

Image retrieval based on fuzzy ontology

Madiha Liaqat¹  · Sharifullah Khan¹ ·
Muhammad Majid²

Received: 30 September 2016 / Revised: 3 May 2017 / Accepted: 3 May 2017 /

Published online: 6 June 2017

© Springer Science+Business Media New York 2017

Abstract Rapid increase in digital images demands effective and efficient image retrieval systems. In text based image retrieval, images are annotated with keywords based on human perception. A user query is composed of keywords according to his/her requirements. Query keywords are matched with the keywords associated with images, for retrieval. This process has been extended with ontology to resolve semantic heterogeneities. However, crisp annotation and retrieval processes could not produce the desired results because both processes involve human perception. To overcome this problem, we have proposed a retrieval system that makes use of fuzzy ontology for improving retrieval performance. For modeling the semantic description of an image, it is divided into regions in our dataset and then regions are classified into concepts. The concepts are combined into categories. The concepts, categories and images are linked among themselves with fuzzy values in ontology. The retrieved results are ranked based on the relevancy between the keywords of a query and images. For evaluating the performance of the proposed methodology, we have used both the objective and subjective measures. Experimental results show that the proposed system performs better than the existing systems in terms of retrieval performance.

Keywords Image retrieval · Text based image retrieval · Fuzzy ontology · Objective evaluation · Subjective evaluation

1 Introduction

Digital images are gaining importance nowadays in different domains, such as medical, education, astronomy, fashion and security [12, 14]. Everyday huge amount of images are

✉ Madiha Liaqat
12phdmliaqat@seecs.edu.pk

¹ School of Electrical Engineering and Computer Science, National University of Sciences & Technology (NUST), Islamabad, Pakistan

² Department of Computer Engineering, University of Engineering & Technology (UET), Taxila, Pakistan

generated by military or civilian equipment that need to be organized for efficient and accurate retrieval [3]. Image retrieval is a science of finding images that fulfill a user specified need [11]. Image retrieval process typically involves two steps: annotation (aka. Indexing) and retrieval. In text based retrieval systems, images are annotated with keywords (i.e., textual descriptors) in a natural language based on human perception [28]. A user specifies his/her requirements through a query comprising keywords. For retrieval, keywords in a query are matched with the keywords associated with images [9]. Text based retrieval computes the relevancy on the basis of lexical matching of keywords. However, it does not consider the meaning of the keywords. It is very difficult for computers to automatically retrieve images with the intended meaning of the associated keywords [30]. This is why the retrieval process has been extended with ontology to resolve the problem of semantic heterogeneity [7, 20, 30, 31].

Ontology is an explicit specification of the terms in a domain and their relations among them [9]. It provides an easy and feasible way of capturing a shared understanding of terms that can be used by humans and computers to exchange information [30]. Ontology based systems, such as OLYBIA and OntoPic have been proposed in [20, 25]. In [20], visual as well as animal ontologies have been built to reduce the semantic gap whereas in [25] better object recognition has been provided using landscape ontology.

In existing image retrieval systems, annotation of images with keywords is binary, i.e., a keyword is either associated with an image or not. However, both the annotation and retrieval processes involve human perception that is mostly approximate or uncertain [5, 37]. Figure 1 shows an image of beautiful view of a sea side where we can see some pink flowers in the bottom left side.

A system analyst might annotate this image with keywords like water, mountain, sand, grass and flowers. If a user is searching for flower images and an image retrieval system retrieves this image as a first result then the user will be surprised to see this result. Since images are annotated using binary model and the search result is a crisp set of images where all images are equally relevant to the given query. We believe that images cannot be precisely represented with keywords using binary model of annotation. Therefore, existing systems could not produce the desired results. The objective of this research is to



Fig. 1 Sea-side view annotated by system analyst

consider the relative importance of a particular keyword in both the annotation and retrieval processes because the importance of the keyword varies from user to user. Relevant retrieved images are essential for the satisfaction of users.

To achieve this objective, we have proposed a fuzzy ontology-based system in this paper. The proposed system makes use of fuzzy ontology to improve the retrieval performance. Images are represented with concepts and categories. In order to annotate an image with all the possible concepts, it is divided into regions in our dataset. Regions are then classified into concepts by adopting the technique proposed in [34]. A concept describes an object that the image contains. The frequency of occurrence of the concepts inside an image determines a category which depicts a scene. This categorization enables the semantic comparison of scenes and also helps in search space reduction while querying for specific concepts inside a category. Concepts, categories and images are linked among themselves with fuzzy values in the ontology. By adding a value for degree of membership to each concept and category in an image, the retrieved images from ontology based search reflect the likely information need. For mapping the query terms and ontology concepts, fuzzy search mechanism is applied that searches and ranks the retrieved results based on the degree of relevancy between the keywords of a query and images. The main contribution of this research is two-fold: (i) a new image retrieval system using fuzzy ontology has been proposed to enhance the retrieval performance and (ii) the proposed system has been subjectively evaluated to ensure its effectiveness.

The remaining paper is organized as follows: Section 2 describes the related work. Section 3 describes the proposed system methodology of image retrieval system. Section 4 contains results and discussion. The paper has been concluded in Section 5.

2 Related work

Image retrieval systems are either content based or text based. In content based image retrieval systems (CBIR), low level features are extracted automatically and images are retrieved based on the features like color, shape and texture [14]. But there is a gap between what image features a system can recognize and what human perceives from the image. The focus of this research is on text based image retrieval systems. Therefore, related work is further categorized as: text based, ontology based and fuzzy ontology based retrieval systems.

2.1 Text based retrieval systems

In text based image retrieval systems, images are annotated with keywords. Image retrieval is based on matching the keywords associated with images with the user specified keywords [28]. Keyword based system proposed in [24] has been built for qualitative spatial relationships like “before and after” or “more and less”. The system has been evaluated using “psychophysical evaluation” [32]. In [16], text based image retrieval system has been combined with content based model for efficient search. Text based search was applied first and then content based filtering was applied on the resulting set. Precision and recall measures were used for system evaluation.

2.2 Ontology based retrieval systems

In [30], an ontology based system has been proposed for exhibition system. The proposed system has been compared with text based approach using objective evaluation measures, such as precision and recall. In [13], authors built the natural scenes ontology to reduce the gap between low level features, such as color, texture, shape and high level semantics. Precision and recall metrics were used to measure the system performance. Keywords and ontology based image retrieval systems have been compared in [35]. Result shows that the ontology based system performed better as compared to the keyword based system in terms of precision. In [25], a supervised learning system, OntoPic, has been proposed that allows the semantic search. It has used DARPA agent markup language and ontology inference layer (DAML+OIL) for domain knowledge but the system performance was not mentioned. In [35], semantic based image retrieval system has been proposed. A domain specific (i.e., flower family) low level feature based ontology has been created. These low level features were considered as a data property in web ontology language (OWL). Users can specify a query in the form of text or image. Features were extracted from a query image and matched with the corpus images through ontology and matched images were shown to the users. Semantic image representation model, containing local and global categorization of scenes, has been proposed in [33].

Ontology based image annotation (OLYBIA) system has been proposed in [20]. Low level features, such as color, shape and texture were extracted to build the visual ontology. Inference engine was used to extract high level concepts, such as “Eagle”, “Cheetah” using visual and animal ontologies and inference rules. The experimental results have not been compared with any other model. In [26], image annotation and retrieval through ontology have been discussed. Ontology was constructed for animal domain. Although the system showed benefits of using ontologies but the burden of manual annotation was still there. In [1], an ontology based image retrieval system has been proposed that utilizes visual features and semantic features. The proposed model has been evaluated using precision and recall.

2.3 Fuzzy ontology based retrieval systems

In image retrieval systems, fuzzy based models have been explored for object recognition [2, 4, 25, 27]. For example, if an object is recognized as *sky* with a value of 0.99, then that means it is 99% sure that the object is *sky*. It does not mean that the image contains 99% *sky*. In retrieval systems, users are not only concerned with object recognition but also want the maximum portion of the object in retrieved images. This has been done in different document search engines using fuzzy ontology.

A document search using fuzzy set theory has been described in [23]. The model considered the importance of keywords in search and their relevancy score between the query and the documents. Highly relevant documents were retrieved based on fuzzy set operations and shown to user. In the Ogawa model [19], a keyword connection matrix has been proposed for computing the relevance of the document with user keywords. In addition, users can enter compound queries containing operators such as and, or, not. In the Horng model [8], a multi-relationship fuzzy concept network has been proposed that shows the fuzzy relations between the concepts and their relevance degree with the documents. An information retrieval model

based on ontology encoded with fuzzy relations has been proposed in [21]. When a user enters a query, composed of concepts, the system performs query expansion and adds new concepts based on ontology knowledge. After expansion, similarity between query and documents is calculated by fuzzy operations. The authors have compared their proposed model with Ogawa ([8, 19] models. Results show that the model proposed in (Pereira, Ricarte & Gomide 2006) gives better retrieval accuracy as compared to Ogawa and Horng models. The above mentioned fuzzy based systems were tested for text documents retrieval.

3 Proposed methodology of the system

In this paper, a fuzzy ontology based image retrieval system has been proposed that uses annotated images as an input. Images were annotated with concepts and categories as shown in Fig. 2 by adopting the technique followed in [34]. An input image was divided into 10×10 grids. Features, such as color and texture, were extracted from each region.

Each region was annotated with one of the concept, such as sky, mountain and water. Each image was assigned a category based on concept occurrence in the image. An overview of the proposed image retrieval system is shown in Fig. 3. Fuzzy knowledge base and fuzzy search mechanism are two main modules of the proposed system. An image along with associated concepts and categories is the input to the fuzzy knowledge base. To conceptually represent the images, fuzzy ontology that utilizes the concepts and categories associated with the images was constructed. The fuzzy values in the ontology were then computed by applying data mining approaches on input images. For image retrieval, users were provided with an interface, where they can input multiple keywords based on their requirements. Fuzzy search mechanism was applied in the proposed system and the retrieved images were ranked and shown to user based on the relevancy degree between an image and the query keywords.

In next subsections, fuzzy knowledge base is discussed in detail that shows step by step construction of fuzzy ontology. The image retrieval algorithm is discussed. This shows how query is processed; and in the end a walk-through example is presented.

3.1 Fuzzy ontology construction

The fuzzy ontology in the proposed model was constructed by adopting the idea of [22], which is used for documents retrieval. The fuzzy ontology shows the relationship between the images

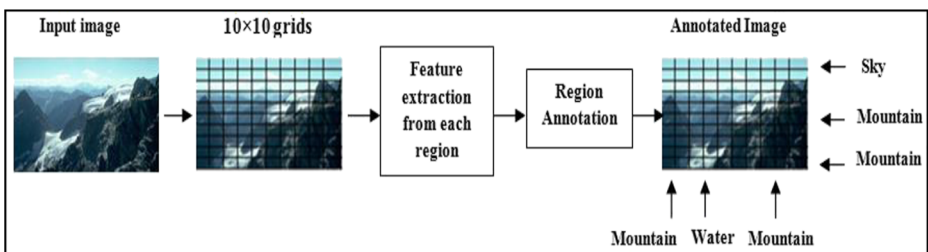


Fig. 2 Image Annotation with concepts using [33]

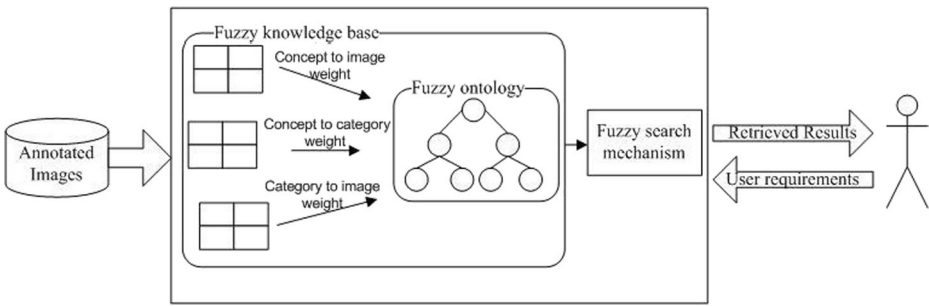


Fig. 3 The proposed image retrieval system

and concepts, concepts and categories and categories and images by values between 0 and 1 (i.e., both 0 and 1 are inclusive). The steps followed for computing the fuzzy values in the ontology are as follows:

Let $I = \{I_1, I_2, I_3, \dots, I_M\}$, $A = \{A_1, A_2, A_3, \dots, A_N\}$ and $B = \{B_1, B_2, B_3, \dots, B_O\}$ are sets of images, concepts and categories consisting of M , N and O number of elements respectively. Let W_{CB} be a matrix representing binary weights for relationship of a category to an image and is written as:

$$W_{CB} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1M} \\ w_{21} & w_{22} & \dots & w_{2M} \\ \vdots & \vdots & & \vdots \\ w_{O1} & w_{O2} & \dots & w_{OM} \end{bmatrix}, \tag{1}$$

where $w_{kj} = 0$ or $w_{kj} = 1$, $1 \leq k \leq O$ and $1 \leq j \leq M$. Let W_{CI} be a matrix representing the frequency of concepts in image and is written as:

$$W_{CI} = \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1M} \\ f_{21} & f_{22} & \dots & f_{2M} \\ \vdots & \vdots & & \vdots \\ f_{N1} & f_{N2} & \dots & f_{NM} \end{bmatrix}, \tag{2}$$

where f_{ij} is the frequency of a concept A_i in an image I , and $1 \leq i \leq N$ and $1 \leq j \leq M$.

The relationship among image content (i.e., concepts, categories and an image itself) was originally a crisp set defined by W_{CB} and W_{CI} . The relationship was made fuzzy by the proposed methodology. In our system, an image content is represented by three matrices namely weight of the concept to image W_A , weight of the category to image W_B , and weight of the concept to category W_{CF} and are defined as:

$$W_A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NM} \end{bmatrix}, \tag{3}$$

where a_{ij} is the relevancy between the concept A_i and the image I , and $0 \leq a_{ij} \leq 1$, $1 \leq i \leq N$ and $1 \leq j \leq M$. Element of weight of a concept to an image matrix a_{ij} is calculated as:

$$a_{ij} = \frac{f_{ij}}{T_j}, \quad (4)$$

where f_{ij} is the frequency of the concept A_i in the image I and T_j is the total number of concepts in the image I . The weight of a concept to a category is a matrix as shown below:

$$\mathbf{WB} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1O} \\ b_{21} & b_{22} & \cdots & b_{2O} \\ \vdots & \vdots & & \vdots \\ b_{N1} & b_{N2} & \cdots & b_{NO} \end{bmatrix}, \quad (5)$$

where b_{ik} is the relevancy between the concept A_i and the category B_k , and $0 \leq b_{ik} \leq 1$, $1 \leq i \leq N$ and $1 \leq k \leq O$. The proposed formula for calculating weight of the concept to the category b_{ik} is as follows:

$$b_{ik} = \frac{\sum_{j=1}^M a_{ij} w_{kj}}{\sum_{j=1}^M w_{kj}}, \quad (6)$$

The relationship between category and image can be obtained implicitly, i.e., through a transitive property, from weight of concept to image and weight of concept to category matrices. The weight of the category to image is a matrix as shown below:

$$\mathbf{WCF} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1M} \\ c_{21} & c_{22} & \cdots & c_{2M} \\ \vdots & \vdots & & \vdots \\ c_{O1} & c_{O2} & \cdots & c_{OM} \end{bmatrix}, \quad (7)$$

where c_{kj} is the relevancy between the category B_k and the image I , and $0 \leq c_{kj} \leq 1$, and $1 \leq k \leq O$. Element of weight of a category to an image matrix a_{ij} is calculated as:

$$c_{kj} = \frac{\sum_{i=1}^N a_{ij} b_{ik}}{F_{ik}}, \quad (8)$$

where F_{ik} is the number of concepts in a category.

3.2 Image retrieval

A user query consists of keywords that can be i) single or combination of concepts, (ii) single or combination of categories and (iii) combination of concepts and categories. The proposed retrieval algorithm is shown as Algorithm 1.

Algorithm 1. The proposed fuzzy ontology based image retrieval algorithm.

<p>Input: Single or combination of concepts, A_N, OR single or combination of categories, B_O, OR combination of concepts and categories, A_N, B_O, retrieval size S, database size D</p> <p>Output: A list of ranked Images R</p> <p>Basic Idea:</p> <ol style="list-style-type: none"> 1. case 1: [query contains single or combination of concepts] 2. [Use procedure 1.1 to retrieve images] Call RetrieveA(A_N, S, D, R) 3. case 2: [query contains single or combination of categories] 4. [Use procedure 1.2 to retrieve images] Call RetrieveB(B_O, S, D, R) 5. case 3: [query contains combination of concepts categories] 6. Split the query into two. $Q_1 = A_N$ and $Q_2 = B_O$ 7. Call RetrieveA(A_N, S, D, R), Call RetrieveB(B_O, S, D, R) 8. Take intersection of images returned by both the procedures and store it in R. 9. Show R to users. 10. Exit
<p>Procedure 1.1: RetrieveA(A_N, S, D, R)</p> <ol style="list-style-type: none"> 1. Repeat step 2 to 3 for $k=1$ to N [where N is the total number of concepts in Q] 2. Repeat step 3 for $p=1$ to D 3. $R_2 [A_k][p] = W_A [A_k][p]$ 4. Rank each row of R_2 in descending order. Take intersection of rows of R_2 and store the result of intersection in R. Return R based on S.
<p>Procedure 1.2: RetrieveB(B_O, S, D, R)</p> <ol style="list-style-type: none"> 1. Repeat step 2 to 3 for $k=1$ to N [where N is the total number of categories in Q] 2. Repeat step 3 for $p=1$ to D 3. $R_2 [B_k][p] = W_{CF} [B_k][p]$ 4. Rank each row of R_2 in descending order. Take intersection of rows of R_2 and store the result of intersection in R. Return R based on retrieval size.

The detail of algorithm is illustrated below through example.

3.3 Walk-through example

Let $I = \{I_1, I_2, I_3, I_4\}$, $A = \{\text{Sky, Foliage, Grass, Water}\}$ and $B = \{\text{Sky_Cloud, Field}\}$ are sets of images, concepts and categories of the image collection. The matrix W_{CB} represents the binary weights of a category to an image and is given as:

$$W_{CB} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

The matrix W_{CI} , represents the frequency of concepts in images and is defined as:

$$W_{CI} = \begin{bmatrix} 80 & 20 & 70 & 40 \\ 0 & 20 & 30 & 0 \\ 20 & 60 & 0 & 60 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The fuzzy weights in matrices W_A , W_B , and W_{CF} were computed using Eq. (4), Eq. (6) and Eq. (8) and are as follows:

$$W_A = \begin{bmatrix} 0.8 & 0.2 & 0.7 & 0.4 \\ 0 & 0.2 & 0.3 & 0 \\ 0.2 & 0.6 & 0 & 0.6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad W_B = \begin{bmatrix} 0.75 & 0.3 \\ 0.15 & 0.1 \\ 0.1 & 0.6 \\ 0 & 0 \end{bmatrix}$$

$$W_{CF} = \begin{bmatrix} 0.31 & 0.08 & 0.285 & 0.18 \\ 0.18 & 0.15 & 0.12 & 0.24 \end{bmatrix}$$

The fuzzy ontology, constructed according to the above computed weights is shown in Fig. 4. The next step in retrieving an image is to take user requirements and retrieval size (as users are interested in top results only) and apply retrieval algorithm to get the list of retrieved images. If a query contains only concept, i.e., $Q = \{A1\}$. From W_A following vector was extracted against the given query:

$$R = [0.8 \ 0.2 \ 0.7 \ 0.4],$$

The above vector was sorted in the descending order as $[0.8 \ 0.7 \ 0.4 \ 0.2]$ and with the retrieval size of 3 out of 4 images, images I_1 , I_3 , and I_4 , corresponding to the vector values, were returned to the user. From Fig. 4, we can see that I_1 and I_3 are highly relevant images to query “sky”, while I_4 is less relevant because it contains small portion of sky. When a query contains a category, i.e., $Q = \{B1\}$, then from W_{CF} the following vector was extracted against the given query:

$$R = [0.31 \ 0.08 \ 0.285 \ 0.18],$$

The above vector was sorted in the descending order as $[0.31 \ 0.285 \ 0.18 \ 0.08]$ and with the retrieval size of 3, the images I_1 , I_3 , and I_4 were returned to the user. When a query contains both the concept and category, i.e., $Q = \{A1, B1\}$, the query is first split into two queries, i.e., $Q1 = \{A1\}$ containing the concept and $Q2 = \{B1\}$ containing the category. $Q1$ returns the following vector from W_A :

$$R1 = [0.8 \ 0.2 \ 0.7 \ 0.4],$$

and $Q2$ returns the following vector from W_{CF} :

$$R2 = [0.31 \ 0.08 \ 0.285 \ 0.18]$$

Sort both the vectors $R1$ and $R2$ in the descending order, take the intersection of images corresponding to these vector values and store the result in R . Keeping the retrieval size of 3, the images I_1 , I_3 , and I_4 are retrieved for user illustration.

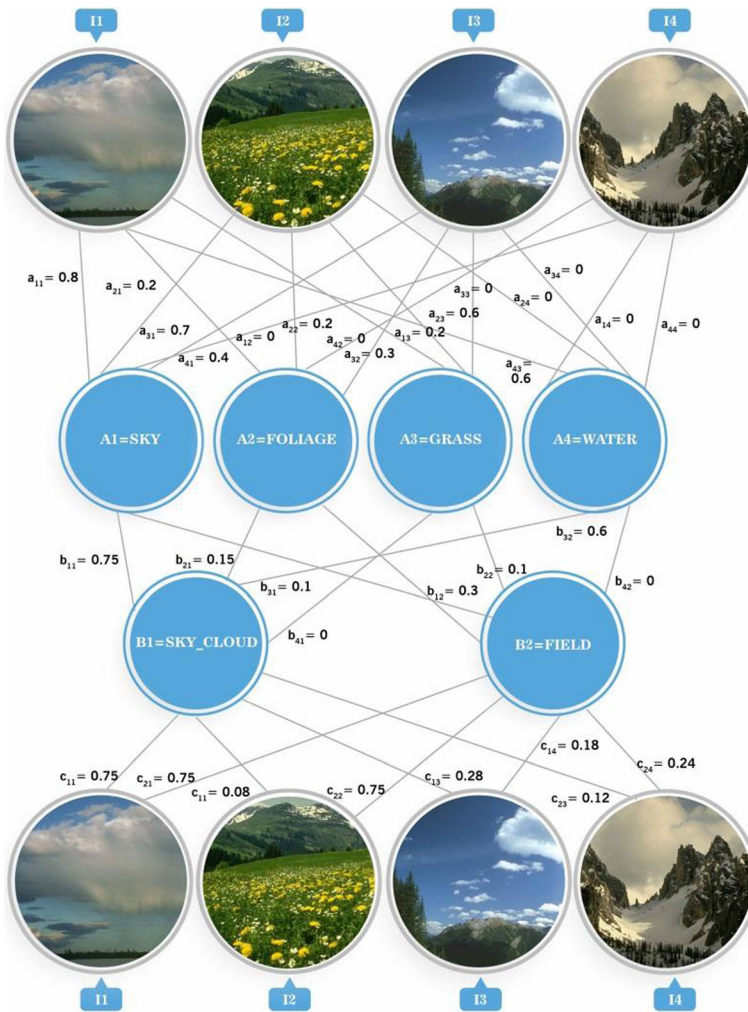


Fig. 4 An example of fuzzy ontology for four images, four concepts and two categories

4 Results and discussion

This section discusses the results achieved in this research. First of all the experimental setup is explained to understand the context of this research. Then two types of evaluation, i.e., objective and subjective, have been carried out to measure the performance of the proposed system.

4.1 Experimental setup

A dataset of seven hundred annotated images (i.e., $M = 700$) about natural scenes [33] has been used to validate the proposed retrieval model. The dataset consists of five categories (i.e., $O = 5$), namely sky_clouds, forest, field, waterscapes and landscape with mountains and ten concepts (i.e., $N = 10$), namely sky, foliage, grass, rocks, mountains, trunks, flower, water, sand

and fields. In order to compare our system, we have selected the fuzzy relational ontological model proposed in [22]. In that system, a user query is composed of concepts, categories or the combination of both. When user enters a query, it is expanded based on ontological knowledge. After expansion, relevancy score is calculated between the query keywords and the ontology concepts based on fuzzy operations.

Figure 5 shows the retrieved results of the proposed system (i.e., on the left side) and the reference system (i.e., on the right side) for three different queries with retrieval size of 15. The first row shows the result of query-by-concept, i.e., “flower”, second row shows the result of query-by-category, i.e., “field” and the third row shows the result of query-by-concept & category, i.e., “flower and field”.

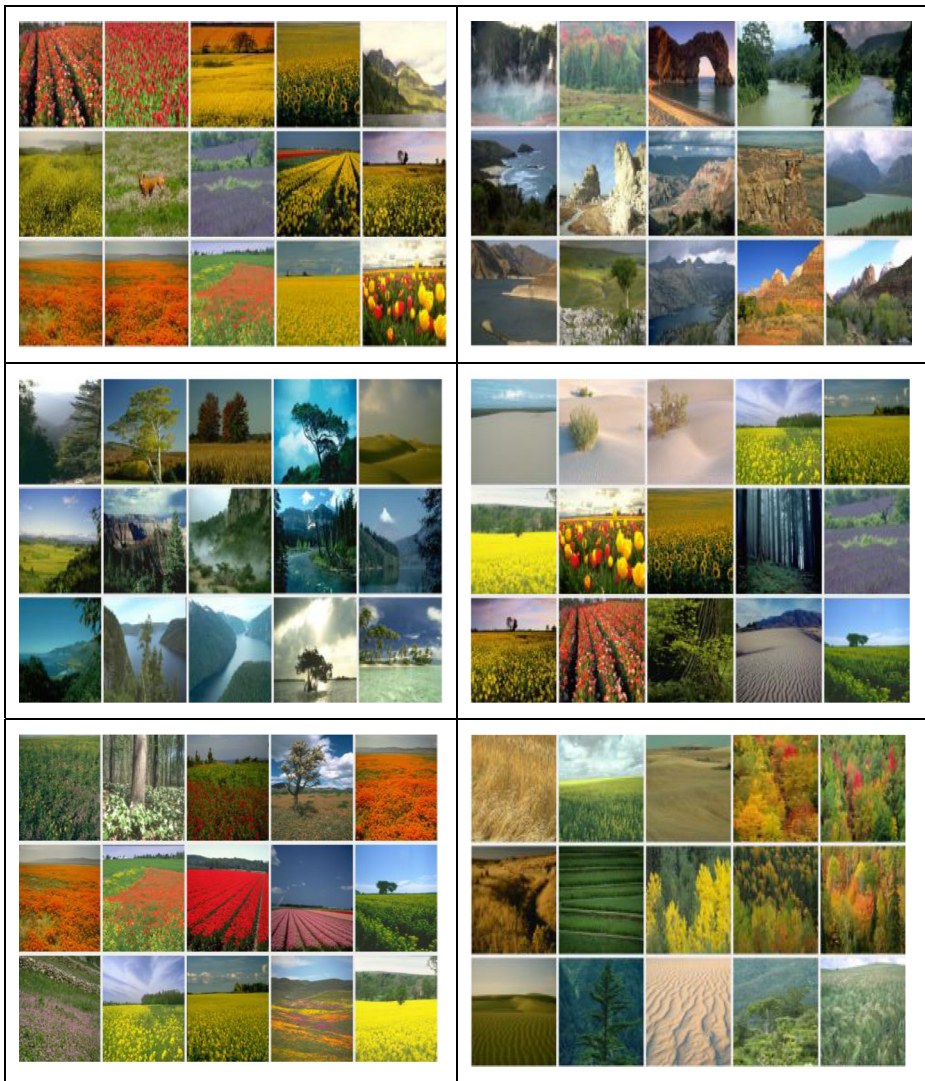


Fig. 5 Retrieval output of the proposed system and reference system for retrieval size 15. Left column: Proposed system, Right column: Reference system [22]

A total of 167 queries have been designed for evaluation purpose in which 10 queries were based on $01 \times$ concept, such as “sky”, 41 queries were based on $02 \times$ concepts, such as “sky and grass” and 71 queries were based on $03 \times$ concepts, such as “sky grass water”. Similarly, 5 queries were based on $01 \times$ category, such as “Sky_Cloud” and 6 queries were based on $02 \times$ categories, such as “Sky_Cloud and Field”. 34 queries were based on $01 \times$ concept & $01 \times$ category, such as “Sky and Field”. The performance of the proposed system and the reference system has been evaluated in two different ways: (i) objective and (ii) subjective.

4.2 Objective evaluation

The system was objectively evaluated using two different approaches: (i) mean and variance and (ii) precision, recall and average normalized modified retrieval rank (ANMRR). Mean and variance indicate the amount of required information in the retrieved images, whereas precision, recall and ANMRR indicate the retrieval of relevant images in the results. Both approaches are described in detail in the next sections.

4.2.1 Mean and variance

Mean is computed by taking the sum of all the values in the dataset and then dividing it by the total number of values in the dataset. On the other hand, variance measures the spread or variability of the values from mean in the dataset. In order to show that the proposed system retrieves images with maximum amount of the user requested information, mean and variance were computed for the top 15 retrieved results. In this evaluation, mean is the average occurrence of particular concept in an image and variance is the variability of the concept's occurrences from the mean. Higher value of mean indicates images with higher amount of concept occurrences are retrieved.

Table 1 Occurrence of “foliage” in the top 15 retrieved results of the proposed and reference system in response to query-by-concept

Retrieved Image	Proposed system Occurrence [0–1]	Reference system Occurrence [0–1]
I1	1	0.09
I2	1	0.09
I3	1	0.34
I4	1	0
I5	1	0.32
I6	1	0.4
I7	1	0.16
I8	1	0.02
I9	1	0.2
I10	1	0.57
I11	0.93	0.15
I12	1	0.4
I13	0.89	0.62
I14	0.89	0.12
I15	0.93	0
Mean	0.967	0.232
Variance	0.0017	0.0372

Table 1 shows the occurrence of foliage computed using Eq. (4) in the top 15 results for only one query, i.e., “foliage”. The mean (i.e., 0.967) of the proposed system shows that the average occurrence of foliage in the top 15 results is around 1. On the other hand, the mean (i.e., 0.232) of the reference system indicates that the retrieved images contain foliage in lesser quantity. Similarly, the variance (i.e., 0.0017) of the proposed system shows that the variation in the occurrence of foliage in all the retrieved images is very small as compared to the variance (i.e., 0.0372) of the reference system.

Figure 6 shows the mean and variance of the proposed and the reference system against ten different $01 \times$ concept based queries, described as: C_1 = “Sky”, C_2 = “Foliage”, C_3 = “Mountain”, C_4 = “Grass”, C_5 = “Field”, C_6 = “Rock”, C_7 = “Water”, C_8 = “Trunk”, C_9 = “Flower”, and C_{10} = “Sand”. It is evident that the proposed system performs better as compared to the reference system.

Table 2 shows the mean and variance of two randomly selected $02 \times$ concepts queries, i.e., Q_1 = “Foliage-Trunk” and Q_2 = “Rock-Water” and $03 \times$ concepts queries, i.e., Q_3 = “Sky-Foliage-Grass” and Q_4 = “Sky-Foliage-Field”. Mean and variance values shown in Table 2 are better for the proposed system as compared to the reference system.

Figures 7 and 8 show the results of mean and variance of different concept occurrences in two randomly selected query-by category, i.e., Q_5 = “Landscape with mountain” and Q_6 = “Forest” respectively. In Q_5 , results of the reference system are slightly better as it shows presence of concepts C_1 , C_2 and C_6 , i.e., sky, foliage and rocks in higher amount in the images as compared to the proposed system, whereas in Q_6 the proposed methodology performs better in terms of mean and variance.

Figure 9 shows mean and variance of 5 randomly selected query-by-concept & category, i.e., Q_7 = “Sky and Field”, Q_8 = “Foliage and Sky_Cloud”, Q_9 = “Flower and Sky_cloud”, Q_{10} = “Flower and Field” and Q_{11} = “Sand and Field”. The plots show the amount of concept present in the top 15 retrieved images from a particular category. From the plots, it is evident that the proposed system performs better as compared to the reference system for all the five queries.

4.2.2 Precision, recall and ANMRR

Three different evaluation measures (i) precision, (ii) recall and (iii) ANMRR [10] were computed for each query. High precision value indicates that more relevant results are

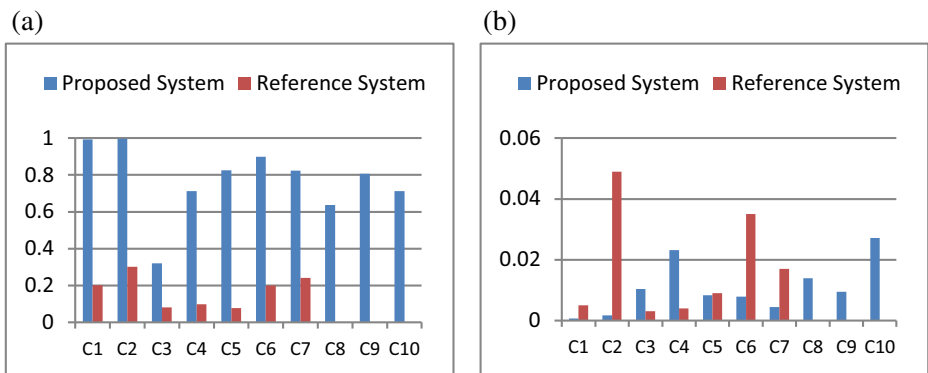


Fig. 6 Comparison of the proposed system with reference model [22] for ten different queries by concept in terms of (a) mean and (b) variance

Table 2 Comparison of the proposed system with reference system [22] for different queries by concept and categories in terms of (a) mean and (b) variance

		Proposed System Mean	Reference System Mean	Proposed System Variance	Reference System Variance
02 × Concepts					
Q ₁	Foliage	0.512	0.301	0.0051	0.049
	Trunk	0.248	0.002	0.030	0.00005
Q ₂	Rock	0.249	0.199	0.0004	0.035
	Water	0.352	0.240	0.051	0.017
03 × Concepts					
Q ₃	Sky	0.572	0.325	0.010	0.008
	Foliage	0.157	0.208	0.019	0.001
	Grass	0.267	0.21	0.023	0.002
Q ₄	Sky	0.527	0.36	0.005	0.0030
	Foliage	0.18	0.20	0.016	0.007
	Field	0.259	0	0.028	0

retrieved, whereas high value of recall indicates that most of the relevant results are retrieved. ANMRR score indicates the performance of algorithms based on ranking of the results. The low value of ANMRR means the algorithm ranked the results in better way. Readers interested in detail of ANMRR may consult [6, 15, 17] for details. Table 3 shows the result of proposed system and reference system in terms of precision, recall and ANMRR for different retrieval sizes, such as 15%, 30%, 50% and 100%. Varying retrieval sizes allow us to judge a system performance at different levels. For example, for the top 15% results of the proposed system when a query contains 01 × concept, precision value 1 indicates that all the retrieved results are relevant, whereas recall value of 0.1425 and ANMRR value of 0.8737 indicate that there are still many relevant images in the ground truth list for that query. As the retrieval size is increased, the recall and ANMRR values for 01 × concept for the proposed system change to 0.2911 and 0.6735 respectively, showing that more relevant images are retrieved. For 100% retrieval size, recall and ANMRR values are changed to 0.9816 and 0.0175 respectively. In case of query-by-concept & category, precision, around 1 for the proposed system shows that all the retrieved results are relevant as compared to 0.7647 value of the reference system. However, the reference system shows slightly better results in the case of a query by category.

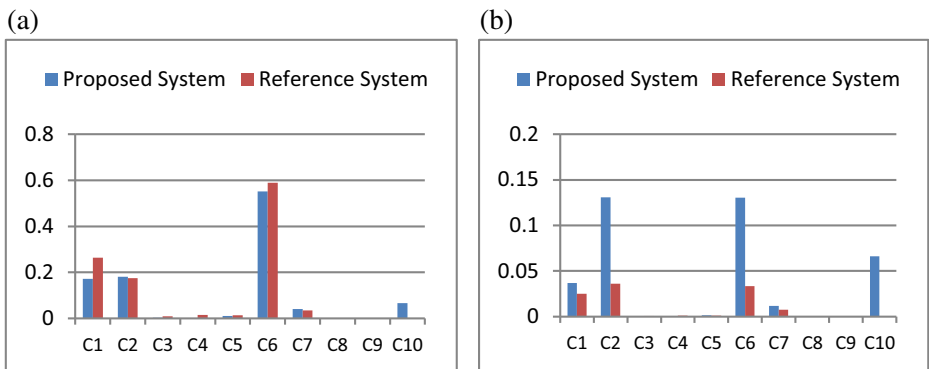


Fig. 7 Comparison of the proposed system with reference system [22] showing occurrence of different concepts in category “Landscape with Mountain” in terms of (a) mean and (b) variance

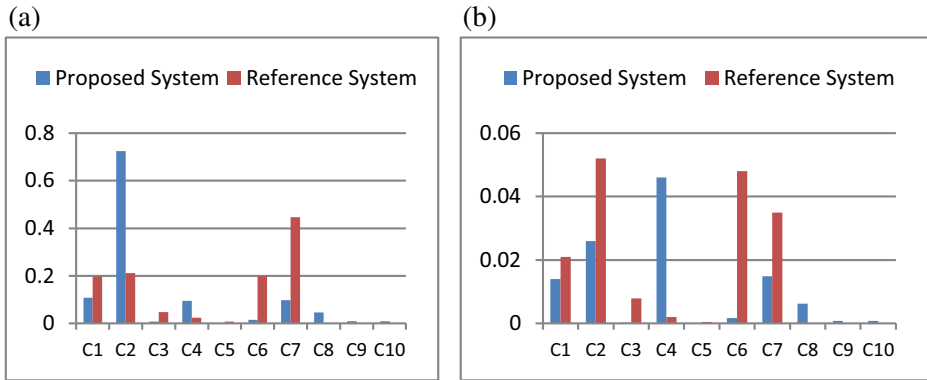


Fig. 8 Comparison of the proposed system with reference system [22] showing occurrence of different concepts in category “Forest” in terms of (a) mean and (b) variance

4.2.3 Discussion

It is evident from the results that the proposed system shows better performance in case of query-by-concept and query by-concept & category. In case of query-by-category, the reference system shows slightly better performance because the dataset contains predefined categories in which the relationship between an image and a category is binary. Also an image can belong to only one category even if the content of an image allows it to belong to different categories with different degrees of membership. For example, see the images in Fig. 10 where the image on the left side belongs to the category “Sky_Cloud” and the image on the right side belongs to the category “Field”. However, they both include similar contents, such as sky, field, mountain and foliage. If the search is based on category (e.g., Sky_Cloud) and retrieval system retrieves image from a different category (e.g., Field) having similar content, then can we judge the system to be successful despite the fact that the evaluation measures compute a very poor result? The proposed system retrieves images based on fuzzy ontology in which the

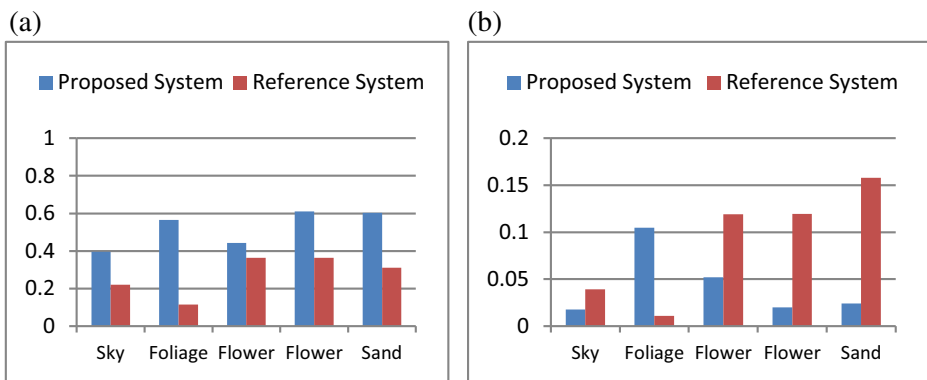


Fig. 9 Comparison of the proposed system with reference system [22] showing occurrence of different concepts in different categories in terms of (a) mean and (b) variance

Table 3 Comparison of the proposed and reference system in terms of precision, recall and ANMRR results for different queries

Query	No. of Queries	Precision		Recall		ANMRR		
		Pro. Model	Ref Model	Pro. Model	Ref Model	Pro. Model	Ref Model	
Retrieval Size: 15%								
Concept	01 × concept	10	1	0.5969	0.1425	0.1212	0.8373	0.8632
	02 × concepts	41	1	0.3626	0.1240	0.2890	0.8700	0.7205
	03 × concepts	71	1	0.2650	0.1130	0.3964	0.8883	0.6666
Category	01 × category	5	0.3909	0.9897	0.0514	0.0576	0.9458	0.925
	02 × categories	6	0.4617	0.1233	0.0514	0.0577	0.9413	0.9311
01 × concept&01 × category	34	1	0.7647	0.0303	0.0294	0.9651	0.5673	
Retrieval Size: 30%								
Concept	01 × Concept	10	1	0.5984	0.2911	0.2447	0.6735	0.7313
	02 × Concepts	41	1	0.3642	0.2717	0.5804	0.7094	0.5123
	03 × Concepts	71	1	0.2663	0.2523	0.7967	0.7366	0.5216
Category	01 × Category	5	0.4476	0.9948	0.1225	0.1241	0.8703	0.8461
	02 × Categories	6	0.4280	0.1363	0.0952	0.1229	0.8913	0.8495
01 × Concept&01 × Category	34	1	0.7647	0.0625	0.2352	0.9273	0.1270	
Retrieval Size: 50%								
Concept	01 × Concept	10	1	0.5990	0.4898	0.4081	0.4670	0.5716
	02 × Concepts	41	1	0.3648	0.4703	0.9689	0.5034	0.3647
	03 × Concepts	71	1	0.2668	0.4531	0.7977	0.5273	0.6522
Category	01 × Category	5	0.4243	0.9969	0.1980	0.2435	0.7968	0.7311
	02 × Categories	6	0.3759	0.1480	0.1344	0.2355	0.8483	0.7133
01 × Concept&01 × Category	34	1	0.7647	0.1061	0.7941	0.8771	0.4573	
Retrieval Size: 100%								
Concept	01 × Concept	10	1	0.5995	0.9816	0.8184	0.0175	0.2448
	02 × Concept	41	0.9661	0.3653	0.9200	0.9690	0.0799	0.6457
	03 × Concept	71	1	0.2672	0.9287	0.7980	0.0777	0.6522
Category	01 × Category	5	0.3802	0.9984	0.4653	0.300	0.5951	0.6541
	02 × Category	6	0.4647	0.2891	0.3335	0.3254	0.6527	0.5459
01 × Concept&01 × Category	34	1	0.7647	0.2140	0.8235	0.7569	0.5600	

relationship between an image and a category is fuzzy. An image belongs to different categories with different degrees of membership based on the frequency of concepts contained in the image. However, the objective measures used for evaluation show the result on the basis of predefined categories in which an image belongs to just one category and this is the reason why the results of query-by-category of the proposed

**Fig. 10** Image on the left side belongs to category “Sky_Cloud” and image on the right side belongs to category “Field”

system are poor. To ensure this problem, the retrieval system performance is subjectively evaluated. The next section shows the subjective results for the same dataset with the same set of queries.

4.3 Subjective evaluation

Subjective evaluation is carried out based on perception of human observers [18]. The problem of retrieval system evaluation is its relevance, which is a subjective notion. For complete evaluation of the system, users' expectation is of vital importance. The ranking of retrieved images varies with users depending on the particular content that the users' attention is currently focused on.

Feedback from 300 observers, 55% male and 45% female, has been recorded in the digital systems laboratory of Computer Engineering Department, University of Engineering and Technology, Taxila (UETT), Pakistan. 280 participants were in the first age group (19–40 years) and their qualification was intermediate or BSc or MSC or PhD. The remaining 20 in the second age group (30–45 years) were the faculty members at UETT. The maximum of three query-result pairs of two different retrieval systems (i.e., proposed and reference) were shown to each user with retrieval size of 15, as shown in Fig. 11. The results are shown in random order, i.e., the user does not know which retrieval system is under evaluation. Each query was evaluated by five users in order to ensure that results are not biased by specific user score. The evaluation process took almost three months to complete the feedback process. The mean opinion score (MOS) in terms of normalized discounted cumulative gain (NDCG) and the mean overall score (O) were recorded

The screenshot shows a feedback form for a subjective evaluation. At the top, it identifies the user as 'User ID 2' and the query as 'Sand'. The form is divided into two columns, 'System 1' and 'System 2', each containing five images. Below each image is a 'User Score' slider with a scale from 0 to 5. The scores for the images in System 1 are: 0, 0, 0, 3, and 5. The scores for the images in System 2 are: 0, 0, 0, 3, and 5. At the bottom of the form, there are two summary fields: 'Overall Score (O)' with a value of 3.4 and 'Cumulative Score (C)' with a value of 3.4. A 'Done' button is located at the very bottom.

Fig. 11 Feedback form for subjective evaluation containing 15 retrieved images for query “Sand”

Table 4 Comparison in terms of Mean overall score and NDCG of the proposed and reference system

Retrieval size = 15

Query contains		No. of Queries	Mean Overall Score (O)		Normalized Discounted Cumulative Gain (NDCG)	
			Proposed System	Reference System	Proposed System	Reference System
Concept	01 × concept	10	4.1826	1.484	0.8476	0.4217
	02 × concepts	41	3.2426	2.6173	0.7116	0.6263
	03 × concepts	71	3.3716	2.7819	0.7083	0.6734
Category	01 × category	5	2.92	2.9493	0.6214	0.5208
	02 × categories	6	3.2488	3.22	0.6551	0.6767
01 × concept & 01 × category		34	3.3019	2.1760	0.7547	0.5012

from users’ feedback. MOS [29, 36] is a commonly used metric in which each retrieved image is evaluated by selecting a score that ranges from 0 to 5 and is defined as follows:

$$MOS = \begin{cases} 0, & \text{when irrelevant image is retrieved} \\ 1, & \text{when slightly relevant image is retrieved} \\ 2, & \text{when somewhat relevant image is retrieved} \\ 3, & \text{when relevant image is retrieved} \\ 4, & \text{when very relevant image is retrieved} \\ 5, & \text{when highly relevant image is retrieved} \end{cases} \quad (9)$$

where $0 \leq MOS \leq 5$ and MOS is taken from a user against each retrieved image. NDCG is defined as:

$$NDCG = \frac{1}{Q} \sum_{i=1}^Q DCG, \quad (10)$$

where Q is the total number of queries, and DCG is defined as:

$$DCG = \frac{1}{U} \sum_{i=1}^U \sum_{j=1}^S \frac{MOS_{ij}}{\log_2 j}, \quad (11)$$

where MOS_{ij} is the relevancy score of the j th retrieved image assigned by the i th user, U is the

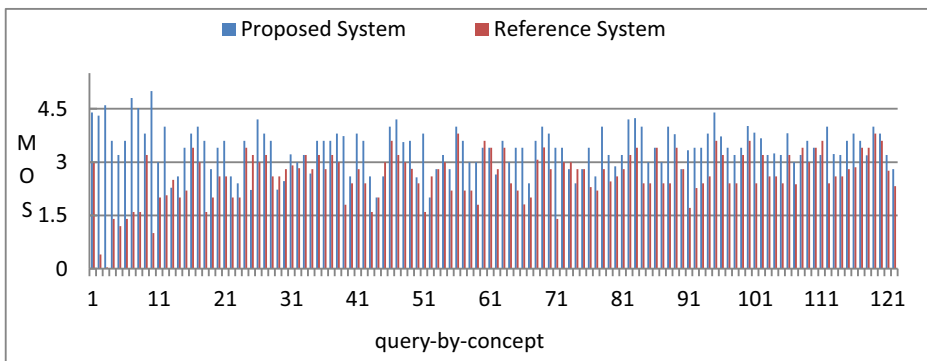


Fig. 12 Comparison of the proposed system with reference system [22] for different queries by concept in terms of mean overall score

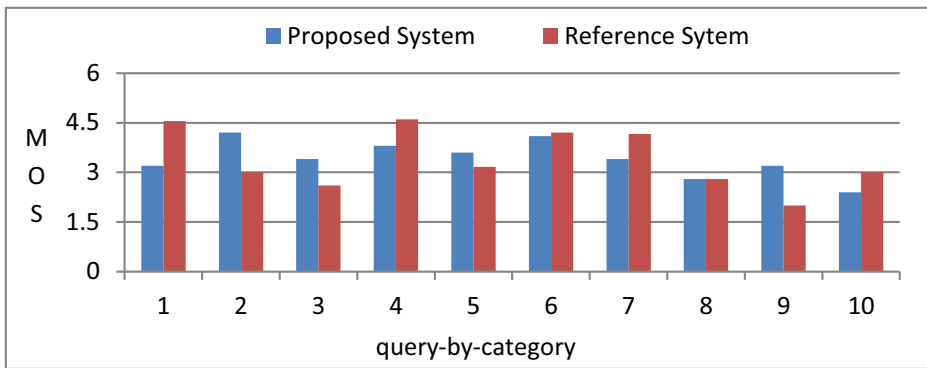


Fig. 13 Comparison of the proposed system with reference system [22] for different queries by category in terms of mean overall score

total number of users, and S is the retrieval size. Similarly, the mean overall score O associated with each query is defined as:

$$O = \frac{1}{U} \sum_{i=1}^U u_i, \tag{12}$$

where u_i is the overall score of the retrieval result by i^{th} user.

NDCG measures the performance of a system based on graded relevance that varies from 0 to 5 (i.e., 0 = not relevant and 5 = highly relevant). The usefulness of a retrieved image is measured based on its position in a retrieved list. Higher NDCG value indicates that the highly relevant images are retrieved at the top of the list. From Table 4, it is evident that the proposed system outperforms the reference system when evaluated subjectively. The proposed system shows higher mean overall score against all the queries, except when a query contains $01 \times$ category where the reference model shows slightly high value with a difference of 0.0293 which is tolerable. Similarly, NDCG values are higher for the proposed system against all the queries except when a query contains $02 \times$ categories where the reference model shows a slightly high value with a difference of 0.0213 which is acceptable.

Figure 12 shows mean overall score of five users for 122 queries-by-concept (i.e., 10 queries contain $01 \times$ concept, 41 queries contain $02 \times$ concepts and the remaining 71 queries contain

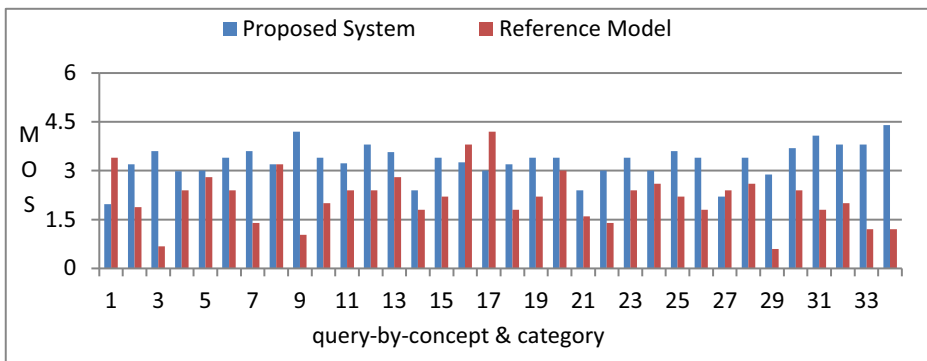


Fig. 14 Comparison of the proposed system with reference system [22] for different queries by concept & category in terms of mean overall score

03 × concepts). From the plot, it is obvious that users are satisfied with the retrieved results of the proposed system as mean overall score for any query lies in the range from 2 to 4.8, whereas for the reference system, it ranges from 0 to 4.

Figure 13 shows mean overall score of five users for 10 queries-by-category (i.e., 4 queries contain 01 × category and 6 queries contain 02 × categories), whereas Fig. 14 shows mean overall score of five users for 34 queries-by-concept-and-category for the proposed and the reference system. It is evident from the plots that the proposed system performs better in most of the queries as compared to the reference system.

5 Conclusion and future work

In this paper, a fuzzy ontology based system has been proposed for improving the performance of image retrieval. First of all, fuzzy ontology was constructed by utilizing the concepts and categories associated with images. Concepts describe the objects that an image contains and category depicts a scene based on the frequency of concepts inside the image. Concepts, categories and images are linked among themselves with fuzzy values in the ontology. Then users are provided with an interface to input keywords that may consist of concepts, categories or both. Retrieved results are ranked based on the relevancy between the keywords of query and images. The advantages of the proposed model are (i) the relationship between an image and concepts, and image and categories are fuzzy values that resolve the problem of binary annotation and retrieval and (ii) it allows an image to belong to different categories with different degrees of membership based on the content of an image. With the help of reasoning through the ontology, a query asking for either concepts or categories can be expanded with their respective categories and concepts respectively for result improvement.

For evaluating the performance of the proposed system, both objective and subjective measures were used. Objective evaluation results show better performance for query-by-concept and query-by-concept & category whereas for query-by-category the reference system shows slightly better performance. To investigate the reason, we have subjectively evaluated the same set of queries with 300 observers of different age groups and qualifications. The experimental results show that the proposed system achieves higher values for MOS in terms of normalized discounted cumulative gain and mean overall score as compared to the reference system for all sets of queries.

Currently, we are improving the ranking of the retrieved results using fuzzy relations in ontology where user requirements include multiple concepts and categories.

References

1. Allani O, Zghal HB, Mellouli N, Akdag H (2016) A knowledge-based image retrieval system integrating semantic and visual features. *Proc Comput Sci* 96:1428–1436
2. Bannour H, Hudelot C (2014) Building and using fuzzy multimedia ontologies for semantic image annotation. *Multimed Tools Appl* 72(3):2107–2141
3. Cai B, Zheng C, Yang S, Zheng JZ (2007) Mixed query image retrieval system. ICIA'07. International conference on information acquisition, 2007
4. Dasiopoulou S, Kompatsiaris I, Strintzis MG (2010) Investigating fuzzy DLs-based reasoning in semantic image analysis. *Multimed Tools Appl* 49(1):167–194
5. Galindo J (2008) Handbook of research on fuzzy information processing in databases. Information Science Reference. <https://www.ibspan.waw.pl/~kacprzyk/papers/Handbook-Galindo-2008.pdf>
6. Guo J-M, Prasetyo H (2015) Content-based image retrieval using features extracted from halftoning-based block truncation coding. *IEEE Trans Image Process* 24(3):1010–1024

7. Hakimpour F, Timpf S (2001) Using ontologies for resolution of semantic heterogeneity in GIS. Proceedings of 4th AGILE Conference on Geographic Information Science
8. Hornq Y-J, Chen S-M, Lee C-H (2001) Automatically constructing multi-relationship fuzzy concept networks in fuzzy information retrieval systems. The 10th IEEE international conference on fuzzy systems, 2001
9. Hyvonen E, Saarela S, Styrman A, Viljanen, K (2003) Ontology-based image retrieval. WWW (Posters)
10. Jarvelin K, Kekalainen J (2000) IR evaluation methods for retrieving highly relevant documents. Proceedings of the 23rd annual international ACM SIGIR conference on research and development in information retrieval
11. Kaur H, Jyoti K (2013) Survey of techniques of high level semantic based image retrieval. IJRCC 2(1):15–19
12. Liaqat M (2013) Image classification and retrieval based on crisp and fuzzy ontology. 2013 3rd international conference on Computer, Control & Communication (IC4)
13. Liu S, Chia L-T, Chan S (2004) Ontology for nature-scene image retrieval. OTM confederated international conferences “on the move to meaningful internet systems”
14. Liu Y, Zhang D, Lu G, Ma W-Y (2007) A survey of content-based image retrieval with high-level semantics. Pattern Recogn 40(1):262–282
15. Long, F, Zhang, H, & Feng, DD (2003) Fundamentals of content-based image retrieval. In: Multimedia Information Retrieval and Management. Springer, Berlin Heidelberg, pp 1–26
16. Luo B, Wang X, Tang X (2003) World Wide Web based image search engine using text and image content features. Electron Imaging 2003
17. Manjunath BS, Ohm J-R, Vasudevan VV, Yamada A (2001) Color and texture descriptors. IEEE Trans Circ Syst video Technol 11(6):703–715
18. Muller H, Muller W, Squire DM, Marchand-Maillet S, Pun T (2001) Performance evaluation in content-based image retrieval: overview and proposals. Pattern Recogn Lett 22(5):593–601
19. Ogawa Y, Morita T, Kobayashi K (1991) A fuzzy document retrieval system using the keyword connection matrix and a learning method. Fuzzy Sets Syst 39(2):163–179
20. Park K-W, Jeong J-W, Lee D-H (2007) OLYBIA: ontology-based automatic image annotation system using semantic inference rules. International Conference on Database Systems for Advanced Applications
21. Pereira R, Ricarte I, Gomide F (2006) Fuzzy relational ontological model in information search systems. Capturing Intell 1:395–412
22. Pereira R, Ricarte I, Gomide F (2009) Information retrieval with FROM: the fuzzy relational ontological model. Int J Intell Syst 24(3):340–356
23. Radecki T (1979) Fuzzy set theoretical approach to document retrieval. Inf Process Manag 15(5):247–259
24. Sarwar S, Qayyum ZU, Majeed S (2013) Ontology based image retrieval framework using qualitative semantic image descriptions. Proc Comput Sci 22:285–294
25. Schober J-P, Hermes T, Herzog O (2004) Content-based image retrieval by ontology-based object recognition. Proc. Workshop on Applications of Description Logics, Ulm, Germany
26. Schreiber AT, Dubbeldam B, Wielemaker J, Wielinga B (2001) Ontology-based photo annotation. IEEE Intell Syst 16(3):66–74
27. Simou N, Athanasiadis T, Stoilos G, Kollias S (2008) 'Image indexing and retrieval using expressive fuzzy description logics'. Signal. Image Video Process 2(4):321–335
28. Smeaton AF (1996) Information retrieval and hypertext. Kluwer, New York
29. Strejil RC, Winkler S, Hands DS (2016) Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives. Multimedia Systems 22(2):213–227
30. Styrman A (2005) Ontology-based image annotation and retrieval. Ph.D. dissertation, master thesis
31. Town C (2006) Ontological inference for image and video analysis. Mach Vis Appl 17(2):94–115
32. Ul-Qayyum Z, Cohn AG, Klippel A (2010) Psychophysical evaluation for a qualitative semantic image categorisation and retrieval approach. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems
33. Vogel J, Schiele B (2007) Semantic modeling of natural scenes for content-based image retrieval. Int J Comput Vis 72(2):133–157
34. Vogel J, Schwaninger A, Wallraven C, Bulthoff HH (2007) Categorization of natural scenes: Local versus global information and the role of color. ACM Trans Appl Percept (TAP) 4(3):19
35. Wang H, Liu S, Chia L-T (2006) Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches. Proceedings of the 14th ACM international conference on multimedia
36. Yang M, Sowmya A (2015) An underwater color image quality evaluation metric. IEEE Trans Image Process 24(12):6062–6071
37. Zadeh LA (1975) Fuzzy logic and approximate reasoning. Synthese 30(3):407–428



Madiha Liaqat holds MSc in Software Engineering from University of Engineering and Technology, Taxila (UET-T), Pakistan. She is serving as an Assistant professor at UET-T. Her main areas of interest include Artificial Intelligence, Semantic Web and Image Processing.



Sharifullah Khan received PhD in Computer Science from the University of Leeds, Leeds, UK in 2002. He is an associate professor at School of Electrical Engineering and Computer Science (SECS), the National University of Sciences and Technology (NUST), Islamabad, Pakistan. He is member of ACM. He leads the Data Engineering for Large Scale Applications (DELSA) Research Lab. at SECS. His major area of research interest is databases, data integration, ontology engineering and information retrieval.



Muhammad Majid received BSc. in Computer Engineering with honors from University of Engineering and Technology (UET) Taxila, Pakistan in 2005, MSc. in Data Communications with distinction and PhD in Electronic and Electrical Engineering from the University of Sheffield, UK in 2007 and 2011 respectively. He is currently an Assistant Professor at Department of Computer Engineering, UET Taxila and heading Signal, Image, Multimedia Processing and LEarning (SIMPLE) research group since January 2012. He was a recipient of Faculty Development Program Scholarship of Higher Education Commission (HEC) Pakistan for his MSc and PhD Studies. His research interests include image and video coding, multimedia signal processing, biomedical signal processing and emotion classification in response to multimedia content.