

# 视频运动捕获系统中多标记点的三维跟踪算法

晏剑云, 郑江滨, 李秀秀

(西北工业大学计算机学院, 西安 710072)

**摘要:** 针对视频运动参数捕获系统中标记点无法被正确匹配及跟踪的问题, 提出一种多目标三维立体跟踪算法。该算法通过二维和三维卡尔曼滤波预测来标记点的位置, 并利用三维立体匹配方法以及外极限约束关系, 解决了多标记点的匹配和跟踪问题。实验结果表明, 该算法是有效的。

**关键词:** 多目标跟踪; 卡尔曼滤波; 三维立体跟踪

## Multi-marks 3D Tracking Algorithm in Video Motion Capture System

YAN Jian-yun, ZHENG Jiang-bin, LI Xiu-xiu

(School of Computer, Northwestern Polytechnical University, Xi'an 710072)

**【Abstract】** Aiming at the problems that the marks can not be matched and tracked in video motion parameter capture system, a 3D stereo multi-targets tracking algorithm is proposed, which locates the position of the marks by using 2D and 3D Kalman filters prediction. This algorithm also adopts 3D stereo matching method and the constrain relationship of epipolar line to complete matching and tracking of multi-targets. Experimental results demonstrate that the algorithm is effective.

**【Key words】** multi-targets tracking; Kalman filter; 3D stereo tracking

### 1 概述

基于标记点的视频运动捕捉<sup>[1]</sup>因其灵活、高效、易于扩充等特点, 已被广泛应用于影视特效、动画制作、虚拟现实、模拟训练以及生物力学研究等诸多方面。

目前已有许多基于特征的标记点三维跟踪方法。文献[2]提出一种基于色块的跟踪方法, 它把不同色块作为标记点, 以 RGB 值和灰度值为匹配特征, 采用基于子块的匹配方法完成对色块中心点的跟踪。该方法利用有明显特征的标记点, 部分解决了自遮挡问题, 但因为标记点为平面色块, 在三维视觉中易发生形变, 且在不同的摄像头间存在成像色调的差异, 所以会导致跟踪的不准确。文献[3]使用少量彩色标记点和人体轮廓图像, 利用双目立体视觉计算标记点的三维位置, 并以人体约束信息辅助解决标记点的遮挡检测问题。该方法用少量的标记点, 结合人体轮廓信息实现了简单人体运动数据的获取, 但适用的动作范围小, 跟踪的准确性还有待改善。

卡尔曼滤波方法被广泛应用于运动目标跟踪中, 它的特点是跟踪精度高且运算量小, 对一些复杂运动目标能进行连续稳定的跟踪。文献[4]使用扩展卡尔曼滤波器预测目标下一时刻的可能位置, 缩小目标搜索范围, 简化了目标匹配算法, 从而实现目标的快速跟踪。

本文结合卡尔曼滤波和外极限约束两种方法, 提出一种新的目标跟踪方案, 有效解决了在多个标记点间缺少明显差异特征的情况下多标记点三维立体的跟踪匹配问题。

### 2 运动捕获跟踪匹配算法

该算法在双目视觉下, 以标记点在单目序列图像中的二维卡尔曼滤波预测跟踪为基础, 进行标记点的三维卡尔曼滤波跟踪, 同时利用三维立体匹配和双目下的外极限约束条件指导标记点的跟踪匹配。该算法的特点是: (1)利用卡尔曼滤

波器预测标记点在下一帧中的位置, 缩小其搜索区域, 减少由于缺乏特征区别所引起的标记点错误跟踪; (2)利用外极限约束条件解决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, 获得正确的三维候选标记点; (3)利用标记点的三维空间位置相异性, 解决在三维跟踪中出现多个候选标记点时的正确匹配问题, 从而指导二维标记点跟踪。

跟踪算法的实现步骤为: (1)手工建立双目视觉下两组视频图像间的同步对应关系。(2)当人体运动趋于稳定后, 对已建立对应关系的两组视频, 手工标注连续 4 帧的人体标记点, 并初始化每个标记点的二维和三维滤波器的位置、速度以及加速度。(3)利用卡尔曼滤波器公式预测二维图像中对应标记点在第  $i+1$  帧中的位置、速度和加速度, 以及三维空间中的标记点在第  $i+1$  帧的位置、速度和加速度。(4)根据预测的位置和范围, 在二维图像中进行标记点跟踪。(5)获得三维候选标记点中三维跟踪结果, 并修正三维卡尔曼滤波器。(6)根据三维跟踪结果, 与其对应的二维候选标记点即为相应的二维跟踪结果, 并修正二维卡尔曼滤波器。

### 3 卡尔曼滤波预测跟踪

#### 3.1 卡尔曼滤波器的组成

在单目情况下对每个标记点在图像中的二维位置建立其二维卡尔曼滤波器。同时, 在双目情况下对标记点的三维空

**基金项目:** 国家“863”计划基金资助项目(2006AA01Z324); 陕西省自然科学基金资助项目(2006F02); 西北工业大学基础研究基金资助项目(2006); 西北工业大学研究生创业种子基金资助项目(Z200752)

**作者简介:** 晏剑云(1981-), 女, 硕士研究生, 主研方向: 图像处理 and 计算机视觉; 郑江滨, 副教授、博士; 李秀秀, 博士研究生

**收稿日期:** 2008-04-30 **E-mail:** jianyun07.yan@Gmail.com

间位置建立其三维卡尔曼滤波器。在下一时刻分别对标记点在图像中的二维位置和空间中的三维位置进行预测，并在预测范围内搜索最优匹配点。

设卡尔曼滤波器由 10 个参数组成：

$$O_n = (p_k, \hat{p}_k, \Delta p_k, v_k, \hat{v}_k, \Delta v_k, a_k, \hat{a}_k, \Delta a_k) \quad (1)$$

其中， $p_k$  表示标记点  $O_n$  在第  $k$  时刻测量到的质心坐标位置； $\hat{p}_k$  表示标记点  $O_n$  在第  $k$  时刻预测到的质心坐标位置； $\Delta p_k$  表示标记点  $O_n$  在第  $k$  时刻的目标搜索范围； $v_k$  表示标记点  $O_n$  在第  $k$  时刻测量到的速度； $\hat{v}_k$  表示标记点  $O_n$  在第  $k$  时刻预测到的速度； $\Delta v_k$  表示标记点  $O_n$  在第  $k$  时刻预测速度的误差； $a_k$  表示标记点  $O_n$  在第  $k$  时刻测量到的加速度； $\hat{a}_k$  表示标记点  $O_n$  在第  $k$  时刻预测到的加速度； $\Delta a_k$  表示标记点  $O_n$  在第  $k$  时刻预测加速度的误差。

### 3.2 卡尔曼滤波器预测

利用以下二维卡尔曼滤波器公式能够预测下一帧中标记点可能出现在图像中的位置和范围。

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

$$\Delta p_{k+1} = \Delta p_k + \Delta v_k \cdot \Delta t + \frac{1}{2} \Delta a_k \cdot \Delta t^2 \quad (3)$$

其中， $\hat{p}_{k+1}$  表示预测位置的坐标； $\Delta p_{k+1}$  表示目标匹配的搜索范围。三维卡尔曼滤波器亦可用式(2)和式(3)预测下一帧中标记点可能出现的空间位置和范围。

### 3.3 目标搜索匹配

根据灰度和面积特征，在双目下的 2 幅图像帧中的搜索区域内提取候选标记点，并进行二维目标匹配：若仅有一个候选标记点，则该标记点位置  $p_{k+1}$  即是所跟踪的标记点；若出现多个候选标记点时，则需结合极限约束关系和三维卡尔曼预测结果进行三维立体匹配；若没有候选标记点，则把预测位置作为当前时刻标记点的位置。

在三维立体匹配中的策略是：对双目下获得的两组候选标记点进行对应，并计算对应关系下的三维坐标，同时由二维候选标记点产生三维候选标记点。根据与三维预测中心的欧式距离获得最佳候选中心。

令标记点  $P$  在第  $k+1$  帧时三维卡尔曼预测位置为  $\hat{q}_{k+1}$ ，预测范围是  $\Delta q_{k+1}$ 。在当前时刻，摄像机  $A$  获取的图像候选标记点位置有  $\{p_{a1}^*, p_{a2}^*\}$ ，摄像机  $B$  获取的图像候选标记点位置有  $\{p_{b1}^*, p_{b2}^*\}$ 。

此时，由两组候选标记点产生对应标记点有 2 种可能的组合： $\{(p_{a1}^*, p_{b1}^*), (p_{a2}^*, p_{b2}^*)\}, \{(p_{a1}^*, p_{b2}^*), (p_{a2}^*, p_{b1}^*)\}$ 。

利用外极限约束关系剔除错误组合  $(p_{a1}^*, p_{b1}^*), (p_{a2}^*, p_{b2}^*)$ ，再根据三维重建函数  $F$  计算出两组三维候选标记点坐标  $\{q_1^*, q_2^*\}$ ：

$$q_k^* = F(R_a, R_b, D_a, D_b, (p_{ai}^*, p_{bj}^*)), \quad i, j = 1, 2 \quad k = 1, 2, 3, 4 \quad (4)$$

其中， $R_a, R_b$  分别是相机  $A$  和相机  $B$  的旋转矩阵； $D_a, D_b$  分别是相机  $A$  和相机  $B$  的畸变矩阵。

然后计算以上三维坐标是否落在三维卡尔曼预测的空间中，并以标记点与预测的三维位置的欧式距离为匹配准则：

$$J_i = [\sum_{i=1}^3 (q_i^* - \hat{q}_{k+1})^2]^{1/2}, \quad i = 1, 2, 3, 4 \quad (5)$$

$$p_k^* = \arg \min_{q_i^*} (J_1, J_2, J_3, \dots)$$

距离  $J_i$  最小的候选标记点即为被跟踪的标记点在该时

刻的三维测量位置。图 1 为目标匹配的示意图， $q_1^*$  为最佳匹配点，其对应的两个二维点  $(p_{a1}^*, p_{b2}^*)$  分别为相机  $A$  和相机  $B$  在该时刻的二维测量中心。

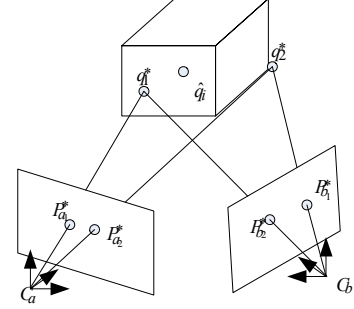


图 1 目标匹配

### 3.4 目标更新

根据匹配得到的最佳三维位置  $q_k^*$  以及与其对应的两个二维位置  $(p_{a1}^*, p_{b1}^*)$ ，更新目标的二维和三维速度和加速度，并用 IIR 滤波器对卡尔曼滤波器的相关参数进行修正：

$$\hat{v}_{k+1} = \alpha \cdot v_{k+1} + (1 - \alpha) \cdot \hat{v}_k; \quad \Delta v_{k+1} = \alpha \cdot |v_{k+1} - v_{k+1}| + (1 - \alpha) \cdot \Delta v_k \quad (6)$$

$$\hat{a}_{k+1} = \beta \cdot a_{k+1} + (1 - \beta) \cdot \hat{a}_k; \quad \Delta a_{k+1} = \beta \cdot |a_{k+1} - a_{k+1}| + (1 - \beta) \cdot \Delta a_k$$

其中， $0 < \alpha, \beta < 1$ 。

### 4 外极线约束指导跟踪

根据外极线理论，点  $P$  在两个摄像机下所成的像为  $p$  和  $p'$ ，点  $P$  与两个摄像机的光学中心点  $O$  和  $O'$  形成两条相交光线  $OP$  和  $O'P$ ，这两条相交光线形成的平面即为外极平面。

第 2 个摄像机的光学中心在第 1 个摄像机下的投影即为第 1 个摄像机的外极点，令双摄像机下的外极点分别为  $e$  和  $e'$ ，外极点与其对应像点的连线称为外极线，即  $l_{ep}$  和  $l_{e'p'}$ 。

外极点与外极线满足关系：

$$l_{ep} = Fp', \quad l_{e'p'} = Fp$$

两个对应的外极点满足：

$$pF^T p' = 0$$

其中， $F$  为基本矩阵。

在相机  $A$  获取的图像  $I_a$  中，候选的标记点位置有  $\{p_{a1}^*, p_{a2}^*\}$ ，候选标记点在相机  $B$  获取的图像  $I_b$  中的外极线分别是  $l_{b1}, l_{b2}$ ，图像  $I_b$  中的候选标记点位置  $\{p_{b1}^*, p_{b2}^*\}$  与极线  $l_{b1}, l_{b2}$  的距离为  $dis_k$ ，且满足：

$$p_{bi}^* (u_i, v_i) = \arg \min_{p_{bi}^* (u_i, v_i)} (dis_1, dis_2, \dots) \quad (7)$$

与极线距离最近的标记点即为与图像  $I_a$  中候选的标记点对应的点。

根据以上距离准则可获得结果：图像  $I_a$  中候选的标记点  $p_{a1}^*$  对应图像  $I_b$  中候选的标记点  $p_{b1}^*$ ；同理， $p_{a2}^*$  对应  $p_{b2}^*$ 。从而剔除了候选标记点组合  $(p_{a1}^*, p_{b1}^*), (p_{a2}^*, p_{b2}^*)$ 。

### 5 实验结果

本文算法由 C/C++ 编程实现，在实验中采用双相机对着 10 个白色标记点的人体作手臂的平举运动进行拍摄，并取  $\alpha = 0.7, \beta = 0.1$ 。图像分辨率为  $752 \times 480$ ，帧率为  $30 \text{ f/s}$ 。

图 2 从左到右依次展示了第 26 帧、第 35 帧、第 39 帧以及第 40 帧时的跟踪结果。图中白色矩形框显示了使用本文算法的跟踪结果。图 3 给出了在第 40 帧时，右时标记点的二维卡尔曼预测范围。（下转第 210 页）