## sensors

ISSN 1424-8220
www.mdpi.com/journal/sensors

## Article

# Non-Rigid Structure Estimation in Trajectory Space from Monocular Vision 

Yaming Wang, Lingling Tong *, Mingfeng Jiang and Junbao Zheng<br>School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China; E-Mails: ywang@zstu.edu.cn (Y.W.); m.jiang@zstu.edu.cn (M.J.);<br>zhengjunbao@zstu.edu.cn (J.Z.)

* Author to whom correspondence should be addressed; E-Mail: 15705811142@163.com; Tel.: +86-571-8684-3328.

Academic Editor: Vittorio M. N. Passaro
Received: 2 July 2015 / Accepted: 6 October 2015 / Published: 12 October 2015


#### Abstract

In this paper, the problem of non-rigid structure estimation in trajectory space from monocular vision is investigated. Similar to the Point Trajectory Approach (PTA), based on characteristic points' trajectories described by a predefined Discrete Cosine Transform (DCT) basis, the structure matrix was also calculated by using a factorization method. To further optimize the non-rigid structure estimation from monocular vision, the rank minimization problem about structure matrix is proposed to implement the non-rigid structure estimation by introducing the basic low-rank condition. Moreover, the Accelerated Proximal Gradient (APG) algorithm is proposed to solve the rank minimization problem, and the initial structure matrix calculated by the PTA method is optimized. The APG algorithm can converge to efficient solutions quickly and lessen the reconstruction error obviously. The reconstruction results of real image sequences indicate that the proposed approach runs reliably, and effectively improves the accuracy of non-rigid structure estimation from monocular vision.


Keywords: non-rigid structure estimation; monocular vision; trace minimization constraint; rank minimization; APG algorithm

## 1. Introduction

Recently, non-rigid structure estimation from monocular vision, which can recover the time varying 3D coordinates of points on a non-rigid object from their 2D places in a video sequence, has become a popular research topic. Generally, two major methods, i.e., the trajectory basis method and shape basis method, are often used to solve non-rigid structure estimation problems. The factorization method was first proposed to recover rigid structure by Tomasi and Kanade [1], and the method was extended to solve the non-rigid structure problem in the seminal paper by Bregler et al. [2]. The core idea is that shapes observed from motion can be represented by the linear combination of a compact set of basis shapes. Each instantaneous structure, such as a running person, can be expressed as a point in the linear space of shapes spanned by the shape basis. A large number of methods have subsequently been developed [3-5], which promoted improved performances of shape basis. However, the shape basis has some limitations since it has a specific nature and can't generally apply to all non-rigid bodies. The shape basis of a dancer moving, for example, cannot be recycled to compactly represent a person running. So, as an alternative to a shape space, Akhter et al. [6,7] proposed to represent the time-varying structure of a non-rigid object by using a linear combination of a set of basis trajectories, which was called Point Trajectory Approach (PTA). The primary advantage of PTA was that the trajectory basis can be predefined to be close to many real trajectories, which resulted in a significant reduction in unknowns, and corresponding stability improvement in estimation. Zhu et al. [8] pointed out that the importance of selecting the number of trajectory basis, rather than the more bases used the better, which varied from different models. On this basis, Gotardo and Matinez [9] combined the shape method and trajectory method, which can further improve the reconstruction performance. Recently, Rehan et al. [10] proposed a novel constraint in the form of local rigidity, which gave stable results in challenging realistic scenarios with small camera motions and shorter sequences. Minsk et al. [11] introduced new constraints that were more effective for non-rigid structure estimation, which constrained the motion parameters so that the 3D shapes were most closely aligned to each other, making the rank constraints unnecessary. Then they proposed a new probabilistic model in [12], which incorporated the smoothness constraint without requiring any prior knowledge. This approach regarded the sequence of 3D shapes as a simple stationary Markov process with Procrustes alignment, whose parameters were learned during the fitting process. Antonio et al. [13] proposed an online solution to estimate non- rigid structure, which modeled non-rigid deformations as a linear combination of some mode shapes obtained using modal analysis from continuum mechanics. However, the underlying principle behind most approaches was to model deformations using a low-rank shape [2,9,14,15], and it improved the accuracy of the non-rigid structure estimation.

In order to further improve the accuracy of non-rigid structure estimation, low rank condition of structure matrix is also investigated in this paper, and the APG algorithm is proposed to optimize the structure matrix, which can quickly converge to an efficient solution. Many trajectory bases can be used to recover the structure [16], such as the Discrete Cosine Transform (DCT) basis, Walsh Hadamard Transform (WHT) basis and Discrete Wavelet Transform (DWT) basis. In this paper, the predefined DCT basis is introduced to recover the motion and structure of the non-rigid object. For the 2D signal of $M \times N$ sample points, the DCT formula can be defined as follows:

$$
\begin{equation*}
y(k, d)=u(k) \sum_{n=1}^{N} u(n, d) \cos \frac{\pi(2 N-1)(k-1)}{2 N} \tag{1}
\end{equation*}
$$

where $k=1,2, \ldots, M$ and $\mu(k)$ is the coefficient:

$$
u(k)=\left\{\begin{array}{l}
\frac{1}{\sqrt{M}}, k=1  \tag{2}\\
\sqrt{\frac{2}{M}}, 2 \leq k \leq M
\end{array}\right.
$$

In the paper, the APG method is proposed to solve the problem of non-rigid structure estimation. A new constraint, called trace minimization constraint of the rectification matrix, is introduced to narrow the solution space and improve the computiational speed of our algorithm. The proposed method can effectively estimate both 3D structures of non-rigid objects and the camera motion. The experimental results on real image sequences indicate that the proposed approach effectively improves the accuracy of non-rigid structure estimation from monocular vision.

This paper is organized as follows: the problems are formally described in Section 2 before briefly introducing how to get the initial structure matrix $S$ by using PTA method in Section 3. In Section 4, the APG algorithm is introduced to optimize $S$. Experimental results are presented in Section 5. Finally, a summary and future works are discussed.

## 2. The Problem of Non-Rigid Structure Estimation

In fact, 3D reconstruction of non-rigid motion from monocular vision is equivalent to the decomposition of the measurement matrix $W$, that is decomposing the $W$ into the rotation matrix $R$ of camera and the structure matrix $S$ of the non-rigid object. This problem can be simplified to estimate the rectification matrix $Q$. The PTA method is implemented to estimate the corresponding unknown parameters by a series of constraints, and to recover the structure $S$ of the non-rigid object. Then, the APG algorithm is used to reconstruct the structure matrix $S$ which is calculated by PTA, and can further improve the accuracy of the non-rigid structure estimation from monocular vision.

After feature point correspondence, the measured 2D trajectories can be included in the measurement matrix $W$, containing the location of $N$ image points across $M$ frames:

$$
W=\left[\begin{array}{c}
X  \tag{3}\\
Y
\end{array}\right]=\left[\begin{array}{ccc}
x_{11} & \cdots & x_{1 N} \\
y_{11} & \cdots & y_{1 N} \\
\vdots & \vdots & \vdots \\
x_{M 1} & \cdots & x_{M N} \\
y_{M 1} & \cdots & y_{M N}
\end{array}\right]
$$

The measurement matrix $W$ can be decomposed as $W=R S$, where $R$ is a $2 M \times 3 M$ matrix, $R_{i}(i=1,2, \cdots M)$ is an orthogonal projection matrix:

$$
R=\left[\begin{array}{lll}
R_{1} & &  \tag{4}\\
& \ddots & \\
& & R_{M}
\end{array}\right]
$$

The structure matrix $S$ is a $3 M \times N$ matrix, and the structure at a time instant $t$ can be represented as follows:

$$
S=\left[\begin{array}{ccc}
X_{t 1} & \cdots & X_{t N}  \tag{5}\\
Y_{t 1} & \cdots & Y_{t N} \\
Z_{t 1} & \cdots & Z_{t N}
\end{array}\right]
$$

## 3. The Calculation of Matrix $S$ Using PTA

The structure matrix $S$ can be decomposed into the trajectory basis matrix $\Theta$ and the coefficient matrix $A, S_{3 M \times N}=\Theta_{3 M \times 3 K} A_{3 K \times N}$. Defining the equation $\Lambda=R \Theta$, the elements of the matrix $\Lambda$ are as follows:

$$
\Lambda=R \Theta=\left(\begin{array}{ccc}
r_{1}^{1} \theta_{1}^{T} & r_{2}^{1} \theta_{1}^{T} & r_{3}^{1} \theta_{1}^{T}  \tag{6}\\
r_{4}^{1} \theta_{1}^{T} & r_{5}^{1} \theta_{1}^{T} & r_{6}^{1} \theta_{1}^{T} \\
& \vdots & \\
r_{1}^{F} \theta_{1}^{T} & r_{2}^{F} \theta_{1}^{T} & r_{3}^{F} \theta_{1}^{T} \\
r_{4}^{F} \theta_{1}^{T} & r_{5}^{F} \theta_{1}^{T} & r_{6}^{F} \theta_{1}^{T}
\end{array}\right)
$$

where:

$$
\Theta=\left[\begin{array}{ccc}
\theta_{1}^{T} & &  \tag{7}\\
& \theta_{1}^{T} & \\
& \vdots & \theta_{1}^{T} \\
\theta_{M}^{T} & & \\
& \theta_{M}^{T} & \\
& & \theta_{M}^{T}
\end{array}\right]_{3 M \times 3 K}
$$

$K$ is the size of the DCT basis. If $K$ is chosen too small, the trajectory is poorly represented, but if it is chosen too large, the system is ill-conditioned and the reconstruction error becomes unlimited, so how to choose a suitable $K$ is very important.

According to the above mentioned, the measurement matrix $W$ is decomposed as:

$$
\begin{equation*}
W=R S=R \Theta \mathrm{~A}=\Lambda \mathrm{A} \tag{8}
\end{equation*}
$$

Factorize $W$ with the Singular Value Decomposition (SVD) method:

$$
\begin{equation*}
W=\hat{\Lambda} \hat{\mathrm{A}} \tag{9}
\end{equation*}
$$

However, the matrices $\hat{\Lambda}$ and $\hat{\mathrm{A}}$ will not be equal to $\Lambda$ and A respectively, because SVD is not unique. Any non-singular orthogonal matrix [17] $Q \in R^{3 K \times 3 K}$ can be inserted between $\hat{\Lambda} \hat{\mathrm{A}}$, and get a new valid decomposition $W=\hat{\Lambda} \hat{\mathrm{A}}=\hat{\Lambda} Q Q^{-1} \hat{\mathrm{~A}}=\Lambda \mathrm{A}$. The matrix $Q$ is called the rectification matrix.

According to reference [6], instead of computing the whole matrix $Q$, only three columns of $Q$ are sufficient to rectify $\hat{\Lambda}$ and $\hat{\mathrm{A}}$. After defining the first, $K+1^{s t}$ and $2 K+1^{\text {st }}$ columns of the rectification matrix $Q$ as $Q_{k}$, we can get the $R$ :

$$
\begin{gather*}
\hat{\Lambda} Q_{k}=\left[\begin{array}{c}
Q_{11} R_{1} \\
\vdots \\
Q_{M 1} R_{M}
\end{array}\right]  \tag{10}\\
\hat{\Lambda}_{2 i-1: 2 i} Q_{k}\left(\hat{\Lambda}_{2 i-1: 2 i} Q_{k}\right)^{T}=\hat{\Lambda}_{2 i-1: 2 i} Q_{k} Q_{k}^{T} \hat{\Lambda}^{T}{ }_{2 i-1: 2 i}=\theta_{i, 1}^{2} I_{2 \times 2}, i=1,2, \ldots, M \tag{11}
\end{gather*}
$$

where $\hat{\Lambda}_{2 i-1: 2 i} \in R^{2 \times 3 K}$ denotes the two rows of matrix $\hat{\Lambda}$ at positions between $2 i-1$ and $2 i$.
Due to the inherent ambiguity of the orthogonal constraint, Xiao et al. [18] found that the above method couldn't obtain a unique solution of the rectification matrix $Q$. However, Akhter et al. [14] showed that the inherent ambiguity did not necessarily lead to a fuzzy shape. Experimental results proved that only using the constraint can also recover the unique structure $S$.

The rectification matrix $Q_{k}$ can be estimated precisely by using the trace minimization constraint of the $Q_{k} Q_{k}^{T}$. Once matrix $Q_{k}$ has been computed, the matrix $R$ can be estimated by using a nonlinear minimization routine.

According to Equation (8), the structure matrix $S$ can be calculated by the pseudo-inverse method. Because $\left(\Lambda^{T} \Lambda\right)^{-1} \Lambda^{T} \Lambda=E, W=\Lambda A$, coefficient matrix $A$ is calculated as follows:

$$
\begin{equation*}
A=\left(\Lambda^{T} \Lambda\right)^{-1} \Lambda^{T} W \tag{12}
\end{equation*}
$$

Then the structure matrix is calculated by the equation $S=\Theta A=\Theta\left(\Lambda^{T} \Lambda\right)^{-1} \Lambda^{T} W$. The $S$ is set as the iterative initial value of the APG algorithm.

## 4. The Optimization of Matrix $S$ Using the APG Algorithm

### 4.1. The Trace-Minimization Problem

The goal of this paper is to solve the structure matrix $S$ through the equation $W=R S$, where the measurement matrix $W$ is known, and the rotation matrix $R$ is calculated. Because $S=\Theta A$, the rank of the matrix should meet the requirement of the low-order linear model:

$$
\begin{equation*}
\operatorname{rank}(S) \leq \min \{\operatorname{rank}(\Theta), \operatorname{rank}(A)\} \leq 3 K \tag{13}
\end{equation*}
$$

The size of DCT basis $K$ is a small constant, so the structure matrix $S$ is a low-rank matrix. Then the low-rank condition is relaxed to a rank-minimization problem [19,20]. Now the structure matrix $S$ will be a solution to the rank minimization problem as follows:

$$
\begin{align*}
& \min \operatorname{rank}(S), s . t \\
& W=R S \tag{14}
\end{align*}
$$

according to Dai et al. [21], because the rank-function itself is not very numerically stable and rank-minimization is an NP-hard problem in general. Relaxing the above rank-minimization to a nuclear-norm minimization form in an effective way [22,23], that is $\min \|S\|_{*}$. In principle, the
nuclear-norm minimization may be solved by a standard SDP solver [15]. In this study, the size of $S$ is $3 M \times N$. However, when the size is large, the SDP technique cannot work well.

### 4.2. The Application of the APG Algorithm

Many efficient convex optimization algorithms could be used to solve the problem. In this paper, an effective iterative algorithm, the APG algorithm [24,25], is proposed to optimize the non-rigid structure estimation from monocular vision. According to this algorithm, a closed form solution of the following Equation (15) can be obtained. In Equation (8), $W$ is a measurement matrix of the signal $S$, which was obtained by using the calculated matrix $R$. The above minimization Equation (14) can be rewritten in Lagrangian form as follows:

$$
\begin{equation*}
\min \frac{1}{2}\left|\mid W-R S\left\|_{F}^{2}+\mu\right\| S \|_{*}\right. \tag{15}
\end{equation*}
$$

where $\mu>0$ is a given parameter, and set $f(S)=\frac{1}{2}| | W-R S\left|\left\|_{F}^{2}, \quad P(S)=\mu| | S \mid\right\|_{*}\right.$, $F(S)=f(S)+P(S)$.

Here, the stopping condition of the APG algorithm is defined as following:

$$
\begin{equation*}
\frac{\left\|\mathrm{S}_{\mathrm{K}+1}-\mathrm{S}_{\mathrm{K}}\right\|_{F}}{L_{f} \max \left\{1,\left\|\mathrm{~S}_{\mathrm{K}}\right\|_{F}\right\}} \leq \text { tol } \tag{16}
\end{equation*}
$$

where $t o l$ is a moderately small positive number, since when $S_{K}$ gets close to an optimal solution $S$, the distance between $S_{K}$ and $S_{K+1}$ should become very small. If tol is too large, the non-rigid structure may not be calculated accurately, if tol is too small, the running time will be too long, so we should choose a suitable tol.
$L_{f}$ is the Lipschitz constant of $\nabla f$ :

$$
\begin{equation*}
\left\|\nabla f\left(S_{1}\right)-\nabla f\left(S_{2}\right)\right\| \leq \mathrm{L}_{f}\left\|S_{1}-S_{2}\right\| \tag{17}
\end{equation*}
$$

where the Lipschitz constant $\mathrm{L}_{f}$ is simply the square of the operator norm of the linear map: $S \rightarrow W$.
Instead of directly minimizing $F(S)$, the APG method minimizes a sequence of separable quadratic approximation to $F(S)$, denoted as $Q\left(S, S_{2}\right)$, formed at specially chosen points $S_{2}$ :

$$
\begin{equation*}
Q\left(S, S_{2}\right)=f\left(S_{2}\right)+\left\langle\nabla f\left(S_{2}\right), S-S_{2}\right\rangle+\frac{\tau}{2}\left\|S-S_{2}\right\|_{F}^{2}+P(S) \tag{18}
\end{equation*}
$$

Because:

$$
\begin{equation*}
\underset{X}{\arg \min _{X}} Q\left(S, S_{2}\right)=\arg \min _{X} \frac{L_{f}}{2}\left\|S-S_{2}+\frac{1}{L_{f}} \nabla f\left(S_{2}\right)\right\|_{F}^{2}+P(S) \tag{19}
\end{equation*}
$$

and then the iterative formula is shown as follows:

$$
\left\{\begin{array}{l}
Y_{k}=S_{k}+\frac{t_{k-1}-1}{t_{k}}\left(S_{k}-S_{k-1}\right)  \tag{20}\\
S_{k+1}=D \frac{\mu}{L_{f}}\left(Y_{k}-\frac{1}{L_{f}} \nabla f\left(S_{k}\right)\right)
\end{array}\right.
$$

The detailed steps of the APG algorithm are summarized in Algorithm 1.
Algorithm 1. The steps of the APG algorithm.
Step 1. Initialization: Given $\mu>0, S_{1}=S_{0}=S, t_{1}=t_{0} \in[1,+\infty), K=1,2,3 \ldots$
Step 2. While not converged do
Step 3. $Y_{K}=S_{K}+\frac{t_{K-1}-1}{t_{k}}\left(S_{K}-S_{K-1}\right)$
Step 4. $G_{K}=Y_{K}-\frac{1}{L_{f}} R^{*}\left(R S_{K}-W\right)$,
Step 5. $S_{L_{f}}\left(G_{K}\right)=U D \operatorname{iag}\left(\left(\sigma-\mu / L_{f}\right)_{+}\right) V^{T}$
Step 6. $S_{K+1}=S_{L_{f}}\left(G_{K}\right), t_{K+1}=\frac{1+\sqrt{1+4 t_{K}^{2}}}{2}$
Step 7. End while.
Step 8. Output $S_{K+1}$, namely the reconstructed structure matrix $S$.
In the Algorithm 1, factorize $G$ with SVD method, $G=U \Sigma V^{T}, \Sigma=\operatorname{Diag}(\sigma)$.

## 5. Experiment Results

### 5.1. The Yoga Sequence Experiment

The experimental dataset consists of a 307 -frame sequence of a human practicing yoga, which comes from http://cvlab.lums.edu.pk/non-rigid structure estimation. The database is observed by a perspective camera orbiting the subject on a horizontal plane at a speed of $5^{\circ}$ per frame. The reconstruction performances of Akhter et al.'s PTA approach and the APG approach are presented in the following Figures 1 and 2, respectively. In the figures, the blue dots are the ground truth 3 D points, and the red circles show the recovered points.

As shown in Figures 1 and 2, in general, the APG method reconstruction is better than that of PTA algorithm. Especially, the reconstruction precision of the APG method is increased significantly when $K=9$.

From Figure 3, it can be found that APG method can improve the reconstruction quality with less reconstruction structure errors than the PTA method with different values of $K$.

From Figure 4, it can be found that APG method can improve the reconstruction quality with less reconstruction rotation errors than the PTA method with different values of $K$. From the above figures, the proposed method performs effectively, and the reconstruction accuracy of non-rigid structure estimation from monocular vision is improved effectively. The reconstruction results on the real yoga
sequence images indicate that not only the structure reconstruction, but also the rotation reconstruction is obviously improved by using the APG algorithm.





(a)

(c)



4



(b)

(d)

Figure 1. Reconstruction of the yoga sequence using the PTA method with four different $K$ values, (a) $K=4$; (b) $K=7$; (c) $K=9$; and (d) $K=11$.


(a)




(c)






(b)

(d)

Figure 2. Reconstruction of the yoga sequence using the APG method with four different $K$ values, (a) $K=4$; (b) $K=7$; (c) $K=9$; and (d) $K=11$.


Figure 3. The structure error of different values of $K$ by using PTA and APG algorithms.


Figure 4. The rotation error of different values of $K$ by using PTA and APG algorithms.

### 5.2. The Pickup Sequence Experiment

In addition, another experimental dataset is proposed to test the proposed APG method for non-rigid structure estimation. The experimental dataset consists of a 357 -frames human pickup sequence, which come from http://cvlab.lums.edu.pk/non-rigid structure estimation. The database is observed by a perspective camera orbiting the subject on a horizontal plane at a speed of $5^{\circ}$ per frame. In this paper, the comparison of non-rigid structure estimation from monocular vision between the proposed APG algorithm and PTA algorithm is given, which are presented in terms of the reconstruction result figures and reconstruction error curves with different $K$ values. In the following figures, the blue dots are the ground truth 3D points, and the red circles show the recovered points.

From Figures 5 and 6, in general, the APG method reconstruction is better than the PTA algorithm one. When $K=3$, the reconstruction precision of the APG method is increased significantly.


Figure 5. PTA reconstruction of the pickup sequence with four different $K$ values, (a) $K=3$; (b) $K=7$; (c) $K=10$; and (d) $K=12$.

(a)

(c)








(b)

(d)

Figure 6. APG method reconstruction of the pickup sequence with four different $K$ values, (a) $K=3$; (b) $K=7$; (c) $K=10$; and (d) $K=12$.

From Figure 7, it can be found that APG method can improve the reconstruction quality with less reconstruction structure errors than the PTA method with different values of $K$.


Figure 7. The structure error of different values of $K$ using the PTA and APG algorithms.
From Figure 8, it can be found that APG method can improve the reconstruction quality with less reconstruction rotation errors than the PTA method with different values of $K$. From Figures 7 and 8, we can see the proposed APG method runs reliably, and the accuracy of non-rigid structure estimation is effectively improved. The reconstruction results on the real pickup image sequences indicate that the APG method outperforms the PTA method in terms of reconstruction accuracy of non-rigid structure estimation.


Figure 8. The rotation error of different values of $K$ using the PTA and APG algorithms.

### 5.3. The Comparison of APG Method and Block Matrix Method

The comparisons of reconstruction results between the APG algorithm and Block Matrix Method (the method of Dai et al. [20]) on the shark sequence are presented in the following figure. The shark sequence contains 240 frames and 91 features.


Figure 9. 3D reconstruction results on the shark sequence. (a) Reconstruction results using the APG method, where the mean structure error is 0.204 ; (b) Reconstruction results using the Block Matrix Method, where the mean structure error is 0.242 .

The comparison of reconstruction errors between the APG algorithm and Block Matrix Method is provided in Table 1, where the number in brackets is the best value of $K$, which is chosen by exhaustively trying out different numeric values between 2 and 13. In our paper, the best value of $K$ is confirmed by the tracked positions of a sequence of non-rigid shapes by using the rank analysis method [26]. As can be seen in Figures 9 and 10 and Table 1, the proposed APG method outperforms the Block Matrix Method when the experiment involves the shark and drink sequences. When the experiment is about other sequences, the APG method is not good as the Block Matrix Method, but the difference is not obvious.

(b)

Figure 10. 3D reconstruction results on the drink sequence. (a) Reconstruction results using the APG method, where the mean structure error is 0.017; (b) Reconstruction results using the Block Matrix Method, where the mean structure error is 0.019 .

All the structure error mentioned in our paper refers to the mean error of each frame, we call it mean structure error $\left(e_{S}(t)\right)$. The computational formula is as follows:

$$
\begin{gather*}
e_{S}(t)=\frac{1}{\sigma N} \sum_{j=1}^{N} e_{t j}, \quad \sigma=\frac{1}{3 T} \sum_{t=1}^{T}\left(\sigma_{t x}+\sigma_{t y}+\sigma_{t z}\right)  \tag{21}\\
e_{t j}=\sqrt{L_{t j x}^{2}+L_{t j y}^{2}+L_{t j z}^{2}}  \tag{22}\\
L_{t j x}=\left|S_{r}(3 t-2, j)-S_{0}(3 t-2, j)\right|  \tag{23}\\
L_{t j y}=\left|S_{r}(3 t-1, j)-S_{0}(3 t-1, j)\right|  \tag{24}\\
L_{t j z}=\left|S_{r}(3 t, j)-S_{0}(3 t, j)\right| \tag{25}
\end{gather*}
$$

where $j=1,2, \ldots, N, \sigma_{t x}, \sigma_{t y}$ and $\sigma_{t z}$ are respectively the standard deviation of the point $X, Y, Z$ coordinates of the $t$-th frame corresponding to the 3D structure. $e_{t j}$ represents the reconstruction error of the j -th 3D point of the $t$-th frame. $S_{r}$ is the structure matrix we reconstructed, while $S_{0}$ is the actual structure matrix.

Table 1. The reconstruction error by using APG method and Block Matrix Method.

| Database | Shark | Drink | Yoga | Dance | Pickup |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Block Matrix Method | $0.242(3)$ | $0.019(4)$ | $0.125(9)$ | $0.171(10)$ | $0.138(7)$ |
| APG method | $0.204(2)$ | $0.017(13)$ | $0.135(9)$ | $0.231(5)$ | $0.202(7)$ |

### 5.4. The Comparison of the APG Method and Existing Methods

Moreover, the comparison of reconstruction error between the APG algorithm and the existing EM-PPCA [17], MP [27], the PTA method [6], and CSF [9] are also provided in Figure 11 and Table 2.


Figure 11. The structure error of five different sequences by using different methods.

The abscissa of Figure 11 shows five sequences: drink, yoga, pickup, shark and dance, respectively. Its ordinate indicates the structure errors using five different methods. It can be found that the reconstruction results using the APG method are obviously better than those obtained with the other methods.

Table 2. APG method compared with other methods in terms of reconstruction error.

| Database | Methods |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | EM-PPCA | MP | PTA | CSF | APG |
| Drink | 0.339 | 0.460 | $0.025(3)$ | $0.022(6)$ | $0.017(13)$ |
| Dance | 0.984 | 0.264 | $0.296(5)$ | $0.271(2)$ | $0.231(5)$ |
| Yoga | 0.810 | 0.804 | $0.162(11)$ | $0.147(7)$ | $0.135(9)$ |
| Shark | 0.050 | 0.157 | $0.312(2)$ | $0.254(2)$ | $0.204(2)$ |
| Pickup | 0.582 | 0.433 | $0.237(12)$ | $0.230(6)$ | $0.202(7)$ |

From Figure 11 and Table 2, the reconstruction results on the most real image sequences indicate that reconstruction accuracy is significantly improved by the APG method, which obviously reduces the structure errors. The APG method can further optimize the structures calculated by the PTA method, which is proved to be an effective approach for non-rigid structure estimation from monocular vision. However, for dramatic movement image sequence, the proposed approach cannot improve the accuracy of the non-rigid structure estimation effectively, such as in the shark sequence.

In our algorithm, the PTA method is first used to calculate the structure matrix, and then the APG method is used to optimize the structure matrix. In the total execution time of our algorithm ( $T$ ), the percentage of the execution time of APG algorithm $(t / T)$ is showedn in Table 3.

Table 3. The percentage of the execution time of APG algorithm.

| Time | Drink | Dance | Yoga | Shark | Pickup |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\boldsymbol{T}$ | 30.073 s | 1.9245 s | 2.4637 s | 0.8516 s | 3.7898 s |
| $\boldsymbol{t} \boldsymbol{\boldsymbol { T }}$ | $4.99 \%$ | $19.1 \%$ | $3.12 \%$ | $23.9 \%$ | $2.62 \%$ |

Table 3 shows that our APG method runs quickly, and can converge to the best solution in a little time.

## 6. Conclusions

In this paper, the APG algorithm is proposed to solve the trace-minimization problem of the structure matrix. The initial value of the APG algorithm can be calculated by using the PTA method. The proposed APG method can further improve the structure performance and converge to the optimal solution. Above experimental databases are applied to test the proposed APG method for non-rigid structure estimation from monocular vision. The experimental results show that the proposed method can improve the reconstruction quality of non-rigid structure estimation with less reconstruction error than the PTA method. The APG algorithm can also converge to the best solution quickly, so the time consumption of the proposed method is near that of the PTA method.

However, the APG method is not available for reconstructing dramatic movement sequences, so in the future, the APG method will be improved to make it suitable for dramatic movement. Moreover, the selection of the initial value plays an important role in the reconstruction efficiency of the proposed algorithm, and how to select the best initial value is another future work. The structure matrix $S$ can also act as a solution to the rank minimization problem, so in the future, some optimization algorithms, such as the Singular Value Thresholding (SVT) algorithm, will be considered to further reduce the reconstruction error of non-rigid structure estimation from monocular vision.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (61272311) and is also supported in part by Natural Science Foundation of Zhejiang Province (LZ15F020004 and LY14F010022). This work is also supported by Science Technology Department of Zhejiang Province (2015C31075) and 521 project of Zhejiang Sci-Tech University.

## Author Contributions

Yaming Wang was responsible for the research management and in charge of revising this manuscript. Lingling Tong was in charge of date analysis and the preparation of this manuscript. And she was in charge of planning and performing experiments. Mingfeng Jiang provided valuable advice about the revised manuscript. Junbao Zheng and Lingling Tong were involved in discussions and the experimental analysis.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Tomasi, C.; Kanade, T. Shape and motion from image streams under orthography: A factorization method. Int. J. Comput. Vis. 1992, 9, 137-154.
2. Bregler, C.; Hertzmann, A.; Biermann, H. Recovering Non-Rigid 3D Shape from Image Streams. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, USA, 13-18 June 2000; pp. 690-696.
3. Torresani, L.; Hertzmann, A.; Bregler, C. Learning Non-Rigid 3D Shape from 2D Motion. In Proceedings of the 17th Annual Conference on Neural Information Processing Systems, Vancouver, Canada, 8-13 December 2003.
4. Brand, M. A Direct Method for 3D Factorization of Nonrigid Motion Observed in 2D. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Cabridge, MA, USA, 20-25 June 2005; pp. 122-128.
5. Xiao, J.; Chai, J.; Kanade, T. A closed form solution to non-rigid shape and motion recovery. Int. J. Comput. Vis. 2006, 67, 233-246.
6. Akhter, I.; Sheikh, Y.; Khan, S.; Kanade, T. Trajectory space: A dual representation for nonrigid structure from motion. IEEE Trans. Pattern Anal. Mach. Intell. 2010, 33, 1442-1456.
7. Akhter, I.; Sheikh, Y.; Khan, S.; Kanade, T. Nonrigid structure from motion in trajectory space. IEEE Trans. Pattern Anal. Mach. Intell. 2011, 7, 1442-1456.
8. Zhu, Y.; Cox, M.; Lucey, S. 3D Motion Reconstruction for Real-World Camera Motion. In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20-25 June 2011; pp. 1-8.
9. Gotardo, P.F.U.; Martinez, A. M. Non-Rigid Structure from Motion with Complementary Rank-3 Spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition , Colorado Springs, CO, USA, 20-25 June 2011; pp. 3065-3072.
10. Rehan, A.; Zaheer, A.; AKhter, I.; Saeed, A. NRSFM Using Local Rigidity. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), Steamboat Springs, CO, USA, 24-26 March 2014; pp. 69-74.
11. Minsik, L.; Jungchan, C.; Chong-Ho, C.; Songhwai, O. Procrustean Normal Distribution for Non-Rigid Structure from Motion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23-28 June 2013; pp. 1280-1287.
12. Minsik, L.; Chong-Ho, C.; Songhwai, O. A Procrustean Markov Process for Non-Rigid Structure Recovery. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23-28 June 2014; pp. 1550-1557.
13. Antonio, A.; Lourdes, A.; Begoña, C.; Montiel, J.M.M. Good Vibrations: A Modal Analysis Approach for Sequential Non-Rigid Structure from Motion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23-28 June 2014; pp. 1558-1565.
14. Akhter, I.; Sheikh, Y.; Khan, S. In Defense of Orthonormality Constraints for Nonrigid Structure from Motion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20-25 June 2009; pp. 1534-1541.
15. Wen, Z.W.; Goldfarb, D.; Yin, W. Alternating direction augmented Lagrangian methods for semi definite programming. Math. Program. Comput. 2010, 2, 203-230.
16. Wang, Y.M.; Cheng, J.M.; Zheng, J.B.; Xiong, Y.L. Analysis of wavelet basis selection in optimal trajectory space finding for 3D non-rigid structure from motion. Int. J. Wavelets Multiresolut. Inf. Process. 2014, 12, 1-14.
17. Torresani, L.; Hertzmann, A.; Bregler, C. Nonrigid structure-from motion: Estimating shape and motion with hierarchical priors. Pattern Anal. Mach. Intell. 2008, 30, 878-892.
18. Xiao, J.; Chai, J.; Kanade, T. A Closed-Form Solution to Non-rigid Shape and Motion Recovery; Springer Berlin Heidelberg: Berlin, Germany, 2004.
19. Wright, J.; Ganesh, A.; Rao, S.; Ma, Y. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. Adv. Neural Inf. Process. Syst. 2009, 1, 2080-2088.
20. Gandy, S.; Yamada, I. Alternating Minimization Techniques for the Efficient Recovery of a Sparsely Corrupted Low-Rank Matrix. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Dallas, TX, USA, 14-19 March 2010; pp. 3638-3641.
21. Dai, Y.C.; Li, H.D.; He, M.Y. A Simple Prior-Free Method for Non-Rigid Structure-fromMotionfactorization. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16-21 June 2012; pp. 2018-2025
22. Do, T.T.; Chen, Y.; Nguyen, N.; Lu, G. A Fast and Efficient Heuristic Nuclear-Norm Algorithm for Affine Rank Minimization. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Taipei, China, 19-24 April 2009; pp. 3393-3396.
23. Recht, B.; Fazel, M.; Parrilo, P. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev. 2010, 52, 471-501.
24. Tremoulheac, B.; Atkinson, D.; Arridge, S.R. Fast Dynamic MRI via Nuclear Norm Minimization and Accelerated Proximal Gradient. In Proceedings of the IEEE 10th International Symposium on Biomedical Imaging: From Nano to Macro, San Francisco, CA, USA, 7-11 April 2013; pp. 35-38.
25. Lu, Y.; Zhang, L.W. The augmented Lagrangian method based on the APG strategy for an inverse damped gyroscopic eigenvalue problem. Comput. Optim. Appl. 2015, 1, 1-36.
26. Wang, Y.M.; Zheng, J.B.; Wang, Y.P.; Shi, X.Z. Estimation of Deformation Degree with Uncertainty and Missing data. In Proceedings of the International Conference on Computational Intelligence and Software Engineering, Wuhan, China, 11-13 December 2009.
27. Paladini, M.; Delbue, A.; Stosic, M. Factorization for Non-Rigid and Articulated Structure Using Metric Projections. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20-25 June, 2009; pp. 2898-2905.
© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).
