

A Literature Review of Image Retrieval based on Semantic Concept

Alaa M. Riad
Vice Dean of Students Affairs
Faculty of Computers and
Information Sciences, Mansoura
University, Mansoura, Egypt

Hamdy.K. Elminir
Head of electronic and
communication Dept, Misr Higher
Institute of Engineering and
Technology, Mansoura, Egypt.

Sameh Abd-Elghany
Faculty of Computers and
information sciences,
Mansoura University,
Mansoura, Egypt.

ABSTRACT

This paper attempts to provide a comprehensive review and characterize the problem of the semantic gap that is the key problem of content-based image retrieval and the current attempts in high-level semantic-based image retrieval being made to bridge it. Major recent publications are included in this review covering different aspects of the research in the area of high-level semantic features. In this paper the different methods of image retrieval systems are described and major categories of the state-of-the-art techniques in narrowing down the 'semantic gap' are presented. Finally, based on existing technologies and the demand from real-world applications, a few promising future research directions are suggested.

General Terms

Image Processing, Computer Vision.

Keywords

Semantic Gap, Image Retrieval, Automatic Annotation and Ontology.

1. INTRODUCTION

The World Wide Web has become one of the most important sources of information due to the fast development of internet technology. Search engines are the most powerful resources for finding visual content (e.g., images, videos, etc) from World-Wide Web, These search engines use the surrounding text near the image for describing the content of an image and rely on text retrieval techniques for searching particular images [1]. However, there are two significant drawbacks of such engines; (a) when the surrounding words are ambiguous or even irrelevant to the image; search results using this method usually contain many irrelevant images. (b) The retrieval of images will be ineffective when different languages are used in the description of the images if this image collection is to be shared globally around the world. It is difficult to map semantically equivalent words across different languages [2, 3].

Content-based image retrieval (CBIR) came into being to solve the problems inherited by text based image retrieval [4]. Under this technique, various low-level visual features are automatically extracted from each image in the database and image retrieval is formulated as searching for the best match to the features that are extracted from the query image.

In general, the bottleneck to the efficiency of CBIR is the

semantic gap between the high level image interpretations of the users and the low level image features stored in the database for indexing and querying. In other words, there is a difference between what image features can distinguish and what people perceives from the image [5]. In order to understand the information of the color image and enhance accuracy of CBIR, research focus has been shifted from designing sophisticated low-level feature extraction algorithms to reducing the 'semantic gap' between the visual features and the richness of human semantics [6]. This paper shows that current techniques that presented to reduce the 'semantic gap' in the last decade.

The remainder of this paper is organized as follows. Section 2 explains in details each method of image retrieval systems. In Section 3, the semantic gap problem and major categories of the techniques used in narrowing down the semantic gap, the current attempts in high-level semantic-based image retrieval being made to bridge the semantic gap are presented. Section 4 presents the key future directions in CBIR. Section 5 concludes this research.

2. IMAGE RETRIEVAL APPROACHES

With the development of the Internet, and the availability of image capturing devices such as digital cameras, huge amounts of images are being created everyday in different areas including remote sensing, fashion, crime prevention, publishing, medicine, architecture, etc [7]. For this purpose, the need for the development of efficient and effective methodologies to manage large image databases for retrieval is urgent so many general-purpose image retrieval systems have been developed. There are three methods for image retrieval: text-based method, content-based method and hybrid method. This section explains in details each method.

2.1 Text-Based Image Retrieval (TBIR)

TBIR is currently used in almost all general-purpose web image retrieval systems today. This approach uses the text associated with an image to determine what the image contains. This text can be text surrounding the image, the image's filename, a hyperlink leading to the image, an annotation to the image, or any other piece of text that can be associated with the image [8]. Google, Yahoo Image Search engines are examples of systems using this approach. These search engines having indexed over one billion images [9]. Although these search engines are fast and robust, they sometimes fail to retrieve relevant images as in figure 1, this for many reasons,

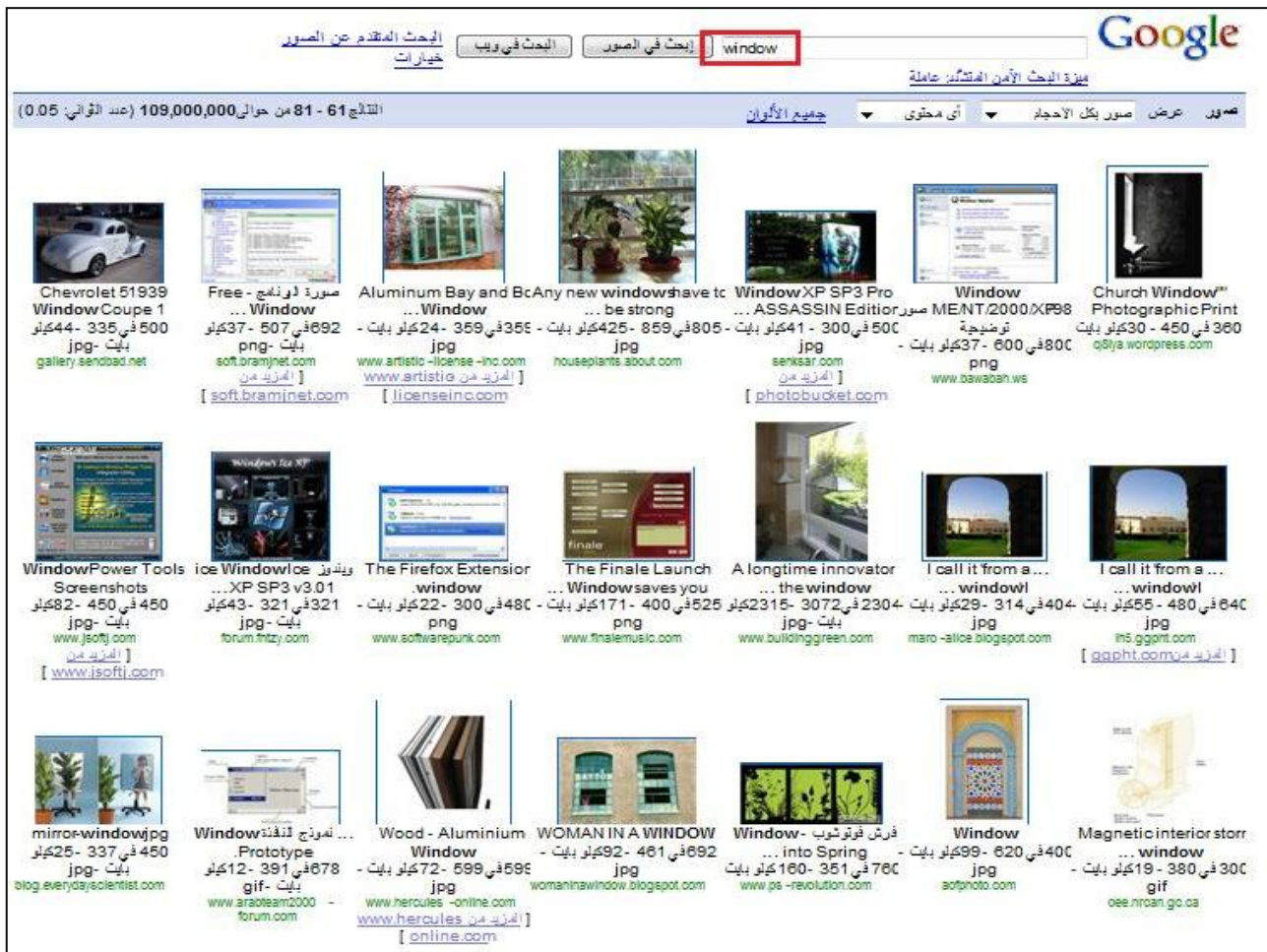


Fig.1: Google image search result for window keyword

Firstly, there are too many irrelevant words in the surrounding textual descriptions, which results in low image search precision rate. Secondly, the surrounding text does not seem to fully describe the semantic content of Web images, which results in low image search recall rate [10]. The third problem is polysemy problem[11] (same word can be used to refer to more than one object). Due to the query polysemy, the results searcher will fail to find images tagged in Chinese, and a Dutch searcher will fail to find images tagged in English. This means the query must match the language of the text associated with the images [12].

2.2 Content-Based Image Retrieval

CBIR is a technique for image retrieval based on extracting and indexing of the low-level features of images, such as color, texture and shape, to support visual queries in an intuitive way, and index images automatically with content descriptors [13].

Most of the existing CBIR systems consider each image as a whole; however, a single image can include multiple regions/objects with completely different semantic meanings. A user is often interested in only one particular region of the query image instead of the image as a whole. Therefore, rather than viewing each image as a whole, it is more reasonable to view it as a set of regions. The features employed by the majority of Image Retrieval systems include color, texture, shape and spatial layout. Such features are apparently not

effective for CBIR, if they are extracted from a whole image, because they suffer from the differing backgrounds, overlaps, occlusion and cluttering in different images and they do not have adequate ability to capture important properties of objects [14, 15], as a result most popular approaches in recent years is to change the focus from the global content description of images into the local content description by regions or even the objects in images. RBIR is a promising extension of the classical CBIR: rather than deploying global features over the entire content, RBIR systems partition an image into a number of homogenous regions and extract local features for each region then features of regions are used to represent and index images in RBIR. For RBIR, The user supplies a query object by selecting a region of a query image and then the corresponding similarity measure is computed between features of region in the query and a set of features of segmented regions in features database and the system returns a ranked list of images that contain the same object.

2.3 Hybrid Method for Image Retrieval

A recent trend for image search is to fuse the two basic modalities of Web images, i.e., Textual context (usually represented by keywords) and visual features for retrieval [16]. It is suggested a joint use existing Textual context and visual features can provide a better retrieval results [10]. The simplest approach for this method is based on counting the frequency-of-occurrence of words for automatic indexing. This simple approach can be extended by giving more weights

to the words which occur in the alt or src tag of the image or which can occur inside the head tag or any other important tags of the HTML document.

The second approach takes a different stand and treats images and texts as equivalent data. It attempts to discover the correlation between visual features and textual words on an unsupervised basis, by estimating the joint distribution of features and words and posing annotation as statistical inference in a graphical model. For example image retrieval system based on decision trees and rule induction was presented in [17] to annotate web image using combination of image feature and metadata, while in [18], a system that automatically integrate the keyword and visual features for web image retrieval by using association rule mining technology. These approaches usually learn the keywords correlations according to the appearance of keywords in the training set or lexicon, and the correlation may not reflect the real correlation for annotating Web images or semantic meaning of keywords such as synonym [19]. As a result, pure combination of traditional text-based and content-based approaches is not sufficient for dealing with the problem of image retrieval on the Web.

3. SEMANTIC GAP PROBLEM

During retrieving images in CBIR system, users feed the retrieval system example image, which are internally transformed into feature vectors and matched against those in the feature vector database [20]. Although, the extracted visual features are natural and objective, there is a significant gap between the high-level concepts (which human perceives) and the low-level features (which are used in describing images)[21].

As example, a red flower may be regarded as the same as a rising sun, and a fish the same as an airplane etc. This means that traditional CBIR system that only depends on extraction and comparing of primitive features has no understanding of the image's semantic contents and cannot meet the users' needs due to the 'semantic gap' between low-level features and the richness of human semantics [22,23]. How to bridge the semantic gap is currently a major research problem in CBIR [24].

Recent research has suggested different approaches to bridging the semantic gap. This survey shows these approaches. The mainly approaches can be categorized into : (1) Relevance feedback (RF), which needs the user to interact with the system for feedback after the initial retrieval, (2) Automatic Image Annotation. The next subsections explain in details each approach.

3.1 Relevance Feedback (RF)

RF is an on-line processing which taken to reduce the gap between the high level semantic concepts and the low-level image features [25]. It refers to the feedback from a user on specific images result regarding their relevance to a target image, the refined query is re-evaluated in each iteration [26]. Basically, the image retrieval process with relevance feedback is comprised of four steps[27], (i) showing a number of retrieved images to the user; (ii) user indication of relevant and non-relevant images; (iii) learning the user needs by taking into account his/her feedbacks; (iv) and selecting a new set of images to be shown. This procedure is repeated until a satisfactory result is reached.

Generally, though efforts in relevance feedback (RF) have been made to bridge this gap [28,29], the technique is a variation of CBIR that involves multiple interactions with a

user at search time, this means lots of time to reach a satisfactory results. The labeled samples provided by the user x are limited, and such small training data set will result in weak classification of database images (as relevant/irrelevant) [30]. The performance of the RF method in Web image retrieval strongly relies on good quality of the top-N images that are used for feedback. When there are only a few relevant images are returned, this method might be ineffective [15]. On the other hand, if the user is searching for a specific object that cannot be sufficiently represented by combinations of available feature vectors, these relevance feedback systems will not return many relevant results even with a large number of user feedbacks [31].

3.2 Automatic Image Annotation

Manual image annotation is a time consuming task and as such it is particularly difficult to be performed on large volumes of content [32]. There are many image annotation tools available but human input is still needed to supervise the process. So, there should be a way to minimize the human input by making the annotation process fully automatic [33]. In Automatic image annotation images are automatically classified into a set of pre-defined categories (keywords). Low-level features of the training images are extracted. Then, classifiers are constructed with low-level features to give the class decision. Lastly, the trained classifiers are used to classify new instances and annotate unlabelled images automatically [34]. Automatic image annotation plays an important role in bridging the semantic gap between low-level features and high-level semantic contents in image access [35].

Many methods have been proposed for automatic image annotation, which can be roughly categorized into two groups: keyword-based methods and ontology-based methods [36]. Keywords-based methods: Arbitrarily chosen keywords from controlled vocabularies, i.e. restricted vocabularies defined in advance, are used to describe the images.

Ontology-based methods: Ontology is a way of describing concepts and their relationships into hierarchical categories [37]. This is similar to classification by keywords, but the fact that the keywords belong to a hierarchy enriches the annotations. For example, it can easily be found out that a car or bus is a subclass of the class land vehicle, while car and bus have a disjoint relationship. The two approaches are discussed in the following subsections.

3.2.1 Keywords-based Annotation

This method defines automatic annotation as a traditional supervised classification problem, which treats each word (or semantic category) as an independent class and creates a different class model for every word [38]. This approach separates the textual components from the visual components, computing similarity at the visual level using support vector machine (SVM) classifiers that classify images by visual features and textual information as in [39,40], or using decision tree (DT) and SVM [10]. Then it annotates a new image by propagating the corresponding class words. The state-of-the-art techniques can be roughly categorized into two different for image annotation; region level and image level. Annotation at the image level [41,42], involves finding words/labels that best describe the entire image. Region level annotation [43,44], requires segmenting the image into objects, regions, or blobs, and annotating each region. Region level annotation allows for direct object searching and in most cases, produces higher retrieval accuracy since what is inside the image can be better represented [45].

This method is useful especially to a user who knows what keywords are used to index the images and therefore can easily formulate queries [46], but naive user does not necessarily know the vocabulary which has been used to annotate an image collection. This makes searching by text input more difficult. Forcing the user to choose from an on-screen list of keywords is a solution to this problem, but this makes the search task more frustrating if the number of keywords is large. Moreover, The way of applying keyword annotation to the image resource has low capability to analyze semantic relations among keywords, such as synonym, hypernymy/hyponymy (ISA) and meronymy (PARTOF) relationships[47]. Taking the topic of images as example, it is nearly impossible to include all the synonyms of the topic keyword in the annotation of every image [48]. This will lead to the consequence that if the user input a keyword which has the exact same meaning to the topic keyword of some certain images in the index database but in different terms, those images are not able to be retrieved.

3.2.2 *Ontology-based Annotation*

Ontology is a specification of a conceptualization. Ontology defines a set of representational terms called concepts; each concept has three basic components: terms, attributes and relations. Terms are the names used to refer to a specific concept, and can include a set of synonyms that specify the same concepts. Attributes are features of a concept that describe the concept in more detail. Finally relations are used to represent relationships among different concepts and to provide a general structure to the ontology[37]. Adding a hierarchical structure to a collection of keywords produces a taxonomy, which is an ontology as it encodes the relationship “is a” (a dog is an animal). Ontology-based image retrieval is an effective approach to bridge the semantic gap because it is more focused on capturing semantic content which has the potential to satisfy user requirements better. The major difficulty in the ontology based approach is the extra work needed in creating the ontology and the detailed annotations.

A hybrid model which Ontology based reasoning Bayesian Network model for relevance ranking was presented in [49].In the retrieval method, the system offers recommendations of search keywords which are semantically related to the user’s query so that it can assist the users to navigate around the relevant images. But this system is based on simple and single keyword query and cannot answer complex query.

Natural Language Processing (NLP) framework is used in [50], to extract concepts from image captions. Next, an ontology-based framework is deployed in order to resolve natural language ambiguities.

While semantically rich ontology addresses the need for complete descriptions of image retrieval and improves the precision of retrieval[51]. However, the lack of text information which affects the performance of keyword approach is still a problem in text ontology approach. In other words, although these approaches could address the mismatch problem between the terms, it is not the most suitable for image retrieval system because it does not consider the features of the image content[52]. Ontology works better with the combination of image visual features.

A system for combining Wikipedia-based high-level text description and visual feature was presented in [53]. The system generates the multi-modality ontology by extracting a set of relevant concepts to the target domain with their associated semantic relations from Wikipedia. Then, other concepts related to the low-level features of the images,

together with their relations are obtained from a set of training images.

The system in [54], presents a hierarchical image semantic structure, which captures image semantics in an ontology tree and visual features in a set of specific semantic domains. The query algorithm works in two phases. First, the ontology is used for quickly locating the relevant semantic domains. Secondly, a query makes visual similarity comparison in the reduced domains.

A simple lexical ontology derived from WordNet and CBIR techniques are combined to automatically associate picture classes to concepts in [9].

A general domains for Combining CBIR and ontology was presented in [55].The limitation of these systems is that no segmentation was performed to extract different objects from image; this lead to the user can not query a region of the image. A new method for semantic annotation algorithm based on image regional object ontology was proposed in [56,57,58]. This system firstly constructs the regional objects ontology, the ontology logical reasoning applies to the image semantic annotation, and an image overall semantic annotation method based on regional object ontology. This works not support spatial relationship query between objects.

An extension of the ontology driven image retrieval system was proposed in [59]. this system provides the possibility to query the system in several languages and obtain consistent sets of results irrespective to the language of the query. It proposed a partial translation of several WordNet to a multilingual ontology.

4. LIMITATIONS OF CURRENT ONTOLOGY BASED IMAGE RETRIEVAL SYSTEMS

Firstly, most of current systems use WordNet as a lexical resource to build concept ontology. However, as the most important feature of WordNet is to group words into synset (set of synonyms) and connect them through hypernymy/hyponymy (ISA) and meronymy (PARTOF) relationships, the ontology generated from WordNet can only come with concepts and their hierarchical relationships [53]. However, There are many other possible relationships between words, for instance between a verb and its typical object (e.g. 'drive' and 'car'), or between an entity and a part of that entity (e.g. 'car' and 'wheel'). These relations are not covered by WordNet. This means that there are many relationships between words that will seem natural to humans that will not be discovered using WordNet. Therefore, ontology with rich relations is more meaningful.

Secondly, these systems do not implies the using of ontology reasoning with the goal of searching and retrieving complex images based on the complex query formulated in a symbolic language. For example, in a traditional system if a user queries “beach without people” the text system looks for the words “beach” and “people” and does not understand the meaning of “without”.

Thirdly, the spatial context of identified regions, objects, scenes and faces is not encoded within the index. This means theses systems cannot return semantically accurate results for queries involving spatial prepositions such as “next to”, “on”, “beside” “against” etc. In addition to querying properties which are in the “top” “bottom” “center” “left” or “right” of an image.

Finally, critical stage of image retrieval concerns the great linguistic diversity that exists on the Web. Yet given this multiplicity, one would expect a similar variety in the languages of texts associated with the images. When using a textual query, the image searcher faces a double challenge when trying to retrieve images: the query must correspond to the text associated with the images and the language of the query must also match the language of the text associated with the images.

5. FUTURE DIRECTIONS FOR IMAGE RETRIEVAL SYSTEMS

Currently, most image retrieval systems use either purely visual features or textual metadata associated with images. They have advantages and disadvantages respectively. To overcome their drawbacks and improve the performance without sacrificing the efficiency, the new web image retrieval systems should pay a great attention for these features:

5.1 Automatic Image Segmentation and Classification

As the image passes through the segmentation and classification process the system automatically identifies regions, scenes, objects, facial aspects and spatial positions of those regions, objects and faces within the image. As part of this process the attributes within the image are given statistical relevancy based on how they typify the concept. Automatic classification of the content of an image lends itself to many applications, combining this with existing metadata allows users to search more accurately, for many more things in an image, in addition to making images with poor or non-existent keywords visible for the first time at a dramatically reduced cost compared with manually adding keywords.

5.2 Semantic Image Annotation

The objective of semantic annotation is to describe the semantic content in images and retrieval queries. Semantic annotation requires some understanding of the semantic meaning in images and retrieval query, and standardization of representation of images. Based on the semantic annotation of images and retrieval queries, we can compare semantic similarity between images and a retrieval query. At present, semantic annotation is implemented by some markup language such as XML based on a shared ontology definition.

5.3 Semantic Image Retrieval

This feature of the system is a the retrieval architecture, which understands the syntax and meaning of a users query and uses a linguistic ontology to translate this into a query against the visual ontology index and any metadata or keywords associated with the image. The retrieval system takes textual queries and reasons about them through understanding their syntax and meaning. For example, in a traditional system if a user queries “beach without people” the text system looks for the words “beach” and “people” and does not understand the meaning of “without”.

5.4 Ontology Reasoning

Ontological reasoning is the cornerstone of the semantic web, a vision of a future where machines are able to reason about various aspects of available information to produce more comprehensive and semantically relevant results to search queries. Rather than simply matching keywords, the web of the future will make use of ontology to understand the

relationship between disparate pieces of information in order to more accurately analyze and retrieve information.

5.5 Multi-Object Search

Object-based image retrieval has recently become an important research issue in retrieving images on the basis of the semantics of images. However, most existing object-based image retrieval systems are based on single object matching, with its main limitation being that one individual image region (object) can hardly represent the user’s retrieval target especially when more than one object of interest is involved in the user query. An important aspect of the system is that users are allowed to formulate a query based on multi objects of an image [60].

5.6 Spatial Search

As part of the classification process, the spatial context of identified regions, objects, scenes and faces is encoded within the index. This means the system can return semantically accurate results for queries involving spatial prepositions such as “with”, “next to”, “on”, “beside” “against” etc. In addition to querying properties which are in the “top” “bottom” “center” “left” or “right” of an image. In other word, enable the user to search for multiple Object criteria as the same as text based information Retrieval.

5.7 Query with Different Forms

The system should enable the user to search for images using text or using image example or using combination of text and image query.

5.8 Text Query with Multilanguages

Current search engines face the problem of- Limited Resource Languages - The lower the Web presence of a language, the fewer hits a speaker of that language gets from a query. A query for ‘grenivka’ (Slovenian for ‘grapefruit’) produces only 24 results, of which only 9 are images of grapefruits. Yet translating the query into English produces tens of thousands of images with high precision [61].

5.9 Semantic Recommendation

Most image retrieval method always assumes that users have the exact searching goal in their mind. However, in the real world application, the case is that users do not clearly know what they want. Most of the times, they only hold a general interest to explore some related images. As a result, building a recommendation system based on the user query is necessary [49].

The system should be able to represent common search terms used in image retrieval. This is used by a keyword-generation tool to expand a user's search keyword. This is achieved by finding which concepts in the ontology relate to a keyword and retrieving information about each of these concepts.

6. CONCLUSION

A wide variety of researches have been made on image retrieval. Each work has its own technique, contribution and limitations. As a review paper, it might not include each and every aspect of individual works, however this paper attempts to deal with a detailed review of the most common traditional and modern image retrieval systems from early text based systems to content based retrieval and ontology based schemes. This paper review those works mainly based on the methods/approaches they used to come up to an efficient retrieval system together with the limitations/challenges. And we tried to give a constructive idea for future work in this field.

7. REFERENCES

- [1] Riad, A. M., Atwan,A., and Abd El-Ghany,S. 2009. Image Based Information Retrieval Using Mobile Agent, Egyptian Informatics Journal, Vol.10, No.1.
- [2] Kherfi, M.L., Ziou, D., and Bernardi,A. 2004. Image retrieval from the World Wide Web: issues, techniques, and systems. ACM Computing Surveys Vol. 36, No. 1, pp.35–67.
- [3] Riad, A. M., Atwan,A., and Abd El-Ghany,S. 2008. Analysis of Performance of Mobile Agents in Distributed Content Based Image Retrieval. In Proc. IEEE international Conference on Computer Engineering & Systems, ICCES 2008.
- [4] Wang, C., Zhang, L., and Zhang, H. 2008. Learning to Reduce the Semantic Gap in Web Image Retrieval and Annotation. In SIGIR'08, Singapore.
- [5] Alemu, Y., Koh, J., and Ikram, M. 2009. Image Retrieval in Multimedia Databases: A Survey. In Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing.
- [6] Men, H. and Chen, J. 2008. A Method of the Extraction of Texture Feature In Proceeding ISICA '08 Proceedings of the 3rd International Symposium on Advances in Computation and Intelligence.
- [7] Liu, Y., Zhang, D., Lu, G., and Ying, W. 2007. A survey of content-based image retrieval with high-level semantics. In Journal of Pattern Recognition Vol. 40, pp: 262– 282.
- [8] Su,J., Wang,B., Yeh, H., and Tseng, V. S. 2009. Ontology-Based Semantic Web Image Retrieval by Utilizing Textual and Visual Annotations. In Web Intelligence/IAT Workshops, pp: 425-428.
- [9] Popescu, A., Grefenstette, G., and Moëllic, P. 2008. Improving Image Retrieval Using Semantic Resources. In Advances in Semantic Media Adaptation and Personalization, pp: 75-96.
- [10] Hou , J., Zhang ,D., Chen ,Z., Jiang, L., Zhang ,H., and Qin ,X. 2010 .Web Image Search by Automatic Image Annotation and Translation. In 17th International Conference on Systems, Signals and Image Processing.
- [11] Saenko,k., Darrell,T. 2008. Unsupervised Learning of Visual Sense Models for Polysemous Word. In Proceedings of the 22nd Annual Conference on Neural Information Processing Systems Vancouver, Canada, pp.1393-1400.
- [12] Ménard, E. 2011. Search Behaviours of Image Users: A Pilot Study on Museum Objects. The Canadian Journal of Library and Information Practice and Research, vol. 6, no. 1.
- [13] Smeulders, A.W.M., Member, S., Worring, M., Santini ,S., Gupta ,A., and Jain, R . 2000. Content-based image retrieval at the end of the early years. In IEEE Transactions on Pattern Analysis and Machine Intelligence, pp: 1349–1380.
- [14] Ting, W. Su, Chen, J., Jier, J. and Lien, J. 2010. Region-based image retrieval system with heuristic pre-clustering relevance feedback. In Expert Systems with Applications Vol 37(7): pp:4984-499.
- [15] Shao, L. 2009. An Efficient Local Invariant Region Detector for Image Retrieval. In Canadian Conference on Computer and Robot Vision ,pp 208:212.
- [16] He, R., Xiong ,N., Yang ,L.T., and Park , J. H. 2010. Using Multi-Modal Semantic Association Rules to fuse keywords and visual features automatically for Web image retrieval. In Information Fusion. Vol 12, Issue 3.
- [17] Wong, R.C.F. and Leung, C.H.C. 2008. Automatic Semantic Annotation of Real-World Web Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 30, NO. 11, PP 1933-1945.
- [18] He, R., Xiong, N., Kim, T.-H. and Zhu ,Y. 2008. Mining Cross-Modal Association Rules for Web Image Retrieval. 2008. International Symposium on Computer Science and its Applications, pp: 393-396.
- [19] Xu,H., Zhou ,X., Lin ,L., Xiang ,Y., and Shi ,B. 2009.Automatic Web Image Annotation via Web-Scale Image Semantic Space Learning. In Proceedings of the Joint International Conferences on Advances in Data and Web Management, pp. 211–222.
- [20] Zheng, Q.-F., and Gao, W. 2008. Constructing visual phrases for effective and efficient object-based image retrieval. ACM Trans. Multimedia Computing and Communication.
- [21] Chandrika, P., Jawahar, C.V., .2010. Multi Modal Semantic Indexing for Image Retrieval. In Proceedings of the ACM International Conference on Image and Video Retrieval, China.
- [22] Huang, R.-B., Dong, S.-L., and Du, M.-H. 2008. A Semantic Retrieval Approach by Color and Spatial Location of Image Regions. In CISP '08, Image and Signal Processing conference, pp: 466- 470, China.
- [23] Nguyen, N.V., Boucher, A., Ogier, J.-M. And TABBONE, S. 2009. Region-based semi-automatic annotation using the Bag of Words representation of the keywords. In Fifth International Conference on Image and Graphics.
- [24] Lin, W.-C., Oakes, M., and Tait, J. 2010.” Improving image annotation via representative feature vector selection”, Neuro computing Vol. 73, pp: 1774–1782.
- [25] Premchaiswadi , W. and Tungkatsathan, A. 2010. On-line Content-Based Image Retrieval System using Joint Querying and Relevance Feedback Scheme. In Journal of World Scientific and Engineering Academy and Society (WSEAS), Vol 9, Issue 5.
- [26] Shen ,H. T. 2009. Speed up interactive image retrieval. In VLDB Journal, Springer, Vol. 18, No. 1, pp: 465-474
- [27] Santos, J. A., Ferreira, C. D., Torres, R., Gonçalves , M. A., and Lamparelli, R. A. C . 2011 .A relevance feedback method based on genetic programming for classification of remote sensing images. In Information Sciences: an International Journal, Vol.181 No.13, pp.2671-2684.
- [28] Zhu, X.S., and Huang, T.S. 2003. Relevance feedback in image retrieval: a comprehensive review”, Multimedia System, pp :536–544.
- [29] Ferreira, C.D., Santos, J.A., Torres, R.S., Gonçalves, M.A., Rezende, R.C., and Fan W. 2011. Relevance Feedback based on Genetic Programming for Image

- Retrieval. In *Pattern Recognition Letters*, Vol. 32, Issue 1, pp: 27-37.
- [30] Modaghegh, H., Javidi, M., Yazdi, H. S., and Pourreza, H. R. 2010. Learning of Relevance Feedback Using a Novel Kernel Based Neural Network, *Australian Journal of Basic and Applied Sciences*, Vol. 4, pp: 171-186.
- [31] Lu, Y., Hu, C., Zhu, X., and Zhang, H., Yang, Q. 2000. A unified framework for semantics and feature based relevance feedback in image retrieval systems. In *ACM International Conference on Multimedia*, pp. 31–37.
- [32] Chatzilari, E., Nikolopoulos, S., Papadopoulos, Z., Zigkolis, C., and Kompatsiaris, Y. 2011. Semi-Supervised object recognition using flickr images. In *9th International Workshop on Content-Based Multimedia Indexing*, Madrid, Spain.
- [33] Muda, Z. 2008. Ontological Description of Image Content Using Regions Relationships. In *European Semantic Web Conference (ESWC2008)*, Tenerife, Spain, pp: 46-50.
- [34] Li, Y., Lu, J., Zhang, Y., Li, R., and Xu, W. 2008. Ensemble of Two-class Classifiers for Image Annotation. In *International Workshop on Education Technology and Training & International Workshop on Geoscience and Remote Sensing*, pp. 763-767.
- [35] Zhang, D., Islam, M. M., Lu, G. 2011. A review on automatic image annotation techniques. In *Journal of Pattern Recognition*, pp: 1-17.
- [36] Feng, S., and Xu, D. 2010. Transductive Multi-Instance Multi-Label learning algorithm with application to automatic image annotation. In *Expert Systems with Applications* 37, pp: 661–670.
- [37] Hyvo, E., Styrman, A., and Saarela, S. 2002. Ontology-based image retrieval. In *Proceedings of XML Finland Conference*, pp: 27–51.
- [38] Li, Z., Shi, Z. 2009. Modeling latent aspects for automatic image annotation. In *16th IEEE International Conference on Image Processing*, Cairo, EGYPT.
- [39] Murdoch, O., Coyle, L. and Dobson, S. 2008. Ontology-based query recommendation as a support to image retrieval. In *Proceedings of the 19th Irish Conference in Artificial Intelligence and Cognitive Science*.
- [40] Qi, X., and Han, T. 2007. Incorporating Multiple SVMs for Automatic Image Annotation. *Pattern Recognition*, Vol. 40, No. 2, pp: 728-741.
- [41] Wong, R. C. F. and Leung, C. H. C. 2008. Automatic Semantic Annotation of Real-World Web Images *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, VOL. 30, NO. 11, pp 1933- 1944.
- [42] Liua, J., Li, M., Liua Q., Lu, H., and Ma, S. 2009. Image annotation via graph learning. In *Journal of Pattern Recognition* Vol. 42, pp: 218 - 228.
- [43] Liu, X., Ji, R., Yao, H., Xu, P., Sun, X., and Liu, T. 2008. Cross-Media Manifold Learning for Image Retrieval & Annotation. In *Proceedings of the 1st ACM international conference on Multimedia information retrieval (MIR'08)*, Vancouver, British Columbia, Canada. , pp: 141- 148.
- [44] Udristoiu, S., and Ion, A. L.. 2010. Image Annotation by Learning Rules from Regions Patterns. In *International Conference on Complex, Intelligent and Software Intensive Systems*, pp : 124-131.
- [45] Song, H., Li, X., and Wang, P. 2009. Multimodal Image Retrieval Based on Annotation Keywords and Visual Content. In *International Conference on Control, Automation and Systems Engineering*, pp: 295-289.
- [46] Caudill, J. D. 2009. "BRIDGING THE SEMANTIC GAP IN CONTENT-BASED IMAGE RETRIEVAL. Doctoral Thesis, Department of Computer Engineering and Computer Science University of Louisville, Louisville, Kentucky.
- [47] Kingshy, E., Sychay, G., and Gang Wu, G. 2003. CBSA: Content-based Soft Annotation for Multimodal Image Retrieval using Bayes Point Machines *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 3 ,issue 1, pp: 26 – 38.
- [48] Fan, L., Li, 2007. A User-Driven Ontology Guided Image Retrieval Model. *Proceeding COGINF '07 Proceedings of the 6th IEEE International Conference on Cognitive Informatics*.
- [49] Fan, L., and Li, B. 2006. A Hybrid Model of Image Retrieval Based on Ontology Technology and Probabilistic Ranking, In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, pp: 477-480.
- [50] Kesorn, K., Poslad, S. 2009. Semantic Restructuring of Natural Language Image Captions to Enhance Image Retrieval. In *Journal of Multimedia*, Vol. 4, No. 5, pp: 284-297.
- [51] Wang, H., Liu, S., and Chia, L-T. 2006. Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches. In *Proceedings of the 14th annual ACM international conference on Multimedia*, Santa Barbara, CA, USA.
- [52] Kong, H., Hwang, M., Na, K. and Kim, Pankoo. 2005. The Study on the Semantic Image Retrieval Using the Cognitive Spatial Relationships in the Semantic Web. *IFIP International Federation for Information Processing*.
- [53] Wang, H., Jiang X., Chia and Tan, L.-T. 2008. Ontology enhanced web image retrieval: aided by wikipedia & spreading activation theory. In *Proceeding of the 1st ACM international conference on Multimedia information retrieval*, Vancouver, British Columbia, Canada.
- [54] Li, X. Y., Shou, L. D., and Chen, G. 2006. A Latent Image Semantic Indexing Scheme for Image Retrieval on the Web. In *International Conference of Web Information Systems*, pp: 315-326.
- [55] Popescu, A., Millet, C., and Moëllic, P.-A. 2007. Ontology driven content based image retrieval. In *Proceedings of the 6th ACM international conference on Image and video retrieval*, Amsterdam, The Netherlands , pp: 387-394.
- [56] Mezaris, V., Kompatsiaris, I., and Strintzis, M. G. 2003. An ontology approach to object-based image retrieval. In *Proc. IEEE International Conference on Image Processing (ICIP 2003)*, Barcelona, Spain, Vol. 2, pp: 511-514.

- [57] Shi, L., GU, G., Liu, H., and Shen, J. 2008. A Semantic Annotation Algorithm Based on Image Regional Object Ontology. In International Conference of Computer Science and Software Engineering, Vol. 4, pp: 540-543, Wuhan, Hubei.
- [58] Ganea, E. 2010. An Object Oriented Graph Approach for Image Representation and Query Based on Content. In International Journal of Computer Science and Applications, Vol. 7 No. 1, pp: 45 – 59.
- [59] Popescu, A., 2007. Image Retrieval Using a Multilingual Ontology. Proceeding RIAO '07 Large Scale Semantic Access to Content (Text, Image, Video, and Sound), pp: 461-474.
- [60] Zhou, L. and Zhang, C. 2009. An Image Clustering and Retrieval Framework Using Feedback-based Integrated Region Matching. In International Conference on Machine Learning and Applications, pp: 422 – 427.
- [61] Etzioni, O., Reiter, K., Soderland, S., and Sammer, M. 2007. Lexical translation with application to image search on the Web, in Proceedings of the in: Machine Translation Summit XI.