

文章编号: 1007-2861(2007)05-0511-05

# 基于聚类肤色模型的人脸检测

丁友东<sup>1</sup>, 杜晓凤<sup>2</sup>, 李晓强<sup>1</sup>

(1. 上海大学 计算机工程与科学学院, 上海 20072; 2. 厦门大学 计算机科学系, 厦门 361005)

**摘要:** 在安全检测、人机交互等人脸识别应用中, 人脸的检测与定位是前期重要工作. 该文提出一种改进的基于亮度聚类的肤色模型, 并提出一种新的最佳阈值选择算法, 设计了肤色检测算法流程. 实验证明该算法提高了人脸检测的速度和鲁棒性.

**关键词:** 人脸检测; 肤色模型; 聚类

**中图分类号:** TP 391.4      **文献标识码:** A

## Human Face Detection Based on Cluster Skin Color Model

DING You-dong<sup>1</sup>, DU Xiao-feng<sup>2</sup>, Li Xiao-qiang<sup>1</sup>

(1. School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China;

2. Department of Computer Science, Xiamen University, Xiamen 361005, China)

**Abstract:** Human face localization and detection is often the first step in applications such as video surveillance and human-computer interface. In this paper, a skin color model based on cluster and a novel selection method of best limit-value is presented. An algorithmic flow for human face skin color detection is proposed. Examples are given to show the improved detection speed of the proposed algorithm and its robustness.

**Key words:** face detection; skin model; cluster

人脸检测作为人脸信息处理的第一步工作, 对后期的人脸精确定位与人脸识别影响较大. 颜色是人脸表面最为显著的特征之一, 利用颜色检测人脸是一个自然的想法. 根据肤色信息定位人脸区域有速度快且对姿态不敏感等优点. 对背景简单的图像, 仅利用肤色分割就可以定位人脸, 而对于复杂背景的彩色图像, 利用肤色分割的方法可以有效地减小进一步检测的搜索空间, 提高检测速度. 但是, 人脸表面的颜色受人类个体差异、光源的颜色以及光照的角度不同所造成的高光和阴影等诸多因素的影响, 为了不漏检人脸, 需要鲁棒性较强的肤色检测方案. 肤色检测的核心工作在于色彩空间的选择、肤色

模型的建立和阈值选择. 目前的肤色模型, 出于计算复杂度的考虑, 基本采用二维肤色模型<sup>[1-7]</sup>, 尽量将色彩分量与亮度分离, 以避免光照的影响, 而忽略了亮度因素造成的肤色分布的不均匀情况, 应用中存在一定程度的漏检. Phung 利用  $k$ -means 聚类建立肤色模型<sup>[8]</sup>, 利用迭代将采样分为高、中、低三类亮度, 提高了算法的鲁棒性, 但该算法存在两个问题: 一是忽略了色度分布, 造成了误检; 二是三维肤色模型计算和存储的复杂度太大, 限制了其在实际中的应用.

我们针对上述两个问题给出解决方案, 提出了一种改进的基于亮度聚类的肤色模型, 并提出了一种新的最佳阈值选择算法, 设计了肤色检测算法流程.

收稿日期: 2007-03-28    基金项目: 国家自然科学基金资助项目(60605012)

通信作者: 丁友东(1967~), 男, 教授, 博士生导师, 博士, 研究方向为图形学、人脸识别、数字水印算法研究. E-mail: ydding@shu.edu.cn

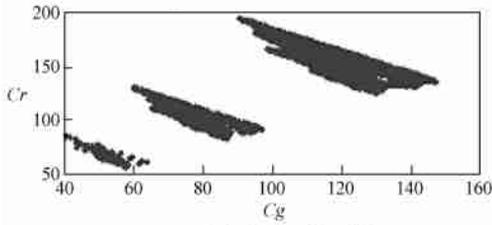
实验证明该算法提高了人脸检测的速度和鲁棒性.

### 1 色彩空间的选择

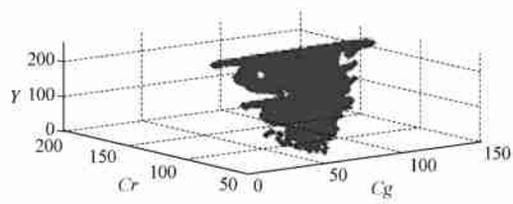
彩色图像处理有许多彩色空间坐标系,最常见的是 RGB 空间,其来源于 CRT 显像设备,是面向硬件设备的.常用的还有 YCbCr 空间、HSV 空间及 TSL 空间等,这些都是从 RGB 空间转换而来.在彩色空间不同色调的各类物体聚类在各自不同的色带中,由此可以利用彩色信息在图像中快速检测出相应的目标.由于 YCbCr 空间中的 Cb 分量是蓝色分量 B 与亮度 Y 的差,而在肤色中,B 分量所占比重较小.因此,De Dios 提出了 YCgCr 空间<sup>[4]</sup>,其中 Cg 采用了绿

色分量 G 与亮度 Y 的差,同时它也是由 RGB 的线性变换得来,但更加有效地表达了肤色的特征.所以本研究采用 YCgCr 空间. YCgCr 色彩空间与 RGB 之间是线性变化的,图 1 显示了 YCgCr 空间中肤色的分布.

$$\begin{pmatrix} Y \\ Cg \\ Cr \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \begin{pmatrix} 65.481 & 128.533 & 24.966 \\ -81.085 & 112 & -30.915 \\ 112 & -93.768 & -18.214 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$



(a) 肤色在 CgCr 空间分布



(b) 肤色在 YCgCr 空间分布

图 1 肤色在 YCgCr 色彩空间的分布

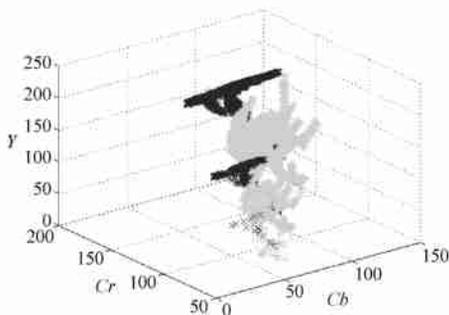
Fig.1 Distribution of skin color in YCgCr space

### 2 改进的基于聚类的肤色模型

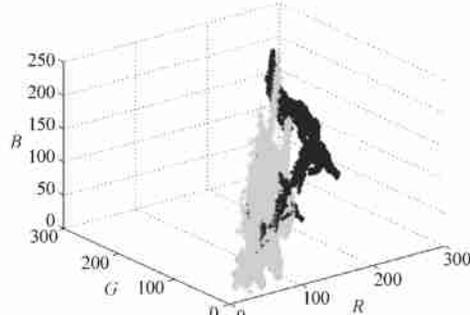
所谓的肤色模型是指一个二类分类器,即如何分类肤色与非肤色.大部分肤色模型是基于统计思想的.现在常用的肤色模型有参数化的肤色模型与非参数化肤色模型,其中前者主要有混合高斯模型,后者主要有统计直方图模型<sup>[2-3]</sup>.

彩色人脸图像易受光照的影响,图像中的人脸会产生高光区、阴影区.这些区域在肤色检测时,往往会产生漏检,干扰了进一步的检测与识别.由图 2

可以看到:(1)肤色样本在常用的色彩空间的分布不均匀,在高亮度处色彩分量分布较宽,密度较大,而在低亮度处分布较窄,密度较小;(2)黑色人种与白色、黄色人种的肤色分布范围有一定的差异;(3)人类对色彩的感知是一个涉及到认知学、心理学的问题,不同的亮度下,人类对色彩的理解也不相同.显然,如果将整个肤色空间作为一个整体来对待,使用线性转换公式或单一的分类器,是无法准确表达肤色的特征的.



(a) YCbCr 空间



(b) RGB 空间

图 2 肤色在 YCbCr 空间和 RGB 空间的分布,其中黑色部分为黑色人种肤色分布,灰色部分为白色人种与黄色人种肤色分布  
Fig.2 Distributions of skin color in YCbCr space and RGB space, where black part for black race and gray part for paleface and yellow race

*YCbCr*、*HSV* 色彩空间的非线性转换的目的是为了让肤色在颜色空间更好地聚在一起,尽量与非肤色分开。但是,在研究中发现,肤色的分布是与亮度有关的。人类的视觉在中等亮度时,对色彩变化敏感,而在高亮度时,看到的往往是一片刺眼的白色;在低亮度时,可分辨的色彩数量也很少,近于灰黑色。无论在 *YCbCr* 还是在 *YCrCb* 空间,色彩分量的分布可以根据 *Y* 分量明显地分类,如果按照亮度划分肤色分布空间,在肤色检测时按亮度分别检测各像素点,则可检测到不同亮度的肤色像素,从而提高肤色检测的鲁棒性,这同时为分析图像的光照环境提供了相关信息。

基于以上分析,Phung 建立了基于 *YCbCr* 色彩空间的肤色模型<sup>[9]</sup>,利用 *k*-means 聚类迭代将采样分成高、中、低 3 类亮度,并建立高斯混合肤色模型用于彩色图像人脸检测。我们将这种肤色模型进行改进,采用更好的表达肤色的 *YCrCb* 色彩空间。同时,为了降低计算复杂度,把亮度 *Y* [0,255]量化为 *Y* [1,16],并在建立肤色模型后,生成一个 *YCrCb*空间到肤色概率的查找表,从而可以大大提高检测的速度。

改进模型的具体步骤如下:

设 *E* 为肤色像素集:  $E = \{X_1, X_2, \dots, X_n\}$ , 其中  $X_i = (Y, Cg, Cr)$ 。

(1) 按 *Y* 大致可划分为 3 类,  $E = \{E_1, E_2, E_3\}$ , 其中,  $E_1 = \{X_j | 1 \leq Y_j < 7\}$ ,  $E_2 = \{X_j | 7 \leq Y_j < 12\}$ ,  $E_3 = \{X_j | 12 \leq Y_j \leq 16\}$ , 即将亮度分成高、中、低 3 类亮度范围;

(2) 分别计算每类中 *Y*、*Cg* 与 *Cr* 的均值与协方差:

$$\mu(k) = \frac{1}{n} \sum_{i=1}^n X_i, \quad (2)$$

$$\sigma(k) = \frac{1}{n} \sum_{i=1}^n (X_i - \mu(k))(X_i - \mu(k))^T; \quad (3)$$

(3) 对 *E* 中每个采样点,逐一计算其在 3 个高斯混合分布中的马氏距离:

$$p(X_i, k) = \frac{1}{2 \sqrt{|(k)|}} \times \exp\left\{-\frac{1}{2} (X_i - \mu(k))^T \cdot (k)^{-1} (X_i - \mu(k))\right\}, \quad (4)$$

将此点分配到概率值最大的类中,

$$X_j \rightarrow E_k, k = \arg \max_k p(X_j, k); \quad (5)$$

(4) 重复(2)、(3),直到达到一定的迭代次数或者某类中像素的个数小于给定值;

(5) 保存每类中 *Y*、*Cg* 与 *Cr* 的均值  $\mu(k)$  与 *YCrCb*协方差  $\sigma(k)$ 。

在实验中,我们选择 282 082 个不同光照、不同人种的肤色采样像素,转换到 *YCrCb* 空间,求出聚类质心与查找表。

为了减少存储空间,同时降低查找时的复杂度,我们利用得到的各聚类的均值与协方差生成查找表,伪代码如下:

```
for Y = 0 to 15
  for Cg = 40 to 160
    for Cr = 50 to 220
      { i 依次取 1,2,3;
        根据保存的均值与协方差,按照(4)计算马氏距离 Mi;
        T(Y, Cg, Cr) = min(Mi(Y, Cg, Cr));
      }
```

最后,生成一个大小为 16 × 120 × 170 的三维表,以存储空间换取计算时间。这样,在后面的检测中,可以直接查找表中相对应的值,不需要每次计算,极大地提高了检测速度。

### 3 最佳阈值选取

阈值分割的目的是将肤色与背景分开,在进行肤色检测时,对图像中的每个像素点,逐一按照式(4)计算像素点到聚类中心的马氏距离,取最小值作为该点到肤色像素类的距离,设定一个阈值,若该距离大于一定值则为非肤色,否则为肤色像素。因此,阈值的设定对肤色检测的精度有很大影响,值太小会漏检属于肤色的像素点,而值太大又会把非肤色误检为肤色。一般有如下 3 种设定阈值的方法:(1)固定阈值法,对所有的图像设定一个阈值;(2)自适应阈值法,根据各图像的亮度等信息,自适应地调整阈值大小;(3)最佳阈值法,针对每幅图像,选择一个最佳分割阈值。

采用固定阈值的方法,效果很不理想,而自适应阈值法在应用中计算速度较慢,不适合于快速检测。对此,我们提出一种最佳阈值的分割方法,这种方法基于以下事实:(1)图像中主要人脸的图像较之图中其他似肤色连通区域,其面积大致相同;(2)在有大片连通的似肤色的背景中,当阈值由大到小递减到

某一阈值时,肤色的像素数量会突然增大;(3)在有大量分散的似肤色的背景中,当阈值由大到小递减到某一阈值时,连通区域的数量会突然增大。

该算法的基本思想是:如果把肤色像素的数目与生成的连通区域的数目的“商”看作图像中人脸的平均面积。那么,在阈值变化中,这个“商”变化最小的时候,也就是人脸的平均面积变化最小时,可以认为此时的阈值为最佳阈值。实验中,我们以0.1为间隔,由0.75至0.05,计算每次肤色像素的数目与生成的连通区域的数目的“商”(相当于人脸的平均面积),求得其中变化最小的一个,作为最佳阈值进行分割,实验证明效果很好。

#### 4 肤色检测流程

图像分割是将一幅图像划分为若干个相互没有重叠区域的图片的过程。每个区域具有一致的“有意义”属性的像素集合。肤色检测的本质就是图像分割,而其最终目的是将图像中的肤色像素点与背景分离。

根据前面提出的肤色模型和最佳阈值选择算法,我们设计出如下的肤色检测流程:

(1) 给定一幅彩色图像,将其转换到  $Y CgCr$  空间;

(2) 对其中的每个点,根据  $Y CgCr$  的值查表,通过计算式(4)得到所有点的马氏距离值  $p(i, j)$ , 其中  $(i, j)$  为图中点的坐标;

(3) 对  $p(i, j)$  进行预处理。因为人脸肤色是连续的,所以生成的距离也应该是连续的,对得到的肤

色距离进行  $3 \times 3$  均值滤波,以去除图像中的噪声;

(4) 对  $p(i, j)$  采取极值规范化,以减少光照与成像设备的影响,

$$p(i, j) = \frac{p(i, j) - \min(p(i, j))}{\max(p(i, j)) - \min(p(i, j))}, \quad (6)$$

式中,  $p(i, j)$  为原点  $(i, j)$  的值,  $p(i, j)$  为规范化后的点  $(i, j)$  的值;

(5) 采用最佳阈值法进行阈值分割,选择阈值并计算

$$\begin{aligned} p(i, j) &= 1, p(i, j) > \tau, \\ p(i, j) &= 0, p(i, j) \leq \tau; \end{aligned}$$

(6) 对得到的二值图像  $p(i, j)$  进行二值形态学的闭处理,其结构元素为  $3 \times 3$  的单位矩阵;然后对其连通区域进行人脸区域选取。去除噪声区域,即面积小于图像总面积一定比例阈值的区域。

#### 5 实验结果与总结

针对本算法,我们在 WinXP 操作系统, P4 2.0G 双 CPU 处理器,内存 504 M 的环境下进行了实验。由实验结果可知,改进的基于亮度聚类的肤色模型在应用中有较好的鲁棒性,能适应于不同光照下的人脸检测,与 Phung 的模型相比降低了在中低亮度的误检错误,缩小了定位人脸的范围,图3是部分实验比较结果。我们在实验中还发现,在复杂背景的彩色图像中,利用亮度聚类的肤色模型,漏检现象相对较少。由于采用查找表与量化的方法,检测的速度有了极大提高。基于面积与连通区域的最佳阈值选择,



图3 原图(左)、Phung 肤色模型检测结果(中)和本算法肤色模型检测结果(右)

Fig. 3 Original image(left), detection result by Phung model(middle) and the improved model(right)

当图像中存在多人脸时,图像分割的效果很好.

我们重点研究了色彩空间与肤色模型.选择了兼顾转换速度与检测精度的色彩空间  $YCgCr$  作为肤色模型建立的色彩空间,并提出了改进的基于聚类的肤色模型,建立了量化亮度的查找表,可以快速区别图像中的肤色区域与非肤色区域.并且提出了基于平均面积的最佳阈值选择方法,实验证明,我们提出的基于亮度聚类的肤色模型有很强的鲁棒性和较高的检测速度.

#### 参考文献:

- [1] TERRILLON J, SHIRAZI M, FUKAMACHI H, et al. Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images [C]// Proceedings of 4th IEEE International Conference on Automatic Face and Gesture Recognition. 2000:54-61.
- [2] VEZHNEVETS V, SAZONOV V, ANDREEVA A. A survey on pixel-based skin color detection techniques [C]// Proceedings on Graphicon. 2003:85-92.
- [3] HSU R L, ABDEL-MOTTALIB M, JAIN A K. Face detection in color images[J]. Pattern Analysis and Machine Intelligence, 2002, 24(5):696-706.
- [4] DE DIOS J, GARCIA N. Face detection based on a new color space  $YCgCr$  [C]// Proceedings of International Conference on Image Processing. 2003, 2:909-912.
- [5] LIU H, YAN S Y, CHEN X L, et al. Rotated face detection in color images using radial template (RT) [C]// Multimedia and Expo, ICME 03. 2003, 3:137-140.
- [6] 张洪明,赵德斌,高文.基于肤色模型、神经网络和人脸结构模型的平面旋转人脸检测[J].计算机学报,2002, 25(11):1250-1255.
- [7] 刘洁,张汗灵.一种新的基于肤色模型的人脸检测算法[J].计算机工程与应用,2006,42(11):70-72;134.
- [8] PHUNG S, BOUZERDOUM A, CHAI D. A novel skin color model in  $YCbCr$  color space and its application to human face detection [C]// Proceedings of International Conference on Image Processing. 2002, 1:289-292.
- [9] PHUNG S, CHAI D, BOUZERDOUM A. Adaptive skin segmentation in color images [C]// Proceedings ICASSP. 2003, 3:173-176.

(编辑:赵宇)

·上海大学计算机工程与科学学院研究机构·

### 智能信息处理研究室

研究室研究方向涉及智能信息处理研究领域的数据挖掘、人工智能应用技术、智能系统等课题.现有 10 名研究人员,其中教授 5 名,副教授 5 名.承担了国家和上海市级科研项目、国际合作项目 10 多项.主要项目是:

- (1) 分布式概念格数学模型角度算法研究(国家自然科学基金项目);
- (2) 智能系统中人机交互理论及其应用研究(国家自然科学基金项目);
- (3) 数据挖掘在地震预报中的应用研究(国家地震联合基金);
- (4) 基于语义特定内容检查与过滤技术及其应用研究(上海市科委重点攻关研究项目);
- (5) 面向本体的形式概念分析扩展模型和算法(国家自然科学基金项目);
- (6) 药物构效关系的集成学习方法(国家自然科学基金项目);
- (7) 地震预报中并行数据挖掘方法研究(上海市科委自然基金).

提出了概念格快速生成和规则生成等算法以及基于容差关系的概念格扩展等模型,是我国开展形式概念分析研究最早的单位,研究成果被国内同行专家较普遍地引用;提出并在超级计算机上实现了多种数据挖掘并行算法;提出了一套反编译原理与技术,开发了我国第一个 C 语言反编译系统.语义网格融合了网格、语义 Web 以及 Web services 的优势,扩充了网格的语义能力,提升了语义 Web 的计算能力.发表学术论文 300 多篇,被 EI 和 SCI 收录 60 多篇,出版著作、教材和译著 8 部,获得省部级科技进步二等奖 4 项,三等奖 2 项.主持召开了第二届国际计算机和信息技术会议(CIT2001).