X-RAY FLUORESCENCE AND RAMAN SPECTROSCOPY DATA FUSION FOR ANALYSIS OF DUCT TAPES: INTRA ROLL AND INTER PRODUCT CORRELATIONS.

S. Mamedov, Horiba Scientific, Edison, NJ 08820 USA

ABSTRACT

X-Ray Fluorescence (XRF) and Raman spectroscopies are useful tools for recognizing substances and confirming their identity with little or no sample preparation. XRF provides information about elemental composition of the material, whereas Raman spectroscopy offers molecular information. The purpose of this paper is to demonstrate the application of Principal Component Analysis (PCA) to classification of duct tapes. We will likewise show that the Data Fusion Technology can be used to improve results of such classification. We will then compare and discuss the results of single source and data fused analysis. For the purpose of this research, we used micro-XRF and micro-Raman spectroscopy to analyze the spectra of duct tapes from seven different sources. To create a PCA model, we applied data fusion technology to the set of combined XRF and Raman data. This study presents the methods that allow one to improve classification of duct tapes based on spectral analysis of micro-XRF and micro-Raman data.

INTRODUCTION

X-Ray Fluorescence (XRF) spectroscopy is a non-destructive analytical technique, which requires little or no sample preparation. XRF provides information about elemental composition of the material. Raman spectroscopy is also a non-destructive analytical technique, yet it provides molecular information about the material. Both techniques enable the recording of a spectrum and a hyperspectral image of any object with high spatial resolution. A hyperspectral image is a set of data that contains information about the position of a particular point along with the full spectrum of such point, meaning that the data can be analyzed for unexpected features after the measurements have been made.

Principle Component Analysis (PCA) is a statistical procedure for identifying patterns in data and expressing such data in a way that allows one to highlight their similarities and differences (Bretscher, 1995; Manly, 1986). PCA reduces a set of correlated variables to a smaller set of

This document was presented at the Denver X-ray Conference (DXC) on Applications of X-ray Analysis.

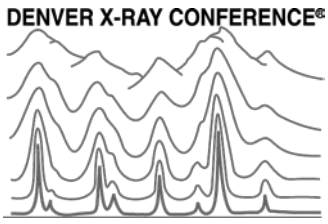
Sponsored by the International Centre for Diffraction Data (ICDD).

This document is provided by ICDD in cooperation with the authors and presenters of the DXC for the express purpose of educating the scientific community.

Usage is restricted for the purposes of education and scientific research.

DXC Website
– www.dxcicdd.com

ICDD Website
- www.icdd.com



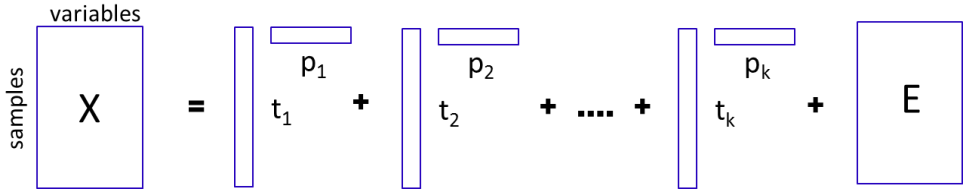
uncorrelated variables, called principal components. The first principal component has the largest possible variance and each succeeding component has the highest possible variance under the constraint that it must be orthogonal to (i.e., uncorrelated with) the preceding components. For a data matrix ***X*** with ***m*** samples and ***n*** variables, the PCA decomposition is as follows (1):

$$X = t_1 p_1^T + t_2 p_2^T + \cdots t_k p_k^T + \cdots + t_q p_q^T; q \leq \min(m, n) \qquad (1)$$

In this case, $t_1, t_2, \ldots, t_k$ are score vectors and $p_1, p_2, \ldots, p_k$ are loading vectors. Superscript T represents the transposition of the loading vectors. The $t_i p_i^T$ pairs are ordered by the amount of variance captured.

Generally, model is truncated, leaving small amount of variance in a residual matrix ***E*** (2):

$$X = t_1 p_1^T + t_2 p_2^T + \cdots t_k p_k^T + \cdots + t_q p_q^T = T^k P_k^T + E \qquad (2)$$



PCA becomes an important method in material characterizations when one needs to answer the following question: does a certain sample belong to a particular class of materials or not? (S. Mamedov, 2018)

Data fusion is the process of integrating two or more data sources to produce results that are more consistent and accurate than the ones provided by any individual data source (Llinas and Hall, 1998). There are different techniques for data fusion, yet only two of such approaches are suitable to be used with the spectral data: 1) data level fusion or low-level fusion; 2) feature data fusion or mid-level fusion. The raw data is fused in low-level fusion. In this paper, a single model of combined data blocks, appropriately scaled and preprocessed, will be used. Figure 1 presents the strategy, in which the spectra from XRF and Raman measurements are fused.

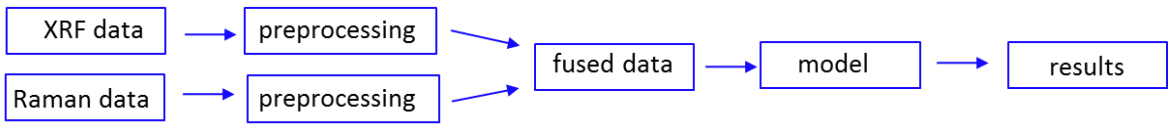


Figure 1. Data Fusion strategy for XRF and Raman spectra.

Fused XRF and Raman data follows the equation below:

$$\boldsymbol{X_{fused}} = [\boldsymbol{X_{XRF}}\boldsymbol{X_{Raman}}]$$

where $\boldsymbol{X_{fused}}$ stands for matrix of fused data or combined data block, $\boldsymbol{X_{XRF}}$ stands for matrix of micro-XRF spectra, and $\boldsymbol{X_{Raman}}$ stands for matrix of micro-Raman spectra. PCA was applied to the matrix of fused (combined) data, the matrix of micro-XRF spectra, and the matrix of Raman spectra.

EXPERIMENTAL

Seven different samples of duct tapes with labels #1020 (3M), #1110 (Scotch), #1130 (Scotch), #1230 (Scotch), #2330 (3M), #03451 (Scotch), and #3920 (3M) were analyzed. Three areas in size of 12 x 18 mm were selected to collect XRF spectra. The average spectrum over the mapped area was calculated from the hyperspectral image; the three average spectra from different areas represent XRF features of a particular tape. For Raman spectra, the three measurements are made over three separate circles on each piece of tape with the diameter of 50 microns, using DuoScan and treated as a single spectrum. These three spectra represent the Raman signature of a tape. In total, twenty-one Raman spectra from 100 cm^{-1} to 3500 cm^{-1} and twenty-one XRF spectra from 1.00 keV to 40.96 keV were collected and used for statistical analysis.

An XGT-7200V (HORIBA) X-ray fluorescence analytical microscope, equipped with a 50W Rh anode X-ray tube, was used in this study. X-ray fluorescence spectra of the materials were collected in vacuum (10^{-2} Pa) under the following conditions: 1) X-ray spot size of 1.2 mm; 2) accelerating voltage of 50 keV; 3) energy resolution of less than 143 eV at Mn Kα; 4) live time per average spectrum of 1000 s. Average spectra were truncated and the analysis was performed over the spectral range of 1 - 12 keV, as no spectral features were observed in the energy range above 12 keV. Truncated spectra were exported into EMSA file format for future analysis. LabRam Evolution (HORIBA) Raman microscope equipped with two gratings (1800 gr/mm and 300 gr/mm), DuoScan, and TE cooled CCD (Synapse, HORIBA) was used in this study. The spectra were collected using He-Ne laser (632.8 nm), Olympus x50 LWD objective, and laser power between 0.93 mW to 1.2 mW (depending on the sample). To avoid burning material under

the laser irradiation, DuoScan with spot size of 50 microns in diameter was used (Mamedov *et al.*, 1995). Acquisition time varied from 1 s to 50 s per spectra.

A matrix of Raman data has dimensions of 1678 (wavenumbers in spectrum) x 21 (7 different tapes and 3 independent measurements in each tape). A matrix of XRF data has a dimension of and 1201 (energy channels) x 21 (7 different tapes and 3 independent measurements in each tape). Fused or combined data matrix has dimensions of 2879 (1678+1201) x 21.

XRF and Raman spectra were loaded into PLS Tool Box (version 8.2.1, Eigenvector, Inc.), running under MatLab 2017a (The Mathwork Inc.).

RESULTS AND DISCUSSION

Figure 2 shows the XRF spectra (mean of 3 spectra) of the samples, where the labels correspond to different duct tapes mentioned previously.

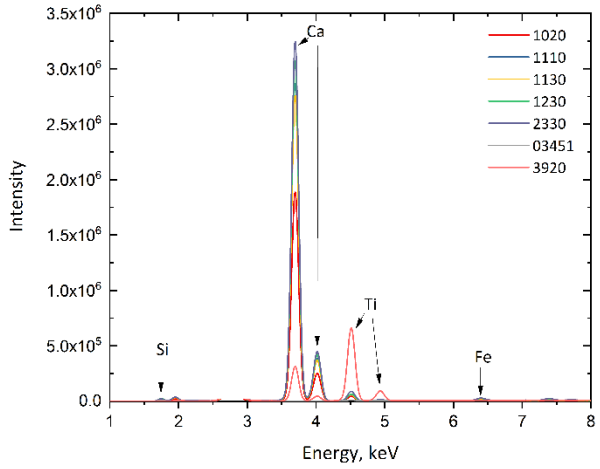


Figure 2. XRF spectra (mean of 3 spectra) of the samples.

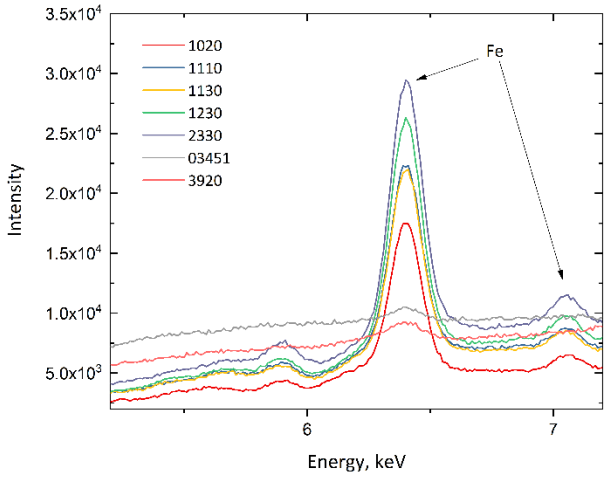


Figure 3. Intensity of Fe $K_{\alpha 1}$ line at 4.84 keV line (mean of 3 spectra) of the samples.

The XRF spectra of the samples are similar and reveal the presence of Si, Ca, Ti, and Fe. The only exception is Sample #1020, which contains significant amount of Ti (K_α line at 4.51 keV and K_β line at 4.93 keV). Ti lines for this sample are marked with arrows, as shown in Figure 2. Figure 3 presents the differences among all spectra in intensity of K_α line of Fe at 6.40 keV. Content of Fe in tapes is an important feature in classification because two tapes (03451 and 3920) do not show any presence of Fe.

PCA is scale-dependent. Autoscaling and normalization were used as preprocessing of XRF data in this study. The data for PCA is arranged in matrices by convention. The rows correspond to

samples and the columns correspond to variables (channel or energy in case of XRF). Autoscaling is mean-centering, followed by the division of each column (variable) by the standard deviation of the column. Mean-centering calculates the mean of each column (energy) and subtracts this from the column. Figure 4 shows PC1-PC2 score plot. The first three PCs captured 97.01% of variance in the spectra. The score plot shows easily identifiable and classified duct tapes with the exception of Samples #03451 and #3920, which are marked inside of the red circle. Scores for these samples overlap, which is why the tapes could not be discriminated based on XRF data.

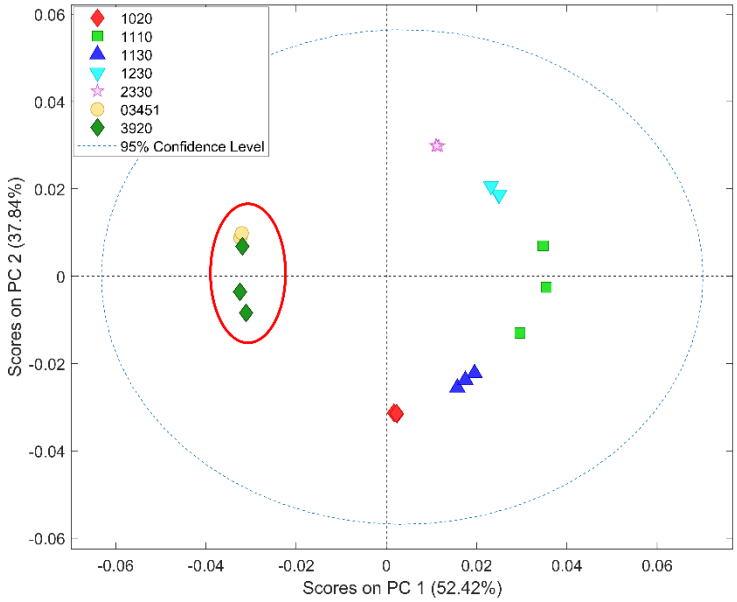


Figure 4. PC1-PC2 score plot for XRF data set.

Figure 5 shows Raman spectra of the samples in spectral range of 100 cm^{-1} to 3500 cm^{-1}. Spectra show sharp features at 142.9 cm^{-1}, 440.3 cm^{-1}, 611.3 cm^{-1}, and 1085.3 cm^{-1} correspond to anatase, rutile (both are crystalline forms of TiO_2), and calcite ($CaCO_3$) respectively (S. Mamedov, 2015). Figure 6 shows the Raman spectra of these oxides. In addition, there are

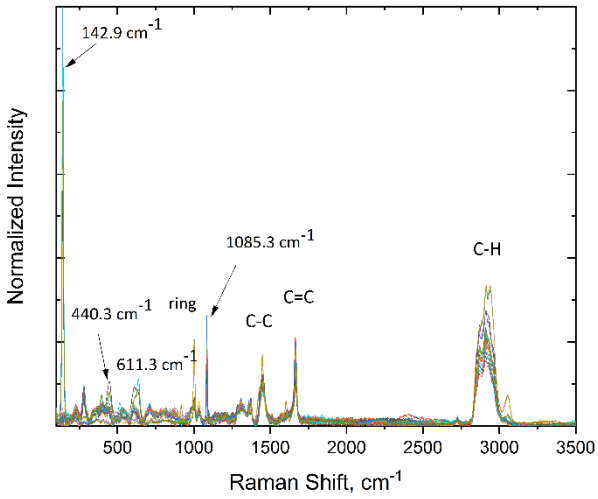


Figure 5. Raman spectra of duct tapes in the spectral range of 100 cm^{-1} to 3500 cm^{-1}.

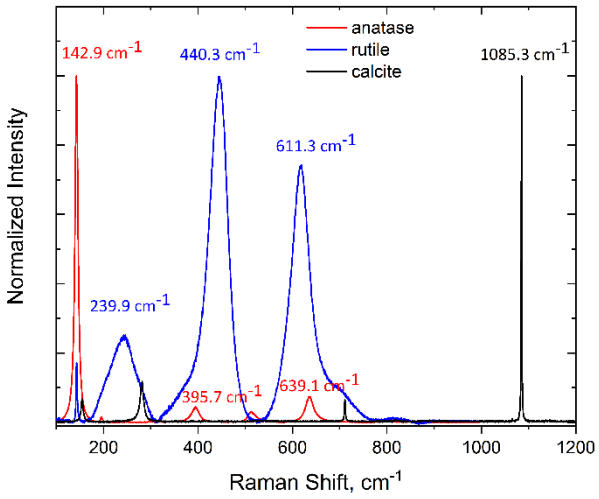


Figure 6. Raman spectra of anatise, rutile and calcite.

several peaks from adhesives that correspond to the ring vibration (around 1000 cm^{-1}), C-C single bond (around 1440 cm^{-1}), C=C double bond (around 1650 cm^{-1}), and C-H stretch region around 2900 cm^{-1}.

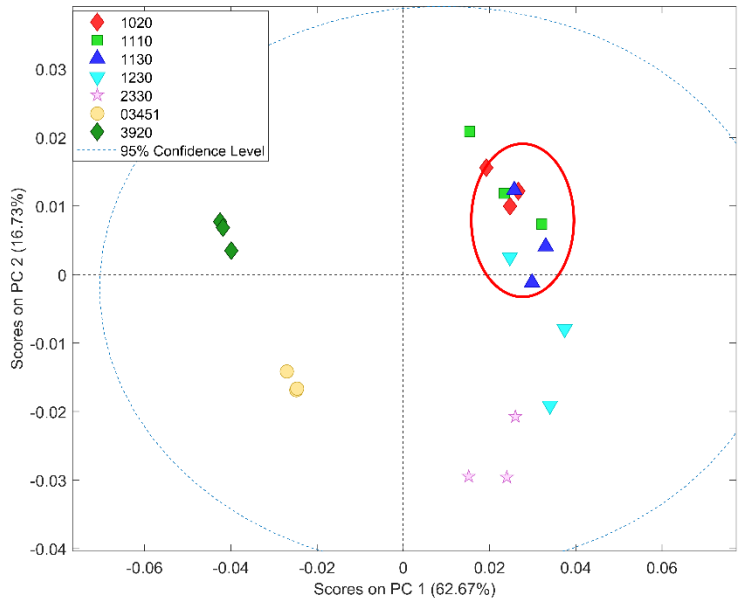


Figure 7. PC1-PC2 score plot for Raman data set.

Mean-centering and normalization preprocessing applied to Raman spectra and PCA analysis were performed on matrix of Raman data. After mean-centering, each row (sample) of the selected data includes only how that row differs from the average sample in the original data matrix. Figure 7 shows PC1-PC2 score plot. The PCA model (based on the three PCs) captures 88.48% of the variance in Raman data set. Red circle in the Figure 7 shows the region, where scores for different tapes overlap.

The tapes that are included in red circle are #1020, #1110, and #1130. Therefore, these duct tapes cannot be fully discriminated based on PCA analysis of Raman spectra.

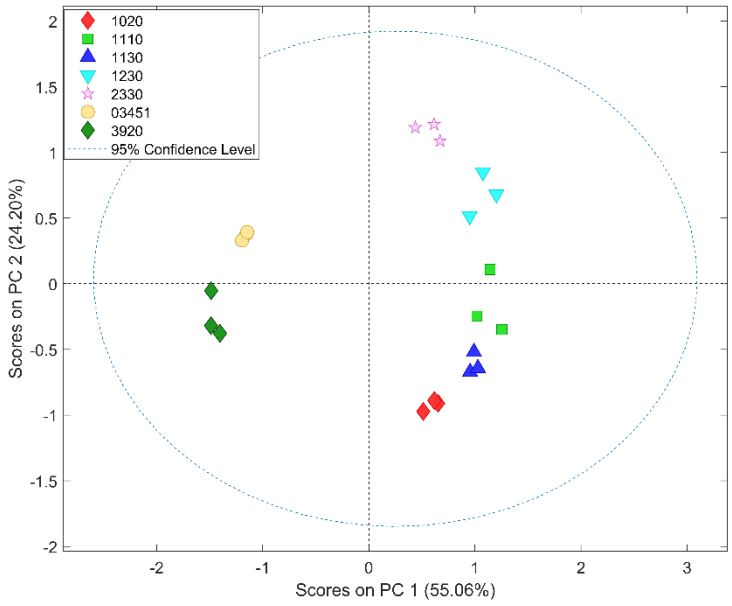


Figure 8. PC1-PC2 score plot for fused XRF-Raman data.

The XRF and Raman spectra were combined according to the methodology described in the introduction. XRF and Raman spectra were preprocessed before the data was fused and the obtained XRF-Raman meta-spectra were processed by PCA. Figure 8 presents PC1-PC2 score plot. The

fused data shows better distribution of classes (duct tapes) when compared with the models obtained using individual methods. With the fused data, all samples for each class (duct tapes) are correctly separated and there are no overlaps between the samples from different duct tapes.

CONCLUSION

Combination of micro-X-ray fluorescence and micro-Raman spectroscopy offers a powerful tool for characterization and identification of the duct tapes. Preprocessing should be performed before concatenation of raw data. Classification of duct tapes was improved using fused data. PCA model performed on fused data are more robust and visual discrimination from class distributions is better with respect to those results obtained by individual classification.

REFERENCES

Mamedov, S. (**2018**). "Industrial and forensic application of micro-XRF: glass analysis", Advances in X-ray Analysis, vol. 61, p. 149-155.

Bretscher, O. (**1995**). *Linear Algebra with Applications* (Upper Saddle River, NJ: Prentice Hall).

Manly, Bryan F. J. (**1986**). *Multivariate Statistical Methods: A Primer.* (Chapman and Hall, London – New York 1986), 159 S. **ISBN** 0-412-28610-6,

Llinas, J., Hall, D.L. (**1998**). "An introduction to multi-sensor data fusion", in: Proceedings of the 1998 IEEE International Symposium on Circuits and Systems 1998, ISCAS '98, vol. 536, 1998, pp. 537–540.

Mamedov, S., Kisliuk, A., Loheider, S., Quitmann, D. (**1995**). "New scanning beam technique for Raman spectroscopy", Applied Spectroscopy, vol. 49, N. 8, p. 1199-1200.

Mamedov, S. (**2015**). "Structural characterization of TiO_2 nanopowders by Raman spectroscopy", in: Mater. Res. Soc. Symp. Proc. Vol. 1806 © 2015 Materials Research Society, DOI: 10.1557/opl.2015.377