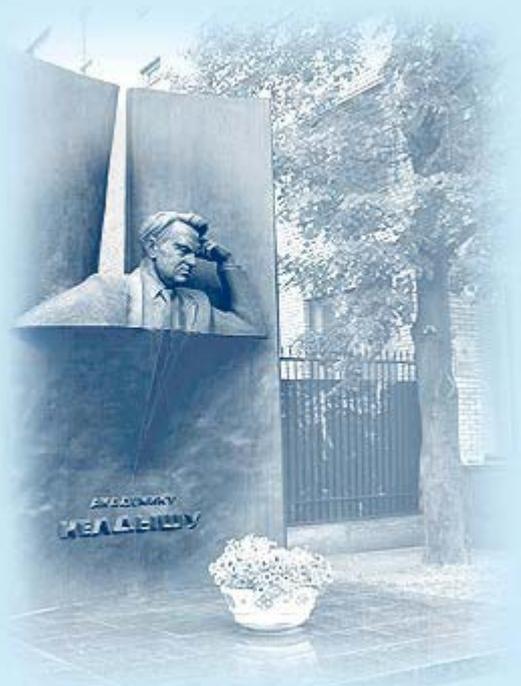




ИПМ им.М.В.Келдыша РАН • [Электронная библиотека](#)

[Препринты ИПМ](#) • [Препринт № 124 за 2019 г.](#)



ISSN 2071-2898 (Print)  
ISSN 2071-2901 (Online)

[Гавриков М.Б.](#), Пестрякова Н.В.

Распознавание символов  
методом полиномиальной  
регрессии

**Рекомендуемая форма библиографической ссылки:** Гавриков М.Б., Пестрякова Н.В. Распознавание символов методом полиномиальной регрессии // Препринты ИПМ им. М.В.Келдыша. 2019. № 124. 14 с. doi:[10.20948/prepr-2019-124](https://doi.org/10.20948/prepr-2019-124)  
URL: <http://library.keldysh.ru/preprint.asp?id=2019-124>

**Ордена Ленина  
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ  
имени М.В.Келдыша  
Российской академии наук**

**М.Б. Гавриков, Н.В. Пестрякова**

**Распознавание символов методом  
полиномиальной регрессии**

**Москва — 2019**

*М.Б. Гавриков, Н.В. Пестрякова*

### **Распознавание символов методом полиномиальной регрессии**

Описывается статистический способ классификации, имеющий вероятностные оценки, в основе которого лежит полиномиально-регрессионный подход. Данный метод обучен и проверен на различных множествах печатных и рукопечатных символов. Он обладает быстродействием, высокой точностью, монотонностью (надежностью) оценок. Его вычислительная структура обеспечивает возможность распараллеливания. Рассматривается проблема устойчивости метода к искажениям.

**Ключевые слова:** классификация, полиномиальная регрессия, печатные и рукопечатные символы

*Mikhail Borisovich Gavrikov,*

*Nadejda Vladimirovna Pestryakova*

### **Recognition of symbols by the method of polynomial regression**

A statistical classification method is described that has probabilistic estimates based on a polynomial-regression approach. This method is trained and tested on various sets of printed and hand-printed characters. It has high speed, high accuracy, monotonicity (reliability) of estimates. Its computational structure provides the possibility of parallelization. A problem of resistance to distortion is considered.

**Key words:** classification, polynomial regression, printed and hand-printed symbols

## **Оглавление**

Введение .....	3
Метод распознавания .....	3
Моделирование искажений изображений символов .....	8
Заключение.....	13
Библиографический список.....	14

## **Введение**

Задача распознавания печатных и рукопечатных символов весьма актуальна для различных видов современных наукоемких технологий, использующих оптический ввод документов.

Существующие классы распознающих систем (программы ввода текстов, системы потокового ввода структурированных документов, видеорегистраторы текстовых объектов) имеют различные стратегии распознавания, но опираются на общие алгоритмы распознавания символов.

С расширением области применения систем распознавания ужесточаются предъявляемые к ним требования. Характеристики качества включают не только традиционные точность и быстродействие, но также свойства оценок распознавания, на основании которых определяется надежность алгоритмов более высокого уровня и всей системы в целом. Большое значение имеет устойчивость метода распознавания к искажениям символов.

Появление и распространение современных многопроцессорных технологий порождает интерес к разработке методов, вычислительная структура которых удобна для распараллеливания.

Несмотря на то что в последние годы поток публикаций по распознаванию символов не только не иссякает, но и даже усиливается, принципиально новые подходы не появляются. В настоящее время наиболее распространенными являются искусственные нейронные сети.

Данная работа посвящена обеспечивающим возможность распараллеливания эффективным приложениям полиномиальной регрессии к задаче распознавания печатных и рукопечатных символов. Описан статистический метод [1-3], генерирующий вероятностные оценки, обученный и проверенный на различных базах символов (печатные прямые буквы и цифры, печатные прямые и курсивные цифры, рукопечатные цифры). Проведен теоретический и численный анализ характеристик качества и свойств разработанного метода при распознавании искаженных изображений.

По точности и быстродействию данный метод практически не уступает искусственным нейронным сетям, а при распознавании загрязненных и частично разрушенных объектов значительно превосходит их.

## **Метод распознавания**

Используется байесовский вероятностный прецедентный подход для задачи классификации на  $K$  непересекающихся классов. Ответом распознавателя считается  $K$ -мерный вектор вероятностей в принадлежности объекта к каждому из классов, по которому можно найти номер класса. Тем самым классификация превращается в специальный случай регрессии.

Разработанный алгоритм позволяет по предъявляемому растру изображения определить, какому символу из некоторого конечного множества

с  $K$  элементами он соответствует. Представлением символа является растр, состоящий из  $N=N_1 \times N_2$  серых пикселей. Перенумеровав все пиксели растра, запоминаем в  $i$ -ой компоненте ( $1 \leq i \leq N$ ) вектора  $\mathbf{v} \in \mathbf{R}^N$  состояние  $i$ -го пикселя (яркость), а именно, значение на отрезке  $[0,1]$  для серого растра. Пусть  $V=\{\mathbf{v}\}$  – совокупность всевозможных растров. Очевидно,  $V \subseteq \mathbf{R}^N$ , причем поскольку пиксели серые, то  $V=[0,1]^N$  –  $N$ -мерный единичный куб в  $\mathbf{R}^N$ .

Отождествим  $k$ -й символ с базисным вектором  $\mathbf{e}_k=(0 \dots 1 \dots 0)$  (1 на  $k$ -м месте,  $1 \leq k \leq K$ ) из  $\mathbf{R}^K$ . Обозначим  $Y=\{\mathbf{e}_1, \dots, \mathbf{e}_K\}$ .

Пусть для предъявляемого растра  $\mathbf{v} \in V$  можно найти  $p_k(\mathbf{v})$  – вероятность того, что растр изображает символ с номером  $k$ ,  $1 \leq k \leq K$ . В рассматриваемой постановке растр не отождествляется с единственным символом. Растр – это изображение, соответствующее каждому из символов заданного конечного набора с некоторой вероятностью, в предположении, что такие вероятности существуют. Тогда результатом распознавания считается символ с порядковым номером  $k_0$ , где

$$p_{k_0}(\mathbf{v}) = \max_{1 \leq k \leq K} p_k(\mathbf{v}). \quad (1)$$

Следовательно, для решения задачи требуется вычислить вектор вероятностей  $(p_1(\mathbf{v}), p_2(\mathbf{v}), \dots, p_K(\mathbf{v}))$ . Он может быть найден на основе метода наименьших квадратов [4]. Приближенные значения компонент  $(p_1(\mathbf{v}), \dots, p_K(\mathbf{v}))$  в соответствии с методом полиномиальной регрессии следует искать в виде разложения по произведениям степеней яркости в различных пикселях:

$$p_k(\mathbf{v}) \cong c_0^{(k)} + \sum_{i=1}^N c_i^{(k)} v_i + \sum_{i,j=1}^N c_{i,j}^{(k)} v_i v_j + \dots, \quad 1 \leq k \leq K. \quad (2)$$

Суммы в (2) конечные и определяются выбором базисных мономов, причем поиск их оптимальной конфигурации (количества и качественного состава) является большой проблемой. Вводится в рассмотрение полиномиальный вектор

$$\mathbf{x}(\mathbf{v}) = (1, v_1, \dots, v_N, \dots)^T.$$

Это конечный вектор размерности  $L$  из приведенных в (2) базисных мономов, упорядоченных некоторым образом и определяющих соответствующее признаковое пространство. Тогда (2) можно записать следующим образом:

$$\mathbf{p}(\mathbf{v}) = (p_1(\mathbf{v}), \dots, p_K(\mathbf{v}))^T \cong A^T \mathbf{x}(\mathbf{v}). \quad (3)$$

Столбцами матрицы  $A$  размера  $L \times K$  являются векторы  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(K)}$ . Каждый такой вектор составлен из коэффициентов при мономах соответствующей

строки (2) с совпадающим индексом  $k$ , упорядоченных так, как это сделано в  $\mathbf{x}(\mathbf{v})$ .

Для вычисления матрицы  $A$  используется обучающая выборка:

$$[\mathbf{v}^{(1)}, \mathbf{y}^{(1)}], [\mathbf{v}^{(2)}, \mathbf{y}^{(2)}], \dots, [\mathbf{v}^{(J)}, \mathbf{y}^{(J)}]. \quad (4)$$

Вектор  $\mathbf{v}^{(j)}$  соответствует растру изображения. Вектор  $\mathbf{y}^{(j)}$  кодирует символ. Все его компоненты нулевые, кроме той, номер которой соответствует номеру символа, – она равна 1.

$$\mathbf{y}^{(j)} = (0, \dots, 1, \dots, 0).$$

Приближенное значение  $A$  ищется следующим образом: (так называемое обучение):

$$A \cong \left( \frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T \right)^{-1} \left( \frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{y}^{(j)})^T \right). \quad (5)$$

Проблема обращения заполненной матрицы большой размерности до сих пор не решена [5]. Чтобы обойти ее, предлагается [6] правую часть (5) вычислить с использованием рекуррентной процедуры:

$$\begin{aligned} A_j &= A_{j-1} - \alpha_j G_j \mathbf{x}^{(j)} [A_{j-1}^T \mathbf{x}^{(j)} - \mathbf{y}^{(j)}]^T, \quad \alpha_j = 1/J, \\ G_j &= \frac{1}{1 - \alpha_j} \left[ G_{j-1} - \alpha_j \frac{G_{j-1} \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T G_{j-1}}{1 + \alpha_j ((\mathbf{x}^{(j)})^T G_{j-1} \mathbf{x}^{(j)} - 1)} \right], \quad 1 \leq j \leq J \end{aligned} \quad (6)$$

$A_0$  и  $G_0$  заданы некоторым образом. Введение матрицы  $G_j$  размера  $L \times L$  помогает избежать обращения матрицы в (5).

Разработанный метод – конкретная алгоритмическая реализация, основанная на описанном математическом подходе к решению задачи. Используется упрощенная модификация процедуры (6):

$$G_j \cong D^{-1}, \quad D = \text{diag} (E\{x_1^2\}, E\{x_2^2\}, \dots, E\{x_L^2\}), \quad 1 \leq j \leq J, \quad (7)$$

где  $x_1, x_2, \dots, x_L$  – компоненты вектора  $\mathbf{x}(\mathbf{v})$ . В этом случае были получены приемлемые практические результаты. Рассматривались серые растры размера  $N=256=16 \times 16$  (рис. 1).

Приведенная математическая постановка задач обучения и распознавания не является полной, поскольку не определен вектор базисных мономов  $\mathbf{x}(\mathbf{v})$ . Достижением данного метода является выбор пространства первичных признаков, определяющего конкретный вид вектора базисных мономов  $\mathbf{x}(\mathbf{v})$ . Его структура оптимизирована по точности распознавания символов обучающей базы графических образов (175 тыс. образов символов для рукопечатных цифр, 1 млн. элементов для печатных букв и цифр), а также

тестовых выборок меньшего объема. Применение обучающих символьных множеств большого объема позволяет избежать проблемы переобучения.

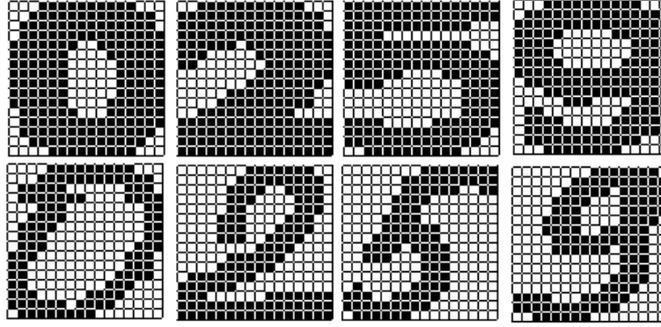


Рис. 1. Образы 16x16 печатных и рукопечатных символов

Предлагаемая в настоящей работе математическая модель распознавания основана на использовании не только самих величин яркости, но и разности их значений в соседних пикселах (перепад яркости), поскольку обнаружение человеком информативных признаков зрительных объектов зависит от фиксации перепадов яркости.

Следует отметить, что разработанная структура вектора не является строго полиномиальной, поскольку в качестве компонент полиномиального вектора фигурируют не только сами мономы, но и линейные комбинации мономов разного порядка, соответствующие различным степеням градиентов яркости по двум ортогональным направлениям ориентации пикселей раstra и их произведениям.

Используются два варианта вектора  $\mathbf{x}$ . Длинный вектор имеет вид:

$$\begin{aligned} \mathbf{x} = & (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_y\}, \{(\delta v_i)_y^2\}, \\ & \{(\delta v_i)_r^4\}, \{(\delta v_i)_y^4\}, \{(\delta v_i)_r(\delta v_i)_y\}, \{(\delta v_i)_r^2(\delta v_i)_y^2\}, \{(\delta v_i)_r^4(\delta v_i)_y^4\}, \\ & \{(\delta v_i)_r((\delta v_i)_r)_l\}, \{(\delta v_i)_y((\delta v_i)_y)_l\}, \{(\delta v_i)_r((\delta v_i)_y)_l\}, \\ & \{(\delta v_i)_y((\delta v_i)_r)_l\}, \{(\delta v_i)_r((\delta v_i)_r)_d\}, \{(\delta v_i)_y((\delta v_i)_y)_d\}, \\ & \{(\delta v_i)_r((\delta v_i)_y)_d\}, \{(\delta v_i)_y((\delta v_i)_r)_d\}). \end{aligned} \quad (8)$$

Короткий вектор составлен из элементов в первой строке (8):

$$\mathbf{x} = (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_y\}, \{(\delta v_i)_y^2\}). \quad (9)$$

Через  $(\delta v_i)_r$  и  $(\delta v_i)_y$  обозначены конечные центральные разности величин  $v_i$  по ортогональным направлениям ориентации раstra – нижние индексы  $r$  и  $y$  соответственно. Так, для горизонтального направления использовалось следующее выражение:

$$\delta v_i = v_{i+1} - v_{i-1}.$$

А для вертикального направления это выражение, соответственно, имело вид:

$$\delta v_i = v_{i+16} - v_{i-16}.$$

Здесь учитывается одномерная нумерация пикселей в растре размером  $16 \times 16$ .

Компоненты вектора  $\mathbf{x}$ , не имеющие индекса  $l$  или  $d$ , вычисляются для всех пикселей растра; с индексом  $l$  – кроме левых граничных, с индексом  $d$  – кроме нижних граничных. Индекс  $l$  (left) или  $d$  (down) означает, что величины относятся к пикселу слева или снизу от данного. Вне растра  $v_i=0$  (используется при вычислении конечных разностей на границе растра).

Поскольку рукопечатные символы имеют меньшую толщину, чем печатные, использовался прием искусственного «уширения» изображения. Суть его заключается в том, что при вычислении компонент полиномиального вектора имеющееся изображение искусственно увеличивается в размере на один пиксел в направлении, ортогональном соответствующему участку границы.

Понятно, что местонахождение границы заранее неизвестно. Поэтому проводилась проверка значения яркости  $v_i$  на превышение порогового значения, равного 0,3. Критерием наличия границы считалось присутствие двух соседних пикселей, таких что для одного из них выполняется условие  $v_i > 0,3$  (это означает, что пиксел условно расположен внутри затемненной области символа), а для другого  $v_j < 0,3$  (пиксел условно находится вне затемненной области символа).

Пороговое значение 0,3 было подобрано в результате численных экспериментов путем оптимизации точности распознавания при обучении и распознавании на фрагментах базы рукопечатных символов.

При вычислении элементов матрицы  $D$  (7) для каждого  $j$ -го элемента базы символов строится вектор  $\mathbf{x}^j$  согласно (8) или (9). Попутно рассчитываются компоненты вспомогательного вектора  $\mathbf{m}^j$  по формуле:

$$m_p^j = (1-1/j) m_p^{j-1} + (x_p^j)^2 / j, \quad j=1, \dots, J, \quad p=1, \dots, L. \quad (10)$$

В конце этой процедуры для последнего элемента имеем согласно (7):

$$G_j \equiv D^{-1} = \text{diag} (1/m_1^j, 1/m_2^j, \dots, 1/m_L^j). \quad (11)$$

После вычисления  $G_j$  для каждого  $j$ -го элемента базы символов строится вектор  $\mathbf{x}^j$  согласно (8) или (9) и находятся элементы матрицы  $A_j$  (5):

$$a_j^{pk} = a_{j-1}^{pk} - \alpha_j x_p^j \left( \sum_{i=1}^L a_{j-1}^{ik} x_i^j - y_k^j \right) / m_p^j, \quad \alpha_j = 1/J, \quad (12)$$

$$A_j = [a_j^{pk}], \quad j=1, \dots, J, \quad p = 1, \dots, L, \quad k = 1, \dots, K.$$

При распознавании по изображению строится вектор  $\mathbf{x}$  согласно (8) или (9). Далее по формуле (4), используя  $A=A_J$  (12), вычисляются оценки для каждого из символов. Затем ищется символ с максимальной оценкой.

Получаемые из-за приближенности метода отрицательные оценки искусственно обнуляли, а превышающие единицу – делали равными 1.

## Моделирование искажений изображений символов

Очевидно, что любое распознаваемое множество может быть получено из обучающего, но неизвестен способ этого преобразования. Для исследования устойчивости метода распознавания к искажениям изображений в качестве базы распознавания используем заданную модификацию базы обучения. Выбор такого универсального способа модификации, который позволил бы получить исчерпывающую информацию о свойствах метода распознавания по отношению к искажениям, причем вне зависимости от обучающего множества, является неразрешенной проблемой. Соответственно, при анализе полученных результатов о чувствительности метода распознавания к искажениям следует делать поправки на конкретный способ модификации и особенности обучающей выборки. Рассмотрены две модели искажения, характеризующиеся увеличением яркости в пикселах (затемнение), а также ее уменьшением (засветление). О важности рассмотрения указанных способов искажения говорит то, что любая модификация исходного изображения представляет собой некоторую композицию затемнения и засветления пикселей.

Моделирование процесса нарастания различия обучающего и распознаваемого множеств проводится следующим образом. При затемнении на этапе распознавания значение яркости для каждого пиксела раstra постепенно увеличивается:  $v_i \rightarrow v_i + 0,01 \cdot n$ , где  $n = 0, 1, \dots, 100$ . Если для каких-то пикселей начиная с некоторого  $n$  имеем  $v_i > 1$ , то считаем  $v_i = 1$ . При засветлении яркость в пикселах уменьшается с ростом  $n$ :  $v_i \rightarrow v_i - 0,01 \cdot n$ , где  $n = 0, 1, \dots, 100$ . Если получено  $v_i < 0$ , то считаем  $v_i = 0$ .

На рис. 2 приведена зависимость от  $n$  доли (в процентном выражении) числа нераспознанных изображений относительно их общего числа (*mis%* – затемнение, *mis%* – засветление). Очевидно, что увеличение  $n$ , соответствующее нарастанию искажения изображений символов, должно приводить к уменьшению числа правильных распознаваний. Однако следует отметить, что особенности динамики этого процесса существенно зависят как от типа написания, так и от направления изменения яркости (ее усиления или ослабления). Для рукопечатных цифр указанная величина при затемнении

растет монотонно, причем на отрезке  $24 \leq n \leq 32$  темпы ее роста стремительно увеличиваются и в дальнейшем остаются высокими. При засветлении рукопечатных в целом наблюдается картина монотонного роста, но на участке  $64 \leq n \leq 72$  имеется небольшая немонотонность. Для печатных  $mis\%t$  и  $mis\%s$  увеличиваются монотонно, причем процесс затемнения характеризуется резким увеличением темпов роста величины  $mis\%t$  при  $48 \leq n \leq 56$ , которые при больших значениях  $n$  еще увеличиваются, аналогично наблюдавшемуся для рукопечатных. Для любого типа написания при засветлении соответствующая кривая является более полой, чем при затемнении. Засветление печатных образов приводит к наименьшему нарастанию доли неправильно распознанных образов по сравнению с засветлением рукопечатных, а также затемнением и печатных, и рукопечатных символов: при  $n = 96$  для печатных цифр  $mis\%s = 2,4\%$ ,  $mis\%t = 82,9\%$ , для рукопечатных  $mis\%t = 81,0\%$ ,  $mis\%s = 46,1\%$ .

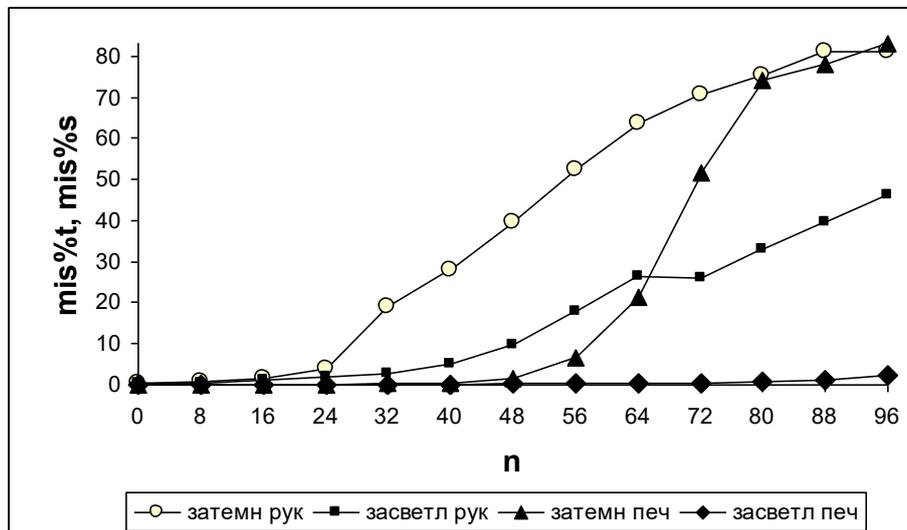


Рис. 2. Доля нераспознанных растров при затемнении и засветлении

Для проведения последующего анализа следует заметить, что важнейшим показателем «правильного» поведения метода является уменьшение оценки при нарастании различия обучающего и распознаваемого множеств.

Средняя оценка распознавания для рукопечатных символов и при затемнении ( $Prb_t$ ), и при засветлении ( $Prb_s$ ) сначала уменьшается, а затем увеличивается (рис. 3). Точка минимума определяет предел достоверности оценок. Для  $Prb_t$  средние темпы и падения, и роста в полтора-два раза выше, чем для  $Prb_s$ . Следует заметить, что минимум  $Prb_t$  достигается при  $n = 32$ , а минимум  $Prb_s$  – при  $n = 64$ , то есть именно на участках особенного поведения величин  $mis\%t$  и  $mis\%s$ . Для печатных цифр  $Prb_s$  стремительно монотонно падает, а  $Prb_t$  при общей тенденции к гораздо более медленному монотонному уменьшению имеет небольшой участок немонотонности (ограничение предела достоверности оценок значением  $n = 72$  определяет локальный максимум при  $n = 80$ ); ранее отмечалось, что здесь также

нарушается гладкость для  $mis\%$ . Итак, пределу достоверности оценки для рукопечатных символов соответствует значение  $n = 32$  при затемнении и  $n = 64$  при засветлении, в то время как для печатных при затемнении ему соответствует значение  $n = 72$ , а при засветлении оценка является «правильной» вплоть до максимальных значений.

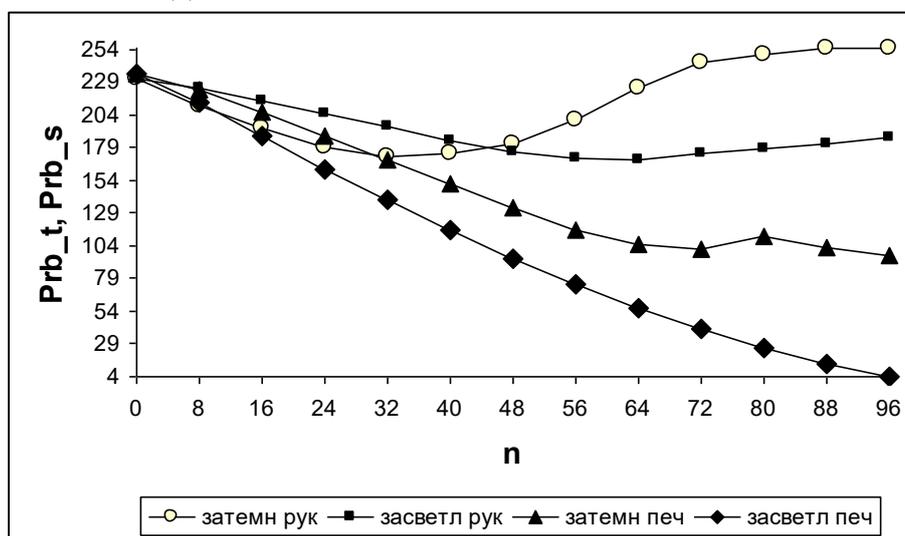


Рис. 3. Средняя оценка распознавания при затемнении и засветлении

Средняя оценка при  $n = 0$  для печатных цифр несколько больше, чем для рукопечатных. Но при засветлении печатных изображений  $Prb_s$  резко падает.  $Prb_s$  при  $n > 0$  меньше  $Prb_t$  для печатных, а при  $n = 3$  становится ниже оценки засветления и при  $n = 12$  – затемнения рукопечатных. Тем не менее, несмотря на падение оценки распознавания при засветлении печатных символов, точность уменьшается медленно. Все это говорит о высокой устойчивости метода к распознаванию печатных символов при уменьшении яркости, а также о достоверности выставляемых оценок.

Зависимость от  $n$  расстояния между среднестатистическими растрами баз обучения и распознавания ( $vv_t$  – затемнение и  $vv_s$  – засветление) приведена на рис. 4, а отклонения между среднестатистическими полиномиальными векторами ( $xx_t$  при затемнении и  $xx_s$  при засветлении) – на рис. 5а, б.

Для любого печатного или рукопечатного символа с ростом  $n$  увеличиваются  $vv_t$  и  $vv_s$  (рис.4). При  $0 \leq n \leq 100$  и для печатных, и для рукопечатных функции  $vv_t$  и  $vv_s$  гладкие и почти линейные, с замедлением роста (увеличивается число пикселей, в которых достигается значение яркости 1 при затемнении и 0 при засветлении). Они располагаются по возрастающей так: засветление (рукопечатные), засветление (печатные), затемнение (печатные), затемнение (рукопечатные). Для  $n > 0$  величина  $vv_t$  превышает  $vv_s$  для рукопечатных цифр более чем вдвое, а для печатных – не более 15%.

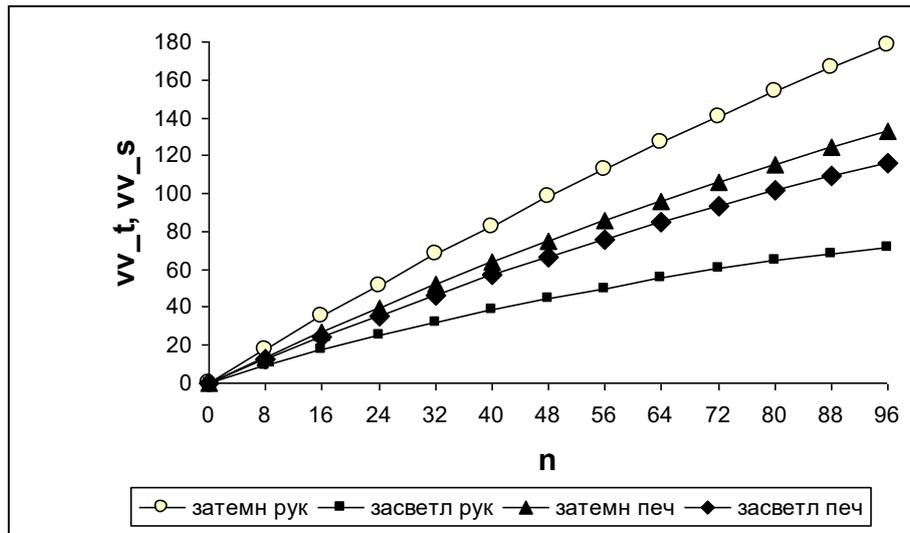


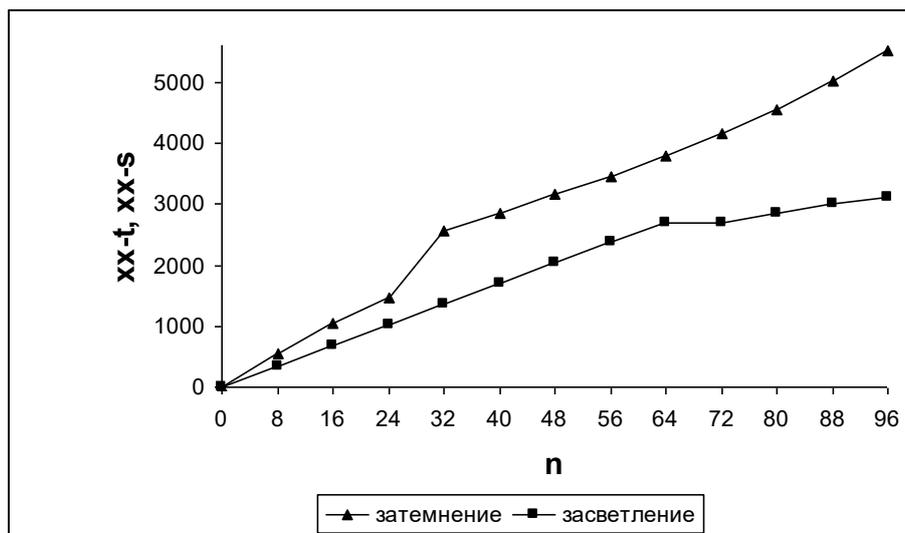
Рис. 4. Расстояние между среднестатистическими растрами баз обучения и распознавания при затемнении и засветлении

С ростом  $n$  для рукопечатных  $xx_t$  строго растет, а  $xx_s$  при  $64 \leq n \leq 72$  немного падает (рис. 5а). Для печатных  $xx_t$  и  $xx_s$  строго растут (рис. 5б).

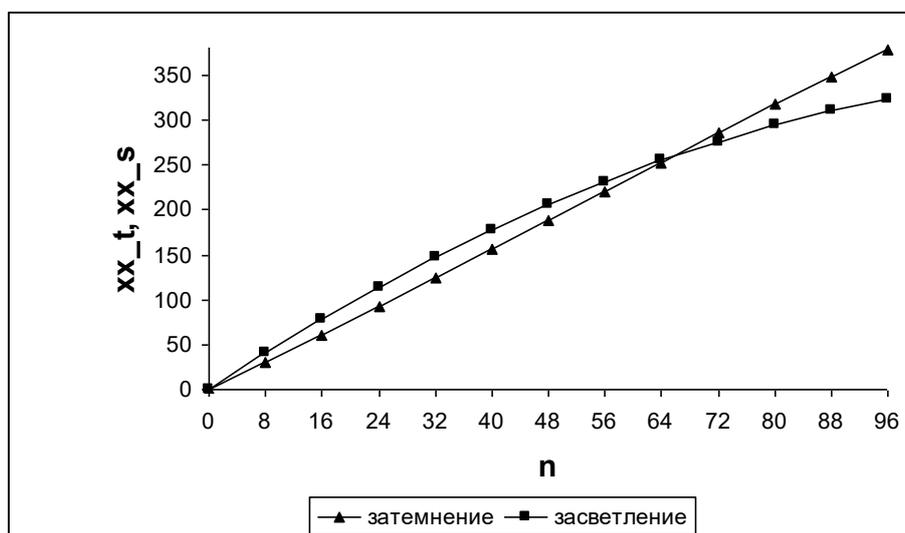
Аналогично  $vv_t$  и  $vv_s$  для  $n > 0$  величина  $xx_t$  более чем в полтора раза превосходит  $xx_s$  для рукопечатных цифр, а для печатных их разница приближается к 20% лишь около  $n = 100$ .

В случае рукопечатных цифр график  $xx_s$  перестраивается при  $64 \leq n \leq 72$  (рис. 5а). Здесь  $xx_s$  незначительно убывает, а левее и правее нарастает с почти постоянной скоростью, которая для  $n > 72$  в два раза ниже, чем для  $n < 64$ . На этом же отрезке перестраивается и график  $mis\%$  (рис. 2). На этом участке  $mis\%$  незначительно убывает, а слева и справа монотонно увеличивается. Справа  $mis\%$  увеличивается с почти постоянной скоростью, которая ниже, чем соответствующая величина вблизи отрезка  $64 \leq n \leq 72$  слева. Итак, для  $n < 64$  и  $n > 72$  нарастанию  $xx_s$  соответствует увеличение  $mis\%$ , а при  $64 \leq n \leq 72$  очень незначительно убывают и  $xx_s$  и  $mis\%$ . При  $64 \leq n \leq 72$  поведение  $Prb_s$  меняется от убывания к нарастанию: при  $n = 64$  имеется минимум  $Prb_s$  (рис. 3), определяющий границу достоверности оценки.

График  $xx_t$  (рис. 5а) – монотонно возрастающий и кусочно-линейный. Производные рвутся при  $n = 24$  и  $n = 32$ . На отрезке  $24 \leq n \leq 32$  темпы роста  $xx_t$  приблизительно втрое больше, чем левее и правее. Здесь же на отрезке перестраивается график  $mis\%$  (рис. 2); темпы роста  $mis\%$  выше, чем слева и справа. Итак, в диапазоне  $0 \leq n \leq 100$  монотонно возрастают и  $xx_t$ , и  $mis\%$ . При  $24 \leq n \leq 32$  темпы их роста намного выше, чем на остальных участках. Именно при  $24 \leq n \leq 32$  поведение  $Prb_t$  коренным образом меняется от убывания к нарастанию: при  $n = 32$  имеется минимум этой величины (рис. 3), являющийся пределом достоверности оценки.



а) рукопечатные цифры



б) печатные цифры

Рис. 5. Расстояние между среднестатистическими векторами баз обучения и распознавания при затемнении и засветлении

Структура вектора для печатных цифр более простая, чем для рукопечатных, но сложнее растрового представления. График оценки распознавания для печатных цифр имеет меньше особенностей, чем для рукопечатных (*печатные*: монотонное убывание для засветления и тенденция к этому для затемнения; *рукопечатные*: выраженная немонотонность (вогнутость) и для затемнения, и для засветления) – рис. 3. Это говорит о большом диапазоне достоверности получаемых оценок при искажении печатных символов (особенно при засветлении) по сравнению с рукопечатными. Поведение доли нераспознанных изображений печатных символов также имеет меньше особенностей (*печатные*: почти монотонный рост, гладкость для засветления и нарушение ее лишь в одной точке для

затемнения; *рукопечатные*: нарушение монотонности для засветления, большее нарушение гладкости для засветления и затемнения) – рис. 2. В отличие от «сложного» рукопечатного вектора, графики отклонения среднестатистического «простого» печатного вектора имеют меньше особенностей: являются гладкими, монотонными – сходны с печатными и рукопечатными растровыми аналогами (рис. 4– рис. 5а, б). Как и последние, они не отражают проявлений немонотонности оценки и нарушения гладкости в графике доли нераспознанных символов (при затемнении печатного образа).

При полном затемнении с  $n = 100$  («черный квадрат») все рукопечатные изображения распознаются как «1» с оценкой 255. Этим и объясняется рост средней оценки  $Prb_t$  для  $n > 32$  при затемнении. Однако при полном засветлении с  $n = 100$  среди цифр от «0» до «9» нет столь уникально подходящей для «белого квадрата». В качестве решения выбираются различные цифры с существенно более низкими оценками, чем при полном затемнении.

Для печатных цифр при полном затемнении все образы, как и для рукопечатных, распознаются как «1», но с меньшей оценкой  $Prb_t = 93$ . Этим обстоятельством с учетом резкого падения темпа роста  $mis\%$  с нарушением гладкости именно в данной точке объясняется незначительное нарушение монотонного убывания  $Prb_t$  при затемнении. При полном засветлении все образы распознаются как «1» с минимальной оценкой  $Prb_s = 1$ , чем и объясняется монотонное убывание  $Prb_s$ .

При  $n = 100$  для рукопечатных цифр выполняются равенства:  $vv_t = 183,8$ , а также  $vv_s = 72,2$ . Следовательно,  $vv_t + vv_s = 256$ . Эта сумма соответствует расстоянию между растрами «белого» и «черного» квадратов. Отклонение среднестатистического растра базы обучения от растра «черного» квадрата в два с половиной раза больше, чем от растра «белого» квадрата. Поэтому  $vv_t$  превышает  $vv_s$  более чем в два раза для каждого  $n > 0$ .

Для печатных цифр при  $n = 100$  выполняются равенства:  $vv_t = 136,6$ , а также  $vv_s = 119,4$ . Следовательно,  $vv_t + vv_s = 256$  (расстояние между растрами «белого» и «черного» квадратов). Отклонение среднестатистического растра базы обучения от растра «черного» квадрата приблизительно на 10% больше, чем от растра «белого» квадрата (для рукопечатных символов – в два с половиной раза). Поэтому для печатных при  $n > 0$  величина  $vv_t$  превышает  $vv_s$  не более чем на 20% (а для рукопечатных – более чем в два раза).

## Заключение

Описан статистический метод распознавания печатных и рукопечатных символов, генерирующий вероятностные оценки. Реализован алгоритм, позволяющий при помощи рекуррентной процедуры обойти проблему обращения матрицы большого размера. Предложена оригинальная конфигурация мономов, позволившая существенно уменьшить размерность признакового пространства и ускорить процедуру обучения. Учтены существенные различия в написании печатных и рукопечатных символов. На

основе разработанной математической модели проведено численное исследование характеристик распознавания, в частности устойчивости к искажениям. Последнее качество объясняет то, что метод хорошо зарекомендовал себя на практике при распознавании сильно зашумленных (загрязненных и в значительной степени разрушенных) изображений.

## **Библиографический список**

1. Гавриков М. Б., Пестрякова Н. В. Метод полиномиальной регрессии в задачах распознавания печатных и рукопечатных символов // Препринты ИПМ им. М.В.Келдыша. 2004. №22. 12 с.

URL: [http://keldysh.ru/papers/2004/prep22/rep2004\\_22.html](http://keldysh.ru/papers/2004/prep22/rep2004_22.html)

2. Гавриков Б.М., Гавриков М.Б., Пестрякова Н.В. Статистический анализ характеристик метода распознавания на обучающем множестве // Информационные технологии и вычислительные системы. 2015, №1. С.82-88.

3. Гавриков Б.М., Гавриков М.Б., Пестрякова Н.В. Статистический метод распознавания и классификации на основе полиномиальной регрессии // Материалы XXI Международной конференции по вычислительной механике и современным прикладным программным системам (ВМСППС'2019), 24–31 мая 2019 г., Алушта. – М.: Изд-во МАИ, 2019. С. 143-145.

4. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. — М.: «Физматлит», 1958.

5. Гавриков М.Б., Локуциевский О.В. Начала численного анализа. — М.: Янус, 1995.

6. Schürmann J. Pattern Classification. — New York: John Wiley&Sons, Inc., 1996.