Cluttered Scene Segmentation Using the Symmetry Constraint
Aleksandrs Ecins, Cornelia Fermüller, Yiannis Aloimonos
University of Maryland College Park

Problem statement
Object segmentation in pointclouds of heavily cluttered tabletop scenes collected from multiple views.

Observation
Almost all of man-made objects have shape that exhibits 3D bilateral symmetry.

Idea
1. Detect bilateral symmetries in the scene.
2. Use symmetries as seeds for segmentation.

Symmetry detection

- Two points in space \( p_1, p_2 \) uniquely define a bilateral symmetry.
- The quality of the symmetric match between two points with normals is measured by the angular difference between the normal of the first point \( n_1 \) and the reflected normal of the second point \( n_2 \).

Naive approach
Find symmetric correspondences between pairs of points in the cloud.

Our approach
Find symmetric correspondences between pairs of surface normal edges in the pointcloud.

Segmentation
Goal: find subsets of points in the observed pointcloud that satisfy the following grouping principles:
- Symmetry consistency
- Convexity
Setup multiple graph based foreground segmentation problems each tuned to a single symmetry hypothesis. Solve using graph cuts.

Symmetry Consistency
Key insight: Symmetry hypothesis that corresponds to one of the objects in the scene is consistent with all of the observed points of that object.

Given a point \( p \) of the input pointcloud and a bilateral symmetry hypothesis \( S \) we analyse \( p \)’s reflection \( p' \) by \( S \):
- \( p \) reflects to a point in the cloud → consistent with \( S \)
- \( p \) reflects into occluded space → possibly consistent with \( S \)
- \( p \) reflects into free space → not consistent with \( S \)

Convexity
Objects tend to be convex. Two adjacent points in a convex configuration are likely to belong to the same object whereas a concave transition between two points indicates a boundary between two objects.

Evaluation
Dataset: 89 tabletop scenes captured by moving a Kinect sensor.
Metric: An object in the scene is considered to be segmented correctly if the segmentation algorithm return at least one mask which overlaps with the ground truth mask by more than 90%.

Compare to two state-of-the-art algorithms:
- Felzenswalb [1]
- Locally Convex Connected Patches (LCCP) [2]

Results

References: