Proactive Multimodal Perception for Feature Based Anchoring of Complex Objects

Martin Weser, Jianwei Zhang

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics

Cross-modal Interactions in Natural and Artificial Cognitive Systems

15-18 December 2007,
IEEE Robio, Sanya, China

Outline

Motivation

Action and Perception - Two Sides of one Coin

- Action needs perception
  - mobile motion requires localization
  - communication requires attention and understanding
  - manipulation requires object recognition
- Perception is highly influenced by action

but . . .

- Robot perception is almost exclusively used for action
- Robot perception still lacks in complex environments

Motivation

Action and Perception - Two Sides of one Coin

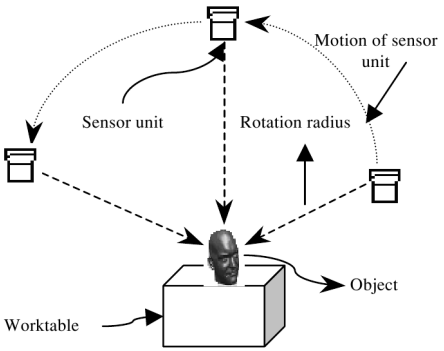
- Action needs perception
  - mobile motion requires localization
  - communication requires attention and understanding
  - manipulation requires object recognition
- Perception is highly influenced by action

but . . .

- Robot perception is almost exclusively used for action
- Robot perception still lacks in complex environments

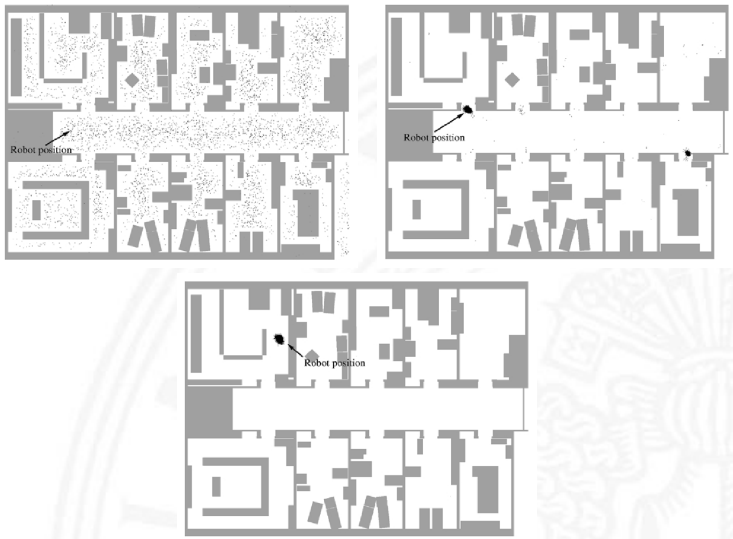
Approaches for Action Oriented Perception

Next best view:



Source: [Li *et all*, Meas. Sci. Technol. 2005]

Localization:

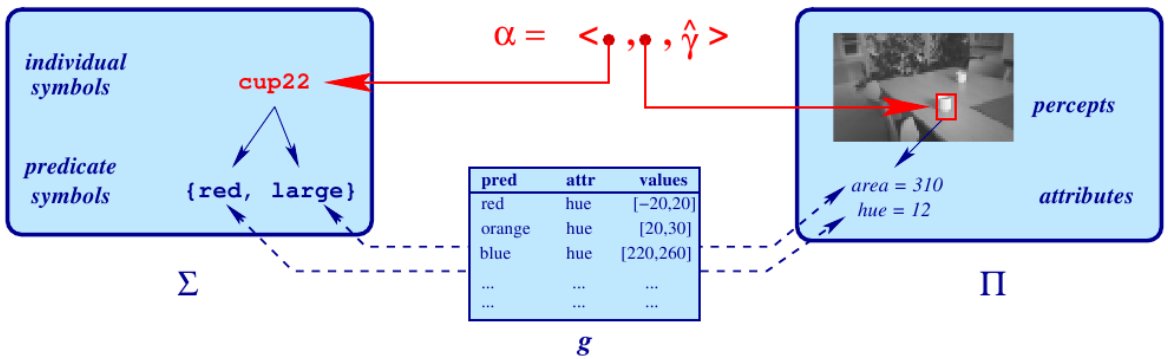


Source: [Thrun *et all*, Artificial Intelligence, 2001]

Review: Anchoring Framework

Coradeschi and Saffiotti (2003)

- Two distinct systems: perceptual and symbol system
- Bottom-Up information flow. Information tokens are
  - percepts: "Subset of perceptual data that is originated by one object"
  - measurable attributes of percepts
  - predicate symbols derived from attributes
- Symbol system keeps track of objects and assign predicate symbols to objects

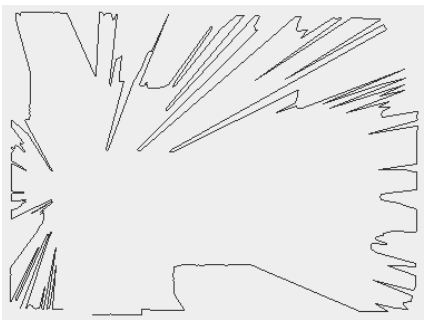
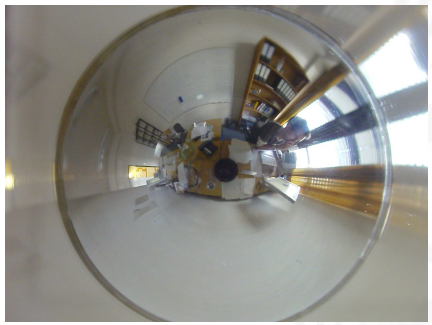


Evaluation / Discussion

Inapplicable in complex systems

In the domain of service robots arise following problems

- Grouping of sensor data to percepts
- Bottom-up approach is not sufficient
- Hierarchical partonomy / taxonomy


pan $= 3^{\circ}$; tilt $= 0^{\circ}$; Task: Find percept of table!

Towards Complex Perception
Contribution

- Constraint: insufficient Sensors and Algorithms
- Conventional solution: more sensors (modalities)

Action oriented feature based approach

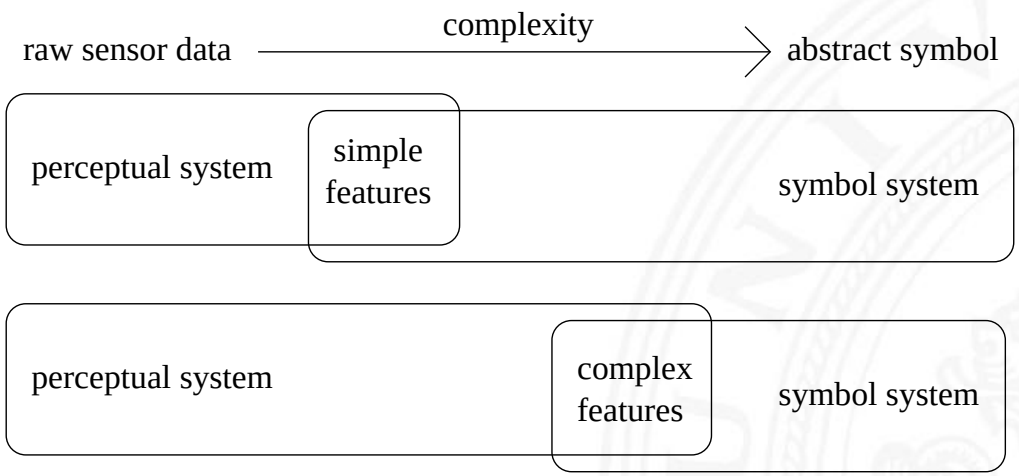
- Use simple features to narrow search space for complex features
- Use directed action for better utilization of given sensors

Outline

Object Perception via Feature Recognition

Objects are perceived via sets of features that may

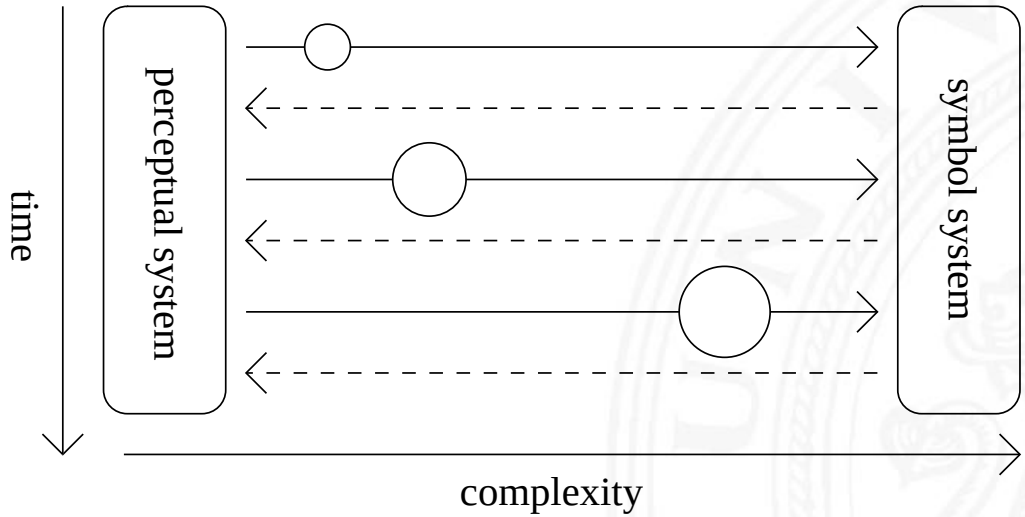
- deliver different amount of information
- differ in complexity



Complexity of features is inverse relational to information gain

Iterative Knowledge Acquisition

1. Recognition of feature
2. Extract assumptions that narrow search space
3. Execute action if necessary
4. Go to step 1



Size of the circles denote the complexity of features measured by the perceptual system

Active Object Inspection

- Changing parameter of
  - camera focus, shutter, ...
  - operation mode of sensors, preprocessing algorithms, ...
- Position and orientation of flexible mounted sensors
  - camera with pan tilt unit
  - force sensor at manipulator
- Reduce distance to target object candidate

Outline

Example Objects

Table

- laserscan:table-leg
- mobile:driveTo
- ptu:lookAt
- ptucam:verifyLeg
- mobile:driveTo
- arm:move
- hand:perceiveForce

Door

- laserscan:indentation
- mobile:driveTo
- arm:move
- handcam:findDoorknob
- arm:move
- hand:pushDoor
- hand:perceiveForce

Example Objects

Table

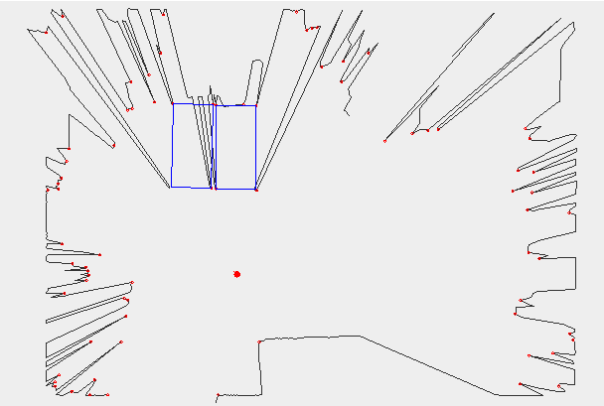
- laserscan:table-leg
- mobile:driveTo
- ptu:lookAt
- ptucam:verifyLeg
- mobile:driveTo
- arm:move
- hand:perceiveForce

Door

- laserscan:indentation
- mobile:driveTo
- arm:move
- handcam:findDoorknob
- arm:move
- hand:pushDoor
- hand:perceiveForce

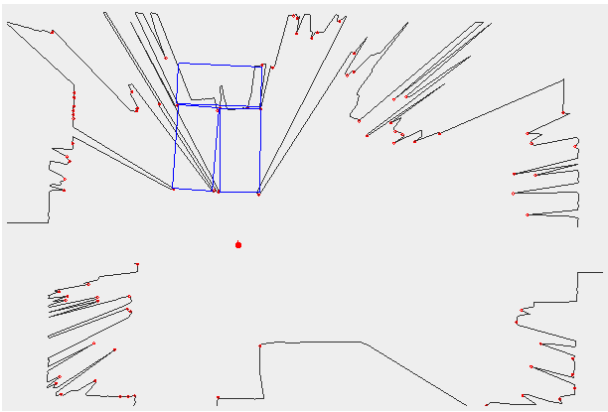
Table Anchoring Experiment

Proof of concept



Somewhere in the lab

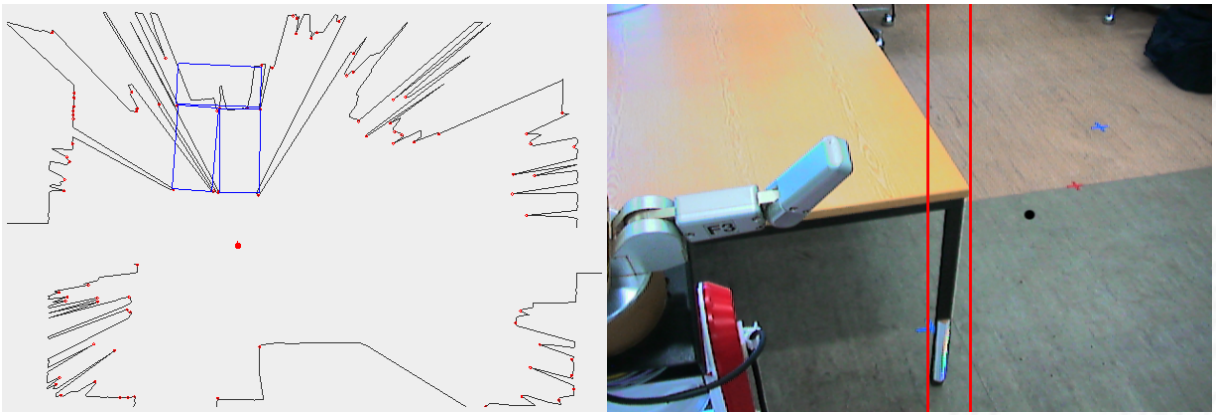
Table Anchoring Experiment
Proof of concept



In front of table

Table Anchoring Experiment

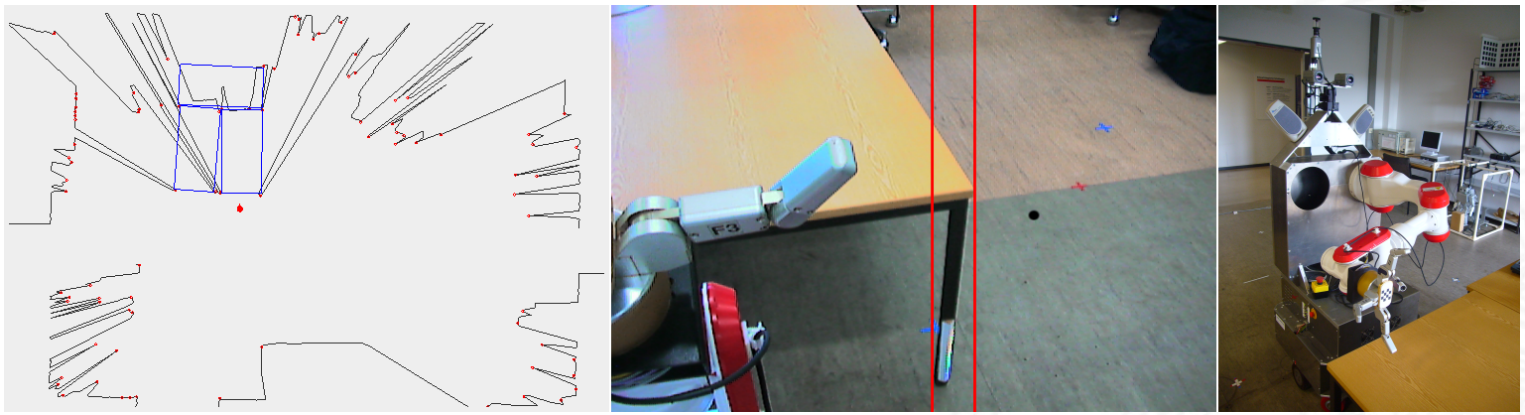
Proof of concept



Grab image

Table Anchoring Experiment

Proof of concept



Push table

Summary

- Current anchoring frameworks are not sufficient for complex multimodal perception
- Sequence of feature recognition allow use of more complex features
- Action in the service of perception improve reliability and robustness of object recognition

Thank you for your attention!

Martin Weser, Jianwei Zhang
{weser, zhang}@informatik.uni-hamburg.de

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics

Cross-modal Interactions in Natural and Artificial Cognitive Systems