Semantic Gap in CBIR: Automatic Objects Spatial Relationships Semantic Extraction and Representation

Hui Hui Wang hhwang@fit.unimas.my
Department of Computing and Software Engineering,
Faculty of Computer Science and Information Technology,
Universiti Malaysia Sarawak,
94300 Kota Samarahan, Sarawak, Malaysia

Dzulkifli Mohamad dzulkifli@utm.my
Department of Computer Graphics and Multimedia,
Faculty of Computer Science and Information Technology,
Universiti Teknologi Malaysia,
81300 Skudai, Johor Bahru, Malaysia

N. A. Ismail azman@utm.my
Department of Computer Graphics and Multimedia,
Faculty of Computer Science and Information Technology,
Universiti Teknologi Malaysia,
81300 Skudai, Johor Bahru, Malaysia

Abstracts

The explosive growth of image data leads to the need of research and development of Image retrieval. Image retrieval researches are moving from keyword, to low level features and to semantic features. Drive towards semantic features is due to the problem of the keywords which can be very subjective and time consuming while low level features cannot always describe high level concepts in the users' mind. This paper is proposed a novel technique for objects spatial relationships semantics extraction and representation among objects exists in images. All objects are identified based on low level features extraction integrated with proposed line detection techniques. Objects are represented using a Minimum Bound Region (MBR) with a reference coordinate. The reference coordinate is used to compute the spatial relation among objects. There are 8 spatial relationship concepts are determined: "Front", "Back", "Right", "Left", "Right-Front", "Left-Front", "Right-Back", "Left-Back" concept. The user query in text form is automatically translated to semantic meaning and representation. Besides, the image similarity of objects spatial relationships semantic has been proposed.

Keywords : Semantic Gap, Objects Spatial Relationships Semantic, Automatic Image Semantic Extraction, Image Retrieval

1. INTRODUCTION

Image retrieval is the field of study concerned with searching and browsing digital images from database collection. This area of research is very active research since the 1970s [1, 2]. Due to more and more images have been generated in digital form around the world, image retrieval attracts interest among researchers in the fields of image processing, multimedia, digital libraries, remote sensing, astronomy, database applications and other related area. Effective and fast retrieval of digital images has not always been easy, especially when the collections grow into thousands. An effective image retrieval system needs to operate on the collection of images to retrieve the relevant images based on the query image which conforms as closely as possible to human perception.

1.1. Evolution of Image Retrieval

The purpose of an image database is to store and retrieve an image or image sequences that are relevant to a query. There are a variety of domains such as information retrieval, computer graphics, database management and user behavior which have evolved separately but are interrelated and provide a valuable contribution to this research subject. As more and more visual information is available in digital archives, the need for effective image retrieval has become clear [3,4]. In image retrieval research, researchers are moving from keyword based, to content based then towards semantic based image retrieval and the main problem encountered in the content-based image retrieval research is the semantic gap between the low-level feature representing and high-level semantics in the images.

1.1.1. Keyword Based Image Retrieval

In 1970s, the conventional image retrieval system used keyword as descriptors to index an image however the content of an image is much richer than what any set of keywords can express.

Text-based image retrieval techniques employ text to describe the content of the image which often causes ambiguity and inadequacy in performing an image database search and query processing. This problem is due to the difficulty in specifying exact terms and phrases in describing the content of images as the content of an image is much richer than what any set of keywords can express. Since the textual annotations are based on language, variations in annotation will pose challenges to image retrieval.

1.1.2. Content Based Image Retrieval

In 1990s, Content-based image retrieval (CBIR) then has been used as an alternative to text based image retrieval. Unlike keywords-based system, visual features for contents-based system are extracted from the image itself. CBIR can be categorized based on the type of features used for retrieval which could be either low level or high level features. At early years, low level features include colour, texture, shape and spatial relations were used. The summary of CBIR researches done in retrieving the image based on their visual content can be found in our paper, ref [5]

Although there are many sophisticated algorithms to describe color, shape and texture features approaches, these algorithms do not satisfied and comfort to human perception This is mainly due to the unavailability of low level image features in describing high level concepts in the users' mind. For an example finding an image of a little boy is playing a ball in the garden. The only way a machine is able to perform automatic extraction is by extracting the low level features that represented by the color, texture, shape and spatial from images with a good degree of efficiency.

1.1.3. Semantic Based Image retrieval

In 2000s, semantic based image retrieval has been introduced. This is due to neither a single features nor a combination of multiple visual features could fully capture high level concept of images. Besides, the performance of image retrieval system based on low level features are not satisfactory, there is a need for the mainstream of the research converges to retrieve based on semantic meaning by trying to extract the cognitive concept of a human to map the low level image features to high level concept (semantic gap). In addition, representing the image content with semantic terms allows users to access images through text query which is more intuitive, easier and preferred by the front end users to express their mind compare with using images. The review and general framework of semantic based image retrieval can be found our paper in ref [6]

1.2. Semantic Gap

Bridging the semantic gap for image retrieval is a very challenging problem yet to be solved [7,8]. Describing images in semantic terms is an important and challenging task that needed to carry out to fulfill human satisfaction besides to have more intelligent image retrieval system.

Human beings are able to interpret images at different levels, both in low level features (colour, shape, texture and object detection) and high level semantics (abstract objects, an event). However, a machine is only able to interpret images based on low level image features. Besides, users prefer to articulate high-level queries [9,10], but CBIR systems index images using low-level features. Hence, introducing an interpretation inconsistency between image descriptors and high-level semantics that is known as the semantic gap [3,10]. The semantic gap is the lack of correlation between the semantic categories that a user requires and the low-level features that CBIR systems offer. The semantic gap between the low-level visual features (color, shape, texture, etc.) and semantic concepts identified by the user remains a major problem in content based image retrieval [8].

Semantic content representation has been identified as an important issue to bridge the semantic gap in visual information access. It has been addressed as a good description and representation of an image, it able to capture meaningful contents of the image. Current researches often represent images in terms of labeled regions or images, but pay little attention to the spatial positions or relationships between those regions or objects [11]. Spatial relationship is needed in order to further increase the confidence in image understanding. Besides, users preferred to express their information needs at the semantic level instead of the level of preliminary image features. Moreover textual queries usually provide more accurate description of users' information needs.

The attempt to overcome the gap between high level semantic and low level features by representing images at the object level is needed [7] as well as the spatial relationship of objects to further increase the image understanding

2. RELATED WORKS

In general, there is no direct link between such high-level semantic concepts and the automatically extracted, low-level image features. Therefore, to support query by semantic concept, there is a compelling need for CBIR systems to provide maximum support towards bridging the 'semantic gap' between the low-level visual features [3,12] and it is a very challenging task to extract and manage meaningful semantics and to make use of them to achieve more intelligent and user friendly retrieval [13].

2.1. Manual Image Semantic Extraction

One conventional and common ways to describe the image in high level is using the manual annotation. Manual annotation needs to annotate every image by human where users enter some descriptive keywords when the images are loaded/registered/browsed. Existing applications that support manual image annotation include Wikipedia image collection [14], lonely planet [15], photoblog [16], fotopages [17], flickr and etc. They allow human to annotate images with some keywords. It is based on whole images and cannot annotated based on the objects or regions of the images. Inotes [18] and facebook [19] are most popular manual image annotation approaches where user can annotate various objects or regions based on selected regions in an image instead of just annotate whole images.

Although manual annotation of image content is considered a "best case" in terms of accuracy, since keywords are selected based on human determination of the semantic content of images, as well as able to support user queries in text. However it is a labor intensive and tedious process. In addition, manual annotation may also introduce retrieval errors due to users forgetting what descriptors they used when annotating their images after a lengthy period of time.

Researchers have explored techniques for improving the process of manual annotation. So, researchers are moving toward automatically automatic extraction of the image semantic content.

2.2. Semi/Automatic Image Semantic Extraction

Reference [20] was the one of the early proposed automatic solutions, where a probabilistic framework based on estimating class likelihoods of local areas, labeled as either man made or natural, inside or outside objects scenes. Zhao and Grosky [13] proposed a method integrating the LSI, normalization and term weighting to obtain the meaningful features mapped to semantic landscapes.

All the methods discussed are only able to retrieve similar images which have the whole semantics and does not indicate which part of the image gives rise to which words, so it is not explicitly object recognition. They are lacking of the ability to find the object semantics in images

Various methods have been proposed to automatically capture region semantic of images instead of image semantic only. Reference [21] introduced the region-based co-ocurrence model to assign the word to the region. Reference [22] was proposed a model of object recognition as machine translation. In this model, the mapping between regions and keywords is learnt using a method based on the EM algorithm. Reference [23] implement a cross-media relevance model and identify 70 object concepts. The model in learns the joint probability of associating words to image features from training sets and uses it to generate the probability of associating a word to a given query image. Reference [24] then improved the reference [38] approach by using continuous probability density functions to estimate the probability of observing a region given an image, it is proven that it can improve on the results obtained in [23]. In [25], the authors propose a framework to link signal-based features to semantic data by mapping the multidimensional space defined by the extracted low-level features to predefined semantic classes through a Bayesian network.

Even though the object/region semantics can be captured but then the extraction of spatial relational semantic descriptors is often neglected. They do not take into account the relational spatial semantics among objects in the images which affects the quality of the retrieval results. In other words, they only able to recognize images which contain a cat and a dog. However it fails to tell the actual relative direction (spatial relationship) between the dog and the cat. Representation of spatial relations semantics among objects are important as it can convey important information about the image and to further increase the confidence in image understanding contribute to richer querying and retrieval facilities.

2.3. Objects Spatial Relationships Semantic Extraction

Some methods has been introduced to capture the spatial relationship semantic. Some researches [11, 26] use ontology method to capture the relation semantics. Reference [26] adding the spatial semantic in image annotation by adopting the spatial concepts from the Suggested Upper Merged Ontology (SUMO). This approach is having high dependency on the semi-annotation process. Reference [26] extending it in both the granularity of the absolute positions, the extraction of combined relations (like above and to the left of) and through the use of object properties in the ontology to infer more complex spatial relations. Reference [42] proposed a unified multi-facetted framework unifying visual semantics and relational spatial characterization for automatic image retrieval that enforces expressivity and computational efficiency through fast lattice processing. The spatial relation of objects in image is represented using conceptual structures. The image is index using conceptual graph. The combined relation also can be captured using the mentioned method.

Even thought simple relation and combined relation has been captured. However there are still having some false objects spatial relationships extraction concept (Example in Figure 1, the object B and C suppose to have Front/Back spatial relationship however ref [26, 27] extracted it as Left/right relation concept and ref [11] extracted it as above-right relation).The combined relation is limited to 2 objects only for query images such as Object A is Left to and in Front object B and also lack of abstract query such as traffic jam situation. The multiple objects with combined relation should be considered such as Object A is Left to Object B and Right-Front Object C. And also, there is none of spatial relation semantic similarity for the spatial semantics description. Besides, It should be fully automatic image and spatial relation semantic extraction without involving any user or relevance feedback during the retrieval process.

3. THE RESEARCH FRAMEWORK

The research framework consists of four main components, which include low level features extraction, object identification and object semantic extraction and representation.

1) The Low Level Features Extraction Component – the low level features of images are extracted and grouped based on their common characteristics to get the regions/segment of images.
2) Object Identification Component- Regions/segment of images are integrated with line detection technique to extract the object of interest in images.
3) Object Semantic Extraction Component - The object identified is indicated using Minimum Bound Region (MBR) and the properties of objects are extracted (eg. Color of car). Each object is represented by a reference coordinate.
4) Objects Spatial Relationship Semantic Extraction Component -This component automatically extracts and identifies spatial information. It captured the spatial relationship among objects in the images.
In this paper, the discussion and concentration is mainly on the objects and their spatial relationship semantics extraction and representation in the image. Traffic images are used as the domain of study.

3.1. Objects Spatial Relationships Semantic extraction

The objects spatial relationships semantic extraction approach has 6 main stages

3.1.1 Determine the Road Slope as Z-axis

The road slope is needed to determine and it is used as the reference slope for getting spatial relationship for all car objects in image. It is actually act as a Z-axis due to the image view is slanted.

Road Slope, $m(R) = \dfrac{y_2 - y_1}{x_2 - x_1}$

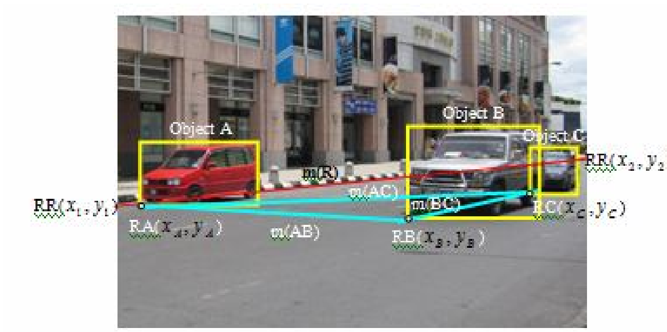


FIGURE 1 : Traffic images

3.1.2 Calculate the Slope of Each Possible Pair of Objects in Image

Each of the objects is represented using Minimum Bound Region (MBR) that indicated using a box as show above. Each object has a reference coordinate (indicated by a small circle, left bottom of MBR)

From Figure 1, Image consists of multiple car objects, I={Object A, Object B, Object C}
The reference coordinate of objects as below

Object A = RA(x_A, y_A)

Object B = RB(x_B, y_B)

Object C = RC(x_C, y_C)

The slope of each possible pair of objects is calculated based on their object's reference coordinate as below.

Object A and Object B = $m(AB) = \frac{y_B - y_A}{x_B - x_A}$

Object A and Object C = $m(AC) = \frac{y_C - y_A}{x_C - x_A}$

Object B and Object C = $m(BC) = \frac{y_C - y_B}{x_C - x_B}$

The spatial relationship of two objects is defined by computing and comparing the slope of two relative objects.

3.1.3. Determine the Relational/Directional Relationship

The basis of interpreting positions in reality is assumed as follows: The positions of left and right when viewing an image is inversed from the positions in real life. This means that when interpreting the image in Figure. 1, the red car is on the left in the image but in actual fact, it is on the right.

There are four basic relative relations of “Left to”, “Right to‘, “Front to” and “Back to” as well as the composite spatial relations of “Right-Front”, “Left-Front”, Right-Back”, Left-Back” will be determined and discussed. Those are the most important relationship in the domain of traffic images.

a) Front / Back Relationship

The positive and negative value of slope between pair of objects will be used as indicator for Front/Back relation concept to ensure object is front or back of another object.

The rules for inferring Front/Back relations are defined and illustrated in Figure 2:

m(AB) > 0, A is Front to B
m(AB) < 0, A is Back to B

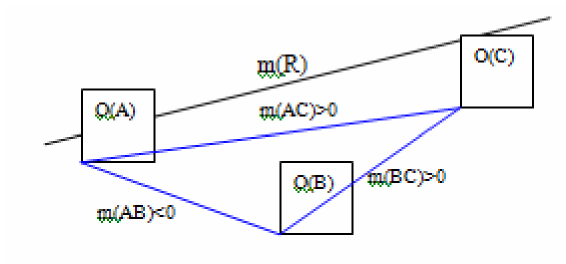


FIGURE 2 : Front/Back relationship determination

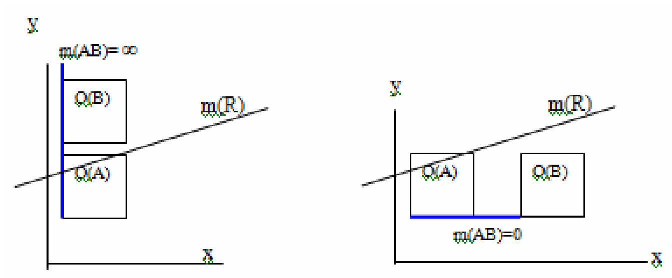
b) Absolute Front/Back and Right/Left Relationship

Two objects are considered absolutely Front/Back relation when their slope value always infinity and absolutely Right/Left relation when their slope value is 0.

The rules for inferring absolutely Front/Back and Right/Left relations are defined and illustrated in figure 3 :

m(AB) = ∞, A is Front/Back to B

m(AB) = 0, A is Right/Left to B



(a) (b)

FIGURE 3 : Absolute Front/Back relationship(a) and Absolute Right/Left relationship (b)

From the Figure 3, it shows that 2 objects only have 1 relation. It's either Front/Back relation for object A and B in Figure 3(a) or Right/Left relation for object A and B in Figure 3(b). However, there are some composite relation exists between objects such as Left-Front, Left-Back, Right-Front and Right-Back.

c) Composite Relationship

In traffic image, the road slope, m(R) value is used as reference line (axis-z) instead of axis-x and axis-y. Given the figure below,

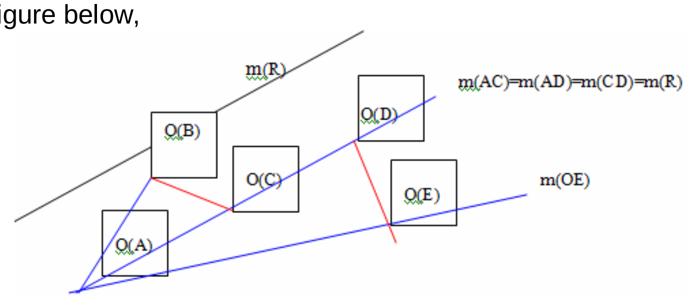


FIGURE 4 : Composite Relation

If the slope of 2 objects is same with slope of road, it means there is only Front/Back relation. The rules for inferring Absolute Front/back relationship are defined as below

m(R) =m $(O_i, O_j \mid i \neq j)$, Object I is absolute Front to Object j

If the slope of 2 objects is same with 0, its mean there is only Right/Left relation.

m $(O_i, O_j \mid i \neq j)$ =0, Object i is absolute Right to Object j

The composite relation is observed when the slope of 2 objects has value either greater or smaller than slope of road,

m $(O_i, O_j \mid i \neq j)$ <m(R), Object i is Right-Front to Object j, m>0
m $(O_i, O_j \mid i \neq j)$ <m(R), Object i is Right-Back to Object j, m<0
m $(O_i, O_j \mid i \neq j)$ >m(R), Object i is Left-Front to Object j, m>0

3.1.4. Distance of Spatial Relationship Semantic

The distance of the spatial relationship among 2 objects are calculated by,

Spatial relationship Distance, d, d $(O_i, O_j \mid i \neq j)$ =|m($(O_i, O_j \mid i \neq j)$ -m(R))|

where m $(O_i, O_j \mid i \neq j)$ is the slope of object O_i and object O_j , m(R) is the slope of the road.

3.1.5. Spatial Relationship Semantics Representation

Spatial relation semantic concept of image can be represented as, Sp

Sp={m(R),[O:(P_{ij} (O_i, O_j) , m_{ij} (O_i, O_j) , d_{ij} (O_i, O_j))]| $i \neq j$, $\forall i, j \in O, O \in I$ }

where m(R) is the slope of road, P_{ij} (O_i, O_j) is pair of objects O_i and O_j , m (O_i, O_j) is the slope between object O_i and O_j and d (O_i, O_j) is the distance of the spatial relation between object O_i and O_j , O is total number of object in image I.

4. USER QUERY

The user query is used to express the user's information need to retrieve images in collection of database that conform to human perception. According to Ref [28], to define a semantic meaning and representation of the input query that can precisely understand and distinguish the intent of the input query are the major challenges. It is difficult and often requires many human efforts to meet all these challenges by the statistical machine learning approaches.

Querying by visual example is a paradigm, particularly suited to express perceptual asre pects of low/intermediate features of visual content [29]. Visual content refer to color, shape and texture features of images. Although promising progresses have been made in image retrieval techniques based on visual features, formulating a query such as submitting an example image or a sketch is sometimes not convenient for users.

Text-based queries are the most popular query method. User usually prefers using keywords to indicate what they want. [9,30]. Textual queries usually provide more accurate description of users' information needs as it allow users to express their information needs at the semantic level and high level abstractions instead of limited to the level of preliminary image features. However, the textual words need to be translated automatically to semantic meaning and representation that are matched in the images semantic representation in database in order to have fully and precisely understand the user input.

4.1 Semantic Extraction and Representation for User Query

The user query in text forms that mainly focus on Object Spatial Relationship is automatically translate to semantics meaning and representation. The object spatial relationship semantic translation from user query has 3 main stages

4.1.1. Determine Rules for Objects Spatial Relationship

Right/Left, Rule 1: (P_{ij} (O_i, O_j), m_{ij} (O_i, O_j)=0, d_{ij} $(O_i, O_j) < 0$), i Right j, j Left i

Front/Back Rule 2: (P_{ij} (O_i, O_j), m_{ij} $(O_i, O_j) = \infty$, d_{ij} $(O_i, O_j) > 0$), i Front j, j Back i, j left-Back i

Left-Front Rule 3: (P_{ij} (O_i, O_j), m_{ij} (O_i, O_j)>0, d_{ij} (O_i, O_j)>0), i Left-Front j

Right-Front Rule 4: (P_{ij} (O_i, O_j), m_{ij} (O_i, O_j)>0, d_{ij} (O_i, O_j)<0), i Right-Front j

Right-Back Rule 5: (P_{ij} (O_i, O_j), m_{ij} (O_i, O_j)<0, d_{ij} (O_i, O_j)<0), i Right-Back j

Logical operation L: AND, OR, NOT

4.1.2. Sub Divided User Query to Sub User Query (if there is any).

Given example user query:

Object A Right Object B AND Right-Front Object C

Q=(QC_1….QC_k),

QC_1 : Sub Query 1 : Object A Right Object B

QC_2 : Sub Query 2 : Object A Right-Front Object C

4.1.3. Assign the Sub Query to Rules for Objects Spatial Relationship

The conversion of the above user query to spatial relationship semantics representation rules is as below,

Q=(QC_1….QC_k),

Where QC_1….QC_k is sub user query and k is the number of sub query for the user query.

QC_1= (P_{AB} (O_A, O_B), $m_{AB}$$(O_A, O_B$=0), d_{AB} $(O_A, O_B) < 0$) Operator

QC_2 = (P_{AC} (O_A, O_C), $m_{AC}$$(O_A, O_C$>0), d_{AC} $(O_A, O_C) < 0$))

There is a operator between sub query, the operator can be logical operation AND, OR or NOT.

5. IMAGE SPATIAL SEMANTICS SIMILARITY MEASUREMENT

Image Spatial semantic similarity is used to define the Spatial similarity between the image in database and user query.

Database Images $D_I = I_{i=0},\ I_1 \ldots\ldots\ldots I_p$

$I_i = (P_o,\ P_1 \ldots\ldots P_n)$

Where database images consists of k number of images, each image i consists of the n number of object pair. The n number is different for the image from i=0 to i=p.

User Query, Q=(QC_1….QC_k),

Where QC_1….QC_k is sub user query and k is the number of sub query for the user query.

The steps of the image semantics similarity is discuss as below

5.1. Query Matching

Query matching is the process to determine total number of matched element in sample image I to the Query Q. It is evaluated in terms of the matching based on m and d characteristic

As an example of Query Q, object A Right Object B is defined as below,

$$Q=(P_{AB}\ (O_A, O_B), m_{AB}(O_A, O_B=0),\ d_{AB}\ (O_A, O_B) < 0),$$

And image I,

$$I=\ (P_{ij}\ (O_i, O_j),\ m_{ij}\ (O_i, O_j),\ d_{ij}\ (O_i, O_j))\ |\ i \neq j, \forall i, j \in I)$$

Where I consists of a list of pair object i and Object j

The matching of Q to image I is defined as a total number of matched pairs in I as show below

Match(Q,I) = number of element for ($Q \cap I$)

5.2. Image Similarity

The Image similarity is used to define the similarity between user query and Image. The image similarity based on Single condition of User Query (without logical operation) and Multiple condition of User query (Logical operation involved) are then determined

5.2.1. Single Condition of User Query (without Logical Operation)

In this single condition of user query, there is no logical operation involved in the user query.

The similarity of image I to query Q is defined as the ratio of number of matched pair, over the total number of pairs in image as show below,

$$S_{sim}(QC_1,\ I_i) = \frac{number\ \ of\ \ element\ \ (Q, I)}{n},$$

Where number of element (Q,I) is the number of matched pair in query matching while n is total number of pairs in images

5.2.2. Multiple Condition of User Query (Logical Operation Involved)

For multiple condition of user query, image Semantics Similarity is determined based on the involvement of logical operation AND, OR and NOT.

5.2.2.1. Logical AND Operation

If there is a or more logical AND operation involved in user query, all of their sub query must be combined to get the object relationship semantic similarity, S_{AND}

$$S_{sim(AND)} = \frac{\sum_{i}^{k}\left(S_{sim}(\mathrm{Q}C_i, \mathrm{I_i}) + S_{sim}(\mathrm{Q}C_j, \mathrm{I_i})\right)}{Total(k)} i \neq j$$

Where $S_{sim}(\mathrm{Q}C_i)$ is semantic similarity of user sub query i and $S_{sim}(\mathrm{Q}C_j, \mathrm{I_i})$ is semantic similarity of user sub query j respectively to the image i while k is total number of sub user query.

5.2.2.2. Logical OR Operation

If there is a logical OR operation involved in user query, the maximum value of the object spatial relationship semantics similarity of sub query is chosen as it represents the closest match between database images and user query.

$$S_{sim(OR)} = \text{Max} \sum_{i}^{k}\left(S_{sim}(\mathrm{Q}C_i, \mathrm{I_i}), S_{sim}(\mathrm{Q}C_j, \mathrm{I_i})\right)$$

Where $S_{sim}\ (\mathrm{Q}C_i, \mathrm{I_i})$ is the semantics similarity for sub query i and k is total number of sub query.

5.2.2.3 Logical NOT Operation

For logical NOT operation, the list of images are the images that are not listed in the candidate objects in Candidate object filters.

$$O_{NOT} = \overline{O_{QC_i}}$$

where $\overline{O_{QC_i}}$ is list of objects that do not found in the O_{QC_i}

5.3. Range of Image Spatial Semantic Similarity

The degree of image spatial similarity ranges from 0 to 1. The value 1 indicates that there is a perfect match of the database images with user query while value 0 indicates that there are dissimilar.

6. CONCLUSION

This paper provides a study of image retrieval work towards narrowing down the ‘semantic gap’. Recent works are mostly lack of semantic features extraction especially in objects spatial relationships semantic and user behavior consideration. Therefore, a new method and approach for extracting objects spatial relationships semantic and representation automatically from images has been proposed in order to bridge the semantic gap besides enhance the high level semantic search and retrieval. There are 8 of main important spatial relationship concept has been introduced. Besides, the objects spatial relationship semantic similarity is also introduced. This work will be enhanced and expanded further to include the object characteristics as well as more abstract high level queries with spatial relationship semantic representation. There is a need of image retrieval system that is capable to interpret the user query and automatically extract the semantic feature that can make the retrieval more efficient and accurate to bridge the semantic gap problem in image retrieval.

7. REFERENCES

1. A. Rosenfeld. “Picture processing by computer”. ACM Computing Surveys, 1(3):147-176, 1969.
2. H. Tamura and S. Mori. “A data management system for manipulating large images”. In Proceedings of Workshop on Picture Data Description and Management, pages 45-54, Chicago, Illinois, USA, April 1977
3. Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying M. “A survey of content-based image retrieval with high-level semantics”. Pattern Recognition, 40(1):262 – 282, 2007
4. R.Datta, D.Joshi, J.Li, J.Z.Wang. “Image Retrieval: Ideas, Inuences, and Trends of the New Age”. ACM Transactions on Computing Surveys, 40(2), 2008
5. Hui Hui Wang, Dzulkifli Mohamad, N.A.Ismail. “Image Retrieval: Techniques, challenge and Trend”. In International conference on Machine Vision, Image processing and Pattern Analysis, Bangkok, 2009
6. Hui Hui Wang, Dzulkifli Mohamad, N.A.Ismail. “Towards Semantic Based Image Retrieval : A Review”. In The 2nd International Conference on Digital Image Processing (ICDIP 2010), Singapore,2010.Proceedings of SPIE, Vol 7546, 2010
7. C.Y.Chang, H.Y.Wang, C.F.Li. “Semantic analysis of real world images using support vector machine”. Expert systems with application : An international journal, 36(7):10560-10569, 2009
8. Najlae Idrissi. “Bridging the Semantic Gap for Texture-based Image Retrieval and Navigation”, Journal Of Multimedia, 4(5):277-283, October 2009
9. M. L. Kherfi, D. Ziou, and A. Benard. “Image retrieval from the World Wide Web: Issues, techniques and systems”. ACM Computing Surveys, 36(1):35-67, 2004
10. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. “Content-based image retrieval at the end of the early years”. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12):1349-1380, December 2000.
11. Zurina Muda, Paul H. Lewis, Terry R. Payne, Mark J. Weal. “Enhanced Image Annotations Based on Spatial Information Extraction and Ontologies”. In Proceedings of the IEEE International Conference on Signal and Image Processing Application, ICSIPA 2009
12. Haiying Guan, Sameer Antani, L. Rodney Long, and George R. Thomas. “Bridging the semantic gap using rnaking SVM for image retrieval”. In Proceedings of the sixth IEEE international Conference on Symposium on Biomedical Imaging: From Nano to Macro, USA 2009.
13. Zhao and Grosky “Bridging the semantic gap in Image retrieval”. Idea Group Publishing Series, Book Chapter 2, pp 14-36 (2002)
14. Wikipedia, http://wikipedia.org. Accessed at June 2010
15. Travel Photography and Professional Stock Phtography-lonely Planet Images, http://www.lonelyplanetimages.com. Accessed at June 2010
16. Photoblog-Your Life in Photos, http://www.photoblog.com. Accessed at June 2010
17. Photo Blogging made Easy-fotopages, http://www.fotopages.com. Accessed at June 2010
18. Inote : Image Annotation in Java, http://www.iath.virginia.edu/inote/. Accessed at June 2010
19. Facebook, http://www.facebook.com. Accessed at June 2010
20. Bradshaw B. “Semantic based image retrieval: A probabilistic approach”. In Proceedings of the Eight ACM International Conference on Multimedia, United States, 2000

21. Y. Mori, H. Takahashi, and R. Oka. “Image-to-word transformation based on dividing and vector quantizing images with words”. In Proceedings of the First International Workshop on Multimedia Intelligent Storage and Retrieval Management, 1999
22. P Duygulu, K Barnard, N de Fretias, and D Forsyth. “Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary”. In Proceedings of the European Conference on Computer Vision, pages 97–112, 2002
23. Jeon,J.,Laverenko,V.,Manmatha, R “ Automatic image annotation and retrieval using cross-media relevance models". In:Proceedings of international ACM Conference on Research and Development in Information Retrieval, 2003
24. Lavrenko,V., Manmatha,R.,Jeon,J:”A model for learning the semantic of pictures”. In proceeding of Advances in Neural Information Processing Systems, NIPS, 2003
25. Mittal A, Cheong LF. “Framework for synthesizing semantic-level indexes”. Multimedia Tools and Applications 20(2):135–158.2003
26. L. Hollink, G. Nguyen, G. Schreiber, J. Wielemaker, and B. Wielinga, "Adding spatial semantics to image annotations," In 4th International Workshop on Knowledge Markup and Semantic Annotation, 2004
27. Belkhatir M. “An operational model based on knowledge representation for querying the image content with concepts and relations”. Multimed Tools Application, 43:1–23. Springerlink Netherlands, 2009
28. Jian Hu, Gang Wang, Fred Lochovsky, Jian-Tao Sun, Zheng Chen. “Understanding User’s Query Intent with Wikipedia”. Proceedings of the 18th international conference on World wide web. ACM. 2009
29. E. Chang, K. Goh, G. Sychay, G. Wu, CBSA: CBSA: “Content-based Soft Annotation for Multimodal Image Retrieval using Bayes Point Machines”, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 26-38, 2003.
30. Hong Fu. “Attention driven image interpretation, annotation and retrieval”. PhD thesis. The Hong Kong Polytechnic University. 2006