

Towards Autonomous Navigation of an UAV-based Mobile Mapping System

This paper was peer-reviewed by the scientific committee.

Lasse Klingbeil¹, Matthias Nieuwenhuisen², Johannes Schneider¹, Christian Eling¹,
David Droeschel², Dirk Holz², Thomas Läbe¹, Wolfgang Förstner¹, Sven Behnke²,
Heiner Kuhlmann¹

¹*Institute of Geodesy and Geoinformation, University of Bonn*

²*Institute of Computer Science, University of Bonn*

Abstract

For situations, where mapping is neither possible from high altitudes nor from the ground, we are developing an autonomous micro aerial vehicle able to fly at low altitudes in close vicinity of obstacles. This vehicle is based on a MikroKopter™ octocopter platform (maximum total weight: 5kg), and contains a dual frequency GPS board, an IMU, a compass, two stereo camera pairs with fisheye lenses, a rotating 3D laser scanner, 8 ultrasound sensors, a real-time processing unit, and a compact PC for on-board ego-motion estimation and obstacle detection for autonomous navigation. A high-resolution camera is used for the actual mapping task, where the environment is reconstructed in three dimensions from images, using a highly accurate bundle adjustment. In this contribution, we describe the sensor system setup and present results from the evaluation of several aspects of the different subsystems as well as initial results from flight tests.

Keywords

Mobile Mapping, GPS, UAV, Inertial Sensors, Laserscanner, obstacle avoidance, stereo vision, trajectory estimation, motion planning

1 INTRODUCTION

Data acquisition from mobile platforms has become established in many communities, because of the possibility to cover wide areas. For mapping from above, unmanned aerial systems (UASs) have been developed in the past ten years (Everaerts 2008). In contrast to other mobile platforms, unmanned aerial vehicles (UAVs) have the advantage of being able to overfly inaccessible and also dangerous areas. Furthermore, they can get very close to objects to achieve high resolution measurements with low resolution sensors. Especially in the field of precision farming such as phenotyping or plant monitoring (Xiang et al. 2011), or in the fields of infrastructure inspection (Merz et al. 2011) and recording of archaeological sites (Eisenbeiss et al. 2005), UAVs are meanwhile often deployed. In most cases, UAVs collect object information via remote sensing, which means that the data is acquired without physical contact to the object. In the following, we will use the term MAV (Micro/Mini Aerial Vehicle) as it is often applied to for UAVs with a weight lower than 5kg.

The goal of the research project *Mapping on Demand* is to develop a lightweight autonomously flying MAV that is able to identify and measure inaccessible three-dimensional objects using visual information. A major challenge within the project comes with the term ‘on demand’. Apart from the classical ‘mapping’ part, where 3D information is extracted from aerial images, the MAV is intended to fly fully autonomous on the basis of a high-level user request, avoiding obstacles and processing mapping data in real-time including the extraction of semantic information (Loch-Dehbi et al 2013). Therefore a precise and robust direct georeferencing is necessary, not only for the real-time processing of the image data, but also for the autonomous navigation of the system.

2 SYSTEM DESIGN AND SENSOR SETUP

Within the project, we have developed two different micro aerial vehicles (MAV), focussing on different aspects of the project goals. In the future, the sensors and properties of the two systems will be combined in a single setup. One MAV (Figure 1, left, referred to as ‘IGGKopter’) is designed to focus on the development of precise real-time trajectory estimation, and delivers directly georeferenced high resolution images, which will be used later to reconstruct surfaces and objects of the environment. The other MAV (Figure 1, right, referred to as ‘AISKopter’) is designed to focus on the autonomous flying aspects of the project, including obstacle

detection, collision avoidance, and path planning.

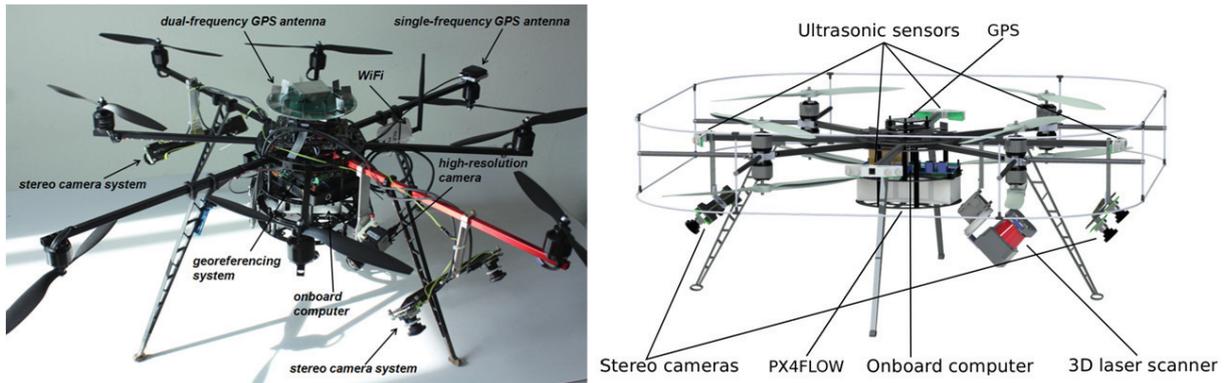


Figure 1: The two MAVs, as they are developed within the project. Left: IGGKopter. Right: AISKopter

Both MAVs are based on a MikroKopter OktoXL assembly kit of HiSystems GmbH, customized to address the specifications of the application. They both contain an onboard computer (Intel Core i7, 8GB RAM), based on an EPI-QM77 embedded PC board, providing sufficient computing power for complex tasks, such as image processing. The computer is connected to a ground station via Wi-Fi. Another component, which is identical in both MAVs, is a system of two stereo camera pairs pointing forward and backward with a pitch angle of 45° . Each camera has a fisheye lens providing a field of view of about 185° . Sampled with a frequency of 10Hz, these stereo cameras contribute to trajectory estimation, as described in Sec. 3, and also provide information about objects and obstacles in the vicinity of the vehicle (see Sec. 4). A major component on both MAVs is the direct georeferencing unit. This unit contains inertial sensors (accelerometer and angular rate sensors), a magnetometer, a barometer, a dual-frequency GPS receiver, a single frequency GPS receiver, and a processing unit in order to provide precise information about the vehicles motion state. Details on the georeferencing unit are given in Sec. 3.1.

The main mapping sensor, a 5 Megapixel camera with global shutter, is attached to the IGGKopter. This camera is mounted to a servo and thus can be pointed downward or sideways. Instead of the high resolution camera, the AISKopter contains a customized continuously rotating 3D laser scanner, measuring distances to objects up to 30m in almost all directions. It is used to build a local map of the environment and to detect obstacles (see Sec. 4). The vehicle also contains eight ultrasonic sensors, also used to detect obstacles, which may not be seen by the laser scanner (e.g. windows).

For the communication between different components of our system (onboard processing and ground station), we employ the communication infrastructure of the Robot Operating System ROS (Quigley et al., 2009). Those processing pipelines with large amounts of data and almost batch-like processing like the laser pipeline and the visual obstacles pipeline presented in the following are implemented using nodelets and efficient communication, respectively.

3 TRAJECTORY ESTIMATION AND DIRECT GEOREFERENCING

A crucial part of the project is the ability to estimate the current position and orientation of the MAV at any time and with a high precision. This is necessary for the reconstruction of surfaces and objects in a global reference frame. In classical airborne photogrammetry the images are oriented using ground control points (and GPS/INS data from the camera mount). We want to avoid the deployment of ground control points and therefore need a precise position and attitude determination of the camera for every taken image. This method is known as direct georeferencing. Within the project, we developed a direct georeferencing multi sensor unit, which is described in this chapter. An important specification of this unit is the ability to provide the motion information of the vehicle in real-time, which is necessary as input to the autonomous navigation software of the MAV. It is also planned to consider a priori map information during the flight, such as LOD2 models of buildings (Level of Detail 2, 3D model including roof structure), which are usually available in global coordinate systems. Since GNSS observations, which provide the main data source for the direct georeferencing, are not necessary available at all times, a trajectory estimation based on the two stereo camera pairs is also developed in order to improve and robustify the georeferencing performance.

3.1 Direct Georeferencing Unit

The sensor unit consists of dual-frequency GPS receiver (Novatel OEM 615), a single-frequency GPS receiver (uBlox Lea6T), a tactical-grade MEMS-based inertial measurement unit (Analog Devices ADIS 16488), a magnetic field sensor, a barometer and a real-time capable processing unit (Nations Instruments sbRIO 9606). It also contains a radio modem, enabling the reception of GPS data from a reference station. Figure 2 shows the prototype version of the sensor unit. It measures 11x10.2x4.5cm and has a weight of about 240g, without the GPS antennas.

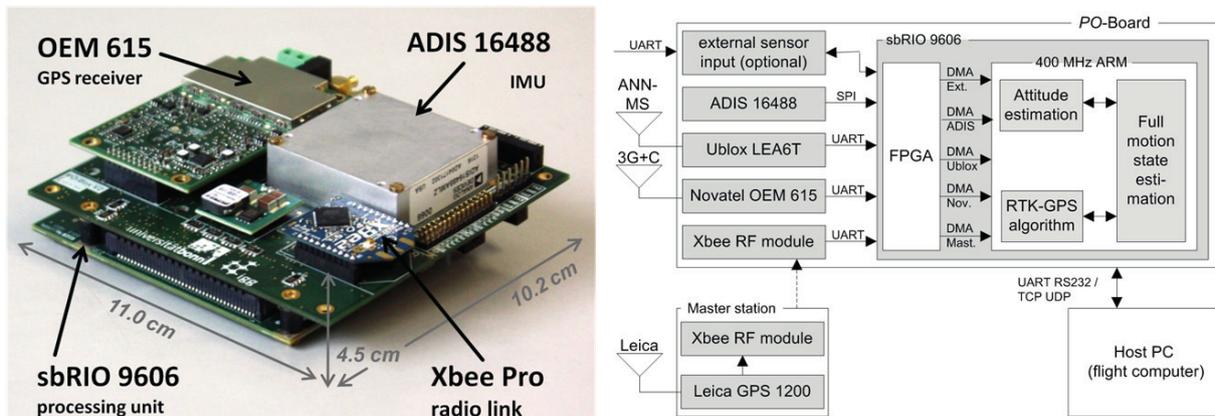


Figure 2: Left: Prototype of the direct georeferencing unit. Right: Block diagram of the sensor unit.

The main georeferencing sensor is the dual-frequency GPS receiver. Together with the GPS observations from a reference station, which are received via the radio modem, a real-time kinematic (RTK) GPS solution with an accuracy of about 1cm is calculated using a custom developed algorithm. The main task in RTK GPS processing is the determination of the carrier phase ambiguities (see Hofmann-Wellenhof et al., 1994 for details). This is done by first estimating real-valued ambiguities using GPS code and carrier phase observations (the so called float solution with an accuracy of about 20cm), and then trying to find the true integer-valued ambiguities using a search algorithm, leading to a so called fixed solution with an accuracy of about 1cm. Figure 3 (left) shows the results of a test flight under ideal conditions. To demonstrate the functionality of the ambiguity search, the ambiguities have been reinitialized in all epochs, still leading to a fix in nearly all cases. The flight trajectory in Figure 6 was performed under non-ideal conditions very close to a building. Here satellite occlusions yield more situations, where only a float solution could be obtained (red circles), leading to systematic errors in the order of 20cm. Currently, the position of the vehicle is estimated using GPS observations only, without any motion model of the device. In the future, the full motion estimation will also use the inertial sensor data and the results from the incremental bundle adjustment (see sec. 3.2) in a tightly coupled way. This and the possibility to adjust the estimation filter optimally to the system and the application are the reasons, why we decided to develop our own RTK GPS processing algorithm (Eling et al., 2013a), instead of using a freely or commercially available one.

The attitude of the vehicle is estimated using the inertial sensors, the magnetometer and a short baseline between the antenna of the single frequency receiver and the antenna of the dual frequency receiver. The magnetometer provides a true north reference, but this can be strongly disturbed by electrical currents on the MAV. So, we decided to implement a short single-frequency baseline of about 50cm on the vehicle, enabling the determination of the vehicles heading angle in the GPS coordinate frame (see Eling et al., 2013b for details). The estimation of position and attitude within a single estimation filter, as it is the subject of current research, will also improve the attitude accuracy, especially in conditions when GPS observations are not always available.

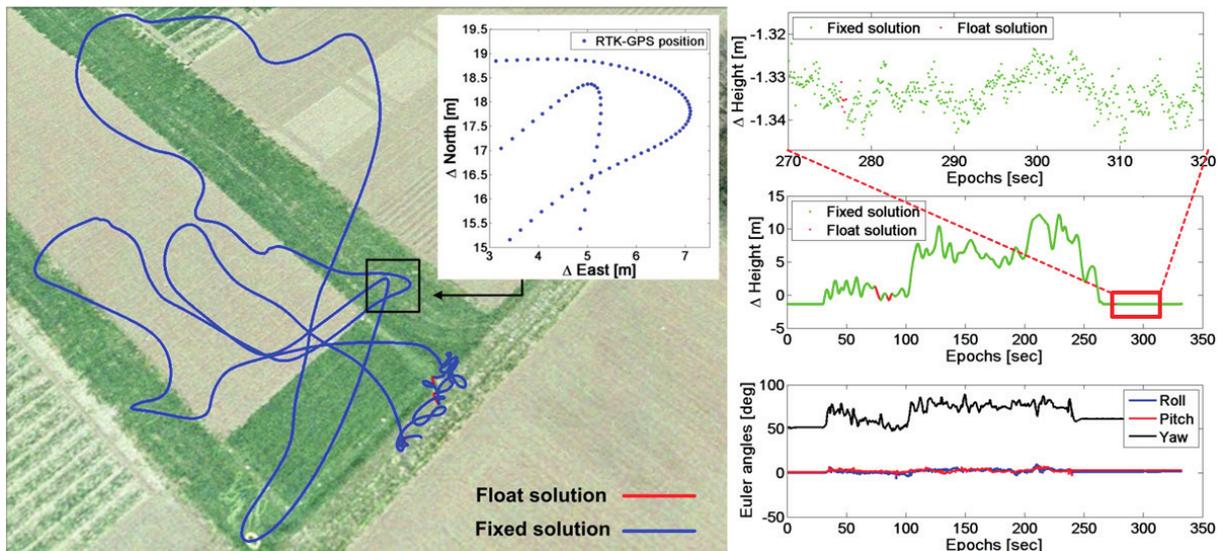


Figure 3: Example trajectory under ideal GPS conditions, calculated onboard during the flight. For nearly all times the carrier phase ambiguities could be fixed. The right side shows the height component and the attitude.

3.2 Incremental Bundle Adjustment

Bundle adjustment is the method of choice for orienting cameras and determining 3D coordinates of object points. It has a number of favourable 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 and variable reordering is done to prevent unnecessary fill-in. However, the computational expense of applying periodic batch bundle adjustments on each image of an image sequence would be too high as it grows with the number of involved camera motions and 3D points. For this reason, we process the images of the multi-camera system in a keyframe-based fast incremental bundle adjustment that makes a visual SLAM (Simultaneous Localization and Mapping) application for real-time on-board ego-motion estimation in an unknown scene feasible. The four cameras with 1,55mm fisheye lenses have a field-of-view of up to 185° and capture four monochromatic image sequences with a frame rate of 10 Hz in a synchronized way with an image resolution of 752×480 pixels. The cameras are mounted as two stereo pairs with a basis between the cameras of 20 cm, one looking ahead and one looking backwards with a pitch angle of 45°, see Figure 4.

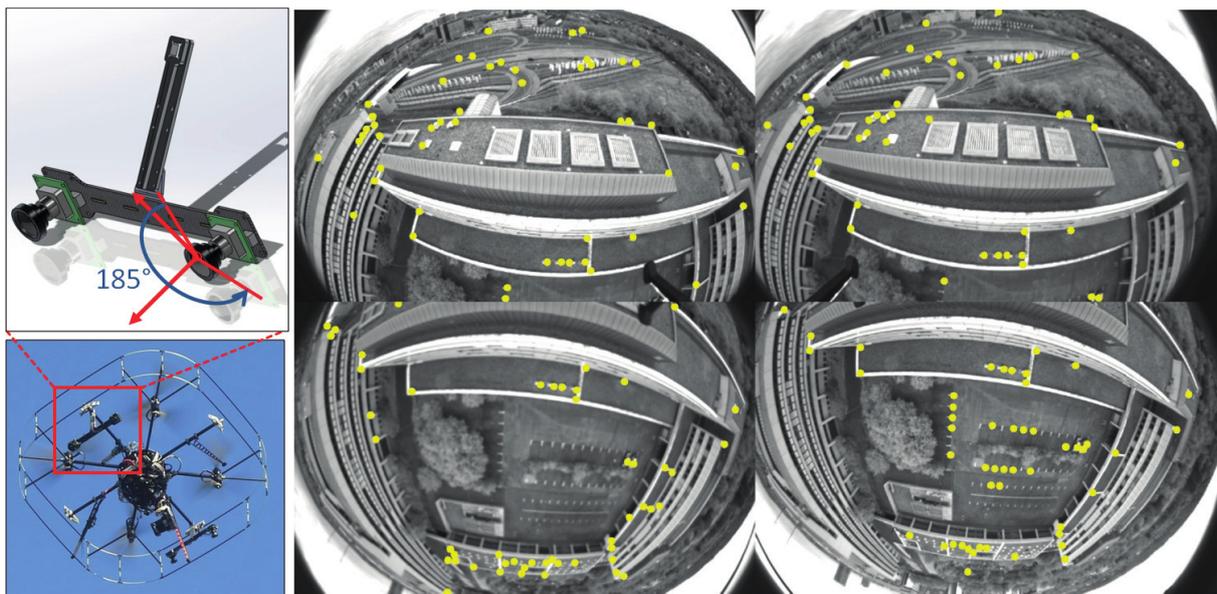


Figure 4: Left: Multi-camera system, as it is mounted to both MAVs. Right: Images of all four synchronized fisheye cameras. The yellow points are features, detected and tracked by the KLT tracker.

This configuration provides highly overlapping views and a large field of view. Figure 4 (right) shows an example frame set. The yellow points are interest points, which are corners in the gradient image with a large smallest eigenvalue of the structure tensor, cf. (Shi and Tomasi, 1994). The interest points are extracted

and tracked in the synchronized video streams of the individual cameras by the OpenCV implementation of the KLT tracker, that implements the Lucas-Kanade method with pyramids according to (Bouquet, 2000), and matched across the cameras, if possible. The feature points are converted into ray directions pointing to the observed scene points in the individual camera systems. For this, we model the fisheye lens with the equidistant-model described in (Abraham and Förstner, 2005), allowing for ray directions with an intersection angle equal or larger than 90° to the viewing direction.

Previous work yielded a rigorous batch bundle adjustment for omnidirectional and multi-view cameras for efficient maximum-likelihood estimation with scene points at infinity, called BACS (bundle adjustment for camera systems), see (Schneider et al., 2012). Classical bundle adjustments are not capable of far or even ideal points with small intersection angles, i.e. points at infinity, e.g. points at the horizon, which have been proven to be effective in stabilizing the orientation of cameras, especially their rotations. The calibration of the multi-camera system, which includes the translation and rotation from each camera into each other, is determined in advance according to (Schneider and Förstner, 2013).

Recently, we made BACS real-time applicable as described in (Schneider et al., 2013). Our keyframe-based method computationally selects only a small number of past frames to process a global incremental bundle adjustment step, which reduces the processing to some geometrically useful, tracked observations. The initiating frame set is chosen as the first keyframe set with a fixed pose \mathcal{M}_k , defining the coordinate system up to scale. The index k denotes a motion of a set of keyframes \mathcal{K}_k of all keyframe sets $\mathcal{K} = \{\mathcal{K}_k, k = 1, \dots, K\} \subset \mathcal{T} = \{\mathcal{T}_t, t = 1, \dots, T\}$, taken out of the set \mathcal{T} of all frame sets \mathcal{T}_t , t being the index referring to the time of exposure of a set of frames taken in a synchronized way. A new keyframe set with motion \mathcal{M}_k is initiated in case a minimal geometric distance to the last keyframe set with motion \mathcal{M}_{k-1} is exceeded, e.g. a translation of 1 m or a rotation of 30° , see Fig. 6.

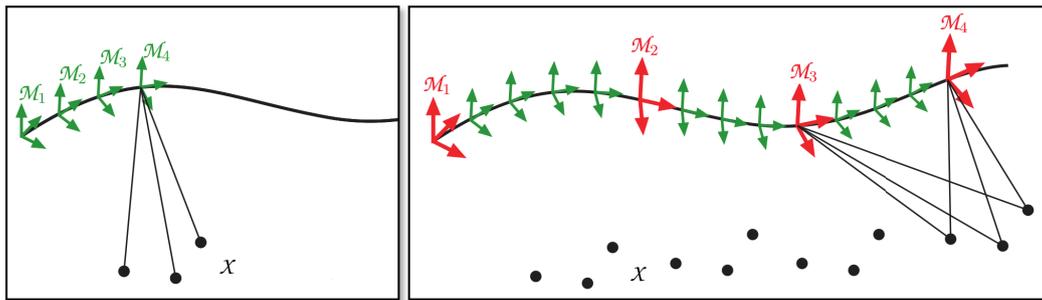


Figure 5: Incremental bundle adjustment. Left: At every frame, a new motion state \mathcal{M}_t (green) is calculated by resection using the scene points in the map \mathcal{X} . Right: After a certain motion distance, e.g. 1m or 30° , a keyframe is initiated (red). At every keyframe, a fast incremental bundle adjustment step is calculated to refine all keyframes \mathcal{M}_k in \mathcal{K} and scene points in \mathcal{X} .

A map in our context is a set of scene points $\mathcal{X} = \{\mathcal{X}_i, i = 1, \dots, I\}$, which is initialized by forward intersecting the matched ray directions in the stereo pairs in the initiating frame set. Robust estimates for the motion \mathcal{M}_t of the camera system in the map are computed at each time of exposure t via simultaneous resection of all cameras by using a generalized camera model with multiple projection centres. We determine the solution for the six pose parameters of \mathcal{M}_t by a robust iterative maximum likelihood-type estimation. Observations with large residuals are weighted down using the robust Huber cost function (see Huber, 1981). Using the pose \mathcal{M}_{t-1} as the initial approximate value, the estimation for \mathcal{M}_t converges in most cases after 10-20 msec (2-3 iterations). This allows the orientation of set of frames taken with a high frame rate. A track of observations getting a low weight is put on the blacklist. This procedure ensures a reliable data association as tracks on the blacklist are not considered in the following frames anymore.

The map is continually expanded as new keyframe sets are added. Initial values for new tracked scene points are obtained by forward intersection. Scene points in the map \mathcal{X} and poses in \mathcal{K} are updated and refined by the new observations in the incremental bundle adjustment step when a new keyframe is initiated. The incremental bundle adjustment step recalculates only entries of the information matrix, i.e. the normal equation matrix or inverse covariance matrix, which are actually effected by new measurements. To this end, we use the sparse non-linear incremental optimization algorithm iSAM2 provided by (Kaess et al., 2012), 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.

Tracking 200 feature points in each camera and setting a convergence criterion for the rotations to 0.5° and for the translations to 3 cm yields a very fast processing of the bundle adjustment that is always faster than one

second on a 3.6 GHz machine. In (Schneider 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.

Figure 7 shows an onboard processed trajectory of the MAV by our visual odometry algorithm (solid line with crosses) and the georeferencing unit (fixed GPS solutions are marked with green dots and float solutions with red dots). 273 keyframes with 17,175 observations of 864 scene points were initiated during the five minute long flight, whereby 4,803 frame sets were orientated by resection (not shown in Figure 7). To compare both trajectories the positions of the incrementally refined keyframes are transformed with a 7-parameter similarity transformation on the GPS positions. The differences between the GPS positions and the transformed positions of the keyframes are up to 60 cm due to the drift effects in the visual odometry. These results illustrate the potential of the visual odometry to bridge GPS losses of lock, to fix the ambiguities of float solutions and to detect GPS cycle slips. The relative accuracy of the estimated rotation parameters between succeeding sets of frames obtained by resection is always between 0.05° to 0.2° and 1 cm to 6 cm in translation.

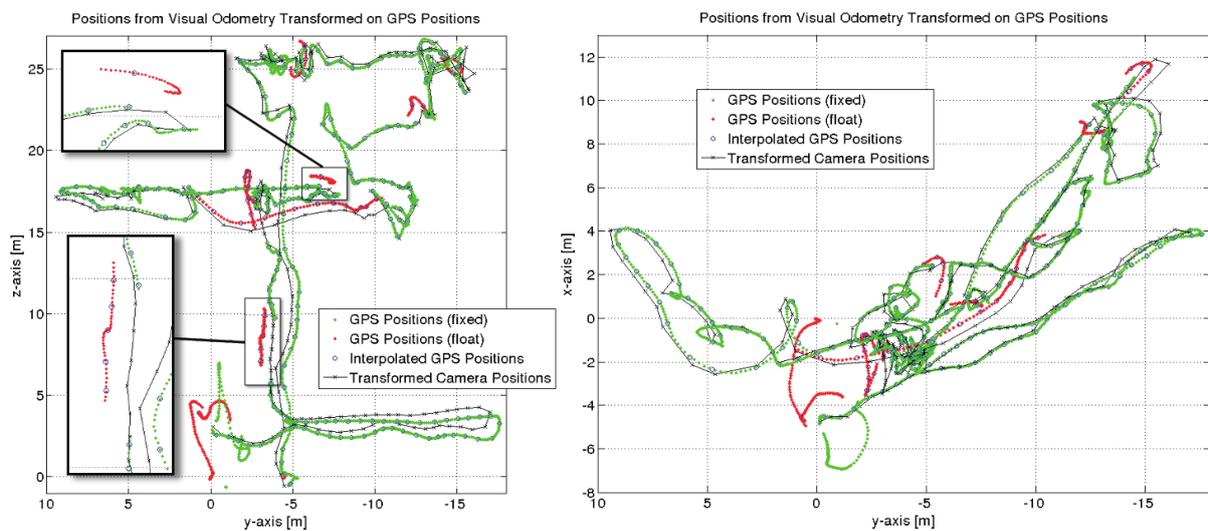


Figure 6: The onboard processed trajectories of the georeferencing unit and transformed incremental bundle adjustment. Left: Lateral view on the trajectories in the yz-plane, note the capability of the visual odometry to bridge float solutions of the georeferencing unit. Right: Top view on the trajectories in the xy-plane.

4 OBSTACLE DETECTION AND COLLISION AVOIDANCE

4.1 Sensors

To fulfil the objectives of mapping and inspection missions, our MAV has to navigate in the close vicinity of human-built and natural structures, e.g., buildings, power lines, and vegetation. These structures have very different characteristics and cannot be perceived reliably by individual sensor modalities, e.g., it is not possible to detect glass surfaces by means of optical sensors. Thus, we equipped our MAV with a multimodal sensor system to exploit the advantages of different sensors and alleviate their drawbacks.

4.1.1 3D Laser Scanner

Our primary sensor for obstacle avoidance is a continuously rotating 3D laser range finder that provides an almost omnidirectional view of the surroundings. Full 3D point clouds are acquired and processed with up to 2Hz. The scanner consists of a Hokuyo UTM-30LX-EW 2D laser range finder, mounted on a bearing (see Figure 7). The bearing is continuously rotated to gain a spherical field of view. The whole setup is mounted on the front side of the MAV and pitched downward by 45° which allows to maximize the field of view and to minimize the blind spot of the sensor, which is caused by the central core of the MAV. The scanner is able to measure up to three echoes of a single emitted light pulse. The number of echoes a light pulse reflects depends on the surface of the object, i.e., shape and reflectivity. For example, transparent material, vegetation or edges of buildings often reflect the light only partially yielding more than one echo. Hence, multi-echo detection is ideal for outdoor applications.

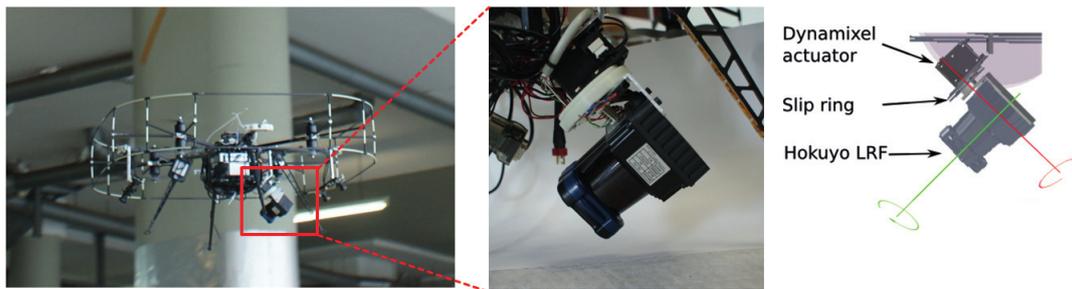


Figure 7: Rotating 3D laser scanner, mounted on the AISKopter. The sensor acquires full 3D point clouds at a rate of 2Hz.

For processing of the acquired 3D data, we form distinct 3D point clouds from the continuous data stream of the rotating laser scanner. We keep track of the rotation angle and start aggregating laser range scans to form a new 3D point cloud every half rotation. Since movement of the sensor during acquisition leads to a distortion of the 3D scan, we use visual odometry to compensate this effect. Figure 8 shows example point clouds from indoor and outdoor environments. We filter out measurements on the MAV itself by applying a simplified robot model for estimating which measurements coincide with the robot's body parts.

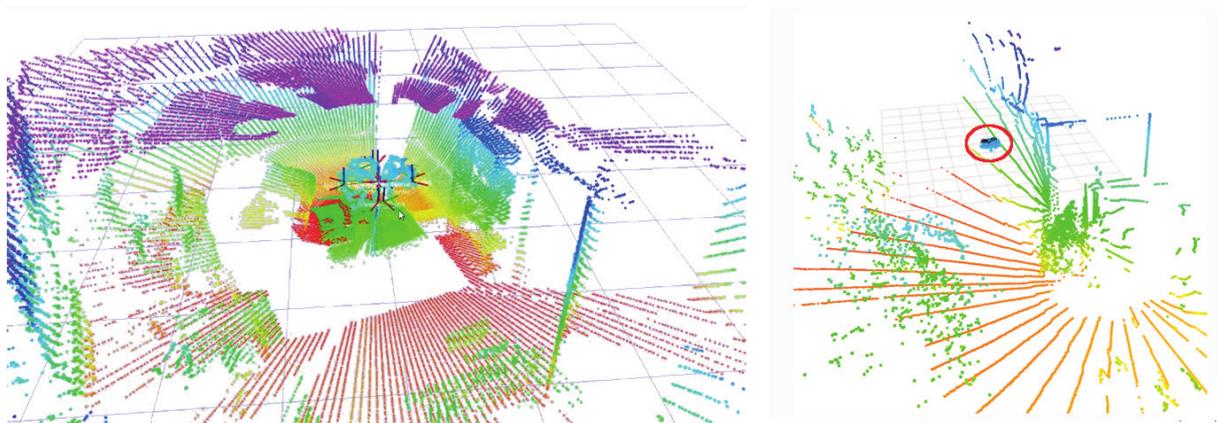


Figure 8: Point clouds acquired by the rotating laser scanner, mounted on the MAV. The colour indicates height. Left: indoor environment. Right: outdoor environment. The MAV is within the red circle.

4.1.2 Visual Obstacles

In addition to the 3D laser scanner, we obtain 3D point clouds of the surrounding environment using the two stereo camera pairs. Sparse visual obstacle detection and tracking is based on interest points, as described in Sec. 3.2. To determine the coordinates of the matched feature points via triangulation, we match tracked keypoints across the cameras using the known mutual orientations between the cameras within a stereo pair. The mutual orientations are determined in advance as described in (Schneider and Förstner 2013).

To match feature points in the overlapping images of a stereo camera pair, we determine the correlation coefficients between the local image patches at the feature points in the left and right images. Using the known relative orientation between the cameras within a stereo pair, we can reduce the amount of possible candidates to feature points lying close to the corresponding epipolar lines.

We assume feature points with the highest correlation coefficient to match, if it is above an absolute threshold, and - if there is more than one candidate close to the epipolar line - the closest-to-second-closest-ratio with the second highest correlation coefficient is lower than an absolute threshold.

The matched feature points are converted into ray directions pointing to the observed scene point in the camera frame system. An unknown scene point can be determined via the intersection of the observing camera rays in the camera frame using the known pose of the right camera in the camera frame defined by the left camera.

4.1.3 Ultrasonic Sensors

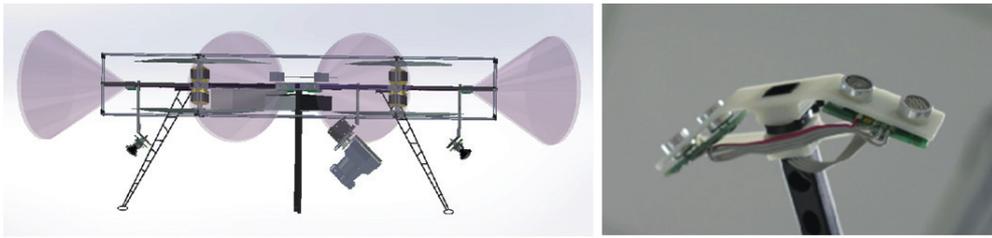


Figure 9: Setup and mounting of the ultrasonic sensors. These sensors are used to reliably detect small obstacles in the close vicinity of the MAV.

As neither the laser point cloud nor the visual obstacles are dense, and transparent obstacles cannot be measured optically, we have equipped our MAV with eight ultrasonic sensors. With a measurement range from 4cm to 6m, these sensors cover the space in the vicinity of the MAV. Ultrasonic sensors are particularly well suited for detecting close obstacles. In our setup, they are used as a fallback for dynamic obstacles suddenly appearing in the MAV's vicinity. Furthermore, their measurement principle with the wide sonar cone allows for perceiving obstacles that are hard to detect otherwise, e.g., wires and tree branches, as well as transparent obstacles such as windows. Referring to Figure 9, the ultrasonic sensors are mounted in a ring around the MAV in a star-like pattern with one pair of sensors at each of the four riggers of the frame. We filter out erroneous measurements by examining a sequence of measurements for each of the ultrasonic sensors, and only take a measurement into account for collision avoidance when it appears stable over several readings. In all our experiments, incorrect measurements were sparse and not persisting over multiple range readings.

4.2 Local Map Building and Collision Avoidance

4.2.1 Map Building

In order to fuse and accumulate measurements, we construct local egocentric obstacle maps. For each measurement and the corresponding 3D point, the individual cell of the map is marked as occupied. An exemplary map from an indoor map is shown in Figure 10. The map is used by our obstacle avoidance algorithm described in the next section. Along with the occupancy information, each cell also maintains its most recent 3D scan points in a ring buffer. These 3D points can be used for point-based scan processing, for example 3D scan registration. We aim for efficient map management for translation and rotation. To this end, individual grid cells are stored in a ring buffer to allow shifting of elements in constant time. We interlace multiple ring buffers to obtain a 3D map. In case of a translation of the MAV, the ring buffers are shifted whenever necessary to maintain the egocentric property of the map. For sub-cell-length translations, the translational parts are accumulated and shifted if they exceed the length of a cell. Since we store 3D points for every cell for point-based processing, single points are transformed in the cell's local coordinate frame when adding, and back to the map's coordinate frame when accessing. Since rotating the map would necessitate accessing of all cells, our map is oriented independent to the MAV's orientation. We maintain the orientation between the map and the MAV and use it to rotate measurements when accessing the map.

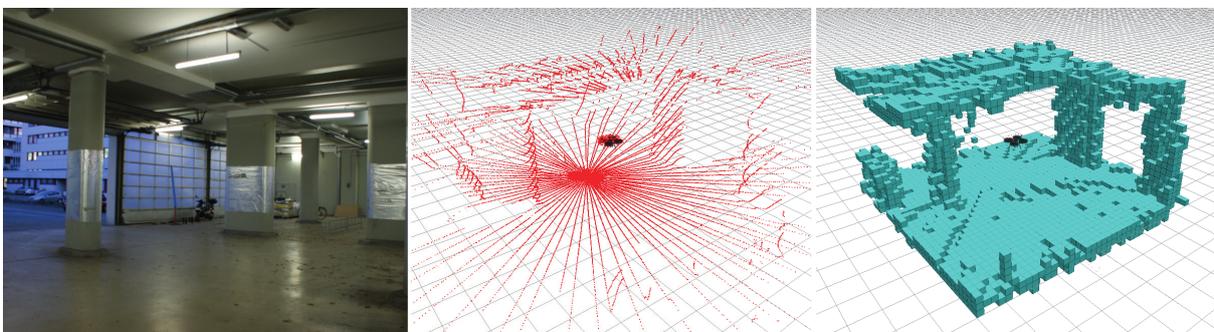


Figure 10: Middle: 3D laser scan of the indoor environment in the left figure (aggregated over 1s). Right: Resulting local grid-based map.

A complete 3D scan is aligned to the map by means of the Iterative Closest Point (ICP) algorithm (Besl and McKay, 1992). The ICP algorithm estimates a transformation between two point clouds, describing the displacement between them. The scan registration is triggered after acquiring a full 3D scan (i.e., a half rotation). When using the ICP algorithm for scan registration, corresponding points between the model and the current

point cloud are assigned, usually by building a space-partitioned data structure from the model point cloud. In contrast, we utilize our grid-based data structure for efficient correspondence assignment. Every point from a newly acquired 3D scan is directly assigned to a map cell in constant time. The closest point in terms of the Euclidean distance from the point list of this cell is initially assigned as corresponding point. Points in the neighbouring cells might be closer to the measured point than the initially assigned point. Consequently, we extend the search to neighbouring cells, if the distance to the initial assignment is larger than the distance to the border of a neighbouring cell.

4.2.2 Potential Field-based Collision Avoidance

Our concept for the navigation of the MAV is based on a multi-layer approach. Between low-level control and high-level planning layers, we employ a fast reactive collision avoidance module based on artificial potential fields (Ge and Cui, 2002) employing our local obstacle map described in the previous section. This enables the MAV to immediately react to nearby obstacles and deviate from a path that was planned based on a static allocentric environment model.

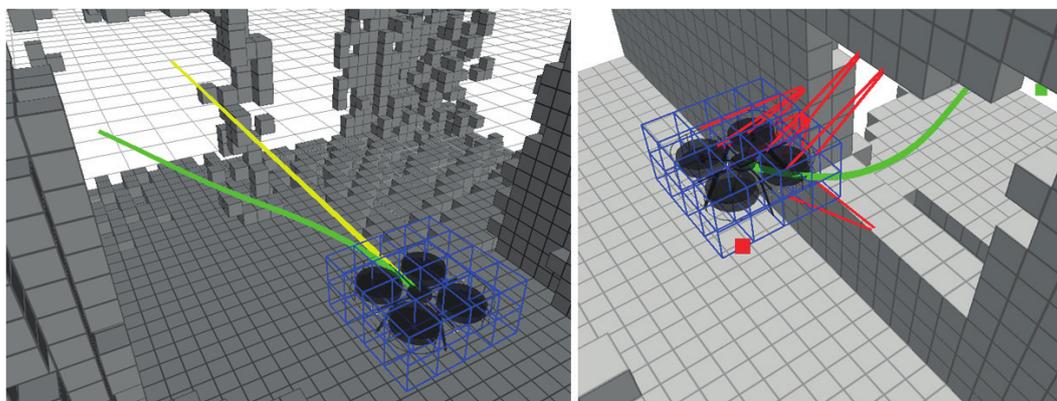


Figure 11: Visualization of our obstacle avoidance using a 3d grid map (left: real MAV, right: simulation). Left: A previously unknown obstacle obstructs the direct connection to the next waypoint (yellow). Right: Obstacles in range induce repelling forces (red lines). The trajectory is predicted into the future (green) given the current dynamic state of the MAV and the potential field. The blue cells depict our discretized robot model.

In contrast to the standard potential field-based approach, we relax the assumption that the robot is an idealized particle. We account for the shape of the MAV by discretizing it into cells of the size of our 3D grid map (see Figure 11). The centre points of these cells are individual particles to the algorithm. Hence, obstacles induce repulsive forces and the target waypoint induces an attractive force on each of these cells. Thus, multiple obstacles can induce forces on different parts of the MAV. Furthermore, we relax the assumption that the motion of a vehicle can be changed immediately. To overcome this limitation, we predict the MAV's future trajectory given the current dynamic state and the probable sequence of motion commands for a fixed discrete-time horizon (Figure 11). This time horizon is tightly bound by the property that MAVs can quickly stop or change their motion. To predict the trajectory, we employ a motion model of the MAV and the estimated resulting forces along the trajectory. The magnitudes of the velocity commands are calculated according to the predicted future forces. If a given force threshold is exceeded at any point of the trajectory, we reduce the velocity of the MAV.

In addition to guiding the MAV in a collision free manner to waypoints, our approach can act as a safety co-pilot to assist a human pilot. As the human pilot sends direct motion commands instead of coordinates relative to the MAV, we omit the attractive force in this case. Instead, we directly influence the control command given by the pilot if the MAV operates in the vicinity of obstacles. Repulsive forces induce a delta command that is added to the original control command, yielding a stop or deviation from the commanded direction.

We tested our collision avoidance in waypoint following scenarios in simulation. The tests include bounded environments with walls and unbounded environments, where the waypoints direct the MAV through window-like openings of different size. These experiments revealed that our collision avoidance approach is able to follow paths, if a relatively sparse trajectory is given that covers only the most crucial navigation points. The simulated MAV was able to fly through all passageways and windows of its size plus a safety margin. The prediction of the near future outcome of motion commands leads to smoother trajectories, keeping the MAV further away from obstacles than the same potential field approach without trajectory prediction. The occurring artificial forces during the flights were reduced to 68% of the forces without trajectory prediction. No collisions

occurred during these test runs. Experiments with the real MAV showed that our approach is able to successfully avoid obstacles, even if the commanded flight direction from an upper layer would lead to a collision.

5 CONCLUSIONS AND FUTURE WORK

In this contribution, we presented the concept and the current status of a project which aims to develop an autonomously flying MAV for online mapping purposes. We showed that a custom developed direct georeferencing unit is able to determine the position and the attitude of the vehicle in real-time with an accuracy of a few centimetres and a few degrees and is therefore able to capture georeferenced high-resolution images without the need for ground control points. We also showed that an incremental bundle adjustment algorithm, based on the images of four fisheye cameras, has the potential to bridge GPS gaps and to improve the attitude estimation of the MAV. Additionally, our robot possesses a number of sensors modalities, such as a rotating 3D laserscanner and ultrasonic sensors, which constantly update an ego-centric obstacle map. This map enables the MAV to operate in a mode, which goes far beyond the usual waypoint flight, as it is common for high-altitude MAV-based aerial photogrammetry. The system will be able to accomplish a higher-level mission, such as ‘fly around the house (which is for example given as a LOD2 model) and count all the windows’, while adapting to locally and temporally changing conditions, such as static or moving obstacles.

Subjects of future work are the integration of the vision based and the GPS/IMU based trajectory estimation to improve and robustify the real-time georeferencing capability of the MAV and the integration of all presented subcomponents and processes into a single mobile mapping workflow. The calibration and the evaluation of the system regarding motion estimation and mapping accuracy are also important topics of ongoing investigation. It should be noted here, that an essential part of the project ‘Mapping on Demand’ has not been discussed within this paper. This part deals with the problem of processing the georeferenced images to depth maps and point clouds, reconstructing objects and surfaces and extracting semantic information from the data. The integration of these tasks into the mapping work flow is also a major subject of future work.

As an intermediate result of the mapping procedure, we present in Figure 12 a georeferenced point cloud, which has been reconstructed from images, recorded by the high-resolution camera on the IGGKopter, which was manually controlled during the flight. Note, that this point cloud has been processed using a freely available software for dense image matching (pmvs2, see Furukawa 2010). More advanced methods will be used in the future, as described above.



Figure 12: Georeferenced point cloud, extracted from images using a freely available software package for dense image matching (Furukawa, 2010).

ACKNOWLEDGEMENTS

This work was supported by the DFG (Deutsche Forschungsgemeinschaft) as research unit FOR 1505 ‘Mapping on Demand’.

REFERENCES

- ABRAHAM, S. and FÖRSTNER, W.: *Fish-Eye-Stereo Calibration and Epipolar Rectification*. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 59(5), pp. 278-288, 2005.
- BESL, PAUL J., NEIL D. MCKAY.: *Method for registration of 3-D shapes*. Robotics-DL tentative. International Society for Optics and Photonics, 1992.
- BOUGUET, J.-Y.: *Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the Algorithm*. Intel Corporation Microprocessor Research Labs, 2002.
- DROESCHEL, D., SCHREIBER, M., BEHNKE, S.: *Omnidirectional perception for Lightweight UAVs using a continuously rotating 3D laser scanner*. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W2, 107-112, 2013.
- ELING, C., KLINGBEIL, L., WIELAND, M., KUHLMANN, H.: *A precise position and attitude determination system for lightweight unmanned aerial vehicles*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W2, 113-118, 2013a.
- ELING, C., ZEIMETZ, P., KUHLMANN, H.: *Development of an instantaneous GNSS/MEMS attitude determination system*. GPS Solution 17, pp. 129–138, 2013b.
- EVERAERTS, J.: *The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B1, Beijing, pp. 1187–1191., 2008.
- FURUKAWA, Y., PONCE, J.: *Accurate, Dense, and Robust Multi-View Stereopsis*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, Issue 8, Pages 1362-1376, August 2010.
- GE, S.S., CUI, Y.J.: *Dynamic Motion Planning for Mobile Robots Using Potential Field Method*. Auton. Robots 13, 3, pp. 207-222, 2002.
- HOFMANN-WELLENHOF, B., LICHTENEGGER, H., COLLINS, J.: *Global Positioning System, Theory and Practice*, S. 353. Springer Verlag, Wien, New York, 1994.
- HOLZ, D., NIEUWENHUISEN, M., DROESCHEL, D., SCHREIBER, M., BEHNKE, S.: *Towards multi-modal omnidirectional obstacle detection for autonomous unmanned aerial vehicles*. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W2, 201-206, 2013.
- HUBER, P. J.: *Robust Statistics*. John Wiley & Sons, New York, 1981.
- KAESS, M., JOHANNSSON, H., ROBERTS, R., ILA, V., LEONARD, J. and DELLAERT, F.: *iSAM2: Incremental Smoothing and Mapping Using the Bayes tree*. International Journal of Robotics Research, Vol. 31, pp. 217-236, 2012.
- LOCH-DEHBI, S., DEHBI, Y., PLUEMER, L.: *Stochastic reasoning for UAV supported reconstruction of 3D building models*. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W2, 257-261., 2013.
- MERZ, T. & KENDOUL, F.: *Beyond visual range obstacle avoidance and infrastructure inspection by an autonomous helicopter*. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), San Francisco, USA, 2011.
- NIEUWENHUISEN, M., SCHADLER, M., BEHNKE, S.: *Predictive potential field based collision avoidance for multicopters*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W2, 293-298, 2013.
- QUIGLEY, M., CONLEY, K., GERKEY, B., FAUST, J., FOOTE, T.B., LEIBS, J., WHEELER, R., NG, A.Y.: *ROS: An open-source robot operating system*. Proceedings of the ICRA Workshop on Open Source Software, 2009.

SCHNEIDER, J., LAEBE, T., FOERSTNER, W.: *Incremental real-time bundle adjustment for multi-camera systems with points at infinity*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W2, 355-360, 2013.

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

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

SHI, J. and TOMASI, C.: *Good Features to Track*. 9th IEEE Conference on Computer Vision and Pattern Recognition, pp. 593-600, 1994.

STEMPFHUBER, W., BUCHHOLZ, M.: *A precise, low-cost RTK GNSS system for UAV applications*. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII (1-C22), pp. 289-293, 2011.

XIANG, H., TIAN, L.: *Development of a low-cost agriculture remote sensing system based on an autonomous unmanned aerial vehicle (UAV)*. Biosystems Engineering 108, pp. 174-190, 2011.

Contact:

Dr. Lasse Klingbeil,

Institute of Geodesy and Geoinformation, Rheinische Friedrich-Wilhelms-Universität Bonn,

Nussallee 17, D-53117 Bonn, Germany

Email: klingbeil@igg.uni-bonn.de