International Journal of Computer Science & Emerging Technologies (E-ISSN: 2044-6004)
Volume 2, Issue 1, February 2011

Semantic Query Expansion Using Knowledge Based for Images Search and Retrieval

Roohullah Khan[1] and J. Jaafar[2]
Department of Computer and Information Sciences
UniversitiTeknologi PETRONAS,
Bandar Seri Iskandar, 31750 Tronoh,Perak, Malaysia.
[1]Roohullah_orc@yahoo.com, [2]jafreez@petronas.edu.my

Abstract—*The falling prices of multimedia and storage devices make almost everyone to act like a professional to capture photo and archive them for later use. Without efficient retrieval methods the search of images in large collections can become a painstaking work. Most of the traditional image search engines rely on keyword-based annotations which lacks the query semantic space equivalent to the annotation semantic space, because of the difficulty in describing the same concepts with other keywords. In this paper, we propose a novel approach for the query expansion using lexical and commonsensical knowledgebases like WordNet and ConceptNet, which will not only fill the gap in the semantic space between user query and annotation but will also provide an opportunity to discard the less important words from the query semantic space. For evaluation we have selected LabelMe datasets, which is openly available for researcher.*

Keywords**:** *Content Based Retrieval, Semantic Gap, Query Expansion, WordNet, ConceptNet.*

1. Introduction

With the increase of the digital media both online and offline, there is a growing increase in the demand for the system that can process, store, organise and manage the digital media efficiently and effectively. Processing and managing such an ever increasing amount of data is a great challenge. Keeping this, it is impossible for the user to manually search the relevant images from the large image corpus. This explosive growth of the digital media [1] without appropriate management mimics its use.

Currently, the multimedia search and retrieval are an active research dilemma among the academia and the industry. The large data is available an online repositories like Google, YouTube, Flicker, etc. provides required images or videos through text based matching techniques [31] (surrounded text around the images likes an image name, metadata) in which the novice user has hardly found and accessing the useful images or videos of interest and becomes difficult. Finding the image of interest becomes harder and harder. The area of the textual information retrieval is matured. Nevertheless, the image retrieval is still worth investigating. Due to this explosive growth, there is a strong urge for the system that can efficiently and effectively interpret the user demand for searching and retrieving the relevant information images.

Based on above reasons new semantic Query Expansion techniques using knowledge based for Images search and retrieval will be proposed and develop. This technique should be able to convert a user demand into set of discrete concepts. The semantic query algorithm which would be automatically interpreting the query according to the user's requirements.

The proposed technique will expand the query and interprets the user query semantically so that it can be further processed for the accurate retrieval. The proposed query engine is the text based query engine that will expand the user query by combining WordNet [19], ConceptNet [20] knowledge bases to retrieve the results semantically with higher accuracy.

The image is ultimately a congregation of objects depicts some concepts. For a computer, an image is just affection a mix of pixels that are characterized by the low-level features like colour, shape, texture, etc. while for the human it is more than that. For human an image is the mix of one or more semantic idea. For them, it refers to, not the content of the image that's appearing, nevertheless rather a semantic idea that it representing. It is worth saying that for the same images dissimilar mankind extractsseveralconcepts depend on the nature and knowledge of the human. Due to the open-ended nature of the human and the hard coded computer nature there appears a problem known as the semantic gap. Which are showing in Figure 1. Semantic gap is one of the key problems between the Computer interpretation and human understanding for visual information.

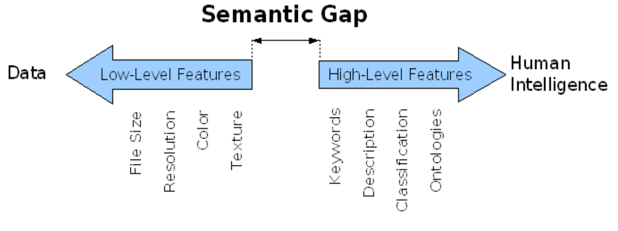

Figure 1. Semantic Gap

According to Smeulders et. al. [32] Semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation".

Although the start-of-the-art image retrieval techniques or the Content Based Images Retrieval (CBIR) system are a day by the day getting more and more powerful and can now achieve the accuracy up to some extent. Extensive research effort was done in retrieving the image assuming their visual content such as Query By Image Content (QBIC) [24], Netra [25], VisualSeek [26], WeebSeek [27], Virage[28], Videoq[29], Multimedia Analysis and Retrieval System (MARS) [30]etc. Indeed, the power of these tools doesn't reduce the semantic gap. And now the trend is completely moved from low-level features to the high-level idea.

The process of extracting the semantic idea from the images through hard coded machine remains a difficult it will be creating the following general research questions.

i. How can we interpret the user requirements that are given to the system in the form of query?

ii. How to reduce the semantic gap for Images of human intelligent and machine interpretation?

iii. How to bring the correctness to the system result along with the semantic proficiency?

2. Related Work

Content Based Retrieval (CBR) system is the widely uses for the Image retrieval but this system has retrieved the images through Content based query. CBIR is the techniques that retrieve the images through low-level feature like colour, shape and texture.

CBIR uses the visual content of image such as color, shape, texture, etc. some fundamental techniques for content based image retrieval include visual content description, similarity distance measures, indexing scheme as shown in the Table 1. Some retrieval system has incorporated user relevance feedback in order to facilitate the retrieval process.

Table 1: Depicts the content based image retrieval by using low-level features

Content Based Image Retrieval	
Feature components	Techniques
Color	Global histogram
	Correlation histogram
	Average color vector
	Color coherence
	Dominant color
	Region histogram
Texture	Wavelet
	Random field
	Automatic texture feature
Shape	Template matching
	Elementary description
	Fourier description

Query by Colour is the most prominent visual feature in CBIR since it is well correlated with human visual perceptions of objects in an image. Retrieving images based on color [7], [8] similarity is achieved by the color histogram for each image. A color histogram is a representation of the distribution of colors in an image, derived by counting the number of pixels of each of a given set of color ranges in a typically two-dimensional (2D) or three-dimensional (3D) color space. Color searches will usually involve comparing color histograms. When comparing is matching then retrieve that type of images. Similarly query by shape CBIR is also matching the Retrieving the images based on the shape include the shape [9], [10] of a particular region that is being sought out. While for the Texture CBIR look for visual patterns in images[11], [12] and how they are spatially defined. Textures are represented by texels or texture elements (also texture pixel), and retrieve the images through texture, wavelets. That the same texture matches to other images. However, the Deok-Hwan Kim andSeung-Hoon Yu [13] have been using CBIR through the region based instead to match the entire image. However, they can also get low level precision from the entire results. However,Ying Liu et al [7] have also been using the region based content matching and get little more accurate results. Nevertheless, the region based query has retrieved the images in a specific part of a region exist so through this query a lot of irrelative images have been retrieved, and the accuracy will be lower.

Query by example [10] is matching the input image as a hole through colour, shape, and texture instead as a part, and retrieve that images which are totally matching to this image. A query – by-example (query-by-image) retrieves the images based on the low level features. However, the visual feature lacks the semantics. Sometimes the images are visually similar, but it cannot contain similar information as shown in the Figure 2. While sometimes the images may be visually different but contain the same information as shown in the Figure 3. All these led to the semantic gap.

Figure 2. Depicts the visually similar but semantically different images

Figure 3. Depicts the visually different but semantically similar images

Furthermore, others query techniques have been using to find images from the large image corpus like querying by visual sketch [14], querying by direct specification for image features [15] and multimodal queries [13] (e.g. that combining touch, voice, etc.). However, all these CBIR techniques have not retrieved the images through semantic similarity? Now the trend has completely changed from the low level feature toward a high level semantic idea.

Semantics defines the concepts at a high level such as the objects, events, scene and the relationship among them.E.g. “Burning of wood in the street”. Where the wood is an object burning is an event, and street is a scene. It tells us that “What is actually happening in the image”.

Query plays a vital role in the performance of the information retrieval systems. Sometimes the user queries cannot define about they actually needs or sometimes the vocabulary in the query is inconsistent [6] with that in the relevant document. In order to solve this problem the general purpose knowledge bases are used such as WordNet, ConceptNet. Query expansion [2]-[5] for text based searching is too much successfully and given more accurate results with higher accuracy. This idea of the WordNet with query expansion is firstly, implemented by Zhiguo Gong et al [16], where WordNetis used as the basic expansion rules and then usesWordNet Lexical Chains and semantic similarity to assign terms in the same query into different groups with respect to their semantic similarities. Yokoyama et al [17], expand the query for the new users surfing the internet, they expand the query terms by WordNet, and the expanded query is then submitting to the search engine for getting most related web results. Ming-Hung Hsu et al. [18] combine the WordNet and ConceptNet knowledge bases for query term expansion.

WordNet is a lexical database for the English language [19]. It has developed by the George A. Miller under supervision. WordNetgroup's English words into a set of synonyms called synsets, provides short, general definitions, and records the various semantic relations between these synonym sets. The purpose is twofold: to produce a combination of dictionary and thesaurus that is more intuitively usable, and to support automatic text analysis and artificial intelligence applications. The latest version ofWordNethasto contain over 155.00 words, grouped into over 117.00 synsets.

ConceptNet is a freely available, machine-usable common sense resource [20]. ConceptNet is developed by MIT Media Laboratory and is presently the largest commonsense knowledgebase. It is the relational semantic network that is automatically generated from about 700,000 English sentences of the Open Mind Common Sense (OMCS) corpus ConceptNet 3 presently consist of over 250,000 elements of common sense knowledge. The ConceptNet aims to give computer access to common-sense knowledge, the kind of information that ordinary people know [21].

3. Propose Method

Taking into account all the above issues we proposed a query expansion technique that will expand the user query lexically as well as conceptually to reduce the semantic gap. The user queries can be expanded lexically by using an open source lexical knowledge base, i.e. WordNet while the conceptual expansion can be done through ConceptNet. The overall framework of the proposed model is shown in Figure 4.

Throughout this model the user will be getting the images through searching from the large image corpus semantically. So this model is suitable to minimize the semantic gap between human understanding and machine interpretation. This model is also help to the user whoselects those words from the expanded query those are more relatives to the user original query and implementation for the searching. While stop those words which are fewer degrees of relevance have related to the original query.

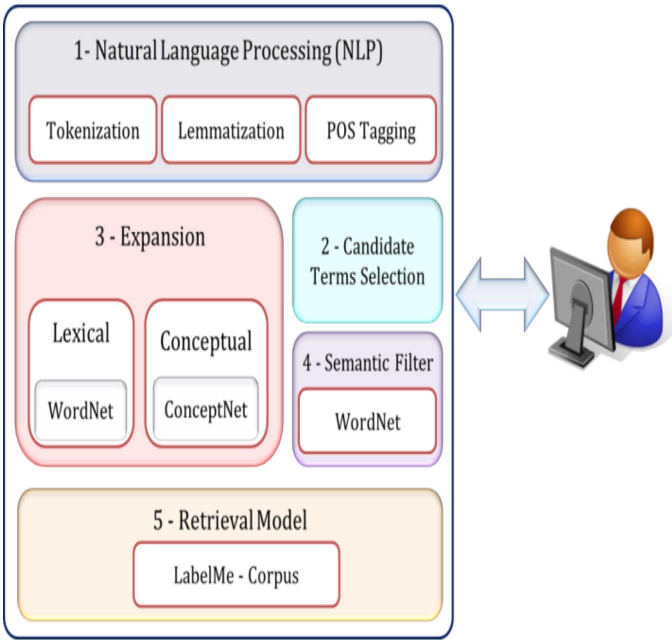

Figure 4. Depicts the overall model of the proposed framework

In this model, there are five (5) parts that get the user query and passing through this each step and after passing these steps the result will be display to a user. These five parts have to discuss details as below:

1. *Nature Language Processing:* Natural Language processing (NLP) is a field of computer science and linguistics concerned with the interactions between computers and human (natural) languages. There are further three parts of this part.
   - *Tokenization:* To break the sentence into words is called tokenization that each word separate from the sentence.
   - *Lemmatization:* In linguistics, lemmatization is the process of grouping together the different inflected forms of a word, so they can be analysed as a single item or based form.
   - *POSTage:* is the process of marking up the words in a text (corpus) as corresponding to a particular part of speech, based on both its definition, as well as its context — i.e. Relationship with adjacent and related words in a phrase, sentence, or paragraph?
2. *Candidate term selection:* Candidate term selection means that just select the noun, verbs and adverb from the user search query for expanding.
3. *Expansion:* To expand the query through lexically from WordNet and Conceptually from ConceptNetto add more relative terms to the user original query.
4. *Semantic filter:* Semantic filters mean that filter the words matching with query with a select degree of relevancies that's stopping the irrelative or unusual words after the expansion and get the semantic relative words from the expanded.
5. *Retrieval Model:* In retrieving a model the search engine will be retrieving the result from the large image corpus that stores the images with relative annotations.

This model will be implemented in the label dataset [22]. Which are containing 31.8 GB datasets that containa total of 181, 932 images with 56946 annotated images, 352475 annotated objects and total of 12126 classes? The result will be containing through Vector Space Model.

Vector space model [23] is the information retrieval for the algorithm and framework.This model is an algebraic model to representing documents as a vector of identifiers. Vector space model is used in information filtering, information retrieval, indexing and relevancy rankings.The terms of a query surrogate can be weighted to take into account their importance, and they are computed by using the statistical distributions of the terms in the collection and in the documents [Salton 1983]. The vector space model can assign a high ranking score to a document that contains only a few of the query terms if these terms occur infrequently in the collection but frequently in the document.

This model represents documents and queries as vectors.

$$d_j = (w_{1,j}, w_{2,j}, \ldots, w_{t,j})$$

$$q = (w_{1,q}, w_{2,q}, \ldots, w_{t,q})$$

Each dimension corresponds to a separate term [33]. If a term occurs in the document, its value in the vector is non-zero.

Relevance rankings of documents in a keyword search can be calculated, using the assumptions of document similarities, by comparing the deviation of angles between each document vector and the original query vector where the query is represented as a same kind of vector as the documents.

To calculate the relevance ranking of document Cosine is very easily calculating the angle between the vectors instead of the angle [34].

$$\cos\theta = \frac{\mathbf{d_2} \cdot \mathbf{q}}{\|\mathbf{d_2}\| \, \|\mathbf{q}\|}$$

If the cosine value is equal to zero means that the query and document vector is orthogonal and have no match, i.e. no documents have match to the query.

The vector space model makes the following assumptions:

i. The more similar a document vector is to a query vector, the more likely it is that the document is relevant to that query.
ii. The words used to define the dimensions of the space are orthogonal or independent. While it is a reasonable first approximation, the assumption that words are pairwise independent is not realistic.

Vector space model is same to Boolean model, but it's the advance model. Which are some advantages?

i. Simple model to represent results based on linear algebra.
ii. Term weights not like a Binary model.
iii. Allow computing a continuous degree of similarity between query terms and documents.
iv. Allow ranking documents according to their possible relevance.
v. Allow partial matching of query terms to the documents.

4. Discussion

The main limitation of the past approaches was that the content based retrieval is an automatic solution for theretrieval, but they rely only on the low level feature's extraction. Now the question is that whether the low level feature extraction alone is enough for efficient searching and retrieval?

The answers might be no, because the low level feature only captures one aspect of the multimedia data. In addition, sometimes the images or videos that look similar are not semantically similar. So the retrieval results that are solely based on low level feature extraction are mostly unsatisfactory and unpredictable. This opens a new era for the research community to diverge from the existing methodologies to new a paradigm or new direction that there is something behind the visual features that need to be considered for accurate searching and retrieval.

That is the semantic of the multimedia data, i.e. high level features. Modeling the high level features are difficult than the low level features as the low level features are totally based on the colour, shape, texture structure while the high level feature isdepending on the semantics.

5. Conclusion

In a nutshell, they propose model use query expansion techniques for extracting the semantics from the user query. The model an effective way interpreting the user demand keeping in view the flexible nature of human as well as hard coded nature of computer. This model will be reducing the semantic gap by interpreting the user demand semantically in order to achieve the semantic accuracy as well as the efficiency in the retrieval.

Reference

[1] IDC white paper, "The diverse and exploding Digital Universe", sponsored by EMC. March 2008.
[2] J.Li, M. Guo and S. Tian "A New Approach To Query Expansion" 4th international conf. on Machine learning and cybernetics, Guangzhou, IEEE, 2005.
[3] Z. Yu, Z. Zheng, S. Tang and J. Guo "Query Expansion for answer document retrieval in chinese question answering system" 4th international conf. and cybernetics, Guangzhou, IEEE, 2005.
[4] H. Cui, J. Wen, J. Nie and W. Ma "Query Expansion by Mining User Logs" IEEE, vol. 15, No 4, pp. 829-839, 2003.
[5] B. Sun, P. Liu and Y. Zheng "Short Query Refinement with Query Derivation" Springer-Verlag Berlin Heidelberg, pp. 620-625, 2008
[6] JinxiXu., "Solving the word mismatch problem through automatic text analysis.", Ph.D. Thesis, Department of Computer Science, University of Massachusetts, Amherst, MA, USA, May 1997.
[7] Y. Liu, X. Chen, C. Zhang and A. Sprague "Semantic clustering for region-based image retrieval" J. Vis. Commun. Image R. 20, pp. 157–166, 2009.
[8] M. ElAlami "Supporting image retrieval framework with rule base system" Knowledge-Based Systems (2010) .
[9] D. He and D. Wu "Enhancing query translation with relevance feedback in translingual information retrieval" Information Processing and Management 47, pp. 1–17, 2010
[10] H. Zhao and W. Jia, "An Adaptive Fuzzy Clustering Method for Query-by-Multiple-Example Image Retrieval" Third

International IEEE Conference on Signal-Image Technologies and Internet-Based System, pp. 997-1004, 2007

[11] H. Imran and A. Sharan "THESAURUS AND QUERY EXPANSION" International Journal of Computer science & Information Technology (IJCSIT), Vol 1, No 2, pp. 89-97, 2009

[12] Z. Zhou, Y. Tian, Y. Li, T. Huang and Wen Gao "Large-Scale Cross-Media Retrieval of WikipediaMM Images with Textual and Visual Query Expansion" Springer-Verlag Berlin Heidelberg, pp. 763-770, 2009

[13] D. Kim and S. Yu "A new region filtering and region weighting approach to relevance feedback in content-based image retrieval" The Journal of Systems and Software 81, pp. 1525–1538, 2008.

[14] M. Eitz, K. Hildebrand, T. Boubekeur, M. Alexa "An evaluation of descriptors for large-scale image retrieval from sketched feature lines" Computers & Graphics 34, pp. 482–498, 2010

[15] L. Cinque, G. Ciocca, S. Levialdi, A. Pelocano and R. Schettini "Color-based image retrieval using spatial-chromatic histograms" Image and computing 19, pp. 979-986, 2001.

[16] Zhiguo Gong, MaybinMuyeba, JingzhiGuo, "Business information query expansion through semantic network", pp. 1-22, Enterprise Information Systems Volume 4 Issue 1, February 2010

[17] Yokoyama, A. Klyuev, V., "Search Engine Query Expansion Using Japanese WordNet", Human-Centric Computing (HumanCom), pp. 1-5, 3rd IEEE international conference, August, 2010

[18] Ming-Hung Hsu, Ming-Feng Tsai, Hsin-Hsi Chen, "Combining WordNet and ConceptNet for automatic query expansion: a learning approach", pp. 213-224, 4th Asia information retrieval conference on Information retrieval technology, AIRS'08, 2008

[19] C. Fellbaum, WordNet: an electronic lexical database. Cambridge, Mass: MIT Press, 1998.

[20] H. Liu and P. Singh, "ConceptNet a practical commonsense reasoning tool-kit," BT Technology Journal, vol. 22, no. 4, pp. 211–226, 2004.

[21] Query Expansion with ConceptNet and WordNet: An Intrinsic Comparison by Ming-Hung Hsu, Tsai and Hsin-His Chen Springer 2006.

[22] B. Russell, A. Torralba and W. Freeman The open annotation tool http://labelme.csail.mit.edu/ Computer Science and Artificial Intelligence Laboratory, University MIT, 2005.

[23] Silva, J. Souza and K. Santos "Dependence Among Terms in Vector Space Model" Proceedings of the International Database Engineering and Applications Symposium, IEEE, 2004

[24] M. Flicker, H. Sawhney, W. Niblack and J. Ashley et al "Query By Image and Video Content: The QBIC System" 11th Annual Computer Security Application Conference, IEEE, pp. 23-32, 1995.

[25] W. Ma and B. Manjunath "NeTra: A toolbox for navigating large image databases" Multimedia Systems 7, pp. 184–198 Springer-Verlag 1999.

[26] J. Smith and S. Chang "VisualSEEk: a fully automated content-based image query system" ACM Multimedia 1996.

[27] J. Smith and S Chang "VisuallySearchingtheWeb for Content"pp. 12-20, IEEE Multimedia, 1997.

[28] Gupta "Visual Information Retrieval TechnologyAVirage Perspective" Virage Image Engine API Specification 1997.

[29] S. Chang, W. Chen, H. Meng, H. Sundaram and D. Zhong "VideoQ: An Automated Content Based Video Search System Using Visual Cues" ACM Multimedia, pp. 313-324, 1997

[30] M. Binderberger, S. Mehrotra, K. Chakrabarti and K. Porkaew " WebMARS: A Multimedia Search Engine" IEEE, 1999

[31] B. Luo, X. Wang and X. Tang "A World Wide Web Based Image Search Engine Using Text andImage Content Features" Proceedings of SPIE-IS&TElectronic Imaging, SPIE Vol. 5018, pp. 123-130, 2003

[32] Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. 22, pp.1349–1380, 2000

[33] I. Silva, J. Souza and K. Santos "Dependence Among Terms in Vector Space Model" Proceedings of the International Database Engineering and Applications Symposium (IDEAS'04), IEEE, 2004.

[34] X. Tai, F. Ren and K. Kita" An information retrieval model based on vector space method by supervised learning" Information Processing and Management 38, Elsevier Science, pp. 749–764, 2002.