

Многослойное древовидное представление объектов многоканальных изображений*

Ланге М. М., Степанов Д. Ю.

lange_mm@ccas.ru

Москва, Вычислительный Центр РАН

На основе обобщения процедуры представления двумерных тел с полутоновой окраской предложен способ описания объектов многоканальных изображений стопками деревьев. Введена мера различия стопок деревьев. Используя пространство трёхслойных древовидных представлений для описания RGB-объектов, продемонстрирован выигрыш в распознавании лиц, заданных цветными изображениями по сравнению с полутоновыми.

Структурные методы представления данных составляют основу для сокращения вычислительной сложности алгоритмов обработки изображений и, в частности, алгоритмов распознавания образов. К таким методам относятся вейвлет-преобразования [1] и методы древовидной декомпозиции изображений [2, 3]. Для широкого класса объектов, заданных на полутоновых изображениях двумерными телами, в работе [4] предложен метод представления таких объектов деревьями эллиптических примитивов. Этот метод позволил построить быстрый классификатор в пространстве древовидных представлений объектов, который продемонстрировал высокие показатели качества распознавания жестов и подписей с вероятностью ошибки порядка 0,01. Естественное развитие этого подхода состоит в построении древовидных представлений для объектов, заданных несколькими (многоканальными) изображениями. Примерами многоканальных изображений являются полутоновые изображения объектов, снятых в различных ракурсах, а также цветные RGB-изображения. Использование многоканальных данных эквивалентно увеличению размерности пространства представлений и в соответствии с фундаментальной теоремой теории информации [5] должно привести к снижению вероятности ошибки распознавания, выполняемого по схеме декодирования.

Представление двумерных объектов стопками деревьев

Способ представления, предложенный в работе [4], переводит любой двумерный объект, заданный полутоновым изображением и имеющий в плоскости изображения идентифицируемую систему собственных координат, в завершённое бинарное дерево эллиптических примитивов. При достаточно малых размерах пикселя такие представления инвариантны к преобразованиям поворота, смещения и масштабирования объектов в плоскости изображений, а также к изменению уровня яркостной окраски объектов. Пример древовидного

представления информативного участка лица, выделенного на полутоновом изображении, показан на рис. 1. В этом представлении дерево содержит девять уровней, образующих пирамиду, в которой каждый уровень $l = 0, 1, \dots, 8$ содержит 2^l эллиптических примитивов.

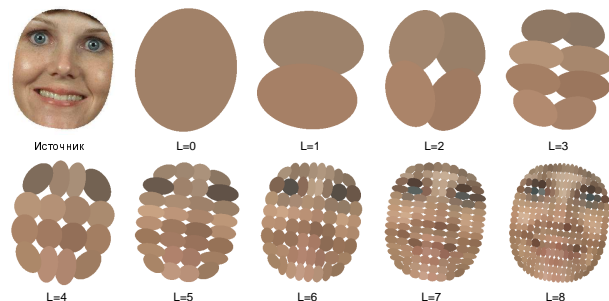


Рис. 1. Древовидное представление лица эллиптическими примитивами.

Формально древовидное представление любого объекта, выделенного на одиночном изображении, задано множеством эллиптических примитивов

$$A = \{a_n, 0 \leq n \leq n_{\max}\}, \quad (1)$$

где a_n — примитив, который соответствует в дереве вершине с номером n , находящейся на уровне $l = \lfloor \log_2(n+1) \rfloor$. Нумерация вершин-примитивов производится в соответствии с рекуррентным правилом $n \rightarrow (2n+1, 2n+2)$, в котором n -я вершина текущего уровня $l \geq 0$ порождает две вершины следующего уровня $l+1$. На уровне $l=0$ корневой примитив имеет номер $n=0$. Каждый примитив в дереве (1) описывается набором параметров, который включает вектор центра эллипса, единичные векторы его главных осей, радиусы вдоль главных осей и среднее значение яркостной окраски примитива. Указанные параметры задаются в собственных координатах объекта и нормализуются относительно параметров корневого примитива.

Для объектов, заданных N -канальными изображениями, представления вида (1) могут быть построены по каждому каналу независимо. В результате для любого объекта, имеющего на всех N изображениях идентифицируемые системы собствен-

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

ных координат, формируется многослойное представление в виде стопки деревьев

$$A^N = (A_1, \dots, A_N), \quad (2)$$

в которой каждый слой A_i , $i = 1, \dots, N$, представлен деревом примитивов вида (1). Примерами трёхслойных древовидных представлений являются стопки деревьев, которыми могут быть представлены цветные объекты, заданные RGB-изображениями. Слои в таких стопках являются независимыми древовидными представлениями объекта по R, G и B каналам.

Мера различия стопок деревьев

Для введения меры различия любой пары представлений (A, \hat{A}) вида (1) в работе [4] использовано пересечение деревьев $A \cap \hat{A}$, образованное всевозможными парами соответственных примитивов $(a_n \in A, \hat{a}_n \in \hat{A})$ с одинаковыми номерами n . Мера различия пары (a_n, \hat{a}_n) введена на множестве

$$\Omega = \{(a_n, \hat{a}_n) \in (A \cap \hat{A})\},$$

в котором по крайней мере один из примитивов пары (a_n, \hat{a}_n) является концевой вершиной соответствующего дерева. Указанная мера имеет вид

$$D(A, \hat{A}) = \sum_{k=1}^3 w_k \sum_{(a_n, \hat{a}_n) \in \Omega} d_k(a_n, \hat{a}_n) 2^{-\lfloor \log_2(n+1) \rfloor}, \quad (3)$$

где w_1, w_2, w_3 – нормированная тройка параметров ($w_1 + w_2 + w_3 = 1$), а функции $d_k(a_n, \hat{a}_n) \geq 0$ при $k = 1, 2, 3$ соответствуют различиям векторов центров ($k = 1$), векторов ориентации и радиусов ($k = 2$) и уровней яркости ($k = 3$) примитивов a_n и \hat{a}_n .

Различие i -х слоёв $A_i \in A^N$ и $\hat{A}_i \in \hat{A}^N$, $i = 1, \dots, N$, в любой паре стопок деревьев (A, \hat{A}^N) вида (2) определяется функциями $D(A_i, \hat{A}_i)$ вида (3). Поэтому для пары многослойных представлений (A^N, \hat{A}^N) вводится мера различия

$$D^N(A^N, \hat{A}^N) = \sum_{i=1}^N \alpha_i D(A_i, \hat{A}_i) \quad (4)$$

с нормализованными параметрами $\alpha_i : \sum_{i=1}^N \alpha_i = 1$. Коэффициенты w_k в (3) и коэффициенты α_i в (4) могут быть оценены на этапе обучения из условия минимизации вероятности ошибки распознавания. В частном случае, оценки коэффициентов α_i при $i = 1, \dots, N$ могут быть получены с использованием нормализованных дисперсий яркостей:

$$\sigma_i^2 = \frac{\sigma^2(A_i)}{\sum_{i=1}^N \sigma^2(A_i)}, \quad (5a)$$

$$\hat{\sigma}_i^2 = \frac{\sigma^2(\hat{A}_i)}{\sum_{i=1}^N \sigma^2(\hat{A}_i)}, \quad (5b)$$

где $\sigma^2(A_i)$ и $\sigma^2(\hat{A}_i)$ – дисперсии яркостей пикселей объектов, имеющих в i -ых слоях представления $A_i \in A^N$ и $\hat{A}_i \in \hat{A}^N$. С учётом соотношений (5), $\alpha_i = (\sigma_i^2 + \hat{\sigma}_i^2)/2$, $i = 1, \dots, N$. В случае RGB-изображений возможен выбор одинаковых коэффициентов $\alpha_i = 1/3$, $i = 1, 2, 3$.

Применение к распознаванию лиц по RGB-изображениям

Концепция многослойных древовидных представлений может найти применение для распознавания образов, заданных цветными изображениями, и, в частности, для распознавания лиц. В рамках настоящей работы выполнены эксперименты по выделению лиц на цветных изображениях, построению их представлений в виде стопок деревьев по RGB-каналам и распознаванию лиц в пространстве многослойных древовидных представлений по критерию ближайшего соседа, использующего меру различия вида (4). На этапе предобработки выполняется поиск координат глаз на изображении. Используя расстояние между центрами глаз r_0 строится система декартовых координат (U, V) с учётом симметрии лица относительно оси U и выбора оси V , параллельной линии глаз и смещённой влево вдоль оси U на величину $\beta_0 r_0$ с параметром $\beta_0 > 0$ (рис. 2).

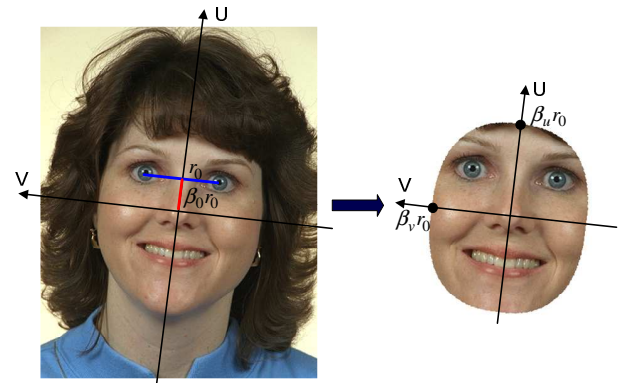


Рис. 2. Выделение информативной части лица на изображении.

В координатах (U, V) информативная часть лица задаётся областью, ограниченной овалом

$$\frac{U^m}{(\beta_u r_0)^m} + \frac{V^m}{(\beta_v r_0)^m} \leq 1 \quad (6)$$

с радиусами $r_u = \beta_u r_0$ и $r_v = \beta_v r_0$ по осям U и V , и параметром формы $m > 1$. Параметр смещения и параметры радиусов овалоида (6) выбирались равными: $\beta_0 = 0,5$, $\beta_u = 0,6$ и $\beta_v = 0,8$. Экспериментально найден наиболее подходящий параметр формы $m = 2,5$. Для объектов, соответствующих участкам лиц, ограниченных на RGB-изображениях фигурами (6), строились представления в виде трёхслойных стопок деревьев вида (2).

Эксперименты проведены с множеством цветных изображений, взятых из базы данных, размещенной на сайте [6]. Исходное множество изображений содержало 108 лиц (18 персон по 6 реализаций), которые были разбиты на две эквивалентные части, соответствующие обучающему и тестовому множествам, по три реализации каждой персоны в обоих множествах. В качестве эталонов использовались представления всех объектов обучающего множества. Распознавание проводилось для объектов тестового множества по критерию ближайшего эталона. Результаты в виде зависимостей доли ошибочных решений (в %) от числа уровней в используемых древовидных представлениях даны на рис. 3 для распознавания по полутоновым (а) и цветным (б) изображениям. В случае (а) использовалась мера вида (3), в случае (б) — мера вида (4) с параметрами: $w_1 = 0,1$; $w_2 = 0,2$; $w_3 = 0,7$. В мере (4) были взяты $\alpha_i = 1/3$ для каналов R , G и B ($i = 1, 2, 3$).

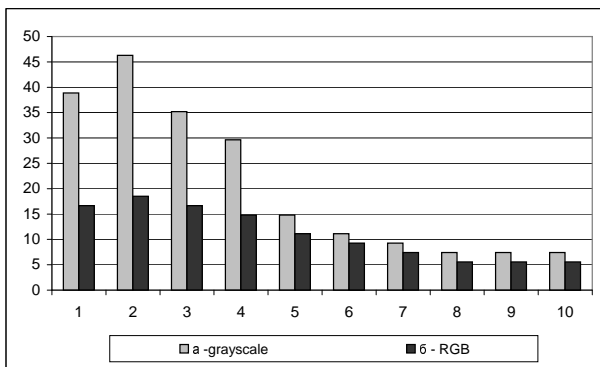


Рис. 3. Экспериментальные зависимости доли ошибок распознавания лиц от уровня древовидных представлений: а — полутоновые изображения; б — RGB-изображения.

Приведённые зависимости демонстрируют снижение доли ошибок при распознавании лиц в пространстве представлений RGB-объектов стопками деревьев по сравнению с пространством древовид-

ных представлений полутоновых объектов. Дополнительное снижение доли ошибочных решений может быть достигнуто за счёт оптимизации процедуры отбора эталонов и введения их сфер влияния.

Выводы

Предложенный способ древовидного описания объектов, заданных многоканальными изображениями, обладает свойством универсальности, что позволяет использовать его для построения иерархически структурированных представлений образов, порождаемых различными источниками. В частности, возможно построение таких представлений для образов, заданных цветными изображениями в стандартах HSI или HSB. Возможно также расширение алфавита признаков примитивов с целью более адекватного описания текстуры образов. Введённая мера различия стопок деревьев допускает обобщение для многослойных пирамидальных представлений. Указанное обобщение может быть использовано для уменьшения вычислительной сложности поисковых процедур в пространстве представлений с многоуровневым разрешением.

Литература

- [1] *Gonzalez R., Woods R.* Digital Image Processing. — New Jersey: Prentice Hall, 2002.
- [2] *Berretti S., Del Bimbo A.* Multiresolution spatial partitioning for shape representation. // IEEE Proc. of ICPR, 2007. — Vol. 2. — Pp. 775–778.
- [3] *Lange M., Ganebnykh S., Lange A.* Moment-based Pattern Representation Using Shape and Grayscale Features. // Lecture Notes in Computer Science — Berlin: Springer, 2007 — Vol. 4477. — Pp. 523–530.
- [4] *Ганебных С. Н., Ланге М. М.* Древовидное представление образов для распознавания полутоновых объектов. // Труды Вычислительного центра им. А. А. Дородницына РАН (отдельный выпуск). — М.: ВЦ РАН, 2007. — 32 с.
- [5] *Gallager R.* Information Theory and Reliable Communication. — New York: Wiley, 1968.
- [6] face.nist.gov/colorferet/— The Color FERET Database, USA. — 2003.