https://doi.org/10.2352/EI.2022.34.6.IRIACV-307

Efficient Landslide Detection by UAV-based Multi-temporal Visual Analysis

Yosuke Yamaguchi [a], ***Kai Matsui*** [a], ***Jun Ohya*** [a], ***Katsuya Hasegawa*** [b], ***and Hiroshi Nagahashi*** [c]
[a] ***Waseda University, Tokyo, Japan***
[b] ***Japan Aerospace Exploration Agency, Institute of Space and Astronautical Science, Tokyo, Japan***
[c] ***Tokyo Institute of Technology, Tokyo, Japan***

Abstract

This paper proposes a landslide detection method by UAV-based visual analysis. The fundamental strategy is to detect ground surface elevation changes caused by landslides. Our method consists of five steps: multi-temporal image acquisition, ground surface reconstruction, georeferencing, elevation data export, and landslide detection. In order to improve efficiency, we use Visual Simultaneous Localization and Mapping for ground surface reconstruction. It can perform faster than conventional methods based on Structure-from-Motion. In addition, we introduce convolutional neural network (CNN) to detect landslides robustly in the multi-temporal elevation data. The experimental results in a simulation environment show that the proposed method runs 5.5 times as fast as the conventional methods. In addition, the CNN-based model achieved F1 score of 0.79-0.84, showing robustness against reconstruction noise and registration error.

Introduction

Landslides have been a great threat to human lives and economy [1]. Due to heavy rainfalls and earthquakes, the number of landslides in Japan is increasing [2]. In addition, because of labor shortage in forestry, landslides in remote areas tend to be left unfound for a long time. Considering the circumstances, fast and accurate detection of landslides is crucial for efficient disaster management.

Research on landslide detection adopts either ground-based methods or remote-sensing methods. The former methods monitor the ground displacement by placing sensors directly in the target regions [3] [4]. On the other hand, remote-sensing methods analyze the ground surface data which is remotely acquired from satellites, aircraft, or unmanned aerial vehicles (UAVs). The main advantage of remote sensing is the capability to provide spatially continuous data with high precision [5]. Among them, UAVs equipped with optical cameras are powerful platforms as they enable high-resolution, low-cost, and flexible visual analysis [6].

Structure from Motion - Multi View Stereo (SfM-MVS) [7] is an important 3D reconstruction technique to analyze an UAV image set. In the SfM process, tie points, such as SIFT [8] features, are first detected and matched within the image set. Then, bundle adjustment estimates the camera models, the camera orientations, and the sparse point-cloud [9]. MVS densifies the point-cloud and reconstructs the 3D terrain model with color. Since the obtained model is in relative coordinates, georeferencing process based on Ground Control Points (GCPs) is applied. The resulting model can generate an orthophoto and a Digital Surface Model (DSM). A DSM represents an elevation data of ground surface, including surface components such as vegetation and buildings. In addition, by filtering out these objects, a Digital Elevation Model (DEM) can be obtained.

James et al. [9] compared 3D models derived from SfM workflow with data acquired by laser scanning, showing the relative precision (measurement precision: observation distance) of 1:1000. D'Oleire-Oltmanns et al. [10] applied SfM workflow to UAV image sets taken from 70 m above ground level(a.g.l.) and achieved 3 cm precision. As it enables highly accurate measurement without expert knowledge, SfM workflow has become the standard tool for landslide detection by UAV analysis [5] [11] [12] [13].

Tanteri et al. [5] [11] used SfM workflow to monitor landslides by multi-temporal UAV surveys. They focused on elevation change caused by landslides. They calculated difference of DEMs between the surveys and detected landslides in the target slope. Xu et al. [13] used a similar approach based on DEM difference to detect landslides in a wide area of 32 km^2.

However, SfM workflow has the major drawback of heavy workload. Rossi et al. [11] reported that it took several hours to apply SfM workflow to a 0.0186km^2 area with a workstation (CPU 2x Xeon 2.93 GHz, 32 GB RAM, GPU Radeon HD 5870). The long computation time makes it difficult to cover large areas efficiently. Other difficulties arise in georeferencing and filtering phase. Xu et al. [13] installed 160 GCPs to perform georeferencing in 32 km^2. However, it is uncertain whether sufficient GCPs can be captured during disasters. Tanteri et al. [5] pointed out that the filtering algorithm [14] is less effective on photogrammetric point-clouds than on laser scanning data, causing 20-30 cm errors in DEMs. As the detection approach based on DEM difference is vulnerable to noise, such errors can lead to misdetection.

Visual Simultaneous Localization and Mapping (Visual-SLAM) and convolutional neural network (CNN) are promising techniques to overcome these problems. Visual-SLAM can map the environment and calculate the camera orientations in parallel with image input. While SfM workflow prioritizes accuracy by using sufficient images, Visual-SLAM focuses on processing the input frames in real time [15]. CNN is known to be effective in change detection between two images [16]. Change detection task is complex in that variations caused by extrinsic events need to be ignored. Compared to difference-based detection, CNN is robust to variations such as noise, registration, and occlusion [17].

In this paper, we propose a landslide detection method based on Visual-SLAM and CNN. Our method has the following advantages.

- Visual-SLAM enables fast reconstruction of 3D terrain models.
- Fusion of trajectories by Visual-SLAM and Global Navigation Satellite System (GNSS) enables georeferencing without GCPs.
- CNN enables robust landslide detection compared to methods based on DEM difference.

Proposed method

The proposed method is shown in Fig. 1. It consists of five steps: image acquisition, ground surface reconstruction, georeferencing, DSM export and landslide detection. The fundamental strategy is patch-based change detection on pre- and post-landslide DSMs. Each step is detailed in the following.

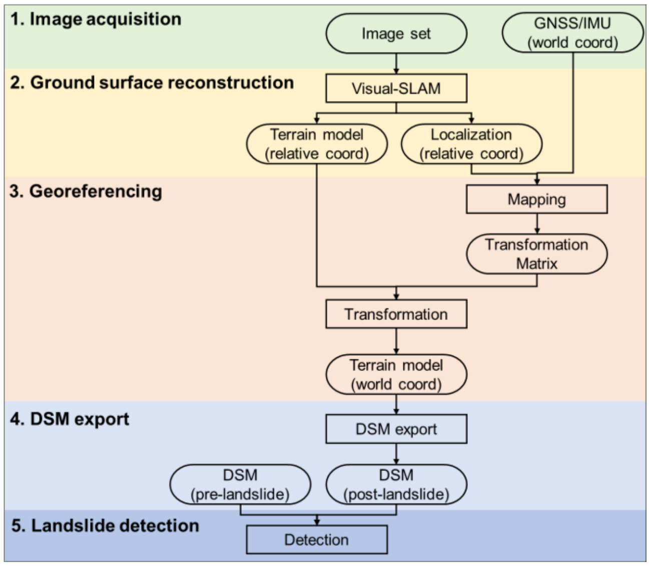


***Figure 1**. Overview of the proposed method*

Image acquisition

UAV surveys are made at fixed intervals to obtain a ground surface image set. In addition, emergency mission is conducted in case of heavy rainfalls and earthquakes, when the risk of landslides is high. This step outputs pre- and post-landslide ground surface image sets.

Ground surface reconstruction

UAV images are resized into 640×360 pixels and processed by Direct Sparse Odometry (DSO) [18], a kind of Visual-SLAM. Direct formulation of DSO leads to robustness in sparsely textured environments. Sparse reconstruction of DSO prioritizes large-scale accuracy over local smoothness [18]. These features are suitable for reconstructing broad ground surface. Since obtained point-cloud contains some noise points caused by calculation errors, statical outlier point removal is applied. This step outputs the colorless 3D terrain model and the localization result in relative coordinates.

Georeferencing

Georeferencing transforms the terrain model into world coordinates. Georeferencing methods without GCPs have been proposed in numerous studies (e.g., [19] [20]). In this research, georeferencing is conducted by simple fusion of trajectories. The detailed process is as follows.

Coordinate transformation is calculated by Eq. (1).

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = T \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{1}$$

where T is the transformation matrix, (x, y, z) is the position in relative coordinates, and (X, Y, Z) is the position in world coordinates. Transformation matrix is a square matrix of four rows and columns. Thus, it can be calculated as shown in Eq. (2) by choosing four corresponding points in both coordinates.

$$T = \begin{pmatrix} X_1 & X_2 & X_3 & X_4 \\ Y_1 & Y_2 & Y_3 & Y_4 \\ Z_1 & Z_2 & Z_3 & Z_4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ 1 & 1 & 1 & 1 \end{pmatrix}^{-1} \tag{2}$$

where the determinant of the matrix in relative coordinates is not 0. This condition means that the selected four points are not to be on a same plane.

Our method uses the trajectory by Visual-SLAM, that is, the localization result, and the trajectory by integrated log of GNSS/IMU (Fig. 2). Corresponding points are selected using time stamps in both trajectories. This step outputs the terrain model in world coordinates.

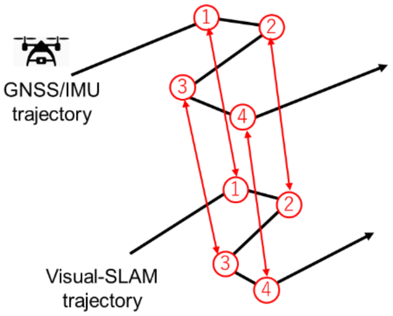


***Figure 2**. Fusion of trajectories*

DSM export

A DSM is generated by extracting elevation data from the terrain model. First, the terrain model is split into grids. Each grid corresponds to a pixel in the DSM. Second, the pixel value is determined by the average height of the corresponding grid. Finally, missing values which occurred in sparse regions in the terrain model are filled by the inpainting technique based on Fast Marching Method [21]. This step outputs the DSM as a gray scale image.

Landslide detection

Previous four steps outputs pre- and post-landslide DSMs. Landslide detection is conducted on them. First, two DSMs are concatenated in the channel dimension. Second, this two-channel data is split into 32 pixel-sized patches. The CNN-based detection model, which is inspired by AlexNet [22], conducts binary classification (landslide or non-landslide) on each patch (Fig. 3). This process is exploited for elevation change detection caused by landslides while ignoring variations caused by reconstruction, georeferencing and vegetation growth.

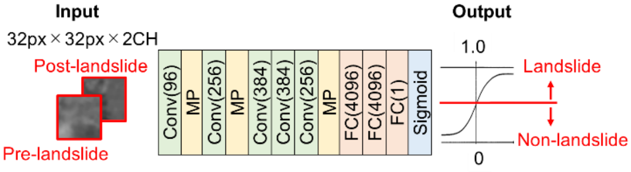


***Figure 3**. Architecture of the detection model. Conv(n) is convolutional layer with n filters of spatial size 3×3 applied with stride 1. MP is max-pooling layer of size 2×2 applied with stride 1. FC(n) is fully connected linear layer with n output units. Batch-normalization follows all the Conv layers and the first and the second FC layers.*

Experiments

We conducted two kinds of experiments to verify the effectiveness of the proposed method. Experimental field was 20 m^2 of sandy area with two mounds of sand (mound-A, mound-B), which were 1.5 m in height (Fig. 4a). Three checkpoints (Fig. 4b) were set in the field. A quadrotor UAV, DJI Phantom 4 Pro, was used for image acquisition. Frame rate was set at 30 fps. Mouse G-Tune HN-Z (CPU Core^{TM} i9-10900K 3.70GHz, GPU NVIDIA GeForce RTX 3070, RAM 32GB) was used for image processing.

(a) mound *(b) checkpoint*

***Figure 4**. Experimental field*

Registration accuracy

In this experiment, we evaluated the registration accuracy of the proposed georeferencing method.

Procedure

First, UAV took a flight on the three-dimensional trajectory as shown in Fig. 5 and acquired an image set. There was 60% overlap between images in the lateral direction. Second, a terrain model was reconstructed and georeferenced by the proposed method. This procedure was repeated three times and three terrain models were obtained. Lastly, registration error was calculated as relative location difference of the three terrain models. To be specific, mean location difference at the three checkpoints was calculated.

Result

Figure 5 shows one of the obtained georeferenced models. The trajectories by Visual-SLAM (green in Fig. 5) and GNSS/IMU (blue) mostly matched. Table 1 shows that the proposed method yielded approximately 1 m and 1.2 m registration error in the horizontal and vertical directions, respectively. The main error factor would be accuracy of GNSS. GNSS data of Phantom4Pro depends on stand-alone GPS, which has positional accuracy of 5-10 m [23]. Using more precise positioning technology such as real-time kinematic GPS could improve the registration accuracy.

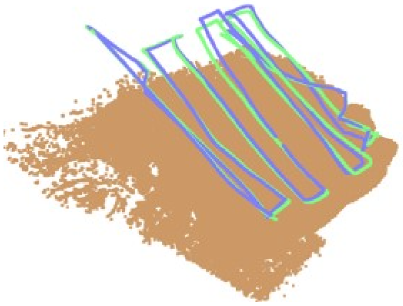


***Figure 5.** Georeferenced model. brown) Terrain model; blue) GNSS/IMU trajectory; green) Visual-SLAM trajectory. UAV was controlled at varying altitudes of 7-20 m a.g.l.*

Table 1. Registration error at the checkpoints

Direction	Registration Error (m)
East	0.77
North	1.06
Elevation	1.22

Efficiency

In this experiment, we evaluated the processing time and the detection ability of the proposed method. Landslides were simulated in one-tenth scale (Table 2). Elevation change includes the effect of fallen trees as well as mass movement caused by landslides. Registration error was set based on the result of the former experiment. The result was compared with the conventional approach which is based on DEM difference derived from SfM workflow.

Table 2. Simulation scale.

Case	DSM resolution (cm/pixel)	Elevation change (m)	Registration error (m)
Actual	20	15 (max.)	Horizontal: 1.0 Vertical: 1.2
Simulated	2	1.5 (max.)	Horizontal: 0.10 Vertical: 0.12

Procedure

(a) Image acquisition

Landslides were simulated by collapsing a part of the mounds in five steps. In each step, the UAV acquired an image set with 60% overlap in the lateral direction. The flight altitude was set at 7 m a.g.l. Six image sets, Step1 (non-collapsed) to Step6 (fully collapsed), were obtained.

(b) Data processing

Each image set was processed by DSO. Terrain models were georeferenced manually, so that they contained no registration error. DSMs were exported at the resolution of 2 cm/pixel and 1.2 cm/intensity.

In addition, SfM workflow was applied using an open source SfM software, OpenDroneMap [24]. Input frames were resized to 2560 ×1350 pixels and extracted at 3 fps, which yielded 90% overlap. The process was accelerated by using GPU for SIFT extraction and pre-matching neighboring 96 frames based on Exif data. DSMs and orthophotos were exported at the same resolution.

(c) Training and Test

Figure 6 shows the process to create training datasets. Step1 was considered as pre-landslide state and other steps were considered as post-landslide state. First, ground truth masks of landslide areas were made based on the SfM-derived orthophotos. Landslide areas were marked in white color. Second, the DSO-derived DSMs and the ground truth masks were divided into 42 pixel-sized patches (including 10-pixel margin) with a stride of 10 pixels. The DSM patches were labeled as 0 (non-landslide) or 1 (landslide) based on the ground truth patches. In detail, when the change rate (the ratio of white color region) in the central 32 pixel-sized area was more than 10%, the patches were labeled as 1. This process generated 1561 patches of datasets for each class.

Table 3 shows the training condition of the detection model. Hyperparameters were optimized by grid search. Two kinds of training, fixed training and loose training, were conducted in order to improve robustness against registration error. In the fixed training, the central 32 pixel-sized area of the DSM patches were cropped and input into the detection model. In the loose training, the pre-landslide patches were randomly cropped using 10-pixel margin. In addition, random intensity ranging between -10 and 10 was added to them (Fig. 7). This random augmentation is used for simulating the registration error shown in Table 2.

The ability of the trained models was evaluated on the test data (step5 of mound-A). The four test patterns (A to D in Table 4) were prepared by manually adding different registration errors.

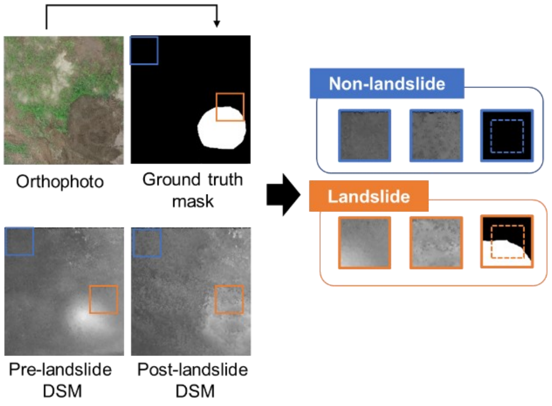


Figure 6. *Dataset creation*

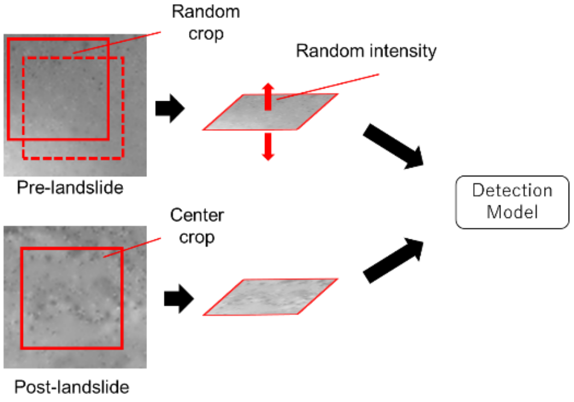


Figure 7. *Augmentation in loose training*

Table 3. Learning condition

Framework	Pytorch
Optimizer	Adam
Loss function	Binary cross entropy
Batch size	32
Weight Initialization	He
Bias Initialization	0
Data Augmentation	flip
Validation data ratio	20%
Learning rate	fixed: 1×10^{-6} / loose: 1×10^{-5}
Regularization	L2-norm (0.01)
Epoch	100

Table 4. Test data

Pattern	Registration error (m)		
	East	North	Elevation
A	0	0	0
B	0.10	0.10	0
C	0.10	0.10	0.12
D	0.10	0.10	-0.12

Result

Table 5 compares the processing times for SfM workflow and DSO. In our settings, DSO processed 5.5 times as fast as SfM on average. Another important difference is that while SfM workflow is applied after the flight mission is finished, DSO can be applied in parallel with image acquisition.

Figure 8 and Fig. 9 show the images in the test case with no registration error. DSM difference can be regarded as DEM difference in this simulation as there was no change in surface components. Concerning the SfM-derived DSM difference, binarization using the threshold of 30 cm successfully removed most noise areas. However, higher threshold led to overlook outer edge of the landslide area. On the other hand, the DSO-derived DSM difference contained many noise areas even with the threshold of 30 cm. This is because the reconstruction accuracy of DSO is less than that of SfM workflow. Considering that registration error can further increase noise areas, it is difficult to detect landslides by simple binarization.

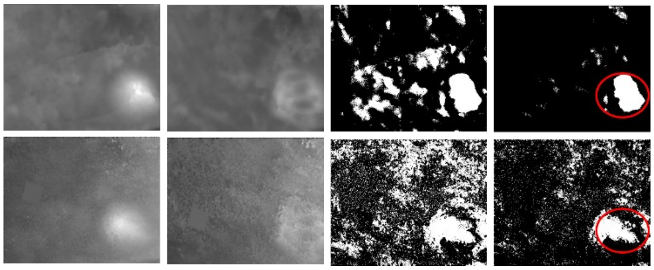
Figures 10, Table 6, and Table 7 show the detection result. The detection ability of the fixed training model dropped on the test pattern with registration error. On the other hand, the loose training model showed higher robustness against registration error. The F1 score of the loose training model was 0.79-0.84. Erroneous detection mostly occurred on the surrounding patches of the landslide. In other regions, the model successfully detected landslides while ignoring variations such as reconstruction noise and registration error.

Table 5. Processing time

Image sets	Processing time (sec)		Ratio SfM/DSO
	SfM	DSO	
Step1	934	155	6.01
Step2	810	203	4.00
Step3	966	101	9.59
Step4	867	155	5.59
Step5	865	184	4.71
Step6	1034	193	5.36
Average	912.7	165.0	5.53

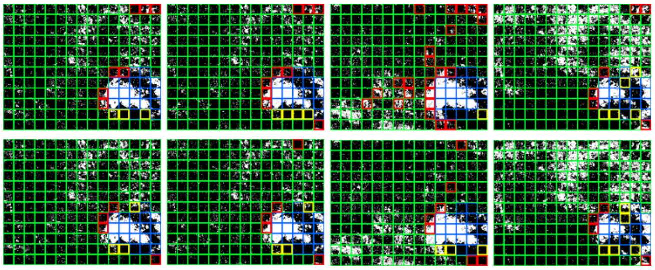
(a) pre-landslide (b) post-landslide (c) ground truth mask

Figure 8. *Orthophotos and a ground truth mask*



(a) pre-landslide DSM (b) post-landslide DSM (c) Difference threshold:20 cm (d) Difference threshold:30 cm

Figure 9. *Comparison between SfM (top row) and DSO (bottom row). The red circles show the landslide area. DSO-derived DSM difference contained more noise areas (white areas exterior to the circle).*



(a) pattern A (b) pattern B (c) pattern C (d) pattern D

Figure 10. *Detection result of fixed training (top row) and loose training (bottom row). blue) Ture positive; green) True negative; red) False Positive; yellow) False negative. The results are projected on the DSM difference with the threshold of 30 cm.*

Table 6. Detection result (fixed training)

Pattern	Precision	Recall	F1
A	0.731	0.864	0.792
B	0.667	0.818	0.735
C	0.477	0.955	0.636
D	0.739	0.773	0.756

Table 7. Detection result (loose training)

Pattern	Precision	Recall	F1
A	0.826	0.864	0.844
B	0.792	0.864	0.826
C	0.731	0.864	0.792
D	0.810	0.773	0.791

Conclusion

This paper has proposed a landslide detection method by UAV-based multi-temporal surveys. The key point is the use of Visual-SLAM and CNN to enable efficient detection.

First, we have confirmed the registration accuracy of the georeferencing method which is based on fusion of trajectories. Second, we have evaluated the detection performance in an outdoor simulation environment. Experimental results showed that our method works faster than the conventional method. In addition, the CNN-based model detects landslides robustly against variations such as reconstruction noise and registration error.

In our future work, we aim to evaluate the proposed method in a practical environment.

References

[1] M. Scaioni, L. Longoni, V. Melillo and M. Papini, "Remote Sensing for Landslide Investigations: An Overview of Recent Achievements and Perspectives," *Remote Sensing,* vol. 6, no. 10, pp. 9600-9652, 2014.

[2] Ministry of Land, Infrastructure, Transport and Tourism, "Occurrence of sediment-related disasters in 2020," 2020.

[3] J. P. Malet, M. Olivier and E. Calais, "The Use of Global Positioning System techniques for the continuous monitoring of landslides. Application to the Super-Sauze earthflow," *Geomorphology,* pp. 33-54, 2002.

[4] I. T. Yang, J. K. Park and D. M. Kim, "Monitoring the symptoms of landslide using the non-prism total station," *KSCE Journal of Civil Engineering,* vol. 11, p. 293–301, 2007.

[5] L. Tanteri, G. Rossi, V. Tofani, P. Vannocci, S. Moretti and N. Casagli, "Multitemporal UAV Survey for Mass Movement Detection and Monitoring," in *Workshop on World Landslide Forum*, Ljubljana, Slovenia, 2017.

[6] C. Li, G. ZHANG, T. LEI and A. GONG, "Quick image-processing method of UAV without control points data in earthquake disaster area," *Transactions of Nonferrous Metals Society of China,* vol. 21, pp. 523-528, 2011.

[7] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey and J. M. Reynolds, "'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications," *Geomorphology,* vol. 179, pp. 300-314, 2012.

[8] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision,* vol. 60, pp. 91-110, 2004.

[9] M. R. James and S. Roboson, "Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application," *Journal of Geophysical Research: Solid Earth,* vol. 117, no. F3, 2012.

[10] S. D'Oleire-Oltmanns, I. Marzolff, K. Peter and J. Ries, "Unmanned Aerial Vehicle (UAV) for Monitoring Soil Erosion in Morocco," *Remote Sensing,* vol. 4, no. 11, pp. 3390-3416, 2012.

[11] G. Rossi, L. Tanteri, V. Tofani, P. Vannocci, S. Moretti and N. Casagli , "Multitemporal UAV surveys for landslide mapping and characterization," *Landslides,* vol. 15, no. 5, p. 1045–1052, 2018.

[12] A. Lucieer, S. Jong and D. Turner, "Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography," *Progress in Physical Geography: Earth and Environment,* vol. 38, no. 1, pp. 97-116, 2014.

[13] Q. Xu, W. Li, Y. Ju, X. Dong and D. Peng, "Multitemporal UAV-based photogrammetry for landslide detection and monitoring in a large area: a case study in the Heifangtai terrace in the Loess Plateau of China," *Journal of Mountain Science,* vol. 17, no. 8, p. 1826–1839, 2020.

[14] N. Brodu and D. Lague, "3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 68, pp. 121-134, 2012.

[15] K. Kado and G. Hirasawa, "Applying pose estimation techniques with Structure from Motion in architectural fields," *AIJ Journal of Technology and Design,* vol. 24, no. 57, pp. 873-876, 2018.

[16] S. R. Klomp and D. W. J. M. van de Wouw, "Real-time small-object change detection from ground vehicles using a siamese convolutional neural network," *Journal of Imaging Science and Technology,* vol. 63, no. 6, 2019.

[17] A. Varghese, J. Gubbi, A. Ramaswamy and B. Purushothaman, "ChangeNet: A Deep Learning Architecture for Visual Change Detection," in *Computer Vision – ECCV 2018 Workshops*, Munich, Germany, 2019.

[18] J. Engel, V. Koltun and D. Cremers, "Direct Sparse Odometry," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* vol. 40, no. 3, pp. 611-625, 2018.

[19] D. P. Shepard and T. E. Humphreys, "High-precision globally-referenced position and attitude via a fusion of visual SLAM, carrier-phase-based GPS, and inertial measurements," in *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, Monterey, CA, USA, 2014.

[20] D. Turner, A. Lucieer and L. Wallace, "Direct Georeferencing of Ultrahigh-Resolution UAV Imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 5, pp. 2738-2745, 2014.

[21] A. Telea, "An Image Inpainting Technique Based on the Fast Marching Method," *Journal of graphics tools,* vol. 9, no. 1, pp. 25-36, 2004.

[22] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems,* vol. 25, no. 2, 2012.

[23] B. Grayson, N. Penna, J. Mills and D. Grant, "GPS precise point positioning for UAV photogrammetry," *The Photogrammetric Record,* vol. 33, no. 84, 2018.

[24] OpenDroneMap, https://github.com/OpenDroneMap/ODM.

Author Biography

Yosuke Yamaguchi is now an undergraduate student of the Department of Modern Mechanical Engineering, the School of Creative Science and Engineering, Waseda University, Japan. His research interests include deep learning, image processing, and robot vision.

Kai Matsui received his Bachelor degree from Department of Modern Mechanical Engineering, Waseda University in March 2021. He is now a master student of the same department and exploring a method for detecting land collapses by analyzing images acquired by a Drone camera.

Jun Ohya received B.S., M.S. and Ph.D. degrees in precision machinery engineering from the University of Tokyo, Japan, in 1977, 1979 and 1988, respectively. He joined NTT Research Laboratories in 1979. He was a visiting research associate at the Computer Vision Laboratory, University of Maryland, USA, from 1988 to 1989. He transferred to ATR, Kyoto, Japan, in 1992. In 2000, he joined Graduate School of Global Information and Telecommunication Studies, Waseda University, Japan as a professor. He was a guest professor at the University of Karlsruhe, Germany, in 2005. Since 2014, he has been a professor of Department of Modern Mechanical Engineering, Waseda University. His research areas include computer vision and machine learning. Dr. Ohya is a Life Member of IEEE, a Fellow of IIEEJ, a member of IPSJ, IIEIC and VRSJ.

Katsuya Hasegawa
B.S., JAPAN Coast Guard Academy; Ph.D., Department of Biomedical Engineering, Kawasaki Medical School; Ph.D., Department of Applied Chemistry, Tokyo University of Agriculture and Technology; Japan Coast Guard officer; appointed as a researcher at the Japan Aerospace Exploration Agency; research fields include metrology and biophysics. He is a part-time lecturer at several universities, teaching mathematics, physics, and life sciences.

Hiroshi Nagahashi received his BEE (1975) and DE in information processing (1980) from Tokyo Institute of Technology, Japan. Since then, he has worked in the fields of image processing, computer vision, computer graphics and machine learning. He is now a Specially Appointed Professor in the School of Computing, Tokyo Institute of Technology.