

基于背景差与 HOG-LBP 的快速运动人体检测

张 慧,郑爱华,涂铮铮,罗 斌

(安徽大学 计算机科学与技术学院,安徽 合肥 230601)

摘 要:针对单目静止摄像机近距离监控的情形,结合运动目标外接矩形长宽比,提出一种 HOG 特征联合 LBP 特征并通过 PCA 降维的快速运动人体检测算法。该方法包含两个步骤:运动目标提取和运动人体检测。使用帧差与背景差相结合的方法提取运动目标,帧差用于更新背景,背景差用于提取运动目标。运动目标判别即人体检测分为两个部分:单运动人体检测以及多运动人体检测。首先根据运动目标外接矩形的长宽比,把目标分为单目标以及多目标;然后,根据肤色的分布判断单个行人。对于多目标,提取 HOG-LBP 特征,用 PCA 降维,结合线性 SVM 进行群目标判定。实验结果表明,该方法不仅提高了人体检测速度,还提高了人体检测率。

关键词:运动人体检测;梯度方向直方图;背景差;局部二元模式

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2015)02-0047-05

doi:10.3969/j.issn.1673-629X.2015.02.011

Rapid Motion Human Detection Based on Background Subtraction and HOG-LBP

ZHANG Hui,ZHENG Ai-hua,TU Zheng-zheng,LUO Bin

(College of Computer Science and Technology, Anhui University, Hefei 230601, China)

Abstract:For the case of close monitoring by monocular static camera, propose a rapid motion human detection algorithm combined with target movement external rectangular length-width ratio, including two steps of moving objects extraction and moving human detection. The moving target extraction is implemented by combining the frame difference and background subtraction, where the frame difference is used to update the background and background subtraction is used to extract the moving target. Moving target discrimination that is human detection can be divided into two parts, the single movement human detection and the crowd detection. Firstly, according to the empirical value of length-width ratio of bounding rectangle of the extracted moving target, the target is divided into single-object and multi-object. Then, single-object is determined according to the skin color distribution. For multi-object, HOG-LBP features are extracted, followed by PCA (Principal Component Analysis) dimensionality reduction. Then the multi-object discrimination is approached by combining with the linear SVM. The experimental results show the remarkable performance of this method is improved on both detection rate and efficiency.

Key words: motion human detection; Histogram of Oriented Gradient (HOG); background subtraction; Local Binary Pattern (LBP)

0 引 言

运动人体检测在智能视频监控应用广泛,是计算机视觉和模式识别领域中的研究重点之一。由于人体是非刚性的目标,受到姿态、外表、视角和衣着、光照变化的影响,加之复杂的背景场景,使得行人检测成为一个难点^[1]。

运动人体检测总体来说可以分为基于静态图像的人体检测以及基于视频的运动人体检测两类。基于静

态图像的人体检测的难点在于特征的提取以及分类器的设计。目前最具影响的人体检测方法是 2005 年 Dalal^[1]等提出的梯度方向直方图(HOG)结合 SVM 分类器的人体检测。HOG 特征还可以结合其他特征来提高人体检测的准确率。例如 Wang Xiaoyu^[2]等提出的 HOG-LBP 特征的人体检测方法解决了人体检测中的部分遮挡问题。其缺点是提取的人体特征向量维数较大,严重影响了系统速度。为了克服此缺点,随后相

收稿日期:2014-03-04

修回日期:2014-06-10

网络出版时间:2014-12-27

基金项目:国家自然科学基金资助项目(61211130309);安徽高校省级自然科学研究重大项目(KJ2010ZD10)

作者简介:张 慧(1987-),女,硕士,研究方向为视频图像处理;罗 斌,博士,教授,研究方向为模式识别与数字图像处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20141227.1343.021.html>

关学者在特征提取、分类器的训练和分类等方面做出了不少改进。Pang Yanwei^[3]等提出快速的搜索窗口级别的特征提取以及 Sub-cell 三线插值特征提取方法加快 HOG 特征提取;黄茜^[4]等在计算 HOG 特征时引入了积分图,节省了特征提取的时间;汪成亮^[5]等提出了一种基于主成分分析 (PCA) 降维的 HOG 描述子,提升了侦测窗口的分类速度。基于视频的运动人体检测不仅可以使基于静止图片的人体检测的方法,还可以利用视频特有的一些信息(如运动信息)进行前景提取,在文献[5-6]中,先进行运动分割即前景提取,再对运动目标进行人体识别。比较流行的前景检测的方法中,帧差法对于动态的场景具有较强的适应性与实时性,但是在运动较慢或者存在相邻帧重叠的情况下,它一般不能提取出完整的运动前景^[7];背景差法具有实现简单、准确性高的特点,但受光线、天气等外界条件变化的影响较大。文中使用帧差与背景差相结合的方法进行前景提取,先使用帧差法进行背景更新,再使用背景差进行前景提取,此方法兼具准确性与实时性的优点。

文中所针对的监控视频是在摄像机固定安装,拍摄场景有限,摄像机近距离监控运动目标的条件下得到的。人体目标在整个监控画面中占有一定的大小,人体较为清晰,其肤色信息可以被准确提取。文中以单目静止摄像机近距离监控情形为前提,提出一种实时的运动人体检测方法:

(1) 利用帧差与背景差相结合的方法进行运动目标分割,根据运动目标轮廓外接矩形的长宽比判别单目标及其他目标;

(2) 对于单目标,结合肤色分布判别目标是否为人;

(3) 对于其他目标,使用 HOG-LBP 特征,及 PCA 降维的方法,利用线性 SVM 进行分类,判别此目标是否为群人。

1 运动目标提取

通过背景差法进行运动分割,提取运动目标前景,重点在于对背景的确定,也就是区分前景与背景。运动目标提取的效果,很大程度地影响了后续运动人体目标检测的精度,运动目标检测的重点是背景的更新要适应监控场景的变化。

1.1 背景更新

对于单目静止摄像机近距离监控而言,由于光线等因素的影响,背景图像并不稳定,背景图像的改变是个渐变的过程。背景差分的选择是选择合适的背景图像,而且背景图像要随着时间的推移进行更新,难点在于如何动态地更新背景模型,背景模型更新算法的准

确与否直接影响到运动目标分割结果的好坏。

静止摄像机近距离监控视频有三个主要特点:背景是静止的,摄像机不会出现大的抖动;运动目标往往比较大,且运动速度变化比较大;背景受光线的影响比较大,但是光线影响是一个渐变的过程。从这三个特点出发,认为背景更新应该满足以下要求:

(1) 能及时感应监控场景的变化,并且能准确地提取出运动区域;

(2) 背景图像要较快地适应环境变化,取代采用长时间的图像帧平均法来实现背景更新法;

(3) 后续进入监控视频的静止物体或局部的背景变化慢慢融入背景图像,消除鬼影。

为满足以上要求,采用帧间差分法来感应光线的变化以及提取运动区域,再结合背景本身进行背景更新,注意下面所说的图像(帧)都是指灰度图像,具体更新过程如下^[8]:

(1) 背景模型初始化,以第一帧作为初始背景模型,记为 B ;

(2) 背景模型的更新,随着时间的推进,背景图像或多或少会发生一些变化,所以要实时地进行背景更新。文中利用帧差法所区分出来的静止区域与运动区域来更新背景模型。具体更新由式(1)决定:

$$B(x, y, i) = \begin{cases} B(x, y, i-1), p(x, y) \in D(x, y, i) \\ \alpha * B(x, y, i-1) + (1 - \alpha) * I(x, y, i), p(x, y) \notin D(x, y, i) \end{cases} \quad (1)$$

式中, $B(x, y, i)$ 表示第 i 帧图像的背景模型; $D(x, y, i)$ 表示由帧差法所确定的运动区域; $I(x, y, i)$ 表示第 i 帧图像的亮度分量; α 表示权值,实验中 $\alpha = 0.85$ 。

1.2 帧差法

相邻帧差法又称图像序列差分法、帧间差分法。当监控场景中出现物体运动时,帧与帧之间会出现较为明显的差别,两帧相减,得到两帧图像亮度差的绝对值,判断它是否大于阈值来分析视频的运动特性,确定图像序列中有无物体运动^[9]。其公式如下:

$$d = |I(x, y, i) - I(x, y, i-1)| \quad (2)$$

$$ID(x, y, i) = \begin{cases} 1, d \geq T \\ 0, d < T \end{cases} \quad (3)$$

式中, ID 表示相邻帧差图; $I(x, y, i)$ 表示第 i 帧图像的亮度分量对应像素的像素值; T 为阈值。

具体算法描述如下:

(1) 使用公式(2)与(3)获得相邻帧差图 ID ;

(2) 利用形态学方法处理相邻帧差图 ID , 获得运动区域连通域,并形成这些连通域的外接矩形集合,用

D 表示。

1.3 运动目标提取

背景差法即用当前帧和背景图像相减,去除背景,提取运动目标。若对应像素值相减所得到的像素值大于某一阈值,则把此像素作为前景。用公式表示如下:

$$d = |I(x, y, i) - B(x, y, i)| \quad (4)$$

$$FB(x, y, i) = \begin{cases} 1, & d \geq T \\ 0, & d < T \end{cases} \quad (5)$$

式中, $B(x, y, i)$ 表示通过背景更新获得的背景模型; $I(x, y, i)$ 表示第 i 帧图像的亮度分量; $FB(x, y, i)$ 表示第 i 帧与对应背景模型差图; T 为阈值。获得 $FB(x, y, i)$ 以后,使用形态学的方法处理,获得运动目标所对应的连通区域的轮廓以及外接矩形,外接矩形所对应的图像即是所需要的运动目标。

图1显示了采用帧差与背景差相结合的方法进行前景提取的效果图。

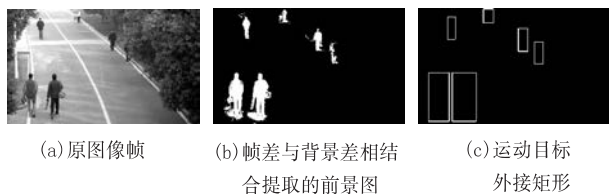


图1 帧差与背景差相结合提取前景效果图

2 运动人体检测

运动目标从背景中被分割出以后,接下来就是人体目标判定。整个判定过程如下:

(1) 计算运动目标外接矩形长宽比,根据长宽比把运动目标分成两部分;

(2) 对于长宽比在一定阈值范围内的目标,使用肤色特征进行人体目标判定;

(3) 不满足(2)的目标使用 HOG-PCA 特征进行判定。

2.1 单运动人体检测

2.1.1 运动目标外接矩形长宽比

对于正常行走的人体或者骑小型车辆的单个运动人体目标来说,其目标连通区域具有一定特征,最常见的特征为目标外接矩形长宽比,文中采用运动目标外接矩形长宽比作为单运动人体目标的判定值之一,即该比例必须在一定范围内,该目标才有可能被判定为单个运动人体目标^[10]:

$$\alpha_1 \leq \frac{W}{H} \leq \alpha_2 \quad (6)$$

式中, W 表示目标外接矩形的宽度; H 表示目标外接矩形的高度; α_1 与 α_2 是单人目标外接矩形长宽比的下限与上限,文中 $\alpha_1 = 0.23$, $\alpha_2 = 0.36$ 。

2.1.2 肤色特征

$YCbCr^{[11]}$ 颜色空间具有与人视觉感知的一致性、肤色聚类效果好及亮度和色度相互独立的特点,所以 $YCbCr$ 颜色空间是经常被使用的建立肤色模型的空间。肤色在 $CbCr$ 空间的值只集中在某一个范围内,可以利用这个特点把它与背景和其他颜色区分开^[12]。文献[12]指出,适合所有输入图像的肤色色度判决范围是 $Cr \in [133, 173]$ 和 $Cb \in [77, 127]$,实验证明该映射表对于不同光照下各种肤色人种判决性能很稳定。 RGB 颜色空间到 $YCrCb$ 颜色空间的映射关系如公式(7)所示:

$$\begin{cases} Y = 0.299 * R + 0.587 * G + 0.114 * B \\ Cb = 0.169 * R - 0.331 * G + 0.5 * B + 128 \\ Cr = 0.5 * R - 0.419 * G - 0.081 * B + 128 \end{cases} \quad (7)$$

结合目标外接矩形长宽比以及肤色特征,单运动人体判定过程如下:

(1) 计算运动分割所获得的运动目标外接矩形的长宽比。

(2) 如果长宽比符合公式(7),则把该目标所在的图像区域转换为 $YCrCb$ 颜色空间图像进入步骤(3);否则,使用2.2中的方法进行人体判定。

(3) 如果外接矩形上 $1/3$ 的子区域内有超过 5% 的像素满足 $Cr \in [133, 173]$ 和 $Cb \in [77, 127]$,则该目标为人体;否则,使用2.2中的方法进行人体判定。

2.2 群人运动人体检测

2.2.1 HOG 特征

HOG 代表的是图像局部区域的特征,而非一整幅图像。HOG 的局部提取范围称作窗(window),需要判断的就是这个窗区域中是否包含直立的行人。在提取的窗中,进一步将窗分割成块(block)和单元(cell)^[1]。HOG 的具体计算过程为:

(1) 输入规范化大小的窗区域图像,窗宽 64 个像素,高 128 个像素,转化为灰度图像;

(2) 通过一维中心模板梯度算子计算每个像素的水平 and 垂直方向上的梯度、梯度幅值和梯度方向;

(3) 把窗区域图像分割为若干个互不重叠的 $8 * 8$ 像素的 cell。由于边缘梯度方向差 180° 完全可以视为同一方向,所以可以把像素的梯度方向范围概括为 $[0^\circ, 180^\circ]$,并且均匀划分成 9 份,用于直方图统计。以各像素点的梯度幅度为权重,cell 内的每个像素为其梯度方向所在的直方图通道进行加权投票,统计各 cell 的梯度方向加权直方图;

(4) 每 $2 * 2$ 个 cell 组成一个块(block),各子 cell 直方图相连得到 block 的直方图向量,进行 L1 范数常规方法归一化处理,以消除光照的影响;

(5) 用块扫描样本图像,步长一个单元,窗区域图

像的 HOG 特征由块特征串联得到。

实验中,cell 大小为 $8 * 8$,block 大小为 $16 * 16$,block 步进为 8 个像素,梯度方向分为 9 个区间,对于大小为 $64 * 128$ 的检测窗口得到的特征向量为 3 780 维。

2.2.2 LBP 特征

HOG 特征描述的是行人轮廓边缘信息,较多地忽略了平坦的表面,导致分类器对于嘈杂的边缘信息的处理能力和鲁棒性较差。LBP^[13]是一种有效的纹理描述算子,它提取和度量了灰度图中局部邻近区域的纹理信息。HOG 特征结合 LBP 特征会提高人体检测的准确率,尤其对于有部分遮挡的人体检测有很大帮助^[2]。

为了和 HOG 特征结合起来使用,使用 block 提取 LBP 特征,不同的是,各 block 之间没有重叠。文中使用 $LBP_{8,1}^2$ 均匀模式的 LBP 编码,具体计算过程如下:

- (1)将规范化大小的检测窗口作为输入,根据 $LBP_{8,1}^2$ 计算每个像素的 LBP 特征值;
- (2)把样本图像分割为若干个 $16 * 16$ 像素的块(block),以各像素点的 LBP 特征值为区间,统计各 cell 的 LBP 直方图,并进行归一化;
- (3)用 block 对样本图像进行扫描,扫描步长为一个 block,最后将所有 block 的特征串联得到人体的纹理特征^[2]。

对整个检测窗口计算其 LBP,block 大小为 $16 * 16$,block 步进为 16 个像素,直方图分为 59 个区间,对于大小为 $64 * 128$ 的检测窗口得到的特征向量为 1 888 维。

由于高维的 HOG 特征向量和 LBP 特征向量存在大量的冗余信息,这些信息不仅可以减缓分类速度,有些维度信息甚至降低了识别的精度^[5]。文中利用 PCA^[14]分别对 HOG 特征向量和 LBP 特征向量进行降维,把降维后的特征向量串联输入线性 SVM 分类器进行学习,即得到了基于 HOG 和 LBP 特征的行人检测分类器。文献[5]指出,在 30~400 维 PCA-HOG 的识别率与 HOG 很接近,部分情况下 PCA-HOG 的分类效果更明显,100 维 PCA-HOG 的最佳识别率比 HOG 最佳识别率约提高了 1.2%^[5]。根据文献[5]的实验结果,文中 HOG 降维维数为 100,LBP 降维维数为 50。

3 实验与分析

3.1 实验环境与样本库

使用 Visual Studio 2005 与 OpenCV2.3 平台验证文中提出的快速运动人体检测算法。分类器为 OpenCV2.3 中的线性支持向量机(lin_SVM)。样本库选择的是 INRIA 行人数据库,使用样本库中的 2 416

个含有人体目标的正样本($64 * 128$)和 3 654 个不含人体目标的负样本($64 * 128$)作为分类器的训练集,该数据库中行人的服饰、姿态、光照和视角等因素的变化很大,是一个难度非常高的数据库^[4]。测试视频是由架设在某园区道路旁的单目静止摄像机近距离所拍摄的 10 段视频,视频中存在单人、群人、小型车辆(摩托车或自行车)以及机动车辆等多种运动目标,且视频中存在部分遮挡,有一定的检测难度。

3.2 实验结果及分析

对采集的 10 段视频进行了测试,视频为室外场景,分辨率为 $800 * 450$ 。这 10 段视频共有 5 921 帧图像出现 10 267 个运动人体目标。下面,把文中提出的快速行人检测方法分别与文献[1]和文献[2]提出的行人检测方法进行比较。

实验中,采用了滑动窗口方式进行选取侦测窗口,横纵方向的滑动间隔为 8 个像素^[5]。实验结果表明,文中提出的快速运动人体检测方法的识别率与文献[2]提出的基于 HOG-LBP 的行人检测方法很接近,比文献[1]提出的基于 HOG 的行人检测方法识别率高。这也验证了文献[2]提出的 HOG-LBP 特征可以提高人体检测的准确性的观点,同时也表明,使用 PCA 降维对人体检测的准确性没有大的影响。表 1 描述了三种方法识别率上的差异。其中,漏检率=漏检总人数/视频出现的总人数,误检率=误检总人数/检测出的总人数。

表 1 三种检测方法在检测准确度上的比较

方法	检测出的 总人数	漏检率 /%	误检率 /%
文中方法	8 963	12.7	7.4
文献[1]方法	8 129	20.8	9.3
文献[2]方法	8 491	17.3	8.85

文中提出的人体检测方法,由于加入了运动信息,通过对运动目标检测和定位,减少了人体检测的检测区域;同时通过肤色信息进行单人检测以及通过 PCA 降低了 HOG-LBP 特征的维度,节省了人体检测的时间,所以文中的改进方法具有较高的检测效率。实验结果表明,文中提出的人体检测方法(平均 70 ms/f)相较于 HOG^[1]滑动窗口方法(平均 850 ms/f)在速度上提高了约 10 倍,比文献[2]的方法(平均 1 s/f)快 13 倍左右,检测速度在 15 帧/s 左右,基本达到了实时性的要求。

图 2 显示了基于 HOG 的行人检测方法^[1]与文中提出的行人检测方法的检测对比效果,人体目标由矩形圈住。实验中,把骑小型车辆(摩托车或自行车)的人定义为行人。



(a) 文中提出的行人检测方法



(b) Dalal 等提出的人体检测方法

图2 检测结果对比效果

4 结束语

针对单目静止摄像机近距离监控的情形,文中提出一种基于背景差与 HOG-LBP 的快速运动人体检测算法。首先,应用帧差与背景差相结合的方法提取视频帧中的运动目标;接着,根据运动目标外界矩形长宽比分别进行单人判定以及群人判定。在群人判定中,文中提出 HOG 特征结合 LBP 特征并通过 PCA 降维的人体检测算法。文中提出的方法在准确性和检测率上都优于传统 HOG 特征法,当视频帧的分辨率为 $800 * 450$ 时,检测速度是传统 HOG 特征法的 10 倍以上。但是,该方法适用于监控场景中目标不是特别拥挤的情形,当目标比较拥挤时检测速度会降低,这将是下一步工作需要解决的问题。

参考文献:

- [1] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//Proc of IEEE computer society conference on computer vision and pattern recognition. San Diego, CA, USA: IEEE Computer Society, 2005: 886-893.
- [2] Wang Xiaoyu, Han T X, Yan Shuicheng. An HOG-LBP human detector with partial occlusion handling[C]//Proc of IEEE international conference on computer vision. Kyoto: IEEE, 2009: 32-39.
- [3] Pang Yanwei, Yuan Yuan, Li Xuelong, et al. Efficient HOG human detection[J]. Signal Processing, 2011, 91(4): 773-781.
- [4] 黄 茜, 顾杰峰, 杨文亮. 基于梯度向量直方图的行人检测[J]. 科学技术与工程, 2009, 9(13): 3646-3651.
- [5] 汪成亮, 周 佳, 黄 晟. 基于高斯混合模型与 PCA-HOG 的快速运动人体检测[J]. 计算机应用研究, 2012, 29(6): 2156-2160.
- [6] Zhang Xin, Gao Yuehua, Wang Xiaotao, et al. A method for detecting pedestrians in video surveillance scenes[C]//Proc of international conference on systems and informatics. Yantai: IEEE, 2012: 2016-2019.
- [7] Lipton A J, Fujiyoshi H, Patil R S. Moving target classification and tracking from real-time video[C]//Proc of IEEE workshop on application of computer vision. Washington DC: IEEE Computer Society, 1998: 19-21.
- [8] 魏 岩, 涂铮铮, 郑爱华, 等. 结合 RGB 颜色特征和纹理特征的消影算法[J]. 计算机技术与发展, 2013, 23(10): 72-74.
- [9] 李 刚, 邱尚斌, 林 凌, 等. 基于背景差法和帧间差法的运动目标检测方法[J]. 仪器仪表学报, 2006, 27(8): 961-964.
- [10] 梁英宏. 基于运动分割和肤色判别的人体目标检测方法[J]. 计算机应用与软件, 2010, 27(6): 111-114.
- [11] Chai D, Ngan K N. Locating facial region of a head-and-shoulders color image[C]//Proc of 3rd IEEE international conference on automatic face and gesture recognition. Nara: IEEE, 1998: 124-129.
- [12] 雷 明, 张军英, 董济扬. 一种可变光照条件下的肤色检测算法[J]. 计算机工程与应用, 2002, 38(24): 123-125.
- [13] Ojala T, Pietikainen M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-988.
- [14] 于成龙. 基于 PCA 的特征选择算法[J]. 计算机技术与发展, 2011, 21(4): 123-125.

基于背景差与HOG-LBP的快速运动人体检测

作者：

张慧， 郑爱华， 涂铮铮， 罗斌， [ZHANG Hui](#)， [ZHENG Ai-hua](#)， [TU Zheng-zheng](#)，
[LUO Bin](#)

作者单位：

[安徽大学 计算机科学与技术学院, 安徽 合肥, 230601](#)

刊名：

[计算机技术与发展](#)

英文刊名：

[Computer Technology and Development](#)

年，卷(期)：

2015 (2)

引用本文格式：[张慧](#). [郑爱华](#). [涂铮铮](#). [罗斌](#). [ZHANG Hui](#). [ZHENG Ai-hua](#). [TU Zheng-zheng](#). [LUO Bin](#) [基于背景差与HOG-LBP的快速运动人体检测](#) [期刊论文]-[计算机技术与发展](#) 2015 (2)