Tree species classification based on 3D spectral point clouds and orthomosaics acquired by snapshot hyperspectral UAS sensor

C. Iseli[1] and A. Lucieer[1]*

[1] TerraLuma research group, Discipline of Geography and Spatial Sciences, School of Technology, Environment and Design, University of Tasmania, Australia – Chris.Iseli@utas.edu.au; Arko.Lucieer@utas.edu.au

KEY WORDS: UAS, UAV, snapshot hyperspectral sensor, SfM, 3D point cloud, random forest classification, ecology

ABSTRACT:

In recent years, there has been a growing number of small hyperspectral sensors suitable for deployment on unmanned aerial systems (UAS. The introduction of the hyperspectral snapshot sensor provides interesting opportunities for acquisition of three-dimensional (3D) hyperspectral point clouds based on the structure-from-motion (SfM) workflow. In this study, we describe the integration of a 25-band hyperspectral snapshot sensor (PhotonFocus camera with IMEC 600 – 875 nm 5x5 mosaic chip) on a multi-rotor UAS. The sensor was integrated with a dual frequency GNSS receiver for accurate time synchronisation and geolocation. We describe the sensor calibration workflow, including dark current and flat field characterisation. An SfM workflow was implemented to derive hyperspectral 3D point clouds and orthomosaics from overlapping frames. On-board GNSS coordinates for each hyperspectral frame assisted in the SfM process and allowed for accurate direct georeferencing (<10 cm absolute accuracy). We present the processing workflow to generate seamless hyperspectral orthomosaics from hundreds of raw images. Spectral reference panels and in-field spectral measurements were used to calibrate and validate the spectral signatures. This process provides a novel data type which contains both 3D, geometric structure and detailed spectral information in a single format. First, to determine the potential improvements that such a format could provide, the core aim of this study was to compare the use of 3D hyperspectral point clouds to conventional hyperspectral imagery in the classification of two Eucalyptus tree species found in Tasmania, Australia. The IMEC SM5x5 hyperspectral snapshot sensor was flown over a small native plantation plot, consisting of a mix of the *Eucalyptus pauciflora* and *E. tenuiramis* species. High overlap hyperspectral imagery was captured and then processed using SfM algorithms to generate both a hyperspectral orthomosaic and a dense hyperspectral point cloud. Additionally, to ensure the optimum spectral quality of the data, the characteristics of the hyperspectral snapshot imaging sensor were analysed utilising measurements captured in a laboratory environment. To coincide with the generated hyperspectral point cloud data, both a file format and additional processing and visualisation software were developed to provide the necessary tools for a complete classification workflow. Results based on the classification of the *E. pauciflora* and *E. tenuiramis* species revealed that the hyperspectral point cloud produced an increased classification accuracy over conventional hyperspectral imagery based on random forest classification. This was represented by an increase in classification accuracy from 67.2% to 73.8%. It was found that even when applied separately, the geometric and spectral feature sets from the point cloud both provided increased classification accuracy over the hyperspectral imagery.

1. INTRODUCTION

Recent advances in sensor technologies have yielded a new breed of hyperspectral snapshot imaging sensors which allow the capture of full-frame, hyperspectral images in a single exposure (Geelen et al. 2015). Up to this point, compact hyperspectral imaging systems have largely been restricted to capturing only two of the three dimensions of a hyperspectral image in a single exposure, requiring an additional scanning process to complete the third dimension (Hagen and Kudenov 2013). The snapshot imaging sensors circumvent the scanning process, by extending the technology used in many modern-day digital cameras. This is characterised by the use of a multispectral filter array (MSFA), in which a mosaic of wavelength specific filters are arranged on the image sensor. This means while each pixel is receptive to only a single band, the repeated arrangement of filters across the sensor allows a full hyperspectral image to be derived via a demosaicking process (Hagen and Kudenov 2013). A key implication of the ability to capture full-frame hyperspectral images in this way, is they retain the geometric constraints of standard optical imagery. This opens the possibility to apply conventional photogrammetric and structure from motion (SfM) principles to images with overlapping extent, in order to derive three dimensional (3D) structural information in the form of point clouds (Aasen et al. 2015). Since the point cloud is directly derived from the hyperspectral imagery, each of the points in space can also be attributed with the hyperspectral information relating to that point. This produces a rich new data source, which effectively combines the desirable aspects of both passive optical and LiDAR point clouds, and has potential to lead to more robust classification methodologies.

Whilst this advancement in sensor technology presents exciting opportunities, there are a number of problematic factors that must also be overcome. Due to the system employed to capture the hyperspectral information, an inherent limitation lies in the fact that effective spatial resolution of each band is inversely proportional to the total number of distinct bands. For example in the case of the 25 band IMEC SM5x5-NIR (SM5x5) image sensor (Geelen et al. 2015) utilised in this study, the full 2048x1088 pixel resolution of the sensor is reduced to an effective resolution of 409x217 pixels for each spectral band. This adversely impacts the level of spatial detail that can be captured, and consequently the range of flying heights and size of regions that a sensor such as this can be efficiently deployed (Yang et al. 2017). Aside from the operational limits, the significant spectral undersampling of the sensor also presents substantial implications in terms of potential spectral accuracy. Since the hyperspectral measurements must be

---

* Corresponding author

reconstructed from a number of independent pixels, each band is sampled from a slightly different location. In environments with highly variable surfaces, this has potential to produce inaccurate spectral readings in which the measurements of individual bands relate to different target objects. Additionally, in the context of the SM5x5 specifically, a degree of filter cross-talk has also been observed (Pichette et al. 2017). This is characterised by the fact that the pixels relating to each band, are not solely responsive to the primary band wavelength.

To coincide with the ability to produce hyperspectral point clouds, a file format and software is also essential for the efficient storage and processing of this information. Many current point cloud formats are designed for LiDAR based data, which are generally well defined with a discrete set of predefined attribute types. The LAS format (American Society for Photogrammetry and Remote Sensing 2013) is one of the most common industry formats, developed by the American Society for Photogrammetry and Remote Sensing (ASPRS) to provide a standardised file format for the storage and exchange of LiDAR data. Whilst it does allow for the storage of spectral attributes, this is currently restricted to only four distinct bands. While ASCII based formats provide a greater level of flexibility, they are inherently inefficient in both processing and storage. A number of open-source and proprietary formats are available, which allow the addition of an arbitrary number of attributes to each point. However, the software and processing tools associated with them are found to be limited for this use case.

The aim of this study is to determine if the rich source of information present within a hyperspectral point cloud can improve the robustness and accuracy of tree species classification, specifically in comparison to standard hyperspectral imagery and geometric point clouds in isolation. In pursuit of this, classification will be targeted at the discrimination of two Eucalyptus species, *E. pauciflora* and *E. tenuiramis*, on a native reforestation plot.

2. METHODS

2.1 Data Collection

All imagery was captured utilising the IMEC CMV2K-SM5x5-NIR snapshot mosaic image sensor deployed in the Photonfocus MV1-D2048x1088-HS02-G2 camera. The image sensor is capable of capturing 25 individual spectral bands, by employing a 5x5 square non-redundant MSFA pattern which is repeated across the image sensor. The camera is externally fitted with a combination of a 595 nm high-pass and 874 nm low-pass filters to restrict the spectral region of transmittance.

For the acquisition of aerial imagery, the camera system is mounted on the DJI S-1000 multi-rotor via the 2-axis gimbal to minimise the effect of flight dynamics on the image acquisition process (Figure 1). On board positioning of the UAV was measured utilising the Trimble BD920-W3G global navigation satellite system (GNSS) receiver module. The camera system was set to trigger the GNSS receiver to log the precise time of each camera exposure, allowing each image to be geotagged in post-processing. A Leica 1200 base station was positioned approximately 300 m from the study plot, to allow a short baseline differential solution to be derived for the Trimble measurements.

A custom hyperspectral point cloud format as well as numerous visualisation and processing modules were developed for the purpose of this study (Figure 2).

Figure 1. The DJI S-1000 multi-rotor flight platform, with downward facing PhotonFocus camera attached to the 2-axis gimbal and Trimble GNSS receiver mounted on top.
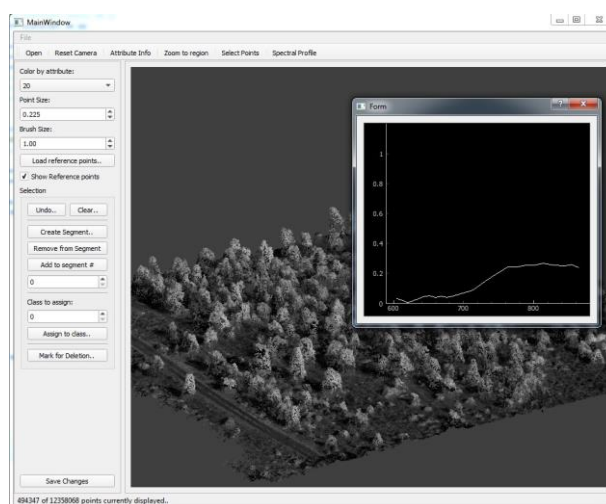

Figure 2. Screenshot of the developed HyPC visualisation software.

2.2 Tree species classification

2.2.1 Study site and data collection

In this study, we collected data from a native restoration plantation located at Dungrove, approximately 13 km NW of the township of Bothwell, Tasmania, Australia. Established in 2010 through a partnership between Greening Australia, the Tasmanian Government and the University of Tasmania, the site consists of plantings of a number of native species on degraded farmland. The specific plot captured in this study covers approximately 2 ha and primarily contains a mixture of *Eucalyptus pauciflora* and *Eucalyptus tenuiramis* plantings, with a number of understory species also present.

Imagery was captured based on two predetermined flight paths, consisting of a series of transects flown in two perpendicular directions forming a grid pattern. Images were captured at an approximate flying height of 50 m above ground level (AGL) resulting in an average ground sampling distance (GSD) of 1.6 cm per pixel and 8 cm for each instance of the filter array. The Photonfocus imagery was automatically captured at a rate of 4 Hz, and the resulting data set was subset by one third to reduce storage and processing demand. Imagery was captured with a side overlap of 70% and a forward overlap equating to approximately 70% after subsetting the captured imagery.

A series of 22 photogrammetric targets were randomly placed throughout the extent of the study region to serve as ground control points (GCPs) to be used for georeferencing of the point cloud and geometric accuracy evaluation. The position of each GCP was precisely measured using the Leica 1200 rover with RTK connection to the Leica 1200 base station located approximately 300 m away. This resulted in reported mean positional accuracy of approximately 0.01 m.

In addition to the acquired field data, an existing dataset of leaf level hyperspectral species measurements was obtained from ongoing research work being undertaken in conjunction with the Dungrove project (Bailey 2013).

2.2.2 Image pre-processing

A number of sequential pre-processing steps were conducted to spectrally calibrate and prepare the imagery for the subsequent SfM workflow.

The initial step in processing was to remove the systematic component of noise from the raw imagery. Utilising the collection of dark reference imagery captured before take-off, the responses of each individual pixel is averaged to derive a representative mean dark current noise image. The white reference correction is performed to simultaneously remove the effects of vignetting and calibrate the raw DN values captured by the sensor. Similarly to the dark reference images, using the collected white reference images, the responses of each individual pixel were averaged to derive a representative mean white reference image.

A demosaicking process was then performed to reconstruct the full hyperspectral data cube from the raw 2D, sub-sampled spectral bands captured by the sensor. This was performed at both the spatial resolution of each band (408x217 pixels) as well as the full native resolution of the sensor (2045x1080). An inverse distance weighting (IDW) algorithm was utilised in both cases. In demosaicking at the lower spatial resolution, the central pixel of each mosaic instance for a given band is interpolated from the surrounding pixels of the same band. The interpolated central pixel values of each mosaic instance are then combined to form a single band image at the spatial resolution of the sparsely sampled spectra. Demosaicking at the full resolution of the sensor is performed using the same algorithm, however rather than interpolating only the central pixel of each mosaic instance, every pixel is interpolated.

2.2.3 SfM point cloud and orthomosaic generation

The Agisoft Photoscan Professional (v1.4.1) (Agisoft LLC 2018) software was used to generate both a point cloud and orthomosaic from the flight imagery. The SfM process comprises of two core stages. First distinctive points are identified and matched across multiple overlapping images. A bundle block adjustment is then performed on the matched points to estimate both the 3D position and orientation of each camera, as well the location of each point. This produces a 3D sparse point cloud representation of the scene. Based on this initial camera alignment stage, a much denser point cloud can then be derived using multi-view stereopsis (MVS) or depth mapping techniques.

Since the images were captured over two distinct flights with independent white reference, dark reference and camera integration times the radiometric calibration was performed separately for each flight. To facilitate this, SfM processing was initially performed with the images of each flight split into independent chunks. From the generated sparse point cloud, a

digital surface model (DSM) was produced with 0.05 m pixel resolution. This was then used to orthorectify the input images and generate an orthomosaic of the study region. The blending mode was set to average with a resolution of 0.05 m per pixel. The orthomosaic was then exported from Photoscan and transferred to the ENVI software, where the 5 spectral reference panels were identified. A region of interest (ROI) was defined for each of the reference panels, which in combination with the lab measured reflectance of each panel was used to perform an empirical line correction.

The coefficients derived from the empirical correction were then applied to the individual images of each flight, radiometrically calibrating them to absolute reflectance. These radiometrically calibrated images were then imported into a new project, upon which the camera positions from the previous alignment step were imported. The two flight chunks were then merged for further processing.

The photogrammetric targets which mark the accurately measured GCP's, were identified and pinpointed in the flight images to allow more accurate orientation of the study site in global space. Of the 22 ground control points 15 were selected to orient and constrain the point cloud. The remaining 7 were used for accuracy validation.

The dense point cloud was then generated utilising the radiometrically calibrated images from both flights. From the dense point cloud, obvious outlying and noise influenced points were removed from both below the ground level and above tree tops. The dense point cloud was then exported in the XYZ ASCII format, ensuring point colours (i.e. spectral reflectance values) were included. The final point cloud consisted of approximately 12 million points. Utilising the generated DSM, an orthomosaic was also produced. The blending mode was set to average, which uses the weighted average pixel value from individual photos to determine the pixel value in the orthomosaic (Agisoft LLC 2018).

2.2.4 Hyperspectral point cloud processing

To prepare the hyperspectral point cloud for classification, processing was completed utilising the developed HyPC processing modules. The point cloud was processed utilising the following steps in sequence:

1. CSF ground estimation
2. DTM estimation
3. Ground removal
4. Point height estimation
5. Crown segmentation

Ground estimation was performed by utilising the python implementation of the CSF algorithm. The CSF algorithm accepts several parameters which influence the characteristics of the surface estimation as well as the interpretation of ground points. The simulated cloth resolution was set to 0.5 m, rigidness to 2 and the threshold to 0.5 m.

After identification of ground points, a digital terrain model (DTM) was derived from the ground labelled points. The DTM was calculated at 1m resolution, with the height derived from the mean ground point height within each cell. All classified ground points were then removed from the point cloud such that only above ground points belonging to shrubs and tree crowns remained.

The height above ground level (AGL) was then calculated for each remaining point based on the DTM surface. All classified ground points were then removed from the point cloud such that

only above ground points belonging to shrubs and tree crowns remained.

2.2.5 Species classification

Classification was performed based on the discrimination of two Eucalyptus species, *E. pauciflora* and *E. tenuiramis*. The locations and extent of crowns for each species of interest were manually digitised from the generated orthomosaic, based on existing reference data from ongoing collaborative research from the University of Tasmania and Greening Australia (Bailey 2013). Crowns that were difficult to identify or determine the extent of were omitted. This resulted in 263 crowns in total, with 139 *E. pauciflora* and 124 *E. tenuiramis*.

Classification was performed utilising the Python based sklearn v.3.2.4.3.1 implementation of the random forest (RF) classifier. The classification result was found to be highly dependent on the selection of training/test splits. Therefore, a simple iterative data splitting approach with k=100 was utilised obtain a more reliable accuracy estimate. For each iteration, 70% of samples were utilised for training with the remaining 30% retained for validation. To maintain consistency across both the orthomosaic and point cloud data sets, identical parameters were used in each instance, along with identical training/test splits. Training samples were randomly selected in a stratified fashion to ensure equal representation of each class. The RF classifier was trained using 1000 trees and unlimited tree depth, as a high number of trees reduces the impact of random variations from changes in parameters (Turner et al. 2018). This provides more accurate out of bag error estimates during the training of the model.

Orthomosaic

For crown classification from the orthomosaic, crowns were delineated utilising the previously defined polygons. For each crown, two spectral feature sets were derived; the mean of the crown spectra and the mean of the 10 brightest spectra. The 10 brightest spectra from the crown are utilised due to the high degree of variation in crown spectra caused by effects such as shadowing.

Hyperspectral point cloud

Individual tree crowns were delineated and segmented utilising the previously defined polygons relating to each crown, with the appropriate species label assigned. Figure 5 shows the final segmented point cloud utilised for the classification process.
For the classification of crowns from the hyperspectral point cloud a number of spectrally and geometrically based feature sets were derived. The derived spectral metrics included identical metrics to those used in the orthomosaic classification. An additional spectral feature set was also calculated, utilising the geometric aspect of points.

Geometric features were chosen based on metrics identified in existing point cloud studies. Specifically, in a study based on tree species classification from photogrammetric point clouds, Nevalainen et al. (2017) found that density and diameter metrics from discrete layers within the crown to provide useful discriminatory information. For this reason, the points within each crown were split into layers based on 5% height intervals of the total crown height. For each of these layers, the point density and diameter were calculated. Classification was performed independently on 3 different sets of the above features: spectral only, geometric only and the combination of all features.

3. RESULTS

3.1.1 Orthophoto classification

RF classification of crowns based on a k=100 iterative data splitting approach produced an overall mean accuracy of 67.2%. Significant variation is observed in the individual classification results, with cross-validation accuracies ranging from 56.9% to 78.4%. This indicates classification accuracy is highly sensitive to the selection of training features. Inspection of the normalised confusion matrix derived from the sum of all iterations shows that *E. pauciflora* was classified with a marginally higher producer and user accuracy. This also corresponds to the slight imbalance of class sample sizes, which may bias the classifier to a degree.

3.1.2 Hyperspectral point cloud classification

Figure 3 shows the produced hyperspectral point cloud of the study site, visualised as a false colour composite of three hyperspectral bands. Figure 4 shows the mean and standard deviation of the spectra for each species, based on the mean spectra of each crown, mean of the brightest spectra of each crown as well as the additional mean spectra of the top 5% of the crown. Notably the brightest spectra also shows lower variance, even though the same number of 10 samples are used as in the orthomosaic. Finally, the spectra from the top layer of the crown is shown to provide a medium between the other spectral features sets, with higher separation and variance than the mean crown spectra, but less than the brightest spectra.
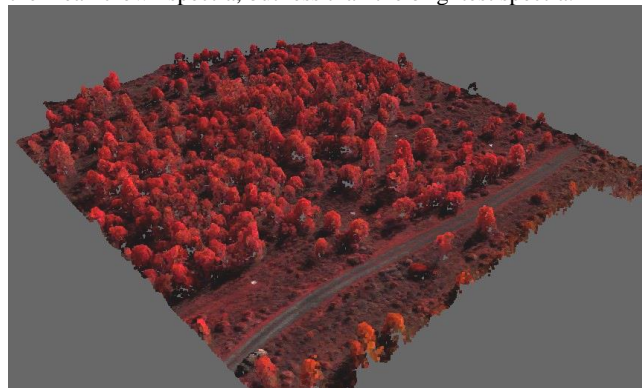

Figure 3. Hyperspectral point cloud of the study site visualised as RGB false colour composite comprising of bands centred at 802, 660 and 605 nm respectively.

Due to the large range of features derived from the hyperspectral point cloud, the relative importance of each feature set was initially analysed utilising the RF based mean decrease impurity metric. Looking at the importance of individual features within each feature set (Figure 6), spectrally we see that wavelengths in the region <750nm are most influential. Additionally, it can be seen that the most influential layer based features are predominantly from the layers of the upper canopy.

RF classification of crowns based the on the 20 most influential features produced a mean overall accuracy of 73.8% (Table 1), a marginal improvement over the orthomosaic result. Once again, the results of individual accuracies for each iteration vary considerably, ranging from 62.1 to 86.1%. Also the producer and user accuracies derived from the normalised confusion matrix are shown to be marginally higher for the *E. pauciflora* class.
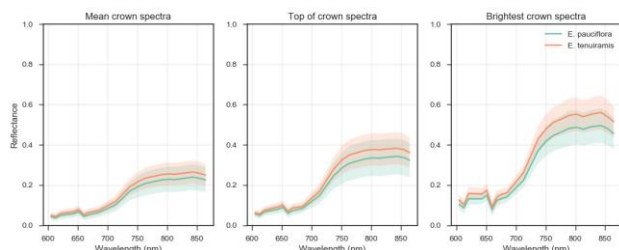
Figure 4. The hyperspectral point cloud based mean and standard deviation of the mean, top and brightest crown spectra for both *E. pauciflora* and *E. tenuiramis*

The classification of crowns was repeated, restricted to spectral only and structural only point cloud features to test the relative discriminatory power each type of feature can provide. Table 1 summarises the classification results based on each set of features as well as the combination of all feature sets. We see that although the spectral feature set contained the majority of the most important features, classification based on the spectral and geometric features in isolation still produced similar accuracies, with 71.9% and 71.1% respectively.

The sensitivity of the classification to the selection of training samples equates to individual classification scores ranging in excess of 10% of the mean classification accuracy. It is based on this metric that greater differentiation between spectral and geometric features is observed, with geometric features producing both reduced minimum and maximum accuracies.

Figure 5. The segmented crowns coloured by their designated class.

Table 1. Mean classification accuracy based on the spectral only, geometric only and combined feature sets of the hyperspectral point cloud

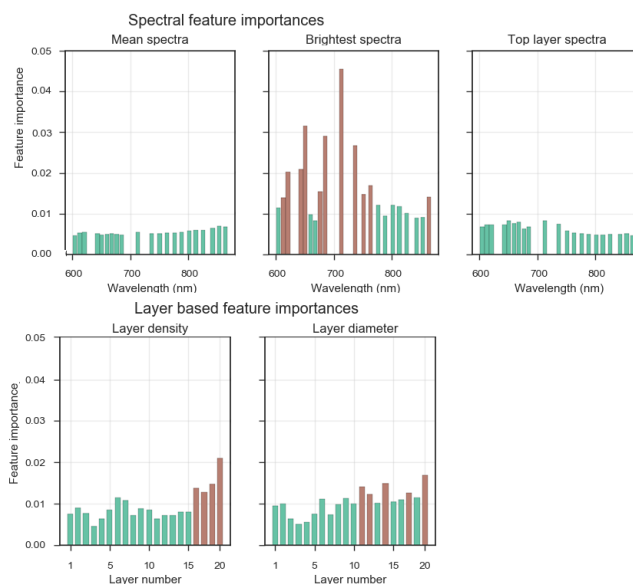| Feature set	Mean classification accuracy	Accuracy range (k=100)
Spectral only	71.9%	60.1-84.6%
Geometric only	71.1%	53.1-79.7%
All feature sets	73.8%	62.1-86.1%

Figure 6. Individual importance for each feature in all feature sets. The 20 most influential features are highlighted.

4. CONCLUSION

The development of UAS compatible hyperspectral snapshot imaging sensors presents the opportunity to produce information rich, 3D spectral point clouds that can provide potential advantages over conventional formats. This study investigated the tree species classification capabilities of a 3D hyperspectral point cloud compared to conventional hyperspectral imagery. This encompassed the characterisation and development of a spectral calibration and correction model for the IMEC SM5x5 hyperspectral snapshot sensor. A spectral point cloud format, visualisation and processing software was also developed to provide a complete workflow for classifying hyperspectral point cloud data.

Based on the classification of two Eucalyptus species, *E. paucilora* and *E. tenuiramis*, the hyperspectral orthomosaic and hyperspectral point cloud produced species classification accuracies of 67.2% and 73.8% respectively. Additionally, it was found in separating the geometric and spectral components of the point cloud and applying them independently, both produced improved accuracies over the orthomosaic.

From the characterisation of the SM5x5 sensor, it was found that a significant degree of random noise was present in the captured measurements. At its most extreme this equated to a SNR which equates to a noise level of 10% of the measured signal. Due to the techniques used for the spectral correction of the SM5x5, this level of noise was found to be substantially amplified after the spectral correction process.

The developed HyPC format implemented an efficient data structure for both the storage and querying of hyperspectral points, and was found to provide similar efficiency to an existing point cloud format of comparable specifications. The newly created visualisation software also provided a practical and intuitive environment for the exploration and manipulation of hyperspectral point clouds.

The analysis and results achieved in this study highlight a number of aspects which could benefit from future research. Of

great significance is the level of noise observed in the SM5x5 sensor measurements. The severity of this noise is such that post-processing solutions are expected to provide limited improvement in spectral quality in the future, and rather it is recommended that the physical source of noise requires identification. The developed software provides an intuitive and practical framework for the visualisation, manipulation and analysis of hyperspectral point cloud data, and is well placed to support future research in this area. It is however recommended that further optimisations are required in some cases to improve efficiency. The demonstrated improvements observed from the hyperspectral point cloud format indicate that it could provide strong potential in the remote sensing of forests in the future.

ACKNOWLEDGEMENTS

REFERENCES

Aasen, H., Burkart, A., Bolten, A., and Bareth, G. "Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 108 (2015), pp. 245–259.

Agisoft LLC. *Dense Cloud Classification & DTM Generation*. 2018. URL: http:// www.agisoft.com/index.php?id=35.

American Society for Photogrammetry and Remote Sensing. *LAS specification*. 2013. URL: https://www.asprs.org.

Bailey, T. "Plantings for carbon, biodiversity and restoration in dry rural landscapes". In: *Australian Forest Grower* 35.4 (2013), p. 39.

Geelen, B., Blanch, C., Gonzalez, P., Tack, N., and Lambrechts, A. "A tiny VIS-NIR snapshot multispectral camera". In: *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VIII*. Vol. 9374. International Society for Optics and Photonics, 2015, p. 937414.

Hagen, N. and Kudenov, M. W. "Review of snapshot spectral imaging technologies". In: *Optical Engineering* 52.9 (Sept. 2013), p. 090901.

Nevalainen, O., Hakala, T., Suomalainen, J., M¨akipa¨¨a, R., Peltoniemi, M., Krooks, A., and Kaasalainen, S. "Fast and nondestructive method for leaf level chlorophyll estimation using hyperspectral LiDAR". In: *Agricultural and Forest Meteorology* 198-199 (Nov. 2014), pp. 250–258.

Pichette, J., Goossens, T., Vunckx, K., and Lambrechts, A. "Hyperspectral calibration method For CMOS-based hyperspectral sensors". In: ed. by Y. G. Soskind and C. Olson. Vol. 10110. International Society for Optics and Photonics, Feb. 2017, 101100H.

Turner, D., Lucieer, A., Malenovsky´, Z., King, D., and

Robinson, S. A. "Assessment of Antarctic moss health from multi-sensor UAS imagery with Random Forest Mod- elling". In: *International journal of applied earth observation and geoinformation* 68 (2018), pp. 168–179.

Yang, G., Li, C., Wang, Y., Yuan, H., Feng, H., Xu, B., and Yang, X. "The DOM Generation and Precise Radiometric Calibration of a UAV-Mounted Miniature Snapshot Hyperspectral Imager". In: *Remote Sensing* 9.7 (June 2017), p. 642.