Forensic Discrimination of Photocopy Toners by FT-Infrared Reflectance Spectroscopy

Edward G. Bartick, Rena A. Merrill
FBI Laboratory, Forensic Science Research and Training Center, FBI Academy, Quantico, VA 22135

William J. Egan, Brian K. Kochanowski, Stephen L. Morgan
Department of Chemistry & Biochemistry, The University of South Carolina, Columbia, SC 29208

Because of the speed, simplicity, and accessibility of photocopying, forensic examiners are encountering photocopies as often as original questioned documents. We investigated the ability of FT-infrared microscopy coupled with pattern recognition methods to discriminate among copy toner samples from a variety of manufacturers. Infrared microscopy is a preferred method due to its non-destructive nature, however, visual comparison is difficult because the observer may not be able to fully utilize the fine structure of the complex patterns. Principal component analysis and canonical variate analysis were used to visualize clustering of samples and to assess the statistical validity of the observed differences. The results illustrate the potential for computer-assisted data interpretation to provide decisive forensic identification of questioned samples.

INTRODUCTION

The use of office and personal photocopying machines has increased dramatically over the last 20 years. As a result, identifying the source of photocopied documents is not an easy task. Forensic scientists are often faced with having to identify photocopy machines involved in counterfeiting, fraud, false documents, anonymous letters, confidential materials, and acts of terrorism (1).

Techniques used for photocopy differentiation include optical microscopy, scanning electron microscopy (SEM), energy dispersive spectrometry (X-ray microprobe analysis), infrared (IR) spectroscopy, and pyrolysis gas chromatography/mass spectrometry (Py-GC/MS) (1-3). IR spectroscopy, in particular, provides information regarding the organic components of toners. Bartick and coworkers at the Forensic Science Research Institute (FBI Laboratory, FBI Academy, Quantico, VA) have performed systematic studies using microscopic infrared spectrometry (IR), diffuse reflectance, attenuated total internal reflectance, and reflection-absorption (4,5). Data from this work was summarized using a flow chart technique that enables discrimination of copy toners based on visual comparison of features in their IR spectra.

The present study is primarily concerned with a group of toners having a poly(styrene:acrylate) base component and which comprise a significant number of the commercial toners used in modern photocopy machines. Visual discrimination of the IR spectra among different samples within this group is difficult. These toners are often the hardest to differentiate by Py-GC/MS due to the complex, yet similar chromatographic patterns. FT-IR spectroscopy offers a powerful means for making comparative measurements of distinguishing characteristics for copy toners; however, an overwhelming amount of data is generated. An infrared spectrum may consist of measured light intensities at several thousand discrete wavelengths. Although an experienced analyst becomes expert at recognizing distinguishing features, the pattern recognition task is subjective and becomes quite difficult when numerous samples are compared. The objective of our preliminary work presented in this paper is the evaluation of multivariate statistical approaches for discriminating spectra of copy toners.

EXPERIMENTAL

Samples were prepared using a heat transfer process to move dry toner from documents to aluminum foil affixed to standard microscope slides as previously described (4,5). Although other materials provide a suitable reflective surface for the reflection-absorption technique, aluminum foil is readily available, inexpensive and permits the sample to be stored for further studies. The sample preparation is simple, fast and essentially non-destructive. The document is still legible after transferring the toner sample and only minimal destruction is visible microscopically. Care should be taken when sampling two-sided documents to avoid direct contact between the soldering iron and toner on the back of the document.

Spectra were collected using one of two systems: a Nicolet (Madison, WI) 20SXC infrared spectrometer with a Spectra Tech (Shelton, CT) IR-Plan microscope or a Nicolet 760 Magna infrared spectrometer with a Nicolet Nic-Plan microscope. Both microscopes utilize medium band Mercury Cadmium Telluride (MCT) detectors. Analyses were conducted at 4 cm^{-1} resolution. Between 128

CP430, *Fourier Transform Spectroscopy:* 11th International Conference
edited by J.A. de Haseth

and 256 scans were acquired for each spectrum over the range of 4000 to 650 cm^{-1}. Baseline adjustment was performed using the Nicolet OMNIC software as needed to flatten the baseline on each spectrum.

DATA ANALYSIS

Principal component analysis (PCA) is commonly used to reduce the dimensionality of data. The decomposition of the data matrix is efficiently accomplished by singular value decomposition (SVD) (6,7). SVD creates a linear combination of the correlated, original variables, forming a new set of variables which are orthogonal, *i.e.*, uncorrelated. This transformation also maximizes the variance of the original data explained by each of the new variables, which are called eigenvectors or PCs. Thus, as much information as possible is packed into the fewest number of unrelated variables. The decomposition of the data matrix allows the projection of points representing the samples into a two- or three-dimensional plot. The clustering of data points in this plot may then be used to base judgments concerning sample comparisons. Decisions about similarity of samples to one another reduce to comparison of distances between points on the plot. The clustering of similar samples can be assessed by comparison to the distances between samples judged different from one another. Whether this information is distributed among the PCs in a useful fashion, *e.g.*, permitting understandable display of the samples on the first 2-3 PCs when the samples had not been easily viewed on the many bi-plots of the original variables, is another question entirely.

Since the sample identities are known for the data presented here, the data was subjected to canonical variate analysis (CVA), also known as discriminant analysis (8). The implementation of CVA is quite similar to that of SVD and is based upon the eigen-decomposition of a matrix derived from the samples. CVA maximizes the ratio of between-group to within-group variance. Thus, CVA generates linear combinations of the original variables that best separate the groups specified by the analyst. The plot of the data points in the two- or three-dimensional space of the canonical variates permits the researcher to visualize clustering and similarity of the data. Recent work (8) has also shown the usefulness of calculating canonical variates that are orthogonal, a property not possessed by CVs in their usual formulation.

RESULTS AND DISCUSSION

FT-IR spectra were taken for each of the copy toners listed in Table 1. Representative spectra are shown in Figures 1 and 2. As can be seen, these spectra are particularly difficult to separate from one another by visual inspection of the spectra. The canonical variates analysis method requires that there be enough samples to make up a group; the 11 groups shown in Table 1 have at least 3 samples each.

TABLE 1. Copy toner samples.

Group	Number of samples
A (AB DICK)	5
B (Brother)	3
C (Citoh)	3
D (Canon)	18
E (Copystar)	5
F (HP)	6
G (Mannesmann)	3
H (Newgen)	5
I (Okidata)	5
J (Qume	4
K (TI)	3

Spectra were preprocessed using the Fast Fourier Transform (FFT). The magnitude of the modulus and the first half of the Fourier coefficients were used as the inputs to orthogonal canonical variate analysis (OCVA). OCVA attempts to find a representation of the data that minimizes the spread within a group and maximizes the spread between different groups. The resulting axes each explain different facets of the information contained in the data, *i.e.*, they are orthogonal. A projection of the samples into a three-dimensional space of the first three orthogonal CVs is shown in Figure 3.

The separation of groups by canonical variate axes are as follows:

Axis 1 & 2	A, B, E, I are well separated C, F, G, H, J are intermingled D and K are intermingled
Axis 1 & 3	E, H, K are well separated D and I are intermingled A, B, C, F, G, J are intermingled
Axis 2 & 3	A, B, E, I, K are well separated C, D, F, G, H, J are intermingled

It is apparent from these plots that samples in groups A, B, D, E, H, I, and K can be discriminated from one another. Replicate samples within these groups are closer together than the distances between groups. Samples in groups C, F, G, and J are not distinguishable by this method. In these cases, replicate samples from these groups overlap with data points representing samples from other groups. Complete differentiation of all the copy toners groups is not achieved, although several groups are clearly distinguishable. These groups include toners from AB Dick (A), Brother (B), Copystar (E), Okidata (I), Newgen (H), and Texas Instruments (K).

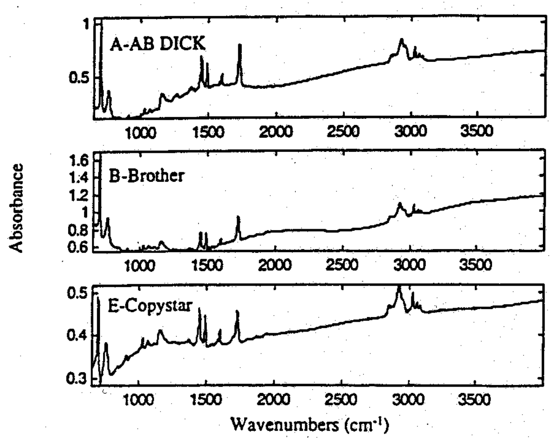


FIGURE 1. Representative FT-IR spectra from AB Dick, Brother, and Copystar toners.

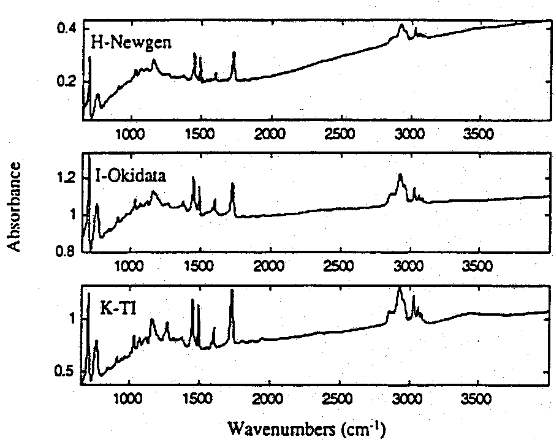


FIGURE 2. Representative FT-IR spectra from Newgen, Okidata, and Texas Instruments toners.

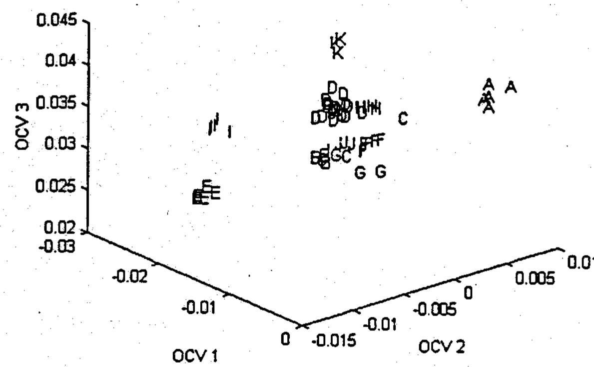


FIGURE 3. Projection of copy toner spectra on first three orthogonal canonical variate axes.

CONCLUSION

The focus of this work is development of automated, statistical-based, strategies for data handling that offer improvements in method validation and ease of interpretation. Although visual examination of these questioned materials by an expert using microscopy can be a powerful tool, such visual examination is subjective and time-consuming. The examples discussed illustrate the potential for computer-assisted data interpretation of forensic analytical data to provide decisive forensic identification of questioned samples. For each set of data, a visually interpretable map displaying the quantitative similarity of the IR spectra of forensic samples can be created. Statistical analysis of the spectra supported the discrimination ability of the IR method for several groups of poly(styrene:acrylate)-based toners.

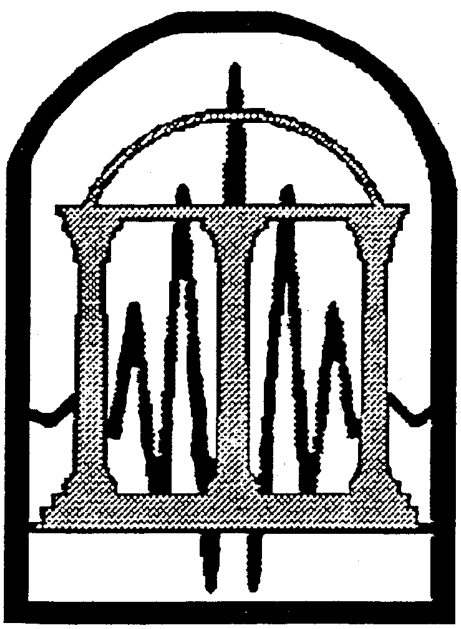
REFERENCES

1. Totty, R. N., *Forensic Science Review*, **2**, 1-23 (1990).
2. Munson, T. O., *J. Forensic Sciences*, **34** (2), 352-365 (1989).
3. Chang, W.-T., Huang, C.-W., and Giang, Y.-S., *J. Forensic Sciences*, **38**, 843-863 (1993).
4. Bartick, E. G., Tungol, M .W., and Reffner, J. A., *Anal. Chim. Acta*, **288**, 35-42 (1994).
5. Merrill, R. A., Bartick, E. G., and Mazzella, W. D., *J. Forensic Sciences*, **41**, 264-271 (1996).
6. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. *Numerical Recipes: The Art of Scientific Computing*, New York: Cambridge University Press, 1986, pp. 52-64.
7. Jackson, J. E., *A User's Guide to Principal Components*, New York: John Wiley & Sons, Inc., 1991.
8. Krzanowski, W. J., *J. Chemom.*, **9**, 509-520 (1995).

FOURIER TRANSFORM SPECTROSCOPY

Eleventh International Conference

Athens, GA August 1997



EDITOR
James A. de Haseth
University of Georgia, Athens

AIP

AIP CONFERENCE PROCEEDINGS 430

American Institute of Physics

Woodbury, New York