Three-dimensional nonrigid reconstruction based on probability model

Yaming Wang, Deming Shen, Wenqing Huang and Yonghua Han

Abstract
Most nonrigid motions use shape-based methods to solve the problem; however, the use of discrete cosine transform trajectory-based methods to solve the nonrigid motion problem is also very prominent. The signal undergoes discrete transformation due to the transform characteristics of the discrete cosine transform. The correlation of the data is well extracted such that a better compression of data is achieved. However, it is important to select the number and sequence of discrete cosine transform trajectory basis appropriately. The error of reconstruction and operational costs will increase for a high value of \( K \) (number of trajectory basis). On the other hand, a lower value of \( K \) would lead to the exclusion of information components. This will lead to poor accuracy as the structure of the object cannot be fully represented. When the number of trajectory basis is determined, the combination form has a considerable influence on the reconstruction algorithm. This article selects an appropriate number and combination of trajectory basis by analyzing the spectrum of re-projection errors and realizes the automatic selection of trajectory basis. Then, combining with the probability framework of normal distribution of a low-order model matrix, the energy information of the high-frequency part is retained, which not only helps maintain accuracy but also improves reconstruction efficiency. The proposed method can be used to reconstruct the three-dimensional structure of sparse data under more precise prior conditions and lower computational costs.

Keywords
Probabilistic trajectory space, nonrigid motion, automatic selection, trajectory basis, matrix normal distribution, low-order model

Date received: 19 July 2019; accepted: 18 December 2019

Introduction
Three-dimensional (3-D) reconstruction encompasses many fields such as image processing, stereoscopic vision, and biological engineering, and has attracted considerable research interest in computer vision. The 3-D motion reconstruction of a nonrigid body is an important technique for virtual representation of the objective world. Generally, 3-D motion reconstruction involves recovering camera rotation matrix \( R \) and 3-D structure \( S \) of a nonrigid body from a given set of 2-D dynamic image sequences.

There are currently four mainstream reconstruction schemes: shape-based 3-D reconstruction,\(^1\) force-based 3-D reconstruction,\(^2\) shape-trajectory-based 3-D reconstruction,\(^3\) and trajectory-based 3-D reconstruction.\(^4\) Shape-based 3-D reconstruction has the advantages of simplicity and convenience, but in the reconstruction process, the shape basis reconstructed for all the different sequences

College of Information, Zhejiang Sci-Tech University, Hangzhou, Zhejiang, China

Corresponding author:
Wenqing Huang, College of Information, Zhejiang Sci-Tech University, Xiasha Campus, Hangzhou, Zhejiang 310000, China.
Email: patternrecog@163.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
will lead to the introduction of a large number of unknowns, which makes the algorithm complex and limits its scope. The advantages of force-based 3-D reconstruction are that it is based on the deformed low-rank force space to formulate the problem, which can better explain the acquired a priori information and more accurately represent the behavior of the actual object but in the process of reconstruction, in addition to determining the force and reality. In addition to rotating the data, it is also necessary to estimate the elastic model of the object, which increases the uncertainty of the reconstruction result. The advantage of the 3-D reconstruction based on the shape trajectory is that it combines the advantages of both the trajectory basis and shape basis; however, the disadvantage is that the a priori unknown is added. Although the reconstruction-based trajectory-based method solves the limitations of the above three methods, the number and type of predefined trajectory basis in this method are difficult to select. The number of trajectory basis and the choice of trajectory basis directly affect the reconstruction accuracy. In particular, the main difficulty in solving the nonrigid motion (NRSFM) problem is that many different 3-D graphics can produce similar observation images, and uniquely considering the re-projection constraints is not sufficient to obtain a single solution of the shape. Therefore, there is a need for more prior knowledge of the deformation of the structure and the motion of the camera. The method based on automatic selection basis-probability model proposed in this article can effectively solve the problems caused by the number of trajectory basis, combination of trajectory basis, and complex prior knowledge.

**Related work**

Most existing methods adopt the matrix decomposition algorithm to decompose a rigid reconstruction and use prior information in the form of low-rank shape basis. Similarly, a low-rank model is proposed to constrain the motion of every point on the object through a predefined trajectory basis. The disadvantage of these methods is that they need to be decomposed into a destination matrix that is proportional to the input points, and they can only be applied to relatively low-resolution shapes.

In the NRSFM, 2-D point tracks obtained only from camera motion and reconstruction of time-varying 3-D shapes are an unconstrained problem. This is because the observation results of different 3-D objects and 2-D images obtained using the camera are similar. A few algorithms have been proposed for solving the rigid constraints in a nonrigid body in the process of solving the 3-D motion reconstruction of a nonrigid body by factorization. Costeira and Kanade constructed an orthogonal projection model and then applied the factorization method to reconstruct the structure and motion of independent moving objects. However, its application scope was limited; thus, it could not satisfy the changeable linear combination. Bascle and Blake proposed the decomposition of the basic shape of a reconstruction target into a group of linear combinations of basic shapes, and the reconstruction problem was simplified into a problem of solving for the basic shape coefficient, which was also the prototype of solving the 3-D motion reconstruction problem of a nonrigid body with the model based on shape basis. Torresani et al. adopted low-rank constraints to track a nonrigid body, that is, a time constraint method combining the simulation of the shape basis coefficient as a linear dynamic system and the establishment of a nonrigid body deformation distribution model with layer factors. Rabaud and Belongie removed the linear basis representation and proposed a method for learning the shape structure from a video. The advantage of this method is that it combines time specifications to prevent the camera and structure from having excessive and blurred changes between frames. Agudo and Moreno-Noguer introduced the force model into 3-D nonrigid reconstruction, which has the advantage of formulating problems based on the low-rank force space of deformation, and better physical interpretation of the obtained prior information. However, in the process of reconstruction, if there is a lack of force and real data, the reconstruction will become difficult. Recently, the research direction of NRSFM introduced the concept of compressibility to enhance the joint of subspaces. Each shape instance uses a different set of shape basis. The application of the above model converts the NRSFM into a three-wire problem, which can be solved using the decomposition technology or optimization strategy, implementing smooth spatial temporal, or tight 3-D shapes. The advantage of this method is that it can reduce excessive and fuzzy changes between the frames of a camera and structure.

Akhter et al. proposed that the trajectory of each point can be limited to a low-dimensional subspace. The advantage of this method based on the trajectory basis is that the target unknown basis, such as discrete cosine transform (DCT), can be used to reduce the number of unknown parameters and improve the accuracy of a 3-D structure. Gotardo and Martinez proposed a shape basis coefficient based on DCT by combining the shape basis and trajectory basis. Zhu et al. pointed out that sequences with poor reconstruction ability could be remedied by adding rigid key frames. They also emphasized the necessity of selecting a trajectory cardinal order instead of using all the DCT basis or applying normalization rules to coefficient vectors to obtain a sparse solution set. However, Zhu et al. could not exploit the known advantage of DCT coefficient distribution in natural signals. In addition, many methods use predefined basis to restrict the trajectory of each target point, thus transforming the trilinear problem into a bilinear problem, which can greatly simplify the trilinear operation problem. In the study of Valmadre and Lucey, a priori on the trajectory is introduced by using the differential of 3-D points. Its advantage is that it combines...
the shape basis and trajectory space. The advantage of combining the shape basis with the trajectory basis is that it can generate a smooth time trajectory of a nonrigid shape in the linear shape space.\(^3\)

In this study, we first use the automatic selection of the trajectory basis, which can not only effectively reduce the large \(K\) value used in the previous method, or the smaller \(K\) value, but can also reasonably select the combination of the trajectory basis. Firstly, the efficiency and accuracy of the reconstructed structure are maximized. Secondly, the matrix normal distribution establishes a model of the known trajectory space,\(^{10,16}\) which combines the spatial smoothness with the inherent temporal smoothness of the subspace. Based on the probabilistic model proposed by Agudo and Moreno-Noguer,\(^{28}\) this study adds accurate prior information and provides more accurate decomposition. The experiment shows the accuracy, versatility, and efficiency of our approach in sparse data sets.

**Low-rank model NRSFM**

The standard matrix decomposition method is generally used for the NRSFM problem. Suppose the motion model being studied consists of \(F\) feature points on each frame, and the 2-D position of each feature point in the image is marked as \(\{u^j_i, v^j_i\} \mid 1 \leq i \leq F; 1 \leq j \leq P\}, \) where \(u^j_i\) is the 2-D horizontal coordinate, \(v^j_i\) is the 2-D vertical coordinate, and \(o^j_i = [u^j_i; v^j_i]^\top\) is the 2-D image projection. All frames and feature points are represented together and arranged into a \(2F \times P\) observation matrix \(W\) as follows

\[
\begin{bmatrix}
\kappa^1_i & \cdots & \kappa^p_i \\
\vdots & \ddots & \vdots \\
\kappa^q_i & \cdots & \kappa^p_i \\
\end{bmatrix}
= \begin{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
R^1 \\
\vdots \\
R^p \\
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
X^1_i \\
\vdots \\
X^p_i \\
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\]

According to Agudo and Moreno-Noguer,\(^{28}\) \(\kappa^j_i\) is the zero mean coordinate, which is obtained by subtracting the average translation vector from the original coordinate, which means \(\kappa^j_i = o^j_i - \eta, \eta = \sum o^j_i / P\). Unlike other sparse\(^22\) technologies, the method proposed in this article can deal with track points lost owing to occlusion or outliers. We initially set the low-rank track prediction model with the 2-D value of \(\kappa^j_i\). These initial predictions are then used to calculate \(R\) and \(\Phi\). \(\Phi\) is a matrix of unknown coefficient vectors \(\phi^j_i \in \mathbb{R}^{3K \times 1}\), for each of the points \(j = \{1, \ldots, P\}\). After estimating \(R\) and \(\Phi\), \(\kappa^j_i\) can be further refined as follows

\[
B = \begin{bmatrix}
I_3 \otimes (b^1)^\top \\
\vdots \\
I_3 \otimes (b^p)^\top
\end{bmatrix}
\]

\[
\kappa^j_i = R^j_i (I_3 \otimes (b^j)^\top \phi^j_i + \eta
\]

where \(\otimes\) represents the Kronecker product. Matrix \(B\) is the trajectory basis matrix obtained after automatically selecting the basis, and matrix \(R\) is a block diagonal, consisting of a \(R^j_i \in \mathbb{R}^{2 \times 3}\) rotation of an orthophoto camera. In short, the NRSFM problem can be expressed as a given 2-D trajectory matrix \(W\), and the attitude parameter \(R\) and the 3-D shape \(S\) are recovered simultaneously. The observation matrix \(W\) can be decomposed into \(W = RS\), including \(R \in \mathbb{R}^{2F \times 3K}\) and \(S \in \mathbb{R}^{3K \times P}\). Then, we will further decompose it into \(W = \tilde{R} \tilde{S}\)

\[
W = RS = R\tilde{S} = \tilde{R} \tilde{S} = \tilde{A} A
\]

**Updating the transition matrix**

Because the SVD decomposition of matrices is not unique, matrices \(\tilde{A}\) and \(\tilde{A}\) are different from \(A\) and \(A\), respectively. For any correction matrix of \(Q \in \mathbb{R}^{3K \times 3K}\), \(\tilde{A} Q\) and \(Q^{-1} \tilde{A}\) are also valid factorizations. Therefore, to restore the transfer structure, we can obtain

\[
\Lambda = \tilde{A} Q, A = Q^{-1} \tilde{A}
\]

Instead of the entire matrix \(Q\), we only need to estimate the three columns of \(Q\) to correct for \(\tilde{A}\) and \(A\). Thus, for the \(F\) frame, there are \(3F\) constraints and \(9K\) unknown parameters in \(Q\). The rotation matrix \(R\) can be estimated by determining \(Q\).

**Automatic base selection**

The type, number, and combination of trajectory basis considerably influence the performance of the NRSFM algorithm, and the DCT trajectory basis is the optimal general trajectory basis. After the types of locus basis are determined, it is important to select the number and combination of locus basis.

In this article, an automatic selection algorithm based on trajectory basis is proposed\(^{29}\) The error in the actual 3-D structure \(S\) and first SVD decomposition \(S_i\) are analyzed in the frequency domain space. In addition, the \(K\) value is expanded and compared with the 3-D shape error of the sequential trajectory basis restoration. The optimal
Automatically select the trajectory basis algorithm

(1) A nonrigid body structure \( S^1 \) is obtained by SVD decomposition of observation matrix \( W \).

(2) Calculate the error of \( S^1 \) obtained from the decomposition of \( S \) and SVD of an actual nonrigid body structure of each frame, where \( p \) is the number of columns in the observation matrix

\[
\text{err}\text{S}(j) = \left( \sum_i \sqrt{\left( S_{[3i-1]j} - S_{[3i-1]j} \right)^2 + \left( S_{[3i-2]j} - S_{[3i-2]j} \right)^2 + \left( S_{3j} - S_{3j} \right)^2} \right) / p, i = 1, \ldots, f, j = 1, \ldots, p
\]  

(7)

(3) 1-D DCT is adopted for error \( \text{err}\text{S}(j) \), where \( z \) is the number of frames

\[
G_i = \frac{1}{2} \text{err}\text{S}(1) + \sum_{k=2}^{z} \text{err}\text{S}(k) \cos \left( \frac{\pi}{z} \left( i + \frac{1}{2} \right) k \right), z = 1, \ldots, f
\]  

(8)

Determine the spectrum of \( G_i \) and the magnitude of \(|G_i|\).

(4) Select a combination of trajectory basis according to the obtained error spectrum amplitude. Then, select \( K \) frequency points with the largest amplitude and use them to represent the actual error level \( \tau \). If the corresponding amplitude of the selected \( K \) points satisfies the following expression

\[
\sum_{i=1}^{K} \frac{|G_i|}{\sum_{i=1}^{\sigma} |G_i|} \geq \tau
\]  

(9)

then the position of the frequency point corresponding to the maximum amplitude of \( K \) points can be determined.

(5) Extract the corresponding \( K \) locus basis from the locus base space according to \( K \) frequency points. The final trajectory basis is further determined in the trajectory space so as to determine the trajectory cardinal order and combination form.

(6) Compare the structural errors after the automatic selection of trajectory basis and sequential reconstruction of trajectory basis, and select the combination of trajectory basis with lowest reconstruction error as the final combination of trajectory basis.

Automatically select trajectory basis-probability model

Solution and improvement of correlation matrix \( C \)

In real life, deformations observed in sports are often not singular, such as the movement of the face or the entire body, and there is similarity between the points of the objects under movement. Therefore, in the normal distribution of a matrix, we utilize a symmetric matrix \( C \) as a covariance matrix. Then, we assume that the observation matrix \( W \) is formed by a low-rank matrix \( C \) combined with the noise term \( E \). Therefore, the following idealized robust principal components analysis (PCA) problem can be obtained.

For the observation matrix \( W, W = C + E \), according to Liu et al.,\(^{30}\) where \( C \) is a low-rank matrix and \( E \) is a sparse matrix. We can obtain a conceptual solution to the above problem, which can be expressed as follows

\[
\min_{C,E} ||C||_1 + \lambda ||E||_1 \text{ subject to } C + E = W
\]  

(10)

where \( ||.||_1 \) is the nuclear norm and \( ||.||_1 \) is the \( L_1 \) norm. \( L_1 \) norm and \( L_0 \) norm can be sparse, and \( L_1 \) is widely used because of its superior optimal solution characteristics compared with those of \( L_0 \). Agudo and Moreno-Noguer\(^{28}\) introduced \( C \) into the covariance matrix of the probability model. According to Costeira and Kanade,\(^{12}\) the method for solving the covariance matrix is modified to: \( X = XC + E \). Therefore, we convert the expression of the robust PCA problem to

\[
\arg \min_{C,E} ||C||_1 + \lambda ||E||_1 \text{ subject to } X = XC + E
\]  

(11)

For this problem, we can use the exact augmented Lagrange multiplier method to solve the problem. However, the operational cost is relatively high. Therefore, an inexact augmented Lagrange multiplier (IALM) method is used to solve the problem. Compared with the exact algorithm, the inexact Lagrange multiplier method\(^{31}\) has a considerable improvement in computing speed while maintaining accuracy.

In the IALM algorithm, when \( \mu_k \) increases linearly, the exact Lagrange multiplier method will converge linearly. When \( \mu_k \) increases rapidly, the convergence speed will also be faster. However, when \( \mu_k \) is very large, the convergence speed of solution will be low. Therefore, the IALM algorithm starts by reducing the computation time of subproblems.

In the precise Lagrange multiplier method, the IT algorithm is used to solve the subproblem
\[(C_{k+1}, E_{k+1}) = \arg \min_{C, E} L(C, E, Y_k, \mu_k)\] iteratively, where \(C_{k+1}\) and \(E_{k+1}\) are, respectively, the values of \(C\) and \(E\), updated after the KTH iteration, \(\mu_k\) is the penalty parameter, and \(Y_k\) is the Lagrange multiplier. According to the experiment,\(^{31}\) for \(C_{k+1}\) and \(E_{k+1}\), we do not need to be precise in the above subproblems. A solution of SVD can obtain a relatively approximate optimal solution, which is sufficient to achieve the desired effect. Therefore, in the IALM algorithm, we removed the iterative solution of the subproblem by using the IT algorithm, and replaced IT with a direct SVD solution, as follows

\[
(U, S, V) = \text{SVD}(W - E_k + \mu_k^{-1} Y_k)
\]

\[
C_{k+1} = U \chi_{\mu_k^{-1}}(S) V^T
\]

\[
E_{k+1} = \chi_{\mu_k^{-1}}(W - C_{k+1} + \mu_k^{-1} Y_k)
\]

Among them

\[
\chi_{\varepsilon}(x) = \begin{cases} 
  x - \varepsilon & \text{if } x > \varepsilon \\
  x + \varepsilon & \text{if } x < -\varepsilon \\
  0 & \text{otherwise}
\end{cases}
\]

where \(\varepsilon\) is represented as a soft threshold, \(x = W - C_{k+1} + \mu_k^{-1} Y_k\).

In this manner, one layer of iterative loop can be removed, considerably increasing the calculation time. After obtaining the accelerated correlation matrix \(C_{k+1}\), we normalize it and force it to be unitary on the diagonal

\[
C^* = C_{k+1} \odot (1_N \frac{1}{\sqrt{N}} - I_N) + I_N
\]

where \(\odot\) represents the hada code product and \(1_N\) represents a vector.

**Adjusting prior rows and column covariance**

Matrix normal distribution uses Kronecker covariance to provide the idea of natural combination around data. Specifically, the normal random variable \(X\) in the matrix represents the matrix itself. The distribution is parameterized by the mean matrix and two covariance matrices, which represent the covariance of the rows and columns of matrix \(X\).

The prior covariance matrix has a considerable influence on the 3-D structure restoration. The processed prior covariance can increase the stability and accuracy of 3-D structure restoration. In other words, prior covariance with higher accuracy can provide better restoration to the 3-D structure movement. Therefore, we adjusted the initial row covariance \(I_{3K}\) and column covariance inverse \(C^*\) accordingly, where \(K\) is the number of trajectory basis

\[
\text{vec}(\Phi) \sim N(\mu; \Sigma_c \otimes \Sigma_s)
\]

Then, the matrix normal distribution of \(\Theta\) is

\[
P(\Phi; \mu, \Sigma_c, \Sigma_s) = (2\pi)^{-pq/2} |(\Sigma_c \otimes \Sigma_s)^{-1}|^{1/2} \times \exp\left\{ -\frac{1}{2} \text{vec}(\Phi - \mu)^T (\Sigma_c \otimes \Sigma_s)^{-1} \text{vec}(\Phi - \mu) \right\}
\]

(18)

According to Gupta and Nagar,\(^{32}\) equation (18) can be simplified as follows

\[
P(\Phi; \mu, \Sigma_c, \Sigma_s) = (2\pi)^{-pq/2} |\Sigma_c|^{-q/2} |\Sigma_s|^{-p/2} \times \exp\left\{ -\frac{1}{2} \text{tr} \left[ \Sigma_c^{-1} (\Theta - \mu) \Sigma_c^{-1} (\Theta - \mu)^T \right] \right\}
\]

(19)

Then, the logarithmic likelihood function of parameter \(\Theta = (\mu, \Sigma_c, \Sigma_s)\) is

\[
\ln P(\Theta; \mu, \Sigma_c, \Sigma_s) = \frac{p}{2} \ln |\Sigma_c| + \frac{q}{2} \ln |\Sigma_s| - \frac{1}{2} D(\Theta, \mu)
\]

(20)

where

\[
D(\Theta, \mu) = \text{tr}\left[ \Sigma_c^{-1} (\Theta - \mu) \Sigma_c^{-1} (\Theta - \mu)^T \right]
\]

According to equation (20), and the maximum likelihood estimation (MLE) can be obtained as follows

\[
\hat{\Sigma}_c = \frac{1}{p} (\Theta - \mu)^T \Sigma_c^{-1} (\Theta - \mu),
\]

\[
\hat{\Sigma}_s = \frac{1}{q} (\Theta - \mu)^T \Sigma_s^{-1} (\Theta - \mu)^T
\]

(21)

However, for any \(k \neq 0\), we define \(\hat{\Sigma}_c = k \Sigma_c, \hat{\Sigma}_s = \Sigma_s / k\) and then \(\hat{\Sigma}_c \odot \hat{\Sigma}_s = \Sigma_c \otimes \Sigma_s\). Both estimates produce the same covariance matrix of the population. To solve this problem, according to Glanz and Carvalho,\(^{33}\) we propose the following modifications to the model

\[
\text{vec}(\Theta) \sim N(\mu(\text{vec}(\mu), \zeta^2 \Sigma_c \otimes \Sigma_s)
\]

(22)

To estimate \(\zeta^2\), similar to equation (20), the new likelihood function can be obtained

\[
\ln P(\Theta; \zeta^2, \mu, \Sigma_c, \Sigma_s) = \frac{p}{2} \ln |\Sigma_c| + \frac{q}{2} \ln |\Sigma_s| - \frac{1}{2} \zeta^2 D(\Theta, \mu)
\]

(23)

The MLE generating \(\zeta^2\) is

\[
\hat{\zeta}^2 = \frac{1}{pq} (\Theta - \mu)^T (\Sigma_c \otimes \Sigma_s)^{-1} (\Theta - \mu)
\]

(24)

However, because the variance scale is determined by \(\zeta^2\), when \(\zeta^2 \leq 1\), the update of matrix \(\hat{\Sigma}_c\) and \(\hat{\Sigma}_s\) will be abnormal. This study uses the empirical value instead. We need to consider the scale constraints of \(\Sigma_c\) and \(\Sigma_s\) when deriving their MLE. For this, we use the results of\(^{33}\)

\[
\hat{\Sigma}_c = \text{ADJUST}\left\{ p, \hat{\zeta}^2, (\Theta - \mu)^T \hat{\Sigma}_s^{-1} (\Theta - \mu) \right\}
\]

\[
\hat{\Sigma}_s = \text{ADJUST}\left\{ q, \hat{\zeta}^2, (\Theta - \mu)^T \hat{\Sigma}_c^{-1} (\Theta - \mu) \right\}
\]

(25)
Parameter solution

The observed 2-D locus point matrix $W$ is accompanied by Gaussian noise, which is represented by the matrix $N \in \mathbb{R}^{2F \times P}$. $W$ can be redefined as follows

$$W = A\Phi + N$$ (26)

According to Agudo and Moreno-Noguer, $^{28}$ we can include the accelerated correlation matrix $C^x$ in the probability model in the form of covariance

$$p(\text{vec}(\Phi)) \sim N(\text{vec}(0); \hat{\Sigma}_c \otimes \hat{\Sigma}_s)$$ (27)

$$p(\text{vec}(N)) \sim N(\text{vec}(0); \sigma^2 I_{2F} \otimes \hat{\Sigma}_s)$$ (28)

where $\text{vec}(\cdot)$ represents the vectorization operator of a matrix.

E step: In step E, we estimate the conditional distribution of potential variable $\Theta$. We apply Bayes’ rule to the equation. According to some properties of matrix variational normal distribution, $^{32,34}$ it can be known that this distribution is also a Gaussian distribution

$$p(\text{vec}(\Phi|W, A, \sigma^2)) \sim N(\text{vec}(\gamma_\Phi); \Sigma_\Phi \otimes \hat{\Sigma}_s)$$ (29)

$$\gamma_\Phi = (A^T A + \sigma^2 \hat{\Sigma}_s)^{-1} A^T W$$ (30)

$$\Sigma_\Phi = \sigma^2 N(A^T A + \sigma^2 \hat{\Sigma}_s)^{-1}$$ (31)

M step: We update model parameters $A$ and $\sigma^2$ and obtain

$$A_{j+1} \leftarrow DA_j \left( \sigma^2 \hat{\Sigma}_c + (A_j^T A_j + \sigma^2 \hat{\Sigma}_c)^{-1} A_j^T \right)^{-1}$$

$$\sigma^2_{j+1} \leftarrow \frac{1}{2T} tr \left( D - DA_j \left( A_j^T A_j + \sigma^2 \hat{\Sigma}_c \right)^{-1} A_j^T \right)$$ (32)

Experimental results

Selection of base

The number and combination of trajectory basis have a considerable influence on the structural errors of reconstructing 3-D nonrigid bodies. This study uses the spectrum analysis method to analyze the frame error of an actual nonrigid structure $S$ of the known data set and the SVD decomposition of the observation matrix $W$ to obtain a nonrigid structure $S^T$. The number of trajectory basis $K$ and the combination form of trajectory basis are determined. Then, the structural errors after the automatic selection of trajectory basis and sequential trajectory basis are compared, and the combination of trajectory basis with smaller reconstruction errors is taken as the final trajectory basis.

The recovery method of Figure 1 is based on previous studies, $^{6-9}$ which also employ a similar nonrigid reconstruction method. Figure 1 shows the trajectory basis combination of the yoga 3-D structure of the nonrigid body restoration shown in the literature. $^{3,5,10-12,14-16,19}$ The first image on the left is the elevation diagram of frame 50, the second image is of frame 140, the third image is of frame 210, and the fourth image is of frame 240.

Figure 2 shows the mean error comparison of 3-D points of the yoga data set of the frames of 50, 100, 150, 200, 250, and 300 for the automatic selection base method and the sequential selection method; the mean errors were obtained for $K$ tracks of the cardinal head. It can be clearly seen that the automatic selection of basis is superior to the sequential selection of basis.

However, different data sets adopt different trajectory basis forms. For example, in the drink data set, the sequential selection of basis performs better than the automatic selection of basis, as shown in Figure 3.

Method comparison

For the quantitative evaluation, we follow the indicators used in the studies of Dai et al. $^{35}$ and Gotardo and Martinez $^{3}$ to show the average rotation error $e_R$ and standardized average 3-D error $e_S$, which are defined as
In frame $f$, $\hat{R}$ is the estimated rotation matrix and $R$ is the corresponding real rotation of the ground. $e_S$ is calculated as

$$e_S = \frac{1}{\nu FP} \sum_{f=1}^{F} \sum_{p=1}^{P} e_p^2 = \frac{1}{3F} \sum_{f=1}^{F} \left( \sigma_x^2 + \sigma_y^2 + \sigma_z^2 \right)$$

(35)

where $e_p^2$ is the 3-D reconstruction error of point $p$ in the $f$ coordinate system. $\sigma_x$, $\sigma_y$, and $\sigma_z$ represent the standard deviation of the $x$, $y$, and $z$ coordinates of the original shape.

**Table 1.** Average rotation error $e_R$.

| Data   | Met    | EM-PPCA$^{10}$ | PTA$^{22}$ | PCPTA$^{28}$ | BMM$^{35}$ | Ours | $e_R$ |
|--------|--------|----------------|------------|--------------|------------|------|-------|
| Drink  |        | 0.186          | 0.006      | 0.006        | 0.007      | 0.004|       |
| Stretch|        | 0.749          | 0.055      | 0.055        | 0.068      | 0.057|       |
| Yoga   |        | 0.688          | 0.162      | 0.106        | 0.088      | 0.105|       |
| Pick-up|        | 0.417          | 0.154      | 0.155        | 0.121      | 0.151|       |
| Dance  |        | ---            | ---        | ---          | ---        | ---  | ---   |

EM-PPCA: expectation-maximization-probabilistic principal components analysis; PTA: point trajectory approach.

$^a$The missing data in Table 1 is due to the lack of reality rotation matrix $R$ of dance data so that the error cannot be analyzed.
in the $f$ coordinate system. When the surface truth 3-D data or rotation data are available, we provide $e_x$ and $e_R$, respectively, where $K$ is the number of trajectory basis. Table 1 shows the errors of the five 3-D reconstruction methods. Because of the lack of codes of comparative papers, the errors of these data are derived from the study of Agudo and Moreno-Noguer.28

In the experiment, the $K$ values of the five data sets in the subspace are yoga ($K = 11$), pick-up ($K = 12$), drink ($K = 13$), stretch ($K = 12$), and dance ($K = 5$).

Table 1 shows average rotation error $e_R$ and Table 2 shows standardized average 3-D error $e_S(K)$.

Further, the stretch is used as the experimental data set. The blue trajectory represents the error curve after the reconstruction of the automatic selection of basis, and the red trajectory represents the error curve after the reconstruction of the sequential selection of trajectory basis. The graph shows the coordinates of each feature point in each dimension. The reconstruction error of the automatic trajectory basis selection method was low. The reconstruction error shown in the figure was obtained using the following equation

\[
\text{err}_X = \sqrt{\sum_{i} \frac{||T_{xij} - \hat{T}_{xij}||^2}{n}, \quad i = 1, \ldots, m; j = 1, \ldots, n}
\]

\[
\text{err}_Y = \sqrt{\sum_{i} \frac{||T_{yij} - \hat{T}_{yij}||^2}{n}, \quad i = 1, \ldots, m; j = 1, \ldots, n}
\]

\[
\text{err}_Z = \sqrt{\sum_{i} \frac{||T_{zij} - \hat{T}_{zij}||^2}{n}, \quad i = 1, \ldots, m; j = 1, \ldots, n}
\]

where $T_{x}$, $T_{y}$, $T_{z}$ are the initial motion trajectories before reconstruction and $\hat{T}_{x}$, $\hat{T}_{y}$, $\hat{T}_{z}$ are the motion trajectories obtained after reconstruction.

Figures 4 and 5 show an analysis of the reconstruction errors of the trajectory basis\textsuperscript{3,5,10–12,14–18,20} and the first 11 trajectory basis in order; moreover, a comparison of the reconstruction errors and overall errors of each point in $X$, $Y$, and $Z$ coordinates is shown. Figure 6 shows the percentage of frames 50, 100, 150, 200, 250, and 300 that are lower than the number of basis points that are selected sequentially. It can be clearly seen that the error level of 11 reconstructed trajectory basis is much lower than that of the trajectory basis selected in sequence. That is, the reconstruction accuracy and time are considerably improved, which demonstrates that the method proposed in this study can improve efficiency on the premise of ensuring reconstruction accuracy.

Table 3 compares the execution time (in second) of probabilistic correlation point trajectory approach (PCPTA)\textsuperscript{28} with that of block matrix method (BMM),\textsuperscript{25} which are two highly accurate and advanced methods to recover 3-D non-rigid body structures. All methods are executed in MATLAB 2018b, where $K$ is the number of locus basis.

From Table 2, we can see that the performance of the proposed method is better in terms of time and accuracy. Before acceleration, the method shows a lower error level. Its emphasis is that the value of $C$ of the correlation matrix has a higher accuracy. However, although the method of determining the acceleration has slightly increased the error, the speed increased significantly. Unfortunately, the source code of the first two methods is not open. Thus, we cannot complete the comparison of noise observations.

The reason why our unaccelerated method is longer than the PCPTA method is that we use the automatic selection base and the ADJUST method.

Figures 7 to 9 show the restored graph of the pick-up data set recovered by the algorithm of the proposed automatically selecting trajectory basis-probability model after acceleration. The sparse 3-D nonrigid body recovery structures of frames 5, 50, 150, and 200 are shown in the order of left to right. Here, $\tau = 0.96$, trajectory basis item $K = 9$, and the combination of the base positions of the obtained trajectory is selected as described in the literature.\textsuperscript{3,5,10,11,14,15,18,19} It can be seen that, compared with the previous methods,\textsuperscript{10,22,28,35} the reduction in trajectory basis purpose does not significantly affect the recovery accuracy, but can accelerate the recovery efficiency.

| Met       | EM-PPCA\textsuperscript{10} | PTA\textsuperscript{22} | PCPTA\textsuperscript{28} | BMM\textsuperscript{35} | Ours |
|-----------|-----------------------------|-------------------------|---------------------------|-------------------------|------|
| Drink     | 0.261 (7)                   | 0.025 (13)              | 0.0274 (13)               | 0.027 (12)              | 0.0277 (10) |
| Stretch   | 0.458 (7)                   | 0.109 (12)              | 0.1620 (12)               | 0.103 (11)              | 0.0785 (11) |
| Yoga      | 0.445 (8)                   | 0.1447 (12)             | 0.1623 (12)               | 0.115 (10)              | 0.1659 (10) |
| Pick-up   | 0.423 (14)                  | 0.237 (12)              | 0.2368 (12)               | 0.173 (12)              | 0.2870 (9)  |
| Dance     | 0.339 (4)                   | 0.296 (5)               | 0.257 (4)                 | 0.188 (10)              | 0.205 (5)   |
| Average error | 0.385                     | 0.166                   | 0.1691                    | 0.121                   | 0.152      |

3-D: three-dimensional.
Figure 4. Errors obtained when stretching at frame 50 at X, Y, and Z coordinates of each feature point and overall reconstruction errors.

Figure 5. Errors obtained when stretching reaches the 200th frame at X, Y, and Z coordinates of each feature point and overall reconstruction errors.
Figure 10 shows an example of the dinosaur toy, the recovery structures of frames 1, 8, 42, and 67. Figure 11 shows an example of cubes toys, the recovery structures of frames 1, 72, 197, and 200, respectively. In the cubes model, trajectory basis item $K = 2$, the trajectory is selected as described in the literature.\textsuperscript{5,10} It should be mentioned that although both cubes are rigid individually, they are connected as a whole by a wire. So, when another wire which connects to one cube is pulled, they move like a no-rigid body as a whole. This kind of movement is simpler than that of the human body or dinosaur toy. Therefore, the number of $K$ here is relatively low. In the dinosaur model, the trajectory is selected as described in the literature.\textsuperscript{3,5,10,12–16,18,19,24} It can be seen from the figure that the method in this article can recover the 3-D nonrigid structure of two toy models well. The data set is from http://mocap.cs.cmu.edu. In particular, the examples of dinosaur toy and cubes toy lack the reality 3-D structure matrix $S$ so that the standardized average 3-D error $e_S$ analysis cannot be analyzed.

| Trajectory of the base | BMM\textsuperscript{15} | PCPTA\textsuperscript{28} | Proposed method without acceleration | Proposed method with acceleration |
|------------------------|--------------------------|---------------------------|------------------------------------|---------------------------------|
| Reconstruct the expected value of the error | 0.103 ($K = 11$) | 0.1620 ($K = 12$) | 0.0785 ($K = 11$) | 0.0998 ($K = 11$) |
| Reconstruction time | 46.23 | 16.32 | 28.34 | 18.45 |

Figure 6. Stretched data set in frames 50, 100, 150, 200, 250, and 300 is better than the percentage graph of sequential DCT basis points. DCT: discrete cosine transform.

Table 3. Reconstruction errors and reconstruction time of the four methods.

Figure 7. Pick-up.

For the dance data set, when the parameter $\zeta^2$ in equations (5) to (13) is obtained as 0.2072, which is less than 1, $\tilde{\Sigma}_c$
and $\hat{\Sigma}_1$ matrix updates are abnormal. At this time, we can manually adjust the value. When setting $\zeta^2 = 1000$, we took the trajectory basis mesh $K = 5$, and selected the sequence for restoration. We obtained the standardized average 3-D error as $e_S = 0.2057$. When $\zeta^2 \leq 1$, the range of artificial regulation $\zeta^2$ is $1 \leq \zeta^2 \leq 1500$.

**Conclusion**

This study adopts a model of automatic selection of trajectory basis combined with a probability framework. The number and combination of trajectory basis are obtained such that the trajectory basis can maximize the reconstruction accuracy for recovering nonrigid structures. The latter
incorporates the low-rank trajectory model into the probability framework of a matrix normal distribution and can also improve the restoration efficiency of a 3-D nonrigid structure within the allowable range of error accuracy by using more precise prior conditions. The combination of the two methods can achieve the accurate reconstruction of sparse data sets. More importantly, the proposed reconstruction method is more accurate and efficient than most previous methods. In the days to come, we want to enhance the solution of the correlation matrix $C$, not only in accuracy, but also in solution speed. In addition, we hope to find or synthesize useful dense data sets for 3-D nonrigid reconstruction. After all, dense data sets are closer to actual daily life activities and more authentic; however, this will be a huge challenge.

Acknowledgement
Thanks to the test data set provided by http://cvlab.lums.edu.pk/nrsfm.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Natural Science Foundation of Zhejiang Province (LZ20F020003, LY17F020034, LSZ19F010001), the National Natural Science Foundation of China (61272311, 61672466), and the 521 Project of Zhejiang Sci-Tech University.

ORCID iD
Wenqing Huang https://orcid.org/0000-0001-6166-5119

References
1. Del Bue A, Llad X, and Agapito L. Non-rigid metric shape and motion recovery from uncalibrated images using priors. In: IEEE conference on computer vision and pattern recognition, New York, NY, USA, 17–22 June 2006, pp. 1191–1198. IEEE.
2. Antonio A and Moreno-Noguer F. Force-based representation for non-rigid shape and elastic model estimation. IEEE Trans Patt Anal Mach Intell 2018; 40(9): 2137–2150.
3. Gotardo PFU and Martinez AM. Computing smooth time-trajectories for camera and deformable shape in structure from motion with occlusion. IEEE Trans Patt Anal Mach Intell 2011; 33(10): 2051–2065.
4. Zaheer A, Akhter I, Baig MH, et al. Multiview structure from motion in trajectory space. In: IEEE international conference on computer vision, Barcelona, Spain, 6–13 November 2011, pp. 2447–2453. IEEE.
5. Tomasi C and Kanade T. Shape and motion from image streams under orthography: a factorization approach. Int J Comput Vis 1992; 9(2): 137–154.
6. de Jesus Rubio J and Enrique JG. Recursive least squares for a manipulator which learns by demonstration [mínimos cuadrados recursivos para un manipulador que aprende por demostración]. Rev Iberoam Autom Inform Ind 2019; 16(2): 147–158.
7. de Jesus Rubio J. Discrete time control based in neural networks for pendulums. Appl Soft Comput 2018; 68: 821–832.
8. Kumar J, Kumar V, and Rana KPS. Design of robust fractional order fuzzy sliding mode PID controller for two link robotic manipulator system. J Intell Fuzzy Syst 2018; 35(5): 5301–5315.
9. Rubio J. Modified optimal control with a back propagation network for robotic arms. IET Control Theory Appl 2012; 6(14):2216–2225.
10. Torresani L, Hertzmann A, and Bregler C. Nonrigid structure-from-motion: estimating shape and motion with hierarchical priors. IEEE Trans Patt Anal Mach Intell 2008; 30(5): 878–892.
11. Bregler C, Hertzmann A, and Biermann H. Recovering non-rigid 3D shape from image streams. In: IEEE conference on computer vision and pattern recognition, Hilton Head Island, SC, USA, 15 June 2000, pp. 690–696. IEEE.
12. Costeira J and Kanade T. A multibody factorization method for independent moving objects. Int J Comput Vis 1998; 29(3): 159–179.
13. Bascle B and Blake A. Separability of pose and expression infacial tracing and animation. In: Proceedings of
14. Torresani L, Yang DB, Alexander EJ, et al. Tracking and modeling non-rigid objects with rank constraints. In: Proceedings of IEEE conference on computer vision and pattern recognition, Kauai, HI, USA, 8–14 December 2001, pp. 493–500. IEEE.

15. Torresani L, Hertzmann A, and Bregler C. Learning non-rigid 3D shape from 2D motion. In: Advances in neural information processing systems, British Columbia, Canada, 13–16 December 2004, pp. 1555–1562.

16. Rabaud V and Belongie SJ. Re-thinking non-rigid structure from motion. In: IEEE conference on computer vision and pattern recognition, Anchorage, AK, USA, 23 June 2008, pp. 1–8. IEEE.

17. Agudo A and Moreno-Noguer F. Learning shape, motion and elastic models in force space. In: IEEE international conference on computer vision, Araucano Park, Las Condes, Chile, 11–18 December 2015, pp. 756–764. IEEE.

18. Kong C and Lucey S. Prior-less compressible structure from motion. In: IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 27–30 June 2016, pp. 4123–4131. IEEE.

19. Agudo A and Moreno-Noguer F. Combining local-physical and global-statistical models for sequential deformable shape from motion. Int J Comput Vis 2017; 122(2): 371–387.

20. Agudo A, Moreno-Noguer F, Calvo B, et al. Real-time 3D reconstruction of non-rigid shapes from single moving camera. Comput Vis Image Underst 2016; 153(12): 37–54.

21. Lee M, Cho J, Choi CH, et al. Procrustean normal distribution for non-rigid structure from motion. In: IEEE conference on computer vision and pattern recognition, Portland, OR, USA, 23–28 June 2013, pp. 1280–1287. IEEE.

22. Akhter I, Sheikh Y, Khan S, et al. Nonrigid structure from motion in trajectory space. In: Advances in neural information processing systems, Vancouver, Canada, 8–10 December 2008, pp. 41–48.

23. Gotardo PFU and Martinez AM. Non-rigid structure from motion with complementary rank-3 space. In: Computer vision and pattern recognition, Colorado Springs, CO, USA, 20 June 2011, pp. 3065–3072. IEEE.

24. Zhu Y, Cox M, and Lucey S. 3D motion reconstruction for real-world camera. In: Computer vision and pattern recognition, Colorado Springs, CO, USA, 20 June 2011, pp. 1–8. IEEE.

25. Gonzalez RC and Woods RE. Digital image processing. 2nd ed. Upper Saddle River: Prentice Hall, 2002.

26. Park HS, Shiratori T, Matthews I, et al. 3D reconstruction of a moving point from a series of 2D projections. In: European conference on computer vision, Heraklion, Crete, Greece, 5 September 2010, pp. 158–171. Springer.

27. Valmadre J and Lucey S. General trajectory prior for non-rigid reconstruction. In: IEEE conference on computer vision and pattern recognition, Providence, RI, USA, 16 June 2012, pp. 1394–1401. IEEE.

28. Agudo A and Moreno-Noguer F. Scalable, efficient, and accurate solution to non-rigid structure from motion. Comput Vis Image Underst 2018; 167: 121–133.

29. Wang Y, Cheng J, Zheng J, et al. Analysis of wavelet basis selection in optimal trajectory space finding for 3D non-rigid structure from motion. Int J Wavelets Multiresolution Inf Process 2014; 12(2): 1450023.

30. Liu G, Lin Z, Yan S, et al. Robust recovery of subspace structures by low-rank representation. IEEE Trans Patt Anal Mach Intell 2012; 35(1): 171–184.

31. Lin Z, Chen M, and Ma Y. The augmented Lagrange multiplier for exact recovery of corrupted low-rank matrices. ArXiv preprint arXiv:1009.5055, 2010.

32. Gupta AK and Nagar DK. Matrix variate distributions. London: Chapman & Hall/CRC, 2000.

33. Glanz H and Carvalho L. An expectation–maximization algorithm for the matrixnormal distribution with an application in remote sensing. J Multivar Anal 2018; 167: 31–48.

34. Tipping ME and Bishop CM. Mixtures of probabilistic principal component analysers. Neural Comput 1999; 11(2): 443–482.

35. Dai Y, Li H, and He M. A simple prior-free method for non-rigid structure from motion factorization. Int J Comput Vis 2014; 107(2): 101–122.