2019 IMEKO TC-4 International Conference on
Metrology for Archaeology and Cultural Heritage
Florence, Italy, December 4-6, 2019

Managing complex Synchrotron radiation FTIR micro-spectra from historic bowed musical instruments by chemometrics

Silvia Grassi[1], Giacomo Fiocco[2,3], Claudia Invernizzi[2,4], Tommaso Rovetta[2], Michela Albano[2], Patrizia Davit[3], Monica Gulmini[3], Chiaramaria Stani[5], Lisa Vaccari[5], Maurizio Licchelli[2], Marco Malagodi[2,6]

[1] Department of Food, Environmental, and Nutritional Sciences (DeFENS), Università degli Studi di Milano, via G. Celoria 2, 20133 Milan, Italy, silvia.grassi@unimi.it
[2] Arvedi Laboratory of Non-Invasive Diagnostics, CISRiC, Università degli Studi di Pavia, Via Bell'Aspa 3, 26100 Cremona, Italy, giacomo.fiocco@unipv.it, claudia.invernizzi@unipv.it, tommaso.rovetta@unipv.it, michela.albano@unipv.it, maurizio.licchelli@unipv.it, marco.malagodi@unipv.it
[3] Department of Chemistry, Università di Torino, Via Pietro Giuria 5, 10125, Torino, Italy, patrizia.davit@unito.it, monica.gulmini@unito.it
[4] Department of Mathematical, Physical and Computer Sciences, Università degli Studi di Parma, Parco Area delle Scienze, 7/A, 43124 Parma, Italy
[5] Elettra-Sincrotrone Trieste S.C.p.A., S.S. 14 km 163.5, 34194 Basovizza, Trieste, Italy, chiaramaria.stani@elettra.eu, lisa.vaccari@elettra.eu
[6] Department of Musicology and Cultural Heritage, Università degli Studi di Pavia, Corso Garibaldi 178, 26100 Cremona, Italy

***Abstract* – The intrinsic complexity of the coating systems on historical bowed string musical instruments complicates the characterization of the different materials that have been subsequently spread on the wood. Even more challenging is the identification of their spatial distribution in the sequence of the finishing layers. To face both issues, in this study we combined Synchrotron radiation micro - FTIR in reflection geometry and Chemometrics for the investigation of five micro-samples mounted in cross-sections, removed from musical instruments made by Antonio Stradivari, Francesco Ruggeri and Lorenzo Storioni. A classification model based on Partial Least Squares – Discriminant Analysis aimed at discriminating the organic materials through infrared signals was developed. High model specificity (> 0.9) was reached in prediction, doing the groundwork for the application of a fast and rigorous methodological approach.**

I. INTRODUCTION

The number of historical bowed string musical instruments investigated by the methods of the chemical and physical sciences dramatically rose in the last decade, and the scientific research has increased its role in detecting materials and methods that were employed in the past to obtain extraordinary masterpieces of craftsmanship. The research mainly focused on the nature of the fine Cremonese varnish [1][2] and on the other materials involved in the finishing treatments [3]. In most cases, the raw wood was in fact treated with proteinaceous materials - such as animal glue or casein - and an oil-resinous varnish. The result is a multiple-layered complex system, which may also contain mineral fillers and colourants of organic and inorganic nature.

The scientific investigation of the finishing layers is normally very challenging because of the intrinsic complexity of the coating system, where a large variety of unknown materials can be present, made even more complex due to modification caused by use and maintenance of the musical instrument. Due to the high value of the historical musical instruments, sampling is seldom allowed. In the rare cases when a (micro)-sample can be obtained, this is normally mounted in cross-section after inclusion in epoxy resin and it represents the (available) record of the material history of the musical instrument. Researchers shall extract from it the maximum information, and then it shall be kept for further analyses to be developed in the future. A number of techniques, both micro-invasive and non-invasive, have been used to variously characterize the materials in historical bowed string instruments. A spectroscopic approach [4] can be supported and integrated by imaging and tomographic techniques [5] while chromatographic techniques may lead – destructively - to the accurate identification of the organic constituents [1]. In the present study we considered five cross-sections obtained from samples detached from four bowed string instruments produced by three ancient masters of violin-making art, that have been analysed by Synchrotron Radiation Fourier Transform Infrared (FRIR) microspectroscopy at Elettra Sincrotrone Trieste (beamline SISSI) in reflection geometry. The constraint imposed by the sampling geometry on cross-sections produced complex spectra. It was clear from the get-go that this analytical approach – albeit promising for achieving the characterization of the coating system - would need a huge effort to obtain a reliable picture of the composition of each layer. In order to support the processing steps, in view of extending the use of this analytical technique to a larger number of samples, and in order to extract the maximum information from the analyses, the spectra have been processed with a chemometric approach. Chemometric approaches are in fact needed to solve data with multivariate nature in order to unravel relevant information hidden in spectroscopic signals [6], such as FTIR data. In particular, through an unsupervised exploratory procedure, namely Principal Component Analysis (PCA), it is possible to understand the relationship between all the FTIR variables and to stand out the sample patterns according to variables' weight in the new reduced space defined by the PC components.

Moreover, supervised classification methods (such as Linear Discriminant Analysis, Partial Least Squares-Discriminant Analysis, Support Vector Machines, Artificial Neural Network) permit to define rules to distinguish objects in classes, such as different materials, allowing fast and rigorous materials identification skipping visual inspection of large number of spectra.

Even if the above mentioned chemometric techniques have been widely used in the cultural heritage field [7], their implementation to musical instrument materials [8] gained relevance in the last years. In this scenario, the present work aims at developing a methodological approach useful to manage and interpret large datasets obtained by SR-FTIR in reflection geometry through PCA and PLS-DA.

II. MATERIALS AND METHODS

A. Materials

The experimental plan encompasses the analysis of 5 sub-millimetric samples collected from 4 different bowed string instruments (Table 1): a fragment of a cello made by Francesco Ruggeri during the 17th Century (FR_c), the "Toscano" violin made by Antonio Stradivari in 1690 (AS_v), the Bracco small-violin (LS_sv1, LS_sv2 and LS_sv3) and a private property violin (LS_sv2), both made by Lorenzo Storioni respectively in 1793 and 1790. The samples were collected under high magnification employing a disposable blade scalpel in selected areas of the instruments. After the sampling, they were embedded in epoxy resin (Epofix Struers and Epofix Hardener, 15:2), and then cut as cross-sections. The surface was then dry-polished with silicon carbide fine sandpaper (500-8000 mesh), obtaining a flat surface. At least two layers (up to 5 μm thick) of organic binders were observed in the coating systems of the five selected samples. Moreover, some heterogeneous sub-micrometric inclusions embedded in the binders were identified.

Table 1. List of the samples involved in the project.

Violin Maker	Instrument	Inventory Name	Referred Name
F. Ruggeri	Cello c17th	Rug_3A	FR_c
A. Stradivari	"Toscano" violin 1690	Stra_1B	AS_v
L. Storioni	Private violin 1790	Sto_2B	LS_v
	"Bracco" small violin 1793	Sto_1B_1	LS_sv1
		Sto_1B_2	LS_sv2
		Sto_1E	LS_sv3

B. SR-FTIR Microspectroscopy

Samples were measured at the SISSI beamline – Chemical and Life Sciences branch at Elettra – Sincrotrone, Trieste (Italy) [9]. Measurements were performed on polished samples in reflection geometry exploiting Infrared Synchrotron Radiation (IRSR) using the Bruker Vertex 70v interferometer coupled with the Hyperion 3000 microscope (Bruker Optik GmbH) and a single point MCT (mercury-cadmium-telluride) detector. 512 scans have been averaged in the acquisition spectral range 4000 – 750 cm^{-1}, with a spectral resolution of 4 cm^{-1},at 120 KHz scanner speed. The size of the acquisition points was set at 10 x 30 μm by closing the knife-edge apertures of the Vis-IR microscope accordingly with the sample stratigraphy. The acquisition was carried out in linear map mode with a vertical step size of 10 μm. For each sample, the background was acquired with the same acquisition parameters on a gold substrate. For the interpretation of the bands produced by organic materials, reflection infrared spectra were transformed to absorbance spectra by applying the Kramers-Kronig (KK) algorithm and smoothed (Savitzky-Golay, 20 wavelengths gap size). Data managing and first spectral transformation were performed by Opus 7.5 software.

C. Chemometrics

The spectra were reduced in the range 2000-1400 cm^{-1} and transformed by smoothing (Savitzky-Golay, 11 wavelengths gap size) followed by first derivative (Savitzky-Golay, 11 wavelengths gap size and 2nd order polynomial) and mean centre. PCA was performed to identify sample grouping to be raised to classes to develop classification models based on PLS-DA algorithm. Indeed, PCA is an unsupervised exploratory procedure with effective graphical outputs to visualise the relationships between objects (scores plot) and the loads of the variables (loadings plot) in a new defined space representing the directions of maximum variation of the original data [9]. For classification purposes, the spectra dataset was divided into a calibration and a test set containing 71 (around 80%) and 26 (20%) spectra, respectively. The calibration set contained spectra referred to samples AS_v, FR_c, LS_sv1, and LS_sv3; whereas the spectra collected for LS_sv2 and LS_v were used as test set. Moreover, the models were cross-validated by venetian blind with eight splits.

The classification by PLS-DA applies the bases of PLS regression to a Y dummy and it completes a rotation of the projection to latent variables searching for the maximum separation among classes [10].

All the data analyses were performed in Matlab environment (v. 2016a, Mathworks, Inc., Natick, MA, USA) and using the PLS toolbox (ver. 8.5, Eigenvector Research, Inc., 130 USA) software package.

III. RESULTS AND DISCUSSION

A. *SR-FTIR Microspectroscopy*

In order to focus the attention on the organic components, the spectral region between 2000-1400 cm^{-1} was selected for identification purposes. Diagnostic bands of each organic chemical class were identified as reported in Table 2.

Table 2. Reflection infrared wavenumber values (cm^{-1}), and their assignment, of the diagnostic bands of the studied organic chemical classes.

Material class	Wavenumber (cm^{-1})	Assignment
Epoxy resin (E)	1510	*v*C-C
Varnishes (V)	1720 - 1700	*v*C=O
Proteins (P)	1665 - 1645	*v*C=O (Amide I)
Wood (W)	-	-

Proteins (P in Fig. 1) were discriminated by the amide I band (*v*C=O) with maximum falling in the region 1665 - 1645 cm^{-1}, while varnishes (V in Fig. 1) through the *v*C=O band produced by the ester and acid contributes of oils and resins from approximately 1720 cm^{-1} to about 1700 cm^{-1} [11][12]. The amide II proteinaceous band at around 1550 cm^{-1}, due to the combination of vC-N and δN-H vibrations, as well as the δCH signal at about 1460 cm^{-1} were not considered as markers because, in the first case, the spectral feature did not regularly occur in the data set whereas, in the second case, the CH band is common to most of organic compounds used in the musical instrument field. The presence of the epoxy resin (E in Fig. 1) was clearly highlighted by the intense and sharp band at 1510 cm^{-1}; compared to this signal, in fact, the primary amine band at 1610 and the δCH band at 1460 cm^{-1} are relatively weak and they contribute only limited information [13]. As the epoxy resin was employed to embed the samples, its occurrence in the spectra can be associated to the acquisition area falling close to the upper or lower boundaries of the stratigraphy or to the contamination possibly arising by polishing. The wood (W in Fig. 1), above which the binders were laid to form sequential stratigraphic layers with different functions, resulted to provide very noisy spectra, with no discriminant features in the considered spectral range. However, some characteristic frequencies of the wood (i.e. un-conjugated *v*C=O, conjugated *v*C-O, *v*(aromatic ring) and δOH) [14][15] fall in the region 1750 to 1550 cm^{-1}, and appear very broad in the collected spectra. In addition to spectra clearly referring to one single organic chemical class, also spectra with signals produced by multiple classes (e.g. varnish + proteins + epoxy resin, (U in Fig. 1) occurred. These expected "mixed" profiles originated from acquisition areas at the interface between layers with different composition.

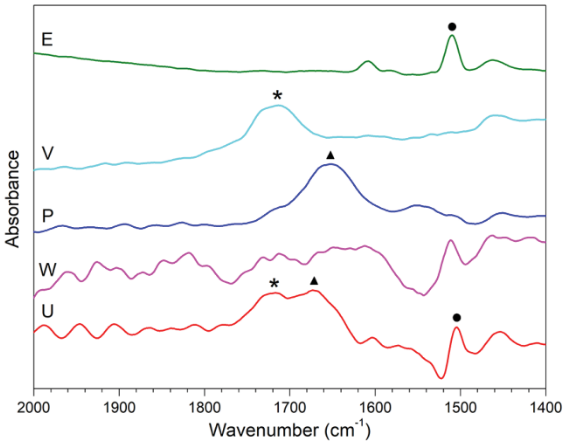

Fig. 1. KK corrected reflection SR-FTIR spectra collected on the LS_sv_2 (E and V), FR_c (P and U) and LS_v (W). Diagnostic bands of epoxy resin (●), varnishes (*) and proteins (▲) are highlighted.

B. *PCA*

In Fig. 2 are reported the results obtained by PCA. Fig. 2a shows the scores plot of PC1 *vs* PC2 accounting for the 54% of total variance. Each object (corresponding to a single spectrum) is coloured according to the most probable material identified through the collected signal. In particular, objects were assigned to a specific group in accordance to the position of the analytical spot in the cross-sections, therefore their assignment mainly reflects the stratigraphy of the sample. It is possible to notice that objects associated to varnish (V) spectral profiles are mostly grouped in the bottom left quarter of the PC1 *vs* PC2 plot, as these objects have both negative PC1 and PC2 scores. Most of the spectra identified as epoxy resin (E) correspond to the objects grouped in the bottom right quarter resulting from positive PC1 combined with negative PC2 values. Objects corresponding to spectra characterized by proteinaceous materials (P) do not form a sharp cluster in the PC1 *vs* PC2 scores plot, however all these objects are characterized by positive PC2 scores and most of them are well separated from the other layers.

Wood (W) groups around the origin of PC1 and PC2 axes close to not identified layers (U, undefined), i.e. layers for which it was not possible to discriminate a predominant component among the identified macro-categories of binding media (varnish and proteinaceous material). Anyway, the U samples seems to be similar to each other in most cases as they cluster in a group with slightly positive PC2 scores and slightly negative PC1 scores, with just few exceptions. In addition, the third PC (accounting for the 11% of the variance) was investigated. Unfortunately, this component does not provide further relevant information.

From Fig. 2b interpretation, the groups previously identified are confirmed, even though some of them are more scattered or form sub-groups, which can be ascribed to the combination of different materials in the same layer, as expected.

From the loadings plot (Fig. 2c) the signals corresponding to the bands used to detect epoxy resin (●), varnishes (*) and proteins (▲) appear the ones mostly influencing the spectra distribution in groups according to the different materials constituting the layers. It should be kept in mind that the spectra were transformed by first derivative, thus the maximum of diagnostic peaks are lost, but they actually correspond to the inflection points of the loading profiles.As loadings can assume values from -1 to +1, variables approaching extreme values in Fig. 2c are the ones with higher influence in constituting the PCs and, thus, they are responsible for spectra distribution in the scores plots (Fig. 2a and b).

PC1 well discriminates varnish (V) from epoxy (E) spectra mainly thanks to the signal corresponding to ~1720 cm^{-1} and ~1510 cm^{-1}. PC2 allows the discrimination of layers rich in proteinaceous material (P) due to the signal related to vC=O of amide I (1665-1645 cm^{-1}).

PCA confirmed the material assignation determined by inspecting the layer position in the cross-sections by highlighting clear samples grouping for varnish (V), epoxy (E), wood (W) and proteinaceous material (P). Moreover, the explorative analysis confirmed that the spectra positioned in the interfaces between different layers actually show signals produced by different materials (varnish + proteins + epoxy resin).

C. *PLS-DA*

The materials identified by layers inspections and confirmed by PCA were used as classes (U, E, P, V and W), thus, constituting the *a-priori* information (Y) to build the PLS-DA classification model (Table 3), able to predict layers predominant material based on the spectral data collected (X).

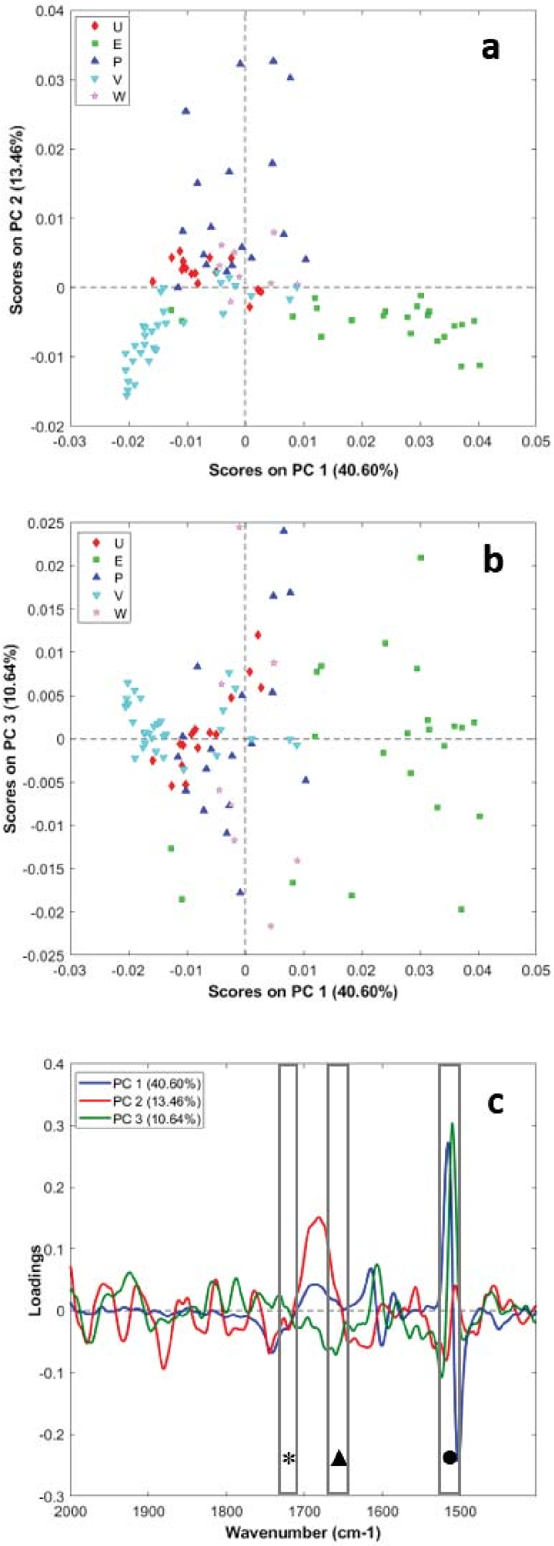

Fig. 2. PCA results: a) scores plot of PC1 vs PC2; b) scores plot of PC1 vs PC3; c) loadings plot of PC1, PC2 and PC3.

Table 3. Figure of merit of the PLS-DA model referred to calibration (Cal), cross-validation (CV) and prediction (Pred) steps. Sens=sensitivity; Spec=specificity.

	U	E	P	V	W
N_{cal}	18	17	12	20	4
N_{pred}	2	7	2	10	5
Sens (Cal)	0.94	1.00	0.83	0.90	1.00
Spec (Cal)	0.91	0.98	0.98	1.00	0.97
Sens (CV)	0.94	1.00	0.75	0.90	0.75
Spec (CV)	0.85	0.98	0.98	1.00	0.97
Sens (Pred)	**1.00**	**1.00**	0.50	**1.00**	0.60
Spec (Pred)	**1.00**	**0.78**	**1.00**	**0.91**	**1.00**

The PLS-DA model was firstly calibrated, i.e. a classification rule (equation) was established based on a representative set of samples. Then the model was internally validated by iterative exclusion of part of the calibration set, i.e. one out of the eight groups of samples (selected by venetian blind procedure) served as a internal test set, whereas the remaining data were used for calibration. The results of the eight tests were averaged and the constituent strategy achieving the highest accuracy was selected. The prediction ability of the optimized model was then tested by an external test set.

The three steps of model development (calibration, cross-validation and prediction) were evaluated for sensitivity and specificity. Sensitivity expresses the model capability to correctly recognize samples belonging to the considered class; whereas specificity describes the model capability to correctly reject samples belonging to all the other classes. The internal validation (cross-validation) of the model well performed for most of the considered classes reaching sensitivity above 0.90 and specificity higher than 0.85. However, the P class reached 0.75 of sensitivity as 4 out of 12 samples were misclassified as A (3) and W (1). The misclassification was expected as the *a-priori* assigned classes referred to the most present component in the layer, however it is unrealistic to assume each single layer, thus each spectrum, as a pure substance. The model prediction abilities resulted optimal for E and V classes with sensitivity of 1.00 and specificity above 0.78. Even though the specificity of P and W classes reached the maximum level (1.00), their sensitivity was poor. One out of two samples defined as protein was classified as undefined; whereas LS_v.25 and LS_v.26 defined as wood were assigned to epoxy class. It is clear that the low performance of the prediction phase is largely related here to the low number of spectra constituting some of the classes, mainly U and P. Indeed, the misclassification of only one spectrum resulted in 0.50 of specificity of class P. However, the prediction phase is the highest strength of the developed model as in most heritage classification cases this phase is missing for difficulties in collecting data from different samples.

IV. CONCLUSIONS

The constructed model reveals the feasibility of the proposed methodological approach aimed at discriminating the constituting materials of bowed string instruments in a fast and rigorous way skipping visual inspection of large number of spectra.

REFERENCES

[1] T. Rovetta, C. Invernizzi, G. Fiocco, M. Albano, M. Licchelli, M. Gulmini, G. Alf, D. Fabbri, A.G. Rombolà, M. Malagodi, "The case of Antonio Stradivari 1718 ex-San Lorenzo violin: History, restorations and conservation perspectives" J. Archaeolog. Sci. Rep., vol. 23, 2019, pp. 443–450.

[2] S.L. Lämmlein, D. Mannes, B. van Damme, I. Burgert, F.W.M. Schwarze, "Influence of varnishing on the vibro-mechanical properties of wood used for violins", J. Mater. Sci., vol. 54, No. 11, 2019, pp 8063–8095

[3] G. Fiocco, T. Rovetta, M. Gulmini, A. Piccirillo, C. Canevari, M. Licchelli, M. Malagodi, "Approaches for detecting madder lake in multi-layered coating systems of historical bowed string instruments" Coatings, vol. 8, No. 5, 2018, pp. 171-187.

[4] G. Fiocco, T. Rovetta, M. Gulmini, A. Piccirillo, M. Licchelli, M. Malagodi, "Spectroscopic analysis to characterize finishing treatments of ancient bowed string instruments", Appl. Spectrosc., vol. 71, No. 11, 2017, pp. 2477-2487.

[5] G. Fiocco, T. Rovetta, C. Invernizzi, M. Albano, M. Malagodi, M. Licchelli, A. Re, A. Lo Giudice, G.N. Lanzafame, F. Zanini, M. Iwanicka, P. Targowski, M. Gulmini, "A micro-tomographic insight into the coating systems of historical bowed string instruments", Coatings, vol. 9, No. 2, 2019, pp. 81-97.

[6] S. Grassi, C. Alamprese, "Advances in NIR spectroscopy applied to process analytical technology in food industries", Curr. Opin. Food Sci., vol. 22, 2018, pp. 17-21.

[7] G. Sciutto, P. Oliveri, S. Prati, E. Catelli, I. Bonacini, R. Mazzeo, "A multivariate methodological workflow for the analysis of ftir chemical mapping applied on historic paint stratigraphies" Int. J. Anal. Chem., 2017, pp. 1-12.

[8] G.V. Fichera, T. Rovetta, G. Fiocco, G. Alberti, C. Invernizzi, M. Licchelli, M. Malagodi, "Elemental

analysis as statistical preliminary study of historical musical instruments", Microchem. J., vol. 137, 2018, pp. 309-317.

[9] S. Lupi, A. Nucara, A. Perucchi, P. Calvani, M. Ortolani, L. Quaroni, M. Kiskinova, "Performance of SISSI, the infrared beamline of the ELETTRA storage ring", J. Opt. Soc. Am. B, vol.24, 4, 2007, pp. 959-964.

[10] P. Oliveri, M. Forina, "Data analysis and chemometrics", in Chemical Analysis of Food: Techniques and Applications, pp. 25-48.

[11] C. Invernizzi, A. Daveri, M. Vagnini, M. Malagodi, "Non-invasive identification of organic materials in historical stringed musical instruments by reflection infrared spectroscopy: A methodological approach", Anal. Bioanal. Chem., vol. 409, 2017, pp. 3281–3288.

[12] C. Invernizzi, T. Rovetta, M. Licchelli, M. Malagodi, "Mid and Near-Infrared Reflection Spectral Database of Natural Organic Materials in the Cultural Heritage Field", Int. J. Anal. Chem., 2018, pp. 1-16.

[13] R.E. Smith, F.N. Larsen, C.L. Long, "Epoxy resin cure. II. FTIR analysis", J. Appl. Polym. Sci., vol. 29, No. 12, 1984 pp. 3713-3726.

[14] B. Mohebby, "Application of ATR infrared spectroscopy in wood acetylation", J. Agric. Sci. Technol., vol. 10, No. 3, 2008, pp. 253-259.

[15] T. Poli, O. Chiantore, M. Nervo, A. Piccirillo, "Mid-IR fiber-optic reflectance spectroscopy for identifying the finish on wooden furniture", Anal. Bioanal. Chem., vol. 400, 2011, pp. 1161-1171.