

## 自适应动态肤色检测技术研究进展

张坤, 王烨东, 李文元, 李昌禄, 雷志春

引用本文:

张坤, 王烨东, 李文元, 等. 自适应动态肤色检测技术研究进展[J]. 液晶与显示, 2021, 36(2): 293–304.

ZHANG Kun, WANG Ye-dong, LI Wen-yuan, et al. Research progress of adaptive dynamic skin color detection technology[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(2): 293–304.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2020-0196>

## 您可能感兴趣的其他文章

Articles you may be interested in

### 一种肤色定位的人脸检测算法

Face detection algorithm based on skin color location

液晶与显示. 2019, 34(1): 70–73 <https://doi.org/10.3788/YJYXS20193401.0070>

### 基于神经网络集成模型的宫颈细胞病理计算机辅助诊断方法

Computer-aided diagnosis of cervical cytopathology based on neural network ensemble model

液晶与显示. 2018, 33(4): 347–355,356 <https://doi.org/10.3788/YJYXS20183304.0347>

### 非凸性优化与动态自适应滤波的湍流退化视频复原

Turbulence-degraded video restoration based on non-convexity optimization and dynamical adaptive filtering

液晶与显示. 2017, 32(6): 482–490 <https://doi.org/10.3788/YJYXS20173206.0482>

### 基于注意力掩模融合的目标检测算法

Object detection algorithm based on attention mask fusion

液晶与显示. 2019, 34(8): 825–833 <https://doi.org/10.3788/YJYXS20193408.0825>

### 基于自适应局部增强的手机TFT-LCD屏Mura缺陷自动检测

Automatic detection of Mura defect in TFT-LCD mobile screen based on adaptive local enhancement

液晶与显示. 2018, 33(6): 475–482 <https://doi.org/10.3788/YJYXS20183306.0475>

文章编号:1007-2780(2021)02-0293-12

## 自适应动态肤色检测技术研究进展

张 坤<sup>1</sup>, 王烨东<sup>2</sup>, 李文元<sup>1</sup>, 李昌禄<sup>1</sup>, 雷志春<sup>1\*</sup>

(1. 天津大学 微电子学院, 天津 300072;

2. 海信视像科技股份有限公司, 山东 青岛 266071)

**摘要:** 肤色检测在人脸检测、手势识别、敏感图像过滤、医疗诊断、图像增强等计算机视觉和图像处理领域具有广泛的应用。由于人体肤色容易受光照条件、种族等因素的影响, 难以在非限制性场景中达到理想的肤色检测效果, 所以自适应肤色检测方法应运而生。本文较全面地对自适应肤色检测方法进行了综述, 重点分析了基于参数动态调整和基于高层语义特征的自适应肤色检测方法, 讨论了它们各自的优势和不足, 并在此基础上展望了未来的发展趋势, 为相关研究人员提供参考。

**关键词:** 肤色检测; 自适应肤色检测算法; 高斯模型; 图像分割; 计算机视觉

**中图分类号:** TP391.41; TP18 **文献标识码:** A **doi:** 10.37188/CJLCD.2020-0196

## Research progress of adaptive dynamic skin color detection technology

ZHANG Kun<sup>1</sup>, WANG Ye-dong<sup>2</sup>, LI Wen-yuan<sup>1</sup>, LI Chang-lu<sup>1</sup>, LEI Zhi-chun<sup>1\*</sup>

(1. School of Microelectronic, Tianjin University, Tianjin 300072, China;

2. Hisense Visual Technology Co., Ltd., Qingdao 266071, China)

**Abstract:** Skin color detection is widely used in face detection, gesture recognition, sensitive image filtering, medical diagnosis, image enhancement and so on. It is an important technology for computer vision and image processing fields. Because human skin color is easily affected by lighting conditions, races, etc., it is difficult to achieve ideal detection results under unconstrained environments. The emergence of adaptive methods makes the above problems improved. In this paper, adaptive skin detection methods are comprehensively reviewed. The paper focuses on the analysis of adaptive skin detection methods based on parameter dynamic adjustment and high-level semantic features. Meanwhile, it discusses their respective advantages and disadvantages. On this basis, the future development trend is prospected, which aims at providing reference for related researchers.

**Key words:** skin detection; adaptive skin detection algorithm; Gaussian model; image segmentation; computer vision

收稿日期: 2020-07-24; 修订日期: 2020-09-18.

基金项目: 天津市科技支撑项目(No.16YFZCGX00760)

Supported by Science and Technology Support Program of Tianjin(No.16YFZCGX00760)

\* 通信联系人, E-mail: zclei@tju.edu.cn

## 1 引言

肤色检测实质上是在数字图像中划分皮肤和非皮肤区域的过程。随着信息社会的发展,图像和视频已经成为人们广泛使用的信息载体,对于肤色检测和处理的需求也日益增多。在各种计算机视觉任务中,肤色检测已经被应用于人脸检测<sup>[1]</sup>、人体追踪<sup>[2]</sup>、表情识别<sup>[3]</sup>、手势识别<sup>[4]</sup>、疲劳驾驶<sup>[5]</sup>、智能监控<sup>[6]</sup>、中医皮肤病诊断<sup>[7]</sup>以及敏感图像过滤<sup>[8-9]</sup>等方面。近年来,“美图秀秀”、“天天 P 图”等人像自动美化软件受到了广大女性用户的喜爱,肤色检测在网络社交媒体中展现出很高的应用价值。2019 年度中国网络视听发展研究报告中<sup>[10]</sup>指出,截止 2018 年底,我国网络视频用户总规模已达 7.25 亿,短视频、网络直播等领域正表现出迅猛发展之势,人们更加喜欢使用皮肤美化功能来使自己的皮肤表现得白皙、有吸引力。同时,“智能电视”也迎来了发展革新的浪潮,肤色呈现作为衡量画质的重要指标自然受到广泛的关注。因此,肤色检测作为肤色美化<sup>[11]</sup>、肤色增强<sup>[12]</sup>的关键步骤,也变得越来越重要。

近 20 年来,从最初的阈值分割法到现在的深度学习语义分割技术,有大量肤色检测相关的方法和理论被提出。文献[13]中根据是否涉及物理成像过程将肤色检测方法分为基于物理的方法和基于统计的方法,但由于需要使用光谱摄影仪等成像设备,基于物理的方法应用范围较小。根据是否需要获取像素的空间分布信息,可以将肤色检测方法分为基于区域的检测方法和基于像素的检测方法两大类型,综述文献[14]中重点对基于像素的各种肤色检测方法及其所涉及的相关技术进行了总结。

上述文献中所介绍的内容主要是静态的肤色检测方法,而这些方法在非限制性场景中的实用性较低。值得注意的是,统计法和机器学习法依赖于一定规模的训练数据并受到数据集质量的影响。虽然深度学习技术可以通过在特征提取网络中增加隐藏层来提高精确度,但与此同时,也将会增加训练时间和计算成本。自适应技术可以在提高准确率的同时,减少或避免对数据集的依赖,维持较低的计算成本,具有极强的实用价值。本文对自适应肤色检测方法

近 20 年来的研究成果进行了归纳总结,主要从基于参数动态调整和基于高层语义特征的自适应肤色检测方法入手,介绍了各类代表性方法,并进行了分析和讨论。

## 2 自适应肤色检测技术研究背景

为了便于理解肤色检测技术的研究背景,本部分简要介绍了传统的静态肤色检测技术以及近些年应用较广的深度学习方法,并分析了这些技术在实际场景中所面临的难题,介绍了几种常用的性能评价指标。

### 2.1 静态肤色检测技术

静态肤色检测方法主要通过使用某种固定规则来实现肤色检测,主要涉及色彩空间的选择、肤色模型的建立两大步骤。本文总结了常见的静态肤色检测方法,详细分类如图 1 所示。

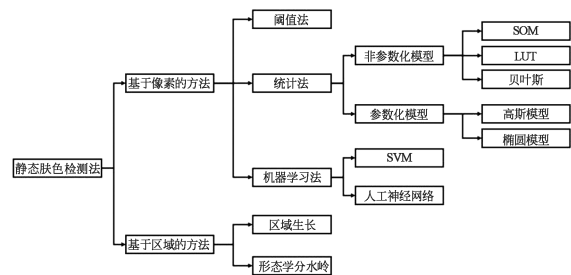


Fig.1 Detection methods of static skin

色彩空间是一种表示颜色的抽象数学模型,常用的色彩空间有 RGB、YCbCr、HSV 等,肤色在不同色彩空间中的分布特征存在差异,如图 2 所示。根据色彩模型构建的方式不同,静态肤色检测方法可以细分为阈值法、统计法、机器学习法、区域生长法等。

阈值法是一种实现简单、计算量小、速度快的分割方法<sup>[15]</sup>,阈值法的性能表现与所设定的阈值密切相关,适用于目标和背景存在明显灰度级差异的场景,而在复杂环境中无法准确地描述肤色的分布。统计法根据肤色样本在色彩空间中的分布构建肤色概率图 (Skin Probability Map, SPM),即在离散化的色彩空间中为每个点分配概率值,然后利用建立起来的 SPM 判断待检测颜色是否属于肤色。非参数化模型通常使用直方图估计样本的 SPM,主要方法包括自组织映射

(Self-Organizing-Map, SOM)、查找表(Look-Up-Table, LUT)、贝叶斯分类器等方法。但是建立可靠的非参数化模型需要较大规模的肤色样本点和较高的存储空间。参数化方法则根据肤色聚类特性,直接使用高斯模型或椭圆模型等已有模型进行拟合即可。由于参数化模型可以在较少训练样本中得到可靠的 SPM,因此常被用于自适应方

法<sup>[16-17]</sup>。机器学习方法与统计法类似,即通过学习训练样本中的规律来实现肤色检测,但与之不同的是机器学习技术不需要建立明确的肤色模型,而是将肤色检测视为一个二分类的过程。其中常用的方法有支持向量机(Support Vector Machine, SVM)和人工神经网络(Artificial Neural Network, ANN)等。

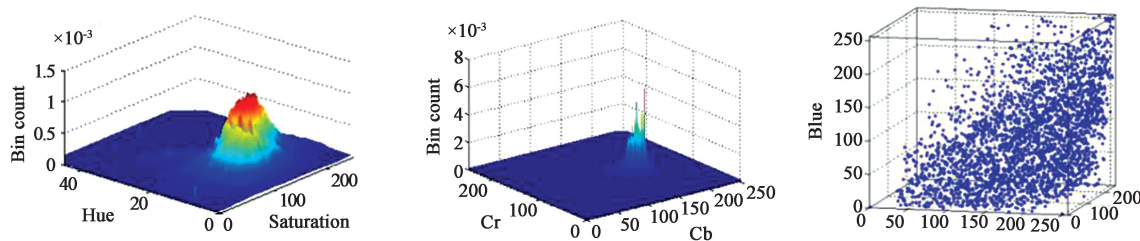


图2 肤色分别在(a) HSV, (b) YCbCr, (c) RGB色彩空间中的分布特征<sup>[15]</sup>。

Fig.2 Skin color distribution feature in (a) HSV, (b) YCbCr and (c) RGB<sup>[15]</sup>, respectively.

上述肤色检测方法主要是通过每个像素点的色彩信息判断其是否属于肤色点,没有考虑像素与像素之间关系,而基于区域的方法则主要通过像素之间的空间关系来进行判断。区域生长法从种子点开始,不断向周围进行扩散,将具有与种子相似属性的邻近像素附加到每个种子上,并继续向周围扩散,直到遇到边缘或者临近像素均不与种子点相似为止。区域生长法可以使用图像的纹理特征,对去除背景有一定帮助,但是区域生长法十分依赖种子点的选取,并且在实际图像尺寸中需要较大的计算量<sup>[15]</sup>。

## 2.2 基于深度学习的肤色检测技术

值得注意的是,近些年来机器学习得到了空前的发展,神经网络的层数越来越多,学习到的特征也越来越丰富,深度学习逐渐发展成为解决图像问题的一个重要的分支,在肤色检测领域也得到了越来越多的关注。

Lumini 等学者<sup>[18]</sup>发现使用 SegNet<sup>[19]</sup>、Unet<sup>[20]</sup>、DeepLab<sup>[21]</sup>等深度学习语义分割网络检测肤色能够获得明显优于传统方法的结果。一方面,借助深层的特征提取网络,深度学习网络可以有效地提取肤色区域的颜色、纹理等特征,有助于减轻复杂背景的干扰,提升检测精确度,因此深度学习方法通常能够得到比传统方法更加出色的检测效果。但另一方面,深度学习方法需要使用大量标注的皮肤数据进行训练,而像素级的准确标注仍

是一项艰巨的任务。尽管 Dourado 等学者<sup>[22]</sup>以及 He 等学者<sup>[23]</sup>已经尝试通过迁移学习和多任务弱监督技术来减弱网络对大量标注肤色数据集的依赖,但是一个性能优异的肤色检测网络通常需要结构复杂的特征提取网络以及较长的训练时间。在深度学习中,检测精度和速度之间的矛盾也在一定程度上制约了深度学习肤色检测在实际中的应用。例如 DeepLab 通用分割网络的深度较大,能得到出色的精度,在本文第3章的实验对比中有所体现。但对于一张  $306 \times 306$  的输入图像,即使在 TITAN X GPU 中也仅能获得 8 FPS (Frames Per Second) 的速度,与实时性要求差距较大。因此在一些实时性要求高、运行设备性能低的情况下,深度学习方法可能不是最佳的选择。

## 2.3 肤色检测方法所面临的挑战

静态和深度学习肤色检测技术主要受到以下因素的影响:

一方面,在非限制性场景中,光照分布和强度的改变都会导致肤色出现变化,是影响肤色检测性能的最重要因素之一;其次,种族、性别等个体差异,化妆等皮肤涂覆手段、相机传感器的性能和参数、图像的尺寸、清晰度、噪声以及调色处理过程等都会对肤色造成影响;此外,去除背景中与肤色相近的颜色干扰也是一个比较困难的问题。虽然深度学习方法可以在一定程度上改善上述问题,但是网络结构复杂度高、训练成本大、不易平

衡精度与速度的特点限制了其实际应用范围。另一方面,数据集是获取先验规则或训练模型的基础,但是标注数据工作量大,目前公开的肤色数据集较少、质量不高,不同种族覆盖程度不均衡等因素也限制着肤色检测技术的发展。

#### 2.4 肤色检测方法使用的评价指标

肤色检测的评价指标主要分为定性和定量两种。其中定性方法将肤色检测结果可视化二值图像,通过主观判断来确定肤色检测的好坏。肤色检测任务实质上是对每个像素进行的二分类任务,因此研究中常用的定量分析指标有:

$$FNR = \frac{FN}{TP + FN}, \quad (1)$$

$$FPR = \frac{FP}{TN + FP}, \quad (2)$$

$$Precision = \frac{TP}{TP + FP}, \quad (3)$$

$$Recall = \frac{TP}{TP + FN}, \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}, \quad (5)$$

$$F\text{-measure} = \frac{2Recall \cdot Precision}{Recall + Precision}, \quad (6)$$

式中,  $TP$  (True positive)、 $TN$  (True negative) 分别代表正确分类为肤色和背景的样本数,  $FP$  (False positive)、 $FN$  (False negative) 分别代表错误分类为肤色和背景的样本数,在肤色检测中总样本数是图像中的总像素数。式(1)~(6)中,  $FNR$ 、 $FPR$ 、 $Precision$ 、 $Recall$ 、 $Accuracy$  以及  $F\text{-measure}$  分别代表漏检率、误检率、精确率、检出率、准确率和  $F$  值,其中  $F$  值常被用作肤色检测器的综合性能指标。除此之外,还有 ROC 曲线<sup>[24]</sup>、混淆矩阵行列式<sup>[25]</sup>等定量分析指标。

### 3 自适应肤色检测方法

在非限制性场景中,肤色容易受到光照变化的影响,从而出现不同程度的亮度和色彩偏移。人眼具有颜色恒常性的特点<sup>[26]</sup>,即人可以在图像整体存在色调偏移时正确识别出图像中物体原有的颜色,而抛开了所受光照的影响。然而,目前人们并未研究透彻颜色恒常性的机制,无法得到完美覆盖不同环境的静态方法,同时在很多场景中对实时性有较高的要求,因此研究学者提出了各

种自适应肤色检测技术来加以解决。在本文中,自适应肤色检测方法主要分为两种:基于参数动态调整和基于高层语义特征的自适应肤色检测方法。

#### 3.1 基于参数动态调整的自适应方法

##### 3.1.1 自适应阈值法

虽然使用特定的阈值方法具有实现简单、易于调整、计算成本低的特点,但是固定阈值无法适应光照变化的环境,针对这种情况,学者们提出使用自适应阈值法来提高检测准确性,光照变化通常会给皮肤带来亮度和颜色上的变化。许多肤色检测算法选择舍弃亮度分量,而 Jayaram 等学者<sup>[27]</sup>在 850 张图像中对比了使用亮度分量和不使用亮度分量的方法,发现亮度分量可以提供不同层次的皮肤颜色信息,因此设定一组随亮度信息变化的规则可以达到自适应肤色检测的目的。Gracia 等学者<sup>[28]</sup>在 YCbCr 和 HSV 色彩空间中分别设定了一组受亮度控制的肤色检测规则,有效提高了不均匀光照环境中的检测准确率。由于该方法在两个色彩空间中使用的是同一种实现方案,因此本文只介绍 YCbCr 色彩空间中的自适应阈值规则,即如果像素满足以下条件,则将其判定为肤色像素:

$$Y \leq 128 \begin{cases} \theta_1 = 6 \\ \theta_2 = 12 \\ \theta_3 = 2 + \frac{Y}{32} \\ \theta_4 = -16 + \frac{Y}{16} \end{cases}, \quad (7)$$

$$Y > 128 \begin{cases} \theta_1 = -2 + \frac{256 - Y}{16} \\ \theta_2 = 20 - \frac{256 - Y}{16} \\ \theta_3 = 6 \\ \theta_4 = 8 \end{cases}, \quad (8)$$

$$\begin{cases} Cr \geq -2(Cb + 24); Cr \geq -(Cb + 17) \\ Cr \geq -4(Cb + 32); Cr \geq 2.5(Cb + \theta_1) \\ Cr \geq \theta_3; Cr \geq 0.5(\theta_4 - Cb) \\ Cr \leq \frac{220 - Cb}{6}; Cr \leq \frac{4}{3}(\theta_2 - Cb) \end{cases}. \quad (9)$$

Cho 等学者<sup>[29]</sup>在 HSV 色彩空间中初始化了

一组阈值  $H$  (0.4, 0.7),  $S$  (0.15, 0.75),  $V$  (0.35, 0.95), 然后根据  $S$ - $V$  平面的直方图信息进行阈值更新。该方法对颜色复杂的场景表现不佳,只适用于背景简单的限制性场景,使用类似方法更新阈值的还有文献[30-31]。

郭耸等学者<sup>[32]</sup>提出了一种基于肤色相似度的自适应阈值方法,通过计算类间方差、类内离散相似度的方式确定分割肤色的最佳阈值,减轻了Adaboost人脸检测算法的计算负担,但是对人脸检测率提升较小。

张明吉等学者<sup>[33]</sup>利用4种线索辅助建立最优阈值。该方法认为最优阈值通常位于肤色概率分布图的某个谷底,且一般靠近肤色峰左侧。基于此规则,作者使用机器学习技术寻找最满足条件的谷底位置作为最优阈值。结果显示,该方法可以获得88.6%的检出率,12.1%的误检率,其中检出率与Cho等学者<sup>[29]</sup>的结果相比较提高了11.76%。Phung等学者<sup>[34]</sup>提出了一种利用人体皮肤纹理特征来选择合适肤色阈值的肤色检测方法,结合非颜色特征能够有效提高肤色模型的性能,该方法得到了4.5%的误检率和4.0%的漏检率。

自适应阈值法虽然能够在一定程度上提升肤色检测效果,但是在复杂背景中不容易得到可靠的阈值范围,同时在某些情况下阈值法难以精确描述肤色的分布情况。

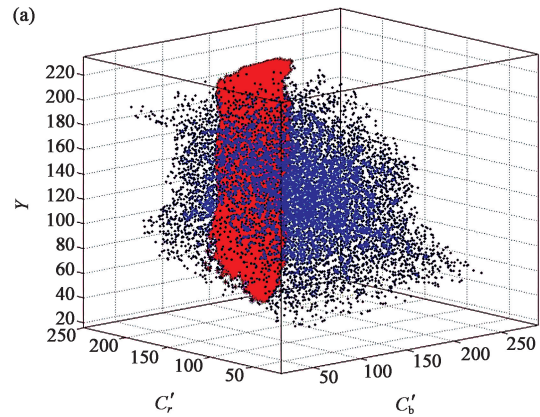
### 3.1.2 自适应椭圆模型

肤色在YCbCr等色彩空间的色度平面中通常呈现椭圆分布,但是在光照变化、不同人种等情况下,固定的椭圆模型不一定能准确描述实际肤色。

李晓光等学者<sup>[35]</sup>提出了一种在DCT压缩域中使用的动态椭圆模型。作者利用图像亮度对色度聚类造成的非线性影响,根据亮度变化采用不同参数的椭圆模型表示肤色,形成了一种三维的椭圆模型,可以适应于不同的光照强度。

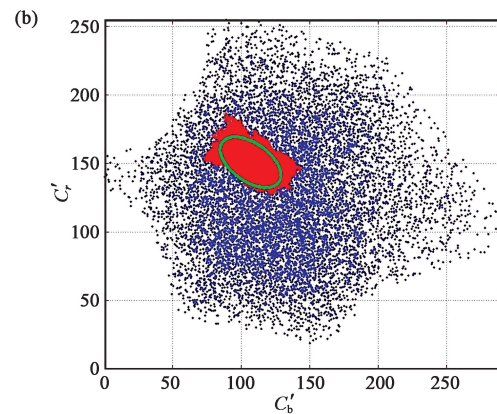
Hsu等学者<sup>[17]</sup>提出了一种使用光照补偿的自适应椭圆模型,如图3所示。首先判断图像中参考白(前50%亮度)的像素个数,如果大于某个阈值,则对RGB图像各分量进行线性调整。同时作者为了消除肤色和亮度分量之间的非线性关联,提出了YCbCr色彩空间,并在CbCr平面建立椭圆模型用于肤色检测。该方法对一般光照强度

变化适应较好,但是在复杂的光照环境中,光照补偿可能使图像更糟,同时也未能很好地解决偏色问题。



(a) YCbCr 色彩空间

(a) Color space of YCbCr



(b) CbCr 子空间投影

(b) Projection of CbCr subspace

图3 YCbCr色彩空间中的肤色分布<sup>[17]</sup>(红色点代表肤色样本点,蓝色点为显示器中的可再现颜色)

Fig.3 Skin color distribution in YCbCr<sup>[17]</sup>(red dots represent skin color sample while blue dots represent the reproduction color on monitor)

夏思宇等学者<sup>[36]</sup>在Hsu基础上提出了一种自适应椭圆模型,使用白平衡法代替Hsu中的亮度补偿,利用退水法从二维直方图中获取新的椭圆模型中心,然后使用调整后的椭圆模型进行肤色检测。该方法中的结果一定程度上依赖于白平衡算法的实际效果以及退水法选择肤色峰的精确度。曾飞等学者<sup>[37]</sup>在Hsu自适应模型基础上,设计了可以随光照环境变化改变长短轴的椭圆模型,以解决人脸检测中的色彩偏移问题。在200

张不同环境的测试图像中,改进的模型相比 Hsu 的椭圆模型提高了 3.5% 的人脸检测精确率,同时漏检率和误检率也得到了一定程度的降低。

椭圆模型可以刻画肤色在色度平面的分布特征,但是难以描述肤色在椭圆区域中的概率分布,此外,在椭圆模型建立不准确时,会产生较差的效果。

### 3.1.3 自适应高斯模型

高斯模型是一种有效描述肤色分布的方法,描述了肤色的概率分布情况,主要使用的是单高斯模型(Single Gaussian Model,SGM)和混合高斯模型(Gaussian Mixture Model,GMM)。Sun 等学者<sup>[16]</sup>提出了一种基于直方图技术和动态 GMM 的视频皮肤区域跟踪方法。通过预训练的肤色模型对检测器进行初始化,然后利用 GMM 估计皮肤像素的分布,对每一幅图像的肤色模型进行微调。

熊霞<sup>[38]</sup>使用相关矩阵法对环境光照进行了估计,主要分为室内白炽灯、室外晴天、室内日光灯、室外阴天,如果检测图像属于某一预设的光照环境,则使用对应单高斯模型进行检测;反之,则使用 4 个 SGM 组成的 GMM 进行肤色检测。

张情等学者<sup>[39]</sup>在移动端掌纹分割应用中,通过“工”字形辅助定位手掌中心的小块区域,得到该子区域 Cb、Cr 的均值  $m_f$ ,根据图像中所有像素点到  $m_f$  的欧氏距离进行排序,选择距离小的点作为肤色参考点生成单高斯模型进行后续肤色检测。江国来等学者<sup>[40]</sup>使用固定阈值获得肤色的高概率区域。然后由此得到高斯模型的均值和方差参数,再使用高斯模型进行肤色检测,对检测结果中位于固定阈值之外的肤色区域,使用贝叶斯分类器做进一步判断。其中高概率生成的高概率区域和最终结果如图 4 所示。

高斯模型作为一种普遍使用的参数化模型,



图 4 文献[40]方法中的(a)原图,(b)肤色高概率区域,(c)最终结果。

Fig.4 (a) Original picture, (b) high probability point of skin color, (c) final result in reference [40].

能够较好地刻画肤色的分布规律,常用于进行肤色概率图的构建,在许多其他自适应模型中也有所体现,单独使用高斯模型的自适应文献相对较少。

### 3.1.4 其他自适应模型

Khan 等学者<sup>[41]</sup>提出了一种自动切换色彩空间的方法。该方法使用马科维兹最优化理论对不同色彩空间在肤色检测中的实际表现进行评估,并由此设定每个色彩分量的权重参数。这种方法可以利用色彩空间的不同特性进行最优组合,但是会大大提高检测方法的计算复杂度,类似的方法还有文献[42-43]等。自动选择色彩空间的方法试图寻找最优空间,但受色彩空间本身的限制,这种方法不一定能够获得明显的优化效果,但却会带来较大的计算负担。

Brancati 等学者<sup>[44]</sup>发现在 YCbCr 的子空间 YCb 和 YCr 中,肤色分布表现为梯形,如图 5 所示。通过方程拟合两个子空间的梯形边界,可以确定肤色的色彩范围。实验结果显示该方法在背景简单的情况下,对高低亮度均可以获得较好的结果,但是容易受大面积明、暗背景的影响而出现错误。

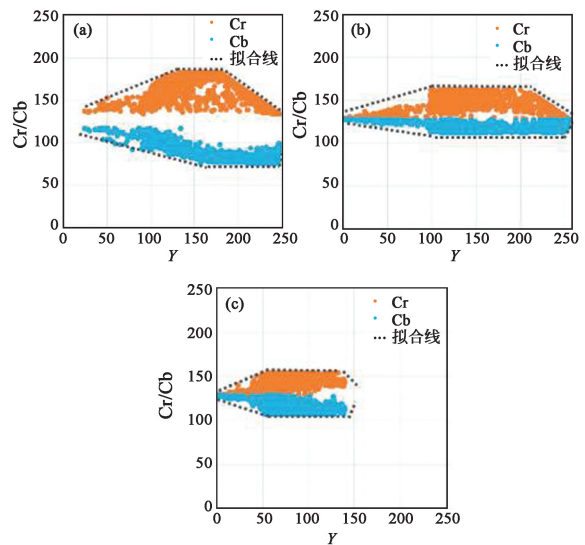


图 5 Brancati<sup>[44]</sup>提出的梯形肤色分布。(a)室内;(b)室外;(c)弱光环境。

Fig.5 Trapezoidal skin color distributions in different lighting conditions proposed by Brancati<sup>[44]</sup>. (a) Indoor; (b) Outdoor; (c) Low lighting.

## 3.2 基于高层语义特征的自适应方法

在实际的肤色检测任务中,由于变化的光照环境、复杂的背景等因素影响,仅依靠色彩信息或

简单纹理构建模型将不可避免地出现性能下降的现象。针对上述问题,学者们利用人脸等高层特征确定肤色区域,并根据可靠的肤色信息建立或调整肤色模型。

由于肤色是特指人体皮肤所表现出的颜色,而人脸作为图像中最常出现的肤色载体,同时人脸也是人眼重点关注的区域,通过建立肤色与人脸之间的联系可以有效提高肤色检测模型的适应性。此类方法通常假设图像中至少有一个可靠的人脸,并且能够被有效地检测到<sup>[45]</sup>。Wimmer 等学者<sup>[25]</sup>首次利用肤色与人脸在同一幅图像中的相关性,提出了可动态调整的肤色模型,混淆矩阵的行列式平均值可达 78.1。同时给出了 3 种对人脸区域肤色建模的常用方法:立方体聚簇、椭圆型聚簇以及基于规则的聚簇表示法。其中,立方体聚簇是通过计算每个分量在样本中的平均值和标准差。由于肤色分布为类高斯分布,因此通常根据 95% 置信区间近似边界设定色彩分量的上下限:

$$l_r = \mu_r - 2\sigma_r, u_r = \mu_r + 2\sigma_r, \quad (10)$$

$$l_g = \mu_g - 2\sigma_g, u_g = \mu_g + 2\sigma_g, \quad (11)$$

$$l_{\text{base}} = \mu_{\text{base}} - 2\sigma_{\text{base}}, u_{\text{base}} = \mu_{\text{base}} + 2\sigma_{\text{base}}, \quad (12)$$

式中,  $\sigma_i$  和  $\mu_i$  分别代表对应分量的标准差和均值,  $l$  和  $u$  分别代表对应分量的上下界。椭圆型聚簇是根据待检测颜色  $c_p$  与人脸区域肤色均值  $\mu$  之间的马氏距离(Mahalanobis distance)  $m_p$  来确定概率的,其中马氏距离  $m_p$  表示为:

$$m_p = (c_p - \mu)^T S^{-1} (c_p - \mu), \quad (13)$$

如果马氏距离  $m_p$  小于阈值  $t$ ,则可以认为当前颜色属于肤色,反之则属于非肤色。

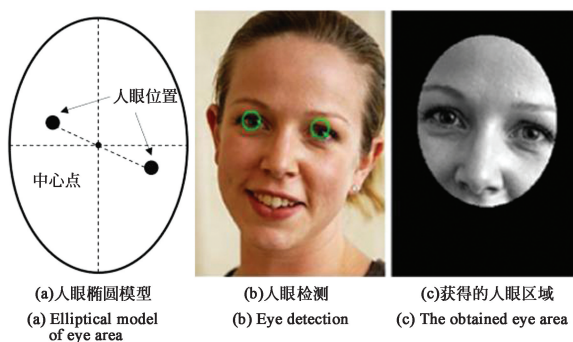


图 6 Tan 等学者<sup>[47]</sup>使用的检测人眼区域的方法  
Fig.6 Method of eye area detection in the work of Tan *et al*<sup>[47]</sup>.

Yogarajah 等学者<sup>[46]</sup>的研究也证实了以人脸颜色信息为依据,构建动态肤色模型的可行性。Tan 等学者<sup>[47]</sup>使用机器感知工具箱(Machine Perception Toolbox)<sup>[48]</sup>中预训练的 Haar-GentleBoost 模型检测人脸和人眼位置,并以两眼中心点为原点获取人眼椭圆区域,如图 6 所示。然后经过 Sobel 边缘检测去除纹理区域,并使用直方图与 GMM 的混合策略来生成肤色模型,需要对每一个检测到的人脸进行一次操作,最终将结果进行融合。但是由于人脸和人眼通常尺寸较小,仅根据其 Haar-like 特征进行检测容易受到画面中其他相似区域的干扰,从而产生较高的误检率。本方法的性能几乎完全取决于人脸和人眼检测的精确度,结果中的漏检和误检将容易导致无法建立准确的肤色模型,如图 7 所示。



图 7 Tan 等学者<sup>[47]</sup>方法中误检人脸时产生的肤色分割结果

Fig.7 Results of skin segmentation for false face region detected in of Tan *et al*<sup>[47]</sup>.

方晶晶等学者<sup>[49]</sup>首先使用 Adaboost 人脸检测器自动获取人脸的位置,并使用固定阈值排除明显不是肤色的区域,然后利用 GMM 来拟合人脸中心区域的色彩信息,从而形成适应性的肤色模型。Luo 等学者<sup>[50]</sup>提出了一种基于人脸定位和面部结构估计的自适应皮肤检测器,其中排除了纹理区域,提取了人脸的中心平滑区域来生成皮肤模型。此方法构建的肤色检测器在 TDSD 数据集(Test Database for Skin Detection)<sup>[51]</sup>中性能表现良好,能得到 93.74% 的准确率和 87.42% 的检出率。但是本方法为了提高精度,引

入了多个高斯模型来获取人脸区域的肤色分布,在对当前图像生成概率图之后,还需要再使用动态阈值法确定最终区别肤色与非肤色的概率阈值,因此该方法在提高结果的同时引入了大量的计算量和时间开销。此外,Mohanty 等学者<sup>[52]</sup>和 Shifa 等学者<sup>[6]</sup>使用了 RGB, HSV, YCbCr 的联合色彩空间来优化基于人脸的肤色模型。

上述研究都论述了如何从人脸区域获取可靠肤色,但并未考虑如何避免与肤色颜色相近的复杂背景对肤色检测造成的干扰。Bianco 等学者<sup>[53]</sup>根据皮肤区域是人体区域的一个子集的事实,提出了使用人体检测来优化肤色检测结果的方法,从而降低复杂背景的干扰,提供相对于单独使用人脸检测方案更可靠的肤色模型,可以在 TDSD 中得到更好的综合性能。该方法检测人体区域的过程如图 8 所示。

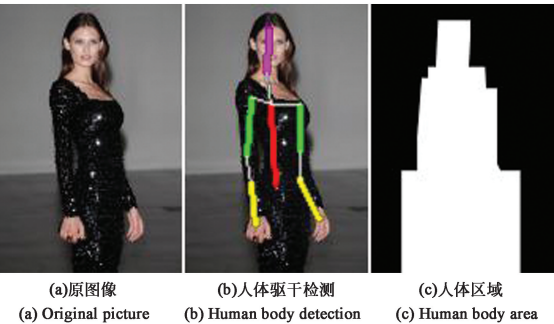


图 8 Bianco 等学者<sup>[53]</sup>所使用的人体区域检测示意图

Fig.8 Schematic diagram of human body area detection used by Bianco *et al*<sup>[53]</sup>

基于人的面部与手部具有相似肤色的假设,Hsieh 等学者<sup>[45]</sup>还提出了一种应用于手势识别的融合模型。该方法通过 Viola-Jones<sup>[54]</sup>人脸检测器定位人脸位置,舍弃总像素数小于阈值的人脸区域以达到减弱误检现象的目的,使用人脸的中心区域生成肤色模型。在归一化和标准 RGB 的混合空间 $[r, g, R]$ 中,依据人脸直方图分布规律形成自适应性的肤色模型,表现出对种族、肤质、偏色、光照环境等良好的适应性,整体过程可以实现  $640 \times 480$  下 39.5 FPS 的处理速度,平均精度达到 95.73%。基于高层语义的自适应方法与深度学习方法相比,在处理速度上和实现复杂度上具有一定优势。

在 TDSD 数据集中,本文定量地分析了上述几种自适应动态肤色检测方法,如表 1 所示。同时在表中也对比了一种常用的深度学习网络 DeepLab v3+,使用文献[18]中的训练方法进行测试分析。实验结果表明,基于高层语义特征的自适应方法通常能获得比基于动态参数调整的自适应模型更高的准确率和综合性能。虽然 Bianco 等人<sup>[53]</sup>的方法可以获得 0.854 0 的综合性能,但是该方法在无法检测到人体时会出现性能衰退。Luo 等人<sup>[50]</sup>的方法与前面的方法相比,可以实现较高的精确度,但会增大计算量,对实时性表现会造成一定的影响。与深度学习方法相比,自适应方法表现出较为接近的综合性能,在训练时间和计算成本上自适应方法占有一定的优势,同时减轻了对大规模肤色数据集的依赖。

表 1 肤色检测方法对比

Tab.1 Comparison of different skin detection methods

| 肤色模型                                   | 精确度     | 检出率     | 准确度     | 综合性能    |
|----------------------------------------|---------|---------|---------|---------|
| 基于自适应阈值的方法 <sup>[28]</sup>             | 0.441 1 | 0.850 1 | 0.675 1 | 0.580 8 |
| 基于梯形边界的自适应方法 <sup>[44]</sup>           | 0.602 9 | 0.819 5 | 0.792 9 | 0.694 7 |
| 基于人脸和人眼检测 <sup>[47]</sup>              | 0.585 7 | 0.859 2 | 0.830 3 | 0.696 6 |
| 基于人脸检测和面部结构 <sup>[50]</sup>            | 0.832 3 | 0.874 2 | 0.937 4 | 0.852 8 |
| 基于人脸和人体检测 <sup>[53]</sup>              | 0.843 9 | 0.864 3 | 0.925 4 | 0.854 0 |
| 基于 DeepLab v3+ 的深度学习方法 <sup>[18]</sup> | 0.874 0 | 0.913 2 | 0.930 1 | 0.889 6 |

## 4 结 论

自适应动态肤色检测方法是传统方法的改进,避免了过高的训练耗时,更加着重于实用性和易操作性,针对不同的应用场景可选择使用不同的自适应方法。基于高级语义特征的方法利用了肤色与人脸等部位之间的相关性,能够构建符合当前图像的肤色模型,但是容易受到人脸检测精确度、人脸非肤色区域的影响。在视频等图像序列中可以通过人脸跟踪建立随时间变化的动态模型。参数自适应调整参数的方法,虽然难以适应各种的复杂环境,但是在大多数应用场景中可以大幅提升传统方法的效率,满足实际使用需求。

在高性能、轻量化的深度学习技术全面产业化之前,实际非限制性场景中的肤色检测主要还是以自适应方法为主。随着硬件水平的不断增长,自适应方法可以结合部分深度学习技术,例如已经较成熟的轻量级人脸检测、跟踪网络来实现辅助检测;此外,可以通过使用更有效的色彩空间、形态学后处理步骤、不同的肤色检测算法组合来提高精度、降低复杂度。未来的肤色检测任务将会更加注重实时性的需求,与显卡等并行计算设备的结合也将是一个研究方向。

本文总结归纳了自适应动态肤色检测方法的研究现状,从基于色彩模型和基于人脸检测的自适应方法方面进行了详细介绍,分析了各种方法的优势和不足,为肤色检测研究人员提供参考。

## 参 考 文 献:

- [1] 林国军,杨明中,陈明举,等.一种肤色定位的人脸检测算法[J].液晶与显示,2019,34(1):70-73.  
LIN G J, YANG M Z, CHEN M J, *et al.* Face detection algorithm based on skin color location [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(1): 70-73. (in Chinese)
- [2] WANG L, HU W M, TAN T N. Recent developments in human motion analysis [J]. *Pattern Recognition*, 2003, 36(3): 585-601.
- [3] 谢非,龚俊,王元祥,等.基于肤色增强和分块 PCA 的人脸表情识别方法[J].南京师范大学学报(工程技术版), 2017,17(2):49-56.  
XIE F, GONG J, WANG Y X, *et al.* A facial expression recognition method based on skin color enhancement and block PCA [J]. *Journal of Nanjing Normal University (Engineering and Technology Edition)*, 2017, 17(2): 49-56. (in Chinese)
- [4] 史东承,倪康.基于 YCbCr 颜色空间背景建模与手势阴影消除[J].中国光学,2015,8(4):589-595.  
SHI D C, NI K. Background modeling based on YCbCr color space and gesture shadow elimination [J]. *Chinese Optics*, 2015, 8(4): 589-595, doi: 10.3788/CO.20150804.0589. (in Chinese)
- [5] WANG R B, GUO K Y, SHI S M, *et al.* A monitoring method of driver fatigue behavior based on machine vision [C]//*IEEE IV2003 Intelligent Vehicles Symposium*. Columbus: IEEE, 2003: 110-113.
- [6] SHIFA A, IMTIAZ M B, ASGHAR M N, *et al.* Skin detection and lightweight encryption for privacy protection in real-time surveillance applications [J]. *Image and Vision Computing*, 2020, 94: 103859.
- [7] 陈梦竹.基于肤色检测的中医面色识别[D].北京:北京交通大学,2018.  
CHEN M Z. TCM complexion recognition based on skin color [D]. Beijing: Beijing Jiaotong University, 2018. (in Chinese)
- [8] 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.
- [9] LEE J S, KUO Y M, CHUNG P C, *et al.* Naked image detection based on adaptive and extensible skin color model [J]. *Pattern Recognition*, 2007, 40(8): 2261-2270.
- [10] 周结.2019 年中国网络视听发展研究报告[EB/OL].(2019-05-27) [2020-07-12]. [https://www.sohu.com/a/316802357\\_728306](https://www.sohu.com/a/316802357_728306).  
ZHOU J. Research report on china's network audio visual development in 2019 [EB/OL]. (2019-05-27) [2020-07-12]. [https://www.sohu.com/a/316802357\\_728306](https://www.sohu.com/a/316802357_728306). (in Chinese)

- [11] ZAFARIFAR B, BELLERS E B, DE WITH P H N. Application and evaluation of texture-adaptive skin detection in TV image enhancement [C]//2013 *IEEE International Conference on Consumer Electronics*. Las Vegas: IEEE, 2013: 88-91.
- [12] CHEN H S, WANG T M, CHEN S H, *et al.* Skin-color correction method based on hue template mapping for wide color gamut liquid crystal display devices [J]. *Color Research and Application*, 2011, 36(5): 335-348.
- [13] 陈敏生, 刘政凯. 肤色检测技术综述[J]. 计算机学报, 2006, 29(2): 194-207.  
CHEN D S, LIU Z K. A survey of skin color detection [J]. *Chinese Journal of Computers*, 2006, 29(2): 194-207. (in Chinese)
- [14] 郑燕, 张霞. 基于像素的肤色检测方法综述[J]. 计算机系统应用, 2020, 29(3): 1-10.  
ZHENG Y, ZHANG X. Survey on pixel-based skin color detection methods [J]. *Computer Systems & Applications*, 2020, 29(3): 1-10. (in Chinese)
- [15] NAJI S, JALAB H A, KAREEM S A. A survey on skin detection in colored images [J]. *Artificial Intelligence Review*, 2019, 52(2): 1041-1087.
- [16] SUN H M. Skin detection for single images using dynamic skin color modeling [J]. *Pattern Recognition*, 2010, 43(4): 1413-1420.
- [17] HSU R L, ABDEL-MOTTALEB M, JAIN A K. Face detection in color images [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 696-706.
- [18] LUMINI A, NANNI L. Fair comparison of skin detection approaches on publicly available datasets [J]. *Expert Systems with Applications*, 2020, 160: 113677.
- [19] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [20] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]//18th *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2015: 234-241.
- [21] CHEN L C, ZHU Y K, PAPANDREOU G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//*Proceedings of the 15th European Conference on Computer Vision*. Cham: Springer, 2018: 801-818.
- [22] DOURADO A, GUTH F, DE CAMPOS T E, *et al.* Domain adaptation for holistic skin detection [J]. *arXiv pre-print arXiv*: 1903.06969, 2019.
- [23] HE Y, SHI J Y, WANG C, *et al.* Semi-supervised skin detection by network with mutual guidance [C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019: 2111-2120.
- [24] SCHMUGGE S J, JAYARAM S, SHIN M C, *et al.* Objective evaluation of approaches of skin detection using ROC analysis [J]. *Computer Vision and Image Understanding*, 2007, 108(1/2): 41-51.
- [25] WIMMER M, RADIG B. Adaptive skin color classifier [C]//*Proceedings of the first ICGST International Conference on Graphics Vision and Image Processing GVIP-05*. Cairo, Egypt: CICC, 2005, 1: 324-327.
- [26] 陶霖密, 彭振云, 徐光祐, 等. 人体的肤色特征[J]. 软件学报, 2001, 12(7): 1032-1041.  
TAO L M, PENG Z Y, XU F Y, *et al.* The features of skin color [J]. *Journal of Software*, 2001, 12(7): 1032-1041. (in Chinese)
- [27] JAYARAM S, SCHMUGGE S, SHIN M C, *et al.* Effect of colorspace transformation, the illuminance component, and color modeling on skin detection [C]//*Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004. Washington: IEEE, 2004: 813-818.
- [28] 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.
- [29] CHO K M, JANG J H, HONG K S. Adaptive skin-color filter [J]. *Pattern Recognition*, 2001, 34(5): 1067-1073.
- [30] 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): 681-90.
- [31] FOOPRATEEPSIRI R, KURUTACH W. A general framework for face reconstruction using single still image based on 2D-to-3D transformation kernel [J]. *Forensic Science International*, 2014, 236: 117-126.
- [32] 郭耸,顾国昌,蔡则苏,等.肤色相似度和动态阈值相结合的肤色分割技术[J].计算机工程与应用,2010,46(18):1-3,12.
- GUO S, GU G C, CAI Z S, *et al.* Skin segmentation using similarity of skin color and dynamic threshold [J]. *Computer Engineering and Applications*, 2010, 46(18): 1-3, 12. (in Chinese)
- [33] 张明吉,王伟强,郑清芳,等.静态图像中的自适应阈值肤色检测算法[J].计算机研究与发展,2006,43(9): 1674-1680.
- ZHANG M J, WANG W Q, ZHENG Q F, *et al.* Adaptive skin-color detection in still images [J]. *Journal of Computer Research and Development*, 2006, 43(9): 1674-1680. (in Chinese)
- [34] PHUNG S L, CHAI D, BOUZERDOUM A. Adaptive skin segmentation in color images [C]//*Proceedings of the 2003 International Conference on Multimedia and Expo*. Baltimore: IEEE, 2003: III-173.
- [35] 李晓光,李晓华,沈兰荪.一种 DCT 压缩域中基于三维椭圆模型的自适应肤色分割算法[J].电子学报,2005,33(12A):2464-2467.
- LI X G, LI X H, SHEN L S. A 3D ellipse model based adaptive skin color segmentation algorithm in DCT compressed domain [J]. *Acta Electronica Sinica*, 2005, 33(12A): 2464-2467. (in Chinese)
- [36] 夏思宇,李久贤,袁晓辉,等.一种改进的自适应肤色检测算法[J].数据采集与处理,2006,21(2):174-178.
- XIA S Y, LI J X, YUAN X H, *et al.* Improved adaptive algorithm for skin color detection [J]. *Journal of Data Acquisition & Processing*, 2006, 21(2): 174-178. (in Chinese)
- [37] 曾飞,蔡灿辉.自适应肤色检测算法的研究[J].微型机与应用,2011,30(4):37-40.
- ZENG F, CAI C H. Study on adaptive skin color detection algorithm [J]. *Microcomputer & its Applications*, 2011, 30(4): 37-40. (in Chinese)
- [38] 熊霞,桑庆兵.光照自适应的肤色检测方法[J].计算机应用,2011,31(5):1233-1236.
- XIONG X, SANG Q B. Illumination-adaptive skin color detection method [J]. *Journal of Computer Applications*, 2011, 31(5): 1233-1236. (in Chinese)
- [39] 张情,黎明,冷璐.结合自适应肤色模型和区域生长的掌纹分割[J].计算机工程与设计,2017,38(10):2794-2798.
- ZHANG Q, LI M, LENG L. Palmprint segmentation combining adaptive skin-color model and region growing [J]. *Computer Engineering and Design*, 2017, 38(10): 2794-2798. (in Chinese)
- [40] 江国来,林耀荣.自适应模型和固定模型结合的肤色分割算法[J].计算机应用,2010,30(10):2698-2701.
- JIANG G L, LIN Y R. Skin color segmentation algorithm combining adaptive model and fixed model [J]. *Journal of Computer Application*, 2010, 30(10): 2698-2701. (in Chinese)
- [41] KHAN R, HANBURY A, STÖTTINGER J, *et al.* Multiple color space channel fusion for skin detection [J]. *Multimedia Tools and Applications*, 2014, 72(2): 1709-1730.
- [42] GUPTA A, CHAUDHARY A. Robust skin segmentation using color space switching [J]. *Pattern Recognition and Image Analysis*, 2016, 26(1): 61-68.
- [43] GOMEZ G. On selecting colour components for skin detection [C]//*Object Recognition Supported by User Interaction for service robots*. Quebec City: IEEE, 2002, 2: 961-964.
- [44] BRANCATI N, DE PIETRO G, FRUCCI M, *et al.* Human skin detection through correlation rules between the YCb and YCr subspaces based on dynamic color clustering [J]. *Computer Vision and Image Understanding*, 2017, 155: 33-42.
- [45] HSIEH C C, LIOU D H, LAI W R. Enhanced face-based adaptive skin color model [J]. *Journal of Applied Science and Engineering*, 2012, 15(2): 167-176.
- [46] YOGARAJAH P, CONDELL J, CURRAN K, *et al.* A dynamic threshold approach for skin segmentation in color images [C]//*2010 IEEE International Conference on Image Processing*. Hong Kong: IEEE, 2010: 2225-2228.
- [47] TAN W R, CHAN C S, YOGARAJAH P, *et al.* A fusion approach for efficient human skin detection [J]. *IEEE*

- Transactions on Industrial Informatics*, 2012, 8(1): 138-147.
- [48] FASEL I, FORTENBERRY B, MOVELLAN J. A generative framework for real time object detection and classification [J]. *Computer Vision and Image Understanding*, 2005, 98(1): 182-210.
- [49] 方晶晶, 李振波, 姜宇. 人体肤色区域的自适应模型分割方法[J]. 计算机辅助设计与图形学学报, 2013, 25(2): 229-234.
- FANG J J, LI Z B, JIANG Y. Human skin color region segmentation based on adaptive model [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2013, 25(2): 229-234. (in Chinese)
- [50] LUO Y, GUAN Y P. Adaptive skin detection using face location and facial structure estimation [J]. *IET Computer Vision*, 2017, 11(7): 550-559.
- [51] ZHU Q, CHENG K T, WU C T. A unified adaptive approach to accurate skin detection [C]//2004 *International Conference on Image Processing*. Singapore: IEEE, 2004: 1189-1192.
- [52] MOHANTY R, RAGHUNADH M V. Skin color segmentation based face detection using multi-color space [J]. *International Journal of Advanced Research in Computer and Communication Engineering*, 2016, 5(5): 470-475.
- [53] BIANCO S, GASPARINI F, SCHETTINI R. Adaptive skin classification using face and body detection [J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 4756-4765.
- [54] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]//*Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Kauai: IEEE, 2001, 1: I.

#### 作者简介:



张 坤(1996—),男,河北石家庄人,硕士研究生,2018年于长春理工大学获得学士学位,主要从事图像处理技术方面的研究。E-mail: zk\_cohen61004@tju.edu.cn



雷志春(1964—),男,河北迁安人,博士,教授,1998年于德国多特蒙德大学获得博士学位,主要从事图像处理、显示技术等方面的研究。E-mail: zclei@tju.edu.cn