TEXTILE TECHNOLOGY

Investigation of the Impact of Instrumental and Software Applications on Cotton and Botanical Trash Identification by Ultraviolet-Visible and Near-Infrared Spectroscopy

Chanel Fortier*, James Rodgers, and Jonn Foulk

ABSTRACT

Given the worldwide production, usage, and manufacturing of cotton, protocols that could identify botanical cotton trash components that can become comingled with cotton could be advantageous for quality assessment prior to spinning. Conventional methods such as the High Volume Instrument (HVI) or Shirley Analyzer do not classify or yield specific trash component information. A program was implemented 1) to determine the efficacy of the ultraviolet-visible (UV-Vis) spectroscopy technique to identify cotton trash types and 2) to compare these identification results to the results of the Fourier-transform near-infrared (FT-NIR) technique to identify cotton and individual cotton trash components. Chemometric routines involve preprocessing methods and evaluation of specific spectral wavelengths to enhance spectral differences among individual pure samples of cotton trash components and cotton fiber. The chemometric software package, Unscrambler, afforded a 67% correct botanical trash identification. The utility of this method to correctly identify cotton fiber and cotton trash components was compared to the FT-NIR spectrometer. Overall, a higher percentage of correct identifications (98%) were observed using the FT-NIR spectrometer coupled with the OPUS IDENT software package. When comparing OPUS IDENT and Unscrambler software packages for sample identification by uploading NIR data into Unscrambler, the FT-NIR identification results with Unscrambler were significantly superior to the UV-Vis identification results. Thus, the FT-NIR technique proved to be a better technique than the UV-Vis technique at identifying cotton trash types.

C. Fortier* and J. Rodgers, USDA-ARS-Southern Regional Research Center, Cotton Structure and Quality, 1100 Robert E. Lee Boulevard, New Orleans, LA 70124; and J. Foulk, USDA-ARS-Cotton Quality Research Laboratory, P.O. Box 792, Clemson, SC 29633

*Corresponding author: Chanel.Fortier@ars.usda.gov
fully identify 100% of the bark and hairy seed coat samples, and had a nearly 95% accuracy of classifying combined hairy and smooth seed coat samples. Although this method was successful, the highest degree of accuracy involved using neural network clustering, which required substantial computational time. Many individuals have evaluated cotton trash using imaging techniques in the visible spectral region, and classification of trash size categories (Siddaiah et al., 2006, 2009; Whitelock et al., 2009). The Cotton Trash Identification System (CTIS) has been used to determine the percent trash, total count, and trash count via image analysis (Siddaiah et al., 2006). The CTIS has been evaluated also for the ability to classify bark/grass, sticks, leaf, and pepper-size trash count (Siddaiah et al., 2009). When comparing CTIS to classifier calls by human classifiers, CTIS agreed with the classifier call 97% of the time. Enhanced thresholding techniques and larger training data sets for neural network algorithms could ultimately increase the accuracy of CTIS trash identifications (Siddaiah et al., 2009).

The near-infrared (NIR) spectral region encompasses 700 to 2500 nm (4000 to 12000 cm⁻¹). The primary absorbencies observed in the NIR spectral region (1100 to 2500 nm) are for the chemical species CH₂, NH₂, and OH (Burns, 1985; Rodgers, 2002). In addition, the first, second, and third overtones in the NIR region can be related to fundamental frequencies in the mid-infrared (MIR) region. Fourier-transform near-infrared (FT-NIR) spectroscopy offers distinct advantages, such as minimal sample preparation, flexibility of multiple sampling systems (e.g., fiber optic probe, rotating sphere), and the option of analyzing powder-size, pepper-size, raw samples (e.g., “sticks”), and large cotton trash samples, which is particularly useful for analyzing a heterogeneous sample such as cotton. These advantages make the study of textiles using FT-NIR, specifically for cotton, attractive (Camjani and Muller, 1996; Montalvo et al., 1991; Rodgers, 2002; Rodgers and Beck, 2005, 2009; Rodgers and Ghosh, 2008; Thibodeaux, 1992; Thomasson and Shearer, 1995).

Preliminary studies have shown that creating a spectral library using Fourier-transform mid-infrared (FT-MIR) spectroscopy has been used successfully to classify cotton trash (Allen et al., 2007; Foulk et al., 2004). This method specifically was able to categorize cotton trash on small-size samples and match unknown cotton trash spectra with reference spectra. Identification of cotton trash has been attempted using NIR (Taylor, 1996). However, this technique was complicated by low levels of trash particle contrast between bark and grass making identification of trash components difficult. Preliminary studies on the identification of individual types of cotton trash using FT-NIR spectroscopy has been used successfully to classify cotton fiber, hull, leaf, seed coat, and stem with a 98% correct identification result (Fortier et al., 2010).

Many new spectrometers include the ability to analyze samples in the ultraviolet (UV) region. The ultraviolet region is on the other end of the visible spectral region from NIR spectroscopy and has the capability of revealing spectral differences not observed in the visible and NIR region. The UV spectral region includes the far-UV region (100 to 200 nm), the mid-UV region (200 to 300 nm), and the near-UV region (300 to 400 nm). At wavelengths below 190 nm, only vacuum measurements are possible because UV radiation is absorbed by atmospheric oxygen (Perkampus, 1995). Ultraviolet-visible (UV-Vis) spectroscopy offers distinct advantages including being rapid, nondestructive, easy to use, and relatively inexpensive compared to other analytical instruments, thus making it an attractive technique for the analysis of textiles (Venkataraman, 1987). The goals of this study were 1) to determine the efficacy of the UV-Vis spectroscopy technique to identify cotton and trash types and 2) to compare these identification results to the identification results of the FT-NIR technique.

**MATERIALS AND METHODS**

**Cotton and Cotton Trash Samples.** A sample of hand-cleaned cotton of variety Paymaster 1218 BG/RR (PVP 2000000213) was used as the cotton reference. Powder-sized and pepper-sized trash samples from nine cotton varieties from three states [Mississippi (MS), New Mexico (NM), and South Carolina (SC)] were used in the FT-NIR calibration and prediction set. The pure cotton trash samples were prepared and separated by hand as outlined in Allen et al. (2007). Samples were ground using a Wiley mill into powder- and pepper-sized samples with 80 and 20 meshes, respectively. The cotton trash varieties were labeled by assigning the first two letters from the state the samples were acquired in and the last letter as the cotton variety (Table 1). The calibration set was different from the prediction
set in that the prediction set was composed of the corresponding powder-sized and pepper-sized trash samples used in the calibration set (e.g., powder-sized MS ‘Deltapine 555 BG/RR’ (PVP 200200047) [DP 555], was included in the calibration set, and pepper-sized MS DP 555 was included in the prediction set). Fewer samples were used in the UV experiments compared to the FT-NIR experiments due to the small quantity of pure samples and to the larger sampling surface for the UV-Vis instrument reflectance mode (15-mm diameter).

**FT-NIR Spectroscopy.** The FT-NIR spectra were acquired using a Bruker MPA instrument fitted with a solid fiber optic probe (Bruker Optics, Billerica, MA). Three replicate spectra were acquired at a resolution of 8 cm\(^{-1}\) and 128 scans for cotton and each cotton trash component. Specific frequency/wavelength regions were investigated to cover the entire spectral range of 800 to 2500 nm (4000 to 12500 cm\(^{-1}\)).

**CAMO Unscrambler Spectral Library Development.** UV-Vis spectral data were first collected in the Scan program. Derivative math (first and second) as well as various spectral regions (200 to 900 nm, 200 to 400 nm, and 250 to 310 nm) were investigated and carried out in the Scan program. The spectral data was then uploaded into the CAMO Unscrambler chemometric software. Sample groups consisting of cotton, hull, leaf, seed coat, and stem were then created and saved as the calibration set. Principle component analysis (PCA) models were generated for cotton and each of the individual cotton trash components, and the classification routines were run to assign identifications to the data in the validation set by projecting new samples onto the PCA models (Bakeev, 2009).

**NIR OPUS IDENT Software.** FT-NIR absorbance spectra were analyzed using the Bruker OPUS IDENT software package. In this software package, the spectral types were separated into groups representing the cotton, hull, leaf, seed coat, and stem. Preprocessing methods such as vector normalization, first derivative, and vector normalization with first derivative were investigated to normalize the spectral data (Optics, 2009). In addition, mathematical algorithms including the standard method and factorization were used to develop the identification models. Specific frequency/wavelength regions (1100 to 2400 nm, 1427 to 1867 nm, 1100 to 1800/2000 to 2400 nm) were investigated to cover the entire spectral range and select spectral ranges, such as those with and without moisture peaks.

**RESULTS AND DISCUSSION**

**FT-NIR Spectroscopy and OPUS IDENT Software.** Our initial evaluations used the NIR spectral region from 1100 to 2400 nm (Fig. 1). FT-NIR spectroscopy was used to generate a spectral library to differentiate between cotton fiber and individual trash components: hull, leaf, seed coat, and stem.

### Table 1. FT-NIR calibration set of cotton trash and cotton samples.

<table>
<thead>
<tr>
<th>Cotton Trash Variety</th>
<th>Trash Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td>Hull, Leaf, Seed Coat, Stem</td>
</tr>
<tr>
<td>NMA</td>
<td>Hull, Leaf, NS, Stem</td>
</tr>
<tr>
<td>NMB</td>
<td>Hull, Leaf, Seed Coat, Stem</td>
</tr>
<tr>
<td>NMC</td>
<td>Hull, Leaf, Seed Coat, Stem</td>
</tr>
<tr>
<td>SCA</td>
<td>Hull, Leaf, Seed Coat, Stem</td>
</tr>
<tr>
<td>SCB</td>
<td>Hull, Leaf, Seed Coat, Stem</td>
</tr>
<tr>
<td>SCC</td>
<td>Hull, NS, Seed Coat, Stem</td>
</tr>
<tr>
<td>SCE</td>
<td>Hull, Leaf, Seed Coat, Stem</td>
</tr>
<tr>
<td>Cotton</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS = no sample. The cotton trash varieties are denoted by the first two letters from the state the samples were acquired in and the last letter as the sample variety (MSA = Mississippi DP 555, NMA = New Mexico DP 555, NMB = New Mexico ‘Acala 1517-99’, NMC = New Mexico experimental fragile seed coat cotton, SCA = South Carolina ‘DP 458’, SCB = South Carolina DP 555, SCC = South Carolina DP555a, SCD = South Carolina ‘FM 989’, SCE = South Carolina PM 1218).
(Fortier et al., 2010). It is readily discerned that the trash components are spectrally different from cotton, but significant overlap is observed between the cotton trash types. Various preprocessing methods and spectral wavelengths (1100 to 2500 nm) and (1427 to 1869 nm) were investigated in OPUS IDENT to yield the largest discrimination of cotton trash. Figure 2 shows the optimized results where individual cotton trash components (leaf and stem) and cotton were identified in the original library with the exception of the hull and seed coat trash types. The optimized conditions in the original library included using first derivative preprocessing, a factorization mathematical algorithm, and a tight spectral range. Because the hull and seed coat spectra were highly overlapping, a sublibrary was created to separate out these trash components. An overall 98% correct identification result proved the efficacy of this method, with Table 2 listing the optimized results.

Table 2. NIR first derivative using OPUS IDENT software identification by cotton trash type for powder- and pepper-size samples.

<table>
<thead>
<tr>
<th>Trash Type</th>
<th>No. of Samples</th>
<th>No. Correct</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hull</td>
<td>27</td>
<td>27</td>
<td>100%</td>
</tr>
<tr>
<td>Leaf</td>
<td>27</td>
<td>27</td>
<td>100%</td>
</tr>
<tr>
<td>Seed Coat</td>
<td>21</td>
<td>19</td>
<td>91%</td>
</tr>
<tr>
<td>Stem</td>
<td>27</td>
<td>27</td>
<td>100%</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
<td>100</td>
<td>98%</td>
</tr>
</tbody>
</table>

% Correct = (no. Correct/no. of samples)*100

UV-Vis Spectroscopy and Unscrambler Software. As can be observed in Fig. 3, the cotton trash data has a high degree of overlapping spectra, whereas the cotton fiber spectrum is easily discerned. The plastic bag spectrum, which was later subtracted from the cotton and cotton trash spectra, did not have a large influence on the UV-Vis cotton trash spectra. The UV-Vis first and second derivative spectra from the Scan program are shown in Figs. 4 and 5, respectively, over the spectral range of 200 to 400 nm with substantial overlapping of the cotton trash spectra. A more narrow spectral range (250 to 310 nm) was applied to enhance spectral differences (see Figs. 6 and 7). The first derivative spectra taken in the Scan program over the narrow spectral range (250 to 310 nm) gave the highest percent accuracy (67%, Table 3). The NIR data had a much higher percent correct identification than the UV-Vis data, 98% compared to 67%, respectively. Thus, the FT-NIR instrument seems more efficient and accurate than the UV-Vis instrument in cotton trash identification. Based on these results, the UV-Vis method is not as robust or as accurate as the FT-NIR method.

Figure 1. FT-NIR absorbance spectra for “clean” cotton and cotton trash samples over entire spectral range (1100 to 2400 nm) where cotton can be identified, but cotton trash components are overlapping. Vector normalization and standard method preprocessing were applied.

Figure 2. FT-NIR first derivative spectra of cotton and cotton trash with factorization preprocessing (1427 to 1869 nm).

Figure 3. Representative UV/Vis percent reflectance spectra of cotton and cotton trash subtracting the plastic bag (200 to 400 nm).
There were several differences between the UV-Vis and FT-NIR methods, with a major difference being the type of software. Besides the instrumental differences, the OPUS IDENT software allowed for the incorporation of sublibraries and mathematical algorithms (factorization) in its chemometric program, whereas the Unscrambler software did not. Therefore, a program was implemented that com-
pared NIR and UV-Vis identification results using a common preprocessing and spectral library package. This study provided a direct comparison of the cotton trash identification capabilities of both the UV-Vis and NIR techniques by use of a common software package permitting an “apple-to-apple” comparison of the data from the two instrumental methods.

**Common Chemometric Software and Preprocessing.** Only the Unscrambler software allowed for the uploading of FT-NIR and UV-Vis raw data into its software to compare the techniques using a common chemometric software and preprocessing technique. An investigation was carried out to observe the software effects on the percent correct identification of cotton fiber and cotton trash components by uploading both the UV-Vis and NIR data in the Unscrambler chemometric software. When uploading the raw UV-Vis data taken with the Scan program into Unscrambler and taking the first derivative in Unscrambler, a classification routine was created with an observed overall 51% correct identification (Table 4). The original raw FT-NIR OPUS IDENT spectral library data also were imported into the Unscrambler program to observe the capability of Unscrambler to classify cotton fiber and the trash components with the FT-NIR data. As performed using OPUS, the Unscrambler first derivative preprocessing method was applied on the FT-NIR data with the spectral classification range of 5350 to 7004 cm\(^{-1}\) (1427 to 1869 nm) This spectral range and preprocessing method were identical to the first tier of the OPUS IDENT spectral library. However, the second tier of the OPUS IDENT library, which included adding a sublibrary to differentiate hull and seed coat trash with its own specific spectral range and mathematical algorithm, could not be performed in Unscrambler. With greater than 84% correct identification, the ability of identifying cotton trash components was demonstrated (Table 4). However, a lower percent overall accuracy was found when importing the raw NIR data from OPUS into Unscrambler compared to the optimized conditions acquired using the OPUS IDENT software, specifically for the hull and seed coat trash types. The lower accuracy results for the NIR data in the Unscrambler software might be due to the inability to prepare sublibraries.

**Sample Size Effects.** The effect of sample size using the powder and pepper-size trash samples on NIR and UV-Vis first derivative spectra was investigated. Figure 8 is representative of the FT-NIR first derivative spectra as a function of varying particle size for the hull trash. Overall, for the FT-NIR trash spectra, changes in peak intensity are observed, but band locations—indicative of key components in the identification of a sample—normally exhibited minimal to no difference. As can be observed in Fig. 8 for the hull trash type, differences in band location were minimal. These observations are in line with previous identification results for other materials (Rodgers, 2002).

![Figure 8. Representative powder-sized and pepper-sized NIR first derivative hull trash spectra (1427 to 1869 nm).](image)

<table>
<thead>
<tr>
<th>Trash Type</th>
<th>UV-Vis-Unscrambler</th>
<th>NIR-Unscrambler</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Samples</td>
<td>No. Correct</td>
</tr>
<tr>
<td>Hull</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td>Leaf</td>
<td>30</td>
<td>11</td>
</tr>
<tr>
<td>Seed Coat</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Stem</td>
<td>27</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
<td>49</td>
</tr>
</tbody>
</table>

% Correct = (no. Correct/no. of samples)\(^*100\)
However, when varying the particle size on the UV-Vis first derivative spectra, there are major and readily observed differences in the spectra of different cotton trash types. Figures 9 and 10 are shown as representatives of the UV-Vis spectra for the hull and seed coat trash types, respectively. Overall, differences in band shifts and in intensities were observed for the different trash types. The spectra in Fig. 9 demonstrate that the powder-sized and pepper-sized trash samples are clearly different for the hull trash type, particularly at 275 nm where the powder sample has a more pronounced intensity compared to the smaller maxima for the pepper trash. In addition, there appears to be a band shift towards longer wavelengths for the powder hull trash. For the seed coat spectra in Fig. 10, there appears to be agreement in the spectral bands at 255 nm, 305 nm, 330 nm, and 340 nm for powder-size and pepper-size samples. However, at wavelengths longer than 350 nm, band shifts do exist. Similar red-shifted bands were observed using UV-Vis spectroscopy and increasing the particle size of other materials (Hammad et al., 2010). When comparing the identification results for the NIR versus the UV-Vis spectra, the high similarity in the NIR powder and pepper spectra coincides with a higher percentage of correct identifications. For the UV-Vis spectra, the higher correct percentage for the seed coat (75%) and stem trash (93%) types also coincides with improved similarities in the powder and pepper spectra for seed coat and stem trash compared to leaf and hull. However, the spectral shift with particle size (powder vs. pepper) is more readily observed for the UV-Vis spectra compared to the FT-NIR spectra. Thus, the particle size does have a definite impact on the UV-Vis identification results.

In conclusion, development of a universal method to analyze botanical cotton trash components could be advantageous to cotton ginners, cotton classification, and might help facilitate the removal of problematic trash comingle with cotton lint. The objectives of this investigation were 1) to determine the efficacy of the UV-Vis spectroscopy technique to identify cotton and trash types and 2) to compare these identification results to the results of the NIR technique to identify cotton and individual cotton trash components. Both the NIR and UV-Vis spectrometers and their respective software packages were capable of identifying cotton and individual cotton trash components. However, the capabilities of creating a sublibrary and applying the factorization mathematical algorithm were not possible using the UV-Vis spectrometer or CAMO Unscrambler software. In addition, more distinct spectral characteristics were observed for the NIR data compared to the UV-Vis data. The UV-Vis spectra coupled with the Unscrambler software gave fair results of identifying the cotton trash components with 67% accuracy, but these results were not as good as the 98% identification results observed with the NIR method using the OPUS IDENT software. Side-by-side comparisons in a common software package revealed that NIR yielded better results than the UV-Vis technique. In terms of the sample size effect, similarity between NIR powder-sized trash and pepper-sized trash sample spectra corresponded with higher percent correct identification results for the NIR technique. Overall, the sample size impact was greater for the UV-Vis trash spectra, yielding larger
spectral shifts in the key band locations in the UV-Vis region compared to the observed spectral shift in the NIR region. These results validated the proof of concept that the NIR technique yielded the best overall cotton and cotton trash types identification results and demonstrated the capability and feasibility of the NIR technique to obtain a high percent accuracy (98%) in the identification of individual cotton trash types. The feasibility for expansion of the library to include more reference samples of different varieties can make the determination of botanical components more robust.

ACKNOWLEDGMENTS

The authors wish to acknowledge the outstanding work by Dr. Sho Yeung Kang, Ms. Claire Kappe, and Ms. Sarah Lillis at SRRC-ARS-USDA in performing the laboratory analysis.

DISCLAIMER

The use of a company or product name is solely for the purpose of providing specific information and does not imply approval or recommendation by the USDA to the exclusion of others.

REFERENCES


