

基于肤色和模板的人脸检测*

艾海舟, 梁路宏, 徐光祐, 张钹

(清华大学 计算机科学与技术系, 北京 100084);

(清华大学 智能技术与系统国家重点实验室, 北京 100084)

E-mail: ahz@mail.tsinghua.edu.cn

http://www.tsinghua.edu.cn

摘要: 针对彩色图像提出了一种基于肤色和模板的人脸检测方法, 由肤色分割、模板匹配和人工神经网络验证3部分组成。首先使用HSI空间的肤色统计模型分割出可能包含人脸的区域, 然后使用平均脸模板匹配和人工神经网络验证的方法在这些区域中搜索人脸。该方法将彩色图像的肤色信息和灰度图像的模板匹配及人工神经网络分类模型综合起来, 既极大地提高了速度, 又具有较强的鲁棒性。实验结果表明, 该算法是快速而有效的。

关键词: 人脸检测; 肤色分割; 模板匹配; 人工神经网络

中图法分类号: TP391 **文献标识码:** A

人脸检测作为人脸信息处理的一个重要内容, 近年来已经成为计算机视觉与模式识别领域的研究热点之一。人脸作为一种自然形体具有很强的共性, 但由于个体的外貌差异、表情变化以及图像在采集时受光照、设备等因素的影响, 使得人脸图像具有比较复杂而细致的模式变化。因此, 人脸检测是一个极具挑战性的模式分类问题, 对这个问题的研究具有十分重要的学术价值。而且人脸检测与许多实际问题紧密相关, 在出入安全检查、视觉监视、智能人机接口、基于内容的检索、数字视频处理等诸多领域有着极为广泛的应用价值。

目前, 在人脸检测方面的研究工作主要有两类方法。一类是针对彩色图像的基于肤色的方法^[1,2], 根据人脸的色彩信息和启发式信息定位人脸区域。其优点是速度快且对姿态不敏感, 缺点是难以区分具有类似肤色的非人脸区域, 只适用于特定环境。实际上, 这类方法也不能对人脸进行精确定位。另一类是仅仅根据灰度信息的模板方法, 包括基于特征脸的方法^[3,4]、基于人工神经网络的方法^[5,6]和基于支持矢量机(support vector machines, 简称SVM)的方法^[7]等。基于模板的方法具有一般性, 有较强的鲁棒性, 但缺点是计算量大、速度慢。很显然, 针对彩色图像, 如果能将上述两种方法结合起来就可以获得较好的鲁棒性和较高的速度。目前这方面深入的工作还不多见, 已有的工作有基于肤色分割和小波特征验证的方法^[8], 计算量也很大。

本文针对一般彩色图像中的正面端正人脸(确切的条件是眼睛基本处于水平位置)的检测问题, 提出了一种将基于肤色的区域分割与基于模板的人脸检测综合起来的算法。首先利用颜色信息分割出可能包含人脸的区域, 然后在各个区域中使用平均脸模板匹配与人工神经网络验证相结合的方法搜索并精确定位人脸。该算法以大量样本的统计训练为依据, 具有较强的鲁棒性。在实验测试

* 收稿日期: 2000-05-09; 修改日期: 2000-07-06

基金项目: 国家863高科技发展计划资助项目(863-805-512-9805-11)

作者简介: 艾海舟(1964-), 男, 江苏人, 博士, 副教授, 主要研究领域为计算机视觉, 模式识别, 人脸信息处理; 梁路宏(1973-), 男, 河北人, 博士, 主要研究领域为人脸信息处理, 模式识别; 徐光祐(1940-), 男, 上海人, 教授, 博士生导师, 主要研究领域为计算机视觉, 人机交互, 多媒体技术; 张钹(1935-), 男, 福建人, 教授, 博士生导师, 中国科学院院士, 主要研究领域为人工智能理论及应用。

集上,该算法对于简单图片(如标准照片等)的检测正确率达到 99%,对于摄像机获取的室内外中等复杂程度图片的检测正确率达到了 95.1%,对于具有更为复杂的光照条件和背景的扫描照片和 Web 上的图片,其检测正确率也达到了 82.7%.

1 人脸检测算法的总体框架

我们提出的算法按照“分割-搜索”的检测模式,将基于肤色的区域分割与基于模板的人脸检测综合起来,算法的流程如图 1 所示,由两个步骤构成:

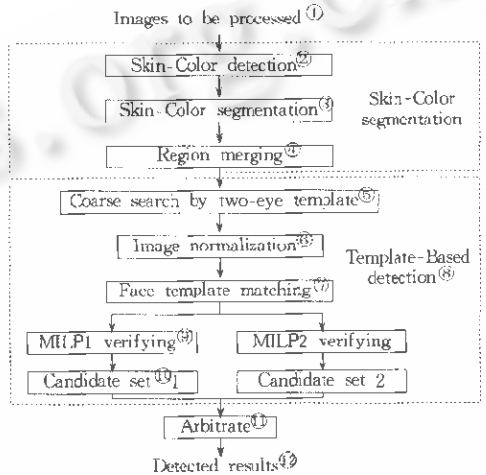
(1) 基于肤色模型的区域分割

为了尽可能精确地分割出肤色区域,在 HSI (hue-saturation-intensity) 颜色空间,我们没有像现有大多数方法那样不考虑亮度分量或仅将色彩相对于亮度标准化后采用二维肤色模型^[1],而是将亮度分量与两个色彩分量一起,根据统计建立了一个更精确的二维模型.在此基础上,根据同一图片中人脸颜色与非人体背景的色彩具有相对差异性以及具体图片中人脸的颜色具有一致性的事实,提出了基于区域颜色一致性准则的快速分割算法.

(2) 基于模板的检测

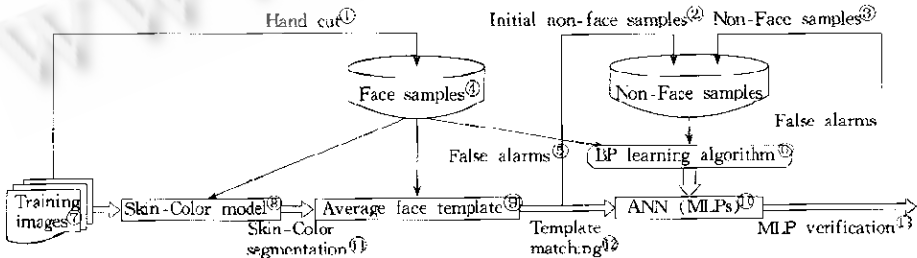
在分割出的肤色区域中,使用平均双眼和平均人脸模板匹配加人工神经网络验证的方法,在一定尺度范围内进行穷举搜索,精确定位出人脸,其中为了减少错误报警,使用了双神经网络仲裁方法.

算法以大量样本为统计数据,分别建立肤色模型、构造平均脸模板和训练人工神经网络.具体而言,如图 2 所示,首先建立了训练用图像库,用手工的方法剪裁出人脸的肤色区域,用于建立肤色模型;用手工标定的方法切割出人脸的五官区域,转换为灰度图像后进行尺度和灰度分布标准化,其中一小部分用于构造平均脸模板^[9],其他的作为训练人工神经网络的“人脸”样本.我们采用了改进的“自举”(bootstrap)方法^[6],用模板匹配加神经网络验证的方法检测训练图像,反复收集错误报警作为“非人脸”样本,用于神经网的再训练.



①待检测图像,②肤色检测,③区域分割,④区域归并,⑤双眼区域粗筛选,⑥图像标准化,⑦人脸模板匹配,⑧基于模板的检测,⑨验证,⑩候选人脸列表,⑪仲裁,⑫检测结果.

Fig. 1 Flow chart of face detection algorithm
图1 人脸检测算法的流程



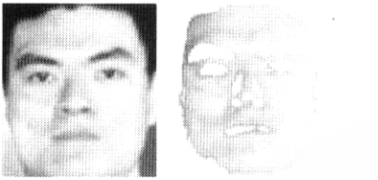
①手工标定,②初始非人脸样本,③非人脸样本,④人脸样本,⑤错误报警,⑥BP算法,⑦训练图像库,⑧肤色模型,⑨平均脸模板,⑩人工神经网络,⑪肤色分割,⑫模板匹配,⑬MLP确认.

Fig. 2 Framework of face detection
图2 人脸检测算法框架

概括地说,我们的方法是基于样本的机器学习方法,具有较强的鲁棒性,而且通过适当地增加样本可望进一步提高检测性能,即具有可扩展性.

2 基于肤色模型的区域分割

2.1 肤色模型



(a) Face image (b) hand cutout face sample
(a) 人脸图片 (b) 剪切出人脸样本

Fig. 3 Face sample

图3 人脸样本

肤色模型是根据大量样本的统计数据建立的.针对东方人种,我们选取了一般室内外照明条件下的 173 幅人脸图片,手工剪裁出肤色区域,共得到 2 375 015 个肤色像素.人脸样本如图 3 所示.

在 HSI(色调-饱和度-亮度)颜色空间,统计样本像素在颜色空间中的频度.如图 4 所示,不同亮度的样本在 H-S 平面上的频度的分布是不同的,因此不考虑亮度分量的模型会有较大的误差.我们将样本按照亮度等间隔地划分为 16 组,然后分别统计,共得到 16 个 H-S 平面上的频度分布.图 4 是亮度分量为 64~79、128~143 和 192~207 的频度分布,图中的亮度代表样本的频度.

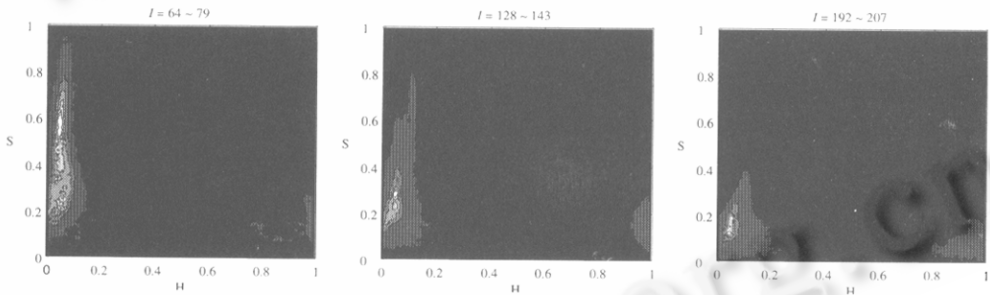


Fig. 4 Skin-Color distribution in HSI color space

图4 肤色样本在 HSI 空间的分布

样本的频度代表了人脸出现某种颜色的概率 $p(\text{color}|\text{face})$,如果可以按照不同条件(如光照)将图片分类,每类的先验概率为 $p(\text{class})$,则有 $p(\text{color}|\text{face}) = \sum_{\text{class}} p(\text{color}|\text{face}, \text{class}) \cdot p(\text{class})$.对于限定条件下获取的图像,可以采用 $p(\text{color}|\text{face}, \text{class})$ 作为似然度来估计像素是否属于肤色.但我们发现,在一般情况下很难对待检测图片进行有效的分类,因而只能用统一的 $p(\text{color}|\text{face})$.进一步实验,我们发现对具体单个图片而言,频度的数值很难反映肤色的似然度.因此,我们仅将肤色样本在 HSI 空间的分布作为一种定性描述,即在检测中简单地判定颜色值在频度分布表中数值为非零的像素就是肤色像素.在算法实现时,将 I, S, H 分别均匀量化为 16、256、256 级,构造一个 $16 \times 256 \times 256$ 的 0-1 查找表,并剔除查找表中孤立的非零点,以减小噪声干扰.检测时逐个考察图像中的各个像素,用查表法判定是否肤色像素.

为了便于肤色分割,需要对肤色检测的结果作平滑处理,抹去大部分噪声点.为此,我们提出了基于肤色像素“密度”的滤波方法:令肤色检测得到的二值图像为 $I_{\text{bin}}(x, y)$,统计 $I_{\text{bin}}(x, y)$ 中每个

5×5 邻域中肤色像素的个数 $d(x, y) = \sum_{i=0}^4 \sum_{j=0}^4 I_{\text{bin}}(x+i, y+j)$,采用类似“卷积”的方法生成新的二

值图像 $I'_{bin}(x,y) = \begin{cases} 1 & d(x,y) > 12 \\ 0 & \text{其他} \end{cases}$. 经过滤波的肤色检测结果如图 5(b)所示.

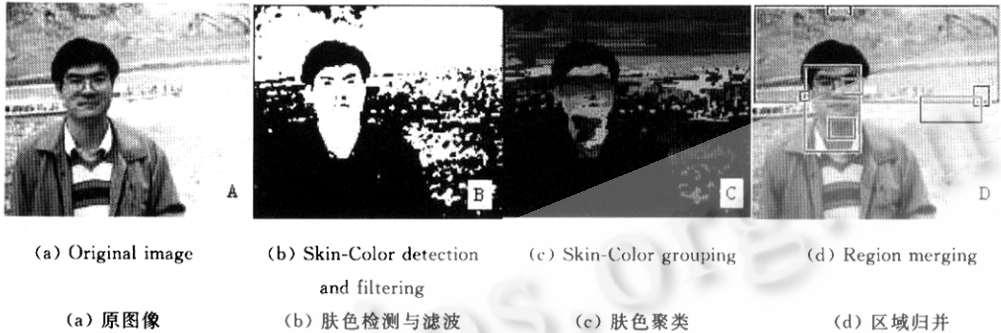


Fig. 5 Skin-Color segmentation

图 5 肤色分割各步骤的处理结果

2.2 区域分割

由于自然界存在大量与人类皮肤颜色相似甚至相同的物体,因此肤色检测的结果中可能包含大量的非皮肤区域.如图 5(a)所示,背景中的黄土和江水也具有人类皮肤的颜色,因而也被检测为肤色(如图 5(b)所示),若直接按照肤色检测的结果进行分割,则得到的区域将远远大于真实的人脸,造成对于人脸尺度的估计错误,最终导致检测失败.

进一步研究发现:(1)在大多数图片中(如图 5(a)所示),某些背景(如黄土)的颜色虽然与其他图片中的人脸颜色相同,但与这幅特定图片中的人脸颜色具有相对的差异性;(2)对训练样本的统计表明,正常光照下的某个特定人脸的绝大部分区域的颜色具有较强的一致性.因此,我们在区域分割中引入了“颜色一致性”判别准则,以期将真实的皮肤区域与类肤色区域分割为不同区域.

算法.

(1)对原始图像进行颜色空间转换和肤色检测,得到 HSI 系统的彩色图像 $I_{HSI}(x,y)$ 和标识相应肤色区域的二值图像 $I'_{bin}(x,y)$,初始化相应的标识矩阵 $I_{mark}(x,y)$.

(2)在 $I'_{bin}(x,y)$ 中由左至右、由上向下扫描,直至找到 3 个连续的未搜索过的肤色点 (x_{init}, y_{init}) , $(x_{init}+1, y_{init})$ 和 $(x_{init}+2, y_{init})$. 将其作为初始新区域

$$R_{new}(x_{left}, y_{top}, x_{right}, y_{bottom}) = (x_{init}, y_{init}, x_{init}+2, y_{init}).$$

初始化 $x_{cur} = x_{init}, y_{cur} = y_{init}, (\bar{H}, \bar{S}, \bar{I}) = (\bar{H}_{init}, \bar{S}_{init}, \bar{I}_{init})$.

(3)从 (x_{cur}, y_{cur}) 在 $I'_{bin}(x,y)$ 中向左搜索;若该点为非肤色点或该点在 $I_{mark}(x,y)$ 中的相应标识为“已搜索”则转步骤(4),否则考察该非零点相应的颜色.若其与 $(\bar{H}, \bar{S}, \bar{I})$ 的差异大于阈值则转步骤(4),否则重新计算 $(\bar{H}, \bar{S}, \bar{I})$. 若新的 $(\bar{H}, \bar{S}, \bar{I})$ 与 $(\bar{H}_{init}, \bar{S}_{init}, \bar{I}_{init})$ 的差异大于阈值则转步骤(4),否则标识当前点为“已搜索”,并记录当前的 x 坐标为 x_d ,继续搜索.

(4)以类似的方法从 (x_{cur}, y_{cur}) 在 $I'_{bin}(x,y)$ 中向右搜索,记录当前的 x 坐标为 x_{cr} .

(5)若(3)、(4)步中均未搜索到符合条件的肤色点,转步骤(6). 否则根据 x_d, x_{cr} 和 y_{cur} 对 R_{new} 进行相应的调整.

令 $x_{cur} = \frac{x_d + x_{cr}}{2}, y_{cur} = y_{cur} + 1$, 转步骤(3).

(6)记录搜索到的区域 $R_{new}(x_{left}, y_{top}, x_{right}, y_{bottom})$, 退回 $(x_{right}+1, y_{top})$ 并转步骤(2)重新搜索,直到搜索完整幅图像.

上述算法第(3)步和(4)步对颜色的判定保证了搜索出的区域中肤色点颜色具有一致性. 图

5(c)显示了搜索的结果,图中不同颜色代表了不同的具有颜色一致性的肤色区域(标识区域的外接矩形没有显示出来).由图 5(c)可以看到,虽然图像中同一人脸的颜色在整体上具有一定的一致性,但由于受阴影和眼镜等附属品的影响,一些局部像素可能与人脸的平均颜色有较大的差异,这时,脸部往往被分割为多个区域.因此,需要对分割出的目标进行归并处理.

2.3 区域归并

归并算法基于这样的事实:(1)图像中的人脸虽然在分割中可能由于局部像素的干扰被分割为多个区域,但这些区域的平均颜色仍比较相近;(2)同一人脸面上分割出的各个区域在位置、尺度和形状上满足一定的几何特征约束.因此,对满足如下颜色和几何条件的区域进行归并:

(1)颜色相近:令 $(\bar{H}_A, \bar{S}_A, \bar{I}_A)$ 和 $(\bar{H}_B, \bar{S}_B, \bar{I}_B)$ 分别为区域A和B中肤色像素在HSI空间各分量的均值,若满足 $|\bar{H}_A - \bar{H}_B| < \delta_H$, $|\bar{S}_A - \bar{S}_B| < \delta_S$ 和 $|\bar{I}_A - \bar{I}_B| < \delta_I$ (其中 $\delta_H = 0.04$, $\delta_S = 0.08$, $\delta_I = 45$,为各个人脸样本像素颜色方差的平均值的两倍),则认为区域A和B的颜色相近;

(2)位置相邻:令 w_1, w_2, h_1, h_2 为区域A,B的宽度和高度, d_x, d_y 为两区域中心分别在x,y方向的距离.若两个区域满足 $d_x < \frac{w_1 + w_2}{2} + \delta_x$ 且 $d_y < \frac{h_1 + h_2}{4}$ (如图6(a)所示),或 $d_y < \frac{h_1 + h_2}{2} + \delta_y$ 且 $d_w < \frac{w_1 + w_2}{4}$ (如图6(b)所示),则认为区域A,B相邻.其中 δ_x, δ_y 取固定值4.

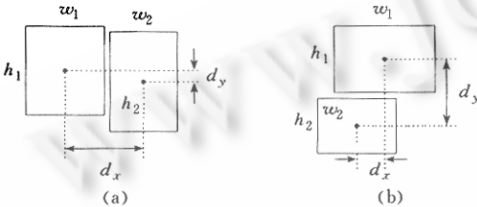


Fig. 6 The adjacent cases in region merging
图6 区域归并中位置相邻的情况

(3)尺度相近:令 w_1, w_2, h_1, h_2 为区域A,B的宽度和高度, $m_1 = \min(w_1, h_1)$, $m_2 = \min(w_2, h_2)$,若有 $\frac{1}{4} < \frac{m_1}{m_2} < 4$,则认为A与B尺度相近.

(4)面积扩大约束:令 $S_A, S_B, S_{A \cup B}$ 分别为区域A,B和两者归并后的面积,若满足 $S_{A \cup B} \leq 1.5 \cdot (S_A + S_B)$,则满足面积扩大约束.

归并操作就是用区域的最小外接矩形替代被归并的区域,并重新计算区域中“肤色”像素的平均颜色.在实现时需要对分割出的区域进行反复的两两归并,直到归并操作不再发生或满足一定的循环次数为止.归并的结果如图5(d)所示,可以看到,虽然许多非人脸的肤色区域也被分割出来,但人脸区域的分割在尺度和位置上都是正确的;而且原图中右下部分的浪花由于不符合颜色一致性和人脸的几何约束,无法被归并成大的区域,最终被滤除.

3 基于模板的人脸检测

经过肤色分割得到了可能包含人脸的区域,现在需要搜索这些区域,以确定人脸的具体位置.我们的方法是对各区域中所有可能位置上尺度在一定范围内的窗口逐个检测,以期得到各个真实人脸的尺度和位置.

3.1 模板匹配

我们用平均脸模板匹配^[9]来进行粗筛选.考虑到双眼在人脸特征中的重要性,将平均脸的双眼部分剪裁出来,与整个人脸一起构成“双眼-人脸”模板对.搜索时,先使用双眼模板,再使用人脸模板,以进一步提高速度并取得更好的筛选性能.在训练图像集中选取了50幅标准证件照,得到 20×20 的平均脸,进一步构造出 20×7 的双眼模板以及 20×20 和 20×24 两种宽长比的人脸模板

(如图 7 所示)。



Fig. 7 Two-Eye template and two face templates of different width to length ratios

图 7 双眼模板和两种宽长比的人脸模板

设 $T[M][N]$, μ_T , σ_T 和 $R[M][N]$, μ_R , σ_R 分别为人脸模板和待检测窗口图像的灰度矩阵、灰度均值与方差, 它们之间的相关系数 $r(T, R)$ 为

$$r(T, R) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (T[i][j] - \mu_T)(R[i][j] - \mu_R)}{M \cdot N \cdot \sigma_T \cdot \sigma_R}$$

对于搜索中遇到的每个待检测窗口, 首先使用双眼模板进行匹配, 若相关系数超过阈值 0.25, 再使用不同宽长比的人脸进行匹配, 若相关系数超过阈值 0.25, 则认为通过了平均脸匹配粗筛选。

3.2 人工神经网络验证

我们采用训练神经网络的方法对通过粗筛选的区域进行分类, 排除误检(错误报警)。

对基于神经网络的人脸检测方法而言, 训练样本(尤其是“非人脸”样本)的选取是一个难题。目前的解决方法是采用“自举”(bootstrap)算法^[6,7], 具有很大的盲目性, 需要大规模的训练样本和计算资源(例如, 文献[6]中每次训练需要在一亿多个非人脸样本中随机选择约 8 000 个, 在超级计算机上训练几天的时间)。我们采用了模板匹配与“自举”算法相结合的“非人脸”样本选取方法, 将非人脸样本的选择限制在一定范围内, 使得在较小的训练集上获得具有很好泛化性能的分类器成为可能。具体过程如下:

(1) 使用单纯的平均脸匹配方法, 对训练集中的(灰度)图片进行检测, 收集错误报警作为初始的“非人脸”样本, 与“人脸”样本一起训练 MLP(three layer perceptron);

(2) 使用平均脸模板匹配与 MLP 验证相结合的方法, 再次对上述图片进行检测, 收集错误报警作为新的“非人脸”样本加入训练样本集, 训练新的 MLP;

(3) 反复进行步骤(2), 直到 MLP 达到预期的性能。

采用上述方法, 训练了两个三层感知器 MLP, 输入节点均为 400 个, 对应于 20×20 图像窗口中按行展开的各个像素; 隐层节点分别为 32 个和 36 个, 与输入层和输出层全链接; 输出节点均为 1 个, 输出范围为 $0 \sim 1$, 理想状态下 1 代表“人脸”, 0 代表“非人脸”; MLP 采用了 Sigmoid 非线性函数, 以下分别简称 MLP1 和 MLP2。在训练过程中使用了相同的“人脸”样本(在训练图像集中手工标出了人脸样本 552 个, 经旋转、镜像等变换生成 3 312 个“人脸”样本), 但使用不同的模板匹配参数收集了不同的初始“非人脸”样本, 并且各自独立地进行训练、检测和收集样本的循环。MLP1 最终使用了 4 065 个“非人脸”样本, MLP2 最终使用了 3 274 个“非人脸”样本。

3.3 搜索与仲裁

考虑到肤色分割得到的区域含有人脸尺度范围的大致信息, 算法中假设可能的人脸宽度在该区域宽度的 $0.3 \sim 1$ 倍之间。对各个区域搜索时采用图像重采样压缩方法, 构造出相应的一系列不同尺度的图像(各个图像尺度差为 1.2 倍), 在压缩图像中逐像素地移动检测窗口进行检测。

在搜索时,仅取局部范围内的最佳匹配位置为最终的检测结果.具体方法是,建立一个候选人脸列表,在搜索过程中将“人脸”窗口逐个加入,若该窗口与列表中已有的候选人脸相重叠,则选择匹配程度高的留在人脸列表中.如图1所示,我们为两个MLP各自建立了候选人脸列表,在搜索过程中分别加入通过了相应MLP验证的区域.当对整个图像的检测搜索结束后,得到两组模板匹配加单个MLP验证的检测结果.

两个MLP使用了相同的“人脸”样本进行训练,因而两个候选人脸列表中正确的人脸在位置、尺度上几乎完全一一对应;另一方面,两个MLP的结构以及使用的“非人脸”样本有所不同,这使得两个候选人脸列表中的错误报警在位置、尺度上大多不同.利用上述特性,对两个人脸列表中的候选人脸进行仲裁,即保留两列表中位置和尺度上近似重叠的候选人脸,删除其他候选人脸.经过仲裁,虽然检测正确率可能有所降低,但错误报警率会降低得更多,总体而言检测性能有所提高.

4 实验结果

我们使用了大量的图像对算法进行测试.根据测试图片的类型和复杂程度,分为3个测试集.(1)简单图像测试集为图像序列Miss America中的150帧图像.其中包含150个正面大致端正的人脸,属于无背景头肩部图像.(2)中等复杂图像测试集为三种不同型号摄像机获取的110幅图像.图像中共有正面大致端正的人脸122个,包括人脸的特写、人物半身和全身多种类型,人脸部的光照比较均匀,背景包括室内和室外两种,约有半数图像的背景中存在类似人脸肤色的木头、土地等物体.(3)复杂图像测试集包括扫描的普通照片以及从Internet下载的图片共111幅.图像中共有正面大致端正的人脸375个,图像均为半身和全身照片,人脸所受光照变化较大.该测试集中的图像具有较为复杂的背景,大多数存在土地等类似肤色的物体.测试实验的统计数据见表1.图8给出了测试集(1)和(2)的部分检测结果,其中A为Miss America序列中的图片,B~D分别为3种不同型号的摄像机获取的图片.图9给出了测试集(3)的部分检测结果,其中A和B为Web上下载和下载的图片,C~E为扫描的普通照片.

Table 1 Face detection on test sets

表1 实验结果统计

Test set ^①	Correct detected ^②	Missed faces ^③	False alarms ^④	Correct rate ^⑤ (%)
Simple image set (150 images, 150 faces) ^⑥	149	1	0	99.3
Mean complexity image set (110 images, 122 faces) ^⑦	116	6	14	95.1
Complex image set (111 images, 375 faces) ^⑧	310	65	119	82.7

①测试集,②正确检测,③漏检,④错误报警,⑤正确率,⑥简单图像测试集(150幅图像,150个人脸),

⑦中等复杂图像测试集(110幅图像,122个人脸),⑧复杂图像测试集(111幅图像,375个人脸).



Fig. 8 Part of detected results on test set (1) and test set (2)

图8 对测试集(1)和(2)进行图片检测的部分结果

如图8所示,本文的算法在肤色分割之后使用了基于模板的搜索,可以对人脸进行精确定位,这是现有的基于“肤色检测-特征验证”的人脸检测方法很难做到的.如图9所示,对于具有更为复

杂的光照和背景条件的图片,本文的算法仍然是有效的,但漏检和错误报警会有所增加.图 9C 中人物的衣服与肤色十分相近,且上面部分褶皱构成了类似人脸的模式,因而被错检为人脸;图 9D 中最右边的人物由于戴眼镜且脸部有亮斑而造成漏检.



Fig. 9 Part of detected results on test set (3)

图 9 对测试集(3)进行图片检测的部分结果

由于合理地使用了肤色信息,我们的算法对于复杂图像的检测速度要比现有算法快得多.以两幅图片为例,在 PC 机(Pentium III600,128M 内存)上,图 8D 尺度为 320×240 像素,处理时间为 2.7s;图 9E 的尺度为 749×517 像素,处理时间为 4.2s.

5 总 结

本文针对彩色图像中的正面端正人脸的检测问题,提出了一种将基于肤色的区域分割与基于模板的人脸检测综合起来的算法.该算法使用按统计方法建立的 HSI 空间的三维肤色模型检测肤色区域,并利用特定人脸肤色本身具有一致性以及人脸与类肤色背景具有相对差异性这一特征,提出了一种鲁棒的人脸肤色分割方法.算法在分割的区域中使用平均脸模板匹配加神经网络验证的方法搜索人脸,有效地提高了检测正确率并降低了误检率.我们的方法是基于样本的机器学习方法,具有较强的鲁棒性,而且通过适当地增加样本可望进一步提高检测性能,即具有可扩展性.

我们将神经网络用于验证,而不是像绝大多数现有工作那样直接用于检测,利用平均脸模板匹配和自举方法搜集非人脸样本,有效地解决了训练神经网络时非人脸样本难以选择的问题.我们认为,这种将模板匹配与神经网络相结合的方法是一种很有潜力的人脸检测新方法.

本文提出的算法也存在一些局限性,例如,没有考虑强侧光、眼睛被头发遮挡以及正面多角度、侧影等情况下的人脸检测.但是不难看出,通过仿射变换生成多角度模板^[9],该算法可以推广到正面多角度人脸的检测,只是计算量将会增大很多.

References:

- [1] Birchfield, S. Elliptical head tracking using intensity gradients and color histograms. In: Proceedings of the IEEE Conference on CVPR'98. Santa Barbara, C A, 1998. 232~237. <http://ied.ihc.com/>.
- [2] Wu, H., Chen, Q., Yachida, M. Face detection from color images using a fuzzy pattern matching method. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999,21(6):557~563.
- [3] Moghaddam, B., Pentland, A. Probabilistic visual learning for object representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997,19(7):696~710.
- [4] Moghaddam, B., Pentland, A. Beyond linear eigenspaces; bayesian matching for face recognition. In: Wechsler, H., Phillips, P. J., Bruce, V., *et al.*, eds. *Face Recognition from Theory to Applications*. New York: Springer-Verlag. 230~243.
- [5] Sung, K., Poggio, T. Example-Based learning for view-based human face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998,20(1):39~51.
- [6] Rowley, H. A., Baluja, S., Kanade, T. Neural network-based face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998,20(1):23~38.
- [7] Osuna, E., Freund, R., Girosi, F. Training support vector machines; an application to face detection. In: Proceedings of the CVPR. Puerto Rico, 1997. 130~136. <http://iel.ihc.com/>.
- [8] 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.
- [9] Liang, Lu-hong, Ai, Hai-zhou, He, Ke-zhong, *et al.* Single rotated face location based on affine template matching. Chinese Journal of Computers, 2000,23(6):640~645 (in Chinese).

附中文参考文献:

- [9] 梁路宏,艾海舟,何克忠,等.基于仿射模板匹配的多角度单人脸定位.计算机学报,2000,23(6):640~645.

Face Detection Based on Skin Color and Template*

AI Hai-zhou, LIANG Lu-hong, XU Guang-you, ZHANG Bo

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China);

(State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China)

E-mail: ahz@mail.tsinghua.edu.cn

<http://www.tsinghua.edu.cn>

Abstract: In this paper, a face detection algorithm for color images based on skin color and template, which is composed of skin color segmentation, template matching, and neural network verifying is presented. First, a skin color model in HSI chrominance space is used for segmenting regions in which may have faces, and then the average-face based template matching and neural network verifying methods are used for searching faces in those regions. This algorithm integrates skin color information in color images with template matching and neural network classification model in gray level images, which results in not only faster speed, but also higher robustness of the algorithm. Experimental results demonstrate the efficiency and feasibility of this algorithm.

Key words: face detection; skin color segmentation; template matching; neural network

* Received May 9, 2000; accepted July 6, 2000

Supported by the National High Technology Development 863 Program of China under Grant No. 863-805-512-9805-11