Exploiting Complementary Resources for Cross-Discipline Multimedia Indexing and Retrieval

Virginia Fernandez, Krishna Chandramouli, and Ebroul Izquierdo

Multimedia and Vision Research Group,
School of Electronic Engineering and Computer Science,
Queen Mary, University of London, Mile End Road, London, E1 4NS, UK
{firstname.lastname}@elec.qmul.ac.uk

Abstract. In recent times, the exponential growth of multimedia retrieval techniques has stimulated interest in the application of these techniques to other alien disciplines. Addressing the challenges raised by such cross-discpline multimedia retrieval engines, in this paper we present a multi-user framework in which complementary resources are exploited to model visual semantics expressed by users. The cross-discpline areas include history of technology and news archives. In the framework presented the query terms generated by historians are first analysed and the extraction of corresponding complementary resources are used to index the multimedia news archives. The experimental evaluation is presented on three semantic queries namely wind mills, solar energy and tidal energy.

Keywords: Cross-discipline retrieval, multimedia indexing, complementary resource analysis, particle swarm optimisation.

1 Introduction

With the advances in information retrieval (IR) techniques the application of multimedia information retrieval (MIR) tools for alien domains has attracted researchers from various domains. Addressing this challenge, the doman of application considered in this paper is "history of technology'. In this field, the users typically also referred to as historians are required to study the evolution of technology. Typical examples include historians composing a documentary regarding the evolution of windmills from 1960's to 2000's. In particular the content for such reports are searched through news archives. Therefore, creating a critical necessity for semantic indexing of news archives with concepts obtained from modeling the domain of history of technology. In this paper, three concepts are chosen to be "wind mill", "solar energy" and "tidal energy". Understandably the news archives content includes multimodal sources which includes video, textual descriptions and images associated with the depiction of different concepts.

P. Daras and O. Mayora (Eds.): UCMedia 2009, LNICST 40, pp. 109–116, 2010.

Although the prototype developed will process all the information, the analysis of video data is considered out of scope of this paper. Therefore it is assumed that, the video data is analysed to extract keyframes and is only available for further processing.

For semantic indexing of the keyframes, it is critical to have training data to construct visual models for individual concepts. However, due to the limited availability of content, complementary resources are exploited to enable historians construct visual models which are used further for enhancing the performing the retrieval system. In order to build efficient retrieval engines different machine learning techniques have been studied and in particular research in developing new kernel methods for SVM [2] classification and ranking and biologically inspired systems [3] have shown to improve the performance of the retrieval systems.

The remainder of the paper is organised as follows. In Section 2, an overview of the proposed framework is presented, followed by the complementary resource analysis in 3. In Section 4, a brief overview of the visual indexing classifier is presented. Section 5 presents the preliminary experimental evaluation of the framework followed by conclusions and future work in 6.

2 Proposed Framework

In Fig. 1, an overview of the multimedia indexing framework exploiting complementary resource is presented. The framework consists of the user (in our case a historian[1]) who provides a query to the system (for example, wind mill) and the query is processed to extract corresponding complementary resources from online. Thus extracted information is further processed by extracting MPEG - 7 low-level features [4]. On the other hand, the database of news archives is considered to contain video material, which are already pre-processed by shot boundary detection module and key frame extraction module. The complementary resources from the web are used to automatically index the new archives and also, the system implements a Relevance Feedback mechanism through which the user correct the automatically corrected indexing schemes.

In Fig. 2, a multi-user environment is considered in which a set of individual users assume ownership of the individual research databases. The database schemes closely follow the MPEG - 7 content access definitions and in addition also, provides copyright protection for the ownership historians. The system infrastructure is configured in such a way that, it will enable a new query to be searched on the proprietary databases. However, the access to the content is limited only to the title. If the user prefers to have access to the content, then the user can make special request to the individual owners for content. Alternatively, if the owner has assigned the content to be public, then the content is freely available within the network.

[1] The term user and historian are interchangeably used in this paper.

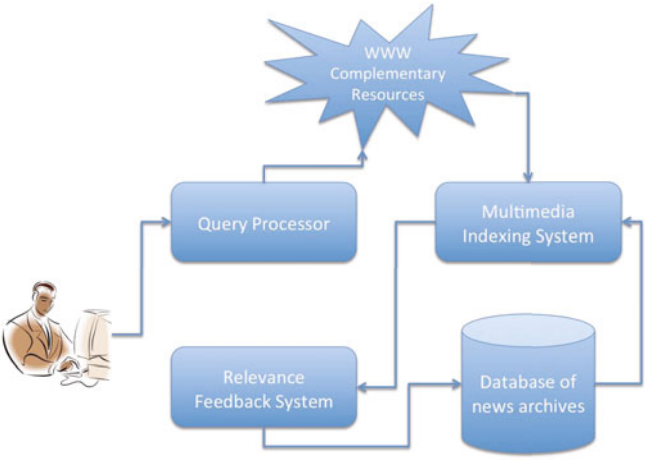

Fig. 1. Complementary Resource Analysis and Multimedia Indexing Framework

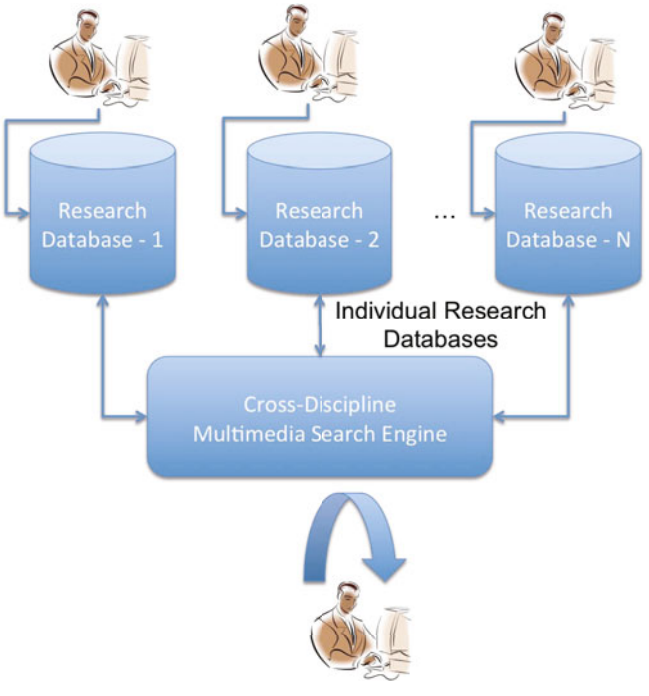

Fig. 2. Multi-User Relevance Feedback Framework from distributed Databases

3 Complementary Resource Analysis

The Flickr[2] website is used as a source of complementary resources in this paper. Flickr is an online photo management and sharing application which allows the user to upload photos, edit photos, organise and share among friends. The popularity of the Flickr could be seen from the 3.2 million items geotagged in one month[3]. Also, Flickr provides webservice wrappers for accessing the photo content directly using an authentication procedure. In order to obtain customised access to the Flickr database a pre-built java webservice wrapper was used with customised functionalities. On an average for extracting 100 images for a given query, the wrapper takes about 10 seconds. Further to the availability of direct access to pictures, the next steps involves the extraction of MPEG - 7 visual features namely Colour Layout Descriptor and Edge Histogram Descriptor. This is achieved by using an open source java library namely Caliph and Emir[4]. As the online resource is bound to get updated, the pictures and therefore the corresponding features are not temporarily stored in any internal databases. In Fig. 3 and Fig. 4, couple of examples for different queries are presented.

Fig. 3. Images downloaded from Flickr for Query: windMill (case sensitive)

4 Particle Swarm Optimisation Based Multimedia Indexing and Retrieval

In the PSO algorithm [5], the birds in a flock are symbolically represented as particles. These particles are considered to be "flying" through the problem space searching for optimal solution [6]. A particle's location in the multidimensional

[2] http://www.flickr.com/
[3] As of 2009-07-14.
[4] http://sourceforge.net/projects/caliph-emir/

Fig. 4. Images downloaded from Flickr for Query: wind mill (case sensitive)

problem space represents one solution for the problem. When a particle moves to a new location, a different solution to the problem is generated. This solution is evaluated by a fitness function that provides a quantitative value of the solution's utility. The velocity and direction of each particle moving along each dimension of the problem space will be altered with each generation of movement. The movement of particles can be influenced in one of two ways. The first is called is the social behavior in which particle gets attracted to the groups center, i.e. following the group, either updating/foregoing the personal best solution. The second is called the cognitive behavior. In this modeling, the particle follows the cognitive experience via personal best solution foregoing the group solution. A more detailed implementation of the RF System is presented in [7].

The PSO based retrieval engine considers as input the MPEG - 7 feature vector extracted from the complementary resources and based on the visual training model provided by the users, the keyframes from the news archives are indexed and retrieved. The ranked list of images with and without the use of the complementary resources are evaluated in the following Section 5.

5 Experimental Results

Flickr Complementary Resource Analysis
The user query search was performed with the open source JAVA api's and for each query the top 100 results are manually annotated for the evaluation. The evaluation is to provide a measure for the performance of Flickr retrieval engine based on textual queries. As discussed previously the results obtained from the Flickr data base are sensitive to query keywords and therefore carefully chosen

query terms are used to retrieve images from the Flickr. In Table 1 performance of the Flickr retrieval engine is presented.

Table 1. Flickr text based retrieval

Query	Retrieval Accuracy (%)
Wind Mill	67
Solar Energy	35
Tidal Energy	44

The results are based on the manual annotation performed on the top 100 documents returned from the Flickr database.

Relevance Feedback Performance measure using Complementary resources

The MPEG - 7 visual descriptors namely Colour Layout Descriptor and Edge Histogram Descriptor are extracted for images from both complementary resources and keyframes. The PSO optimisation implementation includes a combination of cognitive and social behaviour.

In total, 3 different users took part in the user evaluation system and the average performance of the system is presented in Fig. 5 and Fig. 6 with and without the use of complementary resources is presented. Each user interacted with the system As shown in the figures, in both cases PSO based image retrieval provides better results than the SOM and SVM algorithms.

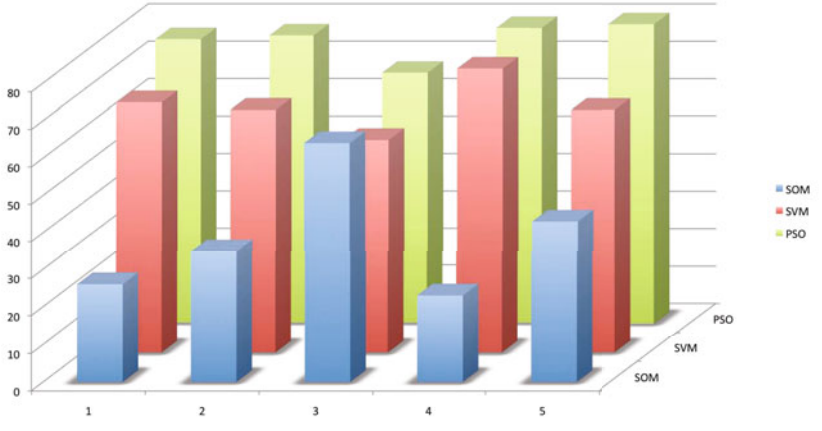

Fig. 5. Average accuracy of relevance feedback results from multiple databases for the queries: Wind Mill, Solar Energy, Tidal Energy without the use of complementary resources

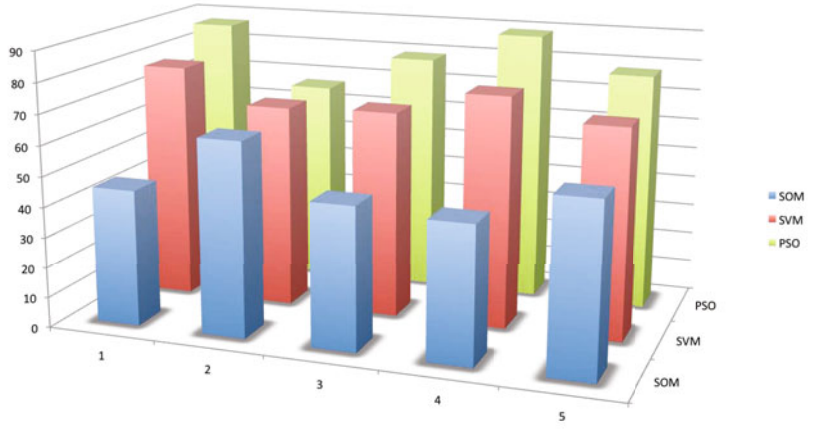

Fig. 6. Average accuracy of relevance feedback results from multiple databases for the queries: Wind Mill, Solar Energy, Tidal Energy with the use of complementary resources

6 Conclusions and Future Work

In this paper, a framework for multimedia indexing is presented using complementary resources. The use of complementary resources are two fold, (i) for building appropriate visual models for the query and (ii) propagating the visual information for indexing keyframes extracted from the video. The evaluation of the relevance feedback algorithm in a multi-user environment is also presented. The experimental results indicate the benefits of creating visual models and propagating the models to index appropriate keyframes. The future work will investigate the possibilities of developing the proposed framework to include P2P network connection. In particular special emphasis will be provided to develop a plugin for Tribler P2P framework. Other possible future work could include the use of SIFT features for image retrieval instead of MPEG - 7 visual features.

Acknowledgement

The research leading to this publication has been partially funded by European Commission under the IST research STREP Papyrus FP7-215874 of the 7th framework programme.

References

1. Goodrum, A.A.: Image information retrieval: An overview of current research. Special Issue of Information Science Research, Information science (2000)
2. Djordjevic, D., Izquierdo, E.: An object- and user- driven system for semantic-based image annotation and retrieval. IEEE Trans. on Circuits and Systems for Video Technology 17(3), 313–323 (2007)

3. Chandramouli, K.: Image classification using self organising feature maps and particle swarm optimisation. In: Doctoral Consortium, Proc. 2nd Int'lWorkshop on Semantic Media Adaptation and Personalization (SMAP 2007), pp. 212–216 (2007)
4. Manjunath, B.S., Ohm, J.R., Vinod, V.V., Yamada, A.: Color and texture descriptors. IEEE Trans. Circuits and Systems for Video Technology, Special Issue on MPEG - 7, 11(6), 703–715 (2001)
5. Eberhart, R., Shi, Y.: Tracking and optimizing dynamic systems with particle swarms. In: Proceedings of the 2001 Congress on Evolutionary Computation, vol. 1 (2001)
6. Reynolds, C.: Flocks, herds and schools: a distributed behavioural model. In: Computer Graphics, pp. 25–34 (1987)
7. Chandramouli, K., Izquierdo, E.: Image Retrieval using Particle Swarm Optimization. CRC Press, Boca Raton (2008) (accepted)
8. Chandramouli, K., Izquierdo, E.: Visual highlight extraction using particle swarm optimisation. In: Latin-American Conference on Networked and Electronic Media (to be published, 2009)