

Распознавание лиц по многослойным древовидным представлениям цветных изображений*

Степанов Д. Ю.¹, Ланге М. М.²

lange_mm@ccas.ru

¹ Москва, Московский государственный институт радиотехники, электроники и автоматики

² Москва, Вычислительный центр им. А. А. Дородницына РАН

Исследуется новый подход к распознаванию лиц по цветным изображениям в рамках модели классификации двумерных объектов, заданных многоканальными изображениями. Рассматриваемая модель использует пространство многослойных представлений лиц деревьями эллиптических примитивов. Предлагается процедура обучения TSC-классификатора на основе построения древовидно-структурированного покрытия (Tree-Structured Covering) обучающего множества шарами по заданной мере различия многослойных представлений. Эффективность TSC-классификатора продемонстрирована экспериментальными результатами распознавания лиц в пространстве представлений цветных HSI-изображений.

В ряде прикладных задач возникает необходимость в классификации объектов по их образам, полученным от нескольких каналов наблюдения. Примером такой модели является многоканальная система биометрической идентификации личности по изображению лица, форме и рисунку ладони, изображению радужной оболочки, подписи и т. п. [1]. Двумерные объекты, заданные цветными изображениями, могут рассматриваться в рамках трёхканальных моделей RGB и HSI [2], причём вторая модель более предпочтительна для анализа цветных изображений в силу меньшей корреляции каналов H, S и I по сравнению с каналами R, G и B. В частности, модель HSI находит применение в системах распознавания лиц [3].

Наряду с многоканальностью исходных изображений, важным фактором является выбор пространства представлений образов на выходе каналов наблюдения. В данной работе предлагается способ построения многослойных древовидных представлений для объектов, заданных многоканальными изображениями. Такие представления обладают многоуровневым разрешением и ориентированы на снижение вычислительной сложности процедуры распознавания за счёт возможности иерархического поиска решения в многоуровневой базе эталонов. Предлагаемое многослойное древовидное представление применяется для описания лиц, заданных цветными изображениями в модели HSI, и является обобщением древовидного представления, предложенного в [4] для описания полутоновых объектов. На множестве обобщённых древовидных представлений введена мера различия наблюдаемых объектов и построен классификатор по критерию ближайшего эталона, в котором предусмотрена функция отказа. Обучение классификатора базируется на построении и оптимизации древовидно-структурированных покрытий (TSC) семантически однородных класте-

ров обучающего множества шарами в выбранном пространстве представлений образов. В результате обучения строятся оценки параметров покрытий и формируется множество эталонов, которое образует многоуровневую сеть, удобную для быстрого поиска решающих эталонов.

Экспериментальная апробация TSC-классификатора проведена с использованием базы цветных изображений лиц в модели HSI. Результатами апробации являются сравнительные оценки вероятности ошибок для разработанного TSC-классификатора и SVM-классификатора из библиотеки OpenCV [5].

Модель классификации и задача исследования

Многослойное древовидное представление двумерных объектов

Пусть \mathbf{A} — множество двумерных объектов, в котором каждый объект $A \in \mathbf{A}$ задан набором образов от N каналов наблюдения. Будем считать, что образ от каждого q -го канала представлен в виде $(L+1)$ -уровневого бинарного дерева [4], а набор представлений образов по N каналам образует многослойное представление

$$A^L = \{A_q^L = (a_q^0, \dots, a_q^l, \dots, a_q^L)\}_{q=1}^N, \quad (1)$$

в котором q -й слой задан деревом A_q^L и

$$a_q^l = \{Q_n | Q_n - \text{концевой узел в } A_q^l\} \quad (2)$$

множество концевых вершин в поддереве $A_q^l \subset A_q^L$. Вершины Q_n в (2) образованы эллиптическими примитивами, которые аппроксимируют сегменты образа в q -ом канале. Пример трёхслойного представления вида (1), построенного для каналов цветного изображения лица в модели HSI, дан на рис. 1. В данном представлении слои с номерами $q = 1, 2, 3$ соответствуют каналам H, S, I, и каждый q -й слой представлен набором множеств примитивов a_q^l , $l = 0, \dots, 8$, мощности $\|a_q^l\| = 2^l$, которые образуют

Работа выполнена при финансовой поддержке РФФИ, проект № 09-01-00573-а.

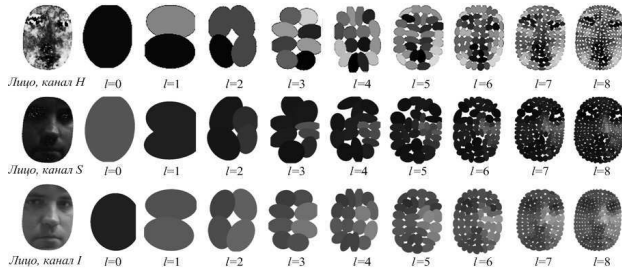


Рис. 1. Пример трёхслойного древовидного представления лица в модели HSI.

полное бинарное дерево A_q^L с параметром $L = 8$. Эллиптические примитивы являются аппроксимациями сегментов, полученных в результате дихотомического разбиения образов в каналах H, S и I.

Для любой пары объектов $A \in \mathbf{A}$ и $\hat{A} \in \mathbf{A}$ примитивы $Q_n \in A_q^L$, $\hat{Q}_n \in \hat{A}_q^L$, находящиеся в вершинах с номером n , считаются соответственными, а подмножество пар соответственных примитивов образует пересечение деревьев q -го слоя

$$A_q^L \cap \hat{A}_q^L = \{(Q_n, \hat{Q}_n)\}.$$

Поскольку описание каждого эллиптического примитива содержит три группы параметров (вектор центра, векторы ориентации и размеров осей, и среднее значение яркости сегмента), то различие соответственных примитивов (Q_n и \hat{Q}_n) в [4] определяется тремя функциями

$$\rho_k^{(q)}(Q_n, \hat{Q}_n) \geq 0, \quad k = 1, 2, 3, \quad (3)$$

выраженными через параметры соответствующей группы. Функции (3) дают три компоненты меры различия пары объектов (A, \hat{A}) по q -му слою их представлений вида (1)

$$d_{Lk}^{(q)}(A, \hat{A}) = \sum_{n: (Q_n, \hat{Q}_n) \in (A_q^L \cap \hat{A}_q^L)} w_n^{(q)} \rho_k^{(q)}(Q_n, \hat{Q}_n), \quad (4)$$

$k = 1, 2, 3$, где $\{w_n^{(q)} \geq 0\}$ — нормированные весовые коэффициенты q -го слоя, определяемые структурой пересечения $A_q^L \cap \hat{A}_q^L$. Тогда с учётом (4) мера различия пары объектов (A, \hat{A}) по q -му слою их представлений определяется функцией

$$d_L^{(q)}(A, \hat{A}) = \sum_{k=1}^3 \omega_k^{(q)} d_{Lk}^{(q)}(A, \hat{A}), \quad (5)$$

а мера различия этой пары по многослойным представлениям вида (1) — функцией

$$d_L(A, \hat{A}) = \sum_{q=1}^N \gamma^{(q)} d_L^{(q)}(A, \hat{A}), \quad (6)$$

где $\{\{\omega_k^{(q)} \geq 0\}_{k=1}^3\}_{q=1}^N$ и $\{\gamma^{(q)} \geq 0\}_{q=1}^N$ — свободные нормированные коэффициенты, оцениваемые на этапе обучения.

Формализация модели классификации

Будем считать, что множество объектов источника $\mathbf{A} = \{\mathbf{A}_i\}_{i=0}^c$ содержит $c + 1$ классов, где каждый класс A_i с номером $i \neq 0$ включает семантически однородные объекты, а класс \mathbf{A}_0 объединяет все прочие объекты. Пусть $\{P_i = P(\mathbf{A}_i)\}_{i=0}^c$ — вероятности классов. Тогда вероятности

$$P_{\text{own}} = \sum_{i=1}^c P_i, \quad P_{\text{alien}} = P_0 = 1 - P_{\text{own}} \quad (7)$$

дают априорное распределение «своих» (own) и «чужих» (alien) объектов на множестве \mathbf{A} .

Для обучения используется множество объектов

$$\mathbf{B} = \{\mathbf{B}_i = \{B_{ij}\}_{j=1}^{m_i}\}_{i=1}^c \subset \mathbf{A}, \quad (8)$$

составленное из семантически однородных кластеров \mathbf{B}_i мощности m_i , $i \neq 0$. На обучающем множестве (8) отбирается множество эталонов

$$\hat{\mathbf{B}} = \{\hat{\mathbf{B}}_i = \{\hat{B}_{ij}\}_{j=1}^{\hat{m}_i}\}_{i=1}^c \subset \mathbf{B}, \quad (9)$$

в котором $\hat{m}_i \leq m_i$ и каждый эталон \hat{B}_{ij} является центром шара

$$S_L(\hat{B}_{ij}, D_L(\hat{B}_{ij})) = \{A \mid d(A, \hat{B}_{ij}) \leq D_L(\hat{B}_{ij})\} \quad (10)$$

с радиусом $D_L(\hat{B}_{ij})$, вычисляемым по мере (6).

Используя множество эталонов (9) и радиусы соответствующих шаров вида (10), вводится мера сходства объекта A с подмножеством эталонов $\hat{\mathbf{B}}_i$:

$$\mu(A, \hat{\mathbf{B}}_i) = \max_{j=1}^{\hat{m}_i} \left(2^{\frac{-d_L(A, \hat{B}_{ij})}{D_L(\hat{B}_{ij})}} [d_L(A, \hat{B}_{ij}) \leq D_L(\hat{B}_{ij})] \right), \quad (11)$$

где $[*]$ — индикатор, принимающий значения 1 или 0 при выполнении указанного или обратного неравенства. При этом $\mu(A, \hat{\mathbf{B}}_i) > 0$, если A принадлежит хотя бы одному из шаров вида (10) с центром $\hat{B}_{ij} \in \hat{\mathbf{B}}_i$ и $\mu(A, \hat{\mathbf{B}}_i) = 0$ в противном случае. Мера сходства (11) даёт следующий критерий установления номера класса i^* для объекта A :

$$\mu(A, \hat{\mathbf{B}}_k) = \max_{i=1}^c \mu(A, \hat{\mathbf{B}}_i), \quad i^* = k[\mu(A, \hat{\mathbf{B}}_k) > 0], \quad (12)$$

в котором значение $i^* = 0$ соответствует классу отказов \mathbf{A}_0 .

Задача исследования

Классификатор по критерию (12) полностью определяется множеством эталонов $\hat{\mathbf{B}}$ вида (9) и множеством радиусов

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

шаров вида (10). Качество классификатора будем оценивать средним значением долей ошибок распознавания «своих» и «чужих» объектов

$$\varepsilon(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \varepsilon^{(\text{own})}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))P_{\text{own}} + \varepsilon^{(\text{alien})}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))P_{\text{alien}} \quad (13)$$

по заданному априорному распределению (7). Задача состоит в выборе стратегии построения покрытия обучающего множества $\mathbf{B} \subset \mathbf{A}$ шарами вида (10) и нахождении в рамках предложенной стратегии пары множеств

$$(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \arg \min_{\hat{\mathbf{B}}', \mathbf{D}_L(\hat{\mathbf{B}}')} \varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}', \mathbf{D}_L(\hat{\mathbf{B}}')), \quad (14)$$

которая доставляет минимум средней доли ошибок $\varepsilon_{\mathbf{B}}(\mathbf{B}, \mathbf{D}_L(\mathbf{B}))$ вида (13) на обучающем множестве объектов \mathbf{B} . Для классификатора, определяемого множествами (14), необходимо найти оценки вероятностей ошибок $\varepsilon_{\mathbf{A} \setminus \mathbf{B}}(\mathbf{B}, \mathbf{D}_L(\mathbf{B}))$ на тестовом множестве $\mathbf{A} \setminus \mathbf{B}$ при различных значениях априорных вероятностей P_{own} и P_{alien} .

В решаемой задаче источник объектов задаётся множеством \mathbf{A} цветных изображений лиц в модели HSI. Однако предлагаемое решение может быть использовано для других источников, допускающих многослойные представления вида (1).

Обучение классификатора

Обучение состоит в получении оценок параметров $\{\{\omega_k^{(q)} \geq 0\}_{k=1}^3\}_{q=1}^N$ и $\{\gamma^{(q)} \geq 0\}_{q=1}^N$ в мерах вида (5) и (6), и построении множеств $(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$, удовлетворяющих условию (14). Ниже предлагается способ построения оптимальной пары множеств эталонов и радиусов по заданной мере различия объектов. Способ применим для построения классификаторов по любым (k, q) -составляющим меры (4) и q -составляющим меры (5), что позволяет найти оценки вероятностей ошибок обучения $\varepsilon_{\mathbf{B}}^{(kq)}$ и $\varepsilon_{\mathbf{B}}^{(q)}$ по указанным составляющим меры различия. При фиксированных значениях k и q в качестве оценок параметров в (5) и (6) выбираются величины

$$\hat{\omega}_k^{(q)} = \frac{\log \varepsilon_{\mathbf{B}}^{(kq)}}{\sum_{k=1}^3 \log \varepsilon_{\mathbf{B}}^{(kq)}}, \quad k = 1, 2, 3, \quad (15)$$

$$\hat{\gamma}^{(q)} = \frac{\log \varepsilon_{\mathbf{B}}^{(q)}}{\sum_{q=1}^N \log \varepsilon_{\mathbf{B}}^{(q)}}. \quad (16)$$

При заданной на множестве \mathbf{B} мере различия $d_L(B, \hat{B})$ пары объектов (B, \hat{B}) , предлагаемый способ построения множеств $(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ состоит в следующем. Для каждого кластера $\mathbf{B}_i \subset \mathbf{B}$ строится древовидно-структурированное подмножество покрывающих этот кластер шаров путём дихотомического разбиения кластера на \hat{m}_i непересекающихся

сегментов

$$\{\mathbf{B}_{ij} \mid \bigcup_{j=1}^{\hat{m}_i} \mathbf{B}_{ij} = \mathbf{B}_i\},$$

и выбора для каждого \mathbf{B}_{ij} шара с центром

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

и радиусом вида

$$D_L(\hat{B}_{ij}) = (1 - \alpha_i)D_L^{(\text{min})}(\hat{B}_{ij}) + \alpha_i D_L^{(\text{max})}(\hat{B}_{ij}),$$

где $\alpha_i \in [0, 1]$ — свободный параметр. Величина $D_L^{(\text{min})}$ определяется расстоянием по заданной мере между центром \hat{B}_{ij} и наиболее удалённым объектом в сегменте \mathbf{B}_{ij} ; величина $D_L^{(\text{max})}$ — расстоянием между \hat{B}_{ij} и наиболее удалённым объектом в подмножестве $\mathbf{B} \setminus \mathbf{B}_i$. На каждом шаге дихотомии выполняется разбиение $\mathbf{B}_{ij} \rightarrow (\mathbf{B}_{ij'}, \mathbf{B}_{ij''})$ сегмента \mathbf{B}_{ij} наибольшей мощности $\|\mathbf{B}_{ij}\|$. Сегменты $\mathbf{B}_{ij'}, \mathbf{B}_{ij''}$ включают объекты сегмента \mathbf{B}_{ij} , которые наиболее близки по заданной мере к наиболее удалённым друг от друга объектам $B_{ij'} \in \mathbf{B}_{ij}$ и $B_{ij''} \in \mathbf{B}_{ij}$. Покрытие кластера $\mathbf{B}_i \subset \mathbf{B}$ даёт пару множеств $(\hat{\mathbf{B}}_i, \mathbf{D}_L(\hat{\mathbf{B}}_i))$. Объединение покрытий кластеров порождает пару $(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \{(\hat{\mathbf{B}}_i, \mathbf{D}_L(\hat{\mathbf{B}}_i))\}_{i=1}^c$.

Оптимизация параметров (\hat{m}_i, α_i) покрытия каждого кластера \mathbf{B}_i производится независимо на основе минимизации доли ошибок скользящего контроля

$$\varepsilon_{\mathbf{B}}^{(\text{cv})}(\hat{m}_i, \alpha_i) = P_i^* \varepsilon_{\mathbf{B}_i}^{(\text{cv})}(\hat{m}_i, \alpha_i) + (1 - P_i^*) \varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{(\text{cv})}(\hat{m}_i, \alpha_i), \quad (17)$$

где

$$\varepsilon_{\mathbf{B}_i}^{(\text{cv})}(\hat{m}_i, \alpha_i) = \text{FRR}(\hat{m}_i, \alpha_i),$$

$$\varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{(\text{cv})}(\hat{m}_i, \alpha_i) = \text{FAR}(\hat{m}_i, \alpha_i)$$

— доли ложных отказов и ложных распознаваний при предъявлении покрытию с параметрами (\hat{m}_i, α_i) «своих» (\mathbf{B}_i) и «чужих» ($\mathbf{B} \setminus \mathbf{B}_i$) объектов из обучающего множества \mathbf{B} , а P_i^* и $(1 - P_i^*)$ — оценки вероятностей «своих» и «чужих» объектов среди общего числа предъявлений. В качестве оценок P_i^* выбираются вероятности P_i , $i = 1, \dots, c$, в (7). При условии, что эти вероятности одинаковы, $P_i^* = P_{\text{own}}/c$ и $1 - P_i^* = 1 - P_{\text{own}}/c$. Оптимизация покрытия кластера \mathbf{B}_i сводится к нахождению пары (\hat{m}_i, α_i) , которая доставляет наименьшее значение функции (17). Для оптимизированных покрытий кластеров оценка вероятности ошибок обучения TSC-классификатора определяется средним значением

$$\varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \frac{1}{c} \sum_{i=1}^c \varepsilon_{\mathbf{B}}^{(\text{cv})}(\hat{m}_i, \alpha_i). \quad (18)$$

Оценки вида (18), полученные при обучении по (k, q) -составляющим меры (4) и q -составляющим меры (5), использованы в (15) и (16) для вычисления соответствующих параметров меры.

Экспериментальные результаты распознавания лиц

Схема проведения экспериментов заключалась в следующем. Исходные объекты из сформированной базы лиц были преобразованы в HSI-изображения размера 800x600 пикселей. Трёхслойные представления информативных объектов строились с использованием $(L + 1)$ -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 объёму обучающие \mathbf{B} и тестовые $\mathbf{A} \setminus \mathbf{B}$ выборки, в которых семантически однородные кластеры имели одинаковые мощности $m_i = 20$, $i = 1, \dots, c$. Априорные вероятности «своих» и «чужих» объектов выбирались равными значениям: $P_{\text{own}} = 1; 0.75; 0.5$ и $P_{\text{alien}} = 1 - P_{\text{own}}$. При каждом фиксированном значении P_{own} , на пяти обучающих выборках строились TSC-классификаторы с помощью покрытий для $c' = \lfloor cP_{\text{own}} \rfloor$ произвольно выбранных кластеров. Для заданной обучающей выборки \mathbf{B} при каждом значении $c' < c$ строилось несколько реализаций таких классификаторов (при $c' = c$ одна реализация для каждой выборки \mathbf{B}). Предъявляя объекты тестовой выборки $\mathbf{A} \setminus \mathbf{B}$ для каждой реализации TSC-классификатора, вычислялась оценка вероятности ошибочных решений (13) и усреднённая оценка вероятности ошибки по всем реализациям классификатора и предъявляемым тестовым выборкам. Полученные для TSC-классификатора доли ошибок распознавания сравнивались с аналогичными показателями SVM-классификатора [5]. В качестве признаков SVM-классификатора использовались попарные расстояния по заданной мере между предъявляемым объектом и объектами обучающей выборки.

Усреднённые оценки вероятностей ошибок для обоих классификаторов представлены в таблице 1. Оценки качества SVM-классификатора приведены для случая $P_{\text{own}} = 1$. Из данных таблицы следует, что оба классификатора демонстрируют сопоставимые показатели качества и обеспечивают уменьшения доли ошибок за счёт использования многослойных представлений модели HSI по сравнению с ошибками по отдельным каналам H, S или I. Необходимо отметить, что применение древовидных представлений с многоуровневым разрешением позволяет организовывать в TSC-классификаторе иерархический поиск решений в многоуровневой базе эталонов. При большом числе классов

Таблица 1. Усреднённые оценки вероятности ошибок распознавания лиц для классификаторов типа TSC и SVM.

P_{own}	1	1	0.75	0.5
Классификатор	TSC	SVM	TSC	TSC
H	0.015	0.011	0.021	0.020
S	0.026	0.018	0.026	0.023
I	0.032	0.078	0.036	0.029
HSI	0.012	0.007	0.019	0.016

с стратегия иерархического поиска сокращает вычислительные затраты в $c / \log c$ раз по сравнению с переборным поиском.

Выводы

Предложен метод построения классификатора по критерию ближайшего эталона в пространстве многослойных древовидных представлений объектов, заданных многоканальными изображениями. Обучение классификатора состоит в отборе эталонов и реализовано на основе построения древовидно-структурированных покрытий семантически однородных кластеров обучающего множества шарами, центры и радиусы которых вычисляются по заданной мере на множестве представлений.

Эффективность предложенного метода продемонстрирована результатами распознавания цветных изображений лиц в трёхканальной модели HSI. Показано преимущество использования многоканальных изображений по сравнению с монохромными изображениями. Полученные оценки качества распознавания лиц сопоставимы с аналогичными показателями классификатора на основе метода опорных векторов.

В перспективе предполагается исследовать модификации древовидно-структурированных покрытий и другие схемы комплексирования многоканальных данных. Планируется также расширение состава биометрических источников.

Литература

- [1] Phillips P. J, Martin A., Wilson L. C, Przybocki M. An introduction to evaluating biometric systems // IEEE Computer, 2000. — Vol. 21, No. 2. — Pp. 56–63.
- [2] Гонсалес Р., Вудс Р., Эддингс С. Цифровая обработка изображений в среде Matlab. — Москва: Техносфера, 2006.
- [3] Zhao W., Chellappa R., Phillips P. J, Rosenfeld A. Face recognition: a literature survey // ACM Computing Surveys, 2003. — Vol. 35, No. 4. — Pp. 399–458.
- [4] Ganebnykh S. N, Lange M. M. Classification of 2D Grayscale Objects in a Space of the Multiresolution Representations // Pattern Recognition and Image Analysis, 2009. — Vol. 19, No. 4. — Pp. 591–602.
- [5] opencv.willowgarage.com — Open Source Computer Vision Library — 2010.