Document Title: Data fusion from infrared, elemental, MSP and Raman analysis techniques to the maximization of the efficiency of the analytical sequence for the forensic examination of paint evidence

Author(s): Patrick Buzzini, James Curran

Document Number: 311037

Date Received: 12/2025

Award Number: 15PNIJ-21-GG-04163-RESS

Final Research Report

Project title: *Data fusion from infrared, elemental, MSP and Raman analysis techniques to the maximization of the efficiency of the analytical sequence for the forensic examination of paint evidence*

Award Number: 15PNIJ-21-GG-04163-RESS

Authors: Patrick Buzzini, James Curran

Purpose of the project

Forensic paint examinations are typically conducted following an analytical sequence that starts with microscopical methods, followed by instrumental analysis methods. The type and number of methods utilized during analyses primarily depend on the quality and quantity of unknown specimens recovered in the investigated case (i.e., multilayered fragments yield more analytical information than abrasive smears resulting from abrasions). While comparing sets of recovered specimens and reference samples, data obtained from the various techniques are usually evaluated *sequentially* by technique, in the sense that when exclusionary differences are noted between compared sets for a given technique, then there is no need to proceed with further testing, and the examiner typically reports an exclusion. In cases where the forensic scientist may not be able to differentiate the compared sets with a method, a subsequent one is applied until all relevant and available methods have been applied. When the compared sets cannot be distinguished at the end of an analytical sequence, an overall evaluation of the data is made to address the question of a common source between them. While evaluating the question of a common source, a critical facet is the estimation of the chance to randomly observe another coated object in a population of interest that exhibits the same properties than those observed within the questioned specimens as a function, for example, of the observed microscopic features *and* infrared profile *and* elemental profile. Currently, this process is carried out subjectively. To date, any effort made to assist practitioners with implementing objective statistical or probabilistic approaches has been limited to considering data from single methods taken in isolation.

However, to properly reflect current practice in trace evidence comparative examinations, these approaches shall include data from all the techniques that trace evidence examiners use in this process. Moreover, different techniques mainly detect analytical information of the same components; for example, pigments can be characterized using multiple methods. This dependence between analytical methods causes redundancy in the collected data. It consists of a large quantity of data that an examiner can only treat manually and sequentially (technique by technique), where the overall picture of the data for a given sample cannot be fully captured objectively. Finally, it is not uncommon that there are instances where a non-differentiation between compared sets using a given technique may be ambiguous due to the limited extent of analytical information resulting from a shift between the quality and quantity of recovered specimens and reference samples (which could also be affected by a poor collection step at the scene). A subsequent technique may often clarify ambiguous (i.e., not clear-cut) results, but this is not always the case, and these uncertainties are carried through the entire adopted analytical scheme. All these limitations affect the step of evaluating the obtained findings in light of a source question, where currently most trace evidence labs rely on the so-called verbal ‘association’ scales to report their findings. This subjective linear process with all its inherent uncertainties results in a difficult task for the examiner (e.g., choosing a level within a verbal scale) where standardization between different examiners may be difficult to guarantee, resulting in a threat to the reliability of the comparative process.

The present research investigated a variety of approaches intended to develop a holistic, objective, and verifiable comparison process based on data fusion methods that efficiently combine analytical data from different relevant instruments. The investigated approach aimed to produce the architecture to assemble ‘packages’ of relevant data from different techniques to be compared with analogous ‘packages’ of relevant data, that is, for example, a package gathered from recovered specimens to be compared with a package gathered from reference samples. The current phase of this project focused on the most efficient way to build the most informative packages.

More specifically, the present study aimed to define criteria to reliably combine analytical data from different methods of paint examination that include the most selective features (i.e., that exhibit the largest between-sample variation) and if data acquired from different techniques could be treated as independent pieces of information or if they displayed redundancy. Different combinations of analytical sequences for paint examination based on the available methods were also considered.

This project aimed to maximize the efficiency of analytical schemes for forensic paint examinations by means of multi-block data fusion methods applied to infrared, elemental, MSP, and Raman analysis techniques. The stated objectives were:

1. Define the features of interest detected with the proposed techniques.
2. Evaluate the ability of each proposed technique to reliably detect minor components resulting from a heterogeneous paint formulation.
3. Identify dependencies or redundancies of analytical information collected from the different analytical techniques.
4. Evaluate the potential interferences of the adjacent layers within multilayered automotive paint systems in an effort to detect features of interest.
5. Compare variants of multi-block exploratory data analysis, feature selection, and predictive modeling methods to identify the most optimal one(s) for the question of interest of this study.

Project design and methods

The study involved collecting different paint samples from various end uses that show properties useful to trace evidence examiners during comparative paint analysis. The sample sets were organized into four sets and included architectural paints, spray paints, automotive refinishing paints, and multilayered automotive paints. The architectural and automotive refinishing paints consisted of binary mixtures of various colors at different controlled concentrations to evaluate the detectability of the various features inherent to a specific instrumental analysis technique.

The analytical methods included light microscopy methods and cross-sectioning, Fourier transform infrared (FTIR) spectroscopy, scanning electron microscopy-energy dispersive spectroscopy (SEM-EDS), UV-Vis microspectrophotometry, and micro Raman spectroscopy (532 nm and 785 nm laser wavelengths).

Data analysis was divided into a qualitative component and a quantitative component. The qualitative component aimed at identifying and defining the properties and the informative extent of the features detected using the various analytical methods (some practitioners refer to these features as the “chemistries”). Paint is known to be a heterogeneous substance. Therefore, aspects such as the repeatability of the measurements, the degree of spectral overlay, position of bands, and elements were first evaluated for each technique based on criteria typically used by trace evidence examiners. For example, for UV-Vis MSP, the overall pattern, the presence of maxima/minima, and points of inflexion were assessed. For FTIR and Raman spectra, the presence of bands and their relative intensity along the considered spectral range was assessed, and for SEM-EDS, the presence/absence of elements and their relative ratios were evaluated. Table 1 summarizes the qualitative assessment from the practitioner’s standpoint.

Table 1 – Summary of detected features typically considered by practitioners during paint examinations for the instrumental analysis techniques utilized in this project

FTIR	SEM-EDS	UV-Vis MSP	Raman

▪ Identification of binders ▪ Identification of extenders ▪ Identification of pigments (rare) ▪ Record minor IR absorption bands	▪ Identification of pigments ▪ Identification of extenders ▪ Record elements not attributable to compounds	▪ Record minima absorbance in the vis range (color) ▪ Record maxima absorbance in the vis range (complementary color) ▪ Record absorbance in the UV range ▪ Record shoulders and points of flexion	▪ Identification of pigments ▪ Identification of extenders ▪ Record minor scattering bands

The quantitative component aimed at sorting relevant information from the data collected using various data fusion approaches that reflected the qualitative component, attempting to consider the same analytical information as practitioners, but in a whole and objective way. The literature distinguishes data fusion approaches into three levels: low-level, mid-level, and high-level. In low-level data fusion, all observations for each instrument are concatenated into a single block after they are scaled (standardized) and then analyzed. In mid-level data fusion, data standardization, data reduction, and classification within each instrument are carried out, followed by a combination of the classification results. Finally, high-level data fusion involves selecting specific features from each instrument and then synthesizing or combining prediction or classification results. The present study explored all three levels. Different exploratory techniques, as well as classification techniques, were evaluated: traditional data fusion methods known as multiblock exploratory and predictive modeling methods the former including common components and specific weight analysis (CCSWA) (i.e., used to assign score plots to each individual block — specific weight — as well as a combined — common — score plot) and the latter including sequential and orthogonalized partial least squares (SO-PLS). Yet, traditional chemometric methods were also evaluated in a data fusion framework such as partial least square – discriminant analysis (PLS-DA), Sparse LDA, Naïve Bayes, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Random Forest, Functional data analysis (dimension reduction) (FDA) followed by classification (i.e., Random Forest and LDA). For classification, ‘simple majority rule’ was often adopted, a decision rule used in ensemble learning, where multiple classifiers (or models) make predictions, and the final classification is determined by the most common (majority) prediction among them.

Project findings

The present study produced 8,575 single measurements consisting of 7 replicates for each of the four sample sets and five instrumental analysis techniques. The data were collated into data frames (.CSV format) of multivariate data for each set and technique and, when applicable, for color block (i.e., binary mixtures). While the details of the results for each set will be disseminated through publications in peer-reviewed journals, the present report aims to summarize the overall findings, highlighting the trends observed across the four sample sets.

Data fusion is successful if it combines data from multiple signals (multimodal), capturing the key features from each signal characteristic of a specimen by minimizing their intra-source variation and maximizing their inter-source variation (i.e., rarity within a relevant population of interest). The key findings of this study were the following:

1. *High-level data fusion performed better than med-level and low-level for classification purposes*

   Information loss was recorded when all measurements for each instrument were concatenated into a single block after they were scaled (standardized). Then they were analyzed, instead of combining all steps of data standardization, data reduction, and classification within each instrument, before combining the classification results. This observation was made across classification methods and data sets. As an example, for blue spray paint, low-level data fusion using PLS-DA achieved a maximum accuracy of 90% when three derived variables/components were used, while applying it on each block before combining the results yielded an accuracy by majority vote of 96%.

2. *Data fusion using chemometric approaches resulted in more efficient outcomes than multiblock approaches*.

   Multiple methods fall under the umbrella of the data fusion literature, and therefore the limitation of this finding is that only a few of them including the popular multiblock common components and specific weight analysis (CCSWA) for data visualization sequential and orthogonalized partial least squares (SO-PLS), and response-oriented sequential alternation (ROSA) were used. The different classification methods typically applied to chemometric problems varied in their performance, depending on the dataset and the sample colors. Overall, ensemble methods, especially random forest, yielded the best results, and the use of majority vote, also an ensemble method approach, was efficient in demonstrating that combining results from the different instruments captured the highest classification accuracy, compared with the ones obtained from single methods. The study showed that combining data compensated for the lower accuracy of certain methods, like MSP and SEM-EDS for the spray paint sample set or FTIR spectroscopy for the automotive refinishing paint sample set.

3. *The evaluated approaches generally mirrored observations of the spectral patterns by visual inspections.*

   In general, the various classification methods reflected the observations of the collected spectral data. In other words, the quantitative approach mainly aligned with the qualitative approach. Most classifiers proved sensitive to small between-sample spectral differences. However, different instances of incoherence between observations made by visual inspections and the classification outcomes were recorded, and varied between sample sets, colors, and analytical methods. Consider the example of the red spray paint samples: using FTIR, all samples were correctly assigned to their manufacturer class. On top of the binder type, additional components such as styrene and barium sulfate constituted added features that the classifier correctly detected. Moreover, the strong signals from *C.I. Pigment Red 254* in Rust Oleum and the minor peaks of *C.I. Pigment Red 170* in the Colorshot brand also contributed to the correct classification. Moreover. On the other hand, for SEM-EDS, lower rates of misclassification were observed. Besides the most prominent peaks from elements C and O present in all samples, other minor peaks of less abundant elements were clearly noticeable in the EDS spectra. Still, the used classifiers could not detect them. Also, the visual inspection of the MSP spectra showed apparent differences between the spectral curves of the different brands. Still, the classifier could not correctly assign the spectra to the appropriate class, despite a relatively low intra-source spectral variation. This observation was also recorded for the other sample sets. Difficulties inherent to the correct classification of spectra were mainly observed for the binary mixtures, and especially for the automotive refinishing paint sample set, which were mixed with nine proportions (90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:80, and 10:90), as opposed to the architectural paints, which were mixed with five proportions (90:10, 75:25, 50:50, 25:75, and 10:90). One of the key features of the architectural paint sample set was the use of white and black paint in the binary mixtures to color desaturation, resulting in the reduced detection of spectral information about the colored components in the binary mixture, and hence adding noise or interferences in the spectra or reducing the intensity of key bands (including flattening curves). This study confirmed that in many cases, chemometric classifiers do not match the level of detail in spectral patterns that visual inspection still enables.

4. *Combining methods resulted in the redundancy of collected analytical data.*

   While combining data from different instruments offers the advantage of assembling complementary analytical information to increase discriminating capabilities, this advantage comes at a cost. Many instances were found where different methods collected information about the same component. This aspect was anticipated since it was known that both FTIR and SEM-EDS detect extenders and that Raman and SEM-EDS, and occasionally FTIR, detect pigments. In general, SEM-EDS appeared to be more expendable when Raman and FTIR spectroscopy are used in tandem. Despite the limitations discussed above, MSP confirmed its effectiveness when analyzing the UV spectral range (240-380 nm) for the clearcoats of the automotive multilayered paint sample set and demonstrated strong complementarity with FTIR spectroscopy. In this application, Raman proved redundant to FTIR. Overall, Raman confirmed its property to offer complementarities when two different lasers (532 and 785 nm) are utilized, yielding enhanced detection capabilities resolving the issue of uninformative spectra (due to fluorescence or thermodegradation) and detecting different pigments through selective resonance effects.

Another anticipated issue related to redundancy, especially the action to decrease it, concerns computational efficiency. This study focused on small size sample sets, and therefore, this issue did not develop. However, computational limitations may arise when used in casework. The adoption of large databases should be implemented for between-source comparisons, along with the consecutive estimation of the rarity of the collected features while addressing source-level questions.

Implications for criminal justice policy and practice in the United States

This study is expected to significantly aid in the adoption of an objective and numerical method for the joint evaluation of findings from various instruments within an analytical scheme, such as forensic paint examinations investigated in this project. It is believed to be an unavoidable direction since the proposed reported schemes in the forensic community, including trace evidence, are geared toward the adoption of verbal scales that include ‘gray shade’ levels that the expert must assign — subjectively — based on criteria (e.g., strength of a ‘match’ or ‘unusualness’ of observed features).

The current project was the initial step in that direction. The study highlighted the benefits and potential of combining data from different methods, as it is currently done subjectively — but linearly — by the trace evidence examiner. The observed limitations, primarily inherent in the ability of the various classifiers tested to detect minor signals, which are critical for differentiating between different samples, must be addressed. Since such limitations were observed across different classifiers, more work needs to be done on how analytical features from the different analytical methods are selected. Also, the present study focused on spectroscopy, including two routine methods (FTIR and SEM-EDS) and two not universally used methods (Raman and MSP). While this approach enabled us to gauge the complementarities between these methods (especially between Raman and SEM-EDS), the use of separation methods, such as pyrolysis gas chromatography with mass spectrometry (pyGCMS), a technique routinely used in paint examinations, shall be investigated and integrated into the approach. Once optimized, the data fusion approaches evaluated in this study can produce probability distributions that can be implemented in probabilistic approaches such as the likelihood ratio.

Scholarly products

While articles expected to be published in scientific peer-reviewed journals are in preparation, the present study has been the object of the following disseminations:

Buzzini P, Carpenter M, Murali S, Curran J. *Data fusion from infrared, elemental, MSP, and Raman analysis techniques to the maximization of the efficiency of the analytical sequence for the forensic examination of paint evidence*. Presented at the 77th Annual Conference of the American Academy of Forensic Sciences, Baltimore, MD, February 17-21, 2025.

Murali S, Carpenter M, Buzzini P. *On the characterization and differentiation of clearcoats in automotive paint using UV-Visible microspectrophotometry, FTIR spectroscopy, and Raman spectroscopy.* Presented at the 50th Annual Meeting of the Northeastern Association of Forensic Scientists (NEAFS), Atlantic City, NJ, October 21-25, 2024

Carpenter M, Buzzini P. *Assessing redundancy of UV-Vis microspectrophotometry, scanning electron microscopy/energy dispersive spectroscopy, Fourier transform infrared spectroscopy, and Raman spectroscopy to the detection of distinguishing features of colored spray paint*. Presented at the Inter/Micro 2022 conference. McCrone Research Institute, Chicago, IL, June 13-16, 2023.

Carpenter M, Buzzini P. *Data fusion from infrared, elemental, MSP and Raman analysis techniques to the maximization of the efficiency of the analytical sequence for the forensic examination of paint evidence*. In-Person meeting of Organization of Scientific Areas Committees (OSAC), Houston, TX, on April 17-21, 2023.

Carpenter M, Estevanes J, Buzzini P. *Assessing the individual and joint use of ultraviolet/visible (UV/VIS) microspectrophotometry and scanning electron microscopy/energy dispersive spectroscopy (SEM/EDS) to the differentiation of colored spray paint*. Poster presented at the 75th Anniversary Conference of the American Academy of Forensic Sciences, Orlando, FL, February 13-18, 2023.

These can be accessed at: http://www.ifrti.org/research.html.

The datasets used in this study can be retrieved at:
https://doi.org/10.5281/zenodo.17887460