Learning To Find Good Correspondences Of Multiple Objects

William Hoff

Colorado School of Mines
Golden, Colorado, USA

My co-authors:

Youye Xie, Gongguo Tang
Colorado School of Mines

Yingheng Tang
Purdue University

Code is available at: https://github.com/youyexie/Learning-To-Find-Good-Correspondences-Of-Multiple-Objects

Outline

- Introduction
- Proposed Method
  - 3D rotation space quantization
  - Learning-based Facet Network
  - Post-processing And Object Detection
- Experiments
  - Finding Correspondences Of One Object
  - Finding Correspondences Of Multiple Objects
  - Object Detection Performance
  - Performance on GMU Kitchen Dataset
- Conclusion

Outline

- Introduction

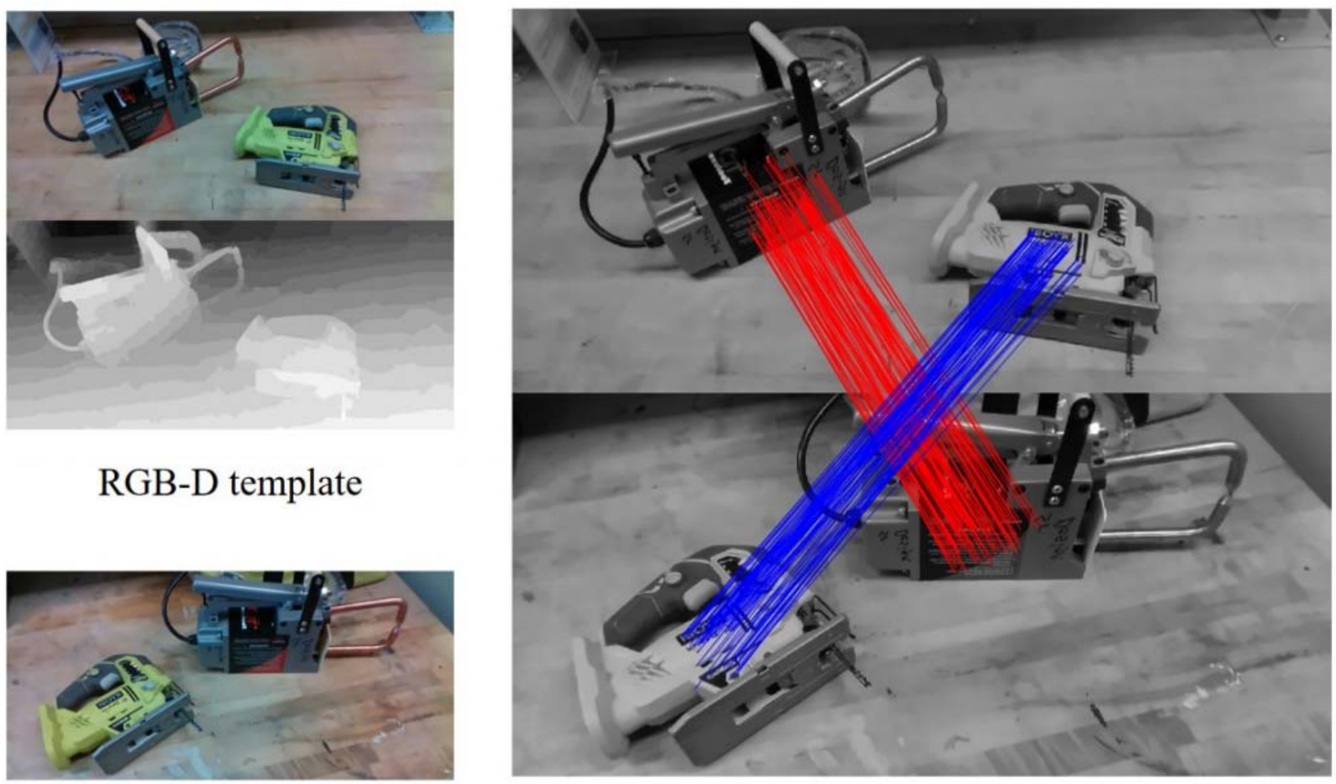
- Proposed Method
  - 3D rotation space quantization
  - Learning-based Facet Network
  - Post-processing And Object Detection

- Experiments
  - Finding Correspondences Of One Object
  - Finding Correspondences Of Multiple Objects
  - Object Detection Performance
  - Performance on GMU Kitchen Dataset

- Conclusion

Introduction

- We address the problem of finding correspondences of multiple objects from a set of 3D to 2D putative matches.
- We want to label the correspondences as inliers or outliers and classify the inliers into multiple objects.



RGB-D template

Test image

Our method output

Introduction

- We can apply RANSAC sequentially. However, this is slow because the outlier portion is very high for each object.
- Yi et al. [1] and Dang et al. [2] proposed deep networks to find inliers but only allow one object.
- Challenges to extend to multiple objects:
  - high outlier portions
  - several potential groups of inliers

[1] Moo Yi, Kwang, et al. "Learning to find good correspondences." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018.
[2] Dang, Zheng, et al. "Eigendecomposition-free training of deep networks with zero eigenvalue-based losses." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

Outline

- Introduction

- Proposed Method
  - 3D rotation space quantization
  - Learning-based Facet Network
  - Post-processing And Object Detection

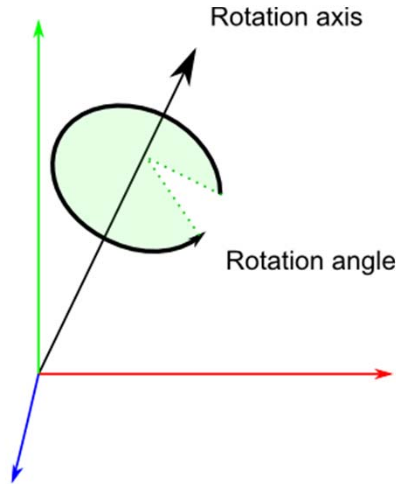
- Experiments
  - Finding Correspondences Of One Object
  - Finding Correspondences Of Multiple Objects
  - Object Detection Performance
  - Performance on GMU Kitchen Dataset

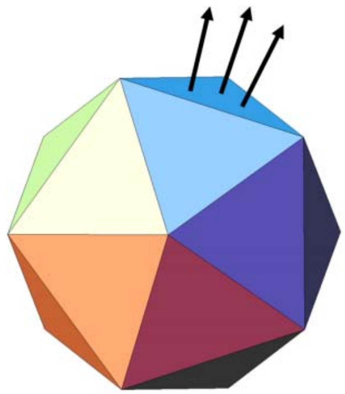
- Conclusion

Proposed Method

3D rotation space quantization

• Any rotation can be represented by a rotation normal vector and a rotation angle around that vector.





Icosahedron
(20 sided shape)

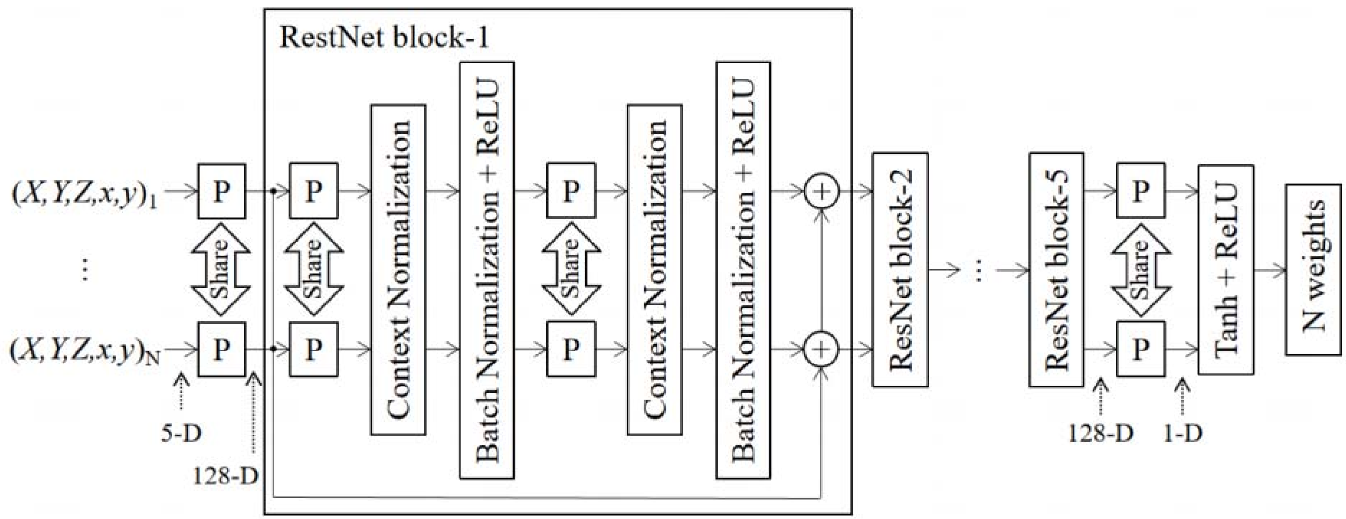
• We discretize the 3D rotation space based on the direction of the rotation vector into 20 bins.

• We say that an object (or a match) belongs to a facet when the pose of the object is associated with a rotation vector pointing towards this facet.

Proposed Method

Learning-based Facet Network

• For each facet, a classifier is trained to identify correspondences that are compatible with a pose whose rotation normal vector points towards this facet.



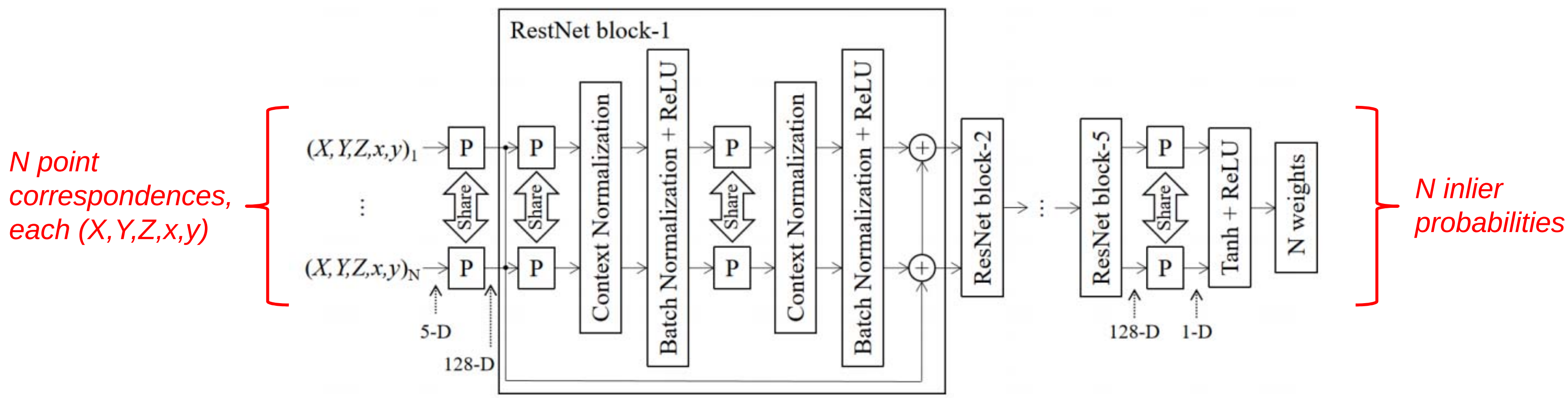
• We use the architecture of Yi et al. [1], but instead of a single classifier we have 20 classifiers, one for each facet.

Proposed Method

Learning-based Facet Network (continued)

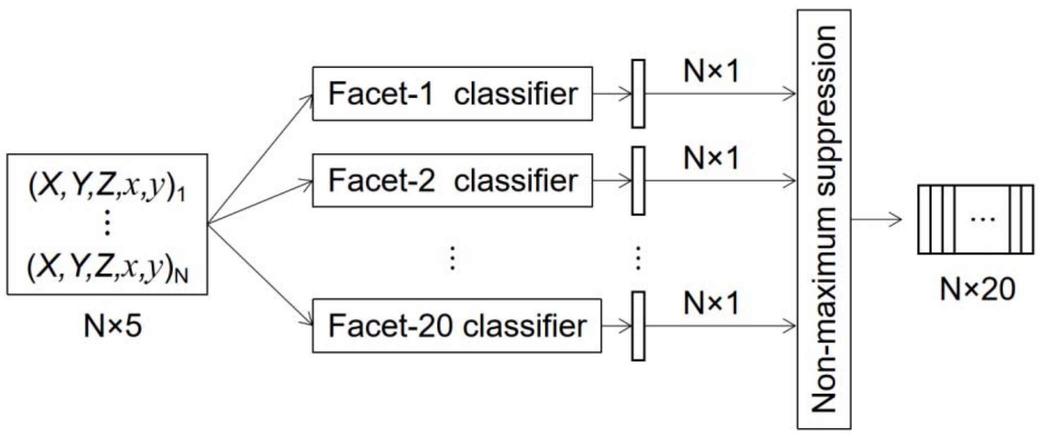
- Each facet classifier is composed of a binary classifier that predicts the inlier probability of matches.
- Classifiers are trained separately, using the loss function:

$$L = -\left[\frac{1}{N_{in}} \sum_{i=1}^{N} \mathbf{1}_i \log(p_i) + \frac{2}{N_{out}} \sum_{i=1}^{N} (1-\mathbf{1}_i) \log(1-p_i) \right]$$

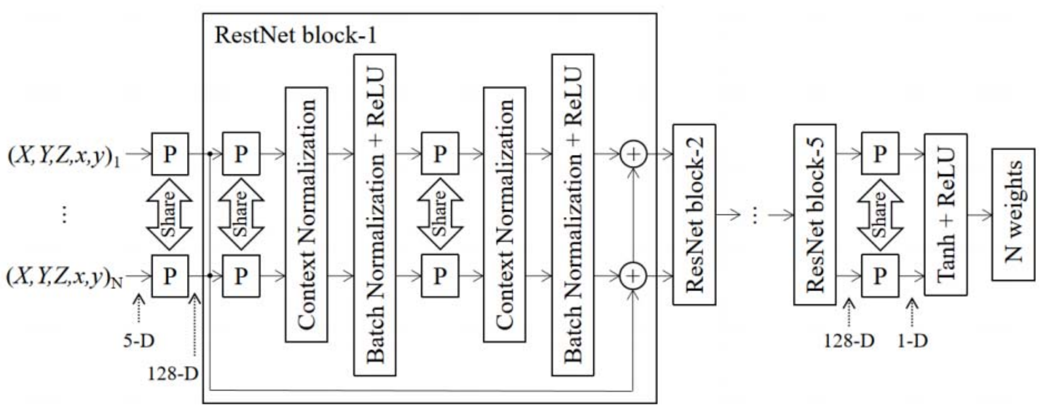


Proposed Method

Learning-based Facet Network (continued)



(a) The structure of the facet network.

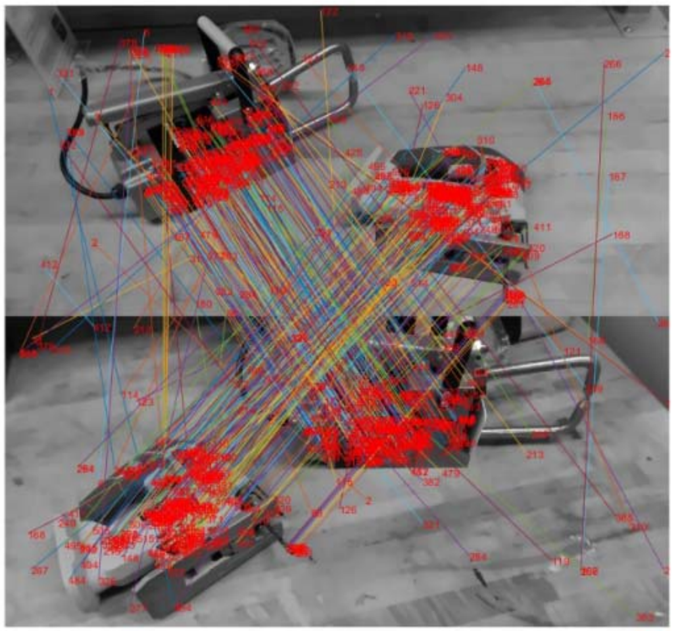


(b) The structure of each facet classifier.

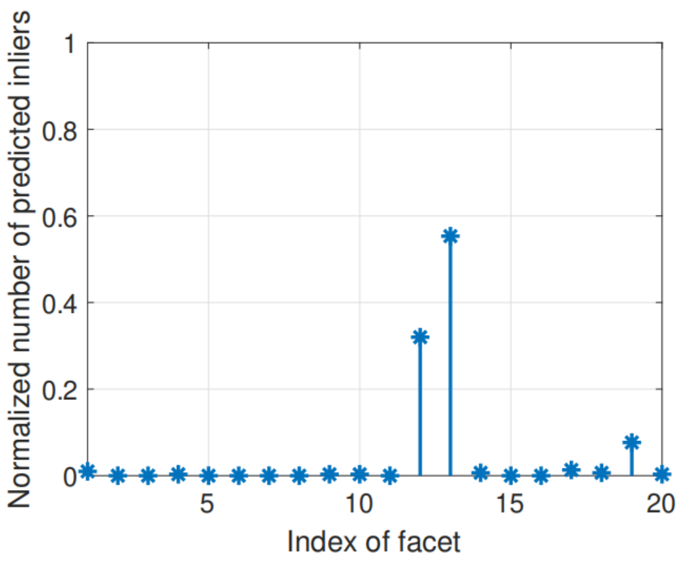
- Outputs from facet classifiers go through a non-maximum suppression block
- That way, each match belongs to only one specific facet.

Proposed Method

- **Post processing: Adaptive thresholding to identify inlier matches**



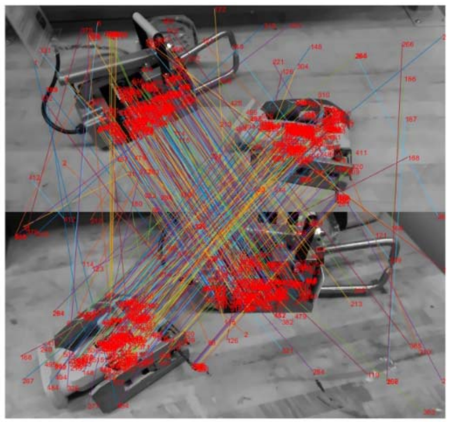
SIFT matching



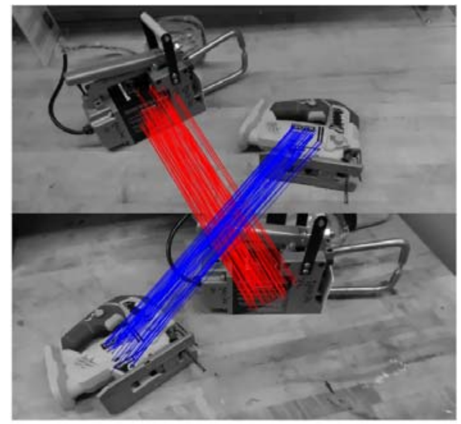
Inlier matches, each facet

Proposed Method

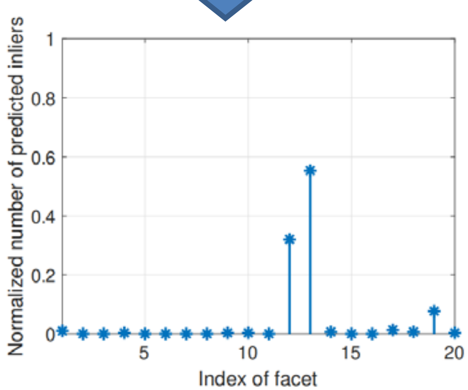
- **Post processing: RANSAC verification of poses**



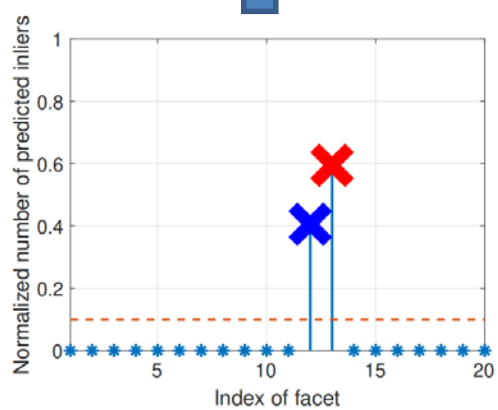
SIFT matching



Final inliers, each object



Inlier matches, each facet



After verification

Proposed Method

Details of Post-processing Steps

- Adaptive thresholding:

Assume there exists at most k objects, we decrease the threshold from 0.9 (inlier probability) with step 0.05 until we have k facets with at least 20 inliers.

- RANSAC-based clustering:

Fit a transformation for the facet with the largest number of inliers and verify this transformation using predicted inliers from all facets. Predicted inliers in other facets that agree with this transformation will be assigned to the current examining facet. Repeat this process for the facet with the second largest number of inliers and so on.

- Object number detection:

A simple thresholding with threshold value 0.1 on the normalized number of predicted inliers for each facet can be applied to detect the number of objects.

- The post processing is very fast because the inlier portion is very high among the estimated inliers of each facet.

Outline

- Introduction
- Proposed Method
  - 3D rotation space quantization
  - Learning-based Facet Network
  - Post-processing And Object Detection
- Experiments
  - Finding Correspondences Of One Object
  - Finding Correspondences Of Multiple Objects
  - Object Detection Performance
  - Performance on GMU Kitchen Dataset
- Conclusion

Experiments

Finding Correspondences Of One Object

Experimental setup:
- Total number of matches: 200
- Inlier portion of each object: 30%
- Gaussian noise standard deviation: 2 pixels

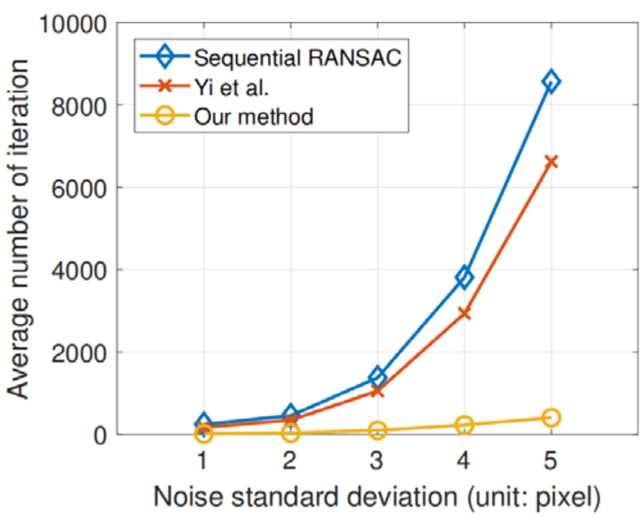
	Precision	Recall	Average number of iterations
RANSAC	99.9%	80.4%	374.3
Yi et al. [13]	99.8%	75.4%	24.4
Our method	99.2%	75.7%	17.5

RANSAC: sequential RANSAC
Yi et al.: Moo Yi, Kwang, et al. "Learning to find good correspondences." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018.

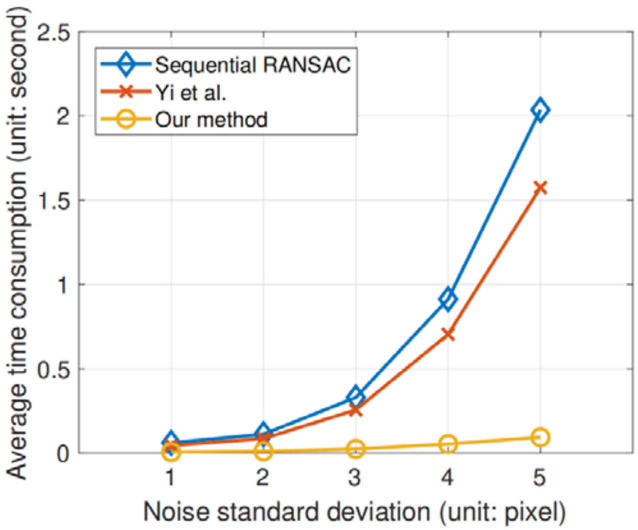
Experiments

Finding Correspondences Of Multiple Objects

Experiment setup:
- Total number of matches: 200
- Number of objects: 3
- Inlier portion of each object: 30%
- Gaussian noise standard deviation: Varying



(a) Average number of iterations.

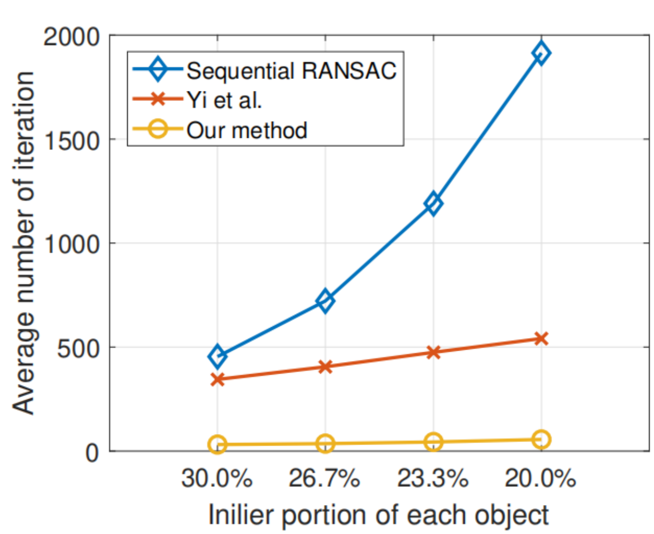


(b) Average time consumption.

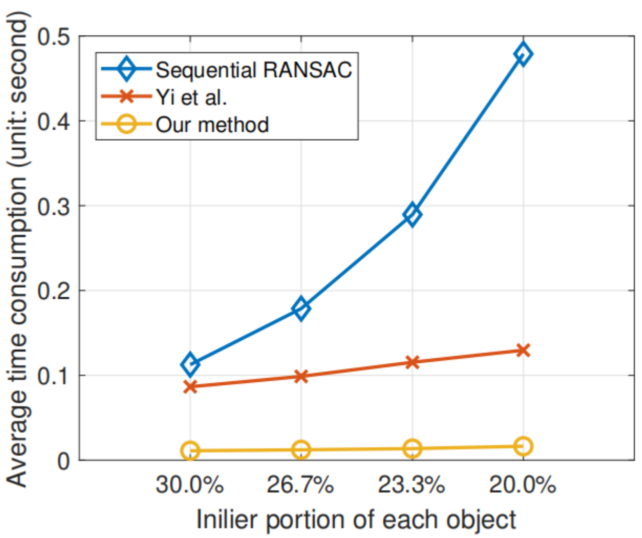
Experiments

Finding Correspondences Of Multiple Objects

Experiment setup:
- Total number of matches: 200
- Number of objects: 3
- Inlier portion of each object: Varying
- Gaussian noise standard deviation: 2 pixels



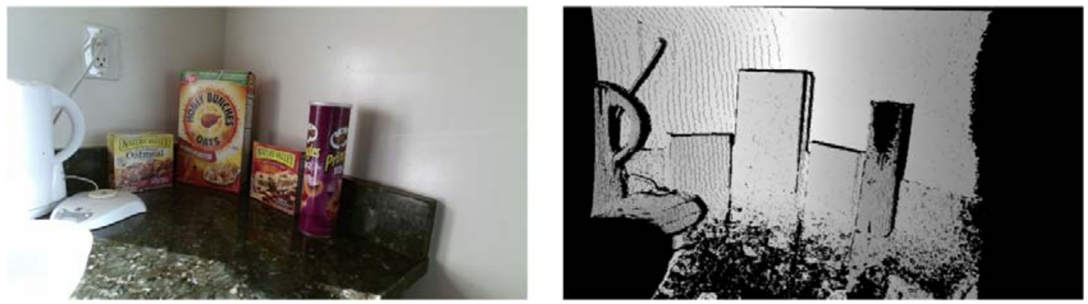
(a) Average number of iterations.



(b) Average time consumption.

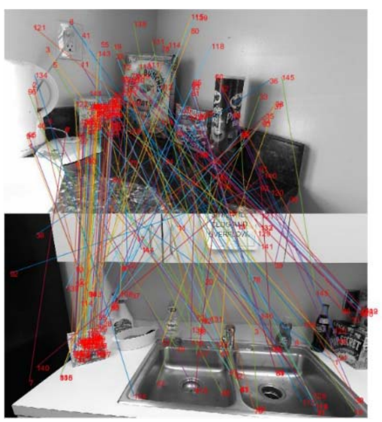
Experiments

Performance on GMU Kitchen Dataset



Template RGB-D image

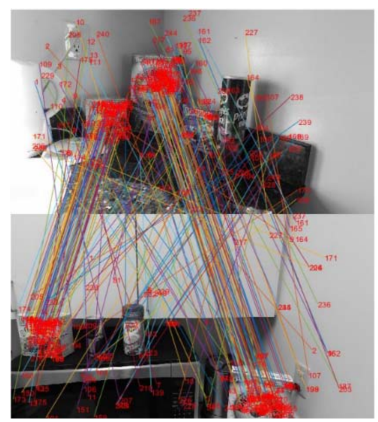
Test image



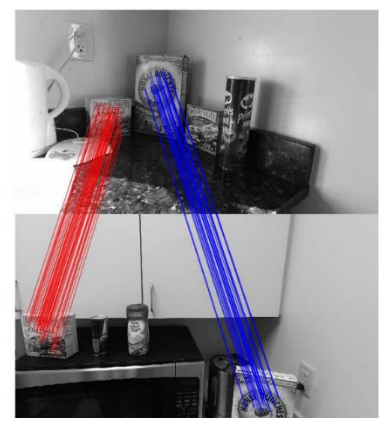
Initial matches

Inlier matches

Test image



Initial matches



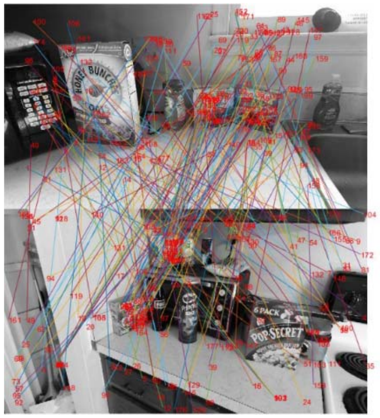
Inlier matches

Experiments

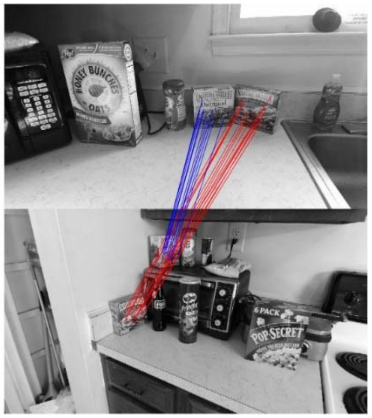
Performance on GMU Kitchen Dataset

Template RGB-D image

Test image



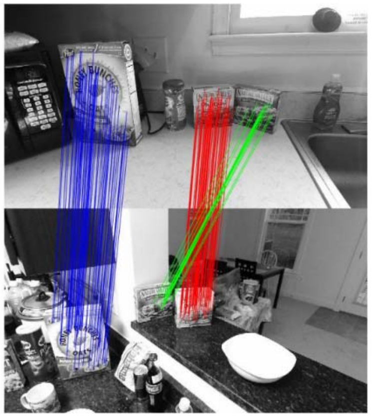
Initial matches



Inlier matches

Test image

Initial matches



Inlier matches

Outline

- Introduction
- Proposed Method
  - 3D rotation space quantization
  - Learning-based Facet Network
  - Post-processing And Object Detection
- Experiments
  - Finding Correspondences Of One Object
  - Finding Correspondences Of Multiple Objects
  - Object Detection Performance
  - Performance on GMU Kitchen Dataset
- Conclusion

Conclusion

- Our method uses a learning-based network and a RANSAC-based post processing step to find inlier correspondences for multiple objects.
- We discretize the 3D rotation space based on the direction of the rotation vector using a regular icosahedron, and train a classifier for each facet of the icosahedron.
- Our proposed method is extremely efficient compared to existing methods.

Thank you!