Real-time Accurate Geo-localization of a MAV with Omnidirectional Visual Odometry and GPS

Johannes Schneider and Wolfgang Förstner

Department of Photogrammetry, University of Bonn
Nußallee 15, 53115 Bonn, Germany
johannes.schneider@uni-bonn.de, wf@ipb.uni-bonn.de

Abstract. This paper presents a system for direct geo-localization of a MAV in an unknown environment using visual odometry and precise real time kinematic (RTK) GPS information. Visual odometry is performed with an multi-camera system with four fisheye cameras that cover a wide field of view which leads to better constraints for localization due to long tracks and a better intersection geometry. Visual observations from the acquired image sequences are refined with a high accuracy on selected keyframes with a incremental bundle adjustment using the iSAM2 algorithm. The optional integration of GPS information yields long-time stability and provides a direct geo-referenced solution. Experiments show the high accuracy, below 3 cm standard deviation in position.

Keywords: visual odometry, incremental bundle adjustment, fisheye camera, multi-camera system, omnidirectional, MAV

1 Introduction

Micro aerial vehicles (MAVs) can operate from above in areas that are inaccessible from the ground such as hazardous environments. They become more and more important for example as low cost and flexible platforms for monitoring changes in agriculture, inspection of buildings and the wide field of surveillance purposes. In order to autonomously navigate in an unknown environment, the MAV must be able to perform self-localization. The position of the MAV is usually determined by a combination of GPS and IMU measurements, in which GPS maintains the long-term stability and provides geo-referencing. But an accurate position can not be guaranteed e.g. when the GPS signal is obscured. In this case a fully autonomous operating MAV has to rely on alternative localization systems. Visual odometry has proven to be very effective on MAVs, as cameras are lightweight and the orientation and position can be recovered using the on-line acquired image sequence.

The MAV we use for our work in the research project *Mapping on Demand* [12] is based on a MikroKopter OktoXL assembly kit. The on-board sensing of a lightweight MAV has to be designed with regards to its limitation in size and weight, and the limited on-board processing power requires highly efficient algorithms. Our platform is equipped with a GPS unit, an IMU, a high

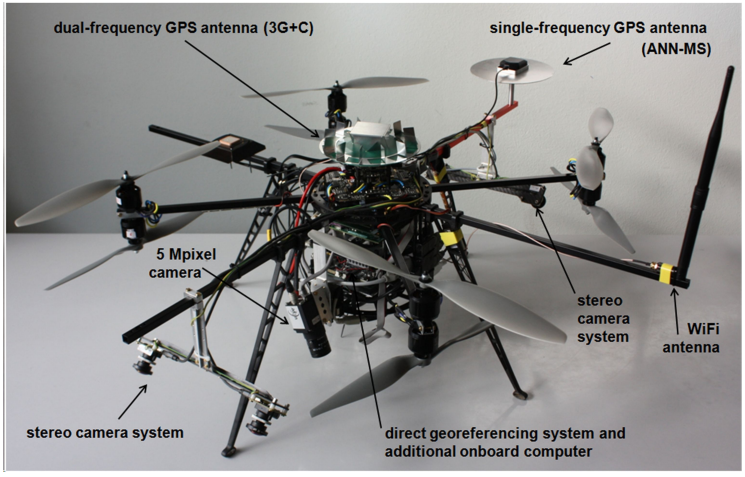

Fig. 1. The MAV based on the Okto XL frame set with its sensor and processing components setup

resolution camera and four fisheye cameras, which are mounted as two stereo pairs, one looking ahead and one looking backwards, providing a large field of view at each time of exposure. Computation power is provided by an onboard computer (Intel Core i7, 8GB RAM) which is based on an EPI-QM77 embedded PC board. The arrangements of the components is shown in Fig. 1. The two cameras are used besides obstacle detection [17] together with the GPS-unit and IMU for ego-motion estimation. The RTK GPS and IMU is processed on-board in a direct geo-referencing unit described in [5]. Under favourable conditions it can provide the position of the MAV with an accuracy of under 5 cm and the orientation under 1°. Our task is to fuse the geo-referenced ego motion with the visual odometry and to use it as an initial estimate for the orientation of the images of the high resolution camera, for near real-time surface reconstruction on a ground station, which is connected to the on-board computer through Wifi.

In contrast to single cameras and traditional stereo setups, an omnidirectional multi-camera setup covers a wider field of view, which leads to better constraints for localization. The four cameras with Lensagon BF2M15520 fisheye lenses with a field angle up to 185° capture four image sequences with a frame rate of 10 Hz in a synchronized way, see Fig. 2. The basis between the cameras of a stereo pair amounts 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 as it is statistically optimal and highly efficient in case sparse matrix operations are used. Factor graph based optimization frame works like g^{2}o have been shown to solve such problems efficiently by exploiting the characteristic structure [13]. Nevertheless, the computational expense rapidly grows with the number of involved images.

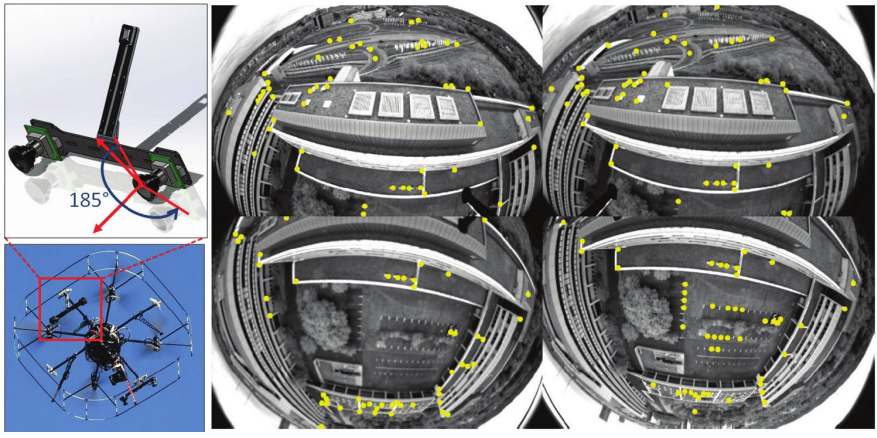

Fig. 2. Left: The omnidirectional multi-camera system as it is mounted on the MAV. Right: An example frame set consisting of four images taken with the four fisheye cameras

Many visual odometry systems uses PTAM [11] which runs in real-time by parallelizing the motion estimation and mapping task simultaneously and by using a keyframe-based bundle adjustment. However, it is not designed for large-scale outdoor environments and it is restricted to monocular and perspective cameras. To overcome the problem of continuously growing optimization problems, so called sliding window filters or local bundle adjustments are used that keep computational cost small enabling real-time applications [16], [6]. More recently the incremental optimization frame work iSAM2 was released [10]. The incremental optimizer 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. Only a subset of all contained variables are relinearized and solved which is realized by using the Bayes tree representation [9]. Fill-in is avoided through incrementally changing the variable ordering.

Multi-camera systems are regularly used for odometry, especially stereo camera systems, e.g. [15], [22] and more than two cameras e.g. in [14] or [8]. Fisheye-Cameras, see e.g. [1], catadioptric cameras, see e.g. [3] or omnidirectional cameras, see [23] ensure stable geometric positioning and full scene coverage due to their large field of view.

In this paper we treat the issue of visual odometry for real-time ego motion estimation using the synchronized images of the omnidirectional multi-camera system in a keyframe based fast incremental bundle adjustment using the iSAM2 algorithm [10]. To obtain longtime stability and to ensure the localization-system from sensor malfunctions, highly accurate GPS information can be integrated.

The paper is organized as follows. In the next section we present our system for visual odometry. The image processing for data acquisition and reliable data

association and the robust orientation of a set of frames taken in a synchronized way is presented. Further, we describe how the sparse non-linear incremental optimization algorithm iSAM2 is applied to avoid periodic batch bundle adjustment steps on sets of keyframes. The paper bases on previous work published in[19] where first results using the iSAM2 algorithm for visual odometry in a Matlab prototype are shown. We present in section 3 first results of the visual odometry and the integration of GPS. Finally we conclude and give an outlook on our future work in section 4.

2 Concept for Visual Odometry

2.1 Overview

Visual odometry is the process of determining the pose of the cameras in real-time by using the associated image sequences. Our real-time system uses feature points and consists of several steps:

1. The data acquisition and association detects feature points, performs the matching and provides camera rays associated with the previous and the other camera images.
2. The fast orientation of individual frames provides a robust solution and approximate values for the subsequent bundle adjustment and allows to select keyframes.
3. The incremental bundle adjustment uses the new information at a keyframe and merges it optimally with the previous information and optionally with GPS/IMU processed pose information.

The last step uses all available data and is therefore the most costly one. To ensure real-time capability it needs to be efficient and the previous steps have to guarantee outlier free information. Therefore we chose for the third step the software package iSAM2 for "incremental smoothing and mapping" and aim at efficient and robust methods for reliable data association. The steps are now described in more detail.

2.2 Data Association

The data association in our visual odometry system is based on interest points, which are tracked simultaneously in the individual cameras by running four threads in parallel. We use the OpenCV implementation of the KLT tracker: Interest points are corners in the gradient image with a large smallest eigenvalue of the structure tensor [21], that are tracked using the Lucas-Kanade implementation with pyramids according to [4]. Fig. 2 shows an example of 50 extracted feature points in the four images of the fisheye cameras.

Each tracked feature point is converted into a ray direction, i.e. a normalized direction vector, that points to the observed scene point in the individual camera system. The fisheye lenses are modelled with the equidistant-model described in [1] which allows for ray directions that have a larger angle than 90°

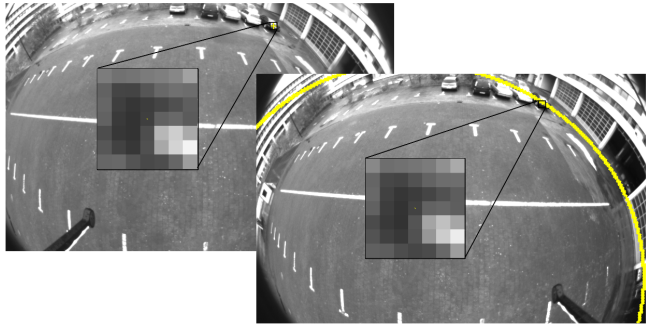

Fig. 3. Two frames taken in the left and right camera of a stereo pair. The extracted feature point in the left image on the rightmost car has the drawn epipolar line in the right image. The matching point in the right image lies on the indicated yellow line and the corresponding local image patches show a high correlation

to the viewing direction. The interior orientation of each camera is determined by camera calibration according to [2] using Chebyshev polynomials. Using the equidistant-projection and applying all corrections to the feature points we obtain image points ${}^e x$. The spherically normalized ray direction $\mathbf{x}^s$ can be derived by using the radial distance $r = |{}^e x|$ that grows with the angle between the viewing direction and the camera ray. The uncertainty of each image point can be transformed to the uncertainty of the associated spherically normalized ray direction $\mathbf{x}^s$ via variance propagation yielding $\Sigma_{\mathbf{x}^s\mathbf{x}^s}$. Note that the covariance matrix of the camera rays is singular, as the normalized 3-vector only depends on two observed image coordinates.

Feature points in the overlapping images of a stereo camera pair are matched by using the correlation coefficients between the local 7×7 image patches at the feature points in the left and right images. The rotation R_1^2 and translation $\boldsymbol{t}_1^2$ from the left into the right camera of the respective stereo pair is determined in advance according to [18]. We can use this information to reduce the amount of possible correspondences to feature points lyingclose to the corresponding epipolar lines, see Fig. 3, by statistically testing the contradiction to the coplanarity constraint $[\mathbf{x}_1^s, R_1^2\boldsymbol{t}_1^2, \mathbf{x}_2^s]$. We assume feature points with the highest correlation coefficient ρ_1 to match, if ρ_1 is above an absolute threshold, e.g. 0.8, and – if there is more than one candidate close to the epipolar line – the closest-to-second-closest-ratio $r = \rho_2/\rho_1$ with the second highest correlation coefficient ρ_2 is lower than an absolute threshold, e.g. 0.7. Finally we counter-check if this criterion holds also for all feature points in the left image if there are more than one feature points fulfilling the coplanarity constraint. In some rare cases this procedure leads to wrong matches, which can be detected later with a third observing ray from another pose.

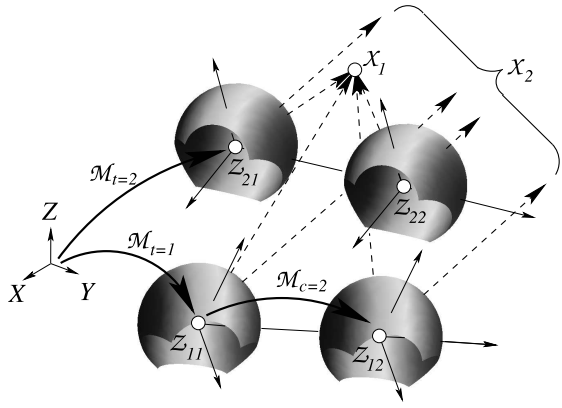

Fig. 4. A two-camera system with fisheye cameras $c = 1, 2$ with projection centers $\mathcal{Z}_{tc}$, rigid motion $\mathcal{M}_c$ and time-varying motion $\mathcal{M}_t$, having a field of view larger than 180° shown at two exposure times $t = 1, 2$ observing two points $\mathcal{X}_i, i = 1, 2$, $\mathcal{X}_1$ being close by and $\mathcal{X}_2$ at infinity

2.3 Localization of each Frame

We use a map which consists in our context of a set of scene points $\mathcal{X} = \{\mathcal{X}_i, i = 1, ..., I\}$. The map is initialized on the initiating frame set by using the matched ray directions in a stereo pair.

The initiating frame set is chosen as the first keyframe with pose $\mathcal{M}_k$. The index k denotes a motion of a set of keyframes $\mathcal{K}_k$ of all keyframe sets $\mathcal{K} = \{\mathcal{K}_k, k = 1, ..., K\} \subset \mathcal{T} = \{\mathcal{T}_t, t = 1, ..., T\}$, taken out of the set $\mathcal{T}$ of all frame sets $\mathcal{T}_t$, t being the index referring to the triggered time of exposure of a set of frames. If a pose from the direct georeferencing unit is available, it is used to define the coordinate system of the map. Otherwise the first keyframe is set with a arbitrary fixed pose $\mathcal{M}_k$.

After initializing the map, robust estimates for the motion $\mathcal{M}_t$ of the MAV in the map are computed in a continuously running thread at each time of exposure t via simultaneous resection of all cameras. We use a generalized camera model with multiple projection centres $c = 1, ..., 4$ and known system calibration, described by the motion $\mathcal{M}_c$ of each single camera from the camera system. The measurement equation, which considers mutually fixed single view cameras, allows the single cameras to be omnidirectional and allows for far or ideal scene points, reads as

$$\boldsymbol{v}_{itc} = J^{\mathsf{T}}(\mathbf{x}_{itc})\mathsf{N}\left([I_3|\mathbf{0}_3]\mathsf{M}_c^{-1}\mathsf{M}_t^{-1}\mathbf{X}_i\right) \tag{1}$$

with the homogeneous scene point $\mathbf{X}_i$, motion matrices M_t and M_c, and the observed and spherically ray direction $\mathbf{x}^s_{itc}$ and residual $\boldsymbol{v}_{itc}$ of scene point i in camera system k at time t, see Fig. 4, whereby $J(\mathbf{x}) = \text{null}(\mathbf{x}^{\mathsf{T}})$ and $\mathsf{N}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|$.

We determine the solution for the six pose parameters of M_t by a robust iterative maximum likelihood-type estimation down weighting observations with large

residuals by minimizing the robust Huber cost function [7]. The rigid motions $\mathcal{M}_c$ are determined in advance using a rigorous bundle adjustment estimating the system calibration [18]. We use the last determined pose $\mathcal{M}_{t-1}$ as the initial approximate value. The estimation of $\mathcal{M}_t$ then converges in most cases after 2-3 iterations. This procedure is very robust and time-effective, and allows to detect wrong data associations and to orientate with high frame rates. A track of observations getting a low weight is excluded from tracking and is not considered in the following frames any more.

2.4 Keyframe-based Incremental Bundle Ajdustment

The applied bundle adjustment refers to sets of keyframes, which reduce the processing to some geometrically useful, tracked observations. A new keyframe set with motion $\mathcal{M}_k$ is initiated in case a geometric distance to the last keyframe set with motion $\mathcal{M}_{k-1}$ is exceeded, e.g. 1 m or 30°. In case a new keyframe set is initiated, the observations x_{ikc} are used to update and refine the scene points in $\mathcal{X}$ and poses in $\mathcal{K}$ in the incremental bundle adjustment, which runs then in a separated thread.

The tracked observations are classified into two sets, $\mathcal{X}_1$ and $\mathcal{X}_2$, where $\mathcal{X}_1$ are the observations of scene points that are already in the map and $\mathcal{X}_2$ denotes those observing new scene points. The map is continually expanded as a new keyframe set is added. Initial values for new tracked scene points are obtained by triangulation with observations $\mathcal{X}_2$. A new scene point has to be observed at least on three keyframe sets to get affiliated in map using forward intersection. Care has to be taken with the sign: We assume the negative Z-coordinate of each camera system to be the viewing direction. The homogeneous representation of the scene points then need to have non-negative homogeneous coordinates $X_{i,4}$. In case of ideal points, we therefore need to distinguish the scene point $[\boldsymbol{X}_i; 0]$ and the scene point $[-\boldsymbol{X}_i; 0]$, which are points at infinity in opposite directions. Intersected scene points that show large residuals in the observations are put on the blacklist and deleted in the data association. Observations $\mathcal{X}_2$ are assumed to be revised from corrupted tracks via the former robust resection.

The map $\mathcal{X}$ and the set of poses in $\mathcal{K}$ is simultaneously refined using the bundle adjustment approach of [20]. The approach uses bundle of rays, allows for multi-camera systems and can numerically deal with points at infinity e.g. points at the horizon. The measurement equation reads as

$$\boldsymbol{v}_{ikc} = \boldsymbol{J}^{\mathsf{T}}(\mathbf{x}'_{ikc})\mathsf{N}\left([I_3|\mathbf{0}_3]\mathsf{M}_c^{-1}\mathsf{M}_k^{-1}\mathbf{X}_i\right) \tag{2}$$

and is not linear in the scene points and pose parameters of $\mathbf{X}_i$ and M_k. The linearization of the non-linear model at the actual linearization points is shown in detail in [20]. We use this model within the incremental solver iSAM2.

If available, pose information from the direct georeferencing unit for $\mathcal{M}_k$ are incorporated as direct observations of the pose parameters. For proper weighting the provided covariance information of the georeferencing unit is used. The offset in rotation and translation between the camera system and the direct

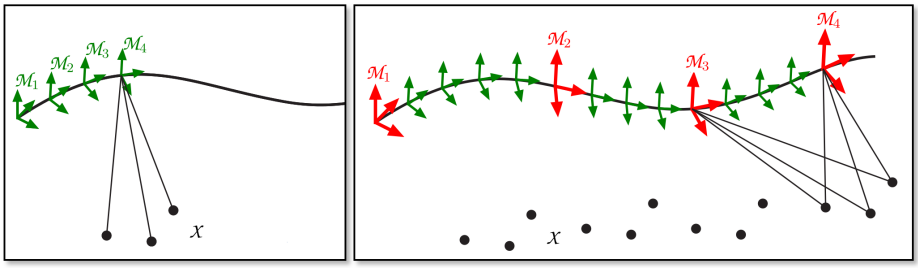

Fig. 5. On each acquired set of frames a new motion $\mathcal{M}_t$ (green) is determined by resection using the observed scene points in map $\mathcal{X}$. After a certain motion distance, e.g. 1 m or 30°, a keyframe is initiated (red). At every keyframe, an incremental bundle adjustment step is performed to refine all keyframes $\mathcal{M}_k$ in $\mathcal{K}$ and scene point in map $\mathcal{X}$

georeferencing unit is determined in advance by using GPS control points in a calibrating bundle adjustment. This way the offset has been determined with an accuracy of under 1 cm and 0.2°.

For our real-time application the processing of a new keyframe set $\mathcal{K}_k$ needs to be finished by the time the next keyframe set is added. For the first keyframe sets we use batch bundle adjustments as the map contains only a small number of scene points yet. After that the new information are incrementally merged with the previous information, yielding a fast optimal solution for the bundle adjustment using iSAM2 [10].

3 Experiments

Sensor data was recorded by the MAV during a 5 min flight in which a building was mapped with the high resolution camera. The visual odometry sets on average each 0.3 to 0.5 sec a new keyframe. This time in most cases is sufficient (1) to detect and track 200 feature points in each of the four cameras with a frame rate of 10 Hz, (2) to determine the spatial resections for each frame set, (3) to revise the tracks from outliers and (4) to execute the incremental bundle adjustment step.

The poses of the direct georeferencing unit are integrated as uncertain prior information on the keyframes to obain long-term stability and a georeferenced ego-motion. Under favorable conditions the incremental bundle adjustment can determine a real-time standard deviation in the position of about 1–2 cm, see Fig. 6.

The theoretical a posteriori estimated uncertainties are in general too optimistic. Therefore we have determined the ego-motion with visual odometry without using prior information from GPS. Using a similarity transformation on the GPS positions we can determine deviations between the independently estimated trajectories. The deviations between the keyframe poses are shown in the

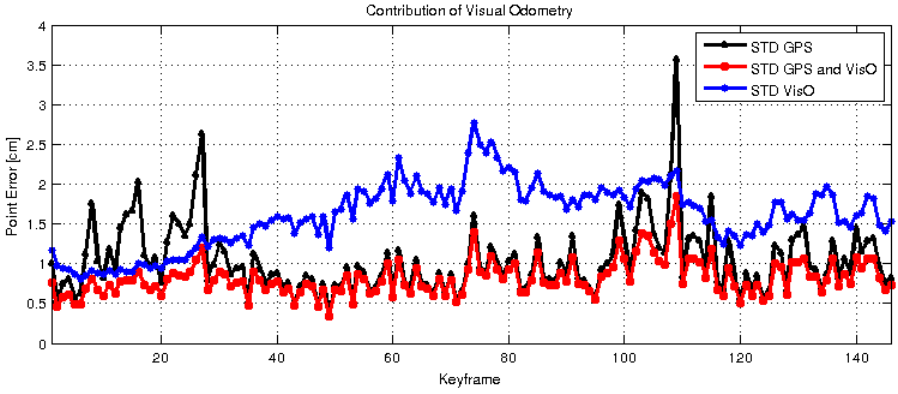

Fig. 6. Accuracy of the positions from GPS and from visual odometry, respectively shown as point errors $\sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_Z^2}$. The accuracy of the GPS measurements (black) and of the visual odometry which integrates the GPS measurements (red). The accuracy of pure visual odometry is derived from difference (blue): Apparently the visual odometry has a standard deviation below 3 cm and on an average is up to twice as uncertain as the GPS measurements, but temporarily it provides more accurate positions. The uncertainty of the integrated position is throughout less than 2 cm

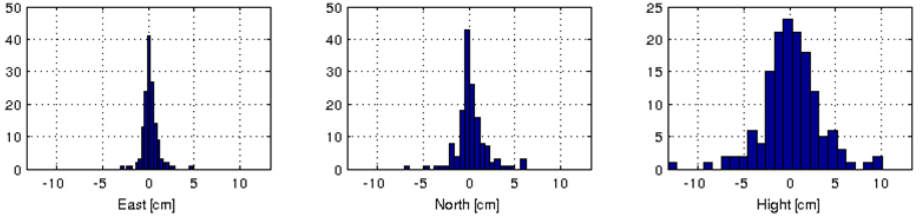

Fig. 7. The deviations between the keyframe positions from visual odometry, that were transformed on the GPS positions, to the GPS coordinates in the east, north and height component

histograms in Fig. 7. The histograms confirm the theoretical standard deviation from Fig. 6.

Figure 8 shows for this flight with 148 keyframes the cumulative percentage of the 2803 track lenghts: Most track lengths contain eight keyframes, 5 % of the tracks contain 35 keyframes and at least twelve tracks contain 100 keyframes. As a consequence we obtain an high long-term stability for the orientation angles. The obtained real-time accuracy of the rotations is throughout in the order of about 0.05–0.1°. Especially scene points close to infinity, i.e. points that are far away relative to the motion of the observing camera system, can be observed for a long time, which increases the accuracy of the camera rotation as shown in [20].

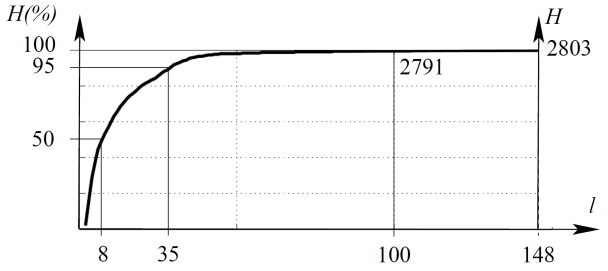

Fig. 8. Cumulative histogram $H(l)$ of the track lengths l of an flight with four fisheye cameras. The median and the 95-%-point is indicated

The iSAM2 algorithm is very efficient as it relinearizes and solves only subproblems that are actually influenced by the added observations. The optimization algorithm provides a solution which is in a statistical sense globally optimal. We compared the result with a offline batch bundle adjustment. The differences in the estimated pose parameters have been throughout in their estimated uncertainty. was tested to be within its uncertainty equals to the solution of a batch bundle adjustment.

4 Conclusions and Future Work

We presented our system for visual odometry performing a keyframe-based bundle adjustment for real-time structure and motion estimation in an unknown scene. Incremental bundle adjustment is performed by using the iSAM2 algorithm for sparse nonlinear incremental optimization in combination with a measurement equations that allows for multi-view cameras, omnidirectional cameras and scene points at infinity. The experiments show that a high accuracy level in the position can be obtained, which is in the order of RTK GPS. Long-time stability and a georeferenced position can be obtained by integrating GPS information. Using fisheye cameras and the inclusion of far points lead to stable poses. The inclusion of GPS is necessary in unknown environments for georeferencing. The visual odometry can bridge gaps due to interruption of the GPS signal with high accuracy.

Future work will focus on the issue of vibrations due to the rotors of the MAV on the mutual orientations of the multi-camera system. Further we have to solve the issue of estimating scene points which are at first near to the camera system and as the camera system moves away lying numerically at infinity.

Acknowledgments

This work was supported by the DFG-Project FOR 1505 Mapping on Demand. We thank the anonymous reviewers for their valuable comments and suggestions.

References

1. Abraham, S., Förstner, W.: Fish-eye-stereo calibration and epipolar rectification. ISPRS Journal of Photogrammetry and Remote Sensing (JPRS) 59(5), 278–288 (Aug 2005)
2. Abraham, S., Hau, T.: Towards autonomous high-precision calibration of digital cameras. In: Proc. of SPIE Videometrics. vol. 3174, pp. 82–93. San Diego, USA (Jul 1997)
3. Aliaga, D.: Accurate catadioptric calibration for real-time pose estimation of room-size environments. In: Proc. of the IEEE Intl. Conf. on Computer Vision (ICCV). pp. 127–134. Vancouver, Canada (Jul 2001)
4. Bouguet, J.Y.: Pyramidal implementation of the lucas kanade feature tracker. Tech. rep., Intel Corporation, Microprocessor Research Labs (2000)
5. Eling, C., Klingbeil, L., Wieland, M., Kuhlmann, H.: Direct georeferencing of micro aerial vehicles – system design, system calibration and first evaluation tests. Photogrammetrie – Fernerkundung – Geoinformation (PFG) (4) (Aug 2014), in press
6. Engels, C., Stewénius, H., Nistér, D.: Bundle adjustment rules. In: Proc. of the ISPRS Conference on Photogrammeric Computer Vision (PCV). Bonn, Germany (Sep 2006)
7. Huber, P.: Robust Statistics. John Wiley, New York (1981)
8. Kaess, M., Dellaert, F.: Visual SLAM with a multi-camera rig. Tech. Rep. GIT-GVU-06-06, Georgia Institute of Technology (Feb 2006)
9. Kaess, M., Ila, V., Roberts, R., Dellaert, F.: The Bayes tree: An algorithmic foundation for probabilistic robot mapping. In: Intl. Workshop on the Algorithmic Foundations of Robotics (WAFR). Springer Tracts in Advanced Robotics, vol. 68, pp. 157–173. Springer Berlin Heidelberg, Singapore (Dec 2010)
10. Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J., Dellaert, F.: iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree. Intl. Journal of Robotics Research (IJRR) 31(2), 217–236 (Feb 2012)
11. Klein, G., Murray, D.: Parallel tracking and mapping for small ar workspaces. In: Proc. of the Intl. Symposium on Mixed and Augmented Reality (ISMAR). pp. 1–10. Nara, Japan (Nov 2007)
12. 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: Proc. of the Intl. Conf. on Machine Control and Guidance (MCG). pp. 136–147. Brunswick, Germany (Mar 2014)
13. Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: G2o: A general framework for graph optimization. In: Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA). pp. 3607–3613. Shanghai, China (May 2011)
14. Maas, H.: Image sequence based automatic multi-camera system calibration techniques. ISPRS Journal of Photogrammetry and Remote Sensing (JPRS) 54(5-6), 352–359 (Dec 1999)
15. Mostafa, M., Schwarz, K.: Digital image georeferencing from a multiple camera system by gps/ins. ISPRS Journal of Photogrammetry and Remote Sensing (JPRS) 56(1), 1–12 (Jun 2001)
16. Mouragnon, E., Lhuillier, M., Dhome, M., Dekeyser, F., Sayd, P.: Generic and real-time structure from motion using local bundle adjustment. Elsevier Journal on Image and Vision Computing (IVC) 27(8), 1178–1193 (Jul 2009)

17. Nieuwenhuisen, M., Dröschel, D., Schneider, J., Holz, D., Läbe, T., Behnke, S.: Multimodal obstacle detection and collision avoidance for micro aerial vehicles. In: Proc. of the European Conf. on Mobile Robotics (ECMR). pp. 7–12. Barcelona, Spain (Sep 2013)
18. Schneider, J., Förstner, W.: Bundle adjustment and system calibration with points at infinity for omnidirectional camera systems. Photogrammetrie – Fernerkundung – Geoinformation (PFG) 4, 309–321 (Aug 2013)
19. Schneider, J., Läbe, T., Förstner, W.: Incremental real-time bundle adjustment for multi-camera systems with points at infinity. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. vol. XL-1/W2, pp. 355–360. Rostock, Germany (Sep 2013)
20. Schneider, J., Schindler, F., Läbe, T., Förstner, W.: Bundle adjustment for multi-camera systems with points at infinity. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. vol. I-3, pp. 75–80. Melbourne, Australia (Aug 2012)
21. Shi, J., Tomasi, C.: Good features to track. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). pp. 593–600. Seattle, USA (Jun 1994)
22. Strasdat, H., Montiel, J., Davison, A.: Visual slam: Why filter? Elsevier Journal on Image and Vision Computing (IVC) 30(2), 65–77 (Feb 2012)
23. Tardif, J.P., Pavlidis, Y., Daniilidis, K.: Monocular visual odometry in urban environments using an omnidirectional camera. In: Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS). pp. 2531–2538. Nice, France (Sep 2008)