

# Ontological Description of Image Content Using Regions Relationships

Zurina Muda

School of Electronics and Computer Science  
University of Southampton, United Kingdom  
{zm06r@ecs.soton.ac.uk}

## Extended Abstract

**Keywords:** Spatial relationships, image annotation and ontology.

## 1 Research Problem And Aim

Rapid growth in the volume of multimedia information creates new challenges for information retrieval and sharing, and thus anticipates the emergence of the Semantic Web [2, 3]. The principal component in most of multimedia applications is the use of visual information and new approaches are essential to improve the inferring of semantic relationships from low-level features for semantic image annotation and retrieval. Much initial research on image annotation represents images in terms of colours, texture, blobs and regions, but pays little attention to the spatial relationships between regions or objects. Annotations are most frequently assigned at the global level [17] and even when assigned locally the extraction of relational descriptors is often neglected. However, current annotation system might recognise and identify a beach and an ocean in an image but fail to represent the fact that they are next to each other. Therefore, to enrich the semantic description of the visual information, it is important to capture such relations.

The aim of this research is an attempt to develop a new approach or technique for enhancing annotation systems, either through automatic or semi-automatic means, by capturing the spatial relationships between labelled regions or objects in images and incorporating such knowledge in a knowledge base such as an ontology. By this means, human users and software agents alike will be able to search, retrieve and analyze visual information in more powerful ways.

## 2 Related Work

Ontologies play an important role for knowledge intensive applications to enable content-based access, interoperability and communication across the Web. These ontologies become the backbone for enabling the Semantic Web [20]. The number of multimedia ontologies available is still rather small, and well-designed ontologies that

fulfill the requirements [5] of reusability, MPEG-7 compliance, extensibility, modularity and interoperability are rare [18]. The COMM ontology which is under development elsewhere and is based on DOLCE ontology as a foundational ontology is of particular relevance.

A pure combination of traditional text-based and content-based approaches is not sufficient for dealing with the problem of image retrieval on the Web, mostly because of the problem of its text based orientation. Some Web images have irrelevant, few or even no surrounding texts. Thus, the problem of limited collateral text for the annotation of images needs to be solved. Besides, manual image annotation is a tedious task and often it is difficult to make accurate annotations on images. There are many annotation tools available but human input is still needed to supervise the process. So, there should be a way to minimize the human input by making the annotation process semi or fully automatic. In the latter case, although there is much research on automatic image annotation, the results often do not really satisfy the retrieval requirements because of the flexibility and variety of user needs.

To date, many content-based image retrieval research systems, frameworks and approaches have been reported. Li et. al[14] presented Integrated Region Matching, a similarity measure for region based image similarity comparison. Ko & Byun[8] used Hausdorff Distance to estimate spatial relationships between regions in their Integrated Finding Region In the Pictures (IFRIP) as extension to their previous FRIP [9]. Laaksonen et. al[10] proposed a context-adaptive analysis of image content, by using automatic image segmentation. Lee et. al[12] proposed a new domain-independent spatial similarity and annotation-based image retrieval system. Zhou et. al [21] proposed an approach for computing the orientation spatial similarity between two symbolic objects in an image. Wang [18] proposed a new spatial-relationship representation model called two dimension begin-end boundary string (2D Be-string), based on previous research in 2D String [11]. Ahmad & Grosky [2] proposed a symbolic image representation and indexing scheme to support retrieval of domain independent spatially similar images.

However, all the research in spatial relationships has been pursued independently without taking into consideration the problems of integrating them with an ontology. Such integration would be valuable in producing high level semantics by making semantic annotation systematically easier and more meaningful. In doing so, existing ontologies such as DOLCE and COMM will be evaluated to identify both their relevance and effectiveness in achieving the research aim.

### **3 Contributions And Evaluation**

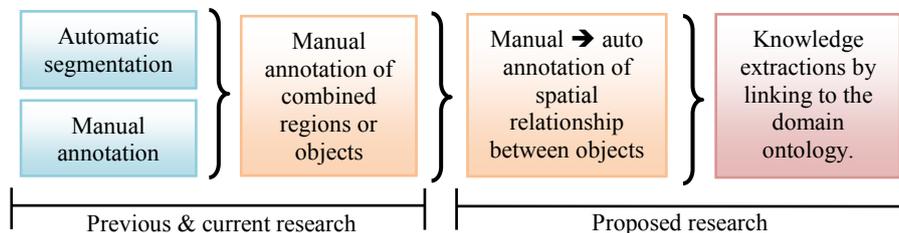
As part of a preliminary experiment, a comparative analysis of three existing annotation tools has been carried out: Caliph & Emir [15], AKTive Media [6], and M-OntoMat-Annotizer [16]. Each of these tools has been explored individually by using a group of images and a comparative study based on an evaluation framework adapted from Lewis[13] and Duineveld[7] has been performed and results obtained. The comparative study investigated image description features (including annotation) and

user interface components to find out the capabilities of existing image descriptions tools and to establish whether the spatial relationships are included and, if so, what the relationships might be. For image description components, follow-up with the developer of the tools has been established to ensure the reliability of the result.

The study shows that, each of the tools offered some special features compared to others and all tools were involved with manual annotations of the whole image. In addition M-OntoMat-Annotizer and AKTive Media allowed segmentation and annotation of the selected regions in images. Caliph & Emir and AKTive Media support some relations but not spatial relationships. Neither of these tools considered the specific locations of objects nor regions in the image for annotation or retrieval.

Based on the study and the previous research, currently, several existing annotation or description tools enable automatic segmentation by grouping multiple regions together and use manual annotation to annotate those regions. By adding the locator description where spatial relationships are considered, the knowledge of the image content becomes more specific and retrieval could be more efficient and performed in an explicit way.

This research will use existing automatic segmentation algorithm when available and manual combining of regions into composite regions for recognised objects. These will be manually annotated in the first instance together with spatial relationships between the objects. From there, an automatic annotation of spatial relationships among the objects in the image plane could be developed based on various available approaches by integrating directional and topological representation of spatial relationships. The process is simplified as illustrated in Fig.2.



**Fig.2.** Research outline.

Therefore the expected contribution will be a new approach or technique to automate spatial relationships extraction between the composite regions or objects in images and linking the knowledge to an extended multimedia ontology. The approach or technique should be reliable in order to counter the uncertainty of matching images with the real world cases. For example, this is how it would work when given an image of a beach:

1. Existing tool would provide the annotation of regions of the image corresponding to: the beach, the ocean, the sky and the coconut tree objects are recognised.
2. Our approach then identify that: a. The coconut tree is within the beach; b. The beach is next to the ocean; c. The ocean is below the sky.

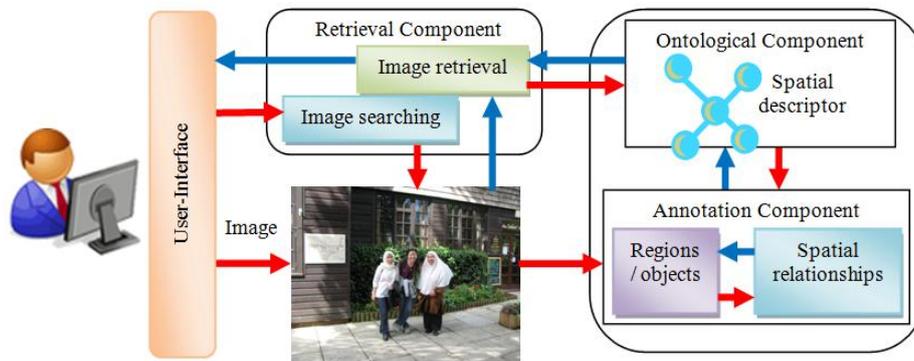
3. By reasoning over appropriate domain ontology, and exploiting the entailed spatial relationships, we would be able to infer that if the beach is in Hawaii, then the ocean must be the Pacific Ocean.

For the time being, the domain of the research would be a subset of everyday scenes such as city scenes or places of interest, but later other domains such as medical domain, may be considered to test the generality of the approach. Evaluation on ground truth with spatial relationships in term of precision and recall test will be made to see how well the automated extraction of spatial relationships has been achieved. The evaluation will use sufficient images such as Corel dataset to ensure statistical significance of the result obtained.

## 4 Work Plan

In order to accomplish the aim, the research plan is assigned into two levels – a macro plan using a Gantt chart for general activities and corresponding timelines, and micro plan using a K-chart [1] for the specific planning and execution of research. The research framework is illustrated in Fig. 3 and consists of:

1. Annotation component – automatically extracts and identifiers spatial relationships between multiple segmented regions or objects.
2. Ontological component – logics and reasoning of the extended existing multimedia ontology specifically in terms of spatial descriptors and locators.
3. Retrieval component – image retrieval mechanisms based on spatial relationships to evaluate the functionality and effectiveness of the approach.



**Fig.3.** Research framework.

So far, the literature reviews and some preliminary experiment have been performed. However further practical works in the research and development phase is now being carried out. As a conclusion, it is hoped that this research will generate a constructive semantics approach in enabling the Semantic Web as well as bridging the Semantic Gap in image retrieval, while at the same time contributing new finding to human knowledge as a whole.

## References

1. Abdullah, M. K., Mohd Suradi, N. R., Jamaluddin, N, Mokhtar, A.S., Abu Talib, A. R. & Zainuddin, M. F.: K-Chart: A Tool for Research Planning and Monitoring. *J. of Quality Management And Analysis*, vol 2(1), 123-130 (2006.)
2. Ahmad, I. & Grosky, W. I.: Indexing and Retrieval of Images by Spatial Constraints. *J. of Visual Communication and Image Representation*, vol. 14(3), Elsevier, 291-320, (2003)
3. Berners-Lee, T., Hendler, J. & Lassila, O.: The semantic web. *Scientific American*, (2001)
4. Berners-Lee, T., Fischetti, M. & Francisco, H.: Weaving the web: The original design and ultimate destiny of the World Wide Web by its Inventor (1999)
5. Bloehdorn, S. Petridis, K. Saathoff, C. Simou, N. Tzouvaras, V. Avrithis, Y. Handschuh, S. Kompatsiaris, I. Staab, S. & Strintzis, M.G.: Semantic Annotation of Images and Videos for Multimedia Analysis. In *Proc. of the 2nd ESWC2005*, (2005)
6. Chakravarthy, A., Ciravegna, F. & Lanfranchi, V.: Cross-media document annotation and enrichment. In: *Proc of the 1st SAAW2006*, (2006)
7. Duineveld, A. J., Stoter, R., Weiden, M. R., Kenepa, B. & Benjamins, V.R.: WonderTools? A comparative study of ontological engineering tools. In: *Proc. of the 12<sup>th</sup> Workshop on Knowledge Acquisition, Modeling and Management*. Alberta, Canada, October (1999)
8. Ko, B. & Byun, H.: Multiple Regions and Their Spatial Relationship-Based Image Retrieval. In: *Proc.of the international CIVR2002*. Lew, et al. (Eds). LNCS, Vol. 2383. Springer-Verlag, London, 81-90(2002)
9. Ko, B. C., Lee, H. S. & Byun, H.: Region-based Image Retrieval System Using Efficient Feature Description. In: *Proc. of the 15th ICPR2000*, vol. 4, 283-286. Spain, Sept., (2000)
10. Laaksonen, J., Koskela, M & Oja, E.: PicSOM-Self-organizing image retrieval with MPEG-7 content descriptions. *IEEE Trans. on Neural Networks*, 13(4), 841-853 (2002)
11. Lee, S.C., Hwang E.J & Lee, Y.K.: Using 3D Spatial Relationships for Image Retrieval by XML Annotation. *ICCSA2004*, LNCS 3046, 838-848 (2004)
12. Lee, S. & Hwang, E.: Spatial Similarity and Annotation-Based Image Retrieval System. *IEEE 4<sup>th</sup> Inter. Sym. on Multimedia Software Engineering*, Newport Beach, CA (2002)
13. Lewis, J.R.: IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instruction for Use. *Inter. J. of HCI*, 7(1), 57-78 (1995)
14. Li, J., Wang, J. Z., Wiederhold G.: IRM: Integrated Region Matching for Image Retrieval. *ACM Multimedia*, pp. 147-156, (2002)
15. Lux, M. Becker, J. & Krottmaier, H. Calph & Emir: Semantic Annotation and Retrieval in Personal Digital Photo Librarian. In: *Proc. of 15<sup>th</sup> CAiSE'03*. pp. 85-89, Austria (2003)
16. Saathoff, C., Petridis, K., Anastopoulos, D., Timmermann, N., Kompatsiaris, I. & Staab, S.: M-OntoMat-Annotizer: Linking Ontologies with Multimedia Low-Level Features for Automatic Image Annotation. In: *Posters of the 3rd ESWC 2006*, Montenegro, (2006)
17. Srikanth, M., Varner, J., Bowden, M. & Moldovan, D.: Exploiting Ontologies for Automatic Image Annotation. In: *Proc. of the 28<sup>th</sup> Annual Inter. ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, 552-558 (2005)
18. Staab, S.: *Multimedia Ontology*. Summer School in Multimedia Semantics (SSMS2007), Glasgow, (2007)
19. Wang, Y. H.: Image Indexing and Similarity Retrieval Based on Spatial Relationship Model. *Inf. Sci. Comput. Sci.* 154, 1-2, Elsevier, New York, USA, pp. 39-58, Aug. (2003)
20. Ying D.: *Ontology: The enabler for the Semantic Web*, <http://citeseer.ist.psu.edu/601004.html> (2002)
21. Zhou X.M., Ang C. H. & Ling T. W.: Image Retrieval based on object's orientation spatial relationship. *Pattern Recognition Letters* 22. Elsevier Science, 469-477 (2001)