

# 复杂背景下的三级级联快速正面人脸检测算法<sup>①</sup>

唐 浩, 李小霞

(西南科技大学 信息工程学院, 绵阳 621010)

(特殊环境机器人技术四川省重点实验室, 绵阳 621010)

**摘要:** 针对复杂背景下的正面人脸检测问题, 提出一种三级级联快速正面人脸检测算法: 第一级使用 HSV 肤色模型, 通过分析最大肤色连通区, 快速排除非人脸区域; 第二级采用 Haar-like 特征结合 AdaBoost 算法定位人脸区域; 最后提出基于局部特征点加权的改进主动形状模型(W-ASM)算法匹配人脸的特征点坐标, 通过 68 个特征点位置判断当前人脸图像是否是正面人脸. 实验结果表明, 本算法能准确识别出垂直于图像旋转不超过 $\pm 5^\circ$ 的正面人脸, 每幅图像(640×480)平均用时仅 52ms, 满足实时性要求.

**关键词:** 人脸识别; 正面人脸检测; 主动形状模型; HSV 肤色模型; 特征点匹配知识表示

## Three Cascaded Fast Front Face Detection Algorithm in Complex Background

TANG Hao, LI Xiao-Xia

(School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China)

(Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, Mianyang 621010, China)

**Abstract:** According to the problem of fast front face detection in complex background, a three cascaded fast front face detection algorithm. was proposed. At the first stage, non face regions were excluded through analysis of the maximum color connected regions using the HSV color model. During the second stage, the face areas were further accurately detected with Haar-like features and Adaboost classification algorithm. At last, an improved active shape model algorithm based on the local feature weighting (W-ASM) was proposed to match face feature points coordinates. Whether the current face image is front face can be judged by the locations of 68 feature points. The experimental results show that, this algorithm can accurately identify the front face perpendicular to the image rotating not more than  $\pm 5^\circ$ . The average detection time of each image (640×480) is only 52ms, which can meet the real-time requirements.

**Key words:** Face recognition; Front face detection; Active Shape Model; HSV color mode; Feature points matching

场景信息的感知、分析和理解是信息社会智能化发展的核心技术, 复杂场景中的人脸识别是场景理解的重要内容, 在机器人自主导航、自动驾驶、自然人机交互、多媒体检索、智能监控等领域具有广泛的应用. 而移动或可变背景的复杂场景下人脸识别存在背景改变、光照和阴影等干扰, 以及人脸角度变化等问题, 都是人脸识别领域的研究难点<sup>[1]</sup>.

基于稀疏编码和深度学习的人脸识别算法是目前的研究热点, 2009 年 Wright 等人<sup>[2]</sup>首次将稀疏表示用到人脸识别问题, 在人脸识别领域取得了突破性进

展. 2012 年 Yang Meng 等人<sup>[3]</sup>在前者基础上提出 MRR 算法, 将人脸对齐中求解变换矩阵 T 分为由粗到细的两个阶段, 进一步提高了识别率. 2012 年 Xudong Cao 等人<sup>[4]</sup>提出改进的主观模型(Active Shape Models, ASM)算法, 其核心思想是通过训练样本的线性组合来约束人脸形状模型. 2014 年 Sun Y 等人<sup>[5]</sup>提出基于深度学习和类内类间联合特征的人脸识别算法, 在 LFW 人脸数据库上检测率为 99.15%. 但是目前大多数算法都是在人脸正对摄像头的前提下进行研究, 对于多角度人脸检测还存在很大困难. 而在实际应用中, 采集到的人

① 收稿时间:2014-08-28;收到修改稿时间:2014-09-30

脸图像往往不是正对摄像头, 因此需要先进行正面人脸检测, 排除非正面人脸情况, 减少人脸识别的错误率.

针对复杂场景下正面人脸检测的精度、速度和鲁棒性问题. 受文献[6]启发, 本文采用三级级联的方式进行正面人脸检测. 首先根据人脸的肤色特性, 将原彩色图像投影到 HSV 颜色空间, 通过肤色统计模型快速过滤掉其它颜色区域, 使用 Seed Filling(种子填充)方法<sup>[7]</sup>确定最大肤色连通区. 其次需要对人脸进行定位, 采用 Haar-like 特征结合 AdaBoost 算法<sup>[8]</sup>训练分类器, 通过多尺度遍历的方式找出最大肤色连通区中人脸的具体位置. 最后本文提出局部特征点加权的改进主动形状模型(W-ASM)算法定位人脸的 68 个特征点坐标, 依靠特征点的位置关系判断当前人脸图像是否是正面人脸. 由于 ASM 算法对人脸特征点具有全局形状约束能力, 对重要特征点进行加权判断后, 使本文算法对人脸局部特征变化有较强的鲁棒性.

### 1 算法概述

为了提高算法的检测速度和检测的准确性, 本文提出三级级联的正面人脸检测方法, 算法流程图如图 1 所示.

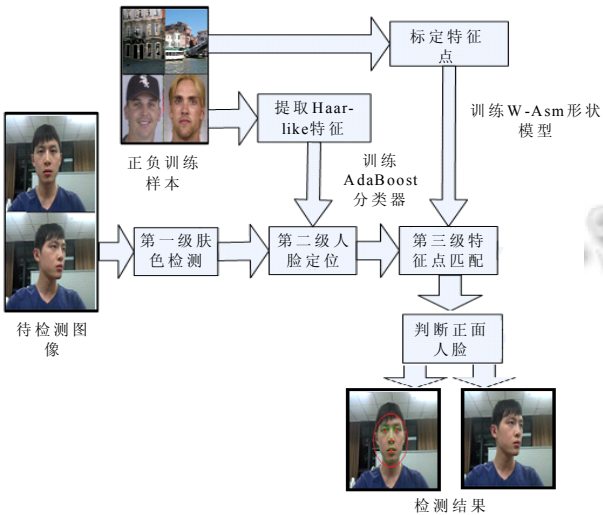


图 1 算法总体框图

图 1 中的检测步骤包括四步:

第 1 步, 肤色检测. 将待检测图像投影到 HSV 颜色空间, 过滤掉非肤色区域, 用种子填充法寻找肤色像素点连通区域, 取最大连通区域为待检测区域.

第 2 步, 人脸定位. 选定 Haar-like 特征模板中的边缘特征模板(如图 2 所示), 通过分别改变四个矩形模板在正负样本图片中的位置和大小提取特征值, 其计算方法为白色矩形像素和减去黑色矩形像素和, 最后将所有特征值级联组成特征向量, 用于人脸识别.

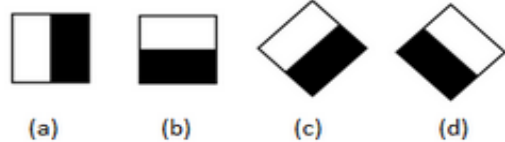


图 2 haar-like 边缘特征

得到人脸样本特征集后, 需要训练 AdaBoost 分类器识别出待检测图像肤色区域中的人脸, 用于定位. AdaBoost 算法的基本思想是: 利用多个分类能力的弱分类器, 提升为一个强分类器. 具体算法描述如下:

```

1:  $w = \{w_j = 1/N | j=1, 2, \dots, N\}$  //初始化样本权值
2: for  $i = 1$  to To do
3: 根据  $w$  对训练样本集  $D$  进行抽样构成训练集  $D_i$ , 在  $D_i$  上训练分类器  $C_i$  用其对  $D$  分类
4:  $\epsilon_i = \frac{1}{N} [\sum_j w_j \delta(C_i(x_j) \neq y_j)]$  //计算错误率
5:  $\alpha_i = \frac{1}{2} \ln \frac{1 - \epsilon_i}{\epsilon_i}$ 
6:  $w_i^{(j+1)} = \frac{w_i^j}{Z_j} \times \begin{cases} e^{-\alpha} & \text{if } C_j(x_i) = y_i \\ e^{\alpha} & \text{if } C_j(x_i) \neq y_i \end{cases}$  //更新权重
7: end for
8:  $C(x) = \arg \max \sum_{j=1}^T \alpha_j \delta(C_j(x) = y)$ 

```

第 3 步, 特征点匹配. 首先对人脸正样本进行标点, 每张人脸图像按固定位置共标记 68 个特征点, 如图 3 所示. 采用改进的局部特征值加权 ASM 算法, 训练 ASM 正面人脸形状模型. 用训练好的人脸模型对人脸区域进行特征点匹配, 得到 68 个特征点在待检测图像的新位置.

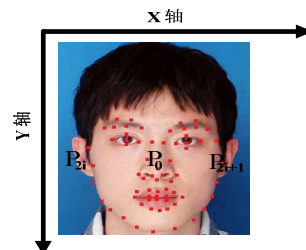


图 3 人脸特征点

第 4 步, 判断正面人脸. 将人脸分为左右两个区域, 计算同一高度左右对应特征点到  $P_0$  的欧式距离之差, 即:

$$g(x) = \delta - \sum_{i=0}^{34} \left| \|P_{2i} - P_0\|_2^2 - \|P_{2i+1} - P_0\|_2^2 \right| \quad (1)$$

式(1)中  $\delta$  表示阈值,  $x$  表示当前人脸, 若  $g(x)$  大于 0 说明  $x$  为正面人脸, 反之亦然.

## 2 肤色检测

### 2.1 HSV 肤色模型

将待检测图像从原始 RGB 色彩空间投影到 HSV(hue, saturation, value)颜色空间. HSV 颜色空间的模型对应于圆柱坐标系中的一个圆锥形子集, 是一类以色调为基础的颜色模型, 其中 H 代表色调, S 是饱和度, V 则表示了色彩亮度. 其色彩 H 由绕 V 轴的旋转角给定, 圆锥切面包含 RGB 模型中的 R, G, B 三个面, 红色对应于角度  $0^\circ$ , 绿色对应于角度  $120^\circ$ , 蓝色对应于角度  $240^\circ$ .

在 HSV 空间模型中, 人脸肤色区域具有线性可分性. 并且由于颜色分量的直觉性以及亮度和色度相分离, 因此将其用于肤色检测有很好的效果. 通过统计发现, 人脸肤色都是分布在一个较窄的频带上, 在色相 H 分量中, 表示为集中在某一小区域范围内, 肤色模型为:

$$\begin{cases} 0 \leq H \leq 25 \\ 26 \leq S \leq 200 \\ 20 \leq V \leq 200 \end{cases} \quad (2)$$

### 2.2 肤色连通区域分析

待检测图像通过肤色检测后, 由于光照等因素的影响得到的往往不是平滑的整块图像, 肤色图像通常成离散的不规则块状分布, 并且含有大量的噪声点, 因此需要通过寻找最大肤色连通区域以确定人脸在待检测图像中可能出现的位置, 检测流程如图 4 所示.

连通区域分析是一种在图像处理应用领域中较为常用和基本的方法, 本文采用 Seed Filling(种子填充)方法遍历肤色二值图像, 快速寻找连通区域. 其基本思想是从二值图像中找出任意一个值为 1 的像素点作为种子, 然后根据连通区域的两个基本条件(像素值相同、位置相邻)将与种子相邻的像素点合并到同一个像素集合中, 最后得到的该像素集合则为一个连通区域. 以此类推, 随机选取没有在像素集合的值为 1 的点作

为下一个种子, 找出该像素所在联通区, 直到没有种子时结束寻找, 具体步骤如下:

① 扫描图像, 直到当前像素点  $B(x,y)=1$ :

- 将  $B(x,y)$  作为种子(像素位置), 并赋予其一个 label, 然后将该种子相邻的所有前景像素都压入栈中;
- 弹出栈顶像素, 赋予其相同的 label, 然后再将与该栈顶像素相邻的所有前景像素都压入栈中;
- 重复 b 步骤, 直到栈为空;

此时, 便找到了图像 B 中的一个连通区域, 该区域内的像素值被标记为 label;

① 重复第 1 步, 直到扫描结束.

② 计算找到的所有连通区域面积, 取面积最大区域为人脸所在区域.

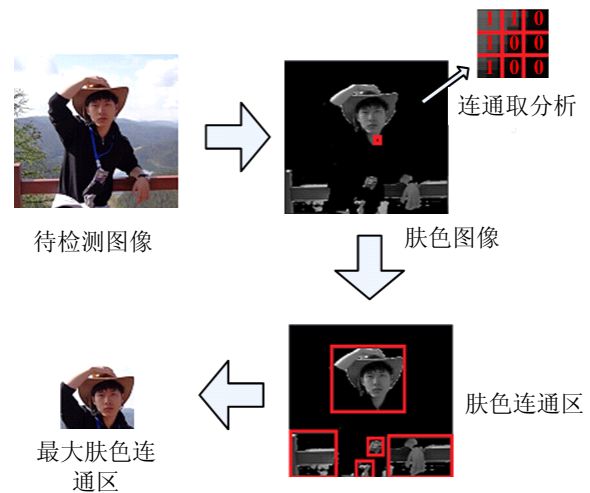


图 4 肤色检测

## 3 局部特征点加权的ASM算法

### 3.1 训练 W-ASM 形状模型

ASM<sup>[9]</sup>是一种基于点分布模型(Point Distribution Model, PDM)的算法. 在 PDM 中, 外形相似的物体, 例如人脸、人手、心脏、肺部等的几何形状可以通过若干关键特征点的坐标依次串联形成一个形状向量来表示.

因此将人脸图像标定的 68 个特征点组成一个形状向量, 即:

$$a_i = (x_1^i, y_1^i, x_2^i, y_2^i, \dots, x_{68}^i, y_{68}^i), i = 1, 2, \dots, n \quad (3)$$

其中  $(x_j^i, y_j^i)$  表示第  $i$  个训练样本上第  $j$  个特征点坐标,  $n$  表示训练样本个数.

得到  $n$  个形状向量后采用 PCA<sup>[10]</sup>主分量提取算法对形状向量进行降维处理, 于是任意一组特征点集可

以看做是主成分向量空间的一个坐标点, 而这个坐标原点就认为是这些点集的平均, 这样任意点位置即是坐标原点加上一个向量. 即:

$$a_i \approx \bar{a} + Pb \quad (4)$$

式(4)中  $\bar{a}$  表示平均形状向量,  $P$  是包含了主成分向量的协方差矩阵,  $b$  是一个小于  $n$  维的向量, 用来控制特征点形状的变化.

由于在人脸图像中眼角和唇角的边缘变化明显, 梯度值大于其余特征点, 在特征点匹配时这些点的位置较为关键, 因此对其进行加权处理, 能使原 ASM 形状模型能匹配的更为准确. 加权后的形状向量表示为:

$$a_i \approx \bar{a} + Pb + \alpha \sum_{m=1}^6 f(x_m, y_m) \quad (5)$$

式(5)为 W-ASM 形状模型, 其中  $\alpha$  表示加权系数,  $f(x_m, y_m)$  分别表示左眼的内外眼角、右眼的内外眼角和左右唇角 6 个特征点.

### 3.2 形状模型特征点匹配

在通过样本集训练得到 W-ASM 模型后, 对其进行仿射变换, 获得在检测图像上的初始形状模型:

$$X = M(s, \theta)[a_i] + X_c \quad (6)$$

其中  $\theta$  表示形状模型逆时针旋转角度,  $s$  表示缩放倍数,  $X_c$  表示移动步长. 将该模型每个点覆盖检测图像, 如图 5 所示:

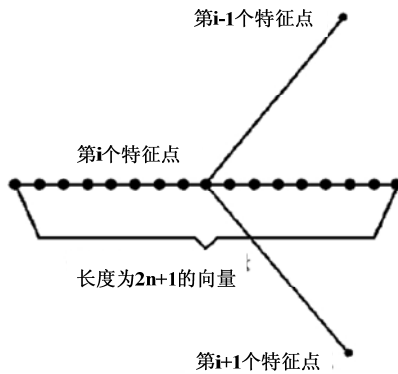


图 5 特征点匹配示意图

对于模型中第  $i$  个特征点, 在其垂直于前后的两个特征点连线方向分别选择  $n$  个像素, 计算每个像素点的灰度值导数作为局部特征, 找出与当前特征点局部特征马氏距离最小的点为当前特征点的新位置. 为所有 68 特征点找到其新位置, 组成一个移动向量:

$$dX = (dX_1, dX_2, \dots, dX_k) \quad (7)$$

调整式(6)参数, 迭代运算找出新的位置  $dX_2$ , 当  $dX_n$  与  $dX_{n+1}$  变化小于一定值时, 认为得到全局最优解, 停止迭代.

## 4 实验分析

为了验证本文算法的性能, 本节分为两个部分: 首先在 ORL 人脸数据库上进行 ASM 算法和本文提出的 W-ASM 算法比较; 然后针对复杂背景下的快速正面人脸检测问题, 选取部分 INRIA 行人数据库中的单个行人图片, 用于仿真实验.

实验平台: 3.4GHz i3-3240 CPU, 2GB RAM, Windows 7 算法编译环境: Visual Studio 2010, OpenCV 2.4

实验数据库: ORL 人脸数据库, INRIA 行人数据库

### 4.1 算法对比

在 ORL 数据库上选取 200 张人脸图像作为正样本, 按图 3 所示的人脸位置对每张图像进行人工标点, 记录每张图像标 68 个特征点的位置信息用于训练两种算法的平均形状模型. 为了便于观察, 在检测结果中将 68 个点用线段连接表示.

对选取的 6 个关键特征点进行加权, 其权系数  $\alpha$  的值直接决定了关键特征点对整个形状模型的影响能力,  $\alpha$  值越大权重越大. 但当  $\alpha$  值过大时, 会大大制约其他特征点对形状模型的约束力, 容易得到局部最优解, 降低算法匹配准确性. 本文通过实验方法取不同  $\alpha$  值, 观察其对算法结果的影响, 选取最优权重. 检测结果统计数据如表 1 所示.

表 1 ORL 数据库结果比较

方法	$\alpha$ 值	平均检测时间(ms)	准确率(%)
ASM[9]		10.86	93.5
W-ASM1	1.5	10.2	93
W-ASM2	2.3	9.84	96
W-ASM3	3.0	9.4	91.5
W-ASM4	4.0	6.4	83

从表 1 中可以看出随着  $\alpha$  值增大检测时间明显减少, 提高了算法收敛速度.  $\alpha$  取 2.3 时本文算法准确率最高, 比原 ASM 算法匹配准确率提升了 2.5%.  $\alpha$  取 2.3 时, 部分检测图像如图 6 所示(其中眼角和唇角特征点加粗的图像为本文算法检测结果).



ASM检测结果



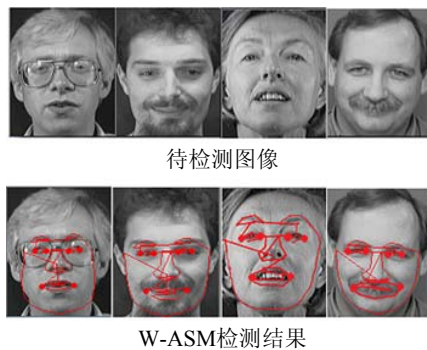


图 6 ORL 数据库检测结果



图 7 INRIA 数据库检测结果

#### 4.2 三级级联正人脸检测算法性能测试

为了检验每级算法的效果和时间,从 INRIA 数据库中随机抽取 50 张单个人物的测试图像,尺度统一为  $640 \times 480$ ,进行分步检测实验,检测结果如表 2 所示:

表 2 分步检测结果

方法	平均检测时间 (ms)	平均处理图 像面积	误检率率(%)
第一级肤色检测	9.02	$640 \times 480$	6.7
第二级人脸定位	32.3	$120 \times 90$	8.1
第三级特征点匹	11.67	$50 \times 50$	3.4

如表 2 所示,第一级肤色检测算法速度最快,算法耗时与处理图像面积成稳定线性关系,处理尺寸为  $640 \times 480$  大小的图片平均耗时仅 9ms,能快速排除非肤色区域,当人脸不在最大肤色连通区中出现误检.第二级由于采用 haar+AdaBoost 分类方法识别定位人脸,需要多尺度遍历图像,随图片尺寸变大计算量成指数递增,耗时越长.因此必须通过第一级肤色检测模块排除大量无关区域,减少计算量.由于第三级只需对第二级检测出的整个人脸区域进行 W-ASM 形状模型特征点匹配,避免了图像扫描,使本文算法满足实时性要求.部分检测结果如图 7 所示,有方框表示检测出是正面人脸.

## 5 结语

本文提出一种三级级联结构的快速正面人脸检测算法,应用于复杂背景下的实时正面人脸检测.算法先通过肤色检测,提取最大连通区,快速排除大量无关区域,减少运算量.然后采用经典的 haar+Adaboost 分类算法进行多尺扫描,度定位人脸位置.最后提出局部特征点加权的改进 ASM 进行特征点匹配,提高了正面人脸检测的准确率.

## 参考文献

- 1 夏佩佩,张莉.一种相似性学习算法及其在人脸识别中的应用.计算机工程,2014,40(6):175-179.
- 2 Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 2009, 31(2): 210-227.
- 3 Yang M, Zhang L, Zhang D. Efficient misalignment-robust representation for real time face recognition. Computer Vision - ECCV 2012, 2012: 850-863.
- 4 Cao XD, Wei YC, Wen F. Face alignment by explicit shape regression. IEEE Conference on Biometrics Compendium. 2012. 2887-2894.
- 5 Sun Y, Wang X, Tang X. Deep learning face representation by joint identification-verification. arXiv preprint arXiv: 1406.773. 2014.
- 6 Wu J, Geyer C, Rehg JM. Real-time human detection using contour cues. International Conference on Robotics and Automation. 2011. 860-867.
- 7 刘万春,刘建君,朱玉文,等.一种实时高速的八连通区域填充算法.计算机应用研究,2006,(6):177-179.
- 8 Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. Proc. of IEEE Conference on Computer Vision and Pattern Recognition 2001. IEEE. 2001. 1-511-1-518.
- 9 Wan KW, Lam KM, Ng KC. An accurate active shape model for facial feature extraction. Pattern Recognition Letters, 2005, 26(15): 2409-2423.
- 10 Weingessel A, Hornik K. Local PCA algorithms. IEEE Trans. on Neural Networks, 2000, 11: 1242-1250.