

JPEG 图像压缩域上的自适应人体皮肤区域检测算法

郑清芳^{1,2} 高文^{1,3}

¹(中国科学院计算技术研究所 北京 100080)

²(中国科学院研究生院 北京 100049)

³(北京大学信息科学技术学院 北京 100871)

(qfzheng@jdl.ac.cn)

Adaptive Skin Detection in JPEG Compressed Images

Zheng Qingfang^{1,2} and Gao Wen^{1,3}

¹(*Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080*)

²(*Graduate University of Chinese Academy of Sciences, Beijing 100049*)

³(*School of Electronics Engineering and Computer Science, Peking University, Beijing 100871*)

Abstract Skin region detection plays an important role in a variety of applications such as face detection and adult image filtering. To improve the accuracy and speed of skin detection, an adaptive skin detection algorithm working on compressed domain of JPEG images is proposed. Two main contributions of the skin detector are: ① It can adaptively control the threshold according to image content; ② It doesn't require full decompression of JPEG images. It directly derives color and texture features of each image block from DCT coefficients and jointly uses both features to classify each image block as human skin region or not; Comparisons with other existing skin detection techniques demonstrate that the algorithm proposed can compute very fast and achieve good accuracy.

Key words JPEG; skin detection; DCT; compress; decompression

摘要 快速而准确地检测图像中的人体皮肤区域在人脸检测、敏感图像过滤等应用中有重要价值。为了提高肤色检测的精度和速度,提出了一种在 JPEG 图像压缩域上的基于自适应阈值的人体皮肤区域检测算法。该算法的优点在于:①检测过程中能根据图像内容自适应地调节检测阈值,有效防止误检和漏检;②无需完全解压缩 JPEG 图像,直接在 DCT 系数域上计算每个图像块的颜色和纹理特征,利用颜色和纹理特征来检测图像块是否为人体皮肤。与现有肤色检测算法的对比实验表明我们的方法具有很好的准确性和很快的速度。

关键词 JPEG; 肤色检测; DCT; 压缩; 解压缩

中图法分类号 TP391.4

1 引言

皮肤是人体最显著的生理特征之一,快速而准确地检测图像中的人体皮肤区域在人脸检测、敏感

图像过滤和手势识别等应用中有重要的理论和实用价值。研究者已经提出了许多肤色检测算法,如参考文献[1~9],这些方法可以概括为基于像素的静态检测方法:首先利用训练数据离线训练一个肤色分布模型,然后对待检测图像中的每个像素进行检

查,如果某像素的特征满足肤色分布模型,则标记该像素为皮肤像素,否则标记该像素为非皮肤像素。

基于像素的静态检测方法取得了某种程度的成功,但仍有其缺陷。人体肤色在不同的人种之间存在差异,而且会随着光照的变化而变化,静态的肤色模型对这些情况缺乏足够的泛化能力,会导致漏检或误检。作为对静态肤色检测方法的改进,有两种思路值得探讨:一种是基于针对不同图像的不同肤色模型的分而治之策略^[10];另一种则是自适应的检测方法^[11~13]。基于分而治之策略的方法需要解决如何将不同图像分类的问题。自适应的检测方法通过对每幅待检测的图像都进行迭代操作,以寻找最适合于该图像的肤色模型:文献[11]的方法在肤色分数图(skin score map)上迭代分割直到每个区域都具有光滑的纹理;文献[12]的方法把肤色的分类阈值从高到低递减,并用每个域值进行肤色检测,当检测到肤色面积出现激增时,认为此时的阈值是最优阈值;文献[13]的方法首先迭代更改肤色在 HSV 颜色空间上 S 和 V 分量的取值范围,并用每次更改后的取值范围进行检测。这些自适应肤色检测方法在提高肤色检测精度的同时也带来了另一个问题——检测速度的降低。这种“以速度换精度”的做法在某些应用中是不可取的。例如,在敏感图像过滤系统中^[6,10],肤色检测往往只作为前期处理步骤,肤色检测的速度很大程度上影响了整个系统总体速度。

目前互联网上大部分的图像都是采用 JPEG 的压缩标准,对这些图像,现有的皮肤检测方法首先把图像完全解压缩,然后对每个像素逐个地进行判别。完全解压缩 JPEG 图像需要很大的计算量,再加上自适应检测算法中的迭代操作,给自适应皮肤检测算法带来了很大的计算负担。本文针对 JPEG 图像,提出了一种在压缩域上的人体皮肤区域自适应检测算法。该算法不需要完全解码 JPEG 图像,而是直接在 DCT 系数域计算每个图像块的颜色和纹理特征,并同时利用颜色和纹理特征来判断每个图像块是否为人体肤色区域。本文所提出的算法的优点在于:

1) 本算法不需要把图像完全解压缩,直接在压缩码流上操作,简化了操作步骤,提高了人体皮肤区域检测的速度。

2) 本算法充分考虑了人体皮肤区域的颜色和纹理信息,在检测过程中根据图像内容自适应调整检测阈值,提高了检测的精度。

2 皮肤的贝叶斯分类器

本节首先介绍考虑皮肤颜色和纹理特征的贝叶斯模型,然后介绍模型中的阈值的自适应选择机制。

2.1 贝叶斯模型

针对 JPEG 图像的文件格式,本文提出的人体皮肤区域检测算法以图像块为基本处理单位,并且用 YCbCr 颜色空间表示图像块的颜色,使检测算法能与 JPEG 压缩标准相一致,尽量减少计算量。

根据贝叶斯规则,一个颜色特征为 $color$,纹理特征为 $texture$ 的图像块,其属于人体皮肤的后验概率为

$$P(\text{skin} | \text{color}, \text{texture}) = P(\text{color}, \text{texture} | \text{skin}) \times \frac{P(\text{skin})}{P(\text{color}, \text{texture})}, \quad (1)$$

其中 $P(\text{color}, \text{texture})$ 是颜色和纹理的联合分布概率。为了在计算中消去 $P(\text{color}, \text{texture})$,我们同时计算该图像块属于非人体皮肤的后验概率:

$$P(\text{skin} | \text{color}, \text{texture}) = P(\text{color}, \text{texture} | \text{skin}) \times \frac{P(\text{skin})}{P(\text{color}, \text{texture})}, \quad (2)$$

则两个后验概率的比值为

$$\frac{P(\text{skin} | \text{color}, \text{texture})}{P(\text{skin} | \text{color}, \text{texture})} = \frac{P(\text{color}, \text{texture} | \text{skin})}{P(\text{color}, \text{texture} | \text{skin})} \times \frac{P(\text{skin})}{P(\text{skin})}, \quad (3)$$

当后验概率的比值大于某一阈值时,标记该图像块为人体皮肤。由于以上公式中, $P(\text{skin})$ 和 $P(\text{skin})$ 都是类概率,是未知的常数,其比值也为某一常数,所以我们把该常数折算在阈值内,则我们的皮肤分类器为

$$P(\text{color}, \text{texture}) = \frac{P(\text{color}, \text{texture} | \text{skin})}{P(\text{color}, \text{texture} | \text{skin})} > \tau. \quad (4)$$

2.2 阈值的自适应选取

通常的肤色检测方法采用固定的分类阈值 τ ,但是阈值 τ 如果选得太小,会导致很多非皮肤图像块被误认为是皮肤;相反地,如果阈值 τ 选得太大,会导致皮肤图像块被误认为是非皮肤。我们注意到在图像中人体皮肤区域具有光滑纹理的特性,皮肤周围的背景尽管可能在颜色上和皮肤相类似,但是

如果把这些背景也当做皮肤,那么整个皮肤区域在纹理上会不一致.根据这一现象,为了使我们的分类器能够适应不同的图像,我们采用自适应阈值的方法:首先阈值 τ 取一个较小的值 τ_0 ,利用这个阈值检测出图像中的候选皮肤块,然后逐渐增大阈值,不断剔除其中的噪声,直到每个检测到皮肤区域光滑为止.

3 压缩域的颜色和纹理特征计算方法

每个图像块的颜色用该图像块内的所有像素的平均颜色来表示: $color_{block} = [y_{block}, cb_{block}, cr_{block}]$.人体皮肤的一个重要特性就是它具有光滑的纹理,我们用方差来刻画纹理的光滑性,方差越小,纹理越光滑.由于人眼对彩色细节的分辨能力远比对亮度细节的分辨能力低,我们只考虑了亮度分量 Y 上的方差,记图像块的纹理为 $texture_{block} = \sigma_{block}^2$,记区域的纹理为 $texture_{region} = \sigma_{region}^2$.考虑一个 $N \times N$ 大小的图像块,其中的某颜色分量上各个像素的值为 $f(i, j)$ ($i=0, \dots, N-1; j=0, \dots, N-1$),该图像块对应的DCT系数为 $F(u, v)$ ($u=0, \dots, N-1; v=0, \dots, N-1$).2维离散余弦变换定义如下:

$$F(u, v) = \frac{2}{N} \alpha(u) \alpha(v) \left[\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \times \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N} \right], \quad (5)$$

它的逆变换为

$$f(i, j) = \frac{2}{N} \alpha(u) \alpha(v) \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \times \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N}. \quad (6)$$

以上两式中:

$$\begin{cases} \alpha(u), \alpha(v) = \frac{1}{\sqrt{2}}, & \text{当 } u, v = 0, \\ \alpha(u), \alpha(v) = 1, & \text{其他.} \end{cases}$$

该图像块中各像素在该颜色分量上的强度值平均为

$$\begin{aligned} \mu_{block} &= \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) = \\ &= \frac{2}{N^3} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u) \alpha(v) F(u, v) \times \\ &= \sum_{i=0}^{N-1} \cos \frac{(2i+1)u\pi}{2N} \sum_{j=0}^{N-1} \cos \frac{(2j+1)v\pi}{2N}. \quad (7) \end{aligned}$$

由于

$$\sum_{i=0}^{N-1} \cos \frac{(2i+1)u\pi}{2N} = \begin{cases} N, & \text{当 } u = 0, \\ 0, & \text{其他,} \end{cases}$$

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

同时,根据Parseval恒等式,可以得到

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [f(i, j)]^2 = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [F(u, v)]^2. \quad (9)$$

图像块的纹理可以计算为

$$\begin{aligned} texture_{block} &= \sigma_{block}^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [f(i, j)]^2 - \\ &= \left[\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f^2(i, j) \right]^2 = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [F^2(u, v)]^2 - \\ &= \left[\frac{1}{N} F^2(0, 0) \right]^2 = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [F^2(u, v)]^2 \quad (10) \end{aligned}$$

其中 $(u, v) \neq (0, 0)$.

对于一个由 m 个图像块构成的图像区域,其纹理可以计算为

$$\begin{aligned} texture_{region} &= \sigma_{region}^2 = \frac{1}{mN^2} \sum_{k=1}^m \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [f_k(i, j)]^2 - \\ &= \left[\frac{1}{mN^2} \sum_{k=1}^m \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f_k(i, j) \right]^2 = \frac{1}{mN^2} \times \\ &= \sum_{k=1}^m \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [f_k(u, v)]^2 - \left[\frac{1}{mN} \sum_{k=1}^m F_k(0, 0) \right]^2 = \\ &= \frac{1}{m} \left\{ \sum_{k=1}^m texture_{block} + \frac{1}{N^2} \sum_{k=1}^m [F_k(0, 0)]^2 \right\} - \\ &= \left[\frac{1}{mN} \sum_{k=1}^m F_k(0, 0) \right]^2. \quad (11) \end{aligned}$$

本文提出的算法适用的图像块大小为 8×8 或者 4×4 或者 2×2 .在JPEG压缩标准中,图像块的大小为 8×8 .我们可以把1个 8×8 的图像块分解为4个 4×4 或者16个 2×2 的小图像块,具体的分解算法见参考文献[14].

4 检测算法

综合以上两节,本文所提出的算法的流程如图1所示并描述如下:

1) 对待检测的JPEG图像进行哈夫曼解码和反量化,得到 Y, Cb, Cr 颜色分量上各个图像块对应的DCT系数.

2) 直接从图像块的DCT系数计算颜色和纹理特征.

3) 检测出图像中的候选人体皮肤块,标记所有满足式(4)的图像块为候选皮肤块.在本文实验中,初始阈值设为 $\tau = 1.0$, $F(color, texture)$ 根据直方图的方法^[4]计算.

4) 图像中相邻的候选肤色块分割成候选皮肤

区域.

5) 按式(11)计算每个候选皮肤区域的纹理 $texture_{region}$, 若 $texture_{region} < \tau_{region}$, 说明该区域光滑, 符合人体皮肤区域的纹理特点, 则标记该候选区域为真正的人体皮肤区域.

6) 若图像中已无候选皮肤区域, 算法结束; 否则, 转 7).

7) 对图像中未标记为真正人体皮肤区域的每个皮肤候选区域, 由于其 $texture_{region} > \tau_{region}$, 说明该区域中含有噪声. 我们通过提高该区域的分类阈值: $\tau = 1.05 \times \tau$, 将该区域中后验概率比值小于更新后阈值的候选人体皮肤块标记为非候选人体皮肤块, 转 4).

本算法综合利用了人体皮肤区域的颜色和纹理特征, 通过动态调整每个候选皮肤区域的分类阈值, 不断剔除候选区域中的干扰图像块, 确保最终得到的区域符合人体皮肤区域特性, 提高皮肤区域检测的精度.

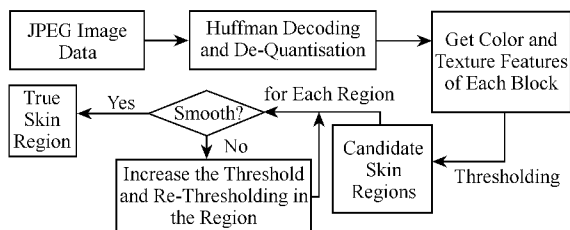


Fig. 1 Schematic description of our skin detection algorithm.

图1 人体皮肤区域检测算法流程图

5 计算量差异比较

本节讨论本文算法与 Phung 等人在文献[11]中所提的方法在计算量上的差异. Phung 的方法首先需要把 JPEG 图像完全解压缩到像素域, 根据每个像素的颜色特征计算其肤色分数 (skin score) 并生成与原图像相对应的肤色分数图 (skin score map), 然后在肤色分数图上重复分割直到每个区域都具有光滑的纹理, 在分割过程中需要对每个像素进行判断. 与 Phung 的方法相比, 本文算法在计算量上明显减少:

1) Phung 的方法需要完全解压缩图像, 对每个 8×8 的图像块, 完全解压缩中的反 DCT 变换需要 4032 次加法运算和 4096 次乘法运算; 而本文算法不需要对图像块进行完全解压缩, 省略了解压缩过

程中反 DCT 变换.

2) Phung 的方法以像素为基本处理单位, 每个像素判断一次, 本文算法以图像块为基本处理单位, 每个 $N \times N$ 的图像块 (N^2 个像素) 只需判断一次, 大大减少了判断次数.

3) 本文算法直接在 DCT 系数上计算颜色和纹理特征. 以 $N \times N$ 的图像块为例, 本算法计算颜色特征只需 3 次乘法运算和 1 次加法运算, 计算纹理特征只需 N^2 次加法和 N^2 次乘法运算; 在已经计算完图像块纹理的基础上, 计算由 m 个图像块组成的区域的纹理特征只需 $m + 5$ 次乘法和 $3m + 1$ 次加法运算; Phung 的方法在计算区域纹理特征时除了需要计算方差还需进行边缘检测, 统计区域内的边缘点个数, 所需计算量大大高于本文算法.

6 实验

为了评测本文提出的算法的精度和速度, 我们设计了两个实验, 与文献[11]中的自适应检测算法和文献[4]中的静态检测算法进行了对比.

6.1 实验 1

在实验 1 中, 我们准备了一个具有 3300 幅含有人体皮肤区域的图像数据集. 该数据集包括 3000 幅来自 ECU 人脸数据库的人脸图像^[11]和 300 幅从网络上收集的成人图像. 所有的图像都经手工标注为人体皮肤区域和非人体皮肤区域. 本文使用检出率 (TP) 和误检率 (FP) 这两个指标来衡量各种算法的精度, 某一算法的检出率越高, 误检率越低, 则说明该算法的精度越好. 检出率和误检率的定义如下:

$$TP = \frac{\text{正确检测到的皮肤像素数}}{\text{真实的皮肤像素总数}} \times 100\%$$

$$FP = \frac{\text{错误检测到的皮肤像素数}}{\text{真实的非皮肤像素总数}} \times 100\%$$

本次实验随机选取了数据集中的 2200 幅图像作为训练数据, 把另外的 1100 幅图像作为测试数据. 测试数据共包括 53412219 个人体皮肤像素和 235305733 个非人体皮肤像素. 对于本文所提出的压缩域自适应检测方法, 我们考察了 3 种不同的图像块大小对检测性能的影响, 本文用 JPEG-8, JPEG-4 和 JPEG-2 分别指代选择以 8×8 , 4×4 和 2×2 大小的图像块作为基本处理单位的压缩域检测方法. 对于文献[4]中的静态方法, 一般采用 ROC 曲线衡量其性能, 为了便于与自适应检测方法相比较,

我们在其 ROC 曲线上分别选择两个特殊点, 其中一个点代表检出率与自适应方法相似时的性能 (Static-2), 另外一点代表误检率与自适应方法相似时的性能 (Static-1). 为了评测各种算法的速度, 本次实验统计了各种算法检测测试集中 1100 幅图像总共所费的时间, 以毫秒为单位. 本次实验在一台配置 Pentium IV 1GHz CPU, 运行 Microsoft Windows 2000 操作系统的 PC 上完成. 表 1 列出了各种算法在测试数据集上的性能对比结果.

Table 1 Performance Comparison Among Different Skin Detection Algorithms

表 1 各种皮肤检测算法的性能对比结果

Method	TP (%)	FP (%)	Time (ms)	Speed (fps)
JPEG-8	77.12	4.12	64140	17.15
JPEG-4	85.24	4.60	85774	12.82
JPEG-2	84.34	4.33	165669	6.64
Static-1 ^[4]	82.27	4.58	166667	6.60
Static-2 ^[4]	85.29	6.60	166163	6.62
Reference [11]	85.09	5.74	472110	2.33

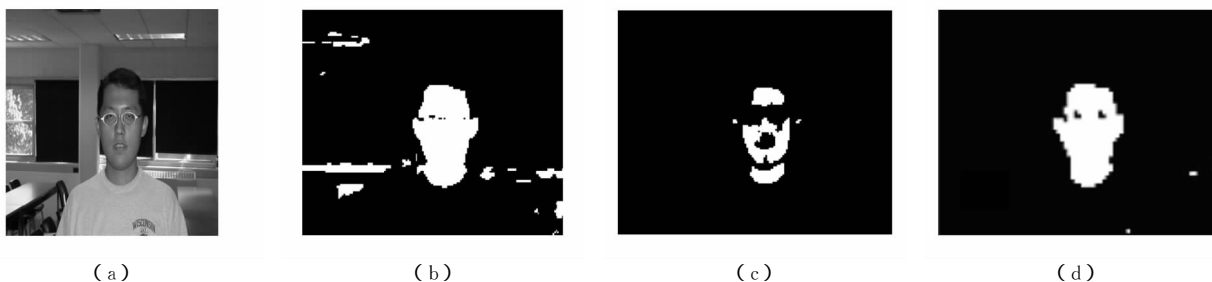


Fig. 2 Detection results of JPEG-4 and non-adaptive skin detection method^[4]. (a) Original image; (b) Detection result when threshold is too low^[4]; (c) Detection result when threshold is too high^[4]; and (d) Detection result by JPEG-4.

图 2 以 4×4 的图像块为基本处理单位时的自适应阈值和固定阈值的检测结果对比图. (a) 原始图像 (b) 阈值过低时的结果 (c) 阈值过高时的结果 (d) 自适应阈值的结果

6.2 实验 2

在实验 2 中, 我们准备了一个包括 100 幅图像的数据集, 该数据集不含有皮肤区域, 而是包含狮子、食物、文字等物体. 这些物体尽管在颜色上与人体皮肤区域相似, 但在纹理特性上与人体皮肤相差很大. 限于篇幅, 本文仅在图 3 中例示 4 幅检测结果对比图. 图 3 中的第 1 行是原始图像; 第 2 行是文献 [11] 的自适应检测方法的检测结果; 第 3 行是 JPEG-4 的检测结果. 图像中白色部分为人体皮肤区域. 实验结果表明, 本文提出的综合利用图像块颜色和纹理特征的皮肤检测方法明显优于基于像素颜色特征的方法.

从表 1 可以看出, 对于压缩域的自适应检测方法 JPEG-8 的检测速度最快, 但检出率不高, 这主要是因为选择以 8×8 大小的图像块作为基本处理单位的做法过于粗放, 忽略了图像细节; JPEG-4 和 JPEG-2 的检测精度相近, 这说明选择以 4×4 的图像块作为基本处理单位, 可以有效地反映图像细节, 而且 JPEG-4 比 JPEG-2 节省了计算量, 检测速度提高了将近 2 倍. JPEG-4 与文献 [4] 中的静态检测方法相比不但检测速度加快, 而且检测精度也有提高: 在检出率相当的情况下, 误检率降低 2%; 在误检率相当的情况下, 检出率上升 2.07%. JPEG-4 与文献 [11] 的自适应方法相比, 不但精度有所提高, 而且检测速度快了将近 6 倍! 图 2 例示了以 4×4 的图像块为基本处理单元时自适应阈值和固定阈值的检测结果对比图. 从图 2 中可以看出: 当阈值过低时, 检测结果中存在误检; 当阈值过大时, 检测结果中存在漏检; 自适应的方法根据图像内容自动选择阈值, 检测结果令人满意.

7 结论及未来研究工作

本文主要解决了静态图像中人体皮肤区域检测算法的精度和速度问题. 为了提高人体皮肤区域检测的速度, 我们提出了直接在 JPEG 图像压缩域进行检测的算法, 该算法无需完全解压缩图像, 而是直接在 DCT 系数上计算图像块的纹理和颜色信息. 为了提高人体皮肤区域检测的精度, 我们的算法综合人体皮肤区域的颜色和纹理特性, 利用纹理和颜色信息来判断图像块是否为人体皮肤区域. 根据候选肤色区域的纹理性质自适应调整分类阈值, 直至

检测出具有光滑纹理的人体皮肤区域. 本文同时考察了不同的图像块大小对肤色检测结果的影响. 在未来的研究工作中,我们将把本文所提出的人体皮

肤区域检测算法应用到敏感图像检测中,构建一个鲁棒且高效的敏感图像过滤系统.

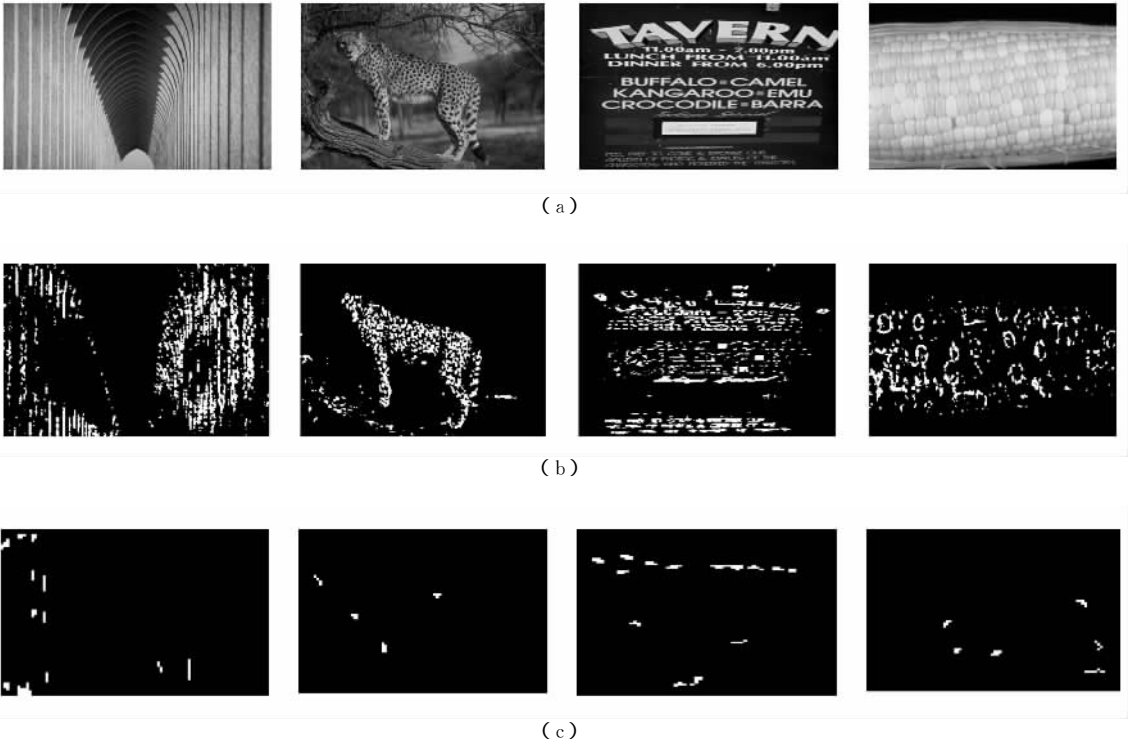


Fig. 3 Skin detection results comparison between JPEG-4 and referenc[11]. (a) Original images ;(b) Results by referenc[11]; and (c) The results by JPEG-4.

图 3 JPEG-4 与文献 [11] 中的自适应方法在非人体图像中的检测结果对比。(a) 原始图像 (b) 文献 [11] 的方法的检测结果 (c) JPEG-4 的检测结果

参 考 文 献

- 1 Tao Linmi , Peng Zhenyun , Xu Guangyou. Features of skin color [J]. Journal of Software. 2001, 12(7): 1032 ~ 1041 (in Chinese)
(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)
- 2 Yao Hongxun, Liu Mingbao, Gao Wen, et al. Method of face locating and tracking based on chromatic coordinates transformation of color images[J]. Chinese Journal of Computers, 2000, 23(2): 158~165 (in Chinese)
(姚鸿勋,刘明宝,高文,等. 基于彩色图像的色系坐标变换的面部定位于跟踪法[J]. 计算机学报, 2000, 23(2): 158~165)
- 3 R. L. Hsu, M. Abdel-Mottaleb, A. K. Jain. Face detection in color images [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2002, 24(5): 696~707
- 4 M. J. Jones, J. M. Rehg. Statistical color model with application to skin detection [J]. International Journal of Computer Vision, 2002, 46(1): 81~96
- 5 H. Greenspan, J. Goldberger, I. Eshet. Mixture model for face color modeling and segmentation [J]. Pattern Recognition Letters, 2001, 22(14): 1525~1536
- 6 M. Fleck, D. Forsyth, C. Bregler. Finding naked people [C]. In: Proc. 4th European Conf. Computer Vision. Berlin: Springer, 1996. 592~602
- 7 S. L. Phung, A. Bouzerdoum, D. Chai. A novel skin color model in YCbCr color space and its application to human face detection [C]. In: Int'l Conf. Image Processing. Los Alamitos: IEEE Computer Society Press, 2002. 289~292
- 8 S. L. Phung, A. Bouzerdoum, D. Chai. Skin segmentation using color pixel classification: Analysis and comparison [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2005, 27(1): 148~154
- 9 J. Yang, T. Tan, W. Hu. Skin color detection using multiple cues [C]. In: Int'l Conf. Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2004. 632~635
- 10 Zeng Wei, Zheng Qingfang, Zhao Debin. Image Guard: An automatic adult image recognition system [J]. High Technology Letters, 2005, 15(3): 11~16 (in Chinese)
(曾炜,郑清芳,赵德斌. 图像卫士: 一个自动成人图像识别系统[J]. 高技术通讯, 2005, 15(3): 11~16)

- 11 S. L. Phung, D. Chai, A. Bouzerdoum. Adaptive skin segmentation in color images [C]. In: IEEE Int'l Conf. Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2003. 353~356
- 12 H. T. Quan, M. Meguro, M. Kaneko. Skin-color extraction in images with complex background and varying illumination [C]. In: Proc. 6th IEEE Workshop on Applications of Computer Vision. Los Alamitos: IEEE Computer Society Press, 2002. 280~285
- 13 K. M. Cho, J. H. Jang, K. S. Hong. Adaptive skin-color filter [J]. Pattern Recognition, 2001, 34(5): 1067~1073
- 14 J. Jiang, A. J. Armstrong, G. C. Feng. Direct content access and extraction from JPEG compressed images [J]. Pattern Recognition, 2002, 35(11): 2511~2519



Zheng Qingfang, born in 1979. Currently Ph. D. candidate in the Institute of Computing Technology, the Chinese Academy of Sciences. His main research interests include multimedia content

analysis, pattern recognition and information retrieval.

郑清芳, 1979年生, 博士研究生, 主要研究方向为多媒体内容分析、模式识别和信息检索。



Gao Wen, born in 1956. Received his Ph. D. degrees from the Harbin Institute of Technology in 1988 and University of Tokyo in 1991, respectively. Now he is professor and Ph. D. Supervisor in the Institute of Computing Technology, the Chinese Academy of Sciences, senior member of CCF. His main research interests include multimedia data compression, image processing, computer vision, multimodal interface, artificial intelligence and virtual reality.

高文, 1956年生, 博士, 教授, 博士生导师, 中国计算机学会高级会员, 主要研究方向为多媒体数据压缩、图像处理、计算机视觉、多模式接口、人工智能和虚拟现实(wgao@jdl.ac.cn).

Research Background

This work is part of the Researches on Key Technologies for Web Image Censorship Project, which is supported by the Chinese National Hi-Tech R&D Program (the 863 Program) under contract No. 2003AA142140. The Internet has currently become an unprecedented large image library, where innocuous and offensive images coexist. To make sure that offensive images are not accessed by the public, especially teenage, web images have to be censored. The project includes three sub-projects: adult image filtering, text detection in images, image matching.

In the adult image filtering sub-project, we are developing image content-based techniques to detect and block pornographic images. Main challenges here are to detect human skin region under different illumination conditions, to select the effective visual features that can discriminate pornographic images from innocuous ones and to design the optimal classifier. In text detection sub-project, we are designing algorithms to automatically extract and read the text components embedded in the images, since these texts can provide crucial cues about the image content. In image matching, we focus our effort on matching a user selected image region among different images to find the most similar region. We are developing effective and efficient local image descriptors that are illumination-invariant, scale-invariant and rotation-invariant.