Fourier Transform Hyperspectral Visible Imaging and the Nondestructive Analysis of Potentially Fraudulent Documents

ERIC B. BRAUNS* and R. BRIAN DYER

Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

The work presented in this paper details the design and performance characteristics of a new hyperspectral *visible* imaging technique. Rather than using optical filters or a dispersing element, this design implements Fourier transform spectroscopy to achieve spectral discrimination. One potentially powerful application of this new technology is the nondestructive analysis and authentication of written and printed documents. Document samples were prepared using red, blue, and black inks. The samples were later altered using a different ink of the same color. While the alterations are undetectable to the naked eye, the alterations involving the blue and black inks were easily detected when the spectrally resolved images were viewed. Analysis of the sample using the red inks was unsuccessful. A 2004 series $20 bill was imaged to demonstrate the application to document authentication. The results argue that counterfeit detection and quality control during printing are plausible applications of Fourier transform hyperspectral visible imaging. All of the images were subjected to fuzzy c-means cluster analysis in an effort to objectively analyze and automate image analysis. Our results show that cluster analysis can distinguish image features that have remarkably similar visible transmission spectra.

Index Headings: **Hyperspectral visible imaging; Document examination; Fourier transform spectroscopy; Fuzzy c-means; Cluster analysis.**

INTRODUCTION

Hyperspectral imaging is the simultaneous acquisition of spectral (>128 resolution elements) and spatial information. Conventionally, an image is composed of a two-dimensional (2D) array of picture elements ("pixels") on a plane. A hyperspectral image is formed when consecutive wavelength-resolved 2D image planes are stacked to form a three-dimensional (3D) image cube. The addition of the third (spectral) dimension tremendously increases the information content of an image. There are numerous methods that can be employed to obtain spectrally resolved images. The simplest solution is to use a series of bandpass filters to acquire an image for each broad spectral region. This method is cumbersome and the spectral resolution is poor. More effective methods that have been used recently incorporate either an acousto optic tunable filter (AOTF) or a liquid crystal tunable filter (LCTF) to select a series of narrow spectral regions.[1–3]

An appealing method that has not yet been explored for visible imaging is to use Fourier transform spectroscopy (FTS) to distinguish spectral bands. In FTS, the illumination source is modulated by passing it through an interferometer to vary the optical retardation, δ. The modulated light intensity is measured as a function of δ, resulting in an interferogram, $I(\delta)$. Fourier transformation of the interferogram gives the frequency spectrum, $I(\nu)$. If an array detector is used rather than a single element detector, each array element records a unique interferogram. The resulting image cube is composed of a stack of 2D image planes at each δ. A hyperspectral image is obtained by Fourier transforming each interferogram. The end result is a complete spectrum at each 2D spatial coordinate in the image. The work presented in this paper describes the development of a novel Fourier transform *visible* imaging instrument.

Fourier transform spectroscopy is well suited for hyperspectral imaging. It differs fundamentally from dispersive or diffractive techniques and presents unique advantages. Since there is no need for entrance or exit slits to control spectral resolution (as in a conventional spectrometer) the light to be spectrally resolved reaches the detector virtually unimpeded, enabling the detection of much weaker signals. This is the so-called "throughput advantage" of FTS. Without the need for a diffractive or dispersive element to separate the light into its component wavelength bands, every wavelength is monitored for the entire duration of the scan. In affect, this increases the integration time over which each wavelength band is measured. This is the "multiplex advantage" of FTS and results in dramatically increased signal-to-noise ratios. Furthermore, most interferometers implement an internal frequency standard to keep track of the mirror position as it scans; consequently, there is no need to calibrate the instrument prior to use. All of these advantages of FTS carry over into imaging applications.

A potentially valuable application of visible hyperspectral imaging is the nondestructive analysis and authentication of written and printed documents. A form of visible spectral imaging is utilized in forensic laboratories to detect document alterations and other features not immediately visible to the naked eye. This method involves visually inspecting a questionable document using a series of colored filters to detect anomalies in inks, paper composition, or some other feature.[4] For example, a legal document might have the value "$40" written in ink. The original marking could easily be altered to read "$400" by the addition of a zero. If the alteration was made at some later time, it is likely that the alteration was made using a different ink than the original. The inks may be indistinguishable to the naked eye, but their visible spectra will be slightly different. If the ink spectra are very similar, the alteration may go undetected if inspected visually under alternate light sources. The same inspection using hyperspectral imaging has obvious benefits. Determining the authenticity of printed documents (e.g., counterfeiting) is also readily amenable to visible hyperspectral imaging. The inks used to print U.S. currency are highly specific in order to hinder counterfeiting efforts.[5] The paper itself has a unique chemical composition in addition to specific morphological features to further deter counterfeiting operations. In principle, all of these features would possess unique visible spectra and should manifest in hyperspectral images. We have obtained

Received 11 February 2005; accepted 11 May 2006.

* Author to whom correspondence should be sent. E-mail: ebrauns@uidaho.edu. Current address: University of Idaho, Department of Chemistry, P.O. Box 442343, Moscow, ID 83844-2343.

0003-7028/06/6008-000000$2.00/0

hyperspectral images of a new 2004 series $20 note to further demonstrate the utility of this new technique.

Hyperspectral images must be combined with careful, detailed analysis to maximize their usefulness. However, the increased information content of a hyperspectral image could easily become overwhelming. In addition, image analysis by visual inspection is subjective and is affected by the experience and expectation of the researcher. Furthermore, this method of analyzing images is time consuming and inefficient. Clearly, the development of fast, efficient, and objective image analysis algorithms is needed. Pattern recognition using cluster analysis is a potentially valuable method to analyze hyperspectral images. "Clustering" is the unsupervised classification of patterns into groups. The goal of clustering in hyperspectral image analysis is to classify each spectrum (pattern) into groups (clusters) that share a degree of similarity. A false color image is constructed that highlights related image features. Cluster analysis has been used to analyze Fourier transform infrared (FT-IR) hyperspectral images of tissue samples with impressive results.[6] In the work presented here, we have used a fuzzy c-means (FCM) clustering algorithm to analyze the FTS images.[7,8]

EXPERIMENTAL

Instrument Setup. The imaging instrument utilizes the Michelson interferometer from a Digilab FTS6000 FT-IR spectrometer to modulate the illumination source (150 W Tungsten lamp). The sample is placed after the interferometer and before the detection optics to measure transmission (Fig. 1). The light that is transmitted through the sample is collected using a 50 mm camera lens and focused onto a charge-coupled device (CCD) array detector (PixelFly 210xs, Cooke Corporation). The FTS6000 is operated in step scan mode at a step rate of 50 Hz. Transistor–transistor logic (TTL) signals from the commercial spectrometer are used to trigger data collection. The camera is composed of 640 × 480 pixels that are 9.9 μm × 9.9 μm each. It utilizes interline transfer architecture permitting a fast acquisition rate without the need for a shutter. Exposures can be as short as 10 μs or as long as 10 ms. The camera has a digital resolution of 12 bits per pixel and a spectral response from ~400 nm to ~850 nm. For the images presented here, the pixels were binned 2 × 2 and a 400 μs exposure was used. At this binning mode, the camera can acquire and read out images at 76 Hz; well matched to the 50 Hz step scan rate of the interferometer. Essentially, the step scan rate is limited by the camera's read-out rate. The camera is controlled using software written in LabVIEW.

Free Spectral Range. Due to a limitation of the commercial spectrometer, the entire visible spectrum cannot be collected in a single scan. The spectral bandwidth of a Michelson interferometer is determined by the step size of the moving mirror and the total mirror travel distance determines the resolution. To track the mirror position, a HeNe laser is passed through the interferometer and travels the same optical path as the light source. The modulated laser signal is detected by a photodiode. The resulting signal is a cosine wave with a wavelength equal to the HeNe wavelength of 632.8 nm. Consequently, the step size of the moving mirror will always be a multiple of 632.8 nm. Typically, in FT-IR spectroscopy, data collection is triggered at specific "zero crossings" of the reference signal. For example, if data are collected at every other zero crossing (an "under-sampling ratio" of 2), the sampling interval is equal to 632.8 nm. This corresponds to a free spectral range of (1/632.8) × 0.5, which is equal to 7901.4 cm^{-1}. The factor of 0.5 is used to account for the Nyquist sampling criterion. To collect the entire visible region in one scan would require triggering data collection between zero crossings (a sampling interval equal to 158.2 nm). This would give a free spectral range (FSR) of 31605.6 cm^{-1}. In step scan mode, the FTS6000 cannot trigger between zero crossings. To account for the larger sampling interval, we must collect our spectra in two portions, one from shorter wavelengths to 632.8 nm and one from 632.8 nm to longer wavelengths. Appropriate bandpass filters are used for each portion to eliminate frequency aliasing. Since we expect little spectral response from the camera below 400 nm or above 850 nm, we can sample at an under-sampling ratio of 2 to reduce the number of points that need to be collected. At this sampling interval, we have two spectral portions, one from 421.9 nm to 632.8 nm and one from 632.8 nm to 1265.6 nm. Collecting 450 points per scan gives a spectral resolution of ~32 cm^{-1}, which is equivalent to approximately 1 nm over the visible spectrum.

Performance and Data Processing. Typically, 25 scans of 450 spectral points each (with a full 2D image array collected at each spectral element) are averaged to obtain the final sample data. The time required to collect a complete data set is ~4 minutes at a 50 Hz step scan rate. Binning the pixels of the 2D image 2 × 2, the total image size for each frame is 153.6 KB (240 × 320 × 2 bytes/pixel). A single scan of 450 images results in a total file size of 69.12 MB/scan. Careful attention to programming nuances and numeric data types is vital for efficient acquisition and processing of these large files. All data were processed using our own software written in LabVIEW. Single beam spectra were calculated for each pixel by Fourier transformation of the corresponding interferogram. The interferograms were processed using the Mertz method and apodized using a triangular function.[9] Transmission data were calculated by dividing the sample by a reference. Calculating transmission against a reference is equivalent to flat field correcting the wavelength resolved images. The wavelength resolved images were subject to fuzzy c-means cluster analysis using a routine written in LabVIEW. Details of the data clustering are given later.

Sample Handling. The document samples were written on standard 20 lb weight white printer paper. A blank sheet of the same type of paper was used as the reference image for transmission calculations. Due to inhomogeneities in the paper morphology, the reference sample is not ideal. A true reference sample would be the same piece of paper with no writing on it. Pens were chosen randomly and quickly tested using thin layer chromatography (TLC) to verify that the inks were spectroscopically different. On each piece of paper the numerals "331" were written in black, blue, or red ink. After several minutes to allow the ink to dry completely, the "331" was altered to "3340" using the alternate pen of the same color. An aliquot of each ink was dissolved in methanol and its solution spectra was recorded. The exception was the red ink used to write the original "331". This ink is insoluble in methanol and water was used instead. A series 2004 $20 note was used for the currency analysis. The bill was obtained from the bank and used prior to circulation. In all cases, the samples were illuminated from the back.

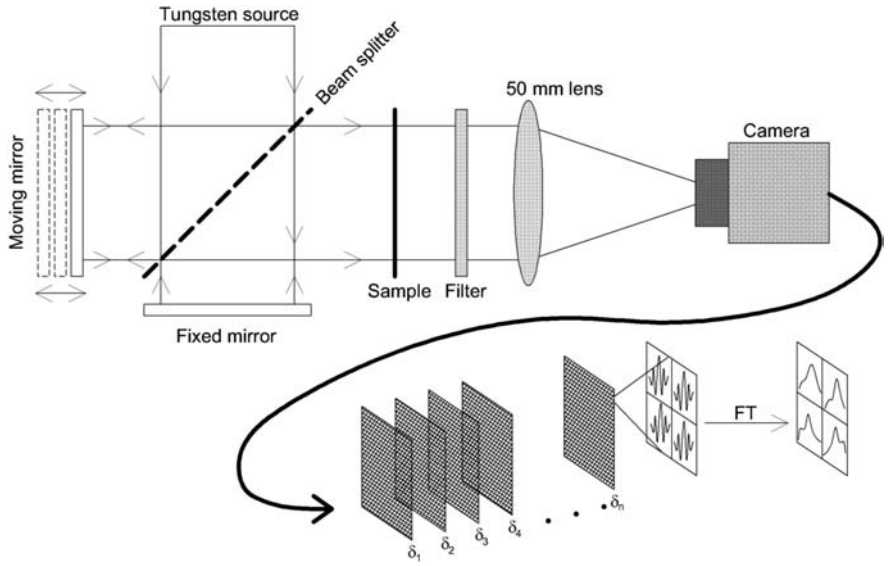

FIG. 1. Instrument setup. A 150 W Tungsten light source is modulated by a Michelson interferometer from an FTS6000 FTIR bench. A sample is placed in the beam at the exit of the interferometer. Light that is transmitted through the sample passes through a cutoff filter to eliminate frequency aliasing. The filtered light is focused onto a CCD camera via a 50 mm camera lens. Image acquisition is triggered at each mirror position, δ, and interferograms are generated at each pixel coordinate. The interferograms are Fourier transformed, resulting in a complete spectrum at each pixel.

RESULTS AND DISCUSSION

Hyperspectral Images. A representative interferogram obtained from an image after averaging 25 scans is shown in Fig. 2; the inset shows the region around the centerburst expanded. The spectral response of the system is illustrated by calculating single beam spectra for both the long and the short wavelength spectral regions; these are shown in Fig. 3. Also included in Fig. 3 are the transmission curves of the filters that were used to prevent frequency aliasing. The long-pass filter used to collect the red portion has a reasonably sharp cut-off and is well suited for use in this spectral region. The drop in response at longer wavelengths is due to the decreasing spectral sensitivity of the camera. Spectra that lie between ~640 nm to ~790 nm are considered reliable. Outside of these limits the reliability is somewhat subjective. In contrast, the short-pass filter used for the blue wavelengths has a fairly broad cut-off and impacts the spectral sensitivity (and reliability) over the long wavelength portion of this region. Although the camera is specified to have spectral sensitivity down to 400 nm, the decreasing response at shorter wavelengths is due to a lack of dynamic range. The camera has a peak sensitivity at ~550 nm and the output of the tungsten lamp drops off sharply below 500 nm. The signal at shorter wavelengths is low at light levels that do not saturate the camera at 550 nm. As such, signals at wavelengths below ~490 nm are not reliable. Overall, images from the long wavelength region are of higher quality that those collected at shorter wavelengths.

Written documents were prepared and later altered with a different ink of the same color. Three different colored inks were studied, black, blue, and red. Two different inks of each color were used for a total of six inks. Each pair of inks was indistinguishable by visual inspection as well as with filtered light using a variety of colored glass filters. For all images the background appears "mottled". This is a direct result of the mismatch between the sample and the reference. The strong scattering efficiency of visible light transmitted through white paper coupled with the paper's lack of uniform thickness results in an inhomogeneous distribution of transmitted light

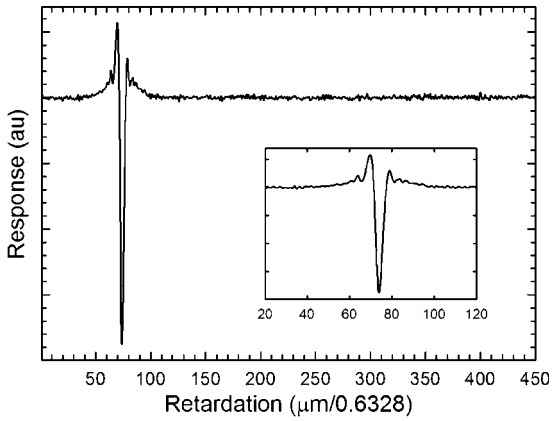

FIG. 2. A representative interferogram extracted from an image after averaging 25 scans. The inset shows the region around the centerburst expanded.

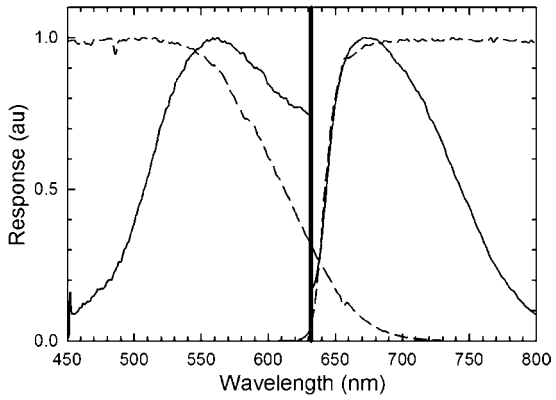

FIG. 3. Single beam spectra (solid lines) after Fourier transformation of the corresponding interferograms. The transmission curves for the filters used to prevent frequency aliasing are shown as dashed curves. The solid vertical line near the center of the figure is placed at 632.8 nm.

intensities. The inhomogeneity is further amplified when the sample response is divided by the response of the reference.

Figure 4 shows images of the black ink comparison (left column). At 640 nm, "3340" appears and the two inks are indistinguishable. At progressively longer wavelengths the transmission of both inks increases. However, at ~700 nm the original "331" fades more rapidly than the added markings. By ~720 nm, the original "331" is no longer visible and it becomes clear that the remaining markings were added using a different ink. Both inks fade completely beyond 800 nm. Below 632 nm, the black inks were indistinguishable. Shown in Fig. 5a are solution-phase UV-Vis transmission spectra of the black inks along with spectra of each ink taken from the hyperspectral images. The solution-phase spectra of the inks are virtually identical. The spectra obtained from the images are also strikingly similar. It would be difficult, if not impossible, to distinguish the inks from the spectra alone. The advantage of using hyperspectral imaging in ink analysis is clearly demonstrated by the ability to tell the different inks apart. Merely extracting a small ink sample and obtaining a solution spectrum would be insufficient.

Wavelength resolved images of the samples with blue inks are also shown in Fig. 4 (right column). The images from 640 nm to 700 nm are quite similar to the results of the black ink comparison. The original "331" begins to fade beyond 700 nm and disappears altogether at ~738 nm. Both are completely transparent beyond ~775 nm. Additionally, it appears that the differences between the blue inks are more dramatic than the black inks. The pigments used to create black inks have much broader absorbances than those used to create colored inks, such as blue. The more subtle differences between the black inks are consistent with this. Solution spectra of the blue inks are shown in Fig. 5b along with the spectra from the images. As with the black inks, the spectra are broad and unstructured

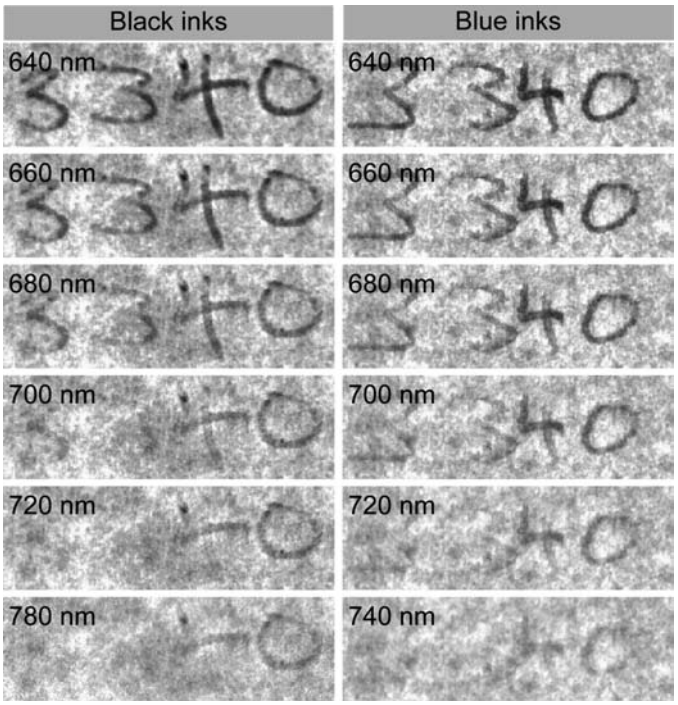

FIG. 4. Wavelength-resolved images for the document samples prepared with black inks (left column) and those prepared with blue inks (right column). Dark features indicate low transmission and light features indicate high transmission.

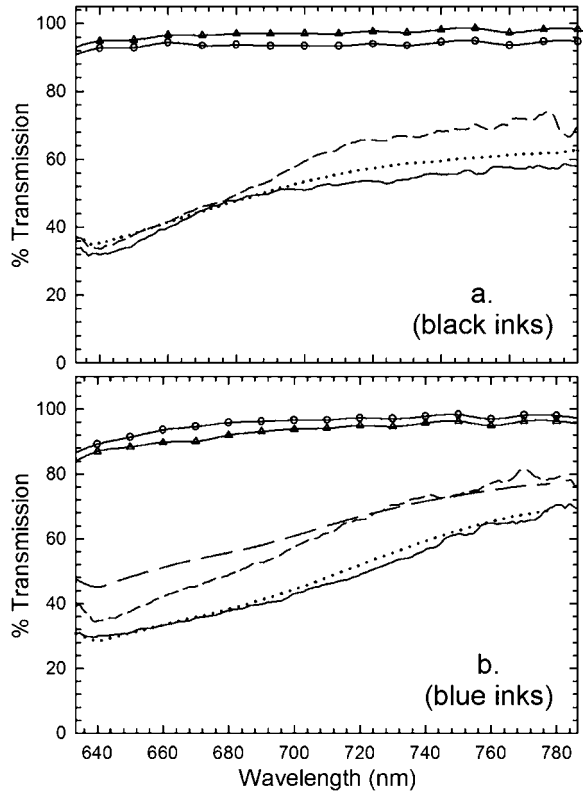

FIG. 5. Transmission spectra and cluster spectra for (**a**) the black ink analysis and (**b**) the blue ink analysis. The solid curves are pixel spectra corresponding to the inks used to alter the document. The dashed curves are pixel spectra corresponding to the original inks. The dotted curve in (a) is the cluster spectrum that is intermediate between the two inks. In (b) the dotted curve is the cluster spectrum corresponding to the alteration ink and the short dashed curve is the cluster spectrum corresponding to the original ink. The two upper curves in both plots show solution spectra (in methanol) of each of the inks; (▲) original ink; (○) alteration ink.

and the differences between them are negligible. The ability to visually distinguish the inks from the wavelength-resolved images is evident. Attempts to distinguish the red inks using imaging or isolated spectra were unsuccessful and data are not shown.

A 2004 series $20 bill was also studied using hyperspectral visible imaging.[4] Wavelength resolved images from 490 nm to 780 nm are shown in Fig. 6. The most striking feature is the serial number. The transmission is high at 490 nm and becomes increasingly transparent between 500 nm and 520 nm. At 540 nm the transmission begins to decrease and reaches a broad minimum near 660 nm. At longer wavelengths, the transmission of the serial number begins to fade and is practically transparent beyond ~780 nm. Also noteworthy are the small "20s" printed on the back of the bill that are visible in the images at 490 nm and 500 nm. The transmission of the large "20" printed on the lower right corner of the face of the bill is low and remains relatively constant over the spectral range studied. This is not surprising since this is a security feature that is printed using a so-called "color shifting ink" that changes color as a function of observation angle. If one looks closely (using a low power magnifying glass) at this feature, it becomes apparent the ink is impregnated with small reflective crystalline particles that reflect different wavelengths with varying efficiencies. The net effect is a "color shifting ink". Regardless, this feature is nearly opaque in the visible and does

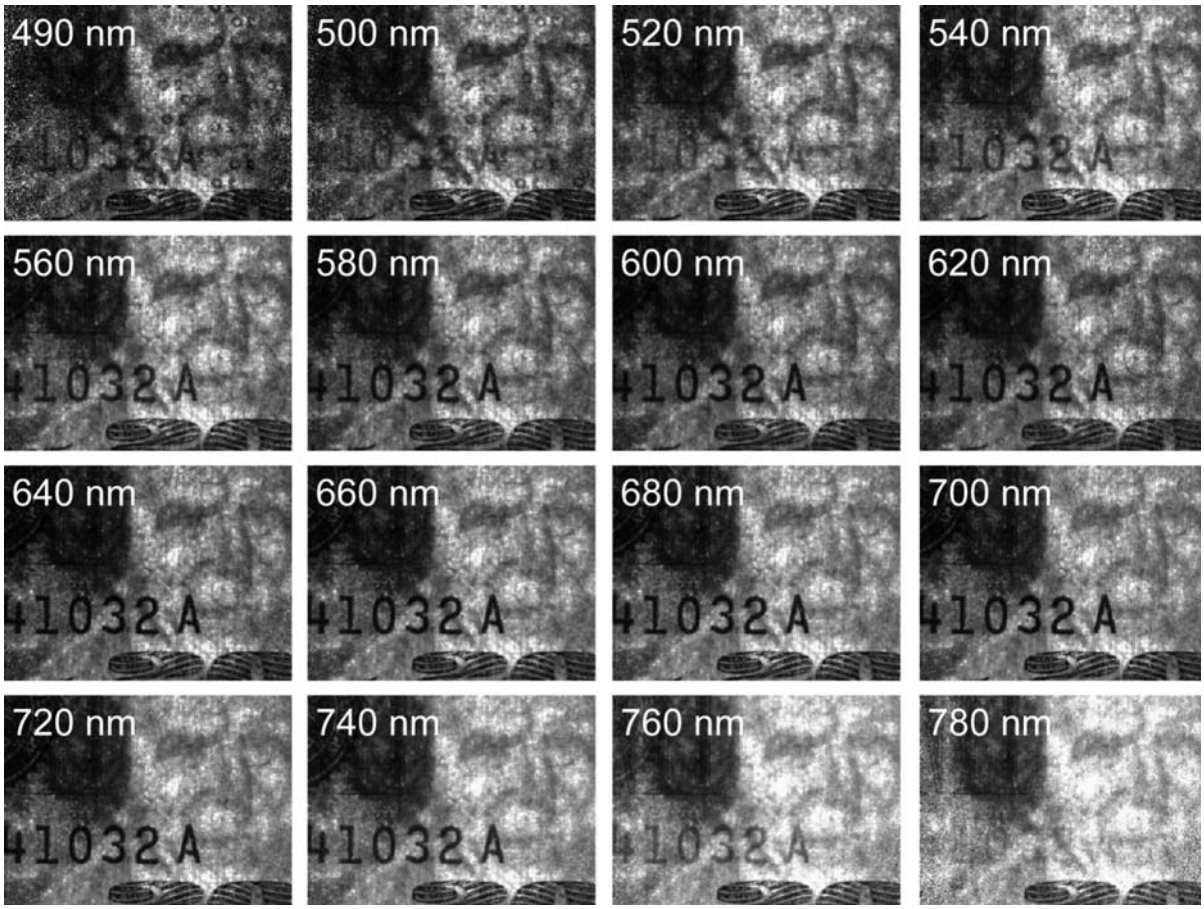

FIG. 6. Wavelength-resolved images of a 2004 series $20 bill. Dark features indicate low transmission and light features indicate high transmission.

not exhibit a wavelength dependence. (This feature is designed to be color shifting in *reflectance*—there should be a spectral and angular dependence of the reflected light—perhaps this gets washed out in transmission, at least in part because so little light is transmitted through this feature.) Also visible at all wavelengths is the watermark that is a component of the paper and not due to pigmentation. In the upper left corner is the seal of the Bureau of Engraving and Printing. The seal is green and lies on top and centered on the word "TWENTY". The seal is most visible in the 580 nm to 740 nm range. As the seal fades at longer wavelengths, the underlying "TWENTY" is still visible. Transmission spectra of the serial number and the large "20" taken from the images are shown in Fig. 7. The serial number spectra (solid curve) is consistent with what we observe from the images. Likewise, the transmission spectra from the large "20" is shown as a dashed line. While not totally opaque, it has a relatively featureless, broad spectrum spanning the entire spectral range.

Cluster Analysis. A fuzzy c-means (FCM) clustering algorithm was used to sort image spectra into related groups. References are provided for a more detailed discussion of cluster analysis.[7,8] In general, clustering is defined as the unsupervised classification of patterns (spectra) into groups (clusters). FCM is a non-hierarchical clustering method that partitions spectra into groups whose members show a degree of similarity. In fuzzy clustering, spectra can be members of more than one cluster as opposed to "hard" clustering, which mandates that each spectra belong to only one cluster. The FCM algorithm is an iterative procedure that updates a membership matrix after each iteration. The algorithm stops when changes in the membership matrix from iteration to iteration become negligible. This procedure is unsupervised with the exception that the number of clusters to classify the spectra into must be chosen initially. The raw data constitutes an $n \times d$ data matrix, where n is equal to the total number of pixels in the image and d is the number of spectral elements. Each row of this matrix is a complete spectrum at that pixel. Once the

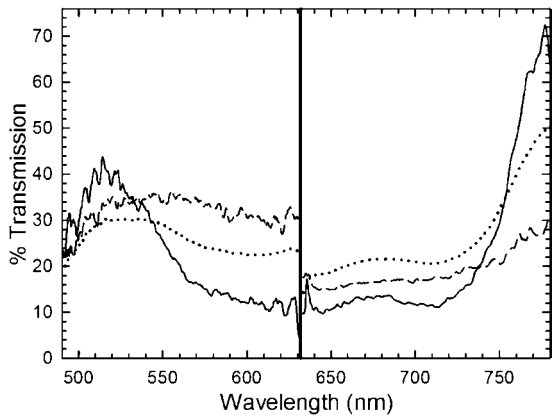

FIG. 7. Transmission and cluster spectra for features on the $20 bill. The serial number pixel spectrum is indicated by the solid line, the dashed line is the pixel spectrum for the large "20" at the bottom of the image, and the dotted line shows the cluster corresponding to both features. The solid vertical line is placed at 632.8 nm.

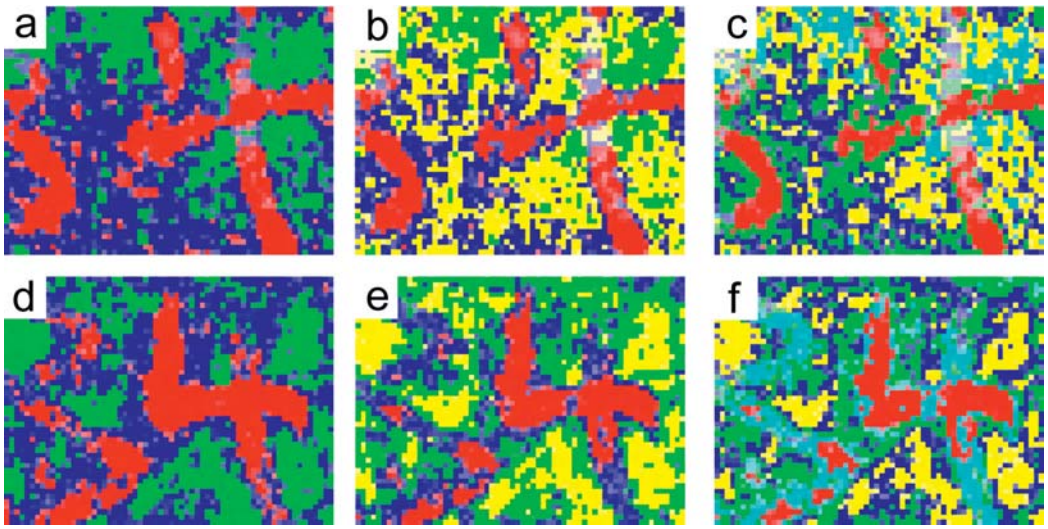

FIG. 8. Cluster images for the document samples prepared with blue and black inks. (**a**) Three, (**b**) four, (**c**) and five cluster images for the black ink documents. (**d**) Three, (**e**) four, and (**f**) and five cluster images for the blue ink documents.

number of clusters is chosen, the algorithm proceeds as follows:

(1) Select an initial fuzzy partition of the n spectra into k clusters by generating a random $n \times k$ membership matrix, **U**. Each element of **U** is a value from 0 to 1 and is the "probability" that spectra n is a member of cluster k. The elements of **U** must satisfy the following criterion: the sum over all clusters for a given spectrum must equal 1.

(2) A $k \times d$ matrix **C** of cluster centroids is calculated. A cluster centroid is a representative spectrum for the kth cluster. The elements of **C** are calculated using

$$c_{kd} = \frac{\sum_n \mu_{kn}^m x_{nd}}{\sum_n \mu_{kn}^m} \tag{1}$$

where μ_{kn} is an element of **U** and represents the membership value of spectrum n in cluster k. The exponent m is a "fuzzy" factor that is set equal to 2.[7,8]

(3) The Euclidean distance of each spectrum from each of the clusters is calculated using **C** and generates an $n \times k$ matrix, **D**:

$$d_{nk} = \sqrt{\sum_n (\vec{x}_n - \vec{c}_k)^2} \tag{2}$$

(4) The membership matrix is updated using the following relationship:

$$\mu_{nk} = \frac{1}{\sum_i \left(\frac{d_{nk}}{d_{ik}}\right)^{2/(m-1)}} \tag{3}$$

(5) Go back to step 2 and repeat until the elements of **U** no longer change significantly.

The algorithm converges rapidly and rarely requires more than 50 iterations. The initial choice of **U** in step 1 does not impact the final result. A false color image is generated using color intensities to represent the degree of membership of a pixel in a particular cluster. The intensity variations of a color representing the cluster illustrates the "fuzziness".

Figures 8a through 8c show the results of the FCM analysis of the black inks. A sub-image containing elements of both inks has been extracted and used for the analysis. The images in Figs. 8a, 8b, and 8c are false color images for the results using three, four, and five clusters, respectively. With few exceptions, each pixel has a high membership value for a particular cluster. That is to say that the degree of certainty for membership in a single cluster is high, while the certainty for membership in the remaining clusters is low. In Fig. 8a, the algorithm was able to easily identify the inks from the background. The "mottled" background is further categorized into two separate clusters. If the number of clusters is increased to 4 (Fig. 8b), no new information is obtained; only the background is further differentiated. Likewise, when 5 clusters (Fig. 8c) are used, the different inks still cannot be distinguished and the background is divided once again. In Fig. 5a, the dotted curve shows the cluster spectra for the inks (red pixels); note that the cluster spectrum is intermediate between the two pixel spectra. There is good correspondence between the ink cluster and both of the ink image spectra despite that fact that they could not be distinguished. Cluster spectra for the background pixels are not shown.

The results of the FCM analysis for the blue inks is shown in Figs. 8d through 8f. Using three clusters (Fig. 8d), the inks are differentiated from the background as with the black inks. When increased to four clusters (Fig. 8e), some of the pixels containing the original inks ("331") show up as low intensity blues, indicating a weak membership. Going a step further, a fifth cluster is added (Fig. 8f) and the original ink (represented as red pixels) can be distinguished from the alteration ink (shown as cyan pixels). Figure 5b shows that there is a strong correspondence between the cluster spectra (short dashed and dotted lines) with their pixel spectra counterparts (solid and long dashed lines).

Fuzzy c-means was also used to analyze the images of the $20 bill. These results are shown in Figs. 9a–9f. The most prominent feature, the serial number, is easily identified using FCM. However, the results also indicate that the serial number

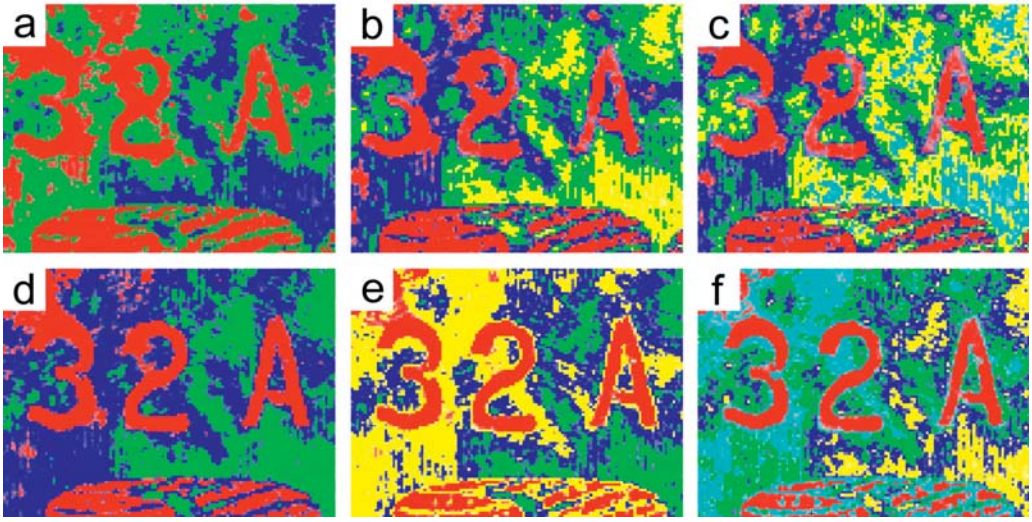

FIG. 9. Cluster images for the 2004 series $20 bill. (**a**) Three, (**b**) four, (**c**) and five cluster images for the short wavelength spectral region. (**d**) Three, (**e**) four, and (**f**) and five cluster images for the long wavelength spectral region.

pixels are members of the same cluster as the large "20" at the bottom of the image and a portion of the seal at the top left. These features could not be distinguished regardless of the number of clusters used. Additional clusters seem to only classify "background" features further. Looking at the cluster and image spectra in Fig. 7, the spectra of the serial number (solid curve) and the large "20" (dashed line) are similar and the resulting cluster (dotted line) is intermediate between the two. These images also suffer from the lack of a suitable reference image and are susceptible to variations in light transmission due to inhomogeneities of the paper.

An issue when analyzing data using FCM is whether or not to scale the data so that each spectrum is weighted equally. When the data are scaled before analysis interesting results manifest. We used a straightforward range scaling procedure that compares each transmission value with the range of all the transmission values for a given spectra according to

$$y_{nd} = \frac{x_{nd} - \min_d x_{nd}}{\max_d x_{nd} - \min_d x_{nd}} \tag{4}$$

so that for all n and d, $0 \leq y_{nd} \leq 1$.[7] Figure 10 is a three-cluster image of the $20 bill where the data were scaled according to the above relationship prior to being subjected to the FCM algorithm. The serial number (the most prominent feature) is easily distinguished and all other features are relegated to the background. The result is the same regardless of the number of clusters. Furthermore, although the data are not shown, the same result is obtained if the data from the ink comparisons are scaled prior to analysis. The inks are readily differentiated from the background and the background becomes a nearly homogeneous distribution of the remaining clusters. Clearly, scaling the data prior to analysis has a significant impact. The affect of scaling seems to favor the more highly structured spectra.

CONCLUSION

The work presented in this paper details the development of a novel hyperspectral visible imaging technique. The spectral discrimination is accomplished using Fourier transform spectroscopy (FTS). Using FTS for visible imaging provides an alternative to other technologies that employ dispersive or diffractive elements. The design is straightforward; the main elements are a Michelson interferometer, a collection lens, and a CCD camera. Images are acquired at each optical retardation introduced by the interferometer. The result is an interferogram at each 2D spatial coordinate on the image. The interferograms are Fourier transformed to give the corresponding visible spectra.

A preliminary application of the new instrument is the nondestructive analysis and authentication of written and printed documents. As an example, we have imaged altered document samples and U.S. currency. Spectrally resolved images of altered documents written in blue or black ink reveal the alterations that were otherwise undetectable to the naked eye. Images of a series 2004 $20 note have shown that it could be a powerful tool for counterfeit detection or the authentication of other printed documents. All of the images were subjected to fuzzy c-means cluster analysis to objectively extract image features. The results indicate that FCM clustering is capable of distinguishing image features that have unstructured and similar visible transmission spectra. In addition to automating image analysis and removing human subjectivity, cluster analysis can also be used to reduce the image file size by distilling the image down to its component features.

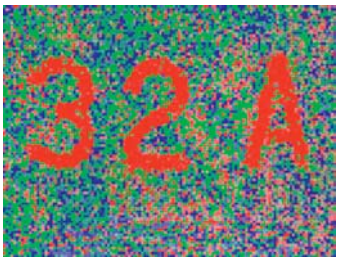

FIG. 10. A three cluster image of the new $20 bill. The data were scaled (Eq. 4) prior to FCM analysis so that all elements would be weighted equally.

1. C. D. Tran, Fresenius' J. Anal. Chem. **369,** 3 (2001).
2. C. Tran, Y. Cui, and S. Smirnov, Anal. Chem. **70,** 4701 (1998).
3. K. J. Zuzak, M. D. Schaeberle, E. N. Lewis, and I. W. Levin, Anal. Chem. **1,** 2021 (2002).
4. A. Giles, *Crime Scene to Court: The Essentials of Forensic Science*, P. White, Ed. (The Royal Society of Chemistry, Cambridge, 1998), p. 123.
5. Bureau of Engraving and Printing, http://www.bep.treas.gov (accessed Nov 2004).
6. P. Lasch, W. Haensch, D. Naumann, and M. Diem, Biochim. Biophys. Acta **2,** 176 (2004).
7. D. L. Massart and L. Kaufman, *The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis* (John Wiley and Sons, New York, 1983), p. 65.
8. J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition* (IEEE Press, Piscataway, 1992).
9. P. Griffiths and J. de Haseth, *Fourier Transform Infrared Spectrometry* (John Wiley and Sons, New York, 1986), p. 83.