

# 运动捕捉及其在动画制作中的应用

黄波士, 陈福民

(同济大学计算机科学与工程系, 上海 200092)

**摘要:** 提出了基于运动捕捉技术的动画制作, 介绍了运动捕捉技术的过程及其3个关键技术的实现。给出了运动捕捉系统的模型和具体的实现方法。对于图像捕捉的同步问题进行了分析和解决。论述了基于运动捕捉技术的动画制作过程中的动画合成和消除滑步等技术。

**关键词:** 运动捕捉; 图像分析; 计算机视觉; 动画制作

## Motion Capture and Its Application in Animation Making

HUANG Boshi, CHEN Fumin

(Department of Computer Science and Engineering, Tongji University, Shanghai 200092)

**【Abstract】** Animation making based on motion capture technique is proposed. Process of motion capture and its three key techniques are introduced. Model of motion capture and its implementation are given, including analysis and solution of the synchronization of image capture. At last, animation synthesis and footskate cleanup during animation making based on motion capture technique is introduced.

**【Key words】** Motion capture; Image analysis; Computer vision; Animation making

在传统的动画制作中,动画师必须根据剧情要求将画面逐帧画出,工作量非常大。引入计算机动画技术后,可以利用计算机先设计造型,再按照剧情确定关键帧,然后动画师调整关键帧的造型姿势,动画软件则根据关键帧生成图像序列。但是,对于一个长的动画作品,逐个确定关键帧仍然是一项相当麻烦的工作。表演动画技术的诞生,彻底改变了这一局面,将动画师从繁重的体力劳动解放出来。它综合运用计算机图形学、光学、计算机视觉、计算机动画等技术,捕捉表演者的动作甚至表情,用这些动作或表情数据直接驱动动画模型。本文论述了动作捕捉技术及其数据在动画制作中的处理和应用。

### 1 运动捕捉技术

#### 1.1 运动捕捉的过程

人体的动作可以看成是人体各个关节点的动作,在运动捕捉系统中,一般把人体看成是由13~19关节点组成的简单模型<sup>[1]</sup>。在运动捕捉之前,人体的各个关节点上固定一个特殊的反光材料,称为标记点(Marker),这些反光材料在外部光源的照射下可以从不同角度反射出RGB值相同的光。利用两台或两台以上的摄像机进行实时视频捕捉,从各个摄像机得到的序列图片中可以看到每一帧中标记点的运动情况。因此可以得到一个特定的点随着时间变化的连续运动轨迹。然后通过三维重建技术将这些点的运动轨迹还原为骨架模型的动作。

#### 1.2 运动捕捉系统的设计

运动捕捉系统的设计如图1所示,运动捕捉的关键技术是标记点跟踪以及空间坐标的三维重建。另外,在计算机视觉中,从二维图像信息计算三维空间结构,需要利用视点的位置信息和视点的朝向信息,这就要用到摄像机的各种参数。空间物体表面某点的三维几何位置与其在图像中对应点之间的相互关系是由摄像机成像的几何模型决定的。这些几何模型参数就是摄像机参数,要通过摄像机标定计算才能得到。摄像机标定也是运动捕捉关键技术之一。

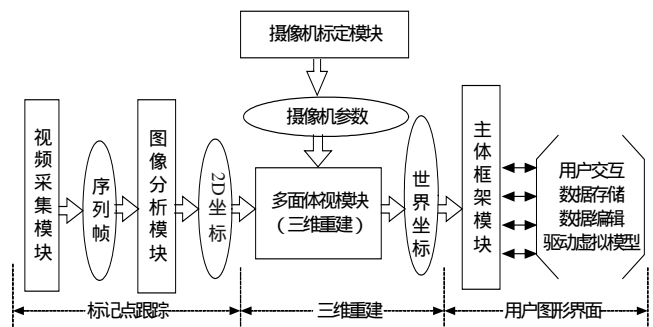


图1 运动捕捉系统模型

采用高性能的视频捕捉卡来捕捉图像,这种捕捉卡可以一帧帧地将视频输入信号直接数字化成YUV 4:2:2信号,并能够将它转换成RGB格式,然后利用PCI总线,把图像数据传到VGA卡显示或计算机内存存储。由于需要得到的是各个摄像机在同一个时刻捕捉的图像,因此不能够直接利用捕捉卡的多路采集功能。这需要在同一台高性能计算机上安装多块捕捉卡,每块捕捉卡对应一台摄像机。系统中捕捉的难度主要在于如何在多块采集卡之间同步采集。

### 2 图像捕捉及标记点跟踪

#### 2.1 图像捕捉及图像分析

图像捕捉最主要的问题在于如何保证多个捕捉线程对每一帧都几乎在同一时间发出捕捉命令,根据计算机视觉原理,计算模块要求多台摄像机几乎在同一时刻拍摄同一场景,否则从2D~3D之间的计算将毫无意义。因此,需要配置性能较高的计算机,起初我们采用定时器(包括多媒体时

**基金项目:** 上海市科技型中小型技术创新基金(种子基金)资助项目(0351H1325)

**作者简介:** 黄波士(1978—),男,博士生,主研方向为网络多媒体技术和虚拟现实技术;陈福民,教授、博导

**定稿日期:** 2004-05-19 **E-mail:** hyuantian@163.com

钟,可等的计时器)都不能满足大约 20FPS 的要求。最后,我们降低了捕捉的频率,采用中间插值的方法使数据达到制作动画的要求。利用事件和信号量内核对象来使视频捕捉线程和计算线程达到同步,而且采集线程在多个地方方式地调用,使得多台摄像机间能在同一时间捕捉图像。

视频捕捉卡提供了两种将数据存储到内存中的方法,这里选取了其中性能较好的双缓存的方法。在内存中开辟两块缓存,一块用于捕捉图像,另一块用来对已经获得的图像数据进行分析。这样,在捕捉的同时也可以对已经采集好的图像数据进行分析处理。在完成一帧图像的采集后,将两块缓存切换。

整个视频捕捉模块如图 2 所示,对每一块捕捉卡,都开了一个视频捕捉线程负责其图像捕捉及分析。另外还开了一个计算线程,这个线程利用分析出来的标记点的图像坐标计算其空间位置。

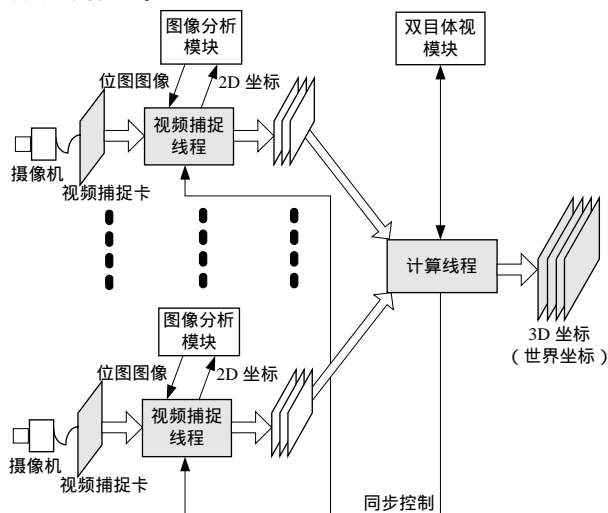


图2 视频捕捉模块

## 2.2 标记点跟踪

为了能使多个标记点一一对应,需要在起始帧处对所有标记点编号,然后,对这些标记点进行跟踪,得到标记点的运动轨迹。系统中采用卡尔曼滤波抢先预测算法来进行跟踪。卡尔曼滤波算法的核心是给所测目标进行建模,根据实际情况,建立如下模型:

$$x((n+1)T) = x(nT) + Tv(nT) + \frac{T^2}{2} a(nT)$$

其中,  $x(n)$  是时刻  $n$  的位置,  $v(n)$  是时刻  $n$  的速度。为了方便起见,改写为  $X(n+1) = FX(n) + Ga(n)$ 。其中,  $a(n)$  是加速度,作为一个高斯白噪声建模的。因为加速度的变化依赖于目标的移动,因此在程序中能修改这个设置是很重要的。一般加速度是被看作时间  $T$  上的一个常数。

在卡尔曼滤波估计的基础上,标记点可以用下面的步骤来跟踪:

- (1) 分析通过不同摄像头捕捉的两幅图片,给所有的标记点进行编号,而且两幅图片的编号要一一对应;
- (2) 对于一个特定的标记点,用传统跟踪方法得出此点在下一帧的位置;
- (3) 以后每帧用卡尔曼滤波估计出一个位置;
- (4) 利用 8 连接搜索法,搜索估计值位置周围的标记点,最先搜索到的一点看成是本帧此标记点的实际位置,转到(6);

(5) 如果超过一个特定的搜索范围仍然没有找到标记点,则认为此点在本帧被遮蔽,将估计值作为跟踪值;

(6) 将跟踪到的数据(标记点的图像坐标)保存,转到步骤(3)。

## 3 摄像机标定及三维重建

由于摄像机成像并不是理想状态的小孔成像,此外,摄像机还存在制造和装配误差,因此,实际成像与理想的小孔成像之间存在差异,这些差异称为畸变。为精确地建立含畸变的图像坐标与实物三维坐标之间的关系,需要在变换关系中引入畸变修正项。摄像机的镜头畸变主要有径向畸变和切向畸变两种,我们在标定时只考虑径向畸变来对摄像机进行标定。径向畸变主要来源于镜片组的各镜片间轴向间距的偏差。

摄像机的待标定参数包括内部参数和外部参数。内部参数包括:

(1) 比例系数  $N_x, N_y$ , 从图像平面坐标到帧缓存中的像素坐标的长宽比例因子;

(2) 图像中心点  $X_c, Y_c$ , 摄像机光轴所经过的摄像机投影平面(也即图像平面)中的位置,一般说来,不是严格地在投影平面的中间;

(3) 摄像机的焦距  $f$ ;

(4) 摄像机径向畸变的系数。

外部参数包括:

(1) 摄像机光轴矢量在世界坐标系中的侧倾角、俯仰角、旋转角;

(2) 摄像机的透镜在世界坐标系中的坐标。

根据通用摄像机模型,建立一个关系等式<sup>[6]</sup>,令  $A=PRT$ ,  $A$  中的元素包括摄像机平移、旋转和投影系数。点  $W_h$  的摄像机坐标系统的坐标  $C_h = AW_h$ , 由于笛卡尔形式的摄像机坐标是  $x = C_{h1} / C_{h4}$  和  $y = C_{h2} / C_{h4}$ , 将上述矩阵展开, 就可以得到共有 12 个未知量的两个方程

$$(a_{11} - a_{41})X + (a_{12} - a_{42})Y + (a_{13} - a_{43})Z + (a_{14} - a_{44}) = 0$$

$$(a_{21} - a_{41})X + (a_{22} - a_{42})Y + (a_{23} - a_{43})Z + (a_{24} - a_{44}) = 0$$

因此,标定的过程就包括:(1) 获得  $M$  6 个具有已知世界坐标  $(X_i, Y_i, Z_i)$ ,  $i = 1, 2, \dots, M$  空间点(2) 用摄像机在给定位置拍摄这些点得到它们对应的像平面坐标  $(x_i, y_i)$ ,  $i = 1, 2, \dots, M$ ; (3) 把这些坐标代入上面两个式子就可以解出未知系数。

三维重建就是从标记点的计算机坐标得到它的世界坐标,过程如图 3 所示。三维重建最关键的问题是如何从两幅图片中找到物体的对应点,不过标记点的跟踪已经解决了这个问题。注意:用三角形法进行三维重建的时候基线不能太小,否则会影响精度。

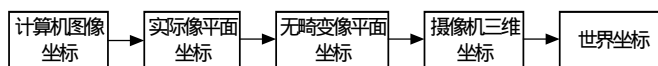


图3 三维重建坐标变换过程

## 4 数据处理及动画制作

动画特别是三维动画的制作过程相当烦琐,需要导演、动画师和表演者的默契配合。基于运动捕捉的动画制作还涉及到对捕捉数据的处理,包括消除滑步、动作匹配与合成等,制作过程如图 4 所示。

运动捕捉数据的一个缺点是不能提供整个运动的连续

控制,而且,运动捕捉是耗时且造价昂贵的一个过程,另外,表演者一般也不可能做到像动画师想像的那样,因此可用的数据只能是一段一段的<sup>[3]</sup>。在实际制作的过程中,动画师需要将有用的运动捕捉数据和所设计的情节进行匹配和合成。人在运动的时候,关节之间有很多联系,例如行走,右脚向前走,左胳膊向前摆动;当臀部角有一个值时,膝盖角就也有一个值,且这两个值之间存在着一定的联系。通过这种联系动画师和所设立的关键帧进行合成,形成连续的动画。

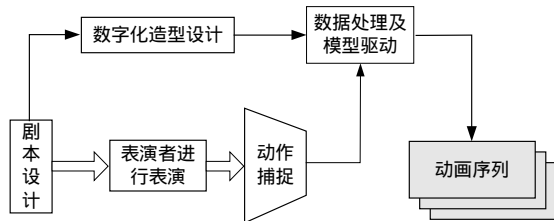


图4 动画制作的工作流程

制作时,动画师需要提供下列信息:(1)哪个连接角需要结构匹配;(2)哪个连接角要人工合成;(3)哪个关节要用来驱动动作。假设动画师画出一个行走的动画且只画了腿部,那就需要去合成上半身的运动。最好的选择是分析臀部角和膝盖角之间的关系,因为它们限制了走的运动。把这些数据分成段,用来发现臀部角和膝盖角运动捕捉数据的段和哪个创建的关键帧相似。相对应上身运动捕捉数据的段可以用来驱动计算机人物的上半身。

为了达到这个要求,需要选择一个对应的数据区域。事实上,由于关键帧数据也许会真实的数据在时间上是不同的,因此问题是很复杂的。另外,选择的段的结束处必须很光滑地连接在一起,避免运动的高频波动。一般需要分成下面4步:(1)频率分析;(2)匹配;(3)路径查找;(4)加入。

最后,为了消除动作的不平滑,在两个加入点之间进行二次拟合,为了提高拟合的精度,需要考虑这两个点之外的N个点。一般N从5~20是比较有效的。对于拟合的结果用二次正弦函数和原始连接数据进行混合。

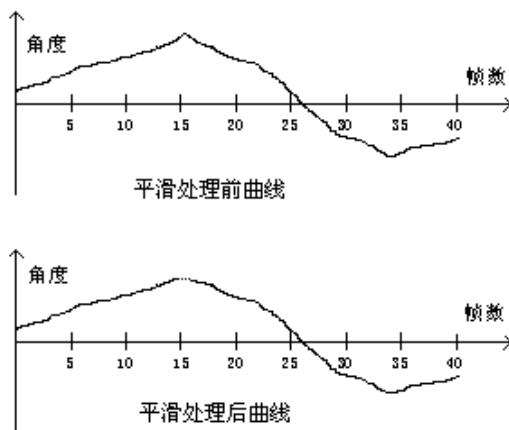


图5 平滑处理前后曲线对照

定义混合参数  $f = (\cos \pi / 2N + 1)^2$ , 这里 N 是要连接的最短段长度的一半, t 是时间,在连接点处为 0。如果定义

q 为从拟合中得到的二次函数, m 为匹配后的数据,那么经过平滑后的数据为:  $s(t) = f(t)q(t) + (1 - f(t))m(t)$ , 平滑处理前后曲线比较如图 5 所示。

滑步的消除是基于骨架运动函数<sup>[5]</sup>  $M(t) = (p_k(t), q_o, \dots, q_k, o_o(t), \dots, o_k(t))$  的,用解析 IK 算法来满足所有的步伐约束。一般的消除滑步的步骤如下:

- (1) 给每一个约束设定一个位置;
- (2) 对于每一个约束帧计算满足前一步约束位置的踝的全局位置和方向;
- (3) 计算脚跟的位置,使得前一帧发现的踝可以在伸腿时达到;
- (4) 对于每一个约束踝,调节腿使踝后部能满足步骤(2)的构造;
- (5) 转到步骤(2),并要确定在调节骨架参数满足约束的时候没有改变动作的平滑性。

在所有的前期处理完成以后,就可以进行模型的驱动了。目前流行三维建模的软件是 Maya,动画师先用 Maya 建立一个满意的模型,然后进行动画的合成。值得一提的是,用 Maya 建立的模型有时和真人一模一样,让观众看不出来那是动画,可以代替电影的实拍过程,这也是现在数码电影制作的一个趋势。

## 5 结论与展望

本文论述了运动捕捉技术及其在电影动画制作中的应用。随着数码电影制作的发展,对运动捕捉的需要也会越来越高。运动捕捉是一个较为复杂的基于计算机视觉的技术,对于系统的准确性和实时性以及实用性还需要进一步的完善。此外,运动捕捉技术还可用于机器人遥控、体育训练、人体生理学、生物力学研究等领域。相信,随着技术本身的发展和相关应用领域技术水平的提高,运动捕捉技术将会得到越来越广泛的应用。

## 参考文献

- 1 Bodenheimer B, Rose C, Rosenthal S, et al. The Process of Motion Capture: Dealing with Data. 8th Eurographics Int. Workshop on Computer Animation and Simulation '97, Budapest, Hungary, Springer-Verlag Wien, 1997:3-18
- 2 Shen Tzungsz, Menq Chiahsiang. Automatic Camera Calibration for a Multiple-sensor Integrated Coordinate Measurement System. IEEE Transactions on Robotics and Automation, 2001,17(4)
- 3 Pullen K, Bregler C. Motion Capture Assisted Animation: Texturing and Synthesis. Proceedings of ACM SIGGRAPH 02, 2002
- 4 Fischer R. Motion Capture and Motion Editing. Chapter 2, Peters A K, 2000
- 5 Kovar L, Schreiner J, Gleicher M. Footskate Cleanup for Motion Capture Editing. Proceedings of the 2002 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2002:97-104
- 6 章毓晋. 图像工程——图像理解与计算机视觉. 北京:清华大学出版社, 2000