文章编号:1001-9081(2010)04-0967-04

基于人脸特征和 AdaBoost 算法的多姿态人脸检测

阮锦新,尹俊勋

(华南理工大学 电子与信息学院,广州 510641)

(ruanjinxin@163.com)

摘要:基于人脸特征和 AdaBoost 算法,提出一种改进的多姿态人脸检测算法。首先利用肤色特征快速排除绝大部分背景区域,然后在肤色区域中搜索眼睛和嘴巴区域,根据眼睛和嘴巴区域的几何特征所确定的人脸方向分割出大致正向的人脸候选区域,最后利用 AdaBoost 算法对候选区域进行分类。实验表明,算法能实现多姿态人脸的快速检测,而且对脸部表情和遮挡有较强的鲁棒性。

关键词:人脸检测;多姿态;肤色区域;人脸特征;AdaBoost 算法;Harr-like 特征

中图分类号: TP391.41 **文献标志码:** A

Multi-pose face detection based on facial features and AdaBoost algorithm

RUAN Jin-xin, YIN Jun-xun

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou Guangdong 510641, China)

Abstract: An improved multi-pose face detection algorithm based on facial features and AdaBoost algorithm was proposed. Making full use of facial skin color information firstly, most of the background regions could be quickly excluded. After detecting eyes and mouth, the approximate frontal face candidate regions were segmented according to face orientation decided by the geometric features of the eyes and mouth regions. At last, the face candidate regions were classified by AdaBoost algorithm. The experimental results demonstrate that the algorithm can further improve the multi-pose face detection accuracy and is highly robust to facial expression and occlusion.

Key words: face detection; multi-pose; skin color region; facial feature; AdaBoost algorithm; Harr-like feature

0 引言

随着人脸识别技术的发展与应用,人脸检测作为人脸技术应用关键和首要的一步,得到了广泛的关注和研究^[1]。现在,在较理想的条件下,正面人脸检测已取得了令人满意的效果。然而,在一般情况下,由于多姿态、遮挡、光照等因素影响,人脸检测仍需进一步研究^[2],而且这也是实际应用中亟须解决的问题。

基于人脸特征是人脸检测的常见方法。人脸肤色^[3]是人脸的重要特征之一。在整个色度空间中,人脸肤色空间只是一个很小的子空间,有较强的聚类特性,肤色特征易于提取。利用肤色进行人脸检测,算法简单、速度快,不易受姿态、表情何旋转等影响,具有相对的稳定性。在彩色图像中,利用肤色特征进行人脸检测预处理,能有效缩小人脸检测的范围。因此,基于肤色特征的人脸检测算法得到了广泛的应用^[4]。但是,在复杂背景下,只基于肤色进行人脸检测并不充分,误检率高,检测精度低,所以常常结合其他方法一起使用。

2001年 Viola 和 Jones 引入积分图概念,提出了基于 Harr-like 特征、级联结构的 AdaBoost 算法,成功应用于物体的快速检测^[5],并在人脸检测领域取得了突破性进展,实现了实时人脸检测。在 Viola 和 Jones 的算法中,主要考虑水平矩形特征,在正面的人脸检测中得到了很高的检测率,但对多姿态人脸检测的检测率不高。近年来,在多姿态人脸检测中,基于 AdaBoost 算法本身提出了很多的改进方法,例如:扩展

Harr-like 特征^[6]、增加姿态检测器^[7]等,都能有效降低误检率,却影响了检测速度。

本文提出了一种基于肤色等人脸特征和 AdaBoost 算法的人脸检测算法,实现多姿态人脸的快速检测。算法首先结合人脸特征,利用眼睛、嘴巴确定人脸方向,按人脸方向分割,得到大致正向的人脸候选区域,然后,用级联结构的 AdaBoost 算法进行验证。这样,可避免由于扩展 Harr-like 特征或增加姿态检测器等原因而增加算法的复杂度,使速度和准确率都有所提高。

1 多姿态人脸检测算法

人脸检测算法流程如图1所示。算法主要由图像预处理,肤色区域检测,眼睛、嘴巴检测,人脸候选区域分割和 AdaBoost 训练与候选区域分类等组成。

图1 人脸检测算法流程

算法主要利用肤色在 $YCbCr$ 色彩空间的色度聚类特性进行肤色检测,从复杂的背景中分割出肤色区域。然后,在肤色

收稿日期:2009-10-09;修回日期:2009-12-07。

作者简介:阮锦新(1971-),男,广东广州人,博士研究生,主要研究方向:人脸检测、表情识别;尹俊勋(1942-),男,广东东莞人,教授,博士生导师,主要研究方向:模式识别、通信与音视频信号处理。

区域搜索眼睛和嘴巴,并按照眼睛、嘴巴区域的几何特征所确定的人脸方向分割出人脸候选区域。最后,利用 AdaBoost 对人脸候选区域进行分类。

1.1 肤色区域检测

1.1.1 色彩空间转换

对于不同种族的人,肤色差别主要是亮度的差别,而色度的差别是很小的^[8]。在 RGB 色彩空间中各分量均含亮度信息,相关性大,易受光照影响。 YC_bC_r 色彩空间是为电视系统开发的亮度分离的色彩空间,广泛应用于 MPEG、JPEG 等图像视频压缩。在 YC_bC_r 色彩空间,肤色有较好的聚类性,使其成为最具有吸引力的肤色建模空间之一。因此,采用目前常用于肤色检测的 YC_bC_r 色彩空间。

YC_bC_r 色彩空间可通过 RGB 色彩空间线性转换得到,转换算法简单。为了减少光照变化的影响,采用分段非线性色彩空间转换^[9],将 YC_bC_r 色彩空间转换成 $YC'_bC'_r$ 色彩空间。

1.1.2 肤色区域检测

为了减小光照、噪声等因素对肤色检测的影响,首先对样本图像和待测图像进行光照补偿,并进行 3×3 的中值滤波处理。在 $YC'_bC'_r$ 色彩空间,根据椭圆聚类方法^[9],按式(1)进行肤色分割后进行形态学处理,得到比较理想的肤色区域检测效果。

$$\begin{cases} \frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} < 1 \\ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} C'_b - c_x \\ C'_r - c_y \end{bmatrix} \end{cases} \quad (1)$$

其中 $a = 25.39$, $b = 14.03$, $ec_x = 1.60$, $ec_y = 2.41$, $\theta = 2.53$, $c_x = 109.38$, $c_y = 152.02$ 。

在 $YC'_bC'_r$ 色彩空间,如果像素满足式(1),则被认为是一个肤色像素,记为 1;否则,认为是一个非肤色像素,记为 0。这样,很容易得到一个二值图像。经过形态学处理,可以准确检测出肤色区域。

根据肤色连通区域的几何特征,如区域长、宽、肤色所占比例等,可以去掉明显的非人脸区域。例如:对于长或宽小于 20 像素的矩形肤色区域不作处理,因为在脸部特征提取的过程中无法提供足够的细节用于人脸识别。而且,在算法中采用 24×20 的图像进行训练。另外,人脸区域中肤色应占一定的比例,根据实验,肤色比例小于 30% 的区域,也认为是非人脸区域。

1.2 人脸候选区域分割

为了能得到大致正向的人脸区域,首先,搜索出眼睛、嘴巴区域,然后,根据其几何特征,确定人脸的方向,最后,按此方向分割出人脸候选区域。由于可能存在遮挡情况,例如头发、眼镜和口红等,这些都影响检测效果。所以,为了增强对遮挡等因素影响的鲁棒性,对所有能检测出眼睛或嘴巴的肤色区域都进行相应的处理。

1.2.1 眼睛检测

由于眼睛的虹膜与皮肤有较大差异,是明显的非肤色区域,所以眼睛是人脸的重要特征之一。在 YC_bC_r 色彩空间中,眼睛与皮肤的 C_b 和 C_r 分量有较大差异。而且,眼睛的灰度值比较低, Y 分量集中在 $(0, 120)$ 内。另外,眼睛的 C_b 分量普遍

比 C_r 分量的值要高。将眼睛的色度和亮度映射结合起来^[9],能准确检测出眼睛区域的大小和位置。

1.2.2 嘴巴检测

嘴巴既是明显的非肤色区域,也是人脸的重要特征之一。嘴巴的方向常常与人脸方向是一致的,具有很强的稳定性。与其他器官相比,嘴巴的颜色中 C_r 分量更高, C_b 分量更低,增大 C_r 和 C_b 分量间的差异^[9],可以准确检测出嘴巴区域的大小和位置。

1.2.3 人脸候选区域分割

在同一个肤色连通区域中搜索眼睛、嘴巴区域,判断眼睛和嘴巴的上下位置关系,计算出每两只眼睛中心连线与水平方向的夹角,以及嘴巴区域与水平方向的夹角。把夹角差小于 10° 的分成一组,计算出每组夹角的平均值。当只搜索到嘴巴或一只眼睛区域时,为了避免由于遮挡等原因而漏检,同样对这些区域作相应处理。最后,根据各组夹角确定人脸的方向,分别分割出相应的人脸候选区域。

1.3 人脸候选区域分类

人脸候选区域是根据人脸方向分割出的,候选区域中人脸的方向大致是正向的,充分利用级联结构的 AdaBoost 算法在正面人脸检测方面快速与准确率高的优势,对人脸候选区域进行分类,可以获得相当好的检测效果。具体算法是:通过对候选区域进行积分图计算,快速计算 Harr-like 特征,利用 AdaBoost 算法将 Haar-like 特征生成的弱分类器叠加成为强分类器,再将多个强分类器级联组成更强有力的人脸检测分类器。

1.3.1 级联结构的 AdaBoost 算法

AdaBoost 算法是具有自适应的 boosting 算法,是 boosting 的有力改进。算法的基本思想是将大量的分类能力一般的弱分类器通过一定方法叠加起来,构成一个具有很强分类能力的强分类器。理论证明,只要每个弱分类器分类能力比随机猜测要好,当弱分类器的个数趋向于无穷时,强分类器的错误率将趋于零。

由于候选区域中人脸的方向大致是正向的,同时,为了保证人脸检测器的速度,检测系统只采用图 2 所示的简单一维 Harr-like 特征。

图2 系统采用的 Harr-like 特征

引入积分图概念后,只需对候选区域进行一次遍历计算,就能以相同时间快速地计算不同尺度下的矩形特征,加快了检测分类器的训练和检测。

级联结构人脸检测分类器的每一级都采用 AdaBoost 算法进行训练。通过 AdaBoost 学习算法将 Haar-like 特征生成的弱分类器叠加成为强分类器,使之具有很强的泛化能力。具体训练算法如下:

1) 给定 n 个样本图像 $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_n, y_n)$, x_i 是输入样本图像, y_i 是类别标志,其中, $y_i = 0, 1$ 分别表示其为非人脸样本和人脸样本。

2) 初始化权重:

$$w_{1,i} = \begin{cases} \frac{1}{2m}, & y_i = 0 \\ \frac{1}{2l}, & y_i = 1 \end{cases}$$

其中, m 和 l 分别是非人脸样本和人脸样本的数量, $n = m + l$ 。

3) For $t = 1, 2, \dots, T$:

a) 归一化权重:

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}, w_t \text{ 为统计分布。}$$

b) 对每个特征 j , 训练一个弱分类器 h_j , 计算对应 w_t 的错误率:

$$\varepsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$

c) 选取最小错误率 ε_t 的弱分类器 h_t 。

d) 更新权重:

$$w_{t+1,i} = w_{t,i} \beta_i^{1-\varepsilon_i}$$

其中, 当 x_i 被正确地分类时, $\varepsilon_i = 0$, 反之, $\varepsilon_i = 1$; $\beta_i = \frac{\varepsilon_i}{1 - \varepsilon_i}$ 。

4) 最终的强分类器为:

$$h(x) = \begin{cases} 1, & \sum_{i=1}^T \alpha_i h_i(x) \geq \frac{1}{2} \sum_{i=1}^T \alpha_i \\ 0, & \text{其他} \end{cases}$$

其中 $\alpha_i = \log \frac{1}{\beta_i}$ 。

级联结构的 AdaBoost 人脸检测器^[7]由多级强分类器组成,每一级的强分类器都采用 AdaBoost 算法进行训练。级联结构人脸检测器是一种从简单到复杂的结构。前面的分类器使用的特征数较少,结构相对简单,能够在高检测率下,快速排除与人脸差异较大的背景区域。后面的分类器则使用更多的特征,结构相对复杂,从而能够将与人脸相似的背景区域区分开来。而且,大部分的背景区域在前面比较简单的分类器中被排除,只有少部分待测区域需要通过后面比较复杂的分类器进行分类,所以提高了检测的速度。级联结构的 AdaBoost 人脸检测器如图 3 所示(T 表示 True, F 表示 False)。

图3 级联结构的人脸检测分类器

1.3.2 训练样本选取

由于基于样本学习的统计方法对训练样本的依赖比较大,所以提高训练样本的有效性显得尤为重要。如图 4 所示,是 Viola 和 Jones 算法采用的部分人脸样本图像^[7],样本图像的长宽比为 1。但是,人脸的长宽比一般约为 1~1.5,而且,人并不总是闭上嘴巴,所以,人脸的长宽比常常大于 1。一般情况下,人脸样本的重要纹理信息主要集中在人脸的中央。如果人为地把人脸的长宽比设定为 1,就会使人脸区域常常包含了人脸两边并不重要的纹理区域,从而降低了样本的有效性。所以,实验时按长宽比为 1.2,图像大小为 24×20 截取人脸样本。如图 5 所示,是部分自建的人脸样本图像。

实验中的样本图像来源于网上下载和生活摄影照片,手

工选取了正面、侧面人脸样本各 2000 多个,其中,一些人脸样本有小部分遮挡,人脸表情丰富;此外,手工选取了 6000 多个“非人脸”样本。人脸检测是属于小样本问题,样本过小易造成过训练,同时为了提高算法的泛化能力,增强对旋转和尺寸微小变化的鲁棒性,对样本进行了如下处理:1) 将每个样本向左或右分别做小于 5° 的旋转;2) 沿 Y 轴方向镜像;3) 按一定尺寸进行缩放。经过这些处理,共得 30 000 多个人脸样本和 50 000 多个非人脸样本。最后,所有样本经过光照补偿和图像去噪处理。

图4 部分 Viola 和 Jones 算法采用的人脸样本图像^[7]

图5 部分自建的人脸样本图像

1.3.3 人脸候选区域分类

为了能检测出人脸候选区域中不同位置的人脸,检测窗口按照从上到下,从左到右,平移窗口逐个位置检测。根据积分图计算,检测窗口在不同尺度下,计算特征值所需时间都是一样的。所以,只对检测窗口进行逐层等比放大。每次放大后,移动检测窗口对人脸候选区域重新检测,就能检测出人脸候选区域中不同大小的人脸。为了加快检测速度,检测窗口在每次放大后作适当的平移。从实验数据分析看,为了同时兼顾检测速度和检测率,采用的检测窗口初始大小为 24×20 ,每次放大 1.2 倍,每次放大后平移 3 个像素,效果比较理想。

由于在一般情况下,同一人脸在不同比例和邻近的位置上会被多次检测到,因此,当人脸窗口面积重叠达到 70% 以上,则认为这两个窗口代表同一个人脸,将两者合并,并将它们位置和大小平均值作为合并后的候选人脸区域的位置和大小。另外,被误检的人脸区域被重复检测到的次数往往较少。

2 实验结果及分析

在实验中,分别采用了两个样本集,一个是与 Viola 和 Jones 算法相同的 MIT + CMU 正面和侧面样本集,另一个是按长宽比为 1.2 选取样本的自建样本集,以测试自建集样本的有效性。

实验采用了 160 张彩色测试图像,图像背景复杂,大小为 320×240 像素,每张图像包含 1 张或多张人脸,共 856 张人脸。大部分人脸姿态各异,包括旋转的正面和侧面人脸,有些人脸还有小部分被遮挡。

测试系统是基于 Microsoft Visual C++ 2005 和 Intel Open CV 编程实现的,在 Intel 酷睿 2 双核 2.83 GHz CPU、2 GB 内存的台式计算机上运行。测试结果如表 1 所示。

从实验数据看,相对于 Viola 和 Jones 算法,本文算法在检测率和检测速度方面都有所提高。而且,采用自己建立的样本集,检测率也得到了进一步的提高,有效地降低了误检

数。图 6 是采用所提算法得到的部分实验结果。

表 1 实验结果

检测算法	训练样本集	误检数	检测率/%	平均速度 ms/帧
Viola 和 Jones 算法 ^[7]	MIT + CMU	645	82.6	113
本文算法	MIT + CMU	186	86.9	38
本文算法	自建样本集	109	91.9	40

(a) 部分实验结果1

(b) 部分实验结果2

图 6 部分实验结果

3 结语

在多姿态人脸检测方面,提出基于人脸特征和 AdaBoost 算法相结合的检测算法。与其他文献相比,算法经过肤色检测、眼睛和嘴巴的搜索,快速有效地缩小人脸检测的范围,提高了检测速度。通过眼睛、嘴巴的几何特征确定人脸方向,按人脸方向分割出大致正向的人脸候选区域,在不扩展 Harr-like 特征或增加姿态检测器等情况下,利用 AdaBoost 算法较好地解决了包括人脸旋转在内的多姿态人脸检测问题。同

时,按长宽比为 1.2 选取样本,提高了样本的有效性。实验表明,算法提高了检测率,具有较高的检测速度和较少的误检数,对光照、遮挡和表情具有较强的鲁棒性。

参考文献:

- [1] YANG M H, KRIEGMAN D J, AHUJA N. Detecting faces in images: A survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(1): 34–58.
- [2] 黄福珍, 苏剑波. 人脸检测[M]. 上海: 上海交通大学出版社, 2006.
- [3] YANG J, LU W, WAIBEL A. Skin-color modeling and adaptation [C]// *Third Asian Conference on Computer Vision*. London: Springer-Verlag, 1998: 687–694.
- [4] KAKUMANU P, MAKROGIANNIS S, BOURBAKIS N. A survey of skin-color modeling and detection methods [J]. *Pattern Recognition*, 2007, 40(3): 1106–1122.
- [5] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2001: 511–518.
- [6] LIENHART R, MAYDT J. An extended set of haar-like features for rapid object detection [C]// *Proceedings of IEEE International Conference on Image Processing*. Piscataway: IEEE Press, 2002: 900–903.
- [7] JONES M, VIOLA P. Fast multi-view face detection [EB/OL]. [2009-02-16]. <http://www.merl.com/reports/docs/TR2003-96.pdf>.
- [8] YANG J, WAIBEL A. Real-time face tracker [C]// *Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision*. Washington, DC: IEEE Computer Society, 1996: 142–147.
- [9] HSU R L, ABDEL-MOTTALEB M, JAIN A K. Face detection in color images [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 696–706.
- [10] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of online learning and an application to boosting [C]// *Proceedings of the Second European Conference on Computational Learning Theory*. London: Springer-Verlag, 1995: 23–37.

(上接第 963 页)

- [10] MOGHADDAM B, JEBARA T, PENTLAND A. Bayesian face recognition [J]. *Pattern Recognition*, 2000, 33(11): 1771–1782.
- [11] LI Y X, TAN C L, DING X. Contextual post-processing based on the confusion matrix in offline handwritten Chinese script recognition [J]. *Pattern Recognition*, 2004, 37(9): 1901–1912.
- [12] HO K T, HULL J J, SRIHARI S A. Decision combination in multiple classifier systems [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(1): 66–75.
- [13] ZHAO W, CHELLAPPA R, PHILLIPS P J, *et al.* Face recognition: A literature survey [J]. *ACM Computing Surveys*, 2003, 35(4): 399–458.
- [14] LEE K C, HO J, YANG M H, *et al.* Video-based face recognition using probabilistic appearance manifolds [C]// *CVPR'03: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2003, 1: 313–320.
- [15] O' TOOLE A J, HARMS J, SNOW S L, *et al.* A video database of moving faces and people [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(5): 812–816.
- [16] SIM T, BAKER S, BSAT M. The CMU pose, illumination, and expression database [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(12): 1615–1618.
- [17] VILAPLANA V, MARQUES F, SALEMBIER P, *et al.* Region-based segmentation and tracking of human faces [C]// *EUSIPCO-98: European Signal Processing Conference*. [S. l.]: IEEE, 1998: 311–315.
- [18] NISTER D, STEWENIUS H. Scalable recognition with a vocabulary tree [C]// *CVPR'06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2006: 2161–2168.
- [19] BELLHUMEUR P N, HESPANHA J P, KRIENGMAN D J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711–720.