Efficient Globally Optimal Consensus Maximisation with Tree Search

Tat-Jun Chin*, Pulak Purkait*, Anders Eriksson† and David Suter*
School of Computer Science, The University of Adelaide*, School of Electrical Engineering and Computer Science, Queensland University of Technology†

Maximum consensus is one of the most popular robust criteria for geometric estimation problems in computer vision. Given a set of measurements $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^N$, the criterion aims to find the estimate θ that agrees with as many of the data as possible (i.e., the inliers) up to a threshold ε

$$\begin{aligned} \max_{\theta,\, \mathcal{I} \subseteq \mathcal{X}} \quad & |\mathcal{I}| \\ \text{subject to} \quad & r_i(\theta) \le \varepsilon \;\; \forall \mathbf{x}_i \in \mathcal{I}, \end{aligned} \tag{1}$$

where $r_i(\theta)$ is the residual of $\mathbf{x}_i$. For example, in triangulation we wish to estimate the 3D point θ seen in N views, where $\mathcal{X}$ contains the 2D observations $\mathbf{x}_i$ of the point and the associated camera matrices $\mathbf{P}_i \in \mathbb{R}^{3\times 4}$. The residual $r_i(\theta)$ is the reprojection error in the i-th view.

$$r_i(\theta) = \frac{\|(\mathbf{P}_{i,1:2} - \mathbf{x}_i \mathbf{P}_{i,3})\tilde{\theta}\|}{\mathbf{P}_{i,3}\tilde{\theta}}, \tag{2}$$

where $\tilde{\theta} = [\theta^T \; 1]^T$, $\mathbf{P}_{i,1:2}$ is the first-two rows of $\mathbf{P}_i$, and $\mathbf{P}_{i,3}$ is the third row of $\mathbf{P}_i$. The additional constraint $\mathbf{P}_{i,3}\tilde{\theta} > 0$ must be imposed such that the 3D point lies in front of the cameras. Another example is homography fitting, where given a set of point matches $\mathcal{X} = \{(\mathbf{u}_i, \mathbf{u}'_i)\}_{i=1}^N$ across two views, we wish to estimate the homography $\theta \in \mathbb{R}^{3\times 3}$ that aligns the points. The residual is

$$r_i(\theta) = \frac{\|(\theta_{1:2} - \mathbf{u}'_i \theta_3)\tilde{\mathbf{u}}_i\|}{\theta_3 \tilde{\mathbf{u}}_i}, \tag{3}$$

where $\tilde{\mathbf{u}}_i = [\mathbf{u}_i^T \; 1]^T$, $\theta_{1:2}$ is the first-2 rows of θ, and θ_3 is the 3rd row of θ.

In this work, we consider residual functions $r_i(\theta)$ that are *pseudo-convex*. This is known to include many other estimation problems in computer vision. Apart from the above two examples, other problems that involve pseudo-convex residuals include linear regression, camera resectioning, SfM with known rotations, etc. [4, 7].

Random sample consensus (RANSAC) [2] has been the dominant approach. The method randomly draws p-tuples from $\mathcal{X}$, where p is the minimum number of data to instantiate θ. The consensus score $|\mathcal{I}|$ of each sample θ is obtained, and the θ with the highest score and its consensus set $\mathcal{I}$ are returned. A probabilistic bound on the number of samples required can be derived as a stopping criterion.

A major shortcoming of randomised methods such as RANSAC is the lack of absolute certainty that the obtained solution is optimal, or indeed whether it represents a satisfactory approximation at all. A less recognised fact is that even if all $\binom{N}{p}$ subsets are examined, we may not find the globally optimal solution θ^* to (1), since θ^* does not generally correspond to an estimate from a p-tuple (see below).

Solving (1) exactly is computationally challenging. Several authors proposed methods based on branch-and-bound (BnB) [5, 10]. However, BnB is typically slow, especially if θ is high-dimensional. In fact, RANSAC is suggested to suboptimally preprocess the data (which may cause genuine inliers to be discarded), before BnB is invoked to refine the result. More fundamentally, the BnB methods are specialised for linear residuals. For many vision problems, this entails linearising $r_i(\theta)$ and adopting algebraic residuals which are not geometrically meaningful.

It has been proven [1, 8] that for various estimation tasks, θ^* can be found as the solution on a subset of $\mathcal{X}$ of size d, where $d \ge p$ and $d \ll N$ (the actual value of d depends on the particular problem). Both works proposed to find θ^* by exhaustively searching over all $\binom{N}{d}$ subsets of $\mathcal{X}$. Although the number of subsets is polynomial w.r.t. N, in realistic problems the number is impracticably large. Olsson et al. [8] also proposed using RANSAC with an optimality verification step. However, the fact remains that an enormous number of subsets may need to be sampled.

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

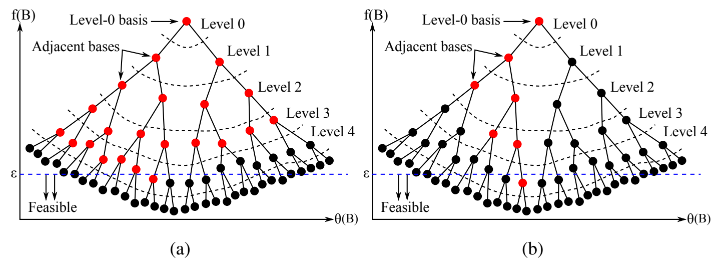

Figure 1: (a) Our work shows how consensus maximisation can be solved as a tree search problem. (b) We propose an efficient consensus maximisation algorithm based on A* search.

Due to the much greater computational expense, currently available global methods are not competitive against RANSAC and its variants. In this paper, we make significant progress towards solving (1) exactly *and* efficiently. Leveraging on the framework of LP-type methods [6, 9], we show how maximising consensus can be casted as a tree search problem. Figure 1(a) illustrates this idea.

We then propose an algorithm based on A* search [3] to traverse the tree. Similar in spirit to [1, 8], we aim to find the optimal data subset. However, instead of sampling or enumerating the subsets, our algorithm deterministically locates the best subset. The A* technique ensures that only a tiny fraction of available subsets need to be explored. Figure 1(b) illustrates our method. Despite its combinatorial nature, our algorithm is fast - on several common estimation problems, our algorithm is orders of magnitude faster than previous exact methods for maximum consensus. Further, our method does not require linearising the residual.

Acknowledgments This work was supported by ARC grants DP130102524 and DE130101775.

[1] O. Enqvist, E. Ask, F. Kahl, and K. Åström. Robust fitting for multiple view geometry. In *ECCV*, 2012.

[2] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 24(6):381–395, 1981.

[3] P. E. Hart, N. J. Nilsson, and B. Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. on Systems Science and Cybernetics*, 4(2):100–107, 1968.

[4] F. Kahl and R. I. Hartley. Multiple-view geometry under the l_∞-norm. *IEEE TPAMI*, 30(9):1603–1617, 2008.

[5] H. Li. Consensus set maximization with guaranteed global optimality for robust geometry estimation. In *ICCV*, 2009.

[6] J. Matoušek. On geometric optimization with few violated constraints. *Discrete and computational geometry*, 14(1):365–384, 1995.

[7] C. Olsson, A. Eriksson, and F. Kahl. Efficient optimization for l_∞-problems using pseudoconvexity. In *ICCV*, 2007.

[8] C. Olsson, O. Enqvist, and F. Kahl. A polynomial-time bound for matching and registration with outliers. In *CVPR*, 2008.

[9] M. Sharir and E. Welzl. A combinatorial bound for linear programming and related problems. In *STACS*, 1992.

[10] Y. Zheng, S. Sugimoto, and M. Okutomi. Deterministically maximizing feasible subsystems for robust model fitting with unit norm constraints. In *CVPR*, 2011.