

# 基于颜色的皮肤检测综述

徐战武 朱森良

(浙江大学计算机学院, 杭州 310027)

**摘要** 在系统回顾和比较了基于颜色的皮肤检测的方法和技术(其中包括:颜色空间选择、肤色建模方法、动态跟踪模型以及光照不变性与自适应模型)的基础上;基于一个包含1894张图片的大样本库,着重比较了肤色在14个3维颜色空间和14个2维色度平面中的分布紧致性、肤色与非肤色类之间的可分辨性,以及肤色概率图(SPM)、高斯混合模型(GMM)、自组织映射图(SOM)和支持向量机(SVM)在这些颜色空间中的皮肤分类性能。比较结果表明:(1)颜色空间的变换并不能改善肤色紧致性、肤色-非肤色可分辨性以及分类等性能,但RGB及线性变换空间却具有较好的类可分辨性和分类性能;(2)去除亮度信息将明显降低肤色和非肤色之间的可分辨性和分类性能;(3)Bayes决策下的3维SPM的分类性能是最优和空间无关的,而其余分类器则普遍存在类似的“空间偏好性”;(4)同时采用肤色和非肤色模型的分分类器的分类性能优于仅使用肤色模型的分分类性能。

**关键词** 皮肤检测 颜色空间 肤色模型 动态模型 光照不变性

**中图分类号**: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2007)03-0377-12

## Color-based Skin Detection: A Survey

XU Zhan-wu, ZHU Miao-liang

(College of Computer Science, Zhejiang University, Hangzhou 310027)

**Abstract** This paper presents a comprehensive review and evaluation for color-based skin detection. The literature is reviewed in four categories: colorspace selection, skin color modeling, dynamic modeling and illumination invariance and adaptation. Using a large data set of 1894 images, we examine whether the colorspace transformation can increase the compactness of skin distribution and the discriminability between skin and nonskin distributions in fourteen 3D colorspace and fourteen 2D chrominance planes. We also evaluate the classification performance of skin probability map (SPM), Gaussian mixture model (GMM), self organizing map (SOM) and support vector machine (SVM) in said colorspace. The results reveal that 1) the colorspace transformation cannot improve performance in general, the discriminability and testing performance in RGB and linear colorspace are better than in other colorspace, 2) the absence of the luminance component decreases discriminability and performance significantly, 3) the performance of Bayes SPM in 3D colorspace is superior to that of others, 4) except 3D Bayes SPM, the 'color preference' of other detector is intrinsic and quite similar, 5) the detector using skin and nonskin model simultaneously is better than the detector using skin model alone.

**Keywords** skin detection, colorspace, skin color model, dynamic model, illuminant invariance

## 1 引言

皮肤检测一方面是人脸检测和跟踪、手势识别等任务的基础和关键步骤;另一方面,由于互联网上大量包含敏感色情内容的图片和视频,使得广大青

少年家长感到非常苦恼,而采用基于内容的图像检索技术来标记裸露内容,以过滤可能的色情图片是有效的解决方案<sup>[1-3]</sup>,其中,皮肤检测是其至关重要的一个步骤。

基于颜色的皮肤检测具有简单快速,不受物体形状变化和视点改变等影响的优点,是皮肤检测领

收稿日期:2005-09-19;改回日期:2006-02-23

第一作者简介:徐战武(1973~),男,助理研究员。2006年获浙江大学计算机学院博士学位。主要研究方向为基于内容的图像与视频检索、模式识别与智能系统等。E-mail: xu.zhanwu@gmail.com

域中最活跃的研究方向。众所周知,皮肤由真皮层和覆盖其上的较薄的表皮层构成,光在表皮层中被黑色素吸收,而在真皮层中则同时发生吸收和散射。不同个体的肤色差异主要表现为由表皮层中黑色素的浓度不同所引起的亮度变化,其真皮层光学特性则基本相同,而且同种族的个体肤色具有较强的共性,并明显区别于大多数背景颜色,在颜色空间中形成一个小而紧致的聚簇<sup>[4]</sup>。Angelopoulou指出,不同种族、性别和年龄个体的皮肤光谱反射率曲线具有很好的相关性<sup>[5]</sup>,人类学家 Jablonski 发现,皮肤反射率与紫外线辐射水平紧密相关,并从生物学上论证了肤色的一致性<sup>[6]</sup>。可见,采用颜色进行皮肤检测具有统计和物理上的有力依据。

本文首先全面比较和分析了基于颜色的皮肤检测技术,包括颜色空间选择、肤色建模方法、基于物理的模型、动态模型、光照不变性和自适应算法等。颜色空间选择和肤色建模方法是其中的两个关键问题。尽管皮肤检测通常需要颜色信息与其他线索,如运动、纹理和形状特征结合,但这里仅分析与颜色特征相关的内容。

## 2 颜色空间选择

### 2.1 皮肤检测中采用的颜色空间

人们在肤色检测中采用了各种各样的颜色空间,包括常用的颜色空间,如 RGB、YCbCr、HSV、 $L^*a^*b^*$  等空间,一些作者则构造了比较特别的颜色空间,如 E-R/G-H<sup>[7]</sup>、H-GY-Wr<sup>[8]</sup> 和 IRgBy<sup>[2]</sup> 空间、离散 KL (Karhunen-Loeve) 变换空间<sup>[9]</sup> 等。其中,RGB、YCbCr、HSV 和归一化 rg 空间是肤色检测中应用最为广泛的颜色空间。根据各颜色空间的特点以及它们与 RGB 空间之间的变换关系可以分为以下主要 5 类:

(1) RGB 及线性颜色空间 RGB 空间是基于视觉三原色原理,直接面向硬件的颜色空间,但各通道间高度相关、亮度与色度混合是其主要缺陷,文献[4,10~12]等采用了该空间。这里将与 RGB 空间具有线性变换关系的颜色空间称为线性颜色空间,包括 XYZ、YIQ、YCbCr、YES 等空间,其中 Y 表示亮度,它们都是亮度分离与设备无关的。XYZ 采用假想的标准原色,其能够完全匹配自然界中所有色光,是色度学的实际应用工具;YIQ 空间很好地匹配了人眼对颜色的不同敏感程度,可应用于电视传

输,并兼容原有的黑白模拟系统;YCbCr 空间是 MPEG 数字视频编码标准建议的颜色空间,由于其在视频压缩域无须解压就能够直接分析其颜色属性,因此成为最具吸引力的肤色空间<sup>[13-19]</sup>。由离散 Karhunen-Loeve 变换得到的线性空间称为 KLT 颜色空间<sup>[9]</sup>,不过该空间的各通道没有明确的物理含义,并具有样本依赖性。

(2) 归一化色品坐标 色品坐标 nrgb 表示了 RGB 空间中 3 个通道之间的比例关系,即

$$r = R/(R+G+B), g = G/(R+G+B), b = B/(R+G+B)$$

由于 nrgb 对亮度的依赖性大大减少,而且对光照方向具有一定的不变性<sup>[20]</sup>,因此获得了研究者的青睐<sup>[21-27]</sup>。不过,在归一化过程中,其不可逆转地损失了亮度信息,而且当亮度很小时,比例值变得不稳定。类似地,归一化坐标 nxyz 表示 XYZ 空间中 3 个通道的比例关系。

(3) 感知均匀空间  $L^*a^*b^*$  和  $L^*u^*v^*$  空间是基于视觉系统的具有拮抗特性的色彩空间,也是国际照明委员会(CIE)建议的感知均匀空间,其中  $L^*$  表示亮度,该空间中任意两点之间的欧氏距离代表了人类观察的感知差别,是与设备无关的。文献[28]选择了  $L^*a^*b^*$  空间,而文献[29~31]则采用了  $L^*u^*v^*$  空间来描述肤色。

(4) 知觉颜色空间 人们更习惯于使用基于知觉感受的颜色系统来描述颜色,如 HSI、HSV 和 HSL 等空间,其中 H 代表色调,对应主波长,S 是饱和度,表示颜色的纯度。HSI、HSV 和 HSL 空间的饱和度和亮度表示方法各异,并具有不同的分布形状和动态范围。由于亮度分离以及颜色的直觉性,使得人们普遍认为知觉颜色系统应该具有较高的分类性能,因此在肤色检测中很受欢迎<sup>[32-37]</sup>。TSL (tint, saturation, lightness)<sup>[38,39]</sup> 空间是由 nrgb 变换引出的空间,并具有较好的知觉特性。知觉颜色空间中灰度轴上的色调没有定义,当亮度接近零时,则色调和饱和度会变得不稳定。

(5) 组合颜色空间 E-R/G-H 和 H-GY-Wr 空间是基于构造性归纳算法自动构造的最优组合颜色系统。文献[7,8]首先选择肤色分布紧密且与非肤色交迭少的颜色分量作为第 1 分量,然后根据互补性原则,选取其余的分量,形成了最优的 E-R/G-H<sup>[7]</sup> 空间和 H-GY-Wr<sup>[8]</sup> 空间。

还有一些颜色空间,如 IRgBy<sup>[2]</sup>、SCT (spherical coordinate transform)<sup>[40]</sup> 和 Farnsworth UCS (uniform

color system)<sup>[41]</sup>等空间也在皮肤检测中得到了一定应用。其中,IRgBy 和 SCT 空间可以看作是一种知觉颜色空间,而 Farnsworth UCS 空间则是一种感知均匀空间,不过,采用这些空间的研究者非常有限。

## 2.2 最优颜色空间

虽然颜色空间的选择对于肤色检测是非常重要的,但是否存在最优的肤色空间呢?大多数研究者直观地选择了“最优”颜色空间而没有给出严格的证明,只有很少的研究者认真考虑了颜色空间选择问题,并对所选空间的最优性进行了论证<sup>[6,39,42,43]</sup>。

一些研究者质疑颜色空间变换对肤色检测结果的明显影响,Shin 以 RGB 空间为基准,比较了 nrgb、XYZ、YUV、YIQ、YCbCr、L\* a\* b\*、HSI 和 SCT 等空间,及其去除了亮度分量的 2 维色度平面的肤色-非肤色的可分辨性<sup>[44]</sup>,结果表明,各颜色空间的可分辨性因指标而异,RGB 空间具有最高的肤色-非肤色类的可分辨性,而去除了亮度分量则明显降低了可分辨性。在此基础上,Jayaram 比较了上述空间中肤色概率图(skin probability map, SPM)和高斯模型的肤色分类性能<sup>[45]</sup>,发现 SCT、HSI 和 L\* a\* b\* 空间中的 3 维 SPM 具有较好的分类性能。Albiol 从理论上证明了最优皮肤检测算法是与颜色空间无关的<sup>[46]</sup>,并验证了这一结论。

笔者首先构造了一个包括不同种族、性别和年龄的个体在不同光照下的 1 894 张图像的样本库,并用手工标记出每个图像中的实际肤色点,共获得超过  $1.04 \times 10^8$  个肤色点和  $2.96 \times 10^8$  个非肤色点;然后以 RGB 空间为基准比较了各颜色空间中的肤色紧致性和肤色-非肤色的可分辨性,并采用配对样本 t 检验来验证去除亮度信息是否能够提高肤色分布紧致性和类间可分辨性;最后考察了各颜色通道的分辨率变化对可分辨性的影响。

### 2.2.1 考察的颜色空间

笔者选择肤色检测中比较常用的颜色空间以及一些作者提出的分类性能较好的颜色空间,包括 RGB、XYZ、YCbCr、YIQ、YES、KLT 等空间(在这里,离散 KL 变换公式为:

$$K = -0.26R + 0.47G - 0.27B$$

$$L = -0.52R - 0.03G + 0.45B$$

$$T = 0.29R + 0.36G + 0.36B$$

其中,  $K$  和  $T$  分别对应最小和最大的特征值)以及 nrgb、nxyz、L\* a\* b\*、L\* u\* v\*、HSI、HSV、HSL、TSL、E-R/G-H 以及 H-GY-Wr 空间共 16 个颜色空间。其

中,nrgb、nxyz 空间本质上是 2 维(可取其中的任意 2 维,这里取  $r, b$  及  $x, z$ )的,而 E-R/G-H、H-GY-Wr 空间则没有明确的亮度通道,除此以外,其余的 12 个 3 维颜色空间,则同时将去除了亮度信息的 2 维色度平面(RGB 空间中的  $G$  可视为亮度,而在 KLT 空间中,则将与最大特征值对应的通道  $T$  当作亮度)作为考察对象。因此,总共包括 14 个 3 维颜色空间和 14 个 2 维“纯色度”平面。

### 2.2.2 采用的性能指标

对所选择的颜色空间的优劣进行评价,通常有以下两类衡量标准:一是在确定基于该空间的模型之后,计算训练和测试集的分类错误,这是有名的分类器性能评估法则;二是肤色与非肤色在颜色空间的交迭以及肤色的紧致性,这类指标是与建模策略无关的,可用来评估颜色空间一般意义上的品质因数。这里采用基于颜色空间本身的第 2 类指标(第 1 类指标将在第 3 节讨论)。

肤色分布熵  $E(E = -\sum_{i=1}^N p_i \ln p_i)$ , 其中  $p_i$  表示颜色  $i$  出现的概率)在一定程度上反映了肤色分布的混乱程度,分布越均匀,熵越大,反之则相反,因而熵  $E$  可以用来衡量肤色分布的紧致性。在一个量化为  $N$  级的空间中,熵的上界为  $\ln N$ 。Shin 提出用 4 个指标来描述肤色与非肤色类之间的可分辨性<sup>[44]</sup>,其中,直方图交迭比(histogram intersection, HI)和直方图平方误差(histogram chi error, HCE)用于衡量肤色与非肤色直方图的相似性; $\text{tr}[S_w]/\text{tr}[S_b]$ 和 $\text{tr}[S_w^{-1}S_b]$ 则是基于离散度矩阵的 2 个指标( $S_w$ 和 $S_b$ 分别是类内离散度矩阵和类间离散度矩阵)。显然,相对较低的熵  $E$ 、直方图交迭比和 $\text{tr}[S_w]/\text{tr}[S_b]$ ,以及较高的直方图平方误差和 $\text{tr}[S_w^{-1}S_b]$ 代表比较理想的类内紧致性和类间可分辨性,这是人们所期望的。

### 2.2.3 各颜色空间的性能指标

表 1 给出了各 3 维颜色空间和 2 维色度平面中的肤色分布紧致性和肤色-非肤色的可分辨性指标,其中每个颜色通道量化成 256 级,粗体字表示其相应的分布紧致性或可分辨性优于原始的 RGB(或 RB)空间。可见,在各颜色空间中,只有肤色紧致性是一致提高的,而各可分辨性指标对颜色空间变换的影响是不同的。具体来说:(1)对直方图交迭比和直方图平方误差而言,3 维空间变换并没有带来任何的性能(指肤色分布紧致性和肤色-非肤色的可分辨性)提升,反而都有不同程度的下降,而在 2 维平面

表 1 各颜色空间的肤色分布紧致性和肤色-非肤色可分辨性

Tab. 1 The compactness of skin distribution and the discriminability between skin and nonskin distributions in fourteen 3D colorspace and fourteen 2D chrominance planes

颜色空间	$E$		直方图交迭比		直方图平方误差		$\text{tr}[S_w]/\text{tr}[S_B]$		$\text{tr}[S_w^{-1}S_B]$	
	3 维	2 维	3 维	2 维	3 维	2 维	3 维	2 维	3 维	2 维
RGB	12.548	9.932	0.344	0.416	1.030	0.841	24.147	17.743	0.331	0.305
XYZ	12.324	9.686	0.346	0.423	1.024	0.827	37.617	38.453	0.331	0.288
YCbCr	12.489	7.779	0.345	0.419	1.025	0.834	20.357	4.471	0.331	0.285
YIQ	12.091	7.394	0.347	0.418	1.021	0.836	21.379	4.349	0.331	0.285
YES	12.173	7.586	0.347	0.414	1.022	0.847	20.613	4.427	0.330	0.284
KLT	12.224	7.713	0.346	0.375	1.022	0.926	20.218	3.769	0.331	0.331
$L^*a^*b^*$	11.343	7.570	0.349	0.473	1.015	0.709	19.589	13.307	0.201	0.110
$L^*u^*v^*$	11.496	7.848	0.348	0.526	1.017	0.601	21.141	13.319	0.215	0.105
HSI	12.209	8.690	0.345	0.452	1.025	0.754	62.521	96.856	0.076	0.024
HSV	12.477	8.867	0.344	0.452	1.027	0.757	20.547	49.299	0.188	0.038
HSL	12.336	8.712	0.345	0.452	1.026	0.754	49.223	79.527	0.106	0.027
TSL	12.128	8.103	0.346	0.457	1.023	0.746	63.323	872.158	0.063	0.003
nrgb		7.420		0.462		0.731		23.615		0.048
nxyz		6.340		0.468		0.719		25.239		0.041
E-R/G-H	9.094		0.362		0.979		29.491		0.315	
H-GY-Wr	9.260		0.369		0.966		9.495		0.226	

中,只有 KL 和 ES 有所提升;(2)对  $\text{tr}[S_w]/\text{tr}[S_B]$  而言,除 XYZ 空间外的线性颜色空间、感知均匀空间及其 2 维色度平面,以及 HSV 和 H-GY-Wr 空间的肤色分类性能均有所提升;(3)除 YCbCr、KLT 和 KL 空间外,其他空间的  $\text{tr}[S_w^{-1}S_B]$  都有不同程度的下降。

归一化坐标和组合空间是以信息损失为代价的,在获得了最优紧致性的同时,类间可分辨性却出现了较大的恶化。线性颜色空间(除 XYZ 空间外)的  $\text{tr}[S_w]/\text{tr}[S_B]$  普遍有所提升,而其他可分辨性指标只有轻微程度的下降,并具有空间变换的无损可逆和速度快的优点,是比较理想的颜色空间;KL 则是可分辨性最好的 2 维色度平面。

#### 2.2.4 去除亮度信息的影响

笔者采用配对样本 t 检验法来对除 nrgb、nxyz 和 E-R/G-H、H-GY-Wr 空间外的 12 对 3 维颜色空间及其 2 维色度平面进行统计检验,以验证去除亮度信息能否提高肤色分布的紧致性以及肤色-非肤色类的可分辨性。对于值越小越好的指标  $\text{tr}[S_w]/\text{tr}[S_B]$ 、熵  $E$  和直方图交迭比,可用 3 维指标减去相应的 2 维指标,而其余 2 个指标则相反,以获得直观一致的结果。表 2 为显著性水平为 0.05 时的双侧检验结果。由表 2 可见,去除亮度信息后明显提升了肤色类的分布紧致性,而直方图交迭比、直方图平方误差和  $\text{tr}[S_w^{-1}S_B]$  的性能则明显恶化,  $\text{tr}[S_w]/\text{tr}[S_B]$  也没有明显差别。

表 2 各指标的配对 t 检验

Tab. 2 Paired samples t-test results

指标	均值	标准差	下界	上界	统计量	显著性
$E$	3.830	0.721	3.372	4.288	18.399	0.000
直方图交迭比	-0.094	0.037	-0.117	-0.070	-8.670	0.000
直方图平方误差	-0.237	0.082	-0.289	-0.184	-9.991	0.000
$\text{tr}[S_w]/\text{tr}[S_B]$	-68.084	234.081	-216.811	80.644	-1.008	0.335
$\text{tr}[S_w^{-1}S_B]$	-0.062	0.040	-0.088	-0.037	-5.413	0.000

### 2.2.5 量化等级对可分辨性的影响

当每个颜色通道量化成 256 级时,3 维直方图将占据很大的存储空间,人们总希望在不明显降低可分辨性的同时,采用较少的量化等级以减少存储代价。通过比较量化等级分别为 256、128、64、32 和 16 的上述 4 个肤色-非肤色可分辨性指标可得知:直方图交叉比和直方图平方误差随着分辨率逐渐变粗而一致下降,当量化级数减少到 16 时,肤色可分辨性加速下降;而基于离散度矩阵的指标则有升有降,分辨率为 32 是一个比较明显的分界点,在这一点上出现了指标加速下降,甚至改变方向或者加速上升的现象。

### 2.2.6 小结

关于颜色空间选择的主要结论有:(1)不同的性能指标对不同的颜色空间具有不同的影响,且普遍意义上的最优空间是不存在的,RGB 空间和线性颜色空间的综合分类性能相对较好,RGB 空间和 KL 空间分别是 3 维和 2 维空间中的最优空间;(2)去除亮度信息可在提升肤色分布紧致性的同时,大大降低了肤色-非肤色的可分辨性;(3)量化等级 32 是肤色和非肤色的可分辨性和存储代价比较好的折中点。

## 3 肤色模型

肤色模型可分为基于物理的模型和统计模型两种,而统计模型还可进一步分为参数、非参数和半参数 3 种类型。参数模型具有明确的函数形式,可通过训练来调整参数以获得与样本数据集相适合的模型,而非参数模型则没有任何的特定形式,半参数方法一般指神经网络方法,它们具有相同的函数形式,以及不同数目的隐式参数。

### 3.1 非参数模型

由一组判别规则构成的阈值边界模型<sup>[13,14,34,35]</sup>是最简单的分类器,其适用于精度要求低、光照条件稳定的场合,而确定合适的颜色空间及良好的决策规则是其主要困难,Gomez 提出了一种自动构建颜色空间以及决策规则的方法<sup>[7,8]</sup>,用来克服这一困难。

SPM 实际上是一个基于归一化直方图的查找表,它的每个柄用于直接存储肤色概率,由于其能够精细地描述肤色分布概率密度而不管分布的内在复杂性,因而得到了非常广泛的应用<sup>[4,10,11,43]</sup>。SPM 的精度受到量化等级的影响,即数目太多的模型非常嘈杂,而数目太少,则导致概率密度过于粗糙,只

有当柄数保持在相对较小并具有足够数量的有代表性的数据时,SPM 才是比较有效的,Jones 认为量化为 32 级的 SPM 综合分类性能最优<sup>[4]</sup>。Wu 将 SPM 当作肤色的离散隶属度<sup>[41]</sup>,采用模糊匹配方法来检测肤色。非参数模型虽具有分类速度快的明显优势,但不能进行内插和数据归纳是其主要缺陷。

### 3.2 参数模型

参数模型提供了一个关于肤色分布的明确而简洁的表示,其不仅占据的空间很小,并能够通过内插来推广不完全的训练数据。参数模型的拟合优度和性能很大程度上依赖于相应颜色空间的肤色分布形状。

高斯模型是最常见的参数模型,其包括单模式高斯模型(single Gaussian model, SGM)和高斯混合模型(Gaussian mixture model, GMM)。个体肤色在 2 维色度平面上集中分布在面积很小的区域,其概率密度函数可用 SGM 来表示<sup>[12,15,16,21,22,47,48]</sup>:

$$p(\mathbf{c} | s) = (2\pi)^{-1} |\Sigma_s|^{-d/2} e^{-(\mathbf{c} - \boldsymbol{\mu}_s)^T \Sigma_s^{-1} (\mathbf{c} - \boldsymbol{\mu}_s) / 2} \quad (1)$$

其中, $\mathbf{c}$  是颜色向量,参数  $\boldsymbol{\mu}_s$  和  $\Sigma_s$  是肤色类  $s$  的均值向量和协方差矩阵,下标  $s$  表示 skin,  $d$  是向量维数。文献[12,16,31]等使用  $\chi^2$  检验对肤色的 2 维高斯分布的假设进行了模型检验。

GMM 是 SGM 的推广,它能够描述更复杂的分布形状,其概率密度表示为

$$p(\mathbf{c} | s) = \sum_{i=1}^k \alpha_i \cdot p_i(\mathbf{c} | s) \quad (2)$$

$p_i(\mathbf{c} | s)$  是与式(1)相同的概率密度函数, $k$  是混合模型的阶数, $\alpha_i$  是混合系数,且满足  $\sum_{i=1}^k \alpha_i = 1$ ,一般可利用 EM 算法来迭代估计模型参数。阶数  $k$  的选取是 GMM 的关键问题, $k$  太小的模型过于概括,而太多则可能引起过拟合,并大大增加训练和分类时间。Yang 证明 2 阶 GMM 能够比较精确地描述肤色分布规律<sup>[31]</sup>,而 Terrillon 则认为  $k = 8$  是真实分布估计的精确性和计算代价的良好折衷<sup>[39]</sup>。Greenspan 提出了一个简单的寻优方法<sup>[48]</sup>,即  $k$  从 1 开始,当  $k$  与  $k + 1$  的对数似然函数幅度变化小于某个阈值时,则停止寻优过程。Raja 假定最优的阶数对应于似然函数的峰值<sup>[24]</sup>,当校验集的对数似然数值出现下降时,则停止学习过程,否则将后验概率最小的成分进行分裂,并重新估计模型参数,该方法还避免了模型参数的初始化过程。假定各子类相对独立,并具有相同的先验概率和协方差矩阵,则 GMM 学习问题得以简化<sup>[21]</sup>,并可采用矢量量化方

法来估计类中心,即以  $\text{tr}[S_w^{-1}S_b]$  为目标函数,在一定范围内来选择最优的  $k$  值。

肤色分布相对其峰值点是非对称的,因此对称的 SGM 显然不够精确。Lee 提出的椭圆边界模型克服了这一缺陷<sup>[49]</sup>,其分类性能超越了 SGM,甚至是部分低阶的 GMM,并具有非常小的计算代价。

### 3.3 半参数模型

以神经网络为代表的半参数模型,其肤色分布隐藏在网络结构和权值中,而且兼具参数模型和非参数模型的优点,即精度高、占据空间小、运行速度较快。基于自组织映射网络(self-organizing map, SOM)的肤色检测器不仅没有“空间偏好性”,而且神经元数目不同(64~196)的整体分类性能几乎没有区别<sup>[32]</sup>,文献[18,19]则使用3层感知机来对肤色进行建模。支持向量机(support vector machine, SVM)是一种新的基于统计学习理论的通用学习方法,由于它可根据结构风险最小化原则尽量提高学习机的泛化能力,因而表现出很多优于现有方法的性能,文献[50]在人脸检测中就使用了一个基于 SVM 的肤色检测器。神经网络本质上是一种矢量量化手段,其每个隐层神经元代表了一个聚类中心,不过,其决策规则隐含在网络结构中,而且神经元数目的确定也需要一定的经验。

### 3.4 基于物理的模型

基于物理的肤色模型是根据皮肤的光学特性、照明条件及成像特性来估计出肤色可能的分布范围,以实现肤色的光照不变性。不同个体的皮肤反射系数曲线在归一化后不仅具有很好的相关性,而且可以使用5阶 GMM 来近似反射曲线<sup>[5]</sup>,或者利用主成分分析法(principal component analysis, PCA)计算其基函数<sup>[51]</sup>。当已知光源和照相机参数时,由于采用反射系数模型能够减少光照变化的影响,从而可获得较好的颜色恒常性。当光源相关色温在从1500K上升到25000K时,不同人种个体的肤色在 rg 平面的分布范围就形成一个紧随黑体辐射普朗克轨迹的小区域,称为“肤色轨迹<sup>[52,53]</sup>”,Storring 进一步考虑了两个光源混合照明条件下的肤色轨迹<sup>[54]</sup>。Soriano 使用该肤色轨迹来约束肤色的搜索范围<sup>[26,27]</sup>,并在不同光照条件下实现了自适应肤色分割,不过,肤色轨迹是与照相机的光学特性紧密相关的。

根据皮肤层状结构及光学属性,基于 Kublenka-Munk 理论估算出的所有正常肤色在3维颜色空间

中处于由两个具有生理意义的轴(一是表皮层中的黑色素含量,二是真皮层中的血液成分)所张成的平面上<sup>[55]</sup>,Cotton 使用59张实际的皮肤图片验证了这一估计。然而,不同部位黑色素含量的差异、皮肤阴影以及皮肤表面的汗渍和油渍引起的高光将会大大降低估计精度。

### 3.5 动态模型

由于视频序列相邻帧之间的皮肤颜色具有一定的连续性,因而可以利用这一事实建立动态模型。动态模型是适应某一特定环境的,其不需具备静态模型的一般性;同时,模型必须能够实时自我更新,以匹配肤色分布变化。由于高斯模型通过调整几个参数就能实现分布变化的更新,因此得到了广泛应用<sup>[12,23,24,29,56]</sup>,而动态直方图<sup>[10,11,26]</sup>则是常见的非参数动态模型。

利用前一帧或几帧图片的模型参数和当前帧估计值的线性加权来作为模型参数更新的递推公式,是常见的参数更新方法。假定检测出的候选肤色仍然服从高斯分布,文献[29]定义了以下分布参数的递推公式:

$$\Theta_{t+1} = (1 - \beta)\Theta_t + \beta\Theta_c \quad (3)$$

$\Theta_t$ 、 $\Theta_{t+1}$  分别是时刻  $t$  和  $t+1$  的肤色分布参数, $\Theta_c$  是当前候选肤色的分布参数, $\beta$  是学习速率( $0 \leq \beta \leq 1$ )。针对固定速率的缺陷,可根据极大似然准则来动态调整预测权值,以实现模型的自适应<sup>[12]</sup>。在参数更新过程中,文献[24,56]采用一个归一化对数似然函数:  $L = \frac{1}{N} \sum \log p(x|O)$  作为突变检测器来进行跟踪,当跟踪器丢失目标时, $L$  值将会突然急剧下降,这时应暂停自适应过程,并通过增加搜索范围来重新开始检测过程,当对象重新以足够高的可能性被跟踪时,则再次启动自适应更新过程。

Soriano 使用前2帧图片的肤色区域与整个图像的比例直方图作为新的模型<sup>[26]</sup>,并使用直方图逆向投影方法来动态定位皮肤区域。假定肤色在 HSV 直方图中的分布变化可用平移、缩放、旋转描述,Sigal 先使用一个显式二阶 Markov 模型来预测直方图的演变<sup>[10]</sup>,然后基于这个预测结果,并结合当前分割结果来动态更新肤色直方图:

$$H_{t+1} = (1 - \beta)H_t + \beta H_p \quad (4)$$

其中,预测部分  $H_p$  使用二阶 Markov 模型估计。该算法虽能够应付比较缓慢的动态场景变换,但结果很大程度上要依赖初始化部分,而对背景的突变则性能有

明显损失。在随后的研究中, Sigal 认为旋转变换并不是一个很重要的因素<sup>[11]</sup>, 可以通过忽略旋转参数来获得更快的计算速度, 而其性能损失可以忽略。

### 3.6 模型性能比较

对于分类来说, 类似问题是: 哪种模型的综合性能最好? 与颜色空间相关吗? 为说明这些问题, 首先将上述 1 894 张图片随机分成训练集和测试集, 其中训练集包含 973 张图片, 约含有  $5.41 \times 10^7$  个肤色点和  $1.53 \times 10^8$  个非肤色点; 然后以 SPM、GMM、SOM 和 SVM 作为考察的肤色模型, 在上述所有空间中训练模型, 并以 ROC (receiver operator characteristics) 和 AUC (area under curve)<sup>[45]</sup> 来衡量模型的综合分类性能; 最后通过测试一系列不同精度的模型参数来分析模型精度对分类性能的影响, 并选择性能最优的参数配置。对于 SPM 和 GMM 分类器, 可采用以下两种不同的决策方式: 一种是仅基于肤色模型的决策 (这里称为 Skinonly 决策), 二是同时使用肤色和非肤色模型的 Bayes 决策。

比较各类模型的综合性能时, 首先根据训练集构造肤色和非肤色类的统计直方图, 并归一化转换成肤色 SPM 和 NSPM (nonskin probability map); 然后基于

归一化直方图训练 1~16 阶肤色和非肤色 GMM, 其中单高斯模型的参数是使用极大似然估计方法直接计算, 2 阶以上的混合模型参数则使用 EM 算法迭代估算 (事先使用  $k$  均值算法初始化), 当相邻两次 EM 迭代结果的似然函数之间的比值差小于  $10^{-7}$  时, 则停止迭代过程, 对于 SPM 和 GMM 分类器, 当待测点的先验概率 (Skinonly 决策) 或后验概率 (Bayes 决策) 大于某个预设阈值时, 就将其划分为肤色, 并通过设定不同的阈值来获取模型的 ROC。对于 SOM, 这里采用六边形拓扑的 2 维结构, 训练时仅使用肤色样本而不考虑非肤色数据, 并以目标点与 SOM 获胜点之间的距离作为判定依据, 由于频率越高的肤色点在 SOM 中分布的范围越大, 与相应的获胜点之间的距离也越小, 可依据这一距离来设定阈值来分类样本点; 最后, 基于 LIBSVM<sup>[57]</sup> 来训练  $\nu$ -SVM<sup>[58]</sup> (采用 RBF 核函数), 并测试分类结果。

表 3 列举了 SPM、SGM、GMM (8 阶)、SOM 和 SVM 在上述所有空间中的综合分类性能。其中, I 代表 Bayes 决策, II 表示 Skinonly 决策。图 1 和图 2 分别是 SPM、SOM 和 SVM 在各颜色空间的 ROC 曲线图 (GMM 的 ROC 曲线图类似于 SPM, 因篇幅限制

表 3 SPM, SGM, GMM, SOM 和 SVM 的 AUC 指标值  
Tab. 3 The AUC of SPM, SGM, GMM, SOM and SVM in all colorspace

颜色空间	不同模型的 AUC 值															
	SPM				SGM				GMM(8 阶)				SOM		SVM	
	3D I	3D II	2D I	2D II	3D I	3D II	2D I	2D II	3D I	3D II	2D I	2D II	3D	2D	3D	2D
RGB	0.873	0.780	0.858	0.809	0.870	0.819	0.840	0.813	0.871	0.712	0.850	0.815	0.829	0.762	0.842	0.840
XYZ	0.874	0.779	0.855	0.810	0.870	0.818	0.837	0.808	0.874	0.793	0.852	0.817	0.820	0.757	0.846	0.836
YIQ	0.875	0.787	0.850	0.711	0.870	0.820	0.854	0.768	0.872	0.792	0.850	0.715	0.829	0.687	0.847	0.830
YCbCr	0.875	0.782	0.849	0.696	0.870	0.820	0.853	0.766	0.874	0.789	0.850	0.708	0.830	0.676	0.843	0.830
YES	0.875	0.780	0.849	0.695	0.870	0.820	0.853	0.766	0.868	0.793	0.850	0.713	0.830	0.679	0.843	0.830
KLT	0.875	0.785	0.870	0.753	0.870	0.820	0.872	0.792	0.872	0.792	0.871	0.752	0.828	0.730	0.842	0.850
L* a* b*	0.875	0.770	0.811	0.706	0.849	0.803	0.803	0.750	0.870	0.763	0.810	0.710	0.816	0.691	0.818	0.754
L* u* v*	0.875	0.769	0.786	0.674	0.853	0.804	0.783	0.717	0.872	0.767	0.785	0.679	0.818	0.674	0.819	0.733
HSI	0.874	0.828	0.817	0.792	0.854	0.843	0.803	0.798	0.867	0.844	0.808	0.804	0.853	0.775	0.710	0.598
HSV	0.874	0.827	0.817	0.800	0.861	0.848	0.809	0.805	0.868	0.849	0.812	0.814	0.854	0.781	0.807	0.636
HSL	0.875	0.828	0.817	0.792	0.857	0.845	0.806	0.799	0.864	0.845	0.811	0.807	0.853	0.777	0.745	0.614
TSL	0.875	0.836	0.817	0.796	0.854	0.852	0.806	0.803	0.867	0.856	0.809	0.807	0.823	0.781	0.769	0.779
nrgb			0.816	0.716			0.788	0.760			0.814	0.720		0.712		0.715
nxyz			0.813	0.702			0.778	0.752			0.811	0.709		0.734		0.708
E-R/G-H	0.873	0.768			0.855	0.818			0.856	0.808			0.751		0.840	
H-CY-Wr	0.870	0.734			0.826	0.822			0.821	0.772			0.764		0.822	
均值	0.874	0.790	0.830	0.747	0.859	0.821	0.820	0.778	0.865	0.798	0.827	0.755	0.821	0.730	0.814	0.754
标准差	0.001	0.029	0.024	0.050	0.013	0.023	0.030	0.028	0.014	0.040	0.025	0.052	0.030	0.043	0.043	0.089

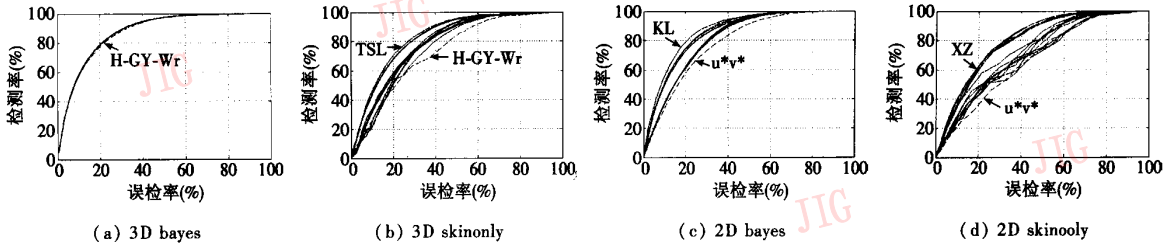


图 1 SPM 在各颜色空间中的 ROC 曲线图

Fig. 1 The ROC curves of SPM in all colorsapces

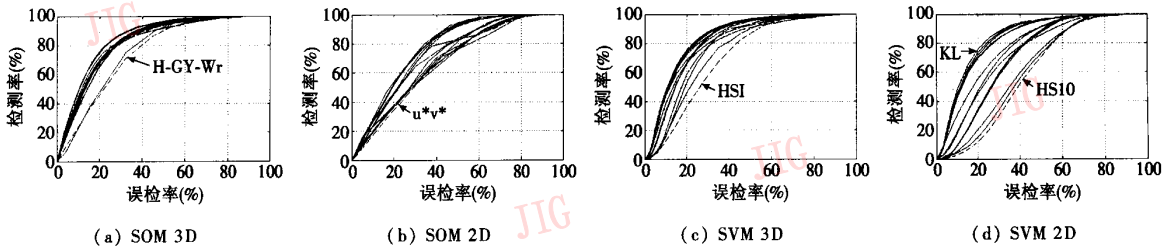


图 2 SOM 和 SVM 在各颜色空间中的 ROC 图

Fig. 2 The ROC curves of SOM and SVM in all colorsapces

予以省略)。从表 3 和图 1、图 2 可以看到,肤色分类器的分类性能与空间的变换关系和决策策略紧密相关。具体来说,各分类器有如下特点:

(1) Bayes 决策下的 3 维 SPM 的分类性能最优,而且不同空间的性能基本重合,只有 H-GY-Wr 的性能略有下降,Bayes 决策下的 3 维 GMM 和 SGM 的分类性能次之,从整体上看,SPM 具有最优的分类性能和较低的计算代价,GMM 的分类性能次之,SOM 最差;

(2) 决策方法是影响 SPM 和 GMM 分类性能的关键因素,Bayes 决策方法优于 Skinonly 决策方法,而且 Bayes 决策下的同一模型在不同颜色空间中的分类性能差别也小于 Skinonly 决策下的性能差别;

(3) 对于所有分类器,当同时采用基于肤色和非肤色模型的决策方法时(SVM 本身包含了肤色和非肤色分布特性),RGB 及线性颜色空间的分类性能最好;而仅使用肤色模型时,知觉颜色空间是分类性能普遍较好的一类空间;

(4) 在相同的决策条件下,3 维空间的分类性能在整体上优于其 2 维色度平面的分类性能,并具有较小的空间性能差别;

(5) 除 Bayes 决策下的 3 维 SPM 外,各模型普遍存在“空间偏好性”,而且该“偏好性”与 RGB 的变换关系紧密相关,并形成以下性能接近的 3 大类:

线性变换空间类、知觉颜色空间类、均匀感知空间类;

(6) 由于组合空间 E-R/G-H 和 H-GY-Wr 并没有表现出最优的性能,因而基于 SPM、GMM 和 SOM 的分类器在这些空间中的性能反而是最差的。

本文还比较了不同的参数精度对模型的性能的影响,各模型的参数名称及取值范围如表 4 所示。在大多数颜色空间中,精度参数在一个相当大的参数范围内对 SPM、SOM 和 SVM 检测器的分类性能影响很小,只有少数颜色空间中的分类性能有所波动。对于 SPM 分类器,当分辨率下降时,只有 Bayes 决策下的组合空间和归一化空间的性能有所下降,而 Skinonly 决策下的组合空间、知觉颜色空间及其色度平面的分类性能反而有所提升。当神经元数目逐步增加到 36 时,只有 HSI、HSV、HSL、 $L^* a^* b^*$ 、

表 4 各肤色模型的精度参数及取值范围

Tab. 4 Values or ranges of precision parameters used to construct skin color models

模型	参数名称	参数范围
SPM	量化级数	256, 128, 64, 32
GMM	混合阶数	1 ~ 16
SOM	神经元数目	20(5 × 4) ~ 132(12 × 11)
SVM	参数 $\nu$	0.2 ~ 0.9

$L^*u^*v^*$  和 KLT 空间中的 SOM 检测性能有明显提升,并趋于平稳;而在其他颜色空间中,不同神经元数目的 SOM 之间的性能则差别很小。GMM 的性能随阶次的变化情况与决策方法有关,基于 Bayes 决策的 GMM 分类器的性能在整体上随阶次的增加而逐渐提升,但在 RGB 空间、线性颜色空间和组合颜色空间以及 IQ、ES、CbCr、KL、 $a^*b^*$  和  $u^*v^*$  平面中,SGM 的分类性能优于低阶 GMM。在 Skinonly 决策条件下,知觉颜色空间及其色度平面中不同阶次的 GMM 性能非常接近,而在其他颜色空间或色度平面中,分类性能随阶次的上升反而有所下降。

## 4 光照不变性与自适应模型

基于大样本的一般分布模型是肤色真实分布的估计,且其推广性受到样本结构的限制,而特定环境中的个体肤色分布仅仅占据非常小的范围,一般模型和特定个体分布之间的不一致性,促使人们寻找更有效的方案。目前,针对肤色检测,主要有光照不变性和自适应两种主要手段,其中,光照不变性用于对光照条件进行补偿,并将其转换到标准光照环境,而自适应方法则兼顾考虑皮肤区域的邻域特征。

### 4.1 光照不变性

皮肤的颜色差异主要是由光照引起的,因此避免光照影响最直接的方法就是寻求一个光照不变的空间。文献[59]提出了一个2维不变性空间: $\xi_1 = \ln(R/G)$ ,  $\xi_2 = \ln(B/G)$ ,在此空间中,光照颜色的改变仅仅使对象颜色分布在2维平面有一个整体平移;Gevers认为在白光照明条件下,nrgb、H和S是光照位置、方向、强度不变的<sup>[20]</sup>。然而,文献[60]的实验结果表明,没有一个颜色空间能够实现光照不变性。

将图像变换到一个相对规范的照明条件,是解决光照不变性问题的另一种常见方法。文献[17]将亮度最大的5%像素视为“参考白点”,并以此按比例调整RGB各通道值,由于该方法没有考虑已饱和通道或者高亮度点属于彩色的可能性,因此仅适用于全局单一光源的照明环境。对应于特征值最大的前3个特征脸(eigenface)保持了脸部和背景的颜色信息,而剩余部分则含有细节信息<sup>[25]</sup>。基于这一点,Soriano用“理想光照”下的最初3个特征脸来代替当前图像的最初3个特征脸来对图像进行重构,从而实现了脸部图像的颜色校正。虽然其在一般照

明环境中取得了很好的颜色恢复效果,但只能用于人脸图像的光照补偿,而对一般肤色区域则没有适用性。

高光和阴影区域的检测和校正是一种局部校正算法。由于含有高光区域的脸部图像在颜色空间中形成一个斜L形或T形,因此,Chen提出了一个基于双色反射模型和序贯PCA的高光自动检测和去除方法<sup>[61]</sup>。该方法假定一组样本的较大的前两个特征值分别为 $\lambda_1$ 和 $\lambda_2$ ( $\lambda_1 > \lambda_2$ ),其在沿亮度增长方向逐步施行PCA过程中,如果特征值比例 $R = \lambda_1/\lambda_2$ 先升后降,且存在峰值,则说明图片中含有高光区域,可通过跟踪最大特征值来估计体反射向量;相反,若沿亮度下降方向实施序贯PCA,则可以估计出表面反射向量。这样将亮度大于设定阈值的样本点沿表面反射向量方向投影到体反射向量轴上,就能达到去除高光的目的。该方法能够最大限度地保留色度差异,并可减少亮度对比,但需要大量的样本数据。

### 4.2 自适应模型

利用肤色在颜色空间和图像空间上分布的紧致性假设来自动调整模型的最优阈值是一种很直观的自适应方法<sup>[15,22]</sup>。假定通用模型能够较好地覆盖所有可能的肤色,则可以针对分割后的候选肤色区域,重新估计模型参数,以实现模型精细化。按照预分类和区域分割顺序不同可分为以下两种方法:一是先使用通用模型粗略地分割出候选肤色区域,再将候选肤色区域细分为肤色和非肤色部分<sup>[3,33]</sup>;二是先将图像按照某种一致性指标分割成几个区域,再确定与肤色最接近的区域<sup>[21,47]</sup>。

假定皮肤区域具有较大的亮度,则候选肤色区域的面积随阈值减小而增大,并且增长的速度逐渐趋缓,但当阈值减小到某个值时,其面积会突然增加。在面积突然增加之前,对应于候选肤色面积增加最少的某个阈值即是最优的阈值<sup>[22]</sup>。类似地,文献[15]根据区域生长终止距离与区域面积的增长速度的关系控制区域生长,从而可精确地定位肤色区域。

文献[33]首先使用HSV空间中的阈值规则来粗略检测肤色,然后利用候选肤色区域的直方图的分布重心位置来调整饱和度(S)和亮度(V)的阈值,以获得一个更合理的分布区域。该方法假定肤色在HSV空间中构成主要的聚簇,以便将阈值调整转化为主聚簇搜索问题。聚类分析使得算法运行速

度比较慢,当类似颜色的背景区域大于肤色区域面积,或者两个聚簇之间的数量差别很大时,该方法将失效,而且不能同时发现不同颜色的皮肤区域也是其缺陷之一。而文献[47]则采用 DFDM (differential feature distribution maps) 算法先将图像粗略地分成颜色均匀的几个区域,然后与通用肤色模型进行比较,以确定可能的皮肤区域,并通过一个最大置信方案调整场景中不同个体的模型参数。Bergasa 介绍了一个快速无监督自适应的 GMM 学习方法<sup>[21]</sup>,该方法假定各类具有相同的先验概率和协方差矩阵,并相对独立,这样就将统计估计问题简化成了基于欧氏距离的聚类问题,其图像中的颜色聚类是将与原型聚类最接近的类被视作皮肤,该方法是一个 Bayes 方法的特殊情形,不仅不存在收敛问题,而且性能也接近一般的 EM 算法。

## 5 结 论

本文全面回顾和系统总结了基于颜色的皮肤检测的方法和技术,并针对颜色空间选择以及肤色建模这两个关键问题,基于一个包含 1894 张图片的样本库,着重比较了 14 个 3 维颜色空间和 14 个 2 维色度平面的肤色分布紧致性和肤色-非肤色的可分辨性,以及基于 SPM、GMM、SOM 和 SVM 的分类器在这些颜色空间中的肤色检测性能。本文所比较的颜色空间包含了所有常用的肤色检测颜色空间,而 SPM、GMM、SOM 和 SVM 则分别是典型的非参数、参数和半参数模型,而且不同精度的参数对肤色-非肤色类的可分辨性和模型分类性能的影响也得到了详尽地分析,是一个比较完善的评估体系。分析结果表明:

(1) 肤色的分类性能完全取决于肤色与非肤色样本的重叠程度,而颜色空间的变换并不能因改变这一基础而改善性能,原始的 RGB 空间及线性颜色空间具有较好的可分辨性和分类性能,是比较理想的一类颜色空间;

(2) 由于个体的肤色差别主要在亮度,因而去除亮度信息会在提高肤色紧致性的同时,大大降低肤色-非肤色的可分辨性,并导致分类性能的恶化;

(3) 量化等级对不同指标的影响并不相同,从 256 下降到 32 是一个比较平稳的过程,而 32 是一个合适的折中点;

(4) 决策策略是影响模型分类性能的重要因

素,同时使用肤色和非肤色模型的分类器的分类性能优于仅基于肤色模型的性能。有趣的是,同时使用肤色和非肤色模型的决策中,RGB 及线性颜色空间的分类性能最优,而在仅采用肤色模型的情况下,则知觉颜色空间的性能最好;

(5) 除 Bayes 决策的 3 维 SPM 外,其他的肤色分类器普遍存在“空间偏好性”,而且不同模型的“空间偏好性”非常类似,且与颜色空间变换规律紧密相关,从而形成性能接近的线性空间、感知均匀空间和知觉颜色空间 3 类。

基于 Bayes 决策的 RGB 及 3 维线性颜色空间 SPM 具有性能好、速度快的明显优势,是比较理想的肤色检测方案。

颜色是皮肤区别于其他物体的重要依据,然而,由于肤色受到光照条件、照相机特性等的影响,肤色与非肤色具有内在的重叠性,因此基于颜色的皮肤检测仍然是一项复杂而艰巨的任务。可见,仅仅依靠颜色来检测和识别皮肤是不够的,皮肤检测研究的趋势是:①利用多种线索<sup>[62,63]</sup>,结合邻域的其他特征(如位置、轮廓、纹理等)进行自适应处理;②与人脸、手部等检测手段及高层特征结合,以形成一个多种信息相互印证,互为目的和手段的启发式综合方法。

## 参考文献 (References)

- 1 Arentz W A, Olstad B. Classifying offensive sites based on image content [J]. *Computer Vision and Image Understanding*, 2004, **94**(1-3): 295 ~ 310.
- 2 Forsyth D A, Fleck M M. Automatic detection of human nudes [J]. *International Journal of Computer Vision*, 1999, **32**(1): 63 ~ 77.
- 3 Zheng H C, Daoudi M, Jedynak B. Blocking adult images based on statistical skin detection [J]. *Electronic Letters on Computer Vision and Image Analysis*, 2004, **4**(2): 1 ~ 14.
- 4 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.
- 5 Angelopoulou E. Understanding the color of human skin [A]. In: *Proceedings of the SPIE Conference on Human Vision and Electronic Imaging* [C], San Jose, CA, USA, 2001, **4299**: 243 ~ 251.
- 6 Jablonski N G, Chaplin G. The evolution of human skin coloration [J]. *Journal of Human Evolution*, 2000, **39**(1): 57 ~ 106.
- 7 Gomez G, Sanchez M, Sucar L E. On selecting an appropriate colour space for skin detection [J]. *Lecture Notes in Computer Science*, 2002, **2313**: 69 ~ 78.
- 8 Gomez G. On selecting colour components for skin detection [A]. In: *Proceedings of International Conference on Pattern Recognition*

- [C], Quebec, Canada, 2002, 2: 961~964.
- 9 Lee C H, Kim J S, Park K H. Automatic face location in a complex background using motion and color information [J]. *Pattern Recognition*, 1996, 29(11): 1877~1889.
- 10 Sigal L, Sclaroff S, Athitsos V. Estimation and prediction of evolving color distributions for skin segmentation under varying illumination [A]. In: *IEEE Conference on Computer Vision and Pattern Recognition* [C], Hilton Head, SC, USA, 2000: 152~159.
- 11 Sigal L, Sclaroff S, Athitsos V. Skin color-based video segmentation under time-varying illumination [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(7): 862~877.
- 12 Yang J, Weier L, Waibel A. Skin-color modeling and adaptation [R]. Technical Report, CMU-CS-97-146, School of Computer Science, Carnegie Mellon University, CA, USA, 1997.
- 13 Chai D, Ngan K N. Face segmentation using skin-color map in videophone applications [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, 9(4): 551~564.
- 14 Garcia C, Tziritas G. Face detection using quantized skin color regions merging and wavelet packet analysis [J]. *IEEE Transactions on Multimedia*, 1999, 1(3): 264~277.
- 15 Gu L F, Bone D. Skin colour region detection in MPEG video sequences [A]. In: *Proceedings of International Conference on Image Analysis and Processing* [C], Venice, Italy, 1999: 898~903.
- 16 Habibi N. Automatic segmentation of the face and hands in sign language video sequences [R]. In: Department of Electrical and Electronic Engineering, University Adelaide, Adelaide, Australia [EB/OL]. <http://www.eleceng.adelaide.edu.au/personal/nha, 2001>.
- 17 Hsu R L, Mohamed A M, Jain A K. Face detection in color images [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 696~706.
- 18 Ishii H, Fukumi M, Akamatsu N. Face detection based on skin color information in visual scenes by neural networks [A]. In: *IEEE International Conference on Systems, Man, and Cybernetics* [C], Nashville, Tennessee, USA, 1999, 5: 350~355.
- 19 Phung S L, Chai D, Bouzerdoum A. A universal and robust human skin color model using neural networks [A]. In: *Proceedings of International Joint Conference on Neural Networks* [C], Washington DC, USA, 2001, 4: 2844~2849.
- 20 Gevers T, Smeulders A W M. A comparative study of several color models for color image invariant retrieval [A]. In: *Proceedings of the First International Workshop on Image Database & Multimedia Search* [C], Amsterdam, Netherlands, 1996: 17~26.
- 21 Bergasa L M, Mazo M, Gardel A, *et al.* Unsupervised and adaptive Gaussian skin-color model [J]. *Image and Vision Computing*, 2000, 18(12): 987~1003.
- 22 Chang H, Robles U. Face detection [EB/OL]. <http://www-cs-students.stanford.edu/~robles/ee368/main.html>
- 23 Liu Ming-bao, Yao Hong-xun, Gao Wen. Real-time human face tracking in color images [J]. *Journal of Computer*, 1998, 21(6): 527~532. [刘明宝, 姚鸿勋, 高文. 彩色图像的实时人脸跟踪方法 [J]. *计算机学报*, 1998, 21(6): 527~532.]
- 24 Raja Y, McKenna S J, Gong S G. Colour model selection and adaptation in dynamic scenes [A]. In: *Proceedings of European Conference on Computer Vision* [C], Freiburg, Germany, 1998, 1: 460~474.
- 25 Soriano M, Marszalec E, Martinkauppi B, *et al.* Making saturated facial images useful again [A]. In: *Proceedings EUROPTO Conference on Polarization and Color Techniques in Industrial Inspection* [C], Munich, Germany: SPIE Press, 1999, 3826: 113~121.
- 26 Soriano M, Martinkauppi B, Huovinen S, *et al.* Skin color modeling under varying illumination conditions using the skin locus for selecting training pixels [A]. In: *Proceedings of Workshop on Real-Time Image Sequence Analysis* [C], Oulu, Finland, 2000: 43~49.
- 27 Soriano M, Martinkauppi B, Huovinen S, *et al.* Adaptive skin color modeling using the skin locus for selecting training pixels [J]. *Pattern Recognition*, 2003, 36(3): 681~690.
- 28 Cai J, Goshtasby A. Detecting human faces in color images [J]. *Image and Vision Computing*, 1999, 18(1): 63~75.
- 29 Fan Yun, Wang Run-sheng. Detect human faces in video [J]. *Journal of Computer Aided Design & Computer Graphics*, 2002, 14(5): 1~7. [樊昀, 王润生. 从视频中检测人脸 [J]. *计算机辅助设计与图形学学报*, 2002, 14(5): 1~7.]
- 30 Yang M H, Ahuja N. Detecting human faces in color images [A]. In: *IEEE International Conference on Image Processing* [C], Chicago, IL, USA, 1998: 127~130.
- 31 Yang M H, Ahuja N. Gaussian mixture model for human skin color and its applications in image and video databases [A]. In: *Proceedings of SPIE Conference on Storage and Retrieval for Image and Video Databases VII* [C], San Jose, California, USA, 1999, 3656: 458~466.
- 32 Brown D, Craw I, Lewthwaite J. A SOM based approach to skin detection with application in real time systems [A]. In: *Proceedings of British Machine Vision Conference* [C], Manchester, UK, 2001: 491~500.
- 33 Cho K M, Jang J H, Hong K S. Adaptive Skin-color filter [J]. *Pattern Recognition*, 2001, 34(5): 1067~1073.
- 34 Herodotou N, Plataniotis K N, Venetsanopoulos A N. Automatic location and tracking of the facial region in color video sequences [J]. *Signal Processing: Image Communication*, 1999, 14(5): 359~388.
- 35 Hsieh I S, Fan K C, Lin C. A statistic approach to the detection of human faces in color nature scene [J]. *Pattern Recognition*, 2002, 35(7): 1583~1596.
- 36 Tao Lin-mi, Peng Zhen-yun, Xu Guang-you. The features of skin color [J]. *Journal of software*, 2001, 12(7): 1032~1041. [陶霖密, 彭振云, 徐光祐. 人体的肤色特征 [J]. *软件学报*, 2001, 12(7): 1032~1041.]
- 37 Yoo T W, Oh I S. A fast algorithm for tracking human faces based on chromatic histograms [J]. *Pattern Recognition Letters*, 1999, 20(10): 967~978.
- 38 Terrillon J C, David M, Akamatsu S. Detection of human faces in

- complex scene images by use of a skin color model and of invariant Fourier-Mellin moments [ A ]. In: Proceedings of International Conference on Pattern Recognition [ C ], Brisbane, Australia, 1998: 1350 ~ 1355.
- 39 Terrillon J C, Shirazi M N, Fukamachi H, *et al.* Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images [ A ]. In: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition [ C ], Grenoble, France, 2000: 54 ~ 61.
- 40 Wark T, Sridharan, S. A syntactic approach to automatic lip feature extraction for speaker identification [ A ]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [ C ], Seattle, WA, USA, 1998: 3693 ~ 3696.
- 41 Wu H Y, Chen Q, Yachida M. Face detection from color images using a fuzzy pattern matching method [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21 ( 6 ): 557 ~ 563.
- 42 Brand J, Mason J S. A comparative assessment of three approaches to pixel-level human skin-detection [ A ]. In: Proceedings of International Conference on Pattern Recognition [ C ], Barcelona, Spain, 2000: 1056 ~ 1059.
- 43 Zarit B D, Super B J, Quek F K H. Comparison of five color models in skin pixel classification [ A ]. In: Proceedings of International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems [ C ], Corfu, Greece, 1999: 58 ~ 63.
- 44 Shin M C, Chang K I, Tsap L V. Does colorspace transformation make any difference on skin detection [ A ]. In: Proceedings of IEEE Workshop on Applications of Computer Vision [ C ], Orlando, FL, USA, 2002: 275 ~ 279.
- 45 Jayaram S, Schmugge S, Shin M C, *et al.* Effect of color space transformation, the illuminance component, and color modeling on skin detection [ A ]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [ C ], Washington DC, USA, 2004, 2: 813 ~ 818.
- 46 Albiol A, Torres L, Delp E J. Optimum color spaces for skin detection [ A ]. In: Proceedings of the International Conference on Image Processing [ C ], Thessaloniki, Greece, 2001: 122 ~ 124.
- 47 Sahbi H, Boujemaa N. Accurate face detection based on coarse segmentation and fine skin color adaption [ A ]. In: International Conference on Image and Signal [ C ], Agadir, Morocco, 2001.
- 48 Greenspan H, Goldberger J, Eshet I. Mixture model for face-color modeling and segmentation [ J ]. Pattern Recognition Letters, 2001, 22(14): 1525 ~ 1536.
- 49 Lee J Y, Yoo S I. An elliptical boundary model for skin color detection [ A ]. In: International Conference on Imaging Science, Systems, and Technology [ C ], Las Vegas, NA, USA, 2002: 579 ~ 584.
- 50 Zhu Q, Cheng K T, Wu C T, *et al.* Adaptive learning of an accurate skin-color model [ A ]. In: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition [ C ], Seoul, Korea, 2004: 37 ~ 42.
- 51 Martinkauppi J B, Soriano M. Basis functions of the color signal of skin under different illuminants [ A ]. In: Proceedings of the 3rd International Conference on Multispectral Color Science [ C ], Joensuu, Finland, 2001: 21 ~ 24.
- 52 Störing M, Andersen H J, Granum E. Skin colour detection under changing lighting conditions [ A ]. In: Proceedings of the 7th Symposium on Intelligent Robotics Systems [ C ], Coimbra, Portugal, 1999: 187 ~ 195.
- 53 Störing M, Kočka T, Andersen H J, *et al.* Tracking regions of human skin through illumination changes [ J ]. Pattern Recognition Letters, 2003, 24(11): 1715 ~ 1723.
- 54 Störing M, Andersen H J, Granum E. Physics-based modelling of human skin colour under mixed illuminants [ J ]. Robotics and Autonomous Systems, 2001, 35(3-4): 131 ~ 142.
- 55 Cotton S D, Claridge E. Developing a predictive model of human skin colouring [ A ]. In: Proceedings of SPIE Conference on Medical Imaging [ C ], Newport Beach, California, USA, 1996, 2708: 814 ~ 825.
- 56 McKenna S J, Raja Y, Gong S G. Tracking colour objects using adaptive mixture models [ J ]. Image and Vision Computing, 1999, 17(3-4): 225 ~ 231.
- 57 Chang C C, Lin C J. LIBSVM: A Library for Support Vector Machines [ CP/OL ]. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- 58 Schölkopf B, Smola A J, Williamson R, *et al.* New support vector algorithms [ J ]. Neural Computation, 2000, 12(5): 1207 ~ 1245.
- 59 Berwick D, Lee S W. A Chromaticity space for specularly illumination color and illumination pose-invariant 3-D object recognition [ A ]. In: Proceedings of International Conference on Computer Vision [ C ], Bombay, India, 1998: 165 ~ 170.
- 60 Martinkauppi J B, Soriano M, Laksonen M H. Behavior of skin color under varying illumination seen by different cameras at different color spaces [ A ]. In: Proceedings of SPIE Conference on Machine Vision Applications in Industrial Inspection [ C ], San Jose, California, USA, 2001, 4301: 102 ~ 113.
- 61 Chen D S, Liu Z K. A method for automatic detection and correction of highlighted area in color face image [ J ]. Journal of Software, 2003, 14(11): 1900 ~ 1906.
- 62 Butler D, McCool C, McKay M, *et al.* Robust face localisation using motion, colour & fusion [ A ]. In: Proceedings of Digital Image Computing-Techniques and Applications [ C ], Sydney, Australia, 2003: 899 ~ 908.
- 63 Saber E, Tekalp A M. Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions [ J ]. Pattern Recognition Letters, 1998, 19(8): 669 ~ 680.