Performance and Analysis of the Automated Semantic Object and Spatial Relationships Extraction in Traffic Images

Wang Hui Hui

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

Abstract

Extraction and 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 which contributes to richer querying and retrieval facilities. This paper discusses the performance of the automated object spatial relationships semantic information extraction as proposed. Experiments have been conducted to demonstrate that the proposed automated object spatial relationship semantic extraction is succeeded to capture the semantic spatial relationship features in the images.

Keywords: *Semantic Extraction, object spatial relationships, image retrieval.*

1. Introduction

There are various ways used to represent the semantic features of the images. However, using proper image representation model and the image contents expression is the premise and the basic need of the image semantic retrieval [1]. 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 (color, 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. Describing images in semantic terms is an important and challenging task that needed to carry out to fulfill human satisfaction [2] and defining a semantic meaning and representation of the input query in describing user's needs remain as major challenges [3].

The semantic content representation has been identified as an important issue to bridge the semantic gap in visual information access [4]. The semantic features especially the semantic object and their semantic spatial relationship features in the images are not fully captured and extracted [5-7]. It often leads to unsatisfactory search results [8]. 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. In addition, the computer usually processes semantic similarity based on low-level feature similarity, however the user queries are supposed to be based on semantic similarity [9]. Current semantic based image retrieval are either based on visual features or it is measured the image similarity based on semantic matching instead than semantic similarity [6, 7, 10 , 11].

2. Related Works

Describing images in semantic terms is an important and challenging task to achieve more intelligent and user friendly system. Besides, representing 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.

Detailed study on the semantic extraction techniques exploring the strength and the weaknesses of the existing semantic extraction techniques has been described in Wang *et al.,* 2010. However, the summary of the comparison and extended review techniques are shown in Table 1.

ACSIJ Advances in Computer Science: an International Journal, Vol. 2, Issue 4, No.5 , September 2013
ISSN : 2322-5157
www.ACSIJ.org

Table 1: Comparison of the various approaches Reviewed

Researchers/ Characteristics	Image semantic	Object/ region semantic	Spatial relationship semantic	Spatial similarity measurement	Multiple Spatial relationship query
Frankel *et al.*, 1995 [12]	Yes	No	No	No	No
Minka, 1996 [13]	Yes	No	No	No	No
Rui *et al.*, 1997[14]	Yes	Yes	No	No	No
Inotes, 1998[15]	Yes	Yes	No	No	No
Laaksonen *et al.*, 1999[16]	Yes	No	No	No	No
Vasconcelos and Lippman, 1999[17]	Yes	No	No	No	No
Mori *et al.*, 1999[18]	Yes	Yes	No	No	No
Bradshaw, 2000[19]	Yes	No	No	No	No
Fotopages, 2000[20]	Yes	No	No	No	No
Photoblog, 2000[21]	Yes	No	No	No	No
MacArthur *et al.*,2000[22]	Yes	No	No	No	No
Tong and Chang, 2001[23]	Yes	No	No	No	No
Wikipedia, 2001[24]	Yes	No	No	No	No
Zhao and Grosky, 2001[25]	Yes	Yes	No	No	No
Duygulu *et al.*, 2002[26]	Yes	Yes	No	No	No
Jeon *et al.*,2003[27]	Yes	Yes	No	No	No
Lavrenko *et al.*,2003[28]	Yes	Yes	No	No	No
Mittal and Cheong, 2003[29]	Yes	Yes	No	No	No
Facebook, 2004[30]	Yes	Yes	No	No	No
Flicker, 2004[31]	Yes	No	No	No	No
Hollink *et al.*, 2004[32]	Yes	Yes	Limited to interclass relationships	No	Limited to 2 objects (one pair of relation) only
Wang and Khan, 2006[33]	Yes	Yes	No	No	No
Liu, 2007[34]	Yes	Yes	No	No	No
Min *et al.*, 2008[35]	Yes	No	No	No	No
Muda, 2009[5]	Yes	Yes	Limited to interclass relationships	No	Limited to 2 objects (one pair of relation) only
Belkhatir, 2009[7]	Yes	Yes	Limited to interclass relationships	No	Limited to 2 objects (one pair of relation) only
Chen, 2010	Yes	No	No	No	No
Proposed	Yes	Yes	Yes	Yes	Yes

For the reviewed semantic extraction, there are 3 categories of semantic extraction, which are manual, semi-automatic (involved human interference or relevance feedback) and automatic. 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 and it is easy to support user queries in text. However it is a labour intensive and tedious process. So, researchers are moving towards automatically automatic extraction of the image semantic content. For the semi automatic semantic extraction approaches that involve human interference, they are time consuming and inconsistent.

Image semantic extraction capability is provided by all of the techniques reviewed where they are only able to retrieve similar images which have the whole image semantics and does not indicate which part of the image give rise to which words, as computer does not indicate which region corresponds to specific semantic concepts. So it is not explicitly object recognition. They lack the ability to find the object semantics in images.

Object/region semantic extraction is provided by researches [14, 15, 18, 25-30, 33-34] from manual semantic extraction [15, 30] and from semi/automatic semantic extraction [14, 18, 25-29, 33]. Even though the object or region semantics can be captured, the extraction of spatial relational semantic descriptors is often neglected. They do not take into consideration the spatial relational semantics among objects in the images that might affect the performance of the image retrieval.

Limited object spatial relationship capability is provided by researches [32, 5, 7,]. However, there are still some false objects spatial relationships extraction concept (Example in Figure 1, the objects A and B are supposed to have Front/Back spatial relationship but it is extracted as Left/right relation concept [32, 7] and it is extracted it as above-right relation [5].

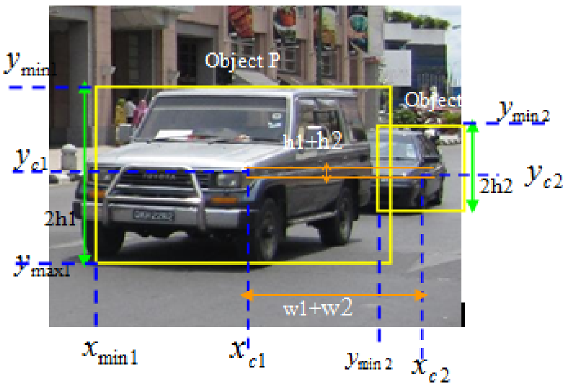


Figure 1: False objects spatial relationships semantic extraction

The semantic knowledge representation models have better representing and matching semantics compared the text representation. However, semantic objects and spatial relationships are not applicable. For ontology and metadata language approaches, there are some limited and some false representation on object spatial relationship semantic representation, besides that it is difficult to do the semantic similarity with this representation for image retrieval.

Representation of spatial relations semantics among objects are also important as it can convey important information about the image and further increase the confidence in image understanding which contributes to richer querying and retrieval facilities. The text representation [12, 15, 19-21, 24, 30, 31] and the semantic knowledge representation models [37-39] do not support semantic objects and spatial relationships. However, there are some false representations on object and spatial relationships for Ontology and metadata language semantic representation and it is hardly to do the semantic similarity with this representation for image retrieval.

3. Proposed Solutions

The automated semantic object relationships extraction has been proposed to extract the semantic object spatial relationship information in the images automatically. The data flow of the Semantic Object Relationship Extraction process is shown in Figure 2.

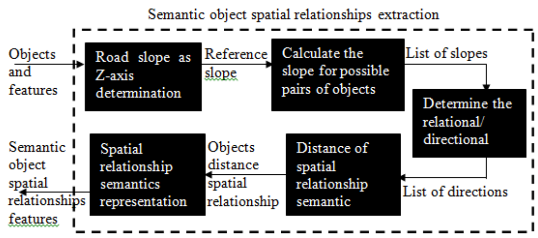


Figure 2 Object Spatial Relationship Semantic Extraction

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 object spatial relationship semantics has been proposed as below,

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

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

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.

The details of the semantic object relationship extraction can be referred to work in [2].

4. Experiments

This experiment shows the ability of the proposed semantic extraction and representation method in extracting the semantic object and their semantic spatial relationships automatically.

4.1 Experiments and Setting

Two experiments were carried out for semantic object and semantic spatial relationship extraction in low complex images. The results are expressed in terms of semantic object and semantic spatial relationships similarity.

To verify the semantic object and their semantic spatial relationship in the images, the extracted semantic features by the proposed semantic extraction and representation were compared with the semantic features interpreted from the user on the same images. An extraction is considered as correct if the semantic features by the proposed algorithms are same as the semantic features defined by users and return with the spatial relationship similarity of 1. The semantic colour of object representation is considered as correct if the semantic colour object by the proposed algorithms is same as the semantic colour object defined by users and also returns with the semantic object similarity of 1.

4.2 Results and Discussion

The results of the object detection experiments are summarized and shown in Table 2.

Table 2: The Experiment Results of the proposed Semantic Extraction and Representation Method

Experiment	Successful Extraction	False Extraction	Successful rate
I (Semantic Objects)	117	13	90%
II(Semantic Spatial relationships)	115	14	88%

The proposed method has proven to be successful in extracting the semantic features of images. The results accuracy is 90% and 88% respectively for the semantic objects and semantic spatial relationships in the images. Thus, the proposed semantic extraction and representation method is acceptable to show the robustness in extracting the semantic features for low complex scenes.

ACSIJ Advances in Computer Science: an International Journal, Vol. 2, Issue 4, No.5 , September 2013
ISSN : 2322-5157
www.ACSIJ.org

Some of the semantic extraction and representation results from proposed method is shown in Figure. 3

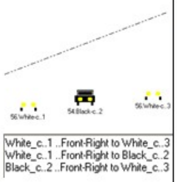
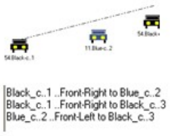
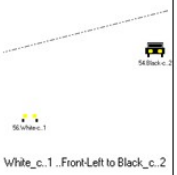
No	Original image	Human Perception	Semantic Features and representation	Image Similarity	
				Semantic Object	Semantic Spatial Relationship
i		1.Black car Front/Right to white car 2.White car front/right to White car 3.white car Front/right black car		1	1
ii		1. Black car front/right to Blue car 2. Black car front/right to Black car 3. Blue car Front/left to Black car		1	1
iii		1. White car front/left to Black car		1	1

Figure 3(a): Experiment results for semantic extraction and representation-successful semantic feature extraction

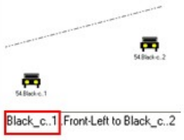
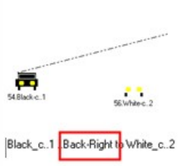
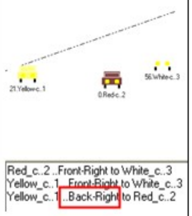
No	Original image	Human perception	Semantic Features and representation	Image Similarity	
				Semantic Object	Semantic Spatial Relation-ship
i		1. Brown car Front/Left to Black car		0.5 (the brown car instead black car)	1
ii		1. Black car front/right to White car		1	0.5 (front-right relation instead black right relation)
iii		1. Red car front/right white car 2. Yellow front –right white car 3. Yellow front/right to Red car		1	0.66 (Front-right relation instead of back right relation)

Figure 3(b): Experiment results for semantic extraction and representation - false semantic features extraction

The results of Figure 3(a) show that the proposed algorithm has successfully extracted the semantic objects and their semantic spatial relationships features automatically and return the similarity value of 1 for both semantic object similarity and spatial relationships similarity.

There are some false semantic features extractions as shown in Figure 3(b), indicated by red rectangle. The false semantic objects are extracted (Figure 3(b),(i)) due to the noise created from the object detection process. The false semantic spatial relationship are extracted (Figure 5.3(b), (ii-iii)) due to the noise created from the object detection.

There is no dataset benchmarking available in the existing semantic extraction and representation. Their approaches are based on interclass relationship (objects of interests with background). Sample datasets of the existing approach [7] with the queries are shown in Figure 4.

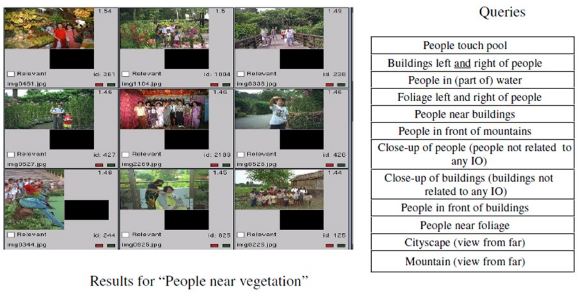


Figure 4 Sample dataset used [7].

The images are only able to indicate that the person (people object class) is near to vegetables (vegetable object class or background of the image) instead of the spatial relationships among people object class. This research focus is on semantic object spatial relationship that exists in the images and the dataset available is not suitable to be used in this research focus. User relevance feedbacks are involved during the retrieval process. In addition, they retrieved images without measuring any similarity.

Even if there were same datasets available, benchmarking could not be applied in image retrieval due to the different parameters and queries used by different researchers in the experiments [34].

5. Conclusions

In conclusion, experiments have been carried out and it is proved the proposed semantic extraction and representation method is acceptable to show the robustness in extracting the semantic features for natural traffic scenes. This method can be integrated with visual query for the semantic based image retrieval to retrieve the images that are conform to human perception.

References

[1] Yu, X. H., Xu, J. H. (2008). The Related Techniques of Content-based Image Retrieval. *Proceedings of International symposium on computer science and computational technology.* 1, 154-158.

[2] Wang,H.H., Dzulkifli Mohamad & N.A. Ismail (2010) Semantic Gap in CBIR: Automatic Objects Spatial relationships Semantic Extraction and Representation. International Journal Of Image Processing (IJIP), Volume (4): Issue (3)

[3] Russ, J. C. (1999). The Image processing handbook. Third edition. A CRC Handbook published in cooperation with IEEE press.

[4] Wang, L., Khan, L. (2006). Automatic image annotation and retrieval using Weighted Feature Selection. Journal Multimedia Tools an Application, 29(1), 55-71.

[5] Muda, Z., Lewis, P. L., Payne, T. R., Weal, M. J. (2009). Enhanced Image Annotations Based on Spatial Information Extraction and Ontologies. Proceedings of the IEEE International Conference on Signal and Image Processing Application, 18-19 November, Kuala Lumpur, Malaysia,173-178.

[6] Hollink, L., Nguyen, G., Schreiber, G., Wielemaker, J., and Wielinga, B. (2004). Adding spatial semantics to image annotations. The 4th International Workshop on Knowledge Markup and Semantic Annotation, 7-11 November, Hiroshima, Japan, 31-40

[7] Belkhatir, M. (2009). An operational model based on knowledge representation for querying the image content with concepts and relations. Multimedia Tools Application, 43(1), 1–23.

[8] Zha, Z. J, Yang, L. J, Mei, T., Wang, M, Wang, Z.F, Chua, T. S., Hua, X.S. (2010). Visual Query Suggestion: Towards Capturing User Intent in Internet Image Search. *ACM Transactions on Multimedia Computing, Communications, and Applications.* 6(3),13(1-19)

[9] Agrawal, R. (2009). N*arrowing Down the Semantic Gap between Content and Context using Multimodal Keywords*. PhD thesis. Wayne State University.2009

[10] Pratikakis, I., Bolovinou, A., Gatos, B., and Perantonis, S. (2011).Semantics Extraction from Images. *Multimedia Information Extraction*, 50–88.

[11] Yoon, J. W. (2011). Searching images in daily life. *Library and Information Science Research 33, 269–275.*

[12] Frankel, C., Swain, M.J., and Athitsos, V.(1996). WebSeer : An image search engine for the world wide web. Technical Report. University of Chicago.

[13] Minka, T.(1996). *An image database browser that learns from user interaction*. MIT Media Laboratory Technical Report #365, Massachusetts Institute of Technology

[14] Rui, Y., Huang, T. S., and Mehrotra, S. (1997). Content-based image retrieval with relevance feedback in mars, *Proceeding of international conference image processing.* 26-29 October. Texas, USA, 815-818.

[15] Inote: Image Annotation in Java (1998), Retrieved 3 June 2010. Website :http://www.iath.virginia.edu/inote/.

[16] Lavrenko,V., Manmatha,R., and Jeon, J. (2003). A model for learning the semantic of pictures. Proceeding of Advances in Neural Information Processing Systems, 8-13 December. British Columbia, Canada, 80-89.

[17] Vasconcelos, N., and Lippman, A. (1999). Learning from user feedback in image retrieval systems. *Proceeding of the 12th Conference on Advances in Neural Information Processing Systems*, 29 November-4 December. Denver, Colorado, 76-89.

[18] Mori, Y., Takahashi, H., and Oka, R. (1999). Image-to-word transformation based on dividing and vector quantizing images with words. Proceedings of the First International Workshop on Multimedia Intelligent Storage and Retrieval Management.

[19] Bradshaw B. (2000). Semantic based image retrieval: A probabilistic approach. Proceedings of the Eight ACM International Conference on Multimedia, 30 October-3 November. USA, 167-176

[20] Photo Blogging made Easy-fotopages (2000), Retrieved 3 June 2010. Website: http://www.fotopages.com/cgi-bin/account.pl?page=3

[21] Photoblog-Your Life in Photos (2000), Retrieved 3 June 2010. Website : http://www.photoblog.com/create/

[22] MacArthur, S.D., Brodley, C.E. and Shyu, C. R (2000). Relevance feedback decision trees in content- based image retrieval. *Proceeding of IEEE Workshop on Content-based Access of Image and Video Libraries*. Hilton Head Island, SC, 68-72.

[23] Tong.S., and Chang, E. (2001). Support vector machine active learning for image retrieval. *Proceeding of ACM Multimedia*, 30 September–5 October. Ottawa, Canada, 101-118

[24] Wikipedia (2001), Retrieved 3 June 2010. Website: http://commons.wikimedia.org/wiki/Special:Upload

[25] Zhao, R., and W. I. Grosky. (2001). Bridging the Semantic Gap in Image Retrieval. Distributed Multimedia Databases: Techniques and Applications, Idea Group Publishing, Hershey, Pennsylvania, 13-36.

[26] Duygulu, P., Barnard, K., Fretias, N., and Forsyth, D. (2002). Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. Proceedings of the European Conference on Computer Vision, 27 May- 2 June. Copenhagen, 97–112.

[27] Jeon,J., Laverenko,V., and Manmatha, R. (2003), Automatic image annotation and retrieval using cross-media relevance models. Proceedings of international ACM Conference on Research and Development in Information Retrieval. 28 July- 1 August. Toronto, Canada,119-126

[28] Lavrenko,V., Manmatha,R., and Jeon, J. (2003). A model for learning the semantic of pictures. *Proceeding of Advances in Neural Information Processing Systems*, 8-13 December. British Columbia, Canada, 80-89.

[29] Mittal, A., Cheong, L. F.(2003). Framework for synthesizing semantic-level indexes. Multimedia Tools and Applications 20(2),135–158.

[30] Facebook (2004), Retrieved 3 June 2010. Website: http://www.facebook.com.

[31] Flicker (2004.), Retrieved 3 June 2010. Website: http://www.flickr.com/photos/upload/done/

[32] Hollink, L., Nguyen, G., Schreiber, G., Wielemaker, J., and Wielinga, B. (2004). Adding spatial semantics to image annotations. The *4th International Workshop on Knowledge Markup and Semantic Annotation,* 7-11 November, Hiroshima, Japan, 31-40

[33] Wang, L., Khan, L. (2006). Automatic image annotation and retrieval using Weighted Feature Selection. *Journal Multimedia Tools an Application*, 29(1), 55-71.

[34] Liu, Y. (2007). Region-Based Image Retrieval with High-Level Semantics. PhD Thesis, Monash University.

[35] Min, Y., Rong, A., and Zhen, Y. (2008). Research on Methodology of Image Semantic Understanding. *Proceeding of SPIE*. 6623(1-6).

[36] Chen, Z. J. (2010). Semantic Research on Content-Based Image Retrieval. Multimedia Technology, *International Conference on Multimedia Technology,* 29-31 October, Ningbo, China. 1-4.

[37] Meghini, C., Sebastiani, F., and Straccia, U.(1997). The terminological image retrieval model. *Proceeding of 9th international conference on image analysis and processing* (ICIAP'97).17-19 September. Florence, Italy, 156–163.

[38] Cavazza, M., Green, R. J., and Palmer, I. J.(1998). Multimedia semantic features and image content description. *Proceeding multimedia modeling*, 12-15 October 1998, Lausanne, 39–46

[39] Colombo, C., Bimbo, D. A., and Pala, P. Semantics in visual information retrieval. *IEEE Multimedia*, 6(3), 38–53.