Real-time Bundle Adjustment with an Omnidirectional Multi-Camera System and GPS

Johannes Schneider, Thomas Läbe, Wolfgang Förstner
Institute of Geodesy and Geoinformation, University of Bonn

Abstract
In this paper we present our system for visual odometry that performs a fast incremental bundle adjustment for real-time structure and motion estimation in an unknown scene. It is applicable to image streams of a calibrated multi-camera system with omnidirectional cameras. In this paper we use an autonomously flying octocopter that is equipped for visual odometry and obstacle detection with four fisheye cameras, which provide a large field of view. For real-time ego-motion estimation the platform is equipped, besides the cameras, with a dual frequency GPS board, an IMU and a compass. In this paper we show how we apply our system for visual odometry using the synchronized video streams of the four fisheye cameras. The position and orientation information from the GPS-unit and the inertial sensors can optionally be integrated into our system. We will show the obtained accuracy of pure odometry and compare it with the solution from GPS/INS.

Keywords
Visual odometry, incremental bundle adjustment, omnidirectional cameras, multi-camera system, UAV

1 INTRODUCTION

In the DFG-project *Mapping on Demand* at the University of Bonn and the Technical University of Munich we are developing a lightweight autonomously navigating unmanned aerial vehicle (UAV). The goal of the research project is to develop and to test procedures and algorithms for the fast three-dimensional semantic mapping of inaccessible objects on the basis of a high-level user request. The UAV flies fully autonomously to capture images in a user specified region even in close vicinity to obstacles. The captured images are directly transferred to a ground station, where they are processed in a fast bundle adjustment, to refine the on-board estimated positions and orientations. The accurately oriented images are subsequently used for dense surface reconstruction (Steinbrücker, F. et al., 2014) and semantic interpretation (Loch-Dehbi, S. et al., 2013).

Lightweight UAVs can operate from above in inaccessible and even dangerous areas. The on-board sensing of a lightweight UAV has to be designed with regards to its limitation in size and weight, and the limited on-board processing power and on demand semantic mapping tasks require highly efficient algorithms. Our sensor setup is carried by an octocopter platform with a maximum total weight of 5 kg. The arrangement of the on-board sensors is shown in Fig. 1. Besides the high resolution camera which is used for the actual mapping task on the ground-station, our octocopter is equipped with four fisheye cameras. The fisheye cameras are used besides (a) ultra sonic sensors and a rotating laser scanner for obstacle perception (Nieuwenhuisen, M. et al., 2013) and (b) a GPS-unit, an IMU and a compass for on-board ego-motion estimation (Eling, C. et al., 2013). The full sensor setup and the integration of the multiple sensors for the autonomous navigation task are described in detail in (Klingbeil, L. et al., 2014).

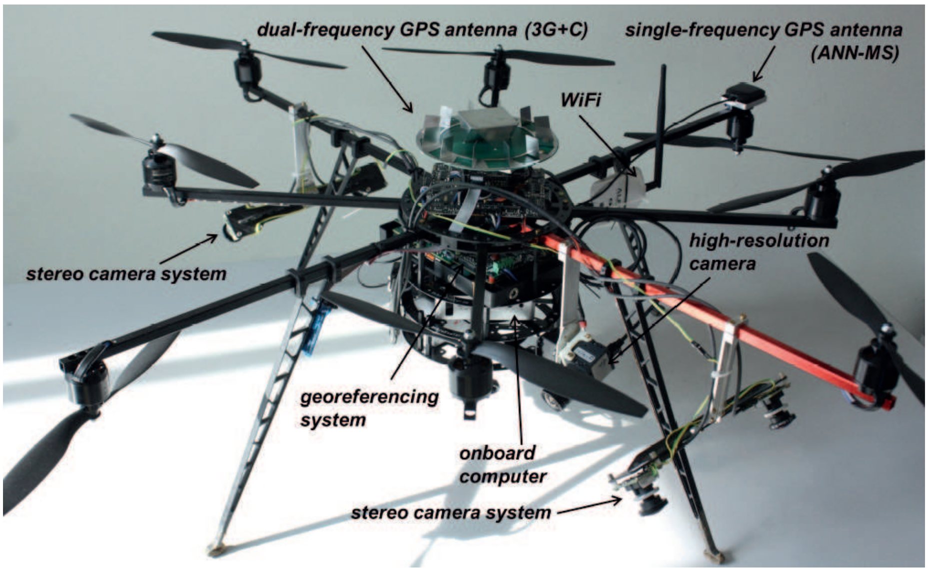

Figure 1: Sensor setup of the UAV (picture taken from Klingbeil, L. et al., 2014).

The four fisheye cameras are mounted as two stereo cameras. One camera pair is looking forward and one backwards with a pitch angle of 45°, covering a large field of view, see Fig. 2. Each camera has a Lensagon BF2M15520 fisheye lens with a field angle up to 185°. The cameras are sampled with a frequency of 10 Hz in a synchronized way. The basis between the cameras amounts to 20 cm providing highly overlapping views at each time of exposure. The monochromatic images have a resolution of 752×480 pixels.

Bundle adjustment is the work horse for orienting cameras and determining 3D points. It has a number of favorable properties, e.g. it is statistically optimal in case all statistical tools are exploited and it is highly efficient in case sparse matrix operations are used. Nevertheless, the computational expense grows with the number of images involved. Incremental bundle adjustment avoids periodical batch steps with recurring calculations by performing only calculations for entries of the information matrix, i.e. the normal equation matrix or inverse covariance matrix, that are actually effected by new measurements. (Kaess, M. et al., 2012) provide a sparse nonlinear incremental optimization algorithm called iSAM2, which is highly efficient, as only variables are relinearized that have not converged yet and as fill-in is avoided through incrementally changing the variable ordering. A bundle adjustment that works with omni-directional and multi-view cameras, that can handle arbitrary bundles of rays and that allows for points at infinity, was realized in former work as a batch version, see (Schneider, J. et al., 2012), and as an incremental version, see (Schneider, J. et al. 2013). In the later paper first experiments w.r.t. time requirements and optimality of the solution were shown by using the iSAM2 algorithm for a keyframe-based incremental real-time bundle adjustment.

The paper is organized as follows. In section 2 we treat the issue of visual odometry for real-time ego-motion estimation using the synchronized image streams of the four omnidirectional fisheye cameras in a keyframe-based fast incremental bundle adjustment that is able to integrate the position and orientation information obtained by a GPS-unit and IMU. In section 3 we will examine the accuracy of pure odometry and compare it with the acquired GPS/INS information. Finally we will conclude and give an outlook on future work.

2 CONCEPT FOR VISUAL ODOMETRY

Visual odometry consists in determining the pose, i.e. the position and orientation, of the cameras in real-time. Our system uses feature points detected and tracked on the synchronously taken frames of the four image streams. The process consists basically of the following steps:

1. The data acquisition and association detects feature points, performs the matching and provides camera rays associated to other cameras.
2. The orientation of the frames of the image streams provides initial values for the subsequent bundle adjustment and is used to select keyframes.
3. The incremental bundle adjustment uses the new information at a keyframe and merges it optimally with the previous information.

We determine the interior orientation of each fisheye camera in advance using the equidistant-model and Chebyshev polynomials according to (Abraham, S. and Hau, T., 1997). The position and orientation between the four cameras is determined by using the system self-calibrating bundle adjustment as described in (Schneider, J. and Förstner, W, 2013).

2.1 Incremental Image Orientation

Our visual odometry method is based on interest points, which are tracked in the image streams from frame to frame using the KLT tracker (Bouguet, J.-Y., 2000) from the OpenCV library. The tracks in the individual cameras are matched across the cameras by using a correlation based matching. An example of extracted feature points in a simultaneously taken frame set of the four fisheye cameras is shown in Fig. 2. Tracked feature points are converted into ray directions by using the interior orientation determined in advance.

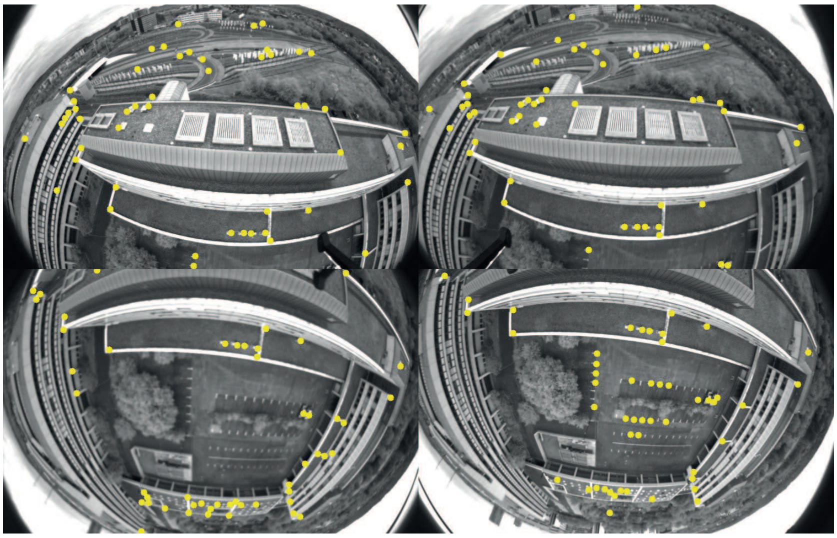

Figure 2: A frame set consisting of four images taken with the four fisheye cameras. Each image contains 50 interest points, which are tracked using the KLT tracker from the OpenCV library.

At the initiating frame set we determine scene point coordinates by forward intersecting the matched ray directions in the stereo pairs. After initialization, each frame set is oriented by computing the motion of the camera system via simultaneous resection of all cameras using a generalized camera model with multiple projection centres and the known system calibration. We use a fast and robust Maximum-Likelihood-type estimation that converges mostly in 2-3 fast iterations when using the last pose as the initial solution. Robustness is achieved by down weighting observations with large residuals. Tracks with observations getting low weights are considered as corrupted. Corrupted tracks are replaced by new tracks by extracting new interest points.

Our incremental bundle adjustment refers to keyframes, which reduce the processing to some geometrically useful, tracked observations. A keyframe consists of four frames taken simultaneously. The initiating frame set is used as the first keyframe. Further keyframes are initiated if a minimal distance or rotation regarding to the last keyframe is exceeded. In case a new keyframe set is initiated, the new observations are used to update and refine the scene points and poses of the keyframes in the incremental bundle adjustment. Initial values for new tracked scene points are obtained by forward intersection, where we claim that each track consists of at least three keyframes. We do not use intersected scene points that show large residuals in the observations and we consider the associated track to be corrupted. Observations of already intersected scene points are assumed to be revised from corrupted tracks via the former robust resection.

With each new keyframe the bundle adjustment is solved including the new observations and variables. The bundle adjustment refines the scene points and poses of the keyframes simultaneously. This step is the most costly one as it uses all available data. For our real-time application the processing of a new keyframe needs to be finished by the time the next keyframe is added. For the first ten keyframe sets we use a batch bundle adjustments as the optimization task includes only a small number of variables yet. Since further bundle adjustment steps grow in complexity they need to be solved efficiently to ensure real-time capability. Therefore we merge new information incrementally with the previous information using the software package iSAM2 for “incremental smoothing and mapping” (Kaess, M. et al., 2012) that yields a fast optimal solution for our bundle adjustment (Schneider, J. et al., 2013).

2.2 Integration of GPS/INS Information

The measured position and orientation by GPS and IMU refers to the body frame of the UAV. We calibrated the four cameras to the body frame by using a highly accurate scanned point cloud of the four cameras and the physical reference of the body frame. The camera directions are determined by fitting cylinders into the scanned camera casings and the projection centres are derived from the centres of the scanned lenses. We transform the mutual positions and orientations derived from the system self-calibrating bundle adjustment on the camera directions and projection centres derived from the scanned point cloud by using a spatial similarity transformation that includes also the rotation parameters (Dickscheid, T et al., 2008). The projection centres derived from the scan are badly determined in the direction of the camera directions and the rotations around the camera directions can only be approximately assumed. We take this into account by lowering the corresponding weights in the similarity transformation.

At the moment we are working on the integration of GPS/IMU information into our real-time visual odometry system. In case all sensors are calibrated in one single frame, the so called body frame, we can use the position and orientation with its covariance as direct observations of the pose parameters on each keyframe. As these poses are oriented in a different coordinate system as the poses of the visual odometry, we transform the poses of all keyframes with a similarity transformation as soon as enough GPS/IMU information is available. From then on new observations can be incrementally added to our bundle adjustment.

3 FIRST EXPERIMENTS

Tracking 200 feature points in each camera and setting the convergence criterion for the rotations to 0.1° and for the translations to 1cm yields a very fast processing of the bundle adjustment that is always faster than one second on the on-board computer (Intel Core i7, 8GB RAM). In (Schneider, J. et al., 2013) we have shown that the required time is independent of the number of new observations added to the optimization problem but rather highly depends on the number of affected variables that need to be relinearized in an incremental optimization step within the iSAM2 algorithm. Further, we have shown that the incremental bundle adjustment provides estimated pose parameters which are in a statistical sense optimal like using a rigorous batch bundle adjustment.

In this paper we want to compare the on-board processed trajectory with our visual odometry algorithm of the UAV with the on-board processed solution of the GPS/IMU. Figure 3 shows positions obtained by our visual odometry system (solid line with crosses) and positions obtained by the georeferencing unit (fixed GPS solutions are marked with green dots and float solutions with red dots) during a five minute long flight. The visual odometry initiated 273 keyframes with 17,157 observations of 864 scene points, whereby 4,803 frame sets were orientated by resection (not shown in Fig. 3). To compare both trajectories the 7-parameter similarity transformation between the poses of incrementally refined keyframes and the poses from GPS/INS is determined, using the positions of both trajectories. The differences to the positions obtained by GPS are up to 60 cm.

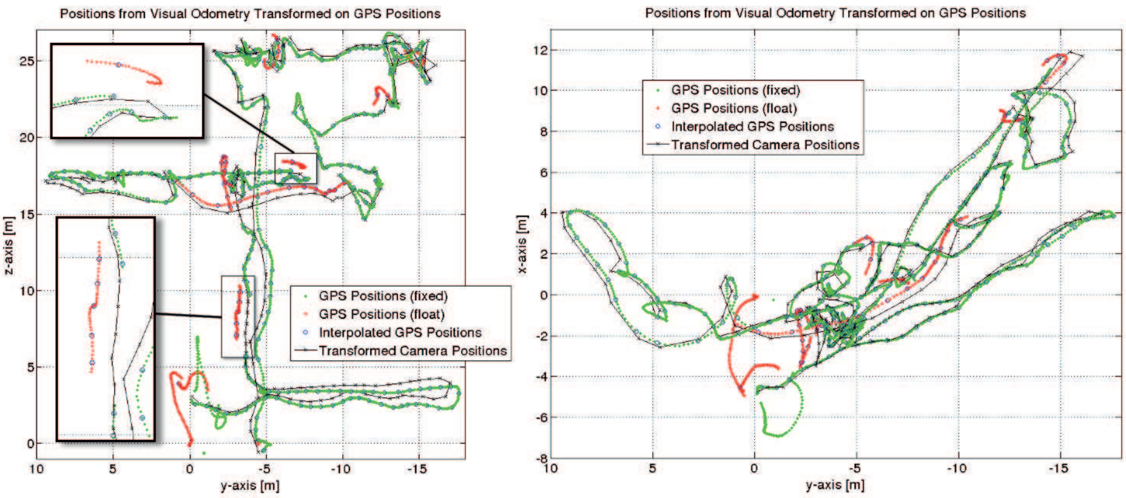

Figure 3: The onboard processed trajectories of the georeferencing unit and transformed incremental bundle adjustment. Left: Lateral view on the trajectories in the yz-plane, note the capability of the visual odometry to bridge float solutions of the georeferencing unit. Right: Top view on the trajectories in the xy-plane.

Figure 4 shows the estimated accuracy of the on-board estimated translation and rotation parameters of the keyframes regarding to the pose of the first keyframe at the origin in Figure 3. Note that the uncertainty in the translation grows in the direction the UAV flies. The uncertainty of the rotations grows linearly. The estimated standard deviations of the translations stay under 30 cm and the uncertainty in the rotation parameters rises up to 0.4°. Note that the estimated uncertainties do not describe drift or systematic errors in the calibration.

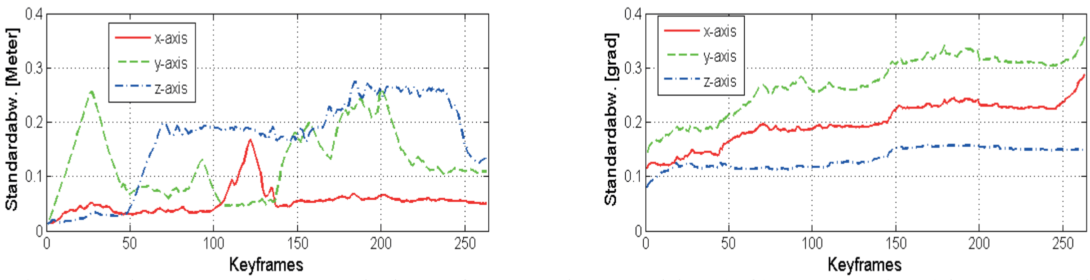

Figure 4: The estimated standard deviations of the obtained position (left) and rotation (right) regarding to the first keyframe. The estimated inner uncertainty is given for 273 keyframes.

4 CONCLUSIONS AND FUTURE WORK

We presented our system for visual odometry performing a keyframe-based bundle adjustment for real-time structure from motion estimation in an unknown scene. Incremental bundle adjustment is performed by using the iSAM2 algorithm for sparse nonlinear incremental optimization. Our bundle adjustment allows for multi-view cameras, omnidirectional cameras. First results show the achieved accuracy of the visual odometry system. Visual odometry is an interesting supplement GPS/INS, as it works in GPS-denied environments and the relative pose estimation is highly accurate.

To overcome the drift, we are integrating at the moment the GPS/INS observations into our visual odometry system. Furthermore we are replacing the drifting frame to frame tracking using the KLT tracker with a keyframe to keyframe tracking and we are thinking of a simple loop closing procedure recognizing already observed environments.

ACKNOWLEDGEMENTS

This work was supported by the DFG-Project FOR 1505 “Mapping on Demand”.

REFERENCES

ABRAHAM, S. and HAU, T.: *Towards Autonomous High-Precision Calibration of Digital Cameras*. In Proc. of SPIE Annual Meeting 3174, Videometrics V, pp. 82–93, 1997.

BOUGUET, J.-Y.: Pyramidal *Implementation of the Lucas Kanade Feature Tracker Description of the algorithm*. Intel Corporation, Microprocessor Research Labs, 2000.

DICKSCHEID, T., LÄBE, T., FÖRSTNER, W.: *Benchmarking Automatic Bundle Adjustment Results*. In ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII-B3a, pp. 7-12, 2008.

ELING, C., KLINGBEIL, L., WIELAND, M., KUHLMANN, H.: *A precise position and attitude determination system for lightweight unmanned aerial vehicles*. In ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-1/W2, pp. 113-118, 2013.

KAESS, M., JOHANNSSON, H., ROBERTS, R., ILA, V., LEONARD, J., DELLAERT, F.: *iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree*. Intl. J. of Robotics Research (IJRR), Vol. 31, pp. 217-236, 2012.

KLINGBEIL, L., NIEUWENHUISEN, M., SCHNEIDER, J., ELING, C., DRÖSCHEL, D., HOLZ, D., LÄBE, T., FÖRSTNER, W., BEHNKE, S., KUHLMANN, H.: *Towards Autonomous Navigation of an UAV-based Mobile Mapping System*. In Proceedings of the 4th international conference on Machine Control and Guidance (MCG), to appear in 2014.

LOCH-DEHBI, S., DEHBI, Y., PLÜMER, L.: *Stochastic reasoning for UAV supported reconstruction of 3D building models*. In ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-1/W2, pp. 257-261, 2013.

NIEUWENHUISEN, M., DROESCHEL, D., SCHNEIDER, J., HOLZ, D., LÄBE, T., BEHNKE, S.: *Multimodal Obstacle Detection and Collision Avoidance for Micro Aerial Vehicles*. In Proceedings of the 6th European Conference on Mobile Robots (ECMR), pp. 7-12, 2013.

SCHNEIDER, J., SCHINDLER, F., LÄBE, T., FÖRSTNER, W.: *Bundle Adjustment for Multi-camera Systems with Points at Infinity*. In ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. I-3, pp. 75-80, 2012.

SCHNEIDER, J., FÖRSTNER, W.: *Bundle Adjustment and System Calibration with Points at Infinity for Omnidirectional Camera Systems*. Z. f. Photogrammetrie, Fernerkundung und Geoinformation, Vol. 4, pp. 309-321, 2013.

SCHNEIDER, J., LÄBE, T., FÖRSTNER, W.: *Incremental Real-time Bundle Adjustment for Multi-camera Systems with Points at Infinity*. In ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-1/W2, pp. 355-360, 2013.

STEINBRÜCKER, J., STURM, J., CREMERS, D.: *Volumetric 3D Mapping in Real-Time on a CPU*. In Proceedings of the Int. Conf. on Robotics and Automation (ICRA), to appear in 2014.

Contact:
Johannes Schneider,
Institute of Geodesy and Geoinformation, Rheinische Friedrich-Wilhelms-Universität Bonn,
Nussallee 15, D-53115 Bonn, Germany
Email: johannes.schneider@uni-bonn.de