

基于 YCbCr 空间的亮度自适应肤色检测

Self-adaptive Skin Color Detection Based on YCbCr Color Space

王金庭 (湖北经济学院 计算机与电子科学系 湖北武汉 430205)

杨 敏 (中国地质大学 信息工程学院 湖北武汉 430074)

摘要:本文针对复杂背景彩色图像提出了一种基于 YCbCr 颜色空间的自适应肤色检测算法。该算法从两个方面考虑了抗光照问题:采用非线性分段色彩变换和基于 Fisher 评价准则的自适应阈值分割方法减小亮度的影响。实验结果表明,该算法复杂度较小,对光照变化具有很好的鲁棒性。

关键词:肤色检测 颜色空间 图像分割 Fisher 准则函数

1 引言

肤色检测主要是根据肤色在颜色空间上的分布特征来检测图像中的肌肤区域。人的肤色是人体表面最为显著的特征之一。肤色是人脸部及手部区别于其他区域的重要特征,同时因其对姿势、旋转、表情等变化不敏感,可有效去除图像中复杂背景的干扰、减小待搜索范围^[1]。因此,利用颜色信息进行皮肤区域探测已引起较多学者关注。利用颜色信息进行皮肤检测在人脸检测与识别、高级人机交互界面、手势识别、基于内容的图像检索和基于内容的不良图像过滤等领域有着重要的地位。

Rossotti 与 Angelopoulou^[2]分别在 1983 年与 1999 年的研究成果证明了在生物和物理上肤色分布的一致性,指出尽管人的肤色因人种的不同而不同,呈现出不同的颜色,但是在排除了亮度、视环境等对肤色的影响后,皮肤的色调基本一致。这也为利用颜色信息进行皮肤检测的可能性提供了有利的证据。

Rossotti 与 Angelopoulou 的研究表明不同人种的肤色在色调上是一致的,并且皮肤的饱和度特征有一定的稳定性,因而采用具有认知属性的、能够将亮度信息与色度信息分开表示的颜色空间对于皮肤的检测能够起到很好的作用。

对于颜色空间的选择问题,或者说在哪个颜色空间中皮肤检测能达到最佳的效果引起了许多学者的广泛关注与研究。目前有许多的颜色空间,陶霖密将颜色空间分为五类^[3],而应用于皮肤检测的颜色空间也

有许多,如归一化 RGB 颜色空间, YCbCr, YUV, YIQ, HIS/ GIHS, CIE Lab 和 CIE Luv 等颜色空间。在这里并不对颜色空间的选择作太多的讨论,因为任何颜色空间在皮肤检测中性能的比较都是在特定的皮肤分布模型下进行的,不具有一般性。而如何有效地检测彩色图像中的皮肤区域并最大限度的滤除掉非皮肤区域仍是皮肤检测算法研究的主要问题。

2 基于 YCbCr 空间的肤色模型

YCbCr 色彩模型是 CCIR601 编码方式的色度模型,被广泛应用在电视的色彩显示等领域中。其中 Y 是亮度,而 Cb 和 Cr 是色度信息。YCbCr 空间具有将色度与亮度分离的特点,在 YCbCr 色彩空间中肤色的聚类特性比较好,而且是两维独立分布,能较好地限制肤色分布区域^[4],并且受人种的影响不大。本研究选择 YCbCr 色彩空间模式进行肤色区域分割。

YcbCr 格式可以从 RGB 格式线性变化得到,转换公式如下:

$$\begin{bmatrix} Y \\ Cb \\ Cr \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 & 0 \\ -0.1687 & -0.3313 & 0.5000 & 0.5 \\ 0.5000 & -0.4187 & -0.0813 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix}$$

通过对大量皮肤像素的统计分析可以看到肤色聚类在色度空间中的一个很小的范围内。传统的肤色分割方法就是在二维的 Cb - Cr 子平面中寻求肤色的聚类区域,根据统计结果设定阈值从而将肤色区域与非

肤色区域分离。

因为 YCbCr 色彩格式直接由 RGB 色彩格式通过线性变换得到, 所以其亮度分量 Y 并不是完全独立于色度信息而存在的, 所以肤色的聚类区域也是随 Y 的不同而呈非线性变化的趋势。

Anil K. Jain 等从 Heinrich - Hertz - Institute (HHI) 图像库中的 137 幅图像中手工选取的 853571 个肤色像素点, 将其绘制在 YCbCr 空间及其二维投影子空间 Cb - Cr 空间中得到的结果如图 1 所示^[5]。从图中可以看到, 在 YCbCr 色彩空间中, 肤色聚类是呈两头尖的纺锤形状, 也就是在 Y 值较大和较小的部分, 肤色聚类区域也随之缩减。由此可见, 在 Y 值不同的地方, 我们取 Cb - Cr 子平面的投影, 得到的结果是不同的。由此得到结论, 简单的排除 Y 分类, 按照传统的做法在二维的 Cb - Cr 子平面中寻求肤色的聚类区域是不可行的, 我们必须考虑 Y 值不同造成的影响, 从而对 YCbCr 色彩格式进行非线性分段色彩变换。

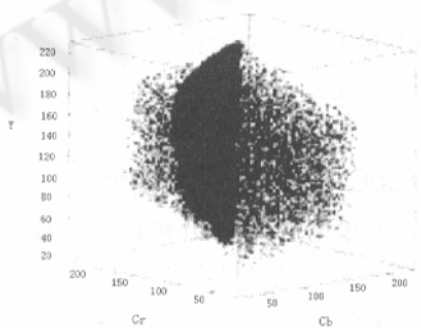


图 1 肤色在 Y - Cb - Cr 空间的分布

Anil K. Jain 等的研究表明, 在肤色聚类的边界上, Cb 和 Cr 两分量随 Y 变化的情况可以用四个边界来限制肤色聚类区域。这样可以很好的适应亮度过明或过暗的区域, 从而使肤色模型的鲁棒性大大提高。

经过了非线性分段色彩变换得到的色彩空间用 YC_bC_r 来表示。 YC_bC_r 坐标空间到 $YC'_bC'_r$ 坐标空间的转换公式为:

$$C'_i = \begin{cases} (C_i(Y) - \bar{C}_i(Y)) \cdot \frac{W_{C_i}}{W_{C_i}(Y)} + \bar{C}_i(Y) & \text{if } Y < K_l \text{ or } Y > K_h \\ C_i(Y) & \text{if } Y \in [K_l, K_h] \end{cases}$$

其中:

$$W_{C_i}(Y) = \begin{cases} W_{LC_i} + \frac{(Y - Y_{\min}) \cdot (W_{C_i} - W_{LC_i})}{K_l - Y_{\min}} & \text{if } Y < K_l \\ W_{HC_i} + \frac{(Y_{\max} - Y) \cdot (W_{C_i} - W_{HC_i})}{Y_{\max} - K_h} & \text{if } Y > K_h \end{cases}$$

$K_l = 125, K_h = 128$ (非线性分段色彩变换的分段值域), $Y_{\min} = 16, Y_{\max} = 235$ (肤色聚类区域中 Y 分量的最小和最大值), $W_{C_b} = 46.79, W_{LC_b} = 23, W_{HC_b} = 14, W_{C_r} = 38.76, W_{LC_r} = 20, W_{HC_r} = 10$ 是试验得到的常数^[5]。

肤色模型是根据大量样本的统计数据建立的。为建立肤色模型, 本人首先手工剪裁出大量肤色区域, 分别来自不同的性别在不同的身体区域 (包括脸部、手臂、四肢和躯干部分)。针对不同种族的人: 亚洲人、高加索人和非洲人, 本文选择在一般室内外照明背景下的人脸图片, 以便所建立的肤色模型具有一般性。然后将皮肤区域中每个像素点的 R, G, B 转换成 Y, Cb, Cr 表示, 得到每个像素点的色度值 (Cb, Cr)。

对剪裁出的肤色区域图像不可避免存在各种噪声, 因为噪声频谱多在高频段, 因此可以采用低通滤波器减小样本的噪声, 进行图像平滑。图像平滑包括空域法和频域法两大类, 在空域法中, 图像平滑一种常用方法是采用均值滤波, 它是用一个有奇数点的滑动窗口在图像上滑动, 将窗口中心点对应的图像像素点的灰度值用窗口内的各个点的灰度值的平均值代替, 如果滑动窗口规定了在取均值过程中窗口各个像素点所占的权重, 也就是各个像素点的系数, 这时候就称为加权均值滤波。作者采用的滤波器的脉冲响应如下。

$$R = \frac{1}{13} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

将皮肤区域中每个像素点通过非线性分段变换, 变换到 $YC'_bC'_r$ 坐标空间。统计色度为 (C'_b, C'_r) 的像素点数, 得到色度分布图。将色度分布图的数值除以其最大值, 得到归一化色度分布图, 即肤色分布可以通过高斯模型 $N(m, C)$ 表示, 其中 m 是均值, C 是协方差阵。

$$m = E\{x\}, x = (C'_b, C'_r)^T$$

$$C = E\{(x-m)(x-m)^T\}$$

通过前面所建立的肤色高斯分布,可以得到某幅彩色图像中任意一个像素点属于皮肤的概率。对于某像素点 s ,从 RGB 空间转换到 YC'_{b}, C' , 颜色空间,得到色度值 (C'_{b}, C') ,那么这个像素的肤色概率密度(即肤色似然度)可以通过下面的式子计算得到:

$$P(C_r, C_b) = \exp\left[-\frac{1}{2}(x-m)^T C^{-1}(x-m)\right]$$

计算被检测的彩色图像各像素点的肤色似然度,并得到整幅图像的最大肤色似然度。每一点像素的肤色似然度除以最大肤色似然度所得到的值,作为该像素点的灰度值(表征这个像素点属于皮肤的概率),从而得到肤色似然图像。如图 2 所示。

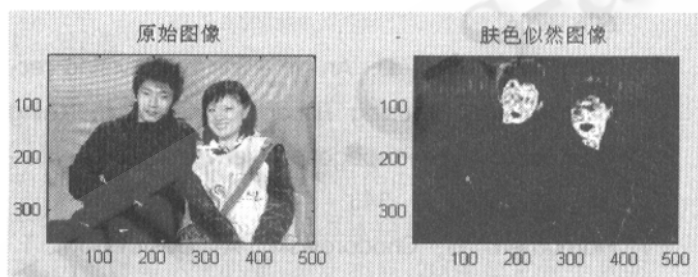


图 2 肤色似然图像

3 自适应阈值分割

图像阈值分割是一种广泛使用的图像分割技术,它利用了图像中要提取的目标和背景在灰度特征上的差异,把图像视为具有不同灰度级的两类区域(目标和背景)的组合,选择一个合适的阈值,以确定图像中每一像素点应属于目标还是背景区域,从而产生相应的二值图像。要从复杂背景中分辨出目标并将其形状完整地提取出来,阈值的选择是关键。如果阈值选取过高,则过多的目标点被误归为背景;阈值选得过低,则会出现相反的情况。

另外,由于不同的图片中人的肤色是不相同的,而且不同人种的肤色有一定的差异,要设置一个确定的阈值是不现实的。故在设定阈值的时候需要对每一张不同的图片对应的阈值计算出自适应阈值,该阈值应该能够最优的区分出肤色与非肤色像素。为此,国内外学者针对这一课题进行了广泛深入的研究和大量实验,提出了多种阈值选择方法。

对比起来,由 Ostu 提出的最大类间方差阈值分割法是建立在最小二乘原理基础上的聚类分割,其方法严密、原理简单,是一种公认有效的传统图像分割方法,其基本思路是将直方图在某一阈值处分割成两组,当被分成的两组的方差为最大时,决定阈值。我们将最大类间方法进行扩展,引入模式识别理论中的 Fisher 评价函数作为判别准则以实现阈值的自动选择。算法核心是利用 Fisher 函数的类间均值最大、总类内方差最小的原则,自动获取某幅图像所对应的最佳的、最易于分类的阈值。

对于两个样本类, Fisher 评价函数为:

$$J_f(T) = \frac{(\bar{m}_1 - \bar{m}_2)^2}{\bar{S}_1^2 + \bar{S}_2^2}$$

$$\text{其中, 各类样本均值 } \bar{m}_i = \frac{1}{N_i} \sum_{y \in \omega_i} y, i=1,2;$$

$$\text{样本类内离散度 } \bar{S}_i^2 = \sum_{y \in \omega_i} (y - \bar{m}_i)^2, i=1,2.$$

可以看出,当两个类别平均值间距离很大,且各类方差很小时, $J_f(T)$ 约取最大值。

Fisher 判别准则已经在字符图像^[6]、医用图像^[7]、显微磨粒图像^[8]等图像领域中开始应用来处理图像分割问题。作者将 Fisher 评价准则用于解决肤色分割问题。

对于本章的研究内容,我们可以把肤色似然图像中的肤色像素和非肤色像素作为使用 Fisher 评价函数分类的两类样本,所解决的问题简单化为:对于图像中的肤色像素和非肤色像素两个类别,利用 Fisher 判别函数在一维空间中找到最好的、最易于分类的阈值。

在一幅图像中,其像素点一般可以分为目标物和背景两类,对于待分割的彩色图像,我们首先根据如前所述的肤色模型得到其对应的肤色似然图像,然后可以设置一定的阈值变化范围,在本文中采用的方法是让阈值从 0.65 开始减少,每次减少 0.1,直到 0.15 为止,计算在不同阈值下图像的 Fisher 评价函数,通过比较其值来确定最佳的阈值来进行分割。显然,图像的 Fisher 评价函数为最大值时,两个类别平均值间距离最大,且两类方差最小,该阈值对图像的分离程度最好。这样就可以实现肤色分割的阈值动态选择,大大提高了分割的效果。

其试验结果如图 3。

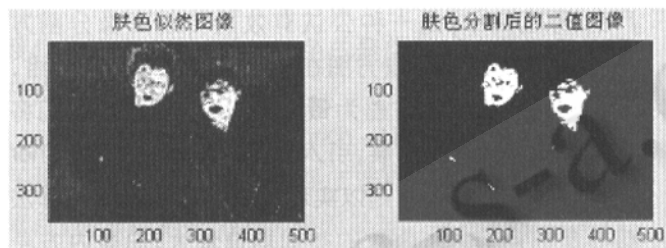


图 3 肤色分割结果

4 结语

肤色的检测在人脸检测和手姿识别方法具有非常重要的应用价值,在人脸检测方面,肤色信息可以用于预处理,也可以用于后期验证。但是,基于肤色检测算法实用化的前提是必须排除光照对肤色特征的影响。本文提出的算法从两个方面考虑了抗光照问题:首先采用非线性分段色彩变换减小亮度的影响,然后提出了基于 Fisher 评价准则的自适应阈值分割方法。不同肤色选用不同阈值进行检测。实验证明该方法对偏黄、黑色肌肤(亚洲、非洲人)和偏白色肌肤(欧美人)均有很好的检测效果,肤色像素检测的正确率达 91.436%,误检率和检测时间相对变化不大,检测效果有明显提高。

参考文献

- 1 Yang M H, Kriegman D, Ahuja N, Detecting faces in images a survey. IEEE Trans. On PAMI, 2002, 24

(1):34-58.

- 2 Elli Angelopoulou. Understanding the Color of Human Skin [C]. Proceedings of the SPIE Conference on Human Vision and Electronic Imaging VI (SPIE), SPIE Press, 2001. 243-251.
- 3 陶霖密、徐光佑, 机器视觉中的颜色问题及应用[J], 科学通报, 2001, 46(3):177-189.
- 4 Y Gong, M Sakauchi. Detection of regions matching specified chromatic features [J]. Computer Vision and Image Understanding, 1995, 61(2): 263-269.
- 5 R-L. Hsu, M. Abdel-Mottaleb and A. K. Jain. Face detection in color images. IEEE Trans. Pattern Analysis and Machine Intelligence, 2002, 24(5): 696-706.
- 6 Jong Oh, Geiger. D. An on-line handwriting recognition system using Fisher segmental matching and Hypotheses Propagation Network. IEEE Conference on, 2000, 2: 343-348.
- 7 Hamarneh. G, Chodorowski, A, Gustavsson. T. Active contour models: application to oral lesion detection in color images. IEEE International Conference on, 2000, 4:2458-2463.
- 8 陈果、左洪福, Fisher 评价准则下的彩色磨粒图像自动分割, 小型微型计算机系统, 2002, 23(6): 721-725.