Recovery of Non-Rigid Structure Based on Sparse Coding

Yaming Wang[1], Xiaomeng Yan[1], Junbao Zheng[1] & Mingfeng Jiang[1]

[1] School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, 310018, China

Keywords: 3D Non-rigid recovery, sparse coding, trajectory space.

Abstract.3D non-rigid recovery is a problem of recovering the real shape of 3D deformable object by using some kinds of algorithms. This paper proposes a novel approach to recovery of non-rigid structure using sparse coding in trajectory space. In order to overcome the difficulty of defining the size of trajectory bases, the sparse coding method is used to construct an over-complete atom dictionary replacing the truncated trajectory bases. Experimental results show that the proposed method does better in solving the 3D non-rigid reconstruction than traditional trajectory basis method and it is also very easy to implement. This paper presents the new thought preliminarily and introduces a simple but effective solution for recovering the structure of the non-rigid objects.

Introduction

3D non-rigid reconstruction is a classical computer vision problem and it is very important when dealing with 3D object simulation problem. Commonly, one prevalent solution for solving this problem is considering the deformable object as a linear combination of some shape bases. And this method worked well in the motion scenes which contain a series of simple action. However, it has shortcomings that it would be hard to recover the structure of a complex motion sequence. For this reason, Akhter et al. [1] introduced a commonly method that reconstructs the 3D object in trajectory space instead of shape space. At the same time, lots of experiments have shown that this approach did well with long complex motion sequences because these measurement trajectory curves were not dependent on each other in trajectory space.

The efficiency of the trajectory basis method relies on two factors: the type of the trajectory basis and the number of the bases. It has been proved that the discrete cosine transform (DCT) basis was more suitable to be defined as a general basis [2]. However, we should know that, though DCT basis has been proved to be better than the others on the whole, it still could not be suitable for every motion sequence. On the other hand, a small size of trajectory basis number K may lead to a big ignorance of much important information of the motion sequences, while a big K may lead to a large number of unknown factors and the system of equations would be ill-posed.

This paper introduces to apply sparse coding algorithms to the non-rigid structure recovery problem. In this process, a set of over-complete bases called atom dictionary is predefined to estimate the deformable object with sparse coefficients. An advantage of the sparse approach is that it is not restricted to only one trajectory basis function and may use two or more incoherent basis functions. It is very useful to recovering the trajectory curves which is complex. What's more, since the goal of sparse approximation is to represent trajectory sequences as a sparse combination of all atoms, it's no need to predefine the number of trajectory bases size.

The paper is organized as follows. In Section 2, a novel method of sparse approximation is proposed, and some definitions and assumptions on this topic are given. Next the Algorithm Solutions are presented in Section 3. And Section 4 shows the experiment results obtained by using the proposed method. Finally, a conclusion of this paper is provided in section 5.

Sparse Approximation Method

The measurement projective trajectories are contained in a 2F×P matrix *W* as follows:

$$W = \begin{pmatrix} x_{11} & x_{12} & & x_{1P} \\ y_{11} & y_{12} & \cdots & y_{1P} \\ & \vdots & \ddots & \vdots \\ x_{F1} & x_{F2} & & x_{FP} \\ y_{F1} & y_{F2} & \cdots & y_{FP} \end{pmatrix} = \left(\frac{X}{Y} \right) \in \square^{2F \times P} \; . \tag{1}$$

W is the measurement matrix generated from the 2D projective coordinates of the feature points got from the deformable object, F indicates the number of the motion frames and P is the number of the feature points. In trajectory space, W can be represented by SVD (Singular Value Decomposition) as

$$W = MS \; ; \tag{2}$$

Where M denotes the camera motion matrix and S denotes the 3D object shape matrix.

Next the matrix S can be represented as follow equation in trajectory space:

$$S = \Theta A \tag{3}$$

The symbol Θ denotes the atom dictionary constructed by some trajectory bases functions and A denotes the corresponding coefficients.

This paper wishes to represent all the trajectory curves of the feature points as a linear combination of a certain waveform. But there's a problem that if the number of trajectory bases was too big, the equations system would be a NP hard problem and if the number of trajectory bases was too small, the recovery would be not accuracy. So a new method is introduced which uses sparse approximation method to estimate the trajectory curves instead of traditional trajectory bases method. An atom dictionary is a concatenation of some orthonormal bases. So the expression is not unique, and many combination methods are proposed. But the goals of these methods are all to get a highly sparse decomposition matrix which contains very few non-zero terms. This puts forward an optimization problem:

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0$$
$$s.t. \; W = \Pi \alpha \; . \tag{4}$$

Where $\Pi = M\Theta$, α denotes a coefficients matrix, $\hat{\alpha}$ denotes an estimated coefficients matrix and the symbol $\|\cdot\|_0$ indicates the L0 norm which constraints the non-sparse items. It has been turned out to be right that if the isometric constant Δ satisfied follow equations:

$$(1 - \Delta)\|\alpha\|_2^2 \leq \|\Pi\alpha\|_2^2 \leq (1 + \Delta)\|\alpha\|_2^2 \; ;$$
$$\Delta < 1. \tag{5}$$

Then a unique K-sparse solution of the vector α will be found. But a drawback to Equation 4 is that it may be NP-hard. Fortunately, one can get a convex relaxation on Equation 4 with a L1 norm as follows:

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1$$
$$s.t. \; W = \Pi \alpha \; . \tag{6}$$

The L1 norm is a very common solution to solving sparse estimation problems and it has been turned out to be effective. A variety of different solution packages can be found to solve the L1 norm problem. This paper will apply L1-regularized least squares, namely the feature-sign search algorithm, to achieve the sparse representation of the object motion trajectory. This algorithm is mentioned in Reference [3].

Algorithm Solutions

In this paper, a semi-definite programming (SDP) of small and fixed size mentioned in Reference [4] is applied to solve camera motion matrix M.

$\min trace(\mathrm{Q}_k)$, such that,

$$Q_k = G_k G_k^T,$$
$$Q_k \succeq 0,$$
$$\begin{bmatrix} (\tilde{M}_i \otimes \tilde{M}_i)(1,:) - \tilde{M}_i \otimes \tilde{M}_i)(4,:) \\ \tilde{M}_i \otimes \tilde{M}_i)(2,:) \end{bmatrix} vec(Q_k) = 0$$

(7)

Where Q demotes a Gram matrix, G demotes a nonsingular matrix and M demotes the camera matrix. Then coefficients matrix A can be solved by a sparse approximation method. Having known the camera motion matrix M, Π can be obtained from the equation $\Pi = M\Theta$. An alternate strategy can be described as follows:

$$\check{\alpha}_i = \arg\min_{\alpha_i} \|\alpha_i\|_1$$

(8)

$$s.t. \ w_i = \Pi\alpha_i.$$

To simplify the expression, it can be represented as the following optimization problem:

$$\min_x f(x) \equiv \|w_i - \Pi\alpha_i\|^2 + \gamma\|\alpha_i\|_1$$

(9)

Where γ is the penalty coefficient and it must be a constant.

An iterative feature-sign search algorithm, which can solve for α in the Fourier domain, can be used to solve this objective efficiently. The details of the algorithm have been introduced in Reference [5, 6]. The coefficients matrix A can be generated from the vector α_i. At last, the shape matrix S will be obtained via the equation $S = \Theta A$.

Experimental Results

The training dataset used in this paper are from CMU Motion Capture dataset, which covering a variety of human actions. The sequence of "Yoga" (43/307) is used in this paper, where (F/P) denotes the number of points (P) and frames (F).

A union atom dictionary of DCT and Dirac basis functions is used in this experiment. Having recovered the 3D coordinates of the deformable object, the corresponding coefficients of the atom dictionary are obtained naturally. And the experimental result is showed in Fig.1.
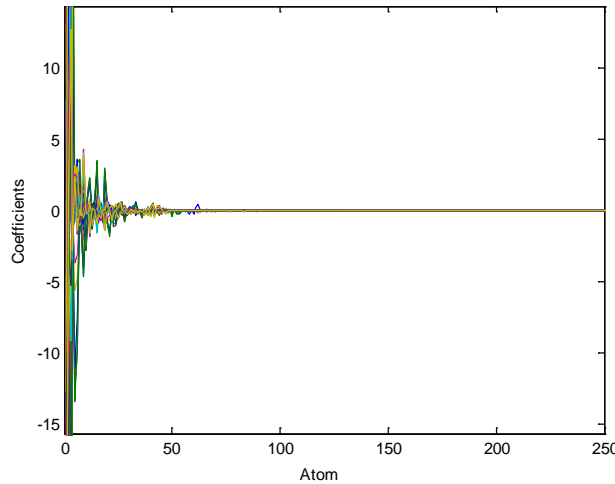

Fig. 1. Sparse coefficients estimations on the "Yoga" sequence (A part, better viewed in color).

From the above graph, one can easily find that most of the coefficients are zero. It proves that the assumption is suitable and the coefficients matrix is a sparse matrix. At the same time, the result also verifies the feasibility of the sparse approximation method.

At last, this paper gives out the shape reconstruction results of the "Yoga" sequences using the DCT method and the sparse approximation method respectively. The truncated size of DCT basis in the "Yoga" experiment is 12, which has been proved to be the most suitable size using in this sequence and the penalty coefficient of the sparse method is 0.1. The experimental results are seen in Fig.2.
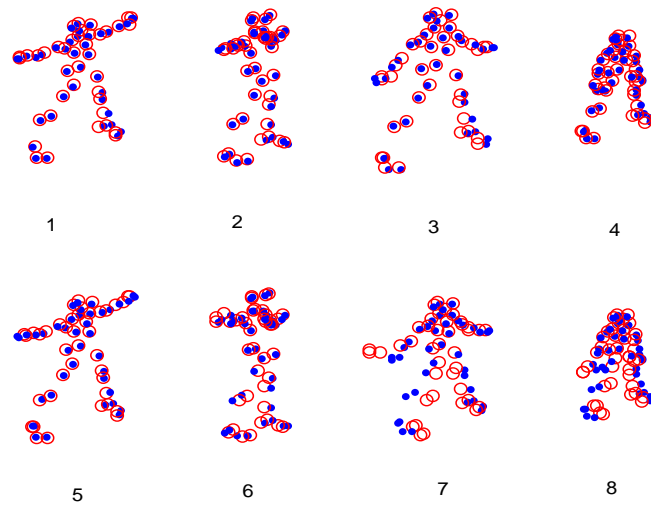
Fig.2. 3D reconstruction results on the "Yoga" sequence with two methods respectively. The red dots denote the ground truth coordinates and the blue circles denote the estimation coordinates. The top row is the reconstruction result with sparse approximation and the bottom row is with the DCT trajectory basis method [7]. Both of the results of these two methods are showed with the 1th, 60th, 120th and 180th frames. (Better viewed in color)

The experimental results also show the advantage of the sparse approximation method. The result obtained by using the approach mentioned in this paper is obviously better than the traditional DCT trajectory basis methods seen from the Fig.2. An important improvement of this method is that it is, unlike the traditional trajectory basis method, useful to all kinds of bases.

Conclusions

This paper introduces a novel sparse coding method to resolve the non-rigid object recovery problem. It is very easy to implement and is guaranteed to get an optimal solution. One need not to consider the size of truncated DCT bases via using the proposed method. In this paper, only a simple atom dictionary is applied into the recovery of 3D deformable object. It is expected that, camera rotation matrix estimation error could be obtained more accurately and the atom dictionary could represent the shape more sparse.

In this paper, the present experiments are all in a situation of orthographic camera models. And thanks to recent progress in signal sparse coding, the proposed solution can be easily applied into solving the 3D object recovery problem. This paper just gives out one thought to solve the size of the trajectory bases and the idea can be expanded to some new solutions. Some other ideas can also to be used in this problem.

Acknowledgements

References

[1] I. Akhter, Y. Sheikh, S. Khan & T. Kanade. Nonrigid structure from motion in trajectory space. *Neural Information Processing Systems.* pp. 41–48, 2008.

[2] Akhter I., Sheikh Y., Khan S. & Kanade T. Trajectory Space: A Dual Representation for Nonrigid Structure from Motion. *Pattern Analysis and Machine Intelligence*. 33, pp.1442-1456, 2011.

[3] Yingying Zhu & Simon Lucey. Convolutional Sparse Coding for Trajectory Reconstruction. *Pattern Analysis and Machine Intelligence*. 6, pp.1-13, 2013.

[4] Yuchao Dai, Hongdong Li & Mingyi He. A simple Prior-free Method for Non-rigid Structure from Motion Factorization. *International Journal of Computer Vision*. 107, pp.101-122, 2014.

[5] Honglak Lee, Alexis Battle & Rajat Raina,Andrew Y. Ng. Efficient sparse coding algorithms. *Neural Information Processing Systems*, pp.801-808, 2007.

[6] M. Chen, G. AlRegib, and B. Juang. Trajectory triangulation: 3d motion reconstruction with L1 optimization. *ICASSP*. pp.4020– 4023. 2011.

[7] J. Valmadre and S. Lucey. General trajectory prior for Nonrigid reconstruction. *Computer Vision and Pattern Recognition*, pp.1394-1401, 2012.

[8] Wang Yaming, Cheng jianmin & Zheng junbao. Analysis of wavelet basis selection in optimal trajectory space finding for 3D non-rigid structure from motion. *International Journal of Wavelets*. 2014.