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

Препринты ИПМ • Препринт № 35 за 2015 г.

Гавриков Б.М., Гавриков М.Б., Пестрякова Н.В.

Исследование характеристик метода распознавания на обучающем множестве

Рекомендуемая форма библиографической ссылки: Гавриков Б.М., Гавриков М.Б., Пестрякова Н.В. Исследование характеристик метода распознавания на обучающем множестве // Препринты ИПМ им. М.В.Келдыша. 2015. № 35. 24 с. URL: http://library.keldysh.ru/preprint.asp?id=2015-35

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

Б.М. Гавриков, М.Б. Гавриков, Н.В. Пестрякова

Исследование характеристик метода распознавания на обучающем множестве

Москва — 2015

Б.М. Гавриков, М.Б. Гавриков, Н.В. Пестрякова

Исследование характеристик метода распознавания на обучающем множестве

Описываются результаты статистического исследования функции оценки, вектора первичных признаков, а также структуры матрицы весовых коэффициентов, полученной на этапе обучения с применением вероятностного высокоточного метода распознавания рукопечатных символов.

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

Boris Mikhailovich Gavrikov, Mikhail Borisovich Gavrikov, Nadejda Vladimirovna Pestryakova

A research of characteristics of the recognition method on the training set

The results of statistical research of the evaluation function, the vector of primary signs and also the matrix of weight coefficients obtained during the training phase with the use of probabilistic high-precision method of recognition of handprinted characters are described.

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

Работа выполнена при поддержке Российского фонда фундаментальных исследований, проекты 13-07-12176-офи_м и 13-07-00262-а.

Оглавление

1. Введение

К настоящему времени разработано большое количество методов распознавания объектов самой разнообразной этимологии. Зачастую изучение свойств этих методов сводилось лишь к сопоставлению с другими подходами по таким характеристикам, как точность и быстродействие [1]. В последние годы были предложены также методики анализа монотонности (надежности) методов и устойчивости к искажениям [3, 4]. Недостаток всех этих исследований заключается в том, что вне области интересов оказалась внутренняя структура самого «устройства» для распознавания. Оно фактически используется в качестве некоего «черного ящика», работающего следующим образом: на его вход поступает кодированное изображение объекта, а на выходе получается номер класса, к которому этот объект относится, а также, возможно, оценка распознавания.

О способе получения этой оценки также имеется самое общее представление. Используемый набор формул, при помощи которых она вычисляется, по сути дает качественное, а не количественное ее описание, полностью лишенное наглядности. Сложность решения обозначенной проблемы обусловлена множественностью как природы распознаваемых образов, так используемых методов. Для описываемого в настоящей статье метода распознавания таким устройством с неизученной структурой является двумерная матрица весовых коэффициентов. При проведении исследований надлежало также учесть особенности объектов, используемых в данной работе как для обучения, так и для распознавания (рукопечатных символов), моделируемых заданным способом в виде векторов признакового пространства. Существенным в описываемом случае являлось то, что множество, их содержащее, имеет большой объем, а сам выбор этих объектов является случайным.

Требовалось провести исследование структуры матрицы весовых коэффициентов, используемой при распознавании, а также статистических свойств векторов первичных признаков и функции оценки на всем обучающем множестве. Для этого необходимо было разработать соответствующую методику, с помощью которой можно было бы решить поставленную задачу. Рассматриваемая проблематика является весьма актуальной, поскольку результаты подобного анализа можно использовать, в частности, для оптимизации матрицы распознавания. Заметим, что пока еще отсутствует некий стандартный способ проведения такого рода исследований. В настоящей работе описаны численные подходы к решению указанной задачи. При этом используется метод распознавания символов, разработанный авторами [2].

2. Метод распознавания

По растру изображения требуется определить, которому из K символов он соответствует. Нормализованный растр состоит из N=N1×N2 серых пикселов, перенумерованных в диапазоне $1\le i\le N$. Вводится в рассмотрение вектор $\mathbf{v}\in\mathbf{R}^N$, i-я компонента которого соответствует яркости i-го пиксела, лежащей на отрезке [0,1] для рассматриваемых серых растров.

Отождествляем k-й символ с базисным вектором $\mathbf{e}_k$=(0…1…0) (1 на k-м месте, $1\le k\le K$) из $\mathbf{R}^K$. Обозначаем $Y=\{\mathbf{e}_1,\ldots,\mathbf{e}_K\}$.

Пусть $\mathrm{p}_k(\mathbf{v})$ – вероятность того, что растр изображает символ с номером k, $1\le k\le K$. Распознанным считается символ с порядковым номером k_o, где

$$\mathrm{p}_{ko}(\mathbf{v})=\max\, \mathrm{p}_k(\mathbf{v}),\ \ 1\le k\le K. \tag{1}$$

Приближенные значения компонент $(\mathrm{p}_1(\mathbf{v}),\ldots,\mathrm{p}_K(\mathbf{v}))$ представляются в виде многочленов от координат $\mathbf{v}=(\mathrm{v}_1,\ldots,\mathrm{v}_\mathrm{N})$:

$$\mathrm{p}_k(\mathbf{v}) \cong c_0^{(k)} + \sum_{i=1}^{N} c_i^{(k)} v_i + \sum_{i,j=1}^{N} c_{i,j}^{(k)} v_i v_j + \ldots,\ 1 \le k \le K. \tag{2}$$

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

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

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

$$\mathbf{p}(\mathbf{v}) = (\mathrm{p}_1(\mathbf{v}),\ldots,\mathrm{p}_K(\mathbf{v}))^\mathrm{T} \cong A^\mathrm{T}\mathbf{x}(\mathbf{v}), \tag{3}$$

где A – матрица размера $L\times K$, столбцами которой являются векторы $\mathbf{a}^{(1)},\ldots,\mathbf{a}^{(K)}$. 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 A.

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

$$A \cong \Big(\frac{1}{J}\sum_{j=1}^{J}\mathbf{x}^{(j)}(\mathbf{x}^{(j)})^\mathrm{T}\Big)^{-1}\Big(\frac{1}{J}\sum_{j=1}^{J}\mathbf{x}^{(j)}(\mathbf{y}^{(j)})^\mathrm{T}\Big). \tag{4}$$

При получении правой части (4) используется следующая рекуррентная процедура, где A_0 и G_0 заданы:

$$A_j = A_{j\text{-}1}\text{-}\alpha_j G_j \mathbf{x}^{(j)}[A_{j\text{-}1}^{\mathrm{T}}\mathbf{x}^{(j)}\text{-}\mathbf{y}^{(j)}]^{\mathrm{T}}, \quad \alpha_j = 1/J;$$

$$G_j = \frac{1}{1\text{-}\alpha_j}[G_{j\text{-}1}\text{-}\alpha_j \frac{G_{j\text{-}1}\mathbf{x}^{(j)}(\mathbf{x}^{(j)})^{\mathrm{T}}G_{j\text{-}1}}{1+\alpha_j((\mathbf{x}^{(j)})^{\mathrm{T}}G_{j\text{-}1}\mathbf{x}^{(j)}\text{-}1)}]; \qquad 1 \le j \le J \tag{5}$$

$$G_j \equiv D^{-1},\ D = \mathrm{diag}\ (\mathrm{E}\{x_1^2\}, \mathrm{E}\{x_2^2\},\ldots, \mathrm{E}\{x_L^2\}).$$

Здесь $x_1, x_2,\ldots,x_L$ – компоненты вектора $\mathbf{x}(\mathbf{v})$. 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.

Обучение и распознавание проводилось на серых растрах, состоящих из N=16×16=256 пикселов. Для рукопечатных цифр использовалась следующая модификация вектора $\mathbf{x}(\mathbf{v})$:

$$\begin{gathered}\mathbf{x} = (1, \{\mathrm{v_i}\}, \{\mathrm{v_i}^2\}, \{(\delta \mathrm{v_i})_r\}, \{(\delta \mathrm{v_i})_r^2\}, \{(\delta \mathrm{v_i})_y\}, \{(\delta \mathrm{v_i})_y^2\},\\ \{(\delta \mathrm{v_i})_r^4\}, \{(\delta \mathrm{v_i})_y^4\}, \{(\delta \mathrm{v_i})_r(\delta \mathrm{v_i})_y\}, \{(\delta \mathrm{v_i})_r^2(\delta \mathrm{v_i})_y^2\}, \{(\delta \mathrm{v_i})_r^4(\delta \mathrm{v_i})_y^4\},\\ \{(\delta \mathrm{v_i})_r((\delta \mathrm{v_i})_r)_L\},\{(\delta \mathrm{v_i})_y((\delta \mathrm{v_i})_y)_L\},\{(\delta \mathrm{v_i})_r((\delta \mathrm{v_i})_y)_L\},\\ \{(\delta \mathrm{v_i})_y((\delta \mathrm{v_i})_r)_L\}, \{(\delta \mathrm{v_i})_r((\delta \mathrm{v_i})_r)_D\}, \{(\delta \mathrm{v_i})_y((\delta \mathrm{v_i})_y)_D\},\\ \{(\delta \mathrm{v_i})_r((\delta \mathrm{v_i})_y)_D\}, \{(\delta \mathrm{v_i})_y((\delta \mathrm{v_i})_r)_D\}).\end{gathered} \tag{6}$$

В (6) выражения в фигурных скобках соответствуют цепочкам элементов вектора, вычисляемым по всем пикселам растра (за исключением указанных ниже случаев). Через $(\delta \mathrm{v_i})_r$ и $(\delta \mathrm{v_i})_y$ обозначены конечные центральные разности величин $\mathrm{v_i}$ 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 пикселов растра; с индексом L – кроме левых граничных; с индексом D – кроме нижних граничных пикселей. Вне растра считаем, что $\mathrm{v_i}$=0 (используется при вычислении конечных разностей на границе растра).

3. Статистический анализ функции оценки

Как обучение, так и распознавание проводилось на базе рукопечатных цифр из 174778 элементов.

Пусть на вход распознавателя поступают перенумерованные изображения ($1 \le j \le J_{k_0}$) только одного определенного k_0-го символа (цифр «0», «1», …, «9»). Каждое из них будет распознано в качестве k-го символа с оценкой p_k^j для $1 \le k \le K$, где K = 10. Согласно (2) и (3),

$$p_k^j = \sum_{l=1}^{L} a^{lk} x_l^j, \qquad 1 \le j \le J_{k_0},\ k = 1,\ldots,K, \tag{7}$$

где a^{lk} – весовой коэффициент, x_l^j – моном.

3.1. Распределение $a^{lk}x_l$. Определим, в каком диапазоне лежат составляющие $a^{lk}x_l^j$, входящие в виде слагаемых в формулу вычисления оценки распознавания, а также как они количественно распределены внутри этого диапазона. В случае когда учитываются все альтернативы распознавания, при поступлении на вход распознавателя k_0-го символа нижняя и верхняя границы этого диапазона (t и T соответственно) будут вычисляться следующим образом:

$$t = \min_{l,j,k} (a^{lk} x_l^j);$$
$$1 \le l \le L,\ 1 \le j \le J_{k_0},\ k = 1,\ldots,K; \tag{8}$$
$$T = \max_{l,j,k} (a^{lk} x_l^j).$$

Для различных входящих символов, рассматриваемых по отдельности (столбцы «0», «1», …, «9»), диапазон совокупности $\{a^{lk} x_l^j\}$ несколько отличается (табл.1).

Таблица 1(1)

	«0»	«1»	«2»	«3»	«4»
$\min_{l,j,k} (a^{lk} x_l^j)$	-0,078	-0,078	-0,066	-0,066	-0,078
$\max_{l,j,k} (a^{lk} x_l^j)$	0,090	0,105	0,062	0,048	0,082

Таблица 1(2)

	«5»	«6»	«7»	«8»	«9»
$\min_{l,j,k} (a^{lk} x_l^j)$	-0,078	-0,070	-0,070	-0,070	-0,078
$\max_{l,j,k} (a^{lk} x_l^j)$	0,082	0,071	0,105	0,082	0,078

Для каждого из символов «0», «1», …, «9» делим соответствующий ему диапазон $[\min_{l,j,k}(a^{lk} x_l^j),\ \max_{l,j,k}(a^{lk} x_l^j)]$ на 10 равных частей и определяем, какое количество значений $a^{lk} x_l^j$ оказывается на каждом таком отрезке (табл.2).

Таблица 2(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	320	320	6090	952434	4336661914
«1»	1607	3103	175049	20191965	7333537653
«2»	1224	419	6348	176963	55398441
«3»	541	62	1675	20160	994145
«4»	820	143	12100	316047	2828627124
«5»	222	972	2751	187268	2882043338
«6»	28	528	3143	198343	2495766294
«7»	566	2132	81656	263090009	3319159966
«8»	38	74	2922	603426	2241768329
«9»	48	123	857	83092	4931824977

Таблица 2(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	5387991	29226	1223	279	2
«1»	1258409	50982	546	181	105
«2»	3933966739	795302	14669	1790	4
«3»	3002280683	24172351	223419	13160	304
«4»	32263394	119271	10053	30	17
«5»	27784202	126983	1447	12	4
«6»	396409862	232389	12732	737	51
«7»	115182	3753	34	9	3
«8»	2553206	10734	273	2	5
«9»	624536441	115560	1181	25	6

Здесь и далее цветом выделены максимумы полученных распределений числа значений $a^{lk} x_l^j$ для различных символов. Каждую строку табл. 2 отнормируем по числу изображений соответствующего символа. Результаты приведены в следующей таблице:

Таблица 2’(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	0,02	0,02	0,3	48,6	221484, 3
«1»	0,05	0,09	5,3	608,9	221156,1
«2»	0,07	0,02	0,4	9,8	3079,4
«3»	0,04	0,005	0,1	1,5	72,8
«4»	0,06	0,01	0,9	24,5	219273,4
«5»	0,02	0,07	0,2	14,3	219667,9

«6»	0,002	0,04	0,2	15,2	191378,4
«7»	0,04	0,1	5,1	16289, 4	205508,0
«8»	0,004	0,007	0,3	59,6	221496,7
«9»	0,0026	0,005	0,03	3,3	196871,4

Таблица 2'(2)

сим вол	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	275, 2	1,5	0,06	0,01	0,0001
«1»	37,9	1,5	0,02	0,005	0,003
«2»	218675,2	44,2	0,8	0,1	0,0002
«3»	219947,3	1770,9	16,4	1,0	0,02
«4»	2501,0	9,2	0,8	0,002	0,001
«5»	2117,7	9, 7	0,1	0,0009	0,0003
«6»	30397, 2	17,8	1,0	0,06	0,004
«7»	7,1	0,2	0,002	0,0006	0,0002
«8»	252, 3	1,1	0,03	0,0002	0,0005
«9»	24930,6	4,6	0,05	0,001	0,0002

Далее для изображений каждого отдельно взятого распознаваемого символа определим, в каком диапазоне лежат значения $a^{lk_0} x_l^j$, а также как они количественно распределены внутри данного диапазона, когда рассматривается единственная альтернатива распознавания, а именно, только сам этот символ. Для такой постановки задачи нижняя и верхняя границы соответствующего диапазона t^{k_0} и T^{k_0} при поступлении на вход распознавателя k_0-го символа будут вычисляться по следующим формулам:

$$t^{k_0} = \min_{l,j} (a^{lk_0} x_l^j);$$

$$1 \le l \le L,\ 1 \le j \le J_{k_0}; \tag{9}$$

$$T^{k_0} = \max_{l,j} (a^{lk_0} x_l^j).$$

В табл. 3 приведены указанные диапазоны:

Таблица 3(1)

	«0»	«1»	«2»	«3»	«4»
$\min_{l,j}(a^{lk_0} x_l^j)$	-0,066	-0,070	-0,035	-0,032	-0,078
$\max_{l,j}(a^{lk_0} x_l^j)$	0,027	0,105	0,036	0,035	0,049

Таблица 3(2)

	«5»	«6»	«7»	«8»	«9»
$\min\limits_{l,j}(a^{lk_0}x_l^j)$	-0,037	-0,047	-0,050	-0,037	-0,062
$\max\limits_{l,j}(a^{lk_0}x_l^j)$	0,048	0,041	0,053	0,051	0,068

Видно, что интервалы, границы которых приведены в табл.3, не выходят за пределы интервалов с границами из табл.2.

Аналогично табл. 2 для каждого из символов делим соответствующий диапазон с границами из табл.3 на 10 равных частей и определяем, какое количество значений $a^{lk_0}x_l^j$ оказывается на каждом таком отрезке (табл.4).

Таблица 4(1)

сим вол	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	316	100	902	976	4367
«1»	335	1201	38626	609398214	125774160
«2»	40	436	1844	69548	356353319
«3»	536	1971	16032	509320	295949638
«4»	25	21	2428	8890	47870
«5»	56	1628	10161	962030	289355849
«6»	46	103	1879	23757	1209802
«7»	2	130	1525	64582	349502628
«8»	104	430	29078	3999319	220190804
«9»	27	36	1331	183014	551197375

Таблица 4(2)

сим вол	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	33395	106599104	423151268	392179	60573
«1»	274180	34974	47	118	105
«2»	42048485	461955	89998	10028	534
«3»	5826163	387427	69390	7305	164
«4»	3779819	282107669	186210	1077	891
«5»	570150	86349	17639	10740	118
«6»	287285192	676689	55972	7000	1972
«7»	8588054	81004	6753	641	7
«8»	260061	13093	908	33	10
«9»	4252822	21500	101	20	2

Каждую строку таблицы отнормировали по числу изображений соответствующего символа и получили следующие результаты:

Таблица 4'(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	0,02	0,005	0,05	0,05	0,2
«1»	0,01	0,04	1,2	18377,5	3792,9
«2»	0,002	0,02	0,1	3,9	19808,4
«3»	0,04	0,1	1,2	37,3	21681,3
«4»	0,002	0,002	0,2	0, 7	3,7
«5»	0,004	0,1	0,8	73,3	22054, 6
«6»	0,004	0,008	0,1	1,8	92, 8
«7»	0,0001	0,008	0,09	4,0	21639,7
«8»	0,01	0,04	2,9	395,2	21755,8
«9»	0,001	0,001	0,05	7,3	22003,0

Таблица 4'(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	1,7	5444,3	21611,4	20,0	3,1
«1»	8,3	1,1	0,001	0,004	0,003
«2»	2337,3	25,7	5,0	0,6	0,03
«3»	426,8	28, 4	5,1	0,5	0,01
«4»	293,0	21868,8	14,4	0,08	0,07
«5»	43,5	6,6	1,3	0,8	0,009
«6»	22029, 4	51,9	4,3	0,5	0,2
«7»	531,7	5,0	0,4	0,04	0,0004
«8»	25,7	1, 3	0,09	0,003	0,001
«9»	169,8	0,9	0,004	0,0008	0,00008

При рассмотрении случая, когда в качестве альтернативы распознавания выбирается только входящий символ, распределение числа значений $a^{lk_0} x_l^j$ для полных диапазонов (полученных с учетом всех альтернатив), указанных в табл.1, будет следующим:

Таблица 5(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	282	161	2053	35263	433437475
«1»	290	403	11129	1571468	733361474
«2»	0	0	43	1805	3414089
«3»	0	0	445	1162	61208
«4»	31	38	7243	65616	281180867
«5»	0	0	50	5463	287674508

«6»	0	6	237	14832	249653135
«7»	0	3	2480	26104541	332106569
«8»	0	0	174	50306	223987063
«9»	0	31	94	19569	468575508

Таблица 5(2)

сим вол	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	818894	9852	0	0	0
«1»	536621	40037	311	122	105
«2»	395411974	206800	1479	0	0
«3»	298343311	4259772	98355	6397	0
«4»	4863632	16573	900	0	0
«5»	3255027	79554	118	0	0
«6»	39528748	63457	2006	0	0
«7»	31505	228	5	0	0
«8»	453801	2541	16	0	0
«9»	87034713	26217	92	4	3

Нормировка каждой строки таблицы по числу изображений соответствующего ей символа приводит к таким результатам:

Таблица 5'(1)

сим вол	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	0,01	0,008	0,1	1,8	22136,8
«1»	0,009	0,01	0,3	47,4	22115,8
«2»	0	0	0,002	0,1	189, 8
«3»	0	0	0,03	0,09	4,5
«4»	0,002	0,003	0,6	5,09	21797,0
«5»	0	0	0,004	0,4	21926,4
«6»	0	0,0005	0,02	1,1	19143,7
«7»	0	0,0002	0,2	1616,3	20562,6
«8»	0	0	0,02	5,0	22130,9
«9»	0	0,001	0,004	0,8	18704,9

Таблица 5'(2)

сим вол	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	41,8	0,5	0	0	0

«1»	16,2	1,2	0,009	0,004	0,003
«2»	21979,5	11, 5	0,08	0	0
«3»	21856,7	312,1	7,2	0,5	0
«4»	377,03	1,3	0,07	0	0
«5»	248, 1	6,1	0,009	0	0
«6»	3031,1	4, 9	0,2	0	0
«7»	2,0	0,01	0,0003	0	0
«8»	44,8	0,3	0,002	0	0
«9»	3474,3	1,0	0,004	0,0002	0,0001

Сопоставление результатов, приведенных в табл.2 и табл.5, показывает сходство распределения числа значений $a^{lk_0} x_l^j$ и $a^{lk} x_l^j$ при распознавании соответственно «своего» и «чужого» символа (аналогично для табл.2’ и табл.5’). При этом наблюдается большое различие в распределении числа значений $a^{lk_0} x_l^j$ внутри «своих» диапазонов (табл.4, 4’).

Возьмем минимальное значение из приведенных в табл.3 левых границ этих «своих» диапазонов и максимальное значение для правых границ:

$$f = \min_{k_0} \min_{l,j} (a^{lk_0} x_l^j);$$
$$F = \max_{k_0} \max_{l,j} (a^{lk_0} x_l^j). \qquad (10)$$

Аналогично табл.4 и табл.5 рассмотрим случай, когда в качестве альтернативы распознавания выбирается только сам символ. Тогда распределение числа значений $a^{lk_0} x_l^j$ на отрезке [f, F] будет следующим:

Таблица 6(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	282	397	4018	352603	433750039
«1»	290	403	11129	1571468	733361474
«2»	0	0	328	282999	398636124
«3»	0	0	809	549409	302092793
«4»	32	47	15794	705280	285319585
«5»	0	0	1206	230743	290656158
«6»	0	7	952	454374	288713075
«7»	0	2	675	581851	357605807
«8»	0	0	220	489185	223970410
«9»	17	29	1189	1639274	553960970

Таблица 6(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	196641	0	0	0	0
«1»	536621	40037	311	122	105
«2»	116313	426	0	0	0
«3»	124874	2765	0	0	0
«4»	93202	960	0	0	0
«5»	116046	10567	0	0	0
«6»	92032	1981	0	0	0
«7»	56343	645	8	0	0
«8»	34033	52	1	0	0
«9»	54654	92	3	3	0

При этом распределение числа рассматриваемых слагаемых, включая расположение отрезков с их максимальным количеством, оказывается полностью синхронным для различных символов.

Каждую строку таблицы отнормируем по числу изображений соответствующего символа. Результаты приведены в следующей таблице:

Таблица 6'(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	0,01	0,02	0,2	18,0	22152,7
«1»	0,009	0,01	0,3	47,4	22115,8
«2»	0	0	0,02	15,7	22158,8
«3»	0	0	0,06	40,3	22131,3
«4»	0,002	0,004	1,2	54, 7	22117,8
«5»	0	0	0,09	17,6	22153,7
«6»	0	0,0005	0,07	34,8	22138,9
«7»	0	0,0001	0,04	36,0	22141,4
«8»	0	0	0,02	48,3	22129,3
«9»	0,0007	0,001	0,05	65,4	22113,3

Таблица 6'(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	10,0	0	0	0	0
«1»	16,2	1,2	0,009	0,004	0,003
«2»	6, 5	0,02	0	0	0
«3»	9,1	0,2	0	0	0
«4»	7,2	0,07	0	0	0

«5»	8,8	0,8	0	0	0
«6»	7,1	0,2	0	0	0
«7»	3, 5	0,04	0,0005	0	0
«8»	3,4	0,005	0,0001	0	0
«9»	2,2	0,004	0,0001	0,0001	0

Числа рассматриваемых слагаемых на совпавших отрезках с их максимальным количеством в значительной степени схожи для различных символов.

Значения $a^{lk_0} x_l^j$ в пятом интервале табл.6 и 6’ лежат в окрестности нуля, причем чистому нулю соответствует следующее количество слагаемых:

Таблица 7(1)

	«0»	«1»	«2»	«3»	«4»
$x_l^j = 0$	199089175	398742920	184680248	141416771	142594020
$x_l^j = 0$	10168,0	12024,8	10265,7	10360,2	11053,8

Таблица 7(2)

	«5»	«6»	«7»	«8»	«9»
$x_l^j = 0$	144694678	127596132	177594198	96439123	244126606
$x_l^j = 0$	11028,6	9784,2	10995,9	9528,6	9745,2

Значения во второй строке табл.7 соответствуют величинам их первой строки, отнормированным по числу изображений входящего символа.

Следует заметить, что количество слагаемых, для которых $a^{lk_0} x_l^j = 0$, совпадает с их числом при $x_l^j = 0$. При этом оказалось, что $a^{lk_0} = 0$ лишь для очень незначительного числа слагаемых. Для каждого из рассматриваемых символов таковых оказалось всего лишь 16. Это служит основанием для того, чтобы сделать вывод об оптимальности полученного набора мономов, за вычетом указанного незначительного количества.

3.2. Распределение вариации $a^{lk} x_l$. При поступлении на вход различных образов одного и того же символа с номером k_0 определим, в каком диапазоне находятся вариации составляющих $a^{lk} x_l^j$ и как они распределены внутри этого диапазона. Для каждого l ($1 \le l \le L$) величина $a^{lk} x_l^j$ варьируется в некотором диапазоне:

$$\sigma_l = \max_{j,k}(a^{lk} x_l^j) - \min_{j,k}(a^{lk} x_l^j) \ge 0, \qquad 1 \le j \le J_{k_0},\ 1 \le k \le K, \tag{11}$$

при рассмотрении всех альтернатив распознавания. Если принимается во внимание только одна альтернатива, совпадающая с входящим символом, имеющим номер k_0, то

$$\sigma_l^{k_0} = \max_j(a^{lk_0} x_l^j) - \min_j(a^{lk_0} x_l^j) \geq 0, \qquad 1 \leq j \leq J_{k_0}. \qquad (12)$$

Определим минимальную и максимальную величину такой вариации

$$r = \min_l(\sigma_l) = 0,\ R = \max_l(\sigma_l) \qquad (13)$$

для всего набора альтернатив или

$$r_{k_0} = \min_l(\sigma_l^{k_0}) = 0,\ R^{k_0} = \max_l(\sigma_l^{k_0}) \qquad (14)$$

для одной альтернативы, совпадающей со входящим символом. В табл. 8 приведены значения R и R^{k_0} – верхняя и нижняя строка соответственно).

Таблица 8(1)

	«0»	«1»	«2»	«3»	«4»
10 альтернатив	0,143	0,147	0,116	0,110	0,143
1 альтернатива	0,066	0,105	0,049	0,043	0,078

Таблица 8(2)

	«5»	«6»	«7»	«8»	«9»
10 альтернатив	0,143	0,126	0,147	0,126	0,143
1 альтернатива	0,048	0,051	0,053	0,051	0,068

Поделим полученный отрезок $[0, R]$ на десять равных частей и выясним, какое количество значений l имеет вариации, относящиеся к каждому такому маленькому отрезку. В табл. 9 приведены результаты для случая, когда рассматриваются все альтернативы распознавания.

Таблица 9(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	18371	3059	563	102	42
«1»	18322	3095	565	110	41
«2»	16830	4008	976	235	64
«3»	16635	4125	1032	260	65
«4»	18278	3137	567	112	40
«5»	18410	3081	533	87	32
«6»	17433	3695	784	177	41

«7»	18496	2995	520	97	36
«8»	17315	3737	829	196	49
«9»	18237	3179	584	107	36

Таблица 9(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	12	4	8	3	1
«1»	10	8	7	2	2
«2»	30	8	7	4	3
«3»	31	8	6	0	3
«4»	13	2	4	2	1
«5»	11	3	6	1	1
«6»	15	9	4	5	2
«7»	7	6	5	2	1
«8»	16	10	4	6	3
«9»	11	2	7	1	1

Аналогичные результаты для интервала [0, R_{k0}] при совпадении альтернативы распознавания и входящего кода приведены в табл.10.

Таблица 10(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	21415	627	73	31	7
«1»	21009	952	150	39	7
«2»	19754	1691	446	147	69
«3»	18241	2436	880	360	130
«4»	20452	1314	290	72	24
«5»	19845	1567	450	178	78
«6»	19709	1670	535	165	40
«7»	19271	2034	568	185	63
«8»	18941	2171	626	259	92
«9»	19403	1960	525	174	65

Таблица 10(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	4	2	4	1	1
«1»	3	3	1	0	1
«2»	39	8	6	1	4

«3»	65	29	17	2	5
«4»	10	1	0	0	2
«5»	26	15	2	0	4
«6»	19	11	9	5	2
«7»	29	6	4	2	3
«8»	43	23	5	2	3
«9»	24	10	2	0	2

Количество значений l с вариациями, относящимися к отрезкам полных диапазонов $[0,\ R]$, приведены в табл.11 для случая, когда альтернатива распознавания совпадает со входящим символом.

Таблица 11(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	22165	84	9	5	2
«1»	21611	487	54	6	5
«2»	21675	422	58	7	3
«3»	21288	767	96	14	0
«4»	21654	459	43	7	1
«5»	21860	284	17	4	0
«6»	21722	395	36	11	1
«7»	21790	348	24	3	0
«8»	21467	618	73	6	1
«9»	21432	651	71	9	2

Таблица 11(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	0	0	0	0	0
«1»	1	0	1	0	0
«2»	0	0	0	0	0
«3»	0	0	0	0	0
«4»	1	0	0	0	0
«5»	0	0	0	0	0
«6»	0	0	0	0	0
«7»	0	0	0	0	0
«8»	0	0	0	0	0
«9»	0	0	0	0	0

Видно, что распределение для случая, когда рассматривается только «своя» альтернатива распознавания, соответствует большей доле значений l с незначительными вариациями компонент $a^{lk}x_l^j$ для элементов обучающего множества. При рассмотрении всех альтернатив распознавания это распределение «вытягивается» и становится более пологим.

Полученные результаты, относящиеся к распределению составляющих $a^{lk}x_l^j$ и их вариации, являются обоснованием устойчивости распознавания. Во-первых, максимум обоих распределений находится в окрестности нуля, а во-вторых, в этом интервале находится подавляющее большинство соответствующих значений. Следовательно, искажение исходных изображений, заключающееся в изменении некоторого количества мономов, приводит к незначительным поправкам в оценке посредством ее составляющих $a^{lk}x_l^j$.

4. Статистический анализ вектора первичных признаков

4.1. Распределение x_l^j. Определим, в каком диапазоне лежат компоненты вектора первичных признаков и как они распределены внутри этого диапазона. Пусть на вход распознавателя поступают изображения одного и того же символа с номером k_0. Величина x_l^j для $1 \le l \le L$, $1 \le j \le J_{k_0}$ находится в некотором диапазоне [s, S]:

$$s = \min_{j,l}(x_l^j),\ S = \max_{j,l}(x_l^j),\ \text{где } 1 \le l \le L,\ 1 \le j \le J_{k_0}. \qquad (15)$$

Для различных входящих символов, рассматриваемых по отдельности, получены одинаковые значения величин s и S:

$$s = -1,\ S = 2. \qquad (16)$$

Поделим полученный отрезок [s, S] на десять равных частей и выясним, какое количество значений x_l^j относится к каждому такому маленькому отрезку (табл. 12) для рассматриваемых по отдельности символов.

Таблица 12(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	536117	3577779	7490094	307059199	32520201
«1»	700651	4135073	7524167	530180585	40664208
«2»	473692	3027347	6637061	284937069	29216294
«3»	374536	2276928	5436444	215501326	22015404
«4»	368767	1956690	3809750	206903513	18765335
«5»	385532	1998741	4688691	209693933	18928628
«6»	371998	2371479	5403767	204635267	22340900
«7»	488475	2511934	5162982	257202330	23837637

«8»	282450	1818259	4396597	158383616	17502304
«9»	757779	4493264	9964870	394569455	43121518

Таблица 12(2)

сим вол	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	19158898	60471815	1917240	1162371	410266
«1»	25026811	122914889	2363891	1393602	618083
«2»	16955523	54507417	1754194	1040486	487107
«3»	12881160	41786841	1357040	784089	356882
«4»	11303132	40956685	1089518	703351	278159
«5»	11164504	42097407	1125208	658349	273727
«6»	13321306	38356767	1344986	779203	336748
«7»	14494593	52147048	1325002	757158	318172
«8»	10459153	29675642	1100160	615148	260572
«9»	25252053	72635986	2665306	1538132	657868

Каждую строку таблицы отнормируем по числу изображений соответствующего символа. Результаты приведены в следующей таблице:

Таблица12’(1)

сим вол	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	27,4	182,7	382,5	15682,3	1660,9
«1»	21,1	124,7	226,9	15988, 6	1226,3
«2»	26,3	168,3	368,9	15838,6	1624,0
«3»	27,4	166,8	398,3	15787,6	1612,9
«4»	28,6	151,7	295,3	16039,0	1454,7
«5»	29,4	152,3	357,4	15982,8	1442,7
«6»	28,5	181,8	414,4	15691,7	1713,1
«7»	30,2	155,5	319,7	15924,9	1475,9
«8»	27,9	179,7	434,4	15649,0	1729,3
«9»	30,2	179,4	397,8	15750,7	1721,3

Таблица12’(2)

сим вол	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	978,5	3088,4	97,9	59,4	21,0
«1»	754,7	3706,7	71,3	42,0	18,6
«2»	942,5	3029,9	97,5	57,8	27,1
«3»	943, 7	3061,3	99,4	57,4	26,1
«4»	876,2	3174,9	84,5	54,5	21,6

«5»	851,0	3208,7	85, 8	50,2	20,7
«6»	1021,5	2941,2	103,1	59,8	25,8
«7»	897,4	3228,7	82,0	46,9	19,7
«8»	1033,4	2932,1	108,7	60,8	25,7
«9»	1008,0	2899,5	106,4	61,4	26, 3

Значения x_l^j в четвертом интервале табл.12 и 12’ лежат в окрестности нуля, причем чистому нулю соответствует количество слагаемых, указанное соответственно в первой или второй строке табл.7.

Распределение величины $a^{lk_0} x_l^j$ (табл.6 и 6’) схоже с распределением x_l^j (табл.12 и 12’), однако последнее является более пологим и имеет менее ярко выраженный максимум.

4.2. Распределение вариации x_l^j. Определим, в каком диапазоне находятся вариации компонент вектора первичных признаков и как они распределены внутри этого диапазона. Пусть на вход распознавателя поступают изображения одного и того же символа с номером k_0. Для каждого l ($1 \le l \le L$) величина x_l^j варьируется в некотором диапазоне:

$$\delta_l = \max_j (x_l^j) - \min_j (x_l^j) \geq 0, \text{ где } 1 \le l \le L,\ 1 \le j \le J_{k_0}. \tag{17}$$

Определим минимальную и максимальную величину такой вариации:

$$d = \min_l (\delta_l),\ D = \max_l (\delta_l), \text{ где } 1 \le l \le L. \tag{18}$$

Для различных входящих символов, рассматриваемых по отдельности, получены одинаковые значения величин d и D:

$$d = 0,\ D = 2. \tag{19}$$

Поделим полученный отрезок $[d, D]$ на десять равных частей и выясним, какое количество значений l имеет вариации δ_l, относящиеся к каждому такому маленькому отрезку (табл. 13) для каждого из символов.

Количество вариаций на каждом маленьком отрезке несколько отличается для различных входящих символов (табл.13). Суммарное количество вариаций в каждой строчке равно L.

Таблица 13(1)

сим вол	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	641	1881	3766	146	857

«1»	519	1706	3804	20	76
«2»	590	1584	4038	132	744
«3»	730	1820	3885	172	903
«4»	572	1596	4021	92	404
«5»	642	1978	3704	146	970
«6»	558	1899	3715	137	647
«7»	629	1754	3899	123	556
«8»	515	1578	3966	69	304
«9»	533	1595	3984	86	448

Таблица 13(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	12987	122	829	70	866
«1»	14159	68	884	55	874
«2»	13185	74	878	40	900
«3»	12801	79	853	57	865
«4»	13580	49	906	33	912
«5»	12873	120	816	81	835
«6»	13355	104	844	61	845
«7»	13349	85	839	72	859
«8»	13849	65	879	51	889
«9»	13646	53	886	42	892

В распределении вариаций мономов для различных символов наблюдается сходство (табл.13), однако оно существенно отличается от распределения вариации $a^{lk} x_l^j$ (табл.10).

5. Анализ структуры матрицы весовых коэффициентов

Определим, в каком диапазоне лежат значения весовых коэффициентов a^{lk}, входящие в виде множителей в слагаемые $a^{lk} x_l^j$ при вычислении оценки распознавания, а также как они распределены внутри этого диапазона. Для вычисления оценок различных символов как альтернатив распознавания, рассматриваемых по отдельности (фиксированные $k = 1,\ldots,K$), диапазон совокупности указанных весовых коэффициентов несколько отличается (табл.14).

Таблица 14(1)

	«0»	«1»	«2»	«3»	«4»
$\min_l (a^{lk})$	-0,139	-0,321	-0,347	-0,210	-0,295

$\max_l(a^{lk})$	0,209	0,244	0,237	0,268	0,302

Таблица 14(2)

	«5»	«6»	«7»	«8»	«9»
$\min_l(a^{lk})$	-0,376	-0,176	-0,295	-0,254	-0,291
$\max_l(a^{lk})$	0,188	0,271	0,213	0,292	0,311

Для каждого из символов делим соответствующий ему диапазон [$\min_l(a^{lk})$, $\max_l(a^{lk})$] на 10 равных частей и определяем, какое количество указанных весовых коэффициентов оказывается на каждом таком отрезке (табл.15).

Таблица 15(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	5	23	79	14748	7251
«1»	3	5	11	34	166
«2»	1	1	2	18	56
«3»	11	25	49	567	21134
«4»	6	9	26	92	18761
«5»	1	1	2	4	23
«6»	7	15	105	18982	2934
«7»	3	4	14	42	157
«8»	6	16	56	256	21251
«9»	4	1	25	173	20483

Таблица 15(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	59	9	4	2	1
«1»	21212	690	45	11	3
«2»	19074	2925	82	7	5
«3»	307	55	22	10	1
«4»	3195	67	17	5	3
«5»	144	21427	542	24	12
«6»	98	24	10	3	2
«7»	20720	1167	51	19	3
«8»	538	29	18	6	4
«9»	1348	97	34	14	1

Суммарное количество весовых коэффициентов в каждой строчке равно L. В распределении этих слагаемых внутри «своих» диапазонов (табл.15) наблюдается асинхронность для разных символов.

Диапазон, в котором лежат значения всех весовых коэффициентов a^{lk} соответствующей матрицы, определяется двумя значениями:

$$\min_{l,k}(a^{lk}) = -0{,}376,\ \max_{l,k}(a^{lk}) = 0{,}311. \tag{20}$$

Делим этот диапазон $[\min_{l,k}(a^{lk}),\ \max_{l,k}(a^{lk})]$ на 10 равных частей и для различных символов по отдельности определяем, какое количество указанных весовых коэффициентов оказывается на каждом из этих отрезков (табл.16).

Таблица 16(1)

символ	1 интервал	2 интервал	3 интервал	4 интервал	5 интервал
«0»	0	0	0	5	116
«1»	1	4	9	30	230
«2»	1	1	3	23	141
«3»	0	0	6	42	214
«4»	0	3	14	38	214
«5»	1	1	3	22	193
«6»	0	0	1	15	201
«7»	0	3	8	40	257
«8»	0	1	7	51	314
«9»	0	3	2	34	406

Таблица 16(2)

символ	6 интервал	7 интервал	8 интервал	9 интервал	10 интервал
«0»	21988	65	6	1	0
«1»	21652	223	25	6	1
«2»	21783	199	23	7	0
«3»	21637	232	41	8	1
«4»	21651	225	27	6	3
«5»	21779	164	17	1	0
«6»	21761	178	17	5	3
«7»	21703	142	28	2	0
«8»	21589	181	25	8	5
«9»	21420	236	55	22	2

Суммарное количество весовых коэффициентов в каждой строчке также равно L.

Распределение числа рассматриваемых слагаемых, включая расположение отрезков с их максимальным количеством в окрестности нуля, оказывается полностью синхронным для различных символов.

6. Заключение

Для разработанного авторами метода распознавания символов предложена и реализована методология анализа механизма классификации. Проведенное на обучающем множестве рукопечатных символов большого объема статистическое исследование функции оценки позволило выявить особенности распределения составляющих $a^{lk} x_l^j$, входящих в виде слагаемых в формулу для ее вычисления, компонент вектора первичных признаков x_l^j, а также структуры матрицы весовых коэффициентов a^{lk}. Подтверждена оптимальность используемого набора мономов. Характер распределения составляющих $a^{lk} x_l^j$ и их вариации является обоснованием устойчивости распознавания.

7. Библиографический список

1. Мисюрев А.В. Использование искусственных нейронных сетей для распознавания рукопечатных символов // Сборник трудов ИСА РАН «Интеллектуальные технологии ввода и обработки информации», 1998, с. 122-127.

2. Гавриков М. Б., Пестрякова Н. В. Метод полиномиальной регрессии в задачах распознавания печатных и рукопечатных символов // Препринты ИПМ им.М.В.Келдыша. 2004. № 22. 12 с. URL: http://keldysh.ru/papers/2004/prep22/prep2004_22.html

3. Об одном методе распознавания символов, основанном на полиномиальной регрессии. / М. Б. Гавриков [и др.] // Автоматика и Телемеханика. 2006, № 2, с. 119-134.

4. Пестрякова Н. В. Динамика качества распознавания при нарастании степени различия баз обучения и распознавания // Информационные технологии и вычислительные системы. 2010, № 2, с. 75-82.