

# 彩色序列图像的人脸检测和识别系统

凌旭峰, 杨 杰, 叶晨洲

(上海交通大学图像处理与模式识别研究所, 上海 200030)

**摘 要:** 文中实现了彩色序列图像的人脸检测和识别的系统. 所用到的关键性技术是肤色运动分析、主成分分析(PCA)和支撑向量机(SVM). 首先根据彩色序列图像中人脸的色度特性、运动特性、几何特性和灰度分布特性完成人脸的检测工作; 其次通过人脸模式之间的相关性进行主成分分析, 提取并且选择特征, 将所选择的特征训练 SVM, 最后用已经训练好的 SVM 完成对人脸的识别任务. 系统算法结构遵循以下原则: 先使用运算量少的简单方法尽可能减少搜索空间, 然后在已经大大减少的空间中再用复杂方法处理, 可以在保持高的检测和识别率的同时, 提高系统响应速度.

**关键词:** 人脸检测; 人脸识别; 肤色分析; 运动分析; 主成分分析; 支撑向量机

**中图分类号:** TP18

**文献标识码:** A

**文章编号:** 0372-2112 (2003) 04-0544-04

## Face Detection and Recognition System in Color Image Series

LING Xu-feng, YANG Jie, YE Chen-zhou

(Institute of Image Processing & Pattern Recognition, Shanghai Jiaotong Univ., Shanghai 200030, China)

**Abstract:** A human face detection & recognition system on color image series is implemented in this paper. The key technologies involved are skin hue analysis, motion analysis, primary component analysis (PCA) and support vector machine. The human faces can be detected through their hue, motion, geometry and grayscale distribution features. PCA analysis can be done through the relation between face patterns, and then the features can be extracted and selected. Using the selected features to train multiple SVMs, which can finally classify human faces. The system structure is according to the following principle: Firstly simple methods are used to reduce the search space, and then more complicated methods are used in the reduced space. So the system can have a quick response speed as well as holding high detection and recognition rate.

**Key words:** Face detection; face recognition; skin analysis; motion analysis; primary component analysis; support vector machine

## 1 引言

人脸检测和识别作为模式识别的一个特定内容, 有着非常广泛的应用前景. 由于彩色序列图像的人脸检测和识别具有很大的实用价值和应用前景, 所以本系统基于彩色序列图像. 人们日益认识到人脸检测的重要性, 相关的研究工作也日益增多, 目前已有的人脸检测方法主要有: 灰度处理方法<sup>[1]</sup>, 色度图模糊技术方法<sup>[2-4]</sup>, 运动估计、运动分割方法等.

传统的人脸识别方法有最近邻法、欧氏距离法、马氏距离法和神经网络方法等. 人脸模式的高维特性使得难以对其建模描述; 人脸作为非刚体使得该模式离散性大; 人脸模式具有很强的非线性. 所以到目前为止, 人们还没有找到一种稳健的人脸识别方法.

相对于其维数而言, 人脸识别是一个小样本问题. 对于人脸识别这个小样本问题, 传统的分类方法一方面容易出现过

学习(overfitting)现象, 算法的推广性差; 并且学习性能差, 无法胜任人脸分类这个非线性很强的模式. 支撑向量机是为了解决小样本问题学习和分类提出的, 一方面可以克服神经网络等方法所固有的过学习和欠学习问题, 另外一方面又有很强的非线性分类能力<sup>[5]</sup>.

## 2 人脸检测

### 2.1 肤色检测

**UCS 颜色体系** 在大多数情况下, 图像信息是以 RGB 的颜色体系保存. 然而在人脸肤色分析中, 由于 RGB 颜色的 R、G、B 三个颜色分量都包含亮度信息, 存在相关性, 不适合进行色度处理. 所以在进行皮肤区域检测之前, 要将以 RGB 颜色体系转换到 UCS 颜色体系中. UCS 颜色体系是 CIE(国际照明委员会)提出的一种均匀色标体系<sup>[6]</sup>, 它与 RGB 基色体系存在着以下关系:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.490 & 0.310 & 0.200 \\ 0.177 & 0.813 & 0.011 \\ 0.000 & 0.010 & 0.990 \end{bmatrix} \begin{bmatrix} R_{CIE} \\ G_{CIE} \\ B_{CIE} \end{bmatrix} \quad (1)$$

$$u = 4X / (X + 15Y + 3Z) \quad v = 6Y / (X + 15Y + 3Z) \quad (2)$$

UCS 体系中的  $Y$  分量表示亮度,  $u$ 、 $v$  表示色度,  $u$ 、 $v$  分量和  $Y$  分量是独立的。

**肤色 uv 概率曲面** 由于  $u \in [0, 1], v \in [0, 1]$ , 我们在统计了大量的皮肤点(约 10,000 个)之后, 将  $u$  和  $v$  所在的  $[0, 1]$  区间等分为 100 份, 用一个  $100 \times 100$  矩阵来表征概率分布, 考虑函数:

$$P(x, y) = n((i-1) < x \leq i, (j-1) < y \leq j) / N \quad (3)$$

式(3)中  $n((i-1) < x \leq i, (j-1) < y \leq j)$  表示在  $i, j$  区间中皮肤点的个数,  $N$  为总的皮肤点的数目, 并且  $0 \leq i \leq 100, 0 \leq j \leq 100$ 。

100。我们把  $P(x, y)$  称作肤色  $uv$  概率曲面, 它表征了肤色在  $uv$  色度空间中的概率分布情况。研究  $uv$  概率曲面, 我们可以看出, 肤色在  $uv$  颜色空间  $u \in [0.33, 0.39], v \in [0.22, 0.28]$  处集中分布。

**皮肤区域检测** 我们将  $uv$  概率曲面当作一个映射函数, 则对于 RGB 彩色图像, 将其映射成为一幅 256 级灰度的概率图像。灰度值越高的像素点, 是皮肤点的概率越大, 反之概率越小。选取一个恰当的门槛  $T_s$  (门槛缺省为 128), 对灰度图像二值化; 在二值图像上通过区域生长法提取出皮肤区域, 再使用形态学滤波器来过滤区域噪声点; 最后可以得到图像上  $k$  的个皮肤区域  $r_1, r_2, \dots, r_k$ , 我们将皮肤区域的集合称为  $R_c$ 。并且  $R_c = R_{cm} \cup R_{cs}$ ,  $R_{cm}$  是运动皮肤区域,  $R_{cs}$  是静止皮肤区域。

## 2.2 运动检测和分析

图像运动分析是在灰度图像上进行的, 将 RGB 彩色图像转换成 256 级灰度图像后, 用图像序列中相邻时间的两幅图像差值来反映运动信息。图像差分运算定义如下:

$$f_{diff}(x, t_1, t_2) = f(x, t_1) - f(x, t_2) \quad (4)$$

式中  $f_{diff}$  是差分图像,  $f(x, t_1)$  是图像序列中  $t_1$  时刻图像,  $f(x, t_2)$  是图像序列中  $t_2$  时刻图像。式中运算只涉及到对应像素的相减运算, 运算量小。差分图像具有边缘图像的性质, 边缘信息不完整。

在所检测出的所有皮肤区域  $R_c$  中统计对应的差分图像  $f_{diff}$  的均值  $m_i$ :

$$m_i = \frac{1}{N} \sum_{x \in r_i} f_{diff}(x) \quad (5)$$

式中  $r_i$  为  $R_c$  中的第  $i$  个皮肤区域,  $N$  为区域  $r_i$  像素点的个数,  $m_i$  为区域  $r_i$  差分图像均值。

选取恰当的门槛  $T_m$  (缺省设为 8, 该门槛取决与运动速度和关照条件, 可以动态设置), 当区域  $r_i$  的  $m_i$  小于  $T_m$ , 则认为该区域为静止区域, 反之, 则认为是运动区域。通过运动检测可以从  $R_c$  中去除  $R_{cs}$ 。

## 2.3 人脸验证

**根据人脸几何特性验证** 人脸区域的类似椭圆, 检测运动皮肤区域集合  $R_{cm}$  的元素  $r_i$  的形状是否为椭圆, 去除  $R_{cm}$  中非人脸皮肤区域。实现方法是: 首先填充  $r_i$ , 使其成为一个非

凹区域  $r_i$ 。如果  $|S(r_i) - S(r_i)| / r_i > 0.2$ , 则可认为  $r_i$  非椭圆, 式中  $S(r)$  为区域  $r$  的面积; 在  $|S(r_i) - S(r_i)| / r_i \leq 0.2$  的情况下, 如果  $l_i^2 / 4S(r_i) > 4.25$ , 则认为是人脸区域, 否则, 认为非人脸区域, 式中  $l_i$  为曲线的周长。

**基于 SVM 的人脸验证方法** 系统用 1,400 个正样本 (70 个人, 每个人 20 张照片), 同时随机抽取了 1,400 个负样本 (非人脸样本) 来训练一个 SVM。每个训练样本是  $19 \times 19$  的 256 级灰度像素点阵, 由于人脸可以区域是一个椭圆形状, 所以在  $19 \times 19$  点阵的左下和右下部分可以不予以考虑, 这样 361 个输入点可以减少为 283 个。这样, 每一个训练样本作为一个 283 维向量, 来训练 SVM。训练好的 SVM 可以用来验证人脸, 如果输入是人脸图像, 则 SVM 输出 +1, 如果非人脸图像, 则输出 -1。

## 2.4 算法流程和实验结果分析

对于序列图像的第一帧图像, 必须进行全图遍历搜索。首先进行肤色分析, 去除图像中颜色与肤色不一致的区域, 保留与肤色相近的区域。接着进行运动检测, 去除图像中静止区域。结合运动检测和肤色识别, 保留运动并且颜色类似肤色的区域。以上处理有虚检的可能, 有可能包含了非人脸区域, 比如运动着手的手等。

对于所有的人脸候选区域, 我们通过几何特性分析 (区域是否类似椭圆等等) 进行第一步验证, 对于使用几何特性无法排除的人脸区域, 通过支撑向量机进行最终验证。该 SVM 输出 +1 则表示为人脸区域, -1 则表示判决为非人脸区域。

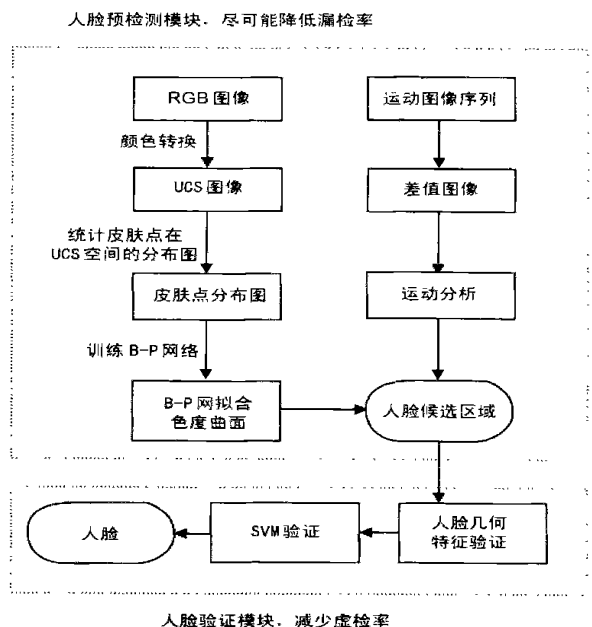


图 1 人脸检测模块流程图

不同的场景下, 我们拍摄了大量的图像序列进行测试。第 1 帧检测计算量大, 第 1 帧检测成功后, 由于运动的连续性, 仅需在人脸位置的周围搜索, 容易地检测第 2 帧、第 3 帧的人脸图像。图 2 显示了每个步骤的结果: 采用以上算法, 在彩色序列图像中人脸检测的成功率在 98% 以上, 并且可以检测出

序列图像中的多个人脸.

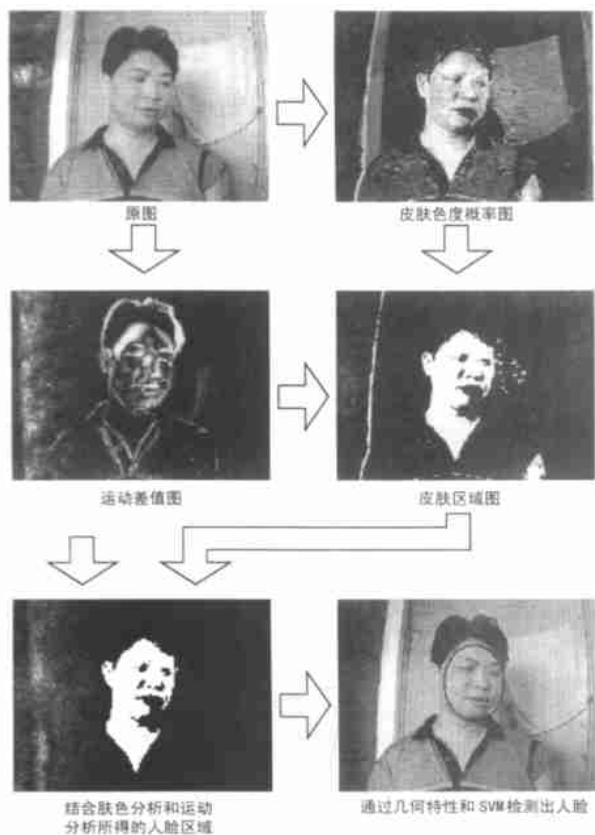


图2 人脸检测结果示意图

### 3 人脸识别

#### 3.1 人脸数据库

文中采用了上海交通大学图像所的人脸数据库. 该数据库包含了 70 个人, 每个人有多于 20 幅的图像. 所有图像在统一的室内光照条件下拍摄. 该图像数据库清晰度高, 不仅可以进行人脸检测、识别, 也满足人脸表情分析的要求.

#### 3.2 使用 PCA 方法进行特征提取和选择

人脸是一个高维的视觉模式, 直接对人脸图像进行识别分类是不可想象的. 本文采用离散 KL 变换<sup>[7]</sup>方法来对人脸图像进行降维. 将每一幅人脸图像行堆叠成为长矢量, 进行以下计算:

$$X = \frac{1}{M} \sum_{i=1}^M [x_i - m][x_i - m]^T / M \quad (6)$$

式中  $X$  为协方差矩阵,  $x_i \in R^N$  为每一幅人脸图像,  $m$  为均值. 求解矩阵  $X$  的特征值和特征向量, 并且由大到小排序. 选择对应于特征值最大的 20~40 特征向量 (又成为特征脸), 就可以达到总能量 70~80% 左右. 每一幅人脸图像都可以表示成特征脸的线性组合, 于是每一幅人脸图像都可以由特征脸的权值来表示. 人脸图像就被降维成一个 20~40 维的矢量.

因为离散 KL 变换没有快速算法, 为了提高减少系统的实时性, 并且使系统更易于扩充性, 文中将 70 个人的图像库分成了 3 个子类: 女性子类 (22 个人), 长脸子类 (22 个人), 圆脸子类 (26 个人). 离散 KL 变换是分别针对每个子类进行的, 不是对每一个人进行的.

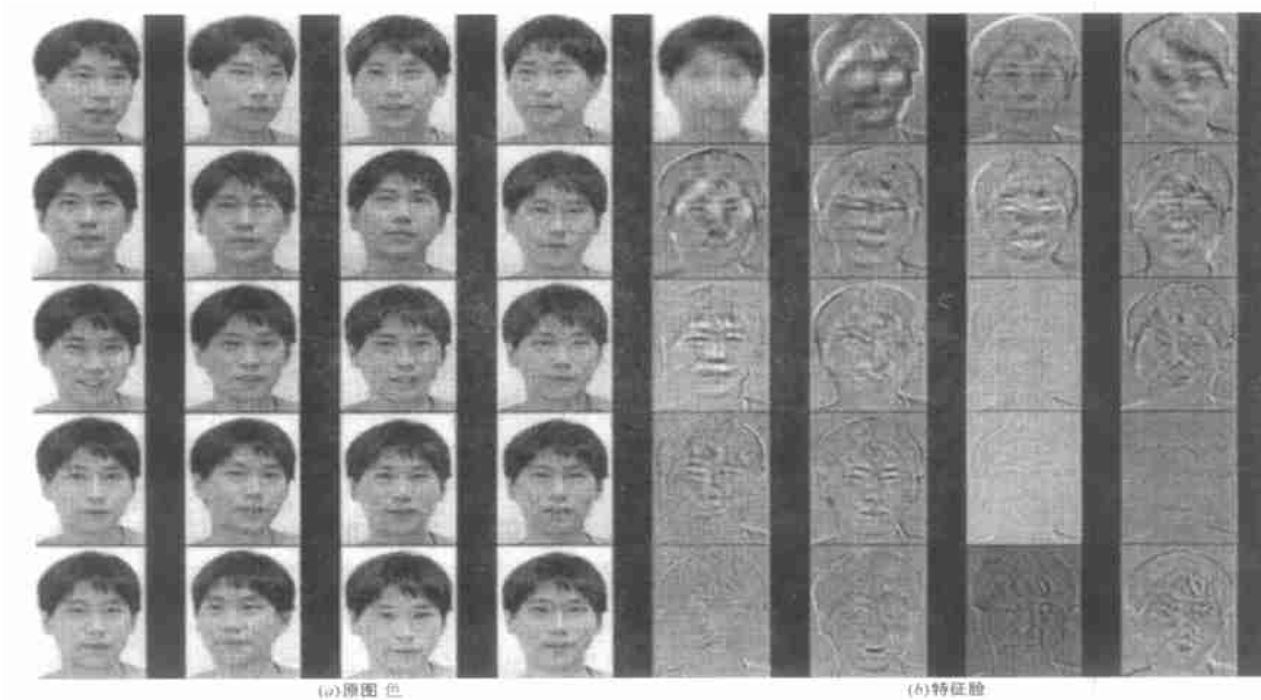


图3 人脸图像及其特征脸, 右边第一幅为均值图像

### 3.3 SVM 的训练

由于实际条件的限制,人们无法针对每个人都采集大量的图像样本.相对于其维数而言,人脸的样本数很少,是一个小样本问题.对于人脸识别这个小样本问题,传统的分类方法一方面容易出现过学习(overfitting)现象,导致算法的推广性差;另一方面就是学习性能差,无法胜任人脸分类这个非线性很强的分类模式.支撑向量机是为解决小样本问题学习和分类提出的,一方面可以克服神经网络等方法所固有的过学习和欠学习问题,另外一方面又有很强的非线性分类能力.但是普通的 SVM 只能对两类进行分类,不能区分多类.为了解决人脸分类这个多类问题,我们针对每一个人训练一个 SVM.首先假定人脸样本库中包含  $n$  个人,具体实现如下所示:

① 从人脸样本库中取出第一个人的所有样本,将其标为类 1;将其他所有人的所有样本标为类 2;将标好分类序号的所有样本作为输入样本,来训练一个 SVM,得到相应的支撑向量和相应的分类平面;将该支撑向量机序号标为 1,表明 1 号 SVM 是用来区分第一个人和其余人的.

② 从人脸样本库中取出第二个人的所有样本,将其标为类 1;将其他所有人的所有样本点标为类 2;接下来的步骤与上一步的操作一样;最后将训练好的 SVM 序号标为 2,该 SVM 是用来区分第二个人和其他所有人的.

③ 重复以上步骤,直到遍历人脸样本库中的所有入.最后得到了  $n$  个支撑向量机.

### 3.4 使用 SVM 分类人脸

SVM 可以将人脸样本分为  $n$  类,而不是只能将人脸分为两类.在分类时,当输入一个新的测试样本,同时将该测试样本交由  $n$  个 SVM 进行分类,则会出现以下几种情况:

① 第  $i$  个 SVM 将该样本分到类 1 中,其余的 SVM 将其分到类 2 中.则判断该样本属于人脸样本库中的第  $i$  个人.

② 同时有多个 SVM 将该样本分到类 1 中,而其余的 SVM 将其分到类 2 中,则判断为错分.

③ 所有 SVM 将该样本分到类 2 中,则判断该样本不属于人脸样本库.

### 3.5 人脸识别实验结果讨论和分析

由于 SVM 对于小样本问题,具有很好的学习性和推广性.与传统的方法相比,该方法显示出了更高的识别率.如下表所示:

表 1 SVM 和传统方法对照表

| 分类方法  | 女性子类正确率 | 长脸子类正确率 | 圆脸子类正确率 |
|-------|---------|---------|---------|
| SVM   | 0.9609  | 0.9649  | 0.9745  |
| 最近邻法  | 0.9348  | 0.9079  | 0.9562  |
| 欧氏距离法 | 0.9261  | 0.8965  | 0.9343  |
| 马氏距离法 | 0.9261  | 0.8965  | 0.9432  |
| 神经网络法 | 0.9410  | 0.9221  | 0.9218  |

上表是在每个人脸图像用 20 维向量表征时的分类结果.可以看出,SVM 方法相对于传统的最近邻法、欧氏距离法、马氏距离法和神经网络法,具有更高的人脸识别率.同样,当人

脸图像由 25 和 30 维表征时,SVM 显示了将近 100% 的识别率,而传统方法的识别率为 96% 左右.实验中,我们选取不同核函数的 SVM 进行了比较,结果发现当选取特征脸数量少时,核函数为 2 阶多项式的 SVM 具有最好的性能,识别率高并且支撑向量数少.

## 4 结论和展望

本文采用肤色分析结合运动分析的方法来检测人脸,并且对所检测出的人脸图像进行验证,达到了很高的人脸检测率.采用主成分分析方法提取特征,然后用 SVM 进行识别,也取得了优于传统方法的识别率.该系统的优势在于两方面:一是系统总运算量少,系统响应速度快;二是针对人脸模式复杂,样本数量少的特点,选用了支撑向量机来实现人脸识别,取得了更高的识别率.下一步的工作将要提高人脸检测和识别的稳健性.在不同光照条件下,并且在复杂背景的情况下,使得该系统具有相对稳定的人脸检测和识别功能.

### 参考文献:

- [1] Rowley H, Baluja S, Kanade T. Neural Network-Based Face Detection [J]. IEEE Trans on PAMI. 1998, 20(1): 23 - 37.
- [2] 卢春雨, 张长水, 闻芳, 阎平凡. 基于区域特征的快速人脸检测方法 [J]. 清华大学学报(自然科学版), 1999, 39(1): 101 - 105.
- [3] 闻芳, 周杰, 张长水, 李衍达. 基于局部线性映射神经网络的亮度补偿的彩色人脸检测 [J]. 清华大学学报(自然科学版), 1999, 39(7): 37 - 40.
- [4] Cai J, Gshtasby A. Detecting Human Faces in Color Images. Image Vision Computing [J]. 1999, 18: 63 - 75.
- [5] Vapnik V. 统计学习理论的本质 [M]. 北京: 清华大学出版社, 2000.
- [6] 崔屹. 图像处理技术与应用 [M]. 北京: 电子工业出版社, 1997.
- [7] 李金宗. 模式识别导论 [M]. 北京: 高等教育出版社, 1994.

### 作者简介:



凌旭峰 男, 1971 年生于安徽屯溪, 2001 年上海交通大学图像所博士毕业, 现任上海市信息化办公室互联网中心软件发展研究室副主任, 已经在国内外学术期刊和会议上发表论文 10 多篇, 攻读博士期间的主要研究方向为: 人脸识别、图像处理和人工智能, 目前的主要研究方向为: 视频图像、软件服务、软件产业等.

杨杰 男, 1964 年生于上海, 1994 年德国汉堡大学计算机系博士毕业, 现任上海交通大学图像处理与模式识别研究所副所长, 教授、博士生导师, 作为项目负责人已承担国家自然科学基金项目、国家 863 计划项目、国防 973 项目等 20 多项国家和省部级科研项目, 在德国出版专著 1 部, 在国内外学术期刊和会议上发表论文 70 多篇, 主要研究领域和方向: 目标检测和识别、数据融合和数据挖掘、智能理论及系统应用、医学图像处理.