Raman spectroscopy for identification of wood species

Journal of Physics: Conference Series **741** (2016) 012131 doi:10.1088/1742-6596/741/1/012131

Raman spectroscopy for identification of wood species

V A Gerasimov[1], A M Gurovich[2], D K Kostrin[1], L M Selivanov[1], V A Simon[1], A B Stuchenkov[2], A V Paltcev[2] and A A Uhov[1]

[1] Department of electronic instruments and devices, Saint Petersburg Electrotechnical University "LETI", 197376, Saint Petersburg, Russia

[2] Research and production center "Innovative engineering and technology", 196084, Saint Petersburg, Russia

Abstract. This article discusses the application of Raman spectroscopy for identification of wood species. Use of Raman spectroscopy allows increasing the certainty of determining the type of wood compared to the analysis of spectra of diffuse reflectance. Raman spectrums of different wood samples when irradiated by laser radiation are shown. Ways to improve the determination reliability of wood species due to the modernization of the identification technique are discussed. The stages of data processing, allowing carrying out correct further analysis are described.

1. Introduction

A number of studies in the field of identification of various natural and artificial objects by spectral methods were conducted in the past years. Instruments for spectral analysis have been successfully used in many areas of scientific research and technical control. However, there are a number of tasks that could be solved by spectral methods, but only in the field, because not every object of study can be delivered to the laboratory. One such area is the identification of wood species for customs control.

One of the first methods proposed for identification of wood species was based on diffuse reflection. The operating experience of this method showed that to improve the accuracy of identification the increase in the signal-to-noise ratio of the photodetector is required, because the resultant spectrums of different wood samples differ very slightly [1]. However, the use of more advanced sensor greatly increases the cost of research equipment. Thus, it is highly relevant to develop new methods to identify wood species in particular with the use of Raman spectrums.

2. Identification of wood species by spectrums of diffuse reflectance

The diffuse reflectance spectrophotometer for the identification of wood species contains a continuous spectrum source (halogen incandescent lamp), diffuse radiation of which falls onto the investigated sample using a photometric sphere. Reflected light through the slit illuminates the dispersing element (diffraction grating) that splits the incident radiation by wavelengths into the spectrum on the photodetector surface. To obtain the background (dark) signal in the scheme a controlled optical shutter is used [2]. To minimize measurement error of the spectral coefficient of diffuse reflectance algorithmic (numerical) compensation that is focused on the readings of temperature sensors located in different parts of the device is implemented [3]. Fragments of the diffuse reflectance spectrums of some wood species are shown in Figure 1.

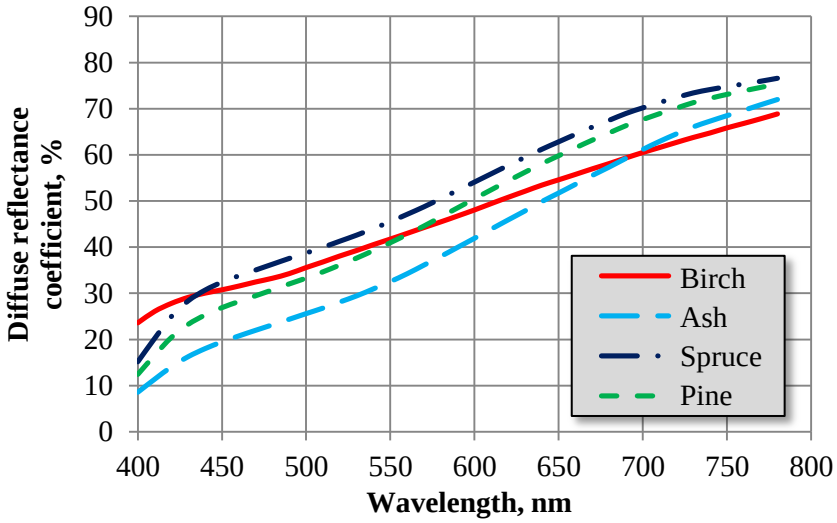

Figure 1. Fragments of the diffuse reflectance spectra of some woods species.

As can be seen from the figure, spectra of diffuse reflectance of wood are continuous. For different species it is impossible to distinguish the individual absorption bands; however the overall level of reflectance and steepness of the spectral functions make it possible to build identification algorithm [4].

The use of this method also imposes a number of restrictions on sample preparation. For confident identification preparation of the log cut ends by machining to achieve low roughness is required. Also for reliable operation of the algorithm the existence and regular updating of a library of spectral information is requested. These factors do not allow using devices based on method of diffuse reflectance effectively in the field.

3. Identification of wood species using Raman spectrums

When using Raman spectroscopy the investigated substance is irradiated with monochromatic radiation. A small portion of the transferred energy is converted into radiation, the characteristics of which are associated with the chemical composition of a substance and the parameters of the incident radiation; meanwhile most of the energy is absorbed or dissipated.

For the laser with a wavelength of 532 nm for the registration of Raman spectrums it is necessary to use a device with the spectral sensitivity region from 532 to 650 nm (this corresponds to Raman shift from 0 to 3400 cm^{-1} relative to the incident radiation). In this work AvaSpec-ULS2048 spectrometer was used. For the laser with a wavelength of 1064 or 785 nm it is possible to examine the spectrums in the direction of decreasing wavelengths from 1064 to 750 nm (corresponding to Raman shift from 0 to 3900 cm^{-1}). In our case Nicolet NXR Raman spectrometer was used. With increasing the wavenumber the spectral sensitivity of the photodetector also increases that enables the registration of the Raman spectrum over a wider range and this corresponds to the greater number of different substances which can be identified.

To separate the reflected radiation in the case of a laser with a wavelength of 532 nm it is possible to use filter glasses OS-13 or OS-14. In the case of the 1064 nm laser it is efficient to use the Semrock notch filter NF03-532/1064E-25.

Instead of the incandescent lamp (as in the devices based on diffuse reflectance method) a laser diode, which is included into the scheme of constant current switching converters, is used. The current value is programmed using the corresponding commands from PC. The current control circuit is required in cases where the laser radiation is able to heat the test samples to temperatures of combustion.

The need for the use of different lasers is related to the ability of substances to fluorescamine under the influence of powerful monochromatic radiation. If the substance is irradiated in the spectral region close to the absorption maximum, then fluorescence will be at its maximum and its value will be

several orders of magnitude greater than the magnitude of the Raman radiation. Figure 2 shows spectra of different wood species when irradiated with the 785 nm laser.

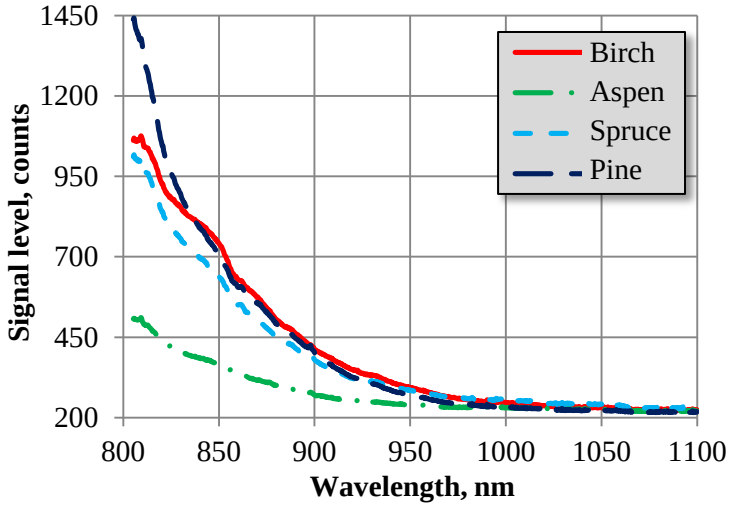

Figure 2. Spectrums of different wood species achieved using laser with a wavelength of 785 nm.

The figure shows that the spectra are also close to continuous, however small local extremes are present. If for each of the spectra the polynomial trend-line of low-order is found and then normalized by the maximum value difference between the spectrum and the trend-line is build [5], extrema can be considered (Figure 3).

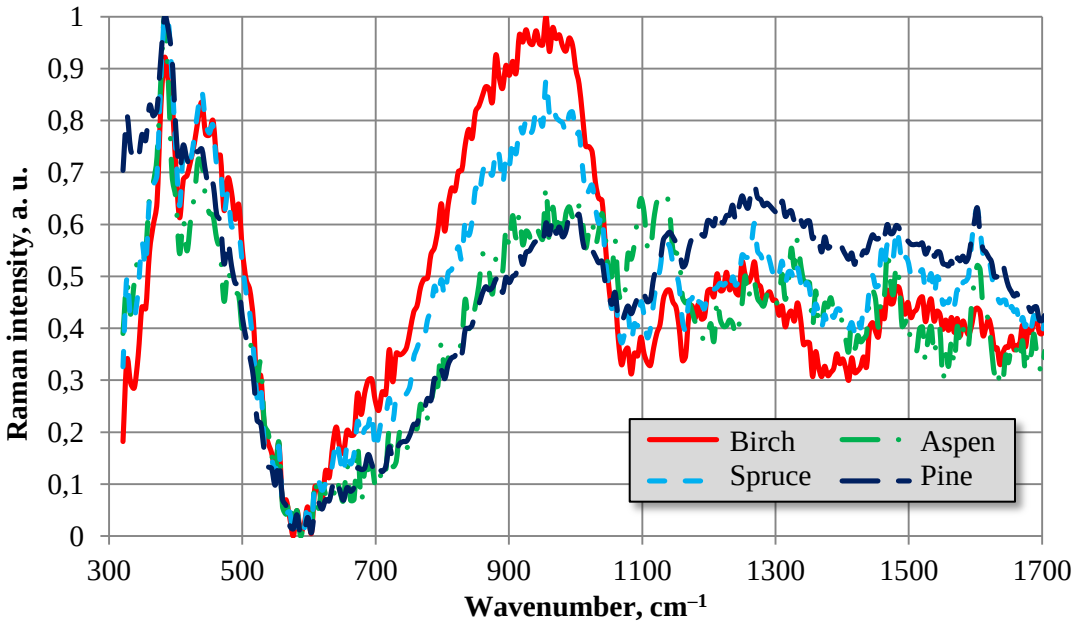

Figure 3. The difference between the Raman spectra and the trend-lines for different wood species.

The figure shows that the differences between the spectra of different wood species are more distinct than in the analysis of spectra of diffuse reflectance.

Spectrum obtained when irradiated by a 1064 nm laser is the most informative as it contains a large number of extrema, characteristic for certain substances. Figure 4 shows a comparison of Raman spectra using irradiation of the object with different lasers. Figure 5 shows spectrums of different wood species under laser irradiation with a wavelength of 1064 nm.

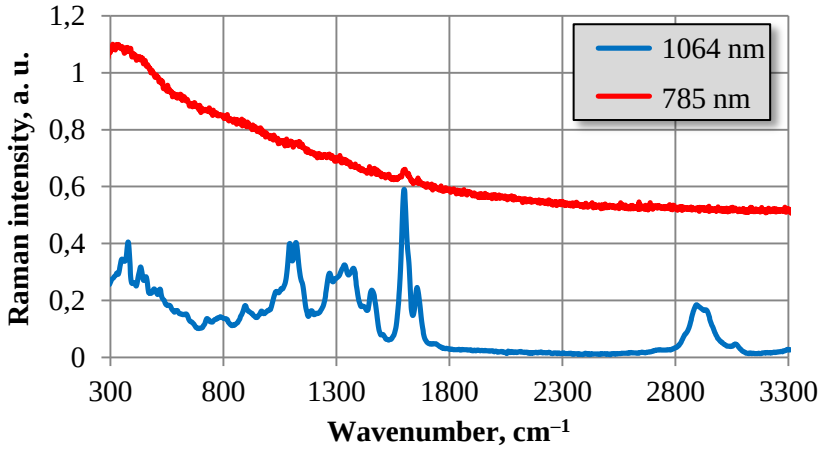

Figure 4. Raman spectra of pine under laser irradiation with different wavelength.

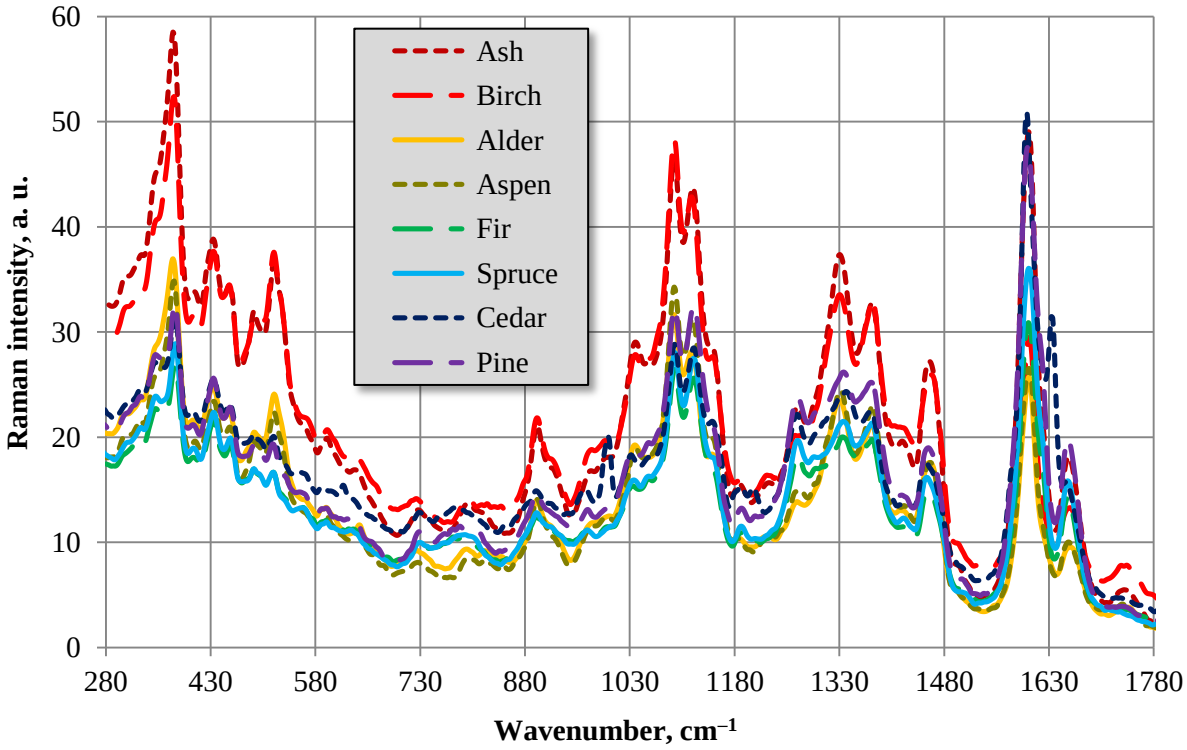

Figure 5. Raman spectrums of wood species achieved using laser with a wavelength of 1064 nm.

The figure shows that the structure of spectrums for different wood species is practically the same. However different concentrations of individual substances in these samples make identification of specific wood species possible.

4. Processing of the obtained spectral data

After receiving the spectrum it is necessary to process the data using the following three stages:

- smoothing;
- subtraction of the baseline;
- normalization.

The source data is always noisy. To reduce the noise effect one of the algorithms of smoothing or their combination should be used. At the moment, authors implemented four smoothing algorithms:

- moving average;
- exponential smoothing;
- median filter;
- Savitzky–Golay smoothing algorithm.

The last type of smoothing is particularly relevant for spectral data, because it carefully processes spectral peaks, not aligning them.

Obtained spectrums almost always have the effect of luminescence, which does not allow direct estimation of the Raman responses. Furthermore, luminescence is different in various samples (and sometimes in different regions of the same sample). This effect greatly complicates the analysis and requires a software solution. The problem is solved by finding the underlying background (base line) on the graph of the spectrum with its subsequent subtraction from the main spectrum. The analysis of the current works in this area has shown the perspectivity of the two approaches:

- modeling of the baseline by a polynomial of fixed order, with evaluation of the quality of approximation by minimizing the assigned cost function [6];
- modeling of the baseline with wavelet functions of different types. In practice two algorithms were implemented:
  - a simple algorithm based on the selection of the wavelet function parameters that use as input the Daubechies 10 wavelet or discrete Meyer wavelet [7];
  - more complex and high-quality algorithm, based on preliminary peak detection, using as basic a mexican hat wavelet [8].

After subtracting the baseline spectra need to be normalized, so that in future, while solving the identification problem, to estimate relative (and not absolute) influence of individual features (the intensity of the spectra at different wavelengths) in the common spectral measurements.

Three methods of normalization were used:

- normalization on the maximum;
- normalization by the area under the spectrum curve;
- normalization on the Euclidean norm (spectrum is considered as a normalized feature vector).

Formally, the problem of identification of the spectra is a problem of classification, that is, it requires preliminary training on pre-prepared sample spectra. The classification is based on a small set (3…5) of the spectra corresponding to multiple measurements of a single substance. An important condition for correct identification is matching of methods and parameters of pre-treatment for the identifiable spectra and spectra of the training samples.

The training sample should include an equal number of spectra of each class and obtained in the same conditions. The order in which the spectra are acquired should not depend on the class (the spectra of one class should not be measured consistently), so that the real differences between classes are not replaced by the possible variation of the measurement conditions.

The analysis is based on applying the principal component analysis (PCA) [9], which allows using the methods of linear algebra to identify in a large data set a relatively small number of parameters, which accurately characterize each of the elements of the set. Thus, reduction of the dimensionality of the processed data (and therefore processing time) can be achieved; also this solution allows to get rid of the slight "noise" effect present in the original data. Modern algorithms for calculating PCA method are based on obtaining singular value decomposition (SVD) of the matrix of original data, which in turn is associated with obtaining the eigenvalues. The quality of the method depends on the number of the selected components. The optimal number of principal components in our case ranges from 11 to 14 (instead of the original 512 points).

After applying the method of principal components and reducing the dimensionality of the original data the method of classification based on linear discriminant analysis (LDA) [10] is used. This method allows constructing linear combinations of available features, which allow maximizing differences between objects from different groups (classes) and minimizing difference between objects from one group. Mathematically the idea of the method is based on selection coefficients, which

maximize the dispersion of the average members of the classes and minimize the variance within classes (Fisher criterion). The adjustment of coefficients for LDA decomposition is performed on the matrix, rows of which are the measurements (spectra) and columns are the main components of each measurement obtained by PCA analysis. The overall results of the LDA analysis are:

- matrix of coefficients that projects the original (training) data into the space with a maximum ratio between the differences amongst classes to intraclass differences;
- set of similar linear functions with different coefficients (linear discriminant functions), each designed for a specific class and depending on the average value of the class and the covariance matrix of its elements.

A necessary step prior to the identification of a new spectrum is to perform preliminary processing identical to that carried out on the spectra of training samples. The task of classification is implemented by the following sequence of steps:

- after preprocessing the spectrum is transformed into the principal component space for which the vector corresponding to the spectrum is multiplied by the matrix of loadings in PCA decomposition of the training sample;
- resulting vector is projected into the principal component space of LDA decomposition for which the vector is multiplied by the projecting matrix, obtained during the LDA analysis;
- for a decision about belonging of the constructed LDA vector to a particular class this vector should be substituted into the linear discriminant function, corresponding to this class. The result of the lookup is a scalar value associated with the probability of belonging of the element to the given class (the larger the function value, the higher the probability that the analyzed spectrum relates to the corresponding class).

Calculating the values of the discriminant functions for all classes, the class for which the result of the function is maximal is selected. When performing classification of a set of spectra evaluation results for each spectrum are summarized. At present, discussed method and implemented system of processing and analysis of the spectra allows to obtain Raman spectra of wood and its identification with a probability of about 90 %.

5. Conclusion

Application of the principles of Raman spectroscopy allows increasing the reliability of the identification method compared to the spectral analysis of diffuse reflectance due to the registration of more informative spectral data. The conducted research gives opportunity to conclude that Raman spectroscopy is a promising method of identifying wood species.

References

[1] Woronin A A, Gerasimov W A, Kostrin D K, Smirnov P E, Turubarov A W, Oukhov A A 2013 *Biotechnosphere* **3** 16–20
[2] Uhov A A, Gerasimov V A, Kostrin D K, Selivanov L M 2014 *Journal of Physics: Conference Series* **567** 012039
[3] Voronin A A, Smirnova E V, Faizov I N 2011 *Journal of Instrument Engineering* **1** 74–78
[4] Voronin A A, Smirnova E V, Smirnov A P 2010 *Scientific and Technical Journal of Information Technologies, Mechanics and Optics* **2** 5–11
[5] Voronin A A, Smirnov P E, Turubarov A V, Gerasimov V A, Kostrin D K, Uhov A A 2014 *Testing. Diagnostics* **5** 35–38
[6] Mazet V, Carteret C, Brie D, Idier J, Humbert B 2005 *Chemometrics and intelligent laboratory systems* **2** 121–133
[7] Galloway C M, Le Ru E C, Etchegoin P G 2009 *Applied Spectroscopy* **12** 1370–1376
[8] Zhang Z-M, Chen S, Liang Y-Z, Liu Z-X, Zhang Q-M, Ding L-X, Yec Fei, Zhou H 2010 *Journal of Raman Spectroscopy* **6** 659–669
[9] Jolliffe I T 2002 Principal Component Analysis (New York: Springer-Verlag)
[10] Perriere G, Thioulouse J 2003 *Computer Methods and Programs in Biomedicine* **70** 99–105