Region-Based Retrieval of Remote Sensing Images Using an Unsupervised Graph-Theoretic Approach

Bindita Chaudhuri, Begüm Demir, *Senior Member, IEEE*, Lorenzo Bruzzone, *Fellow, IEEE*, and Subhasis Chaudhuri, *Fellow, IEEE*

***Abstract*—This letter introduces a novel unsupervised graph-theoretic approach in the framework of region-based retrieval of remote sensing (RS) images. The proposed approach is characterized by two main steps: 1) modeling each image by a graph, which provides region-based image representation combining both local information and related spatial organization, and 2) retrieving the images in the archive that are most similar to the query image by evaluating graph-based similarities. In the first step, each image is initially segmented into distinct regions and then modeled by an attributed relational graph, where nodes and edges represent region characteristics and their spatial relationships, respectively. In the second step, a novel inexact graph matching strategy, which jointly exploits a subgraph isomorphism algorithm and a spectral graph embedding technique, is applied to match corresponding graphs and to retrieve images in the order of graph similarity. Experiments carried out on an archive of aerial images point out that the proposed approach significantly improves the retrieval performance compared to the state-of-the-art unsupervised RS image retrieval methods.**

***Index Terms*—Attributed relational graph (ARG), content-based image retrieval (CBIR), data mining, inexact graph matching, spectral graph embedding, subgraph isomorphism.**

I. Introduction

THE recent advances in satellite systems resulted in a sharp growth of remote sensing (RS) image archives. This has motivated increasing research interest on efficient and effective methods enabling accurate content-based retrieval of RS images from large archives. A simple yet effective method to perform content-based image retrieval (CBIR) on RS image archives is the k-nearest neighbor (kNN) technique. In kNN-based retrieval systems, once the user selects an image for query, the retrieval system automatically ranks the images based on their similarity to the query image and then displays the k most similar images in order of similarity. The similarity of images is estimated based on low-level features that describe and represent the image information content. However, the performance of CBIR is mainly limited by the gap between low-level features and high-level semantic concepts (i.e., semantic gap). Among the different approaches available for the representation of RS image information, compact global image representations that describe a set of local descriptors of each image have been recently found very effective in RS. This is due to the fact that such representations may contribute to fill the semantic gap by characterizing the RS image content in a small neighborhood of interest points. As an example, local descriptors that model portions of images around the interest points extracted by the scale-invariant feature transform (SIFT) have been recently found very effective in the context of RS image retrieval [1]. In order to summarize the local descriptors, their bag-of-visual-words (BoVW) representations (where the distributions of local descriptors are summarized into histograms that count the number of occurrence of visual words) are generally used. In [2], morphological texture descriptors are computed within local subwindows, and then, their bag-of-morphological-words representation is constructed in the framework of CBIR. An image descriptor based on the local binary pattern (LBP) has been recently introduced in [3]. This is defined by initially assigning a binary code to each image pixel by thresholding its neighboring sample values and then computing a histogram of the codes to derive the image descriptor.

Although these local descriptors are effective in modeling RS image content, they do not model the spatial relationships among different regions present in an image (which can be associated to the boundaries, distances, and orientations among regions). Thus, in the case of RS CBIR problems with complex image categories, these features result to be insufficient to provide a complete characterization of the image information content. To address this issue, few region-based image retrieval techniques [4], [5] have been introduced in RS. They represent images by graphs and apply graph matching [6] to retrieve the most similar images to the query image. However, these methods require reliable labeled samples that are representative of all of the land-cover classes within each image in the archive. This is very critical in large RS archives since, in most practical cases, it is time-consuming and almost impossible to get a sufficient number of labeled samples for the archive.

To overcome the limitations of the graph-based methods presented in the RS literature, in this letter, we propose a novel unsupervised graph-based approach to RS image characterization and retrieval. The proposed approach is based on two main steps. In the first step, a graph is constructed for each image by initially segmenting it using an unsupervised image segmentation algorithm and then modeling the image by an attributed relational graph (ARG) that can effectively represent both region characteristics (as nodes) and their spatial relationships (as edges). In the second step, the similarities between the

Manuscript received January 5, 2016; revised March 18, 2016; accepted April 12, 2016. Date of publication June 2, 2016; date of current version June 10, 2016.

B. Chaudhuri and S. Chaudhuri are with the Vision and Image Processing Laboratory (VIPLAB), Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400 076, India (e-mail: binditac@ee.iitb.ac.in; sc@ee.iitb.ac.in).

B. Demir and L. Bruzzone are with the Remote Sensing Laboratory (RSLAB), Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy (e-mail: demir@disi.unitn.it; lorenzo.bruzzone@ing.unitn.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2016.2558289

1545-598X © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

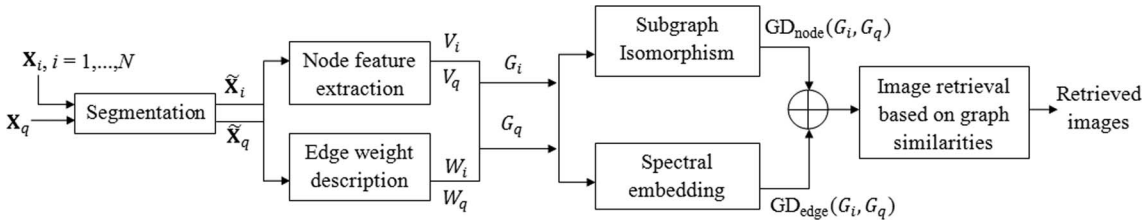

Fig. 1. Block diagram of the proposed graph-theoretic approach to region-based RS image retrieval.

query graph and the graphs of images in the archive are calculated using a novel inexact graph matching strategy that jointly consists of a subgraph isomorphism algorithm (which measures the similarity between graph nodes) and a spectral embedding algorithm (which measures the similarity between graph edges). The most similar images to the query image are then retrieved according to the ranking of graph similarities. The main contribution of the proposed approach is that, according to our knowledge, it is the first entirely unsupervised graph-theoretic approach presented for RS image retrieval problems.

The rest of this letter is organized as follows. Section II defines the considered problem and presents the proposed approach. Section III illustrates the benchmark archive and describes the experimental results. Finally, Section IV draws the conclusion of this letter.

II. Proposed Method

Let $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N\}$ be an archive of N RS images, where $\mathbf{X}_i$, $i = 1, \ldots, N$, is the ith image. Given a query image $\mathbf{X}_q$, we aim to retrieve the most similar images to $\mathbf{X}_q$ from the archive (i.e., images having a similar pattern of regions) without the need of any labeled sample. This is achieved here by a novel unsupervised approach consisting of two main steps: 1) image modeling by graphs and 2) image retrieval based on graph matching. Each step is explained in detail in the following. Fig. 1 shows the general block scheme of the proposed graph-theoretic region-based RS image retrieval approach.

A. Image Modeling by Graphs

This step aims to model each image by an ARG whose nodes and edges represent the image regions and their adjacency relationships, respectively. The ARG of an image $\mathbf{X}_i$ is defined as a tuple $G_i = (V_i, E_i, \mathbf{W}_i)$, where $V_i = \{v_i^1, v_i^2, \ldots, v_i^{n_i}\}$ is the set of n_i nodes, $E_i = \{e_i^{(s,t)} | s, t \in \{1, 2, \ldots, n_i\}\}$ is the set of edges that link the nodes, and $\mathbf{W}_i \in \mathbb{R}^{n_i \times n_i}$ is the weighted adjacency matrix containing edge information. An edge $e_i^{(s,t)}$ exists if the regions corresponding to nodes v_i^s and v_i^t are adjacent to each other.

To define the nodes, the images $\mathbf{X}_i$ and $\mathbf{X}_q$ are initially segmented into a set of disjoint regions to form the segmented images $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_q$, respectively. The task of unsupervised multiregion image segmentation is implemented by using the parametric kernel graph cut approach [7], which can adaptively split an image into semantically meaningful regions. Each region is then associated with a node represented by low-level features, such as intensity, texture, and shape features extracted from the regions to form the node attributes. While intensity features model the spectral distribution of samples (which is invariant to scale and viewpoint changes), texture features model intrinsic spatial structures and local patterns, whereas shape features model the inherent properties of objects (which are invariant to translation and rotation) in an image. The features used in our method and their definitions are given in Section III.

The spatial relationships between two adjacent nodes, such as spatial distance and orientation, are represented by edge attributes. To define the edges, we construct from $\tilde{\mathbf{X}}_i$ an adjacency matrix $\mathbf{W}_i$, whose initial entries are logical 0s and 1s, with $\mathbf{W}_i(v_i^s, v_i^t) = 1$ if $e_i^{(s,t)}$ exists and $\mathbf{W}_i(v_i^s, v_i^t) = 0$ otherwise. Each nonzero value of $\mathbf{W}_i$ is then assigned a weight as follows:

$$\mathbf{W}_i\left(v_i^s, v_i^t\right) = \alpha_1 \left\| c_{v_i^s} - c_{v_i^t} \right\|_2 + \alpha_2 \left| \theta_{v_i^s} - \theta_{v_i^t} \right| \quad (1)$$

where $c_{v_i^s}$ and $c_{v_i^t}$ are the centroids of the pixel coordinates within the regions corresponding to nodes v_i^s and v_i^t, respectively. $\theta_{v_i^s}$ and $\theta_{v_i^t}$ are the orientation angles of those regions (angle between the horizontal axis and the major axis of the ellipse having the same second moments as the region, such that $\theta \in [-90^\circ, +90^\circ]$), and $\| \cdot \|_2$ is the L_2 norm. The orientation angles are included for retrieval of images with high visual similarity, and the choice of using the orientation angles depends on the final goal of retrieval. The weights α_1 and α_2 should be selected after taking into account the importance of the related variables in the retrieval. We have used undirected weighted graphs, i.e., $\mathbf{W}_i(v_i^s, v_i^t) = \mathbf{W}_i(v_i^t, v_i^s)$; hence, $\mathbf{W}_i$ is a real symmetric matrix with all of the diagonal elements 0.

B. Image Retrieval Based on Graph Matching

This step is devoted to find the most similar images to the query image from the archive by matching the corresponding ARGs and evaluating related distances. In this letter, we introduce an inexact graph matching strategy that jointly exploits a subgraph isomorphism algorithm and a spectral embedding algorithm. The proposed strategy does not require identical numbers of nodes or strict correspondence of features between two graphs being matched. Such a strategy is proposed since most of the techniques for finding optimal solutions to weighted graph matching problems have very high computational complexity [6].

The subgraph isomorphism algorithm aims at the following: 1) matching the nodes between the graph G_q and each graph G_i in the archive and 2) estimating the distance between the graphs in terms of their node similarities. A graph G_i is said to be subgraph isomorphic to the graph G_q if a subgraph of G_i is isomorphic to G_q, i.e., there exists a mapping between their nodes. In order to find an accurate mapping, the proposed algorithm assesses the distance (dissimilarity) of each node v_q^r, $r \in \{1, 2, \ldots, n_q\}$ in G_q, to all of the nodes v_i^p, $p \in \{1, 2, \ldots, n_i\}$, in G_i to find the most similar (i.e., matching) node to v_q^r in G_i. To this end, we define a distance matrix $\mathbf{D} \in \mathbb{R}^{n_q \times n_i}$ containing the distances between all pairs of node attributes

of G_q and G_i, such that $\mathbf{D}(r,p)$ is the distance between the node attributes of v_q^r and v_i^p. Different measures are available in literature [8] to compare the distances between node attributes (see Section III for the measures that we have used in this letter). Each node v_q^r, $r \in \{1,2,\ldots,n_q\}$, is then matched with the node v_i^h, $h \in \{1,2,\ldots,n_i\}$, having the smallest $\mathbf{D}(r,h)$ value. Finally, image similarity is evaluated by calculating the graph node distance GD_{node} between G_q and G_i as follows:

$$GD_{\text{node}}(G_q,G_i) = \frac{1}{n_q}\sum_{r=1}^{n_q} \min_{\forall p\in\{1,2,\ldots,n_i\}} \{\mathbf{D}(r,p)\} \quad (2)$$

where $GD_{\text{node}}(G_q,G_i)$ is small if these graphs are similar to each other and vice versa. If $n_q > n_i$, there are $(n_q - n_i)$ nodes of G_q that cannot be matched with any node of G_i. In such cases, node matching should be achieved by estimating $GD_{\text{node}}(G_i,G_q)$ instead of $GD_{\text{node}}(G_q,G_i)$.

The spectral embedding algorithm aims at estimating the distance between two graphs in terms of their edge similarities. Spectral graph embedding involves embedding of graphs in a lower dimensional eigenspace using the eigenvectors of the Laplacian of graph adjacency matrices [9]. The normalized Laplacian of a graph G_i is defined as $\mathbf{L}_i = \mathbf{M}_i^{-1/2}(\mathbf{M}_i - \mathbf{W}_i)\mathbf{M}_i^{-1/2}$, where $\mathbf{M}_i \in \mathbb{R}^{n_i\times n_i}$ is a diagonal matrix such that each diagonal element of $\mathbf{M}_i$ is the sum of all of the elements in the corresponding row of $\mathbf{W}_i$. For G_q and G_i, the eigenvectors corresponding to the K smallest nonnull eigenvalues represent the directions of the principal components of the embedded graphs [9], where $K = \min(n_q,n_i)$. The edge similarity between G_q and G_i can be estimated as $\|\mathbf{P}\mathbf{W}_q\mathbf{P}^T - \mathbf{W}_i\|_F^2$ [9] in the case of $n_q = n_i$, where $\|\cdot\|_F$ is the Frobenius norm and $\mathbf{P}$ is a permutation matrix ranging over the set of all orthogonal matrices. According to [10], the expression is bounded as $\|\mathbf{P}\mathbf{W}_q\mathbf{P}^T - \mathbf{W}_i\|_F^2 \geq \sum_{e=1}^{n_q}(\lambda_q^e - \lambda_i^e)^2$, where λ_q^e and λ_i^e for $e = 1,2,\ldots,n_q$ are the distinct eigenvalues of $\mathbf{W}_q$ and $\mathbf{W}_i$, respectively. However, for large graphs and for $n_q \neq n_i$, the eigenvalues may have multiplicity greater than one (i.e., they are no more distinct), and the expression (or its lower limit mentioned previously) no longer holds [9]. In this case, the minimum edge distance between G_q and G_i is given by the sum of squares of the differences between the K smallest eigenvalues of $\mathbf{L}_q$ and $\mathbf{L}_i$ as

$$GD_{\text{edge}}(G_q,G_i) = \frac{1}{K}\sum_{e'=1}^{K}\left(\tilde{\lambda}_q^{e'} - \tilde{\lambda}_i^{e'}\right)^2 \quad (3)$$

where $\tilde{\lambda}_q^{e'}$ and $\tilde{\lambda}_i^{e'}$ for $e' = 1,2,\ldots,K$ are the K smallest eigenvalues of $\mathbf{L}_q$ and $\mathbf{L}_i$, respectively.

We propose to combine the node distance given by (2) and the edge distance given by (3) as their weighted sum to take into account the similarity both in region characteristics and in their relations. Accordingly, the total distance GD between G_q and G_i is obtained as

$$GD(G_q,G_i) = GD_{\text{node}}(G_q,G_i) + \beta GD_{\text{edge}}(G_q,G_i) \quad (4)$$

where $\beta \in (0,1]$ is the weighting factor, higher value of which gives more importance to the edge information of the graphs compared to the node information. On the basis of this combination, the value of $GD(G_q,G_i)$ is small if the images $\mathbf{X}_q$ and $\mathbf{X}_i$ are similar to each other and vice versa.

III. Experimental Analysis

Experiments have been carried out on an archive of 2100 images from 21 different categories (see Fig. 2) selected from aerial orthoimagery. For further information about the archive, we refer the reader to [1]. The category labels of the images have been used only to evaluate the retrieval performance (they have not been used during the image retrieval process).

A. Attribute Description and Experimental Settings

In our experiments, the regularization parameter required for the parametric kernel graph cut approach is experimentally set to 0.5. After segmenting each image, the regions with less than 10 pixels are initially merged with their adjacent largest regions, and then, morphological opening and closing operations are applied to fill in potential gaps and to smooth the shape contours. Then, each region is described by the following node attributes: 1) *shape features* [8] (which include Fourier descriptors and contour-based shape descriptors like area, convex area, perimeter, extent, solidity, and eccentricity derived from the region contours); 2) *intensity features* (which consist of the mean, standard deviation, and skewness of the samples within each region in each spectral channel); and 3) *texture features* (which include entropy and spectral histogram [11]). The spectral histogram is obtained by applying three types of filters: 1) the intensity filter; 2) the Laplacian of Gaussian filters; and 3) the Gabor filters to each image band. The intensity filter results in the original image itself, the Laplacian of Gaussian filters are obtained by selecting variances of 0.2 and 1, and Gabor filters are defined with orientations of 0°, 45°, 90°, and 135° when variances in both x and y directions are set to 2. In the experiments, the number of histogram bins is set to 10 for each spectral channel. All of the considered features are then concatenated to form a feature vector for each node. Accordingly, the final length of each node attribute is 232. The weights α_1 and α_2 in (1) are set to 0.8 and 0.2, respectively, in order to give more importance to centroid distances in modeling spatial information. The size of the graphs ranges from 5 to 413 depending on the complexity of regions and structures in an image. This variation in the number of graph nodes does not cause any problem because of normalization [see (2)]. The weighting factor β in (4) is set to 1 based on empirical results to give equal importance to the node and edge distances. Finally, according to the literature [8], we have considered different distance measures to evaluate the similarities between node attributes: *city block* distance is considered for Fourier descriptors and intensity features, whereas *Euclidean* distance is considered for entropy and contour-based shape features, and *chi-square* distance is selected for spectral histograms.

B. Results and Discussion

We have compared the proposed approach based on ARG modeling and matching (denoted as ARGMM in the experiments) with two state-of-the-art methods: 1) kNN algorithm applied to the BoVW representations of the local invariant features extracted by the SIFT [1] (denoted as SIFT-kNN) and 2) kNN algorithm applied to the image descriptors obtained based on the LBP [3] (denoted as LBP-kNN). To obtain the BoVW representations of images, k-means clustering is applied to the SIFT descriptors of 50 randomly selected images from

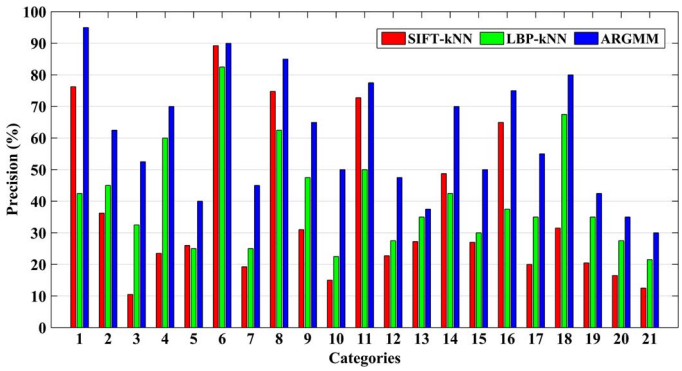

Fig. 2. Average precision (on 20 trials per category) for different categories obtained by the SIFT-kNN, the LBP-kNN, and the proposed ARGMM when the top 20 images are retrieved (1—agricultural, 2—airplane, 3—baseball diamond, 4—beach, 5—buildings, 6—chaparral, 7—dense residential, 8—forest, 9—freeway, 10—golf course, 11—harbor, 12—intersection, 13—medium residential, 14—mobile home park, 15—overpass, 16—parking lot, 17—river, 18—runway, 19—sparse residential, 20—storage tanks, and 21—tennis courts).

each category with $k = 232$. This choice is made to obtain the same length of feature vector as in the proposed method. To obtain image descriptors based on the LBP, we have considered eight neighboring pixels for each sample. The kNN algorithm is applied by using the chi-square distance measure in the SIFT-kNN and the G statistic similarity measure [3] in the LBP-kNN. Experimental results of each method are provided in terms of the following: 1) average precision on the top 20 retrieved images obtained in 20 trials for 20 randomly selected query images from each category; 2) average precision–recall (PR) graphs where precision is plotted as a function of recall; and 3) average normalized modified retrieval rank (ANMRR). Precision is an accuracy measure of a retrieval system given by the ratio of the number of relevant images retrieved to the number of retrieved images. Recall evaluates the ability of the retrieval system to find all of the relevant images in the archive and is defined as the fraction of relevant images that are retrieved. PR graphs can assess the retrieval performance at each point of ranking and are hence widely used in CBIR problems. For details on ANMRR, we refer the reader to [2].

Fig. 2 shows the average precision associated with each category obtained when the SIFT-kNN, the LBP-kNN, and the proposed ARGMM are used. By analyzing the figure, one can observe that the proposed ARGMM results in higher precision values than the SIFT-kNN and the LBP-kNN for all categories. For example, the proposed ARGMM yields a precision of 80% for the runway category, whereas the SIFT-kNN and the LBP-kNN provide only 31.5% and 67.5%, respectively, for the same category. The average precision values using the SIFT-kNN, the LBP-kNN, and the proposed ARGMM are 30.36%, 21.55% and 59.76%, respectively. The differences in the precision values are very high, particularly when the query images are selected from complex image categories, such as medium residential, mobile home park, dense residential, buildings and sparse residential categories (for which retrieval performances interfere with one another due to similar modeling of building structures), baseball diamond, golf course and tennis court categories (which interfere due to landscape similarity), and airplane and runway categories (which interfere due to the presence of common regions with highly similar characteristics in their images). The average ANMRRs of 21 categories using the SIFT-kNN, the LBP-kNN, and the proposed ARGMM are 0.7645, 0.8523 and 0.5748, respectively. All of these results show that the proposed region-based modeling approach is superior for image retrieval in the considered archive compared to the local descriptor-based approaches due to its ability to accurately model the regions and their topographic relationships.

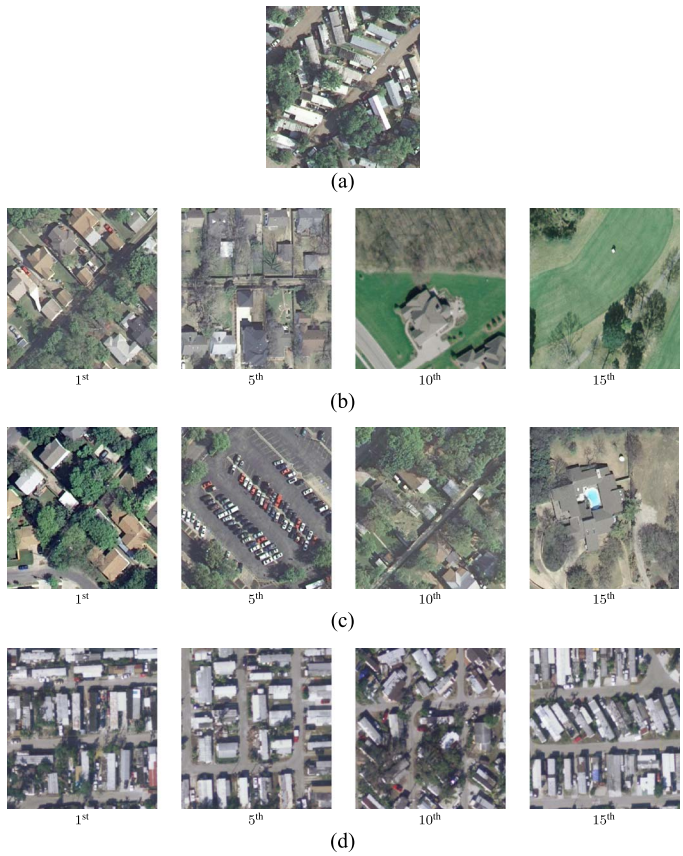

Fig. 3. Mobile home park image retrieval. (a) Query image. (b) SIFT-kNN (retrieval accuracy—48.75%). (c) LBP-kNN (retrieval accuracy—42.5%). (d) Proposed ARGMM (retrieval accuracy—70%).

Fig. 3 shows an example of retrieval results with the corresponding query image selected from the mobile home park category. The retrieval order of each image is given below the related image. From the results, one can see that all of the images retrieved by the proposed ARGMM belong to the mobile home park category [see Fig. 3(d)] and the average precision on the top 20 retrieved images is 70%. On the contrary, not all of the images retrieved by the SIFT-kNN [see Fig. 3(b)] and the LBP-kNN [see Fig. 3(c)] are related to the mobile home park category. For example, the tenth images retrieved by the SIFT-kNN and the LBP-kNN belong to the sparse residential and medium residential categories in the archive, respectively. For these cases, the average precision values on the top 20 retrieved images obtained by the SIFT-kNN and the LBP-kNN are 48.75% and 42.5%, respectively. Fig. 4 shows the behavior of the average PR graphs obtained by the SIFT-kNN, the LBP-kNN, and the proposed ARGMM for the mobile home park category. From Fig. 4, it is evident that the precision-to-recall ratio is higher for the ARGMM compared to the SIFT-kNN and the LBP-kNN. It is worth noting that we also obtained similar behaviors of PR graphs for other categories of images.

Finally, we have analyzed the effectiveness of the proposed approach by selecting a query image acquired by a different sensor having the same spectral characteristics of the images in the

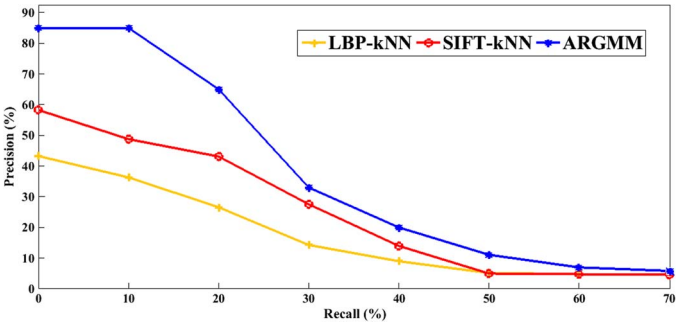

Fig. 4. Precision versus recall graphs obtained by the SIFT-kNN, the LBP-kNN, and the proposed ARGMM for the mobile home park category.

considered archive. To this end, we selected a patch from Google Earth and observed that our approach retrieves very similar images to the query image (due to space limitations, we could not provide them here). These results show that the proposed approach works well as long as the spatial and spectral characteristics of the sensors used to acquire the images remain similar.

The overall computational complexity of the proposed ARGMM is associated to the i) computational time required for graph construction, which depends on the size of the images and the number of spectral bands, and ii) computational time required for image retrieval through graph matching, which depends on both the number of nodes and edges that model the graphs and the number of images in the archive. In our experiments, the average computational time required to construct a graph from an image is about 30 s (which is made up of 15 s required for segmentation, 13 s for node modeling, and 2 s for edge modeling). It takes an average of 81.75 s to match a graph with all of the 2100 graphs in the archive. All of the experiments are implemented via MATLAB on a standard personal computer with Intel Core 2.93 GHz i7 processor and 8-GB RAM. From the experiments, we observed that graph construction and matching in the proposed ARGMM requires higher computation time compared to feature extraction and matching in the SIFT-kNN and the LBP-kNN. However, this time can be significantly reduced by using scalable approximate graph matching algorithms. For example, the random walk algorithm for approximate graph matching [12] has shown promising results by achieving a reduction of about five order of magnitude in execution time on an average. We can also exploit the high redundancy in the graph adjacency matrices and approximate them using compressed representations as described in [13] while per forming spectral embedding. Another method is to use cluster-based parallel computing approach, in which the query graph is initially matched with the graph clusters in order to filter out the largely irrelevant images. Then, a finer search is performed within the most similar cluster to retrieve the desired images.

IV. Conclusion

In this letter, a novel unsupervised graph-theoretic approach has been presented for region-based RS image retrieval. In the proposed approach, the images are modeled as attributed relational graphs and then the graphs of the images in the archive are matched to that of the query image by a novel inexact graph matching strategy. Finally, the most similar images to the query image are ranked and retrieved on the basis of the graph similarity values.

Experimental results show that the proposed approach, when compared to the state-of-the-art unsupervised RS image retrieval methods, allows one to sharply improve the retrieval performance particularly for the complex image categories (due to the effective and accurate characterization of regions and their associations present in the images). However, the improvement in performance is obtained at the cost of slightly increased computational complexity. It is worth noting that the proposed graph-theoretic approach is independent of the selected segmentation algorithm and the considered features used to characterize the images. It is also intrinsically independent of the type of the RS images present in the archive and can be easily adapted to the retrieval problems of multidimensional RS images. This can be done by selecting suitable features for accurately characterizing the images associated to the different spatial resolutions and spectral characteristics, and then using them in the framework of the proposed approach.

As a future development of this work, we are planning to apply fast parallel algorithms in the context of approximate graph matching to speed up the process. In addition, we are planning to include automatic feature learning and feature weighting to further improve the characterization of the image regions.

Acknowledgment

This work was carried out under the India-Trento Program for Advanced Research (ITPAR) during the stay of B. Chaudhuri at the University of Trento, Trento, Italy.

References

[1] Y. Yang and S. Newsam, "Geographic image retrieval using local invariant features," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 818–832, Feb. 2013.

[2] E. Aptoula, "Remote sensing image retrieval with global morphological texture descriptors," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 3023–3034, May 2014.

[3] M. Musci, R. Q. Feitosa, G. A. O. P. Costa, and M. L. F. Velloso, "Assessment of binary coding techniques for texture characterization in remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1607–1611, Nov. 2013.

[4] M. Wang and T. Song, "Remote sensing image retrieval by scene semantic matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2874–2886, May 2013.

[5] S. Aksoy, "Modeling of remote sensing image content using attributed relational graphs," in *Proc. IAPR Int. Workshop Struct. Syntactic Pattern Recognit.*, Aug. 2006, pp. 475–483.

[6] P. Foggia, G. Percannella, and M. Vento, "Graph matching and learning in pattern recognition in the last 10 years," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 28, no. 1, 2014, Art. no. 1450001.

[7] M. Ben Salah, A. Mitiche, and I. B. Ayed, "Multiregion image segmentation by parametric kernel graph cuts," *IEEE Trans. Image Process.*, vol. 20, no. 2, pp. 545–557, Feb. 2011.

[8] F. Long, H. Zhang, and D. D. Feng, "Fundamentals of content-based image retrieval," in *Multimedia Information Retrieval and Management*. Berlin, Germany: Springer-Verlag, 2003, pp. 1–26.

[9] D. Knossow, A. Sharma, D. Mateus, and R. Horaud, "Inexact matching of large and sparse graphs using Laplacian eigenvectors," in *Graph-Based Representations in Pattern Recognition*. Berlin, Germany: Springer-Verlag, 2009, pp. 144–153.

[10] S. Umeyama, "An eigen decomposition approach to weighted graph matching problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 5, pp. 695–703, Sep. 1988.

[11] X. Liu and D. Wang, "Texture classification using spectral histograms," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 661–670, Jun. 2003.

[12] M. Gori, M. Maggini, and L. Sarti, "Exact and approximate graph matching using random walks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 7, pp. 1100–1111, Jul. 2005.

[13] U. Kang, M. Hebert, and S. Park, "Fast and scalable approximate spectral graph matching for correspondence problems," *Inf. Sci.*, vol. 220, no. 20, pp. 306–318, Jan. 2013.