

DOI:10.19651/j.cnki.emt.1802093

基于 TOF 相机的俯视行人检测

廖 畅 马秀丽

(上海大学通信与信息工程学院 上海 200444)

摘要: 在计算机视觉领域,行人检测是一项艰巨的任务。利用 TOF(time-of-flight)相机提供的深度信息进行俯视行人的高精度检测。当 TOF 相机处于俯视角度时,作为人体的一部分,人的头部具有非常丰富的特征信息,并且能在很长的时间里不被遮挡。鉴于此,提出了一种头部检测与注水算法相结合的俯视行人检测方法。该方法首先利用混合高斯背景模型来寻找图像中的感兴趣区域;然后采用注水算法对候选头部区域进行过滤;最后,结合深度信息等先验条件来确定场景中行人的真实头部。为了验证算法的有效性,分别在两种实时采集的深度图像数据集上进行了实验,与其他算法相比,本文所提出的方法具有更好的性能,能够实现实时、准确的行人检测。

关键词: 行人检测;俯视;深度图像;头部检测;注水算法

中图分类号: TN911.73 **文献标识码:** B **国家标准学科分类代码:** 460.40

People detection under an overhead time-of-flight camera

Liao Chang Ma Xiuli

(School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China)

Abstract: People detection is a challenge task in the field of computer vision. This paper aims to detect the pedestrians with high precision by only using the depth information provided by TOF(time-of-flight) camera. When the TOF camera is overhead, the human head, as a part of the human body, has very rich feature information and it can be unobstructed for a long time. In view of this, a people detection method, combing blob detection with water filling algorithm, is proposed. The proposed method first uses the mixture Gaussian background model to find the region of interest in the image; then it uses the water filling algorithm to filter out the candidate heads; at last, it combines with the prior condition to determine the real heads in the scene. In order to verify the effectiveness of the proposed algorithm, two kinds of real-time depth image datasets were tested. The result turns out that, compared with two other algorithms, the proposed method has better performance, and can achieve real-time and accurate people detection.

Keywords: people detection; overhead; depth image; head detection; water filling algorithm

0 引言

当今社会,经济飞速增长,人口数量也与日俱增,由此造成的一系列环境、资源、安全问题也日益凸显。因此,如何平衡庞大的人口数量与各类问题间的矛盾便成了亟需解决的问题。

检测人流量并加以调控便是针对其中某些问题的解决方案之一。对于名胜古迹、博物馆等人流密集的场所,通过人流量监控可以避免拥挤、踩踏等危险情况的发生,保证游客的安全;对于地铁站、公交站点等公交设施,通过人流量监控^[1],可以对于车辆进行更为合理得调度,进行资源最大化利用;而对于商场、超市等商业领域,通过人流量监控,可以更为及时地改变营销策略,进行更优的决策。

作为人流统计的关键部分,行人检测必不可少。在外部光线、阴影、复杂背景等因素的影响下,图像质量较差,更进一步的加大了行人检测的难度^[2]。因此,寻找一种能够应对复杂外界条件并能够快速、准确地检测行人目标的方法成为了研究热点。

Oren 等^[3]运用小波模版并结合样本训练学习,首次将机器学习引入人体目标检测。Dalal 等^[4]提出了一种基于梯度方向直方图(histograms of oriented gradient, HOG)的方法,该算子充分利用人体的边沿形状特征,在相邻的均匀小块中计算梯度直方图,得到局部梯度幅值与方向。该方法在 MIT 行人数据库的测试中取得近乎完美的结果,但由于 HOG 算子计算过程冗长,导致速度较慢,造成人体目标检测的实时性差。Xia 等^[5]设计了一种新型的两级人头检

测算法,其中包括 2D 边缘检测器和 3D 形状探测器。但该方法难以应对行人佩戴帽子的情况。Li 等^[6]提出了基于头-肩 Omega 形状的检测算法。Liu 等^[7]利用 HOG 特征对 Omega 形状进行描述,有效抑制了光线强弱给检测带来的影响。

由于在自然场景中进行行人检测,会受到光照、温度等诸多环境因素的影响,本文只采用 TOF 提供的深度图像,在俯视角度下进行行人检测。由于人的头部在俯视视角下具有类圆形的特征,结合 TOF 相机^[8]获取的深度图像,本文提出了将头部检测与注水法^[9]相结合的方法。实验结果表明,与结合深度信息的基于 21 霍夫变换(Hough transform, HT)^[10]的人头检测方法以及连通域分析^[11]方法相比,本文算法具有较高的检测精度,并且适合多种情况下的俯视行人检测,如有门场景和无门场景等。

1 深度图像数据集

本文的深度图像数据采集系统如图 1 所示。一个 TOF 相机被放置在行人头顶正上方、垂直于地板方向 220 cm 的位置处,从而得到采集的图像数据。为了保护行人隐私,本文只关注 TOF 相机提供的深度图像。该深度图像数据集包含有门和无门的两种场景,图 2 所示为其示例图像。同时,该数据集包含不同密度情况的行人,在相机采集的区域内,行人按照不同的方向行走。在一些简单的场景中,或许只有一个人经过相机采集区域。然而,在大多数实际日常场景中,至少有两个人在同一个方向上或者不同的方向上行走。本文使用的数据集为不同环境和拥挤条件下比较所提出方法的性能提供了保障。

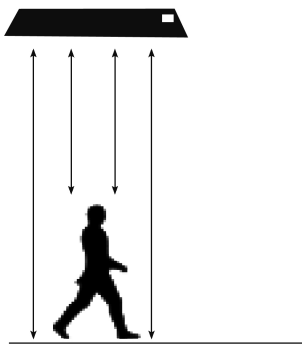


图 1 深度图像采集系统

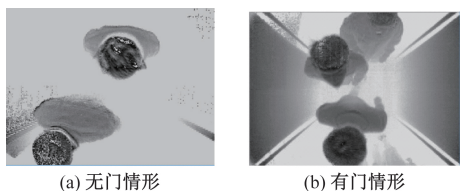


图 2 深度图像采集系统

针对采集到原始的 CSV 数据集,一些初步的数据筛选是十分必要的。首先,删除数据矩阵中左右各 6 列,上下各 4 行的无效数据。同时,根据有效性原则即 TOF 相机的高度为 220 cm,统一将深度数据集中大于 220 的深度值归一化为 220。最后,得到每组 400 帧深度图像,且每张图像的尺寸大小均为 320×240。

2 本文算法

若定义一个函数 f 为深度图像,则 $f(x, y)$ 为在像素 (x, y) 出的深度值,在深度图像中进行行人检测相当于在 f 中找到局部最小区域能满足式(1)。

$$E_A(f(x, y)) \leq E_{N/A}(f(x, y)) \tag{1}$$

式中: A 表示深度图像中某一个局部极值区域; N 是其邻域; $E_r(\cdot)$ 是一种操作,从而得到区域 r 的一个区域极小值。

本文提出将人头检测与注水法相结合的方法对俯视深度图像进行行人检测。该方法的关键在于从深度图像不同的局部最小值中提取出与之相对应的头部。其算法流程如图 3 所示。首先,去除采集到的深度图像的噪声;其次,利用前景提取法去除不相关的数据;然后利用注水法找到与局部最大点相对应的头部;最后,对之前处理的图像进行滤波,从而得到行人检测的结果。

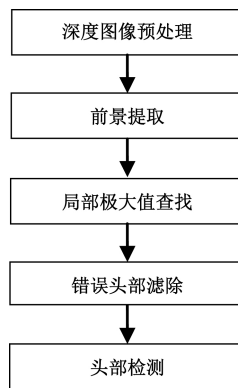


图 3 算法流程

2.1 深度图像预处理

由于采集过程中不可避免地会受外界光线、阴影和噪声等因素的影响而造成深度图像质量下降,对后续的检测效果有很大影响,因此,对原始数据集进行预处理是十分必要的。本文使用一个内核大小为 3 的中值滤波器来去除噪声,并结合形态学处理^[12]中的开运算操作,利用 8×8 矩形结构元素去填充深度图像的空穴(缺失数据)。最后,采用双边滤波器对不均匀的深度图像进行滤波处理。经过预处理的图像为后续处理过程奠定了基础。

2.2 前景提取

首先,利用混合高斯背景模型^[13]提取出深度图像的前景,即对深度图像中的每个像素的高斯分布动态控制,并且

通过在线期望最大化(expectation maximization, EM)算法对高斯分布的各参数进行学习从而自适应提取前景区域。此时提取到的前景部分包含行人的头部以及部分类头部如肩膀等感兴趣区域。其次,应用大津法^[14](OTSU 阈值算法)对行人和背景进行分离,通过计算出分离行人和背景的最佳阈值,从而达到保留前景图像的目的。

2.3 基于注水法查找局部最大值点

这一步的关键是在整个图像中找到对应于头部的局部最大区域。为了达到此目的,本文采用了注水算法,利用该方法的局部属性可以处理不同的情形。注水算法的提出是受到雨天雨滴降落的过程启发,下降的雨滴在地心引力的作用下从空中落下,向地面降落直到重力势能基本保持不变。在深度图像中,所有理想的“雨滴”都同一时间随机的落下,一旦“雨滴”降在图像上,它往往会自动沿着梯度下降最快的方向流动,然后所有的雨滴各自聚集在局部最低处,这就是整个注水过程。“雨滴”落在整个深度图像上,所有找到的由图像像素表示的局部最小区域代表着潜在的头部。

2.4 错误头部滤除

首先,本文通过判断得到的潜在头部区域的深度值是否达到深度阈值,从而得到一系列的待确认的头部区域。随后,通过计算潜在的头部区域的特征来检查每个检测到的潜在头部区域是否对应于真实头部。本文选取 3 个指标作为判断标准:1)检查头部区域的轮廓的半径 R 是否处在合适的范围内(实验中选取 $R \in (11, 17)$); 2)将潜在头部区域的轮廓的周长是否大于设定的阈值作为判别标准(实验中选取 700); 3)计算潜在头部区域的偏心率是否大于某个阈值(实验中选取 7)。将满足上述 3 个约束条件的潜在头部区域提取到最终结果中,通过候选框标注出来。

2.5 头部检测

在这个阶段,每一个候选框都将对应一个真正的头部。在深度图像中,被检测到的人的数量与头部区域的数量相同。同时,将为在深度图像中的检测到的每个头部分配一个 ID。

3 实验和结果分析

本节将对两种深度图像数据集的行人检测结果进行分析。其中,Frame1 数据集包含的是无门场景,Frame2 和 Frame3 数据集包含的是有门的场景。每个数据集均含 400 帧,大小为 320×240 的深度图像。采用 OpenCV 和 VS2013 进行实验,给出了与注水法相结合的头部检测的结果,还将其与结合深度信息的 21HT 人头检测^[15]以及连通域分析的结果进行比较。

此处选择无门场景下的深度图像数据集为例,预处理的结果如图 4 所示。

针对人的头部类圆形特征,在文献^[15]中提出了一种基于深度信息的 21HT 检测算法,将其分别在本文的深度

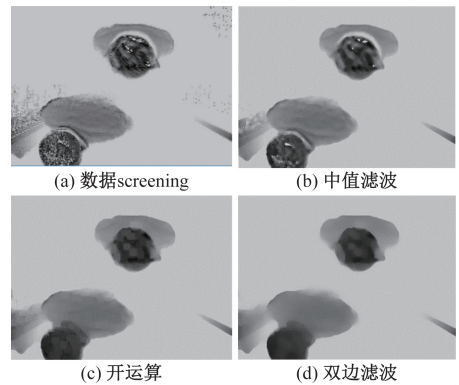


图 4 深度图像预处理结果

数据集两种场景上进行头部检测。首先,采用 21HT 算法对头部进行检测。事实上,行人的肩膀区域也有类似的圆形特征,偶尔也会被误检为头部。然后,滤除那些深度值超出感兴趣区域深度值的伪头部区域,从而得到真实行人的数目,进而达到准确检测效果。该方法的实验结果如图 5 所示。

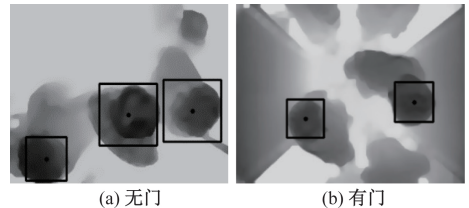


图 5 结合深度信息的 21 检测算法结果

同样,本文也考虑了与连通域分析方法的检测效果进行对比,连通域分析法通过边缘扫描法,将面积作为判决条件在深度图像数据集上进行行人检测,其实验结果如图 6 所示。

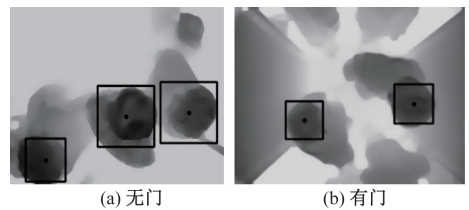
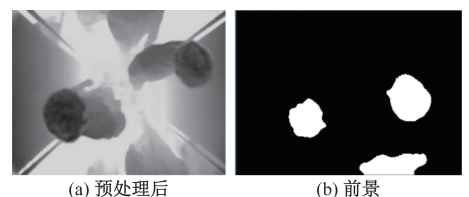


图 6 连通域分析结果

此时,无论数据集中的深度图像场景中是否存在门,采用本文提出的与注水法相结合头部检测的算法都能获得良好的检测结果。以数据集中包含门的场景为例,其实验结果如图 7 所示。



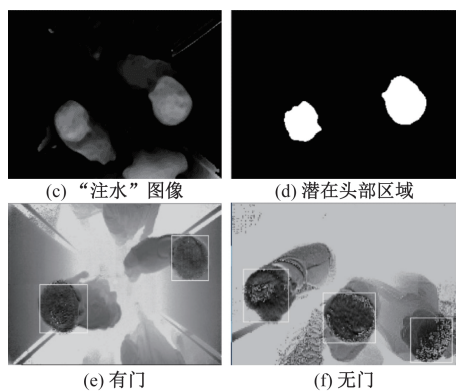


图 7 本文算法的检测结果

从图 8 中可以看出,连通域分析的方法和结合深度信息的 21HT 行人头部检测方法在某些情况下均无法很好地

工作。但本文所提出与注水法相结合的头部检测方法是可处理的。

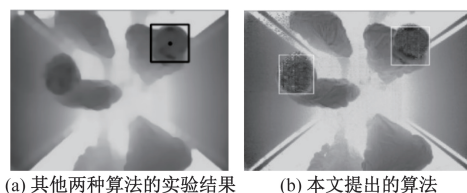


图 8 2 种方法检测结果的比较

在本文实验中,3 组深度图像数据集被采用,包含有门和无门行人通道的场景,3 组深度图像数据集中的行人的密度越来越深。实验结果详见表 1、表 2 和表 3。从中不难看出,本文提出的检测算法不仅有较高的检测精度,而且可以满足实时检测的需求。

表 1 基于连通域分析检测算法的结果

深度图像集	实际人数	检测人数	误检	漏检	平均检测速度/ms
Fram1	69	69	0	0	
Fram2	77	113	45	15	150
Fram3	78	150	71	18	

表 2 结合深度信息的 21HT 检测算法的结果

深度图像集	实际人数	检测人数	误检	漏检	平均检测速度/ms
Fram1	69	72	0	0	
Fram2	77	103	35	10	180
Fram3	78	130	58	12	

表 3 本文提出检测算法的结果

深度图像集	实际人数	检测人数	误检	漏检	平均检测速度/ms
Fram1	69	69	0	0	
Fram2	77	95	18	4	200
Fram3	78	96	18	5	

4 结 论

本文提出了一种头部检测与注水算法相结合的俯视行人检测方法,对 TOF 相机采集的深度图像提取前景,获取潜在的头部区域,将注水算法和头部检测相结合,用于行人真实头部的确定并得到最终的行人检测结果。

为了验证算法有效性,本文对 3 组深度图像数据集在两种场景(有门场景和无门场景)下进行了实验。实验结果表明,文中提出的方法的效果优于结合深度信息的 21HT 检测算法以及连通域分析方法的效果。在后续的研究中,将着重于进一步提高俯视行人检测的准确性,并考虑其他更复杂的场景,如行人戴帽子、背包等。

参考文献

- [1] GUO X J, YE J H. A head detection algorithm using depth information[J]. Journal of Yantai University, 2015.
- [2] SU S Z, LIU Z H, XU S P, et al. Sparse auto-encoder based feature learning for human body detection in depth image[J]. Signal Processing, 2015, 112:43-52.
- [3] OREN M, PAPAGEORGIOU C, SINHA P, et al. Pedestrian detection using wavelet templates [C]. Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 1997:193.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern

- Recognition (CVPR'05), 2005, 1(12):886-893.
- [5] XIA L, CHEN C C, AGGARWAL J K. Human detection using depth information by Kinect [C]. Computer Vision and Pattern Recognition Workshops, IEEE, 2011:15-22.
- [6] LI M, ZHANG Z, HUANG K, et al. Estimating the number of people in crowded scenes by MID based foreground segmentation and head-shoulder detection[C]. International Conference on Pattern Recognition, IEEE, 2009:1-4.
- [7] LIU P, CAI S. Omega-Shape Feature Learning for Robust Human Detection[M]. Pattern Recognition. Springer Singapore, 2016.
- [8] PIATTI D, RINAUDO F. SR-4000 and CamCube3.0 Time of Flight (ToF) Cameras: Tests and Comparison [J]. Remote Sensing, 2012, 4(4): 1069-1089.
- [9] ZHANG X, YAN J, FENG S, et al. Water filling: Un-supervised people counting via vertical Kinect sensor[C]. IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, 2012: 215-220.
- [10] 付琰, 卢荣胜, 夏瑞雪, 等. 基于重要性采样的随机 Hough 变换圆检测算法 [J]. 电子测量技术, 2012, 35(5):45-48.
- [11] SAGHEER M W, NOBILE N, HE C L, et al. A novel handwritten Urdu word spotting based on connected components analysis[C]. 20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, IEEE, 2010:2013-2016.
- [12] LERALLUT R, ÉTIENNE DECENCIÈRE, MEYER F. Image filtering using morphological amoebas[J]. Image & Vision Computing, 2007, 25(4):395-404.
- [13] MA G, PARK S B, IOFFE A, et al. A real time object detection approach applied to reliable pedestrian detection[C]. Intelligent Vehicles Symposium, IEEE, 2007:755-760.
- [14] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems Man & Cybernetics, 2007,9(1):62-66.
- [15] XU X F. The design and implementation of pedestrians recognition and tracking system based on binocular vision [D]. University of Electronic Science and Technology of China, 2013.

作者简介

廖畅, 硕士研究生, 主要研究方向为图像处理、计算机视觉等。

E-mail:757815010@qq.com

马秀丽(通信作者), 博士、副教授, 主要研究方向为模式识别、智能信息处理等。

E-mail:xlma@shu.edu.cn