

采用截面流量统计的高密度人群行人计数^{*}

纪庆革¹, 陈婧¹, 迟锐¹, 方贤勇²

¹(中山大学 信息科学与技术学院, 广东 广州 510006)

²(安徽大学 计算机科学与技术学院, 安徽 合肥 230601)

通讯作者: 纪庆革, E-mail: issjqg@mail.sysu.edu.cn

摘要: 利用摄像头实现行人计数在智能视频监控领域有着重要的价值,但是行人互相遮挡、噪声、摄像机透视效果和图像背景等问题影响了人群计数的准确性.针对高密度人群场景的行人计数准确率的问题,提出了基于截面流量统计的行人计数方法,该方法基于梯度运动历史图像检测前景,并用有效运动图像改进了基于特征提取的行人计数方法,结合运动速度提取方法实现了行人计数.实验结果表明,提出的计数方法在高密度人群场景中具有较高的准确率和实时性,是一种针对高密度人群有效的行人计数方法.

关键词: 行人计数;高密度人群;特征提取;运动图像;运动速度提取

中文引用格式: 纪庆革,陈婧,迟锐,方贤勇.采用截面流量统计的高密度人群行人计数.软件学报,2014,25(Suppl.(2)):258-267. <http://www.jos.org.cn/1000-9825/14043.htm>

英文引用格式: Ji QG, Chen J, Chi R, Fang XY. Counting pedestrians in high-density crowd scenes using cross-sectional flow statistics. Ruan Jian Xue Bao/Journal of Software, 2014, 25(Suppl. (2)): 258-267 (in Chinese). <http://www.jos.org.cn/1000-9825/14043.htm>

Counting Pedestrians in High-Density Crowd Scenes Using Cross-Sectional Flow Statistics

Ji Qing-Ge¹, CHEN Jing¹, CHI Rui¹, FANG Xian-Yong²

¹(School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510006, China)

²(School of Computer Science and Technology, Anhui University, Hefei 230601, China)

Corresponding author: Ji Qing-Ge, E-mail: issjqg@mail.sysu.edu.cn

Abstract: Surveillance cameras have been widely installed in cities all over the world in recent years. Counting pedestrians from cameras has become a very important issue in intelligent video surveillance. However, factors such as occlusions, noise, camera perspective, background clutter may affect the accuracy of pedestrian counting. This paper introduces a pedestrian counting method for high-density crowd scenes using cross-sectional flow statistics. The proposed method consists of a new foreground detection algorithm based on the gradient motion history image, an improved feature-based counting algorithm by an effective motion image, and a moving speed extraction algorithm using optical flows. The experimental results show that the proposed method is robust and effective for counting pedestrians.

Key words: pedestrian counting; high-density crowds; feature extraction; motion image; moving speed extraction

1 引言

1.1 研究背景

行人计数在实际应用中具有十分重要的意义,统计数据不仅可以预防公共场所的安全问题,还可以为管理

* 基金项目: 国家自然科学基金(U0735001, 60473109, 61003131); 广东省自然科学基金博士启动基金(04300602); 安徽省自然科学基金(1408085MF113)

收稿时间: 2014-05-09; 定稿时间: 2014-08-19

者的决策提供数据支持,因此行人计数成为智能视频监控领域的研究热点.学者进行了大量行人计数的研究,高密度人群场景中的行人计数存在以下几个难点:第一,背景被持续遮挡,无法提取真实背景;第二,行人间遮挡现象严重,无法有效检测行人;第三,由于场景复杂、光照变化大,视频图像中存在大量噪声.所以,高密度人群的行人计数一直没有较好的解决办法,需要进一步研究.

1.2 相关工作

现有的行人计数方法可分为3类:基于行人检测跟踪的方法、基于特征提取的方法和基于特征轨迹聚类的方法.

基于行人检测跟踪的行人计数方法的流程为:首先检测场景中的行人,然后对检测到的行人进行跟踪得到运动轨迹,最后分析行人运动轨迹并计数.其中行人检测和行人跟踪是整个算法的关键技术,直接影响计数结果.行人检测方面,有运动目标轮廓^[1]、方向梯度直方图(histogram of oriented gradient,简称 HOG)^[2]、运动目标边缘^[3]、头(肩)检测^[4-8]等方法.为了解决遮挡和透视效果的影响,文献[9]提出了运用多摄像机进行行人检测.行人跟踪方面,文献[10]利用人群目的地、人群中行人间相互作用力、场景边界这3个场景约束,提高了在高密度人群中行人跟踪的鲁棒性.该算法的缺点是:当人群中的行人增加时,行人间存在遮挡现象,导致行人检测算法准确率下降.

基于特征轨迹聚类的行人计数方法:首先,提取图像中的特征点;然后,跟踪特征点,得到特征点的运动轨迹;最后,对特征点的运动轨迹聚类,通过聚类得到的簇数就是行人个数.文献[11]用特征点跟踪的方法得到运动轨迹,然后根据轨迹的长度和空间位置聚类得到行人数.文献[12]利用 KLT 跟踪特征点获得一系列运动轨迹,然后用合并聚类方法对这些特征点轨迹聚类以得出行人个数.为了在高密度人群场景中得到较好的计数结果,文献[13]将基于特征轨迹聚类的行人计数方法和基于特征提取的行人计数方法结合,计算人群的特征点光流并聚类,然后通过学习得到聚类结果与人群行人数之间的关系.这种方法有两个缺点:第一,复杂的轨迹管理;第二,在高密度人群场景中,简单地假设聚类簇数等于行人个数往往是错误的.

基于特征提取的行人计数方法:将人群视为一个整体,利用图像特征和人群人数之间的关系实现行人计数.常用的图像特征有:纹理特征^[14,15]、灰度共生矩阵(GLDM)^[16]、统计地形特征(statistical landscape features,简称 SLFs)^[17]、特征点特征^[18]、方向梯度直方图(HOG)^[19]、运动信息与像素特征^[20]等.基于特征提取的方法的缺点是:没有考虑特征模型的空间特性,即对于同一个目标,随着它与摄像机的距离变化,特征向量也会改变.如果人群密度过大,将整个场景作为计数单位,场景过于复杂,就会产生很多噪声,噪声累积会对计数结果有很大影响.场景中行人速度也存在不同,即同一时间间隔不同区域的行人流量也不同,若将整个场景作为一个计数单位,则无法解决速度不均的问题.

本文在深入研究现有行人计数方法的基础上,针对高密度人群场景提出了基于截面流量统计的行人计数方法,利用梯度运动历史图像改进了前景检测的结果,结合了基于特征提取的行人计数方法与速度提取方法完成行人计数.人群特征提取方面,本文针对高密度人群场景提出了有效运动特征,并将其和梯度像素特征、特征点特征相结合运用在高密度人群的行人计数中.速度提取方面,考虑到鲁棒性和实时性的要求,本文将基于运动历史图像的方法和光流法结合运用到行人计数中.

2 基于特征提取的行人计数方法改进

2.1 基于梯度运动历史图像的前景检测

文献[21]提出了基于时间戳的运动历史图像(timed motion history image,简称 tMHI),利用时间戳更新运动历史图像,去除过期无用的运动信息,tMHI 记录了目标的空间历史信息,建立了目标运动空间上、时间上的相关性.但也存在不足:由于 tMHI 是将帧间差分得到的前景图像作为历史模板,因此无法记录场景中目标的短暂静止信息,文献[22]对这一不足进行了改进,用背景差分前景图像代替帧间差分前景图像,但是在高密度人群场景中,由于背景图像提取困难、前景面积过大等因素,采用背景差分算法无法提取出完整的前景图像.

为了解决文献[22]提出的背景差分算法无法提取出完整的前景图像问题,本文提出用梯度二值化图像作为 *tMHI* 的历史模板,并在此基础上提出了基于梯度运动历史图像的前景检测方法.

计算梯度二值化图像时运用了 Sobel 算子,Sobel 算子利用了邻域像素的灰度信息,对于噪声有一定鲁棒性,因此减少了有效运动图像对噪声的敏感性.另外,由于是用边缘图像作为运动历史图像的模板,因此有效运动图像对光照具有一定的鲁棒性.还有,计算有效运动图像时可以去除一些长期积累的噪声,如光照引起的纹理.

本文提出的前景检测计算方法:

第 1 步:计算图像的梯度二值化图像 $T(x,y)$,

$$T(x,y) = \begin{cases} 255, & T_x(x,y) = 255 \text{ or } T_y(x,y) = 255 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中, $T_x(x,y)$ 和 $T_y(x,y)$ 是图像在 x 和 y 方向的梯度二值化图像.

第 2 步:计算梯度运动历史图像,

$$tMHI(x,y) = \begin{cases} \tau, & \text{if } T(x,y) = 255 \\ 0, & \text{else if } tMHI < (\tau - \delta) \end{cases} \quad (2)$$

其中, $T(x,y)$ 是上一步计算出的梯度二值化图像, τ 为当前时间, δ 为事先定义的时间戳,表示 *tMHI*(基于时间戳的运动历史图像)所记录的时间范围.当前帧的梯度像素在运动历史图像中的灰度值最大,当过去帧与当前帧的时间间隔超过时间戳 δ 后,过去帧的梯度像素在运动历史图像中的灰度值将被清零.这样保证了运动历史图像只记录近期 δ 时间内的运动历史信息.

第 3 步:根据梯度运动历史图像计算出当前帧的有效运动图像,即前景图像,

$$EM(x,y) = \begin{cases} 255, & \min \leq \text{MAX}(x,y) - \text{MIN}(x,y) \leq \max \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

其中, $\text{MAX}(x,y)$ 和 $\text{MIN}(x,y)$ 分别为遍历梯度运动历史图像中点 (x,y) 邻域内所有点后得到的最大最小值, \min 和 \max 为预先设置的阈值.如果 $\text{MAX}(x,y)$ 和 $\text{MIN}(x,y)$ 的差值过大,则说明该点是静止目标或背景的边缘边界点;如果 $\text{MAX}(x,y)$ 和 $\text{MIN}(x,y)$ 的差值过小说明该点是静止目标或背景的边缘内部点.因此可以滤掉这些点,提取出图像中的有效运动点即前景点,如图 1 所示.

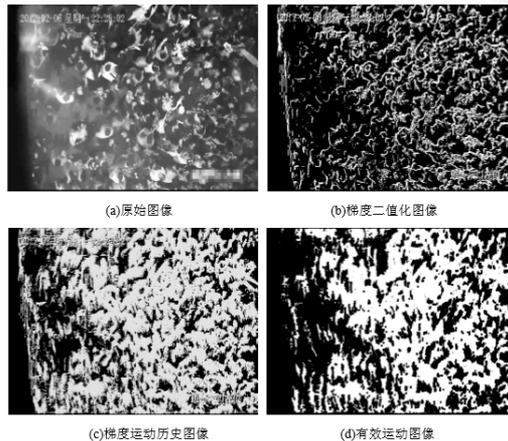


图1 基于梯度运动历史图像的前景检测

2.2 特征选择

本文利用梯度运动历史图像改进了前景检测的结果,并将提取出来的有效运动像素特征和梯度像素特征、特征点特征组合为新的人群特征向量.

2.2.1 有效运动像素特征

特征提取:计算有效运动图像,将截面流量计数块中的有效运动像素数作为其一维特征.

有效运动像素特征计算简单,并且对于噪声有一定的鲁棒性.

2.2.2 梯度像素特征

特征提取:实验中用 Sobel 算子得到边缘图像后,将其进行二值化处理得到边缘二值化图像,将截面计数块中边缘像素个数作为其一维特征.

2.2.3 特征点特征

特征提取:在行人计数算法中,只需提取特征点的数量作为人群的特征,并不利用特征点进行图像匹配、三维重建等高层应用.FAST^[23]特征点只检测出图像的特征点并不对特征点进行描述,其检测速度快而且具有尺度不变性,因此实验中采用 FAST 特征点的数量作为人群的特征.

3 运动速度的提取方式

光流法是一种简单而又实用的图像运动表达方式,但是在实际应用中,由于遮挡、光照不均以及噪声等多种原因,使得光流场基本方程的灰度守恒假设条件不能满足,影响光流法的准确性和有效性.而前文提到的运动像素特征对噪声具有一定的鲁棒性,所以本文将基于运动历史图像的方法和光流法结合,从视频图像中提取出运动速度.

3.1 基于运动历史图像的方法

我们注意到 k 时刻运动历史图像的梯度向量总是与目标边缘垂直.根据光流约束方程可知,运动历史图像在某点 (x,y) 的梯度向量即为该点的运动速度向量,用 Sobel 算子计算 tMHI 中的点 (x,y) 在 x 方向和 y 方向的梯度 $S_x(x,y), S_y(x,y)$, 则点 (x,y) 运动的方向和速度值为^[21]

$$\Phi(x, y) = \arctan \frac{S_y(x, y)}{S_x(x, y)} \quad (4)$$

$$V(x, y) = \sqrt{S_x^2(x, y) + S_y^2(x, y)} \quad (5)$$

注意,只有有效运动点的梯度向量等于该点的运动速度向量,其他点的梯度向量没有意义.当利用运动历史图像计算某区域的全局方向时,需要对 tMHI 进行标准化处理,对区域中的每个点赋权值,当前模板中的运动点权值最大,随着时间的倒推权值逐渐减小.全局方向 $\bar{\Phi}$ 的计算公式为^[21]

$$\bar{\Phi} = \Phi_{ref} + \frac{\sum_{x,y} \text{angDiff}(\Phi(x, y), \Phi_{ref}) \times \text{Norm}(\tau, \delta, \text{tMHI}(x, y))}{\sum_{x,y} \text{Norm}(\tau, \delta, \text{tMHI}(x, y))} \quad (6)$$

其中, Φ_{ref} 为区域中有效运动点的梯度方向直方图的主方向, $\text{Norm}(\tau, \delta, \text{tMHI}(x, y))$ 为 tMHI 在点 (x, y) 处的标准化值, $\text{angDiff}(\Phi(x, y), \Phi_{ref})$ 为点 (x, y) 的梯度方向与 Φ_{ref} 的最小角点差.

速度提取方法是计算截面流量计数块中各有效运动点的运动速度,取加权平均值作为计数块的全局运动速度.

3.2 基于光流的速度提取方法(FAST特征点)

场景的二维运动场是三维运动的投影,光流是时间序列上帧与帧之间的像素变化,它是一种像素级运动.在理想情况下,光流场反映了场景的二维运动场.光流法利用目标运动与图像灰度变化的相关性,得到了场景中的动态信息,被广泛运用在目标检测和目标跟踪等领域.

速度提取方法:先提取 FAST 特征点,然后计算各个特征点的光流,将这些光流的加权平均作为全局运动速度向量.

3.3 速度提取与行人计数

实验采用 SVM(support vector machine)对截面流量计数块进行人数预测.训练部分从训练样本中提取特征

(有效运动像素特征、梯度像素特征和特征点特征),通过学习得到特征向量与行人数之间的关系.预测部分,将从视频帧中提取出的特征向量输入训练好的 SVM 即可预测得到截面流量计数块的行人数.

速度提取与行人数更新:首先提取截面计数块的当前帧的速度,然后更新截面计数块的历史速度和位置,最后判断截面计数块与计数线的位置关系,如果截面计数块超过计数线,则更新流量计数结果并重置截面计数块位置.

4 实验结果及分析

4.1 方法流程

本文提出的基于截面流量统计的行人计数方法,是在视频中设置一条行人流量统计带,并将统计带划分为若干个区域,每个区域为一个截面统计块,提取出每个截面计数块的行人数和速度,累加后得到场景中的人流量数.

如图 2 所示,本文提出的方法流程是:第 1 步,利用基于改进的特征提取的行人计数方法得到截面计数块中的行人个数;第 2 步,提取截面计数块中行人的速度;第 3 步,通过分析截面计数块与计数带的相对位置来统计行人流量.

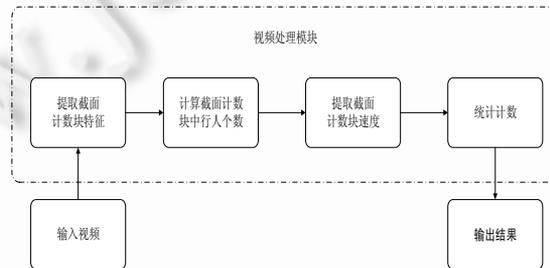


图2 采用截面流量统计的行人计数方法流程图

4.2 实验数据的选择

实验选择 5 段高密度人群视频作为测试视频,视频选择在佛山通济桥牌坊楼梯入口处所拍摄的视频.场景中人群密度极高,光照变化也很大,如图 3 所示.



图3 实验所选择的视频截图

4.3 实验参数

提取边缘图像阶段:梯度二值化阈值 threshold 为 70,由于实验视频是夜间拍摄的,场景中光照变化比较大加上人群密度高,因此要设置一个比较高的阈值来过滤掉噪声,但是同时又要保证历史模板中含有足够的边缘信息.

计算运动历史图像阶段:用当前帧序号代替当前时间.时间戳设置为 10,实验视频中帧间的目标运动比较小,因此需要设置一个较大的时间戳以保证运动历史图像记录了足够的历史运动信息.

计算有效运动图像阶段:阈值 min 和 max 分别设为 2 和 10.

训练与预测阶段:SVM 的类型是 ϵ 类支持向量回归机,采用高斯核函数,惩罚因子 C 取值为 4,核函数参数 γ 取值为 1,参数 ϵ 取值为 1.

4.4 实验结果

实验的训练样本是从视频中随机抽取的 500 个截面计数块,测试集是从视频中随机抽取的 100 帧图像,为了方便对比,每一个实验都采用相同的测试集.实验训练阶段,人工标记 500 个截面计数块内的行人数,将标记结果和提取的特征送入 SVM 中进行训练.实验预测阶段,预测第 i 帧图像中各个截面计数块中的行人数,累加后得到第 i 帧图像中的行人数 $G(i)$,并人工标记第 i 帧图像中的实际行人数 $T(i)$.将 100 帧图像的平均相对错误 (MRE) 作为实验结果的评价标准. MRE 的计算公式为

$$MRE = \frac{1}{N} \sum_{i=1}^N \frac{|G(i) - T(i)|}{T(i)} \quad (7)$$

实验选择了 5 段高密度人群视频作为测试视频,并人工统计通过流量计数带的真实行人人数,计数过程如图 4 所示,真实行人人数与预测行人人数如图 5 所示,评价参数见表 1.



图4 本文提出的行人计数方法过程截图

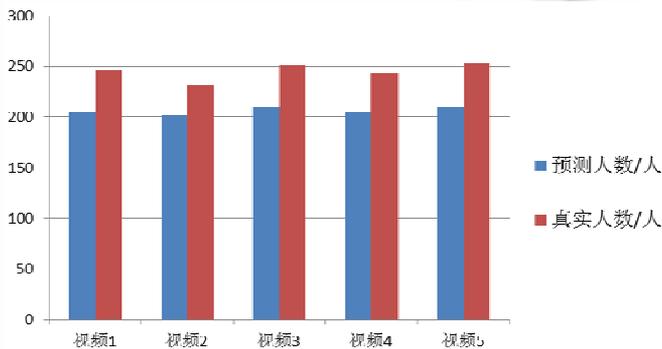


图5 系统预测人数与真实人数对比

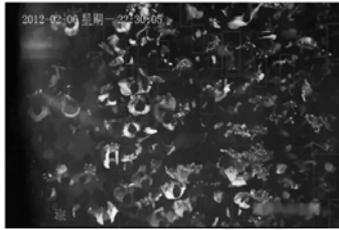
表1 实验评价参数

视频序号	长度/帧	预测人数(人)	真实人数(人)	错误率(%)	准确率(%)	处理时间(ms)
视频 1	748	206	246	16.26	83.74	84.44
视频 2	742	202	231	12.55	87.45	81.63
视频 3	740	210	251	16.33	83.67	82.38
视频 4	697	204	243	16.04	83.96	81.81
视频 5	800	210	253	17.00	83.00	77.78

实验结果表明,本文提出的行人计数方法在高密度人群场景中的准确率可以达到 80%以上,且每帧的处理时间基本可以满足实时的要求,有效地解决了传统计数方法在高密度人群场景中的问题.

4.5 本文方法与其他算法的比较

本文提出的算法较基于行人检测跟踪的文献[2]算法和采用 Haar 特征^[24]的人头检测算法有更高的准确性.图 6 是文献[2]算法和人头检测算法对测试视频人群计数结果截图.



(a) 文献[2]算法在测试视频检验结果



(b) 采用 Haar 特征的人头检测算法在测试视频检测结果

图6 其他算法在测试视频检测结果截图

文献[2]算法基本思想是利用 HOG 特征达到行人检测的目的.可是在高密度人群中,其检测效果很不理想.在图 6(a)中,文献[2]只检测到 27 个人(误检 1 人),而当前帧实际人数大约为 94 人.文献[2]在高密度人群中检测准确率低的原因主要有:摄像机位置、遮挡、行人姿态以及夜间光照不均.

采用 Haar 特征的人头检测算法在高密度人群中检测效果也不是很好,在图 6(b)中,该算法只检测 21 人,当前帧实际人数约为 90 人,人头存在严重的漏检.该算法准确率低的原因主要有:行人头部所占像素比较少,场景复杂、光照变化大,行人手持物遮挡.

4.6 前景提取方法的比较

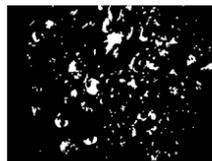
本文将基于梯度运动历史图像的前景检测方法和 Vibe 背景差分法、MOG 背景差分法、帧间差分法作比较,检测结果如图 7 所示.



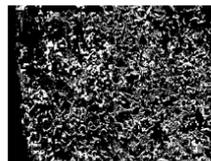
(a) 原始图像



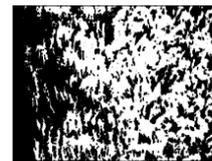
(b) Vibe 背景差分法



(c) MOG 背景差分法



(d) 帧间差分法



(e) 本文提出的前景检测方法

图7 前景提取方法的对比

可以看出,在密度人群场景中,ViBe 背景差分法、MOG 背景差分法和帧间差分法 3 种前景检测方法都无法有效提取出前景图像,总结算法失效的主要原因有如下几点:

- (1) 背景被持续遮挡,无法提取出完整背景图像.
- (2) 场景中 80%以上是运动像素,因此前景像素几乎覆盖了整个图像.这种情况下前景像素无法为后续的处理提供有效信息.
- (3) 场景光照不足,传统的前景提取算法不能有效地检测出运动目标.
- (4) 由于人群密集、场景复杂,导致图像中存在大量噪声.帧间差分法不仅无法提取出完整的前景还会存在大量的误检.

4.7 特征选择的比较

本文分别测试了有效运动像素特征、梯度像素特征、特征点特征和纹理特征在高密度人群场景中的性能,测试结果见表 2.

表2 4个特征实验结果对比

特征提取方法	平均相对错误(MRE) (%)	平均处理时间(ms)
有效运动特征	9.18	79.627 5
梯度像素特征	9.69	7.6153
特征点特征	9.88	60.584 4
纹理特征	10.21	251.87 4

根据对比结果,可以看到在本实验的密集人群场景中纹理特征的错误率比梯度像素特征高,其原因是:第 1,实验视频是俯视拍摄,透视变换小且行人间基本没有遮挡现象,因此梯度像素特征的错误率较小;第 2,实验视频中光照变化大、噪声多对纹理特征的计算影响较大,因此纹理特征的错误率较大.

4.8 速度提取方法的比较

本文还进一步对比了基于运动历史图像的方法、基于光流的方法(Harris 角点)、基于光流的方法(FAST 特征点)和基于跟踪的方法 4 种速度提取方法,平均处理时间测试结果见表 3.

表3 4种速度提取方法平均处理时间

速度提取方法	平均处理时间(ms)
基于运动历史图像的方法	63.374
基于光流的方法(Harris 角点)	343.634
基于光流的方法(FAST 特征点)	5.1427
基于跟踪的方法	343.688

另外,由于实验视频场景复杂,在选用基于跟踪的速度提取方法时,会出现丢失目标的情况,如图 8 所示.

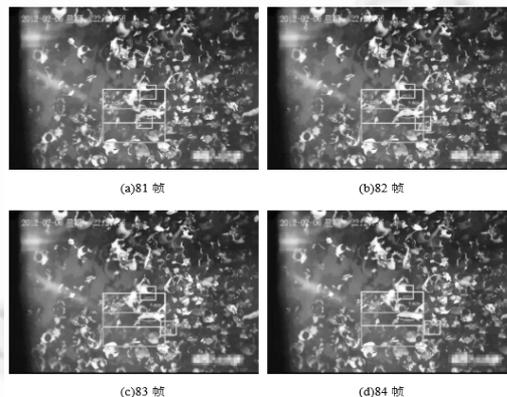


图8 基于跟踪的方法跟踪不准确情况

通过对比可发现,在速度提取方法上,采用基于运动历史图像的方法和基于光流的方法(FAST 特征点)是比较理想的.

5 结 论

本文通过分析现有行人计数方法在高密度人群场景中失效的原因,利用梯度运动历史图像改进了前景检测的结果,并将提取出来的有效运动像素特征和梯度像素特征、特征点特征组合为新的人群特征向量,结合速度提取方法实现了行人计数.实验结果表明,本文提出的计数方法在高密度人群场景中具有较高的准确率和实时性.

本文提出的算法在实际运用中仍有许多不足,今后可以从以下几个方面进行深入的研究:可以根据摄像机标定参数对图像进行透视矫正,从而提高算法的准确率;本文实验采用的视频中行人运动方向一致,因此本文没有考虑对人群进行分割,在今后的研究中可以根据运动方向对提取到的前景图像进行分割,从而使系统可以统计各个方向运动的行人数量;本文提出了有效运动特征并和其他常用特征结合运用到行人流量统计中,但由于光照和手持遮挡物的影响,计数准确率有待进一步提高,需要更深入地研究人群特征的选取.

致谢 感谢郑慧诚博士在数据获取以及修改建议方面提供了帮助.

References:

- [1] Haritaoglu I, Harwood D, Davis LS. Hydra: Multiple people detection and tracking using silhouettes. In: Proc. of the Int'l Conf. on Image Analysis and Processing. Venice: IEEE Computer Society, 1999. 280–285. [doi: 10.1109/VS.1999.780263]
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. San Diego: IEEE Computer Society, 2005. 886–893. [doi:10.1109/CVPR.2005.177]
- [3] Yu SS, Chen XP, Sun WP, Xie DP. A robust method for detecting and counting people. In: Proc. of the Int'l Conf. on Audio, Language and Image Processing. Shanghai: IEEE Computer Society, 2008. 1545–1549. [doi: 10.1109/ICALIP.2008.4590257]
- [4] Lin SF, Chen JY, Chao HX. Estimation of number of people in crowded scenes using perspective transformation. IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans, 2001,31(6):645–654. [doi: 10.1109/TSM.2001.2030598]
- [5] Wen JJ, Xu Y, Zhan YW. People counting based on AdaBoost and inter-frame features. Journal of Image and Graphics, 2011,16(9): 1729–1735 (in Chinese with English abstract)
- [6] Zeng CB, Ma HD. Robust head-shoulder detection by PCA-based multilevel HOG-LBP detector for people counting. In: Proc. of the 2010 Int'l Conf. on Pattern Recognition. Istanbul: IEEE Computer Society, 2010. 2069–2072. [doi: 10.1109/ICPR.2010.509]
- [7] Xu HZ, Lv P, Meng L. A people counting system based on head-shoulder detection and tracking in surveillance video. In: Proc. of the Int'l Conf. on Computer Design and Applications. Qinhuangdao: IEEE Computer Society, 2010. 1: V1-394-V1-398. [doi: 10.1109/ICDDA.2010.5540833]
- [8] Subburaman VB, Descamps A, Carincotte C. Counting people in the crowd using a generic head detector. In: Proc. of the 2012 IEEE Ninth Int'l Conf. on Advanced Video and Signal-Based Surveillance. Beijing: IEEE Computer Society, 2012. 470–475. [doi: 10.1109/AVSS.2012.87]
- [9] Kettner V, Zabih R. Counting people from multiple cameras. In: Proc. of the IEEE Int'l Conf. on Multimedia Computing and Systems. Florence: IEEE Computer Society, 1999. 267–271. [doi: 10.1109/MMCS.1999.778358]
- [10] Ali S, Shah M. Floor fields for tracking in high density crowd scenes. In: Proc. of the European Conf. on Computer Vision. Marseille: Springer-Verlag, 2008. 1–14. [doi: 10.1007/978-3-540-88688-4_1]
- [11] Antonini G, Thiran JP. Counting pedestrians in video sequences using trajectory clustering. IEEE Trans. on Circuits and Systems for Video Technology, 2006,16(8):1008–1020. [doi: 10.1109/TCSVT.2006.879118]
- [12] Rabaud V, Belongie S. Counting crowded moving objects. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2006,1:705–711. [doi: 10.1109/CVPR.2006.92]
- [13] Fujisawa S, Hasegawa G, Taniguchi Y, Nakano H. Pedestrian counting in video sequences based on optical flow clustering. Int'l Journal of Image Processing, 2013,7(1):1–16.

- [14] Chan AB, Liang ZSJ, Vasconcelos N. Privacy preserving crowd monitoring: Counting people without people models or tracking. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. Anchorage: IEEE Computer Society, 2008. 1–7. [doi: 10.1109/CVPR.2008.4587569]
- [15] Qin XH, Wang XF, Zhou X, Liu YF, Li YQ. Counting people in various crowded density scenes using support vector regression. Journal of Image and Graphics, 2013,18(4):392–398 (in Chinese with English abstract).
- [16] Guo S, Liu W, Yan HP. Counting people in crowd open scene based on grey level dependence matrix. In: Proc. of the Int'l Conf. on Information and Automation. Zhuhai, Macau: IEEE Computer Society, 2009. 228–231. [doi: 10.1109/ICINFA.2009.5204926]
- [17] Zhang JP, Tan B, Sha F, He L. Predicting pedestrian counts in crowded scenes with rich and high-dimensional features. IEEE Trans. on Intelligent Transportation Systems, 2011,12(4):1037–1046. [doi: 10.1109/TITS.2011.2132759]
- [18] Albiol A, Silla MJ, Albiol A, Mossi JM. Video analysis using corner motion statistics. In: Proc. of the IEEE Int'l Workshop on Performance Evaluation of Tracking and Surveillance. 2009. 31–38.
- [19] Chan AB, Vasconcelos N. Counting people with low-level features and Bayesian regression. IEEE Trans. on Image Processing, 2012,21(4):2160–2177. [doi: 10.1109/TIP.2011.2172800]
- [20] Gao CW, Huang KQ, Tan TN. People counting using combined feature. In: Proc. of the 2011 3rd Chinese Conf. on Intelligent Visual Surveillance. Beijing: IEEE Computer Society, 2011. 81–84. [doi: 10.1109/IVSurv.2011.6157030]
- [21] Bradski GR, Davis JW. Motion segmentation and pose recognition with motion history gradients. Machine Vision and Applications, 2002,13(3):174–184.
- [22] Yang DD, Chang DH, Han X, Liu Y. Improvement and realization of the moving object detection and tracking. Laser & Infrared, 2010,40(2):205–209 (in Chinese with English abstract).
- [23] Rosten E, Drummond T. Machine learning for high-speed corner detection. In: Proc. of the European Conf. on Computer Vision. Graz: Springer-Verlag, 2006. 430–443. [doi: 10.1007/11744023_34]
- [24] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. Kauai: IEEE, 2001.1:1-511-1-518 vol.1. [doi: 10.1109/CVPR.2001.990517]

附中文参考文献:

- [5] 文嘉俊,徐勇,战荫伟.基于 AdaBoost 和帧间特征的人数统计.中国图象图形学报,2011,16(9):1729–1735.
- [15] 覃勋辉,王修飞,周曦,刘艳飞,李远钱.多种人群密度场景下的人群计数.中国图象图形学报,2013,18(4):392–398.
- [22] 杨冬冬,常丹华,韩夏,刘宇.运动目标检测与跟踪算法的改进与实现.激光与红外,2010,40(2):205–209.



纪庆革(1966—),男,山东临沂人,博士,副教授,主要研究领域为计算机图形学,计算机视觉,虚拟现实.

E-mail: issjqg@mail.sysu.edu.cn



迟锐(1988—),男,硕士生,主要研究领域为计算机视觉.

E-mail: chirui0602@163.com



陈婧(1988—),女,硕士,主要研究领域为计算机视觉.

E-mail: jing.chen.cj1010@gmail.com



方贤勇(1978—),男,博士,教授,主要研究领域为计算机图形学,计算机视觉,模式识别.

E-mail: fangxianyong@ahu.edu.cn