Monocular Range Sensing:
A Non-Parametric Learning Approach

Christian Plagemann, Felix Endres, Jürgen Hess, Cyrill Stachniss, Wolfram Burgard

ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG

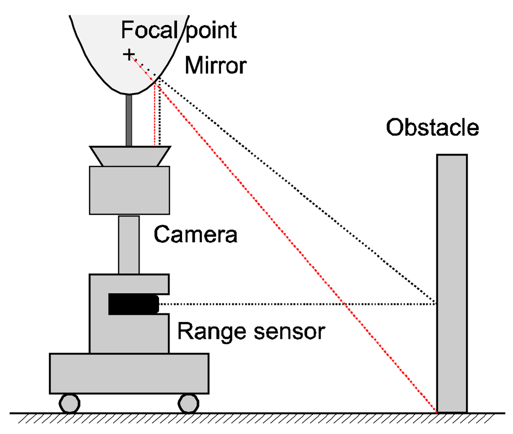
AIS Autonomous Intelligent Systems

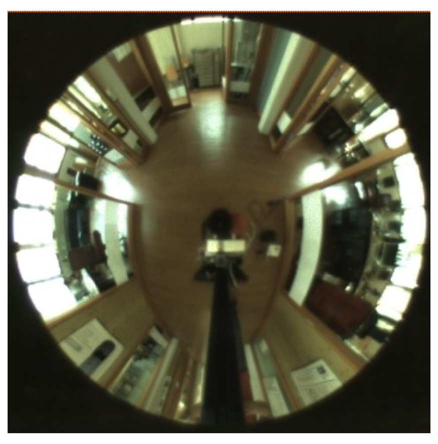
Motivation

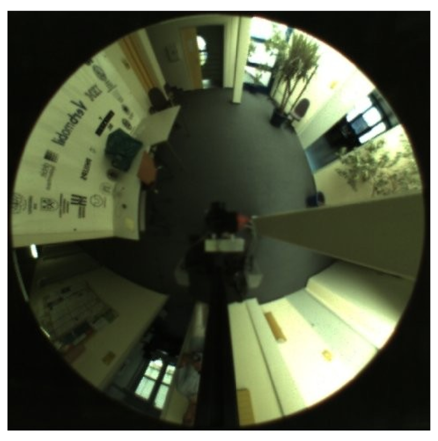
- Mobile robots need to estimate the geometry of the local surrounding area.
- Cameras are cheap and light-weight sensors but do not measure range directly.
- Idea: Learn the relationship between visual features in monocular camera images and range measurements from a laser sensor.

Learning Depth from Monocular Camera Images

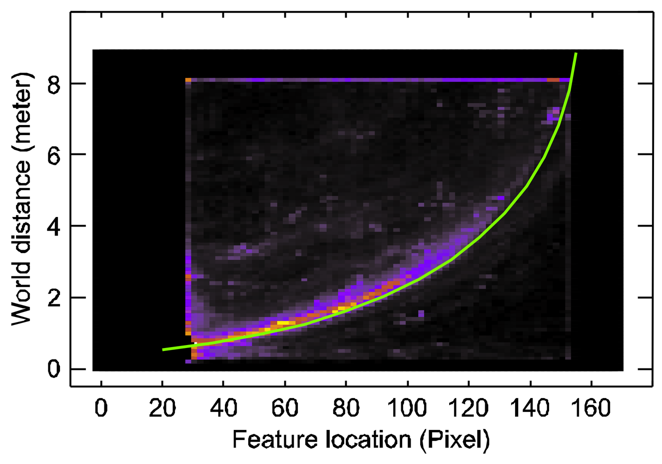
- Training setup: robot + laser + omnidir. camera







- Idea: Learn the range function $r : \mathbb{R}^{420} \rightarrow \mathbb{R}$ that maps (polar) pixel columns $\mathbf{p}$ to ranges.
- Pre-processing: Extract features $\mathbf{x}_i = f(\mathbf{p}_i)$
  - Four types of edge-based features
  - The first six principle components (PCA)



Gaussian Process Regression

- Model all ranges as *jointly* Gaussian distributed

$$r_1, \ldots, r_n \mid \mathbf{x}_1, \ldots, \mathbf{x}_n \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

using a parameterized covariance function, e.g.,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \cdot \exp\left(-\frac{1}{2\ell^2}|\mathbf{x}_i - \mathbf{x}_j|^2\right)$$

Then, *new* ranges can be predicted as

$$r^* \mid \mathbf{x}^*, \mathcal{D} \sim \mathcal{N}(\mu^*, \sigma^*)$$

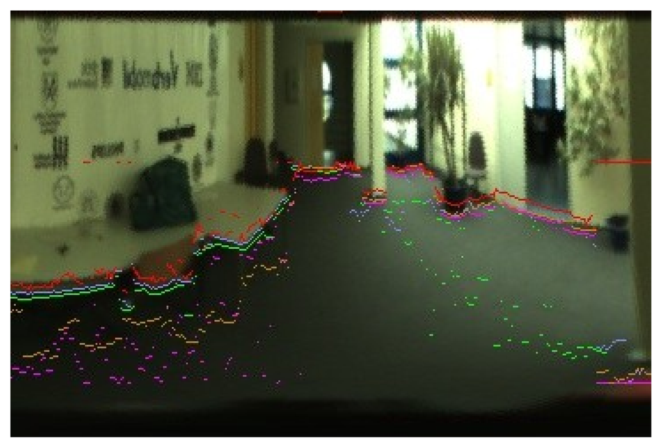
with $\mu^* = \mathrm{E}[r^*] = \mathbf{k}^{*T}\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{r}$

$$\sigma^* = \mathrm{V}[r^*] = k^{**} - \mathbf{k}^{*T}\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{k}^*$$

Training targets

Covariances $k(x^*,x^*)$ $k(x^*,x_i)$ $k(x_i,x_j)$

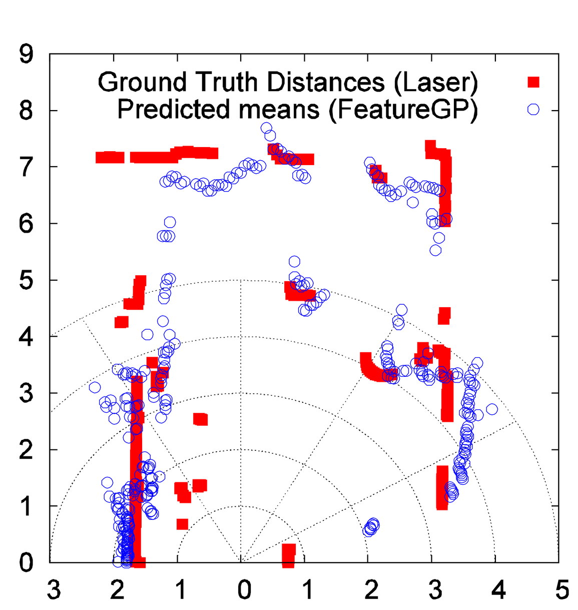
- Example of edge-based features (green, pink, blue, brown) and the GP predictions (red)



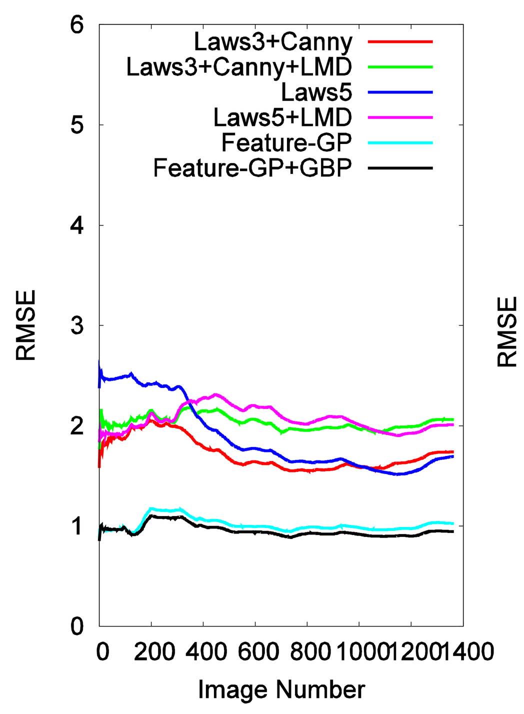
Results

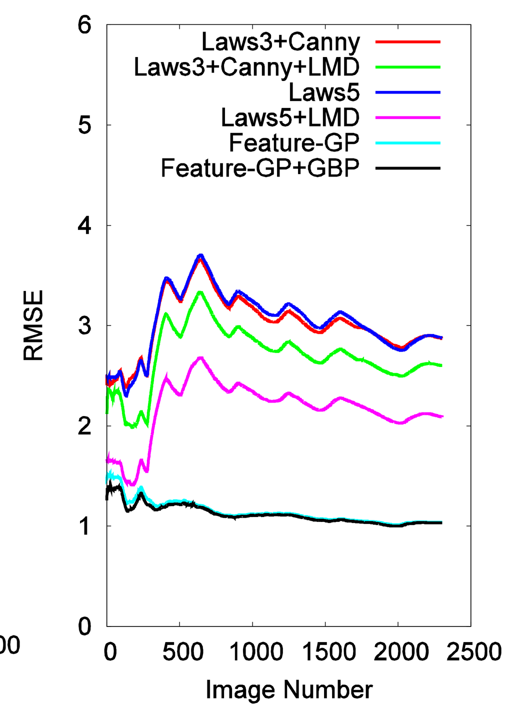
- **Accuracy of range predictions (RMSE on test set)**:

	Saarbrücken	Freiburg
Baseline (Edge-based)	1.70 – 2.06 m	2.08 – 2.87 m
Feature-GP:	1.04 m	1.04 m
Feature-GP + GBP:	1.03 m	0.94 m
LDA-GP + GBP:	1.17 m	1.29 m
PCA-GP + GBP:	1.22 m	1.41 m



Typical 180 degree scan

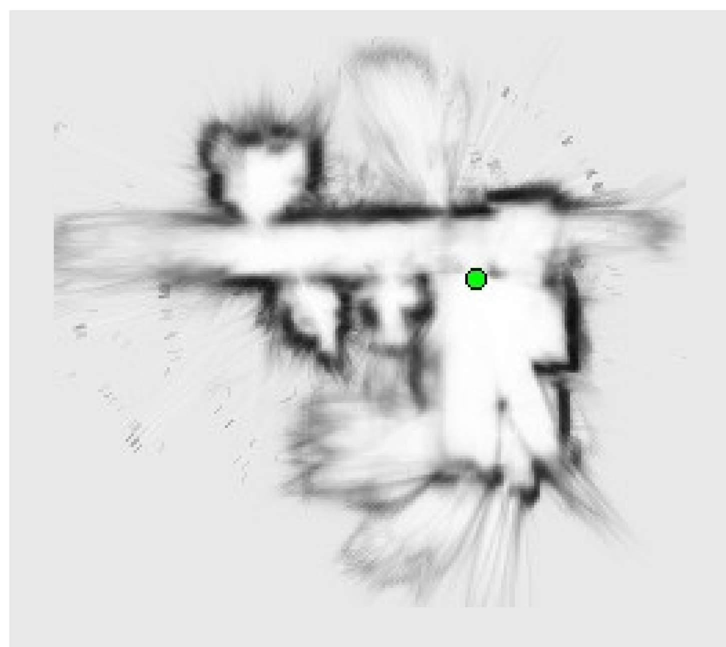




- **Mapping**:
  - Predictive uncertainties of the GP can be used in an extended grid mapping algorithm.
  - Constructed grid maps using the laser sensor directly vs. the GP predictions.

Laser-based grid map

Feature-GP predictions



Conclusions

- Novel approach for predicting range functions from single, monocular camera images.
- Learning framework: Gaussian process regression utilizing edge-based, LDA and PCA-based visual features.
- Accuracy of range predictions is sufficient, e.g., for local obstacle avoidance (comparable to sonar sensor).