Robust Camera Location Estimation by Convex Programming

Onur Özyeşil[1] and Amit Singer[1,2]

[1]Program in Applied and Computational Mathematics, Princeton University. [2]Department of Mathematics, Princeton University.

3D structure recovery from a collection of 2D images requires the estimation of the camera locations and orientations, i.e. the camera motion. For large, irregular collections of images, existing methods for the location estimation part, which can be formulated as the inverse problem of estimating n locations $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n$ in $\mathbb{R}^3$ from noisy measurements of a subset of the pairwise directions $\gamma_{ij} := \frac{\mathbf{t}_i - \mathbf{t}_j}{\|\mathbf{t}_i - \mathbf{t}_j\|}$ (cf. Figure 1), are sensitive to outliers in direction measurements.

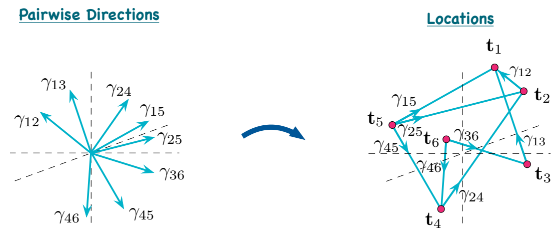

Figure 1: A (noiseless) instance of the location estimation problem in $\mathbb{R}^3$, with $n = 6$ locations and $m = 8$ pairwise directions.

We represent the pairwise directions using a measurement graph $G_t = (V_t, E_t)$, where the i'th node in $V_t = \{1, 2, \ldots, n\}$ corresponds to the location $\mathbf{t}_i$ and each edge $(i, j) \in E_t$ is endowed with the direction γ_{ij}. Provided with the set $\{\gamma_{ij}\}_{(i,j)\in E_t}$ of (noiseless) directions on $G_t = (V_t, E_t)$, we first consider the following fundamental questions: Can we uniquely realize $\{\mathbf{t}_i\}_{i\in V_t}$, of course, up to a global translation and scale? Is unique realizability a generic property of G_t and can it be decided efficiently? The unique realizability of locations (in $\mathbb{R}^d$, $d \geq 2$) was previously studied under the general title of *parallel rigidity theory* (see, e.g., [1, 4, 6]), and was shown to be a generic property of G_t, which admits a complete combinatorial characterization:

Theorem 1. *For a graph $G = (V, E)$, let $(d-1)E$ denote the set consisting of $(d-1)$ copies of each edge in E. Then, G is generically parallel rigid in $\mathbb{R}^d$ if and only if there exists a nonempty set $D \subseteq (d-1)E$, with $|D| = d|V| - (d+1)$, such that for all subsets D' of D, we have*

$$|D'| \leq d|V(D')| - (d+1)\,, \tag{1}$$

where $V(D')$ denotes the vertex set of the edges in D'.

There exist efficient algorithms for testing parallel rigidity (see, e.g., [4] for a randomized spectral test having a time complexity of $\mathcal{O}(m)$). Moreover, for a graph that is not parallel rigid, maximal uniquely realizable subgraphs can be efficiently extracted (see, e.g., [2]). For noisy direction measurements, we consider problem instances on parallel rigid graphs to be *well-posed* instances.

Now, suppose that we are given a set of noisy pairwise direction measurements $\{\gamma_{ij}\}_{(i,j)\in E_t}$, i.e., for each $(i, j) \in E_t$, γ_{ij} satisfies

$$\gamma_{ij} = \frac{\mathbf{t}_i - \mathbf{t}_j}{\|\mathbf{t}_i - \mathbf{t}_j\|} + \varepsilon^{\gamma}_{ij} \tag{2}$$

where, $\varepsilon^{\gamma}_{ij}$ denotes the direction error. Our objective is to estimate the locations $\{\mathbf{t}_i\}_{i\in V_t}$ by maintaining robustness to outliers (i.e., γ_{ij}'s with large $\varepsilon^{\gamma}_{ij}$'s) in a computationally efficient manner. In this respect, we first rewrite (2) as

$$\mathbf{t}_i - \mathbf{t}_j = \|\mathbf{t}_i - \mathbf{t}_j\|\gamma_{ij} + \varepsilon^{\mathbf{t}}_{ij} \tag{3}$$

$$\iff \varepsilon^{\mathbf{t}}_{ij} = \mathbf{t}_i - \mathbf{t}_j - d_{ij}\gamma_{ij} \tag{4}$$

where, $\varepsilon^{\mathbf{t}}_{ij}$ denotes the displacement error, and we define $d_{ij} := \|\mathbf{t}_i - \mathbf{t}_j\|$ to rewrite $\varepsilon^{\mathbf{t}}_{ij}$ linearly in $\mathbf{t}_i$, $\mathbf{t}_j$ and d_{ij}. Observe that, large direction errors $\varepsilon^{\gamma}_{ij}$'s induce large displacement errors $\varepsilon^{\mathbf{t}}_{ij}$'s. Hence, we employ displacement error minimization as a substitute for direction error minimization. Also, to maintain robustness to large $\varepsilon^{\mathbf{t}}_{ij}$'s, we minimize the sum of *unsquared* norms of $\varepsilon^{\mathbf{t}}_{ij}$'s, and for computational efficiency, we drop the intrinsic non-convex constraints $d_{ij} = \|\mathbf{t}_i - \mathbf{t}_j\|$ to obtain the convex "least unsquared deviations" (LUD) formulation

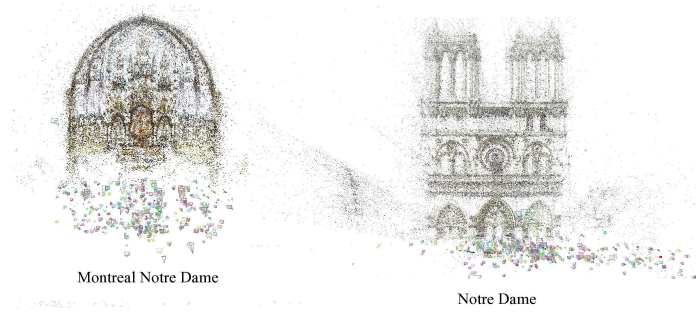

Figure 2: Snapshots of selected 3D structures computed using the initial location estimates of the LUD solver (5) (without bundle adjustment).

$$\begin{aligned} &\underset{\{\mathbf{t}_i\},\{d_{ij}\}}{\text{minimize}} && \sum_{(i,j)\in E_t} \left\| \mathbf{t}_i - \mathbf{t}_j - d_{ij}\gamma_{ij} \right\| \\ &\text{subject to} && \sum_{i\in V_t} \mathbf{t}_i = \mathbf{0}\ ;\ d_{ij} \geq 1,\ \forall (i,j) \in E_t \end{aligned} \tag{5}$$

where the constraints are used to remove the global ambiguities, and to prevent trivial solutions, as well as solutions clustered around a few locations.

In our paper, we also provide a new method for estimating γ_{ij} that is robust to outliers in point correspondences, and an efficient iteratively-reweighted least squares (IRLS) algorithm to solve the LUD problem (5). Additionally, we compare the performance of our formulations to various methods in the literature through experiments on synthetic data sets, and Internet photo collections (cf. Figure 2 for selected 3D structures computed using the LUD solver (5)), which demonstrate the relatively high accuracy and efficiency of our approach.

[1] T. Eren, W. Whiteley, and P. N. Belhumeur. Using angle of arrival (bearing) information in network localization. In *IEEE Conference on Decision & Control*, San Diego, CA, USA, December 2006.

[2] R. Kennedy, K. Daniilidis, O. Naroditsky, and C. J. Taylor. Identifying maximal rigid components in bearing-based localization. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, Algarve, Portugal, October 2012.

[3] P. Moulon, P. Monasse, and R. Marlet. Global fusion of relative motions for robust, accurate and scalable structure from motion. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 3248–3255, Dec 2013.

[4] O. Özyeşil, A. Singer, and R. Basri. Stable camera motion estimation using convex programming, 2015. *SIAM Journal on Imaging Sciences*, accepted. Also available at `http://arxiv.org/abs/1312.5047`.

[5] R. Tron and R. Vidal. Distributed 3-D localization of camera sensor networks from 2-D image measurements. *Automatic Control, IEEE Transactions on*, 59(12):3325–3340, Dec 2014.

[6] W. Whiteley. Matroids from discrete geometry. In J. G. Oxley, J. E. Bonin, and B. Servatius, editors, *Matroid Theory*, volume 197 of *Contemporary Mathematics*, pages 171–313. American Mathematical Society, 1996.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

We note that the least squares version of (5) (i.e., the program with the cost function $\sum_{(i,j)\in E_t} \|\mathbf{t}_i - \mathbf{t}_j - d_{ij}\gamma_{ij}\|^2$, and the same constraints as in (5)), and the ℓ_∞ version of (5), were previously studied (resp.) in [5] and in [3].