A novel approach for vegetation classification using UAV-based hyperspectral imaging

Tetsuro Ishida\textsuperscript{a,*}, Junichi Kurihara\textsuperscript{a}, Fra Angelico Viray\textsuperscript{b}, Shielo Baes Namuco\textsuperscript{c}, Enrico C. Paringit\textsuperscript{b}, Gay Jane Perez\textsuperscript{c}, Yukihiro Takahashi\textsuperscript{a}, Joel Joseph Marciano, Jr.\textsuperscript{d}

\textsuperscript{a}Faculty of Science, Hokkaido University, Sapporo, Japan
\textsuperscript{b}Training Center for Applied Geodesy and Photogrammetry, University of the Philippines Diliman, Quezon City, Philippines
\textsuperscript{c}Institute of Environmental Science and Meteorology, University of the Philippines Diliman, Quezon City, Philippines
\textsuperscript{d}Advanced Science and Technology Institute, Department of Science and Technology, Quezon City, Philippines

*Correspondence to: T. Ishida, Faculty of Science, Hokkaido University, Kita 10 Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan. (ishida.tetsuro@sci.hokudai.ac.jp)
Abstract

The use of unmanned aerial vehicle (UAV)-based spectral imaging offers considerable advantages in high-resolution remote-sensing applications. However, the number of sensors mountable on a UAV is limited, and selecting the optimal combination of spectral bands is complex but crucial for conventional UAV-based multispectral imaging systems. To overcome these limitations, we adopted a liquid crystal tunable filter (LCTF), which can transmit selected wavelengths without the need to exchange optical filters. For calibration and validation of the LCTF-based hyperspectral imaging system, a field campaign was conducted in the Philippines during March 28–April 3, 2016. In this campaign, UAV-based hyperspectral imaging was performed in several vegetated areas, and the spectral reflectances of 14 different ground objects were measured. Additionally, the machine learning (ML) approach using a support vector machine (SVM) model was applied to the obtained dataset, and a high-resolution classification map was then produced from the aerial hyperspectral images. The results clearly showed that a large amount of misclassification occurred in shaded areas due to the difference in spectral reflectance between sunlit and shaded areas. It was also found that the
classification accuracy was drastically improved by training the SVM model with both sunlit and shaded spectral data. As a result, we achieved a classification accuracy of 94.5% in vegetated areas.
Keywords

- Liquid crystal tunable filter
- Unmanned aerial vehicle
- Vegetation classification
- Machine learning
1. Introduction

Spectral reflectance data collected from vegetated areas can provide very valuable information on factors such as the presence or absence of certain tree species, plant growth stages, and plant diseases. The relationships between plant properties and spectral reflectance data have been studied using different types of spectroscopic sensors (e.g., Schmidt and Skidmore, 2003; Kuska et al., 2015). Recently, unmanned aerial vehicle (UAV) technology, which is evolving rapidly, has been applied to survey systems for precision agriculture, and UAVs equipped with spectral sensors have been tested in agricultural applications by many studies. For example, Garcia-Ruiz et al. (2013) conducted a UAV-based aerial survey with multispectral cameras over a citrus orchard, and found that UAV-based datasets yielded better classification accuracy than aircraft-based datasets for diseased citrus trees. Moreover, Peña et al. (2013, 2015) applied UAV-based multispectral imaging for the early detection of weed seedlings in combination with object-based image analysis.

Previous UAV-based spectral imaging systems mostly used multiple imaging sensors with independent optics and different band pass filters to obtain simple vegetation indices such as the normalized difference vegetation index (NDVI).
However, because the number of mountable imaging sensors is limited by the payload weight capacity of the UAV, it is difficult to obtain in-depth spectral information using this type of multispectral imaging system. Hence, we adopted a liquid crystal tunable filter (LCTF) for a UAV-based hyperspectral imaging system. The LCTF is an optical band pass filter whose center wavelength is electrically tunable. Because of the flexibility of spectral bands, hyperspectral imaging with LCTF technology is applicable to a wide variety of remote-sensing applications. This technology was first put into practical use on a space-borne instrument by Hokkaido University, and it has already been mounted on several microsatellites developed by Tohoku University and Hokkaido University (e.g., Sakamoto et al., 2015). More recently, the Philippines’ first microsatellite, DIWATA-1, was built by scientists and engineers from the Philippines, Hokkaido University, and Tohoku University under the PHL-Microsat Program. DIWATA-1 contains a space-borne multispectral imager (SMI) that employs LCTF technology. For calibration and validation of SMI imagery, a field campaign with the UAV-based hyperspectral imaging system using LCTF technology was conducted in the Philippines.

Advanced spectral imaging systems on UAVs can achieve highly efficient data
acquisition. However, large amounts of complex vegetation information are gathered on a daily basis, and this data can be difficult to process using conventional processing techniques. Recently, machine learning (ML) has become one of the most powerful approaches for examining such complex datasets (so-called big data). In general, ML is a data analysis method that can be used to discover underlying structures, similarities, or dissimilarities present in big data. In the case of supervised learning, the ML model is trained by user inputs so that it gains experience throughout the training process. The ML is widely applicable for identification, detection, classification, quantification, estimation, and prediction in precision agriculture (Singh et al., 2016).

The UAV-based hyperspectral imaging system described in this study can realize aerial images with a resolution on the order of tens of millimeters. Thus, it could be useful for leaf-scale plant disease detection when combined with the ML approach. Therefore, this new survey platform using LCTF technology will make a significant contribution to future precision agriculture research. In this paper, we present a UAV-based high-resolution vegetation classification map, and evaluate the validity of this new platform.
2. Observations

The field campaign was conducted in Gerona and Ramos, Tarlac, the Philippines, during March 28–April 3, 2016. In this campaign, UAV-based hyperspectral imaging was carried out in several vegetated areas under clear sky conditions. Each operation was conducted for 10 min because the maximum flight time of the UAV is only approximately 20 min due to the payload weight. Figure 1(a) shows the hyperspectral imager with LCTF technology (hereafter referred to as the LCTF imager) and related devices used in the field campaign. The LCTF imager can capture two-dimensional hyperspectral images ranging from 460 to 780 nm at 1 nm intervals (321 bands in total), and the full width at the half maximum of the spectral band ranges from 5 nm (at 460 nm) to 25 nm (at 780 nm). The maximum switching time of the spectral bands is 300 ms, and the exposure time of the imager typically ranges from 5 to 50 ms, depending on the reflectance of the ground object and the solar zenith angle. The capture time of a single image is approximately 1 s, which includes the switching time, exposure time, and data processing time. Figure 1(b) shows the rotating
wing UAV (DJI’s Matrice 600 Hexacopter) with the LCTF imager. In this field campaign, LCTF-based hyperspectral images of ground objects were recurrently captured from 460 to 780 nm at 10 nm intervals, which equates to 33 images per cycle. Approximately 2,000 images were acquired in all. The flight height was set to 25 m or 50 m in order to capture the target objects at either leaf- or

Figure 1. (a) Liquid crystal tunable filter (LCTF) imager, (b) unmanned aerial vehicle (UAV) with the LCTF imager, (c) mango orchard, and (d) ethylene vinyl acetate (EVA) mat.
canopy-scales. The image resolution was 656 pixels × 494 pixels, and the expected ground sampling distance (GSD) was approximately 57 mm at a height of 25 m, and 115 mm at a height of 50 m. The study area was a mango orchard located in Gerona City (15°34′N, 120°32′E), which mainly contains mango trees, dried grass, and soil (see Figure 1(c)). In the field campaign, we placed white sacks on the ground as landmarks. In proximal sensing, diffuse reflectance standards are often used as the reference for solar irradiance to obtain the spectral reflectance of a ground object. As the size of typical reflectance standards is small compared with the resolution of the imager, we utilized a mat made of ethylene vinyl acetate (EVA) as a substitute for the reflectance standards in this field campaign (see Figure 1(d)). The deployed EVA mat measured 2.4 m × 2.4 m, and its reflectance was independently measured in advance. Using the reflectance of the EVA mat, the spectral reflectance of a target object on the same image at a wavelength of $\lambda$, $R_{obj(\lambda)}$, can be given by the following equation:

$$R_{obj(\lambda)} = \frac{L_{obj(\lambda)}}{L_{eva(\lambda)}} \cdot R_{eva(\lambda)}$$

where $L_{obj(\lambda)}$ and $L_{eva(\lambda)}$ are the measured spectral radiance of the target object and the EVA mat, respectively, and $R_{eva(\lambda)}$ is the reflectance of the EVA mat.
mat. Note that the spectral reflectances measured by the LCTF imager in this study were calculated using this equation.

3. Data processing

Spectroscopic imaging using a conventional hyperspectral sensor employs push-broom line scanning; therefore, the technique is generally combined with a moving platform such as an aircraft or an Earth-observation satellite (e.g., Shaw and Burke, 2003). Hence, accurate GPS measurements and complicated post data processing methods such as georeferencing are necessary to construct so-called spectral data cubes (e.g., Suomalainen et al., 2014). This remains difficult for most users when handling hyperspectral data; thus, simpler approaches should be developed to facilitate the use of hyperspectral data. To address this issue, we developed a snapshot hyperspectral imaging system that does not require accurate global positioning system (GPS) measurements so that the acquired dataset can be processed by a simple image processing technique, detailed below.

In the case of aerial snapshot hyperspectral imaging by the LCTF imager, the captured area shifts slightly from image to image, which is due to small attitude
perturbations of the UAV. In order to obtain a spectral data cube, it is necessary to precisely overlap the time-sequential hyperspectral images using an image processing technique. In this study, we applied Speeded-Up Robust Features (SURF), which is an algorithm that extracts so-called SURF descriptors (in other words, unique features) from images (e.g., Bay et al., 2008). The SURF descriptors are in the form of 64-element vectors that express the local gradation of images. Thus, the features of an image are assumed to be reflected in the SURF descriptors. The similarity of images is quantified by comparing SURF descriptors. Figure 2 shows an example of feature matching results.

Figure 2. Example of feature matching results for two time-sequential hyperspectral images at (a) 680 nm and (b) 690 nm. The color-coded circles indicate the extracted features. Similar features are associated with color-coded lines between the images.
applied to two time-sequential hyperspectral images taken at (a) 680 nm and (b) 690 nm. The color-coded circles indicate the extracted SURF descriptors, and similar features are associated with color-coded lines between the images. Based on the associated pixel locations, a matrix is produced for geometric

Figure 3. Structure of the spectral data cube. In the field campaign, the LCTF imager recurrently captured spectral images ranging from 460 to 780 nm at 10 nm intervals, which equates to 33 images per cycle.
image transformation from one image to another. The area surrounded by green lines in Figure 2(b) is the overlapping area calculated by the transformation matrix. In total, 32 repetitions of the overlapping process from the 460-nm image to the 780-nm image were used to produce a spectral data cube, as illustrated in Figure 3. In the spectral data cube, the spectral reflectance of a captured object can be obtained from each pixel location along the z-axis. With this image processing technique, the spectral reflectances of 14 different ground objects were extracted from the aerial hyperspectral images. Figure 4 shows the mean spectral reflectance of the captured objects. The numbers in parentheses indicate the number of extracted spectral reflectances used for each mean spectrum. The error bars indicate the standard deviation. The figure omits the error bar of the “man-made objects” category for the sake of a better visualization. This is because the mean spectral reflectance of the man-made objects was produced by averaging the spectral reflectances of different types of objects such as rooftops and sacks, and the standard deviation was very large (up to 0.24) compared to the others. It is well known that the spectral reflectance in a shaded area is quite different from that in a sunlit area (e.g., Zhang et al., 2015; Hsieh et al., 2016); therefore, we chose only sunlit areas for
Figure 4. (a) Mean spectral reflectance in the sunlit areas. (b) Comparison of the mean spectral reflectance between sunlit and shaded areas. The numbers in parentheses indicate the number of extracted spectral reflectances used for each mean spectrum. The error bars indicate the standard deviation. Note that the error bar of the “man-made objects” category is not plotted in the figure.

extracting the mean spectra in Figure 4(a).
4. Results and discussion

Using a support vector machine (SVM), which is one type of supervised ML model, we classified the spectral reflectances obtained by the image processing technique described in Section 3, and created a high-resolution classification map of the study area. Figure 5(a) shows a false color image of the study area observed by the LCTF imager, and consists of hyperspectral images taken at 550 nm, 650 nm, and 780 nm, which represent green, red, and near-infrared channels, respectively. These three bands were selected to emphasize the vegetation in red color. Images were taken at an altitude of 50 m at 15:23 PHT on March 29, 2016, when the sun was located in the west with a solar zenith angle of 50.69°. Therefore, the ground under the mango trees was largely

Figure 5. (a) False color image of the mango orchard captured by the UAV-mounted LCTF imager; (b) manual classification result.
shaded. Based on this figure, we manually classified each individual pixel (see Figure 5(b)). Although the manual classification result could contain some mistakes, the ground objects were categorized into the following approximate object categories: (1) mango canopy, (2) dried grass and soil, (3) shaded dried grass and soil, (4) man-made objects, and (5) the EVA mat. For comparison with the classification results obtained by the SVM model, we used Figure 5(b) to elucidate the correct classification according to the discussion below.

Figure 6(a) shows the classification map produced by the SVM model, which was trained with the 420 sunlit spectral reflectance data presented in Figure 4(a). During the training phase, we used 33-bands of each spectral reflectance as 33-dimensional information, and mapped the data onto a 33-dimensional hyperplane using the SVM model. The shape of the decision boundary, which is the determining factor of a classification, can be mapped by the SVM model based on the selected kernel type, the kernel parameters, and the training dataset. Therefore, a proper setting and training dataset should be selected to classify the data at a certain high level (e.g., Huang et al., 2002). In this study, the radial basis function (RBF) kernel was selected, and the cost parameter $C$ and the gamma $\gamma$ of the RBF kernel were tuned to 312 and 0.5, respectively.
order to minimize the cross-validation estimate of the test set error.

As mentioned above, Figure 5(b) shows the results of a manually defined classification for comparison with the classification map produced by the SVM. Accordingly, each classified pixel in Figure 6(a) was compared with the defined category in Figure 5(b), and the classification accuracy was calculated by dividing the number of correctly classified pixels by the total number of pixels. As
a result, the classification accuracy was 81.0%, owing to the large amount of misclassification in shaded areas. Image pixels located in the shaded mango canopy and shaded dried grass and soil areas were misclassified as weeds and man-made objects, respectively. Thus, classification accuracy is highly dependent on the coverage of shaded areas in cases where the spectral images have a high resolution, and where sunlit and shaded areas can be distinguished by the data. Early studies discriminated shaded pixels from non-shaded pixels by using shadow masking, and the pixel data were processed separately to achieve high classification accuracy (e.g., Roussel et al., 2016; Qiao et al., 2017).

In this study, however, shaded and non-shaded pixel data were processed by a common data processing method without shadow masking. In Figure 4(b), the spectral reflectances of the mango canopy and dried grass and soil in sunlit and shaded areas are compared. There was a clear difference between spectral reflectances of the mango canopy and dried grass and soil in shaded areas, as well as between those in sunlit areas. This means that sunlit and shaded spectral reflectances can be classified together in a single process. Therefore, we modified the SVM model by training it simultaneously with both sunlit and shaded spectral data. Figure 6(b) shows the classification map produced by the
modified training dataset. The classification accuracy increased to 94.5%; that is, it was improved by 13.5% from that of Figure 6(a). Thus, this result clearly indicates the importance of using shaded spectra for the training process of the SVM model.

It should be noted that there were still some misclassifications around the boundaries between different objects. Because the structure of flexible objects such as mango canopies is continuously slightly deformed by the wind, its appearance could be different in each image. This small change in appearance leads to feature matching errors; consequently, some non-overlapping regions are created between the transformed images. Hence, the challenge for future research will be to minimize such non-overlapping regions by reducing the feature matching errors. The LCTF imager at present takes ~1 s to acquire a single image, but acquisition time can be shortened by approximately half by upgrading the current system. Therefore, it is expected that misclassifications derived from feature matching errors can be reduced in the near future.

5. Conclusion

Whereas most hyperspectral sensors adopt push-broom line scanning, the
LCTF imager utilizes a snapshot imaging system. This advantage allows users to handle the hyperspectral data more easily and obtain science products more rapidly using a simple image processing technique. With this method, we were able to precisely overlap time-sequential hyperspectral images through image processing using SURF descriptors, and the spectral reflectances of 14 different ground objects were successfully extracted. Additionally, the SVM model trained with spectral reflectances was used for vegetation classification, and a high-resolution classification map produced from the aerial hyperspectral images was presented.

Originally, the SVM model trained with only sunlit spectral data was applied to the mango orchard, but the resulting classification map contained a large number of misclassifications in shaded areas. This was because spectral reflectances in shaded areas deviated from those in sunlit areas. Thus, we trained the SVM model using both sunlit and shaded spectral reflectances, and applied the updated model to the study area. As a result, we achieved a classification accuracy of 94.5%, which represents a 13.5% improvement on results produced using only sunlit spectral reflectance data. Thus, the results clearly indicate that this new survey platform using the LCTF imager is useful for
vegetation classification, and has the potential to make valuable contributions to precision agriculture. Nevertheless, a statistical study is necessary in future work to verify the robustness of the proposed method, because the classification accuracy was estimated from a single event in this study. The results also revealed some misclassification areas around the boundaries between different objects. This was caused by feature matching errors that resulted in the creation of some non-overlapping regions between transformed images. The classification accuracy can be further improved by decreasing the feature matching errors in future research.

Acknowledgements

The Philippine field campaign was conducted as part of the PHL-Microsat Program, which was financially supported by the Philippine Council for Industry, Energy and Emerging Technology Research and Development of the Department of Science and Technology (PCIEERD-DOST). The PHL-Microsat Program is a Philippine government program that aims to build, launch, and utilize microsatellite technology for disaster monitoring and natural resource management. The program was built by PCIEERD-DOST, University of the
Philippines Diliman, Hokkaido University, and Tohoku University. The authors are indebted to Phil-Lidar 1 and PHL-Microsat Projects 4 and 5 team members for providing valuable assistance during this field campaign. Open Source Computer Vision Library (OpenCV), which is an open source computer vision and machine learning software library, was utilized in this study.

References

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-Up Robust Features (SURF). Comput. Vis. and Image Underst. 110 (3), 346–359.

Garcia-Ruiz, F., Sankaran, S., Maja, J.M., Lee, W.S., Rasmussen, J., Ehsani, R., 2013. Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees. Comput. Electron. Agric. 91, 106–115.

Hsieh, Y.T., Wu, S.T., Chen, C.T., Chen, J.C., 2016. Analyzing spectral characteristics of shadow area from ASD-40 high radiometric resolution aerial images. Int. Archives Photogramm., Remote Sens. Spatial Inform. Sci. XLI-B7, 223–227.

Huang, C., Davis, L.S., Townshend, J.R., 2002. An assessment of support vector machines for land cover classification. Int. J. Remote Sens. 23 (4), 725–749.
Kuska, M., Wahabzada, M., Leucker, M., Dehne, H.W., Kersting, K., Oerke, E.C., Steiner, U., Mahlein, A.K., 2015. Hyperspectral phenotyping on microscopic scale: towards automated characterization of plant-pathogen interactions. Plant Methods 11 (28), 1-14.

Peña, J.M., Torres-Sánchez, J., de Castro, A.L., Kelly, M., López-Granados, F., 2013. Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images. PLoS ONE 8 (10), 1–11.

Peña, J.M., Torres-Sánchez, J., Serrano-Pérez, A., de Castro, A.L., Kelly, M., López-Granados, F., 2015. Quantifying efficacy and limits of unmanned aerial vehicle (UAV) technology for weed seedling detection as affected by sensor resolution. Sensors 15 (3), 5609–5626.

Qiao, X., Yuan, D., Li, H., 2017. Urban shadow detection and classification using hyperspectral image. Journal of the Indian Society of Remote Sens. 1–8, https://link.springer.com/article/10.1007/s12524-016-0649-3.

Roussel, G., Ceamanos, X., Briottet, X., 2016. A sun/shadow approach for the classification of hyperspectral data. 16th Onera-DLR Aerospace Symposium (ODAS 2016), 1-7, https://hal.archives-ouvertes.fr/hal-01385498.

Sakamoto, Y., Sugimura, N., Fukuda, K., Kuwahara, T., Yoshida, K., Kurihara, J.,
Fukuhara, T., Takahashi, Y., 2015. Development and flight results of microsatellite bus system for RISING-2. Trans. Jpn. Soc. Aeronaut. Space Sci., Aerospace Technol. Jpn. 14 (30), 89–96.

Schmidt, K.S., Skidmore, A.K., 2003. Spectral discrimination of vegetation types in a coastal wetland. Remote Sensing of Environment 85 (1), 92–108.

Shaw, G.A., Burke, H.K., 2003. Spectral imaging for remote sensing. Lincoln Lab. J. 14 (1), 3–28.

Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S., 2016. Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 21 (2), 110–124.

Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., Kooistra, L., 2014. A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. Remote Sens. 6 (11), 11013–11030.

Zhang, L., Xuejian, S., Wu, T., Zhang, H., 2015. An analysis of shadow effects on spectral vegetation indexes using a ground-based imaging spectrometer. IEEE Geosci. Remote Sens. Lett. 12 (11), 2188–2192.
SVM training data

Sunlit spectra (420)
- Calamansi (20)
- EVA mat (10)
- Man-made objects (20)
- Sugarcane (10)
  - Cashew (55)
  - Ash (40)
  - Monggo (45)
  - Weed (30)
  - Dried grass and soil (20)
  - Mahogany (35)
  - Onion and ash (20)
  - Eggplant (20)
  - Mango canopy (55)
  - Ricefield (40)

Shaded spectra (40)
- Dried grass and soil (20)
- Mango canopy (20)