Unsupervised Image Segmentation based on Texems for Hyperspectral data

Adolfo Martínez-Usó Filiberto Pla Pedro García-Sevilla
Dept. Lenguajes y Sistemas Informáticos. University Jaume I
auso,pla,pgarcia@uji.es www.vision.uji.es

Abstract

There is no doubt about how useful and valuable the information provided by the hyperspectral sensors can be. Image segmentation procedures can take advantage of this information to increase the ability for separating different textures in an image. A multiscale approach for segmenting hyperspectral images is presented in this work. The method is based on the recently proposed texem model which has been extended in this work to spaces of high dimensionality. Furthermore, the hyperspectral extension of the texem-based segmentation would be computationally impracticable without a prior step for reducing the dimensionality. Thus, a band selection process based on the mutual information among bands has also been applied. The complete process is particularly useful in applications for remote sensing or quality inspection tasks.

Introduction

The benefits of hyperspectral imaging in several disciplines are becoming relevant for many emerging applications. Comparing with grey or colour images, multi or hyperspectral sensors acquire data from a much wider range of wavelengths in the spectrum and these data are being introduced in important and demanding application fields like remote sensing, medical imaging, product quality inspection, fine arts, etc. Image segmentation on these data means to detect salient regions in a hyperspectral image and isolate them as accurate as possible. This purpose presents two main problems:

- Firstly, hyperspectral imaging involves a huge amount of information, thus we have to face with the problem of *selecting* only the useful information avoiding redundancies.
- Secondly, the problem of *segmentation* strictly speaking is still a challenging question whatever the input image would be. Thus, it will be particularly hard in hyperspectral data where, in addition to the traditional segmentation difficulties, researchers have to face with the complicated task of managing the amount of information that a hyperspectral image involves.

The aim of the dimensionality reduction techniques is to manage a representation of the hyperspectral data avoiding as many irrelevant and redundant features as possible. In this way, further applications would take profit from a dimensionality reduction in terms of computational or time savings and, in many cases, in terms of accuracy. In this sense, authors in [5] presented a band selection algorithm based on information theory measures. In comparison with other relevant methods in the literature, their algorithm demonstrated a high performance, achieving good classification results and a particularly better consistency among all the tested databases. A more detailed explanation for this band selection technique will be given in the next section.

The recently proposed *texem* (texture exemplar) model has been successfully applied on image segmentation of colour images [8][9]. It performs a multiscale approach which is based on the assumption that a given image can be generated from a collection of image patches. An immediate extension that makes this approach suitable for segmenting high dimensional spaces is proposed in this paper.

This work is focused on solving the whole process, that is, from the data acquired by the hyperspectral sensors, we will firstly obtain a notably reduced subset of bands. A second step will consist of a procedure that segments this reduced representation. Due to the high variability of the input images, where labelled data usually are hard to obtain, the entire process must be done in an unsupervised way. Experimental results will show that the proposed method achieves accurate results with well-outlined boundaries in hyperspectral images.

Band Selection approach

Obviously, hyperspectral images provide much richer spectral information than grey-level or colour images. However, this advantage also comes with the drawback of dealing with large amounts of information. In order to cope with such amount of information, a usual technique consists of applying a dimensionality reduction to select and keep the relevant information. Thus, in this work, the *WaLuMI* method [5] was applied to choose a subset of relevant instances from the initial set of bands.

WaLuMI is a dimensionality reduction method that uses a distance based on mutual information to discriminate among bands. Mutual information is not only widely used as a criterion for measuring the degree of independence between random variables, but it also measures how much a certain variable can explain the information content about another variable, being a generalised correlation measure [1, 7]. Thus, the following measure of similarity between two random variables X_i and X_j (representing the i and j spectral bands) will be used as a relevance criterion,

$$NI(X_i,X_j) = \frac{2 \cdot I(X_i,X_j)}{H(X_i)+H(X_j)}, \tag{1}$$

where $H(\cdot)$ denotes the entropy of a random variable. Since the mutual information I is not a suitable measure of similarity by itself [5], we use NI as a normalised measure of the mutual information. Finally, this NI takes part in the distance measure as follows [2]:

$$D_{NI}(X_i,X_j) = \left(1-\sqrt{NI(X_i,X_j)}\right)^2 \tag{2}$$

Starting from a set of L input bands, a hierarchical clustering process is performed in order to form the Q final subsets of bands ($Q \ll L$). This clustering process is based on a matrix of distances that is updated when two clusters are merged

©2008 Society for Imaging Science and Technology

following the Ward's linkage rule [6]. This linkage strategy has the property of producing minimum variance partitions, grouping the pair of clusters that minimise the increment in the square error of the whole partition. The error used to this calculation is the intra-cluster dispersion. Thus, this method is also called minimum variance clustering, because it pursues to form each possible group in a manner that minimises the loss associated with each grouping (internal cohesion). Several studies point out that this method outperforms other hierarchical clustering methods [3], but, in our case, the process also helps us to form groups with low variance in their level of similarity.

Briefly summarising the linkage strategy, let us suppose that clusters C_r and C_s are merged. The general expression for the distance between the new cluster (C_r, C_s) and any other cluster (C_k) is defined as:

$$\begin{aligned} D[(C_k),(C_r,C_s)] \quad = \quad & \alpha \cdot D(C_k,C_r) \; + \\ & \beta \cdot D(C_k,C_s) \; + \\ & \gamma \cdot D(C_r,C_s) \; + \\ & \delta \cdot |D(C_k,C_r) - D(C_k,C_s)|, \end{aligned} \tag{3}$$

where α, β, γ and δ are the merging coefficients. Ward's inter-cluster distance results from the following coefficients,

$$\alpha = \frac{n_r + n_k}{n_r + n_s + n_k}, \;\; \beta = \frac{n_s + n_k}{n_r + n_s + n_k}, \;\; \gamma = \frac{-n_k}{n_r + n_s + n_k}, \;\; \delta = \emptyset,$$

where n_i is the number of instances in group i.

Thus, summarising the entire strategy in several steps:

1. The algorithm starts with the disjoint partition where each cluster is formed as a single pattern (hyperspectral band). At this step, the distance matrix $M_{L\times L}$ is initialised by means of the dissimilarity measure described on equation (2).
2. After that, the algorithm looks for the two most similar clusters, i.e., those clusters that will have the minimum distance value in the distance matrix.
3. Then, these two clusters are merged into one. Matrix $M_{L\times L}$ is updated using expression (3) and, therefore, $M_{L\times L}$ turns into $M_{(L-1)\times(L-1)}$. The rows/columns corresponding to the merged clusters are deleted and a row/column for the new cluster is added.
4. This process is repeated until the Q number of desired clusters are obtained ($M_{Q\times Q}$). The resulting mutually exclusive clusters represent groups of highly correlated bands, and bands from two different clusters will have low correlation.

The clustering is part of an information reduction process and, at the end of the clustering process, a representative band for each cluster is selected, which will substitute all bands in the cluster, at the lowest possible cost in terms of information loss. Thus, let us consider now a resulting cluster C with R bands. A weight of each band $X_i \in C$ is defined as,

$$W_i = \frac{1}{R} \sum_{j \in C, j \neq i} \frac{1}{\varepsilon + D_{NI}(X_i, X_j)^2} \tag{4}$$

where ε is a very small positive value to avoid singular values. The representative band from each group is selected as the band with the highest W_i in the cluster. In this way, we choose the band in the cluster with the highest average correlation (mutual information) with regard to the other bands in the cluster. It is equivalent to select the band that better predicts the information content of the other bands in the cluster. The more mutual information two random variables share, the better the prediction of one variable about the other one.

The selected representatives will constitute a subset of bands used in representation of the whole original set.

Extending Texems to Hyperspectral data

According to the texem model [9], an image is assumed to be generated by superposition of a small number of image patches of various sizes with certain added variations. In order to extract the texems from a sample image, it is broken down into a set Z of overlapping patches. If just one band is considered (grey-level texems), a texem is defined by a mean μ and a variance ω. However, for colour texems, pixels are assumed to be statistically independent in each texem with Gaussian distribution at each pixel position in the texem. Assuming K texems in the image, $K \ll \|Z\|$, patch i in set Z can be generated from a texem and formulated as a joint probability assuming neighbouring pixels are statistically conditionally independent:

$$p(Z_i|\theta_k) = p(Z_i|\mu_k, \omega_k) = \prod_{j \in S} N(Z_{j,i}; \mu_{j,k}, \omega_{j,k}) \tag{5}$$

where θ_k denotes the μ_k and ω_k parameters for the kth texem. $N(\cdot)$ is a Gaussian distribution over $Z_{j,i}$, S is the patch pixel grid, $\mu_{j,k}$ and $\omega_{j,k}$ denote mean and variance at the jth pixel position in the kth texem. The mixture model is given by:

$$p(Z_i|\Theta) = \sum_{k=1}^{K} p(Z_i|\theta_k)\alpha_k \tag{6}$$

where $\Theta = (\alpha_1, \ldots, \alpha_K, \theta_1, \ldots, \theta_K)$, and α_k is the *a priori* probability of the kth texem. The Expectation-Maximisation algorithm is used to estimate the model parameters. The E-step involves a soft-assignment of each patch Z_i to texems using Bayes' rule. The M-step finishes each iteration updating the parameters. These iterations are repeated until the estimations stabilise or a pre-specified threshold fulfils. Assignations of patches to texems are performed in a three-dimensional space.

An extension of the texem model to multispectral images has been developed. This extension consists of, instead of pixels $Z_{ij} = (R_{ij}, G_{ij}, B_{ij})$ being defined in a three-dimensional colour model, now pixels Z_{ij} are defined in a Q-dimensional space $Z_{ij} = (b^1_{ij}, b^2_{ij}, \ldots, b^Q_{ij})$, being b^k_{ij} the value of the pixel Z_{ij} in the spectral band k. Therefore, normal distributions are now defined in a multispectral space of Q bands[1].

Texem model has been successfully applied as an image segmentation technique [8] and the main objective of segmenting hyperspectral data requires the immediate extension of the colour texem model to multi-band spaces. However, without any previous dimensionality reduction stage, this extension would involve working with hundreds of bands and the algorithm would not be practicable at all. Thus, an input data reduction to a reasonable set of bands should be performed before applying the multi-band extension of the texem algorithm for segmenting hyperspectral images.

Experiments and Results

Three databases have been used to carry out this experimental part (see [5] for a detailed description):

1. Table 1, 1*st* column: **92AV3C** database, 145 X 145 pixels, 220 bands and 17 classes. As described in [4], several

[1] Note that, as the previous section describes, we had reduced the original dimensionality from L to Q bands.

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 signal-noise ratio bands.

2. Table 1, 2*nd* column: **HyMap** database, 700 X 670 pixels, 128 bands and 7 classes. In this case, 126 bands were used, discarding also the lower signal-noise ratio bands.
3. Table 1, 3*rd* column: **VIS** database is an example of the application of hyperspectral imaging to other purposes different from remote sensing. It contains images of orange fruits with several defects (rot, trip, overripe and scratch) in the visible (676 X 516 pixels, 33 bands). No band was discarded in this case.

In order to test the performance of the described process, several subsets of bands have been obtained from the input data by means of the feature selection method. The segmentation algorithm has been applied to each reduced representation of the original set to properly illustrate the results obtained. Thus, in table 1, the first row shows a RGB composition of the input images, the second row shows the results when the reduced subset (RSUB) is made up of four spectral bands and the segmentation algorithm looks for four different regions (REGS), the third row shows the results for RSUB=8 and REGS=6 and, finally, the fourth row shows the results for RSUB=8 and REGS=12. All the regions shown in the results are painted with random colours.

It is worth mentioning that the *WaLuMI* method with the previously described databases were used in [5]. This method reached satisfactory classification rates with four bands and it was consolidated on the plate of the classification curve with eight bands. That is the reason for choosing RSUB=4 and RSUB=8 in the feature selection algorithm.

A $n \times n$ grid is used in the texem model when the Gaussian distributions are calculated. We use a 1×1 grid in order to obtain well-outlined boundaries. However, although we manage accurate edges, the segmentation algorithm seems to be slightly noise-influenced. Better intra-region homogeneities without an important loss of accuracy in the boundaries should be achieved if the grid becomes a bit larger in size.

Conclusions

An unsupervised segmentation procedure for hyperspectral data based on the texem model has been presented. Before segmenting, a selection of the most relevant bands by means of information theoretical-based measures has been done in order to achieve a practical procedure.

The algorithm gives quite satisfactory results, managing accurate edges. However, it seems to be slightly noise-influenced due to the grid size used. This drawback is not very important if the number of input bands and texems is small. Nevetheless, it becomes worse when any of these parametres increases. Future work will take into account this question, producing new segmentation results with larger number of texems and/or bands by means of a bit larger grid.

Due to the fact that grey-level texems are defined by a mean and a variance (in contrast to colour texems where pixels are assumed to be statistically independent), as a future work we are studying the possibility of considering dependencies among the pixels in a texem because, from our point of view, it would obtain a better representation of the image textures.

Acknowledgements

We would like to deeply thank Dr Majid Mirmehdi for their inestimable help with the hyperspectral extension of the texem-based segmentation.

This work was supported by Spanish Ministry of Science and Education under Projects ESP2005-07724-C05-05, CSD2007-00018, PET2005-0643 and HP2005-0095.

References

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991.

[2] Raquel Dosil, Xosé R. Fdez-Vidal, and Xosé M. Pardo. Dissimilarity measures for visual pattern partitioning. *LNCS*, (3523):287–294, 2005.

[3] A. K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice-Hall, 1988.

[4] David A. Landgrebe. *Signal theory methods in multispectral remote sensing*. Wiley, 2003.

[5] Adolfo Martinez-Uso, Filiberto Pla, José M. Sotoca, and Pedro Garcia-Sevilla. Clustering-based hyperspectral band selection using information measures. *IEEE Trans on GRS*, 45(12):4158–4171, December 2007.

[6] John H. Ward. Hierarchical grouping to optimize an objective function. *American Statistical Association*, 58(301):236–244, 1963.

[7] Andrew Webb. *Statistical Pattern Recognition*. Wiley, 2nd edition edition, 2002.

[8] Xianghua Xie and Majid Mirmehdi. Colour image segmentation using texems. *Annals of the BMVA*, 1(6):1–10, January 2007.

[9] Xianghua Xie and Majid Mirmehdi. Texems: Texture exemplars for defect detection on random textured surfaces. *IEEE Trans on PAMI*, 29(8):1454–1464, August 2007.

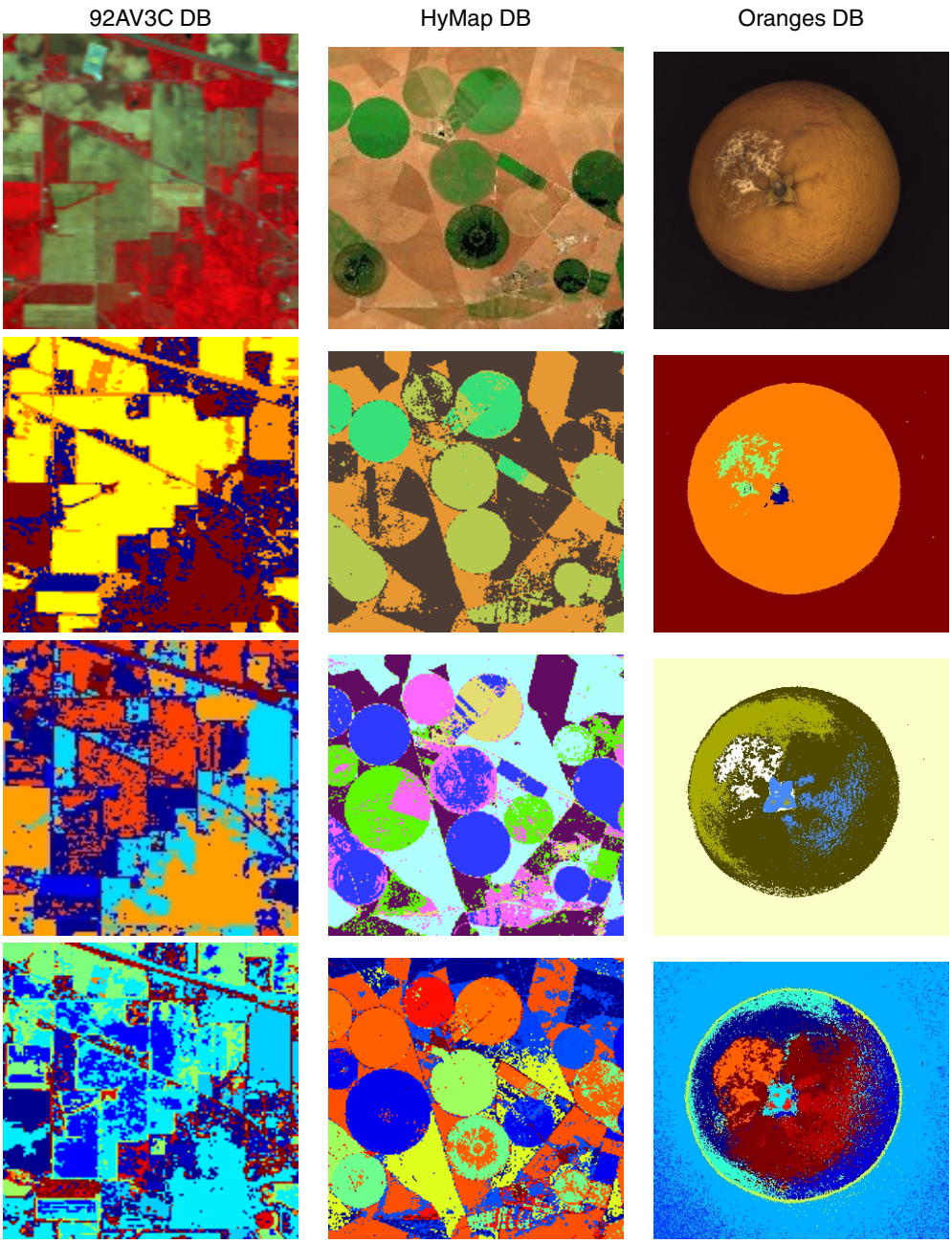
Authors biography

Adolfo Martínez-Usó received the B.Sc. degree in computer science from the Jaume I University of Castellón, Spain. He is currently pursuing the Ph.D. degree in computer science engineering at the Jaume I University. His current research interests are hyperspectral image analysis, particularly, band selection and image segmentation.

Prof. F. Pla is a Full Professor at the Departament de Llenguatges i Sistemes Informatics of the University Jaume I, Castellón, Spain. He received a degree and a Ph.D. in Physics from the University of Valencia in 1989 and 1993, respectively. His current research interests are colour and spectral image analysis, visual motion analysis, active vision and pattern recognition techniques applied to image processing. He is member of the International Association for Pattern Recognition.

Pedro García-Sevilla received the B.Sc. degree in 1990 from the University of Castilla-La Mancha, the M.Sc. degree in 1992 from the University of Malaga, and the Ph.D. degree in 1999 from the Jaume I University (Castellón), all in computer science. Currently, he is with the Computer Vision Group at the Jaume I University as a lecturer. His main research interests are in image processing and pattern recognition, particularly, in texture analysis, hyper-spectral imaging, industrial inspection, and medical imaging.

©2008 Society for Imaging Science and Technology



First column for AVIRIS (92AV3C), second for HyMap spectrometer and third for an orange image from VIS collection. First row shows RGB compositions of each database. Results on second, third and fourth row correspond to RSUB=4/REGS=4, RSUB=8/REGS=6 and RSUB=8/REGS=12 respectively.