From Narrow to Broad Band Design and Selection in Hyperspectral Images*

Adolfo Martínez-Usó, Filiberto Pla, José M. Sotoca, and Pedro García-Sevilla

Dept. Lenguajes y Sistemas Informáticos, Jaume I Univerisity
{auso,pla,pgarcia}@uji.es
http://www.vision.uji.es

Abstract. Selecting the most relevant bands from a hyperspectral image would considerably reduce the amount of data without practically losing relevant information. In addition, if some physical and signal criteria of this selection are taken into account, the obtained results grouping consecutive bands would be useful to design new filters for hyperspectral cameras in order to improve the efficiency of these devices. Starting from certain number of pre-selected bands, intervals of spectrally adjacent instances to these initial bands are considered for calculating new broader bands. Results will show how a weighted average on these intervals can keep, or even improve, the performance respecting to a narrower selection, avoiding, at the same time, common drawbacks from the narrow-band acquisition devices.

1 Introduction

Hyperspectral devices are called to become a generalised technique for many image processing applications since they do not only increase the amount of information related to shapes, colours and chemical composition, but much of this new information reveal important details that are impossible to detect by means of devices used so far. Thus, they are being introduced in important and demanding application fields like remote sensing, medical imaging, product quality inspection, fine arts, etc. However, hyperspectral imaging also entails dealing with large amounts of information with high spectral correlation and, therefore, containing important data redundancies. On the other hand, although a reduced feature representation of the data will lead to a loss of information, keeping only the most discriminative features or data points, data analysis can become simpler, losing no significant performance and, in many cases, results could also become better if the noisy features are avoided by means of the selection process. Thus, the challenging question would be how to select the correct instances from the input set of bands in order to characterise the problem properly. Authors in [3] addressed this question from the point of view of the information theory in an unsupervised scheme.

As figure 1 shows, a hyperspectral camera is basically a device with a filter and a sensor that are usually sensitive to wider ranges in the spectrum than an ordinary colour

* This work was supported by Spanish Ministry of Science and Education under Projects ESP2005-07724-C05-05, CSD2007-00018, PET2005-0643 and HP2005-0095 as well as Generalitat Valenciana under the project GV/2007/105.

A. Campilho and M. Kamel (Eds.): ICIAR 2008, LNCS 5112, pp. 1091–1100, 2008.

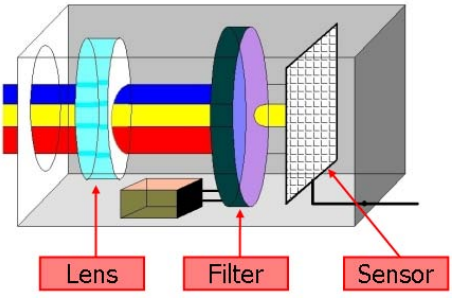


Fig. 1. Hyperspectral camera schema. Filter absorbs the whole spectrum except the spectral band determined by the user and, thereby, just the selected wavelength goes through this filter and reaches the camera sensor.

camera. Nowadays there exist commercial systems that are able to cover ranges from 100nm to 2500nm (visible spectrum starts around the 400nm and extends to 700nm approximately), taking a certain number of bands in a spectrum range. The interval in nanometres between two acquired samples is called spectral resolution or sampling interval. Obviously, selecting an optimal spectral resolution value is not an easy task [5]. The narrower the sampling interval, the more detailed the input information but, however, a sensor with a narrow spectral resolution involves several drawbacks:

1. The amount of data that any further process has to deal with is larger, becoming impracticable in some cases.
2. These kind of sensors are usually quite expensive indeed.
3. Signal-to-noise ratio (*SNR*) is reduced since sensor efficiency decreases at narrow sampling intervals [6].

On the other hand, an important disadvantage of a filter with a wide acquisition range is the effect of the spectral aliasing from the absorption features of different materials. In a broadly acquisition, this effect is a very common drawback which is attenuated by adequate passband filters. The mean-passband frequency of these filters is not usually suitable for all the input frequencies. Therefore, in order to overcome this problem, a weighted-passband filter can be applied (Figure 2).

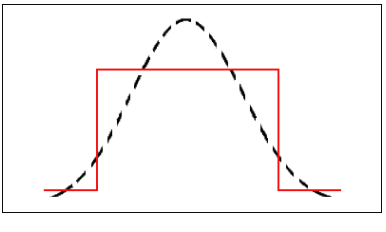


Fig. 2. Gaussian passband filter. Continuous line represents a mean-passband filter whereas the discontinuous one describes a gaussian weighted-passband filter.

Depending on the nature of the problem and/or the available resources, the dimensionality reduction issue can be faced by means of supervised or unsupervised procedures. In many applications, like in remote sensing, it is very difficult (and expensive) to obtain the ground truth. Thus, imaging the cost/difficulty of obtaining the category or parameters for each pixel in an image for a remote sensing application if each represented area has to be visited to properly label the land-use and the measured biophysical parameters (temperature, chlorophyll content, LAI, etc.). Therefore, taking into account our scope which is mainly based on remote sensing, it is important to note that we will focus this work on unsupervised approaches.

This paper describes a new unsupervised broad-band selection algorithm based on the WaLuMI algorithm presented in [3], which was successfully compared with well-known band selection techniques of the literature, demonstrating to be highly competitive and computationally affordable, even for hyperspectral images with a large number of input bands [4]. Around each band selected by the WaluMI algorithm, a continuous interval is defined. A new subset of resulting bands are calculated as a weighted average of these intervals. Thus, this new broad-band selection algorithm presents two main contributions:

- Firstly, experimental results show that the method provides a very suitable subset of broader spectral bands for pixel classification purposes, improving in many cases the classification rates obtained by the original WaLuMI algorithm.
- Secondly, from a physical or signal point of view, group consecutive bands could be useful to design more efficient and less costly sensors for hyperspectral cameras.

2 A Procedure for Selecting Spectral Bands

In [3], authors proposed a method based on a hierarchical clustering structure that groups bands to minimise the intra-cluster variance and maximise the inter-cluster variance. Two distances based on information theory measures were used in this approach in order to discriminate among bands. One distance was based on the mutual information concept and the second on the Kullback-Leibler divergence. Our proposed algorithm is based on the first one, which was called WaLuMI (*Ward's Linkage strategy using Mutual Information*). The methodology of the algorithm can be briefly explained following the next steps:

- A dissimilarity space is defined among image bands, where a distance criterion is defined based on the mutual information concept between any pair of bands.
- From the initial set of bands that form a hyperspectral image, the process starts with a hierarchical clustering in the defined dissimilarity space. In order to progressively construct a hierarchical family of derived clusters the method uses a linkage strategy with an inter-cluster distance as the objective function to optimise.
- Finally, a band representing each final cluster is chosen. Eventually, the K selected bands from the final K clusters will provide an adequate reduced representation in order to manage satisfactory results in any further process. They will be the best representative instances from each cluster, reducing redundancy as much as possible.

2.1 Selecting Cluster Representatives

Let us focus on the last step of the WaLuMI algorithm where a final instance from each cluster is selected. We shall consider a resulting cluster C with R bands. A weight of each band $X_i \in C$ is defined as,

$$W_i = \frac{1}{R} \sum_{j \in C, j \neq i} \frac{1}{\epsilon + D(X_i, X_j)^2} \tag{1}$$

where ϵ is a very small positive value to avoid singular values. Function $D(X_i, X_j)$ returns the distance value between bands i, j and it is the same distance based on mutual information that was also used in the clustering process.

The representative band from each group is selected as the band with the highest W_i in the cluster. A low value of W_i means that the band i has an average large distance from the other bands in the cluster, that is, in this case, the band i will have an average low correlation with regard to the other bands in the cluster. In a reverse way, a high value of W_i means that band i has, in average, a high correlation with regard to the other bands in the cluster. Thus, the band with the highest average correlation with regard to the other bands in the cluster is chosen. That means to select the band that better predicts the information content of the other bands in the cluster, since the more mutual information two random variables share, the more can predict one of the variable about the other one and, in this sense, having a high degree of dependence among them.

3 New Representatives from Subsets of Consecutive Bands

The previous section briefly summarises the original WaLuMI algorithm, paying special attention to its last step where the final set of bands are selected. Thus, this section explains how a new set of bands are calculated from intervals of bands containing those instances selected by the standard WaLuMI algorithm.

Let us consider a cluster C with R bands. Let us suppose a band X_i that has been selected as a representant of the cluster C by the WaLuMI algorithm. We shall define a subset of r contiguous bands $S = \{X_1, \ldots, X_i, \ldots, X_r\} \subseteq C$ so that $r \leq R$ that is, it would be a subset of consecutive bands of C containing the band X_i.

Let us take now the weights calculated by means of (1) for each band in cluster C. A new weight $\hat{W}_i$ for the band $X_i \in S$ is calculated as follows:

$$\hat{W}_i = \frac{W_i}{\sum_{j=1}^{r} W_j} \tag{2}$$

Thus, based on the W values that the WaLuMI algorithm assigned to each band, a new set of weight values are derived. Finally, a new resulting band in representation of the cluster C is obtained, where each pixel $p^{x,y}$ is calculated as follows:

$$p^{x,y} = \hat{W}_1 p_0^{x,y} + \ldots + \hat{W}_i p_i^{x,y} + \ldots + \hat{W}_r p_r^{x,y} \tag{3}$$

x and y are the image coordinates of the pixel p. Therefore, this broad-band selection process chooses a new representative of the cluster C by means of a weighted average of the subset of bands S[1].

[1] A simple average band of the instances in S were also tested resulting in a slight poorer results.

4 Experiments

As a validation criterion, a *k-Nearest Neighbourhood* rule [1] with $k = 3$ has been applied to test the significance of the subsets of selected bands obtained using the hyperspectral images described in table 1. Samples from each database have been divided in training subsets and test subsets. At the same time, these subsets are also divided in

Table 1. Left column shows database examples presented as RGB compositions. From top to bottom, 92AV3C, HyMap, CHRIS-PROBA and VIS collections. Right column offers a description of each database. Bands discarded because of the low *SNR* are also described in order to clarify results in the further graphs.

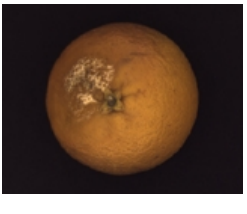
	The $92AV3C$ source of data corresponds to a spectral image that covers from 400 nm to 2500 nm (145 X 145 pixels, 220 bands and 17 classes composed of different crop types, vegetation, man made structures, and an unknown class). It was acquired with the AVIRIS data set and collected in June 1992 over the Indian Pine Test site in Northwestern Indiana (http:/dynamo.ecn.purdue.edu /∼biehl/MultiSpec). As described in [2], several bands should be discarded from this database due to the effect of the atmospheric-absorption. Thus, 185 out of the 220 bands were used, discarding the lower *SNR* bands (0-3, 102-109, 148-164, 216-219).
	$DAISEX'99$ project provides useful aerial images about the study of the variability in the reflectance of different natural surfaces. This source of data, referred as $HyMap$ in figures/tables, corresponds to a spectral image (700 X 670 pixels and 7 classes composed of crops and an unknown class) acquired with the 128-bands HyMap spectrometer during the DAISEX-99 campaign (http:/io.uv.es/projects/daisex/). In this case, 126 bands were used, discarding the lower *SNR* bands (0, 64).
	Satellite PROBA has a positional spectroradiometric system (CHRIS) that measures the spectral radiance. The images used in this study come from the CHRIS-PROBA mode that operates on an area of 15×15 Km, with a spatial resolution of 34 m. obtaining a set of 62 spectral bands that range from 400 to 1050 nm (641 X 617 pixels and 9 classes composed of crops and an unknown class). Concretely, images cover the area known as Barrax (Albacete - Spain). In this case, 52 bands were used, discarding the lower *SNR* bands (25, 33, 36-37, 41-43, 47, 50, 53).
	This is an example of using hyperspectral imaging in other applications different from remote sensing. It contains images of orange fruits obtained by an imaging spectrograph (RetigaEx, Opto-Knowledge Systems Inc., Canada). A spectral range extended from 400 to 720 nm in the visible (VIS, 676 X 516 pixels, 33 spectral bands for each image) is covered. Regarding to the database content, it is composed of several types of unhealthy skin (defects as rot, trip, overripe or scratch). No band was discarded in this case.

Table 2. Training (TRA) and test (TST) size of the sets for 1 partition out of five ($\times 5$) in each database

	HyMap	92AV3C	VIS	Chris-Proba
Pixels per partition (TRA)	$46900(\times 5)$	$2102(\times 5)$	$34882(\times 5)$	$1788(\times 5)$
Pixels per partition (TST)	$46900(\times 5)$	$2102(\times 5)$	$34882(\times 5)$	$1788(\times 5)$

five partitions, assigning the samples to each partition in a random way (cardinality of the different subsets is given in table 2). Classification rates for the standard WaLuMI and our proposal are provided by averaging classification accuracies obtained by the classifier over the five random partitions in each image data set. Samples so distributed satisfies that,

1. The sum of the elements from the different partitions constitutes the entire original set and the *a priori* probabilities for each class in the data sets are preserved.
2. The statistical independence between training and test sets of every partition.

Thus, the performance, both the standard WaLuMI and its modified version presented in this work, has been evaluated by means of a supervised pixel classification process. Table 3 shows the classification rates obtained by each method in the HyMap and 92AV3C databases. As it can be seen, not only the modified version keeps the performance of the standard WaLuMI but, in some cases, the classification rates are even improved.

Figure 3 shows in each row the $K = 5$ groups of bands that have been selected for the 92AV3C (top row), HyMap (middle row) and VIS (bottom row) databases respectively. In the left side, graphs with the spectral resolution (bar width on X-axis) of

Table 3. Classification rates for HyMap and 92AV3C databases

	HyMap DB		**92AV3C DB**	
Num of bands	WaLuMI original	WaLuMI modified	WaLuMI original	WaLuMI modified
1	**18.85**	9.71	**14.00**	13.00
2	42.48	**48.22**	21.80	**43.15**
3	51.94	**58.76**	54.20	**55.09**
4	56.50	**61.51**	68.20	**68.50**
5	61.20	**64.66**	**77.00**	75.63
6	61.87	**64.85**	**79.00**	78.32
7	61.08	**64.62**	80.00	**81.90**
8	61.54	**65.01**	82.40	**83.97**
9	64.15	**65.47**	83.80	**86.14**
10	64.36	**65.76**	85.00	**86.94**
11	64.86	**65.64**	84.00	**87.17**
12	65.17	**65.84**	84.80	**87.96**
13	66.53	**67.24**	85.00	**88.49**
14	66.67	**67.32**	85.20	**88.61**
15	66.70	**67.39**	86.00	**88.66**

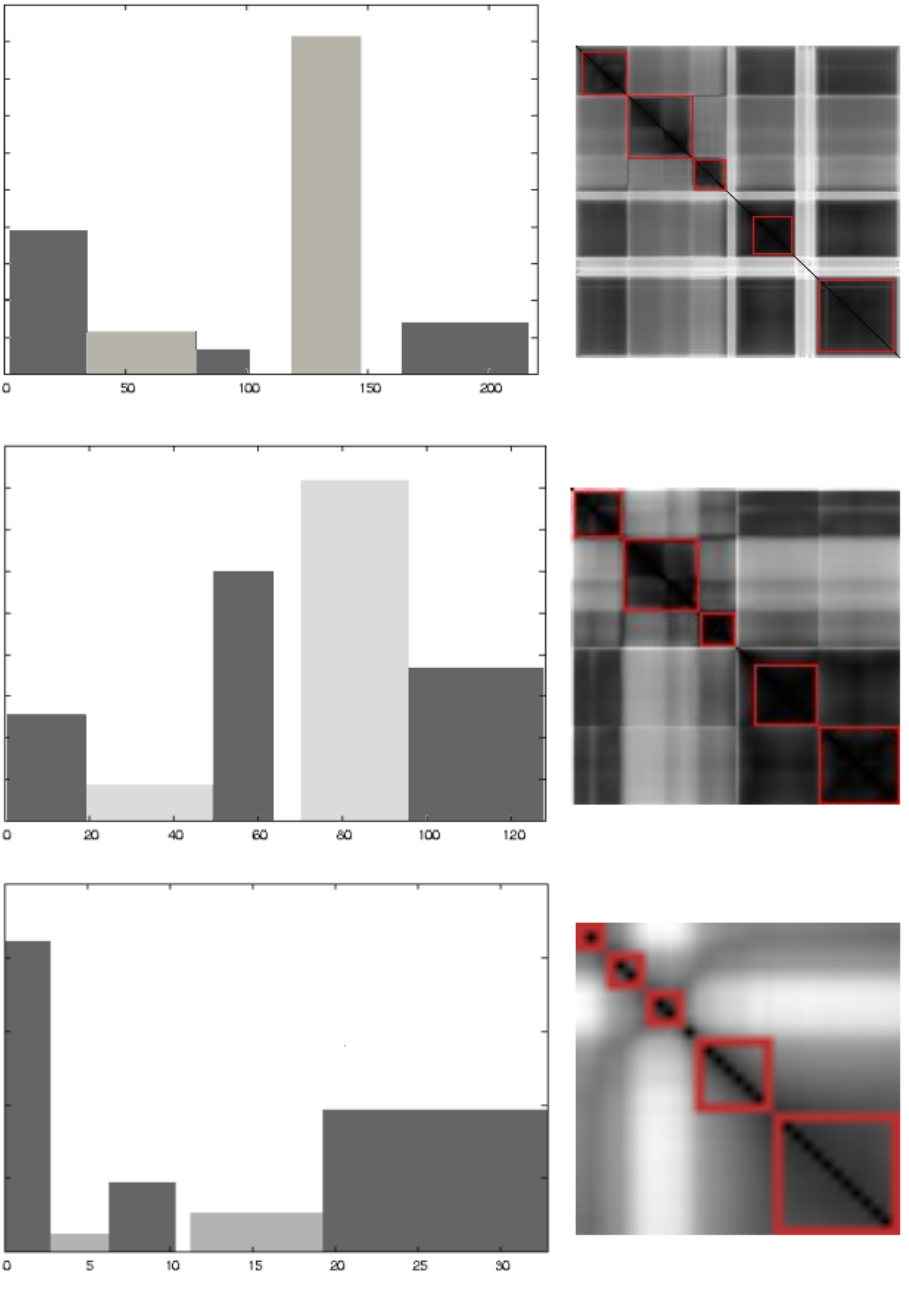


Fig. 3. From top to bottom row, 92AV3C, HyMap and VIS databases respectively. Left-side graph represents the $K = 5$ groups of consecutive bands selected by the algorithm. Right-side picture shows where these groups are located in the distance matrix. Note that figures are scaled in order to show the results properly, however, the number of bands (X-axis) in graphs and the dimension in distance matrices are different.

each interval of continuous bands are shown. Results of the relevance of each cluster are also presented using bars (bar height on Y-axis). The bar height represents the average of the $\hat{W}$ values calculated by means of (2) for each band in the cluster. Thus, the relative relevance of each interval of bands with respect to all the intervals is proportional to the height, i.e., the higher the bar in the graph, the more correlated the bands in the cluster. On the other hand, right side figure shows the distance matrix as an image where each pixel (i, j) represents the distance value calculated between the ith band and the jth band. Thus, the darker the pixel. the closer the bands according to the distance employed. At the same time, by means of red squares, how the intervals of bands selected are distributed in the distance matrix has been enhanced. Note how the selected intervals fit into the darker areas, corresponding to the groups of highly correlated bands.

In [3], authors considered from the first selected band to the 15*th* selected band as the transitory zone of the learning curve, before reaching the flat zone of the curve. In this range, we have checked how the intervals of consecutive bands have been distributed when the number of selected instances is $K = 5$, $K = 10$ and $K = 15$. Figure 4 shows these partitions by means of three graphs of bars. Classification rates obtained by the standard WaLuMI and the proposed technique here are also shown in the top-left graph.

From the graphs shown in figures 3 and 4, it is worth noticing that the intervals of unselected bands mostly correspond to the noise bands discarded in each database (see database details in table 1). Therefore, instances effectively involved in the band selection process tend to form clusters of spectrally adjacent bands.

As an example of application to another field, different from remote sensing, we offer the results obtained using the VIS database, which is involved in a project related to the quality inspection of fruits. Figure 5 shows a graph with the classification rates obtained by the modified WaLuMI against the standard one (left) and the spectral resolutions that are covered by the selected groups of bands (right). It is important to stress that this database has no discarding bands because of poor *SNR* and, as it can be seen, bands are grouped in consecutive clusters covering the entire range of the spectrum.

As it has been described in table 1, all the databases present a spectral resolution of 10 nm. However, graphical results in figures 3, 4 and 5 point out that several bands could be grouped together, meaning that the spectral resolution could be increased in all cases. Thus, on the basis of the smallest interval of continuous bands calculated for each database and knowing that there is a band for each interval of 10 nm in the spectrum, new spectral resolutions could be reasonably moved to the values shown in table 4.

Table 4. New possible values (in nanometres) for the spectral resolution in each database

	HyMap	92AV3C	VIS	Chris-Proba
$K = 5$	140	220	30	30
$K = 10$	20	40	20	20
$K = 15$	20	20	20	10

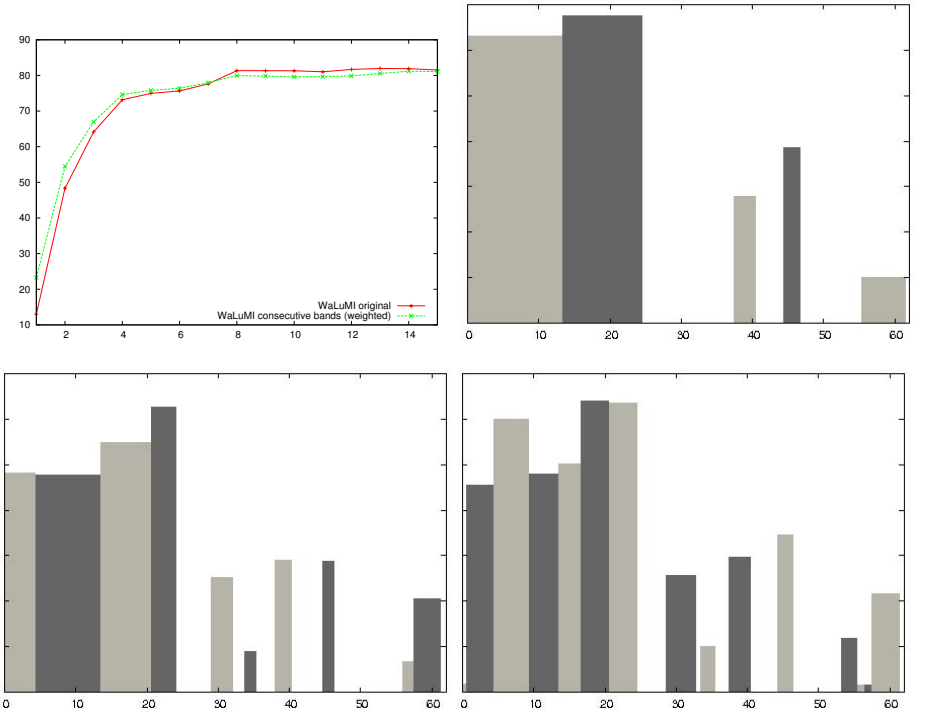


Fig. 4. CHRIS-PROBA database. The top-left graph shows the classification rates obtained by the WaLuMI algorithm and the new proposed variation as a comparative. The bar graphs represent the $K = 5$, $K = 10$ and $K = 15$ groups of consecutive bands selected by the proposed algorithm.

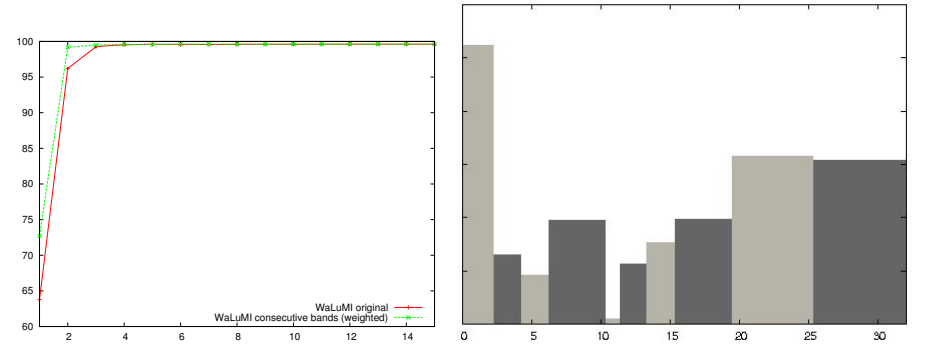


Fig. 5. VIS database. Left graph shows the classification rates obtained by the WaLuMI algorithm and the new proposed variation as a comparative. As an example of the groupings produced on this database, bar graph on the right represents the $K = 10$ groups of consecutive bands selected by the proposed algorithm.

5 Conclusions

Starting from a contrasted band selection technique, we have developed an approach that overcomes the drawbacks described in the introduction section by means of selecting wider intervals of consecutive bands. We have presented a fully unsupervised methodology which is computationally affordable, avoiding the problem of labelling complicated databases. The classification accuracies reached by this new technique are not reasonably affected regarding to the ones obtained by the standard version of the band selection algorithm, being in many cases even better. Therefore, these results demonstrate that this method could be useful in further applications in order to select the appropriate subset of input bands.

Finally, one of the main contributions of this work can be found from the remote sensing point of view, since the classification results obtained averaging consecutive bands are also relevant enough to think about designing sensors with a wider spectral resolution. Thus, further approaches will determine if this optimisation could mean the design of cheaper sensors with more accurate results for classification purposes.

References

1. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Transactions on Information Theory 13(1), 21–27 (1967)
2. Landgrebe, D.A.: Signal theory methods in multispectral remote sensing. Wiley, Chichester (2003)
3. Martinez-Uso, A., Pla, F., Sotoca, J.M., Garcia-Sevilla, P.: Clustering-based hyperspectral band selection using information measures. IEEE Trans on GRS 45(12), 4158–4171 (2007)
4. Martinez-Uso, A., Pla, F., Sotoca, J.M., Garcia-Sevilla, P.: Comparison of Unsupervised Band Selection Methods for Hyperspectral Imaging. In: Martí, J., Benedí, J.M., Mendonça, A.M., Serrat, J. (eds.) IbPRIA 2007. LNCS, vol. 4477, pp. 30–38. Springer, Heidelberg (2007)
5. Price, J.C.: Spectral band selection for visible-near infrared remote sensing:spectral-spatial resolution tradeoffs. IEEE Trans on GRS 35(5), 1277–1285 (1997)
6. Sun, L., Staenz, K., Neville, R., White, H.: Impact of sensor signal-to-noise ratio and spectral characteristics on hyperspectral geoscience products. In: Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on, pp. 2064–2067 (2006)