MVP

Machine Vision and Perception Group @ TUM

Machine Vision and Perception Group
Department of Informatics

Darius Burschka

Technische Universität München

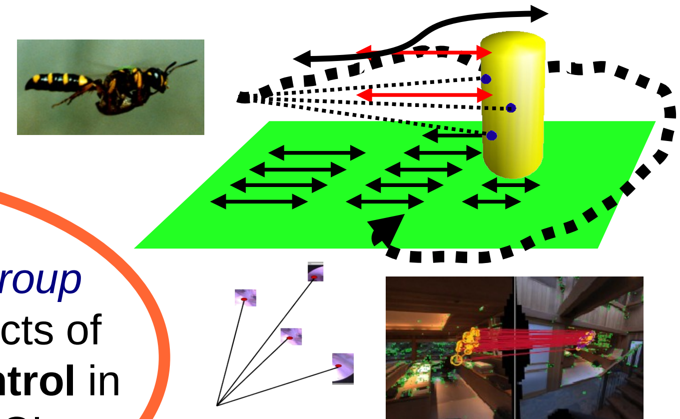
Research of the MVP Group http://www6.in.tum.de/burschka/mvp

MVP Machine Vision and Perception Group @ TUM

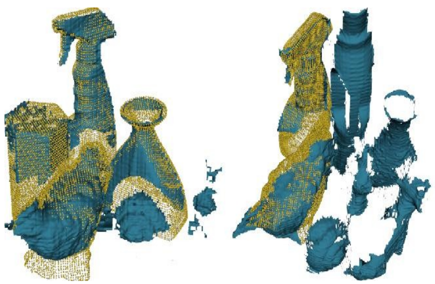
Perception for manipulation

Visual navigation

Biologically motivated perception

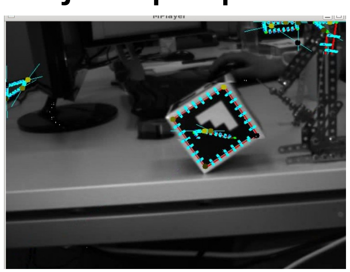


Rigid and Deformable Registration



The *Machine Vision and Perception Group* @TUM works on the aspects of **visual perception** and **control** in medical, mobile, and HCI applications

Exploration of physical object properties

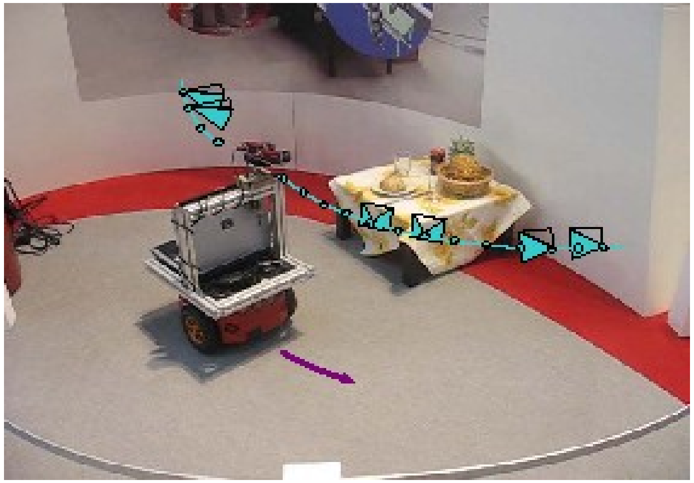


Photogrammetric monocular reconstruction

Visual Action Analysis

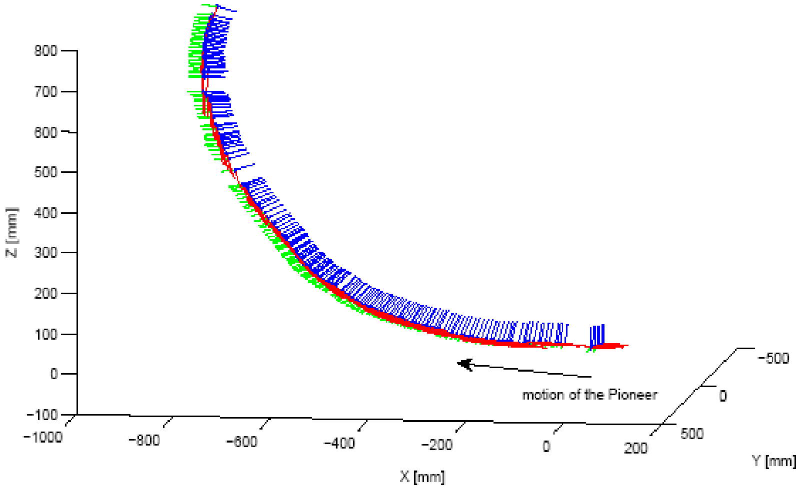
Visual Localization for hybrid Environmental Modeling

Real-Time Localization

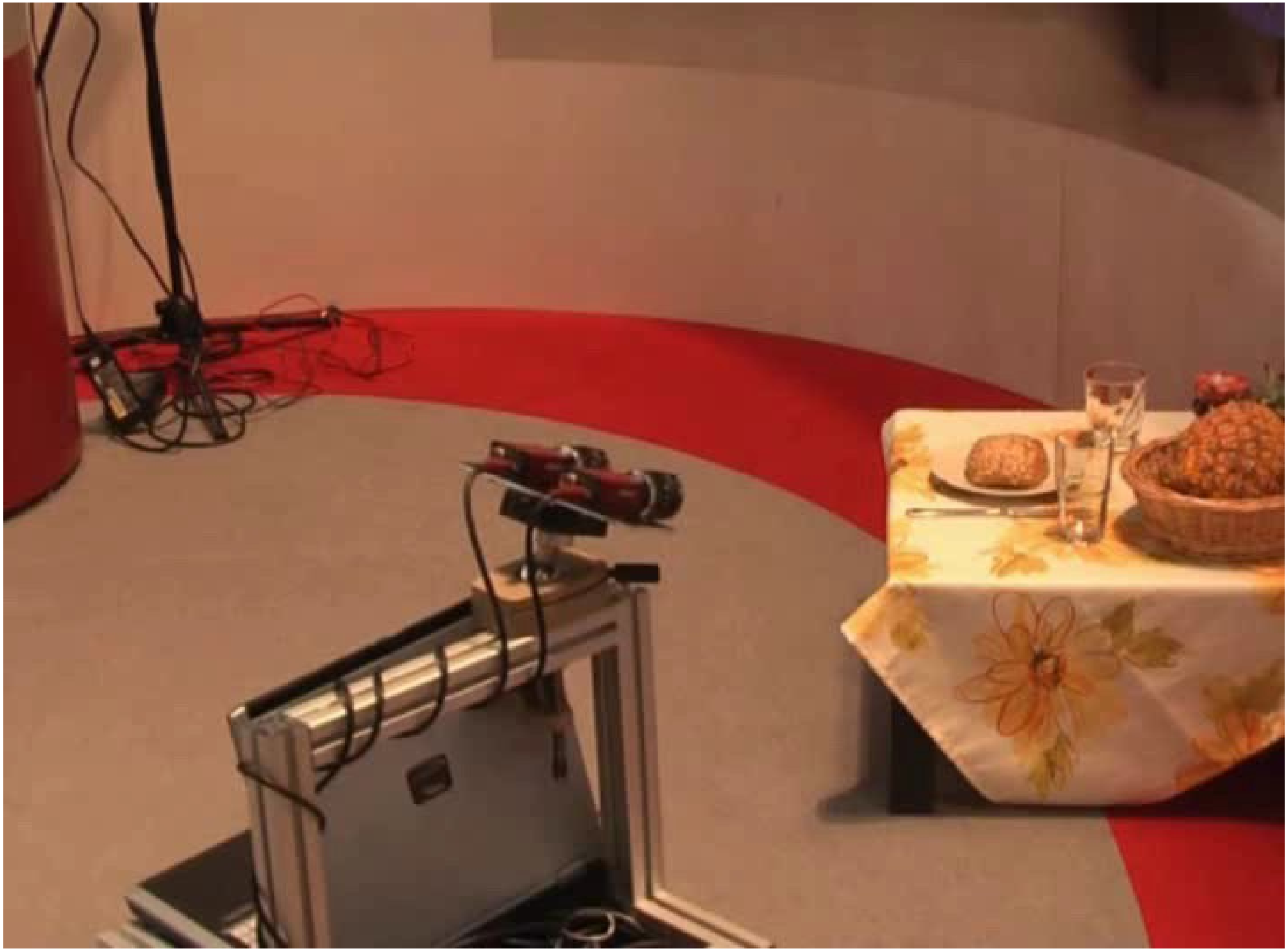


The high accuracy allows direct stiching of images along the trajectory without bundle adjustment

Resulting hybrid (appearance and geometry) model allows path planning and prediction of sensor views for e.g. attention research

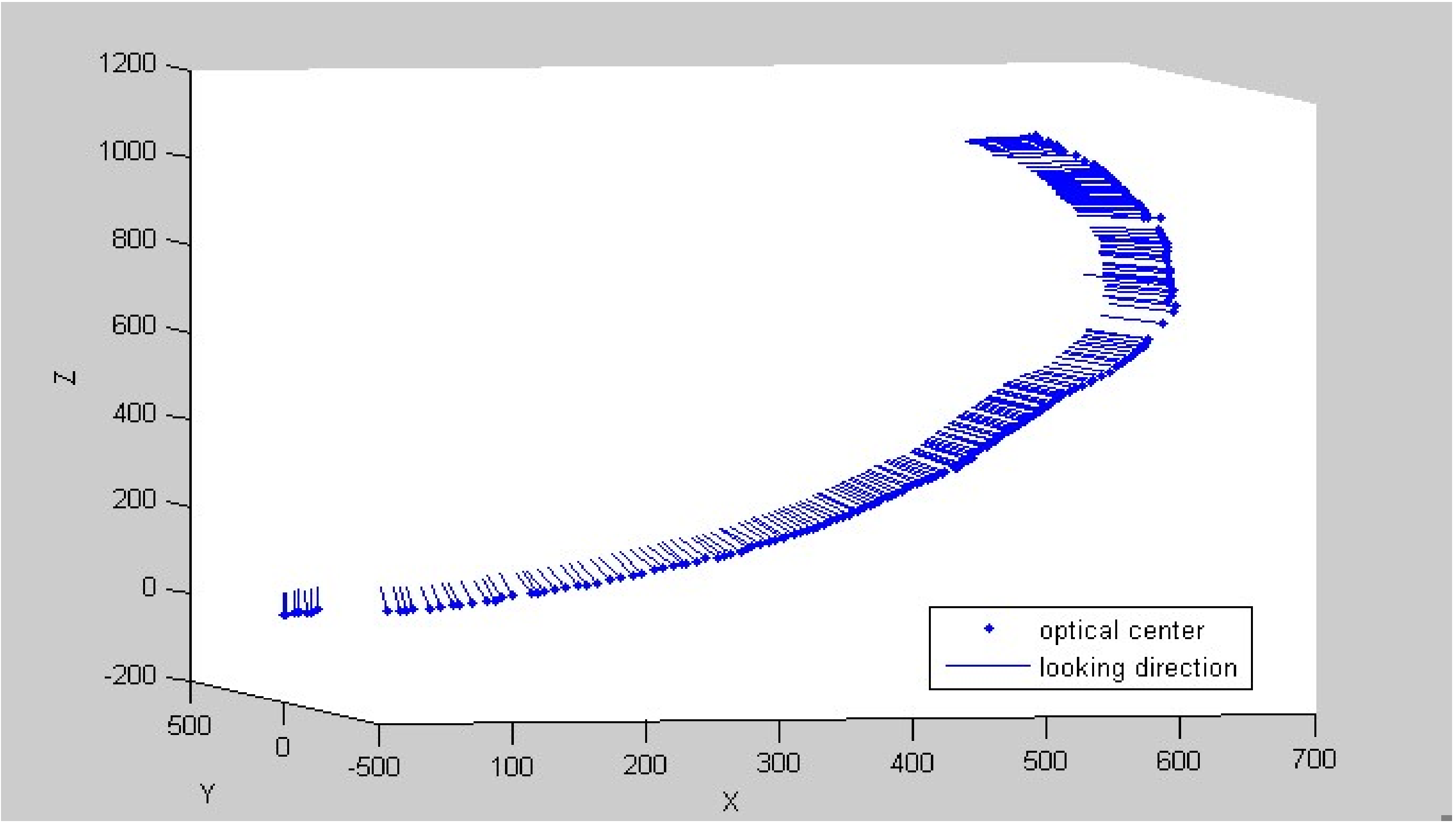


Example of a hybrid model reconstruction

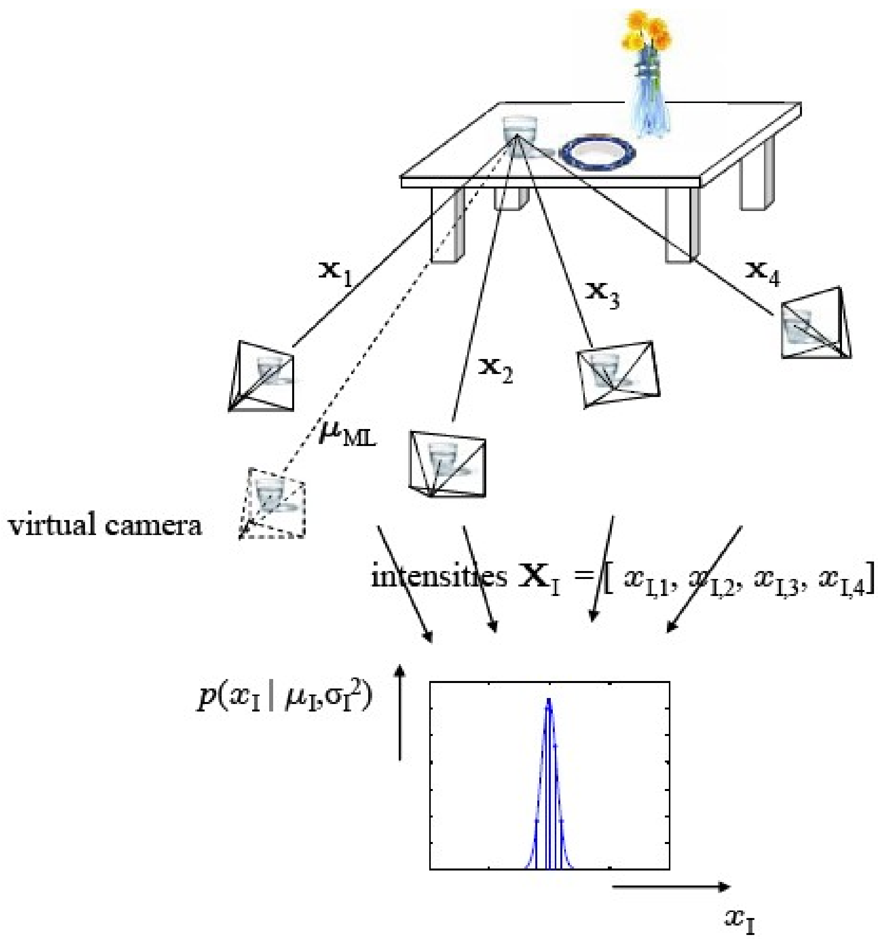


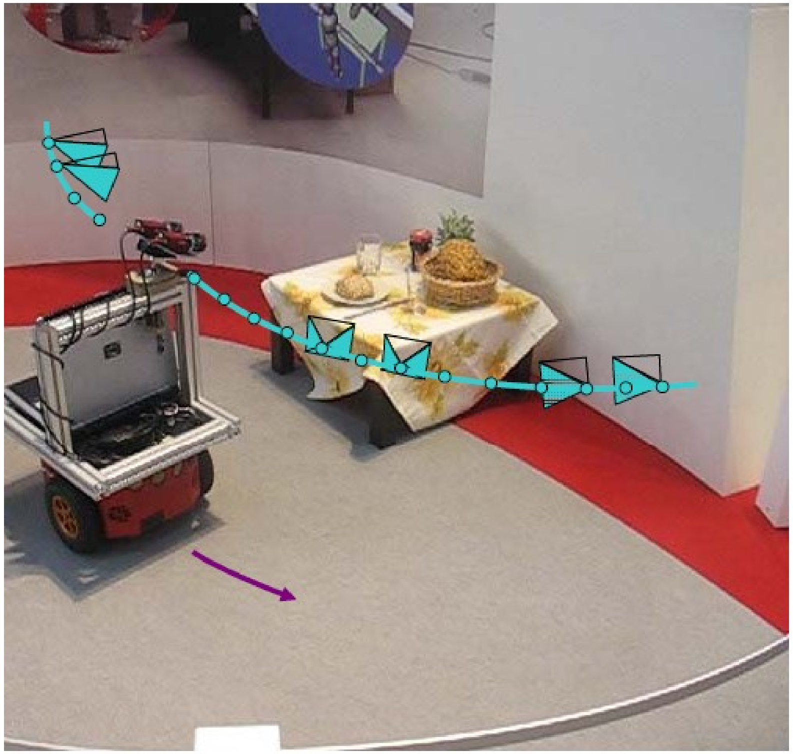
Result of a visual localization

6DoF pose of the robot is calculated fom single camera view

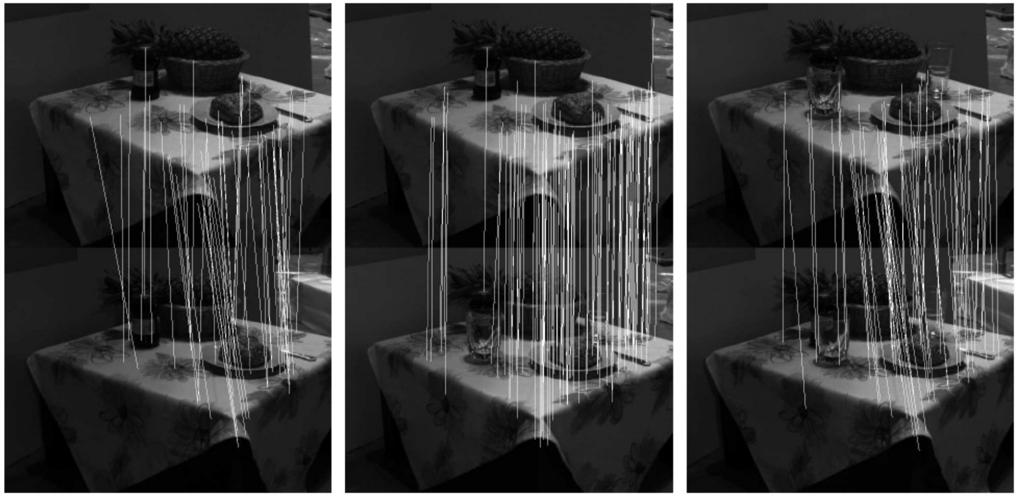


Reconstruction of a hybrid (appearance/geometry) model



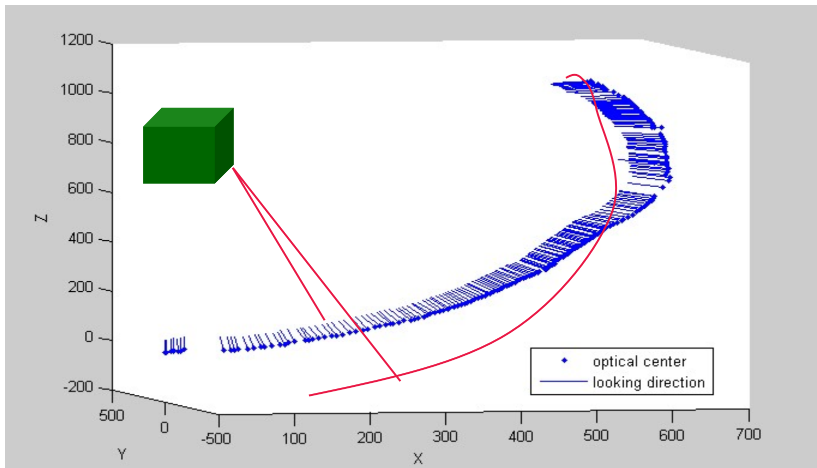


Visual Homing

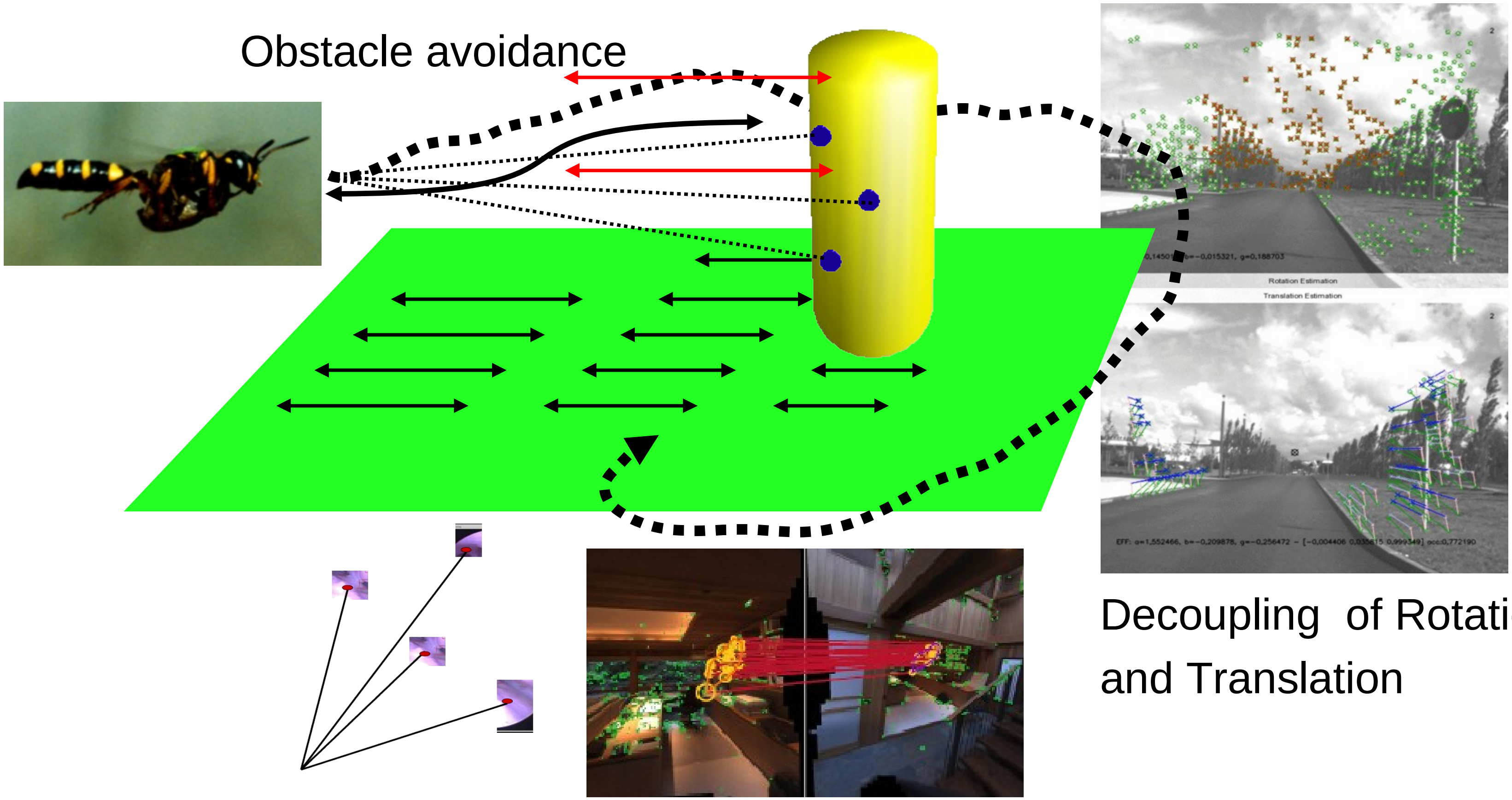


(a) $I_{1.1}$ and $I_{1.2}$ (b) $I_{1.1}$ and $I_{2.1}$ (c) $I_{2.1}$ and $I_{2.2}$

The system is capable of registration to previously observed trajectories



Biologically Motivated Navigation



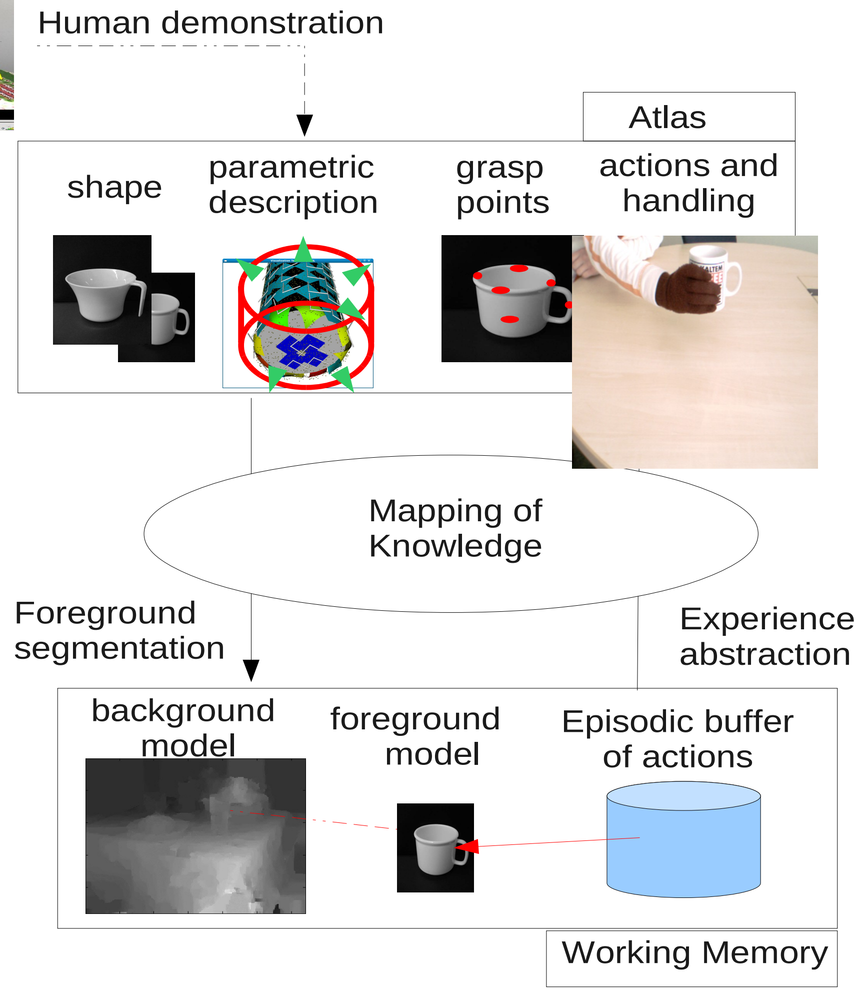
Decoupling of Rotation and Translation

Perception-based loop closure

Work on Optimal Sensor Models

Knowledge Representation

- Atlas:
  - Long-term memory
  - Experience of the system
- Working memory:
  - Short-term memory
  - Experience grounded in a given environment
- Temporal handling information



Indexing of the Atlas information from 3D perception

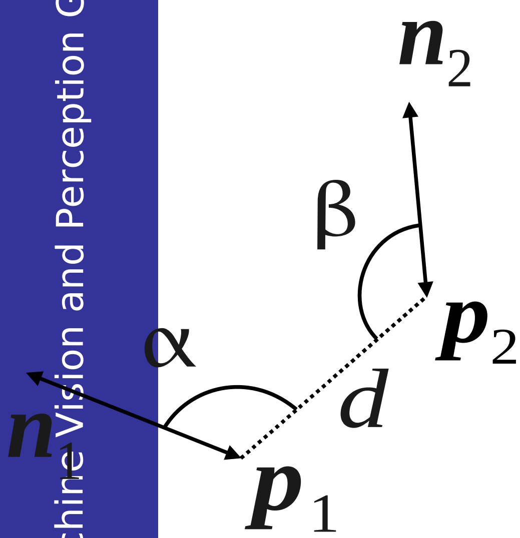
Real-world scenario

scene setup

input point cloud

recognized models

Algorithm Description
(Model Preprocessing Phase)



- For all pairs of surflets at distance *d* insert the triple

$$\left(\alpha\,,\ \beta\,,\ \measuredangle\left(\boldsymbol{n}_1\,,\boldsymbol{n}_2\right)\right)$$

  plus a pointer to its model in a hash-table.
- Do this for all models using the same hash-table.

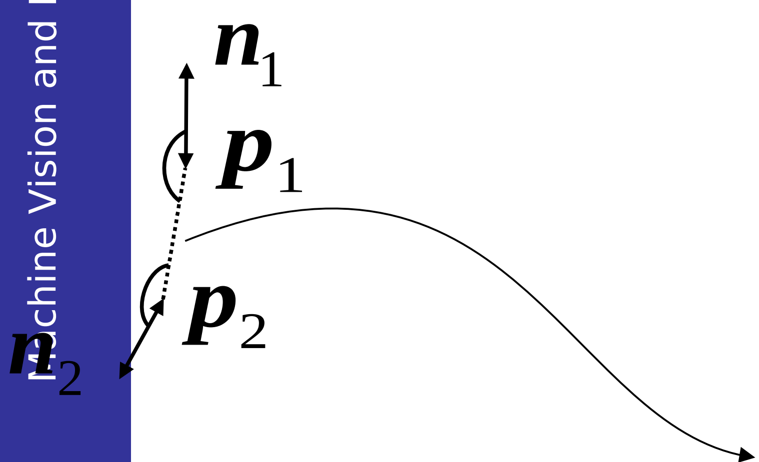
3D Object Recognition

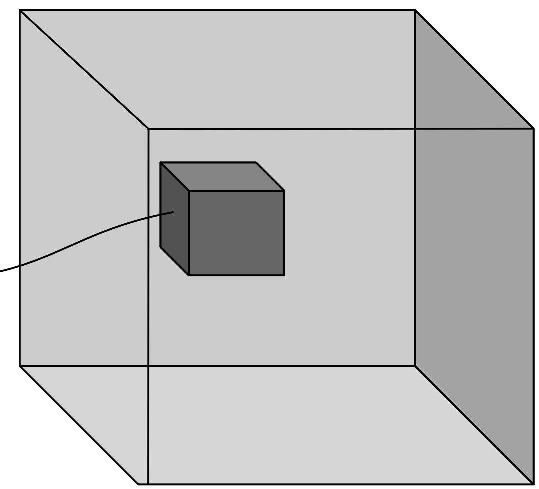
- Challenges
  - Incomplete Data
  - Noise and outliers

Online Recognition Phase

- For each model surflet pair $((\tilde{\boldsymbol{p}}_1, \tilde{\boldsymbol{n}}_1), (\tilde{\boldsymbol{p}}_2, \tilde{\boldsymbol{n}}_2))$ in the hash-table cell:

  Compute the rigid transform *T* that best aligns $((\tilde{\boldsymbol{p}}_1, \tilde{\boldsymbol{n}}_1), (\tilde{\boldsymbol{p}}_2, \tilde{\boldsymbol{n}}_2))$ to $((\boldsymbol{p}_1, \boldsymbol{n}_1), (\boldsymbol{p}_2, \boldsymbol{n}_2))$





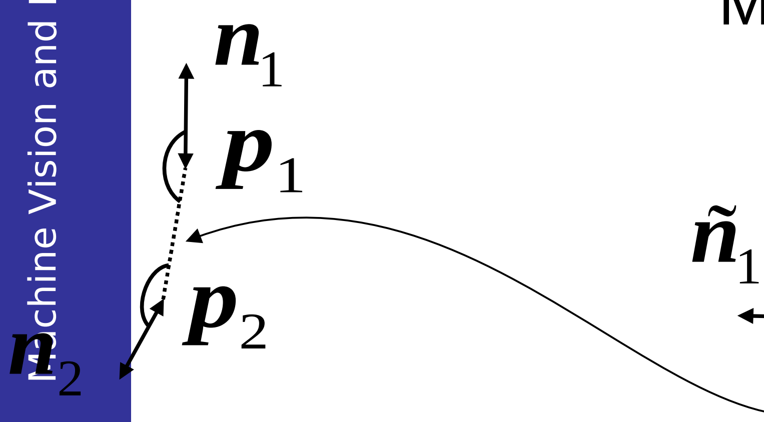
model hash-table

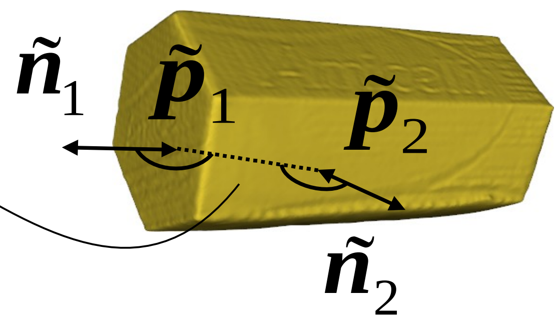
Online Recognition Phase

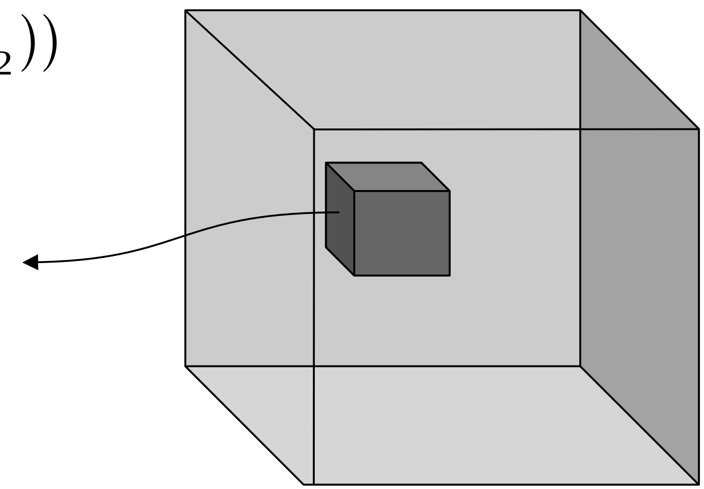
- Check if T(M) matches the scene.

M is the model of

$$((\tilde{\boldsymbol{p}}_1, \tilde{\boldsymbol{n}}_1), (\tilde{\boldsymbol{p}}_2, \tilde{\boldsymbol{n}}_2))$$



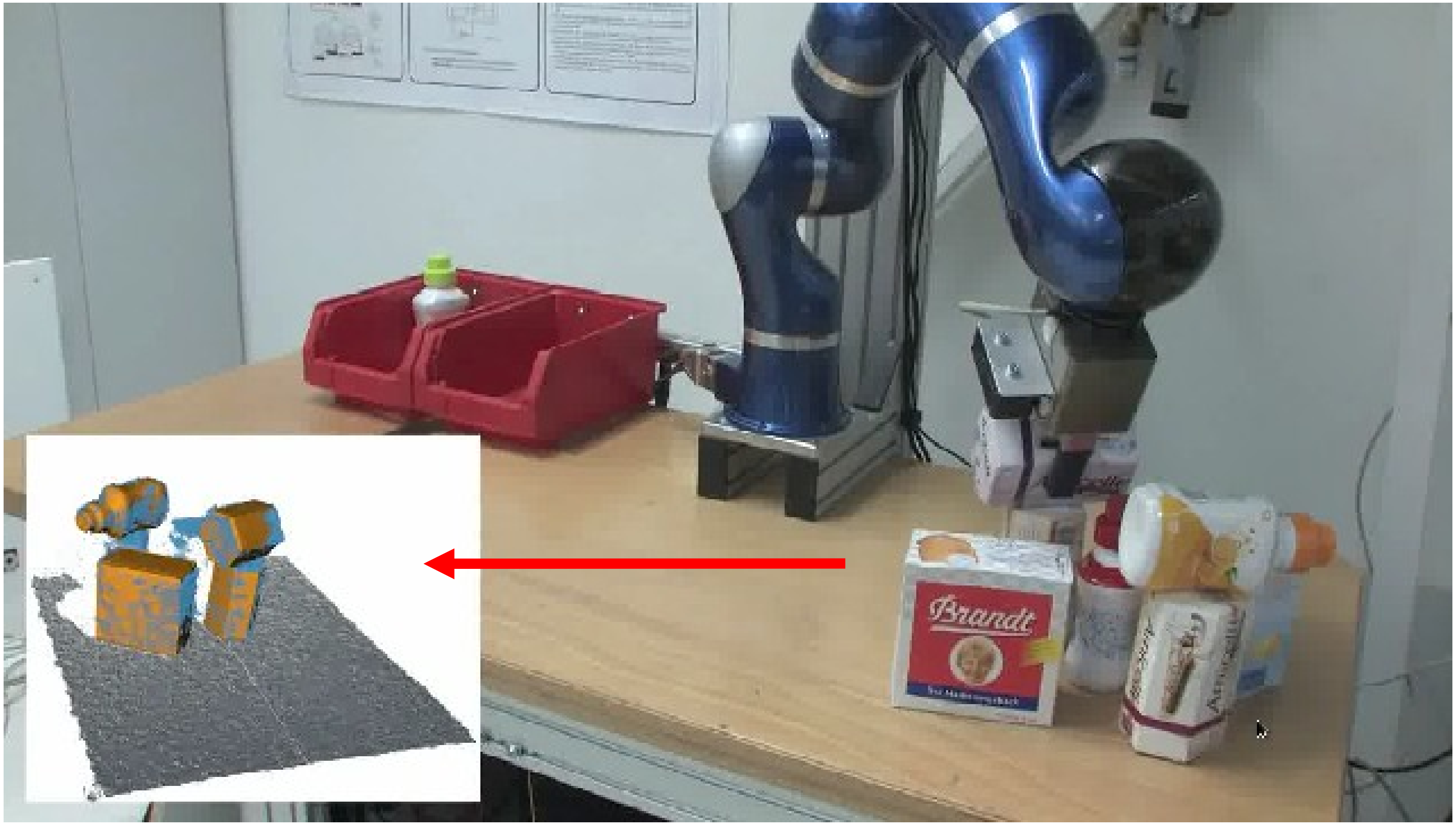




model hash-table

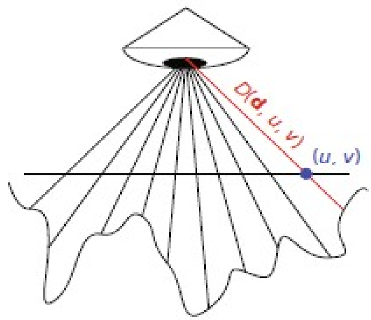
- If yes, accept T(M).

Example of the System in Action

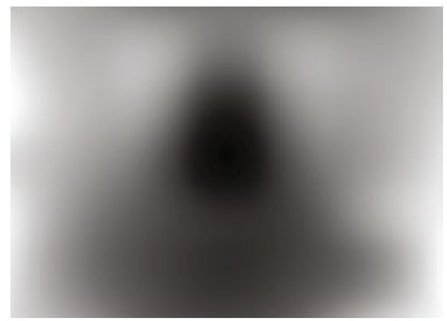


How to reconstruct 3D under poor texture condition?

Problem: texture information is more sparse



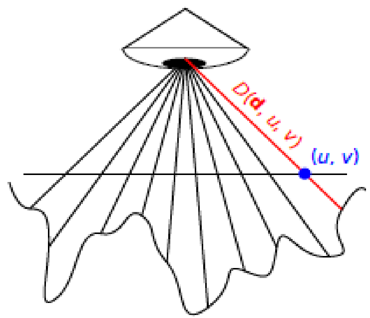
(a) Schematic 2D view of a depth



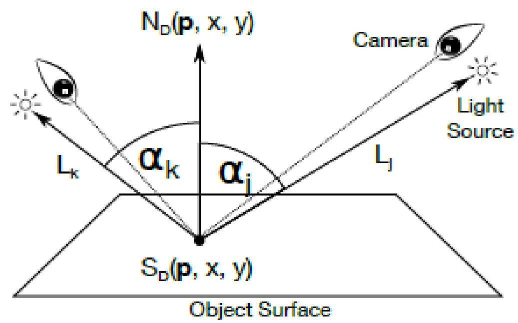
(b) Depth map for reconstructed face

What can we do if the texture information is almost non-existent?

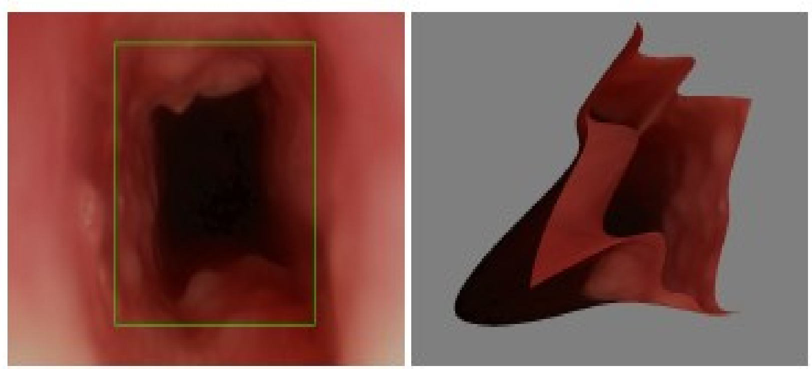
→photogrammetric approach



(a) Concept of a depth map.



(b) Surface illumination.



Intensity-Based Methods

- Established 3D reconstruction methods:
  - Rely on presence of many reliable features
  - May perform very poor if this prerequisite is not met
  - Medical images are typically problematic
- Solution approach: Intensity based methods
  - Do not exclusively rely on salient image regions
  - Shaded regions, e.g., also contain information
  - Use all information that is there

Intensity-Based Bundle Adjustment

- Regular Bundle Adjustment equation:

$$\min_{\boldsymbol{a},\boldsymbol{b}} \sum_{i=1}^{n} \sum_{j=1}^{m} d\left(\boldsymbol{Q}\left(\boldsymbol{a}_{\boldsymbol{j}},\boldsymbol{b}_{\boldsymbol{i}}\right), x_{i,j}\right)$$

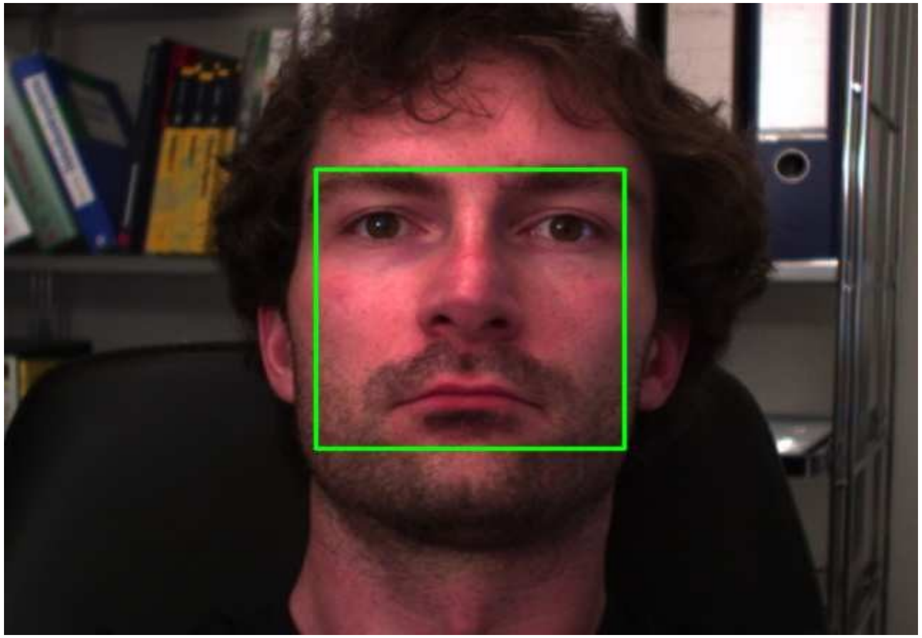
- Finds camera parameters *a* and point coordinates *b*
- Minimizes feature reprojection error

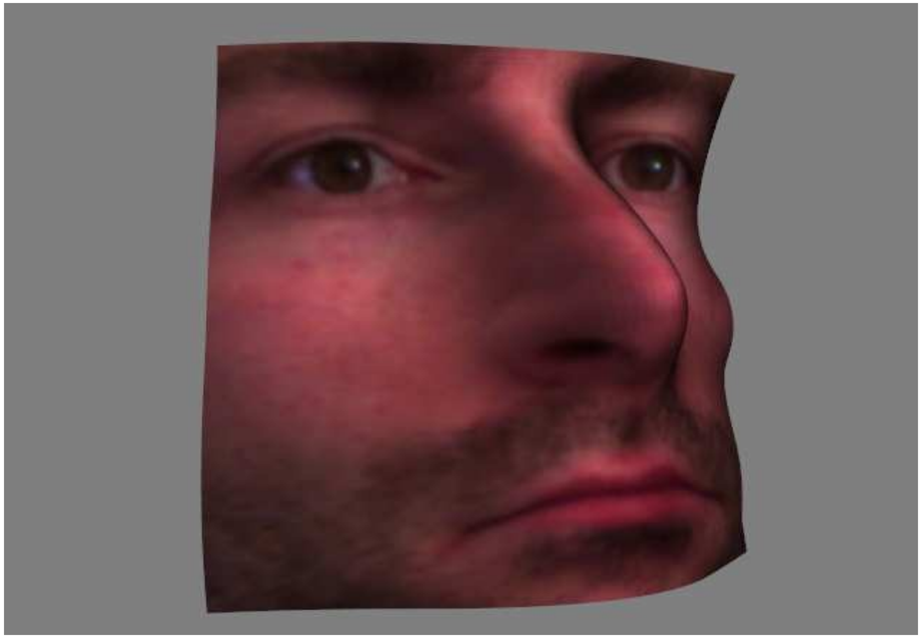
$$\min_{\boldsymbol{a},\boldsymbol{b}} \sum_{i=1}^{n} \sum_{j=1}^{m} d\left(I_j\left(\boldsymbol{Q}\left(\boldsymbol{a}_{\boldsymbol{j}},\boldsymbol{b}_{\boldsymbol{i}}\right)\right), I_0\left(\boldsymbol{p}_{\boldsymbol{i}}\right)\right)$$

- Intensity-Based Bundle Adjustment:
- Uses a regularizer: b now describes a smooth surface
- Apart from that: Same, but minimizes intensity error

Reconstruction Example

- Works well under static lighting conditions and roughly Lambertian surfaces

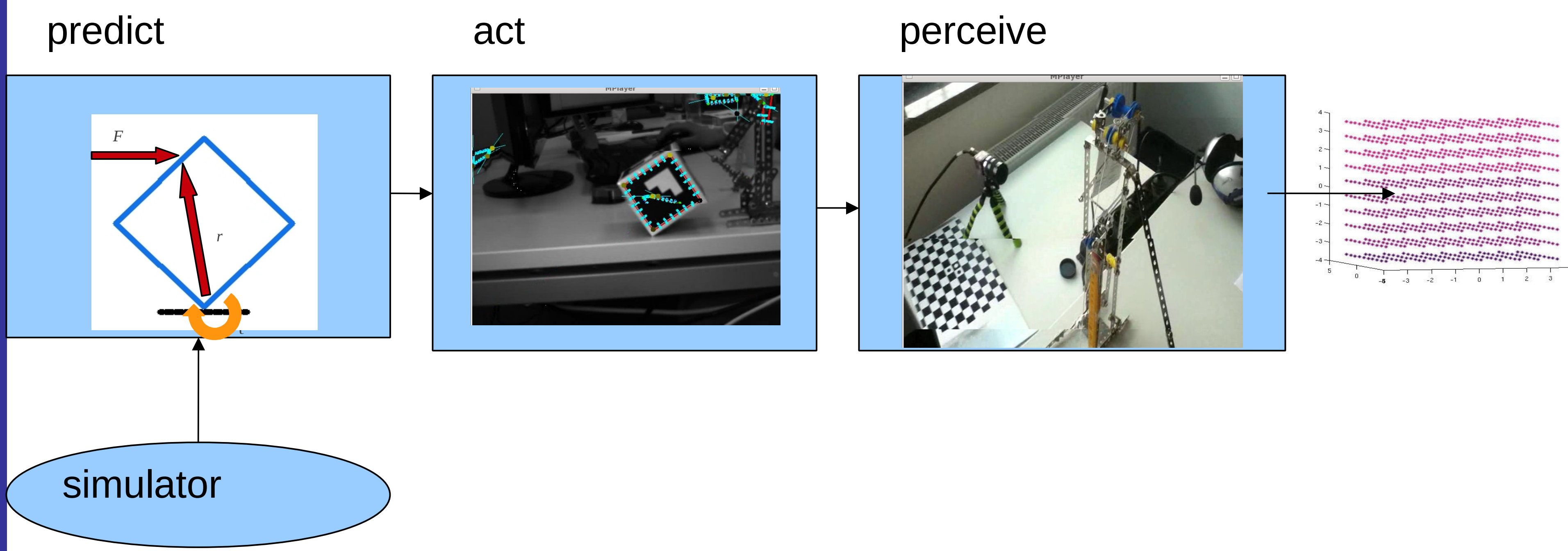




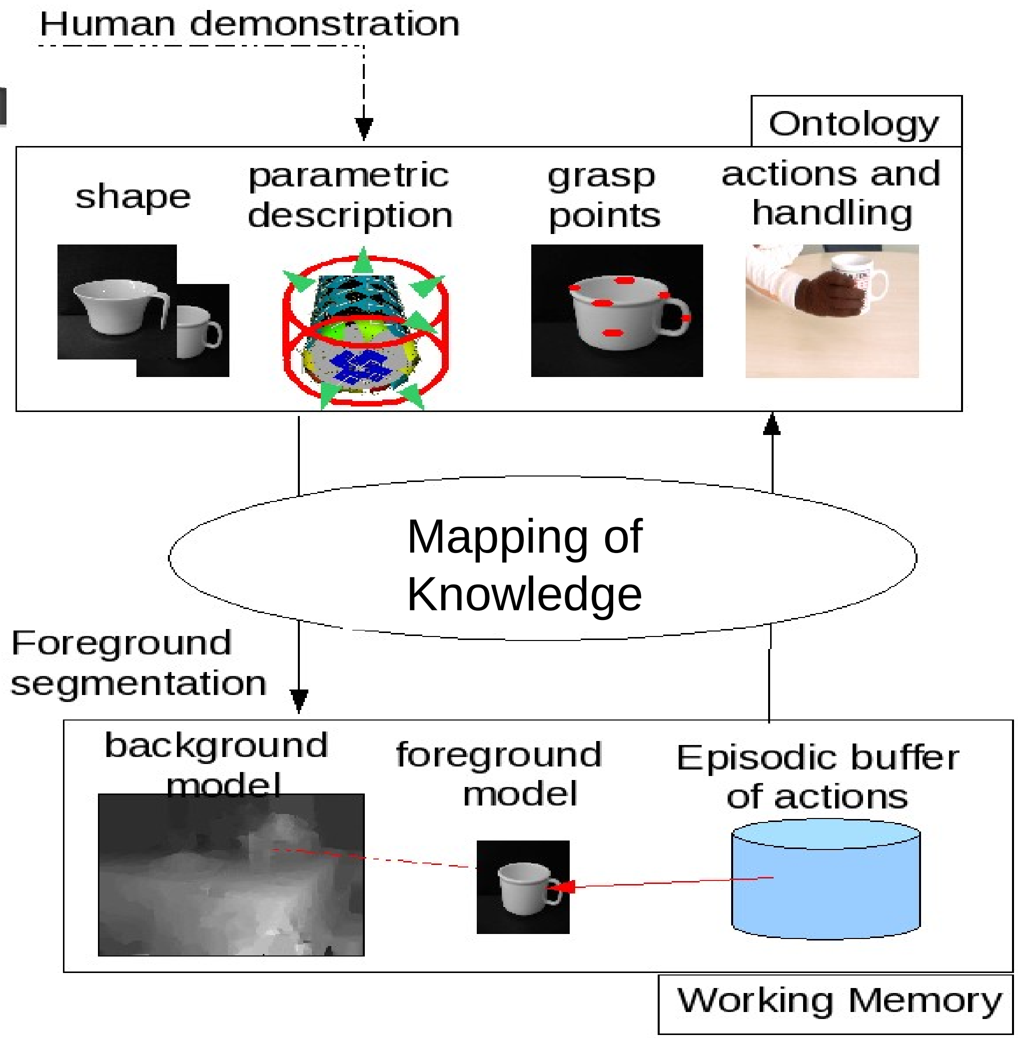
Ruepp and Burschka. Fast recovery of weakly textured surfaces from monocular image sequences. ACCV 2010

Estimation of Physical Properties in Predict-Act-Perceive loop

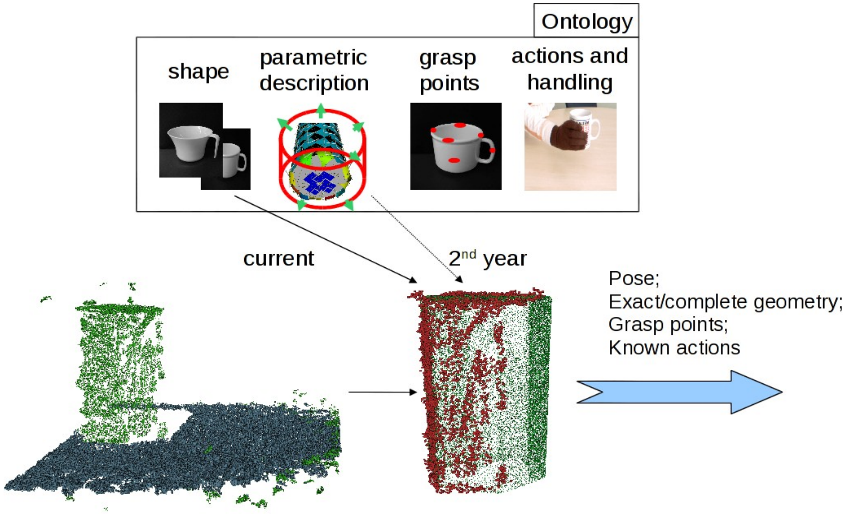
- Estimation of the Center of mass
- Estimation of Mass and Friction Force
- Estimation of Mass Distribution



Knowledge Representation in WP5



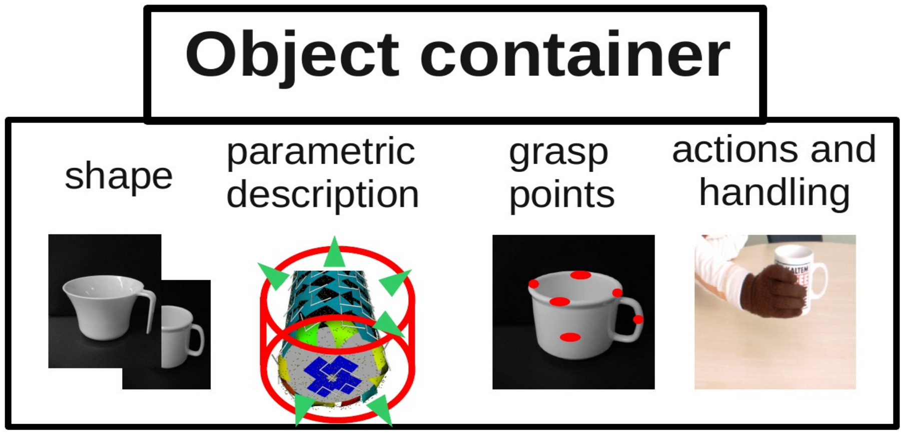
Defining Foreground: Shape Matching



Indexing to the Atlas database needs to be extended to object classes
-> deformable shape registration needed (ACCV2010 - oral)

Object Container

- Orientation
- Max. allowed acceleration
- Mass
- Center of gravity

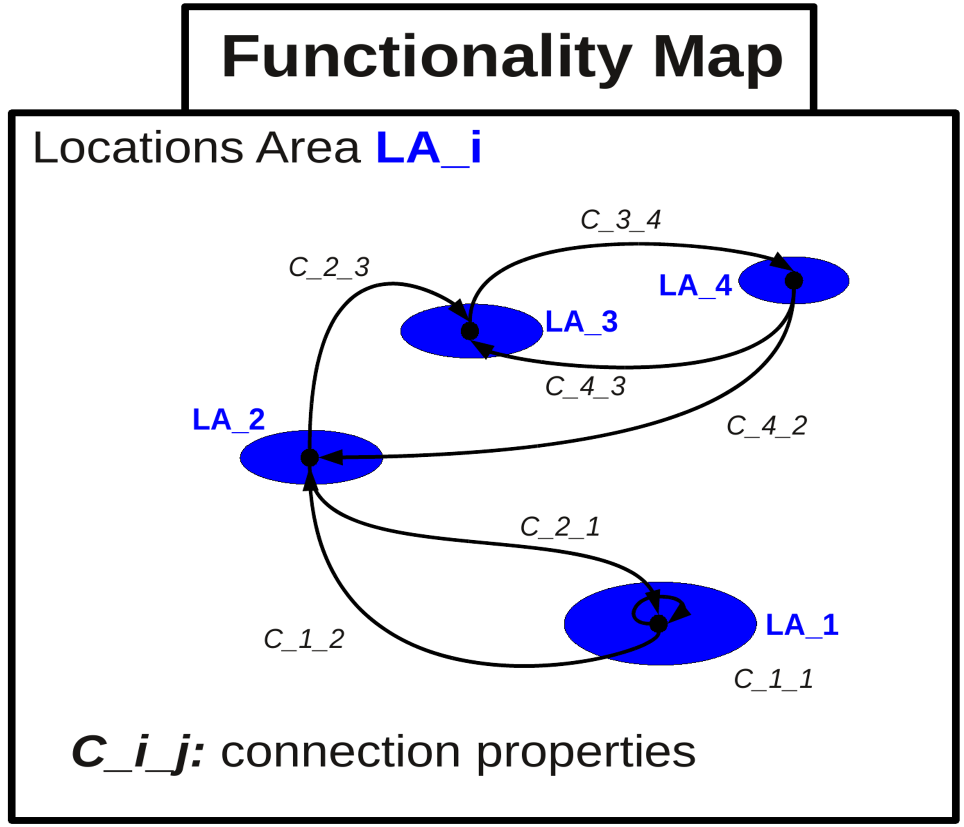


Functionality Map

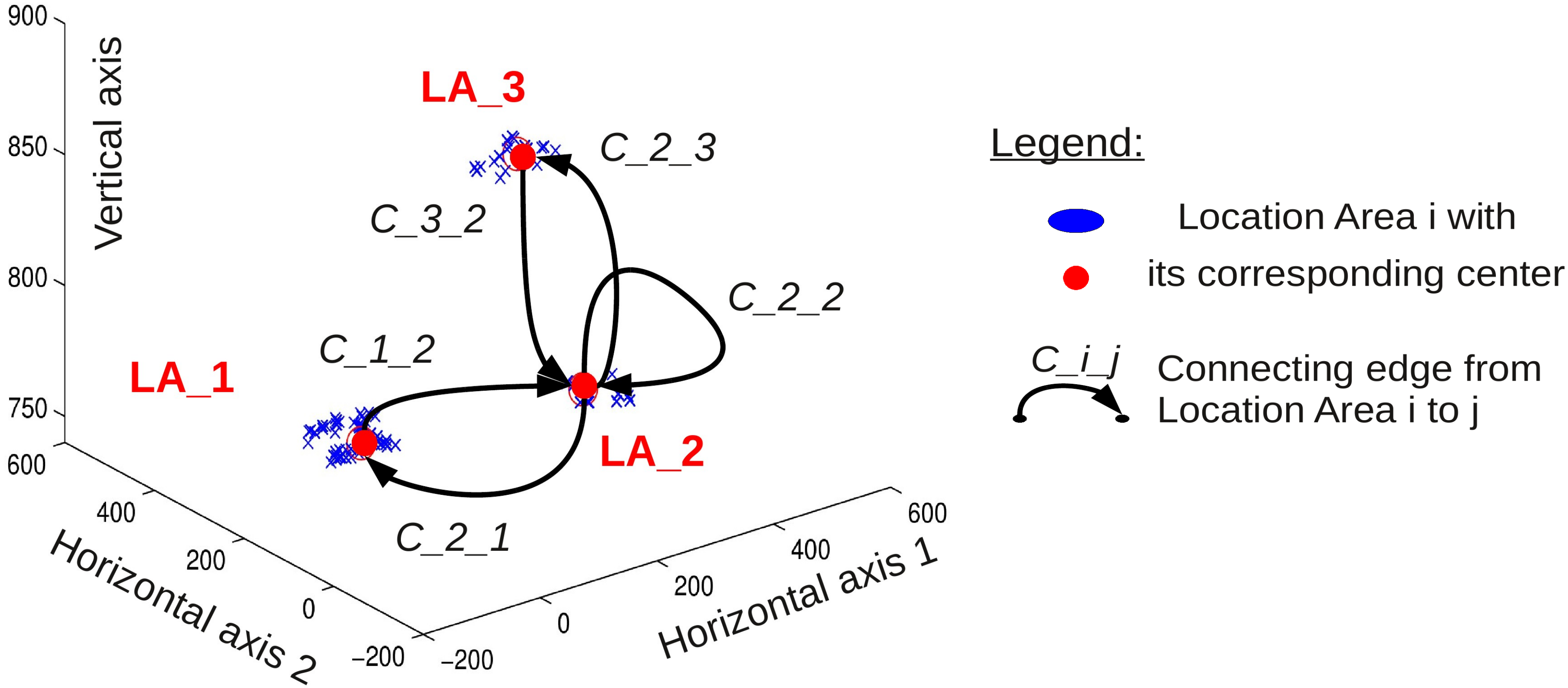
Location Areas

Connection Properties

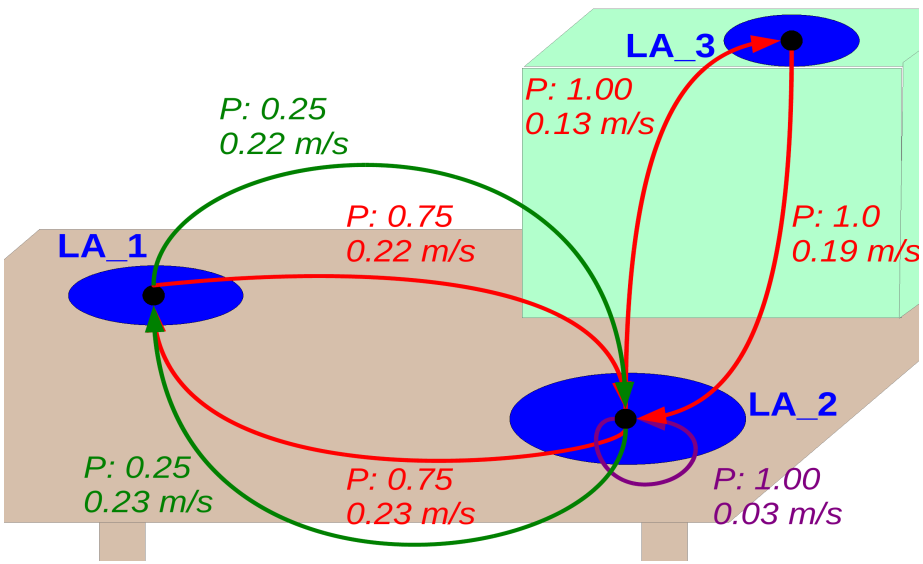
- Pushed vs. lifted objec
- Arbitrary movement vs movement with a goal
- Connection relevance
- Velocity constraint during pick-up
- Grasp taxonomy
- Approach vector



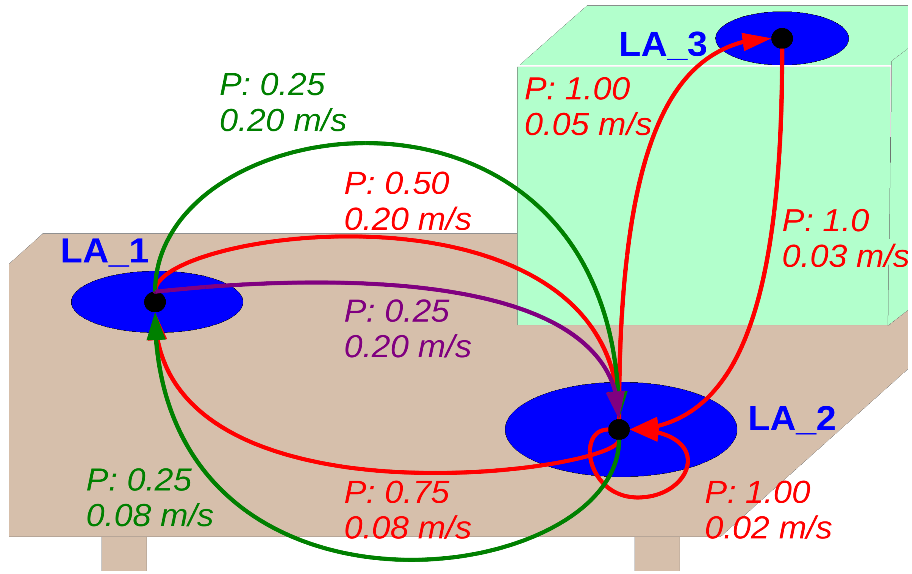
Basic Experiments: Location Areas (Tracking Data)



Basic Experiments: Functionality Maps (Tracking Data)



Milk carton



Cup - Handle

Red = goal, green = push, magenta = arbitrary

Entire System: Exemplary Trajectories