

基于像素的肤色检测方法综述^①

郑燕, 张霞

(武汉大学印刷与包装系, 武汉 430072)
通讯作者: 张霞, E-mail: zxx@whu.edu.cn



摘要: 肤色检测广泛应用于人脸检测、图像过滤、机器视觉等多种领域. 肤色检测的方法有很多种, 本文较全面地综述了基于像素的肤色检测方法, 重点分析了基于统计、基于阈值和基于机器学习等方法的原理和特点, 讨论了其优势和不足, 在此基础上探讨了基于像素的肤色检测方法的发展趋势.

关键词: 肤色检测; 统计法; 阈值法; 机器学习

引用格式: 郑燕, 张霞. 基于像素的肤色检测方法综述. 计算机系统应用, 2020, 29(3): 1-10. <http://www.c-s-a.org.cn/1003-3254/7301.html>

Survey on Pixel-Based Skin Color Detection Methods

ZHENG Yan, ZHANG Xia

(Department of Printing and Packaging, Wuhan University, Wuhan 430072, China)

Abstract: Skin color detection is widely used in many fields such as face detection, image filtering, and machine vision. There are many methods for skin color detection. This paper comprehensively reviews the pixel-based skin color detection methods, focusing on the principles and characteristics of statistics-based, threshold-based, and machine-based learning methods, and discusses their advantages and disadvantages. The development trend of pixel-based skin color detection methods is also discussed.

Key words: skin color detection; statistical method; threshold method; machine learning

1 引言

肤色检测技术在人脸检测与识别^[1-6]、视频监控系統^[7-10]、裸图像检测^[11-17]、基于内容的图像检索^[18,19]、手势识别^[20-26]、肤色病学诊断^[27]、驾驶员疲劳检测^[28]、人机交互^[29,30]等领域有着广泛的应用, 近年来得到了广泛的关注和研究.

肤色检测技术采用的方法有很多, 从不同的角度出发可以分为不同的类别. 根据有无涉及成像过程, 可以分为基于统计的方法和基于物理的方法^[31]; 根据肤色图像数据处理的尺度, 可以分为基于像素的检测和基于区域的检测. 基于像素的肤色检测方法针对单个像素进行肤色和非肤色的分类. 基于区域的肤色检测

方法本质和基于像素的检测方法相似, 但考虑了颜色的空间分布. 本文针对基于像素的肤色检测方法进行综述, 对比分析现有的研究成果, 探讨技术发展的热点和趋势.

基于像素的肤色检测技术通过建立肤色模型来区分肤色像素和非肤色像素, 肤色模型既可以表示为一组规则或阈值, 也可以从机器学习算法中训练出来^[32]. 主要的检测方法可分为基于统计的方法、基于阈值的方法和基于机器学习的方法. 本文第 2-4 部分分别对这 3 类检测方法进行分析讨论, 并在第 5 部分总结了肤色检测方法研究的特点, 探讨了肤色检测领域现存的问题及技术发展的趋势, 最后对全文进行了总结.

① 基金项目: 国家自然科学基金 (61505149)

Foundation item: National Natural Science Foundation of China (61505149)

收稿时间: 2019-07-25; 修改时间: 2019-08-23; 采用时间: 2019-09-04; csa 在线出版时间: 2020-02-28

2 基于统计的肤色检测方法

基于统计的方法将肤色检测视为一个概率问题, 通过在特定的颜色空间中训练数据集来计算肤色像素的概率. 这类方法包括非参数方法和参数方法, 具体分类如图1所示.

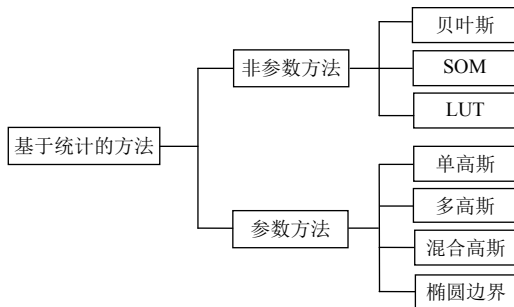


图1 基于统计的肤色检测方法

2.1 非参数方法

非参数方法又称为直方图模型法. 非参数方法通过估计颜色分布和量化模型之间的点对点映射关系, 将肤色概率值分配给颜色模型的每个离散点, 通过概率函数判断像素是否为肤色像素. 此类方法也称为肤色概率图 (SPM) 的构造, 即为离散化颜色空间中的每个点分配概率值. 代表方法主要有3种: 查找表 (LUT) 法、贝叶斯分类器和自组织映射 (SOM) 法.

查找表法是肤色建模常用方法, 该方法通过训练特定组肤色像素, 获得特定颜色空间中的肤色像素分布, 查找表法把 RGB 颜色空间视为最优颜色空间, RGB 直方图由 $256 \times 256 \times 256$ 个单元构成, 每个单元代表一个 $R_i G_i B_i$ 值的肤色概率. 其中每个可能为肤色像素点 $R_i G_i B_i$ 的最终概率值采用函数式 (1) 计算得到. 肤色

$$P(\text{skin}|R_i G_i B_i) = \frac{P(|R_i G_i B_i| \text{skin})P(\text{skin})}{P(|R_i G_i B_i| \text{skin})P(\text{skin}) + P(|R_i G_i B_i| \text{nonskin})P(\text{skin})P(\text{nonskin})} \quad (2)$$

使用肤色和非肤色直方图中的像素分别计算 $P(|R_i G_i B_i| \text{skin})$ 和 $P(|R_i G_i B_i| \text{nonskin})$. 在设计贝叶斯分类器时, 使用 ML (最大似然) 方法令 $P(\text{skin}) = P(\text{nonskin})$, 当两者比值 $> \theta$ (θ 是阈值) 时为肤色像素, 如式 (3) 所示.

$$\frac{P(|R_i G_i B_i| \text{skin})}{P(|R_i G_i B_i| \text{nonskin})} > \theta, \theta = K \left(\frac{P(\text{nonskin})}{P(\text{skin})} \right) \quad (3)$$

其中, K 值是可调参数, 用来消除算法对先验概率的依赖性. θ 的最佳值通过计算 ROC 曲线获得. Brand J 等^[36] 评估了基于像素级的人体肤色检测方法的性能, 其中

概率是肤色训练样本在这种颜色上所出现的相对频数, 如式 (1) 所示.

$$P(R_i G_i B_i) = \frac{R_i G_i B_i \text{出现次数}}{\text{颜色总数}} \quad (1)$$

查找表法学习过程简单, 但是填充直方图需要大量的肤色数据集. 例如, Jones MJ 等^[33] 采用查找表法从网络图像中学习肤色像素和非肤色像素构建数据集, 用于人体检测和成人图像识别, 发现数据集的大小直接影响肤色检测率. Nanni L 等^[34] 改进正面人脸检测 Viola-Jones 算法, 第一步利用查找表去除不包含肤色像素的人脸区域, 实验结果表明, 改进算法在彩色数据集检测率优于 Viola-Jones 算法. 此外可以通过融合不同颜色空间构造多查找表法, 例如, Ibraheem NA 等^[35] 将 RGB、HSV 和 YCbCr 等 3 种颜色空间结合起来, 可以减少光照变化带来的影响, 图 2 是 Ibraheem 结果处理图, 可以看出用于手势识别具有较好的鲁棒性.



图2 手势识别效果图^[35]

与查找表法不同, 贝叶斯分类器考虑不同颜色空间中肤色和非肤色像素之间的重叠问题. 在式 (1) 中, $P(R_i G_i B_i)$ 是已经假设观察到的像素属于肤色的条件概率, 但在 Jones MJ 等^[33] 的研究中, 97.2% 的颜色同时出现在肤色和非肤色上. 贝叶斯规则根据肤色像素直方图和非肤色像素直方图来建立关系检测肤色像素, 如式 (2) 所示.

贝叶斯的检测率高达 95%, 优于查找表检测方法, 证明了贝叶斯分类器在估计肤色概率方面的精确性.

查找表法与贝叶斯分类器可以构造一定精度的肤色检测器, 但不能实现自适应实时环境条件下的肤色检测. 自组织映射法是针对此局限性而提出的方法. SOM 源于 Kohonen 在 80 年代早期的设计, 采用无监督学习训练的人工神经网络, 由输入层和输出层构成. 算法的主要目的是将任意维数的输入信号模式转变为一维或二维的离散映射, 并且以拓扑有序的方式自适

应实现这个变换. SOM 满足非参数法的基本思想, 是特殊的无监督机器学习算法, 所以归为基于统计的肤色检测方法.

SOM 的自适应性在于获胜神经元的特定邻域中的神经元根据参数化学习函数将其码本矢量调整为更接近输入矢量. Brown DA 等^[37]提出的基于 SOM 的肤色分类器中, 对 500 多个彩色图像在 4 个颜色空间中 (HSV、笛卡尔 XY、TSL、归一化 RG) 训练了两个不同的 SOM, 通过训练肤色 SOM 和非肤色 SOM, 发现 SOM 分类器的性能几乎与颜色空间的选择无关, 解决了肤色检测最优颜色空间的选择问题, 并实现了 94% 的高精度检测率. 随后 Yin H 等^[38]提出自组织混合网络 (SOMN) 对此进行了修改, 用以提高 SOM 的稳定性和适用性.

SOM 不需要区分训练集和测试集, 同时具有把高维输入映射到低维的能力, 有自适应性和拓扑保形特性的优点. 但是网络训练时神经元存在不完全抑制情况, 并且当输入模型较少时, 分类结果依赖于模式输入的先后次序.

2.2 参数方法

参数方法通过拟合图像数据建立参数模型来描述肤色的分布. 参数方法与非参数方法相比, 可以采用较少的训练集来建立肤色模型, 同时可以弥补非参数方法的高存储要求. 参数方法主要包括单高斯模型 (SGM)、高斯混合模型 (GMM) 和椭圆模型 (EM) 等.

在 SGM 中, 平均向量周围存在平滑的高斯分布, 通过使用椭圆高斯在特定颜色空间将颜色向量 C 的多元正态分布建模为:

$$P(C; \mu, \Lambda) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Lambda|^{0.5}} \exp\left(-\frac{(C-\mu)^T \Lambda^{-1} (C-\mu)}{2}\right) \quad (4)$$

使用颜色空间为 C_j 的 N 个肤色像素训练数据集, 基于 ML(最大似然) 方法计算分布参数, 即均值向量 (μ) 和对角协方差矩阵 (Λ):

$$\begin{cases} \mu = \frac{1}{N} \sum_{j=1}^N C_j \\ \Lambda = \frac{1}{n-1} \sum_{j=1}^N (C_j - \mu)(C_j - \mu)^T \end{cases} \quad (5)$$

$P(C; \mu, \Lambda)$ 可以直接衡量像素 C 属于肤色的可能性. 也可以通过高斯分布参数计算输入像素 C 与均值 μ 的马氏距离 (MD): $\lambda_c(C_j) = (C_j - \mu)^T \Lambda^{-1} (C_j - \mu)$, 距离大小表示输入像素与肤色模型的接近程度.

SGM 方法比较简单, 具有计算成本低、存储要求低的优点, 但是在不同颜色空间中的性能不同. 例如, Zhu SP 等^[39]在 YCbCr 颜色空间中采用 SGM 进行肤色检测和唇部分割, 实验结果表明, 该方法满足实时性、鲁棒性和有效性的要求. Subban R 等^[40]比较了不同颜色空间中 SGM 的肤色检测性能, 结果显示 SGM 在 YPbPr 颜色空间性能优于 CIE-XYZ, YCC 和 YDbDr 颜色空间. Montenegro J 等^[41]基于自建数据库, 使用 MCC(马修斯相关系数) 作为评估指标在 RGB、HSV、YCbCr、CIE $L^*a^*b^*$ 和 CIE $L^*u^*v^*$ 颜色空间中评估 SGM 肤色检测性能, 得出 CIE $L^*a^*b^*$ 是性能最佳的肤色空间.

GMM 是多个高斯密度函数的标准化加权和^[42]. 它补偿了单高斯模型在处理一般肤色分割问题中误检率较高, 以及其峰值分布的不对称性而无法逼近实际分布的问题. GMM 定义如下:

$$P(C; \mu, \Lambda, w, N) = \sum_{j=1}^N w_j G_j(C; \mu_j, \Lambda_j) \quad (6)$$

其中, w_j 是每个核的权重, N 是单个高斯分量的总数. GMM 的学习过程与 SGM 的学习过程不同, 通常用期望最大化的迭代方法来估计拟合参数. 评估过程类似于 SGM, 可以利用概率本身或贝叶斯规则进行分割. GMM 模型具有评估过程简单和内存成本低的特点, 引起了研究者的广泛关注. Zhu QS 等^[43]在 GMM 的基础上, 提出了一种新的基于 GMM 视频肤色分割算法, 实验表明该算法能有效、快速地分割出运动目标, 具有较强的鲁棒性. Hossain F 等^[44]提出了一种通过估计图像的变化照明和使用 GMM 来进行面部肤色提取的方法, 实验结果表明与传统的基于 GMM 的肤色分割方法相比较, 该方法适用于手术过程等实际应用. 不过, GMM 的计算过程比 SGM 长且算法存在局限性. Caetano TS 等^[45]对高斯分布性能的比较研究表明, 混合模型仅在相关的操作区域 (正检率和误检率较高) 中改善性能, 此外增加内核的数量并不能有效提高模型效率. 图 3 是 8 种高斯模型的 ROC 曲线图, 可以看出所有高斯模型在操作点上表现出非常相似的性能. 因此, 混合高斯模型不一定是肤色建模的最佳选择.

EM 方法是在特定颜色空间中, 如 RGB、YCbCr、CIE $L^*a^*b^*$ 、CIE $L^*u^*v^*$ 、CIE-XYZ 等, 通过分析肤色簇的类椭圆形状来估计肤色分布模型的方法, Lee JY 等^[46]将椭圆模型定义为 $P(C; \mu, \Lambda)$:

$$P(C; \mu, \Lambda) = (C - \mu)^T \Lambda^{-1} (C - \mu) \quad (7)$$

$$\begin{cases} \mu = \frac{1}{N} \sum_{j=1}^N C_j \\ \Lambda = \frac{1}{N} \sum_{j=1}^N f_j (C_j - \Omega)(C_j - \Omega)^T \\ \Omega = \frac{1}{N} \sum_{j=1}^N f_j C_j \end{cases} \quad (8)$$

其中, N 是样本的总数, Ω 是色度矢量的平均值. 给定输入色度矢量 C , 可以使用阈值处理来计算像素的肤色概率^[47].

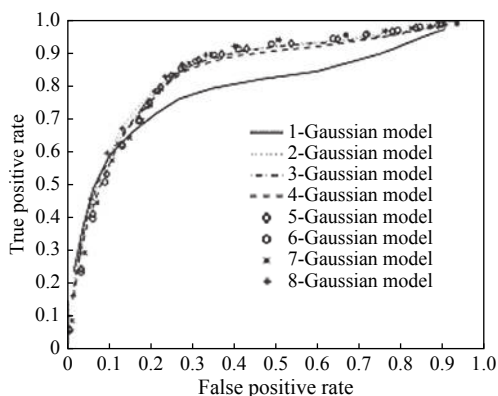


图3 8种高斯模型的ROC曲线图^[45]

Xu D 等^[48]采用椭圆模型结合深度信息, 在光照和复杂背景的变化环境下实现多人脸的定位. Tan WJ 等^[49]针对肤色检测设备成像问题, 在 YCbCr 空间进行椭圆拟合, 并将其用于手势分割, 实验结果表明, 该方法具有良好的精度, 对特定的成像设备具有更好的通用性和环境适应性. Li W 等^[50]于 CbCr-CgCr 空间提出双椭圆模型进行肤色分割和提取肤色区域, 结合 AdaBoost 级联分类器进行人脸检测, 该算法能够有效地应用于单面、多面、不同光照、背景复杂的人脸检测.

Lee JY 等^[46]在多个色度平面中 (rg、CIE-uv、CIE-ab、CIExy、IQ、CbCr) 对 EM 和高斯模型进行评估并得出结论: 在所有颜色空间中, 从检测率的角度来看, EM 优于 SGM 和 6 个内核的 GMM; 从训练的角度来看, EM 比 GMM 计算量少, 参数简单; 在计算速度上, EM 几乎与 SGM 一样快, 但它比 GMM 快.

综上所述, 基于统计的肤色检测方法的主要优势是模型的建立和实施比较容易, 集群形状具有独立性^[51]. 相对而言, 非参数方法的肤色检测率相对更高, 因为在

概率估计的过程中没有拟合误差, 但需要大量肤色像素作为训练集, 当需要精细分辨率时, 需要大型存储器支持^[51,52]. 参数方法不需要太多的训练集, 可以弥补非参数方法的高存储要求, 但是计算成本高.

3 基于阈值的肤色检测方法

基于阈值的肤色检测方法是通过明确规定肤色范围来检测肤色像素, 主要考虑不同种族人群肤色的主要区别在于强度 (亮度) 而不是色度的特征.

基于阈值构建肤色分类器是根据特定颜色空间中的肤色轨迹来定义分类规则, 图 4 是不同颜色空间中的肤色轨迹分布图. 可以看出该方法与颜色空间的选择密切相关^[5,53]. 通常, 分类规则用以确定颜色分量之间的关系, 需要对每个颜色分量确定合适的阈值. 选择阈值的方法有很多, 主要可以分为两种: 一种是手动选择阈值并以交互方式查看结果, 通过不断调试阈值得到最佳阈值, 称为阈值调试法; 一种是使用阈值算法自动计算阈值, 称为自适应阈值处理. 阈值调试法主要研究的是在特定颜色空间中明确定义肤色阈值范围的方法.

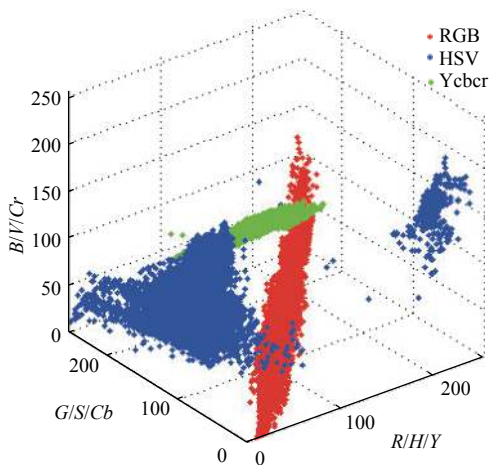


图4 不同颜色空间中肤色分布轨迹图

Kovac J 等^[5]在 RGB 颜色空间提出一种阈值定义方法, 在日光和闪光灯照明条件下, 若满足:

$$\begin{cases} R > 95, G > 40, B > 20, \max\{R, G, B\} - \min\{R, G, B\} > 15 \\ |R - G| > 15, R > G, R > B \end{cases} \quad (9)$$

则认为像素是肤色像素. 在手电筒照明中, 若满足:

$$R > 220, G > 210, B > 170, |R - G| < 15, R > G, R > B \quad (10)$$

则认为像素是肤色像素.

Gomez G 等^[53]提出结合 Harr 特征与肤色进行人脸检测的方法, 后处理采用了明确定义阈值边界的方法, 实验结果表明该方法优于著名的正面人脸检测 Viola-Jones 算法. 阈值调试法经常使用正交颜色空间进行阈值选择, 如 Sagheer A 等^[54]在 YCbCr 颜色空间确定了两种照明条件下 (正常光照和弱光照) 的肤色 CbCr 阈值范围, 式 (11) 和式 (12) 所示.

$$Cb \in \{110, 120\}, Cr \in \{135, 150\} \quad (11)$$

$$Cb \in \{110, 160\}, Cr \in \{120, 150\} \quad (12)$$

自适应阈值方法针对的是可变环境下的肤色检测, 研究的是在特定颜色空间中动态配置规则的方法. 如 Yang GZ 等^[55]为了检测面部肤色像素, 在 YCbCr 颜色空间基于像素亮度值的变化, 定义动态配置规则如下:

$$\text{if } Y > 128 \rightarrow \begin{cases} \theta_1 = -2 + \frac{256-Y}{16} \\ \theta_2 = 20 - \frac{256-Y}{16} \\ \theta_3 = 6 \\ \theta_4 = 8 \end{cases} \quad (13)$$

$$\text{if } Y \leq 128 \rightarrow \begin{cases} \theta_1 = 6 \\ \theta_2 = 12 \\ \theta_3 = 2 + \frac{Y}{32} \\ \theta_4 = -16 + \frac{Y}{16} \end{cases} \quad (14)$$

如果像素的 Cb, Cr 值满足以下条件, 则将其分类为肤色像素:

$$\begin{cases} Cr \geq -2(Cb + 24) \\ Cr \geq -(Cb + 17) \\ Cr \geq -4(Cb + 32) \\ Cr \geq 2.5(Cb + \theta_1) \\ Cr \geq \theta_3 \\ Cr \geq 0.5(\theta_4 - Cb) \\ Cr \leq \frac{220 - Cb}{6} \\ Cr \leq \frac{4}{3}(\theta_2 - Cb) \end{cases} \quad (15)$$

许多研究者采用感知颜色空间, 来定义明确的阈值规则. 如 Zahir NB 等^[56]提出了一种使用 HSV 颜色空间的肤色边界模型, 满足室内和室外照明条件. Pitas I 等^[57]在 HSV 颜色空间定义肤色像素阈值规则如式 (16) 所示.

$$V \geq 40, 0.2 < S < 0.6, (0^\circ < H < 25^\circ \text{ or } 355^\circ < H < 360^\circ) \quad (16)$$

Garcia C 等^[58]基于 HSV 颜色空间设计了更复杂

的规则, 如式 (17) 所示.

$$\begin{cases} V \geq 40 \\ H \leq -0.4V + 75 \\ 10 \leq S \leq (-H - 0.1V + 110) \\ \text{if } (H \geq 0) \rightarrow S < 0.08(100 - V)H + 0.5 \\ \text{if } (H < 0) \rightarrow S < 0.5H + 35 \end{cases} \quad (17)$$

自适应阈值方法的应用研究有很多, 如 Fooprateep-siri R 等^[59]在 HSV 颜色空间结合自适应阈值与面部重建算法, 提高人脸识别的准确性; Mariappan M 等^[60]在 CIE $L^*a^*b^*$ 颜色空间结合自适应阈值与椭圆模型算法来提高面部检测准确性, 精度达到 97% 左右. 此外, 还可以融合不同颜色空间构造检测规则来减少误报率, 例如 Xiang FH 等^[61]提出了一种基于 RGB 和 YUV 组合颜色空间检测肤色的方法, 与归一化 rg、HSV、YUV、YIQ、YCbCr、RGB、YCbCr-YUV 等 7 种颜色空间的检测结果进行比较, 基于 RGB-YUV 组合颜色空间是所提方法的最优颜色空间.

显式定义肤色阈值的方法具有分类规则简单、易于调整、计算成本低等优点, 但是对训练图像和颜色空间有高度依赖性, 另外, 有时阈值的选取不能描述像素的实际分布情况, 规则的有限性会降低分类器的检测性能.

4 基于机器学习的肤色检测方法

基于机器学习的肤色检测方法是从小组训练数据中学习构建肤色分类器, 无需构建明确的肤色模型. 机器学习分为有监督学习和无监督学习, 用于肤色检测的机器学习大多是有监督学习的方法. 比较热门的是采用人工神经网络 (ANN)^[62-65]的方法, 人工神经网络分类如图 5 所示.

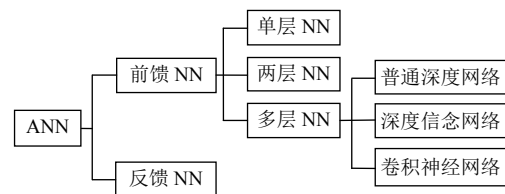


图 5 人工神经网络分类

人工神经网络模型主要考虑网络连接的拓扑结构、神经元特征、学习规则等. 人工神经网络的基本处理单元是神经元, 一般是多输入单输出的单元, 其结构模型如图 6 所示, 每两个神经元间的连接都代表一个对于通过该连接信号的加权重, 称之为权重.

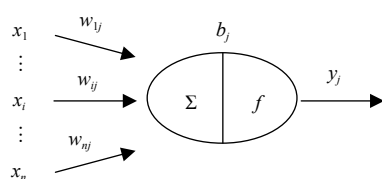


图6 神经元模型

其中, x_i 表示输入信号; n 个输入信号同时输入神经元 j , w_{ij} 表示输入信号 x_i 与神经元 j 连接的权重值, b_j 表示神经元的内部状态即偏置值, y_j 为神经元的输出. 输入与输出之间的对应关系可用式(18)表示:

$$y_j = f(b_j + \sum_{i=1}^n (x_i \times w_{ij})) \quad (18)$$

其中, $f(\cdot)$ 为激活函数, 可以有多种选择, 可以是线性纠正函数(ReLU)、Sigmoid函数、Tanh(x)函数、径向基函数等.

神经网络的输出取决于网络的结构、网络的连接方式、权重和激活函数. 人工神经网络的拓扑结构可以分为前馈神经网络和反馈神经网络, 前馈神经网络在肤色检测问题中被广泛采用^[62].

与传统的前馈神经网络不同, 多层神经网络即深度学习引入了用于训练具有多个隐藏层的神经网络. 深度学习本质上是构建含有多隐层的机器学习架构模型, 通过大规模数据进行训练, 得到大量更具代表性的特征信息. 深度学习算法打破了传统神经网络对层数的限制, 可根据研究者需要选择网络层数. 目前, 用于肤色检测的深度学习算法最具有代表性的有卷积神经网络(CNN)^[63,64]和递归神经网络(RNN)^[65]等.

采用神经网络与深度学习进行肤色检测主要目的有两个: 一是提高算法性能; 二是解决肤色检测存在的多种族肤色、多光照条件等肤色检测挑战问题. 如Yasar A等^[66]用UCI(University of California Irvine)数据库作为数据样本, 15%的编号数据作为测试数据, 15%的随机数据作为验证数据, 剩余的70%作为培训数据, 在RGB颜色空间训练数据集, 确定ANN神经元和偏差权重值. 最终由4个输入层、10个隐藏层和1个输出层构成的神经网络得到了平均值超过99%的验证性能, 同时可以消除年龄与种族带来的肤色检测影响. Vicentini RE等^[67]提出多层结构的前馈神经网络(RNA), 每个RNA的输入层具有与RGB、HSV和YCbCr颜色空间参数相对应的3个接收器节点, 中间层有20个神经元, 每个神经元的输出信号由LOGSIG

激活函数处理, 该函数将其输出范围限制在0(非肤色)和1(肤色)之间; 输出层包含单个神经元, 其激活函数为TANSIG, 结果显示, 在YCbCr颜色空间中对不同的光照条件有较好的检测性能, 一定程度上解决了光照条件对肤色检测效果的影响.

除此之外, 结合不同检测算法与特征提高肤色检测准确性已经成了一种新的研究趋势. 例如, Zuo HQ等^[65]考虑了像素与邻域的关系, 结合全卷积神经网络(FCN)和回归神经网络(RNN)进行人体肤色检测, 采用2个公开的肤色数据集COMPAQ和ECU进行检测, COMPAQ和ECU数据集都提供了具有多种属性的真实肤色数据, 包括肤色背景、不均匀照明、相机特征、种族、年龄等, 实验结果表明此方法有效地提高了复杂背景下肤色检测算法的稳定性.

Al-Mohair HK等^[68]考虑基于像素的检测方法性能会受到特定颜色空间中肤色和非肤色重叠的限制, 肤色点很难从非肤色点中分离出来. 为了提高肤色检测的准确性, 结合纹理特征提出了基于神经网络和 K -均值聚类技术的混合人体肤色检测算法, 实验结果表明, 该方法在ECU数据库(包含相同数量的肤色和非肤色像素)图像可以实现高效率肤色检测, 表1是文献^[68]的检测方法与相关文献中检测方法的结果对比. 可以看出神经网络结合 K -均值聚类方法有更高的检测率.

表1 不同检测方法检测率结果对比(%)

方法	召回率	准确率
MLP ANN+ K -均值聚类 ^[68]	88.00	87.82
基于阈值的方法(YCbCr+YUV) ^[61]	91.16	57.24
基于阈值的方法(RGB+YUV) ^[61]	89.62	71.25
贝叶斯分类器(YCbCr) ^[20]	62.87	49.83

总体而言, 深度学习的分层特性可以获得更多的信息, 在肤色检测精确度方面优于统计和阈值方法, 但是增加隐藏层提高性能的同时也会增加训练时间与计算成本, 确定网络结构与参数是其中的一大难点.

5 存在问题与发展趋势

通过对基于像素的肤色检测方法研究现状的分析, 可以发现, 基于像素的肤色检测方法大部分要考虑检测方法的颜色空间选择问题. 其中, 基于统计的非参数方法训练简单快速, 但强烈依赖于肤色簇的形状, 拟合度与训练集的大小会直接影响参数方法分类器的性能. 基于阈值的方法最容易操作, 但用其难以实现高识别

率, 主要困难是需要根据经验找到适当的颜色空间和决策规则. 与传统的检测方法相比, 机器学习会较少地依赖先验知识, 分类精确度较高, 但机器学习的研究难点是结构复杂度和高训练成本.

从这些技术现状可以看出, 基于像素的肤色检测目前仍然存在如下主要问题.

(1) 检测方法与颜色空间的选择. 基于像素的检测方法大部分都要确认颜色空间的选择, 颜色空间的选择与数据库图像若干因素相关, 例如照明、成像条件、噪声、图像尺寸及清晰度等. 这些图像可以改变特定颜色空间的检测结果. 为了找到检测方法的最优颜色空间, 应考虑可能影响性能的所有因素.

(2) 检测任务的数据样本与评价标准的建设. 目前, 虽然文献中有许多肤色模型, 但是如何测量模型的实用性, 以及该模型的有效性范围都存在限制. 通常, 模型的性能取决于许多因素, 例如使用的颜色空间、分布的形状、使用的参数、数据的性质、训练样本的大小、图像特征、噪声数据等. 为了定量客观地比较各种肤色检测技术, 需要加强公用肤色数据库的建设, 统一方法评测标准和规范.

(3) 光照、种族、成像设备和复杂背景等肤色挑战问题的解决与继续优化. 虽然许多检测方法考虑到肤色检测挑战带来的影响, 并试图解决这些不利因素. 但都是针对单一因素的解决, 需要改进算法解决多因素影响下的肤色挑战任务.

针对这些问题, 本文归纳了未来基于像素的肤色检测方法的一些发展趋势, 供读者参考.

(1) 结合空间信息的肤色检测. 肤色特征通常单独使用效果不佳, 结合与空间分布有关的信息, 如纹理、形状和梯度等, 可以有效提高肤色检测性能.

(2) 自适应动态肤色模型的建立. 手动建立的肤色模型在不同的检测条件下会有局限性. 建立动态肤色模型包括动态直方图、高斯分布自适应等肤色分布动态更新方法, 在提高检测性的同时更好地适应于不同的检测环境.

(3) 预处理与后处理步骤的加强. 预处理对于排除复杂背景以及减少光照影响至关重要, 可以通过照明补偿技术增强检测方法对光照的可控程度. 通过形态学后处理增强肤色分割的效果.

(4) 结合不同检测算法进行肤色检测与识别. 每种肤色检测算法都有各自的优缺点, 可以通过不同方法

的融合实现分类精度高、易操作和降低复杂度及训练成本, 同时提高肤色检测准确度的检测效果.

6 总结

基于像素的肤色检测方法是肤色检测技术采用的基础方法. 其中, 基于统计与基于阈值的肤色检测方法着重于参数与聚类模型的建立, 易操作; 基于机器学习的方法通过训练数据集找到肤色区域, 检测精度高. 它们的局限性都在于不能适应动态实时检测任务, 以及开放式数据库的检测. 本文探讨了基于像素的肤色检测方法相关理论与研究, 归纳了现存研究问题与发展趋势. 最终目的是为肤色检测的方法及其应用的研究提供参考.

参考文献

- 1 Zaqout I, Zainuddin R, Baba S. Human face detection in color images. *Advances in Complex Systems*, 2004, 7(3): 369–383. [doi: [10.1142/S021952590400024X](https://doi.org/10.1142/S021952590400024X)]
- 2 Li XH, Lam KM, Shen LS, *et al.* Face detection using simplified Gabor features and hierarchical regions in a cascade of classifiers. *Pattern Recognition Letters*, 2009, 30(8): 717–728. [doi: [10.1016/j.patrec.2009.03.006](https://doi.org/10.1016/j.patrec.2009.03.006)]
- 3 Verma A, Raj SA, Midya A, *et al.* Face detection using skin color modeling and geometric feature. *Proceedings of 2014 International Conference on Informatics, Electronics & Vision*. Dhaka, Bangladesh. 2014. 1–6.
- 4 Pujol FA, Pujol M, Jimeno-Morenilla A, *et al.* Face detection based on skin color segmentation using fuzzy entropy. *Entropy*, 2017, 19(1): 26. [doi: [10.3390/e19010026](https://doi.org/10.3390/e19010026)]
- 5 Kovac J, Peer P, Solina F. Human skin color clustering for face detection. *Proceedings of the IEEE Region 8 EUROCON: Computer as a Tool*. Ljubljana, Slovenia. 2003. 144–148.
- 6 Hsu RL, Abdel-Mottaleb M, Jain AK. Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 696–706. [doi: [10.1109/34.1000242](https://doi.org/10.1109/34.1000242)]
- 7 Gejguš P, Šperka M. Face tracking in color video sequences. *Proceedings of the 19th Spring Conference on Computer Graphics*. Budmerice, Slovakia. 2003. 245–249.
- 8 Kim MH, Park JB, Joo YH. New fuzzy skin model for face detection. *Proceedings of the 18th Australasian Joint Conference on Artificial Intelligence*. Sydney, Australia. 2005. 557–566.

- 9 Barbu T. Pedestrian detection and tracking using temporal differencing and HOG features. *Computers & Electrical Engineering*, 2014, 40(4): 1072–1079. [doi: [10.1016/j.compeleceng.2013.12.004](https://doi.org/10.1016/j.compeleceng.2013.12.004)]
- 10 Chaichulee S, Villarroel M, Jorge J, *et al.* Multi-task convolutional neural network for patient detection and skin segmentation in continuous non-contact vital sign monitoring. *Proceedings of the 12th IEEE International Conference on Automatic Face & Gesture Recognition*. Washington, DC, USA. 2017. 266–272. [doi: [10.1109/FG.2017.41](https://doi.org/10.1109/FG.2017.41)]
- 11 Perez M, Avila S, Moreira D, *et al.* Video pornography detection through deep learning techniques and motion information. *Neurocomputing*, 2017, 230: 279–293. [doi: [10.1016/j.neucom.2016.12.017](https://doi.org/10.1016/j.neucom.2016.12.017)]
- 12 Carlsson A, Eriksson A, Isik M. Automatic detection of images containing nudity: Image detection using artificial neural networks and statistical methods[Master's thesis]. New Zealand: Chalmers Tekniska Högskola, Chalmers University of Technology, 2008.
- 13 Duan L, Cui G, Gao W, *et al.* Adult image detection method base-on skin color model and support vector machine. *Proceedings of the 5th Asian Conference on Computer Vision*. Melbourne, Australia. 2002. 797–800.
- 14 Rowley HA, Jing YS, Baluja S. Large scale image-based adult-content filtering. *Proceedings of the First International Conference on Computer Vision Theory and Applications*. Setúbal, Portugal. 2006.
- 15 Lee JS, Kuo YM, Chung PC, *et al.* Naked image detection based on adaptive and extensible skin color model. *Pattern Recognition*, 2007, 40(8): 2261–2270. [doi: [10.1016/j.patcog.2006.11.016](https://doi.org/10.1016/j.patcog.2006.11.016)]
- 16 Sevimli H, Esen E, Ateş TK, *et al.* Adult image content classification using global features and skin region detection. In: Gelenbe E, Lent R, Sakellari G, *et al.*, eds. *Computer and Information Sciences*. Dordrecht: Springer, 2011. 253–258.
- 17 Chin TY. Fuzzy skin detection [Master's thesis]. Malaysia: University Teknologi Malaysia, 2008.
- 18 Albiol A, Torres L, Bouman CA, *et al.* A simple and efficient face detection algorithm for video database applications. *Proceedings 2000 International Conference on Image Processing*. Vancouver, Canada. 2000. 239–242. [doi: [10.1109/ICIP.2000.899286](https://doi.org/10.1109/ICIP.2000.899286)]
- 19 Ma WY, Zhang HJ. Content-based image indexing and retrieval. In: Furht B, ed. *Handbook of Multimedia Computing*. London: CPC Press, 1998. 227.
- 20 Rasim A, Alexander T. Hand detection based on skin color segmentation and classification of image local features. *TEM Journal*, 2013, 2(2): 150–155.
- 21 Habili N, Lim CC, Moini A. Segmentation of the face and hands in sign language video sequences using color and motion cues. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, 14(8): 1086–1097. [doi: [10.1109/TCSVT.2004.831970](https://doi.org/10.1109/TCSVT.2004.831970)]
- 22 de Ruijsscher BGC. FPGA based accelerator for real-time skin segmentation [Master's thesis]. Mekelweg: Delft University of Technology, 2006.
- 23 Bretzner L, Laptev I, Lindeberg T. Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering. *Proceedings of fifth IEEE International Conference on Automatic Face Gesture Recognition*. Washington, DC, USA. 2002. 423–428. [doi: [10.1109/AFGR.2002.1004190](https://doi.org/10.1109/AFGR.2002.1004190)]
- 24 Chen FS, Fu CM, Huang CL. Hand gesture recognition using a real-time tracking method and hidden Markov models. *Image and Vision Computing*, 2003, 21(8): 745–758. [doi: [10.1016/S0262-8856\(03\)00070-2](https://doi.org/10.1016/S0262-8856(03)00070-2)]
- 25 Jalab HA, Omer HK. Human computer interface using hand gesture recognition based on neural network. *Proceedings of the 2015 5th National Symposium on Information Technology: Towards New Smart World*. Riyadh, Saudi Arabia. 2015. 1–6.
- 26 Rautaray SS, Agrawal A. Vision based hand gesture recognition for human computer interaction: A survey. *Artificial Intelligence Review*, 2015, 43(1): 1–54. [doi: [10.1007/s10462-012-9356-9](https://doi.org/10.1007/s10462-012-9356-9)]
- 27 Schaefer G, Tait R, Zhu SY. Overlay of thermal and visual medical images using skin detection and image registration. *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society*. New York, NY, USA. 2006. 965–967.
- 28 Devi MS, Bajaj PR. Driver fatigue detection based on eye tracking. *Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology*. Nagpur, India. 2008. 649–652.
- 29 Chai D, Phung SL, Bouzerdoum A. Skin color detection for face localization in human-machine communications. *Proceedings of the Sixth International Symposium on Signal Processing and its Applications*. Kuala Lumpur, Malaysia. 2001. 343–346.
- 30 Jaimes A, Sebe N. Multimodal human-computer interaction: A survey. *Computer Vision and Image Understanding*, 2007,

- 108(1–2): 116–134. [doi: [10.1016/j.cviu.2006.10.019](https://doi.org/10.1016/j.cviu.2006.10.019)]
- 31 陈锻生, 刘政凯. 肤色检测技术综述. 计算机学报, 2006, 29(2): 194–207. [doi: [10.3321/j.issn:0254-4164.2006.02.002](https://doi.org/10.3321/j.issn:0254-4164.2006.02.002)]
- 32 Kawulok M, Nalepa J, Kawulok J. Skin detection and segmentation in color images. In: Celebi ME, Smolka B, eds. *Advances in Low-Level Color Image Processing*. Dordrecht: Springer, 2014. 329–366.
- 33 Jones MJ, Rehg JM. Statistical color models with application to skin detection. *International Journal of Computer Vision*, 2002, 46(1): 81–96. [doi: [10.1023/A:1013200319198](https://doi.org/10.1023/A:1013200319198)]
- 34 Nanni L, Lumini A, Dominio F, *et al.* Effective and precise face detection based on color and depth data. *Applied Computing and Informatics*, 2014, 10(1–2): 1–13. [doi: [10.1016/j.aci.2014.04.001](https://doi.org/10.1016/j.aci.2014.04.001)]
- 35 Ibraheem NA, Khan RZ. Multiple histogram technique for robust skin color based segmentation. *American Journal of Engineering Research (AJER)*, 2013, 2(5): 50–54.
- 36 Brand J, Mason JS. A comparative assessment of three approaches to pixel-level human skin-detection. *Proceedings 15th International Conference on Pattern Recognition*. Barcelona, Spain. 2000. 1056–1059.
- 37 Brown DA, Craw I, Lewthwaite J. A som based approach to skin detection with application in real time systems. *Proceedings of the British Machine Vision Conference*. Manchester, UK. 2001. 491–500.
- 38 Yin H, Allinson NM. Self-organizing mixture networks for probability density estimation. *IEEE Transactions on Neural Networks*, 2001, 12(2): 405–411. [doi: [10.1109/72.914534](https://doi.org/10.1109/72.914534)]
- 39 Zhu SP, Zhang N. Face detection based on skin color model and geometry features. *Proceedings of International Conference on Industrial Control and Electronics Engineering*. Xi'an, China. 2012. 991–994.
- 40 Subban R, Mishra R. Human skin segmentation in color images using gaussian color model. In: Thampi SM, Abraham A, Pal SK, *et al.*, eds. *Recent Advances in Intelligent Informatics*. Cham: Springer, 2014. 13–21.
- 41 Montenegro J, Gómez W, Sánchez-Orellana P. A comparative study of color spaces in skin-based face segmentation. *Proceedings of the 10th International Conference on Electrical Engineering, Computing Science and Automatic Control*. Mexico City, Mexico. 2013. 313–317.
- 42 Yang MH, Ahuja N. Gaussian mixture model for human skin color and its applications in image and video databases. *Proceedings of the SPIE 3656, Storage and Retrieval for Image and Video Databases VII*. San Jose: SPIE, 1999. 458–466. [doi: [10.1117/12.333865](https://doi.org/10.1117/12.333865)]
- 43 Zhu QS, Xie YQ, Gu J, *et al.* A new video object segmentation algorithm by fusion of spatio-temporal information based on GMM learning. In: Lee G, ed. *Advances in Automation and Robotics*, Vol. 2. Berlin, Heidelberg: Springer, 2011. 641–650.
- 44 Hossain F, Shamsi M, Alsharif MR, *et al.* Automatic facial skin detection using Gaussian mixture model under varying illumination. *International Journal of Innovative Computing, Information and Control*, 2012, 8(2): 1135–1144.
- 45 Caetano TS, Olabariaga SD, Barone DAC. Do mixture models in chromaticity space improve skin detection? *Pattern Recognition*, 2003, 36(12): 3019–3021. [doi: [10.1016/S0031-3203\(03\)00116-X](https://doi.org/10.1016/S0031-3203(03)00116-X)]
- 46 Lee JY, Yoo SI. An elliptical boundary model for skin color detection. *Proceedings of the 2002 International Conference on Imaging Science, Systems, and Technology*. Las Vegas, NV, USA. 2002.
- 47 Naji SN, Jalab HA, Kareem SA. A survey on skin detection in colored images. *Artificial Intelligence Review*, 2019, 52(2): 1041–1087. [doi: [10.1007/s10462-018-9664-9](https://doi.org/10.1007/s10462-018-9664-9)]
- 48 Xu D, Chen YL, Wu XY, *et al.* Integrated approach of skin-color detection and depth information for hand and face localization. *Proceedings of IEEE International Conference on Robotics and Biomimetics*. Karon Beach, Thailand. 2011. 952–956.
- 49 Tan WJ, Dai GY, Su H, *et al.* Gesture segmentation based on YCb'Cr' color space ellipse fitting skin color modeling. *Proceedings of the 2012 24th Chinese Control and Decision Conference*. Taiyuan, China. 2012. 1905–1908.
- 50 Li W, Yang QH, He XB. Face detection algorithm based on double ellipse skin model. *Proceedings of the 2011 IEEE 2nd International Conference on Software Engineering and Service Science*. Beijing, China. 2011. 335–339.
- 51 Vezhnevets V, Sazonov V, Andreeva A. A survey on pixel-based skin color detection techniques. In *Proceedings of Graphicon-2003*. Moscow, Russia. 2003. 85–92.
- 52 Shaik KB, Ganesan P, Kalist V, *et al.* Comparative study of skin color detection and segmentation in HSV and YCbCr color space. *Procedia Computer Science*, 2015, 57: 41–48. [doi: [10.1016/j.procs.2015.07.362](https://doi.org/10.1016/j.procs.2015.07.362)]
- 53 Gomez G, Morales EF. Automatic feature construction and a simple rule induction algorithm for skin detection. *Proceedings of the ICML Workshop on Machine Learning in Computer Vision*. Sydney, Australia. 2002. 31–38.
- 54 Sagheer A, Aly S. An effective face detection algorithm

- based on skin color information. Proceedings of 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems. Naples, Italy. 2012. 90–96. [doi: [10.1109/SITIS.2012.25](https://doi.org/10.1109/SITIS.2012.25)]
- 55 Yang GZ, Huang TS. Human face detection in a complex background. Pattern Recognition, 1994, 27(1): 53–63. [doi: [10.1016/0031-3203\(94\)90017-5](https://doi.org/10.1016/0031-3203(94)90017-5)]
- 56 Zahir NB, Samad B, Mustafa M. Initial experimental results of real-time variant pose face detection and tracking system. Proceedings of 2013 IEEE International Conference on Signal and Image Processing Applications. Melaka, Malaysia. 2013. 264–268.
- 57 Pitas I, Sobottka K. A novel method for automatic face segmentation, facial feature extraction and tracking. Signal Processing: Image Communication, 1998, 12(3): 263–281. [doi: [10.1016/S0923-5965\(97\)00042-8](https://doi.org/10.1016/S0923-5965(97)00042-8)]
- 58 Garcia C, Tziritas G. Face detection using quantized skin color regions merging and wavelet packet analysis. IEEE Transactions on Multimedia, 1999, 1(3): 264–277. [doi: [10.1109/6046.784465](https://doi.org/10.1109/6046.784465)]
- 59 Fooprateepsiri R, Kurutach W. A general framework for face reconstruction using single still image based on 2D-to-3D transformation kernel. Forensic Science International, 2014, 236: 117–126. [doi: [10.1016/j.forsciint.2013.12.020](https://doi.org/10.1016/j.forsciint.2013.12.020)]
- 60 Muralindran M, Nadarajan M, Porle RR, *et al.* A LabVIEW design for frontal and non-frontal human face detection system in complex background. Applied Mechanics and Materials, 2014, 490–491: 1259–1266. [doi: [10.4028/www.scientific.net/AMM.490-491.1259](https://doi.org/10.4028/www.scientific.net/AMM.490-491.1259)]
- 61 Xiang FH, Suandi SA. Fusion of multi color space for human skin region segmentation. International Journal of Information and Electronics Engineering, 2013, 3(2): 172–174.
- 62 Deng L, Yu D. Deep learning: Methods and applications. Foundations and Trends® in Signal Processing, 2014, 7(3–4): 197–387. [doi: [10.1561/20000000039](https://doi.org/10.1561/20000000039)]
- 63 Kim Y, Hwang I, Cho NI. A new convolutional network-in-network structure and its applications in skin detection, semantic segmentation, and artifact reduction. arXiv: 1701.06190, 2017.
- 64 Kim Y, Hwang I, Cho NI. Convolutional neural networks and training strategies for skin detection. Proceedings of IEEE International Conference on Image Processing. Beijing, China. 2017. 3919–3923. [doi: [10.1109/ICIP.2017.8297017](https://doi.org/10.1109/ICIP.2017.8297017)]
- 65 Zuo HQ, Fan H, Blasch E, *et al.* Combining convolutional and recurrent neural networks for human skin detection. IEEE Signal Processing Letters, 2017, 24(3): 289–293. [doi: [10.1109/LSP.2017.2654803](https://doi.org/10.1109/LSP.2017.2654803)]
- 66 Yasar A, Saritas I. Managing intellectual capital and innovation for sustainable and inclusive society: Managing intellectual capital and innovation. Proceedings of the MakeLearn and TIIM Joint International Conference. 2015. 2021–2027.
- 67 Vicentini RE, Lotufo ADP. Skin detection in digital images with artificial neural networks. Proceedings of Argentine Conference on Automatic Control. Buenos Aires, Argentina. 2018. 1–6. [doi: [10.23919/AADECA.2018.8577440](https://doi.org/10.23919/AADECA.2018.8577440)]
- 68 Al-Mohair HK, Saleh JM, Suandi SA. Hybrid human skin detection using neural network and k-means clustering technique. Applied Soft Computing, 2015, 33: 337–347. [doi: [10.1016/j.asoc.2015.04.046](https://doi.org/10.1016/j.asoc.2015.04.046)]