Marker-less Object Perception and Articulation Discovery

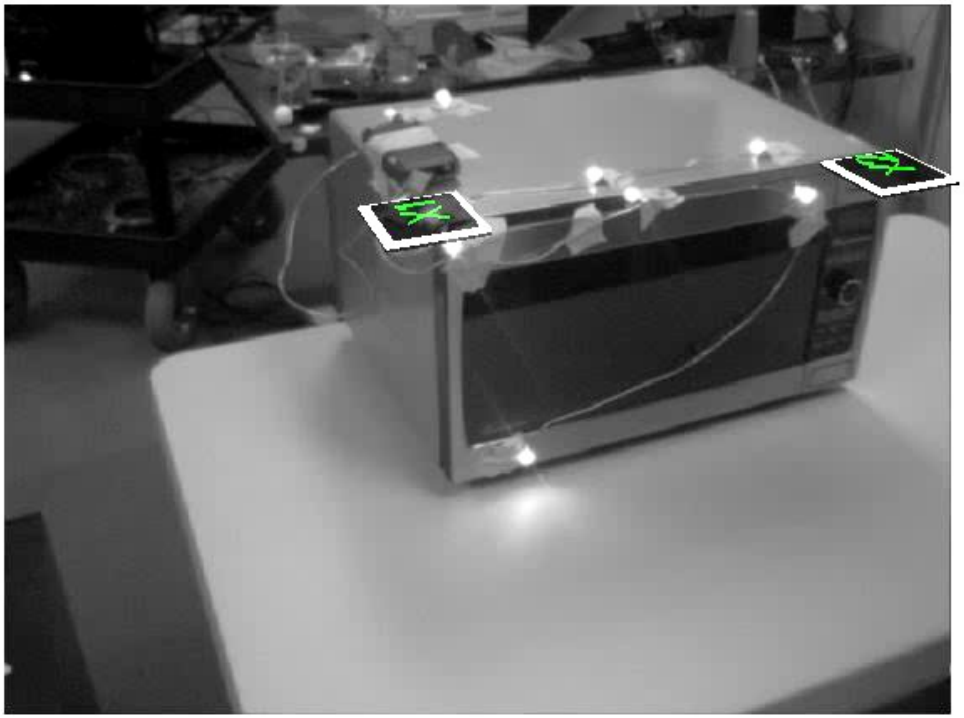
Jürgen Sturm
Kurt Konolige

ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG

Previous work

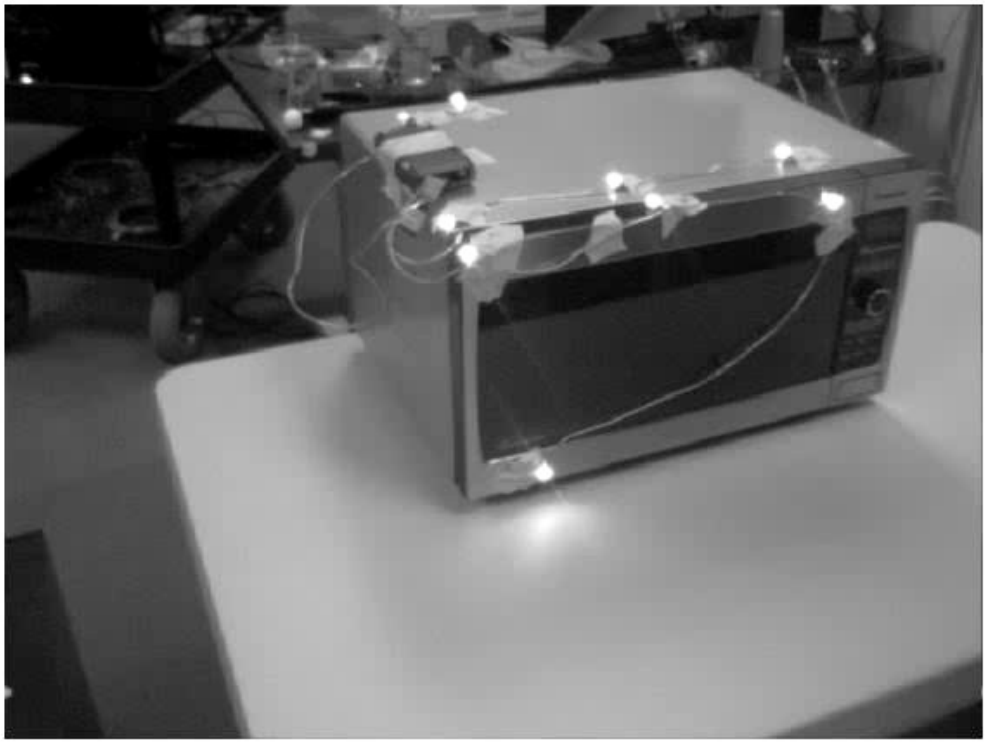
- Learn articulation models from pose observations
- Model selection
  - Rotational model
  - Prismatic model
  - Non-parametric LLE/GP model
- Structure discovery

Microwave Door: Observations



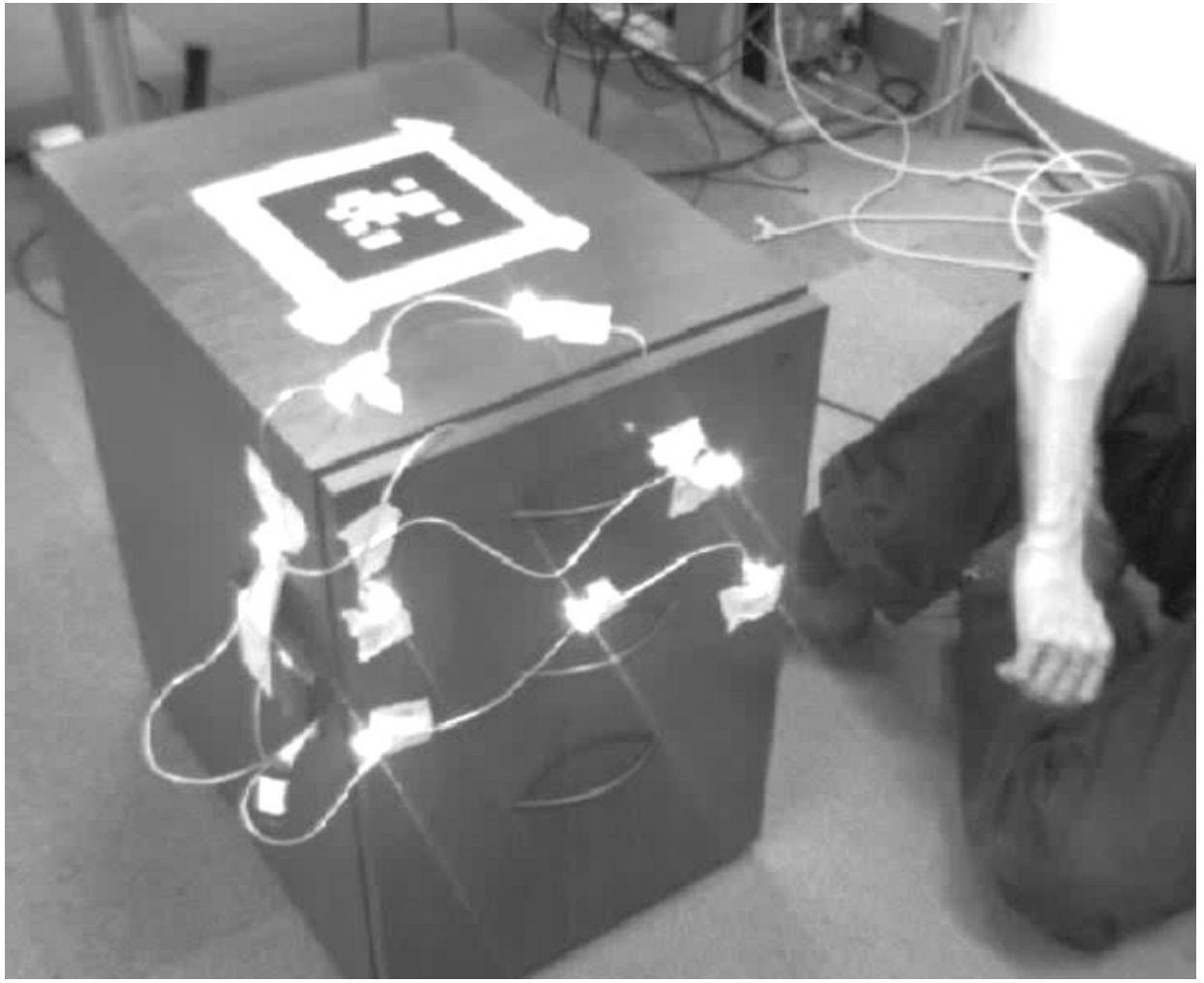
microwave pose observations
from motion capturing studio

Microwave Door: Learned Model



learned model for microwave door

Cabinet with Two Drawers



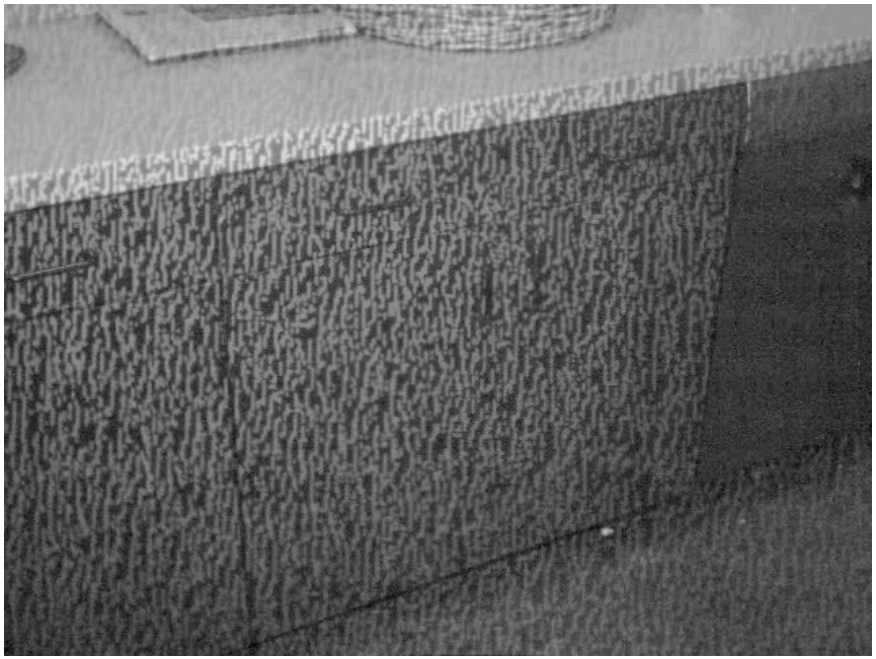
learned models and structure

Research question

- Can we get rid of artificial markers for pose registration?
- Can we learn articulation models in unprepared environments?

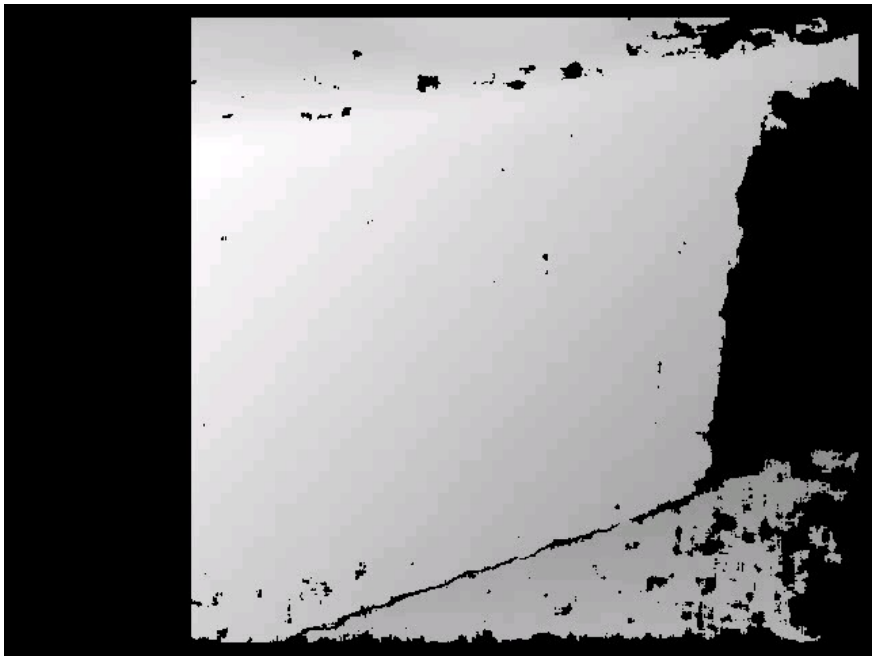
Choosing the sensor

- Use stereo vision?
  - Videre stereo camera
  - Projected light



Stereo vision + structured light

- Structured light projector adds much texture to scene
- Disparity image is dense
- Dense depth video

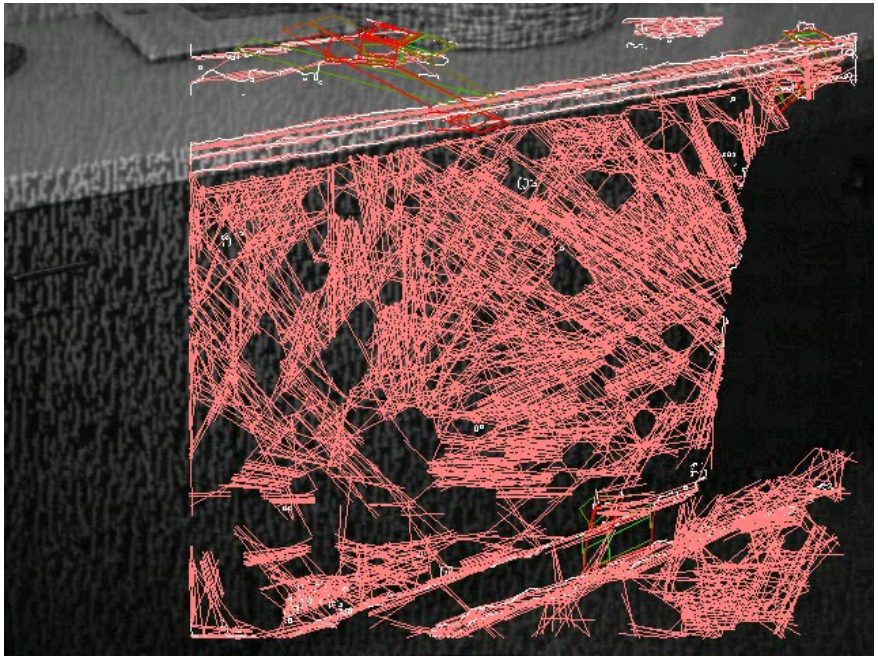


Problem formulation

- Dense stereo data
- Objects have rectangular shape
- Unknown position
- Unknown size
- Unknown orientation

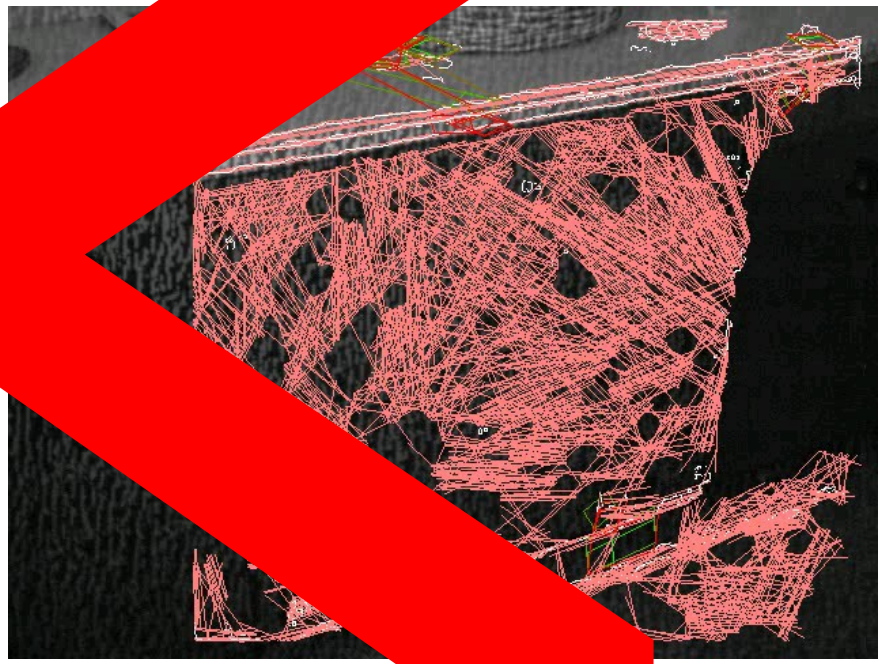
First approach

- Segment planes
- Search for edges (Canny)
- Search for lines (Hough)
- Line intersections
  → corner candidates
- Find width, height
- Optimize fit on distance transform (chamfer matching)



First approach

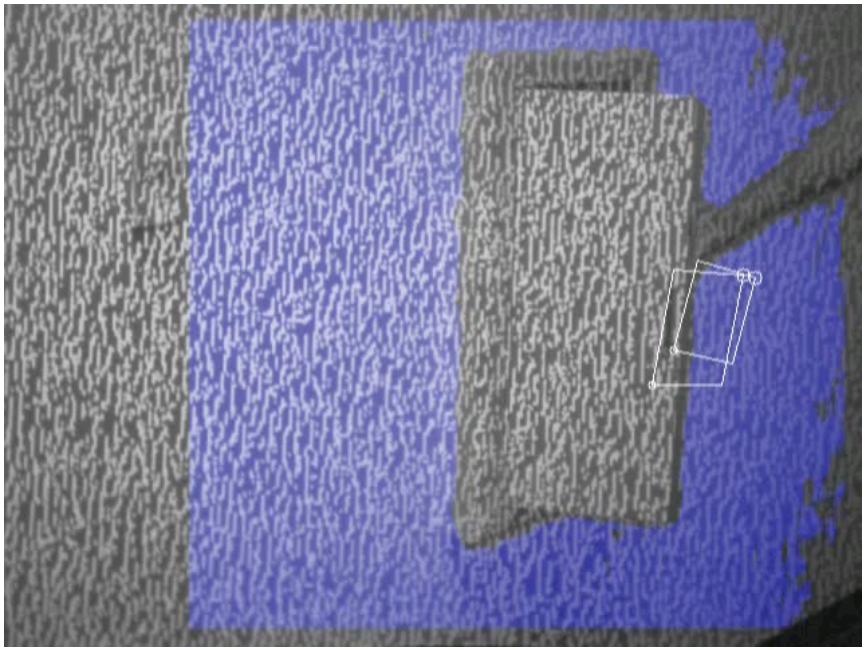
- Segment pl
- Search for edge
- Search for lines (Ho
- Line intersections
  → corner candida
- Find width, h
- Optimize ance
  trans ifer
  m



- Depends on good edge visibility
- Poor performance on doors
- Way too complicated!

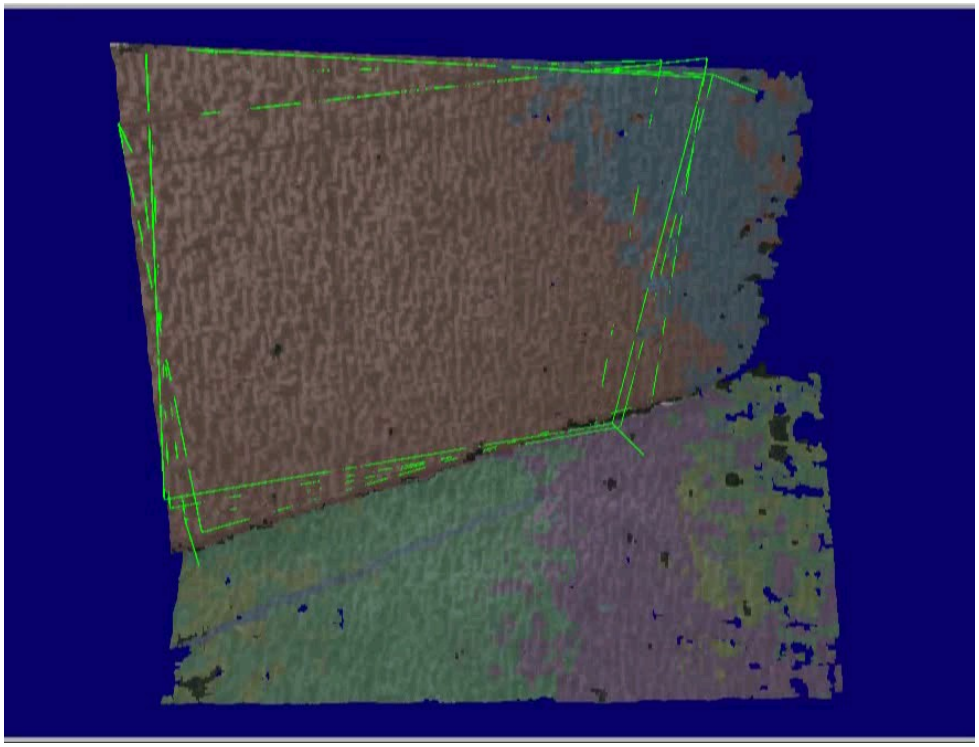
Second approach

- Segment planes
- Pick random seed pixel
- Iteratively optimize in small steps
  - width (from left)
  - width (from right)
  - height (from bottom)
  - height (from top)
  - rotation
- Objective function
  - fill ratio of rectangle
  - slight bias term that favors larger objects



More examples

- Cabinet door
- Cabinet drawer
- Fuse door
- Book
- Carton



Object Tracking

- Track observations over time
- Noise
- Partial observations
- Ambiguities
  - Front/backside flips
  - Rotations of 90/180/270deg
  - Track assignment
- Data association

Discover articulated objects

- Learn articulation models for tracks
- Measure model fit
- Estimate current object configuration
- Make pose predictions for unseen configurations

Conclusions

- simple object detection
- full pose estimates
- articulation model learning on natural features is possible
- (currently) limited to rectangular shaped objects

- implemented as ROS package planar_objects
  - box_detector
  - box_tracker
  - articulation_learner

➡ Demo after this talk in green room

Future work

- ground truth evaluation
- improve objective function (use occ/free/unknown)
- appearance-based matching
- add rotational articulation model
- improve plane extraction using surface normals
- optimize code (currently 1-4s per frame)
- ICRA paper