From Orientation to Functional Modeling for Terrestrial and UAV Images

Helmut Mayer[1]

Andreas Kuhn[1], Mario Michelini[1], William Nguatem[1],
Martin Drauschke[2] and Heiko Hirschmüller[2]

[1]Visual Computing, Bundeswehr University Munich
[2]Institute of Robotics and Mechatronics, German Aerospace Center (DLR)

Introduction

► Images from the ground and from small (about 2 kg) Unmanned Aerial Vehicles (UAVs) – ascend and fly over
► Orientation, dense 3D reconstruction and functional modeling of building walls and roof
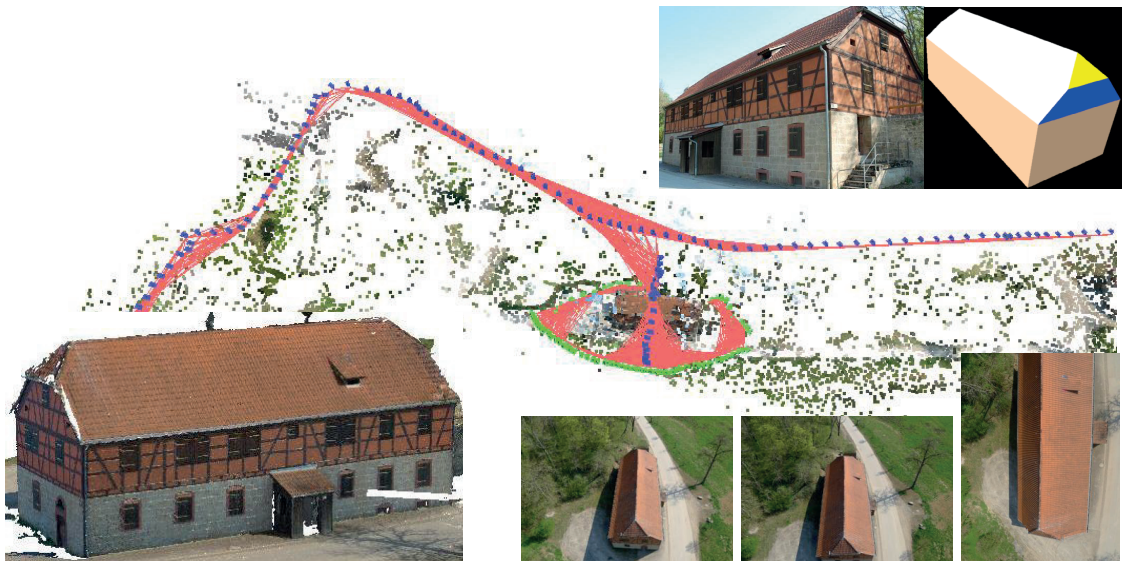
Overview

Orientation

Scalable 3D Modeling

Building Reconstruction

Facade Interpretation

Conclusion

Orientation

- ► Focus on high precision and completeness also for images with strong relative perspective distortion, i.e., wide baseline / very different viewing angle.

- ► Precise and Reliable Orientation (**PRO**) by combination of
  - ► robust direct approaches RANSAC (FISCHLER & BOLLES 1981) and 5-point algorithm (NISTÉR 2003) with
  - ► highly precise least squares image matching and robust bundle adjustment.

- ► Triplets basis for improved reliability.
- ► **Hierarchical linking** (MAYER 2014): Combination of two image sets of similar size and with two common images. Linking of triplets leads to quadruplets ($N = M = 4$) and then to 6-, 10-, 18-, 34-, 66-, 130-, 258-, 514-, 1026-tuples, etc.
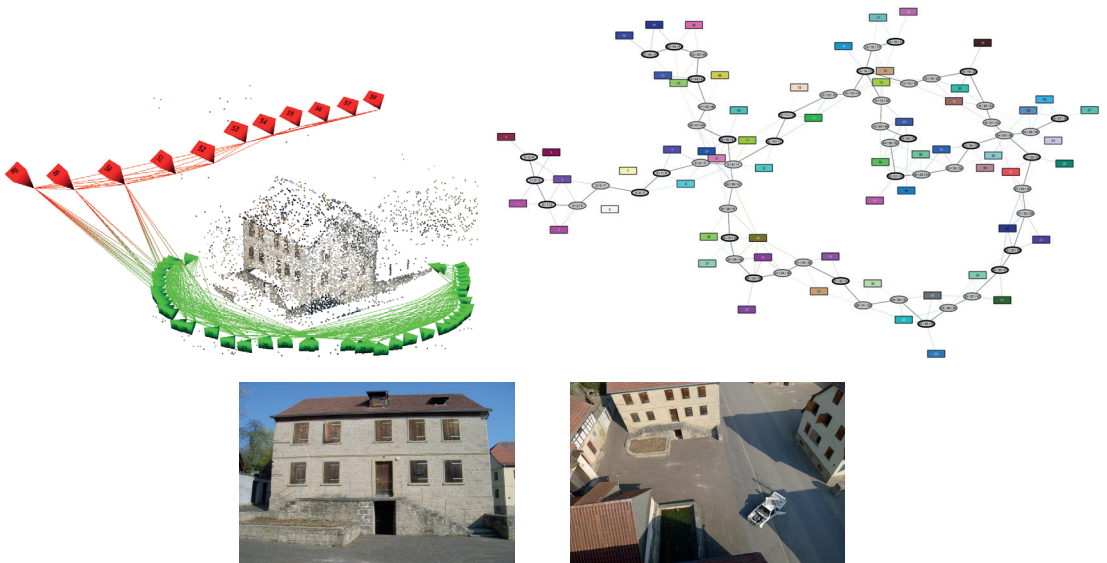
Orientation

Fully automatic linking (Michelini)

▶ Image graph describes similarity of images in image pairs via the number of corresponding points found by SiftGPU (WU 2007).

▶ Minimum number of pairs linking all images found based on Minimum Spanning Tree (MST)

▶ Pairs in MST are verified by PRO. If necessary, additional MSTs are generated.

▶ Verified pairs are basis for terminal Steiner minimal tree (LIN & XUE 2002), i.e., approximately minimal set of triplets linking all images.

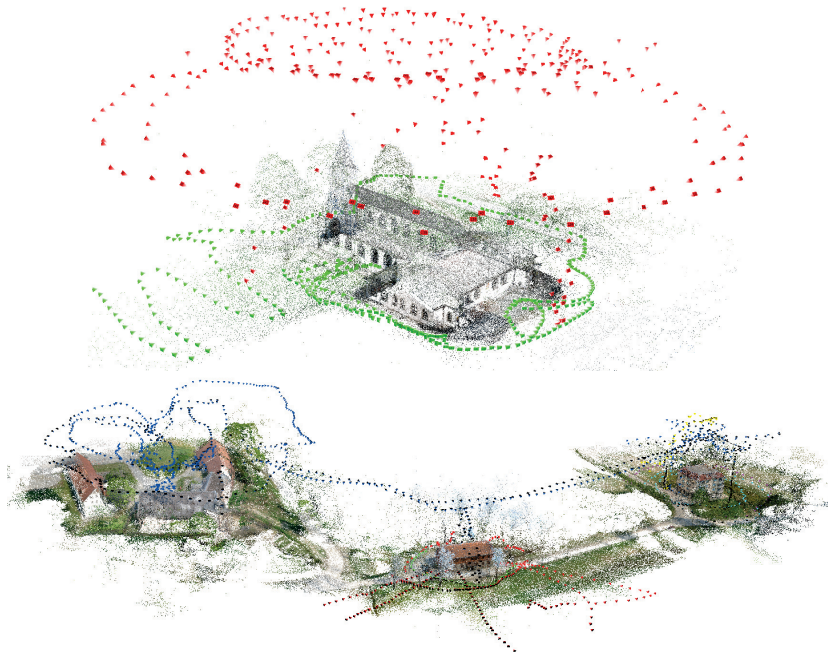▶ Triplets verified by PRO and if necessary, new Steiner tree is built.
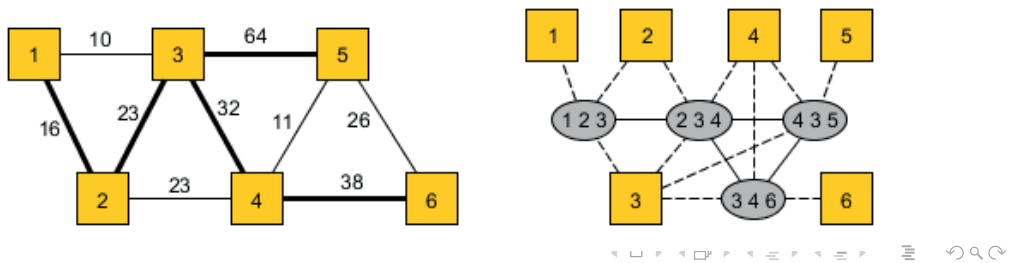
Orientation

3D model House, terminal Steiner minimal tree and two images for which the relative pose could be determined

Orientation

3D models for Church (of Bundeswehr University Munich) and Village


Orientation

Comparison to VisualSFM (WU 2011, WU 2013)

	Image set	House	Church	Village
	# images	59	655	1570
Ours	Runtime [min]	5	32	186
	# points	11 965	141 872	931 424
	Accuracy [pixel]	0.23	0.43	0.29
VisualSFM	Runtime [min]	4	271	1674
	# points	23 735	184 947	729 147
	Accuracy [pixel]	1.03	2.25	2.04

▶ By means of MST and terminal Steiner minimum tree the complexity of the reconstruction can be considerably reduced.

▶ Thus, computationally expensive PRO can be used to obtain highly precise results.

Scalable 3D Modeling

- Goal of (KUHN 2014): High quality fully 3D reconstruction of scenes of unlimited size from possibly high resolution images
- Steps: 1) Orientation $\rightarrow$ 2) Disparity estimation via SGM (HIRSCHMÜLLER 2008) $\rightarrow$ 3) Quality determination $\rightarrow$ 4) 3D probability space $\rightarrow$ 5) Fusion $\rightarrow$ 6) Filtering $\rightarrow$ 7) Triangulation
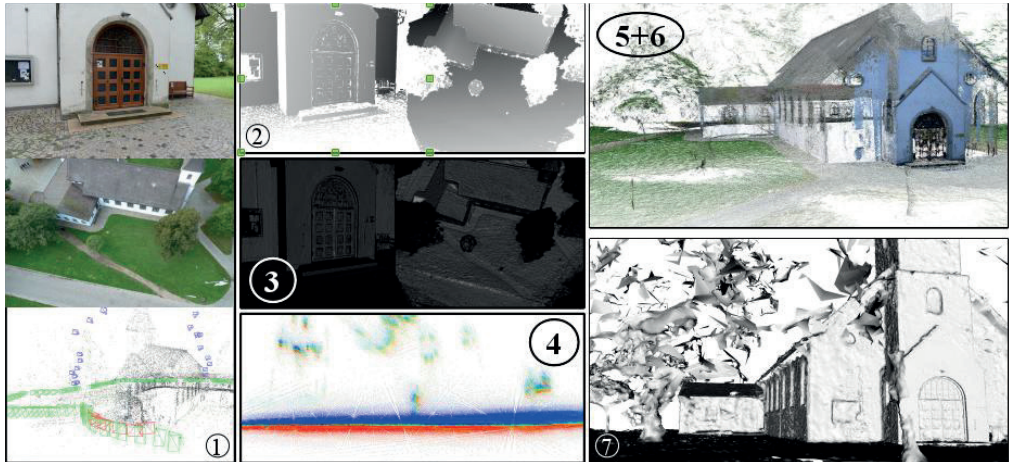
Scalable 3D Modeling

- Divide & Conquer approach with adaptive splitting of 3D voxel space allows for parallel 3D reconstruction for large image sets.
- Local optimization results in fast processing.
- $\Rightarrow$ Processing of unlimited scenes possible.
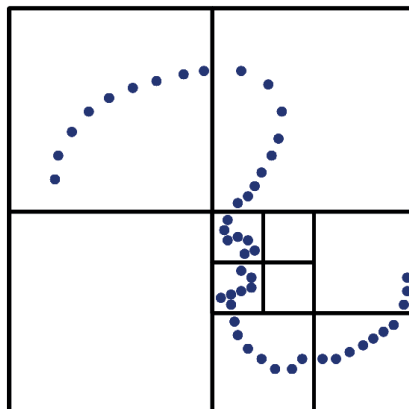
Scalable 3D Modeling

- Important aspect of (KUHN et al. 2014): Modeling of disparity uncertainty and its influence on 3D reconstruction
- Idea: Relation of disparity uncertainty to local total variation (TV)
- Window size determined based on threshold for TV.
- Relation of disparity uncertainty to TV is learned from ground truth data (SCHARSTEIN et al. 2014)

Scalable 3D Modeling

- For 3D integration in voxel space the disparity uncertainties are propagated to the voxels (KUHN et al. 2013).
- For integration a Gaussian distribution and Bayes fusion is used.
- Surface points, which are at the boundary between "in front" and "behind", are estimated by means of Gaussian regression.
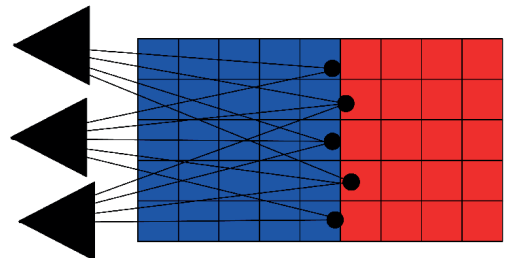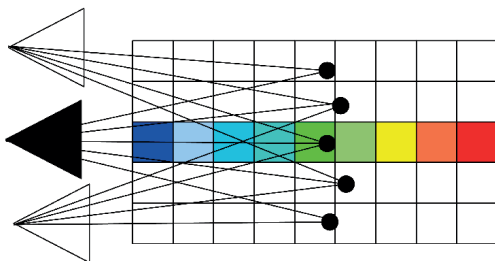
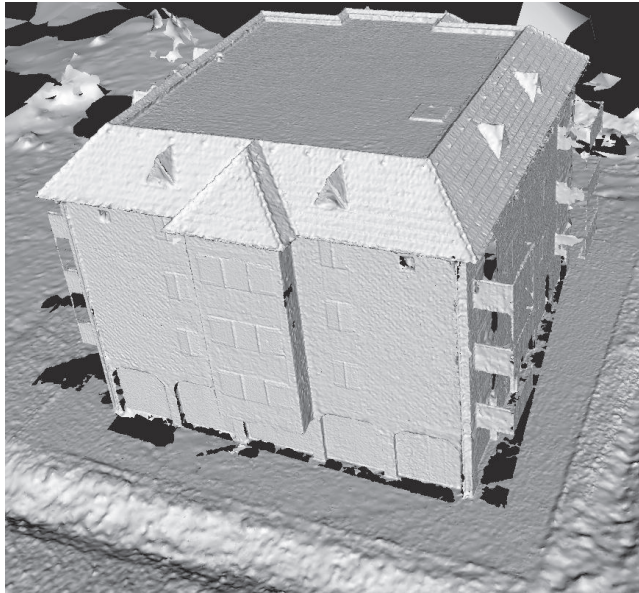# Scalable 3D Modeling

► Herzjesu (STRECHA et al. 2008)

Scalable 3D Modeling

► Church of Bundeswehr University Munich

Scalable 3D Modeling

► Building with mansard roof

Building Reconstruction

► (NGUATEM et al. 2012): Detection of building facades as vertical planes
► Approximate vertical direction by cross product of normals of pairs of points under the assumption that majority of points lies on the facade.
► Vertical direction refined based on edges in point cloud.
► Determination of vertical planes by means of RANSAC: Because vertical direction is known, two points suffice to define plane.
► Adjacent planes are detected and intersected.

Building Reconstruction

- ▸ (NGUATEM et al. 2013): Stochastic sampling of roofs
- ▸ Statistical model selection for determination of roof type
- ▸ Determination of 3D Structures in front of facades
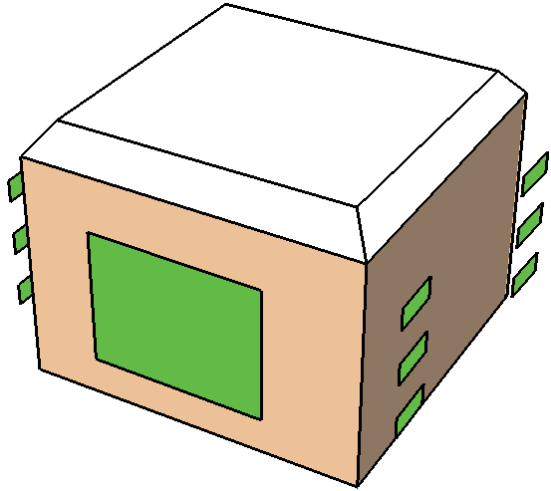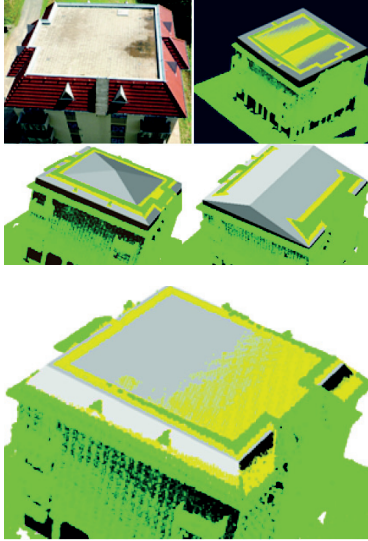
Facade Interpretation

- ▸ (NGUATEM et al. 2014): Localization of windows and doors behind facade in 3D point cloud via model based stochastic search
- ▸ Catalog of doors and windows
- ▸ Generic geometric representation based on splines
- ▸ Reduced stochastical search space using 2D similarity transformation
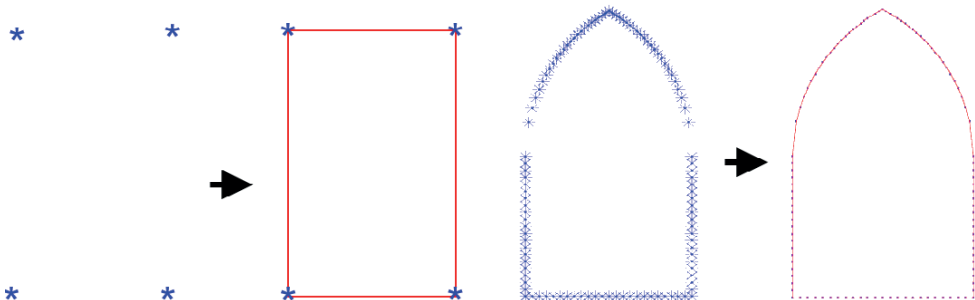
Facade Interpretation

- Stochastical search by Markov Chain Mote Carlo - MCMC for windows and doors
- Window hypotheses and inliers
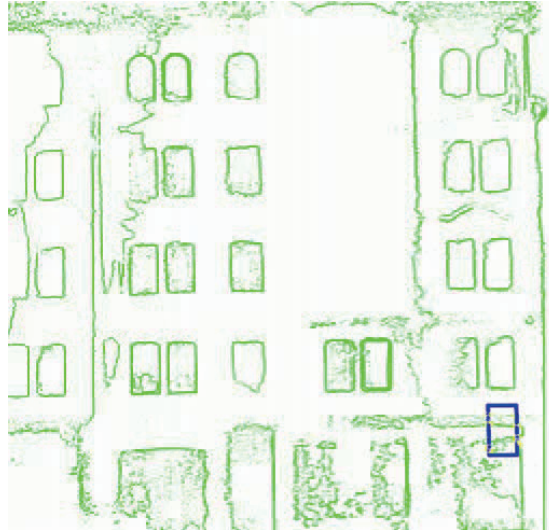
Facade Interpretation

- Small facade

Facade Interpretation

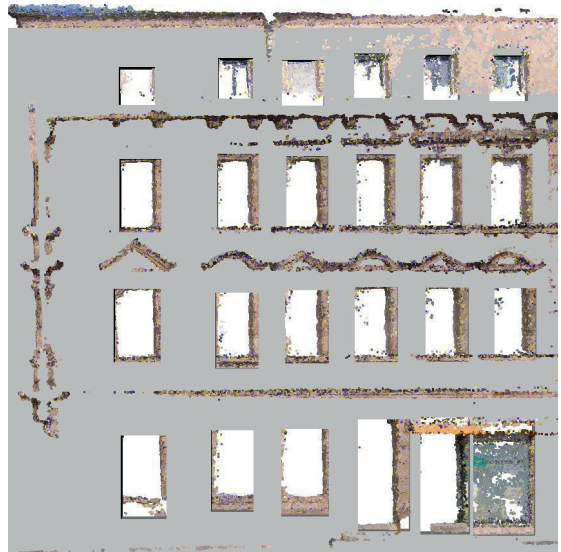▶ Larger facade

---

Conclusion

▶ Presentation of fully automatic approach for precise and reliable orientation for unordered sets of wide baseline images.
▶ Shown in comparison to VisualSFM that Minimum Spanning Trees and terminal Steiner minimal trees can reduce the computational complexity considerably.
▶ Demonstrated that a local statistical volumetric approach taking into account a total variation (TV) measure to model the disparity uncertainty leads to high quality fully 3D reconstruction for large scenes.
▶ Given results for statistical functional modeling for facades, building roofs, windows and doors.

Bibliography I

FISCHLER, M. & BOLLES, R. (1981): Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communications of the ACM* **24**(6): 381–395.

HIRSCHMÜLLER, H. (2008): Stereo Processing by Semiglobal Matching and Mutual Information, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(2): 328–341.

KUHN, A. (2014): *Scalable 3D Surface Reconstruction by Local Stochastic Fusion of Disparity Maps*, Dissertation, Universität der Bundeswehr München.

KUHN, A., HIRSCHMÜLLER, H. & MAYER, H. (2013): Multi-Resolution Range Data Fusion for Multi-View Stereo Reconstruction, *German Conference on Pattern Recognition – GCPR 2013*, Springer-Verlag, Berlin, 41–50.

KUHN, A., HIRSCHMÜLLER, H., MAYER, H. & SCHARSTEIN, D. (2014): A TV Prior for High-Quality Local Multi-View Stereo Reconstruction, *International Conference on 3D Vision (3DV)*, 65–72.

LIN, G. & XUE, G. (2002): On the Terminal Steiner Tree Problem, *Information Processing Letters* **84**(2): 103–107.

MAYER, H. (2014): Efficient Hierarchical Triplet Merging for Camera Pose Estimation, *German Conference on Pattern Recognition – GCPR 2014*, Springer-Verlag, Berlin, 399–409.

Bibliography II

NGUATEM, W., DRAUSCHKE, M. & MAYER, H. (2012): Finding Cuboid-based Building Models in Point Clouds, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume (39) B3B, 149–154.

NGUATEM, W., DRAUSCHKE, M. & MAYER, H. (2013): Roof Reconstruction from Point Clouds Using Importance Sampling, *International Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume II-3/W3, 73–78.

NGUATEM, W., DRAUSCHKE, M. & MAYER, H. (2014): Localization of Windows and Doors in 3D Point Clouds of Fa cades, *International Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume II-3, 87–94.

NISTÉR, D. (2003): An Efficient Solution to the Five-Point Relative Pose Problem, *Computer Vision and Pattern Recognition*, Volume II, 195–202.

SCHARSTEIN, D., HIRSCHMÜLLER, H., KITAJIMA, Y., KRATHWOHL, G., NESIC, N., WANG, X. & WESTLING, P. (2014): High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth, *German Conference on Pattern Recognition – GCPR 2014*, 31–42.

STRECHA, C., VON HANSEN, W., VAN GOOL, L., FUA, P. & THOENNESSEN, U. (2008): On Benchmarking Camera Calibration and Multi-View Stereo for High Resolution Imagery, *Computer Vision and Pattern Recognition*, 1–8.

Bibliography III

Wu, C. (2007): SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT), `cs.unc.edu/~ccwu/siftgpu`.

Wu, C. (2011): VisualSFM: A Visual Structure from Motion System, `ccwu.me/vsfm/`.

Wu, C. (2013): Towards Linear-time Incremental Structure from Motion, *International Conference on 3D Vision (3DV)*, 127–134.