Object recognition and pose estimation using SIFT and MROL techniques

Solutions in Perception Challenge 2011

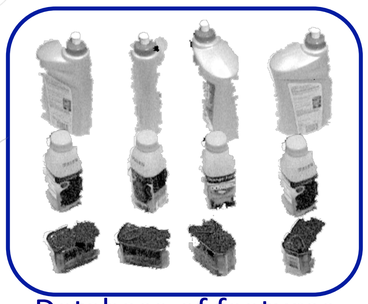
University at Buffalo
The State University of New York

Colin Lea[1] • *Julian Ryde*[1] • *Nick Hillier*[2] • *Jason Corso*[1]

SUNY, University at Buffalo[1] ICT Centre CSIRO[2]

CSIRO

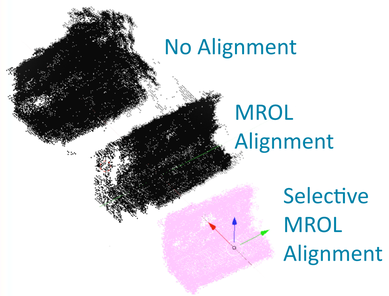
Trainer



Database of features

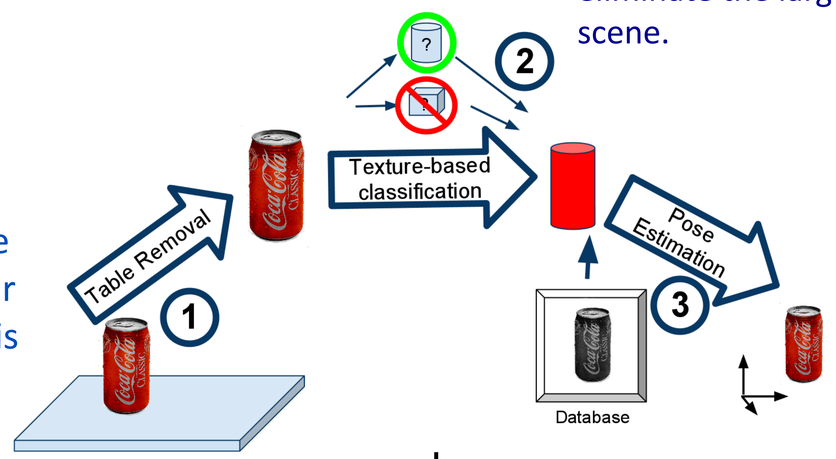
Feature Extraction

SIFT features are extracted from different viewpoints and merged into one dataset.



Model Alignment

Different viewpoints are selectively added if their alignment using MROL is less than a specified threshold.



Objects 1-10

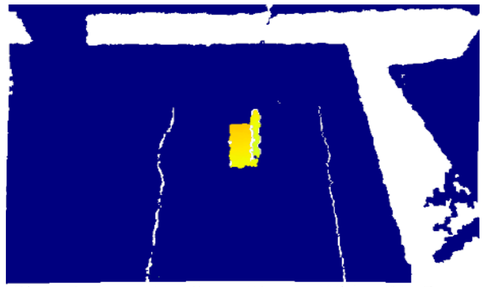
Storage

Texture features and pointclouds are stored in text and Numpy binary files for use in detection.

Detector

Table-based segmentation

Surface normals are calculated using SVD. RANSAC is used on select points to find and eliminate the largest plane in the scene.

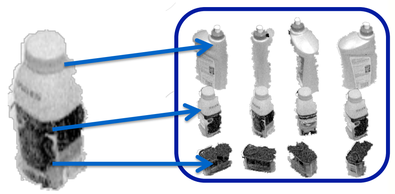


Blue=Removed Yellow=Object White=No Data

Table Removal

Texture classification

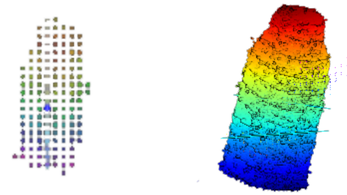
SIFT features are extracted from the scene and compared to all of the models in the database.



Match Descriptors

MROL Pose Estimation

The stored model for the most-likely object is aligned with the pointcloud extracted from the scene.



Scene Aligned Model

MROL$_1$ ***(Multi Resolution Occupied Lists)*** **:** A 6 DoF localization method using a probabilistic multi-resolution approach with occupied voxel lists

SIFT$_2$ ***(Scale-Invariant Feature Transform)*****:** A method for determining scale-invariant features which can be compared between images

Citations:

[1] J. Ryde and N. Hillier. Alignment and 3D scene change detection for segmentation in autonomous earth moving. ICRA 2011.

[2] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 60, 2, pp. 91-110, 2004.

Additional thanks to UB students Kevin Yam, Lai Lee, and Dan Molik