ON THE VALIDATION OF A SPECTRAL/SPATIAL CBIR SYSTEM FOR HYPERSPECTRAL IMAGES

Miguel A. Veganzones, Manuel Grana

Grupo de Inteligencia Computacional
Universidad del País Vasco, Spain

ABSTRACT

We report validation results on a novel Content-Based Image Retrieval system for hyperspectral images that uses both the spectral and spatial features obtained from unsupervised unmixing processes. Spectral features consist of the set of endmembers obtained from the image by an Endmember Induction Algorithm. Spatial features consist of statistics of abundance images. Both pieces of information are combined into a similarity measure defined between the hyperspectral images that guides the search for answers to the query. Such a system allows one to retrieve images containing materials similar to the query image, and in a similar proportion. We provide validation results using synthetic hyperspectral datasets.

Index Terms— Hyperspectral CBIR, unmixing, validation, synthetic images

1. INTRODUCTION

In Content Based Image Retrieval (CBIR) systems [1], the images stored in the database are indexed by feature vectors extracted from the images by means of computer vision and digital image processing techniques. In the query-by-example approach, the interrogation to the database is done through the presentation of a query image, and the answer are the most similar images in the database according to some similarity measure. The increasing amount of Earth Observation data provided by hyperspectral sensors, motivates research in some technological problems related to the scale of the available data. The problem of searching through these huge databases using CBIR techniques has not been properly addressed for the case of hyperspectral images. Approaches to CBIR in remote sensing images proposed up to now are focused on panchromatic, low dimension multispectral images such as LANDSAT or SAR data [2, 3, 4]. There are few works in the literature [5, 6] dealing explicitly with the spectral information to guide the search. These works define the image features as the endmembers induced from the hyperspectral data by some Endmembers Induction Algorithm (EIA). One inconvenience of these CBIR systems is that they can not discriminate among images with the same induced endmembers but very different spatial distributions. In this paper, we propose a spatial-spectral feature extraction process, that makes this discrimination possible.

Section 2 introduces the proposed Spectral-Spatial CBIR system for hyperspectral images. In section 3 we provide validation results using hyperspectral synthetic datasets. Finally, we give some conclusions in section 4.

2. SPECTRAL-SPATIAL CBIR SYSTEM

We first describe the feature extraction process followed to quantify the spectral and spatial information from the hyperspectral images. Then, we introduce the proposed Spectral-Spatial CBIR system for hyperspectral image databases.

2.1. Spectral-spatial feature extraction

Each hyperspectral image is processed to extract the features that quantify the spectral-spatial information in the images. We follow the linear mixing formulation, where a hyperspectral image is the result of the linear combination of the pure spectral signatures of ground components, named endmembers, with a fractional abundance matrix. Let $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ be the pure endmember signatures (normally corresponding to macroscopic objects in scene, such as water, soil, vegetation, ...) where each $\mathbf{e}_i \in \mathbb{R}^q$ is a q-dimensional vector. Then, the hyperspectral signature $\mathbf{r}$ at each pixel in the image is defined by the expression $\mathbf{r} = \mathbf{s} + \mathbf{n} = \sum_{i=1}^{p} \mathbf{e}_i \phi_i + \mathbf{n}$, where the hyperspectral signature $\mathbf{r}$ is formed by the sum of the pixel's signal $\mathbf{s}$ and an independent additive noise component $\mathbf{n}$; and, ϕ is the p-dimensional vector of fractional abundances at given pixel. This equation can be generalized to the full image by $\mathbf{H} = \mathbf{E}\mathbf{\Phi} + \mathbf{n}$ where $\mathbf{H}$ is the hyperspectral image and $\mathbf{\Phi}$ is a matrix of fractional abundances.

We characterize an hyperspectral image H_α by a tuple (E_α, Φ_α), where $E_\alpha = \left\{\mathbf{e}_1^\alpha, \mathbf{e}_2^\alpha, \ldots, \mathbf{e}_{p_\alpha}^\alpha\right\}$ is the set of p_α induced endmembers from the α-th image; and, $\Phi_\alpha = \left\{\phi_1^\alpha, \phi_2^\alpha, \ldots, \phi_{p_\alpha}^\alpha\right\}$ is the set of fractional abundance maps resulting from the unmixing process, where each ϕ_i^α represents the spatial distribution of the induced endmember $\mathbf{e}_i^\alpha$, $i = 1, \ldots, p_\alpha$, in the image. To that effect, an EIA is used to induce the spectral signatures (the endmembers) of the

image; and then, an unmixing method extracts from the image the spatial distribution of each endmember (the fractional abundances).

2.2. Proposed Spectral-Spatial CBIR system

The proposed Spectral-Spatial CBIR system defines a dissimilarity measure between the spectral-spatial feature vectors of hyperspectral images, extracted as described above. The aim is to declare as similar images sharing some of their constituent materials. We detect this similarity through the similarity between image spectra and their relative abundance proportions. The spectral-spatial dissimilarity is a version of the Integrated Region Matching (IRM) dissimilarity function [7] used for region matching-based image retrieval. The overall IRM dissimilarity between two images depends on two aspects: the similarity between each region of the two images and the significance of each region matching for determining the overall dissimilarity. In the proposed Spectral-Spatial dissimilarity function, we calculate the distance between each image endmembers, and provide a significance based on their corresponding average abundances.

Let $E_\alpha = \left\{\mathbf{e}_1^\alpha, \mathbf{e}_2^\alpha, \ldots, \mathbf{e}_{p_\alpha}^\alpha\right\}$ be the set of endmembers induced from the hyperspectral image H_α in the database, where p_α is the number of induced endmembers from the α-th image; and $\bar{\Phi}_\alpha = \left\{\bar{\phi}_1^\alpha, \bar{\phi}_2^\alpha, \ldots, \bar{\phi}_{p_\alpha}^\alpha\right\}$ the average abundances of H_α, where $\bar{\phi}_i^\alpha = \frac{1}{N}\sum_{j=1}^{N}\phi_i^\alpha(j)$, $\phi_i^\alpha(j)$ is the i-th endmember fractional abundance for pixel j, and N is the number of pixels on the image. Given two images, H_α and H_β, we compute the Spectral Distance Matrix, $D_{\alpha,\beta}$, whose elements are the distances between the endmembers of each image as

$$D_{\alpha,\beta} = \left[d_{i,j}; i = 1, \ldots, p_\alpha; j = 1, \ldots, p_\beta\right], \qquad (1)$$

where $d_{i,j}$ is a distance between the endmembers $\mathbf{e}_i^\alpha, \mathbf{e}_j^\beta \in \mathbb{R}^q$. For instance, the euclidean distance or the angular pseudo-distance.

Then, the Spectral-Spatial dissimilarity function, $s(H_\alpha, H_\beta)$, is defined as

$$s(H_\alpha, H_\beta) = \sum_{i,j} r_{i,j} d_{i,j}, \qquad (2)$$

where $d_{i,j}$ is the spectral distance between endmembers $\mathbf{e}_i^\alpha$ and $\mathbf{e}_j^\beta$, and $r_{i,j}$ is the significance associated to $d_{i,j}$.

Thus, the problem reduces to choosing the significance matrix $R_{\alpha,\beta} = [r_{i,j}; i = 1, \ldots, p_\alpha; j = 1, \ldots, p_\beta]$. We followed the most similar highest priority (MSHP) principle as in [7], making use of the average abundances $\bar{\Phi}_\alpha$ and $\bar{\Phi}_\beta$. The average abundances represent "significance credits" assigned to the spectral distances by algorithm 1.

Algorithm 1 Significance credits assignment algorithm.

1. Set $\mathcal{L} = \{\}$ and denote $\mathcal{M} = \{(i,j) : i = 1, \ldots, p_\alpha; j = 1, \ldots, p_\beta\}$.
2. Choose the minimum $d_{i,j}$ for $(i,j) \in \mathcal{M} - \mathcal{L}$. Label the corresponding (i,j) as (i', j').
3. $r_{i',j'} = \min\left(\bar{\phi}_{i'}^\alpha, \bar{\phi}_{j'}^\beta\right)$.
4. If $\bar{\phi}_{i'}^\alpha < \bar{\phi}_{j'}^\beta$, set $r_{i',j} = 0$, $j \neq j'$; otherwise, set $r_{i,j'} = 0$, $i \neq i'$.
5. If $\bar{\phi}_{i'}^\alpha < \bar{\phi}_{j'}^\beta$, set $\bar{\phi}_{i'}^\alpha = 0$ and $\bar{\phi}_{j'}^\beta = \bar{\phi}_{j'}^\beta - \bar{\phi}_{i'}^\alpha$; otherwise, set $\bar{\phi}_{j'}^\beta = 0$ and $\bar{\phi}_{i'}^\alpha = \bar{\phi}_{i'}^\alpha - \bar{\phi}_{j'}^\beta$.
6. $\mathcal{L} = \mathcal{L} + \{(i', j')\}$.
7. If $\sum_{i=1}^{p_\alpha} \bar{\phi}_i^\alpha > 0$ and $\sum_{j=1}^{p_\beta} \bar{\phi}_j^\beta > 0$, go to step 2; otherwise, stop.

3. SYSTEM VALIDATION

We first describe how we construct the synthetic hyperspectral images that serve for the purposes of the computational experiments. Then, we comment on the methodology, giving the definition of the employed performance measures. Finally, we give the results of the experiments.

3.1. Synthetic hyperspectral images

The synthetic hyperspectral images are generated as linear mixtures of a set of spectra (the groundtruth endmembers) according to synthesized abundance coefficients for each pixel. Because the generation of the abundance coefficients follows a spatial distribution, we generate independent images for each abundance coefficient corresponding to an endmember, later we fuse them into a multidimensional abundance image imposing normalization conditions for each pixel independently. The groundtruth endmembers were randomly selected from a subset of the USGS spectral library.

The synthetic groundtruth multidimensional abundance images were generated in a two-step procedure. First, we simulate each abundance image corresponding to each endmember as a Gaussian random field with Matern correlation function of parameters $\theta_1 = 10$ and $\theta_2 = 1$. We applied the procedure proposed by [8] for the efficient generation of Gaussian random fields with large domains. Second, to ensure that there are regions of almost pure endmembers, we selected for each pixel the abundance coefficient with the greatest value and we normalize the remaining to ensure that the abundance coefficients sum up to one.

We have synthesized a total of 18000 hyperspectral images divided in nine datasets of 2000 images each. Each dataset is characterized by the number of endmembers in the collection of groundtruth endmembers and the images spatial size. We defined three collections of groundtruth endmembers, with pools of 5, 10 and 20 endmembers each, repre-

senting an increasing diversity in the materials. We denote the datasets generated by each of the endmembers collections as 5-datasets, 10-datasets and 20-datasets respectively. For each endmember collection we have also defined three collections of spatial sizes, with images having 64×64, 128×128 and 256×256 pixels, representing different spatial scales. All the synthesized hyperspectral images have 269 spectral bands *per* pixel. Each dataset contains 2000 hyperspectral images, where each image is built with 2 to 5 endmembers randomly selected from the corresponding pool of available of groundtruth endmembers.

3.2. Methodology

We have performed independent experiments[1] over each of the nine hyperspectral datasets using the Spectral-Spatial dissimilarity function (2), where the distance between endmembers is measured by the Euclidean distance, s_{euc}, and the Spectral Angle Map pseudo-distance, s_{sam}. For each image, we apply independently the N-FINDER [9] and the EIHA [10] endmember induction algorithms to induce the set of endmembers.

For each image H_α in a dataset we calculate the dissimilarity between H_α and each of the remaining images in the dataset. These dissimilarities are represented as a vector $\mathbf{s}_\alpha = [s_{\alpha 1}, \ldots, s_{\alpha N}]$, where N is the number of images in the dataset (2000 in our experiments) and $s_{\alpha,\beta}$ is the dissimilarity between the images H_α and H_β, with $\alpha, \beta = 1, \ldots, N$. Let us distinguish between $\mathbf{s}_\alpha^{\text{GT}}$, the vector of dissimilarities computed using the known ground truth endmembers, and $\mathbf{s}_\alpha^{\text{IND}}$, the vector of dissimilarities computed using the endmembers induced by one of the EIAs (either N-FINDER or EIHA). We can define the ranking of the dataset relative to one of the images $\Omega_\alpha = [\omega_{\alpha,p} \in \{1, \ldots, N\}; p = 1, \ldots, N]$ as the set of image indices ordered according to increasing values of their corresponding entries in the dissimilarity vector $\mathbf{s}_\alpha$. That is, we sort in increasing order the components of $\mathbf{s}_\alpha$, and the resulting shuffled image indices constitute Ω_α, so that $s_{\alpha,\omega_{\alpha,p}} \leq s_{\alpha,\omega_{\alpha,p+1}}$. We distinguish rankings $\Omega_\alpha^{\text{GT}}$ and $\Omega_\alpha^{\text{IND}}$ corresponding to the ground truth and induced dissimilarities, respectively. A query $Q_k(H_\alpha)$ is formulated as a search for the k most similar (less dissimilar) images H_β in the dataset with respect to the image H_α, with $1 \leq k \leq N$. The set of returned images $T_k(H_\alpha)$ and the set of relevant images $V_k(H_\alpha)$ for a query $Q_k(H_\alpha)$ are defined as follows:

$$T_k(H_\alpha) = \Omega_{\alpha,k}^{\text{IND}} = \left[\omega_{\alpha,p}^{\text{IND}} \text{ s.t. } s_{\alpha,\omega_{\alpha,p}^{\text{IND}}} \leq s_{\alpha,\omega_{\alpha,k}^{\text{IND}}}\right] \quad (3)$$

$$V_k(H_\alpha) = \Omega_{\alpha,k}^{\text{GT}} = \left[\omega_{\alpha,p}^{\text{GT}} \text{ s.t. } s_{\alpha,\omega_{\alpha,p}^{\text{GT}}} \leq t\right] \quad (4)$$

[1]The Matlab code for the hyperspectral image synthesis and endmember induction is available from http://www.ehu.es/ccwintco/index.php/GIC-source-code-free-libre

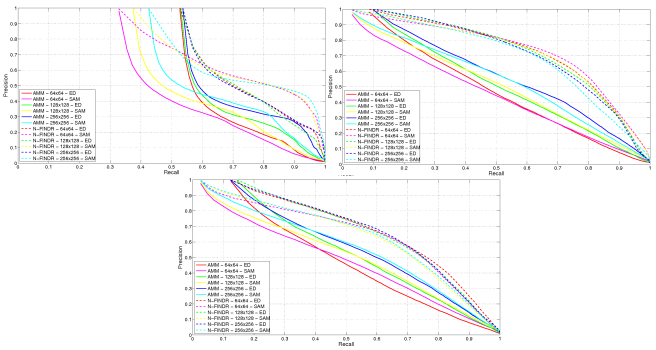

Fig. 1. Precision-recall curves for the 20-datasets (top-left), 10-datasets (top-right), and 5-datasets (bottom).

where $t = \bar{s}_\alpha^{\text{GT}} - 2\sigma_{\mathbf{s}_\alpha^{\text{GT}}}$, and $\bar{s}_\alpha^{\text{GT}}$ and $\sigma_{\mathbf{s}_\alpha^{\text{GT}}}$ are respectively the mean and standard deviation of $\mathbf{s}_\alpha^{\text{GT}}$.

This definition allows for the inclusion in the query answer of images whose dissimilarity is equal to the maximum one, thus allowing that the cardinality of both returned and relevant sets may be bigger than k. The Precision $P_k(H_\alpha)$ and Recall $R_k(H_\alpha)$ for a query $Q_k(H_\alpha)$ are standard performance measures in CBIR literature, they are defined as: $P_k(H_\alpha) = \frac{|V_k(H_\alpha) \cap T_k(H_\alpha)|}{|T_k(H_\alpha)|}$ and $R_k(H_\alpha) = \frac{|V_k(H_\alpha) \cap T_k(H_\alpha)|}{|V_k(H_\alpha)|}$. The average Precision and Recall of the system for a query of size k are defined as: $P_k = \frac{1}{N}\sum_{\alpha=1}^{N} P_k(H_\alpha)$ and $R_k = \frac{1}{N}\sum_{\alpha=1}^{N} R_k(H_\alpha)$. As summary performance quantity, we calculate the normalized average rank of relevant images [11]. The normalized rank for a given image ranking Ω_α, denoted as $Rank\,(H_\alpha)$, is defined as $Rank\,(H_\alpha) = \frac{1}{NN_\alpha}\left(\sum_{i=1}^{N_\alpha} R_i - \frac{N_\alpha(N_\alpha-1)}{2}\right)$, where N is the number of images in the dataset ($N = 2000$ in our experiments), N_α is the number of relevant images for the query, and R_i is the rank at which the i-th image is retrieved. This measure is 0 for perfect performance, and approaches 1 as performance worses. The average normalized rank ANR for the full dataset is given by: $ANR = \frac{1}{N}\sum_{\alpha=1}^{N} Rank\,(H_\alpha)$.

3.3. Results

Figure 1 shows the precision-recall curves for the 5-datasets, 10-datasets and 20-datasets respectively. Table 1 shows the ANR for each dataset. Overall the average normalized rank results for the endmembers/abundances obtained with both the N-FINDER and the EIHA are close to 0, with all combinations of image sizes, number of source endmembers and distances between individual endmembers. The proposed spectral-spatial features are, therefore, good characterization for CBIR systems.

The examination of figure 1 shows a high dependence of the precision-recall of the CBIR system on both the EIA and the underlying endmember diversity. We find that the N-FINDER has obtained better results in the sense that its

			ANR		
Spatial size	Distance	EIA	5-dataset	10-dataset	20-dataset
64×64	Euclidean	EIHA	0.0383	0.0442	0.0356
64×64	Euclidean	N-FINDER	0.0051	0.0120	0.0109
64×64	SAM	EIHA	0.0512	0.0559	0.0558
64×64	SAM	N-FINDER	0.0035	0.0145	0.0318
128×128	Euclidean	EIHA	0.0216	0.0306	0.0228
128×128	Euclidean	N-FINDER	0.0056	0.0108	0.0118
128×128	SAM	EIHA	0.0371	0.0440	0.0458
128×128	SAM	N-FINDER	0.0026	0.0153	0.0340
256×256	Euclidean	EIHA	0.0116	0.0186	0.0189
256×256	Euclidean	N-FINDER	0.0035	0.0119	0.0119
256×256	SAM	EIHA	0.0220	0.0368	0.0412
256×256	SAM	N-FINDER	0.0019	0.0180	0.0316

Table 1. ANR for the 5-datasets , 10-datasets and 20-datasets. Optimal ANR value is zero.

precision-recall curves is systematically above the ones corresponding to the EIHA. The diversity of the undelying pools of endmembers has the effect of giving higher precision under high recall rates as the diversity increases. The effect of the individual endmember distance chosen is negligible.

4. CONCLUSIONS

We introduce and validate a CBIR system for databases of hyperspectral images based on spectral-spatial features. The feature extraction is based on the application of an EIA and the ensuing linear unmixing based on the induced endmembers. We have defined an extensive validation experiment based on a big collection of synthetic hyperspectral images. Validation on synthetic images allows an exact comparison with the known ground truth. We can define precisely the set of relevant images to a query, so that precision and recall are computed exactly. Precision/recall curves show the good response of our feature selection and corresponding distance. We found that the average normalized rank is very close to zero, which is an extraordinarily good performance for CBIR systems. Further work will focus on the validation of this approach on real-life remote sensing hyperspectral images. We are working on the definition of appropriate query results visualization and user interaction for the construction of retrieval feedback to enhance the CBIR system performance.

5. REFERENCES

[1] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 12, pp. 1349–1380, 2000.

[2] M. Datcu and K. Seidel, "Human centered concepts for exploration and understanding of satellite images," 2003, pp. 52–59.

[3] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P.G. Marchetti, and S. D'Elia, "Information mining in remote sensing image archives: system concepts," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 12, pp. 2923–2936, 2003.

[4] H. Daschiel and M. Datcu, "Information mining in remote sensing image archives: system evaluation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 1, pp. 188–199, 2005.

[5] J.O. Maldonado, D. Vicente, M.A. Veganzones, and M. Grana, "Spectral indexing for hyperspectral image CBIR," Torrejon air base, Madrid (Spain), 2006.

[6] Miguel A. Veganzones, Jose O. Maldonado, and Manuel Grana, "On Content-Based image retrieval systems for hyperspectral remote sensing images," in *Computational Intelligence for Remote Sensing*, vol. 133 of *Studies in Computational Intelligence*, pp. 125–144. Springer Berlin / Heidelberg, 2008.

[7] Jia Li, James Z. Wang, and Gio Wiederhold, "IRM: integrated region matching for image retrieval," in *Proceedings of the eighth ACM international conference on Multimedia*, Marina del Rey, California, United States, 2000, MULTIMEDIA '00, pp. 147–156, ACM.

[8] Boris Kozintsev, *Computations with Gaussian Random Fields, PhD Thesis*, University of Maryland, 1999.

[9] M. E. Winter, M. R. Descour, and S. S. Shen, "N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," Denver, CO, USA, Oct. 1999, vol. 3753, pp. 266–275, SPIE.

[10] Manuel Grana, Ivan Villaverde, Jose O. Maldonado, and Carmen Hernandez, "Two lattice computing approaches for the unsupervised segmentation of hyperspectral images," *Neurocomput.*, vol. 72, no. 10-12, pp. 2111–2120, 2009.

[11] Henning Muller, Wolfgang Muller, David McG. Squire, Stephane Marchand-Maillet, and Thierry Pun, "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 593–601, Apr. 2001.