Low cost real time UAV stereo photogrammetry modelling technique – accuracy considerations

Pawel Burdziakowski [*]

Faculty of Civil and Environmental Engineering, Gdansk University of Technology, Narutowicza 11/12 , 80-233 Gdansk, Poland

Abstract. The paper presents accuracy considerations regarding three 3D modelling techniques. The tested new consumer type stereo camera (ZED 3D Stereolabs) has been implemented info an aerial mapping system, on board micro air vehicle MAV) and tested object has been mapped using a real-time photogrammetry with original real-time software application. The evaluated results has been compared with model gained with a state of art unmanned aerial vehicles (UAV) photogrammetry process using commercial UAV and commercial software, and the terrestrial photogrammetry modelling process with other commercial software. Papers concludes the tested real-time technology accuracy with compare to the traditional technique, and discusses real-time photogrammetry modelling in terms of engineering applications.

1 Introduction

A low cost consumer grade range 3D cameras market grew significantly last years. The evolution was a response on a recent demand from the gaming, entertainment and robot industry. The industry has boosted the development of low-cost consumer-grade cameras. The interest in these sensors initially remained at relatively low level, up to the moment when Microsoft company introduced Kinect sensor to their video-game console. The Kinect sensor is used as a motion input device. Over 10 million of Kinect units, flowed video game market, and has had a significant impact on the photogrammetric community. The adorability and versatility of the Microsoft's sensor, and similar devices, led to it's wide range of sensing applications such a 3D scene reconstruction, engineering tasks, robotic navigation, simultaneous localization and mapping techniques, and many similar applications, where time of data processing and acquisition is the main priority. Kinect sensor type (or other named device using the same PrimeSense chipset) have a relatively short (1-6 m) effective maximum working range, with limited accuracy (5-20 mm) that's significantly limiting a engineering measurements requirements. That kind of consumer grade cameras can be divided into two categories: range cameras based on triangulation and range cameras based on time-of-flight (TOF) principle. The Kinect technology can be described as structured light using a fixed pattern projection. The dot pattern is generated by a diffractive optical element and a near-infrared laser diode. TOF flight cameras use a single LED to emit sinusoidal modulated infrared light. A synchronized imagine chips takes four samples on one period of a sinus signal, what allows to compute the phase shift and consequently depth for each pixel [1]. The cameras described above are transmitting any kind of near-infrared pattern or signal into the object of measurement. The technology proved to be very prone to the strong daylight and not as effective in daylight conditions as indoor [2].

An indoor mapping and navigation become a challenge and highly desired technology for smart sites and robot applications [3]. The visual light sensors (RGB-D cameras) are playing a significant role within robotic community, closing any robots (humanoid robots, micro air vehicle (MAV), unmanned aerial vehicle) to humane like environment sensing [4].

Outdoor mapping, as mentioned above, is very challenging for a low cost RGB-D Kinect type camera. In the case of outdoor mapping and navigation a passive 3D cameras can be considered as a desired solution. For the presented research the ZED 3D Stereolabs camera has been chosen. The ZED 3D camera is based on passive stereovision, and consists of two RGB cameras with fixed base distance to 120 mm, which allows to generate depth image up to 20 meters (40 m is the maximum distance on the new updated firmware). The ZED 3D camera is optimized for real-time calculation using Nvidia CUDA technology. The ZED camera (among variety of options) can deliver already rectified stereo pairs in highest resolution (4416x1242) at 15 frames per second (fps). Lower resolution HD720 supports higher frame rate (60 fps) and data stream, what is recommended for visual navigation applications on fast moving robots (UAV, MAV) or mapping from the ground or surface moving platform. The ZED camera provides data via high speed USB 3.0 interface and is

[*] Corresponding author: pawel.burdziakowski@pg.edu.pl

able to GPU's (graphical processing unit) accelerated real-time calculations. The camera outputs a real-time images in the form RGB rectified video (left and right), raw (unrectified) RGB video (left and right), left and right rectified side-by-side video (SBS), 32-bits depth map image, registered point cloud, confidence image and disparity image. The calculation is already done on the hardware level on the camera chip itself. The data can be recorded also on custom video format .svo and can be processed on the accompanying software.

The paper considers the ZED 3D camera accuracy mapping results, using Stereolabs application programming interface (API) and its standard camera calculations algorithms in terms of engineering 3D measurements and modelling, compared to the typical UAV mapping technique and terrestrial photogrammetry with digital single-lens reflex camera (DSLR). All three techniques have been used discreetly for the modelling of the same 3D object in outdoor environment in daylight conditions.

Fig. 1 ZED Stereolabs 3D camera.

2 Data acquisition and processing

The object has been placed on the outdoor environment, marked with 12 bits coded targets as a control points for scale recovery and dimensions control. The first modelling technique utilizes MAV (DJI Mavic Pro) with Bentley Context Capture Software for data processing (georeferenced UAV mapping). The MAV (Fig. 1a). At the moment, this product is the smallest flying and stabilized camera on the commercial market (weight only 734 g). The platform utilizes for navigation one GPS and GLONASS module, two Inertial Measurement Unit (IMU) modules, Forward and Downward Vision System to automatically stabilize itself and navigate between obstacles and track moving objects. The presented MAV is using non-metric cameras with the sensor size 1/2.3`` - 6.17x4.55 mm with the pixel size 1.55 μm. The camera of this MAV tags (into the EXIF metadata) the images witch geolocation data using MAV's GPS (direct image georeferencing). The modelling software, Context Capture, is the reality modelling category application for processing images of the physical environment into 3D representations to provide current context within geospatial modelling environments.

a)   *b)*   *c)*

Fig. 2. a) DJI Mavic Pro, b) Canon EOS 500D, c) AI-GEO MAV with ZED Camera.

The second modelling technique utilises DSLR Canon EOS 500 D camera with standard kit lens (EF-S 18-55mm f/3.5-5.6 IS) (Fig. 2b) (terrestrial photogrammetry). Cannon camera is equipped with CMOS sensor with effective 15.1 million pixels. For this task, the Agisoft PhotoScan Pro has been used, as a data processing software. Agisoft is a stand-alone software product that performs photogrammetric processing of digital images and generates 3D spatial data.

The tested configuration, ZED Stereolabs camera, was implemented on board original developer platform named AI-GEO MAV [5] (Fig. 2c). The platform is designed for real-time photogrammetry applications and employs NVidia TX2 embedded commuter for real-time photogrammetry calculations. The operating system Linux Ubuntu running the original software based on Stereolab's application programming interface (API). In this case API is a set of subroutine definitions, communication protocols and tools for building original software, or generally it is a set of clearly defined methods of communication among various components within the system. API's algorithm have not been modified for this research and used as delivered by the Stereolab's within own original mapping commuter application. It has to be mentioned, that the model has been calculated based on the Stereolab's delivered API without any API's code modification. The code is dedicated to the real-time data processing, so it can be expected that accuracy of the model accuracy part can be lower. The presented approach allows evaluate the quality of the overall real-time mapping process. Oppositely to the research [6] where authors evaluated the same camera but using the photogrammetry software, the quality of the models was comparable to the typical UAV photogrammetry technology. Here, in this research, the camera and its API are evaluated in real-time UAV mapping application.

2.1. Georeferenced UAV mapping

The first model calculated using low altitude photogrammetry methods. The images were taken around the object (Fig. 3a). The MAV size allows to carried out the flight around the object in the presented localization.
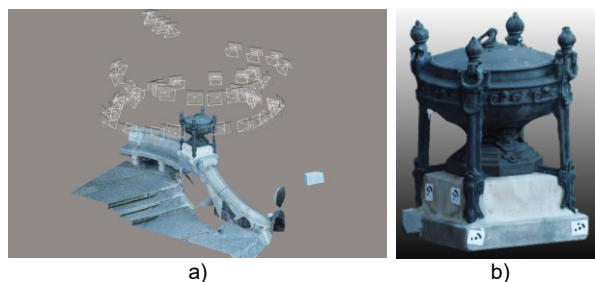

a)   b)

Fig. 3. UAV photogrametry, a) flight patch and extrincis image parameters, b) point cloud model.

The object was modelled (Fig. 3b) form 51 georeferenced images (4000x3000) using Bentley Context Capture. The software reported reprojection error (RMS) is 0.71 pixels, with a minimum resolution

of 0.034 mm/pixel and a maximum of 1.564 meters/pixel. The median resolution equals 0.191 meters/pixel. The model has no shortages, all surfaces has its representation in point cloud. Due to above reason the model form UAV photogrammetry, for this research, has been considered as the referenced model.
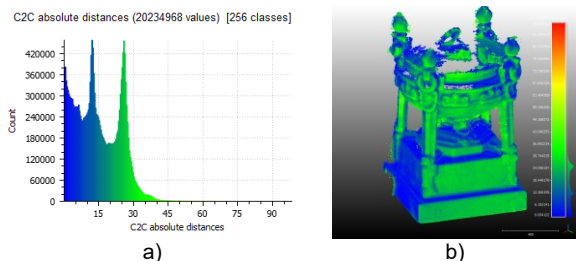
2.2 Terrestrial photogrammetry

The object was situated in the place where the terrestrial photogrammetry technique was inconvenient to use. The outer part of the object was unable to photograph without man lift, nevertheless the images offered the highest resolution, in the presented study, the problems with a camera position caused acquisition problems and some part of the object has not been photographed. The 12 bit circular coded control point has been established and measured. The scale of the object was recovered using marked tie points measurements. The same points was used to the accuracy control in the UAV photogrammetry case.
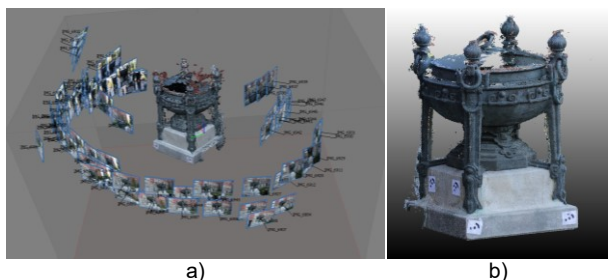
Fig. 4. Terestrial photogrametry, a) camera position and images orientation, b) point cloud model.

The object modelled using terrestrial photogrammetry (Fig. 4b) form 74 images (5184 x 3456) using Agisoft Photo Scan has some shortages, and surfaces has no representation in point cloud. The top of the modelled object has not covered by the source images, due to the object location. All recovered surfaces has a quality and accuracy adequate to the referenced model. Software reported reprojection error (RMS) is 1,99 pixels.


Fig. 5. Cloud to cloud distances comparison (DJI to Canon) a) histogram, b) coloured model and scaled distances in mm.

The comparison of the referenced cloud form DJI Mavic Pro (UAV photogrammetry) to the cloud form Canon camera (terrestrial photogrammetry) (Fig. 5), revealed that the max distances not exceed approx. 35 mm. It confirms the initial assumptions, that the UAV model can be considered as a reference model.

2.3 ZED 3D modelling

The real-time UAV mapping technology has been applied on the original developer UAV platform [5]. The platform allows to deliver the mapping instrument (ZED 3D Camera) to the desired area or object with simultaneous data acquisition and processing on board the platforms commuter. The operator receives processed data in point cloud or meshes (textured, only surface or triangle mesh). The Fig. 6 presents the main interface of the real-time mapping software. The visual light camera image is presented to operator's overall orientation in the vicinity of the object of interest. The recovered depth image in the 12 bit colour scale allows to monitor the quality of the stereo matching process and allows to adjust the camera parameters (the mapping distance). The live meshing process shows to operator already mapped environment and acquired data. In case the environment has not been mapped, there is a possibility to repeat the flight in the not mapped areas.
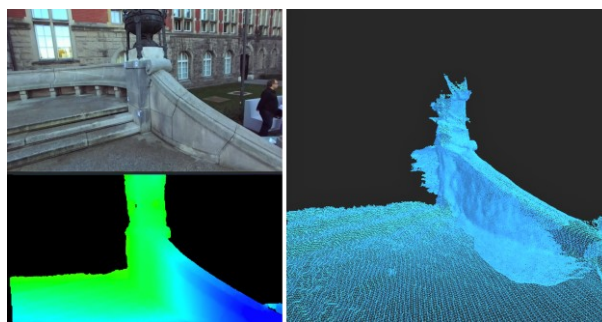
Fig. 6. The real-time mapping software live view (left top: camera view, left bottom: disparity map (depth image), right: life meshing process).
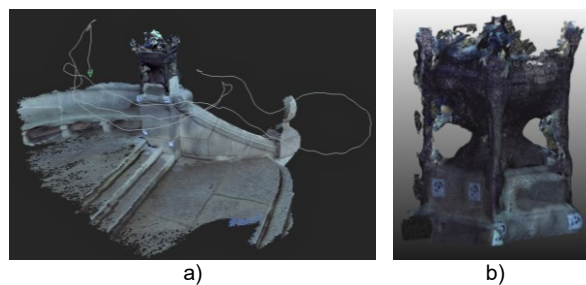

Fig. 7. The platform visual odometry flight patch (a) and textured model of the object and form ZED Stereolabs (b).

The object modelled form the ZED camera using MAV presented on the Fig. 7b. The data has been acquired as a side by side video with resolution HD720 and 60 frames per second. The left and right video frames are synchronized and streamed as a single uncompressed video frame in side-by-side format (resolution for SBS video 2560x720) to the on-board processing commuter.

The texturing process time not exceed 2 minutes, for the mapping area, and starts after the data gathering is finished and the operator confirms the area has been mapped. The point cloud and mesh is received in the real time, means that are created during the flight and results are displayed on the operators screen, but the texturing process stars always after the complete mesh is

recovered. A texture map is generated for the whole area.

The models point to point comparison clarifies the main modelling results. The tested equipment result has been compared to the reference model (Fig. 8).
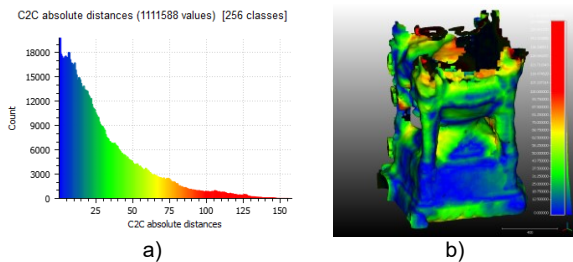
Fig. 8. ZED model vs referenced model (DJI) a) histogram, b)coloured model and scaled distances.

The comparison of the referenced cloud to the cloud form real-time modelling process using ZED Camera (Fig. 8), revealed that the max distances not exceed approx. 125 mm. The biggest deviation from the referenced model can be observed in the areas where the camera did not acquire a sufficient data (above 100 mm), covered by other object elements (approx. 60-75 mm), very detailed object part (approx. 50 mm). Deviations below 25 mm can be observed on the flat and simple shape object parts. It can be noticed, that the object overall shape and middle size details are modelled. The smaller details, are simplified on the mesh and clearly not modelled (50 mm deviation). The mesh generation process simplifies the details, and fast data processing are at the expense of details quality.

4 Conclusions

A real-time photogrammetry applications and technique are still under the development. A fast computation techniques allows to run calculation process, even that complicated, in the real-time on board micro aerial vehicles. The paper considered three different modelling techniques, form high resolution terrestrial photogrammetry, thorough offline commercial UAV photogrammetry, and finally to the real-time stereo photogrammetry form the unmanned aerial vehicles. It can be stated that the presented solution can deliver the spatial data in the real time, means during the flight. The data can be analysed and monitored online, via wireless connection. The spatial data are delivered already in the popular interchangeable file format (*.obj, *.las), and can be used in third party software on in own applications. Form the photogrammetric point of view the quality of calculated the model is affected by the fast data calculation priority. Consequently, this research opens the discussion, if the fast 3D camera volume measurements are meet the geodetic standards. As it can be showed here, the very detailed geometric modelling is rather impossible using the API delivered by Sterolabs at the geodetic grade level. Despite the fact that presented API simplifies the object, and detailed are lost, it would be possible to use either a different software for real-time spatial data calculation for a very detailed object, or

use this process for the larger object i.e. for geodetic volume calculation.

The visual textured model quality, for the user, presents a very attractive form (Fig. 7a). The quality of this process mitigates the geometry weakness. Texturization exposes all small details as a picture on the object, and consequently the user can have an immersion that the details have been modelled and have their geometrical representation in point cloud or mesh.

The presented mapping results have a quality and accuracy on the navigation level, means that the model of the environment can be safely used for the autonomous UAV navigation applications for example on board unmanned platform [7], and for initial and rough geodetic mapping (using standard camera API). While the photogrammetry offline software is engaged, the quality of the results would be on the typical photogrammetric accuracy level, however calculation time will be affected. The good solution would be also to use online photogrammetry technique with live data calculation on remote server and stereo image transmission via internet connection [8] and methods like [9].

The significant problems has been observed while the UAV moved to the shadowed area and the cameras parameters has not been automatically (or wrongly) adjusted. In that case the radiometric correction [10] should be implemented in the real-time data acquisition process. The moving objects with the sensing equipment (cameras) are not always faced the same light condition on the whole object, as well as on the background. That was observed in this case. The online, real-time radiometric correction should be implemented.

Some not modelled parts of the object lead to the conclusion, that the construction of the AI-GEO MAV [5] and 3D camera mount should be modified. The tilting mechanism should be added for mapping an object underneath the MAV while the platform is loitering in the one position or circles around.

References

1. T. Luhmann, S. Robson, S. Kyle, and J. Boehm, *Close-range photogrammetry and 3D imaging*. Walter de Gruyter (2014)

2. F. Endres, J. Hess, J. Sturm, D. Cremers, and W. Burgard, *3-D Mapping With an RGB-D Camera*, IEEE Trans. Robot., **30**, no. 1, pp. 177–187 (2014)

3. P. Burdziakowski, A. Janowski, M. Przyborski, and J. Szulwic, *A modern approach to an unmanned vehicle navigation*, 16th International Multidisciplinary Scientific GeoConference SGEM 2016, Book 2, **2**, pp. 747-758 (2016) ISBN 978-619-7105-59-9 / ISSN 1314-2704

4. T. Gupta and H. Li, *Indoor mapping for smart cities - An affordable approach: Using Kinect Sensor and ZED stereo camera*, 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8 (2017)

5. P. Burdziakowski, UAV design and construction for real time photogrammetry and visual navigation, 2018 Baltic Geodetic Congress (BGC Geomatics) (2018)

6. A. Deris, I. Trigonis, A. Aravanis, and E. K. Stathopoulou, *Depth cameras on UAVs: A first approach*, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, **42**, no. 2W3, pp. 231–236 (2017)

7. A. Stateczny, D. Gronska, and W. Motyl, *Hydrodron- New Step for Professional Hydrography for Restricted Waters*, 2018 Baltic Geodetic Congress (BGC Geomatics), pp. 226–230 (2018)

8. Z. Paszotta, J. Szulwic, and M. Szumilo, *Internet Photogrammetry for Inspection of Seaports*, Polish Marit. Res., **24**, no. s1(93), pp. 174–181 (2017)

9. M. Kedzierski and P. Delis, *Fast orientation of video images of buildings acquired from a UAV without stabilization*, Sensors (Switzerland), **16**, no. 7 (2016)

10. M. Kedzierski and D. Wierzbicki, *Methodology of improvement of radiometric quality of images acquired from low altitudes*, Meas. J. Int. Meas. Confed., **92**, pp. 70–78 (2016)