Probabilistic Scene Understanding using Virtual Reality and Markov Logic Networks

Dominik Dieckmann

Institute for Artificial Intelligence
University Bremen

18.12.2018

Autonomous Robots in Houshold Environments

- perception component
  - detect objects
  - analyse objects
- reasoning component
  - identify/classify objects based on their visual cues
  - needs to be trained

Creation of Training Data

time and resource intensive:

- manually creating scenarios and images
- no groundtruth

$\rightarrow$ synthetic images from a game engine

System Setup

- list of objects, classes and scenarios
- Unreal Engine to create *Unreal Images*
- RoboSherlock analysis the images
- learn a *Markov Logic Network*
- classify the objects in the images

Unreal Engine

- photorealism
- rendering in realtime
- open source

Unreal Images

- Assets
  - scanned 3D-models of the objects
  - kitchen environment
- URoboVision plugin
  - create RGBD image from a scene
  - create *ObjectImage* and *ObjectMap*
  - send them to RoboSherlock
- RSpawnBox class
  - visual representation of scene space
  - rotates camera around the scene

Unreal Images

- 114 scenes
- 2 - 5 objects per scene
- 5 viewpoints per scene
- only objects of one scenario per scene
- total of 570 images

RoboSherlock

- based on *UIMA*
- segments images and creates object hypotheses
- annotates attributes of the hypotheses

Perceptionpipeline

Annotates the following attributes:

- color
- size
- shape
- goggles_{Logo, Text, Product}
- instance
- object

UnrealGTAnnotator

Annotates the groundtruth by:

- counting the color of pixels in the ObjectImage
- looking up the corresponding asset
- looking up the groundtruth for that asset
- setting the groundtruth

Markov Logic Networks

what are they?
advantages for object classification

Experiments

baseline: PR2 paper
show results...

Conclusion

it just works