

# Критерии оценки качества в задаче автоматизированной настройки алгоритмов бинаризации\*

Д. П. НИКОЛАЕВ, А. А. САРАЕВ

**Аннотация.** В данной работе предложен метод оценки качества бинаризации, отличающийся низкой трудоемкостью. Он основан на расчете невязки простейших статистик для ключевых фрагментов изображения. Проведено сравнение свойств метода при использовании MSE метрики и предложенной полуметрики CPM. Кроме того, на примере одной из модификаций алгоритма Ниблэка продемонстрирована пригодность предложенного метода для решения задачи автоматизированного подбора оптимальных параметров бинаризации.

**Ключевые слова:** обработка изображений, бинаризация, оценка качества изображений, OCR.

## Введение

Бинаризация изображений используется при решении широкого круга задач: распознавание документов [1], анализ медицинских снимков [16], неразрушающий контроль [14] и т. д. Большое число различных методов бинаризации и зависимость результата работы многих из них от выбранных параметров приводят к необходимости аккуратного и обоснованного выбора метода бинаризации и способа его настройки.

Подбор параметров алгоритма, оценку результатов бинаризации зачастую осуществляет человек (так называемые психофизические методы) [3, 4]. Психофизические методы трудоемки, с их помощью затруднительно обрабатывать большие объемы информации.

Автоматизация подбора требует выбора информативной функции расстояния между бинарными изображениями (прямые методы) или некоторого целевого функционала, который может быть рассчитан для бинаризованного изображения (косвенные методы).

Среди косвенных методов, можно упомянуть, во-первых, методы, где целевым функционалом является качество работы некоторой системы, частью которой является сама бинаризация. Так, оценка бинаризации на основе распознавания [5, 9] (в роли функционала здесь может выступать точность, полнота или их комбинация) позволяет получить результат, оптимальный для выбранной OCR-системы.

Такой подход может быть лучшим выбором, если нас интересуют только результаты работы выбранной системы распознавания. Иначе обстоит дело, если мы хотим использовать бинаризацию для OCR-системы, недоступной на момент оптимизации параметров, или для задачи мало связанной с распознаванием текста. Тогда сравнение с некоторой «правильной» бинаризацией представляется предпочтительным.

Во-вторых, в процессе оценки можно опираться на некоторые общепризнанные алгоритмы. Например, вычислять расстояние между границей на бинарной картинке и границей, извлеченной из исходного «серого» алгоритмом детектирования границ [3, 15]. Здесь возникают вопросы выбора параметров для самого «эталонного» алгоритма: результат оптимизации алгоритма бинаризации начинает зависеть от параметров детектора краев.

Прямые методы предоставляют больше независимости (нет «посредника» при формировании оценки) и гибкости («правильную» разметку всегда можно подкорректировать). К сожалению, для их работы необходимо создание набора «правильно» бинаризованных изображений («идеальной» разметки, «идеальной карты» бинаризации), что является крайне трудоемким процессом.

В данной работе рассматриваются прямые методы, предлагается эффективный способ создания «эталонной» разметки. На основе введенного критерия оценки качества настраивается модификация алгоритма Ниблэка.

\* Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта №13–07–12172.

### 1. Бинаризация

Будем представлять изображение в виде двумерной матрицы, значения элементов которой определяются отображением вида:

$$f : \langle x, y \rangle \rightarrow e, \tag{1}$$

где  $e \in E$ , множеству, определяющему тип изображения,  $x, y \in Z$ ,  $0 \leq x < m$ ,  $0 \leq y < n$ , где  $n$  и  $m$  — высота и ширина изображения. Для полутонового или «серого» изображения будем считать, что  $e \in [0; 1]$ ; для бинарного —  $e \in \{0, 1\}$ .

Обозначим множество всех полутоновых изображений как  $G$ , а множество бинарных как  $B$ . Тогда алгоритм бинаризации можно определить следующим способом:

$$\langle G, P_B \rangle \rightarrow B, \tag{2}$$

где  $P_B$  — это множество параметров соответствующего метода бинаризации.

Примем, что  $I \in G$ ,  $A \in B$ . В этом случае, любую бинаризацию можно формально свести к правилу:

$$A(x, y) = \begin{cases} 0, & \text{если } I(x, y) \leq T(x, y), \\ 1, & \text{если } I(x, y) > T(x, y), \end{cases} \tag{3}$$

где  $T(x, y) \in \mathcal{R}$  — порог в соответствующей точке изображения.

В простейшем случае порог не зависит от координат точки:

$$g : \langle G, \mathcal{R} \rangle \rightarrow B, \quad T(x, y) = const. \tag{4}$$

Такие алгоритмы бинаризации называются глобальными, в них порог бинаризации может вычисляться

автоматически (как, например, в методе Оцу [11]), или являться внешним параметром алгоритма. Благодаря своей простоте, бинаризация с фиксированным глобальным порогом может быть быстро настроена человеком.

#### 1.1. Алгоритм Ниблэка

Алгоритм Ниблэка [10] является одним из самых известных локальных методов бинаризации. При его применении для каждого пикселя устанавливается локальный порог по формулам:

$$T(x, y) = \mu(x, y) + k\sigma(x, y), \tag{5}$$

$$\mu(x, y) = \frac{1}{(2w_x+1)(2w_y+1)} \sum_{i=-w_x}^{w_x} \sum_{j=-w_y}^{w_y} I(x+i, y+j), \tag{6}$$

$$\sigma(x, y) = \sqrt{D(x, y)} = \sqrt{\mu(I(x, y)^2) - \mu(I(x, y))^2}. \tag{7}$$

Хотя в большинстве случаев алгоритм Ниблэка дает результат, хорошо подходящий для задач распознавания (не теряет контрастные детали), данный алгоритм плохо справляется со слабовыраженными структурами в области фона — может появляться «грязь» (см. рис. 1).

Данный недостаток послужил причиной появления ряда идейно близких алгоритмов, иногда именуемых ниблэковскими: Саволы [13], Вольфа [17], Фенга [6], алгоритм NICK [8]. Обобщая, можно сказать, что они вводят дополнительные параметры, позволяющие устранить указанный дефект.

Мы будем рассматривать простую и вычислительно эффективную модификацию алгоритма Ниблэка:

$$T(x, y) = \mu(x, y) + k\sigma(x, y) + a, \tag{8}$$

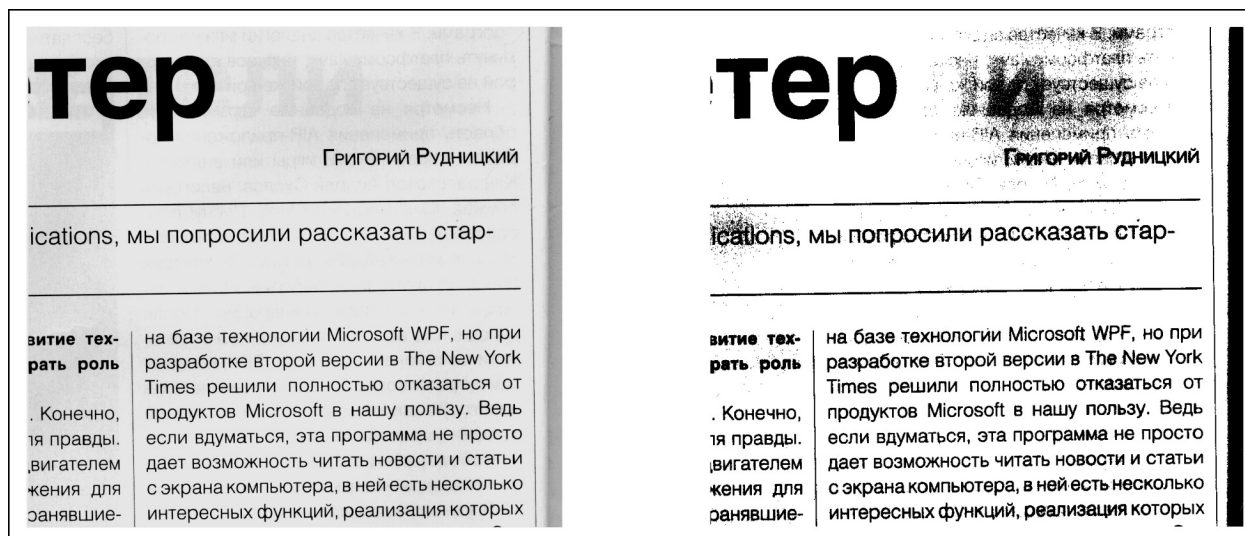


Рис. 1. Изображение размера 1003x940 и результат его бинаризации алгоритмом Ниблэка ( $k = -0,2$ ;  $w_x = w_y = 401$ )

рассматриваемую также, например, в [2]. Здесь за удаление шума в области фона отвечает дополнительный параметр  $a$ .

Для указанной модификации можно записать:

$$nib : \langle G, \langle \mathbb{R}, \mathbb{R}, \mathbb{N}, \mathbb{N} \rangle \rangle \rightarrow B, \quad (9)$$

где  $\langle \mathbb{R}, \mathbb{R}, \mathbb{N}, \mathbb{N} \rangle$  является пространством  $\langle k, a, w_x, w_y \rangle$  — параметров алгоритма.

### 1.2. Оценка значений параметров алгоритма Ниблэка

При решении задачи оптимизации сначала необходимо ограничить область допустимых значений параметров. Сделаем это для рассматриваемой модификации алгоритма Ниблэка. Здесь и далее будем считать размеры окна произвольными, но фиксированными.

Очевидно, что математическое ожидание  $\mu(x, y) \in [0; 1]$ . Кроме того, имеет место следующее утверждение: для заданного значения  $\mu(x, y)$  дисперсия  $D(x, y)$  достигает своего максимума тогда и только тогда, когда все значения пикселей равны нулю и единице, за исключением, возможно, одного.

Предположим противное. Тогда в рассматриваемой области есть хотя бы два пикселя, с определенными значениями. Для определенности примем  $I_1 > I_2$ . Тогда пусть  $I'_1 = 1$ ,  $I'_2 = I_1 + I_2 - 1$ . Очевидно, что  $\mu$  не изменилось. В то же время, учитывая определение дисперсии (7) и тот факт, что  $(1 - \mu)^2 + (I_1 + I_2 - \mu)^2 > (I_2 - \mu)^2 + (I_2 - \mu)^2$ , приходим к заключению, что наше предположение неверно.

Оценим теперь максимальное значение дисперсии. Для достаточно больших окрестностей наличием одного элемента со значением не равным нулю или единице можно пренебречь. В этом случае:

$$\sigma = \sqrt{\mu(I) - \mu(I)^2}. \quad (10)$$

Из (10) следует, что  $\sigma \in [0; 0,5]$  с максимумом при  $\mu = 0,5$  (см. рис. 2).

При достаточном размере окна если  $\sigma = 0$ , то соответствующая точка должна быть классифицирована как фон. Получаем  $\mu(x, y) = I(x, y)$ , откуда  $\mu(x, y) + k\sigma(x, y) + a < I(x, y) \Leftrightarrow \mu(x, y) + a < I(x, y) \Leftrightarrow a < 0$ .

Кроме того, при разумных параметрах бинаризации на множестве  $\mu \in [0; 1]$ ,  $\sigma \in [0; 0,5]$ ,  $v \in [0; 1]$  должны быть выполнимы оба неравенства:

$$\mu + k\sigma + a < v, \quad (11)$$

$$\mu + k\sigma + a \geq v. \quad (12)$$

Из выполнимости (11) следует:

$$k \geq 0 \Rightarrow a < 1, \quad k < 0 \Rightarrow a < 1 - 0,5k. \quad (13)$$

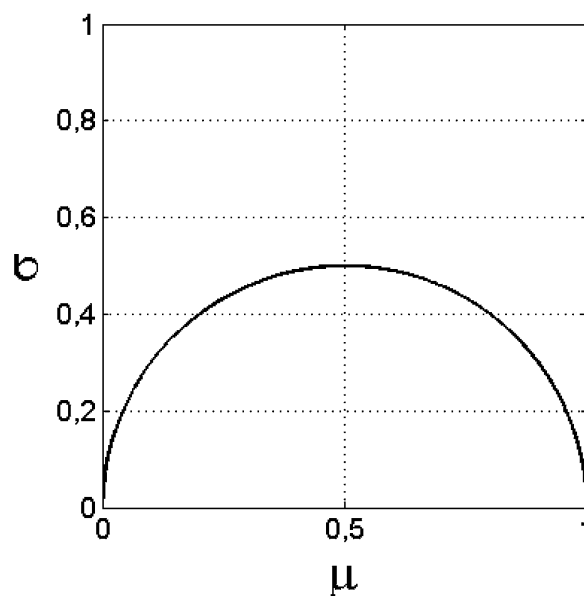


Рис. 2. Зависимость  $\sigma$  от  $\mu$

Из выполнимости второго неравенства следует:

$$k \geq 0 \Rightarrow a \geq -1 - 0,5k, \quad k < 0 \Rightarrow a \geq -1. \quad (14)$$

Тогда, учитывая, что  $a < 0$ , получаем:

$$k \geq 0 \Rightarrow a \in [-1 - 0,5k; 0], \quad k < 0 \Rightarrow a \in [-1; 0]. \quad (15)$$

## 2. Задача оптимизации параметров бинаризации

В данной работе рассматриваются задачи настройки алгоритма бинаризации, а также разработки эффективных критериев качества с низкой трудоемкостью.

В качестве входных данных мы имеем набор разнородных полутоновых изображений печатных документов. У нас имеется репрезентативная обучающая выборка из этого набора, для которой составлена карта идеальной разметки в виде набора прямоугольных областей с заданным оптимальным порогом.

Выходными данными задачи являются значения оптимальных параметров исследуемого алгоритма бинаризации в смысле заданного критерия контроля качества.

## 3. Выбор функции расстояния

### 3.1. Основные положения

Настройка алгоритмов бинаризации связана с некоторым минимизируемым критерием. Логично в виде такого критерия принять расстояние между

получаемыми на текущем этапе образцами и «правильной» разметкой (прямые методы оценки качества), ведь гарантированно верно бинаризовать изображение в общем случае может только эксперт. То есть, если  $s \in \mathbb{R}$ ,  $A \in B$  — оцениваемое изображение, а  $J \in B$  — идеальная бинаризация, то критерием качества бинаризации является функция следующего вида:

$$q : \langle A, J \rangle \rightarrow s. \tag{16}$$

Классическим расстоянием данного вида является  $MSE$  (mean squared error, средняя квадратичная ошибка):

$$MSE = \frac{\sum_{x=1}^M \sum_{y=1}^N (A(x, y) - J(x, y))^2}{MN}. \tag{17}$$

В случае с бинарными изображениями числитель формулы (17) — количество различающихся пикселей, а собственно  $MSE$  — просто доля таких пикселей.

Также отметим несколько расстояний, используемых в ежегодном соревновании алгоритмов бинаризации DIBCO [7, 12]: PSNR (peak signal-to-noise ratio, пиковое отношение сигнала к шуму), precision (точность), recall (полнота), F-measure (F-мера), MPM (misclassification penalty metric, метрика штрафов за ошибки классификации). Рассмотрение свойств этих метрик выходит за рамки данной работы.

### 3.2. Расстояние на основе количества черных пикселей

Самым очевидным недостатком использования прямых методов оценки качества является трудоемкость создания карт «идеальной разметки». Но есть и менее очевидная проблема: идеала как такового не существует. Рассмотрим тривиальный пример, приведенный на рис. 3.

Как видим, имеется два изображения буквы «Т» примерно одинакового качества. Ни для человека, ни для OCR-системы распознать данные растры не сложно. Тем не менее, каждый раз при попиксельном фиксировании идеала мы штрафует одну из приведенных бинаризаций относительно другой. Так устроены все выше приведенные расстояния, которые условно можно назвать XOR-метриками. Данный класс оценок наиболее распространен, и зачастую дает приемлемые результаты, что не отменяет указанный выше дефект.

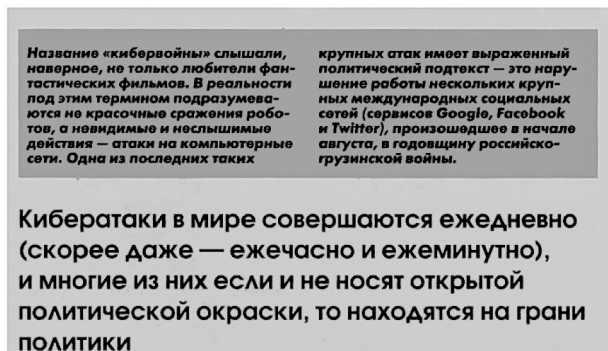
Можно заметить, что мы можем из одного бинаризованного растра получить другой с помощью двух операций: переставить пиксель, и изменить цвет пикселя. Первая не является необходимой, но примечательна тем, что именно так можно из рис. 4а получить рис. 4б. Так мы подходим к идее, что расстояние между двумя изображениями можно измерять как разницу количества черных пикселей. То есть, если функция  $c$  — количество черных пикселей на изображении:

$$c(A) = \sum_{x=1}^w \sum_{y=1}^h \begin{cases} 1, & \text{если } A(x, y) = 0, \\ 0, & \text{если } A(x, y) = 1. \end{cases} \tag{18}$$

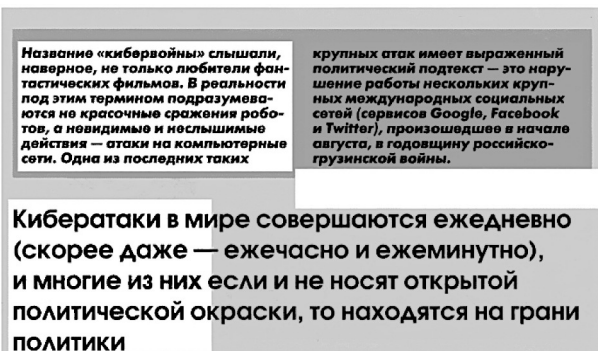


а) б)

Рис. 3. Два бинарных изображения буквы «Т»



а)



б)

Рис. 4. Пример «правильной» разметки: а — исходное изображение; б — идеальная разметка.

Тогда расстояние (точнее, псевдометрику) на основе количества черных пикселей (для удобства обозначим его как CPM — count pseudometric measure) можно выразить следующим образом:

$$CPM(A, J) = |c(A) - c(J)|. \quad (19)$$

На первый взгляд это может показаться наивным, поскольку такое расстояние вовсе не учитывает форму полученного изображения. Но, во-первых, исследуемый метод Ниблэка в определенном смысле «монотонен»: выделить шум, не выделив объекты интереса обычно затруднительно, если только шум не выглядит как объект. А во-вторых, размечая не только текст, но и отдельно область фона мы добьемся того, что если бинаризация добавляет шум в области текста и не штрафует, то в размеченной области фона ошибка будет велика.

### 3.3. Правила разметки изображений

Рассмотрим теперь правила создания целевых значений («идеальной разметки») при подборе параметров бинаризации с использованием CPM. Эти правила были выработаны экспериментально с целью минимизации трудоемкости разметки.

1. Размечать следует небольшие прямоугольные фрагменты. Размеры области должны быть таковы, чтобы можно было получить «правильную» бинаризацию, задавая лишь глобальный порог бинаризации для области. При этом учитываться будет только полученное количество черных пикселей.

В таком случае, разметка  $M$  определяется как:

$$M : \{m \langle r, t \rangle, r \in X, t \in \mathbb{R}\}, \quad (20)$$

где  $t$  — заданный экспертом порог, а  $X$  — множество прямоугольников:

$$X : \{r \langle x, y, w, h \rangle, x, y \in \mathbb{N}_0, w, h \in \mathbb{N}\}. \quad (21)$$

Если  $I \in G$  — исходное изображение, а  $M$  — заданная на нем разметка, то эталонное бинарное изображение  $J \in B$  области изображения  $V \in G$ ,  $V \subset I$  для заданной на ней разметки  $m \in M$ , ограниченной соответствующим прямоугольником  $r$ , может быть получено посредством бинаризации  $V$  с порогом  $t$ :

$$area : \langle I, r \rangle \rightarrow V, \quad g(V, t) \rightarrow J. \quad (22)$$

2. Следует размечать только ключевые, непохожие друг на друга участки изображения, что сильно снижает трудоемкость создания разметки.
3. Помимо текста, полезно разметить участок фона, ведь на исследуемом типе изображений всегда присутствует как минимум два типа областей — фон и текст. Если параметры бинаризации некорректны, то на участках фона может

появиться грязь. Наличие участка фона в разметке даст нам слагаемое с большим показателем ошибки, даже если на участках с текстом количество черного удовлетворительно.

## 4. Настройка алгоритмов бинаризации

### 4.1. Настройка бинаризации в общем виде

Как было отмечено выше, настройку алгоритма бинаризации можно рассматривать с точки зрения минимизации расстояния между опытными данными и заранее размеченными человеком образцами.

Рассмотрим произвольный алгоритм бинаризации  $b \in B$  с параметрами  $p_B \in P_B$ . Принимая во внимание (2), (4), (16), (22) можно выразить расстояние между бинаризацией и «правильной» разметкой:

$$s(I, M, p_B) = \sum_{i=1}^{|M|} q(b(area(I, r_i), p_B), g(area(I, r_i), t_i)). \quad (23)$$

Так как входными данными задачи является обучающая выборка  $T$ :

$$T : \{T_i : T_i \in G, 0 \leq i \leq K\}, \quad (24)$$

состоящая из  $K$  изображений  $T_i$ , для каждого из которых задана разметка  $M_i$ , общая ошибка на всем наборе будет равна:

$$S(T, M, p_B) = \sum_{i=1}^K s(T_i, M_i, p_B). \quad (25)$$

Искомые оптимальные параметры бинаризации алгоритма  $p_B \in P_B$  в этом случае:

$$p_B = \operatorname{argmin}_{p_B} [S(T, M, p_B)]. \quad (26)$$

### 4.2. Настройка алгоритма Ниблэка

Для изображения  $I$  можно высчитать матрицы  $\mu$  и  $\sigma$ . При этом мы знаем, что пиксель классифицируется как черный если и только если:

$$\mu(x, y) + \delta(x, y)k + a \geq I(x, y). \quad (27)$$

Теперь предположим, что мы работаем с матрицей  $H$  — двумерной гистограммой, где, для определенности, оси абсцисс соответствуют значения параметра  $k$ , а оси ординат — значения параметра  $a$ . Учитывая результаты, полученные в части 3.2, и фиксируя  $k_{\min}$  и  $k_{\max}$  получаем ограничение  $a \in [a_{\min}, a_{\max}]$ . Для отображения на целочисленную решетку, параметры дискретизируются с некоторыми интерва-

лами  $a_{step}$  и  $k_{step}$ . Значение в ячейке гистограммы  $H$  обозначим как  $H(k', a')$ , где

$$k' = (k - k_{min})/k_{step}, \quad a' = (a - a_{min})/a_{step}. \quad (28)$$

$H(k', a')$  хранит количество голосов, поданных за соответствующие значения параметров  $k$  и  $a$  точками изображения  $I$ . Голос подается пикселем, когда он соответствует некоторому выбранному критерию.

В данной работе рассматривается два критерия: один соответствует MSE метрике, другой расстоянию на основе количества черных пикселей.

Точка считается соответствующей MSE критерию, когда при выбранных параметрах она классифицируется так же, как и на разметке ( $GT$ ). То есть, если  $GT(x, y)$  белый пиксель, и для параметров  $k, a$  не верно (27), то  $H(k', a') = H(k', a') + 1$ , иначе  $H(k', a')$  останется прежним. Аналогичные рассуждения имеют место, когда  $I(x, y)$  помечен на идеальной разметке как черный пиксель.

Точка удовлетворяет критерию на основе количества черных пикселей, когда при выбранных значениях параметров  $k, a$  выполняется (27). То есть, если (27) верно, то  $H(k', a') = H(k', a') + 1$ , иначе  $H(k', a')$  не меняется.

После того, как гистограмма построена, мы можем выбрать параметры  $a_*, k_*$  оптимальные относительно выбранного критерия.

В случае если гистограмма была построена для MSE метрики:

$$(a'_*, k'_*) = \operatorname{argmax}_{\substack{0 \leq a' \leq a'_{max} \\ 0 \leq k' \leq k'_{max}}} H(a', k'). \quad (29)$$

Если же мы работаем с гистограммой, построенной по критерию количества черных пикселей:

$$(a'_*, k'_*) = \operatorname{argmin}_{\substack{0 \leq a' \leq a'_{max} \\ 0 \leq k' \leq k'_{max}}} |H(a', k') - cO|. \quad (30)$$

где  $c$  — количество черных пикселей в соответствующей области разметки,  $O$  — матрица из одних единиц.

Прямое вычисление  $H_{MSE}$  и  $H_{BCM}$  по приведенным формулам имеет трудоемкость порядка  $O(n_w n_h m_w m_h)$ , где  $n_w, n_h$  — ширина и высота изображения  $I$ , а  $m_w$  и  $m_h$  — размеры  $H$ . Если для наглядности принять  $n_w = n_h = n, m_w = m_h = m$ , то сложность составит  $O(n^2 m^2)$ . Заметим, что вообще говоря, рассмотренная процедура есть обычный перебор в пространстве параметров и может легко быть обобщена на произвольный алгоритм бинаризации.

### 5. Экспериментальные результаты

Предложенные методы разметки и настройки были опробованы на нескольких наборах изображений. Во всех опытах, приведенных ниже,  $-4 \leq k \leq 4, -3 \leq a \leq 0$  с шагом 0,01 по обоим параметрам. Эксперименты выполнялись на процессоре Intel Core i5-2450M с частотой 2,50 GHz в режиме conservative под управлением операционной системы Ubuntu Linux.

#### 5.1. Базовые свойства разметки

Покажем здесь несколько тривиальных свойств разметки на примере одной картинке. Будем последовательно добавлять размеченные прямоугольные



Рис. 5. Области разметки: а — исходное изображение; б — размеченное изображение. Цифрами обозначен порядок добавления областей к разметке

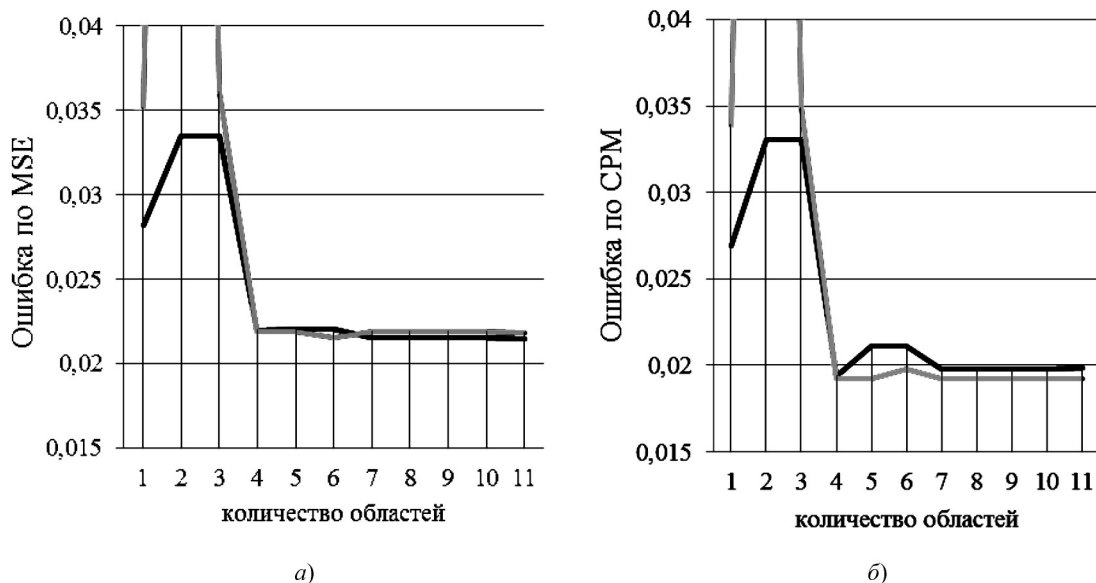


Рис. 6. Зависимость ошибок по CPM и MSE от количества областей. Черная линия соответствует оптимизации, минимизирующей MSE критерий, серая линия — CPM критерий

области (смотрите рис. 5), наблюдая за тем, как будут меняться MSE и CPM результатов бинаризации Ниблэка, обученной на текущей разметке, по сравнению с конечной разметкой по мере добавления областей. Размер окна был взят равным 121 пикселю (смотрите рис. 6).

Как видно из рис. 6, начиная с некоторого количества нанесенной разметки результаты мало меняются. Обратим также внимание на пик, возникающий при оптимизации по CPM критерию при количестве размеченных областей равной 2. Из рис. 7 г явно видно, что деградация возникает в области фона. При добавлении области фона в разметку бинаризация улучшается (смотрите рис. 7 е).

## 5.2. DIBCO2009

Предложенным методом была размечена база DIBCO2009. С помощью программы, поставляемой с DIBCO2011, были рассчитаны ряд характеристик. Выбор DIBCO2009, а не DIBCO2011 был обусловлен присутствием там результатов по алгоритмам Sauvola и Оцу. В то же время, программу расчета расстояний для DIBCO2009 найти не удалось. Размер окна был взят равным 141 пикселю. Результаты эксперимента приведены в таблице 1.

В приведенной таблице «MSE» обозначает настройку алгоритма Ниблэка по MSE критерию, соответственно, «CPM» — по CPM. «base» обозначает, что при обучении в качестве идеальной разметки (GT) использовался результат бинаризации размеченных прямоугольников с указанным человеком порогом, «gef» — для размеченных прямоугольни-

ков в качестве GT использовались данные из предоставленной авторами DIBCO идеальной попиксельной разметки. «abs» обозначает, что для обучения использовалась практически вся площадь изображений (изображения выделялись одним большим прямоугольником) и в качестве GT использовалась идеальная попиксельная разметка.

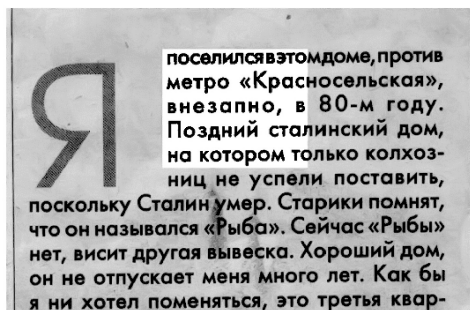
## 5.3. Тестирование на большом наборе с отдельной репрезентативной выборкой

Был проведен эксперимент на наборе из 413 изображений печатных документов, в среднем достаточно хорошего качества. Была сформирована репрезентативная выборка из 15 изображений, на которых были заданы 36 областей с общим количеством пикселей 5 709 395, что составило 13,2% от суммарного количества точек на всех картинках. Разметка заняла 15 минут. Размер окна при обучении был принят равным 121 пикселю.

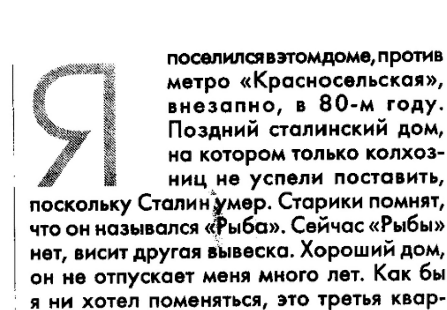
Результаты оптимизации по CPM и MSE и здесь оказались очень близки. Проблемы у них возникли с одними и теми же изображениями, причем, учитывая размер использованной выборки, можно заключить, что данные сложности — функциональное ограничение самого алгоритма бинаризации.

## Заключение

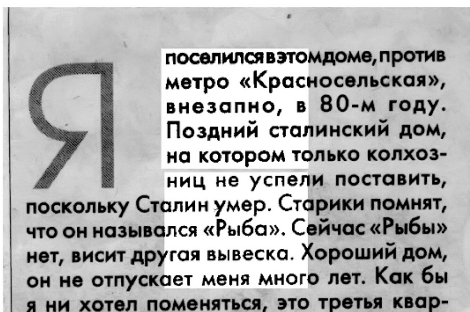
В работе рассматривалась задача оптимального выбора параметров бинаризации на основе обучающей выборки. В ходе проведенных экспериментов



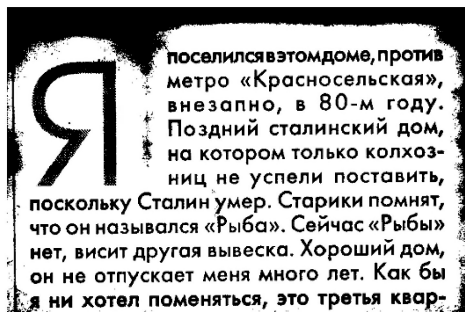
а)



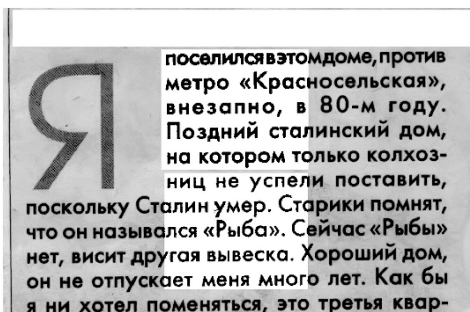
б)



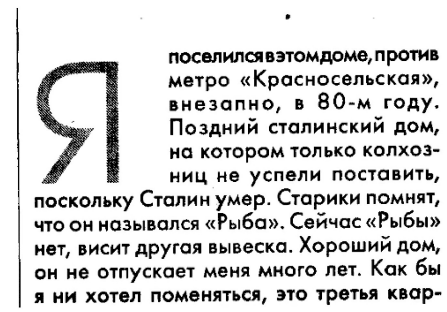
в)



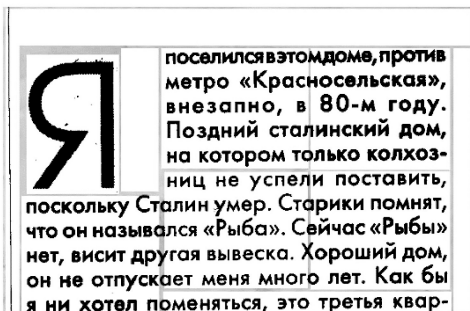
г)



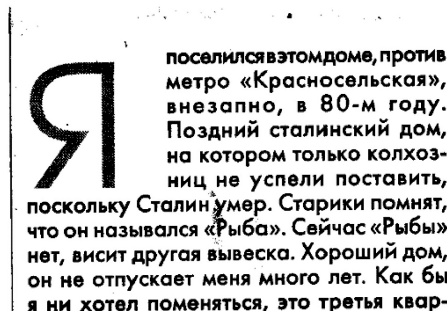
д)



е)



ж)



з)

Рис. 7. Разметка и результат бинаризации алгоритмом Ниблэка с полученными обучением параметрами на указанной разметке: а — 1 область; б —  $a = -0,73, k = 2,67$ ; в — 2 области; г —  $a = 0,37; k = -2,97$ ; д — 3 области; е —  $a = -0,1, k = -0,79$ ; ж — 11 областей; з —  $a = -0,07, k = -0,49$



Таблица 1

Значения критериев качества, использовавшихся в контексте DIBCO, для результатов бинаризации тестовых изображений найденными параметрами

	Recall	Precision	F-measure	PSNR	DRD	MPM	a; k
MSE/abs	84,370	87,976	84,650	16,484	6,045	1,933	-0,21; 0,28
MSE/base	81,431	89,245	83,855	16,027	6,836	2,527	-0,14; -0,32
MSE/ref	83,243	88,621	84,539	16,249	6,448	2,376	-0,16; -0,12
CPM/base	81,408	89,067	83,754	15,985	7,022	2,707	-0,13; -0,39
CPM/ref	85,673	86,189	84,504	16,164	7,660	3,840	-0,11; -0,38
Otsu	...	...	82,170	15,060	...	13,860	...
Sauvola	...	...	87,260	16,690	...	3,390	...

Таблица 2

Результаты работы алгоритма настройки бинаризации на репрезентативной выборке

Критерий минимизации	Общее время, м	Среднее время, с	a; k
MSE	694,66	2778,71	-0,31; 0,97
CPM	273,83	1095,32	-0,32; 1,03

продемонстрирована применимость схемы разметки обучающей выборки набором небольших прямоугольных областей. Показана применимость псевдометрики на основе количества черных пикселей (CPM) для оценки качества бинаризации. Предложен метод настройки произвольного алгоритма бинаризации, рассмотренный подробнее на примере алгоритма Ниблэка.

## Литература

1. Арлазаров В. Л., Славин О. А., Котович Н. В. Адаптивное распознавание // Информационные технологии и вычислительные системы. 2002. Вып. 4. С. 11–23.
2. Миронов А. С., Николаев Д. П. Сравнение вариантов реализации алгоритма Ниблэка бинаризации полутоновых изображений // Информационные технологии и системы (ИТиС'10). М., 2010. С. 138–144.
3. Avci̇bas I., Sankur B., Sayood K. Statistical evaluation of image quality measures // Journal of Electronic Imaging. 2002. № 11. P. 206–223.
4. Badekas E., Papamarkos N. Estimation of appropriate parameter values for document binarization techniques // International Journal of Robotics and Automation. 2009. № 24. V. 24.
5. Chien-Hsing C., Wen-Hsiung L., Fu C. A binarization method with learning built rules for document images produced by cameras // Pattern Recognition. 2009. V. 43. P. 1518–1530.
6. Feng M., Tan Y. Adaptive binarization method for document image analysis // IEEE International Conference. Taipei, 2004. V. 1. P. 339–342.
7. Gatos B., Ntirogiannis K., Pratikakis I. DIBCO 2009: document image binarization contest // Document Analysis and Recognition. Barcelona, 2009. P. 1375–1382.
8. Khurshid K., Siddiqi I., Faure C., e. a. Comparison of Niblack inspired Binarization methods for ancient documents // Document Recognition and Retrieval. 2009. V. XVI.
9. Leedham G., Yan C., Takru K., Nata J. e. a. Comparison of Some Thresholding Algorithms for Text / Background Segmentation in Difficult Document // Images School of Computer Engineering. Washington, 2003. P. 859–864.
10. Niblack W. An Introduction to Digital Image Processing. Englewood Cliffs: PrenticeHall, 1986.
11. Otsu N. A threshold selection method from gray-level histograms // Systems, Man and Cybernetics, IEEE Transactions. 1979. V. 9. P. 62–66.
12. Pratikakis I., Gatos B., Ntirogiannis K., ICDAR 2011 Document Image Binarization Contest (DIBCO 2011) // International Conference on Document Analysis and Recognition. Beijing, 2011. P. 1506–1510.
13. Sauvola J., PietikaKinen M. Adaptive document image binarization // Document Analysis and Recognition. 1997. V. 1. P. 147–152.
14. Sezgin M., Sankur B. Selection of thresholding methods for nondestructive testing applications // International Conference on Image Processing. 2001. V. 3. P. 764–767.

15. *Sezgin M., Sankir B.* Survey over image thresholding techniques and quantitative performance evaluation // *Journal of Electronic Imaging*. 2004. № 13.
16. *Somasundaram K., Kalavathi P.* Medical Image Binarization using Square Wave Representation // *Image Processing Lab. Gandhigram*, 2011. P. 152–158.
17. *Wolf C., Jolion J.-M.* Extraction and Recognition of Artificial Text in Multimedia Documents // *Pattern Analysis and Applications*. 2003. V. 6. P. 309–326.

**Николаев Дмитрий Петрович.** Зав. сектором ИППИ РАН. К. ф.-м. н. Окончил в 2000 г. МГУ им. М. В. Ломоносова. Кол-во печатных работ: 103. Область научных интересов: быстрые алгоритмы обработки изображений. E-mail: [dimonstr@iitr.ru](mailto:dimonstr@iitr.ru)

**Сараев Антон Александрович.** Студент НИТУ МИСиС. Количество печатных работ: 1. Область научных интересов: обработка изображений. E-mail: [antonysaraev@gmail.com](mailto:antonysaraev@gmail.com)