Application Note

Hyperspectral Spectroscopy for Wood Treatment Identification

Authors

Laura Ferlin
Kevin F. Harsh
Duc Nguyen
Sporian Microsystems Inc.

Introduction

Construction and demolition (C&D) waste materials constitute a significant waste stream and many materials retain significant value; thus, increasing the diversion of C&D materials through recycling and reuse is of high interest to government and commercial entities.

Wood is the third most used construction material in the United States, but there are limited options for reusing wood waste due to difficulty in quickly sorting treated versus untreated materials. About three quarters of waste wood is either not being recycled or is being recycled in a low-value use case due to sorting difficulties. Thus a method of identifying wood treatments is needed to improve the wood recycling outcomes.

Hyperspectral imaging (HSI) is emerging as a non-destructive, real-time detection tool for industrial and environmental sensing and inspection processes. This is due to its ability to obtain large amounts of spectral information on the objects being studied in real time. The technology has been applied to a broad range of sensing, agriculture, industrial, process inspection, food safety, mineralogy, astronomy, chemistry, environmental monitoring, biomedical, and surveillance applications [1,2,3,4] and has made possible the rapid spatial assessment of grade, defects, contamination, and constituent characterization of mixed, raw, and processed materials.

In this application note, we'll explore the use of a low-cost hyperspectral measurement system for the real-time accurate identification of wood treatment chemicals on pine samples. Initial system testing/demonstration using relevant materials focused on treated and untreated wood. A range of pine wood samples (Figure 1) with standard (known hazardous) treatment coatings (Table 1) were measured using the prototype hardware.

Experiment/Methods

White pine wood samples with a range of coatings/treatments (shown in Figure 1) were purchased from a local lumber company.

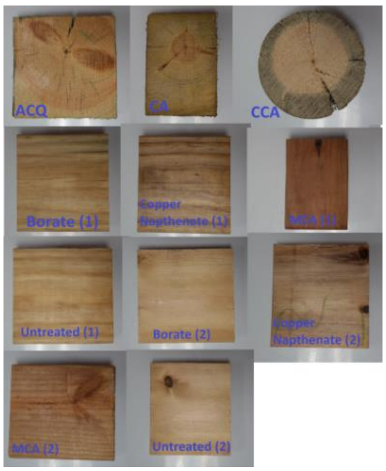

Figure 1: Example treated pine wood samples measured

Table 1: Wood treatments tested

Material	Note
ACQ - Alkaline Copper Quaternary	Copper fungicide with quaternary ammonium compound - preferred replacement for CCA
CCA - Chromated Copper Arsenate	Was widely used as fungicide/insecticide. Banned for most residential use in early 2000s. Leaches arsenic into soil
CA - Copper Azole	Similar to ACQ. Major copper based wood preservative in wide use in Canada, US, Europe, Japan & Australia
CN - Copper Naphthenate	Used since 1911, is registered with the EPA as a non-restricted use pesticide
MCA - Micronized Copper Azole	Particulate preservative technology. Uses an azole biocide/nano-particles of copper oxide or copper carbonate, for which there are safety concerns
Borate Disodium octaborate/ tetraborate	Water-based wood preservative registered by EPA and used throughout Asia, North America & Europe. Borate compounds can be leached out.
Uncoated wood	Uncoated pine wood

Data presented here was collected using a Sporian Microsystems SpecIQ™ BroadSpec Sensor System. The measurement system has the capability to provide real-time measurement of materials by collecting reflected or emitted light through a slit and breaking the light into spectral components for measurement. Whereas a standard camera only records three distinct colors visible to the human eye, hyperspectral imaging records potentially hundreds of bands of color in a range beyond what the human eye can detect. Key to the system's utility for materials identification is the very wide spectral range captured (400 to 2200 nm), ranging from ultra-violet to short wave infra-red, which allows for the collection of spectral information relating to pigmentation, chemical composition, and microstructure of the material imaged.

Key system specifications for the system used in these experiments are shown in Table 2, but Sporian builds and provides similar systems in a range of different formats and configurations for different user applications, including handheld units, lab instruments, and systems for deployment on unmanned aerial vehicles. The system used was a lab configuration, designed to operate connected to a computer, receiving camera settings via a user interface, and outputting data to the database file.

Table 2: Key SpecIQ™ BroadSpec specifications

Spec/Feature	Unit
Wavelength Range	400-2200 nm
Spectral Resolution	<10 UVNIR nm; <25 SWIR nm
Wavelength Reproducibility	<0.1 nm
Wavelength Accuracy	+/-0.2 nm
Instrument Dimensions	~4x4x3 inch

Using a prototype system of the Sporian Microsystems SpecIQ BroadSpec sensor, 20+ scans of each wood sample were taken varied across the surface of the piece. All measurements were done under a stable, broadband (UV to IR) light source built from an array of emitter types to cover the wavelength range, balance the light across the measurement spectra, and reflect off of a diffuse reflection coated dome to produce spatially even, directionally diffuse light onto the sample. Calibrated Spectralon® reflectance standards were used to monitor light source stability, but were not used as part of subsequent data processing and analysis.

A drawback of utilizing hyperspectral measurements is that the resulting data can contain very high dimensionality per measurement. One material measurement can contain hundreds of distinct features from the spectral information of the sample. Depending on the number of measurements and range of materials the dataset can become very large and complex, and analyzing the data such by a human analyst is challenging and time-consuming.

Machine learning (ML), a sub-category of artificial intelligence, is a data-driven approach to analysis where an algorithm is used to leverage relationships within the data to generate correlating models. ML is therefore suitable for a situation where the system outputs are unknown, hard to understand, or time-consuming to process by a human analyst. ML models can learn from spectral data, identify patterns, and be able to make decisions with minimal human intervention. The models can rapidly process large amounts of high-complexity spectral data to be able to classify unknown chemical compounds and/or determine composition.

All spectral data were processed and analyzed by performing a sequence of pre-processing, dimension reduction, and machine learning classification and/or regression. Preprocessing included the stabilization of signal fluctuations and the removal of background signal. Then, Linear Discriminant Analysis (LDA) [5] and Quadratic Discriminant Analysis (QDA) were performed to find the optimal variance in the data as well as transform the data and reduce its dimensionality. Using the optimized data, the data was then split into a training dataset and a smaller test dataset. A supervised ML model (via Scikit-learn) was implemented to train on the known dataset, analyzing their relationships (see below). Scikit-learn is an open source ML library that supports supervised and unsupervised learning. For simple identification of material a classification model was then used that produced a confidence interval (CI) in the classification of the material spectra.

Results

Figure 2 shows an example of the raw data output from the sensor, representing multiple measurements from all sample types. While this data is hard to distinguish and difficult to process with the human eye, the number of discrete spectral bands provides an abundance of key data that a computer-operated, ML-based algorithm is capable of discerning. Linear Discriminant Analysis (LDA) is one way in which computers can tease apart spectral data. LDA assumes each band of data is a distinct 'feature' of that data set. It then looks for linear combinations of these features that can group the data into a smaller-order multi-dimensional space.

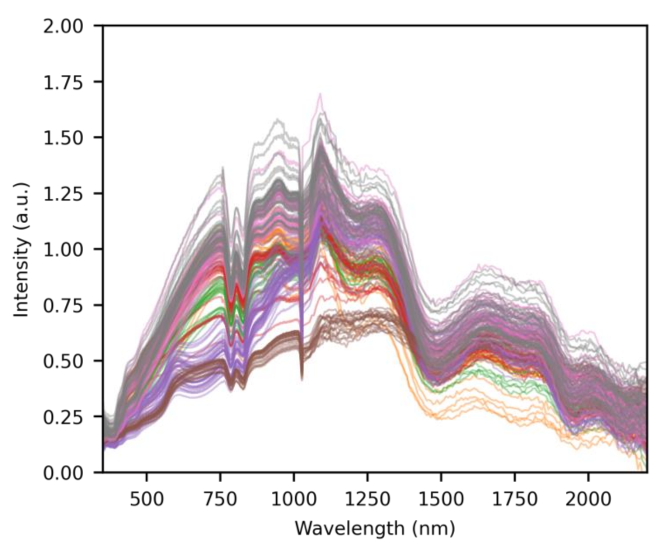

Figure 2 : Raw data from Sporian Microsystems SpecIQ BroadSpec Lab device

SPORIAN® MICROSYSTEMS, INC

In this space, the goal is to group data points with the same classifier – i.e. came from the same piece of wood – closely together and map each of these clusters as far from one another as possible. In this way each classification can be discriminated from one another by setting spatial boundaries around each classification set.

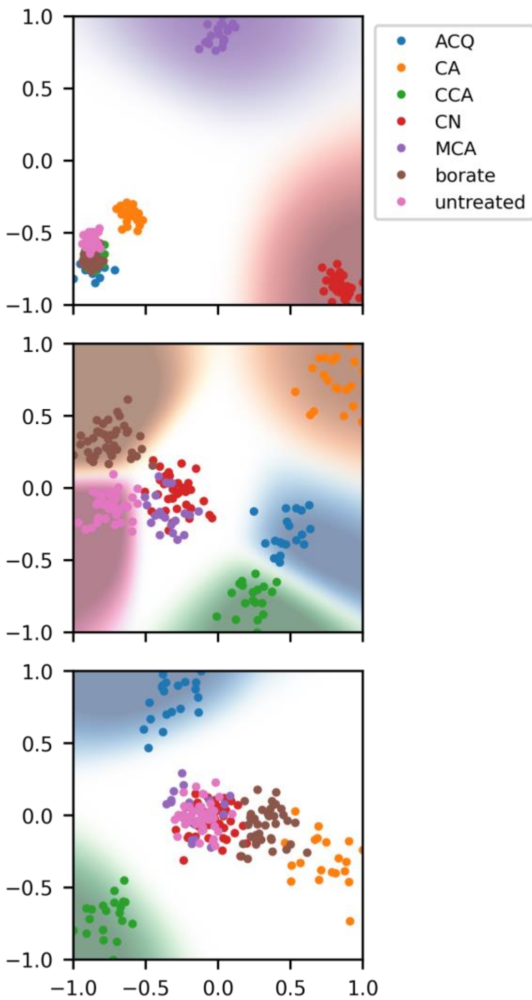

Figure 3 : LDA and QDA transformed data in 6-dimensional space showing discriminated treatment types

Figure 3 shows the same data sets after LDA methods have been applied to categorize the materials by the most meaningful data features. Each point on a plot represents an entire raw measurement scan. The measurements in the scan have been combined through LDA to form a six-dimensional vector, or an ordered collection of numbers. LDA resulted in a six-dimensional vector because the number of LDA components is equal to one less than the number of classes, which is seven here. The axes of each plot in Figure 3 are two of the LDA components, normalized to be between -1 and 1. The dots represent the transformed scan data. The data show a clear grouping by coating type regardless of orientation, grain, and scattering effects. The shading on the graph depicts the results of Quadratic Discriminant Analysis (QDA), a classification algorithm, on the well-separated groups in each plot.

Table 3 shows the confusion matrix for a six-dimensional QDA algorithm applied to the results of the LDA preprocessing. Each row of the table represents the instances in an actual class while each column represents the instances in a predicted class. The name stems from the fact that it makes it easy to see if the system is confusing two classes (i.e. commonly mislabeling one as another) [6]. For this classifier, we observe 100% accuracy on all classes except borate and untreated. Those classes are confused with each other 20% of the time.

Table 3: LDA-QDA Confusion Matrix. Values are percent of true samples in the predicted category.

True label \ Predicted label	ACQ	CA	CCA	CN	MCA	bor.	untr.
ACQ	1	0	0	0	0	0	0
CA	0	1	0	0	0	0	0
CCA	0	0	1	0	0	0	0
CN	0	0	0	1	0	0	0
MCA	0	0	0	0	1	0	0
bor.	0	0	0	0	0	0.8	0.2
untr.	0	0	0	0	0	0.2	0.8

Other algorithms can achieve 100% accuracy on these two groups, but are not very generalizable.

While these results are very good, we are only considering one species of wood, and up to two different wood samples per coating. A well-trained algorithm would need to consider multiple species of wood as well as multiple samples from each species. Exposure to the elements and other chemicals should also be considered, as most samples are not in pristine condition.

Conclusions

Hyperspectral imaging (HSI) is an emerging, non-destructive, real-time detection tool for industrial and environmental sensing and inspection processes. HSI is being applied to a broad range of industrial applications including agriculture, food processing, and food safety inspection. The Sporian Microsystems SpecIQ BroadSpec prototype hardware and the associated software algorithms were able to successfully classify each wood treatment type with 100% accuracy except for borate versus untreated pine wood. While this dataset is limited by the quality of training data and sample sets given to the computer, it shows that hyperspectral spectroscopy is capable of classifying wood treatments and could be further trained to work in more industrial environments.

1 Goetz, Alexander FH. "Imaging spectrometry for remote sensing: Vision to reality in 15 years." Imaging Spectrometry. Vol. 2480. International Society for Optics and Photonics, 1995.

2 Adão, Telmo, et al. "Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry." Remote Sensing 9.11 (2017): 1110.

3 Qin, Jianwei, et al. "Line-scan hyperspectral imaging techniques for food safety and quality applications." Applied Sciences 7.2 (2017): 125.

4 Bellon -Maurel, Véronique, and Nathalie Gorretta. "NIR Hyperspectral Imaging for Food and Agricultural Products." Infrared and Raman Spectroscopic Imaging: Second, Completely Revised and Updated Edition (2014): 295-338.

5 https://en.wikipedia.org/wiki/Linear_discriminant_analysis

6 Wikipedia contributors. "Confusion matrix." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 23 Sep. 2019. Web. 17 Oct. 2019.