

一种基于立体视觉的公交车客流计数方法

朱秋煜¹⁾ 唐利¹⁾ 郁铭²⁾ 江毅凭¹⁾

¹⁾(上海大学通信与信息工程学院,上海 200072) ²⁾(上海交通投资信息科技有限公司,上海 200433)

摘要 为了能实时地采集到公交系统中的客流量,提出了一种基于双目立体视觉的客流计数方法。该方法由3部分组成:运动目标检测,标记,匹配与计数。目标检测是计数问题的核心,该方法首先利用行帧差法检测可能的运动目标,然后基于视差分析判断可能目标的真伪。通过模拟环境和实际公交车环境的测试表明,该方法现场计数准确率达到92%以上,并有效克服了因计数场景中光线突变、阴影对计数准确率的影响。

关键词 客流计数 双目立体视觉 运动目标检测 行帧差 视差分析

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2009)11-2391-05

A Passenger Counting Method Based on Stereovision

ZHU Qiu-yu¹⁾, TANG Li¹⁾, YU Ming²⁾, JIANG Yi-ping¹⁾

¹⁾ (School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

²⁾ (Shanghai Transportation Investment Information Technology CO., LTD, Shanghai 200433)

Abstract A passenger counting method is proposed in this paper based on binocular stereovision for collecting the real-time passenger data in public transport. This counting method involves three blocks: moving object detection, object labeling, object matching and counting. Object detection is the vital issue of the counting system. A row frame difference method is firstly adopted to detect possible moving objects. Then, the facticity of each possible object is judged based on disparity analysis. Experimental results show that the proposed method can obtain an on-spot counting accuracy of 92% tested on a large data set capturing from simulated scenarios in laboratory and on moving buses. This approach is effective to avoid illumination changes and shadow of passengers.

Keywords passenger counting, binocular stereovision, moving object detection, row frame difference, disparity analysis

1 引言

根据公交系统中的客流信息,公交管理部门可以合理地安排公交线路和灵活调度公交车辆以提高公共交通的运营效率,因此对公交客流信息采集技术的研究具有很大的实用价值。

目前已有多种客流采集技术被提出,其中基于图像处理的客流采集技术成为研究重点,该技术有两个分支:一是利用单摄像机获得的平面信息;二是利用两个或多个摄像机获得的立体视觉。前一种技

术主要存在的问题是对计数环境的光线变化比较敏感^[1],不能排除阴影的干扰^[2];当运动目标与背景相近时,不能检测出目标^[1-2]。后一种技术可以较好地解决前一种技术存在的问题,Terada等人提出利用双目立体视觉获取行人的深度信息^[3],并根据深度信息检测目标,实验结果证明该方法的有效性;于海滨等人提出基于行人头部特征并利用头部区域的视差实现运动行人计数^[4];Yahiaoui等人提出了一种基于立体视觉用于公交车上的客流计数系统^[5]。但上述基于立体视觉的方法也存在计算量过大、运动跟踪的可靠性不高等问题。

基金项目:上海市科委重点攻关项目(07DZ1206),上海市教委重点学科项目(J50104)

收稿日期:2009-08-07;改回日期:2009-08-18

第一作者简介:朱秋煜(1964~)男。副教授。2006年于上海大学获工学博士学位。主要研究方向是图像处理与模式识别,主要研究成果为关于集装箱编码、车牌、人脸等的识别算法,图像篡改检测算法,人脸图像压缩算法等。E-mail:zhuqiuyu@staff.shu.edu.cn

本文提出了一种新的适用于公交车上的客流计数方法。该方法基于双目立体视觉实现,由 3 个模块组成,计数算法具有较高的实时性和鲁棒性。

2 系统实现框架

图 1 为提出的基于平行双目立体视觉的客流计数方法的实现装置示意图。系统装置中两个平行同步摄像机被垂直装在计数场所的正上方,这可从最大程度上避免乘客之间的互相遮挡情况。为了判断目标的运动方向,在计数场所中设置与乘客运动方向垂直且相隔一定距离的两条(in 和 out)虚拟计数线。左右摄像机获得的视频序列包含同一计数场所中 in 计数线和 out 计数线上的信息,提取每帧视图中这两条计数线所在行的灰度信息按时间轴依次展开,并取最近的固定行数构成灰度行时空图。如果有目标完全通过计数场所,则一定可以从灰度行时空图中获取该目标信息。获取左右视频序列中 in 和 out 计数线上的行时空图方法相同,本文仅以右视图为例说明。图 2(a) 为右视频序列某一时刻的

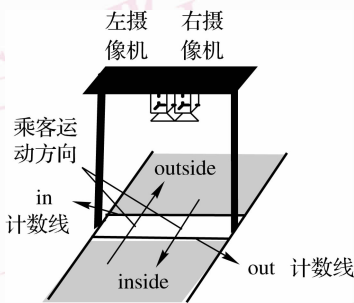
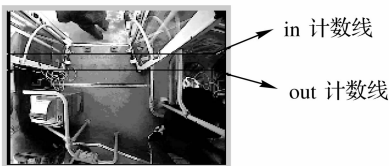


图 1 系统装置示意图

Fig. 1 The schematic device of system



(a) 右视图

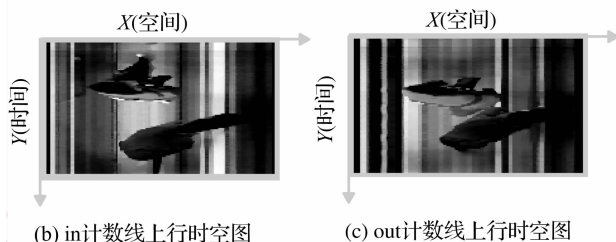


图 2 右视图数据

Fig. 2 The data of right view

图像,图 2(b),图 2(c) 分别为提取右视频序列中 in 和 out 计数线上灰度数据构成的灰度行时空图,系统后续图像处理均以灰度行时空图为数据源。

3 运动目标检测

运动目标检测是自动计数技术中的关键问题。基于立体视觉获取行时空图的视差图运算量大,并且对于缺少特征的区域会出现错误匹配。为解决上述问题,本文提出了首先利用行帧差法检测可能的运动目标,然后计算可能目标的视差并根据视差二次检测运动目标。

3.1 行帧差法检测运动目标

帧差法是传统的运动目标检测算法之一,行帧差法的工作原理和帧差法相似,利用行时空图中连续两行或几行图像数据的差异检测运动目标区域。

本文利用 3 行灰度数据,根据下式定义,在行时空图中检测出可能的运动区域:

$$d_i(j) = \frac{1}{2} (|f_i(j) - f_{i-1}(j)| + |f_i(j) - f_{i-2}(j)|) \tag{1}$$

$$g_i(j) = \begin{cases} 255 & d_i(j) \geq T_d \\ 0 & \text{其他} \end{cases} \tag{2}$$

式中, $f_i(j)$ 表示第 i 帧计数线上的灰度值, $d_i(j)$ 表示利用连续 3 行作行帧差得到第 i 帧计数线上的行帧差图, $g_i(j)$ 为行帧差检测得到的二值化运动目标图。

T_d 是检测目标的阈值,但为了检测出所有可能运动区域,故可将该门限值降低。图 3 为行帧差法检测出的运动目标。

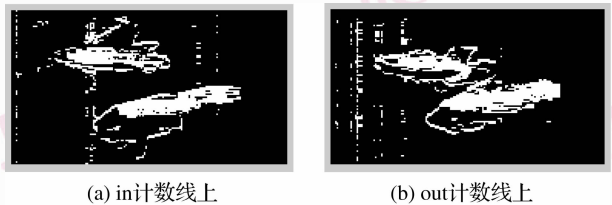


图 3 行帧差结果

Fig 3 The results of row frame difference

3.2 视差图计算

视差是指同一空间点在两个或多个摄像机平面上成像的位置差异,视差计算的关键是图像配准问题。常用配准方法有两种^[6]:基于区域的配准和基

于特征的配准。为得到稠密的视差图,本方法采用基于区域的配准算法。

计算当前处理行的视差,如果像素为可能运动目标点,则计算该像素的视差,否则将其视差当做 0 处理。为了对行视差图进行后续图像处理,将像素的视差做灰度变换,最后获得视差灰度在 0 ~ 255 范围内的行视差。像素匹配是视差计算的关键,根据极线约束条件,右视图计数线上像素点的匹配点在左视图的对应计数线上。视差计算方法如下:

首先在右视图的计数线上检测运动目标点,然后到左视图的对应计数线上搜索最佳匹配目标点,并采用以目标点和待匹配像素点为中心长为 $2W + 1$ 的两个 1 维窗口的灰度互相关函数作为相似度的测度,根据相关系数的极值寻找最佳匹配像素点。其中,为保证灰度平坦区域视差计算的可靠性,根据匹配窗口内像素灰度的方差动态调整 W ,即当窗口灰度方差小于门限值(根据实验得), W 每次以 3 个像素增长。在匹配过程中,根据匹配准则的唯一性约束、连续性约束条件判断所求视差的合理性并根据其邻近像素点的视差进行调整。

将每帧计数线上得到的行视差灰度数据按时间轴依次展开,并取最近的固定行数形成对应 in 和 out 计数线的两幅行视差时空图。为便于后续图像分割处理,本文采用 3×11 矩形窗口对当前行的视差进行平滑滤波处理,3 行视差是当前行视差及其前后一行视差。图 4 为平滑后得到的行视差图。

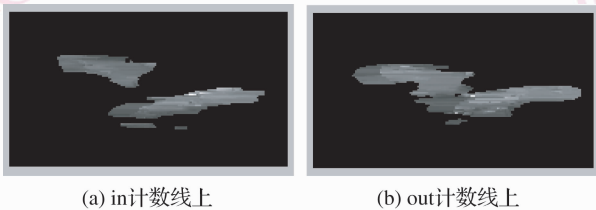


图 4 行视差图
Fig. 4 Disparity maps

3.3 运动目标标记

由于视差和目标场景中的景深成反比,因此对行视差图进行阈值分割可以检测出需要的目标对象。乘客通过计数区域时,其头部离摄像机像平面较近,故前景目标视差比背景视差大。

在行视差时空图中二次检测运动目标,如果像

素视差值大于 P_{min} 则认为是运动目标点,否则就认为是由于光照变化等干扰因素所造成的虚假运动目标点,作为背景点对待,确定 P_{min} 的值和计数环境中摄像机装的高度密切相关。在行视差图中检测出所有目标连通区域,并记录每个目标连通域的位置(起始行 y_T , 结束行 y_B , 左边界 x_L 和右边界 x_R) 并计算其面积 s 。这些连通区域可能存在 4 种情况:(1) 一个完整的运动目标;(2) 两个甚至多个运动目标并在一起的情况;(3) 一个目标分裂成的多个连通域之一;(4) 非运动目标引起的伪目标。根据训练样本统计得出判断一个连通域属于哪种情况的条件为:

$$\begin{cases} (1) & w_{min} \leq w \leq w_{max} \cap s_{min} \leq s \leq s_{max} \\ (2) & w \geq w_{max} \cup s \geq s_{max} \\ (3) & w_{min} \leq w \leq w_{max} \cap s \leq s_{min} \\ (4) & w < w_{min} \end{cases} \quad (3)$$

式中, $w = x_R - x_L$ 为连通域的宽度, w_{min}, w_{max} 表示单一运动目标在水平方向上宽度的变化范围, s_{min}, s_{max} 表示单一运动目标面积的变化范围。

当某一连通域满足上条件时,分别进行下述处理:

满足条件(1):将该连通域的信息存储到目标结构数组中,并在行视差图中用外接矩形标记出该目标。

满足条件(2):首先统计该连通域的视差灰度直方图,利用运动目标的距离信息对该连通域进行细分,并对该区域进行平滑;然后在区域检测连通域,并记录其位置信息和面积;最后判断这些连通域属于种情况。

满足条件(3):将该连通域的信息缓存到另一个目标结构数组中,等待判断是否可以和其邻近的连通域合并。

满足条件(4):认为该目标是干扰物,丢弃该连通域。

图 5 为经上述步骤处理后的行视差图。

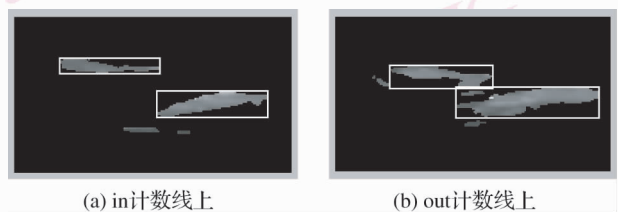


图 5 目标标记结果
Fig. 5 The results of object labeling

4 运动目标匹配及计数

乘客运动方向是根据其到达两条计数线的先后顺序确定的,即目标先到达 in 计数线,则该目标是上车,反之,为下车。为比较同一目标出现在 in 和 out 计数线的先后顺序,问题关键是在两条计数线的行视差图中进行目标匹配。

当在 in 计数线的行视差图中检测到运动目标时,到 out 计数线的行视差图中搜索待匹配目标。若在匹配范围内存在多个待匹配目标,根据代价函数求出最佳匹配目标,代价函数定义如下:

$$F(x, y_n) = aS(x, y_n) + bD_x(x, y_n) + cD_y(x, y_n) \quad (4)$$

该函数用于测量 in 计数线上目标 x 与 out 计数线上待匹配目标 y_n 的相似度。式中, $S(x, y_n)$ 为目标 x 和目标 y_n 的视差灰度平均值之差的绝对值, $D_x(x, y_n)$ 为 x 和 y_n 在水平方向的距离, $D_y(x, y_n)$ 为 x 和 y_n 在时间轴上的距离, a, b, c 分别表示上述 3 个测量标准在代价函数中所占权重,考虑到乘客通过两条计数线时深度信息几乎不变,在水平方向偏移不会太大,但在时间轴上变化不定,故根据它们所占权重确定 a, b, c 的大小,且 $a + b + c = 1$, 本文 a, b, c 取值分别为 0.4, 0.4 和 0.2。 $F(x, y_n)$ 值越小,说明两个目标越匹配,选择使 $F(x, y_n)$ 值最小的目标 y_n 为最佳匹配目标。得到最佳匹配目标时,比较两目标的质心位置即可判断目标的运动方向,进而完成计数。

5 实验结果及分析

为测试本计数方法的可靠性,本文采用在实验室模拟环境下拍摄的拥挤程度不同的视频(1, 2, 3),静止公交车上拍摄的视频(4)和运行中的公交车上拍摄的视频(5)对本文计数算法进行测试,这些场景中都包含了多人(最多 5 人)同时上下车、相互拥挤、光照变化等各种情况,其测试结果如表 1 所示。

测试结果表明,在实验室模拟环境下计数错误率在 $\pm 4\%$ 范围内,在运行公交车上计数错误率在 $\pm 8\%$ 范围内,分析出现错误计数的原因为:(1)客

流非常拥挤时,较难分割紧密依靠在一起的目标,会出现目标漏记;(2)乘客上下车的速度不一,速度慢的目标可能被分割而引起多记,速度快的目标会因运动方向判断错误导致错记;(3)公交车开关门时,会对目标检测带来较大的干扰,会引起目标多记。

表 1 实验结果

Tab. 1 Experimental results

视频序列	帧数	人工客流计数 (上车/下车)	自动客流计数 (上车/下车)	误差率 (%)
1	14 711	88/90	89/92	1.14/2.22
2	10 722	111/114	111/111	0.00/-0.03
3	10 814	200/196	193/189	-3.50/-3.57
4	15 114	132/134	138/129	4.35/-3.73
5	22 953	120/177	129/189	7.50/6.78

6 结 论

公交车上是一种复杂计数场所,为实现客流计数系统的可靠性和实时性,本文提出将运动目标的灰度信息和深度信息相结合实现客流计数的方法。该方法的优点在于:(1)适用于光线突变的计数场景;(2)可排除阴影的干扰;(3)计算量小。通过多组视频数据对本文方法进行测试,结果比较理想,现场计数准确率达到 92% 以上。需进一步研究的方向包括提高致密视差计算的精确度、更好地分割非常拥挤情况下的运动目标,以及解决公交车开关门时对运动目标检测带来的干扰等。同时,为保障该方法计数的可靠性,还需进行大量的现场测试。

参考文献 (References)

- Chen T H, Chen T Y, Chen Z X. An intelligent people-flow counting method for passing through a gate [A]. In: Proceedings of 2006 IEEE International Conference on Robotics, Automation and Mechatronics [C], Bangkok, Thailand, 2006: 1-6.
- Barandiaran J, Murguia B, Boto F. Real-time people counting using multiple lines [A]. In: Proceedings of the Ninth International Workshop on Image Analysis for Multimedia Interactive Services [C], Klagenfurt, Austria, 2008: 159-162.
- Terada K, Yoshida D, Oe S, et al. A method of counting the passing people by using the stereo images [A]. In: Proceedings of the IEEE International Conference on Image Processing [C], Kobe, Japan, 1999: 338-342.

- 4 Yu Hai-bin, Liu Jing-biao, Liu Ji-lin. An target region matching algorithm for the head feature extraction of pedestrain[J]. Journal of Image and Graphics, 2009, **14**(3):482-488. [于海滨, 刘敬彪, 刘济林. 用于行人头部特征提取的目标区域匹配方法[J]. 中国图象图形学报, 2009, **14**(3):482-488.]
- 5 Yahiaoui T, Meurie C, Khoudour L, *et al.* A people counting system based on dense and close stereovision [A]. In: Proceedings of the International Conference on Image and Signal Processing[C], Sanya, China, 2008: 59-66.
- 6 Fua P. A parallel stereo algorithm that produces dense depth maps and preserves image features[J]. Machine Vision and Applications, 1993, **6**(1): 35-49.