Digital Terrain Modeling Using AKAZE Features Derived from UAV-Acquired, Nadir and Oblique Images

H. Kim[a], H. Son[a], and C. Kim[a]

[a]Department of Architectural Engineering, Chung-Ang University, South Korea
E-mail: kkhs1127@cau.ac.kr, hjson0908@cau.ac.kr, changwan@cau.ac.kr

Abstract –

During construction, digital terrain models based on images from unmanned aerial vehicles (UAVs) are useful, as they rapidly provide data for objective volume calculations that can be used to monitor in-progress earthworks. As many curved objects are present on construction sites and as construction materials and equipment can provide partial obstructions, nadir images alone is not adequate to generate accurate digital terrain models during the earthwork phase, thus, it is necessary to acquire both nadir and oblique images. However, it is difficult to extract features from oblique images to determine accurate locations using the traditional method from photogrammetry software, the scale-invariant feature transform (SIFT) algorithm. This study proposes a method for generating accurate digital terrain models of construction sites based on the accelerated KAZE (AKAZE) algorithm using a combination of nadir and oblique UAV images. The proposed method consists of the following steps: 1) feature extraction and matching based on the AKAZE algorithm; 2) three-dimensional (3D) point cloud generation based on the results of the feature matching; and 3) digital terrain model generation based on the resulting 3D point cloud and on the GPS information for the corresponding locations of each ground control point (GCP). Validation of the proposed method involves 100 oblique UAV images of the actual construction site's in-progress earthworks. The experimental results indicate that the AKAZE algorithm can be applied to generate an accurate digital terrain model by extracting and matching features to ensure the accuracy of the edges and corners. Based on these results, the proposed method can be expected to provide accurate volume calculations from a generated digital terrain model, which can enable monitoring of the earthwork phase.

Keywords –

AKAZE; Digital Terrain Model; Earthwork; Oblique Images; Photogrammetry; UAV

1 Introduction

3D digital terrain models are used in fields such as traffic, geology, and archeology, but they are especially common in the construction industry, where they can rapidly provide data for objective volume calculations that can be used to monitor in-progress earthworks. The role of the digital terrain model is regarded as increasing because of the need to periodically monitor the continual terrain changes of the construction site. However, for such digital terrain models to be effectively used in construction project, an approach is needed that can both model the terrain rapidly to provide periodic updates and accurately represent the terrain.

In the construction industry, terrestrial laser-scanning surveys are widely used in digital terrain modeling because of their high accuracy (e.g., [1–3]). However, these surveys' data acquisition is not efficient enough to provide periodic updates because, for each area of the construction site, the laser scanner needs to be moved and installed multiple times before the data can be acquired; this is time consuming and requires at least two people. These surveys also have the disadvantage of being difficult to use on terrain with severe drop-offs and/or inclines. In recent years, digital terrain modeling through the use of photogrammetry based on images from UAVs has gained popularity due to the evolution of UAV technology, which is now possible to improve stability and to allow longer flight time. This approach allows for the fast and easy acquisition of images for a large area, including for terrain with severe variations.

Several researchers have attempted to generate digital terrain models for construction sites using UAV nadir images (e.g., [4,5]) or a combination of UAV and ground images [6]. Previous studies have shown that digital terrain modeling using UAV images has the potential to rapidly generate digital terrain for a large area. However, image-acquisition or -processing methods from other industries are inadequate to be used in construction site without considering the unique characteristics of construction-site terrain [4–6].

Construction sites often have severe terrain variations, and various construction materials and equipment can partially obstruct terrain in the images. In this environment, using nadir and oblique images together (rather than nadir type alone) increases the vertical accuracy of a digital terrain model [7]. Several UAV manufactures (e.g., Leica Geosystems, Pictometry International, and Microsoft) have developed camera sensors that can acquire both nadir and oblique images from four directions, thus solving the obstruction problem. Accordingly, several researchers (e.g., [8–10]) have attempted to generate digital terrain models using nadir and oblique images in combination. The most commonly used feature-extraction and -matching algorithm in photogrammetry software—SIFT algorithm [12], which is employed in Agisoft's PhotoScan and Pix4D's Pix4Dmapper—is suitable for matching parallel nadir images. This is because the SIFT algorithm performs well when matching two images that have been acquired at different distances [12]; in this scenario, it can extract the features of the edges and corners. On the other hand, the SIFT algorithm has a limited ability to accurately extract features when matching two images that have been acquired from different angles [11]. It is still necessary to consider which feature-extraction and -matching method—when considering construction sites' terrain characteristics—will generate the most accurate digital terrain model from nadir and oblique images.

The aim of this study is to propose a method for using nadir and oblique UAV images to generate accurate digital terrain models of construction sites based on the AKAZE algorithm. This paper is organized as follows. Section 2 reviews the studies on digital terrain modeling based on UAV images in the construction industry. Section 3 introduces the proposed method, and Section 4 provides analysis of the experimental results. Finally, Section 5 describes the conclusions and includes suggestions for future research.

2 Literature Review

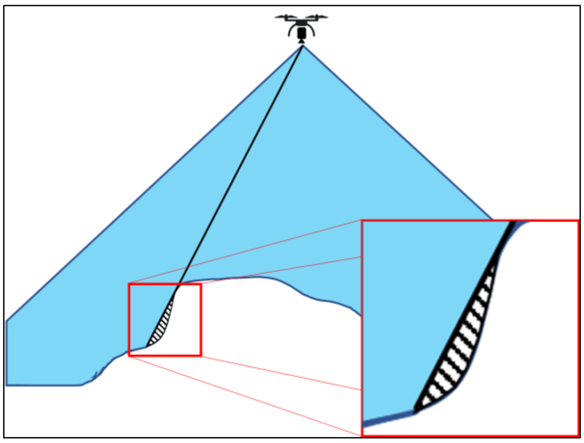
Several researchers (e.g., [4–6]) have attempted to generate digital terrain models for construction sites based on UAV-based nadir image acquisition. Sibert and Teizer [4] and Hugenholtz et al. [5] used UAV flight-plan software provided by manufacturers (e.g., DJI and Intel) in addition to commercial photogrammetry software. The UAV flight-plan software developed by the UAV industry provides the ability to acquire images while maintaining a constant overlap ratio at a certain height. The resulting images can be used to generate digital terrain models using photogrammetry software. Sibert and Teizer [4] used the UAV flight-plan software to collected nadir images from a height of 30 m and then applied commercial photogrammetry software (Agisoft's PhotoScan) to generate a digital terrain model of a construction site. Hugenholtz et al. [5] used AutoGrid's Aeryon Labs flight-plan software to acquire nadir images from a height of 100 m and then applied commercial photogrammetry software (MosaicMill's EnsoMOSAIC) to generate a digital terrain model. In general, commercial photogrammetry software generates digital terrain models based on the SIFT feature-extraction algorithm [12]. Bügler et al. [6] acquired both aerial and ground images and used them to generate a digital terrain model using open-source VisualSFM software [13], which is also based on the SIFT algorithm. The resulting digital terrain model and the activity status of the construction equipment based on video analysis provided information on equipment productivity. In summary, image-acquisition and -processing methods from other industries are inadequate to be used in construction without first considering the characteristics of construction sites' terrain [4–6]). Thus, it is still necessary to consider—based on construction sites' terrain—which method of feature extraction is most appropriate for generating accurate digital terrain models from UAV images.

3 Methodology

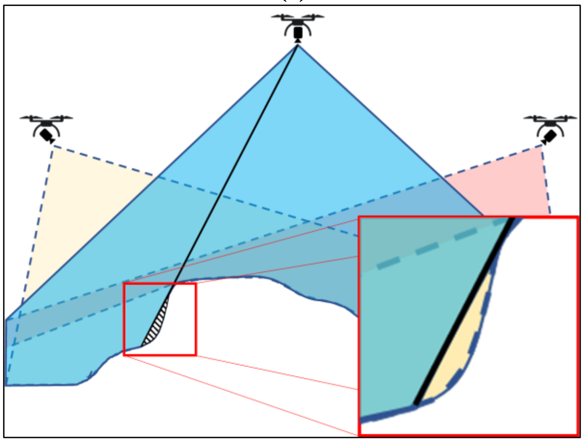
3.1 Image Acquisition

In this study, the images were acquired using DJI's UAV Matrice 100 and Zenmuse Z3 camera, with the point-of-interest function from DJI's Ground Station Pro software used to plan the image-acquisition flight. The flight path had a 90% overlap ratio, using angles of −75°, −60°, −45°, and −30°. This resulted in four flight paths. Each flight path followed a hemispherical orbit with a radius of 30 m around the target terrain; in all, 100 images of 4,000 × 3,000 resolution were acquired.

Figure 1 illustrates the difference between regions for rough terrain when using the nadir images alone (Figure 1[a]) and when using nadir and oblique images simultaneously (Figure 1[b]). As shown in Figure 1(a), when using only nadir UAV images for digital terrain modeling, obstructions can be occurred in areas of severe drop-offs and/or inclines [4,5]. On the other hand, as shown in Figure 1(b), when both nadir and oblique images are used, obstructed areas can be revealed.



(a)



(b)

Figure 1. Image-acquisition approaches: (a) nadir images alone and (b) both nadir and oblique images

3.2 Image Processing

Nadir and oblique images are based on different distances to the same object, as shown in Figure 2. For this reason, the same object appears on a different scale in these images. Under these conditions, determining accurate locations in all images is important because they affect the accuracy of the resulting digital terrain model. The SIFT algorithm that is often used in commercial photogrammetry software applies the Laplacian of Gaussian method, which uses differences between images (after blurring with a Gaussian filter) to extract the same features of an object that is seen from various scales. Equal-sized masks are used on the images of multiple scales to extract the features that are relevant to determining the edges and corners; this process helps identify the corresponding features for the same objects at various distances. However, the Gaussian filter can result in blurred edges and corners, which makes it difficult to extract accurate locations of features [14].

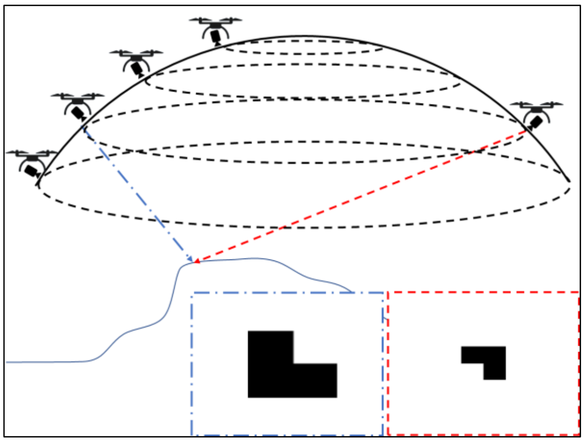


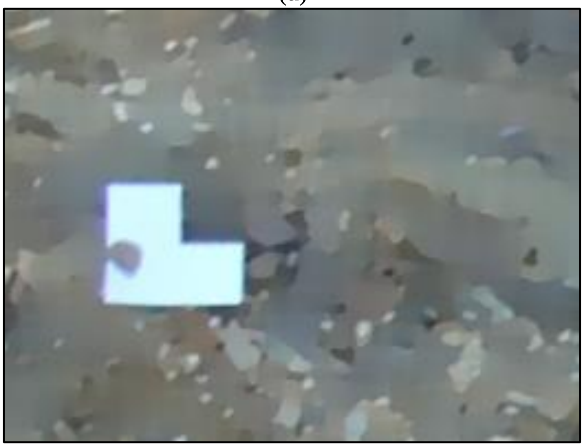
Figure 2. Changes in the scale of the same object in images acquired from different angles

To handle the problem of constructing nonlinear scale space to find the corresponding features for the same object from images obtained at multiple distances, the KAZE algorithm using a nonlinear diffusion filter was proposed by Alcantarilla et al. [14]. The nonlinear diffusion filter is advantageous because the edge information is preserved rather than blurred as it is processed in the Gaussian filter (e.g., [14–16]). Figure 3 shows an image of the GCP placed on the terrain, and the GCP's edges are distinctive in the image.

Figure 3. An example of a UAV image obtained at the actual earthwork site

Figure 4 shows the results of applying the Gaussian filter and the nonlinear diffusion filter, respectively, to Figure 3. Figure 4(a) shows the result of applying the Gaussian filter to construct the scale space in the SIFT algorithm, and this process blurred edge information from the original image. Figure 4(b) shows an example of applying the nonlinear diffusion filter. During the construction of the scale space, the distinctive edges in the image were retained without sacrificing the object's edges. Unlike the SIFT algorithm, this property leads to the identification of exact position of objects' edges and corners at various scales using such scale spaces [14–16]. This study employed the AKAZE [17] algorithm, which uses the same nonlinear diffusion filter as the KAZE algorithm but improves the processing speed.

(a)



(b)

Figure 4. Example of applying (a) Gaussian filter and (b) nonlinear diffusion filter

3.3 3D Terrain Model Generation

Once the features have been extracted from a pair of images, the features from the same locations are matched using the random sample consensus algorithm and transformed into a 3D point clouds. When the position of the extracted feature is more accurate, so too is its corresponding 3D point. A 3D point cloud is then generated based on the extracted feature-matching results of 100 images. In this process, the SIFT and AKAZE algorithms are applied using OpenCV3.4.0 connected to OpenMVG [18]—open-source photogrammetry software—by applying defualt parameters. Then, the digital terrain model is generated using OpenMVS, which is an open-source method for generating dense point clouds and mesh models from 3D point cloud recalibrated through the bundle-adjustment process; this position of the resulting 3D point cloud is based on the GPS information of the corresponding locations of each GCP.

4 Experiments

This section provides a description of the results for each step of the experiment. The experiment was performed at an actual highway construction site with in-progress earthworks in Asan-si, Chungcheongnam-do, South Korea. The target area for generating the digital terrain model included bumpy terrain. The UAV acquired images of the target area using a hemispherical flight path.

4.1 Performance Measures

To evaluate the accuracy of the generated 3D point cloud, the values of X, Y, and Z at each check point (CP) derived from the dense point cloud are compared with GPS information measured at the site. The dense point cloud generated by registering GPS information corresponding to each GCP location contains CPs to be used for evaluation, which are not used in georeferencing, and the values of X, Y, and Z in each CP can be derived using CloudCompare. The measures used for evaluation are four values of root mean square error (RMSE), which represent the error for each axis and the overall error in the 3D space. The equations for the four measures are as follows.

$$\mathrm{RMSE_x} = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(X_{pi}-X_{qi}\right)^2\right]}{n}} \quad (1)$$

$$\mathrm{RMSE_y} = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(Y_{pi}-Y_{qi}\right)^2\right]}{n}} \quad (2)$$

$$\mathrm{RMSE}_z = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(Z_{pi}-Z_{qi}\right)^2\right]}{n}} \quad (3)$$

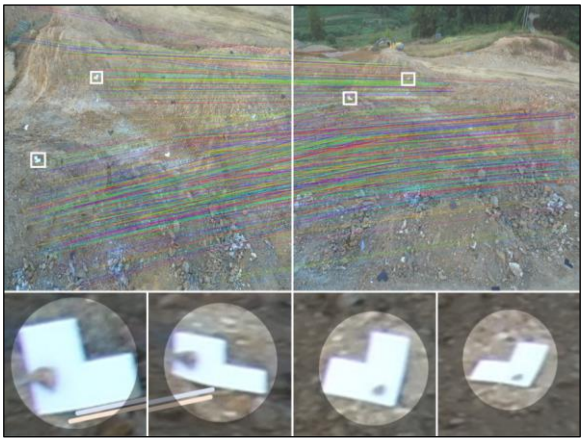
$$\mathrm{RMSE_{xyz}} = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(X_{pi}-X_{qi}\right)^2+\left(Y_{pi}-Y_{qi}\right)^2+\left(Z_{pi}-Z_{qi}\right)^2\right]}{n}} \quad (4)$$

where n is the number of CPs and X_{pi}, Y_{pi}, and Z_{pi} are the X, Y, and Z coordinates derived from the 3D point cloud for the i_{th} CP. X_{qi}, Y_{qi}, and Z_{qi} are X, Y, and Z coordinates of GPS information for the i_{th} CP measured at the site. A smaller value for each measure means the 3D point cloud is generated closer to the actual site.

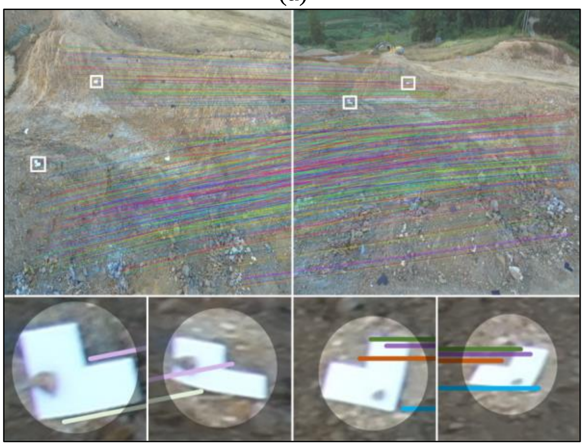
4.2 Results and Discussion

The resulting nadir and oblique images produced a digital terrain model through feature extraction and matching; the resulting point cloud was generated based on GPS information for each GCP. Figure 5 compares the results of the matching features for the SIFT and AKAZE algorithms based on two oblique images acquired from the construction site. The point at the end of each line represents the location of the feature.

The feature-matching result of the AKAZE algorithm in Figure 5(b) was more accurate than the SIFT algorithm in Figure 5(a) in terms of the edges and corners. Figure 6(a) shows a 3D point cloud that was generated using the AKAZE algorithm. The extracted features have two-dimensional location values. Through triangulation based on the camera locations, a single point with a 3D location value can be computed. Finding accurate locations for the edges and corners in each image can reduce errors in the point cloud, thus enabling better matching between pairs of images and improving the accuracy of the resulting digital terrain model. Because the resulting point cloud is generated only using feature matching without GPS information, it is necessary to register GPS information to improve accuracy. Figure 6(b) shows a dense point cloud with GPS information. In Figure 6(b), each pair of a white point and a tag represents where the CP is located.

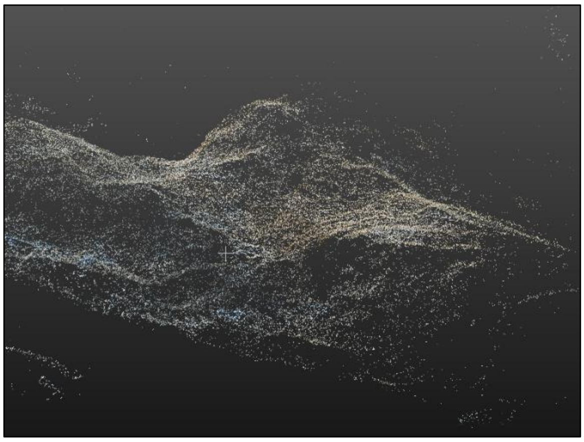


(a)

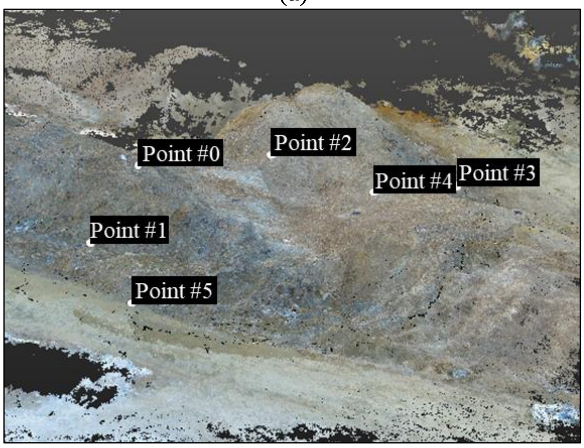


(b)

Figure 5. Comparison of the feature-extraction and -matching results for (a) the SIFT algorithm and (b) the AKAZE algorithm



(a)



(b)

Figure 6. (a) 3D point cloud generated from images using the AKAZE algorithm and (b) dense point cloud based on feature points

Table 1 summarizes the comparison of errors of 3D point cloud generated by the two algorithms. The 3D point cloud generated by applying the AKAZE algorithm showed smaller error values than that generated by applying the SIFT algorithm for all four measures. In other words, the results indicate that a more accurate 3D point cloud was generated when the AKAZE algorithm was applied.

Table 1. Comparison of errors of 3D point cloud generated by various algorithms

Measures (m)	SIFT	AKAZE
$\mathrm{RMSE_x}$	0.044	**0.012**
$\mathrm{RMSE_y}$	0.116	**0.108**
RMSE_z	0.053	**0.038**
$\mathrm{RMSE_{xyz}}$	0.053	**0.032**

Figure 7 shows a digital terrain model that is generated from the same dense point cloud from Figure 6. The digital terrain model is then generated, with open-source code (OpenMVS) as the basis for the dense point cloud and the mesh model. The resulting model is based on accurate feature locations, which were extracted using the AKAZE algorithm and which thus are expected to have higher accuracy than those produced using the SIFT algorithm.

Figure 7. Digital terrain model generated using the proposed method

5 Conclusion

In this study, the AKAZE algorithm was employed to generate an accurate digital terrain model from nadir and oblique UAV images. The results indicate that the features extracted using the AKAZE algorithm are more accurate than those extracted using the SIFT algorithm. This difference implies that the proposed method can generate more accurate digital terrain models than can be generated through conventional method. The model generated by the proposed method can be used to provide quantitative measurements, which can be applied for more accurate construction-site volume calculations. Future research will apply the algorithms based on stable affine invariant features and compare the 3D reconstruction performance in terms of accuracy.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF- 2018R1D1A1B07049846).

References

[1] Jaselskis, E. J., Gao, Z., & Walters, R. C. Improving transportation projects using laser scanning. *Journal of Construction Engineering and Management* 131(3):377–384, 2005.

[2] Du, J. C., & Teng, H. C. 3D laser scanning and GPS technology for landslide earthwork volume estimation. *Automation in Construction* 16(5):657–663, 2007.

[3] Slattery, K. T., Slattery, D. K., & Peterson, J. P. Road construction earthwork volume calculation using Three-dimensional Laser Scanning. *Journal of Surveying Engineering* 138(2):96–99, 2011.

[4] Siebert, S., & Teizer, J. Mobile 3D mapping for surveying earthwork projects using an unmanned aerial vehicle (UAV) system. *Automation in Construction* 41:1–14, 2014.

[5] Hugenholtz, C. H., Walker, J., Brown, O., & Myshak, S. Earthwork volumetrics with an unmanned aerial vehicle and softcopy photogrammetry. *Journal of Surveying Engineering* 141(1):06014003, 2014.

[6] Bügler, M., Borrmann, A., Ogunmakin, G., Vela, P. A., & Teizer, J. Fusion of photogrammetry and video analysis for productivity assessment of earthwork processes. *Computer-Aided Civil and Infrastructure Engineering* 32(2):107–123, 2017.

[7] James, M. R., & Robson, S. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. Earth Surface Processes and Landforms 39(10):1413–1420, 2014.

[8] Fernández-Lozano, J., & Gutiérrez-Alonso, G. Improving archaeological prospection using localized UAVs assisted photogrammetry: An example from the roman gold district of the Eria river valley (NW Spain). *Journal of Archaeological Science Reports* 5:509–520, 2016.

[9] Agueera-Vega, F., Carvajal-Ramirez, F., Martínez-Carricondo, P., López, J. S. H., Mesas-Carrascosa, F. J., García-Ferrer, A., & Pérez-Porras, F. J. Reconstruction of extreme topography from UAV structure from motion photogrammetry. *Measurement* 121:127–138, 2018.

[10] Rusnák, M., Sladek, J., Kidová, A., & Lehotský, M. Template for high-resolution river landscape mapping using UAV technology. *Measurement* 115:139–151, 2018.

[11] Yu, G., & Morel, J. M. ASIFT: An algorithm for fully affine invariant comparison. *Image Processing on Line* 1:11–38, 2011.

[12] Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2):91–110, 2004.

[13] Wu, C. VisualSFM: A visual structure from motion system. Online: http://ccwu.me/vsfm/, Accessed: 27/01/2019.

[14] Alcantarilla, P. F., Bartoli, A., & Davison, A. J. KAZE features. In *Proceedings of the European Conference on Computer Vision*, pages 214–227, Berlin, Heidelberg, 2012.

[15] Fortun, D., Bouthemy, P., & Kervrann, C. Optical flow modeling and computation: A survey. *Computer Vision and Image Understanding* 134:1–21, 2015.

[16] Gastal, E. S., & Oliveira, M. M. Oliveira. Domain transform for edge-aware image and video processing. *ACM Transactions on Graphics* 30(4), 2011.

[17] Alcantarilla P. F., Nuevo J., and Bartoli A. Fast explicit diffusion for accelerated features in nonlinear scale spaces, In *Proceedings of the British Machine Vision Conference*, 2013.

[18] Moulon, P., Monasse, P., Perrot, R., & Marlet, R. OpenMVG. Online: https://openmvg.readthedocs.io/en/latest/#, Accessed: 27/01/2019.