EXAMINATION OF A UAV IMAGE CLASSIFICATION METHOD BY USING MACHINE LEARNING AND WAVELET TRANSFORM

Akito MOMOSE1, Hitoshi MIYAMOTO2, Shuji IWAMI3 and Takayuki NAGAYA4

1Member of JSCE, CTI Engineering Co., Ltd., Former graduate student, Shibaura Institute of Technology
(1-14-6 Kamikizaki, Urawa-ku, Saitama 330-0071, Japan)
E-mail: akt-momose@ctie.co.jp
2Member of JSCE, Professor, Department of Civil Engineering, Shibaura Institute of Technology
(3-7-5 Toyosu, Koto, Tokyo 135-8548, Japan)
E-mail: miyamo@shibaura-it.ac.jp (Corresponding Author)
3Member of JSCE, CTI Engineering Co., Ltd. (1-14-6 Kamikizaki, Urawa-ku, Saitama 330-0071, Japan)
E-mail: iwami@ctie.co.jp
4Member of JSCE, CTI Engineering Co., Ltd. (1-14-6 Kamikizaki, Urawa-ku, Saitama 330-0071, Japan)
E-mail: nagaya@ctie.co.jp

This paper examined a machine learning technique with the wavelet transform for classifying land cover conditions in Unmanned Aerial Vehicle (UAV) images of a riverine landscape. The UAV images were taken in a river course of Kurobe River, Japan. Each UAV image analyzed was composed of RGB, Normalized Difference Vegetation Index (NDVI), and a Digital Surface Model (DSM) of the river geomorphology made from a Structure from Motion (SfM) image processing of the UAV images. In a pre-processing of the machine learning, the DSM was decomposed into low/high wavenumber components through wavelet transform, and its edges were further extracted to effectively utilize the height difference information in DSM. The result of the machine learning showed that the F-measure had high enough above 0.91 in the dataset including all characteristic values from RGB, DMS, and NDVI into the machine learning algorithm.

Key Words: river management, riparian vegetation, UAV, land cover classification, machine learning, wavelet transform

1. INTRODUCTION

Many rivers in Japan have experienced vegetation overgrowth phenomena in the recent several decades1, 2. Excessive vegetation recruitment into river channels could reduce the discharge capacity of floods and significantly transform the river ecosystem of sandbars and gravel beds. Therefore, appropriate control of excessive riparian vegetation is one of the important issues to be solved in river management. The vegetation overgrowth phenomena have been progressing on a decadal time scale. In order to accurately grasp the phenomenon for proper river management, it would be of great importance to continuously monitor the vegetation status with the corresponding hydrological/hydraulic/geomorphological conditions.

One method of continuously and efficiently monitoring changes in river channel conditions is aerial photography using an Unmanned Aerial Vehicle (UAV). It could be because UAVs have become relatively inexpensive and have easily been manipulated for research and technological development. In addition, the dramatic development of Structure from Motion (SfM) technology has made it possible to generate orthoimages and Digital Surface Model (DSM) easier.

Previous studies3-5 using UAV images tried to detect vegetation automatically through information from the image texture, RGB values, and the Normalized Difference Vegetation Index (NDVI). As for tree detection, machine learning with multispectral information extracted the tall coniferous trees from the UAV images3. As for the usage of the RGB values, machine learning with UAV mosaic images used the image segmentation of the vegetation community for vegetation classification5. Deep learning with UAV images automatically detected the presence or absence of willows from riparian vegetation images5. These studies focused on vegetation
classification only, and the other classes on the land use and land cover conditions have not yet been fully investigated.

The authors took UAV aerial photographs of several river channels of the Kinu River and tried to classify the state of the land cover using maximum likelihood estimation (MLE) with their RGB information of the river channels\textsuperscript{6}. However, the MLE had a serious problem. There was a high possibility of misclassification in parts having similar colors, such as woody and herbaceous, dead grass and bare land, and so forth. Therefore, in the previous research works\textsuperscript{7,8}, the authors tried to examine a machine learning method that selected DSM as a feature for machine learning in addition to RGB and NDVI obtained from the UAV aerial imagery for the local spatial range of the Kurobe River.

The present study examines an applied approach of the previous machine learning method\textsuperscript{7,8} to extend it to a broader spatial range of river channels. Specifically, for DSM, which could be significantly affected by the longitudinal gradient of rivers, new features are extracted by performing pretreatment using wavelet transform\textsuperscript{9-12} and Laplacian filter\textsuperscript{13}. Then, the new features are used with RGB and NDVI for the machine learning of riparian land cover classification to achieve higher accuracy. Wavelet transform is applied mathematics for frequency/wavenumber decomposition of signals/images\textsuperscript{9}. It has a lot of advantages in the coherent structure analysis of turbulence in fluid mechanics and hydraulics\textsuperscript{10,11}.

2. STUDY SITE

Figure 1 shows the observation point in the target river, the Kurobe River. The length of the main channel in the Kurobe River is 85 km, and its basin area is 682 km\textsuperscript{2}. The red dot shown in Fig. 1 is the section of the UAV measurement. The analysis target in this study is the 4–6 km section within the alluvial fan from the estuary to the 13 km section. As shown in the RGB image of Fig. 3 (a) below, this section is the section of braided channel where the gravel riverbed develops, and vegetation recruitment into the river channel has been remarkable. The land cover states targeted for machine learning in this study are the water surface, bare land, grass, and tree in the river channel. Typical plant species in the target section are Miscanthus sinensis for grass and Elaeagnus umbellata for trees\textsuperscript{14}.

3. METHODS

(1) Overview

Figure 2 shows the flowchart of UAV image processing examined in this paper. The numbers in Fig. 2 are those of the following subsections. First, the target river channel is measured by UAV to take images of the river channel. Next, the UAV images are processed by SfM to obtain the RGB and NDVI.
orthoimages and DSM. Here we have the following problems: (i) there exist small color fluctuations, which cannot be ignored for analysis, (ii) the longitudinal gradient of the river dramatically affects the spatial distribution of DSM, so that the height differences information of each local site cannot be used for the analysis. This paper tries to introduce the following procedures as preprocessing for machine learning to deal with these problems. One is the removal of the longitudinal river gradient using the Laplacian filter. The other is the edge extraction using the Wavelet transform for DSM. The other is the edge extraction using the Laplacian filter. Furthermore, based on the RGB information, pixels with similar characteristics are grouped into a tiny segment (referred to as an image object) by ArcGIS. This image object is used as the minimum unit for land cover classification. High-frequency noise, such as the small color fluctuations, is removed by averaging the RGB within the image object. Finally, using the preprocessed data, land cover classification using machine learning is performed. We used random forests as a machine learning algorithm. Machine learning attempts to classify the water surface, bare land, grass, and tree. This paper examines the classification accuracy of the machine learning that uses image features based on the RGB, NDVI, and DSM of UAV after the preprocessing mentioned above.

(2) UAV aerial photography and SfM processing

The UAV images for the target section shown in Fig. 1 were taken from 10:00 am to 12:00 am on November 28, 2017. The UAV altitude above ground level was 140m. The UAV used was eBee (manufactured by senseFly). The onboard camera was SEQUOIA (manufactured by Parrot) that can shoot 5-band multispectral and RGB simultaneously. The image resolution is 1280 × 960 pixels for multispectral and 4600 × 340 pixels for RGB. The SfM processing was applied to the captured images to obtain the RGB and NDVI orthoimages and the DSM. The SfM software used was Pix4D Mapper (manufactured by senseFly).

(3) Data preprocessing

a) Wavelet transform

Wavelet transform is one of the applied mathematical methods in image processing. It decomposes the spatial distribution of pixel values into low- and high-wavenumber components. Therefore, it is possible to extract components of any wavenumber scale and resynthesize them. This paper decomposed the DSM into several wavenumber components using a two-dimensional discrete wavelet transform. This paper used the Wavelet Toolbox of MATLAB to implement the wavelet transform. Daubechies (N = 4) was used for the discrete wavelet function. The DSM was decomposed into eight levels, in which low wavenumber components were removed. Then, the remaining high wavenumber components were resynthesized and used for machine learning. Hereafter, the resynthesized images are referred to as the DSM high wavenumber.

b) Laplacian filter

The Laplacian filter is one of the convolution filters that detect edges in an image. This paper used the ArcGIS convolution filter function as the Laplacian filter for calculation. The following equation defines the kernel of the Laplacian filter:

\[
\begin{pmatrix}
0 & -1 & 0 \\
-1 & 4 & -1 \\
0 & -1 & 0
\end{pmatrix}
\] (1)

The image analysis with this kernel is equivalent to taking the second derivative’s summation both in the horizontal and vertical directions. Applying the Laplacian filter to DSM detected the edges of the DSM image. From hereon, the obtained image will be referred to as the DSM edge.

c) Image object

This study used the segment average shift tool in ArcGIS to generate the image objects. This tool made it possible to group adjacent pixels with similar characteristics into one image object. The statistical values of NDVI, the DSM high wavenumber, and the DSM edge in the created image object were calculated and used as features in machine learning. In contrast, the land covers in the image object were manually determined based on the visual inspection of the RGB image. The obtained trues were used for machine learning to verify the model accuracy. In the image output by the SfM processing, error noises were observed around a bridge. Therefore, this study excluded image objects containing the bridge from the machine learning analysis.

(4) Land cover classification by machine learning

a) Feature dataset

Table 1 shows the features used in machine learning. RGB features used its average values of the image objects, while those of the DSM high wavenumber, the DSM edge, and NDVI both the average value and the standard deviation for machine learning. This paper examined the effect of combining the features on the classification accuracy of land covers in machine learning performance.

b) Random forest

This study chose random forests as the machine learning algorithm. Random forests classify land covers using a large number of decision trees in their algorithm.
This study used the scikit-learn\textsuperscript{17}) in Python library to implement the random forests. The evaluation of land cover classification in machine learning employed a cross-validation method that divides the whole dataset into five sub-datasets. F-measure was used as an evaluation index. Equation (4) defines the F-measure, which was the harmonic mean of Precision in Eq.(2) and Recall in Eq.(3):

\begin{equation}
Precision = \frac{TP}{TP + FP} \tag{2}
\end{equation}
\begin{equation}
Recall = \frac{TP}{TP + FN} \tag{3}
\end{equation}
\begin{equation}
F - measure = \frac{2 \cdot Precision \times Recall}{Precision + Recall} \tag{4}
\end{equation}

Here, TP (True Positive) is the number of image objects whose land cover predicted by machine learning is correct; FP (False Positive) is the number of image objects whose land cover predicted by machine learning is incorrect; and FN (False Negative) is the number of image objects whose correct land cover was not predicted by machine learning.

4. RESULTS AND DISCUSSION

(1) RGB and NDVI images from UAV aerial photography
Figure 3 shows the RGB orthoimage processed by SfM and the true values of the land cover determined by the RGB orthoimage. The red frames shown in Fig. 3 (a) are the areas for classifying the land cover. The river stream flows from right to left in Fig. 3. As shown in the true values in Fig. 3 (b), trees were abundant in the gravel riverbed, and grasses were well-grown on the embankment side of the river channel, including the high-water channel. Since the water surface and bare land could be recognized in the orthoimage, RGB was expected to be effective for their classification.

The NDVIs shown in Fig. 4 were high in vegetation and low in water surface. In addition, NDVIs in trees tended to be higher than those in grasses. This was because the NDVIs became higher values depending on the activity of the plant. Therefore, the NDVIs were expected to be effective in classifying trees and grasses.

(2) Feature Extraction from DSM

a) Wavenumber decomposition of DSM using wavelet transform
Figure 5 shows the results of the wavelet wavenumber decomposition for DSM. The original DSM shown in Fig. 5 (a) had high values toward the upstream side, confirming the influence of the longitudinal river gradient. Similarly, the low wavenumber component after the decomposition shown in Fig. 5 (b) also became high as it went upstream. The DSM

| Table 1 Features in each dataset used in machine learning. |
|--------------|---|---|---|---|---|---|---|---|
| Dataset      | A  | B  | C  | D  | E  | F  | G  | H  |
| RGB          | ○  | ○  | ○  | ○  | ○  | ○  | ○  | ○  |
| DSM high wavenumber | ○  |     |     |     |     |     |     |     |
| DSM edge     |     | ○  |     |     | ○  |     |     | ○  |
| NDVI         |     |     |     |     |     | ○  | ○  | ○  |

Fig. 3 UAV orthoimage and true land covers in the target river section of Kurobe River.

Fig. 4 NDVI image of the Kurobe River.
high wavenumber shown in Fig. 5 (c) was obtained by removing the low wavenumber component from the original DSM. Thus, the influence of the longitudinal river gradient shown in Fig. 5 (a) was eliminated in the DSM high wavenumber shown in Fig. 5 (c). In addition, the DSM high wavenumber tended to be high in the trees indicated in Fig. 3. Therefore, the DSM high wavenumber could have information on the relative height difference between trees and the other land covers.

b) Edge detection by the Laplacian filter

Figure 6 shows the DSM edge detection results. The DSM edges were high at the boundaries of trees and other land covers. In particular, as the height difference was remarkable at the boundary between trees and the ground next to them, the DSM edge became higher. Therefore, the DSM edge could be effective for classifying the image objects of trees at the boundary.

(3) Accuracy comparison of the feature datasets for random forests

Table 2 summarizes the F-measure of each dataset in machine learning. Figure 7 shows the classification results of datasets A and H. The upper part of Fig. 7 shows both the RGB image and the true values.

First, the results of the datasets A and H were compared; the former used only RGB while the latter used all features, i.e., the RGB, DSM high wavenumber, DSM edge, and NDVI. Table 2 indicated that the F-measure in the dataset H exceeded 0.91 for all land covers. The classification results of Fig. 7 revealed that trees were often misclassified as water surfaces in dataset A, while they could be accurately classified in dataset H. In addition, an 8%–9% vegetation misclassification found in the dataset H could be due to a mixture of multiple land covers within an image object, as well as the high DSM edge values at the boundaries of grasses and the other land covers.

Next, the contribution of each feature was evaluated; i.e., the DSM high wavenumber, DSM edge, and NDVI, against the classification accuracy of land covers by comparing their F-measures. Comparing the datasets A, B, and C, combining RGB with the DSM high wavenumber or the DSM edge improved the F-measure of grasses and trees to 0.02–0.03. Therefore, the features based on the DSM height difference could contribute to the classification of grasses and trees. Comparing the datasets A and D confirmed that NDVI could contribute to the water surface and vegetation classification. In particular, NDVI had a considerable improvement in the F-measure of trees. This was because NDVI was high enough in the parts where the tree canopy was dense. Finally, comparing the datasets B, C, and E confirmed that using both the DSM high wavenumber and the DSM edge could improve the F-measure compared to using either one only. Thus, the DSM high wavenumber and the DSM edge could complement helpful information for land cover classification in random forests.

This paper did not examine the influences of the quality change of UAV images depending on the weather and time zone and the detailed classification of plant species, etc. These issues are for future research works.
5. CONCLUSIONS

This paper examined a new land cover classification method for UAV river aerial photography using wavelet transform and machine learning. The target river section was a 2-km section of a gravel bed river in an alluvial fan of the lower Kurobe River. The features for machine learning were characterized by using DSM obtained through SfM processing in addition to the RGB image of UAV aerial photography and NDVI. Furthermore, to effectively use the height difference information of DSM, the wavelet decomposition and the edge detection of DSM were performed as preprocessing. Through these treatments, the influence of the longitudinal river gradient was removed from the original DSM, and it became possible to use the local height difference of DSM as the feature quantity. When random forests performed the land cover classification with these features, highly accurate land cover classification resulted from using the dataset with all the features. Future research works plan to examine an algorithm for automatically generating image objects and developing machine learning that can process land cover classification throughout riverscapes.

![Table 2 F-meaures for each dataset in machine learning (random forests).](image1.png)

Note: The circles in the row of Features mean the inclusion of each feature in the dataset, in the order from the left, the RGB, DSM high wavenumber, DSM edge, and NDVI. See details in Table 1.

![Fig. 7 Image comparison of machine learning classification results for datasets A and H.](image2.png)
ACKNOWLEDGMENTS: The research funding support to H.M. from the JSPS KAKENHI Grant Number JP20H02261 and the River Fund of the River Foundation, Japan, is gratefully acknowledged.

REFERENCES
1) Miyamoto, H., Akamatsu, Y. and Toda, Y.: Current status and future prospects for research on riverine forest issues, Advances in River Engineering, Vol. 19, pp. 441-446, 2013 (in Japanese).
2) JSCE Committee on Hydrosceience and Hydraulic Engineering, Subcommittee on Environmental Hydraulics, eds.: Environmental Hydraulics, Chap. 6.3., Riverine forestation, pp. 233–230, JSCE, 2015 (in Japanese).
3) Franklin, S. E., Ahmed, O. S. and Williams, G.: Northern conifer forest species classification using multispectral data acquired from an unmanned aerial vehicle, Photogrammetric Engineering & Remote Sensing, Vol. 83(7), pp. 501–507, 2017.
4) Suzuki, T., Tsuchiya, T., Suzuki, S. and Yamada, A.: Vegetation classification using a small UAV based on superpixel segmentation and machine learning, Journal of the Remote Sensing Society of Japan, Regular Papers, Vol. 36, pp. 59–71, 2016 (in Japanese).
5) Saito, M., Wakuda, Y., Ichikawa, K., Amaya, K., Nasuno, A., Oishi, T., Ikeuchi, K. and Ishikawa, Y.: Development of river channel maintenance via automatic discrimination of vegetation using UAV and deep learning, Journal of JSCE, Ser. B1 (Hydraulic Engineering), Vol. 74(4), pp. I_829–I_834, 2018 (in Japanese).
6) Iimura, H., Miyamoto, H., Inoue, T., Chigasaki, Y. and Hamauchi, K.: Impact assessment of a large flood on riparian vegetation in Kinugawa river channels by using UAV measurement, Journal of JSCE, Ser. B1 (Hydraulic Engineering), Vol. 73(4), pp. I_1069–I_1074, 2017 (in Japanese).
7) Miyamoto, H., Momose, A. and Iwami, S.: UAV image classification of a riverine landscape by using machine learning techniques, EGU2018-5919, EGU General Assembly 2018, Vienna, Austria, 9-13, Apr., 2018.
8) Momose, A., Iwami, S., Nagaya, T. and Miyamoto, H.: Examination of a new land cover classification method for UAV river aerial imagery by using machine learning, Proceedings of 73rd Annual Conference of JSCE, 2018 (in Japanese).
9) Daubechies, I.: Ten lectures on wavelets, CBMS Lecture Notes Series, SIAM, 1991.
10) Farge, M.: Wavelet transforms and their application to turbulence, Annu. Rev. Fluid Mech., Vol. 24, pp. 395–457, 1992.
11) Miyamoto, H. and Kanda, T.: Analysis on hierarchical flow structure in open-channel with concave bed by using a “multiresolutional-proper orthogonal” hybrid expansion, Journal of JSCE, No. 712, pp. 11–23, 2002 (in Japanese).
12) Wavelet Toolbox - MATLAB – MathWorks. Accessed July 07, 2021, https://jp.mathworks.com/products/wavelet.html
13) Convolution function. Accessed July 07, 2021, https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/convolution-function.htm
14) River Environmental Database. Accessed July 07, 2021, http://www.nilim.go.jp/lab/bg/ksnkankyo/ (in Japanese)
15) Segment Mean Shift. Accessed July 07, 2021, https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analysttoolbox/segment-mean-shift.htm
16) Muller, A. C. and Guido, S.: Introduction to Machine Learning with Python, O’Reilly Media, 2016.
17) scikit-learn, Machine Learning in Python. accessed July 07, 2021, http://scikit-learn.org/stable/#

(Received July 12, 2021)
(Accepted July 26, 2021)