

基于矩形特征和改进 Adaboost 的手势检测

张秋余, 姚开博, 吴佩莉

(兰州理工大学计算机与通信学院, 兰州 730050)

摘要: 为了实时、精确地从视频流中检测和识别出特定手势, 提出一种矩形特征描述手势, 给出快速计算方法和手势类可分离性的评价方法。为了避免分类器的过度训练问题, 提出一种基于此方法的改进的 Adaboost 算法。实验结果表明, 矩形特征能够产生可靠的检测器, 对手势的姿态变化较敏感。在摄像头实时捕获视频中, 其检测手势实时性较好, 对复杂背景和噪声有较强的适应能力, 当手势旋转角度小时, 正确检测率可以达到 95% 以上。

关键词: 矩形特征; 改进 Adaboost; 复杂背景; 手势检测; 手势识别

Hand Detection Based on Rectangle Features and Improved Adaboost

ZHANG Qiu-yu, YAO Kai-bo, WU Pei-li

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050)

【Abstract】 In order to detect and recognize hand gesture from video sequence in real time accurately, a set of rectangle features are used to describe the hand characteristics, and the fast calculation methods of their features and the evaluation ways of the separability of hand gesture class are given. The Adaboost algorithm is improved to deal with the excessive training. Experimental results show that rectangle features can obtain the reliable detector and has a good real-time performance and adaptive capacity in complex backgrounds, but it is more sensitive to the changes of hand gesture. When hand gesture rotates small angle, the detection rates are 95% above.

【Key words】 rectangle features; improved Adaboost; complex ground; hand gesture detection; hand gesture recognition

1 概述

手势检测是从静止图像或视频帧中通过预处理、提取特征、匹配特征, 最后判断图像中是否存在待检测的手势, 并给出每个手势在图像中的位置和范围。目前, 从视频中进行手势检测的研究工作大部分是在有约束场景中进行的。基于肤色特征的手势检测要求手能从同色目标中区分出来, 受光线不足和彩色光的影响。动作流信息能弥补肤色检测缺陷, 但对于不稳定的摄像机, 实现的难度增加, 可信度降低。手位置统计信息用作先验概率时有效, 但它要求针对特定的应用进行预先训练。在背景与目标之间有足够差异时, 形状模型一般表现较好, 但在凹体和杂乱背景中也存在问题。可捕获深度信息或热红外图像的摄像机能提供更多消息, 但这需要专门的昂贵硬件。

另一方面, 脸部检测引起了研究人员大量的兴趣, 许多方法依赖于形状、纹理和时间信息, 特别是基于纹理的方法, 由于能处理静止和部分闭包对象, 在变化的图像场景中, 可以产生更好的检测效果。

基于视觉的人机手势界面要求满足实时性、精确性和鲁棒性, 构造一个在背景变化、光线不敏感, 人-摄像机独立的分类器, 去可靠地检测手势具有重要意义。2001年, Paul Viola 和 Michael Jones 提出了一种基于 AdaBoost 的人脸检测算法^[1], 同时建立了第 1 个真正实时的人脸检测系统, 从根本上解决了检测速度的问题, 同时也有较好的识别效果。

2 手势特征描述、快速计算以及类可分离性评价

1998年, Papegeorgiou 在应用 Haar 小波变换从人脸图像中提取特征时, 提出了局部 Haar 特征的概念。2001年 Viola 和 Jones^[1] 研究了最快和高精确的基于学习模式识别方法用于

单目视觉的脸部检测方法。2005年, Wu Bo 等^[2] 为了提高计算效率和降低检测的虚假正确率, 开展了多视角的脸部快速检测方法研究。

2.1 矩形特征定义与快速计算方法

局部 Haar 特征^[3] 的定义是黑色矩形和白色矩形在图像子窗口中对应区域的灰度级总和之差。它反映了图像局部的灰度变化, 包括边界、线、对角线特征。本方法中只用了 Haar 特征中边界、细线特征, 另外加入一种新的特征, 矩形的位置和大小没有限制, 而且可以叠加, 更加灵活, 这里总称为矩形特征。如图 1 所示, 其中图 1(a)、图 1(b) 为边界特征; 图 1(c)、图 1(d) 为线特征; 图 1(e) 为新类型(矩形的位置和大小没有限制, 可重叠)。

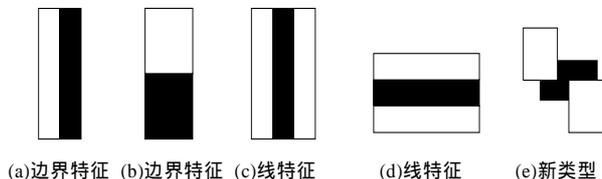


图 1 矩形特征原型

这种矩形特征能够通过积分图在常数时间内快速计算出来。如图 2 所示, 在点 $P(x, y)$ 位置的积分图定义为从点 $(0, 0)$ 至 $P(x, y)$ 的矩形区域内的所有像素灰度值和, 即

基金项目: 甘肃省自然科学基金资助项目(ZS021-A25-018-G)

作者简介: 张秋余(1966 -), 男, 副研究员, 主研方向: 图像处理, 模式识别, 多媒体通信; 姚开博, 助理研究员、硕士研究生; 吴佩莉, 工程师、硕士研究生

收稿日期: 2007-08-30 **E-mail:** eagleflying2007@163.com

$IIS(x,y) = \int_0^x \int_0^y I(x',y') dx' dy'$ ，由此计算公式建立积分表。

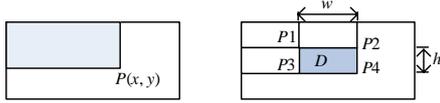


图2 矩形区域面积表计算

积分图可以通过查积分表用迭代方式快速计算求得

$$IIS(x,y) = IIS(x-1,y) + IIS(x,y-1) - IIS(x-1,y-1) + I(x,y)$$

其中， $IIS(-1,y) = IIS(x,-1) = 0$ 。

通过查表，快速计算矩形 $D(x,y,x+w,y+h)$ 内灰度值和

$$RS(x,y) = IIS(x+w,y+h) + IIS(x+w,y) + IIS(x,y+h) - IIS(x,y)$$

为了在检测阶段更容易评估特征值，特征值应规格化为

$$\bar{I}(x,y) = \frac{I(x,y) - \mu}{c\sigma}, c \in \mathbb{R}^+, \text{ 其中, } \mu \text{ 为图像均值; } \sigma \text{ 为方差; } c \text{ 为常数。}$$

一幅 $W \times H$ 大小的图像，矩形特征原型窗口大小为 $w \times h$ ，对每个矩形特征原型会产生特征数为

$$NF_i(W,H,w,h) = \frac{W}{w} \times \frac{H}{h} \times \frac{(W-w+1)}{2} \times \frac{(H-h+1)}{2}$$

图像总的特征数为 $TNF = \sum_{i=0}^{N-1} NF_i(W,H,w,h)$ ，其中， N 为矩形特征原型数。

2.2 手势类可分离性评价

为了减少训练每一个可能手势检测器时所用资源代价，也为了找到有最好表现变化的手势，用一种方法评估每种手势的分类潜力^[4]。先收集每种具有相似表现的手势若干张图片，计算均值图像，其大小为 $W \times H$ ，图像矩阵坐标为 (w,h) ，进行Fourie变换得

$$F(u,v) = \frac{1}{W \times H} \sum_{w=0}^{W-1} \sum_{h=0}^{H-1} I(w,h) e^{-i2\pi(\frac{hu}{H} + \frac{wv}{W})}$$

其中，性色图像的傅氏变换为

$$P(u,v) = \frac{1}{W \times H} \sum_{w=0}^{W-1} \sum_{h=0}^{H-1} \frac{1}{2} e^{-i2\pi(\frac{hu}{H} + \frac{wv}{W})}$$

再计算变换差值并取对数得： $D(u,v) = 1b|F(u,v) - P(u,v)|$ ，将来自图像裁切的频率影响消除掉，产生一个差分变换。最后

计算标准化的幅值总和： $s = \exp(\frac{1}{k} \sum_{u,v} D(u,v))$ ，它是一个图像中手势表现变化量的指标，值越大，手势表现的灰度级变化数量越大，反映了类的可分类能力，即具有越大 s 值的手势能更可靠地被检测，从而有效地指导 Adaboost 算法级联分类器的设计训练。

3 改进的 Adaboost 算法

用一个包含一个矩形特征的子窗口扫描图像，即使对于同一个手势，在不同旋转角度情况下，也会有不同的特征值，如图3所示。



图3 同一手势不同旋转角度的不同矩形特征值

基于每一矩形特征 f_j ，定义一个弱分类器 $h_j(x)$ ，其学习算法如下：

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j < p_j \theta_j \\ 0 & \text{otherwise} \end{cases}$$

其中， p_j 取不同符号改变不等号方向。

3.1 AdaBoost 算法原理

1997年，Freund和Schapire提出了AdaBoost算法^[5]，这是一种迭代机器学习算法。每个训练样本都被赋予一个权值，表明该样本被某个弱分类器选中的概率。每轮训练选出一个具有最小分类错误的弱分类器，由这些弱分类器的线性组合构成一个总的强分类器。

3.2 存在问题

Adaboost 算法的权值更新规则可保证学习算法集中处理比较困难的训练样本，但是如果训练样本集包含噪声样本或者其他复杂的非目标样本，将会给这些样本分配较高的权值，最终可能导致过匹配现象，降低算法性能。

3.3 改进的 Adaboost 算法

对原算法的权值更新规则加以改进：在每轮训练中，定义一个阈值 ϵ_t ，结合样本是否被错误分类以及当前权值是否大于 ϵ_t ，为样本更新权值，即

$$\omega_{t+1}(x_i) = \frac{\omega_t(x_i)}{\sum_{i=1}^N \omega_t(x_i)} \times \begin{cases} \beta_t & \text{if } h_t(x_i) = y_i \\ 1/\beta_t & \text{if } h_t(x_i) \neq y_i, \text{ and } \omega_t(x_i) > \epsilon_t \\ \beta_t & \text{if } h_t(x_i) \neq y_i, \text{ and } \omega_t(x_i) < \epsilon_t \end{cases}$$

其中， $\beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$ ； ϵ_t 是第 t 轮训练中的权值更新阈值，取

该轮训练中所有样本权值的均值作为 ϵ_t ，即 $\epsilon_t = \frac{\sum_{i=1}^N \omega_t(x_i)}{n}$ 。

这样，只有当某个样本被错误分类，且当前权值 ω_t 小于该轮的权值更新阈值 ϵ_t 时，样本的权值才会被增加，否则，其权值都将被减小。这样，即使困难样本在每轮都被错误分类，它们的权值也不会被过分增大，在一定程度上避免了过匹配现象的发生。在此基础上设计分类器。

N 轮训练结束后，最终得到一个强分类器：

$$H(x) = \text{sign}(\sum_{t=1}^N (-\ln \beta_t) h_t(x))$$

4 试验过程、结果及分析

硬件环境为 Celeron 2.8 GHz, 512 MB, USB1.0 Camera；软件环境为 Windows XP, VC++2005, Intel OpenCV1.0；参数控制：特征窗口大小为 25×25 ，比例因子为 1.2，分类器层数设置为 15 层，每层级联分类器可接受的最小检测率为 0.995，最大误检率为 0.4。整个级联分类器的检测率可以达到 0.93，误检率为 1.07×10^{-6} 。采用 Visual C++ 编程，利用 OpenCV library 和 Viola-jones 方法的修改版本进行相应的训练、测试工作。

4.1 同种手势不同旋转角度的分类器训练试验

试验对象：先收集正样本，类似图3中的 Open 手势，手势主要从 Jochen Triesch Static Hand Posture Database II 中选取，取不同复杂背景、不同光线亮度、多位试验者、同一视角的手势图 450 张，其中，350 张作训练样本，100 张作测试样本，此为正样本集 A；再依次对 450 张图旋转 $\pm 30^\circ$ ， $\pm 60^\circ$ ， $\pm 90^\circ$ ，则又得到 6 套不同视角的手势集 B, C, D, E, F, G。统一规格化为 25×25 大小。再通过在网收集和在 Photo CD 素材库中取不包含手势的任何图片 1000 张作负样本。试验结果如图4所示。初步结果显示平均检测正确率为 65%，最高为 83%，最小为 36.7%，最大旋转角度的检测正确率较低，说明训练集没有包含足够的样本变化，这是由于本文的训练样本规模不大。虚假正确率大约 7.7%，结果较好，这说明矩形特征能够产生可靠的检测器，但是对手势的姿态变化较敏感。非不变特征的局限性与不同视角的对象表现不同有

关, 这可以通过训练多级或并行分类器, 处理包含相同目标的各种变化类型图像来克服, 但这样会增加相应训练时间和检测时间。

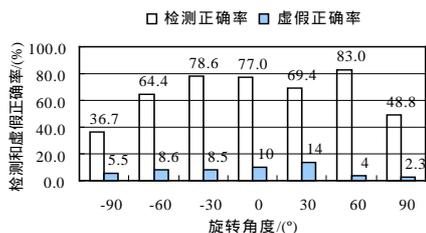


图4 不同旋转角度的手势测试结果

4.2 单手势实时检测试验

进一步作旋转度在 $\pm 10^\circ$ 以内的 open 单手势, 正样本包含光线亮度、复杂背景变化很宽范围的场景 2 300 张, 负样本 3 000 张, 训练 20 小时以上, 得 15 级的级联分类器。用摄像头实时捕获视频帧图像进行实时检测试验, 初步结果表明, 处理一幅 320×240 的视频帧图像平均时间为 47 ms, 平均速度为 21 帧/s, 处理 160×120 帧图像平均需 18 ms, 速度可达 50 帧/s, 具有实时性和明显的实际应用价值。

图 5 示出了单手势实时检测试验结果。



图5 手势视频实时检测和图片检测结果例子

对于图像质量好而且手势旋转角度小的图片, 正确检测率可以达到 95% 以上。初步结果表明, 如果正样本覆盖了光

(上接第 175 页)

协议的计算开销包括证书验证、模指运算和杂凑运算, 每一方都包括验证证书的运算, 以及 2 次模指运算(计算 SK 时可以利用前 2 次的计算结果)和 3 次杂凑运算, 如果证书在先前的通信中已经验证过, 还可以进一步减少运算量。这种计算开销和计算的对称性使得协议适合参与双方处理能力有限的无线通信环境。

在通信开销方面, 由于 g^x 和 g^y 的范围已经确定, 长度小于 $1bq$ bit, 杂凑值的长度也是固定的, 因此这种通信开销是合理的。

通信方的存储开销包括需要的算法和数据。协议中使用了公钥密码算法和杂凑算法, 但避免了文献[2]中的对称密钥算法, 数据主要是保存的自身和对方的数字证书, 用户可以根据自身的存储能力选择保存证书的数量。

综上所述, 协议在计算开销、通信开销和存储开销等方面具有良好的性能和实用性。

5 结束语

本文提出无线网络环境下基于证书的密钥协商协议, 描述了协议的实现并分析了协议的安全性, 证明在 RO 模型下

线亮度、背景变化很宽的场景, 训练出的分类器在识别过程中对背景干扰、各种噪声等都有较强的适应和处理能力, 有很好的鲁棒性。

5 结束语

基于矩形特征的手势检测方法因使用积分图快速计算矩形特征值而具有较快的速度, 虽然训练过程比较复杂费时, 但训练出的多层分类器简单有效。该方法也可作为少量手势的快速跟踪和识别方法。利用改进 AdaBoost 算法训练分类器, 获得了很高的检测速度和检测正确率。对包含光线亮度、复杂背景变化很宽范围的场景手势样本进行训练, 试验表明该方法在识别过程中对背景干扰、噪声有较强的适应和处理能力, 训练出的分类器快速、可靠、鲁棒性好。基于矩形特征的识别方法对手势的姿态变化较敏感, 未来研究包括有效地处理旋转的算法、基于肤色的直方图预处理和应用图像移动特征, 以提高手势识别的性能和精确度。

参考文献

- [1] Viola P, Jones M. Robust Real Time Object Detection[C]//Proc. of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: [s. n.], 2001.
- [2] Wu Bo. A Multi-view Face Detection Based on Real Adaboost Algorithm[J]. Computer Research and Development, 2005, 42(9): 1612-1621.
- [3] Lienhart R, Maydt J. An Extended Set of Haar-like Features for Rapid Object Detection[C]//Proc. of ICIP'02. [S. l.]: IEEE Press, 2002-09: 900-903.
- [4] Kölsch M, Turk M. Robust Hand Detection[C]//Proc. of IEEE Int'l Conference on Automatic Face and Gesture Recognition. [S. l.]: IEEE Press, 2004-05.
- [5] Freund Y, Schapire R E. A Decision-theoretic Generalization of Online Learning and an Application to Boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.

本协议的安全性基于 CDH 问题。本协议不但是可证明安全的, 而且计算开销, 通信开销和存储开销方面都适合对等实体在移动环境下进行身份认证和协商密钥, 实用价值较高。

参考文献

- [1] Bellare M, Rogaway P. Random Oracles Are Practical: A Paradigm for Designing Efficient Protocols[C]//Proc. of the 1st ACM Conf. on Computer and Communications Security. New York, USA: ACM Press, 1993.
- [2] 邓红素, 左益强, 赵一鸣, 等. 移动通信中可证安全的双向认证密钥协商协议[J]. 软件学报, 2003, 14(8): 1489-1494.
- [3] 周永彬, 张振峰, 冯登国. 一种认证密钥协商协议的安全性分析及改进[J]. 软件学报, 2006, 17(4): 868-875.
- [4] Wilson S B, Johnson D, Menezes A. Key Agreement Protocols and Their Security Analysis[C]//Proc. of the 6th Int'l Conf. on Cryptography and Coding. Berlin, Germany: Springer-Verlag, 1997.
- [5] Bellare M, Rogaway P. Entity Authentication and Key Distribution[C]//Proc. of CRYPTO'93. Berlin, Germany: Springer-Verlag, 1993.