Adaptive Semantic Construction for Diversity-based Image Retrieval

Ghada Feki, Anis Ben Ammar and Chokri Ben Amar
REGIM: REsearch Group on Intelligent Machines,University of Sfax, ENIS, BP W, 3038, Sfax, Tunisia

Keywords: Diversity, Semantic, Ambiguity, Image Retrieval, Wikipedia.

Abstract: In recent years, the explosive growth of multimedia databases and digital libraries reveals crucial problems in indexing and retrieving images, what led us to develop our own approach. Our proposed approach TAD consists in disambiguating web queries to build an adaptive semantic for diversity-based image retrieval. In fact, the TAD approach is a puzzle constituted by three main components which are the TAWQU (Thesaurus-Based Ambiguous Web Query Understanding) process, the ASC (Adaptive Semantic Construction) process and the DR (Diversity-based Retrieval) process. The Wikipedia pages represent our main source of information. The NUS-WIDE dataset is the bedrock of our adaptive semantic. Actually, it permits us to perform a respectful evaluation. Fortunately, the experiments demonstrate promising results for the majority of the twelve ambiguous queries.

1 INTRODUCTION

Given the popularity of the internet and the massive growth of document collections, applied researchers have recently become increasingly interested in web search. Consequently, many investigations have lately turned to information retrieval domain what leads to the continuous developing of search engines. Our interest is mainly in image search engines like Google, Bing and Flickr. We notice that for ambiguous query, the returned images belong only to one or two well known meanings.

In this paper, we give preliminary result of our proposed image search approach (TAD approach). We describe our first perception of an adaptive semantic construction for diversity-based image retrieval. Our system distinguishes different meanings derived from the query despites of the ambiguity coming with the user words.

The reminder of this paper is divided into four sections. In section 2, we explore related works for existing image retrieval systems. In section 3, we gather an idea about our strategy in diversity-based image retrieval which is achieved thanks to the query disambiguation and the adaptive semantic construction. In section 4, we show our experimental results.

2 RELATED WORKS: DIVERSITY CHALLENGE FOR IMAGE SEARCH ENGINES

In this section, we will not devalue existing commercial search engines or academic approaches, but we will just discuss about their initial search results organization.

A standard procedure for image retrieval aims to return satisfactory images for the user according to his announced query. The main steps are the query formulation, the query interpretation and the result ranking. The query formulation consists in how the user reveals his need. The query interpretation is the fact of rewriting the query in a more efficient way and estimating the cost of various execution strategies (Fakhfakh et al., 2012). The result ranking consists in classifying at the top, the images that have the highest degrees of similarity according to the query (Feki et al., 2012).

The rate of diversity for a given image retrieval system is revealed in the obtained image organization (Feki et al., 2013). In fact, every system has its own method to declare different meanings which are generally announced through textual or visual suggestions for a next re-ranking step (Ksibi et al., 2013). However, we think that it is necessary to cover all the meanings from the first image ranking which is generally a step of an image pool

Feki G., Ben Ammar A. and Ben Amar C..
Adaptive Semantic Construction for Diversity-based Image Retrieval.
DOI: 10.5220/0005157104440449
In *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval* (KDIR-2014), pages 444-449
ISBN: 978-989-758-048-2

text-based construction (Upadhyay et al., 2014). This pool can be structured by fixing the number of retrieved images for each meaning according to their relatedness score to the main topic (Hoque et al., 2012) and by determining the similarity of a given image to its neighbours. Obtained images are organized founded on their visual similarity on a 2D virtual canvas. As for the re-ranking step, the user can use a concept hierarchy to focus and filter the search results, or perform visual filtering through pan and zoom operations (Hoque et al., 2013).

Google[1] Image search engine ameliorates the query understanding by performing synonyms and searching in the title and the surrounding text (MacKinlay, 2005) in order to cover the majority of the meanings and provide textual suggestions, but it focus on the most well known meaning in ranked images and gives similar images (Chechik et al., 2010). Picsearch[2] offers also textual suggestions to the user to boost the right understanding of the ambiguous query and provides both simple and deep search options, with the possibility of limiting the search by colour, or to one of seven different gradations of size.

Bing[3], the search engine of Microsoft, refines the result with five filters based on image size, aspect ratio, colour or black and white, photo or illustration, and facial features recognition. As well, Exalead[4] Images search engine presents a variety of filters. In fact, it supplies the ability to specify the type of images you wish. Details are also about the size, the format (jpeg, gif or png), the orientation (landscape or portrait), the dominant colour and the content (image contains face) of the image. For ambiguous query, the two last details can improve the understanding of the user intent. Nevertheless, the face filter becomes not efficient if there are many famous persons who refer to the same query.

The Incogna[5] search engine tries from the beginning of the search to cover all the sub-topics. Therefore, unlike the major search engines, the more you put keywords in your query, the more the result is relevant. Then, if the user chooses an image describing his intent, Incogna will provide similar images thanks to its large-scale searchable visual index built by processing the shape in every image.

Finally, having a large number of high quality images, Flickr[6], a very large photo-sharing site, enhances the Yahoo image retrieval process. However, as for scientific images search, Flickr may not be an ideal source and its insertion risks obscuring adequate images. Unfortunately, it often discards the diversity factor and considers only one meaning.

Hence, the image search engines do not always guess the right meaning and follow the user intent. Even if they did, the ranked images would belong always to only one or two meanings. In fact, a text has not always expressed exactly the content of an image especially when the query is specified with few terms or the method of the query understanding seems limited.

3 OUR PROPOSED TAD APPROACH

3.1 Strategy

We propose the TAD approach for image retrieval system. The TAD approach is a puzzle constituted by three main components which are the TAWQU (Thesaurus-Based Ambiguous Web Query Understanding) process, the ASC (Adaptive Semantic Construction) process and the DR (Diversity-based Retrieval) process.

The principal input is a textual query. According to the standard image retrieval process announced previously in the second section, the process must include a query formulation step, but we ignore it as we have only one term and there is no need to filter the query by eliminating the senseless words. Therefore, we turn straight to the query interpretation by executing the TAWQU process in order to extract different meanings for a given query. The obtaining of these weighted subtopics launches the ASC process, where new concepts are added to construct our adaptive semantic. Then, based on extracted weighted subtopics, we return diverse images and textual and visual suggestions for users.

The Figure 1 describes our strategy, which ensures the ability of covering the majority of the possible meanings for a given ambiguous query by providing diverse resulting images.

[1] http://images.google.com
[2] http://fr.picsearch.com/
[3] https://www.bing.com
[4] http://www.exalead.com/search/image/
[5] http://www.incogna.com/#random
[6] https://www.flickr.com/explore

3.2 Thesaurus-based Ambiguous Web Query Understanding Process

The TAWQU process extracts the main topic announced by the user words in a first step. Then, it mines the useful information from the Wikipedia pages. Finally, it illustrates the query understanding.

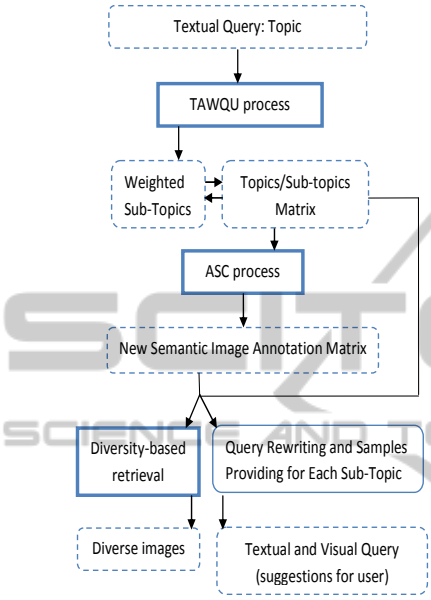

Figure 1: TAD strategy.

To gain time, the TAWQU process identifies the ambiguous queries. If the query is explicit, a simple relevance-based retrieval is running. If the query is ambiguous, and typed for the first time, the reformulation procedure starts to obtain weighted subtopics.

Our image search works in an incremental way. Indeed, the system treats the new queries by browsing the entire algorithm from the phase of the ambiguity identification. After achieving this step, it adds such given query (topic) and its subtopics to the graph which illustrates the relations between a given topic and its associated sub-topics in order to facilitate and speed up the next dealing with this query.

3.3 Adaptive Semantic Construction Process

The ASC process provides a dynamic list of concepts. Two key points constitute the ASC process. First, the online access to the Wikipedia resources allows us to have an updated semantic. Second, the exploitation of the NUS-WIDE data permits us to have a great collection of images and a considerable basic list of concepts. In fact, NUS-WIDE is a real-world web image database from national university of Singapore. This web image dataset is created by NUS's lab for media search. The dataset includes:

- 269,648 images and the associated tags from Flickr, with a total of 5,018 unique tags;
- Six types of low-level features extracted from these images, including 64-D color histogram, 144-D color correlogram, 73-D edge direction histogram, 128-D wavelet texture, 225-D block-wise color moments extracted over 5×5 fixed grid partitions, and 500-D bag of words based on SIFT descriptions;
- Ground-truth for 81 concepts.

As it shown in the Figure 2, we adapt the annotation of images in the NUSWIDE database with the context of the ambiguous query understanding for the diversity-based image retrieval. In fact, images are indexed according to a new list of concepts, which is the fruit of the semantic building through information extracted from NUSWIDE and Wikipedia. The new list of concepts is an adaptive list which suits in the context of the ambiguous query understanding.

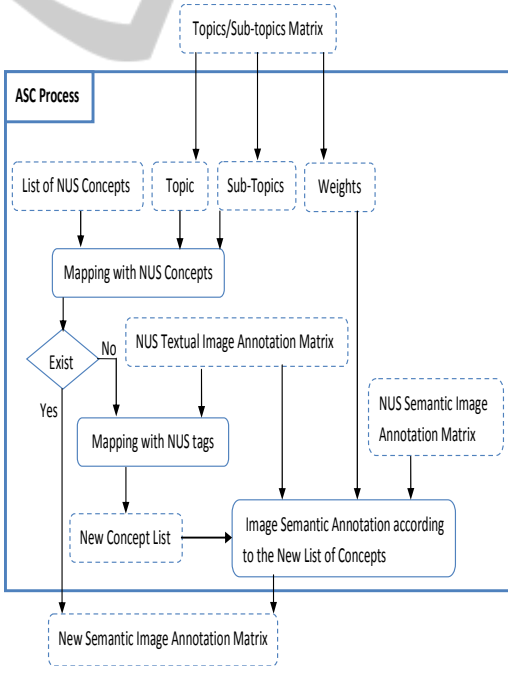

Figure 2: ASC process.

3.4 Diversity-based Retrieval Process

After avoiding the ambiguity from a given query and

indicating the associated concepts for it, the system tries to present diverse images (Figure 3). In order to cover all the meanings which can be mentioned by a given ambiguous query, our system starts by returning images that present intersection between two or more sub-topics. Then, it returns images by alternating between the different lists of images presenting an intersection between the main topic and one from the list of sub-topics.

For example, for an ambiguous query having five sub-topics, the one hundred showed images are returned as follows:

- The first image presents the main topic and an intersection between the first and the forth sub-topics;
- The second image presents the main topic and an intersection between the third and the fifth sub-topics;
- The third image presents the main topic and the first sub-topic;
- The forth image presents the main topic and the second sub-topic;
- The fifth image presents the main topic and the third sub-topic;
- The sixth image presents the main topic and the forth sub-topic;
- The seventh image presents the main topic and the fifth sub-topic;
- The eighth image presents the main topic and the first sub-topic;

...

- The 98th image presents the main topic and the first sub-topic;
- The 99th image presents the main topic and the third sub-topic;

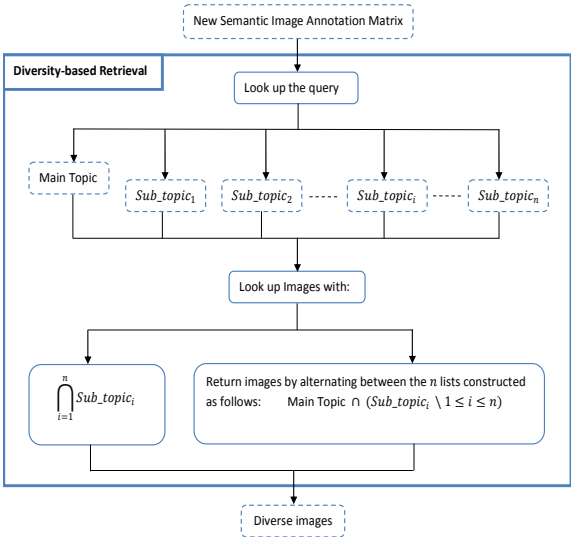

Figure 3: Diversity-based image retrieval.

- The 100th image presents the main topic and the first sub-topic.

We notice that with such strategy, we can achieve an extreme diversity in returned images despites of the disappearance of some sub-topics like the second, the forth and the fifth sub-topics in the example previously mentioned.

Finally, the TAD approach offers to the user a panoramic view describing all the possible intents, which can be mined from his ambiguous query. The achieved result of TAD allows the user to mention easily his need.

4 EXPERIMENTS

Based on a deep literature review and our personal linguistic knowledge, we come up with a list of twelve ambiguous queries.

The Table 1 provides an idea about the meanings detected by our approach TAD for each ambiguous query. We denote by:

- DM: Detected Meanings in showed images for a given query;
- NDM: Number of the Detected Meanings in showed images for a given query;
- DR: Diversity Rate which is the average of the Numbers of the Detected Meanings in showed images for all the queries.

Table 1: Detected meanings for the ambiguous queries.

Queries	**DM**
Apple	Fruit, Inc., cocktail, cake, mythology, tree
Dove	Columbidae
Eagle	Bird, Handbook Birds World, Central America, Claw talons, diffraction, apex predator (dinosaur), Asia (an independent fund of funds management firm), Helm Identification Guides Christopher (book), mule deer, duiker, David Allen Sibley, Pete Dunne (author)
Jaguar	Cat, Cars, Archie Comics, cartoonist, Jagúar band, TV series Banzai!, Fender, Jaguar!, microarchitecture, rocket, software, Mac OS X v 0, Hewlett Packard, Argentina Jaguars, Chiapas F C, IUPUI Jaguars, Jacksonville Jaguars, Racing, South American Jaguars, UHV Jaguars, Armstrong Siddeley, Claas, SEPECAT, Jaguars Mesoamerican culture, Beretta series
Jordan	Country, King, Dead Sea
Pear	Fruit

Table 2: Detected meanings for the ambiguous queries. (cont.)

Washington	George, state, D C, Court House Ohio, Park, Terrace Utah
Cambridge	City status United Kingdom, University Library
Beetle	animal, Carl Linnaeus (botanist, physician, and zoologist),10th edition Systema Naturae (his book), Ancient Greek, Morphology biology (schema), Monophyly (schema), Zoologica Scripta (International Journal of Systematic Zoology)
Fuji	Mount, Harry Fujiwara, Keiko, Sumiko, Shusuke, Xiangcheng City, River, Saga, Shizuoka, Speedway, Mt Jazz Festival, Rock Festival, Mr, Fujifilm, Television, Fujitec, Fujitsu, apple
Jim Clark	United Kingdom British, Team Lotus, Formula One season
Beckham	Surname, County Oklahoma (USA), Bend It Like Beckham (Anglo-German comedy-drama film), rule, Beckingham Palace, Posh and Becks

Based on the popularity of some image search engines (Table 2), we compare our approach (images from NUSWIDE[7]) with Google, Bing and Flickr. The popularity depends on Google PageRank and Alexa Rank[8]. Google PageRank reflects the importance of a giving website by counting the number and quality of links to it. Alexa Rank is how many surfers get in relation to other websites. A giving site aims to gain a high Google PageRank and a low Alexa Rank.

Table 3: Comparison between different search engines.

Image Search Engines	**Alexa Rank**	**Google PageRank**
Google	1	9/10
Bing (Microsoft)	24	6/10
Flickr (Yahoo!)	99	-
Exalead	24 577	6/10
Picsearch	52 102	2/10
Quality Image Search	2 257 653	3/10
Incogna (University of Ottawa, Canada)	3 400 174	4/10

The number of detected meanings reached by our approach is compared with the number of detected meanings attained by Google, Bing and Flickr. For the last image search engine, the evaluation concerns the "pertinent" search.

The Figure 4 shows that the TAD approach achieves best results for the majority of the queries. For some queries like "Jaguar", "Fuji" and "Eagle", the difference is important.

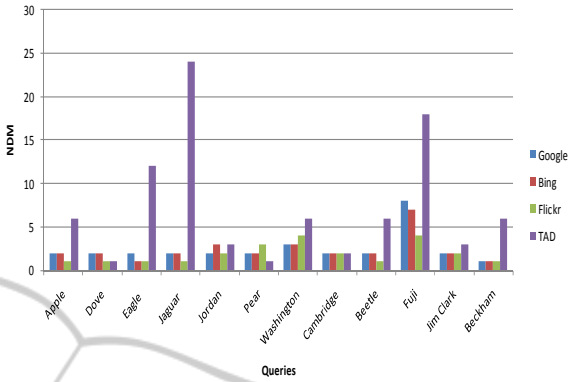

Figure 4: TAD approach versus search engines.

The Table 3 shows that the diversity rates of Google and Bing are almost similar. The diversity rate of TAD approach reflects its ability of covering the majority of the possible meanings.

Although existing search engines detect multiple meanings, the focus remains on one or two well known meanings.

Table 4: Diversity rates comparison.

	Google	**Bing**	**Flickr**	**TAD**
DR	2,5	2,41666667	1,91666667	7,33333333

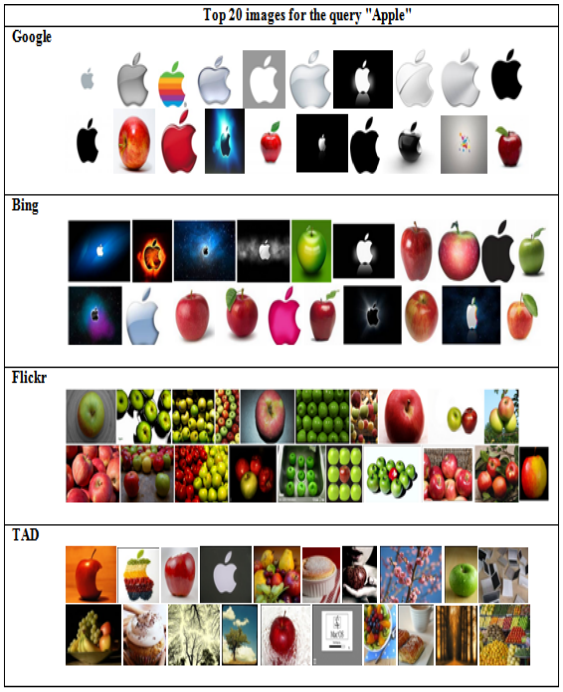

Figure 5: Results for the query "Apple".

[7] http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm
[8] www.seomastering.com/google-alexa-rank-checker.php

The Figure 5 illustrates the top ten images for the query "Apple" which are returned by Google, Bing and Flickr. It shows also the top ten images returned by our approach TAD from the NUSWIDE dataset.

We notice that TAD approach returns for the query "Apple", in the beginning, two images of intersections. Then, it provides one image for each meaning. At the rank nineteen, we detect the disappearance of the meaning tree because there are no more images describing this meaning.

5 CONCLUSIONS

In this paper, we described our first perception of our image retrieval system based on TAD approach. The TAD approach is a puzzle constituted by three main components which are the TAWQU (Thesaurus-Based Ambiguous Web Query Understanding) process, the ASC (Adaptive Semantic Construction) process and the DR (Diversity-based Retrieval) process. Wikipedia pages were a very beneficial source of information for our research for the query disambiguation and the adaptive semantic construction. The NUS-WIDE dataset allowed us to carry out a respectful evaluation. The obtained results were promising for the majority of the twelve ambiguous queries. This work hopes to be a first step toward better dealing with the query ambiguity problem and we are already working on the enhancement of the TAWQU process.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

REFERENCES

Chechik, G., Sharma, V., Shalit, U., Bengio, S., 2010. Large Scale Online Learning of Image Similarity through Ranking. *JMLR*, *Journal of Machine Learning Research*. pp. 1109-1135.

Fakhfakh, R., Feki, G., Ksibi, A., Ben Ammar, A., Ben Amar, C., 2012. REGIMvid at ImageCLEF2012: Concept-based Query Refinement and Relevance-based Ranking Enhancement for Image Retrieval. In *CLEF, Conference and Labs of the Evaluation Forum*. Italy.

Feki, G., Ksibi, A., Ben Ammar, A. , Ben Amar, C., 2013. Improving image search effectiveness by integrating contextual information. In *CBMI*, *11th International Workshop on Content-Based Multimedia Indexing*. 149-154.

Feki, G., Ksibi, A., Ben Ammar, A. , Ben Amar, C., 2012. REGIMvid at ImageCLEF2012: Improving Diversity in Personal Photo Ranking Using Fuzzy Logic. In *CLEF*, *Conference and Labs of the Evaluation Forum*. Italy.

Hoque, E., Hoeber, O., Strong, G., Gong, M., 2013. Combining conceptual query expansion and visual search results exploration for web image retrieval. *Journal of Ambient Intelligence and Humanized Computing*. Volume 4, Issue 3 , pp 389-400.

Hoque, E., Hoeber, O., Gong, M., 2012. Balancing the Trade-Offs Between Diversity and Precision for Web Image Search Using Concept-Based Query Expansio. In *JETWI*, *Journal of Emerging Technologies in Web Intelligence*.Vol. 4, No. 1.

Ksibi, A., Feki, G., Ben Ammar, A. , Ben Amar, C., 2013. Effective Diversification for Ambiguous Queries in Social Image Retrieval. In *CAIP, 15th International Conference Computer Analysis of Images and Patterns*. 571-578.

MacKinlay, A., 2005. Using Diverse Information Sources to Retrieve Samples of Low-Density Languages. In *Proceedings of the Australasian Language Technology Workshop*. Pages 64–70, Sydney, Australia.

Upadhyay, K., Chhajed, G., 2014. Daubchies Wavelet transform and Frei-Chen Edge detector for Intention based Image Search. *International Journal of Computer Science and Engineering*. Volume-2, Issue-5 E-ISSN: 2347-2693.