QUICK 3D WITH UAV AND TOF CAMERA FOR GEOMORPHOMETRIC ASSESSMENT

A. Masiero[1,2,] [*], G. Sofia[3], P. Tarolli[4]

[1] Dept. of Civil and Environmental Engineering, University of Florence,
via di Santa Marta 3, Florence 50139, Italy - andrea.masiero@unifi.it
[2] Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova,
Viale dell'Università 16, Legnaro (PD) 35020, Italy - masiero@dei.unipd.it
[3] Department of Civil and Environmental Engineering, University of Connecticut, Storrs,
CT 06269, USA - giulia.sofia@uconn.edu
[4] Department of Land, Environment, Agriculture and Forestry, University of Padova,
Viale dell'Università 16, Legnaro (PD) 35020, Italy - paolo.tarolli@unipd.it

Commission I, WG I/7

KEY WORDS: Time-of-Flight camera, UAV, Mobile Mapping, Geomorphometry

ABSTRACT:

Most of the high resolution topographic models are currently obtained either by means of Light Detection and Ranging (LiDAR) or photogrammetry: the former is usually preferred for producing very accurate models, whereas the latter is much more frequently used in low cost applications. In particular, the availability of more affordable Unmanned Aerial Vehicles (UAVs) equipped with high resolution cameras led to a dramatic worldwide increase of UAV photogrammetry-based 3D reconstructions. Nevertheless, accurate high resolution photogrammetric reconstructions typically require quite long data processing procedures, which make them less suitable for real-time applications.
This work aims at investigating the use of a low cost Time of Flight (ToF) camera, combined with an Ultra-Wide Band (UWB) positioning system, mounted on a drone, in order to enable quasi real time 3D reconstructions of small to mid-size areas, even in locations where Global Navigation Satellite Systems (GNSSs) are not available.
The proposed system, tested on a small area on the Italian Alps, provided high resolution mapping results, with an error of few centimeters with respect to a terrestrial close-range photogrammetry survey conducted on the same day.

1. INTRODUCTION

The worldwide spread of drones and the development of new remote sensing techniques opened the opportunity for quick high resolution topographic reconstructions and autonomous inspections to better understand Earth surface processes and monitoring areas affected by natural hazards (Tarolli, 2014, Sofia, 2020).

Nowadays, high resolution topographic models are typically obtained either by means of airborne or terrestrial Light Detection and Ranging (LiDAR), or by means of photogrammetry. LiDAR is usually considered as the state-of-the-art for 3D model generation, however the comparable lower cost of photogrammetry makes the latter very attractive in certain applications (Prosdocimi et al., 2015).

Differently from LiDAR, photogrammetric 3D reconstruction typically requires a relatively long processing procedure to produce a high resolution spatial model of the area of interest. Furthermore, image mismatches may lead to reconstruction errors; and camera self-calibration (Fraser, 1997), which is often used in the structure from motion approach, may cause distortions in the produced model, in particular when dealing with UAV photogrammetry with nadir camera (Fraser, 2018).

This work aims at obtaining low cost, high resolution topographic reconstructions for geomorphometric assessment. Differently from the typical relatively long processing of photo-

grammetry, this work aims at quickly obtaining geospatial information by using a Time of Flight (ToF) camera attached to a drone.

Thanks to technological developments over the last decade, ToF cameras have become cheaper, smaller and provide improved performance : the introduction of the Microsoft Kinect v2 marked a significant cost reduction, while ensuring a quite high resolution and frame rate (Lachat et al., 2015, Steward et al., 2015).

This work considers the use of a commercial off-the-shelf ToF camera, conceptually similar to Kinect v2, but smaller and lighter: the Pico Zense DCAM710, shown in Fig. 1.

Figure 1. Pico Zense DCAM710.

The characteristics of the Pico Zense DCAM710 make it an ideal solution for acquiring high resolution spatial information from a drone (ground sample distance (GSD) smaller than 1 cm, which is clearly a sufficiently high resolution for most of the analysis related to geospatial data (Boreggio et al., 2018)). In this work the ToF camera was attached to the bell of a low cost quadcopter, the Parrot Bebop 2 ($\approx$ \$300), and used to simultaneously acquire RGB and depth information flying a few meters

---

* Corresponding author.

above the ground (compatible with the maximum range of the ToF camera).

Thanks to a proper calibration of the camera and of the navigation system (Chiang et al., 2015, Lo et al., 2015), the depth information provided by the ToF camera can be quickly converted into point clouds, which can be aligned by exploiting both the information provided by the drone positioning system and space resection of the RGB camera images.

To such aim, RTK GNSS is typically used in the current generation of commercial systems. Instead, this work investigates the use of an Ultra-Wide Band (UWB) positioning system, which guarantees the usability of the system even when GNSS is not available (Section 3)

The mapping results obtained with the considered system are compared with those of a standard close-range photogrammetry survey conducted on a test site on the Italian Alps, on October 26, 2019 (Section 4).

2. PHOTOGRAMMETRIC SURVEY

A Canon G7X camera (20.2 MPix), with settings fixed at constant values during all the photogrammetric survey (1/200-s shutter speed, f/10 aperture, ISO 125, 8.8-mm focal length, i.e., 35 mm equivalent: 24 mm) was used to acquire images of the area of interest, with resolution of 1.4 mm per pixel, approximately. Ninety images have been collected and successively processed with commercial photogrammetric software (Agisoft Metashape), leading to the production of a cloud with 7 million points. Self-camera calibration, image alignment and dense point cloud generation was performed in about two hours. Self-camera calibration led to a reprojection error of 0.68 pixels.

Model scaling factor and georeferencing was achieved by means of eight ground control points (GCPs): a Topcon HiPer V GNSS receiver, working in Network Real-Time-Kinematic (N-RTK) operative mode, allowed to measure GCP positions with accuracy at centimeter level.

Fig. 2 shows a drone view of the area of interest, along with the GCPs, which are approximately homogeneously distributed.

Fig. 3 shows a top view of the reconstructed area, obtained with the standard close-range photogrammetry approach described in this section. Horizontal and vertical axes in the local reference system used in the figure correspond to Easting and Northing directions, respectively.
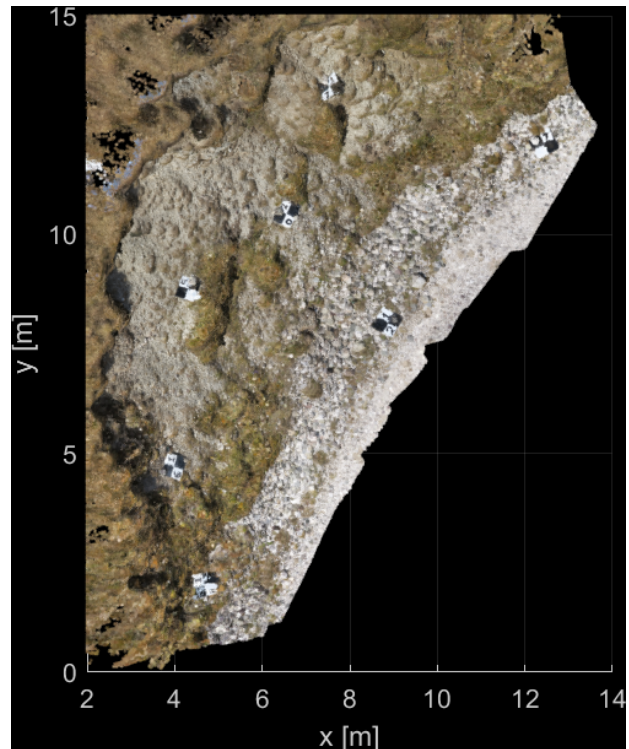

Figure 2. Drone view of the case study area.

Figure 3. Top view of the area reconstructed via close-range photogrammetry.

3. QUADRICOPTER-TOF CAMERA SURVEY

The mapping system described in this section is based on the use of a Pico Zense DCAM710 ToF camera, which was attached to a Parrot Bebop 2.

Pico Zense DCAM710 camera is a low cost ($\approx$ \$200), small (10.3 cm$\times$3.3 cm$\times$2.2 cm) and lightweight device that can simultaneously acquire RGB, IR images and depth information at a 30 frames-per-second frequency, with range accuracy typically at the centimeter level (Masiero et al., 2019), and maximum resolution as reported in Table 1.

information type	resolution
RGB	1920 × 1080
IR	640 × 480
Depth	640 × 480

Table 1. Resolution of the outputs of the Pico Zense DCAM710 camera.

Field of view of the depth sensor is $69° \times 51°$.

Parrot Bebop 2 is an affordable drone, ensuring 25 minutes of flying time with its standard battery in normal flying conditions. However, the battery life is significantly influenced by the payload, leading to about a factor two reduction in the working conditions considered in this paper. Furthermore, the presence of some extra sensors attached to the drone reduces its flying stability. The maximum speed and tilt angle were limited to reduce the issues related to such potential instability during the survey. Clearly, the small drone category may be highly affected by this kind of issues, whereas large drones can usually easily carry some extra sensors.

Given the relatively short maximum range of the ToF camera, the drone is required to fly quite close to the ground. Despite some possible issues due to obstacle avoidance, this allows to acquire high resolution depth and RGB images of the area of interest.

Each depth image can be easily converted to a 3D point cloud, with coordinates expressed according to the ToF camera local reference system. The conversion from camera local coordinate system to the survey reference system is achieved by combining the information provided an Ultra-Wide band positioning system and the Pico Zense standard RGB imagery.

To be more specific, a Pozyx UWB positioning system was used in this work. A set of eight Pozyx anchors was used to track a Pozyx UWB tag (see Fig. 4), attached to the drone during the flight. In this case, the geographic anchor positions were measured with the same Topcon HiPer V GNSS receiver used for the photogrammetric survey. Nevertheless, accurate anchor positions in a local reference system can be obtained also in GNSS-denied environments using standard surveying techniques, e.g. by using a total station. The UWB positioning system provides position estimates of the drone position in real time with accuracy at decimeter level (Goel et al., 2017, Gabela et al., 2019).
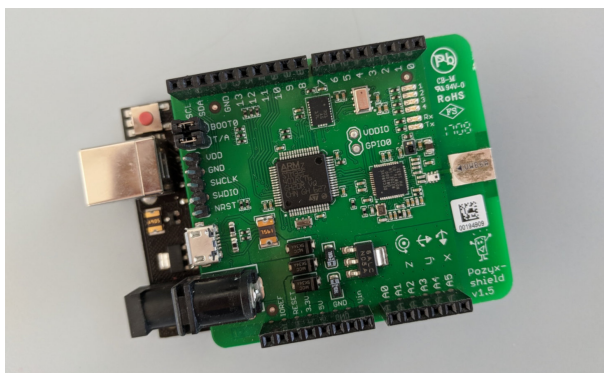

Figure 4. Pozyx tag.

Then, a visual odometry approach can be applied to the images acquired by the RGB camera of the Pico Zense DCAM710 (Huang et al., 2011): the outcome of such visual information processing can be used to determine estimates of the roto-translations (in a local non-metric reference system) of the ToF camera during the survey, if a suitable calibration of the RGB and depth camera system is available.

The combination of the information provided by the UWB positioning system and by visual odometry allows to assess the ToF pose during the depth image acquisitions, and hence can be used to properly combine the point clouds associated to each depth image.

If visual information is acquired at a sufficiently high frame rate, visual odometry typically provides a reliable dead reckoning solution to the drone positioning problem. However, such solution is expressed in a local reference system, e.g. with respect to the local camera reference system of the first image acquisition. Instead, UWB system computes the rover position in accordance to its own coordinate system, which in this case is derived from the GNSS measurements of the anchor locations. Despite the UWB positioning error is usually larger than the visual odometry one, the integration of such two systems

(done making a least squares estimation of the rigid transformation between them) allows to express the vision-based drone positioning results in a properly defined coordinate system, e.g. in geographic coordinates.

It is worth to notice that, even when GNSS is not available, comparison between the results obtained in different surveys can be obtained by referencing the obtained 3D point cloud according to certain points assumed to be at fixed positions during all the time interval of interest.

Point cloud/image alignment is currently done in few minutes, depending on the number of acquired frames, in post-processing right after the drone flight. Nevertheless, if the device used to collect the data acquired by the PicoZense camera is provided with a sufficient computational power, point cloud generation can potentially be performed also in real-time, along with the computation of geomorphometric parameters of the surveyed area, if the device used to collect the data acquired by PicoZense DCAM710 is provided with a sufficient computational power.

To be more specific, ninety-four depth images of the area of interest were acquired by the Pico Zense camera, and combined off-line according to the process described above in approximately two minutes. The outcome of such process was a cloud of 16 million points.

4. RESULTS

First, the trajectory estimated with visual odometry was compared with the one provided by UWB positioning system, aiming at estimating the rigid transformation between their coordinate systems.

The main statistical characteristics of the fitting error between such trajectories are reported in Table 2. In particular the 3D root mean square error (RMSE) was 23.1 cm, whereas the 2D (planimetric) RMS fitting error was 17.3 cm. A significant height error is notable in terms of both RMSE and maximum error.

Table 2. Comparison between UWB and vision-based UAV positioning

	RMSE [cm]	max [cm]
2D	17.3	34.7
height	15.3	49.0
3D	23.1	51.5

Then, Figure 5 shows the error distribution between vision-based and UWB positioning. Figure 5(a) shows the 2D RMSE between them, whereas Figure 5(b) reports their height difference.

The photogrammetric and ToF point clouds were compared computing the point-to-point distance: the derived histogram is shown in Fig. 6, whereas Table 3 reports the comparison between such point clouds in terms of average, root mean square and median absolute deviation (computing also a robust approximation of the mean with the median) of the 2D distances and height differences between closest point. It is worth to notice that, in order to equally weight all the regions in the reconstructed area, the statistics of Table 3 were obtained by dividing the overall area in 5 cm×5 cm subsets and considering 40 (randomly sampled) points per each of such subset.
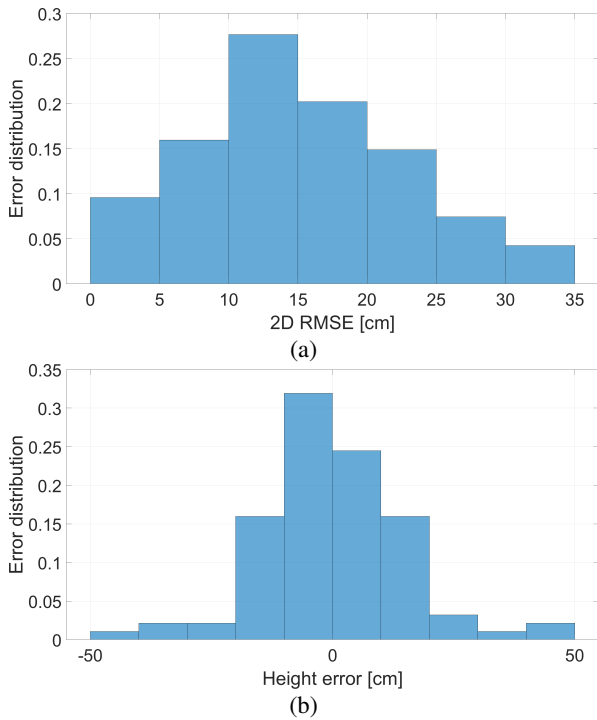
Figure 5. Comparison between UWB and vision-based positioning. (a) 2D root mean square error (RMSE), (b) height error distribution.
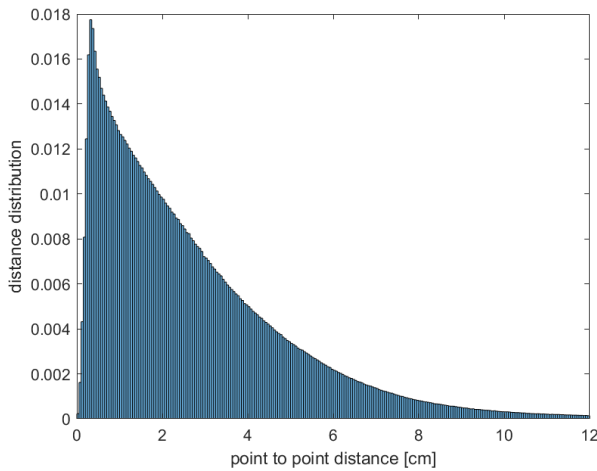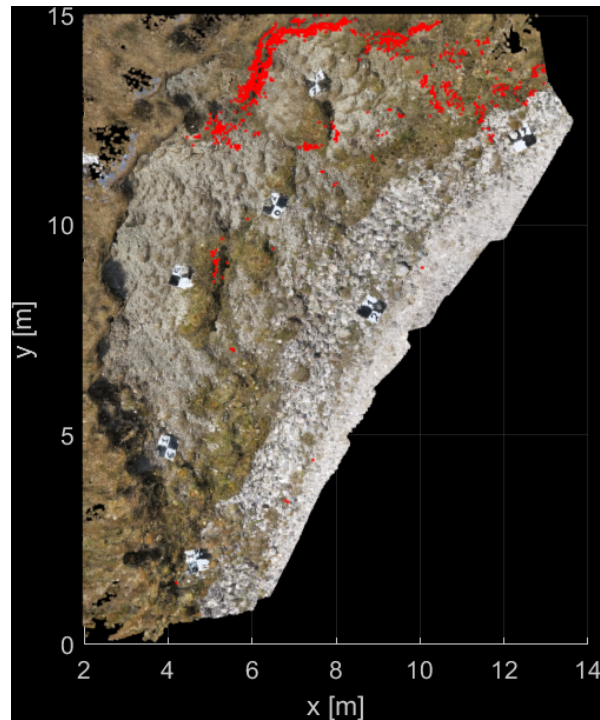

Figure 7. Points (in red) with largest position differences between point clouds generated with photogrammetry and ToF camera.

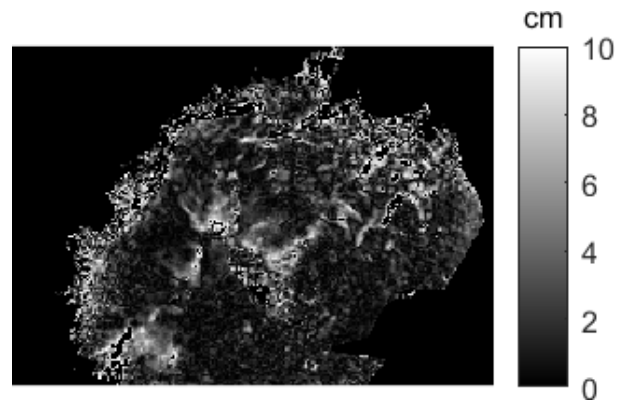ence of the two DEMs in the northern region of the case study area.

Figure 6. Point-to-point distance between point clouds generated with photogrammetry and ToF camera.

Figure 8. DEM absolute difference (northern region of the considered study area): photogrammetric survey vs ToF camera.

Fig. 7 shows in red the areas associated to larger differences between the two point clouds.

Furthermore, Fig. 9 shows the contour lines (with 40 cm intervals) computed by using the two DEMs, in the same area already considered in Fig. 8.

Table 3. Comparison between point clouds generated with close-range photogrammetry and ToF camera

	avg [cm]	RMS [cm]	MAD [cm]
2D	1.1	1.5	0.5
height	0.0	5.3	1.9

High resolution Digital Elevation Models were computed for both the models, with a spatial resolution of 5 cm. Average height difference between the obtained DEMs was 0.1 cm, and standard deviation of 5.7 cm. Fig. 8 shows the absolute differ-

5. DISCUSSION

The proposed method allowed to quickly obtain a high resolution spatial representation of the area of interest. The main advantage with respect to the standard photogrammetric survey is clearly that of notably reducing the processing time to produce the 3D spatial information, e.g. two orders of magnitude smaller, from hours to about two minutes of processing time. Given the dramatic computational burden reduction, the
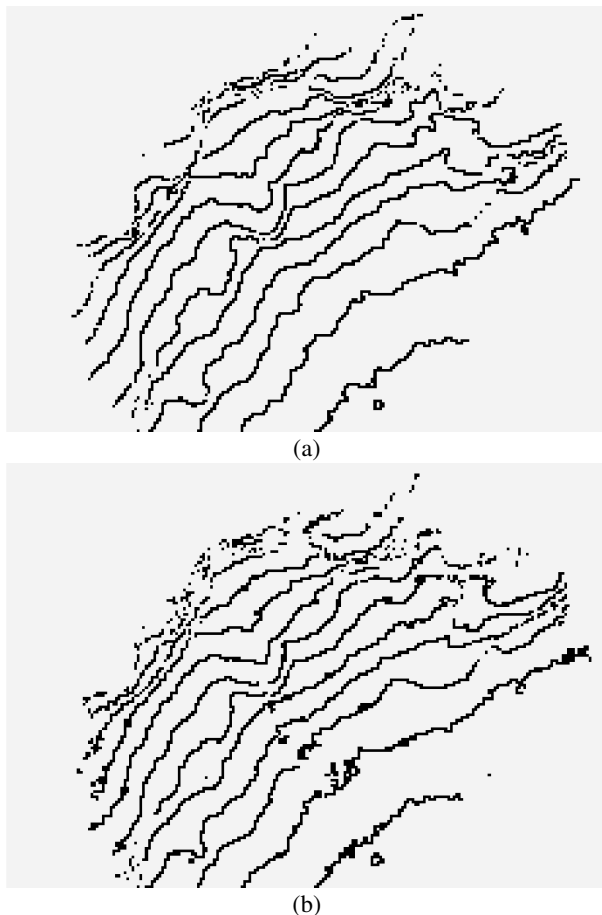
(a)

(b)

Figure 9. Contour lines (40 cm intervals) in the norther region of the considered study area: (a) photogrammetric survey, (b) ToF camera.

proposed method can be of great interest for those applications requiring quick (e.g. (quasi) real-time) 3D reconstructions.

The coordinate reference system of the produced point cloud is determined by assessing a rigid transformation between the vision-based positioning and the UWB estimates of the drone flight track. Table 2 shows that the position differences between such two positioning methods are typically of few decimeters (2D, 3D and height RMS errors are in the 15–25 cm interval), which is in agreement with the typical accuracy of UWB positioning in line-of-sight ranging conditions. Vision-based trajectory is also affected by some estimation error, however, in these working conditions its level is usually smaller than the UWB one.

Furthermore, Fig. 5 shows the planimetric and altimetric positioning error distributions, which show the presence of few quite large height errors. Nevertheless, both planimetric and altimetric errors are with high probability less than 25 cm. Vision based techniques can also be implemented to further improve the rigid transformation such two reference systems.

The produced point cloud and DEM are quite similar to those produced by the photogrammetric survey: Fig. 6 and 8 show that the typical differences between the spatial information generated by the two considered methods is typically of few centimeters. In particular, larger differences (red areas shown in Fig. 7) are usually visible in vegetated areas, which is quite

reasonable, given the different behavior of the two 3D information generation methods with respect to such areas.

Table 3 shows also the results of the comparison between between the photogrammetric and ToF-based point clouds. Planimetric error is slightly smaller than altimetric one. A small bias in the planimetric error (1.1 cm) has a notable effect on the overall planimetric RMS error (1.5 cm). Furthermore, the significant difference between RMS and MAD in the height error shows that the RMS value may also be influenced by the presence of outliers in the reconstruction.

It is also worth to mention that the few centimeter differences between the two point clouds shown in Table 3 are in agreement with previous results on the reconstruction error of the ToF camera employed in this work (Masiero et al., 2019).

The point cloud comparison results reported in Table 3 are quite similar to the height differences obtained comparing the two produced DEMs. Slight differences between such height difference comparisons are probably mostly motivated by the different approach considered in such computation. Indeed, it is also worth to notice that results in Table 3 were obtained by partitioning the case study area in 5 cm×5 cm subsets, and randomly sampling 40 points from each of such subset (clearly only from those which contain at least such number of points).

The contour lines shown in Fig. 9 shows that the ToF-point cloud seems to be affected by a higher of noise with respect to the photogrammetric one. Such result can be due several factors:

- the noise of the ToF camera, which linearly increases with the distance from the object (Masiero et al., 2019), and that was at some centimeter level in this case.

- point clouds obtained by the depth images were aligned exploiting the combination of information provided by the RGB camera and the UWB system, however some noise might be the result of the residual error in such registration process. The development of a further registration error reduction method will be object of investigation in our future works, in particular exploiting a simultaneous localization and mapping (SLAM)-like formulation of the problem (Leonard, Durrant-Whyte, 1991).

- the photogrammetric point cloud, computed with Agisoft Metashape, has been obtained by also applying a mild filtering of the 3D point cloud, which might have had a denoising and regularization effect on the final outcome of the reconstruction algorithm.

The main limitation of the proposed method is probably the restriction on the flight altitude, which is severely limited by the short maximum range of the ToF camera. Such restriction may cause some obstacle avoidance issues.

It is also worth to mention that the proposed method can be used also in GNSS-denied environments: in such case the obtained 3D reconstruction is expressed in the local reference system of the UWB system. In monitoring applications, when surveys are typically done periodically, the results obtained in different surveys can be compared by using certain reference points, which should be chosen in static (time invariant) positions.

6. CONCLUSIONS

Despite LiDAR can probably be considered the state-of-the-art for accurate topographic model generation (Guarnieri et al., 2015), the use of drones, typically provided with high resolution cameras, is a widely spread low cost method for obtaining high resolution models (Lo et al., 2015). Furthermore, recently other mobile devices, such as smartphones, have been also considered for producing 3D spatial information (Prosdocimi et al., 2017, Fissore et al., 2018).

The approach proposed in this paper generalizes the investigation reported in (Nitsche et al., 2013), showing the potential of the use of ToF cameras in UAV surveys, in particular in terms of very low cost, quick and high resolution topographic reconstruction of small to middle-size sites, whose UAV photogrammetric accurate reconstruction (Nex, Remondino, 2014) typically requires much longer post-processing time.

The results obtained in a test site on the Italian Alps, showed that the difference between the generated 3D model and that provided by a standard photogrammetric survey is at few centimeters level. In particular, the model produced by the proposed model seems to be affected by a higher noise level, however it has been computed in few minutes, far less than the time needed to generate the photogrammetric one.

REFERENCES

Boreggio, M., Bernard, M., Gregoretti, C., 2018. Evaluating the differences of gridding techniques for Digital Elevation Models generation and their influence on the modeling of stony debris flows routing: A case study from Rovina di Cancia basin (North-eastern Italian Alps). *Frontiers in Earth Science*, 6, 89.

Chiang, K.-W., Tsai, M.-L., Naser, E.-S., Habib, A., Chu, C.-H., 2015. New Calibration Method Using Low Cost MEM IMUs to Verify the Performance of UAV-Borne MMS Payloads. *Sensors*, 15(3), 6560–6585.

Fissore, F., Masiero, A., Piragnolo, M., Pirotti, F., Guarnieri, A., Vettore, A., 2018. Towards surveying with a smartphone. *New Advanced GNSS and 3D Spatial Techniques*, Springer, 167–176.

Fraser, C., 2018. Camera calibration considerations for UAV photogrammetry. *ISPRS TC II Symposium: Towards Photogrammetry 2020*.

Fraser, C. S., 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(4), 149 - 159.

Gabela, J., Retscher, G., Goel, S., Perakis, H., Masiero, A., Toth, C., Gikas, V., Kealy, A., Koppányi, Z., Błaszczak-Bak, W. et al., 2019. Experimental evaluation of a UWB-based cooperative positioning system for pedestrians in GNSS-denied environment. *Sensors*, 19(23), 5274.

Goel, S., Kealy, A., Gikas, V., Retscher, G., Toth, C., Grejner-Brzezinska, D., Lohani, B., 2017. Cooperative Localization of Unmanned Aerial Vehicles Using GNSS, MEMS Inertial, and UWB Sensors. *Journal of Surveying Engineering*, 143(4), 04017007.

Guarnieri, A., Masiero, A., Vettore, A., Pirotti, F., 2015. Evaluation of the dynamic processes of a landslide with laser scanners and Bayesian methods. *Geomatics, Natural Hazards and Risk*, 6(5-7), 614–634.

Huang, A. S., Bachrach, A., Henry, P., Krainin, M., Maturana, D., Fox, D., Roy, N., 2011. Visual odometry and mapping for autonomous flight using an RGB-D camera. *International Symposium on Robotics Research (ISRR)*, 2.

Lachat, E., Macher, H., Landes, T., Grussenmeyer, P., 2015. Assessment and calibration of a RGB-D camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling. *Remote Sensing*, 7(10), 13070–13097.

Leonard, J., Durrant-Whyte, H., 1991. Simultaneous map building and localization for an autonomous mobile robot. *Intelligent Robots and Systems '91. 'Intelligence for Mechanical Systems, Proceedings IROS '91. IEEE/RSJ International Workshop on*, 1442–1447 vol.3.

Lo, C., Tsai, M., Chiang, K., Chu, C., Tsai, G., Cheng, C., El-Sheimy, N., Habib, A., 2015. The Direct Georeferencing Application and Performance Analysis of UAV Helicopter in GCP-Free Area. *ISPRS - International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(1), 151.

Masiero, A., Guarnieri, A., Vettore, A., 2019. Assessment of a Portable ToF Camera and Comparison with Smartphone Stereo Vision. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W17, 187–193. doi.org/10.5194/isprs-archives-XLII-2-W17-187-2019.

Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: a review. *Applied Geomatics*, 6(1), 1–15.

Nitsche, M., Turowski, J. M., Badoux, A., Rickenmann, D., Kohoutek, T. K., Pauli, M., Kirchner, J. W., 2013. Range imaging: a new method for high-resolution topographic measurements in small-and medium-scale field sites. *Earth Surface Processes and Landforms*, 38(8), 810–825.

Prosdocimi, M., Burguet, M., Di Prima, S., Sofia, G., Terol, E., Comino, J. R., Cerdà, A., Tarolli, P., 2017. Rainfall simulation and Structure-from-Motion photogrammetry for the analysis of soil water erosion in Mediterranean vineyards. *Science of the Total Environment*, 574, 204–215.

Prosdocimi, M., Calligaro, S., Sofia, G., Dalla Fontana, G., Tarolli, P., 2015. Bank erosion in agricultural drainage networks: new challenges from structure-from-motion photogrammetry for post-event analysis. *Earth Surface Processes and Landforms*, 40(14), 1891–1906.

Sofia, G., 2020. Combining geomorphometry, feature extraction techniques and Earth-surface processes research: The way forward. *Geomorphology*, 107055.

Steward, J., Lichti, D., Chow, J., Ferber, R., Osis, S., 2015. Performance assessment and calibration of the Kinect 2.0 time-of-flight range camera for use in motion capture applications. *FIG Working week 2015*, 1–14.

Tarolli, P., 2014. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. *Geomorphology*, 216, 295–312.