



## International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

# A Review on Different Content Based Image Retrieval Techniques Using High Level Semantic Features

Nancy Goyal<sup>1</sup>, Navdeep Singh<sup>2</sup>

<sup>1</sup>Research Scholar, Master of Technology, Department of Computer Engineering, Punjabi University, Patiala, Punjab,  
India.

<sup>2</sup>Assistant Professor, Department of Computer Engineering, Punjabi University, Patiala, Punjab, India.

**ABSTRACT:** The significance of content based image retrieval system (CBIR) depends on the adopted features to represent images in the knowledge base. Using low-level features cannot give satisfactory results in many cases recovery; especially when high-level concepts in the user's mind are not easily expressible in terms of low-level features, ie semantic gap. Semantic gap between visual features and human semantics has become a bottleneck in content-based image retrieval. The need to improve the precision of image retrieval systems and reduce the semantic gap is high in view of the growing need for image retrieval.

In this paper, first introduce semantic extraction methods, and then the key technologies for reducing the semantic gap, ie, object-ontology, machine learning, generating semantic relevance feedback templates and web image retrieval are discussed.

**KEYWORDS:** Semantic gap, Semantic feature extraction, Object Ontology, Machine Intelligence, Relevance Feedback.

### I. INTRODUCTION

With the development of technology, more and more images come into view as and have become a part of our daily existence. There is a wide range of applications which require image processing tools. Some of its examples are crime prevention medicine, Fashion and graphic design, Architectural and engineering design, Publishing and advertising, research, law etc. Thus, there are many technologies have been developed to meet the requirement. There are two kinds of such technologies: text-based image retrieval and content based image retrieval, CBIR. In the text-based approach, images are usually manually searched by text descriptors. Its greatest merit is that when images are recorded correctly, good search results can be achieved. This approach has some limitations. The first is that a considerable amount of human labor for manual annotation is required. The second entry is inaccurate due to the subjectivity of human perception. To overcome the above drawbacks of text-based retrieval of images, CBIR was introduced and has become the predominant technology.

The fundamental difference between content-based and text-based retrieval systems is that the human interaction is an indispensable part of the latter system. Humans tend to use high-level features (concepts), such as keywords, text descriptors, to interpret images and measure their similarity. While the features automatically extracted using computer vision techniques are mostly low-level features (color, texture, shape, spatial layout, etc.). In general, there is no direct link between the high-level concepts and the low-level features [14].

The fundamental difference between content-based systems and text-based content is human interaction. Humans tend to use high-level functions as keywords, text descriptors, to interpret the images and measure their similarity. While the features extracted automatically using computer vision techniques which are basically low-level features (color, texture, shape, spatial design, etc.). In general, there is no direct link between high-level concepts and the low-level features [14].



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

Although many complicated algorithms have been designed to describe the color, shape and texture features, these algorithms cannot sufficiently represent the semantics of image and have many limitations when dealing with large databases of image content. Extensive experiments on CBIR systems show that the low-level content often fail to describe the high-level semantic concepts in the user's mind. Therefore, the performance of CBIR is far from user expectations [3]. There is three levels of queries in CBIR.

Level 1: Retrieval by primitive features such as colour, texture, shape or the spatial location of image elements. Typical query is query by example, 'find pictures like this'.

Level 2: Retrieval of objects which are characterized with the features of images, using some logical implication. For example, 'find a picture of a flower'.

Level 3: Retrieval by abstract attributes, which involve a great deal of high-level thinking about the purpose of the objects or scenes expressed. This includes flashbacks name of pictures with emotional or religious significance, etc. Query example, 'find pictures of a joyful crowd'.

Level 2 and level 3 together are known as semantic image retrieval, and the gap between levels 1 and 2 is named as semantic gap [1]. More specifically, the inconsistency between the limited expressive power of image features low level and richness of user semantics is known as the "semantic gap". Users at Level 1 usually required submitting a sample image or sketching as query. But what if the user does not have a sample image at hand? Semantic image retrieval is more convenient for users, supporting keyword query or texture. Therefore, to allow queries by high-level concepts a CBIR system should provide full support in reducing the "semantic gap" between features of digital images and the richness of human semantics.

## II. SURVEY OF RELATED WORK

The field of image retrieval has been an active research area for several decades and becomes more and more interesting area in recent years. Several reviews on image retrieval have been published. In 2008, Yu Xiaohong describes that the basic components of content-based image retrieval system are color, texture, shape and semantic. The semantic-based image retrieval is a better way to solve the "semantic gap" problem, so the semantic-based image retrieval method is stressed in his paper. Other related techniques such as relevance feedback and performance evaluation also discussed.

In 2008, Prof. Sharvari Tamane, proposed a new system for image retrieval using high level features which is based on extracting low level features such as color, shape and texture features and converse them into high level semantic features using fuzzy production rules, with the help of image mining technique. Main advantage of his proposed method is the possibility of retrieval using high level semantic features.

In 2010, Lijun Zhao and Jiakui Tang compare different visual feature combinations in retrieval experiments. They describe that only low-level features for CBIR cannot achieve a satisfactory measurement performance, since the user's high-level semantics cannot be easily expressed by low-level features. In order to reduce the gap between user query concept and low-level features in CBIR, a multi-round relevance feedback (RF) strategy based on both support vector machine (SVM) and feature similarity is adopted to meet the user's requirement. This implementation can improve the performance by increasing the number of feedback.

In 2012, Anuja Khodaskar proposed effective content-based image retrieval (CBIR) systems are accurate characterization of visual information. Traditional image retrieval method has some limitations. In order to improve the retrieval accuracy of content-based image retrieval systems, 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. His paper presents the technique for efficient CBIR with high level semantic features by using object ontology. Object-Ontology provides a qualitative definition of high-level query concepts.



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

In 2013, Jisha.K.P, Thusnavis Bella Mary. I , Dr.A.Vasuki, focuses on the semantic based image retrieval system using Gray Level Co-occurrence Matrix (GLCM) for texture feature extraction. The images are retrieved according to user satisfaction and thereby reduce the semantic gap between low level features and high level features.

## III. EXTRACTING SEMANTIC FEATURES

Content of the image is blurred, complex, abstract, simply by using low-level features to describe image is far from sufficient, high-level semantics are needed to describe the abstract characteristics of the image [9]. There are three conventional methods for extracting the semantic feature image. First, semantic feature can be extracted based on image processing and domain knowledge. There are three fundamental processes: image segmentation, object recognition, and analysis of the object relation. Secondly, we can obtain the semantics of images from tags manual or human interaction [10]. But in this hard human labor is required. Third, we can extract the semantics from external information, such as file name, URL, image near the text or metadata information.

### A. Semantic Extraction Based on Knowledge

The main feature of semantic extraction based on knowledge is providing the system necessary knowledge in advance. It can be divided into the following methods according to the semantic content extracted and methods used.

#### 1) Semantic extraction based on object recognition

Semantic extraction based on object detection generally uses the traditional outline of computer vision. It has three main processes: image segmentation, object recognition and analysis of the spatial relationship of objects. This is a bottom-up process, and each step of the process is the based on previous step. System like this is intuitive and consistent with the order of the people observing things. The objects and spatial relationships can become the basis for higher semantic level and assistant to obtain semantic scene measured image. Such methods when combined with specific domain knowledge can be successful in specific application areas. Representation of spatial relations and corresponding similarity methods can be used such as strings and 2d graphic spatial orientation. These methods may require a higher level of spatial semantics in practice. Image segmentation and object recognition are great difficulties in the process.

#### 2) Extraction of scene and behavioral semantics

The perceptual thinking is that on the basis of object recognition and spatial relations, abstract scene and behavior semantics can be extracted with the semantic-level knowledge of concept. However, this thinking is only possible in limited by the variety of the scene image areas. There are some other methods are used to obtain the semantics of the scene images in the case of the difficulty of object recognition cannot be solved effectively. The characteristic of these methods is bypassing the process of image recognition. Usually, you can not automatically extract from a single image in terms of semantics and behavior often needs to consider the motion information from the image

#### 3) Extraction of emotional semantics

The highest level of image semantics is emotional semantics of the image. Compared to the other semantics, is more subjective, as it involves subjective models of people, cultural background and aesthetic standards. It must involve feelings of the people, psychological, cognitive, and aesthetic and other factors in the extraction of visual features closely related to emotion present in image and then establish a mapping between feature space and emotional space.

At present, the transformation method based on the emotional semantic boundary is an issue. It needs to study more background combined with knowledge and experience in related fields.

### B. Semantic Extraction Based on Manual Interaction

To access the semantics of images effectively we cannot completely depending on computers according to current position of computer technologies and artificial intelligence. One possible solution would be, to add human into retrieval system so that computers can get better information and performance of system can be improved and retrieve better results. So, make a system to learn and correct the semantic description of the image database with the interaction between the system and the human. This is mainly involved in the pre-image processing and feedback learning. The simplest method of image preprocessing is manual mark, while the more rational method is combined



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

with knowledge-based methods: first, let the system automatically learn objects and describing scenes, and then be modified by persons. The user feedback can play two roles. One is to capture the real needs of the users gradually according to the user operation based on the retrieval results. The other is to establish and modify the relationship between high-level semantics and low-level features of the image.

## C. Semantic Extraction Based on External Information

In this method, semantics are extracted from some other relevant information of images, such as file name, URL, text near the image or metadata information. The most common and widely used method is manual indexing. The advantage of this method is that the expression of information is spontaneous, easily processed in machine and can describe high-level concepts. The disadvantage is that management or organization of information is difficult, and there is subjectivity and incorrectness.

## IV. REDUCING 'SEMANTIC GAP'

Those retrieval systems which are based on low-level features do not match human perception, and considered insufficient and unpredictable. This is known as semantic gap [2]. These systems are not able to find images according to user's requirement such as "Find images with rocks." Some researchers have tried to fill the semantic gap by proposing new techniques. These techniques are explained below:

### A. Object Ontology:

Object Ontology uses semantics of an image to define it. This technique, define different levels for assigning low-level image features. Each level defines the region attribute that is well known to humans and considered as the intermediate level descriptor for an image [5]. For example, sky can be defined as topmost blue region. Similarly, blue region can be defined as 'blue high', 'blue medium', 'blue low'. Images can be classified into different categories by mapping such descriptors to high-level semantics based on our knowledge. For example "sky" can be defined as region of "light blue" color and "upper" spatial location. Quantization of color and texture feature is the key in such systems. To support semantic-based image retrieval, a more effective and widely used way to quantize color information is by color naming. Berk, Brownston and Kaufman proposed a color naming system 'CNS' which quantizes the hue values into set of basic colors such as red, orange, brown, yellow, green, blue and purple, black, grey and white. Compared to color, texture naming system is not yet available.

### B. Machine Intelligence

This technique is suitable for complex semantics but difficult to implement. This technique uses two types of machine learning tools that are supervised or unsupervised, associated low-level features with query concepts. This method predicts the value of production on the basis of inputs that user provide.

In supervised machine intelligence, number of images is collected and on basis of entry measure a binary classifier is trained to detect semantic category label. Bayesian Classifier is an important method in which images database are automatically classified into general types as indoor or outdoor (city / landscape, etc.) Another method is neural network according to which the user selects 11 categories (concepts): brick, cloud, fur, grass, ice, road, rock, sand, skin, trees and water. Then a lot of input is introduced into the neural network classifiers to establish the link between low-level features of an image and its high-level semantics. Decision tree is another technique to get semantic features. Decision tree is constructed on the bases of relevancy of images input to the query and then used as a model to distinguish between images from databases into two classes: relevant and irrelevant.

In unsupervised learning, the aim is to describe how data is organized or grouped that user entered. No outcome; It merely describes the organization of the input data. In this technique, similar images are assigned to groups. Each group is assigned a name, which maximizes the possibility of obtaining similar image of that specific group.

### C. Relevance Feedback

Relevance feedback methods work online and based on intentions of user. Relevance feedback technique is not feasible in some domain but this technique involves user's communication. Relevance feedback mechanism works when the user enters the query in the form of an image, picture or text. The system provides the initial results of user retrieval



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

and make image as "relevant" and "irrelevant." Machine learning algorithm learns user feedback and selector gets other images[4]. The process is repeated until the user is satisfied with the results.

A typical scenario for RF in CBIR is:

- (1) The system provides the initial results of retrieval through query, sketches, etc.
  - (2) User judges the above results as to whether and to what extent they are relevant (positive examples) / irrelevant (negative examples) to the query.
  - (3) Machine learning algorithm is applied to learn user feedback. Then go back to (2).
- (2) & (3) are repeated until the user is satisfied with the results.

## D. Semantic Template

Semantic Template is a set of general characteristics that are calculated from the number of images stored in the database. It traces a link between the characteristics of high and low-level features. It usually defined as a function of a concept representative calculated from a group of sample images. Chang et al. introduced semantic visual template (SVT) to link the low-level image feature to high level. Semantic visual template (SVT) image feature links low-level to high-level concepts for video retrieval. To generate SVT, the user first defines the template for a specific concept by specifying objects and their spatial and temporal constraints, the weights assigned to each feature for each object. This initial stage of consultation is provided to the system. Through user interaction, the system eventually meets to a small set of queries that match the concept in the mind of the user i.e. Recall ratio maximizes. SVT generation depends on the interaction with the user and requires user understanding in depth of image characteristics. This obstructs its application to normal users.

## E. Web Image Retrieval

This system has some technical difference from image retrieval in other application. Some additional information on the Web is available to ease semantic-based image retrieval. For example, the URL of image file often has a clear hierarchical structure including some information about the image such as image category [12]. In addition, the HTML document also contains some useful information in image title, ALT-tag, the descriptive text surrounding the image, hyperlinks, etc. However, such information can only annotate images to a certain extend. Existing Web image searching such as Google and AltaVista search images based on textual evidences only. Though these approaches can find many relevant images, they cannot confirm whether the retrieved images really contain the query concepts so the retrieval precision is poor. The result is that users have to go through the entire list to find the desired images. This is a time-consuming process as the returned results always contain multiple topics which are mixed together. To improve Web image retrieval performance, researchers are making effort to combine the evidences from textual information and visual image contents.

## V. CONCLUSION

In this paper various semantic feature extraction methods and techniques for reducing semantic gap are described. There are three traditional methods to extract the image semantic feature. Firstly, semantic feature can be extracted based on image processing and domain knowledge which contains three key processes that are image segmentation, object recognition, and object relation analysis, Secondly, we can get image semantics from manual labels or human interaction. Thirdly, we can extract semantics from external information such as file name, URL, text near the image or metadata information. Techniques for reducing semantic gap are summarized Ontology-based algorithms are easy to design and are suitable to applications with simple semantic features. Machine learning techniques are used to learn more complex semantics. Due to its simplicity in implementation and the spontaneous mapping from low-level features to high-level concepts using decision tree which is a very effective tool for image retrieval. RF has been proved to be effective in increasing image retrieval accuracy. The problem is that most current systems requires about five or even more iterations before it converges to a stable performance level, but users are usually irritated and may give up after two or three tries. Semantic templates uses characteristics extracted from a group of similar images which ease to retrieve image of same kind. Web image retrieval uses additional information to facilitate the retrieval process and become an active research area. Table 1, shows advantages and disadvantages of different techniques used for reducing semantic gap in retrieval of images.



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

Comparison between various Semantic Reduction Gap Techniques		
Techniques	Advantages	Disadvantages
Object Ontology	Easy to design, suitable to application with simple semantic	No proper system available for texture naming
Machine Learning	Suitable for complex semantics	Low speed of learning.
Relevance Feedback	Effective in boosting image retrieval accuracy	Requires five or more iterations to achieve high performance
Semantic Templates	Maximum Recall Ratio	In depth understanding requires
Web Image Retrieval	Provides additional information on web	Results are not so much accurate

Table 1

## REFERENCES

- [1] Liu, Ying, Zhang, Dengsheng, Lu, Guojun, and Ma, Wei-Ying, "A survey of content-based image retrieval with high-level semantics", Pattern Recognition 40, 262 – 282, 2009.
- [2] Zhou, Xiang Sean, and Huang, Thomas S., "CBIR: From Low-Level Features to High-Level Semantics", Beckman Institute for Advanced Science and Technology University of Illinois at Urbana Champaign, Urbana, IL 61801.
- [3] Tamane, Sharvari, "Content Based Image retrieval Using High level semantic features", Proceedings of the 2<sup>nd</sup> National Conference, 2008.
- [4] Zhao, Lijun, and Tang, Jiakui, "Content-Based Image Retrieval Using Optimal Feature Combination and Relevance Feedback", IEEE, International Conference on Computer Application and System Modeling (ICCASM 2010), 2010.
- [5] Khodaskar, Anuja and Ladke, S.A., "Content Based Image Retrieval with Semantic Features using Object Ontology", International Journal of Engineering Research & Technology (IJERT) Vol. 1 Issue 4, June – 2012.
- [6] Jisha K.P, Thusnavis Bella Mary. I, and Vasuki, A., "An Image Retrieval Technique based on texture features using semantic Properties", International Conference on Signal Processing, Image Processing and Pattern Recognition [ICSIPRI], 2013.
- [7] Jiang, Wei, Chan, Kap Luk, Li, Mingjing, and Zhang, Hongjaing, "Mapping Low-Level Features to High-Level Semantic Concepts in Region-Based Image Retrieval", IEEE, Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [8] HuMin, and Shuangyuan, Yang, "Overview of content-based image retrieval with high-level semantics", IEEE, 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), 2010.
- [9] Li, Qingyong, Hu, Hong, and Shi, Zhongzhi, "Semantic Feature Extraction Using Genetic Programming in Image Retrieval", IEEE, Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04).
- [10] Eidenberger, Horst, and Breiteneder, Christian, "Semantic Feature Layers in Content-based Image Retrieval: Implementation of Human World Features", Vienna University of Technology, Institute of Software Technology and Interactive Systems, Favoritenstrasse 9-11 – 188/2, A-1040 Vienna, Austria.
- [11] Chen, Zhijun, "Semantic Research on Content-Based Image Retrieval", IEEE, 2010.
- [12] Kaur, Harpreet, and Jyoti, Kiran, "Survey of Techniques of High Level Semantic Based Image Retrieval", International Journal of Research in Computer and Communication technology, IJRCCT, ISSN 2278-5841, Vol 2, Issue 1, January 2013.
- [13] Pattanaik, Swapnalini, and Bhalke, D.G., "Beginners to Content Based Image Retrieval", International Journal of Scientific Research Engineering & Technology (IJSRET), Volume 1 Issue2, pp 040-044, May 2012.
- [14] Deb, Sagarmay, and Zhang, Yanchun. "An Overview of Content-based Image Retrieval Techniques", IEEE, 18th International Conference on Advanced Information Networking and Application (AINA'04), 2004.