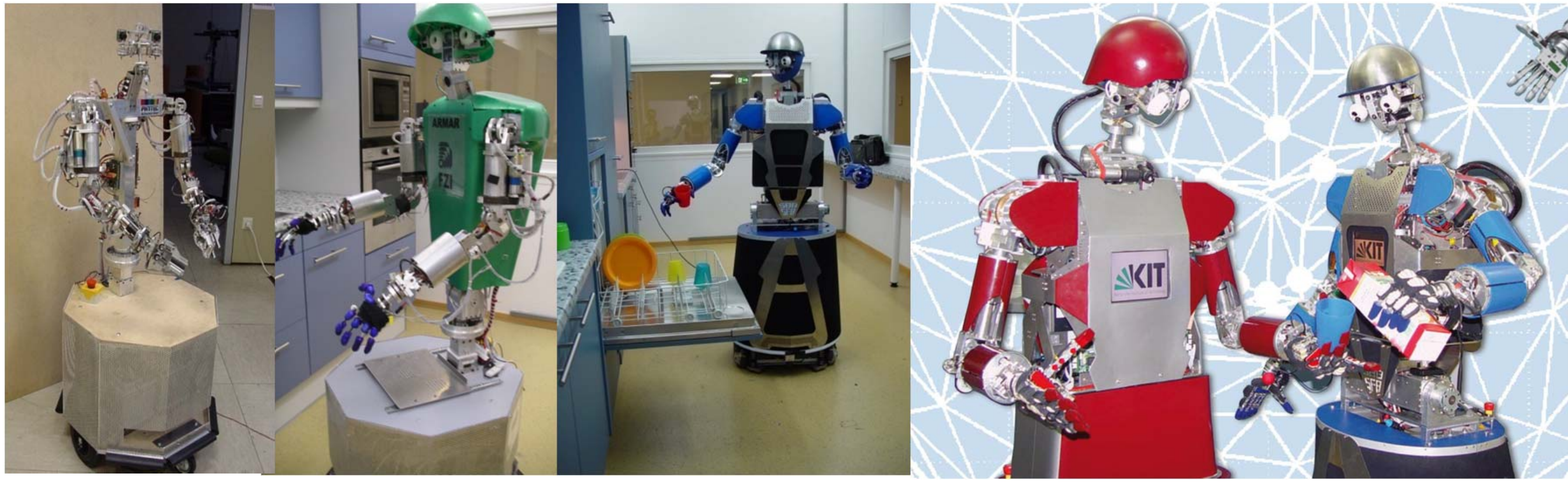
Karlsruhe Institute of Technology

ICRA 2011, Shanghai
13 May 2011

Active Perception for Mobile Manipulation

Tamim Asfour
Humanoids and Intelligence Systems Lab (Prof. Dillmann)

INSTITUTE FOR ANTHROPOMATICS, DEPARTMENT OF INFORMATICS



http://his.anthropomatik.kit.edu http://his.anthropomatik.kit.edu/english/65.php

Three key questions

- Grasping and manipulation in human-centered and open-ended environments
- Learning through observation of humans and imitation of human actions
- Interaction and natural communication

© SFB 588, Karlsruhe

Interactive tasks in the Robo-KITchen

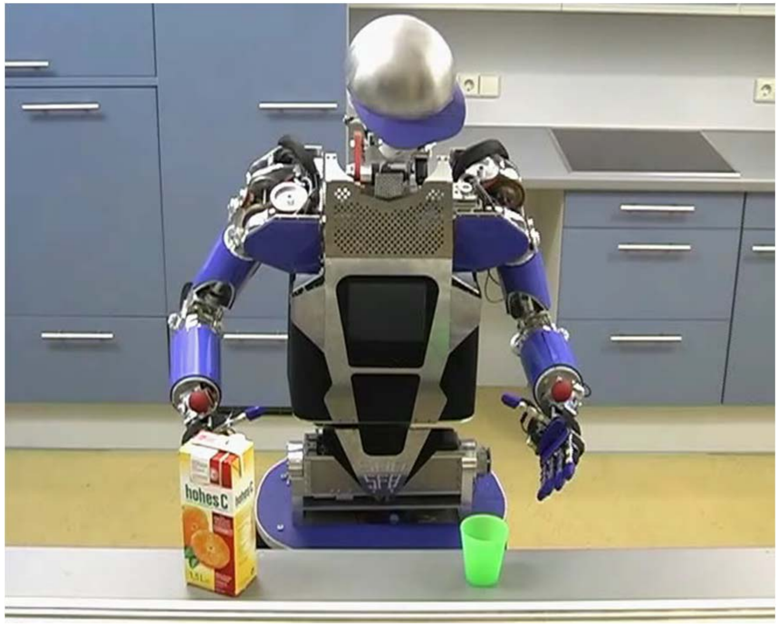
- Object recognition and localization
- Vision-based grasping
- Hybrid position/force control
- Vision-based self-localisation
- Collision-free navigation
- Combining force and vision for opening and closing door tasks
- Learning new objects, persons and words
- Audio-visual user tracking and localization
- Multimodal human-robot dialogs
- Speech recognition for continuous speech

[Humanoids 2006, IROS 2006, IROS 2007, RAS 2008, Humanoids 2008, Humanoids 2009]

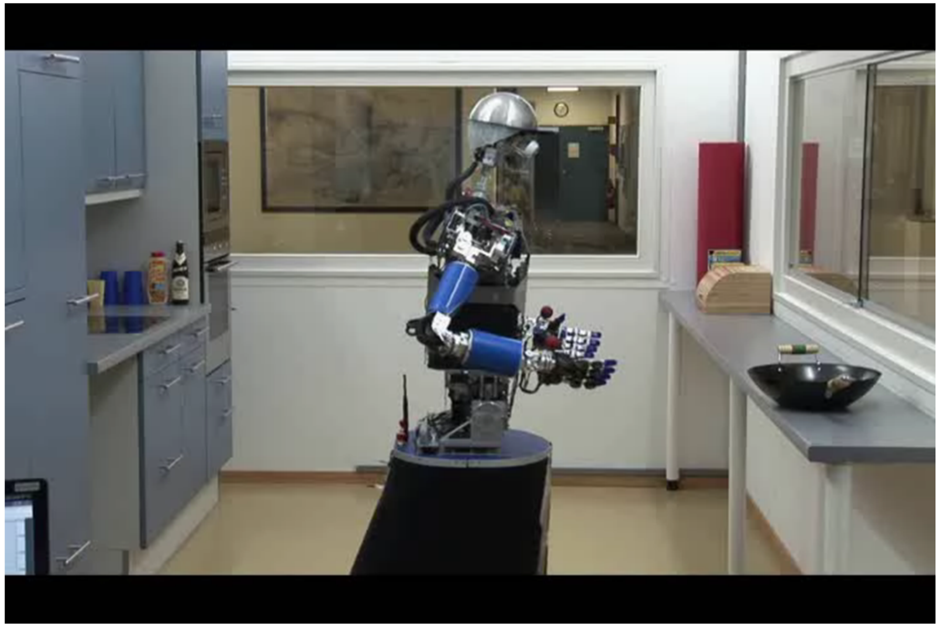
Bimanual grasping and manipulation

- Stereovision for object recognition and localization
- Visual Servoing for dual-hand grasping
- Zero-force control for teaching of grasp poses

Humanoids 2009



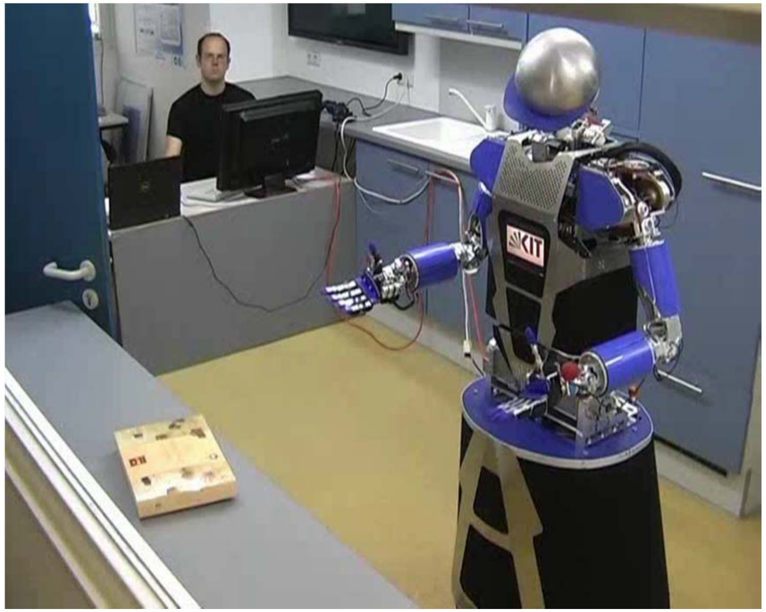
Loosely coupled dual-arm tasks



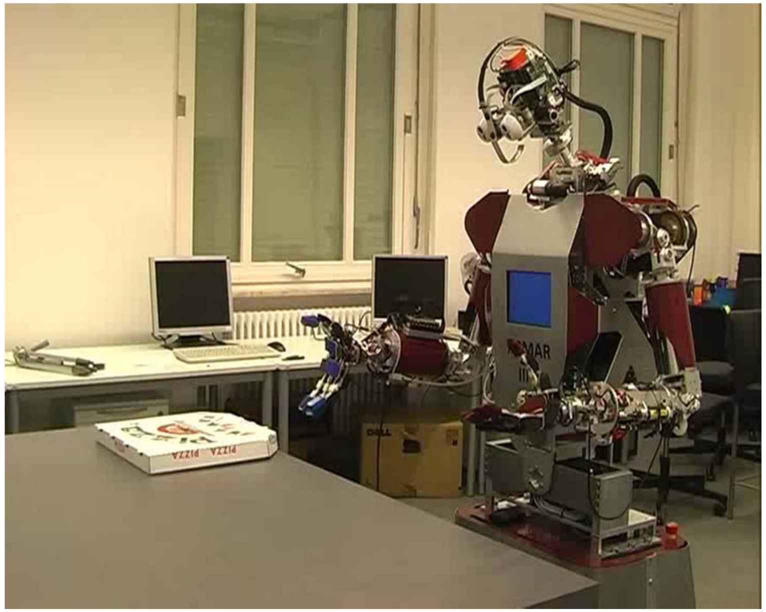
Tightly coupled dual-arm tasks

Pushing for grasping and pre-grasp manipulation

- Flat objects on the table are difficult to grasp → Pre-grasp manipulation to adjust the object before final grasping
- Learning actions on objects



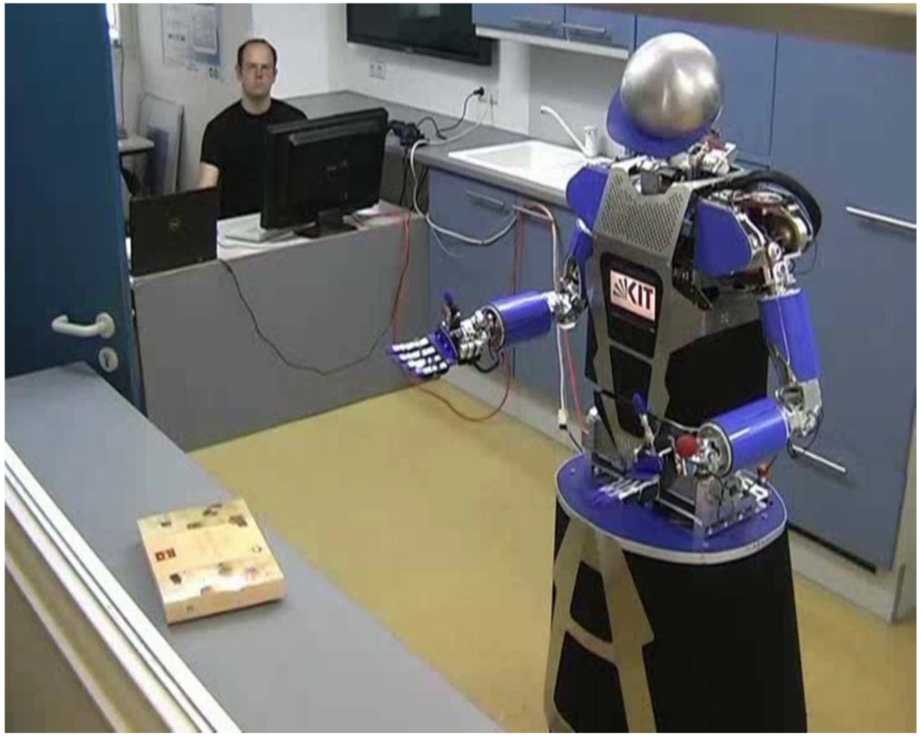
Joint work with Damir Omrcen and Ales Ude; (Humanoids 2009)



Joint work with Lillian Chang and Nancy Pollard (CMU), Humanoids 2010

Pushing for grasping

- Object independent pushing (generalization across objects).
- Learning relationship between point and angle of push and the actual movement of an object
- Direct association between the binarized object image and the response of the object with respect to the applied pushing action.
- Use the knowledge in order to find the appropriate point and angle of push in order to bring an object to a goal



Joint work with Damir Omrcen and Ales Ude; (Humanoids 2009)

- **Workshop Motion Planning on Monday:**

  Grasp and motion planning

- **Today: Mobile manipulation workshop**

  Active Perception for autonomous knowledge acquisition

Limitations and shortcuts

- Objects
  - Complete model knowledge (shape, color, texture)
  - Only visual object representation is used

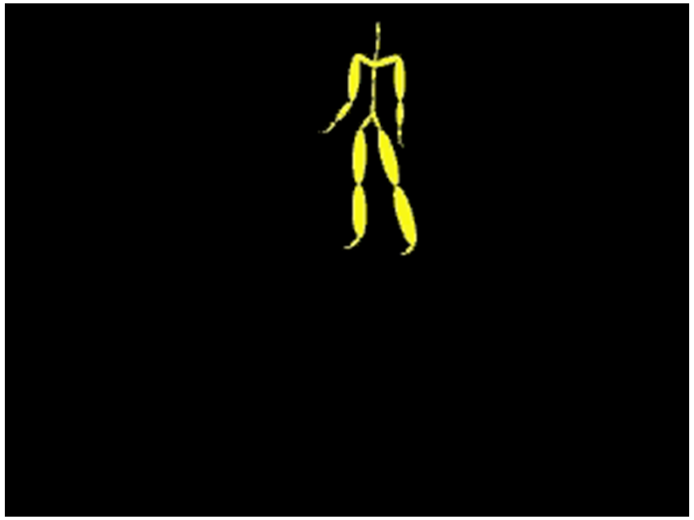
  - How to acquire multi-sensory representations of objects?
  - How to learn objects representations?
  - …

- Actions
  - Kinematic-based approaches as place holders for learned primitive actions

  - How to learn new actions?
  - How to adapt actions to new situations?
  - How to chain different actions to achieve complex tasks?
  - How to learn associations between objects and actions?
  - How to learn consequence of action?
  - …

Objects and Actions are inseparably intertwined

- Visually based object recognition fails
- Visual information is sparse and limited
- Activity involving the object decreases the uncertainty about the object's nature considerably!



CMU Graphics Lab Motion Capture Database
http://mocap.cs.cmu.edu/

Intuitions on Object-Action Complexes

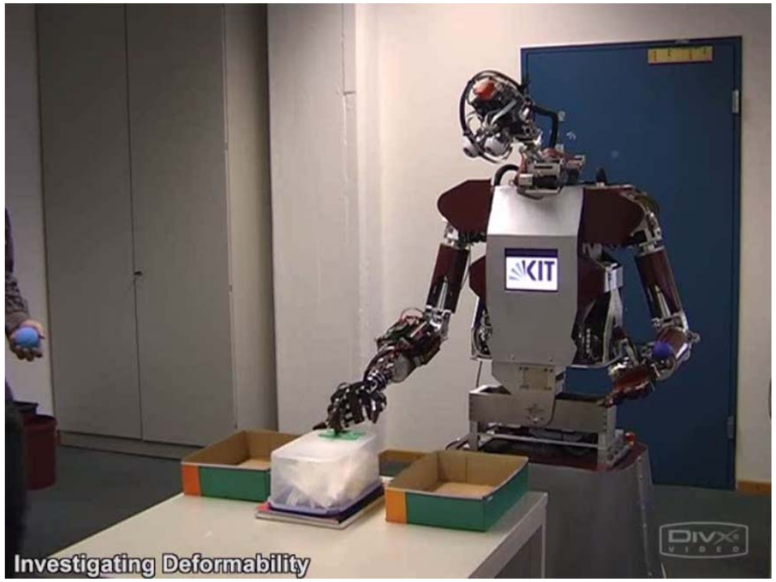
- Affordances (J.J. Gibson)
  - Objects affords actions
- Object-Action Complexes (OACs)
  - Actions define the meaning of Objects and
  - Objects suggest Actions
- OACs are associations of objects and affordances
  - Affordances can be expressed by STRIPS rules comprising:
    - Preconditions and
    - Deletions/additions

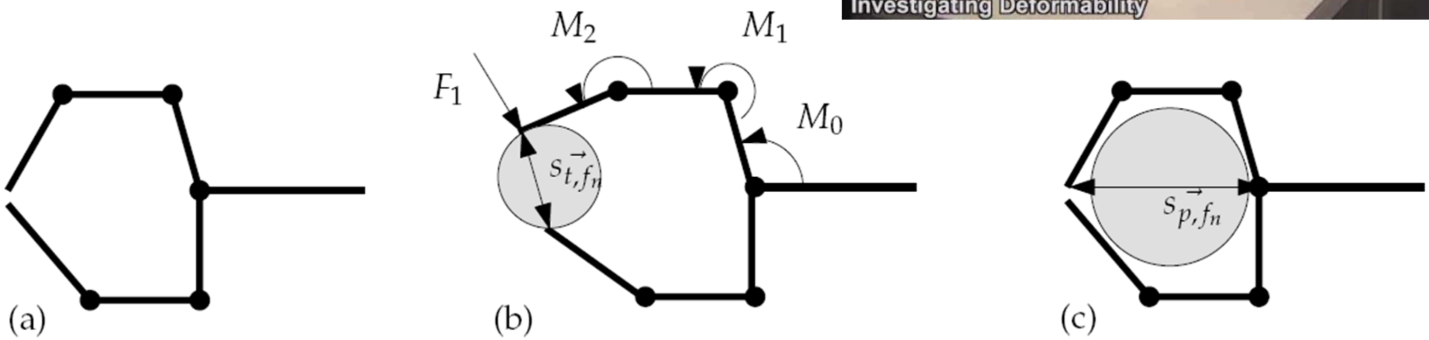
How to relax such limitations?

- Autonomous Exploration:
  - Visually-guided haptic exploration
  - Visual object search
- Coaching and Imitation
  - Learning from Observation
  - Goal-directed Imitation

Hand: available skills

- Direct/Inverse Kinematics
- Position/force control [Humanoids 2009]
- Detection of contact and "objectness"
- Assessment of deformability and object size





Failed grasp "no object"

Precision grasps: Distance between fingertips

Power grasps: Distance between fingertips and palm

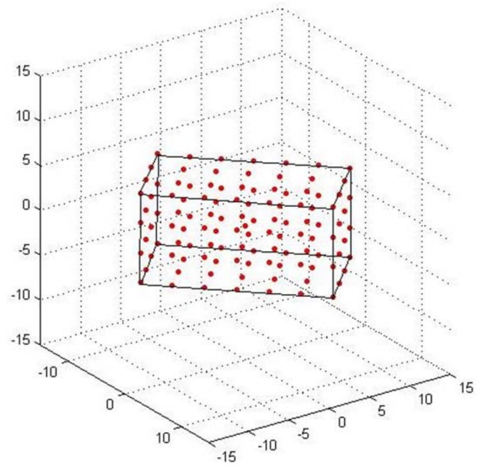
Tactile exploration using potential fields

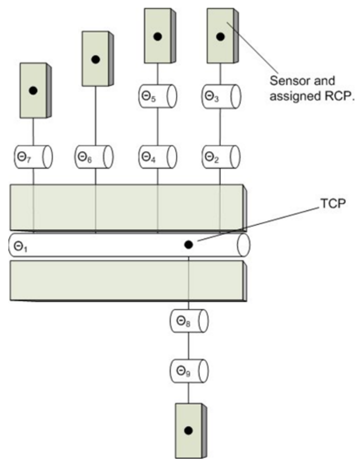
- Potential field in operational space
  - Unknown regions → attractive $\Phi_a < 0$
  - Known regions → repellent $\Phi_r > 0$
- Dynamic adaptation of potential field configuration based on tactile response
- Superposition of individual potential sources

$$\Phi(x) = \sum_i \Phi_{r,i}(x) + \sum_j \Phi_{a,j}(x)$$

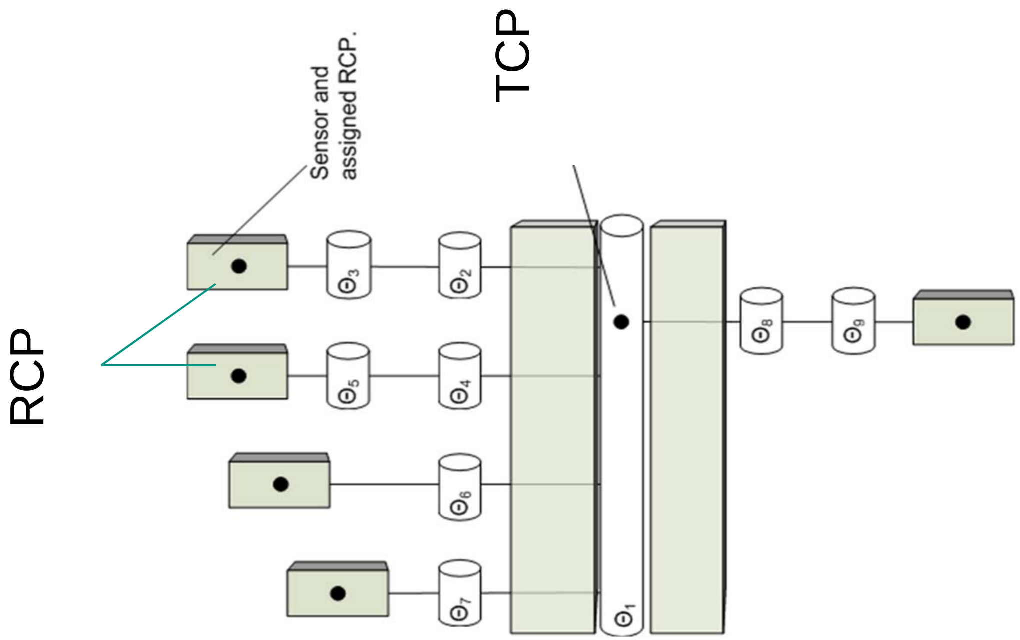
- Generation of trajectories for the fingertips by gradient methods

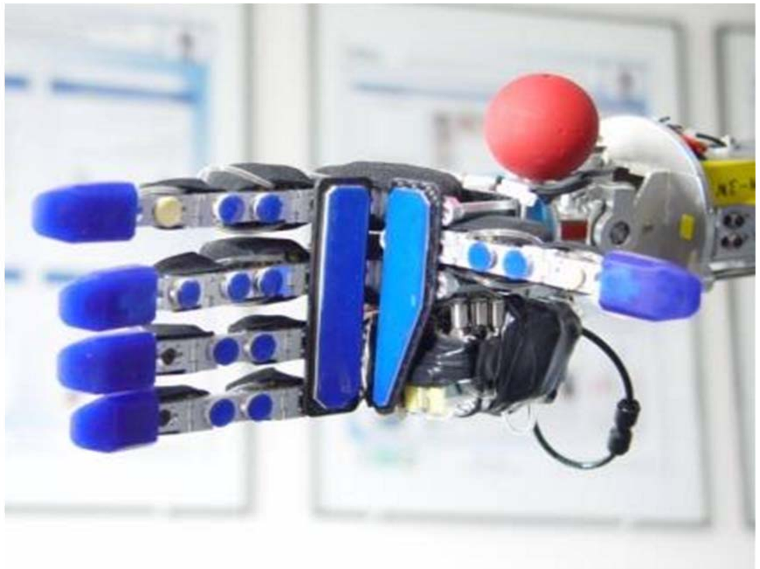
$$F = -\nabla \phi(x)$$





Hand in details

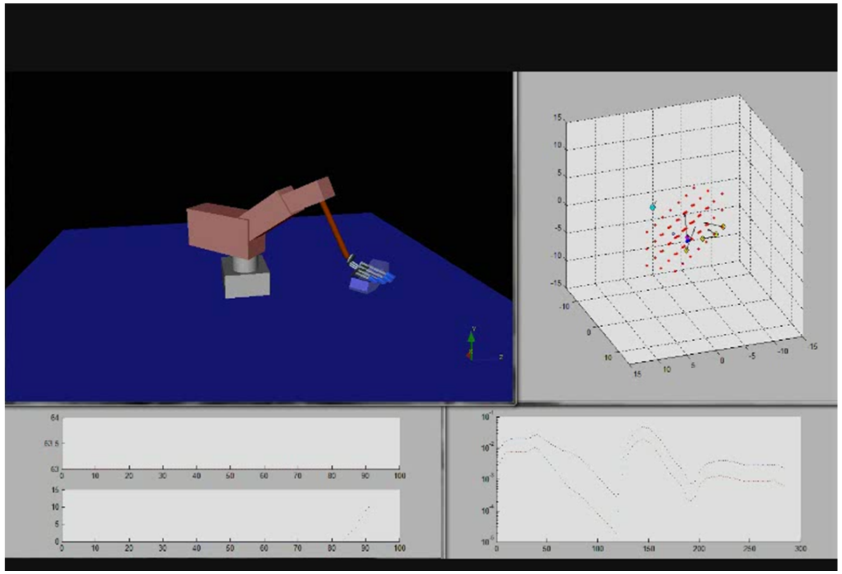


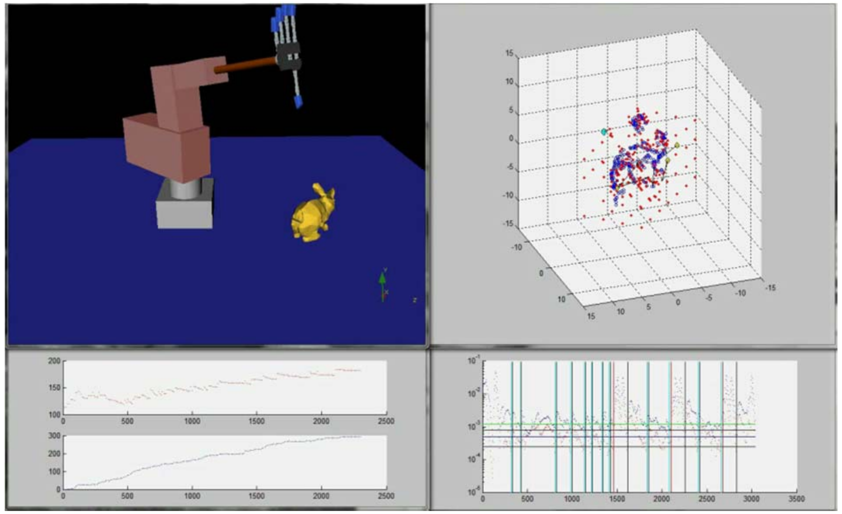


Tactile object exploration

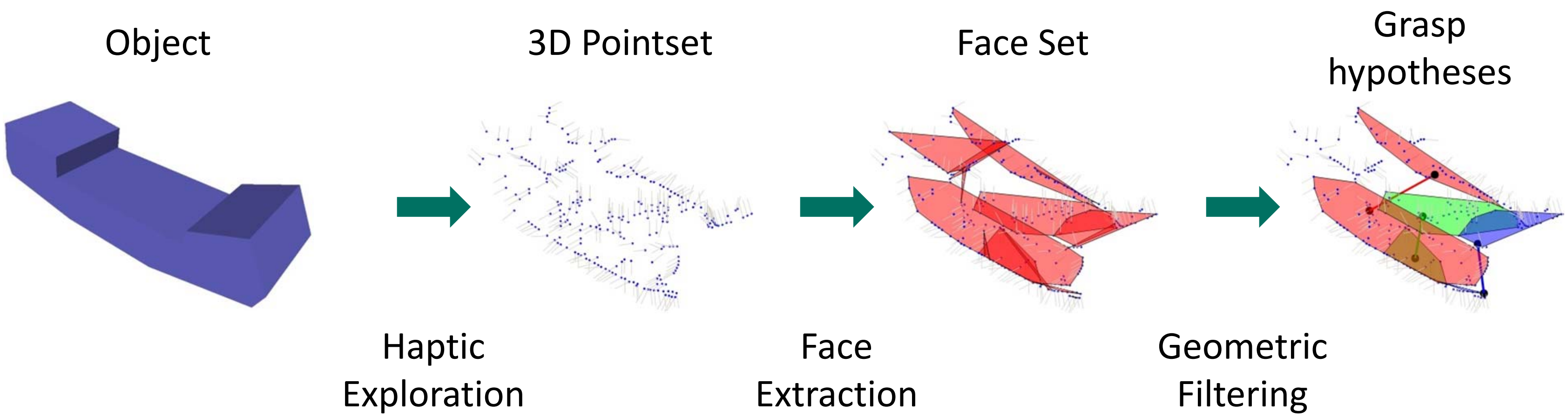
- Potential field guides the robot hand along the object surface
- Oriented 3D point cloud from contact data
- Compute face pairings from 3D point
- Calculate grasping hypotheses using a geometric feature filter pipeline
  - Parallelism
  - Minimum face size
  - Face distance
  - Mutual visibility
- Evaluation of grasp qualities

→ Association between “objects” and actions (grasps) → symbolic grasps (grasp affordances)

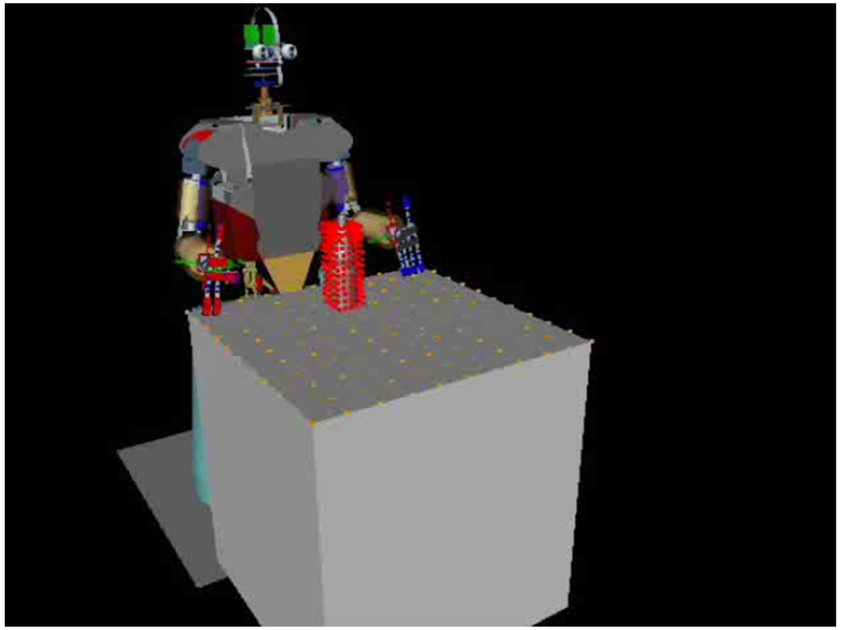


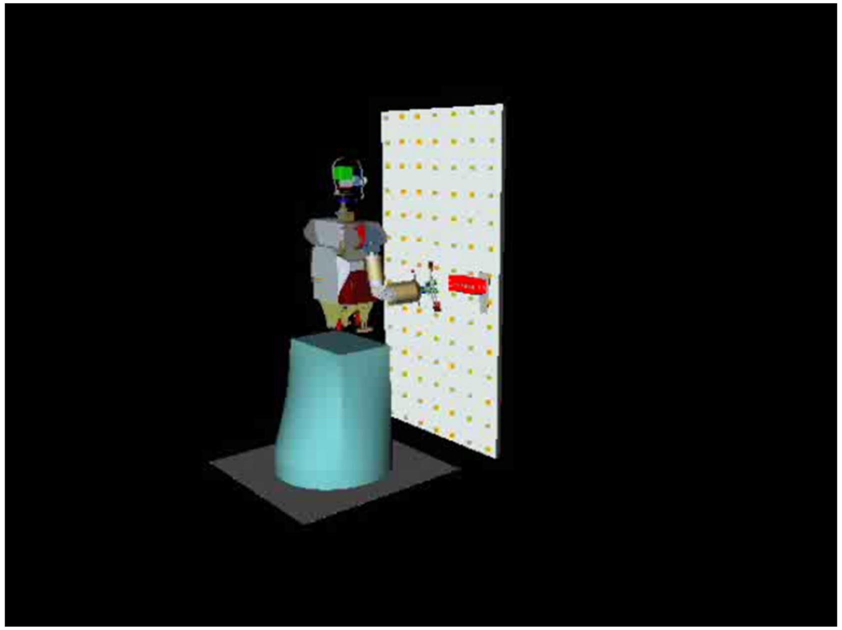


Generation of grasp hypotheses



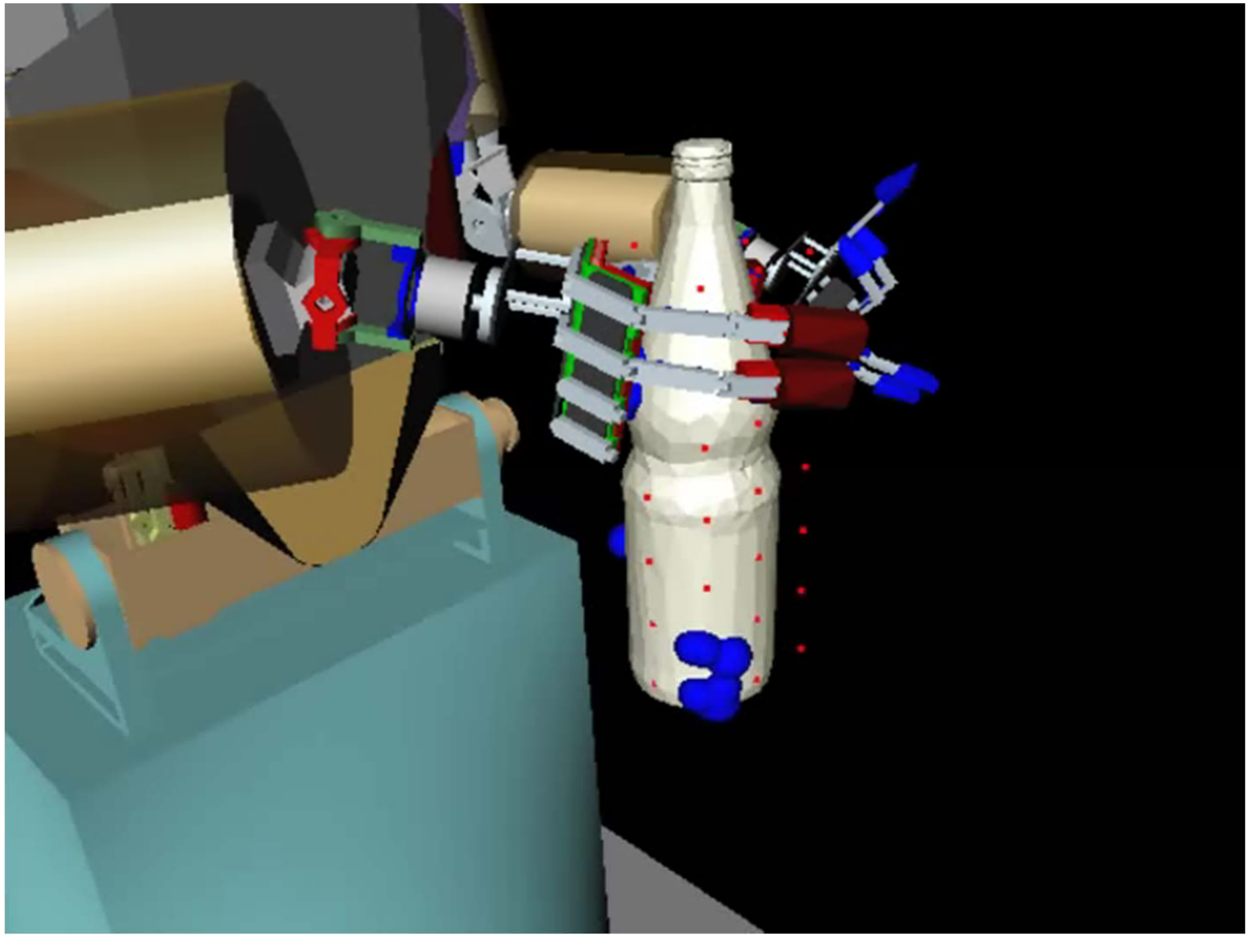
Haptic exploration on ARMAR



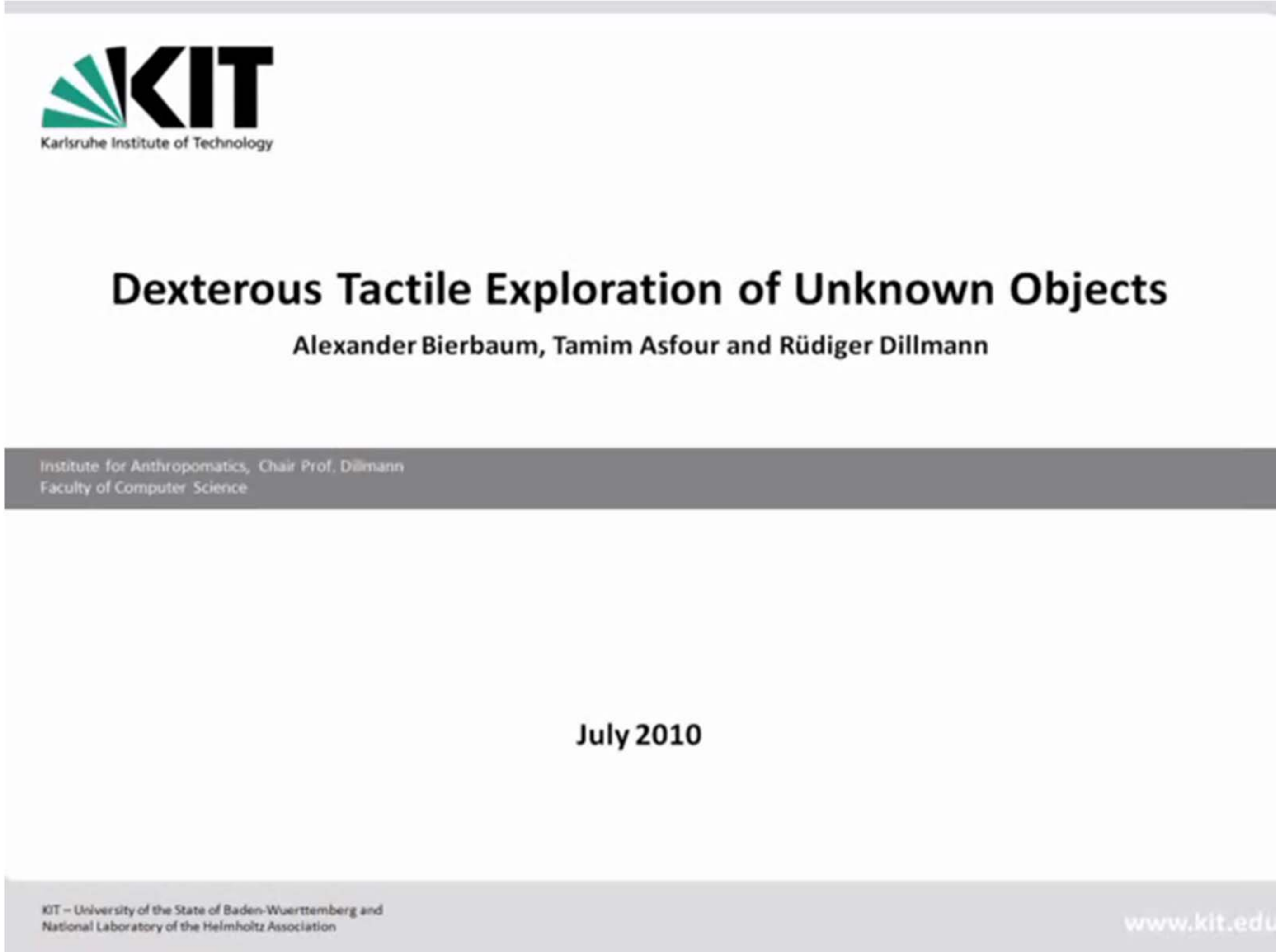


- In simulation
  - Physics extension for Open Inventor/VRML modeling of complex mechanical systems
  - Modeling of virtual sensors
  - VMC-based inverse kinematics

Haptics exploration on ARMAR

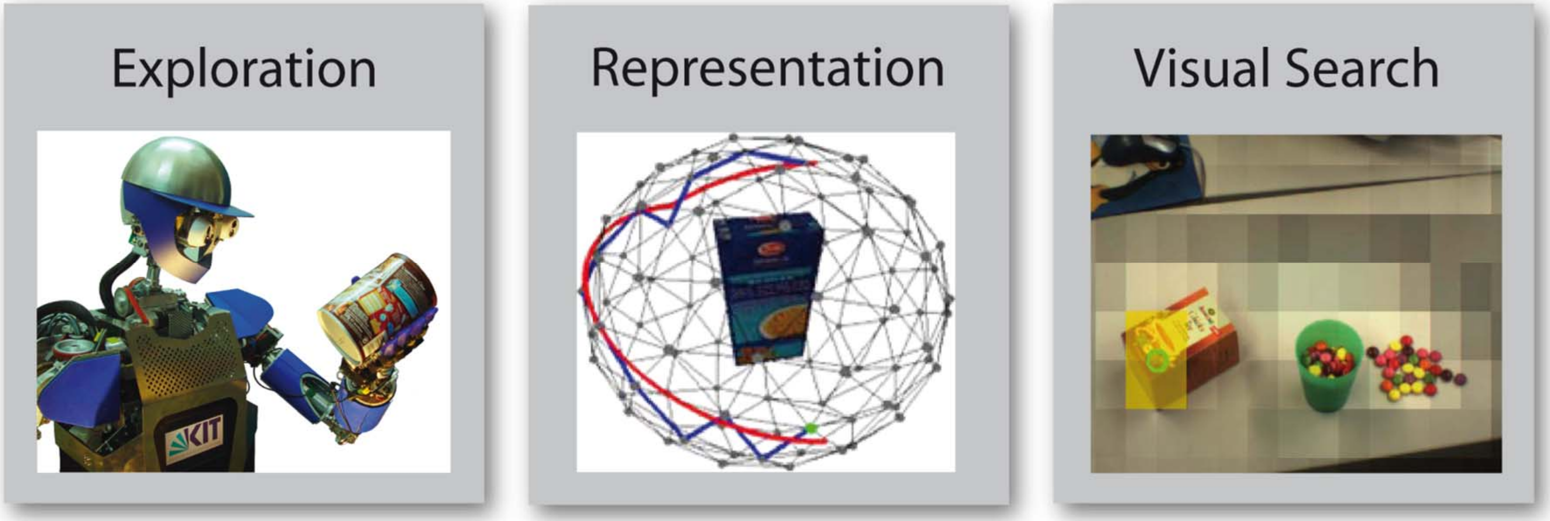


Haptic Exploration using ARMAR-III



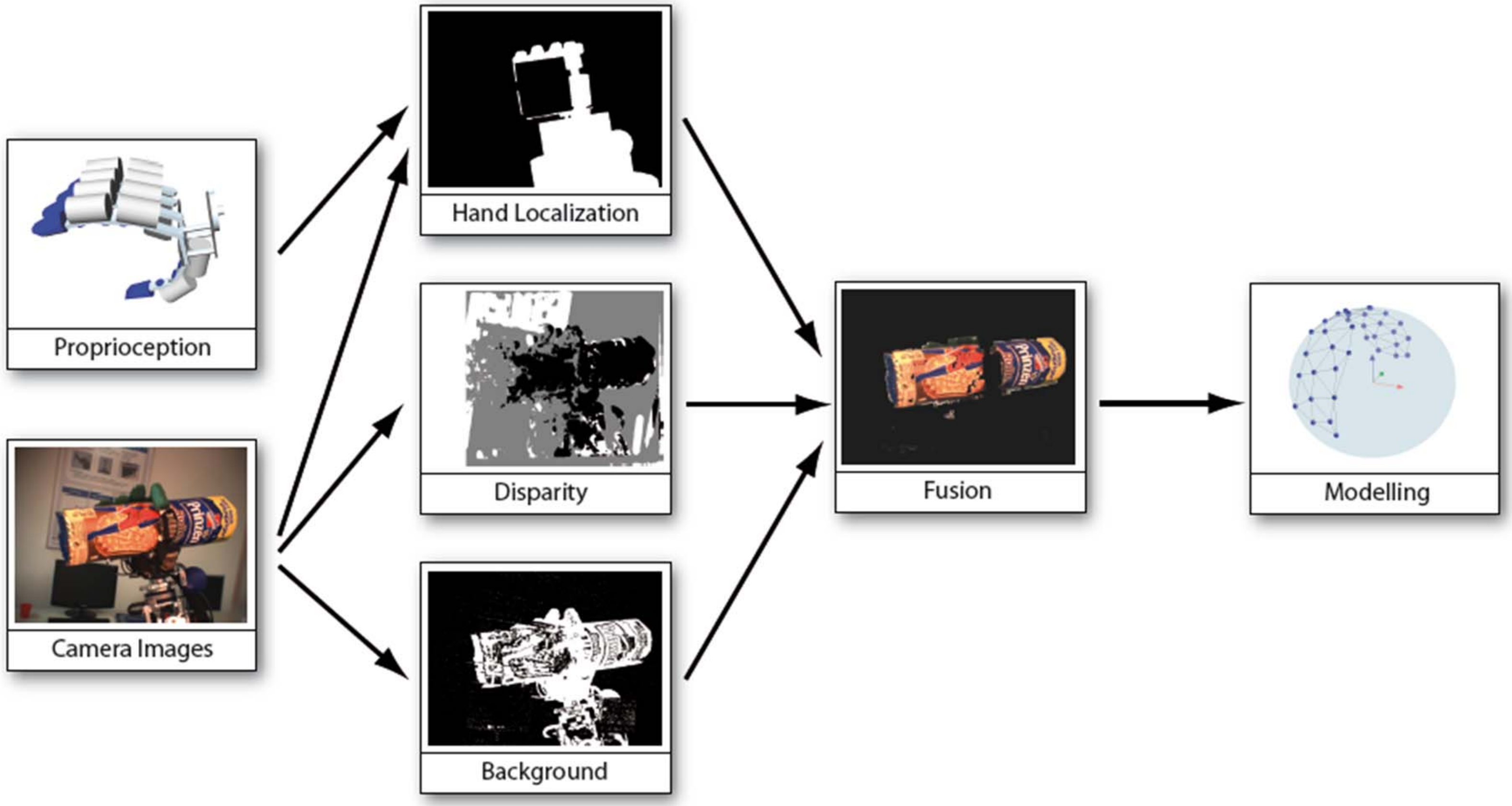
Generation of multi-view object representations

- Generation of visual representations through manipulation
- Application of generated representations in recognition tasks and visual search tasks



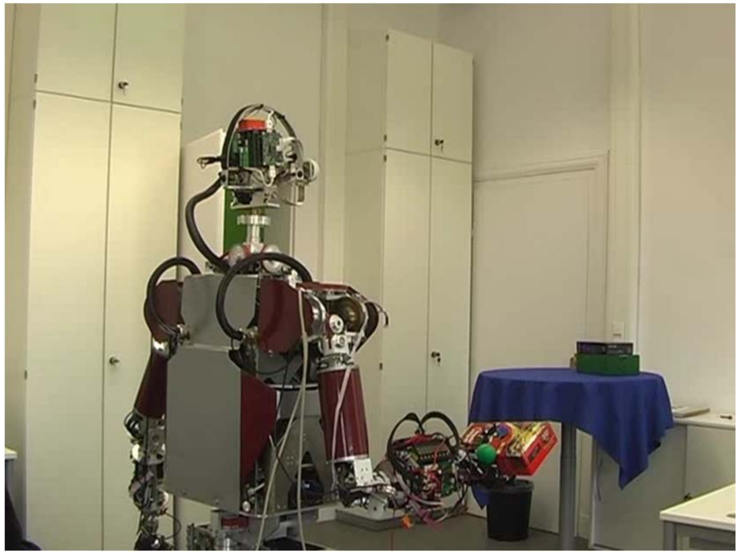
Acquisition of multiple object views

- Background-object and hand-object segmentation



Active visual search

- Generation of different views through manipulation
- Active search using perspective and foveal camera
- Integration of object hypotheses in an ego-centric representation (scene memory)

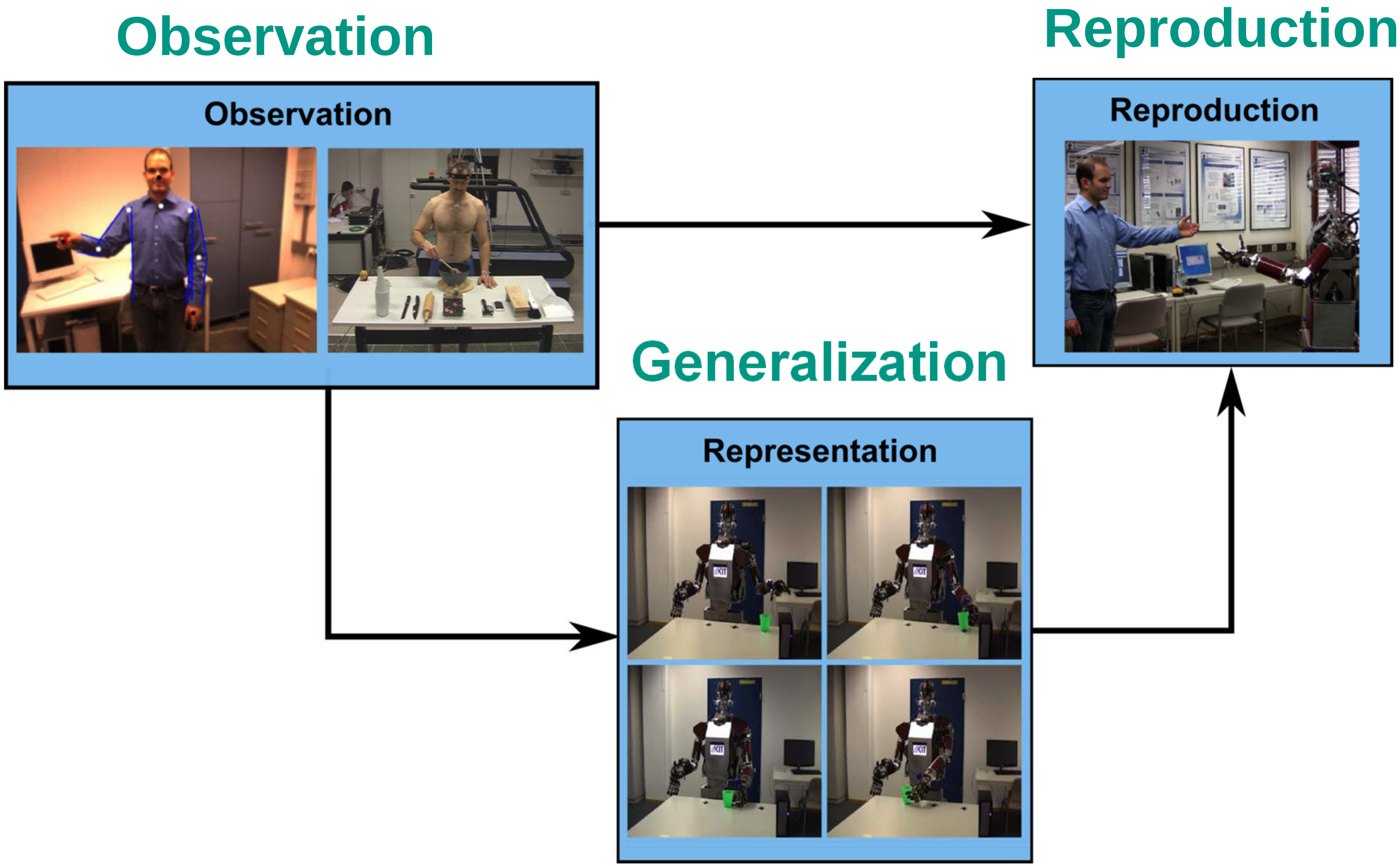


ICRA 2010
Humanoids 2009
ICRA 2009

How to relax the limitations in our scenario?

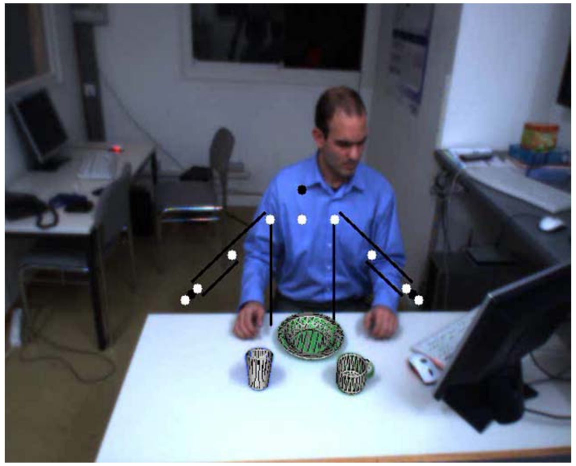
- Autonomous Exploration
  - Visually-guided haptic exploration
  - Visual object search
- Coaching and Imitation
  - Learning from Observation
  - Goal-directed Imitation

Learning from human observation



Stereo-based 3D Human Motion Capture (HMC)

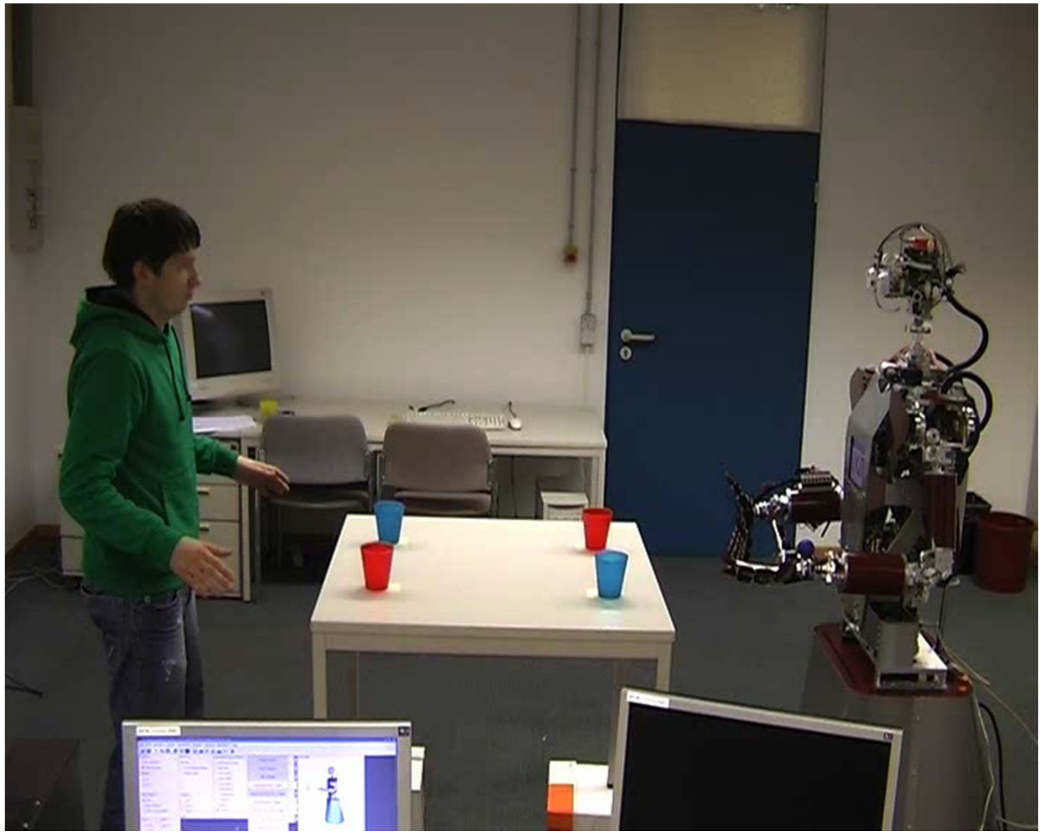
- Capture 3D human motion based on the image input from the cameras of the robot's head only
- Approach
  - Hierarchical Particle Filter framework
  - Localization of hands and head using color segmentation and stereo triangulation
  - Fusion of 3d positions and edge information
  - Half of the particles are sampled using inverse kinematics
- Features
  - Automatic Initialization
  - 30 fps real-time tracking on a 3 GHz CPU, 640x440 images
  - Smooth tracking of real 3d motion



Reproduction on ARMAR

- Tracking of human and object motion
- Visual servoing for grasping

Generalisation?



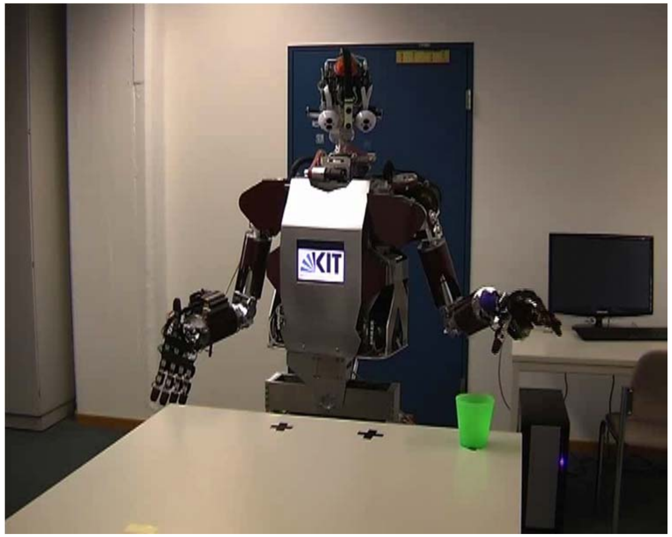
Action representation

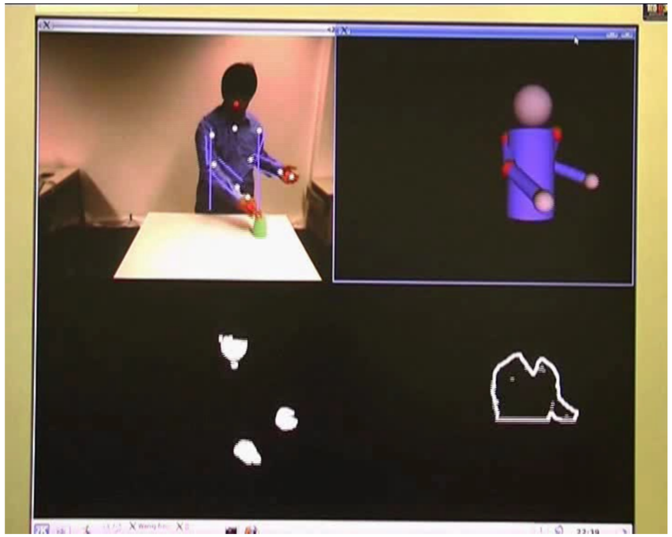
- Hidden Markov Models (HMM) Humanoids 2006, IJHR 2008
  - Extract key points (KP) in the demonstration
  - Determine key points that are common in multiple demonstrations (common key points: CKP)
  - Reproduction through interpolation between CKPs
- Dynamic movement primitives (DMP) ICRA 2009, T-RO 2010
  - Ijspeert, Nakanishi & Schaal, 2002
  - Trajectory formulation using canonical systems of differential equations
  - Parameters are estimated using locally weighted regression
- Spline-based representations Humanoids 2007
  - fifth order splines that correspond to minimum jerk trajectories to encode the trajectories
  - Time normalize the example trajectories
  - Determine common knot points so that all example trajectories are properly approximated. Similar to via-point, key-points calculation.

Learning from Observation

- Library of motor primitives
  - Markerless human motion tracking
  - Object tracking

- Action representation
  - Dynamic movement primitives for generating discrete movements
  - Adaptation of dynamic systems to allow sequencing of movement primitives
  - Associating semantic information with DMPs
  - → sequencing of movement primitives
  - → Planning

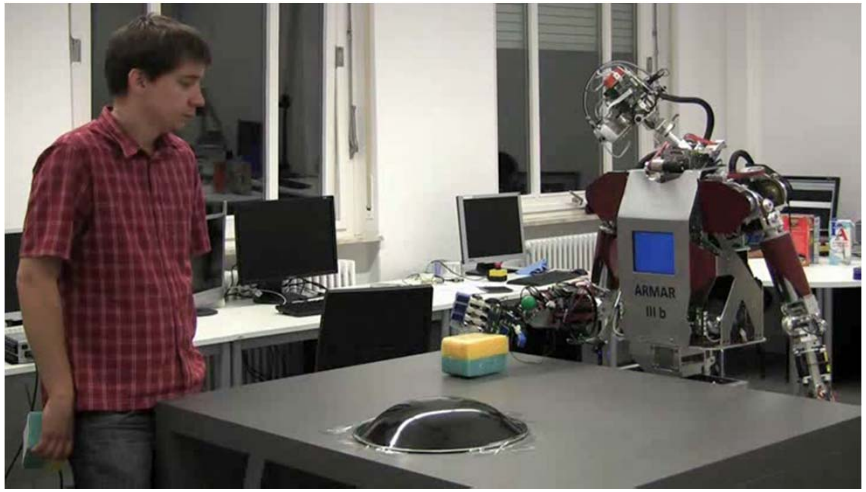




Learning from Observation

- Periodic movements (Wiping)
  - Extract the frequency and learn the waveform.
  - Incremental regression for waveform learning

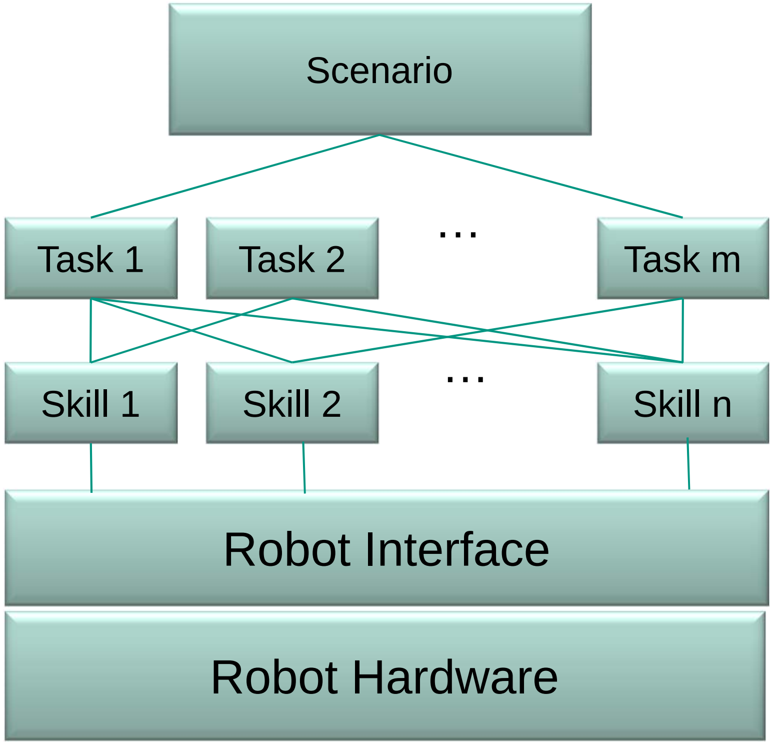
- Adaptation of the learned movement to maintain contact with different surfaces, based on force-feedback.



Joint work with Andrej Gams and Ales Ude, Humanoids 2010

API of the ARMARs

- Robot interface (API) allows access to the robot's sensors and actors via C++ variables and method calls
- Skills: implemented atomic capabilities of the robot
  - SearchForObject
  - Grasp
  - Place
  - HandOver
  - Open/Close door
  - …
- Tasks: combination of several skills for a specific purpose
  - Bring object from …
  - Put object on
  - Stack objects
  - …



Connecting High-level Task Planning to Execution

Karlsruhe Institute of Technology

PKS planner (STRIPS-like planner)

Joint work with Ron Patrick and Mark Steedman, University of Edinburgh

Thanks

- Humanoids@KIT
  - Rüdiger Dillmann
  - Stefan Ulbrich
  - David Gonzalez
  - Manfred Kröhnert
  - Niko Vahrenkamp
  - Julian Schill
  - Kai Welke
  - Ömer Terlemez
  - Alex Bierbaum
  - Martin Do
  - Markus Przybylski
  - Tamim Asfour
  - Pedram Azad (not in the picture)
  - Paul Holz (not in the picture)
  - Ioana Gheta (not in the picture)
  - Christian Böge (not in the picture)
  - Sebastian Schulz (not in the picture)
  - Isabelle Wappler

Karlsruhe Institute of Technology

Thank you ...

... for your attention

- Thanks to the funding agencies
  - German Research Foundation (DFG)
    - SFB 588 www.sfb588.uni-karlsruhe.de
  - European Commission
    - Xperience www.xperience.org
    - GRASP www.grasp-project.eu
    - PACO-PLUS www.paco-plus.org

Deutsche Forschungsgemeinschaft
DFG

CogSys
Cognitive Systems