CVFH

Clustered **V**iewpoint **F**eature **H**istogram

Aitor Aldoma
Vienna University of Technology

04-04-2011

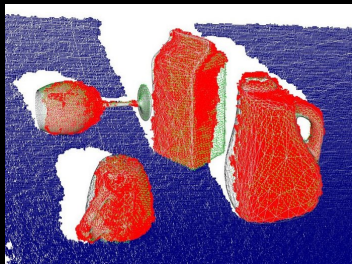pointcloudlibrary

pointcloudlibrary

Introduction

► **Goal**: Object recognition + 6D pose

► **Context**:
  ► object manipulation (0.5m - 1.5m)
  ► efficient
  ► robust to noise and missing parts/partial occlusions
  ► training on synthetic data (CAD models)
  ► recognition using Kinect sensor

 pointcloudlibrary   VFH (1/2)

Viewpoint Feature Histogram

▶ Based our work on VFH
▶ Encodes the surface of the object and the viewpoint
   ▶ Relative to the centroid and the average of the normals
▶ Efficient to compute

pointcloudlibrary
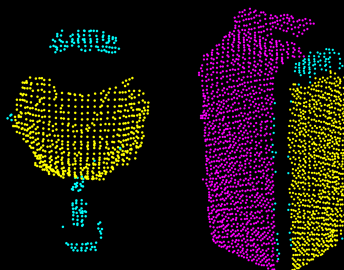
VFH (2/2)
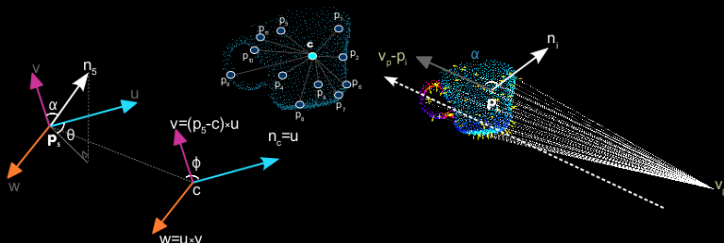
Problems with VFH

- ▶ Invariant to size of the object
- ▶ Invariant to rotation about the roll axis of the camera.
- ▶ Very sensitive to missing parts
  - ▶ Segmentation / Partial occlusions / Noisy parts


- ▶ Our synthetic training data is perfect!

 pointcloudlibrary   CVFH (1/2)

Clustering
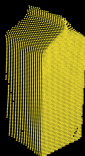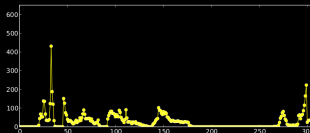
- Parts based model
- Smooth region growing using position + normal information

$$\exists p_j \in C_k : ||p_i - p_j|| < t_d \wedge n_i \cdot n_j > t_n$$

- Use each region's average normal and centroid
- ... to compute a CVFH descriptor

⚬ point**cloud**library

CVFH (2/2)

▶ Histograms are not normalized (variant to objects' size)

▶ Each bin represents number of points
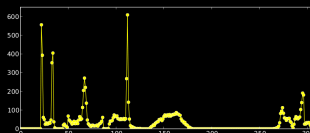
▶ Histogram metric robust to outliers:

$$d(A, B) = 1 - \frac{1 + \sum\limits_{i=1}^{308} \min(A_i, B_i)}{1 + \sum\limits_{i=1}^{308} \max(A_i, B_i)}$$

pointcloudlibrary

Camera's roll histogram

- ▶ Use normal information projected on the view plane
- ▶ Cross correlation maximization problem (FFT)

pointcloudlibrary
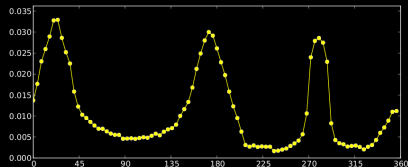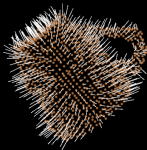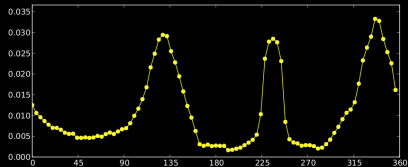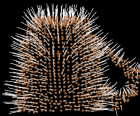
Results (1/3)

- ▶ Recognition results for 18 out of 44 objects in our training set.
- ▶ Improvement over VFH (first - 28.9%, top10 - 43.6%).
- ▶ Reduce from 1704 to top-10 views in 2ms using FLANN.
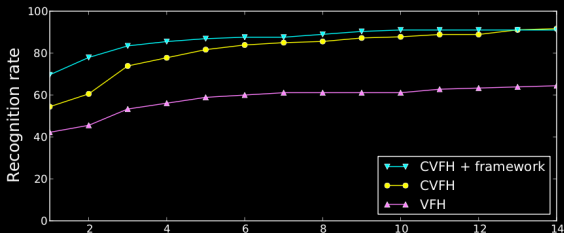
- ▶ Framework involves ICP (10 iter.) for N models.
- ▶ Inliers count and reordering.

pointcloudlibrary

 pointcloudlibrary   Results (3/3)
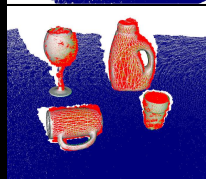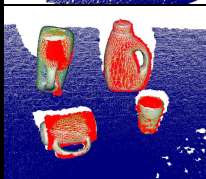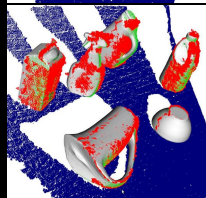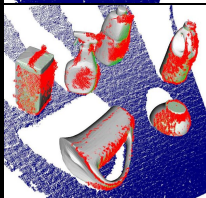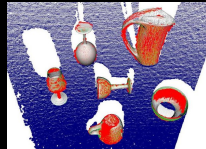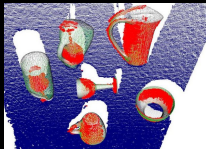
Integration with the manipulation pipeline

▶ Replace current object detection with CVFH.

▶ Object recognition in any pose given segmentation.

▶ Good segmentation improves CVFH. Other way around?