

基于多层码本模型的遗弃行李检测算法

宁文鑫, 王贵锦, 何 礼, 林行刚
(清华大学 电子工程系, 北京 100084)

摘 要: 提出一种在公共场所智能检测遗弃行李的新算法框架。首先使用多层码本模型检测场景中的稳定前景区域作为可疑区域,之后结合梯度方向直方图特征设计一种消除光照影响的机制,最后使用同样基于梯度方向直方图的快速行人检测算法消除行人的影响。实验结果表明,该方法计算复杂度低,在一些标准数据库上能有效检测遗弃行李,并对复杂场景有较好的鲁棒性。提出的算法框架可以应用到公共场所的监控系统中。

关键词: 遗弃行李; 多层码本模型; 梯度方向直方图; 行人检测

中图分类号: TP391;TP301.6 **文献标志码:** A **文章编号:** 1001-3695(2012)01-0386-03
doi:10.3969/j.issn.1001-3695.2012.01.107

Abandoned luggage detection method based on multi-layer codebook model

NING Wen-xin, WANG Gui-jin, HE Li, LIN Xing-gang
(Dept. of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract: This paper proposed a novel framework for detecting abandoned luggage in public places intelligently. Firstly created a multi-layer codebook model to detect the stable foreground of the scene as the suspicion regions. Then designed a new mechanism combined with HOG features to eliminate the influence of illumination change. At last, applied a human detector also based on HOG to eliminate the false alarms caused by pedestrians. Experiments demonstrate the algorithm is with low computing complexity and effective on standard databases. The method is also robust in complex scenes. The algorithm framework proposed in this paper can be applied in the monitoring system of public places.

Key words: abandoned luggage; multi-layer codebook model; histogram of oriented gradients(HOG); human detector

监控视频中的异常事件检测是计算机视觉领域中一个重要的研究课题。随着近年来监控系统广泛地应用在各种场所,工作人员观看监控视频并发现异常事件的难度变得越来越大。监控系统的智能识别功能变得非常必要。遗弃行李是公共场所中的一种重要安全隐患,许多地方都发生过隐藏在遗弃行李中的炸弹爆炸的危险事件。

国内外很多研究者对遗弃行李检测算法进行了研究。一类方法是基于跟踪算法的结果进行检测。Lv 等人^[1]利用跟踪算法确定场景中的行人和物体,并将它们映射到三维地图上,之后再根据它们之间的距离判断是否发生了遗弃行李事件。Smith 等人^[2]利用基于块的跟踪算法判断行人和行李分离的事件,再利用大小、形状等特征区分行人和遗弃行李。基于跟踪算法的检测可以在比较简单的场景中获得很好的结果,但在人群拥挤并有大量遮挡情况的环境中正确率会大幅下降。另一类方法是基于背景减除的。Tian 等人^[3]采用了以 Stauffer 等人^[4]提出的高斯混合模型为主的多种背景模型联合检测场景的稳定前景区域,并利用颜色、形状等多种特征区分遗弃行李。Singh 等人^[5]采用两个更新速率不同的高斯混合模型联合检测遗弃行李区域。基于背景减除的算法通常对光照比较敏感,在光照变化的场景中检测结果较差。Hu 等人^[6]提出了一种基于梯度方向直方图(HOG)特征的适应光照变化的背景模型,但该模型计算复杂度较高,不适用于实时检测的系统。

1 算法框架结构

本文提出了一种基于多层码本模型的新算法框架。该算法使用一个实时更新的多层码本模型对于场景的不稳定前景、稳定前景、次要背景和主要背景分别建模。这种模型的参数设置清晰,适合控制各层模型之间的更新速度,不同层次的模型对场景进行了更精细的建模,有利于后续的分析 and 检测。此外,本文设计了一种结合多层码本模型和 HOG 特征消除光照影响的新机制,利用了图像的空间信息,可以高效、准确地消除光照的影响。本文还使用了基于 HOG 特征的行人检测算法,在减少特征计算成本的同时消除场景中站立行人的影响。算法框架结构如图 1 所示。

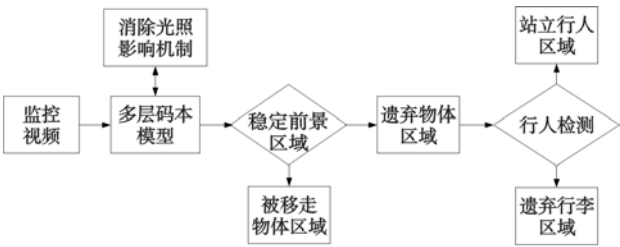


图1 算法框架结构

2 多层码本模型

Kim 等人^[7]提出一种码本建模的思想,他们为每个像素建

立由多个码字 (code word) 组成的码本模型 (codebook model, CBM)。CBM 算法对每帧图像中各个点对应的像素值根据亮度和颜色进行聚类,计算量和内存占用量都比较小。本文在 CBM 算法的基础上提出一个多层码本模型,针对遗弃行李检测的特点,对每个像素点采用多层结构建模,并设计了新的更新策略和检测方法。

2.1 模型结构

现在已有的背景模型都采用了多态建模的思想,用多个稳态值联合描述一个点的像素值,以适应不断变动的背景。但通常来说,场景的背景有一个主要的稳态值,在短时间内可能被不断遮挡,但在长时间内是稳定的。在背景建模时应当对主要背景和次要背景进行不同的处理。背景模型的主要作用是获得场景的前景区域,但在遗弃行李检测中,要获得的是持续一段时间稳定的前景区域,应当在模型结构设计时考虑这个需求。

本文针对以上问题提出一个四层结构的码本模型,分别是不稳定前景层 (unsteady foreground)、稳定前景层 (stable foreground)、次要背景层 (subordinate background)、主要背景层 (main background)。多层码本模型可用 $c = \{C_{UF}, C_{SF}, C_{SB}, C_{MB}\}$ 来描述。模型使用的码字总数为 L ,其中 C_{SF} 和 C_{MB} 通常使用 1 个码字, C_{UF} 和 C_{SB} 可根据应用场景使用不同数量的码字。在本方法中,四层模型分别使用 4、1、3、1 个码字。模型中每一个码字 $w_i (i=1, \dots, L)$, 可以表示为 $w_i = \{v_i, f_i, \lambda_i, \tau_i\}$ 。其中 $v_i = (R_i, G_i, B_i)$ 是一个 RGB 向量, f_i 是码字总出现次数, λ_i 表示码字连续未出现次数, τ_i 表示码字最大连续未出现次数。

2.2 码字匹配和更新

当监控系统获取新的一帧图像,每个像素对应的多层码本模型都应进行更新。通过亮度条件和颜色条件将新输入的 RGB 值进行聚类,并将聚类得到的码字添加到相应层的模型中。聚类过程中需要将输入的 RGB 值与模型中各个码字匹配并对码字进行更新。对于 c 中每一个码字 $w_i (i=1, \dots, L)$, 当输入一个新的 RGB 向量 v_t 时,码字匹配要同时符合以下两个条件:

$$\beta \|v_t\| < \|v_t\| < \gamma \|v_i\| \quad (1)$$

$$\delta = \sqrt{\|v_t\|^2 \cdot \frac{(v_t, v_i)^2}{\|v_i\|^2}} \leq T_s \quad (2)$$

其中,式(1)为亮度条件, β 的取值范围为 0.7 ~ 0.85; γ 的取值范围为 1.1 ~ 1.25; 亮度条件要求 v_t 和 v_i 向量的模之间的比值在一定范围内。式(2)为颜色条件,要求 v_t 与 v_i 方向的投影距离 δ 要在阈值 T_s 之内。同时考虑颜色和亮度信息能对 RGB 值进行有效聚类,颜色信息的使用也使模型在一定程度上能够适应光照变化的影响。

$C_{UF}, C_{SF}, C_{SB}, C_{MB}$ 中只要有一个码字与 v_t 匹配,则该层模型的匹配结果为 1,否则为 0。四层模型的匹配结果分别记为 $M_{UF}, M_{SF}, M_{SB}, M_{MB}$ 。如果码字与 v_t 匹配,则对该码字中的各个元素进行如下更新:

$$v_i = v_t \times (1 - \alpha) + v_i$$

$$f_i = f_i + 1$$

$$\lambda_i = 0$$

$$\tau_i = \tau_i$$

其中: α 是模型更新速率,一般取值在 0.95 ~ 0.98。对于与 v_t

不匹配的码字中的元素,进行如下更新:

$$v_i = v_i$$

$$f_i = f_i$$

$$\lambda_i = \lambda_i + 1$$

$$x_i = \max(\tau_i, \lambda_i)$$

对于一幅分辨率为 $m \times n$ 的图像,最多要进行 $m \times n \times L$ 次码字匹配与更新。但是,通过将式(2)两侧平方,并在计算中全部使用整型数据进行计算,码字匹配和更新的计算效率很高。

2.3 模型更新与检测

在模型中各个码字都进行匹配和更新后,模型各层之间要按照一定的策略进行码字更新和删除。不稳定前景层中出现达到一定次数的码字可以认为是稳定前景;稳定前景层中出现次数达到一定阈值的码字可以认为是次要背景。次要背景层和主要背景层中出现次数最多的码字是主要背景,其余的是次要背景。更新策略如下:

a) 如果 v_t 与所有码字都不匹配,利用 v_t 构建一个新的码字添加到 C_{UF} 中替换该层中 f 最小的码字,新码字的 f, λ, τ 都设置为 0。

b) 当 C_{UF} 中一个码字出现次数达到一定阈值 (人工设置的遗弃行李报警时间门限) 时,将该码字添加到 C_{SF} 中并替换其中 f 最小的码字。

c) 当 C_{SF} 中的一个码字出现次数超过一定阈值 (人工设置的将前景区域吸收到背景模型中的最短时间) 时,将该码字添加到 C_{SB} 中并替换其中 f 最小的码字。

d) C_{SB} 和 C_{MB} 之间进行码字双向更新,保证 C_{SB} 和 C_{MB} 中最大的码字在 C_{MB} 中。

另外,稳定前景层和次要背景层中保存的码字应该是以一定周期出现的,如灯光闪烁、树叶晃动等场景中不断变化的像素。因此 C_{SF} 和 C_{SB} 中的码字应该是持续出现或以一定的周期准周期地出现,如果这两层模型中的码字连续一段时间不出现,即将其从模型中删除。本文中使用的一个阈值 T_τ ,当 C_{SF} 和 C_{SB} 中的码字符合 $\tau \geq T_\tau$ 时则将该码字从该层模型中删除。 C_{MB} 中保存的是模型中出现次数最多的码字,即使一段时间不出现,笔者认为这是由于遮挡造成的,不删除该码字。

如果 M_{SB} 和 M_{MB} 都为 0,则该像素属于前景区域。在 C_{SF} 中有码字存在的情况下,如果 M_{SF} 为 1,则该像素属于稳定前景区域,即遗弃物体区域;如果 M_{SF} 为 0,且 M_{SB} 和 M_{MB} 都为 0,则该像素属于被遮挡的遗弃物体区域像素;如果 M_{SF} 为 0,但 M_{SB} 或 M_{MB} 为 1,则该像素属于被移走物体区域。这些不同的区域为之后的检测提供了足够的信息。

3 消除光照影响的机制

多层码本模型只利用了一个点的像素信息,没有使用不同点之间关系的空间信息,消除光照影响的能力有限。一般来说,图像的纹理具有很好的光照不变性,可以结合块的纹理统计方法消除光照影响。在本文中,结合多层码本模型和 HOG 特征设计了一种以较低计算成本消除光照影响的机制。

利用每个点的多层码本模型中 C_{MB} 保存的像素值构建一幅主要背景图像 I_{MB} ,将 I_{MB} 分成不重合的 $N \times N$ 大小的图像块,计算各个块的 HOG 特征并保存。每隔一段时间,如果某些点的 C_{MB} 中像素值发生变化,则更新该点所在图像块的 HOG

特征。对于多层码本模型检测出的前景区域,计算这些区域对应图像块的 HOG 特征向量 h_i ,并与 I_{MB} 中对应图像块的 HOG 特征向量 h_b 进行比较,比较公式为

$$D(h_i, h_b) = \sum_{i=1}^{I_b} \min(h_{i,j}, h_{i,j})$$

其中, $h_{i,i}$ 和 $h_{b,i}$ 表示特征向量 h_i 和 h_b 的第 i 维数据。
若 $D(h_i, h_b)$ 小于设定的阈值 T_1 ,则认为该块中的前景像素是由于光照变化或阴影导致的,不对这些像素点码本模型中 C_{UF} 和 C_{SF} 进行更新,而是将 C_{MB} 中码字按照匹配的情况进行更新,输出 M_{MB} 为 1, M_{UF} 、 M_{SF} 、 M_{SB} 为 0。若 $D(h_i, h_b)$ 不小于该阈值,则按照原有方式进行模型更新与检测。

由于场景中各个像素点的 C_{MB} 通常不会发生大的变化,前景区域的面积也小于场景整体面积,处理每帧图像平均的 HOG 计算量远少于文献[6]中的模型。

4 基于梯度方向直方图的行人检测

已有的基于底层信息的遗弃行李检测算法采用颜色、形状等特征来判断稳定前景区域是否为遗弃行李。但由于行李的多样性和因拍摄角度导致的形状失真,判断的准确率较低。公共场所背景模型检测出的稳定前景区域中很大一部分是由站立行人导致的,目前已经有一些算法能在监控视频中较为准确地检测出行人。通过行人检测算法消除行人导致的稳定前景区域可以减少大量虚警。

Pang 等人^[8]提出一种基于分层索引的梯度方向直方图(hierarchical HOG index matrices, HHIM)的快速行人检测算法。HHIM 算法首先计算待检测区域中每个 4×4 像素图像块的 HOG 特征,并将所有特征存储在一个矩阵中。从这个矩阵出发可以通过简单的矩阵加法得到各种大小图像块的 HOG 特征,并以块的大小为索引分层存储。采用一个基于此矩阵结构的窗口扫描算法可以利用各种大小的窗口扫描待检测区域,从而检测出区域中的行人。

本文的算法首先将多层码本模型检测得到的稳定前景区域扩展为适合 HHIM 算法检测的矩形区域,再使用 HHIM 算法扫描整个区域。如果在该区域中没有检测出行人,则该区域有很大可能性是由遗弃行李导致的,系统应当发出报警,并用特殊标记标明该区域。

结合 HOG 特征消除光照的机制和行人检测算法都使用了相同的特征。其中消除光照影响的机制计算了多层码本模型得到的前景区域的 HOG 特征,基于梯度方向直方图的行人检测算法使用稳定前景的 HOG 特征。而通常情况下,前景区域包含了稳定前景区域,因此只需要在 HHIM 算法中计算所有前景区域所在图像块的 HOG 特征便可为整个算法提供所需的 HOG 特征数据,降低了算法的计算复杂度。

5 实验结果

实验图像来自于标准数据库 PETS2006,以及专门拍摄的人群拥挤场景、光照变化场景和一段长时间的地下停车场测试数据。

5.1 PETS2006 数据库检测结果

PETS2006 数据库提供了 7 个不同场景的在公共场所遗弃行李的监控视频数据,每个场景中有 1 个遗弃行李事件。实验

结果显示本文提出的算法在 7 段视频中都准确检测出遗弃行李。在第 3 段视频中将一个只有上半身出现在场景中的站立行人被检测为遗弃行李。在第 5、6、7 段视频中,在遗弃行李未被遮挡时可以正确检测,但当行人站立在行李前,并几乎完全遮挡行李时检测结果会受到影响。表 1 是检测结果的统计数据,由于 PETS2006 视频数据的帧率为 25 fps,采用 125 s 作为阈值即代表着一件行李被遗弃 5 s 后被认定为遗弃行李。

表 1 PETS2006 数据库检测结果

序列	帧数	遗弃行李 出现帧数	报警阈值 (帧数)	正确检测 帧数	正确率/%
S1	3020	1133	125	1007	100
S2	2550	1336	125	1209	100
S3	2370	402	125	277	100
S4	3050	2254	125	2129	100
S5	3400	1524	125	1276	91.9
S6	2800	1186	125	1028	86.7
S7	3400	712	125	552	77.5

图 2 是在 PETS2006 数据库上的一些检测结果中,较细边框标记的区域为遗弃行李区域,较粗边框标记出的为站立行人的区域。



图 2 PETS2006数据库检测结果

5.2 复杂场景检测结果

为检测本算法在人群密集场景下的性能,笔者拍摄了两段学校食堂和教学楼人流高峰时间的测试数据,并进行实验,检测结果如图 3 所示。



图 3 人群密集场景检测结果

为测试本算法在光照变化场景下检测结果,笔者拍摄了教学楼中灯光开关变化导致光照变化的场景,并测试了本算法,实验结果如图 4 所示。



图 4 光照变化场景检测结果

图 4 中,第一幅图是顶灯全部打开的场景,第二幅图是关闭部分顶灯但没有使用消除光照影响机制的检测结果,第三幅图是使用了消除光照影响机制的结果。

5.3 长时间测试结果

为了测试算法的稳定性和检测速度,拍摄了一段时长 146 min 的地下停车场视频,并进行了测试,实验结果如图 5 所示。
本文提出的算法框架处理每帧图像的时间(下转第 392 页)

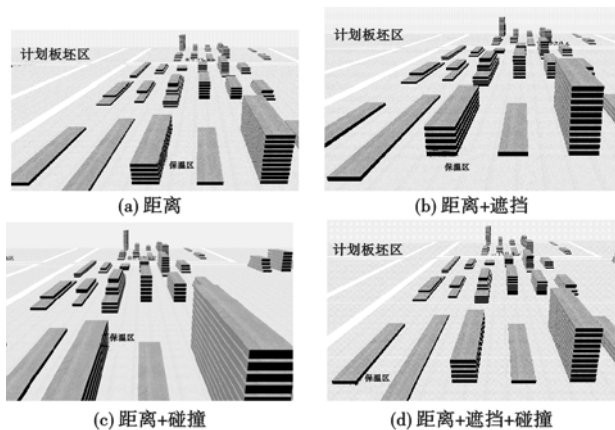


图8 几种不同情况的文字注解结果

4 结束语

本文提出了为板坯库进行动态文字注解的一种模型,这个模型将碰撞检测与三维文字注解联系在一起,综合考虑各种影响因素,能够在视点变换过程中实时计算文字注解的最佳位置,并针对模型间的相互碰撞问题分析了几种碰撞检测的方法。这对于存在大量静态模型少量动态模型的场景注解具有现实性意义。这个模型很好地权衡了距观察者距离、与板坯发生碰撞、与板坯发生遮挡之间的问题。但是由于三维板坯库场景中的某些区域板坯会相对密集,该模型并没有完全消除遮挡的问题。

参考文献:

[1] 赵巍,李苏剑,王刚祥.板坯库天车作业指导系统的设计与应用

[J].物流技术与应用,2007(5):97-99.
[2] 杭开武.板坯库计算机控制和管理[J].冶金自动化,2007(5):15-17.
[3] 谢斌,张晨.2250 热轧板坯库自动化控制系统[J].冶金动力,2010(2):87-89.
[4] 樊红,张祖勋,杜道生,等.基于神经网络模型求取注册配置最优解[J].武汉测绘科技大学学报,1998,23(1):32-35.
[5] 余代俊,耿留勇,兰燕. Delaunay 三角形法实现面状要素自动注册初探[J].测绘科学,2007,32(4):81-82.
[6] 邓红艳,武芳,李铭,等.遗传算法在点注册自动配置中的应用[J].测绘学院学报,2003,20(1):69-72.
[7] 吕国梁,孙敏,周晓光.基于视点的3维模型动态注册方法[J].地理世界信息,2008(2):53-57.
[8] 张云哲.三维建模中碰撞检测技术的基本算法研究[J].科技创新导报,2010(24):15.
[9] 宁涛,郭晨,张升文.用混合包围盒优化碰撞检测方法[J].计算机工程与应用,2011,47(1):1-3.
[10] KIM D J, GUIBAS L J, SHIN S Y. Collision detection among multiple moving spheres [J]. IEEE Trans on Visualization and Computer Graphics, 1998, 4(3):230-242.
[11] 王季,翟正军,蔡小斌.基于球深度纹理的实时碰撞检测算法[J].系统仿真学报,2007,19(11):2503-2506.
[12] 王晓荣,王萌,李春贵.基于 AABB 包围盒的碰撞检测算法的研究[J].计算机工程与科学,2010,32(4):59-61.
[13] LI C F, FENG Y T, OWEN D R J. SMB: collision detection based on temporal coherence [J]. Computer Methods in Applied Mechanics and Engineering, 2006, 195(19-22): 2252-2269.
[14] GANTER M A, ISARANKURA B P. Dynamic collision detection using space partitioning [J]. Journal of Mechanical Design, Transactions of the ASME, 1993, 115(1): 150-155.
[15] 丑强.虚拟环境中基于八叉树的碰撞检测问题[D].长春:吉林大学,2007.

(上接第 388 页)不是固定的。多层码本模型的检测和更新时间取决于使用码字的数量,消除光照影响和行人检测取决于前景区域或稳定前景区域的大小。



图5 长时间测试结果

本段视频的分辨率为 320×240 ,在测试中,平均处理每一帧图像需时约 15 ms,即达到了约 67 fps 的处理速度。在出现大量前景区域或稳定前景区域时,处理一帧图像的时间会迅速上升,实验中峰值为 43 ms。

6 结束语

本文提出了一种基于多层码本模型的遗弃行李检测新算法。该算法以多层码本模型为核心,设计了一种结合 HOG 特征消除光照影响的机制,并采用行人检测算法减少行人造成的虚警。整个算法综合利用了像素点的像素值信息和不同点之间的空间信息,能以较低的计算成本检测出场景中的遗弃行李,并对人群密集和光照变化等复杂场景有较好的鲁棒性。

参考文献:

[1] LV Feng-jin, SONG Xue-feng, WU Bo, et al. Left-luggage detection using Bayesian inference [C]//Proc of the 9th IEEE International

Workshop on Performance Evaluation of Tracking and Surveillance. 2006.
[2] SMITH K, QUELHAS P, GATICA-PEREZ D. Detecting abandoned luggage items in a public space [C]//Proc of the 9th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. 2006:1-8.
[3] TIAN Ying-li, LU M, HAMPAPUR A. Robust and efficient foreground analysis for real-time video surveillance [C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005:1182-1187.
[4] STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking [C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1999:117.
[5] SINGH A, SAWAN S, HANMANDLU M, et al. An abandoned object detection system based on dual background segmentation [C]//Proc of the 6th IEEE International Conference on Advanced Video and Signal Based Surveillance. 2009:352-357.
[6] HU Lei, LIN Wei-bin, LI Bo, et al. Robust motion detection using histogram of oriented gradients for illumination variations [C]//Proc of the 2nd International Conference on Industrial Mechatronics and Automation. 2010:443-447.
[7] KIM K, CHALIDABHONGSE T H, HARWOOD D, et al. Real-time foreground-background segmentation using codebook model [J]. Real-Time Imaging, Elsevier Science, 2005, 11(3): 172-185.
[8] PANG Guan, WANG Gui-jin, LIN Xing-gang. Real-time human detection using hierarchical HOG index matrices [J]. IEEE Trans on Information and Systems, 2010, E93-D(3): 658-661.