Document Title: Application of Laser-Induced Breakdown Spectroscopy to Forensic Science: Analysis of Paint Samples

Author: Michael E. Sigman, Ph.D., Erin M. McIntee, M.S., Candice Bridge, Ph.D.

Document No.: 237839

Date Received: February 2012

Award Number: 2006-DN-BX-K251

FINAL TECHNICAL REPORT

Application of Laser-Induced Breakdown Spectroscopy to Forensic Science: Analysis of Paint Samples

Michael E. Sigman, Ph.D. (Principal Investigator)

Erin M. McIntee, M.S.

Candice Bridge, Ph.D.

[1]National Center for Forensic Science and Department of Chemistry, University of Central Florida, PO Box 162367, Orlando, FL 32816-2367

Award number 2006-DN-BX-K251

Office of Justice Programs

National Institute of Justice

Department of Justice

Abstract:

A comparison has been made between laser induced breakdown spectroscopy (LIBS), X-ray fluorescence (XRF), scanning electron microscopy – energy dispersive X-ray spectroscopy (SEM/EDS) and laser ablation – inductively coupled plasma – mass spectrometry (LA-ICP-MS) for the discrimination of automotive paint samples. All discriminations were performed by hypothesis testing at the $\alpha = 0.05$ significance level using both parametric and nonparametric statistical tests. Discrimination was tested across all paint samples, irrespective of paint color or other features, and in a more forensically relevant fashion, discrimination was determined for only those samples of the same color group, number of paint layers and the presence or absence of effect pigments in the paint. The paint samples came from automobiles manufactured in years 1985 – 2006, representing both original equipment manufacturers paint and repaint samples. The samples came from color groups that included black, blue, green, red, silver, tan and white. A total of 200 paint samples, comprising one group of 110 samples and one group of 90 samples were examined by the different analytical methods. Not all samples were analyzed by each method.

LIBS was determined to have a discrimination power of 90% (10% Type II errors) at a verified 5% Type I error rate. Discrimination was found to be slightly lower (86.6%) among the white color group. Variations in the LIBS signal over time led to same sample discriminations and an artificially high Type I error rate, which was overcome by attention to the sampling protocol and confining spectra collection on samples that were to be compared to a narrow time window. LA-ICP-MS was determined to give the best sample discrimination (100%), with XRF and SEM/EDS giving the lowest discriminations, 85% and 73% total discrimination respectively for each method. The results of this study suggest that LIBS may provide an important screening tool in the analysis of automotive paint samples; however, careful attention to sampling protocols and statistical comparison of samples is recommended. In cases where two samples can not be discriminated, a more accurate comparison method or multiple comparison methods should be sought.

Table of Contents

Executive Summary:

Laser induced breakdown spectroscopy (LIBS) can provide an inexpensive and rapid analytical method for the characterization of a sample's elemental composition. This research examines the application of LIBS for the analysis and pairwise comparison of automotive paint samples for the purpose of discriminating between two samples at a known level of statistical significance. In the absence of a uniquely identifying characteristic (i.e., matching fracture pattern, etc.), items of physical evidence are unlikely to found similar or different with a probability as high as that is enjoyed by biological evidence. This does not preclude the possibility of assessing the similarity of two items of physical evidence on a statistical basis. In order to achieve this goal, the analytical methodology used for the comparison must be characterized as to its ability to discriminate between different samples at a known level of statistical significance. The approach taken in this research was to utilize hypothesis testing to assess the ability of LIBS to distinguish between paint samples that came from different sources.

In the hypothesis testing approach, the null hypothesis (H_0) is that a representative parameter measured for two samples (i.e., $\overline{X}_A$ and $\overline{X}_B$) will be equal ($H_0 : \overline{X}_A = \overline{X}_B$), and the alternate hypothesis (H_A) is that they are not equal. When a Type II error (accepting H_0 when it is false, or false inclusion) holds more serious consequences than a Type I error, as in the case of forensic science, it is common to hold $\alpha = 0.05$. Hypothesis testing is made in this research using parametric tests, which rely on an underlying normal distribution of the parameter of interest, and by a nonparametric permutation method that is free from any underlying assumptions of normality. Parametric tests used

for discrimination were the multivariate analysis of variance (MANOVA) as well as the t-test based on the hit quality index (HQI) and Fisher transformation of the Pearson product moment correlation (Z(r)). The nonparametric method was a permutation test utilizing a test metric based on Z(r).

As a measure of the forensic usefulness of LIBS, the discriminating power of LIBS was compared to that of X-ray fluorescence (XRF), scanning electron microscopy – energy dispersive X-ray spectroscopy (SEM/EDS) and laser ablation – inductively coupled plasma – mass spectrometry (LA-ICP-MS). Automotive paint sample were examined by each analytical method, making use of hypothesis testing at the $\alpha = 0.05$ significance level, which sets the Type I error (same sample discrimination) at 5%. The power of each method is defined as the percent of different sample discrimination, at the defined significance level. The power of any test may be increased (higher different sample discrimination) at the expense of more Type I (false exclusion) errors. While neither one of these two errors is desired, it is important to know both error rates for a given analytical method. The typical approach is to hold the Type I error rate at a specified value while determining the Type II rate.

The discrimination power of each method was evaluated base on several methods of sample comparisons:

1. All pairwise comparisons of the analytical data from different samples (DS), irrespective of the color of the sample, the number of layers of paint or the presence/absence of effect pigments (metal flake, etc.).
2. Pairwise comparisons limited to samples from the same color group, e.g., black samples compared only to other black samples, etc.
3. The most forensically useful approach, where comparisons were limited to samples of the same color group that had the same number of layers of

paint. Samples with effect pigments were not compared to samples that did not contain effect pigments.

All samples came from different sources. Failure to discriminate different samples (DS) constituted a Type II error. The Type I error rate was nominally held at 5%; however, this assumption was tested in several cases by performing repeated analyses on the same sample and checking for statistical discrimination between the same sample (SS) comparisons. The paint samples examined in this research came from automobiles manufactured in years 1985 – 2006, representing both original equipment manufacturers paint and repaint samples. The samples came from color groups that included black, blue, green, red, silver, tan and white. A total of 200 paint samples, comprising one group of 110 samples and one group of 90 samples were examined by the different analytical methods. Not all samples were analyzed by each method. The following table summarized the analytical results.

Summary of different sample (DS) and same sample (SS) discrimination by the analytical and data analysis methods utilized in this research.

				% Discrimination based on:				
Analytical Method	Data Analysis Method	Number of Samples	Number of DS Comparisons	Total†	Color‡	Color + Layers + Effect Pigments	Number of SS Comparisons	Measured Type I Error Rate
XRF	MANOVA	102	5202	83.6	82.4	80.4	-	-
SEM/EDS	MANOVA	102	5202	73.3	71	70.3	-	-
SEM/EDS	HQI	101	5050	84.3	80.5	79.0	-	-
LA-ICP-MS	MANOVA	18	162	100	100	100	-	-
LIBS	MANOVA	20	780	87.3	78.9	84.4	20	5
LIBS	Z(r)	25	300	95.0	-	-	25	12 - 20
LIBS	Z(r)	93	4278	99.2	-	-	-	-
LIBS / log*	nonParametric	90	924	-	-	89.8	90	4.4

* Log transformation of emission intensities used for sample comparisons
† All pairwise comparisons irrespective of color, number of layers of paint or presence/absence of effect pigments
‡ Pairwise comparisons limited to samples of same color group, irrespective of number of paint layers and presence or absence of effect pigments

The results from these studies indicate that LA-ICP-MS analysis of automotive paint samples provides a very high discrimination. Although the total number of samples analyzed by LA-ICP-MS in this study was lower than by the other methods, the different sample discrimination was 100% of 162 different sample comparisons. The high discrimination by this technique is attributed to the reproducibility and accuracy of the method. Due to casework precedence and instrument repair issues, access to this technique was limited; however, the results reinforce the reports from other laboratories. The only issue of potential concern with this method is the untested level of Type I error. When the Type II error goes to 0%, as in this case, caution should be taken to insure that the Type I error is held at the nominal level (5% in our analyses). Although the MANOVA statistical method is considered to be highly robust, this is a potential topic to be considered in future research.

Comparison by MANOVA of the XRF spectra from 102 paint samples (5,202 unique different sample comparisons) resulted in an overall discrimination of 83.6% of the samples. Limiting the different sample comparisons to the more forensically realistic comparison of only samples of the same color, number of layers and presence or absence of effect pigments only lowered the discrimination to 80.4%. A discriminating power of 80% reflects a 20% chance of a committing a Type II error, and the associated social consequences.

SEM/EDS gave only a 73% overall different sample discrimination, based on 5,202 comparisons by hypothesis testing using the MANOVA approach. When the sample comparisons were limited to samples in the same color group, having the same number of layers of paint and the presence or absence of effect pigments, the discrimination remained at 70%. The MANOVA discrimination was based on a series of elements in each of three paint layers, the clearcoat, color and primer. Again, the Type II error is higher than would be acceptable in most cases where the consequences of error can be high.

Analysis of paint samples by LIBS gave better discrimination than XRF, or SEM/EDS, but not a good as LA-ICP-MS; however, in this case the Type I error was verified. LIBS spectra of the paint samples were collected by a drill down technique, whereby spectra from successive laser ablations were collected from a single spot on the sample surface. When a series of 14 emission peaks were chosen from 20 LIBS spectra for the purpose of MANOVA discrimination between the paint samples, an overall discrimination of 87.3% was obtained for 780 different sample pairwise comparisons. Limiting the pairwise comparisons to the forensically more useful comparison to samples in the same color group and having the same number of layers and presence of absence of effect pigments lowered the discrimination slightly, to 84.4%. The Type I error rate was experimentally verified to be 5% based on the duplicate analyses and same sample comparison. Discrimination of paint samples by LIBS full spectral comparison using the HQI or Z(r) similarity metric gave very high different sample discrimination (> 95%) but also resulted in apparent Type I error rates as high as 20%. The apparent high Type I error rates were determined to be due to temporal variations or drift in the LIBS instrument. The problem of high apparent Type I error was corrected by limiting the sample analysis (with duplicate analysis) to a single group of samples, with each group comprised of samples having the same color, number of layers and presence/absence of effect pigments. With this modification in the spectral collection protocol, discrimination of the log transformation of the emission intensities of 90 paint samples gave 89.8% discrimination of the 924 different sample comparisons for samples having the same color, number of layers and presence/absence of effect pigments. The Type I error rate was determined to be 4.4%, very close to the prescribed 5% level. The white paints wiere the least discriminated at 86.6% by the nonparametric permutation test. The Wald test achieved a power of 85.17% and an actual size of Type I error of 3.33%.

These results demonstrate that LIBS spectra obtained by drill down analyses can be used for the discrimination of automotive paint samples. The discriminating power is higher than that obtained by SEM/EDS or XRF; however, drift or temporal fluctuation in the LIBS instrument must be taken into account by a well planned sampling protocol. The results also demonstrate the importance of using hypothesis testing in sample discrimination as a method of controlling the Type I error rate. Likewise, it is important to experimentally know the Type II error rate given a prescribed significance level, e.g., at an α of 0.05, LIBS discrimination of automotive paint samples can be expected to result in approximately 10% Type II errors by the statistical hypothesis testing methods used in this research. The level of Type II errors can be reduced at the expense of greater Type I errors.

While commercially available LIBS instruments are available at substantially lower costs than some other instruments that may be used for automotive paint analysis, these instruments should be carefully assessed for stability and protocols should be put in place to monitor the performance of the instruments over time. Automotive paint discrimination should be conducted with appropriate hypothesis testing and the level of