5

On Content-Based Image Retrieval Systems For Hyperspectral Remote Sensing Images

Miguel A. Veganzones, José Orlando Maldonado, Manuel Graña

Computational Intelligence Group, UPV/EHU,
www.ehu.es/ccwintco

Summary. This chapter includes a review of some key elements of Content Based Image Retrieval Systems (CBIR). We place into this context the the works found in the literature regarding remote sensing images. Our own focus is on hyperspectral images. The approach we are pursuing is that of characterizing the spectral content of the image through the set of endmembers induced from it. We describe some ideas and numerical experiments for a system that would perform CBIR on hyperspectral remote sensing images . We propose a spectral image similarity to guide the search to answer the queries.

5.1 Introduction

The interest in fast and intelligent information retrieval systems over collections of remote sensing images is increasing as the volume of the available data grows exponentially. For instance a single in-orbit Synthetic Aperture Radar (SAR) sensor collects about $10 - 100$ Gbytes of data per day, giving about 10 Tbytes of imagery data per year. New sensing instruments can produce an information amount of one order of magnitude higher than this. Current retrieval systems offer to their users raw images, thematic maps and ancilliary data in response to very precise queries requiring a detailed knowledge of the structure of the stored information. However there is a growing need for more flexible ways to access the information based on the image computed intrinsic properties, i.e. texture and induced semantic contents. 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. We are specially interested in hyperspectral images and the ways to characterize them for CBIR tasks.

The fields of applicability are extensive, geoscience tools are demanded for disasters prevision, enviromental analysis, security. The investigation of new tools capable of predict, monitor and assist in preparation of strategies to minimize the damage produced by natural and human-made disasters, would contribute to a better quality of life and to the reduction of economical losses. The ability to perform automatic and intelligent searches is a key tool for these goals. The field of CBIR is already a mature area of research [33]. Some facts, like the Semantic Gap between low level representations and high level semantic world models, have been acknowledged leading to new design and research paradigms. A fruitful track of research and source of innovative approaches introduces user interaction to obtain some retrieval feedback that allows the building and refinement of user world models [3] [7] [8] [13] [14] [15] [16] [18] [19] [20] [24] [25] [31] [34] [36].

The attempts to define CBIR strategies for remote sensing hyperspectral images are scarce and partial. In [1] 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 [7] some of the texture features from the first spectral 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 [10,11] 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 [10,11] 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.

In section 5.2 we introduce some ideas about the general structure of CBIR systems, referring them to the systems found in the literature of remote sensing images. In section 5.3 we present some ideas about the evaluation of CBIR systems, discussing how systems found in the literature have been validated. In section 5.4 we discuss our own ideas and proposal for a CBIR system based on the spectral characterization of the images.

5.2 CBIR Systems Anatomy

A CBIR system is a query resolving system over image collections that use the information inherently contained in the image. The CBIR system has to be able to extract quantitative features from the images that allow the system to index the image collection and to compute a distance between images. The simplest typical use of CBIR systems is described in figure 5.1. The user interacts with the system by a querying interface, usually a web page, where the query is defined and sent to the CBIR engine. In the process described by figure 5.1 the query is represented by an image provided by the users, asking the CBIR system for a list of the most similar images in the database. To resolve the query, the CBIR engine computes the image features which correspond to a point in the metric space defined by the system. Each image in database has a representative in this metric space so a distance to the query image could be computed for each image, using a similarity (or disimilarity) function. This produces a list ordered by similarity (or disimilarity) to the query image, that is presented to the user as response. In real world applications, computing the distance to each and every image in the database has an unbereable cost. Therefore, the database is indexed grouping the images into clusters. Then heuristic algorithms are used to retrieve the group of most similar images. The retrieved images are inspected by the user, selecting the ones that are relevant to his target search and those that are irrelevant. The sets of relevant and irrelevant images are used to redefine the query so the features quantifying the query information fits better the target in the mind of the user. The search is iterated until the set of retrieved images satisfies the user's expectatives.

From the previous description, it is natural to say that a CBIR system is composed of three subsystems: the query definition, the feature extraction and the image retrieval. Query definition is the proccess by which the user defines the query. It could be a graphical design interface or a web page. It can be as simple as selecting an image or it can be more sofisticated, i.e. the user can select regions of interest on the query image and interact with the retrieval proccess in a more expert mode. The feature extraction is the proccess of quantifying image information as a collection of low level characteristics, denoted as the feature vector. Image Retrieval encompasses the methods that use the feature vectors to select from the database the images that satisfy the query, in the sense of being the most similar ones. To prevent the Semantic Gap, the image retrieval may involve the interaction with the user to improve the query in any way. We review some ideas about CBIR systems, while omitting others, like the visualization function [33], which are not relevant to the work in this chapter.

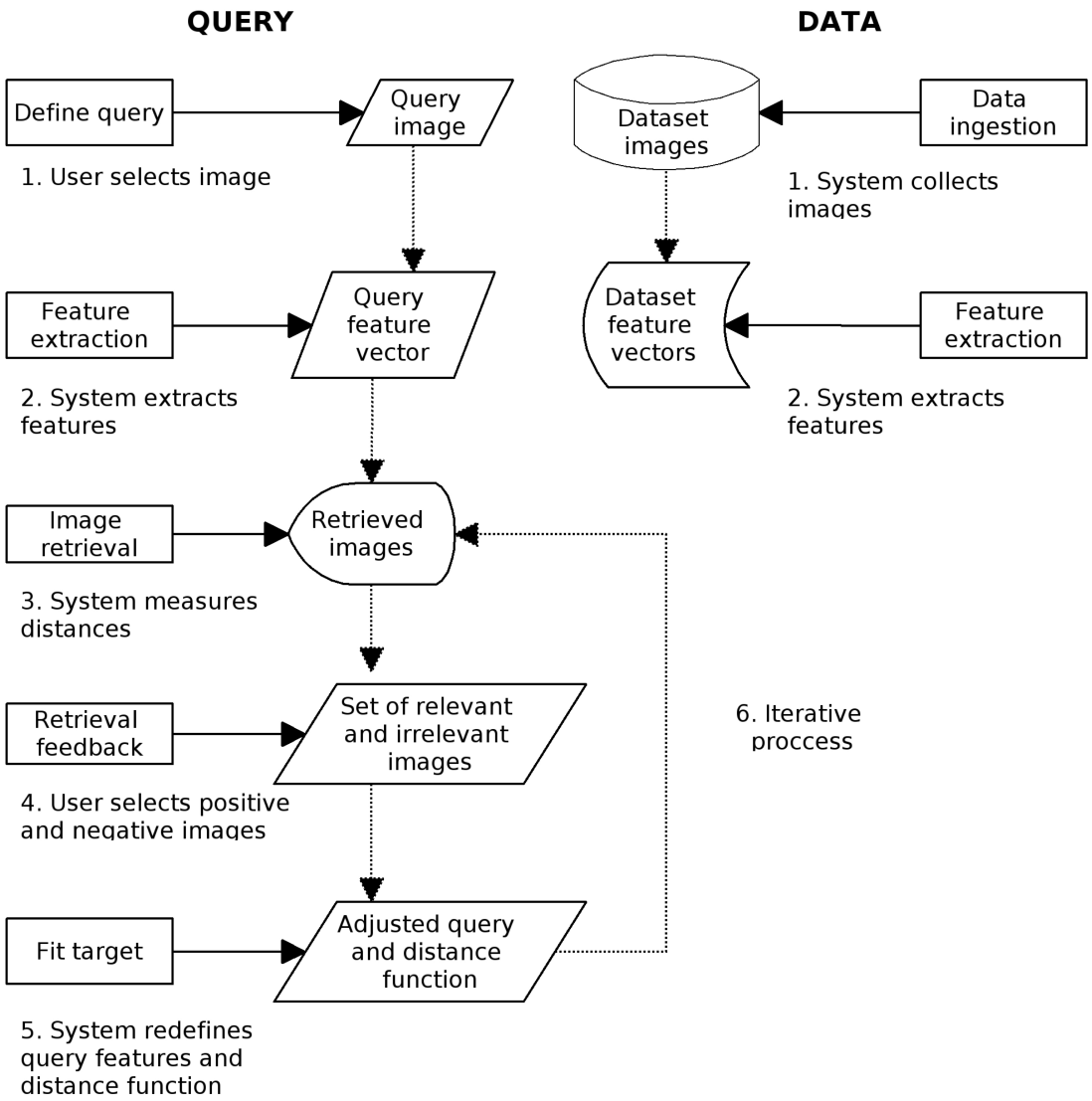

Figure 5.1. Typical use of a CBIR system

5.2.1 Query definition and interaction

In image information retrieval defining the query may be a critical design issue. The most common query specification method in CBIR search engines is the Query By Example(s), where the query is represented by an example image (or a set of images) related to the search target. The *Page Zero Problem* [38] arises when the user has not any image to start querying. When the user is only interested in an specific region of the images, he may specify the query by selecting a region of interest (ROI), which is known as the Query By Region(s). In the Query By Sketch, the user can draw an sketch of the desired target to specify the query. Conventional Query By Text could be usefull in situations where relevant information about images is collected as textual metadata, as for example, geografical coordinates, time of adquisition or sensor type. Text may be used as a complement to the visual query, or may solve the Page Zero Problem.

Formally, the Query Space [33] is the goal dependent 4-Tuple $\{D_Q, F_Q, S_Q, Z_Q\}$ where:

- D_Q is the subset of active images from dataset collection D within which the retrieval system is going to perform the search. For example, the user could select a subset of images belonging to an specific sensor in a concrete date of adquisition instead of querying the full image collection.
- F_Q is a selection of the quantitative image features $F_Q \subset F$. Usually the user doesn't have the knowledge to select the most relevant features for resolving the query, and the system automatically decides what features are going to be used.
- S_Q is the similarity function. This similarity function should be parameterized to allow adaptation to diverse goals and data sets. Very often the user is not capable of decide which similarity function is the most appropiate for the formulated query and S_Q is selected automatically by the system.
- Z_Q is a collection labels to capture the goal dependent semantic concepts. For example, if the user wants to retrieve images containing rivers, Z_Q could be defined as $Z_Q = \{'water', 'river'\}$. Conventional CBIR systems have not the capacity of incorporate semantic concepts to the query resolving system, so they use exclusively the Feature Vector information as the criteria for searching the query's answer.

Queries can be specified either in an *exact* or *approximate* way. The exact specification requires that the answer $A(q) \subseteq I_Q$ to the query contains the images satisfying certain criteria, which can be given by some relational equations or propositions. The approximate specification requires that the answer $A(q)$ to the query provides the images ordered relative to the query satisfaction. Exact specification may consist of spatial, image or group predicates. The approximate specification may consist of spatial relations (skeetches) or image examples.

The interaction of the user with the system can be formalized as a sequence of Query Spaces $\{Q^1, \ldots, Q^n\}$. The transition between those query spaces is the effect of the user interaction, which can take several forms, from indicating the relevance of the answer images up to modifying the distance between images.

Regarding the kind of search, the following types can be distinguished [33]:

1. Search by category: the user wants to retrieve images belonging the same class.
2. Search by target: the user wants to retrieve images containing an object.
3. Search by association: user has not an specific target in mind and its only aim is to browse through the database.

In the context of remote sensing image databases, the proposed systems cover some of these possibilities. In [1, 20] a basic relevance approach is applied to multispectral images. In [18] the topology of the feature space is used for relevance feedback. In [31, 6, 7] a semantic model of the cover type desired by the user is constructed as a bayesian network from the interaction with the user.

5.2.2 Feature extraction

The feature extraction process produces the low dimensional (relative to image size) quantitative description of the image which allows the computation of similarity measures, the definition of an ordering of the images and their indexing for the search processess. For remote sensing images, some approaches use the spectral information, while other emphasize the spatial features. Most of the works reported in the literature have been performed on Landsat images, of similar sensor images, whose number of bands is small compared to modern hyperspectral sensors.

In [12] the image features are computed as multispectral distribution invariants from a physically-based image formation model. The result of dimensional reduction procedures, such as the Orthogonal Subspace Projection [30] can be also used as fatures for classification and search. Eigenviews are proposed by [29] for object recognition. A collection of image features is computed offline in the system structure proposed in [6, 5]. These features are the parameters of Gauss Markov Random Fields modelling texture and spatial fatures in the case of optical and SAR images. In [31] a similar process is proposed to build up a vocabulary of image features, previous to the interactive construction of a Bayesian semantic model of the user target. This vocabulary is the result of a clustering performed over the parameter data from the texture model estimation.

Unmixing algorithms

The hyperspectral imagery could be seen as mixing model where an hyperspectral image is the result of the linear combination of the pure spectral signature of ground components, named endmembers, with abundance matrices. Let $E = [e_1, \ldots, e_n]$ be the pure endmember signatures where each $e_i \in R^d$ is a d-dimensional vector. Then, the hyperspectral signature $h(x, y)$ at the pixel with coordinates (x, y) is defined by the expression:

$$h(x,y) = \sum_{i=1}^{n} e_i \Phi_i(x,y) + \eta_i(x,y) \tag{5.1}$$

where $\Phi(x, y)$ is the n-dimensional vector of fractional abundance at pixel (x, y) and $\eta(x, y)$ is the independent additive noise component. There are two constraints in the equation 5.1, the abundance non-negatice constraint 5.2 and the abundance sum-to-one constraint 5.3, respectivily defined as

$$\Phi_i(x,y) \geq 0, \text{ for all } 1 \leq i \leq n \tag{5.2}$$

$$\sum_{i=1}^{n} \Phi_i = 1 \tag{5.3}$$

These restrictions require careful numerical methods for the computation of the fractional abundances [32] when the endmembers are known. The image spectral signatures could be a good characterization of the image but, additionaly, the abundance matrices could be used as spatial information about the image. The idea of using the spectral unmixing fractional abundances as a kind of feature extraction for classification purposes was introduced in [9]. The feature vector for each pixel (x, y) in the image is the vector $[\Phi_1(x, y), ..., \Phi_n(x, y)]$ which is the set convex coordinates of the pixel relative to the enmembers. We extend this idea to the design of CBIR systems for hyperspectral images. In fact, we will assume the set of endmembers E as the image features. The key problem then is the definition of the set of endmember spectral signatures. A library of known pure ground signatures or laboratory samples could be used. However this poses several problems, such as the illumination invariance, the difference in sensor intrinsic parameters and the *a priori* knowledge about the material composition of the scene. There are diverse approaches in the literature for the estimation of endmembers from the image data. Early approaches [37] searched for the minimum simplex covering the image data. A method based on morphological operators is proposed in [26, 27]. In [2] an statistical approach is proposed. In [10, 11] we proposed a method based on morphological independence and the use of Associative Morphological Memories (AMM). Recently, [4] proposed a fast Pixel Purity Index procedure. An approach based on ICA is proposed in [23].

5.2.3 The similarity function

The retrieval process provides a subset of images answering the query on the basis of the similarity between the images, computed over the feature vector extracted from each image. There are two basic strategies to compute the answer A_Q, either an ε-similarity or an KNN-similarity retrieval. In ε-similarity retrieval, the system selects only those images in D_Q whose similarity to the query criteria falls below a given threshold ε. In KNN-similarity retrieval, the system ranks D_Q according to the similarity to the query given a similarity function, S_Q, and retrieves the first K images. These retrieval strategies are similar to the exact and approximate searches described above. There are two essentially different retrieval paradigms: the metric and the probabilistic.

Metric paradigm

Most of the retrieval systems found in the literature are based in the metric paradigm. The image information is quantified as low level feature vectors in a high dimensional metric space. $F : I \rightarrow \Re^r$. Therefore, an image is described as a point in a r-dimensional space. The distance in feature space corresponds to the image similarity function $S : IxI \rightarrow \Re^+$.

The equation 5.4 expresses the similarity function between two images $S(I_1, I_2)$ as a function of the distance $d(F_1, F_2)$ between their corresponding feature vectors. The function Ψ is a positive, monotonically nonincresing function.

$$S(I_1, I_2) = \Psi(d(F_1, F_2)) \tag{5.4}$$

The Euclidean distance (eq. 5.5) is the simplest and most used distance function. Other useful metrics are the Manhattan (eq. 5.6) and the Maximum (eq. 5.7) distances.

$$d_E(F_1, F_2) = \sqrt{\sum_{i=0}^{r-1}(F_{1i} - F_{2i})^2} \tag{5.5}$$

$$d_M(F_1, F_2) = \sum_{i=0}^{r-1} |F_{1i} - F_{2i}| \tag{5.6}$$

$$d_{Max}(F_1, F_2) = \max_i \{|F_{1i} - F_{2i}|\} \tag{5.7}$$

where F_{1i}, F_{2i} are the i-th feature of the respective feature vectors. When scale invariance is needed, the Correlation-Based distance (eq. 5.8) can be used.

$$d_{Corr}(F_1, F_2) = \frac{\sum_{i=0}^{r-1} F_{1i}F_{2i}}{\sqrt{\sum_{i=0}^{r-1} F_{1i}^2}\sqrt{\sum_{i=0}^{r-1} F_{2i}^2}} \tag{5.8}$$

In a relevance-guided iterative retrieval proccess [1, 8, 14, 15, 19, 20] the user feedback is specified through the identification of a set of relevant and irrelevant images in the answer set $A(q)$, aiming to better approach the target that the user has in mind. In the metric paradigm this goal can be reached from two different points of view. The first modifies the query definition while second one modifies the similarity function. Redefining the query is equivalent to move the query point across the feature vector space. One way of doing that is the following one: in each iteration the user identifies the relevant images and their relevance degree; then the query is redefined as the weighted centroid of the set of relevant images weighted by their relevance degree. Let η_j be the relevance degree of the j-th relevant image. Each component of the new feature vector corresponding to the redefined query, F_i', is given by the following equation:

$$F_i' = \frac{\sum_j \eta_j F_{ji}}{\sum_j \eta_j}$$

Another kind of query redefinition is based in both the relevant and irrelevant set of images selected by the user. In [31] the relevance feedback guides the estimation of a Bayesian Network parameters. Let F_{ji} be the i-th feature of the j-th image on relevant set. Let $\bar{F}_{ki}$ be the i-th feature of the k-th image on irrelevant set. Let N_F and $N_{\bar{F}}$ be the number of images in the relevant and irrelevant set respectively, and α, β, γ are suitable constants; the following equation defines the query movement along the feature space at each retrieval feedback iteration.

$$F_i' = \alpha F_i + \beta \left(\frac{1}{N_F} \sum_{j=0}^{N_F - 1} F_{ji} \right) + \gamma \left(\frac{1}{N_{\bar{F}}} \sum_{k=0}^{N_{\bar{F}} - 1} F_{ki} \right)$$

The second relevance feedback class of methods works with the similarity function, updating it at each relevance feedback iteration having into account the user preferences. Weighting the Euclidean distance so the user's information could be incorporated into the retrieval proccess can be achieved in two different ways, depending on whether there is any interdependency among the feature vector components or not. If the feature vector components are mutually independent, the relevance of each feature element is modelled by a diagonal matrix, $\Lambda = diag \cdot (\lambda_1, \lambda_2, \ldots, \lambda_r)$ where λ_i represents the individual relevance of the i-th feature component. The Weighted Euclidean Distance (WED) is defined by equation 5.9. If the feature components are not independent, we have a Generalized Euclidean Distance (GED), defined by equation 5.10, where W is a symetric matrix representing the mutual dependent of the feature components.

$$d_{WED}(F_1, F_2) = (F_1 - F_2)^T \Lambda (F_1 - F_2) \tag{5.9}$$

$$d_{GED}(F_1, F_2) = (F_1 - F_2)^T W (F_1 - F_2) \tag{5.10}$$

Both WED and GED are instances of the Mahalanobis distance. The values of Λ's diagonal can be estimated as $\lambda_i = \frac{1}{\sigma_i}$, where σ_i is the feature vector component standard deviation computed over the relevant set of images selected by the user in the answer to the query. If the standard deviation for the i-th component is high, this feature component has not information about the user target. Other approach to set Λ is trying to maximize the likelihood of the answer to the query, minimizing the distance between the previous query feature vector and the feature vectors of the selected relevant images. The minimized objective function is defined by $J(\Lambda) = \sum_{j=0}^{K-1} d_{WED}(F_Q, F_j)$, where F_Q is the feature vector of the previous query definition and F_j the feature vector of each one of the images in the relevant set. The $\sum_{i=0}^{r-1} \lambda_i^2 = 1$ constraint is imposed to avoid the trivial all zeros solution, and the estimation problem is stated as the minimization problem, $\hat{\Lambda} = argminJ(\Lambda)$. The estimation of the parameters of the matrix W for the GED is similar to the optimal estimation of the WED parameters. In this case, the minimization problem is defined by $\hat{W} = argminJ(W) = argmin \left\{ \sum_{j=0}^{K-1} d_{GED}(F_Q, F_j) \right\}$ under the constraint $det(W) = 1$. All these ideas can be extended to Correlation Based distances, which we will not elaborate here.

Probabilistic paradigm

In the probabilistic paradigm [31] the user target or semantic interpretation of the search is modelled by a probability distribution defined on the feature space. According to that, the process of image retrieval is an inference proccess that asigns a probability value to each image in database. The retrieval user feedback aims to fit the best probalibity distribution for the query, using the positive and negative relevance sets of images. Then, this probability distribution is used to asign *a posteriori* probability for each image in the database, which will be used as the distance measure for the query answer retrieval. Although in many cases the same strategy is used for positive and negative examples, is known that the negative images are not related to the target distribution, but they are isolated and independent, so different methodologies appear to treat the negative set of images product of the retrieval feedback iterations.

The Gaussian distribution for the semantic target features, $F \rightarrow N(\mu_F, \Sigma_F)$, the probability density function is given by:

$$g(F) = \frac{1}{(2\pi)^{d/2} \left|\Sigma_F\right|^{1/2}} e^{-\frac{1}{2}(F-\mu_F)^T \Sigma_F^{-1} (F-\mu_F)}$$

where d is the feature vector dimension, μ_F is the average feature vector for the Gaussian distribution and Σ_F is the covariance matrix whose determinant$|\Sigma_F|$ is used to normalize. Using the bayesian aproximation, the log posterior probability of image I denoted by its feature vector F belonging to target model M is defined by:

$$p_M(F) = \ln P(M|F) \propto \ln p(F|M) + \ln P(M)$$

$$= -\frac{1}{2}\frac{(F-\mu_F)^T}{\Sigma_F} + \ln P(M) + constant$$

and the distance function can be defined by:

$$d(I, M) = -p_M(F_I)$$

The estimation of the log posterior probability needs the estimation of μ_F, Σ_F and the *a priori* probability of the target model, $P(M)$. Using only the set of positive images, let F_j be the feature vector of the j-th image on relevant set and N_F the number of relevant images in the set. To estimate the parameters needed to value $p_M(F)$ the following conventional equations are used:

$$\mu_F = \frac{1}{N_F}\sum_j F_j$$

$$\Sigma_F = \frac{1}{N_F - 1}\sum_j (F_j - \mu_F)(F_j - \mu_F)^T$$

If the features are supposed to be mutually independent the covariance matrix can be simplicited as a diagonal matrix of standard deviations, $\sigma_F^2 = diag \cdot (\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2)$. The *a priori* probability, $P(M)$, is assumed to be equal for all the possible classes in database son it remainds constant all along the retrieval proccess.

Instead of computing the parameters in each iteration using all the positive images, these can be estimated using the parameters of the previous iteration and the features of the new positive images in relevant set. Supposing features are mutually independent, let F_j^U be the feature vector of the j-th new positive image on the present iteration, U_F the number of new relevant images in the current iteration, N'_F the number of positive images in previous iteration and μ'_F, ${\sigma'}_F^2$ the parameters of the previous gaussian distribution. The new parameters μ_F and σ_F^2 are estimated using:

$$N_F = N'_F + U_F$$

$$\mu_F = \frac{N'_F\mu'_F + \sum_U F_j^U}{N_F}$$

$$\sigma_F^2 = N'_F{\sigma'}_F^2 + \frac{N'_F U_F {\mu'}_F^2 - 2N'_F\mu'_F\sum_U F_j^U - (\sum_U F_j^U)^2}{N_F} + \sum_U (F_j^U)^2$$

If the irrelevat set of images is used to estimate the distribution parameters, this can be seen as the difference of two distribution, the one estimated with the positive examples and the one with the negatives, been the resultant distribution, in the case of the Gaussian distribution, as an hyperbolic in the future space instead of the hiperelypsoid resultant if only the relevant set is used. Instead of using the negative examples to estimate the distribution parameters, the set can be used as a penalization for the distance image. For this, the final distance measure is:

$$d(I, M) = -p_M(F_I) + \sum_V p_V(d(F_I, F_V))$$

where the first term is the distance explained above and the second one is a penalization function where each image, V, in the irrelevant set is modelled as a gaussian distribution, and $p_V(d(F_I, F_V))$ is the log posterior probability of feature vectors distance, euclidean for example.

In [31] a fully Bayesian framework is proposed, that starts from the low level parameter inference (the Markov Random Field texture model) up to the labeling of the cover types in the image, and the labeling of the whole image, where the interaction process helps building a Bayesian model of the image semantics.

5.3 Evaluating CBIR Systems

In CBIR systems, the approapriate evaluation methodology is is an open question [33, 22, 5, 8, 35]. Most of the methods used in CBIR evaluation come from the Information Retrieval (IR) field where the researchers have great experience evaluating information retrieval systems. From this area of knowledge appears two standar methods, the Precision and Recall measures (below defined), that are long used over CBIR systems. But working with image collections instead of textual databases has some inherited difficulties that made IR methods insufficient. First of them is that there is not a standar image collection researchers could work with, so there is not a way to compare the results of different investigations.

Suppose an image database D and a query Q are given and that the database can be divided in two groups, the relevant set of images for the query Q, R_Q, and its complement, the set of irrelevant images $\overline{R}_Q$. Suppose the CBIR System retrieves a set of images in D for the query Q, A_Q. The Precision measure, denoted p, is defined as the fraction of the set of images retrieved by the system that are relevant for the query. Another point of view is the Error measure, denoted e, assumed to be the fraction of the set of images retrieved by the system that are irrelevant for the query. The Recall measure, denoted r, is the fraction of relevant images returned by the query. Usually this information is presented as a Precision (or Error) versus Recall graph, PR (or ER) graphs.

$$p = \frac{|A_Q \cap R_Q)|}{A_Q} \qquad e = \frac{|A_Q \cap \overline{R}_Q|}{A_Q} \qquad r = \frac{|R_Q \cap A_Q|}{R_Q}$$

But PR graphs have limitations and other analyses based on precision and recall measures are used too as, for example, the recall measure where the precision falls below 0.5, the non-interpolated mean average precision or the precision after the first N_K relevant images are retrieved where N_K is the number of relevant images in a K-NN based query. Moreover, the precision and recall measures may be misleading due to the great difficulty of defining the relevant set of images. Because of these limitations other measures have been defined for evaluating CBIR systems. One of them is the First Rank, $rank_1$, and the Average Rank, $\widetilde{rank}$, being respectively, the rank at which the first relevant image is retrieved and the normalized average rank of relevant images defined by equation 5.11 where N_D is the number of images in dataset, N_R is the number of images in the relevant set R_Q and $rank_i$ is the rank at which the ith relevant image is retrieved. A $\widetilde{rank} = 0$ indicates a perfect performance, while the performance goes worst as $\widetilde{rank}$ approximates 1. A random retrieval would give a evaluation of $\widetilde{rank} = 0.5$.

$$\widetilde{rank} = \frac{1}{N_D N_R}\left(\sum_{i=1}^{N_R} rank_i - \frac{N_R(N_R-1)}{2}\right) \tag{5.11}$$

Another point of view is not to use any query answer relevant image set. Given a query Q for the database D, an ideal ordering of the database $Z = [z_1, \ldots, z_k]$ is assumed. An ideal retrieval systems will provide a relevance measure $S(I_j) \in [0, 1]$ for each image. The system under evaluation will retrieve for query Q an ordering $Z_Q = [z_{\pi_1}, \ldots, z_{\pi_k}]$, where $[\pi_1, \ldots, \pi_k]$ is a permutation of $[1, \ldots, k]$. The weighted displacement, denoted ω, can be computed by equation 5.12 and can be used as a measure of system performance without the need of defining a relevant set, althought the problem of obtaining an ideal ordering Z remains open.

$$\omega = \sum_j S(I_j)\,|j - \pi_j| \tag{5.12}$$

5.3.1 Evaluation in the literature of remote sensing

Early works, such as [12] contain as evaluation processes the gathering of a collection of images and the counting of relevant positive answers by the observer. The works of Datcu and colleagues

have been quite influential in the last years. In [31] the evaluation consists on the exploration of a database of some aerial images and a Landsat TM image through a web site. The authors suggest some statistical measures of the labellling process, such as the variance of the labelling probabilities for a given class, or the information divergence. In [5] a complete evaluation protocol is proposed that includes the quality of the queried results by precision and recall values, as well as the probability to overretrieve and forget. They also analyze man-machine interaction. The Human Centered Concepts for remote sensing image database exploration [7], are tested over a collection of heterogeneous data from different sensor types, performing complicated tasks, such as risk analysis in mined areas and detection of objects and structures.

5.4 A case of hyperespectral image retrieval

In this section we will present some ideas about hyperspectral image retrieval which have been already introduced in [21]. The idea is that the endmembers induced from the image by any suitable method are to be used as the image features for image indexing and retrieval. For our own computational experiments we use the method proposed in [10, 11]. This is an appropriate characterization of the spectral image content. We propose an image similarity function defined on the basis of the distance between endmembers. In the next subsection we present this image similarity function, before commenting some computational results over a database of synthetic images.

5.4.1 Endmember based image similarity

Let $E^k = [e_1^k, \ldots, e_{n_k}^k]$ be the image endmembers obtained applying any of the methods discussed in section 5.2.2 over the k-th image $f_k(x, y)$ in the database. Given two hyperspectral images, $f_k(x, y)$ and $f_l(x, y)$, their endmember sets E^k and E^l may have different number of endmembers, i.e. $n_k \neq n_l$. That implies that the feature spaces of the images are different, whithout any functional relation between them. The computation of the similarity between the images starts with the computation of the matrix of Euclidean distances between all the possible pairs of endmembers from each image:

$$D_{k,l} = [d_{i,j}; i = 1, ..., n_k; j = 1, .., n_l]$$

where

$$d_{ij} = \left\| e_i^k - e_j^l \right\|.$$

We compute the vector of row minimum values

$$\mathbf{m}_k = \left[m_i^k = min_j \left\{ d_{ij} \right\} \right],$$

and column minimum values

$$\mathbf{m}_l = \left[m_j^l = min_i \left\{ d_{ij} \right\} \right].$$

We compute the similarity between the hyperspectral images as follows:

$$d(f_k, f_l) = (\|m_k\| + \|m_l\|)(|n_k - n_l| + 1). \tag{5.13}$$

The endmember induction procedure may give different number of endmembers and endmember features for two hyperspectral images. The similarity measure of eq. 5.13 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.

The worst case is when all the endmembers E^k of image f_k are very near to a subset E_*^l of the endmembers E^l of the image f_l. Then the magnitude of the vector of row minimum distances will be very small, because all the rows will have a very low minimum value. However, the magnitude of the vector of column minimum distances will be much greater, because the columns corresponding

to the subset of endmembers $E^l - E^l_*$ will have a relatively large minimum value. Thus the similarity measure of eq. 5.13 can cope with the asymmetry of the situation. Besides, the formulation of eq. 5.13 avoids the combinatorial problem of deciding the optimal matching of endmembers. When the number of endmembers is different from one image to the other, their difference is introduced as an amplifying factor. The measure is independent of image size and, as the endmember induction algorithm described in [10, 11] is very fast, it can be computed in acceptable time. The endmember set poses no storage problem. Our approach does not use spatial features, 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.

5.4.2 Experimental results

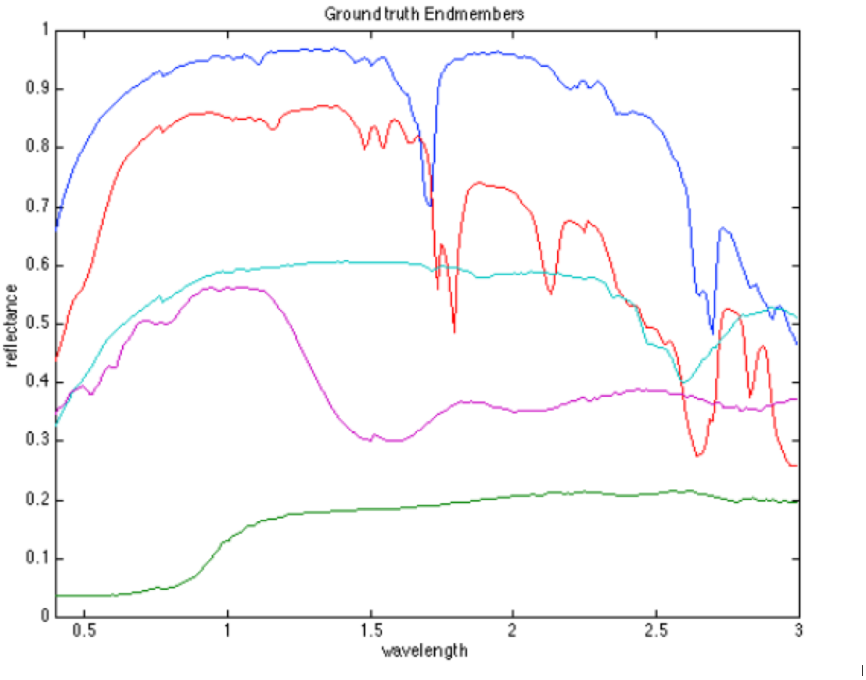

Figure 5.2. Endmembers from the USGS repository used to synthesize an image.

The hyperspectral images used for the experimental results are generated as linear mixtures of a set of spectra (the ground truth endmembers) applying synthesized abundance images. The ground truth endmembers were randomly selected from a subset of the USGS spectral libraries corresponding to the AVIRIS flights. Figure 5.4.2 show some spectra in used in 5 endmember images. The synthetic ground truth abundance images were generated in a two step procedure, first we simulate each as an gaussian random field with Mattern correlation function of parameters varying between 2 and 20. We applied the procedures proposed by [17] 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 region of almost pure pixels, that appear as almost white regions when visualizing the abundance images. Image size is 256x256 pixels of 224 spectral bands . We have generated collections of images with 2 to 5

ground truth endmember/abundances, for total number of 400 images, 100 images for each number of predefined endmembers. This is the biggest synthetic hyperspectral image database we know of.

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

1. Compute the distances between the images in the database, on the basis of eq. 5.13, 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 [10,11] The value of the noise gain was set to a=0.5.
3. Compute the distances between the images in the database, on the basis of eq. 5.13, 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 response images, or correct responses.

In table 5.5 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 restrict the search to the subcollection of images with the same number of endmembers as the query image. Each entry in the table corresponds to the average number of relevant (as defined before) images obtained in the response to the query.

In table 5.5 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 relevant images obtained in the query. This effect can be due to the fact that the morphological independence algorithm may find a number of endmembers different from the ground truth number of endmembers, 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 the same number of ground truth endmembers as the query, 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 eq. 5.13 or the probability of classification 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.

5.5 Conclusions 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) and other algorithms have been favourable to the AMM algorithm [10,11].

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

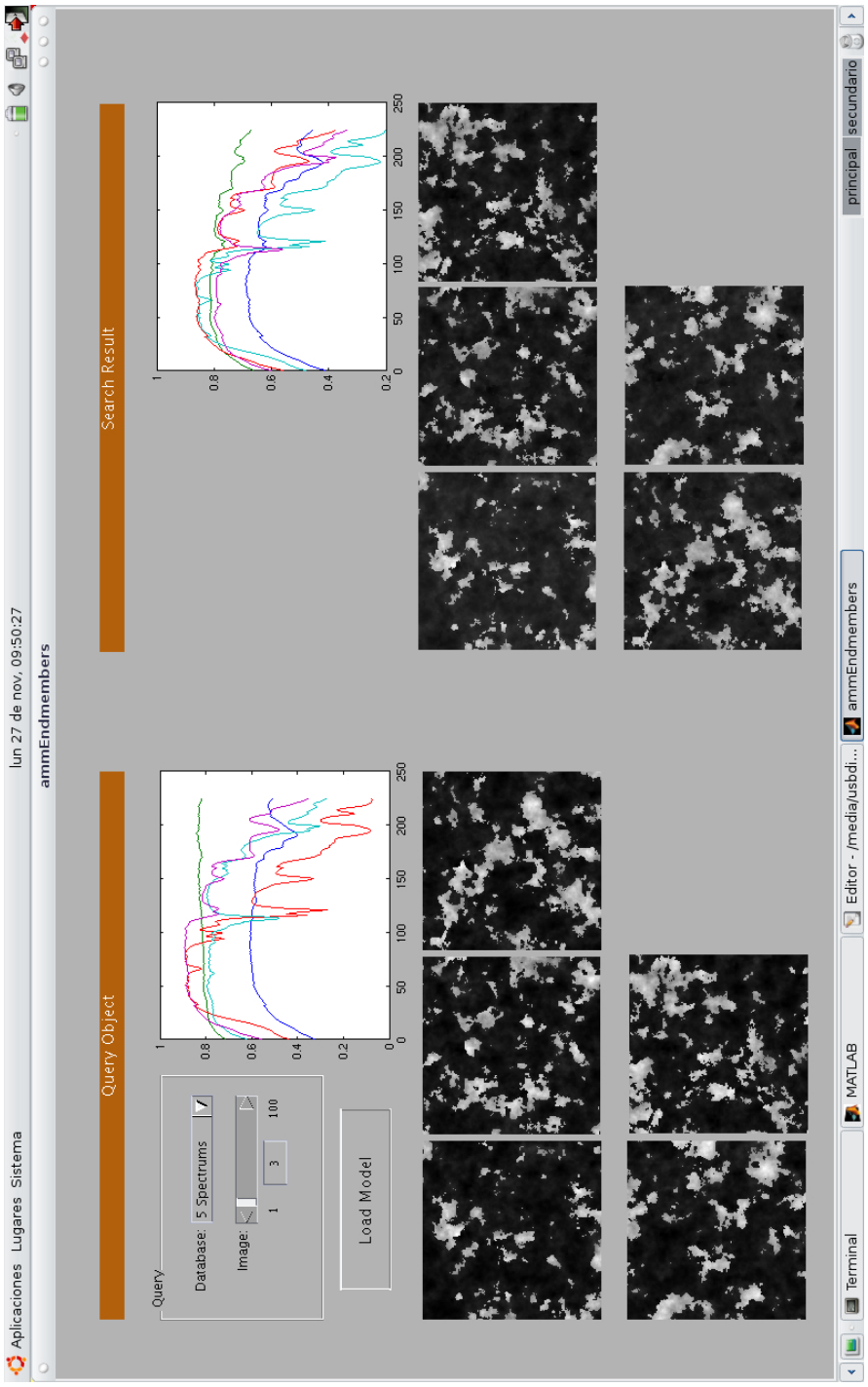

Figure 5.3. An instance of the hyperspectral CBIR interface: a query and its closest image according to the set of induced endmembers.

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 georeferencing coordinates, will serve to reduce this computational problem. The spatial information may also serve as the basis for user relevance feedback. Relevance may be easier to specify for the user on the abundance images, because she can focus on spatial structures.

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).

	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 5.1. Relevance results of the experimentation retrieving images from the synthetics image database

Acknowledgements

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

References

1. 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
2. M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, J. F. Huntington; ICE: A statical Approach to Identifying Endmembers in Hyperspectral Images; IEEE Transactions on Geoscience and Remote Sensing, Vol. 42, N°10, Pag. 2085-2095, October 2004
3. D. Brahmi, D. Ziou; Improving CBIR Systems by Integrating Semantic Features; Proceedings of the First Canadian Conference on Computer and Robot Vision (CRV'04), IEEE, 2004
4. C. Chang, A. Plaza; A Fast Iterative Algorithm for Implementation of Pixel Purity Index; IEEE Geoscience and Remote Sensing Letters, Vol. 3, N°1, Pag. 63-67, January 2006
5. H. Daschiel, M. Datcu; Information Mining in Remote Sensing Image Archives: System Evaluation; IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, N°1, Pag. 188-199, January 2005

6. M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Marchetti, S. D'Elia; Information Mining in Remote Sensing Image Archives: System Concepts; IEEE Transactions on Geoscience and Remote Sensing, Vol. 41, N°12, Pag. 2923-2936, December 2003
7. M. Datcu, K. Seidel; Human-Centered Concepts for Exploration and Understanding of Earth Observation Images; IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, N°3, Pag. 601-609, March 2005
8. N. Doulamis, A. Doulamis; Evaluation of Relevance Feedback Schemes in Content-Based in Retrieval Systems; Signal Processing: Image Communication, Vol. 21, N°4, Pag. 334-357, April 2006
9. D. Gillis, J. Bowles, M. E. Winter; Using Endmembers as a Coordinate System in Hyperspectral Imagery; 16th Annual International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando FL, 2002
10. M. Graña, J.Gallego, C. Hernandez, Further results on AMM for endmember induction, Proc. IEEE Workshop on Adv. Tech. Anal. Remotely Sensed Data, Washington D.C., Oct. 2003, pp.237 - 243
11. 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
12. G. Healey, A. Jain; Retrieving Multispectral Satellite Images Using Physics-Based Invariant Representations; IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, N° 8, Pag. 842-848, August 1996
13. C. Hoi, M. R. Lyu; Group-Based Relevance Feedback With Support Vector Machine Ensembles; Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), IEEE, 2004
14. C. Hsu, C. Li; Relevance Feedback Using Generalized Bayesian Framework With Region-Based Optimization Learning; IEEE TRansactions on Image Processing, Vol. 14, N°10, Pag. 1617-1631, October 2005
15. W. Jiang, G. Er, Q. Dai, J. Gu; Hidden Annotation for Image Retrieval With Long-Term Relevance Feedback Learning; Pattern Recognition, Volume 38, N°11, Pag. 2007-2021, November 2005
16. I. King, Z. Jin; Integrated Probability Function and its Application to Content-Based Image Retrieval by Relevance Feedback; Pattern Recognition, Vol. 36, N°9, Pag. 2177-2186, September 2003
17. Kozintsev B. Computations With Gaussian Random Fields, PhD Thesis, ISR99-3 University of Maryland (1999)
18. J. W. Kwak, N. I. Cho; Relevance Feedback in Content-Based Image Retrieval System by Selective Region Growing in the Feature Space; Signal Processing: Image Communication, Vol. 18, N°9, Pag. 787-799, October 2003
19. B. Li, S. Yuan; A Novel Relevance Feedback Method in Content-Based Image Retrieval; Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04), IEEE, 2004
20. S. D. MacArthur, C. E. Brodley, A. C. Kak, L. S. Broderick; Interactive Content-Based Image Retrieval Using Relevance Feedback; Computer Vision and Image Understanding, Vol. 88, N°2, Pag. 55-75, November 2002
21. O. Maldonado, D. Vicente, M. Graña; CBIR Indexing Hyperspectral Images; Geoscience and Remote Sensing Symposium, IGARSS'05. Proceedings 2005 IEEE International, Vol. 2, Pag. 1202-1205, July 2005
22. H. Müller, W. Müller, D. McG. Squire, S. Marchand-Maillet, T. Pun; Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals; Pattern Recognition Letters, Vol. 22, N°5, Pag. 593-601, April 2001
23. J. M. P. Nascimento, J. M. Bioucas Dias; Does Independent Component Analysis Play a Role in Unmixing Hyperspectral Data; IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, N°1, Pag. 175-187, January 2005
24. C. U. Ng, G. R. Martin; Automatic Selection of Attributes by Importance in Relevance Feedback Visualization; Proceedings of the 8th International Conference on Information Visualisation (IV'04), IEEE, 2004
25. G. Park, Y. Baek, H. Lee; Re-Ranking Algorithm Using Post-Retrieval Clustering for Content-Based Image Retrieval; Information Processing & Management, Vol. 41, N°2, Pag. 177-194, March 2005
26. A. Plaza, P. Martínez, R. Pérez, J. Plaza; Spatial/Spectral Endmember Extraction by Multidimensional Morphological Operations; IEEE Transactions on Geoscience and Remote Sensing, Vol. 40, N°9, Pag. 2025-2041, September 2002
27. A. Plaza, P. Martínez, R. Pérez, J. Plaza; A Quantitative and Comparative Analisys of Endmember Extraction Algorithms from Hyperspectral Data; IEEE Transactions on Geoscience and Remote Sensing, Vol. 42, N°3, Pag. 650-663, March 2004

28. A. Plaza, P. Martínez, J. Plaza, R. Pérez; Dimensionality Reduction and Classification of Hyperspectral Image Data Using Sequences of Extended Morphological Transformations; IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, N°3, Pag. 466-479, March 2005
29. R. Ramanath, W. E. Snyder, H. Qi; Eigenviews for Object Recognition in Multispectral Imaging Systems; Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop (AIPR'03), IEEE, 2003
30. H. Ren, C. Chang; A Generalized Orthogonal Subspace Projection Approach to Unsupervised Multispectral Image Classification; IEEE Transactions on Geoscience and Remote Sensing, Vol. 38, N° 6, Pag. 2515-2528, November 2000
31. M. Schröder, H. Rehrauer, K. Seidel, M. Datcu; Interactive Learning and probabilistic Retrieval in Remote Sensing Image Archives; IEEE Transactions on Geoscience and Remote Sensing, Vol. 38, N°5, Pag. 2288-2298, September 2000
32. Y. E. Shimabukuro, J. A. Smith; The Least-Squares Mixing Models to Generate Fraction Images Derived From Remote Sensing Multispectral Data; IEEE Transactions on Geoscience and Remote Sensing, Vol. 29, N°1, Pag. 16-20, January 1991
33. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain; Content-Based Image Retrieval at the End of the Early Years; IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, N° 12, Pag. 1349-1380, December 2000
34. D. Tao, X. Tang; Nonparametric Discriminant Analysis in Relevance Feedback for Content Based Image Retrieval; Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), IEEE, 2004
35. J. Vogel, B. Schiele; Performance Evaluation and Optimization for Content-Based Image Retrieval; Pattern Recognition, Vol. 39, N° 5, Pag. 897-909, May 2006
36. L. Wang, Y. Gao, K. L. Chan, P. Xue; Retrieval With Knowledge-Driven Kernel Design: An Approach to Improving SVM-Based CBIR With Relevance Feedback; Proceedings of the 10th International Conference on Computer Vision (ICCV'05), IEEE, 2005
37. E. M. Winter; N-FINDR: an Algorithm for Fast Autonomous Spectral Endmember Determination in Hyperspectral Data; Proceedings of SPIE, Vol. 3753: Imaging Spectrometry V, October 1999
38. D. Ziou, S. Boutemedjet, An Information Filtering Approach for the Page Zero Problem, Multimedia Content Representation, Classification and Security, LNCS 4105, pp. 619-626