AN INTERACTIVE IMAGE DATABASE SYSTEM

Jozsef Vass *Jia Yao* *Xinhua Zhuang*

Multimedia Communications and Visualization Laboratory
Department of Computer Engineering & Computer Science
University of Missouri-Columbia, Columbia, MO 65211, USA
E-mail: {vass, jiayao, zhuang}@cecs.missouri.edu

ABSTRACT

The most important features of the developed image database system include compressed domain indexing, searching by using scalable features, and progressive image transmission. User interaction is involved at both the search refinement stage and display of the query results. The indexing and searching algorithms are tightly coupled with the underlying image compression algorithm by means the images are stored in the database. By this scheme, both the complexity and the storage requirements of the database management system are substantially reduced. In this research, we utilize our previously developed very high performance wavelet image coding algorithm termed significance-linked connected component analysis, which not only renders very high compression performance when compared to other top ranked wavelet image coding algorithms and the JPEG standard, but also naturally supports scalable features and progressive transmission. Computer experiments demonstrate the effectiveness of the proposed system.

1. INTRODUCTION

With the decreasing cost and increasing performance of digital image and video capture devices, computing power, and storage capabilities, more and more visual information will be available on-line in large image and video repositories. However, image and video data is much more voluminous than textual data, and visual information cannot be indexed by traditional methods well suited for the indexing of textual information. This makes *effective* and *efficient* visual information search a very difficult task. Here, "effective" means that the user can find exactly what he or she is looking for, and "efficient" stands for the fast response of the system. Thus visual information retrieval has of late attracted significant research interest from the academia as well as industry [1].

The paper aims to design a flexible image database system which uses the highly efficient wavelet-based image compression algorithm termed significance-linked connected component analysis (SLCCA) [2, 3] for both data storage and indexing, and supports user interactivity for both the query by using scalable features and display by using progressive image transmission.

The organization of the paper is as follows. We describe the three main requirements of any visual database system in Section 2. Section 3 reviews the state-of-the-art, and Section 4 presents the proposed image database system. Performance evaluation is addressed in Section 5 and the last section gives conclusions and further research directions.

2. IMAGE DATABASE REQUIREMENTS

To be effective and efficient, we argue that a visual database system should address the following three key issues [4, 5]:

- Efficient search algorithms;
- progressive image transmission; and
- user interactivity.

2.1. Efficient Search Algorithms

Efficient search algorithms are prerequisite of any visual database system due to the large number of images available in the database. In most systems, image features are extracted off-line from the original images, and stored as metadata in the database. In query, the same features are extracted from the query image, and the features of the query image are compared with the features of each target image in the database. To further speed up the search, prefiltering operations can be applied to reduce the number of candidates, and the actual feature matching is only carried out on a subset of target images [6]. In the proposed search engine, *scalable* features are used, i.e., the higher layer feature is enhancement or refinement of the lower layer. Thus the search is done on several *layers* by using scalable features, i.e., at the initial layer only coarse image features are searched which can be done very efficiently. As the search progresses to higher layers, features are refined, or finer features are introduced, and the number of candidate images is gradually decreased. We use both *object-scalable* and *resolution-scalable* features defined later in the paper. In the proposed algorithm, indexing is tightly coupled with the underlying compression algorithm used to store images in the database, and features are directly extracted from the compressed bitstream thereby substantially reducing the storage requirements and complexity of the database management system.

2.2. Progressive Image Transmission

Progressive transmission and display of the search results is becoming a critical issue in the design of a visual database system due to the rapidly increasing number of users on the

Internet which makes bandwidth more and more valuable. In the developed system, the search results are returned in the coarsest resolution, and the user can interactively refine the results. Quite frequently, based on the coarse representation of the image, the user can decide whether to further decode the result, i.e., only images with interest are refined for further evaluation. It is also possible, that the user can examine part of the results while other parts are still downloading and decoding. Most of the search engines lack these capabilities. Usually, results are returned as so-called *thumbnail* images [6, 7], used to specify the original resolution image to be downloaded. In the VideoQ system [8], besides thumbnails, hyperlinks to the original video in different resolution and quality are also provided, which requires the user's knowledge of the available bandwidth and increases server side storage. In the proposed approach, we use SLCCA as the compression scheme which naturally supports progressive transmission, i.e., a higher resolution image makes full use of the previous resolution which allows maximal utilization of the available bandwidth, user's time, and server storage space.

2.3. User Interactivity

User interaction plays a major role for both query and transmission of query results. *Interactive query* is important due to several reasons such as the large number of images available in the image database and the inherently ambiguous nature of image query, i.e., the same description might match several images with no or little interest to the user. Furthermore, a large percent of users may not know what exactly they are looking for, or how to describe the scene. So it is preferable that after a coarse and fast query, the user can preview and narrow down the search by providing feedback based on images returned by the search engine. Interactive search is also closely related to scalable features. Based on a low layer fast and coarse query, the user can refine the search at higher layers. Higher layer search techniques usually give more freedom to the user resulting in more accurate results at the price of higher computational complexity. The main motivation for user *interactivity of transmitting* and displaying search results is the efficient utilization of the available bandwidth as described above.

3. REVIEW OF PREVIOUS WORK

There have been several image and video search algorithms developed in both academia and industry. Histogram-based techniques are far the most popular in image indexing and searching [9, 10]. Their advantages include compact representation of the image and computational efficiency. In [11], by using several moments of the histogram, the authors presented improved results over traditional histogram comparison with reduced storage and computational complexity. Translation and scale invariant moments are proposed in [12]. Jacobs *et al.* [13] introduced a wavelet-based indexing and searching algorithm, where the similarity between two images are measured by comparing the sign of the most significant wavelet coefficients. In Liang *et al.* [14], a wavelet-based image retrieval system is developed, where the translation variance of wavelet transform is combatted by adopting the number of significant coefficients in each subband as the primary feature. In [15], histogram comparison of highpass subbands is proposed by modeling the histogram of each subband by generalized Gaussian distribution and the variance and shape parameters are used as features.

As a summary, we deduce that only some of the proposed algorithms use scalable features [14], progressive display [14], or user interactivity [13], but to our best knowledge none of the existing image database systems integrate scalable features and progressive display with user interactivity by using a unified compression algorithm for both indexing and data storage.

4. PROPOSED IMAGE DATABASE SYSTEM

4.1. System Architecture

The block diagram of the implemented system is shown in Fig. 1. The client is implemented as a Java interface. Requests received from the client are passed to the search engine which executes the search based on the metadata directly extracted from the compressed SLCCA bitstream and stored in the database. The system is made extremely flexible by executing individual search algorithms through common gateway interface (CGI). The state of each client is maintained at the server which is inevitable as a higher layer search is based on the result of a lower one.

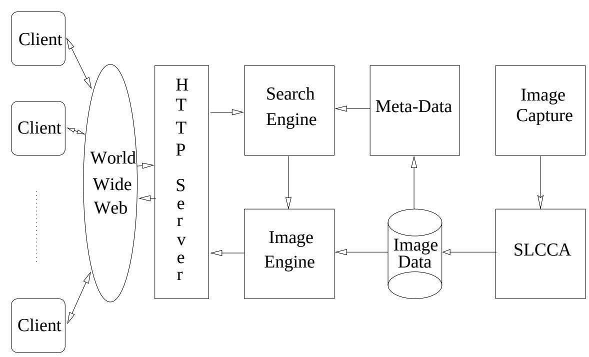

Figure 1: Architecture of the developed image database system.

4.2. SLCCA Image Coding Algorithm

Significance-linked connected component analysis (SLCCA) is a very high performance wavelet image coding algorithm. As opposed to zerotree-based algorithms such as embedded zerotree wavelet (EZW) [16] and set partitioning in hierarchical trees (SPIHT) [17], it organizes significant wavelet data by decomposing them into irregularly shaped connected components or clusters in each subband, a technique called morphological representation of wavelet data (MRWD) [18], and by linking two clusters across subband if one subband contains *parent* and another contains at least one *child*. This exploitation of *cross-scale* dependency among clusters results in the highest peak signal-to-noise ratio (PSNR) in comparison with the other three aforementioned wavelet image coders. SLCCA consistently outperforms JPEG by 3-5 dB in PSNR.

4.3. Features and Implemented Query Types

Most of the image and video search algorithms extensively use the red, green, and blue (RGB) color space [6, 10]. The main drawback of RGB color space lies in that the brightness (luminance) and color (chrominance) information are not treated separately, and all three color components need to be represented at the same resolution, which results in large storages and computational overhead. However, it is well known that most of the image information (about 80%) such as shape and texture are conveyed solely in the luminance component. By using YUV color space, the luminance component is separated from the chrominance components, the later can be downsampled, reducing the storage requirement and accelerating the indexing and searching procedures as well. Since this is the way the image is stored and transmitted, the YUV components are available immediately from the compressed domain with no need of further transformation.

Several global/local and/or scalable/non-scalable features are used for query in the paper. The simplest implemented query is *query by average color*. The average color is extracted from each *original* Y, U, or V-image and stored as metadata in the database. Note that average color is a global, non-scalable feature. *Query by significant coefficients count* [14] and *query by significance map* [13] are also implemented. Both the number of significant coefficients and the significance map, which are global, resolution-scalable features, are provided by SLCCA algorithm, where resolution-scalability means that features can be matched at different scale of the wavelet pyramid. In *query by color layout*, by using a Java drawing tool, users can draw a sketch which consists of possibly several rectangulars each called *object*. The location and extent of each individual object are specified and the color is further layouted over objects; in case of overlapped objects, the color of the top object is used to represent the color of the overlapping part. Apparently, the color layout is a local, object-scalable feature. Object-scalability means that at a lower layer the user may be allowed to specify only few objects to maintain an affordable searching cost, meanwhile, at a higher layer, the user may specify more objects in the sketch which increases the computational complexity for each search. However, as the searching progresses, the number of candidate searches is gradually reduced, thereby keeping the approximate response time constant. In SLCCA, connected component analysis results in an efficient segmentation of wavelet coefficients. Clusters of significant coefficients provide an invaluable information about the structure of the image, i.e., clusters in high frequency subbands correspond to texture and edge regions. This information is readily available from the SLCCA bitstream, and is used for indexing and searching in the following three features. In SLCCA, the relationship between clusters at different scales is represented by the significance-linkage, thus the *number* and *spatial distribution* of significance-links within each subband are used as global, resolution-scalable features. In addition to location, the extent of each cluster provides further insight about the structure of the image, i.e., texture regions from edge regions are likely distinguishable due to the apparent line structure of edges. Thus the last feature is the *clustering* information, where the sign of coefficients within clusters are compared. Clustering is a local, both resolution-scalable and object-scalable feature. Clusters of the query image are compared with clusters of the target images in decreasing size order, i.e., as the search progresses, clusters with smaller sizes are also involved in the comparison with the number of candidates being reduced. Note, that when all the clusters at the full resolution are considered, the algorithm is identical to [13], where the sign of all the wavelet coefficients are used for matching. In *query by example*, the weights of the all the above features can be specified by the user.

5. SYSTEM EVALUATION

The lack of accepted performance metrics of visual retrieval systems makes performance evaluation and comparison of different search engines fairly complicated. This stems from the subjectivity of visual similarity and lack of standard image database that everybody can use. Currently, the developed system contains about 500 images scanned by the authors. The developed image database system is freely accessible on the WWW with client-side Java interface, and can be evaluated at http://meru.cecs.missouri.edu/vdb. The screen captures of the query interface and the progressive display are shown in Figs. 2 and 3, respectively.

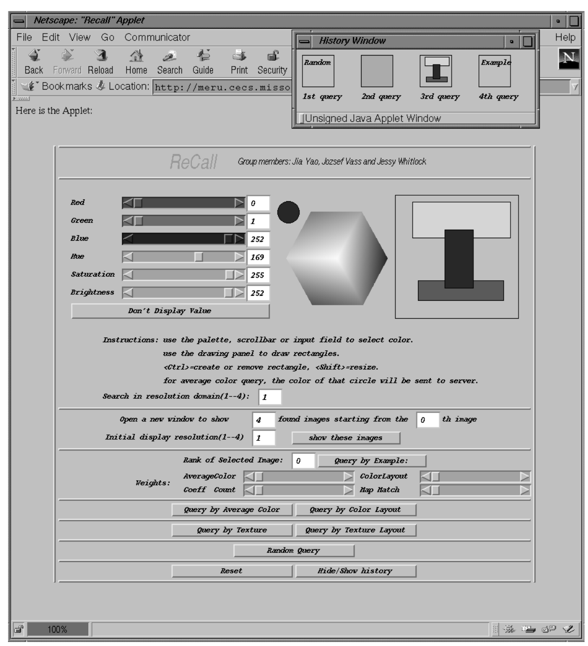

Figure 2: Image query interface.

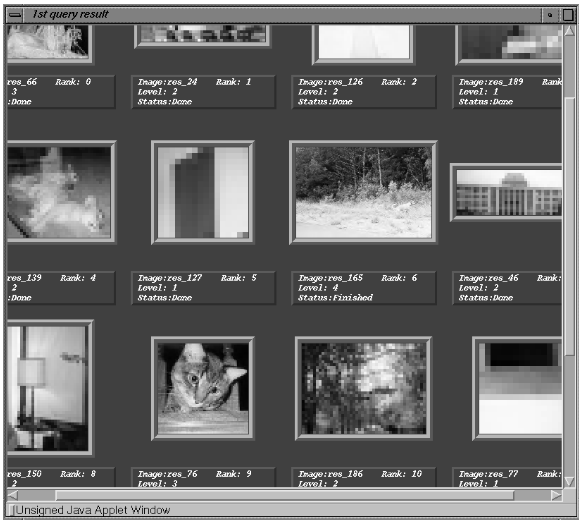

Figure 3: Progressive image display interface.

6. CONCLUSIONS AND FURTHER RESEARCH DIRECTIONS

In the paper, a novel image retrieval system is presented. The proposed image database is tightly coupled with the very high performance SLCCA image compression algorithm that naturally supports scalable features used for both indexing and searching, and progressive transmission. The user is involved for both query refinement and transmission of search results in an interactive manner.

Further research directions include the incorporation of efficient video indexing based on our novel video coding algorithms [19, 20], and by using our automatic spatio-temporal segmentation algorithm [21], the system will be extended to support content-based image and video retrieval as well.

7. REFERENCES

[1] MPEG-7 Requirements Group, "MPEG-7: Context and objectives," July 1997, Doc. ISO/MPEG N1722, MPEG Stockholm Meeting.

[2] B.-B. Chai, J. Vass, and X. Zhuang, "Highly efficient codec based on significance-linked connected component analysis of wavelet coefficients," in *Proceedings of SPIE AeroSense*, 1997.

[3] B.-B. Chai, J. Vass, and X. Zhuang, "Significance-linked connected component analysis for wavelet image coding," *IEEE Transactions on Image Processing, to appear*, 1998.

[4] MPEG-7 Requirements Group, "Third draft MPEG-7 requirements," Oct. 1997, Doc. ISO/MPEG N1921, MPEG Fribourg Meeting.

[5] S.-F. Chang, J.R. Smith, M. Beigi, and A. Benitez, "Visual information retrieval from large distributed on-line repositories," *Communications of ACM*, vol. 40, pp. 63–71, Dec. 1997.

[6] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, and C. Faloutsos, "The QBIC project: Querying images by content using color, texture, and shape," in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, Feb. 1993, vol. 1908.

[7] J.R. Bach, C. Fuller, A. Gupata, A. Hampapur, B. Horowitz, R. Humphrey, R.C. Jain, and C. Shu, "Virage image search engine: An open framework for image management," in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, Feb. 1996, vol. 2670, pp. 76–87.

[8] S.-F. Chang, W. Chen, H. Meng, H. Sundaram, and D. Zhong, "VideoQ: An automated content-based video search system using visual cues," in *Proceedings of ACM Multimedia*, Seattle, WA, Nov. 1997.

[9] M.J. Swain and D.H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[10] A. Berman and L. Shapiro, "Efficient image retrieval with multiple distance measures," in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, 1997, vol. 3022, pp. 12–21.

[11] M. Stricker and M. Orengo, "Similarity of color images," in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, Feb. 1995, vol. 2420, pp. 381–392.

[12] M.K. Mandal, S. Panchanathan, and T. Aboulnasr, "Image indexing using translation and scale-invariant moments and wavelets," in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, 1997, vol. 3022, pp. 380–388.

[13] C.E. Jacobs, A. Finkelstein, and D.H. Salesin, "Fast multiresolution image querying," in *Proceedings of ACM SIGGRAPH*, Aug. 1995, pp. 277–286.

[14] K.-C. Liang, X. Wan, and C.-C.J. Kuo, "Indexing, retrieval and browsing of wavelet compressed imagery data," in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, 1997, vol. 3022, pp. 506–517.

[15] M.K. Mandal, T. Aboulnasr, and S. Panchanathan, "Image indexing using moments and wavelets," *IEEE Transactions on Consumer Electronics*, vol. 42, no. 3, pp. 557–565, 1996.

[16] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[17] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, June 1996.

[18] S. Servetto, K. Ramchandran, and M.T. Orchard, "Wavelet based image coding via morphological prediction of significance," in *Proceedings of IEEE International Conference on Image Processing*, Oct. 1995, pp. 530–533.

[19] J. Vass, B.-B. Chai, and X. Zhuang, "Significance-linked wavelet video coder," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 1998, pp. 2829–2832.

[20] J. Vass, B.-B. Chai, and X. Zhuang, "3DSLCCA – A highly scalable very low bit rate software-only wavelet video codec," in *Proceedings of IEEE Workshop on Multimedia Signal Processing*, Dec. 1998.

[21] J. Vass, K. Palaniappan, and X. Zhuang, "Automatic spatio-temporal video sequence segmentation," in *Proceedings of IEEE International Conference on Image Processing*, Oct. 1998.