

# Observations on Using Type-2 Fuzzy Logic for Reducing Semantic Gap in Content-Based Image Retrieval System

Saad M. Darwish and Raad A. Ali

**Abstract**—Semantic-based image retrieval has been one of the most challenging problems in recent years. Although so many solutions are provided for filling the so-called gap between the content based image retrieval (CBIR) and what human beings expect from the retrieval task; none of them yields satisfactory results and the problem is still open for further research. In this paper, type-2 fuzzy logic (T2FL) framework is considered to alleviate two problems in traditional CBIR systems, including the semantic gap and the perception subjectivity. Employing T2FL has the potential to overcome the limitations of type-1 fuzzy logic and produce a new generation of fuzzy controllers with improved performance for many CBIR applications that require handling high levels of uncertainty. Thus, our contributions in this study are threefold. (1) The proposed system maps low-level visual statistical features to high-level semantic concepts; enabling to retrieve and browse image collections by their high-level semantic concepts. (2) Type2 fuzzy logic has been used to fuse (combine) extracted features as well as to deal with the ambiguity of human judgment of image similarity. (3) The system models the human perception subjectivity with the ability to handle high levels of uncertainties appropriately. A comparative study with the state-of-the-art type-1 fuzzy based image retrieval approaches reveals the effectiveness of the proposed system.

**Index Terms**—Type-2 fuzzy logic, semantic-based image retrieval, soft computing, image processing.

## I. INTRODUCTION

More and more digitized images are gathered and prolonged in our daily life and techniques need to be established on how to organize and retrieval them. Content-based image retrieval (CBIR) refers to the retrieval of images from a database that are similar to a query image, using measures of information derived from the images themselves, rather than relying on the accompanying text or annotation [1]. CBIR depends on the absolute distance of image features like colors, textures, shapes, or the spatial layout of target [2]. Potential uses for CBIR include: architectural and engineering design, art collections, crime prevention, geographical information and remote sensing systems, intellectual property, medical diagnosis, photograph archives, and retail catalogs.

Unfortunately, there are still many problems that hinder CBIR systems from being popular [1], [3]. Two of these problems are: (1) Semantic gap: Users may prefer using

high-level textual concepts, such as words, to interpret an image. However, most of early CBIR systems provide low-level numerical features to represent an image. The semantic gap problem is the lack of coincidence between the image representation and the human interpretation for an image. (2) Perception subjectivity: Different users, or even the same user under different circumstances, may interpret an image differently. In most cases, users have a vague concept of what it looks like.

Recently several approaches have been proposed that attempt to tackle the above problems and particularly bridge the semantic gap. One can classify these solutions into four categories [1], [3], [4]: 1) using classification and clustering techniques to define high-level concepts 2) inducing rational knowledge in the process of image description to reduce the retrieval errors; 3) introducing relevance feedback (RF) into retrieval loop for continuous learning of users intention; 4) making use of both the visual content of images and the textual information obtained from the Web image retrieval. Researches have revealed that there exists a nonlinear relation between the high-level semantics and the low-level features [3].

However, none of the current approaches has achieved satisfactory performance because it has been difficult to infer semantic meaning from low-level features [1]. It has been widely recognized that the family of image retrieval techniques should become an integration of both low-level visual features addressing the more detailed perceptual characteristics and high-level semantic features underlying the more general conceptual aspects of visual data [3], [5], [6]. Although efforts have been devoted to combine these two aspects of visual data, the gap between them is still a huge barrier in front of researchers.

For the sake of alleviating the semantic gap and the perception subjectivity problems, a CBIR system should well characterize a mapping from image features to human concepts, and effectively capture the user's preference on retrieval [6], [7]. In the literature, it is known that fuzzy logic can provide a flexible and vague mapping from low-level numerical features to high-level human concepts [8], [9]. Fuzzy logic deals with reasoning that is approximate rather than fixed and exact. So, it has the ability to deal with the vagueness and ambiguity of human judgment of image similarity. Fuzzy logic has been extensively used at various stages of image retrieval [10]-[13] such as fuzzy region groupings or fuzzy color histogram within the images as a feature extraction technique; for measuring the similarity between the target image and the images in the database (e.g. fuzzy hamming distance).

Following this recent development, this paper presents an improved approach for semantic based image retrieval using type-2 fuzzy logic (T2FL) controller. The proposed system is

Manuscript received February 15, 2014; revised April 21, 2014.

S. M. Darwish is with the Department of Information Technology, Institute of Graduate Studies and Research, Alexandria University, 163 Horreya Avenue, El Shatby 21526, P.O. Box 832, Alexandria, Egypt. (tel: (+203)4295007; e-mail: Saad.darwish@alex-igsr.edu.eg).

Raad A. Ali is with the Department of Computer, Ministry of Education, Iraq. (e-mail: Raad.ali885@yahoo.com).

on top of the previously published type-1 fuzzy-based image retrieval, which is extended via a type-2 membership function representation to handle high level of uncertainties associated with semantic description of the image. In general, type-2 fuzzy systems have outperformed their type-1 counterparts in many challenging real world applications [14]. This is due to the type-2 fuzzy systems ability to handle the high levels of uncertainty as a result of having additional degrees of freedom provided by the footprint of uncertainty (FOU).

The reminder of this paper is organized as follows: Section II provides state-of-the-art semantically CBIR systems. Section III presents our solution for the problem of semantic-based image retrieval and the details of the proposed type-2 fuzzy modeling. The results are discussed using appropriate charts and tables in Section IV. Finally, Section V concludes the paper.

## II. LITERATURE SURVEY

Although many sophisticated algorithms have been designed to describe low level features, these algorithms cannot adequately model image semantics and have many limitations when dealing with broad content image databases [1], [4], [5]. 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.

The state-of-the-art techniques in reducing the semantic gap can be classified into two categories [4], [7]: human-based and machine based methods. For the human-based methods, the system designers have to provide the system with some prior knowledge, such as certain rules to generate semantics from the low-level visual features or direct semantics description manually labeled through certain technology such as ontology and relevance feedback. In this category, it is very expensive and subjective to go through manual annotation with large databases. On the contrary, the machine-based methods automatically extract semantics description by machine learning algorithms such as supervised / unsupervised learning methods or semantic template. Details about employing machine learning in semantic based image retrieval can be found in [6], [7].

Fuzzy modeling has some benefits comparing to other non-fuzzy approaches mentioned above. It does not classify images and hence does not suffer from mistakes arisen from misclassifying query images. Moreover, it does not require learning semantic rules as is done in the first category in human-based methods, but the rules are extracted automatically [2], [8], [9]. There are some related studies that incorporate fuzzy logic into CBIR systems. Some authors proposed a methodology for semantic indexing and retrieval of images based on techniques of image segmentation and fuzzy reasoning. In their algorithm region classification process is employed to assign semantic labels using a confidence degree and concurrently merges regions based on their semantic similarity. This information comprises the assertion component of fuzzy knowledge base that is stored in a semantic repository permitting image retrieval and ranking [1]. Other researchers proposed a fuzzy relevance feedback approach that enables the user to make a fuzzy

judgment. Their system integrates the user's fuzzy interpretation of visual content into the notion of relevance feedback. An efficient learning approach has been offered using fuzzy radial basis function network. The network is constructed based on the user's feedback.

The authors in [4] tried to model the operation of an expert based on the input-output data he/she uses in retrieval of the images using fuzzy rules created by an automatic process. The main idea behind their solution relies on assigning different weights for image regions and features during the retrieval process. Their system learns these weights from the expert's user input-output data. In [8], the scholars defined a query description language to unify the query expression of textual descriptions, visual examples and relevance feedback. They suggested unsupervised fuzzy clustering algorithm that builds a mapping from low-level image features (Tamura feature) to high-level concepts (linguistic terms) in a fully automated way.

The study suggested by M. Ionescu *et al.* [15] presented initial results on a new approach to measure similarity between images using the notion of Fuzzy Hamming Distance (FHD) and its use to CBIR. The main advantage of the FHD is that the extent to which two different images are considered indeed different can be tuned to become more context dependent and to capture (implicit) semantic image information. The study in [16] focused on color semantics and proposed an approach to extract the fuzzy color semantics. According to human color perception model, the authors utilize the linguistic variable to describe the image color semantics, so it becomes possible to depict the image in linguistic expression such as mostly red. Furthermore, they applied the feed forward neural network to model the vagueness of human color perception and to extract the fuzzy semantic feature vector. Another related work in [2] where the authors embedded fuzzy logic into CBIR to deal with the vagueness and ambiguity of human judgment of image similarity. In this case, fuzzy inference is used to construct the weights assignment among various image features. A review of the recent fuzzy-based CBIR methods can be found in [8], [10].

All of the above mentioned approaches assume that fuzzy sets and their corresponding membership functions are given beforehand. In other words, they rely on an expert to manually specify the fuzzy sets. Furthermore, utilizing type-1 fuzzy sets do not have the ability to deal efficiently with the uncertainties associated with description of image features required to reduce semantic gap. From this point, the focus of this paper is to introduce type-2 fuzzy based semantic gap reduction strategy, which ensures that image retrieval system can be used in an efficient manner when dealing with different image with different semantics. The recommended system can be thought of as a compromise exploiting the efficiency of fuzzy logic while avoiding its disadvantage that can divert away from correct solution in existing fuzzy logic CBIR systems.

Our system employs type-2 fuzzy logic instead of traditional fuzzy logic for describing low-level image features similarity in order to build a global semantic model. This employment makes the system has a high degree for semantic interpretation of image features to reduce the semantic gap. In general type-2 fuzzy sets have grades of

membership that themselves fuzzy, it can be used to convey the uncertainties in membership function of type-1 sets, due to the dependence of the membership functions on available linguistic and numerical information [17]. The advised type-2 fuzzy-based CBIR is simple, intuitive, and effective in image retrieval.

### III. PROPOSED SYSTEM

This paper imports the type-2 fuzzy logic into semantic image retrieval to deal with the vagueness and ambiguity of human judgment of image similarity. Our retrieval system has the following properties: firstly adopting the fuzzy language variables to describe the similarity degree of image features, not the features themselves so as to infer the image similarity as human thinking; secondly expressing the subjectivity of human perceptions by fuzzy rules impliedly to diminish the semantic gap. In the introduced system, the final unique distance between any two images is computed by using type-2 fuzzy fusion of the individual feature's similarity. This system can achieve better precisions, since it can model the operation of human expert.

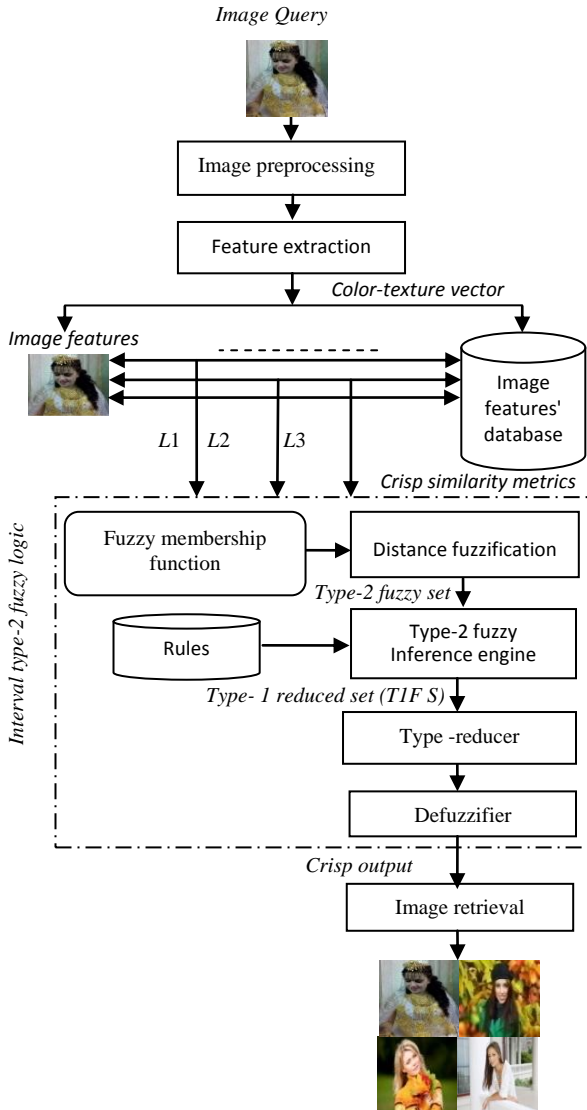


Fig. 1. Overall structure of the proposed type-2 fuzzy image retrieval system.

The overall structure of the proposed type-2 fuzzy semantic retrieval system is depicted in Fig. 1. When an

image is given to the system, it first extracts its predefined features. The resultant distance vector between the image features is used as an input to the fuzzy inference module which creates fuzzy fusion control,  $\mathcal{F}$ , that is used for image retrieval.

$$\mathcal{F} = \bigoplus_{D=1}^3 D_{qt} = \bigoplus_{D=1}^3 |f_i^q - f_i^t|^p \quad (1)$$

$f_i^q$  and  $f_i^t$  denote the  $i^{\text{th}}$  feature element of query and target images respectively,  $p$  denotes the power in distance measure  $D_{qt}$  and is set to 1 in our experiments, and  $|\cdot|$  measures the absolute value of its operand. The following subsections describe the details of the system.

#### A. Preprocessing

Pre-processing becomes necessary when we have images that are corrupted by some kind of distortion such as noise, bad illumination, and blurred. Noise is unwanted information that can result from the image acquisition process. In this work, median filter is often used to remove noise, pixel dropouts and other spurious features of single pixel level while preserving overall image quality [18]. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise [19]. With median filtering, the value of an output pixel is determined by the median of the neighborhood pixels, rather than the mean (average). The median is much less sensitive than the mean to extreme values. Median filtering is therefore better able to remove this outlier without reducing the sharpness of the image.

#### B. Feature Extraction

One of the most important issues in an image retrieval system is the feature extraction process, where the visual content of the images is mapped into a new space, the feature space. Basically, the key to get a successful retrieval system is to choose the right features that represent the images as “strong” as possible [1]. Feature representation of images may include color, texture, and shape. It is very difficult to achieve satisfactory retrieval results by using only one of these feature's categories. So, many techniques adopt methods where more than one feature types are involved so that retrieval performance can be improved [3], [7], [10].

Regarding CBIR, image features can be either extracted from the entire image or from regions. As it has been found that users are usually more interested in specific regions rather than the entire image; most current CBIR systems are region-based [18]. Representation of images at region level is proved to be more close to human perception system. In this paper, we focus on global features to reduce the calculations resulting from the segmentation of regions of interest.

Because the raw data of the image is always too difficult to use and understand, the image color and texture visual features are used as the universe of discourse in the proposed semantic retrieval system. Given a color space  $C$ , the conventional histogram  $H$  of image  $I$  is defined as follow [8]:

$$H_C(I) = \{N(I_i, C_i) | i \in [1, \dots, n]\} \quad (2)$$

where  $N(I_i, C_i)$  is the number of pixels of  $I$  fall into cell  $C_i$  in the color space  $C$  and  $i$  is the gray level.  $H_C(I)$  shows the

proportion of pixels of each color within the image. For representing color, we use HSV (Hue, Saturation, and Value) color model because this model is closely related to human visual perception [10]. Hue is used to distinguish colors and to determine the redness or greenness of the light. Saturation is the measure of percentage of white light that is added to a pure color. Value refers to the perceived light intensity. Color quantization is useful for reducing the calculation cost. Furthermore, it provides better performance for semantic rule because it can eliminate the detailed color components that can be considered as noises [1]. The human visual system is more sensitive to hue than saturation and value so that hue should be quantized finer than saturation and value. In the experiments, we non-uniformly quantized HSV space into 10 bins for hue, three bins for saturation and three bins for value for lower resolution. So each image has 10+3+3(=16) dimensional color vector.

Furthermore, we use Gabor feature, Tamura texture and Gray Level Co- occurrence Matrix (GLCM) to represent the texture feature [3], [7]. As for each image, we filter it in 6 directions and 4 scales, then the mean and standard deviation of the filtering results are computed to form the 48 dimensional Gabor texture feature vector. As for the Tamura texture, it includes coarseness, contrast, directionality, linearity, regularity and waviness of the texture, in a total of a seven dimensional feature. The GLCM feature is also extracted in our work, which is to describe the occurrence frequency of one pair of gray level  $i$  and  $j$ , with a distance  $d$  ( $d=1$ ) in direction  $\theta$  ( $0^\circ, 45^\circ, 90^\circ, 135^\circ$ ). We pick up angular second moment entropy, contrast, homogeneity and correlation, in total 5 different attributes of the texture features. Finally, for each image we obtain a 20-dimension texture feature vector from the 4 GLCM in 4 different directions. So each image has 48+7+20(=75) dimensional texture vector. Readers looking for mathematical background regarding both color and texture features extraction can refer to [3], [7], [10].

### C. Similarity Measure

Similarity measurement plays a vital role in CBIR system. After the process of feature extraction has been carried out on an image database, the stored image feature content must be compared in terms of similarity taking into account either color and texture features. Basically, this procedure is a clustering process where the query image is compared with the stored images and the most similar images are obtained. For this purpose, a computation of a feature-based distance similarity can be obtained for every image in the database that tells us how similar the query image is to each one of the stored images. There are several similarity-checking techniques, but the most usually used are Euclidean distance ( $L_1$ ), city block ( $L_2$ ), and correlation metric ( $L_3$ ). In this research, type-2 fuzzy logic-based similarity is used to quantify the matching between two images.

The rationale of choosing these measures is that they are simple, faster and suitable in most of the cases, and especially when we have a large number of images. Furthermore, correlation distance considers not only the distance between two points but also their relation to the origin (i.e. more precise measure). Given an  $m$ -by- $n$  data matrix  $I$  that represents image, which is treated as  $m$  (1-by- $n$ ) row vectors

$(x_1, x_2, \dots, x_m)$ , the various distances between two images  $I_s$  and  $I_t$  are defined as follows[18]:

$$L_1 = (I_s - I_t)(I_s - I_t)^T \quad (3)$$

$$L_2 = \sum_{j=1}^n |I_{sj} - I_{tj}| \quad (4)$$

$$L_3 = 1 - \frac{(I_s - \bar{I}_s)(I_t - \bar{I}_t)}{\sqrt{(I_s - \bar{I}_s)(I_s - \bar{I}_s)} \sqrt{(I_t - \bar{I}_t)(I_t - \bar{I}_t)}} \quad (5)$$

$\bar{I}$  symbols the matrix mean. To satisfy the requirement of the membership grade function used in the proposed fuzzy system, distance measures for feature vectors need to be transformed into range [0, 1] with Gaussian normalization method as stated in [8]. In our case, the fusion method is based on matching score in which the distances generated by the individual feature are averaged using type-2 fuzzy set and the result is then used for image retrieval.

### D. Interval Type 2 Fuzzy Semantic Mapping

This approach is to make CBIR systems intelligent so that they can interpret human thinking, which not depend on the rigid distance metrics to decide the image similarity. Fuzzy logic(FL) is a powerful tool to realize this goal. FL is probably the most efficient and flexible method available for image retrieval to manage the combination of measurements through their degrees of uncertainty [10]. FL is a theory that allows the natural descriptions, in linguistic terms, of problems to be solved rather than using numerical values. In this research, the attempt is to model the uncertainty of similarity degree of image features through a type-2 fuzzy model.

Similar to a type-1 FL, a type -2 FL includes fuzzifier, rule based, fuzzy inference engine, and output processor. The output processor includes type-reducer and defuzzifier; it generates a type-1 fuzzy set output (from the type-reduce) or a crisp number (from the defuzzifier) [14]. A type-2 FL is again characterized by IF-THEN rules, but its antecedent or consequent sets are now type-2. These fuzzy sets are characterized by their FOU, which in turn are characterized by their boundaries-upper and lower membership functions (MFs) [20]. Type -2 FL can be used when the circumstances are too uncertain to determine exact membership grades such as when training data have the problem of vagueness in human features perception in image retrieval- a case studied in this paper.

In general, type-2 FL is computationally intensive because type-reduction is very intensive. Things simplify a lot when secondary MFs are interval sets (in this case, the secondary memberships are either zero or one and we call them interval type-2 sets (IT2FS)) and this is the case employed in this paper. Formally an IT2FS  $\tilde{A}$  is characterized as [21]:

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x \subseteq [0,1]} 1/(x, u) = \int_{x \in X} \left[ \int_{u \in J_x \subseteq [0,1]} 1/u \right] / x \quad (6)$$

where  $x$ , the primary variable, has domain  $X$ ;  $u$ , the secondary variable, has domain  $J_x$  at each  $x \in X$ ;  $J_x$  is called the primary membership of  $x$ ; and the secondary grades of  $\tilde{A}$  are equal one. Uncertainty about  $\tilde{A}$  is conveyed by the union of all of the primary memberships, which is called FOU of  $\tilde{A}$  i.e.

$$\text{FOU}(\tilde{A}) = U_{x \in X} J_x \quad (7)$$

$$\bar{\mu}_{\tilde{A}}(x) \equiv \overline{\text{FOU}(\tilde{A})} \forall x \in X \quad (8)$$

$$\underline{\mu}_{\tilde{A}}(x) \equiv \underline{\text{FOU}(\tilde{A})} \forall x \in X \quad (9)$$

$\bar{\mu}_{\tilde{A}}(x)$  and  $\underline{\mu}_{\tilde{A}}(x)$  represent upper membership function (UMF) and lower membership function (LMF) respectively. For continuous universes of discourse  $X$  and  $U$ , an embedded interval T2FS  $\tilde{A}_e$ , is defined as [22]:

$$\tilde{A}_e = \int_{x \in X} [1/\theta] x \quad \theta \in J_x \subseteq U = [0,1] \quad (10)$$

$$A_e = \int_{x \in X} \theta/x \quad \theta \in J_x \subseteq U = [0,1] \quad (11)$$

Set  $\tilde{A}_e$  is embedded in  $\tilde{A}$  such that at each  $x$  it only has one secondary variable, and there are an uncountable number of embedded interval T2FS. Set  $A_e$  that acts as the domain for  $\tilde{A}_e$ , is the union of all the primary memberships of the set  $\tilde{A}_e$ . Herein, we focus on an important sub-class of an IT2FS, namely the symmetric IT2FS, for which the FOU is symmetrical about  $x=m$ , i.e. [22].

$$\bar{\mu}_{\tilde{A}}(m+x) = \bar{\mu}_{\tilde{A}}(m-x) \quad (12)$$

$$\underline{\mu}_{\tilde{A}}(m+x) = \underline{\mu}_{\tilde{A}}(m-x) \quad (13)$$

Next, we will fuzzily the input and output of the system. Suppose query image is  $Q$  and image from the database is  $I$ . The similarity distance  $L_1$ ,  $L_2$ , and  $L_3$  between image  $Q$  and  $I$  are three input of the image retrieval system. Three fuzzy variables including "very similar", "similar", "not similar" are used to describe the feature differences and the output of the system (similarity of images). By such description, we can infer the similar of images in the same way as what human think and let us characterize linguistic uncertainty. In this work, we utilize trapezoidal MF to describe all variables that is defined as [22]:

$$\underline{\mu}_{\tilde{A}}(x) = \begin{cases} (x+a)/(a-c), & \text{if } -a \leq x \leq -c \\ 1, & \text{if } -c \leq x \leq c \\ (a-x)/(a-c), & \text{if } c \leq x \leq -a \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

$$\bar{\mu}_{\tilde{A}}(x) = \begin{cases} (x+b)/(b-d), & \text{if } -b \leq x \leq -d \\ 1, & \text{if } -d \leq x \leq d \\ (b-x)/(b-d), & \text{if } d \leq x \leq -b \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

where  $0 \leq c \leq a$  and  $0 \leq d \leq b$ . All of MF' parameters are numerically specified based on the experiences to retrieve images so that if the feature difference of two images is no more than 20%, the two images are very similar, between 30%-50% similar, between 70%-90% or above not similar.

Once the system acquires the fuzzy descriptions of the features distance, the rule base (fuzzy reasoning) can be built to make inference of their similarity. Fuzzy reasoning, which is formulated by group of fuzzy IF-THEN rules, presents a degree of presence or absence of association or interaction between the elements of two or more sets [17]. These rules can be made explicitly by the expert itself as is the case in our

system, or implicitly from the input-output data gathered from system operation. Here, Mandeni's fuzzy interface method is employed; it expects the output membership function to be fuzzy set. In the proposed system, reasoning is carried out through the rules shown in Table I. The nine rules altogether deal with the weight assignments impliedly in the same way as what humans think. The fuzzy inference processes all of the three cases in a parallel manner, which makes the decision more reasonable. For example, if different objects have same features variance scopes, the rule base will be similar and hence our proposed system has a good robustness to images categories (semantics).

TABLE I: FUZZY RULES OF THE PROPOSED RETRIEVAL SYSTEM

Rule	Input			Output
	$L_1$	$L_2$	$L_3$	Similarity Degree
1	Vary Similar	Vary Similar	Vary Similar	Vary Similar
2	Vary Similar	Vary Similar	Similar	Vary Similar
3	Vary Similar	Vary Similar	Not Similar	Similar
4	Similar	Similar	Similar	Similar
5	Similar	Similar	Vary Similar	Similar
6	Similar	Similar	Not Similar	Similar
7	Not Similar	Not Similar	Not Similar	Not Similar
8	Not Similar	Not Similar	Vary Similar	Similar
9	Not Similar	Not Similar	Similar	Not Similar

The outputs of fuzzy values are then defuzzified to generate a crisp value for the output variable. In type-2 fuzzy system, this will be carried in two steps. Type-reduction, the first step of output processing that computes the centroid of an IT2FS. There exist many kinds of type-reduction techniques, for our case, we used center of sets (COS) type reduction,  $Y_{COS}$ , which is expressed as[21], [22]:

$$Y_{COS}(x) = [c_l(\tilde{A})c_r(\tilde{A})] \quad (16)$$

$$c_l(\tilde{A}) = \frac{\int_{-\infty}^{c_l} x \bar{\mu}(x) dx + \int_{c_l}^{\infty} x \underline{\mu}(x) dx}{\int_{-\infty}^{c_l} \bar{\mu}(x) dx + \int_{c_l}^{\infty} \underline{\mu}(x) dx} = \frac{\sum_{i=1}^M \mu_l^i c_l^i}{\sum_{i=1}^M \mu_l^i} \quad (17)$$

$$c_r(\tilde{A}) = \frac{\int_{-\infty}^{c_r} x \underline{\mu}(x) dx + \int_{c_r}^{\infty} x \bar{\mu}(x) dx}{\int_{-\infty}^{c_r} \underline{\mu}(x) dx + \int_{c_r}^{\infty} \bar{\mu}(x) dx} = \frac{\sum_{i=1}^M \mu_r^i c_r^i}{\sum_{i=1}^M \mu_r^i} \quad (18)$$

$M$  presents the number of rules. From the type-reducer we obtain an interval set,  $Y_{COS}$ , to defuzzify it we use the average of  $c_l(\tilde{A})$  and  $c_r(\tilde{A})$  so the defuzzified output of an interval singleton type-2 FLS is calculated as [17].

$$y = \frac{c_l(\tilde{A}) + c_r(\tilde{A})}{2} \quad (19)$$

Finally, semantic image retrieval is performed using nearest neighbor classification. For a query sample  $Q$ , all the other samples  $i \neq Q$  are ordered in a sorted hit list with increasing distance to the query  $Q$  using similarity degree. Ideally the first ranked image should be the query image. If one considers, not only the nearest neighbor (top 1), but also a longer list of neighbors starting with the first and up to a chosen rank (e.g. top 5), the chance of finding the correct hit increases with the list size.

#### IV. EXPERIMENTAL RESULTS

This section will demonstrate all relevant experimental results that were conducted using the proposed type-2 fuzzy system for semantic-based image retrieval. In our experiments 2000 images containing 100 different distinct semantic categories from Corel database [23] and Internet are selected. They contain a wide variety of images, like the images of flowers, mountains, Africa, and so on. Most of the images are in color JPG format. The results are analyzed using precision and recall measures. Suppose  $R(Q)$  is the set of images relevant for the query  $Q$  and  $A(Q)$  is the set of retrieved images. The precision of the result is the fraction of retrieved images that are truly relevant to the query, while the recall is the fraction of relevant images that are actually retrieved [11]:

$$precision = \frac{|A(Q) \cap R(Q)|}{|A(Q)|} \quad (20)$$

$$recall = \frac{|A(Q) \cap R(Q)|}{|R(Q)|} \quad (21)$$

To facilitate the evaluation process, our system automatically selects query images randomly and performs the retrieval process. A retrieved image is automatically classified as relevant if it is in the same semantic category as the query. As feature vector, we used 16-dimensional HSV color vector and 75-dimensional texture vector. Therefore, a feature vector of 91 elements is made for each image.

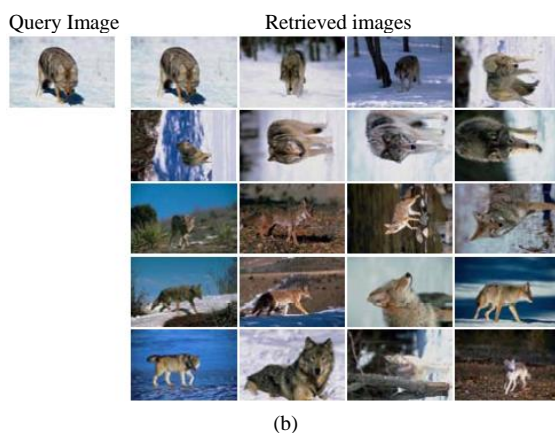
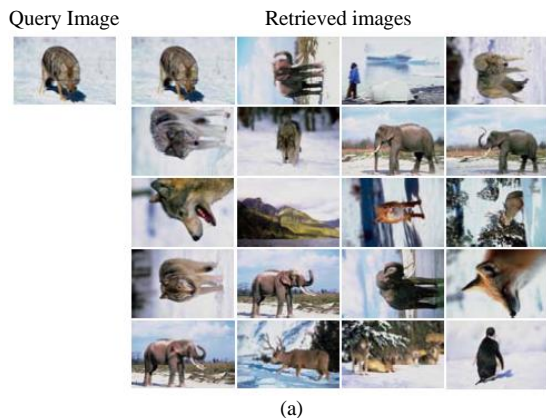


Fig. 2. Retrieval results for a sample query image using different methods; (a) Type-1 fuzzy CBIR; (b) type-2 Fuzzy modeling.

Fig. 2 and Fig. 3 show the retrieval results for two different sample query images using type-1 fuzzy CBIR and the proposed type-2 fuzzy modeling method. In the type-1 CBIR

method, we used the same feature vectors and the  $L_1$ ,  $L_2$ , and  $L_3$  distance measures to determine the similarity between the query and database images. In these figures, the first 20 images retrieved for each query image are shown for comparison purpose. The reduction of semantic gap in the proposed solution is clear in the retrieved images. In Figs. 2(a) and 3(a) the misunderstandings of the images contents from the semantic point of view cause the retrieval of irrelevant images. But, in Figs. 2(b) and 3(b) all first 20 images retrieved are relevant, although the retrieved images are dissimilar from the color and texture features. This shows that the proposed method can enhance the retrieval precision.

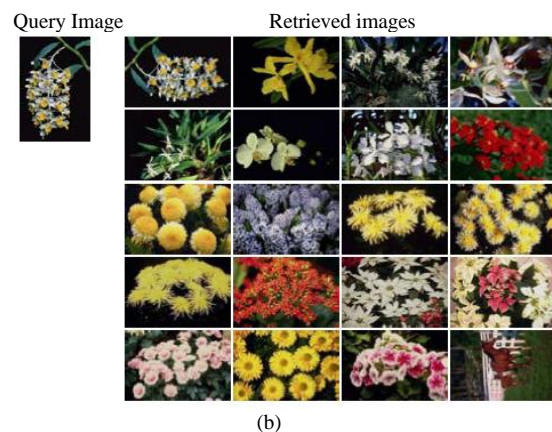
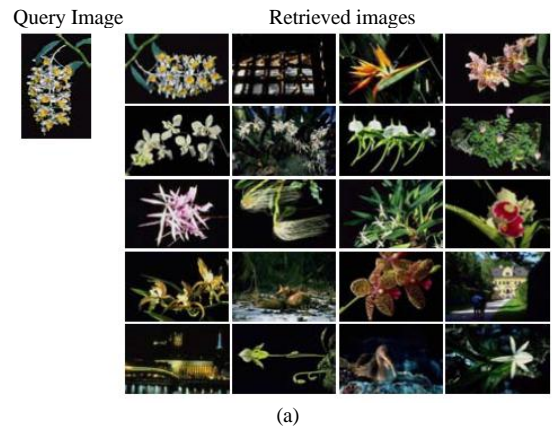


Fig. 3. Retrieval results for a sample query image using different methods; (a) Type-1 fuzzy CBIR; (b) type-2 Fuzzy modeling.

Fig. 4 shows the precision and recall plots for some selected query images that are chosen at random. In these plots, the dotted line indicates the precision-recall values for the type-1 fuzzy CBIR method suggested by W. Xiaoling *et al.* [2] in which the color and shape distances between images are two inputs of the fuzzy inference engine. The dashed line is for the proposed system in which the similarity distance  $L_1$ ,  $L_2$ , and  $L_3$  between two image feature vectors that include color and texture features are three inputs to the type-2 fuzzy inference engine. The experiment is organized as follows. First we submit some selected image in our database as a query to retrieve its relevant images. To construct the graph, we repeat the above process by gradually increasing the number of the retrieved images from 5, 10, 15, 20 to 30. For each number of the retrieve images, we calculate the average precision and recall.

As can be seen, the proposed type-2 fuzzy system improves the precision-recall performance with respect to type-1 fuzzy CBIR for all input queries and all chosen rank.

This phenomenon is due to the fact that using more similar distance function results in feature spaces with type-2 membership function representation to handle high level of uncertainties associated with semantic description of the image and hence, most query image lies in its relevant cluster in the space. For the images with large appearance variety such as Africa, flower and animals etc., our system has an average precision of above 85.3 % vs. the top 30 images, which means that our system has a good robustness to the image categories.

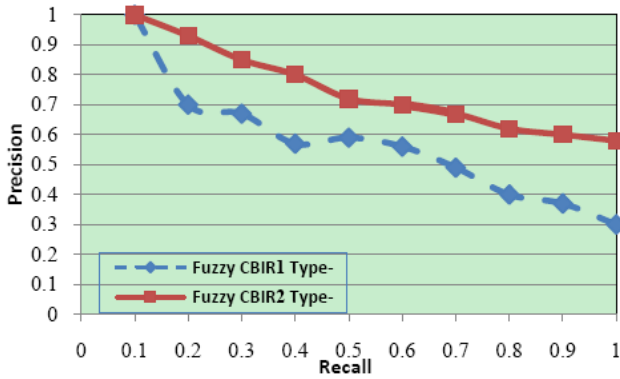


Fig. 4. The average recall-precision plots.

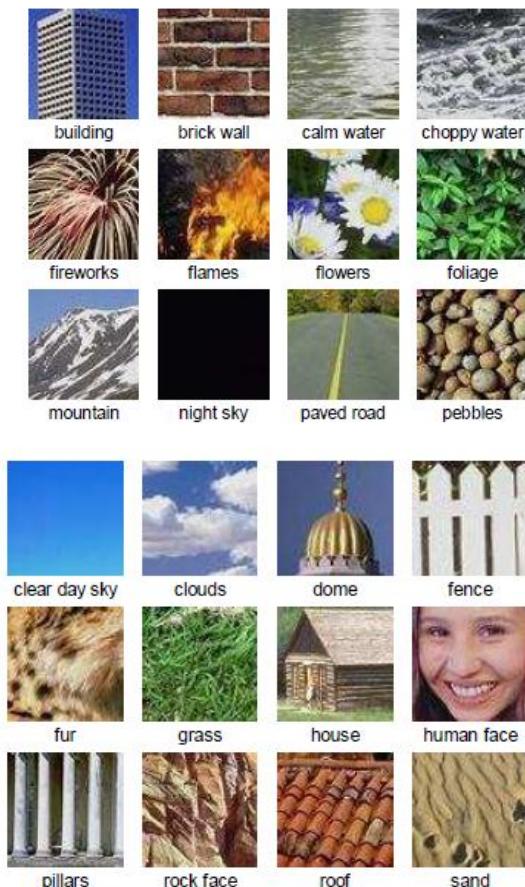


Fig. 5. Sample images of some used semantic categories.

In experiment 3, the classification accuracy of the proposed retrieval system was investigated through the confusion matrix, a matrix used to summarize the results of a supervised classification, as shown in Table II. Entries along the main diagonal are correct classifications. Entries other than those on the main diagonal are classification errors. Fig. 5 shows a variety of some semantic classes that were

identified. The overall recognition accuracy was approximately 95.6% at the top 10 images. As we can see from Table II, water is misclassified to sky with a high rate 5.7%. We have to say that water and sky have the most similar features. We can also find that foliage gets the lower recognition rate of 89.5%. Due to the highly similar features between foliage vs. both grass and rocks, foliage is classified incorrectly as grass with a rate of 3.9 % and rocks of 4.2%. In our experiment, there are some foliage images that contain different color echo, which cannot easily be classified.

TABLE II: CONFUSION MATRIX (%) OF THE PROPOSED RETRIEVAL SYSTEM AT TOP 10 IMAGES

	Foliage	Grass	Rocks	Sky	Water
Foliage	89.5	3.9	4.2	0	2.4
Grass	3.9	96.1	0	0	0
Rocks	4.2	0	95.8	0	0
Sky	0.8	0	0	94.3	4.9
Water	1.6	0	0	5.7	92.7

Finally, our system performs queries in real time and uses less space to save learned knowledge concerning semantic gap. It requires  $O(N^3)$  space used in the memory semantic learning technique, where  $N$  is the total number of images in database.

## V. CONCLUSION

This paper has established and investigated an enhanced fuzzy system for reducing the semantic gap in the CBIR using type-2 fuzzy similarity distance. By predicting the semantic correlation based on the analysis of several similarity measures between low-level features, retrieval effectiveness can be significantly improved. The important contributions to the CBIR field in this work can be summarized as follows: (1) we construct type-2 fuzzy repository to pool the similarity of the extracted features as well as to deal with the ambiguity of human judgment of image similarity. (2) Since the system is dynamically constructed, it is possible to add new images with new concepts to existing image database.

The performance of the system was evaluated on real world images of different categories and promising research results were obtained and described in this paper. We observe that type-2 fuzzy based retrieval excels the traditional fuzzy based one in both precision and recall. The average precision gap is 0.5 % and recall gap is 3.2% at the top 10 images. Therefore, type-2 fuzzy based system can improve the query performance of the CBIR. The system archives remarkably highly retrieval precision at the list of neighbors starting with the first and up to a chosen rank.

Not that our aim is to show that using type-2 fuzzy modeling can improve the performance of CBIR systems and reduce the semantic gap. However, to achieve perfect results, one should use more sophisticated features and similarity measures. In the future, we will test the proposed system on a large database, where each image may have multiple semantic meanings, for its effectiveness and scalability. The semantic clustering technique will be further studied to determine the relations of images to assist in the extraction of semantic rules automatically.

REFERENCES

- [1] Y. Liu, D. Zhang, G. Lu, and W. Ying, "A survey of content-based image retrieval with high-level semantic," *Pattern Recognition*, vol. 40, no. 1, pp. 262-282, 2007.
- [2] W.-X. Ling and X.-K. Lin, "Application of the fuzzy in content-based image retrieval," *Int. Journal of Computer Science*, vol. 5, no. 1, pp. 19-24, 2005.
- [3] N. Idrissi, J. Martine, and D. Aboutajdine, "Bridging the semantic gap for texture-based image retrieval and navigation," *Journal of Multimedia*, vol. 4, no. 5, pp. 277-283, 2009.
- [4] A. Lakdashti, M. Moin, and K. Badie, "Reducing the semantic gap of the MRI image retrieval system using a fuzzy rule based technique," *International Journal of System*, vol. 11, no. 4, pp. 232-249, 2009.
- [5] Q. Li, S. Luo, and Z. Shi, "Semantic-based art image retrieval using linguistic variable," in *Proc. Int. Conf. on Fuzzy System and Knowledge Discovery*, China, 2007, pp. 406-410.
- [6] F. Rajam and S. Valli, "SRBIR: semantic region based retrieval by extracting the dominant region and semantic learning," *Journal of Computer Science*, vol. 7, no. 3, pp. 400-408, 2011.
- [7] L. Yin *et al.*, "Local semantic classification of natural image based on spatial context," in *Proc. Int. Conf. on Image and Graphics*, China, 2011, pp. 482-487.
- [8] C. Chiu, H. Lin, and S. Yang, "A fuzzy logic CBIR system," in *Proc. Int. Conf. on Fuzzy Systems*, China, 2003, pp. 1171-1176.
- [9] A. Lakdashti, M. Moin, and K. Badie, "IRTF: Image retrieval through fuzzy modeling," in *Proc. Int. Conf. on Communication*, China, 2008, pp. 490-494.
- [10] Q. Li and Z. Shi, "Image retrieval based on fuzzy color semantic," in *Proc. Int. Conf. on Fuzzy Systems*, London, 2007, pp. 1-5.
- [11] S. Kulkarni, "Natural language based fuzzy queries and fuzzy mapping of feature database for image retrieval," *Int. Journal of Information Technology and Application*, vol. 41, no. 1, pp. 11-20, 2010.
- [12] N. Prasath, N. Lakshmi, M. Nathiya, N. Bharathan, and P. Neetha, "A survey on the application of fuzzy logic in medical diagnosis," *Int. Journal of Scientific & Engineering Research*, vol. 4, no. 4, pp. 1199-1203, 2013.
- [13] H. Ajorloo and A. Lakdashti, "IRFUM: image retrieval via fuzzy modeling," *Int. Journal Computing and Informatics*, vol. 30, no. 5, pp. 913-941, 2011.
- [14] X. Ye, L. Fu, and Y. Zhang, "Type-2 fuzzy logic system and level set," in *Proc. Int. Conf. on Semantic, Knowledge and Grid*, USA, 2007, pp. 80-85.
- [15] M. Ionescu and A. Ralescu, "Fuzzy hamming distance in a content-based image retrieval system," in *Proc. IEEE Int. Conf. on Fuzzy Systems*, USA, 2004, pp. 1721-1726.
- [16] S.-Z. Ping, L.-S. Wei *et al.*, "Image retrieval based on fuzzy color semantics," in *Proc. IEEE Int. Conf. on Fuzzy Systems*, UK, 2007, pp. 1-5.
- [17] J. M. Mendel and R. I. John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy SYSTEMS*, vol. 10, no. 2, pp. 117-127, Apr. 2002.
- [18] H. Min and Y.-S. Yuan, "Overview of content-based image retrieval with high-level semantic," in *Proc. Int. Conf. on Advanced Computer Theory and Engineering*, China, 2010, pp. 312-316.
- [19] J. Church and S. Rice, "A spatial median filter for noise removal in digital images," in *Proc. IEEE Southeast Conference*, USA, 2008, pp. 618-623.
- [20] J. Mendel and H. Wu, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: part 1 forward problems," *IEEE Trans. on Fuzzy Systems*, vol. 14, no. 6, pp. 718-792, Dec. 2006.
- [21] J. Castro, O. Castro, and L. Martinez, "Interval type-2 fuzzy logic toolbox," in *Proc. Int. Conf. on Fuzzy Systems*, Italy, 2007, pp. 1-6.
- [22] Q. Liang and J. Mendel, "Interval type-2 fuzzy logic systems: theory and design," *IEEE Trans. on Fuzzy Systems*, vol. 8, no. 5, pp. 535-550, Oct. 2000.
- [23] Corel image dataset. [Online]. Available: <http://www.corel.com>



**Saad M. Darwish** received his Ph.D. degree from the Alexandria University, Egypt. His research work concentrates on the field of image processing, optimization techniques, security technologies, computer vision, pattern recognition and machine learning. Dr. Saad is the author of more than 40 articles in peer-reviewed international journals and conferences and served as TPC of many international conferences. Since Feb. 2012, he has been an associate professor in the Department of Information Technology, Institute of Graduate Studies and Research, Egypt.



**Raad A. Ali** received the B.Sc. degree in computer science from the College of Science, University of Kirkuk, Iraq in 2008. Currently he is a M.Sc. student in the Department of Information Technology, Institute of Graduate Studies and Research, Alexandria University, Egypt. His research and professional interests include image processing, fuzzy systems and machine learning.