Australian Journal of Basic and Applied Sciences, 4(2): 171-186, 2010
ISSN 1991-8178

Learning of Relevance Feedback Using a Novel Kernel Based Neural Network

[1]Hamed Modaghegh, [2]Malihe Javidi, [2]Hadi Sadoghi Yazdi and [2]Hamid Reza Pourreza

[1]Electrical Department Ferdowsi University of Mashhad, Iran
[2]Computer Department Ferdowsi University of Mashhad, Iran

Abstract: In this paper, we introduce a novel neural network and fuzzy transaction based image retrieval system. The proposed system is a composite relevance feedback approach for image retrieval using semantic and visual learning. In semantic learning, the system integrates the log information of user feedback using a fuzzy feedback model to construct fuzzy repository. The repository remembers the user's intent and therefore, provides a better representation of each image in the database. The semantic similarity between the query image and each database image can then be computed using the current feedback and the semantic values in the fuzzy repository. In addition, most of the existing approaches assume a linear relationship between different features, and the usefulness of such systems was limited due to the difficulty in representing high-level concepts using low-level feature. This paper presents a novel neural network which is based on the nonlinear kernel Least Mean Square (KLMS). The proposed approach allows the users to select an initial query image and incrementally search target images via relevance feedback. If users aren't satisfied with the retrieved results, relevance feedback method is used to enhance the performance of the proposed system by updating a boundary for separating relevant images from irrelevant ones. These two similarity measures are normalized and combined together to form the overall similarity measure. Experimental results using a COREL database demonstrate the effectiveness of the proposed method.

Key words: Fuzzy Repository, Relevance Feedback, Semantic Image Retrieval, Neural Network Kernel Least Mean Square

INTRODUCTION

With the rapid development of computing hardware, digital acquisition of information has become one popular method in recent years. Consequently, how to make use of this huge amount of images effectively becomes a highly challenging problem (Han, J.H., D.S. Huang, 2005). 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. Early literature emphasizes fully automated systems and tries to find best features. In such approaches, the best features and representations and their corresponding weights are fixed, which can not effectively model high level concepts and user's perception subjectively (Lee, H.K., S.I. Yoo, 2001).

In order to improve the retrieval accuracy of CBIR systems, the focus of research has been shifted from designing sophisticated low-level feature extraction algorithms to reducing the 'semantic gap' between low-level features and high-level semantic concept (Liua, Y., D. Zhang, 2007). So to reduce the gap, different techniques were introduced such as Relevance Feedback (RF). The relevance feedback approach has been proposed to semi-automatic annotation and image retrieval (Asbaghi, S., M.R. Keyvanpour, 2008). Different approaches, studied in the next subsection, are used to learn the user's feedback.

Related Work of Relevance Feedback:

To bridge the gap between low-level features and semantic meaning, relevance feedback techniques have been incorporated in the process of image retrieval. In (Datta, R., D. Joshi, 2008) *Datta* has presented a short review of recent work in RF. According to (Datta, R., D. Joshi, 2008) the various ways of RF can be categorized to: (a) feedback specification, (b) user-driven, (c) probabilistic, (d) region-based, and (e) learning-based.

Corresponding Author: Malihe Javidi, Computer Department Ferdowsi University of Mashhad, Iran
E-mail: malihejavidi@gmail.com

a. Feedback Specification:

Recent work has introduced other paradigms of query specification that have been found either more effective. Feedback based directly on an image semantics characterized by manually defined image labels, appropriately termed *semantic feedback*, is proposed in Yang *et al*. (2005). A well-known issue with feedback solicitation is that multiple rounds of feedback test the user's patience. To overcome this limitation, another method, generally 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. *Han* (2005) and *Zhuang* (2002) introduce a knowledge memory model to store semantic information by simply accumulating user-provided interactions. In the process of image retrieval, the system converts user feedback to semantic relations between images and constructs a semantic relation network. The semantic-space-based (He, X., Q. King, 2003; Xiaojun. Qi, Ran. Chang, 2007; Wacht, M., Shan, 2006) and the log-based (Hoi, S., M.R. Lyu, 2006) are other systems that integrate the log information of user feedback with RF for image retrieval.

b. User-driven:

There has been some interest in design RF paradigms aimed to help users. In some new developments, there have been attempts at tailoring the search experience by providing the user with cues and hints for more specific query formulation (Jaimes, A., K. Omura, 2004; Nagamine, T., A. Jaimes, 2004).

c. Probabilistic:

Probabilistic approaches have been taken in Cox et al. (2000), Su et al. (2001), and Vasconcelos and Lippman (2000). They have been utilized to incorporate user feedback to update the probability distribution of all images in the database. These models were popular in the early years of image retrieval for tackling the basic problem, have found increasing patronage for performing RF in recent years.

d. Region Based:

The region-based retrieval systems (Sumengen, B., B. Manjunath, 2003) are introduced whereas they attempt to overcome the deficiencies of feature-based image retrieval by representing images at the object-level. In Jing et al. (2004) a region-based image retrieval framework is introduced that integrates efficient region-based representation in terms of storage and retrieval and effective on-line learning capability. In this feedback process, the region importance for each segmented region is learned, for successively better retrieval. This core idea, namely that of integrating region-based retrieval with relevance feedback, has been further detailed for RF in (Jing, F., M. Li, 2004).

e. Learning-based:

Based on the user's relevant feedback, learning-based approaches are typically used to appropriately modify the feature set or similarity measure. Artificial neural networks (Laaksonen, J., M. Koskela, 2002; Muneesawang, P., L. Guan, 2002) have been adopted in the relevance feedback process due to their ability to simulate universal mapping. Recently Machine learning techniques such as SVM are also used for concept learning (Tong, S., E. Chang, 2001). SVM is often utilized to capture the query concept by separating relevant images from irrelevant ones. 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 (Tian, Q., Y. Yu, 2000), the D-EM (Discriminant-Expectation Maximization) is used to solve this problem. We propose a new KLMS-based Neural Network (KNN) algorithm and use it in learning of relevance feedback. The proposed KNN algorithm is based on Kernel Least Mean Square (KLMS) algorithm which was proposed in the (Pokharel, P., W. Liu, 2007; Liu, W., P. Pokharel, 2008) and more details will be mentioned in the next section.

The proposed system integrates user feedback from all iterations in order to construct a fuzzy repository. 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, a new neural network based on KLMS learning is applied to the session-term feedback. The retrieval results are returned by combining the normalized similarity scores computed from both techniques. The dominant feature of the proposed systems includes a new learning algorithm over fuzzy relevance. The proposed neural network is applied after each round of labeling to update a boundary for separating accumulated short-term positively and negatively labeled images.

This paper is organized as follows. Section 2 includes studying of Least Mean Square (LMS) and Kernel LMS (KLMS) algorithm. Section 3 appropriate to explanation of the proposed KLMS-based Neural Network (KNN) and experimental results are illustrated in Section 4. Concluding remarks are offered in section 5.

Least Mean Square (LMS) Algorithm and Kernel LMS (KLMS):

In order to clarify the background of the new algorithm, it is necessary to show how they are interrelated and how they have evolved. Firstly, the LMS algorithm is explained briefly. In 1959 the LMS algorithm was introduced as a simple way of training a linear adaptive system with mean square error minimization. An unknown system -*Y(n)*- is to be identified and the LMS algorithm attempts to adapt the filter $\hat{Y}(n)$:o make it as close as possible to *Y(n)*. The algorithm uses *u(n)* as the input, *d(n)* as desired output and *e(n)* as calculated error. LMS uses steepest-descent algorithm to update the weight vector so that the weight vector converges to optimum Wiener solution. Updating weight vector based on the following rule is applied:

$$w(n+1) = w(n) + 2\mu \times e(n) \times u(n) \tag{1}$$

where *w(n)* is the weight vector of the estimator, *u(n)* is the vector of the input data sequence, μ is the step size, *e(n)* is the estimation error, and *d(n)* is the desired response. The filter output *Y* is calculated by:

$$Y(n) = \overline{w} \times u(n) \tag{2}$$

Successive corrections of the weight vector eventually leads to the minimum value of the mean squared error. You can find further information about LMS in (Widrow, B., 1966).

Normalized LMS Algorithm:

The main problem of LMS algorithm is utilizing a constant step-size, which forces the designer to set a small step-size to avoid instability which leads to low convergence speed. NLMS is an updated version of the LMS algorithm that solves this drawback by normalizing the input vector. So a variable step-size could be used instead of a constant one.

NLMS adjusts the step-size parameter so that $w(n+1)$ minimizes the error, thus convergence speed is improved.

$$w(n+1) = w(n) + \frac{2\mu}{< u(n), u(n) >} e(n) \times u(n) \tag{3}$$

Some of NLMS applications can be find in (Carini, A., S. Malatini, 2008; Laska, B.N.M., R.A. Goubran, 2008; Vega, L.R., 2008).

Kernel LMS Algorithm:

Kernel methods are applied to map the input data into a High Dimensional Space (HDS). In HDS a variety of methods can be used to find linear relations in the data. Mapping procedure is handled by Φ functions (Fig. 1). Kernel functions help the algorithm to handle the converted input data in the HDS ever without knowing the coordinates of data in that space; simply by computing the kernel of input data instead of calculating the inner products between images of all pairs of data in HDS. This method is called the kernel trick. You can find further information about KMs in (Taylor, J.S., N. Cristianini, 2004; Scholkopf, B., A.J. Smola, 2002).

As presented in (Pokharel, P., W. Liu, 2007; Liu, W., P. Pokharel, 2008) estimation and prediction of some time-series could be optimized with a new approach, named KLMS. The basic idea is to perform the linear LMS algorithm given by Eq. (1) in the kernel space.

$$\Omega(n+1) = \Omega(n) + 2\mu \times e(n) \times \Phi(u(n)) \tag{4}$$

where Ω*(n)* is weight vector in the HDS. The estimated output *y(n)* will be calculated by:

$$y(n) = < \Omega(n), \Phi(u(n)) > \tag{5}$$

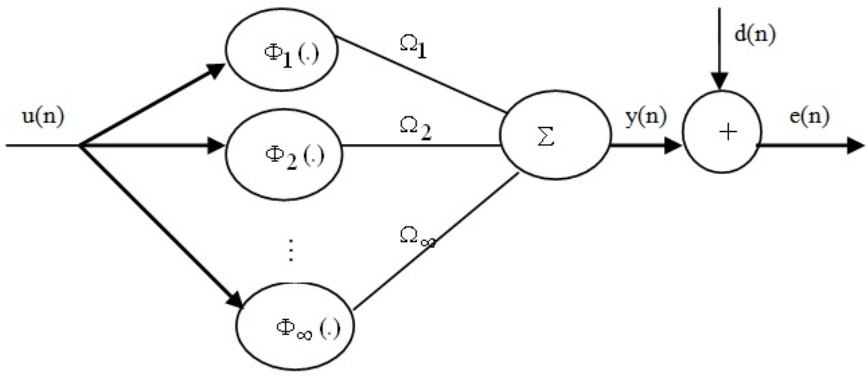

Fig. 1: The block diagram of a simple kernel estimation system

Fig. 1 shows the input vector *u(n)* being transformed to the infinite feature vector $\Phi(u(n))$, whose components are then linearly combined by the infinite dimensional weight vector. Non-recursive type of Eq. (4) can be written as:

$$\Omega(n) = \Omega(0) + 2\mu\sum_{i=0}^{n-1} e(i)\Phi(u(i)) \tag{6}$$

By choosing $\Omega(0) = 0$:

$$\Omega(n) = 2\mu\sum_{i=0}^{n-1} e(i)\Phi(u(i)) \tag{7}$$

Based on Eq. (5) and (7):

$$\begin{aligned} y(n) = < \Omega(n), \Phi(u(n)) > = < 2\mu\sum_{i=0}^{n-1} e(i)\Phi(u(i)), \Phi(u(n)) > \\ = 2\mu\sum_{i=0}^{n-1} e(i) < \Phi(u(i)), \Phi(u(n)) > \end{aligned} \tag{8}$$

We can use kernel trick to calculate *Y(n)*:

$$y(n) = \mu\sum_{i=0}^{n-1} e(i)\, k(u(i), u(n)) \tag{9}$$

Eq. (9) is named Kernel LMS algorithm. As error of system reduces by time, we can ignore the *e(n)* after ξ sample and predict new data with previous error.

$$y(n) = \mu\sum_{i=0}^{\xi} e(i)\, k(u(i), u(n)) \tag{10}$$

This change decreases the complexity of algorithm. After that we can train model system with fewer data and use it for prediction of new data. Good prediction ability in non-linear channels is one of algorithm advantages.

Normalized Kernel LMS Algorithm:

As mentioned in subsection 2.1, NLMS corrects the step-size parameter using Eq. (3). Referring to previous topics, the normalized version of KLMS is now presented. In this method a variable step-size $\mu(n)$ instead of constant one is used. $\mu(n)$ is updated based on input data in HDS, $\Phi(u(n))$. When the $\Phi(u(n))$ value is high, the algorithm decreases the step size and vice versa:

$$\Omega(n+1) = \Omega(n) + 2\mu(n)\, e(n)\, \Phi(u(n)) \tag{11}$$

$$\Omega(n+1) = \Omega(n) + 2\frac{\mu}{\langle \Phi(u(n)), \Phi(u(n)) \rangle} e(n)\, \Phi(u(n))$$

$$\Omega(n) = \Omega(0) + 2\mu \sum_{i=0}^{n-1} e(i) \frac{\Phi(u(i))}{\langle \Phi(u(i)), \Phi(u(i)) \rangle} \tag{12}$$

Whereas $\Omega(0) = 0$, So:

$$\Omega(n) = 2\mu \sum_{i=0}^{n-1} e(i) \frac{\Phi(u(i))}{\langle \Phi(u(i)), \Phi(u(i)) \rangle} \tag{13}$$

Then the Eq. (5) can be rewritten as:

$$y(n) = \langle \Omega(n), \Phi(u(n)) \rangle = \langle 2\mu \sum_{i=0}^{n-1} e(i) \frac{\Phi(u(i))}{\langle \Phi(u(i)), \Phi(u(i)) \rangle}, \Phi(u(n)) \rangle = 2\mu \sum_{i=0}^{n-1} \frac{e(i)}{\langle \Phi(u(i)), \Phi(u(i)) \rangle} \langle \Phi(u(i)), \Phi(u(n)) \rangle \tag{14}$$

With use of kernel trick we have:

$$y(n) = \langle 2\mu \sum_{i=0}^{n-1} \frac{e(i)}{\langle \Phi(u(i)), \Phi(u(i)) \rangle} k(u(i), u(n)) \rangle = 2\mu \sum_{i=0}^{n-1} \frac{e(i)}{k(u(i), u(n)) \rangle} k(u(i), u(n)) \rangle \tag{15}$$

We propose a new neural network based on Normalized KLMS learning algorithm for image retrieval. In the next section, the proposed image retrieval system is introduced with core of KLMS-based neuron.

The Proposed Semantic Image Retrieval System:

The block diagram of the proposed system is shown in Fig. 2. The system first computes the low-level features of the query image and then returns 20 images with the highest similarity scores to the user. 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 fuzzy repository to estimate the query semantic feature. Furthermore, the proposed KLMS-based Neural Network (KNN) uses session-term feedback to learn the KNN for measuring the visual similarity between the query image and each database image. 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.

Image Content Representation:

Two set of features are used to represent the images. The first set contains a 9-dimensional feature vector, including color moments in each color channel (H, S and V). The first order (mean), the second (variance) and the third order (skewness) color moments have been proved to be efficient and effective in representing color distributions of images. Mathematically, the first three moments can be defined as:

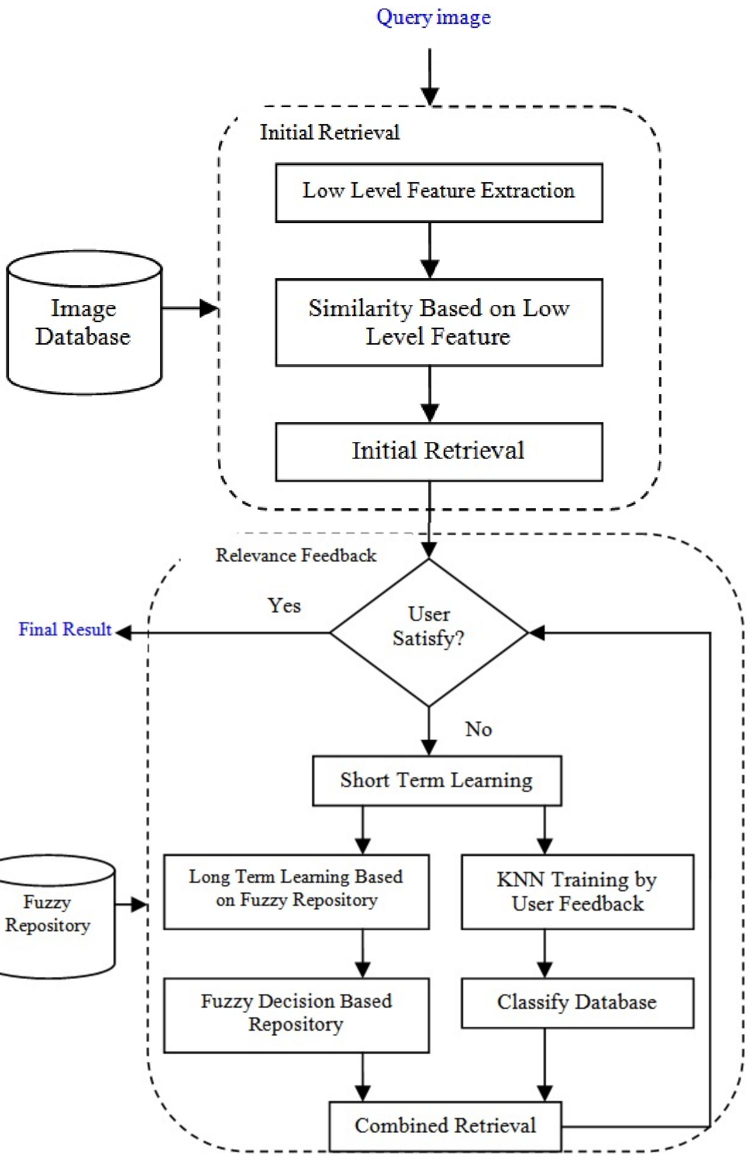

Fig. 2: The block diagram of our proposed system

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij}$$
$$\sigma_i = (\frac{1}{N}\sum_{j=1}^{N} (f_{ij} - \mu_i)^2)^{1/2} \quad (16)$$
$$s_i = (\frac{1}{N}\sum_{j=1}^{N} (f_{ij} - \mu_i)^3)^{1/3}$$

where *fij* is the value of the i^{th} color component of the image pixel *j*, and *N* is the number of pixels in the image.

The second set contains three types of features, color, shape, and texture. The color feature is similar to the first set. 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). 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 noticeable point is that, the variance of feature increases as the same as the feature dimensions and so the average retrieval precision is reduced. To overcome this problem, the feature is normalized, and then it is used to learn the visual similarity.

Semantic Learning Based on Fuzzy Repository:

Fuzzy Feedback:

A novel fuzzy relevance feedback is proposed for user feedback which enables the user to make a fuzzy judgment. The fuzzy labeling 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 fuzzy repository includes fuzzy labels obtained from the relevance feedback procedure. Five types of fuzzy labels are used in our feedback process: Full Irrelevant (FIR), Irrelevant (IR), Don't Care (DC), Full Relevant (FR) and Relevant (R). FIR, IR, FR and R membership functions are defined by a trapezoidal membership function and DC is defined by a triangular membership function. Users judge the relevance degree of the retrieved images. When the image does not retrieve in any iteration, its fuzzy repository 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.

The following linear membership function will be considered for each triangle fuzzy sample (Zadeh, L.A., 1965):

$$\mu_i(y) = \begin{cases} \dfrac{y-(Y_i-d_i)}{d_i} & Y_i - d_i \le y \le Y_i \\ \dfrac{Y_i+d_i-y}{d_i} & Y_i \le y \le Y_i + d_i \\ 0 & y \ge Y_i + d_i, y \le Y_i - d_i \end{cases} \quad (17)$$

where d_i is the tolerance of i^{th} input vector and $d_i \in [0,1]$ which is shown in Fig. 3.

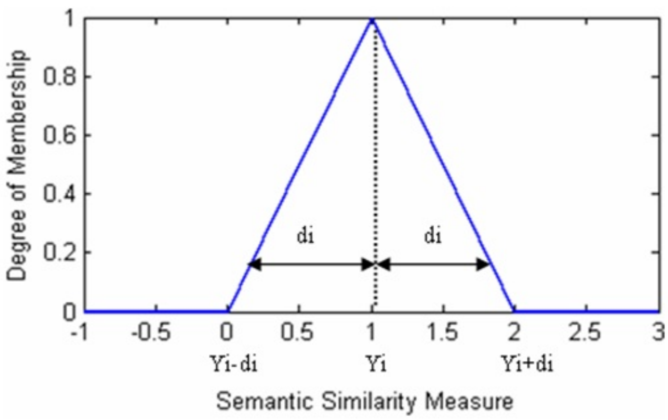

Fig. 3: User feedback to form of triangle according to Eq. (17) with $Y_i = 1$ and $d_i = 1$ Also for

trapezoidal fuzzy samples:

$$\mu_i(y)=\begin{cases}\dfrac{y-(Y_i-d_i)}{d_i} & Y_i-d_i\le y\le Y_i\\ \dfrac{Y_i+b_i+d_i-y}{d_i} & Y_i+b_i\le y\le Y_i+b_i+d_i\\ 0 & y\ge Y_i+b_i+d_i,\ y\le Y_i-d_i\\ 1 & Y_i\le y\le Y_i+b_i\end{cases} \tag{18}$$

where b_i is the width of trapezoidal membership function shown as an example in Fig. 4.

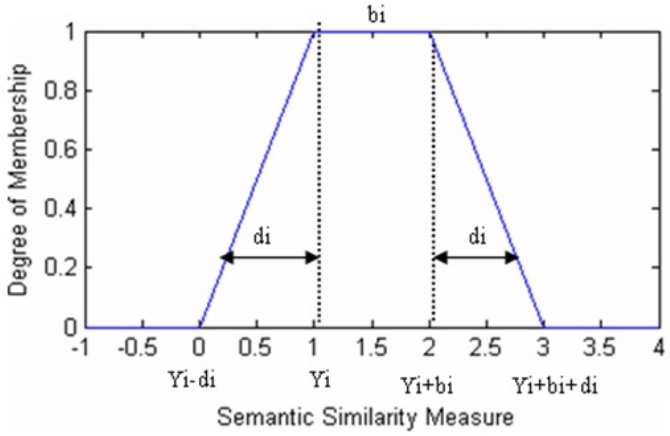

Fig. 4: User feedback to form of trapezoidal according to Eq. (18) with $Y_i=1$, $d_i=1$ and $b_i=1$

Fuzzy Repository Construction:

Fuzzy repository stores user feedback. Each row in the fuzzy repository represents an image in the database and each column corresponds to one semantic group. Initially, fuzzy repository is empty and it is constructed dynamically as follows:

A. For each query image Q

a Append a new column to the fuzzy repository. A new column signifies that a new concept have been added to the database. Moreover if a query is a new image not existed in the database, new row is added to the fuzzy repository.

b Retrieve images using low-level features and return 20 ones which are the most similar to a query Q.

c The relevance feedback mechanism solicits the user to judge the relevance of the retrieved images.

d 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 fuzzy repository have numeric values, but the membership functions, introduced in the proposed system, are used to compute semantic similarity. i.e. the FR and R membership functions are utilized for numeric values 2 and 1 in the repository, but, the FIR and IR are applied for numeric values -2 and -1 respectively.

e Compute the semantic similarity score between a query Q and each database image by using fuzzy transaction-based semantic learning (section 3.2.3).

f Compute the visual similarity score between a query Q and each database image by using KNN learning technique (section 3.3).

g The two similarity measures are combined together to form the overall similarity measure (section 3.4).

h 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 the semantic group. As database images are increased, images that belong to this semantic group are gradually added to the database.

B Finally, the new column is compared with all existing columns in the fuzzy repository 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 fuzzy repository dimensions and search time are reduced.

Semantic Similarity Based on Fuzzy Repository:

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 images. Each semantic vector corresponds to a row vector of fuzzy repository, which is determined by the index number of the labeled images. Then, the high-level feature vector of the query (Q) is initialized as:

$$Q = (q_1, q_2, \ldots, q_n) \tag{19}$$

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

$$q_i = (x_i^1 \vee x_i^2 \vee \ldots \vee x_i^s) \wedge \overline{(x_i^{s+1} \vee x_i^{s+2} \vee \ldots \vee x_i^{s+t})} \tag{20}$$

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 is be treated as 0'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 fuzzy repository, w_i is used for weighting the fuzzy membership function. The result of weighting is:

$$\mu_R(x) = \sum_{i=1}^{n} w_i \mu_i(x) \tag{21}$$

where $\mu_i(x)$ is the fuzzy membership function of the i^{th} element of the semantic feature vector x, w is the weight vector, n equals the number of columns in fuzzy repository and $\mu_R(x)$ is the result of adding the weighted membership function in fuzzy repository. Fig. 5 shows the summation of two weighted membership functions according to the Eq. (21):

Afterwards, results are defuzzified. The Centroid defuzzification is used to defuzzify results.

$$R = \frac{\sum_{(x)} x \mu_R(x)}{\sum_{(x)} \mu_R(x)} \tag{22}$$

For the following feedback iterations, short-term learning refines the query by updating its weight vector using current feedback and the semantic values in fuzzy repository as follows:

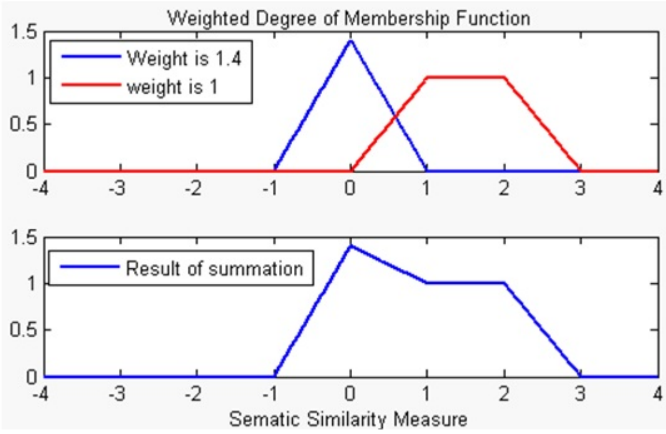

Fig. 5: Example of adding weighted membership functions

Positive Feedback:

$$w_i^{(t+1)} = \begin{cases} \alpha w_i^{(t)}, & if \quad x_i = 1 \quad and \quad w_i^{(t)} \neq 0 \\ 2 * \alpha w_i^{(t)}, & if \quad x_i = 2 \quad and \quad w_i^{(t)} \neq 0 \\ 1, & if \quad (x_i = 1 \quad or \quad 2) \quad and \quad w_i^{(t)} = 0 \\ w_i^{(t)} & if \quad x_i = 0 \\ \dfrac{w_i^{(t)}}{\alpha}. & if \quad x_i = -1 \\ \dfrac{w_i^{(t)}}{2 * \alpha}. & if \quad x_i = -2 \end{cases} \tag{23}$$

Negative Feedback:

$$w_i^{(t+1)} = \begin{cases} 2 * \alpha w_i^{(t)} & if \quad x_i = -2 \quad and \quad w_i^{(t)} \neq 0 \\ \alpha w_i^{(t)} & if \quad x_i = -1 \quad and \quad w_i^{(t)} \neq 0 \\ 1 & if \quad (x_i = -1 \quad or \quad x_i = -2) \quad and \quad w_i^{(t)} = 0 \\ \dfrac{w_i^{(t)}}{2 * \alpha}. & if \quad x_i = 2 \\ \dfrac{w_i^{(t)}}{\alpha}. & if \quad x_i = 1 \\ w_i^{(t)}, & if \quad x_i = 0 \end{cases} \tag{24}$$

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.

The Proposed Kernel-Based Neural Network (KNN):

As mentioned the KLMS algorithm in section 2.2, input data are transferred to HDS then in the created linear space, a linear combiner is used for estimation/prediction task. Of course we see in the KLMS we don't need any weights and output is obtained only by calculation of error. This note is used for KNN construction. In the proposed KNN, at first KLMS is used and a logistic function is added in the end of KLMS structure as shown in the following figure. Some relations are needed for finding output from error term which is discussed here.

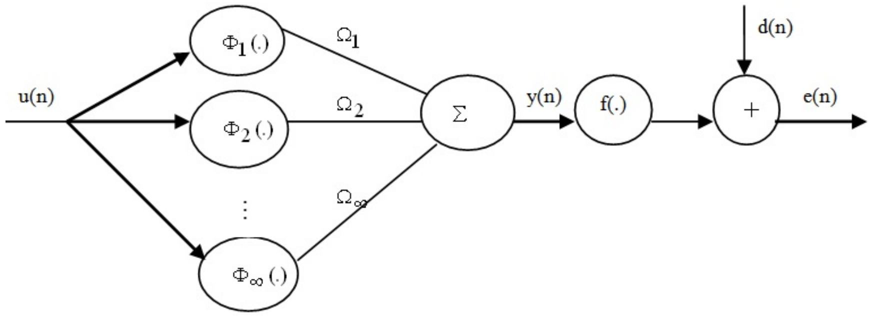

Fig. 6: The block diagram of the proposed KLMS based neural network

The output of KNN is calculated to form of,

$$y(n) = f(<\Omega(n), \Phi(u(n))>) \tag{25}$$

where

$$e(n) = d(n) - y(n) \tag{26}$$

Weight update formula is,

$$\Omega(n+1) = \Omega(n) + \mu\nabla e(n)^2 = \Omega(n) + 2\mu\Phi(u(n)) \times e(n) f'(<\Omega(n), \Phi(u(n))>) \tag{27}$$

Where $E(n)$ is defined as follows,

$$E(n) = e(n) f'(<\Omega(n), \Phi(u(n))>) \tag{28}$$

After substitution,

$$\Omega(n+1) = \Omega(n) + \mu\nabla e(n)^2 \tag{29}$$

$$\Omega(n+1) = \Omega(n) + 2\mu E(n)\Phi(u(n))$$

After simplification of Eq. (29) with substitution of n we have,

$$\Omega(n) = \Omega(0) + 2\mu\sum_{i=0}^{n-1} E_i \Phi(u(i)) \tag{30}$$

With substitution $\Omega(n)$, the Eq. (25) can be defined as:

$$y_n = f(<\Omega(n), \Phi(u(n))>) = f(<2\mu \sum_{i=0}^{n-1} E_i \Phi(u(i)), \Phi(u(n))>) = f(2\mu \sum_{i=0}^{n-1} E_i <\Phi(u(i)), \Phi(u(n))>) \quad (31)$$

We use kernel trick, so the Eq. (31) can be changed as follows:

$$y(n) = f(2\mu \sum_{i=0}^{n-1} E(i) \times K(\Phi(u(i)), \Phi(u(n)))) \quad (32)$$

We name recently relation as KNN algorithm. In the KNN, only $E(i)$ coefficients are needed for finding output and these coefficients are obtained using Eq. (28). In other words, learning of KNN include finding of $E(i)$ coefficients. The kernel method RBF is used in Eq. (32).

In the test procedure for finding results of applying input u_T , $E(i)$ coefficients (obtained in learning phase) and learning samples $u(i)$ are used in the following equation.

$$y(n) = f(u(T)) = f(2\mu \sum_{i=0}^{n-1} E(i) \times K(\Phi(u(i)), \Phi(u(T)))) \quad (33)$$

where u_T is test samples.

The distance of each database image will be computed and normalized to the range of [0, 1] to rank the visual similarity between the query and each database image. If user satisfies, these retrieved images are final result. Otherwise user applies his judgments and run relevance again. By each user feedback, the number of relevant retrieved images is increased and provide more positive sample for better training of the KNN and so, improve precision in the next retrieval results.

Retrieval Results:

After the first retrieval result is achieved, the similarity between the query image and each database image is computed using both normalized KNN ($Score_{low}$) and transaction-based semantic similarity ($Score_{high}$).

$$Score(I, Q) = Score_{high}(I, Q) + Score_{low}(I, Q) \quad (34)$$

where Q and I is the query and each database image.

Experimental Results:

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.

To evaluate CBIR system, two important criterions are utilized: Precision and Recall. Since the proposed system is based on relevance feedback, the Precision criterion is used to evaluate the proposed system.

Fig. 7 summarizes the average retrieval precision for the proposed approach using the first set of feature and the second one. Since the second set of feature involves color, shape and texture, it clearly shows that the precision of the proposed approach has better than the first one.

In order to verify the performance of the proposed system, we also compare the proposed system with other image retrieval systems. The proposed system is a composite relevance feedback approach for image retrieval using transaction-based and new neural network based KLMS learning, whereas learning in (Wacht, 2006) 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 approach and *Wacht* is shown in Fig. 8. It shows that the proposed system is more efficient to retrieve images than the other system.

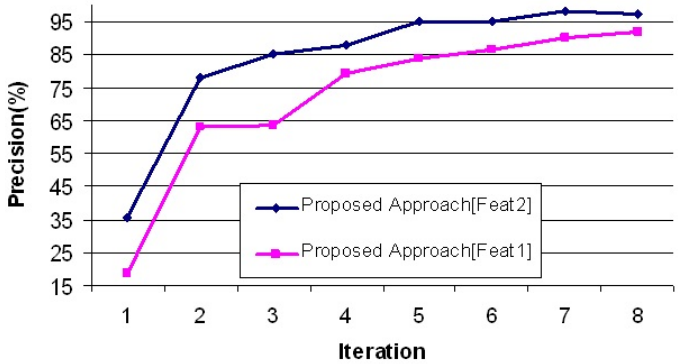

Fig. 7: Comparison of average precision of the proposed approach using different feature

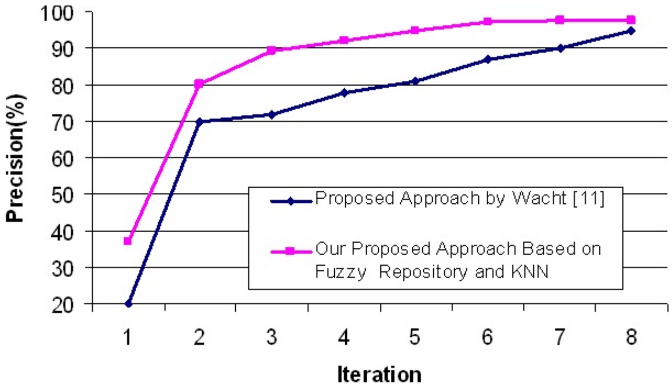

Fig. 8: Comparison of average precision between the proposed system and other system

Recenty machine learning techniques suchas SVM are more used for concept learning. Fig. 9 shows the average retrieval precision of our proposed approach and other systems that used SVM to learn visual similarity. This figure shows that the precision of our proposed approach exceeds 92% after four iteration, whereas, after four iteration the retrieval precision of the SVM based approach only reaches 85%. 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.

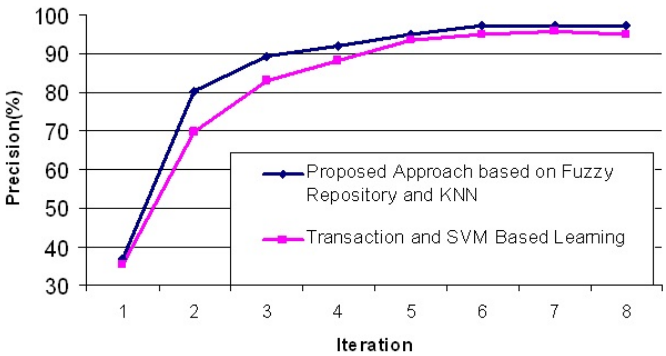

Fig. 9: Comparison of the proposed approach (fuzzy repository and KNN) and other approach (transaction and SVM based)

Experiments further document the retrieval precision increases as the same as the information of the fuzzy repository and so that each iteration leads to a better retrieval precision. For this purpose, Fig. 10 shows the average retrieval precision with an empty repository. Our proposed approach achieves better performance on the 3th iteration and achieves comparable retrieval accuracy for the remaining relevance feedback iterations when compared to the transaction and SVM based learning.

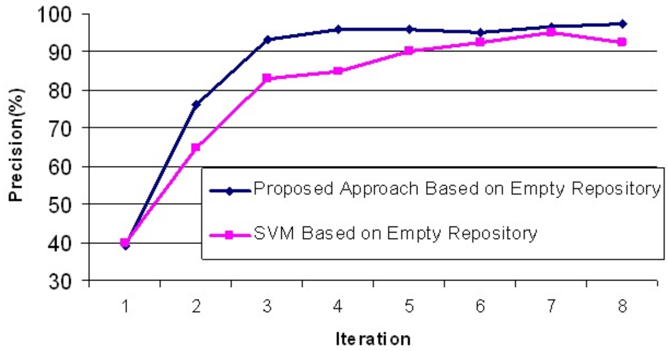

Fig. 10: Comparison of average precision between the proposed approach and other approach based on empty repository

To evaluate the effect of the initial repository, experiment has been tested on the pre-built repository. It means that, 3% of the database images in each category are used as queries to pre-build the repository. It clearly shows that the larger the pre-built repository is the better the average retrieval precision, Fig. 11.

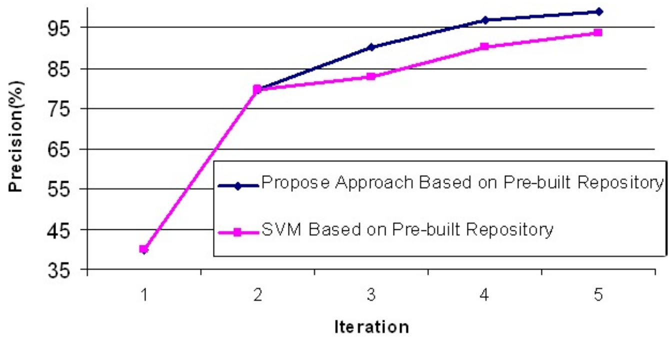

Fig. 11: Comparison of average precision between the proposed approach and other approach based on pre-built repository

Conclusion:

In this paper we presented a novel neural network and fuzzy transaction based image retrieval system. The important contributions of this work can be summarized as follows:

a A fuzzy transaction repository is as dynamically constructed to store user relevance feedback information.
b An incremental method was developed to deal with new log sessions, through updating the information of fuzzy repository in each session. This is important for the purpose of long term learning.
c A fuzzy feedback model used in retrieval sessions allows users to better judge the relevance of the returned results.
d A new neural network based KLMS learning can be applied to finding images which are visually similar to the query image.

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

REFERENCES

Asbaghi, S., M.R. Keyvanpour, A. Amiri, 2008. “Learning-Based Approach for Semantic Image Retrieval By Using A Dynamic Semantic Network”, 19th International Conference on Database and Expert Systems Application, IEEE.

Carini, A., S. Malatini, 2008. “Optimal Variable Step-Size NLMS Algorithms with Auxiliary Noise Power Scheduling for Feedforward Active Noise Control”, IEEE Trans. Audio, Speech, and Language Processing, 16(8): 1383-1395.

Cox, I.J., TP. Minka, T.V. Papathomas, P.N. Yianilos, 2000. “The Bayesian Image Retrieval System, PicHunter: theory, implementation, and psychophysical experiments”, IEEE Trans Image Process.

Datta, R., D. Joshi, J. Li, J.Z. Wang, 2008. “Image Retrieval: Ideas, Influences, and Trends of The New Age” , ACM Computing Surveys, 40(2): 5.

Han, J., K.N. Ngan, M. Li, H.J. Zhang, 2005. “A Memory Learning Framework for Effective Image Retrieval”, IEEE Transactions on Image Processing, 14(4).

Han, J.H., D.S. Huang, T.M. Lok, M.R. Lyu, 2005. “A Novel Image Retrieval System Based on BP Neural Network”, IEEE International Symposium, ISCAS 2005, 2: 1557- 1560.

He, X., Q. King, W.Y. Ma, M. Li, H.J. Zhang, 2003. “ Learning a Semantic Space From User’s Relevance Feedback for Image Retrieval “, IEEE Transactions on Circuits and Systems for Video Technology, 13(1).

Hoi, S., M.R. Lyu, R. Jin, 2006. “A Unified Log-Based Relevance Feedback Scheme for Image Retrieval”, IEEE Transactions on Knowledge and Data Engineering, 18(4).

Jaimes, A., K. Omura, T. Nagamine, K. Hirata, 2004. “Memory Cues for Meeting Video Retrieval”, In Proceedings of the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE) at the ACM International Multimedia Conference.

Jing, F., M. Li, H.J. Zhang, B. Zhang, 2004. “An Efficient and Effective Region-Based Image Retrieval Framework”, IEEE Trans. Image Process, 13(5): 699-709.

Jing, F., M. Li, H.J. Zhang, B. Zhang, 2004. “Relevance Feedback in Region-Based Image Retrieval”, IEEE Trans. Circ. Syst. Video Technol., 14(5): 672-681.

Laaksonen, J., M. Koskela, E. Oja, 2002. “PicSom-Self-Organizing Image Retrieval with MPEG-7 Content Descriptions”, IEEE Trans Neural Netw, 13(4): 841-853.

Laska, B.N.M., R.A. Goubran, M. Bolic, 2008. “Improved Proportionate Subband NLMS for Acoustic Echo Cancellation in Changing Environments”, IEEE Trans. Signal Processing, 15: 337-340.

Lee, H.K., S.I. Yoo, 2001. “Intelligent Image Retrieval Using Neural Network”, IEICE Transaction Information & System, Vol. E84-D, 12.

Liu, W., P. Pokharel, J.C. Principe, 2008. “The Kernel Least-Mean-Square Algorithm”, IEEE Transactions on Signal Processing, 56(2).

Liua, Y., D. Zhang, G. Lua, W.Y. Mab, 2007. “A Survey of Content-Based Image Retrieval with High-level Semantics”, Journal of the Pattern Recognition, pp: 262-282.

Muneesawang, P., L. Guan, 2002. “Automatic Machine Interactions for Content-Based Image Retrieval Using A Self-Organizing Tree Map Architecture”, IEEE Trans Neural Netw, 13(4): 821-834.

Nagamine, T., A. Jaimes, K. Omura, K. Hirata, 2004. “A Visuospatial Memory Cue System for Meeting Video Retrieval”, In Proceedings of the ACM International Conference on Multimedia (demonstration).

Pokharel, P., W. Liu, J.C. Principe, 2007. “Kernel LMS”, ICASSP, 3: 1421-1424.

Scholkopf, B., A.J. Smola, 2002. “Learning with Kernels, Massachusetts Institute of Technology.

Su, Z., H.J. Zhang, S. Ma, 2001. ”Relevant feedback Using A Bayesian Classifier in Content-Based Image Retrieval”, SPIE Electronic Imaging, San Jose, CA.

Sumengen, B., B. Manjunath, S. Kenney, 2003. “Image Segmentation Using Multi-Region Stability and Edge Strength”, Proceedings of IEEE International Conference on Image Processing.

Taylor, J.S., N. Cristianini, 2004. “Kernel Methods for Pattern Analysis”, Cambridge University Press.

Tong, S., E. Chang, 2001. “Support Vector Machine Active Learning for Image Retrieval”; In: Proc. of ACM Int. Conf. on Multimedia, pp: 107-118. ACM Press, New York.

Tian, Q., Y. Yu, T.S. Huang, 2000. “Incorporate Dscriminant Analysis with EM Algorithm in Image Retrieval”, Proceedings of the International Conference on Multimedia and Expo (ICME), pp: 299-302.

Vasconcelos, N., A. Lippman, 2000. “Learning From User Feedback in Image Retrieval Systems”, In Proceedings of the Conference on Advances in Neural Information Processing Systems (NIPS).

Vega, L.R., H. Rey, J. Benesty, S.A. Tressens, 2008. “New Robust Variable Step-Size NLMS Algorithm” IEEE Trans. Signal Processing, 56(5): 1878-18.

Wacht, M., J. Shan, X. Qi, 2006. “A Short-Term and Long-Term Learning Approach for Content-Based Image Retrieval,“ Int. Conf. on Acoustic, Speech, and Signal Processing (ICASSP).

Widrow, B., 1966. “Adaptive Filters I.Fundamentals” (TR 6764-6), Stanford Electronics Laboratories, Stanford, CA, Technical Report.

Xiaojun. Qi, Ran. Chang, 2007. “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.

Yang, C., M. Dong, F. Fotouhi, 2005. “Semantic Feedback for Interactive Image Retrieval”, In Proceedings of the ACM International Conference on Multimedia.

Zadeh, L.A., 1965. Fuzzy Sets. Information and Control, 8: 338-35.

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