

Динамика качества распознавания при нарастании степени различия баз обучения и распознавания¹

Н.В. Пестрякова

Аннотация. Для разработанного автором метода распознавания символов, основанного на полиномиальной регрессии, выполнено исследование зависимости качества распознавания от степени различия между базой обучения и полученной путем ее модификации базой распознавания. Изучено поведение основанных на растрах и полиномиальных векторах характеристик базы распознавания. Результаты получены для печатных и рукопечатных цифр.

Ключевые слова: распознавание символов, база обучения, база распознавания, регрессионный анализ.

Введение

Настоящая работа является очередной в серии [1–10], посвященной методу распознавания символов, основанному на регрессионном анализе [11–20].

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: «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 характери-

стик 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 в работе серых растров.

Пусть для некоторого растра \mathbf{v} можно найти $p_k(\mathbf{v})$ – вероятность того, что растр изображает

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 10-07-00374)

символ с порядковым номером k , $1 \leq k \leq K$. Тогда распознанным считается символ с порядковым номером k_0 , где

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

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

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

$$1 \leq k \leq K.$$

Суммы в правых частях равенств (2) конечные и определяются выбором базисных мономов. Если $\mathbf{x}(\mathbf{v}) = (1, v_1, \dots, v_N, \dots)^T$ - конечный вектор размерности L из выбранных и приведенных в (2) базисных мономов, упорядоченных определенным образом, то в векторном виде соотношения (2) можно записать так:

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

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

Значение \mathbf{A} вычисляется приближенно в процессе обучения с использованием содержащихся в некоторой базе данных наборов 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:

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

$$\alpha = 1/J, 1 \leq j \leq J$$

$$\mathbf{G} \equiv \mathbf{D}^{-1}, \mathbf{D} = \text{diag}(E\{x_1^2\}, E\{x_2^2\}, \dots, E\{x_L^2\}),$$

$$E\{x_i^2\} \cong \frac{1}{J} \sum_{j=1}^J (\mathbf{x}_i^{(j)})^2, 1 \leq i \leq L.$$

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} :

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

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 , вычис-

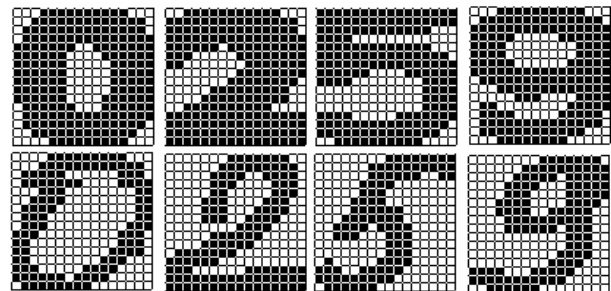


Рис. 1. Образы 16x16 печатных и рукопечатных цифр

ляются для всех пикселей растра, с индексом L – кроме левых граничных, с индексом D – кроме нижних граничных пикселей. Вне растра считаем, что $v_i = 0$ (используется при вычислении конечных разностей на границе растра).

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

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

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

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

Для «среднестатистического» полиномиального вектора конкретного символа по некоторой базе значение в каждой компоненте вектора, имеющей номер i , равно среднему арифметиче-

скому 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:

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

1.4. Модели затемнения и засветления

На этапе распознавания все пиксели растра постепенно затемняются: $v_i \rightarrow v_i + 0,01 \cdot n$, где $n = 0, 1, \dots, 100$. Если для каких-то пикселей, начиная с некоторого n , имеем: $v_i > 1$, то считаем $v_i = 1$. При засветлении аналогично $v_i \rightarrow v_i - 0,01 \cdot n$, где $n = 0, 1, \dots, 100$. Если для каких-то пикселей начиная с некоторого n , имеем: $v_i < 0$, то считаем $v_i = 0$.

2. Качество распознавания

На Рис. 2 приведена зависимость от n доли (%) числа нераспознанных изображений символов относительно общего количества изображений ($\text{mis}\%t$ при затемнении и $\text{mis}\%s$ при засветлении). Для рукопечатных символов эта величина в процессе затемнения увеличивается монотонно, а при засветлении имеется немонотонность при $n = 1$ (незначительная), а также при $64 \leq n \leq 72$. Следует отметить, что разность ($\text{mis}\%t - \text{mis}\%s$) начинает превышать 10% при $n \approx 29$.

Для печатных символов $\text{mis}\%t$ увеличивается монотонно, а $\text{mis}\%s$ имеет немонотонность на отрезке $0 \leq n \leq 8$. Сравнивая с результатом для рукопечатных цифр, имеем: $\text{mis}\%t$ больше, чем

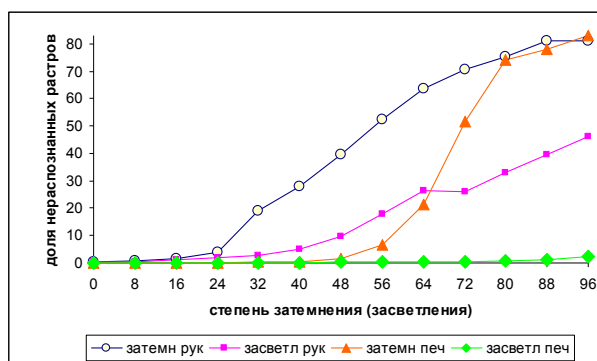


Рис. 2. « $\text{mis}\%t - \text{mis}\%s$ »

$\text{mis}\%s$ при $n > 0$, но разность ($\text{mis}\%t - \text{mis}\%s$) начинает превышать 10% лишь при $n \approx 59$.

Следует отметить «особенное» поведение $\text{mis}\%s$ для печатных цифр, а именно, при $n = 96$ получено $\text{mis}\%s = 2,44\%$. Можно сравнить со следующими результатами: для печатных символов $\text{mis}\%t = 82,9\%$, для рукопечатных $\text{mis}\%t = 81,03\%$, $\text{mis}\%s = 46,12\%$ при этом же значении n .

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 $n \approx 48$. Однако при затемнении печатных цифр доля ошибок для $n = 72$ становится больше, чем при засветлении рукопечатных, а для $n = 96$ – чем при затемнении последних.

Средняя оценка распознавания для рукопечатных символов и при затемнении (Prb_t), и при засветлении (Prb_s) сначала уменьшается, а затем увеличивается (Рис. 3). Для Prb_t средние темпы и падения, и роста в полтора-два раза выше, чем для Prb_s (раздел 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. 3) 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

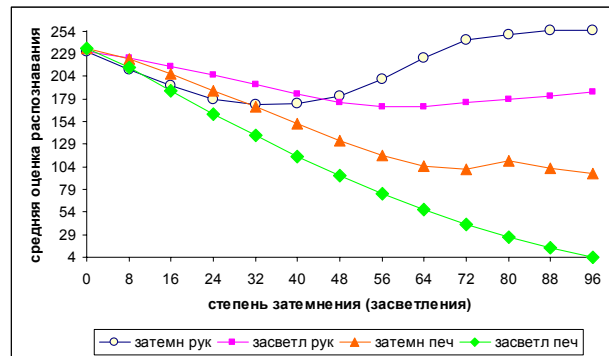


Рис. 3. « $\text{Prb}_t - \text{Prb}_s$ »

изображений печатных символов происходит со стремительным падением оценки распознавания, которая всюду при $n > 0$ меньше соответствующей оценки для затемнения печатных, а 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 $n = 32$.

3. Интегральные характеристики модифицированной базы

На Рис. 4 приведена зависимость от n расстояния между среднестатистическими растрами баз обучения и распознавания (vv_t при затемнении и vv_s при засветлении). На Рис. 5 а,б представлено отклонение между среднестатистическими полиномиальными векторами баз обучения и распознавания (xx_t при затемнении и xx_s при засветлении) как функция n . Все четыре величины получены независимо от правильности распознавания и их следует рассматривать как интегральные характеристики для базы распознавания в целом.

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 (Рис. 4). Во всем 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 . При каждом $n > 0$ они располагаются по возрастающей так: засветление (рукопечатные символы), засветление (печатные), затемнение (печатные), затемнение (рукопечатные). Объяснение приведено в разделе 8. Там же обосновывается следующее: для каждого $n > 0$ величина vv_t превышает vv_s для рукопечатных цифр более чем в два раза, а для печатных – при уже отмеченном мо-

нотонном увеличении расстояния между ними с ростом n – не более чем на 15%.

Расстояние между среднестатистическими полиномиальными векторами рукопечатных символов увеличивается с ростом n для затемнения – монотонно, а для засветления при $64 \leq n \leq 72$ эта величина незначительно уменьшается (Рис. 5а). В случае печатных символов указанное расстояние (Рис. 5б) монотонно нарастает, причем сначала после выхода из точки $(0, 0)$ эта величина при затемнении меньше, чем при засветлении, а при $n \approx 67$ ситуация меняется на противоположную.

Аналогично vv_t и vv_s , для каждого $n > 0$ величина xx_t более чем в полтора раза превосходит xx_s для рукопечатных цифр, а для печатных эти величины отличаются незначительно на большей части отрезка $0 \leq n \leq 100$, причем их разница приближается к 20% лишь около $n = 100$.

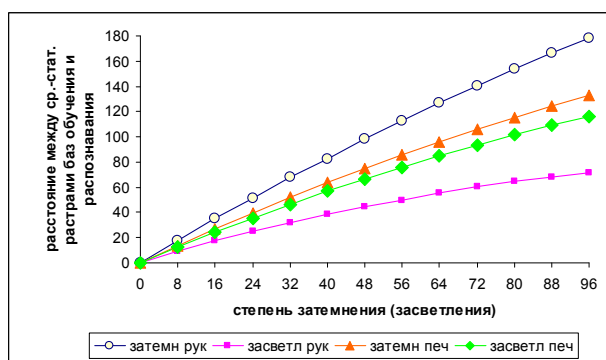
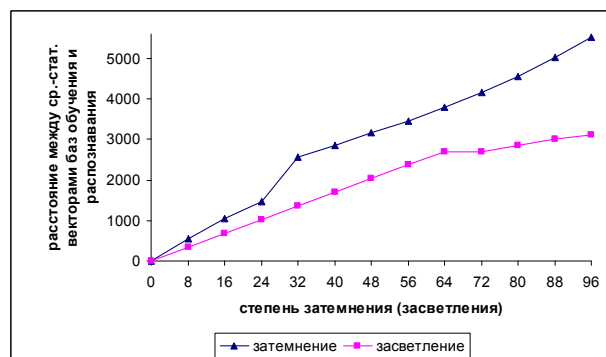
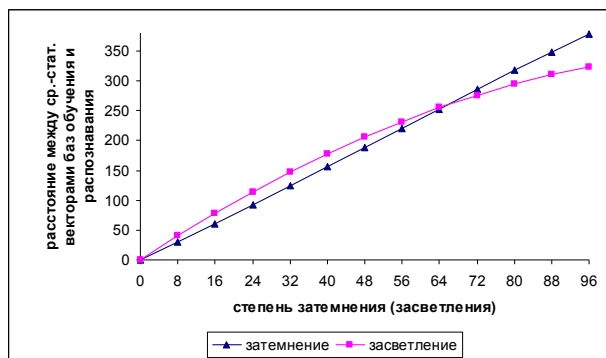


Рис. 4. « $vv_t - vv_s$ »



а) рукопечатные цифры



б) печатные цифры

Рис. 5. « xx_t, xx_s »

4. Корреляционные зависимости режима засветления для рукопечатных символов

По Рис.5а. 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, которая для $n > 72$ приблизительно в два раза ниже, чем для $n < 64$.

На этом же отрезке $64 \leq n \leq 72$ «перестраивается» и график, соответствующий изменению количества неправильно распознанных символов mis_s (Рис.2). 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 $n = 1$), 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.3).

5. Корреляционные зависимости режима затемнения для рукопечатных символов

На Рис.5а. 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.2). 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.3).

6. Особенности засветления и затемнения для печатных символов

Структура вектора для печатных цифр более простая, чем для рукопечатных. Это «промежу-

точная 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 доли нераспознанных изображений печатных символов также имеет меньше особенностей (*печатные*: почти монотонный рост, гладкость для засветления и нарушение ее лишь в одной точке для затемнения; *рукопечатные*: нарушение монотонности для засветления, большее нарушение гладкости для засветления и затемнения).

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). А именно, в разделе 2 отмечено, что 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 .

7. О сильном затемнении и засветлении

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. 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 удачная, 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.

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 матрице. 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.

Для печатных цифр при полном засветлении все образы распознаются как «1», но уже с минимальной оценкой $Prb_s = 1$, чем и объясняется монотонное убывание данной функции. При этом для печатных символов очень высока устойчивость правильного распознавания к засветлению (п.2).

8. Расстояние между растрами «Белого» и «Черного» квадратов

При $n = 100$ для рукопечатных цифр выполняются равенства: $vv_t = 183,8$ (отклонение растра «Черного квадрата» от среднестатистического растра обучающей базы), а также $vv_s = 72,2$ (отклонение растра «Белого квадрата» от среднестатистического растра обучающей базы). Следовательно, $vv_t + vv_s = 256$. Указанная сумма соответствует расстоянию между растрами «Белого» и «Черного» квадратов.

Итак, отклонение среднестатистического растра базы обучения от растра «Черного» квадрата в два с половиной раза больше, чем от растра «Белого» квадрата. Этим объясняется, что для каждого $n > 0$ величина vv_t превышает vv_s более чем в два раза.

Для печатных цифр при $n = 100$ выполняются равенства: $vv_t = 136,6$ (отклонение растра «Черного квадрата» от среднестатистического растра обучающей базы), а также $vv_s = 119,4$ (отклонение растра «Белого квадрата» от сред-

нестатистического растра обучающей базы). Следовательно, $vv_t + vv_s = 256$. 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 этом увеличивается количество пикселей, в которых достигается значение «1» при затемнении и «0» при засветлении.

Заключение

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; они сходны с печатными и рукопечатными растровыми аналогами.

Доля 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 доли ошибок для печатных символов характеризует почти моно-

тонный 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 при соответствующем типе написания.

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

Для печатных символов при засветлении оценка стремительно монотонно падает, а при затемнении наблюдается общая тенденция к гораздо более медленному монотонному уменьшению (исключением является небольшой участок немонотонности). Следует заметить, что именно при этом значении имеется нарушение гладкости функции mis_t .

Хотя средняя оценка распознавания исходной базы для печатных цифр несколько больше, чем для рукопечатных, однако засветление изображений печатных символов происходит с быстрым падением оценки распознавания, которая уже на начальных этапах изменения базы становится сначала ниже оценки засветления, а затем затемнения рукопечатных символов. Оценка при затемнении печатных цифр падает медленнее, чем при засветлении, но и она становится меньше оценки вначале при засветлении рукопечатных, а затем при их затемнении.

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, Фарсоби́на В.В. "Развитие метода полиномиальной регрессии и практическое применение в задаче распознавания", //Препринт ИПМатем. РАН, М., 2006, №25, 21 стр.
3. Гавриков М.Б., Мисюрёв 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., Фарсоби́на 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., Фарсоби́на 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., Фарсоби́на 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., Фарсоби́на 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., Фарсоби́на В.В. "О среднестатистических 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., Фарсоби́на 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. Пестрякова Н.В. "Структуры в распознавании", Информационные технологии и вычислительные системы. 2009, №1, С. 58-71.
11. Sebestyen G.S. Decision Making Processes in Pattern Recognition, MacMillan, New York, 1962.
12. Nilson N. J. Learning Machines, McGraw-Hill, New York, 1965.
13. Schürmann J. Polynomklassifikatoren, Oldenbourg, München, 1977.
14. Schürmann J. Pattern Classification, John Wiley&Sons, Inc., 1996.
15. Albert A.E. and Gardner L.A. Stochastic Approximation and Nonlinear Regression // Research Monograph 42. MIT Press, Cambridge, MA, 1966.
16. Becker D. and Schürmann J. Zur verstärkten Berücksichtigung schlecht erkennbarer Zeichen in der Lernstichprobe // Wissenschaftliche Berichte AEG-Telefunken **45**, 1972, pp. 97 – 105.
17. 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.
18. 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.
19. Дж.Себер. Линейный регрессионный анализ. М.: "Мир", 1980.
20. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. М.: "Физматлит", 1958.