SemView: A Semantic-sensitive Distributed Image Retrieval System*

Wei Wang, Yimin Wu and Aidong Zhang
Department of Computer Science and Engineering
State University of New York at Buffalo
Buffalo, NY 14260 USA

Abstract

State-of-the-art image retrieval techniques have been developed to support high-level (semantics) querying and browsing of images. For certain user queries, it is also very important to know about the information about the remote image databases on the Internet so that the queries can be guided toward the most relevant databases. In this paper, we introduce a novel system, *SemView*, which supports both visual and semantic queries and intelligently ranks the distributed image databases for the queries. Our system summarizes the visual and semantic contents of each database in a metadatabase, which is a collection of automatically generated visual, semantic and statistical metadata. With the metadatabase, our system can effectively guide the user queries toward the most relevant image databases in a distributed environment for better retrieval performance.

1 Introduction

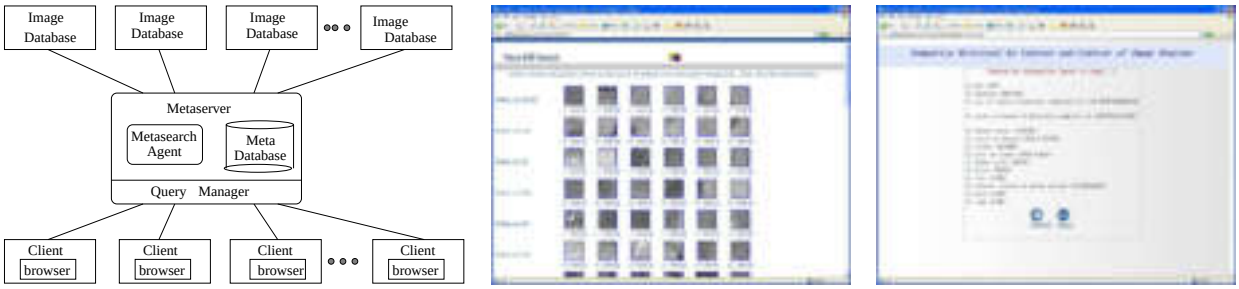
The ultimate goal of image retrieval is to provide the users with the facility to manage large image databases in an automatic, flexible, and efficient way. Therefore, besides current content-based image retrieval (CBIR) techniques which mainly reply on low-level (visual) features, image retrieval systems should also be armed to support high-level (semantic) querying and browsing of images. Although fully automatic semantics extraction for general-domain images is still an open problem due to so-called "semantic gap" between low-level features and high-level semantics, certain methods have been developed to extract the semantics of scenery images (Song *et al*, 2002; Wang *et al* [5], 2002).

Another challenging problem for image retrieval systems is to retrieve the relevant images from remote image databases given user queries. Under such a distributed environment, it is impractical for the system to examine all the images in the databases to match for the relevant images. Instead, based on a particular query, the system should be able to rank the relevant image databases intelligently by the metadata gathered and summarized from respective image databases. The query is then guided to the respective databases, and the relevant images in those databases are returned to users. Webview, a prototype of such a system, has been developed in our previous works (Wang *et al* [4], 2002).

Based on these previous works, *SemView*, a semantic-sensitive distributed image retrieval system, is developed to support both visual and semantics-based retrieval on the distributed image databases. The system includes three major components. The first component is the remote image databases and respective local retrieving servers. The databases, located in the remote host computers connected by the Internet, include geographical and scenery images. The second component is the metaserver, including the metadatabase, the metasearch agent, and the query manager. The third component is the user access applications at the client machines. The components are illustrated in Figure 1(a).

In the following sections, we will focus on the design and implementation of the metaserver, then briefly describe the mechanisms of the local retrieving servers.

*This research is supported by the National Science Foundation Digital Government Grant EIA-9983430.



(a) (b) (c)

Figure 1: (a) System architecture of the *SemView*; Interface of query manager for (b) visual queries, (c) semantic queries

2 Metaserver

The *Metadatabase* of the metaserver houses the *metadata* of both visual and semantic contents of remote image databases. The *Query manager* of the metaserver handles the user queries correspondingly with regard to different types of queries (visual or semantic). When dealing with visual queries, the query manager extracts the visual feature vectors of the queries and match them to the visual part of the metadata (referred as *visual metadata*). If the queries are semantics-based, it matches the queries to the semantic part of the metadata (referred as *semantic metadata*). Figure 1(b) and (c) show the user interface of the query manager. Based on the matching results, the *Metasearch agent* of the metaserver produces a ranked list of the remote databases relevant to the queries, and guides the queries to the selected databases.

We would first briefly introduce how to extract the visual and semantic contents of the images in the databases, then describe the generation of the metadata and the database ranking mechanism.

2.1 Extraction of visual and semantic contents of images

Visual content

Based on different approaches, the visual content of each image in the databases can be represented as multiple visual feature vectors, each corresponding to a different feature class. We use *Wavelet* (Chang *et al*, 1998) and *Keyblock*-based methods (Zhu *et al*, 2000) to extract the visual feature vectors. Refer to (Wang *et al* [4], 2002) for details.

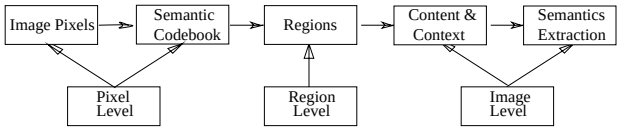


Figure 2: Semantics extraction process

Semantic content

We use the approach proposed in (Wang *et al* [5], 2002) to extract the semantics for the images. The method consists of three levels. At pixel level, a color-texture classification is used to form a semantic codebook based on certain semantic categories and training images. At region level, the semantic codebook is used to segment the images into regions since the basic elements to carry semantic information are the image regions. At image level, the content and context of image regions are defined and represented to support effective semantics extraction from images. The three levels are illustrated in Figure 2. After the extraction process, each region's semantic annotation and its confidence score are stored as the semantics of the image.

2.2 Generation of metadata

As mentioned earlier, the metadatabase contains the *visual metadata* and *semantic metadata* with regard to different aspects for the content summary of remote databases. It also contains the *statistical metadata* associated with both visual and semantic metadata to characterize their similarity distributions with the images in remote databases. We briefly introduce how to generate them as follows:

(1) *Visual metadata*: The visual metadata, which records the summaries of the visual contents for remote databases, is represented by the *image templates*. To find the image templates, we first select sample images from remote databases. By using the hierarchical clustering method, we build a tree-like structure called dendrogram and cut it at different levels resulting in different sets of clusters. We then use the centroids of the resulting clusters as the templates.
(2) *Semantic metadata*: To generate the semantic metadata which records the summaries of the semantic contents for remote databases, we adopt the process similar to the generation of semantic codebook (Wang *et al* [5], 2002). We first define the major semantic categories based on the semantic constraints of the images in the remote databases. For each semantic category, certain number of images are chosen to be the training images such that the semantic metadata generated from them can represent as much as possible the color-texture characteristics of all images in the databases belonging to that semantic category. A color-texture classification is used to classify each pixel of the training images to a certain cell of the color-texture space. For each cell we then count the number of pixels that fall into the cell. The counting results are used to construct a statistical table between cells and semantic categories. In addition, we count the times that two or three different categories present in the same training images. and store the results in a statistics table between semantic categories. In this way, the semantic metadata effectively represent the semantic categories and their relationship, which in turn represents the semantic contents for remote databases.
(3) *Statistical metadata*: The statistical metadata for visual metadata can be computed by the similarity distributions between database images and the image templates. We assume that the similarity of visual feature vectors between a certain image and a template indicates their similarity of visual content. Similarly, the statistical metadata for semantic metadata are computed by the similarity distributions between the semantic contents of the database images and the semantic metadata.

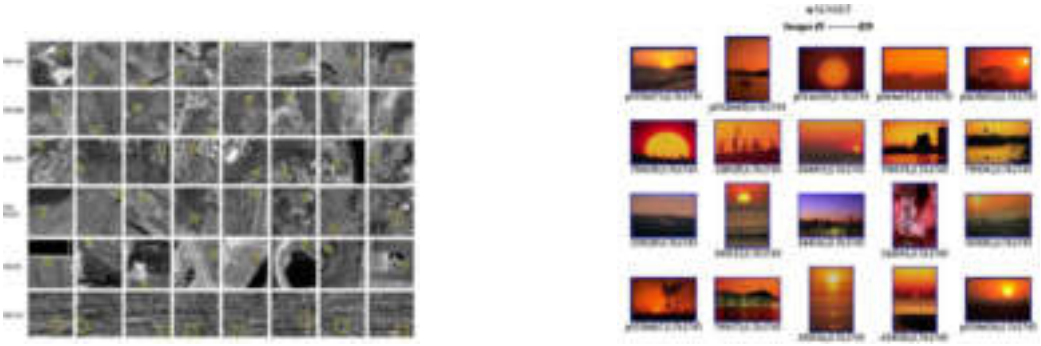
2.3 Remote database ranking mechanism

Based on the metadata, the remote databases can be ranked by the metasearch agent with regard to a particular visual or semantic query. The similarity between the query and the metadata will be calculated to determine the most relevant visual or semantic metadata for the query. Based on the associated statistical metadata, the *Mean-Based approach* (Chang *et al*, 1998) is applied to rank the databases. After the ranked list of the remote databases with regard to a certain query is generated, the metaserver will pose the query to the relevant databases selected by users.

3 Local retrieving servers

Once the query is posed to the remote databases by metaserver, certain relevant images will be retrieved from the remote databases by local retrieving servers. Similar to the metaserver, local retrieving servers would also extract the visual and semantic contents of images for retrieval purpose. Additionally, relevancy feedback mechanism is also applied to improve the performance.

Relevance feedback is a mechanism to interactively learn the user's query preference online. It has been extensively used to improve the performance of image retrieval. We adopt the relevance feedback technique in (Wu *et al*, 2002) to improve the performance of our system. The method evaluates the weights of image features according to their capability to cluster the relevant images together and to discriminate irrelevant images from relevant ones.



(a) (b)

Figure 3: Retrieval results from remote databases for a given (a) visual query, (b) semantic query “sunset”

Figure 3 shows the final retrieval results from the remote databases for certain visual and semantic queries.

4 Conclusion

We have developed the *SemView* to support both visual and semantic queries and retrieve images from distributed image databases over Internet. Based on the queries, the system ranks the remote images databases by the information collected in the metadatabase, then guides the queries to the most relevant databases to effectively retrieve the relevant images back to users.

References

[1] W. Chang, G. Sheikholeslami, J. Wang, and A. Zhang. Data resource selection in distributed visual information systems. *IEEE Transactions on Knowledge and Data Engineering*, 10(6):926–946, November/December 1998.

[2] L. Zhu, A. Zhang, A. Rao and R. Srihari. Keyblock: An approach for content-based image retrieval. In *Proceedings of ACM Multimedia 2000*, pages 157–166, Los Angeles, California, USA, Oct 30 - Nov 3 2000.

[3] Y. Song, W. Wang, and A. Zhang. Automatic annotation and retrieval of images. In *Proc. Sixth IFIP Working Conf. on Vis. Database Sys. pp. 267-280*, 2002.

[4] W. Wang, D. Ma, Y. Wu, A. Zhang, and D. Mark. Webview: A distributed geographical image retrieval system. In *Proc. of 2002 National Conference on Digital Government Research*, 2002.

[5] W Wang, Y Song, and A Zhang. Semantics retrieval by content and context of image regions. In *Proc. of the 15th International Conference on Vision Interface (VI’2002), Calgary, Canada*, pages pp 17–24, May 27-29, 2002.

[6] Yimin Wu and Aidong Zhang. A feature re-weighting approach for relevance feedback in image retrieval. In *Proc. IEEE Int. Conf. on Image Proc.*, 2002.