Multi-body Non-rigid Structure-from-Motion Supplementary Material

Suryansh Kumar[1] Yuchao Dai[1] Hongdong Li[1,2]
[1]Research School of Engineering, Australian National University.
[2] Australian Center for Robotic Vision.

Abstract

In the supplementary material, we provide a detailed derivation of each sub-problems in the formulation. Besides this, we also provide insight into convergence curve and effect of noisy track features on the performance of our algorithm.

1. Sub-problem derivation of the involved optimization

$$\begin{aligned}&\underset{\mathtt{C},\mathtt{S},\mathtt{S}^\sharp}{\text{minimize}}\ \frac{1}{2}\|\mathtt{W}-\mathtt{RS}\|_F^2+\lambda_1\|\mathtt{C}\|_1+\lambda_2\|\mathtt{CD}\|_1+\lambda_3\|\mathtt{S}^\sharp\|_*\\&\text{subject to:}\\&\mathtt{S}^\sharp=g(\mathtt{S}),\mathtt{S}=\mathtt{SC},1^T\mathtt{C}=1^T,\text{diag}(\mathtt{C})=0.\end{aligned}\tag{1}$$

To further decouple the constraint, we introduce an auxiliary variable $\mathtt{E}=\mathtt{CD}$. With these operations, the optimization problem Eq.-(1) can be reformulated as:

$$\begin{aligned}&\underset{\mathtt{E},\mathtt{S},\mathtt{S}^\sharp,\mathtt{C}}{\text{minimize}}\ \frac{1}{2}\|\mathtt{W}-\mathtt{RS}\|_F^2+\lambda_1\|\mathtt{E}\|_1+\lambda_2\|\mathtt{S}^\sharp\|_*\\&\text{subject to:}\\&\mathtt{S}^\sharp=g(\mathtt{S}),\mathtt{S}=\mathtt{SC},\mathtt{CD}=\mathtt{E},1^T\mathtt{C}=1^T,\text{diag}(\mathtt{C})=0.\end{aligned}\tag{2}$$

The Augmented Lagrangian formulation for Eq.-(2) is:

$$\begin{aligned}\mathcal{L}(\mathtt{S},\mathtt{S}^\sharp,\mathtt{C},\mathtt{E},\{\mathtt{Y}_i\}_{i=1}^4)=\frac{1}{2}\|\mathtt{W}-\mathtt{RS}\|_F^2+\lambda_1\|\mathtt{E}\|_1+\\\lambda_2\|\mathtt{S}^\sharp\|_*+<\mathtt{Y}_1,\mathtt{S}^\sharp-g(\mathtt{S})>+\frac{\beta}{2}\|\mathtt{S}^\sharp-g(\mathtt{S})\|_F^2+\\<\mathtt{Y}_2,\mathtt{S}-\mathtt{SC}>+\frac{\beta}{2}\|\mathtt{S}-\mathtt{SC}\|_F^2+<\mathtt{Y}_3,\mathtt{CD}-\mathtt{E}>+\\\frac{\beta}{2}\|\mathtt{CD}-\mathtt{E}\|_F^2+<\mathtt{Y}_4,1^T\mathtt{C}-1^T>+\frac{\beta}{2}\|1^T\mathtt{C}-1^T\|_F^2.\end{aligned}$$

1.0.1 The solution of S:

$$\begin{aligned}\mathtt{S}=\arg\min_{\mathtt{S}}\ &\tfrac{1}{2}\|\mathtt{W}-\mathtt{RS}\|_F^2+<\mathtt{Y}_1,\mathtt{S}^\sharp-g(\mathtt{S})>+\\&\tfrac{\beta}{2}\|\mathtt{S}^\sharp-g(\mathtt{S})\|_F^2+<\mathtt{Y}_2,\mathtt{S}-\mathtt{SC}>+\\&\tfrac{\beta}{2}\|\mathtt{S}-\mathtt{SC}\|_F^2.\end{aligned}\tag{3}$$

The sub-problem for S reaches a least squares problem. The closed-form solution of S can be derived as:

$$\begin{aligned}&\tfrac{1}{\beta}(\mathtt{R}^T\mathtt{R}+\beta\mathtt{I})\mathtt{S}+\mathtt{S}(\mathtt{I}-\mathtt{C})(\mathtt{I}-\mathtt{C}^T)=\tfrac{1}{\beta}\mathtt{R}^T\mathtt{W}+\\&(g^{-1}(\mathtt{S}^\sharp)+\tfrac{g^{-1}(\mathtt{Y}_1)}{\beta}-\tfrac{\mathtt{Y}_2}{\beta}(\mathtt{I}-\mathtt{C}^T)),\end{aligned}\tag{4}$$

which is a Sylvester equation.

1.0.2 The solution of $\mathtt{S}^\sharp$:

$$\mathtt{S}^\sharp=\arg\min_{\mathtt{S}^\sharp}\ \lambda_2\|\mathtt{S}^\sharp\|_*+<\mathtt{Y}_1,\mathtt{S}^\sharp-g(\mathtt{S})>+\frac{\beta}{2}\|\mathtt{S}^\sharp-g(\mathtt{S})\|_F^2\tag{5}$$

A close-form solution exists for this sub-problem. Let's define the soft-thresholding operation as $\mathcal{S}_\tau[x]=\text{sign}(x)\max(|x|-\tau,0)$. The optimal solution to Eq.-(5) can be obtained as:

$$\mathtt{S}^\sharp=\mathtt{U}\mathcal{S}_{\lambda_2/\beta}(\Sigma)\mathtt{V},\tag{6}$$

where $[\mathtt{U},\Sigma,\mathtt{V}]=\text{svd}(g(\mathtt{S})-\mathtt{Y}_1/\beta)$.

1.0.3 The solution of E:

$$\mathtt{E}=\arg\min_{\mathtt{E}}\ \lambda_1\|\mathtt{E}\|_1+<\mathtt{Y}_3,\mathtt{CD}-\mathtt{E}>+\frac{\beta}{2}\|\mathtt{CD}-\mathtt{E}\|_F^2,\tag{7}$$

A close-form solution exists for this sub-problem by using element-wise shrinkage.

$$\mathtt{E}=\mathcal{S}_{\lambda_1/\beta}(\mathtt{CD}+\frac{\mathtt{Y}_3}{\beta}).\tag{8}$$

1.0.4 The solution of C:

$$\begin{aligned}\mathtt{C}=\arg\min_{\mathtt{C}}\ &<\mathtt{Y}_2,\mathtt{S}-\mathtt{SC}>+\tfrac{\beta}{2}\|\mathtt{S}-\mathtt{SC}\|_F^2+\\&<\mathtt{Y}_3,\mathtt{CD}-\mathtt{E}>+\tfrac{\beta}{2}\|\mathtt{CD}-\mathtt{E}\|_F^2+<\mathtt{Y}_4,1^T\mathtt{C}-1^T>+\\&\tfrac{\beta}{2}\|1^T\mathtt{C}-1^T\|_F^2.\end{aligned}\tag{9}$$

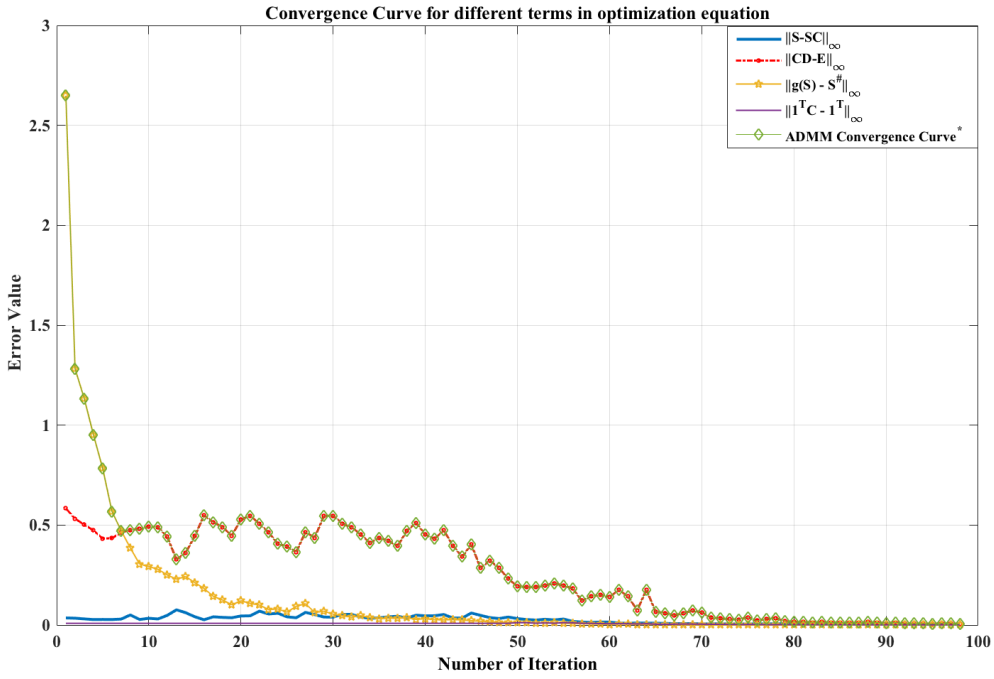

Figure 1: Typical convergence curves of the objective function and the primal residuals $\|\mathtt{S}^\sharp - g(\mathtt{S})\|_\infty, \|\mathtt{S} - \mathtt{SC}\|_\infty, \|\mathtt{CD} - \mathtt{E}\|_\infty$ and $\|1^T\mathtt{C} - 1^T\|_\infty$. The above plot shows the convergence statistics for Dance+Yoga Sequence. *ADMM Convergence Curve = maximum of ($\|\mathtt{S}^\sharp - g(\mathtt{S})\|_\infty, \|\mathtt{S} - \mathtt{SC}\|_\infty, \|\mathtt{CD} - \mathtt{E}\|_\infty$ and $\|1^T\mathtt{C} - 1^T\|_\infty$).

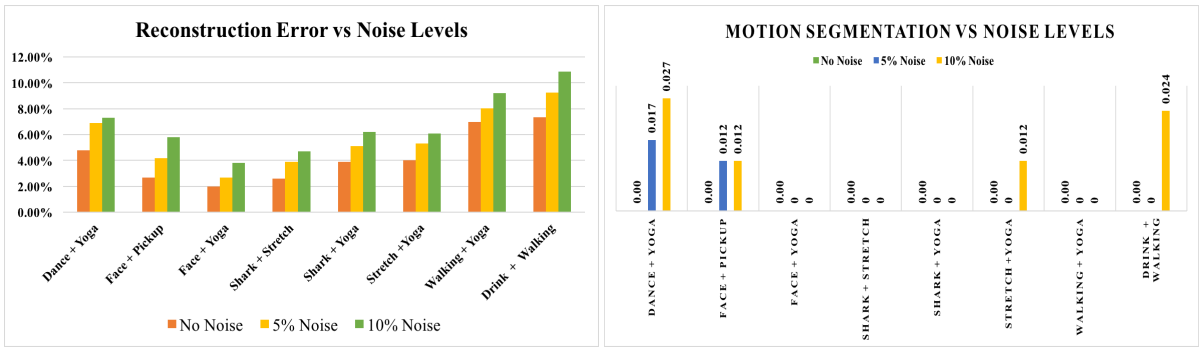

Figure 2: Left: 3D Reconstruction error VS noise levels; Right: non-rigid motion segmentation error VS noise levels.

The closed-form solution of $\mathtt{C}$ is derived as:

$$(\mathtt{S}^T\mathtt{S} + 1\mathtt{1}^T)\mathtt{C} + \mathtt{C}(\mathtt{D}\mathtt{D}^T) = \mathtt{S}^T\mathtt{S} + \mathtt{S}^T\frac{\mathtt{Y}_2}{\beta} + \mathtt{E}\mathtt{D}^T - \mathtt{Y}_3\frac{\mathtt{D}^T}{\beta} + 1\mathtt{1}^T - 1\frac{\mathtt{Y}_4}{\beta}. \tag{10}$$

$$\mathtt{C} = \mathtt{C} - \mathrm{diag}(\mathtt{C}), \tag{11}$$

Finally, the Lagrange multipliers $\{\mathtt{Y}_i\}_{i=1}^4$ and β are updated as:

$$\mathtt{Y}_1 = \mathtt{Y}_1 + \beta(\mathtt{S}^\sharp - g(\mathtt{S})), \mathtt{Y}_2 = \mathtt{Y}_2 + \beta(\mathtt{S} - \mathtt{SC}), \tag{12}$$

$$\mathtt{Y}_3 = \mathtt{Y}_3 + \beta(\mathtt{CD} - \mathtt{E}), \mathtt{Y}_4 = \mathtt{Y}_4 + \beta(1^T\mathtt{C} - 1^T), \tag{13}$$

$$\beta = \min(\beta_m, \rho\beta), \tag{14}$$

1.1. Experiment: Convergence

In this experiment, we would like to study the convergence of our algorithm. Given noise free input, we want to check whether or not our proposed algorithm converge; and if it does converge, whether it converges to the correct

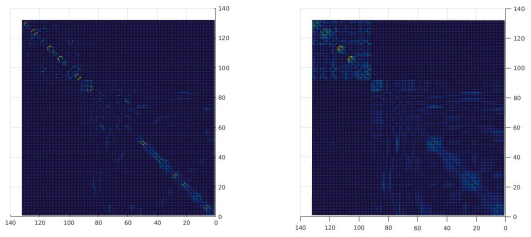

Figure 3: Obtained Affinity Matrix $\mathtt{A} = |\mathtt{C}| + |\mathtt{C}^T|$. a) Affinity matrix from SSC; b) Affinity matrix from our Method. Best Viewed on Screen.

solution. Note that we use the sparse sequences from the CMU MoCap dataset [1] directly without any dimension reduction or projection. Typical convergence curves of the objective function and the primal residuals are illustrated in Fig. 1.

1.2. Experiment: Performance on noisy feature tracks

In the second experiment, we conducted analysis to the performance of our method under different level of noise. In the same manner as above, we generated multi-body non-rigid sequences ("Dance + Yoga", "Face + Pickup", "Face + Yoga", "Shark + Stretch", "Shark + Yoga", "Stretch + Yoga" and "Walking + Yoga"), then zero-mean Gaussian noise with standard deviation σ were added to the feature tracks. For each noisy input, we ran our code for 5 times and recorded the mean 3D reconstruction error and non-rigid motion segmentation error.

In Fig. 2, we illustrated the statistical results of 3D non-rigid reconstruction and non-rigid motion segmentation. From the figures, we conclude that both the 3D reconstruction error and the motion segmentation error increases with the increase of noise level. Our 3D reconstruction based non-rigid motion segmentation achieves smaller motion segmentation error compared with 2D trajectory based motion segmentation methods such as sparse subspace clustering (SSC) [2] and efficient dense subspace clustering (EDSC) [3].

1.3. Affinity Matrix Comparison

In Fig. 3, we compare the affinity matrices from SSC [2] and our method. It is clear that our method outputs an affinity matrix with better structure, which results in better non-rigid motion segmentation performance.

References

[1] I. Akhter, Y. Sheikh, S. Khan, and T. Kanade. Nonrigid structure from motion in trajectory space. In *Advances in Neural Information Processing Systems*, pages 41–48, 2008. 3

[2] E. Elhamifar and R. Vidal. Sparse subspace clustering: Algorithm, theory, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2765–2781, 2013. 3

[3] P. Ji, M. Salzmann, and H. Li. Efficient dense subspace clustering. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 461–468, March 2014. 3