From Sensorimotor Primitives Learned from Humans to Imitation and Manipulation Strategies in Humanoid Robots

Rüdiger Dillmann

Karlsruhe Institute of Technology (KIT)

Institute for Anthropomatics

wwwiaim.ira.uka.de

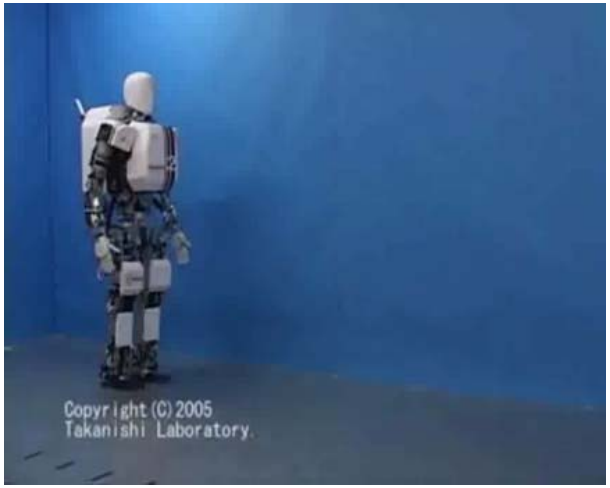
Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Fakultät für Informatik

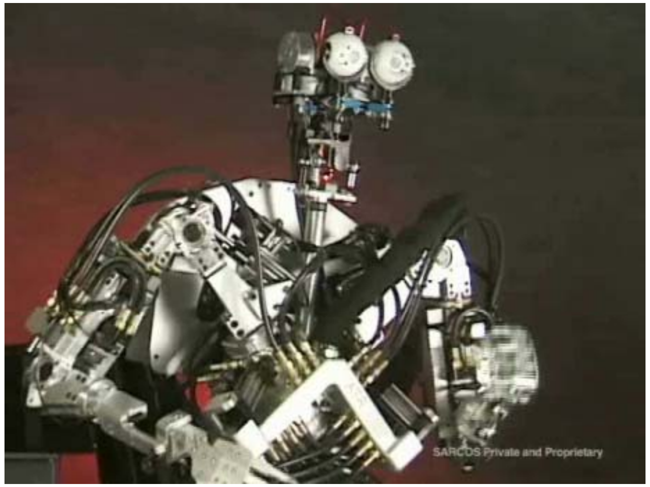
Advances in Humanoids Research

ASIMO, Honda, Japan

HRP-2 at Ikeuchi Lab, Kawada Industries, Japan



Wabian-2, Waseda University, Japan



CB, ATR, Kyoto, Japan

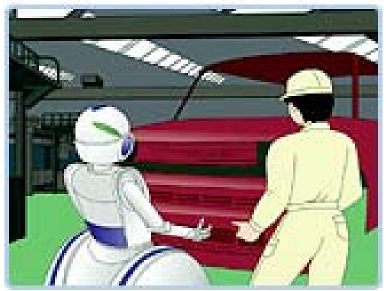
Potential application areas of humanoids

Close interactions with humans supported by natural multimodal communication channels such as gesture, vision, body-language, haptics, speech and natural language

Assistance

Elderly Care

Manufacturing



Mobility

Three key questions

- Grasping and manipulation in open-ended environments that
  - are populated by humans
  - are made for humans
  - should not be modified in any way for the benefit of the robot

© SFB 588, Karlsruhe

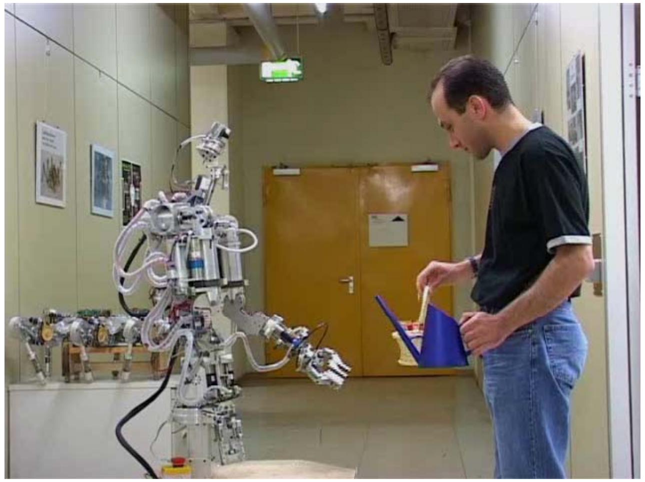
- Learning through
  - Observation of humans
  - Imitation of human actions

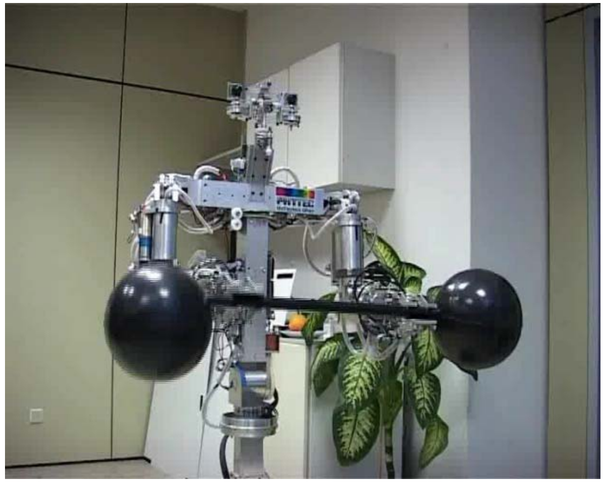
- Natural communication and interaction with humans
  - Multimodal interfaces
  - Natural speech

Humanoid Robots in Karlsruhe

- Collaborative Research Center on Learning and Cooperating Multimodal Humanoid Robots (SFB 588)
  - Humanoid robots for human-centered environments
  - Funded by the German Research Foundation (DFG: Deutsche Forschungsgemeinschaft)
  - 2001 - 2012
  - Interdisciplinary research project
    - University of Karlsruhe
    - Research Center Computer Science (FZI)
    - Research Center Karlsruhe
    - Fraunhofer Institute Karlsruhe (IITB)
  - http://www.sfb588.uni-karlsruhe.de/

ARMAR-I and ARMAR-II





First demonstrator of the SFB 588

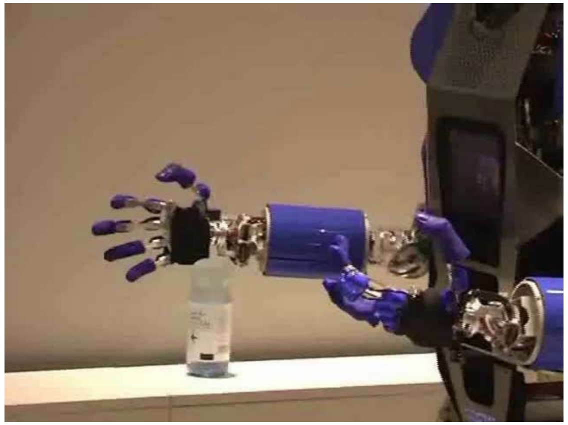
Demo at CEBIT 2006

HURO 1999, ICAR199
Humanoids 2000 , 2001
IROS 2003, 2004
ISER 2004
STAR2005

ARMAR-IIIa

- 7 DOF head with foveated vision
  - 2 cameras in each eye
  - 6 microphones
- 7-DOF arms
  - Position, velocity and torque sensors
  - Force-Torque 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

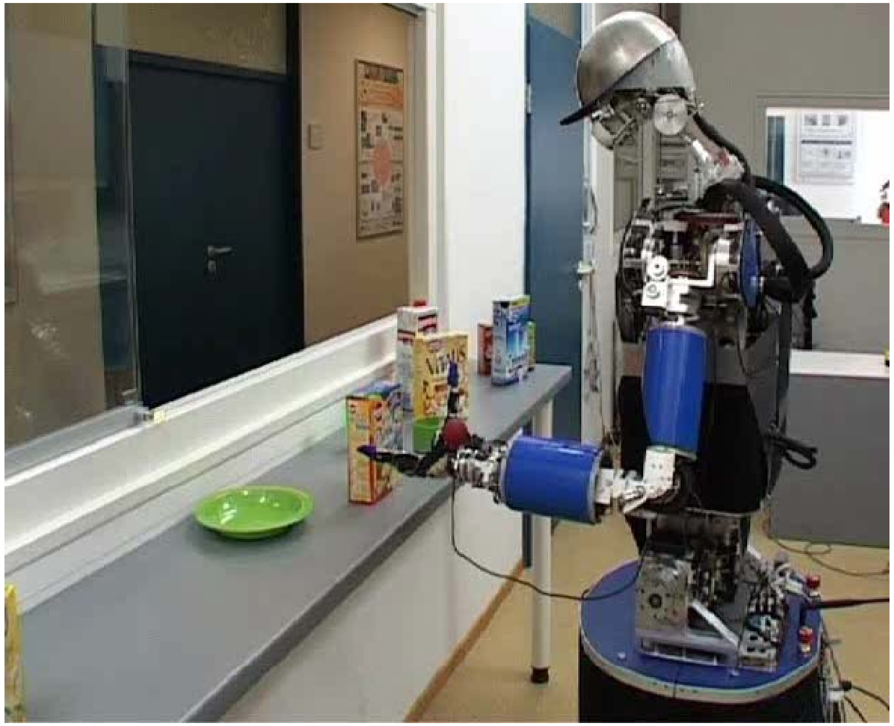


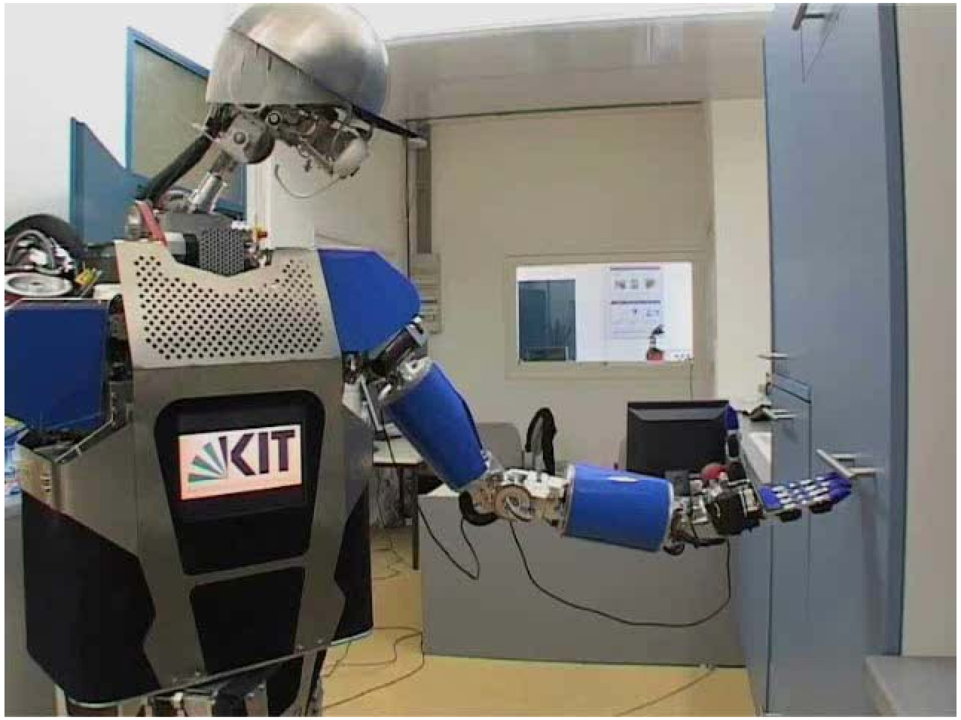



Fully integrated autonomous humanoid system!

Demonstration Scenario

Interactive service tasks in a kitchen environment

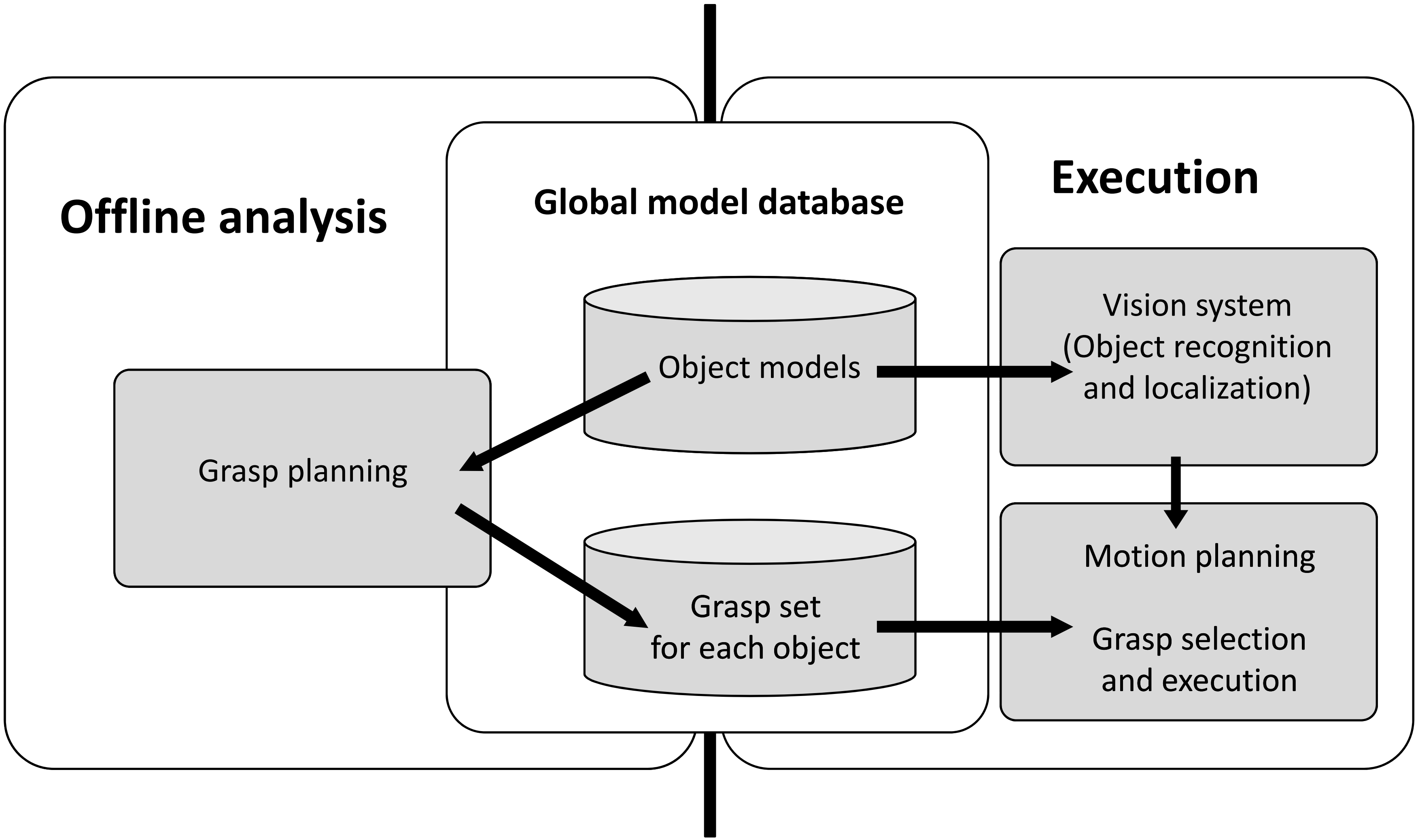




Outline of the talk

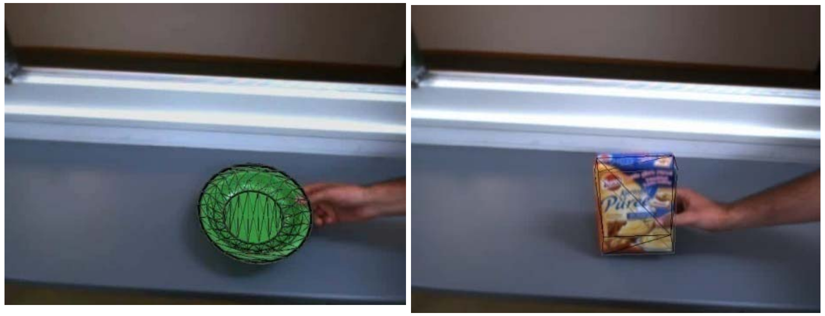
- Key techniques for the grasping and manipulation
  - Object recognition and localisation
  - Collision-free motion planning
  - Grasp planning
  - Kinematics control and execution
- Manipulation and imitation Strategies
  - Visually guided haptic exploration
  - Human and object motion capture
  - Learning from Observation
  - Imitation of human actions

Grasping and manipulation system



Object recognition and localization

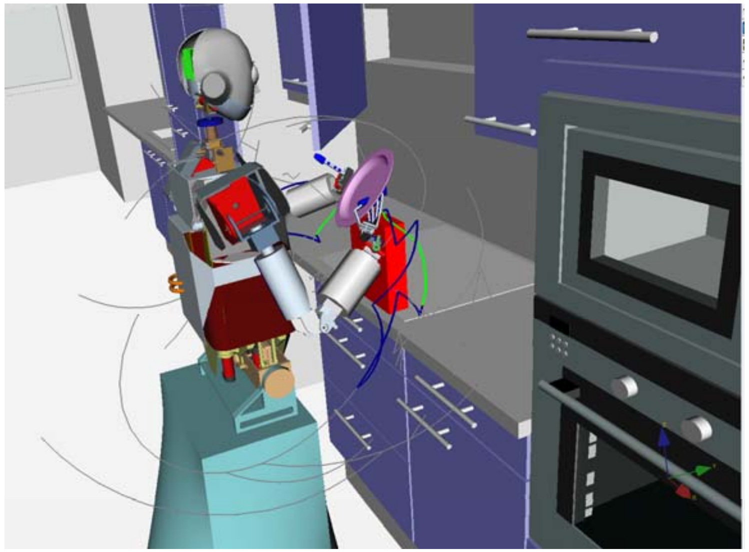
- Recognition of colored objects
  Segmentation by color (IROS 2006)
  - Segmentation by color
  - Appearance-based recognition using a global approach
  - Model-based generation of view sets
  - Combination of stereo vision and stored orientation information for 6D pose estimation
- Recognition of textured objects
  using local features (Humanoids 2006)
  - 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

Collision detection and motion planning

- Reduced 3D models for Collision detection
- Motion planning for dual-arm grasping and re-grasping tasks
  - Probabilistically complete RRT-based algorithms (J^+-RRTs and IK-RRT)

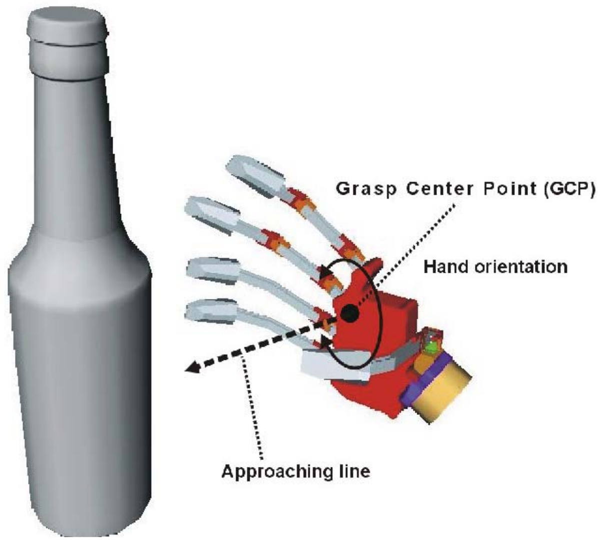


Offline grasp analysis

- Feasible grasps for every object are computed offline and stored together with the object models.

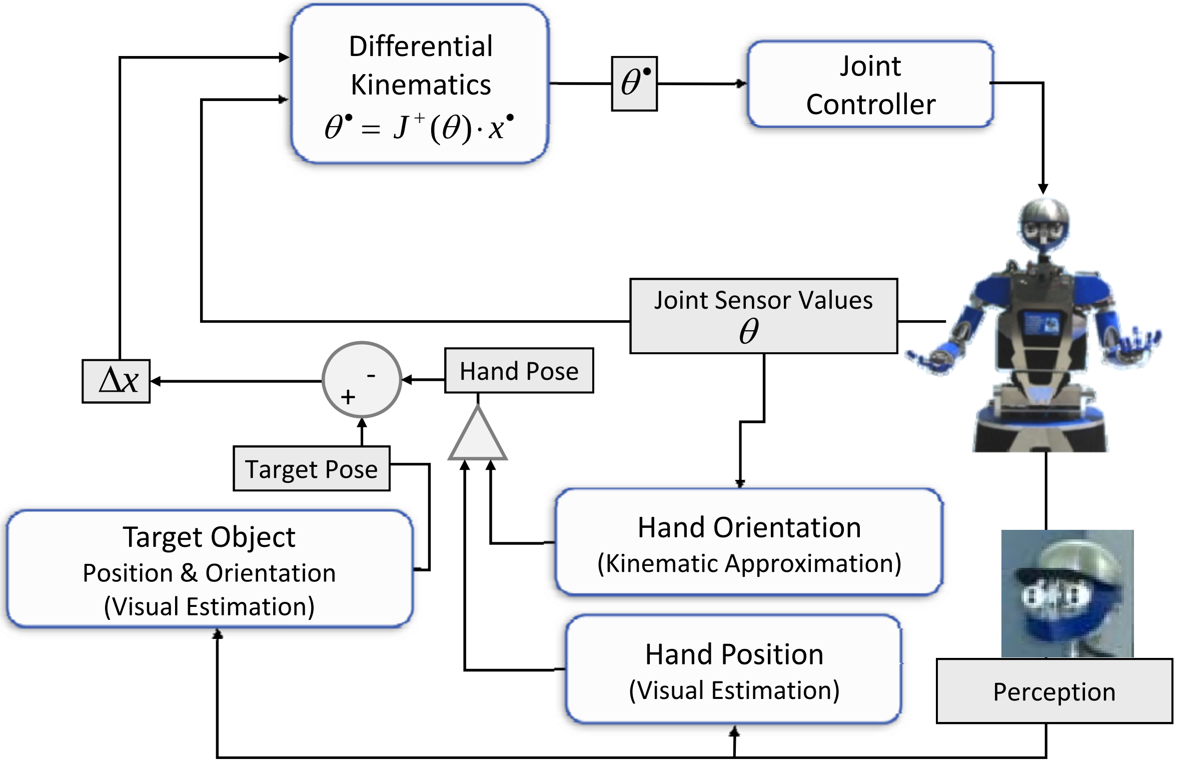
- A grasp is defined by:
  - Grasp type
  - Grasp starting point
  - Approaching direction
  - Hand orientation

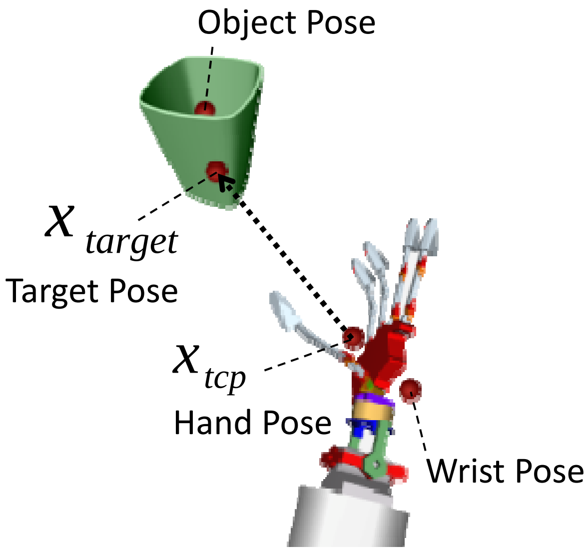
- Simulation environment: GraspIt!



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

Position-based Visual Servoing





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

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

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

First integrated implementation

What can be learned?

- Objects
  - How to learn new objects?
  - How to acquire multi-sensory representations of objects?

- Actions
  - How to build a library of atomic actions?
  - How to learn new actions?
  - How to select proper 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 understand observed and applied actions on objects?

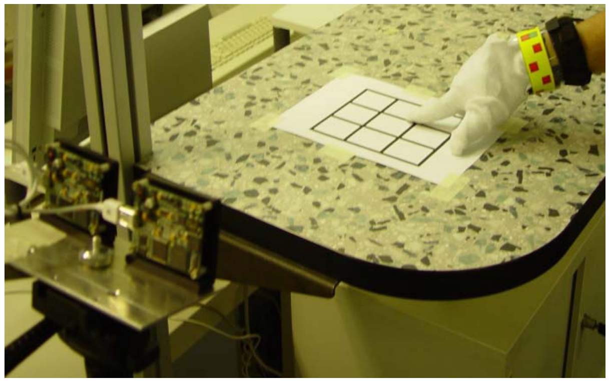
→ grounded symbols for reasoning

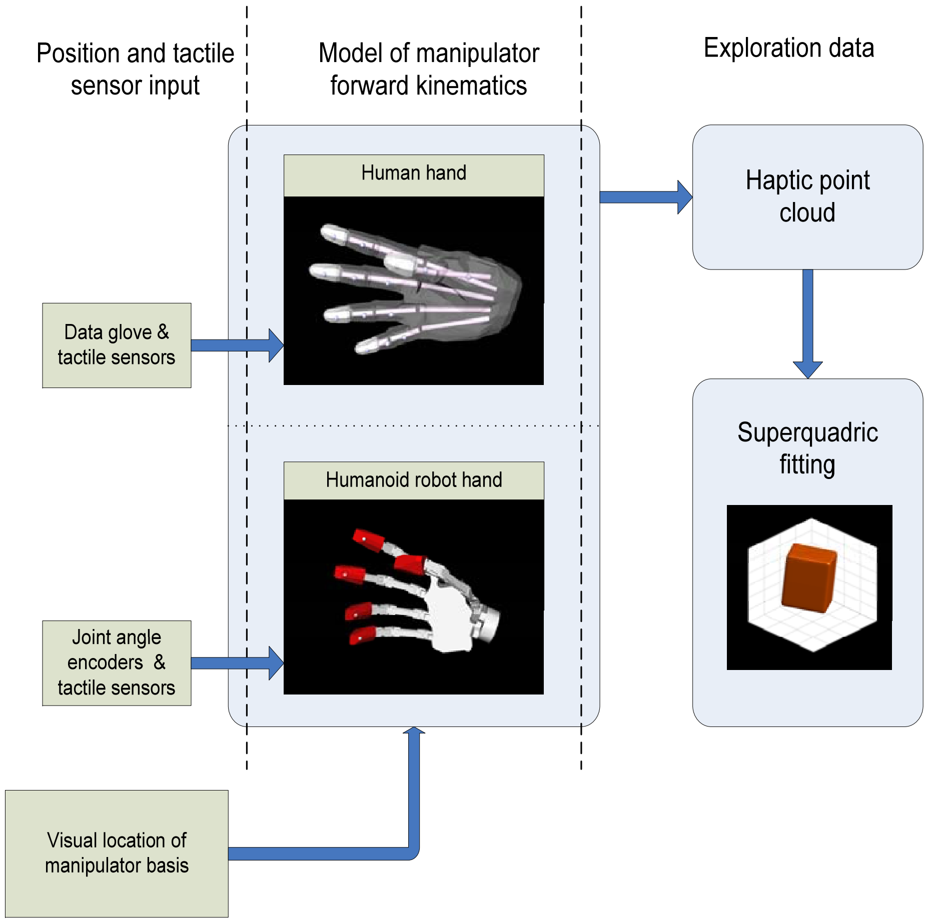
In this talk ...

- Objects:
  - – Multisensorial object representations
- Actions
  - – Learning from Observation
  - – Goal-directed Imitation

Visually-guided haptics exploration of objects

- 3D-Shape reconstruction using Five-Finger hands
- Acquisition and interpretation of 3D shape information from tactile contact and position data.
- Extraction of object shape properties: Super quadric fitting
- Kinematic hand model with 26 DoF s

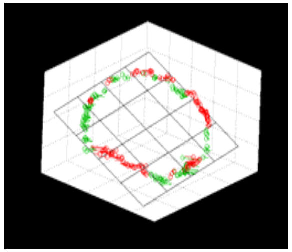


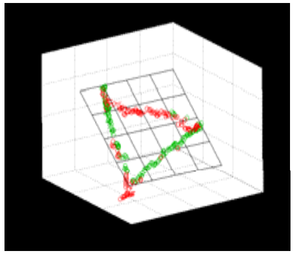


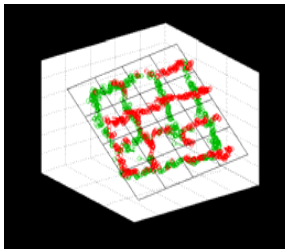
(Humanoids 2007)

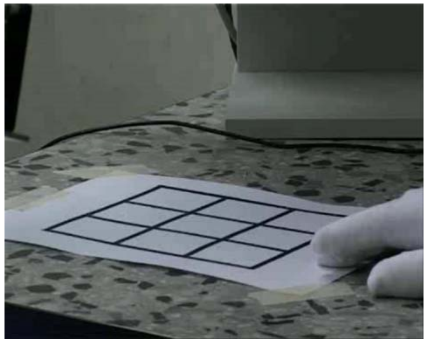
Exploration with dataglove

Contour Following (2D)
of z-plane reference shapes, σ < 6.1mm

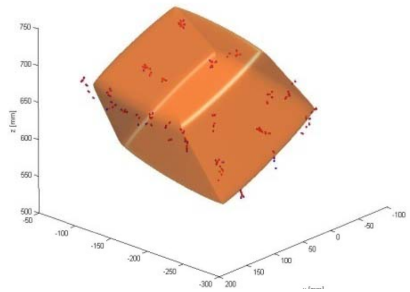


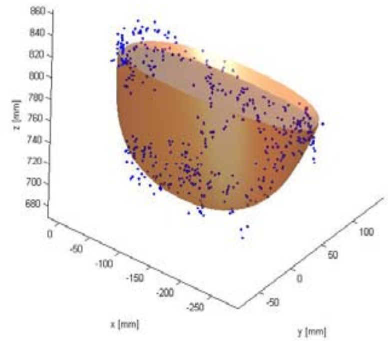






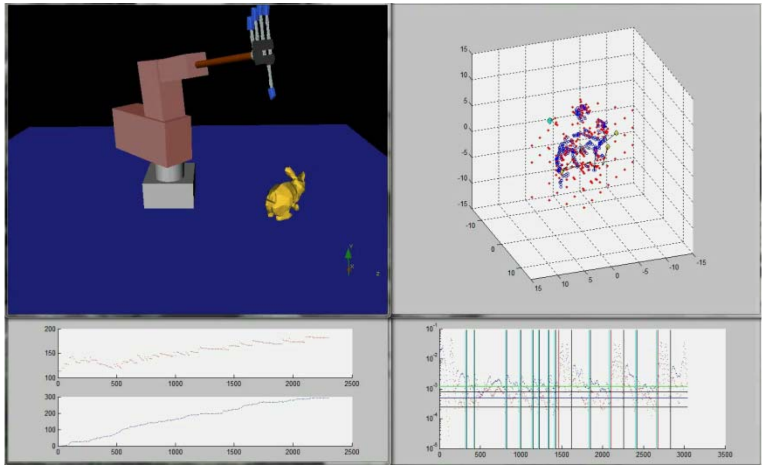
Active Touch Exploration (3D)
of 3D objects and superquadric fitting results.

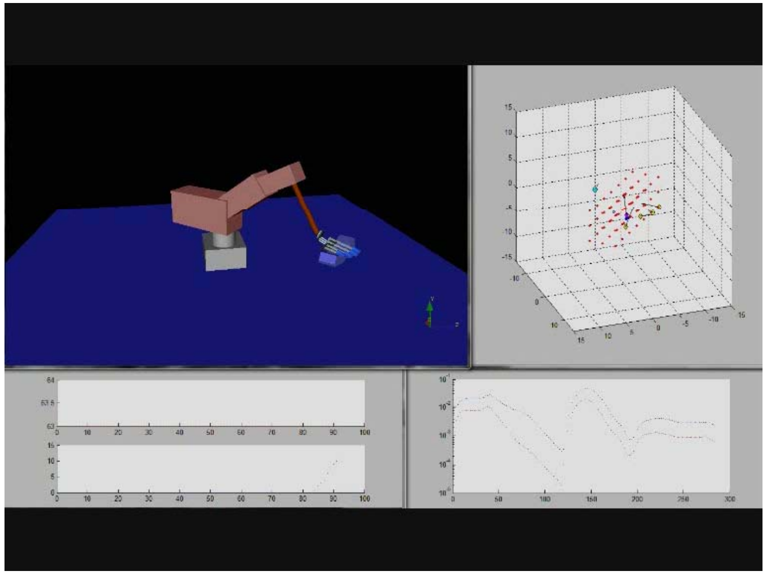




Simulation for multi-finger manipulation strategies

- Coordination strategies for a multi-finger robot hand.
  - Haptic exploration
  - Dynamic grasping
- Potential field control for exploring unknown objects
- Physics extension for Open Inventor/VRML
  - Modelling of complex mechanical systems
  - Modelling of virtual sensors

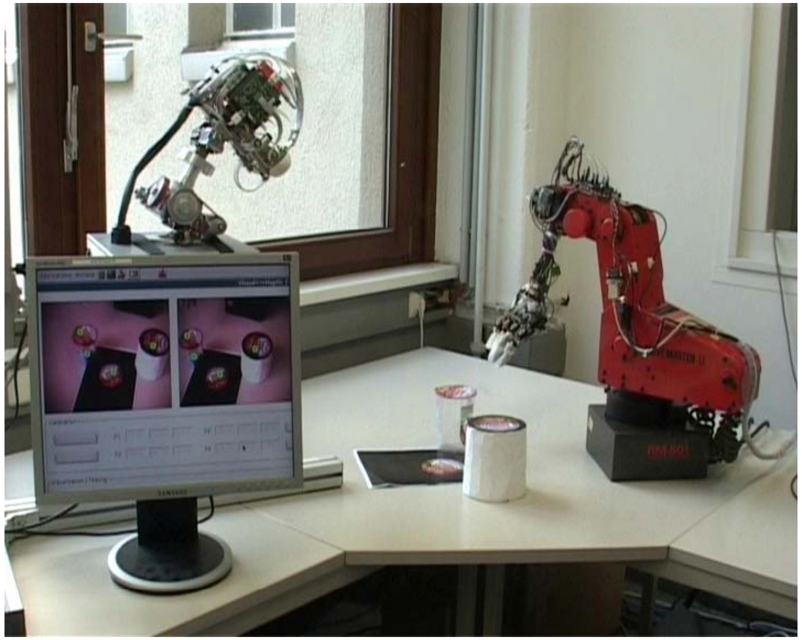




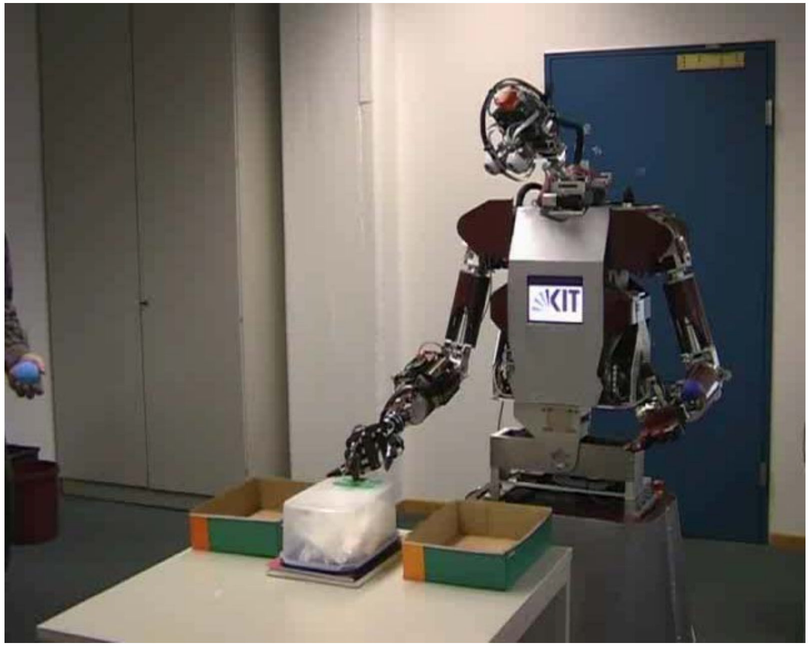
(Humanoids 2008)

Multisensorial object exploration

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



Verification of object size

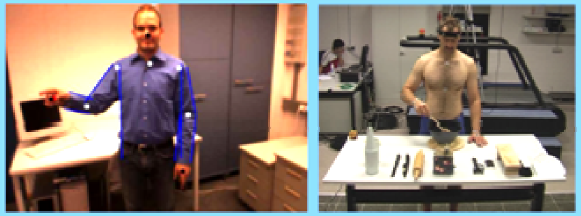


Verification of object deformability

Observation, Generalization and Reproduction

Observation

Beobachtung



- Erfassung
- Erkennung

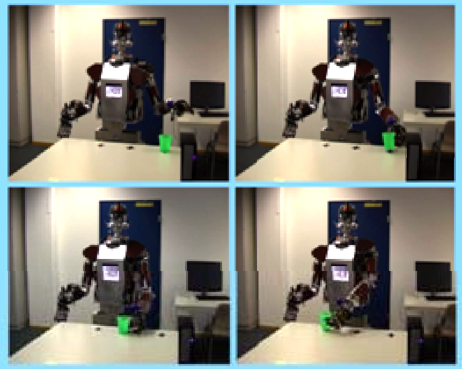
Reproduction

Ausführung

- Greif- und Manipulationsaufgaben
- Online System

Generalization

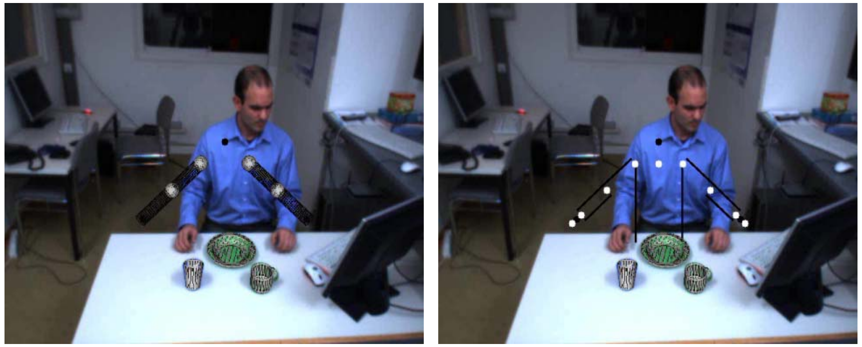
Repräsentation



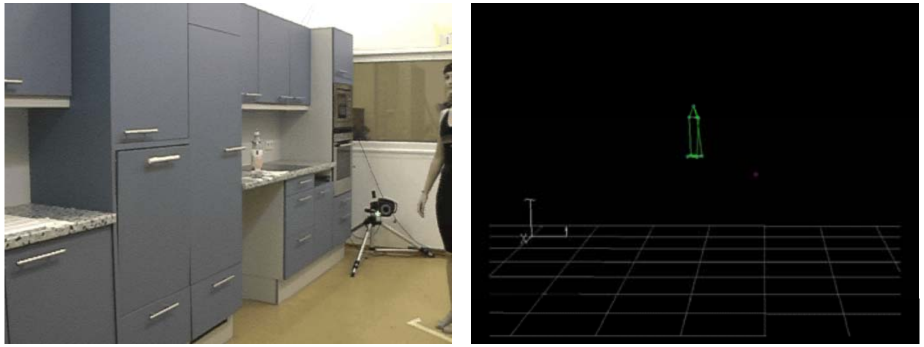
- Trajektorien von Greif- und Manipulationsaufgaben
- Generierung von Bewegungs-primitive
- Bewegungsanalyse

Observation

- Markerless Stereo-based 3D Human Motion Capture
  - Hierarchical particle filter approach
  - Integration of multiple cues
  - Real-time (15 Hz)
  - Integrated with object tracking
  - Image input from the cameras of the robot's head only

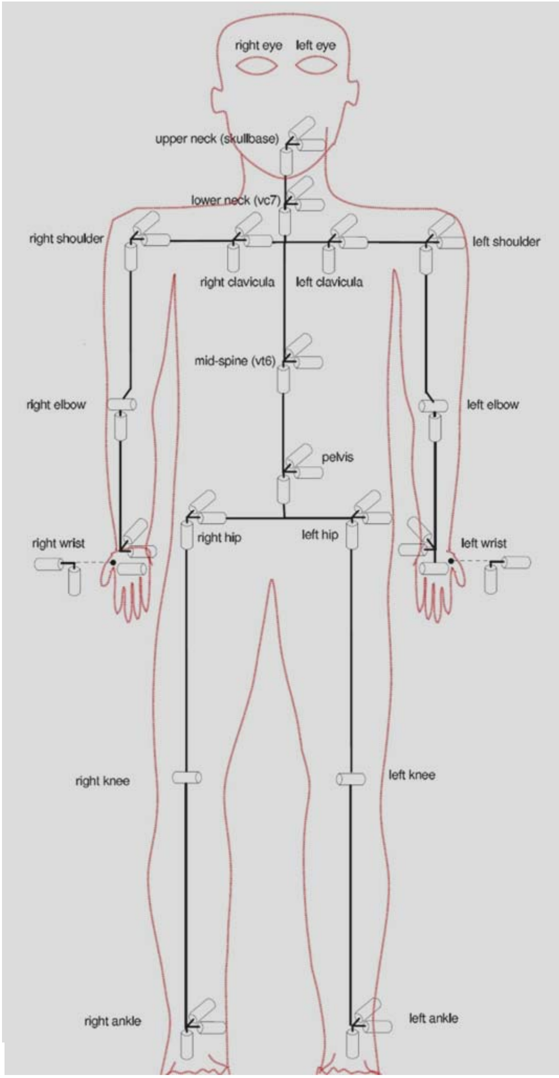


- Marker-based Human Motion Capture
  - Vicon optical motion capture system
  - 10 cameras (120 Hz, resolution: 1000X1000)
  - Standard biomechanical Marker set



Mapping: Master Motor Map (MMM)

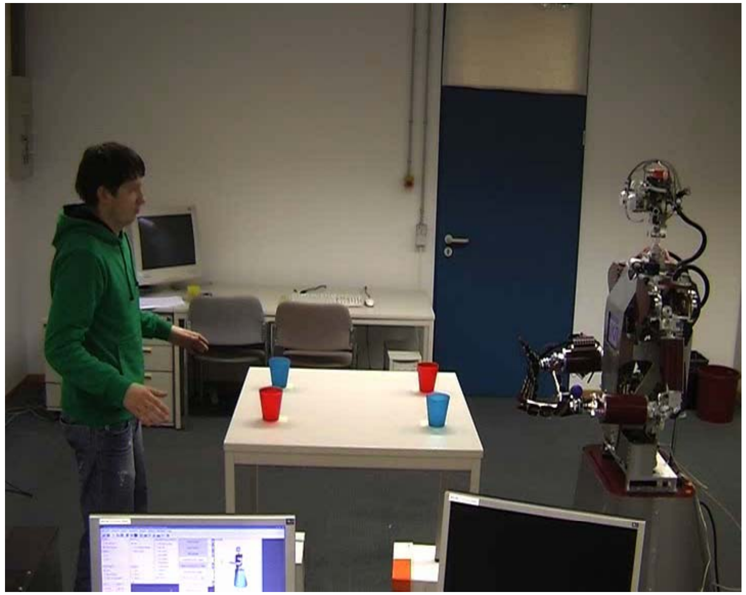
- Various human motion capture systems action recognition systems, imitation systems, visualization modules, and robots → Unified action representation is needed!
- Interface for the transfer of motor knowledge between different embodiments: Master Motor Map (MMM)
- Specification of a reference kinematic model of the human body
  - 52 DoF → does not limit any human motion capture system
  - Kinematics is similar to kinematics of humanoid robot systems
  - File format is fully specified



http://wwwiaim.ira.uka.de/users/asfour/mmm (ICRA 2007)

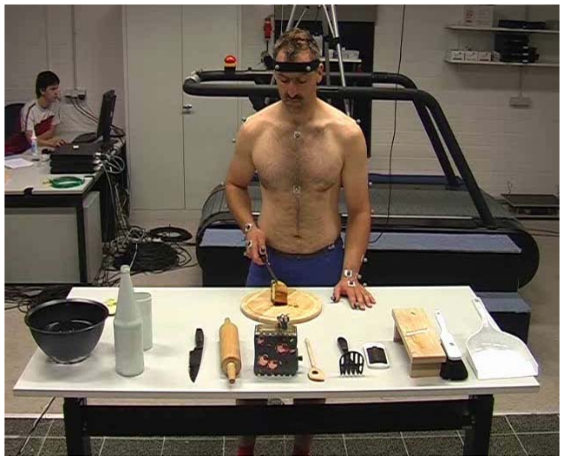
Movement reproduction on the robot

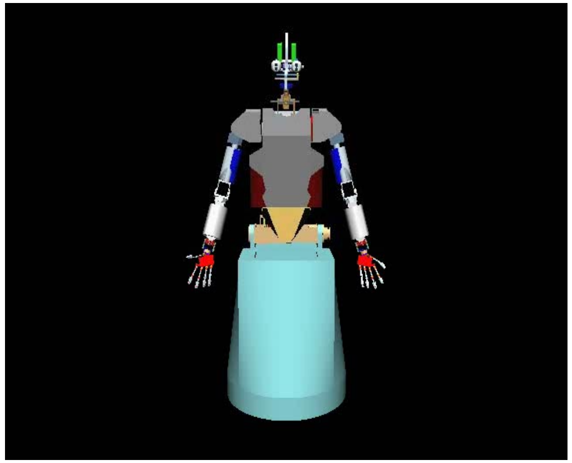
- Data from stereo-based markerless human motion capture system
  - Images captured at 15 Hz
  - Tracking of human and object motion
  - Visual servoing for grasping



- Data from VICON system

Generalisation?



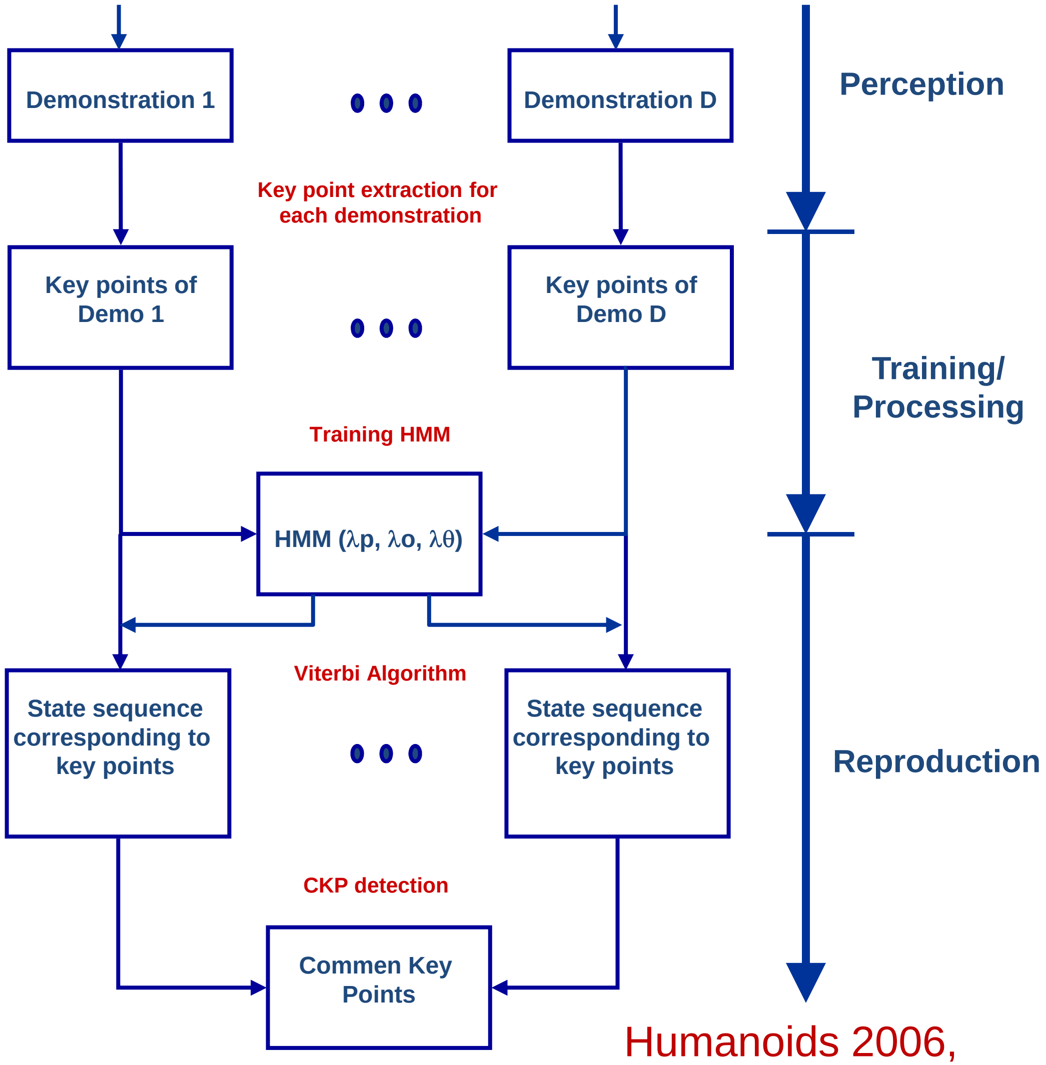


Action representation

- Hidden Markov Models (HMM)
  - 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)
  - Ijspeert, Nakanishi & Schaal, 2002
  - Trajectory formulation using canonical systems of differential equations
  - Parameters are estimated using locally weighted regression
- Spline-based representations
  - 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 HMMs

- Continuous left-right HMMs are used to generalize movements from multiple demonstrations.
- HMMs are trained with key points of each demonstration
- Use HMMs to match key points across demonstrations → Common key points (CPK)
  - means of the output density functions
  - Average of the timestamps
- Movement representation through the resulting CKPs

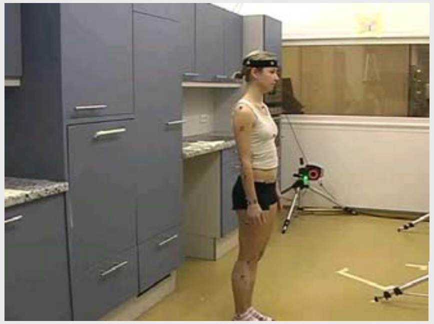


Humanoids 2006, IJHR 2008

Capturing, modeling and recognition of human motion

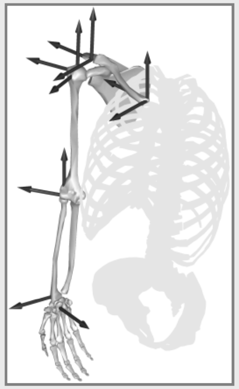
- Gesture

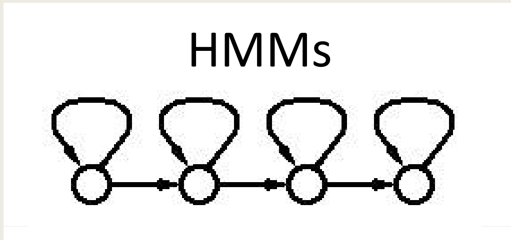
Motion Capture



IR-Tracking

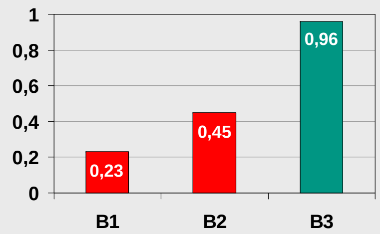
Modelling





Biomechanics and stochastic models

Recognition

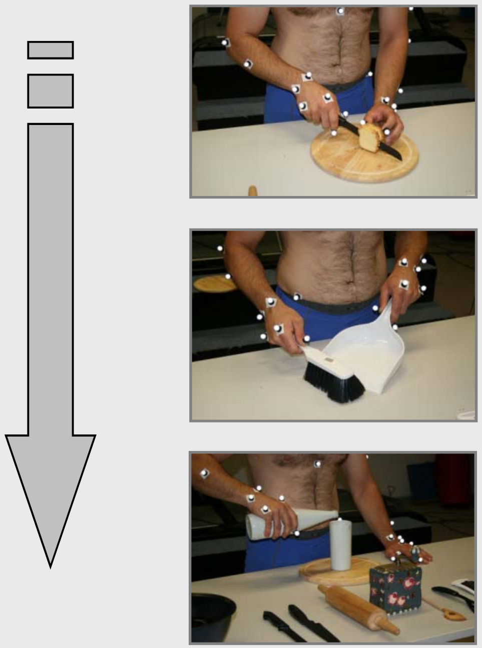


B3 → *Gestures*

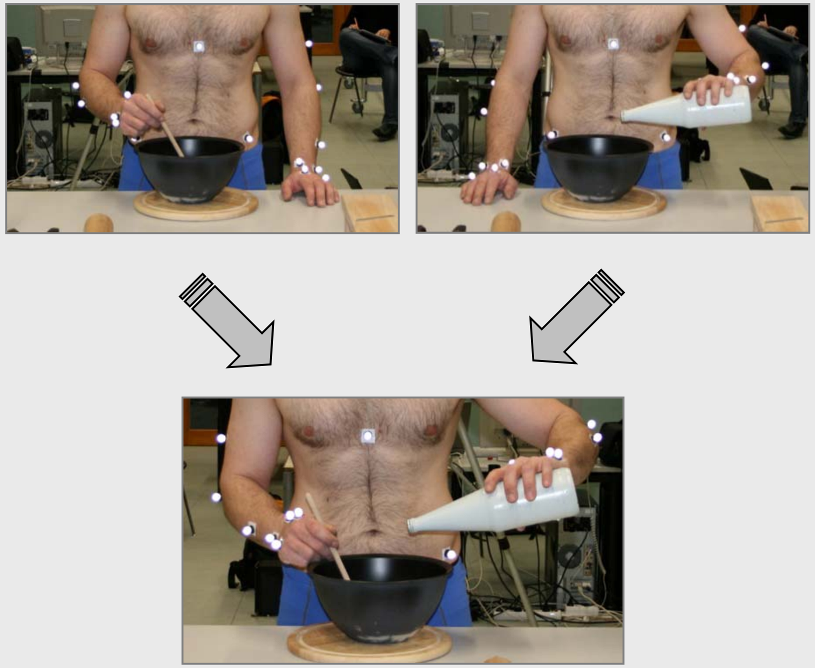
Capturing, modeling and recognition of human motion

- Dual arm manipulation and grasping tasks

Sequential movements



concurrent movements



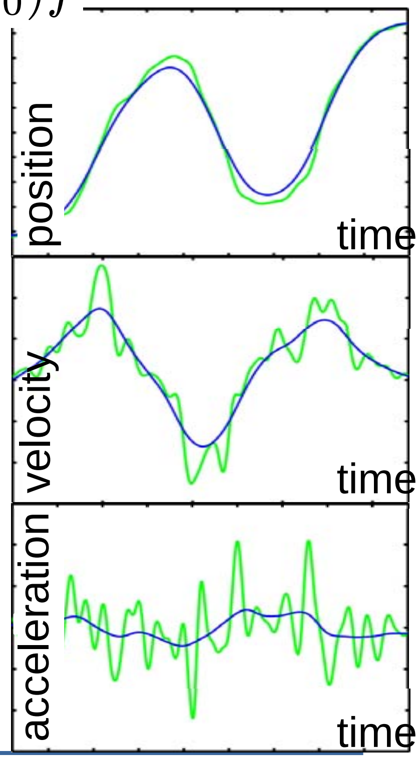
Action representation using DMPs

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

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

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

$\tau\dot{x} = v$

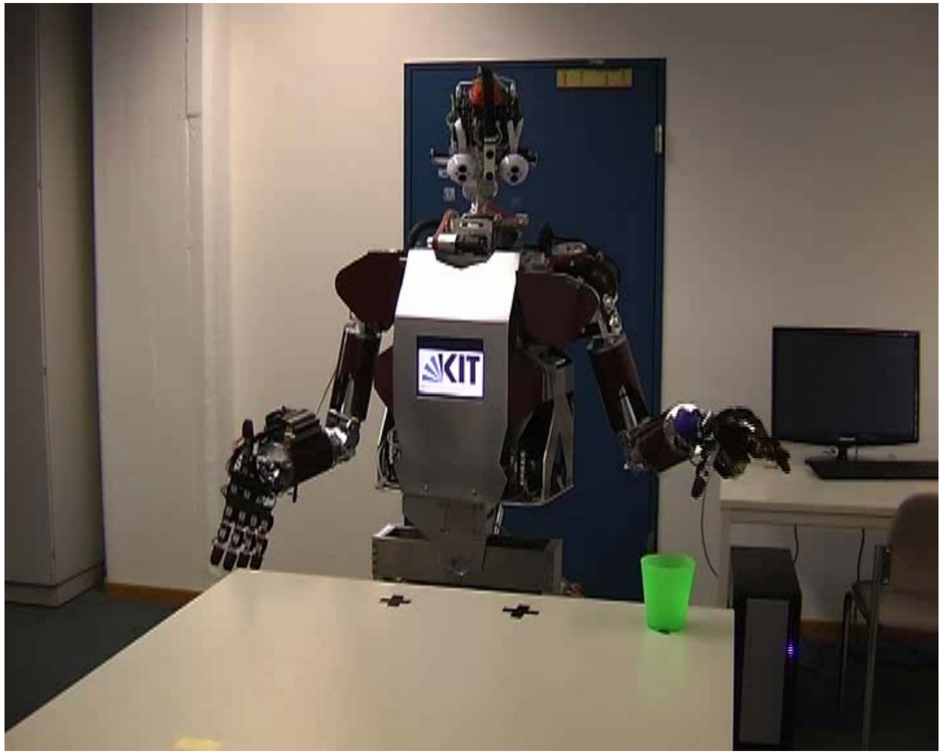


- Canonical system of differential equations for point to point movement
- 1D demonstration (green) and movement reproduction (blue) using a DMP

Action representation using DMPs

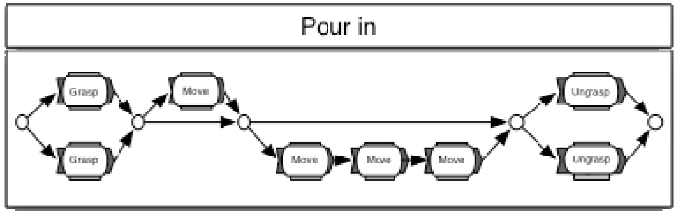
- Generation of object manipulation movements using DMPs
- Adaption to new situations
- Associating semantic information with DMPs

→ sequencing of movement primitives → Planning



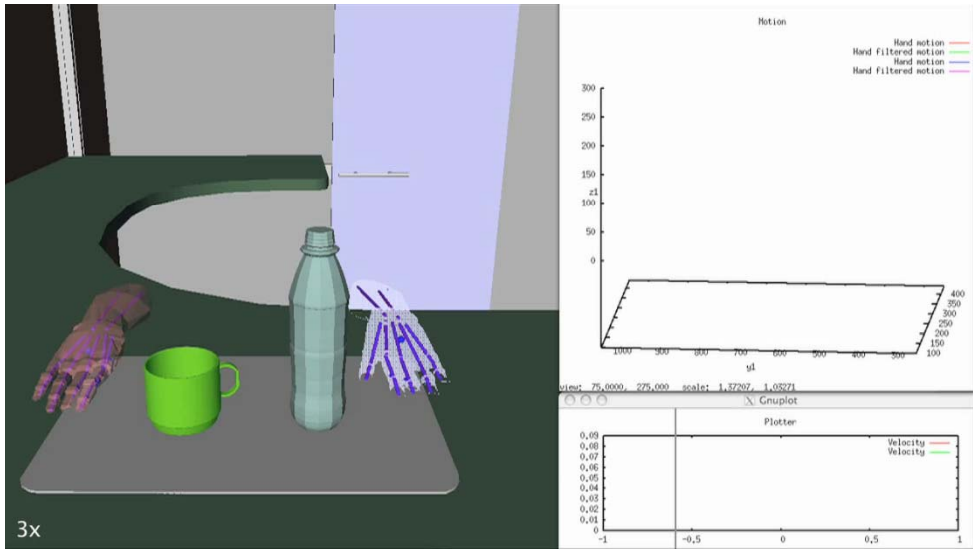
Learning of manipulation tasks

- Observation of human demonstrations of manipulation tasks in a multi-sensor environment
- Learning of symbolic actions and tasks based on the demonstrations (Programming by Demonstration)
- Learning of manipulation constraints and strategies



Representation of the learned "pour in"

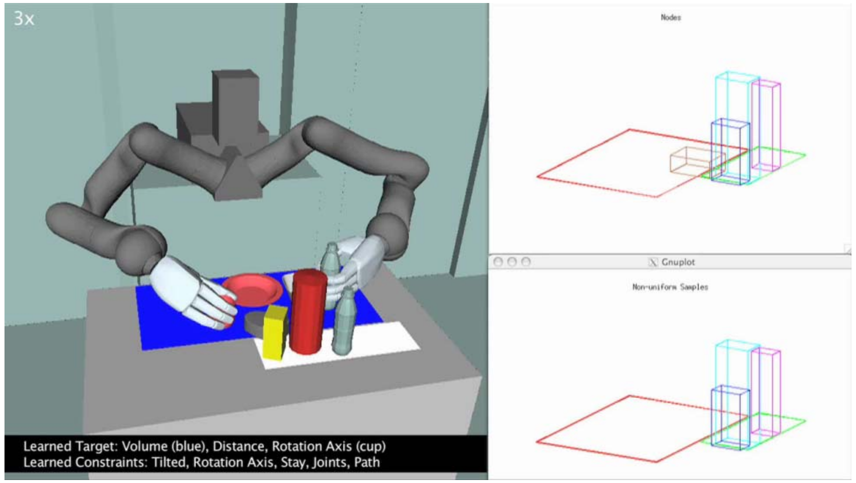
Sensory environment



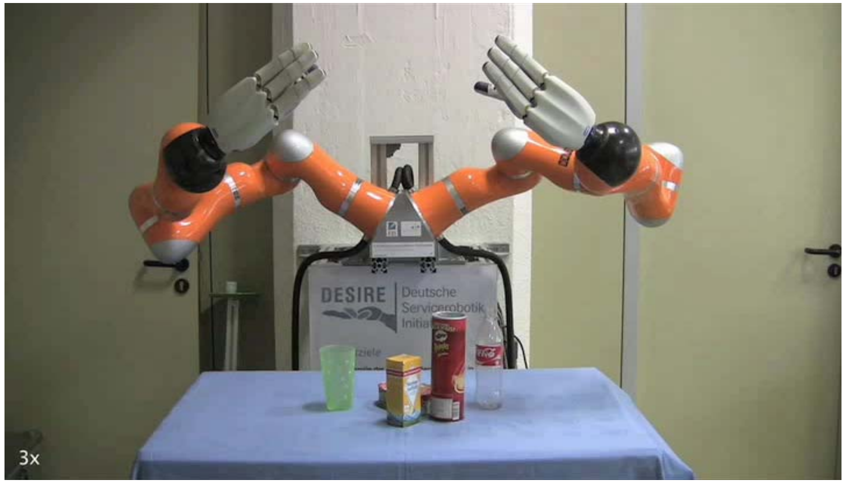
Learning of the "pour in" manipulation task

Execution of learned manipulation tasks

- Formulation of a planning problem based on the task representation with learned constraints and strategies
- Instantiation of the learned constraints and strategies in the execution environment (knowledge transformation)
- Solving of the planning problem using state-of-the-art motion and grasp planning



Knowledge transformation & Planning



Execution of the “pour in” manipulation task

Conclusions

- Fully integrated autonomous humanoid robot ARMAR-III
- First steps toward humanoid robots in human-centered environments
  - Grasping and manipulation
  - Learning from observation
  - Natural interaction
- Strategies for acquisition of object and action representations in the motor and sensor space of the robot
  - Multi-sensorial object representation
  - Markerless human and object movement tracking
  - Action representations, which support adaption to novel situations

Thanks

- Humanoids Group @ IAIM
  - 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
  - James Kuffner (CMU,USA), Ales Ude (JSI, Slovenia), Gordon Cheng (ATR, Japan), Giorgio Metta (IIT, Italy), Stefan Schaal (USC, USA)

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 project PACO-PLUS funded by the European Commission
    www.paco-plus.org
  - the EU Cognitive Systems project GRASP funded by the European Commission
    www.grasp-project.eu
  - the EU Cognitive Systems project DEXMART funded by the European Commission
    www.dexmart.eu