0	0	1	0	0	0	1	1
2	0	0	1	0	0	1	1	1	0
3	0	0	1	1	1	1	0	0	0
4	1	0	0	0	1	0	0	1	1
5	1	0	1	0	1	0	1	0	1
6	0	1	0	1	1	0	1	0	1
7	0	1	1	1	0	0	0	0	0
8	1	1	1	0	1	0	1	1	1
9	1	0	1	0	1	1	0	1	0

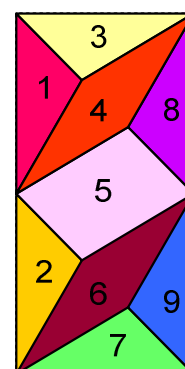
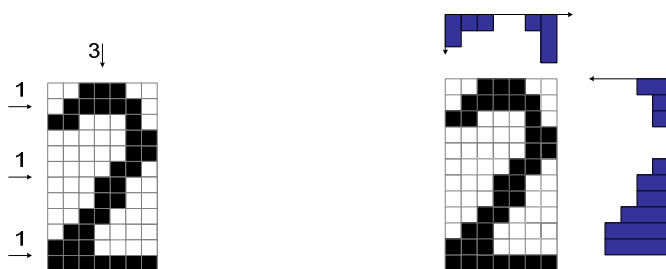


Рис. 11.  
Дискриминантные  
зоны символа

Другой подход заключается в подсчете количества пересечений символа в некоторых строках/столбцах. Также возможно выполнить распознавание на основе анализа профилей символа (левого, правого, верхнего, нижнего). Для выполнения распознавания необходимо чтобы количество признаков сравниваемых объектов совпадало, поэтому при расчете последних признаков либо изображение приводят к некоторому фиксированному размеру, либо изображение разбивается на некоторое количество участков для определения признаков.



количество пересечений символа

профили символа

Рис. 12. Характеристики символов

## ЛИТЕРАТУРА

1. Дьяконов В. MATLAB. Обработка сигналов и изображений. Специальный справочник. – СПб.: Питер, 2002. - 608 с.: ил.
2. Павлилис Т. Алгоритмы машинной графики и обработки изображений. М.: Радио и связь, 1986
3. Прэйт У. Цифровая обработка изображений: в 2-х томах. Перевод с английского под редакцией канд. техн. наук Д. С. Лебедева. М.: Мир, 1982
4. Роджерс Д. Алгоритмические основы машинной графики. М.: Мир. 1989.
5. Рудаков П. И., Сафонов И. В. Обработка сигналов и изображений. MATLAB 5.x / под общ. ред. К. т. н. В. Г. Потемкина. - М.: ДИАЛОГ-МИФИ, 2000. - 416 с. - (Пакеты прикладных программ; Кн. 2)
6. Форсайт, Понс. Компьютерное зрение. Современный подход.: Пер. с англ. – М.: Вильямс, 2004 – 928 с.
7. Бакут П. А., Колмогоров Г. С., Ворновицкий И. Э. Сегментация изображений: методы пороговой обработки // Зарубеж. радиоэлектрон. – 1987. №10. – С. 6-24.
8. Haralick, Robert M., and Linda G. Shapiro, Computer and Robot Vision, Volume I, Addison-Wesley, 1992, pp. 28-48.
9. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. - 2-е изд., испр. - М.: Физматлит, 2003. - 784 с. - ISBN 5-9221-0270-2
10. Fu Chang, Chun-Jen Chen, and Chi-Jen Lu. A linear-time component-labeling algorithm using contour tracing technique. // Computer Vision and Image Understanding 93 (2004) pp. 206–220
11. Otsu, N., "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.
12. Дискретная математика для программистов / Ф.А. Новиков – СПб: Питер, 2000. – 304 с.: ил.