Unsupervised Learning for Multi-Model Consensus Maximization

A novel Deep Learning approach for multi-model fitting problems

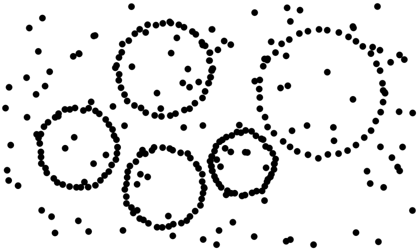
William Bonvini

April 14, 2021

The Multi-Model Fitting Problem

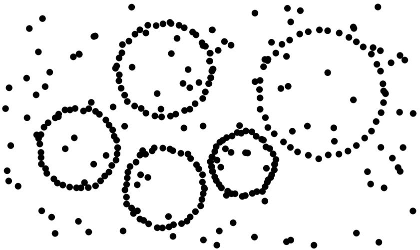
Multi-Model Fitting: Problem Definition

Given a set of data $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$ possibly corrupted by noise and outliers, and a family of geometric models θ,



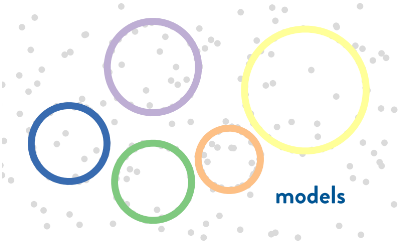
The Multi-Model Fitting Problem

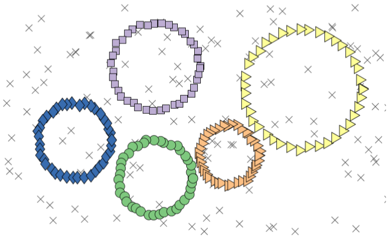
Given a set of data $\mathcal{X} = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ possibly corrupted by noise and outliers, and a family of geometric models θ, automatically estimate the models that best explain the data.



The Multi-Model Fitting Problem

Given a set of data $\mathcal{X} = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ possibly corrupted by noise and outliers, and a family of geometric models θ, automatically estimate the models that best explain the data. Do so by retrieving structures hidden in data.





Applications

Applications: Scan2Bim

Given a scanned point cloud of an interior environment, detect its primary facility surfaces – such as floors, walls, and ceilings.

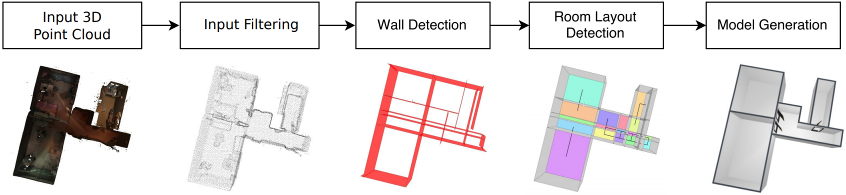


Figure: Plant generation with Wall Detection

$$\mathcal{X} \subset \mathbb{R}^3, \Theta = \text{planes}$$

Applications: Two View Geometry

Figure: $\mathcal{X} \in \mathbb{R}^4, \Theta =$ homographies

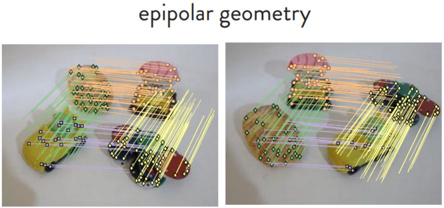
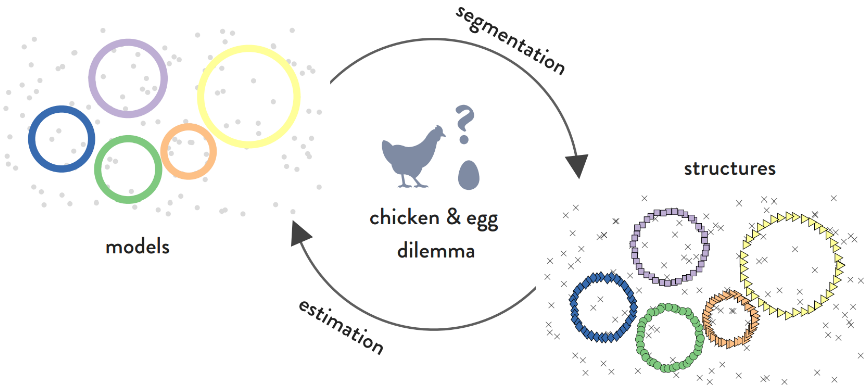


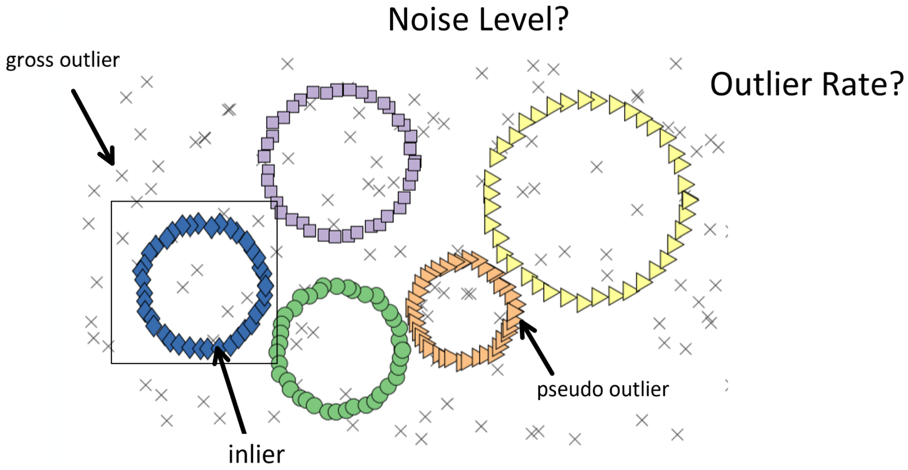
Figure: $\mathcal{X} \in \mathbb{R}^4, \Theta =$ fundamental matrices

Challenges

Multi-Model Fitting Challenges



Multi-Model Fitting Challenges



Thesis Objective

Thesis Objective

Train a neural network to perform robust Multi-Model Fitting in an unsupervised fashion.

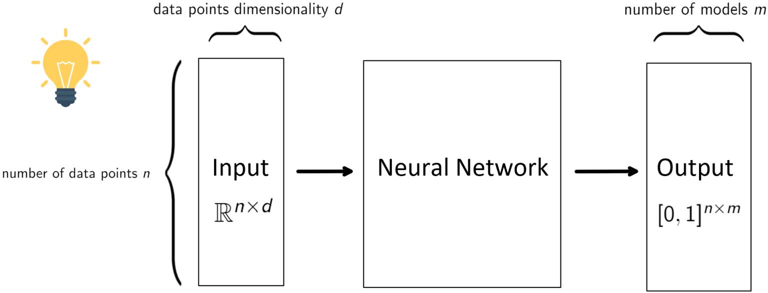


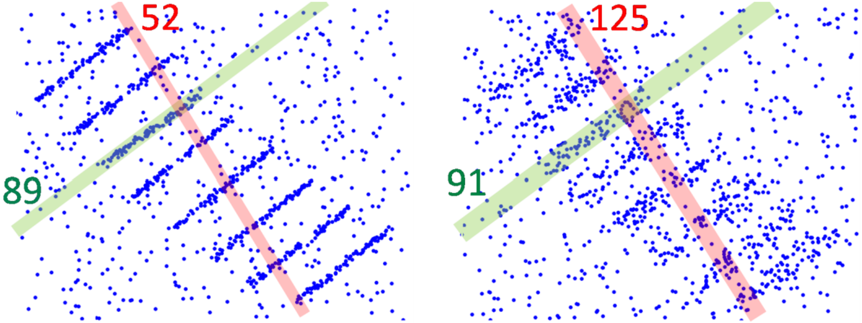
Figure: High Level Neural Network View

Motivations

	RANSAC heuristics	Global methods	Learned Unsupervised
High Outliers	✗	✓	✓
Speed	✓	✗	✓
Generalization	✓	✓	✓
Differentiable	✗	✗	✓
Non-greedy	✗	✓	✓
methods	Vincent and Laganière, 2001 Zuliani et al., 2005	Isack et al., 2012	Ours

- Our approach takes inspiration from the work of Probst et al. (2019), that has proven successful in the single model scenario.
- No end-to-end Deep Learning solutions proposed so far for multi-model fitting.

RANSAC-based heuristics: Greediness Problem



(a) low noise (b) high noise

Results

Experiments have shown that my approach

- exploits the global signature of data $\rightarrow$ non-greedy
- is able to adapt to different noise levels
- outperforms Seq-RANSAC at
  - high outliers rate in the multiple homographies estimation task
  - high noise contamination levels

Mathematical Formulation

Multi-Model Consensus Maximization

Definition

Given a set $\mathcal{X} = \{(u_i, v_i) \subseteq \mathbb{R}^{d_u + d_v}, i = 1, \dots, n\}$ of corresponding measurements, find the largest m subsets $\Omega_j \subseteq \mathcal{X}$, each explained by a parametric transformation $\phi_j : \mathbb{R}^{d_u} \to \mathbb{R}^{d_v}$.

$$\max_{\{\Omega_1, \dots, \Omega_m\}} \sum_{j=1}^{m} |\Omega_j|$$

$$\text{s.t.} \qquad \Omega_j = \{(u_i, v_i) \in \mathcal{X} : \mathbf{d}\big(\phi_j(u_i), v_i\big) < \varepsilon\} \quad \forall j \in \{1, \dots, m\}$$

Multi-Model Consensus Maximization

Definition

Given a set $\mathcal{X} = \{(u_i, v_i) \subseteq \mathbb{R}^{d_u+d_v}, i = 1, \ldots, n\}$ of corresponding measurements, find the largest m subsets $\Omega_j \subseteq \mathcal{X}$, each explained by a parametric transformation $\phi_j : \mathbb{R}^{d_u} \rightarrow \mathbb{R}^{d_v}$.

$$\max_{\{\Omega_1,\ldots,\Omega_m\}} \sum_{j=1}^{m} |\Omega_j|$$

$$\text{s.t.} \qquad \Omega_j = \{(u_i, v_i) \in \mathcal{X} : \mathbf{d}\big(\phi_j(u_i), v_i\big) < \varepsilon\} \quad \forall j \in \{1, \ldots, m\}$$

How do we make it differentiable?

The Ideal $\mathcal{I}(\Omega)$

Definition

Being $\mathcal{X}$ a set of correspondences, **the ideal** $\mathcal{I}(\Omega)$ is the set of polynomials that vanishes on some samples $\Omega \in \mathcal{X}$.

$$\mathcal{I}(\Omega) := \{p(x) \in \mathcal{R}[x] : p(x) = 0, \quad \forall x \in \Omega\}$$

where

- $\mathcal{R}[x] := \mathbb{R}[x_1, \ldots, x_n]_d$ - Ring of multivariate polynomials of degree $\leq d$

The Ideal $\mathcal{I}(\Omega)$

Definition

Being $\mathcal{X}$ a set of correspondences, **the ideal** $\mathcal{I}(\Omega)$ is the set of polynomials that vanishes on some samples $\Omega \in \mathcal{X}$.

$$\mathcal{I}(\Omega) := \{p(x) \in \mathcal{R}[x] : p(x) = 0, \quad \forall x \in \Omega\}$$

where

- $\mathcal{R}[x] := \mathbb{R}[x_1, \ldots, x_n]_d$ - Ring of multivariate polynomials of degree $\leq d$

The ideal $\mathcal{I}(\Omega)$ contains infinite solutions.

The space of valid polynomials $\mathcal{R}_{\mathcal{B}}$

We consider only the set of polynomials $\mathcal{R}_{\mathcal{B}}$ spanned by a known basis $\mathcal{B}$.

Example

if

$$\mathcal{B} = \left\{x^2 \quad xy \quad y^2 \quad x \quad y \quad 1\right\}$$

then

$$\mathcal{R}_{\mathcal{B}} := \{[a_1, a_2, a_3, a_4, a_5, a_6] \cdot \mathcal{B}^T, a \in \mathbb{R}^6\}$$

The polynomial representation of ϕ

The polynomial representation of a model ϕ involves a known number r of linearly independent equations in $\mathcal{R}_{\mathcal{B}}$ that vanish for all $x \in \Omega$.
Therefore we can state that

$$\dim(\mathcal{I}(\Omega) \cap \mathcal{R}_{\mathcal{B}}) = r$$

The polynomial representation of ϕ

The polynomial representation of a model ϕ involves a known number r of linearly independent equations in $\mathcal{R}_{\mathcal{B}}$ that vanish for all $x \in \Omega$.
Therefore we can state that

$$\dim(\mathcal{I}(\Omega) \cap \mathcal{R}_{\mathcal{B}}) = r$$

How can we exploit this result?

Vandermonde Matrix

Definition

$M_d(\Omega) \in \mathbb{R}^{n \times t}$ is a matrix with the terms of a geometric progression's monomials with degree at most d in each row.

Example

- $\Omega = \{(x_i, y_i)\}_{i=0}^{n}$
- $d = 2$
- $n = 3$

$$M_2(\Omega) = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 y_3 & y_3^2 & x_3 & y_3 & 1 \end{bmatrix}$$

Vandermonde Matrix - Property

Definition

The **kernel** of a matrix A are all the solutions to the linear system $A\vec{x} = 0$.

Theorem

The kernel $ker(M_d(\Omega))$ of the Vandermonde matrix $M_d(\Omega)$ equals to the vector space $\mathcal{I}(\Omega) \cap \mathcal{R}_{\mathcal{B}}$. i.e. all polynomials that are linear combinations of $\mathcal{B}$ and vanish on Ω are represented by:

$$ker(M_d(\Omega)) = \mathcal{I}(\Omega) \cap \mathcal{R}_{\mathcal{B}}$$

Reformulation of Multi-Model Consensus Maximization

Definition

$$\max_{\{\Omega_1,\ldots,\Omega_m\}} \sum_{j=1}^{m} |\Omega_j|$$

$$\text{s.t.} \qquad \dim(\ker(M_{\mathcal{B}}(\Omega_j))) = r \qquad \forall j \in \{1, \ldots, m\}$$

Reformulation of Multi-Model Consensus Maximization

Definition

$$\max_{\{\Omega_1,\ldots,\Omega_m\}} \sum_{j=1}^{m} |\Omega_j|$$
$$\text{s.t.} \qquad \dim(\ker(M_{\mathcal{B}}(\Omega_j))) = r \qquad \forall j \in \{1,\ldots,m\}$$

Data is possibly corrupted by noise.

Reformulation of Multi-Model Consensus Maximization

We relax the rank constraint by minimizing the trailing r singular values σ_k of $M_{\mathcal{B}}$:

Definition

$$\max_{\{\Omega_1,\ldots,\Omega_m\}} \sum_{j=1}^{m} \left(|\Omega_j| - \lambda \sum_{k=s-r}^{s} \sigma_k(M_{\mathcal{B}}(\Omega_j)) \right)$$

Reformulation of Multi-Model Consensus Maximization

We relax the rank constraint by minimizing the trailing r singular values σ_k:

Definition

$$\max_{\{\Omega_1,\ldots,\Omega_m\}} \sum_{j=1}^{m} \left(|\Omega_j| - \lambda \sum_{k=s-r}^{s} \sigma_k(M_{\mathcal{B}}(\Omega_j)) \right)$$

NP-hard combinatorial problem, not differentiable.

Reformulation of Multi-Model Consensus Maximization

Soft assignments $0 \leq w_{ij} \leq 1$:

Definition

$$\max_{[\mathbf{w}_1,\ldots,\mathbf{w}_m]\in[0,1]^{n\times m}} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \mathbf{w}_{i,j} - \lambda \sum_{k=s-r}^{s} \sigma_k(\mathrm{diag}(\mathbf{w}_j) M_{\mathcal{B}}(\mathcal{X})) \right)$$

Reformulation of Multi-Model Consensus Maximization

Soft assignments $0 \leq w_{ij} \leq 1$

Definition

$$\max_{[\mathbf{w}_1,\ldots,\mathbf{w}_m]\in[0,1]^{n\times m}} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \mathbf{w}_{i,j} - \lambda \sum_{k=s-r}^{s} \sigma_k(\mathrm{diag}(\mathbf{w}_j) M_{\mathcal{B}}(\mathcal{X})) \right)$$

The objective is now differentiable and can serve as an unsupervised loss.

Loss Function

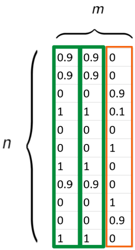
$$\mathcal{L}(\theta, \mathcal{X}) = \sum_{j}^{m} \left(-\lambda_{in} ||\mathbf{w}_j(\mathcal{X})|| + \lambda_{vander} \sum_{k=0}^{r-1} \sigma_{s-k}(\text{diag}(\mathbf{w}_j) M_{\mathcal{B}}(\mathcal{X}) \right)$$

- first term: maximize consensus
- second term: minimize algebraic error

Loss Function

Penalize column similarity

The neural network should not predict multiple times the same model.



Add term to penalize similarity between columns of $\mathbf{W}_\theta$.

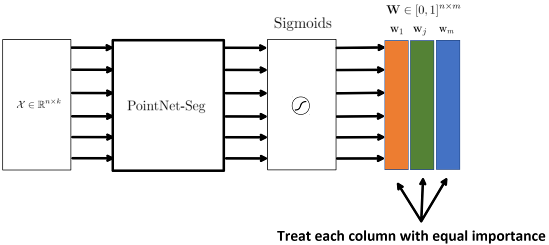
$$\log(1 + ||\hat{\mathbf{W}}_\theta^T \hat{\mathbf{W}}_\theta - I||_2)$$

with

$$\hat{\mathbf{w}}_{\theta,j} = \frac{\mathbf{w}_{\theta,j}}{||\mathbf{w}_{\theta,j}||_2} \forall j \in \{1, \ldots, m\}$$

Loss Function

Avoid zero-columns



This term helps the network to treat equally importantly each model

$$\frac{\sum_j^m (||\mathbf{w}_{\theta,j}(\mathcal{X})||_1 - w_{avg})^2}{m}$$

with w_{avg} being defined as

$$w_{avg} = \frac{\sum_j^m ||\mathbf{w}_{\theta,j}(\mathcal{X})||_1}{m}$$

.

Loss Function

Complete formulation

$$\mathcal{L}(\theta, \mathcal{X}) = \sum_{j}^{m} \left(-\lambda_{in} ||\mathbf{w}_j(\mathcal{X})||_1 + \lambda_{vander} \sum_{k=0}^{r-1} \sigma_{s-k} \big(\text{diag}(\mathbf{w}_j) M_{\mathcal{B}}(\mathcal{X}) \big) \right) +$$

$$\lambda_{sim} \log(1 + ||\hat{\mathbf{W}}_\theta^T \hat{\mathbf{W}}_\theta - I||_2) +$$

$$\lambda_{var} \frac{\sum_j^m (||\mathbf{w}_{\theta,j}(\mathcal{X})||_1 - w_{avg})^2}{m}$$

Architecture

Complete Architecture

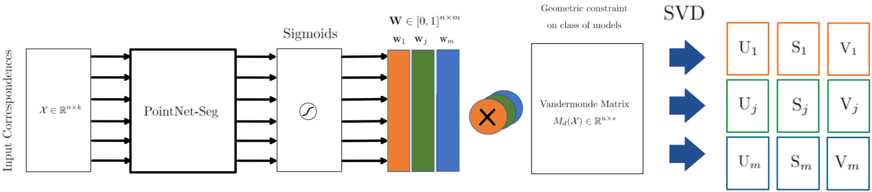


Figure: **Unsupervised Learning for Multi-Model Consensus Maximization**

Architecture

PointNet-Seg

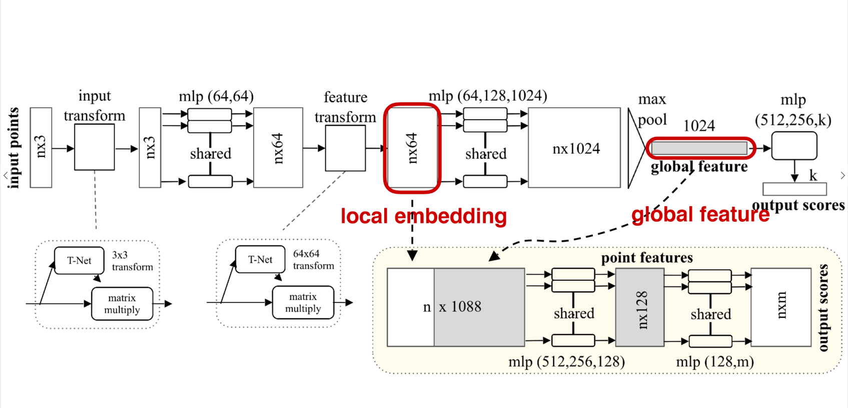


Figure: **PointNet Segmentation Network**

Results

Performance Metrics

Performance was assessed with Accuracy; Inliers Detection Rate; Outliers Detection Rate; F1-Score; and Geometric distance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{IDR} = \text{Recall} = \frac{TP}{TP + FN}$$

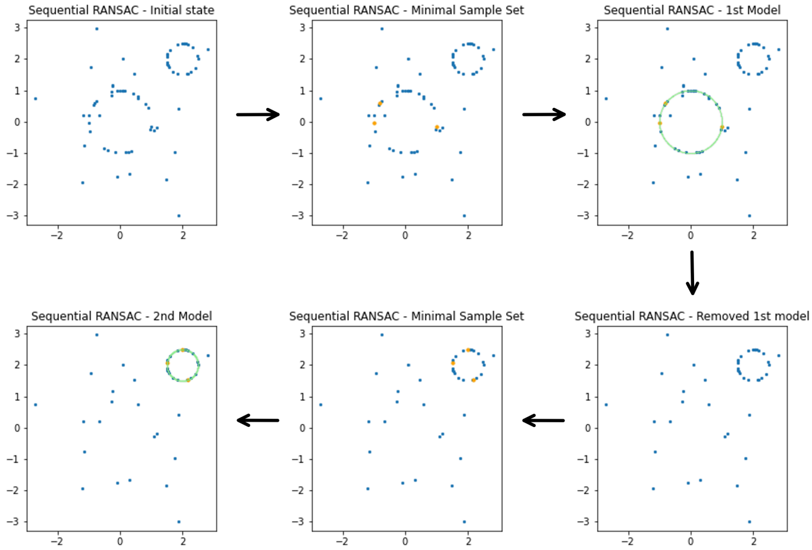
$$\text{ODR} = \frac{TN}{TN + FP}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Geometric Distance} = \sum_{\mathbf{x}_i \in \mathcal{X}} \sqrt{||v_i - \hat{\theta}(u_i)||_2}$$

Competitor - Sequential RANSAC



Seq-RANSAC Algorithm

Algorithm 4 Sequential RANSAC

Input: $\mathcal{X}$ - a set of observations
Output: $\theta^* = \{\mathcal{M}_1^*, \ldots, \mathcal{M}_M^*\}$ - A set of models
Procedure:

```
mincost = ∞
θ* = ∅
m = 0
while m < M do                                M = number of models
    i = 0
    while i < k do                            k = number of iterations
        mss := n randomly selected values from 𝒳
        𝓜 := model fitted on mss
        Ω := {x_j ∈ 𝒳 | d(v_j, 𝓜(u_j))^2 < ε^2}
        cost := Σ_{x_j∈Ω} (d(v_j, 𝓜(u_j))^2) + Σ_{x_j∈𝒳\Ω} ε^2
        if cost < mincost then
            mincost = cost
            𝓜* = 𝓜
        end if
        i = i + 1
    end while                                 update best model
    θ* = θ* ∪ 𝓜*
    m = m + 1
    𝒳 = 𝒳 \ Ω
end while
return θ*
```

$$\Omega := \{x_j \in \mathcal{X} \mid d(v_j, \mathcal{M}(u_j))^2 < \varepsilon^2\}$$

$$\text{cost} := \sum_{x_j \in \Omega} (d(v_j, \mathcal{M}(u_j))^2) + \sum_{x_j \in \mathcal{X} \setminus \Omega} \varepsilon^2$$

Homography Estimation

Definition

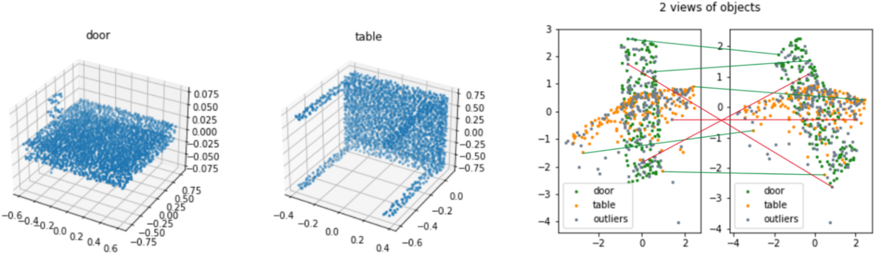
Homography. a *homography* is an invertible mapping h from $\mathbb{P}^2$ to itself such that three points $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_3$ lie on the same line if and only if $h(\mathbf{x}_1), h(\mathbf{x}_2)$ and $h(\mathbf{x}_3)$ do.

An equivalent algebraic definition of a homography is possible, based on the following result.

Theorem

A mapping $h : \mathbb{P}^2 \to \mathbb{P}^2$ *is a homography if and only if there exists a non-singular* 3×3 *matrix* H *such that for any point in* $\mathbb{P}^2$ *represented by a vector* $\mathbf{x}$ *it is true that* $h(\mathbf{x}) = H\mathbf{x}$.

Data Generation



(a) (b) (c)

Figure: **Example of pair of homography correspondences generated from Modelnet-40**. (a) and (b) show the 3D point clouds from which is obtained the first and the second homography respectively. (c) shows the generated matches.

Two Homographies fitting

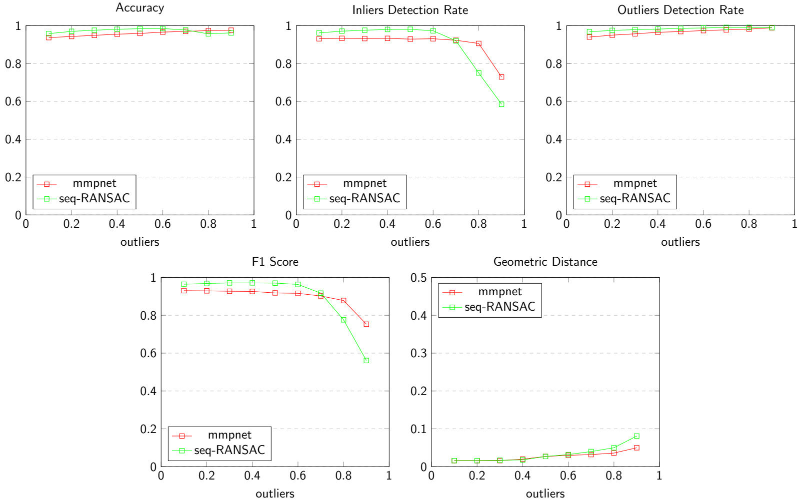


Figure: **Two Homographies estimation task with increasing outliers rate.**

Three Homographies Fitting

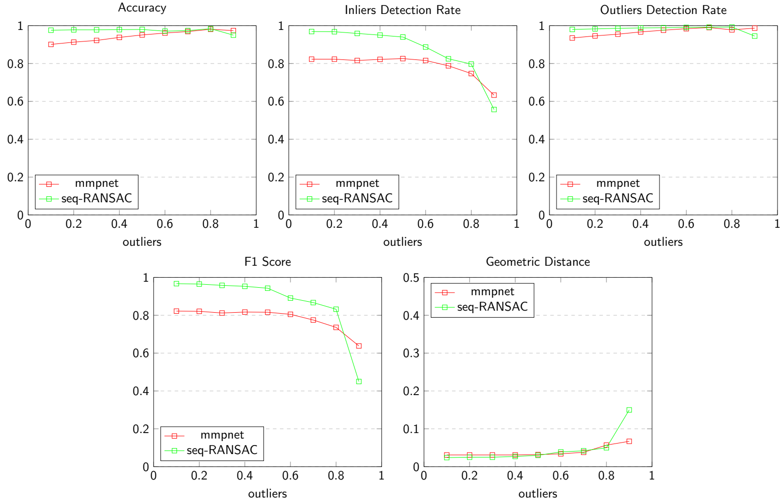


Figure: **Three Homographies estimation task with increasing outliers rate.**

Two Circles Fitting

A more challenging scenario...

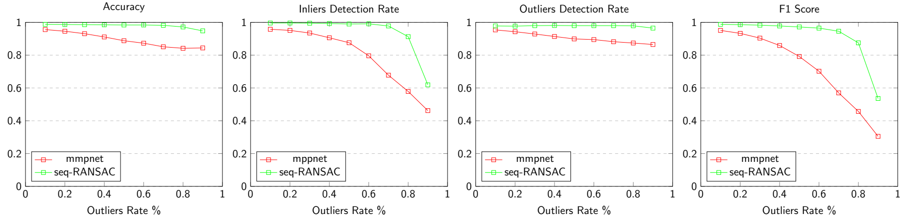


Figure: **Performance on 2 Circles fitting with increasing outliers rate**.

Two Circles Fitting

Model Estimation

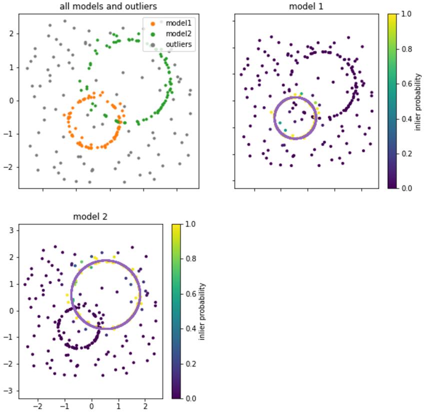


Figure: **Test sample with 60% outliers contamination**

Two Circles Fitting

Comparison

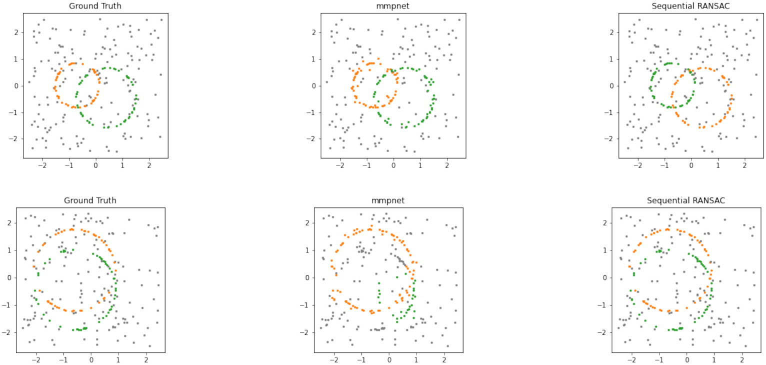


Figure: **Comparison among mmpnet and Sequential RANSAC**.
Top row: 60% outliers; bottom row: 70% outliers.

Two Circles Fitting - Increasing Noise Level

Performance

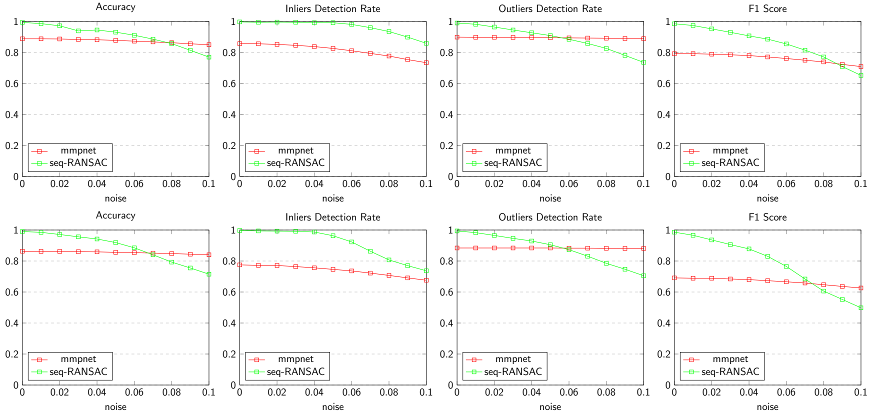


Figure: **Increasing level of noise**. Top row: 50% outliers; bottom row: 60% outliers

Two Circles Fitting - Increasing Noise Level

Comparison

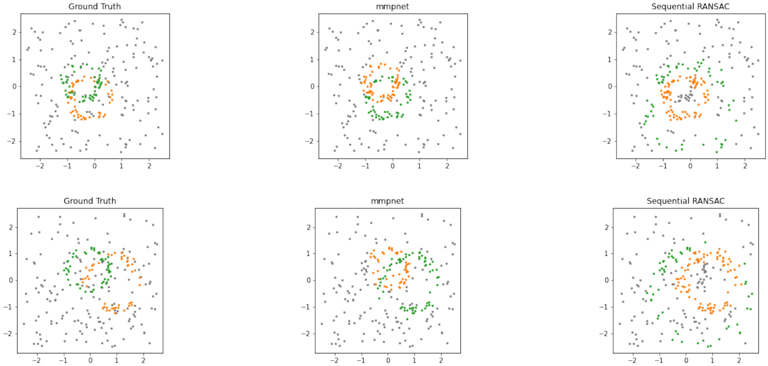


Figure: **Comparison among mmpnet and seq-ransac**. 60% outliers and 9% noise.

The Union Loss

Union Loss

Seek the inlier set $\Omega \in \mathcal{X}$ for which is maximized the consensus of a polynomial q representing the union of multiple models.

Union Loss

Seek the inlier set $\Omega \in \mathcal{X}$ for which is maximized the consensus of a polynomial q representing the union of multiple models.
The task reduces to a single model fitting problem.

Fitting the union of two lines in $\mathbb{R}^2$

Example

Data is sampled from two lines:

- $S_1 = \{\mathbf{x} : a_1 x_1 + b_1 x_2 + c_1 = 0\}$
- $S_2 = \{\mathbf{x} : a_2 x_1 + b_2 x_2 + c_2 = 0\}$

$$
\begin{aligned}
S_1 \cup S_2 &= \{\mathbf{x} : (a_1 x_1 + b_1 x_2 + c_1 = 0) \vee (a_2 x_1 + b_2 x_2 + c_2 = 0)\} \\
&= \{\mathbf{x} : c_1 x_1^2 + c_2 x_1 x_2 + c_3 x_1 + c_4 x_2^2 + c_5 x_2 + c_6 = 0\}
\end{aligned}
$$

Formulation of the union

Definition

The union of m homogeneous polynomials with coefficient vectors $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ is the set of points satisfying at least one element in $\mathbf{B}$

$$q_d(\mathbf{x}) = \prod_{j=1}^{m} (\mathbf{b}_j^T \mathbf{x}) = \mathbf{c}_d^T v_d(\mathbf{x}) = 0$$

where $\mathbf{v}_d(\mathbf{x})$ is the Veronese map of degree $d = \sum_{j=1}^{m} d_{\mathbf{b}_j}$.

Two Circles Fitting

Performance

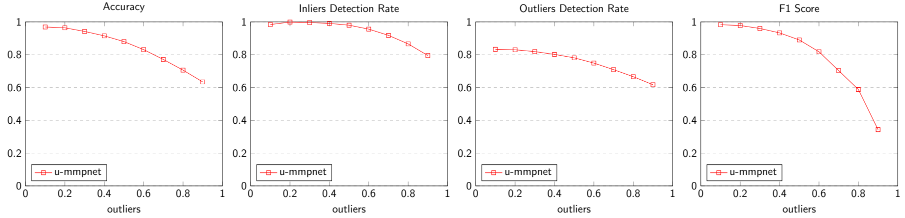
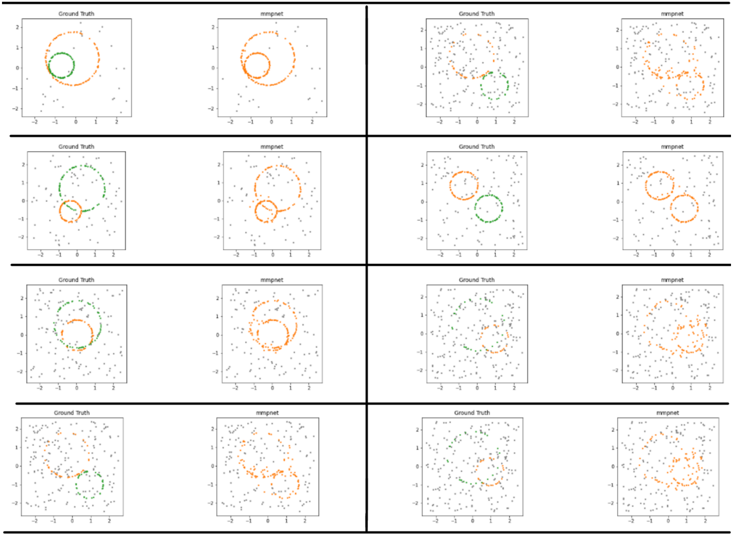


Figure: **Performance of u-mmpnet with increasing outliers rate**.

Two Circles Fitting

Test Samples at increasing outlier rates



Conclusions

Conclusions

Introduced two novel deep learning approaches for multi-model fitting problems.

- advantages
    - robustness at high noise contamination
    - outperform SEQ-RANSAC at high outliers rate
    - execution time independent of data distribution
    - no test-time hyperparameter tuning
- drawbacks
    - number of models m to be known a priori
    - difficult to train the network with $m \geq 4$ models.