A cognitive robot detecting objects and recognizing actions using sound language and perception
Introducing Cognition

Perception
Vision, Sound

Cognition

Action Space
Robots with Language

Robots
- looking at Objects
- humans performing actions
- exploring space
Language

Linguists/ computational linguists have many tools to model lexical semantics
relations such as same meaning
is-a relationship
cause -effect
performs -functions
used for
motivated-by

Examples:
Wordnet: same meaning: “argue” “contend”
is-a-relationship: “car” “vehicle”
Framenet: definition of words as complex structural representation
MindNet, ConceptNet, HowNet
Why language?

Perception was studied bottom up: edges, texture, motion, color, surfaces
Currently: Feature extraction and learning

This is not sufficient!

We need additional knowledge:

We reason while we perceive
Language and Perception

Natural Language Processing
Semantic Nets: Wordnet, Verb net, ConceptNet, etc.
Statistical tools on large corpuses

? Grounding

? Higher level knowledge

Vision:
Actions (motion, change of scene),
Objects( surfaces, color texture, shape)

? Concepts
Challenges

- Object Recognition
- Action Recognition
- Scene Description
for Robot Perception
Object Recognition
Vision: 
Segmentation

Description of objects in terms of attributes:  
- Color 
- Texture 
- Shape properties
The Embodied Lexicon

A lexicon that determines whether a word/concept refers to
(a) A concrete entity (e.g. kitchen knife)
(b) An object category (e.g. knife), or
(c) something more abstract (e.g. cutlery)

It provides information on visual attributes of the denoted objects/concepts (by re-engineering WordNet) and their affordances.

IDEA for Object recognition: use of the hierarchical organisation of this lexicon to predict (in a probabilistic way) an object given its visual attributes.
Frameworks

Implement the interaction: Start with filters of certain attributes
Language tells us what to explore next
Vision computes new attributes

Learn mapping from Language attributes to visual shape representations
Activity Recognition
A girl doing something
Where?

There is a kitchen stand and a cabinet
It’s probably a kitchen.

What’s in her hand?
Looks like a knife.

What do you do with knives?
Cut? Slice?

She is cutting
What is she cutting with the knife?

Cheese
Goals

1. Activity description: the tool/ the object/ the movement/ spatial relationships/ change of object

2. Describe action in terms of sub-events in 3D, the hands touch the object (the grasp) the hands carry out the object the hands release the object
Navigation

Build a map from SLAM
Find objects visually and label them on the map
Summary

Tools prepared

• Robot using ROS tutorial
• Segmentation
• Attention filter
• Skeleton model
• Navigation, SLAM
• Linguistic tools: the Embodied lexicon

Goals for workshop

  development of a number of modules
  discussion of new approaches
Summary

• Robot using ROS - tutorial (Ching Lik Teo, Yezhou Yang, Alex Ecins)
• Segmentation (Ajay Mishra)
• Attention filter (Douglas Summerstay)
• Skeleton model (Yezhou and Ching Lik)
• Navigation (Alex and Austin Myers)
• Linguistic tools: the Embodied lexicon (Katerina Pastra and Eirini Balta)

• During the workshop
  development of a number of modules
A project about Cognition

Knowledge

Vision

Planning

Motor Control

Audition
Concept
dummy_shape_projection#feature

type-token

Concept
dummy_shape_knife_tongue#feature
(Feature#_#)

Concept
knife#abstract

Concept
clapper#abstract

Concept
tongue#abstract

Concept
tongue#entity

Concept
clapper#entity