A Short Term and Long Term Learning Based on Fuzzy Transaction Repository and Feature Re-Weighting

Malihe Javidi, H.R.Pourreza, Hadi Sadoghi Yazdi, Homa Foroughi

Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
ma_ja21@stu-mail.um.ac.ir, hpourreza@um.ac.ir, sadoghi@sttu.ac.ir, ho_fo46@stu-mail.um.ac.ir

Abstract

In this paper, we introduce a combined relevance feedback approach for image retrieval using semantic similarity based on fuzzy transaction repository and feature re-weighting technique. This system accumulates user interactions using soft feedback model to construct Fuzzy Transaction Repository (FTR). The repository remembers the user's intent and therefore provides a better representation of each image in the database in terms of the semantic meanings. The semantic similarity between the query image and each database image can then be computed using the current feedbacks and the semantic values in the FTR. Furthermore, feature re-weighting is applied on the session-term feedback to learn weight of low level features. Then we use the weighted Euclidean distance metric to measure the distance between the query image and each database image. These two similarity measures are normalized and combined together to form the overall similarity measure.

Our experimental results show that the average precision of the proposed system exceeds 83% after three iterations.

1. Introduction

Due to the recent improvements in digital photography and storage capacity, storing large amounts of images has been made possible. As a result, there is an increasing demand for image management tools. Content-based Image Retrieval (CBIR) systems use automatically extracted low-level features such as color, texture and shape for retrieval. Extensive experiments on CBIR systems show that low-level contents often fail to describe the high level semantic concepts in user's mind. Therefore, the performance of CBIR is still far from user's expectations [1].

In order to improve the retrieval accuracy of content-based image retrieval systems, research focus has been shifted from designing sophisticated low-level features extraction algorithms to reducing the 'semantic gap' between the low-level features and high-level semantics [2]. In order to reduce the gap, different techniques were introduced such as Relevance Feedback (RF). RF is a powerful tool traditionally used in text-based information retrieval systems. It was introduced to CBIR, with the intention to bring user in the retrieval loop to reduce the semantic gap. In this technique different approaches are used to learn the user's feedbacks. A typical approach is to adjust the weights of low-level features (feature re-weighting) [3-6]. Feature re-weighting dynamically updates the weights embedded in the query to model the high-level concepts. Another method is called query-point-movement (QPM) [7]. QPM improves the estimation of the query point by moving it towards the positive examples and away from the negative examples. Recently Machine learning techniques such as SVM [8] are also used for concept learning. SVM is often used to capture the query concept by separating the relevant images from the irrelevant images using a hyper-plane in a projected space. Generally, the labeled samples provided by the user are limited, and such small training data set will result in weak classification of database images (as relevant/irrelevant). In [9], D-EM (Discriminant-EM) is used to solve this problem.

Noticeable point is that, purpose of all of these methods is to improve the retrieval performance of the current query session without learning previous user's behaviors. To overcome this limitation, another school of thought generally called long-term learning, has become available in recent years. They memorize and

978-1-4244-2673-7/08/$25.00 ©2008 IEEE

accumulate users' preferences in the RF process. The historical retrieval experience will then be used to guide new users' queries. Indeed, long-term learning algorithms are based on previous users' behaviors, which basically embody more semantic information than low-level features. *Han* [10] and *Zhuang* [11], introduce a knowledge memory model to store the semantic information by simply accumulating user-provided interactions. In the process of image retrieval, system converts user's feedbacks to some semantic relations between images, and then a semantic relation network is constructed. [10], [11] are effective approaches for learning the semantic relationships between the images, but they have some problems: 1-They store the relationship between each two images, so it requires O(n^2) space, where n is the number of images in the database. 2-These learning techniques are susceptible to images being mislabeled by the user, where the relationship between two images will therefore be learned incorrectly.

To address the limitations of current CBIR systems, we introduce an image retrieval system that uses fuzzy decision based on transaction repository and feature re-weighting technique. The proposed system integrates the user's feedback from all iterations to construct a Fuzzy Transaction Repository (FTR). We then refine the query by updating its associated weight vector using current feedbacks and FTR. Then, the semantic similarity between the query and each database image is computed using weighted combination of fuzzy membership function. Furthermore, to improve short term learning, feature re-weighting is applied on the session-term feedback to learn the weights of low level features. We finally return retrieval results by combining the normalized similarity scores computed from both fuzzy transaction-based and feature re-weighting. The semantic-space-based [12, 13] and the log based [14] are the other systems that integrate the log information of user feedbacks with RF for image retrieval. Our proposed model outperforms these systems. Dominant features of the proposed systems are as follows:

First, we dynamically construct a fuzzy transaction repository by recording each session-term feedback. This is important for us because, FTR is not only limited to the existing images in database. So a new image, with a new concept can be added to database in our proposed model, also its semantic information is further added. Indeed, FTR is updated for both of existent and recently added images, continually. Second, while most systems that use RF, want user to mark retrieval images as relevant or irrelevant images and this decision would be difficult for images which are rich in semantic, we use soft feedback model in retrieval process that allows user to judge more truly. According to this feedback, we use fuzzy decision making for computing semantic similarity between query image and each database image. Third, for retrieving results, we integrate both low level features and high level concepts. To this end, we use low level features re-weighting technique.

The remainder of the paper is organized as follows. Section 2 describes our proposed CBIR system in detail. Section 3 illustrates the experimental results. Section 4 draws conclusion.

2. The proposed system

The block diagram of our proposed system is shown in Fig. 1. The system first computes low level features of the query image and returns 20 images with the highest similarity scores to the user. The user labels examples according to the relevance of each retrieved image to the query image. The transaction based semantic learning uses session-term feedback and the FTR to estimate the query semantic feature. The feature re-weighting technique also uses session-term feedback to learn the weight of low level features. The system then returns top 20 images ranked by fusing the normalized scores computed from both techniques. The user labels each returned image for the next iteration. The process will be continued and refined iteratively until the user is satisfied. The following subsections will explain our proposed system in details.

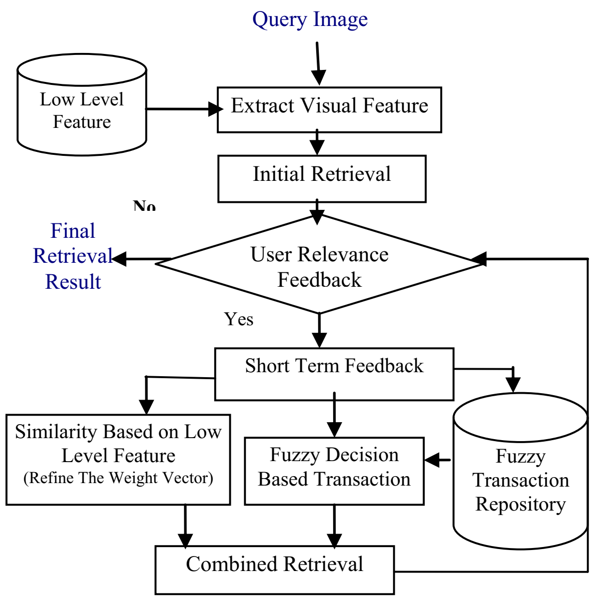

Figure 1. The block diagram of our proposed system

2.1. Feature extraction

We use the sets of features to represent the images. This set contains three types of features, color, shape, and texture. The color feature used in our experiments is color moment. It is a 9-dimension feature vector where 3 moments (color mean, color variance and color skewness) in each color channel (H, S, and V) are extracted from each image. We use edge direction histogram for image shape feature. It is an 18-dimension feature vector. The texture feature used is the co-occurrence feature. It is a 16-dimension feature by analyzing the Co-occurrence matrix. The weighted Euclidian distance is used to measure the similarity between the query and each image in the database. Initially, the weight vector is set to 1 and updated in next iterations according to relation (11) in section 2.4.

2.2. Fuzzy Transaction Repository

2.2.1. Fuzzy labeling. Most existing image retrieval methods assume that images have binary memberships in semantic classes. So they assign a single semantic label -Crisp Label- to each image. While, images may belong to many classes with different degrees of relevance. So, we used fuzzy labeling for user's feedback. The fuzzy labeling or soft feedback model provides more flexibility for users, especially when the query or images are semantically rich.

Our proposed FTR include fuzzy labels that obtained from relevance feedback procedure. Five types of fuzzy labels are used in our feedback process; include Full Irrelevant (FIR), Irrelevant (IR), Don't Care (DC), Full Relevant (FR), and Relevant (R). FIR, IR, FR and R membership function are defined to form of trapezoidal form and DC is defined to triangular form.

We will consider the following linear membership function related to each triangle fuzzy sample: [15]

$$\mu_i(y)=\begin{cases}\dfrac{y-(Y_i+d_i)}{d_i} & Y_i-d_i\le y\le Y_i\\[2mm] \dfrac{Y_i+d_i-y}{d_i} & Y_i\le y\le Y_i+d_i\\[2mm] 0 & y\ge Y_i+d_i,\ y\le Y_i-d_i\end{cases} \quad (1)$$

where d_i is tolerance of i[th] input vector and $d_i\in[0,1]$ which is shown in Fig 2.

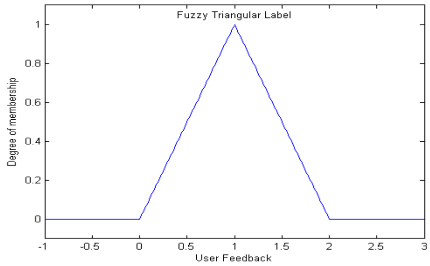

Figure 2. User feedback to form of triangle according to relation (1) with $Y_i=1$ and $d_i=1$.

Also for trapezoidal fuzzy samples we can write:

$$\mu_i(y)=\begin{cases}\dfrac{y-(Y_i+d_i)}{d_i} & Y_i-d_i\le y\le Y_i\\[2mm] \dfrac{Y_i+d_i-y}{d_i} & Y_i+b_i\le y\le Y_i+b_i+d_i\\[2mm] 0 & y\ge Y_i+d_i,\ y\le Y_i-d_i\\ 1 & Y_i\le y\le Y_i+b_i\end{cases} \quad (2)$$

where b_i is width of trapezoidal membership function as an example in Fig 3.

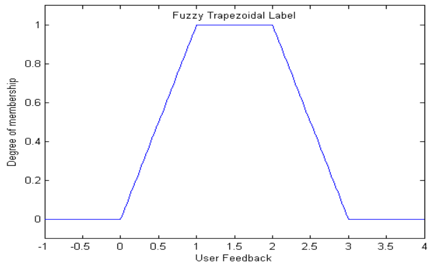

Figure 3. User feedback to form of trapezoidal according to relation (2) with $Y_i=1$, $d_i=1$ and $b_i=1$

2.2.2. Fuzzy Transaction Repository construction. FTR stores user's feedback. Each row in the FTR represents an image in the database and each column corresponds to one semantic group. Initially FTR is empty and dynamically is constructed as follows:

1. For each query image *Q*
   1.1 Append a new row and new column to FTR, new row means that new images is added to database, new column means that a new concept is added to database.
   1.2 Retrieve images using low-level features and return 20 images most similar to query image *Q*.

1.3 The relevance feedback mechanism solicits the user to judge the relevance of the retrieved images.
1.4 According to user's feedback, the elements corresponding to the rows of all full relevant images and relevant images are respectively set to 2 and 1, and the elements corresponding to the full irrelevant and irrelevant images are respectively set to -2 and -1 and the remaining elements are set to 0.
1.5 Compute the semantic similarity score between query image Q and each database image using the fuzzy transaction-based semantic learning (section 2.3).
1.6 Compute the visual similarity score between query image Q and each database image using the feature re-weighting learning technique (section 2.4).
1.7 These two similarity measures combined together to form the overall similarity measure (section 2.5).
1.8 Repeat steps 1.3 through 1.8 until the user is satisfied with the retrieval results or the maximum iteration is reached. If iteration numbers exceeded from maximum, that means, new semantic group has been added to database that previously, no images or few images of the semantic group exist in database. Along database images are increased, images that belong to this semantic group are gradually added to database.

2. Finally, the new column is compared with all existing columns in FTR to determine if there is any image that's full relevant in both of columns. If any match is found-that means the new column has similar semantic to that column-, information of the new column is merged with the information of identified column. Now we can remove the new column. Otherwise new column is remained. This process is useful because, dimensions of FTR and also searching time are reduced.

2.3. Semantic similarity based on Fuzzy Transaction Repository

We introduced our fuzzy decision based transaction repository in this section. As aforementioned, for each new query we retrieve images using low-level features in the first iteration. The user labels retrieval results. Each labeled image is represented by a semantic vector x^j with $j = 1,...,s$ for the relevant images and $j = s+1,...,s+t$ for the irrelevant images. Each semantic vector corresponds to a row vector of FTR, which is determined by the index number of the labeled images. Then the high-level feature vector of the query (Q) is initialized as:

$$Q = (q_1, q_2, ..., q_n) \tag{3}$$

where n equals the number of columns in FTR , q_i is i^{th} element of the query semantic feature vector and defined as follows:

$$q_i = (x_i^1 \vee x_i^2 \vee ... \vee x_i^s) \wedge (\overline{x_i^{s+1} \vee x_i^{s+2} \vee ... \vee x_i^{s+t}}) \tag{4}$$

where x_i^j *is* i^{th} element of the semantic feature vector of the j^{th} image. The x_i^j with a value of -1 or -2 will be treated as 0's in this computation.

Initially, we set $w = Q$, where w represents the weight vector associated with the query.

In the previous systems [12-14], the system calculates the similarity score between the query and each image in the database using (5).

$$h_{score}(x) = (w \cdot x) = \sum_{j=1}^{n} w_j x_j \tag{5}$$

where w represents the weight vector associated with the query, x represents the semantic vector of an image in the database and n is the length of both vectors.

In the proposed FTR w_i is used for weighting of fuzzy membership function. Result of weighting is:

$$\mu_R(x) = \sum_{i=1}^{n} w_i \mu_i(x) \tag{6}$$

where $\mu_i(x)$ is fuzzy membership function of i^{th} element of the semantic feature vector x, w is the weight vector, n equals the number of columns in FTR and $\mu_R(x)$ is result of adding weighted membership function in FTR. Relation (6) is fuzzy form of relation (5). Fig 4 shows summation of two weighted membership function.

Afterwards, results will be defuzzified. We use Centroid defuzzification according follow,

$$R = \frac{\sum_{\langle x \rangle} x \mu_R(x)}{\sum_{\langle x \rangle} \mu_R(x)} \tag{7}$$

As an example, R in Fig 4 is -1.1667.

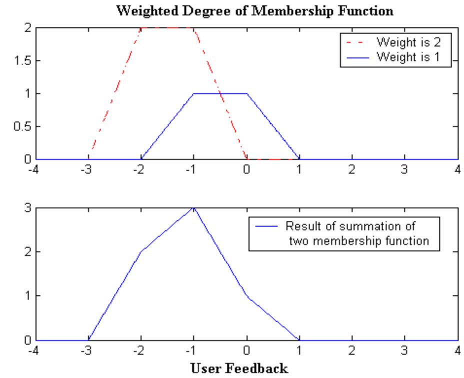

Figure 4. Example of adding weighted membership function

For the following feedback iterations, the short-term learning refines the query by updating its weight vector using current feedbacks and FTR as follows:

Positive Feedback:

$$w_i^{(t+1)} = \begin{cases} \alpha w_i^{(t)}, & if \quad x_i = 1 \quad and \quad w_i^{(t)} \neq 0 \\ 2 * \alpha w_i^{(t)}, & if \quad x_i = 2 \quad and \quad w_i^{(t)} \neq 0 \\ 1, & if \quad (x_i = 1 \quad or \quad 2) \quad and \quad w_i^{(t)} = 0 \\ w_i^{(t)} & if \quad x_i = 0 \\ \dfrac{w_i^{(t)}}{\alpha}. & if \quad x_i = -1 \\ \dfrac{w_i^{(t)}}{2 * \alpha}. & if \quad x_i = -2 \end{cases} \tag{8}$$

Negative Feedback:

$$w_i^{(t+1)} = \begin{cases} 2 * \alpha w_i^{(t)} & if \quad x_i = -2 \quad and \quad w_i^{(t)} \neq 0 \\ \alpha w_i^{(t)} & if \quad x_i = -1 \quad and \quad w_i^{(t)} \neq 0 \\ 1 & if \quad (x_i = -1 \quad or \quad x_i = -2) \quad and \quad w_i^{(t)} = 0 \\ \dfrac{w_i^{(t)}}{2 * \alpha}. & if \quad x_i = 2 \\ \dfrac{w_i^{(t)}}{\alpha}. & if \quad x_i = 1 \\ w_i^{(t)}, & if \quad x_i = 0 \end{cases} \tag{9}$$

where $w_i^{(t)}$ is the i^{th} element of the current weight vector, $w_i^{(t+1)}$ is the i^{th} element of the updated weight vector, x_i is the i^{th} element of the hidden semantic feature vector of the labeled image x, and α is the adjustment rate and is empirically set to 1.1.

2.4. Feature re-weighting

There are several ways for measuring distance between a query and each database image such as Minkowski distance, Quadratic distance and Euclidean distances. Each of them has its own merits and demerits. We chose weighted Euclidean distance because it is computationally very simple and produces fairly good results. If Q is query image and I is database image, weighted Euclidean distance is computed as follows:

$$D(I,Q) = \left(\sum_{i=1}^{M} w_i * \left|f_{iI} - f_{iQ}\right|^2\right)^{1/2} \tag{10}$$

where f_{iI} and f_{iQ} are i^{th} feature component of I and of Q respectively, M is the length of feature vector and w_i is weight factor. When there is no RF, equal weight values are used for each feature component. For the following feedback iterations, these weights are updated as follows:

$$w_i^{t+1} = \frac{\delta_i^t}{\sigma_i^{t,R} + \varepsilon} \tag{11}$$

$$\varepsilon = 0.00001$$

where w_i^{t+1} is weight factor in $t+1^{th}$ iteration, $\sigma_i^{t,R}$ is the standard deviation of i^{th} component of feature vector into the set of relevant images (Full Relevant and Relevant), δ_i^t is the discriminant ratio of the i^{th} component of feature vector and is defined as follows:

$$\delta_i^t = 1 - \frac{\sum_{l=1}^{t} \left|\psi_i^{l,U}\right|}{\sum_{l=1}^{t} \left|F_i^{l,U}\right|} \tag{12}$$

where $\sum_{l=1}^{t} \left|\psi_i^{l,U}\right|$ is the number of non-relevant images (Full Irrelevant and Irrelevant), located inside the dominant range of relevant samples and $\sum_{l=1}^{t} \left|F_i^{l,U}\right|$ is the total number of non-relevant images among the retrieved images, for the i^{th} feature component. The confusion set of i^{th} feature component after the t^{th} iteration is given by:

$$\psi_i^{t,U} = \left\{ f_i^t \mid f_i^t \in \Phi_i^t \quad and \quad f_i^t \in F_i^{t,U} \right\} \tag{13}$$

For all irrelevant images we stack i^{th} feature component into the set $F_i^{t,U}$, the dominant range of relevant images on the axis of the i^{th} feature component are defined as follows:

$$\begin{aligned} \Phi_i^t &= [\phi_i^{t,1}, \phi_i^{t,2}] \\ \phi_i^{t,1} &= \min(F_i^{t,R}) \\ \phi_i^{t,2} &= \max(F_i^{t,R}) \end{aligned} \quad (14)$$

2.5. Retrieval results

After the first retrieval result is prepared, similarity between query image and each database image is computed using both visual similarity and semantic based similarity. Visual similarity ($Score_{low}$) and semantic similarity ($Score_{high}$) -computed using (7)-, are combined to obtain final similarity measure:

$$\begin{aligned} Score_{low}(I,Q) &= -D(I,Q) \\ Score(I,Q) &= Score_{high}(I,Q) + Score_{low}(I,Q) \end{aligned} \quad (15)$$

3. Experimental results

We have tested our proposed system on a general purpose images with one thousand images from COREL. These images have 10 categories with 100 images in each category. Fig. 5 shows a scheme of the proposed interactive retrieval system with first image as a query.

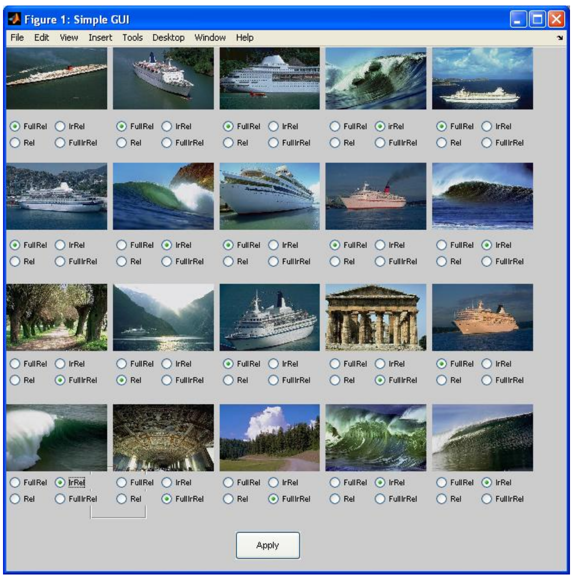

Figure 5. Image retrieval system

In order to evaluate CBIR system two important criterions are Precision and Recall. Precision is the fraction of the relevant images which has been retrieved and recall is the ratio of relevant retrieved images over all relevant images. We use the first criterion to evaluate our system. Fig. 6 summarizes the average retrieval precision for our proposed system (transaction based and feature re-weighting) and feature re-weighting system. It shows that precision of our proposed approach exceeds 91% after 6 iterations. While after 6 iterations retrieval precision of the feature re-weighting approach reaches only 67%. Thus, our approach is able to accomplish the retrieval goal in only a few iterations. This improvement is preferred in image retrieval since the user wants to retrieve the desired images in as few feedback steps as possible.

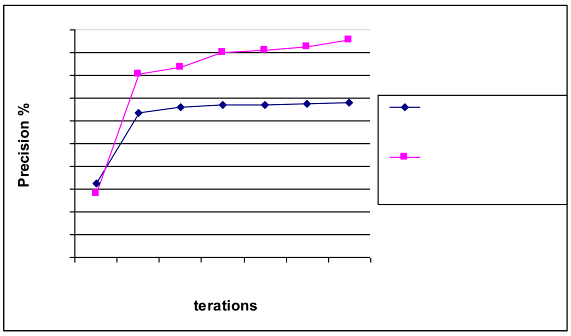

Figure 6. Comparisons of our proposed approach and feature re-weighting approach

Furthermore experiments show that, the retrieval precision increases as the information of FTR increase and each iteration leads to better retrieval precision. For this purpose, we show experiments in three steps. In the first step FTR is empty and we show the average of retrieval precision on query images, in the second step we will start test on a pre-built FTR from step1. Finally, in the third step we will test on a pre-built FTR from step1 and step2. So Fig. 7 compares the retrieval precision for our proposed approach on an empty FTR and a pre-built FTR (step1) and a pre-built FTR (step1 & step2). It clearly shows, a pre-built FTR from step1 & step2 has better precision on the 2th iteration and achieves retrieval precision as high as 95% on the 7th iterations.

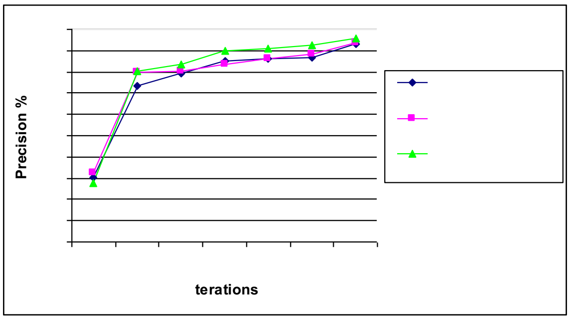

Figure 7. Comparison of retrieval precision on empty repository and pre-build repository

4. Conclusion

We have proposed image retrieval system based on transaction based semantic learning and low level features learning. The important contributions to the field in this work can be summarized as follows:

1. We construct fuzzy transaction repository dynamically to store user's relevance feedback information.
2. Since FTR is dynamically constructed, it is possible to add new images with new concepts to existing database images. Their semantic information is further added to FTR during future sessions.
3. We develop an incremental method to deal with new log sessions, it is important toward a long term learning purpose.
4. Using soft feedback model in retrieval sessions that allows user to judge about the relevance of the returned results more truly.
5. The semantic similarity can be computed using the current feedback and the semantic values in the FTR.
6. Applying the feature re-weighting learning to find images which are visually similar to the query image.

Experiments show that our proposed system has a desirable performance and achieves remarkably high retrieval precision after the first three iterations.

5. References

[1] X.S. Zhou, T.S. Huang, "CBIR: From Low-level Features to High-level Semantics", Proceedings of the SPIE, Image and Video Communication and Processing, San Jose, CA, vol. 3974, pp. 426–431, January 2000

[2] Y. Liua, D. Zhang, G. Lua, W.Y. Mab, "A Survey of Content-Based Image Retrieval with High-level Semantics" ; Proceedings of the Pattern Recognition, pp. 262–282, 2007

[3] S.K. Saha, A.K. Das, B. Chanda, "Image Retrieval Based on Indexing and Relevance Feedback" , Pattern Recognition Letters 28 ,357–366, 2007

[4] M. Lamine Kherfi, D. Ziou, "Relevance Feedback for CBIR: A New Approach Based on Probabilistic Feature Weighting With Positive and Negative Examples", IEEE Transactions on Image Processing, Vol. 15, No. 4 , April 2006

[5] A.K. Jiang, G. Er, Q. Dai, J. Gu, "Similarity-Based Online Feature Selection in Content-Based Image Retrieval", IEEE Transactions on Image Processing, Vol. 15, No. 3, March 2006

[6] Y. Wu, A. Zhang, "A Feature Re-Weighting Approach for Relevance Feedback in Image Retrieval", Special issue on Segmentation, Description, and Retrieval of Video Content, Rochester NewYork (September 2002)

[7] G. Guo, A.K. Jain, W. Ma, H.J. Zhang, "Learning similarity measure for natural image retrieval with relevance feedback", IEEE Trans. Neural Networks 13 (4) (2002) 811–820.

[8] S. Tong, E. Chang, "Support Vector Machine Active Learning for Image Retrieval"; In: Proc. of ACM Int. Conf. on Multimedia, pp. 107–118. ACM Press, New York (2001)

[9] Q. Tian, Y. Yu, T.S. Huang, "Incorporate discriminant analysis with EM algorithm in image retrieval", Proceedings of the International Conference on Multimedia and Expo (ICME), 2000, pp. 299–302.

[10] J. Han, K.N. Ngan, M. Li, H.J. Zhang, "A Memory Learning Framework for Effective Image Retrieval", IEEE Transactions on Image Processing, Vol. 14, No. 4, April 2005

[11] Y. Zhuang, J. Yang, Q. Li, "A Graphic-Theoretic Mode for Incremental Relevance Feedback in Image Retrieval", Proc. Of 2002 Int Conf. on Image Processing, New York, pp.413-416, Sep 2002

[12] X. He, Q. King, W.Y. Ma, M. Li, H.J. Zhang, " Learning a Semantic Space From User's Relevance Feedback for Image Retrieval ", IEEE Transactions on Circuits and Systems for Video Technology , Vol. 13, No. 1, January 2003

[13] Xiaojun. Qi, Ran. Chang, "Image Retrieval Using Transaction-Based and SVM-Based Learning in Relevance Feedback Sessions", ICIAR 2007, LNCS 4633, pp. 638–649, 2007.© Springer-Verlag Berlin Heidelberg 2007

[14] S. Hoi, M.R. Lyu, R. Jin, "A Unified Log-Based Relevance Feedback Scheme for Image Retrieval", IEEE Transactions on Knowledge and Data Engineering, Vol. 18, No. 4, April 2006

[15] L.A. Zadeh, Fuzzy Sets. Information and Control 8 (1965), 338-35