A Three-Layered Graph-Based Learning Approach for Remote Sensing Image Retrieval

Yuebin Wang, Liqiang Zhang, Xiaohua Tong, Liang Zhang, Zhenxin Zhang, Hao Liu, Xiaoyue Xing, and P. Takis Mathiopoulos, Senior Member, IEEE

Abstract—With the emergence of huge volumes of high-resolution remote sensing images produced by all sorts of satellites and airborne sensors, processing and analysis of these images require effective retrieval techniques. To alleviate the dramatic variation of the retrieval accuracy among queries caused by the single image feature algorithms, we developed a novel graph-based learning method for effectively retrieving remote sensing images. The method utilizes a three-layer framework that integrates the strengths of query expansion and fusion of holistic and local features. In the first layer, two retrieval image sets are obtained by, respectively, using the retrieval methods based on holistic and local features, and the top-ranked and common images from both of the top candidate lists subsequently form graph anchors. In the second layer, the graph anchors as an expansion query retrieve six image sets from the image database using each individual feature. In the third layer, the images in the six image sets are evaluated for generating positive and negative data, and SimpleMKL is applied to learn suitable query-dependent fusion weights for achieving the final image retrieval result. Extensive experiments were performed on the UC Merced Land Use–Land Cover data set. The source code has been available at our website. Compared with other related methods, the retrieval precision is significantly enhanced without sacrificing the scalability of our approach.

Index Terms—Expansion query, graph-based learning, image retrieval, query fusion, reranking

I. INTRODUCTION

With the availability of huge volumes of high-resolution remote sensing images produced by all sorts of satellites and airborne sensors, remote sensing image processing and analysis have been an active research topic in the geoscience field [1]–[4] in recent decades. As a result, many applications involving high-resolution satellite image analysis are in need of an image retrieval task. Due to the huge number of images, it is important to develop effective and efficient methods for searching, retrieving, and recognizing candidates within the expanding collections.

Manuscript received October 25, 2015; revised April 21, 2016; accepted June 6, 2016. This work was supported by the National Natural Science Foundation of China under Grant 41371324.

Y. Wang, L. Zhang, L. Zhang, Z. Zhang, H. Liu, and X. Xing are with the State Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China (e-mail: zhangql@bnu.edu.cn; xxgcdxywb@163.com; yammer_0303@126.com; zhenxin066@163.com; apprentice_g@163.com; xysxjing@mail.bnu.edu.cn).

X. Tong is with the School of Surveying and Geo-informatics, Tongji University, Shanghai 200092, China (e-mail: xhtong@tongji.edu.cn).

P. T. Mathiopoulos is with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, 15784 Athens, Greece (e-mail: mathio@di.uoa.gr).

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/TGRS.2016.2579648

A necessity for developing a successful remote sensing image retrieval system is the extraction of representative features to describe the query image and the ones in the database [36]. Local and holistic features are two types of descriptors for representing images. Local features are particularly capable of attending to local image patterns or textures, whereas holistic features describe the overall layouts of an image. One drawback of the two features is that the retrieved images often look alike but may be irrelevant to the query because remote sensing images often represent large natural geographical scenes that contain abundant and complex visual contents [35]. They are usually hindered by the noise from irrelevant content [7]. Clearly, integration of their strengths can greatly improve retrieval precision. Directly combining different feature vectors into one vector is not a good idea for improving image retrieval accuracy because the feature characteristics and the algorithmic procedures are dramatically different [6]. Although query expansion [12], [13] can achieve precise retrieval results, the performance of query expansion tends to degrade because of false-positive search results [5].

Inspired by the strengths of the query expansion and specific rank fusion method [6], we propose a three-layered graph-based learning approach to retrieve remote sensing images. Fig. 1 illustrates the overview of our proposed approach. In this approach, we extract the holistic feature by means of GIST and the local feature through dense SIFT and ScSPM [43]. Next, the image retrieval process is divided into three layers. In the first layer, by using the holistic feature learning algorithm, one image list in which every image is similar to the query is retrieved from the remote sensing image database. Accordingly, we obtain another image list by using the local feature learning algorithm. In the second layer, we first rerank the images in the aforementioned two lists and then obtain three types of graph anchors: \( P_H \), \( P_L \), and \( P_C \). \( P_H \) and \( P_L \) are the top-ranked images of the two lists. \( P_C \) contains common similar images of the two lists. We take \( P_H \), \( P_L \), and \( P_C \) as the queries for retrieving database images by means of the holistic feature or the local feature learning algorithm. Thus, six lists containing retrieved images can be obtained. In the third layer, positive and negative sets are selected by evaluating the images in the six lists. The parameters in SimpleMKL [45] are trained to fuse retrieval results, and we obtain the final retrieved images.

The main contributions of this paper are summarized as follows.

1) A novel three-layered graph-based learning approach for remote sensing image retrieval is developed. The
Fig. 1. Overview of the proposed approach.

approach refines the original input query by combining it with the obtained top-ranked retrieved images. The quality of the retrieval results obtained by using local or holistic feature methods is verified, and the ordered retrieval image sets are fused to generate the final retrieval result. The retrieval precision is significantly enhanced without sacrificing the scalability of the proposed method.

2) A new expansion query method that is robust to unstable feature extraction is introduced. Its main advantage is that it mines multiple relevant images by a single input image rather than requiring users to input multirelevant images. Combining the retrieved images, along with the original query, a richer expansion query image set is formed as prior knowledge for the next retrieval. By means of the set, the accuracy of image retrieval can be improved. The proposed method can eliminate the deficiency of the single image-based retrieval.

3) A novel approach for fusing different image retrieval results is presented to allow accurate evaluation of the quality of each return. SimpleMKL is applied to learn suitable query-dependent fusion weights for holistic and local features. The various image retrieval results are fused to enhance the retrieval accuracy.

II. RELATED WORK

In this section, we will present the related work about feature extraction, similarity measure metrics, and image retrieval in the following sections.

A. Feature Representation and Fusion

Over the last two decades, a variety of feature descriptions has been proposed [14]–[23]. They are generally divided into two categories: holistic and local features. Holistic features often represent the entire feature distribution of images, e.g., extended Gaussian images (EGIs) [14], texture features [15], [16], complex EGI [17], and spherical attribute images [18]. The global morphological texture descriptors, including circular covariance histograms (CCHs) and the rotation-invariant point triplets (RITs) [19], have been developed to retrieve content-based images. The texture descriptors in [19] are then adapted to a local scale by means of the popular bag-of-visual-words paradigm [8]. The performance of such holistic feature descriptors is usually affected by irrelevant parts of the image, thus making it sensitive to noises. On the other hand, the local feature extraction strategy avoids this problem by focusing on detecting and describing salient image regions/points [19]. The spin image descriptor [20] combines the descriptive nature of the global object properties with robustness to partial views and clutters of the local shape descriptions, and it is widely used in classification and recognition. In [21] and [22], quantized local features are employed to detect complex geospatial objects such as nuclear and coal power plants in 1-m resolution digital globe imagery. Yang and Newsam [23] utilized local invariant features for content-based geographic image retrieval, and they concluded that the local invariant features are more effective than features such as color and texture for image retrieval of land-use/land-cover (LULC) classes. Aiptoula et al. [8] later extended the low-level feature analysis in [23] to the “bag of morphological words” approach for content-based geographical retrieval. In scene classification, dense low-level feature descriptors can also be extracted to characterize the local spatial patterns for aerial scene classification [9].

Clearly, combining the complementary advantages offered by both local and holistic features can enhance the overall retrieval precision. However, combining local and holistic cues at the feature level makes it difficult to preserve the efficiency and scalability induced by the vocabulary tree structure and compact hashing [6]. For this reason, some relevant publications mainly focused on one reranked retrieval method [24] or text-based retrieval approach [25]. Xia et al. [26] proposed a class-specific hypergraph to integrate local features and global geometric constraints for object recognition. In a very recent publication [6], a graph-based query-specific fusion approach for fusing the retrieval results based on local and holistic
features was proposed. Although this method has high computational efficiency and improves the image retrieval precision, the retrieval results are significantly affected by the quality of the single input query image.

B. Similarity Measure Approaches

Similarity measure is often used on the extracted features to identify the images similar to the query [19]. Chen et al. [27] proposed a visual similarity-based 3-D model retrieval system using the L1 distance of features to measure model similarity. Zhang et al. [28] presented an upright-based normalization method, which accelerates the matching of the shape descriptors by correctly rotating these building models. In their approach, they also calculated the L1 distance of the features to measure the similarity of the building models. Kang et al. [29] formulated the image similarity assessment in terms of sparse representation. To evaluate the applicability of the proposed feature-based sparse representation using an image similarity assessment approach, they applied FSRISA to three popular image applications (namely, copy detection, retrieval, and recognition) by properly formulating them to sparse representation problems. In [5], a voting-based method was employed to calculate the similarity measure for object retrieval. Note that the similarity measure computation in this work only considered the matched feature pairs with spatial consistency. In [30], an endmember-based distance measure, particularly adaptive for hyperspectral image retrieval, was presented for image retrieval. In another contribution [31], a probabilistic method was used to match multiple views in object retrieval. However, in this approach, the Hausdorff distance fails to describe three or more objects that look similar because the distance only reflects the information of the closest views of a pair of objects. When comparing two objects, the similarity is measured by summing up the similarity from all of the corresponding images in [32]. Li and Fonseca [33] applied different priorities to topology, direction, and distance similarities for qualitative spatial similarity assessment, where they considered commonality and differences in their similarity assessment studies.

C. Image Retrieval

In this section, we briefly review the subjects of image matching and retrieval related to our paper. In [34], a content-based image retrieval system was presented, which aimed at classifying and retrieving oceanic scenes from satellite images. In [35], semantic matching of multilevel image scenes was utilized to retrieve remote sensing images. In [36], the authors described a content-based shape retrieval of objects from a large-scale satellite imagery database. Li and Bretschneider [37] proposed a context-sensitive Bayesian network for semantic inference of the segmented scenes. The semantics were employed for extracting candidate scenes, which were evaluated and ranked in a consecutive step. Schröder et al. [38] developed a probabilistic retrieval methodology based on interactive learning. Ferecatu and Boujemaa [15] developed an active relevance feedback solution based on support vector machines using weighted histograms as descriptors. In [39], a support vector machine approach was employed to recognize land cover information corresponding to spectral characteristics, and Gabor filters were used to extract textural features that characterize spatial information. Integration of the land cover information and textural features led to the retrieval of the spectral and spatial patterns of remotely sensed imagery. In [40], the latent Dirichlet allocation model was utilized to map low-level features of clusters and segment the high-level map labels for remote sensing image annotation and mapping. Finally, Kalaycilar et al. [41] computed the topological and distance-based spatial relationships between the regions. The spatial relationship histograms were constructed by classifying these regions in an image.

III. IMAGE FEATURE DESCRIPTORS

In this section, we will present the procedure for extracting holistic and local feature descriptors.

A. Holistic Features

A GIST feature [42] is a holistic feature descriptor representing the image by defining the spatial envelop with a low-dimensional approach. We calculate the 2048-dimensional GIST descriptors of the query image and every image in the image database. To reduce the computational cost without decreasing the retrieval accuracy, a reduction of the linear-dimensionality is performed on the descriptors to generate the projected data D. Then, (1) is employed to minimize the quantization loss and implement binary quantization in the resulting space

$$\Omega(B, R) = \|B - DR\|^2_F$$  (1)

where B is the GIST feature descriptor and R is the orthogonal matrix. $\|\cdot\|_F$ denotes the Frobenius norm.

Fig. 2 illustrates the GIST features of six categories in the UC Merced image database. It is noted that each category has its own discriminative holistic feature.

B. Local Features

In the following, we first introduce dense SIFT to achieve a representative local feature by means of spatial pyramid maximum pooling with sparse coding.

Dense SIFT is computed from the 16 $\times$ 16 size image patches with a step size of 8 pixels, and it is further encoded by using sparse coding. Similar to the principal component analysis, sparse coding uses a group of “supercomplete” base vectors to represent the sample data, which can find the structures and patterns within the input data [43].

For an image $O = [o_1, \ldots, o_n] \in \mathbb{R}^{m \times n}$, where $o_1, o_2, \ldots, o_n$ are the patches of the feature vectors, $O$ can be factorized into two matrices by matrix factorization: $U \in \mathbb{R}^{m \times k}$ and $T \in \mathbb{R}^{k \times n}$. The basis $U$ can capture the intrinsic geometric structure of $O$, and each column vector of $T$ is a sparse representation. Norms such as the Frobenius are used to
approximate $O$. We make use of (2) to implement the matrix factorization of the image

$$\min_{U,T} \|O - UT\|^2 + \lambda \|T\|_1$$

s.t. $\|U_k\| \leq 1, \quad \forall k = 1, 2, \ldots, K. \tag{2}$

Equations (3) and (4) are adopted to extract more discriminative features [43] through multiscale maximum pooling

$$z = F(U) \tag{3}$$

where $z$ is the local image feature by a prechosen pooling function

$$z_j = \max \{|u_{1j}|, |u_{2j}|, \ldots, |u_{Mj}|\} \tag{4}$$

where $z_j$ is the $j$th element of $z$, $u_{ij}$ is the matrix element at the $i$th row and the $j$th column of $U$, and $M$ is the number of local feature descriptors in the region.

Fig. 3 illustrates the process for extracting the local feature. Because the extracted local feature is highly dimensional, it is difficult to learn the feature joint relevance without considerable data overfitting and to obtain optimized parameters with limited training samples. The dimension reduction method [44] is applied to reduce the dimension.

IV. THREE-LAYERED GRAPH-BASED IMAGE RETRIEVAL

As illustrated in Fig. 4, the process of the proposed image retrieval can be divided into three layers. In the first layer, by using the graph-based holistic feature learning algorithm, one list whose images are similar to the input query is obtained from the image database. Accordingly, we obtain another image list by using the graph-based local feature learning algorithm. In the second layer, we first rerank the images in each of the two lists and then obtain three types of graph anchors $P_H$, $P_L$, and $P_C$. $P_H$ and $P_L$ are the top-ranked images of the two candidate lists, whereas $P_C$ contains the common retrieved images in both of the top candidate lists. We take $P_H$, $P_L$, and $P_C$ as the queries to retrieve database images by means of the graph-based holistic feature learning algorithm and local feature learning algorithm, respectively. Thus, six image retrieval lists can be obtained. In the third layer, positive and negative data are generated by evaluating the quality of these retrieval results. SimpleMKL [45] is introduced to fuse the retrieval results.

A. First-Layer Graph-Based Retrieval

In this layer, a semisupervised learning method [46] is introduced for retrieving images. We only consider the query image as the labeled image, and the remaining images are considered unlabeled. To achieve this goal, we first construct a weighted graph to obtain the optimized ranking function. Afterward, database images are retrieved based on the local and holistic features of these images.

1) Graph Construction: Given an image dataset $X = \{x_1, \ldots, x_i, x_{i+1}, \ldots, x_n\}$. Some of the images are labeled as query images, and the rest of them need to be ranked according to their relevance with the queries. For each query image, a weighted undirected graph $G = (V, E, W)$ is constructed from each individual feature-based retrieval method, where the retrieval quality or the relevance is modeled by the weights on the edges, where $V$ is a set of vertices, $E$ is a set of edges, and $W$ is a set of edge weights. Each database image corresponds to a vertex in $G$. For each image, we identify its $k$ nearest neighbors and connect the corresponding vertices in $G$ with edges that are associated with the distance between the two vertices. $G$ is usually not a connected graph. To obtain a connected graph, the two nearest connected components of $G$ are combined into one by an edge that links their two nearest points. The process ends when $G$ becomes a connected and undirected graph.
Fig. 4. Three-layered graph-based learning framework for image retrieval.

The graph \( G \) corresponds to the similarity among vertices with the weight \( W \) defined by

\[
W_{ij} = \exp \left( -\frac{d^2(V_i, V_j)}{\sigma^2} \right)
\]

where \( d(V_i, V_j) \) denotes the feature distance between the vertices \( V_i \) and \( V_j \), and \( W_{ij} \) is the edge weight of \( E_{ii} \). \( \sigma \) is a constant that controls the strength of the weight, and it is set as the median distance among all images.

If the similarity among vertices is determined by the holistic feature based on pairwise image distances, \( G \) is called the graph \( G_{H} \) of the holistic feature. Otherwise, it is called the graph \( G_{L} \) of the local feature.

2) Learning on the Graph: After the graphs are constructed, the optimal pairwise relevance is obtained through a learning algorithm. Given the initial labeled data (a query image), the regularization term \( \Omega(f) \) and the empirical loss term \( R(f) \) can be combined to form the following optimization function:

\[
\arg \min_f \{ \Omega(f) + \mu R(f) \}
\]

where \( f \) is a to-be-obtained ranking function and \( \mu \) is a weighting factor controlling the balance of the aforementioned two terms, which is mathematically expressed as

\[
\Omega(f) = \frac{1}{2} \sum_{V_i, V_j \in V} W_{ij} \left( \frac{f(V_i)}{d(V_i)} - \frac{f(V_j)}{d(V_j)} \right)^2
\]

\[
= \sum_{V_i, V_j \in V} W_{ij} \left( \frac{f^2(V_i)}{d(V_i)} - \frac{f(V_i) f(V_j)}{d(V_i)d(V_j)} \right)
\]

\[
= f^T(I - \Theta s)f
\]

where \( \Theta s = K^{-1/2} W K^{-1/2}, K \) is a degree matrix, and

\[
R(f) = \|f - y\|^2
\]

The ranking function \( f \) that minimizes (6) is computed by setting the derivative of (6) to zero. It is mathematically expressed as

\[
f = (I - \alpha \Theta s)^{-1} y
\]

where \( \alpha = 1/(1 + \mu) \). A large \( f_i \) indicates the high similarity between the database image \( i \) and the query.

We have obtained two retrieval image lists by using the graph-based holistic and local feature learning algorithms. The two lists represent different retrieval results and have different image arrangements. Next, we select the top-ranked images from the lists as the new input to further refine the retrieval results.

B. Second-Layer Graph-Anchor-Based Retrieval

In this layer, the task is to use the retrieved images as prior knowledge to improve the retrieval precision. To achieve this, we utilize the reranking method to obtain the top-ranked images as graph anchors. Then, these anchors are taken as new queries to further improve the retrieval accuracy.

1) Construction of Graph Anchors: Through the first layer, we have obtained two image retrieval results. They are stored in two lists: \( H_{LQ} \) and \( L_{LQ} \). We separately rerank the images in each of the lists and obtain two reranked image retrieval lists. The top-ranked images and common images are taken as the graph anchors for expanding the query. In the following, we discuss the procedure for generating the graph anchors.

To generate accurate graph anchors for learning the joint relevance of the holistic and local features, it is necessary to precisely measure the similarity among the images. We define the similarity degree of the two images as the relative score (RS). RS can be computed by using the following reranking...
method. This method considers the similarity among image features and the image clustering information.

In Fig. 5, every circle represents one image. Q is the query image. N is used to represent the retrieved image matrix. A₁, A₂, ..., Aₘ are the top-ranked images in the retrieval results in the first layer, which are stored in a list Lₚ. The images in Lₚ are further used as queries to perform the search. For example, N₁, N₂, ..., Nₘ are the query results of A₁. Similarly, each of the other images in Lₚ retrieves the m best matching database images. Finally, m + 1 lists are generated

\[
\begin{pmatrix}
L_Q \\
L_{A_1} \\
L_{A_2} \\
\vdots \\
L_{A_m}
\end{pmatrix} =
\begin{pmatrix}
A_1 & A_2 & \cdots & A_m \\
N_{11} & N_{12} & \cdots & N_{1m} \\
N_{21} & N_{22} & \cdots & N_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
N_{m1} & N_{m2} & \cdots & N_{mm}
\end{pmatrix}
\]

We compute the similarity score (SS) between two images. The SS between the retrieved images Nᵢ and Q is derived by

\[
SS(Q, Nᵢ) = SS(Q, Aᵢ) \cdot SS(Aᵢ, Nᵢ)
\]

where SS(Q, Aᵢ) indicates the similarity between images Q and Aᵢ, and SS(Aᵢ, Nᵢ) represents the similarity between images Aᵢ and Nᵢ.

When SS(Q, Nᵢ) is computed, we first consider the relationship between Q and Aᵢ. If they are very similar, many common images exist in the lists Lₚ and Lₐᵢ, and the spatial distributions of the image features are similar between the two lists. Then, con(eᵢ, Aᵢ) is defined to obtain the contributions of common elements e₁, e₂, ..., eₘₙ to Aᵢ by using the following:

\[
con(eᵢ, Aᵢ) = \ln (1 + \lambda_{con} \cdot SS(eᵢ, Aᵢ)) \cdot p^{index}
\]

where SS(eᵢ, Aᵢ) is computed by (9), jε[1, m], index is the sequence of the retrieved images in the list, 0 < p < 1, and \(\lambda_{con}\) is the contribution constraint coefficient.

After all of the contributions of the common elements between Lₚ and Lₐᵢ are computed, the similarity between Q and Aᵢ is as follows:

\[
SS(Q, Aᵢ) = \frac{\text{num}}{n} \cdot R \left( \sum_{j=1}^{\text{num}} \text{con}(eⱼ, Q), \sum_{j=1}^{\text{num}} \text{con}(eⱼ, Aᵢ) \right)
\]

where R(a, b) = \{ a/b, if a ≤ b, b/a, otherwise. \}

The similarity score SS(Aᵢ, Nᵢ) between Aᵢ and Nᵢ is also solved by (11). Because a retrieved image RI may have several positions in N, we search RI in the lists Lₚ, Lₐ₁, ..., Lₐₘ, and we compute the RS between Q and RI based on the following:

\[
RS(Q, RI) = \sum_{i=1}^{n+1} SS(Q, Aᵢ) \cdot \text{con}(Nᵢ, Aᵢ) \cdot \delta(Nᵢ)
\]

where \(\delta(Nᵢ) = \begin{cases} 1, & \text{if } Nᵢ \in Lₚ, \\ 0, & \text{otherwise.} \end{cases}\)

After the images in \(H_{Lₚ} \) and \(L_{Lₚ} \) are reranked, the top retrieved images are taken as the prior knowledge for further image retrieval. Graph anchors [53] are initially for scalable graph-based semisupervised learning. In our approach, we want to seek a better anchor set to enhance retrieval accuracy. In the following, we obtain three types of graph anchors: \(P_H \), \(P_L \), and \(P_C \). \(P_H \) and \(P_L \) are the top reranked images of \(H_{Lₚ} \) and \(L_{Lₚ} \), respectively, whereas \(P_C \) contains the common images in both of the top reranked lists (\(H_{Lₚ} \) and \(L_{Lₚ} \)). It is easy to obtain \(P_H \) and \(P_L \). The steps for obtaining \(P_C \) are described as follows.

Suppose that \(CI \) is a common image in the reranked \(H_{Lₚ} \) and \(L_{Lₚ} \). Its similarity \(sim_{HL} \) to the query image can be mathematically expressed as follows:

\[
sim_{HL} = \ln (1 + \lambda_{con} \cdot RS_H(Q, CI)) \cdot p^i + \ln (1 + \lambda_{con} \cdot RS_L(Q, CI)) \cdot p^j
\]

where i and j denote the positions of the common image \(CI \) in \(H_{Lₚ} \) and \(L_{Lₚ} \), respectively. RS is the relative score. \(\lambda_{con} \) and \(p \) have been defined in (11).

If \(sim_{HL} \) is no smaller than the given similarity threshold, the corresponding common image in the lists is taken as a graph anchor in \(P_C \). The aforementioned process is continued until all of the common images in \(H_{Lₚ} \) and \(L_{Lₚ} \) are traversed.

2) Retrieval Reinforcement and Retrieval Fusion: We take \(P_H \), \(P_L \), and \(P_C \) as query images to retrieve images from the image database using the learning algorithm of holistic and local features introduced in the first layer. We label these queries as the prior knowledge and modify the vector \(y \) in (9), where the values of the graph anchors are all equal to 1. Thus, we obtain six retrieval lists through the two graphs \(G_H \) and \(G_L \). That is, through the graph \(G_H \), we separately derive three
retrieved results \((L_{HH}, L_{HL}, \text{and } L_{HC})\) corresponding to the queries \(P_H, P_L, \text{and } P_C\). Accordingly, we also obtain other three retrieved results \((L_{HH}, L_{HL}, \text{and } L_{LC})\) by \(G_L\). \(L_{HH}\) and \(L_{LL}\) are the results of reinforcement learning. \(L_{HL}, L_{HC}, L_{LC}, \text{and } L_{LC}\) are the results of the preliminary fusion. We know that \(L_{HL}\) and \(L_{HC}\) are the retrieval results corresponding to the graph \(G_H\) of the holistic feature, but the queries are the graph anchors \(P_L\) and \(P_C\) which are not generated from \(G_H\). Similarly, \(L_{HL}\) and \(L_{LC}\) are the retrieval results corresponding to \(G_L\), but the queries are the graph anchors \(P_L\) and \(P_C\) which are not generated from \(G_L\). \(L_{HC}\) and \(L_{LC}\) are also the results of the graph-based learning algorithm of holistic and local features, which use the common graph anchors as the queries. This approach can reduce the influence of some inappropriate graph anchors on the retrieval result to a certain extent.

C. Third-Layer Fusion-Based Retrieval Results

To generate the fusion retrieval result of different features, we evaluate the performance of the retrieved images in the aforementioned six lists and find similar and dissimilar images. Then, we learn the weights of holistic and local features and related parameters of the fusion. SimpleMKL [45] is introduced to obtain the final fusion result.

For the fusion process, positive data and negative data are needed. Thus, we introduce an image list \((L_G)\) that contains similar images (positive data), as well as a second one \((L_D)\) that contains the dissimilar images (negative data). To create \(L_D\), we randomly selected a certain number of images from the bottom of \(L_{HH}, L_{HL}, L_{HC}, L_{LL}, L_{LC}\), and \(L_{LC}\) separately and stored them in \(L_D\). Clearly, obtaining \(L_G\) is a more difficult task that can be achieved by the following retrieval evaluation procedure.

1) Graph-Based Retrieval Evaluation: The images in \(L_{HH}, L_{HL}, L_{HC}, L_{LL}, L_{LC}\), and \(L_{LC}\) are reranked separately. The top-ranked images are usually very similar to the query image and belong to the same category. We apply the retrieval consistency to evaluate the aforementioned reranked results. For example, we select the \(c_n\) top-ranked images from reranked \(L_{HH}\) as the graph anchors for an expansion query. We modified \(y\) in (9), where the values of the graph anchors are set to 1. The graph \(G_H\) of the holistic feature and \(f\) computed in Section IV-A are employed to obtain the retrieved result \(L_a\). For the top-ranked images in reranked \(L_{HH}\) and \(L_a\), their similarities to the query image and the sequence in the aforementioned two lists are considered. Assume that the position of a retrieved image in reranked \(L_{HH}\) is \(i\) and that in \(L_a\) is \(j\). The consistency degree of the retrieval results is

\[
DC = \frac{\text{num}}{c_n} \prod_{k=1}^{c_n} \frac{1}{1 + i + j} \tag{15}
\]

where \(\text{num}\) is the number of common images in the reranked lists \(L_{HH}\) and \(L_a\). In practice, \(c_n\) is usually set to 20, 30, 40, and 50.

From the retrieved result \(L_{HH}\), we can obtain a certain number of the top-ranked images by using the retrieval result evaluation. These images are stored in the aforementioned image list \(L_G\).

The evaluation processes of other retrieval lists are the same as the process of \(L_{HH}\). To the retrieved results \(L_{HL}, L_{HC}, L_{LL}, L_{LC}, \text{and } L_{LC}\), we compute their consistency degree by (15) and choose the good retrieval results for the fusion of the final result. These images are also stored in \(L_G\).

2) Further Feature Fusion by SimpleMKL: After the evaluation, we need to address the problem of obtaining the best fusion weights. SimpleMKL is applied to learn suitable query-dependent fusion weights for two types of features. In the image lists \(L_G\), the images are taken as the positive data, and the images in \(L_D\) with low similarity are taken as negative data. We train the parameters of SimpleMKL through these data and the two features. SimpleMKL determines the combination of the feature weights by solving a standard SVM optimization problem based on a gradient descent method. The kernel \(K(x, x')\) represents the similarity between two images using one feature. Supposing that there are \(M\) types of features in an image \(x_i\), we formulate the corresponding kernel matrix for the \(m\)th feature

\[
K_m(x_i, x_j) = \exp \left( -\gamma_m \| x_i^m - x_j^m \|^2 \right) \tag{16}
\]

where \(m = 1, 2, \ldots, M\), \(x_i^m\) corresponds to the \(m\)th feature of \(x_i\) and \(\gamma_m\) is a positive parameter that can be computed according to the data distribution.

To combine multiple features, the conventional approach is to combine the basic kernels in a convex way

\[
K(x_i, x_j) = \sum_{m=1}^{M} \omega_m K_m(x_i, x_j) \tag{17}
\]

where \(\omega_m \geq 0\) and \(\sum_{m=1}^{M} \omega_m = 1\).

SimpleMKL is optimized by solving the following minimization problem:

\[
\min_{\{f_m\}, b, \zeta, \omega} \frac{1}{2} \sum_{m=1}^{M} \frac{1}{\omega_m} \| f_m \|^2 H_m + C \sum_{i=1}^{M} \zeta_i
\]

s.t.

\[
y_i \left( \sum_{m=1}^{M} f_m(x_i) + y_i b \right) \geq 1 - \zeta_i \quad \forall i
\]

\[
\zeta_i \geq 0 \quad \forall i
\]

\[
\sum_{m=1}^{M} \omega_m = 1, \omega_m \geq 0 \quad \forall m \tag{18}
\]

where \(f_m\) is a function corresponding to a reproducing kernel Hilbert space \(H_m\) of kernel \(K_m(x_i^m, x_j^m)\). \(\zeta_i\) is the slack variable, \(C\) is the regularization parameter that trades off between the decision function and the slack variable, \(\omega\) is the weight vector, and \(b\) is the coefficients to be learned from the training data.

After the optimized parameters are obtained, we compute the decision function \(f(x) + b = \sum f_m(x) + b\), where each function \(f_m\) corresponds to a different \(H_m\) associated with a kernel \(K_m\). The decision function describes the relevance of the retrieved images to the query image. The higher the relevance
value is, the more similar the two images are. After all of the images in the retrieved image list are sorted in descending order, they represent the final retrieval result.

D. Computational Complexity Analysis

The common way to express the complexity of one algorithm is using big $O$ notation. Assume that we have $n$ database images. The computational cost of the proposed method mainly lies in the three layers. In the first layer, the image retrieval is computed with the cost of $O(n^3 + n^2)$. In the second layer, since the number of reranking images is far less than $n$, the complexity of the reranking process can be negligible. The cost for the second layer is $O(n^3 + n^2)$. In the third layer, the main computational cost is for optimizing the parameters in SimpleMKL with the time complexity of $O(N^3 + lN^2 + dlN)$, where $d$ is the feature dimension, $l$ is the number of training samples, and $N$ is the number of support vectors. Since $d \ll n$ and $l \ll n$, in the third layer, different features can be fused quickly. The whole cost of the proposed method is $O(n^3 + n^2 + N^3 + lN^2 + dlN)$.

V. EXPERIMENTS

In this section, we evaluate the performance of our proposed method for image retrieval on the UC Merced image data set [48]. The data set consists of images of 21 LULC categories with a pixel resolution of 30 cm. Each class contains 100 RGB color samples of size $256 \times 256$ pixels [19]. The example images are shown in Fig. 6. The source code of the proposed method is available at the website: http://geogother.bnu.edu.cn/teacherweb/zhangliqiang/

Fig. 6. Some examples of the UC Merced Land Use–Land Cover dataset. (a) Agricultural. (b) Airplane. (c) Baseball diamond. (d) Beach. (e) Buildings. (f) Chaparral. (g) Dense residential. (h) Forest. (i) Freeway. (j) Golf course. (k) Harbor. (l) Intersection. (m) Medium residential. (n) Mobile home park. (o) Overpass. (p) Parking lot. (q) River. (r) Runway. (s) Sparse residential. (t) Storage tanks. (u) Tennis court.

To evaluate the effectiveness of our method, we conduct extensive experiments on the UC Merced dataset. We first compare our approach in different layers with those by using each individual feature and then further compare our method with other related methods. mAP and precision–recall curves are employed to evaluate the retrieval performance. We also use the average normalized modified retrieval rank (ANMRR) [52] to provide an overall rating of the performance.

A. Validation of Retrieval Performance in Each Layer

We extract the 2048-dimensional GIST feature and apply the algorithm in [49] to reduce the dimension of the feature to 512. The dimension reduction method [44] is used to reduce the local feature to 32. Fig. 7 demonstrates that features whose dimensions have been reduced have little influences on image retrieval.

Fig. 7. Retrieval results by using different bits in the UC Merced data set with the local feature.
accuracy, but the retrieval efficiency is enhanced, saving more than 80% of the computational time in the graph construction.

1) Retrieval Performance in the First Layer: In the first layer, we use the holistic and local features to construct the graph. $\alpha$ is set to 0.1 in the graph of the holistic feature and 0.01 in the graph of the local feature.

To validate the retrieval performance in the first layer, we use two methods to retrieve images. One is the holistic or local feature-based retrieval method with Cityblock distance (the first method). The other is the proposed method in the first layer. The results shown in Fig. 8 are the average values of retrieval performance of 21 categories. The retrieval precision of some categories, such as agricultural, buildings, and harbor, is even higher than the values in Fig. 8. The learned ranking function takes holistic and local features into account and uses the smoothness and fitting constraint to solve the image similarity. Thus, the retrieval results in the first layer are more preferable. From Fig. 8, we notice that our presented method always performs better than the first method.

2) Retrieval Performance in the Second Layer: In the second layer, we rerank the retrieval results of the first layer. The number of top reranked images is set to 10, 20, and 30. mAP is used to evaluate different retrieval results. The first retrieval result is obtained by the holistic or local feature-based retrieval method with Cityblock distance. The second result is from the retrieval of the first layer. The third result is the reranked result by using our method. The three retrieval results are shown in Fig. 9. It is shown that the retrieval performance of the reranked method in the second layer is always better than in the other two methods. Fig. 9(a)–(c) shows the different results obtained by the holistic feature. The method in the first layer and the reranking process can retrieve more accurate images than the holistic feature-based retrieval method with Cityblock distance. When the nearest neighbor number is equal to 3, the retrieval performances of the method in the first layer and the reranking method are nearly the same. As the number of retrieved images increases, the reranking method outperforms the other methods. For the local feature, Fig. 9(d)–(f) shows the different results obtained by the local feature. From the mAP curves, we note that the performance of the reranking method is still higher than those of other methods. The retrieval precision obtained in the first layer is further improved after the reranking.

Assume that $RS_{H_{\text{max}}}$ denotes the maximum $RS$ in the graph-based retrieval method of the holistic feature and $RS_{L_{\text{max}}}$ denotes the maximum $RS$ in the graph-based retrieval method of the local feature. The thresholds are defined as $t_{H_{\text{sim}}} = 0.1 \times RS_{H_{\text{max}}}$ and $t_{L_{\text{sim}}} = 0.1 \times RS_{L_{\text{max}}}$ to extract the graph anchors $P_H$ and $P_L$. Equation (14) and the threshold $0.01 \times \max(\text{sim}_{\text{HL}})$ are utilized to extract the graph anchor $P_C$. In the second layer, the graph anchors $P_H$, $P_L$, and $P_C$ are used to retrieve images from the database.

From Figs. 10 and 11, it is noted that use of the top-ranked images in the expansion query can improve the retrieval accuracy. For example, by using the graph anchors $P_L$, the obtained retrieval result LHL is more accurate than the other retrieval results shown in Fig. 10. The reason is that the reranking process can offer more precise prior knowledge for $P_L$. In Fig. 11, the retrieval results of some categories, such as dense residential, golf course, and tennis courts, may not be very good in LLH. The main reason is that the graph anchors $P_H$ cannot provide more useful information for the local feature learning in the second layer.

3) Retrieval Performance in the Third Layer: We have reranked the retrieval result lists $L_{HH}$, $L_{HL}$, $L_{HC}$, $L_{LH}$, $L_{LL}$, and $L_{LC}$. We usually select the top 50 images in the aforementioned reranked lists to evaluate the retrieval consistency. These good results are considered as the positive data, and the images at the bottom of the reranked lists are selected as the negative data. We remove the duplicate elements of these positive and negative data. The parameters of SimpleMKL are
Fig. 9. Comparisons of different retrieval methods. (a) Retrieval results obtained by using the GIST feature with Cityblock distance. The number of top-ranked images is 10. (b) Retrieval results obtained by using the holistic feature-based retrieval method in the first layer. The number of top-ranked images is 20. (c) Reranking result. The number of top-ranked images is 30. (d) Retrieval results obtained by using the SIFT feature with Cityblock distance. The number of top-ranked images is 10. (e) Retrieval results obtained by using the local feature-based retrieval method in the first layer. The number of top-ranked images is 20. (f) Reranking result. The number of top-ranked images is 30.
trained through the positive data and negative data. The results are shown in Figs. 12 and 13.

As illustrated in Fig. 12, when the nearest neighbor number is equal to 100, the mAPs of the harbors are 13.85%, 15.40%, 14.70%, 21.81%, and 35.48%, corresponding to the retrieval results in the first layer, LHC, LHH, LHL, and third layer, respectively. The mAPs of parking lot are 13.97%, 15.26%, 16.12%, 25.45%, and 38.52%. The retrieval results in the third layer perform best with the holistic feature.

In Fig. 13, by using the local feature-based retrieval method, the mAPs of agricultural are 23.19%, 23.17%, 23.4%, 24.23%, and 42.58%, corresponding to the retrieval results in the first layer, LLC, LLH, LLL, and third layer, respectively. The mAPs of baseball diamond are 34.95%, 39.91%, 42.83%, 42.39%, and 45.45%, and the mAPs of beach are 30.99%, 33.21%, 35.09%, 34.36%, and 41.56%. For most categories, the retrieval results have high precisions. The retrieval processes in the first two layers only consider one single feature type, i.e., the holistic or local feature, whereas we extract positive and negative data in the third layer and train the parameters of the SimpleMKL with holistic and local features at the same time. This approach combines the advantages of the holistic and local feature-based retrieval methods, and it fuses the two retrieval results efficiently. Therefore, the performance in this layer outperforms those in the first two layers, which means that fusion of the ranked retrieval image sets given by the holistic and local feature-based retrieval methods can remarkably enhance the final retrieval accuracy.

From Figs. 12 and 13, we conclude that the retrieval results in the third layer are higher than those of the other two layers.

B. Comparisons of the Related Methods

In this section, we compare our retrieval result with other methods: pyramid of histograms of orientation gradients (PHOG) [50], local invariants [23], pyramid histogram of words (PHOW) [51], and global morphological texture descriptors [19]. In [50], a shape feature descriptor together with a spatial pyramid kernel is introduced to represent the local shape and spatial layout of the objects in the image. They generalize the spatial pyramid kernel and learn its level weighting parameters...
Fig. 13. Comparisons of the retrieval results obtained by using the local feature. LLC denotes the retrieval result when \( P_C \) represents the query images. LLH denotes the retrieval result when \( P_H \) represents the query images. LLL denotes the retrieval result when \( P_L \) represents the query images.

for combining multilevel features. In this experiment, to build pyramid histograms, we discretize gradient orientations into eight bins and set the number of pyramid levels to 3. In [23], the dense SIFT descriptor is extracted to describe the local invariance of images. The space between dense SIFT samples is set to 8, and the size of each patch for the SIFT descriptor is set to 16. The SIFT descriptor is quantized using codebooks resulting from the \( k \)-means clustering. The dictionary size is set to 400, and 50 texton images are used to build the histogram bins. The sum of the square error function is used for \( k \)-means clustering. The point with a local minimum is taken as the center of the cluster. In [51], PHOW partitions the image into fine subregions and computes histograms of local features found inside each subregion. PHOW introduces the spatial pyramid to compute the feature histograms. For PHOW, the space between dense SIFT samples is set to 8, the size of each patch for the SIFT descriptor is set to 16, and the number of pyramid levels is 3. The approach for obtaining each pyramid feature is the same as in [23]. In [19], the CCH [55], RITs [54], and two texture descriptors based on the Fourier power spectrum of an image’s quasi-flat zone representation (i.e., FPS\(_1\) and FPS\(_2\)) are combined into morphological descriptors for retrieving remote sensing images. FPS\(_1\) and FPS\(_2\) have a number of parameters to set, i.e., the number of scales (\( m \)), the step (\( s \)) for computing the QFZ scale space, and, of course, the number of disks and sectors (\( n \)), respectively. Specifically, FPS\(_1\)’s parameters are set as \( m = 2 \) using the scales QFZ\(_{15,15}\) and QFZ\(_{30,30}\) with \( n = 7 \) disks, and FPS\(_2\) is calculated also with two scales QFZ\(_{10,10}\) and QFZ\(_{20,20}\) and \( n = 16 \) sectors.

In our method, the input is still a single query image in our approach, but we can explore more relevant images iteratively to overcome the deficiency of the single image-based retrieval methods. It allows us to accurately validate each return, suppressing the false positives. The top-ranked retrieved images can be used as prior knowledge for the next retrieval. The retrieval results are fused by using SimpleMKL, which greatly improves the retrieval precision. The other four methods use a single query image for retrieving similar images from the image database.

In order to further validate the efficiency and effectiveness of our presented method, we have also computed the average ANMRR, feature extraction time, retrieval time, and overall computational time for obtaining the retrieved images only using the SIFT feature and the first two layers (second layer-SIFT) of our method.

As shown in Fig. 14, the second layer-SIFT achieves a higher retrieval accuracy compared with the methods in [19], [23], [50], and [51]. The first 5% of the retrieval result of three layers is not much better than other methods. The reason is that some inappropriate positive data are brought into the training data. However, our method next achieves the best retrieval precision among all of the methods.

In Fig. 15, the chaparral land-use is easy to be recognized, but some categories are obscured by other categories, such as buildings and overpasses. The performance of our method still outperforms other methods in most categories.

All experiments are performed on a PC with a 3.0-GHz CPU and 8-GB memory. As listed in Table I, it is noted that the overall computational time of our method is approximately equal to those of the methods in [23] and [51] and more than the method in [50]. However, our retrieval accuracy is higher than other methods. In our method, feature extraction consumes most of the overall time, and from Section IV-D, we know that
learning suitable query-dependent fusion weights for holistic and local features by using SimpleMKL costs most of the retrieval time. We also notice that the overall computational time of the second layer-SIFT is less than the methods in [19] and [51], but the average ANMRR is the smallest.

In our image retrieval procedure, most of the steps are parallelizable, such as the holistic and local feature extraction of the query images and the image retrieval by using each individual feature. Therefore, they can be implemented by employing a parallelization scheme to further reduce computation time.

VI. CONCLUSION

In this paper, we have proposed a three-layered remote sensing image retrieval method based on local and holistic features. Unlike previous image retrieval methods, which often concatenate all types of features into one vector to retrieve images, we have extended the query and applied a novel three-layered learning image retrieval approach to fuse various features. To extend the query, the proposed approach refines the original input query by combining it with the top-ranked retrieval results obtained by using the holistic and local feature-based methods. The expansion query images are taken as graph anchors for further retrieving six image sets. The images in each set are evaluated for generating positive data and negative training data through the image retrieval.

Future work will focus on decreasing computational complexity and fusing geometric [56], [57] and spectral features [58] to improve the retrieval performance of remote sensing images.

ACKNOWLEDGMENT

We would like to thank the editors and anonymous reviewers for their valuable comments, and also thank Erchan Aptoula for providing the comparison features.

REFERENCES


Yuebin Wang is currently working toward the Ph.D. degree in the School of Geography, Beijing Normal University, Beijing, China. His research interests are remote sensing imagery processing and 3-D urban modeling.
Liqiang Zhang received the Ph.D. degree in geoinformatics from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2004. He is currently a Professor with the School of Geography, Beijing Normal University, Beijing. His research interests include remote sensing image processing, 3-D urban modeling, and spatial object recognition.

Xiaohua Tong received the Ph.D. degree from Tongji University, Shanghai, China, in 1999. He was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China, between 2001 and 2003. He was a Research Fellow with The Hong Kong Polytechnic University, Hung Hom, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, between 2008 and 2009. He is a Chang-Jiang Scholar Chair Professor appointed by the Ministry of Education, China. He also received the National Natural Science Funds for Distinguished Young Scholar in 2013. He is the author of more than 40 publications in international journals. His current research interests include remote sensing, GIS, trust in spatial data, and image processing for high-resolution and hyperspectral images.

Prof. Tong serves as the Vice-Chair of the Commission on Spatial Data Quality of the International Cartographical Association and the Co-Chair of the ISPRS Working Group (WG II/4) on Spatial Statistics and Uncertainty Modeling. He received the State Natural Science Award (Second Place) from the State Council of the Peoples’ Republic of China in 2007.

Liang Zhang is currently working toward the Ph.D. degree in the School of Geography, Beijing Normal University, Beijing, China. His research interests are remote sensing imagery processing and 3-D urban modeling.

Zhenxin Zhang is currently working toward the Ph.D. degree in geoinformatics in the School of Geography, Beijing Normal University, Beijing, China. His research interests include light detection and ranging data processing, quality analysis of geographic information systems, and algorithm development.

Hao Liu received the Bachelor’s degree from Beijing University of Aeronautics and Astronautics, Beijing, China, in 2012 and the Master’s degree from Kyushu University, Fukuoka, Japan, in 2015. He is currently working toward the Ph.D. degree in the School of Geography, Beijing Normal University, Beijing. His present research interests include machine learning and computer vision.

Xiaoyue Xing is currently working toward the Bachelor’s degree in the School of Geography, Beijing Normal University, Beijing, China. Her research interests are TLS point cloud processing.

P. Takis Mathiopoulos (SM’94) received the Ph.D. degree in digital communications from the University of Ottawa, Ottawa, ON, Canada, in 1989. From 1982 to 1986, he was with Raytheon Canada Ltd., working in the areas of air navigational and satellite communications. In 1988, he joined the Department of Electrical and Computer Engineering (ECE), University of British Columbia (UBC), Vancouver, BC, Canada, where he was a faculty member until 2003, holding the rank of Professor from 2000 to 2003. From 2000 to 2014, he was with the Institute for Space Applications and Remote Sensing (ISARS), National Observatory of Athens (NOA), first as the Director and then as the Director of Research, where he established the Wireless Communications Research Group. As ISARS’ Director (2000–2004), he led the institute to a significant expansion in R&D growth and international scientific recognition. For these achievements, ISARS was selected as a National Centre of Excellence for 2005 to 2008. Since 2014, he has been an Adjunct Researcher with the Institute for Astronomical and Astrophysics, Space Applications and Remote Sensing (IAASARS, NOA). Since 2003, he has also been teaching part-time at the Department of Informatics and Telecommunications, University of Athens, Athens, Greece, where he has been a Professor of digital communications since 2014. From 2008 to 2013, he was appointed as a Guest Professor with Southwest Jiaotong University, Chengdu, China. He has been also appointed by the Government of China as a Senior Foreign Expert at the School of Information Engineering, Yangzhou University (2014–2015), Yangzhou, China, and by Keio University, Minato, Japan, as a Visiting Professor with the Department of Information and Computer Science (2015–2016) under the Top Global University Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Government of Japan. For the last 25 years, he has been conducting research mainly on the physical layer of digital communication systems for terrestrial and satellite applications, including digital communications over fading and interference environments. He is the coauthor of a paper in GLOBECOM’89 establishing for the first time in the open technical literature the link between MLSE and multiple (or multisymbol) differential detection for the AWGN and fading channels. He is also interested in channel characterization and measurements, modulation and coding techniques, synchronization, SIMO/MIMO, OFDM, software/cognitive radios, and green communications. In addition, since 2010, he has been actively involved with research activities in the fields of remote sensing, LiDAR systems, and photogrammetry. In these areas, he is the coauthor of more than 100 journal papers, mainly published in various IEEE and IET journals, 4 book chapters, and more than 120 conference papers. He has been the PI of more than 40 research grants and has supervised the thesis of 11 Ph.D. and 23 Master’s students. He has regularly acted as a consultant for various governmental and private organizations. Since 1993, he has been serving on a regular basis as a Scientific Advisor and a Technical Expert for the European Commission (EC). In addition, from 2001 to 2014, he served as a Greek representative to high-level committees in the EC and the European Space Agency. Dr. Mathiopoulos has been or currently serves on the editorial board of several archival journals, including the IET Communications and the IEEE TRANSACTIONS ON COMMUNICATIONS (1993–2005). He has been a member of the Technical Program Committee (TPC) of more than 70 international IEEE conferences, as well as TPC Vice Chair of the 2006–8 IEEE VTC and 2008–F IEEE VTC as well as the Cochair of FITC2011. He has delivered numerous invited presentations, including plenary and keynote lectures, and has taught many short courses all over the world. As a faculty member at the ECE of UBC, he was elected as an ASI Fellow and a Killam Research Fellow. He was a coeditor of two best paper awards for papers published in the 2nd International Symposium on Communication, Control, and Signal Processing (2008) and 3rd International Conference on Advances in Satellite and Space Communications (2011).