Available online at www.sciencedirect.com
ScienceDirect
Information Processing and Management 44 (2008) 1146–1158
INFORMATION PROCESSING & MANAGEMENT
www.elsevier.com/locate/infoproman

Using information gain to improve multi-modal information retrieval systems

M.T. Martín-Valdivia *, M.C. Díaz-Galiano, A. Montejo-Raez, L.A. Ureña-López

Departamento de Informática, Campus Las Lagunillas, s/n., University of Jaén, Jaén E-23071, Spain

Received 18 July 2007; received in revised form 20 September 2007; accepted 25 September 2007
Available online 19 November 2007

Abstract

Nowadays, access to information requires managing multimedia databases effectively, and so, multi-modal retrieval techniques (particularly images retrieval) have become an active research direction. In the past few years, a lot of content-based image retrieval (CBIR) systems have been developed. However, despite the progress achieved in the CBIR, the retrieval accuracy of current systems is still limited and often worse than only textual information retrieval systems. In this paper, we propose to combine content-based and text-based approaches to multi-modal retrieval in order to achieve better results and overcome the lacks of these techniques when they are taken separately. For this purpose, we use a medical collection that includes both images and non-structured text. We retrieve images from a CBIR system and textual information through a traditional information retrieval system. Then, we combine the results obtained from both systems in order to improve the final performance. Furthermore, we use the information gain (IG) measure to reduce and improve the textual information included in multi-modal information retrieval systems. We have carried out several experiments that combine this reduction technique with a visual and textual information merger. The results obtained are highly promising and show the profit obtained when textual information is managed to improve conventional multi-modal systems.

Keywords: Multi-modal information retrieval; Information gain; Data fusion; Medical database access

1. Introduction

Nowadays, a huge amount of information, accumulated in databases and the Internet, is generated every day. All sort of documents are generated due to the rapid expansion of multimedia technology: plain text, images, videos, source code, etc. This amount of documents makes the use of automatic techniques necessary in order to achieve an efficient access to the information. Specifically, information retrieval (IR) techniques are related to the task of retrieving relevant documents from user queries and a large set of documents. Multi-modal retrieval is a subtask of IR where documents are images, videos, or any kind of multimedia

* Corresponding author. Tel.: +34 953 212898; fax: +34 953 212472.
E-mail address: maite@ujaen.es (M.T. Martín-Valdivia).

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2007.09.014

allows us to select the most significant labels within the corpus or, at least, those that provide better information.

This selection system requires no type of external training or knowledge; it simply studies the importance of each label with regard to all the documents. Furthermore, it is independent from the corpus analyzed since in our experiments the IG calculation has been done separately in each sub-collection.

Likewise, it has been proved that using and combining various 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 with visual information, the results are better than those obtained from independent retrievals.

Future research will attempt to study the incidence of applying this technique in systems that require more information, such as, for example, question answering systems. Furthermore, the results obtained will be applied to other collections with meta-data, such as the TRECVid collections.[13]

Besides, we will also explore the integration of external knowledge from a medical ontology in order to enrich the textual collection.

Acknowledgements

This paper 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 and Carol Peters in particular.

References

Buckley, C. & Voorhees, E. M. (2000) Evaluating evaluation measure stability. In *Proceedings of the 23rd annual international ACM SIGIR conference on research and development in information retrieval* (pp. 33–40).

Clough, P., Müller, H., Deselaers, T., Grubinger, M., Lehmann, T., Jensen, J., et al. (2005). The CLEF 2005 cross-language image retrieval track. In *Workshop of the cross-language evaluation forum (CLEF 2005)*. LNCS 2005.

Clough, P., Grubinger, M., Deselaers, T., Hanbury, A., & Müller, H. (2006). Overview of the ImageCLEF 2006 photographic retrieval and object annotation tasks. In *Evaluation of multilingual and multi-modal information retrieval – Seventh workshop of the cross-language evaluation forum (CLEF 2006)*. LNCS 2006.

Declerck, T., Kuper, J., Saggion, H., Samiotou, A., Wittenburg, P., & Contreras, J. (2004). *Contribution of NLP to the content indexing of multimedia documents*. Image and video retrieval. LNCS 2004 (Vol 3115/2004).

Hirata, K., & Kato, T. (1992). Query by visual example: Content based image retrieval. *Advances in database technology – EDBT92*. Berlin: Springer-Verlag.

Lacoste, C., Chevallet, J. P., Lim, J. H., Wei, X., Raccoceanu, D., Hoang, D. L. T., et al. (2006). IPAL knowledge-based medical image retrieval in ImageCLEFmed 2006. In *Workshop of the cross-language evaluation forum (CLEF 2006)*. LNCS 2006.

Lee, W. & Xiang, D. (2001). Information-theoretic measures for anomaly detection (2001). In *Proceedings of the 2001 IEEE symposium on security and privacy*.

Lewis, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications, 2*.

Mitchell, T. (1996). *Machine learning*. McGraw Hill.

Müller, H., Deselaers, T., Lehmann, T., Clough, P., & Hersh, W. (2006). Overview of the ImageCLEFmed 2006 medical retrieval and annotation tasks. Evaluation of multilingual and multi-modal information retrieval. In *Workshop of the cross-language evaluation forum (CLEF 2006)*. LNCS 2006.

Quinlan, J. R. (1986). *Induction of Decision Trees Machine Learning*(1), 81–106.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423, and 623–656.

Squire, D. M., Müller, W., Müller, H., & Pun, T. (2000). Content-based query of image databases: Inspirations from text retrieval. *Pattern Recognition Letters 21* (13–14), 1193–1198 (Selected papers from the 11th scandinavian conference on image analysis SCIA'99).

Yang, Y. & Pedersen, J. O. (1997). A comparative study on feature selection in text categorization. In *Proceedings of ICML-97, 14th international conference on machine learning*.

[13] http://www-nlpir.nist.gov/projects/trecvid/.