Incremental Mosaicking of Images from Autonomous, Small-Scale UAVs

Saeed Yahyanejad Daniel Wischounig-Strucl Markus Quaritsch Bernhard Rinner
Institute of Networked and Embedded Systems
Klagenfurt University, AUSTRIA
<firstname.lastname>@uni-klu.ac.at

Abstract

Unmanned aerial vehicles (UAVs) have been recently deployed in various civilian applications such as environmental monitoring, aerial imaging or surveillance. Small-scale UAVs are of special interest for first responders since they can rather easily provide bird's eye view images of disaster areas. In this paper we present a hybrid approach to mosaick an overview image of the area of interest given a set of individual images captured by UAVs flying at low altitude. Our approach combines metadata-based and image-based stitching methods in order to overcome the challenges of low-altitude, small-scale UAV deployment such as non-nadir view, inaccurate sensor data, non-planar ground surfaces and limited computing and communication resources. For the generation of the overview image we preserve geo-referencing as much as possible, since this is an important requirement for disaster management applications. Our mosaicking method has been implemented on our UAV system and evaluated based on a quality metric.

1. Introduction

Unmanned aerial vehicles (UAVs) are widely used in the military domain. Advances in technology, material science and control engineering made the development of small-scale UAVs possible and affordable. Such small-scale UAVs with a total weight of approximately 1 kg and a diameter of less than 1 m are getting prominent in civilian applications and pose new research questions. These UAVs are equipped with sensors such as accelerometers, gyroscopes, and barometers to stabilize the flight attitude and GPS receivers to obtain accurate position information. Additionally, UAVs can carry payloads such as cameras, infrared cameras, or other sensors.

Thus, UAVs enable us to obtain a bird's eye view of an area which is helpful in applications such as environmental monitoring, surveillance and law enforcement, and disaster assessment and disaster management [15]. Obviously, each application domain has different requirements. Our goal is to support first responders in disaster assessment and disaster management since this is—in our opinion—the most challenging application domain. In disaster situations such as earthquakes or flooding, first responders can not rely on a fixed infrastructure and the available information (e.g., maps) may no longer be valid. The overall goal, hence, is to provide the first responders a quick and accurate overview of the affected area, typically spanning hundreds of thousands of square meters. This overview image is refined and updated over time and can be augmented with additional information such as detected objects or the trajectory of moving objects. When covering large areas at reasonable resolution from such small-scale UAVs, the overview image needs to be generated from dozens of individual images. Moreover, a number of UAVs equipped with cameras is employed instead of a single UAV to cope with the stringent time constraints and the limited flight time. The UAVs—flying at low altitudes of up to 100 m—provide images of the affected area which are stitched to an accurate overview image.

In this paper we describe different methods and their trade-offs for generating an overview image in order to surveil a certain area. We present a hybrid approach which allows to quickly mosaick the individual images and refine the alignment over time as more images are available. We have implemented and tested the hybrid approach in our UAV system.

The remainder of this paper is organized as follows. Section 2 gives an overview of related work in the domain of image mosaicking. In Section 3 we describe the intended use-case and give a rough overview of the UAV system. In Section 4 we formulate our main goal and identify the challenges we face. Section 5 goes into detail on image mosaicking, ranging from simple position-based mosaicking to pure image-based mosaicking. We propose a hybrid approach for image mosaicking which takes both the position information and the image data into account. In Section 6 we evaluate the described mosaicking approaches using a quality metric which is based on a spatial metric and a correlation metric. Section 7 finally concludes the paper.

2. Related work

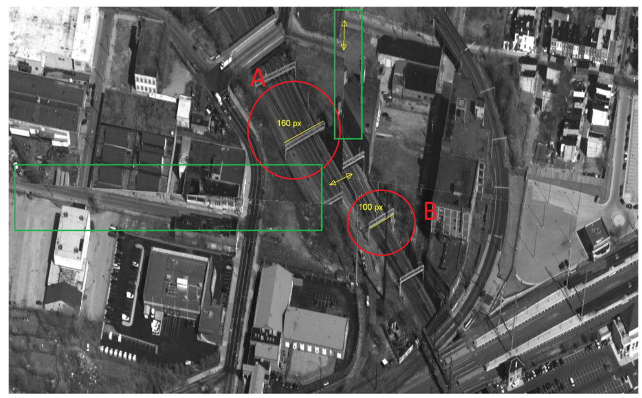
Much research has been done in the area of mosaicking of aerial imagery and surveillance over the past years. Many approaches have been proposed ranging from using low altitude imagery of stationary cameras and UAVs to higher altitudes imagery captured from balloons, airplanes, and satellites. High altitude imagery and on-ground mosaicking such as panoramic image construction are not in our area of interest since they are dealing with different challenges.

There has been a breakthrough regarding the seamless stitching in past years by exploiting robust feature extraction methods [23, 19, 3], depth-maps [10, 5], 3D-reconstruction of the scene, image fusion, and many other approaches (e.g., [20, 17]). Figure 1(a) shows a sample stitching of five sequential images generated by a SURF feature-based algorithm [3]. The result looks seamless at the stitching part but the obvious drawback is that the transformation performed on the images leads to a distortion in scales and relative distances. Such a traditional feature-based approach is difficult for our case because the generation of a geo-referenced image is hardly possible due to the scale and angle distortions as well as the error propagation over multiple images. The non-planar surface is one of the main reasons for this distortion, i.e., by using corresponding points at different elevation levels for the image registration. In principle, it is possible to improve the stitching result by using metadata, global alignment and bundle adjustment [17, 6, 18], but we need to know either accurate IMU (inertial measurement unit) data of the UAV's camera or accurate corresponding feature pairs.

A challenge of low altitude imagery and mosaicking for surveillance purposes is finding an appropriate balance between seamless stitching and geo-referencing under consideration of processing time and other resources. As shown in Figure 1(b), the scale difference as a result of different flying altitude resulted in several stitching errors. A similar error occurred in Figure 1(c) which was caused by inaccurate camera position or rotation.

Many approaches have been proposed to tackle these problems. Examples include the wavelet-based stitching [22], image registering in binary domains [7], automatic mosaicking by 3D-reconstruction and epipolar geometry [12], exploiting known ground reference points for distortion correction [13], IMU-based multi-spectral image correction [9], combining GPS, IMU and video sensors for distortion correction and geo-referencing [4] and perspective correction by projective transformation [21]. Some of these approaches differ from ours in a sense that they are considering higher altitude [4, 7, 12, 13, 21], while others are using different types of UAVs such as small fixed-wing aircrafts [11, 9, 8]. These aircrafts show less geo-referencing accuracy caused by higher speed and degree of tilting (higher amount of roll and pitch). Zhu et al. [24]

(a) Mosaicking five images using SURF features. The path borders (red lines) are supposed to be almost parallel. This type of error accumulates over multiple images if not compensated.



(b) Significant stitching errors induced by scale differences among images. Similar objects have different sizes (circles A and B), and there is a disparity in horizontal and vertical stitching (green rectangles) [2].

(c) Stitching disparities caused by inaccurate camera angle or position (red circles). [1].

Figure 1. Examples of image stitching errors.

performed an aerial imagery mosaicking without any 3D reconstruction or complex global registration. The difference of their approach is that they used the video stream which was taken from an airplane. Huang et al. [8] performed also a seamless feature based mosaicking using a small fixed-wing UAV, but no geo-referencing assessment was conducted.

Roßmann and Rast [16] also used small-scale quadrocopters. Their mosaicking results are seamless but lacking geo-referencing. No details about the mosaicking approach are presented.

3. System overview

The basic idea of our project is to deploy multiple small-scale UAVs to support first responders in disaster assessment and disaster management. In particular we use commercially available quadrocopters since they are agile, easy to fly and very stable in the air due to sophisticated on-board control. Each UAV is equipped with an RGB camera.

The intended use-case can be sketched as follows: The operator first specifies the areas to be observed on a digital map and defines the quality parameters for each area [15]. Quality parameters include the spatial and temporal resolution of the generated overview image, and the minimum and maximum flight altitude, among others.

Based on the user's input, the system generates plans for the individual drones to cover the observation areas [14]. Therefore, the observation areas are partitioned into smaller areas covered by a single picture taken from a UAV flying at a certain height. The partitioning has to consider a certain overlap of neighboring images which is required by the stitching process. Given a partitioning we can discretize the continuous areas to be covered to a set of so-called picture-points. The picture-points are placed in the center of each partition at the chosen height. The pictures are taken with the camera pointing downwards (nadir view).

The mission planner component generates routes for individual UAVs such that each picture-point is visited taking into account the UAV's resource limitations. The images together with metadata (i.e., the position and orientation of the camera) are transferred to the base-station during flight where the individual images are stitched to an overview image.

4. Problem definition and challenges

The major goal is to generate an overall image $I_{res,n}$ of the target area given a set of n individual images $\{I_i\}$. The overall image can be iterativly constructed as follows:

$$I_{res,0} = \mathbf{O}, \quad I_{res,i} = \mathit{Merge}(I_{res,i-1}, T_i(I_i)) \qquad (1)$$

where $\mathbf{O}$ is an empty background matrix, T is a transformation function and the *Merge* function combines the transformed image to the overall image.

This mosaicking can be described as an optimization problem, in which we need to find T_i in a way that it maximizes our quality function $\lambda(I_{res,i})$. This quality function, based on the system use case, balances the visual appearance and the geo-referencing accuracy. While in some applications it is more important to have a visually appealing overview image, other applications may require accurate geo-referencing in the overview image. We use a quality function that is a combination of the correlation between overlapping images and relative distances in the generated overview image compared to the ground truth (cf. Section 6 for more details).

In the following we summarize the most important challenges for solving our problem using images from low-flying, small-scale UAVs:

Low altitude and non-planar surface. When taking images from a low altitude the assumption of a planar surface is no longer true. Objects such as buildings, trees and even cars cause high perspective distortions in images. Without a common ground plane, the matching of overlapping images requires depth information. Image transformations exploiting correspondences of points at different elevations may result in severe matching errors.

Non-nadir view. Due to their light weight small-scale UAVs are vulnerable to wind influences requiring high-dynamic control actions to achieve a stable flight behavior. Even if the onboard camera position is actively compensated, a perfect nadir-view of the images cannot be provided.

Inaccurate position and orientation information. The UAV's auxiliary sensors such as GPS, IMU and altimeter are used to determine its position and orientation. However, such auxiliary sensors in small-scale UAVs provide only limited accuracy which is not comparable with larger aircrafts. As consequence, we can not rely on accurate and reliable position, orientation and altitude data of the UAV. Hence we have to deal with this inaccuracy in the mosaicking process.

Resource limitations. In our application the resources such as computation power and memory on-board the UAVs but also on the ground station are very limited. In disaster situations it is usually not possible to have a huge computing infrastructure available. The base-station typically will consist of notebooks and standard desktop PCs. But at the same time, we want to present the overview image as quick as possible.

Incremental refinement. The individual images are taken from multiple UAVs in an arbitrary order. An incremental approach is needed to present the user the available image data as early as possible while the UAVs are still on their mission. The more images are taken the better the overview image gets. This also means that a new image may require to adjust the position of already processed images to improve the overall quality.

5. Mosaicking approach

As described in Section 4 we must find the appropriate transformation T_i for each image I_i captured at a picture-point in order to solve our mosaicking problem. There are two basic approaches for computing these transformations: The *metadata approach* exploits auxiliary sensor information to derive the position and orientation of the camera which is then used to compute the transformations. In this case we assume that auxiliary sensor data (i.e., GPS, IMU, altitude and time) is provided for each captured image. The *image-based approach* only exploits image data to compute the transformations. In this section we first present the basic approaches considering the challenges of small-scale UAVs and then describe our hybrid approach which enhances metadata-based alignment with image-based alignment. The presented approaches vary in their resource requirements and their achieved results. Thus, they fit nicely to our problem domain.

5.1. Position-based alignment

A very simple and naive approach is to align the images based on the camera's position. Hence, for image alignment the world coordinates of the camera are mapped to corresponding pixel coordinates in the generated overview image.

Defining the origin of the overview image of the observed target area as $\mathbf{o}_{world} = (lat, lon, alt)^T$ in world coordinates, all image coordinates are related to this origin on the local tangential plane (LTP) by approximation to the earth model WGS84.

Given the camera's position we compute the area covered by the picture in world coordinates relative to the origin taking into account the camera's intrinsic parameters. The relative world coordinates are directly related to the pixel coordinates in the generated overview image.

An example of the resulting overview image is depicted in Figure 2(a) utilizing the placement function (Equation 1) with transformation T being a just a simple translation for each image. In this approach we assume reasonably accurate position information and a nadir view but do not take into account the camera's orientation. Obviously, effects introduced by non-planar surfaces can not be compensated with this approach.

5.2. Position- and orientation-based alignment

A more advanced approach is to extend the naive position-based alignment by compensating the camera's orientation deviation (i.e., roll, pitch, yaw angles). The placement function of the individual images to generate the overview image is the same as in Equation 1. But instead of considering only translation, we use a perspective transformation T with eight degrees of freedom.

If we assume a nadir view (i.e., neglecting deviation of roll and pitch angles) the transformation T is reduced to a similarity transformation.

5.3. Image-based alignment

Image-based alignment can be categorized into (i) pixel-based, and (ii) feature-based methods. The idea is to find transformations T_i and consequently the position of each new image which maximizes the quality function:

$$\lambda(\mathit{Merge}(I_{res,i-1}, T_i(I_i))) \tag{2}$$

The pixel-based approaches are computationally more expensive because the quality function is computed from all pixels in the overlapping parts of two images. Feature-based approaches try to reduce the computational effort by first extracting distinctive feature points and then match the feature points in overlapping parts. Depending on the chosen degree of freedom the resulting transformation ranges from a similarity transformation to a perspective transformation.

The benefit of this approach is that the generated overview image is visually more appealing. But on the other hand, the major disadvantages are that the search space grows with the number of images to be stitched and the images may get distorted (cf. Figure 1(a)).

5.4. Hybrid approach

We propose a combination of metadata-based (cf. Sections 5.1 and 5.2) and image-based methods (cf. Section 5.3).

The idea is to first place the new images based on the camera's position and orientation information on the already generated overview image. In the next step, we use image-based methods to correct for inaccurate position and orientation information and at the same time improve the visual appearance. Since we already know the approximate position of the image from the camera's position we can reduce the search-space significantly. Thus, we split the transformation T_i from Equation 2 into two transformation whereas the $T_{i,pos}$ represents the transformation based on the camera's position and orientation and $T_{i,img}$ represents the transformation which optimizes the alignment using the image-based method.

We favor transformations $T_{i,img}$ and $T_{i,pos}$ which maximize the quality function:

$$\begin{aligned}&\lambda(\mathit{Merge}(I_{res,i-1}, T_{i,img} \circ T_{i,pos}(I_i))),\\ &pos \in \{(u,v) | u \in [x_{min}, x_{max}], v \in [y_{min}, y_{max}]\}\end{aligned} \tag{3}$$

We limit the search space to a reduced set of possible positions based on the expected inaccuracy of position and orientation information (cf. Figure 3).

With this proposed approach we can generate an appealing overview image without significant perspective distortions and at the same time maintain the relative distances and geo-references in the overview image. Moreover, this approach can cope with inaccurate position and orientation information of the camera and thus avoid stitching disparities in the overview image.

6. Results

In this section we compare the results of the first three approaches (Sections 5.1, 5.2 and 5.3) with our hybrid approach (Section 5.4). This evaluation mainly focuses on the geospatial accuracy and image correlation which are specified in our quality metric (Section 6.1). We further compare the required computation times of all approaches which have been implemented in Matlab on a standard PC running at 2.66 GHz.

For the evaluation we used a rectangular round trip mission for which 40 picture points have been planned (cp. Figure 4). Images have been captured from a single UAV flying at an altitude of approximately 30 m. The overlap among adjacent images is about 60 %. A subset of 8 images is used to compare the stitching results of the first three mosaicking approaches (cp. Figure 2).

6.1. Quality evaluation

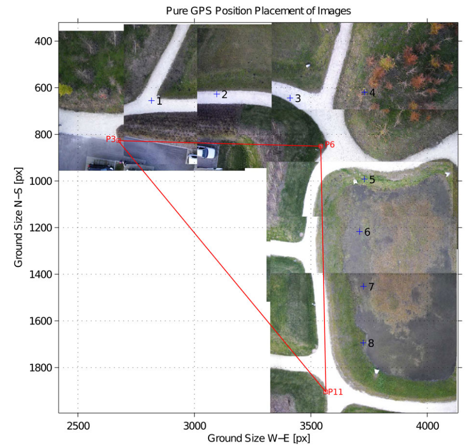
To evaluate the quality of the different mosaicking approaches presented in Section 5 the following metric for the overview image quality λ is defined:

$$\lambda(I_{res}) = \alpha \cdot \lambda_{spat}(I_{res}) + (1-\alpha) \cdot \lambda_{corr}(I_{res}) \qquad (4)$$

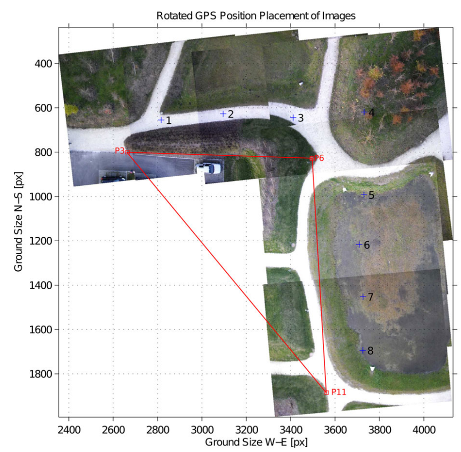
where

$$\begin{aligned} \lambda_{spat} &= \frac{1}{m}\sum_{i=1}^{m}\frac{1}{1+|\frac{d_i-\hat{d}_i}{d_i}|}, \\ \lambda_{corr} &= \frac{1}{n}\sum_{i=1}^{n}\frac{1+CC(\text{Overlaps}(I_{res,i-1},T_i(I_i)))}{2}, \\ CC(X,Y) &= \frac{\text{Covariance}(X,Y)}{\sigma_X\sigma_Y}, \end{aligned}$$

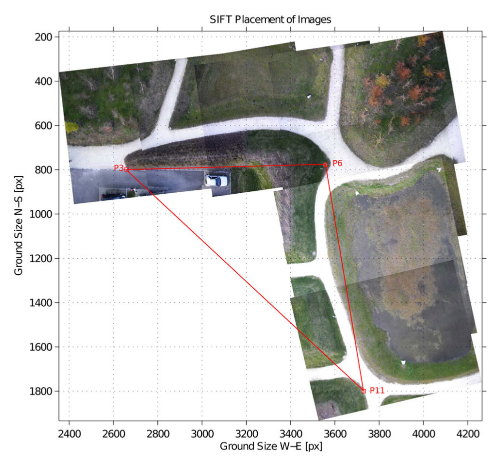
d_i is the actual distance measured between two ground control points, $\hat{d}_i$ is the estimated distance extracted from overview image and m is the number of considered distances. As it can be deduced from the equations, λ, λ_{spat} and λ_{corr} are all in the range of $(0,1]$. The total quality function λ is a weighted combination of λ_{spat} and λ_{corr} ($0 \leq \alpha \leq 1$). λ_{spat} represents the accuracy of spatial distances while λ_{corr} shows the correlation in areas of overlapping images, which is a measure for the seamlessness mosaicking. In our evaluations we set the weight $\alpha = 0.5$.



(a) Position-based alignment.



(b) Position- and orientation-based alignment.



(c) Image-based alignment using SIFT features.

Figure 2. Stitching results of basic image mosaicking approaches using 8 images of the round trip mission. The red triangle depicts the distances to compute the spatial accuracy. The units are given in pixels.

	Reference	Pos	Pos + Rot	Image	Hybrid
$\lvert\overline{P_3P_6}\rvert$ [m]	31	31.54	30.53	30.13	31.30
$\lvert\overline{P_6P_{11}}\rvert$ [m]	37.9	38.17	38.07	38.27	38.19
$\lvert\overline{P_3P_{11}}\rvert$ [m]	51.75	50.61	50.76	50.93	52.40
$\lambda_{spat}(I_{res})$ [%]		95.3	96.1	94.6	96.9
$\lambda_{corr}(I_{res})$ [%]		69.6	74.5	82.4	86.7
$\lambda(I_{res})$ [%]		82.4	85.3	88.5	91.8

Table 1. Spatial accuracy and quality parameters of the three basic and the hybrid mosaicking approaches.

6.2. Metadata and image based approaches

The quality of correlation λ_{corr} can easily noticed in the overview images in Figure 2, that is increasing by the complexity of the approaches.

In our evaluation we chose a triangle, spanning significant points (P_3, P_6, P_{11}) for simplified spatial evaluation in the reduced set of eight images. In Table 1 the measured distances $(\lvert\overline{P_3P_6}\rvert, \lvert\overline{P_6P_{11}}\rvert, \lvert\overline{P_3P_{11}}\rvert)$, the resulting spatial quality and the correlation quality are presented and combined according to Equation 4 to a final quality characteristic to compare the presented approaches.

Metadata-based approaches, like the position-based approach shown in Figure 2(a) and the position-based approach with rotation presented in Figure 2(b) retain georeferencing, if only similarity transformations are used. Image-based approaches even when restricting the matching function to a similarity transformation, as presented in Figure 2(c), show a good correlation quality.

The computation time for the whole set of 37 images in the scaled resolution of 400 by 300 px took $t_{pos} = 17.31$ s for position-based, $t_{pos+rot} = 18.33$ s with rotation, and increased dramatically to $t_{image} = 459.20$ s in the image-based alignment approach.

6.3. Hybrid approach

We used the complete round trip mission to evaluate the hybrid approach (Figure 4). However, three images were lost in the real UAV mission (cp. positions B, C and D in Figure 4) which reduced the overlap in these specific areas to approximately 20%. As shown the deviation of the last image to the starting point is not noticeable which implies that the relative distances are almost kept to a certain extent. The computation time was $t_{hybrid} = 136.28$ s for the whole set of images, which is significantly less than the image based approach (cf. Section 5.3). The total error range in the hybrid approach defines the search space in order to find the estimated position. By estimating the appropriate image position we compensate for the total error (GPS and camera tilting errors). Figure 3 helps to understand this concept better. We search inside this possible error range to find the best estimated position which maximizes our quality function best. In fact considering a case without any

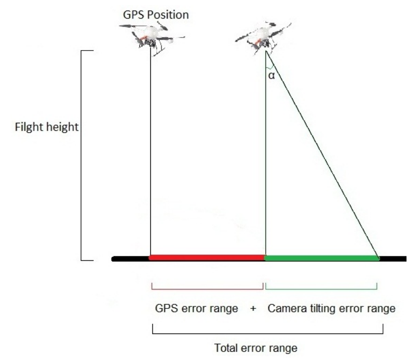


Figure 3. The red line shows the GPS error range (the real position is in this range). The green line shows the tilting error range. The sum of this two errors give us the total positioning error

GPS error and considering all views completely nadir, the hybrid algorithm will be reduced to a simple position based approach. The total error range we have used in Figure 4 is $GPS_{error} + tan(\alpha) \times$ height $\simeq 7m$ in real world distance at the ground level, which is approximately equivalent to $\frac{1}{4}$ of the image width. Yet, in a complete nadir view, orthogonality will be reduced when getting away from the optical axis. Somehow it gives us an idea that the middle parts of an images contain more reliable data. So for more pleasant result in our mosaicking, we make sure that the central part of each image under each picture-point is not masked by the border parts of other images.

In Figure 5 , the upper graph shows the relation between correlation of the overlapping parts of two adjacent images in different approaches. As we see the hybrid approach shows the highest correlation comparing to the others; the lower graph indicates the relative distance from the estimated position to the corresponding GPS position on each image in the hybrid approach. By comparing these two graphs, we see that if the estimated position of an image is close to its indicated GPS position it results in a higher correlation and vice versa.

Figure 6 shows a snapshot from our user interface. The operator defines the target area, then single images are placed iteratively over background map; the green line shows the flight path of the UAV. We can use any kind of existing geo-referenced digital maps (e.g., from Google earth or Microsoft virtual earth) or we can set an empty background.

7. Conclusion

In this paper we presented our system for mosaicking high-resolution overview images of large areas with high geometric accuracy from a set of images taken from small-scale UAVs. Although much research has been done on mosaicking of aerial imagery, the challenges in our application

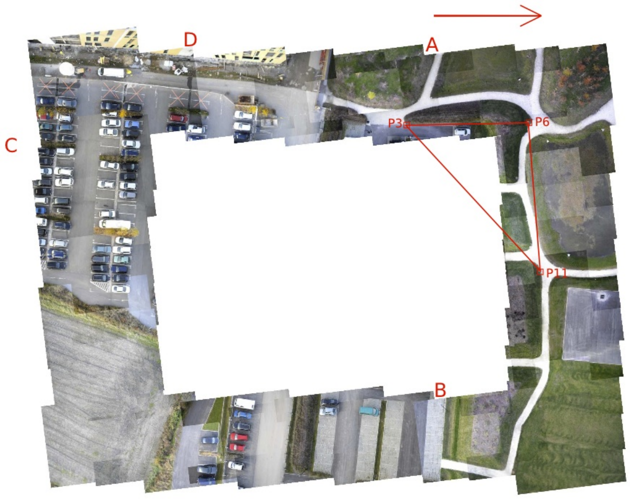


Figure 4. Mosaicking result of images taken from a round trip mission using the hybrid approach.

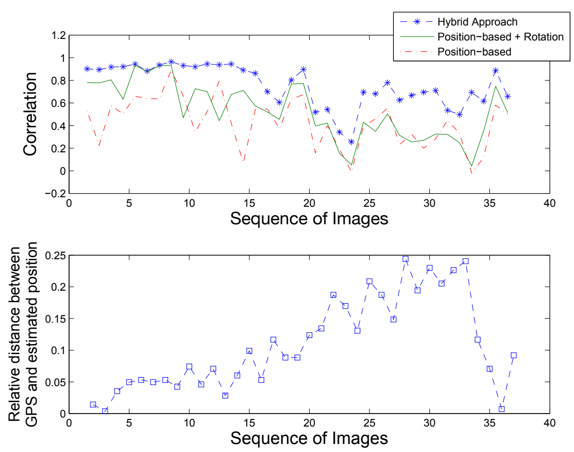


Figure 5. The upper graph depicts a comparison between correlation of the overlapping parts of two adjacent images in different approaches; the lower graph shows the relative distance between the estimated position and the GPS position

are significantly different since small-scale UAVs flying at low altitude pose new problems. We propose a hybrid approach that combines inaccurate information on the camera's position and orientation, and the image data. Thus, we can maintain geometric accuracy and at the same time enhance the visual appearance. The evaluations show that our approach results in a higher correlation between overlapping image regions and retains spatial distances with an error of less than 30 cm. The computation time for a set of 37 images is reduced by approximately 70 % compared to an image-based mosaicking.

Future works may include more dynamic and interactive methods of mosaicking images to increase the quality of the overview image, i.e., as new images are taken the transfor-

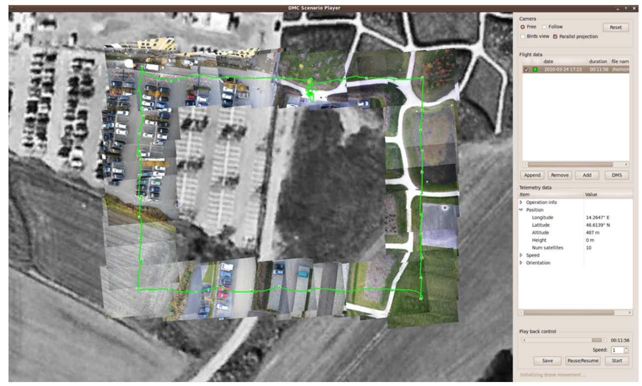


Figure 6. Screen shot of the GUI of our UAV system. The captured images are incrementally stitched over the—partially outdated—background image.

mation of already mosaicked images are refined. Moreover, we want to apply our proposed method also for larger areas and use images from multiple UAVs.

Acknowledgment

This work was performed in the project *Collaborative Microdrones (cDrones)* of the research cluster Lakeside Labs and was partly funded by the European Regional Development Fund, the Carinthian Economic Promotion Fund (KWF), and the state of Austria under grant KWF-20214/17095/24772.

References

[1] Geosense. `http://www.geosense.com.my/sample_uav_image/Index.htm`, March 2010. 2

[2] Persistent surveillance. `http://www.persistentsurveillance.com/graphics/Philadelphia.jpg`, March 2010. 2

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.*, 110(3):346–359, 2008. 2

[4] A. Brown, C. Gilbert, H. Holland, and Y. Lu. Near Real-Time Dissemination of Geo-Referenced Imagery by an Enterprise Server. In *Proceedings of 2006 GeoTec Event*, Ottawa, Ontario, Canada, June 2006. 2

[5] C. Çiğla and A. A. Alatan. Multi-view dense depth map estimation. In *IMMERSCOM '09: Proceedings of the 2nd International Conference on Immersive Telecommunications*, pages 1–6, ICST, Brussels, Belgium, 2009. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). 2

[6] Y. Furukawa and J. Ponce. Accurate Camera Calibration from Multi-View Stereo and Bundle Adjustment. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, number 3, pages 1–8, Hingham, MA, USA, 2008. 2

[7] X. Han, H. Zhao, L. Yan, and S. Yang. An approach of fast mosaic for serial remote sensing images from uav. In *FSKD '07: Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, pages 11–15, Washington, DC, USA, 2007. IEEE Computer Society. 2

[8] Y. Huang, J. Li, and N. Fan. Image Mosaicing for UAV Application. In *KAM '08: Proceedings of the 2008 International Symposium on Knowledge Acquisition and Modeling*, pages 663–667, Washington, DC, USA, 2008. IEEE Computer Society. 2, 3

[9] A. Jensen, M. Baumann, and Y. Chen. Low-cost multispectral aerial imaging using autonomous runway-free small flying wing vehicles. *Geoscience and Remote Sensing Symposium, IGARSS*, 5:506–509, 2008. 2

[10] S. B. Kang and R. Szeliski. Extracting view-dependent depth maps from a collection of images. *Int. J. Comput. Vision*, 58(2):139–163, 2004. 2

[11] G. B. Ladd, A. Nagchaudhuri, M. Mitra, T. J. Earl, and G. L. Bland. Rectification, georeferencing, and mosaicking of images acquired with remotely operated aerial platforms. In *Proceedings of ASPRS 2006 Annual Conference*, page 10 pp., Reno, NV, USA, May 2006. 2

[12] L. Lou, F.-M. Zhang, C. Xu, F. Li, and M.-G. Xue. Automatic registration of aerial image series using geometric invariance. In *Proceedings of IEEE International Conference on Automation and Logistics*, pages 1198 –1203, 2008. 2

[13] P. Pesti, J. Elson, J. Howell, D. Steedly, and M. Uyttendaele. Low-cost orthographic imagery. In *GIS '08: Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, pages 1–8, New York, NY, USA, 2008. ACM. 2

[14] M. Quaritsch, K. Kruggl, D. Wischounig-Strucl, S. Bhattacharya, M. Shah, and B. Rinner. Networked UAVs as Aerial Sensor Network for Disaster Management Applications. *e&i Journal*, 127(3):56–63, March 2010. 3

[15] M. Quaritsch, E. Stojanovski, C. Bettstetter, G. Friedrich, H. Hellwagner, M. Hofbaur, M. Shah, and B. Rinner. Collaborative Microdrones: Applications and Research Challenges. In *Proceedings of the Second International Conference on Autonomic Computing and Communication Systems (Autonomics 2008)*, page 7, Turin, Italy, September 2008. 1, 3

[16] J. Roßmann and M. Rast. High-detail local aerial imaging using autonomous drones. In *Proceedings of 12th AGILE International Conference on Geographic Information Science: Advances in GIScience*, Hannover, Germany, June 2009. 3

[17] H.-Y. Shum and R. Szeliski. Construction and refinement of panoramic mosaics with global and local alignment. In *Proceedings of Sixth International Conference on Computer Vision*, pages 953–956, 1998. 2

[18] G. Sibley, C. Mei, I. Reid, and P. Newman. Adaptive relative bundle adjustment. In *Robotics Science and Systems (RSS)*, Seattle, USA, June 2009. 2

[19] D. Steedly, C. Pal, and R. Szeliski. Efficiently registering video into panoramic mosaics. In *Proceedings of the Tenth IEEE International Conference on Computer Vision*, volume 2, pages 1300–1307, Los Alamitos, CA, USA, 17-21 2005. IEEE Computer Society. 2

[20] R. Szeliski. Image alignment and stitching: a tutorial. *Found. Trends. Comput. Graph. Vis.*, 2(1):1–104, 2006. 2

[21] W. H. WANG Yue, WU Yun-dong. Free image registration and mosaicing based on tin and improved szeliski algorithm. In *Proceedings of ISPRS Congress*, volume XXXVII, Beijing, 2008. 2

[22] C. Yuanhang, H. Xiaowei, and X. Dingyu. A mosaic approach for remote sensing images based on wavelet transform. In *WiCOM '08: Proceedings of the Fourth International Conference on Wireless Communications, Networking and Mobile Computing*, pages 1–4, 2008. 2

[23] Y. Zhan-long and G. Bao-long. Image registration using rotation normalized feature points. In *ISDA '08: Proceedings of the 2008 Eighth International Conference on Intelligent Systems Design and Applications*, pages 237–241, Washington, DC, USA, 2008. IEEE Computer Society. 2

[24] Z. Zhu, E. M. Riseman, A. R. Hanson, and H. J. Schultz. An efficient method for geo-referenced video mosaicing for environmental monitoring. *Mach. Vis. Appl.*, 16(4):203–216, 2005. 2