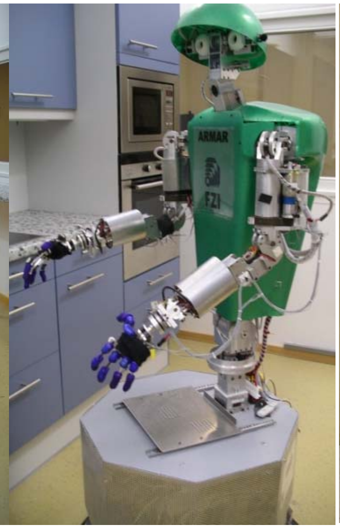
Humanoid Grasping and Manipulation

T. Asfour, A. Ude, N. Krueger, J. Piater, D. Kragic, R. Dillmann
N. Vahrenkamp, D. Omrcen, K. Huebner, M. Popovic, P. Azad, K. Welke, M. Do, J. Romero

Institute for Anthropomatics, Computer Science Department,
Humanoids and Intelligence Systems Lab (Prof. Dillmann)






wwwiaim.ira.uka.de

www.sfb588.uni-karlsruhe.de

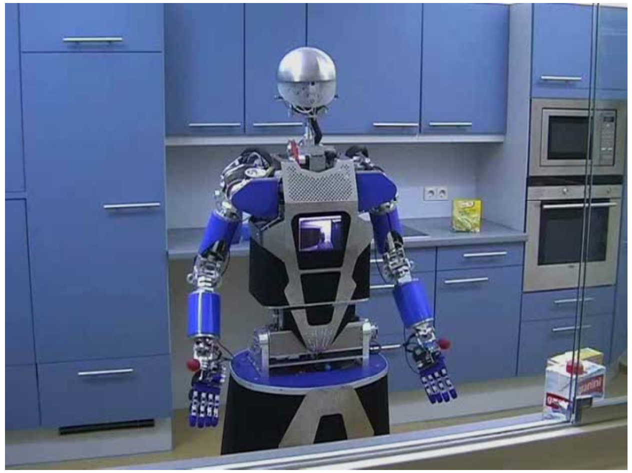
Three key questions

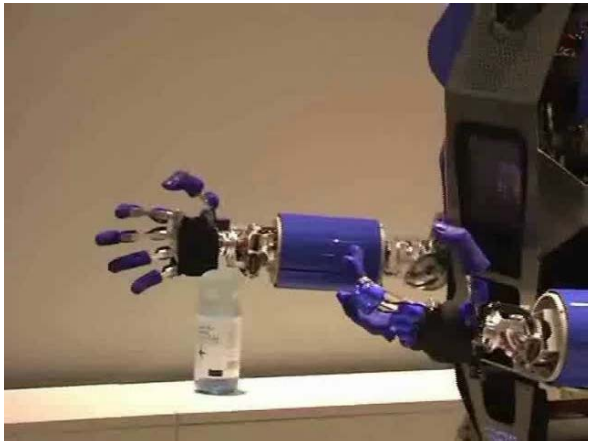
- Grasping and manipulation in human-centered and open-ended environments
- Learning through Observation of humans and imitation of human actions
- Interaction and natural communication

© SFB 588, Karlsruhe

Humanoid Robot ARMAR-III

- 7 DOF head with foveated vision
  - 2 cameras in each eye
  - 6 microphones
- 7-DOF arms
  - Position, velocity and torque sensors
  - 6D FT-Sensors
  - Sensitive Skin
- 8-DOF Hands
  - Pneumatic actuators
  - Weight 250g
  - Holding force 2,5 kg
- 3 DOF torso
  - 2 Embedded PCs
  - 10 DSP/FPGA Units
- Holonomic mobile platform
  - 3 laser scanner
  - 3 Embedded PCs
  - 2 Batteries

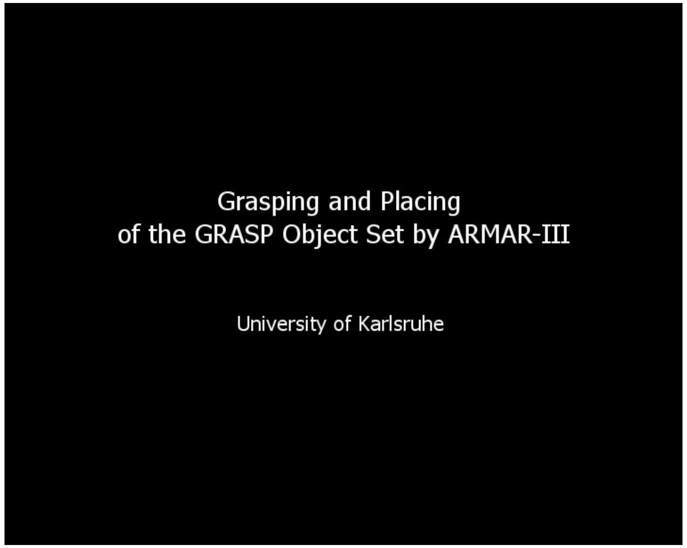




Fully integrated autonomous humanoid system

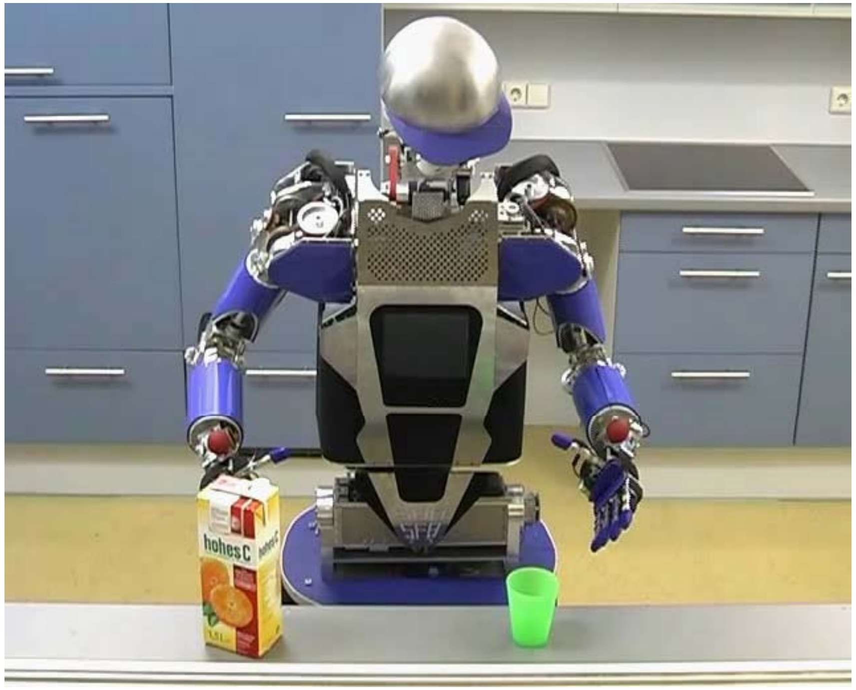
Integrated implementation

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

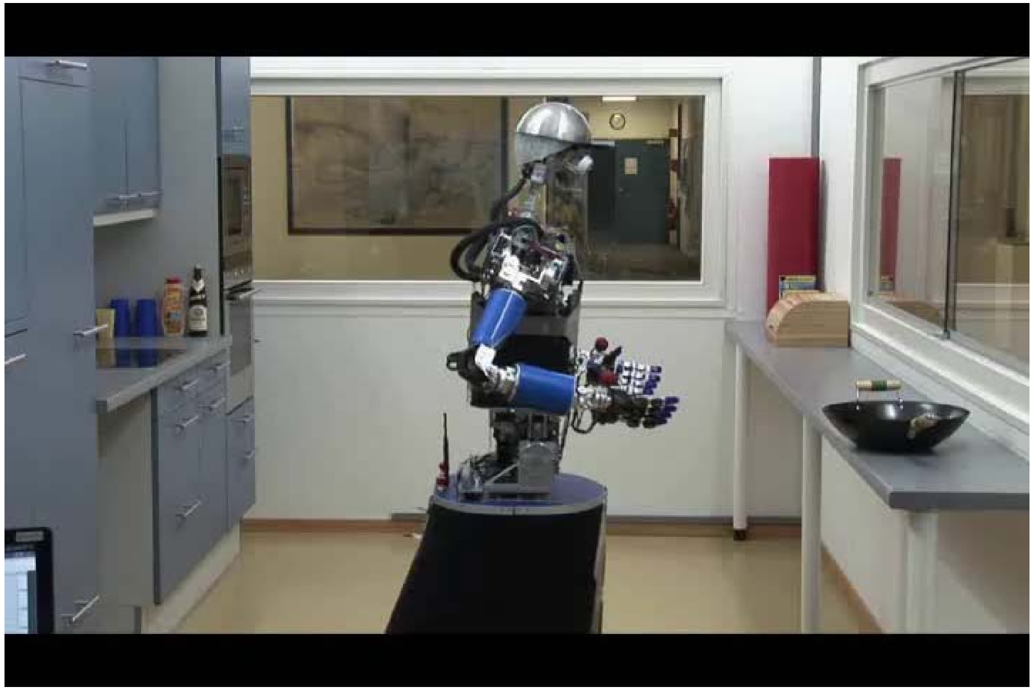


Bimanual grasping and manipulation

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

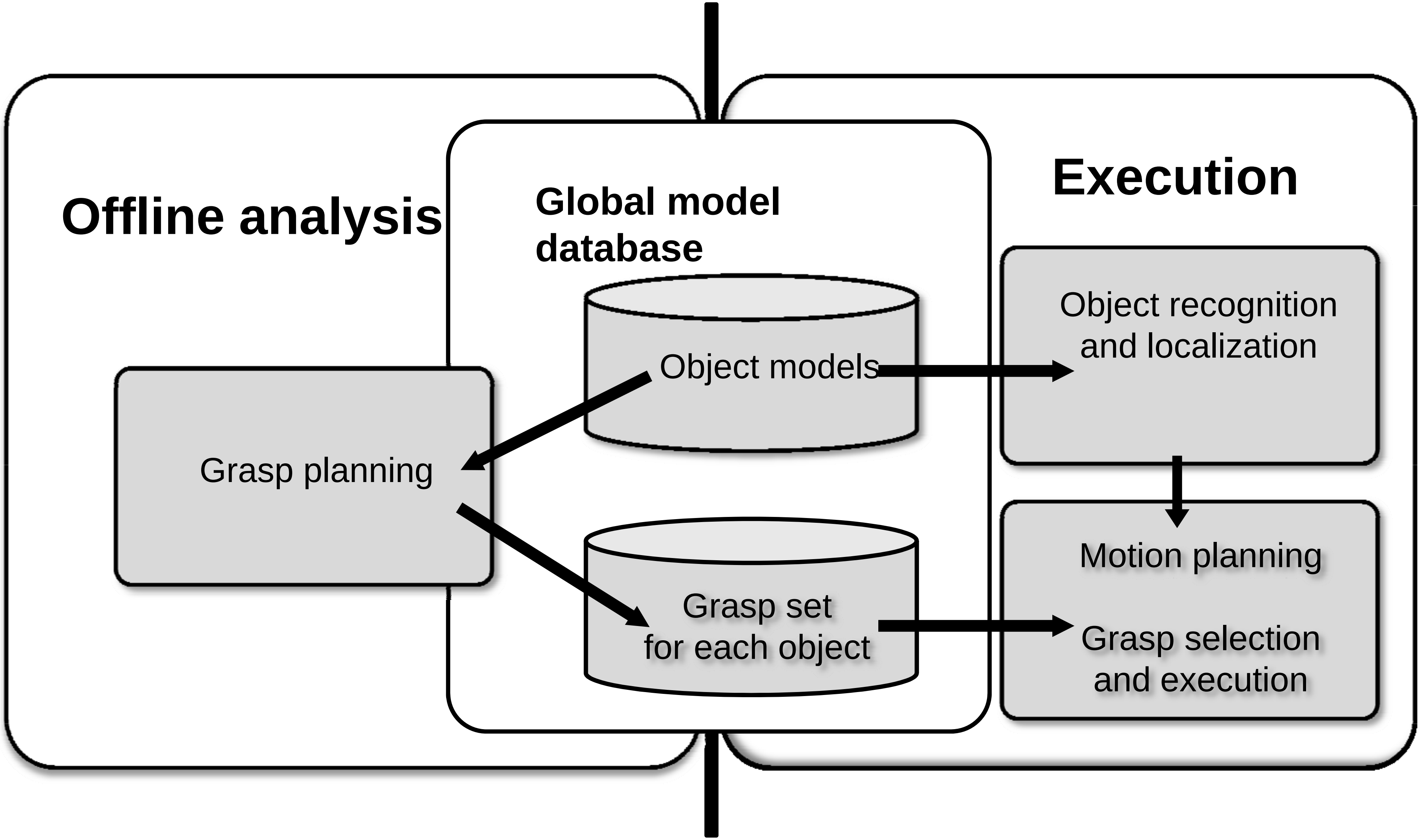


Loosely coupled dual-arm tasks



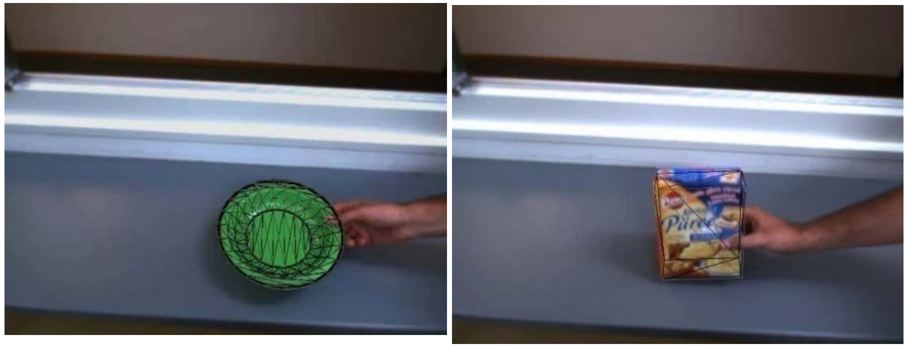
Tightly coupled dual-arm tasks

Grasping known objects: System architecture



Object recognition and localization

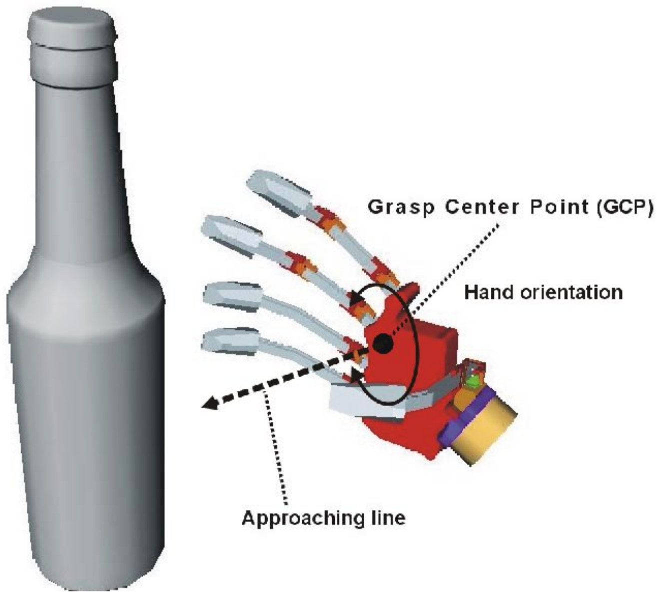
- Colored objects (IROS 2006, IROS 2009)
  - Segmentation by color
  - Appearance-based recognition using a global approach
  - Model-based generation of view sets
  - Combination of stereo vision and stored orientation information for 6D pose estimation
- Textured objects (Humanoids 2006, IROS 2009)
  - Recognition using local features
  - Calculation of consistent features with respect to the pose of the object using the Hough transform
  - 2D-localization using image point correspondences
  - 6D pose estimation using stereo vision



Correspondences between learned view and view in scene

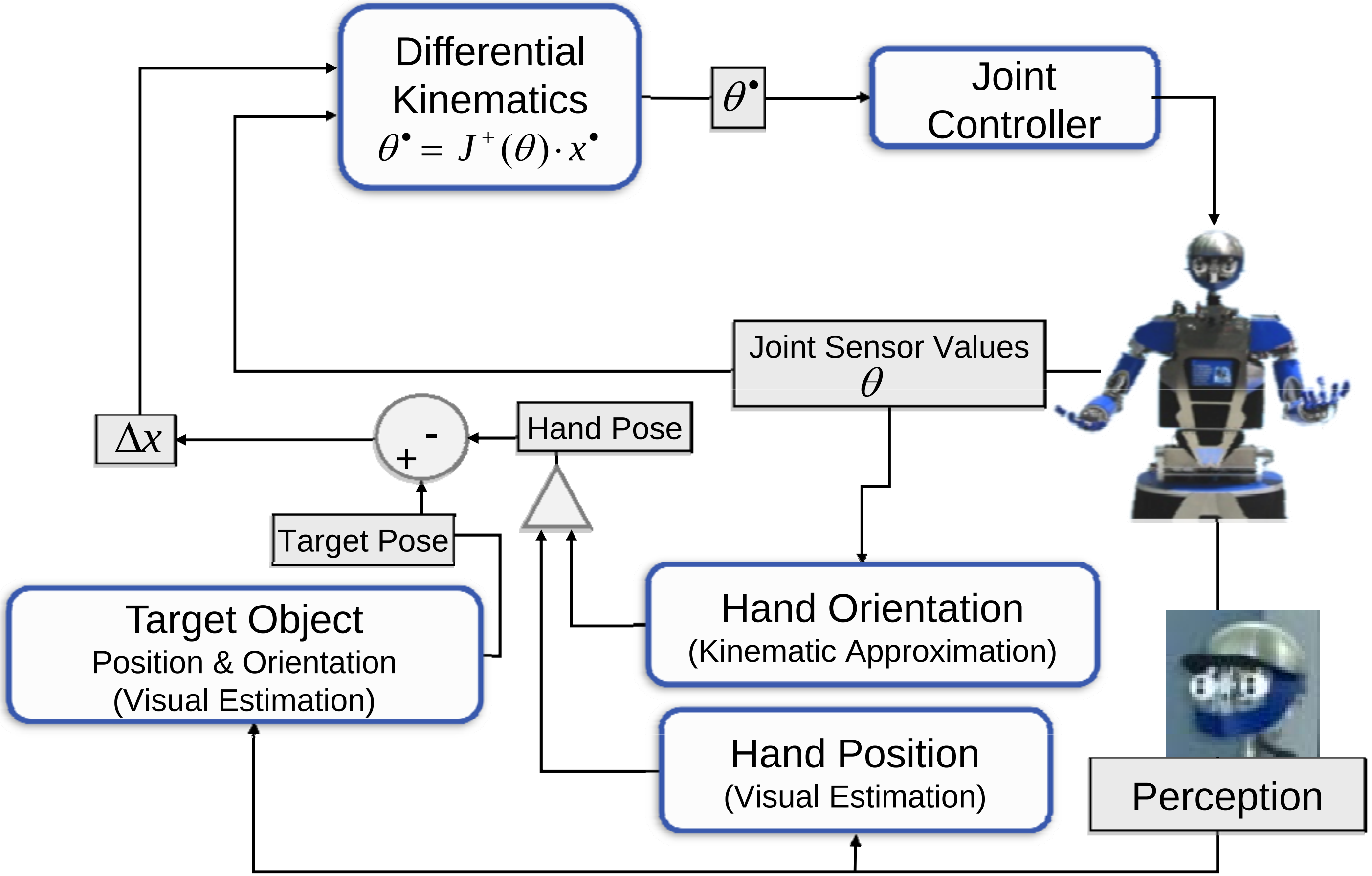
Grasping known objects: Offline grasp analysis

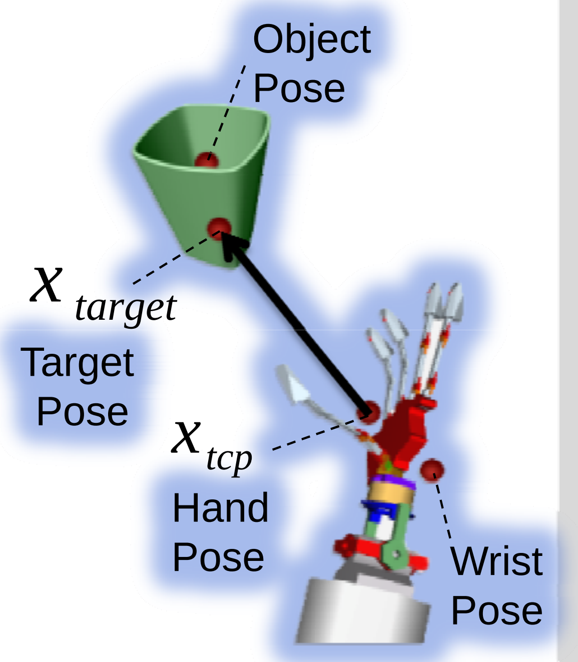
- Feasible grasps for every object are computed offline and stored together with the object models.
- A grasp is defined by:
  - Grasp type
  - Grasp starting point
  - Approaching direction
  - Hand orientation
- Simulation environment: GraspIt!



Hook *Cylindrical* *Spherical* *Pinch* *Tripod*

Position-based Visual Servoing





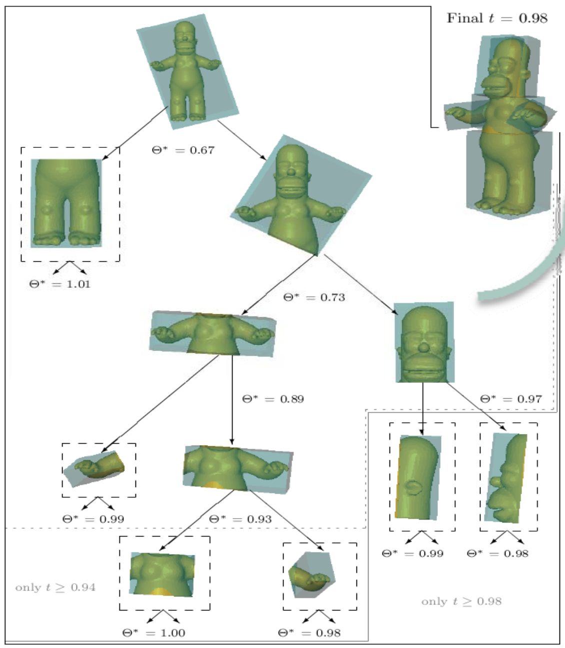
$$\theta^{\bullet} = J^{+}(\theta) \cdot x^{\bullet}$$

$$\partial^{t} = x^{t}_{vision} - x^{t}_{kinematic}$$

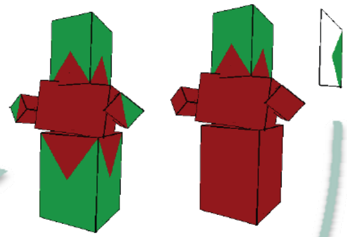
$$x^{t+1}_{tcp} = x^{t+1}_{kinematic} + \partial^{t}_{tcp}$$

Grasping Known Objects: A Box-Based Approach

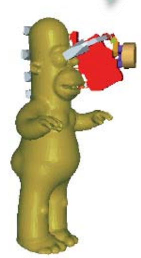
- Object shape approximation through Box Decomposition [Huebner et al. 2008]
- Grasp hypothesis generation on objects, evaluated in GraspIt! – part based grasping :



Shape Approximation



Heuristical Selection (Geometric Reduction & Grasp Quality Learning)



Final grasp for the task "show object"

Joint work with Danica Kragic and Kai Huebner, KTH, Sweden

Limitations and shortcuts

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

  - How to grasp unknown objects?
  - How to learn objects representations?
  - How to acquire multi-sensory representations of objects?

- 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?

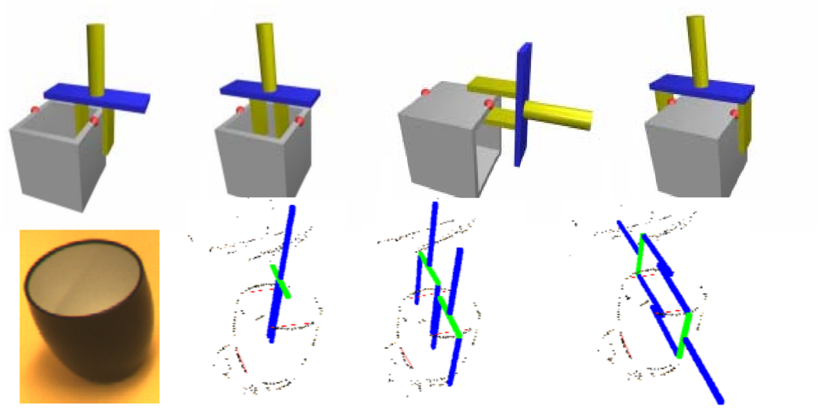
How to relax the limitations?

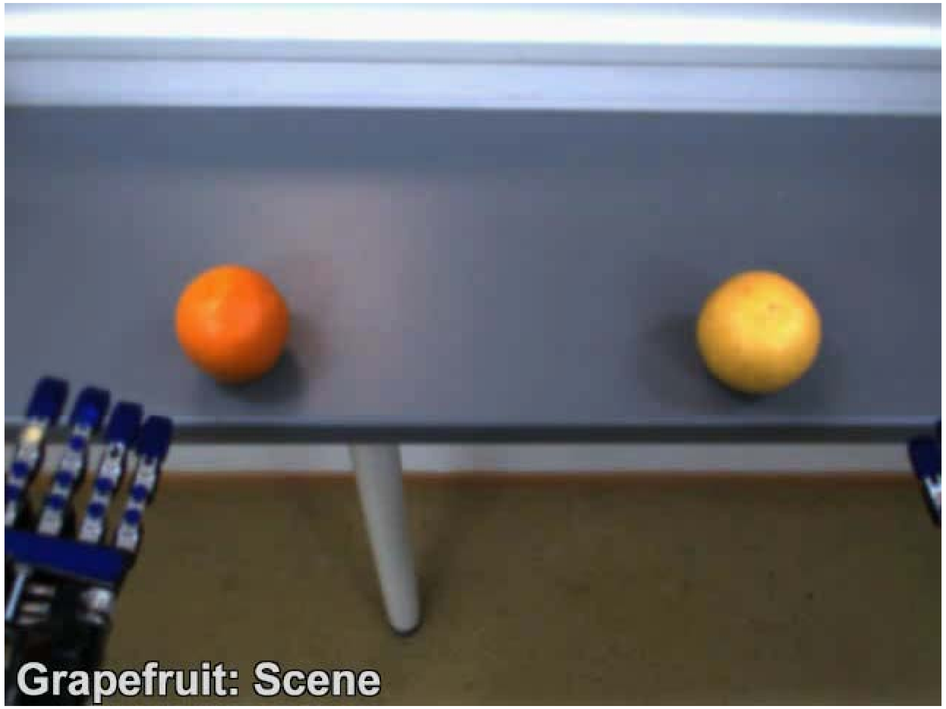
- Autonomous Exploration
  - Grasping unknown objects
  - Learning actions on objects (Pushing for grasping)
  - Visually-guided haptic exploration

- Coaching and Imitation
  - Learning from observation and goal-directed Imitation

Grasping unknown objects

- Co-planarity relation between visual entities define potential grasping affordances
- Surprising result: A success rate between 30-40% is already achievable by such a simple mechanisms.
  - One reason is high level mechansism for hypotheses rejections through motion planning
- There is an *autonomous success evaluation* based on force/haptic information
  - Collision, no success, unstable, stable

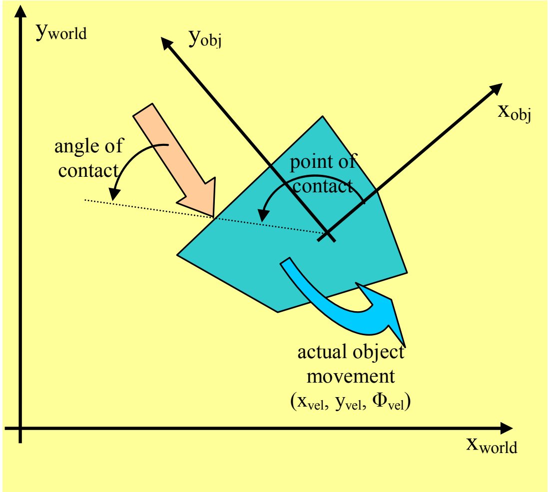


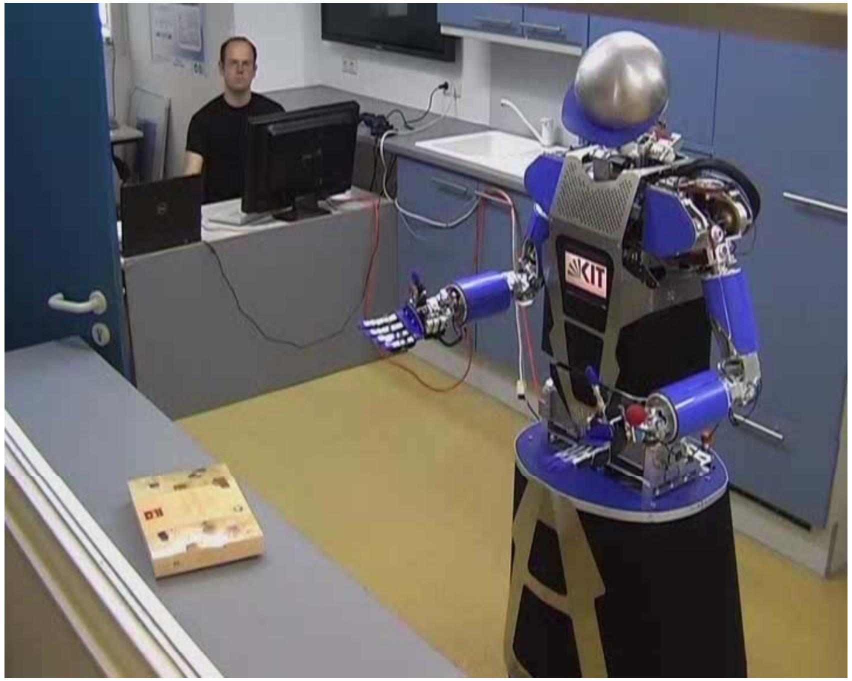


Joint work with Norbert Krüger and Mila Popovic, University of Southern Denmark

Learning of actions on objects

- Pushing for grasping
- Learning relationship between point and angle of push and the actual movement of an object
- Use the knowledge in order to find the appropriate point and angle of push in order to bring an object to a goal



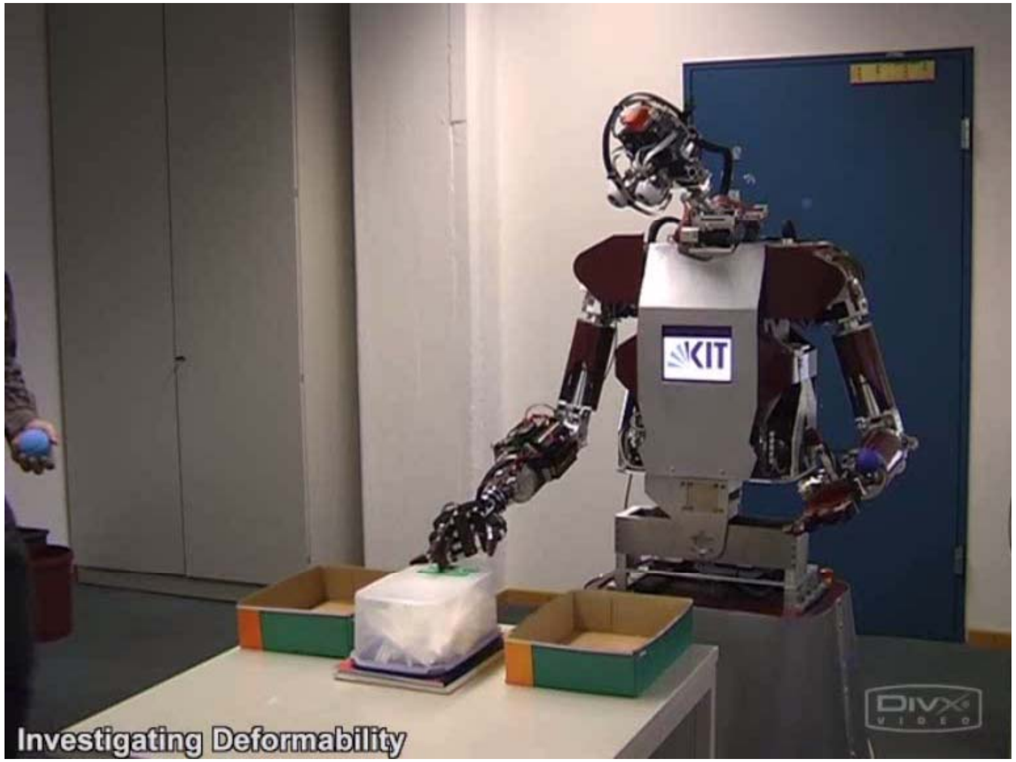


Joint work with Damir Omrcen and Ales Ude

Visually-guided haptic exploration

Fusion of tactile, proprioceptive and visual sensor data with a five-fingered hand

Verification of object size



Verification of object deformability

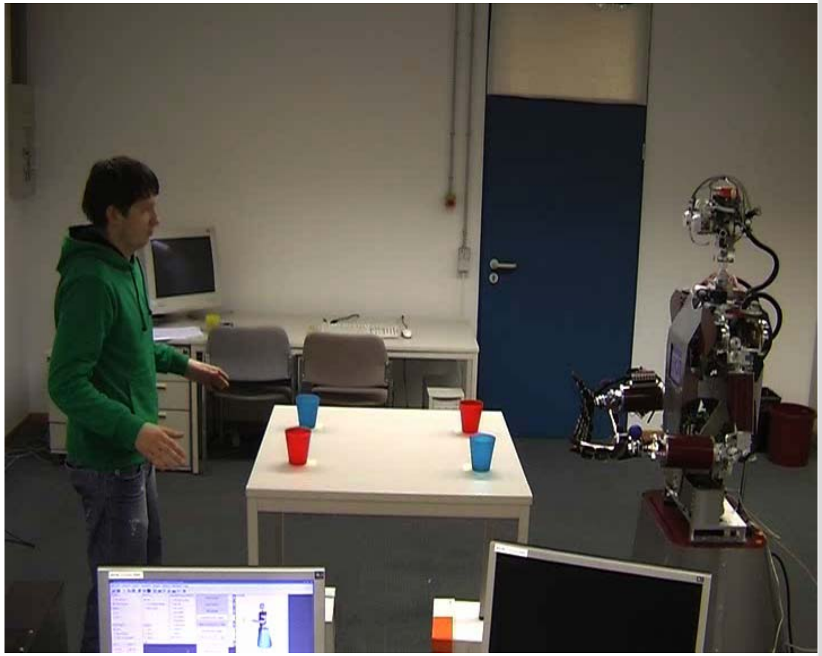
How to relax the limitations?

- Autonomous Exploration
  - Grasping unknown objects
  - Learning actions on objects (Pushing for grasping)
  - Visually-guided haptic exploration

- Coaching and Imitation
  - Learning from observation and goal-directed Imitation

Learning from Observation

- Library of motor primitives
  - Markerless human motion tracking
  - Object tracking
- Master Motor Map (MMM) as an interface for the transfer of motor knowledge between different embodiments
- Action representation
  - Dynamic movement primitives for generating discrete and periodic 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

Joint work with Andrej Gams and Ales Ude

- Human grasp recognition, mapping and reproduction

Joint work with Danica Kragic

Connecting High-level Task Planning to Execution

PKS planner (STRIPS-like planner)

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

Conclusions

Current work on humanoid grasping and manipulation in human-centered environments

- Grasping known objects
  - Appearance-based approaches
  - Shape-based approaches (box decomposition)
- Grasping unknown objects
  - 3D visual primitives (CoVis) → Grasp reflex
- Pushing for Grasping
  - pushing movements encoded by a neural network learned by exploration
- Learning from human observation
  - Action representations, which support adaption to novel situations

Thanks

- Humanoids Group @ KIT
  - Rüdiger Dillmann
  - Tamim Asfour
  - Steven Wieland
  - Alex Bierbaum
  - Paul Holz
  - Martin Do
  - David Gonzalez
  - Kai Welke
  - Niko Vahrenkamp
  - Stefan Ulbrich
  - Pedram Azad
  - Kristian Regenstein
  - Markus Przybylski
  - Stefan Gärtner

- Collaborators
  - Members of the SFB 588
  - Ales Ude (JSI, Slovenia), James Kuffner (CMU,USA), Danic Kragic (KTH, Sweden), Norbert Krüger (Denmark), Mark Steedman (Edinburgh, UK), Stefan Schaal and Peter Pastor (USC, USA), Justus Piater (Liege, Belgium)

Thank you ...

... for your attention.

- This work has been conducted within:
  - the German Humanoid Research project SFB588 funded by the German Research Foundation (DFG) www.sfb588.uni-karlsruhe.de
  - the EU Cognitive Systems projects funded by the European Commission
    - PACO-PLUS www.paco-plus.org
    - GRASP www.grasp-project.eu