A Semantic Feedback Framework for Image Retrieval

Malihe Javidi, Hamid Reza Pourreza, Hadi Sadoghi Yazdi, *Member IACSIT*

***Abstract*—In this paper, a new system of fuzzy relevance feedback for image retrieval is introduced. In conventional CBIR systems, the users are restricted to make a binary labeling on the retrieval results, while this determination is difficult for rich images in semantic. In the proposed system, we accumulate user interactions using a soft feedback model to construct Fuzzy Transaction Repository (FTR). The repository remembers the user's intent and, therefore, in terms of the semantic meanings, provides a better representation of each image in the database. To best exploit the benefits of user feedback, we improved the proposed system, so that the repository remembers the user's intent in a suitable manner (as structure-based fuzzy transaction repository) and provides an accurate representation for each image in the database. The semantic similarity between the query and each database image can then be computed using the current feedback and the semantic values in the FTR. Furthermore, feature re-weighting is applied to the session-term feedback in order to learn the weight of low-level features. 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 systems exceeds 83% after three iterations.**

***Index Terms*— Feature Re-weighting, Fuzzy Relevance Feedback, Fuzzy Transaction Repository, Image Retrieval**

I. Introduction

Content-based image retrieval (CBIR) is a process of retrieving a set of desired images from a collection of images based on visual contents present in the images, such as color, texture, shape or spatial relationship. Extensive experiments on CBIR systems show that the retrieval accuracy of today's CBIR systems remains relatively unsatisfactory [1]. CBIR systems interpret user information needs based on a set of low-level visual features extracted from the images [2]. However, these features may not correspond to the user's interpretation and understanding of image contents. Indeed, the semantic gap between the low-level representation of images and high-level user concepts poses great challenges in a CBIR system design.

To bridge the gap between low-level features and semantic meaning, relevance feedback techniques have been incorporated in the process of image retrieval [3]. Indeed, the relevance feedback approach has been proposed to semi-automatic annotation and image retrieval [4]. Early relevance feedback schemes can be classified into two categories: query point movement (QPM) or query refinement [5], [6] and feature re-weighting method [7]-[10].

The query point movement was proposed by Rocchoio [5]; it improves the estimation of the query point by moving it towards positive examples and away from the negative ones. The best-known implementation is the MARS [11]. It assumes that the query is represented as a point in the multi-dimensional feature space and that all relevant images cluster in the feature domain. However, in most cases, it is neither sufficient nor effective to describe the user information need using a single global model.

The feature re-weighting method considers the discriminating power of different features and enhances the contribution of features that identify the relevant examples marked by the user in the best way. It uses heuristic formulation with empirical parameter adjustment, mainly along the line of independent axis weighting in the feature space.

Recently Learning based approaches are typically used to appropriately modify the feature set or similarity measure. Bayesian or probabilistic learning [12], [13] has been utilized to incorporate user feedback to update the probability distribution of all images in the database. Artificial neural networks [14], [15] have also been adopted in the relevance feedback process due to their ability to simulate universal mapping. Since the number of feedback samples from the users is usually small in respect to feature dimensions, neural networks suffer from this limitation, as they require large volumes of training data. *Tong* and *Chang* in [16] propose a support vector machines (SVM) that is an active learning algorithm for image retrieval. To address the problems of insufficient training samples, the D-EM (Discriminant-Expectation Maximization) algorithm [17] is introduced to incorporate unlabeled data in order to enhance performance in CBIR. The results are promising, but computation complexity can be significant for a large database. In [18] a complete review of recent works in RF has been presented.

Traditionally, the user is restricted to binary classification to determine whether an image is "fully relevant" or "totally irrelevant." Therefore, a single semantic label or namely a crisp label is assigned to each image.

An additional point to note is that all these methods are applied to improve the retrieval performance of the current query session without learns the previous users' behaviors. To overcome this limitation, another method, generally

Manuscript received July 26, 2009; accepted September 24,2009.
M. Javidi is with the Machine Vision Lab from Computer Engineering Department of Ferdowsi University of Mashhad, Mashhad, Iran. (e-mail: malihejavidi@ gmail.com).
H.R. Pourreza is with Computer Engineering Department of Ferdowsi University of Mashhad, Mashhad, Iran. (phone: (+98) 511 876 3306; fax: (+98) 511 862 6541; e-mail: hpourreza@ um.ac.ir).
H. Sadoghi Yazdi Author is with Computer Engineering Department of Ferdowsi University of Mashhad, Mashhad, Iran. (phone: (+98) 511 876 3306; fax: (+98) 511 862 6541; e-mail: h_sadoghi@ um.ac.ir)

called long-term learning, has been recently developed. This involves the memorization and accumulation of user's preferences in the RF process. The historical retrieval experience is then used to guide new users' queries. Indeed, long-term learning algorithms are based on previous users' behavior, which embody more semantic information than low-level features. *Han* [19] and *Zhuang* [20] introduce a knowledge memory model to store semantic information by accumulating user-provided interactions. In the process of image retrieval, the system converts user feedback to semantic relations between images, to do that, it constructs a semantic relation network. [19], [20] are effective approaches for learning about semantic relationships between images. However, these approaches have some problems:

- The relationship is stored between each of the two images and so requires $O(n^2)$ space, where n is the number of images in the database.
- These learning techniques are susceptible to images mislabeled by the user, which result in the relationship between two images to be learned incorrectly.

To address the limitations of current CBIR systems, we introduce an image retrieval system that uses fuzzy decision based on a transaction repository and feature re-weighting technique. The proposed system integrates user feedback from all iterations in order to construct a Fuzzy Transaction Repository (FTR). To optimally exploit the benefits of user feedback, we also improved the proposed system. In the improved system, the repository remembers the user's intent as a new structure and therefore provides an accurate representation for each image in the database. The semantic similarity between the query and each database image is then computed using a weighted combination of the fuzzy membership function. Furthermore, to improve short term learning, feature re-weighting is applied to the session-term feedback in order to learn the weights of low level features. The retrieval results are returned by combining the normalized similarity scores computed from both the fuzzy transaction-based and the feature re-weighting. The semantic-space-based [21]-[23] and the log-based [24] are other systems that integrate the log information of user feedback with RF for image retrieval. Our proposed model improves these systems. The dominant features of the proposed systems are as follows:

First, a fuzzy transaction repository is dynamically constructed by recording each session-term feedback. This is important because FTR is not only limited to the existing images in the database. Therefore, a new image, with a new concept, can be added to the database in the proposed model. Also, its semantic information can be further added. Indeed, FTR is updated continually for both of existing and recently added images. Second, most systems using RF prefer that the user mark retrieval images as relevant or irrelevant, whereas this determination is difficult for images rich in semantic. The fuzzy feedback model is applied in the retrieval process to allow the user to judge more accurately. According to this feedback, the fuzzy decision is used for computing the semantic similarity between the query and each database image. Third, to retrieve the results, both low-level features and high-level concepts are integrated. To this end, feature re-weighting technique is employed. Finally, because of the subjectivity of human annotation, the first system has been improved and the structure of the FTR changed. In the improved system different user's judgments for each image in each semantic group are completely stored.

The structure of the paper is organized as follows. Section 2 describes the proposed system in detail. For better comprehension of recommended method, we describe it into two separate primary and improved phase. Section 3 illustrates the experimental results. Section 4 draws conclusion.

II. The Proposed System

The system first computes the low-level features of the query image and then returns 20 ones with the highest similarity scores to the user. The system solicits the user to judge the relevance of the retrieved images. The user provides his evaluation by labeling each displayed image as 'full relevant,' 'relevant,' 'full irrelevant' and 'irrelevant.' Transaction-based semantic learning employs session-term feedback and the semantic values in FTR to estimate the query semantic feature. The FTR stores the high-level semantic features for each image in the database, which is constructed on-line by collecting user feedback from various query sessions. Furthermore, the feature re-weighting technique uses session-term feedback to learn the weight of low-level features. The system then returns the top 20 images which were ranked by fusing the normalized scores computed from both techniques. The user goes on to label each returned image for the next iteration. The process will be continued and refined iteratively until the user is satisfied. The following subsections explain the proposed system in detail.

A. Feature Extraction

Low-level sets of features are used to represent images. They contain 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 (mean, variance and skewness) in each color channel (H, S, and V) are extracted from each image. Edge direction histogram is used 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 analysing the co-occurrence matrix. The Euclidian distance is used to measure the similarity between the query and each image in the database.

B. Fuzzy Labeling

Most existing image retrieval methods assume that images have binary memberships in semantic classes. Therefore, they assign a single semantic or a crisp label to each image, while images may belong to many classes with different degrees of relevance. Hence, a novel fuzzy relevance feedback is proposed for user feedback which enables the user to make a fuzzy judgment. The fuzzy labelling has more flexibility for users, especially when the queries or images are semantically rich; it provides a natural and flexible way in expressing the user's preferences.

Our proposed FTR includes fuzzy labels obtained from the relevance feedback procedure. Five types of fuzzy labels are used in our feedback process: full irrelevant, irrelevant, full relevant, relevant, and Don't Care (DC). The four primary membership functions are defined by a trapezoidal membership function, while the DC is defined by a triangular one. Users judge the relevance degree of the retrieved images. When the image does not retrieve in any iteration, its FTR element is empty. These elements are called missing values. If the number of missing values is few, instances with missing values can be discarded, while, there are many images of this form, especially in the first query sessions. Therefore, the DC membership function is used to solve the missing value problem in the transaction logs' data.

C. First Phase: Semantic Image Retrieval Based on Fuzzy Transaction Repository

For better comprehension of recommended method, we describe it into two separate primary and improved phase. In this section, we describe the first phase. The improved phase will be detailed in the next section.

The membership functions of the fuzzy labels are determined as follows:

- Five different states are considered for user's feedback. These states are including: Relevant (R), Full Relevant (FR), IrRelevant (IR) and Full IrRelevant (FIR). The 5^{th} state belongs to the image does not appear as a retrieved image in any iteration, i.e. DC.
- The membership functions are shown in Fig. 1.

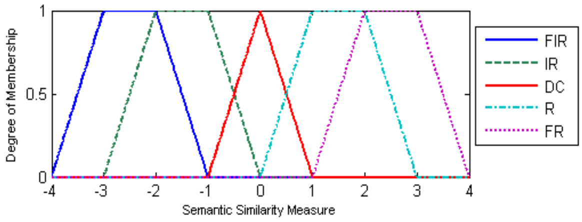

Figure 1. The membership functions of each fuzzy label in the first phase of proposed system

Fuzzy Transaction Repository construction (First Phase): The FTR stores user feedback. Each row in the FTR represents an image in the database and each column corresponds to a semantic group. Initially the FTR is empty and it is constructed dynamically as follows:

1) For each query image Q :

- Append a new column to the FTR. A new column signifies that a new concept have been added to the database.
- Retrieve images using low-level features and return 20 ones which are the most similar to a query Q.
- The relevance feedback mechanism solicits the user to judge the relevance of the retrieved images.
- According to the user feedback, the elements corresponding to the rows of all full relevant and relevant images are set to 2 and 1, but, the full irrelevant and irrelevant ones are set to -2 and -1 respectively. Remaining elements are set to 0.

The elements of the FTR have numeric values, but the membership functions, introduced in the proposed FTR system, are used to compute semantic similarity.

- Compute the semantic similarity score between a query Q and each database image by using fuzzy transaction-based semantic learning.
- Compute the visual similarity score between a query Q and each database image by using feature re-weighting learning technique (section E).
- The two similarity measures are combined together to form the overall similarity measure (section F).
- Repeat steps *c* through *h* until the user is satisfied with the retrieval results or when the maximum iteration is reached. If iteration numbers exceed the maximum, it means that, a new semantic group has been added to the database, which previously contained no images or few ones of a semantic group. As database images are increased, images that belong to this semantic group are gradually added to the database.

2) Finally, the new column is compared with all existing columns in the FTR to determine if there is any correlation between new column and other existing columns. If the correlation value is greater than 80%, it means that the new column is semantically similar to the other. Therefore the information of the new column (fuzzy label of images assigned by user in current session) has been merged with the information of the identified column, so the new column can be removed; otherwise it will remain. This process is effective because FTR dimensions and search time are reduced.

Semantic Similarity Based on Fuzzy Transaction Repository (First Phase): Fuzzy decision-based transaction repository is introduced in this subsection. As mentioned, for each new query, images are retrieved using low-level features in the first iteration. The user then 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 ones. Each semantic vector corresponds to a row vector of the FTR, which is determined by the index number of the labeled image. Then, the high-level feature vector of the query (Q) is initialized as:

$$Q=(q_1,q_2,...,q_n) \tag{1}$$

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

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

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 and -2 is be treated as 0's, but, the x_i^j with a value of 1 and 2 is be treated as 1's in this computation.

Initially, w is equal to Q $(w = Q)$, where w represents the weight vector associated with the query. In the proposed systems, the similarity score between the query and each image in the database calculates using Eq.3.

$$\mu_R(x)=\sum_{i=1}^{n} w_i \mu_i(x) \tag{3}$$

where $\mu_i(x)$ is the fuzzy membership function of the i^{th} element of the semantic feature vector x , w_i is used for weighting the fuzzy membership function, n equals the number of columns in the FTR and $\mu_R(x)$ is the result of adding the weighted membership function in the FTR. Fig. 2 shows the summation of two weighted membership functions according to the Eq. (3):

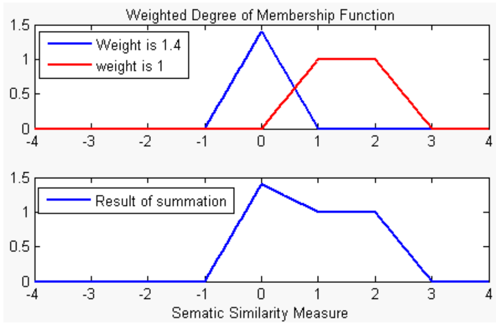

Figure 2. Example of adding weighted membership functions (convex form)

The result of summation of membership functions shown in Fig. 2 becomes non-convex form, but the experiments show that it is appeared in convex form too, Fig. 3. So the output of the Eq. (3) can be convex or non-convex form.

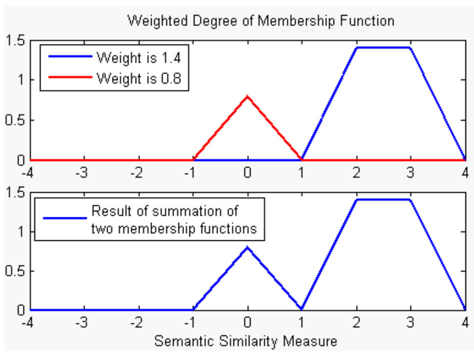

Figure 3. Example of adding weighted membership functions (non-convex form)

Therefore, results are defuzzified. The Centroid defuzzification is used according to the following:

$$R = \frac{\sum_{\langle x \rangle} x \mu_R(x)}{\sum_{\langle x \rangle} \mu_R(x)} \quad (4)$$

For example, R in Fig. 2 is -1.1667.

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

Positive Feedback:

$$w_i^{(t+1)} = \begin{cases} \alpha w_i^{(t)}, & if \quad x_i \in \{R\} \quad and \quad w_i^{(t)} \neq 0 \\ 2 * \alpha w_i^{(t)}, & if \quad x_i \in \{FR\} \quad and \quad w_i^{(t)} \neq 0 \\ 1, & if \quad (x_i = \in \{R \, or \, FR\} \quad and \quad w_i^{(t)} = 0 \\ w_i^{(t)} & if \quad x_i \in \{DC\} \\ \frac{w_i^{(t)}}{\alpha}. & if \quad x_i \in \{IR\} \\ \frac{w_i^{(t)}}{2 * \alpha}. & if \quad x_i \in \{FIR\} \end{cases} \quad (5)$$

Negative Feedback:

$$w_i^{(t+1)} = \begin{cases} \alpha w_i^{(t)} & if \quad x_i \in \{IR\} \quad and \quad w_i^{(t)} \neq 0 \\ 2 * \alpha w_i^{(t)} & if \quad x_i \in \{FIR\} \quad and \quad w_i^{(t)} \neq 0 \\ 1 & if \quad (x_i = \in \{IR \, or \, FIR\} \quad and \quad w_i^{(t)} = 0 \\ w_i^{(t)}, & if \quad x_i \in \{DC\} \\ \frac{w_i^{(t)}}{\alpha}. & if \quad x_i \in \{R\} \\ \frac{w_i^{(t)}}{2 * \alpha}. & if \quad x_i \in \{FR\} \end{cases} \quad (6)$$

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.

D. Improved Phase: Semantic Image Retrieval Based on Structured Fuzzy Transaction Repository (SFTR)

The problem of the first phase is that each element of the FTR determines whether the image in the semantic group is full relevant, relevant, full irrelevant or irrelevant and it is not possible to store different user judgments for each image of each semantic group. This problem motivates us to improve the proposed system. Due to the subjectivity of human annotation, the structure of the FTR has been changed. In the improved system, each element of the FTR is a data structure and it can be possible to store different user judgments for each image of each semantic group. For example, if three users determine that image i is "full relevant" to the query then, the full relevant term of the data structure of image i in the column is equal to 3. The importance feature and difference of the improved system will be detailed in this section.

In the improved system, five different states are considered for user feedback without any overlapping (Fig. 4). These membership functions will be updated in long term learning. So that, the importance of each option; relevant, full relevant, irrelevant and full irrelevant; will be increased by user feedback on it.

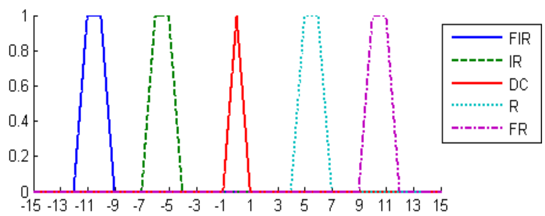

Figure 4. The membership functions of each fuzzy label in the improved system

Construction of SFTR (Improved Phase): Similar to the previous scheme, SFTR stores user feedback. Each row in the SFTR represents an image in the database and each column corresponds to a semantic group. Whereas, each element of the SFTR is a data structure and includes four terms: FIR, IR, R and FR. Initially, the SFTR is empty and, dynamically, it is constructed as follows:

All steps of this algorithm are similar to the previous, but step 'd' in previous is replaced with these two steps:

- Initially, empty the temporary feedback table. According to user feedback, add retrieved images to the temporary feedback table. The elements corresponding to the rows of all full relevant and relevant images are set to 2 and 1, but, the full irrelevant and irrelevant ones are set to -2 and -1 respectively. Remaining elements are set to 0.
- The temporary feedback table is compared with all existing columns in the SFTR to determine if there is any image that is full relevant in both of them. If any match is found, this means the new session is semantically similar to that particular column, and so information of the temporary feedback table is added to the information of the identified column. For example, if image i is identified FR in the current session, the FR element from image i in the identified column is increased by one unit. Similarly, if image j is identified IR, its IR element in the identified column is increased by one unit. In this way, information of the temporary table is added to the identified column. Otherwise, the information of the temporary feedback table is added to the SFTR as a new column or a new semantic group. This process is effective because, the information of the SFTR is continually updated after each iteration.

Similar to the previous algorithm, repeat steps *d* through *h* until the user is satisfied with the retrieval results or the maximum iteration is reached.

Semantic Similarity Based on SFTR (Improved Phase): According to the memorized knowledge in the SFTR, the semantic similarity between a query and each database image is computed in another manner. Also, the weight vector is updated differently. These differences will be illustrated in the following.

In the proposed SFTR, the Eq. (3) is improved as follows;

$$\mu_R(x) = \sum_{i=1}^{N} \sum_{\substack{j \in FR, R, \\ FIrR, IrR, DC}} W_i \, \mu_{ij}(x) \qquad (7)$$

Where $\mu_{ij}(x)$ is the fuzzy membership function of the i^{th} element from the j^{th} term of the semantic feature vector x, W is the weight vector, N equals the number of columns in the SFTR and $\mu_R(x)$ is the result of adding the weighted membership function in the SFTR. Similar to the first phase, the result of summation of membership functions can be convex or non-convex form. Therefore, the results of the Eq. (7) will be defuzzified by Centroid defuzzification to compute semantic similarity between a query and each database image.

An additional point to note is that, in the first phase of the proposed system, the fuzzy membership function is defined to form of trapezoidal and triangular functions, which are fixed during computations. While in the new scheme, according to the knowledge stores in the SFTR, it is possible to update these membership functions. For example, if most of the user determined image *i* is full relevant in the semantic group one, then the full relevant membership function of this image in this semantic group is invigorated. Also, according to user judgement, other membership functions of this image, namely relevance, irrelevance and full irrelevance are invigorated and then the Eq. (7) is computed.

Method of Membership Function Updating: The updating membership functions will be described as follows.

The full irrelevant membership function is:

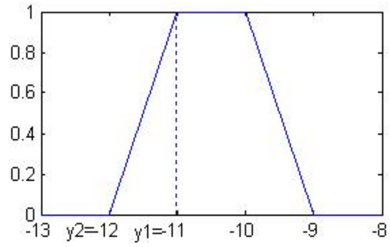

Figure 5: full irrelevant membership function

The width of the membership function is invigorated, and the following Eq. is utilized to update y_2 and y_1:

$$\begin{aligned} y_1 &= a * x + b_1 \\ y_2 &= a * x + b_2 \end{aligned} \qquad (8)$$

where x is the normalized FIR or IR or R or FR term of structure in SFTR. The values a, b_1 and b_2 are determined according to table 1:

TABLE 1: VALUES USED IN EQ. (8)

FIR	IR	R	FR
a=-2	a=-2	a=2	a=2
b1=-11	b1=-7	b1=7	b1=12
b2=-12	b2=-6	b2=6	b2=11

The values of a, b_1 and b_2 are empirically set, so that the center of gravity is moved 65% for each user feedback. e.g. if a user determine that image *i* is irrelevant and two ones specify that image *i* is full irrelevant, so the value *x* in the Eq. (8) for IR label and FIR label are 1/3 and 2/3 respectively. Fig. 6 shows the updated membership functions for two fuzzy labels, i.e. FIR and IR for image *i*.

As shown in Fig. 6, the membership function of FIR and IR are invigorated. Its invigoration is directed to the left side. So, the moment of membership function was appeared according to user feedback.

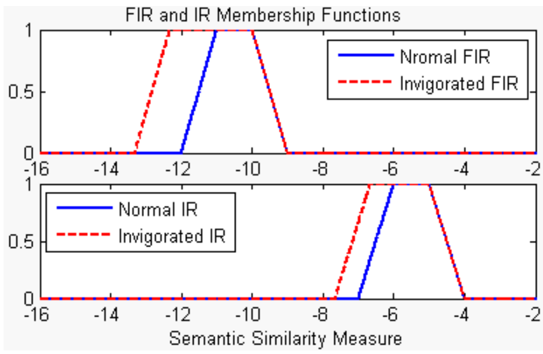

Figure 6: Normal and updated membership functions for FIR and IR

Also, the updated membership functions for two fuzzy labels, i.e. FR and R is shown in Fig. 7, where *x* in Eq. 8 is equal to 0.5 and 0.5 for R and FR, respectively.

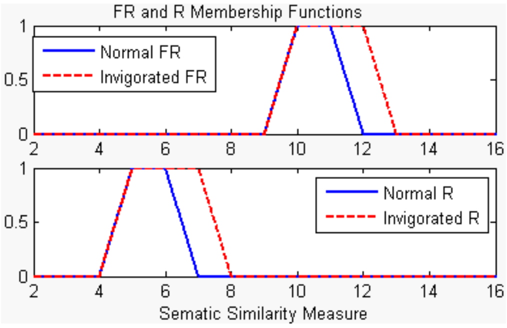

Figure 7: Normal and updated membership functions for R and FR

As shown in above Fig., the membership function of FR and R are invigorated. Its invigoration is directed to the right side. It is necessary to illustrate by an example as follows:

Example of Membership Function Updating: As an example, the following result was obtained for image *i* in the database during the first query sessions:

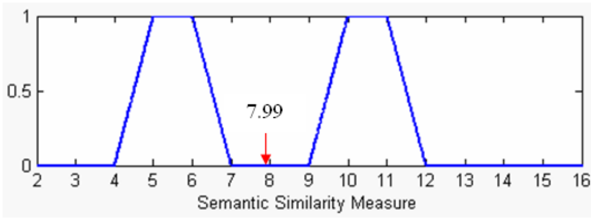

Figure 8: Semantic similarity measure for image i

As shown in Fig. 8, a user specify the image *i* as relevant one to the query and another users determine this image as full relevant one, so the centroid deffuzification for image *i* is equal to 7.99. After long term learning and obtaining different user feedback, the status of image *i* is as follows:

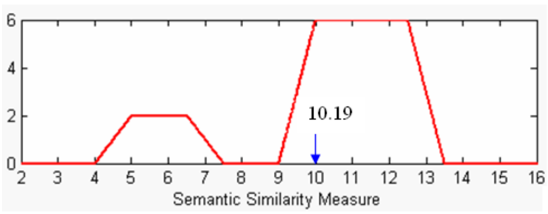

Figure 9: Semantic similarity measure for image i

The membership function of the FR label for image *i* is invigorated to right side more than the R label type, because the most users determine that image *i* is full relevant image. The centroid deffuzification for image *i* is equal to 10.19. Therefore, the way to update the membership functions is exactly based on user's opinion.

Method of Weight Updating: Since the SFTR has perfect knowledge, in comparison with the FTR, the weight vector is effectively updated as follows. For each positive and negative image, the related semantic feature vector in the SFTR is reviewed. For example, for image *i,* the semantic feature vector x_i^j in the SFTR for j=1,…, n is reviewed, where *n* equals the number of columns of the SFTR. Initially the terms of structure x_i^j are normalized and the max term is computed as follows:

$$x_i^j = \begin{cases} FR:2 \\ R:1 \\ IR:1 \\ FIR:0 \end{cases} \xrightarrow{\text{After normalization}}$$

$$x_i^j = \begin{cases} FR:2/4 \\ R:1/4 \\ IR:1/4 \\ FIR:0 \end{cases} \xrightarrow{\text{Max term}} \quad (9)$$

$$\max term = FR:2/4$$

If the max term and user feedback are similar, then this column must be invigorated. This is done as follows:

$$W_i^{(t+1)} = \lambda * W_i^{(t)} + (1-\lambda) * v \quad (10)$$

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, λ is the amplification coefficient equal to 0.95 and v is equal to the max term. For example, if a user determined image *i* as full relevant one, it means that, the user feedback and the max term are similar, for this reason this column is invigorated according to the Eq. (10). Otherwise, this column must be weakened. This is done as achieved:

$$W_i^{(t+1)} = \lambda * W_i^{(t)} - (1-\lambda) * s \quad (11)$$

where λ is similar to the Eq. (10) and s is equal to the term that is related to user feedback**.** For example, if a user determined image *i* as irrelevant one, it means that, the user feedback is not similar to the max term, so this column must be weakened.

E. *Feature Re-Weighting*

There are several ways to measure the distance between a query and each database image, such as the Minkowski distance, the Quadratic distance and the Euclidean distances. Each one has its own merits and demerits. In the proposed system, the weighted Euclidean distance is chooses [26], because it is computationally very simple and produces fairly good results.

F. *Retrieval Results*

After the first retrieval result is achieved, the similarity between the query and each database image is computed using both normalized visual and semantic-based similarity. Visual similarity ($Score_{low}$) and semantic similarity ($Score_{high}$), computed using Eq. 4, are normalized and combined to obtain the final similarity measure:

$$Score_{low}(I,Q) = -D(I,Q)$$

$$Score(I,Q) = NormScore_{high}(I,Q) + NormScore_{low}(I,Q)$$

(12)

III. Experimental Results

More than half of systems use a subset of Corel image dataset to test retrieval performance [3]. Corel image dataset contains a large amount of images of various contents ranging from animals and outdoor sports to natural sceneries. These images are pre-classified into different categories of size 100 by domain professionals. The proposed system has been tested on general purpose images with one thousand images from Corel. These images have ten categories with

100 images in each category.

In order to evaluate CBIR system, two important criterions are utilized: Precision and Recall. Precision is the fraction of the relevant images which have been retrieved while recall is the ratio of relevant retrieved images over all relevant ones. To evaluate our system, it can be realized by two points of view; so two different criterions are introduced. There is not any difference between relevant and full relevant retrieved images, so the first criterion is introduced as Normal Precision. Normal Precision is defined as follows:

$$Normal\ Precision = \frac{|Full\ Relevant + Relevant|}{|All\ Retrieved\ Images|} \tag{13}$$

On the other hand, the precision can be measured with emphasis on the full relevant retrieved images. Therefore, the second criterion is introduced as High Precision. High precision is defined according to Eq. (14):

$$High\ Precision = \frac{|2 \times Full\ Relevant + Relevant|}{|2 \times All\ Retrieved\ Images|} \tag{14}$$

Since the important contribution of the proposed system is the fuzzy feedback model, which enables the user to make a fuzzy judgment about retrieved images, the proposed system is compared with the previous systems introduced in [21], [22], [24]. As mentioned, these systems are based on long term learning similar to our system but with the binary feedback model. Fig. 10 shows the comparison of the proposed system (Semantic Image Retrieval Based on the FTR) and the previous systems based on High Precision criterion.

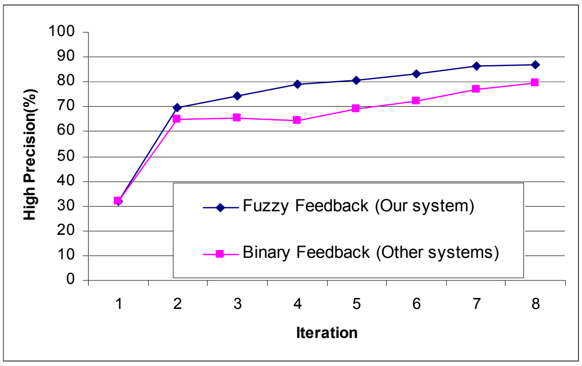

Figure 10: Comparison of the proposed system and other systems based on High Precision

Furthermore, there is another difference between the proposed system and the system introduced by Wacht [23]. Our proposed system is a composite relevance feedback approach for image retrieval using transaction-based and feature re-weighting based learning, whereas learning in [23] is just based on semantic information that is stored in semantic space and the visual similarity is completely ignored. Therefore, the comparison of the proposed system and Wacht system is shown in Fig. 11 based on Normal Precision.

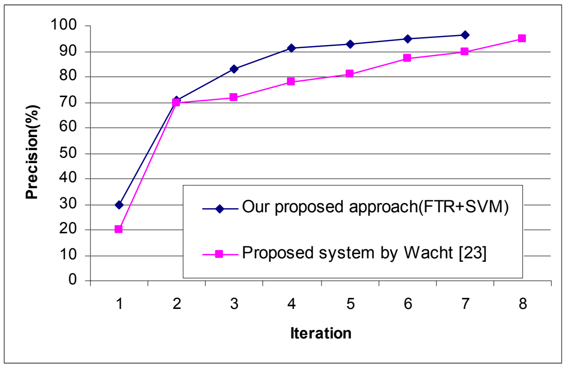

Figure 11: Comparison of the proposed system and Wacht system based on Normal Precision

To evaluate the proposed system, the experimental results can be presented in two stages. First stage is related to evaluation of the short term and the long term learning of the proposed system. In short term learning, the results will be obtained just by using user feedback in current session, whereas long term learning algorithms are based on previous users' behaviors. Finally in the second stage, the retrieval results of the proposed system, first phase and improved one will be compared.

Stage 1: Fig. 12 summarizes the average retrieval precision for the proposed system transaction-based and feature re-weighting (long term and short term learning) and for the feature re-weighting (short term learning) system. This figure shows that the precision of our proposed approach exceeds 94% after seven iterations, whereas, after seven iterations the retrieval precision of the feature re-weighting approach only reaches 67%. Thus, the proposed approach is able to reach the retrieval goal in only a few iterations. This improvement is preferred in image retrieval since the user aims to retrieve the desired images in as few feedback steps as possible.

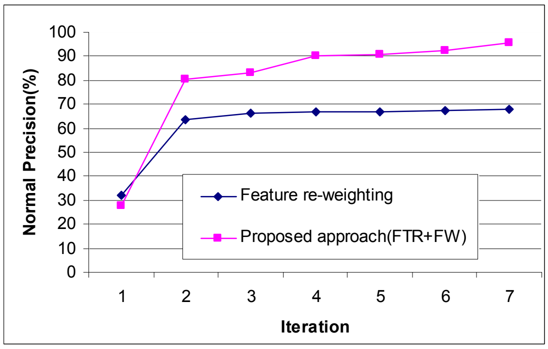

Figure 12: Comparisons of our proposed approach and feature re-weighting approach based on Normal Precision

Experiments further document the retrieval precision increases as the same as the information of the FTR and so that each iteration leads to a better retrieval precision. For this purpose, the experiments are presented in three steps. In the first step, FTR is empty and the average retrieval precision on query images is shown. In the second step, the query session tests will be started on a pre-built FTR from step one. Finally, in the third step we shall test on a pre-built FTR from step one and two. Therefore, Fig. 13 compares the retrieval precision for the proposed approach on an empty FTR, a pre-built FTR (step one) and a pre-built FTR (step one and two) based on

Normal Precision. This clearly shows that a pre-built FTR from step one and two has better precision in the second iteration and achieves a retrieval precision as high as 94% on the seventh iteration.

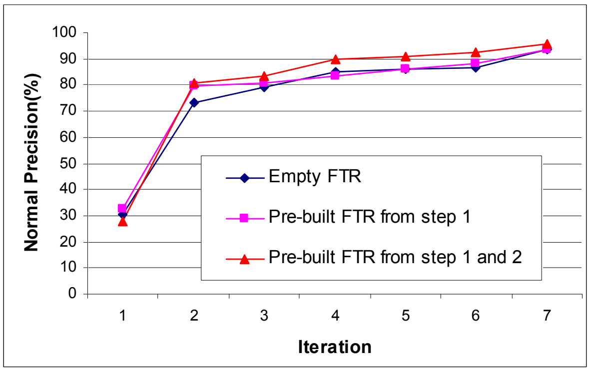

Figure 13: Comparison of our proposed approach on an empty and pre-build repository based on Normal Precision

Stage 2: Finally, the retrieval precision of the proposed system, first phase and improved one are compared. Since decision in the improved system is based on the most users' opinion, the retrieval results of it emphasize on the full relevant images. Therefore, the High Precision criterion is used to compare the two proposed schemes. Fig. 14 demonstrates the average of High Precision of two proposed schemes. As shown in Fig. 14, the average precision of the second scheme is higher than the first scheme in some of the iterations. The reasons for this are:

- The long term information has been stored in the SFTR is effective so that, the full relevant images which will be retrieved are more than the first scheme.
- The long term learning in the improved scheme is possible by updating membership function; because of the type of the long term information has been stored in the SFTR, whereas it is impossible in the first scheme.

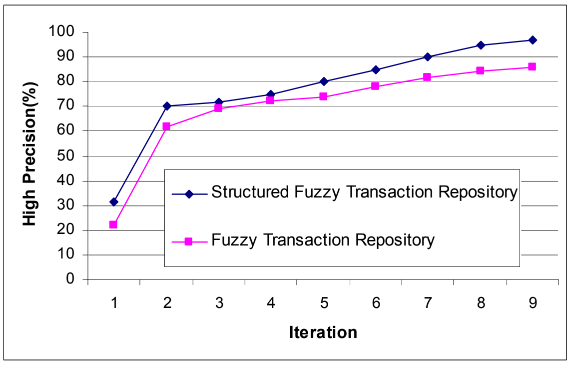

Figure 14: Comparison of two proposed system, first phase and improved phase based on High Precision

Especially in long term, the average retrieval precision of the improved system surpasses that of the first one.

IV. Conclusion

We have proposed an image retrieval system founded on transaction-based semantic learning and low-level features learning. The important contributions of this work can be summarized as follows:

- A fuzzy transaction repository is as dynamically constructed to store user relevance feedback information. An incremental method was developed to deal with new log sessions, through updating the information of the FTR in each session. This is important for the purpose of long term learning.
- Due to the subjectivity of the human annotation, the first scheme was improved and the structure of the FTR was changed. In the new scheme, each element of the SFTR is a data structure. Therefore, in the new scheme, user feedback can be utilized more than the first scheme by storing it accurately.

Experiments show that our proposed systems obtain a desirable performance and achieve remarkably high retrieval precision after the first three iterations.

References

[1] K. Wu, K.H. Yap, "Content based image retrieval using fuzzy perceptual feedback", Multimedia Tools and Applications archive Volume 32 , Issue 3, Pages 235 – 251, March 2007
[2] 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, Jan. 2000
[3] 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
[4] S. Asbaghi, M.R. Keyvanpour, A. Amiri, "Learning-based approach for semantic image retrieval by using a dynamic semantic network", 19th International Conference on Database and Expert Systems Application, IEEE 2008
[5] J. Rocchio, "Relevance feedback in information retrieval", Prentice-Hall, 1971.
[6] 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) 811–820, 2002
[7] S.K. Saha, A.K. Das, B. Chanda, "Image retrieval based on indexing and relevance feedback" , Pattern Recognition Letters 28 ,357–366, 2007
[8] 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
[9] 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
[10] 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, Sep. 2002
[11] Y. Rui, T. Huang, S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS", Proc IEEE Int Conf Image Processing, Washington DC, USA, pp 815–818, 1997
[12] I. Cox, T. Minka, T. Papathomas, P. Yianilos, "The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments", IEEE Trans Image Process, 2000
[13] Z. Su, HJ. Zhang, S. Ma, "relevant feedback using a Bayesian classifier in content-based image retrieval", SPIE Electronic Imaging, San Jose, CA, Jan 2001
[14] J. Laaksonen, M. Koskela, E. Oja, "PicSom-self-organizing image retrieval with MPEG-7 content descriptions", IEEE Trans Neural Netw 13(4):841–853, July 2002
[15] P. Muneesawang, L. Guan, "Automatic machine interactions for content-based image retrieval using a self-organizing tree map architecture", IEEE Trans Neural Netw 13(4):821–834, July 2002
[16] 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
[17] 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), pp. 299–302, 2000
[18] R. Datta, D. Joshi, J. Li, J.Z. Wang, "Image retrieval: ideas, influences, and trends of the new age" , ACM Computing Surveys, Vol. 40, No. 2, Article 5, April 2008

[19] 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
[20] 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
[21] 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, Jan. 2003
[22] X. Qi, R. Chang, "Image retrieval using transaction-based and SVM-based learning in relevance feedback sessions", ICIAR 2007, LNCS 4633, pp. 638–649, 2007
[23] M. Wacht, J. Shan, X. Qi, "A short-term and long-term learning approach for content-based image retrieval," Int. Conf. on Acoustic, Speech, and Signal Processing (ICASSP), 2006
[24] 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
[25] L.A. Zadeh, Fuzzy Sets. Information and Control 8, 338-35, 1965
[26] G. Das, S. Ray, C. Wilson, "Feature re-weighting in content-based image retrieval", 5th International Conference, CIVR 2006 Tempe, AZ, USA, Lecture Notes in Computer Science 4071, July 2006

Malihe Javidi was born in Mashad, Iran, in 1984. She obtained the B.S. and M.S. degrees in computer engineering from Ferdowsi University of Mashad, Iran, in 2006 and 2008, respectively. Her research interests include image and video processing, machine learning, and fuzzy systems. She has some journal and conference publications in subject of interesting area. Email: malihejavidi@gmail.com

Hamid Reza Pourreza was born in Gonabad, Iran, in 1969. He received the B.S. degree in electrical engineering from Ferdowsi University of Mashad in 1989, and then he received to the M.S. and PhD degrees in engineering department from Amirkabir University of Technology, Iran, in 1992 and 2003 respectively. He works in Computer Department as an associate professor at Ferdowsi University of Mashhad. His research interests include machine vision, image and video processing, and signal processing. He has more than 50 journal and conference publications in subject of interesting area. Email: hpourreza@um.ac.ir

Hadi Sadoghi Yazdi was born in Sabzevar, Iran, in 1971. He received the B.S. degree in electrical engineering from Ferdowsi University of Mashad in 1994, and then he received to the M.S. and PhD degrees in electrical engineering from Tarbiat Modarres University of Tehran, Iran, in 1996 and 2005 respectively. He works in Computer Department as an assistant professor at Ferdowsi University of Mashhad. His research interests include adaptive filtering, image and video processing, and optimization in signal processing. He has more than 140 journal and conference publications in subject of interesting area. Email: h-sadoghi@um.ac.ir