Exploiting Symmetry and/or Manhattan Properties for 3D Object Structure Estimation from Single and Multiple Images

Yuan Gao, Alan L. Yuille
1. Tencent AI Lab 2. UCLA 3. Johns Hopkins University
ethan.y.gao@gmail.com alan.yuille@jhu.edu

Motivation
Many objects, especially that made by human, have intrinsic symmetry and Manhattan structures, e.g. cars, aeroplanes, etc.

Our goal is to investigate how symmetry and Manhattan structures can improve 3D object structure reconstruction.

Contributions
1. Single image 3D reconstruction:
   - Manhattan Structure
   - Camera Matrix
   - Symmetry

   However, Manhattan structure may be hard to observe from a single image due to occlusion

2. Multiple images 3D reconstruction:
   - Symmetric Rigid Structure from Motion

   This method can deal with occlusion, which are updated iteratively.

   By exploring symmetry, Sym-RSfM method achieved state-of-the-art performance on Pascal3D+ dataset.

Single-image reconstruction
Let 3D KPs $S_1 - S_2 = [x, y, 0]^T$ $Y_1 - Y_2 = R(S_1 - S_2) = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11}x + r_{12}y + r_{13} \\ r_{21}x + r_{22}y + r_{23} \end{bmatrix}$. Therefore, we have $r_{12}/r_{22}$. Similarly, by exploring the other two Manhattan axes, we have $r_{13}/r_{23}, r_{21}/r_{23}$.

Imposing Orthogonality Constrains $RR^T = I$

Then, assume symmetry is along the x-axis:

$$\begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

where $\mu_1 = r_{11}/r_{21}, \mu_2 = r_{12}/r_{22}, \mu_3 = r_{13}/r_{23}$.

Multiple-image reconstruction
Input: Symmetric 2D keypoint pairs for multiple images for several rigid-deformed objects (e.g. sedan cars of different brand).

Output: The Common 3D structure of all the input images $S$, and the camera matrix for each image $R_n$.

Let $Y_n, Y'_n \in R^{N \times P}$ are symmetric pairs of image $n$:

$$Y_n = R_nS + N_n, \quad Y'_n = R'_nS + N_n.$$ 

The ambiguities can be solved by Orthogonality Constrains $RR^T = I$, where $R = [\hat{R}_1, \hat{R}_2, \hat{R}_3]$, $M = R^T R$.

Final Solution: $\hat{R} = R[0, 0, 0]$, $\hat{S} = [0, 0, 0]$. Ambiguities! Ambiguities!

Our solution is decomposing it into two independent terms to enable SVD.

$$L = R^T S_n = R^T \hat{L} = \hat{R}^T \hat{S}_n, \quad M = R^T R = R^T R^T.$$ 

The ambiguities can be solved by Orthogonality Constrains $RR^T = I$, where $R = \hat{R}$.

Shape error for car
Rotation error for car

Reference