



Math-Net.Ru

Общероссийский математический портал

Б. М. Гавриков, М. Б. Гавриков, Н. В. Пестрякова, Статистический метод распознавания на основе нелинейной регрессии, *Матем. моделирование*, 2020, том 32, номер 4, 116–130

DOI: 10.20948/mm-2020-04-09

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.189.143.127

9 января 2025 г., 10:34:12



СТАТИСТИЧЕСКИЙ МЕТОД РАСПОЗНАВАНИЯ НА ОСНОВЕ НЕЛИНЕЙНОЙ РЕГРЕССИИ

© 2020 г. *Б.М. Гавриков¹, М.Б. Гавриков², Н.В. Пестрякова³*

¹ГБУЗ г. Москвы "Московская городская онкологическая больница №62 Департамента здравоохранения г. Москвы"

²ФГУ "Федеральный исследовательский центр "Институт прикладной математики им. М.В. Келдыша" РАН", г. Москва

³ФГУ "Федеральный исследовательский центр "Информатика и управление" РАН", г. Москва

bmgavrikov@gmail.com, mbgavrikov@yandex.ru, pestryakova@isa.ru

DOI: 10.20948/mm-2020-04-09

Работа посвящена статистическому методу классификации, базирующемуся на нелинейной регрессии. Изложены способы его реализации при решении задачи распознавания печатных и рукопечатных символов, а также впервые для оценивания состояния здоровья систем организма человека по параметрам периферической крови. Предложена оптимальная структура многочленов. Описаны свойства вероятностных оценок, генерируемых методом. Проанализирована структура множеств, используемых для его обучения.

Ключевые слова: распознавание, классификация, статистический метод, полиномиальная регрессия, печатный символ, рукопечатный символ, состояние здоровья человека, система организма, периферическая кровь.

STATISTICAL METHOD OF RECOGNIZING ON THE BASE OF NONLINEAR REGRESSION

B.M. Gavrikov¹, M.B. Gavrikov², N.V. Pestryakova³

¹Moscow State Budgetary Healthcare Institution "Moscow City Oncology Hospital No. 62 of the Moscow City Healthcare Department"

²Federal Research Center "Keldysh Institute of Applied Mathematics" of RAS, Moscow

³Federal Research Center "Computer Science and Control" of RAS, Moscow

The work is devoted to the statistical method of classification based on nonlinear regression. The ways of its implementation in solving the problem of recognition of printed and handwritten characters, as well as for the first time for assessing the health of the systems of the human body according to the parameters of peripheral blood. The optimal structure of polynomials is proposed. The properties of the probability estimates generated by the method are described. The structure of the sets used for its training is analyzed.

Key words: recognition, classification, statistical method, polynomial regression, printed symbol, handwritten symbol, human health condition, body system, peripheral blood.

1. Введение

Рассматриваются два приложения статистического подхода в распознавании, в основе которого лежит метод полиномиальной регрессии. Это распознаватель печатных и рукопечатных символов [1-5], а также классификатор для оценивания состояния здоровья систем организма (СО) человека по результатам анализа периферической крови (из пальца) [6,7].

Задача распознавания символов весьма актуальна для различных видов современных наукоемких технологий, использующих оптический ввод документов. Существующие классы распознающих систем (программы чтения текстов, системы потокового ввода структурированных документов, видеорегистраторы текстовых объектов) опираются на общие алгоритмы распознавания символов [8-10].

Расширение области применения систем распознавания ужесточает предъявляемые к ним требования. Характеристики качества включают не только традиционные точность и быстродействие, но также свойства оценок распознавания, на основании которых определяется надежность алгоритмов более высокого уровня и всей системы в целом. С появлением и распространением новых высокопроизводительных компьютерных технологий возрастает интерес к разработке методов, вычислительная структура которых удобна для распараллеливания.

Проблема оценивания состояния здоровья человека (СЗЧ) относится к понятию гомеостаза (в переводе с греческого *homoios* – подобный, тот же самый, *stasis* – состояние, подвижность), характеризующему относительное динамическое постоянство внутренней среды (крови, лимфы, тканевой жидкости) и устойчивости основных физиологических функций организма: кровообращения, дыхания, терморегулирования, обмена веществ.

При оценке гомеостаза и его динамики (гомеостатической активности) в клинической практике подход, основанный на сопоставлении таких параметров, как температура тела, артериальное давление, пульсации сердца, крайне неточен, поскольку измеряемые величины нестабильны, они могут быстро изменяться под влиянием внешних факторов. В качестве более объективного способа определения уровня гомеостаза и гомеостатической активности предлагается использовать не менее пяти показателей периферической крови [11].

Актуальность задачи предварительной диагностики, контролирования состояния систем организма пациента по результатам анализа крови с использованием данных, полученных в процессе медицинского обследования значительного количества людей, обуславливает необходимость разработки математических методов в доказательной медицине. Применение именно статистического подхода определяется тем, что он является наиболее адек-

ватным, учитывая наличие для каждой системы организма баз значительно объема, устанавливающих соответствие между классами здоровья и наборами показателей крови.

Оптимизация диагностики за счет предварительного выбора направления исследований имеет целью улучшение ее качества, ускорение и удешевление длительных и многоступенчатых медицинских процедур. Также эта методология может использоваться для контроля процесса лечения и предсказания его результатов [12-14].

Представляемый в данной работе статистический метод оказался эффективным при решении задачи распознавания (классификации) для объектов различной природы.

2. Метод распознавания

Используется байесовский вероятностный прецедентный подход для задачи классификации на K непересекающихся классов [15-17]. Ответом распознавателя считается K -мерный вектор вероятностей принадлежности объекта к каждому из классов, по которому можно найти номер класса. Тем самым, классификация превращается в специальный случай регрессии.

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

Пусть $p_k(\mathbf{v})$ – вероятность того, что входящий объект, задаваемый вектором $\mathbf{v} \in \mathbf{R}^N$, принадлежит k -му классу, где $1 \leq k \leq K$. Искомый класс будет иметь порядковый номер r , получивший максимальное значение вероятности (оценка распознавания):

$$p_r(\mathbf{v}) = \max_k \{p_k(\mathbf{v})\}, \quad 1 \leq k \leq K. \quad (1)$$

Следовательно, для решения задачи требуется вычислить вектор вероятностей $(p_1(\mathbf{v}), p_2(\mathbf{v}), \dots, p_K(\mathbf{v}))$. Он может быть найден на основе метода наименьших квадратов [18].

Приближенные значения $p_1(\mathbf{v}), \dots, p_K(\mathbf{v})$ в соответствии с методом полиномиальной регрессии представляются в виде конечных многочленов от координат $\mathbf{v} = (v_1, \dots, v_N)$ и определяются выбором базисных мономов:

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

Представим упорядоченные мономы (2) в виде вектора размерности L :

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

Тогда (2) можно записать в векторном виде:

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

где 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 $\mathbf{x}(\mathbf{v})$.

Вычисляем A приближенно в процессе обучения, используя базу данных: $[\mathbf{v}^{(1)}, \mathbf{y}^{(1)}], \dots, [\mathbf{v}^{(j)}, \mathbf{y}^{(j)}]$. Здесь $\mathbf{v}^{(j)}$ – вектор объекта, соответствующий k -му классу, $1 \leq k \leq K$; $\mathbf{y}^{(j)} = (0 \dots 1 \dots 0)$ – его базисный вектор, где 1 стоит на k -м месте, $1 \leq j \leq J$:

$$A \cong \left(\frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T \right)^{-1} \left(\frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{y}^{(j)})^T \right). \quad (4)$$

Проблема обращения заполненной матрицы большой размерности до сих пор не решена [19]. Чтобы обойти ее, предлагается [20] правую часть (4) вычислить посредством рекуррентной процедуры:

$$\begin{aligned} A_j &= A_{j-1} - \alpha_j G_j \mathbf{x}^{(j)} [A_{j-1}^T \mathbf{x}^{(j)} - \mathbf{y}^{(j)}]^T, \quad \alpha_j = 1/J, \\ G_j &= \frac{1}{1 - \alpha_j} \left[G_{j-1} - \alpha_j \frac{G_{j-1} \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T G_{j-1}}{1 + \alpha_j ((\mathbf{x}^{(j)})^T G_{j-1} \mathbf{x}^{(j)} - 1)} \right], \quad 1 \leq j \leq J. \end{aligned} \quad (5)$$

Здесь A_0 и G_0 заданы некоторым образом. Введение матрицы G_j размера $L \times L$ помогает избежать обращения матрицы в (4).

Доказана работоспособность упрощенной модификации процедуры (5) в обоих приложениях:

$$G_j \equiv D^{-1}, \quad D = \text{diag}(E\{x_1^2\}, E\{x_2^2\}, \dots, E\{x_L^2\}), \quad 1 \leq j \leq J, \quad (6)$$

где x_1, x_2, \dots, x_L – компоненты вектора $\mathbf{x}(\mathbf{v})$. Здесь E – общепринятое обозначение математического ожидания. В этом случае были получены приемлемые практические результаты.

Приведенная математическая постановка не является полной, поскольку не определен вектор базисных мономов $\mathbf{x}(\mathbf{v})$. Они получены авторами для двух рассматриваемых приложений и описаны ниже.

Первое приложение разработанного метода относится к распознаванию печатных и рукопечатных цифр ($K=10$). Алгоритм позволяет определить,

какому из K символов соответствует поступающий на вход распознавателя растр изображения. Нормализованный растр состоит из $N=16 \times 16$ серых пикселей (рис.1), перенумерованных в диапазоне $1 \leq i \leq N$. Вводим вектор $\mathbf{v} \in \mathbf{R}^N$, i -я компонента которого равна яркости i -го пикселя, $0 \leq v_i \leq 1$.

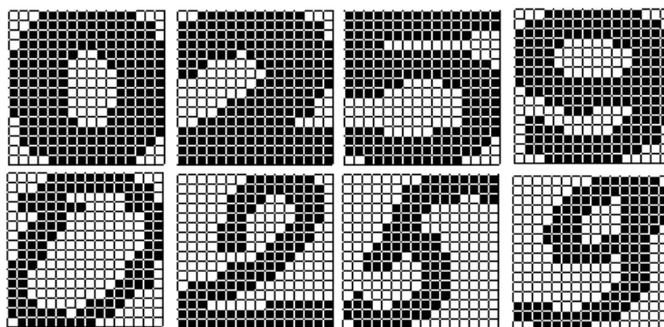


Рис.1. Образы 16×16 печатных и рукопечатных символов.

Здесь k -й символ считаем тождественным базисному вектору \mathbf{e}_k из \mathbf{R}^K ; $p_k(\mathbf{v})$ – вероятность того, что входящий растр изображает k -й символ, где $1 \leq k \leq K$. Искомый символ будет иметь порядковый номер r , соответствующий максимальному значению вероятности (1).

В предлагаемой математической модели распознавания изображений используются не только сами величины яркости, но и разности их значений в соседних пикселях, поскольку человеческий глаз фиксирует информативные признаки зрительных объектов по перепадам яркости. Размерность признакового пространства уменьшается ввиду отсутствия мономов с перекрестными произведениями яркости взаимно удаленных пикселей.

Для рукопечатных цифр использовали следующую модификацию $\mathbf{x}(\mathbf{v})$:

$$\begin{aligned} \mathbf{x}(\mathbf{v}) = & (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_z\}, \{(\delta v_i)_z^2\}, \\ & \{(\delta v_i)_r^4\}, \{(\delta v_i)_z^4\}, \{(\delta v_i)_r(\delta v_i)_z\}, \{(\delta v_i)_r^2(\delta v_i)_z^2\}, \{(\delta v_i)_r^4(\delta v_i)_z^4\}, \\ & \{(\delta v_i)_r((\delta v_i)_r)_1\}, \{(\delta v_i)_z((\delta v_i)_z)_1\}, \{(\delta v_i)_r((\delta v_i)_z)_1\}, \\ & \{(\delta v_i)_z((\delta v_i)_r)_1\}, \{(\delta v_i)_r((\delta v_i)_r)_d\}, \{(\delta v_i)_z((\delta v_i)_z)_d\}, \\ & \{(\delta v_i)_r((\delta v_i)_z)_d\}, \{(\delta v_i)_z((\delta v_i)_r)_d\}). \end{aligned} \quad (7)$$

Для печатных цифр достаточно элементов первой строки (7). Поскольку рукопечатные символы имеют меньшую толщину по сравнению с печатными, при вычислении компонент вектора $\mathbf{x}(\mathbf{v})$ имеющееся изображение искусственно увеличивается в размере на один пиксел в направлении, ортогональном соответствующему участку границы [4].

Выражения в фигурных скобках в (7) соответствуют цепочкам элементов, вычисляемым по всем пикселям растра (исключения указаны ниже). Через $(\delta v_i)_r$ и $(\delta v_i)_z$ обозначены конечные центральные разности величин v_i по ортогональным направлениям ориентации растра – нижние индексы r и z соответственно. Нижний индекс l (left) или d (down) означает, что величины относятся к пикселу слева или снизу от данного. Компоненты вектора $\mathbf{x}(\mathbf{v})$ без индекса l или d вычисляются для всех пикселей растра; с индексом l – кроме левых граничных; с индексом d – кроме нижних граничных пикселей. Вне растра считаем $v_i=0$.

Для печатных прямых и курсивных цифр обучение проводилось на базе в 95 тыс. элементов. При обучении и распознавании между прямыми и курсивными символами не делалось различий. Обучение рукопечатных цифр выполнено по базе в 175 тыс. элементов. Точность распознавания на обучающих множествах печатных цифр после третьей итерации обучения – 99.66% (с 1542 мономами), рукопечатных (сделано четыре прохода по обучающей последовательности) – 99.81% (полином из 22181 элемента).

Второе приложение относится к определению состояния здоровья человека (СЗЧ). По анализу крови производится деление на четыре класса ($K=4$), соответствующих стадиям поражения каждой из семи СО для мужчин и девяти – для женщин: здоровые, начальные, выраженные отклонения состояния здоровья, тяжелое заболевание. Здесь $\mathbf{v} \in \mathbf{R}^N$ – вектор, i -я компонента которого – отнормированная на отрезок $[0,1]$ (по базе рассматриваемой СО и пола) величина i -го показателя крови, причем $N=8$.

Приведем общепринятые обозначения и размерность восьми используемых показателей крови: $RBC [L^{-1}]$ – эритроциты, $HGB [g L^{-1}]$ – гемоглобин, $PLT [L^{-1}]$ – тромбоциты, $WBC [L^{-1}]$ – лейкоциты, $LIMPH [L^{-1}]$, [%] – лимфоциты, $GRAN [L^{-1}]$, [%] – гранулоциты. Последний показатель вычисляется следующим образом: $GRAN = NEUT + EOS + BASO$, где $NEUT [L^{-1}]$, [%] – нейтрофилы, $EOS [L^{-1}]$, [%] – эозинофилы, $BASO [L^{-1}]$, [%] – базофилы. Здесь L означает литр, g – грамм.

Теперь k -й элемент множества градаций СЗЧ тождественен базисному вектору \mathbf{e}_k из \mathbf{R}^K ; $p_k(\mathbf{v})$ – вероятность того, что входящий набор отнормированных показателей крови соответствует k -му элементу СЗЧ, где $1 \leq k \leq K$. Искомый элемент СЗЧ будет иметь порядковый номер r , получивший максимальное значение вероятности (1).

В классификаторах СО использовались следующие модификации вектора $\mathbf{x}(\mathbf{v})$:

$$\mathbf{x} = (1, \{v_i\}, \{v_i v_j\}, \{v_i v_j v_k\}, \{v_i v_j v_k v_l\}, \{v_i v_j v_k v_l v_m\}),$$

$$1 \leq i \leq 8, \quad i \leq j \leq 8, \quad j \leq k \leq 8, \quad k \leq l \leq 8, \quad l \leq m \leq 8. \quad (8)$$

Ниже указаны СО и пол, для которых при $x(v)$ данного вида получены наилучшие результаты, и более сложные модификации не использовались. Точность классификации приведена в случаях, когда она меньше 100%.

- ЦНС, органы чувствительности (мужчины); печень и желчевыводящие пути (мужчины): $L=61$ (мономы степенного вида до четвертого порядка, перекрестные произведения второго порядка).

- Органы дыхания (женщины); ЦНС, органы чувствительности (женщины); гинекологическая система (женщины): $L=69$ (степенные мономы до пятого порядка, перекрестные произведения второго порядка).

- Опорно-двигательный аппарат (мужчины): $L=77$ (мономы степенного вида до шестого порядка, перекрестные произведения второго порядка).

- Печень и желчевыводящие пути (женщины): $L=85$ (степенные мономы до седьмого порядка, перекрестные произведения второго порядка).

- Пищеварительная система (женщины); органы дыхания (мужчины); опорно-двигательный аппарат (женщины); эндокринная система (женщины); грудные железы (женщины): $L=165$ (мономы до третьего порядка, перекрестные произведения второго и третьего порядка).

- Пищеварительная система (мужчины – 98.2%); урологическая система (женщины – 99.2%, мужчины – 97.3%); эндокринная система (мужчины – 95%): $L=1287$ (мономы до пятого порядка, перекрестные произведения второго, третьего, четвертого и пятого порядка).

3. Качество распознавания

Анализируются следующие характеристики качества методов: точность распознавания на тестовых множествах и монотонность оценок.

Точностью распознавания по базе B называется величина

$$1 - \frac{\sum_{b \in B} (1 - \rho(C(b), P(b)))}{|B|},$$

где b – элементы тестовой базы образов B , $|B|$ – число образов в базе B , $C(b)$ – код символа, известный для каждого образа из тестовой базы, $P(b)$ – код символа, полученный в результате распознавания, $\rho(s, t)$ – расстояние между известным и распознанным кодами символа (в качестве такого расстояния берется функция сравнения, равная 1, если s и t неразличимы, и равная 0 в противоположном случае).

Используются целочисленные оценки $1, 2, \dots, 16$. После умножения оценки на 16 старый диапазон оценок $[0, 1]$, вычисленных по формуле (1), переходит в новый $[0, 16]$. Затем проводится дискретизация: $[0, 1] \rightarrow 1, (1, 2] \rightarrow 2, \dots, (15, 16] \rightarrow 16$.

Вводим следующие обозначения: $N(W)$ – количество распознанных образов с оценкой W ; $N_E(W)$ – число неправильно распознанных образов с оценкой W ; $\nu(W)=N(W)/|B|$ – частота распознавания с оценкой W ; $\nu_E(W) = N_E(W)/N(W)$ – частота ошибочного распознавания с оценкой W .

Монотонность оценок – свойство оценок характеризовать надежность распознавания. Исследуем совокупность частот $\{\nu_E(W_{\text{MIN}}), \dots, \nu_E(W_{\text{MAX}})\}$. Случаи с $N(W)=0$ не рассматриваем. Монотонность графика частот ошибок распознавания назовем монотонностью оценок метода. Интересны методы с монотонным убыванием в области высоких оценок.

Распределение оценок метода – совокупность частот $\{\nu(W_{\text{MIN}}), \dots, \nu(W_{\text{MAX}})\}$ появления оценок. Для метода с монотонными оценками высокие оценки характеризуются большей надежностью распознавания, поэтому желательно, чтобы метод чаще распознавал образы с большими оценками.

На приведенных далее рисунках гистограммой обозначено распределение $\nu_E(W)$, а графиком – распределение $\nu(W)$.

Для печатных прямых и курсивных цифр на тестовой последовательности из 12979 образов получена точность 99.56%. Ошибок со старшими оценками (равными 14, 15 и 16) нет, доля этих оценок составляет 0.69 (рис.2).

На тестовой последовательности из 8416 рукопечатных цифр точность равняется 99.73%. Ошибок со старшими оценками (равными 12, 13, 14, 15 и 16) нет, доля этих оценок составляет 0.83 (рис.3). В обоих случаях оценки монотонные.

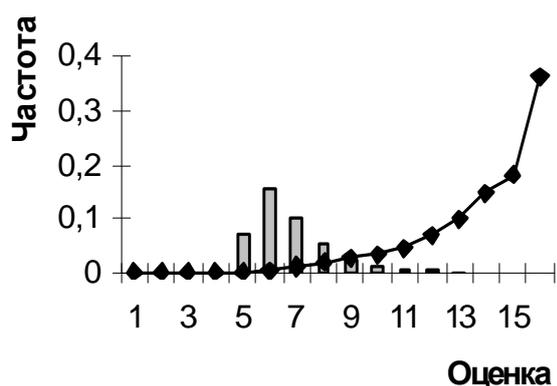


Рис.2. Печатные цифры.

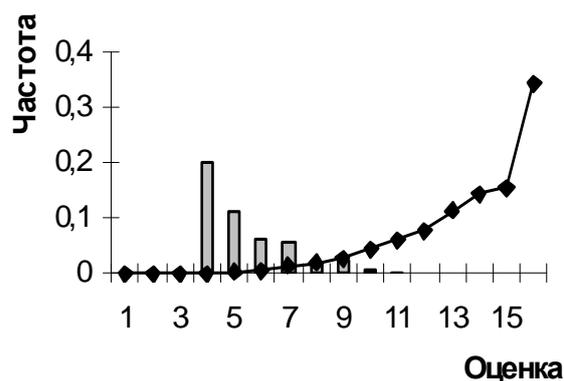


Рис.3. Рукопечатные цифры.

4. Свойства оценки

Выше оценка рассматривалась в качестве независимой величины, и относительно нее определялись число и частота правильного или ошибочного распознавания совокупности изображений всех символов. Теперь исследуем, как она сама зависит от распознаваемых объектов, а именно, растров изображений символов. Для более детального анализа будем использовать

целочисленные оценки 1, 2, ..., 255, полученные после линейного растяжения диапазона оценок $[0,1]$ в $[0,255]$ и дискретизации.

Изучим подбазы имеющихся классов по отдельности.

Для каждого из K символов рассмотрим среднестатистический растр и введем вектор $\mathbf{v}^{cp} \in \mathbf{R}^N$. Иногда такой вектор называют центром масс.

Яркость i -го пиксела центра масс k -го символа равна среднему арифметическому значений яркости i -х пикселов по всем J_k имеющимся в базе растрам изображений этого символа:

$$v_i^{k,cp} = \left(\sum_{j=1}^{J_k} v_i^{k,j} \right) / J_k, \quad (9)$$

где $\mathbf{v}^{k,j}$ – перенумерованные изображения k -го символа: $\{\mathbf{v}^{k,j} = (v_1^{k,j}, \dots, v_N^{k,j}), j=1, \dots, J_k\} = V^k$. Обучающее множество есть $V = \{V^1 \cup \dots \cup V^K\}$.

Для каждого из K символов введем среднестатистический полиномиальный вектор $\mathbf{x}^{cp} \in \mathbf{R}^L$. Иногда такой вектор называют центром моментов.

Значение в i -й компоненте центра моментов k -го символа равно среднему арифметическому i -х компонент полиномиальных векторов, построенных по всем J_k имеющимся в базе изображениям этого символа $\{\mathbf{v}^{k,j}, j=1, \dots, J_k\}$:

$$x_i^{k,cp} = \left(\sum_{j=1}^{J_k} x_i^{k,j} \right) / J_k, \quad (10)$$

где $\mathbf{x}^k = \mathbf{x}(\mathbf{v}^k)$.

По каждому из K символов в отдельности найдем минимальное и максимальное расстояние между центром масс (моментов) этого символа и векторами растров «своих» элементов (их полиномиальными векторами).

Для множества элементов k -го класса минимальное расстояние:

$$D_{-k_{\min}} = \min_{V^k} \{ \|\mathbf{v}^{k,cp} - \mathbf{u}^k\| \}, \quad D_{\mathbf{x}_{-k_{\min}}} = \min_{V^k} \{ \|\mathbf{x}^{k,cp} - \mathbf{x}(\mathbf{u}^k)\| \}; \quad (11)$$

максимальное расстояние:

$$D_{-k_{\max}} = \max_{V^k} \{ \|\mathbf{v}^{k,cp} - \mathbf{u}^k\| \}, \quad D_{\mathbf{x}_{-k_{\max}}} = \max_{V^k} \{ \|\mathbf{x}^{k,cp} - \mathbf{x}(\mathbf{u}^k)\| \}, \quad (12)$$

где $\mathbf{u}^k \in V^k$ – вектор растра, принадлежащий множеству элементов k -го класса V^k , $\mathbf{x}(\mathbf{u}^k)$ – его полиномиальный вектор; $\mathbf{v}^{k,cp}$ – центр масс этого класса, $\mathbf{x}^{k,cp}$ – центр моментов. Используется метрика L^1 .

Делим отрезок $[D_{-k_{\min}}, D_{-k_{\max}}]$ ($[D_{\mathbf{x}_{-k_{\min}}}, D_{\mathbf{x}_{-k_{\max}}}]$) (оси абсцисс на рис.4а,б) на 20 равных частей. Вычисляем среднюю оценку распознавания изображений с растрами (полиномиальными векторами), попавшими в каждый такой участок (оси ординат на рис.4а, б). Для рукопечатной «1» она сначала убывает, а затем увеличивается (рис.4а), а для остальных рукопечатных

цифр, а также всех печатных, убывает монотонно (рис.4б). Имеющиеся шумы можно убрать укрупнением отрезков разбиения и уменьшением их числа. После разделения базы рукопечатной «1» по типу написания [3] на каждом из этих двух подмножеств получили результаты, аналогичные рис.4б.

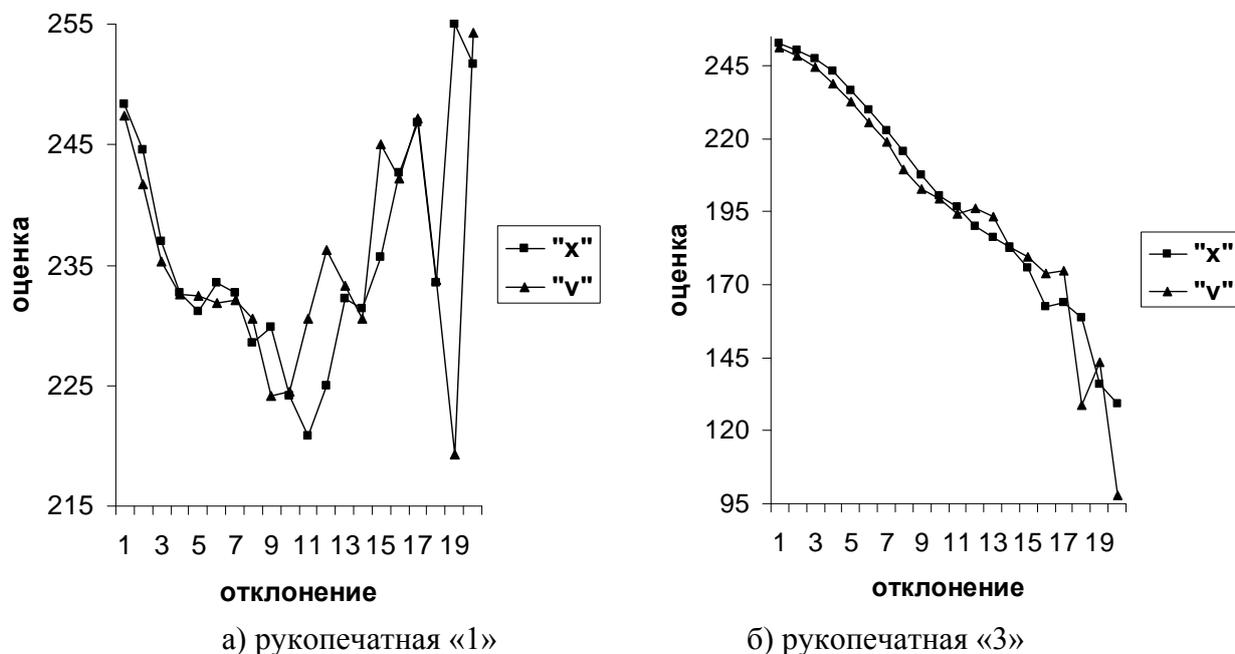


Рис.4. Средняя оценка при удалении от центра масс (моментов).

Для рукопечатного символа «8» показано, как монотонное убывание средней оценки распознавания при удалении от центра моментов соотносится с распределениями числа верно распознанных изображений для различных оценок. Рассмотрены следующие диапазоны оценок: [255, 250), [250, 240), [240, 230), [230, 220), [220, 210), [210, 200), [200, 190), [190, 180), [180, 170), [170, 160), [160, 150), [150, 140), [140, 130), [130, 120). Изображений с более низкими оценками мало. На рис.5а,б приведены распределения числа образов с оценками внутри диапазонов с более высокими оценками (для низких аналогично). На рис.5а также показано распределение для всего спектра оценок [0, 255].

На каждой из частей отрезка $[D_x_{k_{\min}}, D_x_{k_{\max}}]$ монотонно убывающая средняя оценка получается суммированием оценок 1, 2, 3, ..., 254, 255 с весами, определяемыми средней (по этой части) вероятностью оценки.

5. Анализ обучающих множеств

По каждому из K классов аналогично (11),(12) найдем минимальное и максимальное расстояние между центром масс (моментов) и векторами «чужих» растров (полиномиальными векторами).

Минимальное расстояние

$$D_{-kz_{\min}} = \min_{V^{-k}} \{ \| \mathbf{v}^{k,cp} - \mathbf{u}^{-k} \| \}, \quad \mathbf{u}^{-k} \in V^{-k},$$

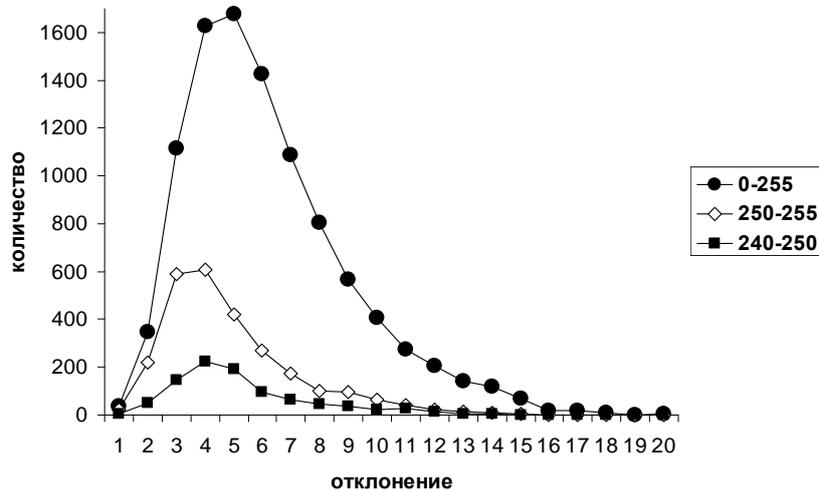
$$D_{\mathbf{x}_{-kz_{\min}}} = \min_{V^{-k}} \{ \| \mathbf{x}^{k,cp} - \mathbf{x}(\mathbf{u}^{-k}) \| \}, \quad \mathbf{u}^{-k} \in V^{-k}.$$
(13)

Максимальное расстояние

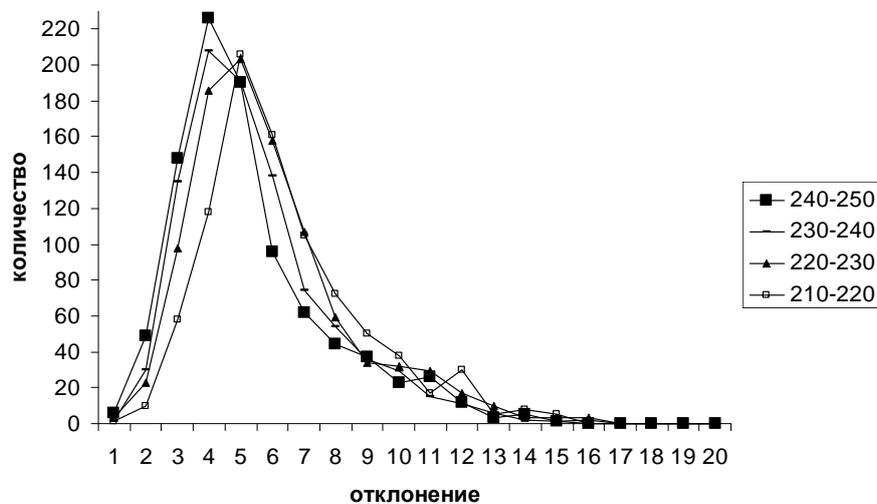
$$D_{-kz_{\max}} = \max_{V^{-k}} \{ \| \mathbf{v}^{k,cp} - \mathbf{u}^{-k} \| \}, \quad \mathbf{u}^{-k} \in V^{-k},$$

$$D_{\mathbf{x}_{-kz_{\max}}} = \max_{V^{-k}} \{ \| \mathbf{x}^{k,cp} - \mathbf{x}(\mathbf{u}^{-k}) \| \}, \quad \mathbf{u}^{-k} \in V^{-k},$$
(14)

где \mathbf{u}^{-k} – вектор, принадлежащий множеству чужих элементов V^{-k} классов, отличных от k -го: $V^{-k} = V \setminus V^k$, $\mathbf{v}^{k,cp}$ ($\mathbf{x}^{k,cp}$) – центр масс (моментов) k -го класса.



а) диапазоны [0, 255], [255, 250), [250, 240)



б) диапазоны [250, 240), [240, 230), [230, 220), [220, 210)

Рис.5. Число распознанных образов «8» для различных оценок.

Получено [3], что минимальное расстояние от центра масс (моментов) каждой из рукопечатных цифр до векторов растров (полиномиальных векторов) своих изображений меньше аналогичной величины, вычисленной для чужих элементов:

$$D_{-k_{\min}} < D_{-kz_{\min}}, \quad D_{x-k_{\min}} < D_{x-kz_{\min}}. \quad (15)$$

Для печатных цифр выполняется более сильное соотношение:

$$\max_k(D_{-k_{\min}}) < \min_k(D_{-kz_{\min}}), \quad \max_k(D_{x-k_{\min}}) < \min_k(D_{x-kz_{\min}}). \quad (16)$$

Для максимальных расстояний (12),(14) по печатным и рукопечатным цифрам выполняется соотношение, аналогичное (15):

$$(D_{-k_{\max}}) < (D_{-kz_{\max}}), \quad (D_{x-k_{\max}}) < (D_{x-kz_{\max}}). \quad (17)$$

Более того, выполняется соотношение, аналогичное (16):

$$\max_k(D_{-k_{\max}}) < \min_k(D_{-kz_{\max}}), \quad \max_k(D_{x-k_{\max}}) < \min_k(D_{x-kz_{\max}}). \quad (18)$$

Для рукопечатной «1» с обучающим множеством из двух подмножеств, различающихся по типу написания, (17) не выполняется [3].

При анализе баз СО понятие центра масс k -го класса аналогично определенному для символов (9); значения яркости в пикселях заменяются на отнормированные параметры крови. Получено [7], что минимальное (максимальное) расстояние между центром масс и своими элементами класса «1» меньше, чем соответствующее расстояние для чужих элементов (15), (17); более того, эти соотношения выполняются, если справа стоят выражения из правой части (16),(18). Однако для классов «2», «3», «4» при рассмотрении ряда СО неравенства (15), (17) нарушаются.

Заметим, что структура множеств изображений символов и класса «1» СО схожа: для каждого i ($1 \leq i \leq N$) выполнено неравенство: $0 \leq v_i \leq 1$ для пикселей, или, аналогично, $0 \leq \text{const}1_i \leq v_i \leq \text{const}2_i \leq 1$ для параметров крови, где $\text{const}1_i$ и $\text{const}2_i$ определяются диапазоном, присущим здоровым людям и коэффициентами нормировки i -го параметра для заданного класса, СО, пола. Элементы класса «1» распознаются без ошибок уже при использовании полиномов более простой структуры с низкими степенями.

6. Заключение

Исследователям и разработчикам в высокотехнологичных отраслях, ориентированных на использование распознавания или классификации, приходится делать выбор в пользу того или другого метода. Данный подход был рассмотрен с той целью, чтобы решение принималось более осознанно.

Показана высокая эффективность метода при распознавании объектов различной природы. Изложено, как следует учитывать особенности элементов предметной области, если они являются изображениями. Описан способ уменьшения размерности пространства полиномиальных векторов, что способствует ускорению обучения.

Метод является точным. Он генерирует монотонные (надежные) оценки. Это особенно важно в случае, когда невозможно использование иных способов проверки, например, со словарем; следовательно, в частности, актуально для распознавания цифр. Он удобен для распараллеливания. Метод устойчив к искажениям: хорошо распознает зашумленные (загрязненные, частично разрушенные) объекты и пригоден для решения реальных задач.

СПИСОК ЛИТЕРАТУРЫ

1. М.Б. Гавриков, Н.В. Пестрякова. Метод полиномиальной регрессии в задачах распознавания печатных и рукопечатных символов. – М.: ИПМ им. М.В.Келдыша, 2004, препринт № 22, 12 с.
M.B. Gavrikov, N.V. Pestriakova. Metod polinomialnoi regressii v zadachakh raspoznavaniia pchatnykh i rukopchatnykh simvolov. – M.: IPM im. M.V.Keldysha, 2004, preprint № 22, 12 s.
2. М.Б. Гавриков, А.В. Мисюрев, Н.В. Пестрякова, О.А. Славин. Об одном методе распознавания символов, основанном на полиномиальной регрессии // Автоматика и Телемеханика, 2006, №2, с.119-134;
англ.пер.: *M.B. Gavrikov, A.V. Misyurev, N.V. Pestriakova, O.A. Slavin. On a Pattern Recognition Method Based on Polynomial Regression // Automation and Remote Control, 2006, v.67, № 2, p.278-292.*
3. Н.В. Пестрякова. Метод распознавания символов, основанный на полиномиальной регрессии. – М.: УРСС, 2011, 141 с.
N.V. Pestriakova. Metod raspoznavaniia simvolov, osnovannyi na polinomialnoi regressii. – M.: URSS, 2011, 141 s.
4. М.Б. Гавриков, Н.В. Пестрякова, Б.М. Гавриков. Об эффективных приложениях полиномиальной регрессии к задаче распознавания символов // Труды ИСА РАН, 2014, т.64, вып.4, с.89-96.
M.B. Gavrikov, N.V. Pestriakova, B.M. Gavrikov. Ob effektivnykh prilozheniakh polinomialnoi regressii k zadache raspoznavaniia simvolov // Trudy ISA RAN, 2014, t.64, v.4, s.89-96.
5. Б.М. Гавриков, М.Б. Гавриков, Н.В. Пестрякова. Статистический анализ характеристик метода распознавания на обучающем множестве // Информационные технологии и вычислительные системы, 2015, №2, с.50-67.
B.M. Gavrikov, M.B. Gavrikov, N.V. Pestriakova. Statisticheskii analiz kharakteristik metoda raspoznavaniia na obuchayushchem mnozhestve // Informatsionnye tekhnologii i vychislitelnye sistemy, 2015, №2, s.50-67.
6. Б.М. Гавриков, М.Б. Гавриков, И.М. Лебеденко, Н.В. Пестрякова, Р.В. Ставицкий. Метод оценивания состояния здоровья человека. – М.: ИПМ им. М.В.Келдыша, 2017, препринт №8, 18 с.

- V.M. Gavrikov, M.B. Gavrikov, I.M. Lebedenko, N.V. Pestriakova, R.V. Stavitskii.* Metod otsenivaniia sostoyaniia zdorovia cheloveka. – М.: ИПМ им. М.В.Келдыша, 2017, preprint № 8, 18 s.
7. *Б.М. Гавриков, М.Б. Гавриков, Н.В. Пестрякова, Р.В. Ставицкий.* Структура базы обучения статистического классификатора состояний систем организма человека. – М.: ИПМ им. М.В.Келдыша, 2018, препринт №255, 40с.
V.M. Gavrikov, M.B. Gavrikov, N.V. Pestriakova, R.V. Stavitskii. Struktura bazy obucheniia statisticheskogo klassifikatora sostoianii sistem organizma cheloveka. – М.: ИПМ им. М.В. Келдыша, 2018, preprint №255, 40s.
8. *В.В. Арлазаров, А.Е. Жуковский, В.Е. Кривцов, Д.П. Николаев, Д.В. Полевой.* Анализ особенностей использования стационарных и мобильных малоразмерных цифровых видео камер для распознавания документов // Информ. технол. и выч. сист., 2014, №3, с.71-81.
V.V. Arlazarov, A.E. Zhukovskii, V.E. Krivtsov, D.P. Nikolaev, D.V. Poleviiy. Analiz osobennostei ispolzovaniia statsionarnykh i mobilnykh malorazmernykh tsifrovyykh video kamer dlia raspoznavaniia dokumentov // Inform. tekhn. i vych. sist., 2014, №3, s.71-81.
9. *О.О. Петрова, К.Б. Булатов.* Методы пост-обработки результатов распознавания машиночитаемой зоны документов // Труды Института системного анализа РАН, Специальный выпуск, 2018, с.43-50.
O.O. Petrova, K.B. Bulatov. Metody post-obrabotki rezultatov raspoznavaniia mashinochitaemoi zony dokumentov // Trudy Instituta sistemnogo analiza RAN, Spetsialnyi vypusk, 2018, s.43-50.
10. *В.В. Арлазаров, К.Б. Булатов, А.В. Усков.* Модель системы распознавания объектов в видеопотоке мобильного устройства // Труды Института системного анализа РАН, Специальный выпуск, 2018, с.73-82.
V.V. Arlazarov, K.B. Bulatov, A.V. Uskov. Model sistemy raspoznavaniia oiektov v videopotoke mobilnogo ustroistva // Trudy Instituta sistemnogo analiza RAN, Spetsialnyi vypusk, 2018, s.73-82.
11. *Р.В. Ставицкий, Л.А. Лебедев, А.Л. Лебедев, А.Ю. Смыслов.* Количественная оценка гомеостатической активности здоровых и больных людей. – М.: ГАРТ, 2013, 131с.
R.V. Stavitskii, L.A. Lebedev, A.L. Lebedev, A.IU. Smyslov. Kolichestvennaia otsenka go-meostaticheskoi aktivnosti zdorovykh i bolnykh liudei – М.: GART, 2013, 131 s.
12. *Н.Ю. Добровольская, Л.А. Лебедев, А.Л. Лебедев, Ю.Б. Новожиллов, Р.В. Ставицкий.* Химио-лучевая терапия рака шейки матки. Методика оценки состояния организма и его систем // Радиология-практика, 2011, №3, с.53-63.
N.IU. Dobrovolskaia, L.A. Lebedev, A.L. Lebedev, IU.B. Novozhilov, R.V. Stavitskii. Khimio-luchevaia terapiia raka sheiki matki. Metodika otsenki sostoianiia organizma i ego sistem // Radiologiiia-praktika, 2011, №3, s.53-63.
13. *И.М. Лебедеенко, Т.З. Чернявская, Р.В. Ставицкий, О.Н. Плаутин.* Технический контроль состояния организма и его систем в процессе химио-лучевой терапии и трансплантации костного мозга при острых лейкозах // Мед. техника, 2014, №5, с.32-36.
I.M.Lebedenko, T.Z.Cherniavskaya, R.V.Stavitskii, O.N.Plautin. Tekhnicheskii control sostoyaniia organizma i ego sistem v protsesse khimio-luchevoi terapii i transplantatsii kostnogo mozga pri ostrykh leikozakh // Meditsinskaia tekhnika, 2014, №5, s.32-36.
14. *Ю.А. Цыбульская, Р.В. Ставицкий, И.М. Лебедеенко, С.В. Смердин, И.В. Шутихина,*

- Л.С. Коков, О.В. Батурич.* Количественный подход к диагностике поражения костно-суставной системы при туберкулезном спондилите // Медицинский вестник Северного Кавказа, 2015, №3, с.212-217.
- I.U.A. Tsybul'skaia, R.V. Stavitskii, I.M. Lebedenko, S.V. Smerdin, I.V. Shutikhina, L.S. Kokov, O.V. Baturin.* Kolichestvennyi podkhod k diagnostike porazheniia kostno-sustavnoy sistemy pri tuberkuleznom spondilite // Meditsinskii vestnik Severnogo Kavkaza, 2015, №3, s.212-217.
15. *А.Б. Мерков.* Распознавание образов. Введение в методы статистического обучения. - М.: URSS, 2019, изд.2, испр., 256с.
- A.B. Merkov.* Raspoznavaniie obrazov. Vvedeniie v metody statisticheskogo obucheniia. - М.: URSS, 2019, izd.2, ispr., 256s.
16. *А.Б. Мерков.* Распознавание образов. Построение и обучение вероятностных моделей. - М.: URSS, 2014, 240с.
- A.B. Merkov.* Raspoznavaniie obrazov. Postroyeniie i obuchenie veroyatnostnykh modelei. - М.: URSS. 2014, 240s.
17. *К.В. Воронцов.* Математические методы обучения по прецедентам // Курс лекций. - Долгопрудный: МФТИ, 2012. [http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_\(курс_лекций,_К.В.Воронцов\)](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_(курс_лекций,_К.В.Воронцов)).
- K.V. Vorontsov.* Matematicheskie metody obucheniia po pretsedentam // Kurs lektsii. - Dolgoprudnyi, MFTI, 2012.
18. *Ю.В. Линник.* Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. - М.: Физматлит, 1958, 336с.
- I.U.V. Linnik.* Metod naimenshikh kvadratov i osnovy matematiko-statisticheskoi teorii obrabotki nabliudeni. - М.: Fizmatlit, 1958, 336s.
19. *М.Б. Гавриков, О.В. Локуцкий.* Начала численного анализа. - М.: Янус, 1995, 581с.
- M.B.Gavrikov, O.V.Lokutsievskii.* Nachala chislennoho analiza. - М.: Ianus, 1995, 581s.
20. *J. Schürmann.* Pattern Classification. - New York: John Wiley&Sons, Inc., 1996, 374p.

Поступила в редакцию 21.10.2019

После доработки 21.10.2019

Принята к публикации 25.11.2019