

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)04-0606-07

论文索引信息: 张莉, 于海滨. 双摄像机下人体遮挡时的跟踪方法 [J]. 中国图象图形学报, 2011, 16(4): 606-612

双摄像机下人体遮挡时的跟踪方法

张莉¹⁾, 于海滨²⁾

¹⁾(杭州电子科技大学计算机学院, 杭州 310018) ²⁾(杭州电子科技大学电子信息学院, 杭州 310018)

摘要: 针对双摄像机下存在人体遮挡情况时的跟踪问题, 提出了利用人体3维位置信息来实现跟踪的方法。该方法首先对其中一个摄像机视频图像中的人体像素抽样, 接着在其他摄像机视频图像中找出抽样像素的匹配点, 计算出每一对匹配点在世界坐标系中所对应的3维点, 然后依据3维位置信息将3维点聚类, 找出每一个聚类区域中的3维点所对应的图像中的一组像素点, 并对其构建高斯平滑直方图模型。在此基础上, 依据直方图模型将互相遮挡的人体分割开来, 最后根据求得的人体像素点的匹配关系来确定不同摄像机中同一个人的对应关系。实验结果表明, 该方法能有效实现遮挡情况下的人体跟踪。

关键词: 人体跟踪; 视频监控; 遮挡; 高斯平滑直方图

Human tracking method under occlusion with two cameras

Zhang Li¹⁾, Yu Haibin²⁾

¹⁾(School of Computer Science, Hangzhou Dianzi University, Hangzhou 310018 China)

²⁾(School of Electronic & Information, Hangzhou Dianzi University, Hangzhou 310018 China)

Abstract: The cameras in this study are used to extend view field. The paper proposed a three-dimensional information based human tracking method. The method firstly samples the pixels within the human body in one of the video images, and then finds the matching points in the other simultaneous video images by polar line constraints. The three-dimensional points in space corresponding to each pair of the matches are obtained by triangulation and are also clustered. The pixels in the image plane corresponding to each clustering region which includes the three-dimensional points are decided. Finally, Gaussian-smoothing histograms are created based on the pixels in the image plane. Occluded people would be divided into each other by the Gaussian-smoothing histograms. The matching relation of each person in different cameras could be found and thus the human tracking is realized. Experimental results show the efficiency of the method.

Keywords: human tracking; video surveillance; occlusion; Gaussian-smoothing histograms

0 引言

由于噪声和遮挡等原因, 不同的人体有可能运动到一起, 合并成互相遮挡的人群, 某一时刻遮挡人群又可能分裂成多个独立的人体, 在多摄像机人体跟踪中经常会遇到这些不利情况, 因此研究遮挡情况出现时的跟踪方法成为人体跟踪技术研究的热点问题。

目前对遮挡情况下人体跟踪问题的研究已取得一定的进展。动态分层表示方法^[1-2]从理论上建立了较完善的解决方案, 但是检测精度有待提高。基于概率统计的跟踪方法^[3-6]有很大的计算量。基于交迭系数的判断法和基于区域的估计面积与观测面积之差的判断方法^[7], 通过建立遮挡判决函数来检测遮挡。但是这类检测方法对目标区域的面积变化较为敏感, 难以区分目标区域变化是因为目标运动

收稿日期: 2009-11-26; 修回日期: 2010-01-04

基金项目: 浙江省自然科学基金项目(Y1090881)。

第一作者简介: 张莉(1979—), 女, 讲师。2008年获浙江大学通信与信息系统专业博士学位, 主要研究方向为模式识别、人工智能。E-mail: dr.zhangliangie@yahoo.com.cn。

造成的还是因为出现遮挡造成的,从而造成很高的误判率。Utsumi 等人^[8]通过选择最好的视点来解决人与人之间的互遮挡问题。

成像角度的限制是造成人体遮挡的主要原因。当从摄像机坐标系中观察,多个人的水平和竖直位置出现重叠时,人体之间就会发生相互遮挡的情况,这时人体只有部分是可见的,为了减少由于遮挡所带来的目标识别歧义性问题,必须采用更好的方法来处理遮挡时特征与身体各部分之间的准确对应问题。由于人体在空间上是实体,所以人体目标虽然在图像平面上相互重叠,但他们与摄像机的距离是不相同的,可以通过距离上的划分来达到遮挡人体目标的分割,进而可以有效地实现人体跟踪。

本文重点讨论如何利用人体 3 维位置信息来实现双摄像机下遮挡情况时的人体跟踪,算法流程如图 1 所示。首先介绍人体区域匹配点的提取方法,它是获取人体 3 维信息的基础。

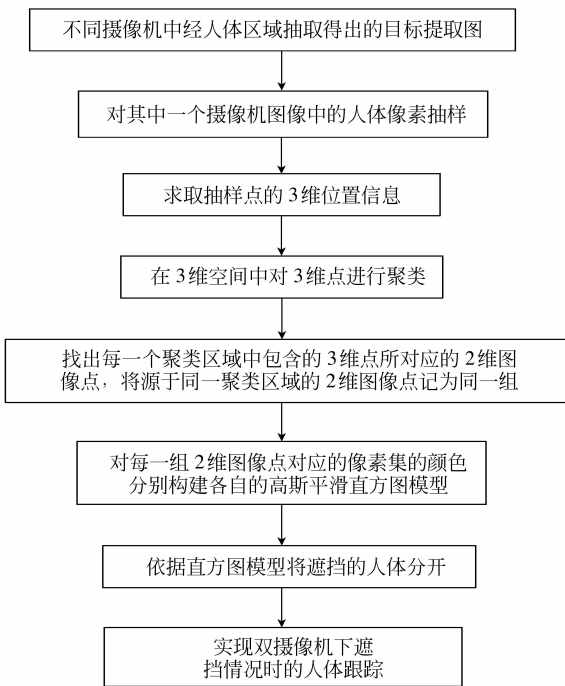
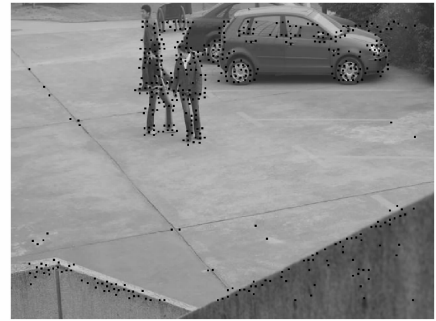


图 1 双摄像机下遮挡情况时的人体跟踪算法的流程图

Fig. 1 Flowchart of human tracking algorithm under occlusion with two cameras

1 人体区域匹配点的获取

实验发现,由 KLT^[9]和 SIFT^[10]等较成熟的方法得到的人体区域的特征点很少(见图 2)。(a)是



第 1 帧

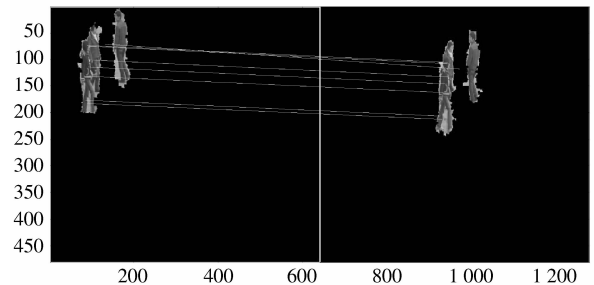


第 9 帧



第 17 帧

(a) 基于 KLT 方法的特征点提取, 人体上的特征点较少



(b) 基于 SIFT 方法的匹配点提取, 匹配点较少

图 2 基于 KLT 和 SIFT 方法的特征点匹配示例

Fig. 2 Feature points matching results using KLT and SIFT

KLT 跟踪示例,黑色点是提取出的特征点。从(a)中可以看到,人体上所能跟踪到的特征点很少,而且过了少数几帧后,已经得不到任何特征点了,因此无法进行特征点匹配,也无法求取人体 3 维坐标。(b)是用 SIFT 方法得出的特征点匹配结果,灰色线

连接的一对点是相互匹配的一对点,可以看出人体上同样得不到足够多的匹配点,这样不可能获得足够多的人体 3 维点来实现相互遮挡的人体在深度方向上(图 5 中的 Z 方向)的分割。

由于视频监控系统中摄像机间的距离通常比较大,不同摄像机的成像角度及光照条件差别较大,匹配点的搜索过程远比短基线图像对的匹配点搜索困难。在这种基线较长的多摄像机系统中,两个匹配

$$S_{ZNCC}(i',j') = \frac{\sum_{(u,v) \in W_s^{(i,j)}} [T(i'+u, j'+v) - \bar{T}(i',j')] \cdot [M(u,v) - \bar{M}]}{\sqrt{\sum_{(u,v) \in W_s^{(i,j)}} [T(i'+u, j'+v) - \bar{T}(i',j')]^2 \cdot \sum_{(u,v) \in W_s^{(i,j)}} [M(u,v) - \bar{M}]^2}} \quad (1)$$

式中, \bar{T} 、 \bar{M} 分别为相应窗口内的像素值的均值。通过搜索相似性测度的极值即可找到匹配点 (i',j') 的位置。

实际操作中,为了减小计算时间,可以适当做抽样处理,对人体区域的抽样像素点用极线约束方法获取其匹配点。本文在其中一个摄像机的人体检测图中按每 3 行 3 列抽取一个像素的采样方式来抽取像素点,示例如图 3 所示。



图 3 人体区域的抽样点示例
Fig. 3 Example of pixel sampling

点的距离可达 100 ~ 400 像素,在这么大的搜索范围内可能存在许多局域极小值,因此本文采用极线约束方法寻找人体目标区域中的匹配点。

基于极线约束的特征点匹配 若设其中一幅视图 M 为基准图,另一幅视图 T 为对准图,为了消除两幅视图之间由光照引起的亮度差异,则 M 中坐标为 (i,j) 的像素点在 T 中对应的外极线上的搜索位置 (i',j') 处的相似测度可采用测度 S_{ZNCC} 。

当选取好抽样点之后,便可以按照式(1)在另一幅图像上搜索匹配点,结果如图 4 所示,其中蓝色线是极线,红色点和黄色点是搜索到的匹配点。



图 4 2 维匹配点及其分组
Fig. 4 Matching points and theirs grouping

2 基于人体 3 维位置信息的跟踪方法

本文利用人体的 3 维位置信息来实现相互遮挡的人体的跟踪,只需要得到遮挡的人体目标到摄像机的距离(Z 方向)的相对划分即可,如图 5 所示。

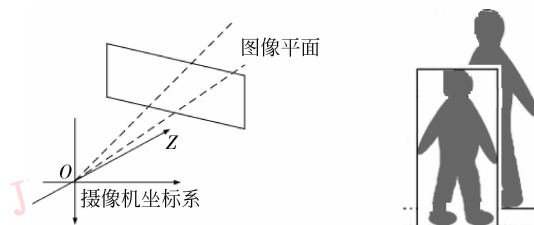


图 5 遮挡人体深度方向上的相对划分的示意图
Fig. 5 Separation of occluded persons in depth (Z direction)

2.1 人体 3 维位置的求解

对于 3 维空间中人体上的任意一点 $M = [X \ Y \ Z \ 1]^T$,它在不同摄像机中的投影点为

$m = [x \ y \ 1]^T, m' = [x' \ y' \ 1]^T$,可以得到投影等式 $m = PM, m' = P'M$,将这两个等式联立起来,可以得到

$$AM = 0, A = \begin{bmatrix} xp_3^T & p_1^T \\ yp_3^T & p_2^T \\ x'p_3'^T & p_1'^T \\ y'p_3'^T & p_2'^T \end{bmatrix} \quad (2)$$

式中, p_i^T 为投影矩阵 P 的第 i 行。

每个 M 都可以建立类似于式(2)的4个等式,在已知 m, m' 的情况下,可以利用SVD分解法,直接解出,据此实现遮挡情况时的人体跟踪。

为了有效实现遮挡时的人体跟踪,首先将人体3维点在3维空间中进行聚类,方法可见文献[11]或文献[12]。如图4所示,红色点和黄色点是经过3维点聚类后,每一聚类区域包括的3维点所对应的2维图像点,它们也是对应于图3中的抽样点的匹配点。

由前文所述,本文并没有对每个人体区域的所有像素点进行匹配,只是对抽样点按照极线约束方法在另一幅图像中寻找与其匹配的2维点,因此本文采用如下3个步骤确定每个像素的归属:

- 1) 将每一个聚类区域中包含的3维点所对应的2维像素点记为同一组;
- 2) 对每一组像素点集的颜色值分别建立各自的高斯平滑直方图模型;
- 3) 依据建立的高斯平滑直方图模型,通过计算人体区域中每个像素点与所有直方图模型的相似度来确定该像素的归属。

2.2 高斯模板平滑的直方图模型的建立

本节利用2维像素点集的颜色值构建的高斯平滑直方图模型^[13]来区分2维平面上相互遮挡的人体,此处是指分别对图4中的黄色点和红色点所对应的像素集的颜色值建立颜色直方图模型。通过对众多视频序列的测试发现,将像素的每1维颜色值分为64段,64³个区间(bin)有利于区分不同的人体。另外,由于相邻像素的颜色一般是相关的,因此本文用大小为7³的高斯模板对直方图进行平滑,平滑的方式如图6所示,为了便于显示,该图显示的是2维平滑直方图,图中白色圆点的位置代表抽样像素的颜色值(此处只显示了红色R和绿色G),该颜色值对其周围不同直方图区间的数值的贡献值不同,图中不同灰度的方格代表的数值表示对周围直

方图区间的数值的贡献值。

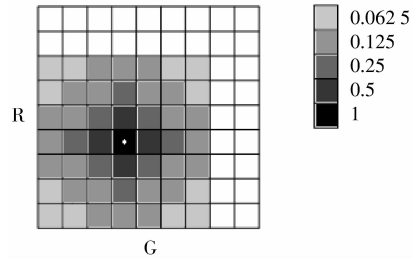


图6 高斯模板(7×7)平滑的2维直方图模型

Fig. 6 Model of Gaussian smoothing two dimensional histogram

2.3 相似性计算

根据2.2节得出的高斯平滑直方图模型,我们可得到一幅与人体目标提取图 O (由于这部分内容不是本文重点,此处不再赘述,详见文献[14])相同尺寸的概率图 P , 概率图 P 中 (x, y) 处的数值表示人体目标提取图 O 中 (x, y) 处的像素的颜色值属于某个人体直方图模型的相对概率,其概率(或相似性)计算如下

$$P(p(R, G, B)) = \frac{m}{m + n} \quad (3)$$

式中, m 和 n 分别为人体上像素 p 的 R, G, B 值落入不同人体的直方图模型中同一区间的值,如图7所示。

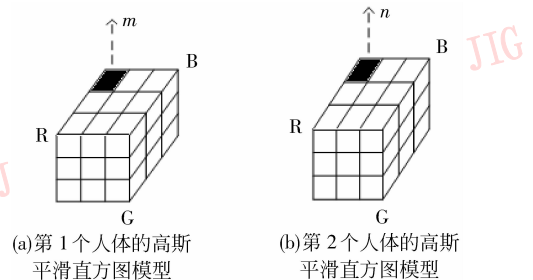


图7 相似性计算

Fig. 7 Example of similarity calculation

相似性计算的示例如图8所示,根据式(3),对原始图像(a)中的人体像素计算出各自的概率,得到的概率图如(b)所示,不同的灰度表示不同的概率。由结果可以看出,已能有效地区分出每个像素所属的人体。

2.4 遮挡人体的分割

至此,已能确定所有像素的归属,但是还必须对概率图进行分割,才能将遮挡人体分割开来,采用基于一致性检验的二值化方法^[15],并结合形态学方

法,即可将图 8(b)划分为两部分,如图 9 所示,由于二值化的原因,图中只能看到其中显示为黑色的像素集,原始图像中与这些黑色像素对应的像素集表示其中的一个人体。

同理,可以得到另一幅视频图像中遮挡人体的分割图像。同一个人体在不同摄像机中的对应关系则由 1.1 节获取的人体区域像素点的匹配关系决定,这样就完成了双摄像机下的人体跟踪。

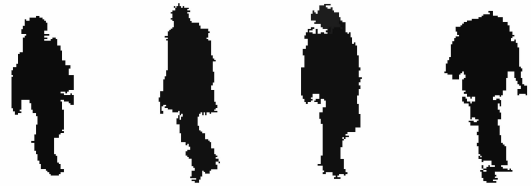
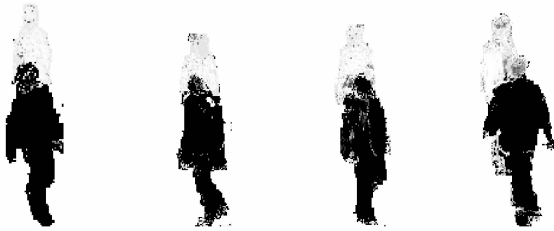


图 9 图 8(b)的二值化图像

Fig. 9 Binarized images of Fig. 8(b)



(a) 原始图像



(b) 概率图

图 8 相似性计算结果

Fig. 8 Similarity calculation results

3 实验结果

为了和其他的方法(如基于颜色模型的跟踪方法)进行对比,本文给出了纯粹用颜色模型进行跟踪得到的结果(由于这部分内容不是本文重点,此处不再赘述,详见文献[15]),结果如图 10 所示。

与图 10(a)比较,该方法能将(a)中深红色矩形框中两个彼此遮挡的人有效分开,如(b)所示,直接证明了该方法的可行性。在利用颜色模型进行跟踪时,不能将深红色矩形框中的遮挡人体分割为两个人,因此在判定最终的颜色模型时,只能将这两个人体归类到同一种颜色模型,否则不可避免地导致跟踪错误。



(a) 基于颜色模型的跟踪方法得到的结果



(b) 利用 3 维信息得到的结果

图 10 遮挡情况时不同方法下的人体跟踪结果比较

Fig. 10 Comparison of human tracking results using different methods under occlusion

最后,本文以两个摄像机构成的视频监控系
统为例,其中摄像机离地面 4.5 m 左右,摄像机间的夹
角为 50°。图 11 给出一组跟踪结果。



图 11 室外环境下同步视频序列的人体跟踪结果

Fig. 11 Human tracking and correspondence results in outdoors

由图 11 可以看出,本文算法能得到较好的跟踪
结果,为得出该方法的跟踪性能,本文对上述视频序

列中不同人体的正确跟踪次数进行了统计,结果如
图 12 所示。

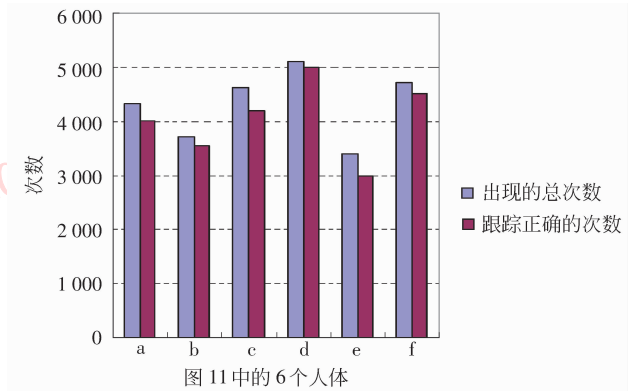


图 12 双摄像机下不同人体被正确跟踪的次数比较

Fig. 12 Comparison between the appearance times and the tracking valid times

图 12 中,出现的总次数是指每个人体在左、右
摄像机中出现的次数之和,跟踪正确的次数是指每
个人体在左、右摄像机中分别被正确跟踪的次数
之和。

4 结 论

讨论了双摄像机下有遮挡情景出现时的人体跟
踪问题,鉴于 2 维成像会造成信息丢失,本文利用了
人体 3 维位置信息来实现遮挡情况下的人体跟踪。
首先对其中一个摄像机视频图像中的人体像素进行
抽样,通过极线约束方法求取抽样像素在另一个摄
像机视频图像中所对应的匹配点,然后求取 3 维点
的坐标,并对 3 维点进行聚类,将每一个聚类区域中
的 3 维点所对应的 2 维匹配点记为同一组,并对其
建立高斯平滑直方图模型,依据该模型将相互遮挡
的人体分割开来。不同摄像机中同一个人体的对应
关系由获得的人体区域像素点的匹配关系确定,从
而实现人体跟踪。该方法合理利用了人体的空间位
置信息,因而对有遮挡情况及人体衣着颜色相近情
况发生的跟踪问题处理的比较好,同时也表明多摄
像机比单摄像机能提供更丰富的信息,从而能更有
效地实现人体跟踪。

参考文献 (References)

[1] Tao H, Sawhney H S, Kumar R. Object tracking with Bayesian

- estimation of dynamic layer representations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(1):75-89.
- [2] Ristivojevic M, Konrad J. Space-time image sequence analysis: object tunnels and occlusion volumes [J]. IEEE Transactions on Image Processing, 2006, 15(2):364-376.
- [3] Zhao T, Nevatia R, Wu B. Segmentation and tracking of multiple humans in crowded environments [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(7):1198-1211.
- [4] Wang Xin, Liu Lei, Tang Zhenmin. Infrared human tracking with improved mean shift algorithm based on multicue fusion [J]. Appl. Opt., 2009, (48): 4201-4212.
- [5] Fleuret F, Berclaz J, Lengagne R, et al. Multi camera people tracking with a probabilistic occupancy map [J]. IEEE Transactions Pattern Anal. Mach. Intell., 2007, 30(2):267-282.
- [6] Yu Y, Harwood D, Yoon K, et al. Human appearance modeling for matching across video sequences [J]. Mach. Vis. Appl., 2007, 18(3-4):139-149.
- [7] Zhang Zexu, Li Jinzong, Li Dongdong. Robust study of multi-feature points object tracking [J]. Journal of Data Acquisition and Processing, 2003, 18(4):423-428. [张泽旭, 李金宗, 李冬冬. 一种运动目标多特征点的鲁棒跟踪方法研究 [J]. 数据采集与处理, 2003, 18(4):423-428.]
- [8] Utsumi A, Mori H, Ohya J, et al. Multiple-view-based tracking of multiple humans [C] // Proceedings of the 14th International Conference on Pattern Recognition. Washington, DC: IEEE Computer Society, 1998: 597-601.
- [9] Shi J, Tomashi C. Good feature to track [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Ithaca, USA, Cornell University, 1994: 593-600.
- [10] Lowe D. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [11] Keck Jr M A, Davis J W, Tyagi A. Tracking mean shift clustered point clouds for 3D surveillance [C] // Proceedings of the International Conference on VSSN. Santa Barbara, California, USA, ACM, 2006: 187-194.
- [12] Xia S X, Li W C, Zhou Y, et al. Improved k-means clustering algorithm [J]. Journal of Southeast University, 2007, 23(3): 435-438.
- [13] Kristensen D. Autonomous Road Following—a Study of Methods for Tracking Unmarking Roads in Image Sequences [D]. Stockholm, Sweden, Royal Institute of Technology, 2004: 30-33.
- [14] Tabbone S, Wendling L. An original multi-scale algorithm to binarize images [C] // Proceedings of the International Conference on Pattern Recognition. Quebec: IEEE Computer Society, 2004: 704-707.
- [15] Zhang Li. Study of Human Tracking with Multi-cameras [D]. Hangzhou: Zhejiang University, 2008. [张莉. 多摄像机人体跟踪技术的研究 [D]. 杭州: 浙江大学, 2008.]