

Статистическое распознавание образов на основе посегментного анализа однородности*

А. В. Савченко

avsavchenko@hse.ru

НИУ Высшая школа экономики, Нижний Новгород, Россия

Исследуется проблема малых выборок в задаче статистического распознавания образов на основе методов ближайших соседей, точность которых во многом определяется выбранной мерой близости, при этом их реализация в режиме реального времени может оказаться невозможной уже при наличии тысяч классов. Для преодоления указанных проблем предложен новый подход к разработке классификаторов с посегментным анализом однородности и быстрой последовательной иерархической обработкой на основе вероятностной модели кусочно-однородного объекта. Экспериментальные исследования в задаче распознавания лиц продемонстрировали повышение точности на 1%–10% по сравнению с традиционными методами (SVM, SIFT, LBP, собственные лица). Вычислительная эффективность оказалась в 2–3 раза выше по сравнению с известным методом Pyramid HOG (Histograms of Oriented Gradients). Показано, что описанная методология посегментного анализа однородности характеризуется высокой точностью и приемлемой производительностью для случая малых выборок и большого числа классов.

Ключевые слова: статистическое распознавание образов; иерархическая классификация; методы приближенного поиска ближайшего соседа; классификация изображений; гистограммы ориентированных градиентов

Statistical pattern recognition based on segment homogeneity testing*

A. V. Savchenko

National Research University Higher School of Economics, N. Novgorod, Russian Federation

Background: This paper is focused on a small-sample size problem in statistical recognition of audiovisual objects with the nearest neighbor method. Its accuracy depends on the applied similarity measure. Moreover, the computing efficiency is insufficient if thousand of classes are available.

Methods: The author introduces an approach to design classifiers of audiovisual objects by testing of segment homogeneity based on the probabilistic model of composite object represented by a sequence of independent identically distributed segments. The asymptotic properties of this approach allow to implement sequential hierarchical classification with approximate search of the nearest neighbor to speed up the decision process.

Results: Experimental study in constrained face recognition with HOG features shows that the proposed approach allows to increase accuracy in 1%–10% in comparison with conventional image recognition techniques (k-NN, SVM, SIFT, histogram of local binary patterns, eigenfaces). Moreover, it is 2–3 times faster than the pyramid HOG (Histograms of Oriented Gradients) hierarchical classifier.

*Исследование выполнено в Национальном исследовательском университете «Высшая школа экономики» за счет средств гранта Российского научного фонда (проект № 14-41-00039)

Concluding Remarks: Described methodology with segment homogeneity testing allows to achieve high accuracy with sufficient performance in case of small-sample-size and medium-sized number of classes.

Keywords: *statistical pattern recognition; hierarchical classification; approximate nearest neighbor methods; image classification; Histograms of Oriented Gradients (HOG)*

1 Введение

Традиционный подход к организации многих систем обработки аудиовизуальной информации состоит в последовательном соединении модулей извлечения признаков, классификации и управления [1, 2]. Попытки построить модели и методы, адекватно описывающие решение человеком таких плохо формализованных задач до настоящего момента не привели к значимым практическим результатам [3]. Поэтому прикладные исследования в области синтеза таких систем в основном сосредоточились на разработке методов принятия решений в модуле классификации, в котором входному объекту X ставится в соответствие один из $C > 1$ заранее точно не определенных классов. Предполагается, что для обучения системы доступна база данных, содержащая $R \geq C$ эталонных объектов (прецедентов) $\{X_r\}, r \in \{1, \dots, R\}$, где класс каждого r -го эталона $c(r) \in \{1, \dots, C\}$ считается известным [1]. В настоящее время считается, что, несмотря на существенные отличия в процедурах извлечения признаков изображений и речевых сигналов, для решения все более сложных задач промышленные системы классификации следует рассматривать в более широком смысле и учитывать доступную структурную и семантическую информацию [4]. На начальном уровне обработки входной объект X и все эталоны X_r разбиваются на несколько сегментов, каждый из которых описывается множеством одинаково распределенных значений признаков, а в новом модуле структурного распознавания осуществляется посегментный анализ результатов классификации сегментов на более низких уровнях. Такой подход является достаточно универсальным в связи с наличием хорошо изученных алгоритмов сегментации как аудио, так и визуальной информации. К сожалению, применение известных способов реализации такого посегментного анализа, основанных на скрытых Марковских моделях [5], возможно лишь в том случае, если для обучения доступна репрезентативная база данных большого объема (сотни эталонов для каждого класса, $R \gg C$). Такое ограничение оказывается слишком жестким для многих промышленных систем, в которых проявляется проблема малых выборок ($R \approx C$): число эталонов для каждого класса недостаточно для обучения сложного классификатора [6]. Проблема усиливается в характерном именно для задач обработки аудиовизуальной информации случае наличия помех как на этапе формирования базы данных эталонов, так и при наблюдении входного объекта X . В случае малых выборок для решения обычно применяются методы ближайшего соседа, точность которых определяется используемой мерой близости [1]. Существующая теория не дает строгого ответа на вопрос о выборе оптимальной меры близости, поэтому на практике исследователи полагаются на свой опыт, что, разумеется, далеко не всегда приводит к наилучшему результату. Более того, в связи с большой размерностью пространства признаков для описания аудиовизуальной информации, реализация такого подхода, основанного на полном переборе базы данных эталонов, в режиме (мягкого) реального времени оказывается во многих промышленных системах невозможной уже при наличии сотни классов [7]. Кажется, что для ускорения процедуры поиска ближайших соседей можно использовать известные приближенные алгоритмы, например рандомизированные kd-деревья [8] или perm-sort [9]. К сожалению,

такие методы разрабатываются специально для автоматизированных систем поиска объектов по содержанию из сверхбольших баз данных (десятки и сотни тысяч классов), поэтому их применение не приводит к существенному снижению времени классификации, если число классов не превышает несколько тысяч единиц.

Со всех перечисленных точек зрения повышенный интерес представляет описанный в работе [10] подход к решению задачи групповой классификации [11], в котором за счет объединения входной и эталонной выборок повышается точность оценки распределения каждого класса для случая малых размерностей эталонных выборок и/или при наличии в них шума. Применению такого подхода при анализе однородности сегментов входного и эталонного объектов для преодоления указанной проблемы недостаточной точности и вычислительной эффективности современных промышленных систем обработки и классификации аудиовизуальной информации при наличии в базе данных большого числа классов (тысячи альтернатив) и малого количества эталонов для каждого класса ($R \approx C$) и посвящена настоящая статья.

2 Посегментный анализ однородности в задаче классификации составных объектов

Прежде всего, сформулируем следующий целевой критерий (ЦК) оценки эффективности методов классификации:

$$\bar{\alpha}_\eta \rightarrow \min; \quad \text{s.t. } t_{\text{avg}} < t_0, \quad (1)$$

согласно которому наилучший классификатор характеризуется минимальной *средней вероятностью ошибки* $\bar{\alpha}_\eta$, где для контроля устойчивости к тестовым данным добавляются помехи η с дисперсией $\sigma_\eta^2 \geq 0$. При этом накладывается ограничение на максимально допустимое среднее время принятия решения t_{avg} , которое не может превышать фиксированный порог t_0 . Проверка такого условия может производиться как с помощью инструментария теории алгоритмов (О-оценок), так и непосредственным сопоставлением среднего времени классификации для различных методов в рамках натурных испытаний. Таким образом, ЦК (1) соответствует Парето-оптимальному решению задачи многокритериальной оптимизации с одновременной минимизацией средней вероятности ошибки и среднего времени классификации. Выражение (1) фактически устанавливает частичный порядок на множестве всех методов классификации. Из двух классификаторов, которые удовлетворяют требованию к среднему времени отклика $t_{\text{avg}} < t_0$, наилучшим является тот, который характеризуется наименьшей средней вероятностью перепутывания $\bar{\alpha}_\eta$. Если один из двух классификаторов в силу своей вычислительной сложности не может быть реализован в режиме «мягкого» реального времени ($t_{\text{avg}} \geq t_0$), то второй классификатор будет являться более предпочтительным, даже если его точность оказывается ниже точности первого метода. Наконец, если два метода характеризуются неприемлемым временем принятия решения, то с точки зрения практической реализации оба таких классификатора считаются неудовлетворительными.

Для решения задачи распознавания входной объект X с помощью известной процедуры сегментации (например, разбиения на фреймы фиксированной размерности или наращивания областей) [5, 12] представляется в виде последовательности из K сегментов $X(k)$, каждому из которых ставится в соответствие совокупность из $n(k)$ векторов значений признаков $\mathbf{x}_j(k), j \in \{1, \dots, n(k)\}$ фиксированной размерности M . Аналогично эталоны X_r разбиваются на K_r частей и k -й сегмент r -го эталона определяется как последовательность векторов значений признаков $\mathbf{x}_j^{(r)}(k), j \in \{1, \dots, n_r(k)\}$, размерности M в k -м

сегменте r -го эталона, а $n_r(k)$ — число признаков. В рамках модели кусочно-однородного объекта [13] предполагается, что векторы $\mathbf{x}_j(k)$ и $\mathbf{x}_j^{(r)}(k)$ являются реализациями случайных векторов, сегменты $X(k)$ — простые выборки (независимых одинаково распределенных) векторов $\mathbf{x}_j(k)$, а эталонные сегменты $X_r(k)$ — простые выборки векторов $\mathbf{x}_j^{(r)}(k)$. Чтобы учесть недостаточную точность детектирования границ, все сегменты $X(k)$ рассматриваются в пределах, близких к k номерам k_r сегментов r -го эталона из множества $N_r(k)$ (окрестность «соседей», определяемая спецификой предметной области). Тогда задача сводится к проверке гипотез $W_r, r \in \{1, \dots, R\}$, о неизвестном законе распределения значений признаков сегментов объекта X .

Рассмотрим случай *дискретных* признаков с конечным множеством из $N > 1$ значений $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$. Тогда сегменты $X(k)$ и $X_r(k)$ полностью описываются своими гистограммами $\{w_i(k)\}$ и $\{\theta_i^{(r)}(k)\}, i \in \{1, \dots, N\}$ соответственно. В таком случае можно показать [14], что для применяемой в известной вероятностной нейронной сети (ВНС) [15] непараметрической оценки неизвестных распределений сегментов с помощью ядра Розенблатта–Парзена $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ решение задачи с помощью байесовского критерия минимума среднего риска сводится к следующему критерию:

$$\nu = \arg \min_{r \in \{1, \dots, R\}} \left(\frac{1}{Kn} \sum_{k=1}^K \min_{k_r \in N_r(k)} \sum_{i=1}^N w_i(k) \ln \frac{w_{K;i}(k)}{\theta_{K;i}^{(r)}(k_r)} - \ln p_r \right). \quad (2)$$

Здесь $\theta_{K;i}^{(r)}(k_r) = \sum_{j=1}^N K_{ij} \theta_j^{(r)}(k_r)$; $w_{K;i}(k) = \sum_{j=1}^N K_{ij} w_j^{(r)}(k)$ — свертки с ядром гистограмм r -й эталонной $X_r(k_r)$ и входной выборки $X(k)$; p_r — априорная вероятность появления класса, определяемого r -м эталоном; $n = \sum_{k=1}^K n(k)/K$ — средняя длительность сегмента.

К сожалению, как известно [10], при подстановке вместо неизвестных распределений их оценок классификатор (2) становится неоптимальным. Для повышения точности в ЦК (1) автором был предложен новый подход к построению классификаторов на основе посегментного анализа однородности (ПАО) признаков входного и эталонного объектов [13], в котором задача сводится к проверке сложных гипотез об однородности сегментов $X(k)$ и $X_r(k_r)$ с оценкой распределения класса по объединенной выборке $\{X(k), X_r(k_r)\}$. Тогда следующий критерий

$$\nu = \arg \min_{r \in \{1, \dots, R\}} (\rho_{PNNH}(X, X_r) - \ln p_r) \quad (3)$$

является при $n(k) \rightarrow \infty$ асимптотически минимаксным для проверки однородности объектов X и r -го эталона. Здесь

$$\rho_{PNNH}(X, X_r) = \frac{1}{Kn} \sum_{k=1}^K \min_{k_r \in N_r(k)} \sum_{i=1}^N \left(w_i(k) \ln \frac{w_{K;i}(k)}{\tilde{\theta}_{\Sigma;i}^{(r)}(k; k_r)} + \frac{n_r(k_r)}{n(k)} \theta_i^{(r)}(k) \ln \frac{\theta_{K;i}^{(r)}(k)}{\tilde{\theta}_{\Sigma;i}^{(r)}(k; k_r)} \right) \quad (4)$$

есть выход синтезированной в работе [14] ВНС с ПАО, а $\tilde{\theta}_{\Sigma;i}^{(r)}(k; k_r) = (n(k)w_{K;i}(k) + n_r(k_r)\theta_{K;i}^{(r)}(k_r))/(n(k) + n_r(k_r))$.

Если ядром K_{ij} является дискретная дельта-функция $\delta(i - j)$, то $\theta_{K;i}^{(r)}(k_r) = \theta_i^{(r)}(k_r)$, $w_{K;i}(k) = w_i(k)$, тогда выражение (2) будет эквивалентно известному принципу минимума информационного рассогласования Кульбака–Лейблера. А если дополнительно предположить, что размеры всех сегментов одинаковы, то рассогласование (4) будет эквивалентно дивергенции Йенсена–Шеннона и, при использовании известной аппроксимации логарифма $\ln(a/b) \approx (a^2 - b^2)/(2ab)$, расстоянию χ^2 . Как видно, применение подхода на основе

посегментного анализа однородности позволяет получить не только новые, но и как частный случай хорошо зарекомендовавшие себя на практике критерии.

3 Способы повышения вычислительной эффективности алгоритма классификации с посегментным анализом однородности

Вычислительная сложность критерия (3) оценивается как $O\left(N \sum_{r=1}^R \sum_{k=1}^K |N_r(k)|\right)$, где $|N_r(k)|$ — число элементов в множестве $N_r(k)$. Как и в любой реализации метода ближайшего соседа, необходимость полного перебора всех эталонов из базы данных может сделать его физически нереализуемым для больших R . Поэтому в настоящем разделе рассмотрены способы повышения вычислительной эффективности критерия (2) на основе посегментного анализа однородности, если они не удовлетворяют ЦК (1) при малом значении порога t_0 . Прежде всего, опишем последовательный способ принятия решений, в котором проводится анализ более детализированного уровня описания аудиовизуальной информации только при получении недостаточно надежных решений на предыдущих, более простых уровнях. На первом уровне такой иерархической системы анализируются наиболее грубые приближения входного и эталонных объектов, например с малым числом однородных сегментов K . Запишем критерий (3) в упрощенной форме [7]:

$$W_\nu : \rho_{\text{PNNH}}(X, X_\nu) < \rho_0. \quad (5)$$

Для заданной вероятности p_{FRR} ложного пропуска объекта порог ρ_0 устанавливается равным p_{FRR} -квантилю множества расстояний между разноименными объектами. Если условие (5) не было выполнено ни для одного из эталонов, то для выделения входных объектов, не принадлежащих ни одному из эталонных классов, применяется критерий сравнения минимального рассогласования с фиксированным порогом:

$$\rho_{\text{PNNH}}(X, X_\nu) > \rho_1. \quad (6)$$

Здесь порог ρ_1 для заданной вероятности p_{FAR} ложного срабатывания устанавливается равным $(1 - p_{\text{FAR}})$ -квантилю множества рассогласований между одноименными объектами из одного класса. Если условие (6) не выполняется, для проверки надежности решения (3) воспользуемся оптимальным в байесовском смысле правилом Чоу [16]:

$$P(W_\nu | X) > p_0, \quad (7)$$

где значение порога p_0 можно выбрать, зафиксировав вероятность ошибки ложного срабатывания. В работе [13] показано, что для предложенного метода с ПАО (3), апостериорная вероятность $P(W_\nu | X)$ гипотезы W_ν оценивается как

$$\hat{P}(W_\nu | X) = \frac{p_\nu \exp(-nK \rho_{\text{PNNH}}(X, X_\nu))}{\sum_{r=1}^R p_r \exp(-nK \rho_{\text{PNNH}}(X, X_r))}. \quad (8)$$

Если для заданного уровня иерархии можно получить достаточно надежное решение (7), то процесс поиска на нем и останавливается. В противном случае, описание входного объекта детализируется (что приводит к увеличению числа однородных сегментов K), и процесс принятия решений (5)–(7) повторяется заново до тех пор, пока не будет получено надежное решение (7) или время обработки не превысит порог t_0 (1). Обычно количество шагов ограничивается исследователем исходя из особенностей конкретной задачи ($J = \text{const}$). Если на J -м шаге надежное решение в смысле (7) не было найдено, среди потенциальных решений, полученных на всех J уровнях, выбирается решение с максимальной

апостериорной вероятностью (8). Заметим, что каждый уровень предложенной иерархической трехпороговой системы можно рассматривать как добавление условия (7) к последовательной тернарной классификации (three-way decisions) [17]. Каждый критерий (5)–(7) определяет приближенное множество с переменной точностью [18], которое выражает степень неуверенности исследователя в принятом решении (3) и задается вероятностями ошибок первого и второго рода. При этом все критерии служат для решения своих задач [19]: выражение (5) является основой для методов сокращения перебора, критерий (6) применяется для отбраковки объектов, не принадлежащих ни к одному классу, а выражение (7) позволяет выбрать решения, наиболее различающиеся от остальных эталонов.

В отличие от известных иерархических методов, таких как Pyramid HOG (PHOG) [20], предложенный подход не требует обязательной обработки всех уровней иерархии, что приводит к повышению вычислительной эффективности классификации. Пусть $K^{(j)}$ — число сегментов во входном объекте на j -м уровне детализации. Вычислительная сложность предложенного последовательного иерархического трехпорогового метода классификации в худшем случае получения ненадежных решений на всех J уровнях совпадает со сложностью традиционных иерархических алгоритмов и может быть оценена как $O\left(N \sum_{j=1}^J \sum_{r=1}^R \sum_{k=1}^{K^{(j)}} |N_r(k)|\right)$. Если предположить, что на каждом уровне иерархии доля ненадежных решений определяется некоторой фиксированной константой $0 \leq \gamma < 1$, то сложность предложенного подхода оценивается как $O\left(N \sum_{j=1}^J \left(\gamma^{j-1} \sum_{r=1}^R \sum_{k=1}^{K^{(j)}} |N_r(k)|\right)\right)$. Если размерность окрестности $N_r(k)$ является константой ($|N_r(k)| = \Delta$), а количество сегментов на $(j + 1)$ -м уровне иерархии в m раз превышает количество сегментов на j -м уровне (как для метода PHOG [20]), то алгоритмическая сложность разработанного метода для худшего случая $O(N\Delta RK^{(1)}(m^J - 1)/(m - 1))$, в то время как средняя сложность для доли γ недостаточно надежных решений на каждом уровне оценивается как $O(N\Delta RK^{(1)}((\gamma m)^J - 1)/(\gamma m - 1))$, т. е. оказывается примерно в $\gamma^{-(J-1)}$ раз быстрее. Например, для средней доли недостаточно надежных решений 20% уже для $J = 2$ уровней среднее время классификации для предложенного подхода оказывается в $0,2^{-1} = 5$ раз ниже по сравнению с более традиционной обработкой всех уровней иерархии.

Для повышения вычислительной эффективности посегментного анализа однородности на каждом уровне иерархии для больших баз данных (несколько тысяч классов) на основе условия досрочного останова (5) может применяться предложенный автором метод максимально правдоподобного направленного перебора (МПНП) [21]. Вначале выбирается один из эталонов X_{r_1} и вычисляется рассогласование $\rho_{\text{PNNH}}(X, X_{r_1})$. Далее последовательно многократно повторяется следующая процедура. Пусть на k -м этапе были проверены эталоны X_{r_1}, \dots, X_{r_k} . Следующий эталон $X_{r_{k+1}}$ ищется по принципу максимального правдоподобия (совместного распределения расстояний $\rho_{\text{PNNH}}(X, X_{r_1}), \dots, \rho_{\text{PNNH}}(X, X_{r_k})$ при справедливости W_ν). Для оценки этого правдоподобия воспользуемся тем фактом, что статистика $nK\rho_{\text{PNNH}}(X, X_{r_i})$ асимптотически (при $n \rightarrow \infty, n_r \rightarrow \infty$) распределена нормально (при больших KN) как

$$\mathcal{N}\left(\rho_{\nu, r_i} + \frac{N - 1}{n}; \left(\frac{\sqrt{4nK\rho_{\nu, r_i} + 2K(N - 1)}}{nK}\right)^2\right), \tag{9}$$

где $\rho_{\nu, r_i} = \rho_{\text{PNNH}}(X_\nu, X_{r_i})$.

В работе [21] показано, что в таком случае для выбора следующего эталона следует воспользоваться выражением:

$$r_{k+1} = \arg \min_{\mu \in \{1, \dots, R\} - \{r_1, \dots, r_k\}} \left(\sum_{i=1}^K \varphi_{\mu}(r_i) - \ln p_{\mu} \right), \quad (10)$$

где в предположении о том, что средний размер сегмента значительно превышает число различных значений признаков (ячеек в гистограмме) $n \gg N$,

$$\varphi_{\mu}(r_i) = \frac{(\rho_{\text{PNNH}}(X, X_{r_i}) - \rho_{\mu, r_i})^2}{\rho_{\mu, r_i}}. \quad (11)$$

Здесь $\varphi_{\mu}(r_i)$ тем меньше, чем ближе между собой рассогласования $\rho_{\text{PNNH}}(X, X_{r_i})$ и ρ_{μ, r_i} и чем больше рассогласование ρ_{μ, r_i} между эталонами X_{μ} и X_{r_i} . При наличии априорной информации о частотах появления каждого класса согласно (10) в первую очередь будут проверяться объекты из классов с большей априорной вероятностью. Если для эталона $X_{r_{k+1}}$ выполняется условие останова (5), то алгоритм завершается на этом этапе $K_{\text{ML-DEM}} = k + 1$. В противном случае процедура направленного поиска наиболее правдоподобного (с точки зрения его согласованности с текущими значениями рассогласований) эталона (10), (11) повторяется.

По сравнению с реализацией критерия (3) метод МПНП требует дополнительной памяти для хранения $R^2 + 2R$ вещественных чисел. Алгоритмическая сложность метода составляет $O\left(K_{\text{ML-DEM}} \left(R + (2N/R) \sum_{r=1}^R \sum_{k=1}^K |N_r(k)|\right)\right)$, что примерно в $R/K_{\text{ML-DEM}}$ раз быстрее полного перебора (3). При наличии параллельной среды выполнения (кластер машин, многоядерный процессор и т. п.) метод МПНП может быть реализован в параллельном варианте [22], в котором все эталоны распределяются на T доступных узлов, после чего на каждом узле применяется МПНП. Узел, первым нашедший эталон, удовлетворяющий условию (5), выдает команду остальным узлам о досрочном окончании перебора. В итоге такая параллельная реализация метода МПНП требует хранения $2R + 2M \sum_{r=1}^R K_r + R^2/T$ чисел, а вычислительная сложность в расчете на одну задачу может быть оценена как $O\left((K_{\text{ML-DEM}}/T) \left(R/T + (2N/R) \sum_{r=1}^R \sum_{k=1}^K |N_r(k)|\right)\right)$.

4 Методология классификации в условиях малых выборок и большого числа классов

На рис. 1 представлена предлагаемая методология ПАО для систем классификации в условиях малых выборок и большого числа классов. Здесь для поступающего на вход классифицируемого объекта X осуществляется его предварительная обработка с целью достижения инвариантности к наиболее характерным для предметной области изменениям условий наблюдения и последующая сегментация с помощью существующих методов, хорошо зарекомендовавших себя в конкретной прикладной области, таких как разбиение на небольшие фреймы фиксированной размерности, наращивание областей, выделение контуров и др. [12]. Далее для каждого выделенного сегмента выполняется извлечение примитивных признаков, таких как яркость или направление градиента пикселя (для изображений) [23] или оценки спектральной плотности мощности, кепстральные коэффициенты (для речевых сигналов) [5]. Рассматривая последовательность примитивных признаков как простую случайную выборку из одной (для фиксированного сегмента) генеральной совокупности, на основе описанных выше методов определяется мера близости



Рис. 1 Методология посегментного анализа однородности для систем классификации

входного и эталонных объектов (4) с ПАО и выравниванием сегментов в некоторой окрестности $N_r(k)$, выбираемой, исходя из особенности конкретной задачи и метода сегментирования. Решение принимается в пользу класса, соответствующего ближайшему эталону (3). Для повышения вычислительной эффективности классификации по обучающей выборке оценивается порог для досрочного останова (5), после чего в рамках иерархической трехпороговой системы [19] для поиска приближенного решения применяется метод МПНП [21]. Если для полученного решения условие (5) не выполняется, последовательно (6) и (7) проверяется его надежность. Если решение оказывается недостаточно надежным (7), то можно либо задействовать режим переспроса, либо при наличии времени для уточнения решения повысить степень детализации входного объекта X (например, снизить порог для объединения нескольких небольших фреймов в один сегмент, что приведет к возрастанию количества сегментов K и уменьшению средней размерности сегмента n), после чего процедура (5)–(7) повторяется. Если надежное решение не было получено на всех уровнях иерархии, итоговое решение принимается с помощью комитета классификаторов [1]. В условиях малых выборок последний строится на основе принципа максимума апостериорной вероятности (8). Далее принятое решение может использоваться в составе интеллектуальной системы в рамках существующих методов [1, 4] для дальнейшей обработки (например, сохранения результатов классификации в базе данных [1]) и для автоматического или автоматизированного принятия управленческих решений в составе системы поддержки принятия решений.

Для примера, рассмотрим применение методологии (см. рис. 1) в задаче распознавания изображений. Прежде всего во входном изображении происходит детектирование классифицируемого объекта X с высотой U и шириной V и осуществляется его предварительная обработка (нормирование по освещенности, эквализация гистограмм, медианная фильтрация и т. п.). Далее выполняется его сегментация на $K_1 \times K_2$ блоков фиксированного размера (по K_1 строк и K_2 столбцов). В модуле извлечения признаков вычисляются направления градиента из хорошо зарекомендовавшего себя на практике метода НОГ [23]. Множество значений ориентации градиента разбивается на N частей и для каждого сегмента (k_1, k_2) вычисляются гистограммы $H(k_1, k_2)$. Каждый выделенный объект классифицируется в модулях выбора эталонов (с помощью метода МПНП) и измерения сходства на основе посегментного анализа однородности (4). Предполагается, что база данных содержит эталоны X_r , которые тоже разбиваются на фрагменты и для каждого блока вычисляются гистограммы $H_r(k_1, k_2)$ направлений градиентов r -го эталона. Согласно модели кусочно-однородного объекта, соседями сегмента (k_1, k_2) считаются участки $(\tilde{k}_1, \tilde{k}_2)$, такие что $|\tilde{k}_1 - k_1| \leq \Delta$, $|\tilde{k}_2 - k_2| \leq \Delta$, где обычно $\Delta = 0$ или 1. Решение задачи ищется в виде, аналогичном (3):

$$\nu = \arg \min_{r \in \{1, \dots, R\}} \left(\sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \rho(H(k_1, k_2), H_r(k_1 + \Delta_1, k_2 + \Delta_2)) - \ln p_r \right), \quad (12)$$

где мера близости $\rho(H(k_1, k_2), H_r(k_1 + \Delta_1, k_2 + \Delta_2))$ определяется аналогично (4). Далее согласно иерархической трехпороговой схеме для ближайшего эталона вычисляется апостериорная вероятность (8), которая используется для принятия окончательного решения в модуле обработки результатов и адаптации. В этом же модуле для каждого объекта из текущего кадра находится ближайший в смысле (12) объект из предыдущего кадра. Если такой объект не найден или минимальное расстояние превышает фиксированный порог (6), то в список изображений идентичных объектов добавляется новый объект. Если объект из списка не был найден в течение некоторого достаточно продолжительного времени, он удаляется из списка. Если для первоначального уровня детализации изображения можно получить достаточно надежное решение (7), (8), то процесс поиска на нем и останавливается. В противном случае количество блоков K_1, K_2 увеличивается и процесс принятия решений повторяется заново.

Аналогичный подход может применяться и в задаче распознавания изолированных слов [5]. Отличие состоит в том, что каждый сегмент входного слова следует сопоставлять не со всеми сегментами всех эталонов (3), а только с L эталонными фонемами (или сенаонами) из фонетической базы данных. Поэтому здесь не требуется наличия речевых сигналов для каждой команды из словаря. Вместо этого оказывается достаточно сохранить в базе данных соответствующие сегментам фонемы, изолированно произнесенные диктором [24]. Так как количество фонем L невелико, применение методов приближенного ближайшего соседа, таких как МПНП, оказывается здесь нецелесообразным. Однако иерархический анализ сигналов с последовательным уменьшением длительности сегментов может представлять существенный интерес.

5 Результаты экспериментальных исследований

5.1 Посегментный анализ однородности

В настоящем разделе результаты экспериментальных исследований представленной методологии (см. рис. 1) с точки зрения сформулированного ЦК (1) в задаче распозна-

Таблица 1 Оценка вероятности ошибки распознавания лиц (в %)

Метод	$x_\eta = 0$	$x_\eta = 3$	$x_\eta = 5$
SIFT	20,4 ± 2,2	22,1 ± 2,2	25,8 ± 2,3
Eigenfaces	27,7 ± 1,3	27,9 ± 1,5	28,5 ± 1,6
LBPН	15,9 ± 1,1	16,1 ± 1,2	18,2 ± 1,3
SVM	12,8 ± 1,3	14 ± 1,3	17,1 ± 1,6
Евклида	11,7 ± 1,6	13,6 ± 1,6	16,8 ± 1,7
ВНС (2), $\Delta = 0$	10,8 ± 1,4	12,1 ± 1,5	14,3 ± 1,6
ПАО (4), (5), $\Delta = 0$	9,4 ± 1,2	10,9 ± 1,2	12,9 ± 1,3
Евклида, $\Delta = 1$	9,9 ± 1,5	12,4 ± 1,5	18,8 ± 1,9
ВНС (2), $\Delta = 1$	8,5 ± 1,3	10,9 ± 1,7	15,6 ± 1,8
ПАО (4), (5), $\Delta = 1$	7,7 ± 1,2	9,4 ± 1,2	14,5 ± 1,4

вания лиц [25, 26]. Все эксперименты проводились на современном ноутбуке (процессор Intel Core i7-2630QM, тактовая частота 2 ГГц, объем ОЗУ 6 ГБ). Лица на фотографиях выделялись с помощью библиотеки OpenCV. Далее выполнялась медианная фильтрация с размером окна 3×3 пикселя [12, 14]. Число элементов в гистограмме ориентированных градиентов было выбрано $N = 8$. Для непараметрических оценок распределений в (2), (4) использовалось гауссовское ядро Парзена–Розенблатта с параметром сглаживания (среднеквадратичным отклонением) $\sigma = 0,577$. Эти значения показали в описываемых экспериментах наилучшую точность для рассматриваемых баз данных. Общее рассогласование между двумя фотографиями рассчитывалось как сумма рассогласований между соответствующими фрагментами (12) с выравниванием сегментов в окрестности $\Delta = 0$ и 1. Кроме мер близости (2) и (4) использовалось традиционное расстояние Евклида и машина опорных векторов (SVM) [1, 2]. Кроме признаков HOG проводилось исследование классического метода SIFT [27], а также известные методы распознавания лиц, реализованные в библиотеке OpenCV: Eigenfaces [25] и LBPН (Local Binary Patterns Histograms) [26]. В настоящем подразделе предположим, что ограничения на максимальное время классификации отсутствуют ($t_0 = +\infty$) и все методы сопоставляются только по значениям средней вероятности ошибки $\bar{\alpha}_\eta$ (1).

Далее представим некоторые результаты для классификации 2720 фронтальных изображений 994 различных людей из базы FERET для случая равенства априорных вероятностей $p_r = 1/R$ (полная априорная неопределенность).

Итоговая близость между изображениями лиц оценивается как взвешенная сумма рассогласований (12) между соответствующими частями. К интенсивности каждой точки всех изображений из контрольной выборки добавлялось случайное число в диапазоне $[-x_\eta; x_\eta]$, где $x_\eta \geq 0$ определяет уровень шума. В табл. 1 приведены оценки средних значений и стандартных отклонения для вероятности ошибки $\bar{\alpha}_\eta$ в ЦК (1), которые оценивались методом кросс-валидации с помощью 20-кратного случайного выбора фиксированного числа эталонов $R = 1370$. Здесь, во-первых, предложенный подход с ПАО оказывается достаточно устойчивым к небольшим искажениям тестовых данных. Так, при выравнивании сегментов в $\Delta = 1$ -окрестности каждого блока для уровня шума $x_\eta \leq 3$ точность уменьшается на 1,5%–1,8%, что несколько лучше по сравнению с другими мерами близости. Однако для большого уровня шума ($x_\eta = 5$) выравнивание ($\Delta = 1$) приводит к росту вероятности

ошибки на 1,3%–2%, поэтому в таком случае следует применять наиболее простые алгоритмы классификации без выравнивания сегментов $\Delta = 0$. Во-вторых, предложенный подход (3), (4), (12) с выравниванием сегментов в Δ -окрестности характеризуется наименьшей вероятностью ошибки при наличии слабых помех ($x_\eta \leq 3$) по сравнению с традиционным суммированием расстояний между гистограммами ($\Delta = 0$). В-третьих, вероятность ошибки традиционных методов SIFT и SVM превосходит даже правило ближайшего соседа с метрикой Евклида, что может быть объяснено недостаточным для анализа числом эталонов для каждого класса (человека), что оказывает негативное влияние на качество обучения SVM. Однако и для этого случая предложенные модификации ВНС с посегментным анализом однородности (3), (4) характеризуются наименьшей вероятностью ошибки. Гистограммы локальных бинарных шаблонов (LBPН) оказались намного более точными (на 11%–12%) по сравнению с собственными лицами (eigenfaces). Однако вероятность ошибки LBPН на 1,6%–4,2% выше вероятности ошибки для признаков HOG даже для простейшей метрики Евклида и отсутствия выравнивания ($\Delta = 0$). А использование ПАО позволяет повысить точность еще на 2,3%–3,9%. При этом во всех случаях критерий Мак-Немара на уровне значимости 0,05 показал, что предлагаемый ПАО приводит к значимому повышению точности распознавания изображений по сравнению как с алгоритмами распознавания лиц из библиотеки OpenCV (Eigenfaces, гистограммы LBP), так и с традиционными подходами (машина опорных векторов, метод ближайшего соседа).

5.2 Метод максимально правдоподобного перебора

В настоящем подразделе рассмотрена возможность учета важных для практики ограничений на среднее время принятия решений в рамках ЦК (1). На рис. 2 представлено среднее время классификации t_{avg} для полного перебора, рандомизированного k-d дерева [8], зарекомендовавшего себя для небольших R метода perm-sort [9] и описанного выше МПНП [21].

Анализ табл. 1 и рис. 2 позволяет сделать следующие выводы. Во-первых, с точки зрения ЦК (1) применение традиционной метрики Евклида является приемлемым только для предельно жестких ограничений на среднее время классификации $t_0 \leq 3$ мс, так как вероятность ошибки здесь превышает аналогичный показатель для предложенного

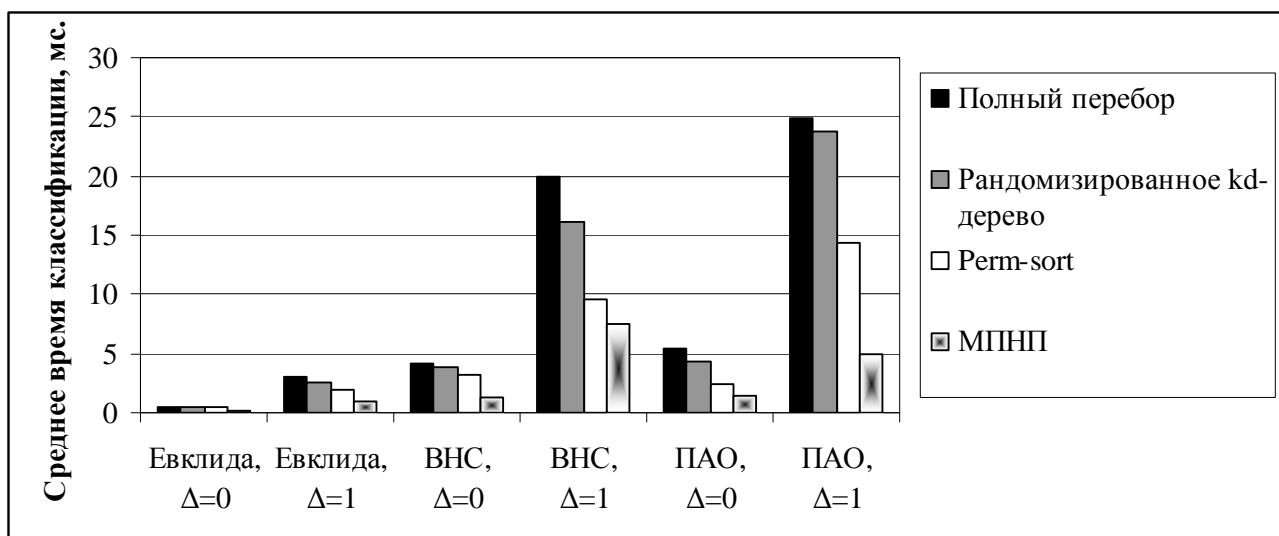


Рис. 2 Среднее время распознавания одного лица

подхода (2), (4) на 2%–4%. При этом для наиболее простого с вычислительной точки зрения случая $\Delta = 0$ применение методов приближенного поиска ближайшего соседа оказывается неоправданным. В то же время с точки зрения ЦК (1) байесовский критерий проверки гипотез о распределении в реализации ВНС (2) во всех случаях оказывается хуже ПАО (3), (4). Во-вторых, реализация предложенного подхода (3), (4) с выравниванием сегментов в Δ -окрестности в рамках полного перебора (12) удовлетворяет ЦК (1) лишь для $t_0 \geq 25$ мс. В-третьих, среднее время поиска для рандомизированного k-d дерева практически не отличается от времени распознавания для полного перебора, что можно объяснить небольшим размером обучающей выборки R . В то же время метод `perm-sort` позволил в 1,3–2,3 раз ускорить классификацию по сравнению с полным перебором. Наконец, основной вывод состоит в том, что предложенный метод МПНП (5), (10), (11) при сопоставимой вероятности ошибки классификации характеризуется наилучшей вычислительной эффективностью: в 2,5–5 раз быстрее по сравнению с полным перебором и в 1,4–3 раза быстрее, чем `perm-sort`. При этом для наиболее актуального для промышленных систем случая $5 < t_0 < 25$ мс наилучшим с точки зрения ЦК оказывается именно применение метода МПНП для ПАО (3), (4) с выравниванием сегментов ($\Delta = 1$). Здесь вероятность ошибки 8,2% всего на 0,5% выше вероятности ошибки полного перебора (12), а время принятия решений $t_{\text{avg}} = 4,9$ мс.

5.3 Иерархическая трехпороговая система

В заключительном эксперименте исследовался иерархический подход (5)–(8). Для повышения быстродействия в пирамиде (иерархии) присутствуют всего два уровня: сетка 5×5 и сетка 10×10 . В предложенном подходе второй уровень иерархии (10×10) анализируется только в том случае, если для первого уровня (10×10) оценка апостериорной вероятности (8) не превышает порог $p_0 = 0,85$ (7). и 4 приведены оценки средней вероятности ошибки и времени классификации для неиерархического распознавания (сетки « 5×5 » и « 10×10 ») в сравнении с предложенным подходом. На рис. 3 Последний был

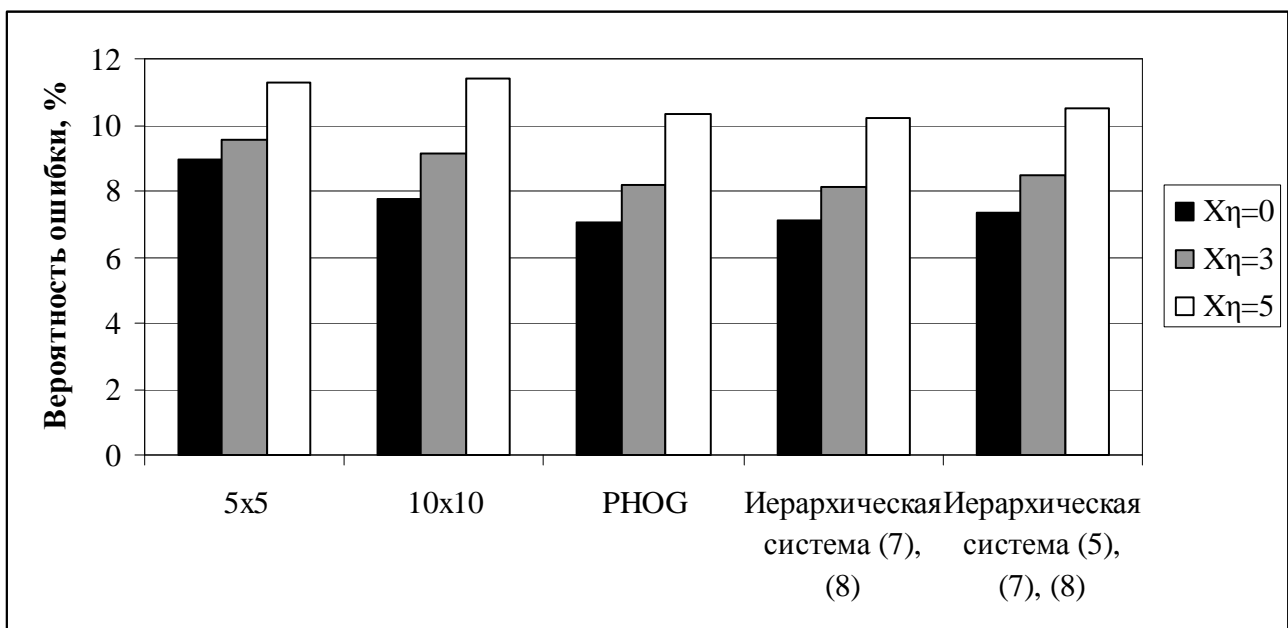


Рис. 3 Оценка вероятности ошибки распознавания для последовательного анализа пирамид НОГ

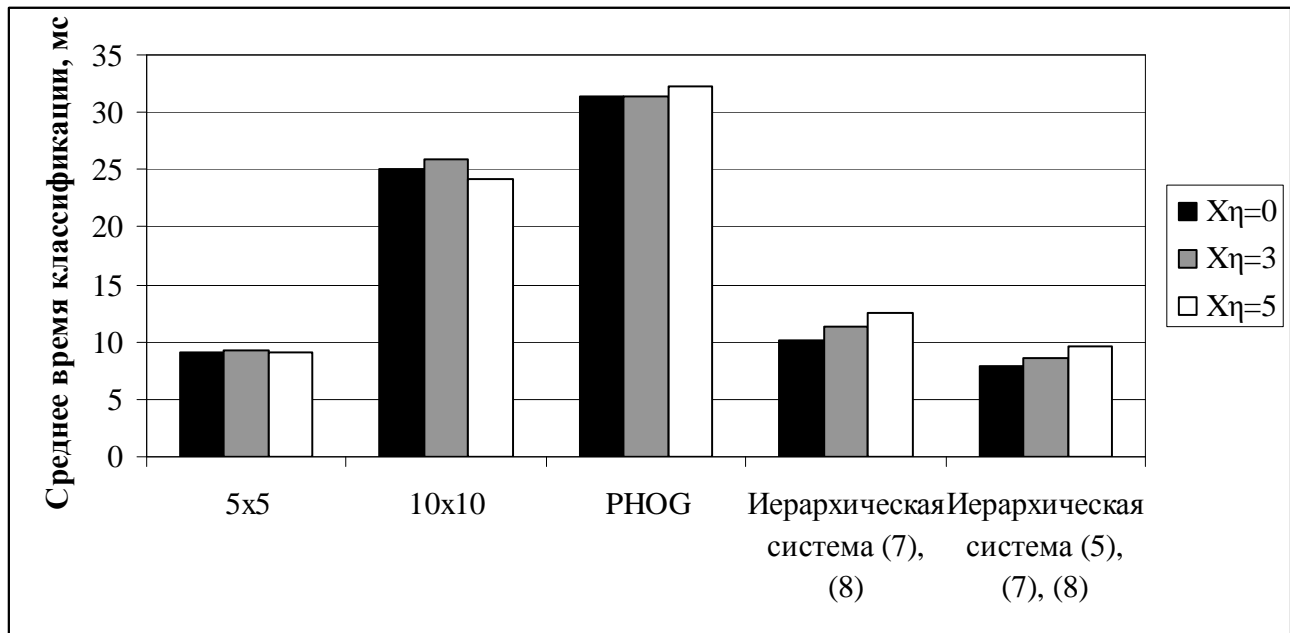


Рис. 4 Среднее время распознавания одного лица для последовательного анализа пирамид НОГ

реализован в двух вариантах: с полным анализом каждого уровня и с досрочным остановом перебора (5). Кроме того, использовался традиционный метод РНОГ [20], в котором решение принимается по минимуму взвешенной суммы расстояний между гистограммами ориентированных градиентов для каждого уровня иерархии.

Здесь точность классификации для иерархического подхода оказывается выше, чем точность каждого уровня пирамиды, что особенно заметно для малых обучающих выборок. Различия в вероятности ошибки РНОГ и предложенного подхода (см. рис. 3) оказываются незначимыми. Однако с точки зрения ЦК (1) разработанная иерархическая система имеет несомненное преимущество с точки зрения вычислительной эффективности. Среднее время распознавания для предложенного подхода оказывается в 2,5–3 раза ниже по сравнению с РНОГ, так как в большинстве случаев уже на первом уровне (сетки 5×5) удается найти достаточно надежное решение (7). А использование досрочного останова перебора (5) позволяет еще на 30% снизить t_{avg} без существенных потерь в точности.

6 Заключение

Таким образом, в настоящей работе подробно описана комплексная методология ПАО (см. рис. 1), позволяющая повысить точность и вычислительную эффективность систем классификации аудиовизуальной информации в условиях малых выборок и большого числа классов. Ее ключевым звеном является подход к классификации аудиовизуальной информации с анализом однородности сегментов и их динамическим выравниванием, который позволил строго обосновать наблюдаемую во многих промышленных системах недостаточную точность классического байесовского классификатора, реализуемого в ВНС, по сравнению с методами ближайшего соседа, обычно объясняемую «наивностью» предположения о независимости признаков. На основе этого подхода разработан иерархический трехпороговый метод классификации, позволивший в 1,5–3,5 раза ускорить процедуру принятия решений по сравнению с традиционными иерархическими методами (такими, как РНОГ). Наконец, рассмотрен приближенный поиск ближайшего соседа в рамках МПНП, основанный на асимптотических свойствах критерия (3), (4). Экспериментально

показано, что метод МПНП ускоряет классификацию лиц в 2–5 раз по сравнению с известными алгоритмами (рандомизированные k-d деревья, perm-sort) для баз данных, содержащих тысячи фотографий. С точки зрения развития данной темы следующим шагом, на наш взгляд, может стать подробный анализ особенностей реализации предложенного подхода (см. рис. 1) к задачам классификации сигналов (прежде всего, распознавания речи), а также реализация ПАО для задач регрессии в виде расширения известного метода GRNN (Generalized Regression Neural Network) [28].

Литература

- [1] Pattern recognition / Eds. S. Theodoridis, K. Koutroumbas. — Academic Press, 2008. 984 p.
- [2] *Haykin S.* Neural networks and learning machines — 3rd ed. — Prentice Hall, 2008. 936 p.
- [3] *Журавлев Ю. И., Рудаков К. В., Гуров С. И., Дюкова Е. В., Кутуков Г. П., Матюнин С. Н., Местецкий Л. М.* Состояние и перспективы развития исследований в области обработки и распознавания видеоинформации. Аналитический обзор // Наука и образование, 2005. Т. 1. <http://technomag.bmstu.ru/doc/48995.html>.
- [4] *Hammerstrom D. W., Rehfuess S.* Neurocomputing hardware: Present and future // Artificial Intell. Rev., 1993. Vol. 7. No. 5. P. 285–300.
- [5] Springer handbook of speech processing / Eds. J. Benesty, M. Sondh, Y. Huang. — Springer, 2008. 1176 p.
- [6] *Tan X., Chen S., Zhou G. H, Zhang F.* Face recognition from a single image per person: A survey // Pattern Recogn., 2006. Vol. 39. No. 9. P. 1725–1745.
- [7] *Savchenko A. V.* Directed enumeration method in image recognition // Pattern Recogn., 2012. Vol. 45. No. 8. P. 2952–2961.
- [8] *Silpa-Anan C., Hartley R.* Optimised KD-trees for fast image descriptor matching // IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR) Proceedings, 2008. P. 1–8.
- [9] *Gonzalez E. C., Figueroa K., Navarro G.* Effective proximity retrieval by ordering permutations // Pattern Anal. Machine Intell., 2008. Vol. 30. No. 9. P. 1647–1658.
- [10] *Боровков А. А.*, Математическая статистика: дополнительные главы. — М.: Наука, 1984. 144 с.
- [11] *Абусев Р. А., Лумельский Я. П.* Статистические модели классификации многомерных наблюдений // Обозрение прикладной и промышленной математики, 1996. Т. 3. С. 7–30.
- [12] *Shapiro L. G, Stockman G. C.* Computer vision. — Prentice Hall, 2001. 608 p.
- [13] *Савченко А. В.* Образ как совокупность выборок независимых одинаково распределенных значений признаков в задачах распознавания сложноструктурированных объектов // Западская лаборатория. Диагностика материалов, 2014. Т. 80. № 3. С. 70–80.
- [14] *Savchenko A. V.* Probabilistic neural network with homogeneity testing in recognition of discrete patterns set // Neural Networks, 2013. Vol. 46. P. 227–241.
- [15] *Specht D.F.* Probabilistic neural networks // Neural Networks, 1990. Vol. 3. P. 109–118.

- [16] *Chow C. K.* On optimum recognition error and reject trade-off // *IEEE Trans. Inform. Theory*, 1970. Vol. 16. P. 41–46.
- [17] *Yao Y. Y.* Granular computing and sequential three-way decisions // *Conference (International) on Rough Sets and Knowledge Technology (RSKT) Proceedings, LNCS/LNAI*, 2013. Vol. 8171. P. 16–27.
- [18] *Dou H., Yang X., Fan J., Xu S.* The models of variable precision multigranulation rough sets // *Conference (International) on Rough Sets and Knowledge Technology (RSKT) Proceedings, LNCS/LNAI*, 2012. Vol. 7414. P. 465–473.
- [19] *Савченко А. В.* Трехпороговая система автоматического распознавания изображений // *Искусственный интеллект и принятие решений*, 2011. № 4. С. 102–109.
- [20] *Bosch A., Zisserman B., Munoz X.* Representing shape with a spatial pyramid kernel // *ACM Conference (International) on Image and Video Retrieval (CIVR) Proceedings*, 2007. P. 401–408.
- [21] *Savchenko A. V.* An optimal greedy approximate nearest neighbor method in statistical pattern recognition // *Conference (International) on Pattern Recognition and Machine Intelligence (PReMI), LNCS*, 2015. Vol. 9124. P. 1–10.
- [22] *Savchenko A. V.* Real-time image recognition with the parallel directed enumeration method // *Conference (International) on Vision Systems (ICVS) Proceedings, LNCS*, 2013. Vol. 7963. P. 123–132.
- [23] *Dalal N., Triggs B.* Histograms of oriented gradients for human detection // *IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR) Proceedings*, 2005.
- [24] *Savchenko A. V.* Phonetic words decoding software in the problem of Russian speech recognition // *Automation Remote Control*, 2013. Vol. 74. No. 7. P. 1225–1232.
- [25] *Turk M. A., Pentland A. P.* Face recognition using eigenfaces // *IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR) Proceedings*, 1991. P. 586–591.
- [26] *Ahonen T., Hadid A., Pietikainen M.* Face recognition with local binary patterns // *European Conference on Computer Vision Proceedings*, 2004. P. 469–481.
- [27] *Lowe D.* Distinctive image features from scale-invariant keypoints // *Int. J. Computer Vision*, 2004. Vol. 60. No. 2. P. 91–110.
- [28] *Specht D. F.* A general regression neural network // *IEEE Trans. Neural Networks*, 1991. Vol. 2. No. 6. P. 568–576.

References

- [1] Theodoridis, S., and K. Koutroumbas, eds. 2008. *Pattern recognition*. 4th ed. Academic Press. 984 p.
- [2] Haykin, S. 2008 *Neural networks and learning machines*. 3rd ed. Prentice Hall. 936 p.
- [3] Zhuravlev, Yu. I., K. V. Rudakov, S. I. Gurov, E. V. Dyukova, G. P. Kutukov, S. N. Matunin, and L. M. Mestetsky. 2005. Sostoyanie i perspektivy razvitiya issledovaniy v oblasti obrabotki i raspoznavaniya videoinformatsii. Analiticheskiy obzor [State and development perspectives of research in video processing and recognition. Analytical survey]. *Nauka i obrazovanie* [Science

- and Education] 1. Available at: <http://technomag.bmstu.ru/doc/48995.html> (accessed June 23, 2015). (In Russian.)
- [4] Hammerstrom, D. W., and S. Rehfuss. 1993. Neurocomputing hardware: Present and future. *Artificial Intell. Rev.* 7(5):285–300.
- [5] Benesty, J., M. Sondh, and Y. Huang, eds. 2008. *Springer handbook of speech processing*. Springer. 1176 p.
- [6] Tan, X., S. Chen, G. H. Zhou, and F. Zhang. 2006. Face recognition from a single image per person: A survey. *Pattern Recogn.* 39(9): 1725–1745.
- [7] Savchenko, A. V. 2012. Directed enumeration method in image recognition. *Pattern Recogn.* 45(8):2952–2961.
- [8] Silpa-Anan, C., and R. Hartley. 2008. Optimised KD-trees for fast image descriptor matching. *IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR) Proceedings*. 1–8.
- [9] Gonzalez, E. C., K. Figueroa, and G. Navarro. 2008. Effective proximity retrieval by ordering permutations *IEEE Trans. Pattern Anal. Machine Intell.* 30(9):1647–1658.
- [10] Borovkov, A. A. 1984. *Matematicheskaya statistika: Dopolnitelnye glavy* [Mathematical statistics: Additional chapters]. Moscow: Nauka. 144 p. (In Russian.)
- [11] Abusev, R. A., and Ya. P. Lumelskiy. 1996. Statisticheskie modeli klassifikatsii mnogomernykh nablyudeniy [Statistical models of classification of multivariate observations]. *Obozrenie prikladnoy i promyshlennoy matematiki* [Review of Applied and Industrial Mathematics] 3:7–30. (In Russian.)
- [12] Shapiro, L. G., and G. C. Stockman. 2001. *Computer vision*. Prentice Hall. 608 p.
- [13] Savchenko, A. V. 2014. Obraz kak sovokupnost' vyborok nezavisimykh odinakovo raspredelennykh znacheniy priznakov v zadachakh raspoznavaniya slozhnostrukturirovannykh ob'ektov [Pattern as a set of samples of independent identically distributed features in the tasks of recognition of complex objects]. *Zavodskaya laboratoriya. Diagnostika materialov* [Industrial Laboratory. Materials Diagnostics] 80(3):70–80. (In Russian.)
- [14] Savchenko, A. V. 2013. Probabilistic neural network with homogeneity testing in recognition of discrete patterns set. *Neural Networks* 46:227–241.
- [15] Specht, D. F. 1990. Probabilistic neural networks. *Neural Networks* 3:109–118.
- [16] Chow, C. K. 1970. On optimum recognition error and reject trade-off. *IEEE Trans. Inform. Theory* 16:41–46.
- [17] Yao, Y. Y. 2013. Granular computing and sequential three-way decisions. *Conference (International) on Rough Sets and Knowledge Technology (RSKT) Proceedings, LNCS/LNAI*. 8171:16–27.
- [18] Dou, H. X. Yang, J. Fan, and S. Xu. 2012. The models of variable precision multigranulation rough sets. *Conference (International) on Rough Sets and Knowledge Technology (RSKT) Proceedings, LNCS/LNAI*. 7414:465–473.

-
- [19] Savchenko, A. V. 2011. Trekhporogovaya sistema avtomaticheskogo raspoznavaniya izobrazheniy [Automatic image recognition three-threshold system]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making] 4:102–109. (In Russian.)
- [20] Bosch, A., B. Zisserman, and X. Munoz. 2007. Representing shape with a spatial pyramid kernel. *ACM Conference (International) on Image and Video Retrieval (CIVR) Proceedings*. 401–408.
- [21] Savchenko, A. V. 2015. An optimal greedy approximate nearest neighbor method in statistical pattern recognition. *Conference (International) on Pattern Recognition and Machine Intelligence (PReMI) Proceedings, LNCS*. 9124:1–10
- [22] Savchenko, A. V. 2013. Real-time image recognition with the parallel directed enumeration method. *Conference (International) on Vision Systems (ICVS), LNCS*. 7963:123–132.
- [23] Dalal, N., and B. Triggs. 2005. Histograms of oriented gradients for human detection. *IEEE Conference (International) Computer Vision and Pattern Recognition (CVPR) Proceedings*. 886–893.
- [24] Savchenko, A. V. 2013. Phonetic words decoding software in the problem of Russian speech recognition. *Automation Remote Control* 74(7):1225–1232.
- [25] Turk, M. A., and A. P., Pentland. 1991. Face recognition using eigenfaces. *IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR) Proceedings*. 586–591.
- [26] Ahonen, T., A. Hadid, and M. Pietikainen. 2004. Face recognition with local binary patterns. *European Conference on Computer Vision Proceedings*. 469–481.
- [27] Lowe, D. 2014. Distinctive image features from scale-invariant keypoints. *Int. J. Computer Vision* 60(2):91–110.
- [28] Specht, D. F. 1991. A general regression neural network. *IEEE Trans. Neural Networks* 2(6):568–576.