

# 基于逻辑相似性的运动检索算法

肖秦汉, 李言俊, 罗艺闯

(西北工业大学航天学院, 陕西西安, 710072)

**摘要** 针对基于内容的人体运动检索问题, 提出了一种基于逻辑相似性的运动数据检索方法。首先, 针对特定人体运动序列数据, 使用自组织网络映射算法对运动数据进行预处理, 找出运动数据帧的最佳匹配单元。然后, 应用数据编码规则对匹配单元进行量化及编码处理, 获得基于符号化表示的运动描述符。基于符号化表示的描述符建立运动序列索引机制, 应用多分辨率框架进行运动模式匹配。最后, 按照模式匹配结果进行运动相似性排序, 输出检索结果。在 CMU 运动数据库中的试验结果表明了算法有效性。

**关键词** 基于内容检索; 人体运动; 运动编码; 模式匹配

**DOI** 10.3969/j.issn.1009-3516.2014.03.020

**中图分类号** TP391 **文献标志码** A **文章编号** 1009-3516(2014)03-0085-04

## A Human Motion Retrieval Algorithm Based on Logical Similarity

XIAO Qin-han, LI Yan-jun, LUO Yi-chuang

(School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract:** A novel logical-based similarity motion retrieval method is proposed for meeting content-based motion retrieval. Firstly, the self-organizing mapping (SOM) algorithm is used to find the best matching units (BMUs) in accordance with the order data specially designed for human body motion, and then, the encoding method is implemented to establish the motion sequences descriptor based on the symbolic aggregate approximation algorithm. Further more, the SAX-based descriptor is used to build the index mechanism and matching. Finally, the proposed method is used to retrieve motion data. The experiment results show that the proposed algorithm is effective.

**Key words:** content-based retrieval; human body motion; motion encoding; pattern matching

近年来, 运动捕获技术得到极大改善, 并建立了大量的三维捕获运动数据库, 这些捕获数据广泛用于计算机动画、游戏、特效制作以及医疗等领域<sup>[1-4]</sup>, 因此, 对数据库中的捕获数据进行有效的检索成为亟待解决的问题。然而, 在进行三维运动检索的时候, 需要解决下面的问题<sup>[5]</sup>:

1) 人体运动数据具有很高的结构复杂度。一个运动序列是由每个关节的时间序列组合而成的, 我们需要抽取其中主要的特性信息来表示整个人体运动序列;

2) 人体运动数据具有很高的空间和时间复杂度, 如果不能有效地对维数进行降维, 将会带来高维

收稿日期: 2013-04-01

基金项目: 国家自然科学基金资助项目(61271362)

作者简介: 肖秦汉(1972-), 男, 湖北汉阳人, 工程师, 博士生, 主要从事图像识别与检索等研究。E-mail: 568441925@qq.com

**引用格式:** 肖秦汉, 李言俊, 罗艺闯. 基于逻辑相似性的运动检索算法[J]. 空军工程大学学报: 自然科学版, 2014, 15(3): 85-88. XIAO Qinhan, LI Yanjun, LUO Yichuang. A human motion retrieval algorithm based on logical similarity[J]. Journal of air force engineering university: natural science edition, 2014, 15(3): 85-88.

度灾难,而且,根据中心极限定理这些数据之间的距离将趋向一样,这样就无法对他们进行区分;

3)在进行运动序列之间的相似性比较时,不仅要考虑空间相似性,也要考虑时间相似性。

为了解决这些问题,Liu Feng 和 Yue-Ting Zhang<sup>[6]</sup>建立起基于层级运动描述的运动索引结构,然后使用最近邻搜索算法来确定检索子集,进而使用弹性匹配算法来计算运动实例和检索子集之间的相似性;Xiang Jian<sup>[7]</sup>提出了基于三维转换规则的运动检索方法来分别处理关节的局部时空特征,这样就避免直接使用原始数据,取得了很好的检索效果,但是所提取的空间特征不能够消除帧长对算法的影响;随着基于内容的检索的算法的提出,动态时间规整<sup>[8]</sup>(Dynamic Time Warping, DTW)被用于人体捕获数据的检索和匹配中,Keogh<sup>[9]</sup>在 DTW 的基础上提出了带边界的动态时间规整方法,极大地加速了匹配速度;Muller<sup>[10]</sup>提出了一种特征提取算法,通过计算关节位置的几何信息来描述运动,得到较好的结果,然而这种方法需要用户根据不同的运动类型选取不同的特征;Chi<sup>[11]</sup>使用自组织网络映射(Self-Organizing Mapping, SOM)来聚类每个关节的旋转角度,检索时首先匹配开始帧和结束帧得到候选序列的集合,进一步使用 DTW 来进行精确匹配得到最终的结果,然而这种方法只能检索到相似的运动的子序列,不能够检索出由许多运动方式组成的复杂运动;Sakamoto<sup>[12]</sup>提出了运动映射的方法,其算法也使用了 SOM 对运动数据进行索引和检索,并用 SOM 节点来表示运动片段来进行检索,但该算法是基于节点的,只能够在节点和节点之间进行比较,缺少对运动在抽象层的表示。

本文在对运动相似性研究的基础上,提出了一种对逻辑相似性运动的检索方法。

## 1 算法流程

文中提出的检索方法首先使用 SOM 算法将高维数据映射到低维空间中,由于 SOM 可以保持拓扑结构的不变性,因此保存着大部分的运动信息,接着使用近似符号算法(Symbolic Aggregate Approximation, SAX)来提取运动特征并用 SAX 字符串来表示,然后利用 SAX 的多分辨率特性对运动数据进行索引,并提取其中的运动的模式来进行检索。整个算法见图 1。

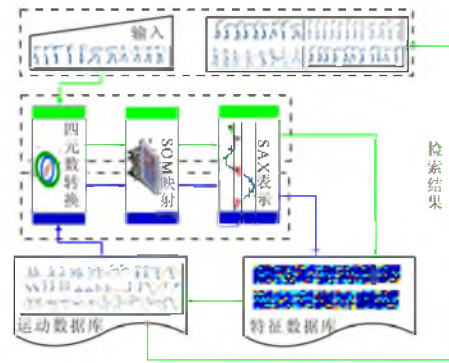


图 1 整体算法框图

Fig. 1 The diagram of whole algorithm

## 2 运动相似性

本文所用数据来自卡梅隆大学运动捕获数据库<sup>[13]</sup>,所用文件格式为 BVH;并使用骨架模型来表示人体结构,我们使用其中的 21 个关节。为了减小计算复杂度同时保持人体骨架的刚性,并去除角色骨架大小的影响,我们将使用四元数来提取初始运动数据特征;如一个  $m$  帧运动序列,其中的一帧可表示为:

$$\mathbf{x}(t) = \{ \mathbf{p}(t), \mathbf{q}_0(t), \mathbf{q}_1(t), \dots, \mathbf{q}_N(t) \},$$

$$t = 1, 2, \dots, m \quad (1)$$

式中:  $\mathbf{p}(t)$  为根节点平移量;  $\mathbf{q}_0(t)$  为根节点朝向;  $\mathbf{q}_i(t)$  为关节对应父节点的四元数表示。因此,每一帧运动可由一个 84 维向量表示。

基于内容的数据检索中,设计合适的相似性度量十分关键。如走和慢跑动作的比较,见图 2,其中左边表示带转弯的慢跑动作,右边表示带转弯的走动动作。初始姿势轨迹很相似,但若观察细节,便发现两运动速度、相对时间、3D 轨迹均不相同。由于方向、速度、方式、和绝对尺度等的不同造成运动差异巨大,因此,逻辑相似的运动不一定是数值相似的。而本文提出的方法将结合逻辑相似及数据相似性,进一步挖掘运动的相关性,寻找二者相关的逻辑相似。

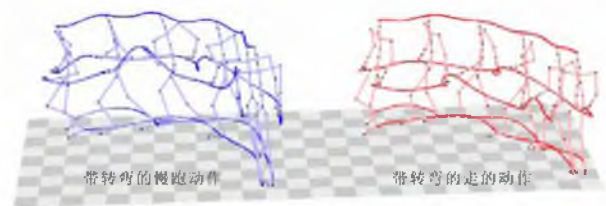


图 2 初始姿态相似的动作比较

Fig. 2 Comparison of initial posture similar action

### 3 运动描述符的建立

#### 3.1 组织网络映射(SOM)

自组织网络映射是一种竞争型神经网络,并使用无监督学习算法来可视化高维数据,它可以将高维数据之间复杂关系转换为低维空间中的几何关系,同时保持原始数据之间的拓扑结构,在这里我们用来对运动数据进行降维处理,以便在低维空间中对数据进行分析。整个算法流程见图 3,其中  $\omega$  表示权值,  $\alpha(t)$  表示学习速率;将运动数据的四元数表示用 SOM 算法进行映射,并不断调整权值和学习速率,直到满足一定的条件( $\alpha t \leq 0.005$ )时,终止整个训练过程,并用得到的最佳匹配单元(Best Matching Units, BMUs)来表示整个运动序列。

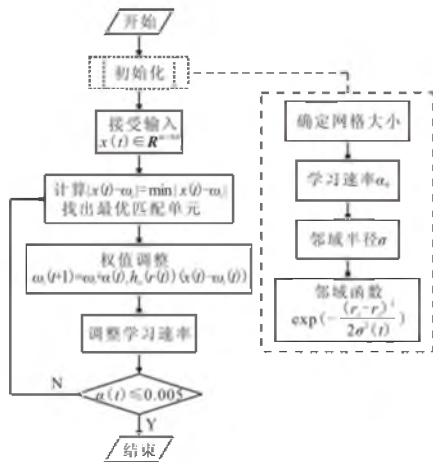


图 3 SOM 算法流程图

Fig. 3 Algorithm flow chart of SOM

#### 3.2 近似符号法(SAX)

由于 SOM 的聚类特性使得所得到的 BMUs 具有很大的离散性,为了使检索得到的序列具有逻辑相似性,我们使用 SAX 对数据进行符号化表示来加入抽象层;由 Liu<sup>[6]</sup>所提出的 SAX 算法广泛用于时间序列处理和数据挖掘领域,并具有很好的索引特性,可以对较大的数据库进行索引,再者其本身还具有多分辨率特性,可以很好地提取数据库中所隐含的模式特征。SAX 的原理为首先在高斯空间中对时间序列进行划分(基于 PAA 算法),然后选取合适的基底和字长来进行表示,对图中的时间序列 T 分别用不同的基底(4 和 2)和相同的字长 4 来进行表示,得到的 SAX 字符串分别为:

$$SAX(T, 4, 4) = T^1 = \{11, 11, 01, 00\}$$

$$SAX(T, 4, 2) = T^1 = \{1, 1, 0, 0\}$$

#### 3.3 多分辨率模式提取

多分辨率运动模式识别的主要思想是将一个用低分辨率表示的字符串扩展到高分辨率。在高分辨

率下,只有子序列非常相似的序列才能聚类,故分辨率越高,聚类越困难。这里我们使用 MrMotif 算法<sup>[15]</sup>来进行运动模式的提取,设置最小的分辨率  $g_{min} = 2$ ,最大的分辨率  $g_{max} = 64$ 。在试验过程中只选取几个特定的分辨率进行分析,并对于所选取的分辨率,找出前 K 个模式。

通过前面的论述可知,如果每一维基底的每一次增加都假设迭代次数倍增,更高分辨率的 SAX 字符串表示层级嵌套在低分辨率的 SAX 表示中;具体来说,在给定基底情况下,随着迭代数倍增,基底增加和接下来的子序列相一致,这时基于最大扇出率就可以建立起一个层级索引系统,并且不会包含重叠的区域<sup>[16]</sup>。

#### 3.4 运动检索

对任一带查询序列,首先将其转换为四元数表示并使用 SOM 进行映射,接着使用 SAX 进行符号化表示,将所得的字符串表示和所提取的模式进行比较,取 K 个模式中和待查询序列距离较小的模式作为候选序列的集合,并在该模式中进行进一步比较,并根据相似度大小并返回相应的运动序列作为检索结果;所提方法不但可以检索单类型的运动序列,也可用于检索由多个运动方式组成的混合运动序列。

### 4 试验结果及分析

使用 CMU 数据库并在 Matlab2010a 环境下进行算法的验证;首先验证不同基底下的 SAX 表示对于聚类精度的影响,见图 4。

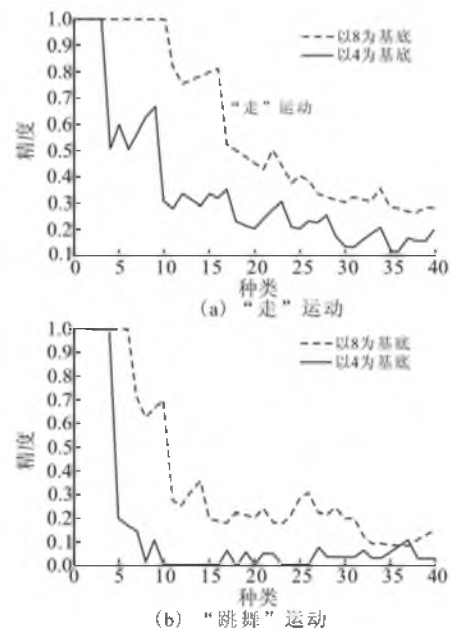


图 4 SAX 符号化时不同基底对于聚类精度的影响

Fig. 4 The influence of SAX based on different clustering accuracy

图4(a)以“走”运动序列为例,分别以8和4为基底,随着返回种类的增加均呈下降趋势,但明显以8为基底的情况下,聚类精度整体较4的情况下高,又如图4(b)的“跳舞”运动,表现出同样的情况。

对“走”运动序列进行检索,首先比较不同基底情况下的PR曲线,可以看出以8为基底的SAX表示较基底4有较好的检索结果,但随着返回数量的增加,两者不相上下。

## 5 结语

为了检索到逻辑相似性的运动数据,本文提出了一种新的检索方法。以往运动数据描述符,由于占用字节高,大大的影响了检索效率,本文针对这一突出问题,首先使用运动序列来训练SOM网络,从而将数据映射到低维空间中进行处理,并用训练得到的最佳匹配单元(Best Matching Units, BMUs)来表示运动序列,从而大大减少了运动描述符的字节数。接着使用SAX算法对所得时间序列进行符号化表示,以在抽象层对数据进行表示,然后,使用SAX的多分辨率特性进行运动模式的提取,来提高检索效率。基于SOM及SAX的运动描述符,在检索效率及存储两方面达到了双重优化的效果。

### 参考文献(References):

- [1] Liu S J, Tang S W, Yang D Q, et al. Technology and integration of Web information extraction based on XML [C]//Proc on 2nd conf on the tenth national database, 2003: 453-457.
- [2] Crescenzi V, Mecca G. Grammars have Exceptions. [J]. Information systems, 1998, 23(8): 539-565.
- [3] Vondrak Marek, Sigal Leonid, Jenkins Odest Chadwicke. Dynamical simulation priors for human motion tracking[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(1): 52-65.
- [4] Zhou Xiaowei, Yang Can, Yu Weichuan. Moving object detection by detecting contiguous outliers in the low-rank representation[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(3): 597-610.
- [5] Xiang J. An efficient three-dimensional motion retrieval method.[J]. Computer science, 2008, 34(3): 84-86.
- [6] Liu F, Zhuang Y T, Wu F, et al. Three-dimensional motion retrieval based on examples computer [J]. Aided design and computer graphics, 2003, 15(10): 1275-1280.
- [7] Xiang J, Zhu HL. Three-dimensional feature extraction based on data analysis and retrieval of human motion.[J]. Computer applications, 2008, 28(5): 1344-1346.
- [8] Kovar L, Gleicher M. Automated extraction and parameterization of motions in large data sets.[J]. ACM transactions on graphics, 2004, 23(3): 559-568.
- [9] Keogh, E, Palpanas T, Zordan V, et al. Indexing large human motion databases. [J]. Proceedings of the 30th VLDB conference, 2004, 30: 780-791.
- [10] Muller M, Roder T, Clausen M. Efficient content-based retrieval of motion capture data [J]. ACM transactions on graphics, 2005, 24(3): 677-685.
- [11] Chiu C, Chao S, Wu M, et al. Content-based retrieval for human motion data.[J]. Journal of visual communication and image representation, 2004, 15(3): 446-466.
- [12] Sakamoto Y, Kuriyama S, Kaneko T. Motion map: image-based retrieval and segmentation of motion data.[C]//Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation. Aire-la-ville, Switzerland: Eurographics association, 2004: 259-266.
- [13] CMU. Carnegie-Mellon MoCap Database [EB/OL]. (2003-03-01) [2013-03-30]. <http://mocap.cs.cmu.edu>, 2003.
- [14] Lin J, Keogh E, Lonardi S, et al. A symbolic representation of time series, with implications for streaming algorithms. In proceedings of the eighth [C]//Proceedings of the 8th ACM SIGMOD workshop on research issues in data mining and knowledge discovery. New York: ACM press, 2003: 2-11.
- [15] Nuno C, Paulo A. Multiresolution motif discovery in times series [C]//Proceedings of the SIAM international conference on data mining. Ohio: [s. n.], 2010: 665-676.
- [16] Lin J, Keogh E, Lonardi S, et al. Finding motifs in time series [C]//Proceedings of the 2nd workshop on temporal data mining. Edmonton, Alberta, Canada: ACM press, 2002: 53-68.

(编辑:徐楠楠)