Monocular Range Sensing: A Non-Parametric Learning Approach

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

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 P to ranges.
- Pre-processing: Extract features $x_i = f(p_i)$
  - Four types of edge-based features
  - The first six principle components (PCA)

Gaussian Process Regression
- Model all ranges as jointly Gaussian distributed $r_1, \ldots, r_n \mid x_1, \ldots, x_n \sim \mathcal{N}(\mu, \Sigma)$ using a parameterized covariance function, e.g.,
  $$k(x_i, x_j) = \sigma^2 \cdot \exp \left(-\frac{1}{2} \|x_i - x_j\|^2\right)$$
- Then, new ranges can be predicted as
  $$r^* | x^*, D \sim \mathcal{N}(\mu^*, \sigma^*)$$
  with $\mu^* = E[r^*] = k^T \left(K + \sigma_n^2 I\right)^{-1} r$
  $$\sigma^* = \sqrt{\text{Var}[r^*]} = k^{**} - k^T \left(K + \sigma_n^2 I\right)^{-1} k^*$$

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

Results
- Accuracy of range predictions (RMSE on test set):
  - Baseline (Edge-based): 1.70 – 2.06 m
  - Feature-GP: 1.04 m
  - Feature-GP + GBP: 1.03 m
  - LDA-GP + GBP: 1.17 m
  - PCA-GP + GBP: 1.22 m

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.

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).