Motion Interpretation and Synthesis by ICA

Renqiang Min

Department of Computer Science, University of Toronto, 10 King's College Road, Toronto, ON M5S3G4, Canada

Abstract. It is known that high-dimensional human motion data lies in a low dimensional space. Looking for independent sub-motions that contribute to generating complicated human motions is an interesting task on its own merits. Independent Component Analysis (ICA) can extract independent factors of an observed signal. In this paper, ICA is applied to the reconstructed human motion data by PCA. Experimental results show that, if the low dimenality is appropriately chosen, ICA can find meaningful sub-motions. Independent sub-motions make it easy to complete missing human motion data. Preliminary results on synthesizing missing human motion data by self-regression using Gassian Process will also be discussed in the paper.

Key words

Human Motion Interpretation, Motion Synthesis, Independent Componenet Analysis, Gaussian Process

1 Introduction

In computer graphics, a lot of motion capture data has been recorded for studying different kinds of human motions such as walking, running, jumping, and dancing ect. These raw data is often high-dimensional in their original formats due to the large number of markers. And it is known that these high-dimensional data often lies in a low dimensional space. PCA is often used to do dimensionality reduction thereby to get a compact representation of the original motion data. However, the low dimensions calculated by PCA only account for the variance of the data on some orthogonal directions. Sometimes, we are interested in a small set of independent factors that contribute to generating different kinds of human motions. For such tasks, ICA becomes a natural choice.

Obtaining independent factors of human motion data is interesting for several reasons. First, independent factors suggest possible independent sub-motions that contribute to generating complicated motions. It gives us meaningful interpretations of human motions. This task is interesting on its own merits.

Second, it makes motion style transfer easy. Given some different styles of motions, we can extract the corresponding independent components that are specific to respective styles. By replacing the set of style-specific independent components in one motion with a desired set of style-specific components, we will get the motion with the desired style, which has been done by [**?**]. We can also arrange the motion data with each arrangement dimension representing different features, for e. g., features can be identity and style. Then we can apply multi-linear ICA to the data to extract feature-specific indpendent factors for style transfer. A similar idea for doing face transfer using multi-linear PCA is described in [3].

Third, obtaining independent factors of human motion data suggests new ways of synthesizing and predicting missing segments of motions. For some complicated motions like dancing, if some segments of the data are missing, it is hard to synthesize and predict them directly using a bruteforce approach. Nevertheless, if the orignal motion data can be decomposed into several independent sub-motions, especially periodic, we can handle each sub-motion seperately, and it will make the synthesis/prediction task much easier. In contrast, most existing approaches to motion synthesis/prediction take the high-dimensional motion as a whole and generate missing motions by approximation. Some recent algorithms for motion synthesis create novel motions by choosing pieces from a motion database and putting them together with smoothing to form a new motion [**?**]. In the picking pieces process, some constraints such as continuity constraints, path constraints, frame constraints, and ect, might be satisfied depending on different goals. All of these systems are based on subsequence matching. A recent system decribed in [6] is very successful in synthesizing motions, it is still a matching based system but implemented in a hierarchical fashion. In [7], a two-layer time-series graphical model based on Gaussian Processes (GP) is used to model the latent low dimensional dynamics of motions, and its application to fitting missing data and to predict short range subsequent motions is discussed. In this paper, some preliminary results for synthesizing missing motion data by self-regression for each independent component will be presented. Although the results are not very good, I believe that powerful methods can be used to produce good results by making better use of the advantage given by the independence of sub-motions.

This paper is organized as follows: in section 2, I will briefly compare PCA and ICA first, then I will describe in details the approach to human motion interpretation based on PCA and ICA and motion synthesis based on independent components; in section 3, I will present some experimental results about the interpretation of different styles of walking; in section 4, I will show the preliminary results on synthesizing missing data using self-regression; I will conclude the paper with some disscussions and future work in section 5.

2 Motion interpretation and synthesis using PCA and ICA

2.1 PCA and ICA

PCA produces a linear mapping by generating orthogonal directions preserving the variance of data as much as possible. Suppose there is a motion capture data set $X = [\mathbf{x^{(1)}}, \cdots, \mathbf{x^{(N)}}]$, where $\mathbf{x^{(i)}}$ is an n-dimensional column vector, n is number of markers, and N is the number of frames. The mean of the data vectors is μ. Suppose V contains the eigenvectors of the covariance matrix of X, we can get the projection $\mathbf{y}$ of a data vector $\mathbf{x}$ as follows:

$$\mathbf{y} = \mathbf{V}(\mathbf{x} - \mu) \tag{1}$$

If $\hat{V}$ contains the eigenvectors corresponding to the largest several eigenvectors, we can reconstruct a high-dimensional data vector $\hat{\mathbf{x}}$ from a low-dimensional vector $\hat{\mathbf{y}}$ as follows:

$$\hat{\mathbf{x}} = V^T \hat{\mathbf{y}} + \mu \tag{2}$$

Standard ICA also gives a linear mapping. Suppose we also have the data matrix X above, ICA assumes that the data matrix X is generated by mixing some latent components having independent non-Gaussian distributions. Unlike that PCA can only find correlations between components of data matrix, ICA can capture higher-order statistics in the data. ICA unmixes the data matrix X as follows:

$$X = AS \tag{3}$$

Where A is the mixing matrix and S contains the independent components. There are several approaches to implementing ICA from different pointviews such as by minimizing mutual information between different components and by projection pursuit which minimizes the non-Gaussianity of each component. In this paper, FastICA [8], which is an iterative procedure combining the above listed pointviews, is used to extract independent factors of human motion data.

2.2 Motion interpretation

ICA can be applied directly to the motion data to extract independent components as described in section 2.1. Due to the high-dimensions of the original data, squre ICA will output a lot of independent factors, in which each factor only conveys a tiny bit of information. Thus, it is very hard to interpretate each independent factor. And because the factors are independent to each other, even if k-mean clustering is run on them, meaningful interpretation for each cluster still cannot always be obtained. If under-complete ICA is applied, the obtained independent components cannot reconstruct the original motion data well, and it makes it hard for motion synthesis to use the obtained independent components.

In this paper, PCA is applied to the motion data X to do dimensionality reduction first. In this phase, the most important information is preserved and some noisy information is removed. The reconstructed motion data is claculated by Equation 2. In the second phase, ICA is applied to the reconstructed motion data generated by PCA. Since the reconstructed data lies in a low dimensional space, it is very easy for ICA to unmix the reconstructed data to extract some independent components S. These independent components often correspond to meaningful interpretations as shown in section 3.

2.3 Synthesize and predict missing motion data

Given a sequence of human motion data with some missing segments, we can synthesize and predict the missing data using the decomposed independent components. Since the components are independent and each component represents a sub-motion, we can consider one component one time during the synthesis. The extracted components are often periodic, thus, the missing segments of each component can be predicted by k-NN or by parametric periodic functions.

In this paper, we will try to use a simple idea to predict the missing segments. On each component, we learn a self-regression model based on Gaussian Process (GP). That is, for each component, we get many input-output pairs (C_t, C_{t+1}) with the component value at time t as input and the component value at t+1 as output, where $t = 1, \cdots, N-1$. N is the total number of frames.

3 Experimental results on human motion interpretation

In the experiments, the CMU motion capture dataset is used. The dimensionality of the original motion data is 62. At each frame, the coordinates and the oritentations of the root and the joint angles are recorded. Experiments on two styles of walking are performed to extract meaningful submotions. When the dimension corresponding to the tranlation of the root along a line is left unchanged, several independent factors seem to explain the main translation of the root. In all the subsequent experiments discussed, the main translation of the root is set to zero in order to focus on the movement of the joints.

Using the approach described in section 2.2, when the dimensionality of a sequence of walking data is reduced to 10, each independent component (IC) extracted by ICA seems to correspond to a meaningful sub-motion of walking. IC1 corresponds to moving the right foot backward periodically. IC2 corresponds to the preparation stage for moving the left foot backward. IC3 corresponds to moving the left feet backward. IC4 and IC9 model the vertical movement of the root during the walking. IC5 and IC6 seem to model the movement of the right hand with the corresponding movement of the feet with the left hand fixed. IC7 seems to model the movement of the right upper leg and IC8 seems to model the movement of the left upper leg. IC10 models the movement of the right hand with the corresponding movement of the legs with the left hand fixed.

When the dimensionality is reduced to 5, then ICA is applied. The extracted components seem to correspond to the combinations of the above extracted components. The extracted ICs still show the independent sub-motions, but they do not reflect the sub-motions of body parts so well as the ICs when the dimensionality is 10.

The movies produced cannot be shown here. Currently, they are still in the amc style.

4 Experimental results on missing motion synthesis and prediciton

Following the discussion in section 2.3, I will use a self-regression model to synthesize and predict the missing motion data. The experiment was performed on a sequence of walking motion data. The total number of frames is 342. The first 280 frames are used as training data, and I try to predict the following 62 frames. The training data is reduced to 3 dimensions by PCA first, and on each of the 3 independent components, a self-regression model based on GP is trained. Figure 1 and Figure 2 shows the components corresponding to PCA and ICA respectively.

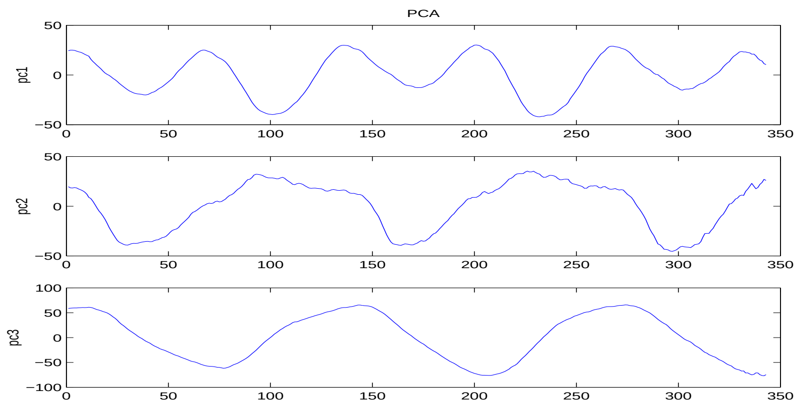

Fig. 1. The principal components of the walking data produced by PCA.

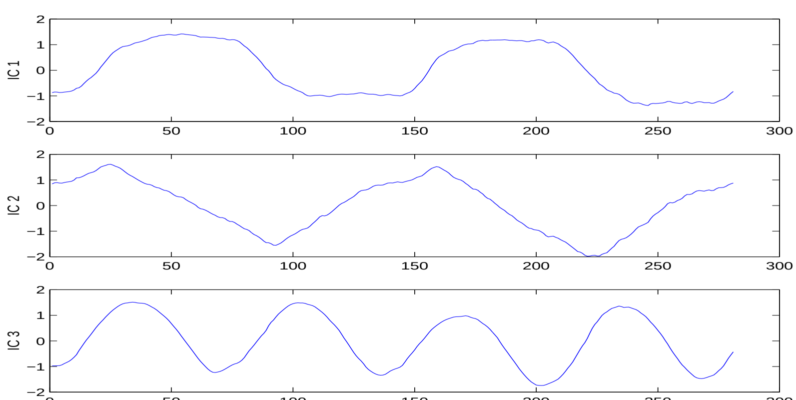

Fig. 2. The independent components of the walking data produced by ICA.

From the above two figures, we can easily find that the components generated by ICA is more periodical than the ones generated by PCA. Thus, it's reasonable to believe that the independent components can provide a big advantage to predicting missing data. In the experiment, the self-regression model based on ICA can predict about half of a walking cycle whereas the self-regression model based on PCA alone can predict nothing.

The above experiment is just a toy experiment. To synthesize and predict missing motion data, the dimensionality should be made slighly bigger, for e. g., 10. But the experimental result still shows that prediction based on ICA is better, and it also shows that the self-regression model fails to work well on missing data synthesis and prediciton. Because the input-output pairs (C_t, C_{t+1}) are used to train the self-regression model, but C_t and C_{t+1}) are often very close to each other, it is very hard for the self-regression model to capture the overall shape of the distribution of the component data. It's very likely that k-NN or parametric periodic functions will do a much better job on synthesizing the missing segments of each independent component.

5 Discussion and future work

In this report, ICA is used to extract the independent sub-motions of human motion data. It shows that it is hard to interpretate each independent component if square ICA or under-complete ICA is applied directly to the motion data; if PCA is used to do dimensionality reduction on the motion data first and the dimensionality is appropriately chosen, the extracted independent components correspond to meaningful sub-motions of body parts. It also shows that the independent components extracted by ICA have some advantages over principal components when they are used for synthesizing or predicting missing motion data.

The self-regression model based on GP fails to give good results in this paper because only C_t is used to regress C_{t+1}. If we use $C_{t-k:t}$ to regress C_{t+1}, the performance might improve. And k-NN and parametric periodic functions might also be promissing candidate tools. Besides, we can also synthesize the missing segments of each component by maximizing the likelyhood of the whole sequence given the observed data. Unlike k-NN, a powerful probabilistic model can be used here to estimate the missing segments of each component simultaneously.

In the future, annotations of motion data can be incorporated to assist motion transfer and motion synthesis based on the independent factors extracted by ICA. A probabilistic model can be used to model each component. When predicting missing data, the anotation labels provide some prior information about the missing segments.

References

1. Y. Cao, P. Faloutsos, and F. Pighin. Unsupervised learning for speech motion editing. SIGGRAPH 2003.

2. A. Shapiro, Y. Cao, and P. Faloutsos. Interactive Motion Decomposition.
3. D. Vlastic, M. Brand, H. Pfister, and J. Popovic. Facer transfer with multilinear models. SIGGRAPH 2005.
4. J. Lee, J. Chai, P. Reitsma, J. Hodgins, and N. Polland. Interactive control of avatars animated with human motion data. SIGGRAPH 2002.
5. Y. Li, T. Wang, and H.-Y. Shum. Motion texture: a two-level statistical model for character motion synthesis. SIGGRAPH 2002.
6. Arikan, D. Forsyth, and J. O'Brien. Motion synthesis from annotations. ACM SIGGRAPH 2003), Vol: 33, No: 3, pp 402–408, 2003
7. J. Wang, D. Fleet, and A. Hertzman. Gaussian process dynamical models. Advances in Neural Information Processing Systems, 2005
8. A. Hyvarinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. IEEE Transactions on Neural Networks 10(3):626-634, 1999.

This article was processed using the LaTeX macro package with LLNCS style