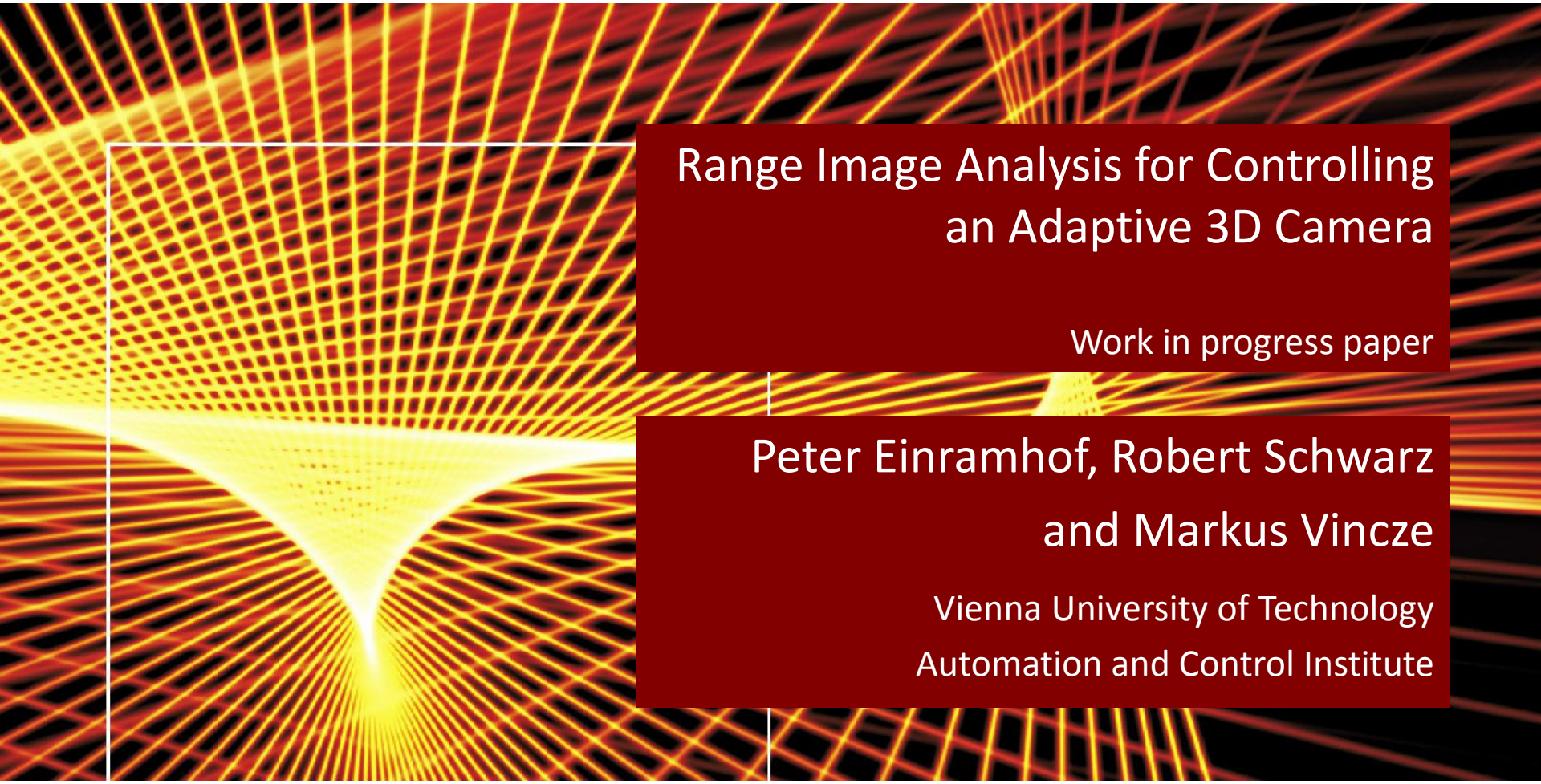


Range Image Analysis for Controlling an Adaptive 3D Camera

Work in progress paper

Peter Einramhof, Robert Schwarz and Markus Vincze

Vienna University of Technology

Automation and Control Institute

Outline

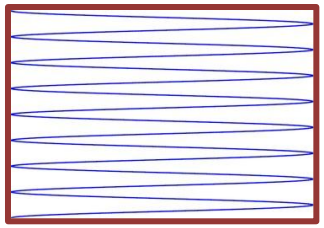
- Background
- Problem statement
- Approach
- Preliminary results
- Future work
- Acknowledgements
- References

Background

A novel 3D camera [1] is being developed in the **context of service robotics** with

- Fast, laser-based single-point time-of-flight distance measurement (1 Msample/s)
- Micro-mechanical scanning elements for two-dimensional laser beam steering (90° x 60°)

Scan plan (stylized)



Range image (250x160)

Background (continued)

Use of **foveation** (inspired by visual attention)

- Uniformly scan the scene
- Identify **task-relevant** regions by means of range image analysis → saliency map
- Scan identified regions at higher resolution

Problem statement

How can we determine **task-relevant** regions?

Basically, we have two options:

- Motion (dynamic features)
  - Deviation from ego-motion [2]
- Scene geometry (static features)
  - What features to extract from the range image?
  - How to combine these features?
  - How to incorporate task knowledge (constraints)?

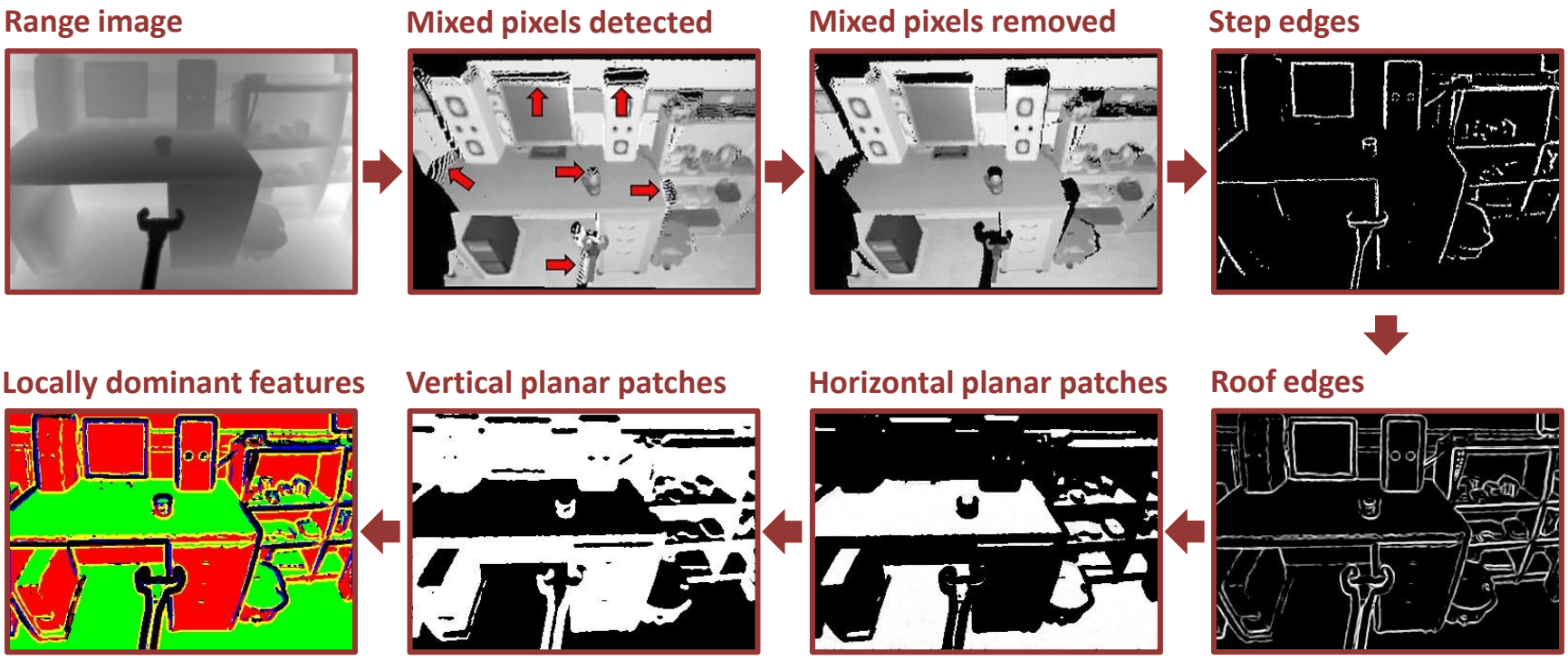
Approach

Static feature extraction from range image

- **Step edges**: Boundaries of objects, separation of foreground and background
- **Roof edges**: Transitions between objects and parts thereof, or between surface patches
- **Planar surface patches**: Boundaries of the environment, support planes, object bodies
- **Mixed pixels**: invalid pixels – to be removed

Approach (continued)

Feature extraction example (for details see [3])



Approach (continued)

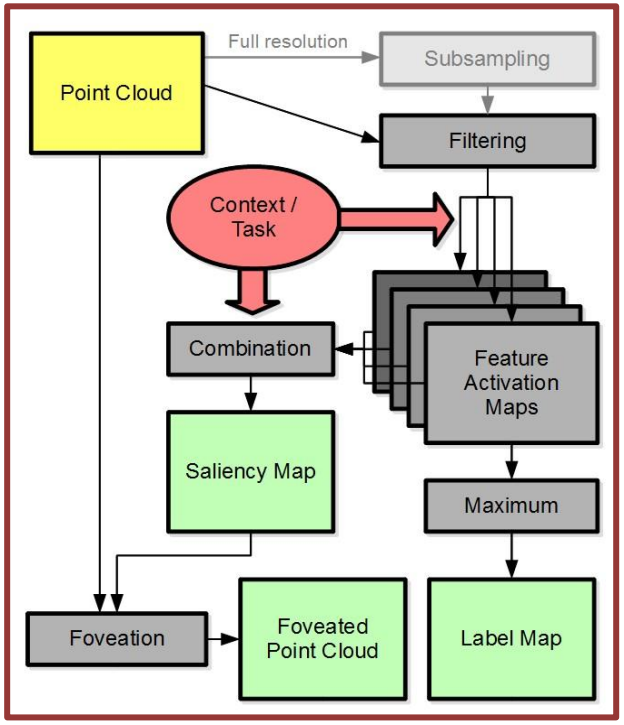
Bottom-up saliency map generation

- Adaption of the well-known bottom-up visual attention approach by Itti and Koch [4] to range images instead of 2D colour images
- Allowing products of feature activation maps to force co-occurrence of features
- Incorporation of spatial constraints (height, range) [5]

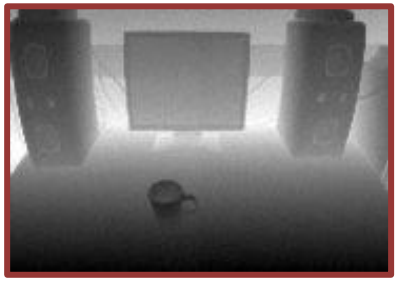
Approach (continued)

Bottom-up saliency map generation – concept

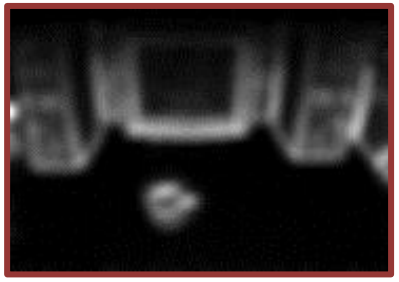
Concept



Range image



Saliency map



Approach (continued)

Bottom-up saliency map generation – drawbacks

- Not straight forward to formulate task-dependent constraints
- Robots interact with the real world on object level rather than on feature level
- Outlines of salient regions [0..1] and overall saliency have to be determined before scan plan generation

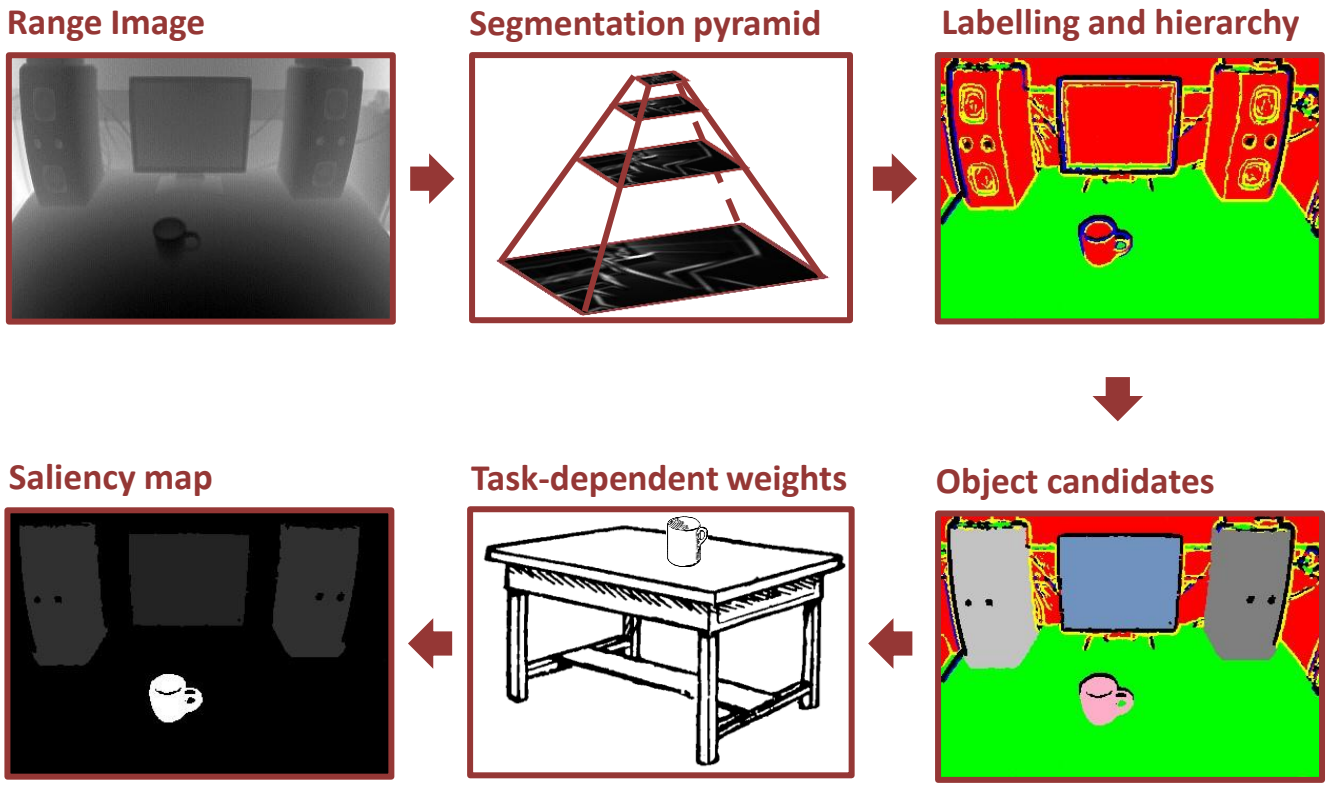
Approach (continued)

Object-based saliency map generation [6]

- Multi-resolution segmentation
- Find support planes and establish BG/FG hierarchy of object candidates with attributes such as 3D bounding box, pose and distance
- Candidates get a saliency value based on how well their attributes fit the target object
- Sequentially attended by descending saliency

Approach (continued)

Object-based saliency map generation – concept

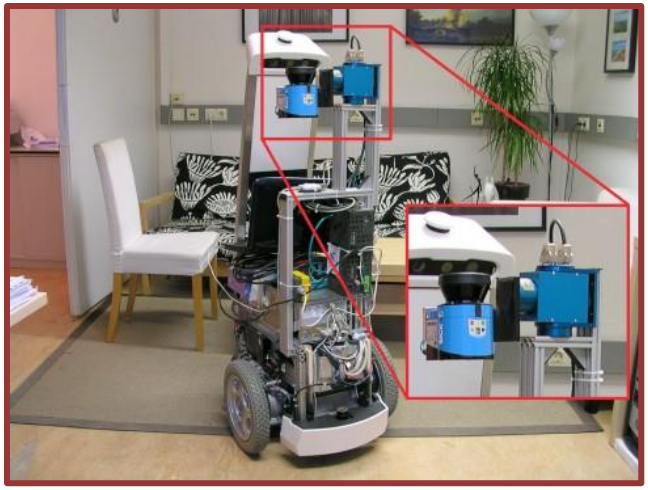


Preliminary results

- Test data used
  - Tilting laser scanner
  - 2,136 range images (360x500)
  - FOV: 90°(H) x 62.5°(V)
- Processing time
  - Notebook Core i5-430M
  - C++, single thread, no SSE
  - 46ms @ 360x250
    22ms @ 250x160

Tilting laser scanner

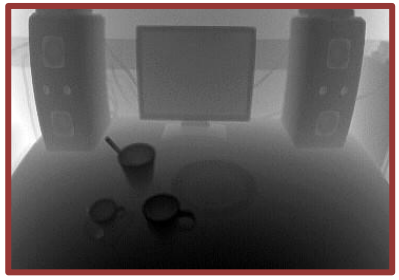
Laser scanner mounted on mobile robot



Preliminary results (continued)

Segmentation for a table scene and a door scene

Range image



Locally dominant features

Range image

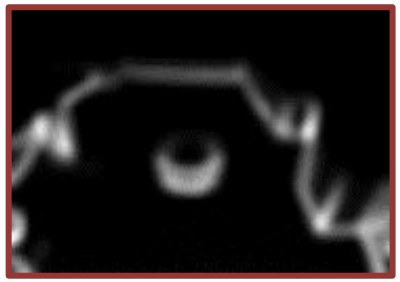
Locally dominant features

Preliminary results (continued)

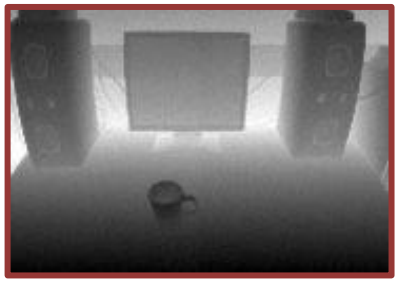
Bottom-up saliency maps for obstacle detection, table scene and door scene

Range image

Saliency map

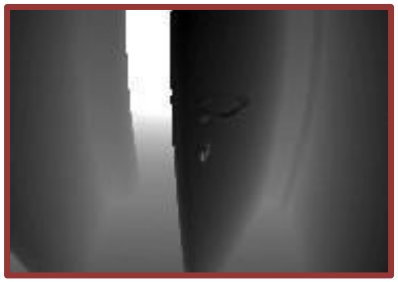


Range image

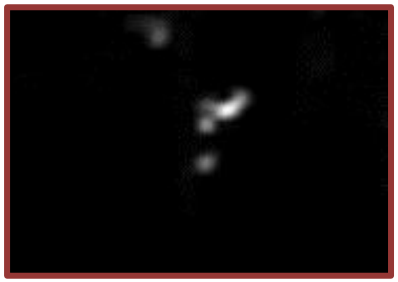


Saliency map

Range image

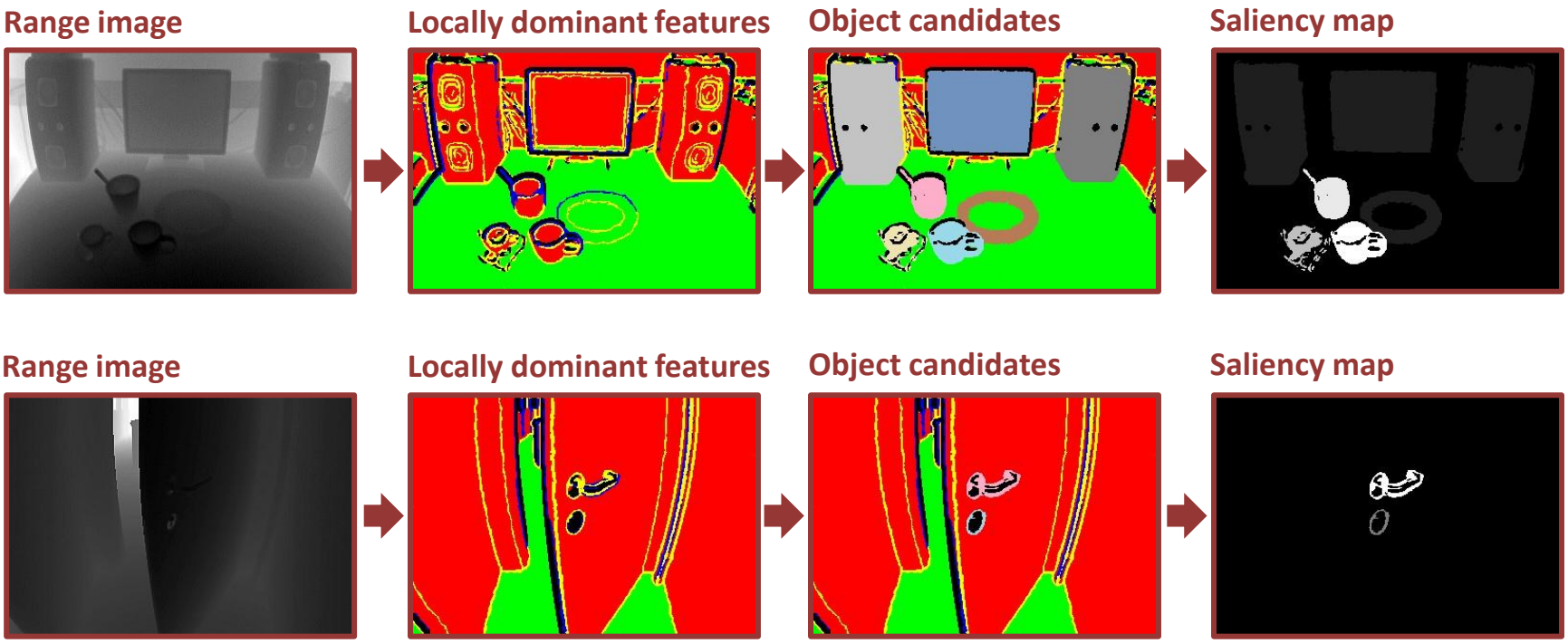


Saliency map



Preliminary results (continued)

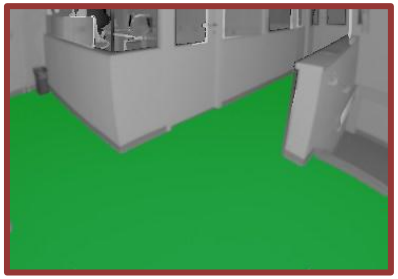
Object-based saliency maps for a table scene and a door scene



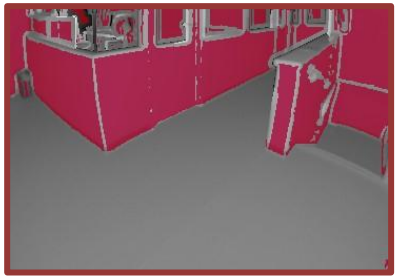
Preliminary results (continued)

Using features (without foveation) for robotic tasks such as navigation and object detection

Floor (planar patch)



Walls (planar patches)



Outlines (edges)

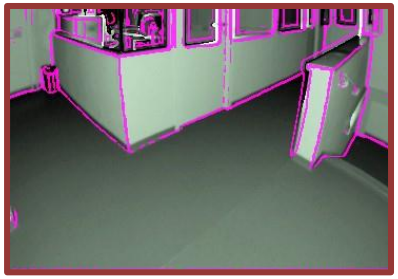


Table top (planar patch)

Walls (planar patches)

Outlines (edges)

Future work

- Improving the quality of extracted features
- Improving the results of object-based saliency maps in cluttered environments
- Embedding the range image analysis software into the sensor hardware

Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°248623.

References

[1] J.T. Thielemann, T. Sandner, S. Schwarzer, U. Cupcic, H. Schumann-Olsen, T. Kirkhus, "*TACO: A Three-dimensional Camera with Object Detection and Foveation*", Smarter sensors, easier processing – SAB 2010 workshops, Paris, France, August 24, 2010.

[2] G.M. Breivik, J.T. Thielemann, A. Berge, Ø. Skotheim, T. Kirkhus, "*A Motion based Real-time Foveation Control Loop for Rapid and Relevant 3D Laser Scanning*", 7th IEEE Workshop on Embedded Computer Vision, ECVW 2011, Colorado Springs, CO, USA, June 20, 2011.

[3] P. Einramhof, R. Schwarz, M. Vincze, "*Fast Range Image Segmentation for a Domestic Service Robot*", 20th International Workshop on Robotics in Alpe-Adria-Danube Region, RAAD 2011, Brno, Czech Republic, October 5-7, 2011.

[4] L. Itti and C. Koch, "*Computational modelling of visual attention*", Nat. Rev. Neurosci, 2(3):194–203, March 2001.

[5] R. Schwarz, P. Einramhof, M. Vincze, "*Real-time Foveation System based on dense 2.5D data*", 20th International Workshop on Robotics in Alpe-Adria-Danube Region, RAAD 2011, Brno, Czech Republic, October 5-7, 2011.

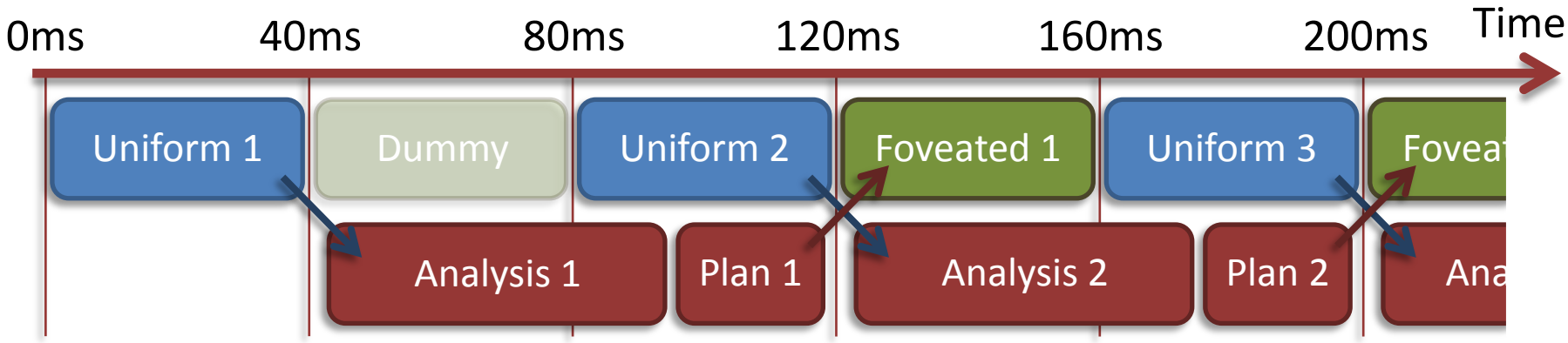
[6] P. Einramhof, R. Schwarz, M. Vincze, "*Range Image Segmentation for Object-based Attention in the Context of Service Robotics*", Poster: Rovereto Attention Workshop (RAW), Rovereto, Italy, October 27-29, 2011.

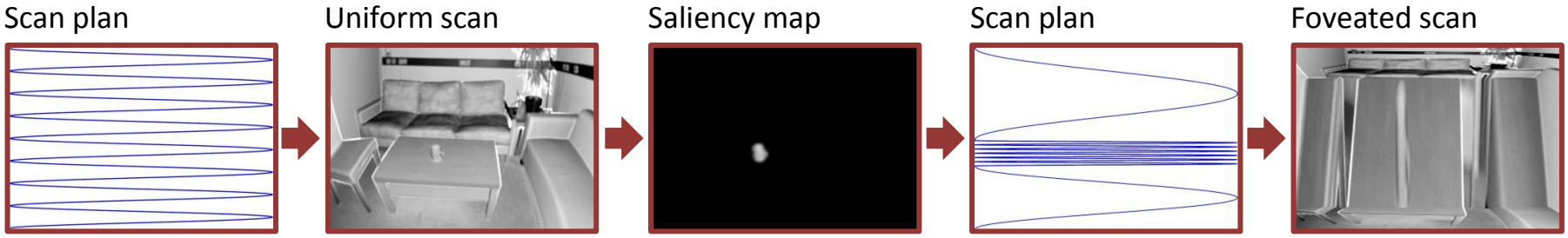
Thank you!

Questions?

Background (auxiliary)

Timeline of foveation

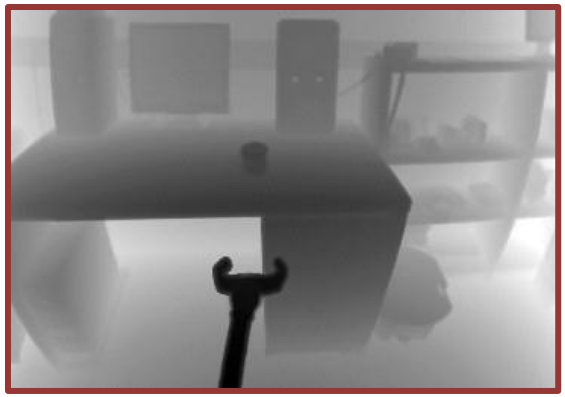


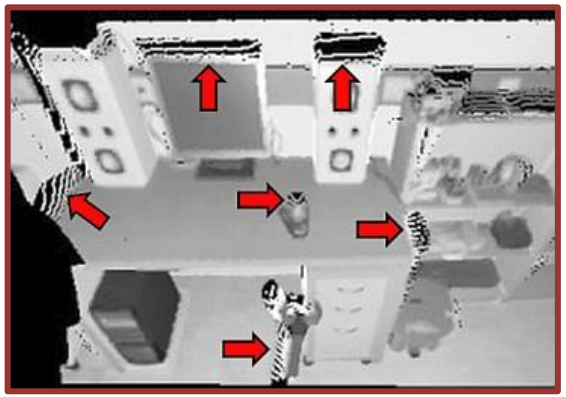


Approach (auxiliary)

Mixed pixels

- Occur at depth discontinuities
- Connect foreground and background
- Don‘t correspond to physical structure

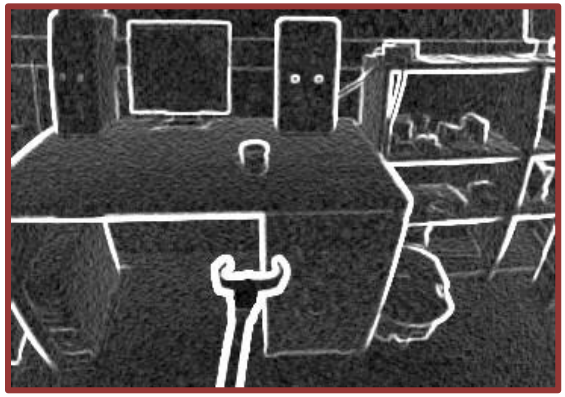


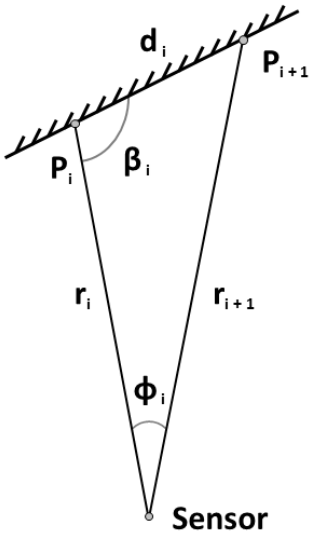


Approach (auxiliary)

Identification of mixed pixels

- High local standard deviation of range values
- Bearing angle close to 0° or 180°





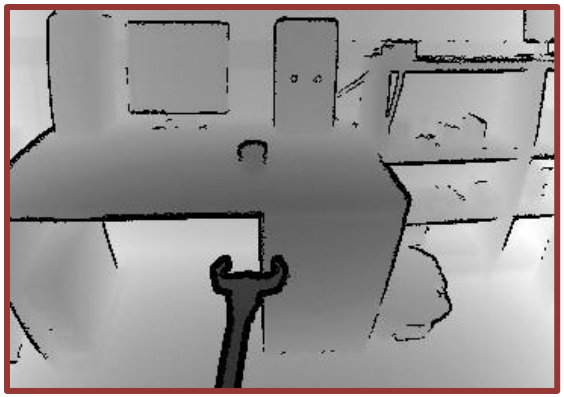
$$d_i = \sqrt{r_i^2 + r_{i+1}^2 - r_i r_{i+1} \cos \Phi_i}$$

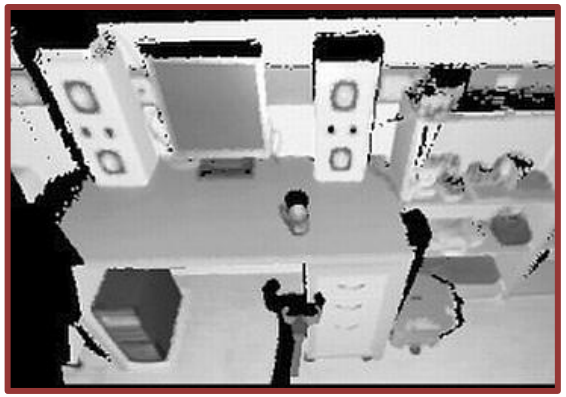
$$\beta_i = arccos\left(\frac{r_i - r_{i+1} \cos \Phi_i}{d_i}\right)$$

Approach (auxiliary)

Removal of mixed pixels

- Histogram over local standard deviations, use peak as σ_{noise} and $3*\sigma_{noise}$ as threshold
- Allow only valid bearing angles $5° \leq \beta_i \leq 175°$

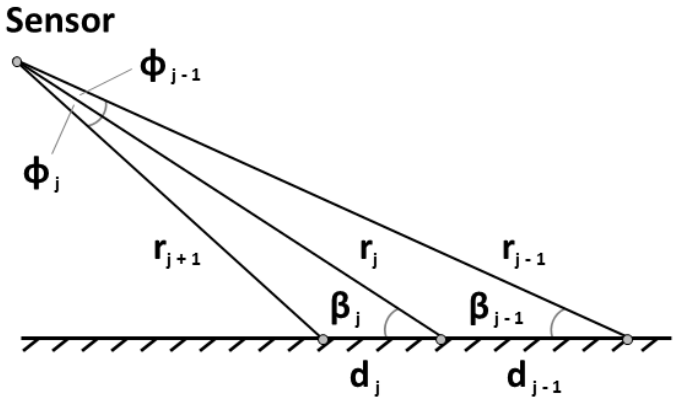




Approach (auxiliary)

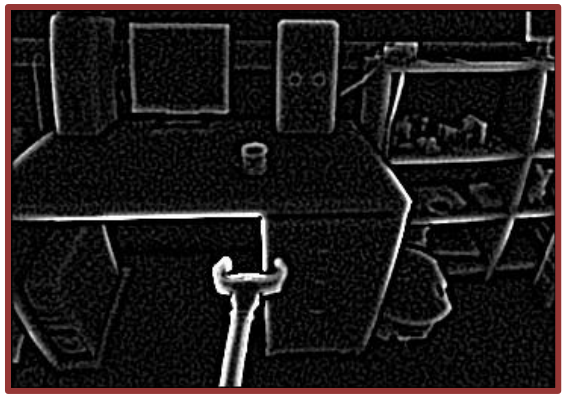
Step edges are valid pixels at depth discontinuities that belong to the foreground

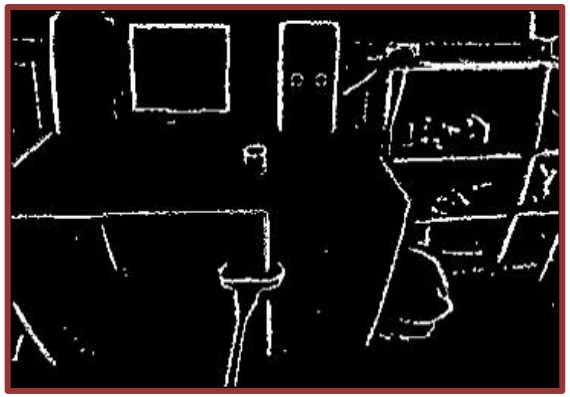
Edge detection based on the first derivative like in 2D images is problematic due to large responses at flat but valid bearing angles



Approach (auxiliary)

- We use the second derivative (3x3 Laplace)
- Sensitive to noise → noise reduction
- Only positive values, histogram, use 3x peak value as threshold, clean-up (single pixels)

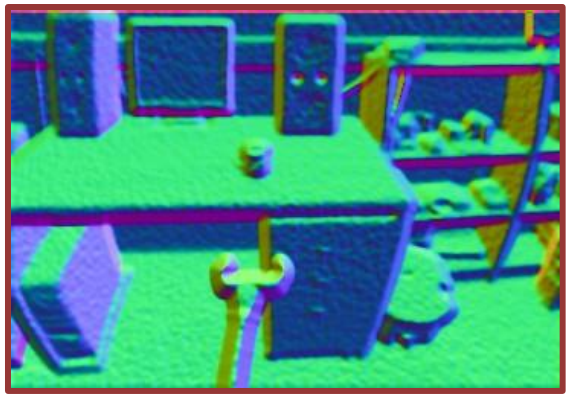


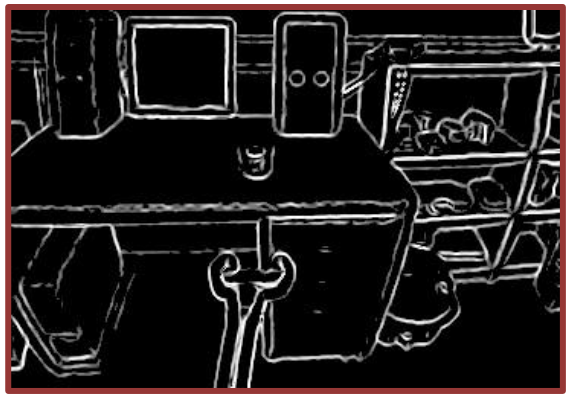


Approach (auxiliary)

Discontinuities in surface normal direction

- 3D data from smoothed range data
- Computation of surface normals (5x5 patches)
- Dot product and thresholding

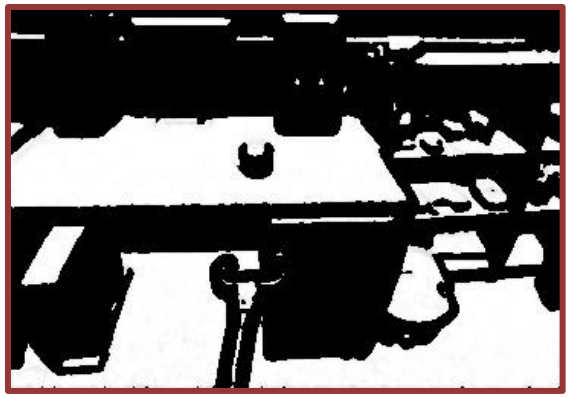


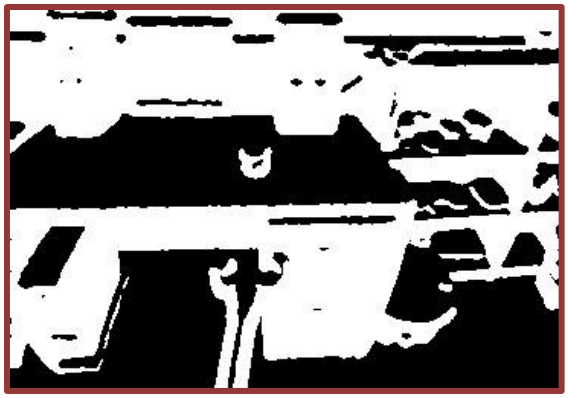


Approach (auxiliary)

Special focus on horizontal and vertical planar patches due to the relevance for robotics tasks

Dot product with vector of vertical direction, allowing a deviation of max. 10°





Approach (auxiliary)

The range image is labelled with the locally dominant feature. If two or more features are locally equally present, there‘s a prioritisation

- Mixed pixels
- Step edges
- Roof edges
- Vertical planar patches
- Horizontal planar patches