



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 9 за 2010 г.



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

О распознавании
модифицированной
обучающей базы (Часть 1)

Рекомендуемая форма библиографической ссылки: О распознавании модифицированной обучающей базы (Часть 1) / М.Б.Гавриков [и др.] // Препринты ИПМ им. М.В.Келдыша. 2010. № 9. 28 с. URL: <http://library.keldysh.ru/preprint.asp?id=2010-9>

Введение

Настоящая работа является очередной в серии [1 – 9], посвященной методу распознавания символов, основанному на регрессионном анализе [10 – 19].

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. Рассматриваются две модели такого преобразования, а именно, «затемнения» и «засветления».

При нарастании степени искажения исходной базы изучается динамика оценки распознавания (средней по всем верно распознанным изображениям), а также числа неправильно распознанных образов. Выявлены особенности корреляции в поведении указанных величин и основанных на растрах и полиномиальных векторах характеристик базы распознавания, осредненных по совокупности изображений символов.

Выполнен сравнительный анализ зависимостей, полученных для печатных и рукопечатных цифр.

1. Постановка задачи и методика решения

1.1. Метод распознавания. Разработанный алгоритм позволяет 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 -й 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 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

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

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, упорядо-

ченных 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 задана:

$$\mathbf{A}_j = \mathbf{A}_{j-1} - \alpha \mathbf{G} \mathbf{x}^{(j)} [\mathbf{A}_{j-1}^T \mathbf{x}^{(j)} - \mathbf{y}^{(j)}]^T, \quad \alpha = 1/J, \quad 1 \leq j \leq J \quad (4)$$

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

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

Для рукопечатных цифр использовался следующий (длинный) вектор \mathbf{x} :

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

Короткий вектор (для печатных символов) составлен из элементов длинного вектора, записанных в первой строке (5):

$$\mathbf{x} = (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_y\}, \{(\delta v_i)_y^2\}). \quad (6)$$

В (5) и (6) 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 \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 пикселей. Вне раstra считаем, что $v_i = 0$ (используется при вычислении конечных разностей на границе раstra).

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.

1.3. Аппарат для анализа метода распознавания символов. Подход к исследованию данного метода распознавания был разработан и впервые описан в работах [4 – 6].

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, $[0, 1] \rightarrow 1, (1, 2] \rightarrow 2, \dots, (254, 255] \rightarrow 255$.

После обучения на базе рукопечатных цифр в 174 778 элементов была получена матрица, с которой ошибочно распознавалось 0,5% элементов (881 изображение) этой же базы. Для печатных цифр на базе в 5496 элементов неверно распознавалось 4 изображения двух символов (0,05%) указанной базы.

Для «среднестатистического» раstra 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:

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

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:

$$\|\mathbf{v}-\mathbf{u}\| = \sum_{i=1}^L |v_i - u_i| \quad (8)$$

2. Модели «затемнения» («засветления»)

2.1. Описание моделей и некоторые замечания. На этапе распознавания все пиксели раstra 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 и Таблице 2 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 Таблиц 1 – 2, 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 Таблицах 1, 2 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.

Таблица 1 «Затемнение и засветление (рукопечатные цифры)»

	0	1	2	3	4	8	12	16
dv_tw	60,7	59,6	58,8	58,2	57,8	57,2	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	57,2	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	66,2
dv_sb	76,4	75,7	75,2	74,7	74,1	72,3	72,2	72,9
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	3140	3175	3231
dx_sw	3220	3205	3191	3176	3165	3115	3072	3036
dx_t	3224	3192	3170	3155	3145	3144	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	3966	3986	4052
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	871	884	892	907	1019	1279	1848
mis%t	0,50	0,50	0,51	0,52	0,53	0,69	1,05	1,69
mis%s	0,50	0,50	0,51	0,51	0,52	0,58	0,73	1,06
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	51,6	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	52,0	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	61,9	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	2998	3009	3061	3147	3253	3381	3452
dx_t	3304	3373	3720	3904	4115	4362	4651	4988
dx_s	3026	3014	3030	3094	3202	3343	3503	3573
dx_tb	3560	3575	3819	3989	4181	4412	4698	5035
dx_sb	3970	3875	3757	3713	3722	3764	3841	3915
xx_t	1252	1458	2568	2860	3154	3460	3790	4153
xx_s	850	1013	1352	1693	2038	2384	2705	2703
mis_t	4772	6972	33309	49051	69225	91758	110884	123073
mis_s	2523	3223	5077	8835	16860	30896	46397	45692
mis%t	2,70	3,99	19,06	28,06	39,61	52,50	63,44	70,42
mis%s	1,44	1,84	2,90	5,05	9,65	17,68	26,55	26,14
Prb_t	186,4	179,3	172,5	174,0	182,5	200,5	224,4	244,1
Prb_s	210,3	205,4	194,7	184,2	175,3	170,6	170,1	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
mis%t	75,39	80,96	81,03	81,03
mis%s	32,95	39,72	46,12	48,77
Prb_t	250,5	255,0	255,0	255,0
Prb_s	178,3	182,5	187,1	189,0

2.2. Интегральные характеристики модифицированной базы. Сравнительный анализ при затемнении / засветлении и различных типах написания символов. В Таблицах 1 и 2 для изучаемого диапазона $0 \leq n \leq 100$ содержатся данные, относящиеся к величинам, осредненным по совокупности изображений символов. Представлено расстояние между среднестатистическими растрами баз обучения и распознавания (vv_t при затемнении и vv_s при засветлении) – см. также Рис. 7а,б,в. Кроме того, приведено отклонение между среднестатистическими полиномиальными векторами баз обучения и распознавания (xx_t при затемнении и xx_s при засветлении) – см. также Рис. 4а,б. 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.

Таблица 2 «Затемнение и засветление (печатные цифры)»

	0	1	2	3	4	8	12
dv_tw	51,2	50,6	50,2	49,9	49,7	49,6	50,4
dv_sw	51,2	50,5	49,8	49,3	48,9	48,0	48,0
dv_t	51,2	50,6	50,2	49,9	49,7	49,7	50,4
dv_s	51,2	50,5	49,9	49,3	48,9	48,0	48,0
dv_tb	85,1	84,1	83,3	82,6	81,3	80,3	79,7
dv_sb	85,1	89,2	88,6	88,0	87,5	86,0	82,2
vv_t	0	1,7	3,5	5,2	6,9	13,7	20,4
vv_s	0	1,6	3,1	4,7	6,3	12,4	18,4
dx_tw	189,9	188,5	187,3	186,3	185,4	182,6	181,0
dx_sw	189,9	186,9	184,2	182,1	180,3	175,8	174,4
dx_t	190,0	188,6	187,4	186,4	185,5	182,7	181,1
dx_s	190,0	187,0	184,3	182,2	180,4	175,9	174,5
dx_tb	287,4	285,6	283,8	282,2	276,5	273,2	270,6
dx_sb	287,4	295,4	292,9	290,6	288,5	281,3	267,3
xx_t	0	3,7	7,4	11,1	14,9	30,0	45,2
xx_s	0	5,2	10,4	15,5	20,6	40,7	60,0
mis_t	4	4	4	4	5	5	5
mis_s	4	3	3	3	3	3	4
mis%t	0,07	0,07	0,07	0,07	0,09	0,09	0,09
mis%s	0,07	0,05	0,05	0,05	0,05	0,05	0,07
false_t	1,3	1,3	1,3	1,3	1,3,5	1,3,5	1,3,5
false_s	1,3	1,3	1,3	1,3	1,3	1,3	1,3
Prb_t	235,3	234,0	232,7	231,3	229,9	223,0	215,0
Prb_s	235,3	233,1	230,7	228,1	225,4	213,4	200,7

	16	20	24	32	40	48	56
dv_tw	51,8	53,7	56,2	62,7	70,3	78,7	87,5
dv_sw	48,8	50,3	52,3	57,7	64,1	71,2	78,7
dv_t	51,8	53,8	56,2	62,7	70,4	78,9	87,9
dv_s	48,8	50,3	52,3	57,7	64,1	71,2	78,8
dv_tb	81,7	81,4	85,4	89,0	91,4	91,6	94,6
dv_sb	81,7	77,4	75,7	77,9	81,4	85,2	90,2
vv_t	26,9	33,3	39,6	51,8	63,6	74,8	85,7
vv_s	24,3	30,1	35,7	46,5	56,8	66,6	76,1
dx_tw	180,5	181,0	182,7	190,2	202,0	217,9	237,6
dx_sw	175,8	179,1	183,8	196,5	212,0	228,7	246,3
dx_t	180,6	181,2	182,9	190,4	202,3	218,7	239,6
dx_s	175,8	179,2	183,8	196,6	212,0	228,7	246,4
dx_tb	274,2	270,7	277,0	281,8	274,2	265,6	267,2
dx_sb	264,3	254,2	250,5	251,4	257,2	263,5	272,3
xx_t	60,6	76,1	91,8	123,6	155,7	187,9	220,4
xx_s	78,8	96,9	114,4	147,4	178,0	206,0	231,7
mis_t	8	9	10	12	23	93	370
mis_s	4	5	6	9	9	11	13
mis%t	0,15	0,16	0,18	0,22	0,42	1,70	6,70
mis%s	0,07	0,09	0,11	0,16	0,16	0,20	0,24
false_t	1,3,5,8	1,3,5,8	1,3,5,8	1,3,5,8	0,3,4,5,8	1,1,6	1,1,6
false_s	1,3	1,3	1,3	1,3,6	1,3,6	1,3,6	1,3,6
Prb_t	206,5	197,6	188,5	170,1	151,2	132,8	116,5
Prb_s	187,9	175,4	162,9	138,8	116,0	94,5	74,8

	64	72	80	88	96	100
dv_tw	96,2	104,7	117,2	123,4	127,9	133,6
dv_sw	86,4	94,1	101,7	109,1	116,1	122,4
dv_t	97,2	106,5	115,6	124,3	132,6	136,6
dv_s	86,4	94,1	101,7	109,1	116,1	119,4
dv_tb	100,9	108,2	115,1	124,5	133,5	137,0
dv_sb	96,7	102,3	107,1	111,0	115,7	119,0
vv_t	96,1	106,0	115,4	124,2	132,6	136,6
vv_s	85,0	93,5	101,5	109,1	116,1	119,4
dx_tw	259,4	280,7	306,4	330,6	351,1	361,1
dx_sw	263,9	280,7	296,5	311,0	324,0	306,3
dx_t	264,1	290,9	319,6	349,0	378,8	393,7
dx_s	263,9	280,8	296,6	311,0	324,0	329,8
dx_tb	281,1	300,6	324,3	354,2	384,5	398,7
dx_sb	285,0	298,1	305,0	312,5	321,9	333,4
xx_t	253,0	285,2	317,1	348,2	378,7	393,7
xx_s	255,0	275,8	294,2	310,2	323,9	329,8
mis_t	1182	2833	4066	4280	4554	4772
mis_s	16	24	33	63	134	4772
mis%t	21,50	51,50	74,00	77,90	82,90	86,83
mis%s	0,29	0,44	0,60	1,15	2,44	86,83
false_t	1,1,6	1,1,6	1	1	1	1
false_s	1,3,4,6,8	1(2,5,7,9)	1(2,7,9)	1(2,9)	1,...,9	1
Prb_t	105,0	101,7	110,8	102,6	97,0	93,0
Prb_s	56,7	40,6	26,4	14,2	4,0	1,0

Таблица 3 «Засветление (печатные)»

	0	8	16	24	32	40	48	56	64	72	80	88	96
0	0	0	0	0	0	0	0	0	0	2	2	2	4
1	2	2	3	3	5	5	6	7	8	8	9	8	7
2	0	0	0	0	0	0	0	0	0	0	0	0	7
3	2	1	1	3	3	3	3	3	2	3	3	6	10
4	0	0	0	0	0	0	0	0	1	3	8	27	64
5	0	0	0	0	0	0	0	0	0	0	2	3	11
6	0	0	0	0	1	1	2	3	4	7	7	7	11
7	0	0	0	0	0	0	0	0	0	0	0	1	1
8	0	0	0	0	0	0	0	0	1	1	2	9	18
9	0	0	0	0	0	0	0	0	0	0	0	0	1
Σ	4	3	4	6	9	9	11	13	16	24	33	63	134

Таблица 4 «Затемнение (печатные)»

	0	8	16	24	32	40	48
0	0	0	0	0	0	2	10
1	2	2	2	1	1	0	0
2	0	0	0	0	0	0	4
3	2	2	3	5	7	11	45
4	0	0	0	0	0	1	8
5	0	1	2	3	3	8	18
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1
8	0	0	1	1	1	1	5
9	0	0	0	0	0	0	2
Σ	4	5	8	10	12	23	93

Сравним поведение указанных величин как в зависимости от характера процесса (затемнение, засветление), так и в зависимости от типа написания цифр (печатные, рукопечатные).

Очевидно, что для произвольного печатного или рукопечатного 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. 7а,б,в).

Во всем исследуемом диапазоне ($0 \leq n \leq 100$) и для печатных, и для рукопечатных символов функции vv_t и vv_s гладкие и почти линейные, с небольшим замедлением темпов их роста при увеличении n . Линии расходятся из точки (0, 0); расстояние между любыми двумя из них увеличивается с ростом n . Следует заметить, что осредненные по совокупности изображений символов отклонения от среднестатистического растра при любом $n > 0$ располагаются по возрастающей в следующем порядке (Рис. 7в):

- 1) засветление (рукопечатные символы),
- 2) засветление (печатные символы),

- 3) затемнение (печатные символы),
- 4) затемнение (рукопечатные символы).

Объяснение этого дано в п.2.14. Там же обосновываются следующие факты: для каждого $n > 0$ величина vv_t превышает vv_s для рукопечатных цифр более чем в два раза (Рис. 7а), а для печатных – при уже отмеченном монотонном увеличении расстояния между ними с ростом n – не более чем на 15% (Рис. 7б).

Расстояние между среднестатистическими полиномиальными векторами рукопечатных символов увеличивается с ростом n при затемнении – монотонно, а при засветлении имеется небольшой диапазон $64 \leq n \leq 72$, где это расстояние незначительно уменьшается (среднее по совокупности символов – Рис. 4а).

Для печатных символов расстояние между среднестатистическими полиномиальными векторами (Рис. 4б) монотонно нарастает, причем сначала после выхода из точки $(0, 0)$ указанная величина при затемнении меньше, чем при засветлении, а затем при $n \approx 67$ линии пересекаются и ситуация меняется на противоположную.

В соответствии со сказанным ранее о соотношении vv_t и vv_s , для каждого $n > 0$ величина xx_t более чем в полтора раза превосходит xx_s для рукопечатных цифр (Рис. 4а). Для печатных символов эти величины отличаются незначительно на большей части отрезка $0 \leq n \leq 100$, причем их разница меняется в знаке (как уже было замечено в предыдущем абзаце) и приближается к 20% лишь около $n = 100$.

2.3. Качество распознавания. Сравнительный анализ при затемнении / засветлении и различных типах написания символов.

В Таблицах 1 и 2 приведены число нераспознанных изображений символов (mis_t при затемнении и mis_s при засветлении) – Рис. 3а,б, их доля относительно общего количества изображений ($mis\%t$ и $mis\%s$) – Рис. 3в, а также средняя оценка распознавания (Prb_t при затемнении и Prb_s при засветлении) – Рис. 2а,б,в.

Количество неправильно распознанных образов рукопечатных символов в процессе затемнения увеличивается монотонно, а для засветления имеется немонотонность при $n = 1$ (незначительная), а также при $64 \leq n \leq 72$. Согласно описанным ранее закономерностям в соотношениях (vv_t, vv_s) и (xx_t, xx_s), имеем: при $n > 1$ значение ($mis_t - 881$) в полтора-два раза больше, чем ($mis_s - 881$), здесь 881 – число нераспознанных изображений при $n = 0$. Кроме того, следует отметить для проводимого далее сравнения, что разность ($mis\%t -$

$\text{mis}\%$ s) начинает превышать 10% от общего количества изображений символов при $n \approx 29$ (Рис. 3а,в).

Число неправильно распознанных растров печатных символов при затемнении увеличивается монотонно, а при засветлении имеется описываемая ниже в абзаце немонотонность на отрезке $0 \leq n \leq 8$ (Рис. 3б,в). Сравнивая с результатом предыдущего абзаца для рукопечатных цифр, имеем: mis_t больше, чем mis_s при $n > 0$, но разность ($\text{mis}\%t - \text{mis}\%s$) начинает превышать 10% лишь при $n \approx 59$. Сопоставлять величины $(\text{mis}_t - 4)$ и $(\text{mis}_s - 4)$, где 4 – число нераспознанных символов при $n = 0$, аналогично тому, как это было сделано в рукопечатном случае, нецелесообразно ввиду незначительного количества ошибочно распознанных изображений при $n = 0$ (при котором минимально возможное нарушение монотонности для засветления на отрезке $0 \leq n \leq 8$ является существенным), а также из-за «особенного» поведения mis_s .

По поводу последнего надо заметить, что для указанного засветления печатных цифр при $n = 96$ получено $\text{mis}_s = 134$ ($\text{mis}\%s = 2,44\%$). Можно сравнить со следующими результатами: для печатных символов $\text{mis}_t = 4554$ ($\text{mis}\%t = 82,9\%$), для рукопечатных $\text{mis}_t = 141617$ ($\text{mis}\%t = 81,03\%$), $\text{mis}_s = 80611$ ($\text{mis}\%s = 46,12\%$).

Эти данные показывают, что засветление изображений печатных символов приводит к наименьшему нарастанию доли неправильно распознанных образов по сравнению с засветлением рукопечатных, а также затемнением и печатных, и рукопечатных символов.

В случае печатного написания при затемнении происходит «запаздывание» в росте доли нераспознанных растров по сравнению с рукопечатным засветлением, а тем более с затемнением. Так для печатных символов $\text{mis}\%t$ начинает превышать 10% от общего количества изображений символов лишь при $n \approx 58$, тогда как для рукопечатных это происходит с $\text{mis}\%t$ уже при $n \approx 28$, а с $\text{mis}\%s$ при $n \approx 48$. При затемнении печатных цифр доля ошибок лишь для $n = 72$ становится больше, чем при засветлении рукопечатных, а для $n = 96$ – чем при затемнении последних (Рис. 3в).

Средняя оценка распознавания для рукопечатных символов и при затемнении, и при засветлении сначала уменьшается, а затем увеличивается (Рис. 2а,в). Для Prb_t средние темпы и падения, и роста в полтора-два раза выше, чем для Prb_s (см. п.2.7). 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.

Для печатных символов (Рис. 2б,в) Prb_s стремительно монотонно падает, а Prb_t при общей тенденции к гораздо более медленному монотонному уменьшению имеет небольшой участок немонотонности, где достигается локальный максимум при $n = 80$. Следует заметить, что именно при этом значении имеется нарушение гладкости функции mis_t .

Итак, средняя оценка распознавания при $n = 0$ для печатных цифр несколько больше, чем для рукопечатных. Однако засветление изображений печатных символов происходит со стремительным падением оценки распознавания, которая всюду при $n > 0$ меньше соответствующей оценки для затемнения печатных, а также при $n = 3$ становится ниже оценки засветления и при $n = 12$ – затемнения рукопечатных символов. Оценка при затемнении печатных цифр падает (с учетом указанного ранее отклонения от монотонности при $n = 80$) медленнее, чем при засветлении; она становится меньше оценки при засветлении рукопечатных лишь для $n = 8$, а при их затемнении – для $n = 32$.

Результаты распознавание печатной и рукопечатной базы обучения отличается существенно. Как уже было показано, на базе рукопечатных цифр в 174 778 элементов 881 изображение различных символов распознано неверно. Для печатных цифр на базе в 5496 элементов ошибочно распознано 4 изображения двух символов. Кажется вполне естественным интерес, изображения каких символов и в каком количестве будут попадать в разряд нераспознанных при затемнении и засветлении печатной базы? В Таблице 2 эти символы перечислены в графах $false_t$ и $false_s$ соответственно. Знак \neg используется перед теми символами, все изображения которых распознаются полностью правильно (в отличие от других символов).

В Таблицах 3 и 4 изображениям разных символов, а также их совокупности (цифры 0, 1, ..., 9 и знак Σ в левом столбце) приводится в соответствие количество ошибок для приведенных в верхней строке степеней n . Таблица 4 ограничивается затемнением со значением $n = 48$, для которого $mis_t = 93$ находится между результатами, полученными для засветления: $mis_s = 63$ при $n = 88$ и $mis_s = 134$ при $n = 96$.

2.4. Корреляция качества распознавания и векторной интегральной характеристики модифицированной базы при засветлении для рукопечатных символов. По Рис.4а. 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 (Рис.3а). 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 (Рис.2а).

2.5. Корреляция качества распознавания и векторной интегральной характеристики модифицированной базы при затемнении для рукопечатных символов. На Рис.4а. 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 (Рис.3а). 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 (Рис.2а).

2.6. О качестве распознавания и интегральных характеристиках модифицированной базы при засветлении и затемнении для печатных символов. В п.п. 2.4 и 2.5 для рукопечатных символов более углубленно излагается материал из п.п.2.2 и 2.3. Здесь потребуются обосновать неприемлемость такого подхода для печатных цифр. Также будут обобщены результаты, изложенные в п.п. 2.2 – 2.3 для печатных цифр.

Структура вектора для печатных цифр более простая, чем для рукопечатных. Это «промежуточная стадия» между растровым представлением и «рукопечатным» вектором.

График оценки распознавания для печатных цифр имеет меньше особенностей, чем для рукопечатных (*печатные*: монотонное убывание для засветления и тенденция к этому для затемнения; *рукопечатные*: ярко выраженная немонотонность, а именно вогнутость и для затемнения, и для засветления).

Поведение количества 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).

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. Как и последние, они не отражают проявлений немонотонности оценки и нарушения гладкости в графике для числа нераспознанных символов (при затемнении печатного образа).

Как уже было отмечено в п.2.3, в случае печатных символов Prb_t при общей тенденции к монотонному уменьшению имеет небольшой участок немонотонности, где достигается локальный максимум при $n = 80$. При этом же значении нарушена гладкость mis_t .

2.7. О сильном затемнении и засветлении. Конечно же, при больших значениях n нельзя говорить о сколь-нибудь приемлемом распознавании, и результаты носят чисто теоретический характер.

В случае рукопечатного написания следует отметить, что ситуация в этой области асимметрична для затемнения и засветления. А именно, при полном затемнении с $n = 100$ («Черный квадрат») все изображения 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, в качестве искомого решения выбираются различные цифры с существенно более низкими оценками, чем при полном затемнении. Иллюстрацией к сказанному является Таблица 5, 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 Σ) приводится в соответствие количество случаев, в которых данное изображение этого символа или их совокупности распознается как один из символов верхней строки 0, 1, ..., 9.

Для печатных цифр при полном затемнении все образы, как и для рукопечатных, распознаются как «1», но уже с меньшей оценкой $Prb_t = 93$. Этим и объясняется тенденция к монотонному убыванию данной функции при затемнении.

Оценка, с которой распознается «Черный квадрат», является свойством, присущим используемой при этом матрицы. Поэтому в данной работе для матриц, соответствующих различным типам написания (рукопечатной или печатной) получены разные оценки. Они говорят не о качестве матрицы, а о ее «восприятии» данного изображения в качестве одного из возможных символов.

Количество (или доля) неправильно распознанных изображений символов при полном затемнении равна общему количеству (или 100%) изображений за вычетом числа (доли) правильно распознанных образов. Поскольку для обоих типов написания «Черный квадрат» распознается как «1», то в Таблицах 1, 2 при $n = 100$ имеем: mis_t равно общему объему базы за вычетом количества изображений «1» ($mis\%t$ равно 100% минус доля «1»).

Для печатных цифр при полном засветлении все образы распознаются как «1», но уже с минимальной оценкой $Prb_s = 1$. Этим и объясняется монотонное убывание данной функции при засветлении.

Тем не менее, результаты, полученные при $n = 88$ (Таблица 6), демонстрируют, насколько высока устойчивость правильного распознавания к засветлению для печатных символов. Приведены результаты при $n = 88$, поскольку с использованием дискретного ряда оценок 1, 2, ..., 255, расчеты удалось выполнить только до $n = 95$. Далее при $n = 96$ и $n = 100$ осуществлялись дополнительные манипуляции по сравнению результатов с низкими оценками.

2.8. Дифференциальные характеристики модифицированной базы. В Таблицах 1, 2 приведены среднее по совокупности изображений символов отклонение растров базы распознавания от среднестатистического растра базы обучения – см. Рис.8а,б – 9а,б, а также, аналогично, среднее по совокупности образов символов отклонение полиномиальных векторов базы распознавания

от среднестатистического вектора базы обучения – см. Рис.5а,б – ба,б. Они определяются следующим образом. Для каждого образа некоторого символа определяется отклонение от среднестатистического растра или вектора этого символа, а затем находится среднее значение по всем изображениям базы.

Эти величины в Таблицах 1, 2 даны без учета правильности распознавания – см. также Рис.5а,б, 8а,б, для правильного распознавания – аналогично изображенному на Рис.5а,б, Рис.8а,б и, наконец, для неправильно распознанных символов – см. также Рис.6а,б, Рис.9а,б. 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 при любом типе написания с ростом n должны увеличиваться (возможно, с некоторыми погрешностями) аналогично vv_t , vv_s , xx_t , xx_s . В дальнейшем изложении приводятся результаты расчетов, а в п.2.13 будут сделаны заключительные замечания по этому вопросу.

2.9. Динамика дифференциальных характеристик модифицированной базы при засветлении для печатных и рукопечатных цифр. Оказалось, что при различных типах написания поведение и векторных, и растровых дифференциальных характеристик отличается значительно меньше, чем векторных интегральных (xx_t , xx_s). Здесь прослеживается аналогия с наблюдающимся и для печатных, и для рукопечатных символов сходством растровых интегральных характеристик (vv_t , vv_s).

При усилении засветления dv_s и dv_{sw} уменьшаются и достигают минимума для рукопечатных цифр при $n = 32$ (для печатных при $8 \leq n \leq 12$), 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.8а,б).

С ростом n величина dv_{sb} сначала уменьшается (за исключением участка незначительного нарастания для рукопечатных символов при $12 \leq n \leq 16$, а для печатных при $n = 1$) и достигает минимума для рукопечатных при $n = 40$ (для печатных при $n = 24$), а затем монотонно увеличивается (Рис.9а,б). Для рукопе-

чатных цифр при $0 \leq n \leq 100$ величина dv_sb существенно больше, чем dv_s и dv_sw . Следует также заметить, что из неравенства $dv_sb > dv_sw$ следует неравенство $dv_s > dv_sw$. В случае печатного написания ситуация отличается: при $0 \leq n \leq 80$ величина dv_sb существенно больше, чем dv_s и dv_sw , а при $n = 88$ первая лишь незначительно больше второй, а при $n = 96$, $n = 100$ значение dv_sb незначительно меньше, чем dv_sw .

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.

При анализе поведения аналогичных величин в терминах полиномиальных векторов оказывается, что при усилении засветления dx_s и dx_sw уменьшаются и достигают минимума для рукопечатных цифр при $n = 24$ (для печатных – при $n = 12$), 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.5а,б).

Величина dx_sb уменьшается (за исключением участка незначительного нарастания для рукопечатных символов при $8 \leq n \leq 16$, а для печатных при $n = 1$) и достигает минимума для рукопечатных при $n = 40$ (для печатных при $n = 24$), после чего монотонно увеличивается. Для рукопечатных цифр dx_sb существенно больше, чем dx_s и dx_sw на всем исследуемом диапазоне изменения n ($0 \leq n \leq 100$) (Рис.6а,б). Кроме того, из неравенства $dx_sb > dx_sw$ следует неравенство $dx_s > dx_sw$. В случае печатного написания ситуация отличается: при $0 \leq n \leq 72$, $n = 100$ величина dx_sb существенно больше, чем dx_s и dx_sw , а при $n = 80$, $n = 88$ первая лишь незначительно больше второй. При $n = 96$ значение dx_sb незначительно меньше, чем dx_sw .

Таблица 5 «Рукопечатные: засветление $n = 100$ »

	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
Σ	16670	32444	30816	70827	1212	18219	721	3858	0	11

Итак, при усилении засветления поведение растровых величин dv_{sb} , dv_{sw} , и dv_s в значительной степени сходно с их векторными аналогами dx_{sb} , dx_{sw} , и dx_s (особенно мало различие для печатных символов). Для обоих типов написания картина изменения этих величин качественно схожа.

Таблица 6 «Печатные: засветление $n = 88$ »

	0	1	2	3	4	5	6	7	8	9
0	53	0	0	0	0	0	2	0	0	0
1	0	716	2	0	0	0	0	6	0	0
2	0	0	967	0	0	0	0	0	0	0
3	0	2	0	574	0	1	0	0	3	0
4	0	26	0	0	590	0	1	0	0	0
5	0	0	0	3	0	804	0	0	0	0
6	0	0	0	0	0	2	275	0	5	0
7	0	0	0	0	0	1	0	452	0	0
8	0	0	0	9	0	0	0	0	445	0
9	0	0	0	0	0	0	0	0	0	557
Σ	53	744	969	586	590	808	278	458	453	557

2.10. Корреляция качества распознавания, векторных интегральной и дифференциальных характеристик модифицированной базы при засветлении рукопечатных символов. Следует отметить такую закономерность. На отрезке $64 \leq n \leq 72$ радикально «перестраиваются» графики отклонения между среднестатистическими полиномиальными векторами баз обучения и распознавания xx_s , количества неправильно распознанных символов mis_s , а также средней оценки распознавания P_{rb_s} (п.п.2.2 – 2.3). 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 незначительно большими значениями n . Величина dx_{sb} приближается к исходному значению, но не достигает его даже при $n = 100$.

По Рис.5а видно, что график dx_s (и аналогично dx_{sw}) «перестраивается» при $64 \leq n \leq 72$, но в меньшей степени, чем xx_s . На самом этом отрезке dx_s нарастает, но очень незначительно, с заметно меньшей скоростью, чем левее и правее указанного участка.

На графиках dx_{sb} (Рис.6а), dv_s (и аналогично dv_{sw}) (Рис.8а), dv_{sb} (Рис.9а) подобная перестройка не наблюдается.

2.11. Динамика дифференциальных характеристик модифицированной базы при затемнении для печатных и рукопечатных цифр. В процессе усиления затемнения расстояния dv_t и dv_{tw} сначала уменьшаются и достигают минимума при $n = 8$ (для рукопечатных цифр это существенно меньшее значе-

нии n по сравнению с засветлением, а для печатных практически то же); затем они 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.8а,б).

С ростом n величина dv_tb сначала уменьшается и достигает минимума при существенно меньшем значении n по сравнению с засветлением (для рукопечатных цифр $n = 16$, для печатных $n = 12$), а затем увеличивается (Рис.9а,б), причем для печатных символов наблюдается незначительный шум. Следует заметить, что для рукопечатных символов при $0 \leq n \leq 16$ (для печатных при $0 \leq n \leq 56$) 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 $n = 20$ для рукопечатных цифр ($n = 64$ для печатных) 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 $n \geq 80$ для рукопечатных и при $n \neq 88$ для печатных), если $dv_tb = dv_tw$, то $dv_t = dv_tw$ (при $n = 48$ для рукопечатных) а если $dv_tb > dv_tw$, то $dv_t > dv_tw$ (при $n < 80$, $n \neq 48$ для рукопечатных и $n = 88$ для печатных).

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 (для рукопечатных символов при $n = 8$, для печатных при $n = 16$ – значения, соответственно меньшее и большее чем при засветлении), 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.5а,б).

Для рукопечатных цифр с ростом n величина dx_tb сначала уменьшается и достигает минимума при существенно меньшем значении n по сравнению с засветлением $n = 20$, а затем увеличивается (Рис.6а). 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, $dx_t > dx_tw$.

Для печатных цифр с ростом n величина dx_tb сначала уменьшается и достигает минимума при $n = 12$, после этого колеблется при $12 \leq n \leq 48$ а затем увеличивается (Рис.6б). Следует заметить, что при $0 \leq n \leq 100$ величина dx_tb существенно больше, чем dx_t и 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 рас-

тров 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).

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 .

Итак, при усилении затемнения поведение растровых величин dv_tb , dv_tw , и dv_t в значительной степени сходно с их векторными аналогами dx_tb , dx_tw , и dx_t . Для обоих типов написания картина изменения этих величин качественно схожа.

2.12. Корреляция качества распознавания, векторных интегральной и дифференциальных характеристик модифицированной базы при затемнении рукопечатных символов. 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.2.4). 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 .

По Рис.5а видно, что график dx_t (и аналогично dx_tw) «перестраивается» при $24 \leq n \leq 32$, но в меньшей степени, чем xx_t . На этом отрезке dx_s нарастает с заметно большей скоростью, чем левее и правее указанного участка.

На графиках dx_tb (Рис.6а), dv_t (и аналогично dv_tw) (Рис.8а), dv_tb (Рис.9а) подобная перестройка не наблюдается.

2.13. Сравнительный анализ интегральных и дифференциальных характеристик при затемнении и засветлении для печатных и рукопечатных символов. Итак, было показано, что дифференциальные характеристики ведут себя отлично от аналогичных интегральных: ни для какого типа написания нет монотонного возрастания, хотя бы и с шумами, по мере роста степени затемнения / засветления. На начальном и среднем участках исследуемого отрезка $0 \leq n \leq 100$ имеется ярко выраженная немонотонность (минимум), а также пульсации. Однако, с учетом обоих типов написания, для dv_t , dv_s (аналогично dv_tw , dv_sw) при $n > 24$ (Рис.8в), а для dv_tb , dv_sb при $n > 43$ (Рис.9в) наблюдается картина, аналогичная изображенной на Рис.7в для vv_t и vv_s . А

именно, указанные функции монотонно возрастают, располагаясь по возрастающей в приведенном в п.2.2 порядке.

Проведем аналитическое исследование, которое имеет самостоятельное значение, а также является иллюстрацией к вышесказанному. В целях простоты выкладки выполнены для изображений одного символа. Пусть при некотором изменении базы (неважно, затемнении, засветлении или еще каком-то) v_i^j – значение в i -м пикселе j -го изображения данного символа (при общем их количестве J) трансформируется в w_i^j . Соответствующую дифференциальную характеристику обозначим DV (по аналогии с dv_t и dv_s), а интегральную – VV (сходно с vv_t и vv_s). Итак, для данного символа

$$DV = \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - v_i^{cp}| \right) = \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - w_i^{cp} + w_i^{cp} - v_i^{cp}| \right) \quad (9)$$

$$VV = \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - v_i^{cp}| \right) \quad (10)$$

Очевидно, что для (9) выполняются два неравенства:

$$DV \leq \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^{cp} - v_i^{cp}| \right) + \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - w_i^{cp}| \right) \quad (11)$$

$$DV \geq \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^{cp} - v_i^{cp}| \right) - \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - w_i^{cp}| \right) \quad (12)$$

С учетом (10) их можно записать так:

$$DV \leq VV + \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - w_i^{cp}| \right) \quad (13)$$

$$DV \geq VV - \frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - w_i^{cp}| \right) \quad (14)$$

Теперь следует учесть тот факт, что начиная с некоторой степени затемнения (засветления) для все большего количества пикселей (i) и аналогично увеличивающегося числа изображений данного символа (j) будет выполняться равенство: $w_i^j = 0$ при засветлении (аналогично $w_i^j = 1$ при затемнении). Соответственно, w_i^{cp} будет стремиться к 0 при засветлении и к 1 при затемнении). Следовательно, величина $\frac{1}{J} \sum_{j=1}^J \left(\sum_{i=1}^N |w_i^j - w_i^{cp}| \right)$, стоящая после знака «+» в неравенстве (13) и после знака «-» в неравенстве (14) будет стремиться к нулю (и достигает его при $n = 100$). Итак, при приближении степени затемнения (засветления) n к максимальному значению 100 имеем: $DV(n)$ приближается к $VV(n)$, и $DV(100) = VV(100)$ – см. Рис.11а,б, а также п.2.14.

Понятно, что терминология выкладок из предыдущего абзаца несколько условна, поскольку речь идет о конечном наборе n . Конечно же, этот момент можно обойти, если перейти от дискретных степеней затемнения (засветления) к непрерывным, но это усложнение излишне.

Для векторных величин для каждого из типов написания в отдельности также наблюдается сходство в поведении дифференциальных и интегральных величин. Так, для рукопечатных символов dx_t , dx_s (аналогично dx_{tw} , dx_{sw}) при $n > 24$ (Рис.5а), а dx_{tb} , dx_{sb} при $n > 43$ (Рис.6а) ведут себя схоже с xx_t и xx_s на Рис.4а. Для печатных символов dx_t , dx_s (аналогично dx_{tw} , dx_{sw}) при $n > 24$ (Рис.5б), а dx_{tb} , dx_{sb} при $n > 50$ (Рис.6б) ведут себя схоже с xx_t и xx_s на Рис.4б.

2.14. Расстояние между растрами «Белого» и «Черного» квадратов. Рукопечатные и печатные символы. По Таблице 1 видно, что при $n = 100$ для рукопечатных цифр 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, $vv_t + vv_s = dv_t + dv_s = 256$ (оба равенства легко получить аналитически). Эта сумма суть расстояние 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, что для каждого $n > 0$ величина vv_t превышает vv_s более чем в два раза.

Для печатных цифр (Таблица 2) при $n = 100$ выполняются равенства: $vv_t = dv_t = 136,6$ (отклонение растра «Черного квадрата» от среднестатистического растра обучающей базы, причем первое равенство легко получить аналитически), а также $vv_s = dv_s = 119,4$ (отклонение растра «Белого квадрата» от среднестатистического растра обучающей базы). Следовательно, $vv_t + vv_s = dv_t + dv_s = 256$ (оба равенства легко получить аналитически). Указанная сумма соответствует расстоянию между растрами «Белого» и «Черного» квадратов.

Итак, отклонение среднестатистического растра базы обучения от растра «Черного» квадрата приблизительно на 10% больше, чем от растра «Белого» квадрата (тогда как для рукопечатных символов – в два с половиной раза). Этим объясняется, что для каждого $n > 0$ величина vv_t превышает vv_s не более чем на 10% (тогда как для рукопечатных цифр – более чем в два раза).

Для обоих типов написания скорость роста этих функций замедляется с ростом n , поскольку при увеличивается количество пикселей, в которых достигается значение «1» при затемнении и «0» при засветлении.

2.15. Выводы. При увеличении степени затемнения / засветления расстояние между среднестатистическими растрами баз обучения и распознавания увеличивается почти линейно, с небольшим падением темпов роста. Это утверждение верно и для печатных, и для рукопечатных цифр. При любой степени затемнения и такого же уровня засветления эти величины располагаются по возрастающей в следующем порядке: сначала засветление (рукопечатных, затем печатных), далее затемнение (печатных, и наконец рукопечатных).

Расстояние между среднестатистическими векторами в случае рукопечатного написания 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. Графики для печатных символов являются гладкими, монотонными, – сходны с печатными и рукопечатными растровыми аналогами.

Количество неправильно распознанных рукопечатных изображений изменяется аналогично расстоянию между среднестатистическими векторами (будучи при этом нелинейной функцией); исключением является небольшое уменьшение числа нераспознанных символов на начальном этапе засветления. Поведение количества ошибок для печатных символов характеризует почти монотонный рост, гладкость для засветления и нарушение ее лишь в одной точке для затемнения. Для любой степени затемнения каждая из этих величин больше, чем для такого же уровня засветления.

Средняя оценка распознавания и при затемнении, и при засветлении рукопечатных сначала убывает, а затем возрастает (при затемнении скорости больше). Минимумы приблизительно равны по величине.

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.

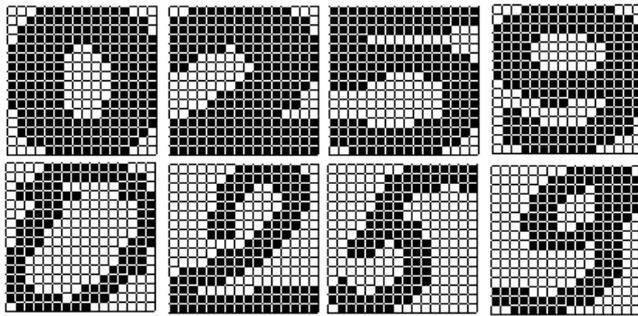
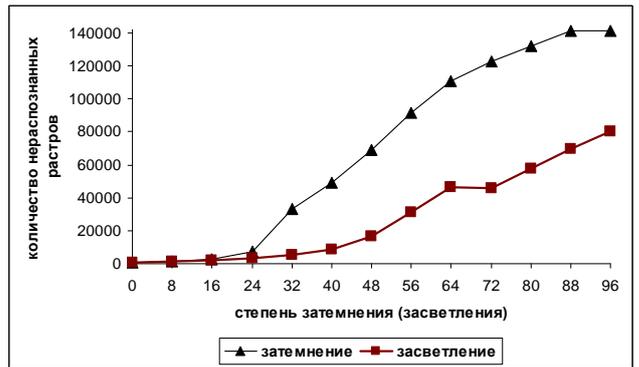
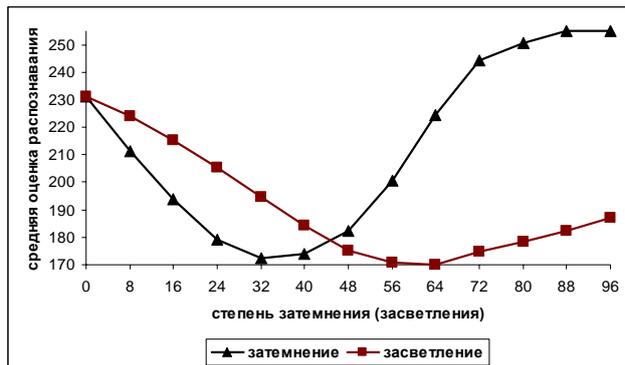


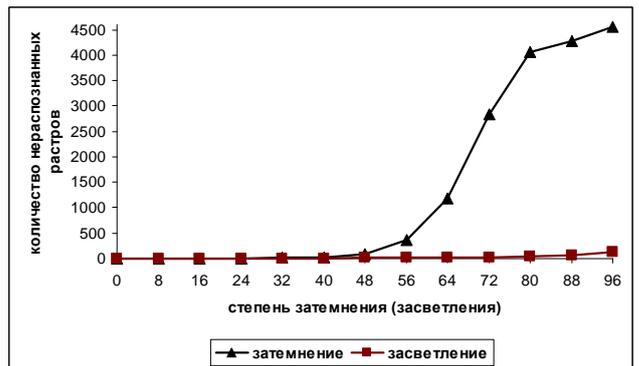
Рис. 1. Образы 16x16 печатных и рукопечатных цифр



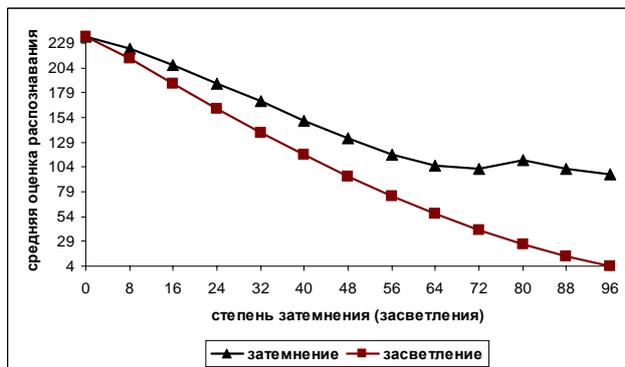
а) рукопечатные цифры (mis_t – mis_s)



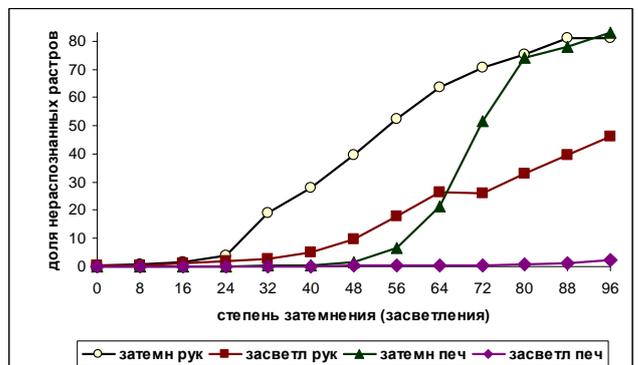
а) рукопечатные цифры



б) печатные цифры (mis_t – mis_s)

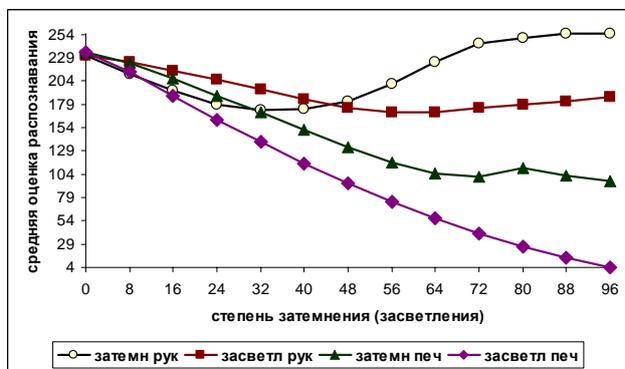


а) печатные цифры



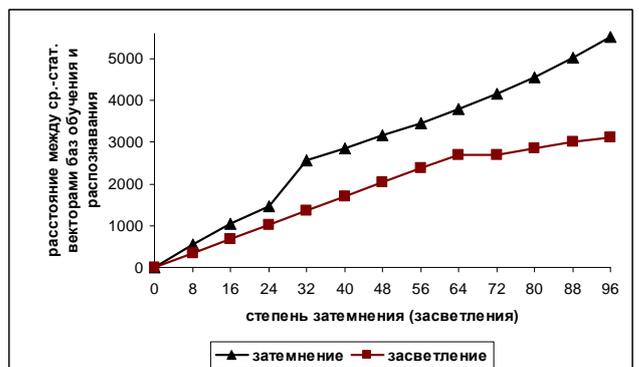
в) рукопечатные и печатные цифры (mis_t – mis_s)

Рис. 3. «(mis_t – mis_s), (mis_t – mis_s)»

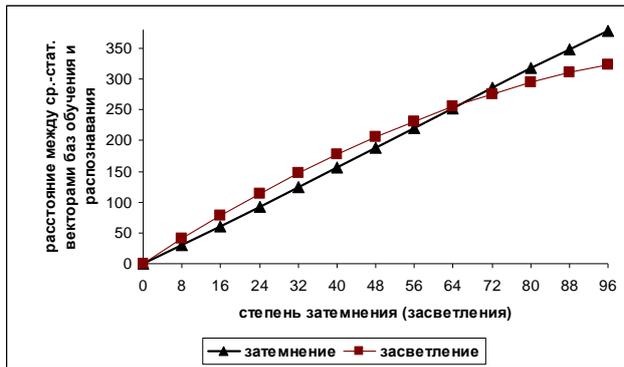


в) рукопечатные и печатные цифры

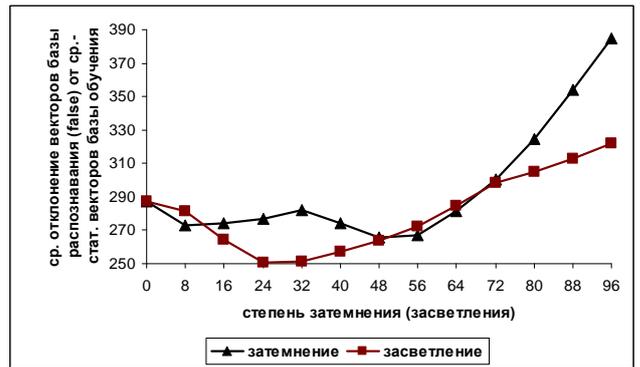
Рис. 2. «Prb_t – Prb_s»



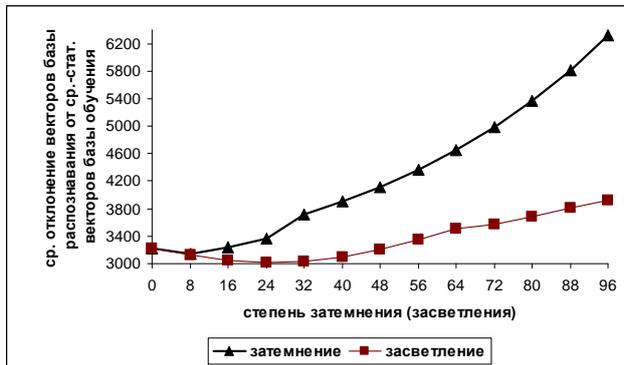
а) рукопечатные цифры



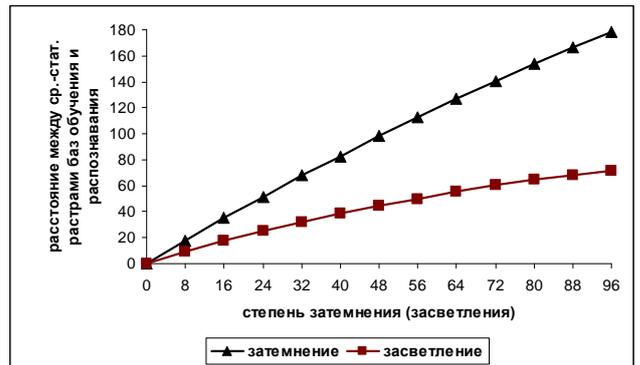
б) печатные цифры
Рис. 4. «xx_t, xx_s»



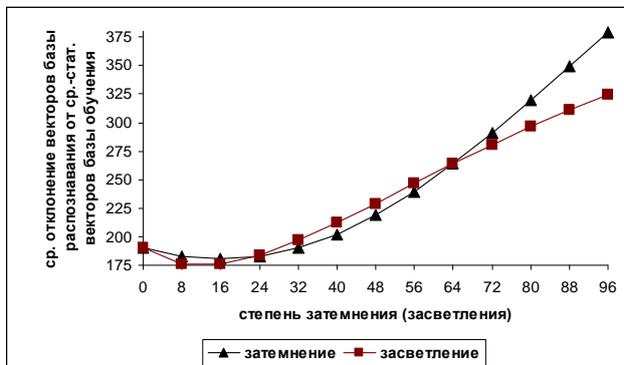
б) печатные цифры
Рис. 6. «dx_tb – dx_sb»



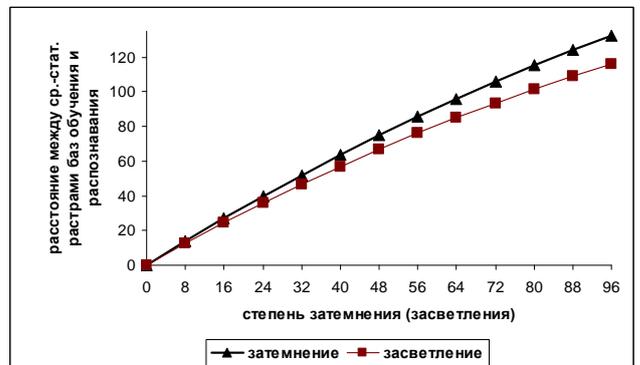
а) рукопечатные цифры



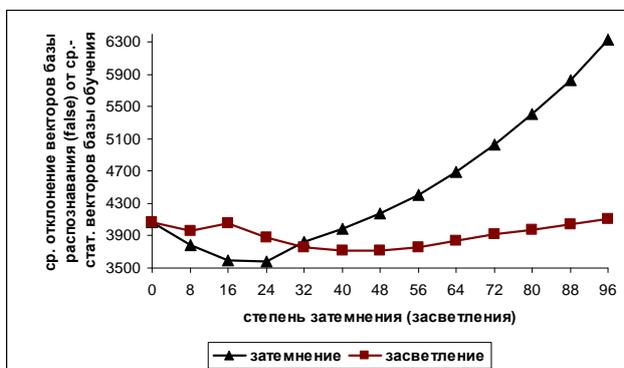
а) рукопечатные цифры



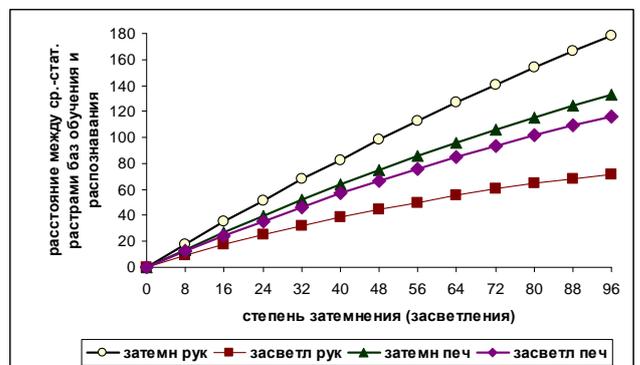
б) печатные цифры
Рис. 5. «dx_t – dx_s»



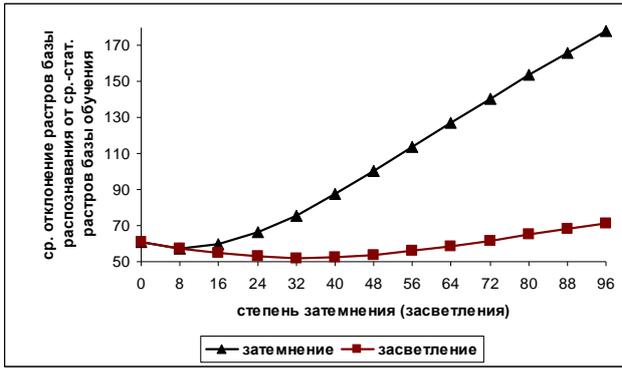
б) печатные цифры



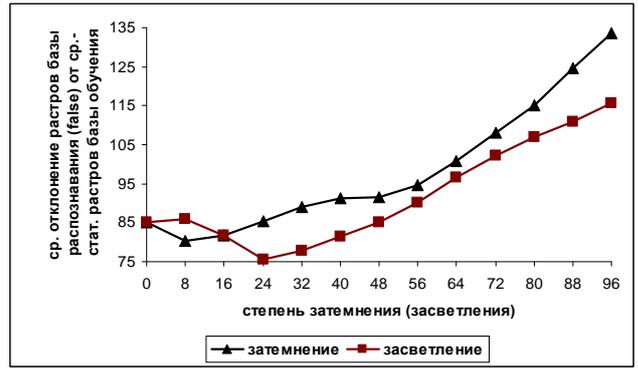
а) рукопечатные цифры



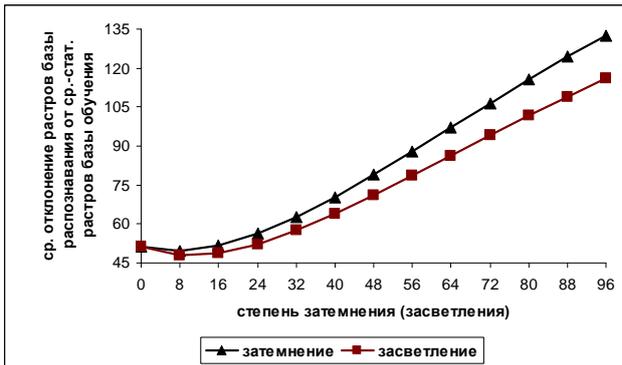
в) рукопечатные и печатные цифры
Рис. 7. «vv_t – vv_s»



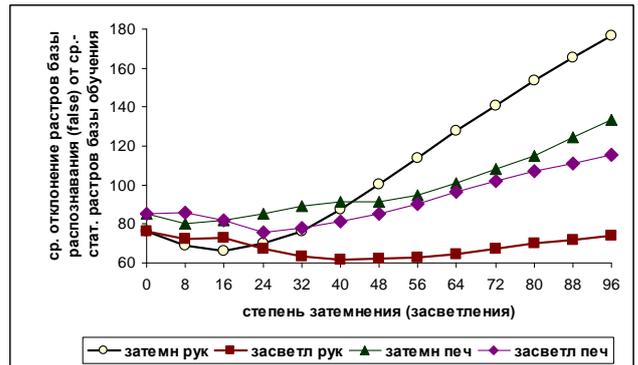
а) рукопечатные цифры



б) печатные цифры

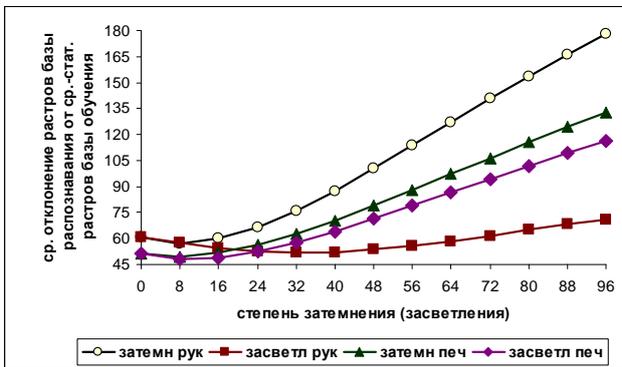


б) печатные цифры



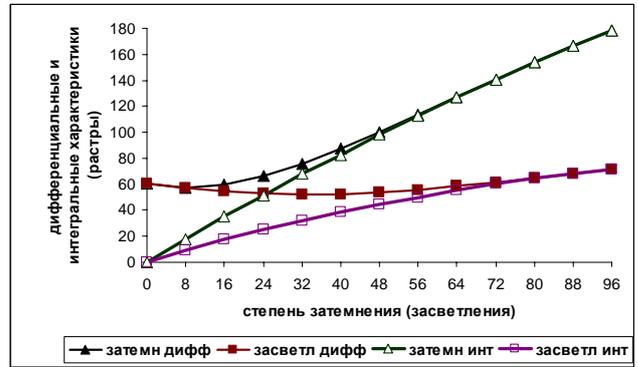
в) рукопечатные и печатные цифры

Рис. 9 «dv_tb – dv_sb»

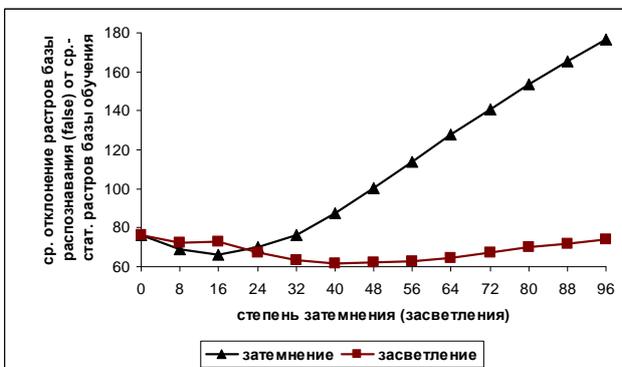


в) рукопечатные и печатные цифры

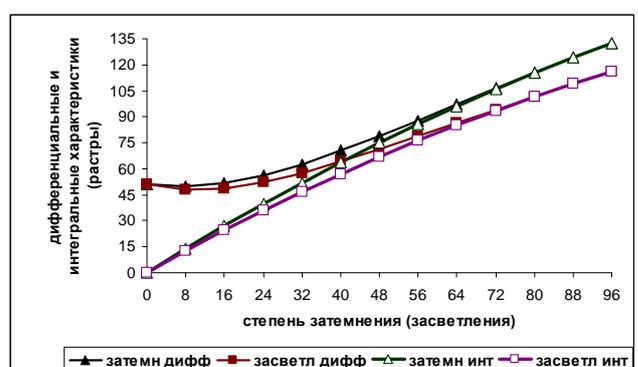
Рис 8 «dv_t – dv_s»



а) рукопечатные цифры



а) рукопечатные цифры



б) печатные цифры

Рис.10 «Dv – Vv»