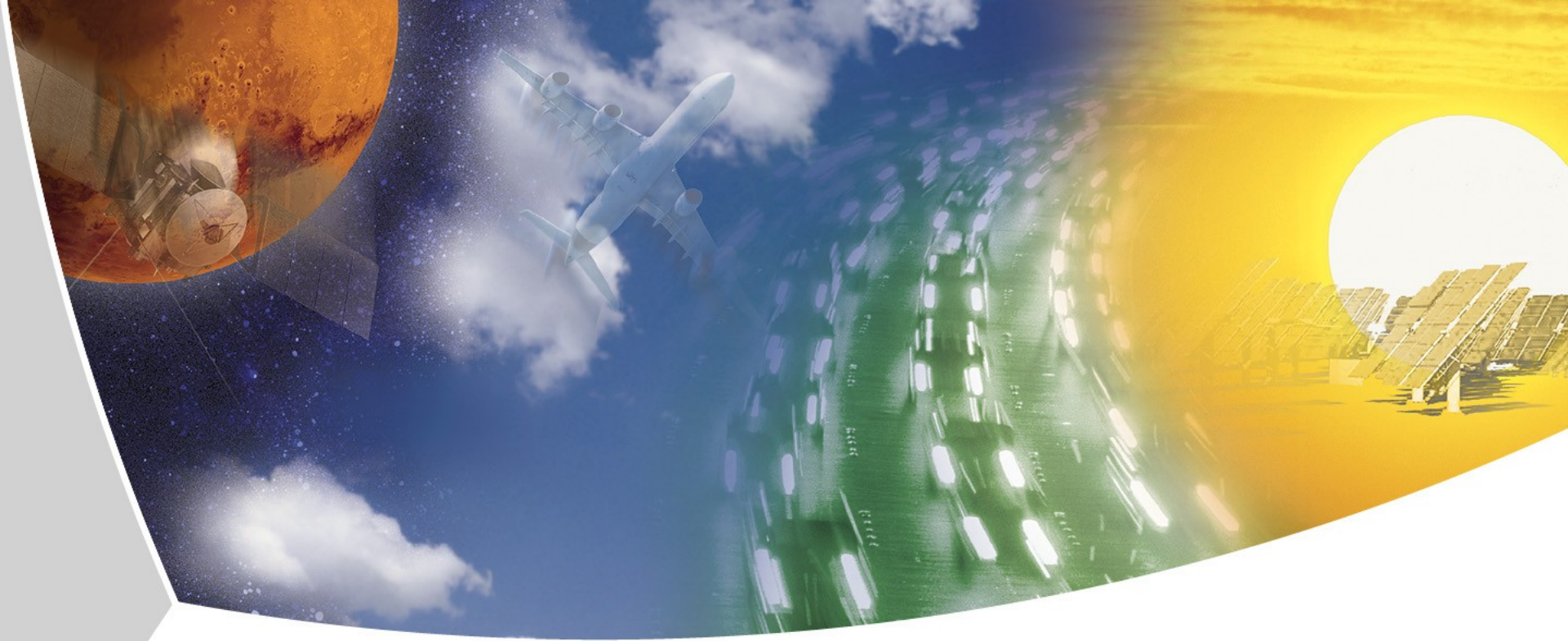




Semi-Global Matching

Dr. Heiko Hirschmüller

Department of Perception and Cognition
Institute of Robotics and Mechatronics



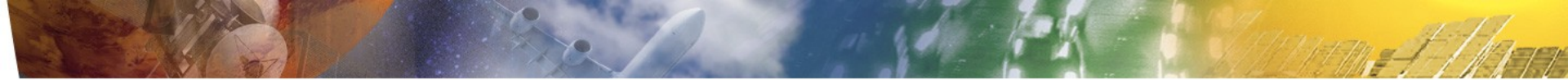
Deutsches Zentrum
für Luft- und Raumfahrt e.V.
in der Helmholtz-Gemeinschaft



Content

- Introduction
- Semi-Global Matching
- Implementations
- Applications

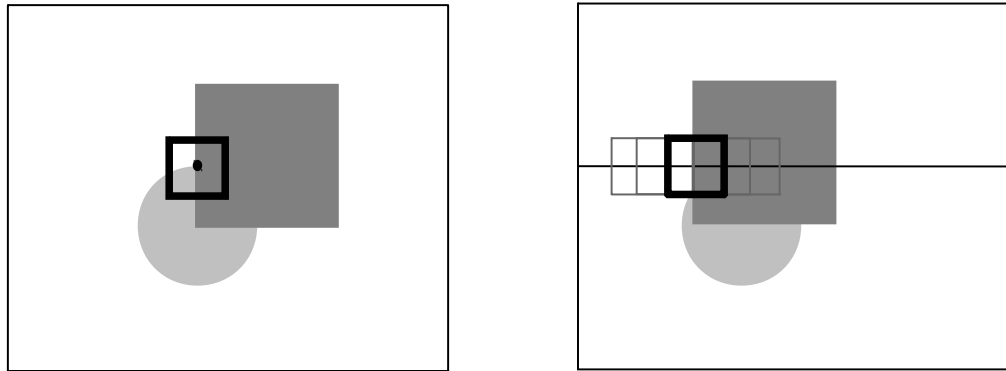




Introduction



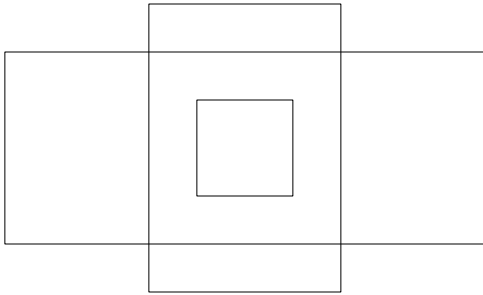
Local Stereo Matching by Correlation



- 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)

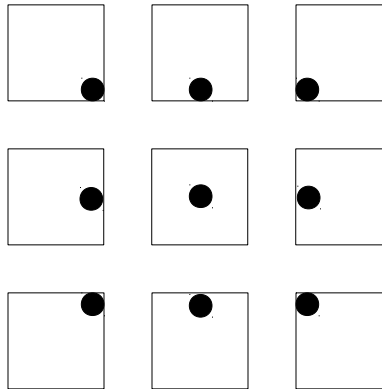
Improving Matching Discontinuities and Slanted surfaces

Algorithmic testing of different shapes and sizes



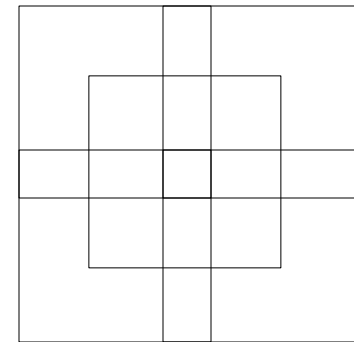
Kanade and Okutomi,
TPAMI 1994

Windows with shifted point of interest



Fusiello et al.,
CVPR 1997

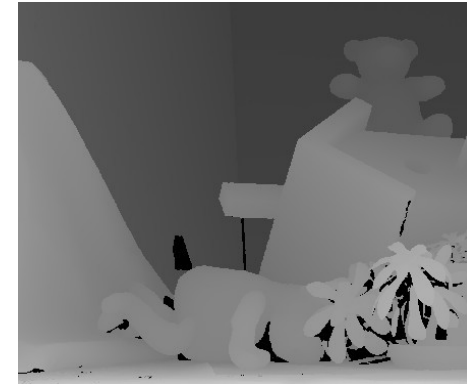
Combination of smaller windows



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

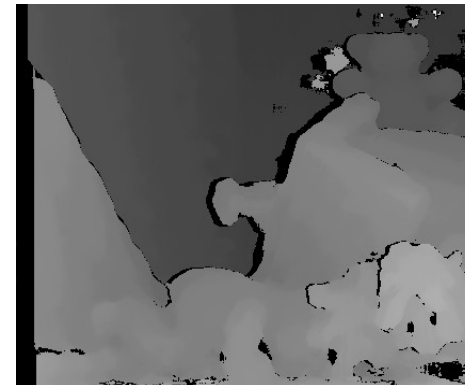
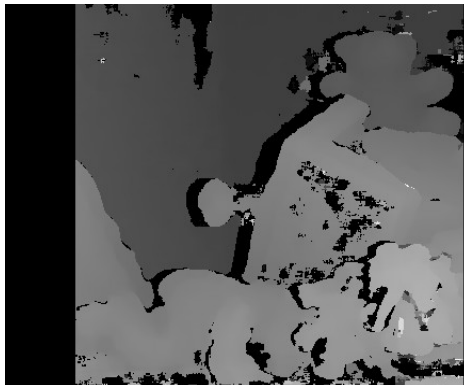
Results of Correlation



Images with dense ground truth (source: vision.middlebury.edu/stereo)

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

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





Global Stereo Matching

- Pixelwise matching (avoiding assumption of locally constant depth)

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

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

$$P T[D_p \neq D_q]$$

- Adding pixelwise matching and smoothness cost for all pixels

$$E(D) = \sum_p (C(\mathbf{p}, D_p) + \sum_{q \in N_p} P T[D_p \neq D_q])$$

- 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)

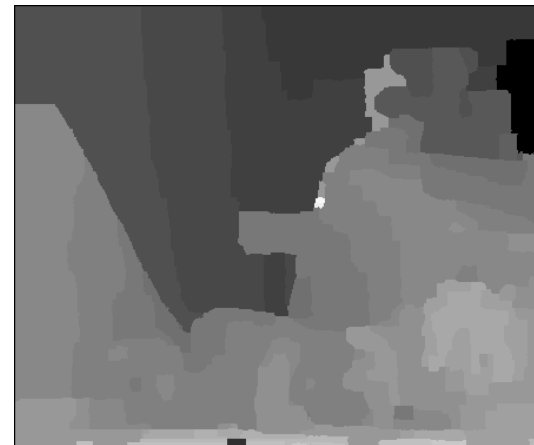
Approximation of Global Minimization

Graph Cuts
(Kolmogorov and Zabih,
ICCV 2001)



$O(?)$ [57s]

Efficient Belief Propagation
(Felzenszwalb and Huttenlocher,
CVPR 2004)

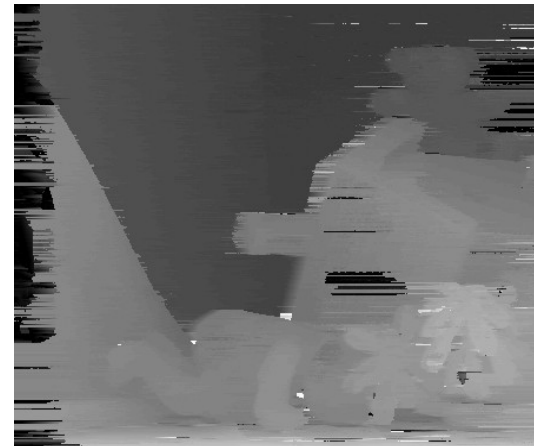
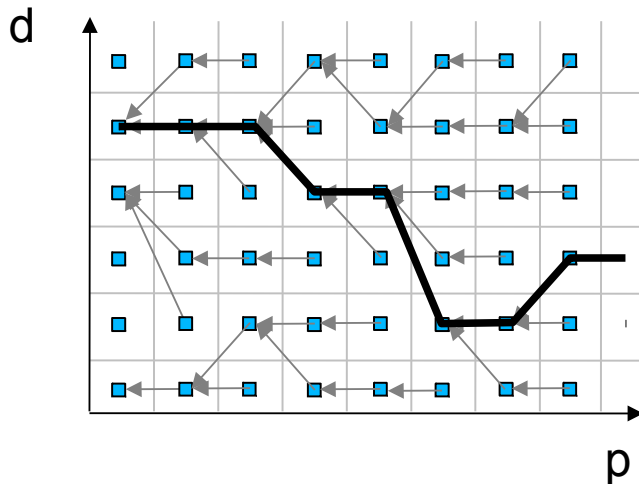


$O(w \cdot h \cdot d)$ [4.5s]

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

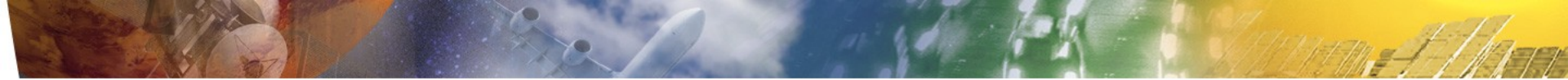
Exact Minimization, but only along Scanlines

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



$O(w \cdot h \cdot d^2)$ oder $O(w \cdot h \cdot d)$

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

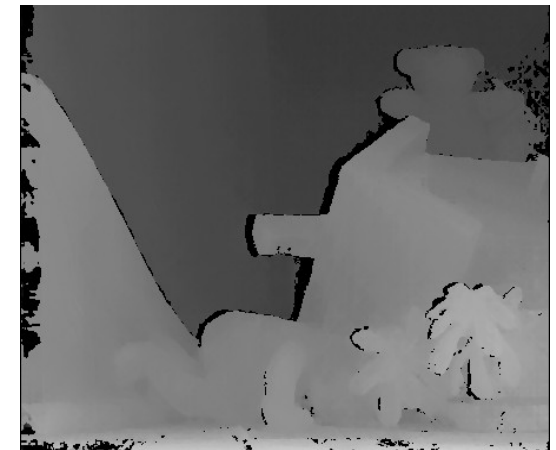
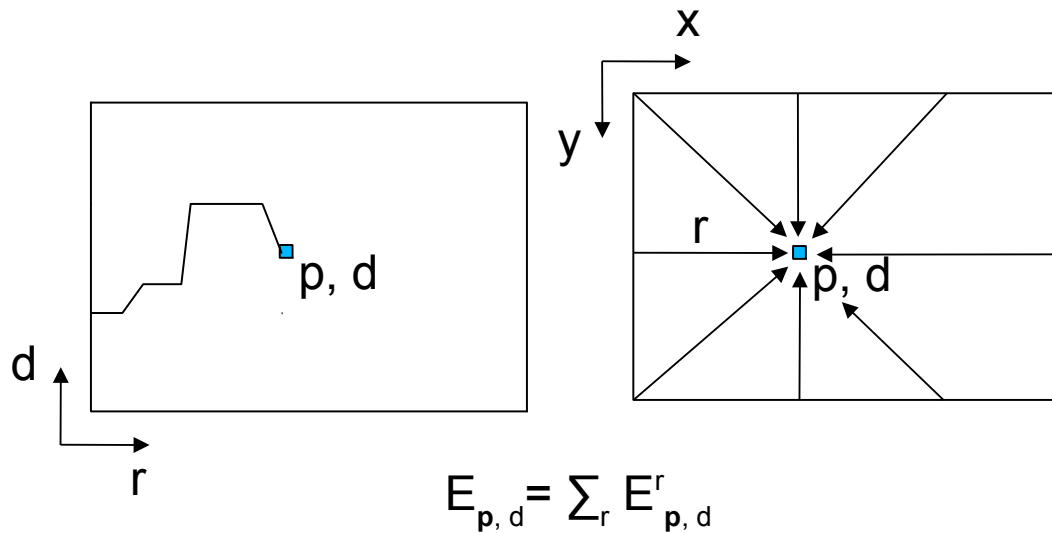


Semi-Global Matching (SGM)



Semi-Global Matching

$$E(D) = \sum_p (C_{BT}(p, D_p) + \sum_{q \in N_p} (P_1 T[|D_p - D_q| = 1] + P_2 T[|D_p - D_q| > 1]))$$

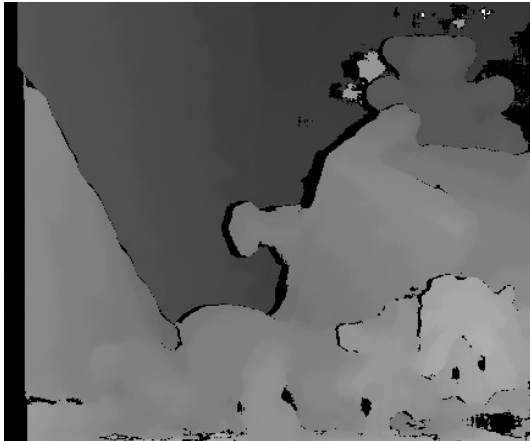


$O(w \cdot h \cdot d)$ [1.06s]

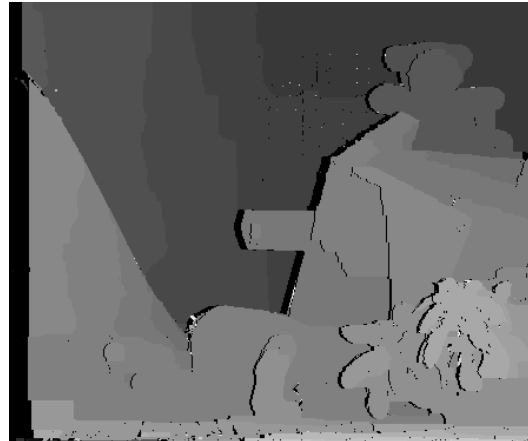
- Disparity of p becomes $\min_d E_{p,d}$
- Additionally consistency check, sub-pixel interpolation, filtering, ...
- Hirschmüller (CVPR 2005 & TPAMI 2008)

Overview of Results

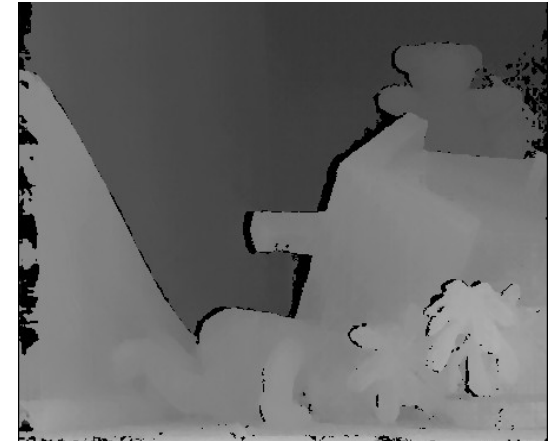
MWMF $O(w \cdot h \cdot d)$ [63 ms]



GC $O(?)$ [57 s]



SGM $O(w \cdot h \cdot d)$ [1.06 s]



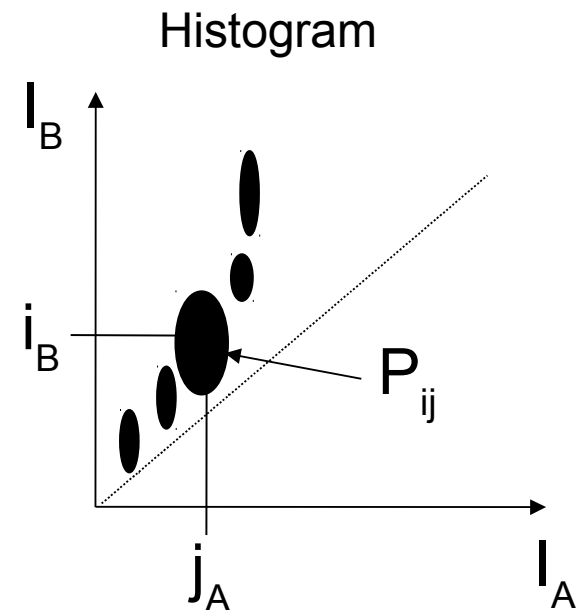


Matching Cost

- 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 reality 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)

Mutual Information (Viola und Wells, IJCV 1997)

- Collecting corresponding intensities (from known disparities) into a histogram
- Computing the probability of each histogram cell
- Few cells with high probability => good matching
- Expressed as entropy: $\int P_{ij} \log(P_{ij}) di dj$
- Hierarchical computation for speed (Hirschmüller, TPAMI 2008)
- Strength:
 - Truly pixelwise matching
 - Models complex radiometric relationships
 - Compensates noise very well
- Weakness:
 - Models global radiometric relationship only



SGM Stereomatching using Hierarchical MI

Left (original)



Right (manipulated)



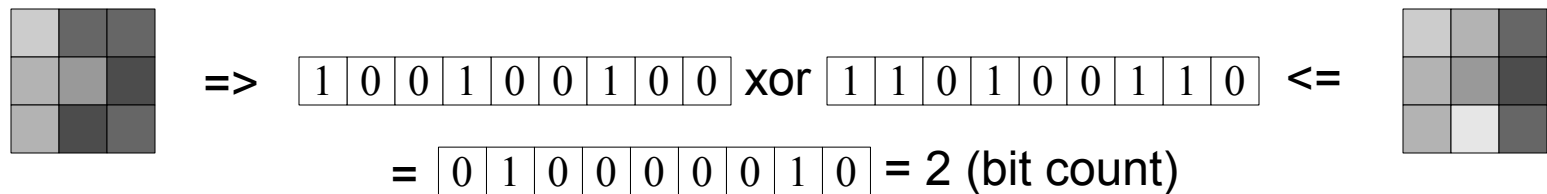
SGM (HMI)



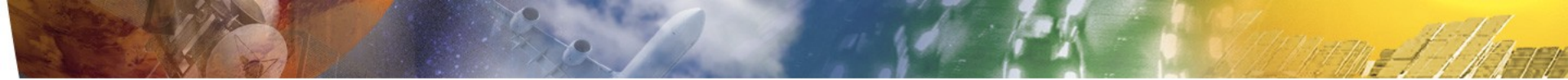
[1.18s]

Census (Zabih und Woodfill, ECCV 1994)

- Encoding the local neighborhood of each pixel into a bit vector
- Hamming distance is used as matching cost



- Originally invented for correlation (sum of hamming distances over window)
- Reduces errors at object borders since outliers have very low influence
- Hamming distance is directly used as cost for pixelwise matching (e.g. SGM)
- Strength:
 - Invariant against all radiometric transformations that maintain the local order of intensities
 - Matches images with global and local radiometric differences very well
- Weakness:
 - Window based (no pixelwise matching)
 - More sensitive to noise than Mutual Information



Implementations





CPU Implementations

- Implemented in C
- Core loop hand coded with SSE 2 vector commands
- 640 x 480 pixel with 128 pixel disparity range requires 3.4 s on Desktop (single threaded)
- Implemented on cluster for aerial and satellite image processing
- Third party CPU reimplementations and modifications:
 - Gehrke et al. (ISPRS 2010), Haller et al. (CITS 2010), Gehrig & Rabe (CVPR 2010), Gallup et al. (CVPR 2008), Gehrig & Franke (ICCV 2007), ...
 - Companies like Leica, ...
 - OpenCV (SGBM $\neq$ SGM !!!)

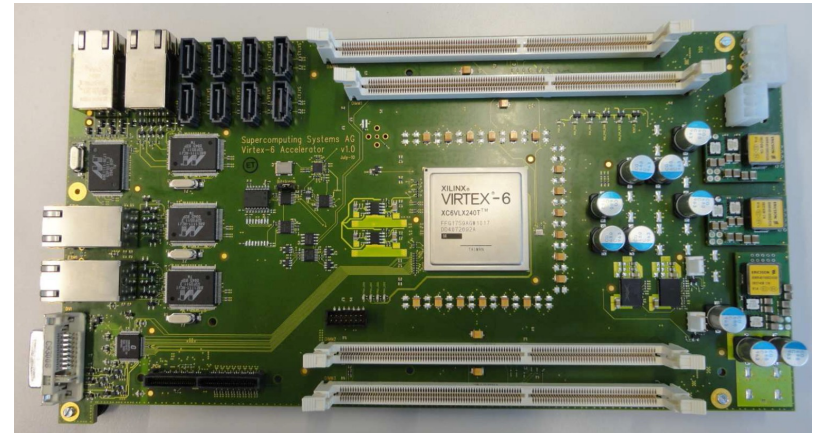


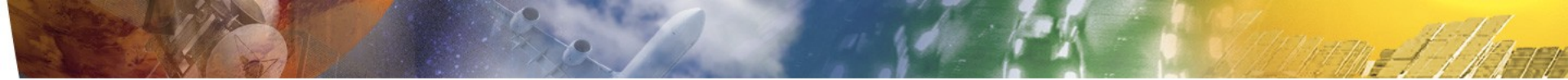
GPU Implementation

- 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)
- Third party GPU reimplementations:
 - Banz et al. (ICCV 2011)
 - Gibson & Marques (CVPR 2008)
 - Rosenberg et al. (SIGGRAPH 2006)
 - ...

FPGA Implementation of SGM

- FPGA accelerator for SGM (Census)
- Communication to PC via GigE
- 16 bit input images
- 512 x 512 - 2048 x 2048 pixels
- Disparity range: 32 - 4096 pixels
- For 2048 x 2048 x 1024 pixel
=> 2.8 s (incl. LR-check)
- Runtime on one Intel X5570 Server CPU core (3 GHz) is 35-68 times higher
- Total power consumption of FPGA: 54 W (including power supply and fan)
- Implemented by Supercomputing Systems (Zürich) on our request
- Hirschmüller (PW 2011)
- Third party FPGA reimplementations:
 - Banz et al. (2010), Gehrig et al. (ICVS 2009), ...





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)



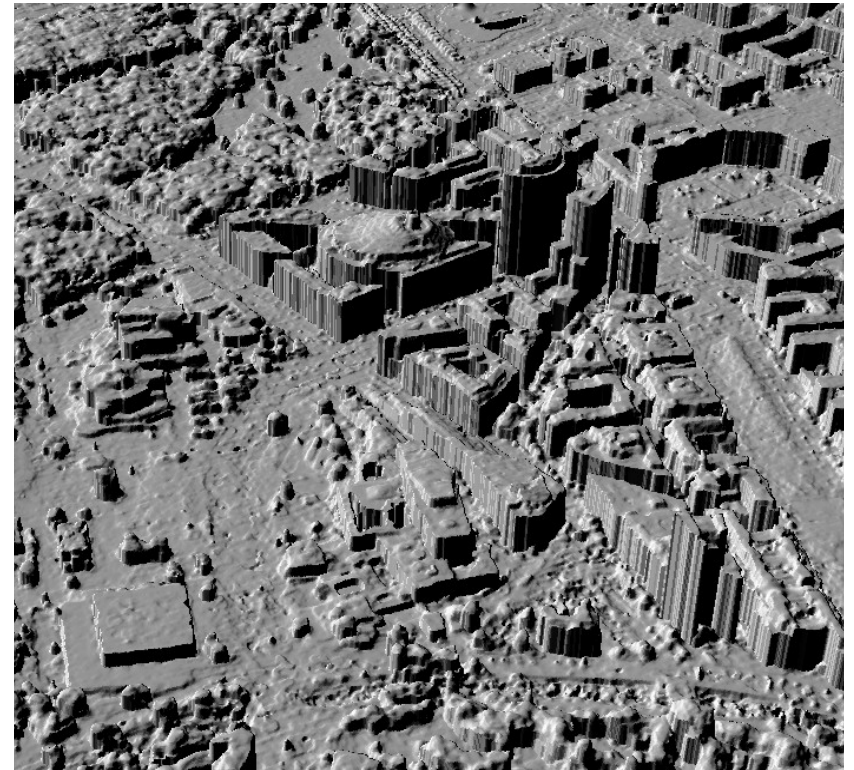
Processing of Aerial Pushbroom Images

- 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



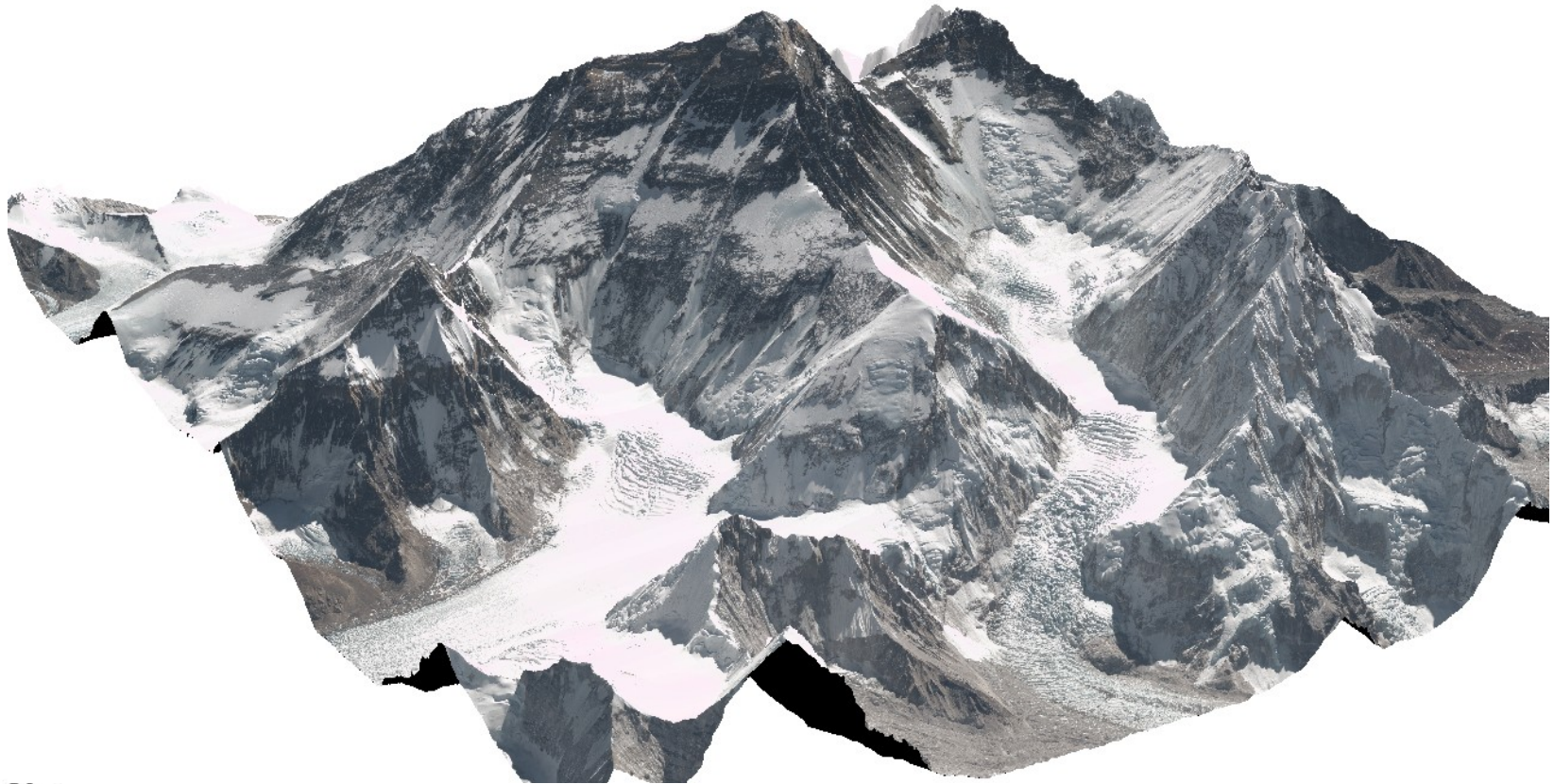
Processing of Satellite Images (1)

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



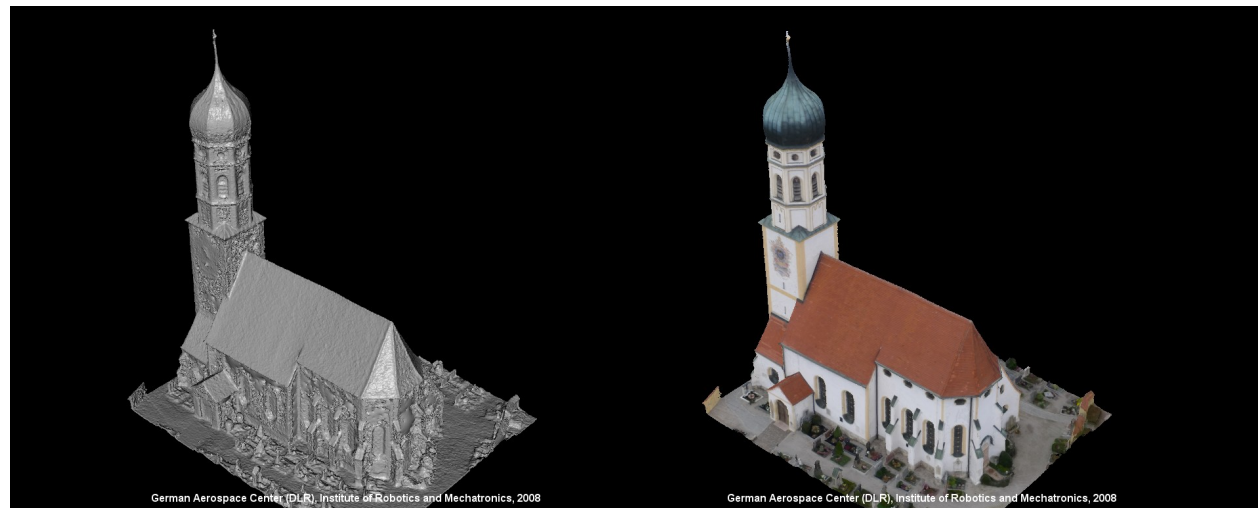
Processing of Satellite Images (2)

- 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



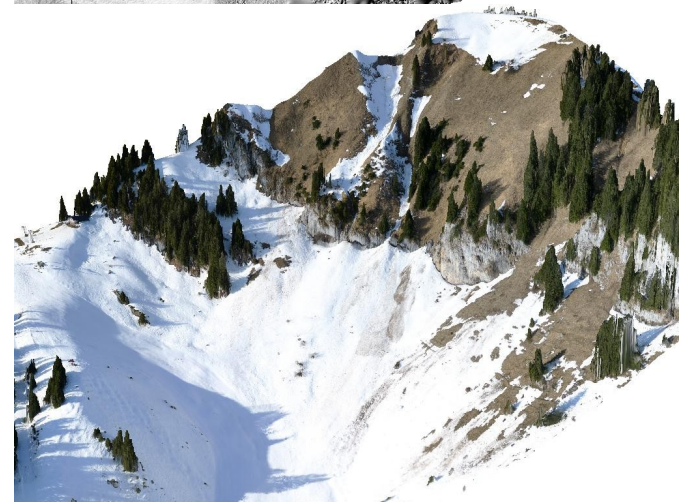
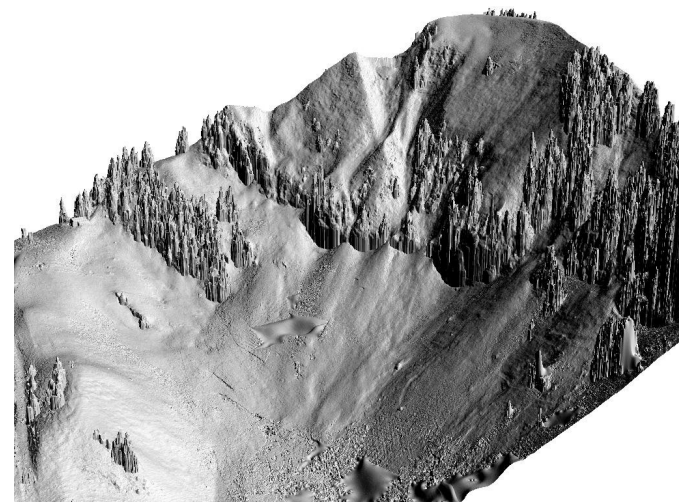
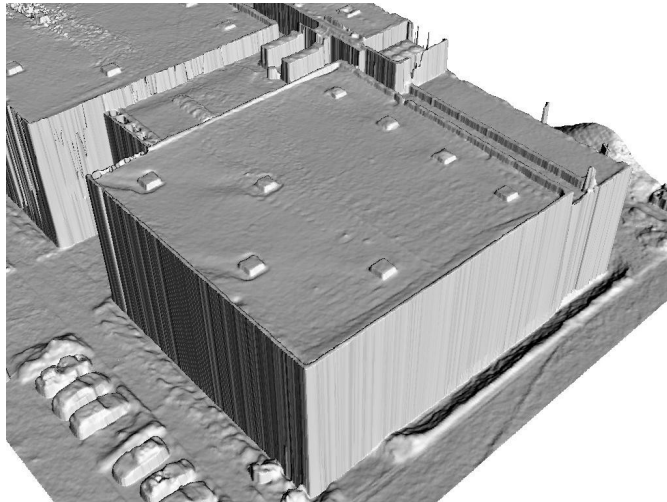
3D Reconstruction from UAVs

- 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



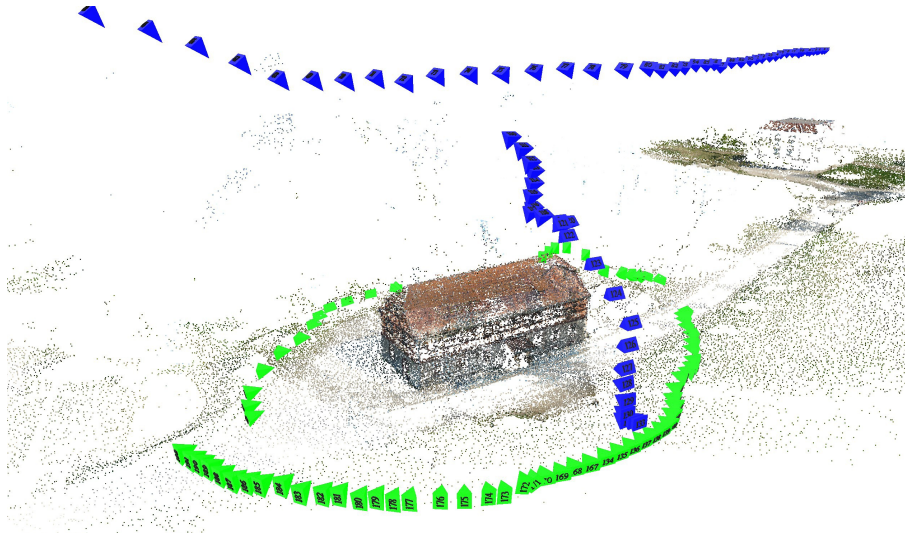
Reconstruction from UAVs

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



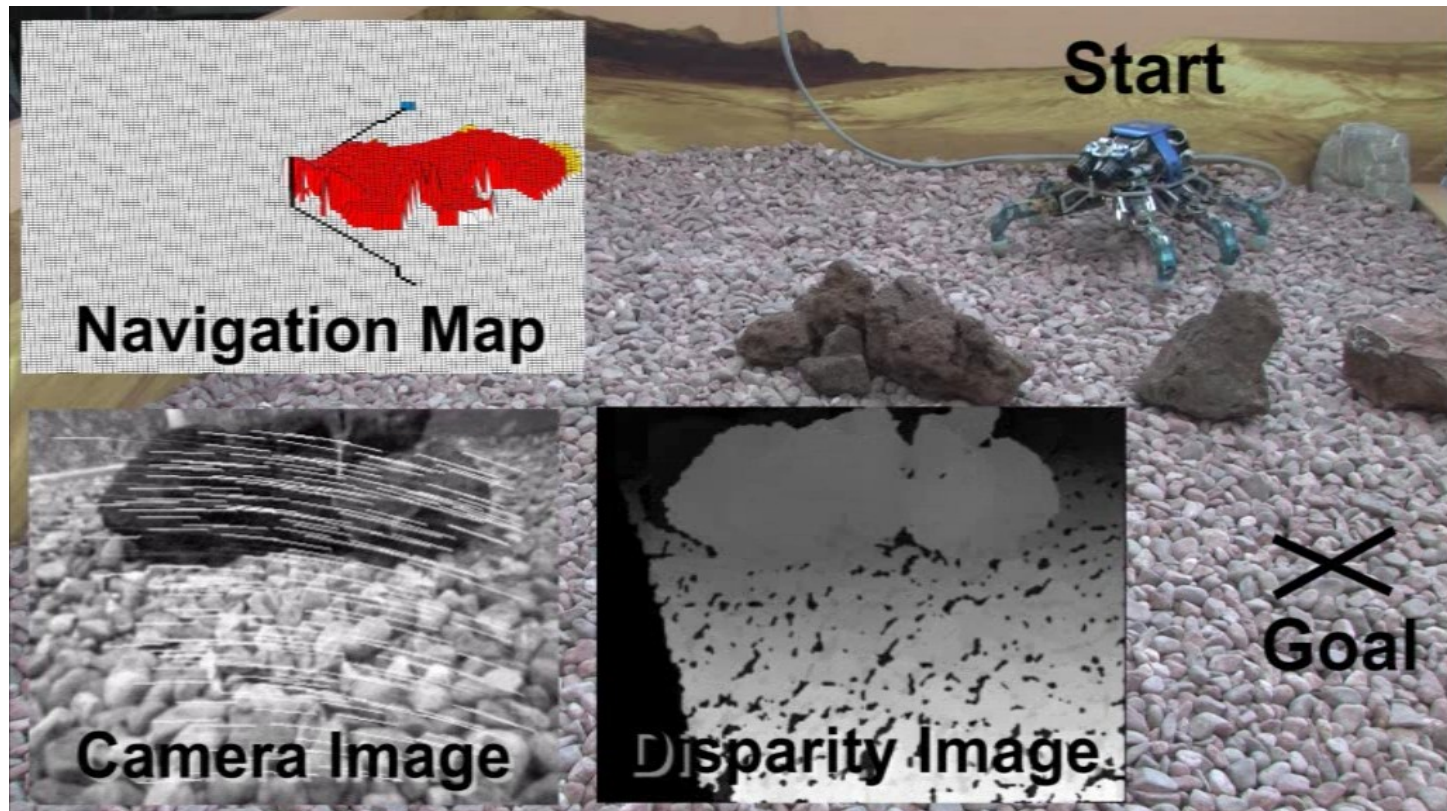
Automatic 3D Modeling from UAV and Ground Images

- Cooperation with Prof. Mayer, Jan Bartelsen and Andreas Kuhn (UniBw)
- Combining robust, automatic wide baseline registration (Bartelsen & Mayer, 2010) with dense, pixelwise matching by SGM
- Current work includes improving 3D fusion of depths images



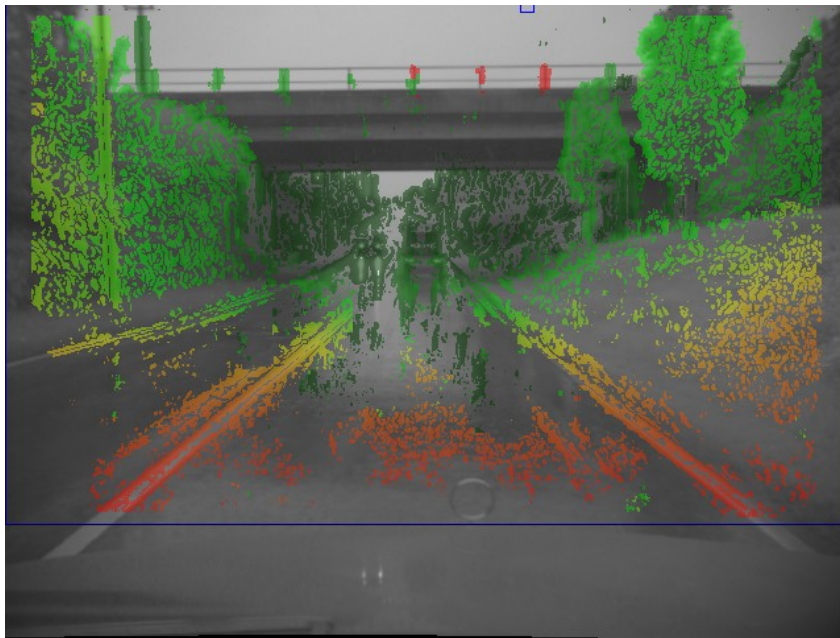
Navigation of Mobile Robots

- 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

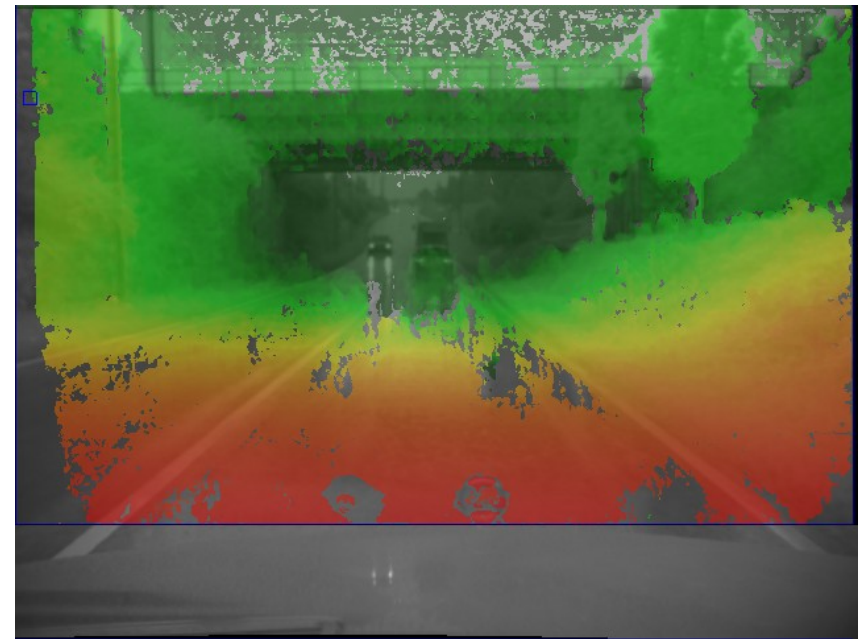


Driver Assistance Systems

- 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)
- Helmut Mayer, Jan Bartelsen and Andreas Kuhn (UniBw)
- Supercomputing Systems (www.scs.ch)
- Stefan Gehrig and Uwe Franke, Daimler AG (www.6d-vision.com)