

MIXTURE ANALYSIS OF PAINTINGS WITH RAMAN SPECTROSCOPIC IMAGING AND CHEMOMETRICS

A thesis presented to the faculty of the Graduate School of Western
Carolina University in partial fulfillment of the requirements for the degree
of Masters of Science in Chemistry.

By

Caitlyn E. Phipps

Advisor: Dr. Scott Huffman
Associate Professor of Chemistry
Chemistry and Physics Department

Committee Members: Dr. David Evanoff, Chemistry and Physics
Dr. Carmen Huffman, Chemistry and Physics

October 2014

ACKNOWLEDGMENTS

Firstly, I would like to thank my thesis advisor, Dr. Scott Huffman, ‘Coach’, who not only introduced me to the field of conservation science, but was also there to answer my never-ending questions, fix any instrument problems, and assist in solving all my coding errors, and was always there to provide his unique style of ‘encouragement’ that pushed me to continue on through this research. I would also like to thank my committee members: Dr. Carmen Huffman for her expert critiquing skills and Dr. David Evanoff for addressing any Raman instrument problems that arose. I would also like to thank the other students in the Chemistry and Physics Department, who were always there to provide a humorous, encouraging, and productive environment through this thesis journey. A special thank you goes out to Lynely Hardie, who was always there to help, assist, and answer any questions that I have had during my time here. Lastly I thank JC and my family for encouraging, supporting, and always being proud of me.

TABLE OF CONTENTS

LIST OF TABLES	iv
LIST OF FIGURES	v
ABSTRACT	vi
1 INTRODUCTION TO CONSERVATION SCIENCE	1
1.1 Mass Spectrometry	2
1.2 Spectroscopy	3
1.3 X-ray Fluorescence	5
2 INTRODUCTION TO CHEMICAL RESEARCH	6
3 DEVELOPMENT OF A CALIBRATED RAMAN LIBRARY AND MATCH- ING PROGRAM	8
3.1 Background	8
3.2 Material Identification and Current Libraries	8
3.3 Problems with Online Databases and the Solutions	8
3.4 Experimental	9
3.4.1 Materials	9
3.4.2 Raman Spectroscopy	10
3.4.3 Data Analysis	11
3.4.4 Modified JCAMP - JDX File Format	11
3.4.5 Calibration Procedure	11
3.4.6 Baseline Correction	12
3.4.7 Intensity Normalization	13
3.4.8 Wavenumber Correction	14
3.4.9 Interval Spacing	16
3.5 Master Library	17
3.6 Matching Procedure	18
3.7 Results & Discussion	18
3.7.1 Testing the Library - In House	18
3.7.2 Testing the HQI Against Each Step of the Calibration Procedure	19
3.7.3 Testing the Library Against an External Database	19
3.8 Conclusion & Future Research	21
4 ALS-MCR AND MIXTURE ANALYSIS	23
4.1 Introduction to Mixture Analysis	23
4.1.1 ALS-MCR	23
4.1.2 Spectroscopic Imaging	26
4.2 Experimental	28
4.2.1 Materials	28

4.2.2	Sample Preparation	30
4.2.3	Instrumentation	30
4.2.4	Data Analysis	30
4.2.5	Spectroscopic Imaging	31
4.2.6	ALS-MCR Octave Code	31
4.2.7	Steps for using the ALS-MCR code	32
4.3	Results and Discussion	35
4.3.1	Sample A: Long Scan	35
4.3.2	Sample A: Quick Scan	42
4.3.3	Sample B: Quick Scan	49
4.3.4	Material Identification with Respect to Image Position	55
4.4	Conclusion & Future Research	58
REFERENCES		59
A Acetaminophen Standard Values		62
B Pigment Measurement Parameters		63
C Calibrated Library Spectra		64
C.1	Acid Red 18 (AR18)	64
C.2	Acid Red 52 (AR52)	65
C.3	Pigment Red 254 (PR254)	66
C.4	Pigment Yellow 83 (PY83)	67
C.5	Pigment Yellow 139 (PY139)	68
C.6	Tartrazine	69
C.7	Sunset Yellow	70
C.8	Acid Green 25 (AG25)	71
C.9	Indigo (NB1)	72
C.10	Pigment Blue 15:1 (PB15:1)	73
C.11	Pigment Blue 15:3 (PB15:3)	74
C.12	Acid Violet 17 (AV17)	75
C.13	Pigment Violet 23 (PV23)	76
C.14	Pigment White 4 (PW4)	77
C.15	Pigment White 6 (PW6)	78
D CODE FOR CALIBRATED RAMAN LIBRARY		79
E <i>hnorm</i> CODE		83
F <i>dsmatch</i> CODE		85
G HQI CODE		86

H	ALS-MCR CODE	88
I	ALS-MCR: K GUESS FUNCTION	94
J	ALS-MCR: NON-NEGATIVITY CONSTRAINT	96

LIST OF TABLES

Table 1:	Pigments for preparing the database with the pigment name used in the thesis, the CAS number, other common names that are associated with the pigment, the Color Index (CI) number and the color of the pigment.	10
Table 2:	HQI scores after each step of calibration procedure for an unknown sample of ‘sunset yellow’.	19
Table 3:	Similarity scores between the library and an external database spectrum.	20
Table 4:	Pigment used in paints for ALS-MCR.	28
Table 5:	Measurement parameters for ALS-MCR Samples.	31
Table 6:	HQI scores for K1, K2, and mean of Sample A: long scan. Highlighted in light blue and dark blue are the highest HQI scores of K1, mean and K2, NB1, NB1 and PB15:1, respectively. The red highlight represents the second highest HQI score for the mean spectra, PB15:3.	41
Table 7:	HQI scores for K1, K2, and mean of Sample A: quick scan. Highlighted in dark blue is the highest HQI score for K1, identified as PB15:1. Highlighted in light blue is highest HQI score of K2 and the highest HQI score in the mean spectrum, NB1. Also highlighted in red is the second highest mean spectrum, identified as AR18.	48
Table 8:	HQI scores for K1, K2, and mean of Sample B: quick scan. Highlighted in dark yellow is the highest similarity score for K1, sunset yellow. Highlighted in orange is the highest similarity score for K2 and the mean spectrum, tartrazine. Also highlighted in dark yellow is the second highest HQI score in the mean spectrum, sunset yellow.	55
Table 9:	HQI scores with respect to each image position for Sample A: long scan.	56
Table 10:	HQI scores with respect to each image position for Sample A: quick scan.	57
Table 11:	HQI scores with respect to each image position for Sample B: quick scan.	58
Table A.1:	NIST acetaminophen standard. ¹⁸	62
Table B.2:	Pigment measurement parameters for library.	63

LIST OF FIGURES

Figure 1:	Raman spectrum of Tartrazine and TiO ₂ as a mixture showing the spectral separation of the two components, an organic and inorganic dye, respectively.	2
Figure 2:	Calibration Procedure.	12
Figure 3:	Original Raman spectrum and baseline corrected Raman spectrum of sunset yellow. The baseline corrected Raman spectrum has an offset by adding a constant to the intensity values of the baseline of the baseline corrected Raman spectrum.	13
Figure 4:	Intensity corrected Raman spectrum of sunset yellow using Equation 4.	14
Figure 5:	Wavenumber corrected Raman spectrum of sunset yellow with an inset showing a magnified portion of the spectrum to better illustrate how the wavenumber correction affected the Raman spectrum. . . .	16
Figure 6:	Interval spacing corrected Raman spectrum of sunset yellow, which is the spectrum used in the library database.	17
Figure 7:	External database spectrum and library spectrum of PW4.	21
Figure 8:	Outline for ALS-MCR.	24
Figure 9:	Spectroscopic image of Sample A, PB15:1 and NB1, with a ‘clean’ interface between the two paints, represented as a 3D surface plot with the image position along the x axis, the wavenumber information along the y axis, and intensity at each x and y, along the the x axis.	27
Figure 10:	Spectral overlap map of each library spectrum compared against every other library spectrum. A color bar is shown with blue representing a low HQI score and dark red representing the highest HQI score. . .	29
Figure 11:	Example of preparation for Sample A and B, with homogeneous paints on either side of a heterogeneous interface.	30
Figure 12:	Procedure for using spectroscopic imaging and the ALS-MCR algorithm for identifying pure components in mixtures.	32
Figure 13:	Before spectroscopic image was done for Sample A: long scan.	35
Figure 14:	After spectroscopic image was done for Sample A: long scan.	36
Figure 15:	Library spectrum for NB1 (blue) and the calculated pure component spectrum of K1 (green). The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline. . .	37
Figure 16:	Library spectrum for PB15:1 (blue) and the calculated pure component spectrum of K2 (green). The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.	38
Figure 17:	Library spectrum for Indigo(blue), PB15:1 (green) and the mean spectrum (red). The calculated pure component spectrum and mean spectrum is offset by adding a constant to the intensity values of the baseline.	39
Figure 18:	Relative concentration profile of NB1 (blue) and PB15:1 (green) with connecting lines for visual aid.	40

Figure 19:	Library spectrum of PB15:1 (red) and PB15:3 (blue) with spectroscopic differences noted by solid black lines.	42
Figure 20:	Approximate area imaged for Sample A: quick scan with x's indicating the approximate image position of each measurement.	43
Figure 21:	Library spectrum for PB15:1 (blue) and the calculated pure component spectrum of K1 (green) for Sample A: quick scan. The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.	44
Figure 22:	Library spectrum for NB1 (blue) and the calculated pure component spectra of K2 (green) for Sample A: quick scan. The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.	45
Figure 23:	Library spectrum for PB15:1 (blue), NB1 (green) and the mean spectrum (red) for Sample A: quick scan. The calculated pure component spectrum and mean spectrum is offset by adding a constant to the intensity values of the baseline.	46
Figure 24:	Relative concentration profile of PB15:1 (blue) and NB1 (green) with connecting lines for visual aid.	47
Figure 25:	Approximate area imaged for Sample B: quick scan with x's indicating the approximate image position of each measurement.	49
Figure 26:	Library spectrum for sunset yellow (blue) and the calculated pure component spectrum of K1 (green) for Sample B: quick scan. The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.	51
Figure 27:	Library spectrum for tartrazine (blue) and the calculated pure component spectra of K2 (green). The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.	52
Figure 28:	Library spectrum for sunset yellow (blue), tartrazine (green) and the mean spectrum (red) for Sample B: quick scan. The calculated pure component spectrum and mean spectrum is offset by adding a constant to the intensity values of the baseline.	53
Figure 29:	Relative concentration profile of sunset yellow (blue) and tartrazine (green) for Sample B: quick scan with connecting lines for visual aid.	54
Figure C.1:	Calibrated Raman spectrum for AR18.	64
Figure C.2:	Calibrated Raman spectrum for AR52.	65
Figure C.3:	Calibrated Raman spectrum for PR254.	66
Figure C.4:	Calibrated Raman spectrum for PY83.	67
Figure C.5:	Calibrated Raman spectrum for PY139.	68
Figure C.6:	Calibrated Raman spectrum for tartrazine.	69
Figure C.7:	Calibrated Raman spectrum for sunset yellow.	70
Figure C.8:	Calibrated Raman spectrum for AG25.	71
Figure C.9:	Calibrated Raman spectrum for NB1.	72

Figure C.10: Calibrated Raman spectrum for PB15:1.	73
Figure C.11: Calibrated Raman spectrum for PB15:3.	74
Figure C.12: Calibrated Raman spectrum for AV17.	75
Figure C.13: Calibrated Raman spectrum for PV23.	76
Figure C.14: Calibrated Raman spectrum for PW4.	77
Figure C.15: Calibrated Raman spectrum for PW6.	78

ABSTRACT

MIXTURE ANALYSIS OF PAINTINGS WITH RAMAN SPECTROSCOPIC IMAGING AND CHEMOMETRICS

Caitlyn Phipps, Masters of Science in Chemistry

Western Carolina University (October 2014)

Advisor: Dr. Scott Huffman

Conservation science is a field that utilizes scientific analysis for material identification to understand, conserve, and preserve culturally and/or historically important objects. Within conservation science, we are focused on material identification using Raman spectroscopy. This work presented is divided into two projects, both utilizing Raman spectroscopy and chemometric algorithms for identification of pigments that are found in paintings. In the first project, we developed a calibrated Raman spectral library and spectral comparison program, which includes 15 pigments that can be found in modern artist materials such as paints, markers, and ink pens. The second project involved the use of mathematical algorithms to digitally extract pure component spectra from Raman spectra of mixtures. Subsequently, by combining the tools developed in these two projects, we were able to identify, without physical separation, the pigments utilized in the creation of paintings.

CHAPTER 1 INTRODUCTION TO CONSERVATION SCIENCE

Conservation science involves the use of scientific techniques to study, protect, and preserve culturally and/or historically important objects (CHOs). A primary research goal in conservation science is the identification of materials that can be used for authentication,¹⁻³ determining optimal display conditions,^{4,5} and degradation kinetics,⁶ with media ranging from textiles, sculptures, paintings and works on paper (books, prints, scrolls).

Measurement of CHOs generally requires that techniques meet two constraints: be (1) non- or minimally destructive to the CHOs and (2) performed *in situ*. *In situ* is defined as the location of the artwork. For paintings this is commonly a gallery. The work presented herein focuses on material identification in paintings, which are inherent mixtures containing pigments, mediums, binders, and fillers. To determine the components of a mixture, physical separation or spectroscopic separation can be used. Some common analytical instrumentation for research with CHOs is discussed in Sections 1.1 - 1.3. Physical separation allows for the individual components to be identified, but requires samples to be taken from the artwork, violating constraint number one by being destructive to the CHO. Spectroscopic separation, the identification of components based on the spectroscopic contributions from each component in a single spectrum, works for elemental analysis, using for example ICP-OES and XRF,⁷ and sometimes molecules such as TiO_2 and tartrazine in a Raman spectrum as shown in Figure 1. While mixtures containing both inorganic and organic pigments can be easily identified with Raman spectroscopy due to the spectral separation of their bands, this is generally not possible with modern organic dyes.

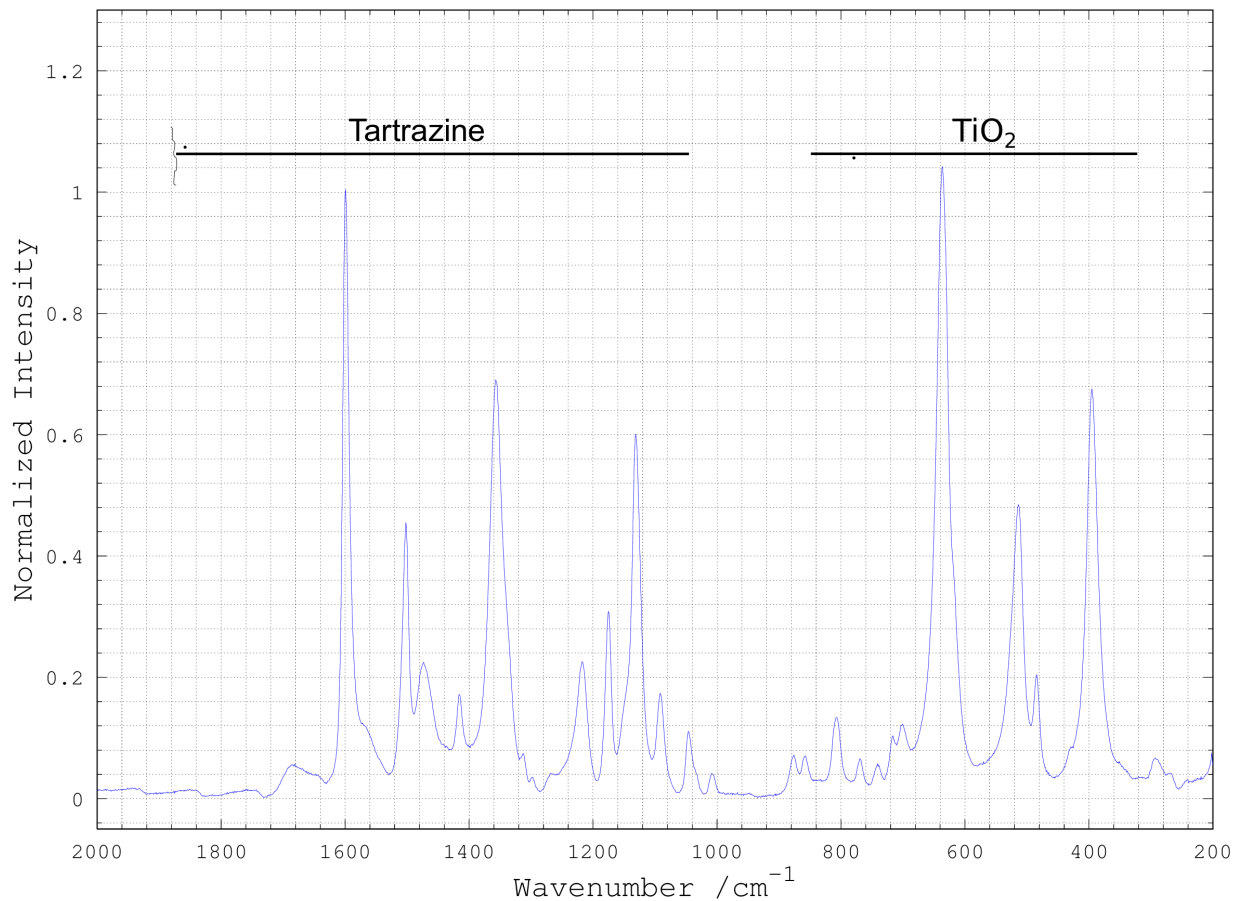


Figure 1: Raman spectrum of Tartrazine and TiO₂ as a mixture showing the spectral separation of the two components, an organic and inorganic dye, respectively.

1.1 Mass Spectrometry

Two common mass spectrometry techniques utilized in conservation science are liquid-chromatography (LC) and gas chromatography (GC), both of which can be used for qualitative and quantitative analysis of various materials used in CHOs. Liquid chromatography - mass spectrometry (LC-MS) separates the components based on their affinity to the stationary phase or mobile phase in the column. These purified components enter a mass spectrometer producing mass-to-charge (m/z) ratio fragmentation patterns, that can be

compared against libraries of known materials. Gas chromatography - mass spectrometry (GC-MS) relies on separation of components based on their volatility in the GC portion, resulting also in fragmentation patterns that are used to identify materials from the MS portion. Both LC-MS and GC-MS are physical separation techniques, that provide high information content, allowing for identification of components in mixtures, but require samples to be removed from the CHO and then destroyed during analysis. While these are both ideal options when performing mixture analysis, they can not be used on materials with limited sample size due to the destructive nature of the analysis.

1.2 Spectroscopy

Infrared (IR) spectroscopy and Raman spectroscopy are complementary techniques giving information about the vibrations and rotations of the molecule. The number of types of vibrations in a molecule, called normal modes, can be calculated using Equation 1 for non-linear (n_{nl}) molecules

$$n_{nl} = 3N - 6 \quad (1)$$

where N represents the number of atoms in the structure. Normal modes have unique vibrational frequencies (ν) related to the change in energy based upon the molecule's interaction with a photon, in which a vibrational transition from the ground state ($\nu = 0$) to the first excited state ($\nu = 1$) occurs. Using the harmonic oscillator approximation, spacing between the energy levels (ΔE) is

$$\Delta E = h\nu \quad (2)$$

where h is Planck's constant. The values of ΔE relate to the specific type of bonds, which is why bonds have a specific ν .⁸

IR spectroscopy involves the absorption of infrared light by the molecule as a function of frequency (ν). The region most commonly associated with IR spectroscopy is mid-infrared,

spanning across a wavenumber range of 400 - 4000 cm^{-1} . While IR spectroscopy is useful for identifying materials that are protein and cellulose based, identification of common pigments used in modern and traditional art such as sulfide pigments (cadmium yellow, red pigments and vermilion) are difficult to identify due to their bands in the far-IR region (below 400 cm^{-1}).⁹ While most materials can be identified with IR spectroscopy, this analytical technique is not commonly thought of as a method for mixture analysis because of overlapping absorption bands from all the components.

Traditionally, IR spectroscopy was the most common vibrational spectroscopic technique used in material identification. Recently, a shift towards using Raman spectroscopy has occurred due to its ability to measure inorganic and organic materials, the ability to perform *in situ* measurements that are non-destructive to the artwork, and the development of portable Raman spectrometers, the stability and ease of use of modern lasers, and the relatively fast measurement time. Raman spectroscopy is a type of vibrational spectroscopy based on inelastic scattering of monochromatic light (laser). Raman spectra consist of sharp bands whose wavenumber position and intensity are characteristic of a specific molecule. The position of the bands in the spectrum are governed by the reduced mass (μ) of the atoms involved in the vibrational mode and the strength of the bonds involved in the vibrations characterized by a force constant (k), as shown in Equation 3.

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \quad (3)$$

Specifically, more massive atoms and weaker chemical bonds result in lower energy bands in the Raman spectrum. Since each molecule has a different atomic arrangement and connectivity, a unique Raman spectrum results. This unique spectrum, therefore, can be used to characterize molecular structures and identify chemical compounds. Chemical identification from a Raman spectrum can be achieved by empirical interpretation of the band positions

or more commonly by comparison to a known set of spectra in the form of a library. Ideally, the library spectra are well-documented (measurement parameters, sample parameters) and low noise spectra.

Although there are many advantages to Raman spectroscopy, some disadvantages have prevented it from becoming the sole analytical tool for identification of materials. One of the most common problems is sample fluorescence, which can be reduced with use of the correct excitation wavelengths, laser power at the sample, acquisition settings, and with SERS, but at times it can not be eliminated.¹⁰ Another problem with Raman spectroscopy is that it does not inherently handle mixture analysis without physical separation.

1.3 X-ray Fluorescence

X-ray fluorescence, XRF, is used for qualitative and quantitative elemental analysis. XRF works by characterizing the unique set of x-rays emitted by an element through electronic relaxations in the inner orbitals of atoms. A sample is bombarded with x-rays from an x-ray tube or a sealed capsule of radioactive material, inducing this electronic transition, and the emission is then measured and the elements spectroscopic “fingerprint” is recorded. An example of XRF usage in CHO analysis is found in dating paintings. TiO_2 was not used in paintings until it became available commercially around the 1950’s, so if present in paintings that are reportedly older than 1940, they are either forged or adulterated.¹ The Ti is relatively easy to identify with XRF, but its polymorphs, rutile or anatase, are indistinguishable with XRF. While XRF can be both a portable and non-destructive technique, it only allows for identification of elements present, not the molecules in which they are found, meaning other techniques, such as the ones mentioned previously, are needed for further identification.

CHAPTER 2 INTRODUCTION TO CHEMICAL RESEARCH

Presented in this thesis is a method to identify modern organic pigments *in situ* on paintings using Raman spectroscopy and chemometric algorithms. This project addresses the need for a mixture analysis method, under the constraints of conservation science. Moreover, the method developed utilizes a digital extraction procedure to automatically identify materials and therefore can be scaled to huge spectroscopic imaging sets that would result from whole painting measurements. One theoretically possible method, would be to record every Raman spectrum of every possible mixture but this is practically impossible because of the amount of possibilities. Listed below are the steps for the mixture analysis developed in this thesis.

1. Spectroscopic mapping or imaging of the surface of the painting.
2. Digital extraction/separation/resolution of the pure component spectra in the mixture using the chemometric algorithm described in Chapter 4.
3. Automatically identifying materials from their pure component spectra using the library searching algorithm described in Chapter 3.

The primary goal of this thesis was mixture analysis and identification of pigments in paintings (Chapter 4) but first a small library of pigments and a spectroscopic matching algorithm had to be developed. The library and matching program is briefly described below and entirely in Chapter 3. The Raman library was developed following a calibration procedure to prevent variations in Raman spectra from day-to-day measurements and to standardize the Raman data, both in wavenumber and intensity. The calibration procedure allows for comparison of measurements made on a research grade Raman spectrometer and measurements from a hand-held Raman spectrometer, both of which are used when working

with CHOs. Once the Raman library was created using the calibration procedure, an automatic identification algorithm was developed. This created a method to automatically search the Raman library spectra and identify the most probable match when trying to identify an unknown pigment. Both the calibrated Raman library and the matching algorithm were both necessary steps for mixture analysis (Chapter 4).

As mentioned previously, paintings are made with multiple components including pigments, media, binders, and fillers. This research focuses on the identification of pigments in mixtures using Raman spectroscopy. Mixtures such as TiO_2 and tartrazine can be identified from the inherent non-overlapping difference of the bands of organic and inorganic compounds, but modern organic pigments all have bands that often overlapping in the same range. The high spectral overlap of organic pigment mixtures requires the use of a chemometric routine to extract the pure component spectra, which are then matched to the library.

We are using spectroscopic imaging to evaluate a chemically heterogeneous areas of the painting. The chemometric routine that was utilized was alternating least squares - multivariate curve resolution (ALS-MCR)¹¹⁻¹⁴ which decomposes the data into pure component spectra and relative concentration profiles from each chemical species, which are then automatically matched to the library of pigments.

CHAPTER 3 DEVELOPMENT OF A CALIBRATED RAMAN LIBRARY AND MATCHING PROGRAM

3.1 Background

As mentioned in Chapter 1, one of the primary duties of conservation scientists is the characterization of materials used in CHOs. Herein a focus on one aspect of this characterization, the identification of pigments used in paintings. While material identification in paintings involves the analysis of mixtures, which is discussed in Chapter 4, first a calibrated Raman library and a matching algorithm that will be used to identify the components from the mixture analysis was developed.

3.2 Material Identification and Current Libraries

Currently, when a Raman spectrum of a substance is measured, an analyst can use an in-house database or an online reference library to identify the material.¹⁵⁻¹⁷ While some in-house libraries may also have an automatic identification algorithm, most identifications are done through visual comparison of the unknown material spectrum with database spectra, which is both time consuming and provides no statistical value for the quality of the comparison.

3.3 Problems with Online Databases and the Solutions

When trying to identify an unknown material, a user will often use an online database of artist materials with visual comparison against the Raman spectrum to identify the unknown material. A problem with the available online databases is the lack of a standardized procedure for measuring the spectra of the materials included in the database. While most databases list the excitation wavelength, other parameters such as the sample state (liquid,

solid, paint (oil+pigment)), measurement parameters (acquisition time, repetitions, laser power at the sample, grating) and data processing procedures (baseline correction, smoothing) are not given. In order to standardize Raman measurements, taking into account the aforementioned parameters, a calibration procedure can be used. The calibration procedure in this thesis adjusts both wavenumber and intensity values, standardizing them with the use of an accepted NIST Raman standard, acetaminophen.¹⁸

Once the calibration procedure was developed for acquiring Raman spectra and the calibrated Raman library was created, visual comparisons of an unknown to the library for spectral matches are possible but become time consuming once the library grows and as the number of unknown Raman spectra grows. To make matching of an unknown to the library faster, a matching algorithm that automatically searches the library was created. While this algorithm not only makes identifying an unknown almost instantaneous, the matching algorithm also provides a confidence associated with the ‘match’.

3.4 Experimental

3.4.1 Materials

Pigments (PR254, PY83, PY139, PB15:1, PB15:3, and PV23) listed in Table 1 were purchased from Karma Pigments (Montreal, Quebec, Canada). Titanium dioxide (anatase) (CAS 13463-67-7) was purchased from Acros Organics (Fair Lawn, NJ), and zinc oxide (CAS 1314-13-2) was purchased from Fisher Scientific (Fair Lawn, NJ). Indigo was synthesized and purified in the laboratory using the procedure described in *Experimental Organic Chemistry*.¹⁹ Pigments (AR52, tartrazine and sunset yellow) were purchased from Abbey Colors (Philadelphia, PA), AV17 was purchased from MP Biomedicals (Santa Ana, CA), and AR18 was purchased from Sigma Aldrich (St.Louis, MO). All were used without further purification. The NIST Raman standard acetaminophen (CAS 103-90-2) was purchased from Mallinckrodt Pharmaceuticals USA (St. Louis, MO) and used without further purification.

Table 1: Pigments for preparing the database with the pigment name used in the thesis, the CAS number, other common names that are associated with the pigment, the Color Index (CI) number and the color of the pigment.

Pigment Name	CAS	Other Names	CI Number	Color
AR18	2611-82-7	Ponceau 4R	16255	Red
AR52	3520-42-1	Sulforhodamine B	45100	Red
PR254	84632-65-5	Fast Brilliant Red S2BL	56110	Red
PY83	5567-15-7	Diarylide HR	21108	Yellow
PY139	36888-99-0	Paliotol Yellow	56298	Yellow
Tartrazine	1934-21-0	FD&C Yellow 5	19140	Yellow
Sunset Yellow	2783-94-0	FD&C Yellow 6	15985	Yellow
AG25	4403-90-1	Alizarin Cyanine Green	61570	Green
NB1	482-89-3	Indigo	73000	Blue
PB15:1	12239-87-1	Copper Phthalocyanine	74160	Blue
PB15:3	147-14-8	Beta Copper Phthalocyanine	74160	Blue
Acid Violet 17	4129-84-4	Food Violet 1	42650	Violet
PV23	6358-30-1	Permanent Violet RL	51319	Violet
PW4	1314-13-2	Zinc White	77947	White
PW6	13463-67-7	Titanium White	77891	White

3.4.2 Raman Spectroscopy

For the creation of the calibrated Raman library, Raman spectra of the the pigments in Table 1 were measured using multiple excitation laser lines to determine the optimal measurement parameters. All Raman spectra were measured on a Horiba Jobin Yvon LabRamHR Raman microscope using an 10x objective with a numerical aperture (NA) of 0.25. The excitation lasers included a Littman/Metcalf diode laser producing 785 nm (Sacher Lasertechnik Group, Marburg, Germany), the 632 nm line from a helium neon (HeNe) laser (Horiba Jobin Yvon, Edison, NJ), a neodymium-doped YAG (Nd-YAG) producing a 561 nm line (Oxxius, Lannion, France), and the 514 nm, 488 nm, and 458 nm lines from an argon ion laser (Spectra-Physics, Santa Clara, Ca). Also, all spectra were measured with 1800 groove mm^{-1} grating, and a TE-cooled CCD detector. The Raman software used was HORIBA Scientific LabSpec 6.

3.4.3 Data Analysis

Octave version 3.2.4 was used for all data processing.²⁰ All data processing was written in house. Also, all code is written to be MATLAB compatible.

3.4.4 Modified JCAMP - JDX File Format

To make this library universal, the file format used for all of the spectra acquired is the JCAMP-JDX version 5.01. JCAMP-JDX file format allows for transfers of spectral information, data and meta-data between users/instruments to be completed error-free. Specifically, we implemented the JCAMP-JDX format proposed by Infrared and Raman Users Group (IRUG) which contains additional meta-data that is important for research with CHOs.²¹

3.4.5 Calibration Procedure

In order for the database to be comparable to spectra acquired from other Raman spectrometers, a calibration is needed to adjust for differences between Raman spectrometers, instrumental setup, sample preparation and other factors that affect Raman spectra.²² This procedure includes the use of the NIST suggested standard reference material for Raman spectroscopy: acetaminophen.¹⁸

The calibration follows the steps shown in Figure 2. An acetaminophen spectrum was acquired daily. The corrections that are used to adjust the acetaminophen spectrum were applied to all of the spectra measured on the same day. This procedure allows the database to be compared with other Raman spectrometers as long as the same calibration procedure and standard is used for sample measurements.

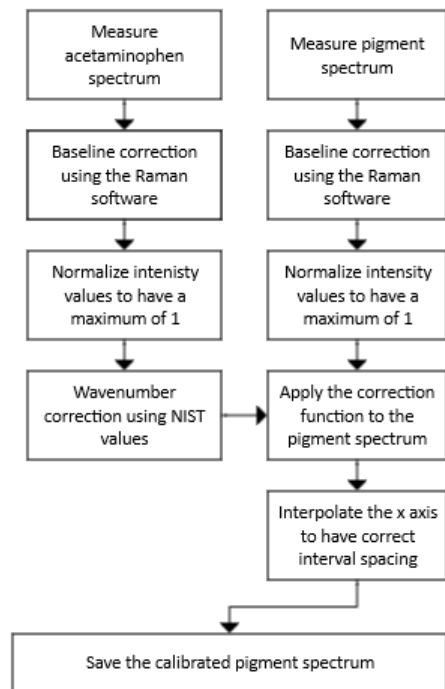


Figure 2: Calibration Procedure.

3.4.6 Baseline Correction

After the spectrum of any sample (including the acetaminophen standard) was measured, its baseline was corrected using the LabSpec software. The baseline correction is necessary to remove undesirable background noise such as fluorescence. The software used a polynomial function with the degree, up to the 24th degree and a maximum of 72 points in the polyfit line, which was optimized to give the best baseline correction for each particular spectrum. Figure 3 shows the original spectrum of sunset yellow and the baseline corrected Raman spectrum of sunset yellow using an 8 degree polynomial fit.

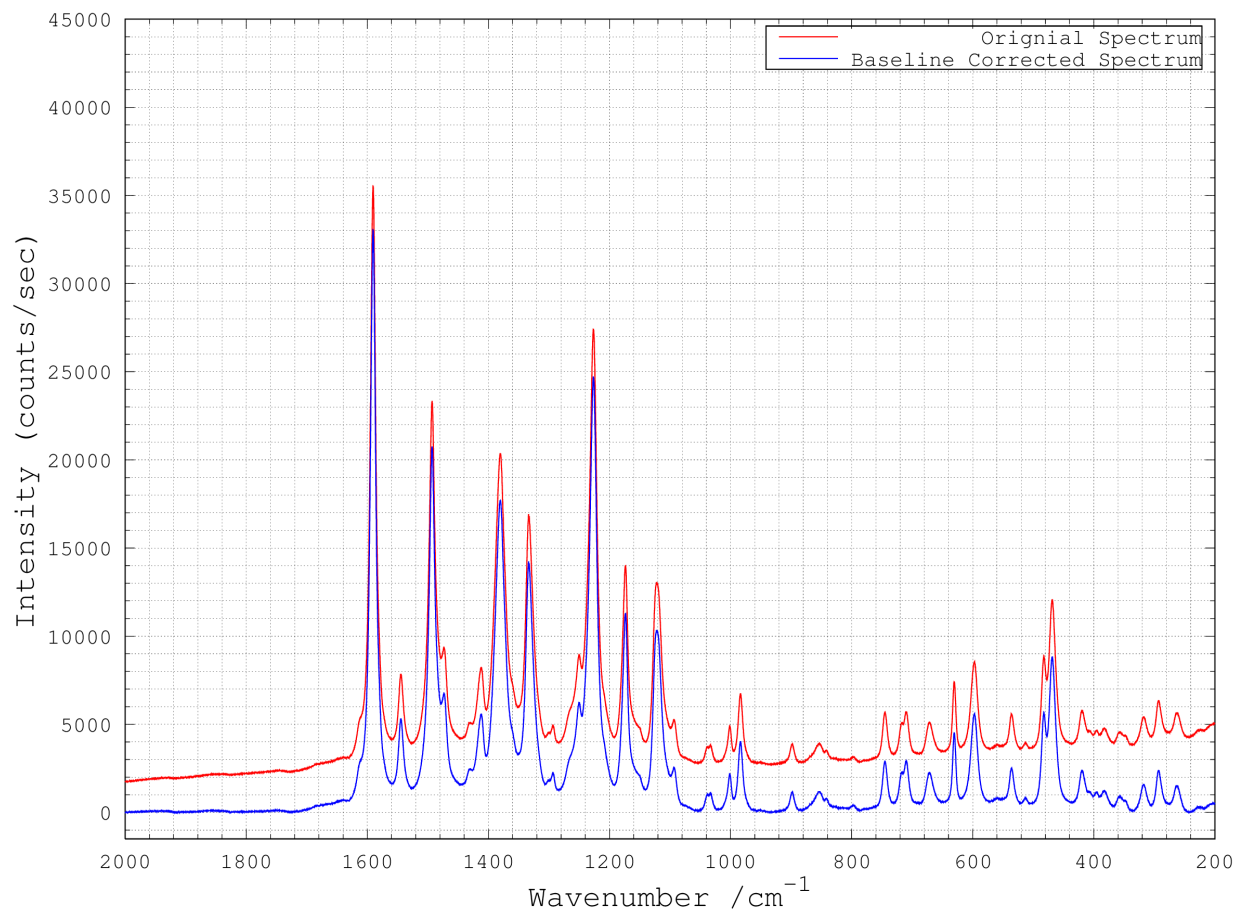


Figure 3: Original Raman spectrum and baseline corrected Raman spectrum of sunset yellow. The baseline corrected Raman spectrum has an offset by adding a constant to the intensity values of the baseline of the baseline corrected Raman spectrum.

3.4.7 Intensity Normalization

Each spectrum is intensity normalized to correct for day to day variations of laser power at the sample and to scale the intensity values for spectra in the library between zero and one. The intensity normalization helps to increase similarity scores between an ‘unknown’ spectrum and a library spectrum. Equation 4 shows the intensity normalization function, with A representing the intensity values of the spectrum, min representing the lowest intensity

value in the spectrum and max representing the highest intensity value in the spectrum. Figure 4 shows the intensity normalized Raman spectrum of sunset yellow.

$$A = \frac{A - \min(A)}{\max(A - \min(A))} \quad (4)$$

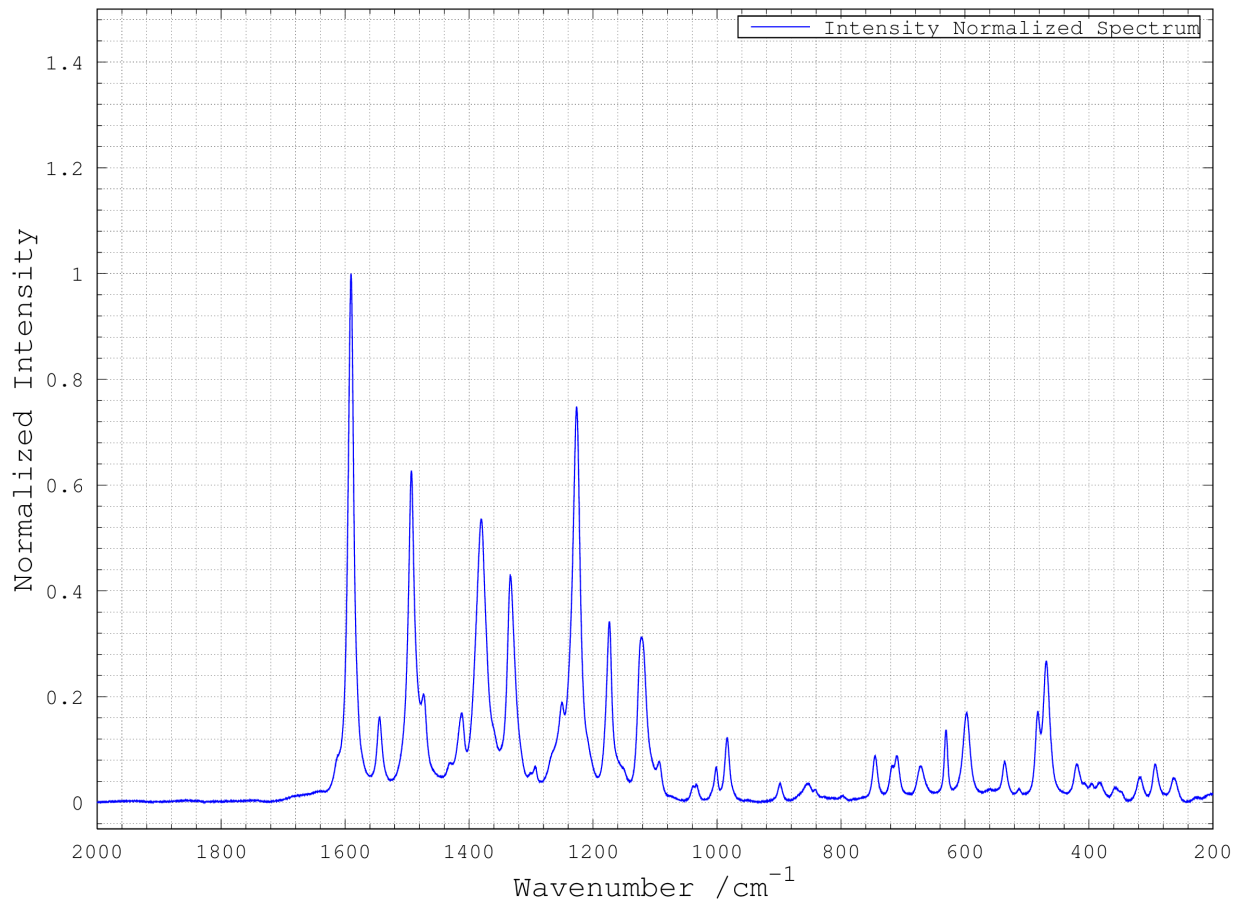


Figure 4: Intensity corrected Raman spectrum of sunset yellow using Equation 4.

3.4.8 Wavenumber Correction

The acetaminophen spectral band positions are corrected to match the NIST band wavenumbers listed in bold in Table A.1 in Appendix A. These bands have been chosen because they

cover a wide wavenumber range, are spectrally isolated and have varying intensities. The wavenumber correction adjusts for thermal drift of optics caused by expansion and contraction of optical components, which can lead to wavenumber drifts as large as $0.05 \text{ cm}^{-1}/^{\circ}\text{C}$.²³ This shift correction involved linearly fitting (using a *polyfit* function code found in Appendix D, lines 79-83) the four bands from a measured acetaminophen spectrum with the NIST band positions. The linear fit gives a correction parameter (slope and intercept, *px* lines 77-79). Using *polyval* (lines 80-81) to apply the linear fit, for both the acetaminophen spectrum and all the spectra measured on that day, are wavenumber calibrated. In Figure 5 is shown the wavenumber corrected Raman spectrum of sunset yellow.

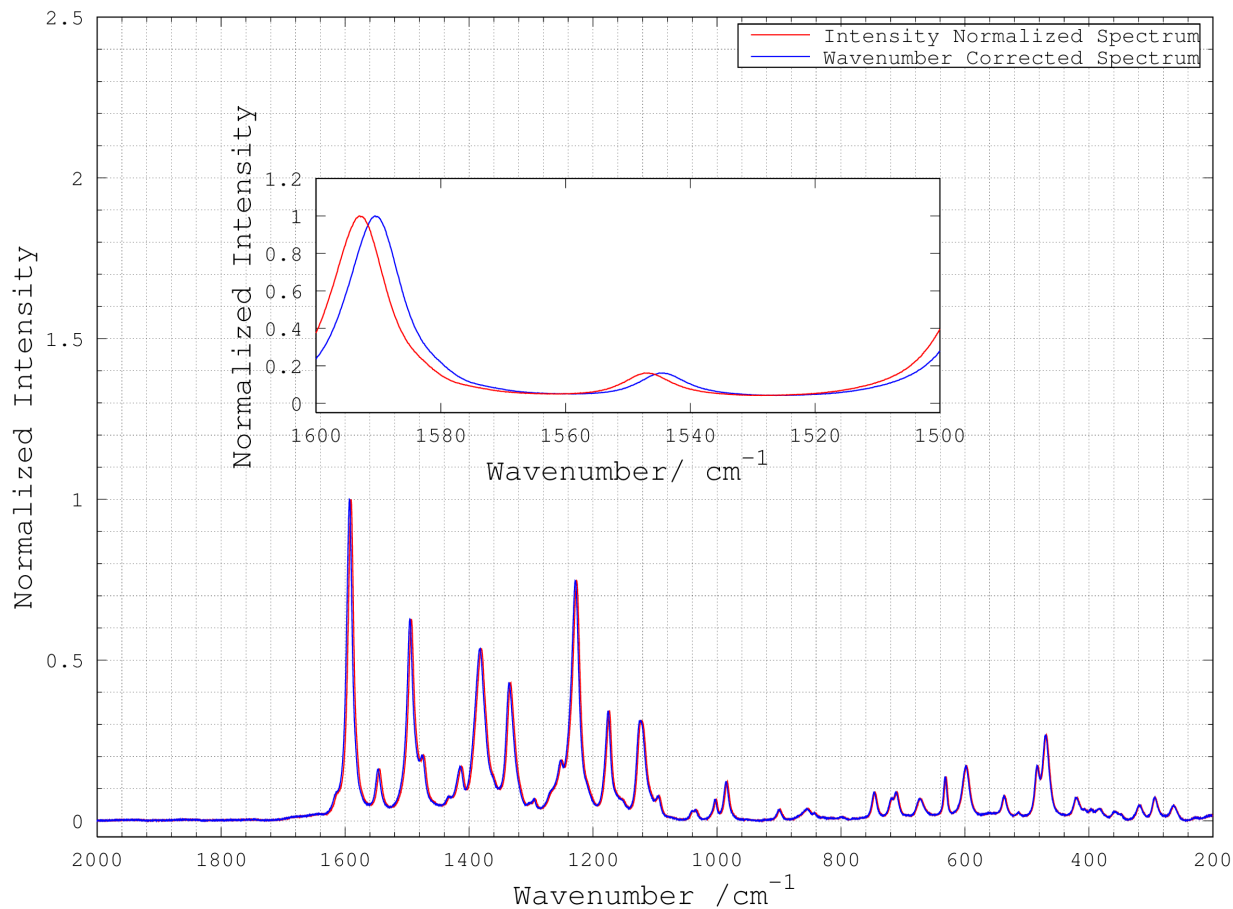


Figure 5: Wavenumber corrected Raman spectrum of sunset yellow with an inset showing a magnified portion of the spectrum to better illustrate how the wavenumber correction affected the Raman spectrum.

3.4.9 Interval Spacing

The reference spectrum and all the spectra included in the pigment database will be interpolated to have an interval spacing of 1 cm^{-1} . This interval spacing standardizes all the data to have a consistent wavenumber range and step size to prevent errors when comparing across gratings, lasers, and systems which all have a different method for saving the total number of data points. The Octave code can be found in Appendix D lines 92-99. Shown in

Figure 6 is the interval spacing corrected Raman spectrum of sunset yellow and the library spectrum.

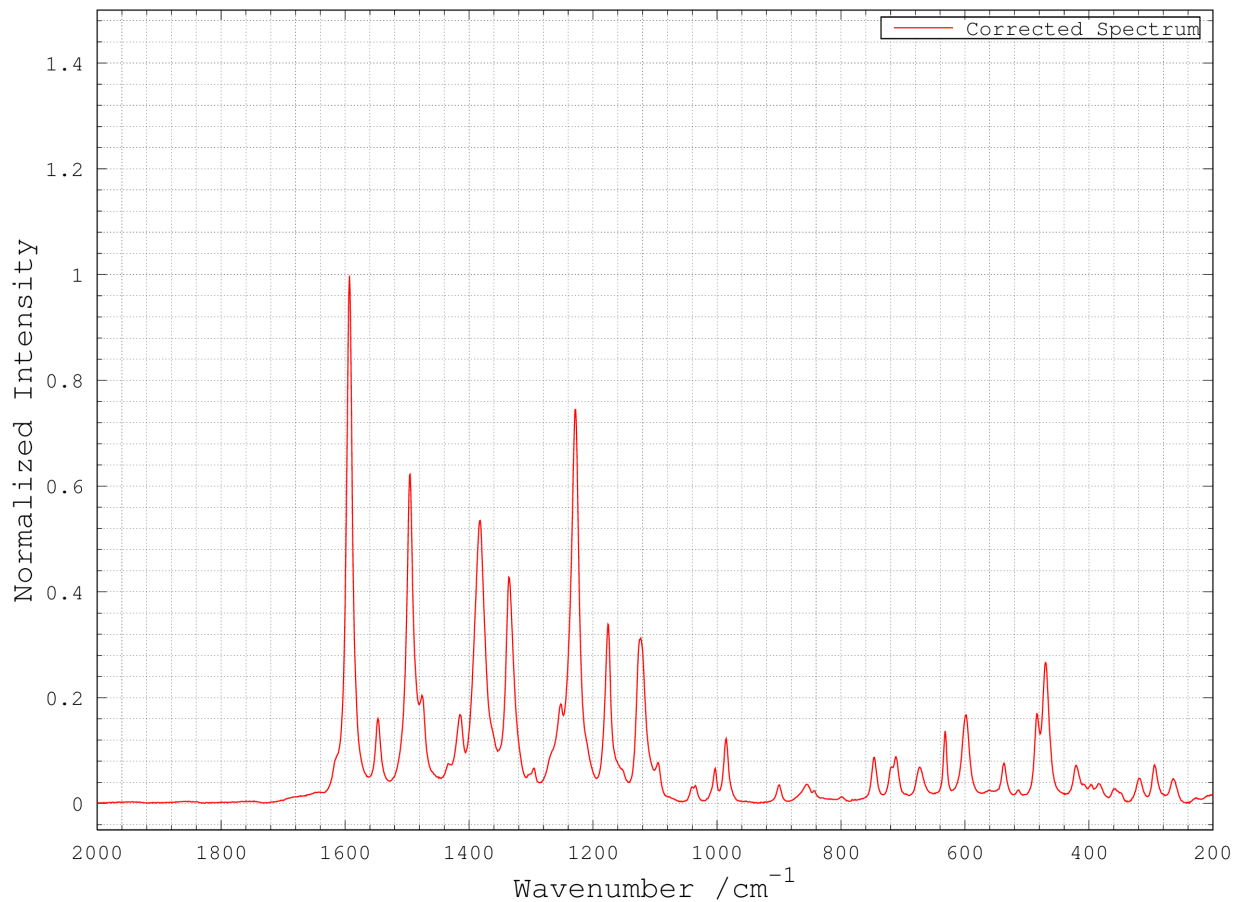


Figure 6: Interval spacing corrected Raman spectrum of sunset yellow, which is the spectrum used in the library database.

3.5 Master Library

Once the spectra of the pigments listed in Table 1 were measured and calibrated, they were saved into the database/master library as a JDX file format, described in Section 3.4.4. This library will be used for comparison when trying to identify an unknown material. In Appendix C are the calibrated Raman spectra of all the pigments and the instrument and

measurement parameters.

3.6 Matching Procedure

For spectral comparison, we used Hit Quality Index (HQI).^{24,25} The spectra in the library with the highest HQI value has the most spectral overlap with the unknown. Using the calibration procedure for both the library components and the “unknown” spectrum enhances the quality of the matches by avoiding problems such as variations in wavenumber, intensity, and peak width/resolution.²²

HQI is a spectral correlation coefficient calculated by taking the dot product of the *library* spectrum and the *test*(unknown) spectrum squared, divided by the dot product of the *library* spectrum multiplied by the dot product of the *test* spectrum.^{24,25}

$$HQI = \frac{(Library \cdot Test)^2}{(Library \cdot Library)(Test \cdot Test)} \quad (5)$$

3.7 Results & Discussion

3.7.1 Testing the Library - In House

To verify that the library matching program from Appendix G works, we first tested a pigment from the library, sunset yellow, against the library. The test unknown of sunset yellow was measured on a separate day than the library spectrum of sunset yellow. In Table 2, the algorithm results are presented in the column labelled ‘Calibrated’ with the HQI scores (0-1) and the pigment names. The match is identified by the highest HQI score (closest to 1) and is sunset yellow with an HQI score of 0.99.

3.7.2 Testing the HQI Against Each Step of the Calibration Procedure

Shown in Table 2 is the HQI score for sunset yellow with each calibration step, although each step required the interval spacing correction for comparison against the library data. This process was done to show that after each step the spectral overlap increases for the ‘unknown’ (sunset yellow), making the calibration procedure an effective tool for improving identification of unknowns.

Table 2: HQI scores after each step of calibration procedure for an unknown sample of ‘sunset yellow’.

Pigment	Original	Baseline Correction	Intensity Correction	Calibrated
PW6	0.07	0.02	0.02	0.02
PY139	0.14	0.08	0.08	0.09
NB1	0.24	0.12	0.12	0.12
PB15.3	0.26	0.15	0.15	0.14
PR254	0.28	0.21	0.21	0.23
PB15.1	0.33	0.18	0.19	0.18
PY83	0.34	0.35	0.35	0.42
AR52	0.34	0.18	0.18	0.19
AG25	0.38	0.26	0.27	0.27
AR18	0.39	0.28	0.28	0.27
PV23	0.42	0.29	0.29	0.31
Tartrazine	0.43	0.4	0.35	0.45
PW4	0.47	0.35	0.35	0.36
AV17	0.59	0.41	0.41	0.39
Sunset Yellow	0.76	0.92	0.95	0.99

3.7.3 Testing the Library Against an External Database

Also tested was the library against a Raman spectrum that is available in an online database of pigments. Because there is only one database that has downloadable data,¹⁷ the only spectrum that was available for comparison was PW4. This spectrum was acquired following no calibration procedure, requiring adjustments to the spectrum to match the library spectra. Listed below are the preprocessing steps that were used

1. The spectrum from the external database had a range of 200-600 cm^{-1} , requiring the spectrum to be zero-filled to 2000 cm^{-1} to match the library's wavenumber range.
2. The external PW4 spectrum was intensity normalized the spectrum to have a maximum intensity of 1.
3. An interval spacing correction was done so that the external PW4 spectrum had a wavenumber range of 200-2000 cm^{-1} with a step size of 1 cm^{-1} .

Once the external database PW4 spectrum was ‘preprocessed’ we were able to determine the similarity of the spectrum to our library. Listed in Table 3 are the results of the HQI between the two. While the scores are extremely low, the highest score (highlighted in blue on the table) did represent the correct match to the library, PW4.

Table 3: Similarity scores between the library and an external database spectrum.

Pigment	External Database HQI Scores
AV17	0.050421
AG25	0.0023202
AR18	0.0019289
AR52	0.0047298
NB1	0.0039407
PB15:1	0.0066395
PB15:3	0.0011582
PR254	0.00072292
PV23	0.006932
PW4	0.098348
PW6	0.018947
PY139	0.0022817
PY83	0.0017149
Sunset Yellow	0.0014234
Tartrazine	0.00092149

Figure 7 shows the reason for the low similarity score. Present in the PW4 library spectra (green) is a broad hump from 1000-1600 cm^{-1} . This comparison shows that the library

spectrum and HQI program could be used with a spectrum acquired without following a calibration procedure and still produce accurate similarity results.

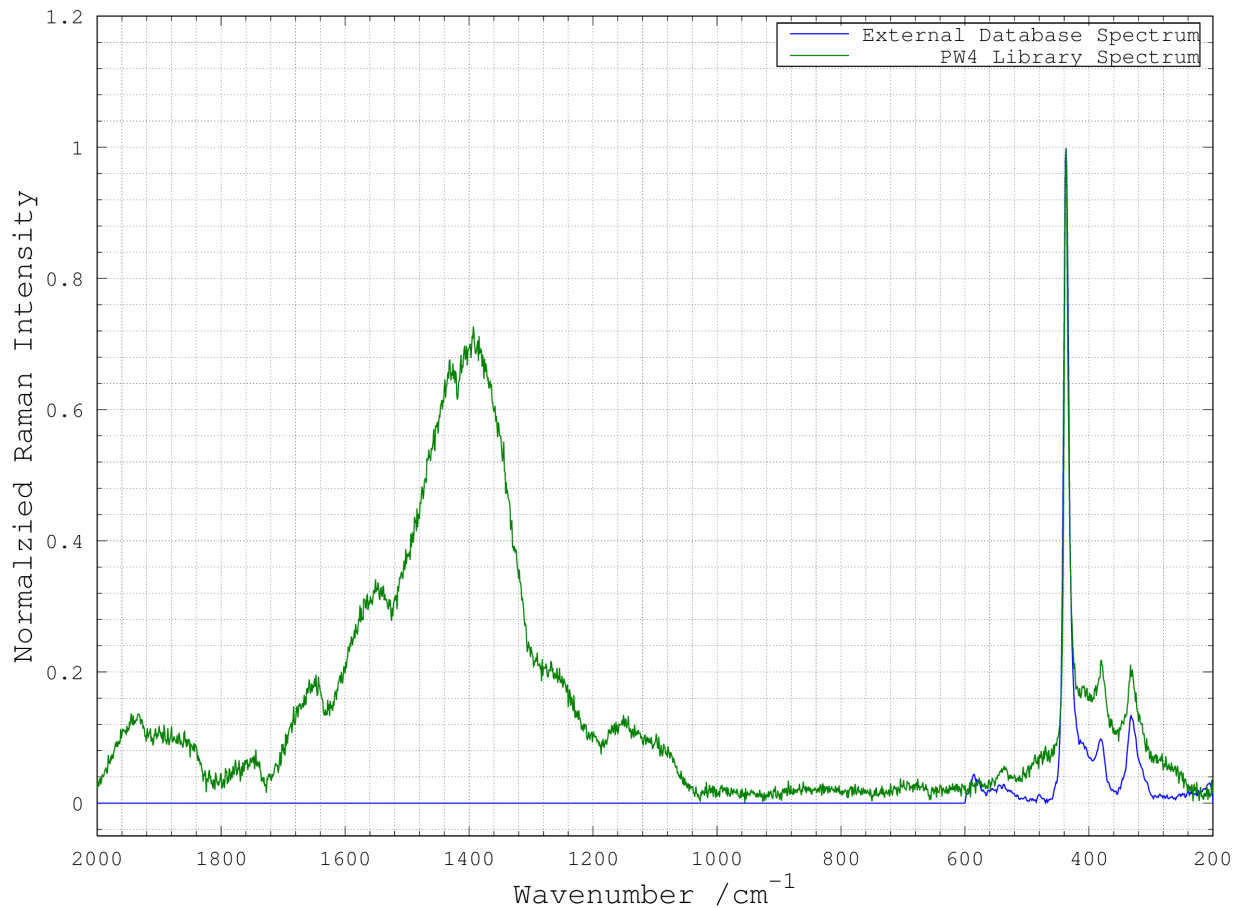


Figure 7: External database spectrum and library spectrum of PW4.

3.8 Conclusion & Future Research

At the end of this project some important goals were reached:

- Developed a program for calibration of Raman spectra. This is now in the form of a GUI to make the calibration more user friendly.
- A spectral library searching program was developed using HQI.

We plan to continue growing the pigment/dye/material database that will include SERS spectra of these materials. The matching program developed will continue to be edited with the addition of matching methods including *Euclidian Distance* and *Mahalanobis Distance* and methods for baseline correction inside the program.²⁶

CHAPTER 4 ALS-MCR AND MIXTURE ANALYSIS

4.1 Introduction to Mixture Analysis

With paintings, which will always be a mixture of components, the automated matching program developed in Chapter 3, can not distinguish spectral contributions of individual components in a single Raman spectrum, leading to misidentification and failure to discriminate between pure components and mixtures.²⁷ Due to this problem, a method utilizing spectroscopic imaging and chemometric algorithms that can identify individual components from mixtures was developed.^{13,14} While this research was done on surrogate paint samples and on a much smaller scale, this method can be applied to a spectroscopic map of an entire painting, providing in theory complete material identification by non-destructive and *in situ* analysis.

4.1.1 ALS-MCR

The chemometric algorithm used to digitally extract pure components in mixtures was alternating least squares - multivariate curve resolution (ALS-MCR). ALS-MCR is used to decompose spectroscopic images (Figure 8 step one), assuming that the spectral contributions from the pure components to the measured signal are dependent upon concentration of the component and are independent of other components.²⁸ If the sample follows this behavior, pure component spectrum \mathbf{K} and a relative concentration profile of the components \mathbf{C} can be resolved.^{11,12,29}

In matrix form, spectroscopic images can be expressed as a matrix \mathbf{A} and resolved into \mathbf{C} concentration profiles and \mathbf{K} pure component spectra as shown in Equation 6

$$\mathbf{A} = \mathbf{C} \mathbf{K} + \mathbf{R} \tag{6}$$

where $\mathbf{A}(m \times n)$ is the spectroscopic image data, with dimensions m samples (spectra) by n wavelengths over which the spectra were collected; $\mathbf{C}(m \times q)$ is the matrix of concentration profiles associated with the number of components q ; $\mathbf{K}(n \times q)$ is the pure component spectral matrix, with q rows containing each pure component's pure spectrum and n wavelengths; and \mathbf{R} is the matrix associated with experimental error and spectral noise, which is ideally negligible.^{11,12,29}

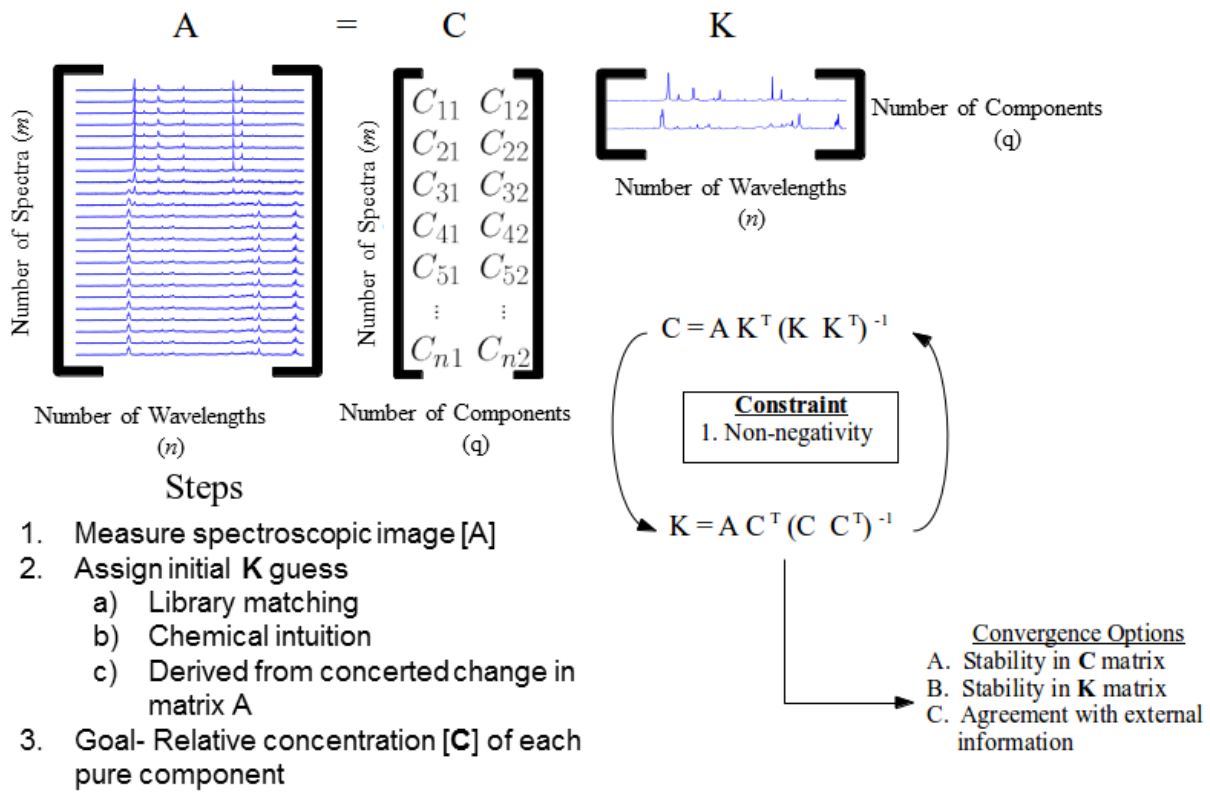


Figure 8: Outline for ALS-MCR.

Figure 8 is a diagram of the ALS-MCR algorithm. Initiating the ALS-MCR algorithm requires an initial guess of either the concentration profile or spectra of the pure components.

For this research a guess of \mathbf{K} , (Figure 8 Step Two) was used. With these mixtures, only two components were present with each individual component being in it's purest form at the ends of the spatial axis (the first and last spectra). Then using Equation 7

$$C = A(K)^+ \quad (7)$$

to calculate the relative concentration values (\mathbf{C}) of the pure components from the measured spectra and the pure component spectra guess, where K^+ is the pseudoinverse of the \mathbf{K} matrix calculated in Equation 8,¹³

$$(K)^+ = K(K^T K)^{-1} \quad (8)$$

To calculate \mathbf{K} Equation 9 is used,

$$K = C^+ A \quad (9)$$

where C^+ is the pseudoinverse of matrix \mathbf{C} , shown in Equation 10

$$C^+ = C(C^T C)^{-1}. \quad (10)$$

Iterations calculating \mathbf{C} and \mathbf{K} are done using the updated \mathbf{C} and \mathbf{K} from the previous calculations until a minimal residual matrix, \mathbf{R} , is reached. The residual matrix is calculated using Equation 11

$$R = A - CK. \quad (11)$$

While performing these iterations, a nonnegativity constraint is applied after each step. The nonnegativity concentration constraint assumes that concentrations of the components in the mixture are always greater than or equal to zero. The nonnegativity spectral constraint

states that the intensity values are always greater than or equal to zero. After each pure component spectra is calculated all values for \mathbf{K} that are below zero are set to zero and after the concentration profiles are calculated, all the \mathbf{C} values below zero, are set to zero.

The ALS-MCR loop is complete once the convergence limit(CV) is met, calculated using Equation 12.^{30,31} For these samples, convergence limit , CV , of 0.01 was used. This CV limit was chosen from an estimation of the noise in the spectra at the region from 1900-2000 cm^{-1} .

$$CV = \sqrt{\frac{\sum(R)^2}{\sum(A)^2}} * 100 \quad (12)$$

Once the CV limit is met, the resolved pure component spectra, \mathbf{K} , and the relative concentration profile of the components, \mathbf{C} , are the best estimated \mathbf{K} and \mathbf{C} . The pure component spectra can then be compared to the library using the matching algorithm from Chapter 3.

4.1.2 Spectroscopic Imaging

Spectroscopic imaging, also called hyperspectral imaging, is needed to identify components in a painting because of the changes in concentration of the pure components across the spatial domain of the painting. Spectroscopic imaging combines both spatial information and spectral information producing a multidimensional data set.^{11,12} Spectroscopic imaging of paintings records the identification of material variances that may appear to be homogeneous to the naked eye, but contain variations of particles on a 1-2 μm scale.³² These variations would most likely be missed with single point Raman acquisition measurements because of the chemical heterogeneity of the sample.

To better explain spectroscopic imaging and the data set that is acquired, shown in Figure 9 is a 3D representation of the spectroscopic image for Sample A, NB1 and PB15:1. The spatial information is usually along the X-Y plane and spectral information is along the

z axis. In this figure the image position is along the x axis, the wavenumber information is along the y axis, intensity at each x and y, is along the the x axis. This 3D plot shows how Raman spectra change with respect to spatial position.

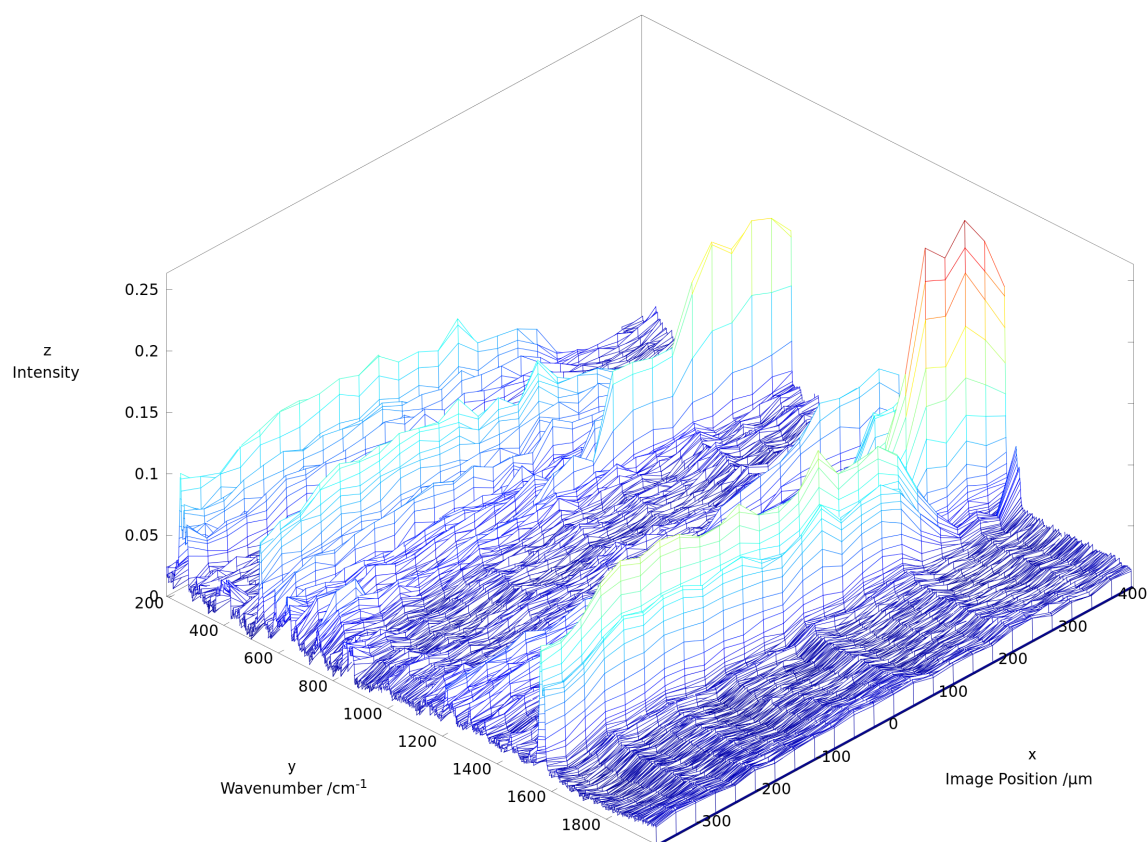


Figure 9: Spectroscopic image of Sample A, PB15:1 and NB1, with a ‘clean’ interface between the two paints, represented as a 3D surface plot with the image position along the x axis, the wavenumber information along the y axis, and intensity at each x and y, along the the x axis.

4.2 Experimental

4.2.1 Materials

For this experiment we made two different samples, Table 4, each containing two paints with one pigment each. Sample A was made with PB15:1 and NB1 mixed with Apple Barrel acrylic white paint, purchased from Karma Pigments (Montreal, Quebec, Canada), was synthesized and purified in the laboratory using the procedure described in *Experimental Organic Chemistry*,¹⁹ and purchased from Wal-Mart (Sylva, NC), respectively. Sample B was made by mixing tartrazine and sunset yellow, purchased from Abbey Colors (Philadelphia, Pa), with Apple Barrel acrylic white paint. Both mixtures were painted on all-cellulose painting paper with no optical brighteners from Dick Blick Art Materials (Galesburg, Il).

Table 4: Pigment used in paints for ALS-MCR.

Sample Letter	Pigment 1	Pigment 2
A	NB1	PB15:1
B	Tartrazine	Sunset Yellow

These pigments were selected because they would test the limits of this algorithm due to spectral similarity or high spectral overlap that could be easily misidentified. Shown in Figure 10 is a spectral overlap map comparing all library components spectra against with blue representing a low HQI score and dark red representing a high score.

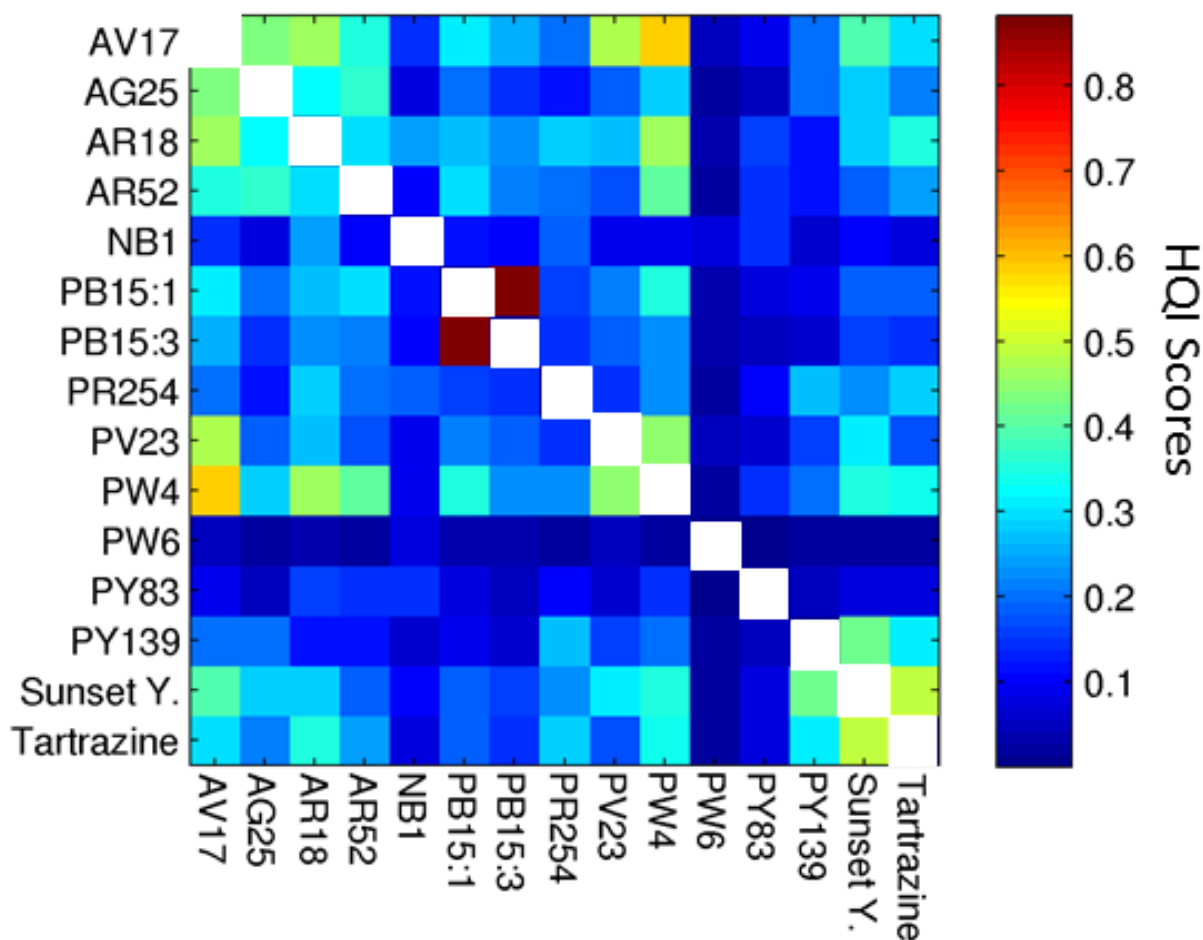


Figure 10: Spectral overlap map of each library spectrum compared against every other library spectrum. A color bar is shown with blue representing a low HQI score and dark red representing the highest HQI score.

For Sample A, we chose to use PB15:1, which has a HQI score of 0.883 when compared to PB15:3. These two pigments are polymorphs, different crystal structures, of the same compound which is discussed in Section 4.3.1. Sample B was chosen to test 2 pigments, sunset yellow and tartrazine, that are spectrally close to each other, with a HQI score of 0.496. The structures of these pigments are found in Appendix C.

4.2.2 Sample Preparation

Visually both of the mixtures had definitive interfaces between the two paints, that we describe as a heterogeneous mixture, as illustrated in Figure 11. Each sample was spectroscopically imaged but under different measurement conditions. Our motivation for the measurement differences are discussed in Section 4.2.5.

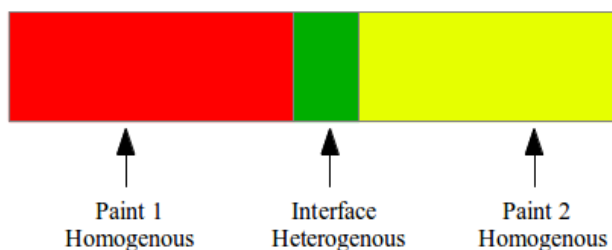


Figure 11: Example of preparation for Sample A and B, with homogeneous paints on either side of a heterogeneous interface.

4.2.3 Instrumentation

Measurements were made using a Horiba Jobin Yvon LabRamHR Raman microscope using a 10x objective with a numerical aperture of 0.25. The excitation laser is a Littman/Metcalf diode laser producing 785 nm (Sacher Lasertechnik Group, Marburg, Germany), measured with 600 groove mm^{-1} grating, and a TE-cooled CCD detector.

4.2.4 Data Analysis

All data analysis was performed using Octave version 3.2.4, with all algorithms being Matlab compatible.²⁰ The ALS-MCR algorithm used, Appendix H, was written in house.

4.2.5 Spectroscopic Imaging

Mapping across the surface of a paint sample is done by using the Raman LabSpec 6 software’s mapping acquisition setting. The user sets the measurement parameters such as acquisition time and repetitions, and also can choose options for spectroscopic imaging including collection of spectra along a line or within a defined shape (rectangle, square or triangle). This project involved mapping along a line with samples taken every 20-50 μm . In Table 5 the acquisition times, repetitions, and array size is shown for each sample.

Table 5: Measurement parameters for ALS-MCR Samples.

Sample Letter	Acquisition Time(sec)	Repetitions	Measurement time(mins)	Measurement Length	Array Size
A	20	3	45	Long Scan	15
A	1	5	6	Quick Scan	15
B	2	5	8	Quick Scan	10

Mentioned earlier, the main difference in the sample (besides the components) was the total acquisition time of the spectroscopic images. When developing this method, measurements were made using longer acquisition times and more repetitions, both increasing the overall measurement time. While these measurements produce spectra with high signal-to-noise ratios, the ratio often associated with spectral quality, the long exposure time damaged paint samples, which is shown in Figure 14 on page 36. Wanting to test the limits if the ALS-MCR algorithm quick scan of Sample A and B were performed. Throughout the rest of this Chapter, we will refer to a long scan (total measurement time 45 minutes) and a quick scan (total measurement time 6-8 minutes).

4.2.6 ALS-MCR Octave Code

The ALS-MCR code developed for this project is found in Appendix H, with additional functions for calculating the **K** and **C** values in Appendix I and for applying the non-

negativity constraint to \mathbf{K} and \mathbf{C} in Appendix J.

4.2.7 Steps for using the ALS-MCR code

Figure 12 outlines the steps for using the ALS-MCR algorithm and spectroscopic image for mixture analysis. To further describe how to use the Octave code developed, step-by-step directions are listed below with line references to the ALS-MCR code in Appendix H.

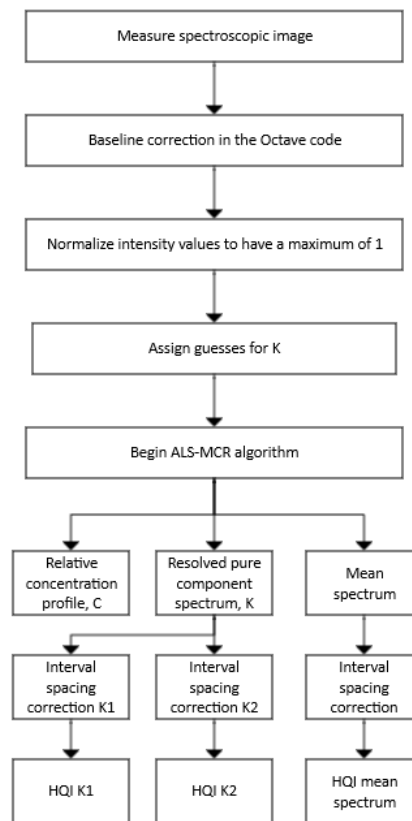


Figure 12: Procedure for using spectroscopic imaging and the ALS-MCR algorithm for identifying pure components in mixtures.

The ALS-MCR code requires the following preprocessing steps (lines of code in Appendix H) in order to process the spectral image into the \mathbf{K} and \mathbf{C} values.

1. Loading the mapping file (line 35) and defining the data matrix into the spatial infor-

mation, intensity values, and wavenumbers (lines 38-50).

2. Baseline correcting the mapping data (lines 56-66), using the points defined in line 59, with a linear fit/ order (line 60). The mapping data is also normalized to have a maximum intensity value of 1, using the *hnorm* function, on line 63.

Once the pre-processing procedure is complete, the ALS-MCR portion of the code can begin, following the step described below.

1. First, the user must assign a starting \mathbf{K} guess for the components, lines 73-80, list guesses that could be used including options for using
 - (a) 4th and last spectrum (line 74)
 - (b) First and random points (line 77)
 - (c) Last and random points (line 80).

An equal number of \mathbf{K} s are required for each component in the mixture.

2. Lines 83-90 define the constraints that were mentioned in Section 4.1.1. This states that \mathbf{C} and \mathbf{K} can not be negative values.
3. The tolerance (line 92) is set as 0.01, which can be adjusted depending on individual mixtures.
4. Once the \mathbf{K} , constraints, and tolerance is set, the ALS-MCR loop can begin, line 93 utilizing the code in Appendix I and J . Again iterations are performed till the CV value is below the set tolerance. Once this is met the ALS-MCR portion is complete, leaving resolved pure component spectra \mathbf{K} and relative concentration values \mathbf{C} of the components.

The steps listed above will produce the resolved pure component spectra of individual components and a relative concentration profile of the spectroscopic image. Also produced is the mean spectra, **A**, which is comparable to a bulk sample measurement instead of a microscopic sample measurement. The mean spectra will be used to show the advantages of measuring on a microscopic level (using a 10x objective) vs. measurements done on lower or without magnification. While the resolved pure component spectra and the mean spectra could be used for spectral comparison to a library visually, processing steps are needed to calibrate the spectra before using the matching algorithm from Chapter 3. The processing steps, similar to the ones in the calibration procedure from Chapter 3, calibrate the wavenumbers and normalize the intensity values. Listed below are the processing steps needed to calibrate the resolved pure component spectra and the mean spectra.

1. The mean spectra is calculated on line 97.
2. A wavenumber range from 200-2000 cm^{-1} is set on lines 98-99.
3. Using code developed to calibrate the wavenumbers in Chapter 3, we interpolate the wavenumbers (line 100) and then extrapolate in line 101, correcting the mean spectra.
4. A similar process is done to the **K** extracts on lines 104-113, using a for loop to correct both the **K** extracts.

Once the calibration steps are performed on the resolved pure component spectra and the mean spectra, we can then use the matching algorithm for comparison against the library. The similarity scores of the resolved pure component spectra and the mean spectra, allow for identification of the components with a confidence associated with the matches.

Listed on lines 126-133 is the code for comparing the resolved pure component spectra to the library, with an easy to read output created using lines 135-147. The output gives an ordered list of the similarity scores, along with the pigment's names. Lines 162-171 show the

similarity comparison between the mean spectra and the library, with an output providing the match that would be given if a single bulk measurement was performed instead of a microscopic measurement.

4.3 Results and Discussion

4.3.1 Sample A: Long Scan

Sample A had a white acrylic base, with two pigments (NB1 and PB15:1), with a clean interface between the two paints, and was measured using acquisition settings that provided a high signal-to-noise ratio spectra. The acquisition settings of the long scan are extremely time consuming and would be nearly impossible to scale up for spectroscopic imaging of entire paintings due to the measurement time that would be needed. Figure 13 shows the sample under 10x magnification before the mapping measurements and Figure 14 shows the sample after the measurements. While the damage was done on a microscopic scale, this type of damage is irreversible and if done on a larger scale could impact the image of the painting. We believe that the damage was caused by the relatively long acquisition time, which was another motivation for the ‘quick’ scan with Sample A and B.

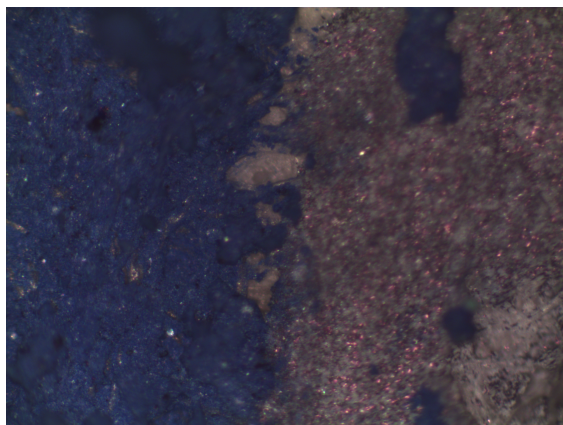


Figure 13: Before spectroscopic image was done for Sample A: long scan.

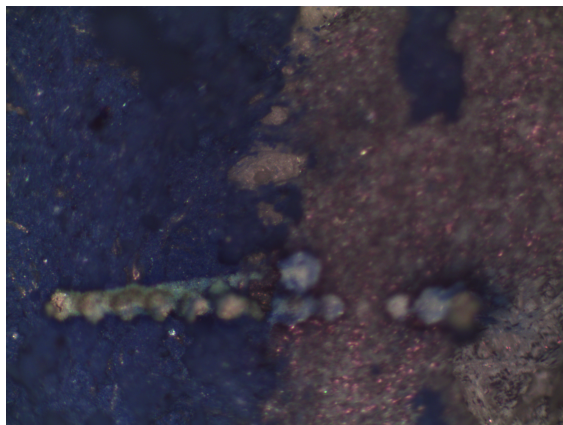


Figure 14: After spectroscopic image was done for Sample A: long scan.

The ALS-MCR code in Appendix H resolved two pure component spectra. Shown in Figure 15 is the resolved pure component spectrum of K1, the first pure spectral component in \mathbf{K} , and the library spectrum of NB1, which was identified as the most similar pigment from the matching algorithm with a HQI score of 0.692. Figure 16 is the resolved pure component spectrum of K2, the second pure spectral component, and the library spectrum of PB15:1, identified as the most similar pigment from the matching algorithm with a HQI score of 0.531. Figure 17 includes the two component's library spectra and the mean spectrum, calculated from the spectroscopic image, which would be produced if a bulk measurement was done with this sample instead of a microscopic one. The relative concentration profile in Figure 18 shows how the two components' concentration varied with respect to image position.

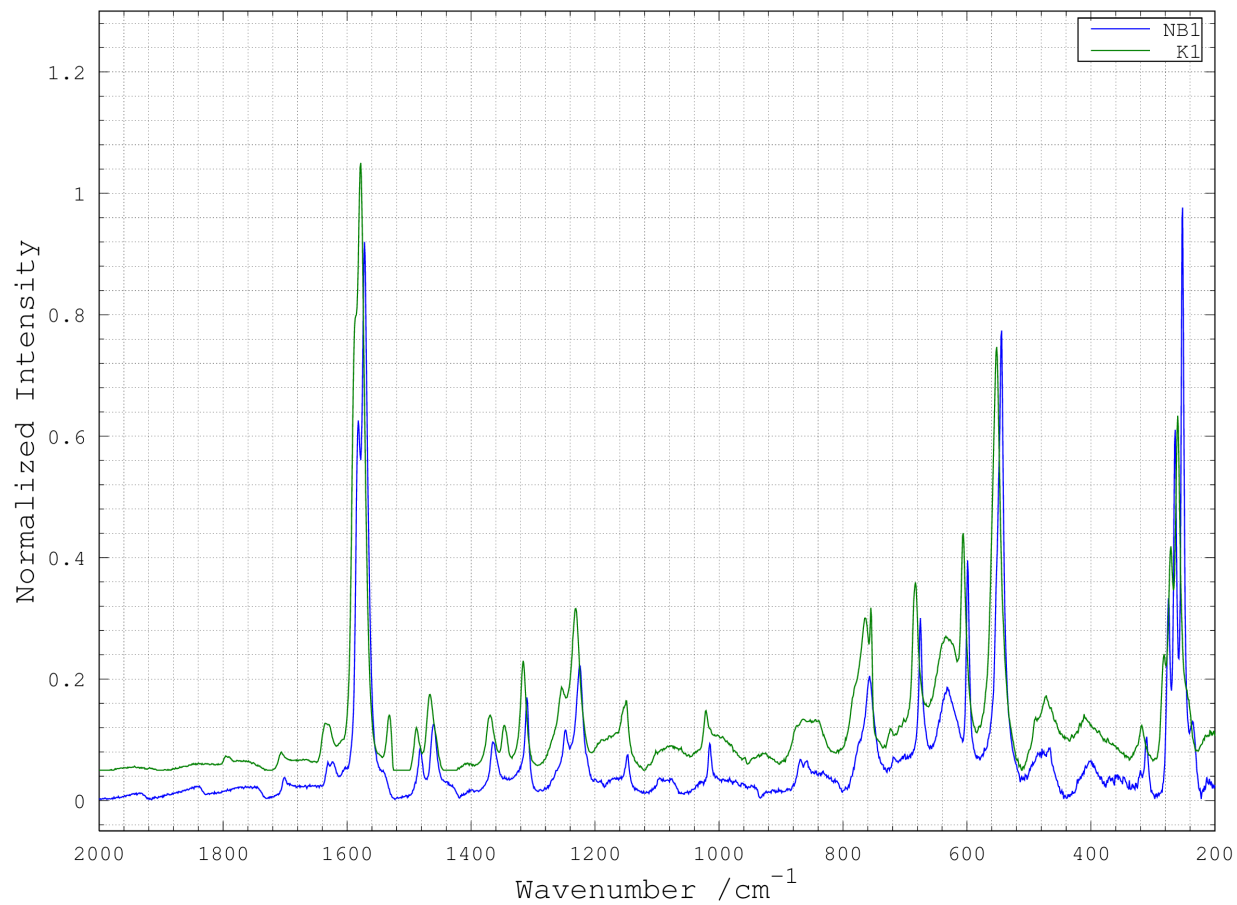


Figure 15: Library spectrum for NB1 (blue) and the calculated pure component spectrum of K1 (green). The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.

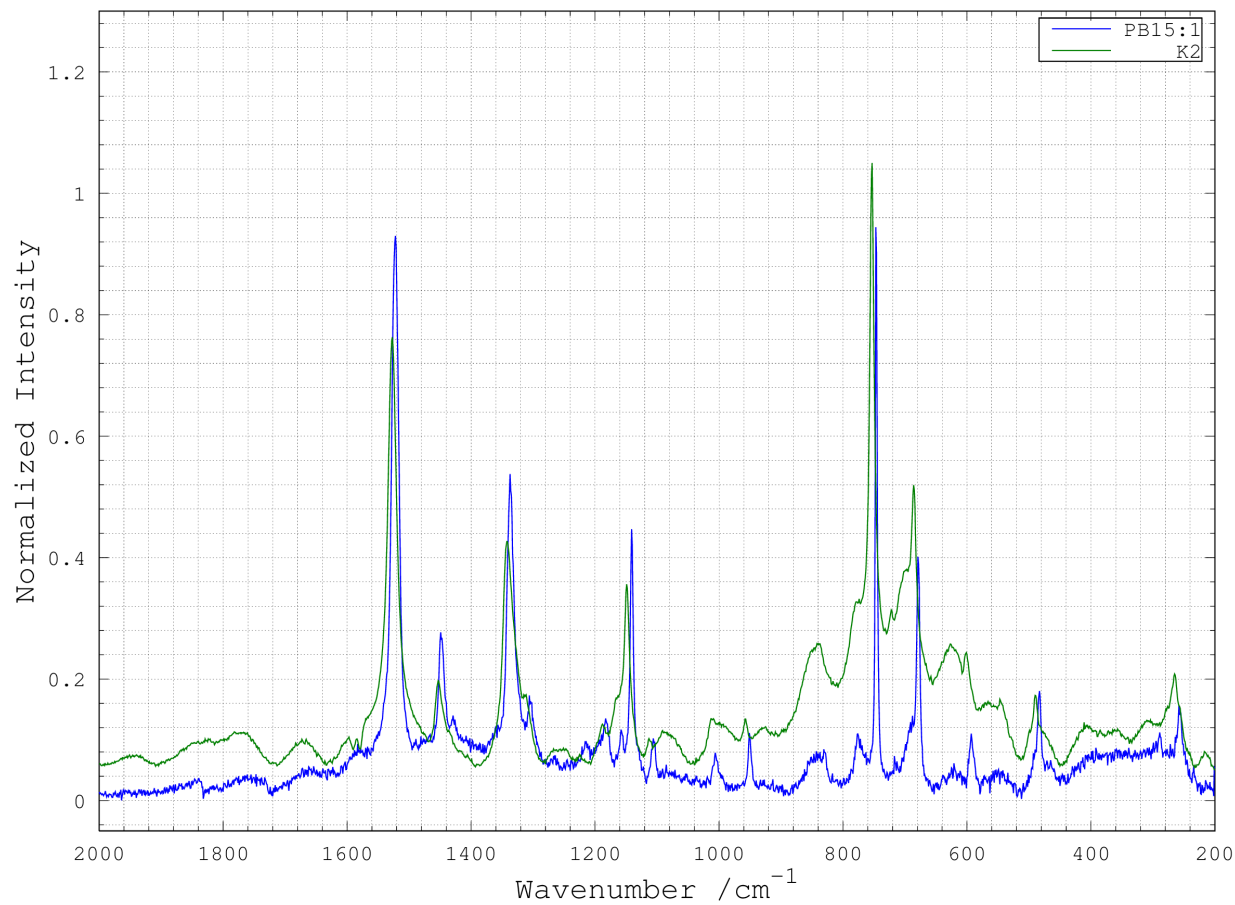


Figure 16: Library spectrum for PB15:1 (blue) and the calculated pure component spectrum of K2 (green). The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.

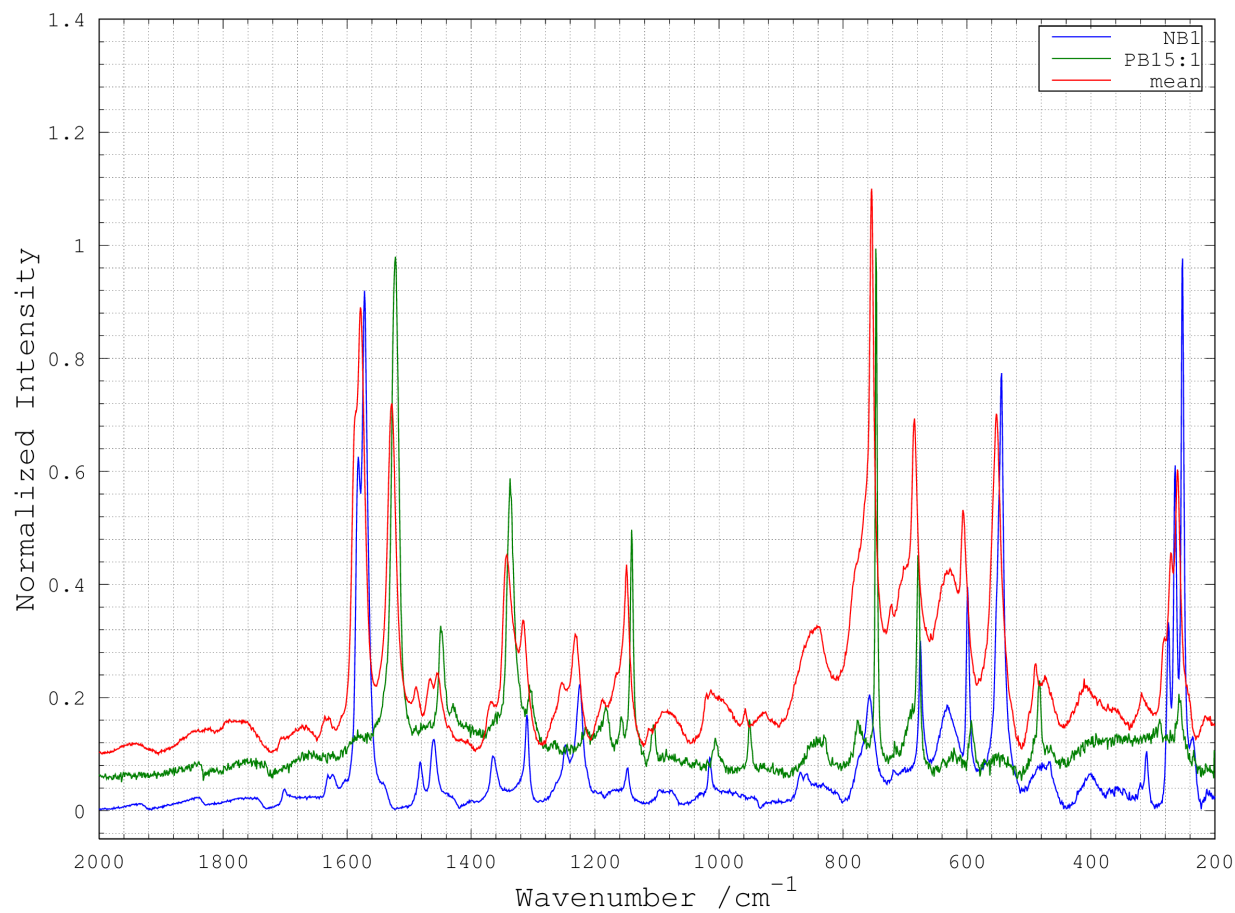


Figure 17: Library spectrum for Indigo(blue), PB15:1 (green) and the mean spectrum (red). The calculated pure component spectrum and mean spectrum is offset by adding a constant to the intensity values of the baseline.

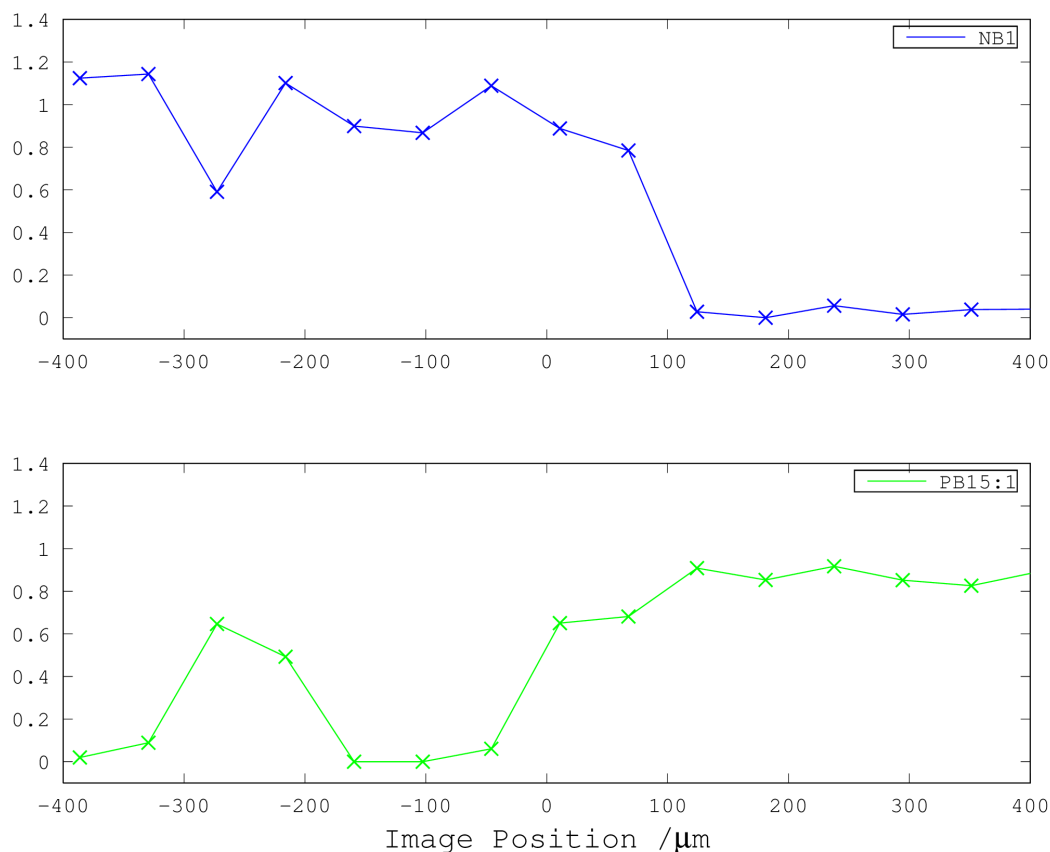


Figure 18: Relative concentration profile of NB1 (blue) and PB15:1 (green) with connecting lines for visual aid.

The similarity scores from the HQI algorithm are shown in Table 6, with HQI scores shown for K1, K2 and the mean spectrum. Highlighted in light blue are the matches for K1 and the mean spectrum, which was NB1. Highlighted in dark blue are the results for K2 similarity, PB15:1. Also highlighted in red is the second highest HQI score for the mean spectrum, which is PB15:3, a polymorph of PB15:1. It's important to note that the spectroscopic mapping and ALS-MCR improved the HQI scores of both K1 (NB1) and K2 (PB15:1) when compared to the mean spectra (measuring at one point) HQI scores.

Table 6: HQI scores for K1, K2, and mean of Sample A: long scan. Highlighted in light blue and dark blue are the highest HQI scores of K1, mean and K2, NB1, NB1 and PB15:1, respectively. The red highlight represents the second highest HQI score for the mean spectra, PB15:3.

Pigment	K1 HQI	K2 HQI	Mean HQI
AR18	0.231	0.179	0.278
AR52	0.109	0.220	0.224
PR254	0.268	0.142	0.271
PY83	0.135	0.044	0.112
PY139	0.111	0.064	0.117
Tartrazine	0.117	0.112	0.165
Sunset Yellow	0.180	0.094	0.182
AG25	0.096	0.111	0.144
NB1	0.692	0.181	0.512
PB15:1	0.113	0.531	0.405
PB15:3	0.114	0.530	0.406
AV17	0.179	0.210	0.271
PV23	0.104	0.157	0.180
PW4	0.088	0.143	0.161
PW6	0.092	0.085	0.120

An interesting outcome with Sample A: long scan, was the correct identification of the second component PB15:1 with the K2 HQI. While the mean HQI resulted in PB15:3 as the second component. While the differences between K2 HQI scores of PB15:1 and PB15:3, 0.531 and 0.530, are minute, have a percent difference of 16.51%, the ALS-MCR method provided the correct match between the two almost identical spectroscopic dyes. Shown in Figure 19 are the library spectra of PB15:1 and PB15:3. The most notable difference is the relative intensity values between the two. While these changes in band intensity impact similarity measurements, so do the band positions, labeled with a solid black line on the graph. There are two doublets (839 and 1215 cm^{-1}) and a band at 233 cm^{-1} which are all distinct in the PB15:3 spectrum (blue), but are distorted or missing in the PB15:1 spectrum (red). The spectral variations are attributed to the variation in the polymorphic crystal structure. These two crystals differ only by the arrangement in the crystal plane, with the

PB15:1 being a “regular” array and PB15:3 form being a “staggered” array.³³

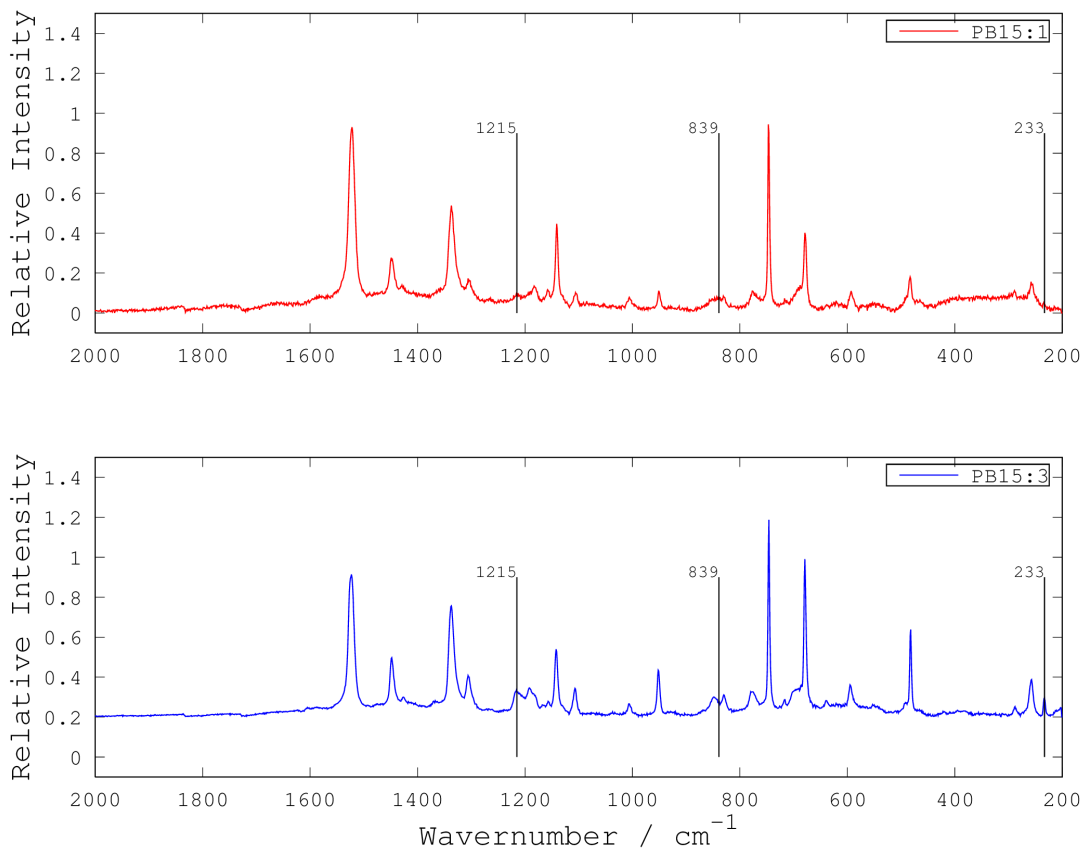


Figure 19: Library spectrum of PB15:1 (red) and PB15:3 (blue) with spectroscopic differences noted by solid black lines.

4.3.2 Sample A: Quick Scan

Since the long scan time for Sample A damaged the painting, but provided a high S/N image, a quick scan time was also tested on Sample A. Figure 20 is showing the region where the spectroscopic image was measured, with a black line indicating the portion which was mapped.

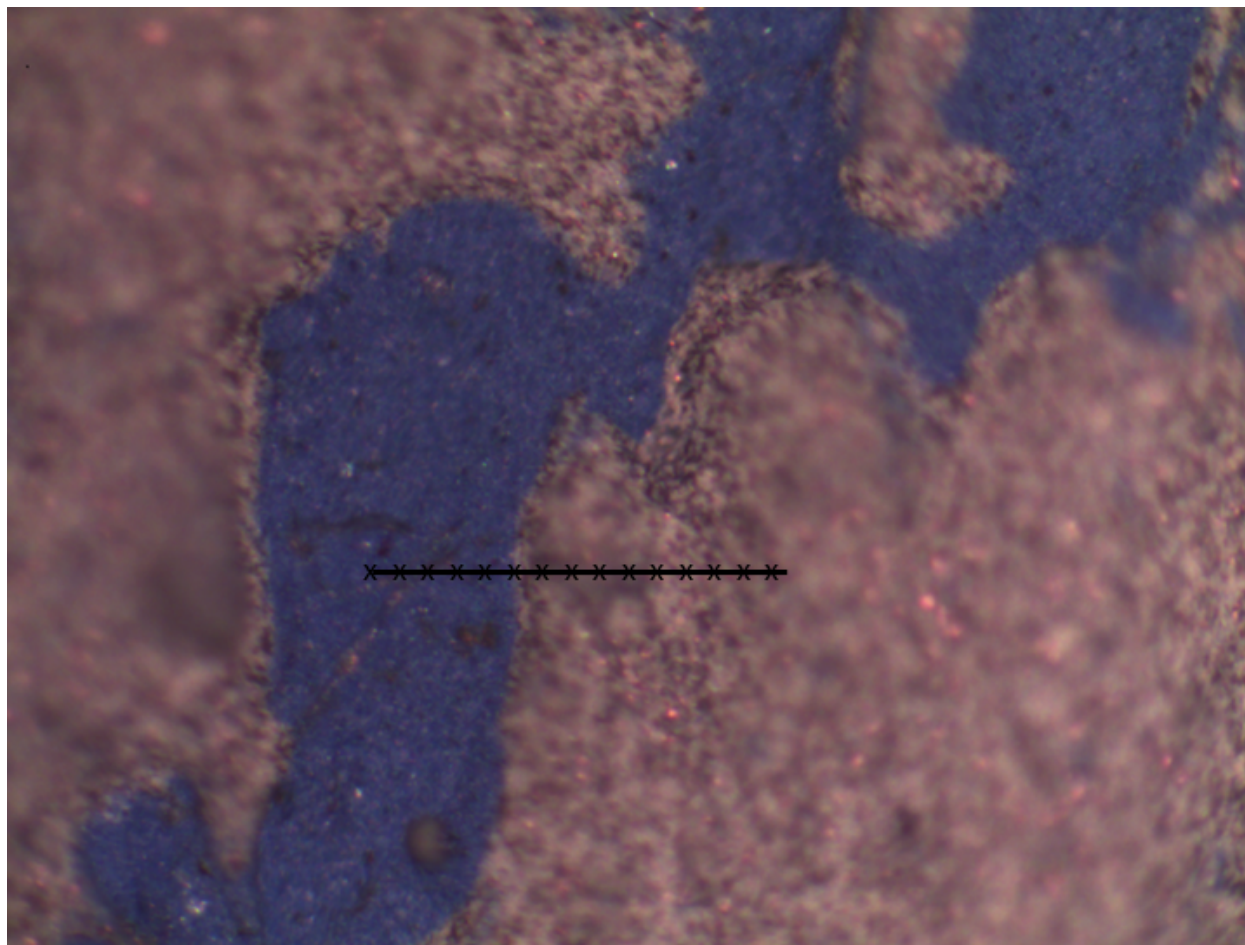


Figure 20: Approximate area imaged for Sample A: quick scan with x's indicating the approximate image position of each measurement.

The ALS-MCR code in Appendix H resolved two pure component spectra and the mean spectrum. Shown in Figure 21 is the resolved pure component spectrum of K1 and the library spectrum of PB15:1, which was identified as the most similar pigment from the matching algorithm with a HQI score of 0.297. In Figure 22 is the resolved pure component spectrum of K2 and the library spectrum of NB1, identified as the most similar pigment from the matching algorithm with a HQI score of 0.618. Figure 23 shows the two components' library spectra and the mean spectrum, which would be produced from a bulk Raman measurement. The HQI scores are shown in Table 7 on page 48. The relative concentration profile in Figure

24 shows how the two components' concentration varied with respect to image position.

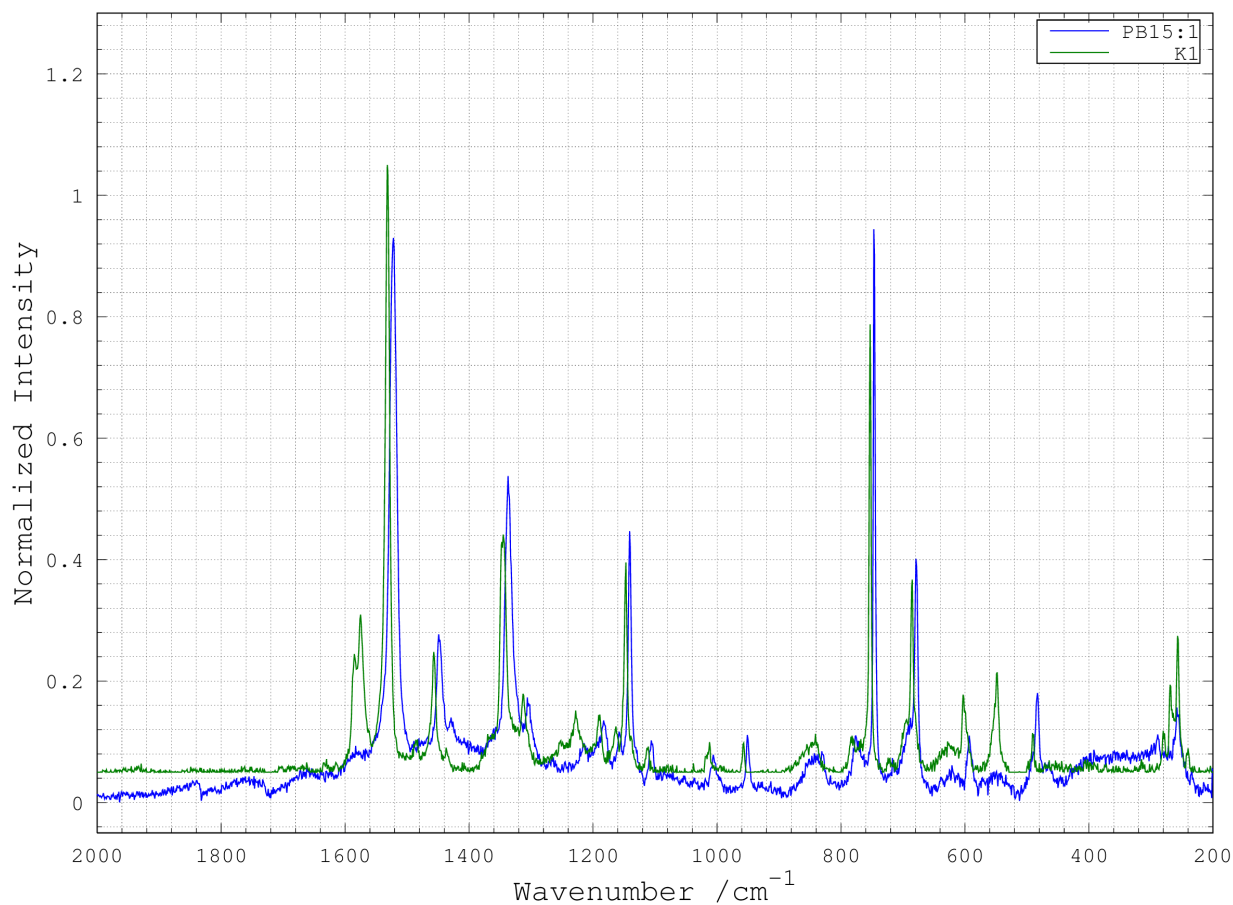


Figure 21: Library spectrum for PB15:1 (blue) and the calculated pure component spectrum of K1 (green) for Sample A: quick scan. The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.

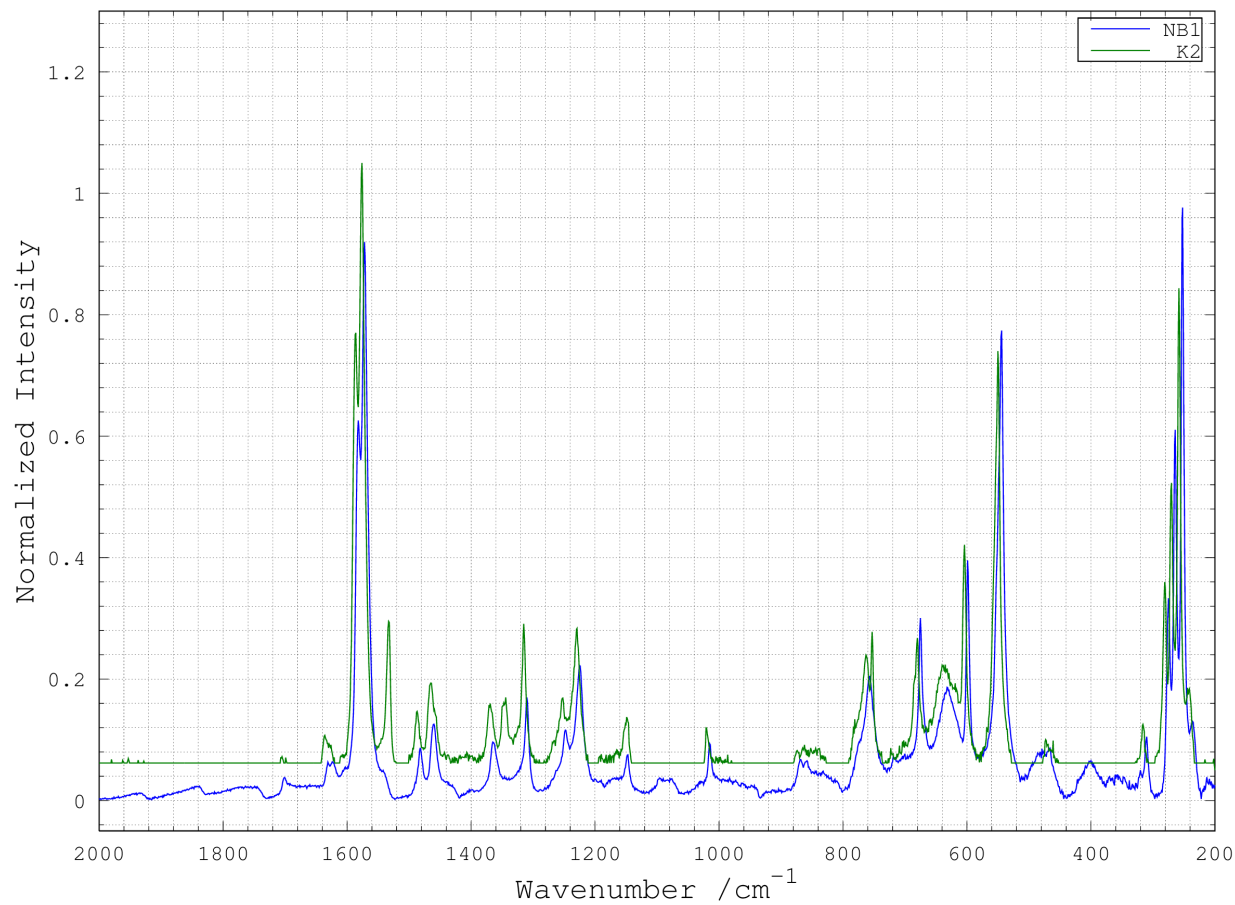


Figure 22: Library spectrum for NB1 (blue) and the calculated pure component spectra of K2 (green) for Sample A: quick scan. The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.

figures/finalfigs/meanquickblue2.png

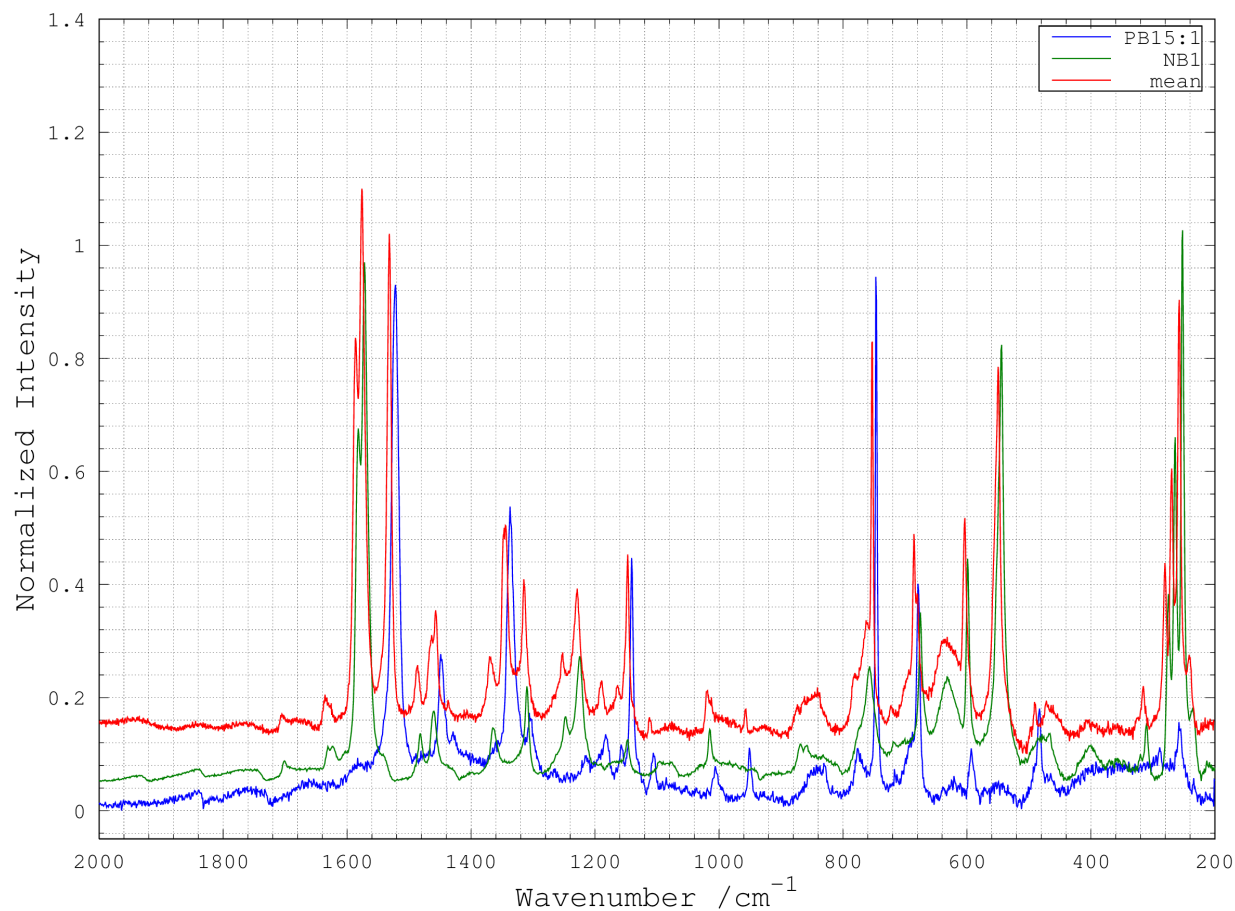


Figure 23: Library spectrum for PB15:1 (blue), NB1 (green) and the mean spectrum (red) for Sample A: quick scan. The calculated pure component spectrum and mean spectrum is offset by adding a constant to the intensity values of the baseline.

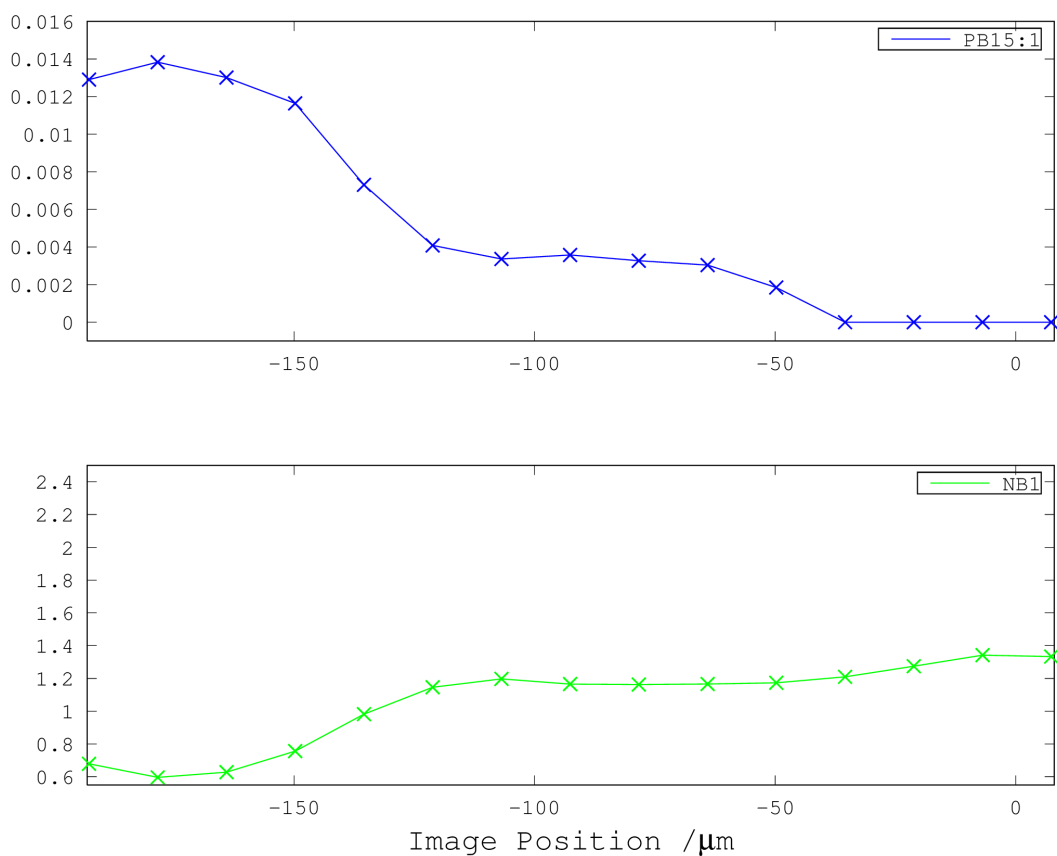


Figure 24: Relative concentration profile of PB15:1 (blue) and NB1 (green) with connecting lines for visual aid.

Shown in Table 7 are the HQI scores between K1, K2 and the mean spectrum for Sample A: quick scan. K1 was identified as PB15:1 with a similarity score of 0.297, highlighted in dark blue in the table. K2 was identified as NB1 with a similarity score of 0.618, highlighted in light blue. The mean spectrum in this mixture identified NB1 with a HQI score of 0.515, highlighted in light blue and the second highest HQI score, PB15:1, highlighted in dark blue with a similarity score of 0.183.

Table 7: HQI scores for K1, K2, and mean of Sample A: quick scan. Highlighted in dark blue is the highest HQI score for K1, identified as PB15:1. Highlighted in light blue is highest HQI score of K2 and the highest HQI score in the mean spectrum, NB1. Also highlighted in red is the second highest mean spectrum, identified as AR18.

Pigment	K1 HQI	K2 HQI	Mean HQI
AR18	0.171	0.190	0.228
AR52	0.161	0.073	0.118
PR254	0.163	0.234	0.257
PY83	0.041	0.091	0.0879
PY139	0.055	0.098	0.102
Tartrazine	0.107	0.083	0.110
Sunset Yellow	0.081	0.126	0.135
AG25	0.072	0.049	0.0662
NB1	0.141	0.618	0.515
PB15:1	0.297	0.083	0.183
PB15:3	0.268	0.081	0.174
AV17	0.161	0.106	0.148
PV23	0.086	0.059	0.082
PW4	0.149	0.073	0.118
PW6	0.009	0.041	0.0266

Using the ALS-MCR algorithm and spectroscopic image, both pure components in the mixture were correctly identified. The HQI scores of Sample A: quick scan for both the resolved pure components is lower than the HQI scores from Sample A: long scan, but both scan times resulted in correct identification of the components. The mean spectrum for the quick scan correctly identified NB1 as a component, but identified the second component as AR18. The misidentification with the mean spectrum, shows that using ALS-MCR and spectroscopic imaging improves identification of the components.

With Sample A: quick scan the limits of the ALS-MCR program and spectroscopic mapping were tested. While the resolved pure component spectrum had a lower signal-to-noise ratio (76 S/N) when compared to the longer scan time from Sample A (196 S/N), the algorithm was still able to successfully identify the two components. With this knowledge spectroscopic images can be acquired with the lower total scan time, meaning shorter expo-

sure to the paints, which helps to prevent damage that would be done done with the longer scans.

4.3.3 Sample B: Quick Scan

Sample B had a base of acrylic white paint mixed with sunset yellow and tartrazine (making two paints), with a visually overlapping interface between the two paints. It was measured with a quick scan, with a total scan time of 8 minutes. Figure 25 is showing the region where the spectroscopic image was measured, with a black line indicating the portion which was mapped.

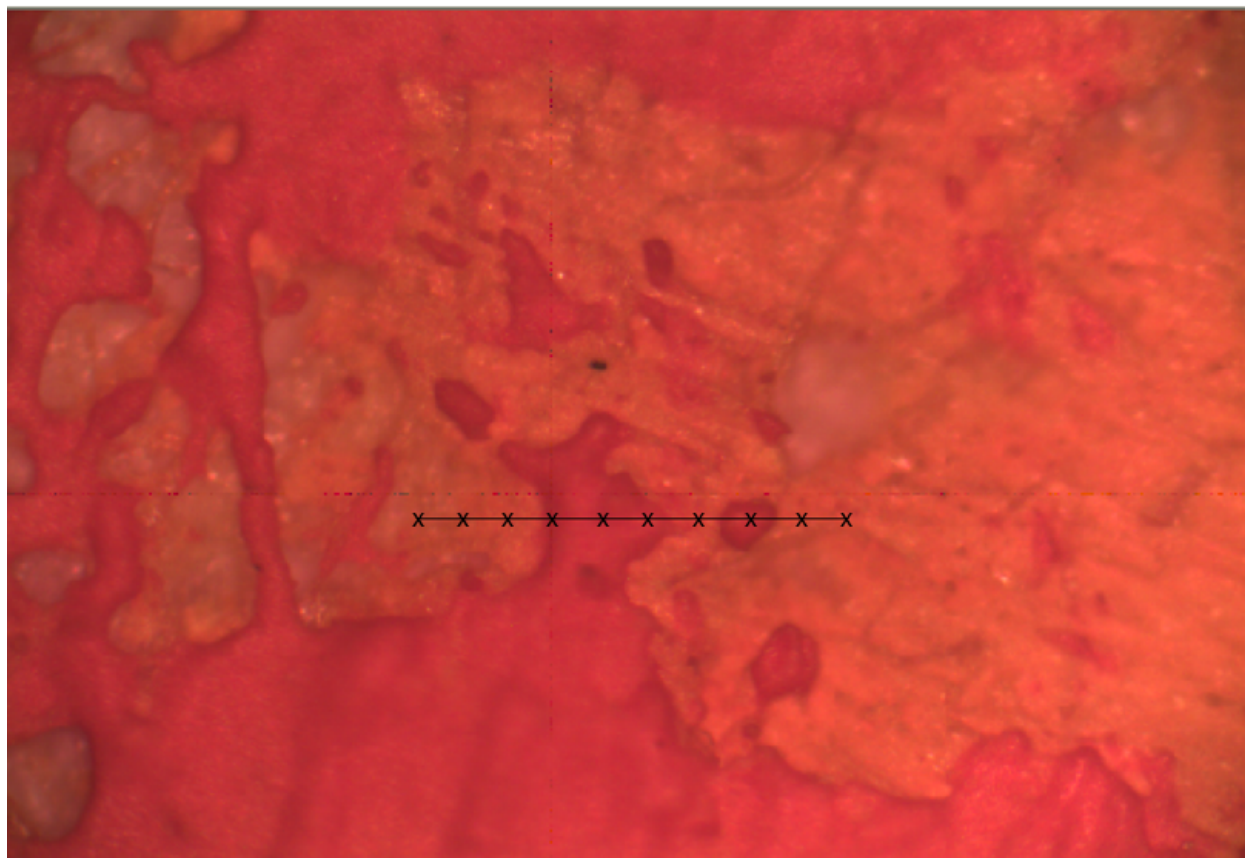


Figure 25: Approximate area imaged for Sample B: quick scan with x's indicating the approximate image position of each measurement.

Shown in Figure 26 is the resolved pure component spectrum of K1 and the library spectrum of sunset yellow, identified as the most similar pigment from the matching algorithm with a HQI score of 0.659. In Figure 27 is the resolved pure component spectrum of K2 and the library spectrum of tartrazine, identified as the most similar pigment from the matching algorithm with a HQI score of 0.638. Figure 28 shows the two components' library spectra and the mean spectrum, which would be produced from a bulk Raman measurement. The HQI scores are shown in Table 8 on page 55. The relative concentration profile in Figure 29 shows how the two components' concentrations varied with respect to image position.

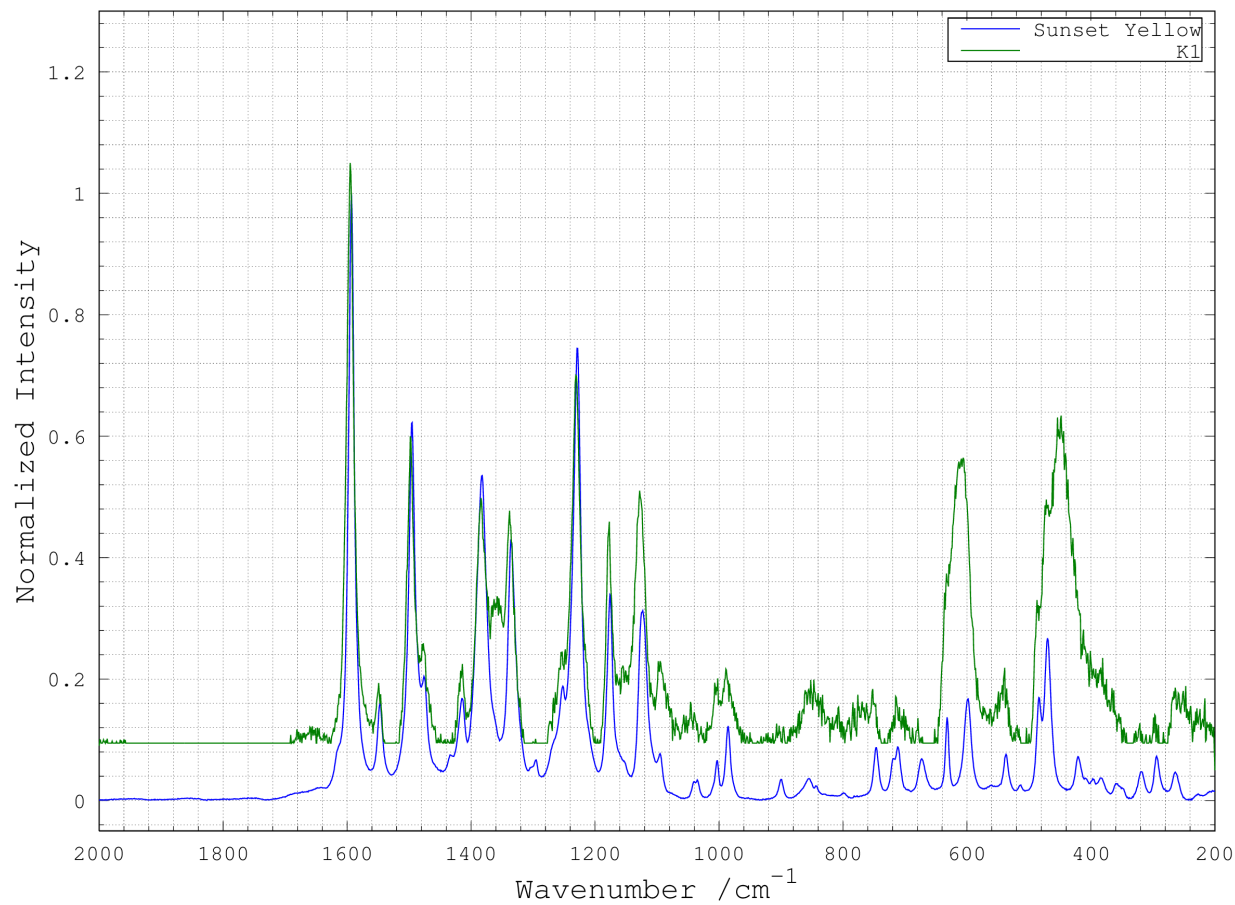


Figure 26: Library spectrum for sunset yellow (blue) and the calculated pure component spectrum of K1 (green) for Sample B: quick scan. The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.

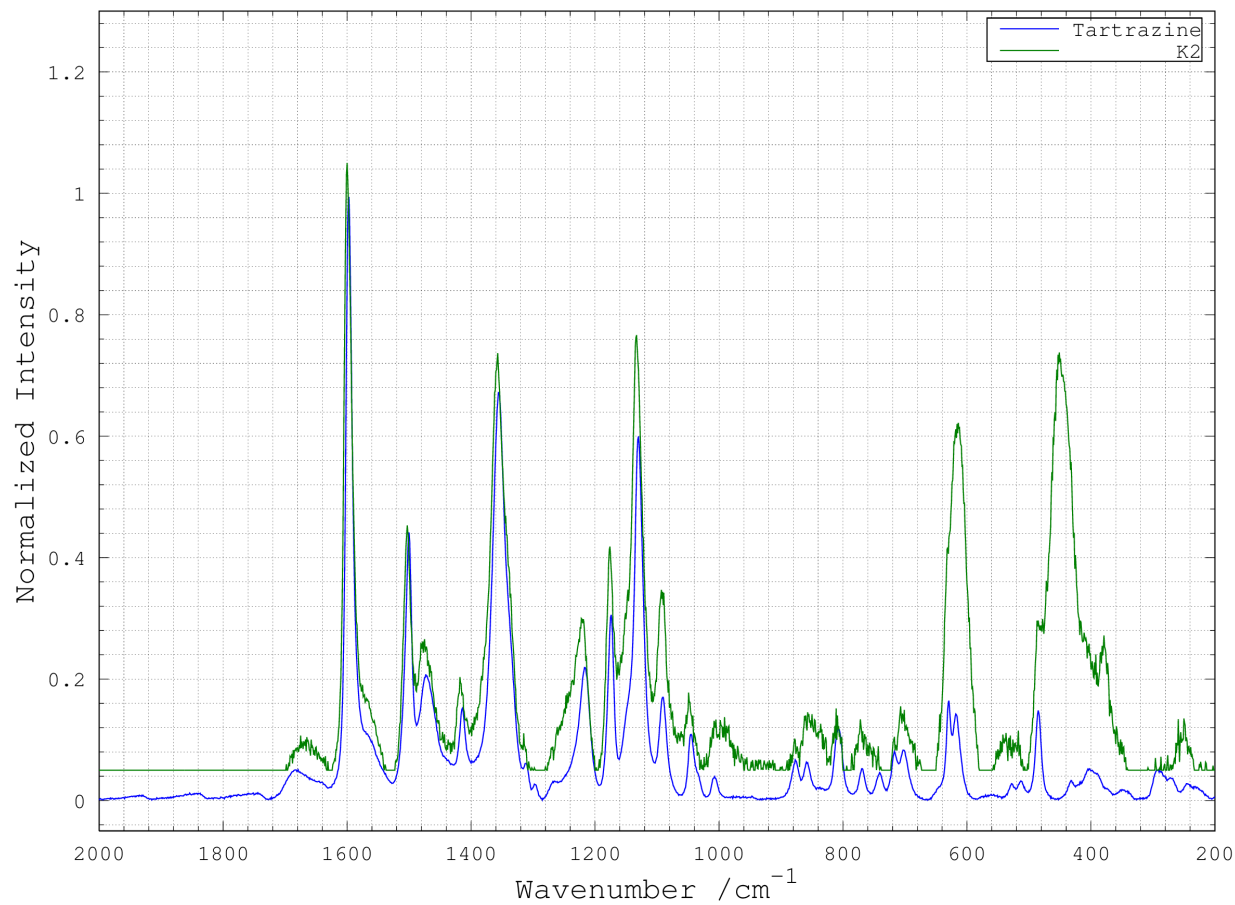


Figure 27: Library spectrum for tartrazine (blue) and the calculated pure component spectra of K2 (green). The calculated pure component spectrum is offset by adding a constant to the intensity values of the baseline.

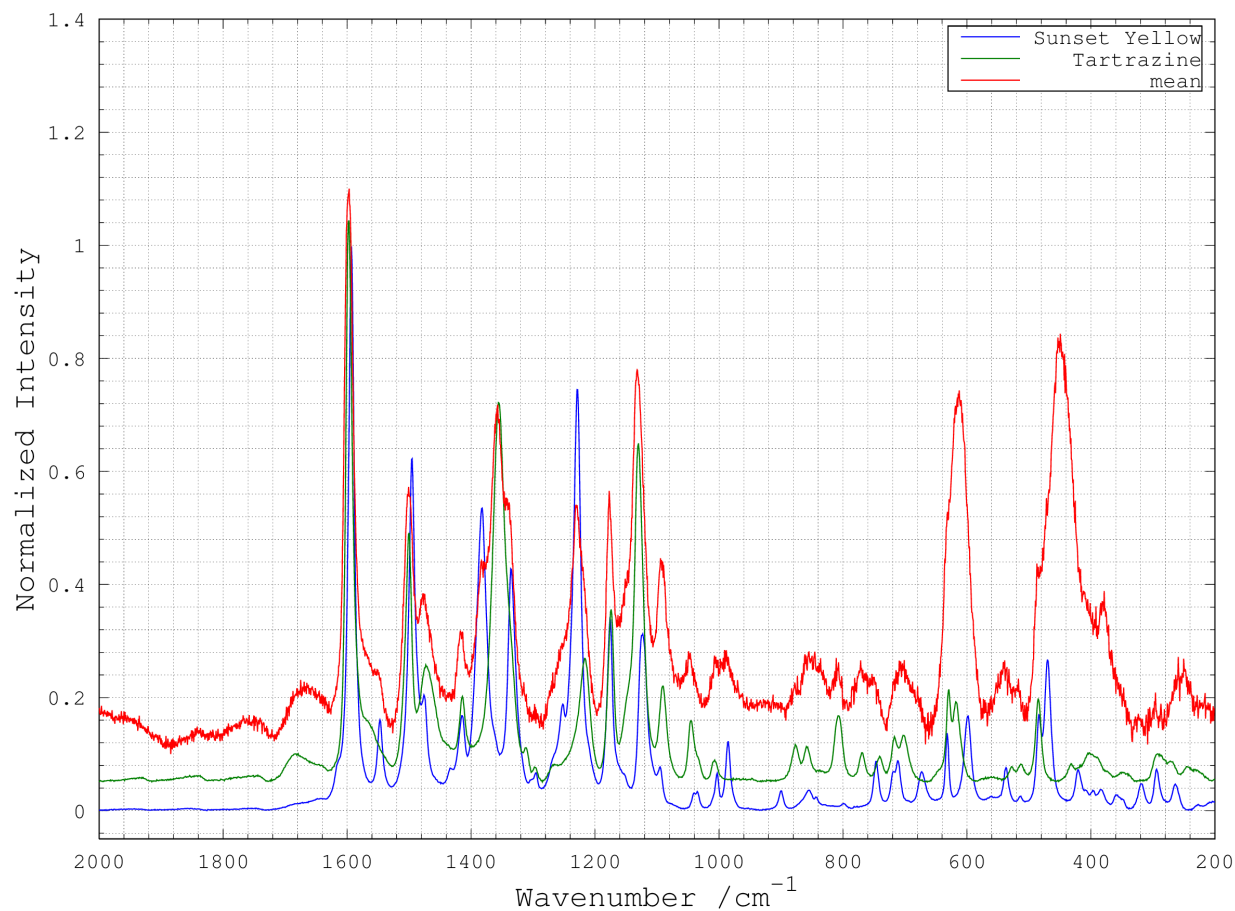


Figure 28: Library spectrum for sunset yellow (blue), tartrazine (green) and the mean spectrum (red) for Sample B: quick scan. The calculated pure component spectrum and mean spectrum is offset by adding a constant to the intensity values of the baseline.

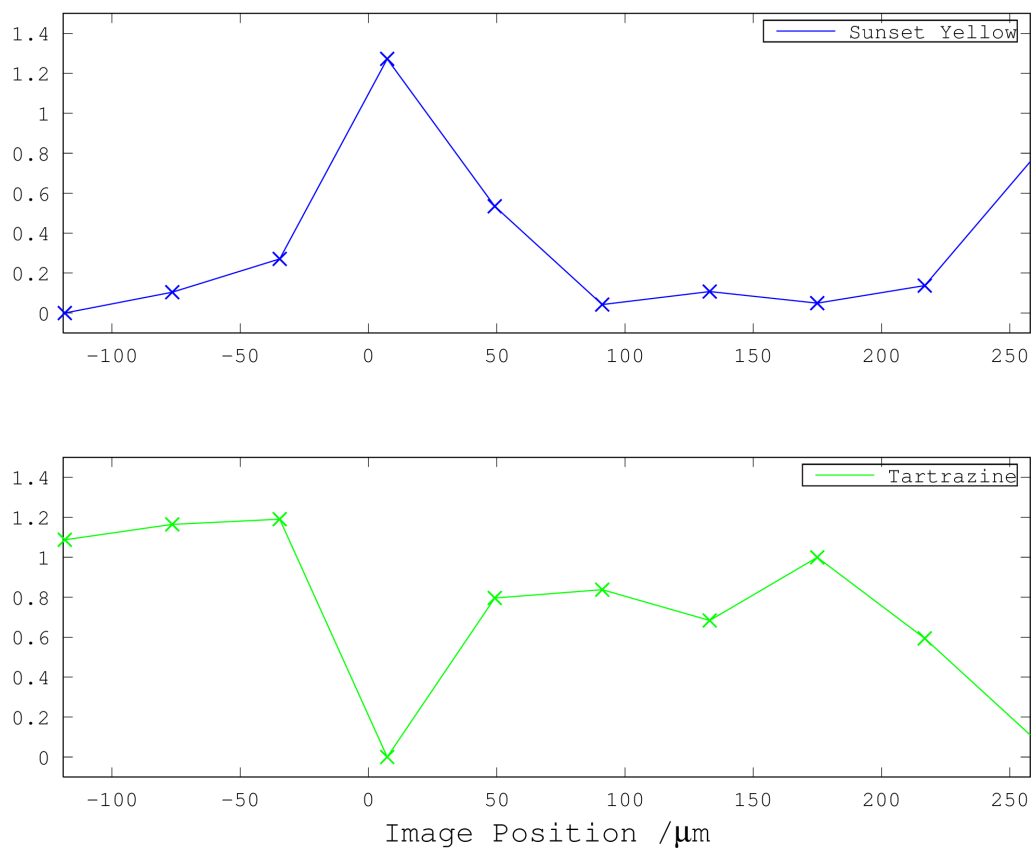


Figure 29: Relative concentration profile of sunset yellow (blue) and tartrazine (green) for Sample B: quick scan with connecting lines for visual aid.

Shown in Table 8 are the similarity results between K1, K2 and the mean spectrum for Sample B: quick scan. K1 was identified as sunset yellow with a similarity score of 0.659, highlighted in dark yellow on the table. K2 was identified as tartrazine with a similarity score of 0.638, highlighted in orange. The mean spectrum in this mixture identified tartrazine as the most similar match with a score of 0.599, highlighted in orange and the second highest score, sunset yellow, highlighted in dark yellow.

Table 8: HQI scores for K1, K2, and mean of Sample B: quick scan. Highlighted in dark yellow is the highest similarity score for K1, sunset yellow. Highlighted in orange is the highest similarity score for K2 and the mean spectrum, tartrazine. Also highlighted in dark yellow is the second highest HQI score in the mean spectrum, sunset yellow.

Pigment	K1 HQI	K2 HQI	Mean HQI
AR18	0.212	0.250	0.229
AR52	0.152	0.183	0.152
PR254	0.187	0.188	0.187
PY83	0.289	0.187	0.216
PY139	0.063	0.060	0.058
Tartrazine	0.495	0.638	0.599
Sunset Yellow	0.659	0.373	0.470
AG25	0.205	0.189	0.186
NB1	0.129	0.098	0.105
PB15:1	0.141	0.153	0.139
PB15:3	0.115	0.115	0.110
AV17	0.368	0.337	0.336
PV23	0.219	0.162	0.174
PW4	0.260	0.289	0.267
PW6	0.0762	0.084	0.079

4.3.4 Material Identification with Respect to Image Position

Before proving that the ALS-MCR algorithm successfully identified the components from a spectroscopic image, identification of the components at each image position along the map using HQI was done. This gives a comparison between ALS-MCR results compared to individual microscopic measurements.

Individual map similarity results are shown in Table 9 for Sample A: long scan. NB1 was the primary component for positions 1-9 with a similarity score of 0.722, matching what was expected from the original image and the area that was mapped. While NB1 was correctly identified as the major component for positions 1-9, positions 10-15 were identified varying between PB15:1 and PB15:3, with a significantly lower similarity score of 0.125 and 0.120, respectively. While this still shows that the primary component can be identified across a spectroscopic map, verification of the second component would require further inquiry. This

could be done by visually comparing the raw data to the pigment spectra, but this would become time consuming as sample size and number of components increase.

Table 9: HQI scores with respect to each image position for Sample A: long scan.

Position Number	Image Position, x axis, (μm)	HQI Score	Pigment Identified
1	-386.212	0.722	NB1
2	-329.485	0.722	NB1
3	-272.757	0.722	NB1
4	-216.030	0.722	NB1
5	-159.303	0.722	NB1
6	-102.575	0.722	NB1
7	-45.848	0.722	NB1
8	10.879	0.722	NB1
9	67.607	0.722	NB1
10	124.334	0.125	PB15:1
11	237.788	0.120	PB15:3
12	294.516	0.125	PB15:1
13	351.243	0.125	PB15:1
14	407.970	0.120	PB15:3

Individual map similarity results are shown in Table 10 for Sample A: quick scan. The only component in the mixture identified was NB1 with a similarity score of 0.307. This results is attributed to NB1 contributing the majority of the bands to the spectrum at each image position. To identify that there were two components in this sample, ALS-MCR would have to be used.

Table 10: HQI scores with respect to each image position for Sample A: quick scan.

Position Number	Image Position, x axis, (μm)	HQI Score	Pigment Identified
1	192.624	0.307	NB1
2	-178.338	0.307	NB1
3	-164.052	0.307	NB1
4	-149.767	0.307	NB1
5	-135.481	0.307	NB1
6	-121.195	0.307	NB1
7	-106.909	0.307	NB1
8	-92.624	0.307	NB1
9	-78.338	0.307	NB1
10	-64.052	0.307	NB1
11	-49.767	0.307	NB1
12	-35.481	0.307	NB1
13	-21.195	0.307	NB1
14	-6.909	0.307	NB1
15	7.376	0.307	NB1

Individual map similarity results are shown in Table 11 for Sample B: quick scan, tartrazine and sunset yellow. The main component in the mixture was identified as tartrazine with a similarity score of 0.605 for 8 out of the 10 points. The second component, sunset yellow, was only identified twice as the primary component with a much lower similarity score of 0.333, which would need to be verified using a similar method to the one described for Sample A: quick scan.

Table 11: HQI scores with respect to each image position for Sample B: quick scan.

Position Number	Image Position, x axis, (μm)	HQI Score	Pigment Identified
1	-118.380	0.605	Tartrazine
2	-76.475	0.605	Tartrazine
3	-34.570	0.605	Tartrazine
4	7.335	0.333	Sunset Yellow
5	49.240	0.605	Tartrazine
6	91.145	0.605	Tartrazine
7	133.050	0.605	Tartrazine
8	174.955	0.605	Tartrazine
9	216.860	0.605	Tartrazine
10	258.765	0.333	Sunset Yellow

Tables 9 - 11 show that to an extent similarity matches can be done without using ALS-MCR from mixtures but would give lower confidence matches with the secondary components. Using ALS-MCR resolves this issue improving the similarity scores of both components as well as indicating the real variation of components along the image.

4.4 Conclusion & Future Research

Overall this project showed that using spectroscopic imaging, ALS-MCR, and the library matching algorithm can be used as a method for identification of pure components in a mixture, such as paintings. While the code developed in Appendix H does require the user to have some background in chemometrics, we hope to make this more ‘user-friendly’ including the possible development of a graphical user interface (GUI) for this type of analysis.

REFERENCES

- [1] Bruni, S.; Maino, G.; Vaccaro, C.; Volpe, L. *Knowledge, Diagnostics and Preservation of Cultural Heritage*, **2012**, *48*, 129-134.
- [2] Li, J.; Yao, L.; Hendriks, E.; Wang, J. *Z. IEEE TR PARTS MATER*, **2012**, *34*, 1159-1176.
- [3] Burgio, L.; Clark, R. J. H.; Hark, R. R. *J. Raman Spectrosc.* **2009**, *40*, 2031-2036.
- [4] Wolbers, R.; Stavroudis, C. *Aqueous methods for the cleaning of paintings*; Archetype Publications: London, 2000; Chapter 30.
- [5] Stoner, J.H.; Rushfield, R. *Conservation of Easel Paintings*; Routledge: London, 2012.
- [6] Hayez, V.; Costa, V.; Guillaume, J.; Terry, H.; Hubin, A. *The Analyst* **2005**, *130*, 550-556.
- [7] Chong, N. S.; Houk, R. S. *Appl. Spectrosc.* **2005**, *41*, 66-74.
- [8] Larkin, P. *IR and Raman Spectroscopy: Principles and Spectral Interpretation*; Elsevier: Amsterdam, 2011.
- [9] Castro, K ; Perez, M ; Rodriguez-Laso, M. D.; Madariaga, J. M. *Anal. Chem.* **2003**, *75*, 214A-220A.
- [10] Platania, E.; Lombardi, J. R.; Leona, M.; Shibayama, N.; Lofrumento, C.; Ricci, M.; Becucci, M.; Castellucci, E. *J. Raman Spectrosc.* **2014**, *45*.
- [11] Zhang, X. ; Tauler, R. *Analytica Chimica Acta.* **2013**, *762*, 25-38.
- [12] Kumar, K.; Mishra, A. K. *Chemometr. Intell. Lab. .* **2012**, *116*, 78-86.

- [13] Azzouz, T.; Tauler, R. *Talanta*. **2008**, *74*, 1201-1210.
- [14] Mendieta, J.; Diaz-Cruz, M.S.; Esteban, M.; Tauler, R. *Biophys. J.*. **1998**, *74*, 2876-2888.
- [15] Downs, R. T. *The RRUFF Project: an integrated study of the chemistry, crystallography, Raman and infrared spectroscopy of minerals..* <http://rruff.info/> (accessed Nov. 2012).
- [16] Castro, K.; Perez-Alonso, M.; Rodriguez-Laso, M. D.; Fernandez, L. A.; Madariaga, J. M. *Anal. Bioanal. Chem.* **2005**, *382*, 248-258.
- [17] Bell, I. M.; Clark, R. J. H.; Gibbs, P. J. *Spectrochimica Acta A*, **1997**, *53A* 53A, 2159-2179.
- [18] McCreery, R. *Raman Spectroscopy for Chemical Analysis*; John Wiley and Sons: New York, 2005.
- [19] Harwood, L. M. ; Moody, C. J. ; Percy, J. M. *Experimental Organic Chemistry, Standard and Microscale*; Blackwell Science: Malden, MA, US, 1999.
- [20] Eaton, J.W. ; Bateman, D. ; Hauberg, S. *GNU Octave Manual*; Network Theory Limited: Boston , 2002.
- [21] Price, B.A; Pretzel, M.; Lomax, S.Q.; Davis, C.; Carlson, J.H. *Revised JCAMP-JDX Spectral File Format for Submissions to the Infrared & Raman Users Group IRUG Spectral Database.* <http://www.irug.org/> (accessed Oct. 2013).
- [22] Rodriguez, J. D. ; Westenberger, B. J. ; Buhse, L. F. ; Kauffman, J. F. *ANALYST*. **2011**, *136*, 4232-4240.
- [23] Bowie, B. T. ; Chase, B. D. ; Griffiths, P. R. ; Fernandez, L. A. ; Madariaga, J. M. *Appl. Spectrosc.*. **2000**, *54*, 164A-173A.

- [24] Smith, B. C. *Infrared Spectral Interpretation: A Systematic Approach* ; CRC Press: Boca Raton, Fl, US, 1998.
- [25] Lo, S.; Brown,C.W. *Appl. Spectrosc.* **1991**, *45*, 1628-1632.
- [26] De Maesschalck, R.; Jouan-Rimbaud, D.; Massart, D.L. *Chemometr. Intell. Lab.* **2000**, *50*, 1-18.
- [27] Nyden, M. R ; Pallister, J. E. ; Sparks, D.T. ; Salari, A. *Appl. Spectr..* **1978**, *41*, 63-66.
- [28] Skoog,D.A; West, D.M.; Holler, F.J. *Fundamentals of Analytical Chemistry, 7th Ed.*; Saunders College Publishing: Philadelphia,Pa,US , 1995.
- [29] Garrido, M ; Rius, F.X. ; Larrechi, M.S.. *Anal. Bioanal. Chem..* **2008**, *390*, 2059-2066.
- [30] Shariati-Rad, M. ; Hasani, M. *Biochimie.* **2009**, *91*, 850-856.
- [31] Tauler, R. ; Barcelo, D. *Trends in Analytical Chemistry.* **1993**, *12*, 319-327.
- [32] Goodall , R. A. ; David, B. ; Kershaw, P. ; Fredericks, P. M. *J. of Archaeological Science.* **2009**, *136*, 2617-2624.
- [33] Kendall, D. *Anal. Chem..* **1953**, *25*, 382-389.

A Acetaminophen Standard Values

Table A.1: NIST acetaminophen standard.¹⁸

Wavenumber / cm^{-1}	Standard Deviation	Relative Intensity
213.3	1.77	17
329.2	0.52	11
390.9	0.76	25
465.1	0.30	11
504.0	0.60	11
651.6	0.50	33
710.8	0.68	17
797.2	0.48	45
834.5	0.46	14
857.9	0.50	82
968.7	0.60	12
1105.5	0.27	7
1168.5	0.65	70
1236.8	0.46	75
1278.5	0.45	42
1323.9	0.46	100
1371.5	0.11	38
1515.1	0.70	9
1561.5	0.52	37
1648.4	0.50	73

B Pigment Measurement Parameters

Table B.2: Pigment measurement parameters for library.

Pigment	Excitation Wave-length	Acquisition Time(sec)	Repetitions	Baseline Correction (polynomial degree)	Power at Sample (mW)
AR18	785	10	20	8	15.2
AR52	488	30	30	6	5.2
PR254	785	3	15	8	15.2
PY83	785	6	25	8	8.2
PY139	785	10	30	8	8.1
Tartrazine	785	15	10	8	15.2
Sunset Yellow	785	10	10	4	15.2
AG25	488	45	20	12	3.5
NB1	785	2	60	8	15.2
PB15:1	785	50	5	8	8
PB15:3	785	50	5	8	8
Acid Violet 17	514	20	20	6	4.3
PV23	785	10	30	8	6.7
PW4	785	10	10	6	15.2
PW6	785	10	10	4	15.2

C Calibrated Library Spectra

C.1 Acid Red 18 (AR18)

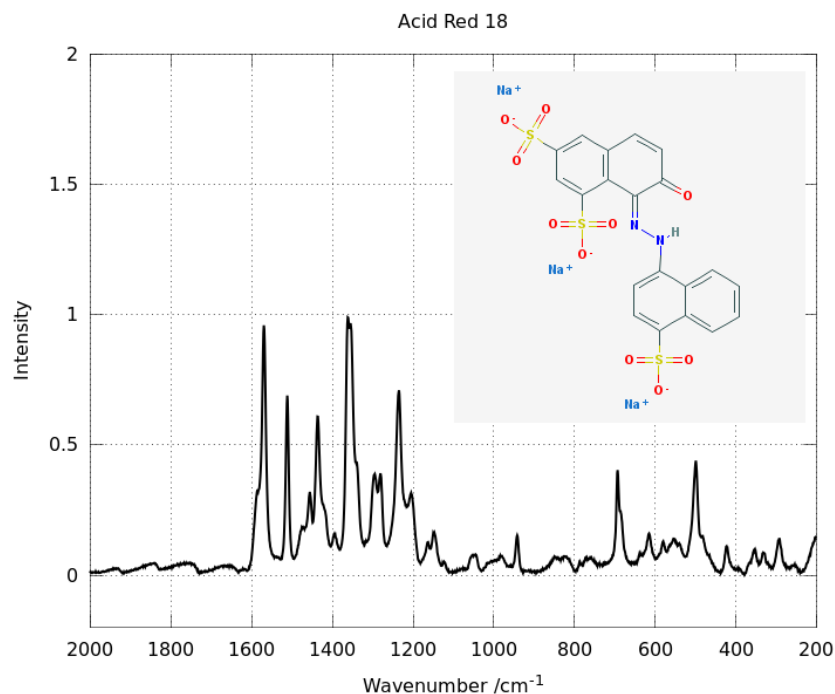


Figure C.1: Calibrated Raman spectrum for AR18.

Common Names Acid Red 18; Ponceau 4R; New Coccine

CAS # 2611-82-7

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 10 seconds (acquisition time); 10 repetitions; 15.2 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.2 Acid Red 52 (AR52)

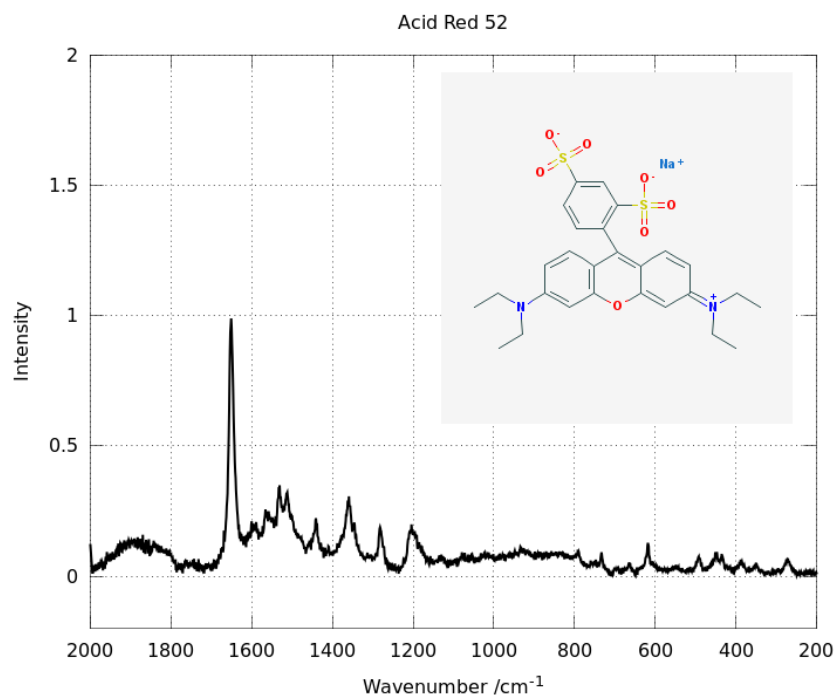


Figure C.2: Calibrated Raman spectrum for AR52.

Common Names Acid Red 52; Sulforhodamine B

CAS # 3520-42-1

Raman Setup 488nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 30 seconds (acquisition time); 30 repetitions; 5.2 mW (Laser Power at Sample); Baseline Correction 6th degree polynomial

C.3 Pigment Red 254 (PR254)

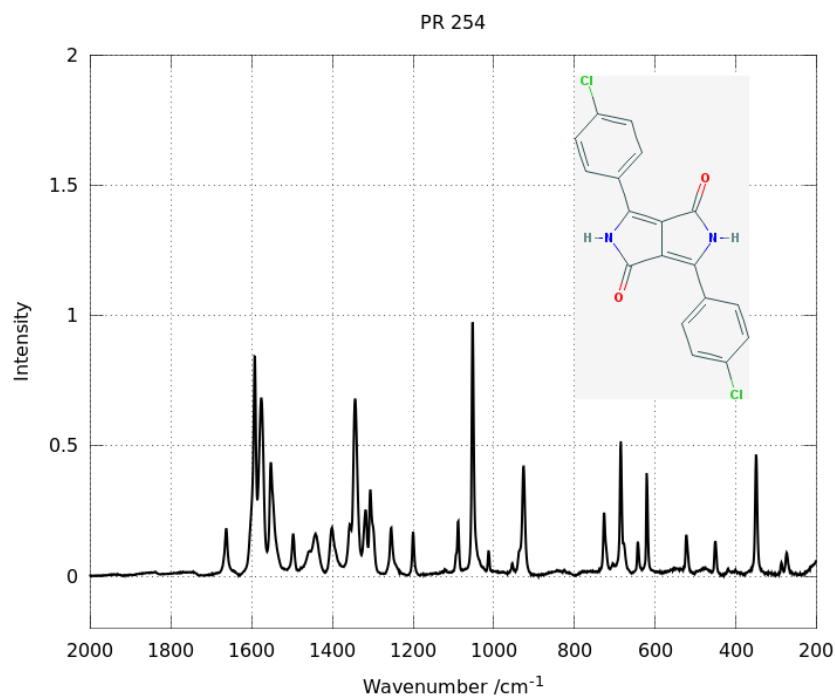


Figure C.3: Calibrated Raman spectrum for PR254.

Common Names Pigment Red 254; Fast Brilliant Red S2BL

CAS # 84632-65-5

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 3 seconds (acquisition time); 15 repetitions; 15.2 mW (Laser Power at Sample); Baseline Correction 6th degree polynomial

C.4 Pigment Yellow 83 (PY83)

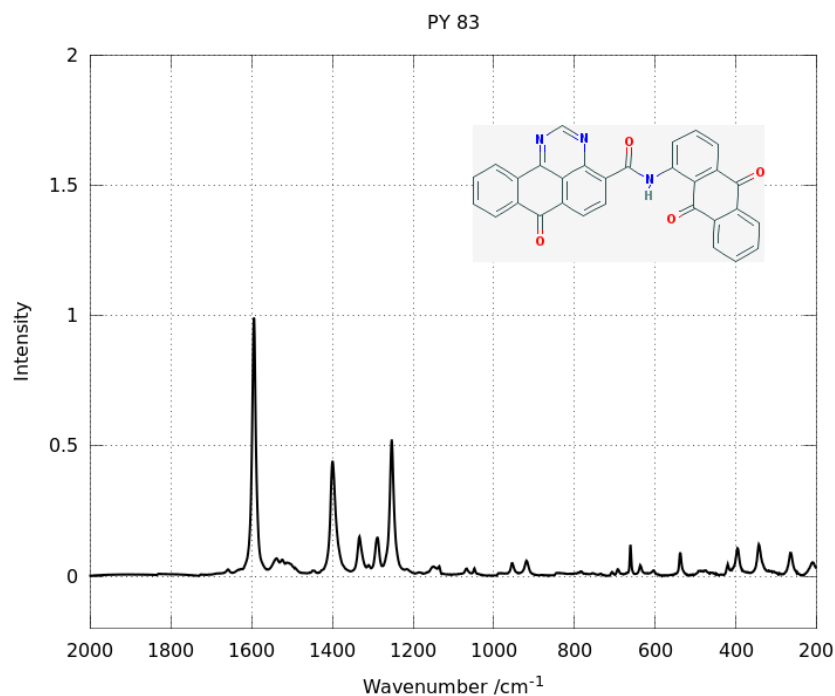


Figure C.4: Calibrated Raman spectrum for PY83.

Common Names Pigment Yellow 83

CAS # 5567-15-7

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 6 seconds (acquisition time); 25 repetitions; 8.2 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.5 Pigment Yellow 139 (PY139)

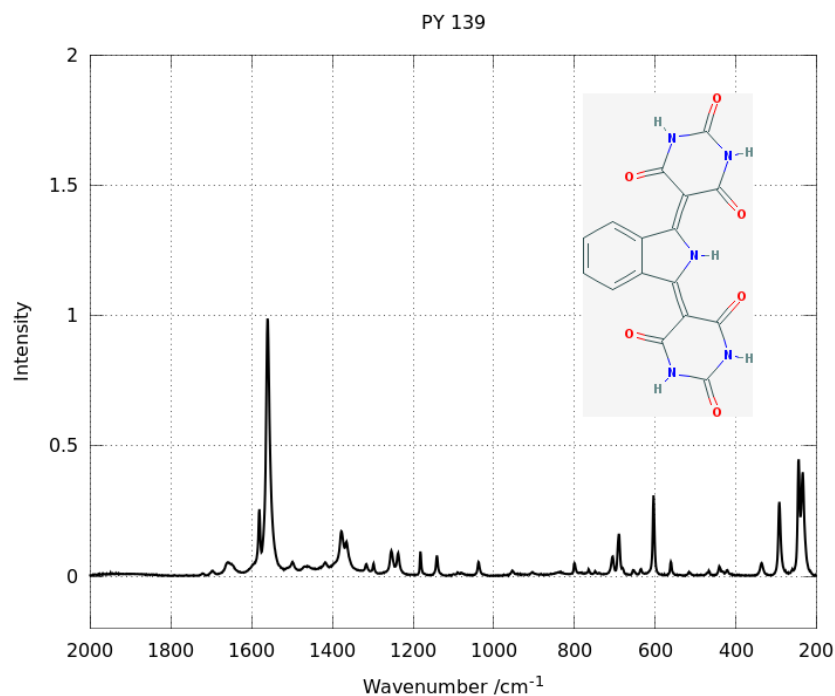


Figure C.5: Calibrated Raman spectrum for PY139.

Common Names Pigment Yellow 139

CAS # 36888-99-0

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 10 seconds (acquisition time); 30 repetitions; 8 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.6 Tartrazine

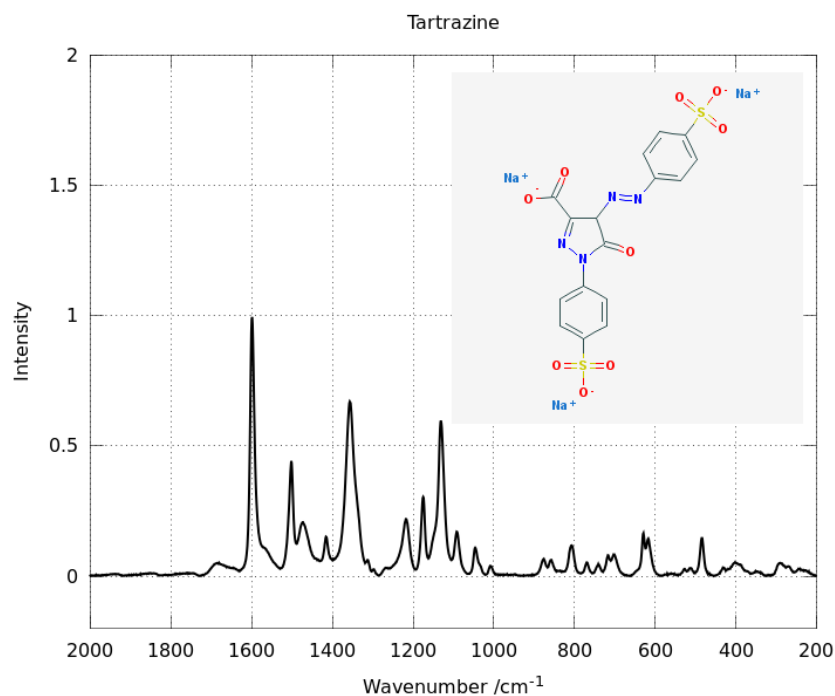


Figure C.6: Calibrated Raman spectrum for tartrazine.

Common Names Tartrazine; FD& C Yellow 5

CAS # 5567-15-7

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 15 seconds (acquisition time); 10 repetitions; 15.2 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.7 Sunset Yellow

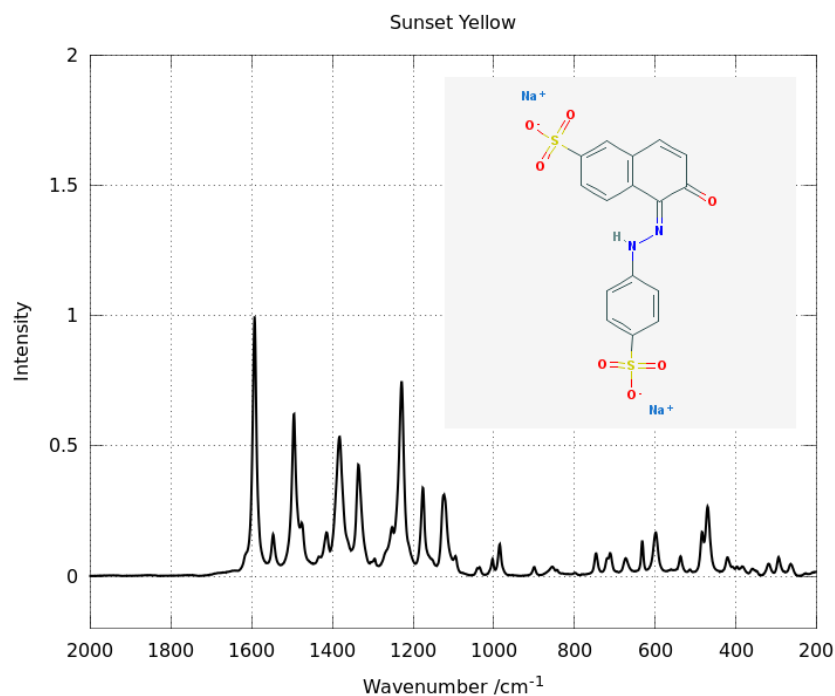


Figure C.7: Calibrated Raman spectrum for sunset yellow.

Common Names Sunset Yellow; FD& C Yellow 6

CAS # 2783-94-0

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 10 seconds (acquisition time); 10 repetitions; 15.2 mW (Laser Power at Sample); Baseline Correction 4th degree polynomial

C.8 Acid Green 25 (AG25)

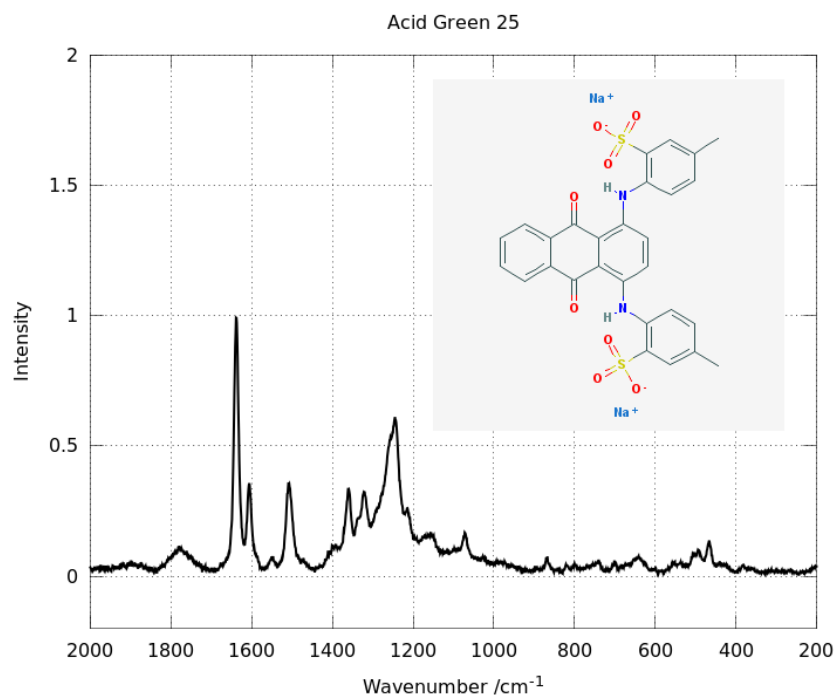


Figure C.8: Calibrated Raman spectrum for AG25.

Common Names Acid Green 25

CAS # 4403-90-1

Raman Setup 488nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 45 seconds (acquisition time); 20 repetitions; 3.5 mW (Laser Power at Sample); Baseline Correction 12th degree polynomial

C.9 Indigo (NB1)

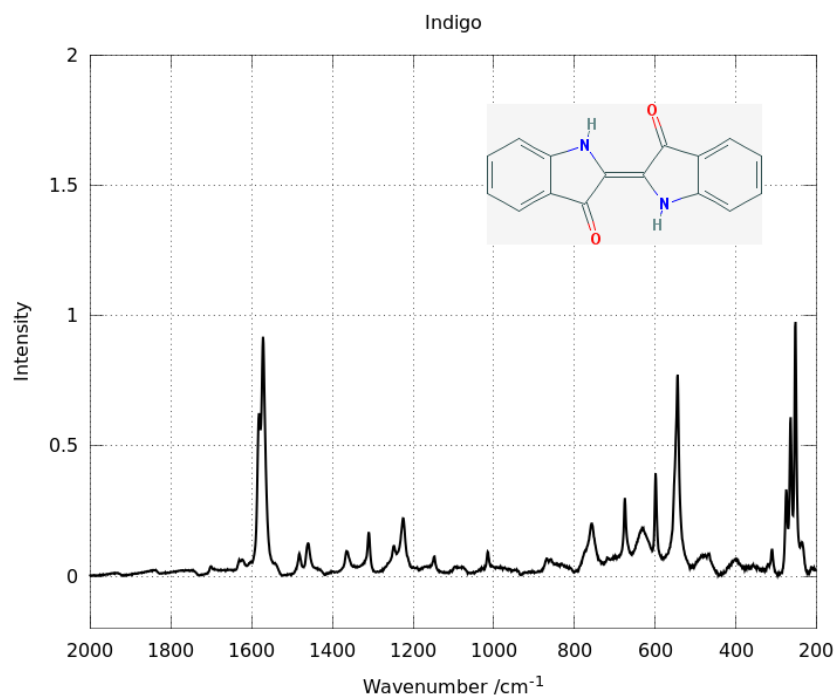


Figure C.9: Calibrated Raman spectrum for NB1.

Common Names Indigo; Natural Blue 1 (NB1)

CAS # 482-89-3

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 50 seconds (acquisition time); 5 repetitions; 15.2 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.10 Pigment Blue 15:1 (PB15:1)

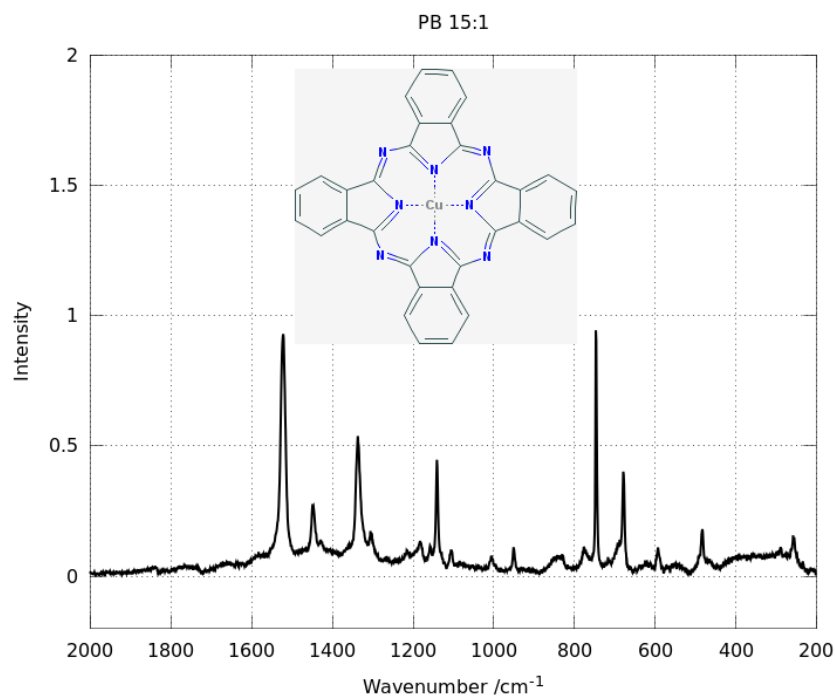


Figure C.10: Calibrated Raman spectrum for PB15:1.

Common Names Pigment Blue 15:1; Copper Phthalocyanine

CAS # 12239-87-1

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 50 seconds (acquisition time); 5 repetitions; 8 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.11 Pigment Blue 15:3 (PB15:3)

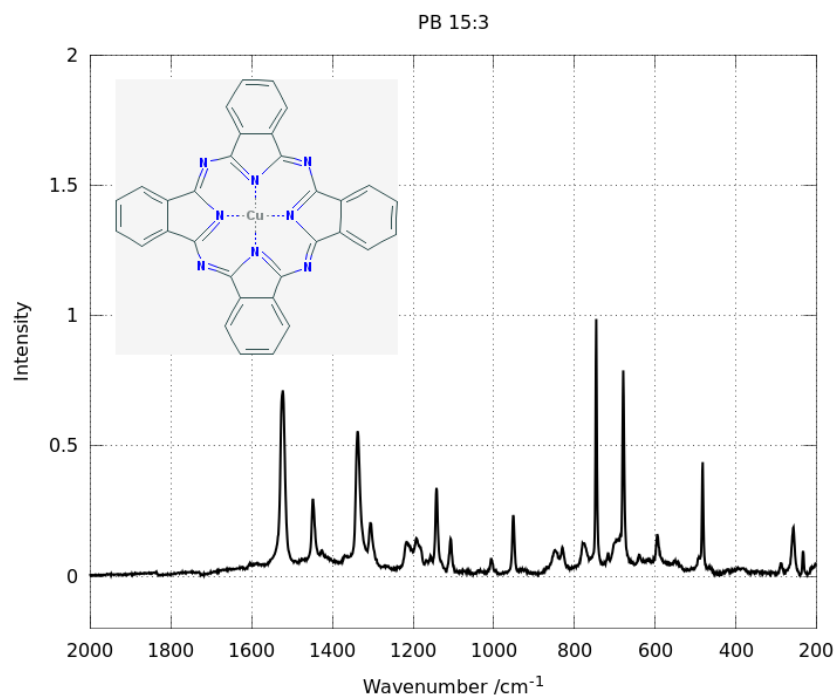


Figure C.11: Calibrated Raman spectrum for PB15:3.

Common Names Pigment Blue 15:3; Beta Copper Phthalocyanine

CAS # 147-14-8

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 50 seconds (acquisition time); 5 repetitions; 8 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.12 Acid Violet 17 (AV17)

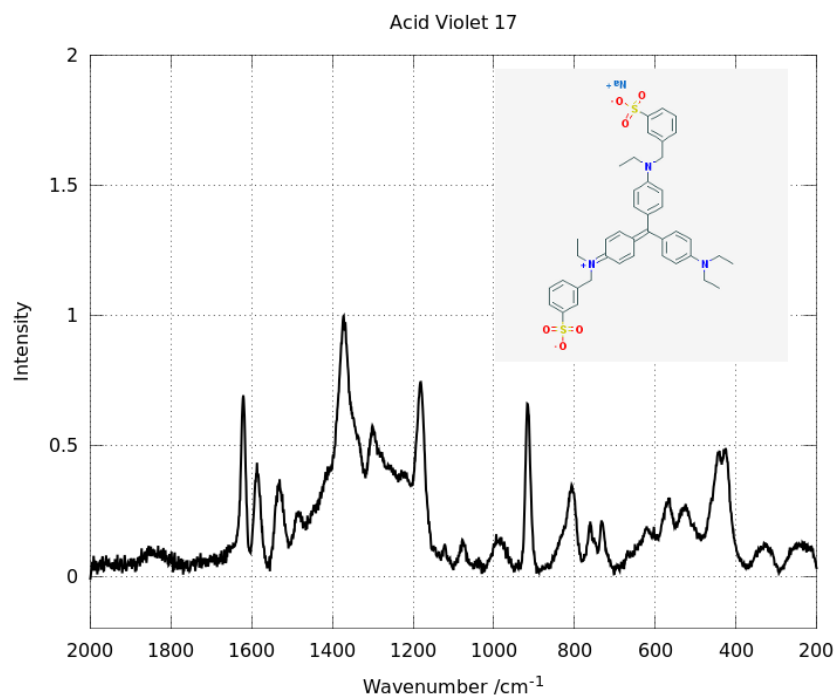


Figure C.12: Calibrated Raman spectrum for AV17.

Common Names Acid Violet 17

CAS # 4129-84-4

Raman Setup 514nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 20 seconds (acquisition time); 20 repetitions; 4.3 mW (Laser Power at Sample); Baseline Correction 6th degree polynomial

C.13 Pigment Violet 23 (PV23)

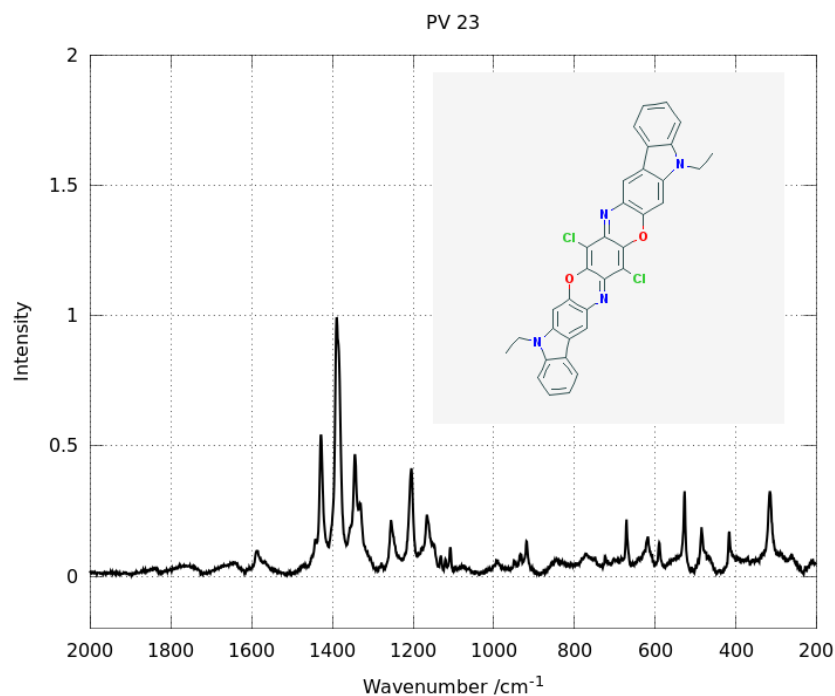


Figure C.13: Calibrated Raman spectrum for PV23.

Common Names Pigment Violet 23; Permanent Violet RL

CAS # 6358-30-1

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 10 seconds (acquisition time); 30 repetitions; 6.7 mW (Laser Power at Sample); Baseline Correction 8th degree polynomial

C.14 Pigment White 4 (PW4)

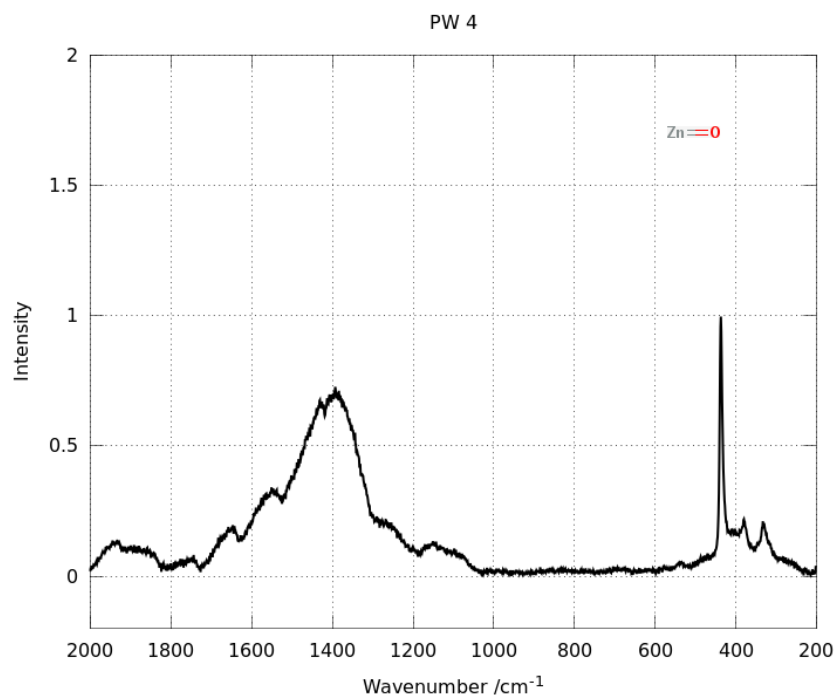


Figure C.14: Calibrated Raman spectrum for PW4.

Common Names Pigment White 4; Zinc White

CAS # 1314-13-2

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 10 seconds (acquisition time); 10 repetitions; 15.2 mW (Laser Power at Sample); Baseline Correction 6th degree polynomial

C.15 Pigment White 6 (PW6)

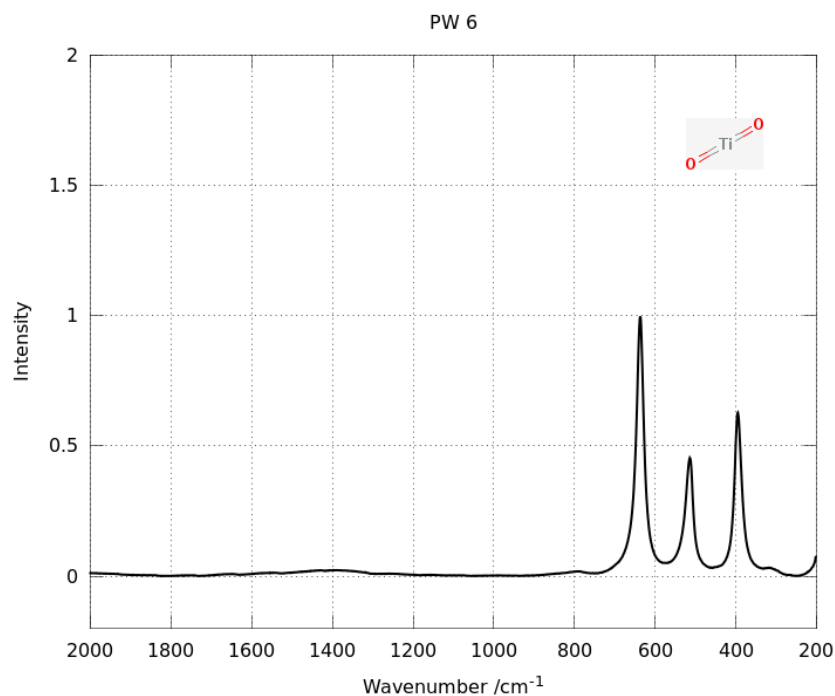


Figure C.15: Calibrated Raman spectrum for PW6.

Common Names Pigment White 6; Titanium White

CAS # 13463-67-7

Raman Setup 785nm(laser line); 10x Obj (Objective used); 1800 lines/mm (grating)

Measurement Parameters 10 seconds (acquisition time); 10 repetitions; 8.2 mW (Laser Power at Sample); Baseline Correction 4th degree polynomial

D CODE FOR CALIBRATED RAMAN LIBRARY

```
1 % This program will convert the Raman spectrum to a calibrated Raman
2 %spectrum that can then be added to a library and used in the matching ...
   program.
3 clear
4 clf
5
6 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
7 %%% User Input %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
8 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
9
10  refspec='~/Desktop/THESIS/cep/codethesis/data/p20140603-001.bc.txt '
11      % Acetaminophen Spectrum (baseline corrected on the Raman ...
      software)
12  samplespec='~/Desktop/THESIS/cep/codethesis/data/p20140603-002.bc.txt '
13      % Pigment Spectrum (baseline corrected on the Raman software)
14  fighole='~/Desktop/THESIS/cep/codethesis/figures '
15      % Where the user wants to save the plot of the 'calibrated' ...
      spectrum
16  figname='pw4final'
17      % Label of the figure name
18  name='Pigment White 4'
19      % Title of the plot
20
21 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
22 %%% Making Figures %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
23 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
24
25  lw = 2 ;          %%% Line Width for figures
```

```

26         ftype = '-dpng';    %% File Format for figures
27         fres = '-r100';    %% Resolution of figures
28
29         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
30         %% Loading the files %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
31         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
32
33         load(refspec);
34         load(samplespec);
35
36         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
37         %% Reading the spectra into the code %%%%%%%%%
38         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
39
40         [xac,specac,hdr]=readdata(refspec ,1);
41         [xunk,specunk,hdr]=readdata(samplespec ,1);
42
43
44         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
45         %% Correction Function: Finding x position %%%%%%%%%
46         %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
47
48         xnist= [390.9 857.9 1323.9 1648.4];%%SRM x
49         nxnist = length(xnist)
50         for i = 1:nxnist
51             tmpidx(i) = findbn(xac,xnist(i));
52         end
53         %%%this is the size of the window that is used to find the
54         %%%band max
55         chunk = 100
56

```

```

57 for i = 1:nxnist
58     rng = (tmpidx(i)-chunk):(tmpidx(i)+chunk);
59     [maxy(i), tmpbnx(i)] = max(ref(rng));%%maxy is the experimental y
60     wntmp(i) = xr(rng(tmpbnx(i)));%experimental x
61     bnx(i) = findbn(xr,wntmp(i));
62 end
63
64
65 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
66 %% hnorm Section Correcting Y %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
67 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
68     newyac=hnorm(newyac)
69     newyunk=hnorm(newyunk)
70     %plot(xac,newyac,'Color','b','LineWidth',lw)
71
72 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
73 %% Polyfit Section Correcting X %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
74 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
75
76     xnist= [390.9 857.9 1323.9 1648.4]';
77     order=1
78     px=polyfit(xac(bnx),xnist,order)
79     newxac = polyval(px,xac)
80     newxunk=polyval(px,xunk)
81     %plot(newxac,newyac+.2,'Color','g','LineWidth',lw)
82     %plot(newxunk, newyunk , 'Color','r','LineWidth',lw)
83
84
85 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
86 %% Correcting the xaxis to 200-2000 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
87 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

88
89     idx=200:2000;
90
91     for i= 1:length(idx)
92         bn(i)=findbn(newxunk,idx(i));
93     end
94     newx= newxunk(bn);
95     nspecunk=newyunk(bn);
96     [xunkmatch,specunkmatch] = dsmatch(newx,nspecunk);
97
98
99     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
100    %% Plotting the corrected spectra %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
101    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
102
103     plot(xunkmatch,specunkmatch, 'Color','r','LineWidth',lw)
104     axis([200 2000 -.2 1.2])
105     set(gca,'XDir','reverse');
106     title(name)
107     xlabel('Wavenumber /cm^-1')
108     ylabel('Intensity')
109     legend('Corrected Spectra')
110     ofile=[fighole '/' figname ]
111     print( ftype, fres, ofile)
112
113
114     return

```

E *hnorm* CODE

```
1 %this program will normalize a row vector
2 %vecout will be between 0 and 1
3 %created Nov 9, 2010
4 %last modified Nov 9, 2010
5 %
6 %usage: vecout = hnorm(vecin, type,{wavelength, x })
7 %
8 %
9 % type=1 => normalize vector by max value
10 % type=2 => normalize by area
11 % type=3 => normalize by specific band (in Wavenumber or wavelength)
12 %
13 % wavelength and x are optional
14 % wavelength is the wavelength that you want ot normalize to
15 % x is the vector of wavelengths that corepsond to the spectrum
16 %
17 function vecout = hnorm(vecin, type,wl,x)
18 v = mkrow(vecin);
19 if (nargin == 1)
20     type = 1;
21 end
22
23 if (nargin < 3)
24     [crp, bn] = max(v);
25 end
26
27 if (nargin ==4)
28     [indxdiff, bn] = min(abs(x-wl))
```

```
29 end
30
31 switch type
32     case(1)
33         tmp = v- min(v);
34         vecout = tmp/max(tmp);
35     case(2)
36         tmp = v- min(v);
37         vecout = tmp/norm(tmp);
38     case(3)
39         tmp = v-min(v);
40         vecout = tmp/tmp(bn);
41     otherwise
42 end
43
44
45 return
```

F *dsmatch* CODE

```
1 %This program will interpolate your data down to single
2 %wavenumber/lengths
3 %
4 %
5 % usage: [xnew,nspec] = dsmatch(old, ospec)
6 %
7 function [xnew,nspec] = dsmatch(xold,ospec)
8     xold = mkrow(xold);
9     ospec = mkrow(ospec);
10    if (xold(1) ~= min(xold));
11        xold = fliplr(xold);
12        ospec = fliplr(ospec);
13    end
14
15    xnew = round(xold):round(xold(length(xold)));
16    nspec = interp1(xold,ospec,xnew,'linear','extrap');
17    return
```

G HQI CODE

```
1 %This program will perform a hit quality measure of an unknown vs. the
2 %library
3 clear
4 clf
5 %% Loading the unknown data, for this program I chose tartrazine as my ...
   unknown.
6 datahole = './data/'
7 tartdat='tart.jdx'
8 tart= [datahole tartdat];
9 [xt,spect,hdr]=readdata(tart,6);
10
11 % Loading the library data, the alllib.mat file is an easy
12 % to load file vs loading each individual jdx library file.
13 ifilelib = [datahole 'alllib.mat']
14 load(ifilelib)
15
16 for i = 1:length(data)
17     Alib(i,:) = data(i).spec;
18     disp(size(data(i).spec))
19     x = data(i).x;
20 end
21
22 [nspec,npts] = size(Alib)
23
24 %% Loop for HQI
25 for i = 1:nspec
26     Ref = spect; %% the intensity values of the tartrazine
27     Kspec = Alib(i,:);
```

```

28         HQIlib(i) = dot(Ref,Kspec)^2/(dot(Kspec,Kspec)+dot(Ref,Ref));
29     end
30     %% This sorts the similarity scores from lowest (0) to highest (1)
31     [srt,idx] = sort(HQIlib);
32     %%% This displays the names from the library jdx files next to the HQI ...
        scores
33     for i = 1:nspec
34         disp([ num2str(HQIlib(idx(i))) ' ' data(idx(i)).names ' ' ...
                num2str(idx(i))])
35     end
36     return

```

H ALS-MCR CODE

```
1 % This is the program that is used to identify pure component spectra
2 % from mixtures using spectroscopic imaging to acquire the data and
3 % ALS-MCR as the method for extracting the pure component spectra.
4
5
6 clear
7 clf
8 % Making publication quality worthy figures
9 labelsandsuch
10 flag = 1%load raw data from text file
11 flag =2 %load matlab file (faster loading)
12 fighole = './figures/' % Where figures will be saves
13 fname = 'test'
14 ftype = '-dpng' % PNG is the file format that the figures will be saved as.
15 fres = '-r300' % Resolution of the figures
16 fext = ['. ' ftype(3:length(ftype))]
17 ofile = [fighole fname fext]
18
19 datahole = './data/' % Folder where all the data is stored
20
21 ifilelib = [datahole 'alllib.mat'] %Library spectra
22 load(ifilelib)
23
24 for i = 1:length(data)
25     Alib(i,:) = data(i).spec;
26     disp(size(data(i).spec))
27     x = data(i).x;
28 end
```

```

29
30
31 %%%%Loading the Data
32 datadir = './data/'
33
34 acfile = 'P20140901.001.txt' % Acetaminophen File as a Raman referenece ...
      standard
35 mixfile = 'sustart_3.txt' % Mixture file
36 ifile = [datadir mixfile]
37
38 % Loading the mixture file
39 d = load(ifile);
40 % Seperating the mixture file into components.
41 [nr,nc] = size(d);
42 xd = d(2:nr,1); % Space
43 y = d(2:nr,2); % Intensity
44 v = d(1,3:nc); % Wavenumber
45 npts = length(v);
46 nx = length(unique(x));
47 A = d(2:nr,3:nc);
48 xm = v;
49
50 save ./data/tmpdata.mat A xm v
51
52
53
54
55
56 %%% Normalize and Baseline correction for the map
57 [nspec, npts] = size(A);
58 %points along the baseline for correction

```

```

59 idx = [1 176 387 480 611 765 905 994 1097 1345 1515 2009 npts];
60 order =1
61
62 for i = 1:nspec
63     A(i,:) = hnorm(A(i,:));% Used to set normalize intenisty to max of 1.
64     fity(i,:) = fitbaseline(xm,A(i,:),xm(idx),order);
65     A(i,:) = A(i,:) - fity(i,:);
66 end
67
68
69 %%%% Start of the ALS-MCR portion of the code
70 % The code that we use requires and initial guess of the K values.
71 % We know that there are 2 components in the mixture, so we guess 2 K ...
    values
72
73 %Guess = 4th spectra in the map and the last
74 Ki = [A(4,:);A(nspec,:)];
75
76 %%%guess = first and random
77 %Ki = [A(1,:);rand(1,npts)];
78
79 %%%guess = last aad random stuff
80 %Ki = [A(nspec,:);rand(1,npts)  ];
81
82 clear K
83 options = [1;%non negativity in C
84           1;%non negativeity in K
85           0;% closer in C
86           1;
87           1;
88           1;

```

```

89         1;
90         1];
91
92 tolerance = 0.01;
93 [K, C] = alsKgf(A,Ki,options,tolerance); % Loop for ALS-MCR
94
95 % Correcting the Mean spectra (A) to have the same interval spacing and
96 % and wavenumbers as the library
97 Amean = mean(A);
98 minwn = 200; % Wavenumber range 200-2000 cm-1
99 maxwn = 2000;
100 tmp = interp1(xm,Amean,minwn:maxwn,'linear','extrap');
101 [xm,Am ] = dsmatch(minwn:maxwn,tmp); % interval spacing of 1
102
103
104 %% dsmatch the K extracts to have the same interval spacing and
105 % wavenumber range as the library
106 [npc,npts] = size(K)
107 for i = 1:npc
108     minwn = 200; % Wavenumber range 200-2000 cm-1
109     maxwn = 2000;
110     tmp = interp1(xm,K(i,:),minwn:maxwn,'linear','extrap');
111     [xm,Kds(i,:) ] = dsmatch(minwn:maxwn,tmp); % interval spacing of 1
112
113 end
114
115
116 [nspec,npts] = size(Kds)
117
118 [nspec,npts] = size(Alib);
119

```

```

120 %%%%Begin the output
121 br = '#####';
122
123 %%%%%%%%% Similarity between the library and the pure component spectra
124 %%%%%%%%% extracted using ALS-MCR
125 disp([br 'Here is the output report' br])
126 for j = 1:npc
127
128     for i = 1:nspec
129         Ref = Alib(i,:);
130         Kspec = Kds(j,:);
131         HQI(i,j) = dot(Ref,Kspec)^2/(dot(Kspec,Kspec)*dot(Ref,Ref));
132     end
133 end
134 for j = 1:npc
135     [srt,idx] = sort(HQI(:,j));
136
137     for i = 1:nspec
138         disp([ num2str(HQI(idx(i),j)) ' ' data(idx(i)).names
139 ' ' num2str(idx(i))])
140     end
141     msg = sprintf('For MCR resolved component number %i\n',j);
142     disp(msg)
143     msg = sprintf('the best match is %i %s with similarity of %f\n',idx(i),
144 data(idx(i)).names,HQI(idx(i),j));
145     disp(msg)
146     disp(br)
147     mch(j) = idx(nspec)
148 end
149 %%% Plot of the library components,the pure component spectra,
150 % and the mean spectra from the map.

```

```

151 plot(x,Alib(mch(1),:),x,Alib(mch(2),:),xmds,hnorm(Kds(1,:),xmds,
152 hnorm(Kds(2,:),xm,hnorm(mean(A))))
153 xlabel('Wavenumber /cm^{-1}')
154 ylabel('Normalized Intensity')
155 axis([200 2000 -0.05 1])
156 grid on grid minor
157 set(gca,'Xdir','Reverse')
158 legend(data(mch(1)).names,data(mch(2)).names,'K1','K2','mean')
159
160 disp('Mean spec of map match')
161 disp(br)
162
163 %%%% Similarity between mean of map data (bulk) and library
164 for i = 1:nspec
165     Ref = Alib(i,:);
166     Kspec = Am;
167     HQImeanSpec(i) = dot(Ref,Kspec)^2/(dot(Kspec,Kspec)*dot(Ref,Ref));
168 end
169 [srt,idx] = sort(HQImeanSpec);
170 for i = 1:nspec
171     disp([ num2str(HQImeanSpec(idx(i))) ' '
172 data(idx(i)).names ' ' num2str(idx(i))])
173 end

```

I ALS-MCR: K GUESS FUNCTION

```
1 %this is the als engine to iteratively refine guesses in C or K matrix
2 % currently this function only works with Ki guesses
3 %
4 %usage: [K, C] = alsKgf(A,Ki,options,tolerance)
5 % notes:
6 % all spectra are row vectors
7 % options is a column vector of flags
8 % $$$ options = [1;%non negativity in C
9 % $$$           1;%non negativeity in K
10 % $$$          1;% closer in C
11 % $$$          1;% unassigned
12 % $$$          1;% unassigned
13 % $$$          1;% unassigned
14 % $$$          1;% unassigned
15 % $$$          1];% unassigned
16
17
18 function [K, C] = alsKgf(A,Ki,options,tolerance)
19 Ki = mkrow(Ki);
20 [npc, npts] = size(Ki);
21 [nspec, npts] = size(A);
22
23 C = zeros(nspec,npc);
24 K = Ki;
25 %begin convergence test loop
26 decide = 10000000;
27 RSDold = 10000000;
28 while abs(decide) > tolerance
```

```

29
30 C = A*K'*inv(K*K');
31 %closer constraint for C
32 if (options(3) == 1)
33     for i = 1:nspec
34         C(i,:) = ccr(C(i,:),1);
35     end
36 end
37 %non negative constraint in C
38 if (options(1) == 1)
39     for i = 1:npc
40         C(:,i) = cnn(C(:,i))';
41     end
42 end
43 K = inv(C'*C)*C'*A;
44 %non negative constraint in K
45 if (options(2) == 1)
46     for i = 1:npc
47         K(i,:) = cnn(K(i,:))';
48     end
49 end
50 Ag = C*K;
51 RSD = sum(sqrt((sum(A-Ag).^2)/(nspec-2)));
52 decide = (RSD-RSDold)*100;
53 disp([num2str(abs(decide(1,1))) ' must be less than ' num2str(tolerance)])
54 RSDold = RSD; %see Anal. Chem. vol 72 no 9 1956-1963
55 end

```

J ALS-MCR: NON-NEGATIVITY CONSTRAINT

```
1 %this function will apply the non-negativity constraint to a row vector
2 %usage: rowvecout = cnn(rowvecin)
3
4 function out = cnn(v)
5 v = mkrow(v);
6 [nr,nc] = size(v);
7 out = v;
8 for i = 1:nc
9     if v(1,i) < 0.001
10         out(1,i) = 0.00000;
11     end
12 end
13 return
```