Darius Burschka – Perception Challenges

Semantic Perception for Seamless Integration of Technical Systems into Everyday Life

Darius Burschka

Machine Vision and Perception Group
Department of Computer Science

Technische Universität München

Where do we need perception?

"known" clutter

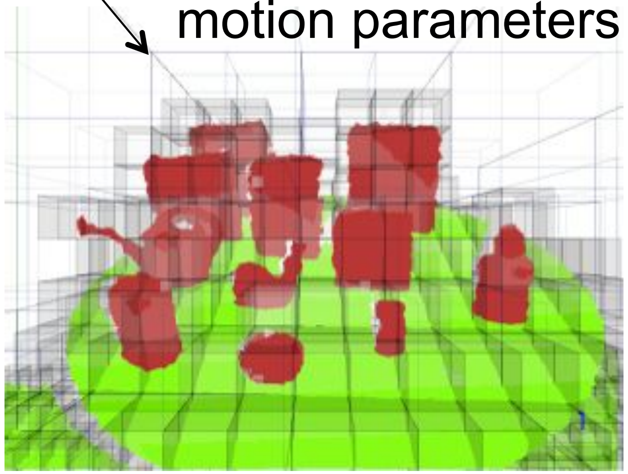
"unknown" clutter

Role of perception in an unknown environment

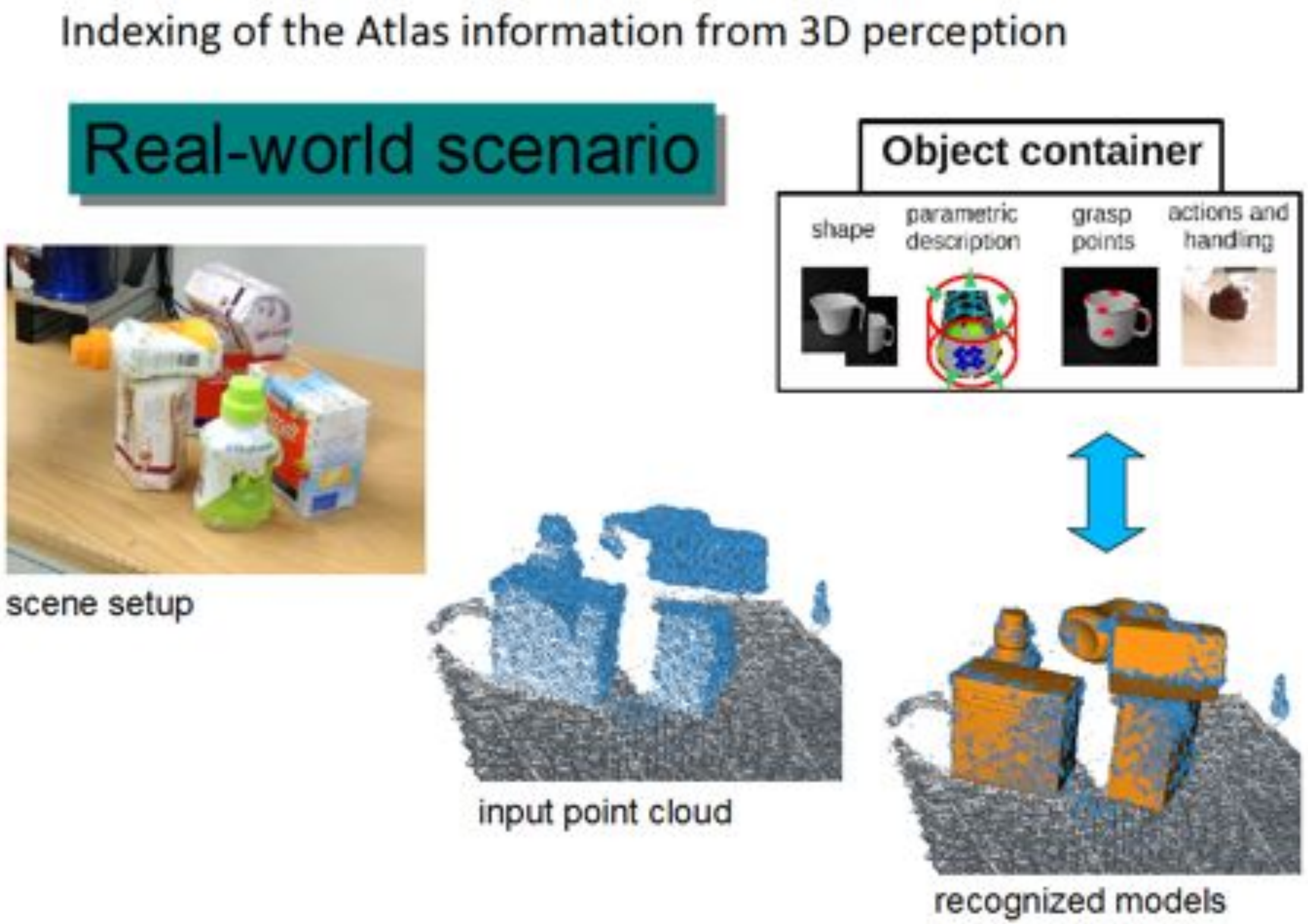
- Acquisition of the scene
- Clustering of raw sensor data
- Identification of independent motion components
- Estimation of action
- Understanding actions

labeling

motion parameters

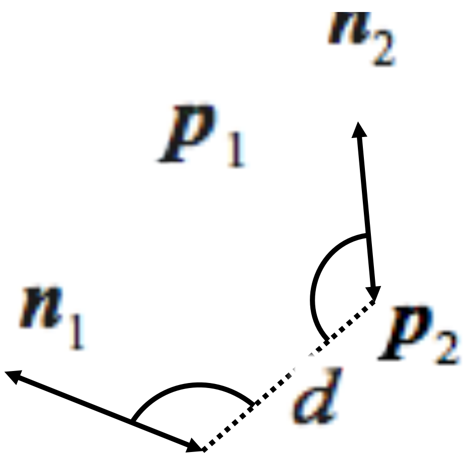


What is in the scene?
(sensor data abstraction)



Feature Extraction

input point cloud

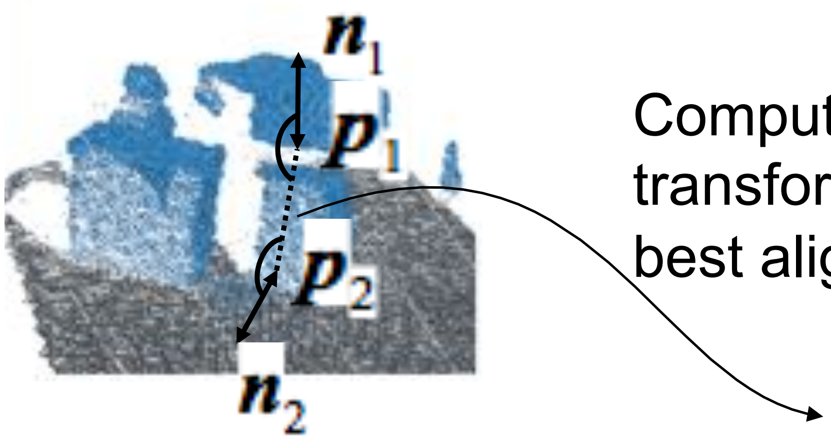


- For all pairs of surflets at distance d insert the triple

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

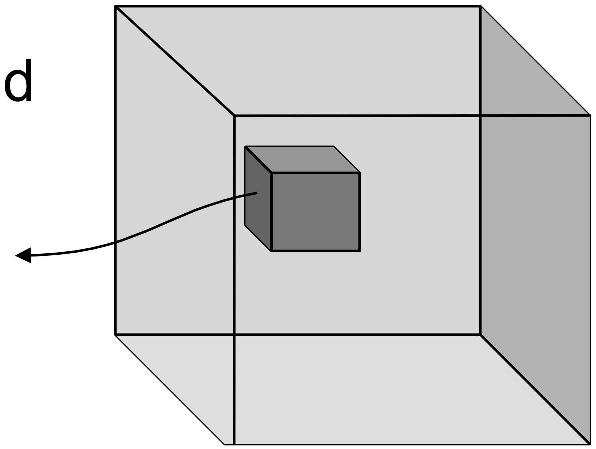
Online Recognition Phase

IJRR 2012 Special Issue, Papazov et al.



- For each model surflet pair in the hash-table cell:

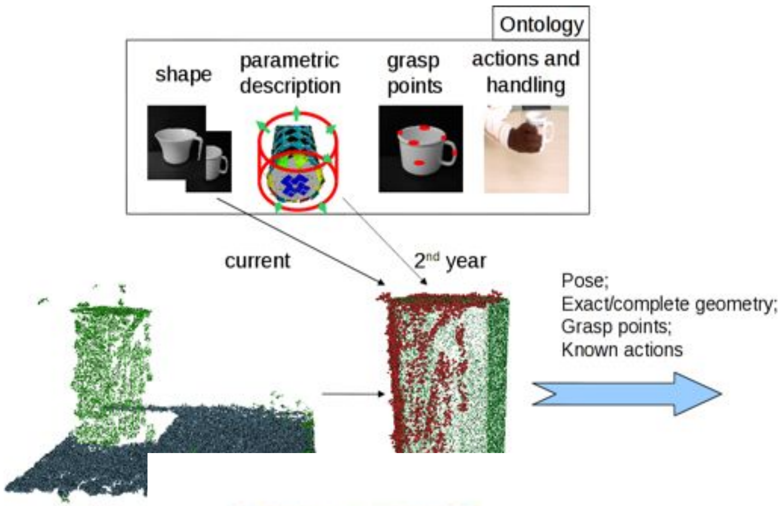
Compute the rigid transform T that best aligns



model hash-table

IJRR 2012 Special Issue, Papazov et al.

What happens if an object is similar to one in the database?



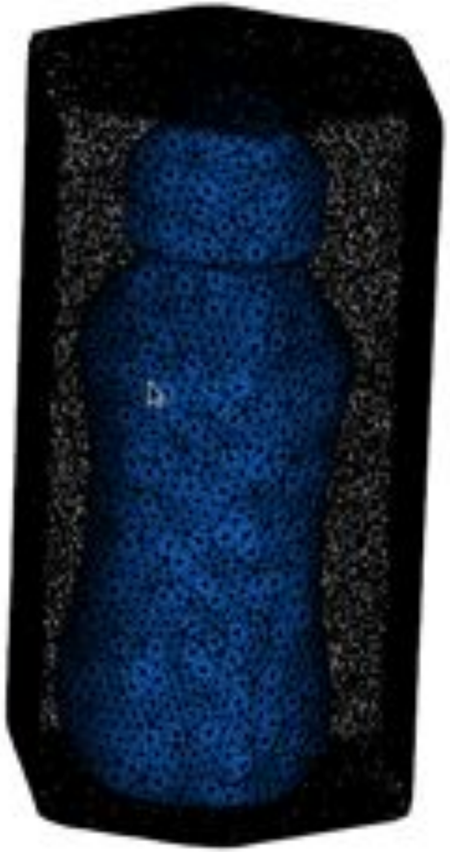
Indexing to the Atlas database needs to be extended to object classes

-> deformable shape registration needed

Atlas information

Observed object

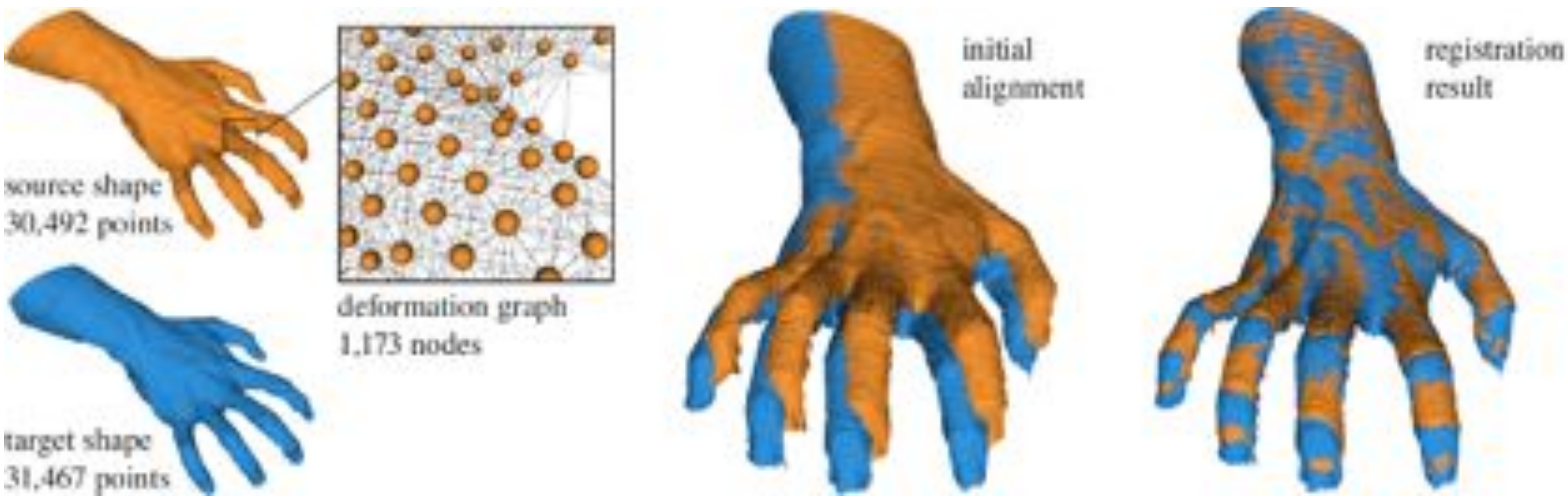
Deformable Registration from generic models (special issue SGP'11 Papazov et al.)



Matching of a detailed shape to a primitive prior

The manipulation “heat map” from the generic model gets propagated

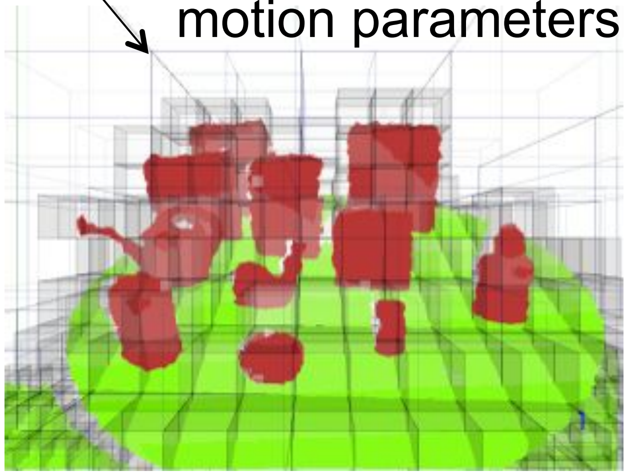
Deformable 3D Shape Registration based on Local Similarity Transforms



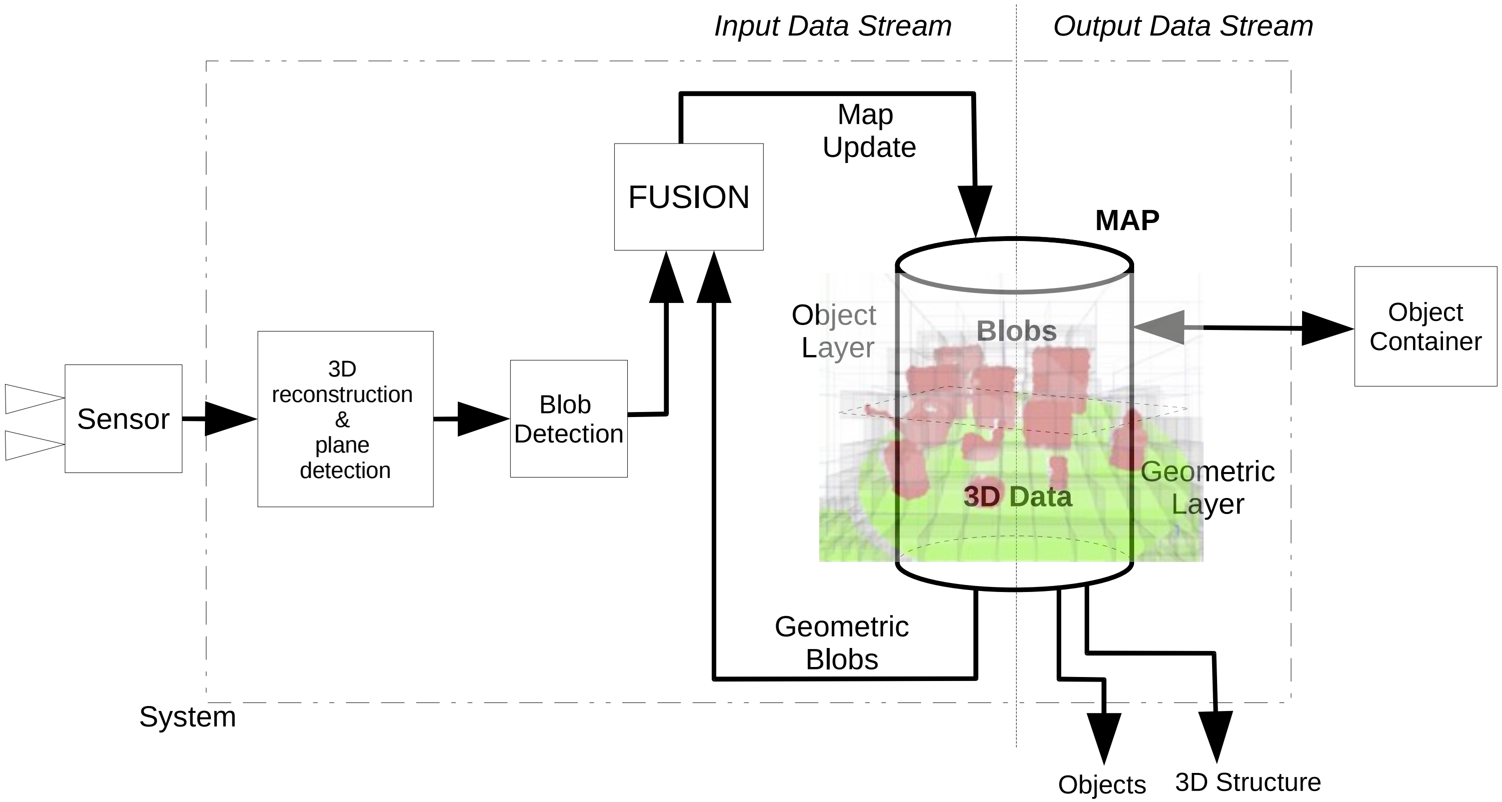
What do we try to extract from the environment?

labeling

motion parameters

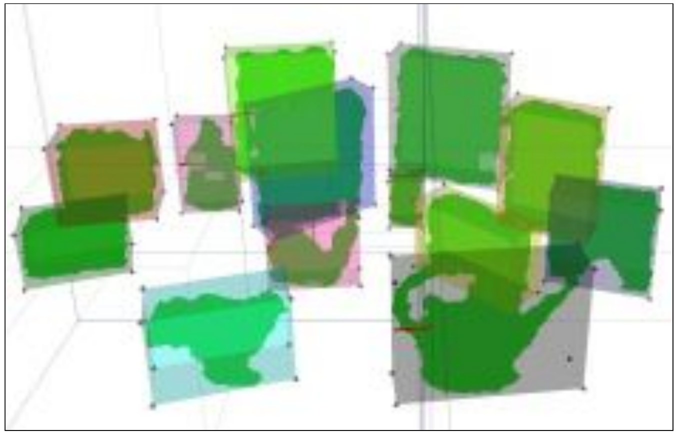


Hybrid Model of the Environment
(JC Ramirez)

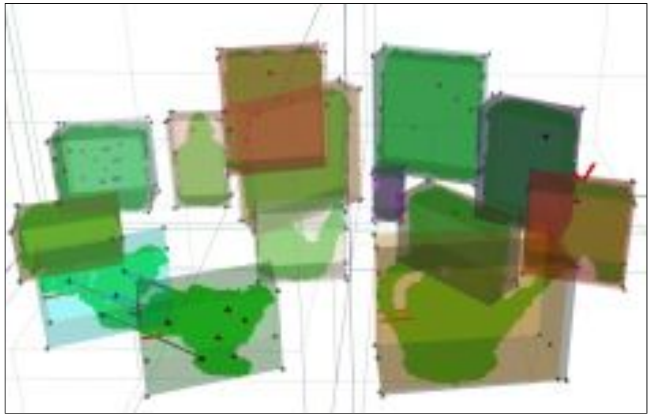


World model saves additional info, like texture, motion, etc

Scene



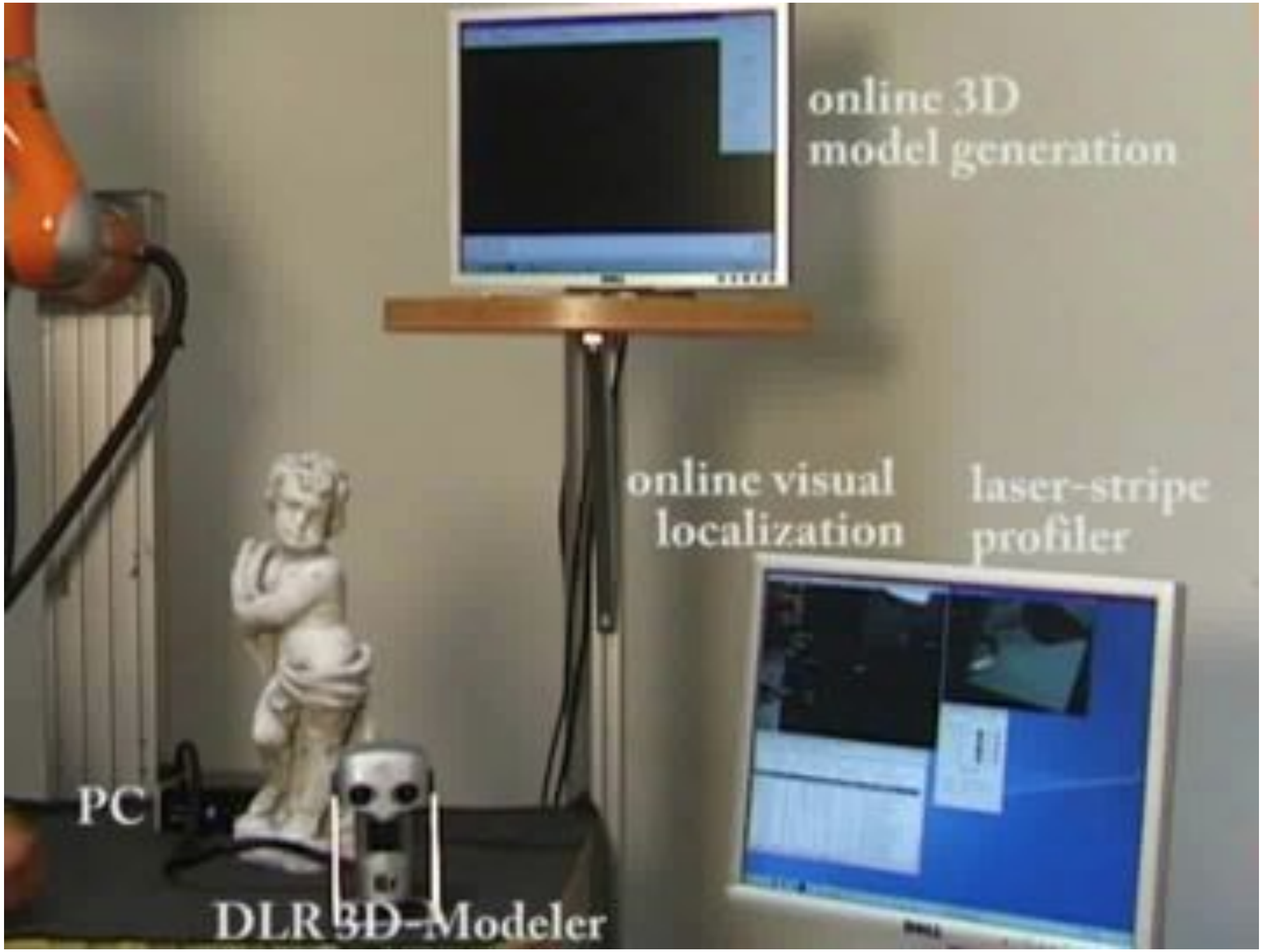
Encapsulated 3D blobs



Motion estimation

VISAPP 2013 J.Ramirez et al.)

Fusion of sensor readings
- Construction of 3D models



Machine Vision and Perception Group @ TUM

MVP

Pose Estimation

Strobl, Mair, Bodenmüller, Kielhofer, Sepp, Suppa, Burschka, Hirzinger
IROS, IEEE/RSJ, 2009, Best Paper Finalist

Mair, Strobl, Bodenmüller, Suppa, Burschka
KI, Springer Journal, 2010

- Two motion prediction concepts
  - 2D feature propagation by motion derivatives
  - IMU-based feature prediction
- Combination of both:
  - translation propagation by feature velocity (2D)
  - rotation propagation by gyroscopes

no feature propagation

Feature Propagation

Strobl, Mair, Bodenmüller, Kielhofer, Sepp, Suppa, Burschka, Hirzinger
IROS, IEEE/RSJ, 2009, Best Paper Finalist

linear feature propagation

Mair, Strobl, Bodenmüller, Suppa, Burschka
KI, Springer Journal, 2010

- Two motion prediction concepts
  - 2D feature propagation by motion derivatives
  - IMU-based feature prediction
- Combination of both:
  - translation propagation by feature velocity (2D)
  - rotation propagation by gyroscopes

Feature Propagation (Data fusion)

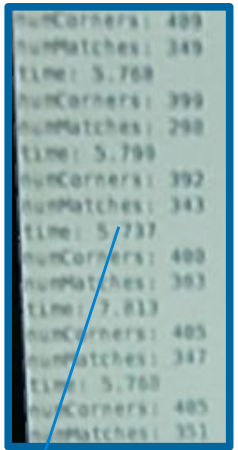
Strobl, Mair, Bodenmüller, Kielhofer, Sepp, Suppa, Burschka, Hirzinger
IROS, IEEE/RSJ, 2009, Best Paper Finalist
Mair, Strobl, Bodenmüller, Suppa, Burschka
KI, Springer Journal, 2010

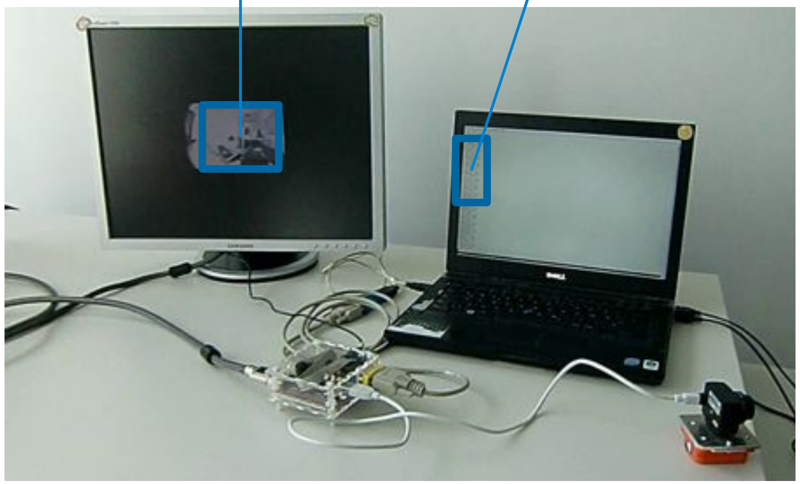
- Two motion prediction concepts
  - 2D feature propagation by motion derivatives
  - IMU-based feature prediction
- Combination of both:
  - translation propagation by feature velocity (2D)
  - rotation propagation by gyroscopes

linear + gyros based prop.

Local Feature Tracking Algorithms

- Image-gradient based → Extended KLT (ExtKLT)
  - patch-based implementation
  - feature propagation
  - corner-binding
  - \+ sub-pixel accuracy
  - algorithm scales bad with number of features
- Tracking-By-Matching → AGAST
  - AGAST corner detector
  - efficient descriptor
  - high frame-rates (hundrets of features in a few milliseconds)
  - \+ algorithm scales well with number of features
  - pixel-accuracy



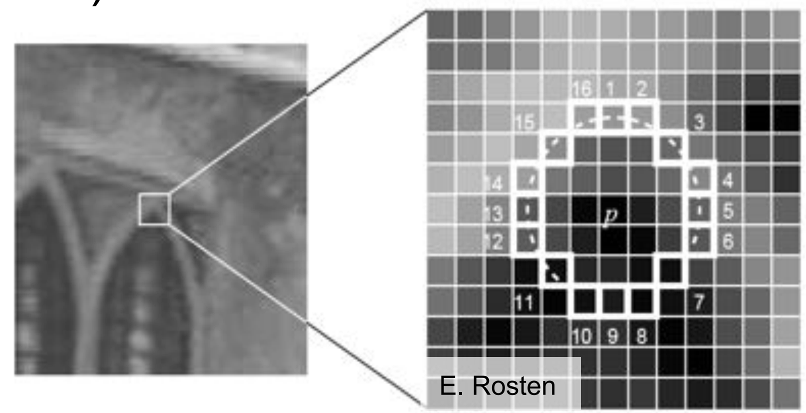


Adaptive and Generic Accelerated Segment Test (AGAST)

Mair, Hager, Burschka, Suppa, Hirzinger
ECCV, Springer, 2010

Improvements compared to FAST:

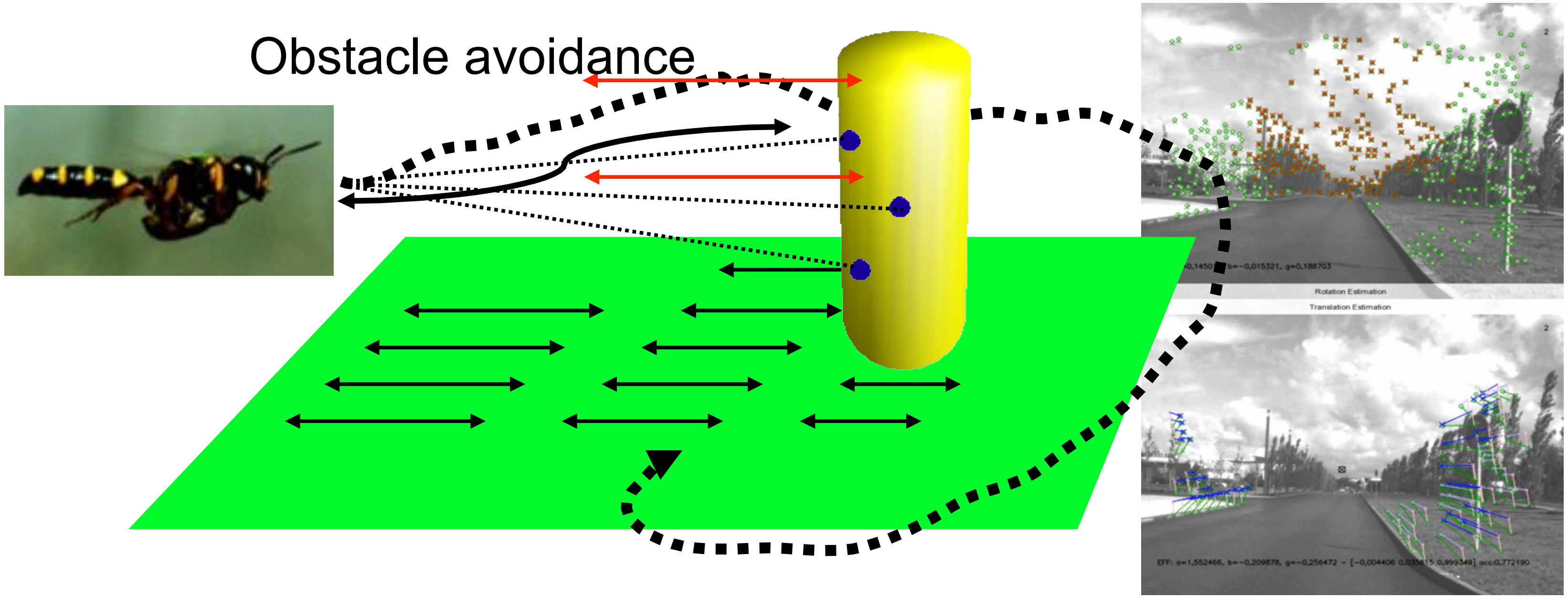
- full exploration of the configuration space by backward-induction (no learning)
- binary decision tree (not ternary)
- computation of the actual probability and processing costs (no greedy algorithm)
- automatic scene adaption by tree switching (at no cost)
- various corner pattern sizes (not just one)



<u>No drawbacks!</u>

Real Time Pose Tracking

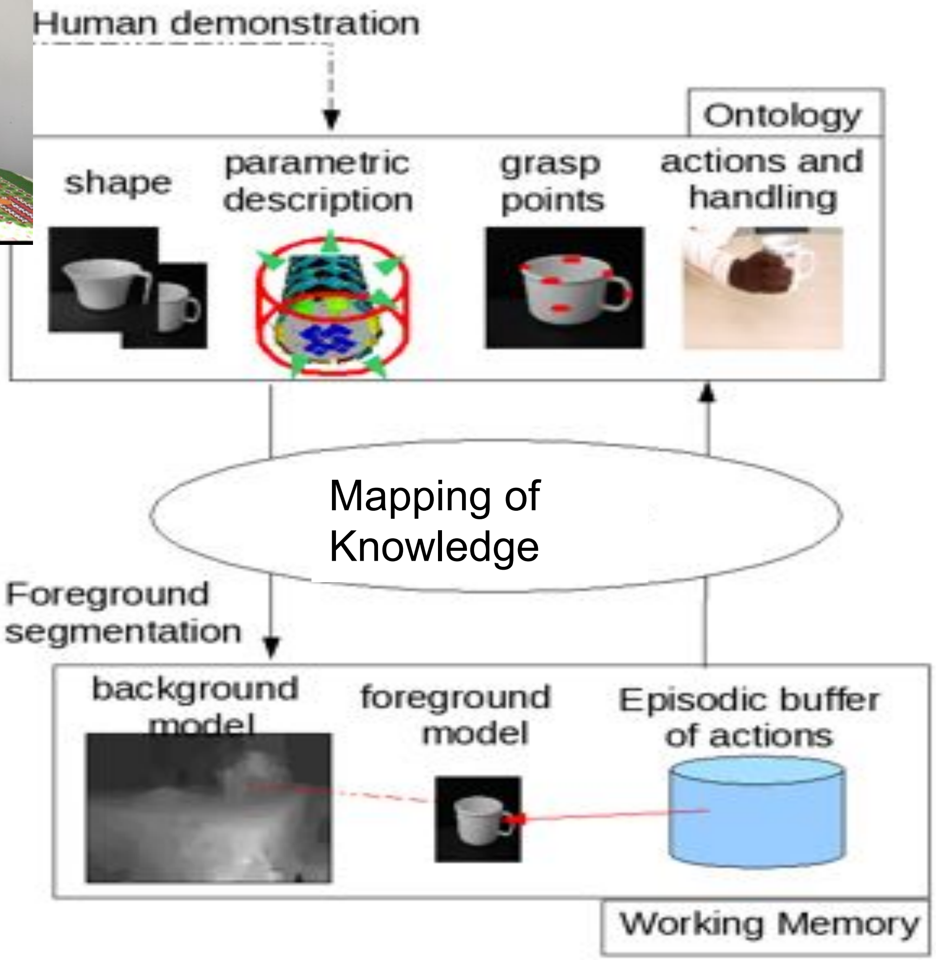
Navigation system with accuracy estimation



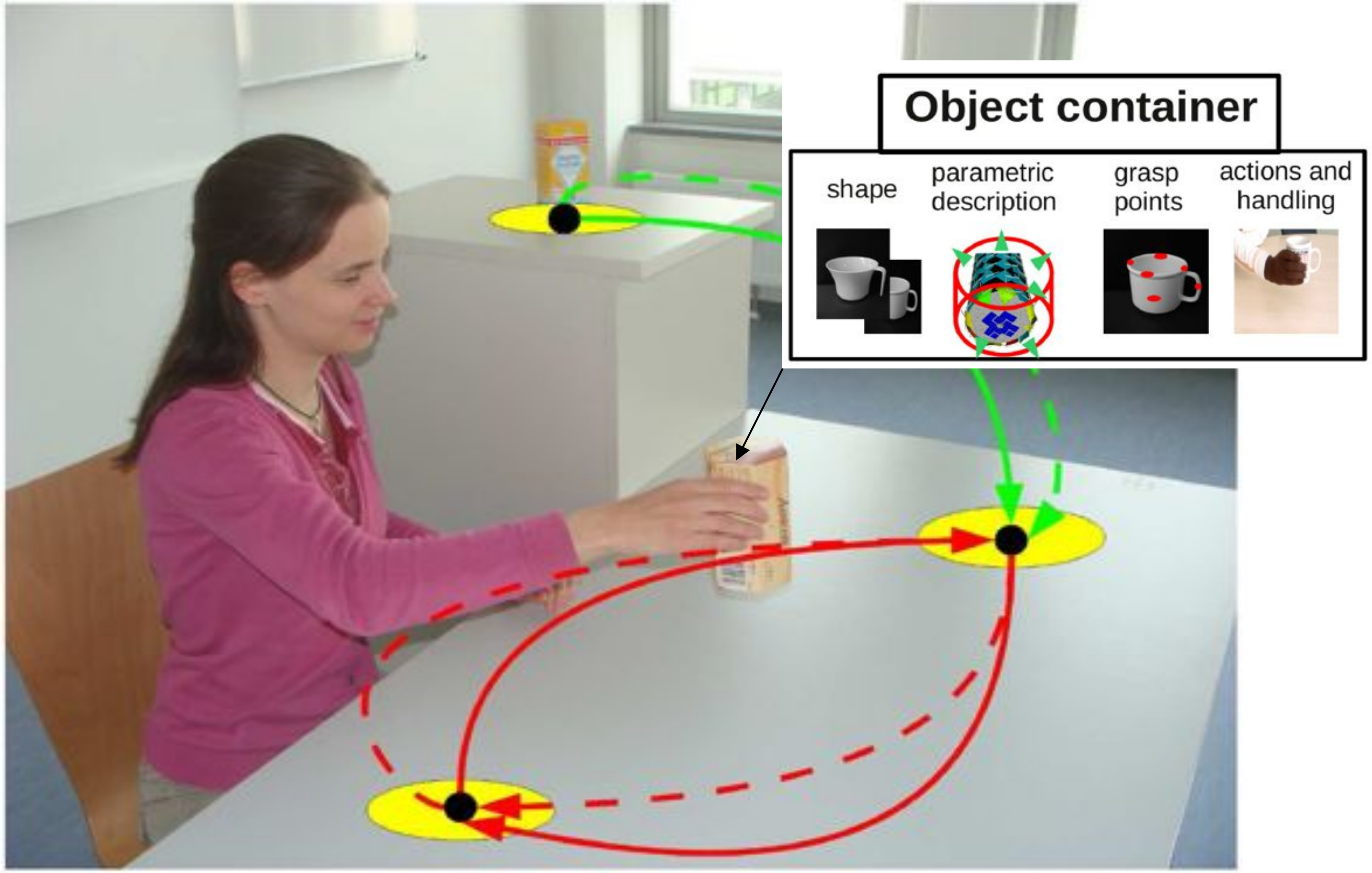
Perception-based loop closure

Navigation example

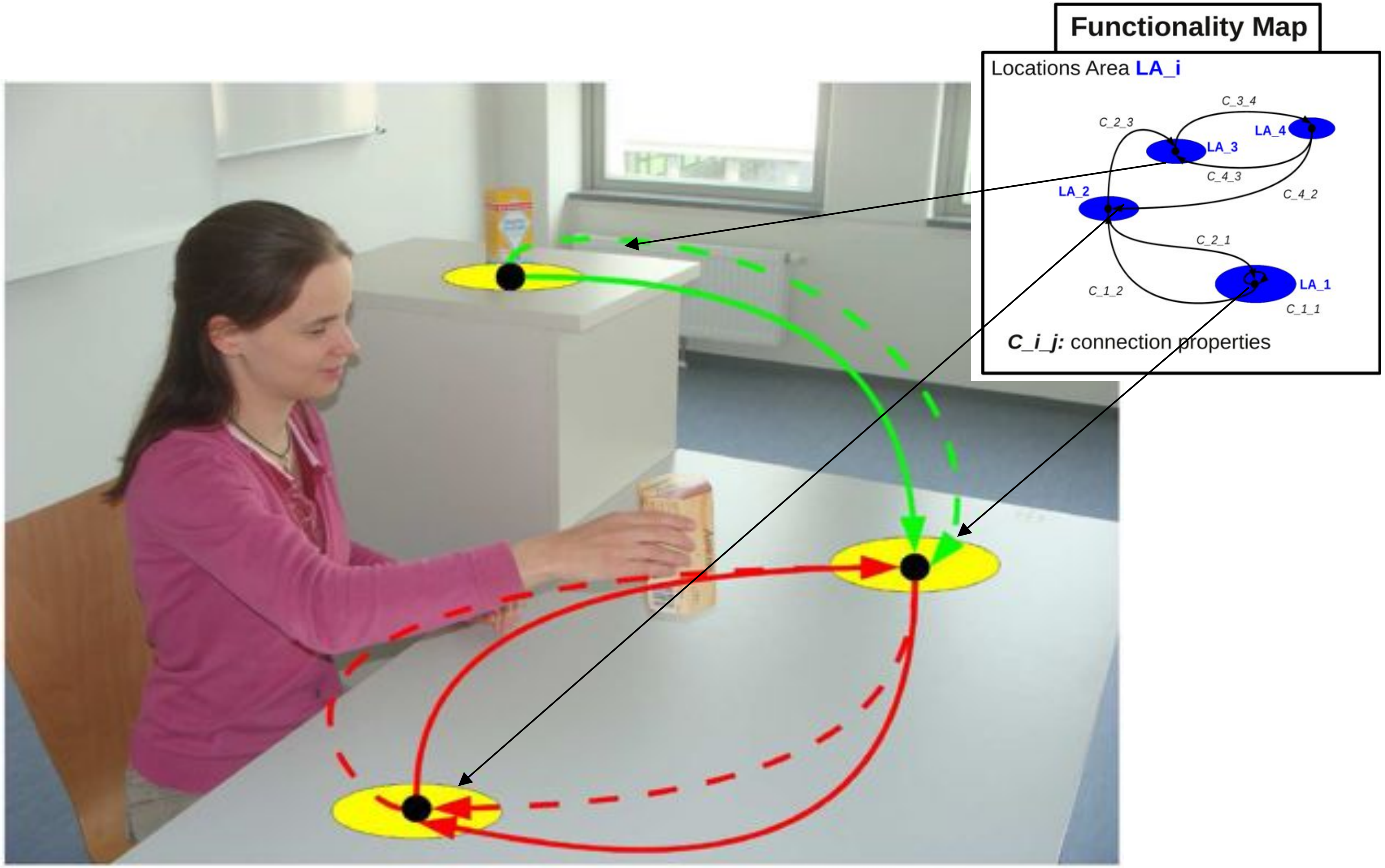
Estimation of unobservable properties



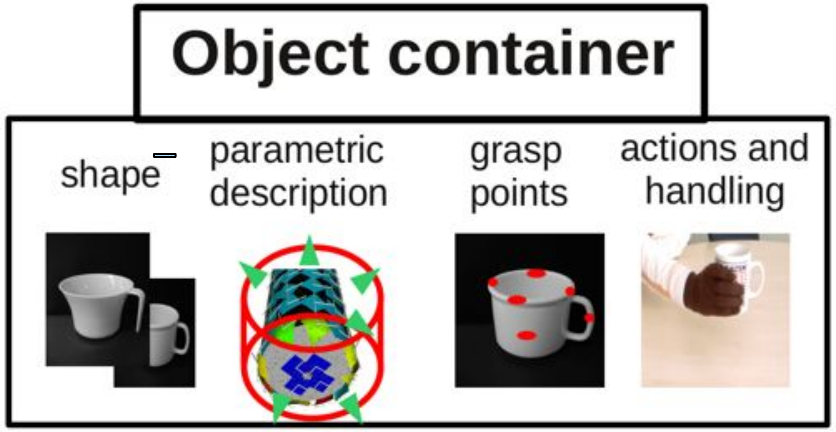
Physical and Geometric Properties of an Object (ICRA 2012 Petsch et al.)



Functional Properties of an Object stored in Functionality Map

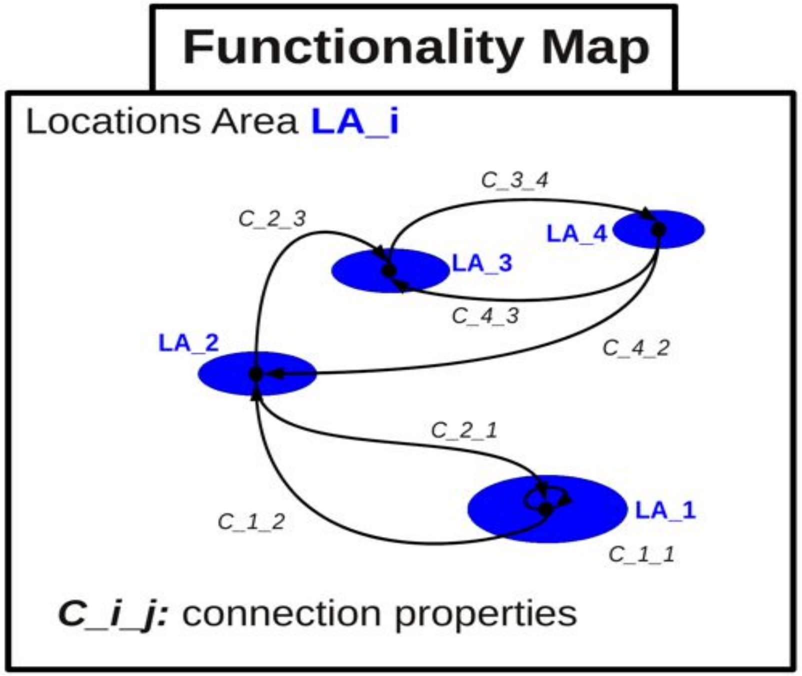


Knowledge Representation



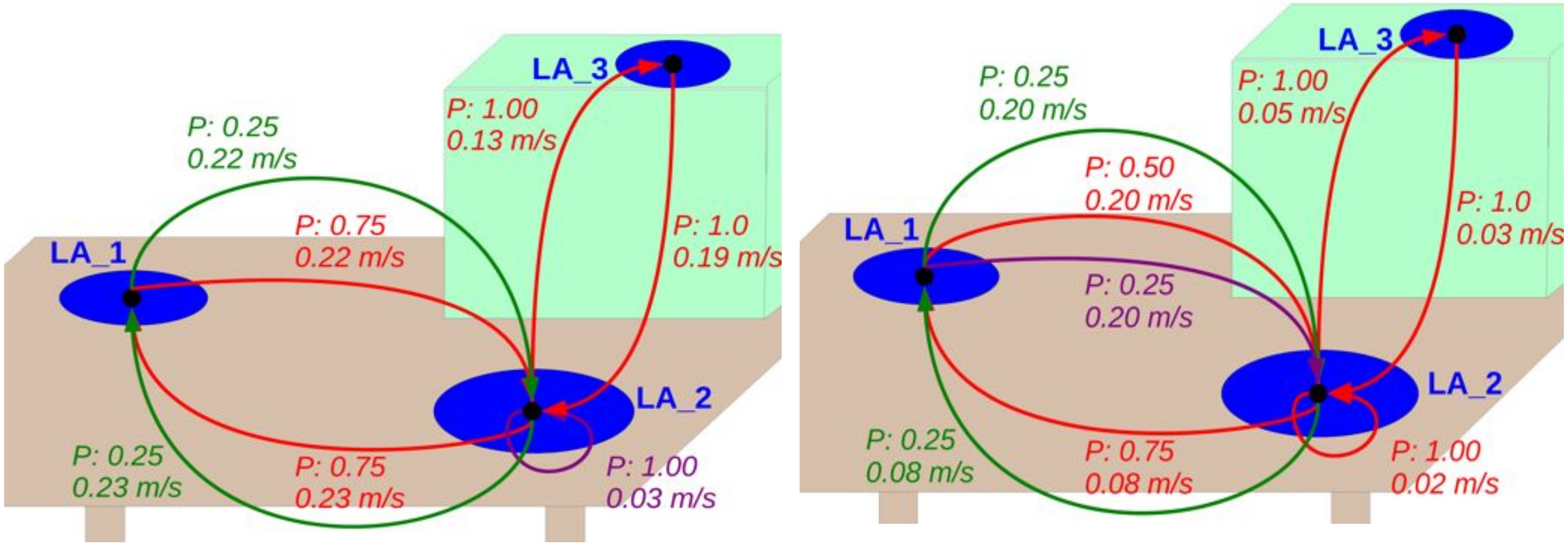
Each tool used in the procedure has its own container describing its shape, handling properties etc.

(Petsch/Burschka IROS2011)



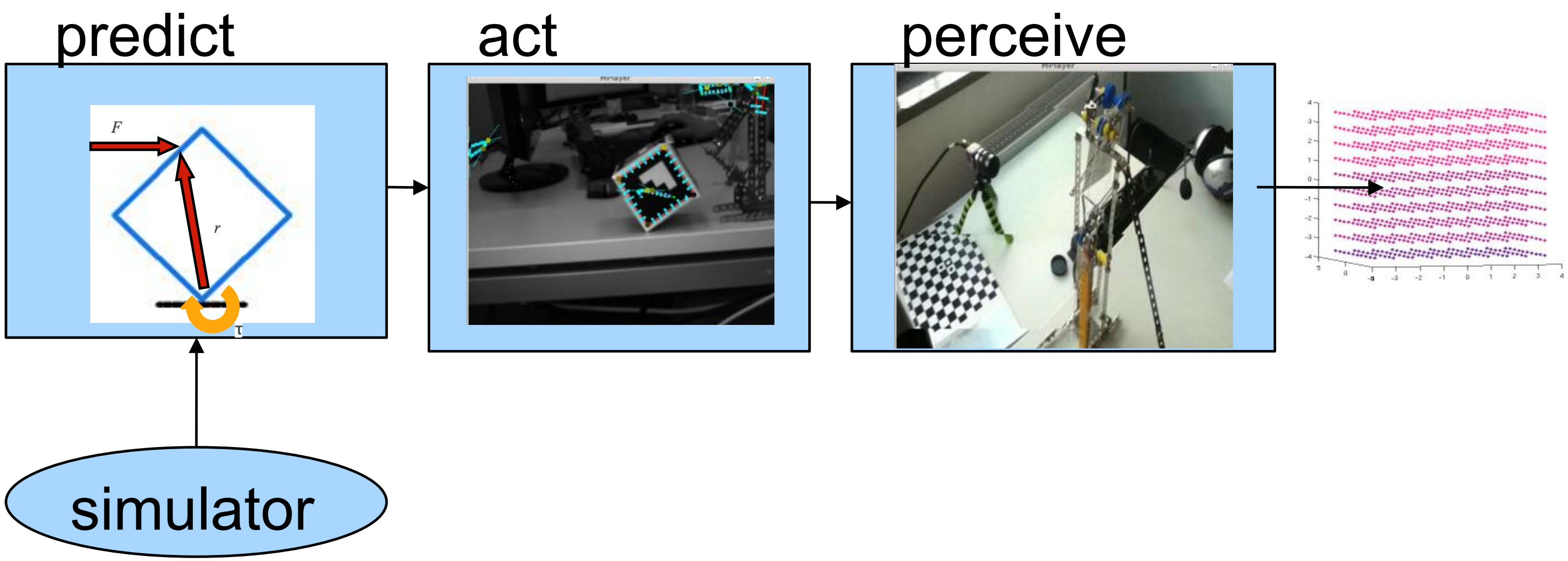
Functionality map for a specific procedure describes the way how the tool was used during the procedure while moved between points in the world

Functionality Maps



Estimation of non-observable Scene properties (robot plays)

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



Challenges for Perception

- Perception is essential to interact with an unknown and changing environments
- Perception allows to compensate own uncertainties of the robot by providing robust reference to the environment
- Level of abstraction of the raw data reduces the amount of information needed to represent a scene at the cost of increased processing – an optimum for a given system needs to be found
- System needs to be able to cope not only with static scenes but also with dynamic changes (actions, failures) in the environment
- Observation of other agents (human other robots) can provide essential data to fill in the unobservable parameters of the scene