Querying Image Ontology

D. N. F. Awang Iskandar *James A. Thom* *S. M. M. Tahaghoghi*
School of Computer Science and Information Technology,
RMIT University,
Melbourne, Australia
{*dayang.awgiskandar, james.thom, seyed.tahaghoghi*}@*rmit.edu.au*

Abstract *Content-based image retrieval has been used in various application domains, but the semantic gap problem remains a challenge to be overcome. One possible way to overcome this problem is to represent the knowledge extracted from the low-level image features through semantic concepts. In this paper we describe how we use an image ontology to this end. We show that we are able to retrieve desired images by using basic ontology queries.*

Keywords Ontology, CBIR, semantic gap

1 Introduction

Advances in digital imaging technology have led to a large volume of digital visual content being produced. However, this has exacerbated the problem of locating a desired image in a large and varied collection. Content-based image retrieval (CBIR) aims to retrieve images on the basis of features automatically extracted from the images themselves.

The main goal of CBIR is to find an image or a set of images that best satisfy a user's information need within an image database or collection. However, the fundamental unsolved problem in CBIR is the semantic gap [14] – the missing link between the semantic categories that a user is looking for and the low-level features that CBIR systems offer.

High-level retrieval involves retrieval of an image based on the name of objects, emotions and actions that can be associated with the image. Low-level retrieval involves retrieval of basic features such as colour, texture, shape and object location. For example, the concepts "president" and "happiness" are considered to be high-level, whereas "red circle" is a low-level concept. Users generally wish to pose high-level queries [8, 9, 14] whereas present CBIR systems can only index and retrieve images based on low-level features.

Approaches to bridge the semantic gap can be top-down, bottom-up or a combination [4] of these. In this work, we adopt a combined approach. For the top-down approach we create an ontology that contains the high level semantic concepts derived from the image regions' low-level features. For the bottom up approach, we automatically learn the semantic concepts using the technique described in earlier work [1]. The main technical contribution of this work is the ontology that we have engineered. We also show that we are able to retrieved desired images using the ontology.

Next, we describe existing research on CBIR and ontology. In Section 3, we describe our approach for querying an image ontology. In Section 4, we explain the ontology's structure and content. In Section 5, we present a query example and the retrieved images. We conclude with a discussion of our findings and suggestions for future work.

2 Background

Most effort to minimize the semantic gap has focused on automatic image annotation [2, 7, 17]. The images are annotated by using keywords or described formally using an ontology [5, 13, 17]. According to Brewster et al. [3], an ontology defines concepts properties and the relationships among the concepts.

Web ontology languages have been proposed as part of research related to the Semantic Web. XML, RDF, RDF Schema, OIL and DAML+OIL are among the earliest web ontology languages, while OWL is the current W3C recommendation.[1] A combination of RDF and OWL (RDF/OWL) can accurately describe the instances and their constraints in an ontology. RDF is used to represent information and to exchange knowledge on the Web. At a higher level, OWL is used to publish and share sets of terms called ontologies, supporting advanced Web search, software agents and knowledge management.

Ontology query languages allow expressions to be written that can be evaluated against an ontology. The queries can be used by knowledge management applications as a basis for inference actions. Existing ontology query languages include OntoQL, SPARQL, DQL (previous version of DAML+OIL), SeRQL, TRIPLE, RDQL, N3, and Versa. The SPARQL query language has been adopted by W3C as the means to query ontologies built using RDF[2] and has been extended to support OWL format. SPARQL is based on SQL and has the ca-

Proceedings of the 12th Australasian Document Computing Symposium, Melbourne, Australia, December 10, 2007.

[1] `http://www.w3.org/TR/owl-guide`
[2] `http://www.w3.org/TR/rdf-sparql-query`

pabilities for querying visual graph patterns along with their conjunctions and disjunctions.

Town [15, 16] shows that the use of ontologies to relate semantic descriptors to their parametric representations for visual image processing leads to an effective computational and representational mechanism. The ontology implements the hierarchical representation of the domain knowledge for a surveillance system. Pre-annotated surveillance video training data and its visual descriptors are incorporated in the ontology. The ontology is used to feed information to the Bayesian inference network for tracking movement. Town also proposed an ontological query language, OQUEL. The query is express using a prescriptive ontology of image content descriptors. Query sentences are grounded through a range of image analysis methods that represent the image content in low, intermediate and high semantic levels. The central role of the ontology is to provide a means for users to define the ontological domain of discourse and for the system to execute the query by grounding and assessing the particular ontological sentence with respect to actual image data.

The query approach using OQUEL is similar to the approach presented by Hyvönen et al. [6] who implement a web-based system to retrieve the images using an ontology — known as Ontagator. Image query is done using a view-based search followed by image recommendations. In the search process, users view the ontology and select the class of interest. The system will return all images related to the class. After finding a class of interest, the semantic ontology model together with image instance data are used to discover the relations between a selected image and other images in the repository. These images are then presented to the user.

Liu et al. [10] also implemented a web-based system to retrieve the images with an ontology. Search for the matching images is done by processing a text-based query. The ontology query engine is written in RDF Data Query Language (RQDL) provided by the Jena Toolkit.[3]

Mezaris et al. [11, 12] propose an approach for region-based image retrieval using an object ontology and relevance feedback. The approach utilises an unsupervised segmentation method for dividing the images into regions that are later indexed. The object ontology is used to represent the low-level features and act as a object relation identifier — for example the shape features are represented as *slightly oblong, moderately oblong, very oblong*. This ontology is not built using any ontology language, but is instead simply a vocabulary listing. The query is done using keywords in the object ontology to provide qualitative information and relationships between objects. The regions that match the query based on the object ontology are retrieved and presented to user. The user can give feedback on the retrieved images and the system will learn using Support Vector Machines (SVMs) and Constraint Similarity Measure (CSM) to filter out the unrelated images.

[3] `http://jena.sourceforge.net`

3 Our Approach

To reduce the problem of object segmentation, we test our approach on a domain where regions are easily separated: a collection of comic strips. In this domain, objects and characters comprise multiple regions of approximately uniform colour.

We have created an image collection that consists of comic strip panels from the Goats comic.[4] These include extended information that describes every panel. The description assists us in performing relevance judgements on our retrieval results. The collection consists of 452 coloured strips, each containing one to five panels. Dividing the strips into panels gives us a total of 1440 panels. We tested the retrieval effectiveness using 1115 regions extracted from 202 panels. The remaining panels are reserved as a validation set for future work. From this point onwards, we will refer to individual panels as images.

The objects in the comics have relatively consistent size and orientation, guiding our choice of the following region-based and contour-based shape features: the region area; the mean grey level value of the pixels in the region; circularity; and shape boundary. We did not use any texture features in this work since colour is a much more prominent feature in the comic image collection we are using.

We adopted the equal-weight linear combination technique from our previous work [1] to recognise and label five concepts representing the main characters in the Goats comic strips — *Bob* (an alien), *Diablo* (a chicken), *Fineas* (a fish), *Jon* (a person) and *Oliver* (a chick). This technique was compared with classification using machine learning algorithms in combining shape features, and we found that an equal-weight linear combination of shape features is simpler and at least as effective as using a machine learning algorithm [1].

4 The Image Ontology

To build the ontology, we incorporate the concepts that were recognised using the above mentioned method. The ontology is automatically augmented when new concepts are derived from the images. The image retrieval is performed using SPARQL.

We divided the ontology structure into two general classes – Concept and Graphic. The class Concept has a subclass Character that further contains character instances – *Bob*, *Diablo*, *Fineas*, *Oliver* and *Jon*. The Graphic subclasses are Strip and Panel. The subclass Panel contains instances that describes the panel sequence (first panel labelled as "a" and so forth) and the character or characters in it. Currently, we have

[4] `http://www.goats.com`

```
<?xml version="1.0"?>
 <rdf:RDF
 xmlns="http://dayang.cs.rmit.edu.au/~dayang/ComicOntology#"
 xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
 xmlns:owl="http://www.w3.org/2002/07/owl#"
 xml:base="http://dayang.cs.rmit.edu.au/~dayang/ComicOntology">
 <owl:Ontology rdf:about="Goats Comic"/>
 <owl:Class rdf:ID="Concept"><rdfs:subClassOf><owl:Class rdf:ID="Comic"/></rdfs:subClassOf></owl:Class>
 <owl:Class rdf:ID="Graphic"><rdfs:subClassOf rdf:resource="#Comic"/></owl:Class>
 <owl:Class rdf:ID="Strip"><rdfs:subClassOf rdf:resource="#Graphic"/></owl:Class>
 <owl:Class rdf:ID="Panel"><rdfs:subClassOf rdf:resource="#Graphic"/></owl:Class>
 <owl:Class rdf:ID="Character"><rdfs:subClassOf rdf:resource="#Concept"/></owl:Class>
 <owl:ObjectProperty rdf:ID="HasCharacter"/>
 <owl:ObjectProperty rdf:ID="PartOf"/>
 <owl:ObjectProperty rdf:ID="InPanel"/>
 <owl:DatatypeProperty rdf:ID="name"/>
 <owl:DatatypeProperty rdf:ID="image_Ref"/>
 <Strip rdf:ID="goats031226.png-a.jpg">
    <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">goats031226.png-a.jpg</name></Strip>
 <Panel rdf:ID="a">
    <PartOf rdf:resource="#goats031226.png-a.jpg"/>
    &lt;image_Ref rdf:datatype="http://www.w3.org/2001/XMLSchema#string"&gt;goats031226.png-a.jpg&lt;/image_Ref&gt;
      <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">a</name>
    <HasCharacter><Character rdf:ID="Oliver">
      <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Oliver</name>
    <InPanel rdf:resource="#a"/></Character>
    <Character rdf:ID="Diablo">
      <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Diablo</name>
    <InPanel rdf:resource="#a"/></Character></HasCharacter></Panel>
    ...
 </rdf:RDF>
```

Figure 1: Snippet of the image ontology in RDF/OWL format.

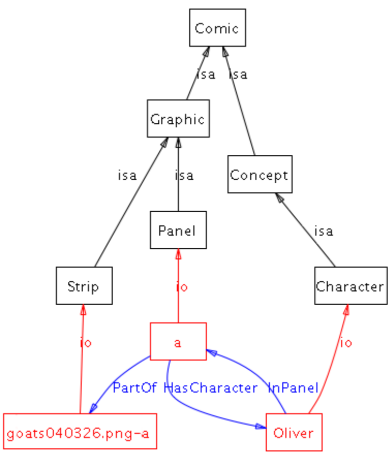

Figure 2: Visual graph of the image ontology. This figure is best viewed in colour. Class nodes are black and instance nodes are red in colour. Relationships — `isa`: is-a; `io`: instance of, `InPanel`, `HasCharacter`, `PartOf`.

successfully recognised the characters. We adopted this ontology structure so the graphic elements of the image collection are separated from the semantic concepts.

Referring to Figure 1, we can see that the character Oliver is in panel “a” of the strip in the RDF/OWL format. The visual graph generated using OntoViz[5] of this RDF/OWL format is depicted in Figure 2. The relation between the strip, panel and character instances are `PartOf` (a panel is part of a strip), `hasCharacter` (panel has character) and `InPanel` (a character is in a particular panel).

5 Queries and Answers

A query to the ontology is done using SPARQL that is available as a built-in component in Protégé.[6] A sample SPARQL query to “Find images depicting Oliver and Diablo” is shown in Figure 3.

The query consists of three parts: the SELECT clause which identifies the criteria to appear in the query results; the WHERE clause specifies the criteria for selecting results from the database; and the FILTER clause restrict the results according the expression. In this case, we want the results that have both characters in the same panel.

This query returns a list of image references that contain the characters Oliver and Diablo. A sample of retrieved images is depicted in Figure 4. This is preliminary work. We have tested the ontology with simple queries and shown that our approach is promising. We plan to extend the ontology to support more complex queries.

6 Conclusion and Future Work

In this paper, we have presented an overview of our approach towards bridging the semantic gap in CBIR. We have built an ontology by incorporating the knowledge

[5] `http://protege.cim3.net/cgi-bin/wiki.pl?OntoViz`

[6] `http://protege.stanford.edu`

```
PREFIX comic:
<http://dayang.cs.rmit.edu.au/~dayang/ComicOntology#>
SELECT ?Panel
FROM
<http://dayang.cs.rmit.edu.au/~dayang/ComicOntology.owl>
WHERE {
?x comic:name ?CharacterName1;
     comic:InPanel  ?Panel.
?y comic:name ?CharacterName2;
     comic:InPanel  ?Panel.

FILTER regex( ?CharacterName1, "^Diablo", "i")
FILTER regex( ?CharacterName2, "^Oliver", "i")
FILTER (?CharacterName1 && ?CharacterName2)
}
```

Figure 3: Example of a SPARQL query to retrieve image panels that contain the character Oliver and Diablo.

learned from our region-based image retrieval method. We have shown that using our image ontology, we are able to pose a query that retrieves desired images containing a particular comic character.

This is a preliminary work; we plan to extend this work to support positioning queries such as “find images of Diablo, where Oliver is stage left”, and to adapt it to accommodate real photographic images. Another interesting direction for this work is to implement a visual region-based ontology querying interface.

Acknowledgements This research was undertaken with a scholarship provided by the Malaysian Ministry of Higher Education. We thank Jonathan Rosenberg, author of the Goats Comic for permission to publish the comic images.

References

[1] D. N. F. Awang Iskandar, J. A. Thom and S. M. M. Tahaghoghi. Content-based image retrieval using image regions as query examples. In *Nineteenth Australasian Database Conference (ADC 2008)*, Volume 75 of *CRPIT*, Wollongong, Australia, 2008. ACS. (To appear).

[2] K. Barnard and D.A. Forsyth. Learning the semantics of words and pictures. In *Proceedings of the 8th International Conference on Computer Vision*, pages 408–415, 2001.

[3] C. Brewster and K. O’Hara. Knowledge representation with ontologies: The present and future. *IEEE Intelligent Systems*, Volume 19, Number 1, pages 72–73, January/February 2004.

[4] J. S. Hare, P. H. Lewis, P. G. B. Enser and C. J. Sandom. Mind the gap: Another look at the problem of the semantic gap in image retrieval. In *Proceedings of Multimedia Content Analysis, Management and Retrieval 2006 SPIE*, Volume 6073, pages 607309–1, 2006.

[5] E. Hyvönen, S. Saarela, A. Styrman and K. Viljanen. Ontology-based image retrieval. In *Proceedings of World Wide Web 2003*, May 2003. Poster papers.

[6] E. Hyvönen, S. Saarela and K. Viljanen. Ontogator: Combining view- and ontology-based search with semantic browsing. In *Proceedings of XML Finland 2003*, October 2003.

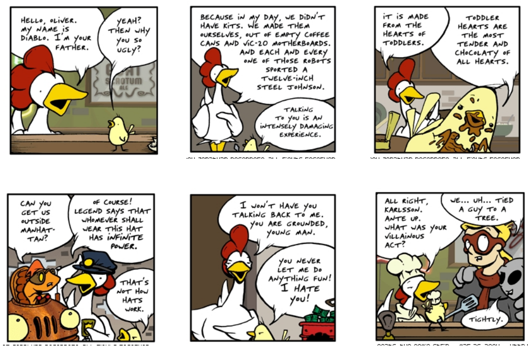

Figure 4: Retrieved images of Oliver and Diablo in the same comic panel. This figure is best viewed in colour.

[7] C. Jelmini and S. Marchand-Maillet. Deva: an extensible ontology-based annotation model for visual document collections. In *SPIE*, Volume 5018, January 2003.

[8] M. L. Kherfi, D. Ziou and A. Benardi. Image retrieval from the world wide web: Issues techniques, and systems. *ACM Computing Surveys*, Volume 36, Number 1, pages 35–67, March 2004.

[9] T. Kurita and T. Kato. Learning of personal visual impression for image database systems. In *Proceedings of the Second International Conference on Document Analysis and Recognition*, pages 547–552, 1993.

[10] S. Liu, L. Chia and S. Chan. Ontology for nature-scene image retrieval. In *CoopIS/DOA/ODBASE (2)*, Volume 3291 of *Lecture Notes in Computer Science*, pages 1050–1061. Springer, 2004.

[11] V. Mezaris, I. Kompatsiaris and M. G. Strintzis. An ontology approach to object-based image retrieval. In *IEEE International Conference on Image Processing (ICIP03)*, pages 511–514, 2003.

[12] V. Mezaris, I. Kompatsiaris and M. G. Strintzis. Region-based image retrieval using an object ontology and relevance feedback. *EURASIP Journal on Applied Signal Processing*, Number 6, pages 886–901, 2004.

[13] A. Th. Schreiber, B. Dubbeldam, J. Wielemaker and B. Wielinga. Ontology-based photo annotation. *IEEE Intelligent Systems*, Volume 16, Number 3, pages 66–74, 2001.

[14] A.W.M Smeulders, M. Worring, S. Santini, A. Gupta and R Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 22, Number 12, pages 1349–1380, Dec 2000.

[15] C. Town. *Ontology based Visual Information Processing*. Ph.D. thesis, University of Cambridge, 2004.

[16] C. Town. Ontological inference for image and video analysis. *Machine Vision and Applications*, Volume 17, Number 2, pages 94 – 115, May 2006.

[17] R. Troncy, J. van Ossenbruggen, J. Z. Pan and G. Stamou. Image annotation on the semantic web, W3C incubator group report, 2007. `http://www.w3.org/2005/Incubator/mmsem/XGR-image-annotation/`.