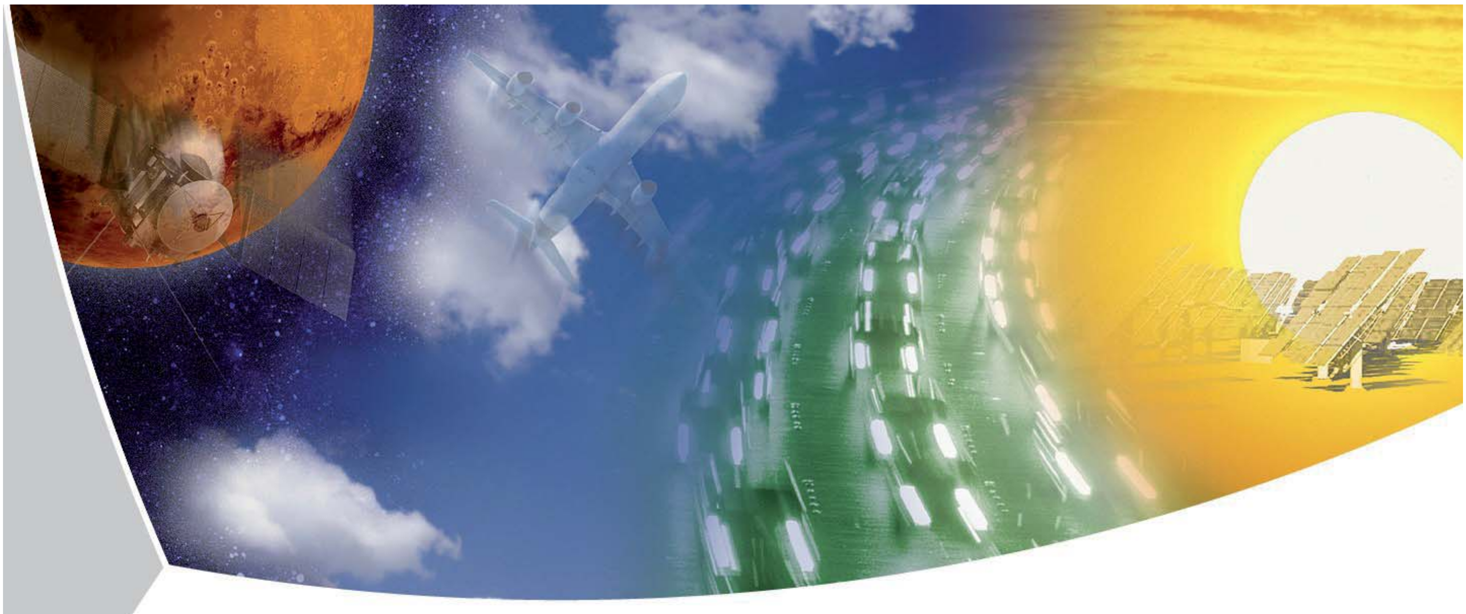


Semi-Global Matching
Motivation, Developments and Applications

Dr. Heiko Hirschmüller

Department of Perception and Cognition
Institute of Robotics and Mechatronics

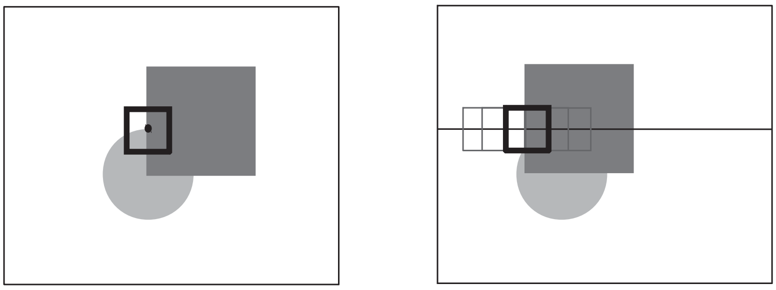
Deutsches Zentrum für Luft- und Raumfahrt e.V. in der Helmholtz-Gemeinschaft

Content

- Motivation
- Semi-Global Matching
- CPU, CPU-Cluster, GPU and FPGA Implementations
- Applications

Motivation

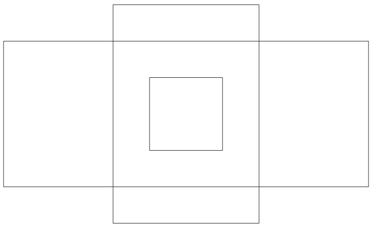
Local Stereo Matching by Correlation



- Matching of (square) windows
- Implicit assumption about constant depth/disparity in the window
- Disparity changes lead to errors
- Results in blurred object borders and loss of small objects
- Foreground objects are mostly fattened (Hirschmüller et al., IJCV 2002)

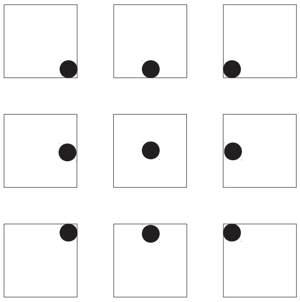
Improving Matching Discontinuities and Slanted surfaces

Algorithmic testing of different shapes and sizes



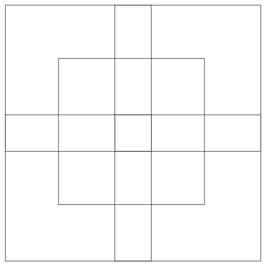
Kanade and Okutomi, TPAMI 1994

Windows with shifted point of interest



Fusiello et al., CVPR 1997

Combination of smaller windows



Hirschmüller et al., IJCV 2002

- Affine transformation for adapting to slanted surfaces (Grün, 1985)
- ...
- All options increase degrees of freedom (more computation time)
- Not all cases are treated at the same time

Results of Correlation

Small Vision System (SVS)
Konolige (RR 1997)
O(w*h*d) [49 ms]

MWMF
Hirschmüller et al. (IJCV 2002)
O(w*h*d) [63 ms]

Global Stereo Matching

- Pixelwise matching (avoiding assumption of locally constant depth)

$$C(\mathbf{p}, D_{\mathbf{p}}) = |L(\mathbf{p}) - R(\mathbf{p}-D_{\mathbf{p}})|$$

- Pixelwise matching is not unique
- Therefore, adding cost for disparity steps (q and p are neighbors)

$$P\ T[D_{\mathbf{p}} \neq D_{\mathbf{q}}]$$

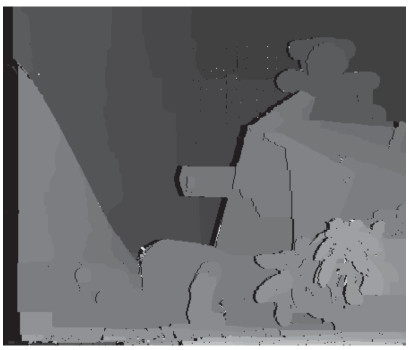
- Adding pixelwise matching and smoothness cost for all pixels

$$E(D)=\sum_{\mathbf{p}}(C(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q} \in N\mathbf{p}} P\ T[D_{\mathbf{p}} \neq D_{\mathbf{q}}])$$

- Goal: Finding the disparity image D that minimizes E(D)
- Smoothness cost connects all pixels (globally)
- Minimization is NP hard (Boykov et al., TPAMI 2001)

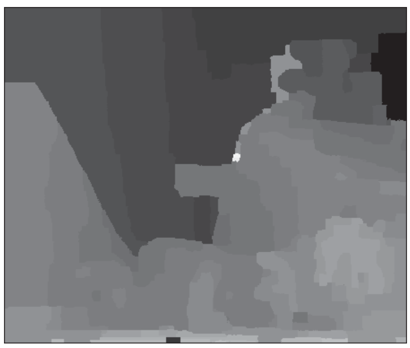
Approximation of Global Minimization

Graph Cuts (Kolmogorov and Zabih, ICCV 2001)



O(?) [57s]

Efficient Belief Propagation (Felzenszwalb and Huttenlocher, CVPR 2004)

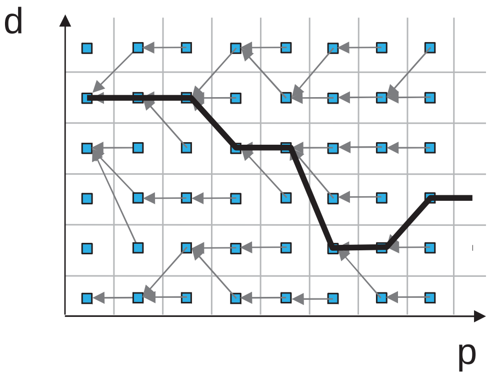


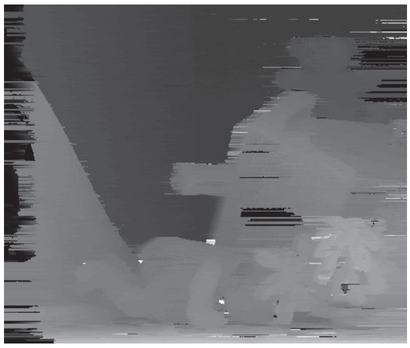
O(w*h*d) [4.5s]

- Problem: Complex, e.g. long run-time
- Approximation, i.e. not optimal solution

Exact Minimization, but only along Scanlines

- Step by step minimization in 1D by $E_{\mathbf{p},d} = C_{\mathbf{p},d} + \min_{d'} (E_{\mathbf{p}',d'} + P\ T[d \neq d'])$
- Can be computed very efficiently by *dynamic programming*





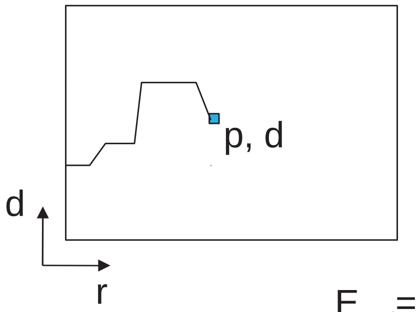
O(w*h*d²) oder O(w*h*d)

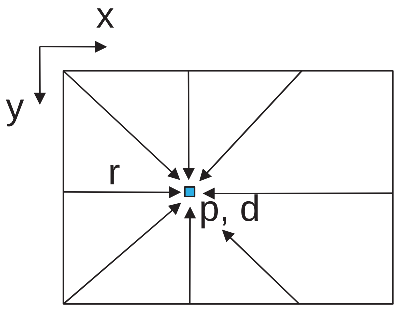
- Problem: Streaking effects
- Many papers in literature about reducing this effect

Semi-Global Matching (SGM)

Semi-Global Matching

$$E(D)=\sum_{\mathbf{p}}(C_{BT}(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q}\in N\mathbf{p}}(P_1\ T[|D_{\mathbf{p}}-D_{\mathbf{q}}|=1] + P_2\ T[|D_{\mathbf{p}}-D_{\mathbf{q}}|>1]))$$





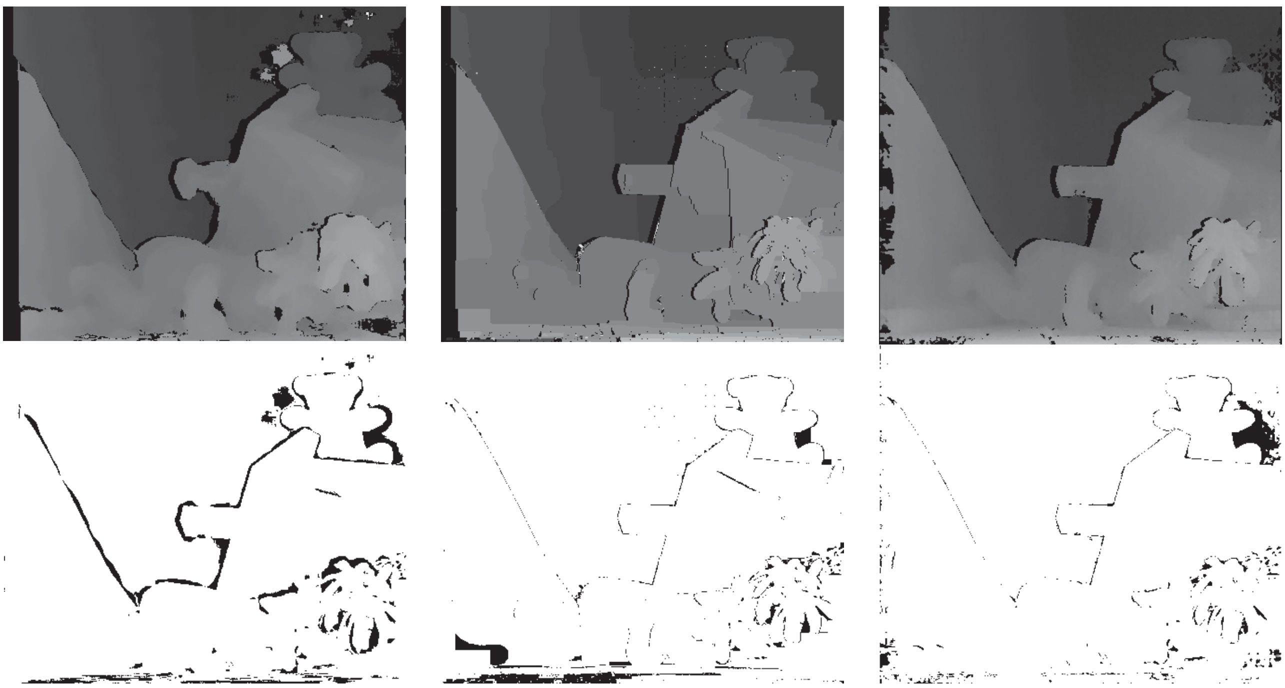
$$E_{\mathbf{p},d}= \sum_r E^r_{\mathbf{p},d}$$

O(w*h*d) [1.06s]

- Disparity of **p** becomes $\min_d E_{\mathbf{p},d}$
- Additionally consistency check, sub-pixel interpolation, filtering, ...
- Hirschmüller (CVPR 2005 & TPAMI 2008)

Overview of Results

MWMF O(w*h*d) [63 ms] | GC O(?) [57 s] | SGM O(w*h*d) [1.06 s]



Matching Cost

- Images have different radiometry due to:
  - Different global and local imaging characteristics (e.g. exposure, vignetting, etc.)
  - Perceived intensity depends on viewpoint (i.e. in realty no pure diffuse reflection)
  - Lighting can be different if there is a time delay between image capture
  - Noise
- Study showed best results with Mutual Information and Census (Hirschmüller und Scharstein, CVPR 2007 & TPAMI 2009)

CPU, Cluster, GPU and FPGA Implementations

Processing Cluster at DLR-RM

- For aerial and satellite image processing
- 3 blade systems with a total of 296 CPU cores
- Each CPU core has 3 GB RAM
- Central files system with a capacity of 55 TB, distributed over 3 RAID systems and accessed through 4 server computer in parallel
- 200 TB archive

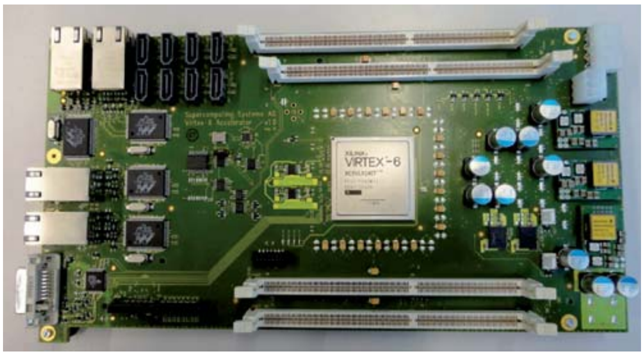
- More than 100 TB of images processed since 2005
- More than 14 TB DSM and ortho images created
- Current yearly throughput is about 50 TB (aerial images)

GPU Implementation

- First implementation in OpenGL/Cg
- Runs on consumer grade graphics cards, e.g. GeForce GTX 275
- Runs with 4.5 Hz (Image size 640 x 480 with 128 pixel disparity range)
- Up to 1024 x 1024 pixel with 200 pixel disparity range possible
- Memory efficient SGM works with up to 2048 x 2048 pixel and 1024 pixel disparity range

- Ernst and Hirschmüller (ISVC 2008)

FPGA Implementation of SGM for Huge Images

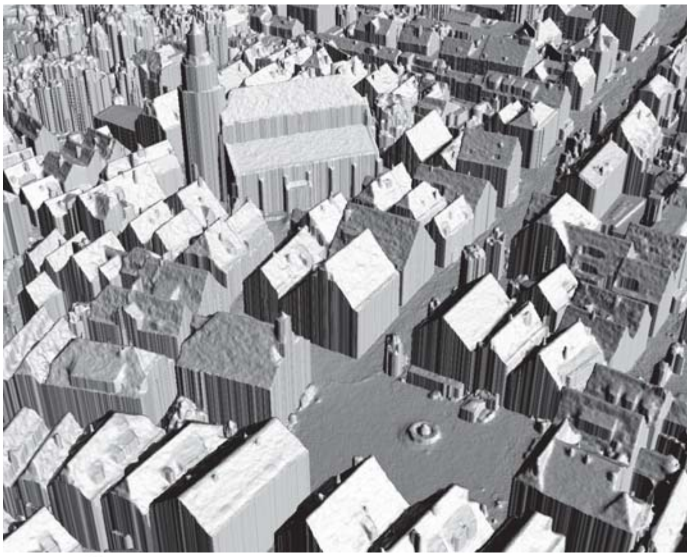


- FPGA accelerator for SGM
- Communication to PC via GigE
- Possibility for creating network of PC's and accelerator boxes
- Input images are 16 bit
- Output disparity is 16 bit with 4 bit for subpixel accuracy
- Image size: 512 x 512 - 2048 x 2048 pixels in steps of 8 pixels
- Right image is extended by disparity range for tilewise matching of arbitrarily large images
- Disparity range: 32 - 4096 pixels in steps of 32 pixels
- Runtime for 2048 x 2048 x 1024 is 4.6 s for matching twice (incl. LR-check)
- Runtime on one Intel X5570 Server CPU core (3 GHz) is 41 times higher
- Total power consumption of FPGA: 54 W (including power supply and fan)
- Implemented by Supercomputing Systems (Zürich) on our request

Applications

Processing of Aerial Pinhole Images

- DGPF UltraCam-X dataset, GSD 8 cm/pixel
- DSM always has GSD of images, i.e. 8 cm/pixel as well!
- Automatically fully textured from aerial images
- Hirschmüller and Bucher (DGPF, 2010)

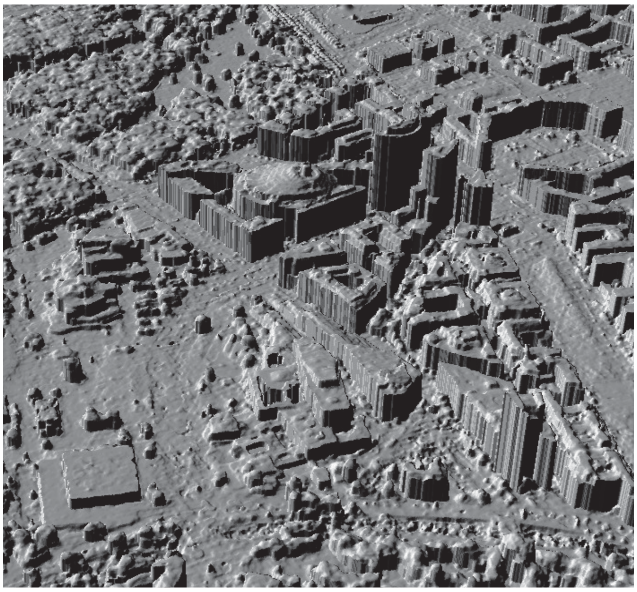


Processing of Aerial Pushbroom Images

- Zugspitze, HRSC-AX (DLR), 20 cm/pixel (Image and DSM resolution)
- Hirschmüller et al. (DAGM 2005)
- Interactive 3D maps (www.reality-maps.de) are based on SGM data sets

Processing of Satellite Images (1)

- Berlin, World View, 0.5 m/pixel (Image and DSM resolution)
- Dataset provided by Digital Globe, image registration by DLR RM-OS (Berlin)



Deutsches Zentrum für Luft- und Raumfahrt e.V. in der Helmholtz-Gemeinschaft

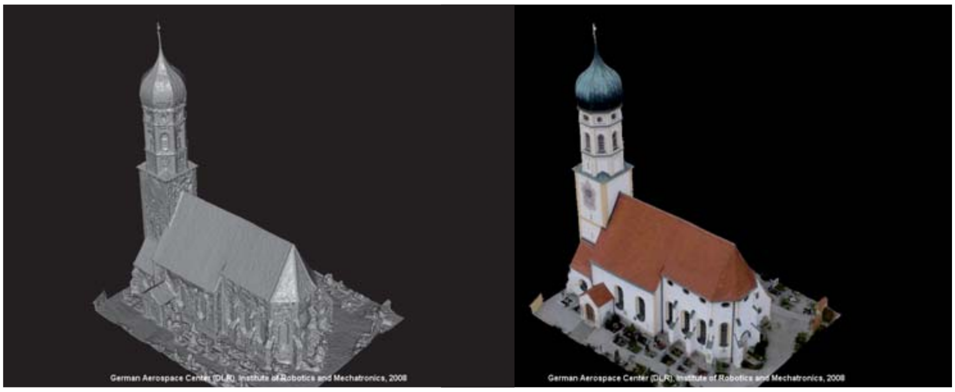
Processing of Satellite Images (2)

- Mt. Everest, World View, 0.5 m/pixel (Image and DSM resolution)
- Almost 5000 m height difference in the data set!
- Dataset provided by Digital Globe, image registration by DLR RM-OS (Berlin)
- See interactive 3D visualization at www.everest3d.de

Deutsches Zentrum für Luft- und Raumfahrt e.V. in der Helmholtz-Gemeinschaft

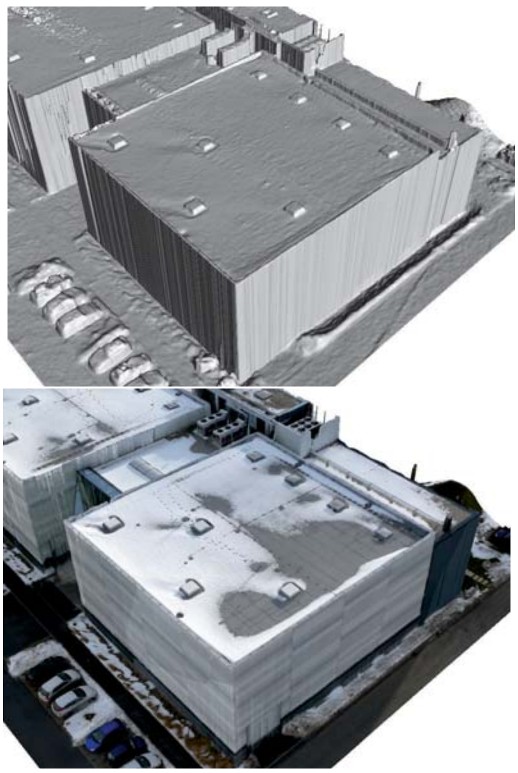
3D Reconstruction from UAVs

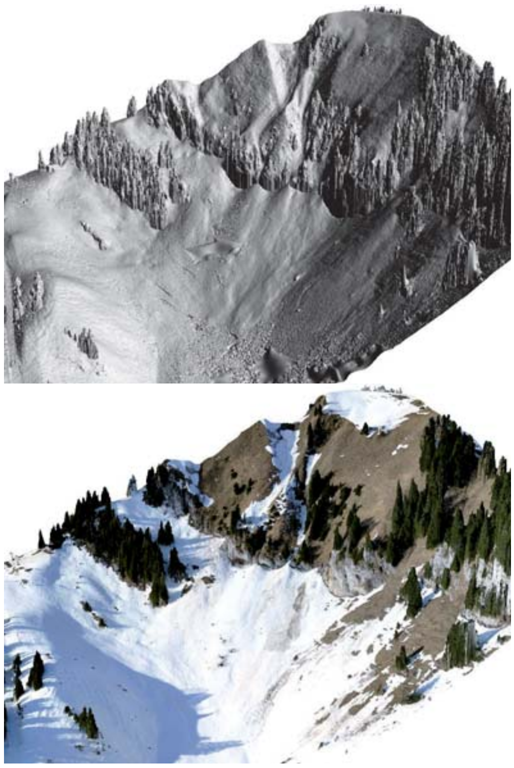
- UAV: AscTec Falcon 8
- Camera: Panasonic Lumix (10 MPixel)
- 60 Images taken in 360 degree cycle (autonomous flight using GPS)
- Registration done by Bundler (Snavely et al., IJCV 2008)
- SGM matching and 3D Reconstruction with 3 cm/pixel



Reconstruction from UAVs

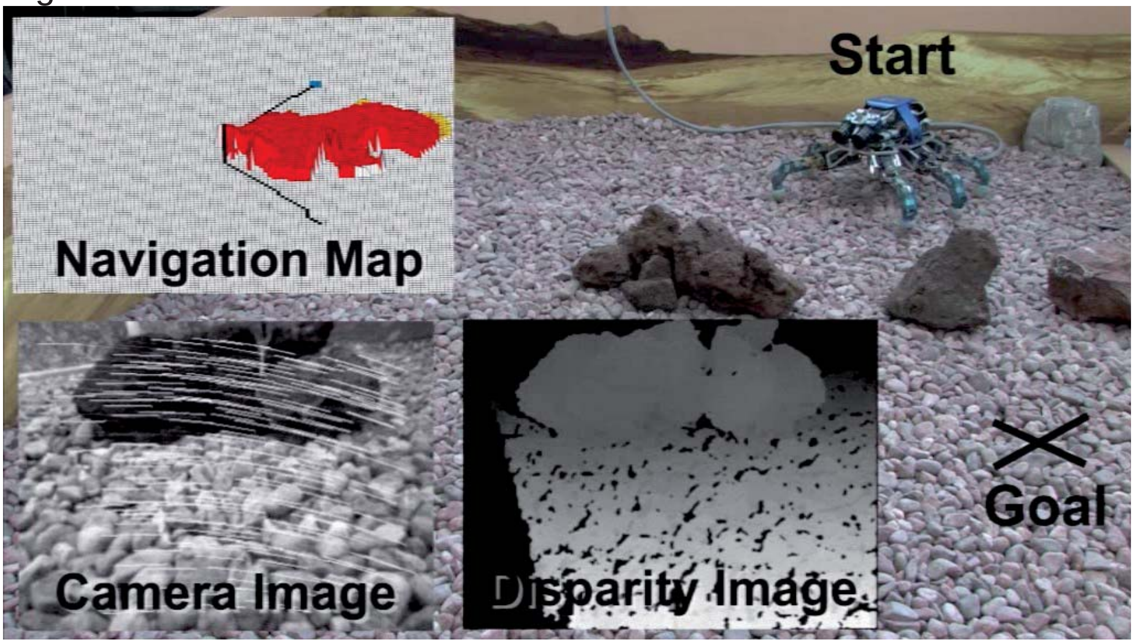
- Autonomous flight planing and reconstruction: Schmid et al. (ICUAS 2011)





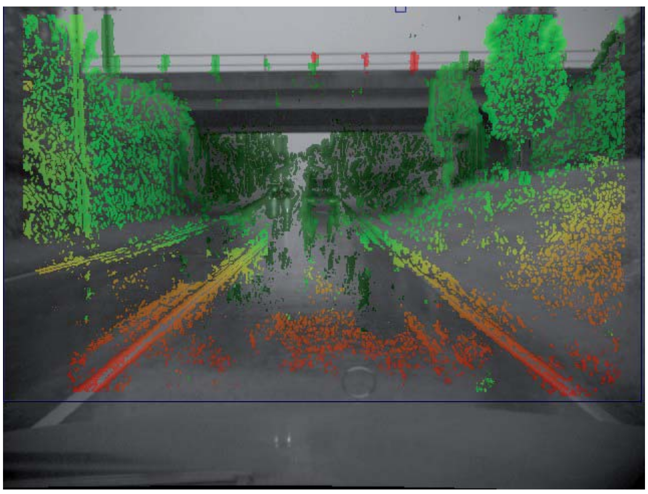
Navigation of Mobile Robots

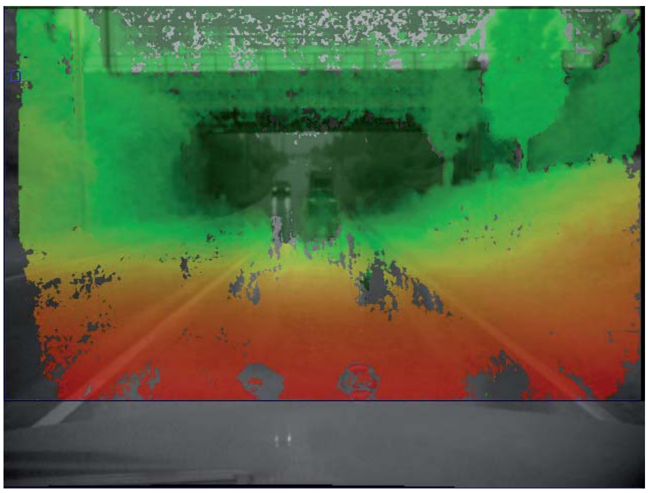
- Navigation on rough terrain with Crawler (Chilian and Hirschmüller, IROS 2009)
- SGM on GPU (Ernst and Hirschmüller, ISVC 2008)
- Visual odometry, sensor fusion, 2.5 D mapping, traversability estimation and path planning in real time



Driver Assistance Systems

- Cooperation with Daimler AG regarding SGM
- Daimler has FPGA implementation of SGM with 27 Hz (Gehrig et al., ICVS 2009)
- Used by Daimler researchers as part of 6D-Vision system (www.6d-vision.com)


Correlation based Stereo


Semi-global matching

The images are courtesy of Stefan Gehrig (Daimler AG)

Acknowledgments

- Colleagues at DLR Institute of Robotics and Mechatronics (Oberpfaffenhofen and Berlin) and DLR Institute of Planetary Research (Berlin)
- DGPF & Vexcel Imaging Graz for UltraCam-X data set of Vaihingen/Enz
- Digital Globe for World View data sets of Berlin and Mt. Everest
- Reality Maps (www.reality-maps.de)
- Supercomputing Systems (www.scs.ch)
- Stefan Gehrig and Uwe Franke, Daimler AG (www.6d-vision.com)