Y-12

**OAK RIDGE
Y-12
PLANT**

LOCKHEED MARTIN

MANAGED BY
LOCKHEED MARTIN ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

UCN-13672 (26 6-95)

RECEIVED

NOV 20 1998

Y/DZ-2074

OSTI

**FEEDSTOCK AND WEB ANALYSIS USING MID-
INFRARED DIFFUSE REFLECTANCE
SPECTROSCOPY AND IMAGING
SPECTRORADIOMETRY**

J. E. Parks II

G. L. Powell

Oak Ridge Centers for Manufacturing Technologies

D. S. Blair

Sandia National Laboratories
Albuquerque, New Mexico

M. M. Szczesniak

M. Dombrowski

Surface Optics Corporation
San Diego, California

Date of Issue: September 15, 1997

Progress Report for submission to
Industry Representatives of:
Agenda 2020 Forest Products
Sensors and Controls Program Group

Prepared by

Oak Ridge Y-12 Plant
Oak Ridge, Tennessee 37831
managed by

Lockheed Martin Energy Systems, Inc.
for the

U. S. Department of Energy
under Contract No. DE-AC05-84OR21400

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Progress Report

Feedstock and Web Analysis Using Mid-Infrared Diffuse Reflectance Spectroscopy and Imaging Spectroradiometry

Dianna S. Blair^a
*Sandia National Laboratory**

Jim E. Parks II^b and G. Louis Powell
*Oak Ridge Centers for Manufacturing Technologies, Oak Ridge Y-12 Plant***

M. Martin Szczesniak and Mark Dombrowski
Surface Optics Corporation^c

^aMailing Address: D. S. Blair, Sandia National Lab, P. O. Box 5800, Dept. 6612/MS-0755, Albuquerque, NM 87185-0755; E-mail Address: dsblair@sandia.gov; Phone: (505)845-8800; Fax: (505)844-0968

^bMailing Address: J. E. Parks II, Oak Ridge National Lab, P. O. Box 2009, Oak Ridge, TN 37831-8087; E-mail Address: parksje@ornl.gov; Phone: (423)574-0928; Fax: (423)574-2102

^cMailing Address: Surface Optics Corporation, 11555 Rancho Bernardo Rd., San Diego, CA 92127

*Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-ACO4-94AL85000.

**Managed for the U. S. Department of Energy by Lockheed Martin Energy Systems, Inc. under Contract No. DE-AC05-84OR21400.

I. INTRODUCTION

Potential applications of mid-infrared (MIR) spectroscopy in the forest products industry include on-line analysis of feedstock and web materials; these applications differ dramatically in purpose, speed, and overall chemical heterogeneity. Characterization of feedstock will enable sorting of the stock and/or wet chemistry adjustment prior to the web stage of paper production. Sorting will require imaging of the stock as well as classification of the wide variety of chemistry found in recycled stock. At the opposite end of the manufacturing process, on-line analysis of the web will enable adjustment of machine parameters to maximize product quality and minimize waste. Spectroscopic requirements for web analysis include high-speed capability and measurement precision. If successful, both applications could result in a reduction of resource waste, a reduction of plant pollution, and a reduction of energy use while simultaneously improving product quality.

Here the progress towards feedstock and web analysis with MIR spectroscopy is presented. To date, work has progressed in three main areas: Diffuse Reflectance mid-Infrared Fourier Transform (DRIFT) spectroscopy of cellulose-based materials, chemometrics analysis, and research of MIR instrumentation for prototype development. The DRIFT spectroscopy data represents a database of the chemistries and spectroscopic signatures of interest to the applications discussed here. Over 50,000 spectra were obtained from cellulose-based materials infused with a wide variety of non-cellulose chemistry. Chemometrics analysis was performed on the DRIFT database to determine the quantitative and qualitative limits of the technique. Emphasis was placed on qualitative evaluation of spectroscopic signatures unique to the particular classes of cellulose-based material; thus, the degree to which classes could be sorted was determined. Finally, investigations of MIR instrumentation suitable for transfer of the technique from the lab-based instrument to a field ready prototype were made.

II. MIR SPECTROSCOPY OF CELLULOSE-BASED MATERIALS

Since mid-infrared (MIR) spectroscopy will be used for web and feedstock analysis, a database of MIR spectra has been obtained from a variety of materials that occur, by design or mishap, in the papermaking process. The MIR spectrum of cellulose differs dramatically from spectra of non-cellulose materials like plastic; the majority of the work presented here regards the more subtle differences in MIR spectra that indicate the presence of non-cellulose materials in cellulose-based materials. Identifying the spectral signatures of non-cellulose materials is particularly important in feedstock analysis and sorting since a large variety of chemical constituents will be present in the feedstock.

The MIR spectra have been obtained with Diffuse Reflectance mid-Infrared Fourier Transform (DRIFT) spectroscopy. DRIFT spectroscopy is sensitive to the chemical nature of the sample surface (with skin depths of $\sim 1\text{ }\mu\text{m}$), and cellulose-based materials are excellent samples for investigation in the diffuse reflectance geometry due to the surface structure of paper, cardboard, and wood. The spectral range of the work presented here is 4500 cm^{-1} to 500 cm^{-1} ; thus, fundamental vibration modes are detected and lead to characterization of the chemical nature of the sample. The materials investigated include lignin and other extractives in wood, ink, paper coatings, adhesives, and moisture. Spectral images showing chemical variation in the sample have been generated by obtaining DRIFT spectra from multiple positions on the sample.

EXPERIMENT DESIGN

Mid-infrared absorbance spectra were obtained with a Fourier transform interferometer-based device (Model SOC-400, Surface Optics Corporation, San Diego, CA 92127). The SOC-400 incorporates an ellipsoidal mirror for collection of the diffusely scattered light. The mirror extends a full 360° angle about the central axis of the ellipse; thus, a large amount of the diffusely scattered light is collected. The sample is positioned at one focal point of the ellipsoid, and the detector is positioned at the other focal point. Diffusely reflected light makes one reflection off of the mirror in route to the detector. The large collection solid angle and the single reflection surface combine to make an extremely efficient system for collecting diffusely scattered light. A diagram of the instrument is shown in **Figure 1**.

The sample was scanned in two dimensions beneath the SOC-400 with translation stages driven by stepper motors. During spectrum acquisition the sample remained still. Array basic programs written in conjunction with the GRAMS software package (GRAMS/32, Galactic Industries Corporation, Salem, NH 03079) controlled positioning of the motors and spectra collection.

No sample preparation was required. However, care was taken to avoid contamination of samples since the technique is sensitive to surface oils and other contaminants. Purge gas (argon) was emitted from the sample aperture of the instrument onto the sample during spectral acquisition.

Spectra in the range of 4500 cm^{-1} to 500 cm^{-1} were collected with 8 cm^{-1} resolution. Typically, 16 interferometer scans were averaged to produce one spectrum. Absorbance spectra were computed with reference to a rough Au sample. Upon completion, an experiment yields a two dimensional array of spectra that correspond to positions in the two-dimensional spatial scan

of the sample. The spectra are analyzed to obtain an image displaying the spatial distribution of a particular chemical moiety on the sample. The maximum spatial resolution obtained by the technique is limited by the 1 mm spot size of illumination on the sample.

SPECTRA AND SPECTRAL IMAGES

Cellulose

Figure 2 shows a DRIFT spectrum of filter paper in the mid-infrared region. Filter paper, which by necessity must be free of extractants and impurities, was chosen as a sample to illustrate the chemical nature of cellulose. Strong absorbance peaks occur at 3450 cm^{-1} and 2900 cm^{-1} due to absorption of light by stretching mode vibrations of O-H and C-H bonds in cellulose, respectively. Moisture in the paper is a primary contributor to the O-H absorbance. Other strong cellulose absorbance features occur in the fingerprint region ($1500\text{--}500\text{ cm}^{-1}$). Since paper is an excellent diffuse scattering medium, virtually no features from spectral reflecting light occur in the spectrum.

Plastics

Plastic items are expected contaminants in recycled feedstock due to the abundance of plastic bottles and packaging material. Since plastic components must be removed from the pulp, removal of plastics from the feedstock prior to pulping is desired. The chemistry of plastics is dramatically different than cellulose; thus, plastics can be identified for sorting with MIR spectroscopy. **Figure 3** shows DRIFT spectra of cellulose and several different types of plastic. The spectral signatures of plastic are quite different than the cellulose spectrum. Most notably, the moisture band is not evident in the plastic spectra, and the absorbance peaks are significantly sharper for the plastics, which is indicative of the controlled and relatively simple polymer chemistry.

In addition to the chemical differences between plastic and cellulose, a physical difference between plastic and cellulose-based materials affects the spectral differences between plastic and cellulose, also. Specifically, the surface structure of most plastic materials is smooth relative to the surface of cellulose-based materials like paper and wood. The rough surface of cellulose-based materials, resulting from wood fiber structure, is an excellent light scattering surface, and therefore, light incident on cellulose materials is primarily diffusely reflected from the surface. In contrast, the smooth surface of most plastics specularly reflects most of the light incident on the surface. Even textured plastic surfaces commonly found on plastic bottles exhibit predominantly specular reflection, since the periodic distal variation of the texture is much larger than the wavelength of light.

For cellulose-based materials the ratio of diffuse to specular reflected light is large, and therefore, the signal obtained with DRIFT spectroscopy is large. For plastics the ratio of diffuse to specular reflected light is relatively small, and a small DRIFT signal is detected (since specular reflected light is generally not detected with the DRIFT technique). Additionally, for diffusely reflected light, reflectance intensities are dominated by the imaginary (extinction, or absorption) coefficient of the index of refraction, but a combination of the real and imaginary parts of the index of refraction affect the reflected intensities for specular reflected light. Thus, a DRIFT spectrum taken from plastic will often have inverted absorbance peaks from increased diffuse reflection at wavelengths where minima in the specular reflection intensities occur.

Figure 4 shows single beam and absorbance spectra for plastic and cellulose. The single beam spectrum is the Fourier transform of the actual signal detected by the detector. A sample

absorbance spectrum ($A(\lambda)$) is generated from the sample single beam spectrum ($s(\lambda)$) by reference to the single beam spectrum of a chemically inactive substrate ($r(\lambda)$) - a rough Au surface is used as a reference in the study presented here. The following formula determines absorbance from the sample and reference single beam spectra:

$$A(\lambda) = -\log[s(\lambda)/r(\lambda)].$$

In **Figure 4.b.** the difference in the diffuse reflected signal intensities between wood and plastic is shown by the single beam spectra. The single beam spectra were taken with the same gain on the detector for this demonstration. In **Figure 4.a.** the absorbance spectra generated from the single beam spectra in **Figure 4.b.** show the diffuse intensity difference, also. A baseline shift of 1.5 absorbance units in the plastic spectrum indicates that only 3.2% of the light hitting the sample contributes to the diffuse reflectance signal, and the noise in the plastic spectrum also indicates a weak diffuse signal.

Adhesives

Adhesives are another class of material that is not desired in feedstock. A wide variety of adhesives exists, and the quantity of adhesive applied for different applications varies. Since adhesives have potential of occurring on a large percentage of recycled stock, quantitation of adhesive content is needed more than identification; thus, parameters in the wet chemistry stage can be appropriately adjusted to remove adhesive chemistries without using excess chemicals or process time.

Figure 5 shows DRIFT spectra of several different adhesives that are commonly found in recycled stock. Tape and glue adhesive spectra are shown, and a paper sample with adhesive is also shown. Similar to the spectra of plastics, the C-H stretching region of adhesives has absorbance features distinctive of the adhesive chemistry. However, significant moisture features exist in the spectra, also. Strong absorbance features in the aliphatic ester carbonyl band ($\sim 1735\text{ cm}^{-1}$) and the aromatic ester band ($\sim 1660\text{ cm}^{-1}$) are typical of adhesives and are also potential interferences for lignin quantitation (see "Lignin Chemistry" section below). However, the adhesive carbonyl, ester, and phenyl bands are relatively sharp and uniquely structured in comparison to natural lignin and extractives. Chemometrics will assist in the separation between adhesives and lignin by detecting the corresponding spectral shapes and using other absorbance peaks (i.e. C-H stretch and fingerprint region) to determine chemical composition.

The amount of adhesive on paper or cardboard can be determined from the magnitude of the adhesive's spectral features. **Figure 6** demonstrates the ability to distinguish different amounts of an adhesive on a paper surface; two spectral images of a common office supply, 3M "Post-It" Notepad, are shown. The false color indicates the absorbance magnitude of the 1738 cm^{-1} peak of the adhesive. **Figure 6.a.** shows the backside of a "Post-It" note; the adhesive strip is clearly evident. **Figure 6.b.** shows the front side of a "Post-It" note; the residue of adhesive from the previous note is detected. The absorbance peak magnitude of the residue adhesive is clearly less than the main adhesive strip; the scale of **Figure 6.b.** is one-sixth of the scale in **Figure 6.a.**

Ink Content

The presence of ink on paper is indicated in DRIFT spectra of paper by a baseline shift and the presence of absorbance peaks due to ink chemistry and/or ink interactions with paper.

The baseline shift is due to absorption of the infrared light by elements in the ink; carbon is the primary absorbing element in black inks. Ink chemistry will vary greatly for the wide variety of available inks, and absorbance bands due to bonds between ink and paper will vary with the ink chemistry.

Figure 7 shows two DRIFT spectra of paper with different black inks applied; a spectrum of paper without ink is shown for reference. The black ink caused a large baseline shift in the spectrum of paper due to the strong absorbing properties of the carbon in the ink. Characteristic absorbance peaks occur at 1729 cm^{-1} for the laser-printer ink and in the fingerprint and C-H stretching regions for both the laser-printer and copier inks.

Figure 8 shows a spectral image of a paper sample with four regions of varying ink content (from complete saturation with ink to no ink). The copier ink shown in **Figure 7** was used to make the sample. The ink content is shown by the magnitude of one of the fingerprint region absorbance bands (758 cm^{-1}). As was observed at the February meeting of the Sensors and Controls group, the ink content could be visually determined; however, the MIR response to ink must be determined in order to avoid interference from ink during measurement of other parameters.

Clay-Based Coatings

Clay-based coatings on paper are detected easily by the DRIFT technique due to the strong absorbance features associated with the clay in the coatings. **Figure 9** shows a spectrum of paper with a clay-based coating and a spectrum of uncoated paper for comparison. The absorbance peaks from 3600 to 3700 cm^{-1} are due to the Group II hydroxyl bonds associated with the clay in the coating material.

Figure 10 is a spectral image of a paper sample with a clay-based coating. The coating thickness varies across the sample. In the lower left corner of the sample, the coating has been removed entirely from the paper. Additionally, a circular region in the center of the sample has had the coating partially removed. The spectral region from 3580 to 3750 cm^{-1} was integrated to generate the spectral image. The spectra were normalized with multiplicative scattering correction prior to integration. The variation in the coating thickness is reflected by the false color magnitude of those regions in the image.

The measurement of coatings on the web is desired to improve product quality, specifically printability of paper and board products. Although clay is a dominant chemistry in the majority of paper coatings, additives and other specialized coatings are extremely important elements affecting printability. The measurement of clay in coatings and other materials in coatings on the web would be beneficial. Since absorbance features in MIR spectroscopy are directly related to chemical bonds in the sample, MIR spectroscopy has potential to directly determine bonding mechanisms between coatings and paper; bond sites for ink attachment are another possible chemical characteristic that MIR spectroscopy may measure.

The DRIFT spectra of several chemically different coatings on paper are shown in **Figure 11**. The coated paper samples are labeled C1 through C7. The clay hydroxyl band (3600 to 3700 cm^{-1}) varies in chemistry and magnitude with different coating thickness and chemistry. Note the sharp absorbance features at 2513 cm^{-1} and 1795 cm^{-1} for all but one of the coatings. The 2513 cm^{-1} band may be a combination band of two peaks (1795 cm^{-1} and 713 cm^{-1}) in the fingerprint region. These absorbance peaks may be due to additives in the coatings and occur in spectral regions not affected by water absorbance features. Latex, a common additive, has characteristic absorbance peaks at 2725 cm^{-1} , 1792 cm^{-1} , and 1742 cm^{-1} . MIR spectroscopy will

be capable of measuring coating thickness as well as coating chemistry. The detection limits will differ with chemistry since the magnitude of absorbance features and the susceptibility to interference from water vary with chemistry.

Lignin Chemistry

The primary constituents of wood are cellulose, hemicellulose, and lignin. The relatively complex nature of lignin with respect to cellulose allows detection of lignin with DRIFT spectrometry. The actual chemistry of the lignin varies with wood type; in addition, the lignin chemistry will vary from tree to tree and within the same tree as environmental factors affect the growth of a tree.

Figure 12 shows DRIFT spectra of four different cellulose-based materials with different lignin content. Detection of lignin content in recycled feedstock is possible by measuring the absorbance strength in the lignin bands ($1500\text{--}1800\text{ cm}^{-1}$). Note that the "cereal-box"-type cardboard spectrum contains small clay peaks, which indicates the presence of recycle material in the board.

The absorbance bands from $1500\text{--}1800\text{ cm}^{-1}$ can potentially be used to distinguish types of wood. **Figure 13** shows DRIFT spectra of four different wood types (elm, holly, mahogany, and oak); a spectrum of filter paper is shown as a reference of cellulose with no lignin. Lignin chemistries present in the spectra include an aliphatic ester carbonyl band at 1743 cm^{-1} , an aromatic ester band at 1663 cm^{-1} , and aromatic bands corresponding to phenyl vibration modes at 1594 cm^{-1} and 1505 cm^{-1} . The relative strengths of the lignin bands vary with wood type.

Figures 14.a., 14.b., and 14.c. show spectral images of a sample made up of the four wood types shown in **Figure 13**. The sample is made up of 1 inch strips of wood; the types of woods from top to bottom in the images are elm, holly, mahogany, and oak, respectively. The grain pattern in each wood runs from top to bottom in the image. **Figures 14.a., 14.b., and 14.c.** were generated by measuring the peak heights of the 1743 cm^{-1} , 1663 cm^{-1} , and 1595 cm^{-1} bands, respectively. The lignin chemistry variation among the different types of wood is evident in the relative differences in the strengths of the absorbance bands analyzed. Variations in chemistry corresponding to grain pattern are also seen in the images.

Moisture

Moisture is a potential interference source in any MIR spectroscopic technique. The strong absorbance of MIR light by moisture makes detection of other absorbance features difficult in conditions of high moisture content. The moisture response of the DRIFT analysis of cellulose-based materials is shown in **Figure 15**. Several spectra with varying moisture content are shown. At high levels of moisture, only strong absorbance features of the cellulose that occur in regions separate from the water absorbance bands (3600 cm^{-1} and 1700 cm^{-1}) are detected in the spectra. At room conditions, moisture is still present in the paper and the DRIFT spectrum of paper.

Certain parameters will be affected by high moisture content, but the degree that moisture will affect a particular measurement can not be determined until a complete analysis incorporating varying moisture levels is completed. It should be noted that Fourier transform instrumentation is particularly affected by moisture content due to the limits of amplification inherent in the detection of the interferogram, which requires a large dynamic detection range. Movement to non-Fourier transform-based measurements should improve the ability to compensate for absorption of MIR light by water; however, since MIR spectroscopy is a surface

spectroscopy, high moisture content that generates a "wet" look will be problematic for any MIR spectroscopic technique.

CONCLUSION

MIR spectroscopy is a sensitive technique for analyzing the chemical nature of cellulose-based materials. The diffuse reflectance geometry is particularly useful for paper and wood due to the scattering properties of those materials. Macroscopic spectral images obtained by scanning the sample beneath the spectrometer allow visual identification of the spatial distribution of chemistries on the sample. The spectroscopy is sensitive to a wide variety of chemistries of interest. Moisture content is a potential interference source; the degree of interference will depend on chemistry and application.

The DRIFT data show encouraging results for two applications of MIR spectroscopy: feedstock analysis and coating analysis. The sensitivity of MIR spectroscopy to a wide variety of chemistries will allow sorting and analysis of recycled stock, which has a broad chemical base. The excellent response of the technique to absorbance features in the phenyl and carbonyl bands will allow identification of wood species based on the spectral signature resulting from the lignin and extractive content of the wood. Chemometrics analysis will be necessary for any feedstock application due to the large variability in the stock. DRIFT spectra of coatings on paper show that coating chemistry and thickness can be determined with MIR spectroscopy. The spectral features from clay in coatings are easily detected and are suitable for thickness analysis. The potential for additive analysis by MIR spectroscopy also exists, especially for additives with spectral features removed from water absorbance bands.

III. CHEMOMETRICS ANALYSIS FOR WOOD IDENTIFICATION

Currently there exist a number of multivariate data analysis techniques that can be used for on-line characterization and monitoring. These techniques provide the means for reducing a large amount of data into its relevant features. Specifically, for analytical technologies like infrared spectroscopy, spectral features that are related to a physical or chemical property of interest can be extracted and modeled using factor analysis techniques like Principle Component Regression (PCR) or Partial Least-Squares Analysis (PLS). These techniques are often used to predict quantitative information. However, for applications requiring qualitative information, i.e., classification schemes, pattern recognition techniques like K-Nearest Neighbor (KNN) and Soft Independent Modeling of Class Analogy (SIMCA) can be used. In this paper, we report on the usefulness and limitations of using both KNN and SIMCA to classify diffuse infrared spectra obtained from a variety of materials of particular interest to Forest and Paper Products Industries.

DATA ANALYSIS BACKGROUND

Both the KNN and SIMCA classification methods used in this study are supervised learning techniques. Models are developed by organizing empirical information based on initial assumptions regarding class membership. Though both KNN and SIMCA are classification schemes, they differ in how they classify individual samples. KNN is a similarity-based method. Its basic premise is that data of samples in the same class will be in close proximity to each other. Therefore, to classify an unknown sample a similarity distance is calculated between that sample and samples of known identity. The smaller the distance, the more similar the materials. Basically, data is collected that contains samples from each class. An interclass distance between samples is calculated,

$$d_{su} = \left[\sum_{j=1}^N (s_j - u_j)^2 \right]^{1/2}$$

where N is the number of variables or components used, and s and u are data related to known and unknown samples, respectively. These values are then used to generate a distance matrix. Distance of the unknown sample to a certain number of its nearest neighbors, designated as K , is then compared. The lowest distance determines class identification. Distances can be based on a single link: the closest member of the class, a complete link: the farthest member of the class, or the centroid: the center of the class cluster.

SIMCA is a classification technique that constructs a principle component analysis (PCA) model for each class of samples. PCA is accomplished by decomposing the data matrix constructed from known samples into a series of eigenvectors, or loadings. These loadings correspond to, in the case of infrared data, spectral features that undergo the greatest amount of variance. This variance is then related to the property of interest through a scores vector. Optimization of a SIMCA model is accomplished by choosing the proper number of loadings that adequately describe the property of interest and best cluster the individual classes. The optimal number of loadings may vary from class to class. Whereas choosing a large number of loading vectors will improve the clustering of a group, choosing too many loadings results in data variances, such as noise, being included that may not be attributable to the property of

interest. The data becomes "overfit", reducing its ability to adequately predict subsequent samples.

The development and evaluation of a classification model contains three data sets. The first is a training set that contains representative samples of known class. It is used to build the classification model. The second is the evaluation set. This set contains known samples that were not used to generate the model but are used to challenge and test the accuracy of the model. The third set contains true unknown samples that the model is intended to predict. Details on how these sets were generated for this study are discussed in the Classifications Section.

PRELIMINARY DATA ANALYSIS

To examine the experimental setup and instrumental variation and stability of the SOC-400 Diffuse Reflectance Infrared Spectrometer, a series of infrared spectra were obtained from an uncoated aluminum plate, or target. Variations in spectral intensities were observed. These variations were further examined by generating a sample matrix containing 501 spectra taken at different positions on the target. These were obtained by the same automated procedure as was followed for collection of the wood and paper sample spectra. Eigenvectors of this matrix, which concisely describe the spectral variations, were then generated. Multiplying eigenvectors with the original data matrix generate a scores vector. Scores can be used to quantitate spectral variation.

The larger the absolute score value or any pattern associated with score intensity can be used as a diagnostic. As can be observed in **Figure 16**, the mean centered raw score intensities plotted as a function of sample number have what appear to be a periodic function superimposed on a negative slope. This suggests that the data matrix has a discernible amount of instrumental drift associated with it. The larger these scores values, the more difficult it is for a classification model to distinguish between physical differences in samples and differences associated with instrumental variation and noise. Therefore, various preprocessing techniques were examined to minimize the effect of instrument drift on the score values.

It was determined that taking the first derivative of the spectra, which removes baseline variations, can successfully remove a large amount of instrument drift. As shown in **Figure 17**, not only are the scores of lower absolute intensity but the periodic function superimposed on the raw data score is not present in this figure. Similar behavior, reduction of spectral variation associated with baseline changes, was observed with infrared spectra obtained from various samples of paper.

CLASSIFICATIONS

The infrared spectra of paper, cardboard, and a series of hardwoods were classified using KNN and SIMCA. All data were preprocessed by taking the 1st derivative of the infrared spectra. A total of 40 spectra were obtained from each of the different samples. Five samples from each class were randomly chosen and withheld during training of the classification model and used as the evaluation set. Sample class and percent correct from a SIMCA classification model, generated using a 9 factor model for each class, and a KNN model, with $K=3$, are shown in **Table 1**. As can be observed, SIMCA could correctly model 86% of the samples (18 of 21) 100% of the time, and 90% of the samples (19 of 21) 80% of the time. The KNN classification model performed significantly poorer with 57% (12 of 21) of the samples correctly predicted with 100% accuracy, and 71% with 80% or better accuracy. The numbers in parenthesis behind the class description in the table are sample class numbers arbitrarily assigned during processing.

As can be seen, mahogany, maple, and oak have multiple numbers associated with them. This is because there were duplicate samples, but different physical samples, run for these classes. Five samples were withheld for each sample and therefore, classification was performed on 15 mahogany, 2 maple, and 2 oak samples.

Table 1. Classification Performance on 130 samples

Class	% Correct Fit SIMCA	% Correct Fit KNN
Paper (1)	100	100
Cardboard (2)	100	100
Birch (3)	60	0
Walnut (4,26)	100	90
Zebrawood (5)	100	40
Cherry (African) (6)	100	100
Cherry (flat) (7)	100	100
Elm (8)	100	0
Eucalyptus (9)	100	100
Hickory (10)	100	20
Holly (11)	100	100
Mahogany(12,13,14)	100	100
Maple (15,16)	80	0
Oak (17,18)	100	90
Oak, brown (19)	100	100
Oak, red (20)	100	100
Padauk (21)	100	100
Rosewood, brazil (22)	100	100
Rosewood, India (23)	100	100
Beech (24)	40	0
Sycamore (25)	100	80

Examination of the various diagnostic tools associated with the SIMCA classification procedure indicates that the fingerprint and group frequency region ($700\text{-}1800\text{ cm}^{-1}$) contains absorption bands with the strongest modeling power. Specifically, features attributable to lignins and carboxyl groups present in aromatic acids have strong modeling power. Absorption features directly associated with cellulose were observed to not vary significantly between the various hardwoods as well as the paper and cardboard.

An infrared spectra 10 wood samples were classified using both KNN and SIMCA models. The results are shown in **Table 2**. As can be seen, SIMCA identified 9 of the 10 woods correctly. Even birch and beech, which were not modeled well in the evaluation set, were correctly identified in this test set. With $K=3$, KNN identified 5 of the 10 samples correctly. Reducing K to 1 improved this identification by 1 but reduced the evaluation set performance.

Table 2. Predicting Classification of Unknown Samples

Sample	SIMCA Best Guess	KNN Guess	Identification
Unknown1	Hickory	Oak	Hickory
Unknown2	Oak	Oak	Oak
Unknown3	Sycamore	Sycamore	Sycamore
Unknown4	Walnut	Walnut	Walnut
Unknown5	Cherry (flat)	Mahogany	Cherry (flat)
Unknown6	Birch	Maple	Birch
Unknown7	Holly	Holly	Holly
Unknown8	Beech	Oak	Beech
Unknown9	Mahogany	Mahogany	Mahogany
Unknown10	Sycamore	Oak	Maple

CONCLUSION

This work clearly shows the power of using diffuse infrared spectroscopy coupled with supervised learning classification methods for the classification of various cellulose based materials. Specifically, SIMCA performed significantly better than KNN during classification of the evaluation set, with 80 % accuracy on 90% and 71% of the sample classes for SIMCA and KNN, respectively. The superior performance of SIMCA was further demonstrated during classification of unknown test samples. SIMCA predicted 90% of the unknown sample classes correctly while KNN predicted 50% of them correctly.

IV. BANDPASS FILTER-BASED INSTRUMENTATION

The Fourier transform(FT)-based instrument used for obtaining the preliminary data on cellulose-based materials is an excellent laboratory device but will not be suitable for use in a factory setting. The two main limitations of a FT-based device are speed and susceptibility to sample motion. Although a FT spectrometer is generally much faster than a scanning grating-based spectrometer, the speed of FT devices is too slow for use on the web or other rapid moving samples. The susceptibility to sample motion during the collection of the interferogram is a greater limitation of FT devices. A FT spectrometer obtains an interference signal (interferogram) that varies with time according to the interference produced by a moving mirror in the interferometer. The FT of the interferogram produces a wavelength-varying signal from the time-varying signal; thus, a spectrum is created. If sample motion occurs during the collection of the interferogram, a time-varying signal other than the interference-based signal can be detected, and a false absorbance feature will result in the spectrum. Thus, sample motion is a potential source of catastrophic interference.

For the applications of interest here, a non-FT instrument will be needed to enable spectroscopy of a large number of samples (or points on a sample) in a short amount of time. The required speed can be obtained with a bandpass filter-based spectrometer. Here a preliminary investigation of bandpass filter-based instrumentation is presented.

VISIBLE/NIR IMAGING SPECTRORADIOMETER

Although no MIR imaging spectroradiometer was available for investigation at this stage in the project, a visible/near-infrared (vis/NIR) imaging spectroradiometer was made available to the project by Surface Optics Corporation, an industry collaborator of Oak Ridge Y-12 Plant. The vis/NIR spectroradiometer is a real-time imaging spectroradiometer (RTISR) that is capable of obtaining images at video rates over the spectral range of 400 to 900 nm. The image size of the instrument is 256 by 192 pixels, and a 30 point spectrum is collected for each pixel in the image.

A schematic of the device is shown in **Figure 18**. The instrument resembles an ordinary video camera, except that a circular variable bandpass filter (CVF) in the optical path allows for wavelength discrimination of the collected light. By rotating the CVF and collecting images at different angular positions of the CVF, a set of images representing reflectance at different wavelengths is generated. Thus, the instrument collects a 30-point spectrum for each pixel in the image plane at video rates (30 Hz). **Figure 19** shows spectra of oak, walnut, and white paper obtained with the instrument.

Four spectral images are shown in **Figure 20**. These images are of a platter of woods that has been used in the DRIFT and chemometrics work of this project. The woods are in 1 by 4 inch strips; 24 woods are apparent in the images. 30 images were taken at 30 different wavelengths in the visible and NIR regions. Four of the images are shown at wavelengths of 468 nm, 555 nm, 641 nm, and 813 nm. The spectral images show the variation of reflected light at those wavelengths.

A schematic of a MIR imaging spectroradiometer is shown in **Figure 21**. This type of instrument will be ideal for feedstock analysis; image and chemical information will be obtained simultaneously for sorting and/or characterization of stock. The difference between the MIR and

vis/NIR instruments is primarily in the optics and components used; the principal is the same. Reflecting optics enable collection of MIR light into the fingerprint region, and the Mercury-Cadmium-Telluride (MCT) CCD detector is sensitive to light throughout the MIR region.

A MIR real-time imaging spectroradiometer is currently being manufactured by Surface Optics Corporation for Oak Y-12 Plant for a separate project. The device will be available in late FY98 for use in this project. It will have response from 2 to 5 μm which will enable investigation of the O-H and C-H stretching regions and most of the chemistry associated with coatings.

A modified version of the instrument shown will be needed for web analysis. The modification that will be needed is the transfer from a circular variable filter to a linear variable filter. The moving web will provide translation of one spatial dimension while the 2-D detector array will collect information from one spatial dimension and the spectral dimension. Thus, the modified device will have no moving parts and be capable of real-time analysis at web speeds.

PROTOTYPE DEVELOPMENT FOR WEB COATING ANALYSIS

Presently, a prototype instrument is being developed for rapid measurement of coating chemistry and thickness on paper. Many of the parts for the prototype have arrived to the lab, and preliminary testing of those components has begun. The first prototype will be capable of collecting one wavelength of information from one spatial point rapidly via a single detector, and a 256-point spectrum will be attainable at slower speeds. Investigation into price and availability of 1-D array detectors for simultaneous measurement of a 256-point spectrum from one spatial point has begun with plans to purchase an array detector with the newly arrived FY98 funds. The final prototype will incorporate a 2-D array detector to measure a 256-point spectrum from 256 spatial points simultaneously at web speeds.

V. CURRENT PROJECT STATUS AND FUTURE PLANS

CURRENT PROJECT STATUS

A large database of MIR spectra has been collected with DRIFT spectroscopy. The data has yielded information on spectral features suitable for the classification and analysis of feedstock and the measurement of coating chemistry and thickness. Potential interference sources have also been identified. Chemometrics analysis has shown methods to reduce instrumentation artifacts in the MIR spectra and shown the ability to classify wood species from evaluation sets of data. Research and development into bandpass filter-based instrumentation for rapid spectroscopic applications has commenced.

FUTURE PLANS

Work will continue in DRIFT spectroscopy of materials. Primarily the work will center on training set development for chemometrics modeling. Chemometrics analysis will be applied to more chemical analysis and classification including the measurement of "stickies" content and other qualities of recycled stock. Development of the first prototype instrument for rapid coating analysis will be completed, and the testing and evaluation of the instrument will begin. The response of the bandpass filter-based prototype will be compared to the Fourier transform-based instrument. The interference from moisture content will be determined.

Figure 1. Diagram of DRIFT instrument. The ellipsoidal mirror allows efficient collection of diffusely scattered light.

Figure 2. DRIFT spectrum of cellulose.

Figure 3. DRIFT spectra of cellulose and common plastics. The sharp features in the C-H stretching region and the fingerprint region are indicative of man-made polymers.

Figure 4. (a) DRIFT absorbance spectra of cellulose and plastic (HDPE). (b) DRIFT Single beam spectra of the same samples. The single beam spectra show the detected signal magnitude for a gain of 1 on the detector preamplifier. The HDPE has a smooth surface that results in a small ratio of diffuse to specular reflected light. Thus, the diffuse reflected signal is small.

Figure 5. DRIFT spectra of adhesives. Note the strong band at 3400 cm⁻¹ indicating moisture content in the adhesives.

Figure 6. Spectral images of the front side (a) and back side (b) of a "Post-It" note. The adhesive content on the paper is shown by the magnitude of the 1738 cm⁻¹ absorbance peak (shown in false color). The scale in (b) is one-sixth of the scale in (a). The residue of adhesive left from the previous note in the pad is detected and shown in (b). The images are 1.5 by 2.5 inches.

Figure 7. DRIFT spectra of paper with two different inks and a spectrum of paper with no ink. The chemistry of the two inks is different as absorbance features in the fingerprint region and at the carbonyl (1735 cm⁻¹) indicate. Both inks absorb MIR light efficiently (due to carbon content) and a baseline shift in the spectra result (note that the curves are only offset by 0.2 absorbance units).

Figure 8. Spectral image of four regions (0.75 inches wide) of a single sheet of paper that has four different levels of ink (copier printed) ranging from complete coverage with black ink (far left region, region 1) to no ink (far right region, region 4). A spectrum of the paper with copier ink is shown in **Figure 7**. The data was preprocessed with multiplicative scattering correction to remove baseline shift. The false color represents the magnitude of the ink absorbance feature at 758 cm^{-1} in the finger print region. The average magnitude difference between regions 3 and 4 is relatively small (0.036 units).

Figure 9. DRIFT spectra of white paper with a clay-based coating and white paper without a coating. The clay in the coating is evident by the Group II hydroxyl bands from 3600 to 3700 cm⁻¹.

Figure 10. Spectral image of white paper with a clay-based coating. The false color corresponds to the integral of the 3580 to 3750 cm^{-1} spectral range and indicates the thickness of coating on the paper sample. The coating has been removed entirely from the paper in the lower left corner of the image. The shaded circular area in the center of the image indicates a region where the coating has been partially removed. The image is 1.25 by 1.25 inches.

Figure 11. DRIFT spectra of water, filter paper, and seven different coated paper samples with various coating chemistry (labeled C1 through C7). Sample C4 does not have strong clay hydroxyl bands (3600 to 3700 cm^{-1}). Many of the coatings have a strong, sharp absorbance band at 2513 cm^{-1} which is an excellent band for coating measurement due to the lack of interference from water in that spectral region.

Figure 12. DRIFT spectra of wood, cardboard, box-type cardboard, and paper. Although the cellulose chemistry is evident in all spectra, the region from 1500 to 1800 cm⁻¹ shows varying levels of lignin and extractives present in natural wood.

Figure 13. DRIFT spectra of elm, holly, mahogany, and oak wood samples, and a DRIFT spectrum of filter paper. Lignin in the wood samples is indicated by the absorbance bands from 1900 to 1500 cm⁻¹.

Figure 14. Spectral images of a sample composed of four wood samples. The woods were cut in 1 inch strips. The wood types from top to bottom are elm, holly, mahogany, and oak, respectively. The images are 3 by 4 inches. The false color images represent the magnitude of the peak heights of the (a) 1743 cm^{-1} , (b) 1663 cm^{-1} , and (c) 1595 cm^{-1} bands, respectively. The magnitudes of the false color images are 0.28, 0.06, and 0.21 absorbance units, respectively.

Figure 15. DRIFT spectra of filter paper with varying degrees of moisture. Even “dry” paper has significant moisture content in typical humidity conditions.

Figure 16. Scores from eigenvector decomposition of raw diffuse infrared spectra obtained from an uncoated aluminum target.

Figure 17. Scores from eigenvector decomposition of preprocessed diffuse infrared spectra obtained from an uncoated aluminum target.

Figure 18. Schematic of the Real-Time Imaging Spectroradiometer (RTISR) manufactured by Surface Optics Corporation (San Diego, CA). The instrument collects spectral images at video rates (30 Hz). A rotating circular variable bandpass-filter provides wavelength discrimination, which enables the instrument to obtain a 30-point spectrum for each pixel in an image.

Figure 19. Visible/NIR reflectance spectra of wood and paper obtained with the visible/NIR real-time imaging spectroradiometer. Each point shown represents one of 30 wavelengths where measurements were obtained to generate the spectrum. The Surface Optics Corporation instrument obtains 49,152 spectra in 1/30 of a second.

(a) 468 nm

(b) 555 nm

(c) 641 nm

(d) 813 nm

Figure 20. Spectral images obtained with the visible/NIR real-time imaging spectroradiometer. The images were obtained with **(a)** 468 nm, **(b)** 555 nm, **(c)** 641 nm, and **(d)** 813 nm light. The sample is a platter of different types of woods that have been used in the MIR studies for classification development. The different colors of the woods are evident from the variation in the grayscale images at 468 nm, 555 nm, and 468 nm. The images show less contrast as the wavelength increases from blue to near-infrared, which is expected since the brown/yellow/red color of the woods is due to absorption of blue light.

Figure 21. Schematic of a MIR version of the real-time imaging spectroradiometer. Here reflective optics and a Mercury-Cadmium-Telluride detector are used to enable imaging and detection of the MIR light, respectively.

Distribution

Sandia National Laboratory

D. S. Blair
D. Horschel
M. Scott

Oak Ridge National Laboratory

R. A. Kisner

U. S. DOE

W. J. Mings

Surface Optics Corporation

M. M. Szczesniak
M. Dombrowski

University of Washington

R. R. Gustafson

Bareiss Associates

B. Bareiss

Oak Ridge Y-12 Plant

J. E. Parks II
G. L. Powell
P. E. McKenzie/M. L. Baker
Y-12 Plant Records Services (2) 9711-5, MS-8169
[2 copies for OSTI]