TELKOMNIKA, Vol.17, No.5, October 2019, pp.2572~2586
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI:10.12928/TELKOMNIKA.v17i5.10202

Understanding user intention in image retrieval: generalization selection using multiple concept hierarchies

Abdelmadjid Youcefa*[1], Mohammed Lamine Kherfi[2], Belal Khaldi[3], Oussama Aiadi[4]
[1]LAGE Laboratory, Université Kasdi Marbah Ouargla, Algeria
[2]LAMIA Laboratory, Université du Québec à Trois-Rivières, 3351, boul. des Forges, C.P. 500, Trois-Rivières, Canada
[2,3,4]LINATI Laboratory, Université Kasdi Marbah Ouargla, Algeria
*Corresponding author, e-mail: abdelmadjidyoucefa@gmail.com

Abstract

Image retrieval is the technique that helps users to find and retrieve desired images from a huge image database. The user has firstly to formulate a query that expresses his/her needs. This query may appear in textual form as in semantic retrieval (SR), in visual example form as in query by visual example (QBVE), or as a combination of these two forms named query by semantic example (QBSE). The focus of this paper lies in the techniques of analyzing queries composed of multiple semantic examples. This is a very challenging task due to the different interpretations that can be drawn from the same query. To solve such a problem, we introduce a model based on Bayesian generalization. In cognitive science, Bayesian generalization, which is the base of most works in literature, is a method that tries to find, in one hierarchy of concepts, the parent concept of a given set of concepts. In addition and instead of using one single concept hierarchy, we propose a generalization so it can be used with multiple hierarchies where each one has a different semantic context and contains several abstraction levels. Our method consists in finding the optimal generalization by, firstly, determining the appropriate concept hierarchy, and then determining the appropriate level of generalization. Experimental evaluations demonstrate that our method, which uses multiple hierarchies, yields better results than those using only one single hierarchy.

Keywords: *bayesian models of generalization, concept hierarchy, generalization of concepts, image retrieval, query expansion, user intention*

1. Introduction

Due to the explosive growth in digital images, there has been an increasing interest in developing techniques to help users retrieving their desired images. These techniques are called “image retrieval” and they can be classified into two main categories which are, content based image retrieval CBIR [1-4] and text-based image retrieval TBIR [5-7].CBIR techniques use the visual content in order to retrieve, for a given query (e.g., image example, sketch, feature vector, etc.) [8], the similar ones. This visual content can be represented in terms of global features [9,10] (color, shape, and texture) or local features [11] (SIFT key points …).

Query by visual example QBVE is one of the most used approaches in CBIR. However, the semantic gap between the low-level visual features and the high-level semantic meaning of images causes a high limitation in CBIR performance. The semantic gap could be defined as the contradiction between the human judgment and CBIR results. In other words, the semantic gap is the discrepancy between two interpretations, one of the user and the other of the machine [12].

TBIR technics use text (e.g., image annotation or text surrounding it) as image descriptor. Due to its simplicity and rapidity, TBIR seems to be more desirable and practical for users. However, the quality of TBIR depends on the quality of the annotations that are often ambiguous and incomplete. For example, the same image may be annotated with two very different annotations based on the interests or the psychological state of the annotator. Additionally the annotations may be incomplete and do not fully describe the content of the image.

Received June 5, 2018; Revised February 10, 2019; Accepted March 12, 2019

In order to eliminate the limitations of QBVE and TBIR, an alternative paradigm has been proposed and denoted as query by semantic example QBSE that combines both techniques [13]. In current work, we are concerned with QBSE. In QBSE paradigm, the query is composed of multiple images, where each image is labeled with different keywords that describe the different visual concepts within the image (e.g., house, rain, sunset, etc.) As a query, the system uses the keywords annotating the images rather than the images themselves. Furthermore, and in order to obtain a better performance, the system should not use these concepts as they are, instead, it has to generalize them to some common or more general concepts (e.g., the user is looking for animals, landscapes, etc.). The process of moving from a set of concepts to a more common or general concept is called “generalization”.

Indeed, using queries that are composed of multiple images (i.e., multiple semantic examples) could significantly improve results. However, finding the appropriate generalization for these semantic examples is a very complicated task. Recently, many studies have been done trying to understand and simulate how humans generalize. Some of those works have used machine vision techniques [14-16], others have opted for Bayesian models of generalization [17-21]. Thus, a great progress has been achieved and generalization methods have been proposed. Starting from one concept hierarchy and a set of given positive concepts, the key idea is to find the appropriate level these concepts need to be generalized to. A concept hierarchy is made up of several abstraction levels where each level holds a set of concepts, which are represented by leaf nodes, as Figure 1 shows. However, one should know that the same set of concepts could be represented by different concept hierarchies based on the selected context. For example, animals could be categorized in a concept hierarchy based on their classes, region of leaving and diet, etc.

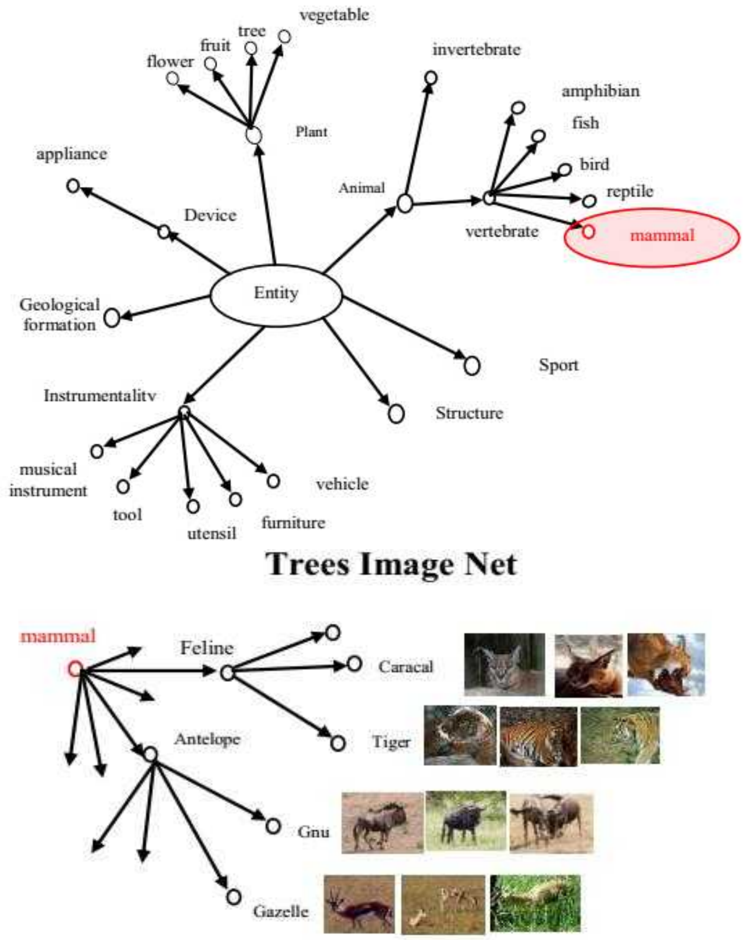


Figure 1. An example of a concept hierarchy and the corresponding images of some leaf nodes (ImageNet as instance)

Despite the great progress achieved by literature works trying to improve generalization, using only one concept hierarchy restricts machine to only one semantic context. Instead, machine should be able to generalize in multiple categories (i.e., contexts) as human does. The generalization should, therefore, be carried out using multiple concept hierarchies. To make things clearer, let us take the example illustrated in Figure 2. Human can generalize the concepts Elephant, Zebra and Giraffe to one hypothesis from the following hypothesis space H= {Mammal, Africa animals, Herbivores}. A hypothesis space is a set of all the possible generalizations obtained from the concepts that compose the given query.

As shown in Figure 2 (a), animals have been categorized according to their family, whereas in Figure 2 (b) according to their diet and in Figure 2 (c) according to their region of living. This means that the generalization in each case will be performed using different concept hierarchy. Therefore, future works should focus on how to combine multiple concept hierarchies to grasp human intention by determining the appropriate context and level of the generalization.

In classical techniques, similarity between images was calculated based on the number of common concepts that annotate those images. However, our approach is not limited to this naïve technique. Instead, it also analyses the semantic relationship between different image concepts. In concept hierarchies, a semantic relationship could be defined as a link that binds two concepts (i.e., father/son nodes). Locating the father node for a given set of nodes is called generalization. It is a very challenging task to determine what relationship assembles a set of given concepts. For example, are Elephant, Giraffe and Zebra Mammals or African animals? Our approach tries to generalize the query concepts by finding the most probable relationship that assembles them. In addition, we extract the concepts that are related to those of the query, which are called "hidden concepts".

To better grasp user intention, in this paper, we propose a method that generalizes user queries using multiple concept hierarchies. In our approach, we, firstly, try to determine the most probable context which corresponds to some concept hierarchy. After determining the appropriate concept hierarchy, we generalize the query concepts and extract the hidden concepts in order to be used in latter retrieval process. In addition, we introduce two new concept hierarchies to be used in our method along with ImageNet. Our paper is organized as follows. In section 2, we overview the related work. In section 3, we describe our proposed solution. Section 4 shows details of the experimentation and the obtained results. Finally, we draw some conclusions.

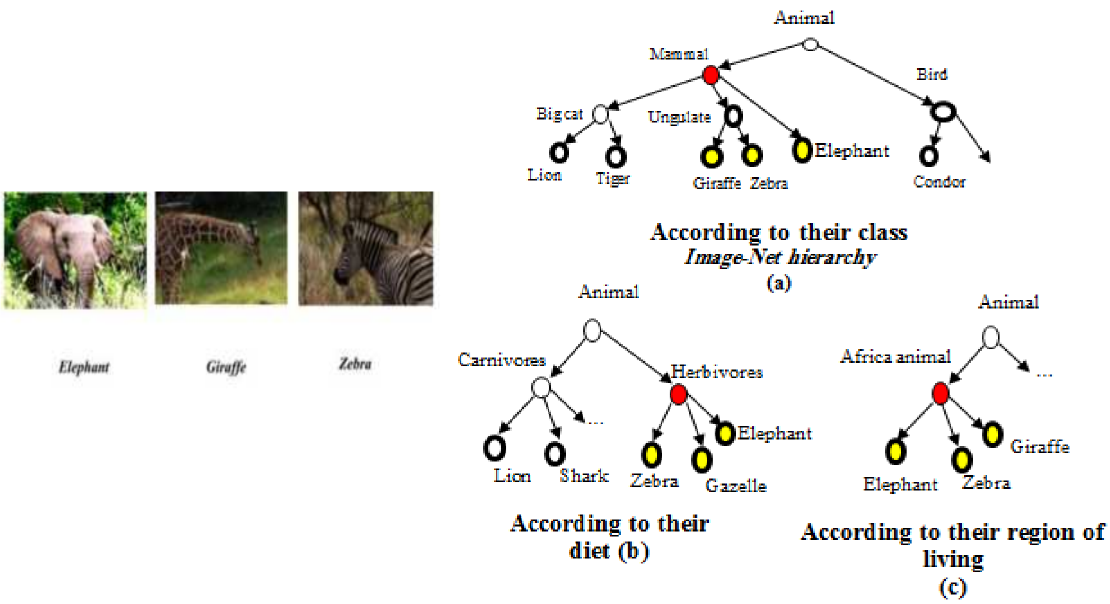


Figure 2. Concepts may be categorized in different ways based on the selected context: (a) animals categorized according to their family, (b) animals categorized according to their diet, (c) animals categorized according to their living region

2. Related Work

Many literature researches in cognitive science have attempted to develop methods that are able to simulate the performance of the human to learn novel visual concepts from positive examples. For example, in [22], the authors have tried to explain how a human child learns new words from a set of pre-provided positive examples. Humans are able to generalize complex sets of images that contain different objects (e.g., natural scene, animals, etc.) very quickly. Besides, humans are able to extract the relationship between a given set of concepts in different contexts. By exploiting concept hierarchies, numerous attempts have been made in the literature, attempting to reach human-like object generalization or categorization. Deng et al. [14], have introduced new classifiers that exploit a concept hierarchy consisting of many levels of abstraction. They have proposed a Dual Accuracy Reward Trade-off Search (DARTS) algorithm that aims to select the appropriate level of categorization in this concept hierarchy. However, these classifiers are not completely accurate in identifying leaf node classes.

Some other works [15, 16] have tried to handle the problem of finding new categories based on pre-provided sets of labeled examples. The main aim in [15] for example was how to learn a new visual category (i.e., generalization) from few positive examples. Salakhutdinov et al [16], have presented a hierarchical classification model that allows rare objects to borrow statistical strength from related objects that have many training examples. However, the former two works have tried to improve the generalization using only the leave nodes of the concept hierarchy. Thus, they did not address the issue of discovering the hidden concepts between the leave nodes, which is a key idea for visual concept learning.

Bayesian models of generalization [17-20] have been extensively used in cognitive science in order to resolve the issue of learning new words or concepts from an initial set of words or concepts. Given a concept hierarchy, Bayesian models of generalization basic idea revolves around finding the optimal degree of generalization, in this hierarchy, for any set of concepts [21]. Tenenbaum and Griffiths, [18] have referred to such an approach as ‘the size principle’ and they have shown how it could potentially explain a wide range of phenomena in category learning, generalization, and similarity judgment. Such phenomena were not previously unified under one single rational-inference. In more recent work, Xu and Tenenbaum [19] have developed a new Bayesian word-learning model. Their model appeared to be capable of mimicking human generalization judgments to create a hypothesis space for three categories (animals, vehicles, and vegetables) with few positive examples. However, their work is too hard to be extended to other categories.

Abbott et al, [20] have proposed a Bayesian-based model for automatically generating hypothesis spaces that are used for generalization. In their model, WordNet database has been used to generate the tree-structured hypothesis space for different concepts. WordNet is a database that encodes the semantic relationships between concepts as a network. On the other hand, ImageNet has been used to indicate the images corresponding to each of these concepts. Unlike the previous works, Abbott’s automatically generated hypothesis space that can be used in any category.

Recent works in visual recognition [23-26] and image retrieval [27] have used hierarchical structures that contain a high number of classes. N. Verma et al. [24] proposed a novel framework to determine the similarity rate between images. In the hierarchy, two images are considered to be similar if the distance between their annotations (i.e., concepts) is minimum and vice-versa. Jia et al. [23] have proposed a system that integrates both Bayesian models of generalization and machine vision techniques. Their main aim was to determine whether a query image is related to a concept generated from some given set of images. Likewise, they have used ImageNet database to build their hypothesis space. In addition to the high performance their system shows, it seems to be similar to human reasoning in generalization. However, all works in the context of concept generalization suffer from one major problem which is performing generalization using only one concept hierarchy. Therefore, they are restricted to only one context of generalization unlike humans. To make this latter point clearer let’s take the example illustrated in Figure 3. In Figure 3 (a), the relationship between the three images comes in terms of Family (i.e., Birds), whereas in Figure 3 (b) another kind of relationship gathers the three images, which is the diet (i.e., omnivores). Finally, in Figure 3 (c), the relationship is living region (i.e., Asia Animals).

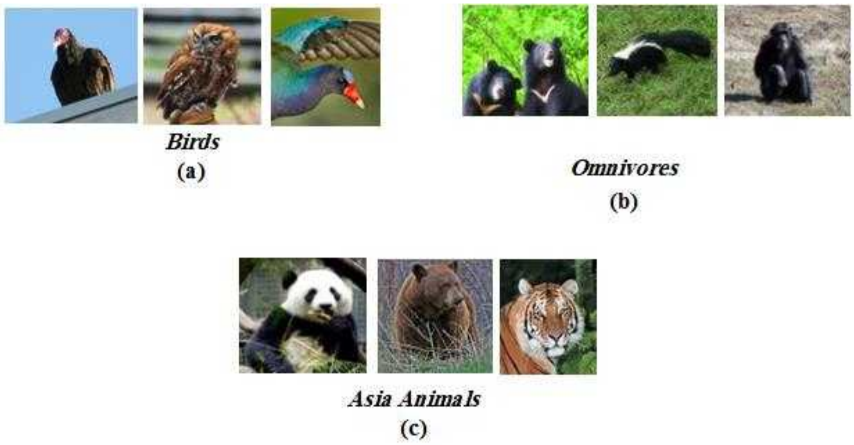


Figure 3. Examples of some generalizations in different contexts (a) generalization by family, (b) generalization by living region, (c) generalization by diet

Supposing that we have a query that contains three concepts: Elephant, Zebra, and Giraffe as we have shown previously in Figure 2. Conventional systems [23] interpret, or rather generalize, this query to the concept Mammal, which is totally correct. However, several other meaningful concepts can be inferred. These concepts, such as African animal, may be closest to the user intention than the concept Mammal. In order to remove this confusion and precisely detects the user intention, we propose to enrich the existing hierarchies with other ones; for example, adding the hierarchy that assembles concepts according to their diet and also according to their region of living. Figure 4 shows the difference between our generalization and that of a conventional system. In the next section, we introduce our method where we try to improve the generalization task by making it able to deal with multiple concept hierarchies.

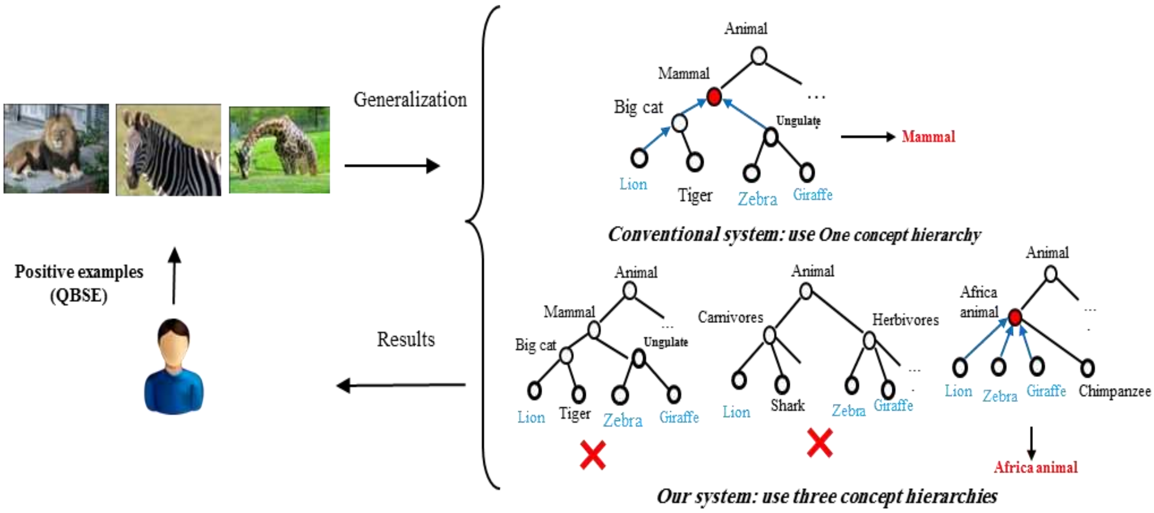


Figure 4. Illustrates the difference between our generalization scheme and a conventional system in [23]

3. Proposed Solution

In this section, we present the details of the proposed method. We start by giving details about the Bayesian Concept Learning, after that we define the concept hierarchies CHs in our framework. Then, we explain our generalization scheme, and finally we present the analysis of some example queries.

3.1. Bayesian Concept Learning

Bayesian framework for concept learning and generalization techniques [19], are particularly useful in the case where learning is performed using only a small number of positive examples. In particular, the problem can be looked to as follows: Given a set of n examples (i.e., images in our case) $X=\{x_1....x_n\}$ which can be grouped under a specific concept C as shown in Figure 3. Given a new example y, the question is: Is y a member of X or not. To answer this question, Bayesian concept learning assumes the existence of a hypothesis space H such that $H=\{h_1....h_n\}$ where the most appropriate hypothesis hi can be considered as C. Each hypothesis hi (e.g., Animal, Mammal, Bird) corresponds to one cluster in the concept hierarchies. An illustration is given in Figure 2. The Bayesian learner evaluates all the hypotheses hi using Bayes rule as follows:

$$P(h|X) \propto P(h)P(X|h) \tag{1}$$

such that *P(h|X)* is the posterior probability, *P(h)* the prior probability and *P(X|h)* the likelihood. The prior *P(h)* of the hypothesis is defined according to the Erlang distribution:

$$P(h)\alpha(|h|/\sigma^2)\exp\{-|h|/\sigma\} \tag{2}$$

where |*h*| is the size of the hypothesis *h* (number of leaf nodes) and σ parameter is the mean size of the basic level hypotheses. The likelihood is determined by the assumption of randomly sampled positive examples.

In the simplest case, each example in *X* is assumed to be independently sampled from a uniform density over the concept *C*. For n examples we then have:

$$P(X|h) = \begin{cases} \left[\frac{1}{|h|}\right]^n & \text{if } x_1, \dots, x_n \in h \\ 0 & \text{if any } x_i \notin h \end{cases} \tag{3}$$

prior work [19] focused on calculating the probability that a new object y is also a member of the concept C by averaging the predictions of all hypotheses weighted by their posterior probabilities:

$$P(y \in C|X) = \sum_{h \in H} P(y \epsilon C|h)\, P(h|X) \tag{4}$$

In our method, however, we focus on finding the hypothesis h that corresponds to the concept C. In particular, we haven't a new example y, but rather a query X. Another substantial difference is that in the previous works [19] the hypothesis space H is generated according to only one concept hierarchy CH, which is not the case in our work because we consider 3 different CHs in generating H.

To determine the most appropriate h from H, we calculate the posterior probability for each h, the appropriate h that corresponds to the concept C is the one having obtained the highest probability score (i.e., Maximum a Posteriori hypothesis h_{MAP}). The h_{MAP} is given by:

$$h_{MAP} = \underset{h \in H}{\operatorname{argmax}}\, P(X|h)P(h) \tag{5}$$

after having determined the most appropriate h, we consider the CH to which h belongs, and we omit the 2 others. Afterwards, we extract the remaining concepts that belong to C (i.e., hidden concepts). Finally, we detect images annotated by concepts contained in C and display them to the user.

3.2. Presentation of Our Concept Hierarchies

In our framework, we use three kinds of concept hierarchies to expand the scope of user understanding, where each hierarchy groups concepts according to a specific relationship. These relationships are: family, diet and living place. Next, we give details about each of them.

3.2.1. Concept Hierarchy According to Family CH_a (ImageNet hierarchy)

We use Image Net hierarchy as the first CH, we denote it by CH_a. Image Net is a large image database which is based on the WordNet hierarchy. Each concept in WordNet is described by multiple words which are called a "synonym set" or "synset". We have chosen Image Net hierarchy because it has a rich hierarchy of concepts and it assembles millions of images (about ten million images that have been manually annotated). In our work, we are interested by the part which categorizes the animals as shown in Figure 5.

3.2.2. Concept Hierarchy According to Diet CH_b

We build this type of relationship based on Wikipedia. Our hierarchy is built based on the food nature of each “synset” as shown in Figure 6. We denote the current CH by CH_b.

3.2.3. Concept Hierarchy According to Region of Living CH_c

The region of living is a synonym set within ImageNet, thus, we adopt a CH that groups concepts according to region of living as shown in Figure 7. We denote it by CH_c.

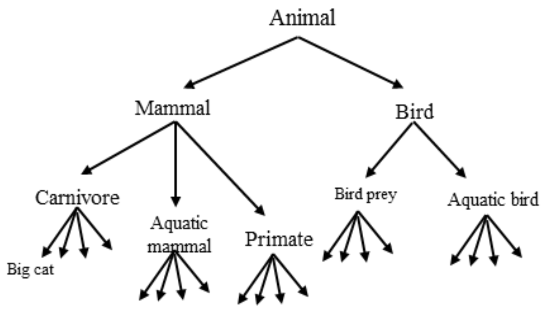


Figure 5. Illustration of the concept hierarchy CH_a

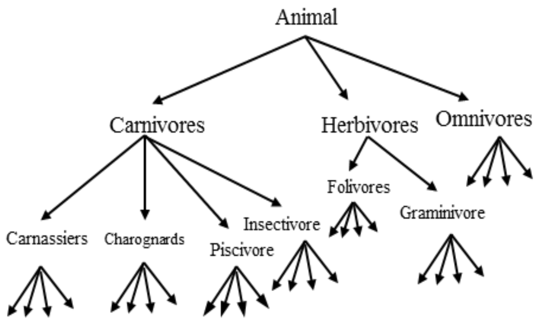


Figure 6. Illustration of the concept hierarchy CH_b

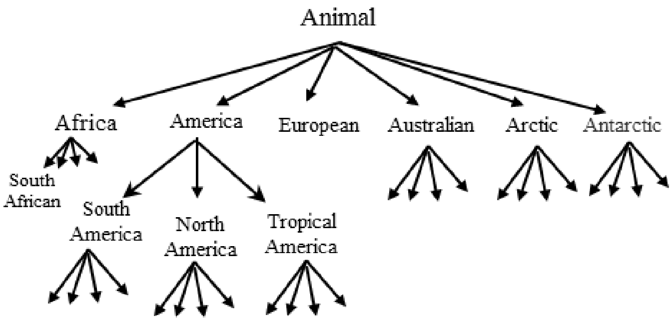


Figure 7. Illustration of the concept hierarchy CH_c

3.3. Our Generalization Scheme

Previous studies [19-21, 23] have attempted to learn concepts using a few number of positive examples. However, these studies have focused only on choosing the appropriate generalization level in a single concept hierarchy. This, in fact, can yield minor or completely irrelevant results. To overcome this problem and improve the results of the engine we use three kinds of CH in the generalization. The details of our generalization scheme are illustrated in Figure 8. There are six main steps to generalize the query which are described as follows.

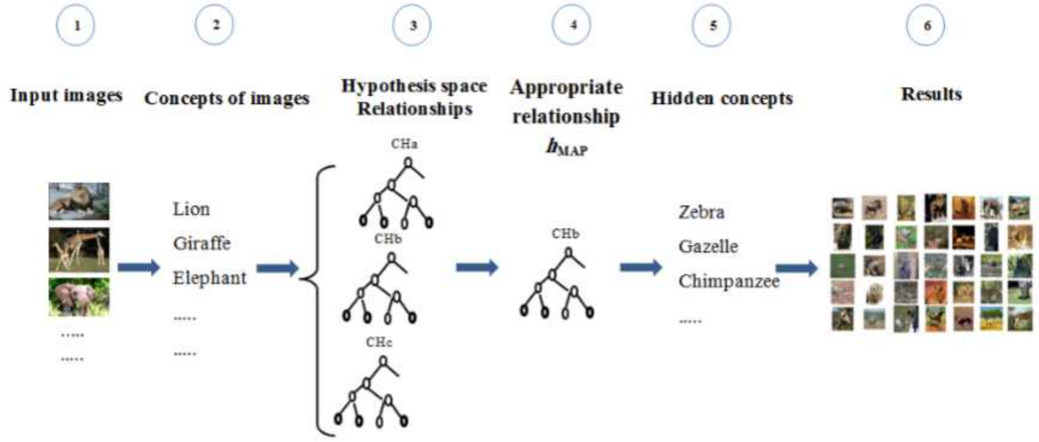


Figure 8. Illustrates the main steps of our generalization scheme

3.3.1. Input Images (Formulation Query)

Our system shows the user some images from dataset (user interface). The user has to select some images example (2-5 images) represent his needs to formulate query, as shown in Figure 9.

Figure 9. User interface of our system

3.3.2. Concepts of Images

In the dataset each image annotated with a concept, we use Image Net dataset this collection annotated from Word Net. After the user formulates his query, our system has to extract the concepts of each image in the query of those concepts which we called concept query as shown in Figure 10.

Figure 10. Images and their corresponding concepts

3.3.3. Hypothesis Space (Relationships)

After finding concepts of the query our system begins searching for all the relationships between concepts query. All relationships in all kinds of concepts hierarchy are called hypothesis space.

3.3.4. Finding the Appropriate Relationship (h_{MAP})

After creating the hypothesis space our system has to find the appropriate relationship gathering those concepts. The max a posterior represents this relationship.

3.3.5. Hidden Concepts

Hidden concepts are the concepts that are linked with concepts query by the appropriate relationships selected in the concept hierarchy.

3.3.6. Results

Finally our system searches all images annotated with concept query and hidden concepts, and shows results to the user. For the sake of clarity, let us illustrate this by a simple example. Suppose that we have 3 positive examples i.e., X= {Lion, Giraffe, Zebra}, the hypotheses we can consider as candidate to be the concept C are: Animal, Mammal, Africa animal. We calculate the posterior probability of each hypothesis according to (1). The hypothesis that obtains the highest score is considered as C.

After determining C, we give back the user the images annotated with all the concepts of leaf nodes under C i.e., we consider the concepts contained in the query together with those which are not contained (i.e., hidden concepts). Indeed, this could help in improving the quality of retrieval results. The steps of our algorithm are summarized in Algorithm 1.

Algorithm 1: Generalization of query
Begin
1: INPUT: X = {x1, x2, …xn}
2: Compute posterior probability P(h|X) of all hypotheses h in CHa, CHb and CHc according to (1) :
3: Find the Max a posteriori hMAP according to (5):
4: Select appropriate CH and the C.
5: Find Hidden Ci (The concepts under C and which didn't appear in the query)
6: OUTPUT: Result of images Ii annotated by all leaf nodes under the concept C.
End

3.4. Examples of Generalizing Queries in Our Formwork

Now, let us explain, by examples, how our proposed method works. In our case, the system presents a sample of images to the user, then the user selects images supposed to be similar to what he is looking for. Our proposed approach uses, thereafter, the annotations assigned with those images in order to discover the hidden relationship between the concepts contained in the query. For example, the generalizing of the queries Q1, Q2 and Q3 shown in Figures 11-13 are respectively demonstrated in Table 1, Table 2, and Table 3.

Figure 11. Illustration of the Query Q1

Figure 12. Illustration of the Query Q2

Figure 13. Illustration of the Query Q3

3.4.1. Generalizing Query Q1

Concepts Query: X= {leopard, ocelot, Tiger, Cougar}
n=3
Hypothesis h

- Animal
- Mammal
- Feline
- Carnivores

Size of each hypothesis |h|

- |Animal|=100
- |Mammal|=86
- | Feline |=5
- |Carnivores|=34

Maximum a Posteriori hypothesis h_{MAP}: Through the results given by Table 1 the Maximum a Posteriori hypothesis of the Query 1 is h_{MAP} = Feline.
Hidden concepts: Lion, Jaguar, Panther, Lynx.

Table 1. Generalization of the Query Q1 in each Concept Hierarchy

Concept hierarchy	Hypothesis h	Posterior probability P(h|X)
	Mammal	0.56
CHa	Animal	0.23
	Feline	0.87
CHb	Carnivores	0.55
CHc	No relation matched	/

3.4.2. Generalizing Query Q2
Concepts Query: X= {Gazelle, Elk, Two-toed Sloth, Elephant, Gnu}
n=5
Hypothesis *h*
- Animal
- Mammal
- Herbivores

Size of each hypothesis |*h*|
- |Animal|=100
- |Mammal|=86
- |Herbivores|=15

Maximum a Posteriori hypothesis h_{MAP} :Through the results given by Table 2 the Maximum a Posteriori hypothesis of the Query 2 is h_{MAP}= Herbivores
Hidden concepts: Giraffe, Hippopotamus, Gorilla, Koala, Caribou, Elk.

Table 2. Generalization of the Query Q2 in each Concept Hierarchy

Concept hierarchy	Hypothesis h	Posterior probability P(h|X)
CHa	Mammal	0.56
	Animal	0.23
CHb	Herbivores	0.86
CHc	No relation matched	/

3.4.3. Generalizing Query Q3
Concepts Query: X= {Lion, Zebra, Elephant}
n=3
Hypothesis *h* (all nodes that gather the concepts)
- Animal
- Mammal
- Africa animals

Size of each hypothesis |*h*| (number of leaf nodes/son nodes)
- |Animal|=100
- |Mammal|=86
- |Africa animals|=15

Maximum a Posteriori hypothesis h_{MAP}: Through the results given by Table 3 the Maximum a Posteriori hypothesis of the Query 3 is h_{MAP}= Africa animals.
Hidden concepts: Lion, Gnu, Gorilla, Chimpanzee.

Table 3. Generalization of the Query Q3 in each Concept Hierarchy

Concept hierarchy	Hypothesis h	Posterior probability P(h|X)
CHa	Mammal	0.56
	Animal	0.23
CHb	No relation matched	/
CHc	Africa animals	0.96

The obtained results of Q1, Q2 and Q3 illustrated in Figure 14. Through the above examples, we find that the algorithm is capable to generalize in any concept hierarchy and find the appropriate relationship between concepts whatever the kind of the relationship.

Figure 14. Results of some example queries: (a) results: generalization in CHa with relationship (Feline); (b) Results: generalization in CHb with relationship (Herbivores animals); (c) results: generalization in CHc with relationship (Africa animals)

4. Experiments
4.1. Dataset
In order to demonstrate the effectiveness of the proposed method, we carried out our experiments on ImageNet dataset [28]. It contains 14,197,122 images in 21,841 categories indexed according to the hierarchy of WordNet [29]. A category in ImageNet corresponds to a synonym set (synset) in WordNet. ImageNet covers a subset of the nouns of WordNet, organized in 12 high level categories, (e.g. animal, Plant, instrumentality...). In the present work, we focus on the animals category. We select, 100 synsets of animals to create our dataset which is made up of 111,135 images (we choose those synset to formulate three concept hierarchy each synset have a relationship in all concept hierarchy). Then, these images are organized according to three hierarchies, one is that of ImageNet According to family CHa and we add two others which are (According to diet CH_b and According to region of living CH_c) as described above.

4.2. Scenario of Experiments
To perform experiments, we have invited 20 participants. Each participant supplies a query made up of 2-5 positive images. This test allows us to truly investigate the ability of our method in meeting human thought and intentions. We compare the performance of the proposed method with another one from the state of the art [23].

4.3. Evaluation Metrics
To evaluate the performance of our method, we use the Precision and Recall metrics, which are given by:

$$\text{Precision} = \frac{Number\ of\ relevant\ retrieved\ images}{Number\ of\ \text{Retrievaed}\ images} \tag{6}$$

$$\text{Recall} = \frac{Number\ of\ relevant\ retrieved\ images}{Number\ of\ relevant\ images\ in\ dataset} \tag{7}$$

to provide an objective evaluation of our algorithm, we launched several queries and measured the precision and Recall of each query. The obtained results are shown in Figure 15 and Figure 16.

4.4. Results and Discussion
Figure 15 shows the number of relevant images. The results show that the proposed method significantly outperforms the method in [23]. This may be attributed to the fact that we haven’t restricted ourselves to only one concept hierarchy and we have generalized the query using 3 different hierarchies.

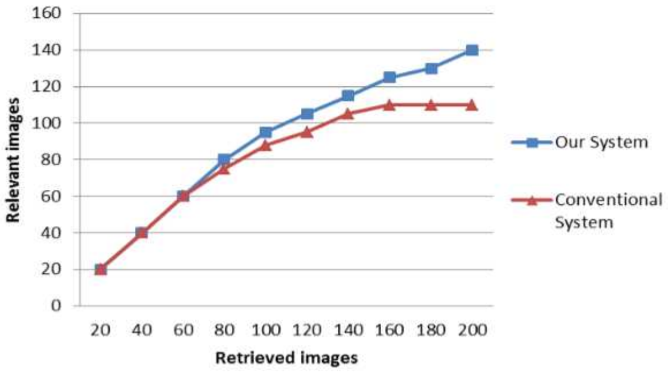
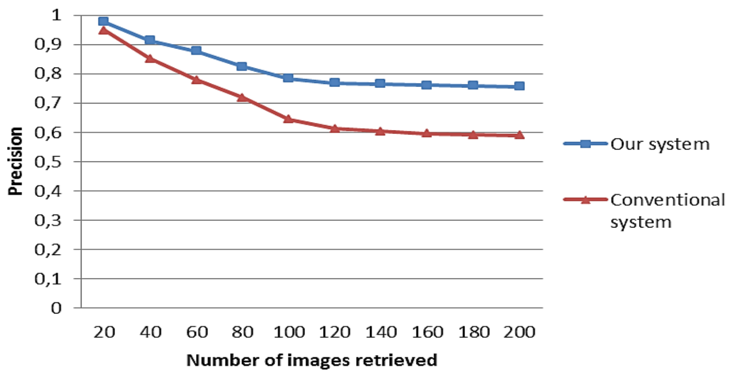
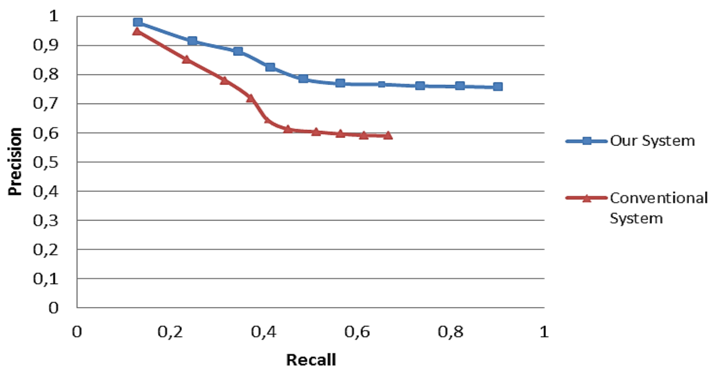


Figure 15. The number of relevant images in our system and conventional system

To confirm the strength of the proposed method, we report the average precision-scope (Figure 16(a)) and precision-recall (Figure 16(b)) curves for both our method and the method of [23]. We can see that the proposed method overcomes the method in [23]. In summary, from the experiments we have conducted, we note that the proposed method is capable to detect the appropriate generalization level in the different hierarchies (CH_a, CH_b and CH_c). This is unlike the existing methods which are limited to a single hierarchy. Therefore, our method is capable to understand the user intention and retrieve the targeted images by the user. Experimental results have demonstrated that our method yields better results than those based on a single hierarchy.



(a)



(b)

Figure 16. The average precision-scope and precision-recall curves of conventional system and our system

5. Conclusion

In this work, we have proposed a new method for image retrieval based on query interpretation. Specifically, we have used multiple concept hierarchies to generalize the query and understand the user intention. Our system is capable to choose the appropriate generalization level among several concept hierarchies and give the best results in image retrieval. Experiments show that the proposed method significantly outperforms the conventional approach. In a future work, one could introduce a pre-processing step which consists in eliminating outliers from the query before performing the retrieval. Considering such an aspect could probably lead to improve the retrieval results.

References

[1] MS Lew, N Sebe, C Djeraba, et al. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*. 2006; 2(1): 1-19.

[2] Yang L, Jin R, Mummert L, et al. A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2010; 32(1): 30-44.
[3] Kherfi ML, Ziou D, Bernardi A. *Learning from negative example in relevance feedback for content-based image retrieval.* International Conference on Pattern Recognition ICPR. 2002: 933-936.
[4] Kherfi ML, Ziou D, Bernardi A. Combining positive and negative examples in relevance feedback for content-based image retrieval. *Journal of Visual Communication and Image Representation.* 2003; 14(4): 428-457.
[5] Chen L, Xu D, Tsang IW, Luo J. *Tag-based web photo retrieval improved by batch mode re-tagging*. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2010: 3440-3446.
[6] Liu Y, Xu D, Tsang IW, Luo J. Textual query of personal photos facilitated by large-scale web data. *IEEE transactions on pattern analysis and machine intelligence*. 2011; 33(5): 1022-36.
[7] Debbagh F, Kherfi ML, Babahenini MC. A semantic relatedness-based solution for reducing missing problem in TBIR. *International Journal of Signal and Imaging Systems Engineering*. 2017; 10(3): 146-156.
[8] Kherfi ML. Review of human-computer interaction issues in image retrieval. In: Pinder S. *Editor*. Advances in Human Computer Interaction. IntechOpen. 2008: 215-240.
[9] Khaldi B, Kherfi ML. Modified integrative color intensity co-occurrence matrix for texture image representation. *Journal of Electronic Imaging*. 2016; 25(5): 053007.
[10] Wu L, Hoi SC, Jin R, Zhu J, Yu N. *Distance metric learning from uncertain side information with application to automated photo tagging*. Proceedings of the 17th ACM international conference on Multimedia. 2009: 135-144.
[11] Lowe DG. *Object recognition from local scale-invariant features*. Proc. of the International Conference on Computer Vision. 1999; 99(2): 1150-1157.
[12] Aiadi O, Kherfi ML. A new method for automatic date fruit classification. *International Journal of Computational Vision and Robotics*. 2017; 7(6): 692-711.
[13] Rasiwasia N, Moreno PJ, Vasconcelos N. Bridging the gap: Query by semantic example. *IEEE Transactions on Multimedia*. 2007; 9(5) 923-938.
[14] Deng J, Krause J, Berg AC, Fei-Fei L. *Hedging your bets: Optimizing accuracy-specificity trade-offs in large scale visual recognition*. 2012 IEEE Conference on Computer Vision and Pattern Recognition. 2012: 3450-3457.
[15] Quattoni A, Collins M, Darrell T. *Transfer learning for image classification with sparse prototype representations*. 2008 IEEE Conference on Computer Vision and Pattern Recognition. 2008: 1-8.
[16] Salakhutdinov R, Torralba A, Tenenbaum J. *Learning to share visual appearance for multiclass object detection*. CVPR 2011. 2011: 1481-1488.
[17] Tenenbaum JB. Rules and similarity in concept learning. *Advances in neural information processing systems*. 2000: 59-65.
[18] Tenenbaum JB, Griffiths TL. Generalization, similarity, and Bayesian inference. *Behavioral and brain sciences*. 2001; 24(4): 629-640.
[19] Xu F, Tenenbaum JB. Word learning as Bayesian inference. *Psychological review*. 2007; 114(2): 245.
[20] Abbott J, Austerweil J, Griffiths T. *Constructing a hypothesis space from the Web for large-scale Bayesian word learning*. Proceedings of the Annual Meeting of the Cognitive Science Society. 2012.
[21] Korichi M, Kherfi ML, Batouche M, Bouanane K. Extended Bayesian generalization model for understanding user's intention in semantics based images retrieval. *Multimedia Tools and Applications*. 2018; 77(23): 31115-31138.
[22] Carey S. The child as word learner. 1978.
[23] Jia Y, Abbott JT, Austerweil JL, et al. *Visual concept learning: Combining machine vision and bayesian generalization on concept hierarchies*. Advances in Neural Information Processing Systems. 2013: 1842-1850.
[24] Verma N, Mahajan D, Sellamanickam S, Nair V. *Learning hierarchical similarity metrics*. 2012 IEEE conference on computer vision and pattern recognition. 2012: 2280-2287.
[25] Yan Z, Zhang H, Piramuthu R, et al. *HD-CNN: hierarchical deep convolutional neural networks for large scale visual recognition*. Proceedings of the IEEE international conference on computer vision 2015: 2740-2748.
[26] Glick J, Miller K. Insect Classification With Heirarchical Deep Convolutional Neural Networks. 2016.
[27] Deng J, Berg AC, Fei-Fei L. *Hierarchical semantic indexing for large scale image retrieval*. CVPR 2011. 2011: 785-792.
[28] Deng J, Dong W, Socher R, et al. *Imagenet: A large-scale hierarchical image database*. 2009 IEEE conference on computer vision and pattern recognition. 2009: 248-255.
[29] Fellbaum C. WordNet. An electronic lexical database. Massachusetts. Cambridge: MIT Press. 1998.