The Accuracy of Automatic Photogrammetric Techniques on Ultra-Light UAV Imagery

CHRISTOPH STRECHA, Lausanne

ABSTRACT

This paper presents an affordable, fully automated and accurate mapping solutions based on ultra-light UAV imagery. Several datasets are analyzed and their accuracy is estimated.
We show that the accuracy highly depends on the ground resolution (flying height) of the input imagery. When chosen appropriately this mapping solution can compete with traditional mapping solutions that capture fewer high-resolution images from airplanes and that rely on highly accurate orientation and positioning sensors on board. Due to the careful integration with recent computer vision techniques, the result is robust and fully automatic and can deal with inaccurate position and orientation information which are typically problematic with traditional techniques.

1. INTRODUCTION

Fully autonomous, ultra-light Unmanned Aerial Vehicles (UAV) have recently become commercially available at very reasonable cost for civil applications.
The advantages linked to their small mass (typically around 500 grams) are that they do not represent a real threat for third parties in case of malfunctioning.
In addition, they are very easy and quick to deploy and to retrieve. The drawback of these autonomous platforms certainly lies in the relatively low accuracy of their orientation estimates. In this paper, we show however that such ultra-light UAV's can take reasonably good images with large amount of overlap while covering areas in the order of a few square kilometers per flight.

Since their miniature on-board autopilots cannot deliver extremely precise positioning and orientation of the recorded images, post-processing is key in the generation of geo-referenced ortho images and digital elevation models (DEMs). In this paper we evaluate an automatic image processing pipeline with respect to its accuracy on various datasets.
Our study shows that ultra-light UAV imagery provides a convenient and affordable solution for measuring geographic information with a similar accuracy as larger airborne systems equipped with high-end imaging sensors, IMU and differential GPS devices.

In the frame of this paper, we present results from a flight campaign carried out with the swinglet CAM, a 500-gram autonomous flying wing produced by senseFly as well as the X100, a stable UAV with 2 kg produced by Gatewing. The images of both are geotagged after flight and form the input to the automated processing developed at EPFL-CVLab. In this paper, we compare two variants:

- The first one consists of an aerial triangulation algorithm based on binary local keypoints. Its output is a geo-referenced orthomosaic together with a DEM of the surveyed area. In its basic form, no ground control point (GCP) is used and the geo-localization process only depends on the GPS measurements (geotags) provided by the UAV. This is a fully automated, ``one click" solution.

- GCPs can be spotted in the original images and automatically taken into account by the algorithm to improve the geolocalization accuracy. The procedure allows removal of the geo-location bias which is due to the geotag inaccuracy. Except for the GCP measurements on the field and determination on the original images, no other manual intervention is needed to produce the results.

Depending on the application, the burden of measuring GCP can be traded against a lower resulting accuracy. This suites various needs in terms of accuracy, time to result and cost.
Growers or people engaged in field mission planning for instance may be interested obtain a quick survey in the form of a georeferenced orthomosaic produced fully automatically within minutes. We show that the accuracy without GCP lies in the range of 2m for low altitude imagery.
With just a little bit more of human intervention, i.e. the designation of a couple of GCPs in the images, an accuracy of 0.05-0.2m can be achieved. This accuracy largely depends on the ground resolution of the original images as will be shown later on.

To the best of our knowledge, this paper presents the first demonstration that the combination of ultra-light UAV imagery and automated processing is possible and yields accurate results, comparable to the ones obtained with traditional photogrammetric systems mounted on airplanes.
The main issue to achieve this is the imprecise measurements for the location and orientation of the individual images (H. Eisenbeiss and W.Stempfhuber, 2009).

Recent techniques rooted in computer vision, their fast and scalable implementation and the robust integration into photogrammetric techniques are the main key to circumvent the lack of precise sensor information.
The presented approach opens the door to a wide range of new applications and users which can now access geographic information at an affordable cost and without any knowledge in photogrammetry. The temporal (4-dimensional) analysis of local areas, as for instance the monitoring of reconstruction sites, becomes on one hand affordable because of the reduced cost of the hardware. Expensive helicopters or airplanes are replaced by ultra-light UAV's. The automated processing on the other hand reduces the labor cost substantially and makes such projects, which would normally require a lot of manual intervention using traditional photogrammetry techniques, feasible for the first time.

2. AUTOMATED DATA PROCESSING BY PIX4D

The web-based service that can automatically process up to 2000 images, is fully automated and requires no manual interaction. Geo-referenced orthomosaic and DEM can be obtained in principle without the need for ground control points. However, as shown in the various examples provided in this section, more accurate results can be obtained by using GCP.
The software performs the following steps:

- The software searches for matching points by analyzing all uploaded images. Most well known in computer vision is the SIFT (Lowe, 2004) feature matching. Studies on the performance of such feature descriptors are given in (K. Mikolajczyk and C. Schmid, 2002). We use here binary descriptors similar to (C. Strecha and A. Bronstein and M. Bronstein and P. Fua, 2011 accepted), which are very powerful to match keypoints fast and accurate.
- Those matching points as well as approximate values of the image position & orientation provided by the UAV autopilot are used in a bundle block adjustment (B. Triggs and P.

McLauchlan and R. Hartley and A. Fitzgibbon, 2000) to reconstruct the exact position and orientation of the camera for every acquired image.

- Based on this reconstruction the matching points are verified and their 3D coordinates calculated. The geo-reference system is WGS84, based on GPS measurements from the UAV autopilot during the flight.
- Those 3D points are interpolated to form a triangulated irregular network in order to obtain a DEM. At this stage, at dense 3D model (D. Scharstein and R. Szeliski, 2002) (C. Strecha and T. Tuytelaars and L. Van Gool, 2003) (H. Hirschmüller, 2008) (C. Strecha and W. von Hansen and L. Van Gool and P. Fua and U. Thoennessen, 2008) can increase the spatial resolution of the triangle structure.
- This DEM is used to project every image pixel and to calculate the georeferenced orthomosaic (C. Strecha and L. Van Gool and P. Fua, 2008).

In order to assess the quality and accuracy of this automated process, we consider here several projects that differ with respect to the coverage area, ground resolution, overlap between original images and the number of images.

For all datasets we measured GCPs, which we then used to evaluate the precision of the automated reconstruction. Thereby we evaluated two different methods.

One is purely based on the geotags of each original image as provided by the UAV autopilot and one which in addition also takes manually designated GCPs into account.

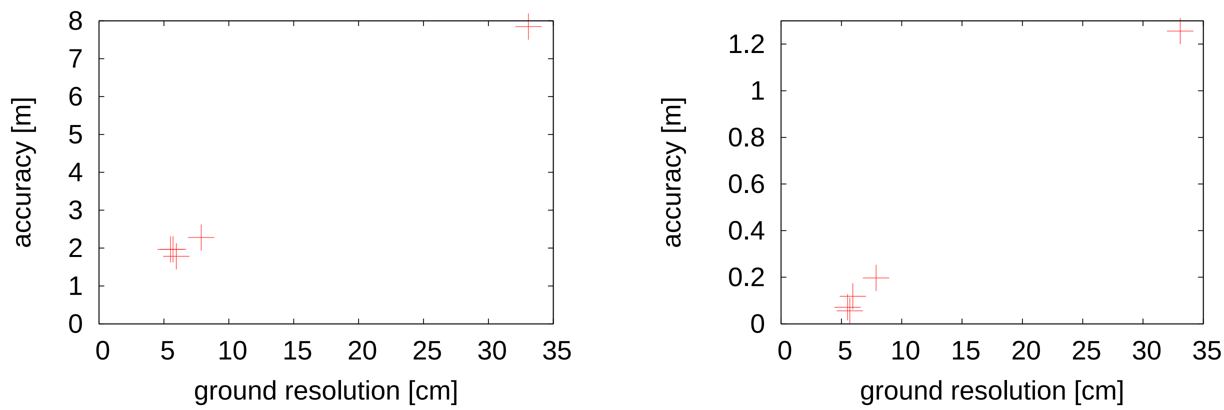
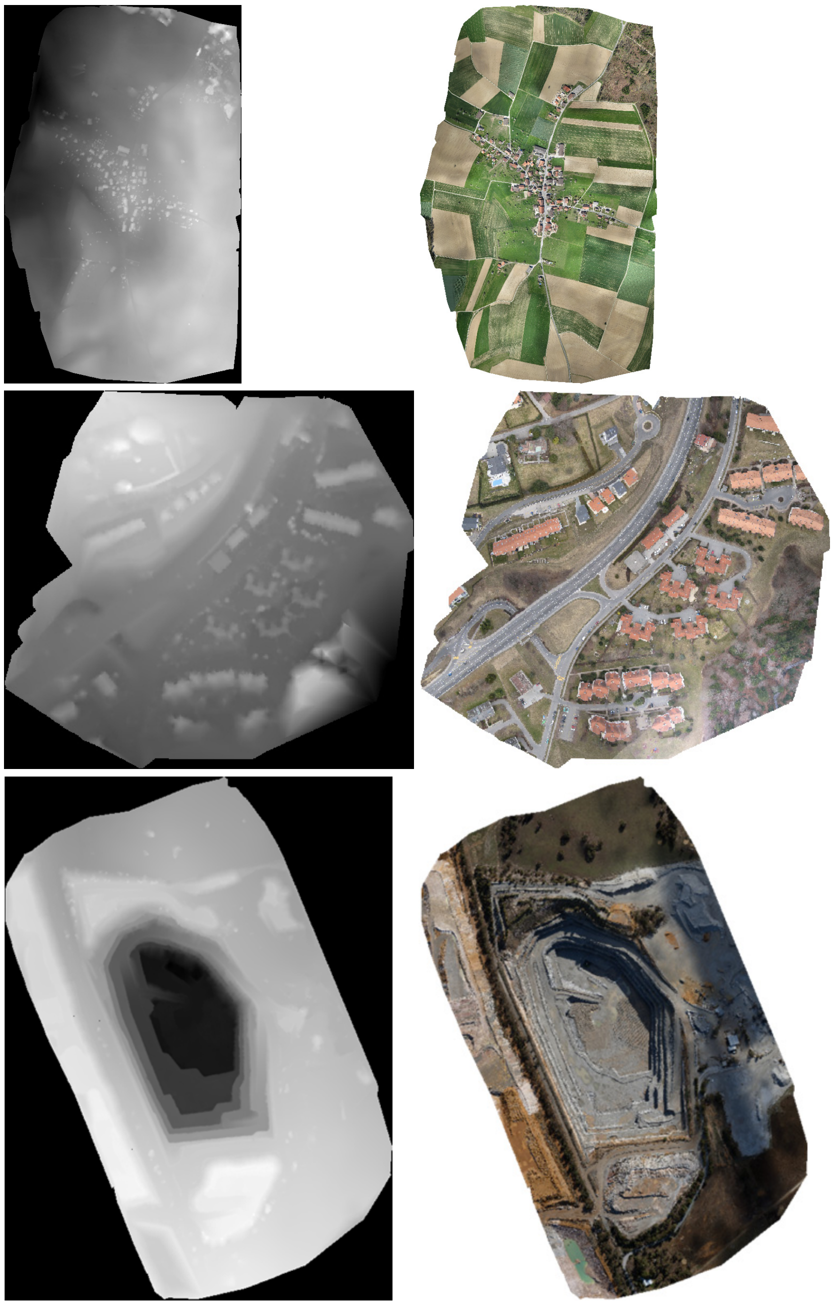


Figure 1: Dependency of the accuracy from the ground resolution (Ground sampling distance) of the original images for various datasets with using GCPs (left) and without using those (right).



	X[m]	Y[m]	Z[m]
RMS	0.11	0.99	0.23
σ	0.064	0.062	0.127

senseFly dataset
523 images
20 GCPs

	X[m]	Y[m]	Z[m]
RMS	0.027	0.029	0.270
σ	0.020	0.020	0.015

senseFly dataset
73 images
12 GCPs

	X[m]	Y[m]	Z[m]
RMS	0.086	0.074	0.053
σ	0.040	0.061	0.053

Gatewing x100 dataset
416 images
11 GCPs

Figure 2: DEM and ortho image for three datasets, one in each row. In the right column one can see the mean errors and their variance in each dimension for all GCPs.

To assess both approaches, we applied our method with and without GCP to three datasets. They are shown in Figures and differ with respect to the ground resolution of the original images, the amount of images and the area they cover. We show in these Figures the resulting orthomosaic and DEM for each of the datasets.

In Figure 1, we plot the accuracy as a function of the image ground resolution for the three datasets above and for two others for which we cannot show detailed results due to space limitation. All experiments confirm the expected dependency of the accuracy on the ground resolution of the original images. We can conclude that the accuracy lies between 0.05-0.2m when including GCPs and 2-8m with the no-manual-intervention variant.

However, this accuracy cannot be achieved for all parts of the orthomosaic. Some areas might not be very well textured or could contain large discontinuities in depth (for instance near building boundaries or thin tree structures). For those areas the accuracy will be slightly worse.
To evaluate this, more experiments with LiDAR as ground truth are necessary (C. Strecha and W. von Hansen and L. Van Gool and P. Fua and U. Thoennessen, 2008).

3. SUMMARY

The experiments on various datasets shown in this paper suggest a linear dependency of the accuracy on the ground resolution of the original images and thus on the flying height of the mini-UAV (see Figure 1). The accuracy, however is bounded by the minimal flying height for which the mini-UAV is able to capture images without motion blur and with sufficient overlap.

We presented a robust and automatic work-flow which can deal with the fact that ultra-light UAVs provide only relatively inaccurate information about the position and orientation of the captured images. This limited accuracy would typically pose a problem to for traditional photogrammetric work-flows which require a lot of manual labor to achieve results.
The presented post processing makes use of recent and robust computer vision techniques to overcome this problem. We believe that this approach will enable a range of decision-makers to create their own maps on the spot and on demand. This can be very useful in many fields such as agriculture, land management, forestry, humanitarian aid, mission planning, mining, architecture, archeology, urban planning, geology, wild life monitoring, forestry and many others.

4. ACKNOWLEDGEMENT

We would like to thank senseFly (www.sensefly.com) and Gatewing (www.gatewing.com) for the collaboration and R-Pod (www.r-pod.ch), Helimap (www.helimap.com) and Position Partners (www.positionpartners.com.au) for the datasets.

REFERENCES

B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon Bundle Adjustment – a Modern Synthesis [Journal] // Vision Algorithms: Theory and Practice. 2000. pp. 298-372.

C. Strecha, A. Bronstein, M. Bronstein, and P. Fua LDAHash: improved matching with smaller [Journal] // TPAMI. 2011 accepted.

C. Strecha, L. Van Gool, and P. Fua A Generative Model for True Orthorectification [Konferenz] // ISPRS. 2008.

C. Strecha, T. Tuytelaars, and L. Van Gool Dense Matching of Multiple Wide-baseline Views [Konferenz] // ICCV. 2003.

C. Strecha, W. von Hansen, L. Van Gool, P. Fua, and U. Thoennessen On Benchmarking Camera Calibration and Multi-View Stereo for High Resolution Imagery [Konferenz] // CVPR. 2008.

D. Scharstein and R. Szeliski A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms [Journal] // IJCV. 2002. p. 2002.

H. Eisenbeiss and W.Stempfhuber Genauigkeitsanalyse der 3D-Trajektorie von Mini-UAVs [Konferenz] // Zukunft mit Tradition: 29. Wissenschaftlich-Technische Jahrestagung der DGPF. 2009. pp. 407-417.

H. Hirschmüller Stereo Processing by Semiglobal Matching and Mutual Information [Journal] // TPAMI. 2008. pp. 328-341.

K. Mikolajczyk and C. Schmid An Affine Invariant Interest Point Detector [Konferenz] // ECCV. 2002. pp. 128-142.

D. G. Lowe Distinctive Image Features from Scale-Invariant Keypoints [Journal] // IJCV. 2004. pp. 91-110.