Support Vector Machines in hyperspectral imaging spectroscopy with application to material identification

P. Beatriz Garcia-Allende*, Francisco Anabitarte, Olga M. Conde, Jesus Mirapeix, Francisco J. Madruga, Jose M. Lopez-Higuera

Photonics Engineering Group, Universidad de Cantabria, Avda. Los Castros s/n, 39005 Santander, Spain

ABSTRACT

A processing methodology based on Support Vector Machines is presented in this paper for the classification of hyperspectral spectroscopic images. The accurate classification of the images is used to perform on-line material identification in industrial environments. Each hyperspectral image consists of the diffuse reflectance of the material under study along all the points of a line of vision. These images are measured through the employment of two imaging spectrographs operating at Vis-NIR, from 400 to 1000 nm, and NIR, from 1000 to 2400 nm, ranges of the spectrum, respectively. The aim of this work is to demonstrate the robustness of Support Vector Machines to recognise certain spectral features of the target. Furthermore, research has been made to find the adequate SVM configuration for this hyperspectral application. In this way, anomaly detection and material identification can be efficiently performed. A classifier with a combination of a Gaussian Kernel and a non linear Principal Component Analysis, namely k-PCA is concluded as the best option in this particular case. Finally, experimental tests have been carried out with materials typical of the tobacco industry (tobacco leaves mixed with unwanted spurious materials, such as leathers, plastics, etc.) to demonstrate the suitability of the proposed technique.

Keywords: Support Vector Machines (SVM), Principal Component Analysis (PCA), Imaging spectroscopy, anomaly detection, material identification.

1. INTRODUCTION

Imaging spectroscopy allows the simultaneous determination of the optical spectrum component and the spatial location of an object in a surface. To simultaneously acquire the information of both domains, spectral and spatial, a passive prism-grating-prism (PGP) device can be used [1]. Imaging spectroscopy based on this kind of hyperspectral device has been widely used in airborne applications. However, its suitability in other fields such as in the processing of tobacco [2] and food [3], or in biological applications [4] has also been reported. This suitability for monitoring industrial scenarios is subject to real-time operation and accuracy constraints. In this paper, Support Vector Machines (SVM) are investigated as the spectral interpretation technique of the hyperspectral images with application to material identification.

A non-destructive system, based on the mentioned optical technique, was previously developed for the detection of non desirable material in an industrial raw material chain [2]. It incorporated an imaging spectrograph able to register the hyperspectral images consisting of the Vis-NIR diffuse reflected spectrum of the material along all the points of an image line. The working material were different kinds of tobacco leaves mixed with typical spurious elements of this field such as plastics, cords, cardboards, papers and leaves of other vegetable material different from tobacco. Image classification, which determines the presence of target material or spurious element for each spatial pixel of the image line, was entirely performed using Artificial neural networks (ANN) [5,6].The learning function utilised was back-propagation with adaptive learning rate. In this paper, Support Vector Machines (SVM) [7] are investigated as an alternative method. The reason behind this investigation is that SVMs have been proved to provide better performance than ANNs in their typical application scenarios as handwritten character classification [8], intrusion detection [9] or image retrieval [10]. In addition, there have been many investigations into SVM applied to spectral data [11-14], some specifically regarding to image classification in remote sensing [12], medicine [13] and the food industry [14], where they have performed very favorably. The aim of this work is to find the more adequate SVM configuration in the

particular scenario of [2] and to make a comparison between ANN and SVM in this application of raw material quality control.

The paper is outlined as follows. Section 2 presents comprehensively the employed materials and methods. Therefore, imaging systems and data acquisition are firstly described. Afterwards, the main principles of SVM are summarized. In Section 3, discrimination capacities of SVM obtained in the two considered spectral ranges are presented and discussed. Finally, several conclusions are extracted in Section 4.

2. MATERIAL AND METHODS

2.1 Hyperspectral reflectance imaging of raw material

A collection of raw material samples belonging to two different classes were considered. The first class is composed by tobacco leaves. And the second class is composed by undesired materials that typically appear in tobacco industry production plants due to the manual harvest procedure of the leaves. This class includes wood, cardboard, different colored cellophanes, leather, foil, paper of sweets, textile threads and brown and green leaves of other vegetable material different from tobacco. From now on the two classes are designed as "*target*" and "*non-target*", respectively.

The image data of the samples was acquired through two different hypespectral imaging systems, working respectively in the Vis-NIR (from 400 to 1000 nm) and NIR (from 1000 to 2400 nm) ranges of the spectrum. Both systems have the same block diagram (depicted in Figure 1) and they are divided into three modules: the sensor module, optics and lighting module. Illumination consists, in both cases, of two halogen floodlights Tasley MX500, with a power rating of 500W each. The optics module includes a spectrograph with a PGP construction and a C-mount lens. The version V10E of the equipment known as Imspector, commercialized by Specim®, is employed as imaging spectrograph in the Vis-NIR system while its version N24E is used in the NIR range. Selected C-mount lens are respectively the Navitar objective lens Zoom 7000 and the Pentax C1614-M (C31634). Finally, the sensor module includes in the Vis-NIR range a high-performance monochrome digital camera, the Pixelink PL-A741 (1280 x 1204 pixels) whose spectral range goes from 350 up to 1000 nm. In the NIR system the monochrome camera XEVA-FPA-2.5-320 (320 x 256 pixels) by Xenics, whose spectral range is from 850 to 2500 nm, has been utilised. Both systems are controlled by a desktop PC(Pentium IV with a 3 GHz processor and RAM-512 MB). Image processing tasks are performed with Matlab® 7.4.0.287 (R2007a) [15] in a Pentium® D CPU 3GHz and RAM-2GHz. Communication interface between computer and cameras is IEEE1394 for the Pixelink PL-A741 and Camera Link for image transference and USB for camera control in the case of the XEVA-FPA-2.5-230.
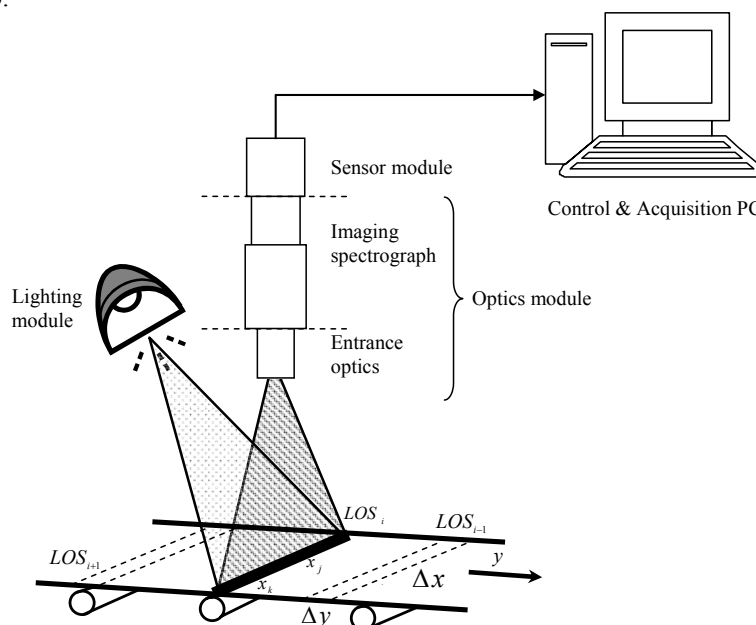

Fig. 1. Block-diagram of the Vis-NIR and NIR hyperspectral systems.

Before data acquisition, several calibration processes (spatial, spectral and responsivity calibration) of the systems are accomplished. Spectral or wavelength calibration requires the employment of different lightsources with known emission wavelength. A monochromatic He-Ne laser of 670 nm and a Hg-Ar lamp with multiple emission lines are employed in the Vis-NIR range, while in the NIR range this procedure is accomplished by using three lasers with wavelengths of 1310 nm, 1550 nm and 1680 nm, respectively. System field of view depends on the length and width of the length and width of slit of each spectrograph, the distance between target and objective lens and the lens focal length. The latter is adjusted to have the sample on focus. Therefore, the length and width of the scene line have to be experimentally determined. A fringe pattern is used for this purpose. Offset compensation, which consists in the subtraction of the dark current, is performed simultaneously with responsivity calibration. The latter takes into account detector responsivity, lens and spectrograph throughput dependency on wavelength, light source color temperature drift and lighting spatial non-uniformity. These variations are calibrated by measuring white and black references. For white reference Spectralon® reflectance standard is used and the black reference is obtained by blocking the sensor input. After that, each time the system gets an image from a sample, its reflectance should be corrected as indicated in (1).

$$R(\lambda_i) = \frac{S(\lambda_i) - D(\lambda_i)}{W(\lambda_i) - D(\lambda_i)} \tag{1}$$

where $S(\lambda_i)$ is the sample value at a wavelength λ_i; $D(\lambda_i)$ is the value for the dark reference and $W(\lambda_i)$ the equivalent for the white reference.

Table 1 summarizes the main specifications of the measurement systems, once the abovementioned calibration procedures have been performed.

Table 1. Summary of the main system specifications

	Vis-NIR (Imspector V10E)	NIR (Imspector N24E)
Spectral range (nm)	400 - 1000	1000 – 2400
Spectral resolution (nm)	2.8	8
Length of the scene line (mm)	20	143
Spatial resolution (µm)	208.833	446.875

Figure 2 shows how the acquired hyperspectral images look like. The vertical axes, along the 1024 pixels of the Pixelink camera in the Vis-NIR range (Figure 2 a)) and along the 256 pixels of the Xeva camera in the NIR range (Figure 2 b)) cover the spectral ranges dispersed by the corresponding PGP devices. The horizontal axes, along the 1280 or 256 pixels of each camera, image the line of vision of the lens. This line of vision is typically located over the conveyor belt which transports the mixed raw materials in an industrial production plant. Bright spots on the image stand for more diffused reflected radiation from the material under analysis. Therefore, each vertical line of the image contains the *spectral signature* of the material located in that specific spatial position.
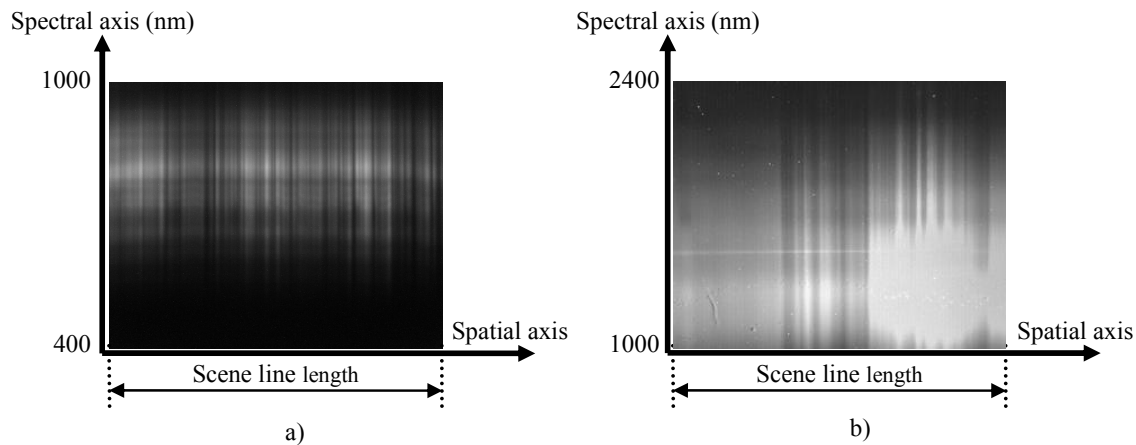

Fig. 2. Hyperspectral images captured by the Vis-NIR (a) and NIR (b) systems.

2.2 Cross-validation

In order to evaluate the performance of the classifier a *3*-fold cross-validation technique [16,17] is applied. A set of 1440 spectra is obtained from the images and it is divided into 3 non-overlapping sets having, therefore, 480 spectra each (240 spectra of target material and 240 of non-target material). Two of these set are employed as the training set to estimate the model parameters in the classifier, and the other one, the so-called validation set, is used to calculate the error of the classifier. This procedure is repeated three times, each time with different training and validation sets. Finally, the estimated performance of the classifier is calculated by averaging the three obtained errors.

2.3 Support Vector Machines for raw material spectral signature classification

Support vector machine is a learning system based on the statistical learning theory [7]. This classification technique aims at finding a separating hyperplane that splits the input data space into two separate regions corresponding to the two classes defined in the discrimination problem. These are target and non-target in the particular case of raw material quality control. There are many possible linear classifiers, but there is only one that maximizes the *margin* or the distance between the classification boundary and the nearest data point of each class. Large separation margin maximizes the generalization ability, i.e. improves classification accuracies for unseen data. The classification problem is defined as follows: x_j is a column vector representing each one of the spectra in the training set, therefore $j = 1,2,...,960$, and $y_j = \{+1,-1\}$ denote class label for x_j, target or non-target respectively. Training data set is said to be *linearly separable* by the separating hyperplane $w^t x + b = 0$ if there exit a vector w and a scalar b such that [13]:

$$w^t x_j + b \geq 1 \quad if \quad y_j = +1 \tag{2}$$

$$w^t x_j + b < -1 \quad if \quad y_j = -1 \tag{3}$$

These can be combined into one set of inequalities:

$$\left[y_j \left(w^t x_j + b \right) - 1 \right] \geq 0 \tag{4}$$

The separating margin between the two classes is, therefore, $2/\|w\|$. This margin is maximized if the norm $\|w\|$ is minimized with the limitation given in (4). To represent this optimization problem, the Lagrange function defined in (5) is used:

$$L(w,b,\alpha) = \frac{1}{2} \|w\|^2 - \sum_{j=1}^{l} \alpha_j \left[y_j \left(w^t x_j + b \right) - 1 \right] \tag{5}$$

where α_j is called a Lagrange multiplier and $l = 960$. The so-called *support vectors* are those spectra of the training set whose associated Lagrange multipliers are non-null; they are consequently the relevant ones in the determination of the separating hyperplane. Once the criterion function has been solved, the discriminate function to classify a new pattern x can be represented by a small subset of support vectors x_j:

$$f(x) = \text{sgn} \left(\sum_j y_j \alpha_j \langle x \cdot x_j \rangle + b \right) \tag{6}$$

As stated in (5), the decision function depends on the inner product between the patterns. For not linearly separable cases, a nonlinear transform function $\Phi(\cdot)$ is used to map the input vectors to higher dimensional feature space, which is more likely linearly separable. Nonlinear decision boundaries in the input space will be mapped to linear decision boundaries in the feature space. The classification function can be represented by the kernel function $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ in the transformed feature space. The nonlinear transform function $\Phi(x_i)$ needs not be specified. Using the kernel function, the discriminant function for a nonlinearly separable problem can be similarly written as:

$$f(x) = \text{sgn} \left(\sum_j y_j \alpha_j K(x, x_j) + b \right) \tag{7}$$

In this application to material identification a Gaussian radial-basis function (g-RBF) kernel is used:

$$K(x, x_j) = \exp\left(-\|x - x_j\|^2 / 2\sigma^2\right) \tag{8}$$

where σ is the parameter that controls the width of the Gaussian kernel function. G-RBF has demonstrated better performance than other kernel function in many applications [14,18]. In addition, the choice of the kernel parameters is a practical problem for training an SVM [13]. The selected kernel has just one parameter simplifying, therefore, this selection. The kernel alignment method [19] is employed to automatically and quickly obtain the value of σ that provides a learning machine with the best performance. This method is a measure of similarity between the distance matrix obtained in the feature space of each data point with respect to the others and a target distance matrix containing the data labels. Let y be a l dimension column vector of labels, where l is the number of spectra in the training data set ($l = 960$), and K is the dot product matrix with, as previously seen, $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. The expression of the alignment between K and the matrix defined by yy' is the following:

$$A = \frac{\langle K, yy' \rangle_F}{\sqrt{\langle K, K \rangle_F \langle yy', yy' \rangle_F}} \qquad with \quad -1 \leq A \leq +1 \tag{9}$$

where $\langle K_1, K_2 \rangle_F$ is the Frobenius product between the two matrices K_1 and K_2 defined as:

$$\langle K_1, K_2 \rangle_F = \sum_{i,j=1}^{l} K_1(i, j) K_2(i, j) \tag{10}$$

Kernel alignment may be, therefore, seen as a correlation coefficient between two descriptions, K and yy', of the same data set. An ideally designed kernel with respect to the alignment will then reveal two independent structures for a two class classification problem.

In Section 3 the discrimination ability of SVM with the Gaussian kernel function is checked and a weak identification capacity is showed. Therefore, a more sophisticated treatment, namely the factor analysis in the feature space (k-PCA) described in the following lines, is justified.

By means of the Principal Component Analysis (PCA), data is expressed in such a way that the new basis vectors are those directions of the data which contains the most relevant information. Since PCA assumes linearity, the new basis is a linear combination of the original. A more detailed description of linear PCA application to captured hyperspectral images can be found in a previous work [2]. The aim of kernel-PCA is to accomplish a linear PCA in the feature space (F). Since the embedding, from input space to feature space, is performed with a non-linear relationship, with the g-RBF, a non-linear analysis is performed respect to input data. As for classical PCA, the new basis vectors are defined by the successive eigenvalues λ and eigenvectors V of the covariance matrix T_F:

$$T_F = \frac{1}{l} \sum_{j=1}^{l} \Phi(x_j) \Phi(x_j)' \tag{11}$$

where $\Phi(x)$ is the mapping function to embed input data into F. The new projection basis vectors are then given by:

$$\lambda V = T_F V \tag{12}$$

Since each vector V is a linear grouping of terms $\Phi(x)$ such as $V = \sum_{j=1}^{l} \mu_j \Phi(x_j)$ and K is the dot product matrix with $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, Eq. (12) becomes:

$$l\lambda\mu = K\mu \tag{13}$$

The problem reduces to find l normalized eigenvectors μ^k of K ($k = \{1,...,l\}$). All these vectors must be normalized, which leads to:

$$l = \lambda_k \left(\mu^k \cdot \mu^k \right) \tag{14}$$

To achieve the k-PCA algorithm, the following steps are performed: first, the dot product matrix K is computed; then, its eigenvectors μ^k are obtained, and, finally, the eigenvectors are normalized. Data projection onto the nth eigenvector in F is computed as

$$\left(V^n \cdot \Phi(x) \right) = \sum_{i=1}^{l} \mu_i^n \left(\Phi(x_i) \cdot \Phi(x) \right) \tag{15}$$

that is to say

$$\left(k - PCA \right)_n (x) = \sum_{i=1}^{l} \mu_i^n K(x_i, x) \tag{16}$$

3. RESULTS AND DISCUSSION

Discrimination capacities of SVM with just the g-RBF kernel, and the latter together with the factor analysis described in Section 2.2, are presented and discussed in this section. In addition, this is accomplished for both Vis-NIR and NIR imaging systems.

3.1 Results in the VIS-NIR range

As mentioned in Section 2.2 results obtained in three independent simulations are averaged. 960 spectra compose the training set and 480 the validation set. In this case, these are obtained from the hyperspectral images captured by one of the systems described in Section 2.1, that one whose spectral range goes from 400 to 1000 nm. To make easier the selection of the tuning parameter of the Gaussian kernel, its bandwidth σ, the previously described kernel alignment method is utilised. The evolution of the alignment A versus σ is depicted in Figure 3.
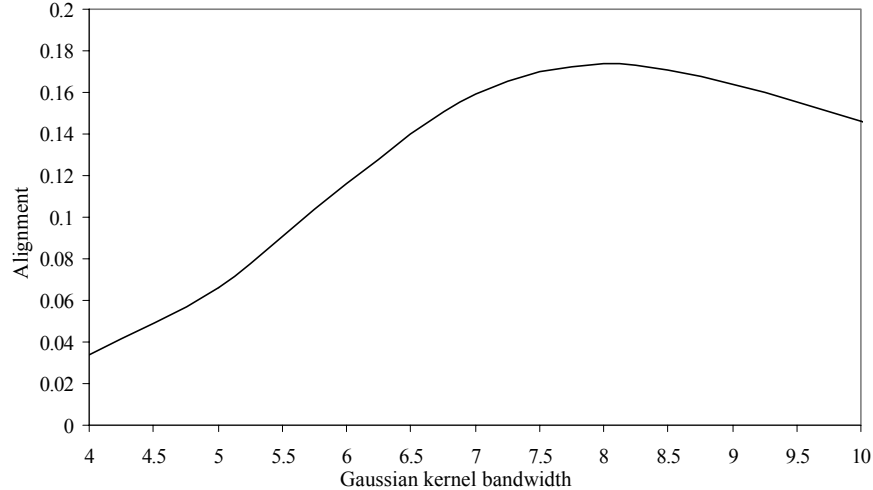

Fig. 3. Behaviour of the alignment versus the Gaussian kernel bandwidth in the Vis-NIR range.

A maximum alignment is obtained for $\sigma = 8$, and, therefore, this would be the employed value of the tuning parameter. Once σ is found, SVM with g-RBF kernel is performed. The classifier cannot even classify data from the training set, 51% misclassified. In addition, a lot of training spectra (933 out of 960) are support vectors, i.e. their corresponding Lagrange multipliers are non-null. This implies a high computational load. Hence, the use of factor analysis is perfectly justified. If k-PCA is applied and all provided non-null eigenvectors are maintained, only 58 training spectra are support vectors and a classification error of 1.11% is obtained. As depicted in Figure 4, if the number of maintained eigenvectors is reduced, classification error increases. In Figure 4 *th* stands for *threshold* and it is obtained by dividing the eigenvalue associated to the last kept eigenvector by the maximum eigenvalue of the matrix K, when eigenvectors are

sorted according to the value of their associated eigenvalues in descending order. Considered values of the threshold and their corresponding number of eigenvectors and support vectors are summarized in Table 2.
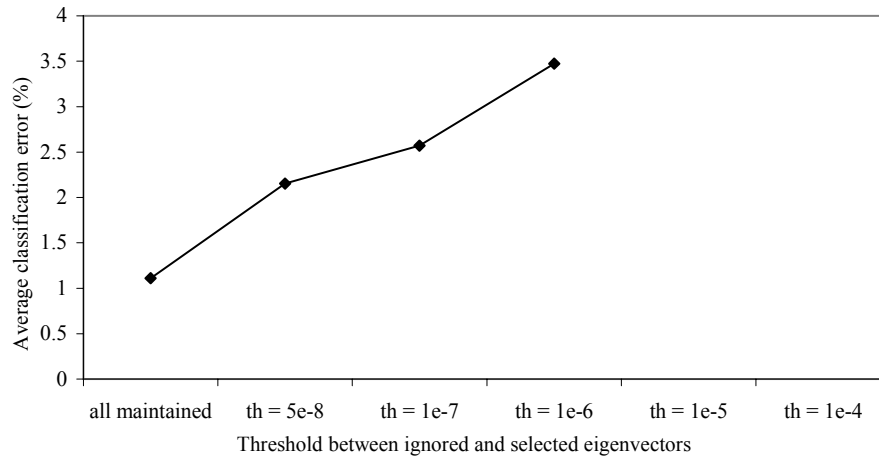

Fig. 4. Behaviour of the classification error when the number of k-PCA maintained eigenvectors is decreased.

Table 2. Corresponding number of eigenvectors and support vectors for each threshold value.

Threshold	Number of maintained eigenvectors	Number of obtained support vectors
0	30	58
5e-8	25	67
1e-7	21	113
1e-6	15	127
1e-5	7	134
1e-4	7	134

As shown, classification error increases as the number of maintained eigenvector decreases. Moreover, for threshold values greater than 1e-6, the system does not converge to a valid solution. This is due to the fact that each time more discriminant information is lost in the transformation, i.e. the classes are closer and system classification performance is worse. If the system does not converge, this means that all discriminant information has been discarded. Finally, Figure 5 depicts simultaneously the achieved execution time and the classification error as a function of the threshold, and it also performs a comparison with SVM with g-RBF kernel, i.e. without the inclusion of k-PCA. Apart from the increase in classification error, classification time also increases as the number of eigenvectors is reduced, because, as also demonstrated in Table 2, the number of data training spectra which are support vector is higher. It is worth noting that a great improvement in system performance has been achieved by employing in the feature space the new vectorial basis given by those directions of maximum variance. Obtained classification error is 1.11% compared with the previously obtained of 50.86%, and execution time is approximately one order of magnitude smaller (0.5457 s with respect to 6.896 s).
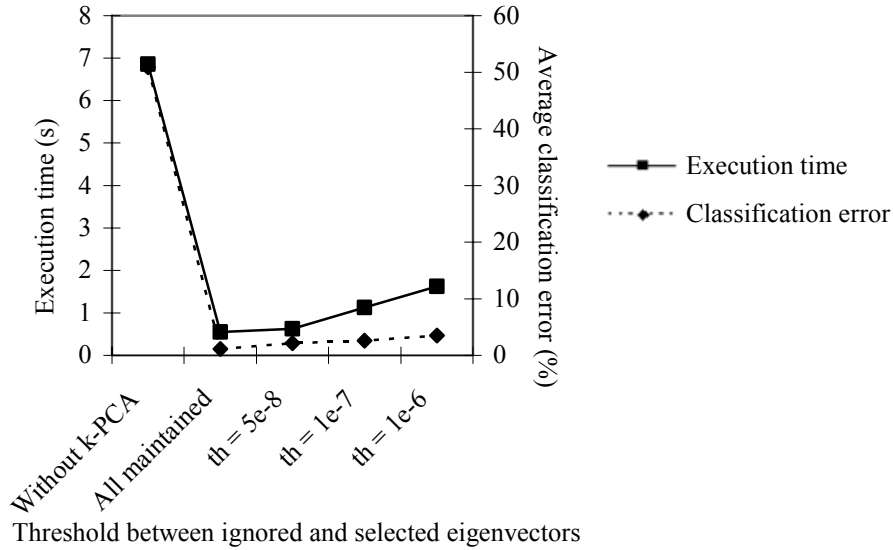
Fig. 5. Qualitative properties of the proposed method (computational burden and error) as a function of the number of maintained eigenvectors.

3.2 Results in the NIR range

The spectra which compose the training and validation sets are now obtained from the hyperspectral images captured by the system whose spectral range goes from 1000 to 2400 nm. As in the Vis-NIR, the kernel alignment method is first applied to determine the value of the kernel bandwidth which provides the best performance. Figure 6 depicts the alignment value versus the Gaussian kernel bandwidth.
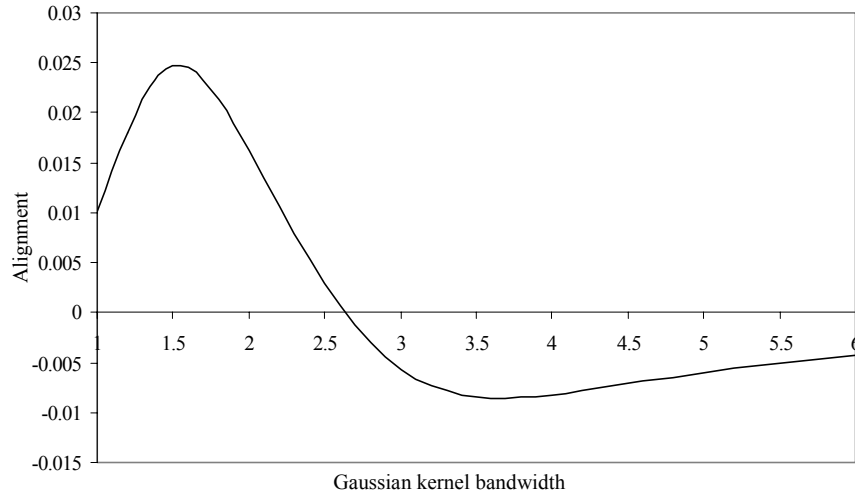

Fig. 6. Behaviour of the alignment versus the Gaussian kernel bandwidth in the NIR range.

The maximum alignment value is now achieved with $\sigma = 1.6$, and, hence, this is the selected value for the kernel tuning parameter in this second spectral range. As in the Vis-NIR range, if k-PCA is not considered, discrimination ability of SVM is very low, with 46% and 47% misclassified data in training and validation procedures, respectively. In addition, the number of support vector is even higher in this range, 958 out to 960. Therefore, the application of k-PCA is again required. Figure 7 depicts the error performance when the threshold between ignored and selected eigenvalues increases. The corresponding number of eigenvectors and support vector in this spectral range are summarized in Table 3. If all non-null eigenvectors are maintained when applying factor analysis, a 2.43% classification error is given. This figure is approximately maintained until the threshold is smaller than 1e-6. This means that some non-null eigenvectors provide

irrelevant information from the material discrimination/identification point of view, and they can, therefore, be discarded in the transformation. The number of required support vectors is also approximately constant if the threshold is under the abovementioned value. This number is also small, i.e. the system is going to be efficient in terms of the time performance.
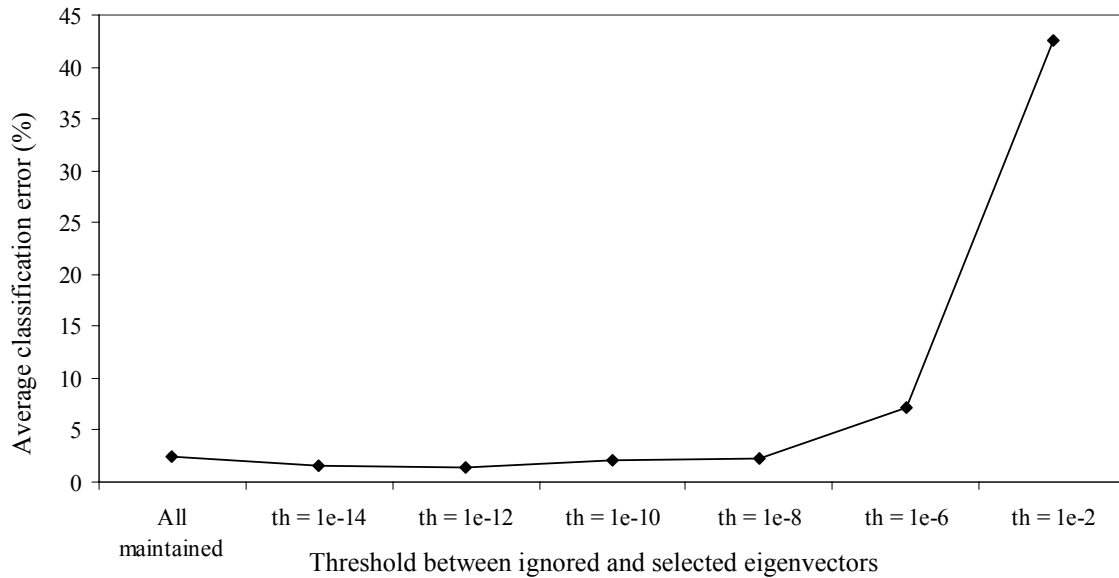

Fig. 7. Behaviour of the classification error when the number of k-PCA maintained eigenvectors is decreased.

Table 3. Corresponding number of eigenvectors and support vectors for each threshold value.

Threshold	Number of maintained eigenvectors	Number of obtained support vectors
0	125	24
1e-14	108	24
1e-12	76	26
1e-10	52	24
1e-8	30	22
1e-6	15	130
1e-2	3	949

Figure 8 presents the behaviour of the qualitative properties of the proposed method when it is applied to the images captured in the NIR range. As shown, both execution time and classification error remains nearly constant if the number of maintained eigenvectors is greater than 15. For 15 and 3 maintained eigenvectors, i.e. when threshold value is 1e-6 and 1e-2 respectively, both of them increase exponentially. A minimum classification error of 1.38% is obtained with 76 eigenvectors with corresponds to a threshold of 1e-12 and 26 required support vectors. Therefore, this would be the employed values in processing scheme design. Again, a great improvement in system performance is achieved by the employment of k-PCA factor analysis, since classification error decreases from 47% to 1.38%. Execution time reduction is less remarkable than the one achieved in the Vis-NIR range, because in this case the number of required support vectors when SVMs with just g-RBF are employed the number of required vectors were of the same order of magnitude than when also k-PCA is applied. This is due to the mentioned fact that, as some non-null eigenvectors provide redundant information, the number of required support vectors does not increase when they are discarded.
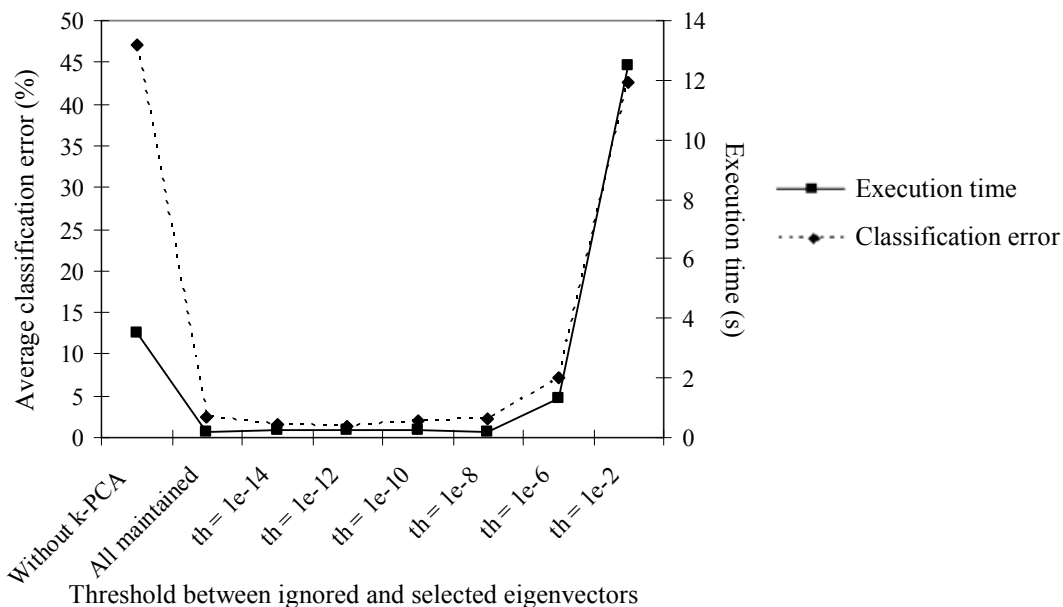
Fig. 8. Qualitative properties of the proposed method (computational burden and error) as a function of the number of maintained eigenvectors.

The NIR imaging system is not optimized from the calibration point of view, because, as commercial available sensors working in this spectral range are very recent, there are not automatical calibration tools. In spite of this fact, it has proved successful in material discrimination when the captured images are processed with a scheme consisting of SVM with g-RBF kernel and k-PCA analysis. Therefore, this spectral range can be concluded as the more convenient for the considered application case.

4. CONCLUSSIONS

The ability of SVM to perform the classification of hyperspectral images from two different spectroscopic sensors with application to material discrimination has been demonstrated. In addition, the employed sensors work in non-overlapping spectral ranges of the spectrum, and, therefore, they capture different spectral signatures of the material under analysis. Classification capability of SVM with Gaussian kernel, which is according the bibliography the one which provides the best performance, has been shown to be insufficient in this particular application case. Therefore, the need of a more sophisticated treatment to achieve acceptable classification errors is justified. The employment of the factor analysis k-PCA has provided the adequately classification of material in both considered spectral ranges.

ACKNOWLEDGEMENTS

REFERENCES

[1]   E. Herrala, T. Hyvarinen, O. Voutilainen and J. Lammasniemi, "An optoelectronic sensor system for industrial multipoint and imaging spectrometry", Sens. Actuator A-Phys. 61, 335-338 (1997).
[2]   P.B. García-Allende, O.M. Conde, A.M. Cubillas, C. Jáuregui, J.M. López-Higuera, "New raw material discrimination system based on a spatial optical spectroscopy technique", Sens. Actuator A-Phys. 135, 605-612 (2007).

[3]   G. ElMasry, N. Wang, A. ElSayed and M. Ngadi, "Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry", J. Food Eng. 81, 98-107 (2007).

[4]   G. Zavattini, S. Vecchi, R.M. Leahy, D.J. Smith, S.R. Cherry, "A hyperspectral fluorescence imaging system for biological applications", IEEE Nuclear Science Symposium, Vol. 2, pp. 942-946, 2003.

[5]   C.H. Chen, *Fuzzy Logic and Neural network Handbook*, McGraw Hill, New York, 1996.

[6]   F. Girosi, J. Makhoul, E. Manalakos, E. Wilson, "Neural networks for signal processing V", Proc. of the 1995 IEEE Workshop, IEEE Press, New York, 1995.

[7]   V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.

[8]   F. Kahraman, A. Capar, A. Ayyaci, H. Demirel, M. Gokmen, "Comparison of SVM and ANN performance for handwritten character classification" Proc. of the IEEE 12th Signal Processing and Communications Applications, IEEE Cat. No.04EX797, 615-18, 2004.

[9]   W.H. Chen, S.H. Hsu, H.P. Shen, "Application of SVM and ANN for intrusion detection", Comput. Oper. Res. 32 (10), 2617-2634 (2005).

[10]   W.T. Wong, S.H. Hsu, "Application of SVM and ANN for image retrieval", Eur. J. Oper. Res. 173, 938-950 (2006).

[11]   B. Peng, L. Junhua, "Application of SVM to mixed gas spectrum analysis", Chinese Journal of Scientific Instrument 27, 1242-7 (2006).

[12]   G. Camp-Valls, L. Bruzzone, "Regularized methods for hyperspectral image classification", Image and Signal Processing for Remote Sensing X, Proc. of SPIE 5573, 226-237 (2004).

[13]   S.G. Kong, Z. Du, M. Martin, T. Vo-Dinh, "Hyperspectral fluorescence image analysis for use in medical diagnostics", Advanced Biomedical and Clinical Diagnostic Systems III, Proc. of SPIE 5692, 21-28 (2005).

[14]   L. Jiang, B. Zhu, X. Rao, G. Berney, Y. Tao, "Discrimination of black walnut shell and pulp in hyperspectral fluorescence imagery using Gaussian kernel function approach", J. Food Eng. 81, 108-117(2007).

[15]   Matlab® Reference Manual, The Mathworks Inc., MA, USA.

[16]   C. Goutte, "Note on free lunches and cross validation", Neural Comput. 9, 1211-1215 (1997).

[17]   H. Zhu, R. Rohwer, "No free lunch for cross-validation", Neural Comput. 8, 1421-1426 (1996).

[18]   G. Camp-Valls, L. Gómez-Chova, J. Calpe-Maravilla, "Robust Support Vector Method for Hyperspectral Data Classification and Knowledge Discovery", IEEE Trans. Geosci. Remote Sensing 42(7), 1530-1542 (2004).

[19]   Y. Langeron, M. Doussot, D.J. Hewson, J. Duchêne, "Classifying NIR spectra of textile products with kernel methods", Eng. Appl. Artif. Intell. 20, 415-427 (2007).