KIT-Department of Informatics
Prof. Dr.-Ing. Tamim Asfour

# Exam Solution Sheet

Robotics II: Humanoid Robotics

am September 20, 2019, 14:00 – 15:00

| Family name: | Given name: | Matriculation number: |
|---|---|---|
| David | Marr | 978-0262514620 |

| | |
|---|---|
| Exercise 1 | 7 out of 7 points |
| Exercise 2 | 6 out of 6 points |
| Exercise 3 | 12 out of 12 points |
| Exercise 4 | 10 out of 10 points |
| Exercise 5 | 10 out of 10 points |

| | |
|---|---|
| **Total:** | 45 out of 45 points |

| | |
|---|---|
| | **Grade:** **1.0** |

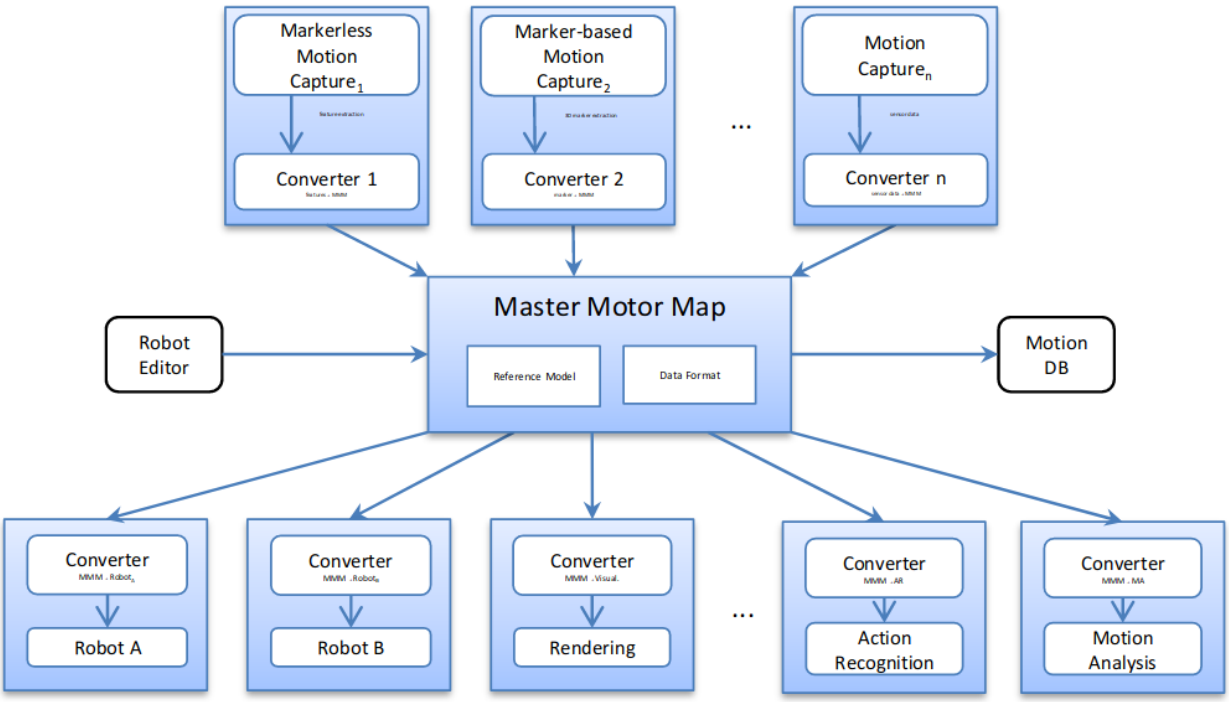
# Exercise 1 *Building Humanoids*

7 p.

1. Motivation behind the MMM framework:

   - The MMM is a unifying framework for capturing, representation, visualization and whole body human motion and mapping/converting to different embodiments
   - It includes interfaces and data structures for the transfer of motor knowledge between different embodiments
   - It allows the replacement of any module (perception, recognition, visualization, reproduction) can be guaranteed by using the MMM as the exchange format
   - All perceptive modules convert their output to the MMM format
   - All recognition and reproduction modules convert the MMM format to their specific internal representation

2. Diagram of the MMM framework:



3. Description of the three main models in the human body reference in the MMM:

   (a) Kinematic Model: joints and segment lengths

   (b) Dynamic Model: segment mass, center of mass and moments of inertia

   (c) Statistic/anthropomorphic model: Segment properties (e.g. length, mass etc.) defined as a function (regression) of global parameters (e.g. body height, weight)

# Exercise 2 *Grasping Synergies and Eigengrasps*

1. fully actuated hand vs. underactuated hand:

   A fully actuated hand has one actuator for each joint, an underactuated hand has less actuators than joints.

   - Fully actuated: Advantage: Each joint can be controlled separately. Disadvantage: Heavy.
   - Underactuated: Advantage: Lightweight, mechanical intelligence. Disadvantage: cannot generate all possible grasps.

2. Grasp study and hand design.

   (a) Minimum number of motors?

   Two motors should be chosen to cover 80% of the variance of the joint angles.

   (b) Method for linear dimensionality reduction:

   Primary Component Analysis (PCA).

   (c) Mechanism from the lecture that realizes mechanical underactuation:

   The KIT Hand V2 realizes mechanical underactuation. The Hand uses a tendon pulley mechanism that distributes the length change of the tendon to the four fingers of the hand. If one finger is blocked the other fingers continue closing, until all fingers are blocked by the object.

   (d) Is dimensionality reduction useful for fully actuated hands?

   Yes, dimensionality reduction is useful even in the case of a fully actuated hand.

   - Using the reduced dimension the control of the hand is easier, since only a two dimensional control is required, instead of a 23 dimensional control.
   - The reduced dimensions can also be used for grasp synthesis, as the problem can be reduced from a 23D problem to a 2D problem.

# Exercise 3 *Grasping*

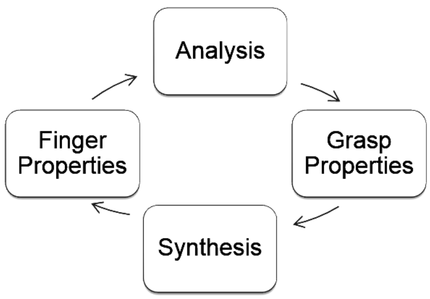
1. Grasp analysis and synthesis

   (a) Properties used for grasp analysis and synthesis:

   - **Dexterity**: How should grasping fingers be configured?
   - **Equilibrium**: How hard to squeeze the grasped object?
   - **Stability**: How to remain unaffected by external disturbances?
   - **Dynamic behavior**: How soft a grasp should be for a given task?

   (b) Description

   - **Analysis**: study of grasp properties for a given set of finger properties.
   - **Synthesis**: determination of the required finger properties in order for the grasp to acquire some desired properties.

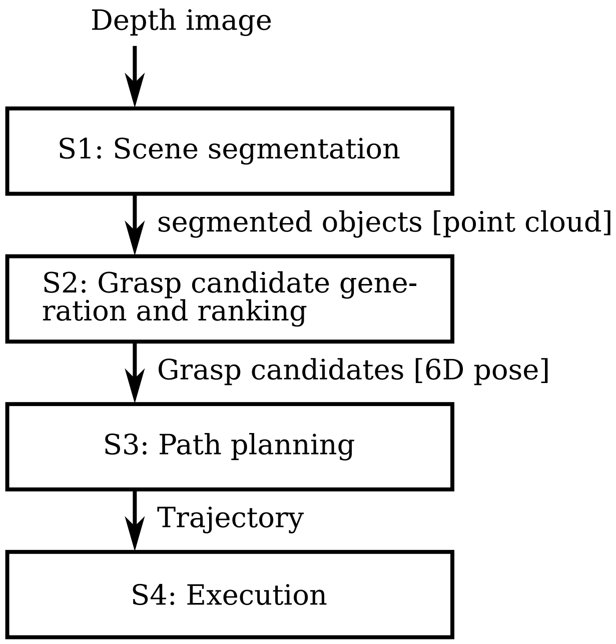



2. Factors relevant for grasp hypothesis generation (1 example per factor required)

   - **Task**: Pick and place, Handover, ...
   - **Hand**: Gripper, Multi-Fingered
   - **Object-Grasp Representation** (Reference Frame): Local, Global
   - **Prior Object Knowledge**: Known, Familiar, Unknown
   - **Grasp Synthesis**: Analytical, Data-Driven
   - **Object Features**: 2D, 3D, Multi-Modal

3. The following application scenario is given: You want to use a humanoid robot to grasp unknown objects on a table and place them in a box.

   (a) Pipeline (Initial Input: PointCloud or RGB-D Image):




   (b) Possible sources of training data:

   - Learning by demonstration: A human teacher demonstrates the task while the robot observes.
   - Training data collection on the target system: The robot tries different policies and updates the policies according to the results.
   - Training data generation in simulation: The robot and scene is simulated in a virtual environment. Actions and respective outcomes are stored as training examples.
   - Hand-labeled data: Create a dataset by hand labeling a set of inputs with the respective outputs.

   (c) Which steps have to be changed in your pipeline?

   Step 1 and step 2 have to be replaced. New Steps:

   - Object detection: Find the known objects in the scene
   - Pose estimation: Estimate the 6D pose for all known objects
   - Grasp selection: Select one or multiple grasps from a grasp database.

# Exercise 4 *Active Perception*

1. The five questions

   Question 1: Why? Description: Determine the next action based on the expectation that its state generates

   Question 2: What? Description: Select the subset of the world that can be sensed.

   Question 3: Where? Description: The sensory elements can only be sensed from a particular viewpoint.

   Question 4: When? Description: When is the action valid and with what duration?

   Question 5: How? Description: What actions must precede the execution of a sensing or perceiving action?

2. Differences:

| | Classical CV | Active Vision | Active Perception |
|---|---|---|---|
| Image Processing | x | x | x |
| Viewpoint Selection | | | |
| Multi-modal sensory input | | | |
| Changing the agent's state | | | |
| Changing the environment | | | |

3. General purpose of the *Iterative Closest Point (ICP)* algorithm:
   Determination of the transformation between two overlapping point clouds. The Transformation between both point clouds should be estimated.

4. Two problems that can occur using ICP:

   - Small objects can vanish in large, complex scenes with many features
   - If object is only partially covered, false points can be included
   - 3D shapes ambiguous, e.g. many planes in most scenes
   - Only finds a local optimum

5. Picking task

   (a) Three heuristics:

   - Planes, cylinders and spheres amongst SIFT features (RANSAC)
   - Unicolored regions of promising size using color MSERs (Maximally stable extremal regions)
   - Visually salient regions (DoG filter)

   (b) Possible approach:

   - Step 1: Push in the region of interest
   - Step 2: Re-localize object hypothesis
   - Step 3: Estimate Motion
   - Step 4: Accept if moved or discard if no movement

# Exercise 5 *Imitation Learning*

1. Four key questions and explanation :

   (a) Who (to imitate): Choosing a demonstrator whose behavior can benefit the robot (teacher selection).

   (b) When (to imitate): The robot has to identify the beginning and end of a demonstrated behavior as well as to decide whether the observation is appropriate in the current context.

   (c) What (to imitate): The robot has to identify, what aspects of the demonstration are of interest or essential for the behavior and which can safely be ignored.

   (d) How (to imitate): The robot has to identify, how the observed behavior can be replicated to its own embodiment, with potentially different kinematic chains, missing limbs (e. g. wheeled robots don't have feet), ...

2. Hierarchical Segmentation:

   Levels of the segmentation: 1) Semantic segmentation and 2) segmentation based on motion profile

   Principle of first layer: Segmentation based on change of object/hand contact relations.

   Principle of second layer: Segmentation based on the change of motion characteristics.

3. Mirror neurons:

   Mirror neurons are nerve cells that have been identified in the brain of humans (and other primates). They are equally active both during observation and during the execution of a particular activity. Mirror neurons thus connect the perception of an action with its execution and are active during the entire observation/execution.

4. Four cognition challenges in Imitation Learning: