

应用于立体车库的人体检测算法

李响

(北京万象国纪投资有限公司, 北京 100029)

摘要: 针对立体车库的安全监控需求, 提出了一种基于机器视觉的人体检测算法, 用于在立体车库内检测非法侵入的人员; 首先, 采用自适应双线性滤波方法滤除视频帧中的噪声; 接着, 差分视频帧与背景图像, 并采用多尺度小波分析法二值化差分图像; 然后, 将二值图像进行形态学滤波以检测出人体; 最后, 利用高斯混合模型 (Gaussian Mixed Model, GMM) 更新背景图像; 实验结果表明, 提出算法可以在立体车库内实时准确地检测出人体, 为立体车库的安全监控提供了可靠的保障。

关键词: 安全监控; 人体检测; 双线性滤波; 多尺度小波分析; 混合高斯模型

Human Detection Algorithm for Stereo Garage

Li Xiang

(Ooepark Corporation, Beijing 100029, China)

Abstract: For the safety monitoring requirements of stereo garage, a machine vision based human detection algorithm is proposed to detect unlawfully intruded persons in the stereo garage. First, an adaptive bilinear filtering method is used to remove the noise of the video frame. Next, a difference image of the video frame and the background image is calculated, and the multi-scale wavelet analysis algorithm is employed to obtain binary image of the difference image. Then, human is detected from the binary image by morphology processing. Final, the Gaussian Mixed Model (GMM) algorithm is adopted to update the background image. Experimental results show that the proposed algorithm can detect human in real time and accurately in the stereo garage. It provides a reliable guarantee for the improved safety monitoring of stereo garage.

Keywords: safety monitoring; human detection; bilinear filtering; multi-scale wavelet analysis; Gaussian mixed model

0 引言

随着城市汽车保有量的快速增长, 有限的停车位已经不能满足泊车的需求, 二者之间的矛盾日益突出^[1]。立体车库将二维停车拓展为三维停车, 使单位面积上的停车位成数倍增加, 很好地提高了土地的利用率, 有效缓解了城市停车难问题。而立体车库使用中如何快速且准确地检测出非法侵入的人员是值得关注和重点研究的问题。现有人体检测算法主要分为基于模板匹配的方法和基于机器学习的方法^[2]。

对于检测非刚性的人体, 基于模板匹配的方法主要利用边缘、纹理和灰度等信息构建可变形的模板进行匹配。根据匹配形式的不同又可分为整体匹配和局部匹配。文献 [3] 提出了基于轮廓的整体匹配方法, 需要构建不同人体姿势的大量轮廓模板进行匹配。模板构建过程繁琐, 鲁棒性不高, 而且不容易包含所有的姿势。文献 [4-5] 提出了一种局部匹配方法, 仅建立人体头肩部模型, 通过匹配头肩模板来检测人体位置, 很大程度上避免了人体非刚性的影响, 也减少了模板个数, 取得了较好的检测效果。

基于机器学习方法提取能够充分表征人体的特征, 并研究有辨别力的分类器, 通过学习的分类模型分类候选区

域, 以检测是否为人体。其中, 特征的选择与分类器构造形式对最终检测效果起着关键性作用^[2]。相对于模板匹配方法, 此类方法在复杂场景下对变姿态人体检测有一定的鲁棒性。最为常用特征是 Dalai 等^[6]提出的梯度方向直方图 (Histograms of Oriented Gradient, HOG) 特征, 该特征在行人检测中得到了较好的结果。Q. Zhu 等^[7]提出了变尺度的 HOG 特征, 并构造级联分类器进行人体检测。文献 [8] 串联使用了 HOG 特征和局部二值模式 (Local Binary Patterns, LBP) 特征, 有效地解决了人体部分遮挡问题。

目前, 针对立体车库的人体检测研究相对较少。为满足立体车库安全监控的需求, 本文提出了一种基于机器视觉的人体检测方法。首先, 采用自适应双线性滤波方法以消除视频帧中的噪声。接着, 差分视频帧与背景图像, 并采用多尺度小波分析法将差分图像进行二值化。然后, 通过形态学滤波从二值图像中检测出非法侵入停车场的人体。最后, 采用混合高斯模型算法更新背景模型。实验证明提出方法得到了较高的人体检测准确率, 且实时性较好, 能够很好地满足立体车库安全监控的要求。

1 自适应双边滤波

作为非线性滤波, 双边滤波器可以理想地滤除视频帧中的随机噪声。而且双边滤波既能降噪, 同时又不会引起较强的模糊效应, 可以较好地保留视频帧的边缘细节^[9]。双边滤波器由两个高斯函数构成, 即空间和值域高斯函数。滤波时, 视频帧中与边缘较远的像素不会较大地影响边缘

收稿日期:2018-08-23; 修回日期:2018-08-29。

作者简介:李响(1982-),男,北京人,本科,工程师,主要从事机械车库安全和机器视觉方向的研究。

像素,因此可以较好地保持边缘特性^[10]。双边滤波定义如式(1)所示:

$$F_t = \frac{\sum_{i,j=-\omega}^{\omega} G_s(x,y,x_i,y_j)G_r(x,y,x_i,y_j)I_t(x_i,y_j)}{\sum_{i,j=-\omega}^{\omega} G_s(x,y,x_i,y_j)G_r(x,y,x_i,y_j)} \quad (1)$$

式中,

$$G_s(x,y,x_i,y_j) = \exp\left[-\frac{((x-x_i)^2+(y-y_j)^2)}{2\sigma_s^2}\right] \quad (2)$$

$$G_r(x,y,x_i,y_j) = \exp\left[-\frac{(I_t(x,y)-I_t(x_i,y_j))^2}{2\sigma_r^2}\right]$$

其中: $I_t(x,y)$ 为时刻 t 采集的视频帧 I_t 中的像素点; (x,y) 为像素点坐标; F_t 为滤波后得到的视频帧; G_s 是以距离作为自变量的高斯滤波函数,表示空间相似度,方差为 σ_s ; G_r 是以像素灰度差值作为自变量的高斯滤波函数,表示像素相似度,方差为 σ_r 。图 1 为不同 σ_r 取值情况下的滤波结果,其中滤波窗口半径 $\omega = 3, \sigma_s = 3$ 。图 1 (a) 为立体车库中摄像机拍摄的 I_t , 其余 3 图分别为 $\sigma_r = 5, \sigma_r = 10$ 和 $\sigma_r = 15$ 时的 F_t 。图 1 (b) 的滤波效果不明显,图 1 (d) 图像细节模糊。 $\sigma_r = 10$ 时的信噪比最大,滤波效果最好,如图 1 (c) 所示。

双边滤波器虽能较好地保证视频帧边缘细节不被模糊化,但依旧会在一定程度上改变边缘像素点的灰度值。噪声点处会存在灰度跳跃,从而使其局部区域灰度跨度较大,而边缘像素点则不是,可以利用这个特性进行自适应的双边滤波。逐行扫描 I_t , 当处理像素点 $I_t(x,y)$ 时,判断 $I_t(x,y)$ 滤波窗口内的像素点灰度方差是否大于经验阈值 T_1 。如果是,则进行滤波;如果不是,则跳过该像素。图 2 为采用改进方法得到的滤波结果,图 2 (a)、图 2 (b) 和图 2 (c) 分别为 $\sigma_r = 5, \sigma_r = 10$ 和 $\sigma_r = 15$ 时的 F_t 。通过对比发现,自适应方法能有效地滤除噪声点,而几乎不模糊边缘细节。

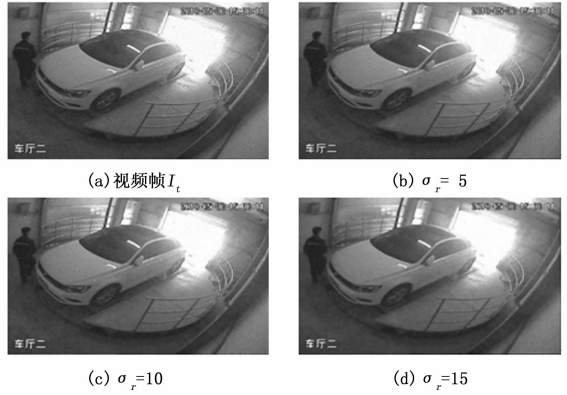


图 2 改进的双边滤波结果

$$D_t = |F_t(x,y) - B_{t-1}(x,y)| \quad (3)$$

$$A_t = \begin{cases} 1, & D_t(x,y) > T_2 \\ 0, & D_t(x,y) \leq T_2 \end{cases} \quad (4)$$

式中, $A_t(x,y)$ 为二值图像, T_2 为二值化阈值。

阈值的自动和最佳选取是进行二值化处理中的关键。二值化阈值存在于灰度图像直方图的波谷处,但在波谷较多时难以确定,而且大小不一毛刺也会影响波谷的准确定位。本文采用多尺度小波分析方法在最佳级数上对灰度直方图进行分解,选出几个波谷点作为候选 T_2 , 然后利用最大类间方差法确定唯一的阈值点^[11]。

信号 $f \in L^2$ 的小波变换^[12]定义为式(5):

$$WT(a,\tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) * \psi\left(\frac{t-\tau}{a}\right) dt \quad (5)$$

式中, a 为尺度变量; τ 为平移变量。式(5)所示的小波变换具有自动变焦的特点,可以通过变换尺度 a 来分析信号的特性:大的变化在大尺度 a 上观察,小的变化在小尺度 a 上观察^[11]。

为了确定最佳尺度 a , 只需对 D_t 的直方图做有限级分解的小波变换。分解级数不能太多,否则计算量会过大;分解级数也不能太低,否则受毛刺的影响,会含有“虚假波谷点”,影响候选 T_2 的确定。根据实验经验,在 256 级灰度图像中,分解级数选 3-4 级为宜^[11]。

D_t 直方图中波谷的位置,即候选 T_2 的位置,为小波变换的负极值点。根据文献 [13],在波谷自适应选取算法中选择了奇对称或偶对称小波,以保证小波变换的负极值点与波谷点的对应。接着采用最大类间方差法来确定精确的 T_2 。最大类间方差法^[14]是基于最小二乘原理的全局搜索方法。本文只在候选 T_2 处进行局部搜索,计算量可以大大降低。

根据式(4)使用阈值 T_2 进行二值化,对二值图像 A_t 进行开运算^[15]和连通分量提取^[16],并统计各连通分量的像素点数目。选择像素点数目最多且大于经验阈值 T_3 的连通分量,可认为该连通分量在 I_t 中对应的区域包含人体目标。二值化图像 A_t 如图 3 所示。

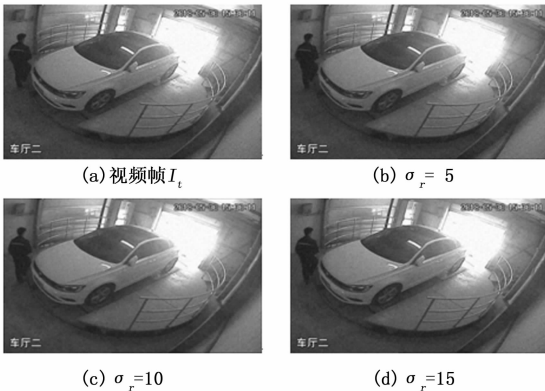


图 1 双边滤波结果

2 二值化分割与人体检测

为了检测出人体,首先要差分 F_t 与背景图像 B_{t-1} , 如式(3)所示,得到差分灰度图像 D_t 。然后对 D_t 进行二值化处理,如式(4)所示。

图 3 二值化图像 D_i

3 背景更新

常用的背景更新算法有多帧平均法^[17]、混合高斯模型^[18]法和 Surendra 法^[19]等等。混合高斯模型法更新背景的精确度高于多帧平均法和 Surendra 法。因此, 本文采用混合高斯模型法。混合高斯模型法可平滑地近似任意形状的密度分布, 即能克服背景模型复杂的场景^[20]。

将背景图像 B_{t-1} 中的每个像素点 $B_{t-1}(x, y)$ 都建立起数量为 N 的高斯模型函数, N 值越大能模拟的场景就越复杂, 但计算量也随之增大。 N 取值一般在 3~5 的范围呢。 $B_{t-1}(x, y)$ 灰度值的混合高斯概率和高斯分布函数如式 (6) 所示:

$$\left\{ \begin{array}{l} p(B_{t-1}(x, y)) = \sum_{n=1}^N \omega_n^{-1}(x, y) \times \eta(B_{t-1}(x, y), \\ \mu_n^{-1}(x, y), \sigma_n^{-1}(x, y)) \\ \eta(B_{t-1}(x, y), \mu_n^{-1}(x, y), \sigma_n^{-1}(x, y)) = \\ \frac{1}{\sqrt{2\pi(\sigma_n^{-1}(x, y))^2}} \exp\left(-\frac{(B_{t-1}(x, y) - \mu_n^{-1}(x, y))^2}{2(\sigma_n^{-1}(x, y))^2}\right) \end{array} \right. \quad (6)$$

式中, μ_n^{-1} , σ_n^{-1} 和 ω_n^{-1} 分别表示第 n 个高斯分布的均值、协方差和权重; η 为当前的高斯分布。

为了得到混合高斯模型首先需要初始化模型参数, 初始化时采用摄像机采集的第一帧视频图像 I_1 。对第一帧视频图像的每个像素点 $I_1(x, y)$ 对应的第一个高斯分布进行初始化, μ_n^{-1} 赋为当前像素点的值, ω_n^{-1} 赋为 1, 其余高斯分布函数的 μ_n^{-1} 和 ω_n^{-1} 都初始化为 0。 σ_n^{-1} 都初始化为相同的值。

初始化得到 I_1 的混合高斯模型后, 当前帧 I_{t-1} 的各像素点 $I_{t-1}(x, y)$ 分别与 $I_1(x, y)$ 的 N 个高斯分布依次进行匹配, 匹配条件为:

$$|I_{t-1}(x, y) - \mu_n^{-1}(x, y)| < D \times \sigma_n^{-1}(x, y) \quad (7)$$

其中: D 为系数, $2.5 \leq D \leq 3.5$ 。若满足式 (7), 则 $I_{t-1}(x, y)$ 与第 n 个高斯分布相匹配, 此时 $I_{t-1}(x, y)$ 为背景点, 需根据式 (8) 对高斯分布函数的参数进行更新:

$$\begin{aligned} \omega_n^{-1}(x, y) &= (1 - \alpha)\omega_n^{-2}(x, y) + \alpha M_n^{-1}, \\ \mu_n^{-1}(x, y) &= (1 - \beta)\mu_n^{-2}(x, y) + \beta I_{t-1}(x, y), \\ \sigma_n^{-1}(x, y) &= (1 - \beta)\sigma_n^{-2}(x, y) + \beta |I_{t-1}(x, y) - \mu_n^{-1}(x, y)|, \\ \beta &= \frac{\alpha}{\omega_n^{-1}}, \end{aligned} \quad (8)$$

其中: 匹配时 $M_n^{-1} = 1$, 不匹配时取 $M_n^{-1} = 0$, α 为权

重更新率, β 为均值和方差的更新率。

按式 (9) 计算各个高斯分布函数的优先级 λ_n^{-1} , 并按从大到小的顺序排序。

$$\lambda_n^{-1} = \frac{\omega_n^{-1}}{\sigma_n^{-1}} \quad (9)$$

选取背景模型时需从优先级 λ_n^{-1} 最高的高斯分布函数选取。当选取 N 个高斯分布函数作为背景模型时, 表达式如式 (10) 所示:

$$B_{t-1} = \operatorname{argmin}_n \left(\sum_{n=1}^N \omega_n^{-1} > T_1 \right) \quad (10)$$

式中, T_1 为背景所占比例。

4 实验结果与分析

根据经验, 实验中参数设置为: 滤波窗口半径 $\omega = 3$; 方差 $\sigma_s = 3$ 和 $\sigma_r = 10$; 背景更新尺度 $\alpha = 0.02$, 阈值 $T_1 = 1.5$, $T_3 = 20$, $T_4 = 0.8$; $D = 3$ 。

为了验证提出方法的准确性, 对从立体车库监控系统采集的 5 段监控视频进行检测, 每段视频 60 s, 帧率为 24 f/s, 分辨率 384×640 。这 5 段监控视频都包含 1 440 帧图像, 共计 7 200 帧图像, 其中 6 693 帧包含人体目标。提出算法和文献 [21] 方法在上述 5 段视频上的检测准确度进行了对比, 结果如表 1 所示。通过对比可以发现, 提出方法在所有视频帧中的平均准确率为 98.6%, 高于文献 [21] 的平均准确率 97.9%。提出方法的准确率满足立体车库安全监控的需要。

表 1 立体车库内人体检测的准确率对比

视频	准确率/%	
	提出方法	文献[21]
视频 1	98.3	98.7
视频 2	98.7	97.3
视频 3	99.3	99.0
视频 4	97.7	98.0
视频 5	99.0	96.4

为了进一步验证提出算法的实时性, 本实验在上述 5 段视频上, 将提出算法与文献 [21] 方法进行了对比。实验所用的计算机配置为: Windows10、CPU Intel Core™ i7-4700MQ、2.4 GHz、RAM 8 GB。采用的开发软件为 Matlab R2012b。实验结果如表 2 所示, 提出方法可以较好地缩短检测时间, 每帧的平均检测时间为 55.5 ms, 比文献 [21] 方法的每帧平均检测时间缩短了 26.2 ms。提出方法的检测效率能够满足立体车库实时安全监控的要求。

表 2 算法实时性对比

视频	平均每帧检测时间/ms	
	提出方法	文献[21]
视频 1	55.7	80.2
视频 2	56.3	82.7
视频 3	54.6	82.8
视频 4	55.1	81.7
视频 5	55.9	80.9

5 结束语

本文提出了一种可用于准确与实时检测非法入侵人体的图像识别方法。采用自适应双线性滤波方法降低视频帧中噪声的影响。采用多尺度小波分析法二值化视频帧与背景图像的差分图像。人体从二值图像中通过形态学滤波检测出来。背景图像在每一帧检测完人体后都利用 GMM 方法进行更新。通过实验可知, 提出算法可以实时和准确地从立体车库中检测人体, 满足立体车库安全监控的需求。

参考文献:

- [1] 方跃胜. 立体车库自动监控系统的设计与开发 [J]. 计算机与现代化, 2009, 12: 201-203.
- [2] 武文静. 监控视频中的人体检测与跟踪算法研究 [D]. 南京: 东南大学, 2015.
- [3] Gavrila D M, Giebel J, Munder S. Vision-based pedestrian detection: The protector system [A]. The IEEE Intelligent Vehicles Symposium [C]. Parma, Italy: IEEE, 2004: 13-18.
- [4] Broggi A, Bertozzi M, Fascioli A, et al. Shape-based pedestrian detection [A]. The IEEE Intelligent Vehicles Symposium [C]. Dearborn, USA: IEEE, 2000: 215-220.
- [5] 林佳, 阮晓钢, 于乃功. 基于头肩模型的人体识别方法 [J]. 计算机测量与控制, 2016, 24 (12): 205-208.
- [6] Dalal N, Triggs B. Histograms of oriented gradients for human detection [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. San Diego, USA: IEEE, 2005: 886-893.
- [7] Zhu Q, Yeh M C, Cheng K T, et al. Fast human detection using a cascade of histograms of oriented gradients [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2006: 1491-1498.
- [8] Wang X, Han T X, Yan S. An HOG-LBP human detector with partial occlusion handling [A]. IEEE International Conference on Computer Vision [C]. Kyoto, Japan: Springer, 2009: 32-39.
- [9] 程淑红, 高许, 程树春, 等. 基于计算机视觉的运动车辆检测 [J]. 计量学报, 2017, 38 (3): 288-291.
- [10] Wang M, Chen W, Li X D. Hand gesture recognition using

valley circle feature and Hu's moments technique for robot movement control [J]. Measurement, 2016, 94: 734-744.

- [11] 王建平, 盛军, 方敏. 基于小波分析的二值化阈值选取 [J]. 合肥工业大学学报, 2001, 24 (4): 559-562.
- [12] 崔锦泰. 小波分析导论 [M]. 程正兴译. 西安: 西安交通大学出版社, 1994.
- [13] Olivo J C. Automation threshold selection using the wavelet transform [J]. Cvgip Graphical Models and Image Processing, 1994, 56 (3): 205-218.
- [14] Otsu N. A threshold selection method from gray-level histograms [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9 (1): 62-66.
- [15] Bana S, Kaur D D. Fingerprint recognition using image segmentation [J]. International Journal of Advanced Engineering Sciences and Technologies, 2011, 5 (1): 12-23.
- [16] Audebert N, Le Saux B, Lefèvre S. Segment-before-detect: vehicle detection and classification through semantic segmentation of aerial images [J]. Remote Sensing, 2017, 9 (4): 368-1-18.
- [17] Cheon M, Lee W, Yoon C, et al. Vision based vehicle detection system with consideration of the detecting location [J]. IEEE Transaction on Intelligent Transportation Systems, 2012, 13 (3): 1243-1252.
- [18] 王丹, 刘怀. 基于改进混合高斯模型的背景提取与更新 [J]. 南京师范大学学报 (工程技术版), 2015, 15 (2): 60-64.
- [19] Cheon M, Lee W, Yoon C, et al. Vision based vehicle detection system with consideration of the detecting location [J]. IEEE Transaction on Intelligent Transportation Systems, 2012, 13 (3): 1243-1252.
- [20] Wu W, Yang J, Xu Z. The detection algorithm of irregular dynamic objects [J]. Multimedia Tools and Applications, 2017, 76 (13): 14599-14615.
- [21] Prakash O, Gwak J, Khare M, et al. Human detection in complex real scenes based on combination of biorthogonal wavelet transform and Zernike moments [J]. Optik, 2018, 157: 1267-1281.

(上接第 88 页)

- [2] 王溢琴, 秦振吉, 芦彩林. 基于嵌入式的智能家居之视频监控系设计 [J]. 计算机测量与控制, 2014, 20 (11): 3623-3626.
- [3] 邓荣钦. 基于 GPRS 的无线监控系统的研究 [D]. 武汉: 武汉理工大学, 2007.
- [4] 胡小颖. 全自动洗衣 PLC 控制系统设计 [J]. 机械工程与自动化, 2015 (3): 163-164.
- [5] 柳胜耀, 李瑛, 赵四海, 等. 基于家用洗衣机的性能测试环

境实验室建设 [J]. 实验室研究与探索, 2014, 33 (11): 248-251.

- [6] 谢自美. 电子线路设计实验测试 [M]. 第三版. 武汉: 华中科技大学出版社, 2006.
- [7] 张哲, 李智, 管四海. 基于 STC89C52 的智能全自动洗衣机控制系统设计 [J]. 微型机与应用, 2017, 36 (7): 102-105.
- [8] 李士勇. 模糊控制、神经网络控制和智能控制论 [M]. 第二版. 哈尔滨: 哈尔滨工业大学出版社, 1998.