Multiresolution Browsing of Pathology Images using Wavelets *

James Ze Wang[†] M.S., John Nguyen[‡] M.S.,
Kin-koi Lo[§] M.S., Christine Law[¶] M.S., Donald Regula[‖] M.D.
Stanford University, Stanford, CA 94305

Digitized pathology images typically have very high resolution, making it difficult to display in their entirety on the computer screen and inefficient to transmit over the network for educational purposes. Progressive zooming of pathology images is desirable despite the availability of inexpensive networking bandwidth. An efficient progressive image resolution refining system for on-line distribution of pathology image using wavelets has been developed and is discussed in this paper. The system is practical for real-world applications, pre-processing and coding each 24-bit image of size 2400×3600 *within 40 seconds on a Pentium II PC. The transmission process is in real-time. Besides its exceptional speed, the algorithm has high flexibility. The server encodes the original pathology images without loss. Based on the image request from a client, the server dynamically generates and sends out the part of the image at the requested scale and quality requirement. The algorithm is expandable for medical image databases such as PACS.*

INTRODUCTION

Medical education is an area that has been revolutionized by the emergence of the World-Wide-Web (Web) and its related technologies. The Web is a rich medium that can increase access to educational materials and can allow new modes of interaction with these materials. Our project utilizes this technology to display images of slide preparations for use in a medical school pathology course. Our goal is to develop a Web-based "virtual microscope" for image viewing that (1) lets students quickly browse and focus on arbitrary regions of the images, (2) minimizes the effort required of course instructors to build medical image libraries, and (3) reduces the network load and storage requirement for the images.

Images captured from slide preparations typically are large and have very high resolution, making it difficult to display them in their entirety on the computer screen. Besides, it is not efficient to transmit the full images over the network for educational purposes. Thus developers of Web-based image libraries have employed a three-step method for displaying large images. First, a small, low-resolution thumbnail is created from the original image. The original image is then divided into regions that are each saved as full-resolution pieces of original. Finally, the thumbnail is hyper-linked to the appropriate pieces of the original image. Clicking on a region of the thumbnail will result in the display of the appropriate region of the original image. This approach is problematic for several reasons: it does not allow students to focus on an arbitrary region of the image, it involves a significant amount of manual effort to build image libraries and it wastes a lot of disk space.

Researchers from University of Maryland [2] have recently developed a virtual microscope system. It is an integrated computer hardware and software system. Using massively parallel computers, the system is shown to be able to serve many users at the same time. However, the algorithm used in the system is not wavelet-based. Therefore, the server is not designed to run on a regular computer such as a Pentium PC or a workstation.

BACKGROUND

Currently, there are many image compression and transmission algorithms available. Interested readers are referred to [3, 4, 9, 10, 11, 18, 14]. There are also commercial software [19] available to compress the entire image using wavelets. Wavelet-based image compression is being considered for JPEG-2000 standard to replace the currently available DCT-based JPEG standard..

Since wavelet transforms decompose images into several resolutions, the coefficients, in their own right, form a localized successive approximation of the original images. Because of this property, wavelet transforms are naturally suited for our virtual microscope progressive image compression algorithms. Many current compression algorithms use wavelet transforms as an initial step [12, 9]. This trend became stronger af-

* *Correspondence to:* wangz@cs.stanford.edu
† Department of Medical Informatics and Department of Computer Science.
‡ Department of Medical Informatics
§ Department of Computer Science.
¶ Department of Electrical Engineering.
‖ Department of Pathology.

ter Shapiro's [12] invention of the zerotree structure for wavelet coefficients. Much subsequent research has taken place based on the zerotree idea, including a very significant improvement made by Said and Pearlman [9], which is referred to as the S & P algorithm. This algorithm was applied to a large database of mammograms [1, 7] and was shown to be highly efficient using real clinical-quality evaluations.

Most wavelet-based progressive transmission algorithms provide successively better approximations to the entire input image as the digital data arrives from the network. However, they are not designed to handle the problem of transmitting a small portion of the image at a user-specified resolution on-demand. For our virtual microscope project, we have designed a new wavelet-based compression algorithm that is suitable to the needs of the pathology image viewing. The algorithm compresses the source image using wavelets. When the image query comes in, the server searches for required information to reconstruct the image at the user specified location and resolution with minimum loss of quality. Then the constructed image is converted to JPEG and transmitted via the Web.

METHODS

The primary goals of our design and implementation of the Web-based virtual microscope system for viewing pathology images are:

1. The system should allow students to browse and expand an arbitrary region of an image at a scale specified by the student. Thus the mode of interaction simulates a standard light microscope.

2. The developers of image libraries should only need to acquire the original high-resolution image. The manipulation of these images is performed by the computer automatically.

3. The system uses wavelet-based compression schemes to minimize the requirements for storage space.

4. The system should be a Web-based client-server model.

Our virtual microscope is a Web-based server-client system. Figure 1 shows the basic structure of the system.

Wavelets

Wavelets are basis functions that have some similarities to both splines and Fourier series [5, 6]. They have advantages when the aperiodic signal contains many

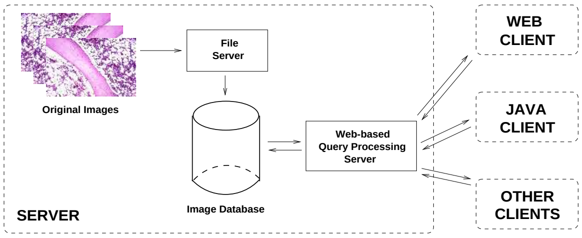

Figure 1: **Basic structure of the system.**

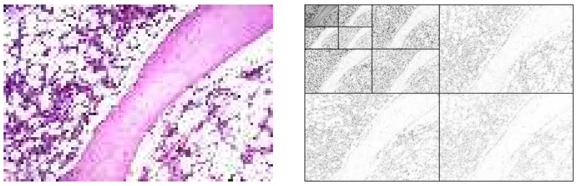

Figure 2: **Wavelet transform of a pathology image.**

discontinuities or sharp changes. It decompose signals into different frequency components and analyze each component with a resolution matching its scale. Applications of wavelets [13, 15, 16, 17] to signal denoising, image compression [8], image smoothing, fractal analysis and turbulence characterization are active research topics. Figure 2 shows the wavelet transform of a pathology image. For the purpose of this project, we selected Haar basis as our wavelet filter because (1) its low filter length results in less boundary effect (2) its wavelet coefficients are scalable to integer which makes lossless compression possible (3) faster computation.

File Server

We have developed an efficient way of storing and transmitting 2D pathology images for the purpose of virtual microscope using wavelet transforms. The file server of the server process, illustrated in Figure 3, performs only once for each image inserted in the database.

We first convert each color image inserted in the database into the RGB color space. Then we convert the RGB image into three gray scale images, each with the red component, the green component or the blue component of the original color image. For the purpose of wavelet transform, we need to pad the border of the image so that the size of the image is suitable to the level of wavelet transform we apply to it.

We apply a 4-level fast wavelet transform (FWT) to each of the three gray images. The wavelet coefficients of each gray image is stored in different files losslessly based on frequency levels for the lowest three levels. Since the last level, i.e. the highest frequency band, is too large and is inefficient for fast retrieval of coeffi-

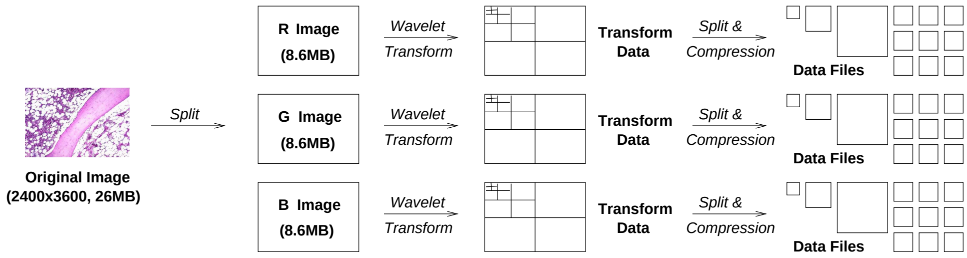

Figure 3: **Structure of the file server.**

cients, we quantize the coefficients and break the coefficients into nine (i.e. 3×3) blocks based on location and store each block of coefficients in a separate file. Huffman coding is used for the lossless compression of coefficients.

Web-based Query Processing Server

The Web-based query processing server is illustrated in Figure 4. Given a query with x-location, y-location and *scale* variables, the Web-based query processing server searches the stored representation of the image, decompresses the data file required, finds necessary coefficients to represent the block of pixels in the specified scale. An inverse wavelet transform is performed to reconstruct the image queried by the user with a maximum error of about 4, out of a 256 scale for each of the three color components of a 24-bit color image.

In order to transmit the image to the Web client, namely the Web browser or a JAVA applet, we convert the constructed image to a JPEG image of very high fidelity to the original. The JPEG image, along with necessary locational information, is then shipped to the client.

Client Program

We have implemented a Web-based user interface that allows the users to magnify any portion of the pathological images in different level of resolutions. We use Web interface and JAVA primarily because the wide acceptance of the Internet and the Web in health care environments.

With the JAVA client program, user is able to draw a rectangle area of interest using the mouse cursor. The client program then composes the corresponding query (i.e. image ID, x1,y1,x2,y2,scale) and send the query to the server to ask for the image for the region at a specified scale. The user is also able to move the current window to nearby window of the same size by simply a click of the mouse button.

We have also created a simple HTML-based client using Web forms. Some people may prefer this client program because of its universal acceptance by browsers. Figure 5 shows sample image query results with the HTML-based client user interface.

RESULTS

The server program is implemented in C on a Pentium II workstation running LINUX. It takes approximately 40 seconds of CPU time for the file server to process a 24-bit color image of 2400×3600 pixels. The coefficients of highest resolution is quantized so that maximum absolute error is 2, out of a scale of 255. After the JPEG compression, we expect the error to get slightly higher. The losslessly compressed files of one such image typically takes 15 MB to 18 MB of disk space, compared to 26 MB for the source image. Huffman coding is used for the lossless compression part.

The query processing is very fast. It takes an average of about one second CPU time to retrieve the image queried by the user, including the time to convert the constructed image to JPEG format. The JPEG compression ratio is approximately 5:1. This ratio can be adjusted by the client for different purposes.

CONCLUSIONS AND FUTURE WORK

In this paper, we have demonstrated an efficient progressive image resolution refining system for online distribution of pathology image using wavelets. The system is practical for real-world applications, pre-processing and coding each 24-bit image of size 2400×3600 within 40 seconds on a Pentium II PC. The transmission process is in real-time. Besides its exceptional speed, the algorithm has high flexibility. The server codes the original pathology images without loss. Based on the image request from a client, the server dynamically generates and sends out the part of the image at the requested scale and quality requirement. The algorithm is expandable for medical image databases such as PACS.

It is possible to improve the coding efficiency of the algorithm by further refining the file server and the query processing server. A more complete user study

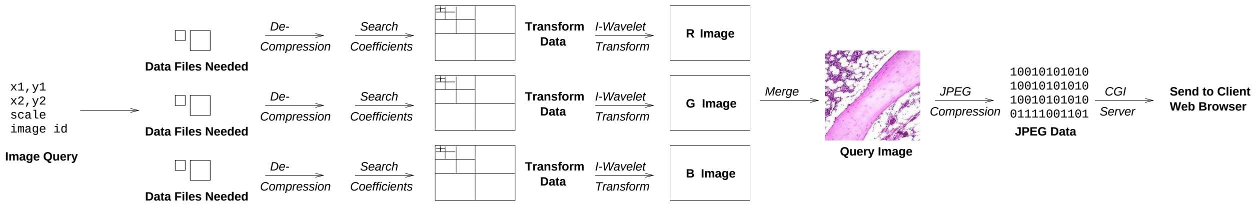

Figure 4: **Structure of the Web-based query processing server.**

Figure 5: **Sample image query results.** HTML-based interface shown.

is also important to make the system more effective for the medical community.

Acknowledgments

We would like to acknowledge the valuable comments and suggestions from Russ Altman, Larry Fagan, Gio Wiederhold and Amit Singh of Stanford Medical Informatics.

References

[1] C. N. Adams et al., Evaluating Quality and Utility of Digital Mammograms and Lossy Compressed Digital Mammograms, *Proceedings of the Third International Workshop on Digital Mammography*, Elsevier, pp. 169-176, 1996.

[2] A Afework, MD Beynon, F Bustamante, S Cho, A Demarzo, R Ferreira, R Miller, M Silberman, J Saltz, A Sussman, H Tsang, Digital dynamic telepathology–the virtual microscope, *Proc AMIA Symp*, pp. 912-6, 1998.

[3] C. Chrysafis and A. Ortega, Efficient context-based entropy coding for lossy wavelet image compression, *Data Compression Conference*, IEEE Computer Society Press, Snowbird, UT, March 25 - 27, 1997.

[4] M. Effros, Zerotree design for image compression: toward weighted universal zerotree coding, *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, November 1997.

[5] Robert M. Gray and Joseph W. Goodman, *Fourier Transforms : An Introduction for Engineers*, Kluwer Academic Publishers, 1995.

[6] Yves Meyer, *Wavelets: Algorithms & Applications*, SIAM, Philadelphia, 1993.

[7] S. M. Perlmutter et al., Image Quality in Lossy Compressed Digital Mammograms, *Signal Processing, Special Issue on Medical Image Compression*, Elsevier, pp. 189-210, January 1997.

[8] A. Said, W.A. Pearlman, An Image Multiresolution Representation for Lossless and Lossy Image Compression, *IEEE Transactions on Image Processing*, Sept. 1996.

[9] A Said and WA Pearlman, A new fast and efficient image codec based on set partitioning in hierachical trees, *IEEE Trans. on Circuits and Systems for Video technology*, vol. 6, no. 3, pp. 243-250, June 1996.

[10] A Said and WA Pearlman, An Image Multiresolution Representation for Lossless and Lossy Image Compression, *IEEE Transactions on Image Processing*, vol. 5, pp. 1303-1310, Sept. 1996.

[11] A Manduca and A Said, Wavelet compression of medical images with set partitioning in hierarchical trees, *SPIE Symposium on Medical Imaging*, Cambridge, MA, March 1996.

[12] Jerome M. Shapiro, Embedded Image Coding Using Zerotrees of Wavelet Coefficients, *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3445-3462, December 1993.

[13] Special Issue on Wavelets and Signal Processing, *IEEE Trans. Signal Processing*, Vol.41, Dec. 1993.

[14] James Ze Wang, Gio Wiederhold, Jia Li Wavelet-based Progressive Transmission and Security Filtering for Medical Image Distribution, *Advances in Biomedical Image Databases*, S. Wong (Ed.), Kluwer, 1999.

[15] James Ze Wang, Gio Wiederhold, Oscar Firschein, Sha Xin Wei, Content-based Image Indexing and Searching Using Daubechies' Wavelets, *International Journal of Digital Libraries(IJODL)*, 1(4):311-328, Springer-Verlag, 1998.

[16] James Ze Wang, Jia Li, Gio Wiederhold, Oscar Firschein, System for Classifying Objectionable Images, *Computer Communications Journal*, 21(15):1355-1360, Elsevier Science, 1998.

[17] James Ze Wang, Jia Li, Robert M. Gray, Gio Wiederhold, Unsupervised Multiresolution Segmentation for Images with Low Depth of Field, submitted for journal publication, 1999.

[18] STC Wong, L Zaremba, D Gooden, HK Huang, Radiologic image compression, *Special issue of Proc. IEEE on Image and Video Compression*, 83(2), 1995:194-219.

[19] Iterated Systems Inc. `http://www.iterated.com`.