UDC 004.652, DOI: 10.2298/CSIS1001099W

A Retrieval Method for Human Mocap Data Based on Biomimetic Pattern Recognition

Xiaopeng Wei[1], Boxiang Xiao[1], and Qiang Zhang[1]

[1]Key Laboratory of Advanced Design and Intelligent Computing(Dalian University),
Ministry of Education, Dalian, 116622, China
zhangq26@126.com

Abstract. A retrieval method for human Mocap (Motion Capture) data based on biomimetic pattern recognition is presented in this paper. BVH rotation channels are extracted as features of motion for both the retrieval instance and the motion data. Several hyper sausage neurons are constructed according to the retrieval instance, and the trained domain covered by these hyper sausage neurons can be considered as the distribution range of a same kind of motions. By use of CMU free motion database, the retrieval algorithm has been implemented and examined, and the experimental results are illustrated. At the same time, the main contributions and limitations are discussed.

Keywords: MoCap (Motion Capture) Data, Retrieval, Biomimetic Pattern Recognition, HSN (Hyper Sausage Neurons).

1. Introduction

In past decades, with the development and popularity of motion capture device, human motion capture including motion data processing and retrieval has attracted increasing attention from many researchers [1-15]. There is a rapidly growing data of human motion, and the reuse of human motion capture data is becoming a practical task. However, because of the lack of general and efficient motion retrieval systems, the reuse of the motion data is limited by many problems and the investigation of motion data retrieval approaches is still a hot issue.

In this paper, we present a retrieval method for human Mocap data based on biomimetic pattern recognition [16], and the method constructs a set of the HSN (Hyper Sausage Neurons) for each kind of motion. Biomimetic pattern recognition also called topological pattern recognition is a novel pattern recognition principles proposed by Wang and colleagues [17] which is based on “matter cognition” instead of “matter classification” in traditional statistical pattern recognition. This recognition model is better closer to the function of human being rather than traditional statistical pattern recognition using “optimal separating” as its main principle. Biomimetic pattern recognition covers the distribution of a feature vectors exactly in high-dimensional space, and has been widely used in face recognition [18], speaker identify [19] and other feature recognition fields.

The aim of this work is to implement an effective retrieval approach for human motion capture data. The flowchart of method in this paper is shown in Fig1. The remainder of the paper is organized as follows: We introduce the feature extraction of motion data in Section 3 and describe our retrieval algorithm based on biomimetic pattern recognition in Section 4. Moreover, in Section 5, the part of experimental results and comparison to other methods are presented and illustrated. Finally, we conclude and explain the main advantages and limitations of this method and discuss the future works in Section 6.

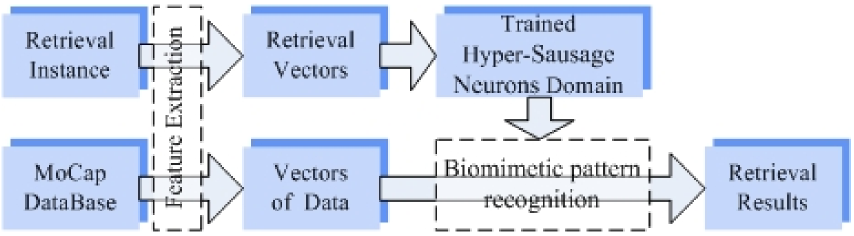

Fig. 1. The flowchart of method in this work.

2. Related Works

Many works in human motion data retrieval have been developed in past decades. In previous studies, one of developed techniques is motion templates based methods for human motion retrieval and classification. Muller and his colleagues [3] proposed a method for automatic classification and retrieval of motion capture data facilitating the identification of logically related motions scattered in some database. Roder introduced templates methods systematically in his doctoral dissertation [4]. Another technique is content-based or index-based methods. For instance, Chiu et.al [2] put forward to a framework for constructing a content-based human motion retrieval system including two major components: indexing and matching. Muller et al [5] presented automated methods for efficient indexing and content-based retrieval of motion capture data. Yamasaki and his colleagues [6] described a content-based cross search scheme for two kinds of three-dimensional (3D) human motion data: time-varying mesh (TVM) and motion capture data. Chao and colleagues [7] presented a simple and effective approach for motion retrieval and synthesis based on posture feature indexing, and posture features of each frame data was extracted by an index function. A 3D motion retrieval method with motion index tree was presented by Liu and colleagues [8]. Feature analysis is also a common used approach by many researchers. A set of relational motion features was been defined in Demuth's work [1]. Xiang and Zhu [9] extracted 3D temporal-spatial features of motion data and automatically constructed data driven decision trees. Lin [10] also defined a kind of motion features in his work. Many researchers used dynamic time warping (DTW) for different extent motion sequences [1,2]. Furthermore, there are several other methods such as probabilistic principal component analysis (PPCA) proposed by Wang [11], movement notation language presented by Yu et.al [12], Ensemble HMM Learning based approach developed by Xiang and Zhu [13], energy morphing based method proposed by Tam et.al [14] and semantic matching based method [15].

3. Feature Extraction of Motion Data

BVH rotation channels are used as features of motion for both the retrieval instance and the motion data in this work. The BVH file format for motion capture is originally developed by Biovision [20], a motion capture services company, as a way to provide motion capture data to their customers. A BVH file has two parts, a header section which describes the hierarchy and initial pose of the skeleton; and a data section which contains the motion data [21]. In BVH format, all joints' position can be calculated by their all parents' local rotation channels and root's translation. The BVH format doesn't account for scales so it isn't necessary to worry about including a scale factor calculation.

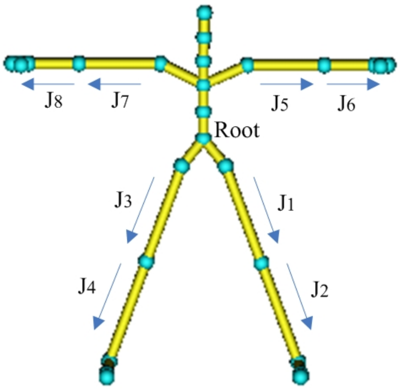

Fig. 2. 8 selected joints including Hips, Knees, Shoulders and Elbows

The BVH conversion of CMU motion data [22] we used in experiments is a standard database and all motions are represented by uniform BVH tree and scales. Here we select 8 joints including Hips, Knees, Shoulders and Elbows as the main feature of motion and extract their correlative values of local rotation channels in Motion sections of BVH files, as shown in Fig.2. Then, we define a 24 dimensional feature vector $M^{24} = (Zrot_{J1}, Yrot_{J1}, Xrot_{J1}, Zrot_{J2}, \cdots, Xrot_{J8})$ by the selected 8 joints' channel values, and all motions and retrieval instances are simplified by feature vectors.

4. Biomimetic Pattern Recognition-Based Retrieval Method

4.1. The Retrieval Instance

The retrieval instance in this work is a set of short query motion clips. Several frames of key pose in query clips are selected by manual operation, and all key pose frames' feature vectors mentioned in Section 3 are subsequently extracted. And then, a Remark Sequence of retrieval instance is constructed at the same time. Fig.3 shows an example of one motion clip of walk, in that the 4 heavy black frames are key poses of a walk clip with Remark Sequence $[1,2,3,4]$, and others are general data.

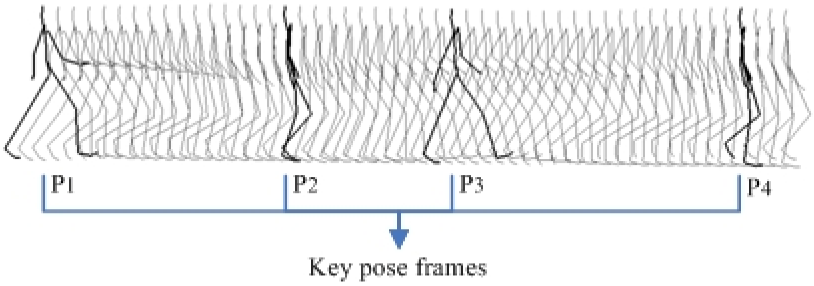

Fig. 3. Key poses of walk (heavy black)

4.2. HSN Covering in High-Dimensional Space

Biomimetic pattern recognition is a novel pattern recognition principles [17][18] which is based on "matter cognition" instead of "matter classification" in traditional statistical pattern recognition. Fig.4 shows the main principle of HSN chains based BPR and the comparison with traditional BP and RBF networks. The triangles represent samples to be recognized, and the circles and crosses represent samples to be distinguished from triangles. Polygonal line denotes the classification boundaries of traditional BP networks. Big circle denotes Radial basis function (RBF) networks.

A hyper sausage neuron can be expressed by topological product of a hyper-sphere and a line segment in high-dimensional space, as shown in Fig.5. For biomimetic pattern recognition, the task of geometrical learning is to cover a given sample set by a chain of hyper sausage units with a minimum sum of volumes via determining the end points of each line segment and the radius of each hyper-sphere.

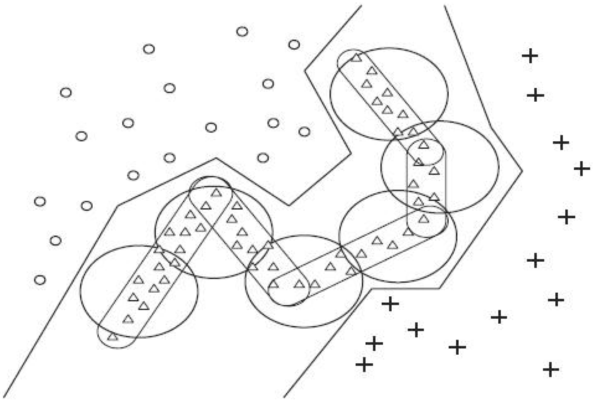

Fig. 4. HSN chains based BPR and comparison with BP and RBF networks.

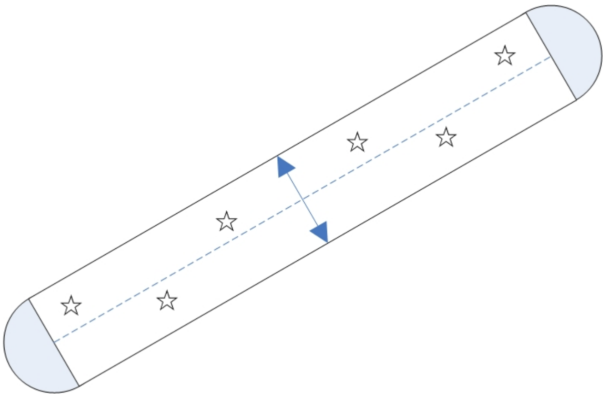

Fig. 5.. A hyper sausage neuron schematic diagram in some 2D section.

And a hyper sausage neuron can be approximately expressed by Eq. 4.1-4.3:

$$f_{HSN}(X) = \mathrm{sgn}(2^{-\frac{d^2(X,\overline{X_1,X_2})}{r^2}} - 0.5) \qquad 4.1$$

which contains a radius parameter r and the distance d between X and the line segment $\overline{X_1,X_2}$ as follows:

$$d^2(X,\overline{X_1,X_2}) = \begin{cases} \|X-X_1\|^2, & q(X,X_1,X_2)<0, \\ \|X-X_2\|^2, & q(X,X_1,X_2)>\|X_1-X_2\|, \\ \|X-X_1\|^2-q^2(X,X_1,X_2), & otherwise \end{cases} \qquad 4.2$$

$$q(X,X_1,X_2) = (X - X_1)\frac{(X_1-X_2)}{\|X_1-X_2\|} \qquad 4.3$$

In motion retrieval, we consider a feature vector of a frame of motion clip in motion database as X , and calculate the distances between X and all feature vectors of a kind of key pose with same mark. And then, we note the two nearest objects in retrieval instance as X_1 and X_2 respectively. Finally, we apply the construction algorithm of hyper sausage neuron with the parameters X , X_1 , X_2 and a threshold r by Eq4.1-4.3. If the sign of calculating result $f_{HSN}(X)$ is plus, the feature vector X can be considered as a similar object to the kind of key pose, and be remarked by same mark of the key pose.

4.3. The Evaluation of Similarity

For motion retrieval, hyper sausage neurons covering in high-dimensional space is used to estimate the distribution of every motion data in the space. If a frame's feature vector in a motion clip have a same mark to pose P_k , it can be considered as a similar object to P_k and be noted as k in Cover Sequence, or else it is noted as 0 . Cover Sequence noted as C is an integer array to record the similar status of a motion sequence, for example in Eq4.4, and it can be simplified. In sequence, zero elements are removed, and continuous same elements are denoted by one element.

$$C = [0,0,0,1,1,0,2,2,2,2,0,0,3,4,4]$$

$$C_{simple} = [1,2,3,4] \qquad 4.4$$

The Cover Sequence of a motion sequence is the unique evaluation for retrieval. If a motion's the Cover Sequence continuously cover the Remark Sequence in order, it is considered as the similar object. Furthermore, part sequence of a sequence is defined to evaluate main similarity, and it is actually the sub sequence of a sequence with one element absent. For example, sequences $[1,2,3]$, $[2,3,4]$, $[1,2,4]$, $[1,3,4]$ are all part sequences of sequence $[1,2,3,4]$.

For instance, if in a Cover Sequence C_{walk} of a walk sequence, a sub sequence equal to the Remark Sequence of retrieval instance such as $[1,2,3,4]$ in Fig.3 exists, the motion can be considered as the similar object to walk and it can be retrieved. If the Cover Sequence of a motion contains part but whole Remark Sequence of retrieval instance, we can consider it as the main similar object which can also be recalled.

5. Experimental Results

To examine our algorithm, we developed a motion retrieval system by C++ program language and OpenGL (Open Graphic Library), and all experimental results were obtained on a 3.0 GHz Pentium 4 with 2 GB of main memory. We evaluated the system on a subset of BVH conversion of the CMU motion database [22], which contains 184 motions about 105 thousands frames sampled at 120 Hz of motion capture data including walk, run, jump, cartwheel, swing and their blends.

The retrieval accuracy of the proposed framework is evaluated by the precision and recall which are commonly adopted by many researchers.

$$\text{precision}=\frac{\#\{\text{relevant}\cap\text{retrieved}\}}{\text{retrieved}}$$

$$\text{recall}=\frac{\#\{\text{relevant}\cap\text{retrieved}\}}{\text{relevant}}$$

where #retrieved is the number of retrieved clips and #relevant is the number of relevant clips. Table 1 shows the part of retrieval results for walk and jump.

Table 1. part of retrieval results for walk and jump

	threshold	precision	recall
walk	30.0	1.000	0.862
	40.0	0.930	0.914
	45.0	0.688	0.914
	50.0	0.578	0.948
	55.0	0.586	1.000
jump	50.0	1.000	0.562
	55.0	0.983	0.795
	60.0	0.848	0.918
	70.0	0.634	0.973
	75.0	0.598	1.000

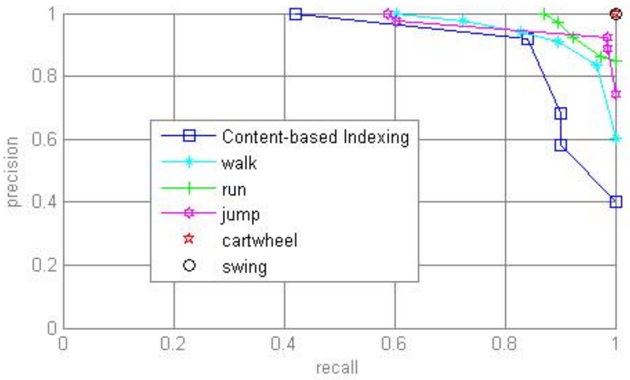

Fig. 6. part results of retrieval accuracy of proposed approach and the comparison to Content-based Indexing method.

The part results are also shown in Fig. 6, and Fig.7 shows part visual retrieval results. In experiments, some motions such as cartwheel and swing achieve results with accuracies at 1.0, because the numbers of samples in database are few respectively at 9 and 10. To evaluate the performance of ours method, we compare our accuracy with the referenced Content-based Indexing method [2], shown in Fig 6.

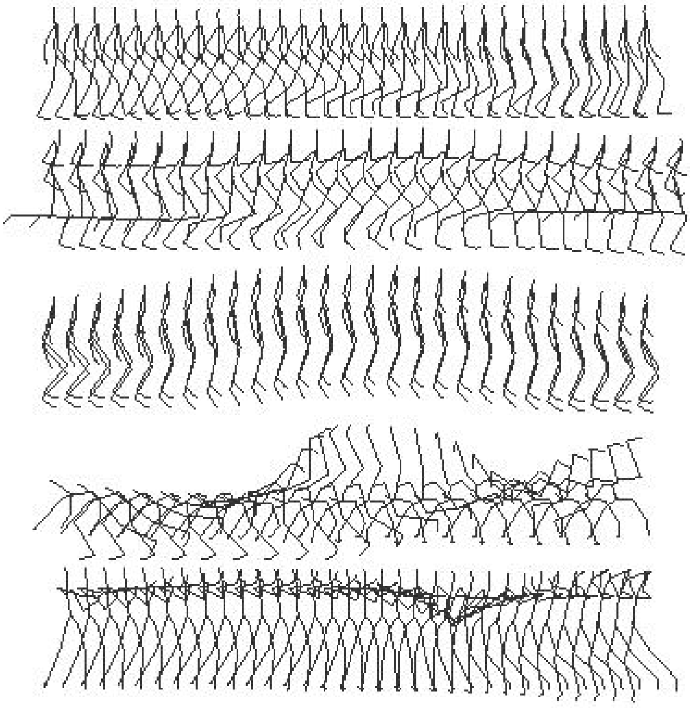

Fig. 7. part visual retrieval results including walk, run, jump, cartwheel, swing from upper to lower.

- On the other hand, time cost of approach is evaluated in ours work. The time cost includes two main parts: feature vectors extraction and computation of HSN covering. A main time-consuming operation is feature vectors extraction for the whole database (about 2 minutes) and the operation needs to be executed only one time in all experiments because the structures of feature vectors are same and uniform for different retrieval instances. Moreover, the time cost of covering computation traversing the whole database about for a retrieval instance is about 25 seconds.

6. Conclusion

In this paper, we presented a retrieval method for human Mocap (Motion Capture) data based on biomimetic pattern recognition is presented in this paper. The retrieval instance in this work is a set of short query motion clips. BVH rotation channels are extracted as features of motion for both the retrieval instance and the motion data. Several hyper sausage neurons are constructed according to the retrieval instance, and the trained domain covered by these hyper sausage neurons can be considered as the distribution range of a same kind of motions. By computation of hyper sausage neurons covering based on biomimetic pattern recognition, relevant motions to the retrieval instance are recalled, and experimental results are illustrated.

The advantage of this work is twofold. Firstly, the proposed algorithm is based on the distribution of similar motions in high-dimensional space, and avoids the influences by differences in extent and velocity of motion sequences. Furthermore, the use of BVH rotation channels of motion data eliminates the differences in geometric parameters including coordinates, position, direction and scales of motions, and the uniform structure of feature vectors ensure the feasibility and efficiency of retrieval in large database.

One limitation of our approach is that the covering domains of hyper sausage neurons of some similar motions such as walk and run may be intersectant in few scenes which can reduce the retrieval accuracy. Another shortage is that some complex freeform motion can not be expressed by distribution in multidimensional space well and it is a difficult task for future work. Moreover, we are also focus on the dynamic distribution of motions in high-dimensional space.

7. Acknowledgment

This work is supported by National Natural Science Foundation of China (Grant No. 60875046), by Program for New Century Excellent Talents in University (Grant No. NCET-06-0298), by the Key Project of Chinese Ministry of Education.(No.209029), by China Postdoctoral Science Foundation ((Grant No. 2009045 1266), by the Program for Liaoning Innovative Research Team in University (Grant Nos.2008T004 and 2009T005), by the Program for Liaoning Science and Technology Research in University (Grant Nos.2009S008,2009S009) and by the Program for Liaoning Excellent Talents in University (Grant No.2006R06).

8. References

1. Demuth, B. and Roder, T. and Muller, M. and Eberhardt, B: An Information Retrieval System for Motion Capture Data. LNCS 3936. 373–384. (2006)
2. Chiu, C. and Chao, S. and Wu, M. and Yang, S. and Lin, H.: Content-based retrieval for human motion data. Journal of Visual Communication and Image Representation. 3, 15, 446–466. (2004) Muller, M. and Roder, T.: Motion Templates for Automatic Classification and Retrieval of Motion Capture Data. Eurographics/ ACM SIGGRAPH Symposium on Computer Animation. 137-146. (2006)

3. Roder, T.: Similarity, Retrieval, and Classification of Motion Capture Data. University of Bonn. 93-110. (2006)
4. Muller, M. and Roder, T. and Clausen, M.: Efficient Content-Based Retrieval of Motion Capture Data. ACM Transactions on Graphics (TOG). 24, 3, 677 - 685. (2005)
5. Yamasaki, T. and Aizawa, K.: Content-Based Cross Search for Human Motion Data Using Time-Varying Mesh and Motion Capture Data. Proceeding of IEEE International Conference on Multimedia and Expo (ICME2007). 2007-2009. (2007)
6. Chao, S. and Chiu, C. and Chao, J. and Ruan, Y. and Yang, S.: Motion Retrieval and Synthesis Based on Posture Feature Indexing. Proceedings of the Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'03). 266 - 271. (2003)
7. Liu, F. and Zhuang, Y. and Wu, F. and Pan, Y.: 3D motion retrieval with motion index tree. Computer Vision and Image Understanding. 92, 2, 265–284. (2003)
8. Xiang, J.and Zhu, H.: Motion Retrieval Based on Temporal-Spatial Features by Decision Tree. LNCS 4561. 224–233. (2007)
9. Lin, Y.: Efficient Motion Search in Large Motion Capture Databases. LNCS 4291. 151–160. (2006)
10 Wang, X. and Yu, Z. and Wong, H.: 3D Motion Sequence Retrieval Based on Data Distribution. ICME 2008. 1229-1232. (2008)
11. Yu, T. and Shen, X. and Li, Q. and Geng, W.: Motion retrieval based on movement notation language. Computer Animation and Virtual Worlds. 16, 273–282. (2005)
12. Xiang, J. and Zhu, H.: Ensemble HMM Learning for Motion Retrieval with Non-linear PCA Dimensionality Reduction.IIHMSP 2007. 604 – 607. (2007)
13. Tam, G. and Zheng, Q. and Rynson, M. and Lau, W.H.: Motion Retrieval Based on Energy Morphing. Ninth IEEE International Symposium on Multimedia 2007. 210-217. (2007)
14. Wei, X. and Xiao, B. and Zhang, Q. And Zhou, D.: Effective approach based on semantic matching for retrieval of MoCap data. Journal of Computational Information Systems. 3, 4, 859-866. (2008)
15. Wang, S. and Lai, J.: Geometrical learning, descriptive geometry, and biomimetic pattern recognition. Neurocomputing. 67, 9–28. (2005)
16. Wang. S.: Bionic (Topological) Pattern Recognition-A New Model of Pattern Recognition Theory and Its Applications. Chinese Journal of Electronics. 30, 10, 1417-1420. (2002)
17. Wang, Z. and Mo, H. and Lu, H. and Wang, S.: A method of biomimetic pattern recognition for face recognition. Proceedings of the International Joint Conference on Neural Networks 2003. 2216-2221. (2003)
18. Qin, H. and Wang, S. and Sun, H.: Biomimetic Pattern Recognition for Speaker-Independent Speech Recognition. International Conference on Neural Networks and Brain2005. 1290-1294. (2005)
19. http://www. Biovision.com.
20. http://www.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/BVH.html
21. http://mocap.cs.cmu.edu/

Xiaopeng Wei is a professor at Dalian University, Dalian, China. His research area include computer animation, intelligent CAD. So far, he has (co-) authored about 160 papers published.

Boxiang Xiao is a lectuer at Dalian University, Dalian, China. He has received his Ph.D degree in 2009 at Dalian University of Technology.

Qiang Zhang is a professor at Dalian University, Dalian, China. His research interests are computer animation, intelligent computing. Now he has served as editorial boards of seven international journals and has edited special issues in journals such as Neurocomputing and International Journal of Computer Applications in Technology.

Received: May 15, 2009; Accepted: October 26, 2009.