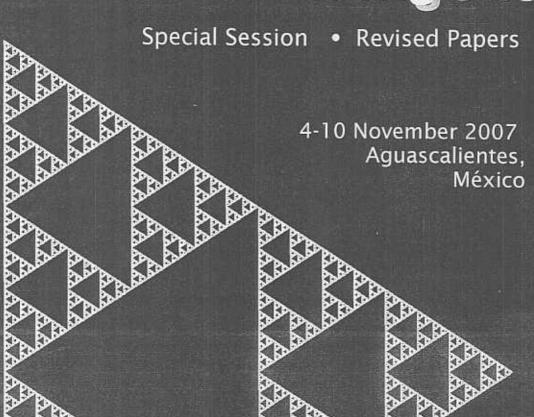
Sixth Mexican International Conference on

Artificial Intelligence



Edited by Alexander Gelbukh Ángel Fernando Kuri Morales

All rights reserved.

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries may photocopy beyond the limits of US copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

IEEE Computer Society Order Number P3124 BMS Part Number CFP0734B-PRT ISBN 978-0-7695-3124-3 Library of Congress Number 2007943618

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
http://computer.org/cspress
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: +81 3 3408 3118
Fax: +81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: <reprints@computer.org>

Editorial production by Patrick Kellenberger Cover art production by Joe Daigle/Studio Productions Printed in the United States of America by Applied Digital Imaging





IEEE Computer Society
Conference Publishing Services (CPS)

http://www.computer.org/cps

Sixth Mexican International Conference on Artificial Intelligence

MICAI 2007

Table of Contents

Kno	wledge Representation and Reasoning
Maintaining Knowledge Bases at the Object Lev J. C. Acosta Guadarrama	el3
An Interface between the Situation Calculus a Pablo Sáez	nd Logic Programming14
A Knowledge Based System Design for the Predi Effects in the Interstellar Environment Iris Iddaly Méndez-Gurrola, Ana Lilia Laure Alfredo J. Santillán-González, and Javier I	26 ano-Cruces,
	Evolutionary Algorithms
Shake - Regicide: A New Heuristic for the Div of Evolutionary Algorithms	ersity Control
of Evolutionary Algorithms	ersity Control
of Evolutionary Algorithms	ersity Control39 Hernandez-Aguirre Machine Learning

Pault Diagnosis in Electrical Power Systems Babilistic Neural Networks
cation of Morphological Feature Extraction and Support lachines in Computerized ECG Interpretation82 Bi Lei, Bing Nan Li, Ming Chui Dong, and Bin Bin Fu
Learning Tools to Time Series Forecasting91 nírez-Amaro and J. C. Chimal-Eguía
Image Processing
N-1
· Mathematical Morphology
ters in Wavelet Domain for Image Processing Applications
atistics Filters in Wavelet Domain for Color Image Processing121 isco Javier Gallegos-Funes, Jesús Martínez-Valdes, osé Manuel De-la-Rosa-Vázquez
exture Histograms for Natural Images Interpretation
um for Affective Pattern Recognition by Means of Use
Initial Momentum
ition and Extraction of Morphologic Features
Riaño-Rojas, F. A. Prieto-Ortiz, L. J. Morantes, nchez-Camperos, and F. Jaramillo-Ayerbe
Natural Language Processing
per Award (Special Session), First Place:
anslation Paraphrases from Trilingual Corpora to Improve Wased Statistical Machine Translation: A Preliminary Report
stically-Based Approach to Detect Causality Relations
tricted Text

Best Paper Award (Special Session), Third Place: Improving Performance of Medical Images Retrieval by Combining Textual and Visual Information
Off-Line Writer Recognition for Farsi Text
Multiagent and Distributed Intelligence
Modelling Intelligent Agents through Causality Theory
Collective Agents and Collective Intentionality Using the EDA Model211 Joaquim B. L. Filipe and Ana L. N. Fred
Uncertain Reasoning in Multi-agent Ontology Mapping on the Semantic Web221 Miklos Nagy, Maria Vargas-Vera, and Enrico Motta
An Intelligent Agent Using a Q-Learning Method to Allocate Replicated Data in a Distributed Database
Optimization Algorithms
A Comparative Study of Three Metaheuristics Applied to the Traveling
Salesman Problem
Incremental Refinement of Solutions for Dynamic Multi Objective
Optimization Problems255 Carlos E. Mariano-Romero and Eduardo F. Morales M.
A Dedicated Genetic Algorithm for Two-Dimensional Non-Guillotine
Strip Packing
Intelligent Contro
intelligent control
PID Controller Optimization Based on the Self-Organization
Genetic Algorithm with Cyclic Mutation
Nonlinear Servo Adaptive Fuzzy Tracking

Hierarchical Fuzzy CMAC Controller with Stable Algorithm for Unknown Nonlinear Systems29 Derto Ortiz, Wen Yu, and Marco Moreno-Armendariz				
nal Neural Network for Nonlinearities Reduction				
, will cludiff cubillos				
Network Model to Control Greenhouse Environment				
.c Digestion Process Identification Using Recurrent				
ba Galvan-Guerra and Ieroham S. Baruch				
Educational Applications				
toaches to Generate Intelligent Teaching-Learning				
Using Artificial Intelligence Techniques				
elevancy Approaches to Improve the Students/ Prediction				
Castro, Francisco Mugica, and Angela Nebot				
ent Tutoring System with Affective Behavior				
er Award (Special Session), Second Place: Reductionist Model for Program Comprehension				
er Environment for Learning about Reactive Systems				
s: A Multi-agent Based Simulation Approach				
Robotics				
spired Method for Friction Estimation				
Evolution of a Neural Network for the Autonomous				
of a Four-Wheeled Robot				
Exploration of Unknown Dynamic Worlds Using Multiple Robots407				
threads Architecture for the Motion Coordination				
erogeneous Multi-robot System				
ndex429				

Improving Performance of Medical Images Retrieval by Combining Textual and Visual Information

M.C. Díaz-Galiano, M.T. Martín-Valdivia, A. Montejo-Ráez, L.A. Ureña-López University of Jaén. Computer Science Department. SINAI Group Campus Las Lagunillas, Ed. A3, E-23071, Jaén, Spain {mcdiaz,maite,amontejo,laurena}@ujaen.es

Abstract

This paper studies the combination of textual and visual information in a database of medical records in order to improve the performance of the multi-modal information retrieval system. The proposed model consists of two subsystems: a Content-Based Information Retrieval subsystem that performs the image retrieval and a textual Information Retrieval subsystem that performs the textual retrieval. The images and text are independently retrieved and then the partial resulting lists are mixed. A study of different weighting schemes has been accomplished and analyzed. The results obtained show that the proper integration of textual information improves conventional multi-modal systems.

1 Introduction

The huge amount of information available electronically, in any format, reveals the need to develop techniques allowing information to be efficiently accessed [3]. Currently, the information available electronically tends to be increasingly multi-modal as it is presented in different formats. Adding image and sound capabilities to computer systems is a great technological step forward from the user's point of view, given that human communication is intrinsically multi-modal (it combines sounds, texts, photographs or even video). However, it would be wrong to think that just by having more information, even multi-modal, access can be efficiently solved. On the contrary, if we do not have systems capable of effective retrieval, the quality of the information available is poor because we will be unable to access it despite its existence.

Image database retrieval has become an important area of interest [5]. An image retriever is an Information Retrieval (IR) system that searches relevant images from a given collection. Content Based Images Retrieval (CBIR) systems index images according to low-level visual features such as color, texture or shape in order to find similar images. In recent years, researchers have been developing CBIR systems. However, despite the progress achieved on this matter, current CBIR systems still have many difficulties to overcome.

In many collections (e.g. historic or stock-photographic archives, medical databases, art collections), images are often associated to some kind of text (e.g. meta-data or captions) semantically related to the image. Retrieval can then be performed either using primitive features based on pixels (using a CBIR system), abstracted textual features associated to the image (using a traditional IR system), or a combination of both. For example, in a

medical domain, it is common to find a clinical case that includes images and the clinical history. In this scenario, traditional text retrieval systems would benefit greatly from a CBIR component. Similarly CBIR systems would take advantage of textual information related to an image [11]. In this paper, we address the problem of how to relate textual and visual information from a multi-modal collection. Our aim is to combine content-based and text-based approaches to multi-modal retrieval in order to obtain better results and overcome the drawbacks of these techniques when they are considered separately.

In our experiments we have used a medical collection with images and multilingual text associated to each image. In particular, we have used the ImageCLEFmed collection [2] which includes images from the MIR, PEIR, PathoPIC and CASImage databases (about 50,000 images). The collection contains textual annotations in XML format.

In order to evaluate the system, we formulate several queries with sample images and a short textual description explaining the research goal. The main goal is, given a query (image and text), to find similar images. This is intended to simulate a medical practitioner needing to find cases similar to the one currently handled in order to assist in professional diagnoses.

The paper is organized as follows. First, we briefly describe the collection used. Then, we discuss the system architecture and the global retrieval process. Section 4 presents the experiments and the results obtained. In the last section, final conclusions and some future research lines are presented.

2 Collection description

For our experiments we have used a multi-modal and multilingual collection. Specifically, we have used the ImageCLEFmed corpora supplied by the Cross Language Evaluation Forum (CLEF)¹ organization for the ImageCLEFmed² task [8]. The CLEF is an international meeting whose purpose is to organize a competition to evaluate different multilingual systems. Furthermore, the ImageCLEF task (the cross-language image retrieval track) includes a multi-modal and multilingual collection and is part of the CLEF campaign.

The ImageCLEF med collection contains about 50,000 images arranged in four data sets: CASImage, MIR, PEIR and PathoPIC. These data sets are very heterogeneous. PEIR and PathoPIC collections contain medical images (scans, x-rays, CT, MRI,) with one textual annotation per image. CASImage and MIR data sets are organized by clinical cases. Each clinical case contains a group of medical images and textual notes in XML format containing information about the illness related to the image.

- The CASImage collection³ contains 8,725 images (scans, x-rays, illustrations, photographs and slides) of 2,076 different cases. Cases are mainly in French with around 20% being in English. Since the queries used are in English, we have translated French annotations into English.
- The MIR data set⁴ (Mallinckrodt Institute of Radiology) contains 1,177 images of 407 cases. It includes nuclear medicine images type. Each case is annotated in English.

¹ http://www.clef-campaign.org/

²http://ir.ohsu.edu/image/

³http://www.casimage.com

⁴http://gamma.wustl.edu/home.html

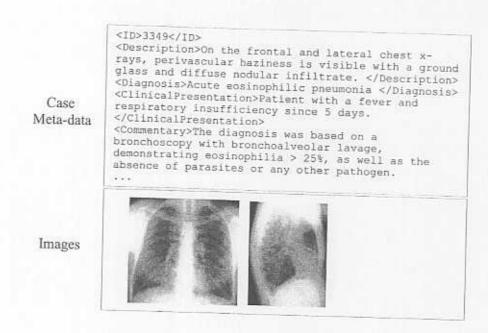


Figure 1: An example of an image and the textual annotations (CASImage data set collection)

- The PEIR database⁵ (Pathology Education Instructional Resource) contains 32,319 images annotated in English. The textual information about these images is very brief but well classified in fields.
- The PathoPIC collection⁶ contains 7,805 pathological images with annotations in two languages: German and English. Since it is a parallel corpus, we only use English annotations.

In a first step, we must pre-process the collection in order to extract the textual information associated to each visual image. The majority of the annotations are in English but a significant number are also in French and German, with a few cases that do not contain any annotation at all. The quality of the text varies among collections and even within the same collection.

We have used English for the document collection as well as for the queries. Annotations in other languages have been translated into English. For French (in CASImage) we have used several Machine Translators (MT) to test the quality of the translation, our final choice being the Reverso Translation Software. For German annotations it was not necessary to translate anything because the information was also available in English.

Figure 2 shows an example of an image and the textual annotations from the CASImage database. For each image, we generate one textual document in order to create the whole textual collection. Note that each case can include more than one image. In this case, we generate more than one textual document per case (one per image) by copying the same text.

Finally, the collections have been pre-processed as usual, applying stop-word removal

⁵http://peir.path.uab.edu

⁶http://alf3.urz.unibas.ch/pathopic/intro.htm

⁷http://www.reverso.net

3 The retrieval process

Given a query composed by one or more images and an associated text, the goal is to find a set of similar images ranked by relevance. For this, we have used two subsystems: a CBIR subsystem that performs the image retrieval and a textual IR subsystem that performs the textual retrieval.

In order to perform the image retrieval process we have selected the Gift software. The Gift (the GNU Image-Finding Tool) is a Content Based Image Retrieval System. This software is the result of a research effort of the Vision Group at the computer science center of the University of Geneva⁹. More information about this system is available in [11].

In order to carry out our textual experiments, we have pre-processed the ImageCLEFmed collection to obtain a textual corpus with a document for every image in the collection. We use the textual Information Retrieval (IR) system LEMUR¹⁰. The LEMUR toolkit is an IR toolkit designed with the idea of language modeling, an attractive new framework for text information retrieval. The toolkit is developed as part of the LEMUR Project, a collaboration between Carnegie Mellon University and the University of Massachusetts Amherst.

We retrieve image and text independently to obtain two partial result lists. Then, the result lists obtained for each subsystem are merged in order to achieve a final single list. Figure 3 shows the architecture designed. The retrieval process can be divided into the following steps:

- 1. The image of the initial query is passed to the CBIR subsystem to obtain a list of relevant images (this list is denoted by L_V -Visual List)
- The text of the initial query is given to the textual IR subsystem to obtain a list of relevant documents (this list is denoted by L_T-Textual List)
- 3. The partial lists are combined using several weighting schemes in order to obtain only one final list with the documents ranked by relevance (L_F-Final List). The experiments are described in the following section.

4 Experiments and Results

For our experiments we have used the queries supplied by the organization of Image-CLEFmed2005. Each query includes one image and a brief textual description of the image. In addition, the CLEF organization also supplied the relevance assessments to evaluate the different system performances. We use the MAP (Mean Average Precision) measure [2] in order to prove the effectiveness of the model proposed.

As we have mentioned above, a list of relevant images is generated for each image by using Gift software. This resulting list is the baseline visual list (L_V) . For each textual query corresponding to each image query, we have generated a list of relevant documents

⁸http://www.gnu.org/software/gift/

http://vision.unige.ch/

¹⁰http://www.lemurproject.org/

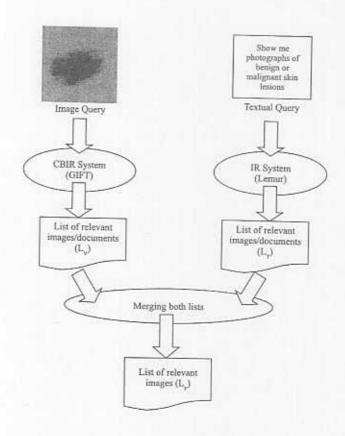


Figure 2: Architecture to merge visual and textual lists

using the Lemur retrieval system. This list represents the baseline textual list (L_T) . These lists contain the most similar cases with regard to the image query, with weighting (the relevance) and ranking values for each image in the list. In order to improve the results, the visual and textual lists are combined in several ways. Thus, we have applied different weighting percentages to each partial list and then we have merged them into one final list re-ranking the new relevance obtained:

$$L_F = L_T * \alpha + L_V * (1 - \alpha) \qquad (1)$$

To verify the expected improvement resulting from the combination of both visual and textual information, we have carried out the following experiments:

- V: The final list obtained using only Visual information. The precision obtained is 0.0941.
- T: The final list obtained using only Textual information. The precision obtained is 0.1614
- TαV: To include both types of information in the final result we have combined
 the textual baseline with the visual baseline, giving different weighting percentage
 to each list. α represents the percentage applied to textual information. We have
 accomplished nine different experiments with α={0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9}

The results obtained are shown in Table 4. The 0 value for α represents the precision obtained with only visual information (results obtained from Gift) and the 1 value for

Table 1: MAP values obtained combining textual and visual information

α	MAP	PIV	PIT
0	0.0941	0%	-72%
0.1	0.1166	19%	-38%
0.2	0.1360	31%	-19%
0.3	0.1572	40%	-3%
0.4	0.1780	47%	9%
0.5	0.1919	51%	16%
0.6	0.1995	53%	19%
0.7	0.1901	51%	15%
0.8	0.1800	48%	10%
0.9	0.1716	45%	6%
1	0.1614	42%	0%

 α represents the precision obtained with only textual information (results obtained from Lemur). The table also includes the percentage of improvement of the proposed model over the visual and textual information retrieval following equations 2 and 3 respectively.

$$PIV = \frac{(MAP - V)}{MAP}$$
(2)

$$PIT = \frac{(MAP - T)}{MAP}$$
(3)

As we can see, these merging strategies improve significantly sthe ystem performance. Those experiments in which the weight given to textual information is greater reported the best results. This is to be expected, since textual retrieval generates better results than visual retrieval with Gift. Nonetheless, those experiments in which the percentage given to the text is between 40% and 90% also manage to surpass the textual baseline. The experiments providing the best results are those in which the contribution of the textual part is not excessively high (50%, 60% and 70%), which empirically demonstrates that combining the two types of results (textual and visual) produces better results than those obtained separately. For these three cases, the improvement percentage over the visual list is higher than 50%. Moreover, the improvement over the textual list is higher than 15%, the best improvement being near 20% (specifically, 19.1% for the T70V experiment). Figure 4 shows graphically the results obtained by the different weighting schemes for textual and visual list.

5 Conclusions and future research

This paper presents a study where visual and textual information are combined in order to improve a multi-modal information system. The results obtained show that the combination of heterogeneous information sources (textual and visual) significantly improves the use of a single source. Although textual retrieval on its own overcomes visual retrieval, when used jointly the results are better than those obtained from independent retrievals.

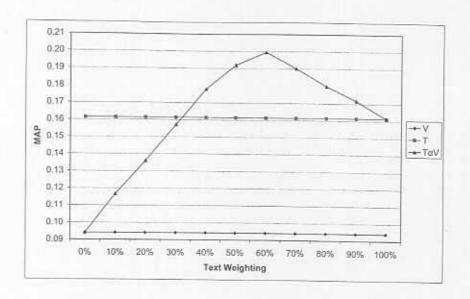


Figure 3: Results obtained with different weighting schemes

Other studies obtain the same conclusions [1][7][10][12]. All of them improve the results of the visual part using differents weighted sum of the scores given by textual and visual retrieval systems.

Future research will attempt to study the incidence of applying some techniques to reduce the textual corpus in order to achieve a better collection (for example, feature selection using information gain). The system proposed will also be applied to other collections containing meta-data information, such as the TRECVid¹¹ collections [9].

6 Acknowledgments

This work has been partially supported by a grant from the Spanish Government, project TIMOM (TIN2006-15265-C06-03). We would like to thank the Cross-Language Evaluation Forum in general, Carol Peters in particular.

References

- Besançon, R., Millet, C.: Using Text and Image Retrieval Systems: Lic2m experiments at ImageCLEF 2006. In Working Notes of the CLEF, LNCS Springer. CLEF, 2006.
- [2] Clough, P., Müller, H., Deselaers, T., Grubinger, M., Lehmann, T. M., Jensen, J., Hersh, W.: The CLEF 2005 Cross-Language Image Retrieval Track, Proceedings of the

¹¹http://www-nlpir.nist.gov/projects/t01v/

- Cross Language Evaluation Forum 2005, Springer Lecture Notes in Computer Science (2006).
- Declerck, T., Kuper, J., Saggion, H., Samiotou, A., Wittenburg, P., Contreras, J.: Contribution of NLP to the Content Indexing of Multimedia Documents. Image and Video Retrieval. Volume 3115/2004. Springer Lecture Notes in Computer Science (2004).
- Fox, E.A., Shaw, J.A.: Combination of multiple searches. Proceedings of the Third Text REtrieval Conference (TREC-1994), pp. 243-252, 1994.
- [5] Lewis, M.S., Sebe, N., Djeraba, C. and Jain, R.: Content-Based Multimedia Information Retrieval: State of the Art and Challenges. ACM Transactions on Multimedia Computing, Communications, and Applications 2 (2006).
- i) Martín-Valdivia, M.T., García-Cumbreras, M.Á., Díaz-Galiano, M.C., Ureña-López, L.A., Montejo-Ráez, A.: SINAI at ImageCLEF 2005. In Proceedings of the Cross Language Evaluation Forum 2005, Springer Lecture Notes in Computer Science (2006).
- McDonald, K., Jones, G.J.F.: Dublin City University at CLEF 2006: Experiments for the ImageCLEF Photo Collection Standard Ad Hoc Task. In Working Notes of the CLEF, LNCS Springer. CLEF, 2006.
- Müller, H., Deselaers, T., Lehmann, T., Clough, P., Hersh, W.: Overview of the ImageCLEFmed 2006 medical retrieval and annotation tasks. Evaluation of Multilingual and Multi-modal Information Retrieval. Workshop of the Cross-Language Evaluation Forum 2006, Springer Lecture Notes in Computer Science (2007).
- Over, P., Ianevay, T., Kraaijz, W., Smeaton, A.F.: TRECVID 2006 An Overview. 2006. Available on-line: http://www-nlpir.nist.gov/projects/tvpubs/tv6.papers/tv6overview.pdf
- [10] Rahman, M.M., Sood, V., Desai, B.C., Bhattacharya, P.: CINDI at ImageCLEF 2006: Image Retrieval and Annotation Tasks for the General Photographic and Medical Image Collections. In Working Notes of the CLEF, LNCS Springer. CLEF, 2006.
- Squire, D.M., Müller, W., Pun., T.: Content-based query of image databases: inspirations from text retrieval. Patter Recognition Letters 21 1193-1198 (2000).
- [2] Wilhelm, T., Eibl, M.: ImageCLEF 2006 Experiments at the Chemnitz Technical University. In Working Notes of the CLEF, LNCS Springer. CLEF, 2006.