Transfer of Human Movements to Humanoid Robots

M. Do, P. Azad, T. Asfour, P. Pastor, D. Gehrig, H. Kühne, A. Wörner, T. Schultz, S. Schaal, R. Dillmann

Universität Karlsruhe (TH)

University of Southern California (USC)

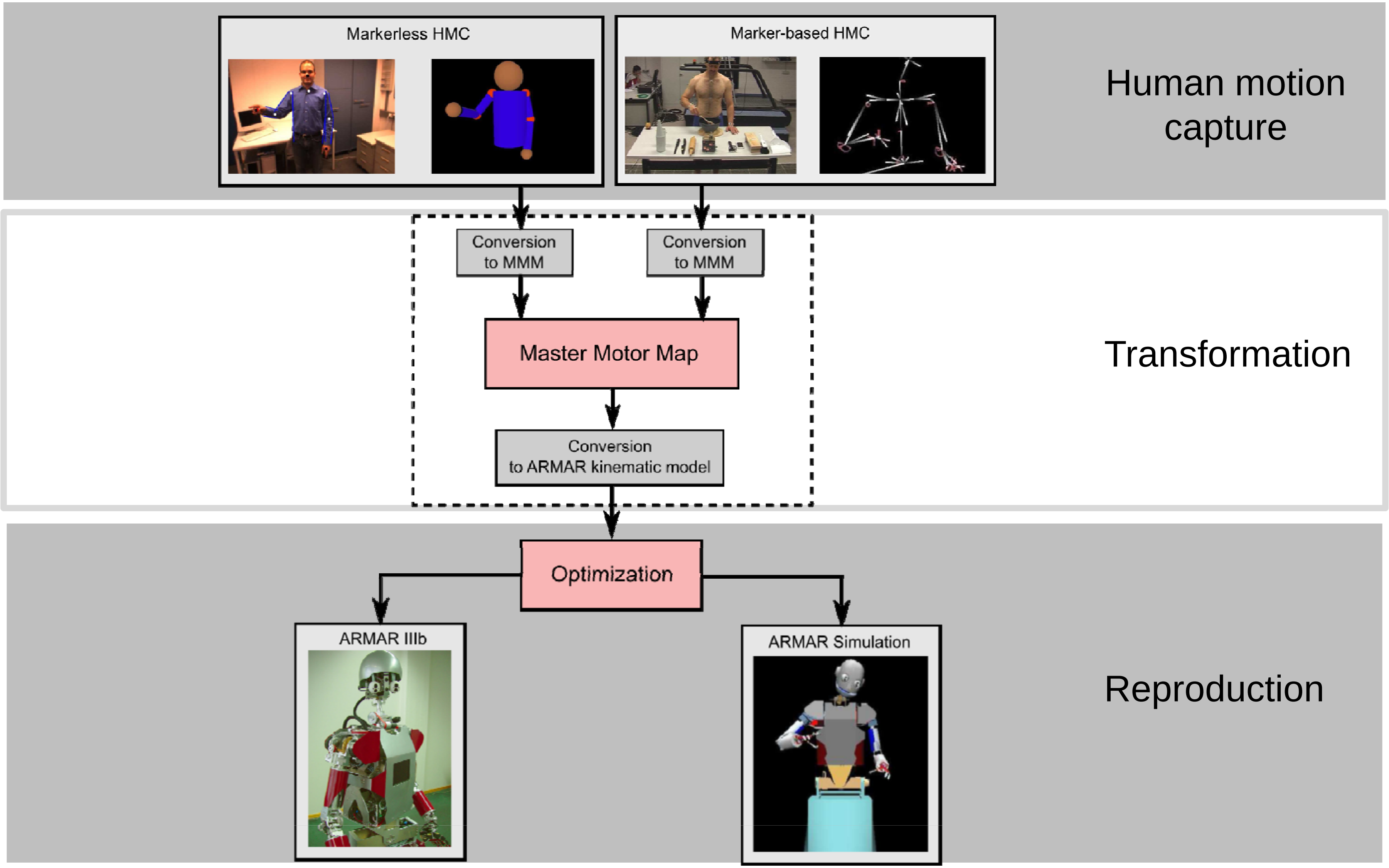
Overview

- Motivation
- Human motion capture
- Transferring human motion data to Robot
  - Master Motor Map
  - Optimization
- Action Representation
  - HMM
  - DMP
- Conclusions

Motivation

- In order to enable a humanoid to interact with humans and human-centered environments, incorporation of human motion is needed
- Imitation of human motion is a promising way of intuitive programming of humanoid robots
- Requirements for successful imitation
  - Human motion capture
  - Mapping to different embodiments with different kinematics
  - Action representation
- The goals is to have a humanoid robot which observes a human performing a specific task and reproduces the task

Framework overview



Markerless HMC

Marker-based HMC

Human motion capture

Conversion to MMM

Conversion to MMM

Master Motor Map

Conversion to ARMAR kinematic model

Transformation

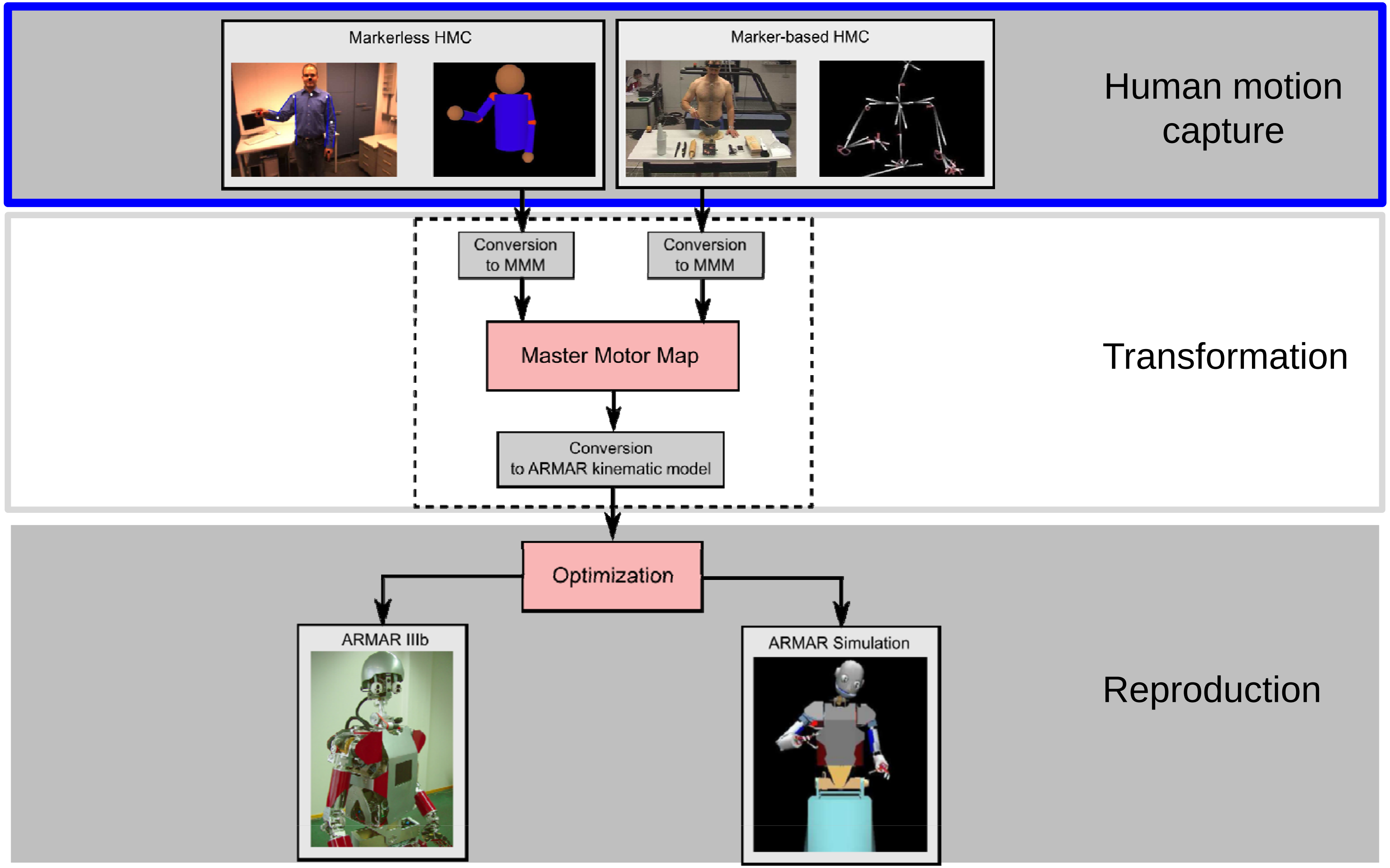
Optimization

ARMAR IIIb

ARMAR Simulation

Reproduction

Framework overview



Markerless HMC

Marker-based HMC

Human motion capture

Conversion to MMM

Conversion to MMM

Master Motor Map

Conversion to ARMAR kinematic model

Transformation

Optimization

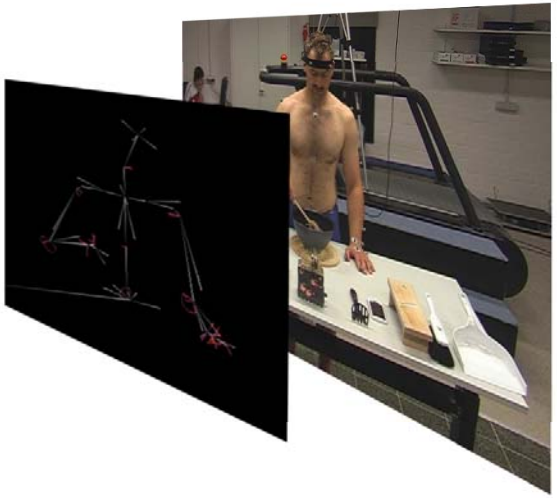
ARMAR IIIb

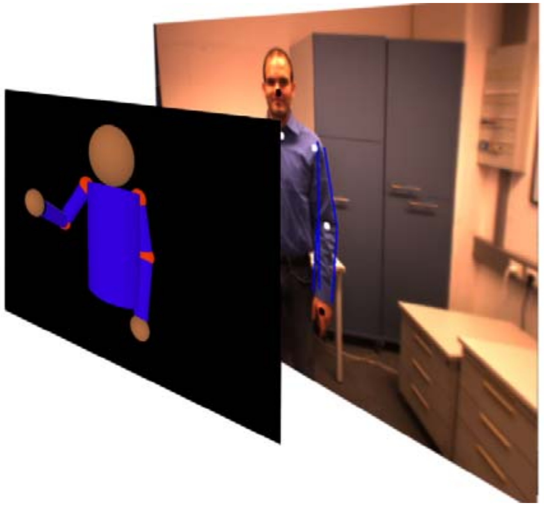
ARMAR Simulation

Reproduction

Human Motion Capture

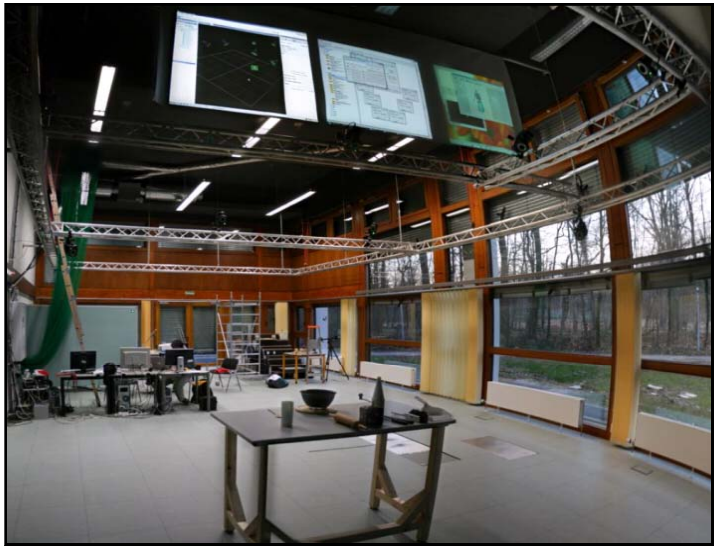
- Human motion is captured with optical systems
- Two different model-based approaches:
  - Marker-based
  - Markerless
- Search for point correspondences between the model and the observed human motion
- Joint angles are obtained from the model
- Trajectory of a human movement is described as a sequence of joint angle configurations





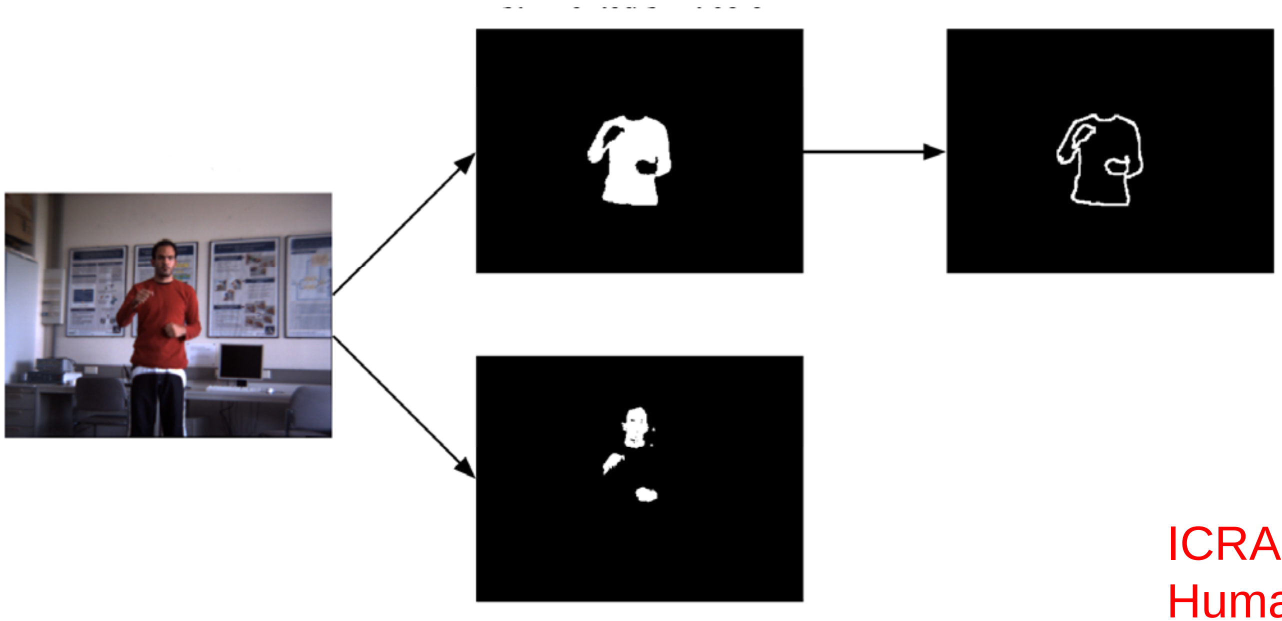
Marker-based Human Motion Capture

- Vicon System
- Using 10 infrared cameras and reflective markers placed on the upper body
- Human Motion is captured with 100 Hz
- Applicable in controlled environments



Markerless Human Motion Capture

- Using the stereo vision system of the humanoid
- Exploits an edge cue and a distance cue
- Particle filter is exploited to track the human upper body
- Human Motion is captured with 15 Hz



ICRA 2007
Humanoids 2008

Comparison

Marker-based

+ High accuracy
+ High Resolution
+ Fast and robust

– High Costs
– Complex handling and preparation
– Expensive post-processing
– In controlled environments

Markerless

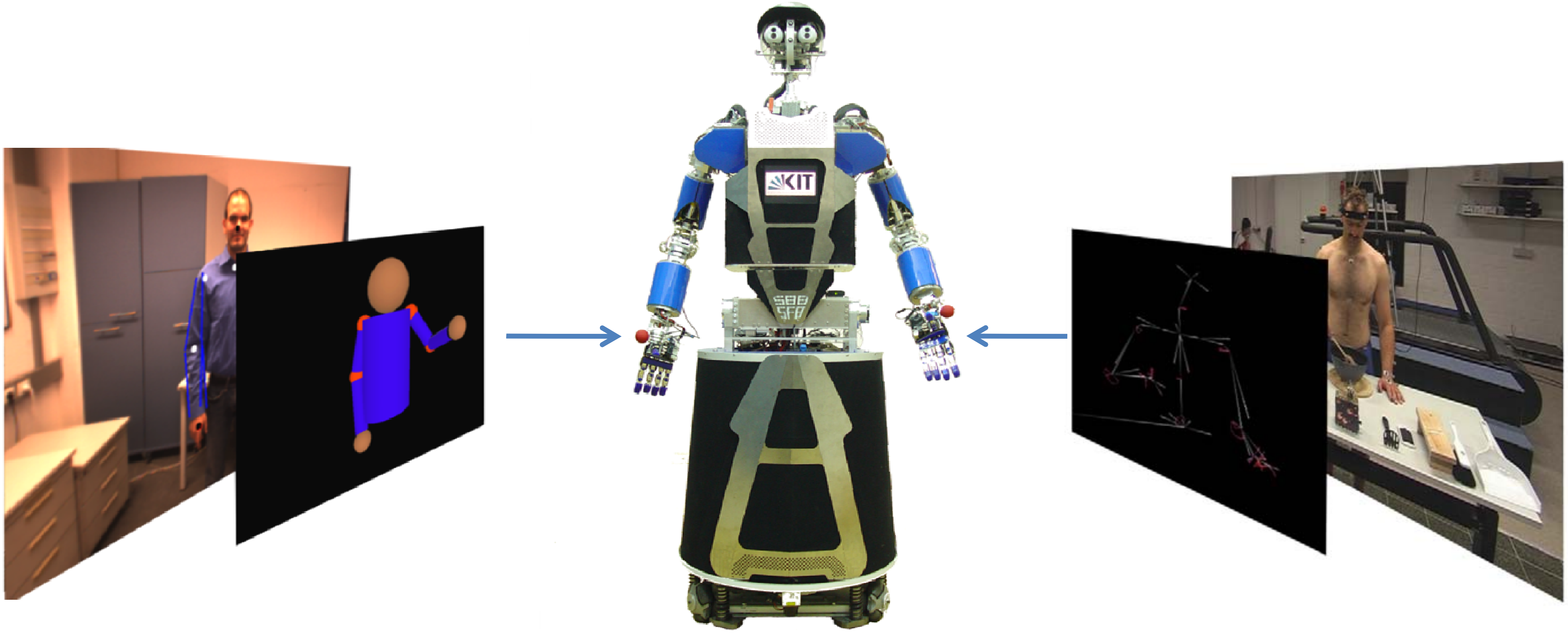
- Low accuracy
- Low Resolution and Occlusion
- Lower frame rates and sensitive to noise and environmental changes

+ Low hardware costs
+ Freedom of movement
+ Easy-to-use
+ Natural way of human motion capture
+ Anywhere and on-site

Transferring human motion to Humanoid

- Human motion data from various systems
- Different robot platforms

How do we transfer human motion data ?



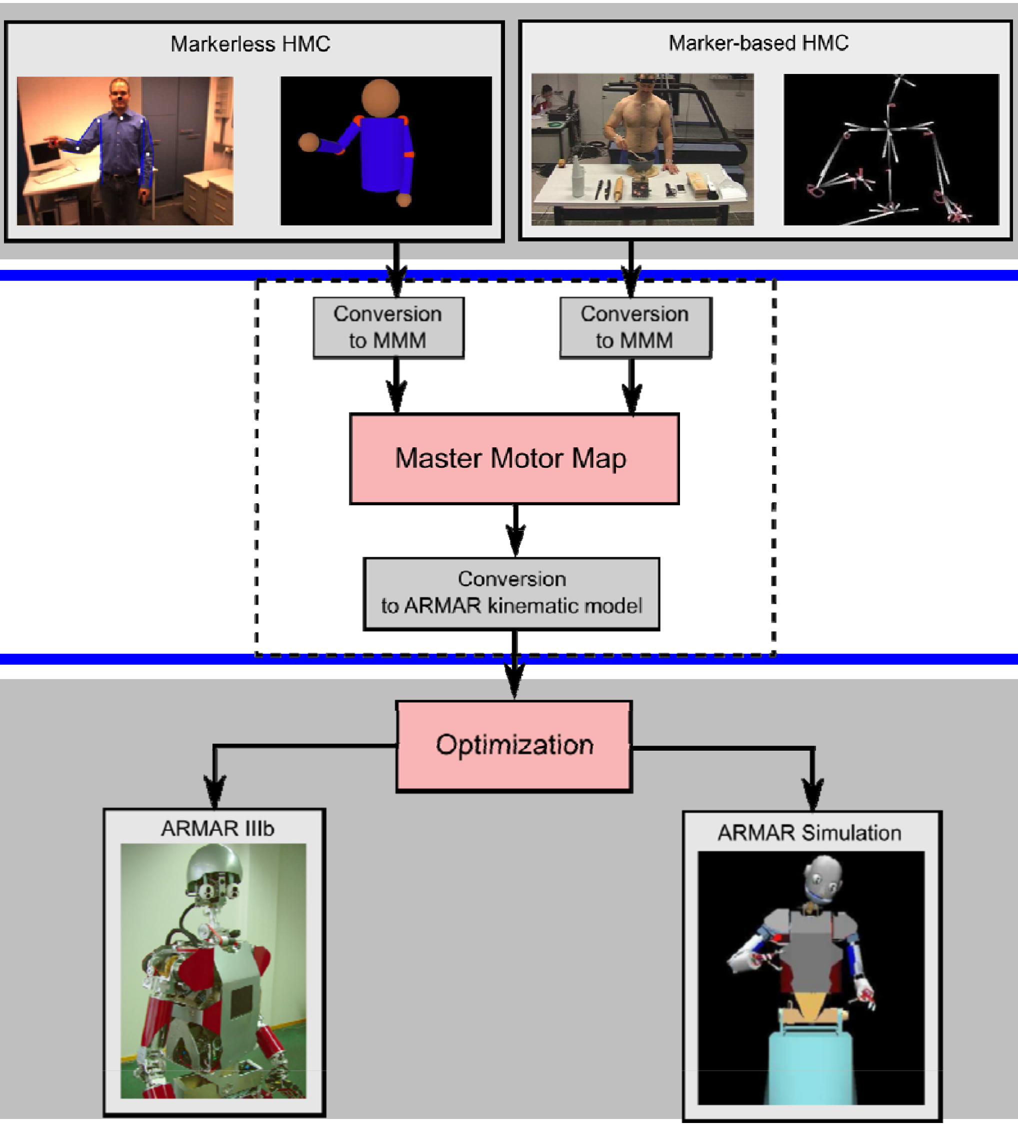
Transferring human motion to Humanoid (2)

Problems

- Motion capture systems with specific models and different file formats
- Various number of joints, which can be captured
- Different kinematic structures between the human model and the Humanoid:
  - Reduction of joints
  - Joint constraints
  - Scaling

Instead of implementing an interface for each human motion capture system and a robot, we use the **Master Motor Map**

Framework overview



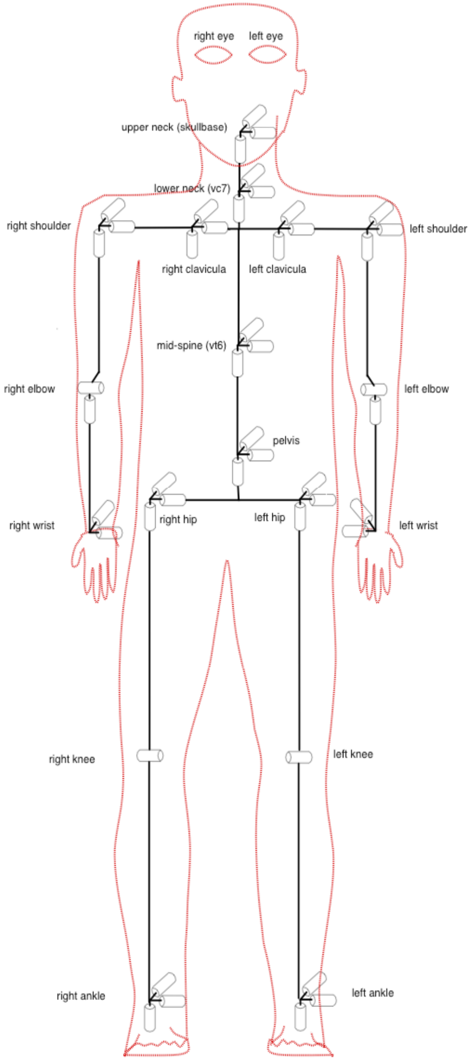
Human motion capture

Transformation

Reproduction

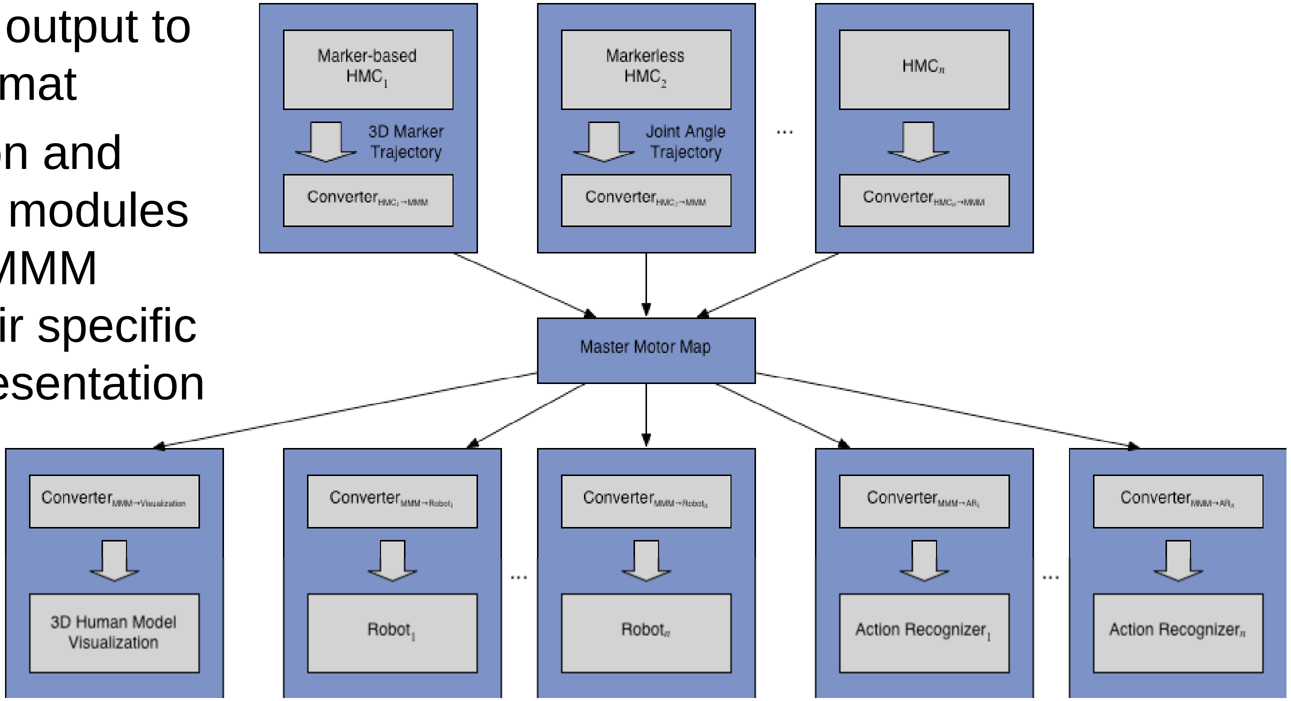
Master Motor Map

- Various human motion capture systems action recognition systems, imitation systems, visualization modules, and robot systems for reproduction → Unified 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+6 DoF → does not limit any human motion capture system
  - Kinematics is similar to kinematics of humanoid robot systems
  - File format is fully specified



Master Motor Map (2)

- Replacement of any module (perception, recognition, visualization, reproduction) can be guaranteed by using the MMM as the exchange format
  - All perceptive module convert their output to the MMM format
  - All recognition and reproduction modules convert the MMM format to their specific internal representation



http://wwwiaim.ira.uka.de/users/asfour/mmm

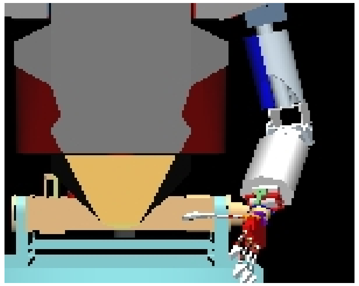
(ICRA 2007)

Master Motor Map (3)

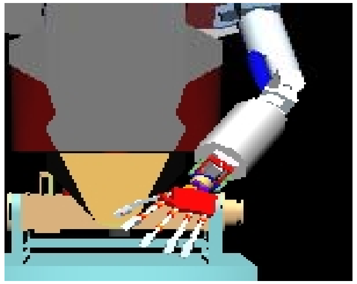
- Trajectory coming from the MMM interface has to be modified in order to:
  - Reproduce a movement on the robot similar to the observed human
  - Retain properties of a movement e.g. goal-directedness
  - Prevent the robot of violating its joint constraints

- Previous approaches
  - Exploit point correspondences between robot and human e.g. with physical markers
  - Minimize the distance between joint angle configurations

Human joint angle configuration

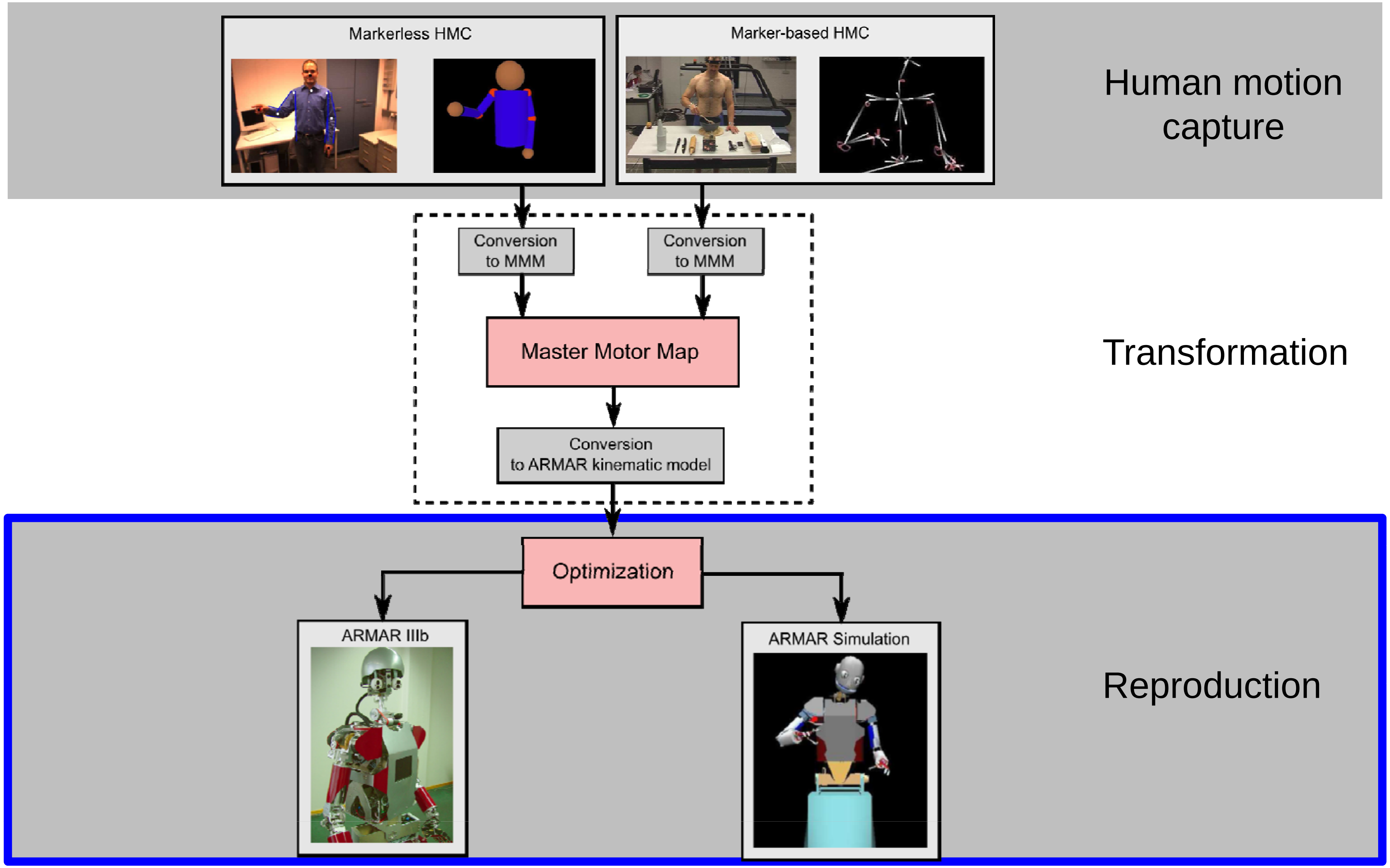


Mapped on robot **without optimization**



Mapped on robot **with optimization**

Framework overview



Human motion capture

Transformation

Reproduction

Similarity as objective function for optimization

- Two main features to determine the similarity between a robot configuration and a human configuration
  - Joint angles
  - Point correspondences
- By incorporating both, Similarity function S is defined:

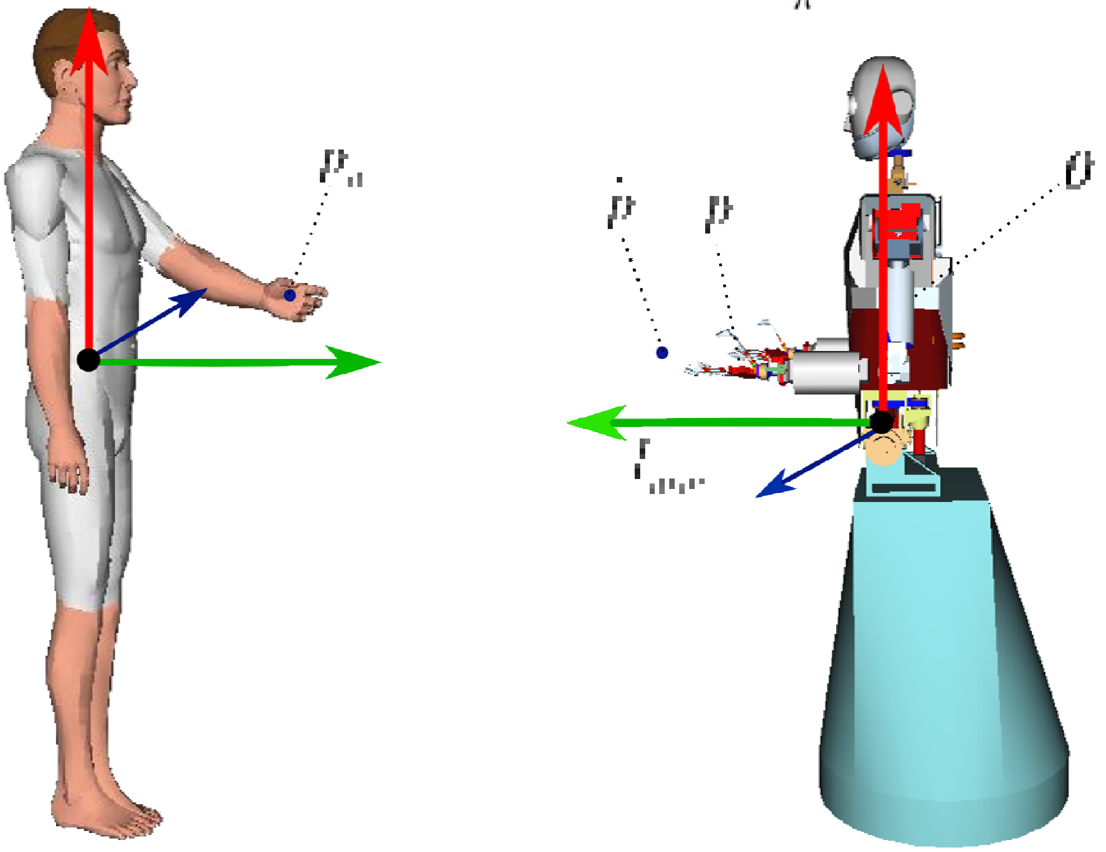
$$S(\theta) = 2 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right)^2}{\pi^2} - \frac{\frac{1}{3}\sum_{k=1}^{3}\left(\hat{p}_k - p_k\right)^2}{\left(2 \cdot l_{arm}\right)^2}$$

with the TCP positions $p, \hat{p}$, joint angle configurations $\theta, \hat{\theta}$

- A solution, which minimizes S can be found by applying a non-linear optimization algorithm

Illustration of the parameters of the similarity function

$$S(\theta) = 2 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\bar{\theta}_i - \theta_i\right)^2}{\pi^2} - \frac{\frac{1}{3}\sum_{k=1}^{3}\left(\hat{p_k} - p_k\right)^2}{\left(2 \cdot l_{arm}\right)^2}$$



Definition of the optimization problem

- Maximization of the similarity function can be described by the optimization problem:

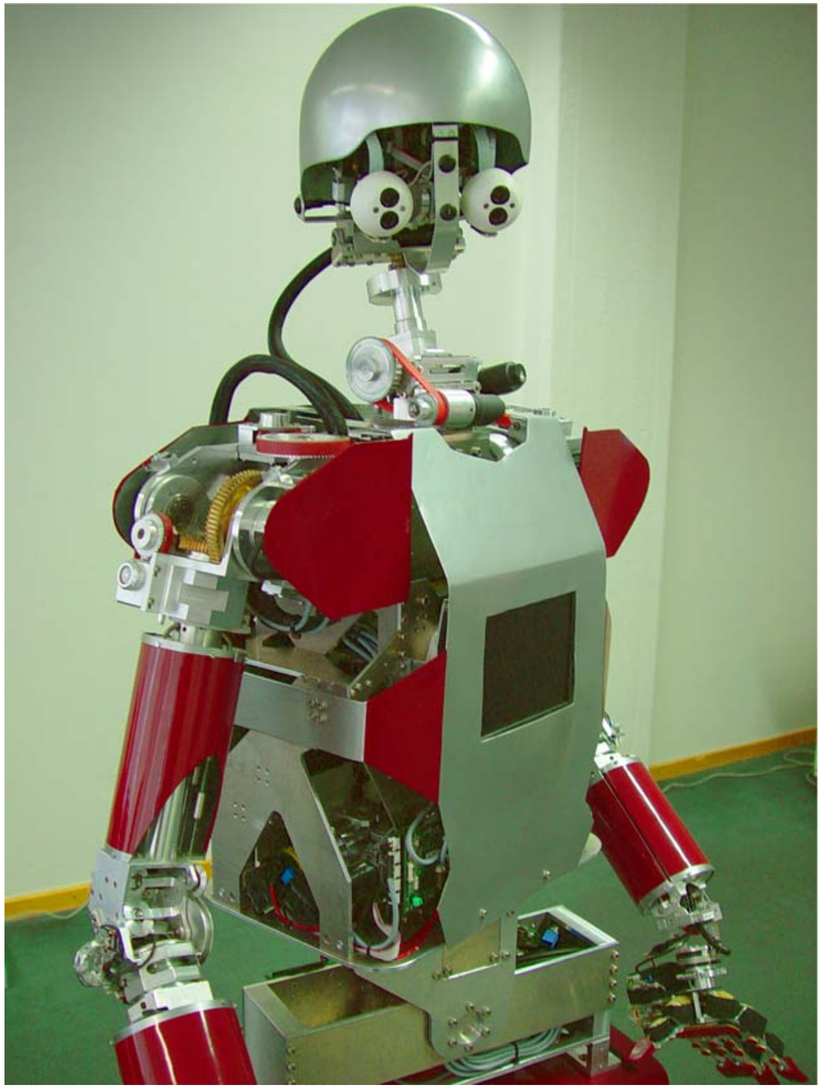
$$\min S'(\phi) = 2 - S(\phi)$$

$$\text{subject to} \quad C_{i_{min}} \leq \phi_i \leq C_{i_{max}}$$

- Levenberg-Marquardt algorithm for solving non-linear least squares problems
  - Combination of Gauss-Newton and steepest descent method
  - Robust and fast convergence
  - Avoidance of local minima by starting optimization with the best initial candidates

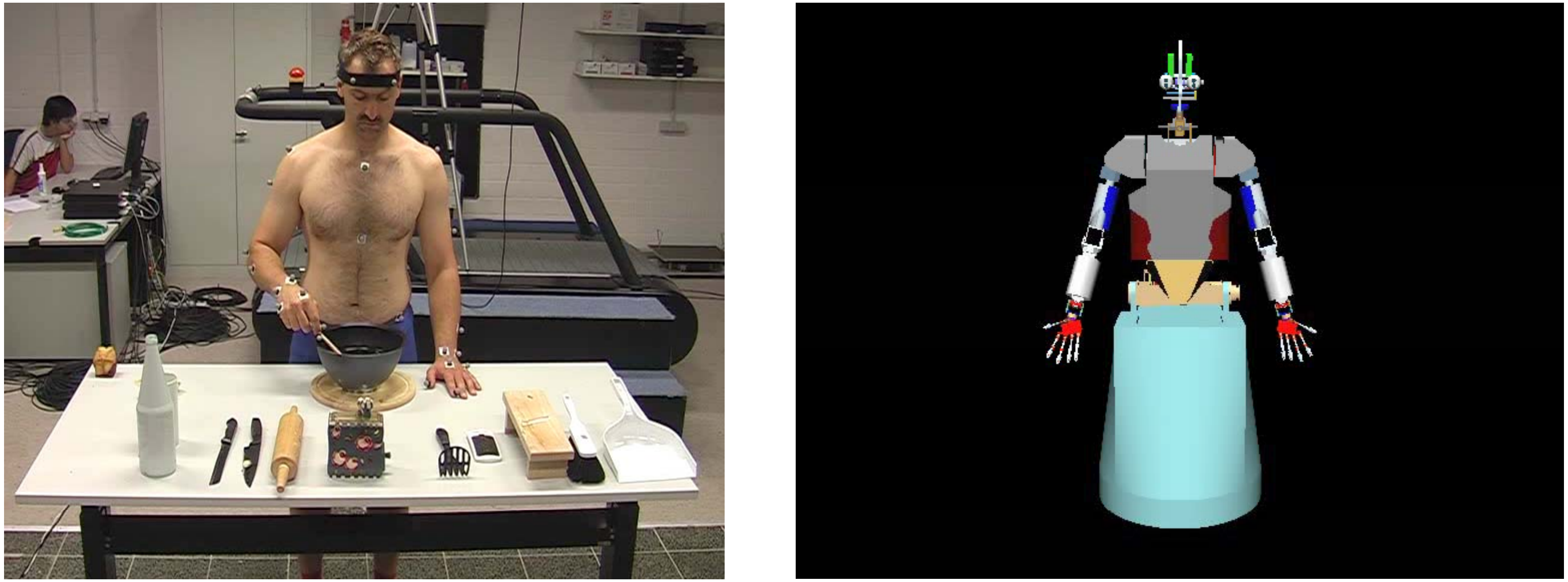
Experimental platform: ARMAR-IIIb

- Humanoid Robot with seven subsystems:
  - Head (7 DoF)
  - Right arm (7 DoF)
  - Left arm (7 DoF)
  - Right hand (8 DoF)
  - Left hand (8 DoF)
  - Torso (3 DoF)
  - Mobile platform
- Head equipped with a stereo camera system



Reproduction of human motion (1)

- **Marker-based**



- Kitchen actions (here “Stirring”) were recorded and reproduced in the ARMAR Simulation
- More joints can be controlled in simulation

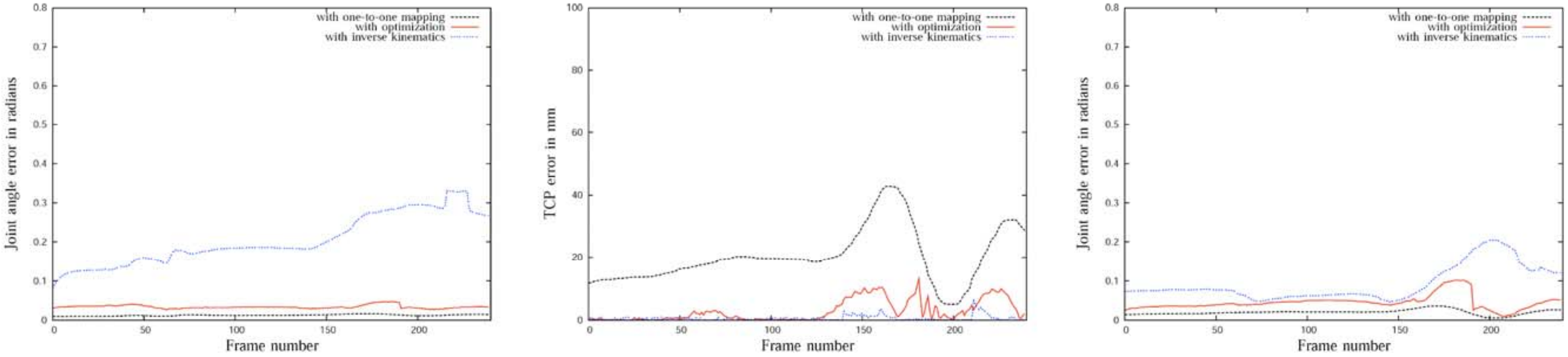
Reproduction of human motion (2)

- **Markerless**

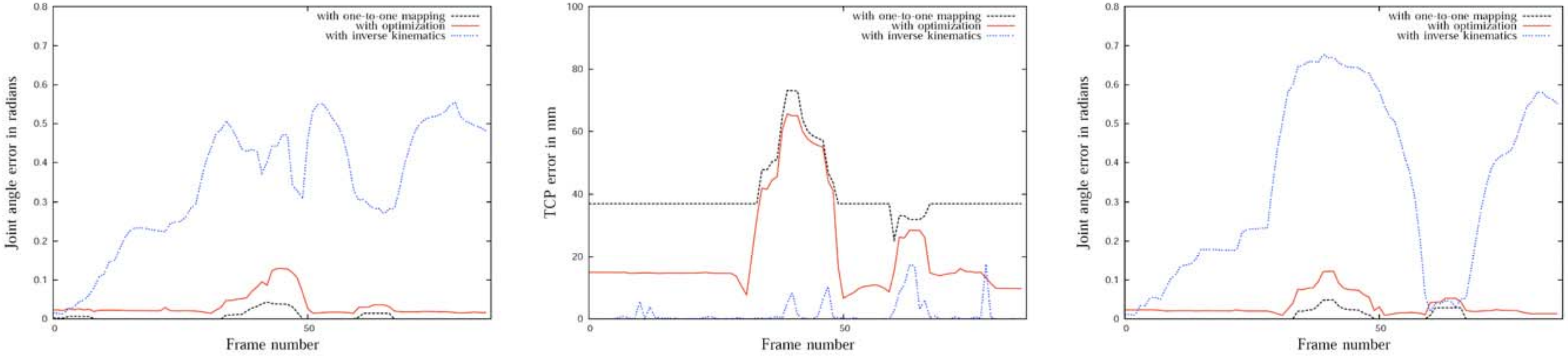
- On-site capturing allows online imitation of observed human motion
- Less joints available

Further examples

- Reproduction of marker-based captured human motion data



- Reproduction of markerless captured human motion data



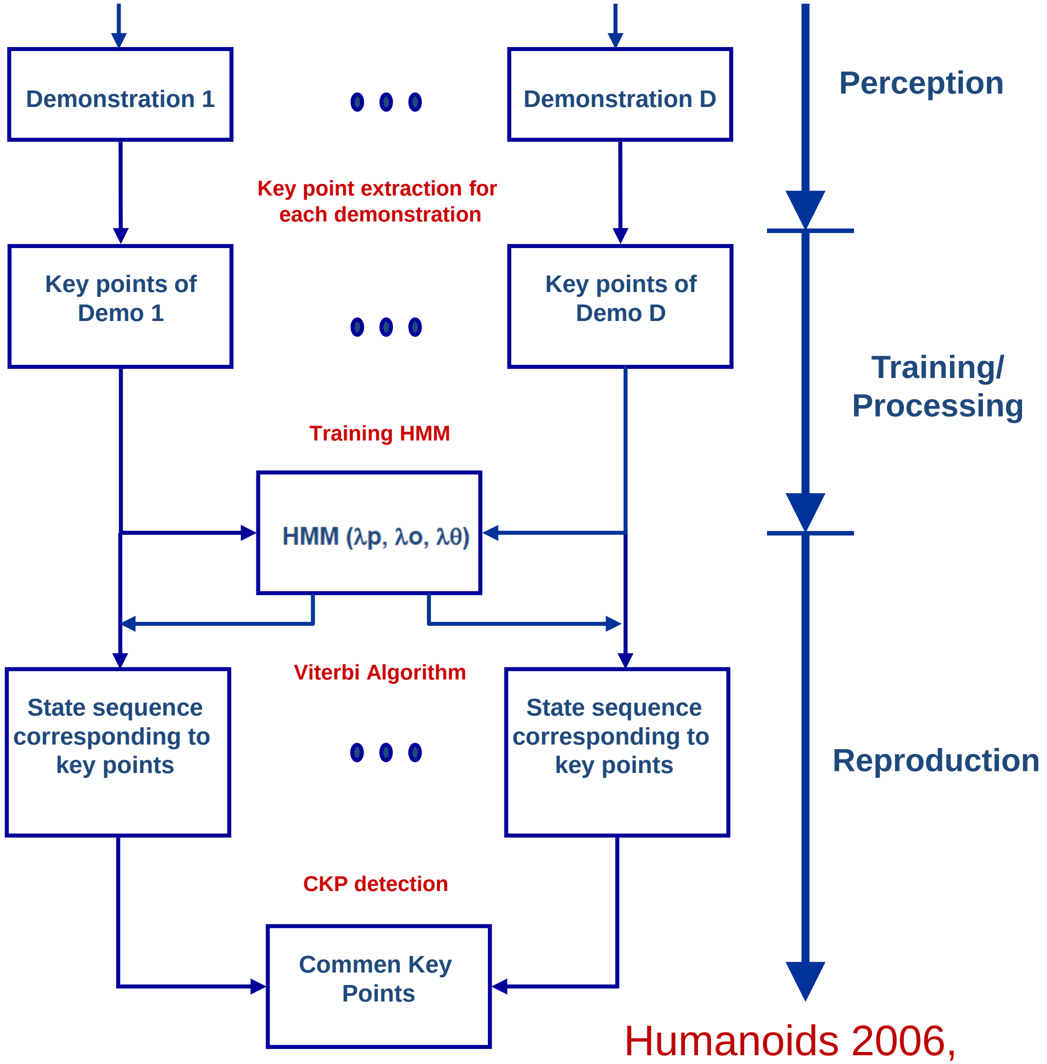
Action representation

- Hidden Markov models (HMM)
  - Extract key points (KP) in the demonstration
  - Determine key points that are common to all demonstrations (common key points: CKP)
  - Reproduction through interpolation between CKPs

- Dynamic movement primitives (Ijspeert, Nakanishi & Schaal, 2002)
  - Trajectory formulation using canonical systems of differential equations
  - Parameters are estimated using locally weighted regression

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



Demonstration 1

• • •

Demonstration D

Perception

Key point extraction for each demonstration

Key points of Demo 1

• • •

Key points of Demo D

Training/ Processing

Training HMM

HMM (λp, λo, $\lambda \theta$)

Viterbi Algorithm

State sequence corresponding to key points

• • •

State sequence corresponding to key points

Reproduction

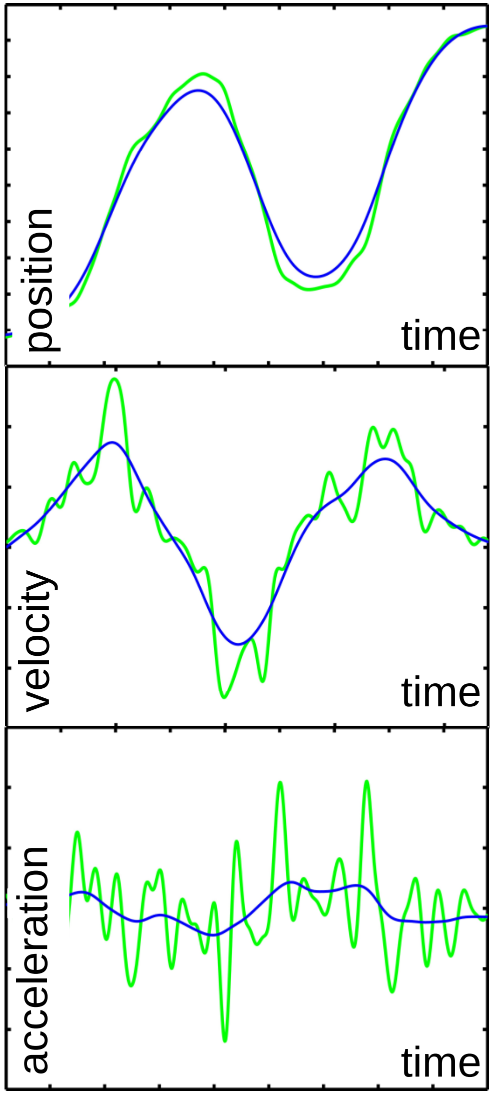
CKP detection

Commen Key Points

Humanoids 2006, IJHR 2008

Action representation using DMPs

- DMP: Dynamic movement primitives
- Canonical system of differential equations for point to point movement
- 1D demonstration (green) and movement reproduction (blue) using a DMP
  - Approximation accuracy depends on the number of basis functions, the more basis functions, the less approximation error - thus, primitive movements can be arbitrarily complex
  - Restricting the learning parameters defines the “length” of a primitive and can be used to segment movement trajectories



Action representation using DMPs

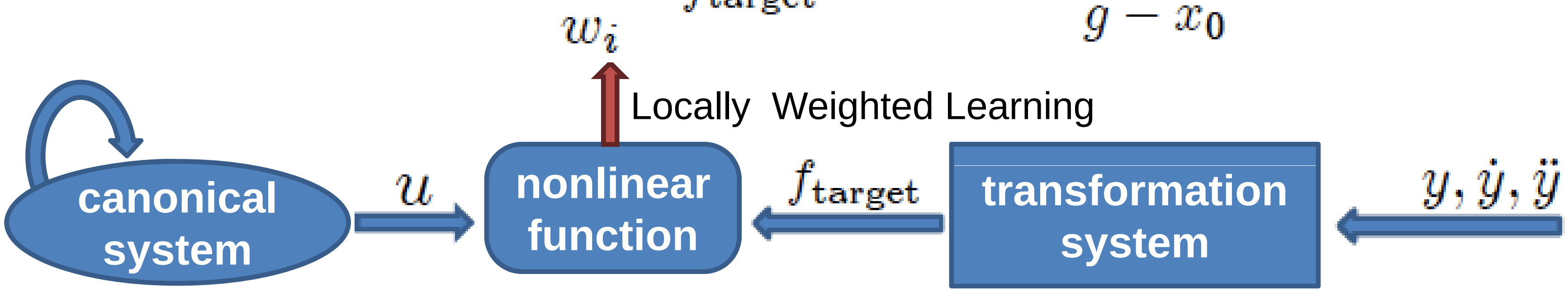
canonical system: $\tau \dot{u} = -\alpha u$

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

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

$\tau \dot{x} = v$

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



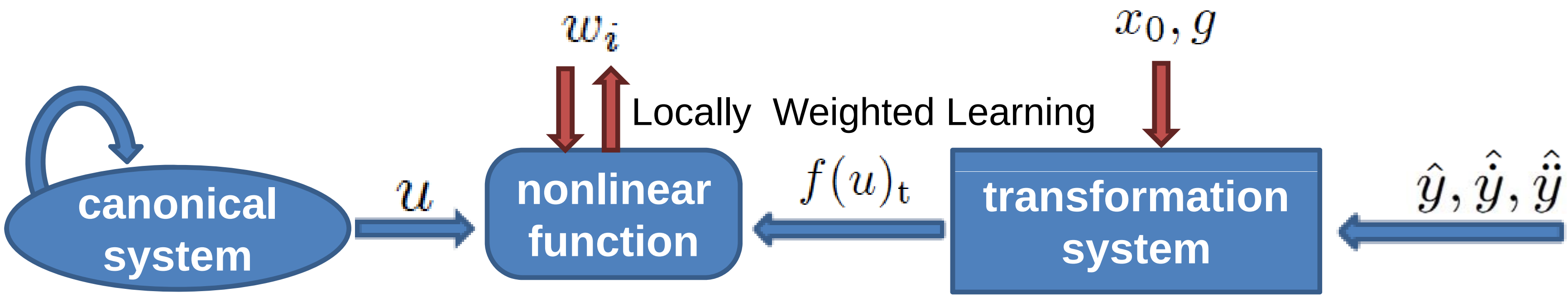
Action representation using DMPs

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

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

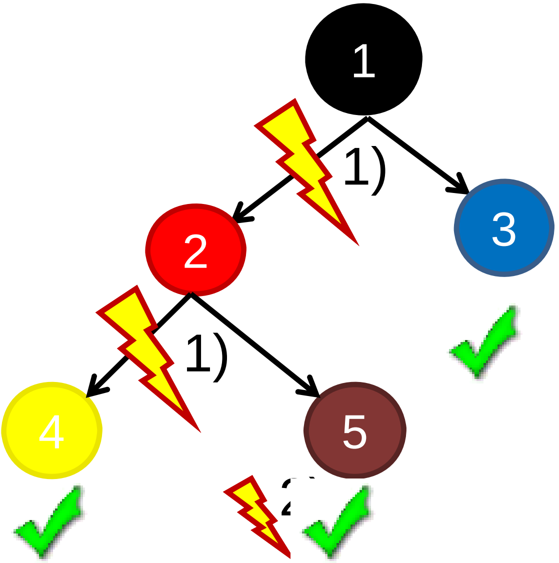
transformation system: $\tau \dot{v} = K(g-x) - Dv + (g-x_0)f$

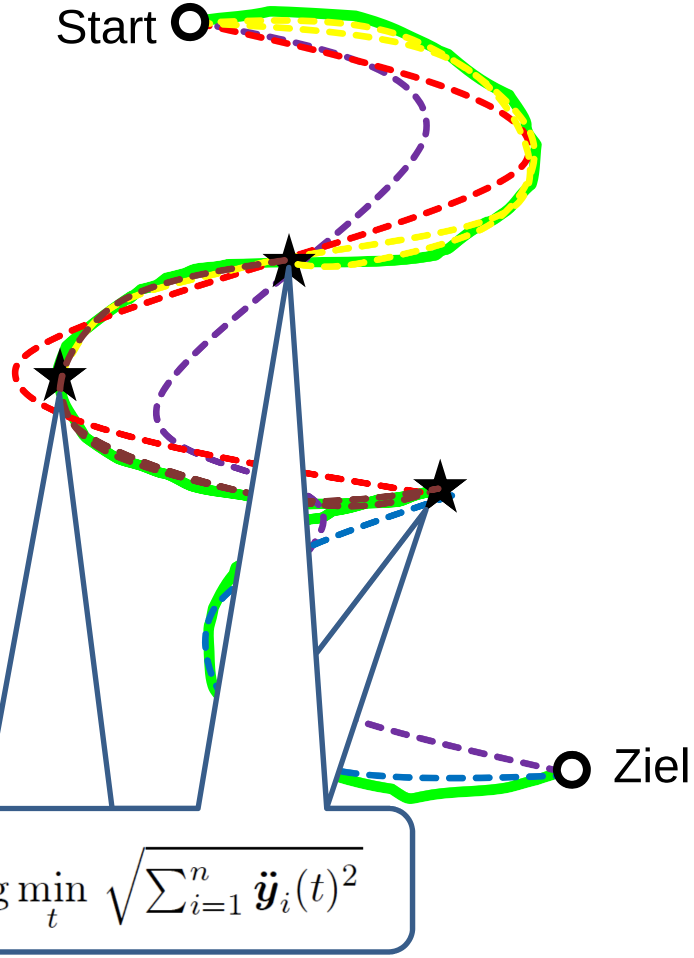
$\tau \dot{x} = v$



Trajectory Segmentation into DMPs

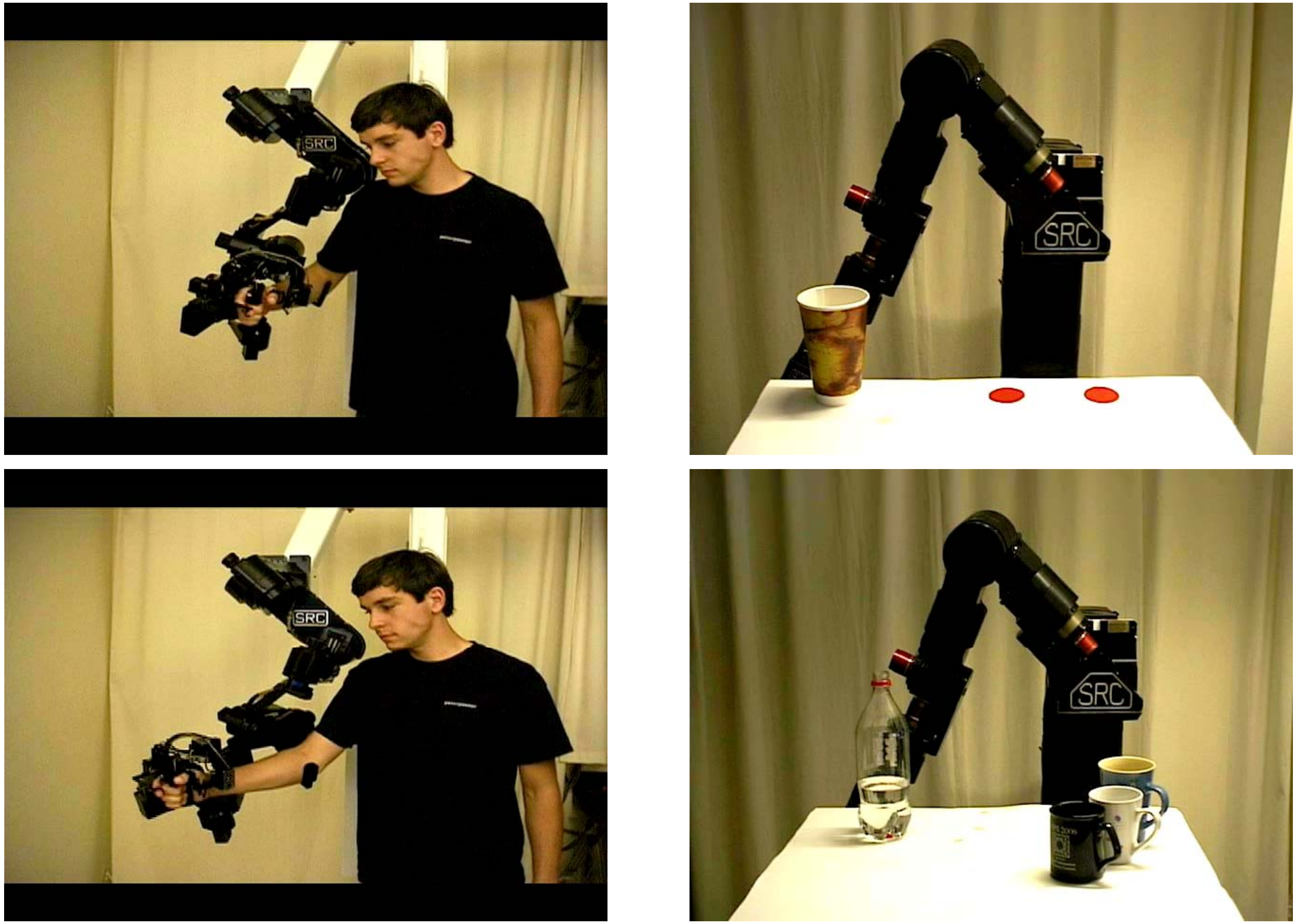
1) Approximation error
2) Minimal movement duration





$$\arg\min_{t} \sqrt{\sum_{i=1}^{n} \ddot{\boldsymbol{y}}_i(t)^2}$$

First implementation at USC



Joint work with Stefan Schaal and Peter Pastor

Conclusions

- Implementation of first components of an imitation system
- Various human motion capture system satisfying different needs according the task
- Transferring different human motion data to the robot using the Master Motor Map
- Reproduction of human-like motion and goal-directedness retained using non-linear optimization
- Action representation using Hidden Markov Models and Dynamic Movement Primitives

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