Journal of Visual Languages and Computing 000 (2017) 1–10

Contents lists available at ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

Behavioral segmentation for human motion capture data based on graph cut method

Xiaomin Yu[a], Weibin Liu[a,1,*], Weiwei Xing[b]

[a] Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
[b] School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China

ARTICLE INFO

Article history:
Received 15 January 2017
Revised 7 September 2017
Accepted 7 September 2017
Available online xxx

Keywords:
Motion capture data
Behavioral segmentation
Graph cut
Spectral clustering

ABSTRACT

With the development of human motion capture, realistic human motion capture data has been widely implemented to many fields. However, segmenting motion capture data sequences manually into distinct behavior is time-consuming and laborious. In this paper, we introduce an efficient unsupervised method based on graph partition for automatically segmenting motion capture data. For N-Frame motion capture data sequence, we construct an undirected, weighted graph $G = G(V, E)$, where the node set V represent frames of motion sequence and the weight of the edge set E describes similarity between frames. In this way, behavioral segmentation problem can be transformed into graph cut problem. However, traditional graph cut problem is NP hard. By analyzing the relationship between graph cut and spectral clustering, we apply spectral clustering to the NP hard problem of graph cut. In this paper, two methods of spectral clustering, t-nearest neighbors and the Nystrom method, are employed to cluster motion capture data for getting behavioral segmentation. In addition, we define an energy function to refine the results of behavioral segmentation. Extensive experiments are conducted on the dataset of multi-behavior motion capture data from CMU database. The experimental results prove that our novel method is robust and effective.

1. Introduction

Human motion capture, once associated with producing special effects for film and television production, entertainment games and virtual reality, is common today in diverse applications ranging from health care to consumer electronics. The data of human motion capture can be extensively applied in many fields, such as producing realistic animation movies [1], physical rehabilitation [2] and ergonomic analysis [3]. However, the cost of capturing motion data is very high. It is necessary to better reuse motion capture data. Motion capture data sequence is usually comprised of multiple types of behaviors. The present work suggests that automatic segmentation of human motion capture data into distinct behaviors based on statistical properties of the motion can be an efficient and quite robust alternative to hand segmentation. This paper focuses on efficient and robust technique which can be able to automatically segment motion capture data sequence into distinct behaviors, as depicted in Fig. 1. According to the behavioral segmentation diagram, it is obvious that both sides of the red human skeleton model behave differently and it shows the main goal of this paper. Human motion capture data is segmented based on their behavioral semantics, such as walking, running, jumping and so on.

The main contributions of this paper can be summarized as follows:

(1) We construct N-Frame motion capture data sequence into an undirected, weighted graph $G = G(V, E)$. In this way, we transform motion capture data segmentation problem into the problem of graph cut.
(2) We employ an algorithm, which can automatically extract the number of behavior and cluster centers, to the motion capture data. Itis based on an unsupervised cluster method. We consider the cluster numbers as the behavior numbers for motion capture data sequence.
(3) Analyzing graph cut and spectral clustering, we transform this NP hard problem into eigenvalues and eigenvectors in spectral space.
Two methods of spectral clustering, t-nearest neighbors and the Nystrom method, are employed to cluster motion capture data for realising behavioral segmentation.

* Corresponding author.
E-mail address: wbliu@bjtu.edu.cn (W. Liu).
[1] DOI reference number: 10.18293/DMS2016-016

http://dx.doi.org/10.1016/j.jvlc.2017.09.001
1045-926X/

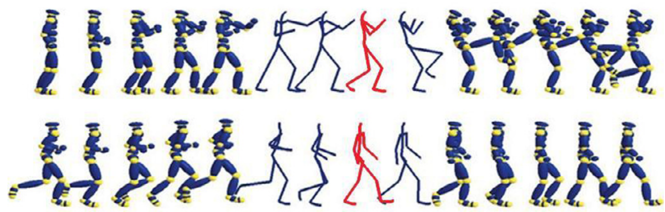

Fig. 1. Segmenting motion capture data into distinct behaviors. It's obvious that both sides of the red human skeleton model behave differently.

(4) For the clustering fragments, we define energy function and use dynamic programming to refine the results of behavioral segmentation.

The remainder of this paper is organized as follows:

Firstly, we introduce works of predecessors for motion segmentation in Section 2. How to calculate similarity of motion capture data is detailed introduced in Section 3. The details for identifying behavior's cluster centers for original motion capture data sequences are presented in Section 4. In Section 5, we analyze in detail the relationship between graph cut and spectral clustering. Then, transforming this NP hard problem of graph cut into spectral clustering problem. In this way, realizing behavioral segmentation for motion capture data. Details of energy function for refining the results of behavioral segmentation are represented in Section 6. In Section 7, we introduce the experiments, which are conducted on the dataset of multi-behavior motion capture data from CMU database. Finally, we provide the discussion and conclusion for our method in Section 8.

2. Related work

Temporal segmentation is related to numbers of different fields such as data mining [4], behavior recognition [5], and so on. Researchers have proposed several techniques to segment motion capture data into distinct behaviors. The methods can be generally categorized into several types: classifier, clustering, machine learning and dimension reduction and so on.

Supervised learning methods generally formalize segmentation motion capture data as a classification problem, where classifiers are trained from a carefully selected training set. In other words, it means that they are usually applying example segments or precomputed templates to test sequences. Muller et al. [6] constructed motion templates to behavioral segmentation. Lv et al. [7] defined Hidden Markov Models (HMM) to deal with behavioral segmentation. Support Vector Machine(SVM), which is based on an annotation training database to segment motion capture data was constructed by Arikan et al. [8]. However, these methods would fail to pick up segments whose corresponding behaviors were not contained in training set.

Another technique, which is named as unsupervised learning method, is used to over come supervised learning method's limitation. It's natural for researchers to use another technique to overcome this limitation, which can be named as unsupervised learning methods. In these methods, motion capture data segmentation is located by clustering motion frames. Zhou et al. [9] used aligned cluster analysis (ACA) to temporally cluster poses into motion primitives which were assigned to different behavior classes. ACA extends standard kernel k-means clustering: the cluster means include a number of features and a dynamic time warping (DTW) kernel is used to achieve temporal invariance. However, ACA method needs users determine the cluster number with respect to temporal constraint. Zhou et al. [10] derived an unsupervised hierarchical bottom-up framework, which is called hierarchical aligned cluster analysis (HACA) to realize behavioral segmentation. HACA provided a crude method to find a low-dimensional embedding for the time series. HACA is efficiently optimized with a coordinate descent strategy and dynamic programming.

Researchers tried to solve the dilemma of non-learning way to behavioral segmentation. Balazia et al. [11] introduced an unsupervised key-pose detection algorithm for segmentation of motion capture data and this proposed algorithm partition motions at the level of gestures. Barbic et al. [12] chosed segments using an indication of intrinsic dimensionality from Principal Component Analysis (PCA). It's based on the observation that simple motions exhibit lower dimensionality than more complex motions. As an extension of the traditional PCA, Barbic et al. [12] defined a proper probability model for PCA which is named probabilistic PCA (PPCA). We can easily know that the directions outside the subspace were discarded, whereas they were modeled with noise in PPCA. Zhanzhan et al. [13] proposed automatic segmentation approaches using Principle Geodesic Analysis (PGA) and using Probabilistic Principle Geodesic Analysis (PPGA). The segmentation method for human motion capture data, which is using PGA, assigned a cut when the intrinsic dimensionality of a local model of the motion suddenly increased. The method using PPGA placed behavioral segmentation when the distribution of poses is changed. Wei et al. [14] proposed behavioral segmentation method for human motion capture data via angle histograms. This method was based on the assumption that the same type motion had similar histogram of angle between bones. Wei et al. [15] presented a novel symbolic representation of human motion capture data, which is called Behavior String (BS). They analyzed behavior strings for original human motion capture data and realized behavioral segmentation and extracted the cycles of motion. Wei et al. [16] realized behavioral segmentation for human motion capture data based on Local Outlier Factor (LOF). This method is based on the assumption that the same type motion can form a cluster and the transition of two behaviors is the outlier of the two clusters. Zhang et al. [17] proposed the motion sequence segmentation based on Locally Linear Embedding (LLE) algorithm. Firstly, this method reduced the dimension of high dimension motion sequence to obtain one-dimension feature curve. Then, using the feature curve to achieve behavioral segmentation for human motion capture data. Bouchard et al. [18] used Laban Movement Analysis (LMA) to realize behavioral segmentation for human motion capture data. This method found motion sequences which exhibit high output similarity from a collection of neural networks trained with temporal variance.

To summarize, we have introduced some existing methods of behavioral segmentation for human motion capture data. To solve the shortcoming and limitations of existing behavioral segmentation methods, this paper proposes an efficient unsupervised method based on graph partition. Generally speaking, existing methods need the user to set the number of behaviors for the sequences of human motion capture data. However, this algorithm of us can automatically extract the number of behaviors for original motion capture data sequences. According to this, this method is to the maximum extent to reduce human intervention. Finally, this paper defines energy function and use dynamic programming to refine the results of behavioral segmentation. It can improve behavioral segmentation performance and robustness of our algorithm. In order to better understand the behavioral segmentation method, which is proposed by this paper, Fig. 2 shows main procedures.