

运动捕获系统中基于多视觉的目标跟踪算法

张欢欢, 郑江滨

(西北工业大学 计算机学院, 陕西 西安 710129)

摘 要: 基于标记点的运动捕获系统中, 由于标记点特征相似, 易出现遮挡, 使得跟踪难度大. 针对上述问题提出了基于多视觉的跟踪算法. 首先进行双目视觉立体跟踪: 采用扩展 Kalman 进行预测, 根据外极线约束剔除预测区域内的错误候选目标; 然后进行多目视觉数据融合: 利用基于可变阈值的最邻近数据融合算法对已获得的多个双目跟踪数据进行处理, 解决了标记点遮挡、丢失问题; 最后得到了标记点的三维运动数据. 实验表明, 提出的跟踪算法可以准确跟踪各个标记点.

关键词: 运动捕获系统; 双目视觉立体跟踪; 多目视觉数据融合

中图分类号: TP391

文献标识码: A

文章编号: 1000 - 7180(2009)06 - 0066 - 04

Multi-view Object Tracking for Motion Capture

ZHANG Huan-huan, ZHENG Jiang-bin

(School of Computer, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: For the tracked multiple markers in a marker-based motion capture system, their images features are quite similar and they are occluded frequently, that make it become very difficult to track these markers at the same time. This paper proposes a novel multi-view object tracking algorithm, which includes binocular stereo tracking and multi-view tracking data fusing. In the procedure of binocular stereo tracking, the extended Kalman is applied and the epipolar geometry restriction is used to eliminate the incorrect matching. During of multi-view tracking data fusing, a nearest neighbor data fusion algorithm based on adaptive threshold is proposed to process these binocular tracking data, that solves the problem of occlusion and missing effectively, and then 3D tracking data are gained. Experiments show that the proposed algorithm can track multiple markers correctly.

Key words: motion capture system; binocular stereo tracking; multi-view tracking data fusing

1 引言

基于标记点的光学运动捕获系统^[1]通过捕获人体上的标记点来重建人体的三维运动, 已经成功地应用于影视特效、动画制作、虚拟现实、模拟训练等研究领域. 现有的捕获系统通常采用基于特征匹配的跟踪方法^[2-4]. 文献[2]中以不同的色块作为标记点, 提出一种基于色块的跟踪方法. 文献[3]中使用少量彩色色带作为标记点, 以人体结构约束信息为辅助进行跟踪. 文献[4]中利用标记点的轮廓特征进行跟踪, 取得了较好的跟踪结果, 但其标记点需要

特殊制作, 不便于实际应用.

上述基于颜色特征的跟踪算法跟踪结果较准确, 且一定程度上可以解决标记点的自遮挡问题. 但因为平面色块在三维视觉中易发生形变, 在不同的摄像头间存在成像色调的差异, 导致跟踪不准确, 且当标记点数量增多时, 可明显区分的颜色有限, 这些都极大地限制了它们的应用. 在文中开发的运动捕获系统中, 采用白色反光小球作为标记点, 使得系统应用起来很便捷, 但是由于这些标记点特征相似, 易出现遮挡等问题, 使得这种基于特征匹配的跟踪算法失效. 为了解决此问题, 文中提出了一种基于多视

收稿日期: 2008 - 09 - 24

基金项目: 国家“八六三”计划项目(2006AA01Z324); 西北工业大学研究生创业种子基金项目(200850)

觉的目标跟踪算法,主要包含双目视觉立体跟踪和多目视觉数据融合这两个模块,其框架如图1所示.

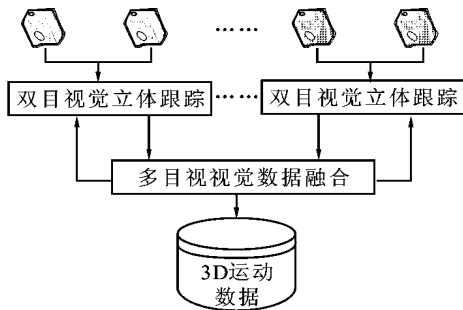


图1 算法框架

首先每个双目模块独立检测和跟踪各个标记点^[5]. 由于基于卡尔曼滤波的跟踪算法精度高,运算量小,对一些复杂运动目标能进行连续稳定的跟踪,文中在跟踪过程中利用卡尔曼滤波器对标记点进行预测,此外,如果在标记点的预测区域内出现多个候选标记点,则利用外极限约束^[6]剔除错误的候选点,使其在双目视觉下能够正确匹配,从而提高跟踪准确度.

然后多目视觉数据融合模块将各个双目的跟踪结果进行处理,最终得到每个标记点的三维轨迹.在处理工程中,尽管各个双目的坐标系已经统一,但是由于图像噪声、标定误差、二维匹配误差等因素的影响使得同一标记点在不同双目下的三维位置并不完全相同甚至相差较大,使得融合难度加大.为解决此问题,文中提出了基于可变阈值的最邻近数据融合算法,很好地解决了双目视觉中出现的遮挡、丢失、丢失后又重新出现等问题.

2 双目立体跟踪算法

由于被跟踪的各个标记点之间无明显特征区别、数量多、易遮挡等特点,传统的匹配跟踪算法已经失效,针对此问题,文中提出了一种新的双目视觉立体跟踪算法,流程如图2所示.

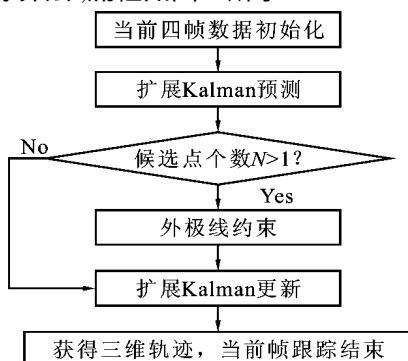


图2 双目视觉立体跟踪算法流程图

首先利用二维扩展 Kalman 分别对两个单目序

列图像中的标记点进行预测跟踪,在此基础上,进行标记点的三维扩展 Kalman 预测跟踪;如果标记点在两个单目序列图像中的预测区域内分别有多个候选标记点出现,则利用双目下的外极限约束条件剔出错误的候选点,找出正确的匹配.

该算法的特点在于:(1)在跟踪过程中,二维扩展 Kalman 和三维扩展 Kalman 同时被用来进行预测,缩小了标记点的搜索区域,减少了由于缺乏特征区别引起的标记点错误跟踪;(2)利用外极限约束条件有效地解决了在双目下多个二维候选标记点的对应问题,剔除错误的对应关系,从而获得正确的候选标记点.

2.1 扩展 kalman 预测

假设当前时刻为第 $k+1$ 时刻,记 p_k, v_k, a_k 为标记点 O_n 在第 k 时刻的位置、速度、加速度.在跟踪过程中,二维和三维扩展 kalman 滤波器同时被用来进行 $k+1$ 时刻的状态预测 ($\hat{p}_{k+1}, \hat{v}_{k+1}, \hat{a}_{k+1}$),如式(1)~式(3).

$$\hat{p}_{k+1} = p_k + v_k \cdot t + \frac{1}{2} a_k \cdot t^2 \quad (1)$$

$$\hat{v}_{k+1} = v_k + (1 - \gamma) \cdot \hat{v}_k \quad (2)$$

$$\hat{a}_{k+1} = a_k + (1 - \gamma) \cdot \hat{a}_k \quad (3)$$

$$\hat{S}_{k+1} = p_k + v_k \cdot t + \frac{1}{2} a_k \cdot t^2 \quad (4)$$

利用式(4)预测标记点的搜索范围 \hat{S}_{k+1} .如果在预测区域 \hat{S}_{k+1} 内检测到标记点 O_n ,则利用 O_n 的新息更新各个状态变量的预测误差 ($p_{k+1}, v_{k+1}, a_{k+1}$),如式(5)~式(7).

$$p_{k+1} = \hat{p}_{k+1} + \gamma (p_{k+1} - \hat{p}_{k+1}) \quad (5)$$

$$v_{k+1} = \hat{v}_{k+1} + \gamma (v_{k+1} - \hat{v}_{k+1}) + (1 - \gamma) \cdot v_k \quad (6)$$

$$a_{k+1} = \hat{a}_{k+1} + \gamma (a_{k+1} - \hat{a}_{k+1}) + (1 - \gamma) \cdot a_k \quad (7)$$

2.2 剔除错误的候选点

如果在标记点 O_n 的预测区域内检测到多个候选标记点,则需要利用外极限约束剔出错误的候选点.

外极线约束原理:假设 $p_1(p_2)$ 为标记点 O_n 在左(右)图像上的位置坐标, F 为此双目的基础矩阵,则 $p_2(p_1)$ 的位置就在与 $p_1(p_2)$ 相关联的外极线 $l_2(l_1)$ 上,如式(8).

$$l_1 = F^T p_2, l_2 = F p_1 \quad (8)$$

这样,对 $p_2(p_1)$ 的搜索就被限制在 $l_2(l_1)$ 上而非整个图像,但是在实际应用中,由于各种噪声及计

算误差的影响,实际检测到的点并不严格满足外极线约束,文中利用候选点与外极线之间的距离作为约束准则,距离越短,二维匹配程度越高,如式(9).

$$\begin{aligned} dis_{ji}(p_{bi}(u_{bi}, v_{bi}), l_j) = \\ |a_j u_{bi} + b_j v_{bi} + c_j| / \sqrt{a_j^2 + b_j^2} \end{aligned} \quad (9)$$

$$q = \operatorname{argmin}(dis_{ji}) \quad i, j = 1, \dots, n.$$

式中, $p_{bi}(u_{bi}, v_{bi})$ 为标记点在左(右)图像上的第个候选点的二维位置坐标. a_j, b_j, c_j 为标记点在右(左)图像上的外极线系数.

3 多目视觉数据融合算法

多目视觉数据融合就是通过综合处理所有双目视觉的跟踪结果,从而得到标记点的三维运动数据.这里只要处理三类标记点的跟踪结果,如图3所示.

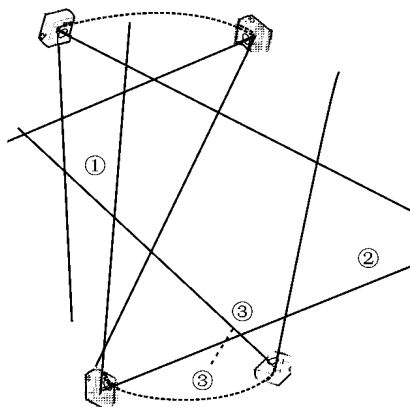


图3 被跟踪标记点的类型

第一类:在双目视觉的公共视野区域内出现且已经正确跟踪的标记点.对于这类标记点,利用式(10)对多个双目跟踪结果加权平均,获得其最终三维运动数据.

$$M_j = \frac{1}{h} \sum_{l=1}^h M_{jl}, j = 1, \dots, n \quad (10)$$

式中, $M_j, j = 1, \dots, n$ 为第 j 个标记点的最终三维运动数据, $M_{jl}, l = 1, \dots, h$ 为其在各个双目视觉下的立体跟踪结果.

第二类:未在任何双目视觉的公共视野区域内出现,但在某两个单目的视野区域内出现且已经正确跟踪的标记点.对于这类标记点,将这两个单目组合成一个双目视觉,获得三维跟踪结果.

第三类:在某个双目视觉下丢失后又重新出现的标记点,称之为新标记点.对于这类标记点,首先需要确定它的标识,即身份.在辨识的过程中,由于图像噪声、标定误差、二维匹配误差等因素的影响使得同一标记点在不同双目下的三维位置坐标不完

全相同甚至相差较大,增加了辨认的难度,针对这个问题,文中提出了基于可变阈值的最邻近数据融合算法,中心思想是:在三维空间中,如果双目 a 中的新标记点与其它双目下的某一个被跟踪的标记点 O_n 的位置或预测位置最近且两者的距离在一定阈值范围内,则认为此新标记点与 O_n 为同一个标记点.与传统的最邻近数据融合算法^[7]相比,此阈值并不是固定不变的,而是随着被跟踪标记点的预测误差变化,即以被跟踪标记点的位置或预测位置为中心,在预测误差的指导下调整标记点的搜索区域,从而确定阈值的取值范围.新标记点一旦被识别出来且连续出现4帧以上,则初始化它的Kalman滤波器,对其进行后续帧的跟踪.

4 实验结果

实验场地四周布置了6台摄像机,由于邻近的摄像机公共视区较大,故将邻近的两两摄像机组成一个双目视觉,这6个摄像机被同时用来进行运动数据的采集.由于原地踏步这一动作有遮挡、有丢失、还有丢失后又重新出现等较难进行跟踪的情况出现,故以原地踏步这一动作为例进行说明.

4.1 双目视觉立体跟踪结果

以两个标记点和在双目视觉中的二维跟踪结果为例进行说明,如图4所示.这两个点在第16到27帧,第45到55帧,第71帧到83帧等时间段发生相互遮挡,然后又彼此分开.从图5中可以看出,文中提出的双目视觉立体跟踪算法在A点和B点发生遮挡时能够准确地跟踪它们,在彼此分开时,也就是在各自的预测区域内有2个候选标记点的情况下,也可以准确地跟踪.

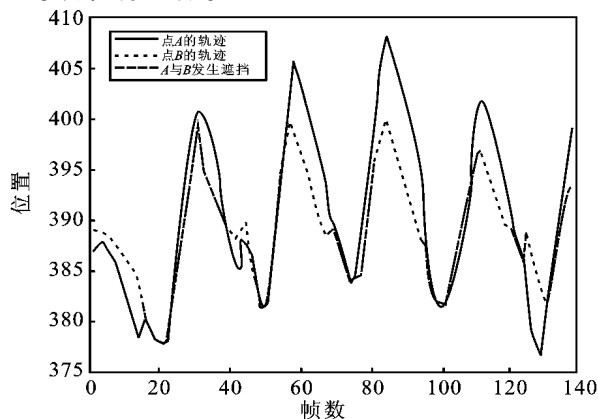


图4 标记点的二维跟踪结果

4.2 多目视觉数据融合结果

以标记点C在双目视觉中a的三维跟踪结果

为例进行说明,如图 5 所示,其中“三维跟踪轨迹 A”是经过可变阈值的最邻近数据融合算法处理后得到的三维运动数据,“三维跟踪轨迹 B”是经过传统的固定阈值的最邻近数据融合算法处理后得到的三维运动数据.标记点从第 30 帧开始丢失,在第 35 帧重新出现直到第 56 帧;从第 57 帧开始丢失,在第 62 帧重新出现直到第 82 帧;在第 83 帧开始丢失,在第 89 帧重新出现直至第 109 帧;从第 110 帧开始丢失,在第 116 帧重新出现直到最后.从图 5 中可以看出,传统算法的识别率极低,文中算法仅在第 63、64、119 帧这三帧中没有识别出来,与传统算法相比,识别率有很大提高.

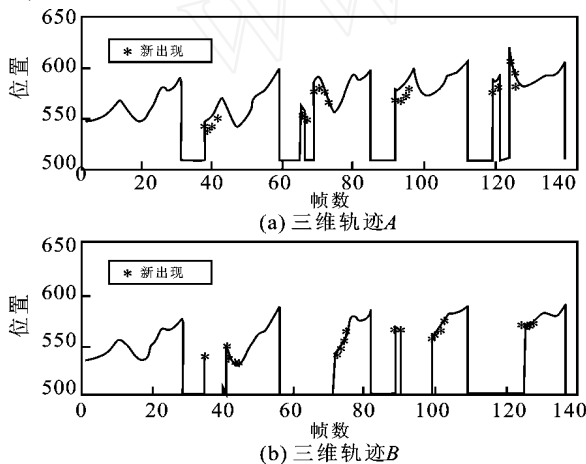


图 5 标记点的三维跟踪结果

4.3 三维人体运动重建结果

整个人体的三维运动重建结果如图 6 所示.从图 6 可以看出,文中提出的基于多视觉的目标跟踪算法可以较逼真地重建人体三维运动.

5 结束语

针对于运动捕获系统中标记点特征相同,易遮挡等问题,文中提出了基于多视觉的目标跟踪算法,包括基于扩展 Kalman 预测和外极限约束的双目视觉立体跟踪算法与多目视觉下基于可变阈值的最邻近数据融合算法.实验结果表明,文中提出的基于多视觉的目标跟踪算法,可以同时精确跟踪多个标记点.在以后的工作中,将实现更多标记点、更复杂的人体运动目标跟踪以及进一步提高新标记点的识别率.

参考文献:

[1] Thomas B Moeslund, Adrian Hilton, Volker Krüger. A

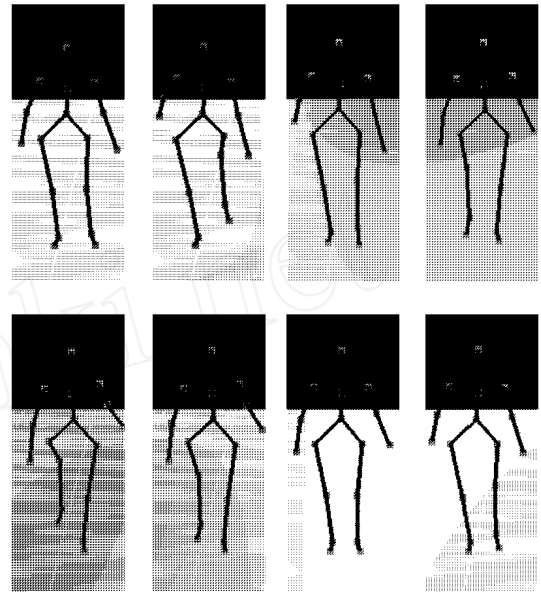


图 6 人体原地踏步运动三维重建结果

survey of advances in vision - based human motion capture and analysis[J]. Computer Vision and Image Understanding, 2006(104): 90 - 126.

- [2] 朱强,庄越挺,陈家实,等. 基于色块的人体运动跟踪[J]. 模式识别与人工智能, 2001, 14(4): 487 - 492.
- [3] Ukida Hiroyuki, Kaji Seiji, Tanimoto Yoshio, et al. Human motion capture system using color markers and silhouette[C]// Proceedings of the IEEE on Instrumentation and Measurement Technology Conference. Japan: Sorrento, 2006: 151 - 156.
- [4] Barca J C, Rumantir G, Koon Li R. A new illuminated contour - based marker system for optical motion capture[J]. Innovations in Information Technology, 2006(11): 1 - 5.
- [5] 晏剑云,郑江滨. 视频三维运动捕获系统中多运动目标跟踪方法[J]. 微电子学与计算机, 2008, 25(8): 88 - 91.
- [6] Miaolong Yuan, Ming Xie, Xiaoming Yin. Robust cooperative strategy for contour matching using epipolar geometry[C]// Asian Conference on Computer Vision. Australia, 2002.
- [7] 于飞,吕冬梅,颜秉勇,等. 分布式传感器的融合及其应用[J]. 微电子学与计算机, 2006, 23(10): 35 - 37.

作者简介:

张欢欢 女, (1983 -), 硕士研究生. 研究方向为语音图像处理与计算机视觉.

郑江滨 男, (1971 -), 博士, 教授. 研究方向为媒体信息处理、计算机视觉.