Robust Model Reasoning and Fitting via Dual Sparsity Pursuit

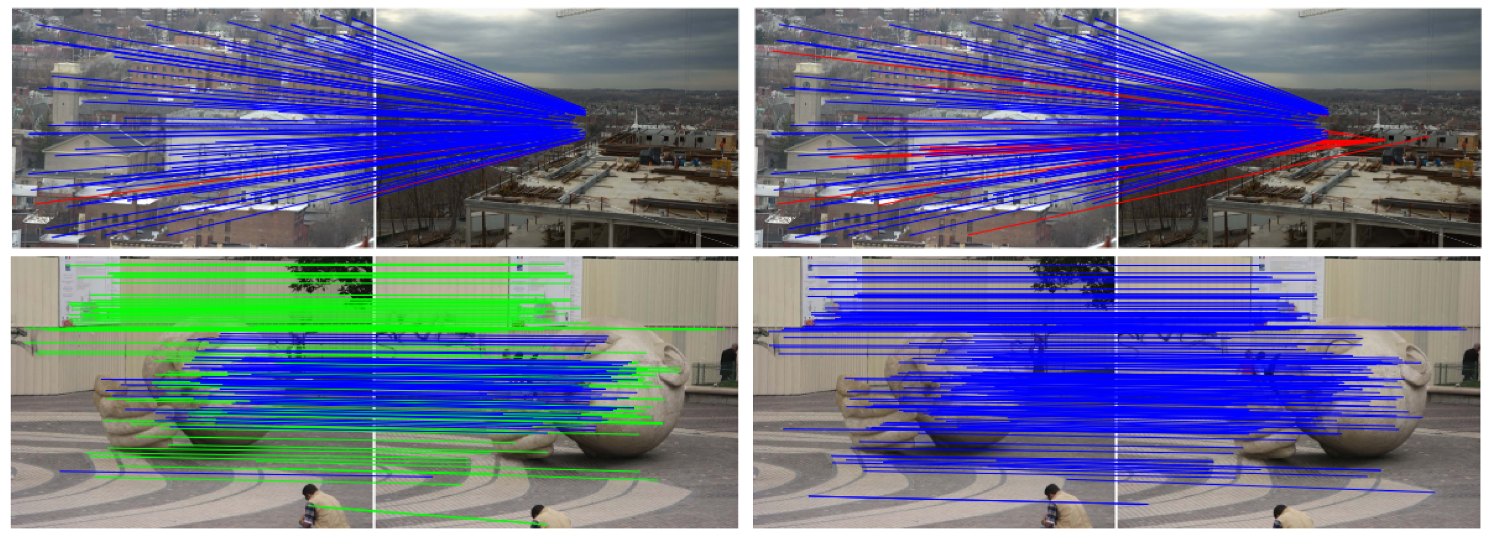
Xingyu Jiang[1,2] Jiayi Ma[2]

[1] Huazhong University of Science and Technology [2] Wuhan University

{jiangxy998, jyma2010}@gmail.com

Project & Code & Demo https://github.com/StaRainJ/DSP

Motivation



(a) Estimate *Homography* Model (b) Estimate *Fundamental* Model

Inlier Detection for Homo.(top) and Fund.(bottom) data. (blue = TP, green = FN, red = FP)

- ***Severe noises*** and ***outliers*** pose great challenges in accurate estimation.
- Existing methods need to ***predefine the true model type*** before fitting .
- Current greedy strategies first fit all possible models, then select the best model with geometric criteria– ***time consuming and inaccurate***.

We try to simultaneously solve i) *outlier rejection*, ii) *true model reasoning* and iii) *parameter estimation* with a unified optimization modeling.

Overview

Contributions:

- We are the first to propose a **continuous optimization modeling** for geometric model fitting with **unknown model type** and **dominant outliers**.
- We introduce **sparse subspace recovery theory**, which is a novel formulation for model reasoning and fitting task.
- We propose an **efficient solution** for our dual sparsity problem with convergence rate of $\mathcal{O}(1/k^2)$, that integrates the approximation strategy and sub-gradient decent method into alternative optimization paradigm.
- Extensive experiments on **known/unknown model fitting** and **two visual applications** are designed to validate the superiority of our method.

Outline of the embedding of data and model for the classical DLT, and our DSP formulation. We conclude the geometry relationship $G_{\mathcal{M}} = (D, d, r, s)$ for each model.

Model (req.)	DLT (data & model type:D,d,r)	DSP for known model fitting (data & model type:D,d,r)	DSP for unknown model fitting (data type:D)
Fund.	$\mathbf{m}_i^\top = (u_i'\mathbf{p}_i^\top, v_i'\mathbf{p}_i^\top, \mathbf{p}_i^\top)_{1\times 9}$ $\mathbf{x} = vec(\mathbf{F})$ $\mathbf{M}^\top\mathbf{x} = \mathbf{0}_{N\times 1}, G = (9,8,1,0)$	$\mathbf{m}_i^\top = (u_i'\mathbf{p}_i^\top, v_i'\mathbf{p}_i^\top, \mathbf{p}_i^\top)_{1\times 9}$ $\mathbf{x} = vec(\mathbf{F})$ $\mathbf{M}^\top\mathbf{x} = \mathbf{0}_{N\times 1}, G = (9,8,1,0)$	$\mathbf{m}_i^\top = (u_i'\mathbf{p}_i^\top, v_i'\mathbf{p}_i^\top, \mathbf{p}_i^\top)_{1\times 9}$ $\mathbf{x} = vec(\mathbf{F})$ $\mathbf{M}^\top\mathbf{x} = \mathbf{0}_{N\times 1}, G = (9,8,1,0)$
Homo.	$\mathbf{m}_i^\top = \begin{bmatrix}\mathbf{p}_i^\top, \mathbf{0}_{1\times 3}, -u_i'\mathbf{p}_i^\top \\ \mathbf{0}_{1\times 3}, \mathbf{p}_i^\top, -v_i'\mathbf{p}_i^\top\end{bmatrix}_{2\times 9}$ $\mathbf{x} = vec(\mathbf{H})$ $\mathbf{M}^\top\mathbf{x} = \mathbf{0}_{2N\times 1}, G = (9,8,1,0)$	$\mathbf{m}_i^\top = (u_i'\mathbf{p}_i^\top, v_i'\mathbf{p}_i^\top, \mathbf{p}_i^\top)_{1\times 9}$ $\mathbf{X} = \begin{bmatrix}\mathbf{0}_{1\times 3}, \mathbf{h}_3^\top, -\mathbf{h}_2^\top \\ -\mathbf{h}_3^\top, \mathbf{0}_{1\times 3}, \mathbf{h}_1^\top \\ \mathbf{h}_2^\top, -\mathbf{h}_1^\top, \mathbf{0}_{1\times 3}\end{bmatrix}^\top$ $\mathbf{M}^\top\mathbf{X} = \mathbf{0}_{N\times 3}, G = (9,6,3,3)$	$\mathbf{m}_i^\top = (u_i'\mathbf{p}_i^\top, v_i'\mathbf{p}_i^\top, \mathbf{p}_i^\top)_{1\times 9}$ $\mathbf{X} = \begin{bmatrix}\mathbf{0}_{1\times 3}, \mathbf{h}_3^\top, -\mathbf{h}_2^\top \\ -\mathbf{h}_3^\top, \mathbf{0}_{1\times 3}, \mathbf{h}_1^\top \\ \mathbf{h}_2^\top, -\mathbf{h}_1^\top, \mathbf{0}_{1\times 3}\end{bmatrix}^\top$ $\mathbf{M}^\top\mathbf{X} = \mathbf{0}_{N\times 3}, G = (9,6,3,3)$
Affine	$\mathbf{m}_i^\top = \begin{bmatrix}\mathbf{p}_i^\top, \mathbf{0}_{1\times 3}, -u_i' \\ \mathbf{0}_{1\times 3}, \mathbf{p}_i^\top, -v_i'\end{bmatrix}_{2\times 7}$ $\mathbf{x} = vec(\mathbf{A})$ $\mathbf{M}^\top\mathbf{x} = \mathbf{0}_{2N\times 1}, G = (7,6,1,0)$	$\mathbf{m}_i^\top = (u_i, v_i, u_i', v_i', 1)_{1\times 5}$ $\mathbf{X} = \begin{bmatrix}a_{11}, a_{12}, -1, 0, a_{13} \\ a_{21}, a_{22}, 0, -1, a_{23}\end{bmatrix}^\top$ $\mathbf{M}^\top\mathbf{X} = \mathbf{0}_{N\times 2}, G = (5,3,2,1)$	$\mathbf{m}_i^\top = (u_i'\mathbf{p}_i^\top, v_i'\mathbf{p}_i^\top, \mathbf{p}_i^\top)_{1\times 9}$ $\mathbf{X} = \begin{bmatrix}0, 0, a_{11}, 0, 0, a_{12}, -1, 0, a_{13} \\ 0, 0, a_{21}, 0, 0, a_{22}, 0, -1, a_{23}\end{bmatrix}^\top$ $\mathbf{M}^\top\mathbf{X} = \mathbf{0}_{N\times 2}, G = (9,7,2,5)$

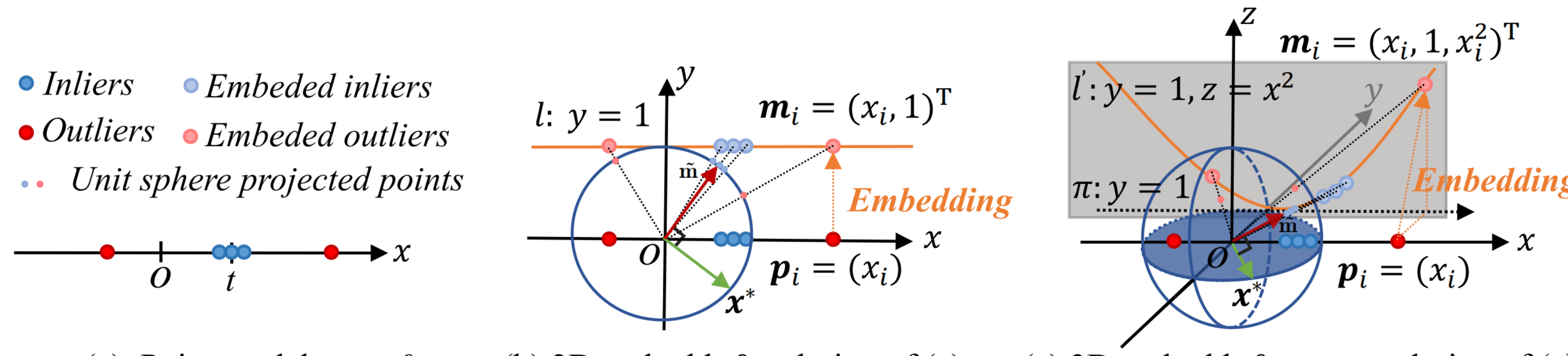
Method

▷**Problem Formulation:** Given a set of feature correspondences $\mathcal{S} = \{\mathbf{s}_i = (\mathbf{p}_i, \mathbf{p}_i')\}_{i=1}^N$, where $\mathbf{p}_i = (u_i, v_i, 1)^\top$ and $\mathbf{p}_i' = (u_i', v_i', 1)^\top$ are keypoints extracted from two-view images. Our goal is to recover the underlying geometric structure including *Fund. matrix*, *Homo.* and *Affine*. This two-view problem is essential in 3D vision, and is the main focus of this paper.

- **Existing Formulation:** DLT via hyperplane fitting:

$$\min_{\theta} \|\mathbf{M}^\top\theta\|_1, \quad s.t. \quad \|\theta\|_2 = 1, \tag{1}$$

where $\mathbf{M}$ is the data embedding for model $\mathcal{M}$, *i.e.*, $\mathbf{m}_i = \Phi_{\mathcal{M}}(\mathbf{p}_i, \mathbf{p}_i')$. $\theta = vec(\mathcal{M}) \in \mathbb{R}^D$ indicates the parameter vector form of $\mathcal{M}$.



(a) Point model: x-t =0 (b) 2D embedd. & solution of (a) (c) 3D embedd. & sparse solution of (a)

Illustration of Sparse Subspace Recovery (SSR) theory. $\mathbf{x}^*$ or $-\mathbf{x}^*$ denotes optimal solution.

- **Our Dual Sparsity Formulation- SSR theory:** formulated as the intersection of multiple sparse hyperplanes under an over embedded data space.

$$\widetilde{\Phi}(\mathbf{p}_i, \mathbf{p}_i')^\top \Psi(\mathcal{M}) = \mathbf{0}, \quad \mathcal{M} \in \{\mathbf{F}, \mathbf{H}, \mathbf{A}\}, \tag{2}$$

where $\mathbf{M} = \{\mathbf{m}_i = \widetilde{\Phi}(\mathbf{p}_i, \mathbf{p}_i')\}_{i=1}^N$ is a common embedding of the model pool, thus avoiding exact $\Phi_{\mathcal{M}}(\cdot)$ for each model. And the solution $X = \Psi(\mathcal{M})$ would directly reason out the model type.

- **Formulation for Unknown Model Fitting with Noises G & Outliers E**

$$\min_{\mathbf{X},\mathbf{E},r} \frac{1}{2}\|\mathbf{G}\|_F^2 + \lambda\|\mathbf{X}\|_0 + \gamma\|\mathbf{E}\|_{2,0} - \tau \mathrm{rank}(\mathbf{X}),$$
$$s.t. \;\; \mathbf{M}^\top\mathbf{X} - \mathbf{G} - \mathbf{E} = \mathbf{0}, \quad \mathbf{X}^\top\mathbf{X} = \mathbf{I}_{r\times r}, \tag{3}$$

▷**Solution:** Convex approximate, then progressively estimate a new sparse basis orthometric to all given bases up to $\mathcal{L}(\mathbf{M}, \hat{\mathbf{x}}_i, \hat{\mathbf{e}}_i) < \tau$ not holds. For any given bases $\{\mathbf{B} = (\mathbf{x}_j)|j = 1, 2, \cdots, i-1\}$ (those have been estimated), a new sparse basis $\mathbf{x}_i (i > 1)$ can be estimated by solving:

$$\min_{\mathbf{x},\mathbf{e}} \mathcal{L}(\mathbf{M}, \mathbf{x}, \mathbf{e}) = \frac{1}{2}\|\mathbf{M}^\top\mathbf{x} - \mathbf{e}\|_2^2 + \lambda\|\mathbf{x}\|_1 + \gamma\|\mathbf{e}\|_1,$$
$$s.t. \;\; \|\mathbf{x}\|_2 = 1, \; \mathbf{x}^\top\mathbf{y} = 0, \; \forall \mathbf{y} \in \mathbf{B}. \tag{4}$$

- **Given $\mathbf{x}^{k-1}$ Update $\mathbf{e}^k$:**

$$\mathbf{e}^k = \mathbf{M}^\top\mathbf{x}^{k-1} - \gamma \mathrm{sgn}(\mathbf{e}^k) = \mathcal{T}_\gamma(\mathbf{M}^\top\mathbf{x}^{k-1}), \tag{5}$$

where $\mathcal{T}_\gamma$ is the standard threshold shrinkage operation.

- **Given $\mathbf{e}^k$ Update $\mathbf{x}^k$:**

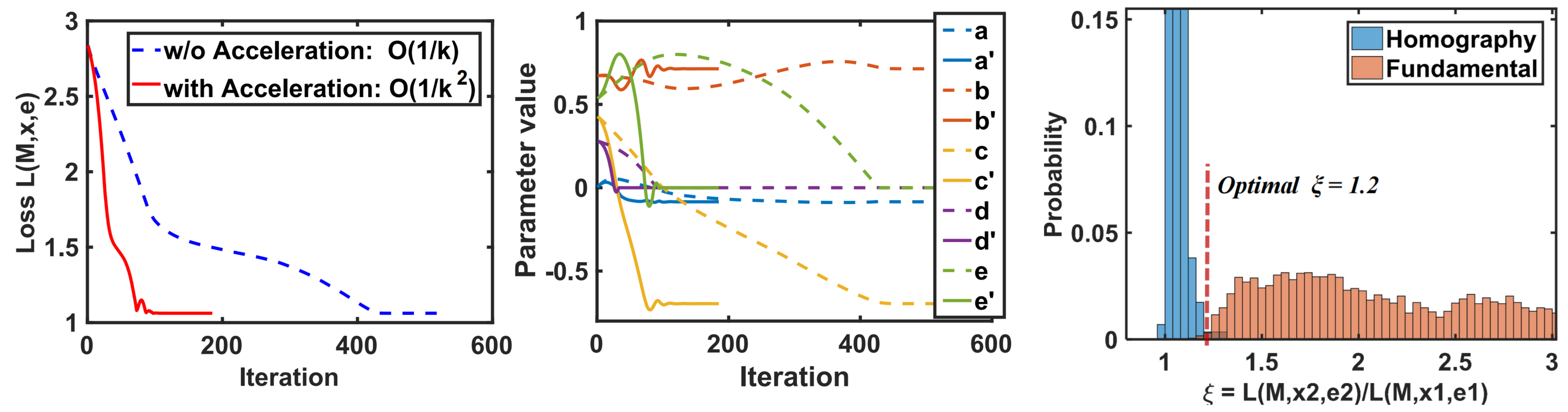
$$\mathbf{x}^k = \arg\min_{\mathbf{x}} \left\{\frac{L}{2}\|\mathbf{x} - (\mathbf{x}^{k-1} - \frac{1}{L}\nabla f(\mathbf{x}^{k-1}))\|_2^2 + \lambda\|\mathbf{x}\|_1\right\},$$
$$\mathbf{x}^k = \mathcal{T}_{\frac{\lambda}{L}}(\mathbf{q}_L(\mathbf{x}^{k-1})), \;\; \mathbf{q}_L(\mathbf{x}^{k-1}) = \mathbf{x}^{k-1} - \frac{1}{L}\nabla f(\mathbf{x}^{k-1}). \tag{6}$$

- **Processing the constraints:** using sphere projection $\mathbf{x} \leftarrow \frac{\mathbf{x}}{\|\mathbf{x}\|_2}$, and orthogonal projection $\mathbf{x} \leftarrow (\mathbf{I} - \mathbf{B}\mathbf{B}^\top)\mathbf{x}$.

Experiments & Results

Mechanism Illustration of Our DSP

Optimization of DSP for 2D Line Fitting: $\mathbf{x} = [a, b, c, d, e]^\top$ Margin of Homo. and Fund.

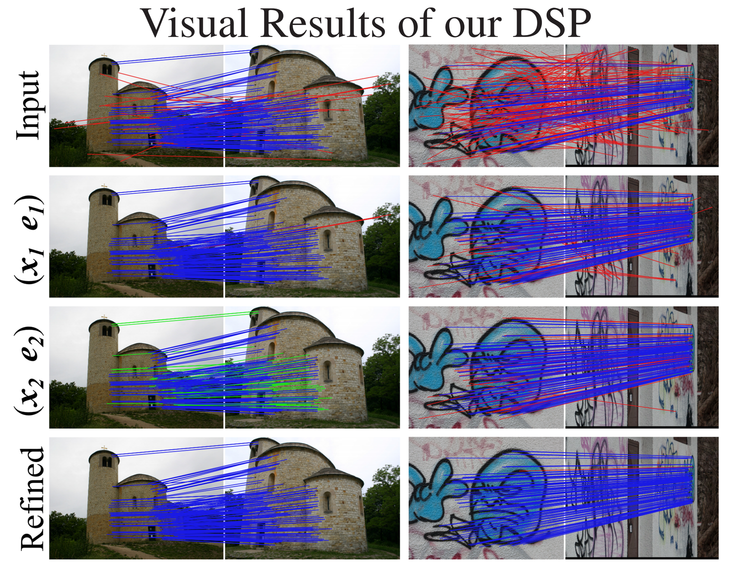


Unknown Model Fitting Test (Model Reasoning)

Reasoning Accuracy on Syn. Data

OR	Data / Model	AIC	BIC	GRIC	DSP (ours)
20%	F,100	100	62	100	100
	H,100	100	98	100	100
	A,100	100	96	100	100
50%	F,100	100	0	100	100
	H,100	100	95	100	100
	A,100	99	97	98	100
80%	F,100	85	0	93	100
	H,100	95	96	98	99
	A,100	96	92	95	98
	all	97.2	70.7	98.4	99.7

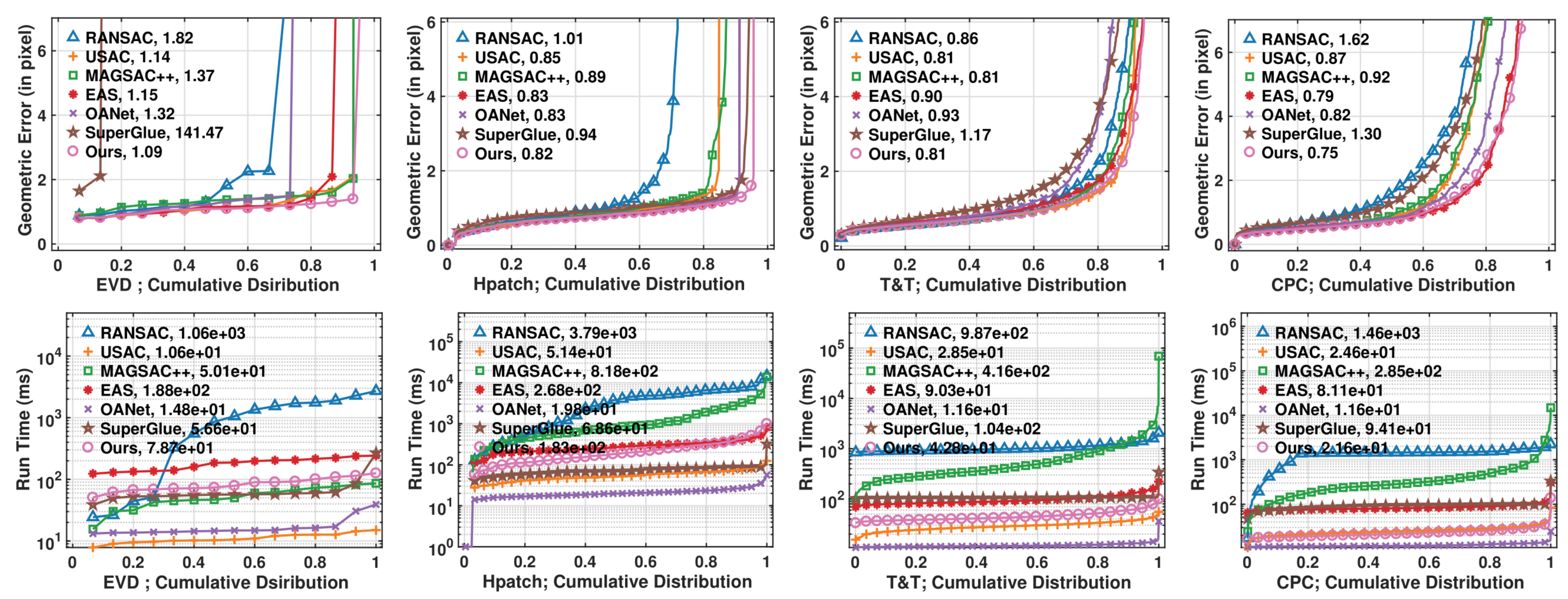
Visual Results of our DSP



Quantitative Results on Real Data

Data\Methods		RANSAC*	USAC*	MAGSAC++*	EAS*	OANet*	SuperGlue*	DSP (ours)
Fund.	*E_med*	0.8478	0.6545	0.6260	0.6900	0.7579	0.8926	**0.6017**
	Time(s)	0.5666	0.0239	0.3918	0.2037	**0.0152**	0.0721	0.0551
	FR	0.2925	0.1992	0.2077	0.2142	0.2632	0.2709	**0.1136**
Homo.	*E_med*	1.0428	0.8744	0.8978	0.8472	0.8480	0.9392	**0.8227**
	Time(s)	2.010	0.0550	1.3419	0.4463	**0.0342**	0.0716	0.2794
	FR	0.3021	0.2604	0.1424	0.0972	0.0903	0.1181	**0.066**
All	*E_med*	0.8794	0.6711	0.6383	0.7091	0.7670	0.9026	**0.6249**
	Time(s)	0.6638	0.0260	0.4558	0.22	**0.0130**	0.0721	0.0703
	FR	0.2932	0.2043	0.2033	0.2064	0.2515	0.2606	**0.1123**

Known Model Fitting Test:



Visual Application:

Multi-model Fitting

	Tlink	RCMSA	RPA	MCT	MLink	RFM-SCAN	Ours
ME_ave	27.65	10.05	5.28	11.36	6.04	2.63	**1.64**
ME_med	27.88	6.087	4.35	1.21	4.22	1.20	**0**
RT(s)	1.95	2.16	10.84	6.44	9.72	**0.01**	0.18

Loop Closure Detection

	RSAC	USAC	MSAC++	EAS	OANet	Ours
K00	0.9112	0.9118	0.9105	0.9086	0.7910	**0.9162**
K02	0.7632	0.7796	0.7757	0.7664	0.6441	**0.7882**

Acknowledgments:

Funding in direct support of this work: NSFC grant 62276192.

References:

[1] Survey [Ma *et al.*, IJCV'21]
[2] RANSAC [Fischler *et al.*, CACM'81]
[3] MAGSAC++ [Barath *et al.*, TPAMI'22]
[4] DPCP [Ding *et al.*, ICML'21]
[5] EAS [Fan *et al.*, TPAMI'22]
[6] GRIC [Philip HS Torr., MPES'98]
[7] TP19 [Probst *et al.*, CVPR'19]
[8] SuperGlue [Sarlin *et al.*, CVPR'20]