

基于视频的角色运动提取与运动合成

苏 洲,王 胜

(河海大学机电工程学院,江苏 常州 213022)

摘要 在基于大量现存的影视片段的基础上,采用一种基于特征的跟踪算法——KL(Kanade-Lucas)算法,对视频中角色的特征关节实现跟踪,得到运动信息不完全的二维角色运动序列,然后进行三维重建生成满足动画要求的三维角色运动序列,并将三维角色运动序列重定向到三维的虚拟角色中,从而得到逼真的角色动画.实际运行表明,生成的角色动画与视频中的角色动作具有良好的吻合性.

关键词 视频 特征跟踪 三维重建 运动合成

中图分类号 TP391.41

文献标识码 A

文章编号 1000-1980(2008)03-0410-04

目前,计算机动画市场潜力非常大,远没有满足广大观众的需求.近年来兴起的运动捕获成为人体动画最有前途的技术之一,代表性的研究成果包括:基于 CCD 摄像机的高度三维运动信息获取系统^[1]、利用视觉技术来进行捕捉的方法^[2]、基于知识的捕捉和描述视频中人行走运动的体系框架^[3]、基于双摄像机的视频动画系统(VBHA)(该系统根据人体骨架模型设计了一套紧身衣,在每个关节处附上不同颜色的色块作为标记,运动捕获的目标就是捕获色块中心点的三维运动序列)^[4-5]等.但这些研究都有 2 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^[6-7].考虑到这些不足,本文提出了利用大量现存的影视片段作为动画创作的素材源,采用特征跟踪算法从视频中提取角色的运动,然后经过三维重建重定向到三维的虚拟角色中,产生高度‘拟人化’的效果.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,而且计算开销低,效率高.

1 基于光流法的特征点跟踪方法

本文采用光流法中的基于特征的跟踪算法——KL(Kanade-Lucas)算法^[8],该算法是由 Lucas 和 Kanade 于 1981 年首先提出的,后来被 Tomasi 和 Kanade 不断完善,目前已成为一种比较经典的特征跟踪算法被广泛使用.该算法首先通过标注的方式来选定跟踪的区域,然后通过特征计算和特征匹配在后续帧中完成自动跟踪.

在特征点的选取上,选取角色关节点作为跟踪目标,为此,通过人体骨架的长短比例关系来建立二维人体骨架模型,如图 1 所示.

将图 1 中的 16 个关节点作为特征点来实现跟踪.在具体的跟踪过程中,为了减少计算量,保证跟踪的精度,将特征点简化为 13 个.

在确定特征点后,就可以采用 KL 算法对其实现跟踪,以下介绍 KL 算法是如何实现跟踪的.假设 2 幅二维的灰度图像分别是 $I(x, y)$ 和 $J(x, y)$,有一个特征点 u 在第一幅图像 $I(x, y)$ 上,现在要解决的问题是如何在 $J(x, y)$ 也找到一个与之相对应的特征点 v .解决的方法是通过求 2 幅图像的偏移量 d 来最终确定 v 的位置关系,其中矢量 $d = (d_x, d_y)^T$ 表示的就是特征点位移,称之为光流.

为了使 d 值误差最小,建立误差函数(式(1)),当 d 最小化时,表明误差值最小,此时得到最佳特征点位移.

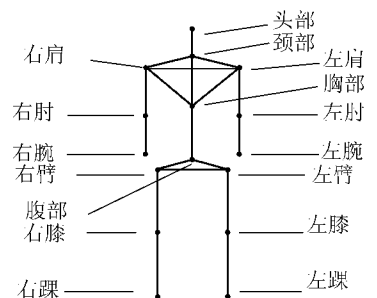


图 1 二维人体骨架模型

Fig.1 2D model of human skeleton

$$\varepsilon(d) = \varepsilon(d_x, d_y) = \sum_{x=u_x-w_x}^{u_x+w_x} \sum_{y=u_y-w_y}^{u_y+w_y} [I(x, y) - J(x + d_x, y + d_y)]^2 \tag{1}$$

式中 w_x, w_y 为正数,一般取值为 2,3,4,5,6,7. 通过给定特征点的初始位置和 2 幅连续的图像帧,就可以利用式(1)来求得该特征点的大致位移. 对于在特征跟踪过程中,由于存在自遮挡现象或者运动变化过大而导致特征点丢失的情况,可采用 Kalman 滤波^[9]的方法来预测特征点的位置,作为特征跟踪的补充和修正,用于后续的跟踪.

通过对一段视频首帧中的角色关节点进行特征点标注,然后在后续帧中做基于特征的跟踪,结果表明,在特征点选取合适的情况下,KL 算法能够保证跟踪的精度,如图 2 所示.



图 2 特征点跟踪结果
Fig.2 Tracking results of feature points

2 图像的三维重建

特征跟踪结果是运动信息不完全的二维角色运动序列,不能满足运动编辑和三维动画制作等应用的需要,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标定确定世界坐标系下空间特征点与图像坐标系下对应的图像特征点之间的关系,然后利用标定结果和二维图像跟踪结果进行三维重建.

2.1 摄像机的标定

在进行摄像机标定之前,必须了解物体的成像过程,即三维空间中的物体到视平面的投影关系. 理想的投影成像模型是光学中的小孔成像模型(pin-hole 模型)^[10]. 图 3 是物体的成像过程示意图. 图中 uO_iv 平面为二维成像平面, O 为视点, O_iuvw 为视觉坐标系, $O_wx_wy_wz_w$ 为空间坐标系. 物体 $P_w(x_w, y_w, z_w)$ 投影到视平面上的理想点为 P_u , 实际情况下图像的畸变等原因产生的实际投影点为 P_d . 根据上述成像过程,摄像机标定的目的就是计算空间坐标系与图像坐标系之间的关系,以及摄像机的内部参数.

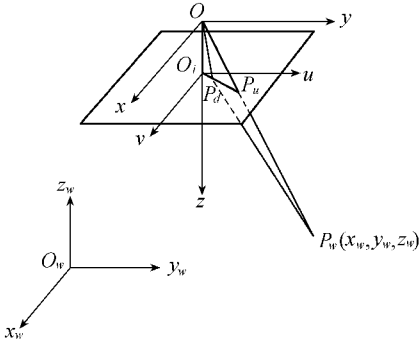


图 3 物体成像模型
Fig.3 A model of imaging

依据小孔成像模型,不考虑摄像机镜头的畸变因素,物体 P_w 从 (x_w, y_w, z_w) 投影到 $P_u(x_d, y_d)$ 的变换过程分为 3 步:(a) 三维世界坐标系到摄像机坐标系的坐标变换;(b) 小孔成像模型下的透视投影变换;(c) 实际图像平面坐标到计算机图像坐标变换.

通过上述变换过程,就可以确定空间坐标系与图像坐标系之间的关系,以及摄像机的内部参数.

2.2 人体骨架的三维重建

通过摄像机标定,得到摄像机的透视投影参数. 下一步工作就是将关节点的二维坐标转换为三维坐标. 从透视投影的小孔成像模型可以看到,连接摄像机光心和成像点得到一条投影直线,该投影直线上的所有点在投影平面上都成像于同一点. 为了在这条投影直线上找到三维关节点,需要利用人体骨骼长度知识. 因为人体的骨骼近似刚体,因而连接骨骼两端的人体关节点间距离近似不变,这就是关节点间的距离约束. 可以从一个与所求点相邻的已知三维关节点出发,寻找投影直线上的一个点,使得该点到已知三维关节点的距离等于 2 点间相应骨骼长度.

假设待求的三维坐标点 $P(x, y, z)$ 与光心间的连线为 m ,求 $P(x, y, z)$ 的过程可以表示为求直线 m 与

以 D (相邻的已知三维关节点)为球心, L (角色模型中关节的长度)为半径的球面的交点的过程.根据投影直线与球面可能出现的相交、相切、相离3种情况,该方程的解也有下列3种可能情况:

- a. 有2个解,即对应于直线与球面相交有2个交点的情况.这说明关节有2种可能位置.这种二义性是由于二维坐标向三维坐标恢复问题本身的病态性造成的.可以利用人体解剖学中的各种人体约束来消除人体三维骨架的二义性^[7],如当上臂前伸时,小臂不可能向后伸等.对于首帧图像,可以通过用户的干预来消除这种二义性.而后续帧图像则可以利用运动的连续性原则,使用户的干预达到最小.
- b. 只有1个解,对应于直线和球面相切的情况.此时,该切点就是所求的三维关节点.
- c. 无解,对应于直线和球面相离的情况.产生的原因有2个:(a)二维特征点坐标存在误差;(b)作为先验知识的人体骨骼长度过短,系统会适当地延伸骨骼长度之后再行计算.实验中延伸骨骼长度,使球面与投影直线相切,然后将切点作为待求点的三维坐标,同时用延伸的骨骼长度来校正利用人体解剖学比例计算的骨骼长度值.

首帧角色运动图像的三维骨架模型建立后,根据人体运动的连续性原理完成后续帧图像角色三维骨架结构的恢复.图4是三维重建的结果.从图4可以看出,捕获到的三维骨架模型自然逼真.

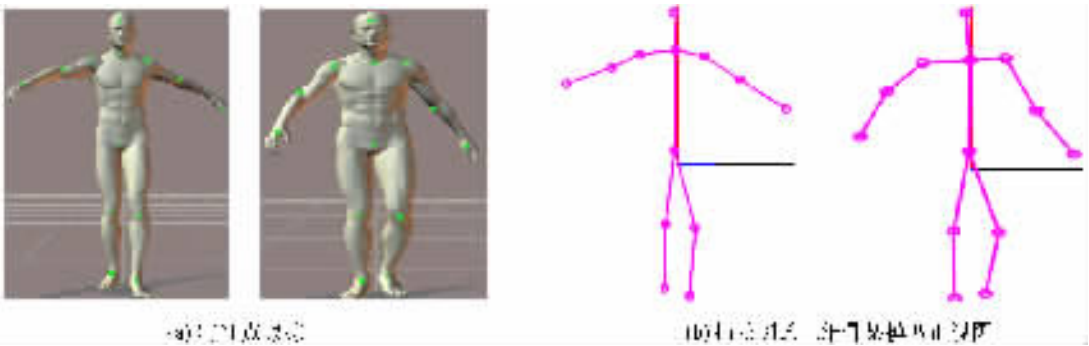


图4 三维重建
Fig.4 3D reconstruction

3 三维动画再生

在获得三维角色运动序列后,就可以将其重定向到三维的虚拟角色中,得到逼真的角色动画,如图5所示,产生的角色动画较好地反映了视频中人体运动的信息.

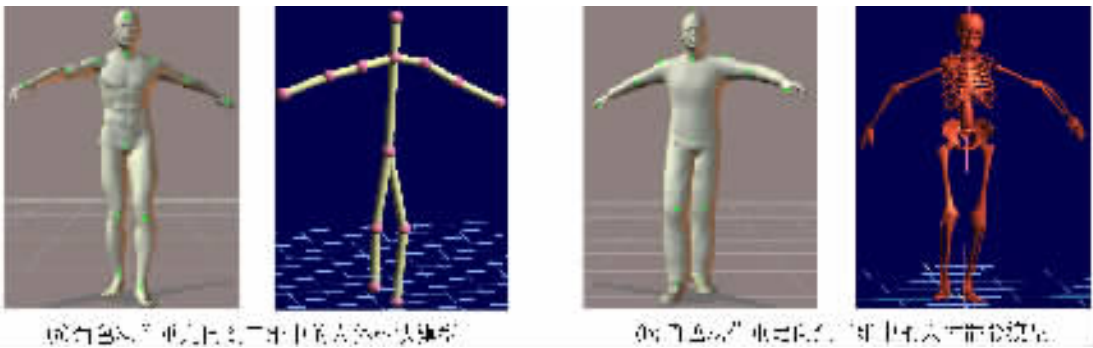


图5 三维动画再生
Fig.5 Regeneration of 3D animation

4 结 语

本文提出了一种新的基于视频的角色运动提取与运动合成方法,并介绍了其中关键技术.在运动捕获过程中,采用基于光流法中的KL算法和Kakman滤波相结合的特征跟踪算法,可以有效地保证跟踪精度,并通过三维重建建立起二维角色运动序列与三维角色运动序列之间的映射关系,最终得到虚拟环境下的角色动画.实际运行表明,与其他的运动捕获方法相比,本文提出的方法不需要在人体关节上附加任何传感器等昂贵的捕获设备,并且能充分利用现有丰富的影视片段作为素材源,保证了人体运动不受限,具有计算成本低、

制作方便、运动真实等特点.

参考文献：

[1] SABEL J C. Optical 3D motion measurement [C]//IEEE Instrumentation and Measurement Technology Conference. Brussels：IEEE Computer Society Press,1996 367-370.

[2] BREGLER C,MALIK J.Tracking people with twists and exponential maps [C]//Proceeding of IEEE Conference Computer Vision Pattern Recognition. Santa Barbara：IEEE Computer Society Press,1998 8-15.

[3] CHENG Jia-ching,MOURA J M F. Tracking human walking in dynamic scenes [C]//Proceeding of ICIP '97. Washington, D. C.：Computer Society Press,1997 137-140.

[4] 罗忠祥. 基于视频的人体动画 [J]. 计算机研究与发展,2003,40(2) 269-276.

[5] 罗忠祥,庄越挺,潘云鹤,等. 基于双摄像机的视频特征跟踪算法研究 [J]. 计算机辅助设计与图形学学报,2002,14(7) 646-650.

[6] BADLER N I. Animation 2000 + + [J]. IEEE Computer Graphics and Application,2000,20(1) 28-29.

[7] LIU Xiao-ming,ZHUANG Yue-ting,PAN Yun-he. Video based human animation technique [C]//Proceeding of the 7th ACM International Multimedia Conference(Multimedia '99). Orlando,Florida：ACM Press,1999 353-362.

[8] BOUGUET J Y. Pyramidal implementation of the Lucas Kanade Feature Tracker [EB/OL]. [2007-10-10]. <http://graphics.Stanford.edu/courses/cs448a-00-fall/bouget00.pdf>.

[9] 黄建新. Kalman 滤波的人体运动位置跟踪算法 [J]. 华侨大学学报：自然科学版,2003,24(3) 254-256.

[10] 高文,陈熙霖. 计算机视觉：算法与系统原理 [M]. 北京：清华大学出版社,1999.

Extraction and synthesis of the motion role based on video

SU Zhou, WANG Sheng

(College of Mechanical and Electrical Engineering, Hohai University, Changzhou 213022, China)

Abstract：The KL (Kanade-Lucas) feature tracking algorithm was used to track feature-articulated points in many videos and provide incomplete 2D character motion sequences. These sequences were reconstructed as 3D character motion sequences so as to meet the requirements of animation. Finally, 3D character motion sequences were mapped with 3D animated models after which they produced natural animation with 3D visual effects. The experimental results indicate that the 3D character animation matches character motion in the videos.

Key words：video；feature tracking；3D reconstruction；motion synthesis