Supplementary Material for the Paper:

Estimation of 3D Category-Specific Object Structure: Symmetry, Manhattan and/or Multiple Images

Yuan Gao Alan Yuille

March 31, 2018

This supplementary material complement our paper by giving more details on six issues:

- In Sect. S1, we give details for the minimization of the energy function *w.r.t.* the camera projection matrix R_n under orthogonality constraints, i.e., update R_n in Eq. (9).
- In Sect. S2, we describe a way to recover a 2×2 matrix B from $BB^\top$ up to a rotation ambiguity.
- In Sect. S3, we show the details for obtaining Eq. (33) from Eqs. (30) - (32).
- In Sect. S4, we show the details for obtaining Eq. (47) from Eqs. (44) and (46).
- In Sect. S5, we give the update of each parameters in **M-step** of *Sym-EM-PPCA* to optimize Eq. (63).
- In Sect. S6, we show our results on the imperfect annotations (*i.e.* the imperfect symmetric pairs).

S1 Update Camera Projection Matrix R_n Under Orthogonality Constraints

The energy function *w.r.t.* R_n is:

$$Q(R_n) = \sum_n ||Y_n - R_n S_n||_2^2 + \sum_n ||Y_n^\dagger - R_n \mathcal{A} S_n||_2^2. \tag{S1}$$

In order to minimize Eq. (S1) *w.r.t.* R_n under the nonlinear orthogonality constraints $R_n R_n^\top = I$, we follow the approach used in [1] which parameterizes R_n as a complete 3×3 rotation matrix Q_n and updates the incremental rotation on Q_n, *i.e.* $Q_n^{new} = e^\xi Q_n$.

The first and second rows of Q_n are the same as R_n, and the third row of Q_n is obtained by the cross product of its first and second rows. The relationship of Q_n and R_n given using a matrix operator $\mathcal{M}$:

$$R_n = \mathcal{M} Q_n, \qquad \mathcal{M} = \begin{bmatrix} 1,0,0 \\ 0,1,0 \end{bmatrix}. \tag{S2}$$

Note that the incremental rotation e^ξ can be approximated by its first order Taylor Series, *i.e.* $e^\xi \approx I + \xi$. Finally, we have:

$$R_n^{new}(\xi) = \mathcal{M}(I + \xi) Q_n. \tag{S3}$$

Setting $\partial \mathcal{Q}/\partial R_n = 0$, replacing R_n by Q_n using Eq. (S31) and vectorizing it, yields:

$$
\begin{aligned}
&R_n = \mathcal{M} e^{\xi} Q_n \approx \mathcal{M}(I+\xi) Q_n \qquad \text{and} \qquad \mathrm{vec}(\xi) = \alpha^{+}\beta,\\
&\alpha = \left(\sum_{p=1}^{P} (S_{n,p} S_{n,p}^{\top} + \mathcal{A} S_{n,p} S_{n,p}^{\top} \mathcal{A}^{\top})^{\top} Q_n^{\top} \right) \otimes \mathcal{M},\\
&\beta = \mathrm{vec}\left(\sum_{p=1}^{P} (Y_{n,p} S_{n,p}^{\top} + Y_{n,p}^{\dagger} S_{n,p}^{\top} \mathcal{A}^{\top}) - Q_n \sum_{p=1}^{P} (S_{n,p} S_{n,p}^{\top} + \mathcal{A} S_{n,p} S_{n,p}^{\top} \mathcal{A}^{\top}) \right),
\end{aligned} \tag{S4}
$$

where the subscript p means the pth keypoint, α^{+} means the pseudo inverse matrix of α, and $\otimes$ denotes Kronecker product.

S2 Recover Matrix B from $BB^{\top}$

In the following, we describe a method for recovering the 2×2 matrix B from $BB^{\top}$ up to a 2D rotation. Let $B = \begin{bmatrix} b_1 \cos\theta, & b_1 \sin\theta \\ b_3 \cos\phi, & b_3 \sin\phi \end{bmatrix}$, then:

$$
BB^{\top} = \begin{bmatrix} (b_1)^2, & b_1 b_3 \cos(\theta - \phi) \\ b_1 b_3 \cos(\theta - \phi), & (b_3)^2 \end{bmatrix} = \begin{bmatrix} bb_1, & bb_2 \\ bb_2, & bb_3 \end{bmatrix}. \tag{S5}
$$

If we assume $\phi = 0$ (due to the "fake" rotation ambiguity on yz-plane) and $b_1, b_3 \geq 0$ (due to the "fake" direction ambiguities of $y-$ and $z-$ axes), all the unknown parameters ($i.e.$ b_1, b_3, θ) can be calculated by:

$$
b_1 = \sqrt{bb_1}, \qquad b_3 = \sqrt{bb_3}, \qquad \theta = \arccos\left(\frac{bb_2}{b_1 b_3}\right) + \phi, \qquad \phi = 0. \tag{S6}
$$

S3 Derivation Details from Eqs. (30) - (32) to Eq. (33)

The details of obtaining Eq. (33) from Eqs. (30) - (32) are given by Eqs. (S7) - (S9):

$$
R_n = [R_n^1, R_n^2] = [\hat{R}_n^1, \hat{R}_n^2] \begin{bmatrix} \lambda, & \mathbf{0} \\ \mathbf{0}, & B \end{bmatrix} = \hat{R}_n \begin{bmatrix} \lambda, & \mathbf{0} \\ \mathbf{0}, & B \end{bmatrix} \tag{S7}
$$

$$
R_n R_n^{\top} = I. \tag{S8}
$$

$$
\hat{R}_n = [\hat{R}_n^1, \hat{R}_n^2] = \begin{bmatrix} \hat{r}_n^{1,1}, & \hat{r}_n^{1,2:3} \\ \hat{r}_n^{2,1}, & \hat{r}_n^{2,2:3} \end{bmatrix} \tag{S9}
$$

Substituting Eqs. (S7) and (S9) into Eq. (S8) yields:

$$
\begin{aligned}
I = & R_n R_n^{\top} = \hat{R}_n \begin{bmatrix} \lambda^2, & \mathbf{0} \\ \mathbf{0}, & BB^{\top} \end{bmatrix} \hat{R}_n^{\top} = \begin{bmatrix} \hat{r}_n^{1,1}, & \hat{r}_n^{1,2:3} \\ \hat{r}_n^{2,1}, & \hat{r}_n^{2,2:3} \end{bmatrix} \begin{bmatrix} \lambda^2, & \mathbf{0} \\ \mathbf{0}, & BB^{\top} \end{bmatrix} \begin{bmatrix} \hat{r}_n^{1,1}, & \hat{r}_n^{1,2:3} \\ \hat{r}_n^{2,1}, & \hat{r}_n^{2,2:3} \end{bmatrix}^{\top} \\
= & \begin{bmatrix} \lambda^2 (\hat{r}_n^{1,1})^2 + \hat{r}_n^{1,2:3} BB^{\top} (\hat{r}_n^{1,2:3})^{\top}, & \lambda^2 \hat{r}_n^{1,1} \hat{r}_n^{2,1} + \hat{r}_n^{1,2:3} BB^{\top} (\hat{r}_n^{2,2:3})^{\top} \\ \lambda^2 \hat{r}_n^{1,1} \hat{r}_n^{2,1} + \hat{r}_n^{2,2:3} BB^{\top} (\hat{r}_n^{1,2:3})^{\top}, & \lambda^2 (\hat{r}_n^{2,1})^2 + \hat{r}_n^{2,2:3} BB^{\top} (\hat{r}_n^{2,2:3})^{\top} \end{bmatrix}.
\end{aligned} \tag{S10}
$$

Thus we have:

$$
\lambda^2 (\hat{r}_n^{1,1})^2 + \hat{r}_n^{1,2:3} BB^{\top} (\hat{r}_n^{1,2:3})^{\top} = 1, \tag{S11}
$$

$$
\lambda^2 (\hat{r}_n^{2,1})^2 + \hat{r}_n^{2,2:3} BB^{\top} (\hat{r}_n^{2,2:3})^{\top} = 1, \tag{S12}
$$

$$
\lambda^2 \hat{r}_n^{1,1} \hat{r}_n^{2,1} + \hat{r}_n^{1,2:3} BB^{\top} (\hat{r}_n^{2,2:3})^{\top} = 0 \tag{S13}
$$

Vectorizing $BB^\top$ of Eqs. (S11)–(S13) by equation $\text{vec}(A\,XB^\top) = (B \otimes A)\text{vec}(X)$, gives the following linear equations:

$$\begin{bmatrix} (\hat{r}_n^{1,1})^2, & \hat{r}_n^{1,2:3} \otimes \hat{r}_n^{1,2:3} \\ (\hat{r}_n^{2,1})^2, & \hat{r}_n^{2,2:3} \otimes \hat{r}_n^{2,2:3} \\ \hat{r}_n^{1,1}\hat{r}_n^{2,1}, & \hat{r}_n^{1,2:3} \otimes \hat{r}_n^{2,2:3} \end{bmatrix} \begin{bmatrix} \lambda^2 \\ \text{vec}(BB^\top) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}. \tag{S14}$$

Note that $BB^\top$ is a symmetric matrix, meaning that the second and third elements of $\text{vec}(BB^\top)$ are the same. In order to enforce this, we sum the third and forth columns of the coefficient matrix of Eq. (S14). Let $\text{vec}(BB^\top) = [bb_1, bb_2, bb_2, bb_3]^\top$, then Eq. (S14) can be rewritten as:

$$\begin{bmatrix} (\hat{r}_n^{1,1})^2, & \hat{r}_n^{1,2:3} \otimes \hat{r}_n^{1,2:3} \\ (\hat{r}_n^{2,1})^2, & \hat{r}_n^{2,2:3} \otimes \hat{r}_n^{2,2:3} \\ \hat{r}_n^{1,1}\hat{r}_n^{2,1}, & \hat{r}_n^{1,2:3} \otimes \hat{r}_n^{2,2:3} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda^2 \\ bb_1 \\ bb_2 \\ bb_3 \end{bmatrix} = A_n\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \tag{S15}$$

where the third and fourth columns of the second matrix are identical (to enforce the symmetry of $BB^\top$) This gives Eq. (33) in the paper.

S4 Derivation Details from Eqs. (44) and (46) to Eq. (47)

Similarly, we describe the derivations from Eqs. (44) and (46) by Eqs. (S16) and (S17):

$$\hat{\Pi}_n^1 h_k^1 = \begin{bmatrix} \hat{\pi}_n^{1,1K} \\ \hat{\pi}_n^{2,1K} \end{bmatrix} h_k^1 = z_{nk} \begin{bmatrix} r_n^{11} \\ r_n^{21} \end{bmatrix}, \qquad \hat{\Pi}_n^2 h_k^2 = \begin{bmatrix} \hat{\pi}_n^{1,K+1:3K} \hat{\pi}_n^{2,K+1:3K} \end{bmatrix} h_k^2 = z_{nk} \begin{bmatrix} r_n^{1,2:3} \\ r_n^{2,2:3} \end{bmatrix} \tag{S16}$$

$$[\hat{\Pi}_n^1 h_k^1, \hat{\Pi}_n^2 h_k^2][\hat{\Pi}_n^1 h_k^1, \hat{\Pi}_n^2 h_k^2]^\top = z_{nk}^2 I \tag{S17}$$

Substituting Eq. (S16) into Eq. (S17) yields:

$$z_{nk}^2 I = [\hat{\Pi}_n^1 h_k^1, \hat{\Pi}_n^2 h_k^2][\hat{\Pi}_n^1 h_k^1, \hat{\Pi}_n^2 h_k^2]^\top = \begin{bmatrix} \hat{\pi}_n^{1,1:K}, & \hat{\pi}_n^{1,K+1:3K} \\ \hat{\pi}_n^{2,1:K}, & \hat{\pi}_n^{2,K+1:3K} \end{bmatrix} \begin{bmatrix} h_k^1 h_k^{1\top}, & \mathbf{0}_{K\times 2} \\ \mathbf{0}_{2K\times 1}, & h_k^2 h_k^{2\top} \end{bmatrix} \begin{bmatrix} \hat{\pi}_n^{1,1:K}, & \hat{\pi}_n^{1,K+1:3K} \\ \hat{\pi}_n^{2,1:K}, & \hat{\pi}_n^{2,K+1:3K} \end{bmatrix}^\top \tag{S18}$$

Thus, we have

$$\hat{\pi}_n^{1,1:K} h_k^1 h_k^{1\top} (\hat{\pi}_n^{1,1:K})^\top + \hat{\pi}_n^{1,K+1:3K} h_k^2 h_k^{2\top} (\hat{\pi}_n^{1,K+1:3K})^\top = z_{nk}^2, \tag{S19}$$

$$\hat{\pi}_n^{2,1:K} h_k^1 h_k^{1\top} (\hat{\pi}_n^{2,1:K})^\top + \hat{\pi}_n^{2,K+1:3K} h_k^2 h_k^{2\top} (\hat{\pi}_n^{2,K+1:3K})^\top = z_{nk}^2, \tag{S20}$$

$$\hat{\pi}_n^{1,1:K} h_k^1 h_k^{1\top} (\hat{\pi}_n^{2,1:K})^\top + \hat{\pi}_n^{1,K+1:3K} h_k^2 h_k^{2\top} (\hat{\pi}_n^{2,K+1:3K})^\top = 0. \tag{S21}$$

Eliminating the unknown value z_{nk}^2 in Eqs. (S19) and (S20) (by subtraction) and rewriting in vectorized form gives the constraints shown in Eq. (S22):

$$\begin{bmatrix} \hat{\pi}_n^{1,1:K} \otimes \hat{\pi}_n^{1,1:K} - \hat{\pi}_n^{2,1:K} \otimes \hat{\pi}_n^{2,1:K}, & \hat{\pi}_n^{1,K+1:3K} \otimes \hat{\pi}_n^{1,K+1:3K} - \hat{\pi}_n^{2,K+1:3K} \otimes \hat{\pi}_n^{1,K+1:3K} \\ \hat{\pi}_n^{1,1:K} \otimes \hat{\pi}_n^{2,1:K}, & \hat{\pi}_n^{1,K+1:3K} \otimes \hat{\pi}_n^{2,K+1:3K} \end{bmatrix} \begin{bmatrix} \text{vec}(h_k^1 h_k^{1\top}) \\ \text{vec}(h_k^2 h_k^{2\top}) \end{bmatrix} = A_n x = 0. \tag{S22}$$

Thus, we obtain Eq. (47) in the paper.

S5 M-Step of Sym-EM-PPCA

This step is to maximize the complete (joint) log-likelihood $P(\mathbb{Y}_n, \mathbb{Y}_n^\dagger, \mathbf{V}|z_n; G_n, \bar{\mathbb{S}}, \mathbf{V}^\dagger, \mathbb{T})$. The complete log-likelihood $Q(\theta)$ is:

$$
\begin{aligned}
\mathcal{Q}(\theta) =& -\sum_n \ln P(\mathbb{Y}_n, \mathbb{Y}_n^\dagger|z_n; G_n, \bar{\mathbb{S}}, \mathbf{V}, \mathbf{V}^\dagger, \mathbb{T}) + \lambda||\mathbf{V}^\dagger - \mathcal{A}_P\mathbf{V}||^2 \\
=& -\sum_n \big(\ln P(\mathbb{Y}_n|z_n; G_n, \bar{\mathbb{S}}, \mathbf{V}, \mathbb{T}) + \ln P(\mathbb{Y}_n^\dagger|z_n; G_n, \bar{\mathbb{S}}, \mathbf{V}^\dagger, \mathbb{T})\big) + \lambda||\mathbf{V}^\dagger - \mathcal{A}_P\mathbf{V}||^2 \\
=& 2PN\ln(2\pi\sigma^2) + \frac{1}{2\sigma^2}\sum_n E_{z_n}||\mathbb{Y}_n - G_n(\bar{\mathbb{S}} + \mathbf{V}z_n) - \mathbb{T}_n||^2 \\
&+ \frac{1}{2\sigma^2}\sum_n E_{z_n}||\mathbb{Y}_n^\dagger - G_n(\mathcal{A}_P\bar{\mathbb{S}} + \mathbf{V}^\dagger z_n) - \mathbb{T}_n||^2 + \lambda||\mathbf{V} - \mathcal{A}_P\mathbf{V}^\dagger||^2 \\
&\text{s. t.} \qquad R_n R_n^\top = I,
\end{aligned} \tag{S23}
$$

where $\theta = \{G_n, \bar{\mathbb{S}}, \mathbf{V}, \mathbf{V}^\dagger, \mathbb{T}_n, \sigma^2\}$. $\mathbb{Y}_n \in \mathbb{R}^{2P\times 1}, \bar{\mathbb{S}} \in \mathbb{R}^{3P\times 1}$, and $\mathbb{T}_n \in \mathbb{R}^{2P\times 1}$ are the stacked vectors of 2D keypoints, 3D mean structure and translations. $G_n = I_P \otimes c_n R_n$, in which c_n is the scale parameter for weak perspective projection, $\mathbf{V} = [\mathbb{V}_1, ..., \mathbb{V}_K] \in \mathbb{R}^{3P\times K}$ is the grouped K deformation bases, $z_n \in \mathbb{R}^{K\times 1}$ is the coefficient of the K bases. $\mathcal{A} = I_P \otimes \mathcal{A}$, and $\mathcal{A} = \mathrm{diag}([-1, 1, 1])$ is a matrix operator which negates the first row of its right matrix.

We first update the shape parameters $\bar{\mathbb{S}}, \mathbf{V}, \mathbf{V}^\dagger$ by maximize the log-likelihood $\mathcal{Q}$. Since these 3 parameters are related to each other in their derivations, thus they should be updated jointly by setting the 3 derivations to 0. According to Eq. (S23):

$$
\begin{bmatrix} A^*, & (B^*)^\top, & C^* \\ B^*, & D^* + 2\lambda\sigma^2 I_{3PK}, & -I_K \otimes 2\lambda\sigma^2\mathcal{A}_P \\ B^*\mathcal{A}_P, & -I_K \otimes 2\lambda\sigma^2\mathcal{A}_P^\top, & D^* + I_K \otimes 2\lambda\sigma^2\mathcal{A}_P^\top\mathcal{A}_P \end{bmatrix}
\begin{bmatrix} \mathbb{S} \\ \mathrm{vec}(\mathbf{V}) \\ \mathrm{vec}(\mathbf{V}^\dagger) \end{bmatrix} =
\begin{bmatrix} \mathrm{vec}\left(\sum_n G_n^\top(\mathbb{Y} - \mathbb{T}_n) + \mathcal{A}_P^\top G_n^\top(\mathbb{Y}^\dagger - \mathbb{T}_n)\right) \\ \mathrm{vec}\left(\sum_n G_n^\top(\mathbb{Y} - \mathbb{T}_n)\mu_n^\top\right) \\ \mathrm{vec}\left(\sum_n G_n^\top(\mathbb{Y}^\dagger - \mathbb{T}_n)\mu_n^\top\right), \end{bmatrix}, \tag{S24}
$$

where we have:

$$
A^* = \sum_n G_n^\top G_n + \mathcal{A}_P^\top G_n^\top G_n \mathcal{A}_P, \quad B^* = \sum_n \mu_n \otimes G_n^\top G_n, \quad C^* = \sum_n \mu_n^\top \otimes \mathcal{A}_P^\top G_n^\top G_n, \quad D^* = \sum_n \phi_n^\top \otimes G_n^\top G_n. \tag{S25}
$$

The camera parameters t_n, c_n, R_n and the variance of the noise σ^2 can be updated similarly as Bregler's method [1]. We first replace some parameters to make the equation to be homogeneous:

$$
\tilde{\mathbf{V}} = [\mathbb{S}, \mathbf{V}], \qquad \tilde{\mathbf{V}}^\dagger = [\mathcal{A}_P\mathbb{S}, \mathbf{V}^\dagger], \qquad \tilde{\mu_n} = [1, \mu_n^\top]^\top, \qquad \tilde{\phi_n} = \begin{bmatrix} 1 & \mu_n^\top \\ \mu_n & \phi_n \end{bmatrix} \tag{S26}
$$

Then the estimations of new σ^2, t_n, c_n are:

$$
\begin{aligned}
\sigma^2 =& \frac{1}{4PN}\sum_n \Big(||\mathbb{Y}_n - \mathbb{T}_n||^2 + ||\mathbb{Y}_n^\dagger - \mathbb{T}_n||^2 - 2(\mathbb{Y}_n - \mathbb{T}_n)G_n\tilde{\mathbf{V}}\tilde{\mu_n} - 2(\mathbb{Y}_n^\dagger - \mathbb{T}_n)G_n\tilde{\mathbf{V}}^\dagger\tilde{\mu_n} \\
&+ \mathrm{tr}(\tilde{\mathbf{V}}^\top G_n^\top G_n \tilde{\mathbf{V}}\tilde{\phi_n}) + \mathrm{tr}(\tilde{\mathbf{V}}^{\dagger\top} G_n^\top G_n \tilde{\mathbf{V}}^\dagger\tilde{\phi_n})\Big)
\end{aligned} \tag{S27}
$$

$$
t_n = \frac{1}{2P}\sum_{p=1}^{P}(\mathbb{Y}_{n,p} - c_n R_n \tilde{\mathbf{V}}_p\tilde{\mu_n} + \mathbb{Y}_{n,p}^\dagger - c_n R_n \tilde{\mathbf{V}}_p^\dagger\tilde{\mu_n}) \tag{S28}
$$

$$
c_n = \frac{\sum_{p=1}^{P}\left(\tilde{\mu_n}^\top\tilde{\mathbf{V}}_p^\top R_n^\top(\mathbb{Y}_{n,p} - t_n) + \tilde{\mu_n}^\top\tilde{\mathbf{V}}_p^{\dagger\top} R_n^\top(\mathbb{Y}_{n,p}^\dagger - t_n)\right)}{\sum_{p=1}^{P}\mathrm{tr}(\tilde{\mathbf{V}}_p^\top R_n^\top R_n \tilde{\mathbf{V}}_p\tilde{\phi_n} + \tilde{\mathbf{V}}_p^{\dagger\top} R_n^\top R_n \tilde{\mathbf{V}}_p^\dagger\tilde{\phi_n})} \tag{S29}
$$

Since R_n is subject to a nonlinear orthonormality constraint and cannot be updated in closed form, we follow an alternative approach used in [1] to parameterize R_n as a complete 3×3 rotation matrix Q_n and update the incremental rotation on Q_n instead, *i.e.* $Q_n^{new} = e^{\xi}Q_n$.

	$\sigma = \mathbf{0.03}\ d_{max}$										$\sigma = \mathbf{0.05}\ d_{max}$				
	I	II	III	IV	V	VI	VII	VIII	IX	X	I	II	III	IV	V
RSfM	0.57	0.69	0.55	1.09	0.68	1.58	0.67	1.45	0.97	0.37	0.59	0.74	0.53	1.11	0.68
CSF (S)	0.95	1.22	1.10	1.01	1.01	1.03	1.16	**1.02**	1.35	1.04	0.95	1.29	1.05	1.02	1.00
CSF (R)	0.95	1.27	1.07	1.05	1.05	0.98	0.95	1.04	1.05	1.16	0.95	1.24	1.07	1.04	1.07
Sym-RSfM	**0.37**	**0.42**	**0.34**	**0.46**	**0.32**	**0.26**	**0.32**	1.06	**0.25**	**0.12**	**0.39**	**0.45**	**0.31**	**0.48**	**0.31**

	$\sigma = \mathbf{0.05}\ d_{max}$ *(cont.)*					$\sigma = \mathbf{0.05}\ d_{max}$									
	VI	VII	VIII	IX	X	I	II	III	IV	V	VI	VII	VIII	IX	X
RSfM	1.54	0.69	1.49	0.99	0.40	0.59	0.72	0.56	1.09	0.64	1.54	0.67	1.55	0.97	0.38
CSF (S)	1.04	1.09	**1.03**	1.37	1.02	1.04	1.25	1.05	1.16	0.98	1.03	0.88	1.02	1.20	1.04
CSF (R)	0.97	0.93	**1.03**	0.89	1.16	0.95	1.22	1.06	1.04	0.88	0.96	1.00	1.04	0.91	1.05
Sym-RSfM	**0.28**	**0.33**	1.13	**0.28**	**0.13**	**0.38**	**0.37**	**0.30**	**0.49**	**0.28**	**0.35**	**0.33**	**1.01**	**0.26**	**0.13**

Table S1: The mean *rotation* errors for the *rigid* SfM methods on *car* with imperfect annotations. The noise is Gaussian $\mathcal{N}(0, \sigma^2)$ with $\sigma = sd_{max}$, where we choose $s = 0.03, 0.05, 0.07$ and d_{max} is the longest distance between all the keypoints (*i.e.* the left/right front wheel to the right/left back roof top). The Roman numerals denotes the index of the subtype. Each result value is obtained by averaging 10 repetitions.

	$\sigma = \mathbf{0.03}\ d_{max}$										$\sigma = \mathbf{0.05}\ d_{max}$				
	I	II	III	IV	V	VI	VII	VIII	IX	X	I	II	III	IV	V
RSfM	1.48	1.48	1.33	1.37	1.45	1.39	1.21	1.82	1.23	1.07	1.48	1.47	1.34	1.39	1.44
CSF (S)	1.06	2.36	1.14	**0.88**	1.37	1.17	**0.77**	1.13	2.00	0.98	1.10	1.22	1.15	**0.90**	1.34
CSF (R)	1.33	0.99	1.02	1.15	1.18	1.25	0.87	**0.90**	1.42	1.12	1.33	0.98	1.01	1.15	1.16
Sym-RSfM	**1.04**	**0.95**	**0.96**	1.08	**0.90**	**1.12**	0.81	1.80	**0.88**	**0.66**	**1.04**	**0.95**	**0.95**	1.08	**0.90**

	$\sigma = \mathbf{0.05}\ d_{max}$ *(cont.)*					$\sigma = \mathbf{0.05}\ d_{max}$									
	VI	VII	VIII	IX	X	I	II	III	IV	V	VI	VII	VIII	IX	X
RSfM	1.43	1.20	1.81	1.21	1.08	1.48	1.46	1.33	1.36	1.43	1.48	1.21	1.79	1.22	1.08
CSF (S)	**1.20**	1.04	1.11	1.96	0.97	3.56	1.28	1.15	1.19	1.38	1.22	2.08	1.06	2.47	1.03
CSF (R)	1.25	0.87	**0.88**	1.41	1.11	1.31	**0.92**	0.99	1.15	1.17	1.25	0.87	**0.88**	1.55	1.11
Sym-RSfM	1.25	**0.81**	1.71	**0.88**	**0.67**	**1.05**	0.95	**0.96**	**1.09**	**0.89**	**1.07**	**0.82**	1.80	**0.88**	**0.68**

Table S2: The mean *shape* errors for the *rigid* SfM methods on *car* with imperfect annotations. Other parameters are the same as Table S1.

Here, the first and second rows of Q_n is the same as R_n, and the third row of Q_n is obtained by the cross product of its first and second rows. The relationship of Q_n and R_n can be revealed by a matrix operator $\mathcal{M}$:

$$R_n = \mathcal{M}Q_n, \qquad \mathcal{M} = \begin{bmatrix} 1, 0, 0 \\ 0, 1, 0 \end{bmatrix}. \tag{S30}$$

Note that the incremental rotation e^{ξ} can be further approximated by its first order Taylor Series, *i.e.* $e^{\xi} \approx I + \xi$. Finally, we have:

$$R_n^{new}(\xi) = \mathcal{M}(I + \xi)Q_n. \tag{S31}$$

Therefore, setting $\partial\mathcal{Q}/\partial R_n = 0$, then replace R_n by Q_n using Eq. (S31) and vectorize it, we have:

$$R_n = \mathcal{M}e^{\xi}Q_n \approx \mathcal{M}(I + \xi)Q_n \quad \text{and} \quad \mathrm{vec}(\xi) = \alpha^{+}\beta \tag{S32}$$

$$\alpha = \left(c_n^2 \sum_{p=1}^{P} (\tilde{\mathbf{V}}_p^\top \tilde{\phi_n} \tilde{\mathbf{V}}_p + \tilde{\mathbf{V}}_p^{\dagger\top} \tilde{\phi_n} \tilde{\mathbf{V}}_p^{\dagger})^\top Q_n^\top \right) \otimes \mathcal{M}; \tag{S33}$$

$$\beta = \mathrm{vec}\left(c_n \sum_{p=1}^{P} \left((\mathbb{Y}_{n,p} - t_n)\tilde{\mu_n}^\top \tilde{\mathbf{V}}_p^\top + (\mathbb{Y}_{n,p}^{\dagger} - t_n)\tilde{\mu_n}^\top \tilde{\mathbf{V}}_p^{\dagger\top} \right) - c_n^2 \mathcal{M} Q_n \sum_{p=1}^{P} (\tilde{\mathbf{V}}_p^\top \tilde{\phi_n} \tilde{\mathbf{V}}_p + \tilde{\mathbf{V}}_p^{\dagger\top} \tilde{\phi_n} \tilde{\mathbf{V}}_p^{\dagger}) \right) \tag{S34}$$

where the subscript p means the pth keypoint, α^{+} means the pseudo inverse matrix of α, $\otimes$ denotes Kronecker product.

Table S3: The mean *shape* and *rotation* errors for the *non-rigid* SfM methods on *aeroplane* with imperfect annotations. The noise is Gaussian $\mathcal{N}(0, \sigma^2)$ with $\sigma = sd_{max}$, where we choose $s = 0.03, 0.05, 0.07$ and d_{max} is the longest distance between all the keypoints (*i.e.* the tip of the nose to the tip of the tail for aeroplane). Each result value is obtained by averaging 10 repetitions.

	$\sigma =$ **0.03** d_{max}								$\sigma =$ **0.05** d_{max}			
	I	II	III	IV	V	VI	VII	mRE	I	II	III	IV
EP	0.34	0.59	0.49	0.45	0.54	0.55	**0.45**	0.33	0.37	0.58	0.51	0.47
PF	0.92	1.01	1.05	1.06	1.13	1.03	1.06	0.52	0.93	1.04	1.05	1.08
Sym-EP	**0.34**	**0.54**	**0.47**	**0.44**	**0.52**	**0.55**	0.46	**0.32**	**0.35**	**0.54**	**0.47**	**0.43**
Sym-PF	0.79	0.93	1.01	0.93	0.91	0.79	0.94	0.60	0.83	0.99	1.09	0.98

	$\sigma =$ **0.05** d_{max}				$\sigma =$ **0.07** d_{max}							
	V	VI	VII	mRE	I	II	III	IV	V	VI	VII	mRE
EP	0.57	0.57	0.47	0.35	0.38	0.61	0.50	0.45	0.61	0.57	**0.47**	0.36
PF	1.15	1.02	1.07	0.54	0.94	1.04	1.08	1.07	1.15	1.03	1.08	0.65
Sym-EP	**0.52**	**0.57**	**0.46**	**0.33**	**0.37**	**0.58**	**0.49**	**0.44**	**0.58**	**0.57**	0.47	**0.35**
Sym-PF	0.94	0.84	1.04	0.63	0.94	1.06	1.15	1.04	1.05	0.89	1.08	0.70

S6 Experimental Results on The Imperfect Annotations

In this section, we investigate what happens if the keypoints are not perfectly annotated. This is important to check because our method depends on keypoint pairs therefore may be sensitive to errors in keypoint location, which will inevitably arise when we use features detectors to detect the keypoints.

To simulate this, we add Gaussian noise $\mathcal{N}(0, \sigma^2)$ to the 2D annotations and re-do the experiments. The standard deviation is set to $\sigma = sd_{max}$, where d_{max} is the longest distance between all the keypoints (*e.g.* for *car*, it is the distance between the left/right front wheel to the right/left back roof top). We have tested for different s by: 0.03, 0.05, 0.07. Other experimental settings are the same as them in the main text, *i.e.* images with more than 5 visible keypoints are used.

For the *rigid* SfM experiments, we show the mean rotation errors and the mean shape errors on *car* with $s = 0.03, 0.05$, 0.07 in Tables S1 and S2. We also show the results for *non-rigid SfM* on *aeroplane* in Table S3. Each result value is obtained by averaging 10 repetitions. The results in Tables S1 - S3 show that the performances of all the methods decrease in general with the increase in the noise level. Nonetheless, our methods still outperform our counterparts with the noisy annotations (*i.e.* the imperfectly labeled annotations).

References

[1] Torresani, L., Hertzmann, A., Bregler, C.: Nonrigid structure-from-motion: Estimating shape and motion with hierarchical priors. IEEE Transactions on Pattern Analysis and Machine Intelligence 30, 878-892 (2008)