

基于序列影像的人体运动信息获取

郑顺义¹

(1 武汉大学遥感信息工程学院, 武汉市珞喻路 129 号, 430079)

摘要: 实现了一种通用的从序列影像中获取任意的人体运动的软件系统。该系统基于一个普通的三维人体模型, 该模型的尺寸可以改变, 以与影像中的人体相匹配。该系统不需要特殊的设备和标志, 对序列影像也没有特殊的要求, 可以处理任意的人体运动, 操作简单, 适合人体运动信息获取的需要。

关键词: 序列影像; 人体运动信息获取; 人体重建; 四元组; 关键帧

中图法分类号: TP751

数字化的虚拟人产生的关键之一在于恢复人体运动, 即运动信息获取。近年来, 许多专家学者探索了利用各种不同的技术来实现人体的逼真重建和运动跟踪。人体运动信息获取可以用于计算机动画、计算机游戏、虚拟现实、电影特技、机器人控制、运动员运动分析等方面。正是由于这些潜在的广泛的应用, 使得该问题的研究成为目前计算机视觉和计算机动画领域内的研究热点之一。但目前要提取人的运动参数和模拟其运动姿态还不是一件简单的事, 要么需要在人体表面布置一些传感器或特殊的标志, 要么需要穿上特殊的服装, 这样就限制了人的运动空间。本文研究目的在于进一步推动人体运动信息获取, 探索一种对硬件环境要求较低, 又可以满足一般动画生成需要的技术。

光学式电磁式运动信息获取被证明是极其有效的人体运动信息获取方法^[1]。近年来, 出现了多种基于序列影像的人体重建与跟踪方法^[2~4], 这些方法非常有效, 但都采用过于简化的人体模型, 影响了关节的精确定位, 从而限制了其在某些领域的应用, 如医学临床、运动员动作分析等。

本文提出了一种基于序列影像的交互式人体运动信息获取方法, 它是一个个人系统, 不需要特殊的装置和设备以及相机检校等复杂的计算。该系统的最主要优点在于, 它对于序列影像的获取没有特殊的要求, 可以来自电视、电影片段、普通摄像等。

1 标准的人体模型

本文选择的模型如图 1 所示^[5], 它可以很好地表达人体的形状、姿势。模型包括人体各关节和躯干各部分的关系, 因而可以提供各种人体内的约束, 如对称约束等。该模型的大小、颜色以及纹理都可以根据具体的需要进行改变。本文没有采用更多的设备或添加更复杂的算法, 而是把重点放在提高操作和控制模型的灵活性和可靠性上。

该模型是一个树形结构, 包括 15 个关节, 每一个关节有 3 个自由度, 可以在局部坐标系内绕 X、Y、Z 坐标轴旋转。整体模型有 6 个自由度(3 个平移、3 个旋转)。通过调整这些自由度对应的参数, 可以为该模型设置各种姿势。

数据结构组织是交互式操作中一个非常关键的问题。本文采用标准的 VRML(virtual reality modeling language)语言组织和描述人体模型, 得到了比较好的效果, 同时, 开发了一套 VRML 语言编译系统。

2 交互式人体运动信息获取

2.1 关键帧获取

对于体育运动(如足球比赛)和体操表演等复杂的运动场景, 目前还没有完全自动的运动信息获取方法。本文选择交互式地提取关键帧姿态。

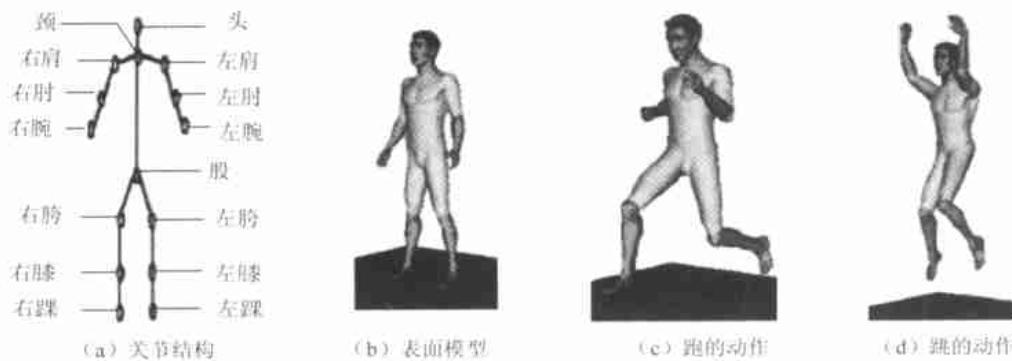


图 1 标准的人体模型

Fig. 1 Standard Human Body Model

然后内插中间帧姿态的方法来完成人体运动信息获取。交互式地提取关键帧姿态,首先要选择合适的关键帧影像,然后参照选择的关键帧影像,人工交互地调节人体模型的位置和躯干各部分的姿势。当人体模型与关键帧影像中的人体影像配准时,记录下当前的模型位置和姿态作为该关键帧对应的运动信息,如图 2 所示。

一般在以下情况需要设置关键帧:①当多个部位的运动开始和结束时;②相机开始移动时;③当旋转超过 180° 时,需要在中间加入一个关键

帧,提示旋转的方向。关键帧的多少取决于运动的复杂性,因为在本文开发的界面下可以很方便地调整模型的姿势,因此,关键帧的数目不是一个很严重的问题。每个关键帧记录了整个人体模型的位置和姿态以及每个动作的关键姿态,这些参数存为一个文件,以后可以重新打开调入该界面,重放该运动过程。通过单张影像确定身体某一部分的姿态时,可能会产生多个解。由于序列影像中影像之间的时间间隔很短,所以完全可以通过前后影像的信息来判断和选择正确的姿态。

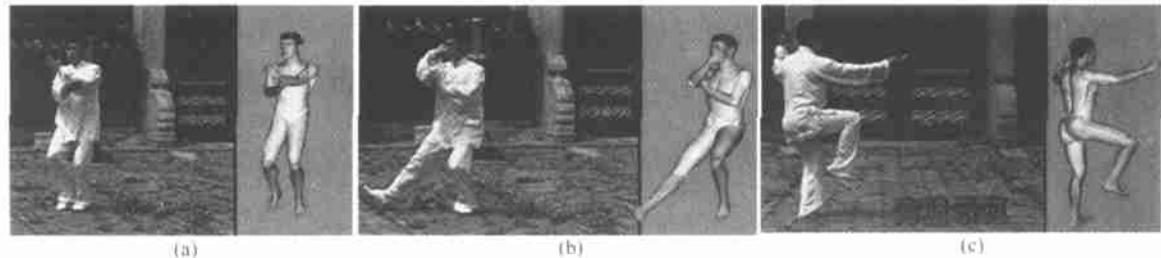


图 2 关键帧对应的运动状态的获取

Fig. 2 Motion Information Capture from Key Frames

2.2 内插中间帧

当获取到关键帧对应的模型运动状态之后,可以通过内插关键帧对应的运动状态来得到中间帧对应的运动状态,从而得到整个连续的运动过程。该方法在计算机动画处理中经常用到,是一种非常简便有效的方法。

根据上述确定关键帧的原则,内插出来的每个关节的运动状态变化应该是光滑连续的。每个关节的运动状态变化涉及到两方面的变化:一是位置的变化,即平移;一是方向的变化,即旋转。位置的变化是一个线性过程,直接通过内插两个位置状态就可以得到中间的位置状态。但方向的变化本身是一个非线性过程,直接简单的线性内插不能得到满意的结果。旋转变化有多种表示方

法,常用的包括:① 欧拉角,即通过绕 X 、 Y 、 Z 三个坐标轴的转角来表示方向;② 轴角,即通过一个轴(空间矢量)和绕该轴旋转的角度来表示方向。实验证明,直接对这两种表示方法中的参数进行线性内插无法得到光滑连续的运动变化。因此,本文在分析比较的基础上,选用了基于四元组(quaternion)的球面内插的方法来实现方向的内插,取得了光滑的运动效果。

四元组是复数的一种扩展^⑥,它通常表示为 $[s, v]$,其中, $s \in R$, $v \in R^3$, s 是比例因子, $v = (x, y, z)$ 是矢量因子。

设一旋转变换的旋转轴为 u (单位矢量),旋转角度为 θ 。则与此旋转变换等价的四元组 q 为:

$$q = (s, v) = (s, x, y, z) \quad (1)$$

式中, $s = \cos(\theta/2)$; $v = \mathbf{u} \sin(\theta/2)$ 。

在方向内插之前, 首先将表示方向的其他形式转化为四元组形式(如式(1)), 然后对以四元组方式表示的方向进行内插, 最后将四元组转化为原来的形式。直接内插公式为:

$$P_{\text{mid}} = k \cdot P + (1 - k) \cdot P' \quad (2)$$

式中, P_{mid} 为内插得到的中间帧; P, P' 为中间帧两端的关键帧; k 为内插系数。

实验证明, 直接对四元组进行线性内插可以得到相对满意的效果, 但可以观察到明显的方向变化速度不均匀, 在两个方向的中间变化剧烈, 而两端变化缓慢, 使得最后的视觉效果还存在缺陷。其原因如图 3 所示, 直接利用四元组对两个方向 P, P' 进行线性内插, 相当于在一个四维空间的单位球上的两个方向之间的连线进行内插, 然后再归一化到单位球上来确定内插的方向, 这样显然会使方向变化速度不均匀, 在中间角度变化速度明显加快。

本文在此基础上又进一步采用了球面线性内插的方法来改进该问题。这里不是直接对两个方向的四元组进行线性内插, 而是先计算两个四元组在单位球上的夹角, 然后线性内插这个夹角来计算新的方向(图 4)。该方法的运用在方向变化轨迹和变化速度两方面都取得了良好的效果。球面内插的具体公式为:

$$P_{\text{mid}} = P \cdot \frac{\sin((1 - k)\bar{\omega})}{\sin \bar{\omega}} + P' \cdot \frac{\sin k\bar{\omega}}{\sin \bar{\omega}} \quad (3)$$

式中, $\bar{\omega}$ 为 P 与 P' 之间的夹角; 其余符号同式(2)。

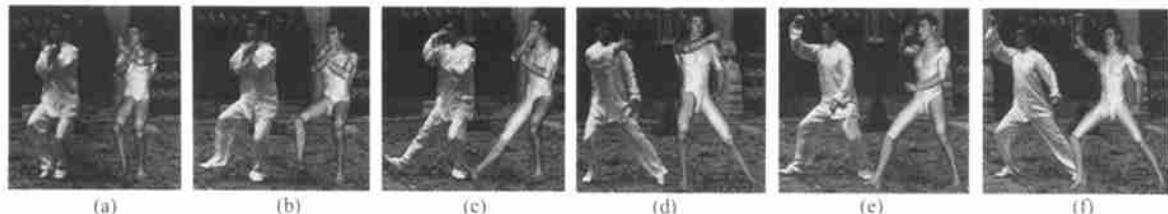


图 5 实验结果

Fig. 5 Results of Experiment

4 结语

本文提供了一个从序列影像中提取人体运动信息的工具, 该工具可以很方便地为计算机动画和其他多媒体应用提供光滑连续的人体运动参数。该方法基于一个标准的人体模型, 该模型自然而然地引入了许多人体中固有的约束, 使得整个运动信息获取减少了很多不必要的计算和人体

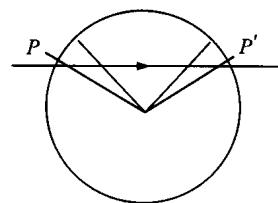


图 3 直接线性内插

Fig. 3 Direct Linear
Interpolation

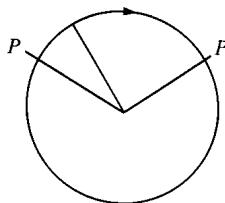


图 4 球面线性内插

Fig. 4 Spherical Linear
Interpolation

3 实验

根据以上所述的方法, 本文基于 OpenGL, 利用 Microsoft Visual C⁺⁺ 实现了整个运动信息获取系统。所有的序列影像贮存在计算机硬盘, 一帧一帧地调入界面。图像和模型的显示利用 OpenGL 图形库实现。利用该系统处理了一套大约 1 400 帧的序列影像(图 5), 选取了其中的 78 帧作为关键帧, 整个运动信息获取过程大约用了 2d 的时间, 这完全依赖于操作员的熟练程度。关键帧的选取与运动类型有关, 如较剧烈的运动就需要选取较多的关键帧。选取的关键帧越多, 获取的运动就越逼真。本文还测试过其他一些影像, 发现该方法是一种实用的运动信息获取方法。由于该研究是在初级阶段, 主要通过视觉检查的方法来判断运动信息获取的质量, 当三维人体模型的运动与序列影像中人的运动在速度和姿态上都基本吻合时, 就认为达到了要求。

运动理解中的不确定性。本文没有考虑采用其他附加设备或开发一些自动算法, 而是重点发展模型控制界面, 使得图形操作更方便、更灵活, 可以很方便地用于人体运动信息的获取。如果和逼真的人体模型重建程序结合, 该系统可以逼真地模拟特定人物的运动, 用于构造数字演员等。另外, 如果能同时处理多个相机获取的同步序列影像, 可以使获取的运动参数更加精确; 以及如何进一步提高自动化的程度等, 这些都是今后研究的重点。

参 考 文 献

- 1 Fua P, Gruen A, Plankers R, et al. Human Body Modeling and Motion Analysis from Video Sequences. International Archives of Photogrammetry and Remote Sensing, 1998, 32(B5): 866~873
- 2 Biegler C, Malik J. Tracking People with Twists and Exponential Maps. Conference on Computer Vision and Pattern Recognition, Santa Barbara, 1998
- 3 Deutscher J, Blake A, Reid I. Articulated Body Motion Capture by Annealed Particle Filtering. Conference on Computer Vision and Pattern Recognition, Hilton Head, 2000
- 4 Gavrila D M, Davis L. 3D Model Based Tracking of Humans in Action: A Multi-view Approach. Conference on Computer Vision and Pattern Recognition, San Francisco, 1996
- 5 Roehl B. Draft Specification for a Standard VRML Humanoid. Version 1.0. <http://ece.uwaterloo.ca/~h-anim/>, 1997
- 6 Barr A H, Currin B, Gabriel S, et al. Smooth Interpolation of Orientations with Angular Velocity Constraints Using Quaternions. Computer Graphics, 1992, 26(2): 313~320

作者简介: 郑顺义, 讲师, 博士。研究方向为计算机视觉、数字摄影测量和虚拟现实等。

E-mail: syzheng@263.net

Human Motion Capture from Video Sequences

ZHENG Shunyi¹

(1 School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: This paper develops a general software to acquire arbitrary motion from a video sequence. A general 3D articulate human model with changeable size, color and surface shape is prepared. This model is then driven either automatically or manually to match with the moving body in the consecutive image frames. This matching starts from key frames that contain key poses. A specific orientation interpolation is used to get the motion status in in-between frames and finally a smooth motion is obtained. This method is suitable for personal use to meet wide needs of human motion acquisition.

Key words: sequences image; human motion information capture; human body reconstruction; quaternion; key frames

About the author: ZHENG Shunyi, lecturer, Ph. D. His research focuses on computer vision, digital photogrammetry and virtual reality.
E-mail: syzheng@263.net

(责任编辑: 晓平)

欢迎订阅《地球空间信息科学学报》

《地球空间信息科学学报》为我国唯一的英文版测绘专业学术期刊。其宗旨是: 立足国内, 面向国际, 通过发表具有创新性和重大研究价值的测绘理论成果, 促进国内外学术交流。本刊内容包括综述与展望、学术论文与研究报告、本领域重大科技新闻, 涉及测绘研究的主要方面, 尤其是数字摄影测量与遥感、全球定位系统、地理信息系统及其集成等。本刊为国际性期刊, 按国际惯例运作, 作者和读者均面向国内外测量界。

本刊国内外公开发行, 读者对象为测绘及相关专业高级研究人员。本刊为季刊, A4 开本, 80 面, 逢季末月 5 日出版, 定价 10 元/册, 邮购价加 20%。本刊为自办发行, 欢迎广大读者直接向本刊编辑部索取订单, 踊跃订阅。