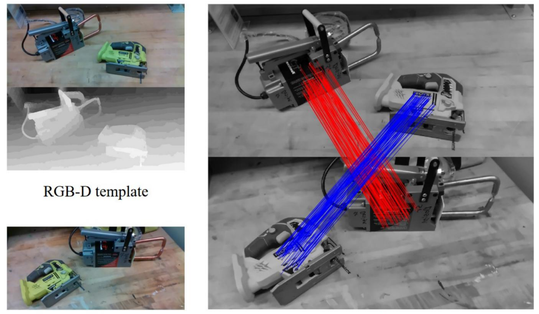
Learning To Find Good Correspondences Of Multiple Objects

Youye Xie[1], Yingheng Tang[2], Gongguo Tang[1], and William Hoff[3]
[1] Department of Electrical Engineering, Colorado School of Mines
[2] School of Electrical and Computer Engineering, Purdue University
[3] Department of Computer Science, Colorado School of Mines

Introduction

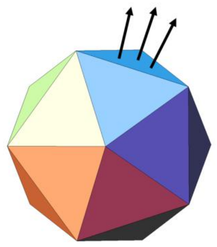
- Given RGB-D templates, how to efficiently detect the object?



- Feature-matching, via SIFT, between the test and template images gives us a set of 3D to 2D putative matches.
- We aim to label the correspondences as inliers or outliers and classify the inliers into multiple objects.
- Challenges:
  - high outlier portions
  - several potential groups of inliers

Rotation Quantization

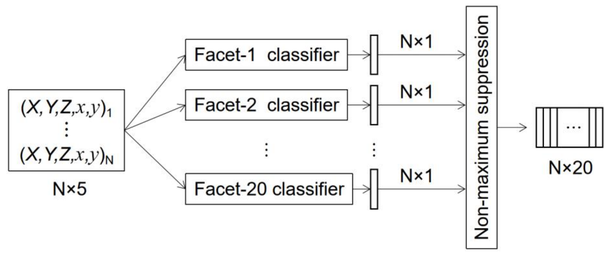
- 3D rotation space quantization via a regular icosahedron.



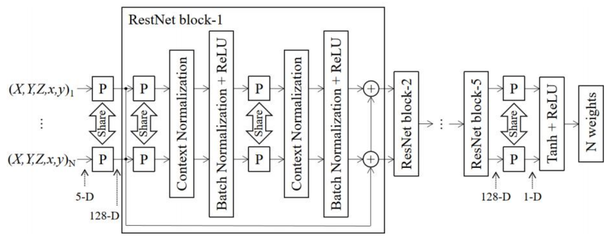
We say that an object belongs to a facet when the pose of the object is associated with a rotation vector pointing towards this facet.

Proposed Method

- Learning-based Facet Network



(a) The structure of the facet network.

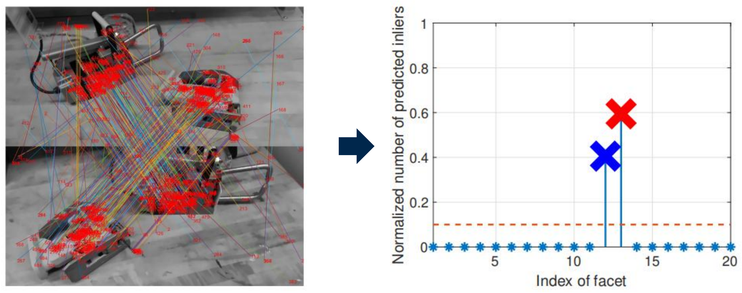


(b) The structure of each facet classifier.

For each facet, a facet classifier is trained to identify correspondences that are compatible with a pose whose rotation normal vector points towards this facet, using loss function:

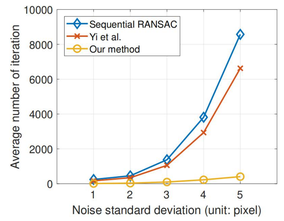
$$L = -\left[\frac{1}{N_{in}}\sum_{i=1}^{N}\mathbf{1}_i \log(p_i) + \frac{2}{N_{out}}\sum_{i=1}^{N}(1-\mathbf{1}_i)\log(1-p_i)\right]$$

- Post-processing And Object Detection
  - adaptive thresholding
  - RANSAC-based clustering

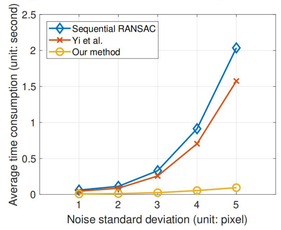


Experiment

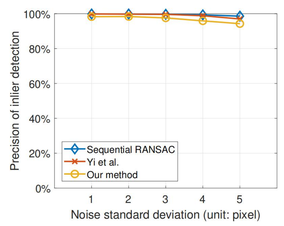
- Comparison to existing methods
  - Sequential RANSAC, Inlier Prediction Network by Yi et al.
- Dataset
  - 3 objects, 30% inlier portion for each object
- Avg. number of iterations



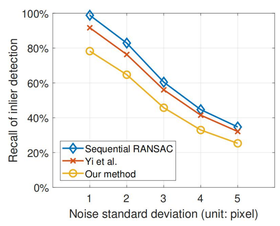
- Avg. time consumption



- Inlier detection precision



- Inlier detection recall



- Performance on GMU Kitchen Dataset