

# A NEW APPROACH FOR PROGRESSIVE DENSE RECONSTRUCTION FROM CONSECUTIVE IMAGES BASED ON PRIOR LOW-DENSITY 3D POINT CLOUDS

Z. Lari \*, N. El-Sheimy

Department of Geomatics Engineering, University of Calgary, 2500 University Dr. NW, Calgary, Canada, T2N 1N4 -  
Zlari, elsheimy@ucalgary.ca

Commission III, WG III/6

**KEY WORDS:** 3D Reconstruction, Point Cloud, Emergency Response, Densification, Triangulation, Matching, Feature Descriptor, Computational Efficiency

## ABSTRACT:

In recent years, the increasing incidence of climate-related disasters has tremendously affected our environment. In order to effectively manage and reduce dramatic impacts of such events, the development of timely disaster management plans is essential. Since these disasters are spatial phenomena, timely provision of geospatial information is crucial for effective development of response and management plans. Due to inaccessibility of the affected areas and limited budget of first-responders, timely acquisition of the required geospatial data for these applications is usually possible only using low-cost imaging and georeferencing sensors mounted on unmanned platforms. Despite rapid collection of the required data using these systems, available processing techniques are not yet capable of delivering geospatial information to responders and decision makers in a timely manner. To address this issue, this paper introduces a new technique for dense 3D reconstruction of the affected scenes which can deliver and improve the needed geospatial information incrementally. This approach is implemented based on prior 3D knowledge of the scene and employs computationally-efficient 2D triangulation, feature descriptor, feature matching and point verification techniques to optimize and speed up 3D dense scene reconstruction procedure. To verify the feasibility and computational efficiency of the proposed approach, an experiment using a set of consecutive images collected onboard a UAV platform and prior low-density airborne laser scanning over the same area is conducted and step by step results are provided. A comparative analysis of the proposed approach and an available image-based dense reconstruction technique is also conducted to prove the computational efficiency and competency of this technique for delivering geospatial information with pre-specified accuracy.

## 1. INTRODUCTION

Over the past two decades, the changes in the climate and associated increasing concentration of atmospheric greenhouse gas emissions have dramatically affected our environment (Van Aalst, 2006). These changes have caused destructive natural disasters which negatively impact all aspects of sustainable development including infrastructures, energy reservoirs, transportation systems, forestry, farming, and agriculture. Moreover, these catastrophes have significantly affected communities' well-being and public health. In order to effectively assess, manage, and diminish the vulnerability and risks associated to these events and improve the resilience of the communities, developing timely disaster management and emergency response plans is needed (Wisetjindawat et al., 2014). The effective implementation of such plans is contingent on having timely access to accurate and up-to-date geospatial data which provide valuable insights into location, nature, extent, and progress of the areas under influence (Zlatanova and Li, 2008). Due to inaccessibility and hazardousness of the affected areas, the required data are usually collected using low-cost imaging and georeferencing sensors onboard unmanned platforms (Giordan et al., 2017). The ability of these platforms to manoeuvre in closer distances and slower speed in these areas and thus providing near real-time high-resolution geospatial data makes them optimal alternatives for rapid-response mapping applications (S. Thenkabail, 2015).

The collected geospatial data onboard unmanned platforms are then processed to provide the required geospatial information (including dense coloured 3D point clouds and textured 3D models) for decision makers and first responders in such emergency situations. So far, different approaches have been proposed and implemented for 3D scene reconstruction which are classified into active and passive methods. Active 3D reconstruction methods are implemented based on laser scanning data (Hernandez et al., 2007). In spite their potential to provide high quality 3D information from the scanned objects, these methods are not always applicable due to costly initialization and operation of data acquisition tools. On the other hand, passive 3D reconstruction approaches only use photometric information to generate 3D reconstructed models. The approaches in this category, which are implemented based on a set of overlapping images captured from a scene, can be categorized into four classes. The methods in first class are implemented by partitioning the scene into a 3D grid (e.g., voxels), computing a cost function for individual voxels, and reconstructing the voxels with costs below a threshold in a same pass (Seitz and Dyer, 1999; Hertzmann and Seitz, 2005). Although these methods are widely utilized due to their simplicity and uniformity, they suffer from two major shortcomings. First, they do not involve a regularization procedure and second, the excluded voxels in a pass cannot be restored later during the reconstruction procedure (Ladikos, 2011). The second class of these approaches is carried out by generating a preliminary mesh or level-set representation

\* Corresponding author

of the scene and iteratively evolving the initial surfaces to decrease or minimize a cost function (Faugeras and Keriven, 2002; Pons et al., 2007). The drawback of these methods is that they are implemented locally and might fall into local minima during the minimization procedure.

The methods in third class, image-space methods, compute a set of depth maps initially. Then, they enforce consistency constraints between the computed depth maps (Szeliski, 1999; Gargallo and Sturm, 2005) or merge them in a post processing step to ensure a consistent 3D representation of the scene (Narayanan et al., 1998). These techniques avoid resampling the geometry on a 3D domain and choose a 2D representation alternatively. Different from the aforementioned approaches, the methods in fourth class firstly detect, extract, and match a set of features points within a scene. The surface patches are then fitted to the reconstructed feature points and grown by adding new patches at the borders of initial ones (Faugeras et al., 1990; Manassis et al., 2000). These approaches suffer from limited feature correspondence in close-range imagery.

In spite of rapid delivery of the collected data onboard unmanned platforms, the existing 3D reconstruction techniques are not yet capable of timely/incrementally data processing and transferring the processed information (3D models and other mapping products) for disaster management and emergency response applications. These methods mainly focus on highly accurate data processing while sacrificing the computational efficiency as they require execution times which are not suitable for these applications (Remondino et al., 2014). Hence, new image-based 3D reconstruction techniques should be developed for timely provision of geospatial information from the affected areas.

In this paper, a new approach for progressive dense 3D scene reconstruction from consecutive images collected onboard unmanned platforms is introduced. To optimize the computational efficiency and reduce the required processing time, the proposed approach is implemented based on prior knowledge of the scene (sparse 3D point clouds collected using a laser scanner or derived using image-based sparse point cloud generation techniques) from the scanned scenes. This approach attempts to achieve a denser overview of the scanned scene in a shorter time and provide incrementally-refined 3D information for different emergency response applications. Furthermore, the proposed approach is capable of delivering 3D point clouds with required accuracy and avoid further unnecessary processing. Figure 1 shows the outline of the proposed method.

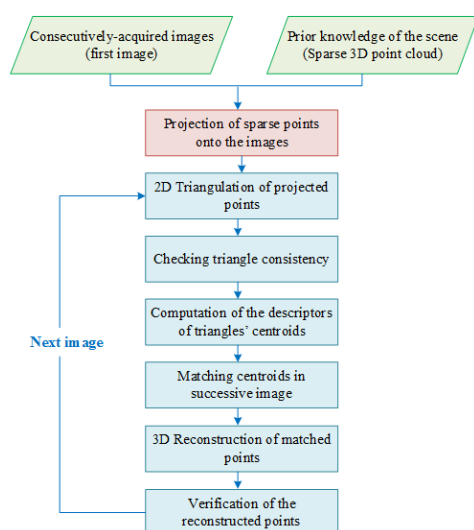


Figure 1. The outline of the proposed incremental 3D dense reconstruction technique

This paper starts by the introduction of the proposed approach for progressive 3D scene reconstruction using consecutively-acquired images and prior knowledge of the scene (sparse 3D point cloud). Afterwards, the feasibility of the proposed approach is verified by conducting an experiment using imagery and low-density LiDAR point cloud collected over the same scene. The achieved experimental results prove the competency of this technique for providing progressively denser representation of the scanned scene. Finally, concluding remarks and recommendations for future research work are presented.

## 2. METHODOLOGY

In this section, the proposed algorithm for progressive 3D scene reconstruction using consecutively-acquired images onboard unmanned platforms is described. Since this approach is implemented based on prior knowledge of the scene, the existing scene points from a laser scanning data or a sparse point cloud achieved through other techniques, are firstly projected on the first captured image. Secondly, a 2D triangulation of the projected points is established and the suitability of the achieved triangles for scene reconstruction is examined. The centroids of triangles passing suitability conditions are then computed to provide scene information in areas between known points instead of their vicinity. In the third step, a computationally efficient descriptor is computed for individual centroid points to find and match their correspondents within successive overlapping images. The matched centroid points are finally reconstructed in 3D space and verified using next images. This procedure is successively implemented for all images to achieve a denser/more accurate 3D view of the scanned scene. The detailed explanation of these steps is provided in following subsections.

### 2.1 Scene-to-Image Point Projection

In the first step of the proposed approach, prior knowledge of the scanned scene is employed as an auxiliary data to optimize the efficiency of 3D reconstruction procedure. The objective of this step is to avoid initial feature detection and matching steps using a previously-acquired low-density laser scanning data or a sparse point cloud over the scene. In order to initiate 3D reconstruction procedure, the 3D points in this point cloud are projected onto the first acquired image – that has certain percentage of overlap and corresponding features in next images – using collinearity equations (Novak, 1992; Habib et al., 2007). This projection is performed using available camera calibration and georeferencing parameters of the first image which are reliable in early mission phase. The projected 2D points will be then used as the basis of 3D reconstruction procedure.

### 2.2 2D Triangulation

In order to identify reasonable points to densify the existing sparse point cloud, a 2D triangulation technique is used to generate a mesh of the projected points and partition the image space into a set of triangular facets. In this approach, a modified version of Delaunay triangulation technique (Delaunay, 1934) is used for the image tessellation due to its computational efficiency, robustness, ease of parallelism, and ability to preserve geometry (Shewchuk, 2012). The suitability of the derived triangles for 3D scene reconstruction is then examined based on a set of conditions as follows:

1. Expected accuracy: A triangle is considered to be suitable if its surface is larger than a predefined threshold. This threshold is specified based on the expected accuracy of the reconstructed 3D model.

2. Interior angles: An obtuse triangle – including an interior angle less than  $15^\circ$  in our case – is not considered a suitable one and will be rejected (Brunnett et al., 2012).

The centroids of the triangles, satisfying the abovementioned conditions, are then computed and selected as the next points to be reconstructed as they are located in maximum distance from triangles' nodes and provide optimal visual entropy (Figure 2).

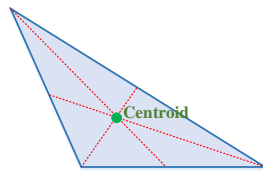


Figure 2: Identification of next points of scene to be reconstructed

The vertices of the achieved triangles are then projected onto the next image to limit the search space for their matching points. The corresponding triangles in an image pair will be matched and considered for 3D reconstruction if they pass the following conditions:

1. Ratio between areas of triangular facets: The corresponding triangular facets will be matched if the ratio between their areas is less than 2 (in our case).
2. Orientation of facet vertices: The corresponding triangles will be matched if the orientation of their vertices is preserved in both images.

The triangles satisfying these conditions in the second image are detected as confined search spaces for matching centroid points. The rejected triangles will be handled later by optimization of the triangular mesh. Figure 3 schematically shows the described 2D triangulation and established centroid points for 3D reconstruction.

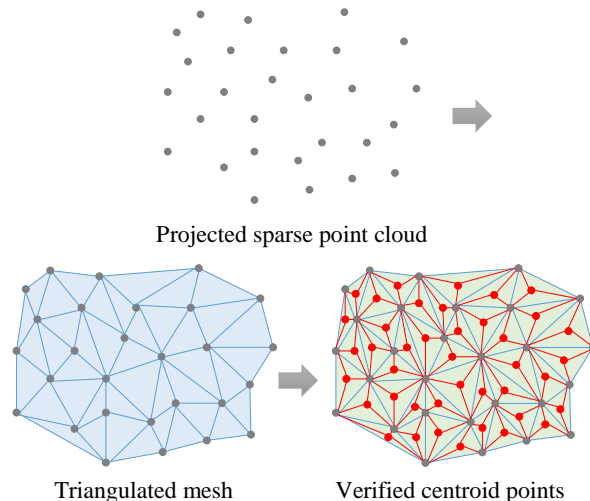


Figure 3: 2D triangulation and identification of candidate points for 3D reconstruction

### 2.3 Computation of Centroids' Descriptors

In order to efficiently find matches for the identified feature points – centroid points of triangles passing suitability conditions, reliable and computationally-efficient feature descriptors need to be selected and accurately estimated. The proposed progressive dense reconstruction procedure employs an efficient feature descriptor, called Synthetic BAis (SYBA), which has been specifically designed for accurate feature matching in real-time applications (Desai, 2015). This feature descriptor provides high matching accuracy and optimized computational efficiency which are essential in our progressive

dense reconstruction procedure. One should note that the utilized feature descriptor requires minimal computational resources and can be optimally implemented on unmanned platform's onboard embedded processors.

The selected feature descriptor is estimated for the centroid points of an established triangles in first image and multiple candidate pixels within the bounded search spaces for their matches (corresponding triangles passing validity conditions) in the next image.

### 2.4 Matching and 3D Reconstruction of Centroids

In the fourth step, the centroid points of suitable triangles in the first image are matched with their correspondents in the next image according to the estimated feature descriptors. To improve the computational efficient of the matching procedure, the search for individual centroid points in first image is limited to the intersection of their associated epipolar lines and corresponding triangles in the next image (Figure 4). As mentioned before, the selected feature descriptor is estimated for all the pixels lying on this intersection line and compared to the estimated descriptor for the investigated centroid point in the first image. The centroid point will be matched one of the pixels on the bounded search space, if their matching quality – computed according to the derived feature descriptors is more than a pre-defined threshold. Furthermore, there should be no other pixels within search region that can be matched to the centroid point with the same quality. In this technique, the comparison between the estimated feature descriptors is performed based on the sum of the absolute differences – instead of Euclidean and Mahalanobis distances – to avoid computational complexity of the feature matching procedure (Desai, 2015).

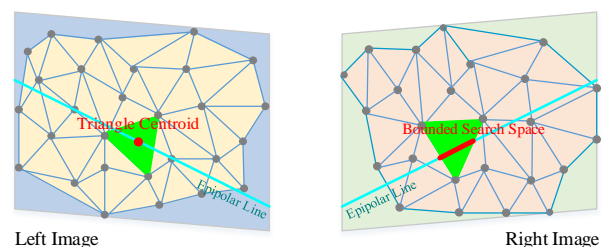


Figure 4: The bounded search space for conjugate feature points

Finally, the 3D coordinates of the matched image points are estimated using a Direct Linear Transformation (DLT). The selection of DLT for 3D triangulation is due to its accepted computational efficiency, accuracy, and simplicity for rapid-response mapping applications. The reconstructed 3D point will be then projected onto the other 3 images which include three nodes of their associated triangles – once they are being processed for point cloud triangulation – to identify possible outliers due to occlusions or false matching problems. The investigated points are rejected if none of these images confirms them at the projected positions. However, these points might be reconstructed in next steps (within other image combinations). The steps 2 - 4 are repeated for consecutive images while adding the reconstructed 3D points onto the modified 2D TIN. The midpoints of the newly generated triangles are then matched within next images and reconstructed using the same procedure. The densification of prior sparse point cloud is continued until the required level of detail and accuracy for decision makers and first responders is achieved.

## 3. EXPERIMENTAL RESULTS

In this paper, we presented a new technique for progressive image-based 3D dense reconstruction based on prior knowledge



of the scene. The proposed approach attempts to address and resolve the computational inefficiency of traditional image-based dense reconstruction methods by avoiding feature detection in images, using low-density point clouds which has been previously collected or generated over the scene, and implementing computationally efficient triangulation, feature descriptor estimation, feature matching and 3D point reconstruction algorithms. In order to evaluate the performance and computational efficiency of the proposed approach, an experiment using a set of consecutive images collected onboard a UAV platform and a sparse point cloud (previously acquired over the scene) is conducted and step by step results are provided. The investigated dataset includes 28 consecutive images acquired using a GoPro Hero 3+ camera mounted on a DJI Phantom 2 UAV at an average flying height of 26m and a low-density airborne laser scanning data collected by an Optech ALTM 3100 laser scanner over a complex building located in Calgary, AB, Canada. Table 1 summarizes the specifications of the used laser scanning point cloud and imagery as presented by data providers. Figure 5.a shows an overview of the provided low-density laser scanning point cloud and figure 5.b shows some of the captured images over the investigated scene.

Dataset	Photogrammetric data	Laser scanning data
System	GoPro HERO 3+ black edition	Optech ALTM 3100
Date acquired	2015	2013
Number of overlapping images/scans	28	4
Average GSD/point density	1.7 cm	3 pts/m <sup>2</sup>
Planimetric accuracy	1.7 cm	13 cm
Vertical accuracy	15 cm	6 cm

Table 1. A summary of the characteristics of imagery and laser scanning point cloud utilized in this experiment

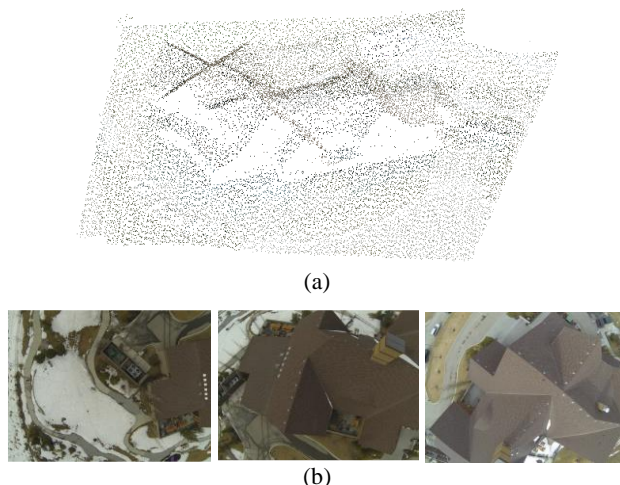


Figure 5. (a) Low-density laser scanning point cloud and (b) some of the captured images over the investigated scene

Figures 6.(a - c) show the progressively densified point clouds after iteration 6, when a complete view of the scene has been achieved. Moreover, Figures 7.(a - c) show the generated triangular mesh for each point cloud which will be used as a basis for reconstruction of new scene points. As demonstrated in provided results, there are some areas in the scene that cannot be successfully reconstructed due to occlusion problems or filtering of the associated triangles. In order to evaluate the computational efficiency of the proposed procedure, the computation time for each iteration was also measured. The proposed approach was

implemented on an Intel Xeon processor with 2.4 GHz, 32 GB of RAM, and nVidia NVS 315 graphics device.

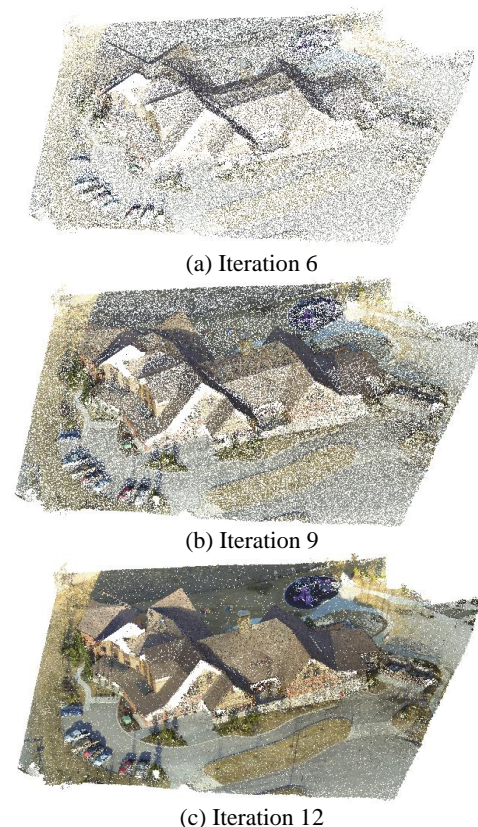


Figure 6. (a – c) Progressive dense reconstruction outcome

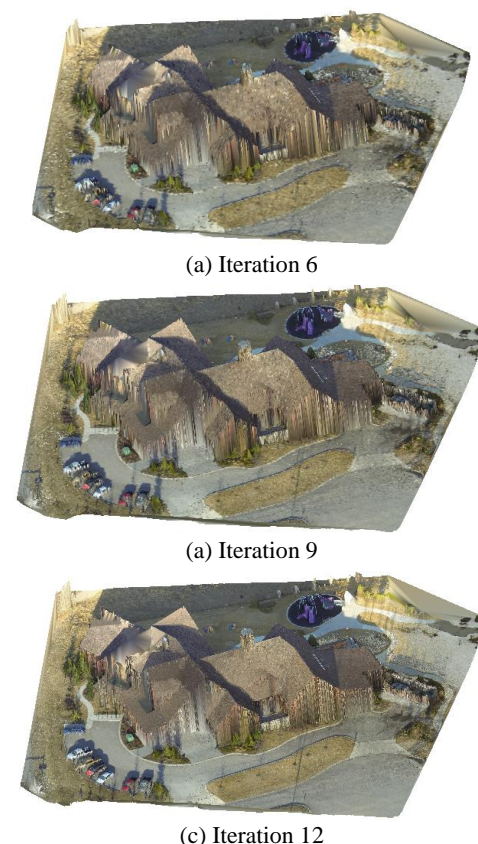


Figure 7. (a – c) Triangular mesh reconstruction from progressive dense reconstruction

Table 2 summarizes the processing time for delivering presented results and a dense point cloud generated using an available image-based dense matching approach (Moulon et al., 2017). In our progressive approach, the required processing time increases in each iteration until the majority of constructed triangles meet the expected accuracy. It will then decrease until the procedure is finalized. As it can be comprehended from Table 2, the required processing time to achieve the expected accuracy (5mm), in our approach, is 123s (almost 2min) which is a quarter of the required processing time to achieve a point cloud with the same accuracy from a sole image-based dense reconstruction technique. Figure 8 provides a visual comparison of the point clouds with expected accuracy which have been attained using the proposed progressive and the investigated image-based dense matching techniques. The inspection of the demonstrated results verifies that the proposed approach is able to provide a dense point cloud with the expected accuracy in much shorter time which is the interest of first responders and decision makers in emergency-response applications.

Dense reconstruction technique	Processing time
Progressive Technique (Iteration 6)	24s
Progressive Technique (Iteration 9)	38s
Progressive Technique (Iteration 12)	61s
Progressive Technique (Iteration 13)	11s
OpenMVG (image-based dense matching technique)	8 min

Table 2. The comparison between the required processing time in successive iterations of the proposed progressive technique and an available image-based dense reconstruction technique

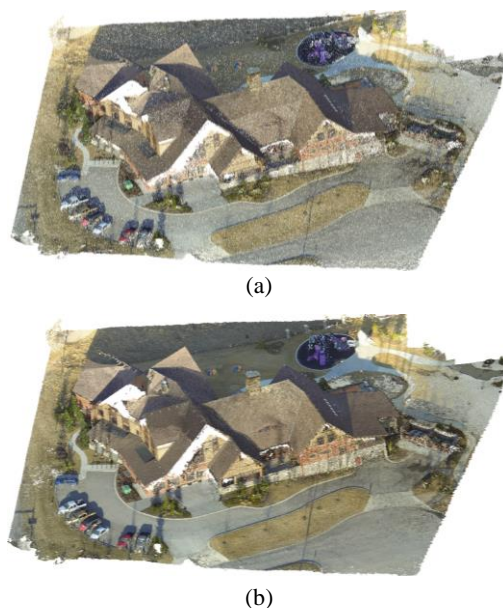


Figure 8: Visual comparison of the dense point clouds achieved using (a) the proposed progressive technique and (b) image-based technique

#### 4. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH WORK

To conclude, we presented a new approach for progressive dense 3D reconstruction of the scanned scenes by consecutive images collected onboard an unmanned mapping platforms which is required for disaster management and emergency response applications in this paper. In order to optimize the computational

efficiency of this technique, prior knowledge of the scene (in the form of previously acquired or computed sparse point cloud) is used as the basis of dense reconstruction procedure. The utilization of this information improves the computational efficiency of 3D reconstruction procedure by avoiding keypoint feature detection within consecutive images. Moreover, this approach optimizes and speeds up the feature matching procedure by bounding the search space for conjugate features within successive images. Finally, this technique employs efficient DLT for 3D triangulation to reconstruct 3D points with reasonable accuracy for emergency response applications. In order to verify the feasibility and computational efficiency of the proposed approach, an experiment using a set of consecutive images collected onboard a UAV platform and a low-density airborne laser scanning point cloud over the same scene were conducted and progressive results were provided. The comparative analysis of the required processing times and qualitative evaluation of the achieved point clouds using the proposed approach and image-based dense reconstruction technique proves feasibility, competency, and computational efficiency of this technique.

Future research work will be concentrated on the development of efficient sparse 3D scene reconstruction techniques to enable implementation of this technique without the need for laser scanning data. Furthermore, real-time context-aware 3D object reconstruction techniques will be investigated and optimized to be used in emergency response and disaster management applications.

#### REFERENCES

- Brunnett, G., Bieri, H., Farin, G., 2012. Geometric Modelling: Dagstuhl 1999 Computing Supplement 14. Springer Science & Business Media, NY, United States.
- Delaunay, B., 1934. Sur la sphère vide. A la mémoire de Georges Voronoï. *Bull. Académie Sci. URSS* 793–800.
- Desai, A., 2015. An Efficient Feature Descriptor and Its Real-Time Applications. Brigham Young University - Provo, Provo, UT, USA.
- Faugeras, O., Keriven, R., 2002. Variational principles, surface evolution, PDE's, level set methods and the stereo problem, in: *5th IEEE EMBS International Summer School on Biomedical Imaging*, 2002. Presented at the 5th IEEE EMBS International Summer School on Biomedical Imaging, 2002., pp. 83 pp.-.
- Faugeras, O.D., Bras-Mehlman, E.L., Boissonnat, J.D., 1990. Representing stereo data with the Delaunay triangulation. *Artif. Intell.* 44, 41–87.
- Gargallo, P., Sturm, P., 2005. Bayesian 3D Modeling from Images using Multiple Depth Maps. In: *Proceedings of IEEE Workshop on Motion and Video Computing, IEEE Computer Society*, pp. 885–891.
- Giordan, D., Manconi, A., Remondino, F., Nex, F., 2017. Use of unmanned aerial vehicles in monitoring application and management of natural hazards. *Geomat. Nat. Hazards Risk* 0, 1–4.
- Habib, A., Kim, E.M., Kim, C., 2007. New methodologies for true orthophoto generation. *Photogramm. Eng. Remote Sens.* 73, 25–36.
- Hernandez, C., Vogiatzis, G., Cipolla, R., 2007. Probabilistic visibility for multi-view stereo, in: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. In: *Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8.

- Hertzmann, A., Seitz, S.M., 2005. Example-based photometric stereo: shape reconstruction with general, varying BRDFs. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 1254–1264.
- Ladikos, S., 2011. Real-time multi-view 3D reconstruction for interventional environments. Technical University Munich.
- Manassis, A., Hilton, A., Palmer, P., McLauchlan, P., Shen, X., 2000. Reconstruction of scene models from sparse 3D structure, in: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000*, pp. 666–671.
- Moulon, P., Monasse, P., Marlet, R., 2017. openMVG: open Multiple View Geometry library. openMVG.
- Narayanan, P.J., Rander, P.W., Kanade, T., 1998. Constructing virtual worlds using dense stereo, in: *Proceedings of Sixth International Conference on Computer Vision*, pp. 3–10.
- Novak, K., 1992. Rectification of digital imagery. *Photogramm. Eng. Remote Sens.* 58, 339–344.
- Pons, J.-P., Keriven, R., Faugeras, O., 2007. Multi-View Stereo Reconstruction and Scene Flow Estimation with a Global Image-Based Matching Score. *Int. J. Comput. Vis.* 72, 179–193.
- Remondino, F., Spera, M.G., Nocerino, E., Menna, F., Nex, F., 2014. State of the art in high density image matching. *Photogramm. Rec.* 29, 144–166.
- S. Thenkabail, P., 2015. Remote Sensing of Water Resources, Disasters, and Urban Studies. CRC Press, FL, USA.
- Seitz, S.M., Dyer, C.R., 1999. Photorealistic Scene Reconstruction by Voxel Coloring. *Int J Comput Vis.* 35, 151–173.
- Shewchuk, J.R., 2012. Lecture notes on Delaunay mesh generation (Lecture Note). University of California at Berkeley, Berkeley, CA, USA.
- Szeliski, R., 1999. A multi-view approach to motion and stereo, in: *Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 163 Vol. 1.
- Van Aalst, M.K., 2006. The impacts of climate change on the risk of natural disasters. *Disasters* 30, 5–18.
- Wisetjindawat, W., Ito, H., Fujita, M., Eizo, H., 2014. Planning Disaster Relief Operations. in: *Proceedings of Soc. Behav. Sci., Eighth International Conference on City Logistics 17-19 June 2013*, Bali, Indonesia 125, pp. 412–421.
- Zlatanova, S., Li, J., 2008. Geospatial Information Technology for Emergency Response. CRC Press, FL, USA.