# A robust inlier identification algorithm for point cloud registration via $\ell_0$-minimization

Yinuo Jiang[1,*] Xiuchuan Tang[2,*] Cheng Cheng[1] Ye Yuan[1,†](email: yye@hust.edu.cn)
[1]Huazhong University of Science and Technology [2] Tsinghua University

NEURAL INFORMATION PROCESSING SYSTEMS

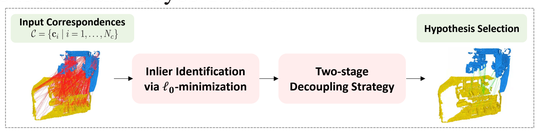
## Motivation and contributions

**Background:** Point cloud registration is a fundamental task in vision and robotics, e.g., 3D reconstruction and autonomous driving.

**Motivation:** Correspondences in point cloud registration are prone to outliers, significantly reducing registration accuracy and highlighting the need for precise inlier identification.

**Key Contributions:**

- A novel robust inlier identification algorithm is proposed by reformulating the conventional registration as an alignment error $\ell_0$-minimization problem.
- A two-stage decoupling strategy is designed to effectively solve the proposed $\ell_0$-minimization problem based on the Bayes Theorem.




## Problem formulation

**Main idea:** We reformulate the conventional registration as an alignment error $\ell_0$-minimization problem

- The conventional registration problem:

$$\min_{\mathbf{R},\mathbf{t}} \sum_{(\mathbf{p}_i,\mathbf{q}_i)\in\mathcal{C}} \|\mathbf{q}_i - \mathbf{R}\mathbf{p}_i - \mathbf{t}\|_2^2 , \tag{1}$$

where $\mathcal{C} = \{\mathbf{c}_i \mid i = 1, \ldots, N_c\}$ is the initial correspondence set.

- Considering only inliers can be fitted by the same transformation, the optimal transformation is estimated as the one that fits the largest number of inliers. Our optimization objective is to maximize the count of zero vectors in the alignment error via $\ell_0$ norm:

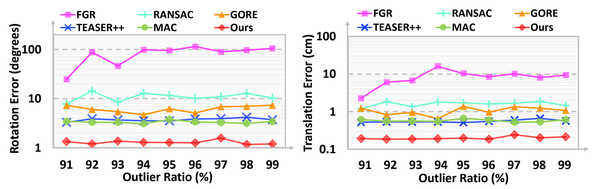
$$\begin{aligned} \mathbf{O}^* &= \arg\min_{\mathbf{O}} \|\mathbf{O}\|_{\ell_0} , \\ &\text{subject to: } \mathbf{o}_i = \mathbf{q}_i - \mathbf{R}\mathbf{p}_i - \mathbf{t} - \xi_k , \end{aligned} \tag{2}$$

where $\mathbf{O} = [\mathbf{o}_1, \mathbf{o}_2, \ldots, \mathbf{o}_{N_c}]$ is the alignment error matrix and and $\xi_k$ is the Gaussian noise.

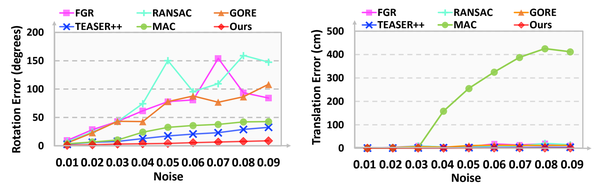
## Robustness to outliers and noise

We evaluate the robustness and accuracy of our algorithm using the Bunny point cloud from the Stanford 3D Scan Repository. To evaluate robustness against outliers, we increase the outlier ratio from 91% to 99%. For robustness to noise, we increase the noise standard deviation from $\sigma = 0.01$ to $\sigma = 0.09$.

**Robustness to outliers:**




**Robustness to noise:**




## Experiments

**Comparisons on the 3DMatch dataset:**

| | FPFH | | | FCGF | | | 3DSmoothNet | | | Time(s) |
|---|---|---|---|---|---|---|---|---|---|---|
| | RR(%)↑ | RE(°)↓ | TE(cm)↓ | RR(%)↑ | RE(°)↓ | TE(cm)↓ | RR(%)↑ | RE(°)↓ | TE(cm)↓ | |
| *i) Traditional* | | | | | | | | | | |
| FGR | 40.91 | 4.96 | 10.25 | 78.93 | 2.90 | 8.41 | 73.26 | 2.51 | 7.45 | 0.89 |
| RANSAC | 66.10 | 3.95 | 11.03 | 91.44 | 2.69 | 8.38 | 92.30 | 2.59 | 7.91 | 2.86 |
| TEASER++ | 75.48 | 2.48 | 7.31 | 85.71 | 2.73 | 8.66 | 92.05 | 2.23 | 6.62 | 0.03 |
| SC$^2$-PCR | 83.90 | 2.12 | 6.69 | 93.16 | 2.06 | 6.53 | 94.82 | 1.76 | 5.98 | 0.12 |
| MAC | 83.90 | **2.11** | 6.80 | **93.72** | **2.04** | 6.54 | 94.57 | 2.21 | 6.52 | 5.54 |
| TR-DE | - | - | - | - | - | - | 91.37 | 2.71 | 7.62 | - |
| TEAR | - | - | - | - | - | - | 94.52 | 2.06 | 6.55 | - |
| *ii) Deep learned* | | | | | | | | | | |
| DGR | 32.84 | 2.45 | 7.53 | 88.85 | 2.28 | 7.02 | - | - | - | 1.53 |
| PointDSC | 72.95 | 2.18 | **6.45** | 91.87 | 2.10 | 6.54 | 93.65 | 2.17 | 6.75 | 0.10 |
| VBReg | 82.57 | 2.14 | 6.77 | 93.53 | **2.04** | 6.49 | 37.09 | 6.15 | 15.65 | 0.20 |
| Ours | **83.92** | 2.12 | 6.64 | 93.28 | **2.04** | **6.48** | **95.07** | **1.75** | **5.97** | 0.36 |

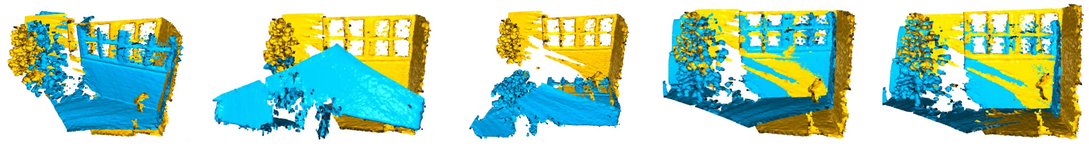
**Comparisons on the KITTI dataset:**

| | FPFH | | | FCGF | | | Time(s) |
|---|---|---|---|---|---|---|---|
| | RR(%)↑ | RE(°)↓ | TE(cm)↓ | RR(%)↑ | RE(°)↓ | TE(cm)↓ | |
| *i) Traditional* | | | | | | | |
| FGR | 5.23 | 0.86 | 43.84 | 89.54 | 0.46 | 25.72 | 3.88 |
| RANSAC | 74.41 | 1.55 | 30.20 | 80.36 | 0.73 | 26.79 | 5.43 |
| TEASER++ | 91.17 | 1.03 | 17.98 | 95.51 | 0.33 | 22.38 | 0.03 |
| SC$^2$-PCR | 99.46 | 0.35 | 7.87 | 98.02 | 0.33 | 20.69 | 0.31 |
| MAC | 97.66 | 0.41 | 8.61 | 97.84 | 0.34 | 19.34 | 3.29 |
| TR-DE | 96.76 | 0.90 | 15.63 | **98.20** | 0.38 | **18.00** | - |
| TEAR | 99.10 | 0.39 | 8.62 | - | - | - | - |
| *ii) Deep learned* | | | | | | | |
| DGR | 77.12 | 1.64 | 33.10 | 96.90 | 0.34 | 21.70 | 2.29 |
| PointDSC | 98.92 | 0.38 | 8.35 | 97.84 | 0.33 | 20.32 | 0.45 |
| VBReg | 98.92 | 0.45 | 8.41 | 98.02 | 0.32 | 20.91 | 0.24 |
| Ours | **99.56** | **0.34** | **7.85** | **98.20** | **0.32** | 20.73 | 0.54 |

**Comparisons on the 3DLoMatch dataset:**

| | 5000 | 2500 | 1000 | 500 | 250 |
|---|---|---|---|---|---|
| | Predator | | | | |
| FGR | 36.4 | 38.2 | 39.7 | 39.6 | 38.0 |
| RANSAC | 62.3 | 62.8 | 62.4 | 61.5 | 58.2 |
| TEASER++ | 62.9 | 62.6 | 61.9 | 59.0 | 56.7 |
| SC$^2$-PCR | 68.9 | 68.4 | 68.7 | 67.1 | 64.9 |
| MAC | 69.4 | 69.3 | 68.4 | **67.7** | 64.6 |
| TR-DE | 64.0 | 64.8 | 61.7 | 58.8 | 56.5 |
| PointDSC | 68.1 | 67.3 | 66.5 | 63.4 | 60.5 |
| VBReg | **69.9** | 69.8 | 68.7 | 66.4 | 63.0 |
| Ours | **69.9** | **69.9** | **69.2** | **67.7** | **65.0** |

**Qualitative results on the 3DMatch dataset:**

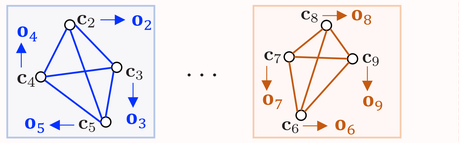


(a) Input (b) MAC (c) SC$^2$-PCR (d) Ours (e) Ground-truth

## Algorithm

### Inlier Identification via $\ell_0$-minimization

Define alignment error $\mathbf{o}_i$ for correspondence $\mathbf{c}_i$ in each local set




$\ell_0$-minimization problem

$\min \|[\mathbf{o}_2, \mathbf{o}_3, \mathbf{o}_4, \mathbf{o}_5]\|_{\ell_0}$ $\min \|[\mathbf{o}_6, \mathbf{o}_7, \mathbf{o}_8, \mathbf{o}_9]\|_{\ell_0}$

### Two-stage Decoupling Strategy

**Stage 1**: Decouple alignment error into **rotation** fitting error $\bar{\mathbf{o}}_{ij}$

$\bar{\mathbf{o}}_{23}$ $\bar{\mathbf{o}}_{25}$ ⋯ $\bar{\mathbf{o}}_{79}$ ⇒ **Stage 2: Bayesian-based** inlier identification ⇒ ⋯ ⇒ $\mathbf{R}_1^*$ $\mathbf{R}_2^*$ ⋯

**Stage 1**: Decouple alignment error into **translation** fitting error $\hat{\mathbf{o}}_i$

$\hat{\mathbf{o}}_2$ $\hat{\mathbf{o}}_7$ ⋯ ⇒ **Stage 2: Bayesian-based** inlier identification ⇒ ⋯ ⇒ $\mathbf{t}_1^*$ $\mathbf{t}_2^*$ ⋯

### Inlier Identification via $\ell_0$-minimization

We identify compatible correspondences to form local sets. For correspondences in the $k$-th local set, the $\ell_0$-minimization problem is defined as:

$$\begin{aligned} \mathbf{O}_k^* &= \arg\min_{\mathbf{O}_k} \|\mathbf{O}_k\|_{\ell_0} , \\ &\text{subject to: } \mathbf{O}_k = \mathbf{Q}_k - \mathbf{P}_k\mathbf{R}_k - \mathbf{t}_k\mathbf{1}^T - \mathbf{\Xi}_k . \end{aligned} \tag{3}$$

### Two-stage Decoupling Strategy

- **Stage 1:** Decouple alignment error into a rotation fitting error and a translation fitting error by calculating relative positions between points, resulting in two fitting error $\ell_0$-minimization problems for rotation and translation.
- **Stage 2:** We introduce null-space matrices to isolate rotation and translation in the $\ell_0$-minimization constraints, where $\bar{\mathbf{\Theta}}_k\bar{\mathbf{P}}_k = \mathbf{0}$ and $\mathbf{\Theta}_k\mathbf{1} = \mathbf{0}$.

Based on the Bayes Theorem and Maximum A Posteriori (MAP) estimate, the unconstrained optimization problems for rotation and translation fitting are formulated as:

$$\min_{\bar{\mathbf{O}}_k} \frac{1}{2} \left\| (\tilde{\bar{\mathbf{Q}}}_k - \bar{\mathbf{\Theta}}\bar{\mathbf{O}}_k)^T \bar{\mathbf{\Pi}}_k^{-1} (\tilde{\bar{\mathbf{Q}}}_k - \bar{\mathbf{\Theta}}_k\bar{\mathbf{O}}_k) \right\|_F^2 + \lambda_R \|\bar{\mathbf{O}}_k\|_{\ell_0}^2 , \tag{4}$$

where $\tilde{\bar{\mathbf{Q}}}_k = \bar{\mathbf{\Theta}}_k\bar{\mathbf{Q}}_k$ and $\bar{\mathbf{\Theta}}_k\bar{\mathbf{\Theta}}_k^T = \bar{\mathbf{\Pi}}_k$.

$$\min_{\hat{\mathbf{O}}_k} \frac{1}{2} \left\| (\mathbf{X}_k - \mathbf{\Theta}_k\hat{\mathbf{O}}_k^T)^T \mathbf{\Pi}_k^{-1} (\mathbf{X}_k - \mathbf{\Theta}_k\hat{\mathbf{O}}_k^T) \right\|_F^2 + \lambda_t \left\| \hat{\mathbf{O}}_k \right\|_{\ell_0}^2 , \tag{5}$$

where $\mathbf{X}_k = \mathbf{\Theta}_k(\mathbf{Q}_k^T - (\mathbf{P}_k\mathbf{R}_k^*)^T)$ and $\mathbf{\Pi}_k = \mathbf{\Theta}_k\mathbf{\Theta}_k^T$.

Rotation fitting error $\bar{\mathbf{O}}_k^*$ and translation fitting error$\hat{\mathbf{O}}_k^*$ can be directly computed due to convex relaxation via the Frobenius norm:

$$\begin{aligned} \bar{\mathbf{O}}_k^* &= (\bar{\mathbf{\Theta}}_k^T\bar{\mathbf{\Pi}}_k^{-1}\bar{\mathbf{\Theta}}_k + 2\lambda_R\mathbf{I})_k^{-1}\bar{\mathbf{\Theta}}_k^T\bar{\mathbf{\Pi}}_k^{-1}\tilde{\bar{\mathbf{Q}}}_k . \\ \hat{\mathbf{O}}_k^* &= ((2\lambda_t\mathbf{I} + \mathbf{\Theta}_k^T\mathbf{\Pi}_k^{-1}\mathbf{\Theta}_k)^{-1}\mathbf{\Theta}_k^T\mathbf{\Pi}_k^{-1}\mathbf{X}_k)^T . \end{aligned} \tag{6}$$