RESEARCH ARTICLE OPEN ACCESS

Mono-Class Oriented Feature Selection and Categorization of Heterogeneous Spatial Images with Background Labels

Buvaneswari B [1], Harini S [2], Kalaivani N [3], Manasa M.S [4]
Associate Professor [1], Student [2], [3] & [4]
Department of Information Technology
Panimalar Engineering College, Chennai
Tamil Nadu - India

ABSTRACT

Performance of content-based image retrieval (CBIR) schemes are mainly improved by Conventional relevance feedback (RF) schemes which demands a large number of image annotation. In this work we present a novel active learning (AL) method to drive RF for retrieving satellite images from large archives with framework of the support vector machine classifier to reduce the labeling effort of the user. The proposed AL method is specifically designed for CBIR with the aim of evaluating threes criteria namely 1) uncertainty; 2) diversity; and 3) density of images in the archive in a two successive steps. The first step involves margin sampling followed by clustering in second step. Experimental results show the effectiveness of the proposed work.

Keywords :- Active learning (AL), content-based image retrieval (CBIR), relevance feedback (RF), remote sensing (RS).

I. INTRODUCTION

With the development of satellite technology, large volume remote sensing (RS) images (i.e., millions of single date as well as time series of Earth observation scenes) become available. Accordingly, one of the most challenging and emerging applications in RS is the efficient and precise retrieval of RS images from such archives according to the users' needs. Conventional RS image retrieval systems often rely on keywords/tags in terms of sensor type, geographical location, and data acquisition time of images stored in the archives. The performance of tag matching-based retrieval approaches highly depends on the availability and the quality of manual tags. However, in practice, keywords/tags are expensive to obtain and often ambiguous. Due to these drawbacks, recent studies have shown that the content of the RS data is more relevant than manual tags.

Accordingly, content-based image retrieval (CBIR) has attracted increasing attentions in the RS community particularly for its potential practical applications to RS image management. This will become particularly important in the next years when the number of acquired images will dramatically increase. Any CBIR system essentially consists of (at least) two modules [1], [2]: 1) a feature extraction module that derives a set of features for characterizing and describing images and 2) a retrieval module that searches and retrieves images similar to the query image. Querying image contents from large RS data archives depends on the capability and effectiveness of the feature extraction techniques in describing and representing the images. The remaining part of this paper is organized as follows.

Section II introduces the existing literature works. Section III describes the proposed AL method, whereas IV illustrates the experimental results. Finally, Section V draws the conclusion of this work.

II. RELATED WORKS

In the RS literature, several primitive (i.e., low level) features have been presented for retrieval purposes, such as the following: intensity features [5], color features [6], [7], shape features [8]–[10], texture features [10]–[16], and local invariant features [17]. However, the low level features from an image have a very limited capability in representing and analyzing the high-level concept conveyed by RS images (i.e., the semantic content of RS images). This issue is known as the semantic gap that occurred between the low level features and the high-level semantic content and leads to poor CBIR performance. Consequently, the semantic gap is the crucial challenge in CBIR applications.

In order to confine the semantic gap, relevance feedback (RF) schemes have been designed to iteratively improve the performance of CBIR by taking user's (i.e., an oracle who knows the correct labeling of all images) feedback into account[3], [4]. At each iteration, the user's feedback is used to provide relevant and irrelevant images to the query image that are positive and negative feedback samples, respectively. RF can be considered as a binary-classification problem: One class includes relevant images, and the other one consists of their relevant ones. Then, any supervised classification method can be used in the context of CBIR by training the classifier with the already annotated images of two classes [3], [4]. Accordingly, during RF, the search strategy is refined iteration by iteration by improving the classification model with the recently annotated images. As mentioned previously, user involvement is required at each RF iteration for annotating images. However, labelling images as relevant or irrelevant is time-consuming and thus costly. Accordingly, despite the retrieval success of RF, the conventional RF schemes are not practical and efficient in real applications, especially when huge archives of RS images are considered.

An effective approach to reduce the annotation effort in RF is active learning (AL) that aims at finding the most informative images in the archive that, when annotated and included in the set of relevant and irrelevant images (i.e., the training set), can significantly improve the retrieval performance [10]. Moreover, selecting the most informative images results in the following: 1) a smaller number of RF iterations to optimize the CBIR and 2) a reduced annotation time due to the optimization of the training set with a minimum number of highly informative images. In the RS community, most of the previous studies in AL have been developed in the context of classification problems for land-cover map generation (see [18] for a comprehensive review on the most relevant techniques). In particular, the unlabeled samples that are highly uncertain and diverse to each other are usually selected as informative samples to be labeled and included in the training set for the classification of RS images [18]. The uncertainty of a sample is related to the confidence of the supervised algorithm in correctly classifying it, whereas the diversity among samples is associated to their correlation in the feature space (i.e., samples that are as distant as possible to each other are the most diverse samples).

AL has been marginally considered in the framework of CBIR problems in the RS community. To the best of our knowledge, only one AL method is presented [10], which is developed in the context of the support vector machine (SVM) classifier and inspired from AL methods used for classification problems [21]. In this method, the uncertainty and diversity criteria have been applied in two consecutive steps. In the first step, the most uncertain images are selected from the archive. To this end, the unlabeled images closest to the current separating hyperplane(those that are the most uncertain) are initially selected by margin sampling (MS) [19], [20]. In the second step, the images that are diverse to each other among the uncertain ones are chosen on the basis of the distances estimated between them. An important shortcoming of the method presented in [10] is that it does not evaluate the representativeness of images in terms of their density in the archive. However, images that fall into the high-density regions of the image feature (descriptor)space are crucial for CBIR problems particularly when a small number of initially annotated images are available. This is due to the fact that they are statistically very representative of the underlying image distribution in the archive. Therefore, the retrieval results on them affect much more the overall retrieval density regions.

III. PROPOSED METHOD

A. Problem Formulation

Let us consider an archive ϒ made up a very large number of R RS { X_1, X_2, ... X_R} images. A general CBIR system with RF driven by AL consists of three modules: 1) the primitive (low level) feature extraction module that is applied to both query image and all images in the archive;2) the initial training set definition module that builds an initial training set T with a small number of relevant and irrelevant images with respect to query; and 3) the RF driven by an AL module that enriches the training set T defined by the previous module and returns the set δ of images from the archive Y. Fig. 1 shows the general block scheme of the CBIR with RF driven by AL. In this paper, we mainly focus on the RF driven by the AL module (see Fig. 2) which is a crucial part for the success of the CBIR system. Then, we briefly present the feature extraction module and the important choices adopted for assessing the similarities of image features in the proposed system.

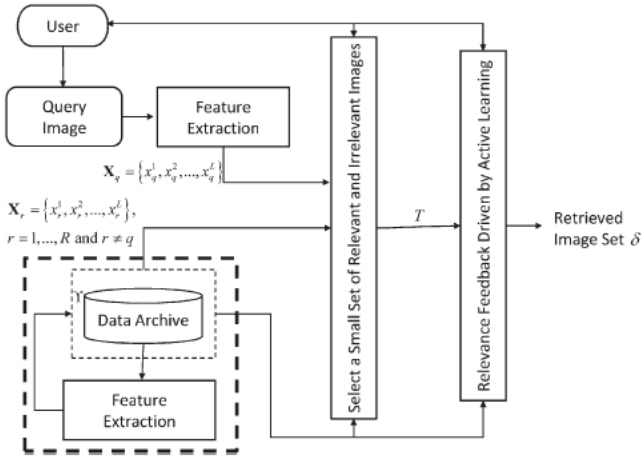

Fig. 1 General architecture of a CBIR system with RF driven by AL.

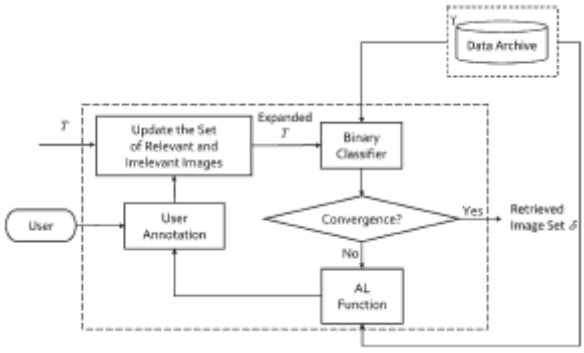

Fig. 2 General flowchart of RF driven by AL.

AL iteratively expands the size of an initial labeled training set T, selecting the most informative images from the archive Y for their annotation. At each RF iteration, the most informative unlabeled images for a given classifier are: 1) selected based on an AL function; 2) annotated by a supervisor (i.e., an oracle); and 3) added to the current training set T. Finally, the supervised classifier is retrained with the images moved from Y to T. It is worth noting that the initial training set T requires few annotated images for the first training of the classifier and then is enriched iteratively by including the most informative images selected from Y. At each iteration, after the classifier is trained, the retrieval of the images under investigation is carried out. These processes are repeated until the user is satisfied with retrieval results. The general flowchart of the AL-based RF approach is given in Fig. 2. The selection of the most informative samples from Y

to be included in the training set T on the basis of AL offers two main advantages: 1) The annotation cost is reduced due to the avoidance of redundant images, and 2) an accurate retrieval accuracy can be obtained due to the improved class models estimated on a high-quality training set on the basis of the classification rule used from the considered classifier (the images to be annotated are selected from the classifier as the most informative for its classification rule). Of course, the success of the RF strongly depends on the capability of the specific AL method considered to select the most informative and representative images to be annotated in order to limit as much as possible the effort of the user for reaching the final relevant result.

B. Proposed AL Method

We propose a novel triple criteria AL (TCAL) method to expand the initial training set during RF rounds in CBIR applications.The aims of the proposed AL method are as follows: 1) to achieve a training set of annotated relevant and irrelevant images with respect to the query image as small as possible within a low number of RF iterations and 2) to retrieve the images similar to the query image with high accuracy. The proposed TCAL method is defined in the context of binary SVM classification and selects a batch $S = \{X_1, X_2, \ldots, X_h\}$ of h images at each RF iteration that are as follows: 1) uncertain(i.e., ambiguous); 2) as more diverse as possible to each other; and 3) located in the highest density regions of the image feature space. The uncertainty of images is assessed according to the MS strategy, whereas the diversity and density of the image are evaluated by a novel clustering-based strategy. At each iteration, the proposed AL method jointly evaluates the aforementioned three criteria by a strategy that is based on two consecutive steps to select the batch S of images.

In the first step, the $m > h$ most uncertain images are selected according to the standard MS technique from Y. In the second step, the most diverse h images among these m uncertain (i.e., ambiguous) images are chosen from the highest density regions of the feature space ($m > h > 1$). The first step is devoted to select unannotated images that have maximum uncertainty on their correct target classes according to the binary SVM classification properties. The basic idea behind this concept is that images, which have the lowest probability to be accurately classified by the considered classifier, are the most beneficial to be included in the training set for separating the two categories of relevant and irrelevant images in an optimal way. For SVM Classification, the images closest to the separating hyperplane (which is the discriminant function) have low confidence to be correctly classified. One of the most popular AL methods in the context of SVM classification is MS, which selects the unlabeled samples closest to the separating hyperplane, as they are the samples considered with the lowest confidence (i.e.those that have the maximal uncertainty on the true information class). Accordingly, we considered this approach to select the most uncertain images in the first step due to its simplicity and effectiveness, and possible fast implementation. To this end, initially, a binary SVM is trained using the existing set of relevant and irrelevant images. Then, the functional distances of the unannotated images to the current SVM hyperplane are estimated.

The second step is devoted to select h images from the uncertain set of the most uncertain images that are diverse to each other, taking into account sample density in the archive image feature space. Selecting the uncertain images from the high-density regions of the image feature space is crucial in the proposed TCAL method. This is due to the fact that, under the reasonable assumption that images in the same region of the image feature space have similar target categories, the selection of images to be annotated from high-density regions is an effective strategy for minimizing the overall retrieval error. This choice aims to reduce errors in the regions of the image feature space where we have many uncertain unlabeled images that can strongly affect the overall retrieval accuracy.

C. RS Image Feature Extraction and Classification in the Context of CBIR

We model RS images by exploiting a bag-of-visual-words (BOVW) representation of the local invariant features extracted by the scale invariant feature transform (SIFT). The SIFT is a translation, rotation, and scale invariant image feature extraction technique and has recently been found very effective and robust in the context of RS image retrieval [17]. The SIFT results in various local interest points within an image and their descriptors (i.e., SIFT descriptors) that characterize portions of images around the interest points. In order to summarize the SIFT descriptors by the BOVW representation (that is generally considered for the local image descriptors), we apply kernel k-means clustering to a subset of randomly selected SIFT descriptors. This process results in a codebook. Then, the descriptors extracted from each image are quantized by assigning the label of the closest cluster [17]. Accordingly, the final representation of an image is the histogram (i.e., frequency) of the codebook entries (known as code-words) in the image [17]. Note that the histogram-based image representation is very popular for the BOVW approaches that result to be the state-of-the-art in many image retrieval problems outside RS.

IV. EXPERIMENTAL EVALUATION

The proposed work is implemented in MATLAB and performance evaluation is done database containing 21 object categories. In the experiments, in order to obtain the BOVW representations of images (which summarizes the SIFT descriptors), kernel k-means clustering was applied to 100 000 randomly selected SIFT descriptors by selecting $k = 150$. Then, the SIFT descriptors are quantized by assigning the label of the closest cluster. The images downloaded from the National Map are in the red–green–blue color space. In order to use SIFT, a coherent way with [17], each image is converted to grayscale. In the experiments, L2 normalized SIFT histogram features have been used, i.e., the components

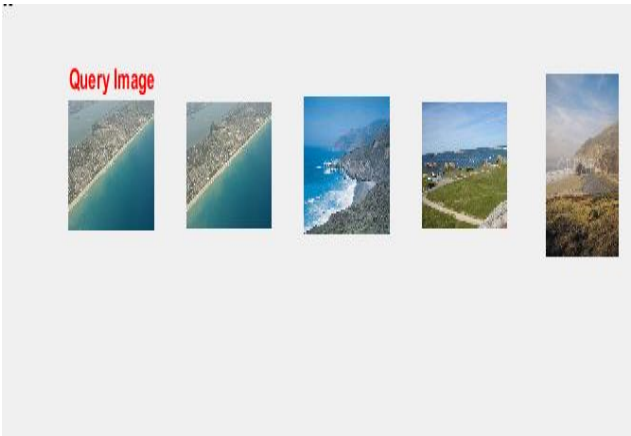

Fig. 3 Results obtained in implementation of proposed work

are normalized so that the feature vectors have length one.

V. CONCLUSION

The proposed AL method selects both informative and representative unlabeled images to be included in the training set at each RF round by the joint evaluation of the uncertainty, diversity, and density criteria. The uncertainty and diversity criteria aim to select the most informative images, whereas the density criterion aims to select the most representative images in terms of prior distribution. In the proposed AL method, the joint assessment of the three criteria is accomplished based on a two-step technique. In the first step, using marginal strategy approach the most uncertain images are selected. In the second step, high density region in image feature space are identified using a novel clustering approach, with these high density region the most diverse images among uncertain images are identified. The limitations in previously presented AL methods in CBIR problems due to 1) unbalanced training sets and 2) biased initial training set are eliminated in this proposed work.

REFERENCES

[1] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content,based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[2] R. Datta, D. Joshi, J. Li, and J.-Z. Wang, "Image retrieval: Ideas, influences, trends of the new age," *ACM Comput. Surveys*, vol. 40, no. 2, pp. 1–60, Apr. 2008.

[3] P. Hong, Q. Tian, and T. S. Huang, "Incorporate support vector machines to content-based image retrieval with relevant feedback," in *Proc. IEEE Int. Conf. Image Process.*, Vancouver, BC, Canada, 2000, pp. 750–753.

[4] X. S. Zhou and T. S.Huan, "Relevance feedback in image retrieval:Acomprehensive review," *Multimedia Syst.*, vol. 8, no. 6, pp. 536–544, Apr. 2003.

[5] Q. Bao and P. Guo, "Comparative studies on similarity measures for remote sensing image retrieval," in *Proc. IEEE Int. Conf. Syst.*, Man Cybern., Hague, The Netherlands, 2004, pp. 1112–1116.

[6] T. Bretschneider, R. Cavet, and O. Kao, "Retrieval of remotely sensed imagery using spectral information content," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Toronto, ON, Canada, 2002, pp. 2253–2255.

[7] T. Bretschneider and O. Kao, "A retrieval system for remotely sensed imagery," in *Proc. Int. Conf. Imag.* Sci., Syst., Technol., Las Vegas, NV, USA, 2002, pp. 439–445.

[8] G. Scott, M. Klaric, C. Davis, and C.-R. Shyu, "Entropy-balanced bitmap tree for shape-based object retrieval from large-scale satellite imagery databases," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1603–1616, May 2011.

[9] A. Ma and I. K. Sethi, "Local shape association based retrieval of infrared satellite images," in *Proc. IEEE Int. Symp. Multimedia*, Irvine, CA, USA, 2005, pp. 551–557.

[10] M. Ferecatu and N. Boujemaa, "Interactive remote-sensing image retrieval using active relevance feedback," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 818–826, Apr. 2007.

[11] Y. Li and T. Bretschneider, "Semantics-based satellite image retrieval using low-level features," in *Proc. IEEE Int. Geosci.* Remote Sens. Symp., Anchorage, AK, USA, 2004, vol. 7, pp. 4406–4409.

[12] Y. Hongyu, L. Bicheng, and C. Wen, "Remote sensing imagery retrieval based-on Gabor texture feature classification," in *Proc. Int. Conf. Signal Process.*, Troia, Turkey, 2004, pp. 733–736.

[13] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

[14] S. Newsam, L.Wang, S. Bhagavathy, and B. S.Manjunath, "Using texture to analyze and manage large collections of remote sensed image and video data," J. Appl. Opt., vol. 43, no. 2, pp. 210–217, Jan. 2004.

[15] A. Samal, S. Bhatia, P. Vadlamani, and D. Marx, "Searching satellite imagery with integrated measures," Pattern Recognit., vol. 42, no. 11, pp. 2502–2513, Nov. 2009.

[16] 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.

[17] L. Bruzzone, C. Persello, and B. Demir, "Active learning methods inclassification of remote sensing images," in Signal and Image Processing for Remote Sensing, 2nd ed, C. H. Chen, Ed. Boca Raton, FL, USA: CRC Press, 2012, ch. 15, pp. 303–323.

[18] G. Schohn and D. Cohn, "Less is more: Active learning with support vector machines," in *Proc. 17th Int. Conf. Mach. Learn.*, Stanford, CA,USA, 2000, pp. 839–846.

[19] S. Tong and D. Koller, "Support vectormachine active learning with applications to text classification," J. Mach. Learn. Res., vol. 2, pp. 45–66, 2001.

[20] K. Brinker, "Incorporating diversity in active learning with support vector machines," in *Proc. Int. Conf. Mach. Learn*., Washington, DC, USA, 2003, pp. 59–66.

[21] R. Zhang and A. I. Rudnicky, "A large scale clustering scheme for kernel k-means," in *Proc. IEEE Int. Conf. Pattern Recog*., Quebec, QC, Canada, 2002, pp. 289–292.

[22] B. Scholkopf, A. Smola, and K. R.Muller, "Nonlinear component analysis as a kernel eigenvalue problem," Neural Comput., vol. 10, no. 5, pp. 1299–1319, Jul. 1998.

[23] C. D. Manning, P. Raghavan, and H. Schutze, Introduction to Information Retrieval, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[24] A. Barla, F. Odone, and A. Verri, "Histogram intersection kernel for image classification," in *Proc. IEEE Int. Conf. Image Process*., Barcelona, Spain, 2003, pp. III-513–III-516.

[25] S. Maji, A. C. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," presented at the *IEEE Conf. Comput. Vis. Pattern Recog*., Anchorage, AK, USA, 2008, pp. 1–8