

融合人脸特征和相关向量机的多姿态人脸检测

阮锦新 尹俊勋

(华南理工大学电子与信息学院, 广州 510640)

摘要 多姿态人脸检测是人脸检测研究领域中的难点和热点之一, 针对这一实际应用中亟待解决的难题, 提出融合人脸特征和相关向量机的检测算法。算法首先利用肤色特征快速排除大部分背景, 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。最后选用分类性能比支持向量机更优的相关向量机对候选区域进行分类。对比实验表明, 算法提高了多姿态人脸的检测率, 对光照、表情和遮挡有较强的鲁棒性。

关键词 多姿态人脸检测 人脸特征 支持向量机 相关向量机

中图分类号 TP391.41; **文献标志码** A

人脸检测是人脸信息处理中的关键技术之一, 有广阔的应用前景。近几年已研究出许多有效的检测方法, 特别是在较理想的条件下, 正面的人脸检测已取得令人满意的效果。然而, 对于多姿态、遮挡、表情和光照等情况效果并不理想, 仍需进一步研究^[1]。

根据利用人脸知识的不同, 人脸检测方法可分为基于人脸特征和基于图像两大类。

人脸肤色是人脸的重要特征之一, 人脸肤色空间只占整个色度空间中很小的一部分, 有较强的聚类特性。在彩色图像中, 利用肤色特征进行人脸检测预处理, 能有效缩小检测范围, 而且算法不易受姿态、表情和旋转等因素影响, 具有相对稳定、算法简单、运行快等优点。因此, 利用肤色特征的人脸检测算法得到广泛应用。Jain 等人^[2]利用肤色特征进行粗检, 根据两眼和嘴巴的几何关系对人脸候选区域进行确认, 对多姿态人脸检测具有很高的鲁棒性。但是, 当眼睛或嘴巴被遮挡时, 容易出现漏检。而且, 在复杂背景下, 算法易受所采用的肤色模型及肤色分割效果的影响, 常常出现误检。

基于图像的人脸检测方法, 主要有神经网络、AdaBoost 和支持向量机 (Support Vector Machine, SVM) 等方法。AdaBoost 算法^[3]在人脸检测领域取得突破性进展, 在正面的人脸检测方面检测率高、速度快, 但由于算法主要考虑水平矩形特征, 对多姿态人脸检测的检测率不高。SVM^[4]是建立在完备的统计学习理论和结构风险最小化原理基础上的机器学习方法, 对小样本的统计估计和预测学习具有出色的性能, 其泛化性能优于神经网络等传统的学习方法, 在人脸检测方面得到广泛重视和应用。相关向量机 (Relevance Vector Machine, RVM) 是 Tipping 在 2000 年提出的一种与 SVM 模型相似、更具稀疏性的概率预测模型, 核函数的选择不受 Mercer 条件限制, 在训练过程中不需要对参数进行调整, 具有更好的泛化能力。

虽然采用基于图像的方法对人脸候选区域进行分类, 可以有效提高检测率。但是, 对多姿态人脸检测, 特别是对旋转人脸的检测鲁棒性不强, 误检数高。如果增加姿态检测分类器, 就会增加算法的复杂度, 影响检测速度。

眼睛、嘴巴也是人脸的明显特征, 利用其几何特征能快速确定人脸方向, 根据人脸方向分割出大致正向的人脸候选区域, 可保证基于图像的人脸检测方法对多姿态人脸检测的检测效果。

2009年12月24日收到

第一作者简介: 阮锦新(1971—), 男, 博士生, 研究方向: 人脸检测、表情识别。E-mail: ruanjinxin@163.com。

本文在增强样本有效性的基础上,针对复杂背景的彩色图像,结合肤色、眼睛、嘴巴等人脸特征和 RVM 算法,实现多姿态人脸检测,得到较好的检测效果。

1 多姿态人脸检测算法流程图

多姿态人脸检测算法(见图 1)主要由图像预处理、肤色区域检测、人脸候选区域分割、样本采集和 RVM 训练与分类等组成。

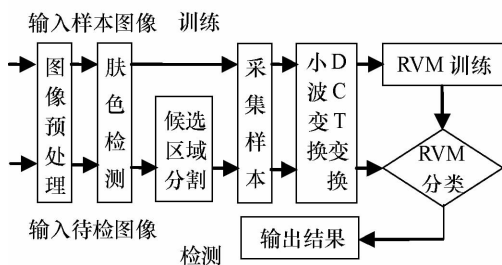


图 1 多姿态人脸检测算法流程图

2 肤色区域检测

2.1 色彩空间转换

不同种族的人在肤色上主要是亮度的差别,而色度的差别是很小的。在 RGB 色彩空间中各分量均含亮度信息,相关性大,易受光照变化影响。 YC_bC_r 是亮度分离的色彩空间,广泛应用于 MPEG、JPEG 等视频压缩,而且可快速从 RGB 色彩空间转换得到。广泛应用于 MPEG、JPEG 等视频压缩。而且,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。

YC_bC_r 色彩空间可通过 RGB 色彩空间按式(1)线性转换得到。为了减少光照变化的影响,采用分段非线性色彩空间转换^[2],将 YC_bC_r 色彩空间转换成 $YC'_bC'_r$ 色彩空间。

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1678 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

2.2 肤色区域检测

对样本图像、待测图像进行光照补偿和去噪预处理后,在 $YC'_bC'_r$ 色彩空间,根据椭圆聚类方法^[2],若像素满足式(2),则被认为是肤色像素,记为 1,否则,认为是非肤色像素,记为 0,这样可得到一个二值图像,再经过形态学处理,可以较准确检测出肤色区域。根据肤色连通区域的几何特征,如区域的长、宽、肤色所占比例等,把明显不可能是人脸的区域去掉。

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} < 1 \quad (2)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} C'_b - c_x \\ C'_r - c_y \end{bmatrix} \quad (3)$$

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

3 人脸候选区域分割

3.1 眼睛和嘴巴区域检测

眼睛的虹膜与皮肤有较大差异,是明显的非肤色区域。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。眼睛的灰度值比较低, Y 分量集中在 $(0, 120)$ 内, C_b 分量普遍比 C_r 分量的值要高。将眼睛的色度和亮度结合起来^[2],能准确检测出眼睛区域的大小和位置。

嘴巴也是明显的非肤色区域,其方向常常与人脸方向一致,具有很强的稳定性。与其它器官相比,嘴巴的 C_r 分量更高, C_b 分量更低,增大 C_r 和 C_b 分量的差异^[2],能准确检测出嘴巴区域的大小和位置。

3.2 人脸候选区域分割

在同一个肤色连通区域中,判断眼睛和嘴巴的上下位置关系,计算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,计算每组夹角的平均值。由于可能存在头发、眼镜和口红等遮挡,影响检测

效果,所以,为了增强对遮挡等影响的鲁棒性和避免漏检,当只搜索到嘴巴或一只眼睛区域时,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。

4 人脸候选区域分类

4.1 相关向量机

RVM 是基于稀疏贝叶斯理论的学习方法,其训练是在贝叶斯框架下进行的,适用于回归问题和分类问题,具有概率预测的优点,其稀疏性和泛化性能等优于 SVM。因此,在人脸检测、移动目标识别、图像检索等领域具有广阔的应用空间。随着 RVM 算法的深入研究,针对算法的计算复杂度高、训练时间比 SVM 长等问题,提出了许多有效的改进^[5,6]。

设 RVM 的模型输出为:

$$y(x;w) = \sum_{i=1}^N w_i K(x, x_i) + w_0 \quad (4)$$

式(4)中, $w = (w_0, w_1, \dots, w_N)^T$ 是可调参数或权值, $K(x, x_i)$ 是核函数。

在贝叶斯框架下,权值 w 可通过最大似然法训练获得,但如果不对权值进行约束,将导致过学习。因此,RVM 为每个权值定义了高斯先验概率分布:

$$p(w | \alpha) = \prod_{i=0}^N N(w_i | 0, \alpha_i^{-1}) \quad (5)$$

式(5)中, α 为 $N+1$ 维超参数。

对于二分类问题,给定训练样本集 $\{x_n, t_n\}_{n=1}^N$, $x_n \in R^d$ 是训练样本向量, $t_n \in \{0, 1\}$ 是相应的预测目标值, N 是样本数。将 Logistic sigmoid 连接函数 $\sigma(y) = \frac{1}{(1 + e^{-y(x;w)})}$ 应用于 $y(x)$, 并设 $P(t|x)$ 服从伯努利分布,则数据集的似然函数为:

$$p(t | w) = \prod_{n=1}^N \sigma\{y(x_n;w)\}^{t_n} [1 - \sigma\{y(x_n;w)\}]^{1-t_n} \quad (6)$$

在分类问题上, $p(t|w)$ 无法解析地积分出权值,因此,采用拉普拉斯方法逼近:

1) 对当前固定的 α 值,求“最大可能”权值 w_{MP} 。因为 $P(w|t, \alpha) \propto P(t|w)P(w|\alpha)$, 所以通过

式(7),采用二次牛顿算法,得到最大值 w_{MP} 。

$$\lg\{P(t | w)p(w | \alpha)\} = \sum_{n=1}^N [t_n \lg y_n + (1 - t_n) \times \lg(1 - y_n)] - \frac{1}{2} w^T A w \quad (7)$$

式(7)中, $y_n = \sigma\{y(x_n;w)\}$ 。

2) 利用拉普拉斯方法,对对数后验概率进行二次逼近。将式(7)求导两次得:

$$\nabla_w \nabla_w \lg(w | t, \alpha) |_{w_{MP}} = -(\Phi^T B \Phi + A) \quad (8)$$

其中, $B = \text{diag}(\beta_1, \beta_2, \dots, \beta_n)$ 是对角矩阵, $\beta_n = \sigma\{y(x_n)\} [1 - \sigma\{y(x_n)\}]$ 。对式(8)右边取负号再求逆,可得协方差矩阵 ε 。

3) 利用高斯逼近的统计量 ε 和 w_{MP} ,对超参数 α 进行更新,直到满足收敛条件:

$$\alpha_i^{\text{new}} = \frac{\gamma_i}{w_{MP}^2} \quad (9)$$

$$\gamma_i \equiv 1 - \alpha_i \varepsilon_{ii} \quad (10)$$

$$\varepsilon = (\Phi^T B \Phi + A)^{-1} \quad (11)$$

$$w_{MP} = \varepsilon \Phi^T B t \quad (12)$$

经足够多的更新,大部分 α_i 趋于无穷大,相应的 w_i 为零,而不为零的 w_i 所对应的学习样本被称为相关向量 (Relevance Vector, RV), 相当于 SVM 中的支持向量 (Support Vector, SV)。对同一训练样本集, RV 数量比 SV 少得多。随着训练样本增大, SV 线性增长,而且比 RV 增长快得多。因此, RVM 比 SVM 更稀疏,所需测试时间更短,更适用于实时检测。

4.2 训练样本选取

基于样本学习的统计方法对训练样本的依赖比较大。传统 SVM 检测方法的样本从 MIT、CMU 等大型的人脸数据库选取,部分非人脸样本常常从风景图像中随机或用自举方法选取,样本图像是灰度图像,而且样本图像的长宽比一般为 1。

人脸的长宽比一般约为 1~1.5,且通常大于 1。另外,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,样本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。

样本从彩色图像的肤色区域选取,保留了有用的色彩信息。同时,非人脸样本限制在肤色区域中选取,提高了样本选择的针对性和训练的有效性。

4.3 样本特征提取

为了减少 RVM 特征向量的维数,对样本做如下处理:(1)样本经过 Haar 小波变换后,选取低频部分。(2)每个样本经过掩模去除四个角的背景干扰。(3)样本图像经 DCT 变换后,对各分量系数进行有区别的选择,选取 Y 分量前 35 个系数, C_b 、 C_r 分量的前 49 个系数共 133 个系数作为样本特征,使样本的特征矢量从 1 440 维降到 133 维。

4.4 人脸候选区域分类

为了能检测出不同大小、位置的人脸,检测时将人脸的候选区域进行金字塔分解。将候选区域以 1.2 倍逐级缩小,直到候选区域大小为 24×20 。每次缩小后,对候选区域用 24×20 的窗口按照从上到下、从左到右扫描,截取彩色图像作为检测样本。检测样本经过小波变换、掩模和 DCT 变换等处理后,输入到 RVM 分类器进行分类判别。

一般情况下,同一人脸在不同比例及邻近位置上会被多次检测。当检测窗口重叠面积达到 70% 以上,则认为这两个窗口代表同一个人脸,将其合并。而被误检的人脸区域重复检测到的次数往往较少。

5 实验结果及分析

在实验中,共选取了 8 500 个正面和侧面人脸样本,以及 9 500 个非人脸样本。大部分人脸表情丰富、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。为了提高算法的可比性,各算法都结合肤色特征进行预检。

AdaBoost 算法采用的样本图像为 24×24 像素的灰度图,选用 5 个 Harr-like 特征(见图 2),由 2 783 个弱分类器组成 23 层级联分类器。SVM 算法采用的样本图像为 24×24 像素的灰度图,选用 RBF 核函数,采用交叉验证选择最佳的惩罚因子 C 和高斯核参数 γ 。而 RVM 算法采用的样本图像为 24×20

像素的彩色图像,选用 RBF 核函数,核参数 $\gamma = 1$ 。



图 2 5 个 Harr-like 特征

采用了 160 张大小为 320×240 像素的彩色测试图像,图像背景复杂,每张图像包含 1 张或多张姿态各异的人脸,共 856 张人脸。

算法基于 Microsoft Visual C++2005 和 Intel Open CV 编程实现,在 Intel 酷睿 2 双核 2.83 GHz CPU、2 GB 内存的台式计算机上运行。检测结果如表 1 所示。

表 1 实验结果

检测算法	SV	RV	误检数	检测率	平均速度 ms/帧
AdaBoost 算法	—	—	679	81.1%	27
SVM 算法	4 087	—	527	86.6%	62
本文算法	—	973	92	93.1%	45

从表 1 可以看出,在检测率方面,本文算法比 AdaBoost、SVM 算法明显提高,误检数少。在检测速度方面,所需相关向量比支持向量少,本文算法速度比 SVM 有较大提高,但是比 AdaBoost 算法慢。图 3 是运用本文算法进行检测的部分实验结果。



图 3 部分实验结果

6 结论

本文把 RVM 算法运用到人脸检测中,作了有益的尝试。按长宽比 1.2 选取彩色样本图像,提高了样本的有效性和分类器的性能。利用小波变换和 DCT 变换,有效减少样本特征向量的维数,加快 RVM 的训练和分类。通过眼睛、嘴巴确定人脸方

向,分割出大致正向的人脸候选区域,选用比 SVM 分类性能更优的 RVM 进行人脸候选区域分类,提高了包括旋转在内的多姿态人脸检测的准确率。实验表明,算法具有较高的检测率和较少的误检数,对光照、表情和遮挡具有较强的鲁棒性,基本达到实时人脸检测要求。下一步工作,重点对算法进行改进,提高算法的运算速度,并在大规模样本数据库中仿真。

参 考 文 献

- 1 黄福珍,苏剑波.人脸检测.上海:上海交通大学出版社,2006
- 2 Hsu R L, Abdel-Mottaleb M, Jain A K. Face detection in color images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002,

24(5): 696—706

- 3 Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington, USA: IEEE Computer Society, 2001: 511—518
- 4 Vapnik V N. *The nature of statistical learning theory*. New York: Springer-Verlag, 1995
- 5 Tipping M E, Faul A C. Fast Marginal Likelihood Maximization for Sparse Bayesian Models. <http://www.miketipping.com/papers/met-fastsbl.pdf>, 2003-01-06
- 6 Yang Z R. A fast algorithm for relevance vector machine. *the 7th Intelligent Data Engineering and Automated Learning*. Berlin Heidelberg: Springer-Verlag, 2006: 33—39

Combining Facial Features and Relevance Vector Machine for Multi-pose Face Detection

RUAN Jin-xin, YIN Jun-xun

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, P. R. China)

[Abstract] Multi-pose face detection has been one of difficult and hot issues in face detection research. For this practical application problem urgently need to be solved, a multi-pose face detection algorithm based on facial features and relevance vector machine algorithm is introduced. Making full use of skin color information firstly, the most background regions can be quickly excluded. After detecting eyes and mouth in the skin color regions, according to face orientation decided by the geometric features of the eyes and mouth region, the approximate frontal face candidates are segmented. At last, the face candidates are classified by relevance vector machine algorithm, which its classification performance is better than support vector machine. The experimental results demonstrate that the algorithm can further improve the multi-pose face detection accuracy and is highly robust to lighting condition, facial expression and occlusion.

[Key words] multi-pose face detection facial feature support vector machine relevance vector machine