Perception-enabled Knowledge
IAS Internal Workshop

Dejan Panceric, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
Intelligent Autonomous Systems Group
Technische Universität München
June 14, 2012
Outline

1. Semantic Mapping
   Semantic Mapping

2. Perception of Objects of Daily Use
   Interactive Segmentation
   Object Categorization - Graph-Parts
   Object Categorization - Hough Voting
   ODUFinder
   Object Recognition using Barcodes

3. Dynamic Scenes and Spatio-Temporal Memory (Nico)
   GPU-based Perception
   Compression of Point Clouds
   Unstructured Information Management applications

4. Misc
Semantic Mapping

Semantic map representation

Abstract knowledge about object classes

Object instances and component hierarchy

Poses in the environment and their changes over time

Related: TBOX/SBOX, Galindo et al (RAS 2008)
Semantic Mapping

System Version 1

DATA ACQUISITION
- Point cloud Data
- Camera Images

EXPLORATION OF UNKNOWN ENVIRONMENT
- Determining Unmapped Regions
- Computing Candidate Robot Poses
- Pose Validation

POINT CLOUD DATA INTERPRETATION
- Recognition of Planes and Fixtures
- Generation of Door and Drawer Hypotheses
- Validation of Hypotheses

SEMANTIC MAP GENERATION

Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
Semantic Mapping

Acquisition of Sensor Data

360deg Scans plus ICP optimizations:
- overlapping regions
- zeroed roll and pitch
Semantic Mapping

Next Best View - Visibility Kernel

Costmap-based approach
1. max fringe points costmap, 2. occupied voxels costmap $\rightarrow$ minimum intersection
Semantic Mapping

Point Cloud Data Interpretation - Floor, Ceiling

- classifying points close to min/max along Z axis as floor and ceiling
Semantic Mapping

Point Cloud Data Interpretation - Walls, ROIs

• classifying planes perpendicular to X and Y axis orientation of the whole point cloud and close to min/max along those axes as walls
• classifying tabletops as those between 0.75m-1m height
Semantic Mapping

Point Cloud Data Interpretation - Fixtures

- remaining vertical surfaces and tabletops classified as areas of interest
- finding fixtures - handles and knobs
- various handle and knob detectors
Semantic Mapping

Door and Drawer Hypotheses

- using region growing
- median intensity and median average distance
Semantic Mapping

System Version 2

Acquisition

Pre-processing → Registration (III-A1) → Surface Reconstruction (III-A2) → Texture Re-projection (III-A3) → Next Best View Planning (III-A4)

Interpretation

Ask-Tell Interface → SOM+ Map Implemented in KnowRob (Section II) → Detection of Relevant Planes (III-B1)

Object of Daily Use Detection And Recognition (Pangercic, IROS2011) → Door and Drawer Hypotheses Validation (III-B5) → Detection of Handles (III-B2)

Door and Drawer Hypotheses (III-B4) → Articulation Model Learning (III-B3)

*
Semantic Mapping

Registration, Reconstruction, Texture Re-projection

Testbed Kitchens  Poisson surface reconstruction  Blending-based texture reprojection
Semantic Mapping

Learning of Articulation Models - Arm Control

- Using impedance controller from Willow Garage (pr2_cockpit stack)
- Initialization: gently pull the handle backwards by moving the Cartesian equilibrium point towards the robot
- Record the trajectory of robot’s gripper $y_{1:n}$ with $y_i \in \text{SE}(3)$

VIDEO: *

Perception

Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
Semantic Mapping

Learning of Articulation Models - Model Fitting

Iteratively:
- (Re-)estimate the kinematic model $\mathcal{M} \in \{\text{rigid, prismatic, rotational}\}$
- Estimate model-specific parameter vector $\theta \in \mathbb{R}^d$ (encoding radius, rotation axis) of the articulated object:
  $\hat{\mathcal{M}}, \hat{\theta} = \arg \max_{\mathcal{M}, \theta} p(\mathcal{M}, \theta \mid y_{1:n})$
- Fit the parameter vector of all model candidates using an MLESAC estimator
Semantic Mapping

Learning of Articulation Models - Model Selection

- Select the best model according to the Bayesian information criterion (BIC)
- Use the model to predict the continuation of the trajectory and to generate the next Cartesian equilibrium point \( \mathbf{x}_{n+1}^{CEP} \).
- Finally, determine opening angle / opening distance
- Output: Kinematic model
Semantic Mapping

Door and Drawer Hypotheses Validation through Interaction
Semantic Mapping

Final Segmented Map
Semantic Mapping

Final KnowRob Map

Acquisition Code:
bosch_registration, bosch_surface_reconstruction, bosch_texture_reconstruction, mapping

Representation Code:
http://www.ros.org/wiki/knowrob
Semantic Mapping

Future Work

In collaboration with KTH:

- classification of rooms using shape, size and set of objects
- language for probabilistic representation and reasoning
- room and object novelty detection
- integration of planning
Semantic Mapping

Tutorial 1

In progress ... - by end of June 2012.

- RGB + D + Plane-based + Edge-based Registration
- Automatic segmentation of fixtures and dimensions
- Automatic learning of articulation models
Outline

1. Semantic Mapping
   Semantic Mapping

2. Perception of Objects of Daily Use
   Interactive Segmentation
   Object Categorization - Graph-Parts
   Object Categorization - Hough Voting
   ODUFinder
   Object Recognition using Barcodes

3. Dynamic Scenes and Spatio-Temporal Memory (Nico)
   GPU-based Perception
   Compression of Point Clouds
   Unstructured Information Management applications

4. Misc
Interactive Segmentation

System
Interactive Segmentation

Contact Point and Push Direction

- Segmentation of cluttered scenes
- Contour estimation
- Finding concave corners
- Bisector gives a push direction
Interactive Segmentation

Clustering Algorithm

- Features $u$ and $v$, depicted as red circles, are randomly selected.
- From their trajectories $S_u$ and $S_v$, a rigid transformation $A_t$ is calculated.
- If $u$ and $v$ are on the same object, all other features will move according to the sequence of $A_t$. 
Interactive Segmentation

Results
Interactive Segmentation

Future Work

• dealing with textureless and transparent objects
• real-world scenes
• integration of an arm planner
• constrained spaces
Interactive Segmentation

Tutorial 2


- Form 3 groups with 1 Kinect, 1 tripod and 2 objects each.

Thanks to Karol Hausman!
Object Categorization - Graph-Parts

Introduction

- Object categorization in cluttered scenes where accurate segmentation can be difficult to achieve.
- Over-segmentation and multiple hypotheses better than relying on a single, possibly erroneous segmentation.
- Approach based on scene- or part-graphs, using additive RGBD feature descriptors and hashing.

Figure: Overview of the process of scanning segmenting and categorizing objects in clutter.
Object Categorization - Graph-Parts

Objectives

- an efficient part-based object classification method for cluttered scenes, taking into account relations between parts;
- a graph-theoretic hashing method that allows model refinement while having competitive classification performance;
- evaluation of geometric, color, and multi-modal features and classification approaches;
- exemplifying the advantages of multiple views, multiple segment groupings, and domain adaptation;
Object Categorization - Graph-Parts

Radius-based Surface Descriptor (RSD)

Theory and Approximation

- Estimate minimum and maximum curvature radius from angle/distance pairs:

\[ d(\alpha) = \sqrt{2r} \sqrt{1 - \cos(\alpha)} \Rightarrow d(\alpha) = r\alpha + \frac{r\alpha^3}{24} + O(\alpha^5) \Rightarrow d = r \cdot \alpha \]
Object Categorization - Graph-Parts

Radius-based Surface Descriptor (RSD)

Principle Radii

- Local variation of normal angles by distance (similar to PFH and “spin images with normals”):

The tilt angles of the lines starting from bottom left corner correspond to the physical radii:
smallest tilt that still covers occupied cells to the min. radius, while the biggest to the max.
Object Categorization - Graph-Parts

Static Scenes

(a) Scene 1
(b) Scene 2
(c) Scene 3
Figure: As the camera is moved (left), multiple frames can be captured that cover different parts of the objects in the scene (right).
Object Categorization - Hough Voting

Part-Based Object Detection using Hough Voting and Model Fitting

- Same task as before, with larger objects: furniture pieces have to be identified for which (similar, but not exactly matching) CAD models are available.

- Here pose estimation is added by model matching and geometric verification.

- Parts are considered separately, and only 3DOF are possible, so more descriptive features can be used (statistics about the distribution of the 3D points).

- Integration is achieved by parts being categorized (unsupervised) and voting for likely positions.
Object Categorization - Hough Voting

Part Codebook of Example CAD Models

- Realistic scan simulation, unsupervised segmentation, building a codebook of parts, learning of spatial relations

- Probabilistic Hough voting to obtain likely object locations and a weighted list of their parts
Object Categorization - Hough Voting

Model Fitting for Verification and 3DOF Pose

Assigning parts/points to objects and rejecting false positives
Object Categorization - Hough Voting

Real-world Results: Office
Object Categorization - Hough Voting

Real-world Results: Seminar Room

- remaining false matches due to high occlusions
Object Categorization - Hough Voting

Improving Results by Taking Multiple Scans

- Decreasing occlusion, increasing number of parts
- Similarly as for small objects: merging the votes from multiple scans

- The model fitting and verification steps also do not assume one viewpoint or one segmentation per scan
Object Categorization - Hough Voting

Future Work

- test on objects of daily use
- include color features
- explore the “power-of-data”
- integrate online building of models and sharing (e.g. via RoboEarth)
Object Categorization - Hough Voting

Tutorial 3

ODUFinder
Object Recognition with ODUFinder - Idea

"The time I had to call because I goofed on my order, they were awesome and understanding. I will always shop with GermanDeli."

- Stella M., TN
"Couldn't have asked for better service or communication."

- Karle L., CA

**Description**


**Ingredients:**
- 200g flour, 40g sugar, 10g baking powder, 10g margarine, 50g milk, 2 eggs, 50g chocolate chips.

**Nutrition Facts:**
- Calorie Content (per serving): 200 cal
- Total Fat: 5g
- Total Carbohydrates: 50g
- Protein: 8g

**Tip:**

1. Preheat the oven to 180°C (350°F).
2. Add the dry ingredients to a bowl and mix well.
3. Add the wet ingredients and stir until well combined.
4. Pour the mixture into a greased baking dish and bake for 25 minutes or until golden brown.
5. Serve warm with a dollop of whipped cream.
ODUFinder

System

In-hand Object Modeling

Barcode Localization and Decoding

Barcoo Website

SIFT Extraction and Vocabulary Tree Training

Knowledge Base

[Perception Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi]
ODUFinder

Vocabulary Tree and SIFT

+ Metric Score (e.g. L1) + TF-IDF

Perception

Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
ODUFinder

Over-Segmentation-Based Object Candidate Detection

\[ r^2(x) = (r_{max}^2 - r_{min}^2)(K(1 - \text{logsig}(x - A))) + r_{min}^2 \]  \hspace{1cm} (1)

where \text{logsig} is defined as:

\[ \text{logsig}(x) = \frac{1}{1 + e^{-x}}, \]  \hspace{1cm} (2)
ODUFinder

Over-Segmentation-Based Object Candidate Detection - cont.
ODUFinder

Semantic Map Prior
ODUFinder

Topic-based Integration with KnowRob

- Automatically created ontology of >7500 objects from the online shop germandeli.com
- Class hierarchy from categories + perishability, weight, price, origin, ...
- Code:
Object Recognition using Barcodes

Idea

Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
Object Recognition using Barcodes System
Object Recognition using Barcodes

Barcode Decode

video input → scan passes → linear scanner → element decoder

image stream → intensity sample stream → element width stream → decoded data stream

ZBar image scanner

EAN-13: 987654321012

Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
Object Recognition using Barcodes

Server Communication

- **Request:** `http://www.barcoo.com/api/get_product_complete?pi=4101530002475&pins=ean&format=xml&source=ias-tum`

- **Response:**

```xml
<root>
  <request>
    <name>format</name><value>xml</value>
    <name>qi</name><value>1</value>
    <name>heavy_load</name><value>false</value>
    <name>brand</name><value>barcoo</value>
    <name>action</name><value>get_product_complete</value>
    <name>pi</name><value>4101530002475</value>
    <name>uuid</name><value>9405f6ff-39ec-4611-8a12-2abd7cc419a</value>
    <name>pins</name><value>ean</value>
    <name>controller</name><value>api</value>
    <name>source</name><value>ias-tum</value>
  </request>
</root>
```

- **Information parsed:** image, category, product name, brand

---

Perception

Dejan Pangercic, Nico Blodow, Zoltan-Csaba Marton, Ferenc Balint-Benczedi
Object Recognition using Barcodes

Integration with KnowRob

- OWL file-based
- Topic-based using json server (TBD)
Object Recognition using Barcodes

Tutorial 4

For this tutorial you will need to connect to the “lab” wireless network (pw: artificialintelligence) and disable your wired connection.
Thanks to Nacer Khalil.
Outline

1. Semantic Mapping
   Semantic Mapping

2. Perception of Objects of Daily Use
   Interactive Segmentation
   Object Categorization - Graph-Parts
   Object Categorization - Hough Voting
   ODUFinder
   Object Recognition using Barcodes

3. Dynamic Scenes and Spatio-Temporal Memory (Nico)
   GPU-based Perception
   Compression of Point Clouds
   Unstructured Information Management applications

4. Misc
Outline

1. Semantic Mapping
   Semantic Mapping

2. Perception of Objects of Daily Use
   Interactive Segmentation
   Object Categorization - Graph-Parts
   Object Categorization - Hough Voting
   ODUFinder
   Object Recognition using Barcodes

3. Dynamic Scenes and Spatio-Temporal Memory (Nico)
   GPU-based Perception
   Compression of Point Clouds
   Unstructured Information Management applications

4. Misc
Misc

List of things we did not talk about ...:

- object tracking
- http://ros.org/wiki/cop
- plethora of algorithms in PCL (pointclouds.org)
- people/hands tracking
- DAFT feature
- ...
Thank You

Contact:
http://ias.cs.tum.edu/people/[pangercic|marton|blodow|balintbe]