

从运动捕获数据中提取关键帧

沈军行 孙守迁 潘云鹤

(浙江大学计算机科学与技术学院 杭州 310027)

(浙江大学 CAD & CG 国家重点实验室 杭州 310027)

摘 要 在四元数表示下导出了两个旋转之间差异的一种简单形式,将人体各关节上总的旋转变化作为帧间距,设计了一种从运动捕获数据中提取关键帧的高效算法.通过向量的线性插值和四元数的球面线性插值重建了原始动画.实验表明,该算法具有很好的数据压缩效果,而且提取出的关键帧在视觉上对原始动画具有概括力.

关键词 计算机动画;运动捕获;运动编辑;四元数;关键帧

中图法分类号 TP391.41

Key-Frame Extraction from Motion Capture Data

Shen Junxing Sun Shouqian Pan Yunhe

(College of Computer Science and Technology, Zhejiang University, Hangzhou 310027)

(State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310027)

Abstract A compact form of the difference between two rotations represented by quaternion is derived, and the total rotation differences in all joints are taken as the distance between two frames. Based on this frame distance, an efficient algorithm is presented to extract key-frames from motion capture data. The whole animation process is reconstructed by simple linear interpolation of root joint translation and spherical linear interpolation of all joint rotations. Running experiment show that the original motion capture data can be compressed in a high ratio, the extracted key frames give a good visual summarization of the original animation and are useful for storing, transferring and browsing the motion data.

Key words computer animation; motion capture; motion editing; quaternion; key-frame

1 引言

由于通过运动捕获技术获得的动画具有高度的逼真性,因此该技术目前在计算机游戏、影视制作等领域得到了广泛的应用.但同传统的关键帧技术和动态仿真技术相比,运动捕获记录的是一堆庞大的

原始数据,缺少结构化信息,因此对其进行编辑具有较大难度.Bruderlin 等^[1]将运动数据作为信号来处理,并采用多分辨率思想在不同细节层次上对运动进行控制. Motion Displacement Mapping 技术^[1-3]将由于编辑调整所引起的运动变化同原始运动分离开来,独立选择表示变化的插值曲线,从而避免了引入不必要的高频分量,使原始运动的细节得以保留.

原稿收到日期:2003-04-01;修改稿收到日期:2003-06-26. 本课题得到国家重点基础研究发展计划(2002CB312106)和广东省自然科学基金(033323)资助. 沈军行,男,1978年生,硕士研究生,主要研究方向为虚拟人、计算机图形学、计算机视觉. 孙守迁,男,1963年生,博士,教授,博士生导师,主要研究方向为计算机辅助工业设计与概念设计、网络化媒体设计、虚拟人、数字化产品开发等. 潘云鹤,男,1946年生,教授,博士生导师,中国工程院院士,主要研究方向为 CAD&CG、人工智能.

Witkin 等提出了时空约束方法^[4],该方法运用非线性最优化求解生成满足用户所指定约束的动画. Gleicher 将时空约束方法应用于运动编辑^[5]和运动重定向^[6],并通过简化提高了问题求解速度. Rose 等应用时空约束方法产生了两段运动之间的平滑过渡^[7]. 信号处理方法虽然思想简单,但无法从高层次上对运动进行编辑,在很多情况下不能预知参数的调整会对运动产生什么样的影响;时空约束方法求解困难,一般很难达到交互的要求.

在传统的键帧动画中,由动画师先编辑一些关键帧,然后通过插值生成一段连续的动画;另外,关键帧提取在基于内容的视频检索中有着很广泛的应用^[8-9],而三维动画和二维视频之间存在一定的相似性. 受上述两个关键帧应用技术的启发,本文研究了从运动捕获得到的三维人体动画中提取关键帧的算法.

2 四元数及其间距

人体是一种关节链结构,其状态可用根关节的平移和所有关节的旋转表示. 平移即是一般的三维向量,旋转可用旋转矩阵、欧拉角或四元数来表示^[10]. 由于四元数在插值方面的优越性,目前广泛应用于计算机动画中.

2.1 四元数的定义

设 $s = w \in \mathbb{R}$, $V = (x, y, z) \in \mathbb{R}^3$, 四元数 $q \in S^4$ 定义为 $q = [s, V] = [w, x, y, z]$. q 的共轭 $\bar{q} = [s, -V]$, q 的模 $\|q\| = \sqrt{w^2 + x^2 + y^2 + z^2}$, q 的逆 $q^{-1} = \bar{q} / \|q\|$. 记 $q_1 = [s_1, V_1]$, $q_2 = [s_2, V_2]$, 四元数乘法定义为 $q_1 q_2 = [s_1 s_2 - V_1 \cdot V_2, s_1 V_2 + s_2 V_1 + V_1 \times V_2]$, 不难证明四元数乘法可结合但不可交换,且满足 $(q q_2)^{-1} = q_2^{-1} q^{-1}$. 若 $\|q\| = 1$, 则称 q 为单位四元数,两个单位四元数相乘结果仍为单位四元数. 所有单位四元数组成的集合记为 S^3 , 不难想像 S^3 应该是 4 维空间上的一个单位超球面.

若有单位四元数 $q = [\cos \theta, u \sin \theta]$, 其中 u 为 \mathbb{R}^3 上的一个单位向量. 设有向量 $P, P' \in \mathbb{R}^3$, 满足

$$[0, P'] = q[0, P]q^{-1} \quad (1)$$

则可证明 P' 与 P 绕轴 u 旋转 2θ 角度的结果一致^[10-11], 式(1)即为用四元数表示旋转的形式. 从旋转意义上讲, q 与 $-q$ 是等同的, 即 $q[0, P]q^{-1} = (-q)[0, P](-q)^{-1}$.

2.2 四元数之间的距离

设参考向量 $P_0 \in \mathbb{R}^3$, 单位四元数 q_1 将 P_0 旋

转到 P_1 , q_2 将 P_0 旋转到 P_2 (如图 1 所示), 即有 $[0, P_1] = q_1[0, P_0]q_1^{-1}$ 和 $[0, P_2] = q_2[0, P_0]q_2^{-1}$. 设单位四元数 $q_2 q_1^{-1} = [w, x, y, z] = [\cos \theta, \sin \theta(a, b, c)]$ ($\|(a, b, c)\| = 1, \theta = \arccos w \in [0, \pi]$), 现将其作用于向量 P_1 :

$$\begin{aligned} & (q_2 q_1^{-1})[0, P]q_1 q_1^{-1} = \\ & (q_2 q_1^{-1})(q_1[0, P_0]q_1^{-1})(q_2 q_1^{-1})^{-1} = \\ & (q_2 q_1^{-1})(q_1[0, P_0]q_1^{-1})(q_1 q_2^{-1}) = \\ & q_2(q_1^{-1} q_1)[0, P_0](q_1^{-1} q_1)q_2^{-1} = \\ & q_2[0, P_0]q_2^{-1} = [0, P_2] \end{aligned} \quad (2)$$

由式(2)可见, $q_2 q_1^{-1}$ 将 P_1 转到 P_2 (即 P_1 绕轴 (a, b, c) 旋转 2θ 角到 P_2), 这就说明为了将 q_1 产生的旋转变换到 q_2 产生的旋转须通过 $q_2 q_1^{-1}$ 作一额外旋转. 因此 θ 自然就可用来衡量 q_1 和 q_2 之间的距离, 同时表示 q_1, q_2 所产生的旋转之间的差异 (如图 1 所示).

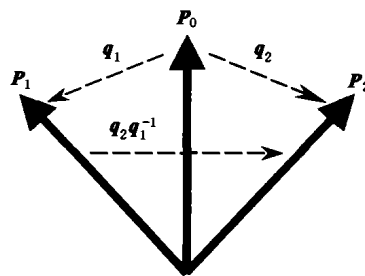


图 1 四元数之间的距离

定义 1. q_1, q_2 为两个单位四元数, 若 $q_2 q_1^{-1} = [w, x, y, z]$, 则 q_1, q_2 之间的距离定义为

$$d(q_1, q_2) = \arccos w \quad (3)$$

3 从运动捕获数据中提取关键帧

为了连续地记录表演者的动作, 要求运动捕获设备有很高的采样速率, 一些光学运动捕获设备的采样速率可达 1000 帧/s 以上. 如此高的速率, 以至记录的数据量过于庞大, 对其进行编辑也显得非常困难. 一种简单的方法是对原数据进行等间隔重采样, 但是通常表演者在整个表演过程中其动作节奏或者说频率不是恒定的, 如一段舞蹈表演时而缓慢抒情, 时而激烈奔放. 等间隔重采样必然在缓慢部分过采样, 造成数据冗余; 而在激烈部分欠采样, 丢失动作细节.

为了能够进行插值重建, 希望两个关键帧之间呈线性变化. 经典的曲线分裂算法^[9]检测曲线中的“尖角”作为关键帧, “尖角”之间近似线性, 算法递归

进行,如图 2 所示。但是该算法计算量比较大,且很难

推广到高维。为此我们设计了一种高效的近似算法。

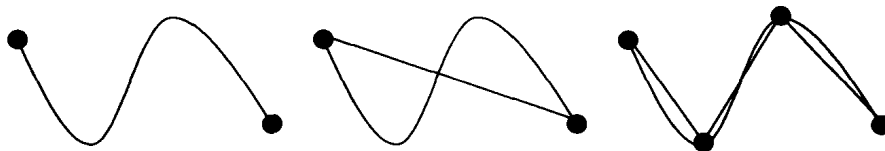


图 2 曲线分裂算法

人体运动可由一离散时间向量函数 $m(t) = [p(t), q_1(t), q_2(t), \dots, q_n(t)] (1 \leq t \leq T)$ 表示, 其中, $p(t) \in \mathbb{R}^3$ 表示根关节的平移运动, $q_i(t) \in S^3 (i=1, 2, \dots, n)$ 表示关节 i 的旋转运动。下面引入帧间距用来衡量两帧之间人体姿势的差异。

定义 2. 设 $m(t_1), m(t_2)$ 分别为不同时刻 t_1, t_2 上的两帧, 它们之间的距离定义为

$$D(t_1, t_2) = \|m(t_1) - m(t_2)\| = \sum_{i=1}^n \alpha_i d(q_i(t_1), q_i(t_2)) \quad (4)$$

其中, $\alpha_i (i=1, 2, \dots, n)$ 为加权系数, 表示各关节的重要度; $d(q_i(t_1), q_i(t_2))$ 根据式(3)算得。因为人体结构是一种树状的关节链, 父关节的运动带动子关节一起运动, 所以从运动上来讲父关节显得更加重要。

算法 1. 设动态数组 **KeyArray** 用来记录提取出的关键帧, *last* 表示目前为止最后找到的关键帧, T 表示总帧数, δ 为事先设定的一阈值。

Step1. 将第一帧作为关键帧加入到 **KeyArray**, 并置 $last = 1$, 初始化循环变量 $t = 2$ 。

Step2. 若 $t > T$, 则算法结束; 否则, 按式(4)计算帧间距 $dist = D(t, last)$ 。

Step3. 若 $dist > \delta$, 则第 t 帧作为关键帧加到 **KeyArray**, 并置 $last = t$ 。

Step4. $t = t + 1$, 返回 Step2。

算法结束时 **KeyArray** 记录了所有提取出的关键帧。

因为运动捕获具有较高的采样速率, 所以数据在局部上是近似线性变化的。算法不断考察当前帧同最后一个关键帧之间的差异情况, 若超过阈值 δ 则认为已超越线性变化范围并作为新关键帧加入。

4 重建动画

重建即是在关键帧之间的插值过程, 根关节的平移用一般的线性插值, 而所有关节的旋转用四元

数的球面线性插值 Slerp 算法^[10]。若 $q_1 = [w_1, x_1, y_1, z_1], q_2 = [w_2, x_2, y_2, z_2]$ 为两个单位四元数, 它们之间的球面线性插值为^[10-11]

$$Slerp(q_1, q_2; u) = \frac{\sin(1-u)\theta}{\sin\theta} q_1 + \frac{\sin u\theta}{\sin\theta} q_2 \quad (5)$$

其中 $\theta = \arccos(w_1 w_2 + x_1 x_2 + y_1 y_2 + z_1 z_2)$ 。

算法 2. 算法 1 中得到的结果 **KeyArray** 记录了所有提取出的关键帧, 记 *num* 为关键帧数目, 对两相邻关键帧之间的非关键帧进行线性插值重建步骤如下:

Step1. 初始化循环变量 $k = 0$ 。

Step2. 若 $k \geq num - 1$, 算法停止; 否则, 取其中两相邻关键帧 $t_1 = \text{KeyArray}[k], t_2 = \text{KeyArray}[k + 1]$ 。

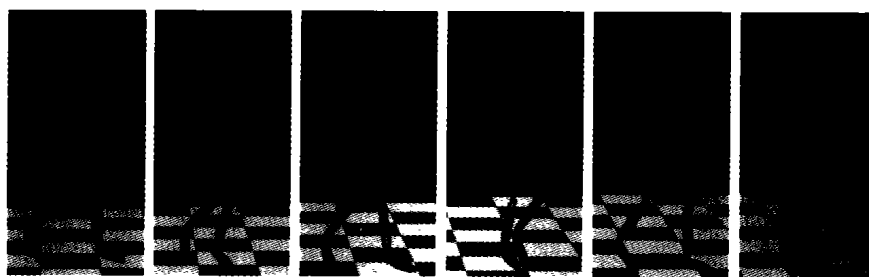
Step3. 对 t_1 和 t_2 之间的所有非关键帧 $t (t_1 < t < t_2)$ 插值 $u = (t - t_1) / (t_2 - t_1)$; 根关节平移 $p(t) = (1 - u)p(t_1) + (u)p(t_2)$; 所有关节的旋转按式(5) $q_i(t) = Slerp(q_i(t_1), q_i(t_2); u), i = 1, 2, \dots, n)$ 。

Step4. $k = k + 1$, 返回 Step2。

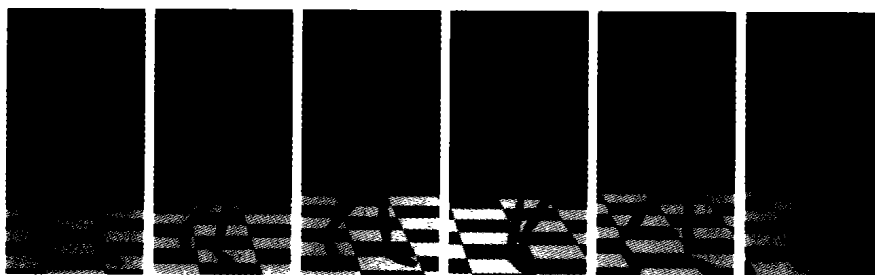
5 实验结果

我们用 Visual C++ 6.0 和 OpenGL 开发了一个运动编辑系统, 命名为 Eidolon, 其中包含了在引言中提到的一些运动编辑技术。本文提出的算法已在 Eidolon 中实现和验证。如图 3 所示, 其中人体行走动画原始共有 163 帧, 速率为 30 帧/s, 表演者开始快速行走, 然后慢慢停下来。用算法 1 在 $\delta = 1.105$ 下共提取出 18 帧关键帧, 压缩率达 11%。为了作比较, 用等间隔重采样法也提取了 18 帧。图 3 所示为原始动画和插值重建结果, 可见关键帧提取的重建结果基本没有失真, 而等间隔重采样的重建结果有明显失真。

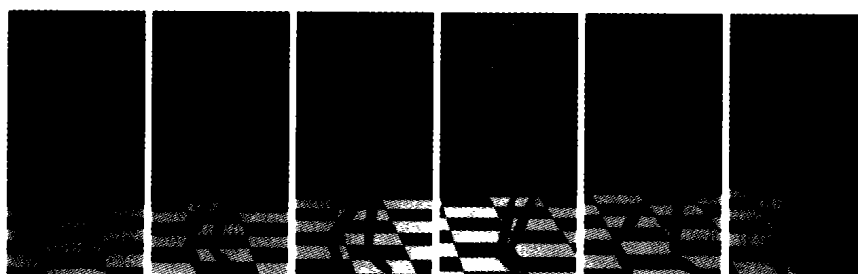
图 4 所示为从一段具有 198 帧的运动捕获动画中用本文方法提取出的 8 帧关键帧 ($\delta = 4.0$), 可见这些关键帧在视觉上对原始动画具有概括力。



a 原始动画中的第 5,15,25,35,45,55 帧



b 用本文方法进行关键帧提取后插值重建结果中的第 5,15,25,35,45,55 帧



c 等间隔重采样后插值重建结果中的第 5,15,25,35,45,55 帧

图 3 原始动画和插值重建结果

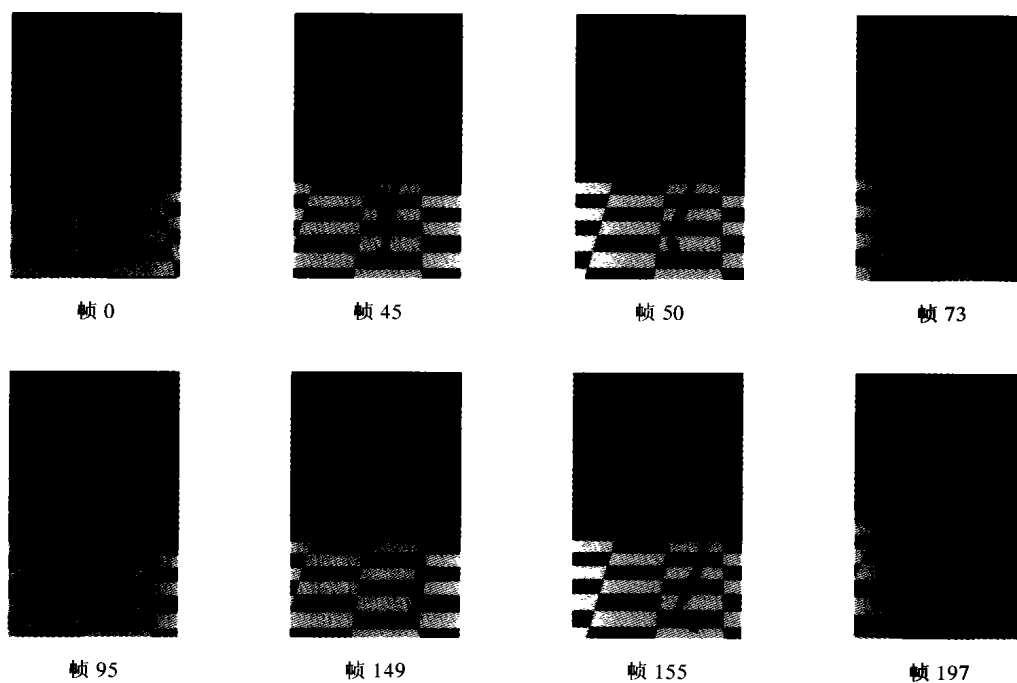


图 4 提取出的关键帧在视觉上对原始动画具有概括力

6 总 结

算法 1 作为对曲线分裂算法的一种近似,有效地克服了等间隔采样的缺点. 但算法 1 中的阈值 δ 由用户指定,不同的 δ 值给出不同的关键帧. 我们将进一步研究使算法具备自适应能力,即用户只需给出所要提取的关键帧数目,算法自动给出相应的阈值 δ .

本文所做的从运动捕获数据中提取关键帧的研究工作具有如下三个方面的意义:

(1) 运动编辑可集中在关键帧上,降低了运动编辑复杂度;

(2) 通过关键帧之间的插值可重建原始动画,起到数据压缩的作用;

(3) 关键帧对原始动画具有概括力,可用于三维动画片断的检索.

参 考 文 献

- [1] Bruderlin A, Williams L. Motion signal processing [A]. In: Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Los Angeles, California, 1996. 97~115
- [2] Witkin Andrew, Popovic Zoran. Motion warping [A]. In: Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Los Angeles, California, 1995. 105~108
- [3] Lee J H, Shin S Y. A hierarchical approach to interactive motion editing for human-like figures [A]. In: Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Los Angeles, California, 1999. 39~48
- [4] Witkin A, Kass M. Spacetime constraints [J]. Computer Graphics, 1988, 22(4): 159~168
- [5] Gleicher M. Motion editing with spacetime constraints [A]. In: Proceedings of Symposium on Interactive 3D Graphics, Providence, Rhode Island, 1997. 139~148
- [6] Gleicher M. Retargeting motion to new characters [A]. In: Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Orlando, Florida, 1998. 33~42
- [7] Rose C, Guenter B, Bodenheimer B, *et al.* Efficient generation of motion transition using spacetime constraints [A]. In: Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, New Orleans, Louisiana, 1996. 147~154
- [8] deMenthon D, Kobla V, Doermann David. Video summarization by curve simplification [A]. In: Proceedings of ACM Multimedia'98, Bristol, UK, 1998. 211~218
- [9] Zhao Li, Qi Wei, Li Stan Z, *et al.* Key-frame extraction and shot retrieval using nearest feature line (NFL) [A]. In: Proceedings of the 2000 ACM workshops on Multimedia, Los Angeles, California, 2000. 217~220
- [10] Shoemake K. Animating rotation with quaternion curves [J]. Computer Graphics, 1985, 19(3): 245~254
- [11] Bao Hujun, Jin Xiaogang, Peng Qunsheng. Fundamental Principles for Computer Animation [M]. Hangzhou: Zhejiang University Press, 2000. 196~222(in Chinese)
(鲍虎军, 金小刚, 彭群生. 计算机动画的算法基础[M]. 杭州: 浙江大学出版社, 2000. 196~222)