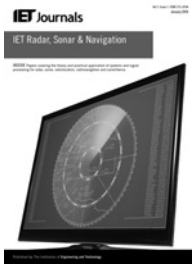
www.ietdl.org

Published in IET Radar, Sonar and Navigation
Received on 3rd December 2009
Revised on 11th May 2010
doi: 10.1049/iet-rsn.2009.0294



ISSN 1751-8784

Active sample-selecting and manifold learning-based relevance feedback method for synthetic aperture radar image retrieval

R. Chen Y.F. Cao H. Sun*

Signal Processing Lab., School of Electronic Information, Wuhan University, Wuhan, People's Republic of China
**E-mail: yongfengcao.cyf@gmail.com*

Abstract: Content-based image retrieval (CBIR) provides an effective way to address the increasing need for intelligent data access to synthetic aperture radar (SAR) image repositories. In CBIR, a critical component is relevance feedback (RF), which is used to bridge the 'semantic gap'. This study proposes a new RF method with active sample-selecting and manifold learning for CBIR of SAR images. In this method, the authors adopt a modified maximum margin projection with a new neighbourhood estimation criterion to discover both the geometrical structure and discriminant structure of the underlying data manifold. In order to achieve a satisfactory performance with limited feedback samples, the authors also propose an active sample selection strategy with which the diversity of feedback samples can be increased while the redundancy is decreased. The authors test our method on a TerraSAR-X image database and compare it with four other state-of-the-art RF methods. The superiority and validity of the authors' method is proved by the retrieval results and the computing cost is acceptable for image retrieval applications.

1 Introduction

Synthetic aperture radar (SAR) has been widely used in remote sensing because of its excellent properties, such as working in all-time, all-weather conditions. With the development of SAR technology, more and more SAR sensors have been launched, such as ALOS, ENVISAT-1, ERS-2 and the newly launched TerraSAR-X. There is a mass of SAR images obtained from these satellites every day. How to manage such large SAR image databases and mine useful information from them have become big challenges.

An effective way to cope with these challenges is content-based image retrieval (CBIR) [1]. In a CBIR system, low-level features (e.g. visual or electromagnetic features) are first extracted from a query image and all images in the database. Then the similarities between the query image and the database images are calculated by a certain metric. The database images are sorted by their similarities and displayed as the retrieval result, so that users can find images that are the most similar to the query image. SAR image retrieval has been applied successfully for many purposes, for example, biomass investigating [2], wet snow detecting in mountain areas [3] and land use classification [4] etc. In our work, it is used to detect land-cover types in urban areas.

In the CBIR system, the gap between low-level features and high-level semantic concepts usually leads to poor performance; this is the so-called 'semantic gap'. Relevance feedback (RF) provides a powerful tool to bridge this gap [5, 6]. In RF, users choose some samples from the current retrieval result, and label them manually as 'relevant' or 'irrelevant', then feed these labelled samples back to the CBIR system to modify the similarity metric (or in other fashions), so that the retrieval results after RF can be much closer to users' desires. The flow diagram of an interactive image retrieval system is shown in Fig. 1.

At early years, many RF methods were designed for finding a better similarity measure to evaluate the relevance between the query image and images in the database. Two simple and popular approaches for RF are query vector modification [7] and feature relevance estimation [5]. These techniques are efficient for target search and mono-modal concept retrieval; they hardly track complex image concepts [8].

Performing an estimation of the query concept can be seen as a statistical learning problem. Recently, machine learning has been widely used in RF and has had great success. Zhang *et al.* [9] use feedback samples to train an support vector machine (SVM) classifier; the relevance of the database images are calculated according to the distance to the boundary. Tieu and Viola [10] use feedback samples to train an Adaboost classifier for RF.

RF is different from the traditional learning problem because the user is not likely to label a large number of images. Thus the number of feedback samples is limited. But the dimension of the feature space is usually very high. So there exists a crucial problem called the 'curse of dimensionality' [11]. The performance of statistical learning techniques may be influenced by this problem. Moreover, the computation will be time consuming when the feature space's dimension is hundreds or even higher.

In much of the work on RF, the images for which the user is asked to provide feedback at the next round were simply those that were currently considered by the learner as the most

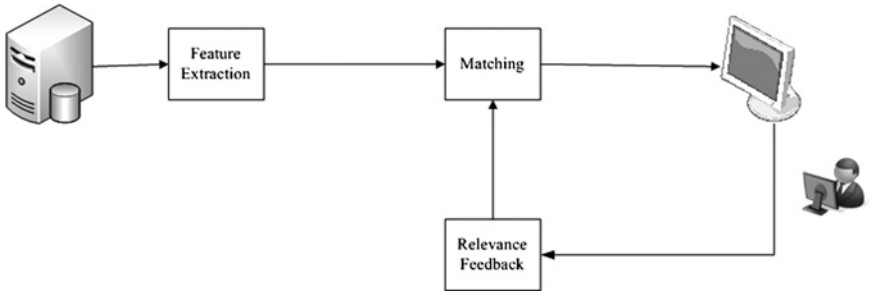


Fig. 1 *Framework of an interactive image retrieval system*

relevant; also, in some cases, these images are randomly selected [12]. In these styles, the feedback samples may not be so informative for RF.

In order to address the problems mentioned above, we propose a manifold learning method and a pool-based active learning strategy for RF in our image retrieval system.

The originality of our method is based on the association of two components:

1. *Active feedback sample selection:* an active sample selection strategy to increase the diversity of the feedback samples while decreasing the redundancy between them.
2. *Finding an optimal feature subspace:* both local geometrical and discriminant structure of the underlying data manifold are discovered by a modified maximum margin projection (MMP) algorithm, in which we propose a new neighbourhood estimation criterion to construct within-class and between-class adjacency graphs. The images are projected from their original high-dimensional feature space to a low-dimensional subspace, in which the margin between relevant images and irrelevant ones is maximised. The similarity of the database images and the query image is re-calculated in the new feature space.

The structure of this paper is as follows: Section 2 presents an overview of related work, including some approaches for finding an optimal feature subspace and some active learning algorithms. Section 3 gives a detailed description of our work. The experimental results on SAR image retrieval and analysis will be given in Section 4. Finally, Section 5 presents the conclusions of the paper.

2 Related work

This section gives a brief survey of recent work on finding an optimal feature subspace and active learning. As stated in the introduction, our method is built up by these two components, which are used in our SAR image retrieval system to deliver a better performance.

2.1 Finding an optimal feature subspace

It is necessary to find a low-dimensional subspace to alleviate the influence of the 'curse of dimensionality' in image retrieval, so as to improve the subsequent learning performance. The most popular dimensionality reduction algorithms include principal component analysis [13, 14] and linear discriminant analysis (LDA) [11, 15, 16] and also many variations of LDA, such as BiasMap [17] for non-linear problem and regularised linear discriminant analysis (RLDA) [18, 19] for 'small sample size' problem. However, all the methods above only discover the global geometrical structure, whereas the local manifold structure is ignored [20].

Most recently, manifold learning has been paid more and more attention by researchers for image retrieval, such as locally linear embedding (LLE) [21], ISOMAP [22] and Laplacian eigenmaps. Among all the manifold learning methods, Laplacian eigenmaps has been proved successful in transductive inference [23], dimensionality reduction [24], clustering [13, 25] and also in image retrieval [26, 27]. The Laplacian eigenmaps manifold learning method is based on spectral graph theory [28]. It aims to discover the underlying manifold of the data. To preserve the local geometrical structure of the data manifold, the samples that are close in the original feature space will still be kept close in the embedding subspace. However, Laplacian eigenmaps is an unsupervised learning method, the labels that have been given by users in the interactive procedure cannot be used easily in Laplacian eigenmaps-based RF. It considers only the local geometrical structure of data, ignoring the discriminant structure given by labelled samples, which is also important for image retrieval.

To get use of the label information in subspace construction, MMP [20] considers within-class and between-class relationship according to the label information. Since the neighbourhood relationship between data points are estimated in Euclidian space, the performance of MMP may be influenced by complex data distribution (details will be shown in Section 3.1).

2.2 Active selection for feedback samples

Active sample selection has become a critical component in a CBIR system for achieving a better performance with limited feedback samples. Most active learning algorithms are conducted in an iterative fashion. In some algorithms, the example with the highest classification uncertainty is chosen as the feedback sample in each round, such as support vector machine active learning [29], query by committee [30] etc. Some other algorithms have been proposed to reduce the redundancy in batch mode sample selection. In [12, 31], the author employs the inner product to measure the redundancy between two data points. The inn-product criterion is simple and convenient for sample selection, but the selected samples may not be so representative.

3 RF method with active sample-selecting and manifold learning

3.1 Brief introduction to our RF method

In our RF method, we employ a modified MMP to find an optimal feature subspace. In MMP [20], both geometrical

and discriminant structure of the data can be discovered with the label information obtained in each RF round. One problem existing in MMP is that the k nearest neighbours of a data point x are calculated by Euclidian distance. So it only considers the position of the data points in feature space, which may not reflect the real structure of the data. The data points that are measured as neighbours by this metric may locate far away on the data manifold (see Fig. 2 for an example).

Considering the problem mentioned above, we employ a new criterion to estimate the neighbourhood of each data point in constructing the k nearest neighbour adjacency graph. In this new criterion, we use the transition probability between two data points to measure their 'distance'. The transition probability considers not only the distance between two points, but also the data distribution. So it can express the neighbourhood relationship more exactly.

On the other hand, we use an active batch mode selection strategy to increase the diversity of feedback samples, and to decrease the information redundancy between them. By this strategy, the most informative and representative images are chosen as feedback samples with high probability.

The three steps in our RF method are listed as follows:

1. *Active feedback sample selection:* In each feedback round, some images are chosen automatically from the current retrieval result as feedback samples by our active sample selection strategy. Then the user labels them manually and sends them back to the system to supervise the process of search.
2. *Neighbourhood estimation based on random walk graph Laplacian (RWGL):* An important step in the spectral graph theory-based manifold method is constructing the neighbourhood adjacency graph. We propose a new neighbourhood estimation criterion (called KNN^{TP}) based on transition probability between two data points using RWGL [26], so as to reflect the underlying data manifold more exactly. The k nearest neighbours of each data point are determined by KNN^{TP}.
3. *Finding an optimal subspace to maximise the margin between relevant images and irrelevant ones:* We adopt a modified MMP to discover both the geometrical and discriminant structure of the data manifold. The adjacency graphs in MMP are constructed based on our new neighbourhood estimation criterion. In the projected subspace, the local geometrical structure can be preserved while maximising the margin between relevant images and irrelevant ones at each local neighbourhood.

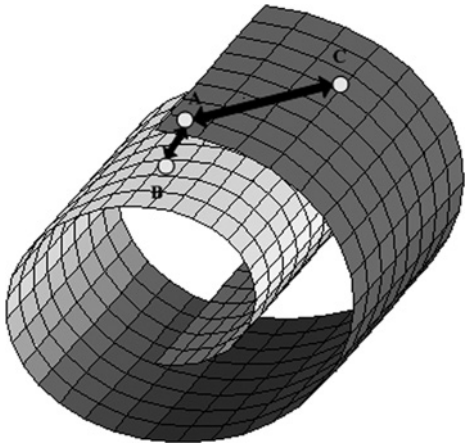


Fig. 2 *Illustrating the 'false neighbour' using Euclidian distance measure*

Point B is the neighbour of A by Euclidian distance measure, whereas point C is much closer to A than point B on the data manifold

A complete flowchart of interactive image retrieval with our RF method is given in Fig. 3.

A detailed description of these three steps will be given in the following subsections.

3.2 Active selection of feedback samples

Since the number of feedback samples is limited in image retrieval, it is necessary to choose samples for feedback that are the most informative and representative. We employ an active selection strategy to select feedback samples automatically in an interactive process. The selection strategy follows the two rules below:

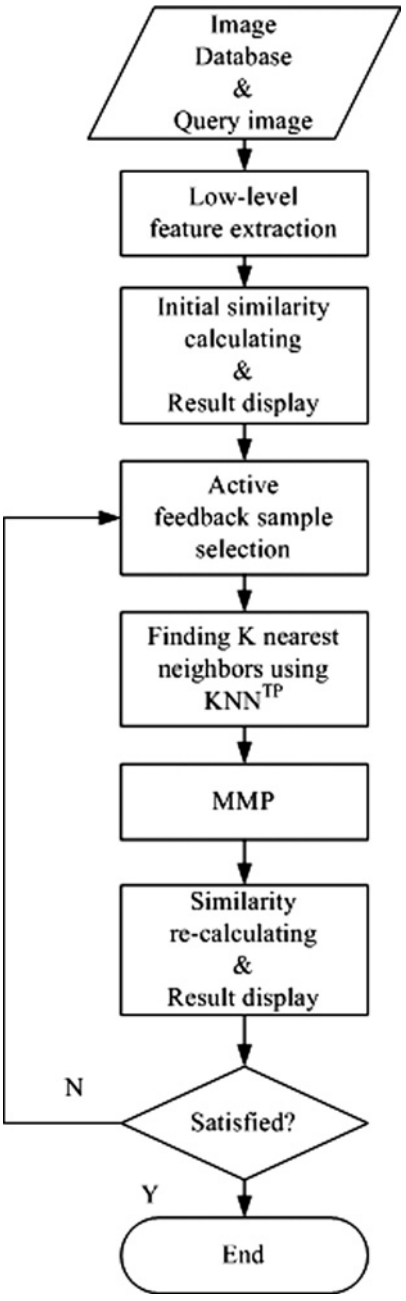


Fig. 3 *Flowchart of our interactive image retrieval framework*

doi: 10.1049/iet-rsn.2009.0294

1. the feedback samples should be highly diverse, to be more representative of the whole data set, and
2. the information redundancy between feedback samples should be low.

We choose the first N_{pool} images of the current retrieval result as the image pool at first. Then the images in the pool are clustered into K clusters by K-means, where K is the number of samples to feedback in each round. In each cluster, we select the image that minimises the upper bound of redundancy (UBR) as a new feedback sample. The UBR is defined as follows

$$\text{UBR}(\boldsymbol{x}) = \max_{\boldsymbol{x}_i \in S_{\text{labelled}}} \langle \boldsymbol{x}, \boldsymbol{x}_i \rangle \tag{1}$$

where S_{labelled} is the feedback sample set, which has been labelled in RF. The clustering step in our algorithm is used to ensure the diversity of the feedback samples. The inner-product $\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle$ is used to measure the redundancy between two images $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. The images with a low inner product (called the most 'orthogonal') are regarded as having low redundancy [12, 32]. The procedure of our active sample selection strategy is shown in Fig. 4.

3.3 New criterion to estimate the k nearest neighbours

In this section, we propose a new criterion for finding the k nearest neighbours. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ denote the data points and $\tilde{K}$ be a kernel function, which is symmetrical non-negative to measure the distance between two input points. In our work, we choose the Gaussian kernel $\tilde{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \text{e}^{-(|\boldsymbol{x}_i - \boldsymbol{x}_j|^2/2\sigma^2)}$ as the kernel function, where σ is a scale parameter. The degree function is defined as $\tilde{d}(\boldsymbol{x}) = \sum_{i=1}^{n} \tilde{K}(\boldsymbol{x}, \boldsymbol{x}_i)$. The normalised kernel used in general graph Laplacian is defined as

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{\tilde{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)}{[\tilde{d}(\boldsymbol{x}_i)\tilde{d}(\boldsymbol{x}_j)]^{\lambda}}$$

where λ is a constant and $\lambda \geq 0$.

Accordingly, the degree function associated with the normalised kernel K is $d(\boldsymbol{x}) = \sum_{i=1}^{n} K(\boldsymbol{x}, \boldsymbol{x}_i)$. This normalised kernel K can be used to construct a weighted undirected graph $G\langle V, E \rangle$ in which the vertices are $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ and the weight on the edge that links vertices $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$. The degree of a vertex $\boldsymbol{x}$ is defined as the sum of the weights of the edges at the vertex, which equals $d(\boldsymbol{x})$.

Let $\boldsymbol{W}$ be a $n \times n$ matrix with $\boldsymbol{W}(i, j) = K(\boldsymbol{x}_i, \boldsymbol{x}_j)$, and $\boldsymbol{D}$ be a $n \times n$ diagonal matrix with $D(i, i) = d(\boldsymbol{x}_i)$. The matrix $L_{\text{rw}} = D^{-1}W$ defines the RWGL, where $L_{\text{rw}}(i, j)$ characterises the transition probability of a walk from vertex $\boldsymbol{x}_i$ to $\boldsymbol{x}_j$ [26]

$$L_{\text{rw}}(i, j) = p(\boldsymbol{x}_j | \boldsymbol{x}_i) \tag{2}$$

This transition probability considers not only the distance measure between points (W), but also the data distribution (D). So it will be more suitable for neighbourhood estimation on a data manifold than the Euclidian distance.

We define the new k nearest neighbour estimation criterion KNN^{TP} according to the transition probability

$$\begin{aligned} \text{KNN}^{\text{TP}}(\boldsymbol{x}) &= \arg\max_i p(\boldsymbol{x}|\boldsymbol{x}_i), \quad \text{for } k = 1 \\ \text{KNN}^{\text{TP}}(\boldsymbol{x}) &= \{\boldsymbol{x}_{\text{NN}}^1, \ldots, \boldsymbol{x}_{\text{NN}}^k\}, \quad \text{for } k > 1 \end{aligned} \tag{3}$$

where $\{\boldsymbol{x}_{\text{NN}}^1, \ldots, \boldsymbol{x}_{\text{NN}}^k\}$ are data points with the k largest values of $p(\boldsymbol{x}|\boldsymbol{x}_i)$. (We set $k = 10$ in the following experiments.)

3.4 Feature space projection with maximum margin

We employ a modified MMP with our new neighbourhood estimation criterion for feature space projection. It is developed based on Laplacian eigenmaps and provides a good way to use the label information of the feedback samples. With the labels that have been given by the users in RF, the k nearest neighbours of each data point, which has been found by KNN^{TP}, can be split into two sets: between-class neighbours N_{b} and within-class neighbours N_{w}

$$\begin{aligned} N_{\text{b}}(\boldsymbol{x}) &= \{\boldsymbol{x}_i | \boldsymbol{x}_i \in \text{KNN}^{\text{TP}}(\boldsymbol{x}) \text{ and } l(\boldsymbol{x}_i) \neq l(\boldsymbol{x})\} \\ N_{\text{w}}(\boldsymbol{x}) &= \text{KNN}^{\text{TP}}(\boldsymbol{x}) \backslash N_{\text{b}}(\boldsymbol{x}) \end{aligned} \tag{4}$$

Algorithm Active feedback sample selection
Input: The most relevant N_{pool} images in current retrieval result;
Batch size for each feedback round K;
Feedback sample set which has been labeled already $S_{labeled}$
$S_{pool} \leftarrow$ the most relevant N_{pool} images
$\{S_{cluster_1}, \ldots, S_{cluster_K}\} \leftarrow$ kmeans(S_{pool})
for $t = 1$ to K
  for $i = 1$ to $\#(S_{cluster_t})$
    for $j = 1$ to $\#(S_{labeled})$
      $\mathbf{x}_{pool} = S^i_{cluster_t}$
      $\mathbf{x}_{labeled} = S^j_{labeled}$
      Calculate $\langle \mathbf{x}_{pool} \cdot \mathbf{x}_{labeled} \rangle$
    end for
  end for
  $\mathbf{x}_{new\ select} = \arg \min_{x_{pool} \in S_{kluster_t}} \max_{x_{selected} \in S_{selected}} \langle \mathbf{x}_{pool} \cdot \mathbf{x}_{selected} \rangle$
  update feedback sample set:
  $S_{selected} \leftarrow S_{selected} \cup \mathbf{x}_{new\ select}$
end for

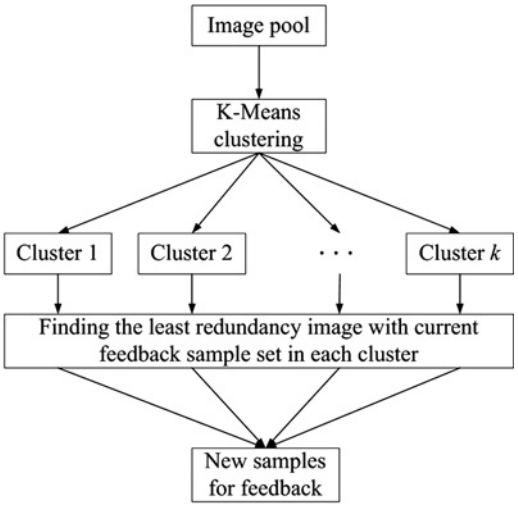


Fig. 4 *Pseudo-code and flow diagram of our active selection algorithm*

With N_b and N_w, we have the between-class neighbour adjacency graph G_b and within-class neighbour adjacency graph G_w. The weight matrices associated with G_b and G_w are given as follows

$$W_{b,ij} = \begin{cases} 1 & \text{if } x_i \in N_b(x_j) \quad \text{or} \quad x_j \in N_b(x_i) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$W_{w,ij} = \begin{cases} \gamma & \text{if } l(x_i) = l(x_j) \\ 1 & \text{if } x_i \quad \text{or} \quad x_j \text{ is unlabelled} \\ & \text{but } x_i \in N_w(x_j) \quad \text{or} \quad x_j \in N_w(x_i) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

The parameter γ should be relatively high to reflect the strong connection between two points with the same label. Since it is not sensitive as stated in [20], we set $\gamma = 50$ empirically in our experiments.

The aim of MMP is to find a projection with a maximum margin between relevant samples and the irrelevant samples at each local neighbourhood. The points in the within-class neighbour set of data point x should still keep close to x, whereas the ones in the between-class neighbour set should be far away from x. Let $X = (x_1, x_2, \ldots, x_m)$, d is the dimension of the original feature space, a is a projection vector, $y = (y_1, \ldots, y_m)^T$ be such a map and $y_i = a^T x_i$. The objective function of MMP is given as follows

$$\begin{cases} \min \sum_{ij} (y_i - y_j)^2 W_{w,ij} \\ \max \sum_{ij} (y_i - y_j)^2 W_{b,ij} \end{cases} \tag{7}$$

The first objective function in (7) is used to preserve the local geometrical structure of the data, whereas the second is used to maximise the margin between relevant samples and irrelevant ones.

The optimisation problem above can be integrated into a unified form (see [20] for details)

$$\arg\max_{\substack{a \\ a^T X D_w X^T a = 1}} a^T X(\alpha L_b + (1-\alpha) W_w) X^T a \tag{8}$$

α $(0 \le \alpha \le 1)$ is a constant (we set $\alpha = 0.5$ in our experiments), which is used to adjust the importance of the two parts in (7). D_w and D_b are diagonal matrices, $D_{w,ij} = \sum_j W_{w,ij}$, $D_{b,ij} = \sum_j W_{b,ij}$ and $L_b = D_b - W_b$ is the Laplacian matrix of G_b.

Then (8) can be transformed into a generalised eigenvalue problem

$$X(\alpha L_b + (1-\alpha) W_w) X^T a = \lambda X D_w X^T a \tag{9}$$

The solutions $a_1, \ldots, a_n (n \le d)$ are eigenvectors of (9), which are associated with the largest n eigenvalues.

Thus, the projection is given as follows

$$y_i = A^T x_i \tag{10}$$

where $A = (a_1, \ldots, a_n)$. Using A as the transform matrix, the query image q and database image x can be mapped into a low-dimensional space from the original feature space

$$q' = A^T q$$

$$x'_i = A^T x_i$$

The distance between q and x_i is re-calculated in the new space

$$\text{Distance}(q, x_i) = \text{Euclidian}(q', x'_i) = \sqrt{(q' - x'_i)^T (q' - x')} = \sqrt{(q - x_i)^T A A^T (q - x_i)} \tag{11}$$

4 Experimental results

In this section, we report the results of a large number of experiments on a TerraSAR-X image database using the CBIR system with our RF method, and compare its performance with four other RF methods. We will give a short introduction to our experiments first, and then give the experimental results on our image database. In the following experiments, we set parameters $k = 10$, $\gamma = 50$ and $\alpha = 0.5$ (see Sections 3.3 and 3.4 for details).

4.1 Image database and experimental setup

The database includes 511 images belonging to five pre-defined concepts (see Fig. 5 for some examples): forest (100 images), grass (102 images), industrial area (113 images), bare land (96 images) and residential area (100 images), which are sampled from an image (20 000 × 20 000 pixels) of Guangzhou in China from TerraSAR-X with a resolution of 1.25 m. The average size of each image is 128 × 128 pixels. The ground truth of the database is obtained based on high-resolution optical photographs of the same site, a priori knowledge of the region, and our knowledge of the SAR response to various surface features.

We use a 256-d grey-level histogram as a low-level feature, which is a simple and effective statistical feature for SAR images. Since the focus in our experiments is RF, we do not spend much more time on extracting complex features. We employ the Euclidian distance in this original feature space as an initial similarity measure. In our image retrieval system, we compare our method with four other RF algorithms: SVM (using Libsvm toolkit [33]), RLDA, RWGL and MMP. The dimension of the embedding subspace is set to five in our method.

We designed two sample selection modes in our experiments:

- *Normal selection mode:* choosing the first ten relevant and the first ten irrelevant images of the current retrieval result as the feedback samples.
- *Active selection mode:* choosing 20 samples (including relevant and irrelevant samples) through our active sample selection algorithm.

We iterate RF three times. Selecting those unlabelled samples, which are currently considered by the learner as the most relevant in the early stage of RF, can sometimes produce a faster convergence to a satisfactory performance [31]. In our method, we use the normal selection mode

doi: 10.1049/iet-rsn.2009.0294

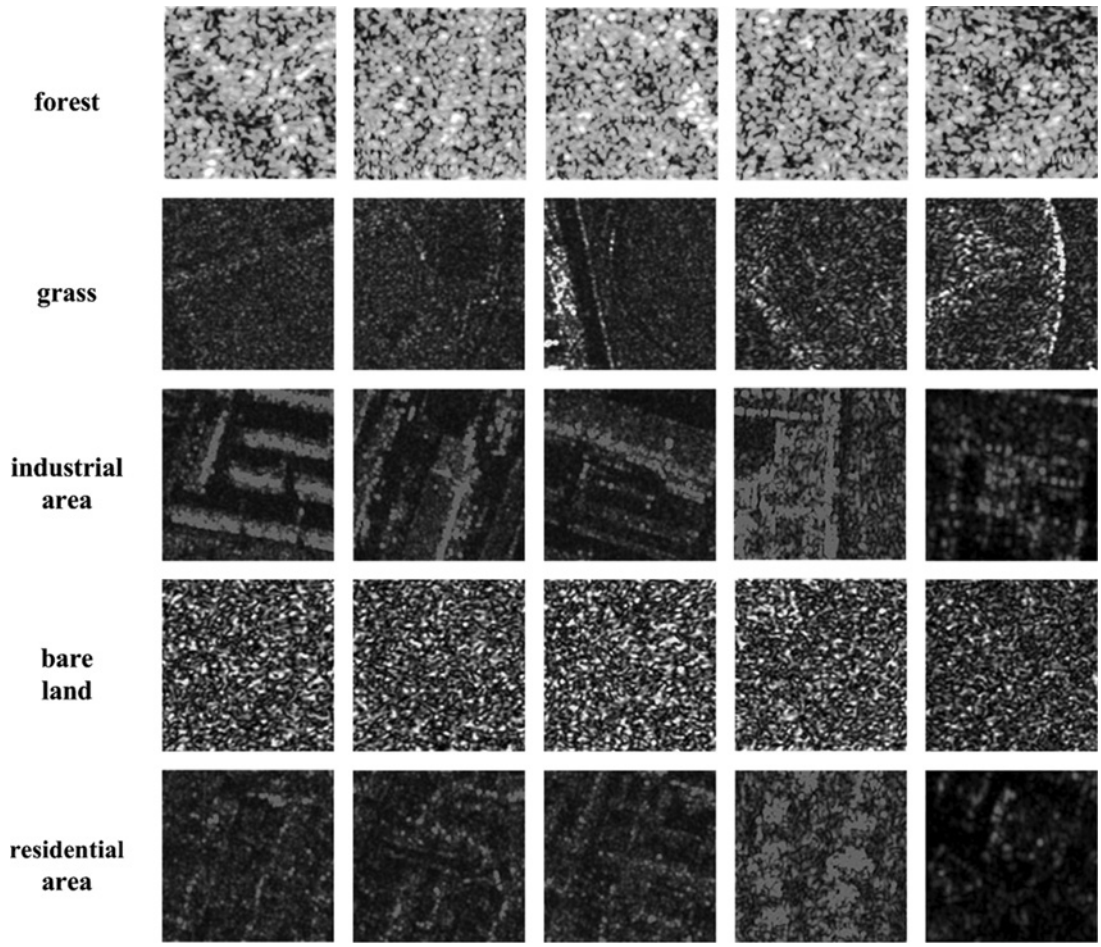


Fig. 5 *Examples from the five concepts in our SAR image database*

instead of the active selection mode in the first round. In order to show the merit of the active sample selection strategy in our method, we also perform the same experiments using the normal selection mode from round1 to round3 for comparison.

4.2 Evaluation metrics

We use three metrics in our experiments to evaluate the performance of the retrieval result. Precision–Recall curve (P–R curve) [34, 35] is used to evaluate the performance of a specific query session and the evolution of the CBIR system while RF iterates. Average precision (AP) and mean average precision (MAP) [36, 37] are used to evaluate the performance of the system statistically during one (or multi)-query session(s). And precision VS scope [38, 1] is used to evaluate the performance of the first few pages of the retrieval result, which are the most significant to some users. The definitions of these metrics are given below.

4.2.1 P–R curve: Suppose that the number of images in the current retrieval result is N_S, the number of targets in the result (the ones which belong to the 'query class') is n and the number of all targets in the database is N_T. Precision and Recall are defined as

$$\begin{aligned} \text{Precision} &= \frac{n}{N_S} \\ \text{Recall} &= \frac{n}{N_T} \end{aligned} \tag{12}$$

The x-axis of the P–R curve represents the Recall, whereas the y-axis represents the Precision. The higher the curve is, the better is the performance.

4.2.2 AP and MAP: AP is defined as the average precision when each target image has been found,

$$\text{AP} = \sum_{D_i \in S_T} P(D_i)/|S_T| \tag{13}$$

where S_T is the set of all target images in the database, D_i is the ith target image, $P(D_i)$ is the retrieval precision when D_i has been found and $|S_T|$ is the size of S_T.

AP can be seen as the average height of the P–R curve. In our experiments, we use the average precision when Recall = 0.1, 0.2, L, 1 to approximate AP.

MAP is defined during multi-query sessions. Assuming that $\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\}$ are n query sessions, then

$$\text{MAP} = \frac{1}{n}\sum_{i=1}^{n} \text{AP}(\boldsymbol{q}_i) \tag{14}$$

4.2.3 Precision against scope: Precision against scope is the retrieval precision in the top N_{scope} images of the current result.

We compute $P30$, $P50$ and $P100$ in our experiments to evaluate the precision in the top 30, 50 and 100 images.

4.3 Retrieval results

We choose one image from the database randomly as the query image in each query session, and implement a total of 100 query sessions in our experiments. The APs in three feedback rounds on these 100 queries are shown in Fig. 6.

From Fig. 6 we can see that our method outperforms the other four RF algorithms on AP. Among all the 100 queries, our method performs best on 74 queries in the first feedback round, and performs best on 90 and 88 queries in the second and third feedback rounds, respectively. The superiority of our method could be due to the joint work by both KNN^{TP} and active sample selection.

To exclude the influence of active sample selection, we also perform the same experiments with normal sample selection. The performance of our method with normal sample selection is lower than that with active sample selection, but is still better than the other four RF algorithms. With normal sample selection, our method performs best on 74, 76 and 66 queries among the 100

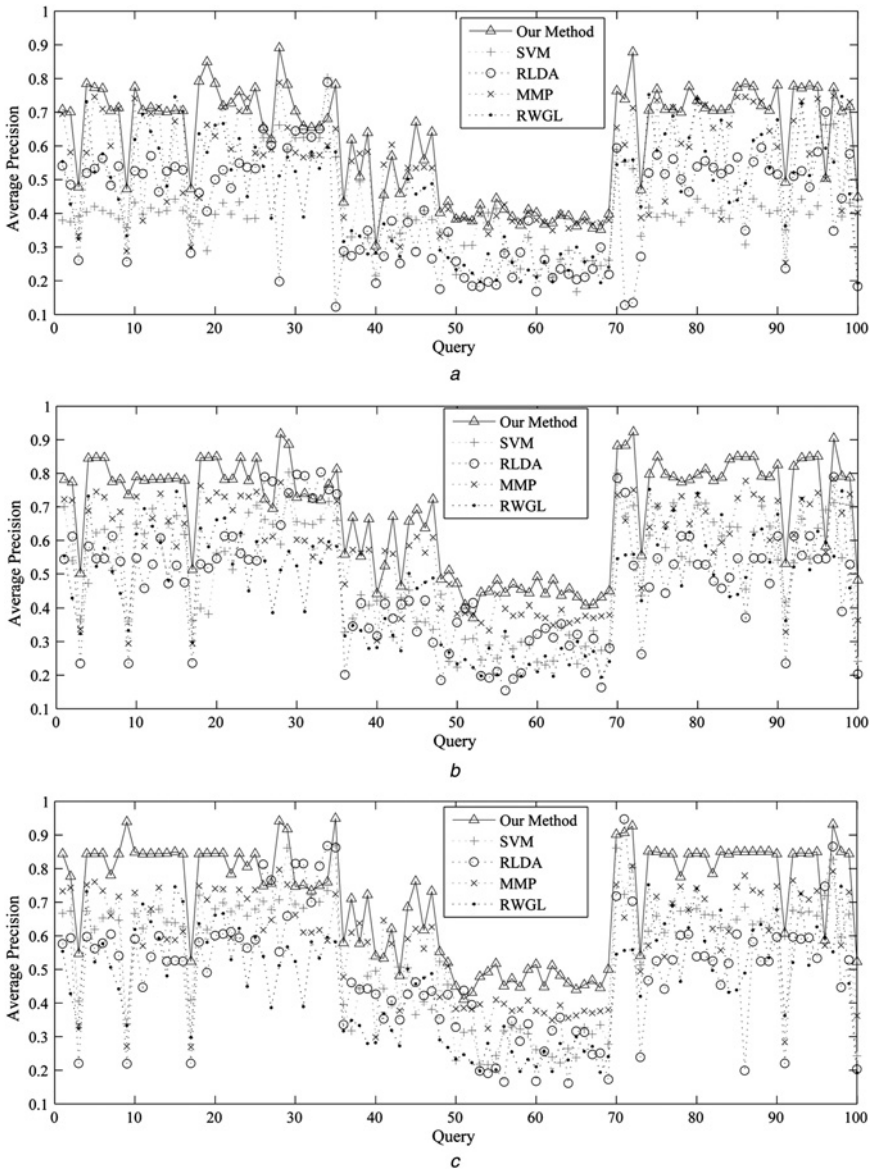


Fig. 6 *Average precision on 100 queries in the first three feedback rounds, X-axis refers to the serial number of the experiment, Y-axis refers to the corresponding average precision*

a Feedback round1
b Feedback round2
c Feedback round3

Table 1 MAP of five RF methods in three feedback rounds

	Our method (active/normal)	SVM	RLDA	MMP	RWGL
round1	–/0.6178	0.4030	0.4108	0.5413	0.4642
round2	0.6872/0.6398	0.5068	0.4668	0.5787	0.4642
round3	0.7170/0.6460	0.5506	0.4867	0.5840	0.4642

doi: 10.1049/iet-rsn.2009.0294

queries in the first three feedback rounds. The values of MAP for all five RF methods on these 100 queries are given in Table 1.

Tables 2–4 show the average retrieval precision of top-ranked images on all the 100 queries in the first three feedback rounds. The size of scope is set to 30, 50 and 100. We find that our method achieves a higher precision on $P30$, $P50$, $P100$ than the other four methods. In the active sample selection mode, $P30$ can reaches 0.9710 within three feedback rounds, which prove that our method can converge quickly to a satisfactory performance. It is very important to the CBIR system, since the top-ranked images

Table 2 Average retrieval precision on 100 queries in the first feedback round

	Our method (active/normal)	SVM	RLDA	MMP	RWGL
top 30 images ($P30$)	–/0.8667	0.6903	0.6267	0.8187	0.6653
top 50 images ($P50$)	–/0.8100	0.5908	0.6188	0.7198	0.5752
top 100 images ($P100$)	–/0.5977	0.3691	0.3777	0.4941	0.4051

Table 3 Average retrieval precision on 100 queries in the second feedback round

	Our method (active/normal)	SVM	RLDA	MMP	RWGL
top 30 images ($P30$)	0.9350/0.8870	0.8353	0.7680	0.8437	0.6653
top 50 images ($P50$)	0.8652/0.8270	0.6910	0.6778	0.7684	0.5752
top 100 images ($P100$)	0.6510/0.6039	0.4331	0.4304	0.5317	0.4051

Table 4 Average retrieval precision on 100 queries in the third feedback round

	Our method (active/normal)	SVM	RLDA	MMP	RWGL
top 30 images ($P30$)	0.9710/0.8907	0.8663	0.7870	0.8577	0.6653
top 50 images ($P50$)	0.8956/0.8286	0.7678	0.7316	78.18	0.5752
top 100 images ($P100$)	0.6802/0.6045	0.4725	0.4679	0.5219	0.4051

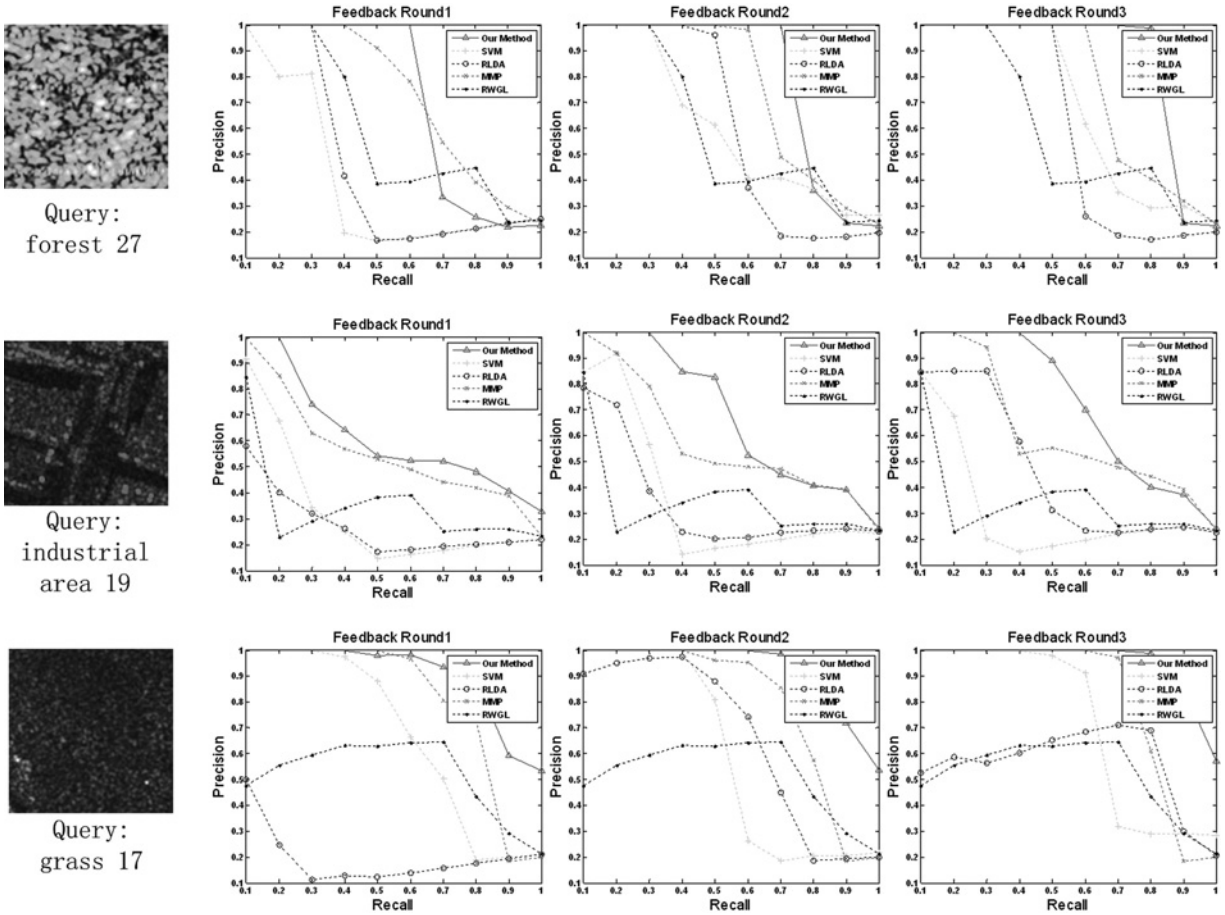


Fig. 7 *P–R curves for three different queries*

The P–R curves from feedback round1 to feedback round3 of each query are shown from left to right

are usually of most interest to users. By comparing the performance of the active selection mode and the normal selection mode used in our method, an improvement of 4–8% on average precision can be achieved by using active selection.

The performance achieved by our method (with normal selection mode in all three feedback rounds) is superior to that achieved by MMP in Tables 1–4, which implies that the new neighbourhood estimation criterion KNN^{TP} proposed in this paper reflects the real data manifold more exactly, since it is the main difference between them.

We drew P–R curves on some specific queries to see the evolution of the CBIR system in the first three feedback rounds. Fig. 7 gives the P–R curves for three query sessions. We can observe the evolution of the system by the development of the P–R curves from round1 to round3. The P–R curve of our method becomes higher as RF iterates. This means that our retrieval system is becoming much closer to the user's intention when more label information is given to the system to supervise it. And the P–R curves of our method are higher than the other four methods in all three feedback rounds, which means that our method can reach a higher precision when the same number of target images are found by the five RF methods.

5 Conclusions

In this paper, we present an active RF method for SAR image retrieval. In this method, we employ a modified MMP with a new neighbourhood estimation criterion based on transition probability to discover both the geometrical structure and discriminant structure of the data manifold with the help of the labels obtained in the interactive procedure. Our neighbourhood estimation criterion can represent the data manifold better since it not only considers the relative position of the data points, but also considers the distribution of the data. It has been proved valid by our retrieval experiments in which our method outperforms the other four algorithms in AP (MAP), *P*30, *P*50, *P*100 and also in P–R curves. Another important component in our method is active sample selection. With active sample selection, the diversity of feedback samples increases while the redundancy between them decreases. The performance of our system can deliver a remarkable improvement using our active sample selection strategy compared with normal sample selection. The average precision on the top 30 can reach 0.9710 within three feedback rounds, which shows that the retrieval result on our SAR image database is reliable using our RF method. The average running time of our RF method (carried out in MATLAB 7.6.0., and performed on a Core2 Duo-2.4 GHz Windows XP machine with a 2-Gbyte memory) is less than 3 s/round, satisfies the time requirement well for a real image retrieval system.

6 Acknowledgments

This work was Supported by the National High Technology Research and Development Program of China (No.2007AA12Z155), National Natural Science Foundation of China (No.40901207) and LIESMARS Special Research Funding.

7 References

1 Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: 'Content-based image retrieval at the end of the early years', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (12), pp. 1349–1380

2 Kurvonen, L., Pulliainen, J., Hallikainen, M.: 'Retrieval of biomass in boreal forests from multitemporal ERS-1 and JERS-1 SAR images', *IEEE Trans. Geosci. Remote Sens.*, 1999, **37**, (1), pp. 198–205

3 Nagler, T., Rott, H.: 'Retrieval of wet snow by means of multitemporal SAR data', *IEEE Trans. Geosci. Remote Sens.*, 2000, **38**, (2), pp. 754–765

4 Jain, A., Zongker, D.: 'Feature selection: evaluation, application, and small sample performance', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1997, **19**, (2), pp. 153–158

5 Rui, Y., Huang, T.S., Ortega, M., Mehrotra, S.: 'Relevance feedback: a powerful tool in interactive content-based image retrieval', *IEEE Trans. Circuits Syst. Video Technol.*, 1998, **8**, (5), pp. 644–655

6 Rui, Y., Huang, T., Mehrotra, S.: 'Relevance feedback techniques in interactive content-based image retrieval', *Proc. SPIE, Storage Retr. Images Video Databases VI*, 1998, **3312**, pp. 25–36

7 Rui, Y., Huang, T.S., Mehrotra, S., Ortega, M.: 'A relevance feedback architecture for content-based multimedia information systems'. Proc. IEEE workshop on Content Based Access of Image and Video Libraries, Proto Rico, 1997, pp. 82–89

8 Gosselin, P.H., Cord, M.: 'Active learning methods for interactive image retrieval', *IEEE Trans. Image Process.*, 2008, **17**, (7), pp. 1200–1211

9 Zhang, L., Lin, F., Zhang, B.: 'Support vector machine learning for image retrieval'. Proc. IEEE Int. Conf. on Image Processing, 2001, pp. 721–724

10 Tieu, K., Viola, P.: 'Boosting image retrieval', *Int. J. Comput. Vis.*, 2004, **56**, (1–2), pp. 17–36

11 Duda, R.O., Hart, P.E., Stork, D.G.: 'Pattern classification' (Wiley-Interscience, 2000, 2nd edn.)

12 Ferecatu, M., Boujemaa, N.: 'Interactive remote-sensing image retrieval using active relevance feedback', *IEEE Trans. Geosci. Remote Sens.*, 2007, **45**, (4), pp. 818–826

13 Spielman, D., Teng, S.: 'Spectral portioning works: planar graphs and finite element meshes'. 37th Ann. Symp. on Found. of Comp. Science (FOCS), 1996, pp. 96–105

14 Su, Z., Li, S., Zhang, H.-J.: 'Extraction of feature subspace for content-based retrieval using relevance feedback'. Proc. Ninth Ann. ACM Int. Conf. on Multimedia, (Multimedia'01), 2001, pp. 98–106

15 Swets, D.L., Weng, J.: 'Using discriminant eigenfeatures for image retrieval', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1996, **18**, pp. 831–836

16 Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: 'Eigenfaces vs. fisherfaces: recognition using class specific linear projection'. Proc. of the Fourth European Conf. on Computer Vision, (ECCV'96), Cambridge, UK, 15–18 April 1996, pp. 45–58

17 Zhou, X.S., Huang, T.S.: 'Small sample learning during multimedia retrieval using BiasMap'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, (CVPR'01), 2001, vol. 1, pp. 11–17

18 Lu, J., Plataniotis, K.N., Venetsanopoulos, A.N.: 'Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition', *Pattern Recognit. Lett.*, 2005, **26**, (2), pp. 181–191

19 Lu, J., Plataniotis, K.N., Venetsanopoulos, A.N.: 'Face recognition using LDA based algorithms', *IEEE Trans. Neural Netw.*, 2003, **14**, (1), pp. 195–200

20 He, X., Cai, D., Han, J.: 'Learning a maximum margin subspace for image retrieval', *IEEE Trans. Knowl. Data Eng.*, 2008, **20**, (2), pp. 189–201

21 Roweis, S., Saul, L.: 'Nonlinear dimensionality reduction by locally linear embedding', *Science*, 2000, **290**, (5500), pp. 2323–2326

22 Nguyen, G.P., Worring, M.: 'Optimizing similarity based visualization in content based image retrieval'. IEEE Int. Conf. on Multimedia and Expo., 2004, vol. 2, pp. 759–762

23 Belkin, M., Niyogi, P.: 'Semi-supervised learning on manifolds', *Mach. Learn.*, 2004, **56**, pp. 209–239

24 Belkin, M., Niyogi, P.: 'Laplacian eigenmaps for dimensionality reduction and data representation', *Neural Comp.*, 2003, **15**, (6), pp. 1373–1396

25 Belkin, M., Niyogi, P.: 'Laplacian eigenmaps and spectral techniques for embedding and clustering', *Adv. Neural Inf. Process. Syst.*, 2001, **14**, (14), pp. 585–591

26 Sahbi, H., Etyngier, P., Audibert, J.-Y., Keriven, R.: 'Manifold learning using robust graph laplacian for interactive image retrieval'. Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2008, pp. 1–8

27 He, X., Ma, W.-Y., Zhang, H.-J.: 'Learning an image manifold for retrieval'. Proc. 12th Ann. ACM Int. Conf. on Multimedia, 2004, pp. 17–23

28 Chung, F.R.: 'Spectral graph theory'. Proc. CBMS Regional Conference, Series in Mathematics, 1997, vol. 92

29 Tong, S., Koller, D.: 'Support vector machine active learning with applications to text classification'. Proc. 17th Int. Conf. on Machine Learning, (ICML'00), 2000, pp. 999–1006

doi: 10.1049/iet-rsn.2009.0294

30 Freund, Y., Seung, H.S., Shamir, E., Tishby, N.: 'Selective sampling using the query by committee algorithm', *Mach. Learn.*, 1997, **28**, (2/3), pp. 133–168
31 Ferecatu, M., Crucianu, M., Boujemaa, N. 'Reducing the redundancy in the selection of samples for SVM-based relevance feedback'. Technical report, INRIA, Report, France, July 2004
32 Steven, C.H., Hoi, R.J., Lyu, M.R.: 'Batch mode active learning with applications to text categorization and image retrieval', *IEEE Trans. Knowl. Data Eng.*, 2009, **21**, (9), pp. 1233–1248
33 Chang, C.-C., Lin, C.-J.: 'LIBSVM: a library for support vector machines', http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2000
34 Costache, M., Maitre, H., Datcu, M.: 'Categorization based relevance feedback search engine for earth observation images repositories'. IEEE Int. Conf. on Geoscience and Remote Sensing Symp., 2006, pp. 13–16
35 Li, C.-Y., Hsu, C.-T.: 'Image retrieval with relevance feedback based on graph-theoretic region correspondence estimation', *IEEE Trans. Multimed.*, 2008, **10**, (3), pp. 447–456
36 Hu, Y., Li, M., Yu, N.: 'Multiple-instance ranking: learning to rank images for image retrieval'. IEEE Conf. on Computer Vision and Pattern Recognition, 2008, pp. 1–8
37 Okabe, M., Yamada, S.: 'Semisupervised query expansion with minimal feedback', *IEEE Trans. Knowl. Data Eng.*, 2007, **19**, (11), pp. 1585–1589
38 Lu, Y., Zhang, H., Wenyin, L., Hu, C.: 'Joint semantics and feature based image retrieval using relevance feedback', *IEEE Trans. Multimed.*, 2003, **5**, (3), pp. 339–347