

# 人脸图像检测及分类系统的研究

陈茂林, 戚飞虎

(上海交通大学计算机科学与工程系, 上海 200030)

**摘 要:** 本文提出了一个人脸图像检测及分类系统的快速算法. 任意输入图像, 系统首先通过查询颜色表的方法快速提取出人脸区域. 在姿势分类时, 本文使用了分类压缩人脸本征空间的方法, 降低了对人脸图像训练和分类的计算量. 实验结果证明, 本算法具有很好的检测和分类性能.

**关键词:** 人脸检测; 颜色模型; 本征脸; 视分类

中图分类号: TN391.4 文献标识码: A 文章编号: 0372-2112 (2001) 12A-1917-03

## Research on the Human Face Detection and Its Classification

CHEN Mao lin, QI Fei hu

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract:** In this paper a fast algorithm is proposed to detect and classify human faces. Querying the color table, the system firstly selects out the face area from the image. During the pose classification each eigenspace is compressed individually to reduce the computation of training and classification of faces. Experiments testify the good performance of detection and classification.

**Key words:** face detection; color model; eigenface; view classification

### 1 引言

本文提出了一个人脸图像检测及分类系统的快速算法. 它能自动检测和定位人脸图像, 并将人脸图像从源图像中分割出来. 由于图像中的人脸可能处于各种不同的姿势, 它又进一步将人脸按照姿势进行分类得到三种视, 即正面视、左侧面视和右侧面视. 因为在各种应用中需要不同视的样本或者需要得到给定输入的人脸图像的视.

当前人脸检测方法在国际上得到了广泛的研究<sup>[1-6]</sup>. 现有的人脸检测系统主要是利用了人脸形状、面部特征点、人脸模式和人脸颜色特征. 其中, 利用人脸颜色特征的多是实时性较强的系统. 人脸分类主要是利用人脸面部特征和神经网络的方法. 前者是利用人脸面部特征的对称性, 如眼部特征. 后者则是用大量的样本训练生成不同视下的子网, 进行人脸视的识别. 本文介绍的方法是利用了颜色的统计信息来进行人脸区域的快速检测, 然后用分类压缩本征空间的方法得到不同视下的本征模型来识别人脸视. 本征模型方法利用人脸总体的结构特征, 不象人脸形状、面部特征点方法那样过分依赖于低层局部特征的检测, 其抗噪声性能较强.

### 2 人脸检测

使用皮肤的颜色模型, 就可以从图像中检测出似皮肤的像素. 因为对于皮肤来说, 其似皮肤像素会聚类一起形成一个区域, 而不是象复杂背景中的似皮肤像素是离散的不能聚类

成一个族. 因此, 该方法可以应用于复杂背景的场所. 而且基于颜色统计的方法不受皮肤区域的形状变化的影响.

#### 2.1 颜色空间模型<sup>[2]</sup>

利用颜色信息的一个缺点是, 其受光照变化的影响很大. 当光照变化时, 图像的像素值也随着变化. 同一目标在不同光照下得到的图像差别很大. 所以, 选择合适的颜色空间, 来降低这种影响就非常重要. 本文使用了 HSV 颜色空间, 它能够将亮度信息和色度信息从输入图像中分离出来并单独表示. 在 HSV 颜色模型中, H 分量代表图像的纯色彩信息, S 代表图像的饱和度信息, V 代表图像的亮度信息. H 分量只表示目标的彩色信息, 受光照变化的影响缓慢, 如果只利用它所代表的信息, 就降低了光照的影响作用. 所以第一步是要将源 RGB 三基色分量转换到 HSV 颜色空间中.

#### 2.2 皮肤颜色模型<sup>[4, 5]</sup>

人类某一人的皮肤颜色分布在一个较窄的频带上. 通过对大量的人体皮肤区域的统计分析, 可以得到其肤色的分布范围, 以及在该范围内各个点上的概率, 由此可以构成肤色分布的颜色表. 在颜色表的每一个入口, 表示该像素值在皮肤图像中出现的概率. 在后期的输入图像中检测皮肤区域时, 就是通过快速查表的方法, 得到源图像中各像素点属于皮肤的概率, 所有像素点的概率组织在一起, 成为新的矩阵, 我们称为相似图.

设用来训练颜色模型的有 S 幅样本皮肤图像, 其宽度和

高度分别是  $M_j, N_j$ , 其颜色分布模型(*Skin Color Model, SCM*)如式(1)所示.

$$P(i) = \frac{f_i}{\sum_{j=1}^s M_j N_j} \quad (1)$$

其中,  $P(i)$  表示颜色表中第  $i$  个颜色量值出现的概率,  $f_i$  表示第  $i$  个颜色量值的出现频度. 在该表中, 峰值对应于皮肤图像中出现次数最多的像素值, 并且表中的数值越大, 该点像素值出现概率就越大. 检查输入图像中的各像素, 就可以得到每个像素对应的皮肤相似度, 并将这些相似度值组成一个新相似图矩阵, 此图中, 对应于皮肤区域的块, 其值很大; 对应于非皮肤区域的块, 其值很小. 对于复杂的背景, 背景图像中存在似皮肤的离散像素. 由于其离散度很大(可连通度低), 不能聚成一区域块, 可以通过滤波的方法去除, 这里不再详细叙述.

### 2.3 基于二维直方图的人脸分割方法

在皮肤区域的相似图中, 在有人脸出现的地方是一些离散的大于零的数值, 它代表该点像素和皮肤颜色的相似程度. 为了将这些人脸区域分割出来, 我们使用了基于二维直方图的分割方法. 通过向  $x$  轴和  $y$  轴方向投影的方法, 得到皮肤区域的最小包围矩形框. 在有目标出现的地方, 其一维直方图应该成单峰状态, 其它地方是零区. 通过几次投影, 可以得到目

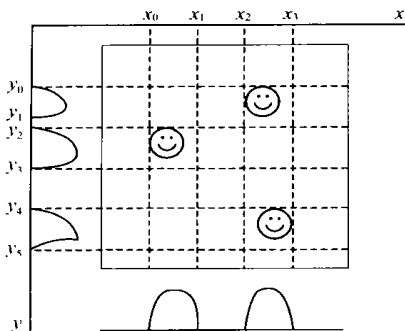


图 1 相似图中似皮肤区域的分割方法示意图

标所在的矩形坐标. 如图 1 所示, 为人脸域分割示意图.

## 3 人脸视分类<sup>[7]</sup>

经过以上颜色信息统计检测方法可以得到人脸图像区域. 由于人脸有其固有的结构特征, 不同方向视下, 其本征模型不一样, 所以我们通过建立不同视下的本征脸模型的方法, 来最终确定人脸目标的方向视.

### 3.1 本征方法

假设含有各类样本的目标图像有  $N$  个, 用向量表示为  $X_1, X_2, \dots, X_N$  (向量维数设为  $L$ ),  $X_{ave}$  为平均图像,  $X'_i$  表示每幅图像的均差,  $C$  为其协方差矩阵. 求出  $C$  的本征值和对应本征值, 分别记为  $\lambda_k$  和  $u_k$ . 这些求出的本征向量所形成的向量空间就可表示这些图像的主要特征信息. 将训练所用的  $N$  个图像均向此空间投影, 得到各自的投影向量  $Y_1, Y_2, \dots, Y_N$ , ( $Y_i = [y_{i1} y_{i2} \dots y_{iL}]$ ),  $i = 1, 2, \dots, N$ :

$$y_{ij} = (u_i)^T X'_j, j = 1, 2, \dots, L \quad (2)$$

对于待识别图像  $X$ , 将其与  $X_{ave}$  先求差, 然后向本征空间

中投影得到投影向量  $Y$ :

$$y_j = (u_i)^T (X - X_{ave}), j = 1, 2, \dots, L \quad (3)$$

将式(3)中得到的投影向量  $Y$  与  $N$  个训练样本图像对应的投影向量  $Y_1, Y_2, \dots, Y_N$  比较, 按照一定的距离准则完成识别.

### 3.2 本征空间的分类压缩方法

将各类样本混合在一起, 训练所得到的本征空间, 是一个高维空间, 其计算量非常大. 为了减少计算量, 本文采用了分类空间压缩的方法对以上的本征方法进行了改进, 将参与训练的各种视下的样本图像分别进行训练, 得到各类别样本的本征空间  $\Omega$ . 我们将它的本征向量按照对应的本征值的大小进行重新排序, 选取排在前面  $k$  个大本征值所对应的本征向量构成本征空间, 所以它是一个降维的空间:

$$\Omega = \left\{ \left\{ X_{ave}^j \right\}, \left\{ U_i \right\}; j = 1, 2, \dots, C \right\}$$

$j$  为类别号,  $C$  为总的类别数

$k$  满足如下条件:

$$\frac{\sum_{i=0}^k \lambda_i}{\sum_{i=0}^{M-1} \lambda_i} \geq a \quad (4)$$

其中,  $a = 99\%$ , 这保证样本集在前  $k$  个轴上的能量占整个能量的 99% 以上.

### 3.3 视判断

对于新的输入  $I, X$  为其图像矩阵, 其概率密度为式(5):

$$p(I|\Omega) = \frac{\exp\left[-\frac{1}{2}(X - \bar{X})^T \Sigma^{-1}(X - \bar{X})\right]}{(2\pi)^{N/2} |\Sigma|^{1/2}} \quad (5)$$

Mahalanobis 距离定义为式(6):

$$\begin{aligned} d(X) &= (X - \bar{X})^T \Sigma^{-1} (X - \bar{X}) = \tilde{X}^T [\Phi \Lambda^{-1} \Phi^T] \tilde{X} \\ &= y^T \Lambda^{-1} y = \sum_{i=1}^N \frac{y_i^2}{\lambda_i} = \sum_{i=1}^k \frac{y_i^2}{\lambda_i} + \sum_{i=k+1}^N \frac{y_i^2}{\lambda_i} \end{aligned} \quad (6)$$

因此, 概率密度值完全取决于 Mahalanobis 距离. 判断  $I$  属于哪个子空间, 可以通过比较和选取具有最大概率密度值的特征空间, 即和  $I$  具有最小 Mahalanobis 距离的空间.

## 4 实验结果

### 4.1 人脸检测

在皮肤颜色模型训练阶段, 我们采用手工分割的方法得到 100 幅皮肤图像作为训练样本. 这里需要特别指出的是, 手工分割出来的人脸图像中, 不包含眼睛和眉毛部分, 参考文献中都忽略了这一点. 首先将它们从 RGB 空间变换到 HSV 空间中, 然后对其 H 分量进行统计, 可以得到皮肤的颜色模型. 由于图像的过亮或者过暗, 都不能正确的表达皮肤的颜色, 所以参加统计的像素必须满足式(7).

$$V_{min} < V < V_{max} \quad (7)$$

其中  $V_{min}$  和  $V_{max}$  是常数, 对应于亮度过暗和过亮的情况.

我们使用了 3000 幅含有人脸的图像(不包含人体其它皮肤区域), 作为检测输入, 检测正确率为 99%.

### 4.2 人脸分类

为了选取合适的样本,这里先对三种视进行定义.将人脸正对摄像机,然后左右分别作 90 度的 pan 运动,这样就得到一个 180 度张角的视空间.将此空间均分成三等份,则得到的视空间即分别为左侧面视、正面视和右侧面视空间.为了训练本征模型,选取的样本来自于三个不同的视,分别含有 50 幅.首先将他们大小归一化为  $24 \times 28$ ,再分别训练得到三个不同视下的本征空间.在多视识别时,共选取了 657 幅样本图像,其中正面视包含 208 幅样本图像,左侧面视包含 221 幅样本图像,右侧面视包含 228 幅样本图像.在经过上面的人脸检测和人脸图像归一化后,再进行视判断.实验结果为,正面视识别率为 98.5%,左侧面视识别率为 97.2%,右侧面视识别率为 97.6%.

## 5 讨论

本文在颜色信息统计的基础上,首先检测出输入图像中的人脸图像,然后用本征投影的方法,将人脸进一步分类成三个方向视.实验结果表明,本文的方法是有效的.颜色统计模型虽然易受到光照的干扰,但是如果采用 HSV 颜色空间模型,这种干扰将得到很好的抑制.与底层特征匹配的方法不同,颜色的统计特征具有相对的稳定性,它不随人脸的表情和平面内旋转而变化,而且能实现快速算法.为了进一步将人脸分类成不同的方向视,本文在对本征空间进行分类压缩的基础上,采用了本征空间投影距离度量方法.采用分类压缩的优点是,大幅度的降低的计算量.综上所述,本文介绍的快速算法具有很大的应用价值.

### 参考文献:

- [ 1 ] V Govindaraju. Locating human faces in photographs [J]. Int. J. Comp. Vision, 1996, 19( 20 ): 129- 146.

- [ 2 ] J Cai, A Goshtasby. Detecting human faces in color images [J]. Image and vision computing 1999, 18: 63- 75.
- [ 3 ] Haiyuan Wu, Qian Chen and Masahiko Yachida. Face detection from color images using a fuzzy pattern method [J]. IEEE Trans., 1999, PAMF 21( 6 ): 557- 563.
- [ 4 ] Toshiaki Kondo, Hong Yan. Automatic human face detection and recognition under non uniform illumination [J]. Pattern Recognition, 1997, 32: 1707- 1718.
- [ 5 ] Douglas Chai, King N. Ngan. Face segmentation using skin color map in videophone applications [J]. IEEE Trans. circuits and systems for video tech., 1999, 9( 4 ): 551- 564.
- [ 6 ] M A Turk, A P Pentland. Face recognition using eigenfaces [A]. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition [C], Hawash, 1991: 586- 591.
- [ 7 ] Liu Rujie, Yuan Baozong. Eigenspace based human face detection [A]. Proceedings of ICSP Beijing [C], 2000: 1305- 1308.

### 作者简介:



陈茂林 男, 1974 年 10 月生于江苏泰州. 现为上海交通大学计算机系博士研究生. 主要研究方向为视频内容分析, 智能目标检索、目标识别和机器学习.

戚飞虎 男, 1938 年 12 月生于上海. 现为上海交通大学计算机系博士生导师. 主要研究方向为视频处理和神经计算. 已在国内外公开发表学术论文百余篇.