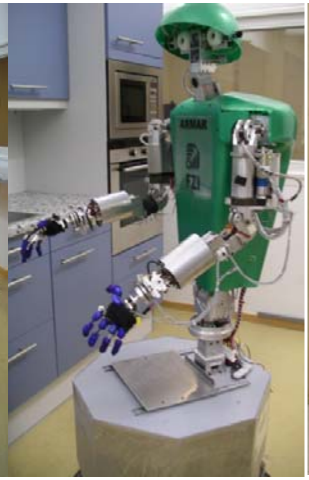
ShanghAI Lectures 2010
16. December 2010

Robots think with their hands

Tamim Asfour
Karlsruhe Institute of Technology, Institute for Anthropomatics
Humanoids and Intelligence Systems Lab

INSTITUTE FOR ANTHROPOMATICS, DEPARTMENT OF INFORMATICS






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

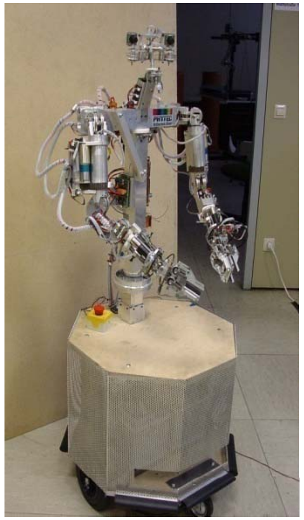
Building Humanoids = Building Human-Centered Technologies

- Assistants/companions for people in different ages, situations, activities and environments in order to improve the quality of life
- Key technologies for future robotic systems
- Experimental platforms to study theories from other disciplines

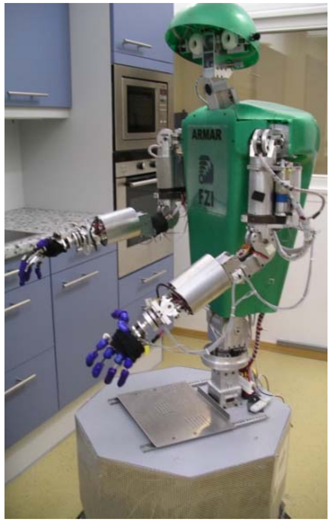
Ultimate goals

- 24/7 integrated complete humanoid robot systems able to act and interact in human-centered environments and to perform a variety of tasks
- Robots with rich sensorimotor capabilities as an indispensible requirement to implement cognitive capabilities in technical systems
- Reproducible complete humanoid systems in terms of mechanical design, mechantronics, hardware and software architecture
- Autonomous humanoid systems

Humanoid Robots @ KIT

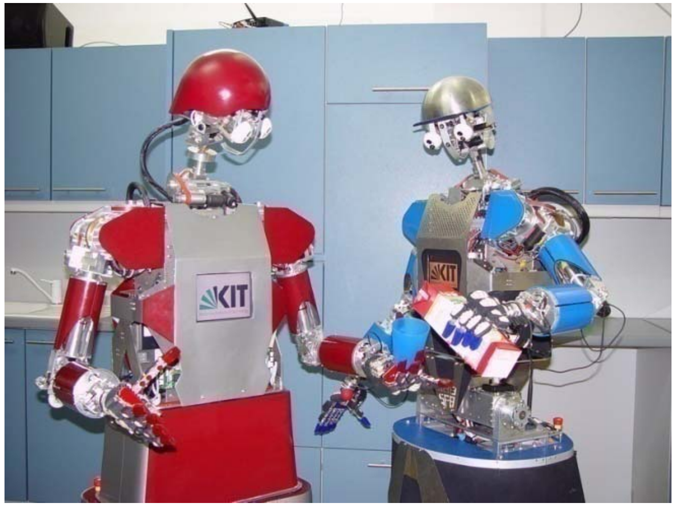


ARMAR, 2000



ARMAR-II, 2002

ARMAR-IIIa, 2006

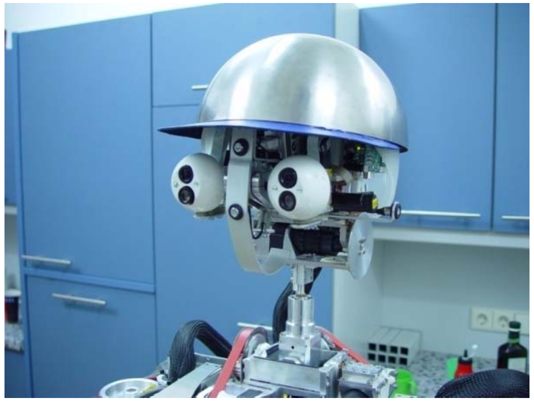


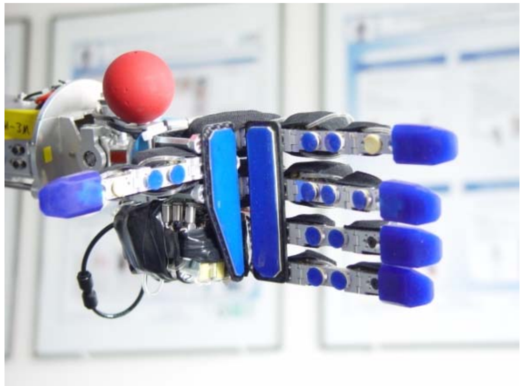
ARMAR-IIIb, 2008

- Collaborative Research Center 588: Humanoid Robots - Learning and Cooperating Multimodal Robots (SFB 588)
  - Funded by the German Research Foundation (DFG: Deutsche Forschungsgemeinschaft)
  - 2001 – 2012
  - http://www.sfb588.uni-karlsruhe.de/

ARMAR-IIIa and ARMAR-IIIb

- 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



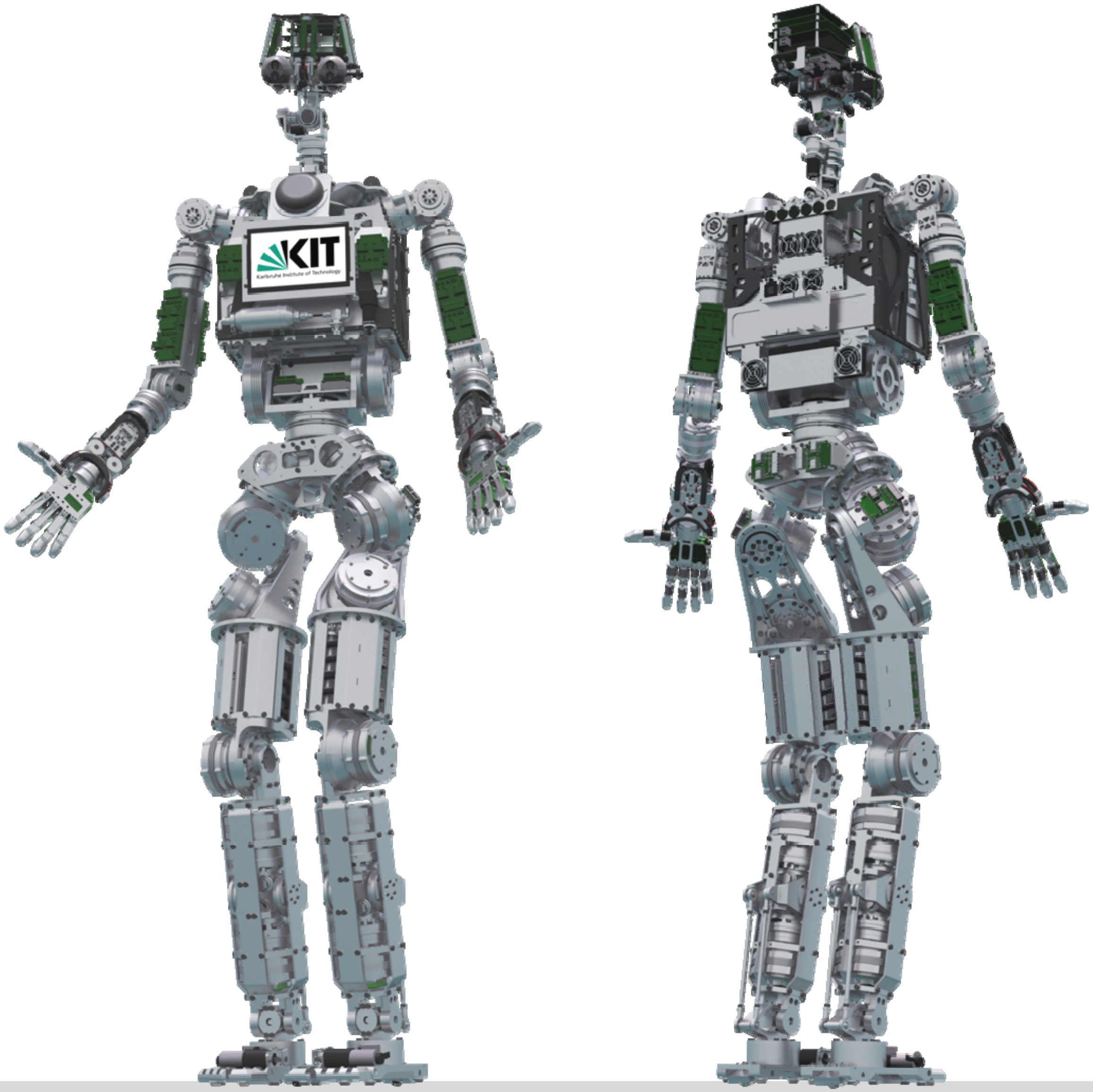


Fully integrated autonomous humanoid system

Table 1
Specification of ARMAR-III

Weight		135 kg (incl. 60 kg battery)	
Height		175 cm	
Speed		1 m/s	
DOF	Eyes	3	Common tilt and independent pan
	Neck	4	Lower Pitch, Roll, Yaw, upper Pitch
	Arms	2×7	3 DOF in each shoulder, 2 DOF in each elbow, and 2 in each wrist
	Hands	2×8	Five-fingered hands with 2 DOF in each Thumb, 2 DOF in each Index and Middle, and 1 DOF in each Ring and Pinkie.
	Toros	3	Pitch, Roll, Yaw
	Platform	3	3 wheels arranged in angles of 120°
Actuator		DC motors + Harmonic Drives in the arms, neck, eyes, torso and platform. Fluidic actuators in the hand.	
Sensors	Eyes	2 Point Grey (www.ptgrey.com) Dragonfly cameras in each eye, six microphones and a 6D inertial sensor (http://www.xsens.com).	
	Arms	Motor encoders, axis sensors in each joint, torque sensors in the first five joints and 6D force–torque sensor (http://www.ati-ia.com) in the wrist.	
	Platform	Motor encoders and 3 Laser-range finders (http://www.hokuyo-aut.jp).	
Power supply		Switchable 24 V Battery and 220 V external power supply.	
Operating system		Linux with the Real-Time Application Interface RTAI/LXRT-Linux.	
Computers and communication		Industrial PCs and PC/104 systems connected via Gigabit Ethernet and 10 DSP/FPGA control units (UCoM) which communicate with the control PC via CAN bus.	
User interface		Graphical user interface (GUI) connected to the robot via wireless LAN and natural speech communication.	

ARMAR-IV



- 63 DOF
- 170 cm
- 70 kg

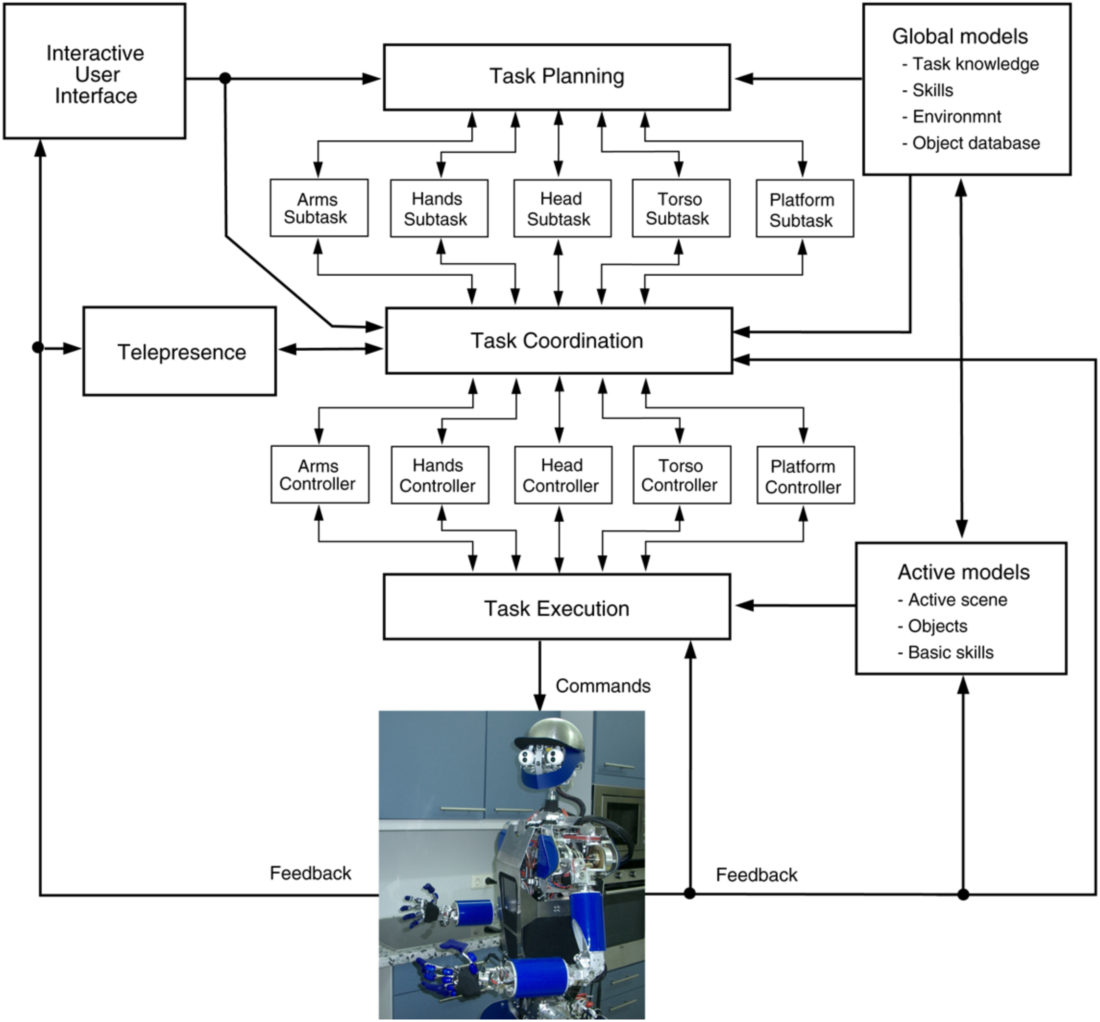
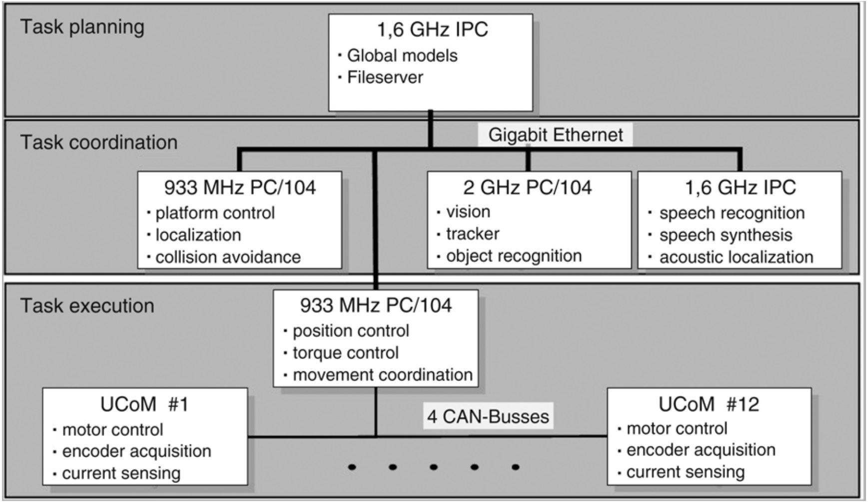


Fig. 2. Hierarchical control architecture for coordinated task execution in humanoid robots: planning, coordination and execution level.



Task planning
1,6 GHz IPC
· Global models
· Fileserver
Gigabit Ethernet
Task coordination
933 MHz PC/104
· platform control
· localization
· collision avoidance
2 GHz PC/104
· vision
· tracker
· object recognition
1,6 GHz IPC
· speech recognition
· speech synthesis
· acoustic localization
Task execution
933 MHz PC/104
· position control
· torque control
· movement coordination
UCoM #1
· motor control
· encoder acquisition
· current sensing
4 CAN-Busses
UCoM #12
· motor control
· encoder acquisition
· current sensing

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 Video

Interactive tasks in kitchen environment

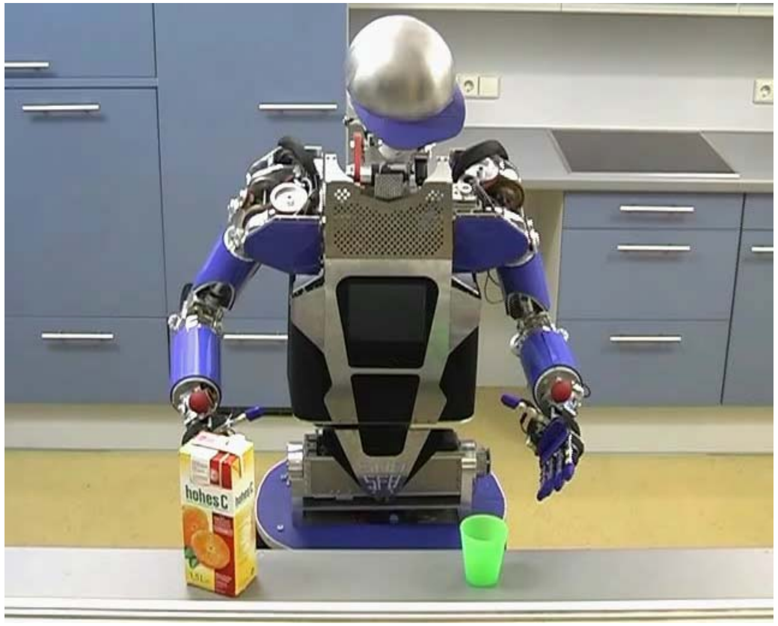
- Object recognition and localization
- Vision-based grasping
- Hybrid position/force control
- Vision-based self-localisation
- Combining force and vision for opening and closing door tasks
- Learning new objects, persons and words
- Collision-free navigation
- Audio-visual user tracking and localization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

Video

Bimanual grasping and manipulation

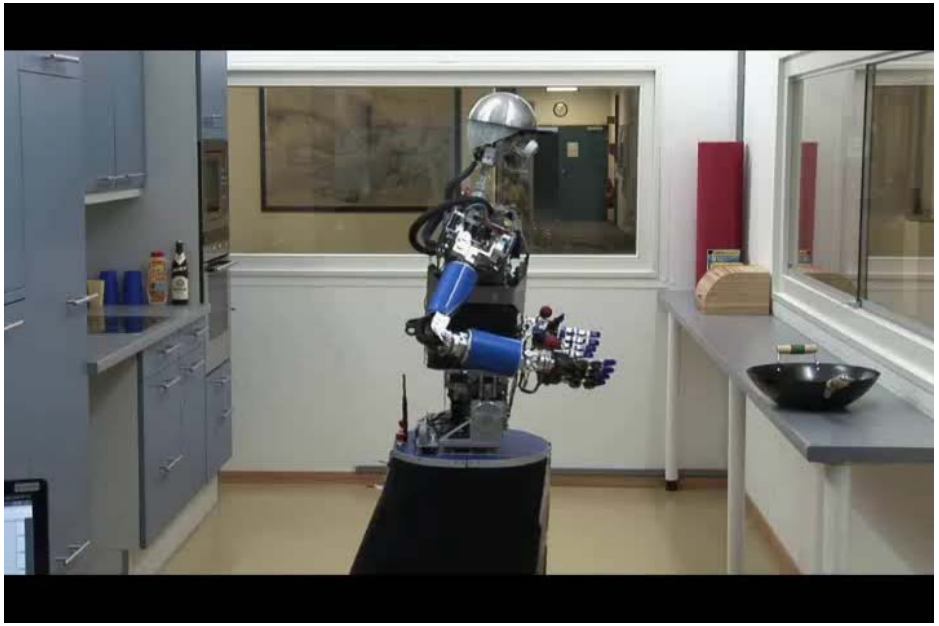
- Stereovision for object recognition and localization
- Visual Servoing for dual-hand grasping
- Zero-force control for teaching of grasp poses

Video



Loosely coupled dual-arm tasks

Video



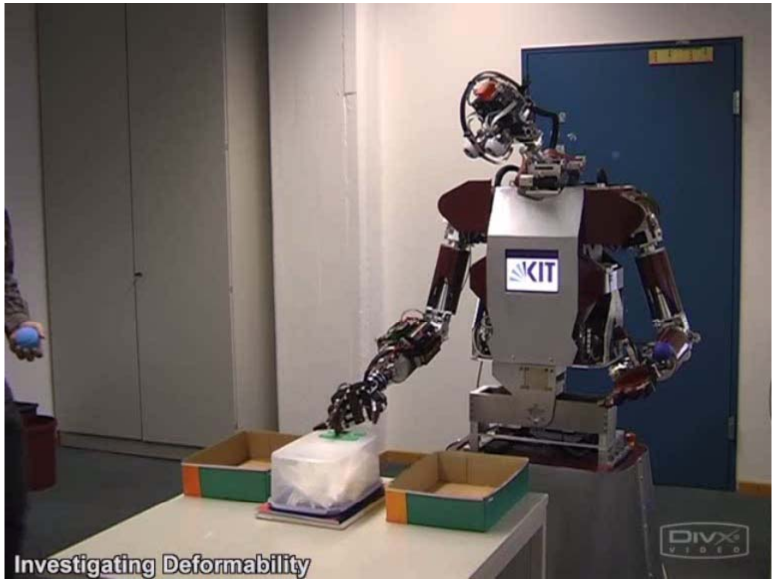
Tightly coupled dual-arm tasks

How to relax the limitations in our scenario?

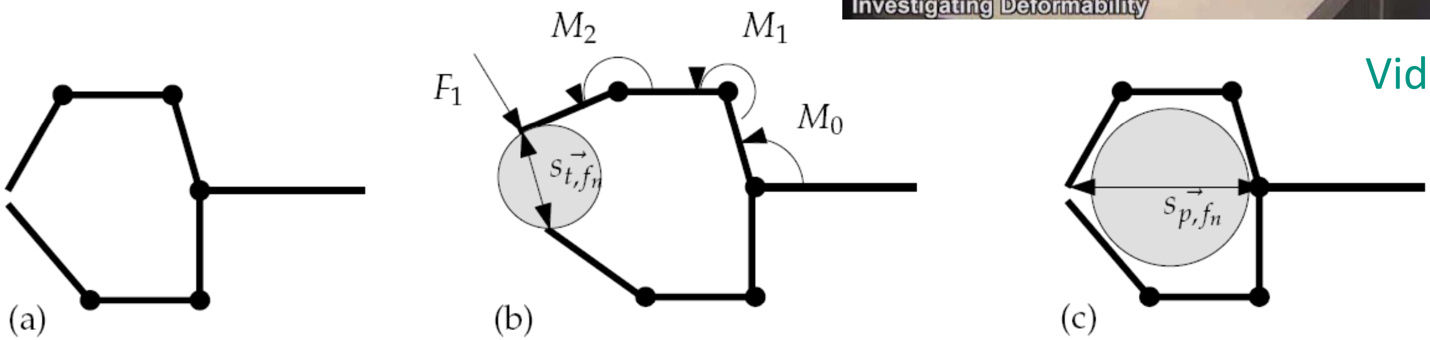
- Autonomous Exploration:
  - Visually-guided haptic exploration
  - Visual object exploration and search
  - Learning actions of objects
- Coaching and Imitation
  - Learning from Observation
  - Goal-directed Imitation

Hand: available skills

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



Video



Failed grasp "no object"

Precision grasps: Distance between fingertips

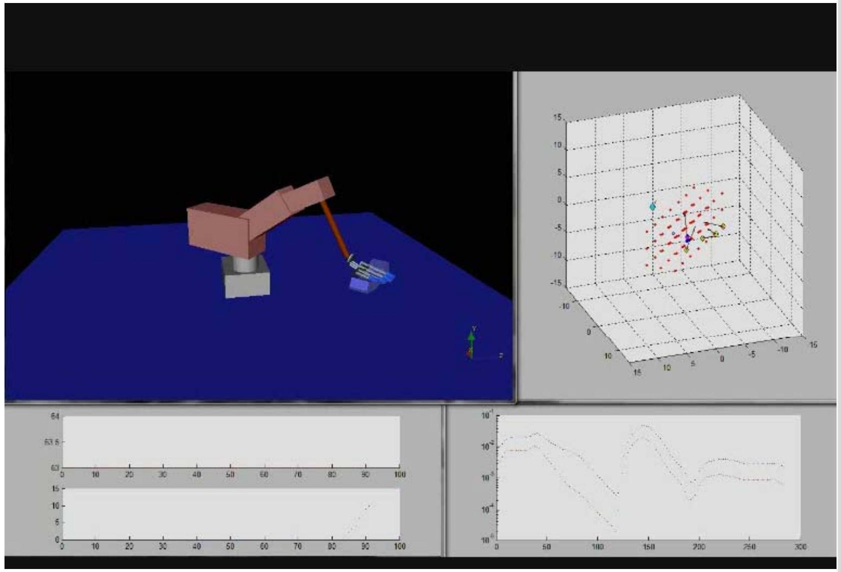
Power grasps: Distance between fingertips and palm

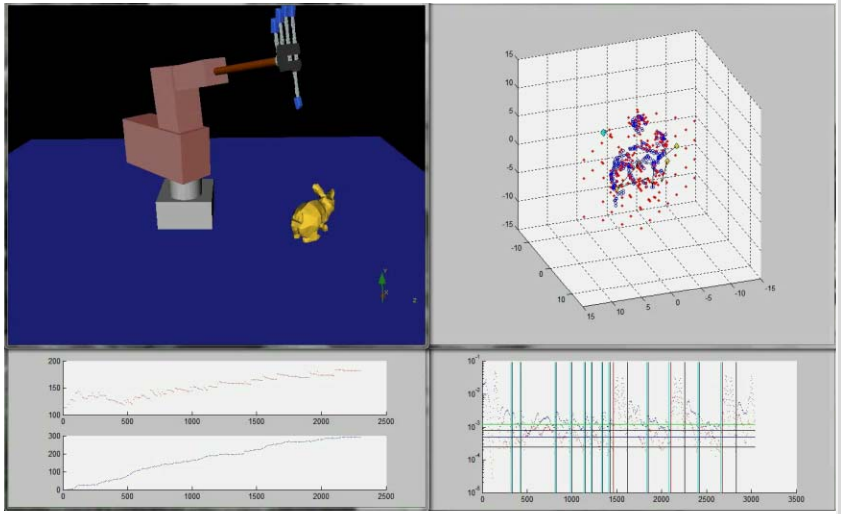
Tactile Object Exploration

Videos

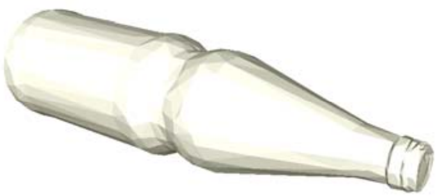
- Potential field approach to guide the robot hand along the object surface
- Oriented 3D point cloud from contact data
- Extract faces from 3D point cloud in a geometric feature filter pipeline
  - Parallelism
  - Minimum face size
  - Face distance
  - Mutual visibility

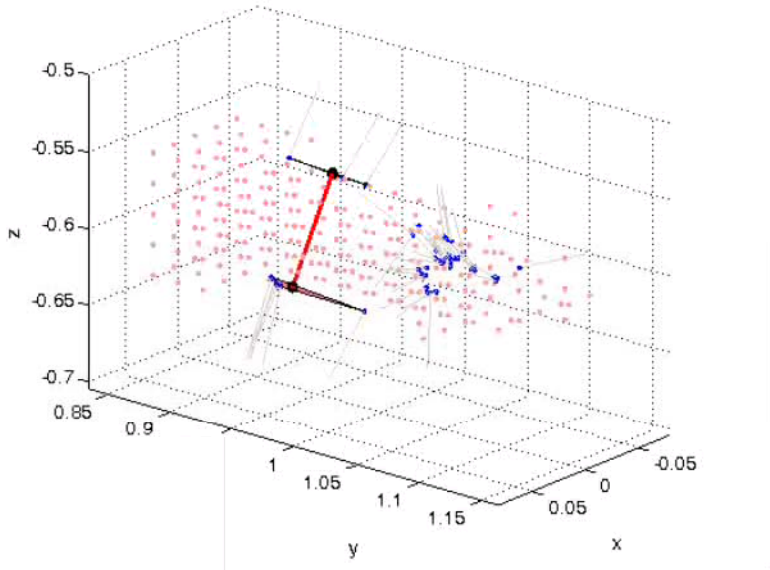
→ Association between objects and actions (grasps) → Symbolic grasps (grasp affordances)

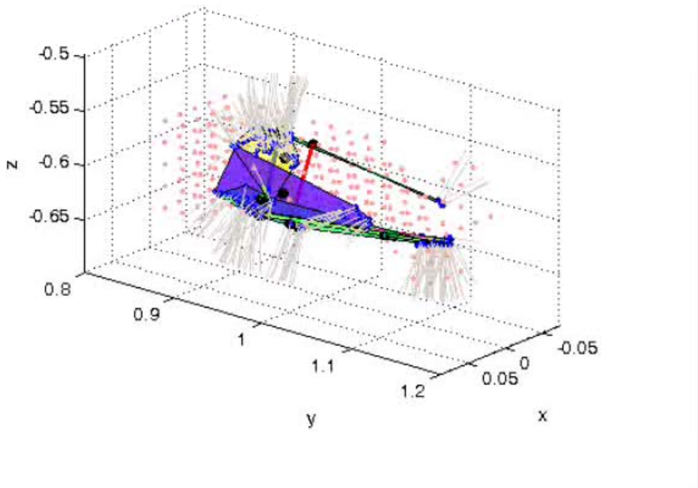




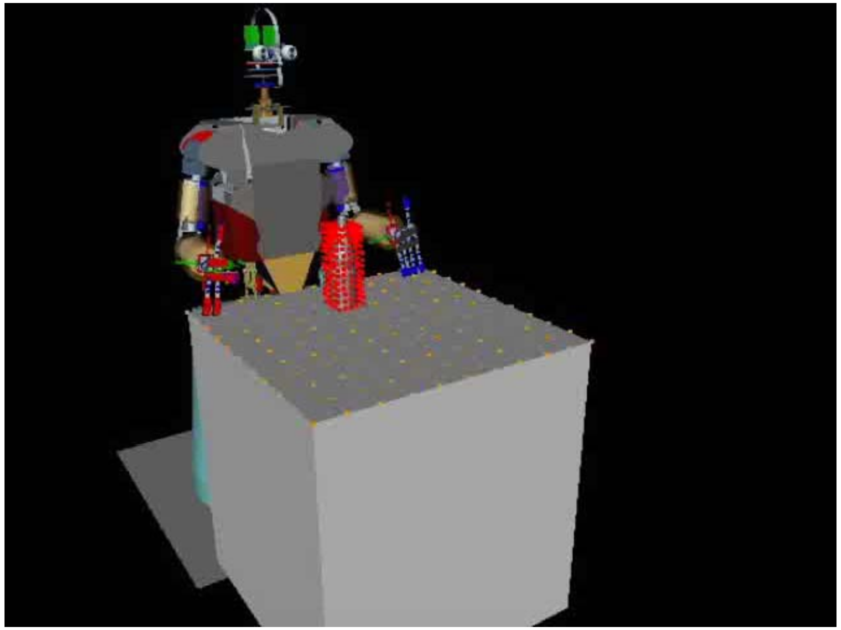
Examples: Bottle

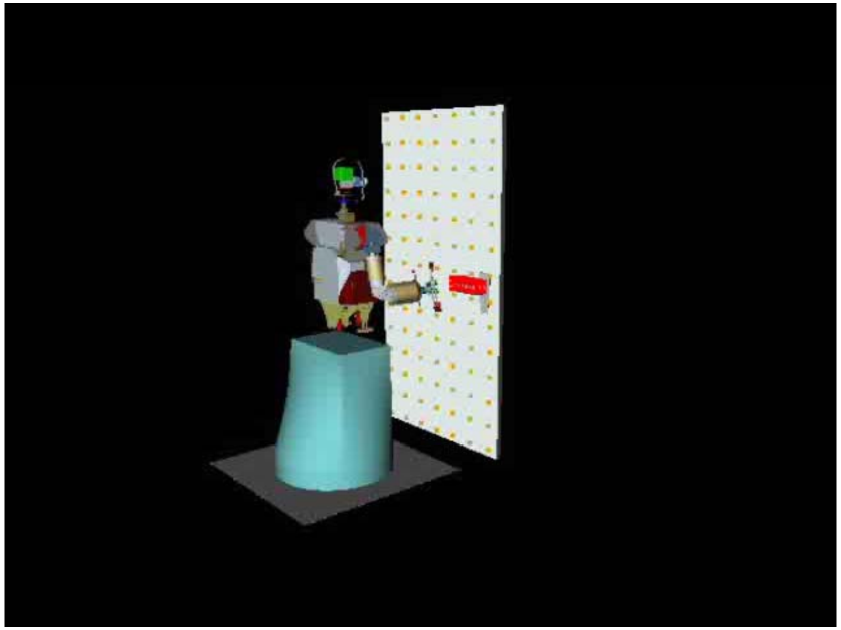






Visually guided haptic exploration on ARMAR





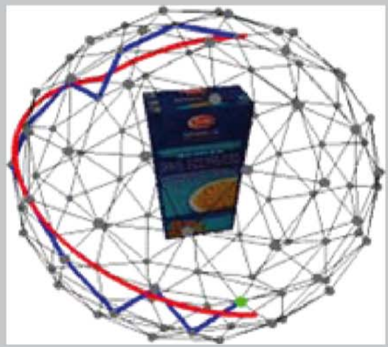
Videos

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

Active Visual Object Exploration and Search

Exploration

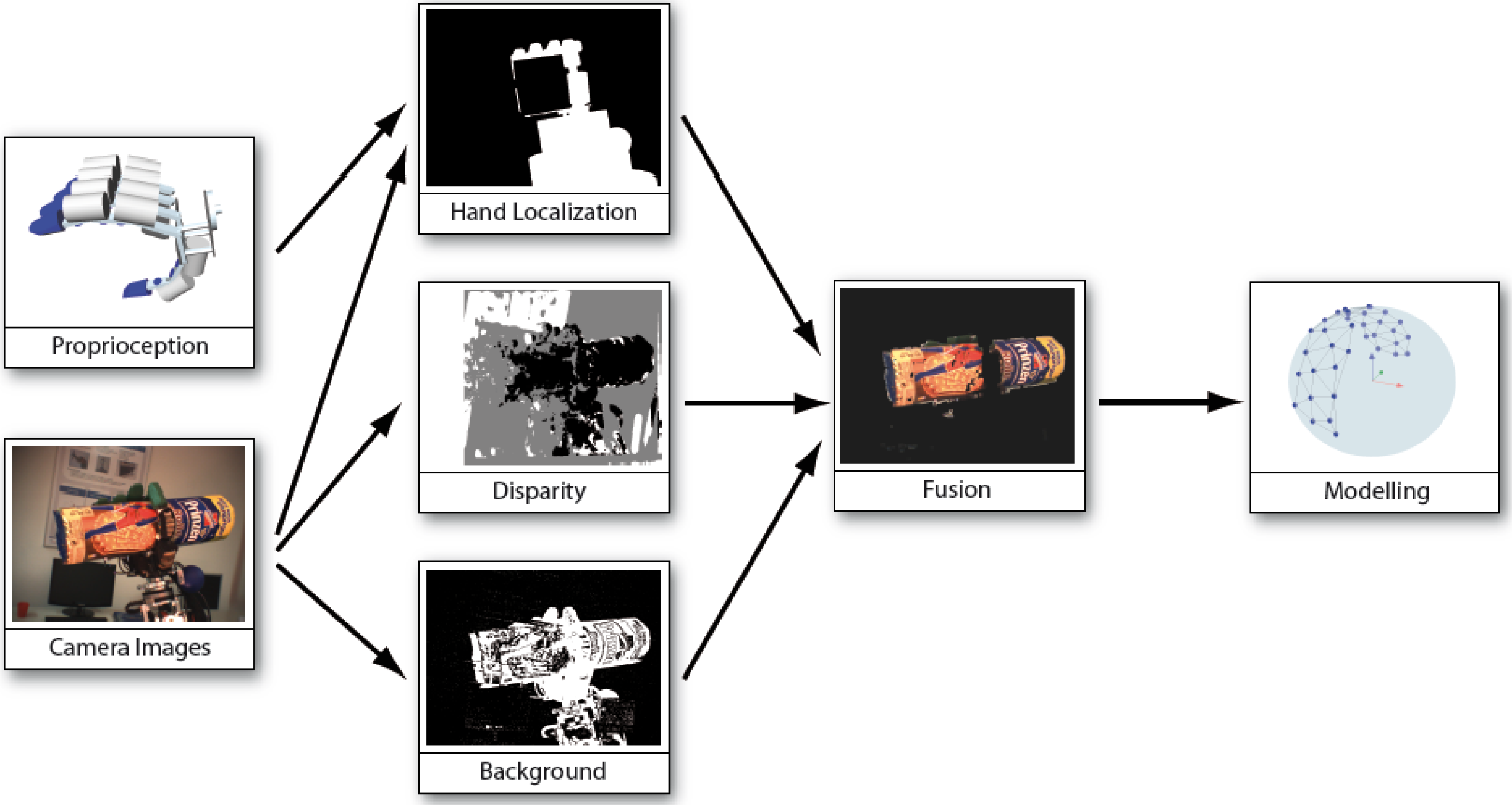
Representation



Visual Search

- Generation of visual representations through exploration
- Application of generated representations in recognition tasks.

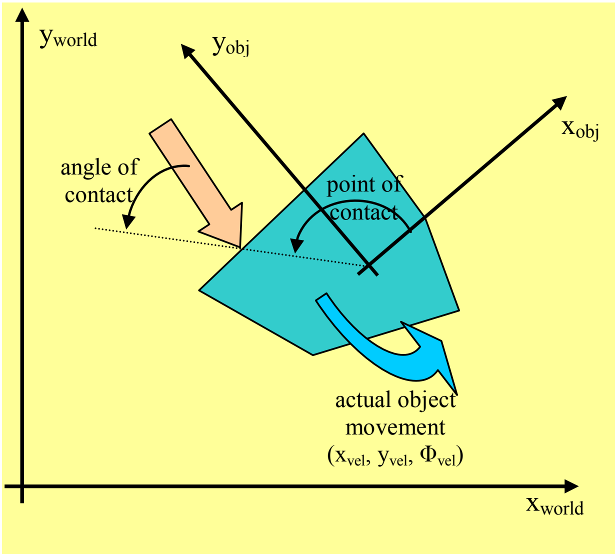
Generation object views through manipualtion

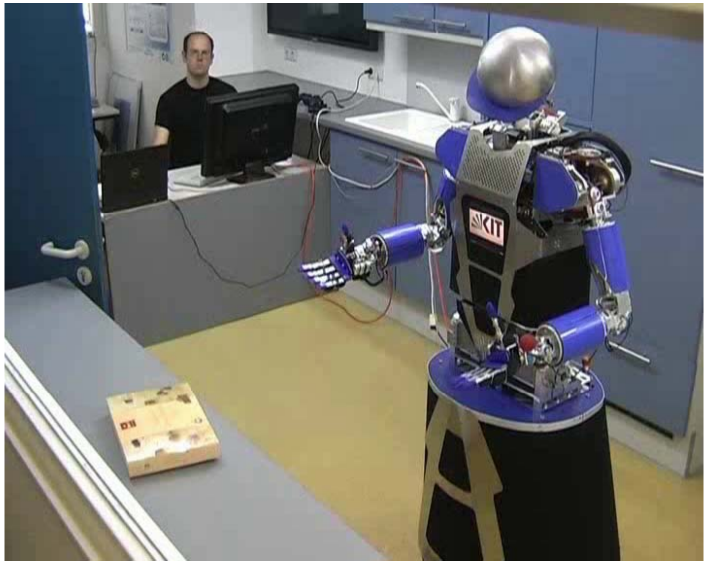


Video

Learning by Autonomous Exploration: Pushing

- Learning of actions on objects (Pushing)
- 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





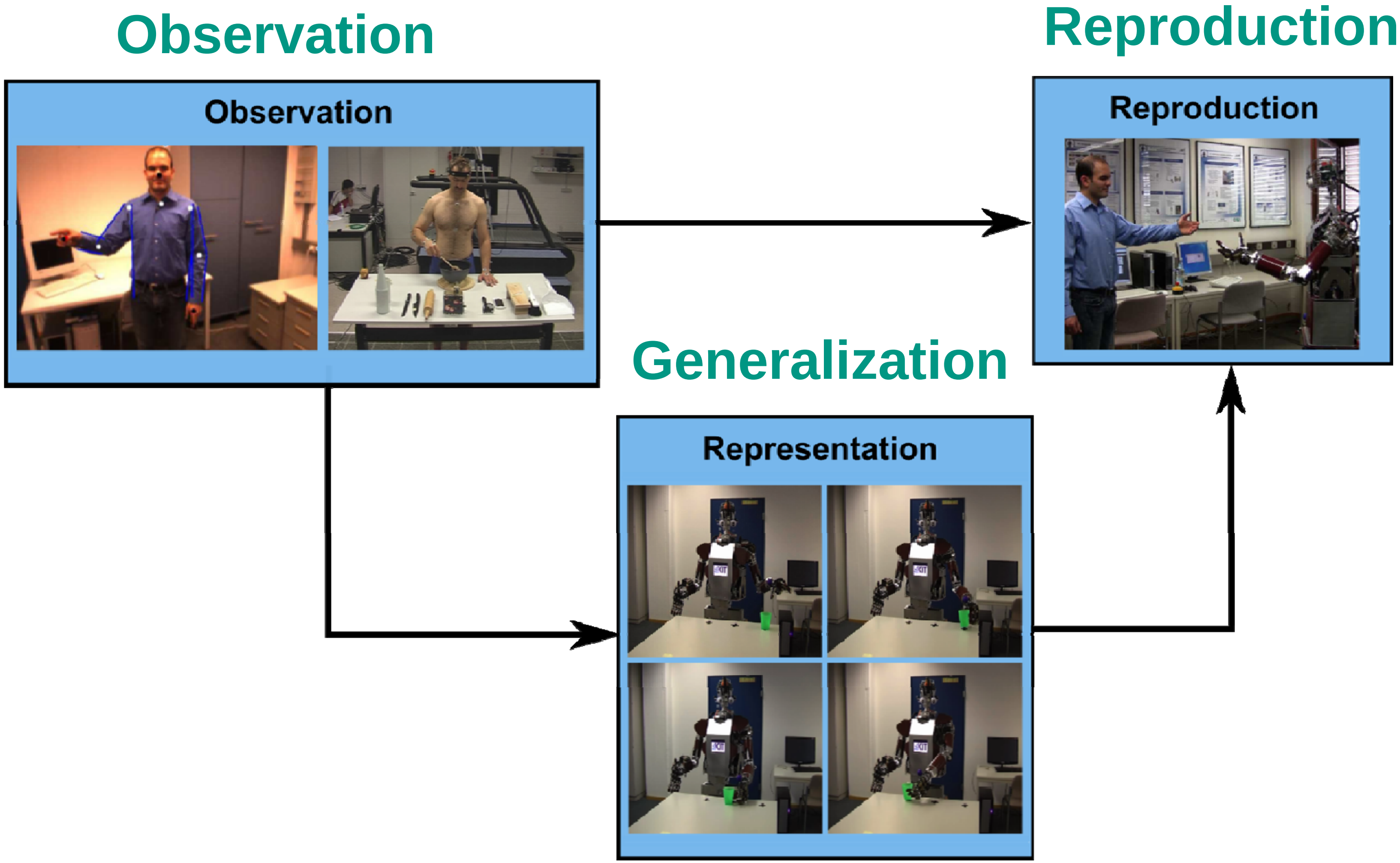
Work with Damir Omrcen and Ales Ude

Video

How to relax the limitations in our scenario?

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

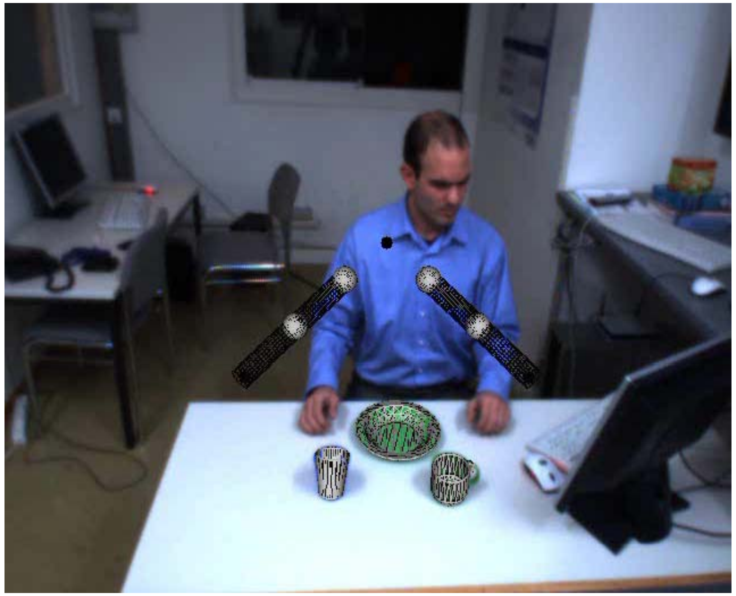
Learning from human observation



Stereo-based 3D Human Motion Capture

- 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

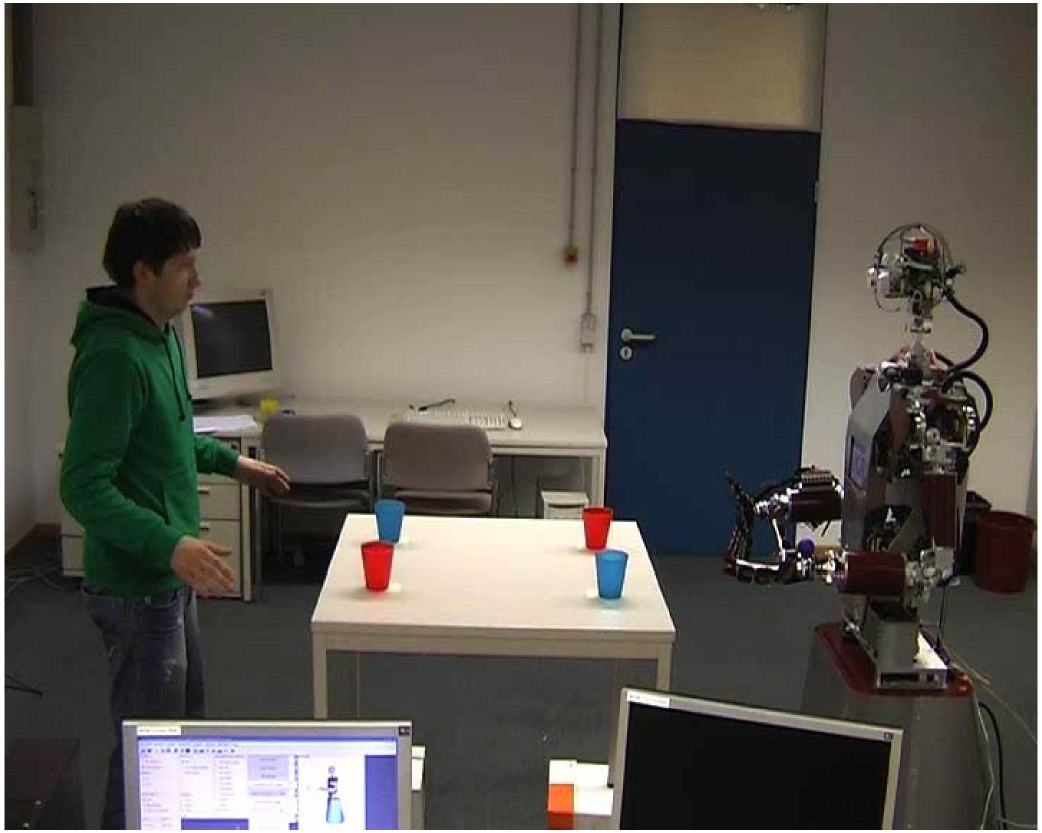
video



Integrating Vision Toolkit: http://ivt.sourceforge.net/

Reproduction

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



Generalisation?

Video

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
  - 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.

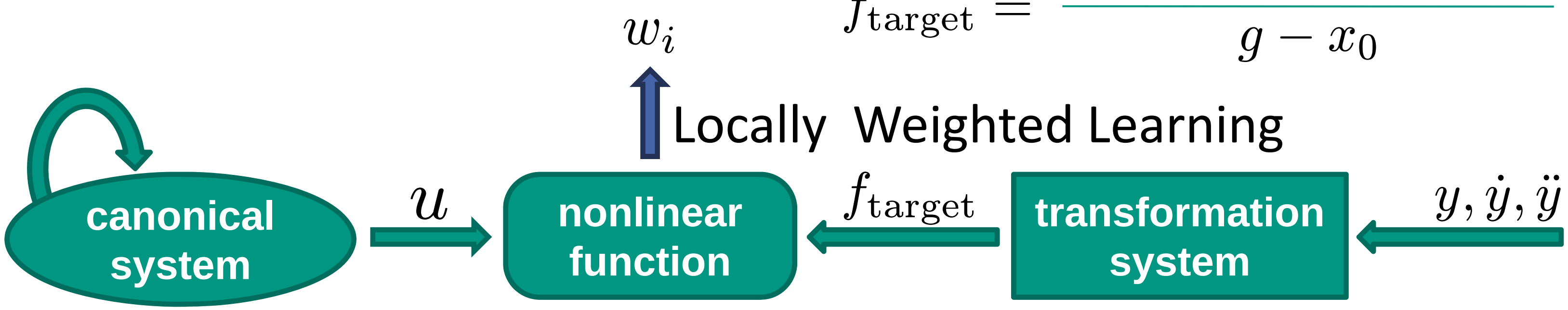
Action representation using DMPs

canonical system: $\tau\,\dot{u} = -\alpha\,u$

nonlinear function: $f(u) = \dfrac{\sum_i \psi_i(u)\,w_i\,u}{\sum_i \psi_i(u)}$ $\qquad \psi_i(u) = e^{-h_i(u-c_i)^2}$

transformation system: $$\tau\dot{v} = K(g-x) - Dv + (g-x_0)f$$ $$\tau\dot{x} = v$$

$$f_{\text{target}} = \frac{-K(g-y) + D\dot{y} + \tau\ddot{y}}{g - x_0}$$



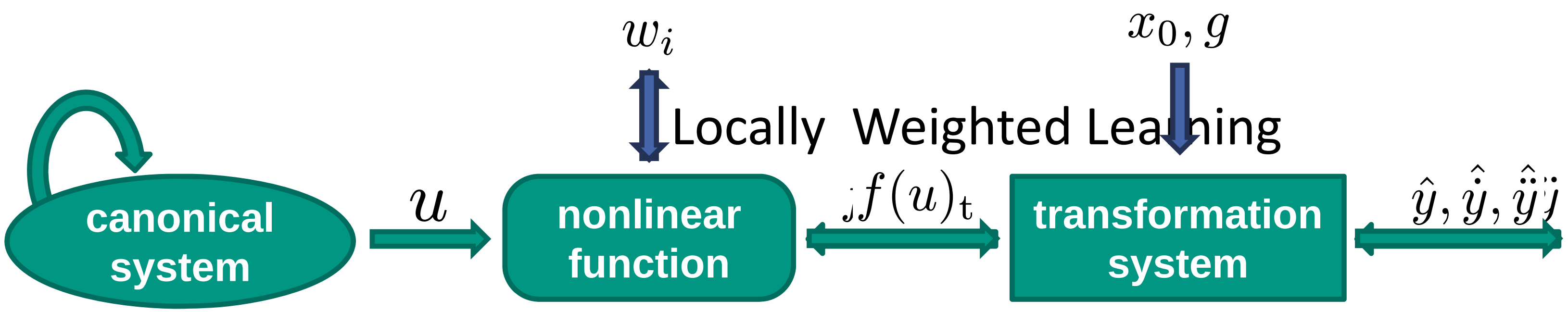
Action representation using DMPs

canonical system: $\tau \dot{u} = -\alpha u$

nonlinear function: $f(u) = \frac{\sum_i \psi_i(u) w_i u}{\sum_i \psi_i(u)}$ $\psi_i(u) = e^{-h_i(u - c_i)^2}$

transformation system: $\tau \dot{v} = K(g - x) - Dv + (g - x_0) f$

$\tau \dot{x} = v$



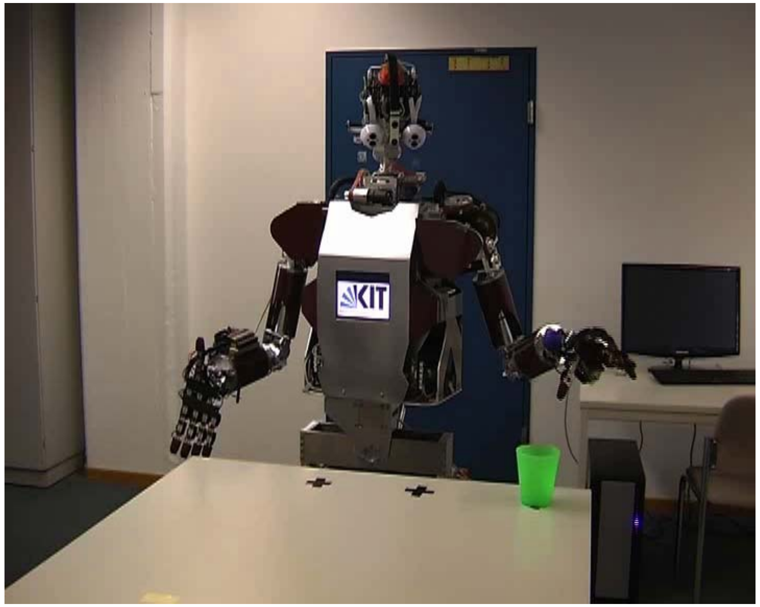
Learning from Observation

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

- 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



Video

Learning from Observation

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

Video

Joint work with Andrej Gams and Ales Ude

A. Ude, A. Gams, T. Asfour, J. Morimoto. Task-Specific Generalization of Discrete and Periodic Dynamic Movement Primitives. IEEE Trans. on Robotics, vol. 26, no. 5, pp. 800 – 815, October, 2010

Thanks to the PACO-PLUS Consortium

Karlsruhe Institute of Technology

University of Karlsruhe, Germany
R. Dillmann, T. Asfour

ATR, Kyoto, Japan
M. Kawato & G. Cheng

Kungliga Tekniska
Högskolan, Sweden
J. O. Eklundh, D. Kragic

University of Liege
Belgium
J. Piater

PACO-PLUS

Perception, Action and Cognition through Learning of Object-Action Complexes

www.paco-plus.org

University of
Göttingen
Germany
F. Wörgötter

University of
Southern Denmark
N. Krüger

Aalborg University
Denmark
V. Krüger

University of Edinburgh
United Kingdom
M. Steedman

Jozef Stefan Institute
Slovenia
A. Ude

Leiden University
Netherlands
B. Hommel

Consejo Superior de Investigaciones
Científicas, Spain
C. Torras & J. Andrade-Cetto

Thanks

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

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