

Filling the holes – Potential of UAV-based photogrammetric façade modelling

Martin Reich, Manfred Wiggenhagen, Daniel Muhle

Institute of Photogrammetry and GeoInformation,
Nienburgerstr. 1, 30167 Hannover, Germany
eMail: reich@ipi.uni-hannover.de
URL: <http://www.ipi.uni-hannover.de>

Abstract. In this paper we examine the photogrammetric workflow for façade reconstruction based on imagery captured by an Unmanned Aerial Vehicle (UAV)-system¹. The UAV we use is a quadcopter which is able to capture the scene from many different perspectives. This makes the system a perfect device for imaging complex façade structures like balconies, jutties and even roof contours. The challenge for the processing is to deal with a not well structured block of images, which are taken by a low-cost consumer camera on a relatively unstable platform. Therefore, in the sense of correctness and completeness, the main advantage of UAV-acquired images lies on the latter, which is often the more important criterion for architectural photogrammetry. We compare the performance of the UAV-system with the results of classical terrestrial photogrammetry and laserscanning using the example of the façade of the Welfenschloss in Hannover. For camera calibration, orientation of the bundle and dense matching we use the open source software APERO/MicMac. Besides that, the near-real-time ability of this workflow for the purpose of a semi-automatic control of the UAV is evaluated in a future prospect.

1 Introduction

For civil applications, Unmanned Aerial Vehicles (UAV) are more and more interesting. Thanks to light weight positioning and navigation units as well as the improvements in control engineering, the UAV is able to follow beforehand planned paths nearly autonomously. Even manual steering does not need specific expertise. Due to the effort embedding these systems into civil aviation by changing the German Air Traffic Act, specific close-range photogrammetric tasks can be completed faster and more flexible, especially with a non-fixed wing UAV.

The production of modelled façades, based on three dimensional point clouds or orthographically straightened orthophotos such as façade plans, is of main interest in the field of architectural photogrammetry. For image acquisition next to UAV-based only terrestrial photogrammetry can be deployed, which often leads to occlusions due to the limitations in perspective. Using a crane or a forklift is expensive and data acquisition takes a long time.

Usually precision and completeness of reconstructed objects benefit from a higher

¹ Microdrone md4-200 <http://microdrones.de/index-de.php>

diversity of the acquired image data in terms of perspective and object distance. One of the first examples integrating UAVs into architectural photogrammetry is given by [11]. However, the usage of UAV-systems implies a significant drawback basically owing to a limited payload. In fact, today's (micro) UAVs are able to carry equipment with a total weight of just a few hundred grams. Therefore the user is straitened to a selection of small, low-cost consumer cameras. In [3, 4, 5] it is shown that low-cost cameras can be used for photogrammetric purposes if just moderate accuracy is required. In addition to that high frequency movements during exposure caused by vibrations of the system lead to a lower image quality which reflects on the results of the photogrammetric processing. Normally in architectural photogrammetry an accuracy of a few cm is sufficient, hence for evaluation this is the magnitude of precision we are aiming at.

The remainder of this paper is structured as follows. In section 2 we will give an overview of the used hardware, namely the UAV-platform and its equipment. Section 3 addresses the photogrammetric workflow and describes how the individual steps are solved in the course of this paper. The test data and its evaluation are examined in section 4. Section 5 summarises the results and offers a short future prospect in the sense of a near real-time application.

2 Hardware

2.1 UAV Platform

For our investigation we use a Microdrones md4-200 shown in figure 1. This vertical takeoff and landing quadrocopter has a carbon fibre body with a weight of less than one kilogram and a maximal payload of about 300 grams. Hence only small consumer cameras can be carried on the electrically driven mounting. Depending on takeoff-weight the potential flight duration is up to 30 minutes with a cruising speed of about 5 m/s, which allows covering a reasonable area for the purpose of façade modelling.



Fig. 1 Microdrones md4-200

The system either flies manually steered via remote control or follows given GPS waypoints autonomously. GPS as well as barometric and inertial measurements support the user in the manner of stabilising the platform. Position hold and dynamic position hold allow in-place hovering and GPS-based moving respectively, which is especially useful under windy conditions. For the purpose of flight planning Microdrones developed the software mdCockpit. However, since on the one hand we continuously shoot pictures with an interval of two seconds (see section 2.2.)

and therefore are not able to plan photo-stops. On the other hand, taken a small object distance of several meters into account, we prefer manual control to avoid a collision, since a GPS path does not fulfil this accuracy purpose. So we do not make use of a pre-flight path planning.

Ground control is accomplished via a specifically designed ground station and a laptop. The station is able to receive and convert the down streamed analogue PAL signal of the carried camera and further flight parameters like positioning accuracy and battery level so that it can be displayed by mdCockpit on the laptop. With this information the user is able to navigate the UAV safely and to adjust the view of the camera with respect to the object of interest.

2.2 Camera

The first camera we tried is a Pentax Optio S1, which has 14 megapixels on a 1/2.3" CCD, resulting in a pixel size of $1.4 \times 1.4 \mu\text{m}^2$. For a remote access of the camera the infrared interface is triggered, hence capturing and zooming can be performed via the remote control of the UAV. Unfortunately after several test flights we were not satisfied with the quality of the resulting imagery. Most of it was blurred and unfeasible for matching. Due to the absence of an editable exposure time we investigated for an alternative camera.

In the course of this paper we use a Canon Digital IXUS 100 IS with a 12 megapixel 1/2.3" CCD sensor with a focal length of about 5.9 mm. Next to its feasible weight of less than 150 grams the main reason we rely on this camera is the availability of the Canon Hack Development Kit (CHDK). This feature is able to handle scripts for automatic shooting and moreover allows us to specify exposure time. Furthermore [3] evaluates the camera with respect to the generation of Digital Surface Models (DSM) out of fixed-wing UAV photographs with satisfying results.

For the reason of navigation and camera adjustment during the flight we installed a small PAL camera for control purposes.

3 Photogrammetric workflow

The photogrammetric workflow handles every processing step that is necessary for the generation of 3d information and orthophotos out of imagery. This is namely the calibration of the camera, the relative orientation of each viewpoint with homologous points of neighbouring image pairs, the dense matching and the projection of the images onto the 3d model for achieving an orthophoto.

Recently there were developed several freely available solutions for this workflow by the computer vision community, for instance Bundler [1, 10] in combination with PMVS [2]. Bundler is able to orient a great amount of highly redundant images with diverse camera geometries. However, since Bundler uses a very basic implementation for modelling the camera distortions and further we use primarily the same camera, we picked an alternative solution. The open source software bundle APERO and MicMac developed by IGN [7, 8] offers a variety of tools to cope with this workflow. Interaction with the software works via xml-files, which provides the opportunity to adjust important parameters.

In the following the most important steps of the workflow are examined in more detail. A schematic view of the workflow with input- and output-data can be seen in figure 2.

Calibration

The intrinsic parameters describe the imaging process of a specific camera. Next to focal length and principle point of the optical axis, specifying the pinhole model, there are several parameters describing the lens distortion. The number of the latter depends on the distortion model and varies between two for a basic correction of radial components and more than ten for high accuracy purposes.

Usually the camera geometry of low-cost consumer cameras is unstable and not suited for high precision measuring. Therefore we use a self-calibration included in the bundle adjustment for image orientation. The parameters are initialised as follows: From the exif-header of the images the nominal value for the focal length and the centre of the image for the principle point are used, the distortion parameters are initialised as zero. First the intrinsic parameters are frozen. Once all the images are oriented the intrinsic parameters are released and estimated in a final iteration.

Orientation

The calculation of the orientation of each camera station is based on homologous tie-points seen in several images, which deal as observations in the adjustment. Parameters are camera positions and viewing directions, object coordinates of tie-points and intrinsic parameters of the camera geometry if they are not estimated in advance. To handle singularities in the definition of the coordinate system, which lead to a datum defect, one can include exterior orientation parameters like GPS positions of the camera or terrestrial measurements. Otherwise the coordinate system is adopted to be random, based on the first image that is oriented. More details of the algorithm can be found in [6].

Dense Matching

The dense matching step needs the oriented images and generates a dense point cloud by pixel wise correlation out of regions that are seen in at least two different images. MicMac uses a pyramidal multi-resolution approach which is based on the idea that real corresponding image regions should appear independently from the scale of the image. The user for instance controls the size of the correlation window, the number of different pyramidal layers and can choose between different regularisation algorithms.

Orthophoto

Caused by central perspective capturing, images per se are distorted, possess occlusions and cannot be used as façade plans. Thus they have to be rectified using the information gathered from orientation and dense matching steps. With the three dimensional data of the object one is able to project each image onto the modelled object, detecting occluded parts. Using information from all the images in terms of stitching them together, occluded parts can be minimised. Obviously, the higher the diversity of views the lower the number and size of occlusions.

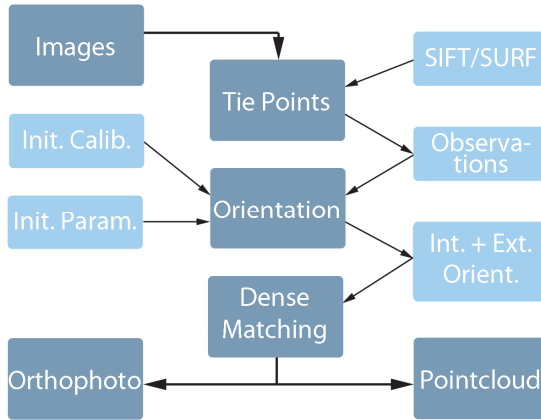


Fig. 2 Schematic view of the photogrammetric workflow

4 Experimental results

4.1 Test data

Since we are aiming at the examination of the potential of UAV-based façade modelling, we firstly wanted an object that would be quite ambitious to model with terrestrial images only. Secondly, flying with the UAV should not be a problem, so for legitimate reasons we preferred a public building. In conclusion we selected the Welfenschloss, main building of the University of Hannover. The building's façade has a lot of different structures like balconies, frets and diversely formed windows on a total length of about 150 m, which provides a good target for evaluating the results.

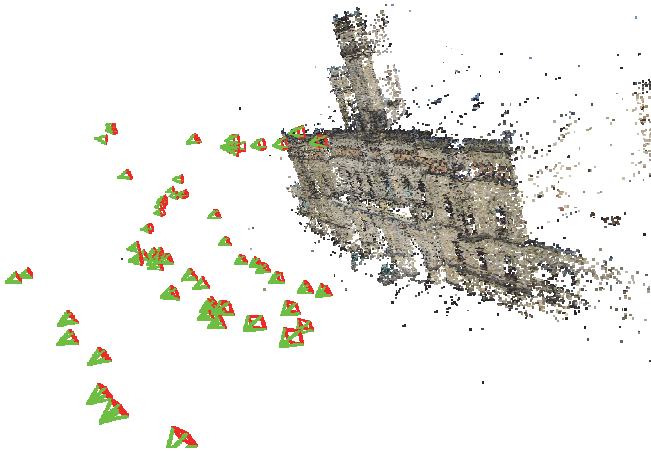


Fig. 3 Exterior orientation and sparse point cloud of the central part of the Welfenschloss

4.2 Flight configuration

As said in section 2.1 we did not employ pre-flight planning due to the automatic imaging mode and a small object distance. We performed imaging in three different object distances. In the closest distance we flew in two different heights to capture every part of the building. For the gain of a higher diversity in perspective we took several convergent photos. Figure 3 shows this configuration for the central part of the building. With a maximal drifting speed parallel to the façade of approximately 1m/s and a minimal object distance of about 5m, the picturing rate of two seconds guarantees an overlap of at least 60%.

In a first configuration an exposure time of 1/500 s was chosen, which appeared to be too long because at least 50 % of the imagery was blurred. In a second attempt we halved that time and reached at least 90 % images with a satisfying quality in terms of passable sharpness and visible noise. However, due to the short exposure the gradation of the imagery had to be corrected. Figure 4 illustrates that there is still a higher level of blur compared to the terrestrial image, which we want to reduce in future flights.

Although the UAV records GPS positions during flight, at the moment we did not merge this information with the images to ensure comparability with the terrestrial images.



Fig. 4 Detail of the façade captured terrestrial (a) and from UAV (b) in the same resolution

4.3 Photogrammetric Processing

We end up with nearly 150 images in a resolution of 8MP, which we do not process at once. It shows that dividing the imagery into three different parts (left, central and right wing of the building) leads to a more stable orientation. The algorithm does not get any further information about image order, so orientation is computed within an unordered block with a multi-scale approach, including camera calibration. In the course of this paper we use a radial calibration model with three coefficients for radial distortion and two coefficients for decentring and affine parameters respectively.

The coordinate system of the relatively oriented block is ambiguous, since we do not imply exterior information. Because we do not have any ground control points (GCP), we create our own coordinate system based on image measurements with the Z-axis orthogonal to the main plane of the building. Anyhow, since MicMac is designed for aerial photogrammetry, this step is essential to avoid singularities. Furthermore it helps analysing the results, as it implies a physical meaning to the model and constitutes the orthophoto aligned to the main plane of the façade. For dense matching we mask the interesting parts of the façade and use a correlation window of 11x11 pixels. Smaller windows result in a perforated orthophoto, because the algorithm cannot find reliable matches. For radiometric equalisation of the rectified orthophoto linear polynomials P and R are computed, holding

$$P_i(x, y)O_i(x, y) = P_j(x, y)O_j(x, y) \quad (1)$$

for intersecting individual orthophotos O_i, O_j and

$$P_i(x, y)O_i(x, y)R(x, y) = O_i(x, y) \quad (2)$$

to avoid radiometric drifting in the overall image [9].

4.4 Evaluation of the results

The final point cloud, stitched together from five individual parts, consists of more than 11M points (figure 5). The point density varies between 2000 and more than 5000 points per square meter, for instance in merging regions (detailed view in figure 5).

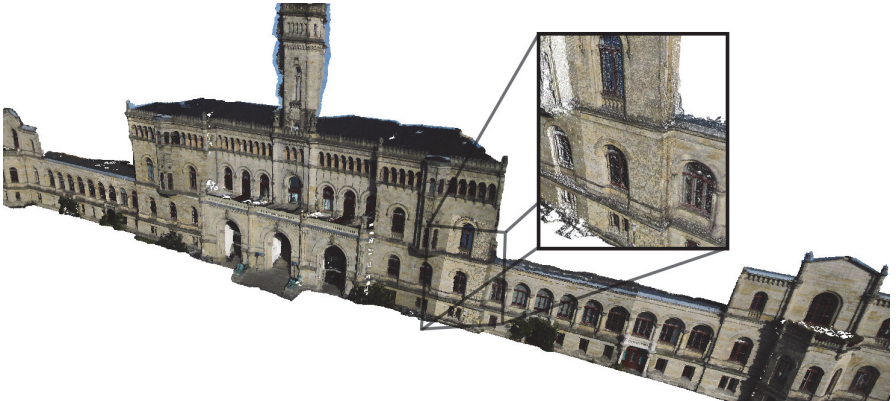


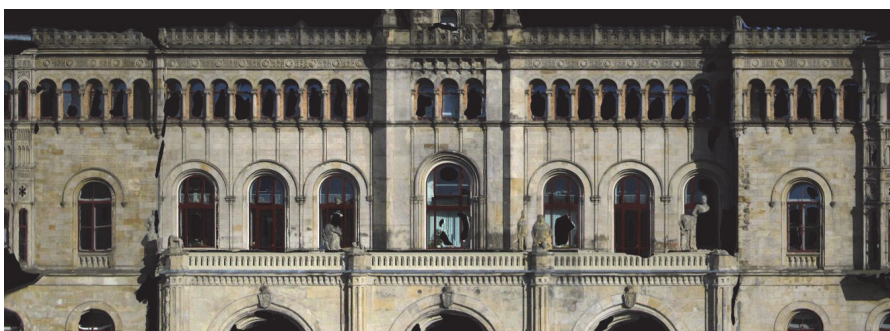
Fig. 5 3d point cloud of the Welfenschloss computed from UAV-imagery

However, in a few parts points are distributed very sparsely, for instance at the arch of the main entrance or other parts orthogonal to the façade plane like the floor of the balcony. This is due to the selected mask for dense matching, which is chosen

for achieving a good orthophoto. The model can be completed with further matching of other parts of the building.

The resulting orthophoto is depicted in figure 6(a). As can be seen it is geometrically correct since there are no distortions due to different depths. Admittedly, at regions of depth jumps like edges at windows or balcony, small parts of missing information depicted as black holes appear. Compared to the orthophoto generated from terrestrial imagery (figure 6(b)) the influence of occlusions due to perspective limitation becomes obvious, which also apply to the computed point cloud. In fact a completion of the model is not possible with terrestrial images only. For a numerical evaluation we compare the photogrammetric result with a reference point cloud from terrestrial laserscanning (achieved by the Institute of Cartography and Geoinformatics²).

In a first example we pick the same part of both point clouds which is supposed to be plain. The two point clouds are scaled with the iterative closest points (ICP) algorithm. We fit a plane through the points of the laserscan via linear regression



(a)



(b)

Fig. 6 Orthophoto of the central part of the Welfenschloss, achieved by UAV- (a) and terrestrial photographs (b)

² <http://www.ikg.uni-hannover.de/>

and get a standard deviation of about 1.5 cm. The mean deviation between this plane and the points of the photogrammetric result is about 2 cm and uniformly distributed, which is the magnitude we expected and which is sufficient for architectural photogrammetry purposes. In a second evaluation we compare a considerable part of the central façade (figure 7). Since the surface of this part of the façade is not plain, estimating a plane does not make any sense. So the point clouds are compared directly, i.e. the distance to the closest reference point is estimated. Due to a different point density, absolute distances are generally higher than in the first example; hence they have to be interpreted just relatively. As can be seen in figure 7 maximal discrepancies between the point clouds are at the windows. This is expectable because firstly, points behind the windows, whether they are triangulated or scanned, are highly dependent on the viewpoint, which are different, considering terrestrial laserscanning and UAV-photographs. Secondly, objects behind the windows, like for instance curtains or plants, may have changed during times of acquisition. Anyhow, there is no systematic error, just random uncertainty.

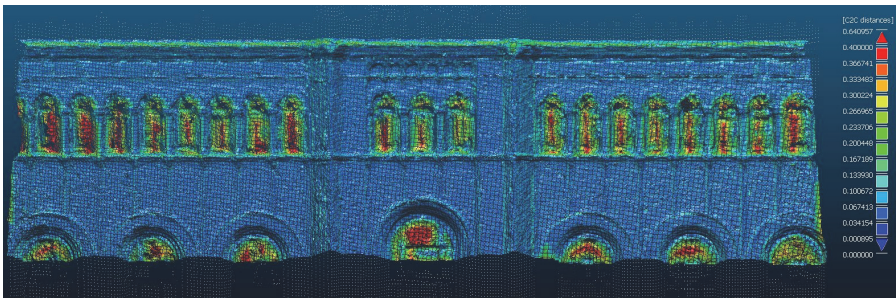


Fig. 7 Differences between laser and photogrammetric point cloud (red: >40cm, blue: <6cm)

5 Conclusion and future prospects

Irrespective of the type of image acquisition, we employ the photogrammetric workflow to both, terrestrial and UAV-based image data. With the results we demonstrate the ability of the UAV to overcome a significant limitation of terrestrial approaches in façade modelling: missing image information due to occlusions. Despite of lower image quality, the gain of the diversity of viewpoints results in a more complete model and orthophoto. An effect of a reduced resolution in terrestrial images for the upper parts of the façade due to a higher object distance cannot be seen in the orthophoto. For higher buildings this effect is supposed to be conceivable. The comparison of the point cloud with data of a terrestrial laserscan shows that in terms of accuracy the model can be used for the purposes of architectural photogrammetry.

The use of a UAV for façade modelling implies the attendance of a laptop during acquisition. Covering a huge object like the Welfenschloss is quite ambitious. To assist the user we think of a near real-time processing of the data achieved by the UAV. In the course of realising this, we implemented a tool that computes a sparse orientation with Bundler once at least two images have been taken. For the

communication between camera and laptop we use a Wi-Fi-SD-card, which directly broadcasts the image to the laptop for processing. The sparse point cloud can be used to get a better overview, which parts of the object need further examination. First tests have shown the applicability for indoor environments. In an outdoor setting the loss of the Wi-Fi-signal regularly constrained the processing. Once we get rid of this problem with the used hardware, we think about using the in-time oriented model to compute a reasonable path for the UAV automatically. Taking this into account a highly efficient exploitation is imaginable.

Literature

1. Agarwal, S., Snavely, N., Simon, I., Seitz, S. M. and Szeliski, R., 2009. Building rome in a day. *ICCV*
2. Furukawa, Y., Ponce, J., 2007, Accurate, dense and robust multiview stereopsis. *CVPR*
3. Haala, N., Cramer, M., Weimer, F., Trittler, M., 2011. Performance test on UAV-based photogrammetric data collection, *International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g), IAPRS Volume XXXVIII-1/C22*.
4. Irschara, A., Kaufmann, V., Klopschitz, M., Bischof, H., Leberl, F., 2010. Towards fully automatic photogrammetric reconstruction using digital images taken from UAVs, *International Society for Photogrammetry and Remote Sensing Symposium, 100 Years ISPRS - Advancing Remote Sensing Science*.
5. Läbe, T. And Förstner, W., 2004. Geometric stability of lowcost digital consumer cameras. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35(5), pp 528-535, Istanbul, Turkey*.
6. Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91-110.
7. Pierrot Deseilligny, M., Clery, I., 2011. APERO, an open source bundle adjustment software for automatic calibration and orientation of set of images. *IAPRS&SIS Volume XXXVIII-5/W16*
8. Pierrot Deseilligny, M., Paparoditis, N., 2006. A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery. *IAPRS&SIS Volume XXXVI-1/W41*
9. Pierrot Deseilligny, M., September 2012. MicMac, Apero, Pastis and Other Beverages in a Nutshell! *Documentation of Apero and MicMac*
10. Snavely, N., Seitz, S. M., Szeliski, R., 2006. Photo Tourism: Exploring photo collections in 3D. *TOG*, 25(3):835-846
11. Zischinsky, T., Dorfner, L., Rottensteiner, F., 2000. Application of a new model helicopter system in architectural photogrammetry, *Amsterdam, IAPRS Volume XXXIII. B5/2*