

# THE STUDY ON THE SEMANTIC IMAGE RETRIEVAL USING THE COGNITIVE SPATIAL RELATIONSHIPS IN THE SEMANTIC WEB

Hyunjang Kong<sup>1</sup>, Myunggwun Hwang<sup>1</sup>, Kwansang Na<sup>3</sup> and Pankoo Kim<sup>2</sup>  
*<sup>1</sup>Dept. of Computer Science, Chosun University, Gwangju 501-759, Korea; <sup>2</sup>Corresponding Author, Dept. of Computer Engineering, Chosun University, Gwangju 501-759, Korea; <sup>3</sup>Dept. of Computer Science and Engineering, Korea University, Seoul 136-701, Korea*

**Abstract:** In present day, there are a number of image data in the web because of the development of the image acquisition devices. So, many researchers have been studying about the image retrieval and management. Keyword matching, contents-based and concept-based methods are the basic studies for the image retrieval. In this paper, we suggest the new image retrieval methodology using the cognitive spatial relationships between the objects in the image. There were the similar studies already using the spatial relationships. However, the studies have the limitations and don't give the good search results. We think to need the new methodology for representing the spatial relationships. It is the cognitive spatial relationships. In our study, we newly define the cognitive spatial relationships and apply it to the image retrieval system. At the result, we realized that our methodology makes possible the semantic image retrieval.

**Key words:** Cognitive Spatial Relationships, Ontology, Semantic Web, Image Retrieval

## 1. INTRODUCTION

There are a huge number of data in the web. Nowadays, users want to search the information more rapidly and correctly. Until now, most information is the document types so the methodology for the information retrieval is based on the text matching.[2][3][4] The methodology provides the sufficiently correct information to users from the web. However, as the image acquisition technologies such as digital camera, a scanner, cellular

phone camera, etc, have improved many images have being existed in the web environment. And the existing retrieval system just using the keyword matching about the image has the limitation. Because most images have the contents, they are generally stored with simple annotations. Thus, many researchers have been studying the methodologies for image retrieval. The basic methodologies studied until now are the contents-based, concepts-based and ontology-based methodologies.[5][6][7] However, the methodologies still don't give the good results to the users. So, we suggest the new methodology for the semantic image retrieval in this paper. Our idea is the cognitive spatial relationships. We focus on the spatial relationships and we define the cognitive spatial relationships. And then we build the spatial ontology based on the cognitive spatial relationships, user's research, WordNet and OXFORD dictionary. Finally we designed the new image retrieval system using the spatial ontology.

In the 2nd chapter, we introduce the related works – information retrieval system, semantic ontology-based image retrieval system and spatial description logics. Then in chapter 3, we explain the background knowledge of our study and the cognitive spatial relationship, and how to build the spatial ontology. In chapter 4 and 5, we describe our system based on the cognitive spatial relationship and experimental results and evaluation of our system. In the end of this paper, we conclude our study and suggest the future works.

## **2. RELATED WORKS**

### **2.1 Information Retrieval System**

Information Retrieval systems are used to store, maintain, search, and retrieve the information items. The information items could be text documents, images, sounds or videos. In the information retrieval system, it is very important to have efficient data structures, fast search tools, and effective information retrieval methods, especially if the amount of data is large. Generally, most information retrieval systems utilize indexing methods to improve the search efficiency. Indexing is the process of assigning descriptive terms to information items for retrieval purposes. Indexing is a very important and difficult task and every information item is stored in a traditional information retrieval system with the index. However, documents often lose their semantics when represented by just simple index terms. Therefore, normally search of the documents using simple keywords

results in retrieving irrelevant documents, which is the case with most web search engines.[8]

## 2.2 Ontology-based Image Retrieval

As mentioned in above section, the traditional information retrieval systems have the mismatch problem between the terminologies. For solving the problem, many researchers have studied to apply the ontology theory. A great many works show that ontologies could be used not only for annotation and precise information retrieval [9], but also for helping the user in formulating the information need and the corresponding query. This is important especially in applications where the domain semantics are complicated and not necessarily known to the user. Furthermore, the ontology-enriched knowledge base of image metadata can be applied to constructing more meaningful answers to queries than just hit-lists.

The major difficulty in the ontology-based approach is the extra work needed in creating the ontology and the detailed annotations. We believe, however, that in many applications this price is justified due to the better accuracy obtained in information retrieval and to the new semantic browsing facilities offered to the end-user. We are trying to implement semantic techniques to avoid so much hard work with the ontology building--the trade--off between annotation work and quality of information retrieval can be balanced by using these less detailed ontologies and annotations. Although this approach could address the mismatch problem between the terms, it is still not suitable for image retrieval system because they did not consider the features of the image data. Therefore, we will not get good results in the ontology-based image retrieval system.

## 2.3 The Description Logic $ALC(D_{rcc8})$

The *Region Connection calculus*  $RCC-8$ [1] is a language for qualitative spatial representation and reasoning where the spatial regions are regular subsets of a topological space. The regions themselves do not need to be internally connected i.e. a region may consist of different disconnected pieces.

As the concrete domain in  $ALC(D_{rcc8})$ ,  $\Delta^D_{rcc8}$  is the set of all the non-empty regular closed subsets of the topological space  $R^2$ .  $\Phi^D_{rcc8}$  is obtained by imposing a union, intersection, composition and converse operations over the set of the elementary binary relationships between the regions i.e.(PO, NTPP, TPP, EQ,  $TPP^{-1}$ ,  $NTPP^{-1}$ , EC, DC) where the intended meaning of the elements are respectively Proper Overlap, Non Tangential Proper Part, Tangential Proper Part, External Connection, and DisConnected.

### 3. OUR APPROACH

#### 3.1 Background Knowledge of the Cognitive Spatial Relationships

In the existing image retrieval system, if the system uses the spatial relationships between object, the system firstly extracts the edge of the objects. Secondly, the system use the spatial relationships based on the regions of the objects. In such a case, the biggest problem is either the spatial relationships do not have the semantic meaning or system defines the spatial relationships incorrectly. Figure 1 explains the general spatial relationships using in the existing image retrieval system.

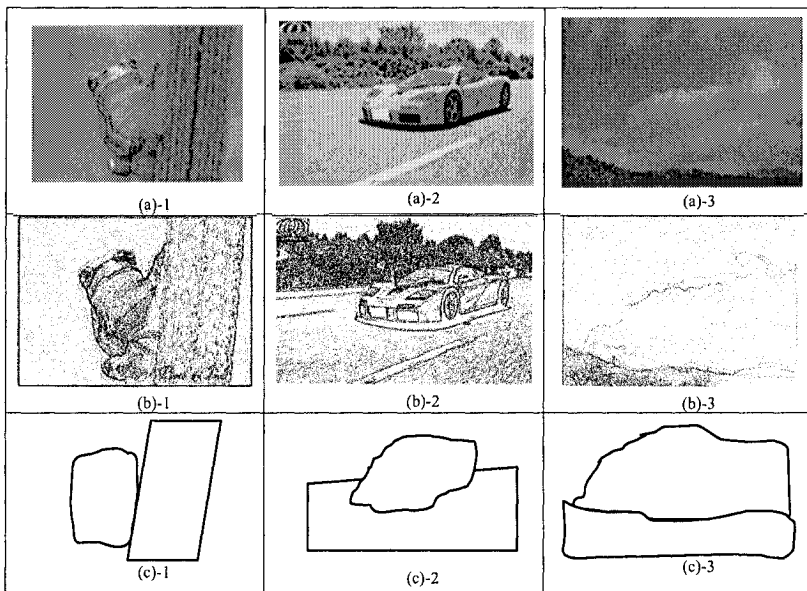


Figure 1. (a) images are the original images, (b) images are the edges of the objects, (c) images are the regions of the objects

In above three images, if the system represents the spatial relationships, it is like as:

- First images((a)-1,(b)-1,(c)-1) are represented the 'flog-connected-tree'.
- Second images((a)-2,(b)-2,(c)-2) are represented the 'car-part of-road'.
- Third images((a)-3,(b)-3,(c)-3) are represented the 'mountain-connected-cloud'.

However, the users recognize different with the system. The cognitive spatial relationships recognized by users are like as:

- First image((a)-1) means the 'flog-connected-tree'
- Second image((a)-2) means the 'car-connected-road'
- Third image((a)-3) means the 'cloud-disconnected-mountain'

Nowadays, the users want that the machine also thinks and processes like the human. It is the basic idea of the semantic web. In the semantic web environment, users make search requests for images based on their visual impressions. If the system stores the image metadata using the region-based spatial relationships, the system will provide the wrong and senseless results to users. Therefore, we try to define the cognitive spatial relationships newly and design the image retrieval system based on our study.

### 3.2 Definition of the Cognitive Spatial Relationships and Construction of the Spatial Ontology

In our study, we used the research for defining the cognitive spatial relationships. We prepare the 200 images containing the objects and spatial relationships between them. And then, we examine the spatial relationships recognized by users when users look at the images. At the result of the research, the cognitive spatial relationships represent the basic three kinds of relationships. Throughout the research, we realized that most images are represented by the 'connect', 'disconnect' and 'partof' relationships. Figure 2 illustrates the model of the cognitive spatial relationships comparing with the RCC-8.

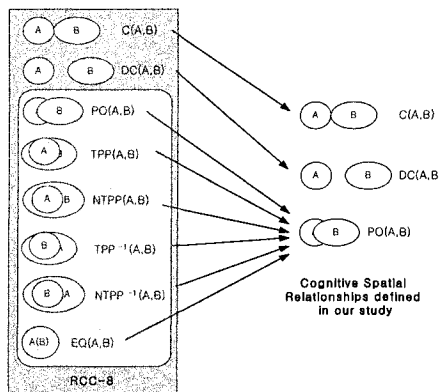


Figure 2. The model of the cognitive spatial relationships

In our study, another significant feature is to build the spatial ontology based on the cognitive spatial relationships and user's research. For grasping the spatial verbs, we examine them using the experimental document containing the sample images. The contents and results of the research are like as:




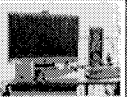

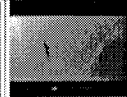
										
image #3	image #33	image #40	image #55	image #59	image #141					
Research results about the images(spatial relationships 1)connect 2)partof 3)disconnect, C_SR:cognitive spatial relationships, sp:spatial relationships)										
number	C	SR	sp	Answers of the Researchers						
3	1	2		lying	lying	lie	lying	lying	lie	sit
33	2	2		swim	search	swim	swim	cross	look pretty	swim
40	1	1		kiss	kiss	kiss	kiss	kiss	love	play
55	3	3		on the left	bigger	beside	behind	left of	beside	beside
59	3	2		wait	in front	waiting	stand	in front	look	stand
141	2	2		fly	soars	fly	fly	fly	fly	fly

Figure 3. The parts of the Experimental Document

In the results of the research, we realized that most users have the similar feeling and use the similar spatial verbs to represent the images.

Thus, we build the spatial ontology based on the cognitive spatial relationships and spatial verbs. The most important parts in the spatial ontology are the cognitive spatial relationships defined in section 3.1 and the spatial verbs. And we adopt the WordNet and OXFORD Dictionary to make more semantic ontology. Figure 4 shows the architecture of the spatial ontology based on the cognitive spatial relationships.

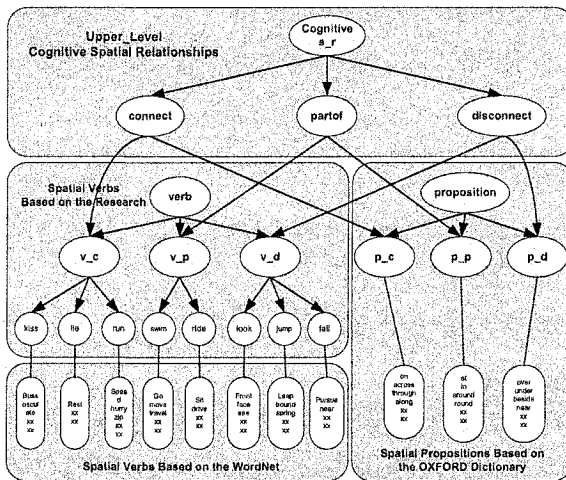


Figure 4. The architecture of the spatial ontology

In figure 4, the cognitive spatial relationships are situated in the top level and the second level consists of the two parts – spatial verbs and spatial propositions. The bottom level is added by the instances based on the WordNet and OXFORD Dictionary. The significant fact newly known throughout the research is that not only the verbs but also the propositions are very important to present the cognitive spatial relationships. Therefore, we made the spatial verbs part based on the WordNet and the spatial propositions part using the OXFORD Dictionary. The table 1 and 2 show the instances used in spatial ontology based on the WordNet and OXFORD Dictionary.

Table 1. Research words match with the WordNet words

Cognitive spatial relationships	Research words	WordNet matching words
connect	Attach	Connect, link, tie, link up, fasten, touch, adjoin, meet, contact
connect	Kiss	Buss, osculate
disconnect	Chase	Chase after, trail, tail, tag, give chase, god, go after, pursue, follow
disconnect	Jump	Leap, bound, spring
partof	Float	Drift, be adrift, blow, swim, transport
partof	Hide	Conceal, shroud, enshroud, cover, obscure, blot out, obliterate, veil

Table 2. The spatial propositions

The spatial propositions based on OXFORD Dictionary	
connect	On, along, across, through
disconnect	Over, under, above, below, by, beside, near, before, behind
partof	At, in, around, round

In our study, firstly we defined the cognitive spatial relationships, and secondly built the spatial ontology based on the cognitive spatial relationships. The cognitive spatial relationships are written by the ontology language for applying to the image retrieval system. In the image retrieval system, the super user can store the image data using three cognitive spatial relationships and the end user can search the image accessing the spatial ontology. We designed the system to be able to query using the natural language. We serialize the spatial ontology using OWL and table 3 shows the OWL syntax about the spatial ontology.

Table 3. The part of the spatial\_ontology.owl

```

.....
<owl:Class rdf:ID="CognitiveSpatialRelationship">
  <rdfs:subClassOf rdf:resource="#SpatialRelationship" />
</owl:Class>
<owl:Class rdf:ID="connect">
  <rdfs:subClassOf rdf:resource="#CognitiveSpatialRelationship" />
</owl:Class>
.....
<owl:Class rdf:ID="verb">
  <rdfs:subClassOf rdf:resource="#word" />
</owl:Class>
<owl:Class rdf:ID="verbConnect">
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#verb" />
    <owl:Class rdf:about="#connect" />
  </owl:intersectionOf>
</owl:Class>
<owl:Class rdf:ID="Kiss">
  <rdfs:subClassOf rdf:resource="#verbConnect" />
</owl:Class>
<Kiss rdf:ID="buss" />
<Kiss rdf:ID="osculate" />
.....
<owl:Class rdf:ID="proposition">
  <rdfs:subClassOf rdf:resource="#word" />
</owl:Class>
<owl:Class rdf:ID="propositionConnect">
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#proposition" />
    <owl:Class rdf:about="#connect" />
  </owl:intersectionOf>
</owl:Class>
<propositionConnect rdf:ID="on" />
<propositionConnect rdf:ID="across" />
.....

```

#### 4. THE IMAGE RETRIEVAL SYSTEM APPLYING THE COGNITIVE SPATIAL RELATIONSHIPS

Our system consists of three parts.

- Content provider interface part – Content provider stores and manages the images
- End user interface part – End user retrieval the images
- Ontology part – Domain and spatial ontologies are in this part.

Figure 5 illustrates the architecture of our system.

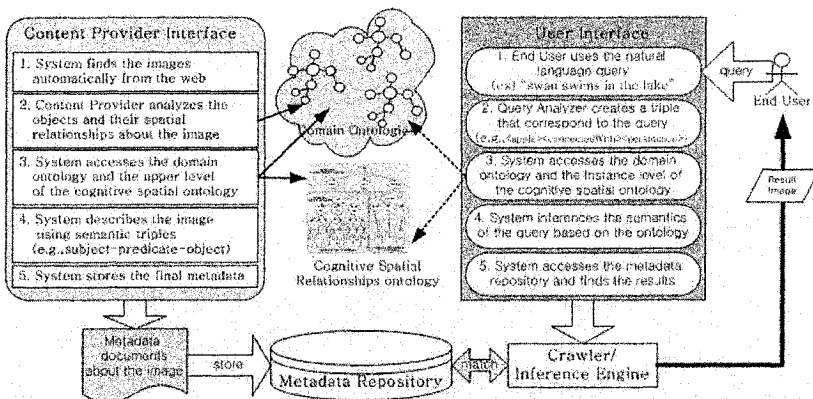


Figure 5. The architecture of the experimental image retrieval system



Our system has two significant features. One is the application of the spatial ontology constructed based on the cognitive spatial relationships. Another is the capability to process the user query using the natural language. We expect improvement of semantic image retrieval throughout our study.

## 5. EXPERIMENTAL RESULTS AND EVALUATIONS

For evaluating our system, we test our system using several test beds. The test systems are Google, Yahoo and our system. Google and Yahoo use the big category that is the kind of the ontology. The sample queries for testing are like as:

1. *Only one word query – e.g. swan*
2. *Two words query – e.g. swan and lake*
3. *Query containing the spatial relationships – e.g. swan in the lake*
4. *Natural Language query containing the spatial verbs – e.g. swimming swan*
5. *Natural Languages query containing the spatial verbs and proposition – e.g. swan swims in the lake*

For testing, we prepared the related sample images. We measure the precision of the search results in three test systems. Because three systems have the different image resources, we measure the accuracy of each system. For measuring the accuracy, we use the simple formula showing as below:

$$\text{Accuracy} = \frac{\text{Correct images matched with the query}}{\text{All images searched throughout the system}}$$

Table 4 shows the result about the test.

Table 4. Expremental results

	1	2	3	4	5
Google	27/200 : 0.135	20/200 : 0.100	67/200 : 0.335	10/30 : 0.333	25/52 : 0.481
Yahoo	72/200 : 0.360	17/48 : 0.354	49/200 : 0.245	7/25 : 0.28	12/26 : 0.462
Our System	42/109 : 0.385	37/60 : 0.617	37/39 : 0.949	35/37 : 0.946	35/35 : 1.000

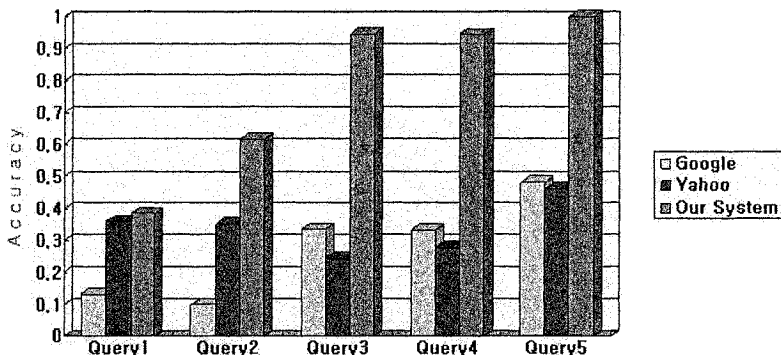


Figure 6. The graph representation of the experimental results

In table 4, we realized that there is no big difference between our system and other systems about the simple text query – query 1 and query 2. However, we could know that our system give the excellent result about the complex queries – query 3, 4 and 5 comparing with other systems. At the result, we can approach more semantic image retrieval and the natural language query processing. Figure 7 and Figure 8 show the results of our system about the test query 4 and 5.

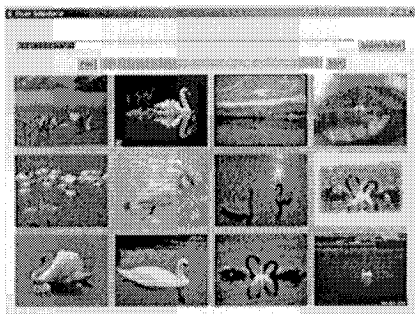


Figure 7. The result about the query 4

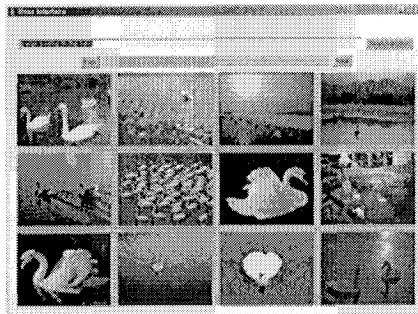


Figure 8. The result about the query 5

## 6. CONCLUSION

The main features of our study are the definition of the cognitive spatial relationships and construction of the spatial ontology using the spatial verbs and propositions. Thus, we realized that the natural language query processing and the semantic image retrieval are possible based on our idea.

However, we have also the limitation that the content provider needs much time to annotate the image semantically throughout our system. It remains our future study. In conclusion, our study presents the vision of the semantic image retrieval and the natural language query processing.

## ACKNOWLEDGEMENT

This study was supported(in part) by research funds from Chosun University, 2003.

## REFERENCES

1. Deborah L. McGuinness, Daniele Nardi, Peter F. Patel-Schneider, "The Description Logic Handbook: Theory, implementation, and applications", Cambridge University Press, January 2003.
2. Arnold W.M. Smeulders, "Content-Base Image Retrieval at the End of the Early Years", IEEE Transactions on pattern analysis and machine intelligence, Vol. 22, No. 12, December, 2000.
3. Ching-chih Chen, and James Z. Wang, "Large-scale Emperor Digital Library and Semantics-sensitive Region-based Retrieval", Proceeding of the International Conference on Digital Library -- IT Opportunities and Challenges in the New Millennium, National Library of China, Beijing, China, pp. 454-462, July 9-11, 2002.
4. Huamin Feng, Rui Shi, Tat-Seng Chua, "A bootstrapping framework for annotating and retrieving WWW images", Proceedings of the 12th annual ACM international conference on Multimedia, October 2004.
5. Von-Wun Soo, Chen-Yu Lee, Chung-Cheng Li, Shu Lei Chen, Ching-chih Chen, "Automatic metadata creation: Automated semantic annotation and retrieval based on sharable ontology and case-based learning techniques", Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries, May 2003.
6. HuaMin Feng, Tat-Seng Chua, "Image retrieval: A bootstrapping approach to annotating large image collection", Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval, November 2003.
7. Meng Yang, Barbara M. Wildemuth, Gary Marchionini, "Technical poster session 1: multimedia analysis, processing, and retrieval: The relative effectiveness of concept-based versus content-based video retrieval", Proceedings of the 12th annual ACM international conference on Multimedia, October 2004.
8. Carmine Cesarano, Antonio d'Acerno, Antonio Picariello, "Intelligent web information access: An intelligent search agent system for semantic information retrieval on the internet", Proceedings of the 5th ACM international workshop on Web information and data management, November 2003.
9. G. Schreiber, I. Blok, D. Carlier, W. van Gent, J. Hokstam, and U. Roos. A mini-experiment in semantic annotation. In I. Horrocks and J. Hendler, editors, *The Semantic Web – ISWC2002*. First international semantic web conference, number LNCS 2342, pages 404–408. Springer-Verlag, Berlin, 2002.