

## Эффективность иерархической классификации в терминах соотношения качество–сложность\*

*М. М. Ланге, С. Н. Ганебных*

lange\_mm@ccas.ru

Вычислительный центр им. А. А. Дородницына РАН, Москва

Исследуется многоклассовая метрическая модель распознавания объектов в пространстве древовидных представлений с многоуровневым разрешением. На множестве представлений построен иерархический ансамбль эталонов в виде наборов нарастающей мощности, которые образуют многоуровневую сеть. Рассмотрено параметрическое семейство решающих алгоритмов по критерию ближайшего эталона на основе процедур иерархического и переборного поиска решения в сети эталонов. Для решающих алгоритмов получены аналитические оценки вычислительной сложности и экспериментальные оценки вероятности ошибки распознавания объектов составного источника, заданного полутонными изображениями лиц, жестов руки и подписей. Эффективность иерархических и переборного алгоритмов продемонстрирована сравнительными зависимостями доли ошибок от вычислительной сложности.

**Ключевые слова:** метрический классификатор; изображение; представление с многоуровневым разрешением; многоуровневая сеть эталонов; вычислительная сложность

## An Efficiency of Hierarchical Classification in Terms of Fidelity-Complexity Ratio\*

*M. M. Lange, S. N. Ganebnykh*

Dorodnicyn Computing Centre of RAS, Moscow

**Background:** Most of existent approaches are limited by two-class classification models using exhaustive search scheme for the decision. The present approach is intended for multiclass classification model and is based on guided search scheme for the decision.

**Methods:** The examined classifier is constructed in a space of original tree-structured object representations and uses a multilevel network of template objects taken at sequential resolution levels. A parametric family of decision algorithms based on hierarchical and exhaustive search for the nearest template decision is examined.

**Results:** For the above decision algorithms, analytical estimates of the computational complexities are obtained. Also, experimental estimates of recognition error rates are calculated using a composite image source of faces, hand gestures, and signatures. The comparative dependences of the error rate on the computational complexity are shown.

**Concludibg Remarks:** The proposed methodology yields a new approach to efficient metric classification of the image-based objects provided that the number of classes is sufficiently large and this approach can be applied for recognition of biometric images.

**Keywords:** metric classifier; image; multiresolution representation; multilevel network of templates; computational complexity

## Введение

Проблема эффективной классификации данных в реальном времени связана с разработкой решающих алгоритмов, способных обеспечить малую долю ошибок при сравнительно небольшой вычислительной сложности. Актуальность таких алгоритмов особенно высока при большом числе классов, когда алгоритмы полного перебора не пригодны из-за чрезмерной вычислительной сложности. В общем случае эффективность решающего алгоритма характеризуется его вычислительной сложностью и долей ошибок, которые связаны обменным соотношением: с ростом сложности уменьшается доля ошибок и наоборот. В терминах соотношения качество–сложность задача классификации объектов в заданном пространстве их представлений сводится к построению решающего алгоритма, который при заданной вычислительной сложности обеспечивал бы возможно меньшую долю ошибок.

В работе [1] отмечалось, что в такой постановке задача классификации с решающим правилом на основе сравнения с эталонами [2] имеет сходство с известной из теории информации задачей кодирования источника с допустимой погрешностью, в которой на вход кодера поступают представления сообщений, получаемые с выхода канала наблюдения, вносящего искажения [3]. Для заданного канала наблюдения указанная задача сводится к минимизации скорости кодирования при заданной допустимой погрешности.

Пусть  $\mathbf{A} = \{\mathbf{A}_i\}_{i=1}^c$  – множество кластеров объектов источника (образов, заданных изображениями), где каждый кластер содержит объекты одного из  $c \geq 2$  классов;  $\mathbf{A}^* = \{\mathbf{A}_i^*\}_{i=1}^c$  – множество кластеров, образованных представлениями образов на выходе канала наблюдения;  $\mathbf{B} = \{\mathbf{B}_i \subset \mathbf{A}_i\}_{i=1}^c$  – множество обучающих кластеров;  $\hat{\mathbf{B}} = \{\hat{\mathbf{B}}_i \subseteq \mathbf{B}_i\}_{i=1}^c$  – множество наборов эталонных объектов;  $\hat{\mathbf{B}}^* = \{\hat{\mathbf{B}}_i^* \subseteq \mathbf{B}_i^* \subset \mathbf{A}_i^*\}_{i=1}^c$  – множество наборов представлений эталонов.

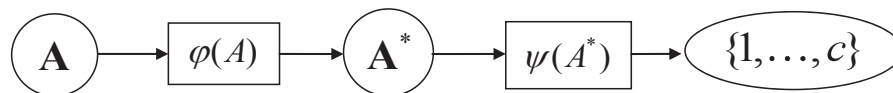


Рис. 1. Блок-схема модели кодера/классификатора в пространстве представлений

Блок-схема исследуемой модели классификатора с числом классов  $c \geq 2$  дана на рис. 1. Преобразования  $\varphi$  и  $\psi$  соответствуют функциям канала наблюдения и блока принятия решения. Функция  $\varphi : \mathbf{A} \rightarrow \mathbf{A}^*$  определяется способом представления объектов источника. Функция  $\psi : \mathbf{A}^* \rightarrow \{1, \dots, c\}$  задается решающим алгоритмом, который каждому объекту  $A \in \mathbf{A}$ , представленному описанием  $A^* \in \mathbf{A}^*$ , присваивает номер  $i \in \{1, \dots, c\}$  набора представлений  $\hat{\mathbf{B}}_i^* \subset \hat{\mathbf{B}}^*$ . Характеристиками эффективности классификатора являются доля ошибок  $\varepsilon_\psi(\hat{\mathbf{B}}^*)$  на множестве объектов  $\mathbf{A} \setminus \mathbf{B}$  и вычислительная сложность  $C_\psi(\hat{\mathbf{B}}^*)$  (в операциях на объект) решающего алгоритма  $\psi$ . Для заданного источника, при выбранном преобразовании  $\varphi$  и заданной допустимой сложности  $C > 0$  решающего алгоритма, оптимальному классификатору соответствует следующее соотношение качество–сложность

$$\varepsilon^*(C) = \min_{\hat{\mathbf{B}}^*, \psi: C_\psi(\hat{\mathbf{B}}^*) \leq C} \varepsilon_\psi(\hat{\mathbf{B}}^*) \quad (1)$$

Функция (1) представляет гипотетическую нижнюю границу зависимости доли ошибок от сложности, получение которой является открытой проблемой. В случае, когда преобразо-

вание  $\varphi$  и алгоритм  $\psi$  заданы, оптимизация классификатора при допустимой сложности  $C > 0$  алгоритма  $\psi$  сводится к нахождению множества эталонов, на котором достигается соотношение качество-сложность следующего вида

$$\varepsilon_{\psi}^*(C) = \min_{\hat{\mathbf{B}}^*: C_{\psi}(\hat{\mathbf{B}}^*) \leq C} \varepsilon_{\psi}(\hat{\mathbf{B}}^*). \quad (2)$$

Необходимо отметить, что для модели кодирования источника с допустимой погрешностью  $\varepsilon \geq 0$  аппроксимации объектов источника эталонами, в случае «бесшумного» канала наблюдения  $\varphi : \mathbf{A} \rightarrow \mathbf{A}$  и решающего алгоритма  $\psi : \mathbf{A} \rightarrow \hat{\mathbf{B}}$  по критерию ближайшего соседа (эталона), найдена зависимость минимальной скорости кодирования от  $\varepsilon$  в форме эпсилон-энтропии [4]. Эпсилон-энтропия достигается на множестве эталонов  $\hat{\mathbf{B}}$ , образованном центрами сфер радиуса  $\varepsilon$ , которыми может быть покрыто множество  $\mathbf{A}$  при условии, что число таких сфер минимально. В [3] получено обобщение эпсилон-энтропии для случая канала с искажениями  $\varphi : \mathbf{A} \rightarrow \mathbf{A}^*$  и алгоритма  $\psi : \mathbf{A}^* \rightarrow \hat{\mathbf{B}}^*$  по критерию ближайшего соседа при допустимой погрешности  $\varepsilon \geq \varepsilon_{\varphi}$  ( $\varepsilon_{\varphi}$  – погрешность канала). Подходы, развитые в работах [3, 4], могут быть полезны для построения оценок функций вида (1) и (2) в аналитической форме.

Сравнительное поведение функций  $\varepsilon^*(C)$  и  $\varepsilon_{\psi}^*(C)$  дано на рис. 2. Наименьшая доля ошибок  $\varepsilon = \varepsilon_{\psi}$  может быть достигнута при наибольшей сложности решающего алгоритма, а наибольшая доля ошибок реализуется при произвольном выборе решения, которому соответствует вычислительная сложность  $C = 0$ . В терминах функции  $\varepsilon_{\psi}^*(C)$ , эффективность алгоритма  $\psi$  тем выше, чем меньшую долю ошибок он обеспечивает при фиксированной сложности.

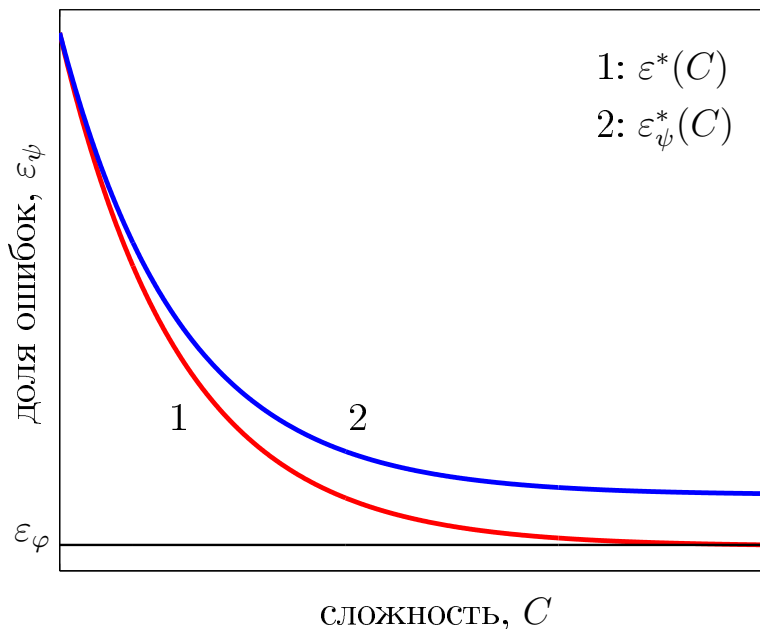


Рис. 2. Вид функций качество-сложность  $\varepsilon^*(C)$  и  $\varepsilon_{\psi}^*(C)$

Решающие алгоритмы, обеспечивающие достаточно малую долю ошибок при фиксированной вычислительной сложности, могут быть построены с использованием представлений образов с многоуровневым разрешением [5, 6, 7]. Древоподобная структура таких представлений позволяет сформировать иерархически структурированную сеть эталонов,

которая допускает применение быстрых решающих алгоритмов на основе направленного поиска [1, 8].

В настоящей работе исследуется задача распознавания полутоновых объектов в пространстве древовидно структурированных представлений эллиптическими примитивами [8]. Рассматриваются множества эталонов  $\hat{\mathbf{B}}$ , составленные из наборов нарастающей мощности, которые образуют ансамбль [1]. Используя ансамбль древовидно структурированных наборов представлений эталонов  $\hat{\mathbf{B}}^*$ , исследуются решающие алгоритмы двух типов: переборный ES алгоритм  $\psi_{\text{ES}}$  на основе переборного поиска (Exhaustive Search) решения в сети эталонов и иерархические GS алгоритмы  $\psi_{\text{GS}}$  на основе направленного поиска (Guided Search) решения. Для указанных алгоритмов получены аналитические оценки сложностей  $C_{\psi_{\text{ES}}}(\hat{\mathbf{B}}^*)$  и  $C_{\psi_{\text{GS}}}(\hat{\mathbf{B}}^*)$  и экспериментальные оценки долей ошибок  $\varepsilon_{\psi_{\text{ES}}}(\hat{\mathbf{B}}^*)$  и  $\varepsilon_{\psi_{\text{GS}}}(\hat{\mathbf{B}}^*)$  классификации объектов составного источника, заданного изображениями лиц, подписей и жестов руки. Используя полученные оценки сложности и долей ошибок, для функции качество-сложность вида (2) построены оценки  $\tilde{\varepsilon}_{\psi_{\text{ES}}}(C)$ ,  $\tilde{\varepsilon}_{\psi_{\text{GS}}}(C)$ , демонстрирующие преимущество иерархического алгоритма  $\psi_{\text{GS}}$  по сравнению с переборным алгоритмом  $\psi_{\text{ES}}$ .

### Структуры данных и формализация задачи

Пусть  $\mathbf{A}^* = \mathbf{A}^L$  — множество представлений, в котором представление  $A^L \in \mathbf{A}^L$  объекта  $A \in \mathbf{A}$  задано преобразованием

$$\varphi(A) = A^L = (a^0, \dots, a^l, \dots, a^L) \quad (3)$$

Последовательность  $A^L$  в (3) является многоуровневым описанием объекта  $A$  в форме полного бинарного дерева глубины  $L$ , а подпоследовательность  $A^l = (a^0, \dots, a^l)$  образует поддереву глубины  $l$ , в котором  $l$ -й уровень  $a^l$  содержит  $2^l$  вершин [8]. Вершины с номерами  $n = 0, 1, \dots, 2^{L+1} - 2$  соответствуют эллиптическим примитивам

$$Q_n = (\mathbf{r}_n, \mathbf{u}_n, \mathbf{v}_n, z_n) \quad (4)$$

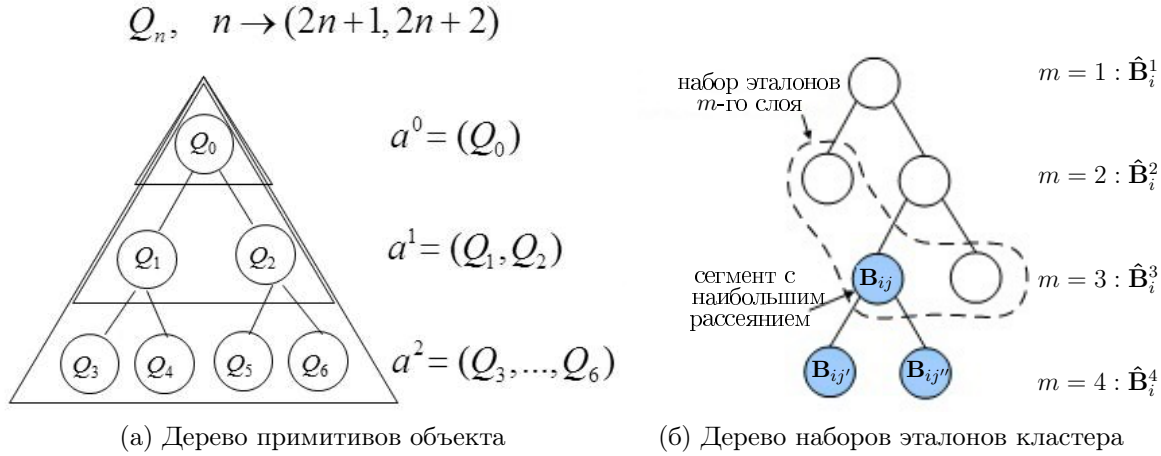
с вектором центра  $\mathbf{r}_n$ , векторами большой и малой полуосей  $\mathbf{u}_n, \mathbf{v}_n$  и уровнем яркости  $z_n$ .

Примитивы (4) вычисляются как эллипсы рассеяния соответствующих сегментов образа, которые формируются в результате его дихотомического разбиения. Яркости примитивов вычисляются как средние значения пикселей соответствующих сегментов. Пример структуры представления объекта в форме полного бинарного дерева глубины  $L = 2$  дан на рис. 3а.

Примитивы  $Q_n$ ,  $n = 0, 1, \dots, 2^{L+1} - 2$ , формируются в системе собственных координат объекта и нормируются относительно параметров корневого примитива с номером  $n = 0$ , что обеспечивает инвариантность представления (1) к повороту, смещению, масштабу и уровню яркости объекта при достаточно малых размерах пикселя и шага квантования яркости. Класс образов, допускающих такое представление, включает всевозможные объекты, заданные односвязными или многосвязными двумерными твердыми телами с полутоновой окраской, которые имеют однозначно идентифицируемую систему собственных координат [1].

На множестве представлений  $\mathbf{A}^L$  вида (3) вводится семейство мер различия объектов  $A \in \mathbf{A}$ ,  $\hat{A} \in \mathbf{A}$  с использованием пар поддеревьев  $A^l, \hat{A}^l$  глубины  $l = 1, \dots, L$ . Мера различия  $l$ -го порядка для пары объектов  $A$  и  $\hat{A}$  имеет вид:

$$d^l(A, \hat{A}) = d(A^l, \hat{A}^l) = \sum_{n=1}^{2^{l+1}-2} w_n \rho(Q_n, \hat{Q}_n), \quad (5)$$



**Рис. 3.** Примеры дерева представления объекта (а) и дерева наборов эталонов кластера (б)

где  $\rho(Q_n, \hat{Q}_n) \geq 0$  – функция различия соответственных (с одинаковыми номерами  $n$ ) примитивов  $Q_n \in A^l$  и  $\hat{Q}_n \in \hat{A}^l$ , заданных описаниями вида (4), а

$$w_n = \frac{[\log_2(n+1)] 2^{-\lfloor \log_2(n+1) \rfloor}}{\sum_{n=1}^{2^{l+1}-2} [\log_2(n+1)] 2^{-\lfloor \log_2(n+1) \rfloor}}$$

вес примитивов с номером  $n$  в поддеревьях  $A^l$  и  $\hat{A}^l$ .

Будем считать, что кластеры  $\mathbf{B}_i \subset \mathbf{A}_i, i = 1, \dots, c$ , обучающего множества содержат по  $M$  объектов. В каждом кластере формируются наборы эталонов

$$\hat{\mathbf{B}}_i^1, \dots, \hat{\mathbf{B}}_i^m, \dots, \hat{\mathbf{B}}_i^M, \quad (6)$$

где  $m = \|\hat{\mathbf{B}}_i^m\|$  – мощность набора  $\hat{\mathbf{B}}_i^m$ . Совокупность наборов (6) по всем кластерам образует ансамбль эталонов

$$\left\{ \hat{\mathbf{B}}_i^1, \dots, \hat{\mathbf{B}}_i^m, \dots, \hat{\mathbf{B}}_i^M \right\}_{i=1}^c. \quad (7)$$

При любом фиксированном  $m = 1, \dots, M$  множество  $\left\{ \hat{\mathbf{B}}_i^m \right\}_{i=1}^c$  из ансамбля (7) образует множество эталонов  $\hat{\mathbf{B}}^m$ , составленное из наборов мощности  $m$ .

Для каждого кластера  $\mathbf{B}_i, i = 1, \dots, c$ , наборы эталонов (6) строятся на множестве  $\mathbf{B}_i^L$  его древовидных представлений вида (3) ( $\varphi: \mathbf{B}_i \rightarrow \mathbf{B}_i^L$ ) с использованием дихотомического разбиения кластера по мере (5) наибольшего порядка  $L$  [1]. Пример структуры таких наборов с параметром  $M = 4$  дан в виде бинарного дерева на рис. 3б. Вершины в дереве соответствуют сегментам разбиения, на  $m$ -м шаге которого ( $m = 1, \dots, M$ ) кластер  $\mathbf{B}_i$  представлен набором сегментов  $\{\mathbf{B}_{ij}\}_{j=1}^m$  (на каждом шаге разбиения принята независимая нумерация сегментов). Набор  $\{\mathbf{B}_{ij}\}_{j=1}^m$  образует слой из  $m$  конечных вершин в дереве разбиения кластера  $\mathbf{B}_i$  на  $m$ -м шаге дихотомии. В каждом сегменте  $\mathbf{B}_{ij}$  выбирается эталон

$$\hat{B}_{ij} = \arg \min_{\hat{B} \in \mathbf{B}_{ij}} \max_{B \in \mathbf{B}_{ij}} d^L(B, \hat{B}) \quad (8)$$

относительно которого объекты этого сегмента имеют наименьшее рассеяние по мере  $d^L(B, \hat{B})$ . Эталоны, отобранные на  $m$ -м шаге дихотомии, образуют набор  $\hat{\mathbf{B}}_i^m = \left\{ \hat{B}_{ij} \right\}_{j=1}^m$ .

Набор эталонов  $\hat{\mathbf{B}}_i^{m+1}$  ( $m+1$ )-го шага строится путем замещения в наборе  $\hat{\mathbf{B}}_i^m$  одного из эталонов парой новых эталонов. Для этой цели среди сегментов  $\{\mathbf{B}_{ij}\}_{j=1}^m$   $m$ -го шага выбирается сегмент (вершина)  $\mathbf{B}_{ij}$  с наибольшим рассеянием  $\max_{B \in \mathbf{B}_{ij}} d^L(B, \hat{B}_{ij})$  относительно эталона  $\hat{B}_{ij} \in \mathbf{B}_{ij}$  и производится разбиение выбранного сегмента  $\mathbf{B}_{ij} \rightarrow (\mathbf{B}_{ij'}, \mathbf{B}_{ij''})$ , порождающее пару новых сегментов. Сегменты  $\mathbf{B}_{ij'}$  и  $\mathbf{B}_{ij''}$  формируются из объектов, ближайших к опорным объектам  $B_{ij'} \in \mathbf{B}_{ij}$  и  $B_{ij''} \in \mathbf{B}_{ij}$ , которые образуют наиболее удаленную друг от друга пару объектов в сегменте  $\mathbf{B}_{ij}$ . Пара сегментов  $(\mathbf{B}_{ij'}, \mathbf{B}_{ij''})$  дает пару новых эталонов, выбираемых согласно (8).

Используя меру (5) и наборы эталонов  $\hat{\mathbf{B}}_i^m$  из (6), вводятся разделяющие функции

$$g^l(A|\hat{\mathbf{B}}_i^m) = \max_{\hat{B} \in \hat{\mathbf{B}}_i^m} (1 - d^l(A, \hat{B})) = 1 - \min_{\hat{B} \in \hat{\mathbf{B}}_i^m} d^l(A, \hat{B}), \quad i = 1, \dots, c \quad (9)$$

порядка  $l = 1, \dots, L$ . Решение по объекту  $A \in \mathbf{A}$  строится на функциях (9) порядка  $L$ , вычисляемых по представлениям  $A^L, \hat{B}^L$  для объекта и эталонов, и дает номер класса

$$\psi(A) = \arg \max_{i=1}^c g^L(A|\hat{\mathbf{B}}_i^m). \quad (10)$$

Решение вида (10) может быть получено с помощью переборного алгоритма  $\psi_{\text{ES}}$  и с помощью иерархического алгоритма  $\psi_{\text{GS}}$ . В следующем разделе рассмотрено параметрическое семейство этих алгоритмов, приведены аналитические оценки их сложностей  $C_{\psi_{\text{ES}}}(L, m)$ ,  $C_{\psi_{\text{GS}}}(L, m)$  как функции глубины  $L$  представляющих деревьев и мощности  $m$  наборов эталонов, и экспериментальные оценки долей ошибок  $\varepsilon_{\psi_{\text{ES}}}(L, m)$ ,  $\varepsilon_{\psi_{\text{GS}}}(L, m)$ , полученные при выбранном значении  $L$  для различных значений  $m$ . Используя эти оценки, для алгоритмов  $\psi_{\text{ES}}$  и  $\psi_{\text{GS}}$  построены оценки функции качество-сложность (2) в виде зависимостей  $\tilde{\varepsilon}_{\psi_{\text{ES}}}(C)$  и  $\tilde{\varepsilon}_{\psi_{\text{GS}}}(C)$ .

## Многоуровневая сеть эталонов и решающие алгоритмы

Рассмотрим ансамбль эталонов  $\hat{\mathbf{B}} = \left\{ \hat{\mathbf{B}}_i^1, \dots, \hat{\mathbf{B}}_i^m, \dots, \hat{\mathbf{B}}_i^M \right\}_{i=1}^c$  вида (7) и ансамбль их древовидных представлений  $\hat{\mathbf{B}}^L = \left\{ \hat{\mathbf{B}}_i^{1L}, \dots, \hat{\mathbf{B}}_i^{mL}, \dots, \hat{\mathbf{B}}_i^{ML} \right\}_{i=1}^c$ , полученный преобразованием  $\varphi : \hat{\mathbf{B}} \rightarrow \hat{\mathbf{B}}^L$  вида (3). Последовательность представлений ансамбля эталонов  $\hat{\mathbf{B}}$  поддеревьями примитивов глубины  $l = 1, \dots, L$  образует многоуровневую сеть

$$\hat{\mathbf{B}}^1, \dots, \hat{\mathbf{B}}^l, \dots, \hat{\mathbf{B}}^L, \quad (11)$$

в которой  $l$ -й уровень задан множеством представлений  $\hat{\mathbf{B}}^l = \left\{ \hat{\mathbf{B}}_i^{1l}, \dots, \hat{\mathbf{B}}_i^{ml}, \dots, \hat{\mathbf{B}}_i^{Ml} \right\}_{i=1}^c$ . В данной работе для поиска решения (10) используются подмножества представлений сети (11), которые образованы наборами эталонов фиксированной мощности  $m = 1, \dots, M$  для всех классов с номерами  $i = 1, \dots, c$ . В этом случае на последовательных уровнях сети (11) обрабатываются подмножества представлений эталонов

$$\hat{\mathbf{B}}^{ml} = \left\{ \hat{\mathbf{B}}_i^{ml} \right\}_{i=1}^c, \quad l = 1, \dots, L. \quad (12)$$

Для описания решающих алгоритмов на подмножествах (12) вводится параметрическая стратегия сужения зоны поиска решения, которая задается функцией

$$c_l = \lfloor c 2^{-\alpha(l-1)} \rfloor, \quad l = 1, \dots, L \quad (13)$$

с параметром  $\alpha = (L - 1)^{-1} \log_2(c/c^*) \geq 0$ , где  $c^* = 1, \dots, c$  – фиксированное целое [1]. Согласно (13), решающий алгоритм на  $l$ -м уровне вычисляет  $c_l$  разделяющих функций (9) и отбирает среди них  $c_{l+1}$  функций с наибольшими значениями, которые соответствуют номерам классов, для которых вычисляются разделяющие функции на  $(l + 1)$ -м уровне. На последнем  $L$ -м уровне принимается решение (10) по  $c_L = c^*$  разделяющим функциям, которые вычисляются для классов, отобранных на  $(L - 1)$ -м уровне. Следует отметить, что с ростом  $c^*$  показатель  $\alpha$  экспоненты (13) уменьшается и в случае  $c^* = c$  значение  $\alpha = 0$  исключает сужение зоны поиска на последовательных уровнях в (12).

В случае  $c^* < c$  стратегия (13) соответствует иерархическому алгоритму  $\psi_{GS}$ , в случае  $c^* = c$  рассмотренная стратегия эквивалентна переборному алгоритму  $\psi_{ES}$ . В силу аддитивности меры (5), оба алгоритма на каждом  $l$ -м уровне вычисляют значения  $d^l(A, \hat{B})$  в (9) для классов, отобранных на  $(l - 1)$ -м уровне, путем суммирования значений  $d^{l-1}(A, \hat{B})$ , вычисленных на  $(l - 1)$ -м уровне, и значений  $w_n \rho(Q_n, \hat{Q}_n)$ , вычисляемых для соответственных примитивов  $Q_n \in A^l$ ,  $\hat{Q}_n \in \hat{B}^l$  в представлениях  $l$ -го уровня. В результате, для вычисления разделяющих функций на парах  $(A, \hat{B}_i^m)$ ,  $i = 1, \dots, c$ , переборный алгоритм  $\psi_{ES}$  обрабатывает все пары соответственных примитивов в многоуровневых представлениях  $A^L$  и  $\hat{B}^{mL} = \left\{ \hat{B}_i^{mL} \right\}_{i=1}^c$ , а число пар соответственных примитивов, обрабатываемых иерархическим алгоритмом  $\psi_{GS}$  существенно меньше.

Полученные на представлениях вида (3) оценки вычислительной сложности алгоритмов  $\psi_{ES}$  и  $\psi_{GS}$  в терминах числа обрабатываемых пар примитивов на объект приводятся в следующем утверждении.

**Утверждение 1.** При значениях  $c \geq 2$ ,  $L > 1$ , и  $m = 1, \dots, M$  вычислительные сложности решающих алгоритмов  $\psi_{ES}$  и  $\psi_{GS}$  удовлетворяют оценкам

$$C_{\psi_{ES}}(L, m) = cm \sum_{l=1}^L 2^l = 2cm(2^L - 1); \quad (14)$$

$$C_{\psi_{GS}}(L, m) = m \sum_{l=1}^L c_l 2^l \leq 2cmL, \quad (15)$$

где неравенство в (15) справедливо при  $\alpha \geq 1$ .

Оценки (14) и (15) базируются на вычислении числа вершин в деревьях, представляющих эталоны, которые обрабатываются соответствующими решающими алгоритмами.

В алгоритме  $\psi_{ES}$  — это число вершин в представляющих деревьях, которые составляют подмножество  $L$ -го уровня в (12). Число вершин на  $l$ -м уровне представляющих деревьев по каждому классу равно  $m2^l$ , а суммирование этих чисел по  $l = 1, \dots, L$  и по  $c$  классам дает оценку (14).

В алгоритме  $\psi_{GS}$  используются подмножества всех уровней в (12). С учетом стратегии (13) число вершин, обрабатываемых на  $l$ -м уровне сети эталонов, равно  $mc_l 2^l$ . Суммирование этих чисел по  $l = 1, \dots, L$  с учетом  $\alpha \geq 1$  дает оценку (15).

**Следствие.** Согласно оценкам (14) и (15) вычислительный выигрыш иерархического алгоритма  $\psi_{GS}$  с параметром  $\alpha \geq 1$  относительно переборного алгоритма  $\psi_{ES}$  определяется глубиной представляющих деревьев и удовлетворяет оценке

$$\frac{C_{\psi_{ES}}}{C_{\psi_{GS}}} \geq \frac{2^L - 1}{L}.$$

## Экспериментальные оценки функции качество-сложность

В этом разделе приводятся экспериментальные оценки долей ошибок  $\varepsilon(L, m, c^*)$  и численные оценки вычислительной сложности  $C(L, m, c^*)$ , полученные при глубине представляющих деревьев  $L = 11$ , числе эталонов в классах  $1 \leq m \leq 20$  (соответствует числу конечных вершин в деревьях наборов эталонов) и значениях параметра решающего алгоритма  $1 \leq c^* \leq c$ , где  $c$  — число классов анализируемого источника. В случае  $c^* = c$  эти оценки соответствуют характеристикам переборного алгоритма  $\psi_{ES}$ , в случае  $c^* < c$  — характеристикам иерархического алгоритма  $\psi_{GS}$ . Максимальное значение  $m$  ограничено размером кластеров обучающей выборки. Глубина представлений  $L = 11$  обусловлена отсутствием уменьшения доли ошибок распознавания при  $L > 11$  для анализируемого источника объектов.

Эксперименты проводились с составным источником полутоновых объектов (8 бит/пиксель), который включает изображения подписей от 40 персон, изображения жестов руки от 25 персон и изображения лиц от 25 персон, по 40 изображений от каждой персоны. Общее число классов определяется числом персон и равно  $c = 90$ , общее число объектов определяется числом изображений и равно 3600.

Разделяющие функции (9) по мере (5) реализованы с использованием функции различия примитивов в форме среднего значения

$$\rho(Q_n, \hat{Q}_n) = \frac{1}{6} \sum_{k=1}^6 \rho_k(Q_n, \hat{Q}_n),$$

с компонентами

$$\rho_1(Q_n, \hat{Q}_n) = |z_n - \hat{z}_n|;$$

$$\rho_2(Q_n, \hat{Q}_n) = |r_n^u - \hat{r}_n^u|;$$

$$\rho_3(Q_n, \hat{Q}_n) = |r_n^v - \hat{r}_n^v|;$$

$$\rho_4(Q_n, \hat{Q}_n) = \frac{|u_n - \hat{u}_n|}{\max(u_n, \hat{u}_n)};$$

$$\rho_5(Q_n, \hat{Q}_n) = \frac{|v_n - \hat{v}_n|}{\max(u_n, \hat{u}_n)};$$

$$\rho_6(Q_n, \hat{Q}_n) = \left(1 - \frac{|\mathbf{u}_n \hat{\mathbf{u}}_n|}{u_n \hat{u}_n}\right) \frac{\min(u_n - v_n, \hat{u}_n - \hat{v}_n)}{\max(u_n, \hat{u}_n)},$$

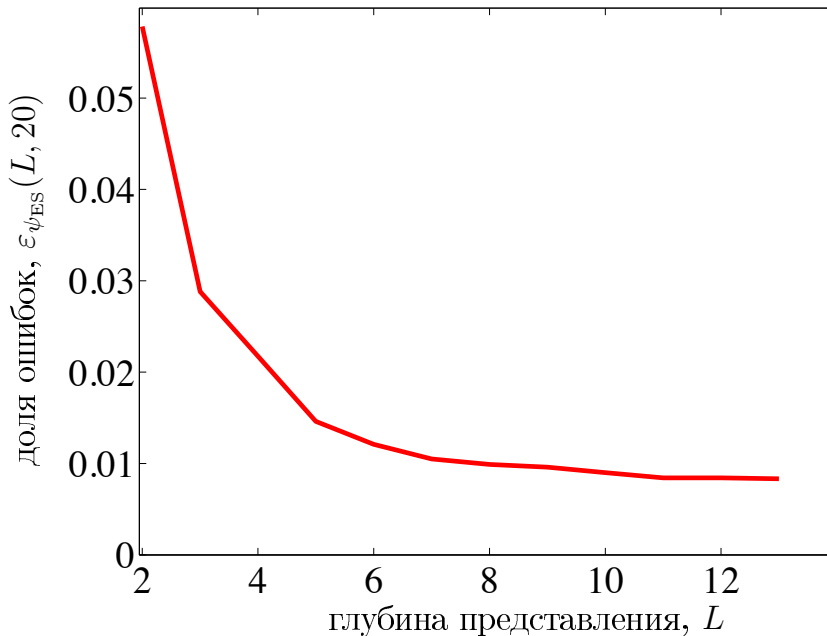
где  $(r_n^u, r_n^v)$  и  $(\hat{r}_n^u, \hat{r}_n^v)$  — проекции векторов центров  $\mathbf{r}_n$  и  $\hat{\mathbf{r}}_n$  примитивов на главные оси корневых примитивов  $Q_0$  и  $\hat{Q}_0$ ;  $u_n = \|\mathbf{u}_n\|$ ,  $v_n = \|\mathbf{v}_n\|$  и  $\hat{u}_n = \|\hat{\mathbf{u}}_n\|$ ,  $\hat{v}_n = \|\hat{\mathbf{v}}_n\|$  — эвклидовы нормы векторов полуосей;  $\mathbf{u}_n \hat{\mathbf{u}}_n$  — скалярное произведение векторов полуосей примитивов  $Q_n$  и  $\hat{Q}_n$  в их описаниях вида (4). Компонента  $\rho_6(Q_n, \hat{Q}_n)$  учитывает различие ориентаций векторов  $\mathbf{u}_n$ ,  $\hat{\mathbf{u}}_n$  и убывает, когда, по крайней мере, один их пары  $Q_n, \hat{Q}_n$  эллиптических примитивов приближается к сфере ( $u_n = v_n$  или/и  $\hat{u}_n = \hat{v}_n$ ).

Для решающих алгоритмов  $\psi_{ES}$  и  $\psi_{GS}$  оценки сложностей  $C_{\psi_{ES}}(L, m) = C(L, m, c^* = c)$  и  $C_{\psi_{GS}}(L, m) = C(L, m, c^* < c)$  получены из соотношений (14) и (15) путем прямого вычисления суммы в (15). Оценки долей ошибок  $\varepsilon_{\psi_{ES}}(L, m) = \varepsilon(L, m, c^* = c)$  и  $\varepsilon_{\psi_{GS}}(L, m) = \varepsilon(L, m, c^* < c)$  этих алгоритмов получены методом кросс-тестирования по схеме 200 кратного разбиения множества изображений источника на 2 части (200 times 2 fold cross-validation) [9]. На каждом разбиении с помощью случайного отбора половины изображений



каждого класса двукратно формировались обучающая и тестовая выборки одинакового объема, которые содержали по 20 изображений (объектов) от каждого класса. Текущая обучающая выборка использовалась для построения ансамбля эталонов вида (7) с параметром  $M = 20$  в пространстве древовидных представлений с числом уровней  $L = 11$  и для формирования на этом ансамбле многоуровневой сети вида (11). Текущая тестовая выборка использовалась для вычисления долей ошибок  $\varepsilon(L, m, c^*)$  при значениях  $L = 11$ ,  $m = 1, 5, 10, 15, 16, 17, 18, 19, 20$  и  $c^* = 5, 10, 15, 20, 90$ . Оценки  $\varepsilon_{\psi_{ES}}(L, m)$  и  $\varepsilon_{\psi_{GS}}(L, m)$  вычислялись усреднением соответствующих долей ошибок  $\varepsilon(L, m, c^* = 90)$  и  $\varepsilon(L, m, c^* < 90)$  по всем тестовым выборкам.

Выбор глубины представляющих деревьев для используемого источника изображений иллюстрируется поведением зависимости  $\varepsilon_{\psi_{ES}}(L, 20)$  от  $L$ , полученной при значениях  $m = 20$ ,  $c^* = 90$ . График этой зависимости дан на рис. 4 и демонстрирует отсутствие уменьшения доли ошибок при  $L > 11$  для анализируемого источника объектов.



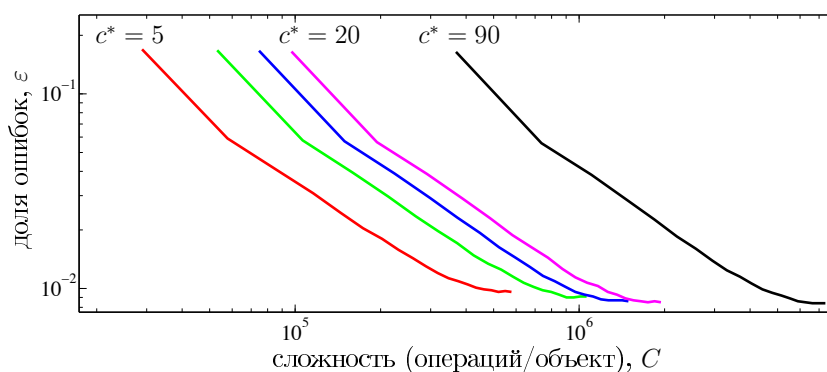
**Рис. 4.** Зависимость доли ошибок  $\varepsilon_{\psi_{ES}}(L, 20)$  от глубины представляющих деревьев  $L$

Полученные при  $L = 11$  средние доли ошибок и значения вычислительной сложности рассмотренных решающих алгоритмов приведены в таблице 1. Столбцы с параметром  $c^* = 5, 10, 15, 20$  содержат усредненные доли ошибок  $\varepsilon$  и оценки сложности  $C$  иерархического алгоритма  $\psi_{GS}$ ; столбцы с параметром  $c^* = 90$  — аналогичные характеристики переборного алгоритма  $\psi_{ES}$ . Построенное по этим оценкам параметрическое семейство функций  $\varepsilon_{\psi_{GS}}(C)$  (с параметром  $c^* = 5, 10, 15, 20$ ) и  $\varepsilon_{\psi_{ES}}(C)$  (с параметром  $c^* = 90$ ) в терминах зависимости доли ошибок от числа операций на объект даны на рис. 5.

Для GS и ES алгоритмов выявлены мощности наборов эталонов, доставляющие наименьшие доли ошибок, которые достигаются при значениях  $m < 20$ . Минимумы долей ошибок, достигаемые на переборном ES алгоритме с параметром  $c^* = 90$  и иерархических GS алгоритмах с параметрами  $c^* = 20$  и  $c^* = 5$  составляют 0,0084, 0,0085 и 0,0096 соответственно. При указанных значениях  $c^*$  вычислительный выигрыш GS алгоритмов по сравнению с ES алгоритмом составляет 3,8 и 12,8. В общем случае вычислительные

**Таблица 1.** Оценки вычислительной сложности и средние доли ошибок для GS алгоритмов ( $c^* = 5, 10, 15, 20$ ) и ES алгоритма ( $c^* = 90$ ) при глубине представляющих деревьев  $L = 11$ 

$m$	$c^* = 5$		$c^* = 10$		$c^* = 15$		$c^* = 20$		$c^* = 90$	
	$\varepsilon$	$C$	$\varepsilon$	$C$	$\varepsilon$	$C$	$\varepsilon$	$C$	$\varepsilon$	$C$
1	0,1695	28896	0,1675	53148	0,1667	74608	0,1651	97064	0,1643	368460
5	0,0245	144480	0,0235	265740	0,0231	373040	0,0230	485320	0,0226	1842300
10	0,0130	288960	0,0125	531480	0,0116	746080	0,0114	970640	0,0115	3684600
15	0,0101	433440	0,0096	797220	0,0091	1119120	0,0089	1455960	0,0089	5526900
16	0,0099	462336	0,0093	850368	0,0088	1193728	0,0087	1553024	0,0086	5895360
17	0,0097	491232	0,0090	903516	0,0088	1268336	0,0086	1650088	0,0085	6263820
18	0,0096	520128	0,0090	956664	0,0087	1342944	0,0085	1747152	0,0084	6632280
19	0,0097	549024	0,0091	1009812	0,0087	1417552	0,0086	1844216	0,0084	7000740
20	0,0097	577920	0,0091	1062960	0,0087	1492160	0,0085	1941280	0,0084	7369200

**Рис. 5.** Семейство зависимостей средней доли ошибок от вычислительной сложности для GS и ES алгоритмов при глубине представляющих деревьев  $L = 11$ 

сложности и доли ошибок указанных алгоритмов удовлетворяют соотношениям

$$C_{\psi_{GS}}(L, m) \leq C_{\psi_{ES}}(L, m)$$

$$\varepsilon_{\psi_{GS}}(L, m) \geq \varepsilon_{\psi_{ES}}(L, m).$$

Уменьшение доли ошибок при заданной вычислительной сложности решающего алгоритма может быть достигнуто за счет оптимизации мощностей  $m_1, \dots, m_c$  наборов эталонов по классам, выполняемой на этапе обучения. Оптимизация может быть выполнена с использованием процедуры скользящего контроля по схеме leave-one-out [9] и в случае фиксированных параметров  $L$  и  $c^*$  сводится к минимизации доли ошибок  $\varepsilon^{\text{LOO}}(L, m_1, \dots, m_c, c^*)$  по наборам  $m_1, \dots, m_c$  при ограничении на сложность решающего алгоритма  $C(L, m_1, \dots, m_c, c^*) < C^*$ , где  $C^*$  — заданное допустимое значение сложности. Такая оптимизация позволит найти более точные оценки  $\tilde{\varepsilon}_{\psi_{ES}}(C^*)$ ,  $\tilde{\varepsilon}_{\psi_{GS}}(C^*)$  функции качество-сложность для рассмотренного параметрического семейства решающих алгоритмов.

## Заключение

Сформулирована задача многоклассовой метрической классификации в терминах минимизации вероятности ошибки при заданной вычислительной сложности решающего алгоритма. Постановка задачи предполагает использование схемы классификации, которая сходна с известной схемой кодирования источника при наличии в ней канала наблюдения, вносящего искажения. Для объектов, заданных изображениями, предложен подход

к построению эффективного классификатора в пространстве древовидных представлений объектов с многоуровневым разрешением. Используемое представление эквивалентно каналу наблюдения, искажения которого влияют на вероятность ошибки классификатора. Исследованы два типа решающих алгоритмов по критерию ближайшего эталона, на основе полного перебора эталонов и на основе иерархического поиска решения в многоуровневой сети эталонов. Для указанных алгоритмов получены аналитические оценки вычислительной сложности и экспериментальные доли ошибок распознавания объектов составного источника, заданного полутонными изображениями лиц, жестов руки и подписей. Используя указанные оценки сложности и качества, построено параметрическое семейство зависимостей доли ошибок от вычислительной сложности решающих алгоритмов. При сопоставимых долях ошибок продемонстрирован выигрыш в быстродействии иерархического решающего алгоритма по сравнению с переборным алгоритмом. В рамках используемой схемы отбора эталонов и параметрического семейства решающих алгоритмов намечен путь оптимизации классификатора по числу отбираемых эталонов в классах. Предложенный подход может найти применение для распознавания в режиме реального времени биометрических образов, когда число классов достаточно велико. Планируется исследовать решающие алгоритмы, построенные на разделяющих функциях более общего вида, и допускающие отказы от распознавания.

## Литература

- [1] *Lange M. M., Ganebnykh S. N.* Hierarchical data structures and decision algorithms for efficient image classification // Machine Learning Data Analysis, 2013. Vol. 1. No. 6. P. 725–733. <http://jmla.org/papers/doc/2013/no6/Lange2013ImageClassification.pdf>.
- [2] *Duda R. O., Hart P. E., Stork D. G.* Pattern classification. — 2nd ed. — Wiley, 2001. 672 p.
- [3] *Dobrushin R. L., Tsybakov B. S.* Information transmission with additional noise // IRE Trans. Inform. Theory. 1962. Vol. 8. No. 5. Pp. 293–304.
- [4] *Kolmogorov A. N., Tikhomirov V. M.*  $\varepsilon$ -Entropy and  $\varepsilon$ -capacity of sets in functional spaces // AMS Transl., 1961. Vol. 17. No. 2. Pp. 277–364.
- [5] *Rosenfeld A.* Quadrees and pyramids for pattern recognition and image analysis // 5th Conference (International) on Pattern Recognition Proceedings, 1980. Pp. 802–811.
- [6] *Torsello A.* Matching hierarchical structures for shape recognition. PhD Thesis. — University of York, 2004. 197 p.
- [7] *Elfiky N. M., Khan F. S., Weijer J., Gonzalez J.* Discriminative compact pyramids for object and scene recognition // Pattern Recognition, 2012. Vol. 45. No. 4. Pp. 1627–1636.
- [8] *Ganebnykh S. N., Lange M. M.* Metric classifier using multilevel network of templates // Pattern Recognition Image Analysis, 2012. Vol. 22. No. 2. Pp. 265–277.
- [9] *Bishop C. M.* Pattern recognition and machine learning. — Springer, 2006. 738 p.