

3D MODELLING USING AERIAL OBLIQUE IMAGES WITH CLOSE RANGE UAV BASED DATA FOR SINGLE OBJECTS

Tomas Orlik^{a*}, Eatay Ben Shechter^b, Gerhard Kemper^c

^a SkylineGlobe GmbH / Germany - torlik@skylinesoft.com

^b Skyline Software Systems / USA - ebshechter@skylinesoft.com

^c GGS GmbH, Speyer / Germany – info@ggs-speyer.com

Comission III

KEY WORDS: Oblique Aerial Missions, Multi-Sensor System, UAV, Video-processing, Multi-resolution

ABSTRACT

The request for 3D Data for the use of 3D city-models is increasing rapidly. More and more tools are able to deal with data of several sensors, out of video-streams, oblique camera setups with huge overlaps as well as terrestrial data. To achieve high accuracy of the data and a fast processing pipeline, a smart workflow has to be defined and established. However, mixed data sources are still a challenge especially if different sensors with an extremely different GSD are used. This abstracts demonstrates such a workflow, the processing pipeline and the challenges in a mixed data processing. Special calibration and co-calibrating procedures have been applied to get model in model solution managed to solve the dual task of 3D city mapping and cultural heritage conservation. Especially the sensor setup directly influences the geometric accuracy of the product. To do missions for 2-5 cm GSD, metric systems are indispensable while for non-metric applications also simple and cheaper sensors do their job. Besides the different data-sources and sensors, the way of capturing and the related projection is a critical issue. While the classical oblique imaging is a standardized airborne application, captures with UAVs are more like close range photogrammetry on the facades. The combination requests specific pre-processing and definition and transformation steps.

1. INTRODUCTION

The combination of aerial oblique images e.g. with 5 cm GSD is turning up to be a standard. As 10 cm GSD was a typical urban mapping resolution, since 10 years the request for aerial oblique images forced a better representation of facades and newer and faster capture rates on the sensor side brought up 5cm GSD as a standard today. But still some objects especially the area below the roof is not perfectly reconstructed due to the limited visibility out of the aircraft's perspectives. Wide angle oblique cameras can assist for better and more horizontal perspectives, in the same matter the GSD on the facades is reduced in the outer parts of the image. Even the perspective is worth to be considered, the facades are more frequently hidden by other buildings. In addition, due to the limits of data volume, often a compromise of lower GSD is applied which on the other hand does not properly represent more complex objects e.g. historical buildings.

Wissembourg, situated in the north east of France, Region Grand Est, Département Alsace, is a historical significant city close to the border of Germany. Founded in the 7th century, Wissembourg has many important old buildings and is rich of cultural heritage objects. Besides that Wissembourg played a key role after the second world war to proclaim unified Europe as a vision for the future. Wissembourg is a useful test site since it covers urban area, cultural heritage objects, hilly terrain with vineyards and agriculture, mountains of the Biosphere Reservation Alsace / Palatine and rivers/cannels floating around and through the city. The Palais Stanislas, that was captured with a UAV mission, is a building from 1722 where prince Stanislas stayed in exile. It is actually under a reconstruction and one of the major historical monuments in Wissembourg.

The reconstruction of historical buildings, especially if they are listed in the inventory of cultural objects in France, has to be done with respect to the cultural heritage and must be documented after each reconstruction step. The UAV Mission is a new feature for such object documentation and be more and more used also for metric measurements.

Data captured for this Palais by UAV have a resolution better than 1 cm GSD (in average 4 mm) that enables the extraction of details that can be used also for other conservation purposes. UAV data can be captured more frequently than aerial missions and support the documentation of an ongoing reconstruction. Principally also terrestrial taken images can assist to solve this task.

2. SENSOR-SETUP

Different oblique sensors systems are available on the market which are capable for producing data in a quality usable for 3D Data processing.

Metric and semi-metric sensors in different price ranges deliver output in different resolution and accuracy. Low priced systems that make use of consumer SLR cameras (e.g. Canon, Nikon) are defined as semi metric. (Kemper, 2016).The focal plane shutter causes distortions during exposure and this effect becomes bigger with higher resolution and flight speed. They are typically used for lower accurate projects e.g. 10 cm GSD and /or for resilience non metric tasks.

Metric cameras use leaf shutters and are usually calibrated for photogrammetric use. Fast capture rates and the reduction of mechanical parts support a stabile calibration and make them a real photogrammetric sensor. (Tölg, Kemper, Kalinski, 2016) Such cameras can be applied to

* Corresponding author

very high resolution projects with an accuracy < 5 cm. Oblique Sensor that uses GNSS-INS and a gyro-stabilized mounts improve the processing speed significantly and supports a fast and accurate workflow with a low number of ground control points.

The project area of Wissembourg was flown with 5 cm GSD, 80% overlap in flight direction and 60% sidelap between the flight-lines.

The focal length, camera orientation and tilt angles have been applied to the mission parameters in order to achieve the optimal result related to the building structure and the terrain. Besides that, the sensor setup can assist for specific flight conditions e.g. regulations of minimum AGL surveys.

The relation between the nadir viewing camera and the oblique's should result in a similar GSD. That way the focal length of the oblique cameras are typically by factor 1.1-1.3 longer than the nadir ones.

For the mission over Wissembourg we used the Oblique Imaging System OIS-L manufactured by GGS. This sensor is equipped with one PhaseOne iXM-RS150 camera with 150 MPix and 70 mm focal length for nadir view, mounted in landscape format, and four iXM-100 100 MPix cameras for oblique views each equipped with a 80 mm lens.



Figure 1: The OIS-L with 550 MPix

The sideward looking cameras were mounted in portrait orientation with an tilt angle of 41 degree while the forward and backward cameras were mounted as landscape with 45 degree tilt angle. This setup results in an oblique GSD of 4.7-7.8 cm. The footprint that is used also for the mission planning is shown in Figure 2.

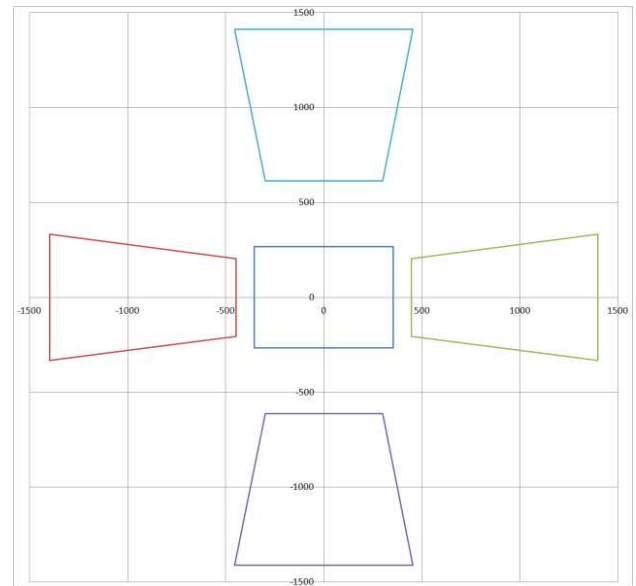


Figure 2: The calculated footprint of the OIS-L

The mission was controlled by the flight management software AeroTopoL. All images are geo-tagged with x, y, z, roll, pitch and heading, and additionally stabilized with the gyro stabilized mount AeroStab XL. Besides that the precise external orientation was achieved with a L1/L2 GNSS and a fiber-optical IMU which results in a few cm accurate position and attitude values at about 10 mgon.

The UAV Mission was flown with an AeroSpector-S quadcopter and a 17 MPix semi metric camera. In this case and due to the slow movement, distortion effects of the rolling shutter and distortions by forward motion was below the final resolution of 0.3 cm.

The widest angle of the camera was used to prevent refocusing during the mission and to capture a big area especially when needed to fly close. Zooming was blocked to the widest opening angle because changes cause big problems in the on-the-fly camera calibration during AT. Big changes in the refocusing also cause small changes in the focal length calibration, that is especially in small opening angles the case.

The images captured are all geo-tagged and have directly been used in the first alignment process. This geo-tags are not as accurate as from the aerial mission but fine to get the images relatively sorted and adjusted already close to the real position.

The use of Video HD streams would have been an alternative, the number of images would increase but the missing geo-tagging of the used sensor would cause other problems in the alignment. On one facade we captured both but used only the geo-tagged images.

The availability of a LiDAR sensor, which we had very much liked to fly in a combination with a camera, was not possible also due to the regulations inside the city. The quadcopter AeroSpector, with a payload of 6.5 kg, is capable to carry both in one mission, but the legalization procedure especially in the pandemic period is a long lasting process. With a takeoff weight of 20 kg there are many restrictions especially in dense and very populated urban areas.



Figure 3: AeroSpector with Camera and Lidar combination

3. AERIAL MISSION

Capture images with 5 or 6 cameras with big overlap and high resolution requires a fast and robust data storage. Losing any single image creates much work in sorting the dataset. A proper and redundant data-handling is necessary. Writing GPS Tags into the image information is a must. Mission planning with an aircraft based Oblique Imaging System (OIS) is a very specific task since flight lines must be planned with an offset to the area of interest to capture the required scene in all four oblique views. Especially for small areas the expenditure increases dramatically. Besides that, the proper calculation of the floating GSD and the estimation of the overlap of the trapezoid footprints causes some mathematical challenges. The situation becomes even more complex if a terrain model has to be applied for mountainous areas. Calculating proper overlaps, taking care of the GSD and to keeping the flight costs in mind require Oblique Imaging Systems operators to find a good compromise.

The area was defined as a polygon and flown by taking the major terrain structure into account.

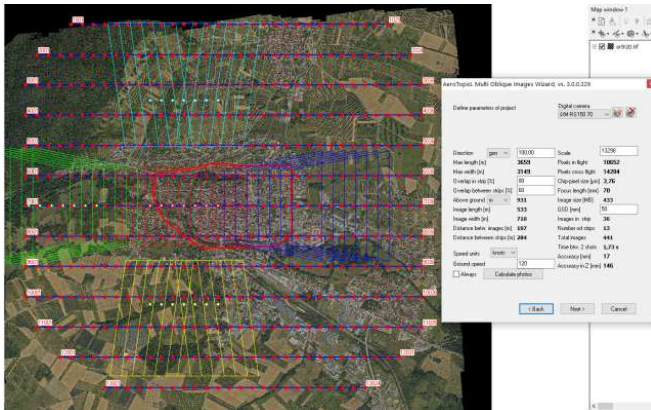


Figure 4: Mission plan with AeroTopoL

The Camera was installed in a Cessna 206 designed for aerial survey with a big hole. The entire flight time over the area was about 50', mainly because of the turns into the new mission line. In average each 1.7 sec a image-set was captured with 550 MPix. In total $441 * 550 \text{ MPix} = 242 \text{ GPPix}$ have been used to model the inner city. A controller with 4 Tbyte SSD Hard drives stored the images fast and save in the raw format.

4. PROCESSING AND CHALLENGES

For the aerial mission data two processing steps were needed before entering the processing pipeline. The conversion from raw to tiff was performed with Capture One software while the raw GNSS-INS Data had been post processed with Inertial Explorer. Ideally all data is stored with proper GeoTags on SSD drives and can be directly loaded to the post processing pipeline. If there has to be a GNSS-INS post-processing done, the image-name and the post-processed EO has to be synchronized. For some camera setups that is a simple process, for others additional steps have to be used to get the data ready for the image processing. Here the bore side calibration and the relative orientation of the sensor-head has been applied, that way we received for each camera its own EO already taken the calibrated tilt angles into account. The processing of the AT of the aerial images with precise EO is a quick and very straight forward step. In some cases the orientation of the camera and the rotation definition must be considered. Rotations around 180 degree may have to be applied either during the processing setup of the AT or already on the post-processed GNSS-INS data together with the bore side values before exporting.

We used PhotoMesh™ from Skyline Software Systems to process the data. Using a first iteration of tie point matching highlights if there are issues of the pre-processed EO. Typically, it is just a kind of definition in the rotation matrix resulted out of a specific camera orientation. Performing this, a first camera calibration is done on the flight during the AT process. To refine this, the result can be exported to Bingo to have a very accurate camera calibration that also can be applied for other projects as well. (Kemper, Melykuty, Yu, 2016). Thus done, a fully automated processing of point clouds, mesh and texturing can be preformed, in most cases just a matter of CPU Power.

While the processing of aerial surveyed data is a standards and fast workflow, UAV data need a special handling. Using UAV images with Geotags but without IMU data, the better way is to process this data-source separately and then join the data with the aerial surveyed result. The best approach to UAV Images with only GPS Coordinates (XYZ) and no orientation (Yaw/Pitch/Roll or Omega/Phi/Kappa) is a 3 step Aerial Triangulation process:

1. Perform Aerial Triangulation with full automated Tie-Point Extraction and Bundle Block Adjustments of only the UAV Images
2. Utilizing the input GPS as a factored observation allows the Bundle Block Adjustment to calculate all of the orientations and camera model calibration (focal length, principal point, radial and tangential distortion)
3. The calculated AT results function as an input to the full Aerial Triangulation with the Aerial Images
After that a combined Aero Triangulation process using the calculated AT input of the UAV Images and Aerial Images can be applied.



Figure 5: Results of the image orientation process on the UAV data, performed on the Palais Stanislas in Wissembourg. The camera positions are displayed.

In the first step the automated Tie-Point Extraction between the UAV images and Aerial Images is done. In order to ensure high level co-registration between the UAV Images and Aerial Images, supplemental control points from the resulting 3D Mesh from the Aerial data are used.



Figure 6: Selecting control points in the aerial 3D model for fine adjustment of the UAV data

This process ensures that the resulting 3D Textured Mesh from the combination of the UAV images and Aerial Images will be both aligned well to one another, and to be able to supplement the Aerial Images in the edges of the data in order to ensure smooth transition between the different meshes in terms of matching color scheme and geometry.

Is in case of a high difference between both GSD values, a interim dataset for AT is recommended, at least a core UAV level in order to support the tie point matching on the facades.

The maximal GSD difference between datasets which can be utilized together in a single 3D output is X4 difference. Therefore, if the aerial images are at 5cm/pixel GSD, a medium altitude data collection from the UAV at 1.25-2cm/pixel range will connect well with the aerial images, and then a low altitude UAV flight can be added as well in the sub-centimeter GSD range (0.3-0.5cm/pixel range) which will be able to be seamlessly integrated with the complete dataset.

One aspect to be taken into account is the temporal latency or time difference between data collections. Mapping is a discipline which attempts to document geographic features in the 3D dimension (X, Y, and Z), though given the dynamic nature of earth and all of its developed surfaces, the temporal dimension is a critical to ensure best co-registration.

With the fourth dimension (time) having limited latency between the different resolution data collections can ensure that objects such as vegetation (seasonal canopy), street furniture (benches, poles, monuments), and temporary traffic conditions (parked cars, road construction) have limited variation between the datasets.

Other dynamic objects are almost impossible to avoid, though due to the fact that they are in continuous motion, they usually do not affect the 3D reconstruction process. These include moving vehicles, bicycles, people, active construction equipment, and other moving objects.

The main challenges are overcoming different scenes in the images due to different times, and the variance in resolution. In cases where the GSD is closer (under X4) and the temporal latency is minimal, it is possible for all images to be solved together without Control Points to co-register the 2 datasets.

In cases of temporal latency is larger, and the resolution difference is difficult to reconcile, it is always possible to extract control points from the Aerial 3D Textured Mesh and measure them in the UAV Imagery.



Figure 7: Example of a produced Orthophoto captured by UAV of the "Maison du Sel" before the reconstruction of the roof with a resolution of a 5 mm GSD.

5. FINAL PROCESSING – OUTCOME

PhotoMesh selects the entire area in tiles and processed them with a certain overlap in parallel - that increases the processing speed of the hardware.

The first step after the fine adjustment of calibration and aero triangulation is the generation of the dense point cloud which is related to the image resolution, means for the UA data in a much denser way than for the aerial surveyed images.

Generating a mesh and fitting the texture is done the next steps. For a good integration of the dens model, it is defined by seam lines for a smooth integration without showing artifacts. The control points help for a perfect correspondence of the 3D models

The resulting high dense model of the UAV mission were cut and inserted into the core model as shown in figure 8 and 9:



Figure 8 and 9: The aerial model above and the UAV model below of the southern facade of the Palais Stanislas

Figure 10 and 11 show the difference in the GSD very detailed. The roof with this old tiles (Biberschwanz-Ziegel) have a very variable structure and the point cloud and model generated out of this perfect target is close to be optimal. The quality of the reconstruction on the facades is worse due to missing structure.

The combination of the models enables also the analytics of the cultural object in its environment. This can assist in further planning, analytics in the visibility and for preparing additional city planning tasks. Also to generate a virtual touristic guide, these highlighted cultural objects are of importance. The office for cultural preservation in France is dealing with so called preservation zones, their appearance and definition nicely can be analyzed and modified using this type of combined model.



Figure 10 and 11: Details of the roof on the aerial model above at 5 cm GSD and the UAV model below at 0.5 cm GSD.

6. DISCUSSIONS

Giving the input GPS accuracy, PhotoMesh is capable of utilizing the input GPS observation with a high confidence factor and translate the accuracy of the Aerial GPS to the accuracy of the final 3D output. If the Airborne GPS accuracy is better than 1 pixel of GSD (e.g. GPS circular accuracy of 3cm in XYZ and image GSD of 5cm/pixel), then the general accuracy of the final product will be up to X2 the GSD, in this case 10cm accuracy. In case where the GPS accuracy is lower than 1 pixel of GSD (E.G. GPS circular accuracy of 50cm in XYZ and image GSD of 5cm/pixel), then the general accuracy of the final product will be on average X2 the GPS accuracy, in this case 1m accuracy. This relates to the absolute accuracy not taking GCPs into account. The relative accuracy typically is below 1 pixel. Often it is defined by 1/2 pixels that is for the airborne data around 2 cm and for the UAV data in average about 0.3 cm. Using GPPs with sub pixel accuracy improves the absolute accuracy of the model significantly while the relative accuracy and resolutions does not improve in the same way.

What would be the benefit of LiDAR data which obviously are less dense than the point cloud but with lower noise more reliable?

The most significant effect LiDAR has on the 3D Reconstruction process is the ability to provide multiple returns and penetrate tree canopy. This ensure that areas that have tree canopy coverage can still be modeled underneath the vegetation. An additional benefit of LiDAR is that it is more precise in areas with limited texture. While the image-based 3D correlation relies on the recognition of common features to triangulate a 3D point, the LiDAR as a light ranger, has consistent density in areas which are rich in features as well as areas which are relatively homogenous. This is most apparent in roads which are mostly uniform in color, and buildings which are homogeneously colored (all white, or all brown). The LiDAR will supplement the areas in which the image correlation did not extract high quality and dense points.

