GRAPH-BASED APPROACH FOR MOTION CAPTURE DATA REPRESENTATION AND ANALYSIS

Jiun-Yu Kao, Antonio Ortega, Shrikanth S. Narayanan

University of Southern California
Department of Electrical Engineering

ABSTRACT

Providing better representation methods for motion capture data can lead to improved performance in terms of classification, recognition, synthesis and dimensionality reduction. In this paper, we propose a novel representation method inspired by algebraic and spectral graph theoretic concepts. Our proposed method represents motion data in a space constructed with bases for skeleton-like graphs. We introduce two criteria and as well as its influences on the generated bases. With experiments on CMU *MoCap* database, we will also discuss how this method may act as a good preprocessing tool in order to enhance the further analysis steps on *MoCap* data.

Index Terms— Human motion, motion capture, graph-based approach, gait analysis, dimensionality reduction

1. INTRODUCTION

For human motion capture data, as well as for many other real-world data, dimensionality reduction is a useful preprocessing tool that leads to more robust analysis. Dimensionality reduction should be a transformation of high-dimensional data space into a lower-dimensional space constructed with meaningful representations for the data. An ideal reduced dimensionality should be the intrinsic dimensionality of the data i.e., the minimum number of parameters needed for explaining properties of the data [1].

Dimensionality reduction methods can be divided into two major types, linear and non-linear. Traditionally, principal component analysis (PCA) is considered to be the best, in the mean-square sense, linear dimensionality reduction technique [2]. Dimensionality reduction with PCA and its variants perform well in many applications using motion capture (*MoCap*) data, such as motion classification, similar motion search [3] and segmentation [4]. However, PCA does not perform as well in some other kinds of applications, such as synthesis of motion sequences [5].

Our proposed method is a novel linear approach to perform dimensionality reduction for human *MoCap* datasets. Our goal is to construct a transformation from high-dimensional data space to lower-dimensional feature space. We specifically take the nature of human skeleton and human motion into consideration. We focus on a new representation for motion data that leads to an interpretation of observed characteristics, such as rigidity and coordination of the body parts within captured motions. It is also worth mentioning that the set of features proposed possesses a skeleton-like structure that can help understanding motions as well.

1.1. Related work

MoCap data can be used to construct a PCA space in some applications such as motion synthesis [5][6][7]. However, in these cases, the construction of PCA space does not take into account the nature of human motion or the natural constraints induced by human skeleton [3]. Skocaj and Leonardis present a weighted PCA scheme for cognitive vision system, wherein different weights may be applied to the subareas of an image [8][9]. Inspired by Skocaj and Leonardis's work, Forbes and Fiume apply weighted PCA to motion data [3]. They use weights derived from the relative amount of body mass that is influenced by the movement of each joint, and manipulate specific weights to change the properties of the resulting space. As an extension of PCA, Tournier *et al.* [15] proposed a method based on principal geodesics analysis (PGA) which exploits spatial and temporal redundancy from motion data. Li *et al.* [16] models the motion data as a sparse linear combination of a set of learned basis functions activated at different time. Hou *et al.* [17] represents the motion data as 3rd-order tensors which are generated by stacking mocap sequence locally. To sum up, the feature space construction in all of the above methods is learned from the motion data itself.

What we propose here is to construct the space without applying motion data but applying natural structure in human skeleton and motions. The motion data is regarded as the graph signal carried by the human skeleton (graph). This method constructs the reduced feature space in a completely non data-driven manner. Thus, a noisy dataset will not affect the construction of the reduced feature space. Its advantage will be lower computational cost because no updates are required when the input motion dataset changes. It is also advantageous for analyzing the interactions between body parts in some specific motions.

Fig. 1. (a) The skeletal structure used. (b) Graph construction. Numbers are indices of vertices. A black line between s and t means that an edge exists between the corresponding vertices.

2. PROPOSED METHOD

To state the problem mathematically, assume a given motion sequence represented as a sequence of frames. Each frame is represented by specifying the three-dimensional coordinates for all the selected joints in the body. In the following sections, each frame M_i $(i = 1, 2, ..., n)$ is represented as a 15-by-3 matrix because the coordinates of 15 joints, as marked in Fig.1, are measured and collected. $M_{i,j}$, the j^{th} row of M_i, is the three-dimensional coordinates of the j^{th} joint at that particular time i. This is our current approach and dataset, but this could be extended to a larger number of points, e.g., some not representing joints.

2.1. Graph construction and spectrum

The motion sequence described is a multi-dimensional dataset that naturally resides on the vertices (joints) of the skeletal structure (graph). Inspired by [10], we model the human skeletal structure shown in Fig.1 as a fixed undirected graph $G = (V, E)$. The vertex set is $V = \{v_1, v_2, ..., v_{15}\}$ because the coordinates of 15 joints are specified in each frame. As for the set of edges, we can make various choices in order to select an appropriate formulation for G. Here we assume that the appropriate set of edges E has been selected. We will discuss the selection for an appropriate edge set with the criteria stated in *Section 2.2*. After G is decided, the adjacency matrix A and the degree matrix D of G can be computed. The normalized Laplacian matrix L can be computed as well with $L = I - D^{-1/2}AD^{-1/2}$.

The spectral domain of graph G can be described by its spectral basis and corresponding spectral frequencies. The spectral basis is the eigenvectors of L denoted by $U = \{u_k\}, k = 1, ..., 15$. The spectral frequency is the eigenvalues of L associated with U denoted by $\sigma(G) = \{\lambda_1, \lambda_2, ..., \lambda_{15}\}$ where $0 = \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{15}$. The eigen-pair system $\{(\lambda_k, \mu_k)\}$ provides Fourier interpretation for graph signals on G.

2.2. Criteria for choices of edge set

First, without prior knowledge about natural structures in specific motions, we can choose to connect the vertices according to the natural skeleton structure shown in Fig.1. The existence of edge a_{st} between vertex s and t in G corresponds to the connection between joint s and t in the human skeletal structure. For instance, there is a natural connection between left shoulder and left elbow, so that an edge exists between v_4 and v_5, i.e., $a_{45} = a_{54} = 1$. It is worth mentioning that the constructed graph is unweighted.

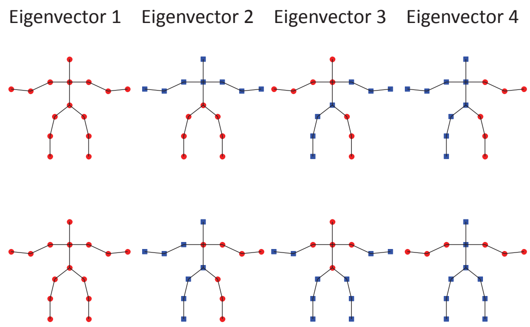

Fig. 2. Sign value of eigenvectors on the constructed graph G : blue-square (+), red-dot (-). Top row: graph formulation as shown in Fig.1. Bottom row: graph formulation as shown in Fig.5(a). Notice that zero-crossings between neighboring vertices increase as the eigenvector corresponds to higher eigenvalue (frequency).

Alternatively, we can use prior knowledge about the motion of interest, we can choose the edge set according to it. For example, bilateral symmetry is regarded to be common in normal human gait [13]. To better take advantage of this, we add more symmetric connectivity between body parts. Specifically take walking motion as an example, given the bilateral symmetry existing between the left and right part of the body, we can construct the edge set as shown in Fig. 5, which makes vertices more connected in either left or right body part.

This alternative fomulation for G can lead to change in basis structure. It will encourage more sign changes symmetrically accross the center.It The bases also have better symmetry, which will be shown with experiments in *Section3.3*.

2.3. Proposed features

The spectral basis $\{u_k\}, k = 1, ..., 15$ forms a basis for any graph signal residing on G. That is, any graph signal can be represented as a unique linear combination of $\{u_k\}$ as in (1).

$$\mathbf{M_i} = \sum_{k=1}^{15} \alpha_{k,i}\mu_\mathbf{k} \tag{1}$$

where

$$\alpha_{k,i} = \mathbf{M_i} \cdot \mu_\mathbf{k} = \sum_{k=1}^{15} m_{i,j}\ \mu_{kj} \tag{2}$$

μ_{kj} is the j^{th} entry in μ_k while $\alpha_{k,i}$ is a vector with length 3.

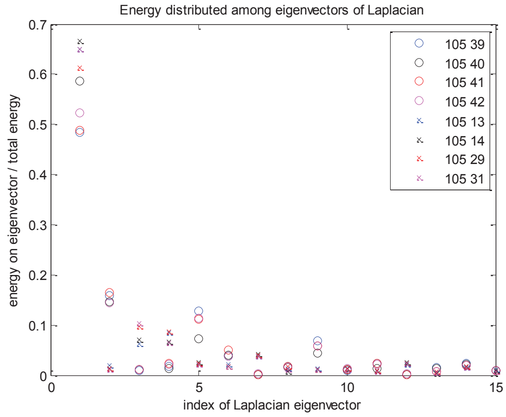

Fig. 3. Energy distribution over Laplacian eigenvectors in each motion task; walking:*cross*; jumping:*circle*.

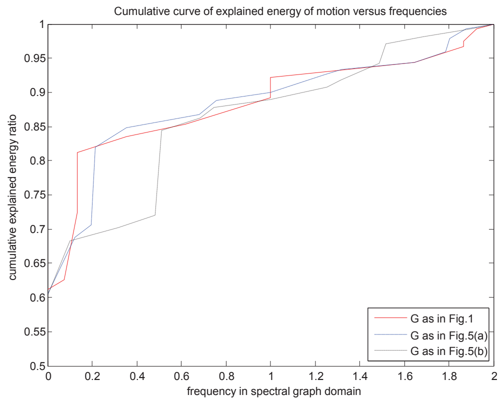

Fig. 4. Cumulative curve of ratio of explained energy versus frequencies. a: with formulation in Fig.1; b: with formulation in Fig.5(a); c: with formulation in Fig.5(b).

Therefore, $(\alpha_{1,i}, \alpha_{2,i}, ..., \alpha_{15,i})$ can act as an unique representation for a given frame m_i.

The basis vectors $\{u_k\}$ used here will vary with different choices for E according to either criterion stated in *Section 2.2*. The sign values of the first 4 out of totally 15 eigenvectors are plotted in Fig. 2. The top row of eigenvectors is associated with graph constructed from natural human skeleton (as in Fig.1) while the bottom row is associated with graph constructed with prior knowledge in bilateral symmetry in walking (as in Fig.5(a)). We can notice that, by connecting vertices and formulating graph according to prior bilateral symmetry assumption, we change the structure of eigenvectors. For example, the 2^{nd} eigenvector, related to the min-cut of graph, changed to be symmetric between left and right body rather than between upper and lower body, which can capture more bilateral symmetric movements within body during walking. The experiments with the usage of each criterion are presented in *Sections 3.1* and *3.2*, respectively.

Fig. 5. Two different constructions for G.

3. EXPERIMENTS

We use Carnegie Mellon University Graphics Lab Motion Capture Database [14] to evaluate our representation. We choose the motions under "*Locomotion*" category to use. Specifically, we use "*walking*" and "*jumping*" motions to examine our method. The following table shows the index of motion clip and its correspondent motion task.

Walking	105_13	105_14	105_29	105_31
Jumping	105_39	105_40	105_41	105_42

The first experiment is applying our method to both "*walking*" and "*jumping*", to see if the key information embedding under different motions can be reserved. We will also discuss whether this method can be a good preprocessing method for the possible classification steps going next. The second experiment is discussing how the important step in our method, the formulation of graph G, may change the effective basis and interpretations for body correlation within motions. We will see several different formulations for graph G and examine the influence of them on the representation basis as well as the provided interpretations.

3.1. Separation between different tasks

First we construct the graph G as shown in Fig.1. With this construction, the corresponding Graph Fourier Transform basis is illustrated in Fig.2. In the experiment, we take 8 motion clips from CMU *MoCap* database, 4 with walking tasks and another 4 with jumping tasks as indexed in table above. Within each task, we project each frame M_i onto the basis $\{u_k\}, k = 1, ..., 15$ with (1). The projection on each out of the 15 eigenvectors results in a vector with dimension three. We take the length square of these projections and measure the ability of each eigenvector to explain the original motion data.

The results in Fig.3 shows that all the walking tasks are separated quite well from all the jumping tasks. It also worth mentioning that the two different motions can be well separated by only the first several eigenvectors such as the $2^{nd}, 3^{rd}, 4^{th}$ and 5^{th} one. Notice that as shown in Fig.2, the 2^{nd} eigenvector can reflect the correlation between upper and lower body while the 3^{rd} and 4^{th} eigenvector can reflect the symmetry or bilateral symmetry embedded under human motions. Take the walking tasks for example. Work in the literature has shown that natural human walking motion possesses bilateral symmetry [11][12] while jumping motion

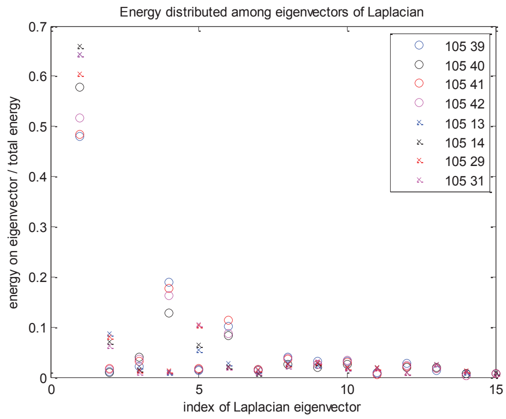

Fig. 6. Energy distribution over each Laplacian eigenvector within each motion task with graph formulation as Fig.5(a)

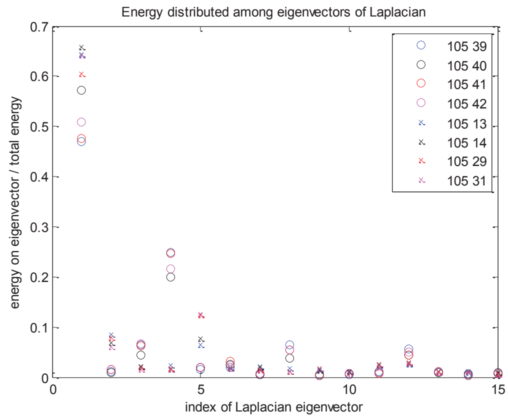

Fig. 7. Energy distribution over each Laplacian eigenvector within each motion task with graph formulation as Fig.5(b)

does not. This fact corresponds to the results shown in Fig.3 which indicate that walking motions have a larger component projected on the 3^{rd} and 4^{th} eigenvectors. The method shows to be able to preserve and reveal the coordination mechanism underlying natural human motions. At the same time, this representation method acts as a preprocessing step for further dimensionality reduction and plausible classifying needs. Fig.4 shows that almost 90% energy in movements can be captured by the several low-frequency eigenvectors.

3.2. Influences from various choices of G

In the second experiment, we examine the influence to the results from using alternative formulations for G. Follow by the same experiment as in *Section 3.1*, we use two alternative formulations for G as shown in Fig.5. We add more extra edges symmetrically to both body sides.

Let us take a look at the results of cumulative curve of explained movement energy versus graph spectrum in Fig.4. With the same graph signals, i.e. motion data, the result with graph formulation c contains more components in higher frequencies. This makes sense because this formulation (Fig.5(b)) adds more connections within the same side of the body, between upper and lower limbs. Since these in a natural walking motion, tend to move oppositely, this leads to a larger component in higher frequency when representing motion data in a space constructed by formulation c. Fig.6 and Fig.7 also show that,with modification on graph structure, we can rearrange the distribution of energy over eigenvectors in a way better for the following analysis steps. For example, we make energy concentrated more in the 4th eigenvector in Fig.7 than in Fig.6.

3.3. Comparison with PCA

In order to compare the degree of symmetry within the generated bases between our approach and PCA, we define a metric $symm$ to measure the level of symmetry in the basis vectors, as follows:

$$symm = \sum_{\substack{(i,j)\in pair\ of \\ symmetric\ nodes}} \left| \frac{(V(i)-\bar{V})(V(j)-\bar{V})}{Var(V)} \right| \quad (3)$$

where V is the basis vector. The maximal value of metric $symm$ would be 6 because there are 6 pairs of symmetric vertices in our setting.

The mean of $symm$ from all 15 basis vectors generated with our approach with graph formulated respectively as Fig.1, Fig.5(a) and Fig.5(b) are 5.3236, 5.5245 and 5.4215, which reveals a highly symmetric bases generated with our approach. The mean of $symm$ from all basis vectors generated with PCA approach in the use of motion data "*105_29*" from CMU *MoCap* database is 3.8622, 3.9818 and 4.1016, corresponding to 3 dimensions in movements respectively. The results show that our approach can generate a more symmetric basis to represent *MoCap* data when taking into account the nature in human motion.

4. CONCLUSIONS

In this paper, we propose a novel representation for motion capture data. We apply graph theory and treat the motion data as graph signals resided on the graph we construct. The basis we use here is the set of Laplacian eigenvectors which can preserve and reveal the embedding structure in natural human motions. The preprocessing for the motion data can lead to a better input for the following processing steps such as dimensionality reduction and classification. Furthermore, we can modify our graph formulation in order to improve the interpretation of the basis in terms of motion and also enhance performance in the following analysis steps.

This novel representation may be applicable to many scenarios in human motion and gait analysis. It can act as a good preprocessing tool for the following analysis steps. It can also help researchers in bio-kinesiology to explore and demonstrate the embedding structure under specific motions.

5. REFERENCES

[1] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press Professional Inc., San Diego, CA, USA, 1990

[2] I.T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986

[3] K. Forbes, E. Fiume, "An Effcient Search Algorithm for Motion Data Using Weighted PCA," *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, pp.67-76, 2005

[4] J. Barbic, A. Safonova, J.Y. Pan, C. Faloutsos, J.K. Hodgins, N.S. Pollard, "Segmenting Motion Capture Data into Distinct Behaviors," *Proceedings of Graphics Interface 2004*, pp. 185-194, July 2004

[5] L. M. Tanco, A. Hilton, "Realistic Synthesis of Novel Human Movements from a Database of Motion Capture Examples," In *IEEE Workshop on Human Motion*, 2000

[6] Y. Li, T. Wang, H.-Y. Shum, "Motion Texture: A Two-level Statistical Model for Character Motion Synthesis," In *Proceedings of ACM SIGGRAPH 2002*, ACM Press, New York, USA, pp.465-472, 2002

[7] M. P. Johnson, *Exploiting Quaternions to Support Expressive Interactive Character Motion*, PhD thesis, MIT, 2003

[8] D. Skocaj, A. Leonardis, "Weighted Incremental Subspace Learning," In *Proc. Workshop on Cognitive Vision*, Zurich, Switzerland, 2002

[9] D. Skocaj, A. Leonardis, "Weighted and Robust Incremental Method for Subspace Learning," In *ICCV'03*, pp. 1494-1501, 2003

[10] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, P. Vandergheynst, "The Emerging Field of Signal Processing on Graphs: Extending High-dimensional Data Analysis to Networks and Other Irregular Domains," In *IEEE Signal Processing Magazine*, pp.83-98, May 2013

[11] D. J. Goble, G. W. Marino, J. R. Potvin, "The Influence of Horizontal Velocity on Interlimb Symmetry in Normal Walking," *Human Movement Science*, pp. 271-283, Aug. 2003

[12] J. M. Haddad, R. E. van Emmerik, S. N. Whittlesey, J. Hamill, "Adaptations in Interlimb and Intralimb Coordination to Asymmetrical Loading in Human Walking," *Gait and Posture*, pp. 429-434, June 2006

[13] D. J. Goble, *The Influence of Horizontal Walking Velocity on the Bilateral Symmetry of Normal Ground Reaction Force Parameters*, Thesis, University of Windsor, 2002

[14] The data used in this project was obtained from mocap.cs.cmu.edu. The database was created with funding from NSF EIA-0196217.

[15] M. Tournier, X. Wu, N. Courty, E. Arnaud, L. Reveret, "Motion Compression using Principal Geodesics Analysis," *Comput. Graph. Forum*, vol. 28, no. 2, pp. 355-364, Apr. 2009

[16] Y. Li, C. Fermuller, Y. Aloimonos, H. Ji, "Learning Shift-Invariant Sparse Representation of Actions," *Proc. IEEE CVPR*, pp. 2630-2637, 2010

[17] J. Hou, L. P. Chau, N. Magnenat-Thalmann, Y. He, "Scalable and Compact Representation for Motion Capture Data Using Tensor Decomposition," *IEEE Signal Processing Letters*, vol. 21, no. 3, pp. 255-259, March 2014