

Multi-body NRSfM

Suryansh Kumar, Yuchao Dai, Hongdong Li

July 26, 2017



Australian
National
University

- 1 Introduction
- 2 Spatial-Temporal Representation
- 3 Joint Optimization Formulation
- 4 Experiments and Results

Introduction

Why Multi-body NRSfM Representation?

- Real-world scene consist of multiple deforming objects. For example: pedestrians, soccer match, human interaction and etc.

Goal:

- To segment and reconstruct multiple deforming objects in a scene.

Baseline strategy:

- Two-stage approach:
 - motion segmentation followed by non-rigid reconstruction
 - non-rigid reconstruction followed by motion segmentation.

Why unified approach?

- To better exploit the inherent structure of the problem
 - ⇒ Motion segmentation benefits reconstruction
 - ⇒ Reconstruction benefits motion segmentation
- Both tasks can be solved efficiently within a single optimization.
- Computationally and numerically efficient.

Spatial-Temporal Representation

To exploit the intrinsic structure both spatially and temporally, we propose the spatial-temporal representation for complex non-rigid reconstruction.

- Spatial Clustering \Rightarrow Provides motion segmentation cues
- Temporal Clustering \Rightarrow Benefits 3D reconstruction
- Spatial Clustering exploits **Trajectory space**.
- Temporal Clustering exploits **Shape space**.

Trajectory Space

Classical NRSfM Representation

$$\mathbf{W} = \mathbf{R}\mathbf{S}, \text{ where } \mathbf{R} \in \mathbb{R}^{2F \times 3F}, \mathbf{S} \in \mathbb{R}^{3F \times P} \quad (1)$$

$\mathbf{W} \in \mathbb{R}^{2F \times P} \Rightarrow$ Measurement matrix.

$\mathbf{S} \Rightarrow$ Shape matrix.

$\mathbf{R} \Rightarrow$ Rotation matrix (Orthographic Camera Model).

Trajectory Space

Representation of multiple non-rigid deformation in the trajectory space.

$$S = SC_1, \text{diag}(C_1) = 0, 1^T C_1 = 1^T. \quad (2)$$

$$S \in \mathbb{R}^{3F \times P}, C_1 \in \mathbb{R}^{P \times P}.$$

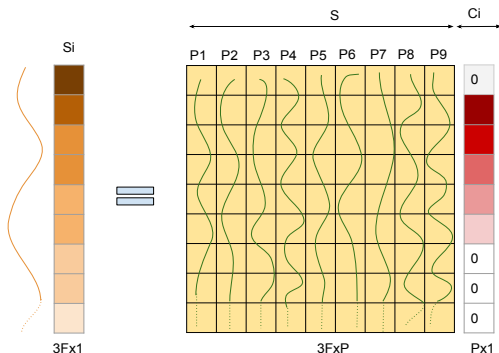


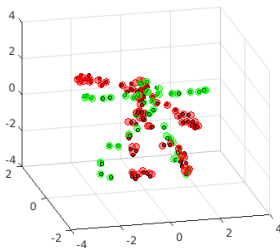
Figure: Illustration of trajectory space

Shape Space

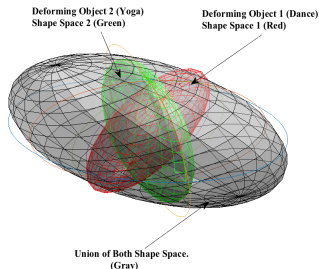
Representation of multiple non-rigid deformation in the shape space.

$$\begin{aligned} S^\sharp &= S^\sharp C_2, \text{diag}(C_2) = 0, 1^T C_2 = 1^T. \\ S^\sharp &\in \mathbb{R}^{3P \times F}, C_2 \in \mathbb{R}^{F \times F}. \end{aligned} \quad (3)$$

⇒ Intuition [Cluster distinct activity (Ex: Dance, Yoga)]



(a)



(b)

Visual illustration

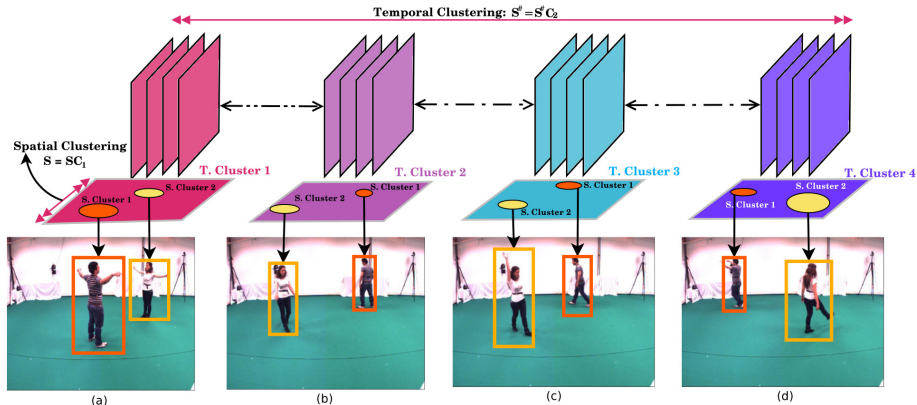


Figure: Intuition of spatial-temporal clustering.

Joint Optimization Formulation

- Objective from the **trajectory space**

$$\begin{aligned} & \underset{C_1}{\text{minimize}} \quad \lambda_1 \|C_1\|_1 + \frac{(1 - \lambda_1)}{2} \|C_1\|_F^2 \\ & \text{subject to:} \\ & S = SC_1, \text{diag}(C_1) = 0, 1^T C_1 = 1^T, \lambda_1 \in [0, 1]. \end{aligned} \tag{4}$$

- Objective from the **shape space**

$$\begin{aligned} & \underset{C_2}{\text{minimize}} \quad \lambda_3 \|C_2\|_1 + \frac{(1 - \lambda_3)}{2} \|C_2\|_F^2 \\ & \text{subject to:} \\ & S^\# = S^\# C_2, \text{diag}(C_2) = 0, 1^T C_2 = 1^T, \lambda_3 \in [0, 1]. \end{aligned} \tag{5}$$

Joint Optimization Formulation

- Overall Objective \Rightarrow solved using ADMM

$$\begin{aligned} \underset{S, C_1, C_2}{\text{minimize}} \quad & \frac{1}{2} \|W - RS\|_F^2 + \lambda_1 \|C_1\|_1 + \frac{1 - \lambda_1}{2} \|C_1\|_F^2 + \lambda_2 \|S^\# \|_* + \\ & \lambda_3 \|C_2\|_1 + \frac{1 - \lambda_3}{2} \|C_2\|_F^2. \end{aligned}$$

subject to:

$$S = SC_1, S^\# = S^\# C_2,$$

$$1^T C_1 = 1^T, 1^T C_2 = 1^T,$$

$$\text{diag}(C_1) = 0, \text{diag}(C_2) = 0,$$

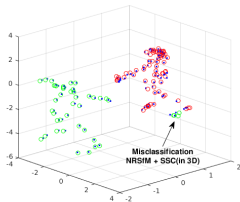
$$\lambda_1, \lambda_3 \in [0, 1].$$

(6)

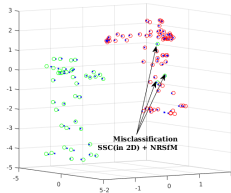
where $S^\# \in \mathbb{R}^{3P \times F}$, $C_1 \in \mathbb{R}^{P \times P}$, and $C_2 \in \mathbb{R}^{F \times F}$ and $\lambda_1, \lambda_2, \lambda_3$ are the trade-off parameters.

Experiments and Results

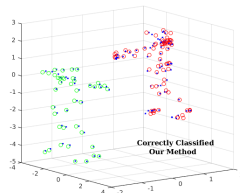
- Advantage over two stage approach



(a) $\text{NRSfM} \Rightarrow \text{SSC}$



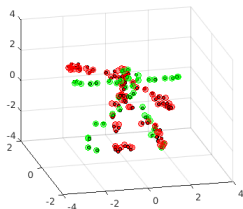
(b) $\text{SSC} \Rightarrow \text{NRSfM}$



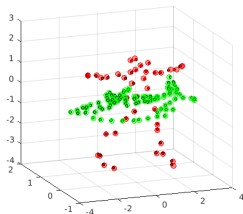
(c) Our approach

Qualitative results on synthetic sequence

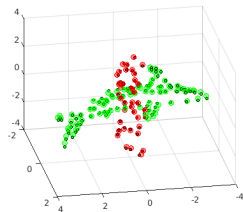
- Two deforming objects are intersecting each other.



(d) Dance-Yoga



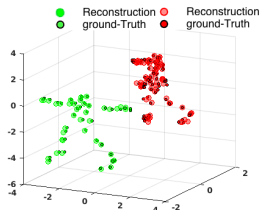
(e) Shark-Stretch



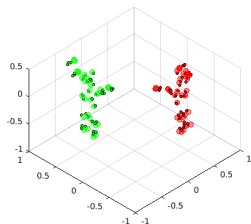
(f) Shark-Yoga

Qualitative results(Cont.)

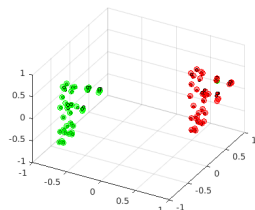
- Two deforming objects are well separated in space.



(g) Dance-Yoga



(h) UMPM p3_ball_1



(i) UMPM p4_meet_12

UMPM is a dataset composed of real-image tracks.

Quantitative Results on benchmark real-dataset

Datasets	BMM	PND	Zhu et al.	Kumar et al.	Ours
p2_free_2	0.1973	0.1544	0.1142	0.1992	0.1171
p2_grab_2	0.2018	0.1570	0.0960	0.2080	0.0822
p3_ball_1	0.1356	0.1477	0.0832	0.1348	0.0810
p4_meet_12	0.0802	0.0862	0.0972	0.0821	0.0815
p4_table_12	0.2313	0.1588	0.1322	0.2313	0.0994

Table: Performance comparison on real benchmark UMPM dataset (showing relative 3D reconstruction error).

Quantitative Results on benchmark real-dataset

Datasets	BMM	PND	Zhu et al.	Kumar et al.	Ours
Face Seq. 1	0.078	0.077	0.082	0.075	0.073
Face Seq. 2	0.059	0.062	0.063	0.050	0.052
Face Seq. 3	0.042	0.051	0.057	0.038	0.039
Face Seq. 4	0.049	0.041	0.056	0.044	0.040

Table: Performance comparison on real benchmark dense face dataset of Garg et. al.(showing relative 3D reconstruction error).

Evaluation result on NRSfM challenge dataset for test frame.

- Mean RMS (in mm) for orthogonal category.

Datasets	Articulated	Balloon	Paper	Stretch	Tearing
Our Method	10.15	10.64	15.78	9.96	14.17

Table: Performance on the NRSfM challenge dataset on all provided sequence for *single* test image provided by the challenge organizers.

- Note: We submitted results of two methods. Numerically both methods provide results that are very close to each other.

Performance comparison with other top 3 performing algorithms on NRSfM challenge dataset.

- Mean RMS (in mm) for orthogonal category.

Datasets	Articulated	Balloon	Paper	Stretch	Tearing	Mean
Multibody	45.51	14.55	22.88	18.30	21.98	24.64
CSF2	35.51	19.01	33.95	23.22	18.77	26.09
RIKS	42.11	18.45	32.18	22.88	18.12	26.75
KSTA	36.63	24.88	31.96	24.25	17.59	26.86

Table: Note: These evaluations were done by the organizers of NRSfM challenge at CVPR 2017.

- Kumar. S, Dai. Y, Li. H “Spatio-Temporal Union of Subspaces for Multi-body Non-rigid Structure-from-Motion”, 428 – 443, Elsevier, *Pattern Recognition* 2017.
- Kumar. S, Dai. Y, Li. H “Multi-body Non-rigid Structure from Motion”, 148–156 *3DV* 2016.
- P. F. U. Gotardo and A. M. Martinez, “Non-rigid structure from motion with complementary rank-3 spaces”, CVPR, 2011. (CSF2)
- Hamsici, Onur C and Gotardo, Paulo FU and Martinez, Aleix M., “Learning spatially-smooth mappings in non-rigid structure from motion”, ECCV, 2012 (RIKS).
- P. F. U. Gotardo and A. M. Martinez, “Kernel non-rigid structure from motion”, ICCV, 2011(KSTA).

Thanks