Visual Terrain Estimation For Legged Robots

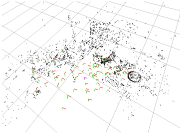
Frederike Dümbgen

Supervisors: Michael Blösch, Philipp Krüsi, Dominik Schindler

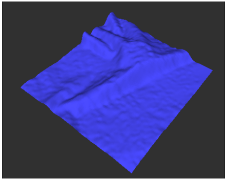
July 26, 2016

Visual SLAM Approaches
Motivation

Sparse SLAM (PTAM, [1])



Dense SLAM (DTAM, [2])



Dense SLAM with B-Splines [3]

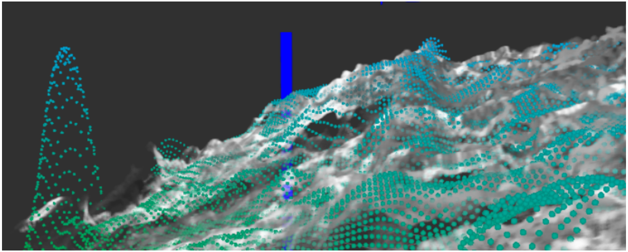
$\Longrightarrow$ **Choice: Dense SLAM with B-Splines**

Visual Terrain Estimation For Legged Robots

Project Goals

Framework for dense stereo camera SLAM

- Static stereo camera $\Longrightarrow$ spline surface (map)
- Map $\Longrightarrow$ new camera position



Visual Terrain Estimation For Legged Robots

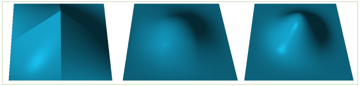
Methods

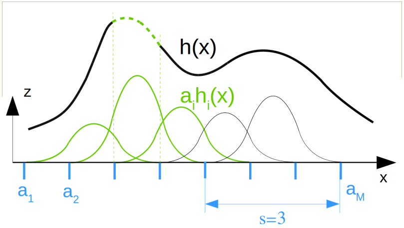
- **Simulation environment**: `ROS` with `rviz` for pointcloud and `opencv` for image handling.
- **Optimization**: own implementation of optimization algorithm using `Eigen` 's sparse matrix solvers.
- **Hardware**: `rovio` sensor for stereo camera data, *MacBook Pro* with Intel Core 2.7MHz, 4 cores.

B-Splines for Surface Representation

Methods

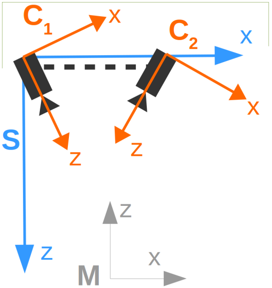
Splines: piecewise polynomial function of degree $<d$.
B(asis)-Splines: Subset of finite-support splines.





$$h(x_j) = \sum_{i=0}^{M} a_i h_i(x_j)$$
$$= \sum_{i=k}^{k+s} a_i h_i(x_j), \text{ for } j = 1 \ldots N$$

Stereo Camera Setup
Methods



Camera poses described by ${}_M\boldsymbol{r}_{MC_k}$ and $\boldsymbol{C}_{C_kM}$ for $k = 1, 2$ or

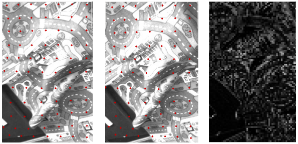
$$\boldsymbol{\xi}_{C_k} := \left[{}_M\boldsymbol{r}_{MC_k}, \boldsymbol{\Phi}_{C_kM}\right]^T$$

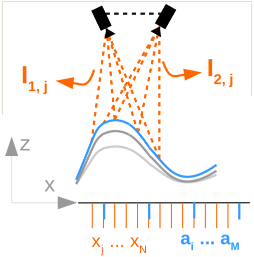
$$\boldsymbol{\xi}_S := \left[{}_M\boldsymbol{r}_{MS}, \boldsymbol{\Phi}_{SM}\right]^T$$

with $\boldsymbol{\Phi}_{C_kM}$, $\boldsymbol{\Phi}_{SM} \in \mathbb{R}^3$ [4]

Relative position of $\{\boldsymbol{\xi}_{C_1}, \boldsymbol{\xi}_{C_2}, \boldsymbol{\xi}_S\}$ stays fixed.

Photometric errors for mapping
Methods





Photometric error of grid point x_j, y_j:

$$r_j(a) = I_1(\boldsymbol{u}_{j,1}(a)) - I_2(\boldsymbol{u}_{j,2}(a)) ,$$

with I_1, I_2 interpolated intensities at the locations $\boldsymbol{u}_{j,k}$ in camera $k = 1$ and $k = 2$.

$$\boldsymbol{u}_{j,k} = \boldsymbol{K}_k D_k(\boldsymbol{T}_k({}_M\boldsymbol{r}_{MX_j})))$$

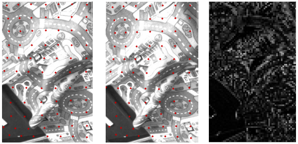
$$\boldsymbol{T}_k({}_M\boldsymbol{r}_{MX_j}) = \pi(\boldsymbol{C}_{C_k M}({}_M\boldsymbol{r}_{MX_j} - {}_M\boldsymbol{r}_{MC_k}))$$

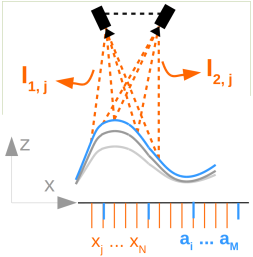
3D point given by spline map:

$${}_M\boldsymbol{r}_{MX_j} = \left[x_j, y_j, h(x_j, y_j)(a)\right]^T$$

Photometric errors for mapping

Methods





$$J_r(\boldsymbol{a}) = \frac{\partial \boldsymbol{r}(\boldsymbol{a})}{\partial \boldsymbol{a}} \in \mathbb{R}^{NxM}$$

$$\begin{aligned} J_r(\boldsymbol{a}) = (&J_{pixel,1} J_{camera,1}({}_M\boldsymbol{r}_{MX_j}) \\ &- J_{pixel,2} J_{camera,2}({}_M\boldsymbol{r}_{MX_j})) J_{splines} \ . \end{aligned}$$

$$J_{pixel,k} = \frac{\partial I_k(\boldsymbol{u}_k)}{\partial \tilde{\boldsymbol{u}}_k}, \quad J_{camera,k}({}_M\boldsymbol{r}_{MX_j}) = \frac{\partial \tilde{\boldsymbol{u}}_k}{\partial {}_M\boldsymbol{r}_{MX_j}}$$

$$J_{splines} = \frac{\partial {}_M\boldsymbol{r}_{MX_j}}{\partial \boldsymbol{a}}$$

Optimization problem for mapping
Methods

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}\in\mathbb{R}^M} f(\boldsymbol{a}) = \arg\min_{\boldsymbol{a}\in\mathbb{R}^M} \frac{1}{2}\Big(\sum_{j=0}^{N} w_j r_j(\boldsymbol{a})^2 + \beta \boldsymbol{a}^T \boldsymbol{B}\boldsymbol{a} + \gamma \boldsymbol{a}^T \boldsymbol{G}\boldsymbol{a}\Big) ,$$

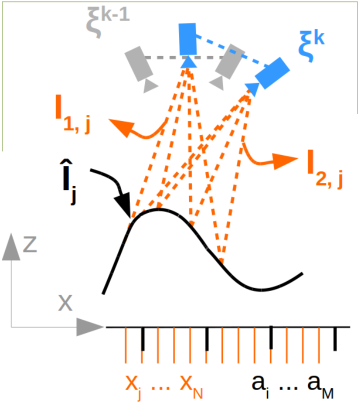
with

- bending and gradient energy regularization terms and
- weight representing the average visibility of point j.

Solved using Gauss-Newton iterations:

$$\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \mathbf{p}_k^{GN}$$
$$\boldsymbol{J}_r(\boldsymbol{a})^T \boldsymbol{J}_r(\boldsymbol{a})\mathbf{p}_k^{GN} = -\, \boldsymbol{J}_r(\boldsymbol{a})^T \mathbf{r}(\boldsymbol{a})$$

Photometric errors for localization
Methods

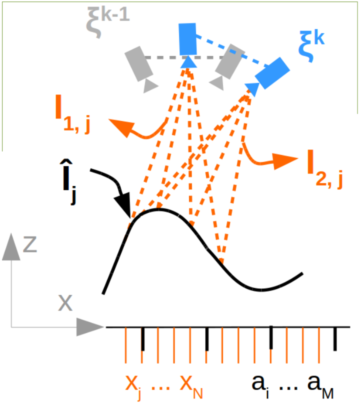


Assume fixed intensities $\hat{I}$ at $\left[x_j, y_j\right]$ from previous mapping step.

$$r_{j,1}(\xi) = I_1(\boldsymbol{u}_{j,1}(\xi)) - \hat{I}(x_j, y_j)$$
$$r_{j,2}(\xi) = I_2(\boldsymbol{u}_{j,2}(\xi)) - \hat{I}(x_j, y_j)$$

Photometric errors for localization

Methods



$$J_r(\xi) = \frac{\partial \boldsymbol{r}(\xi)}{\partial \xi} \in \mathbb{R}^{N\times 6}$$

$$J_r(\xi) = J_{pixel} J_{camera}(\xi)$$

$$J_{pixel} = \frac{\partial I(\boldsymbol{u})}{\partial \tilde{\boldsymbol{u}}}, \quad J_{camera}(\xi) = \frac{\partial \tilde{\boldsymbol{u}}}{\partial \xi}$$

Optimization problem for localization
Methods

$$\hat{\boldsymbol{\xi}} = \underset{\hat{\boldsymbol{\xi}} \in \mathbb{R}^6}{\arg\min} \frac{1}{2} \sum_{j=0}^{N} r_j(\boldsymbol{\xi})^2$$

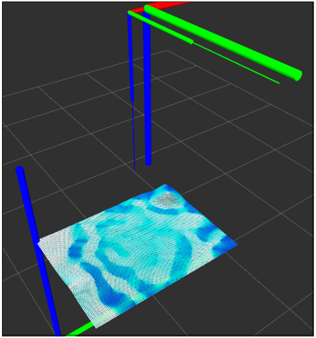
Solved using Gauss-Newton iterations:

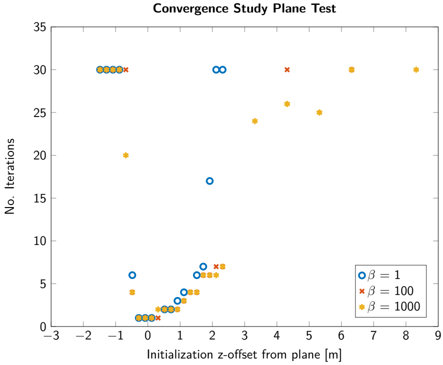
$$\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi}_k \boxplus \mathbf{p}_k^{GN}$$
$$\boldsymbol{J}_r(\boldsymbol{\xi})^T \boldsymbol{J}_r(\boldsymbol{\xi}) \mathbf{p}_k^{GN} = - \boldsymbol{J}_r(\boldsymbol{\xi})^T \mathbf{r}_k(\boldsymbol{\xi})$$

Overview Results
Mapping And Localization Datasets

Dataset	Plane test	Middlebury [5]	Inhouse
Ground truth	analytical	structured light	pattern matching
Images	rectified	rectified	non rectified
Calibration	+++	++	+
Mapping	yes	yes	yes
Localization	yes	no	no
Map dimensions [m]	0.9 x 1.2	1.5 x 2.0	
Spline resolution	20 x 20	75 x 100	
Residuals resolution	90 x 120		

Plane Test
Mapping Results

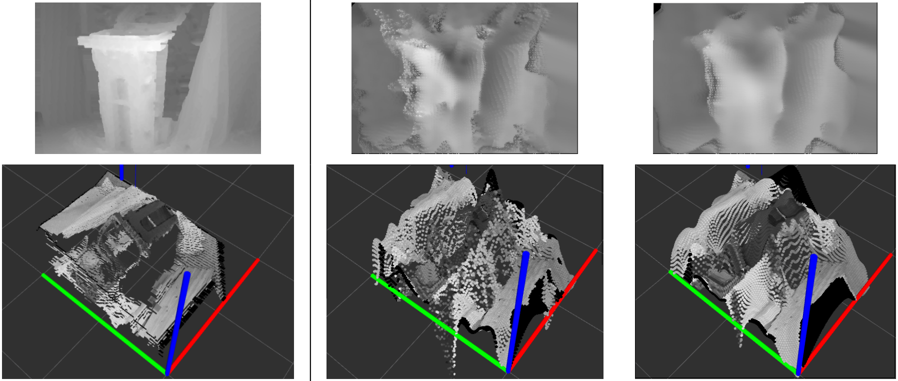




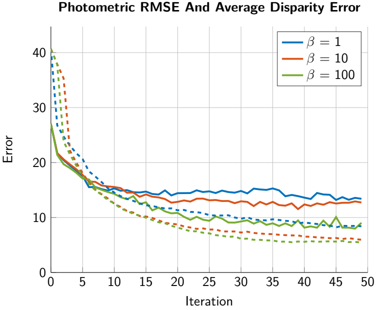
Middlebury Dataset

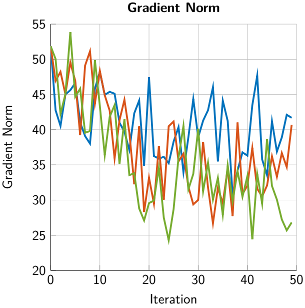
Mapping Results

groundtruth | $\beta = 1$ | $\beta = 10$



Middlebury Dataset
Mapping Results

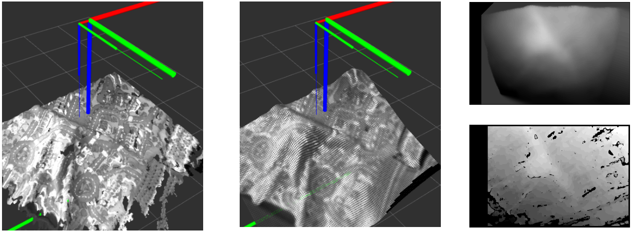




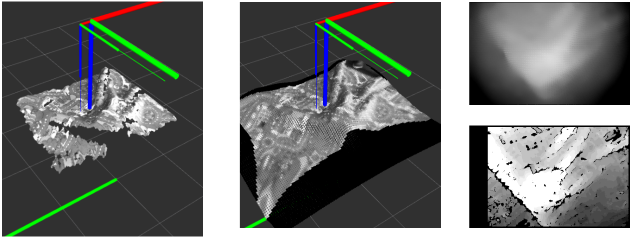
Inhouse Dataset

Mapping Results

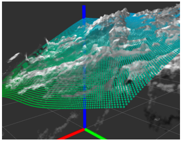
"far "
$\beta = 10$
$\gamma = 1e6$

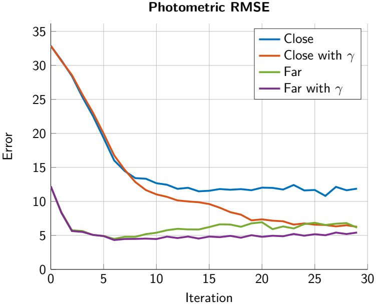


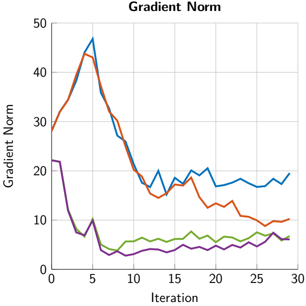
"close"
$\beta = 10$
$\gamma = 1e5$



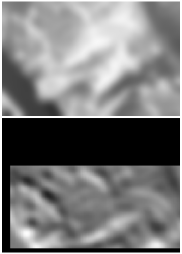
Inhouse Dataset
Mapping Results

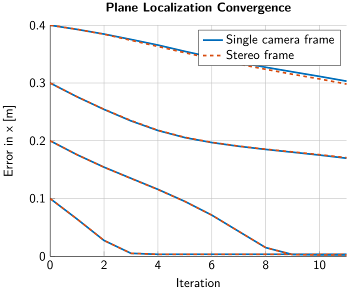


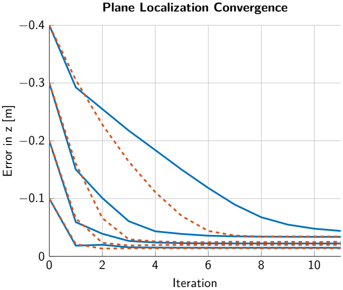




Plane Test
Localization Results







Achievements

- Created versatile `stereo surface reconstruction` package for
  - variable spline degrees and resolution,
  - entirely customizable optimization parameters and
  - rectified and unrectified images.
- Implemented photometric localization algorithm based on one stereo measurement.
- Tested functionalities on real and simulated datasets.

Future Work

- Implement sequence of mapping and localization steps to improve map accuracy by solving recursively over multiple measurements.
- Integrate measurements to extend map and create wider camera baseline.
- Test framework in realistic sceneries.

Questions?

References I

- Georg Klein and David Murray.
  Parallel tracking and mapping for small AR workspaces.
  In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pages 225–234. IEEE, 2007.
- Richard A Newcombe, Steven J Lovegrove, and Andrew J Davison.
  DTAM: Dense tracking and mapping in real-time.
  In *2011 international conference on computer vision*, pages 2320–2327. IEEE, 2011.
- Dominik Schindler.
  Visual Terrain Estimation for Legged Robots.
  Master Thesis, 2015.

References II

- Michael Bloesch, Hannes Sommer, Tristan Laidlow, Michael Burri, Gabriel Nuetzi, and R O Jun.

  A Primer on the Differential Calculus of 3D Orientations.

  2016.

- Daniel Scharstein, Heiko Hirschmüller, York Kitajima, Greg Krathwohl, Nera Nešić, Xi Wang, and Porter Westling.

  High-resolution stereo datasets with subpixel-accurate ground truth.

  In *German Conference on Pattern Recognition*, pages 31–42. Springer International Publishing, 2014.