Spectral Indexing for Hyperspectral Image CBIR

Jose Orlando Maldonado, David Vicente, Miguel A. Veganzones, Manuel Graña

Computational Intelligence Group, Bask Country University (UPV/EHU)

Abstract. In this paper we deal with the problem of indexing hyperspectral images as an special case of Content Based Image Retrieval (CBIR) systems. We define a similarity measure between hyperspectral images based on the image endmembers, which are induced from the images. For this induction we use Associative Morphological Memories (AMM).

1 Introduction

The growth in multimedia information, especially images, is driving the development of the field of Content Based Image Retrieval (CBIR) [1]. Remote sensing and earth observation are a big source of images that may benefit from automated exploration (data mining) and retrieval systems. In CBIR systems, the images stored in the database are labeled by feature vectors, which are extracted from the images by means of computer vision and digital image processing techniques. In CBIR systems, the query to a database is specified by an image. The query's feature vector is computed and the closest items in the database, according to a similarity metric or distance defined in feature space, are returned as the answers to the query. This is the low level, semantic free, definition of the CBIR systems, that does not take into account the semantic gap between the user expectations and the system response. Recent approaches [6] apply relevance interactions with the user to build up semantic models upon to drive the information retrieval and data mining processes to the user expectations. Our work, however, is at the low level, without relevance iteration mechanisms.

In hyperspectral images each pixel contains a fine sampling of the visible and near infrared spectrum, represented by a high dimensional vector. New space missions, like Hyperion, include hyperspectral sensor, so it can be expected a growing need for the maintenance of large collections of hyperspectral images, and for the automated search within these collections. The attempts to define CBIR strategies for them are scarce and partial. In [3] the authors use the spectral mean and variance, as well as a texture feature vector, to characterize hyperspectral image tiles. The approach searches for specific phenomena in the images (hurricanes, fires, etc), using an interactive relevance strategy that allows the user to refine the search. In [6] some of the first PCA and texture features from the first PCA image are used as feature vectors.

We propose the characterization of the hyperspectral images by their so-called endmembers. Endmembers are a set of spectra that are assumed as vertices of a convex hull covering the image pixel points in the high dimensional

spectral space. Endmembers may be defined by the domain experts (geologists, biologists, etc.) selecting them from available spectral libraries, or induced from the hyperspectral image data using machine learning techniques. In [4,5] we have proposed an automated procedure that induces the set of endmembers from the image, using AMM to detect the morphological independence property, which is a necessary condition for a collection of endmembers. The goal in [4,5] was to obtain a method for unsupervised hyperspectral image segmentation. There the abundance images were of interest. In this paper, the goal is to obtain a characterization of the image that can be useful for CBIR. The collection of endmembers serves as a summary of the spectral information in the image.

2 Spectral mixing and endmember extraction

Linear mixing models assume the knowledge of a set of endmembers $S = [s_1, s_2, ..., s_n]$, where each $s_i \in R^d$is a d-dimensional vector. Then, one pixel of a hyperspectral image can be expressed as $f(x, y) = S \cdot a(x, y) + \eta(x, y)$, where $\eta(x, y)$ is the independent additive noise component and $a(x, y)$ is the n-dimensional vector of endmember fractional abundance in the pixel. In other words, $a(x, y)$ are the convex coordinates of the pixel relative to the convex hull defined by the vertices in S. In [4,5] we were interested in $a(x, y)$ as unsupervised segmentations of the hyperspectral image. Here we consider that the set of endmembers S may be *per se* a good characterization of the hyperspectral image S, if it has been obtained from it. The method proposed in [4,5] to obtain the set of endmembers S is based on the notion of morphological independence [8]. Given a vector set $X = \{x_1, x_2, ..., x_m\}$, a new vector y is morphologically independent in the erosive sense from X if $\nexists x \in X | y \leq x$ and it is morphologically independent in the dilative sense from X if $\nexists x \in X | y \geq x$. The partial order defined over the vectors is the one induced from the order of their components: $y \leq x \Leftrightarrow \forall i, y_i \leq x_i$. The vector set X is said to be morphologically independent when all the vectors in the set are independent of the remaining ones in either sense. A set of morphological independent vectors defines a high dimensional box.

In [4,5] we propose the use of AMMs for the detection within a hyperspectral image of a set of morphological independent spectra whose corresponding high dimensional box may be taken as a good approximation to the convex hull of the image data. The algorithm is shown in Algorithm 1. These spectra are the induced endmembers that will be used to characterize the image. In brief the proposed method consists in the following:

1. A seed pixel is taken as the initial set of endmembers.
2. Erosive and dilative AMMs are built from the current set of endmembers.
3. Each pixel is examined as a candidate endmember testing the response of the erosive and dilative AMMs to it.
4. Morphologically independent pixels are added to the set of endmembers. The process takes into account the variance of the spectra at each band to enhance the robust detection of endmembers, adding and subtracting the

Algorithm 1 The induction of the endmembers from the image in one pass over the image

1. Compute the zero mean image: $f^c(i,j) = f(i,j) - \mu; i = 1,...,n; j = 1,...,m$
2. Initialize the set of endmembers $E = \{e_1\}$ with a pixel spectra randomly picked from the image. Initialize the set of morphologically independent binary signatures $X = \{x_1\} = \{e_k^1 > 0; k = 1,...,d\}$
3. Construct the AMM's based on the morphologically independent binary signatures: $M_{XX}$and W_{XX}.
4. For each pixel $f^c(i,j)$:
   (a) Compute the vector of the signs of the Gaussian noise corrections $f^+(i,j) = (f^c(i,j) + \alpha\sigma > 0)$ and $f^-(i,j) = (f^c(i,j) - \alpha\sigma > 0)$
   (b) Compute $y^+ = M_{XX} \wedge f^+(i,j)$
   (c) Compute $y^- = W_{XX} \vee f^-(i,j)$
   (d) if $y^+ \notin X$ or $y^- \notin X$ then $f^c(i,j)$ is a new endmember to be added to E, go to spet 3 and resume the exploration of the image
   (e) if $y^+ \notin X$ and $f^c(i,j) > e_{y^+}$ the pixel spectral signature is more extreme than the stored endmember, then substitute e_{y^+} with $f^c(i,j)$
   (f) if $y^- \notin X$ and $f^c(i,j) < e_{y^-}$ the pixel spectral signature is more extreme than the stored endmember, then substitute e_{y^-} with $f^c(i,j)$

per band standard deviation multiplied by a gain parameter to the pixel spectrum before performing the tests with the dilative and erosive AMMs, respectively. This gain parameter is set by default to 2. The process goes one time over the image data. If the region is very homogeneous, the process may stop without adding any new endmember, besides the seed pixel taken. Then, the gain parameter is reduced and the process is repeated until the number of endmembers is 2 or more.

3 Distance between images

Let it be $S_k = s_1^k, s_2^k, ..., s_{p_k}^k$ the set of endmembers, obtained as described before from the k-th image $f_k(x,y)$ in the database, where p_k is the number of endmembers detected in this image. Given two images $f_k(x,y)$ and $f_l(x,y)$, we compute the following matrix whose elements are the Euclidean distances between the endmembers of each image:

$$D_{k,l} = d_{i,j}; i = 1,...,p_k; j = 1,...,p_l \tag{1}$$

where $d_{i,j} = |s_i^k - s_j^l|$.

We compute the vectors of the minimal values by rows and columns, $m_k = \left[m_i^k = \min_j\{d_{i,j}\}\right]$ and $m_k = \left[m_i^l = \min_i\{d_{i,j}\}\right]$ respectively. Then the similarity between the images is given by the following expression:

$$d(f_k, f_l) = (|m_k| + |m_l|)(|n_k - n_l| + 1) \tag{2}$$

4 Discussion of distance properties

The endmember induction procedure may give different number of endmembers and endmember features for two hyperspectral images. The similarity measure is a composition of two asymmetrical views: each vector of minimal distances measures how close are the endmembers of one image to some endmember of the other image. Suppose that all the endmembers S_k of an image are close to a subset of the endmembers S_l of the other image. Then the vector of minimal distances m_k will be very small, not taking into account the unlike endmembers in the second image. However, the vector of minimal distances m_l will be larger than m_k because it will take into account the distances of endmembers in S_l which are unlike to those in S_k. Thus the similarity measure can cope with the asymmetry of the situation. It avoids the combinatorial problem of trying to decide which endmembers can be matched and what to do in case that the number of endmembers is different from one image to the other. The difference in the number of endmembers is introduced as an amplifying factor. The measure is independent of image size and, as the endmember induction algorithm is very fast, it can be computed in acceptable time. Also the endmember set poses no storage problem.

Our approach does not use spatial features, such as the textures in [3], but the endmembers give a rich characterization of the spectral content of the image. A further work on our approach may be the study of spatial features computed on the abundance images produced by the spectral unmixing, solving the question of band selection or dimension reduction prior to spatial feature computation.

5 Experimental results on simulated data

The hyperspectral images used for the experimental results are generated as linear mixtures of a set of spectra (the ground truth endmembers) with synthesized abundance images. The ground truth endmembers were randomly selected from a subset of the USGS spectral libraries corresponding to the AVIRIS flights. Figures 1 and 2 show some spectra used in the 2 and 5 endmembers images.

The synthetic ground truth abundance images were generated in a two step procedure, first we simulate each as an gaussian random field with Matern correlation function of parameters θ_1, θ_2 varying between 2 and 20. We applied the procedures proposed by [7] for the efficient generation of big domain gaussian random fields. Second, to ensure that there are regions of almost pure endmembers we selected for each pixel the abundance coefficient with the greater value and we normalize the remaining to ensure that the abundance coefficients in this pixel sum up to one. It can be appreciated on the abundance images that each endmember has several regions of almost pure pixels, viewed as brighter regions in the images. Image size is 256x256 pixels of 224 spectral bands each. We have generated collections of 100 images with 2 to 5 ground truth endmember/abundances, for total number of 400 images.

The experiment performed on these images consists on the following steps:

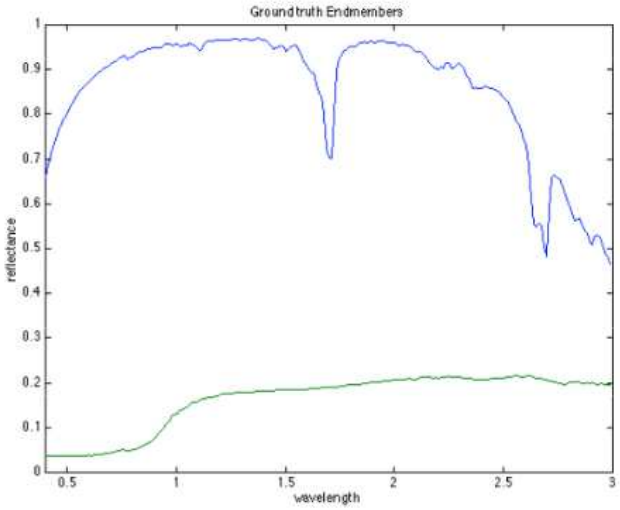

Fig. 1. Ground truth endmembers extracted from the USGS library used in one distance of a 2-endmembers synthetic image

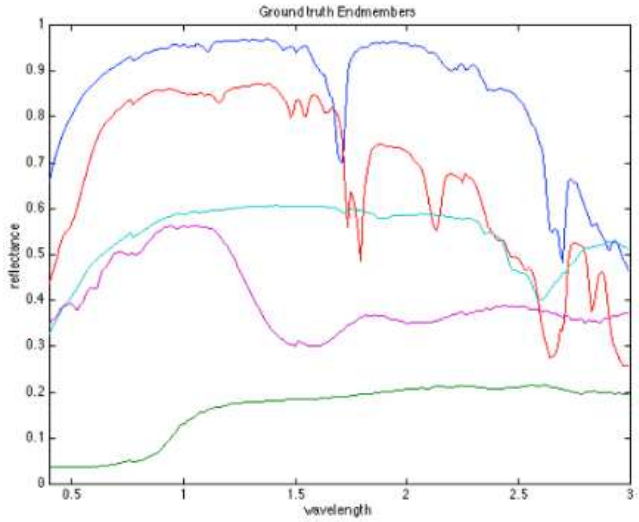

Fig. 2. Ground truth endmembers extracted from the USGS library used in one distance of a 5-endmembers synthetic image

1. Compute the distances between the images in the database using the ground truth endmembers. The distances are computed between images with the same number of ground truth endmembers, and with all the remaining images.
2. Extract the endmembers from the images using the approach described in section 2.
3. Compute the distances between the images in the database using the morphologically independent induced endmembers. The distances are computed between images with the same number of ground truth endmembers, and with all the remaining images.
4. We consider the R closer images to each image in each case (ground truth and morphologically independent induced endmembers) as the responses to a potential query represented by the image.
5. The images that appear in both responses (based on the ground truth and the morphologically independent induced endmembers) are considered as relevant images, or correct responses.

In table 1 we present the results from the experiment with the 400 images, in terms of the average number of correct responses. First row presents the results when we pool together all the images, regardless of the number of ground truth endmembers. The next rows present the results when we only try to search in the subcollection of images with the same number of endmembers as the query image. Each row corresponds to a different number of images in the response to the query. The value of the noise gain was set to $\alpha = 0.5$.

	R=1	R=3	R=5	R=10
All images	0.94	1.21	1.61	2.96
2 Endmembers	0.81	1.55	2.27	4.67
3 Endmembers	0.98	1.44	2.21	4.96
4 Endmembers	0.99	1.53	2.36	4.81
5 Endmembers	1.00	1.57	2.37	4.74

Table 1. Average number of relevant images per query

In Table 1 it can be appreciated that the consideration of all the images as responses to the query introduces some confusion and reduce the average number of correct images obtained in the query. This effect can be due to the fact that the morphological independence algorithm can find a number of endmembers different from the ground truth making it possible for the image to match with images outside its natural collection of images. Then images with different ground truth numbers of endmembers may become similar enough to enter in their respective response sets.

When we restrict the search to the collections with identical number of ground truth endmembers, all the results improve, except when R=1. We have that near 50% of the responses are significative when R>1. The case R=1 can be

interpreted as the probability of obtaining the closest image in the database according to the distance defined in section 3, or the probability of success. It can be seen that it is very high, close to 1 for all search instances, except for the case of 2 ground truth endmembers.

6 Conclussions and further work

We have proposed an approach to CBIR in homogeneous databases of hyperspectral images based on the collection of endmembers induced by an algorithms that searches for morphologically independent vectors. We have performed an experiment of exhaustive search on a collection of simulated hyperespectral images. The results are encouraging: almost 100% success in providing the closest image in terms of the ground truth endmembers.

However these results only confirm the ability of the AMM based algorithm to find a good approximation to the ground truth endmembers. We have still to test the effect of additive noise on the results, and maybe to perform comparison with other endmember extraction algorithms. Previous experiments comparing the AMM based algorithm with Independent Component Analysis (ICA) [9,10] and other algorithms have been favourable to the AMM algorithm [4,5].

It is possible to define other distances based on the endmembers extracted by the AMM (or any alternative algorithm). For example, the Euclidean distance between individual endmembers may be substituted by max/min distances. The whole set of endmbembers may be evaluated according to the Haussdorf distance. There are also diverse ways to evaluate the diverse number of endmembers found in the images, introducing penalization terms.

We have not included yet any spatial element in the distance. One of the obvious ways to do it is to compute the correlation between the abundance images, matched according to the relative match of their corresponding endmembers. The distance between images would include a spatial correlation term, with an specific weight that must be tuned. If the images are homogeneous in their definition of the image domain, this correlation can be computed in the straightforward manner. However if they are of different sizes and/or the original capture can not be assumed to be exactly registered, there is a combinatorial problem to be solved, that of finding the optimal image domain transformation (translation, rotation, scaling) to register the two images being spatially matched. The cost of this operation may be prohibitive in the context of large database indexing. Additional information, such as image coordinates, will serve to reduce this computational problem.

Assuming that the image collection is homogenous we have avoided the problem of matching images from different sensors or images with different frequency bands missing due to image preprocessing and noise. This is a main difficulty for our algorithm, because we can not deal at present with missing data. The only available way is to interpolate assuming local continuity of the observed spectra. However, this solution may be hardly accepted by the remote sensing user (geologist, agriculture engineer, biologist, etc).

There is a strong trend to introduce the human factor in the index construction loop. Some semantic information can be introduced in the system that way. We have yet to devise a methodology to introduce the interaction with the user when dealing with spectral information.

Acknowledgements

The Spanish Ministerio de Educación y Ciencia supports this work through grant VIMS-2003-20088-c04-04.

References

[1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain. Content-based image retrieval at the end of the early years. IEEE Trans. Pat. Anal. Mach. Intel., 2000, 22 (12), pp.1349-1380

[2] D. A. Landgrebe. Signal Theory Methods in Multispectral Remote Sensing. John Wiley & Sons, Hoboken, NJ, 2003

[3] I. E. Alber, Ziyou Xiong, N. Yeager, M. Farber, W. M. Pottenger. Fast retrieval of multi and hyperspectral images using relevance feedback. Proc. Geosci. Rem. Sens. Symp., 2001. IGARSS '01, vol.3, pp. 1149-1151

[4] M. Graña, J. Gallego, C. Hernandez. Further results on AMM for endmember induction. Proc. IEEE Workshop on Adv. Tech. Ana. Remotely Sensed Data, Washington D.C., Oct. 2003, pp. 237-243

[5] M. Graña, J. Gallego. Associative morphological memories for endmember induction. Proc. Geosci. Rem. Sens. Symp., IGARSS '03. Tolouse, Jul. 2003, vol.6, pp. 3757-3759

[6] M. Datcu, K. Seidel. Human-centered concepts for exploration and understanding of Earth observation images. IEEE Trans. Geoscience and Remote Sensing, 43(3):601-609, Mar 2005

[7] B. Kozintsev. Computations With Gaussian Random Fields. PhD Thesis, ISR99-3 University of Maryland (1999)

[8] G. X. Ritter, G. Urcid, L. Iancu. Reconstruction of patterns from moisy inputs using morphological associative memories. J. Math. Imag. (2003)

[9] A. Hyvarynën, J. Karhunen, E. Oja. Independent Component Analysis. John Wiley & Sons, New York, 2001

[10] A. Hyvarynën, E. Oja. A fast fixed-point algorithm for independent component analysis. Neural Comp. 9:1483-1492, 1999