

Схемы коллективных решений для классификации с отказом на ансамбле источников изображений

М.М. Ланге¹, С.Н. Ганебных¹

¹Федеральный исследовательский центр "Информатика и управление" РАН, ул. Вавилова, 40, Москва, Россия, 119333

Аннотация. Исследуются две схемы построения метрических многоклассовых классификаторов с отказом на ансамбле изображений от источников различной модальности. Классификаторы принимают коллективные решения по составным объектам, заданным наборами изображений, по одному от каждого источника. Разделяющие функции порождаются элементарными классификаторами типа класс против всех. Рассматривается известная схема на основе взвешенного голосования решений по объектам индивидуальных источников (Weighted Majority Vote) и оригинальная схема на основе обобщенной меры различия составных объектов в виде взвешенной суммы мер различия соответствующих компонент составных объектов (Weighted General Measure). Приводятся экспериментальные оценки вероятностей ошибок классификации, которые демонстрируют преимущество схемы WGM по сравнению со схемой WMV.

1. Введение

Один из фундаментальных результатов теории классификации заключается в уменьшении вероятности ошибок путем применения композиции алгоритмов, при условии, что вероятность ошибок каждого алгоритма меньше вероятности ошибок случайного выбора решения. Общие основы теории композиции алгоритмов сформулированы в работе [1] и подробно изложены в лекциях по алгоритмическим композициям [2]. Идея композиции алгоритмов лежит в основе процедуры последовательного обучения зависимых классификаторов методом адаптивного бустинга [3]. Композиции классификаторов используются также для распознавания образов от ансамбля источников различной модальности [4]. Ансамбль порождает составные объекты, заданные наборами объектов одного класса, по одному от каждого источника. Построение коллективных решений по предъявляемым составным объектам базируется на использовании композиций данных, которые порождают различные схемы комплексирования. Широкое распространение получила схема комплексирования решений независимых классификаторов по объектам источников на основе взвешенного голосования по большинству (Weighted Majority Vote) [5,6]. В схеме WMV используются композиции решений классификаторов по объектам индивидуальных источников в форме взвешенных сумм голосов классификаторов по классам, и коллективное решение по составному объекту принимается в пользу класса, получившего наибольший суммарный голос. В работах [7,8] предложена альтернативная схема комплексирования, в которой коллективные решения по составным объектам принимаются единым классификатором с решающим правилом на основе взвешенной обобщенной меры

(Weighted General Measure) различия составных объектов. В схеме WGM используются композиции объектов источников в форме взвешенных сумм различий компонент составного объекта с объектами обучающей выборки. Такие композиции порождают на множестве составных объектов разделяющие функции по классам, которые дают решающее правило в пользу класса с наибольшим значением разделяющей функции.

Применяя теоретико-информационную характеристику классификатора в виде средней взаимной информации между множеством объектов и множеством классов, для ансамбля источников с весами, зависящими от соответствующих значений средней взаимной информации, доказано, что в случае классификации без отказов средняя взаимная информация в схеме WGM превосходит среднюю взаимную информацию в схеме WMV [8]. Следствием этого результата является возможность достижения меньшей вероятности ошибок в схеме WGM по сравнению со схемой WMV.

Цель настоящей работы заключается в исследовании сравнительной эффективности схем комплексирования WGM и WMV в случае многоклассового распознавания составных объектов с отказом. Исследуемая многоклассовая модель сопоставима с моделью, предложенной в работе [9] для распознавания объектов отдельных источников. Отличительной особенностью исследуемой модели классификации является возможность распознавания составных объектов, порождаемых ансамблем источников. Эффективность исследуемой модели оценивается в терминах однопараметрической ROC кривой (receiver operating characteristic) [13], параметр которой определяет пороги отказа по классам.

Рассматриваются классификаторы, образованные наборами элементарных NN и SVM классификаторов типа "класс против всех", в которых разделяющие функции вычисляются соответственно по ближайшим соседям в классах [10] и по алгебраическим расстояниям до разделяющих гиперплоскостей [11]. Предлагается процедура обучения элементарных классификаторов и приводятся результаты тестирования многоклассовых классификаторов на цветных изображениях лиц [12], которые образуют трехкомпонентный ансамбль декоррелированных компонент RGB изображений. Результатами тестирования являются доли ошибок классификаторов с параметрами, оптимизированными на этапе обучения, и ROC кривые с соответствующими значениями AUC (area under curve). Указанные характеристики демонстрируют более высокое качество распознавания для схемы WGM по сравнению со схемой WMV.

2. Формализация задачи исследования

Пусть $\mathbf{X}^M = (\mathbf{X}_1, \dots, \mathbf{X}_M)$ – ансамбль источников, в котором $\mathbf{X}_m, m=1, \dots, M$ – множество объектов m -го источника, принадлежащих классам из множества $\Omega = \{\omega_i, i = 0, 1, \dots, c\}, c \geq 1$, с априорными вероятностями $P(\omega_i) > 0: \sum_{i=0}^c P(\omega_i) = 1$. Любая последовательность $\mathbf{x}^M = (\mathbf{x}_1, \dots, \mathbf{x}_M)$ объектов $\mathbf{x}_m \in \mathbf{X}_m, m = 1, \dots, M$, принадлежащих одному и тому же классу, образует составной объект $\mathbf{x}^M \in \mathbf{X}^M$, класс которого $\omega_i \in \Omega$ необходимо найти. Как правило, ансамбль порождается наборами объектов различной модальности, например, биометрическими изображениями лица, отпечатка пальца, радужки, ладони, подписи и т.п., полученными от персон, которые образуют множество классов.

Классы подмножества $\Omega_c = \{\omega_i, i = 1, \dots, c\} \subset \Omega$ включают свои объекты, а класс ω_0 включает все прочие чужие объекты. Априорные вероятности классов дают вероятности $P(\Omega_c) = \sum_{i=1}^c P(\omega_i) = \theta$ и $P(\omega_0) = 1 - \theta$ своих и чужих объектов, в которых θ – случайный параметр, принимающий значения на отрезке $[0, 1]$ с единичной плотностью $p(\theta) = 1$. Предполагается, что известно априорное распределение классов на подмножестве $\Omega_c: \{P(\omega_i | \Omega_c) = P(\omega_i) / \sum_{i=1}^c P(\omega_i), i = 1, \dots, c\}$.

Будем считать, что на множестве объектов каждого источника для любой пары объектов

$\mathbf{x}_m \in \mathbf{X}_m$ и $\hat{\mathbf{x}}_m \in \mathbf{X}_m$ определена мера различия

$$d(\mathbf{x}_m, \hat{\mathbf{x}}_m) \geq 0, m = 1, \dots, M \quad (1)$$

и на ансамбле для любой пары составных объектов $\mathbf{x}^M \in \mathbf{X}^M$ и $\hat{\mathbf{x}}^M \in \mathbf{X}^M$ введена обобщенная мера различия в форме линейной композиции

$$D(\mathbf{x}^M, \hat{\mathbf{x}}^M) = \sum_{m=1}^M w_m d(\mathbf{x}_m, \hat{\mathbf{x}}_m) \quad (2)$$

с весами источников

$$W = \{w_m > 0, m = 1, \dots, M\} \quad (3)$$

На множествах объектов источников $\mathbf{X}_m, m = 1, \dots, M$, вводятся разделяющие функции

$$g_i(\mathbf{x}_m) = P(\omega_i | \Omega_c) p(\mathbf{x}_m | \omega_i), i = 1, \dots, c \quad (4)$$

а на ансамбле $\mathbf{X}^M$ – разделяющие функции

$$g_i(\mathbf{x}^M) = P(\omega_i | \Omega_c) p(\mathbf{x}^M | \omega_i), i = 1, \dots, c \quad (5)$$

по классам подмножества Ω_c , где $p(\mathbf{x}_m | \omega_i)$ и $p(\mathbf{x}^M | \omega_i)$ – условные по классу $\omega_i \in \Omega_c$ плотности вероятностей объектов $\mathbf{x}_m \in \mathbf{X}_m$ и $\mathbf{x}^M \in \mathbf{X}^M$, оценки которых строятся соответственно с использованием мер различия (1) и (2). Согласно формуле Байеса, разделяющие функции вида (4) и (5) пропорциональны соответствующим апостериорным вероятностям $P(\omega_i | \mathbf{x}_m, \Omega_c)$ и $P(\omega_i | \mathbf{x}^M, \Omega_c)$ и, следовательно, позволяют минимизировать вероятность ошибки классификации [11].

Оценки плотностей $p(\mathbf{x}_m | \omega_i)$ и $p(\mathbf{x}^M | \omega_i)$ в (4) и (5) строятся на этапе обучения по откликам элементарных классификаторов типа ”класс ω_i против всех $\Omega_c \setminus \omega_i$ ”. Отклики формируются в виде неотрицательных функций сходства объектов $\mathbf{x}_m$ или $\mathbf{x}^M$ с классом ω_i , которые вычисляются соответственно с использованием мер различия вида (1) или (2) и принимают значения на отрезке $[0,1]$. Предполагается, что вычисляемые по этим мерам функции сходства

$$s_i^d(\mathbf{x}_m) \geq 0, \mathbf{x}_m \in \mathbf{X}_m, m = 1, \dots, M \quad (6)$$

$$s_i^D(\mathbf{x}^M) \geq 0, \mathbf{x}^M \in \mathbf{X}^M \quad (7)$$

с точностью до мультипликативных констант дают оценки условных по классам плотностей

$$p(\mathbf{x}_m | \omega_i) \propto s_i^d(\mathbf{x}_m),$$

$$p(\mathbf{x}^M | \omega_i) \propto s_i^D(\mathbf{x}^M),$$

которые используются в разделяющих функциях (4) и (5).

При любом заданном множестве объектов $\mathbf{X}$, наборы разделяющих функций $\{g_i(\mathbf{x}), i = 1, \dots, c\}$ и их пороговых значений $\{\Delta_i \geq 0, i = 1, \dots, c\}$ определяют преобразование классификатора $F: \mathbf{X} \rightarrow \Omega$ с решающим правилом

$$i_F(\mathbf{x}) = \max_{i=1}^c [g_i(\mathbf{x}) \geq \Delta_i] \arg \max_{i=1}^c (g_i(\mathbf{x}) [g_i(\mathbf{x}) \geq \Delta_i]) \quad (8)$$

для номера класса объекта $\mathbf{x} \in \mathbf{X}$, где $[f]$ – индикатор f . Пороговые значения в (8) определяют условия отказа по классам, согласно которым решение соответствует классу из подмножества Ω_c при единичном значении максимума индикаторов и классу ω_0 при нулевом значении указанного максимума. Эффективность классификатора с решающим правилом (8) оценивается средней долей ошибок по классам

$$\begin{aligned}\varepsilon_F(\mathbf{X}, \Omega) &= \sum_{i=0}^c P(\omega_i) \varepsilon_F(\mathbf{X}, \omega_i) \\ &= \theta \sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon_F(\mathbf{X}, \omega_i) + (1-\theta) \varepsilon_F(\mathbf{X}, \omega_0)\end{aligned}\quad (9)$$

где $\varepsilon_F(\mathbf{X}, \omega_i)$ – доля ошибочных решений на преобразовании F по объектам выборки $\mathbf{X}$, принадлежащим классу $\omega_i \in \Omega$. Составляющие правой части в (9) дают доли ошибок по своим и чужим объектам

$$\begin{aligned}\sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon_F(\mathbf{X}, \omega_i) &= 1 - \text{TPR} \\ \varepsilon_F(\mathbf{X}, \omega_0) &= \text{FPR}\end{aligned}\quad (10)$$

где TPR (true positive rate) и FPR (false positive rate) – доли правильных и ложных решений соответственно по своим и чужим объектам. Вариация пороговыми значениями в (8) позволяет построить ROC кривую (receiver operating characteristic) в виде зависимости TPR от FPR [10].

Использование в решающем правиле (8) разделяющих функций (4) с откликами $s_i^d(\mathbf{x}_m)$, $i=1, \dots, M$, порождает набор классификаторов, выполняющих преобразования $F_m^d: \mathbf{X}_m \rightarrow \Omega$, $m=1, \dots, M$, по объектам источников ансамбля. Комплексирование решений таких классификаторов по схеме WMV с использованием весовых коэффициентов (3) дает для предъявляемого составного объекта $\mathbf{x}^M \in \mathbf{X}^M$ коллективное решение в виде номера класса

$$i_{\text{WMV}}(\mathbf{x}^M) = \arg \max_{i=0}^c \sum_{m=1}^M w_m [F_m^d(\mathbf{x}_m) = \omega_i] \quad (11)$$

где $[z]$ – индикатор z . Применение в решающем правиле (8) разделяющих функций (5) с откликами $s_i^D(\mathbf{x}^M)$ порождает классификатор $F_W^D: \mathbf{X}^M \rightarrow \Omega$, преобразование которого зависит от набора весов (3) и дает коллективное решение для объекта $\mathbf{x}^M \in \mathbf{X}^M$ по схеме WGM

$$i_{\text{WGM}}(\mathbf{x}^M) = \arg \max_{i=0}^c [F_W^D(\mathbf{x}^M) = \omega_i] \quad (12)$$

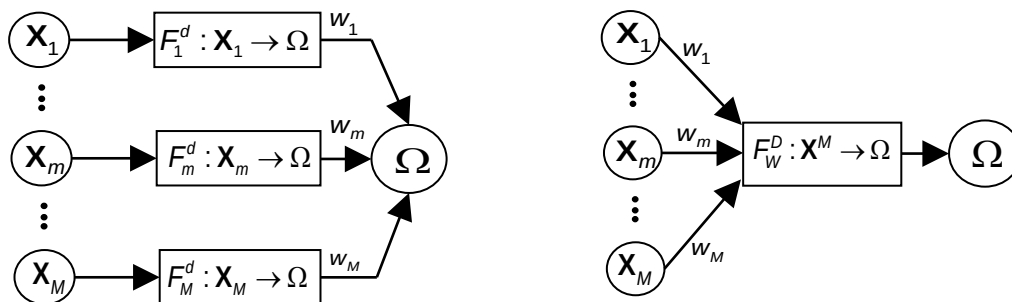


Рисунок 1. Блок-схемы классификаторов на основе схем комплексирования WMV (слева) и WGM (справа).

Схематически классификаторы, формирующие коллективные решения вида (11) и (12), даны на рисунке 1.

Согласно [11], для средней взаимной информации справедливо соотношение

$$I(\mathbf{X}^M; \Omega) = H(\Omega) - H(\Omega | \mathbf{X}^M),$$

где энтропия и условная энтропия определяются функционалами

$$\begin{aligned}
H(\Omega) &= -\sum_{i=0}^c P(\omega_i) \log P(\omega_i) \\
&= -\theta \sum_{i=1}^c P(\omega_i | \Omega_c) \log P(\omega_i | \Omega_c) - (1-\theta) \log(1-\theta) - \theta \log \theta \\
&= \theta H(\Omega_c) + h(\theta)
\end{aligned} \tag{13}$$

$$\begin{aligned}
H(\Omega | \mathbf{X}^M) &= -\int_{\mathbf{X}^M} p(\mathbf{x}^M) d\mathbf{x}^M \sum_{i=0}^c P(\omega_i | \mathbf{x}^M) \log P(\omega_i | \mathbf{x}^M) = \\
&= \int_{\mathbf{X}^M} p(\mathbf{x}^M) d\mathbf{x}^M \left(-\theta_{\mathbf{x}^M} \sum_{i=1}^c P(\omega_i | \mathbf{x}^M, \Omega_c) \log P(\omega_i | \mathbf{x}^M, \Omega_c) - (1-\theta_{\mathbf{x}^M}) \log(1-\theta_{\mathbf{x}^M}) - \theta_{\mathbf{x}^M} \log \theta_{\mathbf{x}^M} \right) \\
&= \int_{\mathbf{X}^M} p(\mathbf{x}^M) d\mathbf{x}^M \left(\theta_{\mathbf{x}^M} H(\Omega_c | \mathbf{x}^M) + h(\theta_{\mathbf{x}^M}) \right)
\end{aligned} \tag{14}$$

Здесь $\theta_{\mathbf{x}^M} = \sum_{i=1}^c P(\omega_i | \mathbf{x}^M)$ и $P(\omega_i | \mathbf{x}^M, \Omega_c) = P(\omega_i | \mathbf{x}^M) / \sum_{i=1}^c P(\omega_i | \mathbf{x}^M)$. Поскольку независимые случайные переменные θ и $\theta_{\mathbf{x}^M}$ принимают всевозможные значения на отрезке $[0,1]$, то усредняя правые части выражений (13) и (14) по плотности $p(\theta, \theta_{\mathbf{x}^M}) = 1$ и учитывая равенство

$$\int_0^1 h(\theta) d\theta = \int_0^1 h(\theta_{\mathbf{x}^M}) d\theta_{\mathbf{x}^M},$$

получим среднее значение

$$\begin{aligned}
\bar{I}(\mathbf{X}^M; \Omega) &= \int_0^1 I(\mathbf{X}^M; \Omega) p(\theta, \theta_{\mathbf{x}^M}) d\theta d\theta_{\mathbf{x}^M} \\
&= \int_0^1 (\theta H(\Omega_c) + h(\theta)) d\theta - \int_{\mathbf{X}^M} p(\mathbf{x}^M) d\mathbf{x}^M \int_0^1 (\theta_{\mathbf{x}^M} H(\Omega_c | \mathbf{x}^M) + h(\theta_{\mathbf{x}^M})) d\theta_{\mathbf{x}^M} \\
&= \frac{1}{2} H(\Omega_c) - \frac{1}{2} \int_{\mathbf{X}^M} p(\mathbf{x}^M) d\mathbf{x}^M H(\Omega_c | \mathbf{x}^M) \\
&= \frac{1}{2} (H(\Omega_c) - H(\Omega_c | \mathbf{X}^M)) = \frac{1}{2} I(\mathbf{X}^M; \Omega_c)
\end{aligned} \tag{15}$$

где $I(\mathbf{X}^M; \Omega_c)$ – средняя взаимная информация между ансамблем $\mathbf{X}^M$ и множеством классов Ω_c в случае модели классификации без отказов, когда $\Omega = \Omega_c$.

Аналогичное усреднение доли ошибок вида (9) по случайному параметру θ с равномерной плотностью $p(\theta) = 1$ дает для модели с отказами следующую среднюю долю ошибок на ансамбле $\mathbf{X}^M$

$$\begin{aligned}
\bar{\varepsilon}_F(\mathbf{X}^M, \Omega) &= \int_0^1 \left(\theta \sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon_F(\mathbf{X}^M, \omega_i) + (1-\theta) \varepsilon_F(\mathbf{X}^M, \omega_0) \right) p(\theta) d\theta \\
&= \int_0^1 \theta \varepsilon_F(\mathbf{X}^M, \Omega_c) d\theta + \int_0^1 (1-\theta) \varepsilon_F(\mathbf{X}^M, \omega_0) d\theta \\
&= \frac{1}{2} (\varepsilon_F(\mathbf{X}^M, \Omega_c) + \varepsilon_F(\mathbf{X}^M, \omega_0))
\end{aligned} \tag{16}$$

где $\varepsilon_F(\mathbf{X}^M, \Omega_c) = \sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon_F(\mathbf{X}^M, \omega_i)$. Необходимо отметить, что при классификации без отказа $\Omega = \Omega_c$, $\theta = 1$ и, следовательно, $\varepsilon_F(\mathbf{X}^M, \Omega) = \varepsilon_F(\mathbf{X}^M, \Omega_c)$.

В работах [7, 8] показано, что в случае классификации без отказа при использовании меры различия (1) в форме квадратичного махаланобисова расстояния и при соответствующем выборе весовых коэффициентов источников, средняя взаимная информация $I(\mathbf{X}^M; \Omega_c)$ между

ансамблем $\mathbf{X}^M$ и множеством классов Ω_c в схеме WMV не превосходит аналогичную взаимную информацию в схеме WGM. Согласно [15], из указанного соотношения средних взаимных информаций следует, что для такой модели минимально достижимая доля ошибок $\min_F \varepsilon_F(\mathbf{X}^M, \Omega_c)$ в схеме WGM не превосходит минимальную долю ошибок в схеме WMV. В случае классификации с отказом, $\Omega = \Omega_c \cup \omega_0$ и, с учетом (15), выполняется аналогичное соотношение средних взаимных информаций $\bar{I}(\mathbf{X}^M; \Omega)$ между ансамблем и множеством классов Ω для схем комплексирования WMV и WGM. В этом случае минимально достижимая доля ошибок $\min_F \bar{\varepsilon}_F(\mathbf{X}^M, \Omega)$ вида (16) в схеме комплексирования WGM не должна превышать минимальную долю ошибок в схеме WMV.

Задача настоящей работы заключается в экспериментальном подтверждении указанного соотношения долей ошибок для реальных многоклассовых классификаторов, использующих схемы комплексирования WGM и WMV. Исследуются классификаторы, строящиеся на наборах элементарных классификаторов типа NN (Nearest Neighbor) и SVM (Support Vector Machine). Обучение и тестирование выполнено на ансамбле изображений, образованном тройкой декоррелированных компонент $I_1(R,G,B)$, $I_2(R,G,B)$, $I_3(R,G,B)$ цветных RGB изображений лиц полученных с помощью функции *decorrstretch* из пакета программ MATLAB [<http://www.mathworks.com/help/images/ref/decorrstretch.html>]. Результаты тестирования даются в виде средних долей ошибок, полученных для индивидуальных источников I_1 , I_2 , I_3 и для ансамбля (I_1, I_2, I_3) при оптимизированных на этапе обучения значениях параметров классификаторов. Для рассматриваемых классификаторов демонстрируется соотношение долей ошибок по своим и чужим объектам в виде соответствующих ROC кривых.

3. Отклики элементарных классификаторов

Будем считать, что заданы обучающие выборки

$$\mathbf{X}_m^{\text{train}} \subset \mathbf{X}_m, \quad m = 1, \dots, M \quad (17)$$

по всем источникам, последовательность которых порождает обучающую выборку

$$\mathbf{X}^M_{\text{-train}} = (\mathbf{X}_1^{\text{train}}, \dots, \mathbf{X}_M^{\text{train}}) \subset \mathbf{X}^M \quad (18)$$

составных объектов ансамбля $\mathbf{X}^M$. В случае однозначного соответствия объектов источников в ансамбле, размеры обучающих выборок источников и ансамбля одинаковы и равны $K = \|\mathbf{X}^M_{\text{-train}}\| = \|\mathbf{X}_m^{\text{train}}\|$, $m = 1, \dots, M$. Предполагается, что обучающие выборки (17) разбиты на непересекающиеся кластеры так, что

$$\mathbf{X}_m^{\text{train}} = \bigcup_{i=1}^c \mathbf{X}_{im}^{\text{train}}, \quad m = 1, \dots, M$$

и последовательность $\mathbf{X}_i^{M\text{-train}} = (\mathbf{X}_{i1}^{\text{train}}, \dots, \mathbf{X}_{iM}^{\text{train}})$ образует i -й кластер обучающей выборки (18).

Элементарные NN классификаторы типа " ω_i против $\Omega \setminus \omega_i$ " вычисляют функции сходства объектов источников или составных объектов ансамбля с классом ω_i по отклонениям предъявляемых объектов $\mathbf{x}_m \in \mathbf{X}_m$ или $\mathbf{x}^M \in \mathbf{X}^M$ от ближайших представителей кластеров $\mathbf{X}_{im}^{\text{train}}$ или $\mathbf{X}_i^{M\text{-train}}$. Указанные отклонения вычисляются соответственно по мерам различия вида (1) и (2). Элементарные SVM классификаторы вычисляют аналогичные функции сходства в евклидовом пространстве вторичных признаков [13]. В таком пространстве любой объект задается вектором признаков

$$\mathbf{x}_m \rightarrow \mathbf{y}^d(\mathbf{x}_m) = (d(\mathbf{x}_m, \hat{\mathbf{x}}_{m1}), \dots, d(\mathbf{x}_m, \hat{\mathbf{x}}_{mK}), \hat{\mathbf{x}}_{mk} \in \mathbf{X}_m^{\text{train}}, \quad k = 1, \dots, K,$$

вычисляемых по мере различия (1) со всеми объектами обучающей выборки (17). Аналогично, любой составной объект $\mathbf{x}^M \in \mathbf{X}^M$ задается вектором признаков

$$\mathbf{x}^M \rightarrow \mathbf{y}^D(\mathbf{x}^M) = D(\mathbf{x}^M, \hat{\mathbf{x}}_1^M), \dots, D(\mathbf{x}^M, \hat{\mathbf{x}}_K^M), \hat{\mathbf{x}}_k^M \in \mathbf{X}^{M-\text{train}}, k=1, \dots, K,$$

вычисляемых по мере различия (2) со всеми объектами обучающей выборки (18). Отклик элементарного SVM классификатора типа " ω_i против $\Omega_c \setminus \omega_i$ " вычисляется по алгебраическому отклонению предъявляемого объекта $\mathbf{y}$ от разделяющей гиперплоскости, которое согласно [14] определяется величиной $\mathbf{a}_i^t \mathbf{y} + a_i$, где $\mathbf{a}_i$ и a_i – соответственно вектор нормали и параметр смещения, оцениваемые по обучающей выборке. С учетом сделанных замечаний, введенные в (6) и (7) отклики элементарных классификаторов определяются следующими экспоненциальными и сигмоидальными функциями

$$s_i^d(\mathbf{x}_m) = \begin{cases} \exp\left\{-\min_{\hat{\mathbf{x}}_m \in \mathbf{X}_{im}^{\text{train}}} d(\mathbf{x}_m, \hat{\mathbf{x}}_m)\right\} & \text{для NN} \\ \left(1 + \exp\left\{-\left(\mathbf{a}_{im}^t \mathbf{y}^d(\mathbf{x}_m) + a_{im}\right)\right\}\right)^{-1} & \text{для SVM} \end{cases} \quad (19)$$

$$s_i^D(\mathbf{x}^M) = \begin{cases} \exp\left\{-\min_{\hat{\mathbf{x}}^M \in \mathbf{X}_i^{M-\text{train}}} D(\mathbf{x}^M, \hat{\mathbf{x}}^M)\right\} & \text{для NN} \\ \left(1 + \exp\left\{-\left(\mathbf{a}_i^t \mathbf{y}^D(\mathbf{x}^M) + a_i\right)\right\}\right)^{-1} & \text{для SVM} \end{cases} \quad (20)$$

Функции (19) и (20) используются как ядра оценок условных по классам плотностей распределения в разделяющих функциях вида (4) и (5).

4. Представление и мера различия объектов

В этом разделе рассматривается способ древовидного представления объектов, порождаемых источниками изображений. В пространстве древовидных представлений с многоуровневым разрешением вводится семейство мер различия на множестве объектов каждого источника и семейство мер различия составных объектов на ансамбле источников. Указанные семейства мер различия позволяют вычислять отклики элементарных классификаторов вида (19) и (20), и соответствующие этим откликам разделяющие функции вида (4) и (5) на последовательных уровнях представления предъявляемых объектов.

Будем считать, что множеством объектов $\mathbf{X}$ любого источника являются образы, выделенные на полутонных изображениях и удовлетворяющие некоторым ограничениям. Налагаемые ограничения формируют множество допустимых образов, которые заданы наборами пикселей в виде компактных или распределенных двумерных тел, имеющих идентифицируемую систему собственных координат. Учитывая однотипность представлений для источников с номерами $m = 1, \dots, M$, в последующих обозначениях номер источника опущен.

В работе [16], предложен способ представления допустимых образов наборами эллиптических примитивов, структурированными в бинарные деревья. Способ основан на дихотомическом разбиении любого образа $\mathbf{x} \in \mathbf{X}$ на сегменты (наборы пикселей) и аппроксимации каждого сегмента с номером n эллиптическим примитивом Q_n , который соответствует вершине бинарного дерева. Каждый делимый сегмент с номером n дает пару новых сегментов и соответствующую пару аппроксимирующих примитивов с номерами $2n+1, 2n+2$. Нумерация сегментов и примитивов производится с учетом ориентации главной собственной оси делимого сегмента. Выбор ориентации оси выполняется с учетом асимметрии делимого сегмента. Корневой примитив Q_0 аппроксимирует исходный образ $\mathbf{x}$ и имеет номер $n=0$. Концевые вершины дерева примитивов соответствуют неделимым сегментам с числом пикселей не более заданного значения (например, сегменты, состоящие из одного пикселя). Аппроксимирующие примитивы строятся в декартовой системе собственных координат аппроксимируемых сегментов. В случае отсутствия у сегмента с номером n однозначно

идентифицируемой системы собственных координат, примитив Q_n строится в системе координат сегмента-родителя с номером $\lfloor (n-1)/2 \rfloor$. Таким образом, для построения дерева примитивов, достаточно существования идентифицируемой системы собственных координат исходного образа, который является корневым сегментом с номером $n=0$.

В соответствии с выбранным способом представления, для любого допустимого образа (объекта) $\mathbf{x} \in \mathbf{X}$ строится последовательность представлений

$$\mathbf{x}^L = (x_0, K, x_1, K, x_L) \quad (21)$$

образующих полное бинарное дерево глубины L , в котором любая промежуточная вершина имеет две исходящие ветви. Любая подпоследовательность $\mathbf{x}^l = x_0, \dots, x_l, l=0, \dots, L$ в (16) является поддеревом глубины l , в котором набор концевых вершин-примитивов

$$x_l = \{Q_n, n - \text{номер концевоегo примитива в } \mathbf{x}^l\}$$

образует l -й слой полного дерева глубины L . Каждый эллиптический примитив определяется набором параметров

$$Q_n = (\mathbf{r}_n, \mathbf{u}_n, \mathbf{v}_n, z_n), \quad (22)$$

где $\mathbf{r}_n$ – вектор центра эллипса; $\mathbf{u}_n, \mathbf{v}_n$ – векторы большой и малой полуосей; z_n – средний уровень яркости пикселей в аппроксимируемом сегменте. Координаты центра, определяющие вектор $\mathbf{r}_n$, и радиусы, определяющие векторы $\mathbf{u}_n, \mathbf{v}_n$, вычисляются как параметры эллипса рассеяния соответствующего сегмента. Векторы $\mathbf{r}_n, \mathbf{u}_n, \mathbf{v}_n$ задаются в собственной системе координат образа $\mathbf{x}$. Разбиение делимого сегмента с номером $n \geq 0$ производится осью, проходящей через центр этого сегмента и параллельной собственной оси сегмента с номером $n=0$ (соответствует образу $\mathbf{x}$), относительно которой момент инерции объекта максимален. При указанной нумерации примитив Q_n находится в полном дереве на уровне $l = \lfloor \log_2(n+1) \rfloor$. Параметры примитивов (22) с номерами $n \geq 0$ нормируются относительно корневого примитива с номером $n=0$. На рисунке 2 дан пример древовидных представлений глубины $L=8$ образов лица, заданных декоррелированными компонентами $I_1(R,G,B)$, $I_2(R,G,B)$, $I_3(R,G,B)$.

$I_1(R,G,B)$	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$
$I_2(R,G,B)$	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$
$I_3(R,G,B)$	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$

Рисунок 2. Пример древовидных представлений образов лица для декоррелированных компонент $I_1(R,G,B)$, $I_2(R,G,B)$, $I_3(R,G,B)$.

Введем для образов $\mathbf{x} \in \mathbf{X}$ и $\hat{\mathbf{x}} \in \hat{\mathbf{X}}$, представленных деревьями $\mathbf{x}^L$ и $\hat{\mathbf{x}}^L$ вида (21), меру различия $d(\mathbf{x}, \hat{\mathbf{x}}) \geq 0$. Указанная мера различия вводится на пересечении $\mathbf{x}^L \mid \hat{\mathbf{x}}^L$, которое образовано множеством пар соответственных примитивов $Q_n \in \mathbf{x}^L$ и $\hat{Q}_n \in \hat{\mathbf{x}}^L$ в n -х вершинах представляющих деревьев $\mathbf{x}^L$ и $\hat{\mathbf{x}}^L$. Используя описания (22) для соответственных примитивов Q_n и $\hat{Q}_n$, определим для деревьев $\mathbf{x}^L$ и $\hat{\mathbf{x}}^L$ следующие функции различия по параметрам примитивов

$$\begin{cases} \rho_r(\mathbf{x}^L, \hat{\mathbf{x}}^L) = \sum_{n: (Q_n, \hat{Q}_n) \in \mathbf{x}^L \mid \hat{\mathbf{x}}^L} \lambda_n \|\mathbf{r}_n - \hat{\mathbf{r}}_n\| \\ \rho_{uv}(\mathbf{x}^L, \hat{\mathbf{x}}^L) = \sum_{n: (Q_n, \hat{Q}_n) \in \mathbf{x}^L \mid \hat{\mathbf{x}}^L} \lambda_n (\|\mathbf{u}_n - \hat{\mathbf{u}}_n\| + \|\mathbf{v}_n - \hat{\mathbf{v}}_n\|) \\ \rho_z(\mathbf{x}^L, \hat{\mathbf{x}}^L) = \sum_{n: (Q_n, \hat{Q}_n) \in \mathbf{x}^L \mid \hat{\mathbf{x}}^L} \lambda_n \|z_n - \hat{z}_n\| \end{cases}, \quad (23)$$

где

$$\lambda_n = \frac{\lfloor \log_2(n+1) \rfloor 2^{-\lfloor \log_2(n+1) \rfloor}}{\sum_{n: (Q_n, \hat{Q}_n) \in \mathbf{x}^L \mid \hat{\mathbf{x}}^L} \lfloor \log_2(n+1) \rfloor 2^{-\lfloor \log_2(n+1) \rfloor}}$$

веса соответственных примитивов, а нормы заданы в метрике L1. Функции (23) позволяют найти средние отклонения образов обучающей выборки $\mathbf{X}^{\text{train}}$ относительно образа $\hat{\mathbf{x}}$ в форме следующих соотношений

$$\begin{cases} \sigma_r(\hat{\mathbf{x}}^L) = \frac{1}{\|\mathbf{X}^{\text{train}}\| - 1} \sum_{\mathbf{x}^L: \mathbf{x} \in \mathbf{X}^{\text{train}}} \rho_r(\mathbf{x}^L, \hat{\mathbf{x}}^L) \\ \sigma_{uv}(\hat{\mathbf{x}}^L) = \frac{1}{\|\mathbf{X}^{\text{train}}\| - 1} \sum_{\mathbf{x}^L: \mathbf{x} \in \mathbf{X}^{\text{train}}} \rho_{uv}(\mathbf{x}^L, \hat{\mathbf{x}}^L) \\ \sigma_z(\hat{\mathbf{x}}^L) = \frac{1}{\|\mathbf{X}^{\text{train}}\| - 1} \sum_{\mathbf{x}^L: \mathbf{x} \in \mathbf{X}^{\text{train}}} \rho_z(\mathbf{x}^L, \hat{\mathbf{x}}^L) \end{cases}, \quad (24)$$

Тогда с учетом (23) и (24) мера различия пары образов $\mathbf{x}$ и $\hat{\mathbf{x}}$ определяется функцией

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\rho_r(\mathbf{x}^L, \hat{\mathbf{x}}^L)}{\sigma_r(\hat{\mathbf{x}}^L)} + \frac{\rho_{uv}(\mathbf{x}^L, \hat{\mathbf{x}}^L)}{\sigma_{uv}(\hat{\mathbf{x}}^L)} + \frac{\rho_z(\mathbf{x}^L, \hat{\mathbf{x}}^L)}{\sigma_z(\hat{\mathbf{x}}^L)}. \quad (25)$$

Полагая, что источники ансамбля $\mathbf{X}^M = (\mathbf{X}_1, \dots, \mathbf{X}_M)$ допускают древовидные представления образов в форме (21), функции вида (25), вычисляемые на множествах $\mathbf{X}_m$, $m = 1, \dots, M$, дают введенные в (1) меры различия объектов для отдельных источников, а сумма значений таких мер, взятых с весами источников, дает обобщенную меру различия вида (2) для составных объектов на ансамбле источников.

5. Схема обучения и тестирования классификаторов

Обучение и тестирование классификаторов выполняется на ансамбле источников изображений, допускающих древовидное представление в форме (21). Применяется процедура разбиения множества составных объектов ансамбля $\mathbf{X}^M$ на обучающую $\mathbf{X}^{M_{\text{train}}}$ и тестовую $\mathbf{X}^{M_{\text{test}}}$ выборки по схеме "N – fold cross-validation"[17]. В таком разбиении группы (folds)

формируются из представителей всех классов: число представителей от класса в каждой группе одинаково и пропорционально мощности класса в $\mathbf{X}^M$, представители классов в каждой из N групп различны. На каждом шаге скользящего контроля с номером $n = 1, \dots, N$ разбиение дает пару выборок $(\mathbf{X}^{M_train}, \mathbf{X}^{M_train})$, в которой обучающая выборка образована $N - 1$ группами, а тестовая – одной группой. Благодаря соответствию между объектами источников, любая пара $(\mathbf{X}^{M_train}, \mathbf{X}^{M_train})$ порождает для всех источников пары выборок $(\mathbf{X}_m^{train}, \mathbf{X}_m^{test})$ мощности $\|\mathbf{X}_m^{train}\| = \|\mathbf{X}^{M_train}\|$ и $\|\mathbf{X}_m^{test}\| = \|\mathbf{X}^{M_test}\|$ при $m = 1, \dots, M$.

Обучение включает получение оценок параметров (24) меры различия (25), оценок параметров $(\mathbf{a}_{im}^t, a_{im})$ и $(\mathbf{a}_i^t, a_i)$ в откликах элементарных SVM классификаторов вида (19) и (20), оценок порогов Δ_i в решающем правиле (8) и весов источников w_m , $m = 1, \dots, M$, используемых в схемах комплексирования WMV и WGM. Оценки параметров меры различия $d(\mathbf{x}_m, \hat{\mathbf{x}}_m)$ вычисляются согласно (24) на обучающей выборке $\mathbf{X}_m^{train} \subset \mathbf{X}_m$. Оценки параметров $(\mathbf{a}_{im}^t, a_{im})$, $(\mathbf{a}_i^t, a_i)$ строятся для элементарных классификаторов " ω_i против $\Omega_c \setminus \omega_i$ ", $i = 1, \dots, c$, в которых свои и чужие объекты заданы соответствующими парами кластеров $(\mathbf{X}_{im}^{train}, \mathbf{X}_m^{train} \setminus \mathbf{X}_{im}^{train})$ и $(\mathbf{X}_i^{M_train}, \mathbf{X}^{M_train} \setminus \mathbf{X}_i^{M_train})$. Для нахождения параметров гиперплоскостей используется функция *fitcecoc* версии SVM из пакета MATLAB (<http://www.mathworks.com/help/stats/fitcecoc.html>). Пороговые значения Δ_i , $i = 1, \dots, c$ определяются из условия минимизации средней доли ошибок вида (8), усредненной по значениям $0 \leq \theta \leq 1$ с равномерной плотностью $p(\theta) = 1$. Полагая

$$\Delta_i = \delta g_i^{\max} = \delta P(\omega_i | \Omega_c) s_i^{\max}, \quad i = 1, \dots, c, \quad (26)$$

нахождение оптимальных порогов Δ_i^* , $i = 1, \dots, c$, сводится к поиску значения $\delta^* \in [0, 1]$, доставляющего минимум средней доли ошибок при заданных вероятностях $P(\omega_i | \Omega_c)$ и выбираемых значениях $s_i^{\max}$ откликов вида (19) или (20). Используемые в эксперименте величины $s_i^{\max}$ приведены в следующем разделе.

Используя представления фиксированной глубины L для объектов источников, на каждом шаге $n = 1, \dots, N$ минимизируется функция

$$\begin{aligned} \bar{\varepsilon}_{F(\delta)}^n(\mathbf{X}^{train}, \Omega) = & \frac{1}{2} \left(\sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon_{F(\delta)}^n(\mathbf{X}^{train}, \omega_i) + \varepsilon_{F(\delta)}^n(\mathbf{X}^{train}, \omega_0) \right) \\ & \frac{1}{2} \left(\sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon^n(\mathbf{X}_i^{train}, \delta) + \sum_{i=1}^c \varepsilon^n(\mathbf{X}^{train} \setminus \mathbf{X}_i^{train}, \delta) \right) \rightarrow \min_{\delta} \end{aligned} \quad (27)$$

где $\varepsilon^n(\mathbf{X}_i^{train}, \delta)$ и $\varepsilon^n(\mathbf{X}^{train} \setminus \mathbf{X}_i^{train}, \delta)$ – доли ошибок классификатора " ω_i против $\Omega_c \setminus \omega_i$ " по выборкам $\mathbf{X}_i^{train}$ и $\mathbf{X}^{train} \setminus \mathbf{X}_i^{train}$ n -го шага, причем для m -го источника: $\delta = \delta_m$, $F(\delta) = F_m^d(\delta_m)$, $\mathbf{X}^{train} = \mathbf{X}_m^{train}$ и $\mathbf{X}_i^{train} = \mathbf{X}_{im}^{train}$, а для ансамбля источников: $F(\delta) = F_W^D(\delta)$, $\mathbf{X}^{train} = \mathbf{X}^{M_train}$ и $\mathbf{X}_i^{M_train} = \mathbf{X}_i^{M_train}$. Минимальные значения долей ошибок (27), найденные для всех источников при значениях δ_m^* , используются для вычисления весов источников

$$w_m = -\log \bar{\varepsilon}_{F_m(\delta_m^*)}^n(\mathbf{X}_m^{train}, \Omega), \quad m = 1, \dots, M. \quad (28)$$

Тестирование производится путем предъявления объектов тестовой выборки $\mathbf{X}^{\text{test}}$ классификаторам, составленным из всевозможных $(c-1)$ элементарных классификаторов, обученных на n -м шаге, и вычисления доли ошибок

$$\bar{\varepsilon}_{F(\delta^*)}^n(\mathbf{X}^{\text{test}}, \Omega) = \frac{1}{2} \left(\sum_{i=1}^c P(\omega_i | \Omega_c) \varepsilon_{F(\delta^*)}^n(\mathbf{X}^{\text{test}}, \omega_i) + \varepsilon_{F(\delta^*)}^n(\mathbf{X}^{\text{test}}, \omega_0) \right) \quad (29)$$

с последующим усреднением ошибок (29), вычисленных на N шагах скользящего контроля. Здесь для m -го источника: $F(\delta^*) = F_m^d(\delta_m^*)$, $\mathbf{X}^{\text{test}} = \mathbf{X}_m^{\text{test}}$, а для ансамбля: $F(\delta^*) = F_W^D(\delta^*)$, $\mathbf{X}^{\text{test}} = \mathbf{X}^{M\text{-test}}$. В (29) доли ошибок по своим и чужим объектам вычисляются с накоплением ошибочных решений на всевозможных допустимых наборах из $(c-1)$ элементарных классификаторов. С учетом последнего замечания

$$\varepsilon_{F(\delta^*)}^n(\mathbf{X}^{\text{test}}, \omega_i) = N_i^{n\text{-err}}(\delta^*) / N_i^n, \quad i = 0, \dots, c, \quad (30)$$

где N_i^n и $N_i^{n\text{-err}}(\delta^*)$ – соответственно суммарное число предъявляемых объектов тестовой выборки $\mathbf{X}^{\text{test}}$, принадлежащих классу ω_i , и суммарное число ошибочных решений при значении δ^* по предъявляемым объектам на n -м шаге. При этом объекты класса, отсутствующего в наборе элементарных классификаторов, образуют класс чужих объектов ω_0 . Усреднение значений (29) дает среднюю долю ошибок скользящего контроля

$$\varepsilon_{F(\delta^*)}^{\text{cv_test}} = \frac{1}{N} \sum_{n=1}^N \bar{\varepsilon}_{F(\delta^*)}^n(\mathbf{X}^{\text{test}}, \Omega), \quad (31)$$

где множество $\mathbf{X}^{\text{test}}$ и преобразование $F(\delta^*): \mathbf{X}^{\text{test}} \rightarrow \Omega$ определяются источником или ансамблем источников и типом тестируемого классификатора.

Оценки вида (30), полученные при фиксированных значениях параметра $\delta \in [0, 1]$ на N шагах скользящего контроля, дают следующие параметрические оценки определенных в (10) долей верных и ложных решений

$$\begin{aligned} \text{TPR}(\delta) &= 1 - \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^c P(\omega_i | \Omega_c) N_i^{n\text{-err}}(\delta) / N_i^n \\ \text{FPR}(\delta) &= \frac{1}{N} \sum_{n=1}^N N_0^{n\text{-err}}(\delta) / N_0^n \end{aligned}, \quad (32)$$

Оценки (32) использованы для построения ROC кривых исследуемых классификаторов.

6. Экспериментальные результаты

Эксперименты выполнены на цветных RGB изображениях лиц. Множество исходных изображений содержит 1000 объектов от 25 персон, по 40 изображений от каждой персоны, полученных при различном ракурсе съемки. Персоны образуют множество Ω_c , содержащее $c = 25$ классов с равномерным условным распределением $P(\omega_i | \Omega_c) = 1/c$, $i = 1, \dots, 25$. Декоррелированные компоненты $I_1(R, G, B)$, $I_2(R, G, B)$, $I_3(R, G, B)$ образуют источники $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ и ансамбль $\mathbf{X}^3 = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ изображений одинаковой мощности $\|\mathbf{X}_1\| = \|\mathbf{X}_2\| = \|\mathbf{X}_3\| = \|\mathbf{X}^3\| = 1000$.

Информативными объектами изображений являются образы лица, ограниченные параметрическим овалом, симметричным относительно нормали к линии глаз. Представления образов заданы наборами эллиптических примитивов, структурированных в бинарные деревья глубины $L = 8$. Эксперименты проведены с использованием процедуры скользящего контроля с параметром $N = 40$, при котором каждая группа разбиения включает 25 объектов, по одному представителю от каждого класса. Для назначения порогов вида (26) использованы

максимальные значения откликов $s_i^{\max} = 1$, $i = 1, \dots, 25$, которые при равномерном распределении $P(\omega_i | \Omega_c) = 1/c$ дают одинаковые пороги по классам. Оптимизация порогов выполнена согласно условию (27), веса источников вычислены согласно (28). Усредненные доли ошибок тестирования вида (31) по отдельным источникам (компонентам изображений) и по ансамблю источников, полученные для классификаторов типа NN и SVM и схем комплексирования WMV и WGM, даны в таблице 1. Построенные на оценках (32) ROC кривые классификаторов для индивидуальных источников I_1 , I_2 , I_3 и для ансамбля (I_1, I_2, I_3) с использованием схем комплексирования WMV и WGM даны на рисунке 3.

Приведенные в таблице 1 значения долей ошибок при оптимальных порогах демонстрируют преимущество классификаторов составных объектов по сравнению с классификаторами объектов отдельных источников. Сравнение ошибок, полученных на двух схемах комплексирования, показало более высокую эффективность схемы WGM относительно схемы WMV. Аналогичный эффект продемонстрирован сравнительными значениями AUC для ROC кривых на рисунке 3.

Таблица 1. Средние доли ошибки распознавания лиц по компонентам I_1, I_2, I_3 цветных RGB изображений лиц и по ансамблю (I_1, I_2, I_3) с использованием схем комплексирования WMV и WGM.

классификатор	компоненты			схема комплексирования	
	I_1	I_2	I_3	WMV	WGM
NN	0,0055	0,0145	0,0295	0,0035	0,0030
SVM	0,0190	0,0500	0,0540	0,0080	0,0025

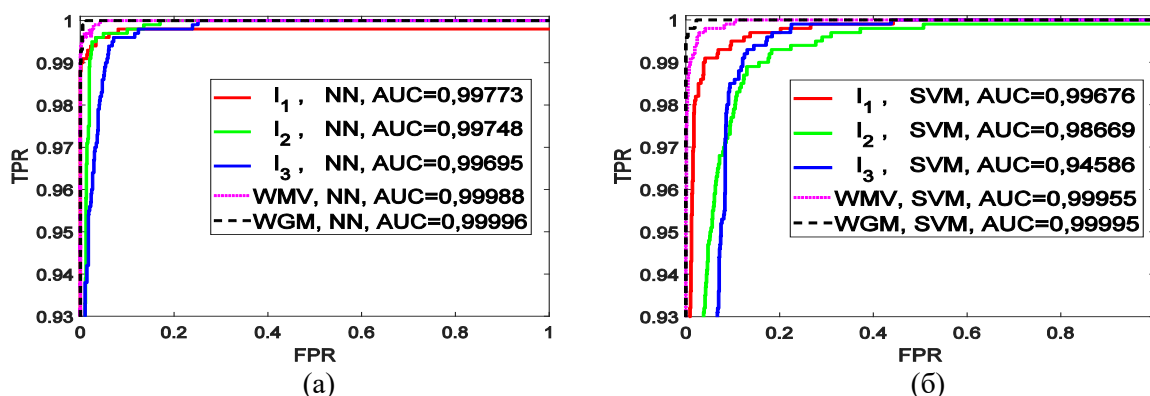


Рисунок 3. ROC кривые классификаторов для источников I_1 , I_2 , I_3 и ансамбля (I_1, I_2, I_3) (а) – классификаторы типа NN; (б) – классификаторы типа SVM.

7. Заключение

В настоящей работе исследованы многоклассовые метрические классификаторы на ансамбле изображений, порождаемых источниками различной модальности. Рассмотрены классификаторы, в которых коллективные решения по предъявляемым наборам изображений (составным объектам) от ансамбля источников строятся на основе традиционной схемы комплексирования решений по источникам (Weighted Majority Vote) и на основе оригинальной схемы комплексирования источников (Weighted General Measure). В обеих схемах предусмотрена возможность отнесения составного объекта к классу отказов. Схема WMV использует взвешенную композицию решений по изображениям отдельных источников и реализуется на наборе классификаторов, по одному на каждый источник. Схема WGM использует взвешенную композицию мер различия объектов источников с представителями

обучающей выборки и реализуется на едином классификаторе составных объектов ансамбля. Для заданного ансамбля источников отмечена зависимость наименьшей вероятности ошибки классификации для заданного ансамбля источников от средней взаимной информации между множеством составных объектов и множеством классов. Показано, что в модели классификации с отказом указанная средняя взаимная информация равна половине средней взаимной информации в модели без отказа. Ранее для модели без отказа было доказано, что средняя взаимная информация между множеством составных объектов и множеством классов в схеме WGM превосходит среднюю взаимную информацию в схеме WMV. Следствием этого результата является возможность достижения меньшей вероятности ошибок в схеме WGM по сравнению со схемой WMV для модели классификации без отказа. С учетом полученного соотношения средних взаимных информаций для случаев с отказом и без отказа, аналогичное соотношение минимальных вероятностей ошибок в схемах WGM и WMV должно сохраняться и для модели с отказом. Этот результат подтвержден результатами распознавания цветных изображений лиц, образующих составные объекты от трехкомпонентных источников. Проведены эксперименты с декоррелированными компонентами цветных RGB изображений лиц, которые образуют трехкомпонентный ансамбль слабо зависимых источников. В терминах ROC кривых исследованы многоклассовые классификаторы с отказом, построенные на наборах элементарных NN и SVM классификаторов, обучаемых по схеме "класс против всех". В обоих случаях схема комплексирования WGM продемонстрировала меньшую долю ошибок по сравнению со схемой WMV. Планируется исследовать классификаторы с указанными схемами комплексирования на ансамблях с более широким составом источников биометрических изображений.

8. Благодарности

Работа поддержана проектами РФФИ №15-01-04671, №18-07-01231.

9. Литература

- [1] Журавлев, Ю.И. Об алгебраическом подходе к решению задач распознавания или классификации / Ю.И. Журавлев // Проблемы кибернетики. – 1978. – Т. 33. – С. 5-68.
- [2] Воронцов, К.В. Лекции по алгоритмическим композициям / К.В. Воронцов. – Электронный ресурс: machinelearning.ru
- [3] Freund Y. Experiments with a new boosting algorithm / Y. Freund, R.E. Schapire // Proceedings of the International Conference on Machine Learning. – 1996. – P. 148-156.
- [4] Kuncheva, L.I. Combining Pattern Classifiers, Methods and Algorithms, Second Edition / L.I. Kuncheva // Wiley and Sons, 2014.
- [5] Lam, L. Application of majority voting to pattern recognition: An analysis of its behavior and performance / L. Lam, C.Y. Suen // IEEE Transactions on Systems, Man, and Cybernetics. – 1997. – Vol. 27(5). – P. 553-568.
- [6] Kuncheva, L.I. Limits on the majority vote accuracy in classifier fusion / L.I. Kuncheva, C.J. Whitaker, C.A. Shipp, R.P.W. Duin // Pattern Analysis and Applications. – 2003. – Vol. 6. – P. 22-31.
- [7] Ланге, М.М. Информационный критерий для сравнения классификаторов на ансамбле источников / М.М. Ланге // Тезисы докладов 11-й Международной конференций "Интеллектуализация обработки информации", Москва-Барселона. – М.: Торус Пресс, 2016. – С. 32-33.
- [8] Lange, M.M. Information Criterion of Efficiency for Metric Classifiers in Ensemble of Sources / М.М. Ланге // Communications in Computer and Information Science, 2018 (принята к печати).
- [9] Tax, D.M.J. Growing a multiclass classifier with a rejection option / D.M.J. Tax, R.P.W. Duin // Pattern Recognition Letters. – 2008. – Vol. 29. – P. 1565-1570.
- [10] Ланге, М.М. Многоклассовое распознавание образов в пространстве представлений с многоуровневым разрешением / М.М. Ланге, С.Н. Ганебных, А.М. Ланге // Машинное обучение и анализ данных. – 2016. – Vol. 2(1). – P. 70-88.

- [11] Duda, R.O. Pattern Classification, 2nd ed. / R.O. Duda, P.E. Hart, D.G. Stork // Wiley and Sons, 2001.
- [12] База цветных изображений лиц. Электронный ресурс: <http://sourceforge.net/projects/colorfaces/>
- [13] Duin, R.P.W. Experiments with a featureless approach to pattern recognition / R.P.W. Duin, D. Ridder, D.M.J. Tax // Pattern Recognition Letters. – 1997. – Vol. 18. – P. 1159-1166.
- [14] Gallager, R.G. Information Theory and Reliable Communication / R.G. Gallager // Wiley and Sons, 1968.
- [15] Добрушин, Р.Л. Передача информации с дополнительным шумом / Р.Л. Добрушин, Б.С. Цыбаков // Проблемы передачи информации. – 1963. – Т. 14. – С. 21-42.
- [16] Ganebnykh, S.N. Metric Classifier Using Multilevel Network of Templates / S.N. Ganebnykh, M.M. Lange, D.Y. Stepanov // Pattern Recognition and Image Analysis. – 2012. – Vol. 22(2). – P. 265-277.
- [17] Bishop, C.M. Pattern recognition and machine learning / C.M. Bishop. – Springer, 2006. – 738 p.

Fusion schemes for classification with reject in ensemble of image sources

M.M. Lange¹, S.N. Ganebnykh¹

¹Federal Research Center «Computer Science and Control» of RAS, Vavilova street 40, Moscow, Russia, 119333

Abstract. In an ensemble of image sources of different modalities, the fusion schemes for multiclass pattern recognition by a metric decision rule with reject are studied. The suggested classifiers make collective decisions about composite objects that are produced by the collections of the images with one from each source. The discriminant functions of the multiclass classifiers are produced by the binary "class-vs-all" NN or SVM classifiers. Two fusion schemes are investigated namely the known WMV scheme based on weighted majority vote of the decisions for the objects of the individual sources and the original WGM scheme which uses the weighted general measure in the set of the composite objects. This measure is given by a composition of values of the appropriate dissimilarity measures between the corresponding components in any pair of the composite objects. The comparative efficiency of the above fusion schemes is supported by the error rates for recognition of the RGB face images given by the ensemble of the decorrelated components of the color images. For both NN and SVM classifiers, the evaluations of the error rates as well as AUCs of the ROC functions show a profit of the WGM scheme in comparison with the WMV scheme.

Keywords: ensemble of sources, fusion schemes, multiclass classifier, discriminant functions, weighted majority vote, weighted general measure.