

## Research Article

# Projection Analysis Optimization for Human Transition Motion Estimation

Wanyi Li , Feifei Zhang, Qiang Chen, and Qian Zhang

*Department of Computer Science, Guangdong University of Education, Guangzhou, Guangdong 510303, China*

Correspondence should be addressed to Wanyi Li; [luther1212@163.com](mailto:luther1212@163.com)

Received 7 November 2018; Revised 18 March 2019; Accepted 9 April 2019; Published 2 June 2019

Academic Editor: Jintao Wang

Copyright © 2019 Wanyi Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is a difficult task to estimate the human transition motion without the specialized software. The 3-dimensional (3D) human motion animation is widely used in video game, movie, and so on. When making the animation, human transition motion is necessary. If there is a method that can generate the transition motion, the making time will cost less and the working efficiency will be improved. Thus a new method called latent space optimization based on projection analysis (LSOPA) is proposed to estimate the human transition motion. LSOPA is carried out under the assistance of Gaussian process dynamical models (GPDM); it builds the object function to optimize the data in the low dimensional (LD) space, and the optimized data in LD space will be obtained to generate the human transition motion. The LSOPA can make the GPDM learn the high dimensional (HD) data to estimate the needed transition motion. The excellent performance of LSOPA will be tested by the experiments.

## 1. Introduction

3-dimensional (3D) human motion animation is applied in many fields, such as video game and movie. It is necessary to estimate the human transition motion for making all kinds of 3D animations [1–4]. Estimating the human transition motion is crucial to making the smooth animation of the 3D human motion; it is a branch of human motion estimation. As to technologies of human motion estimation, there are some advanced methods in recent years, such as multiview image segmentation [5], sparse presentation [6], and convolutional neural network (CNN) coupled with a geometric prior [7]. These methods focus on the reconstruction of the 3D human motion from the 2-dimensional (2D) image sequence. The needed data is the high dimensional data of the 3D human motion model. The mapping will be built between the model and the 2D image for each frame. However, it is difficult to build the mapping without the overcomplete prior information, as a result of the data complexity during the optimization. Some human poses contain the ambiguity of limbs; for example, it is hard to determine which thigh is in the front from the silhouette. The right thigh or left thigh cannot be confirmed. The problem is that if we have enough prior information to distinguish the ambiguity, the

reconstruction will be achieved easily. Thus the generated model of 3D human motion is necessary; the samples of the model can be obtained to construct the prior information. Besides, the generated model can also generate the 3D human motion for making 3D character movie. Then the generated model can be built through the unsupervised learning. The unsupervised learning of the model is the necessary supplement of the advantaged methods above. In this paper, how to generate the human transition motion will be mainly discussed in the following sections

If there is a method that can estimate the valid human transition motion, the animation making time will cost less, and the work will get easier. Thus a new method called latent space optimization based on projection analysis (LSOPA) is proposed to estimate the human transition motion. LSOPA needs to combine Gaussian process dynamical models (GPDM) [8] to process the low dimensional (LD) data. GPDM is the derivation of some dimension reduction models [9–13]; it can provide the prediction of the LD data. After the dimension reduction, LD data will be optimized by LSOPA, so that the valid human transition motion can be generated to achieve the estimation. The human motion is described by the high dimensional data. If the HD data sample is searched in its own dimensional space, the invalid data will be generated;

it means the generated human motion in 3D will be abnormal [14]. GPDM is an unsupervised learning model; it can learn the high dimensional data (HD) sample and estimate the new one, but it needs to process the LD data in the LD space. In the LD space, the LD data can be searched, and the corresponding valid HD data can be generated by the mapping from LD data to HD data. Some methods [15–17] can process the LD data, but the generated LD data are all unreliable and undetermined during the optimization, as a result of the randomization of these methods. The LSOPA can do this work to process the LD data better and ensure the valid transition motion can be generated. The excellent performance of LSOPA will be tested by the experiments in the corresponding section.

The human motion is described by a 3D human motion model. The model has some markers to show the limbs of human motion; it is the HD data. When we make the 3D human motion animation and only have the samples of the two irrelevant human motions, it is necessary to use the transition motion to connect the two irrelevant human motions, so that the complete and smooth human motion in 3D is constructed. Meanwhile, the transition motion consists of many poses, and the poses are all the HD data samples of 3D human model, thus how to estimate the valid transition motion is a challenged task. However, the LSOPA can take the advantage of the GPDM to generate the valid transition motion and avoid generating the invalid pose of the transition motion. The proposed method will be discussed in the following sections.

## 2. Dimension Reduction

When we have the sequence of HD data samples  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_i, \dots, \mathbf{y}_N]^T \in \mathbb{R}^{N \times D}$ ,  $\mathbf{y}_i \in \mathbb{R}^D$ , the corresponding LD data  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N]^T \in \mathbb{R}^{N \times q}$ ,  $\mathbf{x}_i \in \mathbb{R}^q$  can be obtained as the following equations from GPDM [8]:

$$p(\mathbf{Y} | \mathbf{X}, \bar{\boldsymbol{\beta}}, \mathbf{W}) = \frac{|\mathbf{W}|^N}{\sqrt{(2\pi)^{ND} |\mathbf{K}_Y|^D}} \exp\left(-\frac{1}{2}\right. \quad (1)$$

$$\left. \cdot \text{tr}(\mathbf{K}_Y^{-1} \mathbf{Y} \mathbf{W}^2 \mathbf{Y}^T)\right),$$

$$p(\mathbf{X} | \bar{\boldsymbol{\alpha}}) = \frac{p(\mathbf{x}_1)}{\sqrt{(2\pi)^{(N-1)q} |\mathbf{K}_X|^q}} \exp\left(-\frac{1}{2}\right. \quad (2)$$

$$\left. \cdot \text{tr}(\mathbf{K}_X^{-1} \mathbf{X}_{2:N} \mathbf{X}_{2:N}^T)\right),$$

$$(\mathbf{K}_Y)_{i,j} = k_Y(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\beta_1}{2} \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) + \beta_2^{-1} \delta_{\mathbf{x}_i, \mathbf{x}_j}, \quad (3)$$

$$(\mathbf{K}_X)_{i,j} = k_X(\mathbf{x}_i, \mathbf{x}_j) = \alpha_1 \exp\left(-\frac{\alpha_2}{2} \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) + \alpha_3 \mathbf{x}_i^T \mathbf{x}_j + \alpha_4^{-1} \delta_{\mathbf{x}_i, \mathbf{x}_j}, \quad (4)$$

$$p(\mathbf{W}) = \prod_{m=1}^D \frac{2}{\kappa \sqrt{2\pi}} \exp\left(-\frac{w_m^2}{2\kappa^2}\right), \quad (5)$$

$$p(\alpha) \propto \prod_i \alpha_i^{-1}, \quad (6)$$

$$p(\beta) \propto \prod_i \beta_i^{-1}, \quad (7)$$

$$\begin{aligned} \mathbf{X}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}, \mathbf{W} &= \arg \min_{\mathbf{X}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}, \mathbf{W}} (-\ln p(\mathbf{X}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}, \mathbf{W}, \mathbf{Y})) \\ &= \arg \min_{\mathbf{X}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}, \mathbf{W}} (-\ln (p(\mathbf{Y} | \mathbf{X}, \bar{\boldsymbol{\beta}}, \mathbf{W}) \\ &\quad \cdot p(\mathbf{X} | \bar{\boldsymbol{\alpha}}) p(\bar{\boldsymbol{\alpha}}) p(\bar{\boldsymbol{\beta}}) p(\mathbf{W}))), \end{aligned} \quad (8)$$

Equations (1)-(7) are used to computing (8), then the  $\mathbf{X}$  and the other parameters can be got. In (1) and (2), we have the sequences  $\mathbf{X} \in \mathbb{R}^{N \times q}$ ,  $\mathbf{Y} \in \mathbb{R}^{N \times D}$ , respectively,  $\mathbf{Y}$  denotes the HD data samples of 3D human motion, and  $\mathbf{X}$  denotes the LD data of  $\mathbf{Y}$  in the LD space after the dimension reduction. In (3) and (4),  $\mathbf{K}_Y \in \mathbb{R}^{N \times N}$  and  $\mathbf{K}_X \in \mathbb{R}^{(N-1) \times (N-1)}$  are kernel matrices, respectively,  $\bar{\boldsymbol{\beta}} = [\beta_1, \beta_2]$  and  $\bar{\boldsymbol{\alpha}} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$  are their corresponding kernel parameters, and they satisfy the relation of (6) and (7), respectively. Equations (3) and (4) are showing the method of computing the two kernel matrices  $\mathbf{K}_Y$  and  $\mathbf{K}_X$ .  $\mathbf{W} = \text{diag}(w_1, \dots, w_m, \dots, w_D) \in \mathbb{R}^{D \times D}$  is a scale diagonal matrix with the preset parameter  $\kappa$ ,  $\mathbf{x}_1$  conforms Gaussian distribution of  $q$  dimensions, and  $\mathbf{X}_{2:N}$  is the sequence  $\mathbf{X}_{2:N} = [\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N]^T \in \mathbb{R}^{(N-1) \times q}$ . The mapping from LD data to HD data can be built as follows:

$$\begin{aligned} \mathbf{y}^* &= \mathbf{f}(\mathbf{x}^*) = \boldsymbol{\mu}_Y(\mathbf{x}^*) \\ &= \mathbf{Y}^T \mathbf{K}_Y^{-1} [k_Y(\mathbf{x}_1, \mathbf{x}^*), k_Y(\mathbf{x}_2, \mathbf{x}^*), \dots, k_Y(\mathbf{x}_N, \mathbf{x}^*)]^T \quad (9) \\ &= \mathbf{Y}^T \mathbf{K}_Y^{-1} \mathbf{k}_Y(\mathbf{x}^*). \end{aligned}$$

The GPDM has a dynamic process. It can predict the LD data in the latent space (also called LD space) and generate the needed HD data of human motion, so that it has better performance than other dimension reduction models. Thus GPDM can be selected to learn the samples of the two different human motions; then the LD space can be built to find the needed LD data of the transition motion, so that corresponding poses can be generated through the mapping of (9).

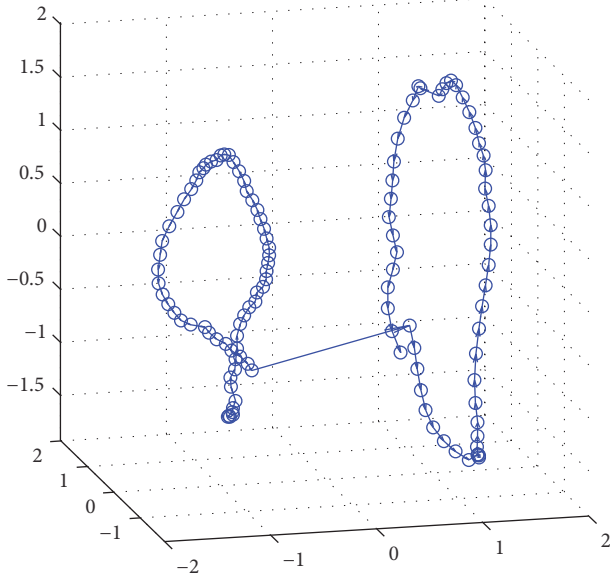


FIGURE 1: The LD data of the two irrelevant human motions.

### 3. Latent Space Optimization Based on Projection Analysis

**3.1. The LD Data of the Transition Motion in the LD Space.** After the dimension reduction, the HD data samples of the two irrelevant human motions can be seen in the LD space. The LD data of two difference sequences can denote the corresponding poses of the two irrelevant human motions. It can be found that there is an obvious distance between the two LD data as Figure 1 shows. From the two sets of LD data, we can know the needed transition motion can be generated through the LD space, but it needs to construct an appropriated curve to connect the two sets of LD data in the LD space. Thus, the optimized work will be started.

**3.2. Optimization in the LD Space.** Assume that  $\mathbf{X}_1 = [\mathbf{x}_1^1, \mathbf{x}_2^1, \dots, \mathbf{x}_{N_1}^1]^T \in \mathbb{R}^{N_1 \times q}$ ,  $\mathbf{X}_2 = [\mathbf{x}_1^2, \mathbf{x}_2^2, \dots, \mathbf{x}_{N_2}^2]^T \in \mathbb{R}^{N_2 \times q}$  are, respectively, the LD data of the two irrelevant human motions after dimension reduction, the first motion has  $N_1$  frames, and the second motion has  $N_2$  frames. The positions of the corresponding LD data can be seen in Figure 2. In Figure 2, the derivations of  $L$  can be got as follows:

$L$

$$L = \left\| \frac{(\mathbf{x}'_i - \mathbf{x}_{N_1}^1) - (\mathbf{x}'_i - \mathbf{x}_{N_1}^1)^T (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1) (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)}{\|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|^2} \right\|, \quad (10)$$

$$\mathbf{A} = \frac{(\mathbf{x}'_i - \mathbf{x}_{N_1}^1)^T (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)}{\|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|}. \quad (11)$$

From Figure 2, we can see  $L$  denotes the distance from the LD data sample  $\mathbf{x}'_i \in \mathbb{R}^3$  to its projection  $\mathbf{A}$  in (10), and we can also see the vector  $(\mathbf{B} - \mathbf{x}_{N_1}^1)$  is the projection of the vector  $(\mathbf{x}'_i - \mathbf{x}_{N_1}^1)$

in the plane  $\omega$ . Assume the set of bases  $\mathbf{B}^* = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_q] \in \mathbb{R}^{q \times q}$  in the plane  $\omega$ ,  $\mathbf{b}_i \in \mathbb{R}^3$ ,  $i = 1, 2, \dots, q$  is the vector from the plane  $\omega$ ; we can get

$$\mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1) = (\mathbf{B} - \mathbf{x}_{N_1}^1), \quad (12)$$

$$(\mathbf{x}'_i - \mathbf{x}_{N_1}^1) - \mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1) = \mathbf{Proj}_{\omega^\perp} (\mathbf{x}'_i - \mathbf{x}_{N_1}^1). \quad (13)$$

$\mathbf{Proj}_H$  denotes the projective operation, which means obtaining the projective vector or plane of  $H$ .  $\mathbf{Proj}_{cH}$  denotes the complement of the projective operation  $\mathbf{Proj}_H$ , which means obtaining the vector perpendicular to projective vector from the operation  $\mathbf{Proj}_H$ . In (12) and (13), the complement of the projection vector is perpendicular to the plane  $\omega$ ; thus the equation can be got from (13) as follows:

$$\begin{aligned} \mathbf{B}^{*T} ((\mathbf{x}'_i - \mathbf{x}_{N_1}^1) - \mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1)) \\ = \mathbf{B}^{*T} \mathbf{Proj}_{\omega^\perp} (\mathbf{x}'_i - \mathbf{x}_{N_1}^1) = \mathbf{0}. \end{aligned} \quad (14)$$

In (14),  $\mathbf{B}^* \in \mathbb{R}^{q \times q}$ ; then  $\mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1)$  is a vector of the plane  $\omega$ , and (15) can deduced from this:

$$\mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1) = \mathbf{B}^* \tilde{\mathbf{y}} = y_1 \mathbf{b}_1 + y_2 \mathbf{b}_2 + \dots + y_q \mathbf{b}_q. \quad (15)$$

In (15),  $\mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1) \in \omega$  can be denoted by the bases  $\mathbf{b}_i \in \mathbb{R}^3$ ,  $i = 1, 2, \dots, q$  and coefficients  $\tilde{\mathbf{y}} = [y_1, y_2, \dots, y_q]^T \in \mathbb{R}^q$ ; then  $\mathbf{B}^* = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_q] \in \mathbb{R}^{q \times q}$ . Substituting (15) into (14), the following equation can be deduced:

$$\mathbf{B}^{*T} ((\mathbf{x}'_i - \mathbf{x}_{N_1}^1) - \mathbf{B}^* \tilde{\mathbf{y}}) = \mathbf{0}. \quad (16)$$

Simplify (16), compute the least squares estimator of  $\tilde{\mathbf{y}}$ , and get the equation as below:

$$\tilde{\mathbf{y}} = (\mathbf{B}^{*T} \mathbf{B}^*)^{-1} \mathbf{B}^{*T} (\mathbf{x}'_i - \mathbf{x}_{N_1}^1). \quad (17)$$

Substitute (12) and (17) into (15); the following equation can be obtained:

$$\begin{aligned} \mathbf{Proj}_\omega (\mathbf{x}'_i - \mathbf{x}_{N_1}^1) &= (\mathbf{B} - \mathbf{x}_{N_1}^1) \\ &= \mathbf{B}^* (\mathbf{B}^{*T} \mathbf{B}^*)^{-1} \mathbf{B}^{*T} (\mathbf{x}'_i - \mathbf{x}_{N_1}^1). \end{aligned} \quad (18)$$

From (15) to (18), we have  $\tilde{\mathbf{y}} \in \mathbb{R}^q$ , and the bases  $\mathbf{b}_i \in \mathbb{R}^3$ ,  $i = 1, 2, \dots, q$  are linearly independent.

Our task is to find the optimal LD data sequence of the transition motion between  $\mathbf{x}_{N_1}^1$  and  $\mathbf{x}_1^2$ ; thus we can build the constraint functions to describe position of the needed LD data sequence  $\mathbf{X}' = [\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_{N'}]^T \in \mathbb{R}^{N' \times q}$  as follows:

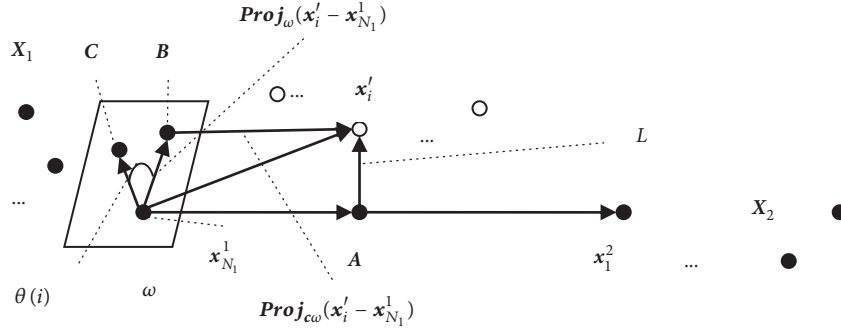


FIGURE 2: The analysis of the optimized models.

$$\frac{(\mathbf{x}'_i - \mathbf{x}_{N_1}^1)^T (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)}{\|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|} = \frac{i \|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|}{(N' + 1)}, \quad (19)$$

$$i = 1, 2, \dots, N'.$$

Equation (19) denotes confirming the projection distance in the vector  $(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)$  from each  $\mathbf{x}'_i \in \mathbb{R}^q$ ,  $i = 1, 2, \dots, N'$ ; then, we have

$$\left\| \frac{(\mathbf{x}'_i - \mathbf{x}_{N_1}^1) - (\mathbf{x}'_i - \mathbf{x}_{N_1}^1)^T (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1) (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)}{\|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|^2} \right\| \quad (20)$$

$$= \frac{l(i) \|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|}{(N' + 1)} = L, \quad i = 1, 2, \dots, N'.$$

Equation (20) denotes confirming the distance between  $\mathbf{x}'_i$  and  $\text{Proj}_{(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)} \mathbf{x}'_i$ ; meanwhile, we can get

$$\arccos \frac{\mathbf{B}'^T \mathbf{C}}{(\|\mathbf{B}'\| \|\mathbf{C}\|)} < \theta(i), \quad i = 1, 2, \dots, N'. \quad (21)$$

In (21),  $\mathbf{B}' = (\mathbf{B} - \mathbf{x}_{N_1}^1)$  is a vector of  $\text{Proj}_\omega(\mathbf{x}'_i - \mathbf{x}_{N_1}^1)$  which is perpendicular to the plane  $\omega$ .  $\mathbf{C}$  is a preset vector in the plane  $\omega$ ; it can be seen as the reference of computing the angle of (21). In (19)-(21), the LD data of the transition motion can be described in the LD space and  $l(i)$  is the preset values of adjusting the value of  $L$ .  $\theta(i)$  is preset angle which is describing the deviation of the LD data position. Then the projection and corresponding vector can be used to build these constraint functions, so that an object function with constraints can be obtained as follows:

$$\underset{\mathbf{x}'_i}{\text{argmin}} \quad \|\mathbf{y}'_i - \mathbf{y}'_{i-1}\| = \underset{\mathbf{x}'_i}{\text{argmin}} \|\mathbf{f}(\mathbf{x}'_i) - \mathbf{f}(\mathbf{x}'_{i-1})\|, \quad \mathbf{x}'_0 = \mathbf{x}_{N_1}^1, \quad i = 1, 2, \dots, N',$$

$$\text{s.t.} \quad \begin{cases} \frac{(\mathbf{x}'_i - \mathbf{x}_{N_1}^1)^T (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)}{\|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|} = \frac{i \|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|}{(N' + 1)} \\ \left\| \frac{(\mathbf{x}'_i - \mathbf{x}_{N_1}^1) - (\mathbf{x}'_i - \mathbf{x}_{N_1}^1)^T (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1) (\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)}{\|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|^2} \right\| = \frac{l(i) \|(\mathbf{x}_1^2 - \mathbf{x}_{N_1}^1)\|}{(N' + 1)}, \quad i = 1, 2, \dots, N'. \\ \arccos \frac{\mathbf{B}'^T \mathbf{C}}{(\|\mathbf{B}'\| \|\mathbf{C}\|)} < \theta(i) \end{cases} \quad (22)$$

The optimized object function of (22) ensures that the frames of the transition motion can be varying regularly and the valid transition motion will be seen in the vision. The solution of (22) can be obtained by the method in [18].

3.3. *The Procedure of the LSOPA.* When the two irrelevant human motion samples are obtained, the two samples are

denoted by  $\mathbf{Y}_1 = [\mathbf{y}_1^1, \mathbf{y}_2^1, \dots, \mathbf{y}_{N_1}^1]^T \in \mathbb{R}^{N_1 \times q}$ ,  $\mathbf{Y}_2 = [\mathbf{y}_1^2, \mathbf{y}_2^2, \dots, \mathbf{y}_{N_2}^2]^T \in \mathbb{R}^{N_2 \times q}$ . Then, we have the following.

LSOPA

- (1) Let  $\mathbf{Y} = [\mathbf{Y}_1^T, \mathbf{Y}_2^T]^T$ ; use the method called principal component analysis (PCA) to initiate  $\mathbf{Y}$  and get the

```

Input:  $Y_1, Y_2$ ; Output:  $Y_{\text{new}}, X_u, \bar{\alpha}_u, \bar{\beta}_u, W_u$ 
Preset  $\theta(i), l(i), i = 1, 2, \dots, N'$ ,  $C, \bar{\alpha}, \bar{\beta}, W$ ;
 $Y = [Y_1^T, Y_2^T]^T$ ;
Use PCA to process  $Y$ , and get  $X = [X_1^T, X_2^T]^T$ ;
 $X, \bar{\alpha}, \bar{\beta}, W = \arg \min_{X, \bar{\alpha}, \bar{\beta}, W} (-\ln p(X, \bar{\alpha}, \bar{\beta}, W, Y))$ ;
For  $i=1$  to  $N'$ 
   $\arg \min_{x'_i} \|y'_i - y'_{i-1}\| = \arg \min_{x'_i} \|f(x'_i) - f(x'_{i-1})\|, \quad x'_0 = x'_{N_1},$ 
  s.t.  $\begin{cases} \frac{(x'_i - x'_{N_1})^T (x'_i - x'_{N_1})}{\|(x'_i - x'_{N_1})\|} = \frac{i \|(x'_1 - x'_{N_1})\|}{(N' + 1)} \\ \left\| \frac{(x'_i - x'_{N_1}) - (x'_i - x'_{N_1})^T (x'_1 - x'_{N_1}) (x'_1 - x'_{N_1})}{\|(x'_i - x'_{N_1})\|^2} \right\| = \frac{l(i) \|(x'_1 - x'_{N_1})\|}{(N' + 1)} \\ \arccos \frac{B^T C}{(\|B\| \|C\|)} < \theta(i) \end{cases}$ 
   $y'_i = f(x'_i)$ ;
End For
 $X' = [x'_1, x'_2, \dots, x'_{N'}]^T$ ;
 $Y' = [y'_1, y'_2, \dots, y'_{N'}]^T$ ;
 $X_{\text{new}} = [X_1^T, X'^T, X_2^T]^T$ ;
 $Y_{\text{new}} = [Y_1^T, Y'^T, Y_2^T]^T$ ;
 $X_u, \bar{\alpha}_u, \bar{\beta}_u, W_u = \arg \min_{X, \bar{\alpha}, \bar{\beta}, W} (-\ln p(X_{\text{new}}, \bar{\alpha}, \bar{\beta}, W, Y_{\text{new}}))$ ;
End the procedure.

```

ALGORITHM 1: The brief procedure of LSOPA.

initiated LD data  $X = [X_1^T, X_2^T]^T$ . Preset the values of  $\theta(i), l(i)$ , and  $C$ .

- (2) Preset  $\bar{\alpha}, \bar{\beta}, W$ . Then, compute  $X, \bar{\alpha}, \bar{\beta}, W = \arg \min_{X, \bar{\alpha}, \bar{\beta}, W} (-\ln p(X, \bar{\alpha}, \bar{\beta}, W, Y))$ , update  $X$ , and get  $\bar{\alpha}, \bar{\beta}, W$ .
- (3) Compute  $\arg \min_{x'_i} \|y_i - y_{i-1}\| = \arg \min_{x'_i} \|f(x'_i) - f(x'_{i-1})\|$ ,  $x'_0 = x'_{N_1}$ ,  $i = 1, 2, \dots, N'$  and the corresponding constraint functions in (22), and get the LD data  $X' = [x'_1, x'_2, \dots, x'_{N'}]^T \in \mathbb{R}^{N' \times q}$  of the transition motion.
- (4) Compute the transition motion  $y'_i = f(x'_i)$ ,  $i = 1, 2, \dots, N'$ , let  $Y' = [y'_1, y'_2, \dots, y'_{N'}]^T$ , and output the 3D transition motion model. Finally, let  $Y_{\text{new}} = [Y_1^T, Y'^T, Y_2^T]^T$ ,  $X_{\text{new}} = [X_1^T, X'^T, X_2^T]^T$ , compute the  $X_u, \bar{\alpha}_u, \bar{\beta}_u, W_u = \arg \min_{X, \bar{\alpha}, \bar{\beta}, W} (-\ln p(X_{\text{new}}, \bar{\alpha}, \bar{\beta}, W, Y_{\text{new}}))$ , and update  $X_u, \bar{\alpha}_u, \bar{\beta}_u, W_u$ . End the whole procedure.

The method LSOPA can be written as pseudocode in Algorithm 1 for the transition motion estimation.

## 4. Experiment and Evaluation

We test the proposed method LSOPA in the performance of the vision and error test. We choose the line initialization [8] (LI) and random initialization [17] (RI) as the compared methods. LI means that the LD data of the transition motion is obtained linearly between the LD data of the two different motions. RI means that the LD data of the transition motion is obtained randomly between the LD data of two different motions. The LD data samples (green ones) in the LD space can be seen in Figures 3(a)–3(c), which denotes the transition motion. The tested transition motion is a total of 15 frames and the tested database is the Carnegie Mellon University (CMU) motion capture database which is shortened for CMU database.

The three methods can estimate the corresponding transition motions, but the transition motions are obviously different. Let us see the effect in the vision from Figures 4(a)–4(c). In the estimation of the transition motion between walking and playing golf, we can find that the transition motion generated by the LSOPA is smoothly and naturally in vision. The transition motions generated by the other two methods are invalid and messy; they are abnormal human transition motions obviously. In a word, the transition motions of LI and RI are not like playing the golf. It is demonstrated that the performance of LSOPA is the best in the visual effect among the three methods.

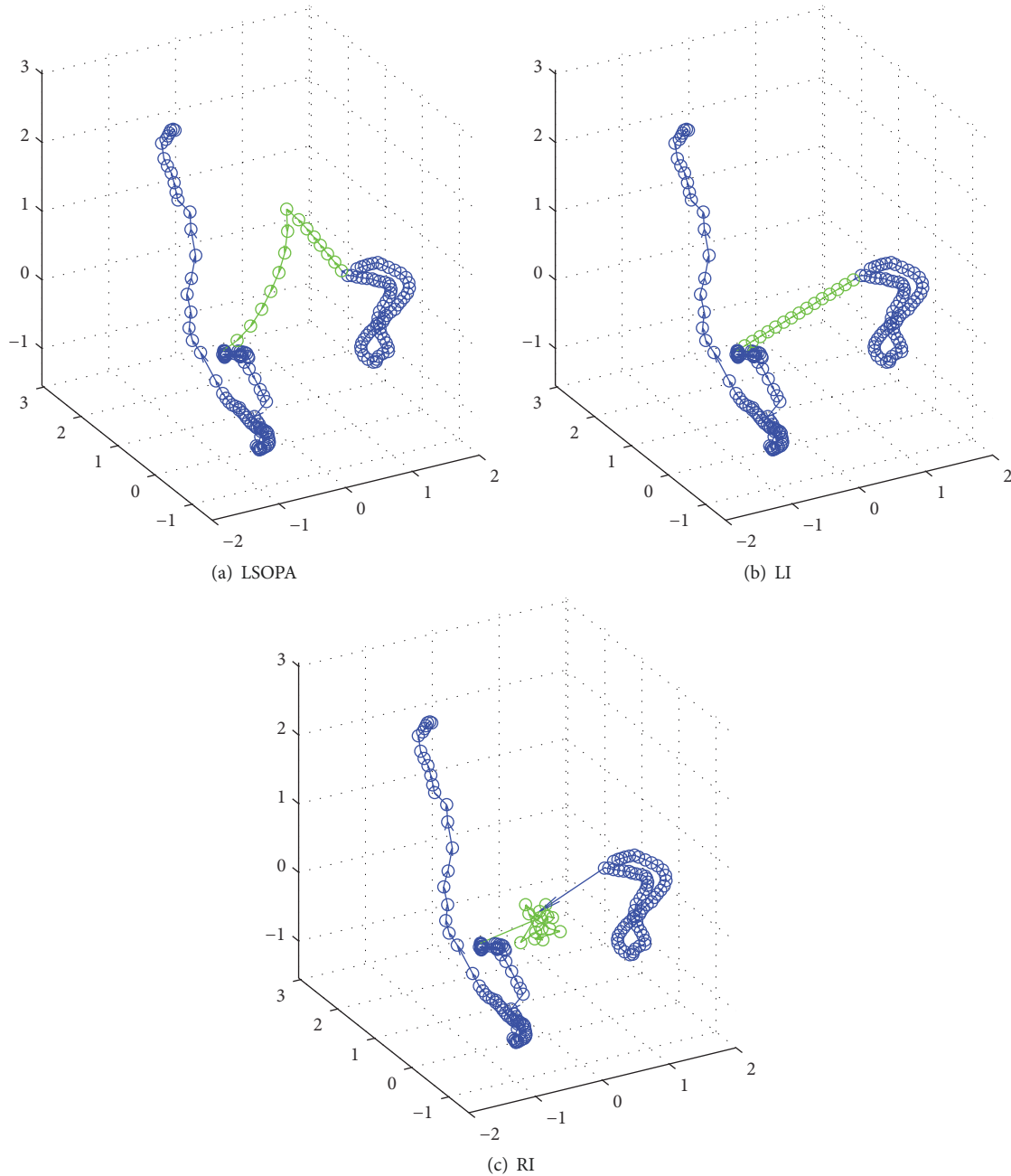


FIGURE 3: The LD data samples of the LSOPA, LI, and RI.

Then, the error and mean error of estimating the transition motion are tested. How to compute the error can be seen in [19], but the scale of markers in the 3D model must be considered. The error can be computed as follows:

$$Er = \frac{1}{Mar} \sum_{ma=1}^{Mar} \|\mathbf{Bx}_{ma} - \mathbf{Bx}'_{ma}\|. \quad (23)$$

In (23),  $\mathbf{Bx}_{ma}$ ,  $\mathbf{Bx}'_{ma} \in \mathbb{R}^3$ ,  $\mathbf{Bx}_{ma}$ , and  $\mathbf{Bx}'_{ma}$  are, respectively, the true position and the estimated position of the marker in

the 3D model; both are the joint marker in the 3D model.  $Mar$  is the number of the markers in the 3D model.

The transition motion includes four transition motions; they are walking and playing golf, walking and dancing, walking and swimming, and walking and playing football, respectively. The contrast of the true data and the estimated data can be seen in Figures 5(a)–5(c). The error of each frame in the transition motion generated by LSOPA is the lowest among the three methods, so is the mean error. Generally speaking, the error of some frames will be close, but the



FIGURE 4: The visual evaluation of the methods.

error of LSOPA will be kept low and stable on the whole. The tested result can also reveal that the LSOPA has the best performance in the error test.

## 5. Conclusion

The LSOPA is proposed to solve the problem of estimating the human transition motion. The LSOPA is tested in the experiments of the vision and error test, and we can find that the performance of the LSOPA is the best among the three methods. However, the LSOPA still cannot process some complex human transition motion well, and its initialization of optimization cannot be random; thus the LSOPA need to be improved in the future research. Then, the improved method can carry out the unsupervised learning of the two complex irrelevant human motions and estimate their smooth and valid transition motion. The 3D human motion

model will also be replaced by other exquisite models [20–22] in the future research.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work is supported by the University Young Creative Talent Project of Guangdong Province (no. 2016KQNCX111); Science and Technology Plan Project of Guangzhou (no.

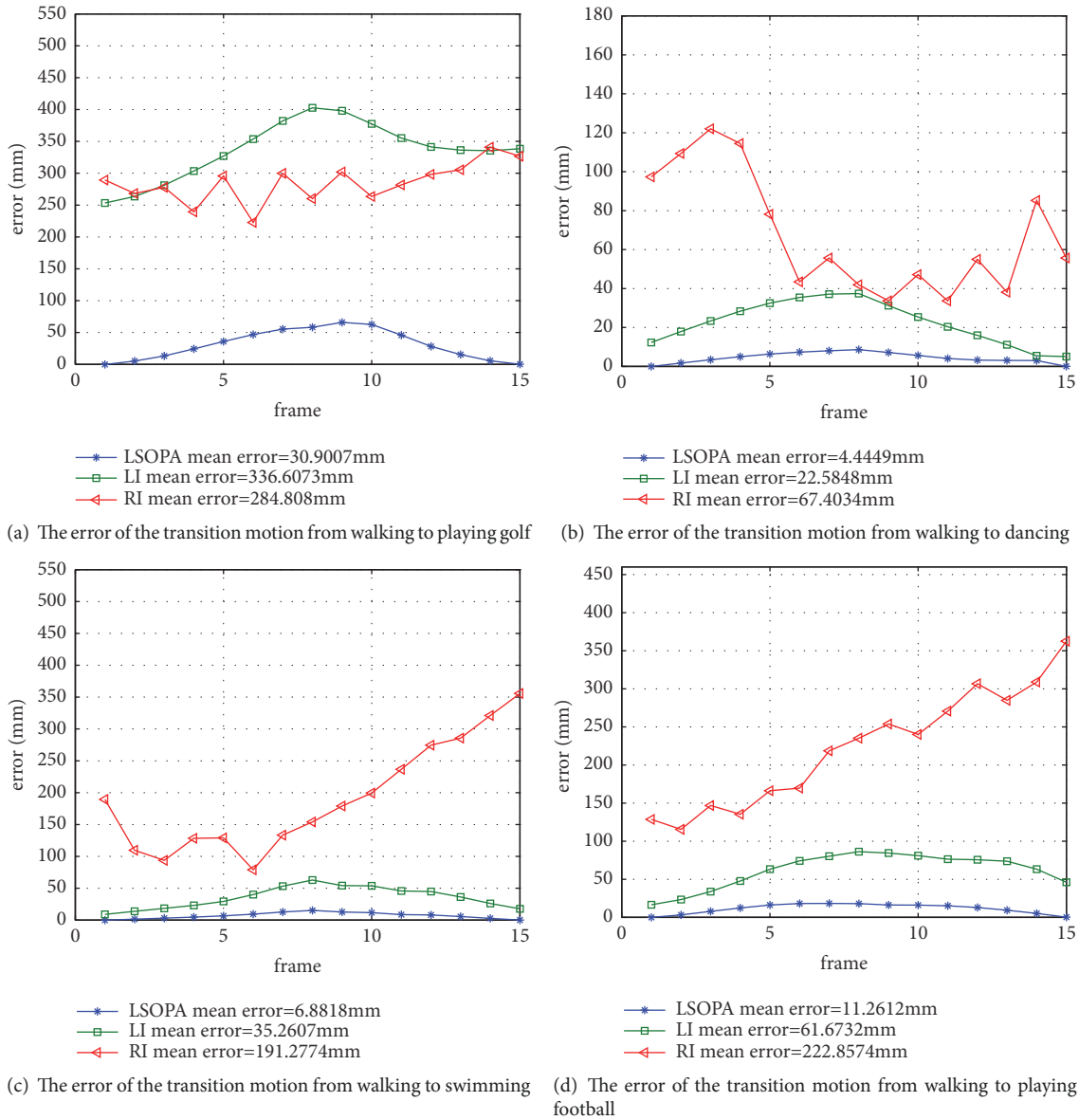


FIGURE 5: The error of several transition motions.

201804010280); National Natural Science Foundation of China (no. 61772140); Science and Technology Planning Project of Guangdong Province (no. 2017A010101021); Appropriative Researching Fund for Professors and Doctors, Guangdong University of Education (no. 2015ARF17, no. 2015ARF25); Key Disciplines Of Network Engineering of Guangdong University of Education (no. ZD2017004); and Computer Practice Teaching Demonstration Center of Guangdong University of Education (no. 2018sfzx01).

## References

- [1] Z. Liu, L. Zhou, H. Leung et al., "High-quality compatible triangulations and their application in interactive animation," *Computers and Graphics*, vol. 76, pp. 60–72, 2018.
- [2] A. Sujar, J. J. Casafra, A. Serrurier et al., "Real-time animation of human characters' anatomy," *Computers and Graphics*, vol. 74, pp. 268–277, 2018.
- [3] B. Walther-Franks and R. Malaka, "An interaction approach to computer animation," *Entertainment Computing*, vol. 5, no. 4, pp. 271–283, 2014.
- [4] D. Archambault and H. C. Purchase, "Can animation support the visualisation of dynamic graphs?" *Information Sciences*, vol. 330, pp. 495–509, 2016.
- [5] L. Yebin, G. Juergen, S. Carsten et al., "Markerless motion capture of multiple characters using multiview image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2720–2735, 2013.
- [6] X. Zhou, M. Zhu, S. Leonardos, and K. Daniilidis, "Sparse representation for 3D shape estimation: a convex relaxation



- approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1648–1661, 2017.
- [7] X. Zhou, M. Zhu, G. Pavlakos, S. Leonardos, K. G. Derpanis, and K. Daniilidis, “MonoCap: monocular human motion capture using a CNN coupled with a geometric prior,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 4, pp. 901–914, 2019.
- [8] J. M. Wang, D. J. Fleet, and A. Hertzmann, “Gaussian process dynamical models for human motion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 283–298, 2008.
- [9] S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [10] N. Lawrence, “Probabilistic non-linear principal component analysis with Gaussian process latent variable models,” *Journal of Machine Learning Research (JMLR)*, vol. 6, pp. 1783–1816, 2005.
- [11] C. H. Ek, *Shared Gaussian Process Latent Variables Models*, Oxford Brookes University, Oxford, UK, 2009.
- [12] J. Xinwei, G. Junbin, W. Tianjiang et al., “Supervised latent linear Gaussian process latent variable model for dimensionality reduction,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 6, pp. 1620–1632, 2012.
- [13] J. B. Tenenbaum, V. de Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [14] W.-Y. Li and J.-F. Sun, “Human motion estimation based on gaussian incremental dimension reduction and manifold boltzmann optimization,” *Acta Electronica Sinica*, vol. 45, no. 12, pp. 3060–3069, 2017.
- [15] W. Li, J. Sun, X. Zhang, and Y. Wu, “Spatial constraints-based maximum likelihood estimation for human motions,” in *Proceedings of the 2013 IEEE International Conference on Signal Processing, Communications and Computing, ICSPCC 2013*, pp. 1–6, IEEE, KunMing, China, 2013.
- [16] W. Li, J. Sun, and Z. Song, “Manifold latent probabilistic optimization for human motion fitting based on orthogonal subspace searching,” *Journal of Information and Computational Science*, vol. 11, no. 15, pp. 5357–5365, 2014.
- [17] W. Li, J. Sun, and X. Zhang, “Learning for human transition motions estimation based on feature similarity optimization,” *Journal of Computational Information Systems*, vol. 10, no. 5, pp. 2127–2136, 2014.
- [18] R. H. Byrd, M. E. Hribar, and J. Nocedal, “An interior point algorithm for large-scale nonlinear programming,” *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1999.
- [19] L. Sigal, A. O. Balan, and M. J. Black, “HumanEva: synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion,” *International Journal of Computer Vision*, vol. 87, no. 1-2, pp. 4–27, 2010.
- [20] M. Sandau, H. Koblauch, T. B. Moeslund, H. Aanæs, T. Alkjær, and E. B. Simonsen, “Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane,” *Medical Engineering & Physics*, vol. 36, no. 9, pp. 1168–1175, 2014.
- [21] E. Yeguas-Bolivar, R. Muñoz-Salinas, R. Medina-Carnicer, and A. Carmona-Poyato, “Comparing evolutionary algorithms and particle filters for markerless human motion capture,” *Applied Soft Computing*, vol. 17, pp. 153–166, 2014.
- [22] J. Gall, B. Rosenhahn, T. Brox et al., “Optimization and filtering for human motion capture - a multi-layer framework,” *International Journal of Computer Vision*, vol. 87, no. 1-2, pp. 75–92, 2010.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

