

Human Motion Behavior Segmentation Based on Local Outlier Factor

Xing Weiwei^{*}, Sun Liya and Wei Xiang

School of Software Engineering, Beijing Jiaotong University, Beijing, P.R. China

Abstract: Motion segmentation, which is a crucial technology in reuse of motion capture data, means automatically dividing a long motion sequence into several motion clips which have different semantics. In this paper, a new motion segmentation method based on Local Outlier Factor (LOF) is proposed. Our method is based on the assumption that motions with same type can form a cluster and the transition of two behaviors is the outlier of the two clusters. We use LOF to find a “bridge” between different semantic motion clips. The presented method mainly consists of four stages as follows: firstly, bone angle is employed to represent the motion characteristic in preprocessing stage; secondly, sliding window is used as the statistical unit and statistic histogram is built for all bone angles in sliding window; thirdly, LOF is calculated for every sliding window; finally the segmentation point can be obtained by locating the peak of LOF curve. The proposed method is evaluated on the Carnegie Mellon Motion Capture database. The experimental results show that the method can achieve the automatic segmentation for human motion capture data, and has better performance compared with the classical PPCA based segmentation method. Additionally, the proposed method shows high efficiency even dealing with long motion sequences.

Keywords: Motion segmentation, Local Outlier Factor, Motion capture data, Sliding window.

1. INTRODUCTION

With the advancement of motion capture techniques and the popularity of motion acquisition systems, the motion capture data has been widely used in various applications, such as computer animation and movie industry. More and more motion capture databases are available to the public, such as the CMU MoCap database [1] and the HDM05 database [2]. Since motion capture data could reflect the naturalness and subtle details in human behavior, the usage of pre-recorded motion capture data has become an important issue in computer animation. Many researches have been done in reuse of the motion capture data for motion synthesis [3-5], motion retrieval [6-8] and motion classification [9, 10] and most researches focus on the small motion sequences which only have single semantic meaning like running, walking, etc. However, the recorded motion capture data usually turns out to be long and extended sequence. On one hand, the actor could perform more naturally and comfortably while doing a series of behaviors for one time. On the other hand, the longer motion sequence could save more information like the natural transitions between two different human behaviors. Thus motion behavior segmentation is crucial for reusing the motion capture data more efficiently. It is noted that the motion segmentation here is high-level segmentation, which aims at cutting the long original motion sequence into semantically different motion clips. This is a labor-intensive and error-prone process and is only feasible for a limited set of gestures. However, for the ever-growing amount of

motion capture data, it is time-consuming and tedious for user to segment the motion capture data manually.

In this paper, an unsupervised automatic motion behavior segmentation method based on Local Outlier Factor (LOF) is proposed, and it can create a segmentation without any prior or training models. Local Outlier Factor proposed by Markus [11] is an anomaly detection algorithm by comparing the local density of a point's neighborhood with that of its neighbors. The LOF is a degree to indicate one object's outlier-ness from a cluster. Since the different small semantic motion clips can be considered as different clusters, and the transition between two different motions is a “bridge” which consists of all the outlier of the two clusters. Only if we find the outlier which has the biggest LOF, the most reasonable cutting frame will be obtained. Our goal is to segment the motion sequence to get the semantically different motion clips even when we lack the prior known information of behaviors. In our view, a small motion clip may be composed of more than one consecutive motion cycle and it needn't to be divided further, e.g., one successful divided clip has three consecutive similar jumps and we needn't to segment it into three smaller clips further. Additionally, the method omits the absolute body position and orientation so that the motion isn't cut when it appears some motion like turning back or turning around in the same behavior. Experiments are carried out to evaluate the performance and efficiency of our method.

The remainder of this paper is organized as follows. In Section 2, we review the related work. Section 3 overviews the framework of our method. Section 4 describes the proposed method in detail, including feature extraction, building statistic histogram in sliding window, LOF computation and cutting frame selection. The experimental results are shown

^{*}Address correspondence to this author at the School of Software Engineering, Beijing Jiaotong University, Haidian District, Beijing 100044, P.R. China; Tel: +86-10-51683866; E-mail: wwxing@bjtu.edu.cn

in Section 5. The Section 6 concludes the paper and outlines future work.

2. RELATED WORK

In view of massively growing motion capture data, the reuse of such data has become a challenging issue. Motion segmentation, which aims at cutting human motion sequence into distinct semantic units, facilitates the further application of motion capture data. As an important step of preprocessing motion capture data, motion segmentation has attracted many researchers from different disciplines and remains a difficult task in computer animation for decades. In order to implement motion segmentation, most method and techniques have been studied and proposed.

Dimensionality reduction techniques usually aid in clustering, modeling, and other processing of motion. Here, dimension reduction based approach is also a typical branch in dealing with motion segmentation problem. Considering the different types of human behavior data should have different intrinsic dimensionalities, which mainly includes the situation of linear and nonlinear, Barbic [12] used the Principal Component Analysis (PCA) to cut the motion by detecting the sudden increase of the intrinsic dimensionality on the joints of the body. While PCA considers only the change in dimensionality and encodes only correlations between joints, Barbic [12] further proposed probabilistic Principal Component Analysis (PPCA) to improve the problem. PPCA assigned a cut when the distribution of poses is observed to change, based on the assumption that frames of different motions form different Gaussian distributions. Since the joint range of motion limits are implicit in PPCA and single distribution for each individual behavior is estimated in PPCA, the experiment shows that PPCA approach provide the better performance than PCA. However, with the increasing number of frames in a long motion sequence, the time consuming become higher.

Model based method is another useful branch in motion segmentation, typical examples include Hidden Markov Model (HMM) [13] and Gaussian Mixture Model (GMM) [12]. Inspired by natural language processing where words are extracted from long sentences using HMM, Wang [14] used HMM to do motion segmentation based on that the gesture primitives with a high-level structure are controlled by the temporal ordering. The sequences of human motion are firstly over segmented into atomic components, and then HMM is used to cluster them to extract lexicon from the discrete sequences. However different training data built different HMM model, which could influence the cutting result. In the application of GMM [12], the only assumption is that different kinds of motions can form different clusters which can be described reasonably well by a Gaussian distribution. Thus the reasonable cutting frame can be found whenever two consecutive sets of frames belong to different Gaussian distributions. In this GMM-based approach, the cluster number must be pre-set, so the different cluster numbers will lead to different results. Wang [14] used the Kernel Dynamic Texture (KDT) model to segment motion data,

because motion capture data of human essentially has the traits of common dynamic textures, such as the same type behaviors are composed by many similar elements and exhibit certain stationary properties in time. But when the motion sequence is quite long, the efficiency of learning KDT is very low.

The third main branch in motion segmentation is classifier based approach. Kanav [15] used the naive Bayesian classifier to derive creator profiles from empirical data which is the low-level motion parameters for hierarchical layered structure. Then these profiles were used to predict how the creators segment gestures in other motion sequences. Support Vector Machine (SVM) [16] classifier was also introduced to generalize the user annotations from the entire database. The coordinate vector they choose for each frame is the joint positions for one second of motion centered at the frame being classified. Zhou [17] propose Aligned Cluster Analysis (ACA), an extension of kernel k-means clustering, to temporally segment streams of motion capture data into actions. Xiao [18] introduced an auto-segmentation method which firstly uses a non-linear dimensionality reduction technique to map original motion sequences into low-dimensional manifold and then clustering techniques are applied to segment primitive actions apart. Obviously more training data and empirical data must be needed so that these classifiers could work well.

In this paper, we propose a motion segmentation method based on LOF which does not need any training data and cluster number in advance. And the experiment turns out that the results and the efficiency of our approach is better than PPCA even when dealing with longer motion sequences.

3. FRAMEWORK OVERVIEW

Fig. (1) presents an overview of our framework. The original long motion capture data are preprocessed, and the angle between different bones is extracted as a feature in order to reduce the dimensionality and simplify the computation. Then the sliding window is introduced as the statistical unit to build statistic histogram for every bone angles. LOF is calculated for each sliding window and LOF curve for the motion sequence will be obtained. Thus the reasonable cutting point, which is the frame whose LOF value is at the peak of the LOF curve, is obtained by curve simplification algorithm.

4. MOTION SEGMENTATION METHOD

4.1. Motion Feature Extraction

The original long motion capture data, which is represented by pose sequences of the body joint configurations over a period of time, can be defined as $M = \{m(i) | i \in [1, Framenum]\}$ where *Framenum* is the total number of frames in motion *M*, and $m(i)$, which consists of the position and orientation of the root joint and the Euler angle of all the rest joints of the body, is the pose of the character in the *i*th

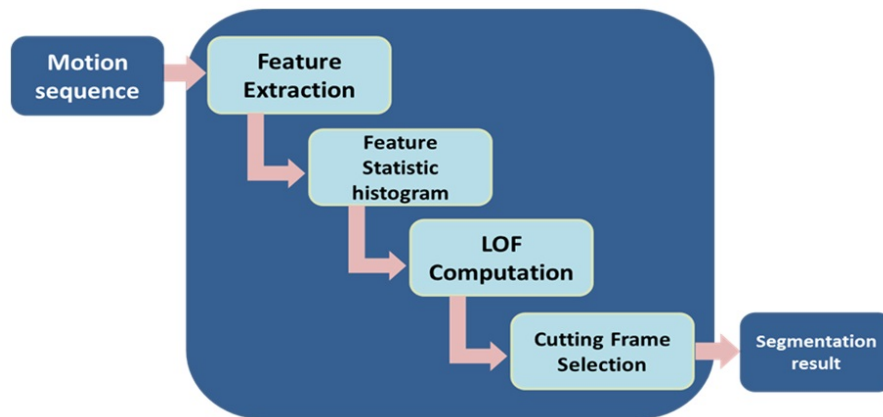


Fig. (1). Framework of our motion segmentation method.

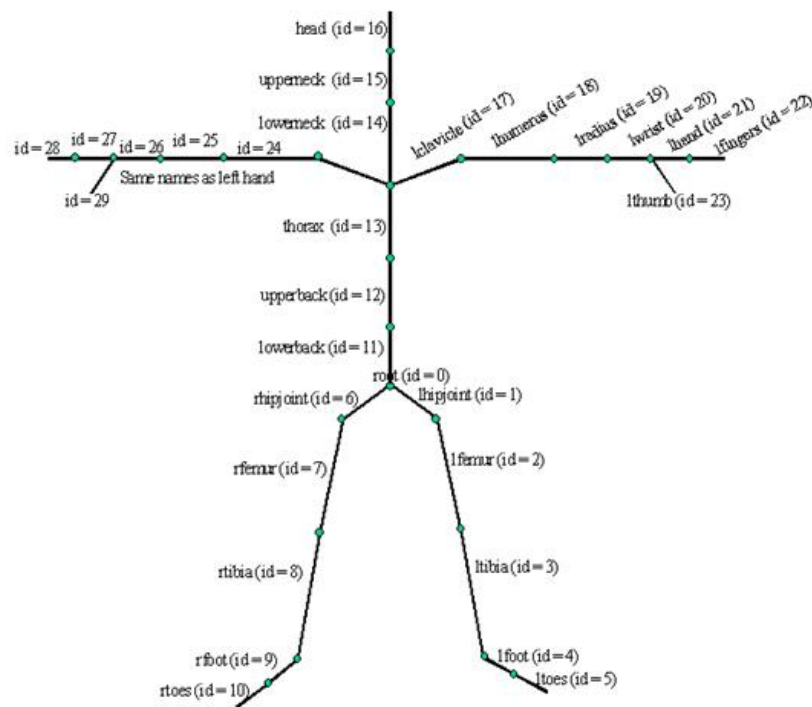


Fig. (2). Human skeleton model.

frame in the motion. In this paper, we use human skeleton model in Fig. (2), in which the joint number is 31 and degree of freedom (DOF) is 62. Thus the frame $m(i)$ can be represented in the 62-dimensional space which is too high and complicated to implement the similarity measurement etc. It is crucial to find an effective way for motion feature extraction and representation meanwhile keeping the intrinsic motion characteristics.

Inspired by bone angle proposed by Xiao [19], we extract the angle between different bones as the motion feature. Bone angles not only reflect the pose intuitively, but also reduce the dimensionality and improve the efficiency. Additionally, bone angles without position and orientation information of the body are effective in ignoring the turning back motions existing in the same behaviors. Considering that the

human limbs have the most obvious changes in human motion and play the decisive role in verdict of the type of motion, we choose the bones of torso, arm and leg in Table 1. The bones are defined as vectors pointing from parent joint position to the child. For example, $Bone_1$ can be defined as

$$Bone_1 = P_{root} - P_{lowerback} \tag{1}$$

where P_{root} and $P_{lowerback}$ are the position of root joint and lower back joint which is needed to calculate firstly.

Since the human skeleton is a hierarchical model and the low-level child bone has a very tight association with its' high-level father bone, the bone angle we used is the angle between the child bone and its father bone which is shown in Table 2. For example, the angle between $Bone_1$ and $Bone_2$ is defined as:

Table 1. Human bones used in this paper.

Bone	Vector Definition
Bone1	root->lowerback
Bone2	lclavicle->lhumerus
Bone3	lhumerus->lradius
Bone4	rclavicle->rhumerus
Bone5	rhumerus->rradius
Bone6	lhip->lfemur
Bone7	lfemur->ltibia
Bone8	rhip->rfemur
Bone9	rfemur->rtibia

Table 2. Bone angles used in this paper.

Bone Angle	Definition
BAngle1	(Bone1, Bone2)
BAngle2	(Bone1, Bone4)
BAngle3	(Bone1, Bone6)
BAngle4	(Bone1, Bone8)
BAngle5	(Bone2, Bone3)
BAngle6	(Bone4, Bone5)
BAngle7	(Bone6, Bone7)
BAngle8	(Bone8, Bone9)

$$BAngle_1 = \arccos\left(\frac{Bone_1 \cdot Bone_2}{|Bone_1||Bone_2|}\right). \quad (2)$$

Thus, instead of 62-dimensional original data, we get the 8-dimensional feature for each pose which can be defined as a vector $BA = (BAngle_1, BAngle_2, \dots, BAngle_8)$.

4.2. Feature Statistic Histogram

After extracting the motion feature, each frame of a motion sequence can be represented as:

$$BA(i) = (BAngle_1^i, BAngle_2^i, \dots, BAngle_8^i) \quad (3)$$

where $i \in [1, Framenum]$. This feature representation not only keeps the physical meaning of motion but also benefits from the later computation. However, when encountering a fierce behavior or a rapid behavior, the variation of bone angle will become dramatical (e.g. in a fierce kicking motion, the angular velocity of some leg joints change much faster and the angular acceleration becomes larger suddenly).

Therefore, those kinds of motion frames in the long sequence are very possible to be mistaken for the cutting frames. Fig. (3) shows the LOF curve of a fierce kicking motion in which frames whose joint has very high accelerated velocity can be detected as cutting points while all the frames belong to the same behavior kicking. In order to deal with this case, sliding window is introduced to build bone angle statistic histogram. We set the sliding window as statistical unit and then build statistic histogram for each bone angle in the sliding window.

The sliding window is defined as SW , which slides from the start to the end of the motion sequence M , and will be treated as the statistical unit. Given the length of the sliding window $SWlength$, the step size Δ for the sliding window moving and the total number of sliding windows $SWnum$, the w th sliding window can be represented as:

$$SW(w) = \{m(i) | i = (w - 1)\Delta + 1, \dots, (w - 1)\Delta + SWlength\} \quad (4)$$

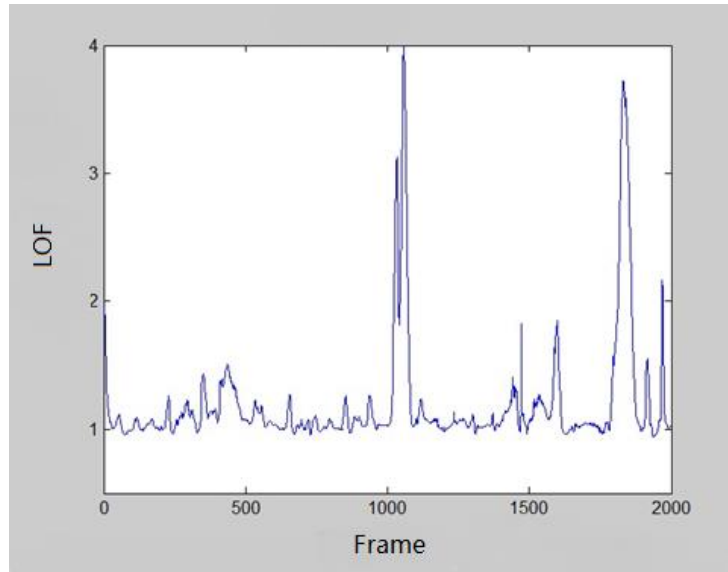


Fig. (3). LOF curve of a fierce kicking motion.

where $w \in [1, SWnum]$.

For each $SW(w)$, we do statistic for the bone angle feature of all frames in it. In details, each bone angle will have a statistic histogram which is used to model the distribution in a sliding window. We assume that the distributions of bone angles in same kind of motions are similar. Considering that the values of each bone angle in different kind of motions have different maximum and minimum, and the range of angle is undoubtedly from 0 to π , we divide the bone angle's interval $[0, \pi]$ into L equidistant intervals. The boundary of the j th ($j \in [0, L]$) interval is:

$$b_j = \frac{\pi}{L} * j. \quad (5)$$

And the j th interval is:

$$\theta_j \in [b_{j-1}, b_j]. \quad (6)$$

The statistic histogram of $BAngle_k$ where $k \in [1, 8]$ is defined as:

$$H_k = (h_1 = \frac{c_1}{SWlength}, \dots, h_L = \frac{c_L}{SWlength}) \quad (7)$$

where c_j ($j \in [0, L]$) is the total number of frames within the sliding window whose $BAngle_k$ value falls into the interval θ_j .

After building the statistic histogram for the 8 bone angles, we get a sliding window feature vector $SH(w) = (H_1, \dots, H_8)$ for the sliding window $SW(w)$. Therefore, the intrinsic motion characteristics are saved in the histogram, but the dramatic variation in fierce behaviors can be weakened to a certain degree.

4.3. LOF Computation

Local Outlier factor (LOF) was introduced by Markus [11], which introduced the measurement that quantifies how

much an object is outlying with respect to other objects in a space. In the following definition parameter k , which is the number of nearest neighbors, is supplied by user.

Definition 1: k -distance of an object p

For any positive integer k , the k -distance of object p , denoted as k -distance(p), is defined as the distance $d(p, o)$ between p and an object $o \in D$ such that:

- (i) for at least k objects $o' \in D \setminus \{p\}$ it holds that $d(p, o') \leq d(p, o)$, and
- (ii) for at most $k-1$ objects $o' \in D \setminus \{p\}$ it holds that $d(p, o') < d(p, o)$.

Definition 2: k -distance neighborhood of an object p

Given the k -distance of p , the k -distance neighborhood of p contains every object whose distance from p is not greater than the k -distance, i.e.

$$N_{k\text{-distance}(p)}(p) = \{q \in D \setminus \{p\} | d(p, q) \leq k\text{-distance}(p)\}. \quad (8)$$

These objects q are called the k -nearest neighbors of p .

Definition 3: Reachability distance of an object p w.r.t. object o

Let k be a natural number. The reachability distance of object p with respect to object o is defined as:

$$reach\text{-}dist_k(p, o) = \max\{k\text{-distance}(o), d(p, o)\}. \quad (9)$$

Definition 4: Local reachability density of an object p

The local reachability density of p is defined as:

$$lrd_k(p) = 1 / \left(\frac{\sum_{o \in N_k(p)} reach\text{-}dist_k(p, o)}{|N_k(p)|} \right). \quad (10)$$

Definition 5: (local) outlier factor of an object p

The (local) outlier factor of p is defined as:

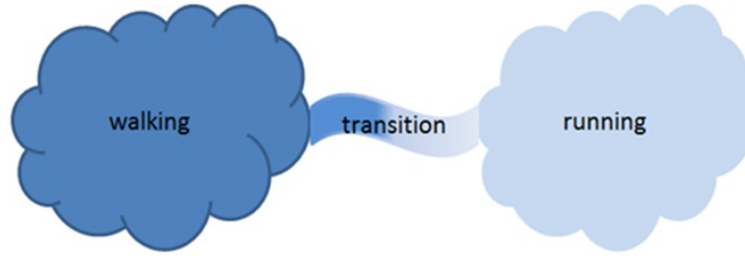


Fig. (4). Example for demonstrating our method.

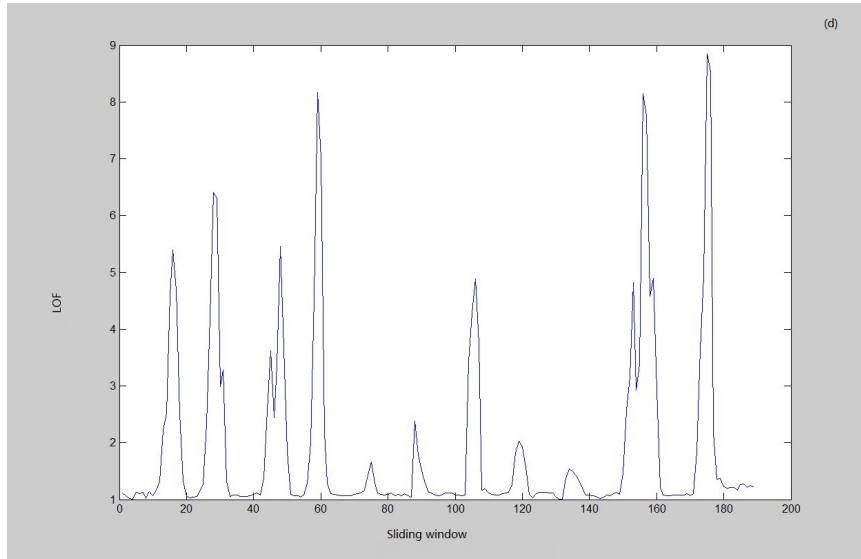


Fig. (5). LOF curve of a motion.

$$LOF_k(p) = \frac{\sum_{o \in N_k(p)} \frac{lr d_k(o)}{lr d_k(p)}}{|N_k(p)|} \quad (11)$$

According to the definition above, LOF is the degree to which the object is isolated from its surrounding neighborhood. For the objects deeply inside a cluster, the LOF value is approximately 1. For the other objects, the LOF will become higher when it is far away from the cluster.

The idea of introducing LOF to do motion segmentation is inspired by GMM based approach [12], which is based on the assumption that different kinds of motions can form different clusters and can be described reasonably well by a Gaussian distribution. Thus the reasonable cutting frame can be found whenever two consecutive sets of frames belong to different Gaussian distributions. Taking the walking and running as example in Fig. (4), the two behaviors form two different clusters respectively while the transition frames, which is the candidate of segmentation point, is reasonable close to the two cluster and the left part of transition frames obey the walking distribution while the right part of transition frames obey the running distribution. However, actually there are some kinds of human behaviors that can't be described well by Gaussian distribution. Additionally, the transition frames between two different kinds of behaviors mainly just have some local changes compared with its adjacent frames. LOF, which is a local density-based algorithm, is

very effective to find the “bridge” of the two motions by judging the outlier degree of the frames in motion sequence. Imaging that the LOF of frames in walking and running are approximately 1, the transition frames' LOF almost will monotonically increase from 1 to a peak value and then monotonically decrease to 1. So we can calculate LOF value for the frames, and the higher the LOF is, the more possible the cutting frame is.

Based on the obtained dimensionality reduced feature vector $\{SH(w) | w \in SWnum\}$ for the original motion sequence, we treat each $SH(w)$ as an object instead of frames and compute the LOF for them. Firstly, we compute the distance by Cosine distance, which is a similarity measurement by measuring the angle between the two vectors. The distance between two sliding windows $SH(i)$ and $SH(j)$ is defined as:

$$d(i, j) = \frac{SH(i) \cdot SH(j)}{|SH(i)| |SH(j)|} \quad (12)$$

Secondly, we manually set k and use KNN algorithm to find the k -distance neighborhood of each sliding window meanwhile getting the k -distance of each sliding window. Finally, LOF of each sliding window is calculated and then we can get the LOF curve in Fig. (5). obviously, the peak is the reasonable segmentation point.

Table 3. Motion data used in experiment.

File	Total Frame	Total Behavior
86_03.amc	8401	9
86_04.amc	10078	10
86_05.amc	8340	11
86_06.amc	9937	12
86_07.amc	8702	10
86_08.amc	9206	11

In this paper, we calculate the LOF for every sliding window instead of frame. The sliding window has many advantages for our method. It not only greatly reduces the computation and time cost, but also can avoid the influence of noisy frames and fierce behaviors. Furthermore, the sliding window keeps the temporal characteristics of motion sequences to a large extent.

4.4. Cutting Frame Selection

In order to get the peak frame, we first need to get the sliding window number of the peak in LOF curve, and then the cutting frame can be obtained through the sliding window number.

Firstly, our goal is to locate the maximum point in every peak. However, it should be noted that not every peak is exactly sharp, and most of them are zigzag. For example, the third and tenth peaks are not very easy to find the exactly sliding window number in Fig. (5). So only getting extreme point is far away from enough. Curve simplification [20] is an algorithm which generates a polygonal approximation of a given curve with a chain of line segments recursively until the error rate is small enough for each line segment. In this paper we use curve simplification to get the exact segmentation sliding window number set $\{S_i | i \in [1, Snum]\}$ where $Snum$ is the total number of peaks in LOF curve.

After obtaining the segmentation sliding window number S_i , the middle of the segmentation sliding window is the reasonable cutting point. Thus the cutting point can be defined as:

$$cp = (S_i - 1/2)SWlength. \quad (13)$$

5. EXPERIMENTS

We evaluate our motion segmentation method on the experimental dataset of motion capture sequences collected from CMU Graphics Lab Motion Capture Database [1]. The motion data format is AMC [2] in which the data is represented in the form of Euler joint angles. We choose 8 motions sequences performed by the subject 86, each contains approximately 9000 frames. Each sequence is a combination of 10 natural behaviors (e.g. walking, running, punching,

climbing). The length of each type behavior in a motion sequence is in the range of 80 to 300 frames. The detail information is shown in Table 3.

After extracting out the bone angle feature for each frame, further dimensionality reduction will be done by sliding window. In our experiment, each sliding window consists of 500 frames ($SWlength=500$) which guarantees that the sliding window include one motion period at least. The sliding window moves forward at the speed of 45 frames per time ($\Delta=45$), because the experimental motion sequence averagely has 9000 frames and the total number of sliding windows is desired to be 200 ($SWnum=200$). Then we built the histogram for every bone angle which is divided into 10 interval ($L=10$). Thus we can calculate the LOF for each sliding window. In LOF, k is the only parameter to be set manually and the higher the value of k , the more similar the reachability distances for objects within the same neighborhood. So here k is set to 20 while we have 200 sliding windows ($k=20$). The final results are the peaks of the LOF curve.

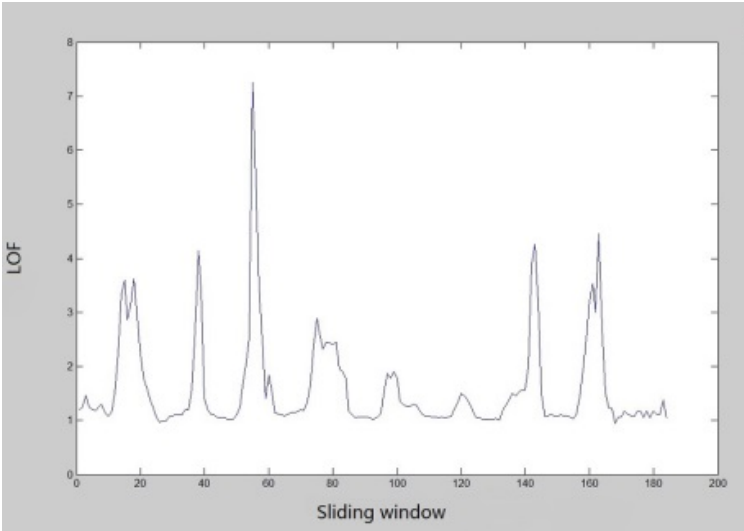
Our experiment results are shown as follows. In order to evaluate the segmentation performance of our proposed method, the correct cut points of motion sequences are labeled manually in advance. In Fig. (6), the LOF curve is shown where the segmentation sliding window will be extracted. And Fig. (7) shows the segmentation results by our method compared with PPCA segmentation proposed by Barbic [12] and the manual labeling.

Fig. (7) demonstrates the segmentation results by our method, which is close to the manual labeling and better than PPCA. And our method is robust in dealing with position and orientation. For example, in Fig. (7a), the PPCA method has tested the orientation change and detected an unnecessary cutting point at the beginning of the motion sequence while our LOF based method segmented correctly when the turning back motion exists in the same behavior.

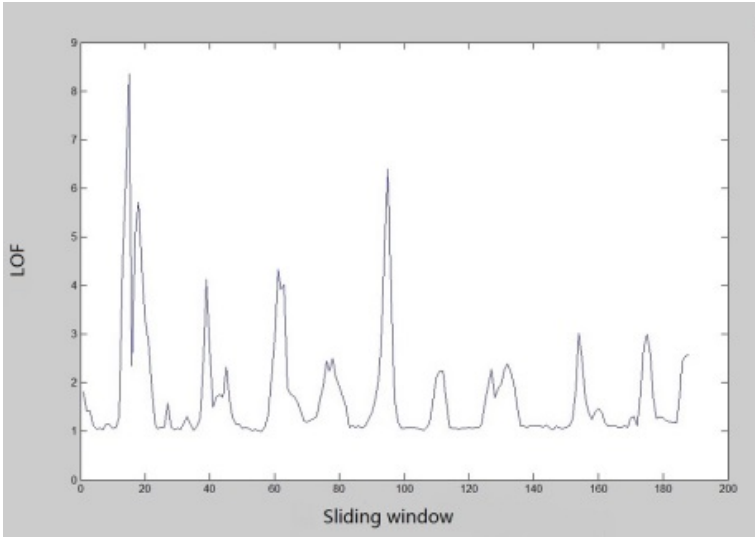
In order to further evaluate the performance of our method, precision and recall rate are used in our experiment. Precision and recall rate are defined as follows:

$$precision = \frac{N_{correct}}{N_{report}}, \quad (14)$$

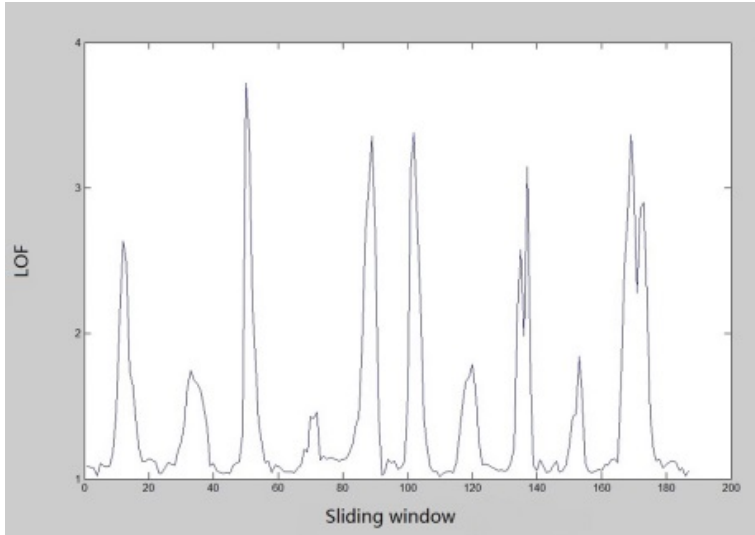
$$recall = \frac{N_{correct}}{N_{human}}. \quad (15)$$



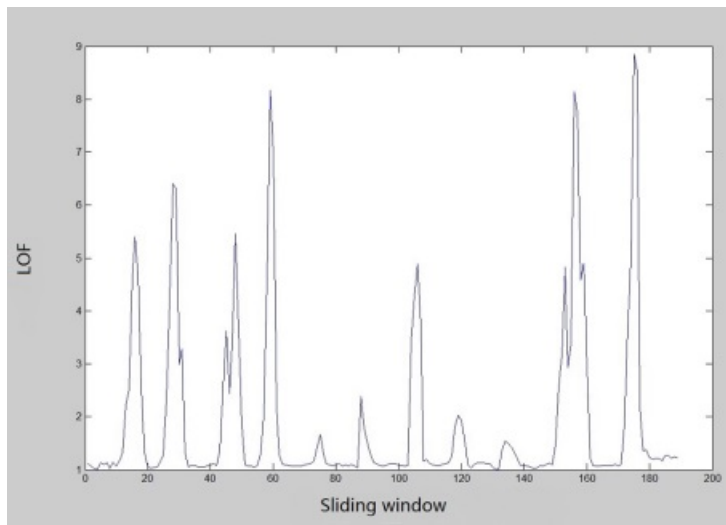
(a)



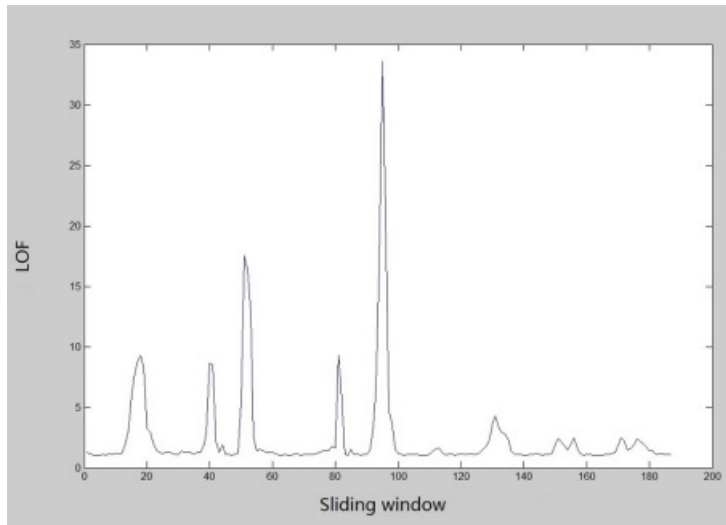
(b)



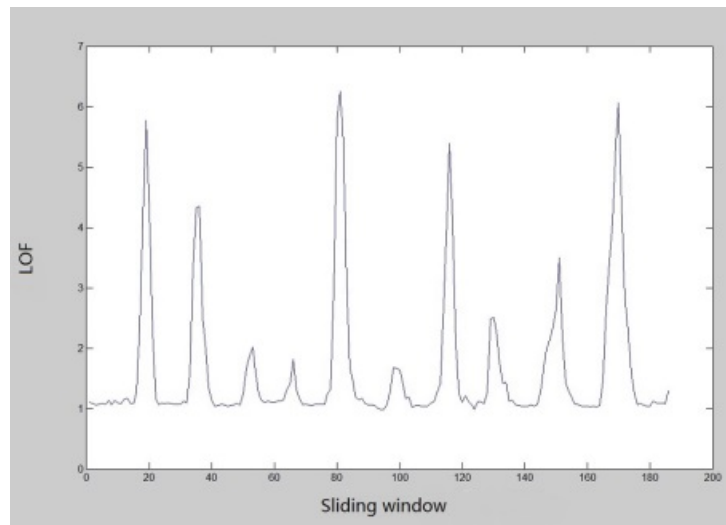
(c)



(d)



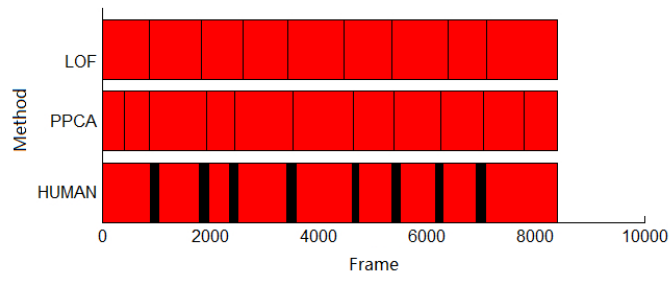
(e)



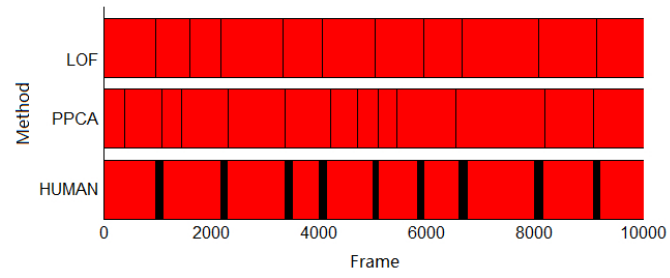
(f)

Motion sequences: (a) 86_03.amc; (b) 86_04.amc; (c) 86_05.amc; (d) 86_06.amc; (e) 86_07.amc; (f) 86_08.amc

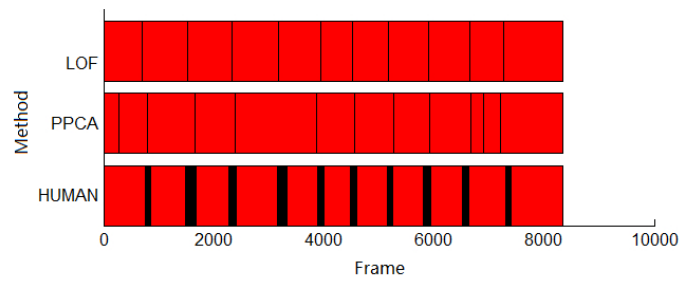
Fig. (6). LOF curve results.



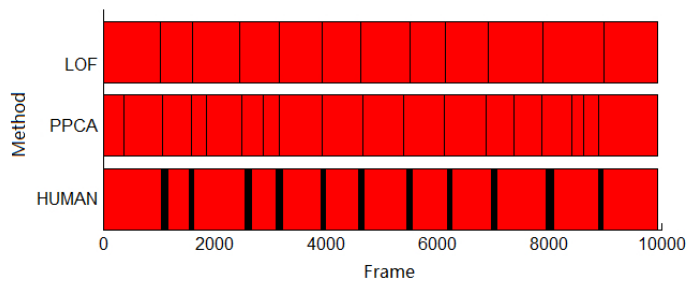
(a)



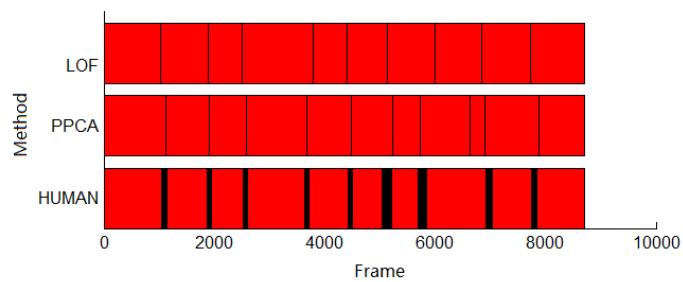
(b)



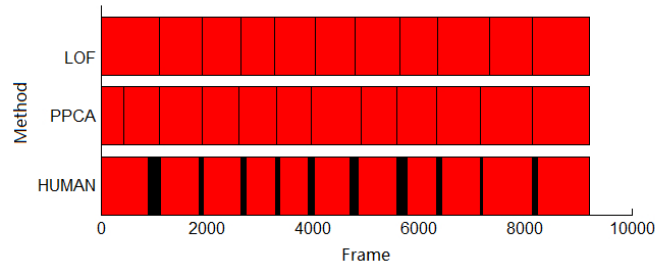
(c)



(d)



(e)



(f)

Motion sequences: (a) 86_03.amc; (b) 86_04.amc; (c) 86_05.amc; (d) 86_06.amc; (e) 86_07.amc; (f) 86_08.amc

Fig. (7). Comparison experiment results.

Table 4. Precision and recall of our method compared with PPCA.

Method	Precision	Recall
LOF	0.93	0.94
PPCA	0.82	0.92

Table 5. Time cost of our method compared with PPCA

File Time Method	86_03.amc	86_04.amc	86_05.amc	86_06.amc	86_07.amc	86_08.amc
LOF	1.16s	1.40s	1.16s	1.36s	1.27s	1.26s
PPCA	12.95s	19.60s	13.54s	24.89s	13.66s	14.37s

Where $N_{correct}$ is the total correct cutting point number of experiment results, and N_{report} is the total cutting point number of experiment results, and N_{human} is the actual cutting point number by manual labeling.

Table 4 gives the precision and recall rates of our experiments, which show that our method can achieve better performance than PPCA.

Additionally, the efficiency of our method is higher than PPCA in our experiment, especially in the long motion sequences. Table 5 shows the time cost of the two methods. Since there are two stages to do dimensionality reduction in our method while the distribution for each individual behavior needs to be estimated in PPCA, our method is less time consuming especially for long sequence.

6. CONCLUSION

In this paper, we propose a LOF based automatic motion segmentation method, which segment motion sequences into distinct motion clips automatically. In essence, it mainly contains two dimensionality reduction processes, extracting the bone angles for each frame and building the statistic

histogram for each sliding window. Then the follow-up stage of calculating the LOF for each sliding window is implemented and the reasonable cutting frame by peak of the LOF curve is obtained. This presented method is able to create a segmentation without any prior or training models, and thus is a kind of unsupervised learning method. Due to the two dimensionality reduction processes, the efficiency of our method is better even when dealing with very long motion sequences. LOF as a local density-based algorithm shows its effectiveness in finding the “bridge” between two successive kinds of behaviors. This method can be used whether a human behavior can be described by a Gaussian distribution or not. Experiments are conducted in CMU dataset, and the results show that our method can achieve better performance compared with the classical PPCA based segmentation method.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This work is supported in part by the National Natural Science Foundations of China (No.61100143, 61370128, and 61272353), Program for New Century Excellent Talents in University (NCET-13-0659), Fundamental Research Funds for the Central Universities (2014JBZ004), Beijing Higher Education Young Elite Teacher Project (YETP0583).

REFERENCES

- [1] CMU Graphics Lab, *CMU Motion Capture Database*. <http://mocap.cs.cmu.edu/>.
- [2] M. Müller, T. Roder, M. Clausen, B. Eberhardt, B. Krieger, and A. Weber, *Documentation Mocap Database HDM05*, Technical Report CG-2007-2, Universität Bonn, 2007.
- [3] L. Kovar, M. Gleicher, and F. Pighin, "Motion graphs," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 473-482, 2002.
- [4] K. Yin, K. Loken, and M. van de Panne, "SIMBICON: simple biped locomotion control," *ACM Transactions on Graph. (TOG)*, vol. 26, no. 3, Article 105, 2007.
- [5] Y. Lee, K. Wampler, G. Bernstein, J. Popovic, and Z. Popovic, "Motion fields for interactive character locomotion," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 6, pp. 1-8, 2010.
- [6] M. Müller, T. Roder, and M. Clausen, "Efficient content-based retrieval of motion capture data," *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 677-685, 2005.
- [7] M. G. Choi, K. Yang, T. Igarashi, J. Mitani, and J. Lee, "Retrieval and visualization of human motion data via stick figures," *Computer Graphics Forum*, vol. 31, no. 7, pp. 2057-2065, 2012.
- [8] M. Kapadia, I. Chiangy, T. Thomasz, N. I. Badlerx, and J. T. Kider Jr, "Efficient motion retrieval in large motion databases," In: *Proceedings of the Symposium on Interactive 3D Graphics*, Orlando, pp. 19-28, 2013.
- [9] M. Müller, and T. Roder, "Motion templates for automatic classification and retrieval of motion capture data," In: *Proceedings Eurographics/ ACM SIGGRAPH Symposium on Computer Animation*, Vienna, pp. 137-146, 2006.
- [10] J.D. Amor, and C.J. James, "Behavioral pattern detection from Personalized Ambient Monitoring," In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, pp. 5193-5196, 2010.
- [11] M.M. Breunig, H.P. Kriegel, R.T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Dallas, pp. 93-104, 2000.
- [12] J. Barbic, A. Safonova, J. Y. Pan, C. Faloutsos, J.K. Hodgins, and N.S. Pollard, "Segmenting motion capture data into distinct behaviors," In: *Proceedings of Graphics Interface*, Ontario, pp. 185-194, 2004.
- [13] T.S. Wang, H.Y. Shum, Y.Q. Xu, and N.N. Zheng, "Unsupervised analysis of human gestures," *Advances in Multimedia Information Processing*, vol. 2195, pp. 174-181, 2001.
- [14] M.J. Wang, W. Wang, Y. LI, and B. Qi, "Motion capture data segmentation using kernel dynamic texture," In: *International Conference on Audio Language and Image Processing*, Shanghai, pp. 592-596, 2010.
- [15] K. Kahol, T. Priyamvada, P. Sethuraman, and T. Rikakis, "Gesture segmentation in complex motion sequences," In: *Proceedings of IEEE International Conference on Image Processing*, Barcelona, pp. II-105-108, 2003.
- [16] O. Arikian, D.A. Forsyth, and J.F. O'Brien, "Motion synthesis from annotations," *ACM Transactions on Graphics(TOG)*, vol. 22, no. 3, pp. 402-408, 2003.
- [17] F. Zhou, F. Torre, and J. K. Hodgins, "Aligned cluster analysis for temporal segmentation of human motion," In: *Proceedings of International Conference on Automatic Face & Gesture Recognition*, Amsterdam, pp. 1-7, 2008.
- [18] J. Xiao, "Research on Some Key Techniques of Intelligent Human Motion," Ph.D Dissertation, Zhejiang University, Zhengjiang, China, 2007.
- [19] J. Xiao, Y.T. Zhuang, and F. WU, "Feature visualization and interactive segmentation of 3D human motion," *Journal of Software*, vol. 19, no. 8, pp. 1995-2003, 2008.
- [20] W.W. Xing, L.M. Tong, Y. Zhang, and C. Ren, "Human behavior segmentation method based on cosine distance for motion data," *Journal of Central South University (Science and Technology)*, vol. 45, no. 4, pp. 1128-1136, 2014.

Received: September 16, 2014

Revised: December 23, 2014

Accepted: December 31, 2014

© Weiwei et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.