



Гавриков М.Б., Пестрякова Н.В.,  
Усков А.В., Фарсобина В.В.

Зависимость качества  
распознавания от взаимного  
расположения  
среднестатистических  
растров и векторов баз  
обучения и распознавания

**Рекомендуемая форма библиографической ссылки:** Зависимость качества распознавания от взаимного расположения среднестатистических растров и векторов баз обучения и распознавания / М.Б.Гавриков [и др.] // Препринты ИПМ им. М.В.Келдыша. 2009. № 19. 29 с.  
URL: <http://library.keldysh.ru/preprint.asp?id=2009-19>

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

**М.Б. Гавриков, Н.В. Пестрякова, А.В. Усков, В.В. Фарсобина**

**Зависимость качества распознавания от взаимного  
расположения среднестатистических растров и  
векторов баз обучения и распознавания**

**Москва – 2009**

**М.Б. Гавриков, Н.В. Пестрякова, А.В. Усков, В.В. Фарсобина. Зависимость качества распознавания от взаимного расположения среднестатистических растров и векторов баз обучения и распознавания.** Препринт Института прикладной математики им. М.В. Келдыша РАН, 2009, 29 страницы, 17 рисунков, библиография: 19 наименований.

Для метода распознавания символов, основанного на регрессионном анализе, разработан подход к исследованию зависимости качества распознавания от степени различия между базой обучения и полученной путем ее модификации базой распознавания.

Изучена корреляция динамики среднестатистических величин (растров и полиномиальных векторов), а также качества распознавания в процессе изменения базы рукопечатных цифр.

**M.B.Gavrikov, N.V.Pestryakova, A.V.Uskov, V.V.Farsobina. Dependence of the quality of recognition of mutual arrangement of average rasters and vectors of bases of learning and recognition.** Preprint, Inst. Appl. Mathem., Russian Academy of Sciences, 2009, 29 Pages, 17 Figures, 19 References.

For the character recognition method based on regression analysis, an approach to the examination of the quality of recognition according to the degree of difference between the bases of learning and received through the modification of recognition is developed.

A study of the correlation in the dynamics of average values (rasters and polynomial vectors) as well as the quality of recognition in the process of changing the base of hand-printed digits is carried out.

## Введение

Настоящая работа является очередной в серии [1 – 9], посвященной современному точному методу распознавания символов, основанному на регрессионном анализе [10 – 18]. Разработанный метод обладает высоким качеством распознавания и монотонными оценками [2, 3].

Исследуется зависимость качества распознавания от степени различия между базами обучения и распознавания. Интуитивно понятно, что любая база распознавания представляет собой каким-либо образом «испорченную» базу обучения. На практике бывает трудно определить, существует ли некая закономерность этого «превращения». Было предложено использовать в качестве базы распознавания заданным образом модифицированную базу обучения. Анализируются три модели такого преобразования, а именно, «затемнения», «засветления» и «дискретизации».

При нарастании степени искажения исходной базы рукопечатных цифр изучается динамика оценки распознавания (средней по всем верно распознанным изображениям), а также числа неправильно распознанных образов. Ставится задача выявить особенности корреляции в поведении указанных величин и базирующихся на полиномиальных векторах и растрах [4 – 8] характеристик базы распознавания.

## 1. Метод распознавания

**1.1. Постановка задачи и методика решения.** Задача распознавания символов состоит в разработке алгоритма, позволяющего по растру изображения определить, какому символу из некоторого конечного множества с  $K$  элементами он соответствует. Представлением символа является растр, состоящий из  $N = N_1 \times N_2$  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, а именно, значение на отрезке  $[0, 1]$  для рассматриваемых в работе серых растров.

Пусть для некоторого растра  $\mathbf{v}$  можно найти  $p_k(\mathbf{v})$  – вероятность того, что растр изображает символ с порядковым номером  $k$ ,  $1 \leq k \leq K$ . Тогда распознанным считается символ с порядковым номером  $k_0$ , где

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

Для решения задачи следует вычислить вектор вероятностей  $(p_1(\mathbf{v}), p_2(\mathbf{v}), \dots, p_K(\mathbf{v}))$ . Приближенные значения компонент вектора  $(p_1(\mathbf{v}), \dots, p_K(\mathbf{v}))$  будем искать в виде многочленов от координат  $\mathbf{v} = (v_1, \dots, v_N)$ :

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

Суммы в правых частях равенств (2) конечные и определяются выбором базисных мономов. А именно, если

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

конечный вектор размерности  $L$  из выбранных и приведенных в (2) базисных мономов, упорядоченных определенным образом, то в векторном виде соотношения (2) можно записать так:

$$\mathbf{p}(\mathbf{v}) = (p_1(\mathbf{v}), \dots, p_K(\mathbf{v})) \cong \mathbf{A}^T \mathbf{x}(\mathbf{v}) \quad (2')$$

где  $\mathbf{A}$  – матрица размера  $L \times K$ , столбцами которой являются векторы  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(K)}$ . Каждый такой вектор составлен из коэффициентов при мономах соответствующей строки (2) (с совпадающим верхним индексом), упорядоченных так же, как в векторе  $\mathbf{x}(\mathbf{v})$ . Следовательно, приближенный поиск вектора вероятностей  $\mathbf{p}(\mathbf{v})$  сводится к нахождению матрицы  $\mathbf{A}$ .

Значение  $\mathbf{A}$  вычисляется приближенно в процессе обучения, используя содержащиеся в некоторой базе данных наборы пар векторов  $[\mathbf{v}^{(1)}, \mathbf{y}^{(1)}], \dots, [\mathbf{v}^{(J)}, \mathbf{y}^{(J)}]$  ( $\mathbf{v}^{(j)}$  образ символа с каким-либо номером  $k$  ( $1 \leq k \leq K$ ) и его базисный вектор  $\mathbf{y}^{(j)} = (0 \dots 1 \dots 0)$ , где 1 стоит на  $k$ -м месте,  $1 \leq j \leq J$ ):

$$\mathbf{A} \cong \left( \frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T \right)^{-1} \left( \frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{y}^{(j)})^T \right) \quad (3)$$

При получении правой части (3) используется следующая рекуррентная процедура, где  $\mathbf{A}_0$  и  $\mathbf{G}_0$  заданы:

$$\begin{aligned} \mathbf{A}_j &= \mathbf{A}_{j-1} - \alpha_j \mathbf{G}_j \mathbf{x}^{(j)} [\mathbf{A}_{j-1}^T \mathbf{x}^{(j)} - \mathbf{y}^{(j)}]^T, \quad \alpha_j = 1/J \\ \mathbf{G}_j &= \frac{1}{1 - \alpha_j} \left[ \mathbf{G}_{j-1} - \alpha_j \frac{\mathbf{G}_{j-1} \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T \mathbf{G}_{j-1}}{1 + \alpha_j ((\mathbf{x}^{(j)})^T \mathbf{G}_{j-1} \mathbf{x}^{(j)} - 1)} \right], \quad 1 \leq j \leq J \end{aligned} \quad (4)$$

$$\mathbf{G}_j \equiv \mathbf{D}^{-1}, \quad \mathbf{D} = \text{diag}(E\{x_1^2\}, E\{x_2^2\}, \dots, E\{x_L^2\}), \quad E\{x_i^2\} \cong \frac{1}{J} \sum_{j=1}^J (x_i^{(j)})^2, \quad 1 \leq i \leq L$$

**1.2. Практическая реализация.** Изображения рукопечатных цифр представляли собой серые растры размера  $N = 256 = 16 \times 16$  (Рис.1).

Использовался следующий вектор  $\mathbf{x}$ :

$$\begin{aligned} \mathbf{x} = & (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_y\}, \{(\delta v_i)_y^2\}, \\ & \{(\delta v_i)_r^4\}, \{(\delta v_i)_y^4\}, \{(\delta v_i)_r (\delta v_i)_y\}, \{(\delta v_i)_r^2 (\delta v_i)_y^2\}, \{(\delta v_i)_r^4 (\delta v_i)_y^4\}, \\ & \{(\delta v_i)_r ((\delta v_i)_r)_L\}, \{(\delta v_i)_y ((\delta v_i)_y)_L\}, \{(\delta v_i)_r ((\delta v_i)_y)_L\}, \\ & \{(\delta v_i)_y ((\delta v_i)_r)_L\}, \{(\delta v_i)_r ((\delta v_i)_r)_D\}, \{(\delta v_i)_y ((\delta v_i)_y)_D\}, \\ & \{(\delta v_i)_r ((\delta v_i)_y)_D\}, \{(\delta v_i)_y ((\delta v_i)_r)_D\}) \end{aligned} \quad (5)$$

В (5) выражения в фигурных скобках соответствуют цепочкам элементов вектора, вычисляемым по всем пикселям растра (за исключением указанных ниже случаев). Через  $(\delta v_i)_r$  и  $(\delta v_i)_y$  обозначены конечные центральные разности величин  $v_i$  по ортогональным направлениям ориентации растра – нижние индексы  $r$

и  $y$  соответственно. Если имеется нижний индекс  $L$  (left) или  $D$  (down), то это означает, что соответствующие величины относятся к пикселю слева или снизу от рассматриваемого. Компоненты вектора  $x$ , не имеющие индекса  $L$  или  $D$ , вычисляются для всех пикселей растра; с индексом  $L$  – кроме левых граничных; с индексом  $D$  – кроме нижних граничных пикселей. Вне растра считаем, что  $v_i = 0$  (используется при вычислении конечных разностей на границе растра).

Получаемые оценки могут выходить за рамки отрезка  $[0, 1]$  из-за того, что используемый метод является приближенным. Отрицательные значения искусственно обнулялись, а те, которые были больше 1, делались равными 1.

## **2. Аппарат для анализа метода распознавания символов**

**2.1. Некоторые замечания.** Аппарат анализа вышеприведенного метода распознавания был разработан и впервые описан в работах [4 – 6], где как обучение, так и распознавание осуществлялось на одной и той же графической базе данных, содержащей изображения и коды символов. Это делалось для обеспечения «чистоты эксперимента», поскольку результаты распознавания на различных символьных последовательностях могут сильно отличаться.

Дальнейшее описание подходов к анализу метода распознавания в пп.2.2 – 2.4 будет сопровождаться кратким изложением ранее полученных с помощью этого анализа результатов [4 – 8].

Как в только что указанных публикациях, так и в настоящей работе при обучении и распознавании в качестве вектора  $x$  использовалась модификация (5). После многократного обучения по базе в 174 778 элементов была получена матрица, с которой распознавалось 99,5% элементов (173 897 изображений распознавалось правильно, 881 – неверно).

Результатом распознавания образа является код символа и его целочисленная оценка, лежащая в диапазоне  $[1, 255]$  (оценка 255 является наилучшей). Эта новая оценка получается следующим образом. В результате умножения оценки на 255 старый непрерывный диапазон оценок  $[0, 1]$  (см. (1)) переходит в новый (также непрерывный)  $[0, 255]$ , после чего проводится дискретизация, а именно,  $[0, 1] \rightarrow 1, (1, 2] \rightarrow 2, \dots, (254, 255] \rightarrow 255$ .

**2.2. Растры изображений символов и оценки распознавания.** Введение понятия «среднестатистического» растра (как и «среднестатистического» полиномиального вектора ниже) было обусловлено вызывающим несомненный интерес решением вопроса о том, имеются ли какие-нибудь закономерности в выставлении оценок распознавания изображений символов, или оно носит абсолютно хаотический характер.

Для «среднестатистического» растра конкретного символа по некоторой базе значение в пикселе, имеющем номер  $i$ , равно среднему арифметическому значений  $i$ -х пикселей по всем изображениям данного символа в этой базе.

Расстояние между растрами  $\mathbf{v} = (v_1, \dots, v_N)$  и  $\mathbf{u} = (u_1, \dots, u_N)$  определяется так: вычисляется модуль разности значений в  $i$ -х пикселях, затем производится суммирование по всем  $N$  пикселям:

$$\|\mathbf{v}-\mathbf{u}\| = \sum_{i=1}^N |v_i - u_i| \quad (6)$$

Диапазон отклонений между растрами распознанных верно изображений символа и его «среднестатистическим» растром находится от минимального  $v\_true\_min$  до максимального  $v\_true\_max$  [4 – 8].

Отрезок  $[v\_true\_min, v\_true\_max]$  был поделен на 20 равных частей. Затем для совокупности изображений, попадающих в каждый такой участок, вычислялась средняя оценка распознавания. Для цифр 0, 2, 3, 4, 5, 6, 7, 8, 9 она имеет тенденцию к убыванию, но велик уровень шумов, а для «1» принимает значение 255 на наибольшем удалении от «среднестатистического» растра.

**2.3. Полиномиальные векторы и оценки распознавания.** Для «среднестатистического» полиномиального вектора конкретного символа по некоторой базе значение в каждой компоненте вектора, имеющей номер  $i$ , равно среднему арифметическому значений  $i$ -х компонент по всем изображениям рассматриваемого символа в этой базе.

Расстояние между векторами  $\mathbf{v} = (v_1, \dots, v_L)$  и  $\mathbf{u} = (u_1, \dots, u_L)$  определяется так: вычисляется модуль разности значений в  $i$ -х компонентах и проводится суммирование по всем  $L$  (п.1.1) компонентам:

$$\|\mathbf{v}-\mathbf{u}\| = \sum_{i=1}^L |v_i - u_i| \quad (7)$$

Диапазон отклонений между полиномиальным вектором распознанного верно изображения символа и «среднестатистическим» вектором этого символа по рассматриваемой базе лежит от минимального  $x\_true\_min$  до максимального  $x\_true\_max$  [5 – 8].

Отрезок  $[x\_true\_min, x\_true\_max]$  был поделен на 20 равных частей. Затем для совокупности изображений, имеющих полиномиальные векторы, попадающие в каждый такой участок, вычислялась средняя оценка распознавания. Средняя оценка распознавания для каждого из рассматриваемых символов на соответствующем этому символу отрезке  $[x\_true\_min, x\_true\_max]$  убывает монотонно (с некоторыми шумовыми погрешностями) по мере «удаления» от

«среднестатистического» вектора, а для «1» сначала монотонно убывает, а затем монотонно увеличивается и принимает максимальное значение 255 на предпоследнем интервале удаления от «среднестатистического» вектора (также с некоторыми погрешностями). Уровень шумов в этих зависимостях существенно ниже, а степень монотонности выше, чем в аналогичных зависимостях для средней оценки распознавания при отклонении от «среднестатистического» раstra.

Поскольку для «1» поведение оценки отлично от имеющего место для других цифр, возникла гипотеза о том, что база единиц составлена из двух (или более) подбаз. В результате численного эксперимента удалось разделить базу единиц на две подбазы. Для каждой из них был найден свой «среднестатистический» вектор, при отклонении от которого поведение средней оценки распознавания было аналогичным полученному для символов 0, 2, 3, 4, 5, 6, 7, 8, 9.

**2.4. Структурированное расположение символов.** В этом пункте кратко излагаются некоторые результаты работы [6].

Диапазон отклонений между полиномиальным вектором неправильно распознанного изображения символа и «среднестатистическим» вектором этого символа по рассматриваемой базе находится от минимального  $x\_false\_min$  до максимального  $x\_false\_max$ .

Для произвольного символа  $G$  из рассматриваемых цифр 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 диапазон отклонений между полиномиальными векторами изображений всех символов, кроме  $G$ , и «среднестатистическим» вектором цифры  $G$  по рассматриваемой базе находится от минимального  $\neg x\_min$  до максимального  $\neg x\_max$ .

Отметим, что для каждой из рассматриваемых цифр  $x\_true\_min < \neg x\_min < x\_false\_min$ ; кроме того,  $x\_true\_max < \neg x\_max$ ; а также  $x\_false\_max < \neg x\_max$ .

Следовательно, для каждого символа  $G$  можно выделить полуинтервал  $[x\_true\_min, \neg x\_min)$ , в котором находятся только правильно распознанные изображения этого символа; внутри полуинтервала  $[\neg x\_min, x\_false\_min)$  имеются правильно распознанные изображения символа  $G$  и изображения символов, отличных от  $G$ ; в полуинтервале  $(MAX(x\_true\_max, x\_false\_max), \neg x\_max]$  лежат только изображения символов, отличных от  $G$ ; внутри отрезка  $[x\_false\_min, MIN(x\_true\_max, x\_false\_max)]$  содержатся любые изображения (и правильно, и неправильно распознанные изображения символа  $G$ , а также отличные от  $G$  символы); в полуинтервале  $(MIN(x\_true\_max, x\_false\_max), MAX(x\_true\_max, x\_false\_max)]$  – отличные от  $G$  символы, а также при  $x\_true\_max > x\_false\_max$  (для цифр 0, 1, 3, 5, 6, 8, 9) – правильно, а при



$x\_true\_max < x\_false\_max$  (для 2, 4, 7) – неправильно распознанные изображения символа G.

Итак, при удалении от «среднестатистического» полиномиального вектора любого рассматриваемого символа выявлено структурированное расположение правильно и неправильно распознанных изображений рассматриваемого символа, а также отличных от него символов.

Аналогичная закономерность отсутствует в терминах растров.

**2.5. Особенности настоящей работы.** Аппарат анализа изучаемого метода распознавания [4 – 8] был разработан с целью исследования качества распознавания для каждого из символов в отдельности. В настоящей работе приведены данные по совокупности всех символов. Пусть  $F_k$  – среднее значение какой-либо величины  $F$  для всех  $n_k$  образов  $k$ -го символа ( $k = 1, \dots, K$ ). Тогда средняя по совокупности символов величина  $F_\Sigma$  вычисляется по формуле:  $F_\Sigma = (\sum_{k=1}^K F_k n_k) / (\sum_{k=1}^K n_k)$ . Как показали расчеты, указанные далее закономерности имеют место также и для любого символа. В терминах полиномиальных векторов степень синхронности по различным символам выше, чем для растров.

Таблица 1

	0	1	2	3	4	8	12	16
dv_tw	60,7	59,6	58,8	58,2	57,8	<b>57,2</b>	57,9	59,7
dv_sw	60,7	60,2	59,8	59,3	58,9	57,3	55,8	54,5
dv_t	60,7	59,7	58,9	58,3	57,9	<b>57,2</b>	58,0	59,9
dv_s	60,7	60,3	59,8	59,4	59,0	57,4	55,9	54,7
dv_tb	76,4	75,1	74,0	73,2	72,3	68,8	66,7	<b>66,2</b>
dv_sb	76,4	75,7	75,2	74,7	74,1	72,3	<b>72,2</b>	<b>72,9</b>
vv_t	0	2,3	4,6	6,8	9,1	18,0	27,0	35,2
vv_s	0	1,2	2,3	3,5	4,6	9,0	13,2	17,3
dx_tw	3220	3188	3166	3151	3141	<b>3140</b>	3175	3231
dx_sw	3220	3205	3191	3176	3165	3115	3072	3036
dx_t	3224	3192	3170	3155	3145	<b>3144</b>	3180	3238
dx_s	3224	3209	3195	3180	3169	3120	3078	3046
dx_tb	4064	4020	3980	3956	3919	3783	3671	3593
dx_sb	4064	4047	4037	4027	4012	<b>3966</b>	3986	<b>4052</b>
xx_t	0	76	150	222	292	558	805	1036
xx_s	0	44	86	128	175	344	513	680
mis_t	881	882	891	905	927	1200	1843	2952
mis_s	881	<b>871</b>	884	892	907	1019	1279	1848
Prb_t	231,3	229,1	226,6	224,0	221,5	211,5	202,2	193,9
Prb_s	231,3	230,6	229,8	228,9	228,1	224,1	219,7	215,2

	20	24	32	40	48	56	64	72
dv_tw	62,6	66,2	75,7	87,3	100,3	113,6	126,8	140,2
dv_sw	53,4	52,5	<b>51,6</b>	51,7	52,7	54,2	56,4	59,7
dv_t	62,7	66,4	75,8	87,4	100,3	113,7	127,3	140,8
dv_s	53,7	52,8	<b>52,0</b>	52,2	53,6	55,8	58,5	61,7
dv_tb	67,4	70,3	76,3	87,7	100,3	113,8	127,6	141,0
dv_sb	70,3	67,3	63,4	<b>61,9</b>	62,0	62,9	64,4	67,3
vv_t	43,5	51,6	67,9	82,7	98,6	112,7	127,2	140,7
vv_s	21,2	24,9	32,1	38,7	45,0	50,0	55,8	60,5
dx_tw	3297	3365	3697	3872	4072	4307	4570	4875
dx_sw	3012	<b>2998</b>	3009	3061	3147	3253	3381	3452
dx_t	3304	3373	3720	3904	4115	4362	4651	4988
dx_s	3026	<b>3014</b>	3030	3094	3202	3343	3503	3573
dx_tb	<b>3560</b>	3575	3819	3989	4181	4412	4698	5035
dx_sb	3970	3875	3757	<b>3713</b>	3722	3764	3841	3915
xx_t	1252	1458	2568	2860	3154	3460	3790	4153
xx_s	850	1013	1352	1693	2038	2384	<b>2705</b>	<b>2703</b>
mis_t	4772	6972	33309	49051	69225	91758	110884	123073
mis_s	2523	3223	5077	8835	16860	30896	<b>46397</b>	<b>45692</b>
Prb_t	186,4	179,3	<b>172,5</b>	174,0	182,5	200,5	224,4	244,1
Prb_s	210,3	205,4	194,7	184,2	175,3	170,6	<b>170,1</b>	174,9

	80	88	96	100
dv_tw	154,5	170,4	183,5	189,9
dv_sw	62,8	65,8	68,6	69,7
dv_t	153,8	166,3	178,1	183,8
dv_s	65,1	68,3	71,1	72,2
dv_tb	153,5	165,3	176,9	182,3
dv_sb	70,0	72,0	74,0	74,8
vv_t	153,8	166,3	178,1	183,8
vv_s	64,6	68,2	71,1	72,2
dx_tw	5262	5752	6285	6585
dx_sw	3549	3655	3764	3814
dx_t	5375	5818	6322	6598
dx_s	3689	3808	3922	3974
dx_tb	5412	5833	6331	6600
dx_sb	3974	4040	4107	4142
xx_t	4558	5013	5523	5800
xx_s	2857	2999	3123	3176
mis_t	131767	141501	141617	141617
mis_s	57590	69416	80611	85241
Prb_t	250,5	255,0	255,0	255,0
Prb_s	178,3	182,5	187,1	189,0

### 3. Модели «затемнения» («засветления»)

**3.1. Описание моделей и некоторые замечания.** На этапе распознавания все пиксели растра постепенно затемняются:  $v_i \rightarrow v_i + 0,01 \cdot n$ , 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 аналогично  $v_i \rightarrow v_i - 0,01 \cdot n$ , где  $n = 0, 1, \dots, 100$ . Если для каких-то пикселей начиная с некоторого  $n$  имеем:  $v_i < 0$ , то считаем  $v_i = 0$ .

В Таблице 1 приведены результаты расчетов не для всех целых неотрицательных  $n$ , принадлежащих исследуемому отрезку  $0 \leq n \leq 100$ . Выбор  $n$  определялся условием, чтобы табличные данные достаточно подробно отражали выявленные закономерности (поэтому шаг по  $n$  непостоянен). Графики на Рис.2 – 9 построены с использованием только части результатов из Таблицы 1, а именно, для значений  $n$ , кратных 8. В тексте при ссылке на графики специально оговариваются случаи, когда по сравнению с табличными данными имеется потеря информации (незначительная).

В Таблице 1 утолщением и затемнением выделены экстремальные значения. Понятно, что табличные данные являются огрубленными. Реальные экстремумы могут находиться на шаге левее или правее указанного  $n$  и несколько отличаться от табличного значения.

**3.2. Качество распознавания и интегральные характеристики модифицированной базы. Сравнительный анализ при затемнении и засветлении.** В Таблице 1 для изучаемого диапазона  $0 \leq n \leq 100$  содержатся данные для осредненных по совокупности символов величин. Представлено отклонение между среднестатистическими растрами баз обучения и распознавания ( $vv\_t$  при затемнении и  $vv\_s$  при засветлении) – см. также Рис.7, кроме того, приведено отклонение между среднестатистическими полиномиальными векторами баз обучения и распознавания ( $xx\_t$  при затемнении и  $xx\_s$  при засветлении) – см. также Рис.4. Все четыре величины получены независимо от правильности распознавания, поскольку их следует рассматривать как некие интегральные характеристики для базы распознавания в целом. Именно таким образом (но для каждого из символов в отдельности) среднестатистические растры / векторы были введены в работах [4, 5]. Там же было показано, что эти величины являются ключевыми характеристиками базы (кратко изложено в пп.2.2 – 2.3).

Очевидно, что для произвольного символа как при затемнении, так и при засветлении рост  $n$  сопровождается увеличением расстояния между среднестатистическими растрами в базах обучения и распознавания (среднее по совокупности символов – Рис.7). Для среднестатистических полиномиальных векторов при затемнении сохраняется указанная закономерность, а при засветлении име-

ется небольшой диапазон степеней  $n$ , где это расстояние незначительно уменьшается (среднее по совокупности символов – Рис.4).

В Таблице 1 для  $0 \leq n \leq 100$  приведены число нераспознанных символов ( $mis\_t$  при затемнении и  $mis\_s$  при засветлении) – Рис.3, а также средняя оценка распознавания ( $Prb\_t$  при затемнении и  $Prb\_s$  при засветлении) – Рис.2.

Сравним поведение указанных величин при затемнении и засветлении.

Во всем исследуемом диапазоне ( $0 \leq n \leq 100$ ) функции  $vv\_t$  и  $vv\_s$  гладкие и почти линейные, с небольшим замедлением темпов их роста при увеличении  $n$ ; кроме того, для каждого  $n$  величина  $vv\_t$  более чем в два раза превышает  $vv\_s$  (Рис.7) – объяснение этих фактов дано в п.3.11. В соответствии с указанным, для каждого  $n$  величина  $xx\_t$  более чем в полтора раза превышает  $xx\_s$  (Рис.4). Согласно этим закономерностям, для количества неправильно распознанных символов при любом  $n$  из указанного интервала значение  $mis\_t$  в полтора-два раза больше, чем  $mis\_s$  (Рис.3).

Средняя оценка распознавания и при затемнении, и при засветлении сначала уменьшается, а затем увеличивается (Рис.2). Но для  $Prb\_t$  средние темпы и падения, и роста в полтора-два раза выше, чем для  $Prb\_s$  (см. п.3.5). При затемнении минимум средней оценки распознавания достигается при меньшем  $n$ , чем при засветлении; кроме того, минимальное значение  $Prb\_s$  меньше, чем минимальное значение  $Prb\_t$ , однако их различие незначительно.

**3.3. Корреляция качества распознавания и интегральной характеристики модифицированной базы при засветлении.** На Рис.4. видно, что график  $xx\_s$  «перестраивается» при  $64 \leq n \leq 72$ . На самом этом отрезке  $xx\_s$  убывает, правда, очень незначительно, а левее и правее указанного участка  $xx\_s$  нарастает с почти постоянной скоростью, причем для  $n > 72$  она приблизительно в два раза ниже, чем при  $n < 64$ .

На этом же отрезке  $64 \leq n \leq 72$  «перестраивается» и график, соответствующий изменению количества неправильно распознанных символов  $mis\_s$ . На указанном участке  $mis\_s$  убывает (незначительно, по сравнению с количеством неправильно распознанных символов). Слева и справа от этого отрезка  $mis\_s$  монотонно увеличивается. Темп роста  $mis\_s$  повышается по мере приближения к указанному участку слева. Справа от него  $mis\_s$  увеличивается с почти постоянной скоростью, которая ниже, чем соответствующая величина в непосредственной близости от отрезка  $64 \leq n \leq 72$  слева.

Исключая из рассмотрения локальное и очень небольшое нарушение монотонности величины  $mis\_s$  на начальном участке (минимум при  $n = 1$ ) – см. Таблицу 1, в целом следует отметить наличие следующей закономерности: при  $n <$

64 и  $n > 72$  нарастанию  $xx_s$  соответствует увеличение  $mis_s$ , а при  $64 \leq n \leq 72$  очень незначительно убывают и  $xx_s$ , и  $mis_s$ .

Именно в области  $64 \leq n \leq 72$  поведение средней оценки распознавания  $Prb_s$  коренным образом меняется от убывания к нарастанию: при  $n = 64$  наблюдается минимум этой величины.

**3.4. Корреляция качества распознавания и интегральной характеристики модифицированной базы при затемнении.** На Рис.4. видно, что график  $xx_t$  является монотонно возрастающим и кусочно-линейным. Точки, где меняются углы наклона (рвутся производные), соответствуют  $n = 24$  и  $n = 32$ . На отрезке  $24 \leq n \leq 32$  величина  $xx_t$  нарастает, причем здесь темпы роста приблизительно в три раза больше, чем левее и правее указанного участка (где они примерно одинаковы).

На этом же отрезке  $24 \leq n \leq 32$  радикально «перестраивается» график, соответствующий изменению количества неправильно распознанных символов  $mis_t$ . На указанном участке темпы роста  $mis_t$  выше, чем на любой части отрезка  $0 \leq n \leq 24$  (и приблизительно в десять раз больше, чем в среднем на этом отрезке). При  $32 \leq n \leq 100$  значения  $mis_t$  монотонно увеличиваются, но темпы роста на любой части этого отрезка меньше, чем при  $24 \leq n \leq 32$ . Кроме того, следует отметить, что вблизи  $n = 0$  и  $n = 100$  скорости увеличения  $mis_t$  низкие.

В общем, наблюдается следующая закономерность: в диапазоне  $0 \leq n \leq 100$  монотонно возрастают и  $xx_t$ , и  $mis_t$ . При  $24 \leq n \leq 32$  темпы увеличения указанных величин намного выше, чем на остальных участках.

Именно в области  $24 \leq n \leq 32$  поведение средней оценки распознавания  $Prb_t$  коренным образом меняется от убывания к нарастанию: при  $n = 32$  имеется минимум этой величины.

**3.5. О сильном затемнении и засветлении.** Конечно же, при больших значениях  $n$  нельзя говорить о сколь-нибудь приемлемом распознавании, и результаты носят чисто теоретический характер. Следует отметить, что ситуация в этой области асимметрична для затемнения и засветления. А именно, при полном затемнении с  $n = 100$  («Черный квадрат») все символы распознаются как «1» с максимальной оценкой 255. Этим и объясняется стабильное нарастание средней оценки распознавания  $Prb_t$  для  $n > 32$  при затемнении. Однако, при полном засветлении с  $n = 100$  ситуация не такая замечательная, поскольку среди цифр от 0 до 9 нет столь уникально подходящей для «Белого квадрата». Соответственно, наблюдается нестабильность: в качестве искомого решения выбираются различные цифры с существенно более низкими оценками, чем при полном затемнении. Иллюстрацией к сказанному является Таблица 2, в которой

при  $n = 100$  предлагаемым для распознавания одинаковым изображениям («Белый квадрат») разных символов, а также их совокупности (перечисленные в левом столбце цифры 0, 1, ..., 9 и знак  $\Sigma$ ) приводится в соответствие количество случаев, в которых данное изображение этого символа или их совокупности распознается как один из символов верхней строки 0, 1, ..., 9.

**3.6. Дифференциальные характеристики модифицированной базы.** В Таблице 1 приведены среднее по совокупности символов отклонение растров базы распознавания от среднестатистического растра базы обучения – см. также Рис.8, 9 и, аналогично, среднее по совокупности символов отклонение полиномиальных векторов базы распознавания от среднестатистического вектора базы обучения – см. также Рис.5, 6. Они были введены в работе [6] для символов в отдельности следующим образом. Для каждого изображения некоторого символа определяется отклонение от среднестатистического растра (6) или вектора (7) этого символа, а затем находится среднее значение по всем изображениям этого символа по базе. Далее проводится осреднение в соответствии с определением  $F_{\Sigma}$  в п.2.5.

Эти величины в Таблице 1 даны без учета правильности распознавания – см. также Рис.5, 8, для правильного распознавания – аналогично изображенному на Рис.5, 8 и, наконец, для неправильно распознанных символов – см. также Рис.6, 9. Используемые обозначения представлены в табличном виде:

	затемнение	засветление
растры без учета правильности распознавания	$dv_t$	$dv_s$
растры, правильное распознавание	$dv_{tw}$	$dv_{sw}$
растры, неправильное распознавание	$dv_{tb}$	$dv_{sb}$
векторы без учета правильности распознавания	$dx_t$	$dx_s$
векторы, правильное распознавание	$dx_{tw}$	$dx_{sw}$
векторы, неправильное распознавание	$dx_{tb}$	$dx_{sb}$

Казалось бы, все эти отклонения от среднестатистических растров / векторов с ростом  $n$  должны увеличиваться (аналогично  $vv_t$ ,  $vv_s$ ,  $xx_t$ ) или с некоторыми погрешностями (как  $xx_s$ ). Однако расчеты показали нечто иное.

**3.7. Динамика дифференциальных характеристик модифицированной базы при засветлении.** При усилении засветления  $dv_s$  и  $dv_{sw}$  уменьшаются и достигают минимума при  $n = 32$ , после чего увеличиваются. Для всего рассматриваемого отрезка изменения  $n$  ( $0 \leq n \leq 100$ ) различия  $dv_s$  и  $dv_{sw}$  незначительны, поэтому приведен график только для первой из них (Рис.8). С ростом  $n$  величина  $dv_{sb}$  сначала уменьшается (за исключением участка незначитель-

ного нарастания при  $12 \leq n \leq 16$ ) и достигает минимума при  $n = 40$ , а затем увеличивается (Рис.9). Следует заметить, что при  $0 \leq n \leq 100$  величина  $dv\_sb$  существенно больше, чем  $dv\_s$  и  $dv\_sw$ ; и кроме того, как нетрудно понять, из неравенства  $dv\_sb > dv\_sw$  следует неравенство  $dv\_s > dv\_sw$ .

При анализе поведения аналогичных величин в терминах полиномиальных векторов оказывается, что при усилении засветления  $dx\_s$  и  $dx\_sw$  уменьшаются и достигают минимума при  $n = 24$ , после чего увеличиваются. Для всего рассматриваемого интервала изменения  $n$  ( $0 \leq n \leq 100$ ) различия  $dx\_s$  и  $dx\_sw$  незначительны, поэтому приведен график только для первой из них (Рис.5). Величина  $dx\_sb$  уменьшается (за исключением участка незначительного нарастания при  $8 \leq n \leq 16$ ) и достигает минимума при  $n = 40$ , после чего увеличивается, оставаясь существенно больше, чем  $dx\_s$  и  $dx\_sw$  на всем исследуемом диапазоне изменения  $n$  ( $0 \leq n \leq 100$ ) (Рис.6). Кроме того, из неравенства  $dx\_sb > dx\_sw$  следует неравенство  $dx\_s > dx\_sw$ .

Итак, при усилении засветления поведение величин  $dv\_sb$ ,  $dv\_sw$ , и  $dv\_s$  в значительной степени сходно с поведением  $dx\_sb$ ,  $dx\_sw$ , и  $dx\_s$ .

В качестве критерия того, что одна величина существенно больше другой, а какие-то величины сопоставимы, для определенности использовалось сравнение с 5% от их среднего арифметического.

Таблица 2

	0	1	2	3	4	5	6	7	8	9
0	16645	17	9	1184	2	1721	0	1	0	0
1	0	29428	460	3002	1	176	93	1	0	0
2	0	60	17576	233	0	121	0	0	0	0
3	0	71	491	12462	0	574	44	10	0	0
4	8	2100	3715	5408	1177	455	31	8	0	0
5	0	5	17	4786	0	8310	8	1	0	0
6	3	157	315	10583	0	1869	119	0	0	0
7	2	376	1346	7047	32	3139	400	3809	0	0
8	0	162	2529	6837	0	577	17	1	0	0
9	12	68	4358	19285	0	1277	9	27	0	11
$\Sigma$	16670	32444	30816	70827	1212	18219	721	3858	0	11

**3.8. Корреляция качества распознавания, интегральной и дифференциальных характеристик модифицированной базы при засветлении.** Следует отметить такую закономерность. На отрезке  $64 \leq n \leq 72$  радикально «перестраиваются» графики отклонения между среднестатистическими полиноми-

альными векторами баз обучения и распознавания  $xx_s$ , количества неправильно распознанных символов  $mis_s$ , а также средней оценки распознавания  $Prb_s$  (см. п.3.3). Именно вблизи этого участка после этапа убывания возрастают до своих начальных значений  $dv_{sb}$ ,  $dv_{sw}$ ,  $dv_s$ ,  $dx_{sw}$ ,  $dx_s$ . Более точно, с  $dx_{sw}$ ,  $dx_s$  это происходит с немного меньшими значениями  $n$ , с  $dv_{sw}$ ,  $dv_s$  на самом этом участке, а  $dv_{sb}$  с немного большими значениями  $n$ . Величина  $dx_{sb}$  приближается к исходному значению, но не достигает его даже при  $n = 100$ .

**3.9. Динамика дифференциальных характеристик модифицированной базы при затемнении.** При усилении затемнения расстояния  $dv_t$  и  $dv_{tw}$  сначала уменьшаются и достигают минимума при существенно меньшем значении  $n$  по сравнению с засветлением ( $n = 8$ ), а затем возрастают. Величины  $dv_t$  и  $dv_{tw}$  отличаются очень мало на всем исследуемом диапазоне значений  $n$  ( $0 \leq n \leq 100$ ), поэтому приведен график только для первой из них (Рис.8). С ростом  $n$  величина  $dv_{tb}$  сначала уменьшается и достигает минимума при существенно меньшем значении  $n$  по сравнению с засветлением ( $n = 16$ ), а затем увеличивается (Рис.9). Следует заметить, что при  $0 \leq n \leq 16$  величина  $dv_{tb}$  существенно больше, чем  $dv_t$  и  $dv_{tw}$ ; кроме того, из неравенства  $dv_{tb} > dv_{tw}$  следует неравенство  $dv_t > dv_{tw}$ . При  $n = 20$   $dv_{tb}$  становится сопоставимым с  $dv_{tw}$  и  $dv_t$ , и дальнейшее нарастание  $dv_{tb}$ ,  $dv_t$  и  $dv_{tw}$  идет уже «синхронно», причем если  $dv_{tb} < dv_{tw}$ , то  $dv_t < dv_{tw}$  (при  $n \geq 80$ ), если  $dv_{tb} = dv_{tw}$ , то  $dv_t = dv_{tw}$  (при  $n = 48$ ) а если  $dv_{tb} > dv_{tw}$ , то  $dv_t > dv_{tw}$  (при  $n < 80$ ,  $n \neq 48$ ).

Если проанализировать поведение аналогичных величин в терминах полиномиальных векторов, то оказывается, что при усилении затемнения расстояния  $dx_t$  и  $dx_{tw}$  сначала уменьшаются и достигают минимума при  $n = 8$ , а затем увеличиваются. Величины  $dx_t$  и  $dx_{tw}$  отличаются очень мало на всем исследуемом диапазоне значений  $n$  ( $0 \leq n \leq 100$ ), поэтому приведен график только для первой из них (Рис.5). С ростом  $n$  величина  $dx_{tb}$  сначала уменьшается и достигает минимума при  $n = 20$ , а затем увеличивается (Рис.6). Следует заметить, что при  $0 \leq n \leq 96$  величина  $dx_{tb}$  существенно больше, чем  $dx_t$  и  $dx_{tw}$ ; кроме того, из неравенства  $dx_{tb} > dx_{tw}$  следует неравенство  $dx_t > dx_{tw}$ . При  $n = 100$  значение  $dx_{tb}$  становится сопоставимым с  $dx_{tw}$  и  $dx_t$ , причем  $dx_{tb} > dx_{tw}$ , и следовательно,  $dx_t > dx_{tw}$ .

Следует еще раз подчеркнуть тот факт, что для базы обучения ( $n = 0$ ) среднее отклонение растров неправильно распознанных символов от среднестатистического раstra ( $dv_{tb}$ ) существенно больше, чем среднее отклонение растров правильно распознанных символов от среднестатистического раstra ( $dv_{tw}$ ), но при  $20 \leq n \leq 100$  они уже различаются мало. В то же время для всех



$n$  (кроме  $n = 100$ ) среднее отклонение векторов неправильно распознанных символов от среднестатистического полиномиального вектора ( $dx_{tb}$ ) существенно больше, чем среднее отклонение векторов правильно распознанных символов от среднестатистического вектора ( $dx_{tw}$ ). Выстраивается аналогия с результатами по имеющемуся для векторов и отсутствующему для растров структурированному расположению символов в обучающей базе (см. [6] и п. 2.4 настоящей работы).

В остальном при усилении затемнения поведение величин  $dv_{tb}$ ,  $dv_{tw}$ ,  $dv_t$  в значительной степени сходно с поведением  $dx_{tb}$ ,  $dx_{tw}$ ,  $dx_t$ .

**3.10. Корреляция качества распознавания, интегральной и дифференциальных характеристик модифицированной базы при затемнении.** Следует отметить еще одну важную закономерность. На отрезке  $24 \leq n \leq 32$  радикально «перестраиваются» графики отклонения между среднестатистическими полиномиальными векторами баз обучения и распознавания  $xx_t$ , количества неправильно распознанных символов  $mis_t$ , а также средней оценки распознавания  $Prb_t$  (см. п.3.4). Именно вблизи этого участка после убывания возрастают до своих начальных значений  $dv_{tb}$ ,  $dv_{tw}$ ,  $dv_t$ ,  $dx_{tb}$ ,  $dx_{tw}$ ,  $dx_t$ ; точнее, с  $dv_{tw}$ ,  $dv_t$ ,  $dx_{tw}$ ,  $dx_t$  это происходит с немного меньшими значениями  $n$ , с  $dv_{tb}$  на самом этом участке, а  $dx_{tb}$  с немного большими значениями  $n$ .

**3.11. Расстояние между растрами «Белого» и «Черного» квадратов.** По Таблице 1 видно, что при  $n = 100$  выполняются равенства:  $vv_t = dv_t = 183,8$  (отклонение растра «Черного квадрата» от среднестатистического растра обучающей базы, причем первое равенство легко получить аналитически), а также  $vv_s = dv_s = 72,2$  (отклонение растра «Белого квадрата» от среднестатистического растра обучающей базы, первое равенство получается аналитически). Следовательно,  $vv_t + vv_s = dv_t + dv_s = 256$  (все эти равенства легко получить аналитически). Указанная сумма соответствует расстоянию между растрами «Белого» и «Черного» квадратов.

Итак, отклонение среднестатистического растра базы обучения от растра «Черного» квадрата в два с половиной раза больше, чем от растра «Белого» квадрата. Этим объясняется аналогичное различие в темпах роста  $vv_t$  и  $vv_s$ .

Скорость роста обеих функций замедляется с ростом  $n$ , поскольку при этом увеличивается количество пикселей, в которых достигается значение «1» при затемнении и «0» при засветлении.

## 4. Модель «дискретизации»

**4.1. Описание модели. Представление дискретизации в виде комбинации затемнения и засветления.** В рассматриваемых «серых» растрах для каж-

дого пикселя  $0 \leq v_i \leq 1$ . Поделим отрезок  $[0, 1]$  на 256 равных по длине частей – отрезок и 255 полуинтервалов:  $[0, dv]$ ,  $(dv, 2 \cdot dv]$ , ...,  $(255 \cdot dv, 256 \cdot dv]$ , где  $dv = 1/256$ . Осуществим для всех пикселей раstra следующее преобразование: если  $0 \leq v_i \leq dv$ , то  $v_i \rightarrow dv/2$  (иначе  $[0, dv] \rightarrow dv/2$ ); в полуинтервале  $k \cdot dv < v_i \leq (k+1) \cdot dv$ , где  $k = 1, \dots, 255$ , производится замена:  $v_i \rightarrow (k + 1/2) \cdot dv$  (иначе  $(k \cdot dv, (k + 1) \cdot dv] \rightarrow (k + 1/2) \cdot dv$ ). Тем самым реализуется дискретизация бесконечного множества значений  $0 \leq v_i \leq 1$ , в результате которой  $v_i$  может принимать только 256 значений:  $\{dv/2, (1 + 1/2) \cdot dv, \dots, (255 + 1/2) \cdot dv\}$ . Выполним распознавание полученной базы символов, которая очень незначительно отличается от исходной базы.

Произведем аналогичную дискретизацию с делением отрезка  $[0, 1]$  на разное количество отрезков дискретизации  $N_{\text{дискр}}$ , а именно, на 128 частей, затем на 64 части, далее на 32, на 16, на 8 и, наконец, на 4. От базы к базе количество отрезков дискретизации уменьшалось в 2 раза, Каждая последующая база все больше отличается от исходной.

Нетрудно понять, что дискретизация является комбинацией затемнения и засветления. Для заданного количества отрезков дискретизации  $N_{\text{дискр}}$  получим, какое максимальное  $n$  может быть в рамках моделей «затемнения» и «засветления». В данном анализе будем считать  $n$ , как и раньше, неотрицательным, но не целым, как в п.3, а действительным. От этого не возникает никаких принципиальных расхождений с изложенным в п.3, поскольку там тоже можно было считать  $n$  действительным, и только для простоты использовались целые значения.

В рамках модели «дискретизации» в каждом пикселе раstra выполняется следующее преобразование (для диапазона, внутри которого находится значение в пикселе): либо  $(k \cdot dv, (k + 1) \cdot dv] \rightarrow (k + 1/2) \cdot dv$  для  $k = 1, \dots, N_{\text{дискр}} - 1$ , либо  $[0, dv] \rightarrow dv/2$  для  $k = 0$ , причем  $dv = 1/N_{\text{дискр}}$ . Это можно представить иначе (для значения в пикселе):  $F \cdot dv \rightarrow (k + 1/2) \cdot dv$ , причем либо  $k < F \leq (k+1)$ , для  $k = 1, \dots, N_{\text{дискр}} - 1$ , либо  $0 \leq F \leq 1$  для  $k = 0$ . В рамках модели «затемнения» используется следующее превращение (для значений в пикселе):  $v_i \rightarrow v_i + 0,01 \cdot n$  (для «засветления»  $v_i \rightarrow v_i - 0,01 \cdot n$ ), где  $n = 0, 1, 2, \dots$ . Рассматриваются только такие пиксели, в которых после затемнения  $v_i \leq 1$  и после засветления  $v_i \geq 0$ . Абсолютная величина разницы модифицированного и исходного значений в пикселях для затемнения и засветления при одном и том же  $n$  равны  $0,01 \cdot n$ . Для дискретизации эту величину можно представить так:  $|(k + 1/2) - F| \cdot dv$ , причем либо  $k < F \leq (k + 1)$ , где  $k = 1, \dots, N_{\text{дискр}} - 1$ , либо  $0 \leq F \leq 1$  для  $k = 0$ . Итак,  $0,01 \cdot n = |(k + 1/2) - F| \cdot dv = |(k + 1/2) - F| / N_{\text{дискр}}$ . Получим ограничение для величины в правой части равенства:  $|(k + 1/2) - F| / N_{\text{дискр}} \leq 1/(2 \cdot N_{\text{дискр}})$ . Следовательно,  $0,01 \cdot n \leq 1/(2 \cdot N_{\text{дискр}})$ . Окончательный результат:  $n \leq 50/N_{\text{дискр}}$ . В

Таблице 3 представлены значения величины  $f_{\text{дискр}} = 50 / N_{\text{дискр}}$  в зависимости от  $N_{\text{дискр}}$ .

**4.2. Анализ табличных и графических данных. Согласованность характеристик качества распознавания для дискретизации с соответствующими величинами для затемнения и засветления.** В Таблице 3 содержатся данные для осредненных величин по совокупности символов при  $N_{\text{дискр}} = 256, 128, 64, 32, 16, 8, 4$ . Приведены отклонения между среднестатистическими растрами  $vv$  и между среднестатистическими полиномиальными векторами  $xx$  баз обучения и распознавания. Обе величины получены независимо от правильности распознавания, как и в п. 3. В Таблице 3 также приводятся число нераспознанных символов  $mis$  и средняя оценка распознавания  $Prb$ .

Очевидно, что для произвольного символа уменьшение  $N_{\text{дискр}}$  должно сопровождаться увеличением расстояния между среднестатистическими растрами в базах обучения и распознавания  $vv$  (Рис.15). Для среднестатистических полиномиальных векторов  $xx$  сохраняется указанная закономерность (Рис.12). Заметим, что темпы роста  $vv$  и  $xx$ , незначительные вблизи  $N_{\text{дискр}} = 256$ , нарастают с уменьшением  $N_{\text{дискр}}$  – графики являются вогнутыми.

Особенность Рис.12 – 17 в том, что расстояния на оси абсцисс не соответствуют меткам делений – цифрам 256, 128, 64, 32, 16, 8, 4, которые образуют убывающую геометрическую прогрессию с коэффициентом  $\frac{1}{2}$ . Для того чтобы установить это соответствие, следует длину отрезка [128, 64] сделать в два раза меньше, чем для отрезка [256, 128], длину отрезка [64, 32] сделать в два раза меньше, чем для отрезка [128, 64] и т.д. Это аналогично сжатию вдоль этой оси фрагментов графиков для каждого последующего отрезка по сравнению с предыдущим в два раза. При этом графики станут еще более крутыми.

Можно заменить метки делений  $N_{\text{дискр}}$  на их «предельные» степени затемнения / засветления, а именно,  $f_{\text{дискр}} = 50/N_{\text{дискр}}$ . Расстояния на оси абсцисс не будут соответствовать новым меткам делений – числам 50/256, 50/128, 50/64, 50/32, 50/16, 50/4, которые образуют возрастающую геометрическую прогрессию с коэффициентом 2. Чтобы установить это соответствие, следует длину отрезка [50/128, 50/64] сделать в два раза больше, чем для отрезка [50/256, 50/128], длину отрезка [50/64, 50/32] сделать в два раза больше, чем для отрезка [50/128, 50/64] и т.д. Это соответствует растяжению вдоль этой оси фрагментов графиков для каждого отрезка по сравнению с предыдущим в два раза. Графики станут более пологими и сходными с полученными при затемнении / засветлении.

Дальнейший анализ проводится для шкалы, изображенной на Рис.12-17. Для линейной по  $N_{\text{дискр}}$  шкалы он качественно не изменится, а скорости или

темпы роста, о которых будет говориться далее, увеличатся. Сравнение с затемнением и засветлением будет проведено отдельно и более точно, с использованием табличных данных.

Рост (с ускорением)  $vv$  (Рис.15) и  $xx$  (Рис.12) при уменьшении  $N_{\text{дискр}}$  соответствует уменьшению средней оценки распознавания  $P_{rb}$  (Рис.10), причем темпы убывания последней также увеличиваются с уменьшением  $N_{\text{дискр}}$ . При этом количество неправильно распознанных символов  $mis$  (Рис.11) с уменьшением числа отрезков дискретизации от 256 до 128 – очень незначительно увеличивается, от 128 до 64 – уменьшается, а при дальнейшем падении числа отрезков дискретизации от 64 до 4 – увеличивается с нарастающими темпами.

Отметим, что указанные в предыдущем абзаце графики качественно соответствуют аналогичным для затемнения и засветления с диапазоном степеней  $n$ :  $0 \leq n \leq 50/4$ .

Величина  $mis$  достигает минимального значения при  $N_{\text{дискр}} = 64$ , для которого  $f_{\text{дискр}} = 0,8$ . Ограничение на степени затемнения / засветления будет следующим:  $n \leq 0,8$ . По Таблице 1 видно, что при  $n = 1$  для засветления также зафиксирован близкий по значению минимум  $mis$ . Нетрудно проверить, что для любого  $N_{\text{дискр}}$  из всего исследуемого набора интервалов дискретизации выполняется следующее: отклонение между среднестатистическими растрами / полиномиальными векторами баз обучения и распознавания, количество нераспознанных символов и средняя оценка распознавания лежат в диапазоне между соответствующими результатами для затемнения и засветления при степени  $n=f_{\text{дискр}}$  (или близки к этим результатам).

В Таблице 3 приведены среднее отклонение растров базы распознавания от среднестатистического растра базы обучения (Рис.16, 17), а также среднее отклонение полиномиальных векторов базы распознавания от среднестатистического вектора базы обучения (Рис.13, 14). Эти величины в Таблице 3 даны без учета правильности распознавания (Рис.13, 16), для правильно распознанных (аналогично Рис.13, 16) и для неправильно распознанных символов (Рис.14, 17). Используемые обозначения представлены в табличном виде:

растры без учета правильности распознавания	$dv$
растры, правильное распознавание	$dvw$
растры, неправильное распознавание	$dvb$
векторы без учета правильности распознавания	$dx$
векторы, правильное распознавание	$dxw$
векторы, неправильное распознавание	$dxb$

Расчеты показали, что при уменьшении  $N_{\text{дискр}}$  расстояния  $dvb$ ,  $dvw$ ,  $dv$ ,  $dxb$ ,  $dxw$ ,  $dx$  монотонно уменьшаются с увеличивающимися темпами. На всем исследуемом диапазоне значений  $N_{\text{дискр}}$  ( $4 \leq N_{\text{дискр}} \leq 256$ ) имеют место следующие закономерности: величины  $dv$  и  $dvw$  отличаются очень мало; величина  $dvb$  существенно больше, чем  $dv$  и  $dvw$  (аналогично затемнению и засветлению для  $0 \leq n \leq 12,5$ ); из неравенства  $dvb > dvw$  следует неравенство  $dv > dvw$ . В терминах полиномиальных векторов величины  $dx$  и  $dxw$  отличаются очень мало; величина  $dxb$  существенно больше, чем  $dx$  и  $dxw$  (аналогично затемнению и засветлению для  $0 \leq n \leq 12,5$ ); из неравенства  $dxb > dxw$  следует неравенство  $dx > dxw$ .

Таблица 3

	256	128	64	32	16	8	4
dvw	60,3	60,0	59,4	58,4	56,7	54,1	51,0
dv	60,4	60,1	59,5	58,5	56,7	54,1	51,1
dvb	75,9	75,7	75,1	73,3	70,6	66,3	59,1
dxw	3210	3199	3181	3156	3106	3022	2965
dx	3214	3203	3185	3160	3110	3027	2972
dxb	4051	4041	4025	3972	3887	3725	3438
vv	0,25	0,5	1,0	2,0	4,0	8,0	16,3
xx	10,1	22,2	41,3	82,2	155,2	310,7	563,5
mis	875	<b>879</b>	<b>868</b>	881	915	1081	2906
Prb	230,7	230,2	229,0	226,6	221,1	209,3	189,3
$f_{\text{дискр}}$	0,2	0,4	0,8	1,6	3,1	6,3	12,5

### Выводы

«Превращение» базы обучения в базу распознавания формализовано в виде трех моделей, а именно, «затемнения», «засветления» и «дискретизации».

При увеличении степени затемнения / засветления расстояние между среднестатистическими растрами баз обучения и распознавания увеличивается почти линейно, с небольшим падением темпов роста.

Расстояние между среднестатистическими векторами изменяется кусочно-линейно, график состоит из трех частей, причем в крайних частях наблюдается монотонное возрастание, а в центральной при затемнении – стремительный рост, а при засветлении – незначительное убывание.

Количество неправильно распознанных изображений изменяется аналогично расстоянию между среднестатистическими векторами (будучи при этом не-

линейной функцией); исключением является небольшое уменьшение числа нераспознанных символов на начальном этапе засветления.

Для любой степени затемнения каждая из этих величин больше, чем для того же уровня засветления.

Средняя оценка распознавания и при затемнении, и при засветлении сначала убывает, а затем возрастает (при затемнении скорости больше). Минимумы приблизительно равны по величине и приходятся на центральные части двух указанных ранее зависимостей.

Посчитанные для базы обучения средние отклонения растров / векторов от среднестатистических растров / векторов «очерчивают» вокруг последних некоторую окрестность. Для засветления, как и для затемнения при переходе через эту окрестность «перестраиваются» графики отклонения между среднестатистическими полиномиальными векторами баз обучения и распознавания, количества неправильно распознанных символов, а также средней оценки распознавания.

Показано, что дискретизацию можно представить в виде комбинации затемнения и засветления. Интегральные характеристики модифицированной базы и величины, определяющие качество распознавания, ведут себя согласованно: отклонение между среднестатистическими растрами / полиномиальными векторами баз обучения и распознавания, количество нераспознанных символов и средняя оценка распознавания при дискретизации лежат в диапазоне между соответствующими результатами для затемнения и засветления (или близки к этим результатам).

### Литература

- [1] *Гавриков М.Б., Пестрякова Н. В.* "Метод полиномиальной регрессии в задачах распознавания печатных и рукопечатных символов", //Препринт ИПМатем. РАН, М., 2004, №22, 12 стр.
- [2] *Гавриков М.Б., Пестрякова Н. В., Славин О.А, Фарсобина В.В.* "Развитие метода полиномиальной регрессии и практическое применение в задаче распознавания", //Препринт ИПМатем. РАН, М., 2006, №25, 21 стр.
- [3] *Гавриков М.Б., Мисюрев А.В., Пестрякова Н.В., Славин О.А.* Об одном методе распознавания символов, основанном на полиномиальной регрессии. Автоматика и Телемеханика. 2006, №2, с. 119-134.
- [4] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "О некоторых свойствах метода распознавания символов, основанного на полиномиальной регрессии", //Препринт ИПМатем. РАН, М., 2007, №69, 20 стр.
- [5] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "О некоторых свойствах оценки метода распознавания символов, основанного на полиномиальной регрессии", //Препринт ИПМатем. РАН, М., 2008, №7, 28 стр.

- [6] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "Анализ метода распознавания символов, основанного на полиномиальной регрессии", //Препринт ИПМатем. РАН, М., 2008, №25, 28 стр.
- [7] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "Оценка распознавания символов для метода, основанного на полиномиальной регрессии", // Труды ИСА РАН под ред. чл.-корр. РАН, проф. Арлазарова В.Л. и д.т.н., проф. Емельянова Н.Е. М., «ЛИБРОКОМ», 2008, с. 194-219.
- [8] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "О средне-статистических растрах и векторах метода распознавания символов, основанного на полиномиальной регрессии", // Труды ИСА РАН под ред. чл.-корр. РАН, проф. Арлазарова В.Л. и д.т.н., проф. Емельянова Н.Е. М., «ЛИБРОКОМ», 2008, с. 220-232.
- [9] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "Зависимость точности и оценки распознавания от степени различия между базами обучения и распознавания", // Труды ИСА РАН под ред. чл.-корр. РАН, проф. Арлазарова В.Л. и д.т.н., проф. Емельянова Н.Е. М., «ЛИБРОКОМ», 2008, с.233-240.
- [10] *Sebestyen G.S.* Decision Making Processes in Pattern Recognition, MacMillan, New York, 1962.
- [11] *Nilson N. J.* Learning Machines, McGraw-Hill, New York, 1965.
- [12] *Schürmann J.* Polynomklassifikatoren, Oldenbourg, München, 1977.
- [13] *Schürmann J.* Pattern Classification, John Wiley&Sons, Inc., 1996.
- [14] *Albert A.E. and Gardner L.A.* Stochastic Approximation and Nonlinear Regression // Research Monograph 42. MIT Press, Cambridge, MA, 1966.
- [15] *Becker D. and Schürmann J.* Zur verstärkten Berücksichtigung schlecht erkennbarer Zeichen in der Lernstichprobe // Wissenschaftliche Berichte AEG-Telefunken **45**, 1972, pp. 97 – 105.
- [16] *Pao Y.-H.* The Functional Link Net: Basis for an Integrated Neural-Net Computing Environment // in Yoh-Han Pao (ed.) Adaptive Pattern Recognition and Neural Networks, Addison-Wesley, Reading, MA, 1989, pp. 197-222.
- [17] *Franke J.* On the Functional Classifier, in Association Francaise pour la Cybernetique Economique et Technique (AFCET), Paris // Proceedings of the First International Conference on Document Analysis and Recognition, St. Malo, 1991, pp.481-489.
- [18] *Дж.Себер.* Линейный регрессионный анализ. М.:”Мир”, 1980.
- [19] *Линник Ю.В.* Метод наименьших квадратов и основы математико - статистической теории обработки наблюдений. М.:”Физматлит”, 1958.

## Рисунки

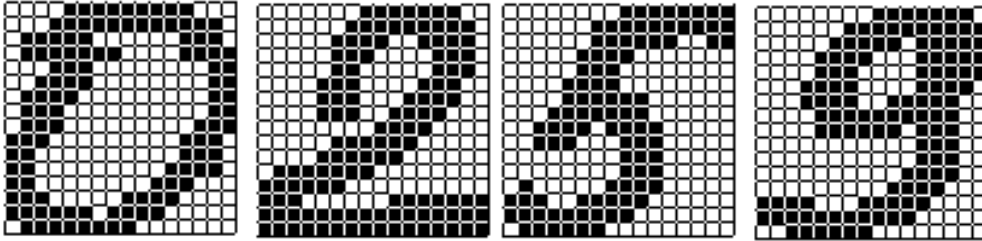


Рис. 1. Образы 16x16 рукопечатных цифр

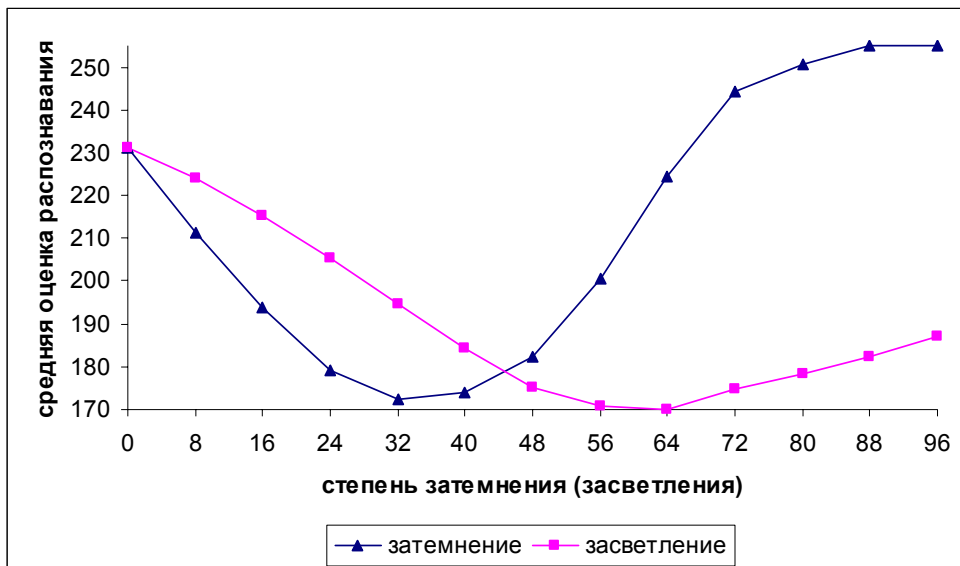


Рис 2.

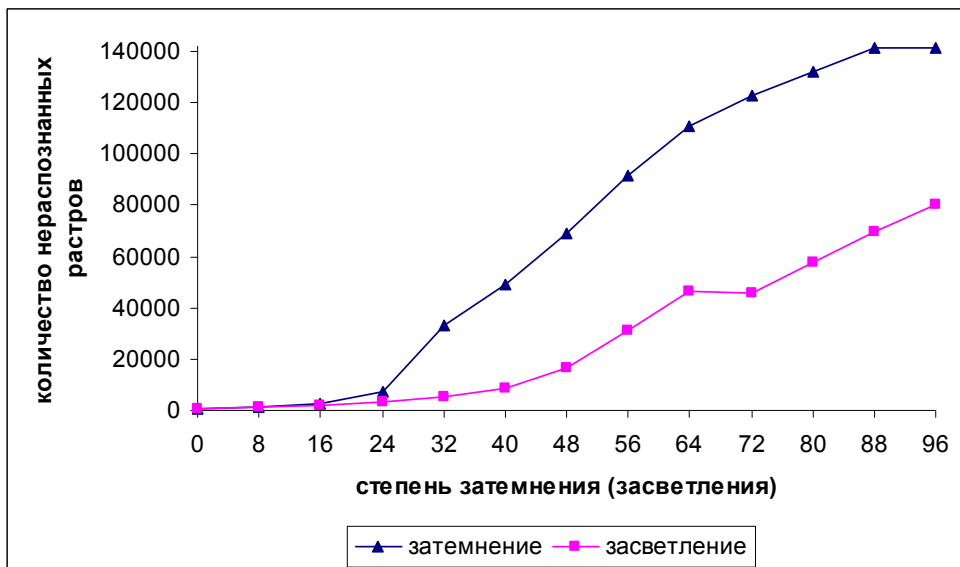


Рис 3.



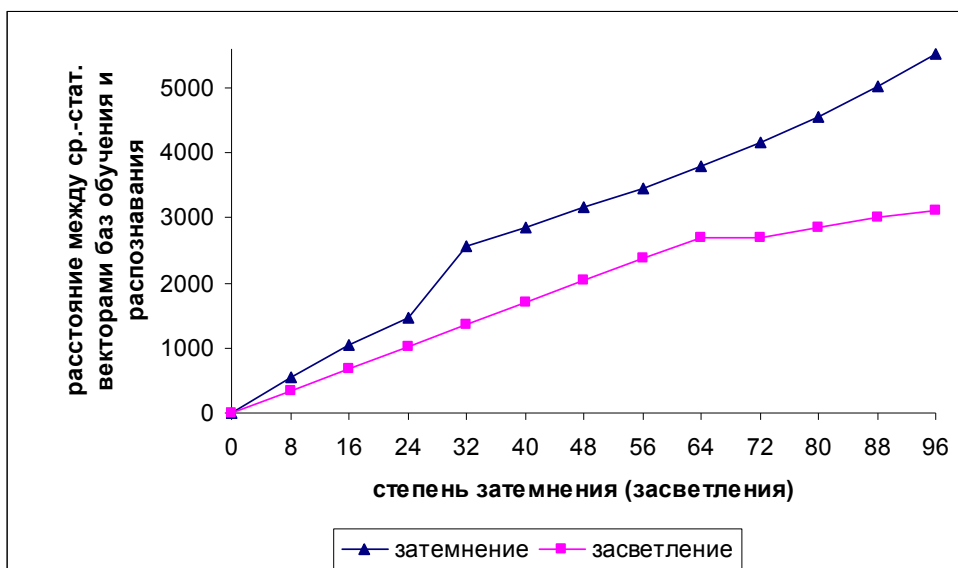


Рис 4.

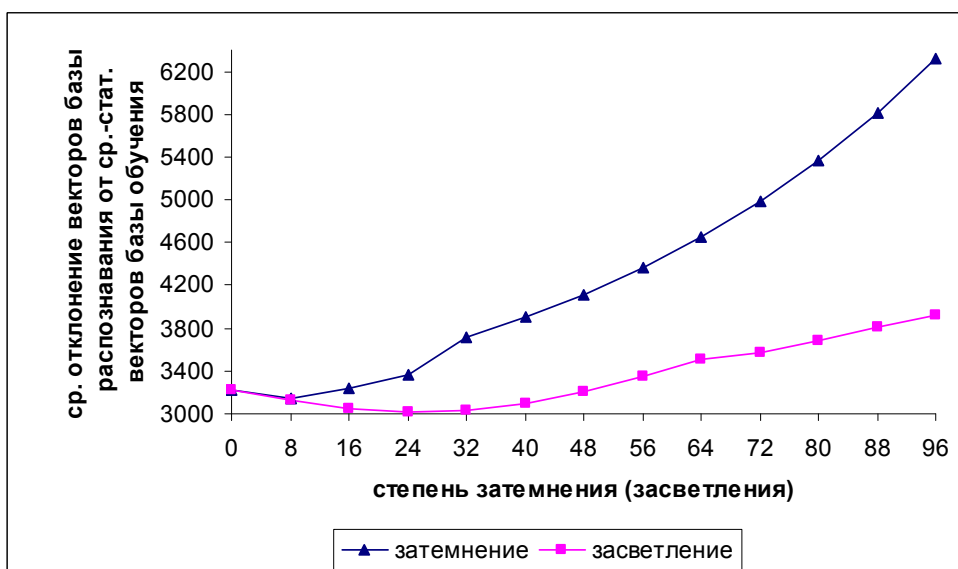


Рис 5.

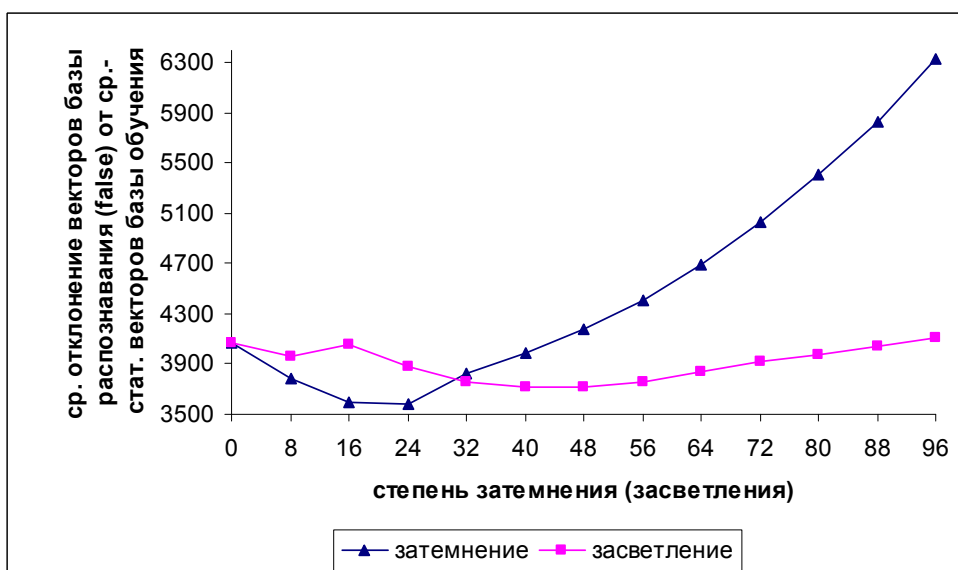


Рис 6.

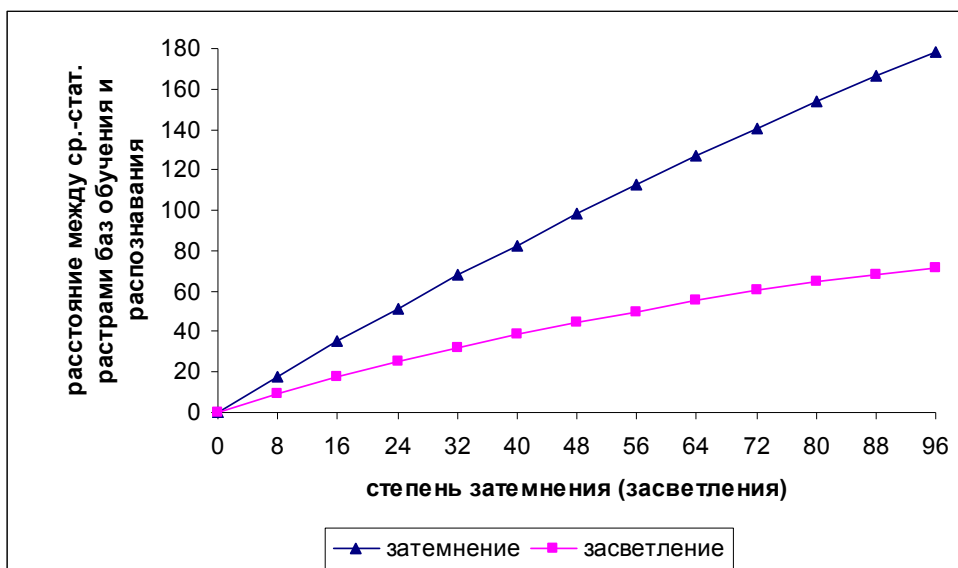


Рис 7.

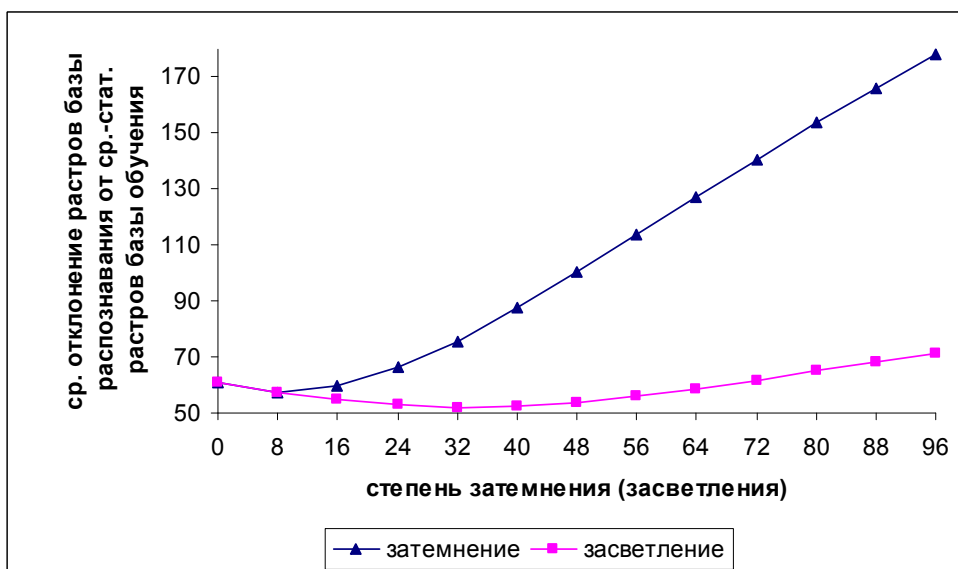


Рис 8.

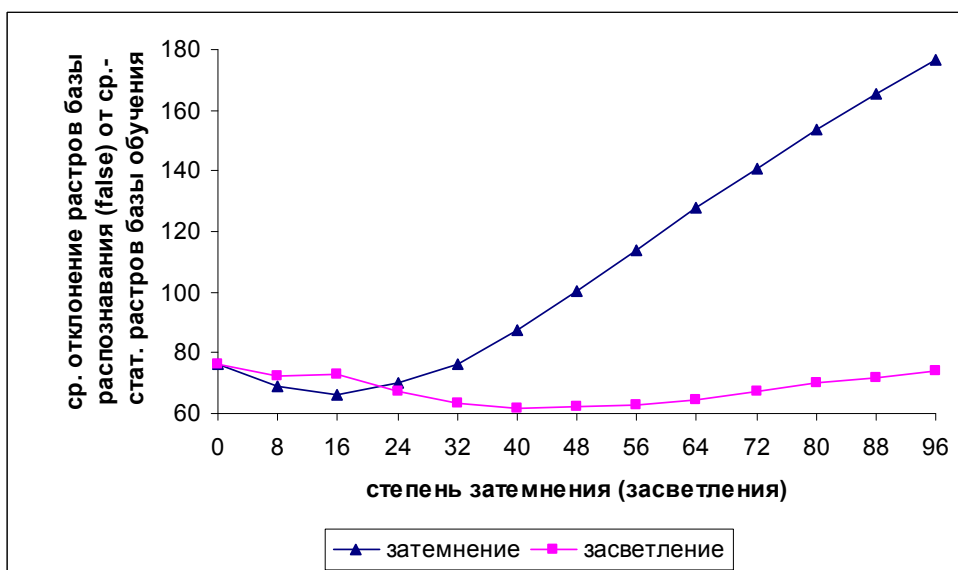


Рис 9.

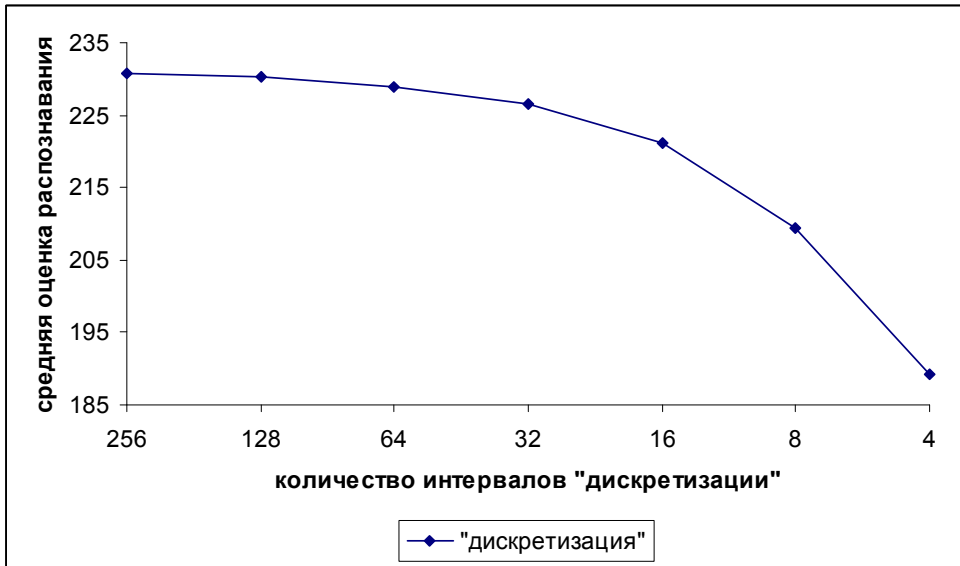


Рис 10.

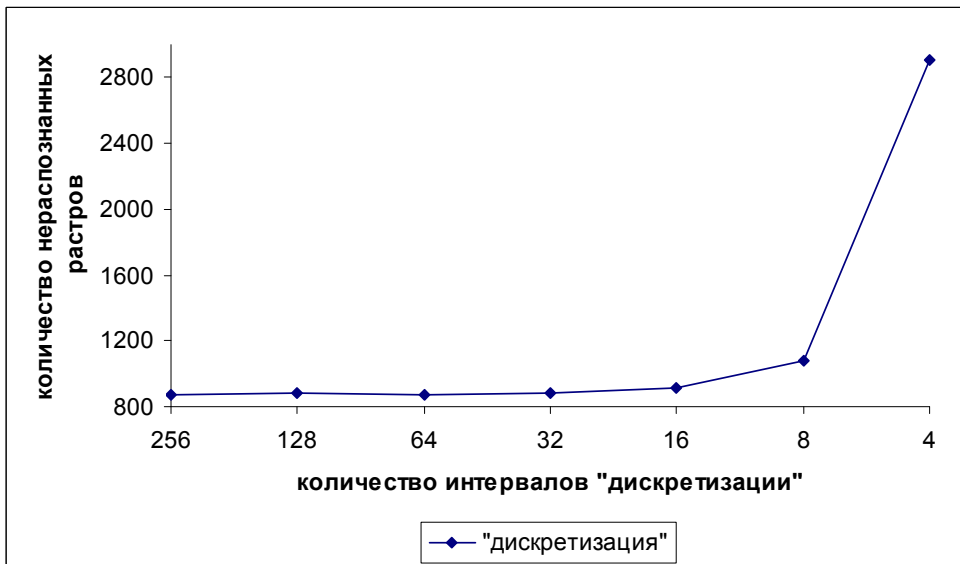


Рис 11.

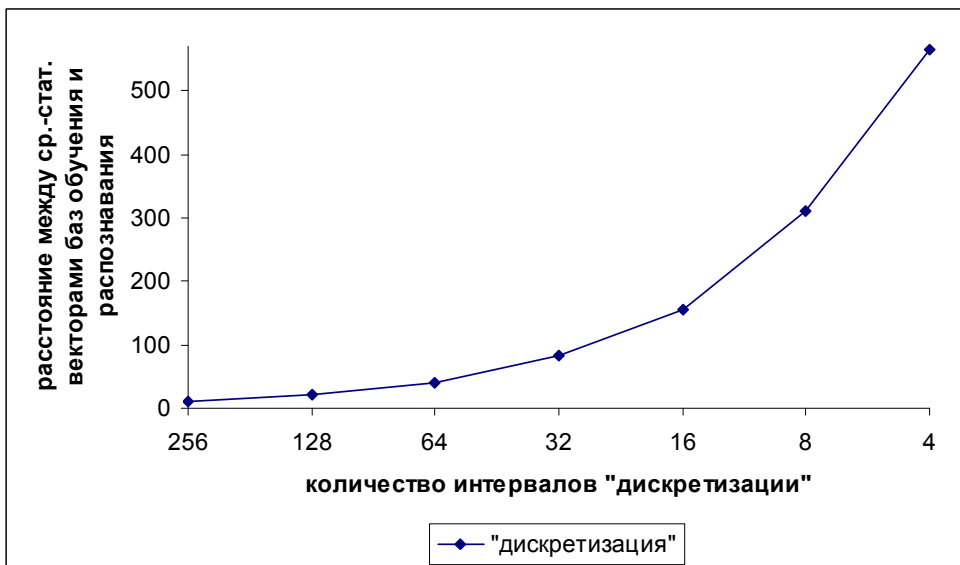


Рис 12.

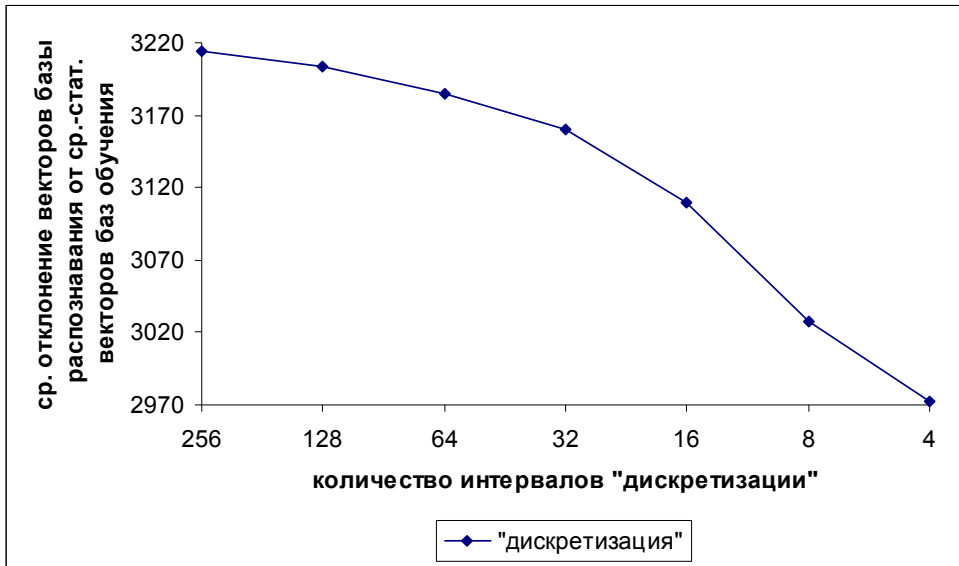


Рис 13.

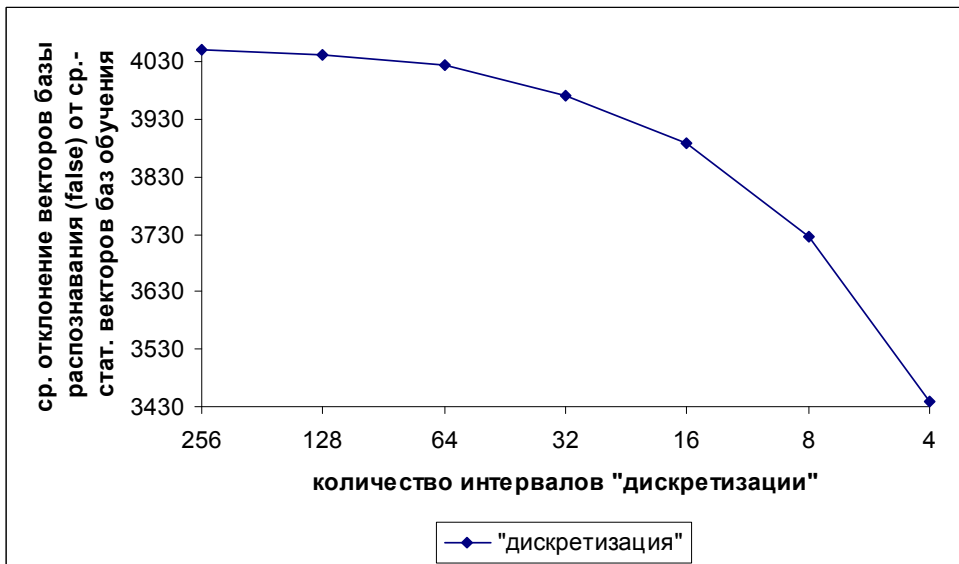


Рис 14.

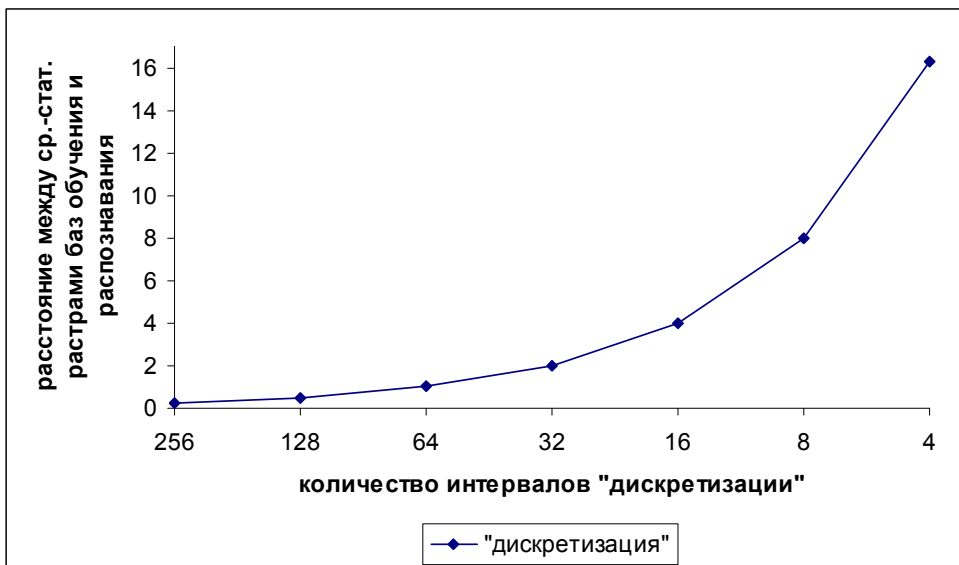


Рис 15.

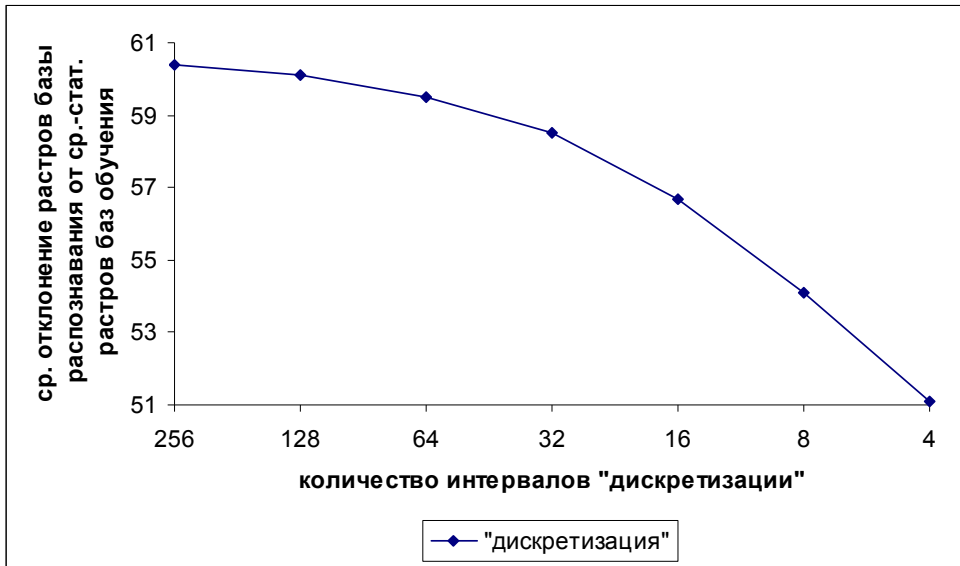


Рис 16.

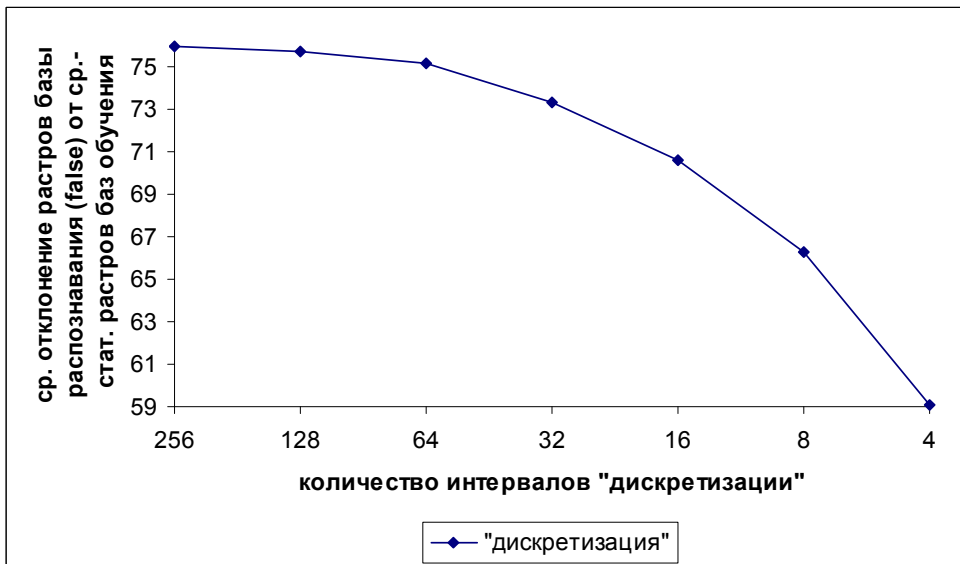


Рис 17.