Design of an Automatic Wood Types Classification System by Using Fluorescence Spectra

Vincenzo Piuri and Fabio Scotti

***Abstract*—The classification of wood types is needed in many industrial sectors, since it can provide relevant information concerning the features and characteristics of the final product (appearance, cost, mechanical properties, etc.). This analysis is typical in the furniture industries and the wood panel production. Usually, the analysis is performed by human experts, is not rapid, and has a nonuniform accuracy related mainly to the operator's experience and attention. This paper presents a methodology to effectively cope with the design of an automatic wood types classification system based on the analysis of the fluorescence spectra suitable for real-time applications. This paper presents an experimental set up based on a laser source, a spectrometer, and a processing system, and then, it discusses a set of techniques suitable to extract features from the spectra and how to exploit the extracted feature to train an inductive classification system capable to properly classify the wood types. Obtained experimental results show that the proposed approach can achieve a good accuracy in the classification and requires a limited computational power, hence allowing for the application in real-time industrial processes.**

***Index Terms*—Automatic spectra analysis, automatic wood classification, computational intelligence.**

I. Introduction

THE AUTOMATIC wood classification is a problem that is present in many industrial contests such as the furniture industries and the wood panel production [1]. Different woods have different aspects, properties, and costs. The correct classification of the wood type is very important to guarantee that the final product has the required features and characteristics. For example, in the production of wood panels, the wood type influences the quantity of the glue that must be used in the panel to guarantee the proper mechanical properties. On the other side, the glue has a great impact on the final cost of the panel and effects the overall environmental impact. In the paper industry, the wood type influences the final quantity of the cellulose in the paper, and hence, the quality of the paper [2].

Usually, the analysis of the wood type in the wood industries is performed by human experts by visual inspection, but this procedure is not rapid and presents a nonuniform accuracy due to the operator's capabilities and tiredness. More expensive chemical tests are available, but they are slow and can be done only on small samples of the production. It is known that the wood kind can be estimated from the wood emitted spectrum, but, unfortunately, the interpretation of the wood fluorescence spectrum is not a simple task that can be easily achieved by an operator in real-time industrial application. In fact, in the case of the wood-type identification, there are no unique set of peaks in the spectrum for each wood type and even small differences in the spectrum pattern are meaningful. As a results, the human identification of the wood type is not accurate and/or repeatable. In addition, more and more applications in the wood panel industry requires the usage of enormous quantities of recycled wood as a basic material, even in chopped and mixed slices (a recent plant can produce up to 3000 m^3/day of panels). In this applicative context, this kind of basic material is not suitable to be classified with a manned system, and the usage of a fast and accurate automatic identification system capable to control a continuous flow of samples coming from the feeding line of the plant is strongly required. To the best of our knowledge, no commercial systems are available in the market and no studies on automatic wood classification systems based on fluorescence spectra are present in the literature.

This paper describes the design of an accurate and continuously uniform method for the automated classification of the wood types by a contactless measurement and classification of the wood by visible and near-IR (NIR) spectra. Experiments has been made in order to test a classification capability up to 21 different wood types. Preliminary results has been presented in [3].

This paper is structured as follows. Section II describes the state of the art of automatic wood analysis and classification, while Section III focuses on the proposed design methodology and the creation of each module composing the system. The section describes how to extract a proper feature set, and create and train different models of inductive classification systems, as well as the creation of the training and testing datasets of wood spectra. The presented methodology has shown effective results with different classification models such as the k-nearest neighbor classifiers, linear, quadratic Bayesian classification systems, and supported vector machines (SVMs). Finally, it follows with the discussion of the accuracy and the performances of the overall system.

II. State of the Art of the Automatic Wood-Type Analysis

In the literature, the automatic analysis of timbers is achieved by following two main approach: image-based processing systems and spectrum-based processing systems. In the first category, the specific wood patterns, colors, and surface properties are analyzed by means of a specific illumination system and one or more cameras. In the latter category, a proper source of radiation is used to excite the wood surface in order to analyze the emitted spectrum.

In the literature, most systems based on *image processing* techniques deal with the identification of surface defects (for example, knots, resin drops, cracks, broken board edges, etc.) in order to classify the boards in different quality classes during the production [4]–[7]. A system for tracking single wood board during the production is presented in [8]. This system exploits a source of light and a color camera to acquire and store the specific wood pattern of the surface of each single wood board. This system processes the images in a very similar fashion to the systems used for the human fingerprints, since the input image patterns are normalized, warped, and then converted into a compress representation (called image packet) capable to identify the board in a large repository. a similar approach capable to identity a single board based on the peculiar signal pattern produced by the local microwave absorption is presented In [9]. The paper demonstrated that this signal is discriminant, since each board has different knots and densities along the three dimensions.

The identification of the wood species by image processing has been effectively addressed by the classification of the wood surface patterns. For example, the species of rain forest woods are recognized in [10] by using extracting textural wood features by using a co-occurrence matrix approach, then a trained neural network is used to achieve the final classification of the sample. The system had been tested with a dataset of 20 species obtaining a classification accuracy of 95% (tests had been carried out with only ten samples for each wood kind during the validation). A very similar approach had been tested in [11] on

Manuscript received July 5, 2009; revised October 19, 2009. First published February 5, 2010; current version published April 14, 2010. This work was supported by the Regional Programme for Industrial Research, Innovation and Technology Transfer (PRRIITT) project Emilia-Romagna nr. 703. This paper was recommended by Associate Editor M. Last.

The authors are with the Department of Information Technologies, Università degli Studi di Milano, Crema CR 26013, Italy (e-mail: fabio.scotti@unimi.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCC.2009.2039479

a dataset composed of six wood species, achieving a classification accuracy of 80%.

A second approach for wood-type analysis is based on the the *spectrum analysis*. In this paper, we present an innovative methodology for the automatic characterization of wood samples has been carried out by fluorescence spectroscopy. Such approach presents some advantages in comparison with those proposed in literature that are mainly based on vibrational spectroscopic methods: such NIR [15], mid-IR (MIR) [16], [17], and Fourier-transform Raman spectroscopies [18], [19]. In particular, the Raman spectrum had been processed in [18] with genetic algorithms in order to identify the relevant lines in the Raman wood spectrum, and differentiate softwoods, hardwoods, and tropical woods, while in [19], a simple neural network was used to classify temperate hardwoods from softwoods. Unfortunately, despite the richness of the providing information, the vibrational spectroscopy presents several drawbacks for the industrial application both for costs and experimental difficulties for its implementation in online and real-time measurement systems operating in industrial environment. For example, all these techniques require normally long integration times and expensive cooled detectors. In particular, IR measurements are affected by some environmental variables that are not easily controllable in the production line, such as the presence of thermal sources, dust, and humidity. As a matter of fact, the tail of the thermal radiation in the detectors sensitivity range produces a noisy background, the random presence of dust produces several artifacts in a long-time measurement, and some water vapor absorption bands overlap the wood spectral features useful for the recognition.

Differently from the techniques based on vibrational spectroscopy, fluorescence spectroscopy, on other hand, working in the visible spectral region has a higher SNR, which is unaffected by thermal noise or water absorption, and furthermore, the high sample rate, for example, allows to reject measurement on flying particles. Fluorescence also represents convenient choice due to the availability of lower cost components, such as the modern high-performance silicon-based charge-coupled device (CCD) detector and high-power diode-pumped solid-state (DPSS) laser. The use of the modulation capability DPSS laser (up to 100 KHz), together with a synchronous detection, allows a further improving of the SNR, fast measurements, and subtraction of the environmental light. All these features make fluorescence spectroscopy particularly suited for real-time measurement system operating in an industrial environment. In the next section, we propose a new methodology based on the fluorescence spectroscopy to effectively cope with the automatic wood-type classification.

III. Measurement System

The proposed method is sketched in Fig. 1. The acquisition system adopts a laser source in order produce fluorescence in the wood sample placed/transported on a flat surface (for example, a conveyor belt). The emitted radiation is then acquired by a spectrometer that produces in output the measured spectrum with a fixed frequency. The obtained spectrum is then preprocessed by the prefiltering module, and then, salient features are extracted. The last module perform the classification task of the extracted features and produce the final classification of the wood type. The method is general, since there are no particular constrains with respect to the wood type, the shape of the analyzed sample, and samples can be also in movement. In the following, the design of each module and parameters of the overall system are discussed.

A. Acquisition System

The prototype measurement system, we set up, consists of a miniature spectrometer and a DPSS laser operating at 473 nm, respectively,

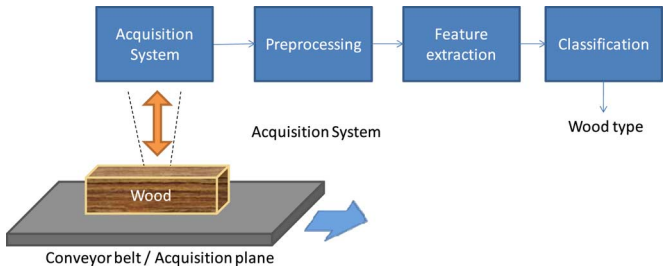

Fig. 1. General scheme of the proposed method.

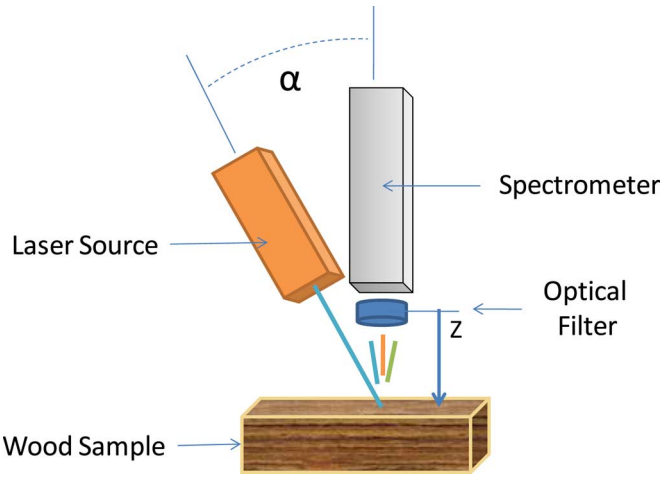

Fig. 2. Structure of the acquisition system.

for fluorescence detection and excitation. Fig. 2 shows in detail the composition of the acquisition system. The exciting laser beam, with modulable optical power up to 50 mW, impinges on samples at α angle with respect to the vertical direction corresponding to the axis of the collection optics. The spectrometer, provided with an adjustable objective lens focusing the collected light into the entrance slit, is orthogonally positioned z millimeters above the sample. A long-pass filter inserted between the sample and spectrometer removes the laser line from the collected light. The geometry of the setup (mainly the parameters α and z) must be tuned properly in order to achieve the maximum signal in input to the spectrometer. In our setup, the tuning has been achieved manually, but, of course, it is possible to adopt an autofocus system capable to vary the distance z of the lens accordingly to the position the wood sample if its position or thickness can change in time.

The spectrometer must be capable to acquire the fluorescence intensity in the correct band of wavelengths, which can be estimated as ranging from 500 up to 1000 nm. Also, the spectral resolution of the spectrometer plays an important role in classification accuracy: a resolution of 1 nm can be considered as largely sufficient to detect all salient patterns and transitions in the fluorescence spectrum. The spectrometer must also be chosen by taking into account the minimum integration time. In the laboratory setups, where the wood sample is not in movement, any particular integration time is not required. Conversely, in the industrial setups, the wood samples are often moving with respect the acquisition system (conveyer belts, chains, etc.). In this case, it is crucial to ensure that the integration time is sufficiently short in order to acquire the fluorescence spectrum of the same moving sample.

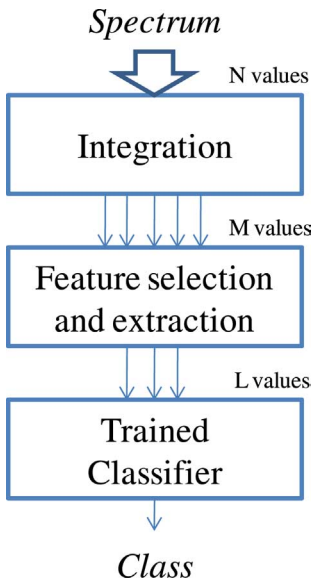

Fig. 3. Structure of the proposed classification system.

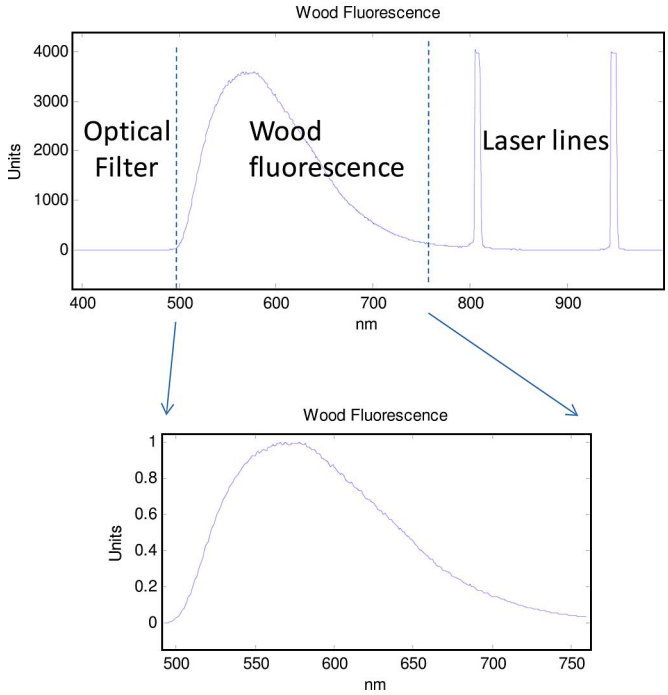

Fig. 4. Prefiltering of the input signal by cropping and normalization.

An optical long-pass filter has been inserted between the sample and spectrometer (see Fig. 2) in order to remove the laser line from the collected light and avoid saturation effects in the spectrometer. As such, only the fluorescence spectrum and two others lines (the subharmonic of the laser and the line produced by the pumping system) will be present in the collected radiation by the spectrometer allowing for a better exploitation of the dynamics of the device.

B. Preprocessing System

The exact relationship between the shape of the spectra and the wood types is not well known, hence it is not possible to directly design an algorithm for a classification system. On the contrary, the capability of the inductive classifiers to learn input–output relationships from examples can be exploited to create a proper classification system [20], [21]. The usage of an inductive classifier also implies the choice of a proper method to extract salient features from the input signal to be used as an input to the classifier.

In this paper, we propose a structure of the classification system partitioned in four main modules (see Fig. 3) achieving the following phases:

1) acquisition of the input spectra and preprocessing (N sample vectors);
2) integration of the spectrum in M contiguous bands;
3) feature selection/extraction of the L values;
4) classification of the wood using the L values.

In the following section, the design steps required to achieve the modules are given and compared with different techniques known in the literature.

C. Preprocessing Algorithms

Under the described assumptions, the typical output spectra obtained by the described acquisition system is composed by three main parts (see Fig. 4). On the bottom part of the wavelength axis, no information is present, since the optical filter cuts the principal laser line and all other contributions. In the second part, which starts from about 500 nm, the wood fluorescence is present and it represents the salient part of the signal. In the third part of the acquired spectrum, a laser subharmonic can be present, typically superimposed to the right tail of the wood spectrum (together with lines of the laser optical pumping system). The preprocessing module aims to extract and normalize the segment containing the wood fluorescence from the overall acquired spectrum.

Since the wavelength range of the wood fluorescence does not depend on experiment conditions, but is related only on the material, no particular algorithms have to be used to adaptively select the partition of the spectrum. In our experiments, we fixed the region of interest of the spectrum from $\lambda_1 = 490$ nm to $\lambda_2 = 750$ nm.

All parts of the spectrum signal can also be differently employed as feedback for an autofocus system, or can be used to manually tune the setup during the experiment, as described in the previous section, in order to maximize the spectral power of the fluorescence.

In any case, to correctly classify the wood sample, the intensity of the fluorescence spectra must be normalized in order to correct the effects related to the sample absorption, which can be very different from point to point. A first approach encompassed the usage of the intensity signal of laser subharmonic (at 946 nm) or of the pumping diode (at 808 nm) as reference for a normalization operation. As such, given the output of the spectrometer $I(\lambda)$ at different wavelength λ, the normalized spectral power $P_1(\lambda)$ is given by

$$P_1(\lambda) = \frac{I(\lambda)}{\max_{\lambda_2 \leq \lambda}(I(\lambda))}. \tag{1}$$

Another approach that can be used to normalize the input signal $I(\lambda)$ uses as reference the maximum value of the fluorescence range. In this case, the normalized spectral power $P(\lambda)$ is given by

$$P_2(\lambda) = \frac{I(\lambda)}{\max_{\lambda_1 \leq \lambda \leq \lambda_2}(I(\lambda))}. \tag{2}$$

Fig. 5 plots the results of the application of this prefiltering method applied to the input spectra of the following wood types: Wild Cherry, Oak Chestnut, Walnut, and Larch. The main assumption that we assume

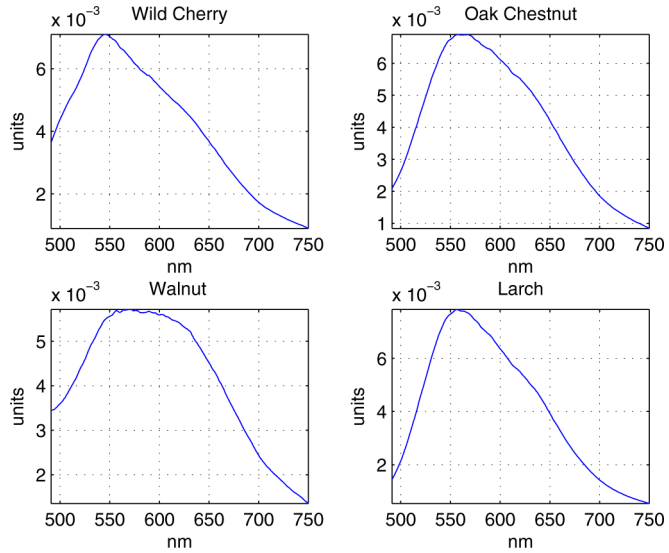

Fig. 5. Examples of prefiltered inputs spectra.

is that the emitted spectra of the different wood types are enough different to be classified with accuracy.

The obtained vectors $P(\lambda)$ can be directly used to extract features, or—differently—it can be applied as a further preprocessing method in order to better enlighten the differences between the patterns of the different woods. This method subtracts to the values of $P(\lambda)$ the mean pattern obtained by averaging K testing measures $\{P^1(\lambda), P^2(\lambda), \ldots, P^L(\lambda)\}$ taken with different wood types, hence producing the transformed vector $P'(\lambda)$ where

$$P'(\lambda) = P'(\lambda) - \frac{1}{K}\sum_{i=1}^{K} P^i(\lambda). \tag{3}$$

D. Feature Extraction and Selection

For any of the presented preprocessing methods, the acquisition system produces a vector of N samples for each wood acquisition. If the spectral resolution of the spectrometer is elevated, the cardinality of the input can be very high ($N = 242$ in our experiments). It cannot be considered as adequate to be directly used as input to the classifiers, and hence, a reduction of the number of the input features has to be considered (features selection/extraction) [21].

In supervised machine-learning problems, the task to identify m of the N inputs that are more relevant/significant for the final classification system is called *feature selection*. The most straightforward approach to the this problem requires to examining all $\binom{m}{N}$ possible subsets of size m with the largest value of a feature of merit $J()$ (for example, the accuracy of the classification system). Unfortunately, this simple approach is most of the times unfeasible due to the combinatorially grow of the number of subsets to be tested even with small values of m and N. A wide literature on this methods is available, for example, in [20], an extensive review is published for further reference. A particular approach suitable to reduce the input cardinality in the feature-selection framework encompasses the usage of the *wrappers* algorithms [12], [13]. This task is accomplished by means of a module called *inducer* that processes the whole set of N inputs on a training partition of the available dataset estimating the optimal (or suboptimal) subset of m inputs capable to maximize a defined figure of merit $J()$. In this paper, we will apply both classical greedy feature selection algorithm like sequential forward selection (SFS) sequential backward selection (SBS) and custom wrappers in order to identify which bands are more relevant for the wood classification problem. Further details will be given in the experimental section.

The reduction of the dimensionality of the input space can be achieved by using different methods [20] such as the *feature-extraction* framework. The most popular technique is the principal components analysis (PCA) that can compress most of the variation measured in the overall spectrum into a minor number of components [22]. Since the PCA-like mapping mixes the input components into a reduced set of new features, the direct relationship between the regions of the spectra and their importance in the wood classification is less explicit [23]. Similar approaches in the literature are based on the neural networks [19], and the genetic algorithms [18]. The linear prediction models were produced in [24] by using multivariate analysis and regression methods on a very specific application: the compression wood in Norway spruce (Picea abies). The spectra coming from a satellite spectrometer has been classified in [25] by using self-organizing maps.

This feature-extraction approach can be considered as functions $f()$ capable to map the N prefiltered input samples vector $\mathbf{P} = \{P(\lambda_1), P(\lambda_2), \ldots, P(\lambda_N)\}$ into a vector of L new features $\mathbf{F}$. For instance, in the case of the PCA mapping, the function $f()$ is given the following linear operation:

$$F = f(\mathbf{P}) = \mathbf{W}\mathbf{P} \tag{4}$$

where the matrix $\mathbf{W}$ is obtained by the PCA by considering only the first L singular vectors

$$\mathbf{W} = PCA(\mathbf{D}, L). \tag{5}$$

The NxK matrix $\mathbf{D}$ is a data matrix where columns are vectors of prefiltered input samples obtained in K experiments.

The approach that we propose aims to integrate the spectral energy into L fixed bands in order to produce a vector of L elements, which can be used as input to the classification system (*feature vector*). This approach very easily permits to test the functioning of the system with different spectral definitions of the spectrometer, and to directly identify which bands of the spectrum are more relevant into the classification problem. Under this assumption, given the N preprocessed normalized spectral power $\mathbf{P}' = \{P'(\lambda_1), P'(\lambda_2), \ldots, P'(\lambda_N)\}$ obtained by uniform sampling, the L extracted feature elements of vector $\mathbf{F}_1 = \{f_1, f_2, \ldots, f_L\}$ are given by

$$f_i = \frac{1}{\Delta}\sum_{j=\Delta(i-1)+1}^{\Delta i} P'(\lambda_j) \tag{6}$$

where λ_1 is the wavelength where the preprocessed input spectrum begins (see Fig. 4, bottom subplot) and $\Delta = N/L$ is the downsampling rate of the spectral power. Without loss of generality, L can be fixed in order to ensure that Δ is an integer. Differently, the downsampling of the spectral power $\mathbf{P}'$ to the feature vector $\mathbf{F}$ can be done by classical techniques. In the following, we refer to this feature-extraction method as *band integration*.

A completely different approach aims to extract the features form the $P(\lambda)$ by mapping the input vector $P(\lambda)$ as a polynomial $\hat{P}(\lambda)$ of degree L with $L > 1$, where

$$\hat{P}(\lambda) = \sum_{i=1}^{L} \alpha_i \lambda^i + \alpha_0 \tag{7}$$

and then exploiting the polynomial coefficients as a feature vector of L values $\mathbf{F}_2 = \{\alpha_0, \alpha_1, \ldots, \alpha_L\}$. In our experiments, the coefficients have been obtained by the least-squares method. This approach can very

effectively reduce the dimensionality of the input space, but, otherwise, it is not possible to further extract the significance of each extracted wavelength band since the information present in the input sample has been embedded in the L coefficients. Large values of L should be avoided due to the high possibility to have a bad-conditioned inversion matrix in the least-square procedure. In the following, we refer to this feature-extraction method as *polynomial coefficients*.

A large number of features extraction and selection methods are available in the literature, but their comprehensive description is beyond the scope of the presented paper. A very good review of such methods can be found in [20], where main available approaches are categorized and compared. In the experimental section, the three presented extraction methods are compared and discussed when adopted to identify the different wood types by using a supervised classifier.

E. Creation and Testing of the Classification System

In this section, the last four phases of the design will be described: the creation of the dataset, the creation of the classifiers, the training phase of the classifiers, and the accuracies estimation of the proposed classification systems.

As previously described, since a model for classifying the wood types from their spectrum patterns is not available, the usage of inductive classifiers has been introduced. Such classifiers can learn the classification rules from a proper dataset of examples previously classified from a supervisor by using a learning algorithm. The estimation of the accuracy of the inductive classifiers is mostly achieved by cross-validation techniques [27], and, as a consequence, it is required to divide the examples dataset in (almost) two partitions.

The first partition (the training dataset) is used to tune the system's parameters and train the inductive classifier. In some approaches, the training dataset is split in more parts, each used for one of these tasks: parameters tuning and classifiers learning will be performed by using separated subsets of data [28]. For example, a subset of the training dataset can be used to test the generalization capability of the classifier directly during the learning phase, as it is typical in neural networks applications [28]. The second partition (validation dataset) is used only once to estimate the system classification error. More accurate techniques for classification error estimation can be used (e.g., the N-fold validation and the leave-one-out), but their computational complexity becomes very high for large datasets [20], [28].

It is worth noting that, in order to guarantee the generality of the classification results, all the described operations that extract parameters from the available dataset (in this paper, the prefiltering method described in (3) and the PCA method) must be processed with a separate dataset or by considering only the training partition of the available dataset. Differently, any usage of the validation dataset to calibrate the system or to process parameters of the algorithms can lead to poor generality of the results, in particular, to an optimistic estimation [29] of the classification error. In the literature, a great number of inductive classification systems are available [20] with very different peculiarities. In this paper, we propose and compare classifiers belonging to four main families: the nearest neighbor classifiers, the linear and quadratic classifiers, and the SVMs.

The first model that we adopted is the linear Bayes normal classifier (LDA in the following), a method that builds a linear classifier between the classes of the dataset by assuming normal densities with equal covariance matrices in the input data [26]. Based on similar hypothesis, but, instead, using a second-order mapping of the input, we considered the quadratic Bayes normal classifier (QDC) [12], [16]. As third family of classifier systems, we adopted the well-known k-nearest neighbor classifier with odd values of the parameter k (1,3, and 5). These classifiers store in their memory the training samples and, each time that a unclassified sample is put in input, the corresponding class is estimated by selecting the k stored samples that are closer to the input regarding the selected metric (the Euclidean metric in our experiments). A voting process is then performed, and the unclassified sample is assigned to the class with the majority within the classes of the k selected samples.

A different approach can be considered by adopting SVMs [30]. SVMs are capable to separate a training-labeled dataset with a hyperplane that is maximally distant from the different classes present in the feature space. In case of dataset that are not linearly separable, kernel functions are adopted in order to realize a nonlinear mapping of the feature space capable to separate the classes. By this point of view, the SVM hyperplane expressed in the transformed feature space can be considered as a nonlinear boundary function in the input space. Without loss of generality, let us assume that the training dataset contains n features vectors $\mathbf{F}$ labeled with values $y_i \in \{-1, +1\}$, $k()$ is a kernel function capable to map the input features vectors F into a transformed feature space $T \in \Re^N$, and b is a constant real value, then the decision boundaries of the SVM can be represented as follows:

$$\mathrm{SVM}(\mathbf{F}) = \mathrm{sign}\left(\langle \mathbf{w}, k(\mathbf{F})\rangle - b\right) \tag{8}$$

where $\langle , \rangle$ is the inner product. The parameters of the SVM (the hyperplane $(\mathbf{w}, b)$) can be found by maximizing the following quantity called *margin*

$$\gamma = \min_i y_i \left(\langle \mathbf{w}, k(\mathbf{F})\rangle - b\right) \tag{9}$$

where the quantity $(\langle \mathbf{w}, k(\mathbf{F})\rangle - b)$ represents the distance between the feature points $\mathbf{F}$ and the decision boundary of the SVM. In (9), the product $y_i \left(\langle \mathbf{w}, k(\mathbf{F})\rangle - b\right)$ is positive for correctly classified sample and *vice versa*, and hence, the maximization operation allows for finding the coefficients that better separate the classes by searching the larger margin value.

SVMs have the characteristic to successfully deal with high dimensional feature spaces better than traditional learning paradigms [30]. This property can be of great help in our context, since the extracted feature elements (vector $\mathbf{F}$) have high dimensionality, especially by using the feature-extraction method in (6). On the contrary, the training procedures are very slow, and the choice of parameters and kernels of the SVMs are not trivial. In our experiments, we focused on SVMs based on Gaussian kernels, and considered the linear and the quadratic form.

The classification error of the presented classifiers has been estimated by using the cross-validation technique with R rotations [15]. In this case, the error estimation is produced by averaging the error obtained with R different training phases where only a portion of $1/R$ of the sample is used in validation and the remaining samples in the training phase. After each training–validation phase, the portion of $1/R$ sample used in the training phase is rotated in validation and new samples will be used in the training phase. After R rotations, the process is then stopped and the mean classification error is computed. The advantages of this method are that all samples have been used to test the systems in training and validation, and the obtained error gives a realistic estimation of the final performance of the classification system when it will work in similar conditions in the applicative context.

If the final operative conditions will be different, a new error-estimation procedure should be considered in order to control the variations in performances occurred. In this case, a new dataset of classified sample has to be collected and a new cross-validation session is required. It is worth noting that R different classification systems will be created for each of the R phases of the test, and hence, a criterion to select the final classifier to be embedded in the classification system is required. Unfortunately, the theory of inductive classifiers gives no

general method in order to select the "proper" classifier between the R that have been tested [29]. As a rule of thumb, the one with the lower classification error can be selected. Notably, when R is high (more then 10, for example), the size of the dataset is very large (hundreds of samples) and when the samples are randomly mixed in the dataset, the differences between the trained classifiers tend to be strongly reduced. In order to test the real advantage of using feature-extraction techniques in this applicative case, we considered the application of the PCA technique as preprocessing for the LDA, QDC, and kNN (k = 1, 2, and 3) classifiers. In the following, we refer to these systems as PCA+LDA, PCA+QDC, and PCA+1NN, respectively.

IV. Experimental Results

A. Setup of the Acquisition Module

The prototype measurement system that we have set up consists of a miniature spectrometer (Ocean Optics USB2000) and a frequency-doubled DPSS laser operating at 473 nm, respectively, for fluorescence detection and excitation. Fig. 2 shows in detail the composition of the acquisition system. The exciting laser beam, with modulable optical power up to 50 mW, impinges on samples at α angle with respect to the vertical direction corresponding to the axis of the collection optics. The spectrometer, provided with an adjustable objective lens focusing the collected light into the entrance slit, is orthogonally positioned z millimeters above the sample. A long-pass filter, with cutoff wavelength of 500 nm inserted between the sample and the spectrometer, removes the laser line from the collected light.

The fluorescence intensity is then measured in the spectral range between 500 and 1000 nm at 1 nm resolution and with 10 ms of integration time. The intensity of fluorescence spectra have been normalized using as reference signal the intensity of laser subharmonic (at 946 nm) or of the pumping diode (at 808 nm), in order to correct the effects due to sample absorption. Such reference signals, which will be employed as feedback for an autofocus system, have been manually maximized during the experiment. The procedure aims to empirically identify the better angle α and distance z capable to maximize the overall power of the signal fluorescence captured by the spectrometer according to the focal length of its objective lens. In our experiments, best results have been obtained with $\alpha = 45^{\circ}$ and $z = 12$ mm. Fig. 6 reports the final acquisition system used for the experiments.

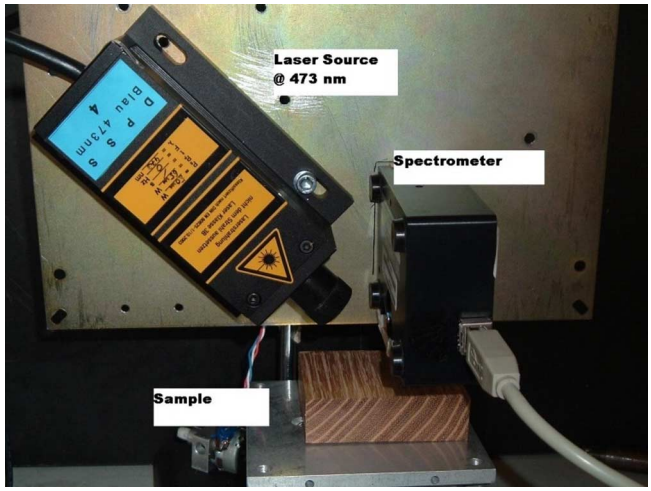

Fig. 6. Picture of the acquisition system. (Left) DPSS laser of 473 nm excites (bottom right) fluorescence of wood sample, while the miniature spectrometer above acquires the emission spectrum. The laser line is blocked by an optical long-pass filter placed on the spectrometer.

B. Datasets and Classification Systems

The datasets have been created by using a set of the 21 different wood types (of certified origin purchased at Woodtechnology Gmbh) belonging to the most common species. Twenty spectra for each sample have been acquired in different points by moving the samples under irradiation. During the measurements, we ensured to probe all wood zones, namely, heartwood, sapwood, and growth ring. The 21 wood types belonging to the dataset with the caption number assigned by the provider are the following: 1) Wild Cherry; 2) Oak Chestnut; 3) Walnut; 4) Larch; 5) Wild Pear Tree; 6) Poplar; 7) Cembar Pine; 8) Beech Tree; 9) *Alnus incana*; 10) Linden Tree; 11) *Fraxinus xanthoxyloides*; 12) Scots Pine; 13) Oak Tree; 14) Spruce; 15) Maple; 16) *Taxus baccata*; 17) Elm; 18) Silver Fir; 19) Birch Tree; 20) Black Locust; 21) *Carpinus betulus*.

The first classification problem that we considered-problem A—is the binary classification between the conifer and broad-leaved wood spectra. This problem is related to the fact that, in some specific applications such as the wood panel production, the properties of the wood types belonging to the same class (conifer or broad-leaved) can be considered as similar. The second classification problem that we considered—problem B—is the classification of the 21 different wood types. Problem B can be considered as more difficult than problem A, since the number of classes is ten time more with the same input data.

The classification error of the cited systems has been estimated using the cross-validation technique (using ten rotations). The test has been applied to all cited classifiers producing the mean classification error and its standard deviations. For the sake of comparison, in the results of tables and figures, when a classifier X has been trained by exploiting the principal component analysis, it is reported as a new different classifier with the label PCA+X.

As first step, we applied three different methods for feature selection directly on the row dataset produced from the sensor. The goal was to identify if some specific bands in the spectrum are more relevant to the classification problem. In particular, the dataset were processed with the SFS and SBS algorithms described in [20] (adopting as kernel the k-NN classifier with $k = 1$, 3, and 5 tested with the ten-fold cross-validation technique), and with the wrapper algorithm (W) specifically designed for classification in industrial application proposed in [14]. Results indicate that there is not a specific subset of bands that are particularly relevant for the classification task; in fact, many different sets of features produce classifiers with very similar final accuracies, and, in the case of dataset mixing, the selected subsets are different. Results indicate that the relevant information is probably distributed along the entire spectrum range and is not concentrated in specific bands or subbands.

As a second step, we considered the effect of the two different prefiltering methods presented in (1) and (2). Experiments showed that the latter normalization method tend to better preserve the peculiar patterns of the same wood sample during different acquisitions, and hence, it has been preferred, since it tends to reduce the intraclass separation of the samples without particularly worsening the interclass separation.

The application of the method expressed in (3) aims to better separate the spectrum patterns of different woods by subtracting the mean spectrum processed on a subset of wood samples. The application of this method allows for enlarging the interclass separation of the wood sample, although a small worsening of the intraclass neighborhood is added. Fig. 7 shows the application of the two method (also in different order) to five examples of wood spectra. As can be seen, the application of the prefiltering methods can greatly improve the separation of

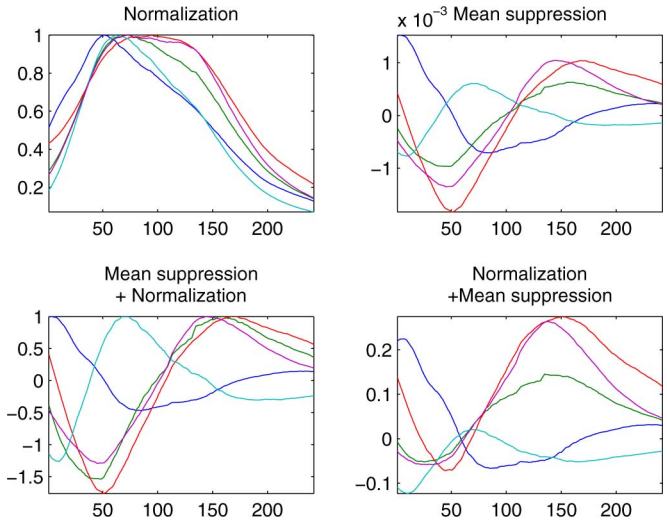

Fig. 7. Application of the techniques of normalization and suppression of the mean wood spectrum to five different wood spectra. The methods labeled as normalization and mean suppression corresponds to (2) and (3), respectively. The application of the prefiltering methods increases the interclass separation of the wood samples.

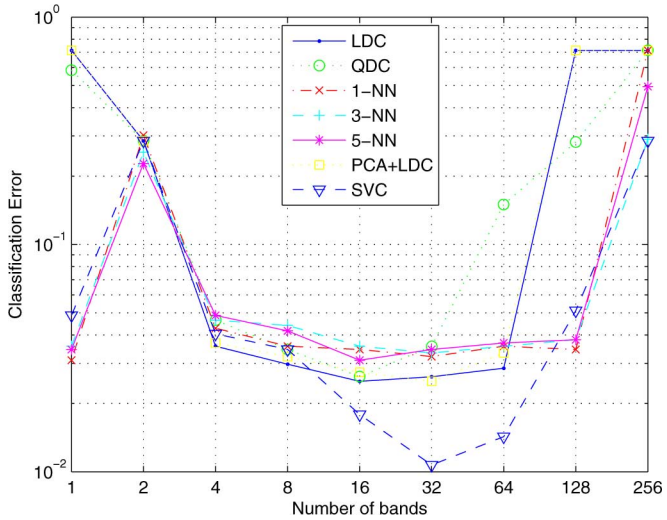

Fig. 8. Classification error of the proposed classifiers on the conifer/broad-leaved problem (dataset A) by using the prefiltering methods (2) and (3) with feature extraction (6).

interclass separations of the wood samples. Experiments showed that the overall effect obtained by combining the prefiltering (2) and (3) can be considered as positive, since the accuracy of all classifiers tested with and without these two prefiltering methods has been enhanced or, at least, did not worsen. For this reason, in the following, we will adopt this configuration of the prefiltering module. The feature selection analysis (SFS, BFS, and W methods) has been repeated again after the application of the prefiltering methods (1), (2), and (3), without any significative enhancement in the final accuracy of the classifier.

1) Feature Extraction by Band Integration: In order to understand the effect of the spectral resolution of the available power spectra, the values of spectral power have been integrated in M bands of the same size, as discussed in the previous section in (6). Figs. 8 and 9 plot the accuracy results of the tested classifiers with respect to the number bands M used as input to the classification systems. Concerning the dataset A (the conifer/croad-leaved problem), the best classifier has

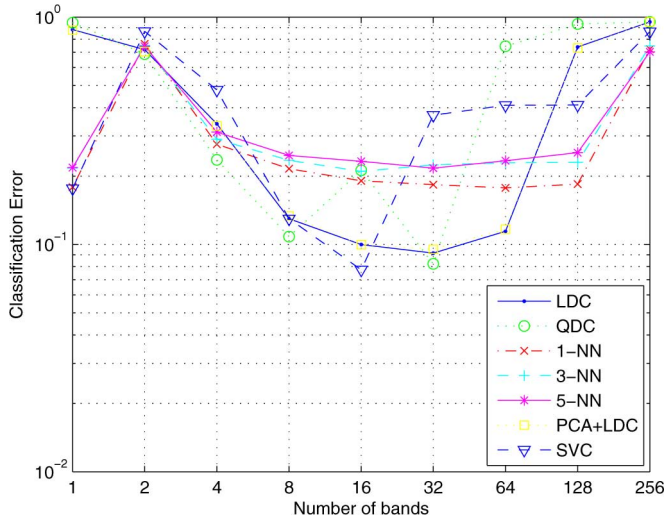

Fig. 9. Classification error of the proposed classifiers on the 21-types problem (dataset B) by using the prefiltering methods (2) and (3) with feature extraction (6).

been the SVM model with a 1.07% accuracy on dataset A with an error standard deviation of 0.2% when 32 spectral bands are used. Notably, most tested classifiers show comparable performance in the range of 2%–4% classification error for a spectrum bar of N sample, where N ranges from 5 to 25 for dataset A. Figs. 5 and 6 do not report the classification error of the PCA+LDA and PCA+1NN classifiers, since the application of the PCA does not significantly affect the classification errors with respect to the LDA and 1NN classifiers.

Results indicate that the classification between conifer and broad-leaved woods can be suitably achieved with different classification systems with the proposed method with a remarkable accuracy. The results of the classification of the 21 wood types show that the classification can be achieved with a classification error of 6.4% with a 0.9% standard deviation by the QDC algorithm when 16 spectral bands are used.

The errors related to the two classification problems are very promising, since they are obtained by using a single spectrum acquisition. A second method can also be considered: more than one spectrum acquisition can be taken from the same point (or considering points that are in a narrow neighborhood of the same wood sample). In this case, it is possible to achieve different operations of classification from the same points/area of the sample, and then, to process an average/voting operation on the class outputs. The averaging/voting method can probably further reduce the classification errors effectively.

2) Feature Extraction by Polynomial Interpolation: The results of the feature-extraction method based on the polynomial representation of the input spectra in (7) are plotted in Figs. 10 and 11, where the classification error of the tested classifiers is plotted against the number L of the interpolation coefficients. This compact spectral representation offers a good classification error in the conifer/broad-leaved problem (dataset A) arriving to the minimum value of 2.9% by using the QDC classification system with an error standard deviation of 0.2%. In particular, three classification system presented similar accuracy (QDC, PCA+LDC, and LDC classifiers). This best three classifiers produced an error that is comparable to the one achieved with the feature extraction by band integration. The approach of feature extraction by polynomial interpolation is less accurate in the 21-wood-types problem (dataset B) achieving a classification error of 8.9% with a error standard deviation of 0.2%. In this case, most classifiers produced an accuracy above 10%, probably because this

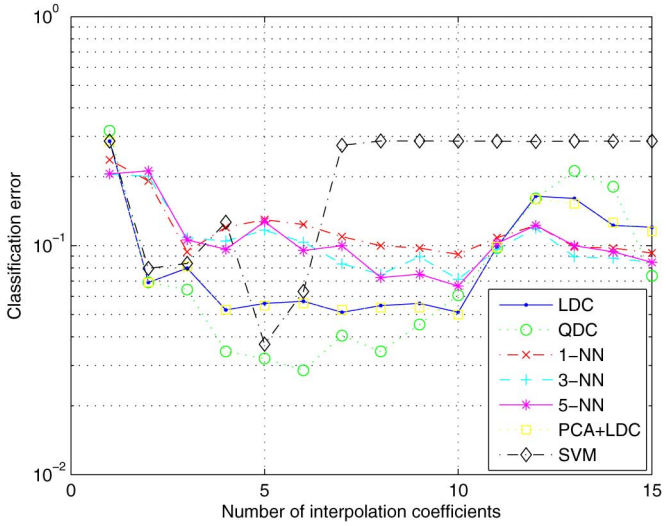

Fig. 10. Classification error of the proposed classifiers on the conifer/broad-leaved problem (dataset A) by using the prefiltering (2) and (3) with feature extraction (7).

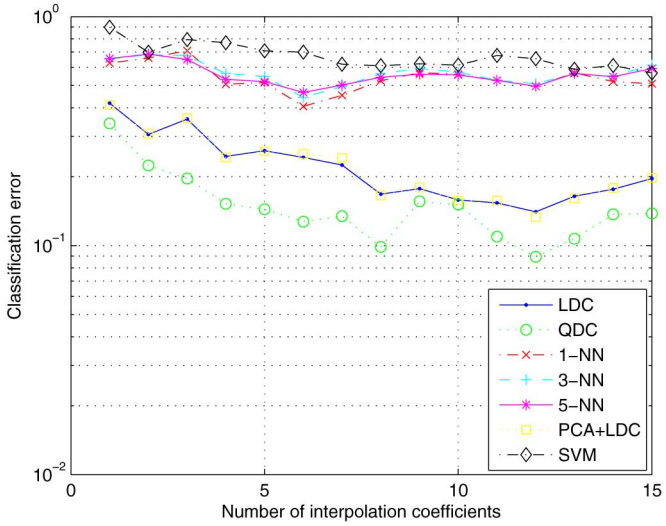

Fig. 11. Classification error of the proposed classifiers on the 21-types problem (dataset B) by using the prefiltering (2) and (3) with feature extraction (7).

feature-extraction method is less performing than the other two methods that we presented in the identification of the complex class boundaries that are present in dataset B.

Table I resumes the obtained results on the different datasets and experimental conditions. All the tested classifiers achieve the classification in a computational time, which ranges between 1 and 45 ms, depending on the number of inputs (the M bands) and the complexity of the algorithms. LDCs and 5NNs classifiers have the minimum computational times and the maximum computational times, respectively. All tests have been performed using a Pentium 1, 7-GHz, 1-GB RAM, using Windows XP Professional. The whole system has been implemented in MATLAB by exploiting the available toolboxes. The obtained computational times suggest that is possible to adopt the proposed classification method in real-time applications.

TABLE I
MAIN EXPERIMENTAL RESULTS

Dataset	Feature Extraction	Best Classifier	Classif. Error	STD
A	Band Int.	SVM	1.1%	0.2%
A	Poly. Coeff.	SVM	2.9%	0.2%
B	Band Int.	QDC	6.4%	0.9%
B	Poly. Coeff.	QDC	8.9%	0.2%

Notes. Feature-extraction methods: Band integration (Band Int.) and Polynomial Coefficients (Poly. Coeff.); Dataset A (Conifer/Broad-leaved problem); Dataset B) 21-types wood problem).

V. CONCLUSION

This paper presented a method for the automated classification of wood types based on the analysis of fluorescence spectra. The proposed method partitions the input spectra in different bands equally spaced. The energy contained in each band is used in input to an inductive classifier. Results show a good classification accuracy up to 21 different wood types. The presented approach has a general validity, and it can be used with spectrometers of different resolutions and with different classification systems, encompassing k-nearest neighbor classifiers, and linear and quadratic Bayesian classification systems. The simple experimental set setup and the limited overall computational complexity permit the adoption of the proposed method in real-time applications.

ACKNOWLEDGMENT

The authors would like to thank IMAL s.r.l., San Damaso (Modena), Italy, and Dr. P. Camorani (University of Parma, Italy) for their profitable cooperation in the data provided and the careful classification of samples.

REFERENCES

[1] S. Radovan, P. George, M. Panagiotis, G. Manos, A. Robert, and D. Igor, "An approach for automated inspection of wood boards," in *Proc. Int. Conf. Image Process.*, 2001, pp. 798–801.
[2] H. Furumoto, "Method and device for process control in cellulose and paper manufacture," U.S. Patent 6 398 914, Apr. 6, 2002..
[3] P. Camorani, M. Badiali, D. Francomacaro, M. Gamassi, V. Piuri, F. Scotti, and M. Zanasi, "A classification method for wood types using fluorescence spectra," presented at the IMTC 2008, Victoria, BC, Canada, May 12–15.
[4] S. Radovan, P. George, M. Panagiotis, and G. Manos, "An approach for automated inspection of wood boards," in *Proc. Int. Conf. Image Process.*, 2001, vol. 1, pp. 798–801.
[5] P. Cavalin, L. S. Oliveira, A. L. Koerich, and A. S. Britto, Jr., "Wood defect detection using grayscale images and an optimized feature set," in *Proc. IEEE Ind. Electron. (IECON 2006)*, pp. 3408–3412.
[6] Z. Zhang, N. Ye, D. Wu, and Q. Ye, "Locating the wood defects with typical features and SVM," presented at the Int. Conf. Comput. Vis., Pattern Recognit. Image Process., Shenzhen, China, 2008.
[7] G. A. Ruz, P. A. Estevez, and P. A. Ramirez, "Automated visual inspection system for wood defect classification using computational intelligence techniques," in *Int. J. Syst. Sci.*, Bristol, PA, vol. 40, pp. 163–172, 2009.
[8] G. M. Carman, P. S. Freeman, O. Heyman, and W. J. Briskey, "Wood tracking by identification of surface characteristics," U.S. Patent 7 426 422, Sep. 16, 2008..
[9] C. Fuentealba, C. Simon, D. Choffel, P. Charpentier, and D. Masson, "The k-nearest neighbor method for automatic identification of wood products," presented at the Int. Conf. Electron., Commun., Comput., Los Alamitos, CA, 2004.
[10] M. Khalid, E. L. Y. Lee, R. Yusof, and M. Nadaraj, "Design of an intelligent wood species recognition system," *Int. J. Simul. Syst., Sci. Technol.*, vol. 9, no. 3, pp. 9–19, 2008.
[11] J. Y. Tou, Y. H. Tay, and P. Y. Lau, "Rotational invariant wood species recognition through wood species verification," in *Proc. Asian Conf. Intell. Inf. Database Syst.*, pp. 115–120, 2009.
[12] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1/2, pp. 273–324, Dec. 1997.

[13] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2010.
[14] C. Alippi, P. Braione, V. Piuri, and F. Scotti, "A methodological approach to multisensor classification for innovative laser material processing unit," in *Proc. IEEE Instrum. Meas. Technol. Conf. (IMTC 2001)*, Budapest, Hungary, vol. 3, pp. 1762–1767.
[15] S. Tsuchikawa, Y. Hirashima, Y. Sasaki *et al.*, "A near-infrared spectroscopic study of the physical and mechanical properties of wood with meso- and micro-scale anatomical observation," *Appl. Spectrosc.*, vol. 59, pp. 86–93, 2005.
[16] M. H. Nuopponen, G. M. Birch, R. J. Sykes, S. J. Lee, and D. J. Stewart, "Estimation of wood density and chemical composition by means of diffuse reflectance mid-infrared Fourier transform (DRIFT-MIR) spectroscopy," *Agric. Food Chem.*, vol. 54, pp. 34–40, 2006.
[17] C. R. Orton, D. Y. Parkinson, P. D. Evans, and N. L. Owen, "Fourier transform infrared studies of heterogeneity, photodegradation, and lignin/hemicellulose ratios within hardwoods and softwood," *Appl. Spectrosc.*, vol. 55, pp. 960–966, 2001.
[18] B. K. Lavine, C. E. Davidson, A. J. Moores, and P. R. Griffiths, "Raman spectroscopy and genetic algorithms for the classification of wood types," *Appl. Spectrosc.*, vol. 55, pp. 960–966, 2001.
[19] H. Yang, I. R. Lewis, and P. R. Griffiths, "Raman spectrometry and neural networks for the classification of wood types," *Spectrochim. Acta A, Mol. Biomol. Spectrosc.*, vol. 55, no. 14, pp. 2783–2791, Dec. 1999.
[20] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.
[21] A. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artif. Intell.*, vol. 97, no. 1/2, pp. 245–271, 1997.
[22] L. R. Schimleck, A. J. Michell, and P. Vinden, "Eucalypt wood classification by NIR spectroscopy and principal components analysis,," *Appita J.*, vol. 49, no. 5, pp. 319–324, 1996.
[23] A. Webb, *Statistical Pattern Recognition*. New York: Wiley, 2002.
[24] J. Nyström and O. Hagman, "Real-time spectral classification of compression wood inPicea abies," *J. Wood Sci.*, vol. 45, no. 1, pp. 30–37, Jan. 1999.
[25] T. L. Roush and R. Hogan, "Automated classification of visible and near-infrared spectra using self-organizing maps," in *Proc. IEEE Aerosp. Conf.*, 2007, pp. 1–10.
[26] L. Rueda and B. J. Oommen, "On optimal pairwise linear classifiers for normal distributions: The two-dimensional case," *Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 2, pp. 274–280, 2002.
[27] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley, 2001.
[28] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York: Academic, 1990.
[29] C. Alippi and P. Braione, "Classification methods and inductive learning rules: What we may learn from theory," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 36, no. 5, pp. 649–655, Sep. 2006.
[30] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.