Semantic Image Annotation and Retrieval with IKen

Dominic Mainz
Heinrich-Heine-University Düsseldorf
Institute of Computer Science
Universitätsstr. 1, D-40225 Düsseldorf
0049 211 8113716
mainzd@uni-duesseldorf.de

Katrin Weller
Heinrich-Heine-University Düsseldorf
Dept. of Information Science
Universitätsstr. 1, D-40225 Düsseldorf
0049 211 8110803
weller@uni-duesseldorf.de

Jürgen Mainz
VARION GmbH
Agency for Communication Design
Dießemer Bruch 112i, D-47805
Krefeld
0049 2151 646840
jmainz@varion.de

ABSTRACT
In this paper we introduce IKen, a platform for image retrieval enhanced by semantic annotations. The paper discusses work in progress and reports the current state of IKen. This comprises the functionalities for annotating photos with an underlying ontology, search features based on these annotations and the development of the domain ontology used for annotation.

Categories and Subject Descriptors
H.3.3 Information Search and Retrieval; H.3.5 Online Information Services

General Terms
Algorithms, Management, Human Factors.

Keywords
Semantic indexing, semantic search, information retrieval, ontology, image retrieval, knowledge representation, semantic application

1. INTRODUCTION
In this paper we present our approach on how ontologies can support annotation and search in a photo collection. Recently, folksonomies have established themselves as popular means for indexing large document collections, with Flickr being the most popular example for social tagging in photo collections. Yet, user-generated annotations cause some problems regarding inconsistent vocabularies, varieties of synonyms, spelling variants, misspellings, language variants etc. [1]. In the long run, the user-centric approach of social tagging will have to be combined with richer semantics of ontologies.

Aims of IKen are: a) to provide a usable application which makes benefits of ontologies visible for internet users; b) to enable easily usable functionalities for ontology-based photo annotations; c) to allow browsing a document collection based on domain semantics.

The IKen system was designed to enable users, who are not experts in knowledge engineering and ontology modelling, to apply semantic annotations and make use of semantic information retrieval in image collections. As a first use case it focuses on applications in the tourism sector. Public relations and advertising agencies in this industry as well as travel organizations should be enabled to search for available photos. This mainly effects the choice of the images included in the database and the ontology in use; the underlying principles can be applied to other and less-specified domains as well.

Related work has been done in the field of semantic image retrieval, e.g. [2], [3], [4]. Broadly related are studies in content-based image retrieval (CBIR) [5], covering e.g. automatic pattern recognition and object recognition, which may be combined with semantic image retrieval [6], [7].

2. SEMANTIC CONTENT ANNOTATION
Semantic content annotation is the basis for semantic content retrieval. Annotating (or indexing) is the process of adding content-descriptive keywords to documents – in this case to photos. If no underlying vocabulary is provided (as in social tagging systems), the semantics of keywords, their meanings and connections, are not represented. An ontology can capture these semantics, for example words that have the same meaning (synonyms), concepts that are parts of others (meronymy, part-of-relation), subconcepts (hyponymy, is-a-relation) or various self-defined relations (e.g. is_located_in, has_colour).

In IKen the concept interrelations defined in the underlying ontology can be exploited through the annotation user interface. Seperated fields have been created for the annotation of the location where the image has been taken and for emotional impressions. The main field captures the content-descriptive annotations. In a first step, concepts can be added to a photo by entering a word – the system will look for a matching ontology concept and add this to the photo. This matching process uses the rdfs:label assertions in the ontology which also include synonyms and spelling variants. If no matching concept is found, the annotation term is added as a simple tag (in a future version these unrelated tags can be used as suggestions for new ontology concepts). In a next step, details for this concept can be displayed and the annotation may be refined. For example, the concept "bus" has been entered. One may now choose the "details" button to see how this concept is related to other aspects in the ontology, for example that it is a kind of vehicle. Clicking on „refine" opens the option to directly specify the nature of the concept bus – in this case one may choose to specify the "color" as being "red".

3. SEMANTIC CONTENT SEARCH
Semantic Image Retrieval promises more precise search for documents than retrieval based on non-semantic keywords or full texts. In traditional keywords based systems searching for "blue car", would retrieve all images somehow labeled with "blue" as well as with "car" (e.g. including a picture of a blue bicycle next to a red car). Yet, an underlying ontology can capture the fact that

blue is a color which may be a property for several objects such as cars. Thus, an ontology-based retrieval system like IKen can select only those pictures, for which the annotation concept "car" has been specified with the color property "blue". To make use of the semantic annotations for precise document retrieval in the IKen user interface, search terms can be added in the same way as annotation concepts. A user may enter several search terms (which will again by replaced by corresponding ontology concepts) and may refine each of them with the available properties. Details regarding the interrelations to other aspects can be shown for every concept. This may offer suggestions for manual query reformulation and refinement.

4. THE IKen ONTOLOGY

The IKen Ontology has been formalized in OWL Lite. It was developed from scratch to exactly fit into the new system. The ontology comprises three main types of concepts: a) geographical concepts like <Country>, <City> and <District> as well as <Street>, <Building> and <Park>; b) concepts that directly relate to objects visible in the pictures like <Building>, <Vehicle>, <Clothing> or <Animal>; c) others are used to capture the sets of individuals that can be used as values of property axioms that themselves define attributes of the visual objects – examples are: <Colour> and <FacialExpression>.

During the first test runs with our system we discovered that trying to describe every visual object in the picture in detail is neither a feasible nor a desirable approach. We had to decide which kind of objects should be described in which detail. For the current version we identified the concept <Person> as being most important for the IKen application. It has the direct subclasses <Adult>, <Adolescent> and <Child> which in turn have further subclasses. Persons can be described using several object properties like <wearsClothing>, <hasHairColour> and <holdsInHand>.

The annotations of a photo are stored using instances of the concept <Photo>. For each photo one instance of <Photo> is created. Annotations belonging to a photo are asserted using the object property <hasAnnotation> with its sub properties <hasGeographicalSpaceAnnotation>, <hasEmotionAnnotation> and <hasCategoryAnnotation> which all have <Photo> as their domain. The property <hasGeographicalSpaceAnnotation> is functional and has the concept <GeographicalSpace> as its range. It is used to unambigously identify where a certain photo has been taken. With <hasEmotionAnnotation> general emotional features of a photo – like <Peace> or <Romantic> - can be added to the description. Likewise <hasCategoryAnnotation> is used to assert the photo to certain related categories like <Nature>, <Sports> or <Culture>. <hasCategoryAnnotation> assertions are automatically added to the photos using SWRL rules like the following:

Photo(?x) ∧ hasAnnotation(?x, ?y) ∧ Plant(?y) → hasCategoryAnnotation(?x, Nature)

This rule encodes the following: If x is an instance of <Photo> and x has an annotation of type <Plant> than x gets a <hasCategoryAnnotation> assertion with the instance <Nature> as value.

Metrics of the IKen ontology (25nd of July): 360 concepts/classes, 70 object properties, 6 datatype properties, 222 intial instances and 8 SWRL rules.

5. CONCLUSION & FUTURE WORK

First test runs with a small set of reference users showed that ontology supported image annotation and retrieval can significantly improve user experience. To empirically support these first results, tests to provide the evaluation of precision, recall and usability are currently prepared.

Terms used for photo annotations that cannot be mapped to the ontology constitute a valuable resource for ontology extension. Therefore future work on the IKen system will include functionalities to incorporate unrelated tags into the used ontology.

6. ACKNOWLEDGMENTS

Our thanks to the Ministry of Economic Affairs and Energy of the State of North Rhine-Westphalia, Germany, for funding the IKen project. Furthermore we would like to thank the following individuals who contributed advice and opinion: Nahal Ahmadinejad, Horst Hallmann, Klaus Hallmann, Indra Mainz, Birgit Mainz and Ingo Paulsen.

7. REFERENCES

[1] Peters, I. and Stock, W.G. 2007. Folksonomy and Information Retrieval. In Proceedings of the 70th ASIS&T Annual Meeting. Vol. 44, 1510-1542.

[2] Hollink, L., Schreiber, G., Wielemaker, J. and Wielinga, B. 2003. Semantic Annotation of Image Collections. In Proceedings of the KCAP'03 Workshop on Knowledge Capture and Semantic Annotation, Florida, October 2003.

[3] Hyvönen, E., Styrman, A. and Saarela, S. 2002. Ontology-based Image Retrieval. Number 2002-03 in HIIT Publications, Helsinki Institute for Information Technology (HIIT), Helsinki, Finland, 15-27.

[4] Schreiber, A. Th. (Guus), Dubbeldam, B., Wielemaker, J. and Wielinga, B. 2001. Ontology-Based Photo Annotation. IEEE Intelligent Systems, 16(3), 66–74.

[5] Lew, M., Sebe, N., Djeraba, C. and Jain, R. 2006. Content-based multimedia information retrieval: State of the art and challenges. ACM Transactions on Multimedia Computing, Communications, and Applications, (2)1, 1-19.

[6] Walter, A. and Nagypal, G. 2008. The Combination of Techniques for Automatic Semantic Image Annotation Generation in the IMAGINATION Application. In 5th European Semantic Web Conference Proceedings, ESWC 2008, Tenerife, Canary Islands, Spain, 879-883.

[7] Popescu, A., Moellic, P.-A. and Millet, C. 2007. SemRetriev: An Ontology Driven Image Retrieval System. In Proceedings of the 6th ACM Conference on Image and Video Retrieval, Amsterdam, The Netherlands. ACM Press, New York, 113–116.