

ИДЕНТИФИКАЦИЯ ЛИЧНОСТИ НА ОСНОВЕ РАСПОЗНАВАНИЯ ЛИЦ ПО ДРЕВОВИДНЫМ ПРЕДСТАВЛЕНИЯМ

Степанов Д.Ю.

*Московский государственный институт радиотехники, электроники и автоматики
(технический университет)*

В работе предложен новый метод распознавания объектов, заданных многоканальными изображениями. Распознавание выполняется в пространстве многослойных представлений образов стопками деревьев, слои которых соответствуют каналам изображений. Введена мера стопок деревьев и сформулирован критерий голосования на основе ближайшего эталона. Описана процедура обучения классификатора на основе построения покрытий кластеров обучающего множества в пространстве объектов, представленных стопками деревьев. Предложен способ оптимизации параметров классификатора с использованием процедуры скользящего контроля. Эффективность построенного классификатора продемонстрирована результатами распознавания лиц по изображениям в цветовой модели *HSI*.

ВВЕДЕНИЕ

Активное развитие систем анализа изображений и распознавания образов обусловлено ростом быстродействия и уменьшением стоимости вычислительных мощностей, что позволяет обрабатывать большие массивы данных за короткий временной интервал [1]. Существует большое число методов распознавания, разработанных как зарубежными [2-4], так и отечественными учеными [5-7].

В общем случае, задача идентификации личности решается средствами распознавания образов от источников биометрической информации, заданных изображениями лиц, отпечатков пальцев, подписей, радужной оболочки глаза, ладони и другими [8]. Одним из методов, способных эффективно и быстро выполнять распознавание образов независимо от природы исходных объектов, является метод классификации в пространстве древовидных представлений образов эллиптическими примитивами [9]. Этот метод ориентирован на работу с объектами, заданными полутоновыми изображениями. Для части источников, таких как отпечатки пальцев, подписи, речевые сигналы использование информации, содержащейся в полутоновых изображениях, обеспечивает вполне приемлемые показатели качества распознавания. Однако, для таких объектов как лица, радужная оболочка глаза и ряда других полутоновые изображения не дают полной информации об этих объектах [1]. Для эффективного распознавания таких объектов целесообразно воспользоваться цветными изображениями. Цветные изображения объектов в модели *HSI* или *RGB* могут

рассматриваться как трехканальные изображения. Поэтому, располагая такими изображениями, необходимо выделить в них информативные объекты (образы) и построить для выделенных объектов представления, удобные для решения задач анализа и распознавания.

В настоящей работе исследуется задача построения классификатора в пространстве многоканальных древовидных представлений. Описывается формирование базы изображений лиц, предлагаются методы выделения информативных объектов лиц, вводится мера различия стопок деревьев, процедура построения и обучения классификатора. Демонстрируются результаты классификации одноканальных и многоканальных объектов лиц в цветовой модели *HSI* предлагаемым методом и методом опорных векторов.

1. ОДНОКАНАЛЬНЫЕ ДРЕВОВИДНЫЕ ПРЕДСТАВЛЕНИЯ ОБЪЕКТОВ И ПОСТАНОВКА ЗАДАЧИ

Пусть \mathbf{A} – множество объектов, каждый объект $A \in \mathbf{A}$ представлен бинарным $(L+1)$ -уровневым деревом:

$$A_L = (a_0, \dots, a_l, \dots, a_L), \quad (1)$$

где $a_l : l+1 \leq \|a_l\| \leq 2^l$ – представление l -го уровня, образованное концевыми вершинами $(l+1)$ -уровневого поддерева $A_l = (a_0, \dots, a_l)$.

Множество \mathbf{A} содержит объекты, принадлежащие $c+1$ классам:

$$\mathbf{A} = \{\mathbf{A}_i\}_{i=0}^c,$$

где каждый класс \mathbf{A}_i с ненулевым номером включает семантически однородные объекты, а нулевой класс \mathbf{A}_0 объединяет все прочие объекты. На множестве представлений вида (1) в работе [10] введена мера различия пары объектов $(A, \hat{A}) \in \mathbf{A}$:

$$d_L(A, \hat{A}) = d(A_L, \hat{A}_L). \quad (2)$$

Мера (2) позволяет ввести на множестве \mathbf{A} сферу:

$$S_L(\hat{A}, D_L(\hat{A})) = \{A : d_L(A, \hat{A}) \leq D_L(\hat{A})\}, \quad (3)$$

где \hat{A} – центр сферы, а $D_L(\hat{A})$ – радиус сферы, определяемый по мере (2).

Для обучения используется обучающее множество объектов:

$$\mathbf{B} = \{\mathbf{B}_i = \{B_{ij}\}_{j=1}^{m_i}\}_{i=0}^c \subset \mathbf{A}$$

где \mathbf{B}_i – i -й кластер мощности m_i . В пространстве представлений вида (1) множество центров сфер вида (3):

$$\mathbf{S}_l(\hat{\mathbf{B}}, \mathbf{D}_l(\hat{\mathbf{B}})) = \{ \{ S_l(\hat{B}_{ij}, D_l(\hat{B}_{ij})) \}_{j=1}^{\hat{m}_i} \}_{i=1}^c, \quad (4)$$

образует сеть эталонов:

$$\hat{\mathbf{B}} = \{ \hat{\mathbf{B}}_i = \{ \hat{B}_{ij} \}_{j=1}^{\hat{m}_i} \}_{i=1}^c \quad (5)$$

и множество радиусов их сфер влияния:

$$\mathbf{D}_L(\hat{\mathbf{B}}) = \{ \mathbf{D}_L(\hat{\mathbf{B}}_i) = \{ D_L(\hat{B}_{ij}) \}_{j=1}^{\hat{m}_i} \}_{i=1}^c. \quad (6)$$

Множества (5) и (6) позволяют ввести меру сходства между любым объектом $A \in \mathbf{A}$ и подмножеством эталонов $\hat{\mathbf{B}}_i$. Эта мера определяется величиной:

$$\mu_L(A, \hat{\mathbf{B}}_i) = \max_{j=1}^{\hat{m}_i} 2^{-d_L(A, \hat{B}_{ij}) / D_L(\hat{B}_{ij})}, \quad (7)$$

которая достигается на «ближайшем» эталоне $d_L(A, \hat{B}_{ij}) / D_L(\hat{B}_{ij})$, обеспечивающим минимум $\hat{\mathbf{B}}_i$. Мера сходства (7) дает неотрицательный голос подмножества эталонов $\hat{\mathbf{B}}_i$ в пользу принадлежности объекта A классу \mathbf{A}_i . Правило классификации по критерию ближайшего эталона сводится к нахождению подмножества эталонов $\hat{\mathbf{B}}_m$, для которого достигается максимальное значение меры сходства:

$$\hat{\mathbf{B}}_m = \arg \max_{i=1}^c \mu_L(A, \hat{\mathbf{B}}_i) \quad (8)$$

и выбору решающего класса \mathbf{A}_m , если классифицируемый объект A принадлежит хотя бы одной сфере в множестве $\mathbf{S}_L(\hat{\mathbf{B}}_m, \mathbf{D}_L(\hat{\mathbf{B}}_m))$, либо выбору нулевого решающего класса \mathbf{A}_0 в противном случае. Формально, номер решающего класса i^* определяется следующим критерием:

$$i^* = m \max_{j=1}^{m_i} [d_L(A, \hat{B}_{mj}) \leq D_L(\hat{B}_{mj})], \quad (9)$$

где $[d_L(A, \hat{B}_{mj}) \leq D_L(\hat{B}_{mj})] = 1$ или 0 , если $d_L(A, \hat{B}_{mj}) \leq D_L(\hat{B}_{mj})$ или неверно.

Пусть P_{own} и P_{alien} – доли «своих» (own) и «чужих» (alien) объектов в множестве \mathbf{A} относительно множества эталонов вида (5). Если $\varepsilon_{\mathbf{A}}^{(\text{own})}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ и $\varepsilon_{\mathbf{A}}^{(\text{alien})}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ – доли ошибок классификации по критерию (9), полученные для объектов из

соответствующих подмножеств $\bigcup_{i=1}^c \mathbf{A}_i$ и \mathbf{A}_0 , то оценка вероятности ошибок классификации определяется математическим ожиданием следующего вида:

$$\varepsilon_{\mathbf{A}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \varepsilon_{\mathbf{A}}^{(\text{own})}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))P_{\text{own}} + \varepsilon_{\mathbf{A}}^{(\text{alien})}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))P_{\text{alien}}. \quad (10)$$

Для оценок вероятности ошибок вида (10) по объектам обучающего множества \mathbf{B} и тестового множества \mathbf{A}/\mathbf{B} приняты обозначения $\varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ и $\varepsilon_{\mathbf{A}\setminus\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$. При выбранном значении L , основная задача заключается в построении множества сфер (4), которое обеспечивает минимальное значение оценки вероятности ошибок обучения $\varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$. Решение этой задачи строится на древовидно-структурированном покрытии (*Tree Structured Covering*) обучающего множества сферами, оптимизация которого дает множество эталонов $\hat{\mathbf{B}}^*$ вида (5) и множество их параметров $\mathbf{D}_L(\hat{\mathbf{B}}^*)$ вида (6) такие, что:

$$(\hat{\mathbf{B}}^*, \mathbf{D}_L(\hat{\mathbf{B}}^*)) = \arg \min_{\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})} \varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})). \quad (11)$$

Пара множеств (11) и критерий (9) порождают *TSC*-классификатор. Необходимо показать, что *TSC*-классификатор с параметрами (11) способен обеспечить достаточно малую величину переобучения:

$$|\varepsilon_{\mathbf{A}\setminus\mathbf{B}}(\hat{\mathbf{B}}^*, \mathbf{D}_L(\hat{\mathbf{B}}^*)) - \varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}^*, \mathbf{D}_L(\hat{\mathbf{B}}^*))|.$$

для реальных объектов, заданных цветными изображениями лиц.

В последующем разделе рассматривается процедура формирования базы цветных изображений лиц. В разделе 3 описывается процедура выделения информативной части лица на изображении. Способ представления многослойных древовидных представлений лиц в цветовой модели *HSI* и мера сравнения многослойных представлений представлена в разделе 4. Раздел 5 содержит описание процедуры обучения *TSC*-классификатора и оптимизации параметров древовидно-структурного покрытия с использованием метода скользящего контроля. В разделе 6 приведены оценки параметров используемой меры различия многослойных представлений лиц. Раздел 7 демонстрирует результаты распознавания лиц с использованием *TSC*-классификатора и *SVM*-классификатора на основе метода опорных векторов.

2. ФОРМИРОВАНИЕ БАЗЫ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ ЛИЦ

В рамках данной работы были проанализированы базы изображений (табл.1) на предмет соответствия требованиям:

- Проведения фото/видео съемки. Съемка объекта должна выполняться при тщательном контроле внешних условий:
 - равномерная освещенность, обеспечивающая отсутствие теней на изображении лица;
 - отсутствие артефактов на лице человека, таких как шапка, шарф, очки и др., исключающее необходимость предобработки данных;
 - различная удаленность камеры от объекта съемки, позволяющая выполнять классификацию объектов различных масштабов;
- Формирования базы изображений лиц. База данных лиц должна содержать следующие объекты:
 - цветные изображения лиц людей, обеспечивающие информативность данных классификации;
 - изображения людей мужского и женского пола, позволяющие проверить обобщающую способность системы распознавания;
 - большое число изображений каждой персоны, обеспечивающее качественное проведение процедур обучения и тестирования классификатора;
 - различные ракурсы съемки каждой персоны, позволяющие проверить обобщающую способность системы распознавания;
 - различный интервал проведения съемки персоны, обеспечивающий качественное проведение процедур обучения и тестирования классификатора;
- Международного стандарта ISO/IEC 19794-5 [11];
- Лучших мировых практик по подготовке данных распознавания [12].

Таблица №1

Перечень общедоступных баз изображений лиц

№	Наименование базы	Ссылка
1	FERET	http://www.itl.nist.gov/iad/humanid/feret/
2	MIT	ftp://whitechapel.media.mit.edu/pub/images/
3	M2VTS	http://www.tele.ucl.ac.be/M2VTS/

№	Наименование базы	Ссылка
4	Yale	http://www.cvc.yale.edu/projects/yalefaces/
5	Yale B	http://www.cvc.yale.edu/projects/yalefacesB/
6	XM2TVS	http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/
7	BioID	http://www.bioid.com/support/downloads/

По результатам проведенного анализа был сделан вывод о необходимости формирования собственной базы изображений, удовлетворяющей ранее указанным требованиям, так как рассмотренные базы этим требованиям не удовлетворяли.

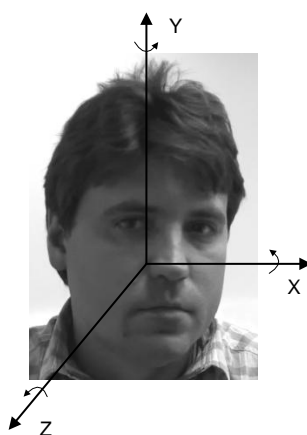


Рис.1. Определение положения объекта в координатном пространстве X, Y, Z

Положение объекта в пространстве описывается координатными осями X, Y, Z . Координатная ось X определяет поворот объекта вокруг горизонтальной плоскости, а ось Y – вертикальной. Ось Z задает положение объекта при его движении против часовой стрелки часов (рис.1).

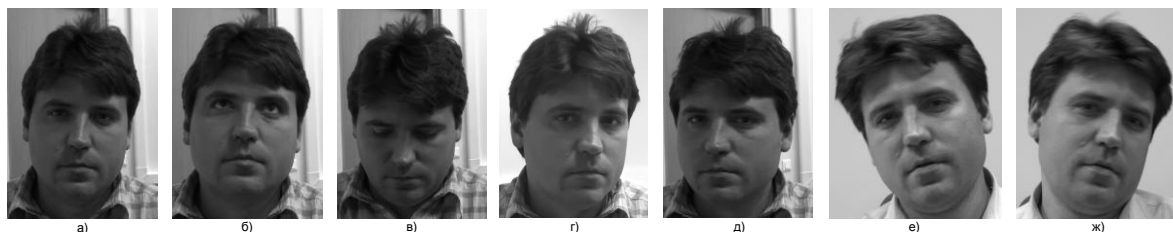


Рис.2. Различные углы поворота объекта по осям X, Y, Z :
 а) $(0,0,0)$; б) $(+45,0,0)$; в) $(-45,0,0)$; г) $(0,+45,0)$; д) $(0,-45,0)$; е) $(0,0,+45)$;
 ж) $(0,0,-45)$

Сформированная база изображений представлена объектами мужского и женского пола, число персон $c = 25$, по $m_i = 40$ реализаций каждой персоны. Цвет волос и глаз

людей не фиксирован. Ракурсы объекта ограничены небольшим углом поворота от 0 до 10 по осям X , Y и Z при нейтральном выражении лица человека (рис.2). Положение глаз объекта относительно угла съемки не задано. Статические изображения лиц получены цифровым фотоаппаратом и видеокамерой в разный временной интервал при искусственном освещении, в виду последнего на изображениях присутствует тень.

3. ВЫДЕЛЕНИЕ ИНФОРМАТИВНЫХ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

Для выделения информативной части лица выполняется поиск координат левого и правого глаза $(x_L, y_L)(x_R, y_R)$ в координатах изображения (X, Y) (рис.3.а). Найденные координаты используются для нахождения расстояния:

$$r = \sqrt{(x_R - x_L)^2 + (y_R - y_L)^2} \quad (12)$$

и вычисления параметров маски-овала, ограничивающей информативную область лица на изображении. Маска задается в координатах (U, V) (рис.3.б), выбираемых следующим образом: ось U проходит через центр линии глаз и перпендикулярна этой линии; центр координат задается точкой на оси U , отстоящей от линии глаз на величину βr с параметром $\beta > 0$; ось V проходит через выбранный центр и опережает ось U .

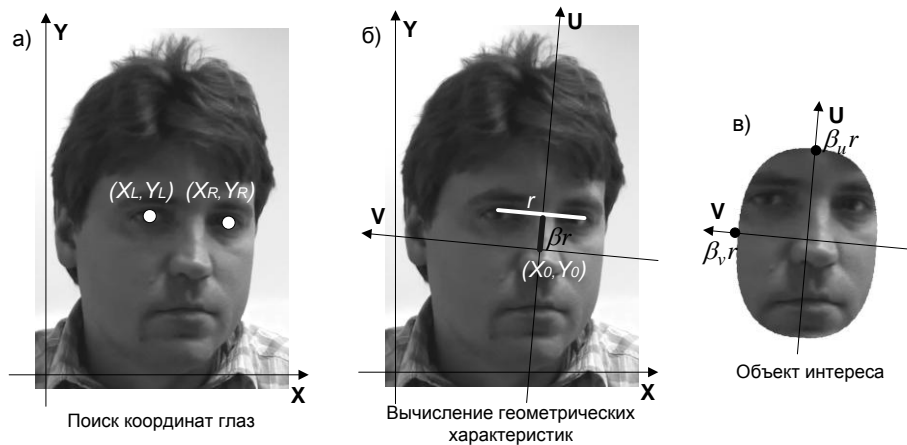


Рис.3. Основные шаги выделения информативной области лица: а) поиск координат глаз; б) вычисление геометрических характеристик; в) выделение объекта интереса

В координатах (U, V) границы маски задаются параметрическим овалом:

$$\frac{U^m}{(\beta_u r)^m} + \frac{V^m}{(\beta_v r)^m} \leq 1 \quad (13)$$

с параметром формы $m \geq 2$, радиусами $\beta_u r$ и $\beta_v r$, определяемыми параметрами $\beta_u > 0$, $\beta_v > 0$ и расстоянием r вида (12). Координаты маски (U, V) связаны с координатами изображения (X, Y) преобразованием поворота и смещения:

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} -\sin \theta & \cos \theta \\ -\cos \theta & -\sin \theta \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}, \quad (14)$$

где $x_0 = \frac{1}{2}(x_R + x_L) + \beta r \sin \theta$, $y_0 = \frac{1}{2}(y_L + y_R) - \beta r \cos \theta$, $\sin \theta = \frac{y_R - y_L}{r}$,

$$\cos \theta = \frac{x_R - x_L}{r}.$$

При заданных параметрах β , β_u , β_v соотношения (12) – (14) позволят выделить на изображении информативную область лица, ограниченную овалом (рис.3.в). В соотношениях (12) и (13) выбраны следующие оптимальные параметры $m = 2.5$, $\beta = 0.5$, $\beta_u = 0.8$, $\beta_v = 0.6$.

Выделенная область инвариантна к поворотам и смещениям изображения. Инвариантность к масштабу и уровню окраски обеспечивается на последующем этапе древовидного представления выделенной области за счет применения соответствующих нормировок. При программной реализации метода выделения информативной части лица использовалась функция определения координат глаз библиотеки *OpenCV* [13].

4. МНОГОСЛОЙНОЕ ДРЕВОВИДНОЕ ПРЕДСТАВЛЕНИЕ ИЗОБРАЖЕНИЙ ЛИЦ В ЦВЕТОВОЙ МОДЕЛИ HSI

Представление объектов, заданных единственным каналом I выполняется методом древовидных описаний, изложенным в [10]. В соответствии с этим методом, выполняется рекурсивная процедура декомпозиции исходного изображения на сегменты и аппроксимация сегментов эллиптическими примитивами, которые характеризуются направляющими осями, центрами и интенсивностью. На этапе классификации объектов система распознавания не использует всю информацию об изображении, так как представление объекта ограничено одним каналом I . Для повышения информативности классификатора каждый пиксель изображения задается вектором значений H, S, I цветовой модели *HSI* [14]:

$$\begin{cases} H = \begin{cases} \sigma, B \leq G \\ 360 - \sigma, B > G \end{cases} \\ S = 1 - \frac{3}{(R+G+B)} \min(R, G, B), \\ I = \frac{1}{3}(R+G+B), \end{cases} \quad (15)$$

где угол $\sigma = \arccos \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{\frac{1}{2}}} \right\} \cdot \frac{180}{\pi}$ определяется в градусной мере, R, G, B – нормализованные значения красного, зеленого и синего цвета пикселя в диапазоне $[0, 1]$. Пример многослойного представления информативной части лица показан на рис.4. Слои представления соответствуют каналам H, S и I .

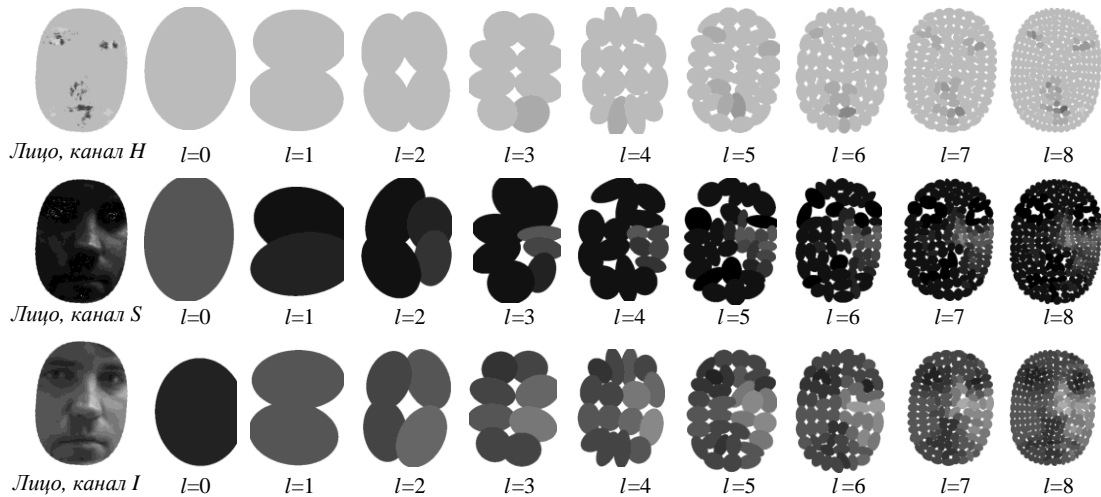


Рис.4. Представление информативной части лица в цветовой модели HSI

Представление каждого слоя построено в виде $(L+1)$ -уровневого бинарного дерева, узлы которого содержат эллиптические примитивы [10]. Стопка таких деревьев соответствует каналам H, S и I .

Используя [10] для параметрического описания эллиптических примитивов, трехслойное представление объекта A деревьями вида (1) определяется тройкой множеств:

$$A_L^H = \{Q_n^H = \{\mathbf{r}_n^H, \mathbf{u}_n^H, \mathbf{v}_n^H, \mathbf{z}_n^H\}\}, \quad (16a)$$

$$A_L^S = \{Q_n^S = \{\mathbf{r}_n^S, \mathbf{u}_n^S, \mathbf{v}_n^S, \mathbf{z}_n^S\}\}, \quad (16b)$$

$$A_L^I = \{Q_n^I = \{\mathbf{r}_n^I, \mathbf{u}_n^I, \mathbf{v}_n^I, \mathbf{z}_n^I\}\}, \quad (16b)$$

где Q_n^H, Q_n^S, Q_n^I - эллиптические примитивы с номером n в слоях H, S, I ; $\mathbf{r}_n^{(H)}, \mathbf{r}_n^{(S)}, \mathbf{r}_n^{(I)}$ - векторы центров, $(\mathbf{u}_n^{(H)}, \mathbf{v}_n^{(H)}), (\mathbf{u}_n^{(S)}, \mathbf{v}_n^{(S)}), (\mathbf{u}_n^{(I)}, \mathbf{v}_n^{(I)})$ - векторы полуосей; а $\mathbf{z}_n^H, \mathbf{z}_n^S, \mathbf{z}_n^I$ - уровни тональности, насыщенности и интенсивности n -ых эллиптических примитивов в соответствующих слоях представления.

На множестве однослойных представлений в [10] введена мера различия любой пары объектов $A \in \mathbf{A}, \hat{A} \in \hat{\mathbf{A}}$, заданные их представлениями $A_L = \{Q_n\}, \hat{A}_L = \{\hat{Q}_n\}$. Эта мера определена на пересечении представлений $A_L \cap \hat{A}_L = \bigcup_n (Q_n, \hat{Q}_n)$, образованном множеством пар соответственных примитивов с одинаковыми номерами. Используя меру, предложенную в [10] введем для пары объектов (A, \hat{A}) меру различия по слоям представления (16):

$$d_L^H(A, \hat{A}) = \sum_{n: (Q_n, \hat{Q}_n) \in (A_L^H \cap \hat{A}_L^H)} \rho^H(Q_n, \hat{Q}_n) \omega_n^H, \quad (17a)$$

$$d_L^S(A, \hat{A}) = \sum_{n: (Q_n, \hat{Q}_n) \in (A_L^S \cap \hat{A}_L^S)} \rho^S(Q_n, \hat{Q}_n) \omega_n^S, \quad (17b)$$

$$d_L^I(A, \hat{A}) = \sum_{n: (Q_n, \hat{Q}_n) \in (A_L^I \cap \hat{A}_L^I)} \rho^I(Q_n, \hat{Q}_n) \omega_n^I, \quad (17b)$$

где нормализованные весовые коэффициенты $\omega_n^H, \omega_n^S, \omega_n^I$ имеют вид:

$$\omega_n^* = \frac{\lfloor \log_2(n+1) \rfloor 2^{-\lfloor \log_2(n+1) \rfloor}}{\sum_{n: (Q_n^*, \hat{Q}_n^*) \in (A_L^* \cap \hat{A}_L^*)} \lfloor \log_2(n+1) \rfloor 2^{-\lfloor \log_2(n+1) \rfloor}},$$

а функции различия $\rho^H(Q_n, \hat{Q}_n), \rho^S(Q_n, \hat{Q}_n), \rho^I(Q_n, \hat{Q}_n)$ соответственных примитивов (Q_n, \hat{Q}_n) для любого слоя их описаний (Q_n^*, \hat{Q}_n^*) в (16) заданы трехкомпонентной суммой:

$$\rho^*(Q_n, \hat{Q}_n) = \sum_{k=1}^3 \omega_k^* \rho_k^*(Q_n, \hat{Q}_n). \quad (18)$$

Компоненты определяются параметрами примитивов в (16):

$$\begin{aligned} \rho_1^*(Q_n, \hat{Q}_n) &= \frac{\|\mathbf{r}_n^* - \hat{\mathbf{r}}_n^*\|}{\max(\|\mathbf{r}_n^*\|, \|\hat{\mathbf{r}}_n^*\|)}, \\ \rho_2^*(Q_n, \hat{Q}_n) &= \frac{1}{2} \left(\frac{\|\mathbf{u}_n^* - \hat{\mathbf{u}}_n^*\|}{\max(\|\mathbf{u}_n^*\|, \|\hat{\mathbf{u}}_n^*\|)} + \frac{\|\mathbf{v}_n^* - \hat{\mathbf{v}}_n^*\|}{\max(\|\mathbf{v}_n^*\|, \|\hat{\mathbf{v}}_n^*\|)} \right), \\ \rho_3^*(Q_n, \hat{Q}_n) &= \frac{|\mathbf{z}_n^* - \hat{\mathbf{z}}_n^*|}{\max(\mathbf{z}_n^*, \hat{\mathbf{z}}_n^*)}, \end{aligned}$$

а нормализованные весовые коэффициенты $\{\omega_k^* \geq 0\}_{k=1}^3$ оцениваются на этапе обучения.

Меры различия вида (17), заданные для древовидных слоев представления (16), позволяют ввести обобщенную меру различия пары объектов (A, \hat{A}) :

$$d_L(A, \hat{A}) = \gamma^H d_L^H(A, \hat{A}) + \gamma^S d_L^S(A, \hat{A}) + \gamma^I d_L^I(A, \hat{A}), \quad (19)$$

в которой параметры $\gamma^H, \gamma^S, \gamma^I \geq 0$ образуют нормированную тройку весовых коэффициентов, оцениваемых на этапе обучения классификатора.

5. ОБУЧЕНИЕ И ПОСТРОЕНИЕ КЛАССИФИКАТОРА НА ОСНОВЕ ДРЕВОВИДНО СТРУКТУРИРОВАННЫХ ПОКРЫТИЙ

Построение древовидно структурированного покрытия

Обучение классификатора заключается в нахождении множества эталонов (5) и множества радиусов сфер влияния эталонов (6). Указанные множества формируются в результате построения покрытий обучающего множества сферами вида (6) по мере (19). Построение покрытия требует знания весовых коэффициентов $\{\omega_k^* \geq 0\}_{k=1}^3$ в функциях (18) и весовых коэффициентов $\{\gamma^H, \gamma^S, \gamma^I\}$ в мере (19). Поэтому целесообразно рассматривать способ построения покрытия при условии, что указанные параметры меры заданы, а затем изложить способ получения оценок весовых коэффициентов.

Для каждого кластера $\mathbf{B}_i \subset \mathbf{B}$ размера $\|\mathbf{B}_i\| = m_i, i = 1, \dots, c$, подмножество сфер покрытия:

$$\mathbf{S}_L(\mathbf{B}_i, \mathbf{D}_L(\mathbf{B}_i)) = \{S_L(B_{ij}, D_L(B_{ij}))\}_{j=1}^{\hat{m}_i} \quad (20)$$

строится путем дихотомического разбиения подмножества объектов \mathbf{B}_i на

непересекающиеся сегменты \mathbf{B}_{ij} размера $\|\mathbf{B}_{ij}\| = m_{ij}$ такие, что $\mathbf{B}_i = \bigcup_{j=1}^{\hat{m}_i} \mathbf{B}_{ij}, m_i = \sum_{j=1}^{\hat{m}_i} m_{ij}$

и выбора покрывающих сфер для указанных сегментов разбиения. Радиус сегмента \mathbf{B}_{ij}

относительно объекта $\hat{B} \in \mathbf{B}_i$ определяется величиной:

$$\hat{d}_L(\mathbf{B}_{ij}, \hat{B}) = \max_{B \in \mathbf{B}_{ij}} d_L(B, \hat{B}). \quad (21)$$

Тогда центр сферы покрытия сегмента $\hat{\mathbf{B}}_{ij}$ задается объектом:

$$\hat{\mathbf{B}}_{ij} = \arg \min_{\hat{B} \in \mathbf{B}_{ij}} \hat{d}_L(\mathbf{B}_{ij}, \hat{B}), \quad (22)$$

а радиус этой сферы определяется функцией:

$$D_L(\hat{B}_{ij}) = (1 - \alpha_i) D_L^{(1)}(\hat{B}_{ij}) + \alpha_i D_L^{(2)}(\hat{B}_{ij}) \quad (23)$$

с параметром α_i принимающим значения на отрезке $[0,1]$ и величинами $D_L^{(1)}(\hat{B}_{ij}), D_L^{(2)}(\hat{B}_{ij}) \geq 0$, выбираемыми согласно следующей стратегии. Величина $D_L^{(1)}(\hat{B}_{ij})$ определяется «расстоянием» между центрами \hat{B}_{ij} и наиболее удаленным объектом в сегменте \mathbf{B}_{ij} . В терминах соотношений (21) и (22):

$$D_L^{(1)}(\hat{B}_{ij}) = \hat{d}_L(\mathbf{B}_i, \hat{B}_{ij}).$$

Для определения величины $D_L^{(2)}(\hat{B}_{ij})$ вводится «расстояние» между центром \hat{B}_{ij} и наиболее удаленным объектом среди объектов множества $\mathbf{B} \setminus \mathbf{B}_i$:

$$d_L(B^*, \hat{B}_{ij}) = \min_{B \in \mathbf{B} \setminus \mathbf{B}_i} d_L(B, \hat{B}_{ij}).$$

Тогда

$$D_L^{(2)}(\hat{B}_{ij}) = \max\{D_L^{(1)}(\hat{B}_{ij}), d_L(B^*, \hat{B}_{ij})\}.$$

На очередном шаге дихотомии разбиению подвергается сегмент наибольшего размера. В разбиваемом сегменте \mathbf{B}_{ij} выделяется пара опорных объектов:

$$(\hat{B}', \hat{B}'') = \arg \max_{B' \in \mathbf{B}_{ij}, B'' \in \mathbf{B}_{ij}} d_L(B', B'').$$

Сегмент \mathbf{B}_{ij} текущего уровня разбивается на два сегмента $\{\mathbf{B}_{ij'}, \mathbf{B}_{ij}''\}$, $j'' = j' + 1$, следующего уровня, так что $\mathbf{B}_{ij'}$ и \mathbf{B}_{ij}'' содержат наиболее близкие объекты по заданной мере различия к соответствующим опорным объектам \hat{B}' и \hat{B}'' . Число разбиений кластера \mathbf{B}_i ограничено сверху величиной $m_i - 1$. Иллюстрация дихотомической схемы построения покрытия кластера дана на рис.5. Множество сфер такого покрытия образует законченное бинарное дерево.

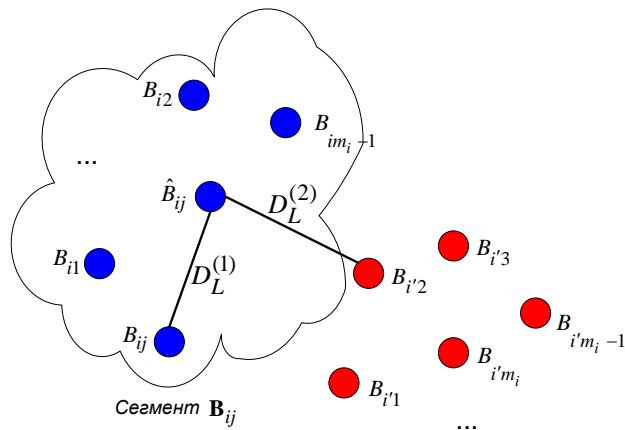


Рис.5. Дихотомическое построение покрытия кластера

Оптимизация параметров покрытия

Процедура построения покрытия (20) на основе дихотомического разбиения кластера на сегменты и выбора сфер с центрами (22) и радиусами (23) производится для всех классов обучающего множества, и требует предварительного задания параметров $\{\hat{m}_i, \alpha_i\}_{i=1}^c$. Эти параметры определяются на этапе оптимизации покрытий.

Оптимизация покрытий (20) производится на основе процедуры скользящего контроля [15], для каждого кластера $\mathbf{B}_i \subset \mathbf{B}, i=1, \dots, c$ и сводится к нахождению числа покрывающих сфер \hat{m}_i и параметра α_i , при которых достигается минимальное значение доли ошибок скользящего контроля $\varepsilon_{CV}(\hat{m}_i, \alpha_i)$. Последняя вычисляется в ходе предъявления всех объектов обучающего множества \mathbf{B} всевозможным покрытиям, строящимся на подмножествах кластера \mathbf{B}_i . Для текущей пары параметров (\hat{m}_i', α_i') оценка ошибки скользящего контроля имеет вид:

$$\varepsilon_{CV}(\hat{m}_i', \alpha_i') = \frac{1}{c} \text{FRR}(\hat{m}_i', \alpha_i') + \frac{c-1}{c} \text{FAR}(\hat{m}_i', \alpha_i'),$$

где $\text{FRR}(\hat{m}_i', \alpha_i')$ и $\text{FAR}(\hat{m}_i', \alpha_i')$ – доли ложных отказов и ложных распознаваний при предъявлении «своих» и «чужих» объектов относительно i -го кластера, а $\frac{1}{c}$ и $\frac{c-1}{c}$ – доли «своих» и «чужих» объектов среди общего числа предъявлений. При этом на каждом шаге скользящего контроля число предъявляемых объектов от каждого кластера фиксировано и одинаково. Оптимизация сводится к нахождению пары свободных параметров:

$$(\hat{m}_i, \alpha_i) = \arg \min_{1 \leq \hat{m}_i \leq m_i; 0 \leq \alpha_i \leq 1} \varepsilon_{CV}(\hat{m}_i, \alpha_i),$$

которая задает оптимальное покрытие (20) с параметрами $\hat{\mathbf{B}}_i, \mathbf{D}_L(\hat{\mathbf{B}}_i)$. Тогда объединение покрытий по всем кластерам обучающего множества порождает множество сфер:

$$\mathbf{S}_L(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \{ \{ S_L(\hat{B}_{ij}, D_L(\hat{B}_{ij})) \}_{j=1}^{\hat{m}_i} \}_{i=1}^c. \quad (24)$$

Множество (24) совместно с критерием (9) порождает TSC-классификатор с множеством эталонов $\hat{\mathbf{B}}_i$ и множеством радиусов $\mathbf{D}_L(\hat{\mathbf{B}}_i)$ сфер влияния эталонов.

6. ОЦЕНКИ ПАРАМЕТРОВ МЕРЫ РАЗЛИЧИЯ

Получение оценок параметров $\{\omega_k^* \geq 0\}, k=1,2,3$ в функциях (18) базируется на построении покрытий $\mathbf{S}_L(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))$ вида (24) по k -ым компонентам меры различия:

$$d_{Lk}(A, \hat{A}) = \sum_{n: (Q_n, \hat{Q}_n) \in (A_L \cap \hat{A}_L^*)} \rho_k^*(Q_n, \hat{Q}_n) \omega_n$$

и вычисление долей ошибок $\varepsilon_{\mathbf{B}}^*(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))$ вида (10) для соответствующих этим покрытий TSC-классификаторов. Индекс (*) соответствует слоям представлений H, S, I . В качестве оценок выбраны величины:

$$\hat{\omega}_k^* = \frac{\log_2 \varepsilon_{\mathbf{B}}^*(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))}{\sum_{k=1}^3 \log_2 \varepsilon_{\mathbf{B}}^*(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))}, k=1,2,3, \quad (25)$$

которые образуют распределение информационных весов для трех компонент меры различия по выбранному слою представлений, где символ (*) означает H, S или I . Оценки вида (25) строятся независимо для каждого слоя представления H, S и I .

Оценки $\{\omega_k^H, \omega_k^S, \omega_k^I\}, k=1,2,3$ полностью определяют меру различия вида (17) по слоям представлений для каналов модели HSI. Эти меры позволяют аналогичным способом получить оценки коэффициентов $\gamma^H, \gamma^S, \gamma^I$ обобщенной меры (19). Оценка каждого коэффициента γ^* строится независимо с использованием покрытий $\mathbf{S}_L^*(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ по соответствующей мере $d_L^*(A, \hat{A})$ и вычисления доли ошибок классификации $\varepsilon_{\mathbf{B}}^*(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))$. В качестве оценок выбраны величины:

$$\hat{\gamma}^* = \frac{\log_2 \varepsilon_{\mathbf{B}}^*(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))}{\log_2 \varepsilon_{\mathbf{B}}^H(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) + \log_2 \varepsilon_{\mathbf{B}}^S(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) + \log_2 \varepsilon_{\mathbf{B}}^I(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))}, \quad (26)$$

где символ (*) соответствует H , S или I слову в представлениях вида (16).

Оценки весовых коэффициентов (25) и (26) определяют обобщенную меру (19), которая используется для построения TSC-классификатора в соответствии с процедурой, описанной в разделе 4.

7. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Схема организации экспериментов заключалась в следующем. Исходные объекты из сформированной базы лиц были преобразованы в изображения размера 800x600 пикселей. Представления информативных объектов строились в форме $(L+1)$ -уровневых полных деревьев с параметром $L=10$. В эксперименте использовано $\|\mathbf{A}\|=1000$ изображений лиц от $c=25$ персон (классов), по $m_i=40$ объектов в каждом классе. Множество лиц разбивалось пятикратно на равные по объему обучающие \mathbf{B} и тестовые $\mathbf{A} \setminus \mathbf{B}$ выборки, которые обеспечивали одинаковые размеры кластеров $m_i=20$, $i=1, \dots, c$. Априорные вероятности предъявления «своих» и «чужих» объектов выбирались равными: $P_{own}=1, 0.75, 0.5$, $P_{alien}=1-P_{own}$.

Обучение TSC-классификатора выполнялось на обучающих выборках в соответствии с процедурами, изложенными в разделах 5 и 6. Для указанных значений P_{own} и пяти обучающих выборок строились TSC-классификаторы, содержащие покрытия для $c'=\lfloor cP_{own} \rfloor$ произвольно выбранных классов. При значении $c'<c$ заданной обучающей выборке \mathbf{B} строилось несколько реализаций таких классификаторов (при $c'=c$ одна реализация каждой выборки \mathbf{B}). Предъявляя объекты тестовой выборки $\mathbf{A} \setminus \mathbf{B}$ для каждой реализации TSC-классификатора, вычислялась оценка вероятности ошибочных решений и усредненная оценка вероятности ошибки по всем реализациям классификатора и предъявляемым тестовым выборкам. Доли ошибок распознавания для TSC-классификатором сравнивались с аналогичными результатами SVM-классификатора, реализованным в библиотеке *OpenCV* [13]. Усредненные результаты распознавания представлены в табл.2. При $P_{own}<1$ использованная версия SVM-классификатора не позволила корректно классифицировать «чужие» объекты из-за отсутствия в этой версии функции отказа.

Усредненные вероятности ошибок
распознавания лиц (%) для классификаторов типа *TSC* и *SVM*

Источник	$P_{own} = 1$		$P_{own} = 0.75$		$P_{own} = 0.5$	
	<i>TSC</i>	<i>SVM</i>	<i>TSC</i>	<i>SVM</i>	<i>TSC</i>	<i>SVM</i>
<i>H</i>	3.7	19.2	6.1	-	4.3	-
<i>S</i>	2.6	2.1	3.3	-	2.9	-
<i>I</i>	2.0	7.4	2.8	-	2.9	-
<i>HSI</i>	1.6	1.0	1.8	-	1.7	-

Из данных табл.2 следует, что оба классификатора (*TSC* и *SVM*) дали улучшение показателей качества за счет использования многослойных представлений модели *HSI* по сравнению с однослойными представлениями для отдельных каналов этой модели. При этом *TSC*-классификатор несколько уступает *SVM*-классификатору (1.6% против 1%). Однако, *TSC*-классификатор способен обеспечивать корректные решения по «чужим» объектам благодаря реализованной в нем функции отказа. Кроме того, применение древовидных представлений с многоуровневым разрешением позволяет организовывать иерархический поиск решений в многоуровневой базе эталонов. Такая стратегия поиска решений приводит к существенному уменьшению вычислительной сложности *TSC*-классификатора по сравнению с *SVM*-классификатором, когда число классов достаточно велико.

ЗАКЛЮЧЕНИЕ

В данной работе предложен новый подход к распознаванию образов, заданных цветными изображениями в модели *HSI*. Подход базируется на трехслойных древовидных представлениях образов, слои которых соответствуют каналам *H*, *S* и *I*. Разработана процедура выделения на *HSI*-изображениях информативной части лица (образа) и построение для него трехслойного древовидного представления эллиптическими примитивами, образующих стопку бинарных деревьев. Создана база изображений лиц для проведения экспериментов с различными классификаторами.

Введена мера различия многослойных древовидных представлений образов и *TSC*-классификатора на основе построения древовидно структурированных покрытий (*TSC*-покрытий) кластеров обучающего множества. Выполнены эксперименты по распознаванию лиц с использованием черно-белых и цветных изображений лиц. Получены сравнительные оценки качества распознавания для *TSC*- и *SVM*-классификаторов. Результаты экспериментов для обоих классификаторов демонстрируют

улучшение характеристик качества при использовании многослойных представлений в модели *HSI* по сравнению с одноканальными древовидными представлениями отдельных каналов.

В перспективе предполагается обобщить предлагаемый подход для классификации объектов от других биометрических источников [16]. Работа выполнена при поддержке РФФИ, проект № 09-01-00573-а.

СПИСОК ЛИТЕРАТУРЫ

1. Phillips P.J., Martin A., Wilson C. L., Przybocki M. An introduction to evaluating biometric systems // *Computer*, 2000 – vol. 33.
2. Jain A.K. Statistical pattern recognition: a review // *IEEE Transactions on pattern analysis and machine intelligence*, 2000 – vol.22.
3. Turk M., Pentland A. Eigenfaces for recognition // *J. Cogn. Neurosci.* – vol. 3, 1991.
4. Cootes T., Walker K., Taylor C. Viewbased active appearance models. In *Proceedings // International Conference on Automatic Face and Gesture Recognition*, 2000.
5. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: Издательство института математики, 1999.
6. Журавлев Ю.И. Избранные научные труды. – М.: Магистр, 1998.
7. Воронцов К.В. Комбинаторная теория надежности обучения по прецедентам: автореф. дис. док.физ.-мат. наук / ВЦ им.А.А.Дородницына РАН. – М.: 2010.
8. Zhao W., Chellappa R., Phillips P.J., Rosenfeld A. Face recognition: a literature survey, *ACM Computer Surveys*, 2003 – vol. 35.
9. Ganebnykh S., Lange M. Classification of 2D Grayscale Objects in a Space of the Multiresolution Representations. // *Pattern Recognition and Image Analysis*. Moscow: MAIK Nauka/Interperiodica, 2009.
10. Ganebnykh S., Lange M. Classification of 2D Grayscale Objects in a Space of the Multiresolution Representations. // *Pattern Recognition and Image Analysis*. Moscow: MAIK Nauka/Interperiodica, 2009 - Vol.19, no.4.
11. ГОСТ Р ИСО/МЭК 19794-5-2006. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. Часть 5. Данные изображения лица. Дата введения 01.01.2007.
12. R. M. McCabe, “Best practice recommendation for the capture of mugshots version 2.0,” 1997, <http://www.nist.gov/itl/div894/894.03/face/face.html>.
13. <http://opencv.willowgarage.com/wiki/>

14. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде Matlab. – М.: Техносфера, 2006.

15. Иванов М. Н., Воронцов К. В. Отбор эталонов, основанный на минимизации функционала полного скользящего контроля // Доклады всероссийской конференции «Математические методы распознавания образов-14». – М.: МАКС Пресс, 2009.

16. Ross A., An introduction to multibiometrics // IEEE Proc. of 15th European Signal Processing Conference. – Poland: 2007.