

基于视频的运动捕获

罗忠祥 庄越挺 潘云鹤 刘 丰

(浙江大学计算机科学与工程系, 浙大-微软视觉感知联合实验室, 杭州 310027)

摘 要 现有的运动捕获方法大都存在运动捕获设备昂贵、演员运动受限等缺点, 为此, 提出了一种利用视觉技术从视频中提取人体运动的方法, 并对其中的特征跟踪和三维运动序列恢复等关键技术进行了深入研究. 基于人体模型的特征跟踪算法利用卡尔曼滤波和极线方程, 能精确地跟踪比较大的人体运动; 采用不共面的非线性定标模型和考虑运动不确定性的三维重建方法, 能恢复逼真的三维人体骨架模型. 实验结果验证了基于视频的运动捕获方法的可行性和有效性.

关键词 视频动画 运动捕获 特征跟踪 三维重建

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2002)08-0752-07

Video Based Motion Capture

LUO Zhong-xiang, ZHUANG Yue-ting, PAN Yun-he, LIU Feng

(Computer Science and Engineering Department, Zhejiang University,

Microsoft visual Perception Laboratory of Zhejiang University, Hangzhou 310027)

Abstract Motion capture has been one of the most promising technologies in character animation in recent years. However, most currently available motion capture approaches suffer from costly equipment and motion restriction caused by markers. To overcome these problems, a novel approach to extract 3D motion from video using vision technology is presented. The key issues, such as feature tracking and 3D reconstruction, are deeply studied. A model-based feature tracking algorithm, which utilizes Kalman filter to predict coordinates of image features and epipolar line equation to aid tracking, is presented to track human motion with great variety. A non-coplanar nonlinear calibration model and a reconstruction approach taken uncertainty into consideration are applied to restore 3D human skeleton model. At last the experimental results and analysis of our VBHA (Video Based Human Animation) system demonstrate the feasibility and effectiveness of our approaches.

Keywords Video based animation, Motion capture, Feature tracking, 3D reconstruction

0 前 言

近年来兴起的运动捕获成为人体动画中最有前途的技术之一. 根据捕获的运动来制作动画具有高逼真度、高效率的特点, 运动捕获涉及的关键技术主要有摄像机定标、特征跟踪、特征对应和三维重建等. 代表性的研究成果包括: 基于 CCD 摄像机的高速三维运动信息获取系统^[1], 其先用标定物来求摄像机的内参数, 然后利用定标结果重建图象中的三

维运动信息; 利用多个摄像机采集图象信息, 对采集的图象信息采用基于色块的特征点跟踪方法来跟踪二维特征点, 然后利用计算机视觉技术重构二维特征点的三维运动信息, 用其开发的实时多相机运动捕捉系统中, 系统将获得的运动信息输入图象绘制模块, 即可生成满足要求的动画^[2]; 利用视觉技术来进行运动捕捉的方法, 其不要求标记附加在人体上, 就能够恢复复杂视频序列中多自由度人体骨架模型的运动结构^[3]; 基于知识的捕捉和描述视频中人行走运动的体系框架, 其利用的知识主要包括由关节

基金项目: 国家自然科学基金(69803009); 教育部优秀青年教师基金以及高等学校骨干教师资助计划项目

收稿日期: 2001-07-16; **改回日期:** 2002-01-11

组成的人体模型以及行走的先验知识,视频经过数据预处理、识别、跟踪和三维恢复后,产生逼真的人体行走动画^[4]。

目前捕获运动的方法主要包括基于机械、电磁、声学、光学和图象等设备的运动获取方法,但这些方法都有两个共同的缺点:需要昂贵的运动捕获设备;附加在身上的传感器限制了演员的自由表演^[5,6]。考虑到这些不足,提出了一种基于视频的运动捕获方法,它利用计算机视觉的基本原理,从视频中直接提取三维人体运动序列。这种方法不需要在人体关节上附加任何传感器,保证了人体不受限的自由运动,而且造价低,效率高,只需在视频首帧标注人体各个关节点,就能得到三维人体运动序列。

1 特征跟踪

在运动捕获中,将人体看成由关节点连接的刚体的集合,将人体运动简化为人体骨架的运动。人体骨架模型如图1所示,它由16个关节特征点组成(在下文中,关节特征点又称为特征点)。运动捕获首先需要跟踪各帧图象中人体骨架模型的各个特征点,然后恢复特征点的三维信息,从而得到人体骨架模型的运动序列。

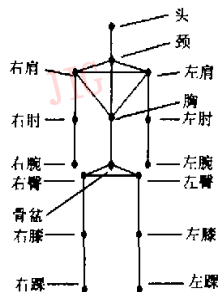


图1 人体骨架模型

特征跟踪是将视频序列中各帧的各个特征点匹配起来,从而产生各个特征点的连续运动轨迹,即图象运动模型。没有任何领域知识约束的图象运动模型估计是一个欠约束问题,因此所有已有的运动模型估计算法都是通过对假设的运动模型增加约束来解决欠约束问题的。运动模型估计算法主要有:基于参数化的跟踪方法,其通过求解过约束的线性方程组来估计运动模型参数;基于色块的跟踪方法;基于运动模板的跟踪方法;基于三维人体的非刚体模型的跟踪方法^[7~9]。在运动捕获中,根据人体骨架模型

设计了一套紧身衣,在每个关节点处附上不同颜色的色块作为标记,特征跟踪的目标就是跟踪色块的中心点。

1.1 基本思想

人体运动跟踪一直是一个具有挑战性的课题,其主要原因是:有高精度要求;在运动过程中,经常存在四肢的自遮挡;图象中四肢的边界与其他部分的视觉特征缺乏有效的对比度。在运动捕获中,利用图象单个特征(颜色、纹理和形状等)的跟踪算法不能解决运动变化大、存在自遮挡以及精度要求高的特征跟踪问题,只有通过引入其他约束才能提高特征跟踪的效率和精度。经过分析发现:(1)视频序列中,前后帧对应特征点的运动是相关的,利用卡尔曼滤波预测可以减少图象搜索范围,提高特征跟踪的速度和效率;(2)在双相机的运动捕获系统中,根据双相机定标结果求得的极线约束方程,可以指导特征跟踪,从而尽可能减少跟踪误差。

根据以上分析,提出了一种基于人体模型的特征跟踪算法,其基本思想是:首先利用卡尔曼滤波预测特征点的位置,然后以RGB和灰度为匹配特征,采用基于子块的模式匹配方法搜索种子点,根据种子点聚类色块,将色块的中心点作为特征跟踪结果,在跟踪到一帧图象的特征点后,利用极线约束方程指导同一时刻另一帧图象的特征跟踪。

1.2 特征跟踪算法

$C_1(i)$ 表示第1个摄像机视频的第 i 帧, $C_2(i)$ 表示第2个摄像机视频的第 i 帧。算法具体实现如下:

(1) 首先找到同一时刻从不同角度拍摄的两段视频中的同步点标志,建立两段视频间的同步关系。

(2) 对已建立同步关系的两段视频,首帧采用手工标注图象中人体骨架模型特征点的位置,并给各个特征点的速度赋初值。

(3) 根据第 i 帧的跟踪结果,采用卡尔曼滤波预测第 $i+1$ 帧中对应特征点的位置和速度。

(4) 根据预测的结果,跟踪第 $C_1(i+1)$ 帧中各个特征点,方法如下:

① 以预测到的特征点为中心点,按顺时针方向搜索周围的8邻点,以RGB值和灰度值为匹配特征,采用基于子块的模式匹配方法搜索与第 $C_1(i)$ 帧中对应特征点匹配的点;

② 以搜索到的近似匹配点作为种子点,以小于某个阈值(根据子块的平均灰度值设定)作为聚类标准,在一个包围盒内聚类色块;

③ 计算色块的中心,该中心点为第 $C_1(i+1)$ 帧的跟踪结果;

④ 根据跟踪结果修改状态向量。

(5) 利用极线约束方程指导第 $C_2(i+1)$ 帧中对应特征点的跟踪,方法如下:

① 利用第4步中的方法跟踪第 $C_2(i+1)$ 帧中的特征点,跟踪结果作为候选特征点;

② 计算候选特征点与极线的距离,如果距离小于阈值,则候选特征点作为跟踪结果,继续第7步;

③ 根据极线方程决定的搜索路径,搜索与种子点匹配的近似点;

④ 计算匹配点与极线的距离,以与极线距离最小的点作为跟踪结果;

⑤ 如果不能跟踪到满足要求的点(如存在自遮挡),则继续第6步;

⑥ 修改状态向量,继续第7步。

(6) 在存在自遮挡时,采用预测机制预测点的大概位置,将预测结果作为特征跟踪结果。

(7) 如果继续跟踪下一帧,则继续第3步,否则结束。

1.3 基于子块的特征匹配

从视频序列中跟踪特征点,就是要从每帧图象中,找到对应的特征点位置。在卡尔曼滤波预测的搜索区间内,搜索聚类的种子点时,因为人体运动大而引起视频序列中对应特征点的RGB值和灰度值变化比较大,所以不采用当前帧与第1帧的特征点,而采用与前一帧的对应特征点进行匹配。首先计算搜索区间内任意点 p 决定的子块,与上一帧的对应特征点决定的子块的灰度和RGB值之差的平方和

$$\epsilon(d) = \sum_{x=u_x-1}^{u_x+1} \sum_{y=u_y-1}^{u_y+1} (I(x,y,t) - I(x+d_x, y+d_y, t+\tau))^2$$

$$\epsilon_R = \sum_{x=u_x-1}^{u_x+1} \sum_{y=u_y-1}^{u_y+1} (R(x,y,t) - R(x+d_x, y+d_y, t+\tau))^2$$

$$\epsilon_G = \sum_{x=u_x-1}^{u_x+1} \sum_{y=u_y-1}^{u_y+1} (G(x,y,t) - G(x+d_x, y+d_y, t+\tau))^2$$

$$\epsilon_B = \sum_{x=u_x-1}^{u_x+1} \sum_{y=u_y-1}^{u_y+1} (B(x,y,t) - B(x+d_x, y+d_y, t+\tau))^2$$

$I(x,y,t)$ 表示 t 时刻图象点 (x,y) 的灰度值, $R(x,y,t)$, $G(x,y,t)$ 和 $B(x,y,t)$ 分别表示 t 时刻图象点 (x,y) 的RGB值, (u_x, u_y) 表示匹配子块的中心点。

视频序列中,对应特征点的图象特征是相同的或者最接近的,理想情况下,特征跟踪就是找到对应的

d_x 和 d_y 使 $\epsilon(d)$ 、 ϵ_R 、 ϵ_G 和 ϵ_B 都为零。但是由于特征点的运动以及其他噪声的影响,实际跟踪中,只要灰度值和RGB值比较接近就能满足要求,因此,特征跟踪的目标就是找到同时满足以下两式的 d_{x_0} 和 d_{y_0} 。

$$\min(\epsilon_R \cap \epsilon_G \cap \epsilon_B) \quad (1)$$

$$\epsilon(d) \leq D \quad \text{其中, } D \text{ 为阈值} \quad (2)$$

式(1)表示不存在使RGB值同时小于点 $(x+d_{x_0}, y+d_{y_0})$ 的RGB值的 d_x 和 d_y 。之所以不采用RGB之差的平方和的原因是:在许多情况下,某点的RGB中的某个值(R或G或B)与上一帧对应特征点的RGB值相差很大,而RGB值之差的平方和却比较小,这样得到的匹配点是噪音引起的“假点”。式(2)表示匹配点的灰度值在由色块灰度平均值确定的某个范围内,它能够保证匹配点在色块区域内。只有同时满足式(1)和式(2)时,该点才是聚类的种子点。

在聚类色块时,只要该点与色块的灰度平均值接近,就认为该点是属于色块区域中的一点,采用的聚类标准是式(2)。

1.4 预测机制

在运动捕获中,人体运动比较大,从而设定的搜索区间也就比较大,因此,跟踪的速度和效率很难保证。为了克服这个缺陷,采用图象坐标系下的卡尔曼滤波预测来指导特征跟踪,能够取得满意的效果。可以将人体运动序列看作为一个动态系统,前后两帧中对应关节点的运动可以看成是匀速的。在采用卡尔曼滤波进行预测时,系统的状态矢量为 $[u(k), v(k), \dot{u}(k), \dot{v}(k)]^T$, k 表示时刻, $u(k)$ 和 $v(k)$ 分别表示图象坐标系中 u 和 v 方向的坐标,而 $\dot{u}(k)$ 和 $\dot{v}(k)$ 分别表示特征点在 u 和 v 方向上的速度分量。因此,动态系统的状态方程可以描述为

$$X(k+1) = FX(k) + \eta(k) \quad (3)$$

而系统的观测方程表示理想跟踪结果与实际跟踪结果之间的关系,因此动态系统的观测方程可以描述为

$$Z(k) = HX(k) + W(k) \quad (4)$$

其中

$$F = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Δt 为帧与帧之间的时间间隔, $\eta(k)$ 为零均值的白噪声,协方差矩阵为 Q , $W(k)$ 为零均值的白噪声,协方差矩阵为 R 。动态系统状态的预测和修改可以

参考文献[2].

对不存在自遮挡的特征点跟踪,采用卡尔曼滤波预测可以提高跟踪速度.但是当出现自遮挡或者其他不完全运动轨迹时,需要采用其他的预测机制,主要包括:(1)当存在自遮挡时,即人体骨架模型中特征点之间的距离小于一定阈值时,预测在一定阈值范围内会再次出现特征点;(2)当跟踪到的特征点的灰度值和RGB值与理想值相差比较大时,在特征点可能出现的范围内搜索特征点,如果搜索到的特征点灰度值和RGB值更接近理想值,则将该特征点作为跟踪结果.

2 三维人体运动序列

特征跟踪结果是运动信息不完全的二维人体运动序列,不能进行运动编辑和制作动画,因此需要恢复三维的运动序列.在此过程中,首先需要通过摄像机,定标确定空间特征点与对应的图象特征点之间的关系,然后利用定标结果和跟踪结果,采用非线性优化方法求出三维人体运动序列.

2.1 非共面非线性摄像机定标

在 Tsai 线性定标模型^[10]的基础上,考虑采用图象非线性畸变的不共面非线性摄像机定标模型,图象点坐标 (u', v') 与空间点 (X_w, Y_w, Z_w) 之间的关系可表达为

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = MX \quad (5)$$

$$u = u'(1 + k_1 r^2) \quad v = v'(1 + k_1 r^2) \quad (6)$$

$$r^2 = x^2 + y^2$$

$$x = d_x(u' - u_0)/s_x \quad y = d_y(v' - v_0) \quad (7)$$

式中, (u, v) 为考虑图象存在非线性径向畸变的情况下,经过非线性变形处理得到的图象坐标. d_x 和 d_y 分别为摄像机的 x 方向和 y 方向的单位像素物理尺寸, f 为摄像机的焦距. d_x 、 d_y 和 f 是摄像机的内参数. Z_c 为空间点在摄像机坐标系中的 z 方向长度, (X_w, Y_w, Z_w) 为空间点在世界坐标系中的位置.

$R = (R_x, R_y, R_z)$ 是旋转矩阵,为 3×3 正交单位矩阵, $t = (T_x, T_y, T_z)$ 为三维平移向量, $0^T = (0, 0, 0)$, M 为变换矩阵. k_1 是考虑存在非线性畸变的径向畸变系数, s_x 为图象坐标系中 x 方向的不确定性比例因子, (u', v') 为实际观察到的图象坐标, (x, y) 为考虑 x 方向不确定性的图象物理坐标, (u_0, v_0) 为图象物理坐标系的原点在图象坐标系中的位置.

基于双相机的视频动画(VBHA)系统中采用的定标物如图2所示,是包含了12个点(空间点和图象点)的立体框架,这些点分布在不同的平面上.图象中特征点的跟踪采用以灰度值为基准的基于摄像机模型的自动跟踪,空间点的坐标采用精确的自动定位仪进行测量.

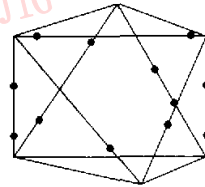


图2 定标物(球为定标点)

根据空间点 (X_w, Y_w, Z_w) 和对应的图象点 (u', v') 求11个未知数(包括 $f, k_1, u_0, v_0, s_x, R_x, R_y, R_z, T_x, T_y, T_z$),从式(5)知道,每个点对应2个方程,因此至少需要6个点,一般采用大于7个定标点.

在求解非线性定标模型的过程中,先利用该非线性模型对应的线性定标模型求解初始值,然后以求出的结果作为初始值,采用非线性优化方法求出精确的结果.具体实现方法为:

(1) 采用 Tsai 提出的线性摄像机定标模型求解方法求出初值^[8];

(2) 采用优化算法 Levenberg-Marquart 对 f, k_1, T_x 进行优化;

(3) 将第2步的结果作为初始值,利用非线性优化方法,即 Levenberg-Marquart 优化算法对除图象中心坐标参数 (u_0, v_0) 外的所有参数进行优化;

(4) 利用第3步的结果作为初始值,对所有参数进行非线性优化,求出精确的定标参数.在定标过程中,有两个问题需要注意,一是在式(5)中, (u, v) 为考虑图象存在非线性径向畸变的情况下经过非线性变形处理得到的图象坐标, (u', v') 才是实际跟踪到的图象坐标,因此进行三维重建时,需要先利用式(6)和式(7)求出 (u, v) ;二是当世界坐标系的原点位于摄

像机的视区内时,实验结果表明,采用非线性的不共面摄像机定标模型定标的误差比较大,此时可以通过将空间坐标系的原点平移到视区外来解决。

2.2 三维运动序列重建

三维重建的目的是恢复跟踪到的图象特征点的三维运动信息。传统的方法一般是利用定标结果直接重建三维运动信息,其基本思路是根据投影理论利用空间点与图象点之间的变换关系恢复图象点的三维坐标,可以表达为

$$Z_{c_1} u_1 = M_1 X_P \quad (8)$$

$$Z_{c_2} u_2 = M_2 X_P \quad (9)$$

Z_{c_1} 和 Z_{c_2} 分别为空间点在两个摄像机坐标系中 z 方向的坐标, X_P 为空间点 P 在世界坐标系下的齐次坐标; u_1 和 u_2 分别是点 P 投影到图象平面上的点 P_1 和 P_2 经过非线性畸变处理的齐次坐标, M_1 和 M_2 分别为两个摄像机的投影变换矩阵。根据式(8)和式(9),采用最小二乘法可以求出三维坐标的解析解。

对这种三维重建算法进行了实验,结果表明,这种方法能够满足一般的应用要求,但是当将恢复的三维特征点再根据投影方程投影到图象平面时,误差就会很大,这种误差对于图象的跟踪来说,是不能接受的,因此考虑采用存在运动不确定性的三维重建方法。这种方法的优化准则是,使三维重建的特征点再投影到图象平面上的点和实际跟踪的图象点的距离最小作为约束条件,即,使下式最小:

$$d = \sum_{i=1}^2 \left(\left(u_i - \frac{M_{11}x + M_{12}y + M_{13}z + M_{14}}{M_{31}x + M_{32}y + M_{33}z + M_{34}} \right)^2 + \left(v_i - \frac{M_{21}x + M_{22}y + M_{23}z + M_{24}}{M_{31}x + M_{32}y + M_{33}z + M_{34}} \right)^2 \right) \quad (10)$$

式中, (x, y, z) 为空间点坐标,分别对 x, y, z 求偏导数,要使 d 最小,则要求偏导数为零,即

$$\begin{cases} f_1(x, y, z) = \partial d / \partial x = 0 \\ f_2(x, y, z) = \partial d / \partial y = 0 \\ f_3(x, y, z) = \partial d / \partial z = 0 \end{cases} \quad (11)$$

式(11)所得的方程组是一个非线性方程组,可以采用牛顿迭代法求解空间点 (x, y, z) 。

3 三维运动数据处理

采用上述方法恢复的人体骨架模型比较精确,但由于特征跟踪结果存在误差,其恢复的三维人体运动序列存在抖动。为了减少特征跟踪误差引起的三维人体运动序列的抖动,需对恢复的人体骨架模

型中的各个特征点序列进行平滑。

在特征点序列中,噪声是由色块大小的变动而引起的,因此它是与特征点位置无关的加性噪声,这里带噪声的特征点位置 g 可描述为理想无噪声的特征点位置 f 与噪声 n 之和

$$g = f + n$$

采用整个时间域的平滑来去除噪声。特征点序列中,尖顶等突变细节比较少,为了保持特征点序列的运动特性,采用时间域加权平均的非递归数字滤波器来滤除噪声。滤波器的输入与输出可以表达为

$$y(t) = x(t) * h(t) = \sum_{r=-n}^n h(r)x(t-r)$$

$$\sum_{r=-n}^n h(r) = 1$$

式中的 $x(t)$ 表示 t 时刻输入的三维数据, $y(t)$ 为输出结果, $h(r)$ 是数字滤波器的单位样值响应,这里采用的单位样值响应函数为偶函数,且为有限长。

实验结果表明,经过滤波处理的人体运动序列,抖动明显减少,运动序列比较平滑,同时运动信息又没有失真。图3是关节点 R_{wrist} 在 x, y, z 方向,滤波前、滤波后的运动轨迹。从图中可以看出,原有的尖顶细节已经平滑了,而整个运动轨迹又没有发生改变。

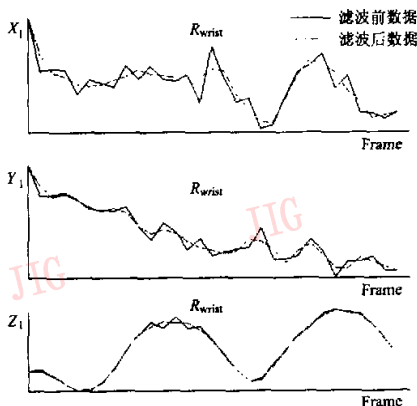


图3 关节点滤波前后的运动轨迹

4 实验结果

在以上研究与分析的基础上,研制了一个基于双相机的视频动画原型系统(VBHA)。系统采用 Visual C++ 实现,运行在 Windows X 平台上。图4是对实验室拍摄的一段视频的跟踪结果及其恢复后

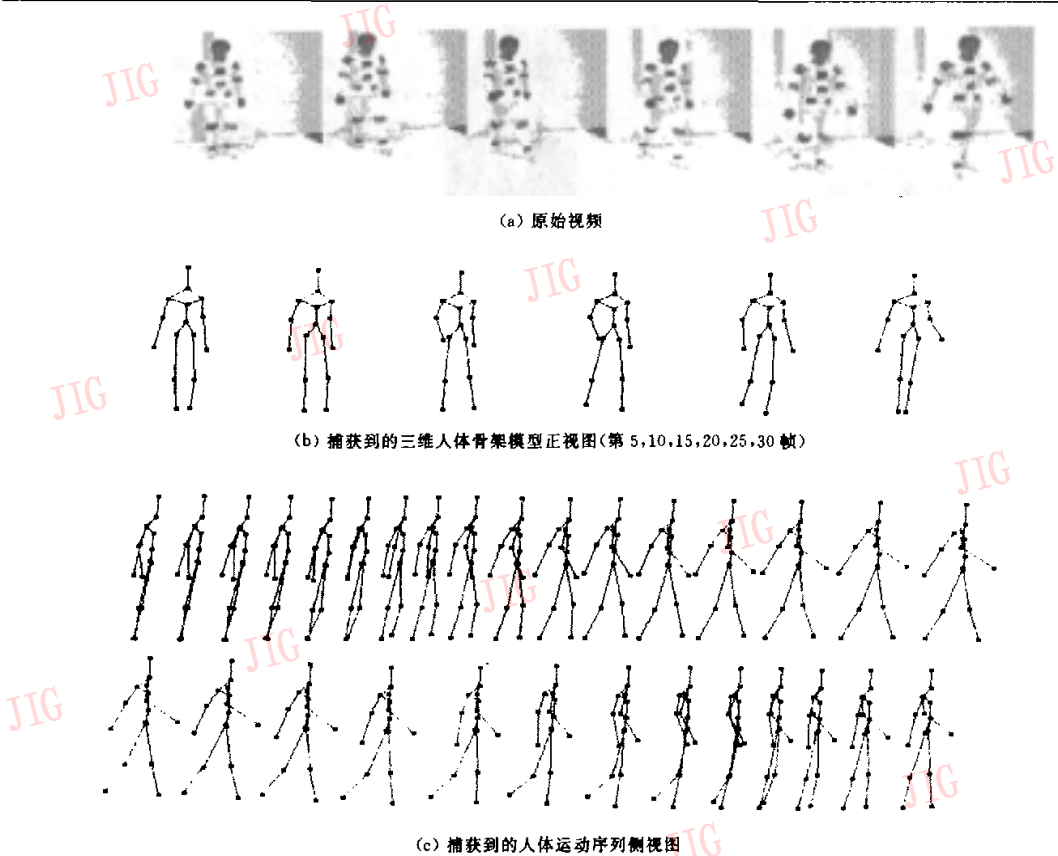


图 4 Walking 运动捕获结果

的三维人体骨架模型.在图 4 中,(a)是在实验室拍摄的演员表演 Walking 动作的视频关键帧,(b)是利用定标技术由上述跟踪算法跟踪的结果恢复的三维人体骨架模型正视图,(c)是恢复的三维人体运动序列侧视图.从图中可以看出,存在自遮挡的特征点都能恢复精确的三维结果.

因为跟踪结果的精确度和效率直接影响着将跟踪结果重定向到动画角色上动画作品的精确度,对人体行走运动序列中一段 60 帧的视频进行跟踪实验,并对基于光流场的跟踪、基于属性的跟踪和本文提出的跟踪算法进行了对比实验,跟踪结果统计如表 1.

表 1 跟踪效率对比

特征	算法		
	基于光流场的跟踪	基于属性的跟踪	本文算法
孤立特征 (正确率%)	40	90	90
自遮挡特征 (正确率%)	0	0	70
时间复杂度(s)	10	25	13

从表中可以看出,由于人体运动比较大,基于光流场的跟踪算法效率很低,而且不能解决自遮挡问题;基于属性的跟踪,由于是先提取色块的属性信息,然后进行特征对应,因而能较好地跟踪孤立特征,但是,对于自遮挡特征,该算法也无能为力,而且该算法的计算复杂度非常高,本文提出的算法,由于采用了预测机制和极线方程,故跟踪效率明显提高,而且时间复杂度比基于属性的跟踪算法低.

5 结论与展望

详细介绍了基于双相机的运动捕获系统中采用的关键技术.采用卡尔曼滤波预测和基于子块模式匹配指导的特征跟踪算法能够跟踪运动变化大、存在自遮挡的特征点.采用非线性的不共面摄像机定标模型,并采用考虑了存在运动不确定性的三维重建方法恢复和优化人体三维运动序列,可以大量减少人体运动序列中,对应特征点的抖动.系统运行结

果表明,捕获到的运动比较精确,可以作为生成动画的原始素材.在下一阶段的工作中,除了完善现有的运动捕获方法外,还需要对捕获的人体运动序列进行编辑处理,生成满足动画师要求的动画.

参 考 文 献

- 1 Sabel J C. Optical 3D motion measurement [A]. In: IEEE Instrumentation and Measurement Technology Conference[C]. Brussels; IEEE Computer Society Press, 1996:367~370.
- 2 Yonemoto S, Tsuruta N, Taniguchi R. Tracking of 3D multi-part objects using multiple viewpoint time-varying sequences [A]. In: Proceedings of International Conference on Pattern Recognition [C]. Santa Barbara, IEEE Computer Society Press, 1998:490~494.
- 3 Bregler C, Malik J. Tracking people with twists and exponential maps[A]. In: Proceeding of IEEE Conference Computer Vision Pattern Recognition[C], Santa Barbara; IEEE Computer Society Press, 1998:8~15.
- 4 Cheng J C, Moura J M F. Tracking human walking in dynamic scenes[A]. In: Proceedings of ICIP'97[C]. Washington, DC: Computer Society Press, 1997:137~140.
- 5 Badler N I. Animation 2000++ [J]. IEEE Computer Graphics and Applications. February 2000:28~29.
- 6 Liu Xiao-ming, Zhuang Yue-ting, Pan Yun-he. Video based human animation technique[A]. In: Proceeding of the 7th ACM International Multimedia Conference[C]. Orlando; ACM Press, 1999:353~362.
- 7 Wren C, Azarbayejani A, Darrell T *et al.* Pfister: Real-time tracking of the human body[J]. IEEE Transaction on pattern analysis and machine intelligence, 1997,19(7):780~785.
- 8 Shi Jian-bo, Tomasi Carlo. Good features to track [A]. In: Proc. of IEEE Computer. Soc. Conf. Computer Vision and Pattern Recognition [C]. Seattle; IEEE Computer Society Press, 1994:593~600.
- 9 Fusiello A, Trucco E, Tommasini T *et al.* Improving feature tracking with robust statistics [J]. Pattern Analysis & Applications. 1999,2:312~320.
- 10 Tsai R Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses [J]. IEEE Journal of Robotics and Animation. 1987,3(4):323~344.



罗忠祥 1976年生,浙江大学计算机科学与工程系博士研究生.主要研究方向为网络多媒体技术和视频运动捕获.



庄越挺 1965年生,浙江大学计算机科学与工程系教授,博士生导师.主要研究方向为网络多媒体技术、智能CAD.



清云鹤 1946年生,浙江大学计算机科学与工程系教授,博士生导师,中国工程院院士.主要研究方向为计算机美术、形象思维、智能CAD系统、GIS、计算机动画和多媒体技术等.



刘丰 1979年生,硕士研究生.主要研究方向为网络多媒体技术.