

基于四元数描述和 EMD 的人体运动捕获数据检索

肖秦琨¹, 李俊芳¹, 肖秦汉²

(1. 西安工业大学 电子信息工程学院, 陕西 西安 710021;

2. 中国电子科技集团公司第20研究所军代室, 陕西 西安 710021)

摘要:针对运动捕获数据的高效匹配问题,提出了一种新的基于四元数描述和 EMD(Earth Mover's Distance)的人体运动检索算法。该算法主要包括特征提取和运动匹配两部分。在特征提取部分,为了解决高维数据检索效率低的问题,引入了四元数描述符对关节点的数据信息特征进行描述,通过映射姿态分布的原始数据,并采取 K-means 聚类方法对待查询动作和运动数据库的特征数据进行降维并归类。在运动匹配部分,根据聚类结果,建立每个特征数据集的距离矩阵,将匹配问题转换为运输优化问题。然后,用 EMD 算法度量待查询动作和数据库动作之间的相似值。仿真实验结果证明了提出的算法是有效的。

关键词:运动捕获;四元数;K-means 聚类;EMD

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2014)03-0090-04

doi:10.3969/j.issn.1673-629X.2014.03.023

Human Motion Capture Data Retrieval Based on Quaternion and EMD

XIAO Qin-kun¹, LI Jun-fang¹, XIAO Qin-han²

(1. College of Electronic Information and Engineering, Xi'an Technological University, Xi'an 710021, China;

2. Air Force Military Representative Office of 20th Research Institute of CETC, Xi'an 710021, China)

Abstract: A novel retrieval approach has been put forward based on quaternion and EMD (Earth Mover's Distance) for efficient matching motion capture data. The method mainly contains two steps: feature extracting and motion matching. In feature extracting part, for solving the problem of low efficiency retrieval of high-dimensional data, introduce the quaternion to represent key-joints rotation data information, through the mapping original data of distribution, take K-means clustering methods to reduce the dimension and classify for the feature data in database of query motion and movement. In motion matching part, according to the clustering results, the distance matrix of each feature dataset was established, converting the matching problem of motion capture data into a optimization transportation problem. Then, the EMD measure algorithm was employed to query the similar value between query motions and database motions. Experiment results show that the proposed approach is efficient.

Key words: motion capture; quaternion; K-means clustering; EMD

0 引言

针对运动捕获技术的快速发展,人类运动序列的分析和识别在计算机动画^[1]、虚拟训练^[2]领域变的及其重要。众所周知原始运动数据具有很高的维度,很难描述运动特征,而且运动特征的表示对于检索却有着至关重要的作用。

大多采用三维坐标表示、关键帧、几何特征、场景描述等方法。文献[3]提出描述各关节点之间三维时

空关系的语义特征,能很好地描述运动内部的几何结构,但原始数据存储空间很大且不能描述骨架的朝向,比如“前进”还是“倒退”。为了采用运动数据集,文献[4]提出一种用骨骼夹角对原始运动数据的特征表示并提取运动片段的关键帧数据集,并建立距离矩阵进行运动序列的相似度匹配,解决了“数值上”相似的运动数据在“逻辑上”不一定相似的问题,不足就是针对少量关键帧进行的,当随着新数据的频繁加入,该索

收稿日期:2013-05-07

修回日期:2013-08-14

网络出版时间:2014-01-07

基金项目:国家自然科学基金资助项目(60972095,61271362);陕西省教育自然科学基金专项(12JK0510,12JK0727);陕西省自然科学基金基础研究计划(2012JM8028)

作者简介:肖秦琨(1973-),男,教授,研究方向为模式识别、计算机动画、多媒体分析与处理;李俊芳(1987-),女,硕士研究生,研究方向为计算机动画、图像检索。

网络出版地址:http://www.cnki.net/kcms/detail/61.1450.TP.20140107.1522.020.html

引的检索效率就会下降。文献[5]描述了一种定性的几何特征提取算法,解决了“逻辑上”相似的运动未必在“数值上”相似的问题,但其缺点是对于每个查询用户必须选择合适的特征才能获得高质量的检索结果。随着基于内容的检索技术研究,动态时间规整^[6-7] (Dynamic Time Warping, DTW) 被用于人体捕获数据匹配中,加速了匹配速度。然而相似的动作可能在很多方面不同,如长度、局部偏移、局部及全局缩放,大多数方法使用欧氏距离和 DTW 来解决匹配问题,但往往不能满足人类的需求。

鉴于此,文中首先用四元数描述运动数据库的骨架数据信息,接着基于 EMD 度量算法在全局范围内进行相似性匹配,将模型匹配问题转化为优化运输问题。

1 相关原理介绍

文中提出一个相对高效的运动检索方法,首先将 CMU 数据库骨架关节的旋转数据信息用四元数描述,以对数据进行抽象表示,然后将所提取特征进行 K-means 聚类,将相同的对象进行归类并建立距离矩阵,最后用 EMD 算法度量待查询动作特征集和 CMU 运动数据特征集的相似值。

1.1 四元数描述

采用卡内梅陇基大学 (CMU^[8]) 数据库的 BVH (BioVision Hierarchy) 格式文件,实验采用 21 个骨架关节代表人体模型。人体运动捕获数据看作是由离散的时间点采样得到的人体运动姿势序列,把每个采样点看作是一帧,这里每一帧的姿势采用 21 个关节信息共同决定,所以,在任意时刻 i , 人体运动的姿势都可以表示为:

$$F_i = (p_i^{(1)}, r_i^{(1)}, r_i^{(2)}, \dots, r_i^{(21)}) \quad (1)$$

这里 $p_i^{(1)} \in R^3$ 和 $r_i^{(1)} \in R^3$ 分别表示 root 关节的位置和方向,也就是平移量和角度旋转量; $r_i^{(j)} \in R^3$, $j=2, \dots, 21$, 表示非 root 关节的方向 (即角度旋转量)。

对动作检索技术来说,为了很好地将动作复杂度和效果表现出来,文中采用四元数^[9] 记录关节的旋转数据信息。在动作数据里每个关节旋转对应的四元数都是单位四元数,从而保证了动作数据对骨架结构的变换是刚性的 (rigid)。四元数公式:

$$q = (w, v) = w + xi + yj + zk \quad (2)$$

其中 w 为实部, v 为虚部。 i, j, k 满足 $i^2 = j^2 = k^2 = -1$, $i * j = j * k = k * i = -1$, 如果要保证变换是线性的, 即不改变向量的长度, 要求 q 为单位四元数, 即 q 满足:

$$\|q\| = \text{Norm}(q) = \sqrt{w^2 + x^2 + y^2 + z^2} = 1 \quad (3)$$

紧接着对所得大型特征数据集进行 K-means 聚类, 其比层次聚类算法快得多^[10]。

1.2 EMD 相似性度量

由于人类运动固有的变异性, 欧氏距离的局限性对时间轴扭曲非常敏感, 动态时间规整只能局部调整时间轴。而人类动作在模型上十分相似, 但是他们之间存在局部偏移或全局缩放。将匹配动作模型转化为运输问题, 即计算一个动作变为另一个动作所需的最小能量。一个有效计算这种转换的方法称作 EMD。EMD 已经成功用于图像^[11] 及视频检索^[12], 文中将进行时间轴的全局调整, 采用 EMD 匹配方式进行相似性度量。EMD 是基于分布的距离度量方法, 作为度量概率数据相似性的标准, 具有抗噪好, 对概率分布间的微小偏移不敏感等优良特性。

计算 EMD 基于优化运输问题^[13], 可以用下面线性规划问题描述, 第一个特征具有 m 类, $P = \{(p_1, w_{p1}), \dots, (p_m, w_{pm})\}$, 这里 p_i 是聚类代表, w_{p_i} 是聚类代表的权值; 第二个特征具有 n 类, $Q = \{(q_1, w_{q1}), \dots, (q_n, w_{qn})\}$, 这里 q_j 是聚类代表, w_{q_j} 是聚类代表的权值, $D = [d_{ij}]$ 是 p_i 和 q_j 的地面距离矩阵。基于上述, EMD 计算公式如 (4) 和 (5)。

$$\text{EMD}(P, Q) = \frac{\sum_{i=1}^m \sum_{j=1}^n f_{ij} d_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}} \quad (4)$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min \left(\sum_{i=1}^m w_{p_i}, \sum_{j=1}^n w_{q_j} \right) \quad (5)$$

2 实验检索结果

文中采用 3 种相似度度量算法包括欧氏距离、DTW、EMD。这里选择一个普通的右腿向前“跑”动作为待查询动作, 如图 1。采取 3 种不同的度量算法得到的结果, 如图 2 ~ 图 4, 可以看到系统分别返回了 9 种不同风格的相似运动。被检索出的相似运动数据按照相似度从小到大、从左到右排列。参考图片下面的数值, 由此看出, 数值越小检索结果越相似, “+”号后面的值是运动数据库里的类别号。从实验结果可以看出, 采用文中方法不仅能够检索出在“数值”上比较相似的运动, 而且能够检索出“逻辑”上比较相似的运动数据。明显的是, EMD 方法检索效率明显提高。因此, 文中方法较前两种与待查询动作更为接近, 可以检索到多种视角下的动作, 提高了查准率。

分析实验结果, 图 1 作为给定的输入, 第一个运动和待查询运动相近, 数值 0 代表待查询动作和最相似动作的相似值, 数值 11 是最相似运动在数据库的编

号。同理,其他 8 个相似度度量值从小到大,相似度度量值越小,两个动作越相近。从总体结果来看,与其他两种算法相比,EMD 算法有更高的相似度。

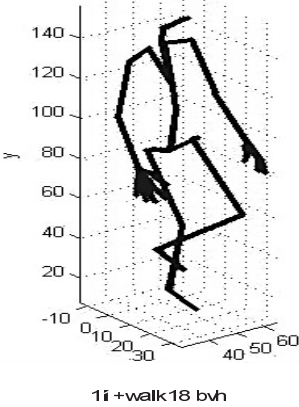


图 1 待检索动作

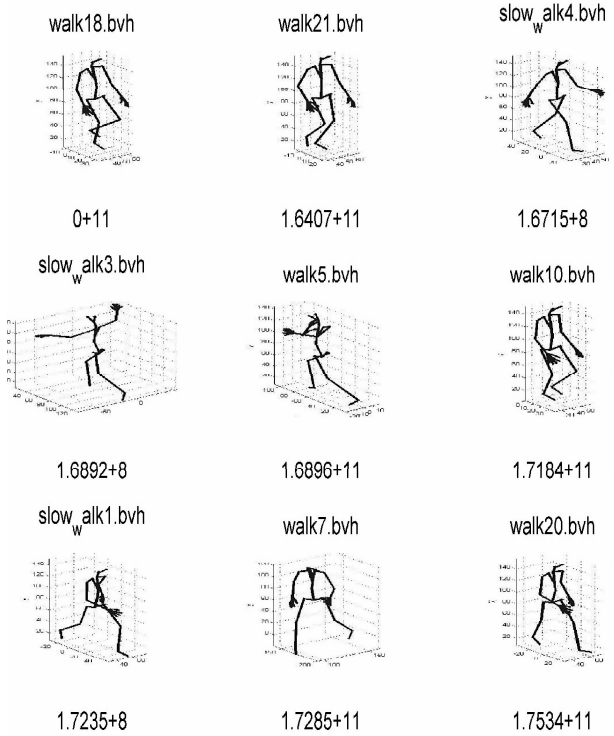


图 2 欧氏距离检索结果

3 实验结果分析

文中方法采用 Matlab2012a 软件验证一个基于内容的运动数据检索系统,实验基于同样的数据库,表 1 对比了三种度量算法的运行时间及精度,很明显 EMD 算法在两个方面都优于欧氏距离和 DTW 算法。

表 1 三种度量算法的运行时间和精度

算法	运行时间/s	精度
欧氏距离	0.338 243	0.700 0
DTW	0.349 399	0.633 3
EMD	0.314 796	0.600 0

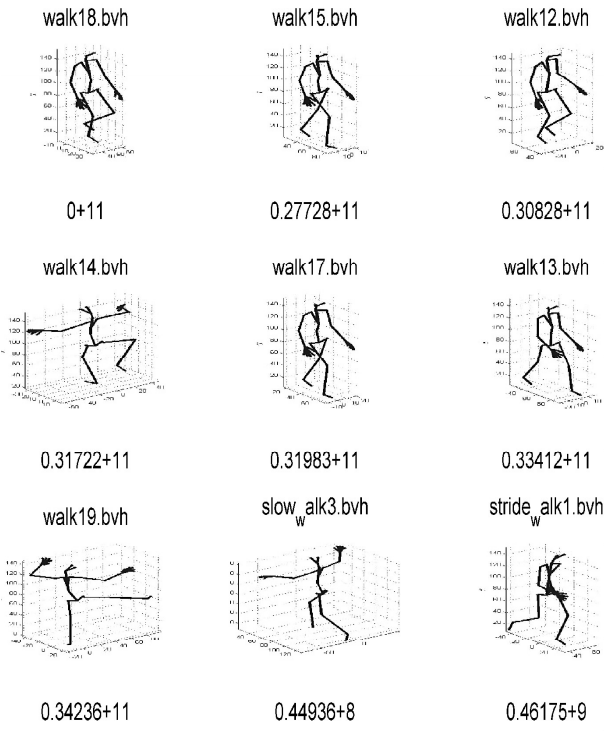


图 3 DTW 检索结果

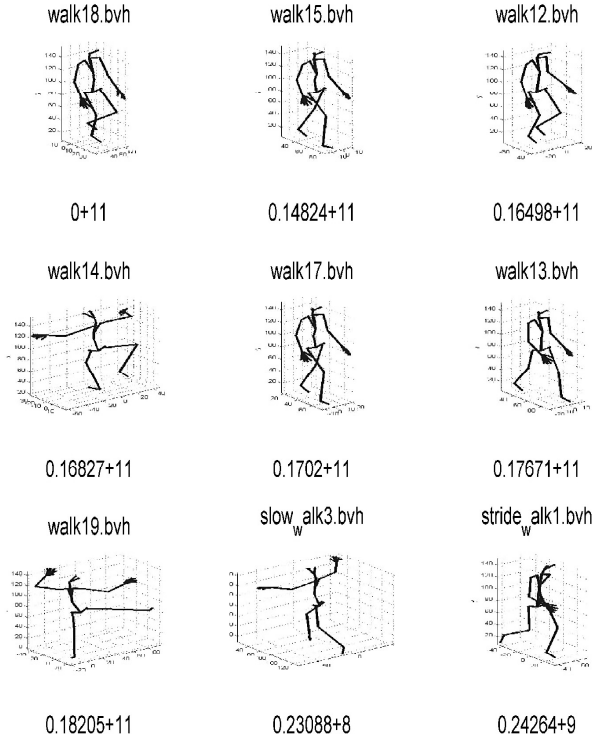


图 4 EMD 检索结果

而图 5 为 PR 曲线,所示为不同算法在测试数据库上得到的 PR(Precision Recall)值,可以看到,文中方法相对其他方法而言取得了较好的结果。文中算法的实验结果优于 Euclidean distance 和 DTW 算法,因为这两种算法在相似度匹配时忽略了动态时间轴全局匹配的时序关系。例如采用 Euclidean distance 和 DTW 算法就无法区分“正走”与“倒走”运动。文中采用 EMD

算法能够更好地检测到运动数据片段之间的全局时间对应关系,因此能够得到更加准确的检索结果。EMD 能检索全局范围,相比于其他算法有着更小的波动幅度。

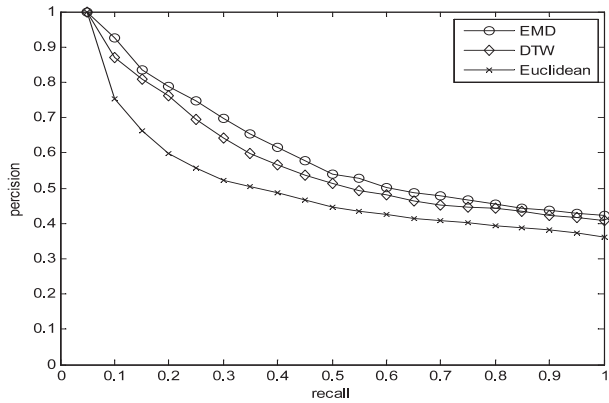


图5 PR 曲线

4 结束语

提出一种新的基于四元数和 EMD 度量的人体运动分析及识别技术,其贡献主要不依赖于预先设定的参数,且相对于现有运动数据检索方法而言,其能够获得更好的索引技术,且该索引的时间效率与检索效果优于欧氏距离和 DTW。今后的主要工作是:

- 1) 研究基于关节旋转信息的运动数据能够随着新数据的加入而频繁更新,不影响检索效率。
- 2) 在运动检索过程中考虑用户反馈信息,引入相关反馈机制,以提高检索结果的查准率与查全率。

参考文献:

[1] Haritaoglu I, Harwood D, Davis L S. W⁴: Real-time surveillance of people and their activities[J]. IEEE transactions on pattern analysis and machine intelligence, 2000, 22(8): 809-830.

[2] Liu Ling, Pu Calton, Han Wei. An XML-enabled data extraction tool kit for web[J]. Information systems, 2001, 26(8): 563-583.

[3] 向坚, 朱红丽. 基于三维特征提取的人体运动数据分析和检索[J]. 计算机应用, 2008, 28(5): 1344-1346.

[4] 潘红, 肖俊, 吴飞, 等. 基于关键帧的三维人体运动检索[J]. 计算机辅助设计与图形学学报, 2009, 21(2): 214-222.

[5] Muller M, Roder T, Clausen M. Efficient content-based retrieval of motion capture data[J]. ACM transactions on graphics, 2005, 24(3): 677-685.

[6] Vlachos M, Gunopulos D, Keogh E. Indexing multidimensional time-series[J]. The VLDB journal, 2006, 15(1): 1-20.

[7] Keogh E, Palpanas T, Zordan V B. Indexing large human-motion databases[C]//Proc of VLDB. [s.l.]: [s.n.], 2004: 780-791.

[8] CMU, Carnegie-Mellon MoCap Database [EB/OL]. 2003. <http://mocap.cs.cmu.edu>.

[9] 连德忠. 四元数向量和矩阵的实表示[J]. 厦门大学学报(自然科学版), 2003, 42(6): 704-708.

[10] 孙吉贵, 刘杰, 赵连宇. 聚类算法研究[J]. 软件学报, 2008, 19(1): 48-61.

[11] Rubner Y, Tomasi C, Guibas L J. Metric for distributions with applications to image databases[C]//Proceedings of the IEEE international conference on computer vision. [s.l.]: [s.n.], 1988: 59-66.

[12] Xu Jia, Zhang Zhenjie. Efficient and effective similarity search over probabilistic data based on earth mover's distance[C]//Proceedings of the VLDB Endowment. [s.l.]: [s.n.], 2010: 758-769.

[13] Rubner Y, Guibas L J, Tomasi C. The earth mover's distance as a metric for image retrieval[J]. International journal of computer vision, 2000, 40(2): 99-121.

(上接第 89 页)

报, 2009, 20(5): 1349-1361.

[4] Lun D S, Medard M, Ho T, et al. Network coding with a cost criterion[C]//Proceeding of the 2004 int'l symposium on information theory and its applications. [s.l.]: [s.n.], 2004: 1232-1237.

[5] Lun D S, Ratnakar N, Koetter R, et al. Achieving minimum-cost multicast: A decentralized approach based on network coding[C]//Proceeding of the IEEE INFOCOM 2005. [s.l.]: [s.n.], 2005: 1607-1617.

[6] Lun D S, Ratnakar N, Medard M, et al. Minimum-cost multicast over coded packet networks[J]. IEEE transaction on information theory, 2006, 52(6): 2608-2623.

[7] 王庆斌, 梅中辉. 无线网络中基于网络编码的最小能量多播[J]. 计算机技术与发展, 2013, 23(1): 150-153.

[8] Chou P A, Wu Y, Jain K. Practical network coding[C]//Proc of allerton conf comm cont comp. [s.l.]: [s.n.], 2003.

[9] Ghasvari H, Raayatpanah M A, Khalaj B H, et al. Optimal sub-graph selection over coded networks with delay and limited-size buffering[J]. IET communications, 2011, 5(11): 1497-1505.

[10] Wu Y. A trellis connectivity analysis of random linear network coding with buffering[C]//Proceeding of IEEE international symposium on information theory. [s.l.]: [s.n.], 2006: 768-772.

[11] Ho T, Lun D S. Network coding: An introduction[M]. New York: Cambridge University Press, 2008.

[12] Wieselthier J E, Nguyen G D, Ephremides A. Energy-efficient broadcast and multicast trees in wireless networks[J]. Mobile networks and applications, 2002, 7(6): 481-492.

基于四元数描述和EMD的人体运动捕获数据检索

作者:	<u>肖秦琨, 李俊芳, 肖秦汉, XIAO Qin-kun, LI Jun-fang, XIAO Qin-han</u>
作者单位:	<u>肖秦琨, 李俊芳, XIAO Qin-kun, LI Jun-fang(西安工业大学 电子信息工程学院, 陕西 西安, 710021), 肖秦汉, XIAO Qin-han(中国电子科技集团公司第20研究所军代室, 陕西 西安, 710021)</u>
刊名:	<u>计算机技术与发展</u>
	<div>ISTIC</div>
英文刊名:	<u>Computer Technology and Development</u>
年, 卷(期):	<u>2014(3)</u>

本文链接: http://d.wanfangdata.com.cn/Periodical_wjfz201403023.aspx