Comparison of Different Ontology-based Query Expansion Algorithms for Effective Image Retrieval

C. H. C. Leung, Yuanxi Li
Department of Computer Science, Hong Kong Baptist University
{clement, yxli}@comp.hkbu.edu.hk

Abstract

We study several semantic concept-based query expansion and re-ranking scheme and compare different ontology-based expansion methods in image search and retrieval. In particular, we exploit the two concept similarities of different concept expansion ontology-WordNet Similarity, Wikipedia Similarity. Furthermore, we compare the keywords semantic distance with the precision of image search results with query expansion according to different concept expansion algorithms. We also compare the image retrieval precision of searching with the expanded query and original plain query. Preliminary experiments have been able to demonstrate that the two proposed retrieval mechanism has the potential to outperform unaided approaches.

1. Introduction

The presence of particular objects in an image often implies the presence of other objects. If term $U \rightarrow V$, and if only U is indexed, then searching for V will not return the image in the result, even though V is present in the image. The application of such inferences will allow the index elements T_i of an image to be automatically expanded according to some probability which will be related to the underlying ontology of the application. There are two types of expansion:

(a) Aggregation hierarchical expansion

This relates to the aggregation hierarchy of sub-objects that constitute an object.

The objects can be classified as:

(i) Concrete, where the relevant objects are well-defined (e.g. an orchestra expanded to conductors, violins, trumpets, clarinets etc, and see Fig. 1)

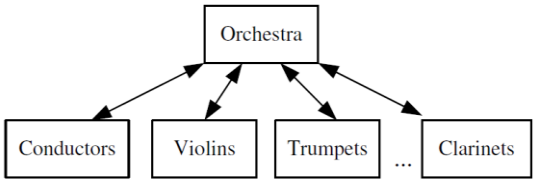

Fig. 1. The concrete aggregation hierarchical expansion, where the relevant objects are well-defined. In this example, an orchestra is expanded to conductors, violins, trumpets, clarinets etc.

(ii) Abstract, where the objects are not concretely defined (e.g. although 'conflict' is not a definite visual object, it does contain certain common characteristics).

Associated with each branch is a tree traversal probability t_{ij} (Fig. 2) which signifies the probability of occurrence of the branch index given the existence of the parent index. In general, the traversal probabilities of different object classes exhibit different characteristics, with $t_{ij} > t'_{mn}$ for t_{ij} belonging to the concrete object class, and t'_{mn} belonging to the abstract object class.

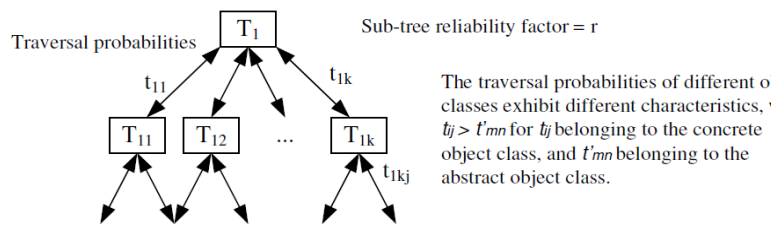

Fig. 2. A tree traversal probability t_{ij} which signifies the probability of occurrence of the branch index given the existence of the parent index.

(b) Co-occurrence expansion

This relates to the expectation that certain semantic objects tend to occur together. The relevant weighting is expressed as a conditional probability given the presence of other objects. An expansion to associate an image object O_j given the presence of object O_i is taken to be indexable when

$$Prob\ [O_j\ |O_i] \geq h' \qquad (1)$$

where h' is a preset threshold value that depends on the tradeoff between precision and recall performance of the system. More generally, complex probabilistic rules taking the form

$$Prob\ [O_j\ |O_1,\ldots, O_n] \geq h' \qquad (2)$$

will be applied. The ontology expansion tree is traversed bi-directionally in the course of the expansion. Top-down traversal will lead to an expansion factor > 1, while bottom-up traversal will have an expansion factor < 1 at each level of expansion.

There are, in general, many sub-trees whose roots are the nodes of the ontology expansion tree. Each sub-tree has been fully expanded, and it has an expansion reliability factor $0 < r < 1$, which signifies the dependability and completeness of the associated expansion. For high precision retrieval ($\pi \approx 1$), only sub-trees, which are having a significant reliability factor, need to be traversed; and nodes with a small value for r will be pruned. Decision rules linking expandability with π and r can be determined. [1]

In this paper, we mainly focus on WordNet Similarity and Wikipedia Similarity. Details of the two will be discussed in this paper later.

This paper is organized as follows. Related work will be introduced in section two; our ontology-based query expansion similarity measures will be presented in detail in section three; then the experiments, as well as results will be reported in section four, followed by a conclusion in the last section.

2 Related Work

2.1 Semantic Knowledge from Ontology Approach

As for redefining image indexing, the most popular way is to simplify the semantic knowledge into the semantic similarity between concepts. Semantic similarity [2] or semantic relatedness is a concept whereby a set of documents or terms within term lists are assigned a metric based on the likeness of their meaning / semantic content.

Concretely, this can be achieved for instance by defining a topological similarity, by using ontologies to define a distance between words (a naive metric for terms arranged as nodes in a directed acyclic graph like a hierarchy would be the minimal distance—in separating edges—between the two term nodes), or using statistical means such as a vector space model to correlate words and textual contexts from a suitable text corpus (co-occurrence).

2.2 Semantic Knowledge from Probabilistic Approach

Probabilistic query expansion is usually based on calculating co-occurrences of terms in documents and selecting terms that are most closely related to query terms [3]. In [4], it categorizes existing approaches for visual concept-based retrieval similarly into three categories for text-based approaches − lexical, global statistical, and local statistical approaches. Lexical approaches for visual concept-based query expansion are based on textual descriptions of the visual concepts, which essentially reduce the problem to that of lexical text-based query expansion. Typically, each concept is represented with a brief description or a set of representative terms (e.g., synonyms). Given a textual query, the query words are then compared against the concept descriptions, and any matched concepts are used for refinement, where the matching may be exact or approximate. Global techniques extract their co-occurrence statistics from the whole document collection and might be resource intensive, although some of the calculations can be performed offline. Local techniques extract their statistics from the top-n documents returned by an initial query and might use some corpus wide statistics such as the inverse document frequency, but they must be fast because they delay the response of the system. All calculations for local methods are done online; just after the user supplies the query and before presenting the results to the user [3].

3 Concept-based Similarity Distance

In our previous research [5], web images could be mainly categorized into the following four different types: (i) images without any caption nor annotation (ii) images with caption but no annotation nor tags (iii) images annotated with keywords and tags (iv) images provided with the full MPEG-7 annotation. For those with some basic caption, such primitive level of information may be exploited to carry out inferential reasoning based on domain content. It has been found in our study in [6] that captions may sometimes harm annotation correctness, and QBE techniques can be additionally deployed to attempt to filter out the misleading captions and provided keywords.

3.1 WordNet Similarity

WordNet [7], is one of these applications of semantic lexicon for the English language and is a general knowledge base and commonsense reasoning engine.

The purpose of the work is both to produce a combination dictionary-and-thesaurus that is more intuitively usable, and to support automatic text analysis and artificial intelligence applications.

For example, by using WordNet, ‘downtown’ has been expanded to ‘business district’, ‘commercial district’, ‘city centre’ and ‘city district’, while ‘city district’ has been expanded to ‘road’, ‘building’, ‘architecture’, ‘highway’ and ‘hotel’. The semantic knowledge is hierarchically expansible from the query terms and concepts and knowledge can be expanded extensively. The more extensive and complete such hierarchies, the greater the scope for rich semantic manipulation.

Recent research [8] on the topic in computational linguistics has emphasized the perspective of semantic relatedness of two lexemes in a lexical resource, or its inverse, semantic distance.

The first line of research [9], which brings together ontology and corpus, tries to define the similarity between two concepts c_1 and c_2 lexicalized in WordNet, named WordNet Distance (WD). It indicates by the information content of the concepts that subsume them in the taxonomy. Formally, define:

$$sim(c_1, c_2) = max(c \in S(c_1, c_2))[-logp(c)] \tag{3}$$

where $p(c) = \frac{\sum_{n \in words(c)} count(n)}{N}$ and N is the total number of nouns observed. And $S(c_1, c_2)$ is the set of concepts that subsume both c_1 and c_2. Moreover, if the taxonomy has a unique top node, then its probability is 1. In practice, we often measure word similarity rather than concept similarity. Using $s(w)$ to represent the set of concepts in the taxonomy that are senses of word w, define

$$wsim(c_1, c_2) = max(c \in S(c_1, c_2))[sim(c_1, c_2)] \tag{4}$$

where c_1 ranges over $s(w_1)$ and c_2 ranges over $s(w_2)$.

It defines two words as similar if near to one another in the thesaurus hierarchy. For example, refer to Fig. 3, ‘entity’ can expand to ‘inanimate-object’. Then ‘inanimateobject’ can expand to both ‘natural-object’ followed by ‘geological-formation’. It then expands to both ‘natural-elevation’ and ‘shore’. The former can expand to ‘hill’ while the latter expands to ‘coast’.

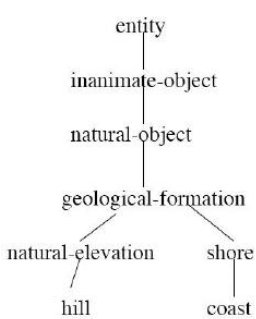

Fig. 3. Example demonstrate the way WordNet Distance defines two words are similar if nearby in thesaurus hierarchy

Beside the original work, researchers [10] propose a similarity measure between arbitrary objects. It uses the same elements but in a different fashion:

$$d(c_1, c_2) = \frac{2 \times \log p(lso(c_1, c_2))}{\log p(c_1) + \log p(c_2)} \tag{5}$$

3.2 Wikipedia Similarity

Wikipedia is the world’s largest collaboratively edited source of encyclopedic knowledge. In spite of its utility, its contents are barely machine-interpretable. Each article in Wikipedia describes a single topic; its title is a succinct, well-formed phrase that resembles a term in a conventional thesaurus. Meanwhile, each article must belong to at least one category of Wikipedia. Hyperlinks between articles keep many of the same semantic relations as defined.

WikiRelate [11] was the first to compute measures of semantic relatedness using Wikipedia. Their approach took familiar techniques that had previously been applied to WordNet and modified them to suit Wikipedia. Implementation of WikiRelate follows hierarchical category structure of Wikipedia.

The Wikipedia Link Vector Model (WLVM) [12] uses Wikipedia to provide structured world knowledge about the terms of interest. Their approaches are using the hyperlink structure of Wikipedia rather than its category hierarchy or textual content [13].

Probability of WLVM is defined by the total number of links to the target article over the total number of articles. Thus if t is the total number of articles within

Wikipedia, then the weighted value w for the link $a \rightarrow b$ is:

$$w(a \rightarrow b) = |a \rightarrow b| \times \log(\sum_{x=1}^{t} \frac{t}{|x \rightarrow b|}) \tag{6}$$

where a and b denotes the search terms.

For example, calculate the similarity between *Israel* and *Jerusalem*, one would consider only the nation and its capital city. The commonness of a sense is defined by the number of times the term is used to link to it: e.g. 95% of Israel anchors link to the nation, 2% to the football team, 1% to the ancient kingdom, and a mere 0.1% to the Ohio township. According to Equation 4.11, WLVM value of both terms *Israel* and *Jerusalem* is 0.994, which is completely reasonable.

4 Preliminary Experiments and Results

We compute the relatedness of the searching keywords and their related concept using WordNet Similarity and Wikipedia Similarity, and then search images by the plain keywords as well as the related words expanded from the WordNet Similarity computation and Wikipedia Similarity computation. As shown in Fig.4 and 5, for Example, the word "downtown" can be inferred to "Brooklyn", "Maryland", and so on according to WordNet Ontology Similarity; while "downtown" can be inferred to "driving", "tower", and so on according to Wikipedia Ontology Similarity. The blue lines (Relatedness line) indicate the performance of relatedness rates between concept 'downtown' and other 327 concepts computed by the WordNet distance/Wikipedia distance. The red lines (Precision 1 line) show the image search precision with only the expanded keyword such as search by "Brooklyn". The green lines (Precision 2 line) represent the image search precision with the expanded keyword combined with the original plain keyword such as search by "downtown + Brooklyn". We use Flickr as the Web Image Database the test the precision.

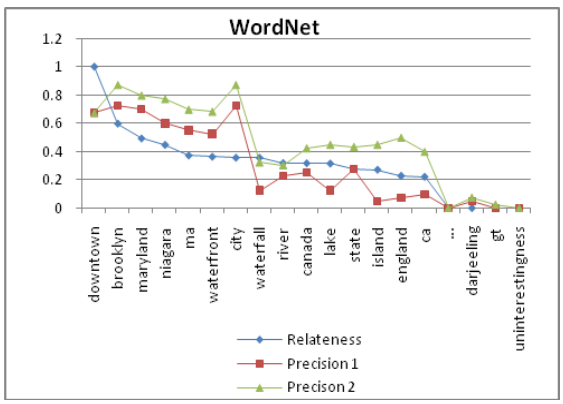

Fig. 4. Performance of relatedness rates between concept 'downtown' and other 327 concepts computed by the WordNet distance and image seaching precision

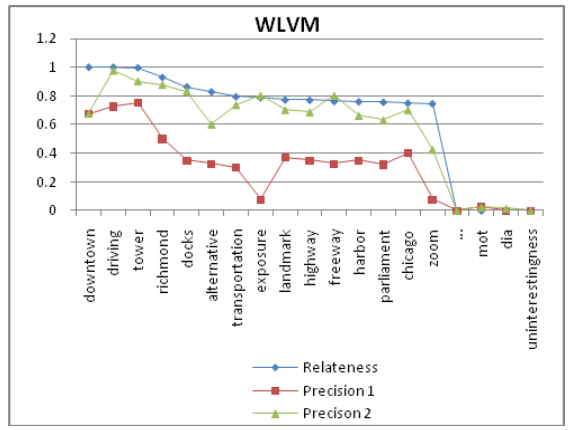

Fig. 5. Performance of relatedness rates between concept 'downtown' and other 327 concepts computed by the Wikipedia distance and image seaching precision

The precision is calculated as follows:

$$\Pr ecision = \frac{|\{relevant\_images\} \cap \{retrieved\_images\}|}{|\{retrieved\_images\}|} \quad (7)$$

We also do the experiment on other 100 concepts, 10,000+ web images. And the results are showing in table 1 and 2:

TABLE 1. PERFORMANCE OF SEMANTIC RELATEDNESS MEASURES OF WD ALGORITHMS WITH THEIR STANDARD DEVIATIONS

WD	Relatedness	Precision 1	Precision 2
Tagged Term	1	0.630	0.630
word 1	0.687	0.786	0.905
word 2	0.582	0.730	0.827
word 3	0.498	0.629	0.752
word 4	0.467	0.560	0.710
word 5	0.401	0.541	0.633
word 6	0.393	0.429	0.575
word 7	0.352	0.425	0.514
word 8	0.325	0.325	0.405
word 9	0.308	0.270	0.388
word 10	0.301	0.125	0.345
word 11	0.265	0.124	0.235
word 12	0.239	0.05	0.210
…	...	…	…
uninterestingness	0	0	0

TABLE 2. PERFORMANCE OF SEMANTIC RELATEDNESS MEASURES OF WLVM ALGORITHMS WITH THEIR STANDARD DEVIATIONS. THE TARGETED TERMS IS UNDERLINED.

WLVM	Relateness	Precision 1	Precison 2
Tagged Term	1	0.630	0.630
word 1	0.995	0.790	0.956
word 2	0.980	0.764	0.920
word 3	0.932	0.733	0.885
word 4	0.887	0.634	0.820
word 5	0.813	0.625	0.752
word 6	0.789	0.598	0.630
word 7	0.781	0.521	0.615
word 8	0.774	0.450	0.568
word 9	0.770	0.350	0.485
word 10	0.762	0.315	0.418
word 11	0.750	0.308	0.366
word 12	0.736	0.300	0.354
…	…	…	…

In these two tables, tagged term is the original keyword, such as "downtown". Word 1-word 12 are the expended words/concepts base on WD (WordNet Distance) and WLVM (Wikipedia Link Vector Model). Precision 1 column shows the image search average precision with only the expanded keywords of all tested tags. Precision 2 column represents the image search precision with the expanded keyword combined with the original plain keywords of all tested tags.

As we can see in Table 1 and 2, average precision are all satisfactorily higher compared with the only searching by the original terms. The value of precision increases with the relatedness increases. The performance of semantic relatedness of WD and WLVM could affect the image search performance positively. Meanwhile, the image search precision line goes with the WLVM semantic similarity line a little better than WD semantic similarity line. Furthermore, the precision of searching with original concept combined with the expanded concepts are higher than searching with only expanded keywords. These results indicate that significant improvement in performance may be attained from using the keywords expansion approach base on WordNet or Wikipedia ontology.

5 Conclusion

We explore numbers of semantic relatedness measure algorithms and developed an ontology-based, i.e. WD, WLVM. Our system is evaluated quantitatively, and experimental results indicate that this approach is able to deliver highly competent performance. Our approach, both search web image with the additional concepts expanded by WordNet Similarity and Wikipedia Similarity, not only demonstrates the applicability of ontology to the image annotation problem, but also using the sub-objects as surrogate terms for general queries improves the precision in the image sets.

However, there still exist some limitations in our proposed methods. This thesaurus based approach has its limitations: lots of detailed semantic knowledge present has been lost and it does not provide a quantitative measure of the semantic similarity between concepts. Our future work consist of developing algorithms for increasing image retrieval precision of our model, and working out more accurate and fast re-ranking scheme, to achieve greater user satisfaction.

6 Reference

1. C.F. Wong, Automatic Semantic Image Annotation and Retrieval. PhD Thesis, Hong Kong Baptist University, August 2010.
2. Wikipedia Article website- Semantic similarity: http://en.wikipedia.org/wiki/Semantic_similarity.
3. A.Andreou, Ontologies and Query Expansion. 2005.
4. A. Natsev, A. Haubold, J. Tesic, L. Xie, R. Yan, Semantiv Concept-Based Query Expansion and Re-ranking for Multimedia Retrieval. Proceedings of the 15th ACM international conference on Multimedia, page 991-1000. New York, NY, USA, 2007.
5. Y.X. Li and C.H.C Leung, "Multi-level Semantic Characterisation and Refinement for Web Image Search," The 2nd International Conference on Innovative Computing and Communication, pp. 70-73, 2011.
6. R. C. F. Wong and C. H. C. Leung, "Automatic Semantic Annotation of Real-World Web Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 11, pp. 1933–1944, 2008.
7. G. A. Miller. Wordnet: a lexical database for english. Communications of the ACM, 38(11):39–41, 1995.
8. A. Budanitsky and G. Hirst. Semantic distance in wordnet: An experimental, application-oriented evaluation of five measures. In Proceedings of Workshop on WordNet and Other Lexical Resources, page 641, Pittsburgh, PA, USA, 2001. North American Chapter of the Association for Computational Linguistics.
9. P. Resnik. Using information content to evaluate semantic similarity in a taxonomy. In Proceedings of the 14th International Joint Conference on Artificial Intelligence, pages 448–453, 1995.
10. D. Lin. An information-theoretic definition of similarity. In Proceedings of the 15th International Conference on Machine Learning, pages 296–304. Morgan Kaufmann, 1998.
11. M. Strube and S. P. Ponzetto. WikiRelate! computing semantic relatedness using wikipedia. In Proceedings of the

Twenty-First National Conference on Artificial Intelligence. AAAI Press, July 2006.
12. D. Milne and I. H. Witten. An effective, low-cost measure of semantic relatedness obtained from wikipedia links. In WIKIAI'08: Proceedings of first AAAI Workshop on Wikipedia and Artificial Intelligence, Chicago, IL, USA, 2008.
13. M. V¨olkel, M. Kr¨otzsch, D. Vrandecic, H. Haller, and R. Studer. Semantic wikipedia. In WWW '06: Proceedings of the 15th International Conference on World Wide Web, pages 585–594, New York, NY, USA, 2006. ACM.
14. C.H.C Leung and Y.X. Li, "CYC Based Query Expansion Framework for Effective Image Retrieval," The 4th International Congress on Image and Signal Processing 2011.
15. C. H. C. Leung, W.S. Chan, A. Milani, J. Liu and Y.X. Li, "Intelligent Social Media Indexing and Sharing Using an Adaptive Indexing Search Engine", ACM Transactions on Intelligent Systems and Technology (To appear in 2011).
16. A. M. Tam, C. H. C. Leung, "Semantic Content Retrieval and Structured Annotation: Beyond Keywords", ISO/IEC JTC1/SC29/WG11 MPEG00/M5738, Noordwijkerhout, Netherlands, March 2000.