

# An Integrative Semantic Framework for Image Annotation and Retrieval

Taha Osman<sup>1</sup>, Dhavalkumar Thakker<sup>1</sup>, Gerald Schaefer<sup>2</sup>, Phil Lakin<sup>3</sup>

<sup>1</sup> *School of Computing & Informatics, Nottingham Trent University, Nottingham, NG11 8NS, UK  
{taha.osman, dhavalkumar.thakker}@ntu.ac.uk*

<sup>2</sup> *School of Engineering & Applied Science, Aston University, Aston Triangle, Birmingham B4 7ET, UK  
g.schaefer@aston.ac.uk*

<sup>3</sup> *PA Photos, Pavilion House, 16 Castle Boulevard, Nottingham, NG7 1FL, UK  
phil.lakin@paphotos.com*

## Abstract

*Most public image retrieval engines utilise free-text search mechanisms, which often return inaccurate matches as they in principle rely on statistical analysis of query keyword recurrence in the image annotation or surrounding text. In this paper we present a semantically-enabled image annotation and retrieval engine that relies on methodically structured ontologies for image annotation, thus allowing for more intelligent reasoning about the image content and subsequently obtaining a more accurate set of results and a richer set of alternatives matchmaking the original query. Our semantic retrieval technology is designed to satisfy the requirements of the commercial image collections market in terms of both accuracy and efficiency of the retrieval process. We also present our efforts in further improving the recall of our retrieval technology by deploying an efficient query expansion technique.*

## 1. Introduction

Affordable access to digital technology and advances in Internet communications have contributed to the unprecedented growth of digital media repositories (audio, images, and video). Retrieving relevant media from these ever-increasing repositories is an impossible task for the user without the aid of search tools. Most public image retrieval engines rely on analysing the text accompanying the image to matchmake it with the user query. Various optimisations were developed including the use of weighting systems where for instance higher regard can be given to the proximity of the keyword to the image location, or advanced text analysis techniques that use term weighting method, which relies on the proximity between the anchor to an image and each word in an HTML file [1]. Despite the optimisation efforts, these search techniques remain hampered by

the fact that they rely on free-text search that, while cost-effective to perform, can return irrelevant results as it primarily relies on the recurrence of exact words in the text accompanying the image. The inaccuracy of the results increases with the complexity of the query. For instance, while performing this research we used the Yahoo™ search engine to look for images of the football player Zico returns some good pictures of the player, mixed with photos of cute dogs (as apparently Zico is also a popular name for pet dogs), but if we add the action of scoring to the search text, this seems to completely confuse the Yahoo search engine and only one picture of Zico is returned, in which he is standing still!

Any significant contribution to the accuracy of matchmaking results can be achieved only if the search engine can “comprehend” the meaning of the data that describes the stored images, for instance, if the search engine can understand that scoring is an act associated with sport activities performed by humans. Semantic annotation techniques have gained wide popularity in associating plain data with “structured” concepts that software programs can reason about [2]. This effort presents a comprehensive semantic-based solution to image annotation and retrieval as well as deploying query expansion techniques for improving the recall rate. It specifically targets the commercial image collections market and acknowledges their requirements for high quality recall without sacrificing the performance of the retrieval process.

The paper begins with an overview of the Semantic web technologies. In section 3 we review the case study that was the motivation for this work. Sections 4, 5, 6, and 7 detail the implementation roadmap of our semantic-based retrieval system, i.e. ontology engineering, annotation, retrieval, and query expansion. We present our conclusions and plans for further work in section 8.

## 2. Overview of the semantic web

### 2.1. Ontologies (domain conceptualisation)

The fundamental premise of the semantic web is to extend the Web's current human-oriented interface to a format that is comprehensible to software programmes. Naturally this requires a standardised and rich knowledge representation scheme or *Ontology*.

One of the most comprehensive definitions of ontologies is that expressed in [3]: "Ontology is a *shared conceptualisation of a domain* and typically consists of comprehensive set of concept classes, relationships between them, and instance information showing how the classes are populated in the application domain. This comprehensive representation of knowledge from a particular domain allows reasoning software to make sense of domain-related entities (images, documents, services, etc.) and aid in the process of their retrieval and use.

### 2.2. Caption-based semantic annotation

Applied to image retrieval, the semantic annotation of images allows retrieval engines to make more intelligent decisions about the relevance of the image to a particular user query, especially for complex queries. For instance to retrieve images of the football star David Beckham expressing anger, it is natural to type the keywords 'David Beckham angry' into the Google™ Image Search engine. However, at the time of the experiment, the search engine returned 14 images of David Beckham and he looks upset in only two of them. The other retrieved images were completely irrelevant with one of them displaying an angry moose!

The use of Semantic technologies can significantly improve the computer's understanding of the image objects and their interactions by providing a machine-understandable conceptualisation of the various domains that the image represents. This conceptualisation integrates concepts and inter-entity relations from different domains, such as Sport, People and Emotions relation to the query above [4], thus allowing the search engine to infer that David Beckham is a person and thus likely to express emotions and that he is also an English footballer playing for Real Madrid FC.

### 2.3. Content-based semantic annotation

The success of caption-based semantic image retrieval largely depends on the quality of the semantic caption (annotation) itself. However, the caption is not always available largely because the annotation is a labour intensive process. In such situations, image

recognition techniques are applied, which is better known as content-based retrieval. However, the best content-based techniques deliver only partial success as image recognition is an extremely complex problem [5], especially in the absence of accompanying text that can aid inferring in the relationship between the recognized objects in the image. Moreover, from a query composition point of view, it is much easier to use a textual interface rather than a visual interface (by providing sample training image or sketch) [6].

## 3. Case study for semantic image retrieval

An opportunity to experiment with our research findings in semantic-based search technology was gratefully provided by PA Photos™. PA Photos is a Nottingham-based company which is part of the Press Association Photo Group Company [7]. As well as owning a huge image database in excess of 4 million annotated images which date back to the early 1900's, the company processes a colossal amount of images each day from varying events ranging from sport to politics and entertainment. The company also receives annotated images from a number of partners that rely on a different photo indexing schema.

More significantly, initial investigation has proven that the accuracy of the results sets matching the user queries do not measure up to the rich repository of photos in the company's library.

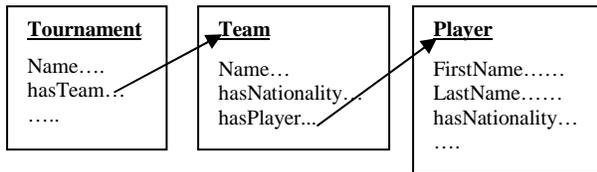
The goal of the case study is two-fold. Initially, we intend to investigate the use of semantic technology to build a classification and indexing system that critically unifies the annotation infrastructure for all the sources of incoming stream of photos. Subsequently, we'll conduct a feasibility study aiming to improve the end-user experience of their images search engine. At the moment PA Photos search engine relies on Free-Text search to return a set of images matching the user requests. Therefore the returned results naturally can go off-tangent if the search keywords do not exactly recur in the photo annotations. A significant improvement can result from semantically enabling the photo search engine. Semantic-based image search will ultimately enable the search engine software to understand the "concept" or "meaning" of the user request and hence return more accurate results (images) and a richer set of alternatives.

It is important here to comment about the dynamics of the retrieval process for this case study as it represents an important and wide-spread class of application areas where there is a commercial opportunity for exploiting semantic technologies:

1. The images in the repository have not been extracted from the web. Consequently the extensive research into using the surrounding



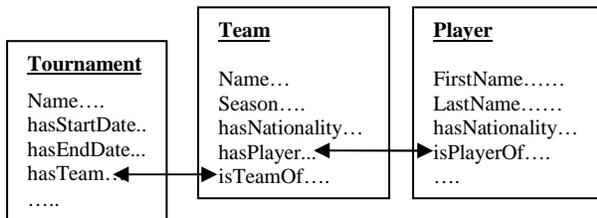
However logical the above description appears at first sight, further analysis reveals inconsistency problems. When a player plays for two different teams at the same time (e.g. his club and his national team) or changes clubs every year, it is almost impossible to determine which team the player plays for. Hence, the order of definition (relationship direction) should always be the reversal sequence of the *part-of* relationship as redesigned below:



**Figure 3 Re-organization of the player classification**

#### 4.3. Coverage

Although consistent, the structural solution in Figure 3 is incomplete as players' membership is temporal. The same problem occurs with tournaments as from one year to another, teams taking part in the tournament change. This problem can be solved by adding a start and end date for the tournament (see Figure 4), rather than by engineering more complex object property solutions. Hence, as far as the semantic reasoner is concerned, the "FIFA World Cup 2004" is a different instance from "FIFA World Cup 2008". The same reasoning can be applied to the class team, as players can change team every season. These considerations, although basic for a human reasoning, need to be explicitly defined in the ontology.



**Figure 4 Resolving Coverage problems in ontology**

#### 4.4. Normalisation: reducing the redundancy

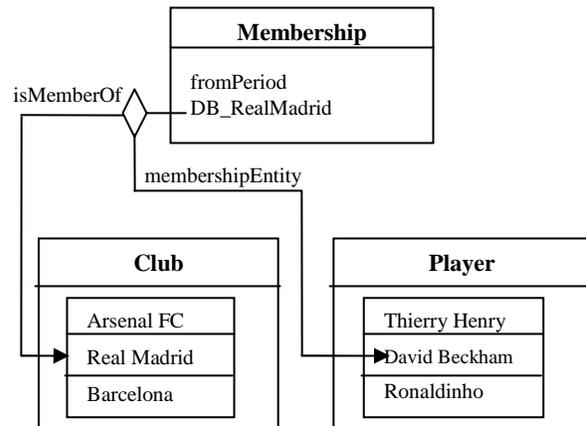
The objective of normalisation is to reduce redundancy. In ontology design, redundancy is often caused by temporal characteristic that can generate redundant information and negatively affect the performance of the reasoning process.

Direct adoption of the ontology description in Figure 4 above will result in creating new team each season, which is rather inefficient as the team should be a non-temporal class regardless of the varying player's membership or tournament participation every

season. Hence, Arsenal or Glasgow Rangers Football clubs need to remain abstract entities. Our approach was to introduce an intermediary temporal *membership* concept that serves as an indispensable link between teams and players, as well as between teams and tournaments as illustrated in Figure 5 below.

The temporal instances from the Membership class link instances from two perpetual classes as follows:

- memberEntity links to a person (Player, Manager, Supporter, Photographer, etc.)
- isMemberOf refers to the organisation (Club, Press Association, Company, etc.)
- fromPeriod and toPeriod depict membership temporal properties

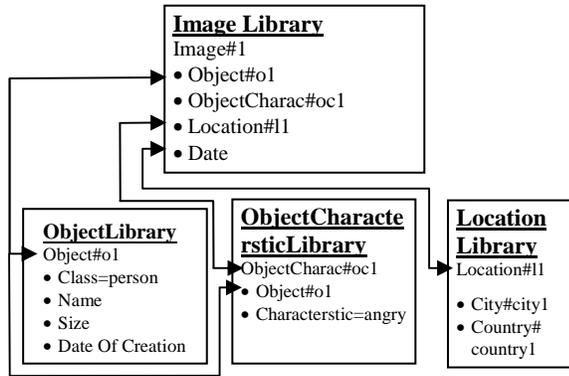


**Figure 5 Membership class in the final ontology**

## 5. Image Annotation

The Protégé® ontology editor that was utilised to construct the sport domain ontology. Protégé uses frame-based knowledge representation [11] and adopts OWL as the ontology language. The Web Ontology Language (OWL) [12] has become the de-facto standard for expressing ontologies, it adds extensive vocabulary to describe properties and classes and express relations between them (such as disjointness), cardinality (for example, "exactly one"), equality, richer typing of properties, and characteristics of properties (such as symmetry). The Jena [13] java API was used to build the annotation portal to the constructed ontology.

The central component of the annotation are the images stored (as OWL descriptions) in image library as illustrated in Figure 6. Each image comprises an object, whose main features are stored within an independent object library. Similarly are the object characteristics, event location, etc. distinct from the image library. This highly modular annotation model facilitates the reuse of semantic information and reducing redundancy.



**Figure 6 Architecture of the annotation**

Taking into account the dynamic motion nature of the sport domain, our research concluded that a variation of the sentence structure suggested in [14] is best suited to design our annotation template. We opted for an “Actor – Action – Object” structure that will allow the natural annotation of motion or emotion-type relationships without the need to involve NLP techniques [15]. For instance, “Beckham – Smiles – null”, or “Gerrard – Tackles – Henry”. An added benefit of the structure is that it simplifies the task of the reasoner in matching *actor* and *action* annotations with entities that have similar characteristics.

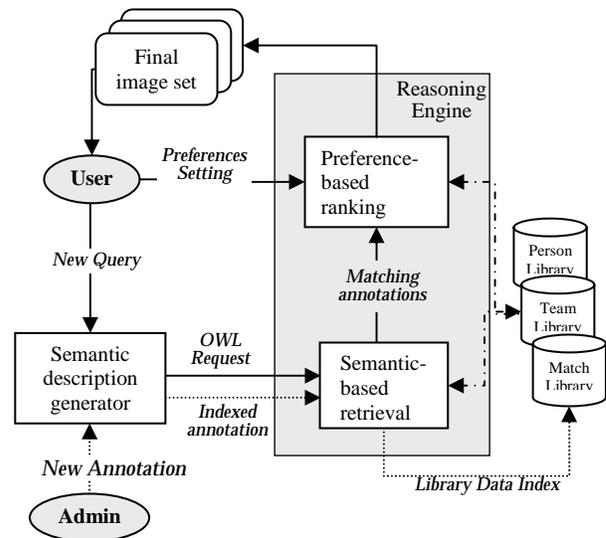
## 6. Image Retrieval

The image retrieval user interface is illustrated in Figure 7. The search query can include sentence-based relational terms (Actor-Emotion/Action-Object) and/or key domain terms (such as tournament and team). In case multiple terms were selected for the query, the user needs to specify which term represents the *main* search preference (criterion).

**Figure 7 Snapshot of the retrieval interface**

For instance, in Figure 7 the relational term (Gerrard Tackles Rooney) is the *primary* search term and team Liverpool is the *secondary* search term. The preference setting is used to improve the ranking of retrieved images.

Figure 8 gives a high level view of the annotation and retrieval mechanism. The semantic description generator allows the annotator to transparently annotate new images and also transforms the user query into OWL format. The semantic reasoning engine applies our matchmaking algorithm at two phases: The first phase retrieves images with annotations matching all concepts in the query, while in the second phase further matchmaking is performed to improve the ranking of the retrieved images in response to user preferences.



**Figure 8 Schematic diagram of the Semantic Web Image Retrieval software**

Our reasoning engine uses a variation of the nearest neighbour matchmaking algorithm [16] to serve both the semantic retrieval and the ranking phases. Our algorithm continues traversing back to the upper class of the ontology and matching instances until there are no super classes in the class hierarchy, i.e. the leaf node for the tree is reached, giving degree of match equal to 0. The degree of match (*DoM*) is calculated according to the following equation:

$$DoM = \frac{MN}{GN} \quad \text{Equation 1}$$

Where the *MN* is the total number of matching nodes in the selected traversal path, and *GN* the total number of nodes in the selected traversal path. This is exemplified in Figure 8. Then the comparison values

are weighted using the user preferences according to the formula [16]:

$$m = |l_r - l_a|; \quad \forall p \in [0,1], \quad v = pm \quad \text{Equation 2}$$

$v$ : value assigned to the comparison;  
 $m$ : matching level of the individuals;  
 $p$ : user preference setting;  
 $l_r$ : level of the request;  
 $l_a$ : level of the annotation.

For example, if the query is *Object-hasCharacteristic-happy*, and *image1* and *image2* are annotated with *Object-hasCharacteristic-happy* and *Object-hasCharacteristic-smile* respectively, the DoM for *image1* is 1 as the instances match to the level of the leaf node (Figure 9). However, for *image2* instances match to the level of *Positive Feeling- Mild* class and is one layer lower than the leaf node giving DoM = 0.5.

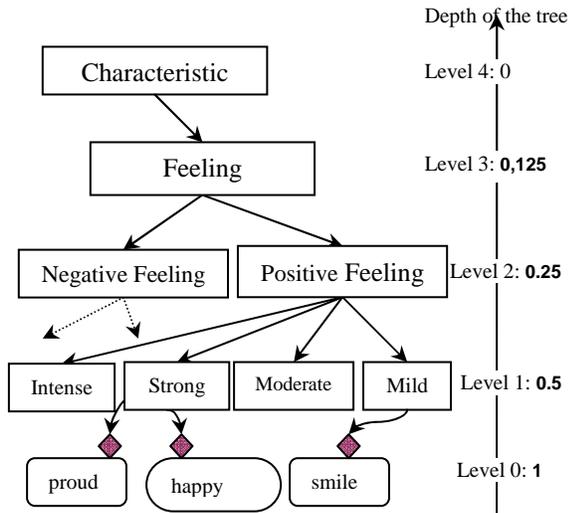


Figure 9 Traversing the Ontology Tree

## 7. Semantic Web based Query Expansion to achieve better precision and recall

Lately query expansion (QE) techniques have gained a lot of attention in attempting to improve the recall of document and media queries. QE methods fit naturally into our image retrieval technology as we rely on computing the aggregate degree of match (ADoM) for the semantic relations describing a particular image to determine its match to the original query. Hence, we can easily determine the quality of the returned results in terms of accuracy and volume and decide whether to apply QE techniques to replace or improve the query concepts to improve the quality of the recall. This is particularly feasible for semantic-based knowledge bases as they provide language expressiveness for

specifying the similarity of the concepts (Implicit and Explicit) at different granularity.

Query expansion techniques can be broadly classified into two categories: the first category uses statistical and probabilistic methods [17] to extract frequently occurring terms from successfully recalled documents and image annotations. These terms are then used to expand the keyword set of similar future queries. The Main shortcoming of the statistics-based QE techniques is that they are as good as the statistics they rely on and have similar disadvantages as free-text based search engines in that they lack structure and are difficult to generalize or to reuse for other domains. The second category [18] utilises lexical databases to expand user queries. A lexical database similar to WordNet [19] is employed, in which language nouns, verbs, adjectives and adverbs are organized into synonym sets that can potentially replace or expand the original query concepts. However, lexical database lack the semantic conceptualisation necessary to interrelate concepts in complex queries and render them comprehensible to search engines.

Semantic relations-based QE technique expands the query with related concepts rather than simple terms. Next we discuss the semantic-based QE algorithm we designed to expand our image retrieval technology.

*Step1: If query has concept  $C^p$  as the primary search concept and  $C^s$  as the secondary search concept provided by the searcher then we apply query expansion on  $C^p$  as follows:*

$C^p$  is the original concept,  $C^{p'}$  is the alternative concept,  $\delta$  is the distance between two concepts and  $\Psi$  is the expected distance between two concepts implying them related. Hence the expansion function is:

$$\sum_{i=1}^n (C^p \xrightarrow{\delta_i, \Psi_i} C_i^{p'}) , \quad \delta_i \geq \Psi_i \quad \text{Equation 3}$$

### 7.1. Formalizing relatedness between two concepts

A major concern in QE techniques is the formalization of relatedness between two concepts in order to select an optimal set of alternatives.

For the benefit of the discussion, we feel it is necessary to revisit the following components of Semantic web formalism and their representation in the OWL ontology language:

*Taxonomy Relationships (TR):* Taxonomy is the concepts classification system facilitated by Semantic Web. *Class* and *Individual* are the two main elements of this structure where a class is simply a name and

collection of properties that describe a set of individuals. Examples of relationships between concepts at the taxonomy level are *class*, *subclass*, *superclass*, *equivalent class*, *individual*, *sameAs*, *oneOf*, *disjointWith*, *differentFrom*, *AllDifferent*.

**Rules based relationships (RR):** Semantic Web Rule Language (SWRL) defines rule based semantics using subset of OWL with the sublanguages of Rule Mark-up Language. SWRL extends OWL with horn-like First Order Logic rules to extend the language expressivity of OWL.

We use this relationship formalism to identify explicit and implicit relatedness of concepts. To evaluate implicit relationships we use *subsumption* and *classification* to perform semantic tree traversal and compare the concepts with respect to the semantic network tree as detailed in our image retrieval algorithm earlier. Contrarily, explicit relationship between two concepts always has a Degree of Match (DoM) of 0 or 1 as they explicitly equate or distinct two individuals. For example the *owl:sameAs* equates two individuals to unify two distinct ontology elements while *owl:differentFrom* has exact opposite effect where it makes individuals mutually distinct.

If the taxonomy and rule based implicit and explicit relationship results in  $n$  number of equivalent concepts represented by  $\{C_1, C_2, C_3, \dots, C_n\}$  or  $C^p$ , then to calculate *DoM* for these likely replacement concepts we employ another semantic web relationship formalism, which we will refer as property based relationship.

**Property Relationships (PR):** Properties can be used to state relationships between individuals or from individuals to data values. These relationships are achieved through the data or object type properties. (i.e., *hasTeam*, *hasTournament*, *isMemberOf*)

*Step2: Assuming Query preference concept  $C^p$  has properties  $R_i$  which has value instances  $I_i^R$  and the annotation matching the alternative concept  $C^p$  has properties  $R'_i$  and the value instances  $I_i^{R'}$ , then we can compare  $I_i^R$  and  $I_i^{R'}$  semantically using Equation 2*

## 7.2. Illustrative example

In this section we illustrate how our QE algorithm works by discussing the following case. If a user is searching for pictures with *England Team* possibly in the *2006 FIFA World Cup tournament*, the system treats *England Team* as user's primary search criterion and *2006 FIFA World Cup Tournament* as secondary search criterion in the query.

Without expanding the query, the retrieval algorithm returns zero results if there are no images annotated with *Team England* (Table 1). The following

section explains the process of expanding query under these circumstances using our algorithm.

England Team ( $C_p$ )	
( $C_p$ has properties $R_i$ )	$I_i^R$ (properties value)
Has Nationality	Country (England)
Has Sport	Sport (Football)
IsWinnerOf	Tournament(Fifawc66)
hasNationalTeamTournament	Fifawc66, 70, ...

**Table 1 Preference Concept**

In our sports domain ontology implicit subsumption relationship is applied to find relevant primary concepts. For instance, to find alternative terms for *Team England*, the reasoner first retrieves siblings of the *National Team* such as *Team Brazil*, *Team Spain*, and then less adjacent siblings of the *Team* instances such as *Team Chelsea* and *Team Barcelona*.

In the following step we compare the relationship as defined in step 2 as illustrated in the Table 2 below:

	Query	Team Brazil		Team Chelsea	
hasNationality	England	Brazil	0	England	1
hasSport	Football	Football	1	Football	1
isWinnerOf	Fifawc 06	Fifawc70	0.5	Prem. 06	0
hasNationalTeamTourna..	Fifawc 66, 70, ...	Fifawc 66,70, ...	1	Prem. 93, 94, ...	0
<b>DoM</b>		Brazil	2.5	Chelsea	2

**Table 2 Comparing relationship**

*Step3: If the ranked images in stage 2 are  $\{X_1, X_2, X_3, \dots\}$ ,  $C^s$  is the secondary search term in the query provided by the searcher, these ranked images have  $C^s$  present in their annotation  $C_x^s$  or  $\{C_{x1}^s, C_{x2}^s, \dots\}$ , then repeat step 2 where  $C^p = C^s$  and  $C^p = C_x^s$*

In our image database this results in images retrieved for the first stage associated with the relevant concepts and they are: Image 1 (Image with Team Brazil in 2006 FIFA world cup), Image 2(Chelsea – Premiership 2007).

	Query	Image 1		Image 2	
hasTournament	Fifawc 06	Fifawc 06	1	Prem. 07	0
<b>DoM</b>		Team Brazil	2.5	Chelsea	2

**Table 3 analyzing secondary terms in the query**

## 7. Conclusions

In this paper we presented a comprehensive solution for image retrieval applications that takes full advantage of advances in semantic web technologies to coherently implement the annotation, retrieval and query expansion components of the integrative

framework. We claim that our solution is particularly attractive to commercial image providers where emphasis is on the efficiency of the retrieval process as much as on improving the accuracy and volume of returned results. For instance, we shied from employing expensive content-based recognition techniques at the retrieval stage and deployed public ontology caching to reduce the reasoning overhead, while designed an efficient query expansion algorithm to improve the quality of the image recall.

The first stage of the development was producing ontologies that conceptualise the objects and their relations in the selected domain. We methodically verified the consistency of our ontology, optimised its coverage, and performed normalisation methods to rid of concept redundancies. Our annotation approach was based on a variation of the “sentence” structure to obtain the semantic-relational capacity for conceptualising the dynamic motion nature of the targeted sport domain.

The retrieval algorithm is based on a variation of the nearest-neighbour search technique for traversing the ontology tree and can accommodate complex, relationship-driven user queries. The algorithm also provides for user-defined weightings to improve the ranking of the returned images and was extended to embrace query expansion technology in a bid to improve the quality of the recall.

Although we recognize that image analysis techniques might have a large time overhead for the on-line retrieval process, we intend to research utilizing advances for in semantically-enabled content recognition technology to aid in semi-automating the annotation process of legacy caption-poor images.

## 8. References

- [1] A. Fujii, T. Ishikawa, “Toward the Automatic Compilation of Multimedia Encyclopaedias: Association Images with Term Descriptions on the Web”, In Proceedings of the 2005 International Conference on Web Intelligence – WIC05, Compiègne, France, September 19-22, 2005, pp. 536-542.
- [2] H. Wang, S. Liu and L-T. Chia, “Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches”, Proceedings of the 14th annual ACM international conference on Multimedia, Hawaii, USA, 2006, pp. 109 – 112
- [3] J.S. Hare, P.G.B. Enser and C.J. Sandom, “Mind the gap: another look at the problem of the semantic gap in image retrieval” *Multimedia Content Analysis, Management, and Retrieval*, Vol. 6073, No. 1, 2006 pp. 607309-1.
- [4] T. Berners-Lee, “Weaving the Web: the original design of the World Wide Web by its inventor” eds. T. Berners-Lee with M. Fischetti. Harper Collins, 2000. pp 157-160.
- [5] T. Lam and R. Singh, “Semantically Relevant Image Retrieval by Combining Image and Linguistic Analysis”, Proc. International Symposium on Visual Computing (ISVC), Lecture Notes in Computer Science Vol. 4292, pp. 1686 - 1695, Springer Verlag, 2006
- [6] E. W. Maina M. Ohta, K. Katayama, I. Hiroshi, “Semantic Image Retrieval Based On Ontology and Relevance Model: A Preliminary Study”, Digital Engineering Workshop, Tokyo, Japan, 24-25 February, 2005, pp. 331-339.
- [7] PA Photos. 2007. <http://www.paphotos.com/>
- [8] M. Roach, J. Mason, N. Evens, L. Xu, F. Stentiford, “Recent Trends in Video Analysis: A Taxonomy Of Video Classification Problems”, *Internet and Multimedia Systems and Applications*, Kauai, Hawaii, USA, 2002, pp.348-353.
- [9] <http://www.aktors.org/ontology/portal#>
- [10] A. Rector, “Modularisation of domain ontologies implemented in description logics and related formalisms including OWL” Proceedings of 2<sup>nd</sup> international conference on Knowledge capture, pp.121 - 128
- [11] N.F.Noy, M. Crubezy, R.W. Ferguson., H.Knublauch , and .M.A. Musen. “Protege-2000: An Open-source Ontology-development and Knowledge-acquisition Environment”, *AMIA Annual Symposium Proc.*, 953.
- [12] OWL Web Ontology Language Overview. <http://www.w3.org/TR/owl-features>
- [13] J.J. Carroll, D. Reynolds, I. Dickinson, A. Seaborne, C. Dollin and K. Wilkinson, “Jena: implementing the semantic web recommendations”, Proceedings of the 13th international World Wide Web conference, New York, USA, ACM Press, pp. 74-83.
- [14] L. Hollink, A.Th. Schreiber, J. Wielemaker, and B. Wielinga, “Semantic annotation of image collections. In Workshop on Knowledge Markup and Semantic Annotation”, *KCAP'03*, Florida, USA, 2003.
- [15] H. Chen , “Machine Learning for information retrieval: Neural networks, symbolic learning and genetic algorithms”, *Journal of the American Society for Information Science and Technology*, 46(3), April 1995, pp. 194-216.
- [16] T. Osman, D. Thakker, D. Al-Dabass D, “Semantic-Driven Matchmaking of Web Services Using Case-Based Reasoning” In proceedings of IEEE International Conference on Web Services (ICWS'06), Chicago, USA, September 2006. pp. 29-36.
- [17] J. Xu, and W. Croft, “Improving the effectiveness of information retrieval with local context analysis” *Transactions on Information Systems (ACM TOIS)*, 18(1), 2000, pp.79–112.
- [18] E. Voorhees, “Query expansion using lexical-semantic relations” In the Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, NY, USA, Springer-Verlag, 1994, pp. 61–69.
- [19] C. Fellbau, *WordNet- An Electronic Lexical Database*, The MIT press, Cambridge, MA, USA, May 1998.