计算机研究与发展
Journal of Computer Research and Development

ISSN 1000-1239/CN 11-1777/TP
47(8): 1321-1328, 2010

基于关键姿态分析的运动图自动构建

宗 丹[1,2] 李淳芃[1] 夏时洪[1] 王兆其[1]
[1](中国科学院计算技术研究所前瞻研究实验室 北京 100190)
[2](中国科学院研究生院 北京 100049)
(zongdan@ict.ac.cn)

Key-Postures Based Automated Construction of Motion Graph

Zong Dan[1,2], Li Chunpeng[1], Xia Shihong[1], and Wang Zhaoqi[1]
[1](*Advanced Computing Research Laboratory, Institute of Computing Technology, Chinese Academy of Sciences, Beijing* 100190)
[2](*Graduate University of Chinese Academy of Sciences, Beijing* 100049)

Abstract STM graph (snap together motion graph), with high degree of polymerization nodes, is a structured graph to describe the relationship of the motion segments in character animation. Nodes in a motion graph serve as postures and edges between these nodes correspond to motion clips. Each node in an STM graph is connected with multiple edges. Many different approaches have been proposed to construct motion graphs from the existing motion capture data, which gives the user a flexible way to synthesize natural looking motion and control the character. So it becomes a hot topic to construct motion graphs automatically. However, the current methods of constructing STM graph depend largely on the experience and manual manipulations. Focusing on the problems mentioned above, a novel method to create motion graph automatically is proposed in this paper. Dimension reduction and nonparametric density estimation analysis are adopted to extract the key postures from motion capture data. The segments are obtained to construct the motion graph with high degree of polymerization nodes. The method not only improves the accuracy of extraction of key postures and reduces the subjective factor, but also improves the flexibility of controlling the virtual characters. Experiments have been done on taekwondo motion clip with 934 frames and badminton motion clip with 1798 frames. The results show the effectiveness of the method.

Key words character animation; motion graph; key posture; dimension reduction; probability density

摘 要 高结点聚合运动图(snap together motion graph, STM graph)是刻画虚拟角色运动序列关系的一种结构化运动图.其特点是图中每个结点都包含多条与之相连的边,能够实现对虚拟角色的灵活控制.然而现有的高结点聚合运动图构建方法存在手工标注任务繁重、关键姿态提取结果不准确等问题.针对上述问题,提出了一种基于关键姿态分析的运动图自动构建新方法:通过维度约简和非参数密度估计分析样本数据的概率密度,获得一组关键姿态;然后通过分割获得运动片段,最后构建高结点聚合运动图.该方法不仅提高了关键姿态的提取精度,减少了构图过程的主观因素,同时提高了对虚拟角色控制的灵活性.实验结果表明了该方法的有效性.

关键词 角色动画;运动图;关键姿态;维度约简;概率密度

中图法分类号 TP391.41

收稿日期:2009-01-08;**修回日期**:2009-09-02

基金项目:国家自然科学基金项目(60603082,60703019);北京市自然科学基金项目(4062032)

0 引 言

随着运动捕获技术的发展，许多利用真实数据合成虚拟角色运动的方法被提出. 然而，这些数据通常是以无序方式存储在数据库里的，如何重用运动捕获数据成为虚拟人运动合成中的研究热点. Kovar 等人[1]于 2002 年提出运动图的概念. 运动图是一种基于图的运动片段组织方式，本质上是由结点和边组成的有向图，其中的“结点”对应于姿态，而“边”对应于连接姿态的运动片段. 用户通过搜索运动图就能合成出逼真的运动序列.

然而传统的运动图是非结构化的，难以满足交互应用的需求，这主要体现在以下两个方面：首先，非结构化运动图由于没有预先考虑运动的连接关系，无法保证快速地从一个运动到达另一个运动；另外，现有的交互式运动控制方法都是建立在已知运动图结构的基础上的，然而在非结构化运动图中要预先知道所有可能的运动是困难的. Gleicher 等人[2]提出了 STM 运动图(snap together motion graph)的概念，通过构建结构化的运动图实时合成运动，能够满足交互应用的需求. 该类运动图中的每个结点都包含多条与之相连的边，称之为高结点聚合度的运动图. 然而，在构建该类运动图时系统需要用户交互式地指定关键姿态同时手动指定跳转阈值，以此构建运动图. 这种方式依赖于用户的主观经验，导致关键姿态的提取精度不高，计算效率低. Park 等人[3-4]也是采用手工标注的方式从周期运动序列中分割出规整的周期运动数据. 朱登明等人[5]结合 DTW 技术和误差平方和准则对时序运动序列进行自动切分和标注，从而提取出重复的基本动作单元. 但是直接捕获或者手工处理得到规整的周期运动都是费时费力的. Kim 等人[6]针对有节奏的规律运动提出了运动节拍的概念，应用运动规律的时序模式来标记有节奏的运动数据，从而达到对运动数据进行分割的目的. 然而手工的方法需要用户提供大量的参数，操作过程较为繁琐. Lee 等人[7]采用统计学习的方法对运动图进行建模，在数据库底层把一段序列看成简单的一阶 Markov 过程，高层通过统计模型对运动数据库进行聚类分析. 然而在对运动数据进行聚类时需要手工指定起始点，否则无法保证聚类的质量. 运动图构建过程的自动化程度不高，大大限制了运动图的应用范围.

本文提出的运动图自动构建方法是基于关键姿态分析的. 虽然目前也存在一些从原始运动片段中提取关键帧的方法，但是这些方法并不适合构建运动图. 例如 Lim 等人[8]选取一定的运动特征表示姿态，将原始运动捕获数据中每一帧看成高维空间中的一点. 通过顺序连接运动序列中对应的数据点将原始运动视为高维空间中的一段曲线，然后采用曲线简化算法提取分割点作为关键帧. Assa 等人[9]采用复制多维尺度分析(replicated multi-dimensional scaling，RMDS)方法将原始高维姿态空间中的一个运动序列表示成低维空间中的一条曲线，通过寻找低维曲线的极值点得到关键姿态. 但是上述方法关注于运动数据压缩和运动序列的表征，提取的关键姿态不是多个运动片段的交汇点，不适合作为运动图的结点.

基于以上分析，本文提出了一种基于关键姿态分析的运动图自动构建方法. 在对运动捕获数据构成的高维姿态空间进行维度约简的基础上，利用非参数密度估计分析样本数据的概率密度获得关键姿态，通过分割得到运动片段，最后生成运动图以合成虚拟人运动. 该方法的特点是：1. 在概率密度空间中提取关键姿态，提高了关键姿态的提取精度；2. 在此基础上自动分割得到运动片段作为运动图的边，提高了构图过程的自动化程度；3. 生成了一种高结点聚合度的运动图，即图中每个结点都包含多条与之相连的边. 合成运动时能方便地在不同的运动序列之间跳转，提高了虚拟人运动合成的灵活性和可控性.

1 算法思想

在文献[2]中，Gleicher 等人提出 STM 运动图的概念，通过构建结构化的运动图实现运动的在线合成. 首先由用户手动输入关键姿态，然后系统在运动捕获数据中自动找到该姿态的一个匹配集并由此生成运动图的结点，通过运动分割得到运动片段作为运动图的边，进而构建完整的运动图. 在对运动数据进行观察的基础上，我们发现该运动图中结点姿态应当具有以下一些特性：它是运动图中运动片段之间的过渡姿态，即存在多个运动片段以该姿态为起始姿态或终止姿态，例如单脚着地的姿态可以看成从“跑步(run)”到“行走(walk)”的过渡姿态，也可以看成从“踢腿(kick)”到“后撤(backward)”的过渡姿态. 由于关键姿态具有这些特性，因此在姿态空间中的概率密度较大. 另外，当虚拟角色运动到关键姿态时的瞬时速度往往较低，甚至会出现短暂停留的现象. 例如向前踢腿时双脚支撑的姿态可以看作是

关键姿态,而踢到最远点的姿态不应当被视为关键姿态.若将这样的关键姿态提取出来作为运动图的结点,并将以关键姿态为起始姿态和终止姿态的运动片段作为运动图的边,就能得到具有高结点聚合度的运动图.利用这样的运动图合成运动序列,可以方便地在不同的运动之间跳转,大大提高虚拟人运动控制的灵活性和可控性.

算法流程如图1所示.整个处理过程包括关键姿态的提取和运动序列的生成两个阶段.在关键姿态的提取阶段,采用多维尺度分析方法得到高维姿态空间的一个低维描述,采用非参数密度估计得到样本的概率密度分布并获得关键姿态.在运动序列的生成阶段,根据提取的关键姿态找到运动序列的分割点,从而获得更小的运动片段作为运动图的边.然后得到具有高结点聚合度的运动图,从而实现对虚拟角色的灵活控制.

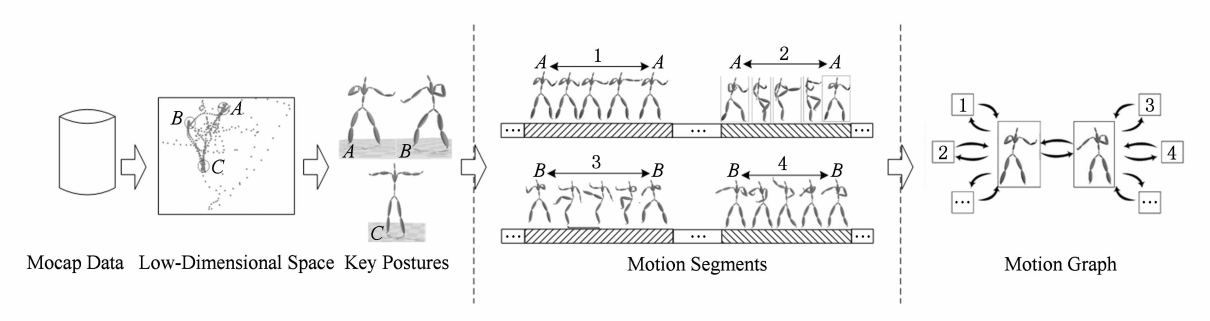

Fig. 1　The algorithm flow chart.

图1　算法流程图

2　姿态空间分析

为了发现高维向量之间的内在联系,需要对高维的姿态空间进行维数约简,从而在低维的特征空间中对运动数据进行分析.文献[10]指出,多维尺度分析方法能够较好地保持样本在原始空间和降维空间中距离关系的一致性,从而获得较好的维度约简效果,故采用多维尺度分析方法对姿态空间进行降维.

2.1　姿态距离的度量

原始的运动捕获数据是用参数向量表示的,但是直接使用向量范数表示样本之间的不相似性无法体现对输入运动作刚性变换的不变性,并且也没有考虑关节的速度、加速度等高阶连续性.为此,采用Lee等人提出的度量运动帧相似性的方法[7],具体公式如下:

$$D_{ij} = d(\boldsymbol{p}_i, \boldsymbol{p}_j) + \lambda \times d(\boldsymbol{v}_i, \boldsymbol{v}_j), \tag{1}$$

式(1)中第1项 $d(\boldsymbol{p}_i, \boldsymbol{p}_j)$ 用于描述姿态的差异,第2项 $d(\boldsymbol{v}_i, \boldsymbol{v}_j)$ 用来描述运动速度之间的差异,加权系数 λ 用于平衡由姿态项和速度项带来的相似性差异.在计算姿态差异时,采用基于点云(point cloud)[1]的方法来度量两个姿态之间的距离.姿态 $\boldsymbol{p}_i$ 对应点集 $\boldsymbol{a}_i$,姿态 $\boldsymbol{p}_j$ 对应点集 $\boldsymbol{b}_i$,姿态 $\boldsymbol{p}_i$ 和 $\boldsymbol{p}_j$ 之间的差异表示为

$$d(\boldsymbol{p}_i, \boldsymbol{p}_j) = \min_{\boldsymbol{R},\boldsymbol{t}} \| \boldsymbol{b}_i - \boldsymbol{R} \times \boldsymbol{a}_i - \boldsymbol{t} \|^2, \tag{2}$$

其中 $\boldsymbol{R}$ 表示对点集 $\boldsymbol{a}_i$ 所作的旋转操作,$\boldsymbol{t}$ 表示对点集 $\boldsymbol{a}_i$ 所作的平移操作.显然姿态相似但速度差异较大的运动片段不能直接拼接,否则会生成不合理的运动.为此将速度项引入到姿态相似性度量中.在描述姿态 $\boldsymbol{p}_i$ 与姿态 $\boldsymbol{p}_j$ 之间的速度差异时,考虑以 $\boldsymbol{p}_i$ 和 $\boldsymbol{p}_j$ 为中心前后 L 帧的窗口内所有帧的姿态差异.这样,式(1)可以表示成如下的形式:

$$D_{ij} = \sum_{k=-L}^{L} \lambda_k \times d(\boldsymbol{p}_{i+k}, \boldsymbol{p}_{j+k}) = \lambda_k \min_{\boldsymbol{R},\boldsymbol{t}} \sum_{k=-L}^{L} \| \boldsymbol{b}_{i+k} - \boldsymbol{R} \times \boldsymbol{a}_{i+k} - \boldsymbol{t} \|^2. \tag{3}$$

2.2　姿态空间的降维

在得到姿态距离矩阵之后,需要对高维的姿态空间进行维度约简.维度约简就是对 D 维空间中的 n 个样本点 $\boldsymbol{x} = \{\boldsymbol{x}_i \mid i = 1, \cdots, n, \boldsymbol{x}_i \in \mathbb{R}^D\}$,通过一定的映射 $f: \mathbb{R}^D \to \mathbb{R}^d (d \ll D)$ 得到样本在 d 维空间中的表示.

为了保持样本在原始空间和降维后空间中距离关系的一致性,采用多维尺度分析(multi-dimensional scaling, MDS)方法进行降维.多维尺度分析方法是一种传统的能够保持数据点之间差异性(或相似性)的维数约简方法.该方法使得原数据集中的点经过变换后,其相对距离关系尽可能保持不变.通过适当定义准则函数来体现在低维空间中对高维距离的重建误差,对准则函数用梯度下降法求解可以推导出解析解[11].经典MDS方法的解析步骤参考文献[12].

在降维过程中，首先计算姿态的相似度矩阵，同时给定维度约简后的维数. 系统自动给出高维姿态空间对应的低维数据，区域中数据的稠密和稀疏与高维观测空间的运动数据分布一致，较好地保持了样本数据在高维空间和低维空间中的一致性.

3 关键姿态的提取

3.1 关键姿态的定义

根据前面的分析，下面给出关键姿态 $\boldsymbol{p}_{\text{key}}$ 的定义：

$$\boldsymbol{p}_{\text{key}} = \underset{\boldsymbol{p}}{\operatorname{argmax}}\{f(\boldsymbol{p}) \mid \boldsymbol{p} \in N(\boldsymbol{p}_0, \delta)\}, \tag{4}$$

其中，$N(\boldsymbol{p}_0, \delta)$表示以 $\boldsymbol{p}_0$ 为中心、δ 为邻域半径的闭区间，$f(\boldsymbol{p})$为低维姿态空间中样本的概率分布密度函数. 由于关键姿态是运动图中多个运动片段的过渡姿态，即存在多个片段以关键姿态为起始或结束姿态，因此在给定区间范围内，关键姿态 $\boldsymbol{p}_{\text{key}}$在样本空间中的概率密度较大.

3.2 关键姿态的提取

关键姿态的提取是多峰值的搜索问题. 传统的局部极值搜索算法往往只能找到一个局部极小值，并且需要显式给出概率密度函数的解析表达式. 然而对于人体运动数据，直接给出其解析表达式是困难的. 均值漂移算法(mean shift)作为一种有效的自适应步长迭代算法，能保证在密度估计的零梯度附近收敛[13]. 该算法在样本概率密度空间中搜索局部极值，与样本数据的维度没有关联，计算效率较高. 均值漂移算法具体做法如下：在概率空间中任意给定一个起始姿态 $\boldsymbol{x}$，计算该样本点一定邻域内均值的偏移量 δ，移动该样本点到其偏移均值；然后以此为新的起始姿态继续移动. 反复执行这一操作，直到达到预先定义次数或误差偏移量小于某一阈值. 这个过程的形式化描述如下：

1. 任意给定一个初始姿态 $\boldsymbol{x}$，核密度估计为 $\hat{f}_n(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{N}\frac{1}{h}K\left(\frac{\boldsymbol{x}-\boldsymbol{X}_i}{h}\right)$，其中核函数选取高斯核 $K(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{x^2}{2}}$，核函数的窗宽为 h.

2. 计算窗口内的均值 $m_h(\boldsymbol{x})$：

$$m_h(\boldsymbol{x}) = \frac{\sum_{i=1}^{n} G\left(\frac{\boldsymbol{x}_i - \boldsymbol{x}}{h}\right) w(\boldsymbol{x}_i)\boldsymbol{x}_i}{\sum_{i=1}^{n} G\left(\frac{\boldsymbol{x}_i - \boldsymbol{x}}{h}\right) w(\boldsymbol{x}_i)}, \tag{5}$$

其中 $w(\boldsymbol{x}_i)(i=1,\cdots,n;\ \sum_{i=1}^{n} w(\boldsymbol{x}_i)=1)$为给定窗宽范围内各个样本点 $\boldsymbol{x}_i$ 的权重.

3. 计算均值漂移量 $\delta = m_h(\boldsymbol{x}) - \boldsymbol{x}$，如果$|\delta| < \varepsilon$($\varepsilon$ 为系统容差)或达到预先定义的次数，结束循环；否则将 $m_h(\boldsymbol{x})$作为新的起始姿态，继续执行步骤 2.

均值漂移算法能够不断地沿着概率密度的梯度方向移动，同时自适应调整步长大小. 移动的步长不仅与梯度大小有关，也与该姿态的概率密度有关. 在满足一定条件下，均值漂移算法一定会收敛到该样本点附近的峰值，具体证明参见文献[14].

值得注意的是，若直接在降维后的样本空间中寻找关键姿态，可能会由于空间中样本数目太少而导致求解结果不正确. 为此需要对原始样本空间进行基于重要性的重采样，产生一定数量的伪样本，从而在稠密的样本空间中寻找关键姿态，增加提取过程的鲁棒性.

4 运动序列的生成

4.1 运动序列的生成

在找到一系列关键姿态之后，得到每一个关键姿态与运动序列中各帧的距离关系. 将原始运动序列中与关键姿态之间的距离小于某给定阈值的帧抽取出来，形成对应关键姿态的匹配集 S. 那么，运动序列就可以看作是按时间先后从一个匹配集 S 进入另一匹配集 S'的过程. 标记开始进入某匹配集的起始帧为 $\boldsymbol{F}_a$，远离该匹配集的结束帧为 $\boldsymbol{F}_b$，并找到 $\boldsymbol{F}_a$ 和 $\boldsymbol{F}_b$ 之间概率密度最大的一点 $\boldsymbol{F}_c$ 为分割点：

$$\boldsymbol{F}_c = \underset{\boldsymbol{F}}{\operatorname{argmax}}\{d(\boldsymbol{p}_{\text{key}}, \boldsymbol{F}) < \varepsilon'\}, \tag{6}$$

其中，匹配集 $S = [\boldsymbol{F}_a, \boldsymbol{F}_b]$. 关键姿态 $\boldsymbol{p}_{\text{key}}$和分割点 $\boldsymbol{F}_c$ 的距离 $d(\boldsymbol{p}_{\text{key}}, \boldsymbol{F}_c)$小于某给定阈值 ε'. 当 $\boldsymbol{F}_a$ 和 $\boldsymbol{F}_b$ 之间的距离大于给定帧数时，找到 $\boldsymbol{F}_a$ 和 $\boldsymbol{F}_b$ 之间概率密度较大的两点 $\boldsymbol{F}_c$ 与 $\boldsymbol{F}_d$. 对 $\boldsymbol{F}_d$ 的定义同上. 由于关键姿态是基于姿态距离关系和概率密度分析得到的，因此 $\boldsymbol{F}_c$ 与 $\boldsymbol{F}_d$ 对应的姿态与对应的关键姿态 $\boldsymbol{p}_{\text{key}}$非常相似. 用同样的方法找到下一个分割点 $\boldsymbol{F}_{c'}$. 不失一般性，假设它与关键姿态 $\boldsymbol{p}'_{\text{key}}$对应，则运动序列$\{\boldsymbol{F}_d, \boldsymbol{F}_{d+1}, \cdots, \boldsymbol{F}_{c'}\}$就是以 $\boldsymbol{p}_{\text{key}}$为起始姿态，以 $\boldsymbol{p}'_{\text{key}}$为末尾姿态的运动片段.

4.2 偏置映射

在构建运动图时，若不加处理地将输出的运动片段作为图的边，将出现不合理的结果. 例如在两个片段的衔接部分出现抖动，从而使合成的运动失真. 为此需要对运动片段两端附近的若干帧数据作平滑操作，这可以通过偏置映射的方法实现.

运动的偏置映射是对一个运动片段的局部作一些修改,尽量保留原始运动的细节以及运动的连续性[15].以改变某运动片段的最后一帧为例,说明偏置映射的工作原理:对于原始运动序列 $C=\{\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n\}$,欲改变该运动片段的最后一帧为新姿态 $\boldsymbol{p}_t$,若直接在最后一帧上变化,将导致运动的不连续.这里我们先对与变化帧相邻的 $L(L<n)$ 帧作简单的线性插值,生成一组偏置映射 $d=\{\boldsymbol{d}_{n-L},\boldsymbol{d}_{n-L+1},\cdots,\boldsymbol{d}_n\}$,其中 $d_i=(\boldsymbol{p}_n-\boldsymbol{p}_t)\times\frac{i-(n-L)}{L}$ $(i=n-L,n-L+1,\cdots,n)$.然后,将 C 与 d_i 逐帧相加得到的新运动序列:$C'=\{\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_{n-L}+\boldsymbol{d}_{n-L},\cdots,\boldsymbol{p}_n+\boldsymbol{d}_n\}$.

新的运动序列 C' 将调整所引起的运动变化同原始运动分离开来,不仅能够满足约束,而且使修改后的运动数据保持一阶连续性.

5　结果分析

实验是在一台 P4 3.0GHz、内存为 1GB 的 PC 机上进行的,编程环境为 Windows/Matlab/Microsoft Visual Studio 2008 环境.实验采用的人体模型具有 24 个骨骼,共 75 个自由度.所有的实验数据都是通过 Vicon 光学运动捕获设备采集,采样频率均为 120 Hz.

我们首先选取了两种类型的运动进行实验,分别是 934 帧的跆拳道运动和 1798 帧的羽毛球运动.其中跆拳道运动包含左踢腿、右踢腿和左右腿的交叉换步等多种动作,羽毛球运动包含正手前场击球、反手前场击球和网前球等.系统在对原始运动数据进行姿态相似性度量之后,将高维姿态空间降维到低维空间,图 2 给出的是降维后的样本空间,其中图 2(a)为跆拳道数据,用 A,B,C 标出了概率密度较大的 3 个区域.图 2(b)为羽毛球数据,也标出了概率密度较大的区域.通过对照查看原始高维姿态空间中的样本和降维后的低维空间中的样本,可知降维结果保持了样本数据在原始空间和降维空间中距离关系的一致性,即降维后区域中样本的稠密和稀疏与高维姿态空间中的数据分布一致.

图 3 以跆拳道运动为例,对比分别采用本文算法、基于曲线简化算法[8]和基于运动摘要算法[9]提取出的关键姿态.其中第 1 行为本文算法提取的结果,得到的关键姿态如图 3(a)所示,依次为 T 型预备姿态、右脚在前的预备姿态和左脚在前的预备姿态.第 2 行为曲线简化算法提取的关键姿态.我们将四肢骨骼与根关节之间的夹角作为运动特征,将运动序列视为空间中的曲线,采用曲线简化算法得到

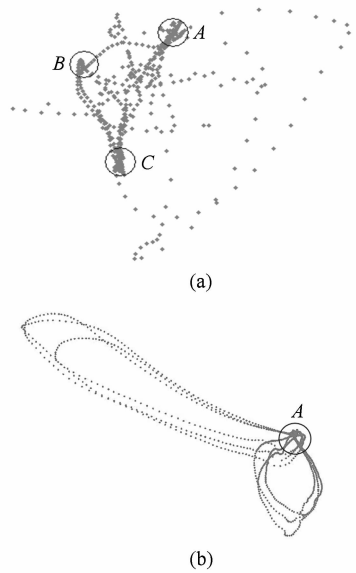

Fig. 2　The posture space after dimension reduction. (a) Taekwondo and (b) Badminton.

图2　降维后的姿态空间.(a) 跆拳道数据;(b) 羽毛球数据

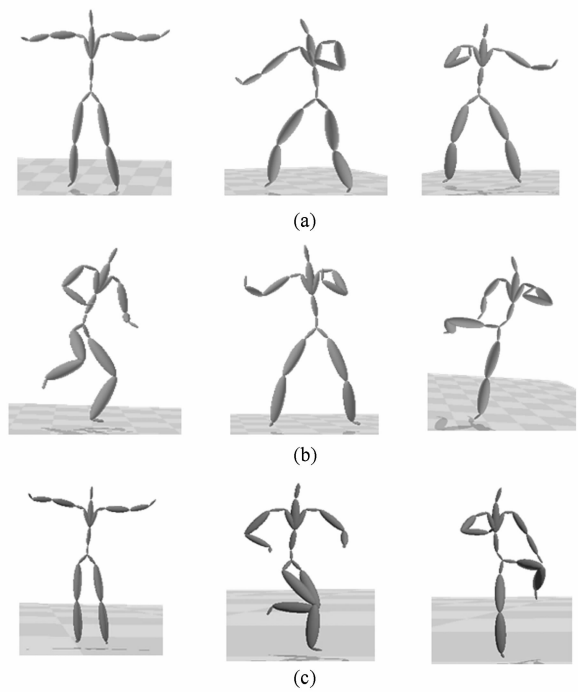

Fig. 3　Key postures. (a) The key postures extracted by our method; (b) The key postures extracted by curve simplification based method; and (c) The key postures extracted by action synopsis based method.

图 3　跆拳道运动的关键姿态提取结果对比图.(a) 本文方法;(b) 曲线简化方法;(c) 运动摘要方法

的关键姿态如图 3(b)所示. 第 3 行为运动摘要算法提取的关键姿态. 这里采用 RMDS 方法将原始高维姿态空间中的一个运动序列表示成低维空间中的一条曲线，然后通过寻找低维曲线的极值点得到关键姿态如图 3(c)所示. 不难看出，利用本文方法提取的关键姿态与跆拳道的实际运动规律相符，具有多个以此为起始和终止姿态的运动片段. 而曲线简化算法和运动摘要算法由于没有考虑样本在空间中的分布，提取出的关键帧关注于运动数据压缩和运动摘要表征，不适合构建运动图.

提取关键姿态之后，我们根据原始运动序列中姿态距离关系和概率密度分布对原始运动序列进行分割. 图 4(a)是一个跆拳道运动的片段分割结果，其中左斜线区域是一个左踢腿片段，右斜线区域是一个右踢腿片段. 同样，图 4(b)给出的是一个羽毛球运动的片段分割结果，其中左斜线区域是一个前场反手接球片段，右斜线区域是一个前场正手接球片段. 可以看到，分割出来的运动片段以提取出来的关键姿态为起始姿态和终止姿态，是独立且有意义的动作单元.

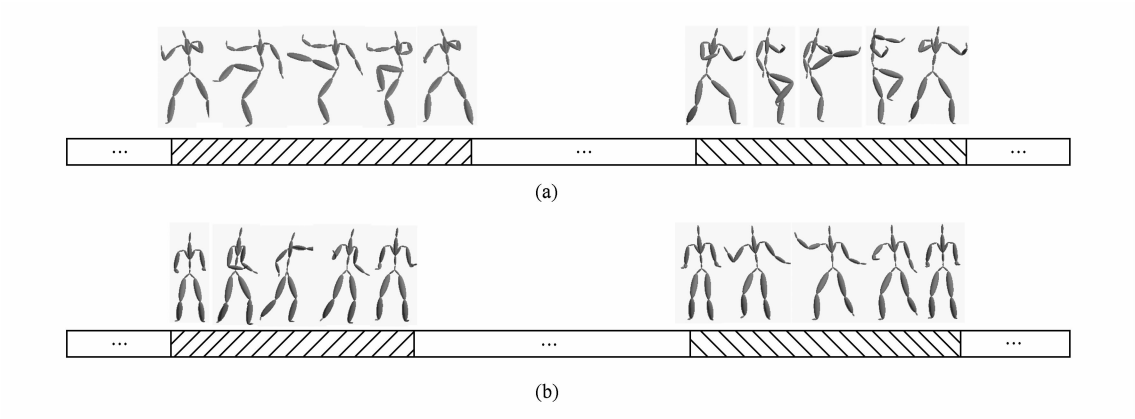

Fig. 4 The partition of motion segments. (a) Taekwondo sequence and (b) Badminton sequence.

图 4 分割后的运动片段示意图. (a)跆拳道运动序列；(b)羽毛球运动序列

为了验证算法的有效性，用一个跆拳道片段对比本文算法、基于曲线简化算法[8]和基于运动摘要算法[9]提取出的关键姿态. 如图 5 所示，其中图 5(a)为本文算法提取出的 3 种关键姿态；图 5(b)为

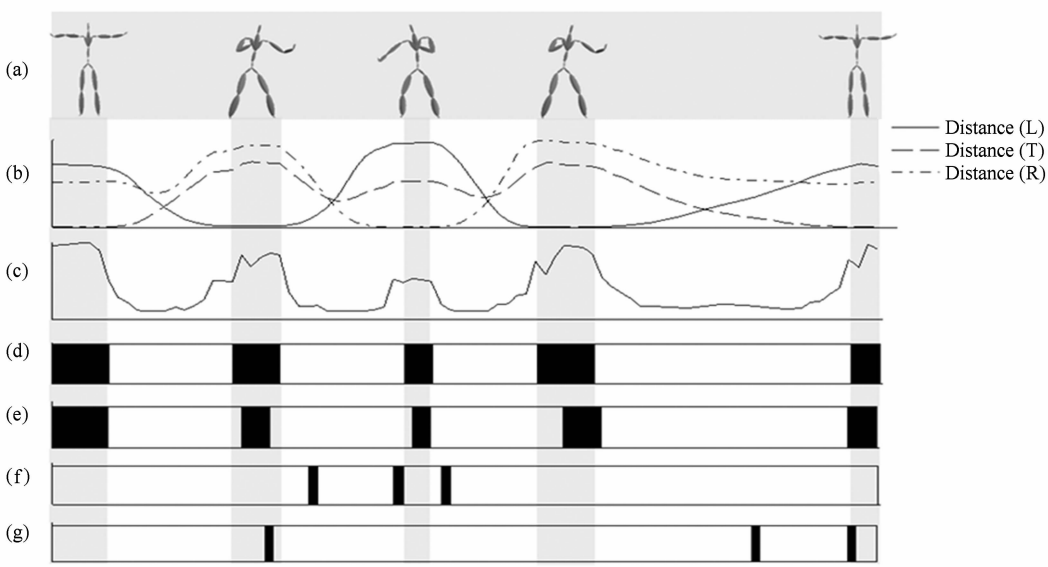

Fig. 5 Comparison diagram. (a) Key postures; (b) The distance between each posture and three key postures respectively; (c) The probability density of each posture; (d) The "ground truth" segments; (e) Segments split by our method; (f) Curve simplification-based method; and (g) Action synopsis-based method.

图 5 跆拳道运动的关键帧对比图. (a) 关键姿态；(b) 各个姿态与 3 个关键姿态的距离；(c) 各个姿态的概率密度；(d) 人工方法的片段分割结果；(e) 本文方法的片段分割结果；(f) 曲线简化算法的关键帧；(g) 运动摘要算法的关键帧

原始运动片段中各帧与 3 个关键姿态的距离;图 5(c)为概率密度空间中各个样本的概率密度;图 5(d)为人工方法得到的片段分割结果;图 5(e)为本文方法得到的片段分割结果;图 5(f)为曲线简化算法得到的关键帧,用黑色线条标出;图 5(g)为运动摘要算法得到的关键帧,也用黑色线条标出.

可以看出,对一些包含以相似姿态为起始和终止的片段的运动序列,本文算法提取的关键姿态的概率比基于曲线简化算法和基于运动摘要算法要大,且用本文方法得到的片段分割结果与人工方法基本一致.这是由于本文算法是在概率密度空间上提取关键姿态的,使得提取出的关键姿态有多个与之相连的运动片段,能较好地反映动作单元的本质;而基于曲线简化的算法和基于运动摘要的算法仅仅考虑了局部姿态的运动特征,提取的关键帧没有多个运动片段与之相连,不适合用于运动图的构建.

6　结论与进一步研究

文中提出了一种自动构建运动图的方法:在对大量运动捕获数据降维的基础上,利用非参数密度估计分析概率密度并自动抽取关键姿态作为运动图的结点,通过运动分割获得运动图的边,进而构建结构化的高结点聚合度运动图.这种运动图的自动构建方法不仅能提高关键姿态的提取精度,而且提高了构建过程中的自动化程度.实验结果也表明了该方法的有效性.

然而,本文提出的运动图的自动构建方法具有一定的局限性.在对姿态空间进行降维的过程中,多维尺度分析方法能保持样本之间的距离关系和度量矩阵的一致性,但是没有考虑运动数据的时序关系.因此,如何在降维的同时考虑样本的空间结构和时序关系,从而有效地揭示出运动数据之间的时空联系,这将作为下一步的研究内容.

参　考　文　献

[1] Kovar L, Gleicher M, Pighin F. Motion graphs [J]. ACM Trans on Graphics, 2002, 21(3): 473-482

[2] Gleicher M, Shin H J, Kovar L, et al. Snap-together motion: Assembling run-time animations [C] //Proc of the Symp on Interactive 3D Graphics. New York: ACM, 2003: 702

[3] Park S I, Shin H J, Shin S Y. On line locomotion generation based on motion blending [C] //Proc of ACM SIGGRAPH/Euro Graphics Symp on Computer Animation. New York: ACM, 2002: 105-111

[4] Park S I, Shin H J, Kim T H, et al. On line motion blending for real time locomotion generation [J]. Computer Animation and Virtual World, 2004, 15(3): 125-138

[5] Zhu Dengming, Wang Zhaoqi. Human animation synthesis based on primitive movement [J]. Journal of Computer Research and Development, 2009, 46(4): 610-617 (in Chinese)
(朱登明,王兆其. 基于动作单元分析的人体动画合成方法研究[J]. 计算机研究与发展, 2009, 46(4): 610-617)

[6] Kim T H, Park S I, Shin S Y. Rhythmic motion synthesis based on motion-beat analysis [J]. ACM Trans on Graphics, 2003, 22(3): 392-401

[7] Lee J, Chai J X, Reitsma P, et al. Interactive control of Avatars animated with human motion data [J]. ACM Trans on Graphics, 2002, 21(3): 491-500

[8] Lim I S, Thalmann D. Key-posture extraction out of human motion data by curve simplification [C] //Proc of the 23rd Annual Int Conf of the IEEE Engineering in Medicine and Biology Society. New York: Springer, 2001: 1167-1169

[9] Assa J, Caspi Y, Cohen-Or D. Action synopsis: Pose selection and illustration [C] //Proc of ACM SIGGRAPH. New York: ACM, 2005: 667-676

[10] Shin H J, Lee J. Motion synthesis and editing in low-dimensional spaces [J]. Computer Animation and Virtual Worlds, 2006, 17(3/4): 219-227

[11] Kruskal J B, Wish M. Multi-Dimensional Scaling [M]. Beverly Hills: SAGE Publications, 1978

[12] Borg I, Groenen P. Modern Multi-Dimensional Scaling: Theory and Application [M]. New York: Springer, 1997

[13] Fukunaga K, Hostetler L. The estimation of the gradient of a density function, with applications in pattern recognition [J]. IEEE Trans on Information Theory, 1975, 21(4): 32-40

[14] Comaniciu D, Ramesh V, Meer P. Real time tracking of non-rigid objects using mean shift [C] //Proc of Computer Vision and Pattern Recognition. Los Alamitos, CA: IEEE Computer Society, 2000: 142-149

[15] Bruderlin A, Williams L. Motion signal processing [C] //Proc of the 22nd Annual Conf on Computer Graphics and Interactive Techniques. New York: ACM, 1995: 97-104

Zong Dan, born in 1984. PhD candidate. Her main research interests include human motion analysis and synthesis.

宗　丹,1984 年生,博士研究生,主要研究方向为虚拟人运动合成与控制.

Li Chunpeng, born in 1980. PhD and assistant professor. Member of China Computer Federation. His main research interests include human motion analysis and synthesis.

李淳芃，1980 年生，博士，助理研究员，中国计算机学会会员，主要研究方向为虚拟人运动合成与控制(cpli@ict.ac.cn).

Xia Shihong, born in 1974. PhD and associate professor. Senior member of China Computer Federation. His main research interests include motion capture, character animation, and human motion analysis and crowd simulation.

夏时洪，1974 年生，博士，副研究员，中国计算机学会高级会员，主要研究方向为运动捕获、角色动画、人体运动分析、群体行为仿真等(xsh@ict.ac.cn).

Wang Zhaoqi, born in 1966. Professor and PhD supervisor of the Institute of Computing Technology, the Chinese Academy of Sciences. Senior member of China Computer Federation. His main research interests include virtual reality and intelligent human computer interaction.

王兆其，1966 年生，研究员，博士生导师，中国计算机学会高级会员，CAD&CG 专委会委员，计算机动画与数字娱乐专委会委员，主要研究方向为虚拟现实与智能人机交互技术(zqwang@ict.ac.cn).

Research Background

Recently, synthesizing natural-looking human motions is a hot topic in the area of character animation. The flexible control of virtual human motion is an important problem, since there are so many applications of character animation including education, entertainment, training and so on.

Motion graph is an effective way to control the character, which has already been widely used. With the rapid development of motion capture technique, human motion data can be easily obtained, while constructing motion graph from the data is a difficult matter. In this paper, a method to construct motion graph automatically is proposed. The method is based on dimension reduction and probability density analysis. Key postures and segments are obtained automatically to construct the motion graph with high degree of polymerization nodes. It has been proved that the method not only improves the accuracy of extraction of key postures, but also improves the flexibility of controlling the virtual characters. It can be applied to many kinds of motion.

This work is supported by the National Natural Science Foundation of China (No. 60603082 and No. 60703019) and the Beijing Natural Science Foundation (No. 4062032).