RGB-D at the Robotics and State Estimation Lab

Kevin Lai

Collaborators: Dieter Fox, Liefeng Bo, Marc Deisenroth, Peter Henry, Evan Herbst, Michael Krainin, Jinna Lei, Cynthia Matuszek, Xiaofeng Ren
RGB-D Mapping

Henry, Krainin, Herbst, Ren, Fox (ISER ‘10)
Object Modelling

Krainin, Henry, Ren, Fox (IJRR ‘11)  
Krainin, Curless, Fox (ICRA ‘11)
Object Discovery

Herbst, Henry, Ren, Fox (ICRA ‘11)
Game Playing Robot

Matuszek, Mayton, Aimi, Deisenroth, Bo, Chu, Kung, LeGrand, Smith, Fox (ICRA ‘11)
RGB-D Object and Pose Recognition

L. Bo, D. Fox, B. Harrison, K. Lai, J. Lei, P. Powledge, X. Ren, S. Grampurohit, R. Ziola
Object Recognition and Pose Estimation System

Training objects (poses)

System

Unknown object

Category
Instance
Pose
Object Recognition Pipeline

Visual and depth features

Bag of Words

Support Vector Machine

Recognition

image

patch features

image feature
Outline

- RGB-D Object Dataset
- Object Recognition and Detection
- Ongoing work: Object-Pose Tree
RGB-D Object Dataset

- 300 objects in 51 categories
- 250,000 640x480 RGB-Depth frames total
- 8 natural scenes (offices, kitchens, meeting rooms)

K. Lai, L. Bo, X. Ren and D. Fox (ICRA ’11)
Download RGB-D Object Dataset

The RGB-D Object Dataset provided here is for non-commercial research/educational use only. For commercial use, please contact us for more information.

To obtain the RGB-D Object Dataset, please click the "Send Email" link below and enter your full name, email, and affiliation (in the message portion). We ask for this information only to keep track of who is using the dataset. We will not give this to third parties or publish it publicly anywhere. Once your information has been verified, we will send you a link to the dataset via the email address that you entered.

Send Email

RGB-D Object Dataset Evaluation Procedure

We will add details about the evaluation procedure for object recognition and object detection here in the near future. In the meantime, please see our paper for a description.
Segmented Objects
Natural Scenes
Segmentation and Labeling
Outline

- RGB-D Object Dataset

- Object Recognition and Detection

- Ongoing work: Object-Pose Tree
Instance and Category Recognition

Query | Instance Recognition | Category Recognition
# Features

## Feature Set
- Dense SIFT features (RGB)
- Color histograms (RGB)
- Texton histograms (RGB)
- Spin image features (Depth)
- 3D bounding box (Depth)

## KDES + Pyramid EMK
- Gradient kernel descriptor (RGB)
- Shape kernel descriptor (RGB)
- Color kernel descriptor (RGB)
- Gradient kernel descriptor (Depth)
- Shape kernel descriptor (Depth)

L. Bo, X. Ren and D. Fox, NIPS ‘10.
## Results on RGB-D Object Dataset

**Classifier: linear SVM**

<table>
<thead>
<tr>
<th></th>
<th>Depth</th>
<th>RGB</th>
<th>RGB+Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature Set</td>
<td>51.7%</td>
<td>72.7%</td>
<td>81.9%</td>
</tr>
<tr>
<td>KDES + Pyramid EMK</td>
<td>72.7%</td>
<td>76.3%</td>
<td>85.6%</td>
</tr>
<tr>
<td><strong>Instance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature Set</td>
<td>42.3%</td>
<td>59.3%</td>
<td>73.9%</td>
</tr>
<tr>
<td>KDES + Pyramid EMK</td>
<td>49.7%</td>
<td>74.9%</td>
<td>82.5%</td>
</tr>
</tbody>
</table>
## Results on RGB-D Object Dataset

Classifier: linear SVM

<table>
<thead>
<tr>
<th>Category</th>
<th>Depth</th>
<th>RGB</th>
<th>RGB+Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Set</td>
<td>51.7%</td>
<td>72.7%</td>
<td>81.9%</td>
</tr>
<tr>
<td>KDES + Pyramid EMK</td>
<td>72.7%</td>
<td>76.3%</td>
<td>85.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instance</th>
<th>Depth</th>
<th>RGB</th>
<th>RGB+Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Set</td>
<td>42.3%</td>
<td>59.3%</td>
<td>73.9%</td>
</tr>
<tr>
<td>KDES + Pyramid EMK</td>
<td>49.7%</td>
<td>74.9%</td>
<td>82.5%</td>
</tr>
</tbody>
</table>

Category: 3% higher than feature set
## Results on RGB-D Object Dataset

**Classifier: linear SVM**

<table>
<thead>
<tr>
<th>Category</th>
<th>Depth</th>
<th>RGB</th>
<th>RGB+Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Set</td>
<td>51.7%</td>
<td>72.7%</td>
<td>81.9%</td>
</tr>
<tr>
<td>KDES + Pyramid EMK</td>
<td>72.7%</td>
<td>76.3%</td>
<td>85.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instance</th>
<th>Depth</th>
<th>RGB</th>
<th>RGB+Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Set</td>
<td>42.3%</td>
<td>59.3%</td>
<td>73.9%</td>
</tr>
<tr>
<td>KDES + Pyramid EMK</td>
<td>49.7%</td>
<td>74.9%</td>
<td>82.5%</td>
</tr>
</tbody>
</table>

**Instance: 10% higher than feature set**
RGB-D Object Detection

- Example images with detected objects and precision-recall curves for different categories and instances.
Outline

- RGB-D Object Dataset
- Object Recognition and Detection
- Ongoing work: Object-Pose Tree
Object-Pose Tree

The goal: recognize objects and their poses jointly

Joint optimization of parameters of the entire tree based on loss function introduced by Bengio et al. NIPS ’10.
Preliminary Results
Summary

- RGB-D Object Dataset: a large scale, publicly available dataset of objects and natural scenes with RGB+depth videos

- Depth is a very useful cue for object recognition and pose estimation!

- Kernel descriptors are a powerful feature for extracting gradient, shape, and color information from RGB and depth images.

- Our approach has been applied to an interactive LEGO playing scenario that was shown live at CES 2011
Questions?
Thanks for your attention!
RGB-D Recognition: Demo

Category recognition
RGB-D Recognition: Demo

Online Learning
RGB-D Recognition: Demo

Category recognition
Object Hierarchy (WordNet/ImageNet)