

人体的肤色特征*

陶霖密, 彭振云, 徐光祐

(清华大学 计算机科学与技术系, 北京 100084)

E-mail: tao_linmi@yahoo.com

http://www.cs.tsinghua.edu.cn

摘要: 肤色特征在 MPEG4、基于内容的图像检索、虚拟现实等领域有着广泛的应用。目前肤色聚类的方法有很多, 结果各异, 但都没有涉及颜色现象的本质。探讨了人体肤色的基本问题: 影响肤色聚类的颜色空间、光源颜色对聚类的影响以及颜色恒常性。提出了用于肤色聚类的新的颜色空间, 给出了光源变化和图像数据之间的关系, 证明了图像中高光和阴影只影响肤色的饱和度。在此基础上, 还给出了人体肤色的聚类结果以及肤色特征在复杂环境中的人脸检测和手势识别中的应用。

关键词: 肤色; 颜色空间; 颜色恒常性; 人脸检测; 手势识别

中图法分类号: TP391 **文献标识码:** A

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。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, 因此, 利用颜色分割人脸的任务成为非常困难的问题。R. Chellappa 的综述文章^[1]被 J. Esch^[2]推荐为关于人类和机器人脸认知的“关键的综述(critical survey)”, “总结了 30 年来生物物理学和神经科学关于人类人脸认知的研究成果”。然而, 这篇文章中却没有提到颜色在人脸认知和机器人脸识别中的应用。

直到最近几年, 颜色在机器人脸认知中的应用才逐渐多起来。C. H. Lee^[3]报道了基于运动和颜色的复杂环境下的人脸定位工作。他们先利用速度场的阈值化来分割运动的人脸, 然后基于人脸中各部位的色调不同来分割眼、眉毛、嘴等区域。M. Kapfer^[4]利用颜色、运动和形状模型检测序列图像中的人脸, 用于极低比特率的可视电话编码。他们首先将图像分割成不同的色块, 再用几何模型从中找出类似于人脸形状的色块, 最后根据候选色块的不同运动特性选出人脸。Y. Dai^[5]根据空间灰度相关矩阵的特征参数建立了人脸质地模型, 并将该模型结合 YIQ 颜色空间中的 I 组份, 用于检测彩色图像中近似于橘黄色部分的人脸。M. H. Yang 和 K. Schwerdt 等人分别根据 LUV 颜色模型^[6]和 RGB 颜色模型^[7]来检测彩色图像中的人脸。Garcia^[8]则采用向量量化的方法, 先将颜色量化, 再检测图像中的肤色区域。这种方法的肤色聚类效果较好, 而分割效果并不好, 因为颜色量化就是一个颜色分割过程, 已经将相似的颜色进行了归类。

所有的这些算法都取得了某种程度的成功, 但也有共同的缺点, 就是没有考虑到环境光照对人脸识别的影响。因此, 这些系统的应用范围都有较大的限制。在最近的文献中, D. Jacobs^[9]和 A. Georghiades^[10]考虑了光照方向对人脸认知的影响, 并在相应的实验中取得了较好的结果。

* 收稿日期: 1999-09-20; 修改日期: 2000-07-31

基金项目: 国家 863 高科技发展计划资助项目(865-306-03-01-1)

作者简介: 陶霖密(1962-), 男, 浙江黄岩人, 博士, 主要研究领域为计算机视觉, 人类视觉认知; 彭振云(1962-), 男, 湖北仙桃人, 博士, 高级工程师, 主要研究领域为计算机视觉, 模式识别, 计算机图形处理; 徐光祐(1940-), 男, 浙江绍兴人, 教授, 博士生导师, 主要研究领域为计算机视觉, 多媒体。

事实上,光照对人脸识别系统的影响远不止点光源的光照方向这一个因素.人脸跟踪和识别系统在不同环境下的应用,影响其识别率的主要因素是光照的变化.因此,在实际应用中,关于人体的肤色特征的一些基本问题依然有待解决:人体的肤色在颜色空间中能否良好聚类?聚类的条件、散布的情况以及光源颜色、高光和阴影对聚类结果的影响如何等等.基于这些考虑,本文介绍了一个新的系统,用于复杂环境下的人脸检测.

1 聚类的颜色空间

在应用肤色特征分割图像之前,首先要回答这样的问题:人类的肤色能不能在颜色空间中聚成单独一类?在什么样的颜色空间中具有更好的聚类结果?是一个自由度的还是两个自由度的聚类?为了实际应用,我们希望在某一颜色空间中,人类的肤色能够聚成密集分布的一类.为了提高基于颜色分割的分辨率,我们希望肤色能够用颜色的两个自由度来聚类.

现有的颜色空间有很多,如 CIE^[11]的 Luv, Lab, XYZ, 计算机图像处理中常用的 YUV, VHS (value, hue, saturation)^[12], IHS (intensity, hue saturation)^[13]以及 GLHS (generalized LHS)^[14]等.在这些颜色空间中, YUV, YIQ, XYZ 颜色空间与 RGB 颜色空间之间是线性关系,而 VHS, IHS, GLHS 颜色空间与 RGB 颜色空间之间是非线性关系,为了使肤色的聚类结果和直接聚类 RGB 空间的结果有明显的改善,我们选择了与 RGB 颜色空间之间有非线性关系的 GLHS 颜色空间,并与肤色聚类实验相结合,提出了新的亮度-饱和度-色调算法.

由于摄像机得到的 RGB 值不能直观地反映人类观察颜色的结果,即亮度、色调、饱和度等颜色属性. LHS (lightness, hue, saturation) 颜色空间建立了 RGB 图像数据和颜色属性之间的映射,又称为认知颜色空间.这一类颜色空间还包括 IHS, VHS 等. Levkowitz 将其统一表达为 GLHS 颜色空间,其基础是以下关于亮度的定义:

$$l(c) = w_{\min} \cdot \min(c) + w_{\text{mid}} \cdot \text{mid}(c) + w_{\max} \cdot \max(c), \quad (1)$$

其中: $w_{\min}, w_{\text{mid}}, w_{\max}$ 均为正数; $\min(c), \text{mid}(c), \max(c)$ 定义为

$$\min(c) = \min(R, G, B),$$

$$\text{mid}(c) = \text{mid}(R, G, B),$$

$$\max(c) = \max(R, G, B).$$

色调是指颜色刺激的主波长,用角度表示.一般定义纯红色 (RGB: 255, 0, 0) 的色调角为 0 度,白色、黑色以及灰色系列的色调角未定义.最原始的色调计算公式由 Tenenbaum 和 Garvey 等人于 1974 年在一份技术报告中给出 (Technical Report 87, AI Center, Stanford Research Institute, 1974). 具体将 RGB 直角坐标系转换为极坐标系的推导见文献 [13, 15]. 许多人提出了简化算法,如 Bajon 提出的 RGB 到 IHS 映射中关于色调的算法^[16].

色调的定义域是: $H \in [0, 360]$. 按照 $\max(c)$ 的不同,可以将其分为 3 个区域:红色区 ($\max(c) = R, H \in [300, 60)$), 绿色区 ($\max(c) = G, H \in [60, 180)$) 及蓝色区 ($\max(c) = B, H \in [180, 300)$). 每一个区域都可以进一步细分为两个小区.这种分区的思想对于颜色聚类很重要,因为肤色的特征色调在红色区内.这一特征表现在图像中,就是肤色的 RGB 值符合下式:

$$R > G > B. \quad (2)$$

肤色的这一特征相当稳定,无论是人种的变化还是环境光照的变化,都不足以造成式 (2) 变成 $G > R$ 或 $B > G$. 因此,肤色色调的聚类结果简单而稳定.同时,这一特征也表明,只要能够实现上述映射的色调算法,对肤色的聚类结果影响就不大. GLHS 模型定义的色调就是在颜色分区的基础上进

一步定义了颜色小区内的角度,具体的公式为

$$h(c) = (k(c) + f(c)) * 60. \tag{3}$$

其中 $k(c)$ 是颜色小区的序号, $f(c)$ 用于在颜色小区内计算角度. 不管这些公式是否符合将 RGB 的直角坐标系转换为极坐标系的要求, 我们的实验则表明, $k(c)$ 的颜色分区对于肤色聚类是一个有效的算法.

$$f(c) = \begin{cases} \frac{\text{mid}(c) - \text{min}(c)}{\text{max}(c) - \text{min}(c)} & \text{if } k(c) \text{ is even} \\ \frac{\text{max}(c) - \text{mid}(c)}{\text{max}(c) - \text{min}(c)} & \text{if } k(c) \text{ is odd} \end{cases}$$

$$k(c) = \begin{cases} 0 & \text{if } R > G > B \\ 1 & \text{if } G > R > B \\ 2 & \text{if } G > B > R \\ 3 & \text{if } B > G > R \\ 4 & \text{if } B > R > G \\ 5 & \text{if } R > B > G \end{cases}$$

饱和度相当于颜色刺激的纯度. 灰色的饱和度定义为 0, 因此, 饱和度是一个与亮度正交的量. 将所有 RGB 向量在亮度正交方向的投影归一化, 就得到 RGB 向量的饱和度. 亮度、色调、饱和度三者之间的关系如图 1 所示. 目前比较通用的饱和度算法是 Foley 和 van Dam 在 HSV 和 HLS 颜色模型中提出的算法^[12].

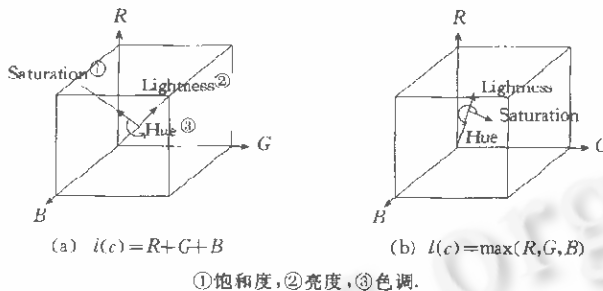


Fig.1 The relationship of lightness, saturation and hue in RGB cube
图1 RGB立方体中亮度、饱和度、色调三者之间的关系

GLHS 模型中定义的饱和度算法为

$$S = \begin{cases} \frac{l(c) - \text{min}(c)}{l(c)} & \text{if } l(c) \leq l(x) \\ \frac{\text{max}(c) - l(c)}{M - l(c)} & \text{if } l(c) > l(x) \end{cases} \tag{4}$$

其中, $l(c)$ 是式(1)所定义的亮度, x 是 GLHS 中定义的和 c 色调相同而亮度不同的一个颜色. 这个饱和度算法有些复杂, 而且肤色的聚类结果也不如上述色调算法那样简单而稳定. 我们用 IHS, LHS 模型中的饱和度算法进行了类似的实验, 结果也不好. 但是, 肤色的饱和度特征是很显著的, 亦即肤色的饱和度偏低. 鉴于此, 我们设计了一个模拟实验来检测饱和度算法的性能, 并在此基础上改进了现有的饱和度算法.

$$S = \begin{cases} \frac{\max(c) - \min(c)}{\max(c) + \min(c)} & \text{else} \\ 0 & \text{if } \max(c) + \min(c) = 0 \end{cases} \quad (5)$$

其中

$$\max(c) = \max\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right),$$

$$\min(c) = \min\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right).$$

该饱和度算法中隐含的亮度定义是

$$l(c) = \max(c) + \min(c). \quad (6)$$

此亮度的物理意义如图2所示. 在RGB立方体中, 等亮度区由6个三角形组成, 不是一个平面. 而在RGB颜色模型中, 等亮度平面和RGB立方体的交平面是正三角形.

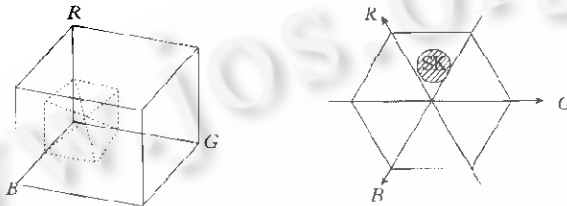


Fig. 2 The equal lightness area in RGB color space and its equal lightness plane (SK is the skin color area)
图2 RGB颜色空间中的等亮度区域及展开后的等亮度平面(SK为肤色区)

在RGB颜色向量中, $\max(R, G, B)$, $\text{mid}(R, G, B)$ 和 $\min(R, G, B)$ 三者之间的关系在GLHS模型中被定义为该颜色向量的色调, 而我们则进一步将 $\max(R, G, B)$ 和 $\min(R, G, B)$ 之间的关系定义为该颜色向量的饱和度. 因此, 式(3)、式(5)、式(6)构成了一个完整的新的LHS颜色模型或颜色空间, 我们称之为TLHS(Tsinghua LHS)颜色模型. 实验表明, 基于TLHS颜色模型的肤色聚类更加稳定和高效.

2 肤色聚类

肤色聚类的任务是在选定的颜色空间内求出肤色的分布. 这一工作由两部分组成: (1) 比较肤色在各颜色空间中的聚类结果, 测试所选择的颜色空间是否合适; (2) 如果颜色空间适合于肤色聚类, 则通过大量的实验求出肤色在该颜色空间中的分布.

测试颜色空间的工作, 是先选择10幅典型的人脸图像, 要求颜色接近, 脸部没有高光和阴影区域. 再在脸部任取一些像素(50~80个), 计算人脸中与该像素“颜色”相近的像素的数目. 以CIE XYZ颜色空间为例, 为了对其中的X维进行聚类, 需从脸部取出50~80个能代表肤色的像素(高光和阴影区除外), 统计脸部与该像素“颜色”相近的像素的数目. 同时, 改变符合条件的像素的颜色, 以确认这些像素是否位于脸部. 具体方法如下.

设 $I(x, y)$ 为图像 I 中的像素, (x, y) 表示该像素在图像中的位置, $X(x, y)$ 表示该像素在CIE XYZ颜色空间中X维的值. $I_j(x, y)$ 和 $X_j(x, y)$ 分别是任一选中的代表肤色的像素及该像素在CIE XYZ颜色空间中X维的值, 则有

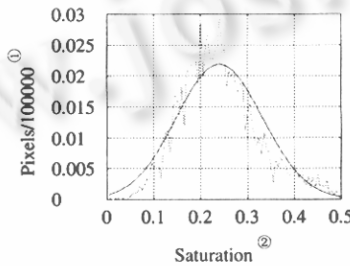
$$\left. \begin{array}{l} \forall \quad I(x, y) \in I, \\ \text{if} \quad X(x, y) \in [X_j(x, y) * (1-k), X_j(x, y) * (1+k)], \\ \text{then} \quad \text{Counter} += 1. \end{array} \right\} \quad (7)$$

其中 k 为计算的宽容度, $k \in [0, 1]$. Counter 是计数器, 计算符合条件的点数. 重复上述计算, 可以得到人脸表面颜色分布的直方图. 对所有 10 个图像的结果进行统计, 可以得到肤色在该颜色空间中的分布, 比较这些结果, 可以从中选择一个较好的颜色空间.

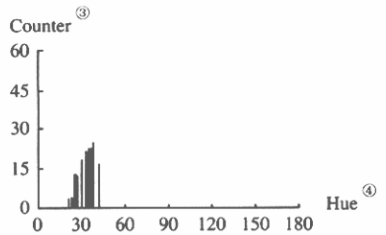
我们采用上述 TLHS 颜色空间, 对人脸数据库中的图像进行了聚类. 实验表明, 肤色在 S 维 (如图 3 所示, 图 3(c) 中横坐标表示所选择的 66 个点的色调 (Hue 值), 纵坐标表示人脸中颜色相近的点的数目 ($Counter \times 160$)) 和 H 维 (如图 4 所示) 都得到了较好的结果. 其中 S 维的聚类结果为正态分布 (如图 3(b) 所示), 而 H 维的典型聚类结果如图 4 所示 (其中横坐标表示所选择点的颜色 (Hue 值), 纵坐标表示人脸中颜色相近的点的数目 ($Counter \times 160$)). 图中 H 维的分布曲线的左半部 (红端) 近似于正态分布, 而右半部 (绿端) 则没有分布. 这个曲线与文献报道的各维均为正态分布的结果^[6,17] 有所不同, 但却符合我们日常的观察结果, 即正常人脸的肤色都带有不同程度的红色, 却不能发青、发绿. “脸都绿了” 表示人脸的一种不正常的状态.



(a) The face image used in the classification
(a) 用于聚类的人脸图像

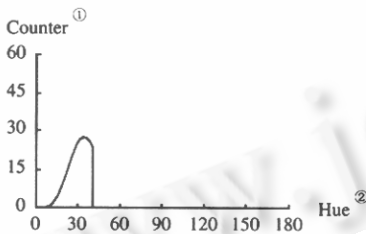


(b) The saturation classification result of 80 face images
(b) 80 幅人脸图像中肤色饱和度的聚类结果
①像素/100000, ②饱和度, ③计数器, ④色调.

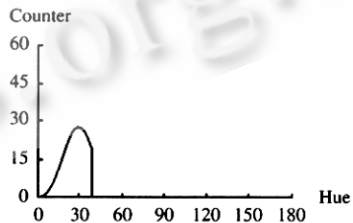


(c) The hue classification result of Fig. 3(a)
(c) 图 3(a) 的色调聚类结果

Fig. 3 Face image and skin color classification
图 3 人脸图像及肤色聚类



(a) High color temperature illumination
(a) 高色温光源



(b) Low color temperature illumination
(b) 低色温光源

①计数器, ②色调.

Fig. 4 A typical result of the skin color classification
图 4 典型的肤色聚类结果

3 颜色恒常性问题

在实际应用上述颜色聚类的结果之前, 我们还必须考虑颜色恒常性问题. 事实上, 目前一般的工作都没有考虑颜色恒常性问题, 文献^[5, 6, 17] 所讨论的颜色空间都是 RGB 颜色空间 (YIQ 和 YUV 颜色空间与 RGB 颜色空间成线性关系). 在该空间内, 肤色值和光源的光谱成分密切相关.

因此,肤色聚类结果只能应用于特定的光照条件下,并且不能有高光和阴影的干扰。

颜色恒常性问题包括两个子问题:(1)人脸表面的高光和阴影部分,其颜色的不同与正常光照下的颜色需要另作处理;(2)上述实验基于人脸在阳光和日光灯下的图像,光源的色温比较高(5500~9500度),不能直接用于低色温的光源(低于5500~8500度)下的图像。

3.1 高光和阴影的消除

对于上述的第1种情况,我们发展了二步法颜色恒常性算法^[18],用于处理人脸中高光和阴影部分。我们假设:

- (1)人脸图像中的高光和阴影部分与其余部分的肤色是相同的,即肤色恒常性。
- (2)入射于人脸表面各部分的光照的光谱成分是相同的,即入射光的颜色是相同的。

非电物物体产生高光的原因是物体表面存在着镜面反射或界面反射,而产生阴影的原因是入射光中的一部分被遮挡,入射于物体表面的光强低于正常光强。在这些情况下,物体表面的反射率可以表示为物体的本体反射 $R_b(\lambda)$ 和反射因子 $R_m(\lambda)$ 两部分。其中

$$R_m(\lambda) = C, \quad (8)$$

C 是常数,表示该反射因子与光的波长无关。如果物体表面存在界面反射,入射光中的一部分被镜面反射,反射光可以表示为

$$I(\lambda)R(\lambda) = \alpha I(\lambda)R_b(\lambda) + (1-\alpha)I(\lambda)R_m(\lambda), \quad \alpha \in [0, 1], \quad C=1. \quad (9)$$

其中 $R(\lambda)$ 是物体的等效反射率, $I(\lambda)$ 是入射光, α 的大小表示镜面反射的强度。此时, $\alpha=1$ 表示没有镜面反射, $\alpha=0$ 表示只有镜面反射。物体表面的阴影也可以用式(10)来表示, α 的大小表示阴影遮挡的强度。此时, $\alpha=1$ 表示没有阴影遮挡, $\alpha=0$ 表示只有阴影遮挡。其中

$$C=0. \quad (10)$$

上述两种情况以及正常的反射,都可以以等效反射率表示为

$$R(\lambda) = \alpha R_b(\lambda) + (1-\alpha)R_m(\lambda). \quad (11)$$

因此,各种情况下的物体表面的反射率函数,都可以统一地用式(11)来描述。镜面反射和阴影都没有改变反射光的主波长,即物体表面颜色的色调。有镜面反射的颜色和另外一种饱和度更低、亮度更高的颜色不可区分。阴影区的颜色和另外一种饱和度更低、亮度更暗的颜色不可区分。换言之,高光和阴影都改变了颜色的饱和度。上述3种情况可以看做是3种不同饱和度、不同亮度、相同色调的颜色在相同光照下的结果。但是,在RGB颜色空间内,包括YIQ和YUV都表现为 (R, G, B) 值的剧烈变化,从而导致聚类的失败。

基于上述分析,我们将原RGB空间中的亮度归一化,改善高光和阴影所造成的饱和度的变化,以消除亮度对聚类的影响。目前,亮度的定义有许多种,除了上文提到的GLHS亮度算法以外,比较常见的还有YUV模型中的亮度 Y :

$$Y = 0.299R + 0.587G + 0.114B. \quad (12)$$

直接应用这些定义是有困难的。因为亮度归一化将影响下一步的饱和度与色调的计算,从而影响整个聚类结果。新的亮度算法既要影响饱和度与色调的计算,又要最大限度地消除亮度的影响。基于这些考虑,我们将亮度定义为 (R, G, B) 向量的模长,即

$$L = (R^2 + G^2 + B^2)^{1/2}. \quad (13)$$

根据这个定义,亮度归一化就是求 (R, G, B) 向量和 (R, G, B) 空间中的单位圆的交点,其结果是将 (R, G, B) 向量归一化到该空间中的单位球面上。这样做的好处是相对于式(6)定义的亮度而言,图

像的亮度依然是不等的,但其动态范围降低了.

3.2 光源色温的处理

光源色温的降低,就是光辐射中高频成分的相对减少以及低频成分的相对增加.这一变化经物体表面的反射,反映在摄像机的采集结果中,就是物体表面的颜色变红.因此,光源色温的降低,经同一物体表面(人脸)反射后,反映在图像数据中就是物体表面的颜色发生红移(脸发红).

这一结论也可以由颜色恒常性的理论计算而得到^[14].图像数据之间的颜色恒常性映射关系为线性关系:

$$\begin{bmatrix} \hat{R} \\ \hat{G} \\ \hat{B} \end{bmatrix} = \hat{T} T^{-1} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (14)$$

其中

$$T = \begin{bmatrix} \sum_{j=1}^3 \alpha_j \Psi_{rj1} & \sum_{j=1}^3 \alpha_j \Psi_{rj2} & \sum_{j=1}^3 \alpha_j \Psi_{rj3} \\ \sum_{j=1}^3 \alpha_j \Psi_{kj1} & \sum_{j=1}^3 \alpha_j \Psi_{kj2} & \sum_{j=1}^3 \alpha_j \Psi_{kj3} \\ \sum_{j=1}^3 \alpha_j \Psi_{bj1} & \sum_{j=1}^3 \alpha_j \Psi_{bj2} & \sum_{j=1}^3 \alpha_j \Psi_{bj3} \end{bmatrix},$$

$$\hat{T} = \begin{bmatrix} \sum_{j=1}^3 \Psi_{rj1} & \sum_{j=1}^3 \Psi_{rj2} & \sum_{j=1}^3 \Psi_{rj3} \\ \sum_{j=1}^3 \Psi_{kj1} & \sum_{j=1}^3 \Psi_{kj2} & \sum_{j=1}^3 \Psi_{kj3} \\ \sum_{j=1}^3 \Psi_{bj1} & \sum_{j=1}^3 \Psi_{bj2} & \sum_{j=1}^3 \Psi_{bj3} \end{bmatrix},$$

其中 Ψ_{ijk} 为常数, α_j 为与光照有关的参数.这是未知环境光照下的图像数据和标准环境光照条件下的图像数据之间的映射关系.高、低色温图像数据之间的关系也可以用上式来表示.换言之,光源色温的变化对图像数据的影响,就是变化前后坐标系的移动.因此,环境光照对聚类结果的影响,既可以通过颜色恒常性算法计算,也可以由实验、学习算法得到.

色温的高低在图像数据中是不可知的.而且,脸发红的原因也不一定是色温降低,也可能是人脸的个体差异、喝酒等等.但是,有一点是可以肯定的,就是“红移”.“红移”幅度的大小和色温改变的多少有关.实验表明,在从蜡烛光下采集到的图像中,其聚类范围和正常光照下的聚类结果的重叠部分较小,影响实际使用.而在普通的白炽灯下采集到的图像与灯的功率有关,其聚类范围和现有的结果有较大的重叠范围,并不影响使用(如图5所示).

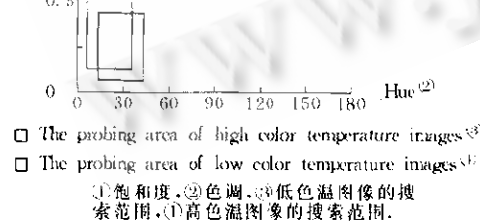


Fig. 5 The changing of the probing area
图5 搜索范围的变化

影响使用(如图5所示).

4 肤色分割实验

在上述恒常性变换的基础上,我们再进行颜色空间的转换和计算,找出符合要求的点,以实现基于肤色的人脸区域分割.以图3为例,将图3(a)作颜色空间的转换和基于肤色的二值化.符合要求的点被置为黑色,不符合要求的点被置为白色,实际运行结果如图6(a)所示.从图中可以看出,脸部有V字形的高光区,鼻梁上也有高光区,颜色与正常的肤色有较大的差别,基于肤色的二值化时,这些区域被作为非肤色区.在增加了颜色恒常性算法^[19]以后,我们可以看到,脸部和鼻梁上的高光区已得到了恢复(如图6(b)所示).鼻梁上依然残留的一些高光区,经检查原始数据发现,其光强已超过摄像机的动态范围,所记录的数据已不包含肤色信息,即其原始数据为(255,255,255).

上述算法适用于色温相近的光源,如阳光、日光灯以及摄影用的碘钨灯等.以采集的电视图像为例,上述算法可以跟踪电视中的人脸.实际的实验结果如图7和图8所示.

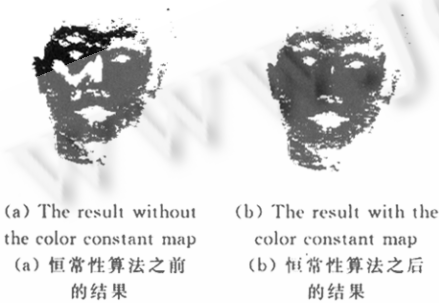


Fig. 6 Face segmentation based on skin color

图6 基于肤色的人脸区域分割



Fig. 7 The experimental result with the image captured from TV (news, outdoor)

图7 电视图像的实验结果(新闻,阳光下)

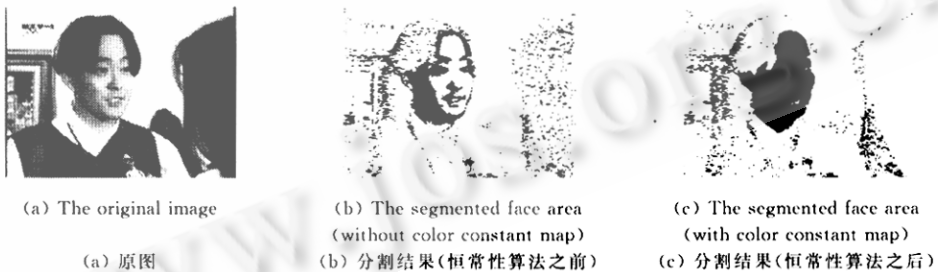


Fig. 8 The experimental result with the image captured from TV (TV play, indoor)

图8 电视图像的实验结果(电视剧,摄影灯光下)

5 结论

本文全面讨论了肤色聚类所涉及到的颜色空间、物体表面的反射性能、环境光照变化的影响以及聚类方法的问题.我们的分析表明,非均匀颜色空间有其特殊的应用场合,基于文中所定义的颜色空间,我们对肤色的聚类得到了较好的结果.

本文证明了物体表面的高光和阴影对聚类的影响只表现在肤色的亮度和饱和度的变化上,而其色调却不变.基于这一结果和肤色恒常性,我们设计了亮度归一化算法,消除了图像中高光和阴影对于肤色聚类的影响.

为了使该算法能够适应广泛的光源颜色,从而具有广泛的应用范围,本文还证明了环境光照变化对肤色聚类结果的影响只限于图像数据坐标系的变换.这一结果为肤色聚类的颜色恒常性算法、学习算法提供了坚实的基础.

我们的实验表明,人类的肤色在一定的颜色空间内是能够聚成一类的.基于这个结果,我们发展了颜色恒常性算法,用于处理人脸中高光和阴影的问题,以恢复人脸表面的本体反射颜色.尽管这个结果适合于比较广泛的场合,对光源的颜色也并不敏感,但是也只能适合于有限的场合,对于光源色温的大幅度变化、人脸颜色的异常变化等特殊任务,仅仅依靠这个算法是不够的.更重要的是,这个算法分割的结果只是分割出图像中颜色和人类肤色相近的区域,并不能确认这些区域是否是人脸.

从上述实验图像可以看出,分割的结果中包含有许多噪声.同时,从图像数据中计算出光源色温的高低是很困难的,在有些图像中是不可能的.为了使上述算法能应用于不同色温场合以及不同的个体,我们需要基于特征的人脸识别技术^[18,20],并从中得到反馈,以改变搜索范围,从而使算法能够根据光源色温的变化而变化,成为自适应的算法.同样,运动也是人的一个重要特征,在连续采集的图像序列中,人总是处于运动之中的.我们将基于特征的人脸识别技术、基于图像序列的运动检测技术与颜色分割技术相结合,形成稳定的、多特征的人脸检测系统.将上述方法用于基于颜色和运动的手势分割实验^[21],也得到了良好的结果.

总之,我们的工作表明,复杂环境下的人脸检测系统必须利用人脸的多种特征,在多种数据融合的基础上,才有可能得到稳定的、复杂环境下的人脸检测系统.

References:

- [1] Chellappa, R., Wilson, C. L., Sirobey, S., Human and machine recognition of faces: a survey. *Proceedings of the IEEE*, 1995, 83(5): 705~740.
- [2] Esch, J. Prolog to human and machine recognition of faces: a survey. *Proceedings of the IEEE*, 1995, 83(5): 704.
- [3] Lee, C. H., Kim, J. S., Park, K. H. Automatic human face location in a complex background using motion and color information. *Pattern Recognition*, 1996, 29(11): 1877~1889.
- [4] Kapfer, M., Benois-Pineau J. Detection of human faces in color image sequences with arbitrary motions for very low bit-rate videophone coding. *Pattern Recognition Letters*, 1997, 18(14): 1503~1518.
- [5] Dai, Y., Nakano, Y. Face-Texture model based on SGLD and its application in face detection in a color scene. *Pattern Recognition*, 1996, 29(6): 1007~1017.
- [6] Yang, M. H., Ahuja, N. Detecting human faces in color images. In: *Proceedings of the International Conference on Imaging Processing*, Vol 1. 1998. 127~130.
- [7] Schwerdt, K., Crowley, J. L. Robust face tracking using color. In: *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*. 2000. 90~95.
- [8] Garcia, C., Tairitas, G. Face detection using quantized skin color regions merging and wavelet packet analysis. *IEEE Transactions on Multimedia*, 1999, 1(3): 264~277.
- [9] Jacobs, D. W., Belhumeur, P. N., Basri, R. Comparing images under variable illumination. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998. 610~617.
- [10] Georgiades, A. S., Kriegman, D. J., Belhumeur, P. N. Illumination cones for recognition under variable lighting: faces. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998. 52~58.
- [11] CIE. *Colorimetry*. 2nd, Vienna: CIE, 1986.
- [12] Foley, J. D., van Dam, A. *Fundamentals of Interactive Computer Graphics*. Reading, MA: Addison-Wesley, 1982.
- [13] Gonzalez, R. C., Woods, R. E. *Digital Image Processing*. 3rd ed. Reading, MA: Addison Wesley, 1992.
- [14] Levkowitz, H. *Color Theory and Modeling for Computer Graphics, Visualization, and Multimedia Applications*. Boston: Kluwer Academic Publishers, 1997.
- [15] Ledley, S., Baas, M., Golab, T. Fundamentals of true-color image processing. In: *Proceedings of the 10th International*

- Conference on Pattern Recognition, 1990, 791~795.
- [16] Bajon, J., Cattoen, M., Kim, S. D. Real-time colorimetric transformations used in robot vision. In: Proceedings of the MICAD, 1985, 76~86.
- [17] Yang, J., Lu, W., Waibel, A. Skin-Color modeling and adaptation, Technical Report, CMU CS 97-146, School of Computer Science, Carnegie Mellon University, 1997.
- [18] Peng, Zhen-yun, Tao, Lin-mi, Xu, Guang-you. The self-adaptive face detection based on skip color and features. In: Gao, Wen, Qian, Yue-liang, eds. Proceedings of the 4th National Conference on the Computer Intelligent Interface and Intelligent Application. Beijing: Publishing House of Electronics Industry, 1999, 212~217 (in Chinese).
- [19] Tao, Lin-mi, Xu, Guang-you. Computer color constancy in two steps. Journal of Tsinghua University (Science and Technology), 2000, 40(7): 101~104 (in Chinese).
- [20] Tao, Lin-mi, Peng, Zhen-yun, Xu, Guang-you. The features of skip color and the face detection in complex background. In: Gao, Wen, Qian, Yue-liang, eds. Proceedings of the 4th National Conference on the Computer Intelligent Interface and Intelligent Application. Beijing: Publishing House of Electronics Industry, 1999, 205~211 (in Chinese).
- [21] Zhu, Yuan-xin, Ren, Hai-ben, Xu, Guang-you, et al. A panoramic browser based on hand gesture input. In: Gao, Wen, Qian, Yue-liang, eds. Proceedings of the 4th National Conference on the Computer Intelligent Interface and Intelligent Application. Beijing: Publishing House of Electronics Industry, 1999, 218~223 (in Chinese).

附中文参考文献:

- [8] 陶霖密, 彭振云, 徐光祐. 人类的肤色特征及复杂环境下的人脸检测. 见: 高文, 钱跃良, 编. 第4届全国计算机智能接口与智能应用学术会议论文集. 北京: 电子工业出版社, 1999, 205~211.
- [17] 祝远新, 任海兵, 徐光祐, 等. 基于手势输入的全景图浏览器. 见: 高文, 钱跃良, 编. 第4届全国计算机智能接口与智能应用学术会议论文集. 北京: 电子工业出版社, 1999, 218~223.
- [19] 陶霖密, 徐光祐. 计算机视觉的三步法颜色恒常性. 清华大学学报(自然科学版), 2000, 40(7): 101~104.
- [21] 彭振云, 陶霖密, 徐光祐, 等. 基于颜色和特征的自适应人脸检测. 见: 高文, 钱跃良, 编. 第4届全国计算机智能接口与智能应用学术会议论文集. 北京: 电子工业出版社, 1999, 212~217.

The Features of Skin Color*

TAO Lin-mi, PENG Zhen-yun, XU Guang-you

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

E-mail: tao_linmi@yahoo.com

http://www.cs.tsinghua.edu.cn

Abstract: The features of skin color are wide used in the areas of MPEG4, content based image retrieve and virtual reality. Presently, there are plenty of the methods of the skin color clustering and of the varieties of the clustering results. Some elementary problems of the features of skin color, such as the effect of light source, of color constancy and of color space in which the skin color clustering is performed, are investigated in this paper. A new color space is constructed for the skin color clustering and the correlation between the image data and the color of illuminants is introduced. It is proved that high lights and shading only changes the saturation of the skin color. In this framework, both the results of the skin color clustering and their applications in the facial detection and hand gesture recognition in complex background are presented.

Key words: skin color; color space; color constancy; face detection; hand gesture recognition

* Received September 20, 1999; accepted July 31, 2000

Supported by the National High Technology Development Program of China under Grant No. 863-306-03-01-1