Camera Resectioning from Image Edges with the L∞-Norm Using Linear Programming

Li Yang
yangli@cis.pku.edu.cn
Xianghua Ying*
http://www.cis.pku.edu.cn/vision/Visual&Robot/people/ying/
Lulu Hou
huyanlulu@163.com
Jing Kong
superkj@pku.edu.cn
Hongbin Zha
http://www.cis.pku.edu.cn/vision/Visual&Robot/people/zha/

Key Lab of Machine Perception, MOE
School of EECS, Peking University
Beijing, 100871 P.R. China

Camera resectioning is known as the computation of the camera projection matrix from corresponding 3D space and image entities. Most related work makes use of the point to point correspondences. Under the circumstances, there are two different ways to process the problem. One way is based on algebraic distance [1]. For example, using Direct Linear Transform(DLT) algorithm gets a solution, then it is used as a start point for a non-linear minimization of a geometric or statistical cost function to get a better solution. The other way is based on geometric distance. One of the recent fashions is the L_∞-Norm minimization framework [2, 3, 4]. A variety of structure and motion problems can be solved using Second Order Cone Programming(SOCP) within the framework.

In this paper, there are several ways different from previous work about camera resectioning. Fisrt, our method is based on line segment correspondences instead of point ones. Because low-level vision routines are relatively good at finding the middle parts of lines due to occlusions, shadows etc., discussing camera resectioning using line segment correspondences have a significant meaning. Second, our solution uses linear programming(LP) instead of using SOCP utilized by most previous works. Even though previous work [5] provides a LP method, it is not rotation-invariant since its residual error is defined by L_∞-Norm. While our proposed definition is based on Euclidean distance keeping rotation invariance.

Given known 3D line segment $\mathbf{X}_1\mathbf{X}_2$ and its corresponding image line segment $\mathbf{x}_1\mathbf{x}_2$, the projection matrix of camera $\mathbf{P}$ should satisfy the constraint that if we project the line segment $\mathbf{X}_1\mathbf{X}_2$ using it, the line segment $\hat{\mathbf{x}}_1\hat{\mathbf{x}}_2$ we get, where $s_1\hat{\mathbf{x}}_1 = \mathbf{P}\mathbf{X}_1, s_2\hat{\mathbf{x}}_2 = \mathbf{P}\mathbf{X}_2$, should "equal" the given image line segment $\mathbf{x}_1\mathbf{x}_2$. When noise is present in practice due to occlusion, shadows and so on, the $\hat{\mathbf{x}}_1\hat{\mathbf{x}}_2$ should be "nearly equal" to $\mathbf{x}_1\mathbf{x}_2$. To measure the error associated with a single correspondence, we define the residue, or distance

$$d(\mathbf{x}_1\mathbf{x}_2, \hat{\mathbf{x}}_1\hat{\mathbf{x}}_2) = \max\{S_{\Delta\mathbf{x}_1\mathbf{x}_2\hat{\mathbf{x}}_1}(\mathbf{P}), S_{\Delta\mathbf{x}_1\mathbf{x}_2\hat{\mathbf{x}}_2}(\mathbf{P})\}, \tag{1}$$

where $S_{\Delta\mathbf{uvw}}$ represents the area of the triangle $\Delta\mathbf{uvw}$ in the image plane. The individual functions $S_{\Delta\mathbf{x}_1\mathbf{x}_2\hat{\mathbf{x}}_k}$ for $k = 1,2$ have following form

$$S_{\Delta\mathbf{x}_1\mathbf{x}_2\hat{\mathbf{x}}_k}(\mathbf{P}) = \frac{\left|\left(d_1\mathbf{X}_k^T, d_2\mathbf{X}_k^T, d_3\mathbf{X}_k^T\right)\alpha\right|}{\left(\mathbf{0}_{1\times4}, \mathbf{0}_{1\times4}, \mathbf{X}_k^T\right)\alpha}, \tag{2}$$

where α is the rearranged vector from the camera projection matrix $\mathbf{P}$

$$\alpha = (a_{11}, a_{12}, a_{13}, a_{14}, a_{21}, a_{22}, a_{23}, a_{24}, a_{31}, a_{32}, a_{33}, a_{34})^T,$$
$$d_1 = y_1 - y_2, d_2 = x_2 - x_1, d_3 = x_1y_2 - x_2y_1.$$

Let us rewrite the formula as a simplified one:

$$\frac{|\mathbf{c}^T\alpha|}{\mathbf{f}^T\alpha} \text{ and } \mathbf{f}^T\alpha \geq 0. \tag{3}$$

With the definition of distance between two line segments, we could formulate this problem. Let $\mathbf{X}_1^i\mathbf{X}_2^i$ denote a set of 3D line segments and let $\mathbf{x}_1^i\mathbf{x}_2^i$ be the corresponding image line segments for $i = 1\cdots m$. With all those correspondences exerting constraints to the projection, obtaining the best estimate of the $\mathbf{P}$ means that the mapped line segments $\hat{\mathbf{x}}_1^i\hat{\mathbf{x}}_2^i$ are most closely to the image line segments $\mathbf{x}_1^i\mathbf{x}_2^i$.

In our context, the words "most closely" means to find the matrix $\mathbf{P}$ that minimizes the following cost function:

$$\max_i\{d(\mathbf{x}_1^i\mathbf{x}_2^i, \hat{\mathbf{x}}_1^i\hat{\mathbf{x}}_2^i)\}, \tag{4}$$

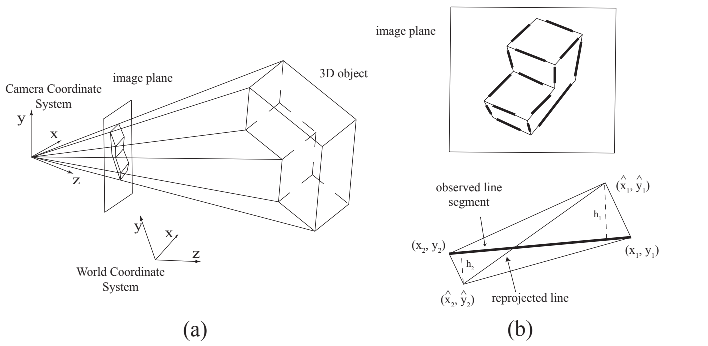

Figure 1: (**a**) Projection of an object onto a camera's image plane (**b**) Giving our definition of the residual error between the reprojected line segment and the observed line segment as the maximum area of the two different colored triangles

Namely, the problem is called the minimax (L_∞) optimization problem:

$$\min_{\mathbf{P}} \max_{i,k} \left\{S_{\Delta\mathbf{x}_1^i\mathbf{x}_2^i\hat{\mathbf{x}}_k^i}\right\}, \tag{5}$$

Finally, the problem is recast as a quasi-convex optimization problem, which has the following form:

$$\min_{\gamma,\alpha} \gamma \text{ subject to } \frac{|\mathbf{c}_j^T\alpha|}{\mathbf{f}_j^T\alpha} \leq \gamma \text{ and } \mathbf{f}_j^T\alpha \geq 0 \tag{6}$$

where $j = \langle i,k\rangle$. The quasiconvex optimization problem could be resolved using *Bisection* method via the LP feasibility problems:

$$\begin{array}{ll} \text{find} & \alpha \\ \text{subject to} & \begin{bmatrix} -\mathbf{c}_j^T - \gamma\mathbf{f}_j^T \\ \mathbf{c}_j^T - \gamma\mathbf{f}_j^T \\ -\mathbf{f}_j^T \end{bmatrix} \alpha \leq \mathbf{0} \qquad j = 1,\cdots,2m \end{array} \tag{7}$$

Our conclusion is that based on our proposed definition of the geometric distance between two line segments in a 2D plane, camera resectioning can be recast as a quasi-convex optimization problem, which can be solved by LP instead of SOCP, which means we simplify the problem using more complex elements, line segments rather than points. This paper may probably provide a special view to considering this problem.

[1] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[2] F. Kahl and R. Hartley. Multiple view geometry under the L_∞-norm. *PAMI*, 30(9):1603 – 1617, September 2008.

[3] Fredrik Kahl. Multiple view geometry and the L_∞-norm. In *International Conference on Computer Vision*, 2005.

[4] Qifa Ke and Takeo Kanade. Quasiconvex optimization for robust geometric reconstruction. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29:1834–1847, October 2007.

[5] Yongduek Seo and Richard Hartley. A fast method to minimize l error norm for geometric vision problems. In *International Conference on Computer Vision*, pages 1–8, 2007.