

多信息融合的实时人脸检测算法

陈齐松 陈锻生

(华侨大学信息科学与工程学院, 福建 泉州 362021)

摘要 提出一种综合使用灰度、梯度和肤色信息的实时人脸检测方法, 使用类 Haar 特征描述人脸模式的灰度差、梯度差和肤色差, 构造相应的特征集. 用 AdaBoost 算法从特征集中学习区分人脸与非人脸模式的有效规则, 构成人脸检测级联分类器. 实验表明, 综合使用多信息的人脸检测器性能, 比单独使用灰度信息的检测器有显著的提高.

关键词 人脸检测, 类 Haar 特征, 多信息融合, Adaboost 算法, 级联分类器

中图分类号 TP 391.41; TP 301.6

文献标识码 A

近 20 年来, 不断有各种人脸检测技术的检测方案提出^[1], 并先后从检测率和检测速度上取得较大突破. 特别是 Viola 等提出的, 使用 Adaboost^[2] 学习算法构建级联分类器的检测算法^[3], 检测率不低于基于支持向量机^[4] 或神经网络^[5] 等最优秀的检测方法, 但速度比上述两种检测器快 1~2 个数量级. Rainer 等人^[6] 扩展了类 Haar 特征的形状, 引入了倾斜 45° 的类 Haar 特征, 训练出的分类器性能比 Viola 的有所提高. Zhang 等人^[7] 对 AdaBoost 算法进行改进, 使用 FloatBoost 算法减小弱分类器的使用量, 使检测速度有所提高. 本文的算法是在 Viola 检测系统的基础上, 综合利用了人脸模式的灰度、梯度和肤色信息做人脸检测, 在检测率和检测速度上都有较大提高.

1 特征表示与提取

人脸模式包含众多的特征, 充分利用多种特征进行人脸检测, 可能有利于提高检测器的性能. 基于这个想法, 本文的检测方法尝试融合使用 3 种类型特征, 即灰度差、梯度差与肤色差特征. 它们分别是在灰度图、梯度图和肤色距离图上用类 Haar 特征^[3] 描述的. 类 Haar 特征是一种“矩形”对特征, 可用于描述图像模式相邻区域的特征差异, 可用积分图^[3] 快速计算矩形区域的特征值总和. 这就使类 Haar 特征值的计算十分迅速.

1.1 梯度图

图像梯度是通过梯度算子与图像进行卷积运算得到. 对于灰度图像来说, 比较常用的梯度算子有一阶的 Robert 算子、Sobel 算子和二阶的 Laplacian 算子, 而对彩色图像的梯度计算, 也有不同方法^[8]. 出于计算量考虑, 我们设计了最简单的纵向(dx)和横向(dy)梯度算子, 如图 1 所示. 对像素求梯度值时, 取 dx, dy 算子计算得的梯度最大值作为该像素的梯度值, 我们称这种算子为 dmax 算子. 各种算子求取梯度图的方法为

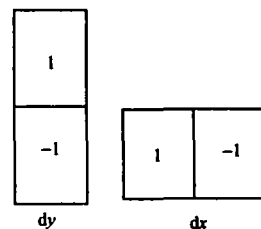


图 1 使用的梯度算子

$$G_x = |dx * A| = |[1, -1] * A|, \quad (1)$$

$$G_y = |dy * A| = |[1, -1] * A|, \quad (2)$$

$$G_{\max}(x', y') = \max(G_x(x', y'), G_y(x', y')), \quad (3)$$

收稿日期 2005-09-11

作者简介 陈齐松(1979-), 男, 硕士研究生, 主要从事模式识别与数字图像处理的研究; 通信作者, 陈锻生(1959-), 男, 教授, E-mail: dschen@hqu.edu.cn

基金项目 福建省科技计划国际合作重点基金资助项目(2004I014)

在式(3)中, $G(x', y')$ 为图像像素在坐标 (x', y') 处的梯度值. 通过实验发现, d_{\max} 算子计算的梯度图生成的梯度差特征比 d_x, d_y 算子更有效. 因为, 前者产生的梯度差特征训练出来的级联分类器, 使用的特征数量最少, 使得在同等的检测率下检测速度最快. 因而, 本文使用 d_{\max} 算子来计算梯度图. 对彩色图像, 我们的处理方法是先将其转化为灰度图, 然后用 d_{\max} 算子计算相应灰度图.

1.2 肤色距离图

利用人脸模式肤色分布的相关性特征, 即在肤色距离图上用类 Haar 特征描述相邻区域肤色特征的差异, 通过 AdaBoost 算法, 从中挑选出最显著的肤色差异特征作检测. 选择一个二维色度模型, 通过实验取得肤色在这个模型的中心值, 将彩色图像的每个像素的值变换到这个色度平面, 计算它们与肤色中心的距离值, 即可得出相应的肤色距离图. 显然, 应该选择一个肤色聚类紧密的色度模型计算肤色距离图. 本文从自制的包含 6 400 个样本的彩色人脸样本集中, 挑选了肤色上有代表性的 118 个人脸样本, 从中取得 296 128 个肤色像素制成肤色样本集. 在该样本集上对肤色在 rg (归一化 RGB), hs (hsv), $CbCr$ ($YCbCr$) 3 种色度平面的上聚类情况进行分析, 发现肤色在 rg 色度平面有最紧密的聚类, 如表 1 所示. 在 hs 色度平面, 肤色的 h 分量值主要处在 $[0, 0.1]$ 和 $[0.9, 1.0]$ 区间, 为使 h 分量值连续, 本文将

表 1 肤色样本在 3 类色度平面中的分布情况

项 目	肤色分布中心的坐标	肤色样本与肤色中心的平均距离(马氏)	与肤色中心距离(马氏)小于 2.0 的样本/(%)
rg 色度平面	$(r, g) = (0.47, 0.31)$	0.825 6	99
hs 色度平面(调整后)	$(h, s) = (0.54, 0.36)$	0.877 1	95
$CbCr$ 色度平面	$(Cb, Cr) = (0.40, 0.60)$	0.867 4	98

h 分量值调整为 $h = (h + 0.5) \bmod 1.0$. 表 1 的实验数据是基于这个调整的. 从表 1 可看出, rg 色度平面在这两个参数上均有最好的性能, 同时从 RGB 到 rg 的变换计算量较小. 因此, 本文选择在 rg 色度平面上计算肤色距离图. 即对彩色图像上的每个像素, 先计算其 (r, g) 值, 然后计算这个 (r, g) 值与肤色中心 (r_0, g_0) 的马氏距离值. 有

$$D_M(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}, \quad \mu = \begin{bmatrix} 0.47 \\ 0.31 \end{bmatrix}, \quad \Sigma^{-1} = \begin{bmatrix} 196.74 & 132.03 \\ 132.03 & 746.71 \end{bmatrix}. \quad (4)$$

在式(4)中, x 表示彩色像素的 (r, g) 值向量, μ 表示肤色在 rg 色度平面上的中心点坐标, Σ 为肤色在 rg 色度平面上的马氏距离分布矩阵. 图像上所有像素的距离值构成其肤色距离图.

2 人脸检测器

本文使用了级联分类器^[3]作为判断窗口类别的二分类器, 它能快速排除大部分非人脸窗口, 加快了检测速度. 级联分类器由层分类器构成, 每个层分类器通过 AdaBoost 算法训练^[3]得到. 利用类 Haar 特征在灰度图、梯度图、肤色图上, 可以描述出相应的灰度差、梯度差、肤色差规则, 构成 3 种类型的规则集. 我们既可以单独使用这 3 种特征集训练分类器, 也可以将它们联合起来使用. 本文分别对单独使用灰度差特征, 联合使用灰度差、梯度差特征, 联合所有的 3 类特征训练出 3 个级联分类器 $G(\text{Gray})$, $GG(\text{Gray} \& \text{Grads})$, $GGS(\text{Gray} \& \text{Grads} \& \text{Skin})$, 来分析比较多特征融合在人脸检测中的利弊. 检测器系统分为训练子系统与检测子系统. 前者通过训练取得分辨人脸/非人脸模式的级联分类器, 后者利用级联分类器进行人脸检测. 3 种检测器系统的结构类似, 其中 GGS 检测系统工作原理, 如图 2 所示. 使用了包括普渡大学、CIT, FBI, IMDB 人脸库及从网上收集图像自建的人脸库, 制作出

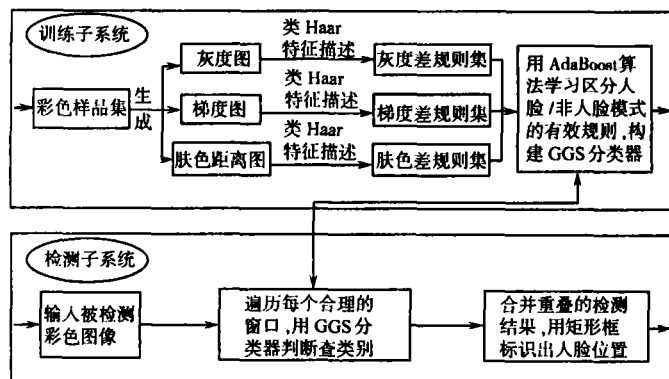


图 2 使用 GGS 级联分类器的人脸检测器工作原理图

包含 6 400 个样本的彩色人脸样本集. 样本是正面无深度旋转的, 因而训练出来的级联分类器适合于正面无深度旋转的人脸检测. 在训练级联分类器时, 非人脸样本用 Bootstrap 法, 从不包含人脸的背景图中通过程序自动切取, 本文使用了约 12 000 张背景图.

3 实验结果

3 种级联分类器皆由 20 个层分类器组成, 在训练时设置层分类器最小检测率为 0. 999 5, 最大误警率为 0. 5. 每个级联分类器使用的弱分类器数越少, 则表明弱分类器越有效, 用于生成弱分类器的特征集越有效. 各个级联分类器使用的弱分类器数, 如表 2 的所示.

表 2 3 种级联分类器的弱分类器使用数对比

分类器	第 i 层								总数
	1	2	3	5	7	10	16	20	
G	5	13	16	28	37	67	197	289	2 094
GG	3	9	13	19	27	45	97	187	827
GGs	3	6	7	9	11	21	48	85	660

从表 2 可以看出, 加入梯差特征后, GG 分类器比 G 分类器使用的弱分类器数下降了不少一半. 进一步加入肤色差特征的 GGS 分类器, 比 GG 分类器使用了更少的弱分类器. 这是由于训练 GGS 分类器使用的特征数量(每个特征对应一个弱分类器)比 G 分类器多得多(约 3 倍), 因而功能较强的弱分类器也就相应更多, 使得只用较少数量的弱分类器出能构造出满足要求的级联分类器. 级联分类器的分类速度由前面几个层分类器使用的弱分类器数目决定(如前面 5 层可排除约 97% 的非人脸窗口), 显然数目越小, 分类速度越快. 对 G, GG 分类器, 我们使用由灰度图构成的 MIT-CMU 测试集作测试. 其中包含较多极模糊的图像和一些旋转角度较大的人脸, 以及手工绘制的人脸轮廓图. 在 130 张的测试图像中, 含有 507 张人脸. GGS 分类器必须在彩色图上作测试. 由于得不到可用的知名彩色测试集, 本文自建一个彩色图像测试集, 对比 3 种分类器在彩色图像检测中的性能. 该彩色图像测试集包含 50 张从互联网下载的各种环境下彩色图像, 共有 342 个人脸.

在该 MIT-CMU 测试集上, 检测器一共检测了 63 887 257 个窗口. G 分类器检测到 436 个人脸, 误警的窗口数为 78 个, 平均每窗口使用了 17 个弱分类器进行判断(检测速度与平均每窗口使用的弱分类器数成反比). GG 共检测到 452 个人脸, 误警的窗口数为 52 个, 平均每窗口使用了 9 个弱分类器进行判断, 如表 3 所示. 表中, η 为检测率, ϵ 为误警率. 在上述的彩色图测试集中, 对 3 个检测器作测试. 检测的窗口数为 28 394 723 个. GGS 分类器检测到的人脸数为 325 个, 误警窗口数为 16 个, 平均每窗口使用了 10 个弱分类器进行判断; GG 分类器检测到的人脸数为 321 个, 误警窗口数为 29 个, 平均每窗口使用了 11 个弱分类器进行判断; G 分类器检测到的人脸数为 311 个, 误警窗口数为 34 个, 平均每窗口使用了 19 个弱分类器进行判断(表 3).

表 3 分类器在测试集上的测试结果

分类器	MIT-CMU 测试集		
	$\eta/(%)$	$\epsilon \times 10^{-5}/(%)$	判别窗口平均使用的弱分类器数
G	85. 9	12. 0	17
GG	89. 2	8. 7	9

分类器	彩色图像测试集		
	$\eta/(%)$	$\epsilon \times 10^{-5}/(%)$	判别窗口平均使用的弱分类器数
G	90. 9	12. 0	19
GG	93. 9	10. 0	11
GGs	95. 0	5. 6	10

抽掉级联分类器最后几个层分类器, 可使得级联分类器的检测率略有上升, 同时误警窗口数增加. 通过这种办法, 我们取测试了 3 种分类器在该彩色测试集上的 ROC 曲线, 如图 3 所示. 测试结果表明, 3 种分类器的检测率/误警率性能从好到次顺序: GGS, GG, G. 与只使用灰度差特征的 G 分类器(Viola-Jones 的检测方法)相比, 加入肤色差和梯度差特征的 G, GGS 分类器的检测率分别约提高了 3%, 5%, 在速度上, 后两者也比前者快约一倍. 因而, 梯度差和肤色差特征的使用皆有利于提高分类器的分类性能, 联合使用这灰度差、梯度差和肤色差特征作训练, 的确能大大提高分类器的性能.

使用 GGS 检测器作视频人脸检测, 视频窗口大小(像素)为 320×240, 最小检测窗口大小(像素)为

72×72. 在 P4 1.7G 的 PC 机上, 一帧图像的检测时间约为 40 ms, 也就是说每秒钟能检测 25 帧图像, 达到了视频实时检测的速度要求. 若对程序做进一步的优化, 检测速度将更快.

4 结束语

本文研究综合利用多种信息实时检测人脸的方法, 在灰度图、梯度图和肤色距离图上, 分别用类 Haar 特征描述人脸模式的灰度差、梯度差和肤色差特征. 分别对单独使用灰度差特征, 联合使用灰度差和梯度差特征, 联合使用灰度差、梯度差和肤色差特征训练出 3 种级联分类器 G, GG, GGS. 通过测试发现, GG 分类器比 G 分类器快约一倍, GGS 分类器略快于 GG 分类器; 在误警率相等情况下, GG, GGS 分类器的检测率分别比 G 分类器高出约 3%, 5%. 将 GGS 分类器用于视频人脸检测, 每秒钟能检测 25 帧 320 像素×240 像素大小的图像, 达到的视频实时检测的要求. 测试结果表明, 无论在速度上还是检测率上, 联合使用灰度、梯度和肤色信息的级联分类器, 性能比只使用灰度信息的 Viola-Jones 方法的级联分类器有较大的提高.

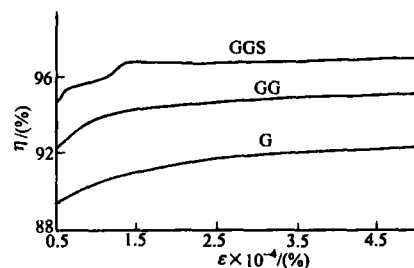


图3 3种分类器在自制彩色测试集上的 ROC 性能图

参 考 文 献

- 1 Yang M H, Kriegman D, Ahuja N. Detecting faces in images: A survey[J]. IEEE, Trans Actions on Pattern Analysis and Machine Intelligence, 2002, 24(1): 34~58
- 2 Freund Y, Schapire R E. A short introduction to boosting[J]. Journal of Japanese Society for Artificial Intelligence, 1999, 14(5): 771~780
- 3 Viola P, Jones M. Robust real-time object detection[J]. International Journal of Computer Vision, 2004, 57(2): 137~154
- 4 Osuna E, Freund R, Girosi F. Training support vector machines: An application to face detection[J]. IEEE, Conference on Computer Vision and Pattern Recognition, 1997, 130~136
- 5 Rowley H A, Baluja S. Neural network based face detection[J]. IEEE, Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(1): 23~38
- 6 Lienhart R, Maydt J. An extended set of Haar-like features for rapid object detection[J]. IEEE ICIP, 2002, (1): 900~903
- 7 Zhang Zhenqiu, Li Mingjing. Multi-view face detection with floatboost[J]. WACV, 2002, 184~188
- 8 陈锻生, 刘政凯. 彩色图像边缘特征及其人脸检测性能评价[J]. 软件学报, 2005, 16(5): 727~732

Real Time Face Detection Algorithm Based on Multi-Information Fusion

Chen Qisong Chen Duansheng

(College of Information Science and Engineering, Huaqiao University, 362021, Quanzhou, China)

Abstract In this paper, we put forward a method of real time face detection that using gray level, gradient and skin color information. Haar-like features describing the differences of gray level, gradient and skin color of face pattern are used to construct the feature set for face detection. Efficient rules differentiating face and non-face pattern are learned with AdaBoost algorithm, and a cascade classifier is constructed to detect faces. Experiments indicate that the feature of the face detector using multiple-information is superior on the face detector using only gray level information.

Keywords face detection, Haar-like feature, multi-information fusion, adaboost algorithm, cascade classifier