Combining CBIR and NLP for Multilingual Terminology Alignment and Cross-Language Image Indexing

Diego A. Burgos Herrera
Translation and New Technologies Group
University of Antioquia
Calle 67 No. 53-108 – Bloque 11
burgos.diego@gmail.com

Abstract

In this paper, an overview of an approach for cross-language image indexing and multilingual terminology alignment is presented. Content-Based Image Retrieval (CBIR) is proposed as a means to find similar images in target language documents in the web and natural language processing is used to reduce the search space and find the image index. As the experiments are carried out in specialized domains, a systematic and recursive use of the approach is used to align multilingual terminology by creating repositories of images with their respective cross-language indices.

1 Introduction

Images, as representation of real world entities, constitute a *sine qua non* prerequisite for a number of language tasks. For instance, children as well as foreign language learners often resort to images in order to concretize lexical learning through associative processes (cf. Bloom, 2000: 57).

Likewise, human translators particularly benefit a lot from images when dealing with specialized texts. For example, a word-based image search is a very useful technique to enhance understanding of the source text and achieve precision in the target text. In the context of online resources, a site with the image of a device provides the translator not only with an illustration of the object, but also with hyperlinks to websites containing relevant information.

However, for an integral usage of images as a supportive resource for automated language processes, comprehensive indexed image databases as well as wide-coverage lists of suitable index terms are required. The availability of such lists and the material to index images are language dependent. For instance, for English, considerably more resources are available than for Spanish. A study carried out by Burgos (2006) with bilingual Spanish-English terminological dictionaries revealed that the average of retrieved Spanish documents per term from the web was dramatically lower (7,860) than the average of retrieved English documents (246,575). One explanation to this is the huge size of the web search space for English and the little search space for Spanish. However, another reason is that Spanish terms found in traditional terminological dictionaries could not be of conventional usage among experts and do not represent what is actually contained in the search space. Therefore, more suitable index terms must be looked for.

In the present work, content-based image retrieval (CBIR) is proposed as a means for multilingual terminology retrieval from the web with the purpose of aligning a multilingual glossary and building up an image index. The main goal of this research is to exploit the co-occurrence of images and terms in specialized texts which has been called the bimodal co-occurrence (BC). Experiments have been done so far for English and Span-

Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas, pages 76–83, Los Angeles, California, June 2010.

ish with a few observations in other languages, e.g., Portuguese. Figure 1 shows a forecast of the whole system.

The following section provides references on previous work and suggests that the use of terminology for indexing specialized domain images in a bilingual or multilingual setting has not been discussed in previous literature. Section 3 describes the bimodal co-occurrence (BC) hypothesis with more detail. Section 4 provides an overview of how CBIR supports image indexing and term alignment and includes an outline of the procedure to select candidate indices through concrete / abstract discrimination. Section 5 presents the current appeals and needs of this research and section 6 sketches the future work.

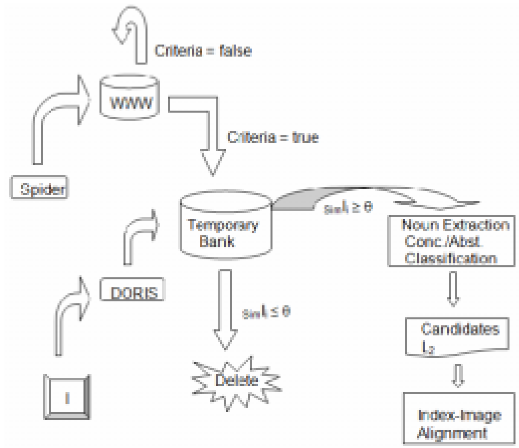


Figure 1. Forecast of the system. A spider is launch to the Internet. Websites fulfilling predefined criteria are temporarily saved and their images analyzed by DORIS. If an image in the website presents feature values within a threshold determined by the example image features, nouns are extracted and classified from the surrounding text to make up a list of candidate target terms which could designate the object in the website's image. Finally, index-image alignment is carried out.

2 Related Research

The particular nature of this research where linguistic and visual representations converge to make up a bimodal co-occurrence which is intended to be exploited for multilingual term retrieval from the web requires the support of diverse specialized knowledge to be applied along the image-based multilingual term retrieval proposed here. As a consequence, the required processes will be framed within or related to the fields and subfields of cross-language information retrieval, cross-language retrieval from image collections, image-term alignment, image annotation and content-based image retrieval.

Many of the latest contributions on the above mentioned fields have been presented in widely known events such as the Text Retrieval Conference (TREC), the Cross-Language Evaluation Forum (CLEF), the Language Resource Evaluation Conference (LREC), the Special Interest Group in Information Retrieval (SIGIR) Conference or the Symposium on String Processing and Information Retrieval (SPIRE), among others.

For work related to cross-language image retrieval which deals with the problem of retrieving images from multilingual collections, see Clough et al. (2006), Clough et al. (2005), Clough (2005), Bansal et al. (2005), Daumke et al. (2006), Izquierdo-Beviá et al. (2005) or Peinado et al. (2005).

Likewise, for standard and alternatives proposals for Content-Based Image Retrieval systems, the reader can check DORIS (Jaramillo and Branch, 2009b), CIRES[1] (Iqbal and Aggarwal, 2003), QBIC[2] (Flickner *et al.*, 1995), PHOTOBOOK[3] (Pentland *et al.*, 1996), IMATCH[4] and VisualSEEk[5] (Smith and Chang, 1996), Nakazato et al. (2003) or Iqbal and Aggarwal (2003). On the other hand, for a detailed description of the CBIR standard technology, see Urcid Pliego (2003), Geradts (2003) or Rui et al. (1999) who present concrete information on the main features for CBIR as well as on some related systems and research. For web-based CBIR related work, see Carson et al. (2002), Yi et al., (2000), Chang et al. (1997), Tollmar et al. (2004) or Drelie et al. (2007). An updated review, compilation of CBIR techniques, real world applications, evaluation techniques and interesting references can be found in Datta et al. (2008).

Content and Text-Based Cross-Language Image Retrieval works can be found in Alvarez et al. (2005), Besançon et al. (2005), Besançon and Mil-

[1] http://amazon.ece.utexas.edu/~qasim/research.htm

[2] http://domino.research.ibm.com/comm/pr.nsf/pages/rsc.qbic.html

[3] http://vismod.media.mit.edu/vismod/demos/photobook/

[4] http://www.photools.com/

[5] http://www.ctr.columbia.edu/~jrsmith/html/pubs/acmmm96/acmfin.html

let (2006), Chang and Chen (2006) or Deselaers et al. (2006).

Image Annotation contributions can be reviewed in Barnard et al. (2003), Cheng et al. (2005), Liu et al. (2006), Qiu et al. (2006), Rahman et al. (2005), Florea et al, (2006), Güld et al. (2006), Petkova and Ballesteros (2005), Müller et al. (2006) or Li and Wang (2003).

Finally, some image-term alignment work has been presented in Burgos and Wanner (2006), Declerck and Alcantara (2006); Li and Wang (2003); Barnard and Forsyth (2001); Pastra (2006) and Wang et al. (2004).

3 BC Hypothesis

The starting point of this proposal is the BC hypothesis which can be defined as follows.

We assume language independent bimodal co-occurrence of images and their index terms in the corpus. This implies that if an image occurs in a document of the corpus, the corresponding index term will also occur in the same document (see Figure 2).

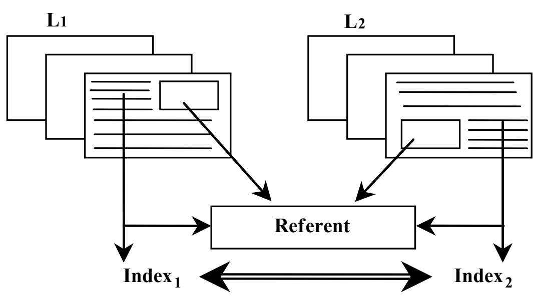


Figure 2. Representation of the BC-hypothesis

Figure 2 also suggests the BC in a bilingual setting. That is, when there is an image of an object in the source language corpus along with its index term there should also be an image of the same object along with its index term in the target language corpus. This means that matching both images would get the two equivalent terms closer. Table 1 shows an example of the bilingual setting of the BC. Both bimodal pairs (image and term) were extracted from manually tracked websites. It is an example of two manually matched images taken from two different language websites which also serve to illustrate how cross-language equivalences between index terms can be established.

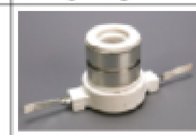
	Source (English)	Target (Spanish)
Image		
Index	Slip-Ring FD 3G 26.9 mm	Colector Ford 3G

Table 1. BC-hypothesis for indexing in a bilingual setting.

In order to prove this BC assumption with some more representative data, a preliminary empirical study (carried out initially for English) was carried out. A sub-corpus of 20 noun phrases[6] designating concrete entities from the automotive industry was extracted from an issue of the Automotive Engineering International Online[7] journal's Tech Briefs section and used to retrieve documents from the web. The first 10 results (i.e., web pages) for each term were stored. Each of the web pages was manually analyzed to check the BC. The result was that the 20 terms confirmed the BC-hypothesis in 145 sites (out of 200) which means a 72.5% of positive cases.

4 CBIR-Based Image indexing

In order to make the most of the BC, it is necessary to automate the process of image matching and image indexing. The fact of matching two images coming from different language documents generates comparable corpora (i.e., topic related) and increases the probability of aligning two equivalent terms by reducing the search space. To do so, we use DORIS, a Domain-ORiented Image Searcher (Jaramillo and Branch, 2009a). DORIS is a JAVA application to retrieve visual information which uses both geometric and Zernike moments based on texture and shape information contained in images. DORIS performance reaches a 90% of precision (Jaramillo and Branch, 2009b).

For the image indexing, we first start from a source language indexed image. An internet segment in the target language is delimited as the search space whose images are compared with the source language image using DORIS. When a

[6] See (Quirk et al., 1985: 247) or (Bosque, 1999: 8-28, 45-51) with respect to the interpretation of the concept 'concrete noun'.

[7] Cf. http://www.sae.org/automag/, state January, 2006.

positive image matching occurs, the target language document containing the matched image is marked as a potential location of the target language index term.

Given that more noise results from a large search space, the size of the image database is usually one of the major concerns in CBIR applications. In our work, we observed that the first problem to tackle is the appropriate definition of the web segment that will constitute the search space. Therefore, scalability and quality issues will be initially addressed by systematically predefining the websites which could contain the image and therefore the target term. In this regard, and as a starting point, the Open Directory Project[8] is used to define our search space. This way, not only categories but also languages can be filtered. For example, the url http://www.dmoz.org/Business/Automotive/ leads to the *automotive* category which contains subcategories and sites in English. On the other hand, following the url http://www.dmoz.org/World/Español/Negocios/Industrias/Automotriz/ which specifies the language, the user finds subcategories and sites of the category *automotriz* for Spanish.

The image database size and quality will depend on this definition. Uniformity is more likely, for example, within the photographs of the same site than between the images of two or more sites. Likewise, there will be greater variance of image characteristics between the images of two different domains than within the images of the same domain, and so on.

Current results were achieved using DORIS. The observations made so far with respect to matching of images on the web suggest that some positive matches in rather homogeneous search spaces provided enough target index term locations to pursue index candidate selection.

4.1 Index Candidate Selection

As it has been suggested, BC can be used for monolingual or bilingual indexing. Once this setting has been decided and the target image has been located as described in the previous section, the index candidate selection can be carried out but, before, it is possible to reduce even more the search space for the index term location by parsing the text surrounding the target image and extracting the noun phrases (NP).

We distinguish NPs from other sort of phrases by means of a chunker. Once all NPs have been extracted, some normalization is done in order to optimize the coming noun classification stage. The cleaning consists of removing determiners at the beginning of the phrase; lemmatization (if appropriate); discarding NPs whose head noun is an acronym[9]; splitting Saxon possessives, and deleting proper nouns and numbers:

three development objectives --> development objective
FSE's single direct injector --> single direct injector

Given the nature of the association, we are focusing, that is image-term alignment, the list of remaining NPs can be additionally pruned by classifying nouns intro concrete and abstract[10].

Classifying nouns as denoting an *abstractum* or a *concretum* is not a trivial task and cannot be widely covered in this paper because of the limited space. It can be said, however, that for noun classification, some approaches have been considered here. For example, remarkable contributions were made particularly by Bullinaria (2008), Katrenko and Adriaans (2008), Peirsman et al. (2008), Shaoul and Westbury (2008), Van de Cruys (2008) and Versley (2008). They use word space and syntactic models which, in some cases, behave very well.

As for the present study experimentation concerning noun classification, three approaches were tested. The number one used non-linguistic variables, the number two was based on syntactic patterns and the number three used lexical semantics information taken from WordNet (Fellbaum, 1998). The automatic semantic annotation was done by the SuperSenseTagger (Ciaramita, 2006). In fact, it is the latter approach the one that yielded the best results with a precision of 88.6% (for detailed information, see Burgos, 2009).

[8] http://dmoz.org/

[9] NPs with acronyms as HN are not included at this stage of the work since often do not reveal whether they designate concrete or abstract entities – which could hinder further validation.

[10] The experiments in this stage so far have been done for English.

	Concrete	Abstract	No annota-tion	No ana-lysis
Concrete	81	14	1	4
Abstract	8	90	0	2

Table 2. Results of noun classification for 100 concrete nouns and 100 abstract nouns. The first two columns/rows show the confusion matrix

These figures suggest that out of 95 concrete nouns, 81 were correctly annotated, and that out of 98 abstract nouns, 90 were annotated with the right sense.

4.2 Index-Image Alignment

With a 90% of precision in image matching and an 88.6% of precision in the noun classification task, we assume a high probability of having the right image with a reduced list of index candidates.

Now, the indexing process can be simplified if the image file name matches any of the candidates. For cases where such matching does not occur, the following procedure is proposed.

For indexing the target image, each candidate is used to query the image database (e.g., Google) for images. For each candidate, the 20 first retrieved images are compared with the target image using DORIS. When a positive image match occurs, the original image is indexed with the candidate that was used to retrieve from the web the image that yielded the positive image match. Table 4 illustrates this procedure by an example. In the example, the images retrieved by *steering wheel* and *air filter* did not match with the original image, but one of the images retrieved by *cylinder head* did. Therefore, the original image is indexed as *cylinder head*.

NP	**Google Images**	**Original image**	**Matching (+/-)**	**New index**
steering wheel	☼ → ☼ →	⇮	– –	–
cylinder head	◙ → ⇮ → OBJ →	⇮	– + –	⇮ cylinder head
air filter	۞ → ۞ →	⇮	– –	–

Table 3. Illustration of the monolingual image-index alignment procedure.

5 Discussion

The approach shows that image indices can be assigned taking into account usage, specificity and geographical variants. The fact of indexing the image with a term retrieved from its context assures that the index term is being used. Moreover, this technique tries to retrieve the appropriate degree of specificity that the index of a specific domain image is expected to present – which is often determined by the number of modifiers of multiword expressions. Likewise, even for specialized discourse, indices should respond to geographical variants. This aspect can be controlled by specifying country domains.

6 Appeals and needs

This work could be incorporated with projects dealing with the access to existing information bases by providing multilingual and multimodal extensions to them. For instance, assistive technology databases (e.g. EASTIN) or patent retrieval engines (cf. Codina et al., 2008) which contain a great deal of visual content.

Content-Based Image Retrieval (CBIR) is an important contribution to multimodal information retrieval. In addition, pairing images with equivalent multilingual terminology has become a matter of interest, particularly in specialized domains. This work could integrate CBIR and natural language processing (NLP) techniques so that images can be used as language independent representations to help in finding documents of textual or ontology descriptions.

Our approach can be especially useful for web users who do not know the structure of the classification system to successfully search or when they do not know the language and special terminology of the information base.

Thus, this work can be integrated to other systems in order to provide cross-lingual retrieval and machine translation for both queries and documents and to enable visualization support for query formulation and document content presentation.

Given the nature of this research's products, they can be included into the scope of multilinguality by combining CBIR and cross-language information retrieval technology. A link to terminological databases can also be established so they can be automatically fed with entries and visual content.

As for this research needs, an adaptation of the SST to Spanish would be really valuable. The SST has already been ported to Italian which represents an interesting experience to take into account.

On the other hand, optimization and integration of the research modules such as a web crawler and an interface for CBIR and noun classification are still pending.

7 Future work

Given that not all process stages of the proposal presented in this paper have been completely integrated and automated, an overall evaluation has not been possible so far. Future work aims at integrating DORIS in modules for index candidate selection and index-image alignment. The goal is to be able to compile multilingual specialized glossaries after systematic and recursive exploration of well delimited web segments and storage of images with their respective cross-language indices. Likewise, some other methods to improve discrimination between concrete and abstract nouns will be researched. The above cited related works in this line have not been tested yet for our proposal, but, for future work, they will be taken into account provided that these models rely on local information and it certainly represents an advantage for this specific task[11]. Even if linguistic specific features are hard to find in both classes of nouns, they are not completely discarded. Finally, further experiments will be carried out with other domains than automotive engineering.

Acknowledgments

This study is part of a wider research work being carried out by the author within the framework of his PhD thesis at the IULA, Universitat Pompeu Fabra, Barcelona, Spain. It was partially supported by a grant from the Government of Catalonia according to resolution UNI/772/2003 of the Departament d'Universitats, Recerca i Societat de la Informació dated March 10th, 2003.

The author is very grateful with the anonymous reviewers of this paper as well as with Leo Wanner and Stefanos Vrochidis for their valuable comments.

[11] From a theoretical and experimental point of view, Altarriba et al. (1999) provide concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. These ratings may be used to further research in areas such as retrieval of abstract and concrete nouns.

References

Altarriba, J.; Bauer, L. M. & Benvenuto, C. (1999), 'Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words', *Behavior Research Methods, Instruments, & Computers* **31**(4), 578-602.

Alvarez, C.; Oumohmed, A. I.; Mignotte, M. & Nie, J.Y. (2005), *Multilingual Information Access for Text, Speech and Images*, Springer Berlin / Heidelberg, Berlin, chapter Toward Cross-Language and Cross-Media Image Retrieval, pp. 676-687.

Bansal, V.; Zhang, C.; Chai, J. Y. & Jin, R. (2005), *Multilingual Information Access for Text, Speech and Images*, Springer Berlin / Heidelberg, Berlin, chapter MSU at ImageCLEF: Cross Language and Interactive Image Retrieval, pp. 805-815.

Barnard, K. & Forsyth, D. (2001), Learning the semantics of words and pictures, *in* 'Proceedings of the International Conference on Computer Vision', pp. 408--415.

Barnard, K.; Duygulu, P.; Forsyth, D.; de Freitas, N.; Blei, D. M. & Jordan, M. I. (2003), 'Matching Words and Pictures', *Journal of Machine Learning Research* **3**, 1107–1135.

Besançon, R. & Millet, C. (2006), Using Text and Image Retrieval Systems: Lic2m Experiments at ImageCLEF 2006, *in* 'Working notes of the CLEF 2006 Workshop'.

Besançon, R.; Hede, P.; Moellic, P.A. & Fluhr, C. (2005), *Multilingual Information Access for Text, Speech and Images*, Springer Berlin / Heidelberg, Berlin, chapter Cross-Media Feedback Strategies: Merging Text and Image Information to Improve Image Retrieval, pp. 709-717.

Bloom, P. (2000), *How Children Learn the Meanings of Words*, MIT Press.

Bosque, I. (1999). El nombre común. In Bosque, I., Demonte, V. (eds) *Gramática descriptiva de la lengua castellana*. Madrid: Espasa Calpe, pp. 3-75.

Bullinaria, J. A. (2008), Semantic Categorization Using Simple Word Co-occurrence Statistics, *in* Baroni Marco; Evert Stefan & Lenci Alessandro, ed.,'ESSLLIWorkshop on Distributional Lexical Semantics'.

Burgos, D. & Wanner, L. (2006), Using CBIR for Multilingual Terminology Glossary Compilation and Cross-Language Image Indexing, *in* 'Proceedings of the Workshop on Language Resources for Content-based Image Retrieval', pp. 5-8.

Burgos, D. (2006). Concept and Usage-Based Approach for Highly Specialized Technical Term Translation. In Gotti, M., Sarcevic, S. (eds) 2006. *Insights into Specialized Translation*. Bern: Peter Lang.

Burgos, D. (2009) “Clasificación de nombres concretos y abstractos para extracción terminológica”. In *La terminología y los usuarios de la información: puntos de encuentro y relaciones necesarias para la transferencia de la información*. 4, 5 and 6 of May, 2009. Medellin, Colombia. ISBN: 978-958-714-251-8 .

Carson, C., Belongie, S., Greenspan, H., Malik, J. (2002). Blobworld: Image Segmentation Using Expectation-Maximisation and its Application to Image Querying. *IEEE Trans. Pattern Analysis and Machine Intelligence* 24(8), pp. 1026-1038.

Chang, S., Smith, J. R., Beigi, M., Benitez, A. (1997). Visual Information Retrieval from Large Distributed Online Repositories. *Communications of the ACM* 40(12). 63-71.

Chang, Y.C. & Chen, H.H. (2006), Approaches of Using a Word-Image Ontology and an Annotated Image Corpus as Intermedia for Cross-Language Image Retrieval, *in* 'Working notes of the CLEF 2006 Workshop'.

Chen, F., Gargi, U., Niles, L., Schutze, H. (1999). Multi-Modal Browsing of Images in Web Documents. *Document Recognition and Retrieval VI, Proceedings of SPIE* 3651, pp. 122-133.

Chen, Y., Wang, J. Krovetz, R. (2003). CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning. *IEEE Transactions on Image Processing*, Vol. 14 (8) pp. 1187-1201.

Cheng, P.C.; Chien, B.C.; Ke, H.R. & Yang, W.P. (2005), NCTU_DBLAB@ImageCLEF 2005: Automatic annotation task, *in* 'Working Notes of the CLEF Workshop 2005'.

Ciaramita, M. & Altun, Y. (2006), Broad-Coverage Sense Disambiguation and Information Extraction with a Supersense Sequence Tagger, *in* 'Proceedings of the Conference on Empirical Methods in Natural Language Processing'.

Clough, P. (2005), *Multilingual Information Access for Text, Speech and Images*, Springer Berlin / Heidelberg, Berlin, chapter Caption and Query Translation for Cross-Language Image Retrieval, pp. 614-625.

Clough, P.; Grubinger, M.; Deselaers, T.; Hanbury, A. & Müller, H. (2006), Overview of the ImageCLEF 2006 photographic retrieval and object annotation tasks, *in* 'Working notes of the CLEF 2006 Workshop'.

Clough, P.; Müller, H. & Sanderson, M. (2005), *Multilingual Information Access for Text, Speech and Images*, Springer Berlin / Heidelberg, Berlin, chapter The CLEF 2004 Cross-Language Image Retrieval Track, pp. 597-613.

Codina, J.; Pianta, E.; Vrochidis, S.; Papadopoulos, S. (2008) ‘Integration of Semantic, Metadata and Image search engines with a text search engine for patent retrieval’, Semantic Search 2008 Workshop, Tenerife, Spain, 2 June.

Datta, R.; Joshi, D.; Li, J. & Wang, J. Z. (2008), 'Image retrieval: Ideas, influences, and trends of the new age', *ACM Comput. Surv.* **40**(2), 1--60.

Daumke, P.; Paetzold, J. & Markó, K. (2006), Morphosaurus in ImageCLEF 2006: The effect of subwords on biomedical IR, *in* 'Working notes of the CLEF 2006 Workshop'.

Declerck, T. & Alcantara, M. (2006), Semantic Analysis of Text Regions Surrounding Images in Web Documents, *in* 'Proceedings of the Workshop on Language Resources for Content-based Image Retrieval', pp. 9-12.

Deselaers, T.; Weyand, T. & Ney, H. (2006), Image Retrieval and Annotation Using Maximum Entropy, *in* 'Working notes of the CLEF 2006 Workshop'.

Fellbaum, C. (1998), *WordNet: An Electronic Lexical Database*, MIT Press, Cambridge.

Gelasca, E. D.; Ghosh, P.; Moxley, E.; Guzman, J. D.; Xu, J.; Bi, Z.; Gauglitz, S.; Rahimi, A. M. & Manjunath, B. S. (2007), 'CORTINA: Searching a 10 Million + Images Database'.

Güld, M. O.; Thies, C.; Fischer, B. & Lehmann, T. M. (2006), Combining global features for content-based retrieval of medical images, *in* 'Working notes of the CLEF 2006 Workshop'.

Iqbal, I. & Aggarwal, J. K. (2003), Feature Integration, Multi-image Queries and Relevance Feedback in Image Retrieval, *in* '6th International Conference on Visual Information Systems (VISUAL 2003)', pp. 467-474.

Izquierdo-Beviá, R.; Tomás, D.; Saiz-Noeda, M. & Vicedo, J. L. (2005), University of Alicante in ImageCLEF2005, *in* 'Working Notes of the CLEF Workshop 2005'.

Jaramillo, G. & Branch, J. (2009), 'Recuperación de Imágenes por Contenido Utilizando Momentos', *Revista Iteckne* **5**(2).

Jaramillo, G. E. & Branch, J. W. (2009), Recuperación Eficiente de Información Visual Utilizando Momentos, *in* 'XXXV Conferencia Latinoamericana de Informática - CLEI 2009'.

Katrenko, S. & Adriaans, P. (2008), Qualia Structures and their Impact on the Concrete Noun Categorization Task, *in* Baroni Marco; Evert Stefan & Lenci Alessandro, ed.,'ESSLLIWorkshop on Distributional Lexical Semantics'.

Li, J. & Wang, J. Z. (2003), 'Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach', *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* **25**(9), 1075-1088.

Liu, J.; Hu, Y.; Li, M. & Ying Ma, W. (2006), Medical Image Annotation and Retrieval Using Visual Features, *in* 'Working notes of the CLEF 2006 Workshop'.

Müller, H.; Gass, T. & Geissbuhler, A. (2006), Performing image classification with a frequency–based information retrieval schema for ImageCLEF 2006, *in* 'Working notes of the CLEF 2006 Workshop'.

Pastra, K. (2006), Image-Language Association: are we looking at the right features?, *in* 'Proceedings of the Workshop on Language Resources for Content-based Image Retrieval', pp. 40-43.

Peinado, V.; López-Ostenero, F. & Gonzalo, J. (2005), UNED at ImageCLEF 2005: Automatically Structured Queries with Named Entities over Metadata, *in* 'Working Notes of the CLEF Workshop 2005'.

Peirsman, Y.; Heylen, K. & Geeraerts, D. (2008), Size Matters: Tight and Loose Context Definitions in English Word Space Models, *in* Baroni Marco; Evert Stefan & Lenci Alessandro, ed.,'ESSLLIWorkshop on Distributional Lexical Semantics'.

Petkova, D. & Ballesteros, L. (2005), Categorizing and Annotating Medical Images by Retrieving Terms Relevant to Visual Features, *in* 'Working Notes of the CLEF Workshop 2005'.

Qiu, B.; Xu, C. & Tian, Q. (2006), Two-stage SVM for Medical Image Annotation, *in* 'Working notes of the CLEF 2006 Workshop'.

Quirk, R., Greenbaum, S., Leech, G. Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Rahman, M. M.; Desai, B. C. & Bhattacharya, P. (2005), Supervised Machine Learning based Medical Image Annotation and Retrieval, *in* 'Working Notes of the CLEF Workshop 2005'.

Routledge English Technical Dictionary. Copenhaguen: Routledge. 1998.

Shaoul, C. & Westbury, C. (2008), Performance of HAL-like word space models on semantic clustering, *in* Baroni Marco; Evert Stefan & Lenci Alessandro, ed.,'ESSLLIWorkshop on Distributional Lexical Semantics'.

Shen H.T., Ooi B.C., Tan K.L. (2000). Giving Meanings to WWW Images. In: *Proceedings of the 8th ACM international conference on multimedia*, 30 October - 3 November 2000, Los Angeles, pp 39-48

Tsai, C. (2003). Stacked Generalisation: a Novel Solution to Bridge the Semantic Gap for Content-Based Image Retrieval. *Online Information Review*, Vol. 27 (6), pp. 442-445.

Van de Cruys, T. (2008), A Comparison of Bag of-Words and Syntax-based Approaches for Word Categorization, *in* Baroni Marco; Evert Stefan & Lenci Alessandro, ed.,'ESSLLIWorkshop on Distributional Lexical Semantics'.

Versley, Y. (2008), Decorrelation and Shallow Semantic Patterns for Distributional Clustering of Nouns and Verbs, *in* Baroni Marco; Evert Stefan & Lenci Alessandro, ed.,'ESSLLIWorkshop on Distributional Lexical Semantics'.

Wang, X. J.; Ma, W.Y. & Li, X. (2004), Data-driven approach for bridging the cognitive gap in image retrieval, *in* 'Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME 2004)', pp. 2231-2234.

Yeh, T., Tollmar, K., Darrell, T. (2004). Searching the Web with Mobile Images for Location Recognition. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR'04), Vol. 2, pp. 76-81.