EPIC: A PHOTOGRAPH RETRIEVAL SYSTEM BASED ON EVIDENCE COMBINATION APPROACH

JOEMON M JOSE AND DAVID J HARPER
The Robert Gordon University
Aberdeen, UK, AB25 1HG.
{j.jose,d.harper}@scms.rgu.ac.uk.

Abstract

In this paper, we describe a system which applies an evidence combination approach for the retrieval of photographs from large repositories. To achieve effective retrieval we use features derived from various components (such as text, and image) of the photographic document. As photographic features we use spatial features, which are objects and their locations in the photograph. For the combination of evidence we apply the Dempster-Shafer theory. In this combination scheme, we achieve the integration of database and information retrieval mechanisms into one computational model.

The architecture of the system is based on a client-server architecture. The server is built using FLAIR, a flexible architecture for information retrieval developed on top of an object-oriented database management system. The client is built using the Java programming language. The architecture of the server allows the integration of DB and IR approaches at the implementation level. The usefulness of the spatial features is evaluated using practicing graphic designers and has given very promising results.

INTRODUCTION

A large amount of pictorial information is produced in domains like the newspaper industry, the advertising business and the publishing industry. Existing retrieval tools have yet to catch up with the need to support greater expressive power in photograph description and retrieval tasks [9]. The important research question is to decide what kind of base-level features should be used in enabling the retrieval. Content-based retrieval research has provided some solutions [1, 2]. However, the applicability of these techniques for large scale applications have not yet been proved. Moreover, it has been observed that to achieve effective retrieval, integration of multiple search modalities, (SQL, text, keywords, content-based retrieval) both at the level of algorithms, and the user interface, are needed [4, 5].

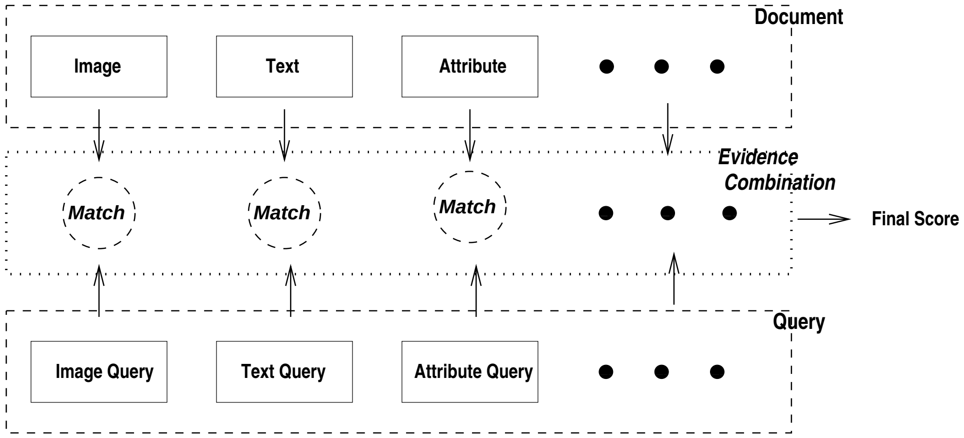

Figure 1: Combining different retrieval mechanisms.

Our approach is depicted in Figure 1. A multimedia collection can be queried in a variety of ways. The idea is to match a query component with a corresponding component of the document (eg. image query with the image component of the document) and arrive at a similarity value. This activity will result in one or more similarity values for the documents in the collection. What we need next is a flexible and efficient method for combining these similarity values to arrive at a final score. This combination method should also have the power to combine exact (DB) and inexact (IR) match retrieval. We propose to use the powerful Dempster-Shafer framework for combining evidence from multiple sources. We have developed a retrieval model based on the Dempster-Shafer approach which satisfies all the above requirements. This retrieval model is applied for photograph retrieval and a prototype system is implemented and described in this paper.

The goal of this paper is to introduce a new retrieval system that can be deployed for photograph retrieval. The outline of the paper is as follows. We describe photograph indexing in the next section. After that we briefly describe the retrieval model used. Following that, we describe the architecture of the system. Then, we discuss the evaluation of the system followed by conclusion.

INDEXING

There are two main sources for deriving index features from a document. The first source is the textual information associated with the photograph, which is indexed by removing stop terms, and stemming all remaining words. The second source is the photograph and we propose to use objects in the photograph and their locations as the indexing features. We refer to these features as the spatial features from a photograph.

Generally, photographs are retrieved by associating a set of keywords with them. Then, this is used for similarity based retrieval. Our idea is, instead of associating a set of keywords with the entire image, to associate it with objects in the photograph. For example, consider a photograph of a car in front of a house. Manually, one can identify the area in which the car or the house appears. Then it can be described using plain text. From this information, the system can generate the location of the object in the image, which can later be used for retrieval. The objects and their location are known as spatial features.

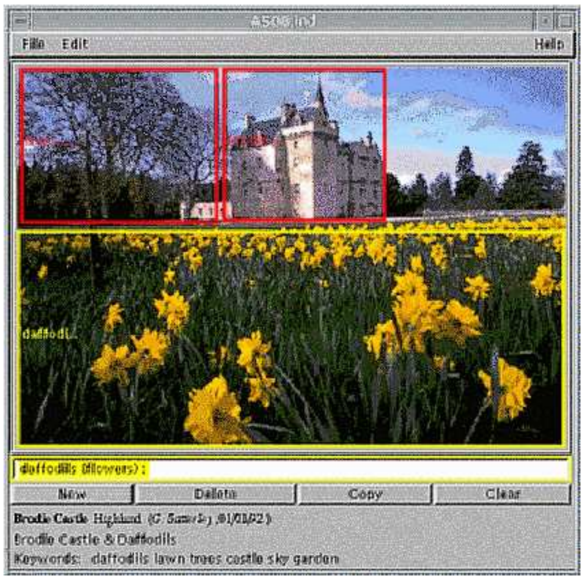

Figure 2: Indexing Interface.

Spatial features are generated manually and computing tools are provided to extract these features with the minimum intervention from the user. Our indexing tool is shown in Fig. 2. To index an image, the indexer identifies objects and specifies the locations of these objects in the image. This is done by drawing a box (rectangle) around the object using the mouse. Then, the indexer describes this object by naming it. From these descriptions, the system generates spatial features. In comparison with the keyword based retrieval approach, the user is performing minimal additional work by identifying the area in which objects appear in an image.

For querying, a query canvas of the size of a photograph is provided [7]. Searchers can draw rectangles on this query canvas and label them. From these, the system automatically derives spatial query features.

RETRIEVAL MODEL

In order to support our generic approach for multimedia retrieval, we need to have a retrieval model which can combine evidence from multiple sources in a flexible and effective way. This combination scheme should support the integration of DB and IR functionalities. We have applied the Dempster-Shafer theory of evidence combination for the development of the retrieval model. The Dempster-Shafer evidence combination method is simply a rule for computing a new belief function based on two or more distinct belief functions [10].

A query may consist of multiple components (e.g. text, spatial query and/or exact match query). Each query component is considered as a source of evidence and a function (probability) representing this source of evidence is computed. The Dempster-Shafer mechanism allows the representation of the uncertainty involved in a particular piece of evidence. Evidence from various query components are combined using the Dempster-Shafer combination scheme. In [6], we have discussed the retrieval model based on a simplified Dempster-Shafer computation scheme and have shown that the behaviour of this computation scheme fits well with our intuition in combining evidence. It is also shown in [6] how this scheme achieves integration of exact and inexact match retrieval in one computational framework.

SYSTEM ARCHITECTURE

A prototype system, called Epic, has been implemented and is based on a client-server architecture and communication between the client and the server takes place using the HTTP protocol. The server system is built using FLAIR, an extensible architecture for building information retrieval applications [8, 3]. The FLAIR architecture is developed to provide a system in which multiple feature extraction and representation techniques can be supported in addition to the support for the integration of DB and IR functionalities. The salient features of the server architecture are: a light-weight data model to support the modelling of complex documents, indexing and representation of multiple features based on an external index specification, flexibility in incorporating new matching criteria.

The client side of the application, namely the user interface, has been built using the JAVA programming language, and a snapshot of this interface is shown in Figure 2. The interface functionality comprises roughly of three categories. Firstly, a facility for constructing and browsing query representations. Secondly, a facility for viewing micro views of the retrieved result sets. Thirdly, a facility for examining specific documents from the result list.

The upper left part of the figure is a multi-modal query interface where a searcher can draw and label query objects. The result viewer is at the bottom left part. A document viewer is provided at the right hand side of the interface. The query interface supports a visual and textual query mechanisms and the result viewer provides a thumb view of the retrieved documents. A searcher can select a particular document from the result viewer and view them in the document viewer.

EVALUATION OF THE SYSTEM

We have tested the conjecture that spatial query interface together with the use of spatial features will help searchers in query specification and hence reduce uncertainty in representing their information need, using professional graphic designers. We used a photographic collection of 800 photographs from the National Trust for Scotland photographic library. The photographs are in general about castles, houses, garden, landscapes etc.

The experimental design is based on a within-subject latin-square design. Eight designers

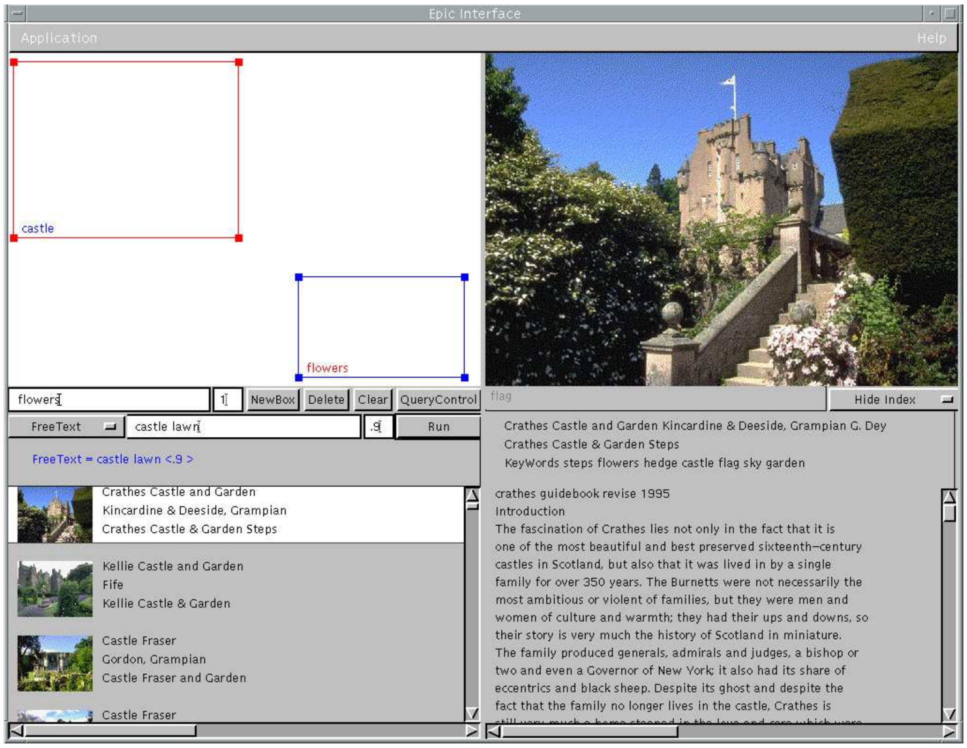

Figure 3: EPIC Interface

were asked to design a leaflet for two different tasks. They were provided with two systems; one with key-word based retrieval and the other only with the spatial interface. This allowed us to compare the usefulness of spatial features in comparison with the key-word based system. To simulate a real work task situation, we have designed real design tasks: design a leaflet for the Scottish Tourist Board. The theme of the leaflet is “Scenic Splendour of the Scottish Countryside”. The collection contains photographs to satisfy this need. Data is collected using questionnaires, think-aloud and automated logging.

The initial analysis has given encouraging results. Out of the eight designers six favoured system and the two preferred the keyword based system. The designers find the *Epic* approach very interesting and are of the view that this influences the way they approach the design task. The two searchers who preferred the keyword based system, were from the same work environment and use keyword based search to browse through their small image collection. Detailed analysis of the experiment is yet to be done.

CONCLUSIONS

The main contributions of our work are the following: we have introduced a new approach to multimedia information retrieval in which we combine multiple sources of evidence; and we have developed a retrieval model based on the Dempster-Shafer theory. This model provides a framework for retrieving information from semi-structured collections and integrates exact and inexact match retrieval. We have introduced a novel approach for picture retrieval which has the potential to provide more precise retrieval. We have also developed a visual query interface which is simple and easy to use for photographic retrieval. A prototype system which is extensible and reusable has been built for photographic information retrieval. The architecture of the system supports the integration of various retrieval modalities at the implementation level. An evaluation of the system has been performed and has given positive results.

References

[1] GUDIVADA, V. N., AND RAGHAVAN, V. V. *IEEE Computer: Special Issue on Content-Based Image Retrieval Systems*, vol. 28 of *9*. IEEE Press, September 1995.

[2] GUPTA, A., AND JAIN, R. Visual information retrieval. *Communications of the ACM 40*, 5 (May 1997), 71–79.

[3] HARPER, D. J., AND WALKER, A. D. M. ECLAIR: an extensible class library for information retrieval. *The Computer Journal 35*, 3 (June 1992), 256–267.

[4] JAIN, R., PENTLAND, A., AND PETKOVIC, D. Workshop report: NSF-ARPA workshop on visual information management systems. Tech. rep., held before ICCV 1995, http://www.virage.com/vim/vimsreport95.html, June 1995.

[5] JOSE, J. M., AND HARPER, D. J. An integrated approach to image retrieval. *The New Review of Document and Text Management 1* (1995), 167–181.

[6] JOSE, J. M., AND HARPER, D. J. A retrieval mechanism for semi-structured photographic collections. In *LNCS 1308 (Proceedings of DEXA 97)* (September 1997), Springer, pp. 276–292.

[7] JOSE, J. M., HARPER, D. J., AND HENDRY, D. G. A spatial feature based photograph retrieval system. In *Proceedings of the Nineteenth Annual International SIGIR Conference on Research and Development in Information Retrieval* (August 1996), H. P. Frei, D. Harman, P. Schauble, and R. Wilkinson, Eds., ACM Press, p. 341.

[8] JOSE, J. M., HENDRY, D. G., AND HARPER, D. J. FLAIR: a flexible architecture for information retrieval. Tech. Rep. TR-97/3, School of Computer and Mathematical Sciences, The Robert Gordon University, Aberdeen, November 1997.

[9] ROMER, D. Image and multimedia retrieval. Discussion note, Research Agenda for Cultural Heritage on Information Networks, The getty art history information program, http://www.ahip.getty.edu/agenda/image.html, 1996.

[10] SHAFER, G. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.