

Multispectral's Three-Dimensional Model Based on SIFT Feature Extraction

Shaharom, M. F. M.,¹ Abd Mukti, S. N.,² Raja Maharjan, G.³ and Tahar, K. N.^{1*}

¹School of Geomatics Science and Natural Resources, College of Built Environment, Universiti Teknologi MARA, 40450 Shah Alam, Selangor Darul Ehsan, Malaysia, E-mail: khairul0127@uitm.edu.my

²Dewan Bandaraya Kuala Lumpur (DBKL), Department of Infrastructure Planning, Level 10 North, Menara DBKL 2, Jalan Raja Laut, 50350 Kuala Lumpur, Malaysia

³Central Department of Geography, Tribhuvan University, P.O.Box 8613, Kirtipur, Kathmandu, Nepal

*Corresponding Author

DOI: <https://doi.org/10.52939/ijg.v19i5.2649>

Abstract

Recently, multispectral images can be captured not only from satellite sensors but also from cameras. Hence, using the photogrammetric approach, multispectral images can be manipulated to generate a three-dimensional model. The main issues regarding multispectral images were the low visibilities of the image features. Moreover, the tie point extractions on multispectral images were still in doubt. Hence, this paper examines the capabilities of the SIFT algorithm to extract feature points from multispectral images and generate the point cloud from the extracted feature points. This study chose a pothole as the subject of this research. The red, red edge, green, and near-infrared bands from the Parrot Sequoia camera were used to generate the pothole model. All captured images were processed using structure-from-motion (SfM) with Multi-View Stereo (MVS) technique. This study records the feature points extraction result and analysis of the pothole model and discuss it in this paper.

Keywords: 3D Model, Band, Image, Processing

1. Introduction

Digital photogrammetry technique changes rapidly as the ability to offer better three-dimensional products is successfully developed. Photogrammetry can be categorised into two major methods of data capture. First, aerial photogrammetry survey uses airborne platforms, such as Unmanned Aerial Vehicle (UAV), helicopter and fixed-wing plane. Second, terrestrial photogrammetry, where the data is captured from the ground. Any data captured within a distance less than 300 meters from the surveyed object is Closed Range Photogrammetry (CRP). The photogrammetry technique produces Digital Surface Model (DSM), orthophoto, three-dimensional model, and point clouds. Those products benefit many fields, such as survey and mapping, medical, military, and automotive industries. Compared to other survey techniques, photogrammetric survey delivers the end product much faster (in terms of large cover area) and is very cost effective.

Additionally, the data covers anything that is displayed in the image. The photogrammetry survey accuracy was proven by millimetre to centimetre

levels [1] and [2] when RGB images were applied. As a result, many researchers from different backgrounds benefit from this technique for different purposes [3]. However, photogrammetry accuracy depends on the image resolution, GCP selection and accuracy, determination of the flying height, and image overlapping during data acquisition [4] and [5]. Nowadays, the use of multispectral images is offered by satellite sensors and drones. It opens a new opportunity to explore the generation of a three-dimensional model using multispectral images.

2. SIFT Image Matching

A feature-based technique consists of three steps: feature extraction, feature description, and feature matching. The main issue for Multispectral (MS) is missed registration due to non-linear intensity spectral response [6]. Scale Invariant Feature Transform (SIFT) is one of the feature-based techniques for image matching. One of this detector's specialties was invariant to image rotations, translations, scaling, and brightness of images [7].

The previous study was used MATLAB software and tested on three satellite images, namely World-View2, Spot6, and TerraSAR-X. Image matching from satellite images was challenging because of the image characteristic itself. Satellite images were taken on different dates and sensor sources. At the same time, the satellite was rotated and translated, which led to changing scale and image brightness. The use of SIFT algorithm solves these issues. The transformation of these images by SIFT was smoothed using a Gaussian filter and sampling to generate a smoother version than the original image [8]. Figure 1 illustrates the SIFT techniques.

However, the limitation of SIFT is the operator extracts round feature points better than corner or edges point [9]. As a result, SIFT is more suitable for obtaining images in forestry, shrubbery, and grassland. Another drawback is SIFT requires large complex computations, resulting in lower image-matching performance [10].

3. Materials and Method

3.1 Data Acquisition

This study used six camera stations to acquire data for each band and took all images simultaneously. A Parrot Sequoia camera was attached to a UAV drone to capture multispectral images. Then, set the altitude for image capturing to approximately two meters from ground level (Figure 2). The study area was a bus parking lot at UiTM Shah Alam, Selangor. Because the image taken was a pothole in the middle of a paved road, the selected location far from busy roads and human activity so that the captured data were free from cars or blocked views. All captured data were processed using structure-from-motion (SfM) concept, where the images were taken using multi-view stereoscopic (MVS) technique. There was a total of 30 Multispectral images, which included the RGB band. This study examined and processed all images using Visual SFM software to obtain feature points and its descriptions using SIFT.

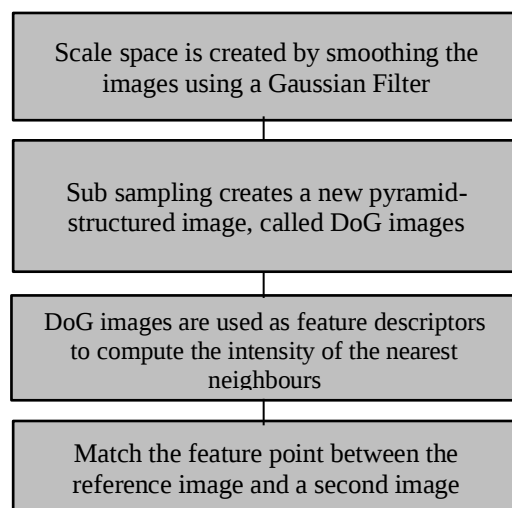


Figure 1: Procedure in SIFT image matching

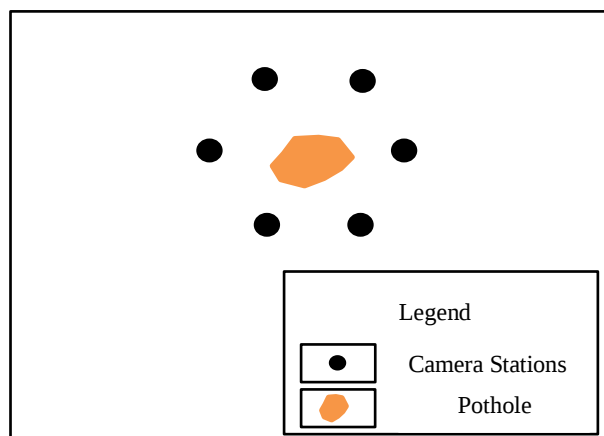


Figure 2: Plan view of camera stations and pothole position

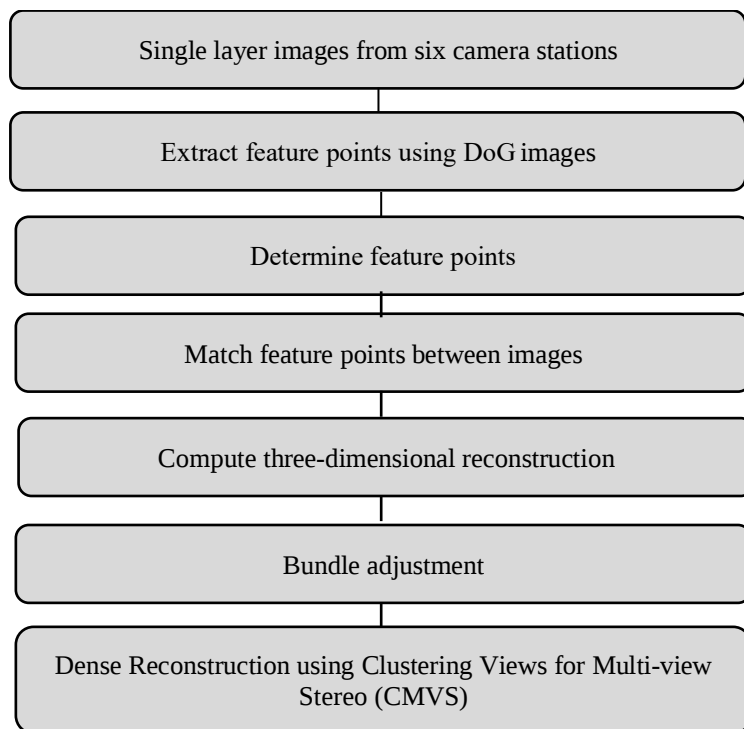


Figure 3: Procedure of point cloud determination in visual SfM software

3.2 Data Processing

This study obtained the feature point and descriptors of single band images by processing them using the Visual SfM software (SIFT algorithm). All feature point candidates were matched between the images to obtain the position of the cameras when the pothole image was captured. Figure 3 shows the procedures for point cloud generations in Visual SfM software. Feature point extraction using Dog images consists of two stages. First, the point containing low contrast characteristics must be excluded. Second, all remaining feature points were processed again, but this time, by discriminating weak points at the edges. The descriptions for the feature points were processed after the above procedure. A three-dimensional reconstruction was chosen to generate point clouds using the Visual SfM software. Then, the final dense point clouds were produced using clustering views for multi-view stereo (CMVS) tools. The final point clouds were in “.ply ” format, which was generated using the MeshLab software. All potholes model from four bands were converted into the raster format of the Digital Elevation Model (DEM) using GIS software for area and volume computation later on.

3.3 Data Validation

The validation process is divided into two categories: validation of the pothole area and validation of the

pothole volume. The pothole area samples were calculated using an area of irregular estimation by a square grid. Samples of pothole volumes were validated using sand and biscuit containers. Figure 4 shows the square grid technique this study used to calculate the pothole area. To calculate the Pothole’s area, the grid was divided into four fractions:

- a) Full grid = 1
- b) Half grid = $\frac{1}{2}$
- c) Three quarter grids = $\frac{3}{4}$
- d) Quarter grid = $\frac{1}{4}$

This study obtained the total pothole area by counting the sum of the grid categories, as shown in Table 1 (grid a). Each fraction was calculated in the metric unit multiplied by the number of grid fractions to obtain the total pothole area. The total pothole area was calculated using the square grid method shown in Table 1 (grid b). The calculated pothole volume as shown in Table 1 (grid c). The full volume of the biscuit container was 5729.76 cm^3 and the next volume was taken by measuring the height of the biscuit container to obtain the remaining value. The quantity of the biscuit container volume was 11459.52 cm^3 . The remaining sand volume was 1245.6 cm^3 . The total pothole volume was 12705.12 cm^3 .

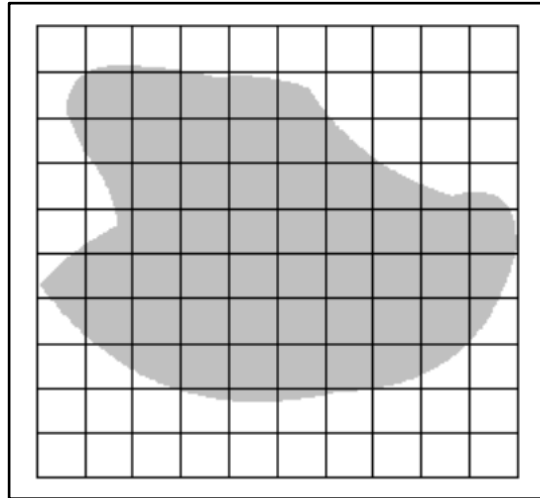


Figure 4: Area of irregular shape by square grid method

Table 1: Categories of grid size and its area; total pothole area using the square grid method and pothole volume

Grid		Mathematical Representations	Size of area (cm^2)
a	Full grid	1	217.56
	Three quarter grids	$\frac{3}{4}$	163.17
	Half grid	$\frac{1}{2}$	108.78
	Quarter grid	$\frac{1}{4}$	54.39
b	Grid	Quantity	Size of area (cm^2)
	Full grid	11	2393.16
	Three quarter grids	5	815.85
	Half grid	3	326.34
	Quarter grid	2	108.78
	Total pothole area		3644.13
c	Biscuit's Container Dimension	Quantity	Volume (cm^3)
	17.3 cm x 14.4 cm x 23 cm = 5729.76 cm^3	2	11459.52
	17.3 cm x 14.4 cm x 5 cm = 1245.6 cm^3	1	1245.60
	Total Volume		12705.12

4. Result and Analysis

4.1 Feature Points Extractions Using SIFT Algorithm

This study successfully extracted the SIFT feature point using Visual SFM. Using this software, the same methodology was used to process all images from the red, green, red edge, and near-infrared bands as shows the results in Table 2. Six images were processed from the red band spectral to extract the feature points. The most points were extracted from image IMG_220923_020652_0000_RED (13629 points). The least feature points were extracted from image IMG_220923_020553_0000_RED (8280 points). This study successfully extracted tens of thousands feature points from three out of six images.

Overall, red band images gave the best result, where this study extracted more than 8000 feature points for each image. Six green band images were processed using SIFT to obtain the feature points. More feature points were extracted from image IMG_220923_021034_0000_GRE (9094 points). The least feature points were obtained from image IMG_220923_020652_0000_GRE (2673 points). The least feature points were extracted from green band images compared to the other bands. A good amount of feature points were extracted from red edge images, i.e., an overall of more than 7000 points for each image. More feature points were extracted from image IMG_220923_020553_0000_REG (10670 points).

Table 2: Feature point extraction using VisualSFM (red band image), (green band image), (red edge band image) and (near infrared band image)

No.	Image ID	Number of Feature point extracted
a	IMG_220923_020652_0000_RED	13629
	IMG_220923_020811_0000_RED	13524
	IMG_220923_020905_0000_RED	12592
	IMG_220923_020553_0000_RED	8280
	IMG_220923_021034_0000_RED	8618
	IMG_220923_020511_0000_RED	8646
b	IMG_220923_020811_0000_GRE	3755
	IMG_220923_020905_0000_GRE	4468
	IMG_220923_020652_0000_GRE	2673
	IMG_220923_020553_0000_GRE	5339
	IMG_220923_020511_0000_GRE	5455
	IMG_220923_021034_0000_GRE	9094
c	IMG_220923_020511_0000_REG	11038
	IMG_220923_020811_0000_REG	9039
	IMG_220923_021034_0000_REG	11070
	IMG_220923_020553_0000_REG	10670
	IMG_220923_020652_0000_REG	7776
	IMG_220923_020905_0000_REG	8885
d	IMG_220923_020553_0000_NIR	10927
	IMG_220923_021034_0000_NIR	11567
	IMG_220923_020511_0000_NIR	11199
	IMG_220923_020905_0000_NIR	7971
	IMG_220923_020811_0000_NIR	8753
	IMG_220923_020652_0000_NIR	8752

The least number of feature points were extracted from image IMG_220923_020652_0000_REG (7776 points). Six images were processed from the near-infrared band. The most feature points were extracted from image IMG_220923_020553_0000_NIR (10927 points). Only 7971 points were extracted from image IMG_220923_020905_0000_NIR. This study successfully extracted tens of thousands feature points from three out of six images. Overall, NIR band images gave a good result, where more than seven thousand feature points were extracted for each image.

4.2 Pothole's Three-dimensional Points cloud

This study produced a three-dimensional pothole model in point cloud form, which involved two stages: two preliminary point cloud formations (point cloud) and dense point cloud, produced from the multi-view stereoscopic (MVS) technique. This study was able to produce a pothole point cloud from all four band images (Figure 5). Red band and red edge band point clouds produced 27,013 and 35,269 of vertices respectively. The densest point cloud was successfully produced using green band images.

There were a total of 41,412 vertices, which made the pothole model look smooth and obvious. On the other hand, the NIR band failed to give a good result, generating only 1730 vertices. As a result, the pothole model produced from NIR images cannot be identified. Further quantitative analyses were performed using GIS software, where this study calculated the area and volume of the potholes and compared them with conventional results.

4.3 Pothole Area and Volume

The pothole's area and volume were determined from the four bands from its DEM using GIS software (Figure 6). Table 3 shows the area and volume of each pothole model. The benchmark of this study was the conventional measurement of pothole area and volume. Our results show that the red edge band produced the closest area value compared to the conventional measurement. On the other hand, the green band pothole's volume was approximately the same as the conventional measurement, a difference of only 8.12 cm³. The highest area and volume difference was for the near-infrared band because its point cloud density was very low, causing difficulty in extracting the pothole shape in the GIS software.

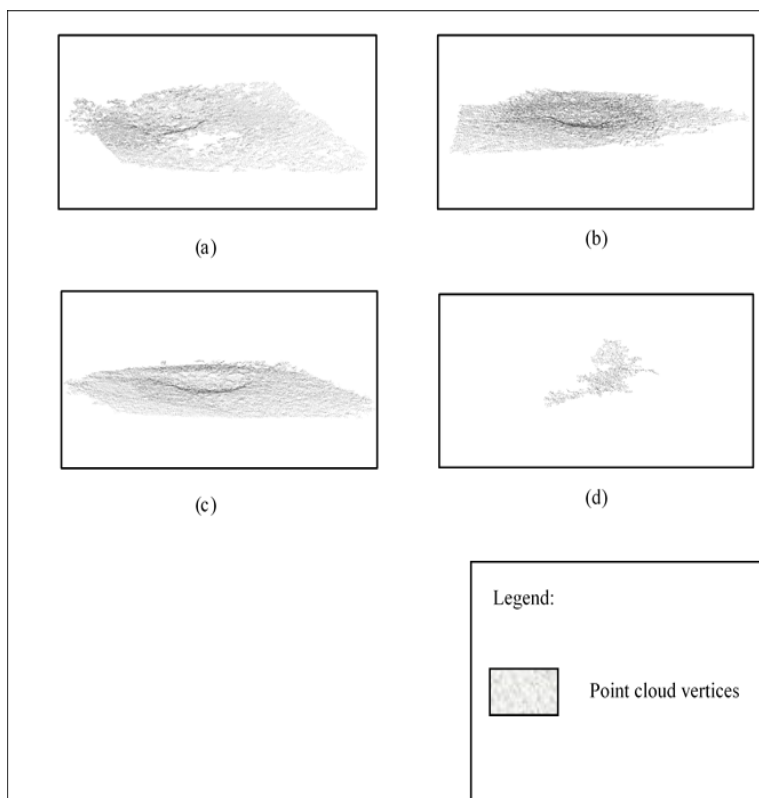


Figure 5: Pothole point cloud from four bands; (a) Red band, (b) Green band, (c) Red edge band, (d) Near infrared band

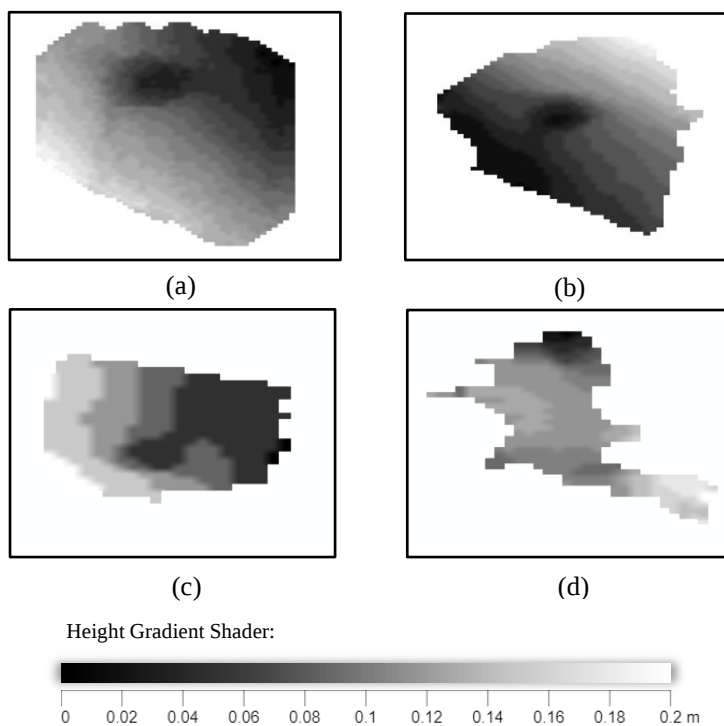


Figure 6: Digital Elevation Model from Four Single bands; (a) Red, (b) Green, (c) Red Edge, (d) Near Infrared

Table 3: Area and volume of the potholes (red, green, red edge and near infrared band)

No.	Bands	Area (cm^2)	Diff from the conventional method (cm^2)	Volume (cm^3)	Diff from the conventional method (cm^3)
1	Red	3137.77	506.36	14580	1874.88
2	Green	3200.00	444.13	12697	8.12
3	Red edge	3625.00	19.13	8956	3749.12
4	Near-infrared	705.00	2939.13	36335	23629.88

5. Conclusion

Multispectral images produce low feature image visibilities. SIFT image matching depends on the visibility of the features because it extracts the features from the image edges. The lower the visibility, the harder the extraction of the feature point candidates. Moreover, points matching to the feature point candidate between images also play a significant role because the matched feature points between images were used to compute the camera pose. The camera position value is used to reconstruct the three-dimensional model. Using a suitable model fitting analysis, such as Nearest Neighbor Distance Ratio (NNDR) and Random Sample Consensus (RANSAC), can eliminate the outliers during feature point matching. In conclusion, SIFT successfully extracted the feature point candidates from four bands' images. However, the point cloud generated for the near-infrared band produced very minimal vertices, leading to a large difference in area and volume values compared to the results of the conventional method. On the other hand, the remaining three bands, red, red edge and green, successfully produced a high-density point cloud. The differences in the pothole's area and volume values for these three bands were small compared to the conventional measurement values. For our future work, this study plan to use other algorithms, such as SURF and established photogrammetry software, to analyse and compare the accuracy of the produced potholes.

Acknowledgements

Ministry of Higher Education (MOHE) are greatly acknowledged for providing the Fundamental Research Grant Scheme, Grant No. FRGS/1/2021/WAB07/UITM/02/2 and GPK fund (Grant No. 600-RMC/GPK 5/3 (223/2020)), College of Built Environment, Universiti Teknologi MARA (UiTM) and Research Management Centre (RMC) to enable this research to be carried out. The authors would also like to thank the people who were directly or indirectly involved in this research.

References

- [1] Lim, C. H., Zhang, L. and Amaludin, A. E., (2021). Topographic Survey and Modelling Using Photogrammetry: A Comparison against Electronic Distance Measurement (EDM) Method. *ASM Sci. J.*, Vol. 16(3), 1-9. <https://doi.org/10.32802/asmscj.2021.720>.
- [2] Elkhachy, I., (2021). Accuracy Assessment of Low-Cost Unmanned Aerial Vehicle (UAV) Photogrammetry. *Alexandria Eng. J.*, Vol. 60(6), 5579–5590. <https://doi.org/10.1016/j.aej.2021.04.011>.
- [3] Ge, Y., Liu, Y. and Liu, X., (2022). Knowledge Mapping Analysis of Digital Photogrammetry Research Using CiteSpace. *Stavební obzor - Civil Engineering Journal*, Vol. 31(1), 181–195. <https://doi.org/10.14311/cej.2022.01.0014>.
- [4] Berra, E. F. and Peppia, M. V., (2020). Advances and Challenges of UAV SFM MVS Photogrammetry and Remote Sensing: Short Review. *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS 2020)*, 22–26 March 2020, Santiago, Chile, 267-272. <https://doi.org/10.1109/LAGIRS48042.2020.9285975>.
- [5] Saad, A. M. and Tahar, K. N., (2019). Identification of Rut and Pothole by using Multirotor Unmanned Aerial Vehicle (UAV). *Measurement: Journal of the International Measurement Confederation*, Vol. 137, 647–654. <https://doi.org/10.1016/j.measurement.2019.01.093>.
- [6] Deliry, S. I. and Avdan, U., (2021). Accuracy of Unmanned Aerial Systems Photogrammetry and Structure from Motion in Surveying and Mapping: A Review *J. Indian Soc. Remote Sens.*, Vol. 49(8) 1997–2017. <https://doi.org/10.1007/s12524-021-01366-x>.

- [7] Jhan, J. P. and Rau, J. Y., (2019). A Normalized Surf for Multispectral Image Matching and Band Co-Registration. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* Vol. 42(2/W13) 393–399. <https://doi.org/10.5194/isprs-archives-XLII-2-W13-393-2019>.
- [8] Abd Mukti, S. N. and Tahar, K. N., (2022). Detection of Potholes on Road Surfaces Using Photogrammetry and Remote Sensing Methods (Review). *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, Vol. 22(3), 459–471. <https://doi.org/10.17586/2226-1494-2022-22-3-459-471>
- [9] Wang, S., Guo, Z. and Liu, Y., (2021). An Image Matching Method Based on SIFT Feature Extraction and FLANN Search Algorithm Improvement *J. Phys. Conf. Ser.*, Vol. 2037(1), 1-6. <https://doi.org/10.1088/1742-6596/2037/1/012122>.
- [10] Xi, W., Shi, Z. and Li, D., (2017). Comparisons of Feature Extraction Algorithm Based on Unmanned Aerial Vehicle Image. *Open Phys.*, Vol. 15(1), 472–478. <https://doi.org/10.1515/phys-2017-0053>.