12 WAVELET-BASED PROGRESSIVE TRANSMISSION AND SECURITY FILTERING FOR MEDICAL IMAGE DISTRIBUTION

James Ze Wang*, Gio Wiederhold†, Jia Li‡

Abstract: With the rapid expansion of computer networks, communication and storage of medical information has entered a new era. Teleclinical practice and computer digital storage are two important medical information technologies that have made the issue of data compression of crucial importance. Efficiently compressing data is the key to making teleclinical practice feasible, since the bandwidth provided by computer media is too limited for the huge amount of medical data that must be transmitted. Because of the high compressibility of the medical images, data compression is also desirable for digital storage despite the availability of inexpensive hardware for mass storage. This chapter addresses the family of progressive image compression algorithms. The progressive property is preferred because it allows users to gradually recover images from low to high quality images and to stop at any desired bit rate, including lossless recovery. A progressive transmission algorithm with automatic security filtering features for on-line medical image distribution using Daubechies' wavelets has been developed and is discussed in this chapter. The system is practical for real-world applications, processing and coding each 12-bit image of size 512×512 within 2 seconds on a Pentium Pro. Besides its exceptional speed, the security filter has demonstrated a remarkable accuracy in detecting sensitive textual information within current or digitized previous medical images.

12.1 INTRODUCTION

Health care is exceptionally information intensive, and the United States spends hundreds of billions of dollars each year in processing and managing such information [16]. However, it is becoming increasingly difficult to maintain and retrieve health care information manually as more and more advanced medical

*Medical Information Sciences Program, Stanford University, Stanford, CA 94305
†Department of Computer Science, Stanford University, Stanford, CA 94305
‡Department of Electrical Engineering, Stanford University, Stanford, CA 94305

equipment is used in diagnosis and management of disease. Besides the traditional textual data such as patient reports, health care records are being filled with X-ray images, MRI scans, CT scans, 3-D volume reconstructions, and video streams. Efficient lossless medical image compression and progressive transmission is becoming increasingly important.

As the demand for greater accessibility to health care information grows, medical institutions are being urged to make information available to legitimate external parties in a timely fashion (*e.g.*, on-line) while protecting the privacy of patient data [2]. It is therefore crucial that health care institutions be provided with on-line tools that allow them to disseminate medical information without compromising data privacy [22, 3]. In this chapter, we present a wavelet-based lossless medical image compression and progressive transmission algorithm that also detects textual information (including identifying information) from some current or digitized previous medical images. Textual terms not known to be innocuous are eliminated [35]. The resulting processed images can then be made available to medical researchers, second-opinion physicians, students, and other legitimate users after being processed by our algorithm. Such a tool could be used by health care institutions and other repositories of medical images as part of a real-time medical image distribution system.

12.1.1 Related Work

Image Compression and Progressive Transmission. Currently, there are many compression and progressive transmission algorithms available. Interested readers are referred to [5, 23, 24, 25, 36]. Progressive transmission algortihms rapidly provide successively better approximations to the input image as the digital data arrives from the network. The progressive property allows users to truncate an image data sequence and still get a reasonably good recovery of the image under this bit rate. Fully embedded compression algorithms generate image data sequences that can be truncated anywhere. Because successive approximation of images solves the conflict of different requirements on image qualities, it is valuable in many applications, especially for medical images because of the wide range of quality requirements in retrieval and diagnosis. GIF89a, a predominant graphics format on the Web, has build-in interlaced progressive transmission and lossless compression algorithms. It is not suitable for medical image transmission due to the high blocky effect introduced in images prior to the last frame.

Since wavelet transforms decompose images into several resolutions, the coefficients, in their own right, form a successive approximation of the original images. Because of this property, wavelet transforms are naturally suited for progressive image compression algorithms. As a result, many current progressive compression algorithms use wavelet transforms as an initial step [26, 23]. This trend became stronger after Shapiro's [26] 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 [23], which is referred to as the S & P algorithm. This algorithm was applied to a large database of mammograms [1, 21] and was shown to be highly efficient using real clinical-quality evaluations. An important characteristic of the S & P algorithm and many other progressive compression algorithms is the encoding and decoding simplicity from the decorrelation of images into different frequency bands brought by the wavelet transforms. No training is needed since trivial scalar quantization of the coefficients is applied in the compression process. However, by sacrificing simplicity, recent research [12] has made improvements on the S & P algorithm by tuning the zerotree structure to specific data.

Image Security: Eye Detection and Text Detection. Before digital medical images in computer-based patient record systems can be distributed online, it is necessary for confidentiality reasons to eliminate patient identification information that appears in the images.

Face and eye detection algorithms [4, 15, 19] are at a mature stage of development in the computer vision community. For photographic medical images, it is necessary to use such an algorithm to detect and eliminate human eyes in order to protect patient privacy. Since the system we are developing deals with radiological images, such an algorithm is not necessary.

With the DICOM standard, it is easy to eliminate textual information such as patient name and ID. However, for digitized films or previous history images, a computerized detection and elimination algorithm is needed. The problem of text identification [17, 27] arises in many applications other than medical security. Document understanding systems locate text and figure captions on a page for processing by optical character recognizers. The detection of text in scanned maps and mechanical, electrical, and piping drawings is important for converting the paper form to computer-analyzable form. Work done by University of Maryland [11, 13] uses neural network, texture and multiresolution analysis to segment the documents into areas of text and areas of image or graphics. However, the algorithms used in such systems are not designed to handle superimposed text because it is difficult to differentiate the edges of text from the edges of the medical objects in the image, as illustrated in Figure 12.2.

12.1.2 Overview of Our Work

In this project, we have developed a wavelet-based progressive lossless compression and transmission algorithm that takes advantage of the Daubechies' wavelet transforms we used in the textual information detection and elimination (TIDE) module of the system. The current implementation of this algorithm is much simpler and hence faster than that of the top-of-the-line algorithms mentioned above. The tradeoff is the recovering image quality at a given bit rate. The quality loss in terms of PSNR compared with the S & P algorithm is within $2dB$. However, our purpose of presenting it is mainly to explain the basic ideas of progressive transmission based on wavelet transforms and to demonstrate an efficient combination of text detection and progressive lossless compression and transmission of digital medical images in our TIDE system.

The TIDE module in the system consists of an efficient and accurate algorithm to distinguish areas with and without textual information in digital or digitized medical images. Areas with text can then be blurred or striped. Because variations in the diagonal directions can be found in almost all Roman characters or Arabic numbers, we use Daubechies' wavelets and analysis techniques to detect the high frequency variation in the diagonal direction that is indicative of text. A mask is used to preserve the losslessness of non-textual areas. With some basic knowledge of the machine used to create the image, we are able to eliminate only sensitive patient identification information while retaining the medical information in the image. Excellent results have been obtained in experiments using a large set of real-world medical images, many with superimposed text.

12.2 BACKGROUND ON DAUBECHIES' WAVELETS

Discrete Fourier Transforms are currently used effectively in signal and image processing because of the frequency domain localization capability. They are ideal for analyzing periodic signals because the Fourier expansions are periodic. However, because of their infinite extensibility, they do not have the spatial localization property needed to locate accurately areas with text within a medical image.

Two mathematical methods are available for non-periodic signals, the Windowed Fourier Transform (WFT) and the wavelet transform. The WFT analyzes the signal in both spatial and frequency domains simultaneously by encoding the signal through a scaled window related to both location and local frequency. Therefore, signals are easily underlocalized or overlocalized in the spatial domain if the spatial behavior is inconsistent with the frequency of the signal. Since we do not know the exact size of the fonts used in the medical images we are dealing with, flexible and adaptive localization is necessary.

Wavelets are basis functions that have some similarities to both splines and Fourier series. They have advantages when the aperiodic signal contains many discontinuities or sharp changes. Wavelets, developed in mathematics, quantum physics, and statistics, are functions that decompose signals into different frequency components and analyze each component with a resolution matching its scale. Applications of wavelets to signal denoising, image compression, image smoothing, fractal analysis and turbulence characterization are active research topics.

Wavelet analysis is based on an approach developed by Haar [20]. Haar found orthonormal bases defined on $[0,1]$, namely $h_0(x), h_1(x), ..., h_n(x), ...$, other than the Fourier bases, such that for any continuous function $f(x)$ on $[0,1]$, the series

$$\sum_{j=1}^{\infty} < f, h_j > h_j(x) \tag{12.1}$$

converges to $f(x)$ uniformly on $[0,1]$. Here, $< u, v >$ denotes $\int_0^1 u(x)\overline{v(x)}dx$ and $\overline{v}$ is the complex conjugate of v.

One version of Haar's construction [20] can be written as follows:

$$h(x) = \begin{cases} 1, x \in [\ 0, 0.5\ ) \\ -1, x \in [\ 0.5, 1\ ) \\ 0, elsewhere \end{cases} \tag{12.2}$$

$$h_n(x) = 2^{j/2} h(2^j x - k) \tag{12.3}$$

where $n = 2^j + k$, $k \in [\ 0, 2^j\ )$, $x \in [\ k2^{-j}, (k+1)2^{-j}\ )$.

One drawback of using Haar's base function to decompose images is that the Haar transform cannot efficiently separate image signals into low frequency and high frequency bands. From the signal processing point of view, since the wavelet transform is essentially a convolution operation, performing a wavelet transform on an image is equivalent to passing the image through a low-pass filter and a high-pass filter. The low-pass and high-pass filters corresponding to the Haar transform do not have a sharp transition and fast attenuation property. Thus, the low-pass filter and high-pass filter cannot separate the image into clean distinct low frequency and high frequency parts.

Another basis for wavelets is that of Daubechies wavelet transform with longer length filters [9] that has better frequency properties. For each integer r, Daubechies' orthonormal basis [10, 18] for $L^2(\mathbb{R})$ is defined as

$$\phi_{r,j,k}(x) = 2^{j/2}\phi_r(2^j x - k),\ \ j, k \in \mathbb{Z} \tag{12.4}$$

where the function $\phi_r(x)$ in $L^2(\mathbb{R})$ has the property that $\{\phi_r(x-k) | k \in \mathbb{Z}\}$ is an orthonormal sequence in $L^2(\mathbb{R})$.

Then the *trend* f_j, at scale 2^{-j}, of a function $f \in L^2(\mathbb{R})$ is defined as

$$f_j(x) = \sum_k < f, \phi_{r,j,k} > \phi_{r,j,k}(x). \tag{12.5}$$

The *details* or *fluctuations* are defined by

$$d_j(x) = f_{j+1}(x) - f_j(x). \tag{12.6}$$

To analyze these details at a given scale, we define an orthonormal basis $\psi_r(x)$ having properties similar to those of $\phi_r(x)$ described above.

The functions $\phi_r(x)$ and $\psi_r(x)$, called the *father wavelet* and the *mother wavelet*, respectively, are the wavelet basis functions required by the wavelet analysis. Figure 12.1 shows some popular mother wavelets. The family of wavelets such as those defined in Eq. 12.4 are generated from the father or the mother wavelet by change of scale and translation in time (or space in image processing).

Daubechies' orthonormal basis has the following properties:

- ψ_r has the compact support interval $[0, 2r+1]$;
- ψ_r has about $r/5$ continuous derivatives;

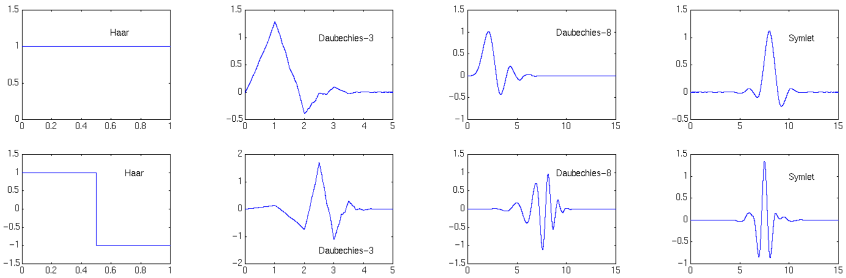


Figure 12.1 Plots of some analyzing wavelets. First row: father wavelets, $\phi(x)$. Second row: mother wavelets, $\psi(x)$

- $\int_{-\infty}^{\infty} \psi_r(x)dx = ... = \int_{-\infty}^{\infty} x^r \psi_r(x)dx = 0.$

Daubechies' wavelets give remarkable results in image analysis and synthesis due to the above properties. In fact, a wavelet function with compact support can be easily implemented by finite length filters. This finite length property is important for spatial domain localization. Furthermore, functions with more continuous derivatives analyze continuous functions more efficiently and avoid the generation of edge artifacts. Since the mother wavelets are used to characterize details in the signal, they should have a zero integral so that the trend information is stored in the coefficients obtained by the father wavelet. A Daubechies' wavelet representation of a function is a linear combination of the wavelet function elements.

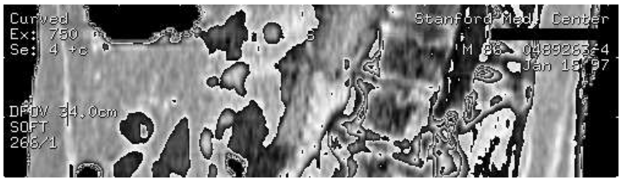
Daubechies' wavelets are usually implemented in numerical computation by quadratic mirror filters [20]. Multiresolution analysis of trend and fluctuation is implemented using convolution with a low-pass filter and a high-pass filter that are versions of the same wavelet. For example, if we denote the sampled signals as $x(n), n \in \mathbb{Z}$, then

$$F_0(x(n)) = \frac{1}{\sqrt{2}}(x(n) + x(n+1)) \tag{12.7}$$

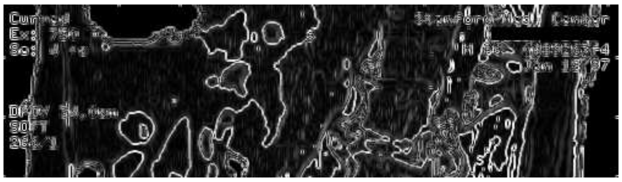
$$F_1(x(n)) = \frac{1}{\sqrt{2}}(x(n) - x(n+1)) \tag{12.8}$$

are quadratic mirror filters for Haar's wavelet.

Besides the advantage of the multiresolution approach, resulting a highly efficient algorithm design, Daubechies' wavelets offer a wide range of flexibility. We may select the appropriate wavelet basis to obtain the exact amount of fluctuation we desire in the high-frequency bands. Clean separation of high-frequency and low-frequency information is essential for an efficient compression algorithm. We have used Daubechies' wavelets also in content-based image retrieval [29, 31] and a system for screening objectionable images [30, 32, 33]. Our group is currently considering how to integrate a wavelet-based medical image indexing and retrieval algorithm with the TIDE system for on-line medical image distribution.



Source image



Traditional edge detection

HH band of Daubechies-4 wavelet transform

Figure 12.2 **Comparison of traditional edge detection and Daubechies-4 wavelet transform in distinguishing areas with and without textual information.** Name of the patient was manually blackened in the source image.

In textual information detection, especially for superimposed text, we want to distinguish areas with and without textual information as effectively as possible. When using the Haar wavelet, we obtain too much noise in the high-pass bands within the non-text areas. Traditional edge detection algorithms have the same problem, as illustrated in Figure 12.2. For both of the two algorithms, it would be difficult to differentiate the edges of text from the edges of the objects in the image. In our TIDE system, we use Daubechies-4 wavelet basis to obtain a clean separation of a medical image into low frequency and high frequency bands. In Figure 12.2, clusters of bright pixels, corresponding to large coefficients, occur in regions of the image that contain text.

12.3 THE ALGORITHM

12.3.1 Overview

We have developed a new progressive transmission algorithm with textual information detection and elimination for digital medical images using Daubechies' wavelet transforms. Figure 12.3 shows the basic structure of the algorithm.

We apply an N-level fast wavelet transform (FWT) with Daubechies-4 wavelet to each medical image, where N is determined adaptively by the image size. If the image is of DICOM standard, we may eliminate the patient identification information without processing the image content. The wavelet transform code can be directly passed to the progressive transmission and lossless compression coding routine.

For non-DICOM images, we extract and analyze the lower right-hand corners of each level of the transform matrix, where the diagonal directional high frequency information is located, to obtain a mask containing only the areas with textual information. Once such a mask is computed, we apply it to all the high-frequency bands to eliminate the text within areas with textual data. Or, we may apply the mask selectively to all the frequency bands to block the areas with text. Knowledge of the rough location (e.g., which one of four corners) of the critical patient identification information of certain type of medical images or the TIHI system [35] is used to eliminate only information needed to be deleted while preserving the rest. When we do not have knowledge of the rough location of patient identification information, we may apply the mask to eliminate all textual information within the medical image.

Then we pass the processed wavelet transform code to a new adaptive progressive transmission and lossless compression coding routine. To achieve progressive secure transmission, we may also apply a PGP encryption [14] to the code segments before sending the data to the TIDE client via public network.

Our text detection algorithm has several immediate advantages.

1. Unlike traditional approaches, such as the neural network, our algorithm does not depend on the actual font size, font type and style of the text in the medical image. Experiments indicate that the algorithm is capable of handling images with superimposed hand-written text and even foreign languages. Figure 12.12 shows one example with hand-written text.

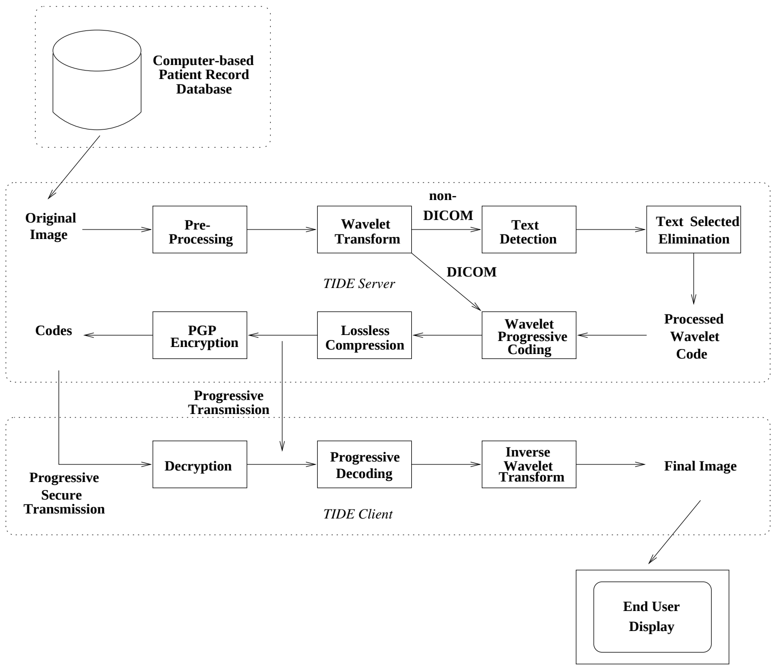


Figure 12.3 **Basic structure of the algorithm used in TIDE.**

2. We used Daubechies' wavelets rather than a traditional edge detector to capture the high frequency information in the images. This reduced the dependence of the results on the quality or the sharpness of the images.

3. The algorithm does not rely on the color of the image or the text. It also has minimum dependence on the contrast between text and background objects.

4. It is faster than other algorithms due to our adaptive multiresolution approach.

5. Wavelet-based algorithm using Daubechies' wavelets can be easily integrated with cutting-edge image compression, compressed-domain indexing and processing algorithms.

6. Our wavelet-based compression and progressive transmission algorithm does not have edge artifacts found in many non-wavelet-based compression algorithms.

12.3.2 Pre-processing

Many medical image formats and bit rates are currently in use. DICOM, PPM, GIF, JPEG and TIFF are the most widely used formats. Because the images

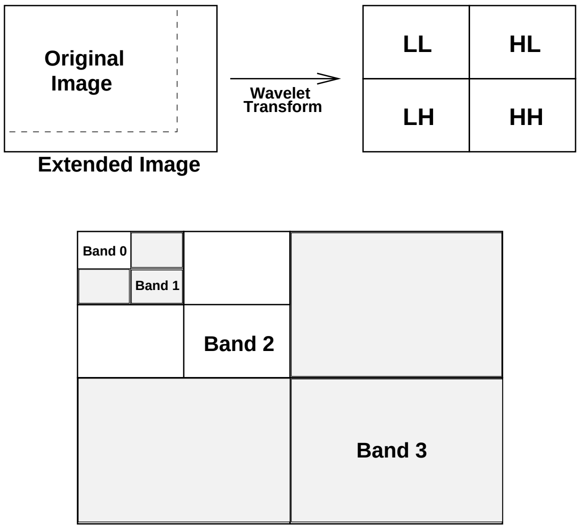


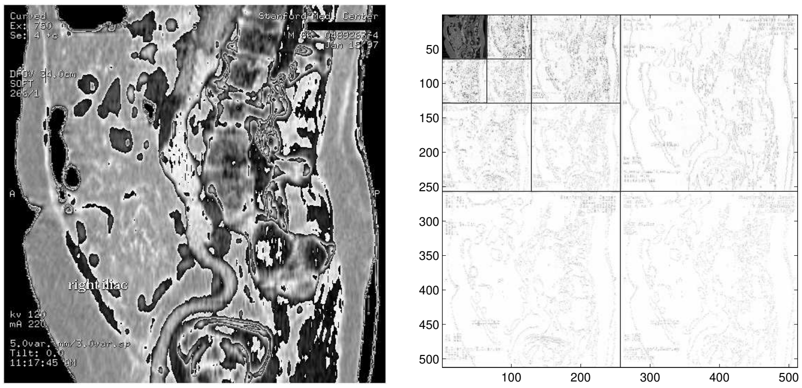
Figure 12.4 **Naming convention for a wavelet transform.**

may have different format, we must first normalize the data for computation. A gray-scale PPM image of any size is adequate for our textual information detection algorithm. A color medical image may be considered equivalent to three gray-scale images. The range of the values of each pixel, or the number of bits per pixel, for the PPM image is not limited for our algorithm. Usually it is adequate to use 12 bits per pixel to store a reasonably clear gray-scale medical image. If the core memory of the computer cannot handle some very large medical images, we must partition the image and perform the same operations on each portion of the original image.

In order to perform an N-level wavelet transform to a medical image, we must make sure the size of the image is suitable, i.e., the number of pixels on each side of the image must be divisible by 2^N. This is due to the downsampling process of the wavelet transform. For images that do not satisfy this constraint, we may simply extend the boundary row or column of the original image to form a larger image that satisfies this constraint.

12.3.3 Wavelet Transform

In this step, we apply an N-level wavelet transform to the image obtained from the pre-processing step. Here, N is determined adaptively so that the smallest band obtained in the highest level wavelet transform contains at least 8×8 pixels. We do not need a high quality edge detection algorithm for this application. Since the goal here is to effectively distinguish the areas with and without textual information, it is not necessary to produce a perceptually pleasant edge image. Consequently, we try to keep the algorithm simple to achieve a fast computation speed.



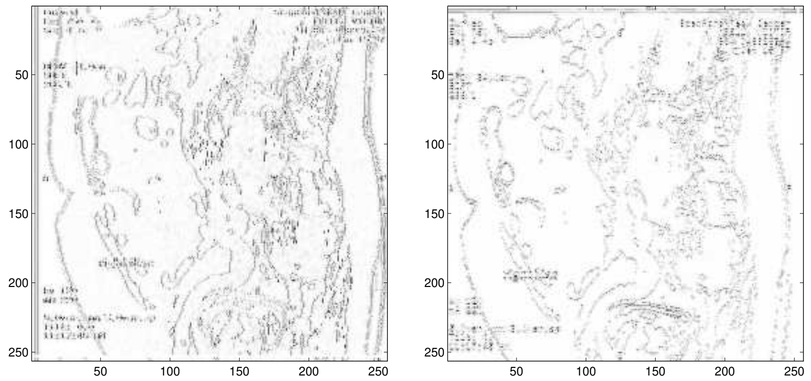
Source image *Wavelet transform*

Figure 12.5 **Daubechies' wavelet transform in TIDE.** Name of the patient was manually blackened in the source image.

We start the process by transforming the gray-scale PPM image converted from the pre-processing using the Daubechies-4 wavelet basis. Figure 12.5 shows a 3-level wavelet transform on a medical image with superimposed text. The image is decomposed into four frequency bands with corresponding names marked in Figure 12.4. Then the low-pass band is decomposed into four smaller frequency bands to replace the low-pass band in the first level decomposition, and so on.

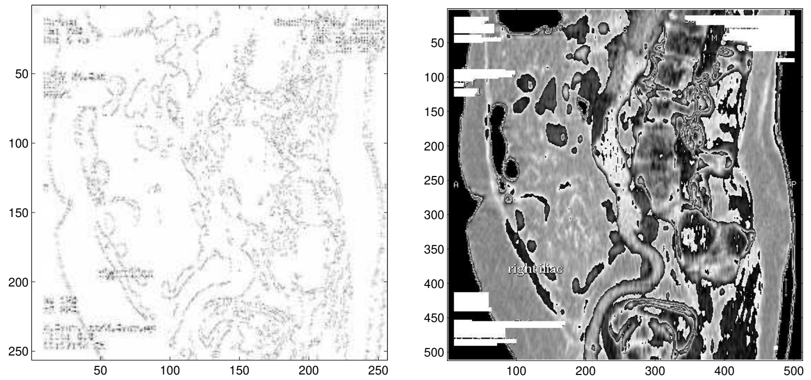
For simplicity, we borrow the notation from the filtering literature [28]. The letter 'L' stands for low frequency and the letter 'H' stands for high frequency. The left upper band is called 'LL' band because it contains low frequency information in both the row and column directions. We avoid the details of explaining the filtering terminologies here; interested readers are referred to [28].

The LH frequency band in each level is sensitive to the horizontal edges for that level of scale, the HL band is sensitive to the vertical edges, and the HH band is sensitive to the diagonal edges [10]. For the medical images that our system is designed for, the HH bands are much better for dealing with the distinctions between areas with and without text. In fact, variations in the diagonal directions can be found in almost all Roman characters or Arabic numbers. Such variations are detected much more frequently in areas with textual information than those with only medical objects, if we make a reasonable assumption that, in general, the text in the medical image is small compared to the objects in the image. The LH bands and the HL bands are not useful for distinguishing areas with and without textual information. As shown in Figure 12.6, vertical clusters of points can be found in both text areas and non-text areas in the HL band, and horizontal clusters of points can be found for the LH band. Therefore, we consider only the HH bands in the lowest 3 to 4 levels for textual information detection and elimination module.



HL band *LH band*

Figure 12.6 HL band and LH band are not suitable for detecting textual information within medical images.



HH band *Text areas detected*

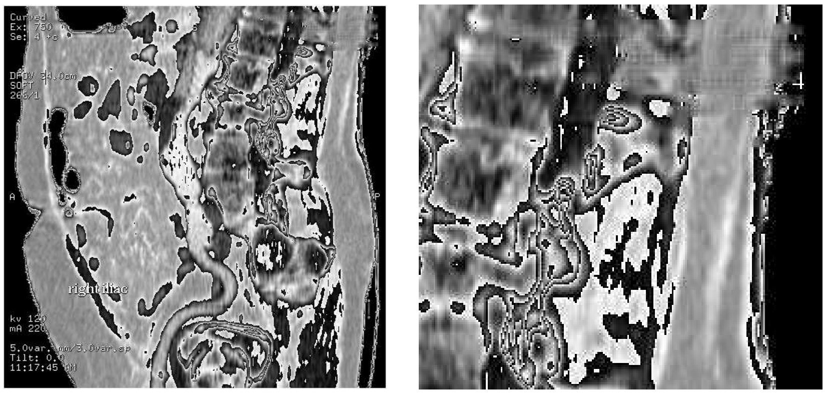
Figure 12.7 Wavelet coefficient analysis in the TIDE algorithm.

Analysis of the HH bands is required to avoid the incorrect elimination of diagonal-wise variations in areas of the image without text. Without loss of generality, we assume that the original image after pre-processing is of size $2n \times 2n$. Then the wavelet transform $W(1 : 2n, 1 : 2n)$[1] is a matrix of size $2n \times 2n$. The process is similar if the image is not a squared image.

In this step, we process the HH matrices in a few frequency bands, depending on the size of the original image. For an image of size close to 512×512, we consider the $HH_1(1 : n, 1 : n) = W(n + 1 : 2n, n + 1 : 2n)$ matrix and the

[1]Here we use MATLAB notation. That is, $A(m_1 : n_1, m_2 : n_2)$ denotes the submatrix with opposite corners $A(m_1, m_2)$ and $A(n_1, n_2)$.

$HH_2(1:\frac{n}{2},1:\frac{n}{2}) = W(\frac{n}{2}+1:n,\frac{n}{2}+1:n)$ matrix obtained from the previous step. Binary matrices, denoted as $B_1(1:n,1:n)$ and $B_2(1:\frac{n}{2},1:\frac{n}{2})$, are constructed from the HH_1 and HH_2 matrices so that the largest $k = O(n)$ coefficients in magnitude in the HH matrix are replaced by 1 and all other coefficients are replaced by 0. k is determined adaptively by the size of the HH matrices. Then we use a moving square window matrix of about $w \times w$ pixels to determine the isolated points in the binary matrix by setting a threshold for the minimum number of non-zero points in such a moving window. For a B matrix of size 256×256, $w = 20$ is appropriate. The isolated points in B_1 and B_2 matrices are then deleted because they represent diagonal-wise variations in the areas without text. Note that we consider HH matrices of lower scales if the original image is a lot larger than 512×512.



Full image *Upper-right corner*

Figure 12.8 Final images after wavelet reconstruction. After text detection, we use the TIHI process or some other basic knowledge of the location to eliminate any text not known to be innocuous.

Denote $B_1'(1:n,1:n)$ and $B_2'(1:\frac{n}{2},1:\frac{n}{2})$ the matrices without the isolated points, converted from $B_1(1:n,1:n)$ and $B_2(1:\frac{n}{2},1:\frac{n}{2})$, respectively. Then we group up the remaining points in the matrices B_1' and rescaled B_2' to form a matrix $Mask(1:n,1:n)$ containing detected textual areas. We amend the mask using the TIHI or some other basic knowledge of the rough location of the patient identification information. Finally, we rescale the amended mask appropriately and apply it to the high-frequency bands of the wavelet transform to eliminate the text. Alternatively, we may apply the amended mask to all the frequency bands to obtain a blocked final image. Figures 12.7 shows the HH band, and the result after isolated points are eliminated to detect text areas. Figure 12.8 shows the image with text eliminated.

12.3.5 Wavelet-based Lossless Compression and Progressive Transmission

Since wavelet transforms using Daubechies' advanced wavelet basis are used in the textual information detection module, we have designed and implemented

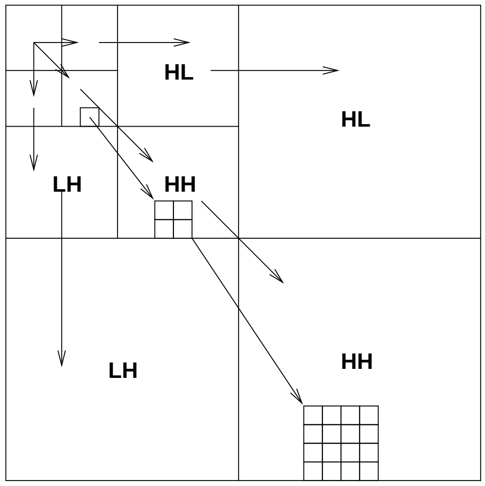


Figure 12.9 **Cross-band dependencies in the zerotree algorithm.**

a wavelet-based lossless compression and progressive transmission algorithm to incorporate compression with the TIDE module for secure and efficient medical image distribution.

We compared the performance of the current implementation of our algorithm with one of the best available image compression algorithms on a few common benchmark images. We have achieved comparable results, and we expect the performance to be improved if we take into consideration the cross-band correlation of wavelet coefficients. The structure of cross-band correlation captured in the zerotree algorithms is illustrated in Figure 12.9.

Our coding algorithm at the server side is outlined as follows:

1. Assume the processed wavelet transform from textual information detection part contains N levels. Group the wavelet coefficients into $N+1$ bands. Band 0 is the LL band of level 0 and band i is the combination of HL, LH and HH bands at the decomposition level i. Fig. 12.4 shows the details of this naming convention. We write N to the channel. For simplicity, we say output N.

2. Calculate and output necessary header parameters

   (a) Output the size of the image

   (b) Calculate the mean value of the coefficients in Band 0, denote the mean value by $\bar{c}$ and output $\bar{c}$. Update the coefficients by subtracting the coefficients in Band 0 by $\bar{c}$.

   (c) Calculate the maximum absolute value of the coefficients, denote it by c_{max} and output c_{max}. Also calculate the maximum absolute values of the coefficients in all the bands, denote them by $c_0, c_1, ..., c_N$.

(d) Calculate the variances of the coefficients in the $N+1$ bands separately. Denote the variances by $v_i, i = 0, ..., N$. Denote the starting encoding bit plane for band i as $S_i, i = 0, ..., N$, which is determined as follows,

$$\begin{aligned} S_0 &= 0; \\ S_i &= max(0, \left[\log\left(\frac{v_0}{v_i}\right)\right]). \end{aligned}$$

The base of the logarithm is 2. Output $S_i, i = 1, ..., N$.

3. Encode the coefficients as specified below. The output bits are then losslessly encoded by run-length coding. For notational simplicity, we omit pointing out that the output bits are being encoded by run-length coding and possibly PGP secure coding later. Hence, when we say 'output a bit', it does not mean that the bit is sent out directly, but rather that the bit is fed into a run-length encoder and the run-length encoder sends out a bit stream after lossless compression is done.

   (a) $c_{max} \rightarrow q$

   (b) $q/2 \rightarrow q$, $0 \rightarrow bitplane$.

   (c) i. Scan the coefficients in Band 0 row by row.

      ii. If the absolute value of a coefficient is smaller than q, output bit 0.

      iii. Otherwise:
      Output bit 1.
      If the coefficient is negative, update the coefficient by adding q. If the coefficient is positive, update the coefficient by subtracting q. If the absolute value of the coefficient is detected greater than q for the first time, i.e., the absolute value is smaller than all the previous q's, output bit 1 if the coefficient is positive and output bit 0 if the coefficient is negative.

   (d) $1 \rightarrow i$

   (e) If $S_i > bitplane$, $i + 1 \rightarrow i$, go back to 3e.

   (f) If $S_i = bitplane$ and $S_i > 0$, encode Band i at all the previous bitplanes as follows, i.e., when quantization value is $c_{max}/2, c_{max}/4, ..., 2q$.

      i. $q' = c_{max}$

      ii. $q'/2 \rightarrow q'$, if $q' = q$, go to 3g. If $c_i < q'$, output bit 0, go to 3(f)ii. Otherwise, output bit 1.

         A. Scan the coefficients in Band i row by row, the HL band is scanned first, then the LH band, and finally the HH band.

         B. If the absolute value of a coefficient is smaller than q, output bit 0.

         C. Otherwise:
         Output bit 1.

If the coefficient is negative, update the coefficient by adding q. If the coefficient is positive, update the coefficient by subtracting q'. If the absolute value of the coefficient is detected greater than q' for the first time, i.e., the absolute value is smaller than all the previous qs, output bit 1 if the coefficient is positive and output bit 0 if the coefficient is negative.

D. Go to 3(f)ii

(g) Encode the coefficients in Band i with quantization value q.

i. Scan the coefficients in Band i row by row, the HL band is scanned first, then the LH band and the HH band.

ii. If the absolute value of a coefficient is smaller than q, output bit 0.

iii. Otherwise:
Output bit 1.
If the coefficient is negative, update the coefficient by adding q. If the coefficient is positive, update the coefficient by subtracting q. If the absolute value of the coefficient is detected greater than q for the first time, i.e., the absolute value is smaller than all the previous qs, output bit 1 if the coefficient is positive and output bit 0 if the coefficient is negative.

(h) $bitplane + 1 \rightarrow bitplane$

(i) Check whether the stopping criterion is met, if so, end. Otherwise, go back to 3b.

The encoding stopping criterion at the server can be specified by the client. Some examples are the number of output bits exceeding a threshold and the distortion of the lossy compressed image being less than a threshold. Because floating point operations are involved in the encoding process of the wavelet transform, we usually cannot get a perfect reconstruction of the original image, although we can make the distortion as small as we want. In the case that perfect reconstruction of the image is required, we can calculate the error image and compress the error image losslessly and transmit it at the end. As the errors are small, the error image is much more losslessly compressible than the original image.

12.4 RESULTS

12.4.1 Experimental Performance

This algorithm has been implemented on a Pentium Pro 200MHz workstation. We have tested about 100 medical images of different modalities, collected from different sources. Some of them are downloaded from the world-wide web and medical imaging newsgroups, while others are provided by the Stanford Medical Center.

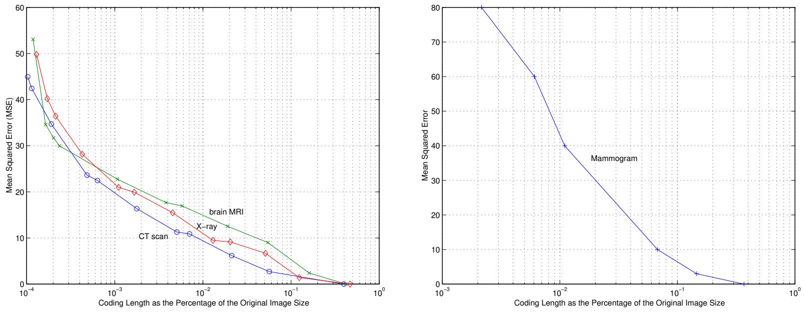


Figure 12.10 **Progressive lossless compression coding efficiency.** CT scan, chest X-ray and brain MRI images are 8-bit images. Mammogram images are 12 bit images.

Modality	CT scan	X-ray	Mammogram	MRI
size/img (bits)	$512^2 \times 8$	$2048^2 \times 8$	$4096^2 \times 12$	$256^2 \times 8$
img/exam	40	2	4	60
size/exam	10 MB	8 MB	128 MB	4 MB
$MSE < 10$	0.07MB (0.7%)	0.10MB (1.3%)	8.65MB (6.7%)	0.22MB (5.5%)
$MSE < 3$	0.56MB (5.6%)	0.61MB (7.6%)	18.6MB (14%)	0.64MB (16%)
$MSE = 0$	4.00MB (40%)	3.73MB (47%)	47.2MB (37%)	1.68MB (42%)

Table 12.1 Performance of progressive coding in TIDE.

Lossless Compression and Progressive Coding Module. The encoding and decoding programs are both very fast. It takes less than 2 seconds of CPU time to fully encode a 512×512 CT scan. The decoder takes about 2 seconds of CPU time to recover to the original image with Mean Squared Error (MSE) less than 3. To fully recover the original image without loss, it takes about 3 seconds of CPU time.

The efficiency of the lossless compression and progressive transmission module in the TIDE system is summarized in Table 12.1 and Figure 12.10. Since we used Daubechies continuous wavelet for compression, edge artifact is minimized. When MSE is lower than 3 for an 8-bit image, the distortion is barely noticeable to the human eye.

Textual Information Detection Module. The textual information detection and elimination module takes about 1 second of CPU time to process a 12-bit medical image of size 512×512. The algorithm is a linear algorithm

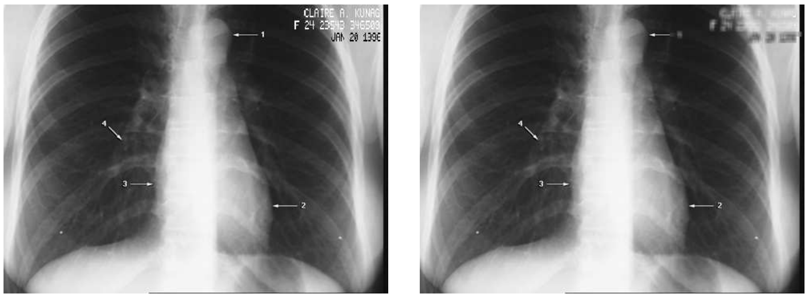


Figure 12.11 **A chest X-ray image processed by the TIDE system.** Note that patient identity information on the upper-right corner is eliminated while the arrows and annotations within the image are preserved. Text Information simulated.

with respect to the size of the image. Besides the fast speed, the algorithm has achieved remarkable accuracy. It successfully detected and eliminated all of the critical textual information within the corners of the medical images.

Figure 12.11 and Figure 12.12 show some sample results on gray scale medical images processed by the TIDE system. The areas without text are maintained without loss.

12.4.2 Limitations of the Text Detection Algorithm

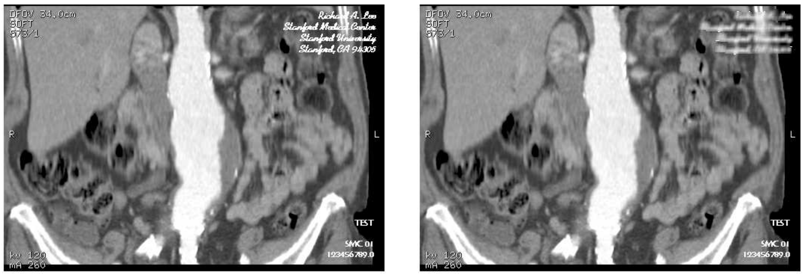


Figure 12.12 **A medical image with hand-written text processed by the TIDE system.** Text Information simulated.

The text detection algorithm in the TIDE system assumes that the text fonts are smaller in size than major medical objects that appear in the images, and small medical objects such as tissue are not clustered together as is the case with text characters. This is true for a vast majority of medical images. However, the algorithm is not suitable for images with text superimposed on a large area of clustered medical objects of similar size as the text characters.

Since the algorithm detects clustered diagonal directional variation, it does not detect printed capitalized Roman characters such as "H", "E" and "F" because no such variation presents in such characters. The algorithm makes corrections by assuming the characters in medical images are in general aligned to form text strings rather than randomly printed in the image. As noticed on examples provided, the algorithm works the best when detecting hand-written text strings. However, if the text strings start or end with characters without diagonal directional variation as discussed above, the algorithm would not be able to perform the deletion of such characters.

Finally, the algorithm assumes some basic knowledge of each modality. For example, if the image to be processed is a CT scan, it is likely that we can assume the sensitive textual information, such as the patient name and patient ID, can be found in upper-right corner of the image due to our knowledge of the machine that creates the image or the digitizer that digitized the film. If we do not know exactly which corner(s) the sensitive textual information is printed, we may still use the algorithm on all four corners of the image or the entire image if necessary. In this case, sensitive textual information will be deleted at the cost of the deletion of non-sensitive textual information. Often, some non-sensitive textual information found on a medical image, such as the character "L" or "R" on X-ray images and the serial number and study protocol for MRI and CT images, can be useful for radiologists to read and understand the medical image. Missing such information in consultation and interpretation may result in misdiagnosis. At this stage, we are unable to distinguish different types of textual information based on semantics. A robust OCR algorithm with some text understanding capability will be needed to make this possible.

12.5 CONCLUSIONS AND FUTURE WORK

In this chapter, we have discussed recent developments in wavelet-based image compression and progressive transmission algorithms. An efficient wavelet-based progressive transmission algorithm with security filtering features for on-line medical image distribution has been demonstrated. The algorithm uses Daubechies' wavelets to detect and eliminate textual information within digitized medical images, while maintaining non-textual areas lossless.

It is possible to improve the coding efficiency of the algorithm by including cross-band analysis of wavelet coefficients. We are also working on spatial thresholding in addition to the traditional bit-rate or quality-based thresholding. Integrating our system into an on-line real-time secure medical information retrieval system such as the TIHI system [34, 35] would eliminate dependence on assuming a fixed location for the identifying information. A user study is also important to make the system more effective for the medical community.

12.6 ACKNOWLEDGMENTS

We would like to thank Oscar Firschein and Michel Bilello of Stanford University, and Steve Wong of University of California at San Francisco for valuable help and discussions. We would also like to thank the Stanford University Libraries and Academic Information Resources (SULAIR) for providing computer equipment during the development and testing process.

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] RC Barrows Jr. and PD Clayton, Privacy, confidentiality, and electronic medical records, *J. of the American Medical Informatics Association*, 3:139-148, 1996.

[3] SH Becker and RL Arenson, Costs and Benefits of Picture Archiving and Communication Systems, *Yearbook of Medical Informatics*, 345-355, 1995.

[4] V Bruce and M Burton (eds), *Processing Images of Faces*, Ablex Pub. Corp., N.J., 1992.

[5] 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.

[6] Charles K. Chui, *An Introduction to Wavelets*, Academic Press, Inc., San Diego, 1992.

[7] Charles K. Chui, *Wavelets: A Tutorial in Theory and Applications*, Academic Press, Inc., San Diego, 1992.

[8] Ruel V. Churchill, *Fourier Series and Boundary Value Problems*, McGraw-Hill, New York, 1987.

[9] Ingrid Daubechies, Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, 41(7):909-996, October 1988.

[10] Ingrid Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1992.

[11] David Doermann, Document Understanding Research at Maryland, *DARPA Image Understanding Workshop Proc.*, Vol 2, p.817-826, Science Applications, Inc, 1994.

[12] 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.

[13] K Etemad, D Doerman, R Chellappa, Page Segmentation Using Decision Integration and Wavelet Packet Basis, *Proc. Int. Conference on Pattern Recognition*, IEEE Computer Society Press, 1994

[14] Simson Garfinkel *PGP : Pretty Good Privacy*, O'Reilly & Associates, Inc., CA, 1995.

[15] V Govindaraju, Locating Human Faces in Photographs, *Int Journal of Computer Vision*, Vol. 19, No. 2 Aug 96 p.129-146.

[16] National Health Expenditures for 1995, Health Care Financing Administration (HCFA), `http://www.hcfa.gov/`, Jan. 1997.

[17] AK Jain, S Bhattacharjee, Text Segmentation Using Gabor Filters for Automatic Document Processing, *Machine Vision and Applications*, 5:169-184, 1992.

[18] Gerald Kaiser, *A Friendly Guide to Wavelets*, Birkhauser, Boston, 1994.

[19] T Kanade, *Computer Recognition of Human Faces*, Birkhauser, 1977.

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

[21] 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.

[22] O Ratib, Y Ligier and JR Scherrer, Digital image management and communication in medicine, *Computerized Medical Imaging and Graphics*, 18(2):73-84, 1994.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] Yuan Y. Tang et al., Quadratic Spline Wavelet Approach to Automatic Extraction of Baseline from Document Images, *Proc. 4th International Conference on Document Analysis and Recognition*, IEEE, August 1997.

[28] Martin Vetterli, *Wavelets and Subband Coding*, Prentice Hall, N.J., 1995.

[29] James Ze Wang, Gio Wiederhold, Oscar Firschein, Sha Xin Wei, Wavelet-Based Image Indexing Techniques with Partial Sketch Retrieval Capability, *Proceedings of the Fourth Forum on Research and Technology Advances in Digital Libraries (ADL'97)*, Washington D.C., May 1997.

[30] James Ze Wang, Gio Wiederhold, Oscar Firschein, System for Screening Objectionable Images Using Daubechies' Wavelets and Color Histograms, *Proc. IDMS'97*, Springer-Verlag LNCS 1309, September 1997.

[31] 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.

[32] James Ze Wang, Jia Li, Gio Wiederhold, Oscar Firschein, System for Classifying Objectionable Images, to appear in *Computer Communications Journal*, Elsevier Science, 1998.

[33] James Ze Wang, Jia Li, Gio Wiederhold, Oscar Firschein, System for Classifying Objectionable Websites, submitted for conference publication, 1998.

[34] Gio Wiederhold, Interoperation, Mediation, and Ontologies, *Workshop on Heterogeneous Cooperative Knowledge-Bases*, Vol.W3, pages 33-48, ICOT, Tokyo, Japan, Dec. 1994.

[35] Gio Wiederhold, Michel Bilello, Vatsala Sarathy, XiaoLei Qian, A Security Mediator for Health Care Information, *Proceedings of the 1996 AMIA Conference*, Washington DC, Oct. 1996, pp.120-124.

[36] 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.

[37] STC Wong, HK Huang, A hospital integrated framework for multimodal image base management, *IEEE Trans Systems, Man, Cybernetics*, 26(4), 1996:455-469.