

基于数据挖掘的图像压缩域肤色检测算法

赵士伟,卓 力,王素玉,沈兰荪

(北京工业大学信号与信息处理研究室,北京 100124)

摘 要: 提出了一种直接在 JPEG 图像压缩域进行肤色检测的算法. 该算法首先在熵解码后的 DCT 系数中提取图像块的颜色特征和纹理特征,然后利用数据挖掘建立用于表征压缩域图像特征和肤色检测结果之间关系的肤色模型,并利用该模型进行初步肤色检测,最后利用区域生长的方法分割出图像中的肤色区域. 实验结果表明,与像素域的 SPM (Skin Probability Map) 肤色检测算法相比,本文方法可以获得更高的检测准确率和更快的检测速度.

关键词: 肤色检测; 压缩域; 数据挖掘; 区域生长

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2010) 03-0605-06

Data-Mining Based Skin Color Detection Method in Compressed Domain

ZHAO Shi-wei, ZHUO Li, WANG Su-yu, SHEN Lan-sun

(Signal and Information Processing Lab, Beijing University of Technology, Beijing 100124, China)

Abstract: A novel skin color detection method in JPEG compressed domain has been proposed. Color and texture features of the image blocks are extracted from the entropy decoded DCT coefficients firstly. Then, data mining method is applied to set up the skin color model to describe the relationship between the image block features and the skin detection results, and the initial skin image blocks are detected based on the model. The skin color regions are finally segmented using region growing method. Experimental results show that, compared with the SPM (Skin Probability Map) skin color detection algorithm in pixel domain, the proposed method can achieve higher detection accuracy and faster detection speed.

Key words: skin color detection; compressed domain; data mining; region growing

1 引言

肤色检测是人脸检测与识别、手势识别、敏感图像过滤等多种应用的关键技术. 肤色检测的准确率和速度直接影响到这些应用系统的性能.

目前肤色检测方法总体可分为两类: 基于像素域的方法和基于压缩域的方法. 其中基于像素域的方法一般利用单像素的颜色信息或者结合在变换域提取的纹理信息检测肤色^[1,2]. 但是当前图像大多以压缩形式存储和传输, 像素域的肤色检测方法需要首先将压缩码流全部解码, 这不仅加重了系统的计算和存储负担, 也大大增加了肤色检测所需的时间. 与像素域的肤色检测方法不同, 压缩域的肤色检测在不解码或部分解码的码流中提取图像特征, 并建立肤色模型检测肤色. 由于无需对码流完全解码, 因此压缩域的肤色检测方法在很大程度上提高了肤色检测的速度^[3,4].

JPEG 是网络传输和存储中广泛应用的图像压缩格式之一. 本文首先在 JPEG 压缩码流中提取颜色、纹理等

特征, 然后利用数据挖掘探寻这些特征与肤色检测结果之间的规律, 并据此建立肤色模型用于对图像中的肤色区域进行初步检测, 最后采用区域生长的方法自适应地分割图像中的肤色区域. 实验结果表明, 与现有的具有代表性的像素域肤色检测算法—SPM (Skin Probability Map) 算法相比, 本文提出的压缩域肤色检测算法可以获得更高的检测准确率和更快的检测速度.

2 基于数据挖掘的压缩域肤色检测

本文提出的压缩域肤色检测方法如图 1 所示, 可分为训练和检测两个阶段. 在训练阶段, 首先对 JPEG 压缩码流进行熵解码, 然后从中提取颜色、纹理以及其他一些图像特征. 接下来利用数据挖掘技术探寻这些特征与肤色判断结果之间的内在规律, 并建立肤色模型. 在检测阶段, 首先从熵解码之后的数据中提取与肤色模型相关的图像特征, 并初步检测图像中的肤色区域, 然后采用区域生长算法完成肤色区域的最终检测.

从图 1 中可以看出, 特征提取和肤色建模是本文提

出算法中的两个关键问题.接下来,本文基于 JPEG 压缩码流的特点,首先介绍压缩域特征提取和基于数据

挖掘的肤色建模两个关键技术,最后介绍所采用的区域生长算法.

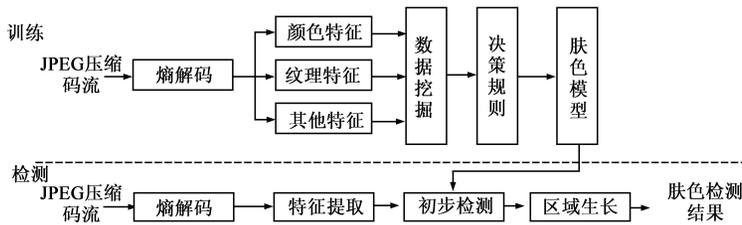


图1 基于压缩域的肤色检测方法流程图

2.1 压缩域特征提取

JPEG 压缩标准采用了 YCbCr 颜色空间,首先将图像划分为 8×8 大小的图像块,然后经过 DCT 变换、量化和熵编码等步骤对图像进行压缩.压缩后的 JPEG 码流中,可用信息主要包括图像的宽、高、压缩时采用的量化表以及熵编码后的 DCT 系数等.其中熵编码后的 DCT 数据是非结构化的,不利于直接分析,因而本文方法在熵解码后的 DCT 系数中提取颜色和纹理等图像特征.

下面介绍本文算法所采用的压缩域颜色和纹理特征提取方法.

2.1.1 压缩域颜色特征提取

本文提取的压缩域颜色特征与文献[3]类似,利用图像块内所有像素的颜色平均值表示.对于 8×8 大小的图像块,其颜色平均值 μ_{p_color} 可由式(1)获得,具体推导过程可参见文献[3].

$$\mu_{p_color} = \frac{1}{8} F(0,0) \quad (1)$$

其中 $F(0,0)$ 为 8×8 图像块经 DCT 变换后的 DC 系数.

由于压缩过程中 DCT 变换后的系数需要经过量化处理,因而式(1)可由量化后的 DC 系数与量化因子的乘积近似得出,如式(2)所示:

$$\mu_{p_color} = \frac{1}{8} F(0,0) \approx \frac{1}{8} F^Q(0,0) * Q(0,0) \quad (2)$$

式(2)中 $Q(0,0)$ 是量化表中 $(0,0)$ 处的值,可在 JPEG 码流中直接获取, $F^Q(0,0)$ 是熵解码后的 DC 系数,可在熵解码后的码流中获取.因而本文将式(2)定义为压缩域的颜色特征,用 μ_{c_color} 表示,如式(3)所示:

$$\mu_{c_color} = \frac{1}{8} F^Q(0,0) * Q(0,0) \quad (3)$$

文中将利用式(3)求得的各个颜色分量的均值作为图像块压缩域的颜色特征,并分别记为 YDC 、 $CbDC$ 、 $CrDC$.

2.1.2 压缩域纹理特征提取

文献[3]利用 AC 系数的能量均值作为图像块的纹理特征,这种方法在计算纹理特征时需要较多的计算量.文献[5]研究发现 JPEG 码流中非零 AC 系数的个数与 AC 系数的能量具有很强的相关性,也能够较好的表

示图像块的纹理特征,并且计算速度快.因而本文采用了类似文献[5]的方法,但为了消除不同量化系数对非零 AC 系数个数的影响,在计算图像块的纹理特征时引入了与量化系数相关的阈值,如式(6)所示.

为了充分考察不同频率的 AC 系数对肤色检测的影响,本文首先根据 AC 系数的频率特性,按图 2 所示的 2 种方式对 8×8 的 DCT 系数块进行划分,分别得到了 7 个不同频率的区域,即 $S_1 \sim S_7$.其中图 2(a)是文献[6]中采用的 DCT 系数划分方式,图 2(b)是文献[7]中采用的划分方式.接下来,统计各区域内绝对值大于阈值的 AC 系数的个数作为压缩域的纹理特征,如式(4)~(6)所示.

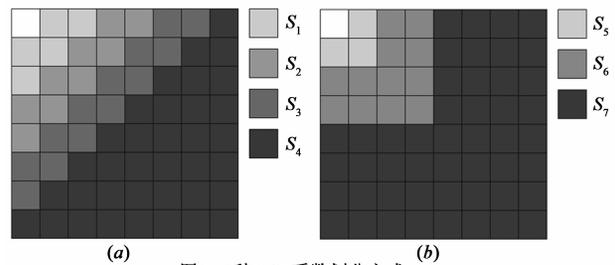


图2 2种DCT系数划分方式

$$\begin{cases} V_i = \sum_{S_i} U(F^Q(u, v)) \\ V_8 = \sum_{S_2+S_3+S_4} U(F^Q(u, v)) \\ V_9 = \sum_{S_5+S_6} U(F^Q(u, v)) \\ V_{10} = \sum_{S_5+S_6+S_7} U(F^Q(u, v)) \\ V_{11} = \sum_{S_6+S_7} U(F^Q(u, v)) \end{cases} \quad (4)$$

其中 S_i 为图 2 中标出的各个区域.函数 $U(F^Q(u, v))$ 定义为:

$$U(F^Q(u, v)) = \begin{cases} 1 & , (|F^Q(u, v)| > TH_AC_{u,v}) \\ 0 & , (|F^Q(u, v)| \leq TH_AC_{u,v}) \end{cases} \quad (5)$$

式(5)中 $TH_AC_{u,v}$ 为:

$$TH_AC_{u,v} = \frac{Q_s(u, v)}{Q(u, v)} \quad (6)$$

其中 $Q_s(u, v)$ 为 ITU-81(1992)附录 K 中列出的量化表中 (u, v) 处的系数值; $Q(u, v)$ 为从待检测图像 JPEG 码流中提取的量化表中 (u, v) 处的值. 由于人眼对亮度分量的变化比较敏感, 因而本文仅计算了图像块中 Y 分量的 $V_i (i = 1 \cdots 11)$ 值作为纹理特征.

在 JPEG 码流中还可以直接提取图像的宽度 (W)、高度 (H), 以及图像的压缩比 ($CR, \text{Compression Ratio}$) 等可用信息, 这些信息都能在一定程度上说明图像的性质, 文中也提取并考察了这些特征对肤色检测的影响.

2.2 基于数据挖掘的肤色建模

采用 2.1 节所述的方法, 本文在 JPEG 压缩码流中共提取了包括颜色、纹理、图像宽度与高度等在内的 17 个特征, 如表 1 所示. 其中颜色特征可由式(3)得出, 纹理特征可由式(4)得出, 其他基本特征可直接从 JPEG 码流中获取.

这些特征与肤色检测结果之间的内在规律无法仅仅通过观测或者经验轻易判定, 但是可以在网络中获取大量的样本数据. 由于数据挖掘技术能够在海量数据中提取有效、可信和可行的信息^[8], 因而本文采用数

据挖掘的方法在大量的训练样本中探寻上述特征与肤色检测结果之间的潜在规律.

表 1 压缩域中提取的特征

特征类型	特征变量
颜色特征	$YDC, CrDC, CbDC$
纹理特征	$V_1 \sim V_{11}$
基本特征	W, H, CR

Analysis Services 作为 Microsoft 公司提供的数据挖掘平台, 集成了多种数据挖掘算法, 其中 Microsoft 决策树是一种被广泛应用的用于分类的数据挖掘技术. 应用决策树算法建立的模型易于理解, 并且容易从中提取“IF-THEN”规则, 便于计算机实现, 因而本文采用该算法进行数据挖掘.

通过对大量样本数据的挖掘获得了利用上述特征判断肤色的决策树模型, 如图 3 所示. 经过对该模型的分析, 发现上述提取的 17 个特征中有 5 个特征与肤色检测结果的相关性最大, 按照相关性程度这 5 个特征依次排列为 $CrDC, V_1, V_2, YDC, CbDC$.

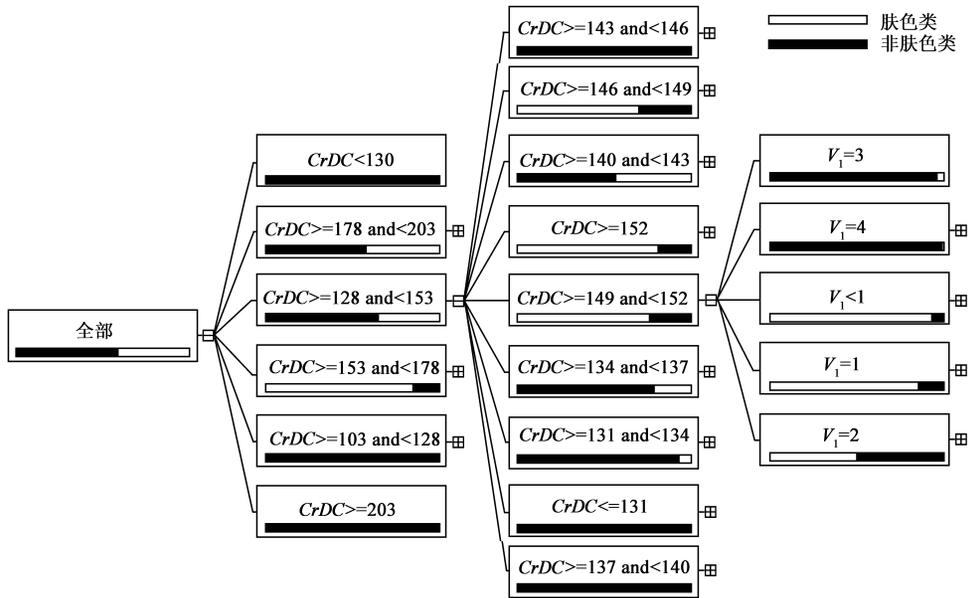


图 3 基于决策树的肤色模型

本文从图 3 所示的决策树中提取了 N 条肤色判断规则, 并将每个规则对应的肤色类和非肤色类的类条件概率密度的似然比作为规则的判断结果. 似然比 $r_i(\mathbf{x})$ 的计算方法如式(7)所示:

$$\begin{cases} r_i(\mathbf{x}) = \frac{p_i(\mathbf{x} | w_s)}{p_i(\mathbf{x} | w_n)} \\ p_i(\mathbf{x} | w_{skin}) = \frac{S(i)}{C(w_s)} \\ p_i(\mathbf{x} | w_{nonskin}) = \frac{N(i)}{C(w_n)} \\ i = 1 \cdots N \end{cases} \quad (7)$$

其中 \mathbf{x} 为 8×8 图像块的特征向量; w_s, w_n 分别表示肤色类和非肤色类; $p_i(\mathbf{x} | w_s), p_i(\mathbf{x} | w_n)$ 分别为符合规则 i 的肤色类和非肤色类的类条件概率密度; $S(i), N(i)$ 分别是训练样本中符合第 i 条规则的肤色样本和非肤色样本的数目; $C(w_s), C(w_n)$ 是训练样本中总的肤色样本和非肤色样本的数目.

为了便于统计图像中肤色概率分布, 本文按式(8)所示将似然比离散化为 48 级, 并用 $f_i(\mathbf{x})$ 表示:

$$\hat{r}_i(\mathbf{x}) = \begin{cases} \left[\frac{r_i(\mathbf{x})}{0.1} \right] & , r_i(\mathbf{x}) \in [0, 2) \\ \left[\frac{r_i(\mathbf{x}) - 2}{0.2} \right] + 20 & , r_i(\mathbf{x}) \in [2, 4) \\ \lfloor r_i(\mathbf{x}) - 4 \rfloor + 30 & , r_i(\mathbf{x}) \in [4, 10) \\ \left[\frac{r_i(\mathbf{x}) - 10}{5} \right] + 36 & , r_i(\mathbf{x}) \in [10, 30) \\ \left[\frac{r_i(\mathbf{x}) - 30}{10} \right] + 40 & , r_i(\mathbf{x}) \in [30, 100) \\ 47 & , r_i(\mathbf{x}) \in [100, \infty) \end{cases} \quad (8)$$

2.3 基于区域生长的肤色分割

图像中的肤色通常成片出现,判断图像块是否属于肤色区域不仅与待检测图像块的特征有关,而且与相邻图像块的类别也有一定关系^[9].区域生长是一种能够考虑邻域信息,将具有相似性质的像素集合起来构成区域的一种图像分割技术.利用区域生长检测肤色区域是一种有效的方法,其难点在于种子点的自适应选择和生长准则的建立^[9].

2.3.1 种子点的自适应选择

根据贝叶斯决策的原理肤色类与非肤色类的类条件概率密度的似然比越高,则待检测图像块为肤色的概率越高.但一般情况下,即使是在高似然比的图像区域也极有可能混淆着高度类似肤色的背景.因此,在种子点选择过程中如何去除背景区域的干扰是影响区域生长方法性能的关键因素之一.

在同一幅图像中肤色区域往往具有相似的颜色,而与背景区域的颜色差别较大.如果能够在图像中找到少量的真正的肤色区域,则可通过颜色差异消除部分背景区域干扰.基于这一思想本文采用了如图4所示的种子点选择方法.

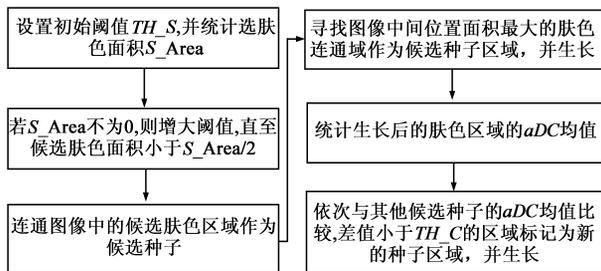


图4 种子点选择流程框图

首先设置初始似然比阈值 TH_S_1 , 并计算图像中候选肤色区域(将图像中似然比大于阈值的区域称为候选肤色区域)的面积 S_Area . 若面积为 0, 则认为图像中无肤色区域, 否则增大阈值直至候选区域面积小于 $S_Area/2$. 这样既可排除部分类似肤色区域的背景噪声, 而又不致于对真正的肤色区域产生大的影响. 然后

连通图像中的候选肤色区域作为候选种子, 并从中选择面积最大的连通域作为最佳种子区域并生长. 最后依次比较其他候选种子区域与由最佳种子区域生长完成的肤色区域的颜色差异, 若差异较小则将候选种子区域标记为新的种子区域. 由于 2.2 节指出颜色特征 $CrDC$ 与建立的肤色模型相关程度最大, 因而此处利用该特征衡量各区域颜色的差异.

2.3.2 区域的生长准则

选定种子后, 如何控制区域的生长(即如何制定区域生长和停止准则)也是至关重要的问题. 考虑到图像中的肤色区域在颜色和似然比方面具有连续性, 本文采用了如下的生长准则:

- ① 如果待生长点的 8 邻域中有不少于 7 个已被检为肤色, 则生长;
- ② 如果待生长点的 8 邻域中有不少于 4 个已被检为肤色, 且该点的似然比大于最小似然比阈值 TH_S_{min} , 则生长;
- ③ 如果待生长点与相邻的种子点颜色差小于阈值 TH_C_2 , 且该点的似然比与相邻的种子点差值小于阈值 TH_DS , 且该点的似然比大于阈值 TH_S_{min} , 则生长.

生长准则①②主要考虑了图像中的肤色区域一般成片出现的特点; 生长准则③则综合考虑了图像中的同一肤色区域具有相似颜色及肤色似然比的特点.

3 实验结果与分析

我们在网络中随机抓取了 4639 幅尺寸在 400×600 至 2832×2128 之间的 JPEG 格式图像进行实验, 其中 2530 幅图像含有人体肤色区域. 为了准确的提取肤色样本和非肤色样本, 我们在含有大片人体肤色区域的图像中手工标注出 1,049,584 个 8×8 的肤色图像块作为正样本, 并在不含人体肤色区域的图像集中均匀采集了 1,452,528 个 8×8 的非肤色图像块作为负样本. 实验中采用其中 575,972 个肤色图像块和 803,061 个非肤色图像块作为训练集, 其余的作为测试集.

实验中, 我们将本文算法与文献[10]中提出的 SPM 算法进行了对比. SPM 算法已被证明是像素域基于颜色的肤色检测算法中效果较好的一种, 并且具有极低的计算复杂度. 在对比实验中, SPM 算法是在解压后的 RGB 颜色空间采用 32 级量化级数建立颜色直方图的基础上实现的. 本文算法中用于区域生长的各阈值根据经验设置为: $TH_S_1 = 10$, $TH_C_1 = 12$, $TH_C_2 = 12$, $TH_S_{min} = 6$, $TH_DS = 5$.

本文采用文献[10]中 ROC(Receiver Operating Characteristic)曲线形式衡量算法性能, 如图 5 所示. 其中正

检率指已标记为肤色的样本被正确识别的比例,误检率指已标记为非肤色的样本被误识为肤色的比例。

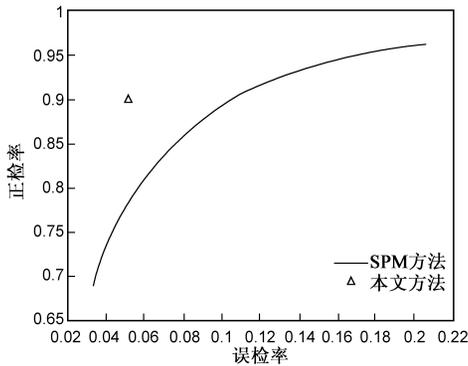


图5 SPM算法与本文算法在测试数据集上的ROC曲线对比

由于本文算法仅产生 1 组正检率和误检率,因而在图 5 中用点表示.从图 5 中可以看出本文算法与 SPM 算法相比能够在正检率和误检率两个指标下均获得较好的性能.这主要是因为 SPM 算法仅利用了像素的颜色信息,而本文算法综合利用了图像块的颜色和纹理信息.

为了公平衡量两种算法的检测速度,本文算法及 SPM 算法中涉及到的 JPEG 压缩图像的解码操作均采用 JPEGLib 库,该库可在 <http://www.ijg.org> 下载.SPM 算法在 IDCT 阶段采用了快速的 AA&N 算法.两种算法检测的平均时间对比如表 2 所示.

表 2 本文算法与 SPM 算法的检测速度对比实验结果

比较算法	检测时间(s)
SPM 算法	0.227
本文算法	0.106

注:PC 配置为奔腾双核 3.0GHZ CPU, Microsoft Windows XP 操作系统

从表 2 中可以看出,与 SPM 算法相比,本文算法的检测速度提高了 2 倍以上.分析其原因,主要在于:(1)本文提出的算法直接在熵解码后的 DCT 系数中提取颜色特征和纹理特征,不需要进行 DCT 反变换,因此节约了大量的图像解码时间;(2)本文提出的算法以 8×8 图像块为基本的检测单元,而像素域的 SPM 算法以像素为基本的检测单元,因而本文算法需要的检测次数是 SPM 算法的 $1/64$.

图 6 所示的是分别采用 SPM 算法和本文算法得到的部分图像的肤色检测结果.对于 SPM 算法分别采用了两个阈值进行肤色检测,其中阈值取 1.05 时,肤色正检率与本文方法相近,而阈值取 4.1 时误检率与本文方法相近.从 6 图中可以看出,对于与肤色接近的背景区域,采用 SPM 算法容易造成误检,而本文算法则会较好地消除背景干扰.但由于本文算法以 8×8 图像块为基本的检测单元,并采用了纹理特征,这导致纹理较丰富的肤色区域会产生漏检,例如图 5 中第 2 幅图像的手部

肤色区域.



(a) 原始图像 (b) SPM算法阈值取1.05 (c) SPM算法阈值取4.1 (d) 本文算法

图6 部分肤色检测结果

4 结论

本文首先在 JPEG 压缩码流中提取了颜色、纹理等多种特征,然后利用数据挖掘的方法建立了决策树肤色模型,并从中发现了与肤色检测密切相关的 5 个特征.在肤色检测过程中,首先利用决策树肤色模型进行初步检测,然后利用区域生长的方法最终检测出图像中的肤色区域.实验结果表明,与基于像素域的 SPM 算法相比,本文提出的压缩域肤色检测算法可以提高肤色检测的准确率和速度.

参考文献:

- [1] P Kakumanu, S Makrogiannis, N Bourbakis. A survey of skin-color modeling and detection methods[J]. Pattern Recognition, 2007, 40(3): 1106 - 1122.
- [2] Z Jiang, M Yao, W Jiang. Skin detection using color, texture and space information[A]. Proc of the 4th International Conference on Fuzzy System and Knowledge Discovery[C]. Haikou, China: FSKD, 2007. 366 - 370.
- [3] Q F Zheng, W Gao. Fast adaptive skin detection in JPEG images[A]. Advances in Multimedia Information Processing[C]. Berlin/Heidelberg: Springer, 2005. 595 - 605.
- [4] 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检测算法[J]. 电子学报, 2005, 33(12A): 2464 - 2467.
X G Li, X H Li, L S Shen. 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)
- [5] R L D Queiroz, R Eschbach. Fast segmentation of the JPEG compressed documents [J]. Journal of Electronic Imaging, 1998, 7(2): 367 - 377.

- [6] H J Bae, S H Jung. Image retrieval using texture based on DCT [A]. Proc of Information, Communications and Signal Processing [C]. Singapore: IEEE, 1997. 1065 – 1068.
- [7] Y L Huang, R F Chang. Texture features for DCT coded image retrieval and classification [A]. Proc of IEEE International Conference on Acoustic, Speech, and Signal Processing [C]. Phoenix, America, 1999. 15 – 19.
- [8] Microsoft. Data Mining Concepts [OL]. [http://msdn.microsoft.com/en-us/library/ms174949\(SQL.90\).aspx](http://msdn.microsoft.com/en-us/library/ms174949(SQL.90).aspx), 2008 – 11.
- [9] L Gu, D Bone. Skin colour region detection in MPEG video sequences [A]. Proc of International Conference on Image Analysis and Processing [C]. Venice, Italy: IEEE Computer Society, 1999. 27 – 29.
- [10] M J JONES, J M REHG. Statical color models with application to skin detection [J]. International Journal of Computer Vision, 2002, 46(1): 81 – 96.

作者简介:



赵士伟 男, 1980 年出生, 山东高唐人. 博士研究生. 主要研究方向为网络多媒体信息处理. E-mail: zhao_shiwei@emails.bjut.edu.cn

卓力 女, 1971 年出生, 江苏徐州人. 博士, 北京工业大学教授. 主要研究方向为图像/视频编码与传输, 网络多媒体分析, 无线视频传感器网络.

王素玉 女, 1976 年出生, 河北丰润人. 讲师, 在站博士后. 主要研究方向为图像/视频信号压缩编码和超分辨率复原, 智能视觉监控等.

沈兰荪 男, 1938 年出生, 江苏苏州人. 教授, 博士生导师. 主要研究方向为图像/视频信号处理、传输、压缩与应用.