



## 基于高斯肤色模型和图像分割的人脸检测

柳映辉, 任德华, 陈少杰

(天津科技大学计算机科学与信息工程学院, 天津 300222)

**摘要:** 提出了一种基于 HSV、YCgCr 颜色空间和图像区域分割的人脸检测方法. 首先,将规范化的 rgb 肤色转化到 HSV、YCgCr 颜色空间,对向量  $(H, Cg, Cr)^T$  进行统计,得到肤色的三维高斯概率模型. 然后,根据该模型计算待检图像的肤色相似概率,采用最大类间方差的方法计算二值化阈值,根据阈值将相似度图像二值化. 为了抵消人脸区域亮度变化和阴影的影响,以二值图像为模板,在原待检图像的饱和图中根据 3 像素×3 像素区域的平均 R、G 值进行图像区域分割. 在分割得到的区域中,根据人脸特征检测出图像中的人脸. 实验表明,提出的方法能在复杂背景的情况下检测出人脸图像.

**关键词:** 人脸检测; 高斯模型; 图像分割

中图分类号: TP391.41

文献标志码: A

文章编号: 1672-6510(2010)05-0064-04

## Face Detection Based on Skin Color Gaussian Model and Image Segmentation

LIU Ying-hui, REN De-hua, CHEN Shao-jie

(College of Computer Science and Information Engineering, Tianjin University of Science & Technology, Tianjin 300222, China)

**Abstract:** Based on HSV and YCgCr color space as well as image segmentation, an approach for face detection was put forward. First, the normalized rgb color was converted into HSV and YCgCr color space. Meanwhile the 3D Gaussian model was obtained by statistics of the vector  $(H, Cg, Cr)^T$ . According to this model, the image similarity of skin color that is ready for being inspected could be calculated, and binary threshold could be attained through Otsu method. As a result, the threshold could binary the similarity image. In the application of binary image, the saturate image was segmented on the basis of the R, G component obtained by the 3 pixel×3 pixel region so as to offset the effect of brightness variation and shadow. In the areas of image segmentation, face was detected from the image by facial features. Experiment shows that the method can detect the image of face accurately in complicated background.

**Keywords:** face detection; Gaussian model; image segmentation

人脸检测 (Face Detection) 的目的是判断图像中是否存在人脸, 并要给出人脸的位置、大小、是否旋转及旋转角度等信息. 随着大量研究工作的展开, 国内外研究者提出了各种各样的人脸检测方法. 大部分的人脸检测算法复杂度都较高, 为了提高运算效率应尽可能地减少检测的范围.

基于颜色特征的方法作为人脸检测的前期步骤可以快速地缩小检测的区域, 是一种有效的预处理方

法<sup>[1]</sup>. 因此, 基于肤色信息的人脸检测方法无论是作为检测前期的预处理, 还是作为检测后期的结果确认, 都是一种有效的手段.

Michael 和 James 统计了大量肤色、非肤色图像, 根据 RGB 空间下的直方图得到一个基于概率比值的肤色分类器<sup>[2]</sup>. 但是在 RGB 颜色空间中, 其颜色与 RGB 三个分量都有关系, 无法做到颜色、亮度和饱和度信息的分离. 刘万春等<sup>[3]</sup>利用规范化 rgb 颜色空间

收稿日期: 2010-01-26; 修回日期: 2010-05-27

基金项目: 天津科技大学引进人才科研启动基金资助项目 (20060416)

作者简介: 柳映辉 (1979—), 男, 山东栖霞人, 讲师, liuyinghui@tust.edu.cn.

中的分量  $r, g$  建立二维高斯模型来检测人脸区域,并建立了一个自适应模型调整肤色模型参数,然后采用扩展卡尔曼滤波器预测运动轨迹,在预测轨迹附近的一个区域进行肤色匹配以减少检测范围,最后采用特征匹配的方法检测人脸. 规范化的  $rgb$  颜色空间虽然能一定程度上减少亮度差异的影响,但是肤色在  $rgb$  空间中的分布离散性还是较大,当待检测图像之间的差异较大时采用该方法不容易取得好的效果. 温静等<sup>[4]</sup>则把  $RGB$  图像转换到  $YCbCr$  颜色空间,并对  $Cb, Cr$  分量建立高斯混合模型(Gaussian Mixture Model, GMM)来初步检测人脸,并配合基于人脸、非人脸特征统计模型的 Bayes 分类器进行人脸检测.  $YCbCr$  空间虽然做到了亮度信息的分离,但是高斯混合模型参数的确定却过于麻烦,并且非人脸的特征统计模型也是不容易获得的.  $YCgCr$  颜色空间中肤色类聚性优于  $YCbCr$  颜色空间,张争珍等<sup>[5]</sup>用  $YCgCr$  颜色空间中的  $Cg, Cr$  分量建立二维高斯模型进行人脸检测. 该方法仅利用了  $Cg, Cr$  分量,背景干扰较大.

本文在文献[5]的基础上,提出了采用  $(H, Cg, Cr)^T$  向量建立高斯模型,先进行初步人脸区域检测,然后进行图像区域分割进一步确定出人脸,最后再利用特征判断的方法得到人脸位置. 由于加入了  $H$  分量,其与亮度信息无关,只与颜色相关,故可以在一定程度上减少背景干扰和亮度差异的影响,并且利用人脸区域连续性的特点,联合采用图像区域分割的方法,可以在复杂背景中检测出人脸位置.

## 1 检测原理

### 1.1 颜色空间转换

在将  $RGB$  颜色空间的数值转化到  $HSV, YCgCr$  颜色空间之前,为了进一步减少皮肤亮度的差异对检测的影响,先将  $RGB$  颜色进行规范化,得到规范化的  $rgb$  值,转换方法与标准方法<sup>[6]</sup>略有不同:

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \frac{255}{R+G+B} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

在此基础上将规范化的  $rgb$  颜色转化到  $HSV$  和  $YCgCr$  空间中,获取  $(H, Cg, Cr)^T$  三个分量. 从  $rgb$  空间转换到  $HSV$  空间的转换公式为

$$\begin{aligned} k_{\max} &= \max(r, g, b) \\ k_{\min} &= \min(r, g, b) \\ \Delta &= k_{\max} - k_{\min} \\ V &= k_{\max} \end{aligned}$$

$$S = \begin{cases} 0 & \Delta = 0 \\ 100\Delta / k_{\max} & \Delta \neq 0 \end{cases}$$

$$H = \begin{cases} 0 & S = 0 \\ (g-b)/\Delta & S \neq 0, r = k_{\max} \\ 2+(b-r)/\Delta & S \neq 0, g = k_{\max} \\ 4+(r-g)/\Delta & S \neq 0, b = k_{\max} \end{cases}$$

$$H = \begin{cases} 60 \times H & H \geq 0 \\ 60 \times H + 360 & H < 0 \end{cases}$$

$$H \in [0, 360), S \in [0, 100], V \in [0, 255]$$

从  $rgb$  空间转换到  $YCgCr$  空间的转换公式为

$$\begin{pmatrix} Y \\ Cg \\ Cr \end{pmatrix} = \begin{pmatrix} 0.2568 & 0.5041 & 0.0979 \\ -0.3180 & 0.4392 & -0.1212 \\ 0.4392 & -0.3678 & -0.0714 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix} + \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix}$$

经过上述步骤以后,可以将任意一个像素经过  $RGB \rightarrow rgb \rightarrow HCbCr$  转换后得到混合颜色空间的肤色向量. 将肤色向量中各个分量看作是随机变量,即可以根据肤色样本图像计算得到高斯模型.

### 1.2 高斯肤色模型

肤色向量为三维向量,故最终统计得到的高斯模型是多随机变量的高斯模型. 下面根据单随机变量的协方差给出多随机变量协方差矩阵的推导过程:

设  $X_1, X_2, \dots, X_n$  为一组随机变量,这些随机变量构成随机向量  $\mathbf{X} = [X_1, X_2, \dots, X_n]^T$ , 每个随机变量有  $m$  个样本,则有样本矩阵

$$\mathbf{M} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = (\beta_1 \ \beta_2 \ \dots \ \beta_n)$$

其中  $\alpha_i (i = 1, 2, \dots, n)$  行向量为随机变量  $X_i$  所有样本值,  $\beta_i (i = 1, 2, \dots, n)$  列向量为随机向量  $\mathbf{X}$  的单个样本值.

随机变量  $X_i, X_j$  之间的协方差可以表示为<sup>[7]</sup>

$$\begin{aligned} C_{ij} &= E[(X_i - E(X_i))(X_j - E(X_j))] = \\ &= E(X_i X_j) - E(X_i)E(X_j) = \\ &= \frac{1}{m} \alpha_i \alpha_j^T - \frac{1}{m^2} \left( \sum_{k=1}^m M_{ik} \right) \left( \sum_{k=1}^m M_{jk} \right) \end{aligned}$$

故协方差矩阵可写为

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{pmatrix} = \frac{1}{m} \mathbf{M} \mathbf{M}^T - \frac{1}{m^2} \begin{pmatrix} \sum_{k=1}^m M_{1k} \\ \sum_{k=1}^m M_{2k} \\ \vdots \\ \sum_{k=1}^m M_{nk} \end{pmatrix} \times \begin{pmatrix} \sum_{k=1}^m M_{1k} & \sum_{k=1}^m M_{2k} & \dots & \sum_{k=1}^m M_{nk} \end{pmatrix} =$$

$$\frac{1}{m} \sum_{i=1}^m [\beta_i - \bar{\beta}][\beta_i - \bar{\beta}]^T$$

其中,  $\bar{\beta} = \frac{1}{m} \sum_{i=1}^m \beta_i$

令  $X = (H \ Cg \ Cr)^T$ , 则协方差矩阵  $C = \frac{1}{m} \sum_{i=1}^m$

$[X_i - \bar{X}][X_i - \bar{X}]^T$ , 其中  $\bar{X}$  是  $X$  所有样本的均值.

得到样本均值和协方差矩阵后, 令

$$p(x) = 255 \times \exp[-\frac{1}{2}(x - \bar{X})^T C^{-1}(x - \bar{X})]$$

则  $p(x)$  可以作为肤色相似度的衡量参数. 进行人脸检测时, 把待检测图像中的每一个像素都进行计算, 最后得到一副 256 灰度级别的肤色相似度图像. 图像中每个像素的亮度值表示该像素与肤色的相似程度大小: 亮度值越高, 则表示原图像中该像素越接近皮肤的颜色.

根据肤色相似度图像, 采用最大类间方差的方法 (Otsu 法) 进行图像二值化, 在二值图像中白色的像素表示肤色点, 黑色的像素表示非肤色点. 从而, 使用高斯肤色模型完成了肤色检测的工作. 一般情况下, 人脸的区域会被检测为肤色区域, 只需在二值图像中像素为白色的区域中进一步检测就可以获得人脸. 当然, 由于待检测图像的多样性, 例如光照不均匀、拍摄条件不一致、背景干扰等原因, 采用高斯肤色模型得到的结果只能作为初选条件, 结合其他检测方法才能得到理想的结果.

协方差矩阵和样本均值决定着人脸检测的最终效果, 而肤色样本的选择直接影响着上述协方差矩阵的值和样本均值, 只有当样本图像的  $(H, Cg, Cr)^T$  向量真正符合高斯分布时, 该模型才能获得好的检测结果. 因此在选取肤色样本时, 应将样本图像中明显差异度较大的图删除.

### 1.3 图像区域分割

二值化肤色相似度图后, 人脸区域可能不连续 (人脸被分裂为多个部分) 或者不完整 (人脸只保留一小部分), 单纯使用二值图像不容易得到完整、准确的人脸位置. 由于原始图像中的人脸区域一般是连续的, 因此利用区域连续性可以从原始图像中得到较为完整的人脸区域.

采用基于区域增长的方法来进行图像分割, 如何判断区域连续性是一个关键问题. 通过实验发现, 利用原始 RGB 颜色作为判断依据时, 照片拍摄条件的不一致导致很难找到一个通用的阈值来适应绝大多数的图像, 人脸区域分割的效果不理想. 而将肤色转换到 HSV 颜色空间时,  $H$  分量为红色偏绿色, 可以以

此作为区域连续性判断的依据, 并且将饱和度  $S$  规范化为 100%, 有利于减少饱和度对连续性判断的影响. 饱和图像中人脸颜色不包含蓝色  $B$  分量, 通过  $R$ 、 $G$  分量来判断某个点是在区域内还是在区域外. 采用计算图 1 所示 3 像素×3 像素区域的  $R$ 、 $G$  均值来判断其中心点是否在区域内, 判断方法为

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \frac{1}{9} \begin{pmatrix} R_p \\ G_p \\ B_p \end{pmatrix} + \frac{1}{9} \sum_{i=1}^8 \begin{pmatrix} R_{p_i} \\ G_{p_i} \\ B_{p_i} \end{pmatrix}$$

当  $(R, G, B)^T$  满足  $B < 0.01, R > 100, G > 30, R > 1.5G$  时,  $P$  点是区域内的点; 否则  $P$  点是区域边缘的点. 实验发现通过如此处理的饱和图 RGB 颜色, 可以得到满意的分割效果.

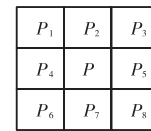


图 1 检测区域举例

Fig.1 Example of detection area

区域增长算法如下:

- (1) 将肤色二值化图像滤除干扰, 得到相对干净的二值化图  $I_B$ .
- (2) 将原图中每个像素从 RGB 空间转换到 HSV 空间, 并令其  $H = 100$ , 再将 HSV 值反算转换到 RGB 空间, 得到饱和图像  $I_S$ .
- (3) 以  $I_B$  中白色点为起始点, 在  $I_S$  中按照上述的判断方法进行区域搜索, 得到区域 Region.
- (4) 如果 Region 区域大于 400 像素, 且其宽高比例介于  $[0.33, 3]$  时, 则将 Region 视为人脸区域, 加入候选区域列表; 否则转到 (3) 继续搜索下一个区域.
- (5) 当  $I_B$  中所有白色点都搜索完毕时, 区域增长算法结束, 输出候选区域列表.

这样在二值图像中不完整的人脸区域就被完整的查找出来了, 当然以上粗略的判断条件无法完全滤除非人脸区域, 因此还要将候选区域中的非人脸滤除, 才能得到最终结果.

### 1.4 非人脸区域滤除

从候选人脸区域中将不符合人脸特征的区域滤除, 主要是根据人眼、人嘴的颜色特征和位置特征. 候选区域先把  $H \in (0, 30) \cup (330, 360)$  的部分保留, 其他部分涂黑, 再计算  $Cr^2/Cg^2$ , 其中  $Cr^2/Cg^2$  要规范化到区间  $[0, 255]$ . 最后将  $Cr^2/Cg^2$  二值化并腐蚀两次, 得到的二值图中白色区域的包围矩形宽高比位于区间  $(1, 10)$  时认为是人嘴的位置. 根据嘴和眼睛的

位置比例,计算区域中属于人眼部分  $E_1$  和  $E_2$  的亮度值,如果二者亮度相近,且比周围区域亮度低,则认为  $E_1$  和  $E_2$  就是人眼的位置. 否则认为候选区域中不存在人脸,将其滤除.

## 2 实验

实验图像集来自于网上图片和自拍照片,共计300副. 样本图像是从图像集中随机选取50副,从50副图中手工切割出200副小的人脸样本图像(删除差异明显的图像). 然后用这200副人脸样本图像建立三维高斯肤色模型,得到的结果为

$$\bar{X} = (23.6982 \quad 121.0278 \quad 141.3382)^T$$

$$C^{-1} = \begin{pmatrix} 0.0009 & 0.0022 & 0.0011 \\ 0.0022 & 0.8209 & 0.4004 \\ 0.0011 & 0.4004 & 0.2117 \end{pmatrix}$$

以单人和多人两种情况为例,图2为原始图像,图3为使用本文方法得到的肤色相似度图. 可以看出人脸区域亮度较高,具有较高的肤色相似度. 区域分割图见图4,图中的人脸在候选区域中被完整检出,但是同时也有一些非人脸区域被检出. 图5是滤除非人脸区域后的最终检测结果.



(a) 单人 (b) 多人

图2 原始图像  
Fig.2 Original image



(a) 单人 (b) 多人

图3 肤色相似度图

Fig.3 Skin color similarity image



(a) 单人 (b) 多人

图4 区域分割图

Fig.4 Result of image segmentation



(a) 单人 (b) 多人

图5 检测结果

Fig.5 Result of detection

本文方法与文献[5]检测方法的结果对比见表1. 可以看出本文方法在单纯的使用高斯肤色模型基础上,增加了图像区域分割的处理,因此检测正确率有较大提高.

表1 实验结果

Tab.1 Result of face detection %

检测方法	正确率	漏检率	误检率
文献[5]	80.3	19.7	3.0
本文	90.6	9.4	3.3

## 3 结语

提出了一种基于 HSV、YCgCr 颜色空间和图像区域分割的人脸检测方法,先进行初步人脸区域检测,然后进行图像区域分割进一步确定出人脸,最后再利用特征判断的方法得到人脸位置. 由于加入了  $H$  分量,可以在一定程度上减少背景干扰和亮度差异的影响. 利用人脸区域连续性的特点,采用图像区域分割的方法,可以在复杂背景中检测出人脸位置. 但是在实验的过程中,也发现了一些问题:

(1)采用 Otsu 算法进行图像二值化的效果不理想,可能的原因是肤色相似度图中灰度分布在低亮度区域过于集中,前景、背景之间无明显分类差异,使本应归于背景的一部分区域也变为了前景.

(2)基于肤色模型的检测方法针对某类图像具有较好的效果,当图像成像条件过于多变时,正确检测率会急剧降低.

以上问题尚需进一步的深入研究,以取得更好的检测结果.

### 参考文献:

[1] 黄福珍,苏剑波. 人脸检测[M]. 上海:上海交通大学出版社,2006:24-27.  
 [2] Jones M J,Rehg J M. Statistical color models with application to skin detection[J]. International Journal of Computer Vision,2002,46(1):81-96.

(下转第78页)