The effect of wavelet bases on the compression of digital mammograms

Zesheng Yang,[1] Maria Kallergi,[2] Ronald A. DeVore,[1] Bradley J. Lucier,[3]

Wei Qian,[2] Robert A. Clark,[2] and Laurence P. Clarke[2]

[1]Department of Mathematics, University of South Carolina, Columbia, SC 29208

[2]Department of Radiology, University of South Florida, Tampa, FL 33612-4799

[3]Department of Mathematics, Purdue University, W. Lafayette, IN 47907-1395

Acknowledgments: This work was supported by the H. Lee Moffitt Cancer Center & Research Institute at the University of South Florida and the Office of Naval Research Contracts N0014-91-J-1076 and N0014-91-J-1152.

Correspondence: Laurence P. Clarke, Ph.D.

Department of Radiology, USF

12901 Bruce B. Downs Blvd., Box 17

Tampa, FL 33612-4799

ABSTRACT

An initial evaluation of various types of wavelet bases is presented in this study of compression of digital mammograms. Screen/film mammograms (23 cm × 28 cm) digitized at a spatial resolution of 105 µm per pixel and a dynamic range of 12 bits per pixel (4096 grayscales) were compressed at different rates (from 5 to 195 to 1) using standard, hyperbolic, and adaptive wavelet bases. Images containing a cluster of microcalcifications, one of the smallest signs of breast cancer, were selected for this study to facilitate the visual evaluation of the magnitude of losses produced by the wavelet coding. The quality and content of the compressed reconstructed images were evaluated quantitatively and qualitatively by calculating the mean square error of the difference between the original and the compressed reconstructed images and by visual examination of the compressed reconstructed data and comparison to the original. Our preliminary results indicated that several wavelet bases could offer high compression rates at relatively low losses, the nature of which depends on the type of wavelet family used for compression. Some of the methods, for example,

retain all diagnostic features, in our case the calcifications, at the expense of some obvious artifacts in the smoothness of the surrounding normal tissues. Others, retain both the diagnostic features and data smoothness at the expense of lower compression rates. The hyperbolic Daubechies-4, Biorthogonal-1-3, Daubechies-2, adaptive Haar, and Haar wavelet bases were consistently rated as the top six methods for compression with the smaller mean square errors and visually better image quality. These methods will, therefore, be used for further research on medical image compression.

Key words: wavelets, image compression, multiresolution analysis, digital mammography.

I. INTRODUCTION

We have previously reported on the use of Haar wavelets for the lossy compression of digital mammograms.[1] Motivated by the results of our first study, we experimented with a variety of wavelet structures for the same purpose. We implemented various types of digital mammogram compression using standard, hyperbolic, and adaptive wavelet bases and we experimented with various compression rates in order to better understand the compression effects on the nature and quality of the images and the types of losses.

The need for high-rate compression algorithms in medical imaging has been repeatedly stated. It is generated by the requirements for high spatial resolution and dynamic range data which need to be stored effectively, accessed, and transmitted at high rates.[2],[3] It is also well known that screening digital mammography implies real-time availability to the radiologist of a series of large images per patient for comparative study and accurate diagnosis.[4] Data storage, transmission and display technologies are challenged by these requirements. High-rate compression methods offer them a possible *cost-effective* solution.

Both lossless and lossy compression methods have been investigated for medical imaging applications.[5],[6] Lossless methods have the advantage of errorless reconstruction but the disadvantage of small compression rates, on the order of 2 or 3 to 1.[7] In contrast, lossy techniques can achieve very high compression at the expense of errors in the reconstructed images.[1],[8],[9],[10] Receiver operating characteristic (ROC) studies of compressed hand and chest images have shown that, in some contexts, approximately 10:1 lossy compression does not compromise diagnostic accuracy while 20:1 compression may be adequate for the transfer and archiving of similar data.[8],[11]

The lossy compression techniques reported until now are mostly based on cosine transforms. For example, the Joint Photographic Experts' Group (JPEG) still-picture standard of compression is based on block cosine transforms. As a result, blocking artifacts occur at the borders of the subimages in which the original image is partitioned. Experimental evidence indicates that for compression rates higher than 10:1, wavelet-based methods (with an appropriate encoder) substantially outperform JPEG. For this reason, we directed our studies toward wavelet-based algorithms, which have shown promising performance in the areas of medical image enhancement[12] and segmentation,[13] and equal promise in image compression. applications.[1],[14],[15],[16] The present study is a continuation of our initial work on Haar wavelet compression of mammograms,[1] the results of which indicated that Haar wavelets preserve breast features such as calcifications, even at high compression rates (≤50:1), but degrade the smoothness of the surrounding normal tissues, generating blocking artifacts similar to the JPEG technique. Based on these observations, we decided to pursue the implementation of smoother wavelet families that may improve the texture of the reconstructed images while keeping all important diagnostic information. This of course entailed the risk of creating different types of artifacts that would need to be identified.

The purpose of this paper is four-fold: (a) develop a variety of wavelet methods including standard, hyperbolic, and adaptive wavelet bases for the compression of high-resolution digital mammograms, (b) apply different levels of compression to the images and determine the dependence of the information loss on the compression ratio, (c) determine qualitatively the types of artifacts and errors introduced to the compressed reconstructed images by the various wavelets and make a relative comparison, and (d) discriminate among the various wavelet bases and select those worthy of further analysis and study.

II. WAVELET METHODS

Wavelet bases

Lossy wavelet-based compression algorithms have the following main ingredients: (a) a choice of wavelet bases to represent the image; (b) the transformation of pixel values to wavelet coefficients; (c) the quantization of the wavelet coefficients; and (d) the encoding of the quantized coefficients. The rendering of the compressed file then consists of decoding the encoded coefficients and then reconstructing pixel values from the quantized coefficients. The compressed image can be viewed as an approximation to the original.

There is a choice in steps (a), (c), and (d) as to the best methods and to a large extent these steps are interdependent and should be viewed together. However, the main purpose of the present paper is to investigate various choices of wavelet bases in step (a). We single out this step in order to focus on the most promising bases for further study. We shall briefly describe the wavelet bases that we study in this investigation; a more detailed description of these wavelet bases can be found in the survey article by DeVore[17] or one of the several papers we cite in this section. A discussion of steps (b), (c) and (d) will also be given here.

The construction of wavelet bases is usually implemented by using multiresolution analysis as introduced by Mallat[18] and Meyer[19]. In one spatial dimension, the starting point is a ladder of function spaces $V_k \subset L_2(\mathbb{R})$, $k \in Z$ (the set of integers), which are assumed to have certain properties. The most important of these is the nesting assumption:

$$V_k \subset V_{k+1}, \quad k \in Z.$$

One also assumes that each space V_k is obtained from each other by dilation. Thus V_k consists of the dilates $v(2^k x)$ of functions v from V_0. The usual examples in multiresolution analysis obtain the spaces V_0 from linear combinations of a function ϕ (called a scaling function) and its integer translates $\phi(x - j)$, $j \in Z$. Wavelets describe how to obtain the finer space V_1 from the coarser space V_0. There are several possibilities but we limit our consideration to the following two.

Orthogonal wavelets: In this case, we consider the orthogonal projector

$P := P_0$ from $L_2(\mathbb{R})$ onto V_0. By dilation, we obtain the projector P_1 from $L_2(\mathbb{R})$ onto V_1. The residual projector $Q := P_1 - P_0$ describes the information that must be added to V_0 to obtain the space V_1. Namely, for any $f \in V_1$, we have

$$f = P_1 f = P_0 f + (P_1 - P_0) f = P_0 f + Q_0 f$$

The functions $P_0 f$ and $Q_0 f$ are orthogonal, i.e.

$$\int_{\mathbb{R}} P_0 f(x) Q_0 f(x) dx = 0.$$

In terms of spaces, this means that V_0 is orthogonal to $W := V_1 \ominus V_0$ (the orthogonal complement of V_0 in V_1) and Q_0 is the orthogonal projector onto W.

The space W is called the wavelet space and it represents the information that must be added to V_0 to obtain V_1. An orthogonal wavelet is a function ψ whose integer translates $\psi(x-j)$ from an orthonormal basis for W. By dilation, the functions

$$\psi_{j,k}(x) := 2^{k/2}\psi(2^k x - j), \quad j \in \mathbb{Z} \tag{1}$$

are an orthonormal basis for $W_k := V_k \ominus V_{k-1}$. We thus obtain that $L_2(\mathbb{R})$ is the orthogonal sum of the W_k, $k \in Z$:

$$L_2(\mathbb{R}) = \bigoplus \sum_{k \in Z} W_k$$

and that the functions $\{\psi_{j,k}\}_{j,k \in Z}$ are an orthonormal bases for $L_2(\mathbb{R})$. This means that each function $f \in L_2(\mathbb{R})$ has the representation

$$f = \sum_{j,k \in Z} \langle f, \psi_{j,k} \rangle \psi_{j,k}$$

with

$$\langle f,g \rangle := \int_{\mathbb{R}} f(x)g(x)dx$$

the inner product of the two real-valued functions f and g.

The simplest example of an orthonormal basis that can be constructed in this manner is the Haar basis, which starts with the generating function $\phi = \chi_{[0,1]}$ (the characteristic function of the interval [0,1]) which is one on this interval and zero outside of it. The orthogonal wavelet ψ in this case is the Haar function

$$H(x) := \begin{cases} +1, & 0 \le x \le 1/2 \\ -1, & 1/2 < x \le 1. \end{cases}$$

While the Haar function has many remarkable properties, it is not continuous, which limits some of its potential applications in image processing. One of the remarkable achievements of Daubechies[20] was to construct a family of orthogonal wavelets D_k that have compact support (i.e. vanish outside of a finite interval) and have smoothness (differentiability). The wavelet D_1 is the Haar wavelet. Increasing the parameter k increases the smoothness and approximation capability of the wavelet but also increases the size of its support. Daubechies wavelets will be used in our compression algorithms for digital mammography.

Biorthogonal wavelets. One of the deficiencies of orthogonal wavelets is their lack of symmetry (only in the case of the Haar wavelets do the Daubechies wavelets have any kind of symmetry). Biorthogonal wavelets were constructed in part to remedy this deficiency.[21],[22] They can be constructed from multiresolution by using oblique projectors (in place of the orthogonal projectors) for P and Q. The result is a function ψ such that the family of functions $\psi_{j,k}$ of (1) forms a non-orthogonal basis for $L_2(\mathbb{R})$. Each $f \in L_2(\mathbb{R})$ has the representation

$$f = \sum_{j,k \in Z} \langle f, \eta_{j,k} \rangle \psi_{j,k} .$$

with the $\eta_{j,k}$ obtained as in (1) from an appropriate function $\eta \in L_2(\mathbb{R})$.

There are many possibilities for biorthogonal wavelets but we shall limit our consideration to a very special class of biorthogonal wavelet obtained from spline functions.[23]. They are indexed on two integer parameters k and l that relate to the size of the decomposition and reconstruction filters. The integer k indicates that the dual wavelet function η is a spline function of order k (degree k-1) with breakpoints at the half integers. The dual wavelet η has l vanishing moments. A discussion of these wavelets and graphs of typical examples can be found in Ref.23.

Bivariate wavelets

Images are bivariate and need to be represented using bivariate wavelet bases. There are two natural ways to obtain bivariate wavelet bases from univariate wavelets.

Regular bases. The usual method for representing bivariate functions with wavelets is as follows. From the univariate scaling function ϕ and the corresponding wavelet ψ, we for the set Ψ consisting of the three functions

$$\phi(x)\,\psi(y),\ \ \psi(x)\,\phi(y),\ \ \psi(x)\,\psi(y).$$

We denote the function in Ψ by $\lambda := \lambda(x,y)$. The functions

$$\lambda_{j,k} := 2^k\lambda(2^kx-j_1,2^ky-j_2),\ \ j=(j_1,j_2),\ j_1,j_2,k\in\mathbb{Z},\ and\ \lambda\in\Psi$$

form a basis for $L_2(\mathbb{R}^2)$. In this way, we can obtain the standard Daubechies D_k bivariate bases or the (l,k) biorthogonal bivariate bases.

The function $\lambda_{j,k}$ is a shifted dilate of the function λ. Thus, what is characteristic of the standard bases is that each of the basis functions $\lambda_{j,k}$ is a scaled version of one of the three functions in Ψ. The support of $\lambda_{j,k}$ is also a shifted-dilate of the support of λ.

Hyperbolic bases. There is another natural bivariate wavelet basis which can be formed from the univariate wavelet ψ; it consists of the tensor products

$$\psi_{j,k}(x)\psi_{j',k'}(y),\ \ j,j',k,k'\in\mathbb{Z} \tag{2}$$

We call the functions in Eq. (2) hyperbolic basis to make clear its distinction from the standard basis.

One should note that in the hyperbolic basis, the dilation parameters 2^k and $2^{k'}$ are different in the two coordinate directions. The function $\psi_{j,k}(x)\psi_{j',k'}(y)$ is, therefore, not simply a shifted dilate of the function $\psi(x)\psi(y)$ since the scaling is different in the two coordinate directions. Similarly the support of $\psi_{j,k}(x)\psi_{j',k'}(y)$ is scaled differently in the two coordinate directions. This results in rectangular supports which may be much longer in one coordinate direction than in the other.

Adaptive bases. The standard and hyperbolic bases are not the only two possible bivariate bases that can be obtained from a univariate pair ϕ, ψ. Suppose for example that the integer translates of ϕ and ψ are orthogonal. We consider the totality of functions

$$\phi_{j,k}(x)\phi_{j',k'}(y), \quad \phi_{j,k}(x)\psi_{j',k'}(y), \quad \psi_{j,k}(x)\phi_{j',k'}(y), \quad \psi_{j,k}(x)\psi_{j',k'}(y), \tag{3}$$

where $j, j', k, k' \in \mathbb{Z}$. The two wavelet bases already mentioned are only special cases of orthogonal bases for $L_2(\Omega)$, $\Omega := [0,1]^2$, that can be selected from the totality of functions in Eq. (3). We mention some examples in the special case of the Haar wavelets, which are useful in partitioning (segmenting) images. We start with any partition $\mathfrak{R}$ of $\mathbb{R}^2$ into dyadic rectangles R. On each rectangle R, we can complete $\phi_{j,k}(x)\phi_{j',k'}(y)$ to an othogonal bases for $L_2(R)$ by adjoininga standard or hyperbolic Haar basis of R. Doing this for each rectangle R results in an orthogonal basis for $L_2(\mathbb{R})$.

In image processing, the variety of orthonormal bases which can be selected from Eq. (3) allows one to adaptively choose a basis for the image at hand. While it is theoretically possible to describe best bases of this type (assuming that the error is measured in $L_2(\mathbb{R})$), it is not efficient to implement these numerically. Instead one settles for a "good basis", selected according to some criteria, which can be found numerically.

Wavelet representation of an image

The utilization of wavelet bases in image compression begins with the representation of the image as a function f defined on the unit square $[0,1]\times[0,1]$. The function f is then expanded in its wavelet decomposition

$$f=\sum_{k}\sum_{j}\sum_{\lambda\in\Psi} c_{j,k,\lambda}\lambda_{j,k}. \tag{4}$$

This is particularly simple in the case of the standard Haar basis in which case $\phi = \chi_{[0,1]}$ and $\psi=H$ the Haar function. Suppose that the image is given be a $2^m\times 2^m$ array of pixel values (p_j), $j = (j_1,j_2)$, $0 \le j_1,j_2 < 2^m$. The function

$$f(x,j) = \sum_{j_1=0}^{2^m-1}\sum_{j_1=0}^{2^m-1} p_{(j_1,j_2)}\phi(2^m x-j_1)\phi(2^m x-j_2)$$

is piecewise constant on the unit square $[0,1]^2$ taking the value p_j on the square $[2^{-m} j_1, 2^{-m} (j_1+1)] \times [2^{-m} j_2, 2^{-m} (j_2+1)]$.

The wavelet representation of Eq. (4) results from a change from the basis

Eq.(3).

Wavelet compression of an image

Calculating the coefficients of a wavelet expansion of a function f that represents a mammographic image does not in any way compress the image. Compression is achieved by finding a wavelet approximation $\tilde{f}$ to f where $\tilde{f}$ has few nonzero wavelet coefficients, using the wavelet coefficients of f to determine the wavelet coefficients of $\tilde{f}$. This is generally accomplished by first *quantizing* the coefficients of f to obtain the wavelet coefficients of $\tilde{f}$, and then *coding* these coefficients. We briefly discuss these aspects of wavelet image compression in this section.

The approximation $\tilde{f}$ is a wavelet expansion of the same type as used for f, but one attempts to have few nonzero terms in $\tilde{f}$. Which terms one chooses will depend on how one intends to use the compressed image $\tilde{f}$. DeVore *et al* [16] proposed the selection of one of a family of error metrics given by L_p norms; if the pixels of f are denoted by p_j and the pixels of $\tilde{f}$ are denoted by $\tilde{p}_j$, then the L_p norm of the difference between f and $\tilde{f}$ is given by

$$\left(\frac{1}{n}\sum_{j}|p_j-\tilde{p}_j|^p\right)^{1/p} \tag{5}$$

where there are n pixels in the image. For large families of orthogonal and biorthogonal wavelets and for large ranges of p (in particular, always for $1 < p < \infty$), DeVore *et al* provided algorithms to make near-optimal choices of the wavelet coefficients used in $\tilde{f}$. We restrict attention in this summary to $p = 2$, the value of p that we found most useful for compressing digital mammograms.

When f and $\tilde{f}$ are expanded in terms of orthogonal wavelets, i.e.

$$f = \sum_k \sum_j \sum_{\psi\in\Psi} c_{j,k,\psi}\psi_{j,k} \text{ and } \tilde{f} = \sum_k \sum_j \sum_{\psi\in\Psi} \tilde{c}_{j,k,\psi}\psi_{j,k} ,$$

then a simple calculation shows that the least error between f and $\tilde{f}$ for a given number of nonzero coefficients is achieved when one chooses a parameter ε and sets

$$\tilde{c}_{j,k,\psi} = \begin{cases} c_{j,k,\psi}, & \|c_{j,k,\psi}\psi_{j,k}\|_{L_2} \geq \varepsilon , \\ 0, & otherwise. \end{cases}$$

It is impractical and unnecessary to keep all the digits of nonzero coefficients $\tilde{c}_{j,k,\psi}$; in fact, one chooses as compressed coefficients any $\tilde{c}_{j,k,\psi}$ that satisfy

$$\|(c_{j,k,\psi}-\tilde{c}_{j,k,\psi})\psi_{j,k}\|_{L_2}\leq\varepsilon.$$

One can apply *scalar quantization*, for which one finds a parameter q, determined by ψ and ε, called the *quantization interval* and determines integers $code_{j,k,\psi}$ such that for all j and k

$$\tilde{c}_{j,k,\psi} = code_{j,k,\psi}\, q\,.$$

It is these integers $code_{j,k,\psi}$ that are coded. Some researchers have also used a technique called *vector quantization* to quantize the coefficients of f. The same algorithm works with biorthogonal wavelets, and $p = 2$ can be replaced by any value of p strictly between 1 and ∞.

It is of interest to characterize the images that can be compressed well by this algorithm. One may wish to know how the error between $\tilde{f}$ and f decays as the number N of nonzero terms in the expansion of $\tilde{f}$ increases. DeVore *et al* showed that $\|f-\tilde{f}\|_{L_p}\leq CN^{-\alpha/2}$ if and only if, roughly speaking, f has α derivatives in L_q for $1/q = \alpha/2 + 1/p$. This result is of practical interest in that one may discover that a certain collection of images, in our case mammograms, have a characteristic smoothness that can be used to estimate the efficiency of the compression algorithm before it is applied.

One can apply the same algorithm to functions expanded in orthogonal hyperbolic wavelet bases. The same algorithm for choosing near-best coefficients $\tilde{c}_{j,k,\psi}$ applies, but there is not yet a complete characterization of images that can be compressed well using hyperbolic wavelet bases, as is described in the previous paragraph for the usual wavelet bases.

The use of adaptive bases introduces yet more flexibility in the choice of approximation $\tilde{f}$, since not only are the coefficients $\tilde{c}_{j,k,\psi}$ at our disposal, but we can choose which functions to include in our basis. Although heuristic methods achieve rather high compression (in that the error between f and $\tilde{f}$ is rather small for a given number of nonzero coefficients in $\tilde{f}$), there is not yet a near-optimal method for choosing either basis elements or coefficients. Of course, near-optimal methods are not really needed if heuristic methods yield better results than near-optimal methods applied to other types of wavelet bases.

We turn now to the question of how to code the integers $code_{j,k,\psi}$. One simple way of coding these integers is to order them in some arbitrary order (either all together, or separating them into groups determined by their dyadic level and possibly the function ψ) and applying an entropy coder to these long list of coefficients. Examples of entropy coders are Huffman coders and arithmetic coders. Huffman coders are simple, but since Huffman coders output at least one bit per code symbol, one cannot achieve high compression rates simply with Huffman coding when large fractions of coefficients are zero. Several researchers have achieved good results by combining Huffman coding of the nonzero coefficients with run-length coding of the zero coefficients. Since arithmetic coders can encode extremely common symbols in less than one bit, coders based on arithmetic coders can achieve relatively good efficiency without run-length coding of the zero coefficients of $\tilde{f}$. These techniques can be applied to any of the expansions mentioned above; however, since the total pool of possible bases elements for the adaptive basis is three times as large as for the other expansions, entropy coding of lists of zero and nonzero coefficients together requires more bits to encode each nonzero coefficient than with the other bases.

Recently several researchers have achieved good results by exploiting the natural tree-structure of the coefficients in the multiresolution expansion of f .[24],[25] These so-called *zero-tree* coders exploit the oft-found spatial coherence of nonzero coefficients in f at the same multiresolution level and between levels. These algorithms apply equally well to orthogonal and biorthogonal wavelet expansions, but not immediately to hyperbolic wavelet bases or to adaptive wavelet bases. There is some hope that information about the adaptive basis selection algorithm could be passed on to a tree-structured coder to achieve good compression results for adaptive wavelet bases.

III. EVALUATION PROCEDURES

Images

Eight screen/film mammograms were considered for a first evaluation test of the compression algorithms. They all contained one biopsy-proven malignant cluster of calcifications embedded into parenchymal tissues of varying density. Calcification cases were selected because, by nature, they are more appropriate for studying the effects of lossy compression algorithms on the preservation of very small indications of cancer. All mammograms were digitized at a resolution of 105 μm/pixel and a dynamic range of 12 bits per pixel (4096 grey levels) with a

105 μm/pixel and a dynamic range of 12 bits per pixel (4096 grey levels) with a DBA digitizer (Melbourne, FL). The optical density range selected for the digitization was different for each mammogram and was determined from the optical densities of the brightest and the darkest spots on the films. This allowed the maximization of the intensity differences between the various breast features and thus maximum separation in pixel values between the microcalcifications and the surrounding tissues. At this point, sections of the images (512×512 pixels) containing the calcification cluster were used for compression. So the number of image pixels n in Eq. (5) is 262,144. This dataset was also part of our initial study on mammogram compression where only Haar wavelets were used.[11] Relative comparisons were thus possible.

Evaluation criteria

A traditional and widely accepted way of evaluating medical image quality has been the subjective visual examination by an observer or ROC analysis. The logistics of this approach, however, are not suitable for algorithm discrimination and preliminary comparative studies among a variety of techniques. Quantitative criteria are necessary for this purpose. In the signal processing area, several quantitative measures are proposed for the evaluation of the quality of compressed data with the most common being the mean square error (MSE) and its variants between the original and the compressed data. We have tested several L_p metrics in this study and the L_2 error metric (MSE) was finally selected for comparisons. This is defined by

$$\|f-\tilde{f}\|_2 := \left[\frac{1}{n}\sum_i |p_i-\tilde{p}_j|^2\right]^{\frac{1}{2}}$$

where the pixel coordinate $j=(j_1, j_2)$ with $1 \leq j_1, j_2 \leq 512$. There is, in general, poor correlation between the MSE and the visual, subjective quality measures but it is a criterion commonly used in signal processing that also facilitates the decision of the optimal quantization strategy in compression.[1],[26] We complement the MSE calculation by the difference image, which is determined by subtracting the compressed image from the original. The appearance of any structure in the difference data indicates loss of information from the original data due to the compression method. A featureless difference image with unstructured noise patterns indicates compression with minimum information loss.

IV. RESULTS AND DISCUSSION

The images were compressed in L_2 with the wavelet methods of Table 1 with compression rates from 5:1 to 195:1 using an increment of 5; compression rates were calculated in terms of the nonzero wavelet coefficients. Note that the compression and reconstruction time of a 512×512 pixel image took about 10 seconds on a Sun SPARCstation 10. Adaptive wavelet compression runs at speeds inversely proportional to the selected compression ratio; for 50:1 compression, the current algorithm required about 50 seconds of computation time.

The MSE was calculated at every compression step and plots were generated of the MSE versus the compression rate for every method. A representative studied case is shown in Fig. 1. This mammogram contains numerous calcifications (>40) compactly clustered in a 1.5×1.5 cm^2 area (about 140×140 pixels). The pixel values along an image row, indicated by arrow, are plotted at the bottom of the figure to indicate the signal smoothness and continuity prior to compression and reconstruction. Figure 2 shows these curves for the mammogram of Fig. 1. Only the best six wavelet methods are presented here.

The relative ranking of these six methods differs from one image to the other although not significantly; their performances are for most images comparable and overlapping. They consistently outperform, however, all other wavelet methods of Table 1 for the eight images tested in this study. According to Fig. 2, hyperbolic Daubechies-4 basis is the method of choice at any compression rate. It is followed by the Biorthogonal-1-3, the Daubechies-2, the adaptive Haar, the hyperbolic Haar, and the standard Haar. The hyperbolic Daubechies-4 basis showed in general the smallest MSE for rates up to 100:1 approximately. In some images, a sharp increase of the MSE was observed for rates greater than 100:1 and this method was outperformed by the Biorthogonal-1-3 and the adaptive Haar bases. A similar change was observed for the standard Daubechies-2 basis. Figure 3 presents an example of this behavior for one of the digital mammograms in our set. Note that the rate of change of the MSE was image dependent and a slightly different behavior was observed for each studied case. There may be a pattern for this behavior related to parenchymal tissue densities and degree of uniformity but a larger dataset need to be studied for definite conclusions.

The compression results from the hyperbolic Daubechies-4, the

Biorthogonal-1-3, the hyperbolic Haar and the adaptive Haar bases are presented in Figs. 4-7 respectively for the mammogram of Fig. 1. The same compression ratio of 50 to 1 is used for these displays. As in Fig. 1, the pixel values of the same image row are plotted on the compressed data to indicate the effect of compression on signal smoothness and continuity. The same scale is used in all plots. These results can be compared with those obtained previously from the regular Haar wavelets.[1] As expected, Daubechies (hyperbolic or standard) and Biorthogonal wavelets result in smoother images than the Haar wavelets. However, the hyperbolic and adaptive Haar bases, although not entirely avoiding the so-called "block-effect", are significantly better than the standard Haar wavelets in terms of signal smoothness.

The visual evaluation of the quality of the compressed images by an expert mammographer indicated that images compressed with hyperbolic and adaptive Haar bases have better contrast than the smoother wavelets from a global point of view. Actually, in two of the cases, for compressions up to 25:1, the compressed reconstructed images were rated as having better contrast than the original. The Biorthogonal-1-3 and hyperbolic Haar compressions, however, seem to preserve better locally the sharpness and the relative intensities of the calcification signals as well as the signal variations due to changes in parenchymal tissue density. At very high compression rates (greater than 100:1), image degradations caused by the smoother wavelets were of course more pronounced. They were manifested as losses of edge sharpness, contrast, or calcification morphology and they were occasionally worse than the degradations caused by the hyperbolic or adaptive Haar wavelets. This effect appears in the difference images where more structured noise is observed in the case of the smoother wavelets particularly in the area of the calcification cluster. In contrast, Haar wavelets, standard, hyperbolic, or adaptive, show more blocking artifacts but their difference images exhibit a more random structure.

V. CONCLUSIONS

The purpose of this work was to investigate cost-effective means for the storage, retrieval, and transmission of digital mammograms, as required for efficient computer assisted diagnosis in central and remote locations. Image compression is an essential element of such means offering realistic solutions within a cost-effective mammography screening framework.

Based on the results of this study and our previous experience, we can conclude that wavelet-based methods are very promising for *visually lossless* digital mammogram compression at high rates. Wavelets such as hyperbolic Daubechies-4 and Biorthogonal-1-3 offered, as expected, higher smoothness in the compressed reconstructed images than the previously used Haar wavelets. The latter, however, still offer the advantage of higher contrast combining enhancement and compression processes in one step. New groups of wavelet bases have emerged from this study with significant theoretical advantages and very promising for medical image compression, namely the hyperbolic and adaptive wavelet bases. Only the Haar family was implemented in the adaptive mode but other wavelets are under investigation. Considering the variety of images encountered in mammography and the fact that the performance of a compression algorithm depends significantly on the contents of a mammogram, an adaptive selection of the optimum wavelet basis using individual image characteristics seems a promising approach to the compression of such data.

The MSE calculations offered a means for a first discrimination among the various wavelet compression techniques studied here although the MSE criterion is not sufficient for a full evaluation of the final candidates. Based on MSE values, however, complemented by difference images and visual evaluation of the quality of the compressed data by a radiologist, we were able to select the best methods for further clinical applications and investigations.

Table 1: Wavelet families, base types, and quantization strategies studied for digital mammogram compression.

Wavelet	Types of Bases	Quantization
Haar	Regular	All L_p
	Hyperbolic	All L_p
	Adaptive	L_2
Daubechies 2, 3, 4, ...	Regular	All L_p
	Hyperbolic	All L_p
Biorthogonal-1-3 Biorthogonal -1-3" etc	Regular	All L_p
	Hyperbolic	All L_p

REFERENCES

1. **Lucier BJ, Kallergi M, Qian W, DeVore RA, Clark RA, Saff EB, and Clarke LP**: Wavelet compression and segmentation of digital mammograms. *J Dig Imag* 6(1):1-13, 1994.

2. **Chan HP, Vyborny CJ, MacMahon J, et al**: Digital mammography ROC studies of the effects of pixel size and unsharp-mask filtering on the detection of subtle calcifications. *Invest Radiol* 22:581-589, 1987.

3. **Davies DH and Dance DR**: The automatic detection of subtle calcifications in high resolution, high quality mammograms. *Phys Med Biol* 37:1385-1390, 1992.

4. **Bassett LW and Gambhir S**: Breast imaging for the 1990s. *Seminars in Oncology* 18:80-86, 1991.

5. **Wilson DL**: Compression for radiological images. *SPIE* 1654:130-139, 1992.

6. **Kuduvalli GR and Rangayyan RM**: Performance analysis of reversible image compression techniques for high-resolution digital teleradiology. *IEEE Trans Med Imag* 11:430-445, 1992.

7. **Lelewer D and Hirshberg D**: Data compression. *ACM Computing Surveys* 19:261-296, 1991.

8. **Sayre J, Aberle DR, Boechat MI, Hall TR, Huang HK, Ho BK, Kashfian P, Rahbar G**: Effect of data compresion on diagnostic accuracy in digital hand and chest radiographs. *SPIE* 1653: 232-240; 1992.

9. **MacMahon H, Doi K, Sanada S, et al**: Data compression: effect of diagnostic accuracy in digital chest radiography. *Radiology* 178:175-179, 1991.

10. **Chen J and Flynn MJ**: The effect of block size on image quality for compressed chest radiographs. *SPIE* 1653:252-260, 1992.

11. **Aberle DR, Gleeson F, Sayre JW, Brown K, Batra P, Young DA, Stewart BK, Ho BKT, and Huang HK**: The effect of irreversible image compression on diagnostic accuracy in thoracic imaging. *Invest Radiol* 28:398-403, 1993.

12. **Qian W, Clarke LP, Kallergi M, Li HD, Clark RA, and Silbiger ML**: Adaptive order statistic filtering and wavelet transform for feature enhancement in mammography. *Proc IEEE EMB 15th Annual Conf*, October 28-31, San Diego, California; pp. 62-63, 1993.

13. **Qian W, Clarke LP, Kallergi M, et al.** Tree-structured nonlinear filter and wavelet transform for microcalcification segmentation in mammography. *IS&T/SPIE Annual Symposium on Electronic Imaging, Science & Technology*, San Jose, California; January 31 - February 5, Proceedings, 1993.

14. **Richardson WB**: Wavelets packets applied to mammograms. *IS&T/SPIE Annual Symposium on Electronic Imaging, Science & Technology*, San Jose, California; January 31 - February 5, Proceedings, 1993.

15. **Magnin IE, Baudin O, Baskurt A, et al**: An adaptive coding algorithm for a mammogram image database. *IS&T/SPIE Annual Symposium on Electronic Imaging, Science & Technology*, San Jose, California; January 31 - February 5, Proceedings, 1993.

16. **DeVore R, Jawerth B, and Lucier B**: Image compression through wavelet transform coding. *IEEE Trans Inf Theory* 38:719-746, 1992.

17. **DeVore R**: Image processing with hyperbolic and adaptive wavelet bases. *Proc Chamonix Conf on Curves and Surfaces*, 1994.

18. **Mallat SG:** Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$. *Trans Amer Math Soc* 315:69-87, 1989.

19. **Meyer Y:** *Ondelettes et opérateurs*. Volumes I, II, and III (Hermann ed.), Paris, 1990.

20. **Daubechies I:** Ten lectures on wavelets. *SIAM*, 1992.

21. **Cohen A, Daubechies I, and Feauveau JC:** Biorthogonal base of compactly supported wavelets. *Commun Pure and Applied Math* XLV:485-560, 1992.

22. **Vetterli M and Herley C:** Wavelets and filter banks: Theory and design. *IEEE Trans Signal Proc* 40(9):2207-2232, 1992.

23. **Daubechies I:** The wavelet transform: A method for time-frequency localization. Chapter 8 of *Advances in Spectrum Analysis and Array Processing: vol. 1*. (Haykin S. ed.) Prentice Hall, Englewood Cliffs, NJ, 366-417, 1991.

24. **Shapiro JM:** Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans Signal Proc* 41:3445-3462, 1993.

25. **Orchard MT and Ramchandran K:** An investigation of wavelet-based image coding using an entropy-constrained quantization framework. *Proc DCC '94: Data Compression Conference, IEEE Computer Society Press* (Storer JA and Cohn M eds.), Snowbird, Utah, March 1994, p. 341-350, 1994.

26. **Eskicioglu AM and Fisher PS.** A survey of quality measures for gray scale image compression. *Proc. Space and Earth Science Data Compression Workshop, Snowbird, April 1; NASA Conf. Publication* 3191:49-61; 1993.

FIGURES

1. Section of unprocessed digital mammogram (512×512 pixels) with a cluster of more than 40 calcifications indicated by arrow; the pixel values at a row indicated by the same arrow are plotted at the bottom of the image.

2. Plots of the MSE or L_2-error metric calculations versus compression rate (number of nonzero wavelet coefficients) for the mammogram of Fig. 1; the results of the best six wavelet methods are plotted. The compression rate was changed in increments of five.

3. Plots of the MSE or L_2-error metric calculations versus compression rate (number of nonzero wavelet coefficients) for a second mammogram from the studied set. The six wavelet methods with the smallest MSE are plotted. Note that the methods are the same as in Fig. 2 but their rank changes as the compression rate increases.

4. Hyperbolic Daubechies-4 wavelet compression (50:1) of the mammogram in Fig. 1; the values of the pixels at the same row as in Fig. 1 (indicated

by arrow) are plotted at the bottom of the image.

5. Biorthogonal-1-3 wavelet compression (50:1) of the mammogram in Fig. 1; the values of the pixels at the same row as in Fig. 1 (indicated by arrow) are plotted at the bottom of the image.

6. Hyperbolic Haar wavelet compression (50:1) of the mammogram in Fig. 1; the values of the pixels at the same row as in Fig. 1 (indicated by arrow) are plotted at the bottom of the image.

7. Adaptive Haar wavelet compression (50:1) of the mammogram in Fig. 1; the values of the pixels at the same row as in Fig. 1 (indicated by arrow) are plotted at the bottom of the image.

AUTHORS' BIOGRAPHICAL PARAGRAPHS

Zesheng Yang received his M.S. degree in Operations Research from the College of William & Mary in 1990. Currently he is working towards the Ph.D. degree in Mathematics under the supervision of Professor Ronald A. DeVore at the Department of Mathematics at the University of South Carolina. His current research interests include wavelet theory and its application to medical image compression and processing.

Maria Kallergi is Assistant Professor of Radiology at the University of South Florida and a Member of the H. Lee Moffitt Cancer Center and Research Institute. She received a B.S. degree in Physics from the University of Athens, Greece, in 1982; a D.E.A. degree in Energy Physics from the Institut National Polytechnique de Grenoble, France, in 1983; and M.S. and Ph.D. degrees in Materials science from USF in 1987 and 1990, respectively. After completion of her postdoctoral studies on medical physics and imaging at the H. Lee Moffitt Cancer Center and Research Institute, she joined the faculty of the Department of Radiology at USF continuing research in the same area.

Ronald A. DeVore is the Robert L. Sumwalt Professor of Mathematics at the University of South Carolina. He received his Ph.D. degree from Ohio State University in 1967 and has been at the University of South Carolina since 1977. He has published over 80 research articles and three research monographs in the areas of approximation theory, harmonic analysis, nonlinear partial differential equations, numerical analysis, and image processing. His recent research interests center on fast wavelet based algorithms for image processing and numerical methods for nonlinear partial differential equations.

Bradley J Lucier received the Honours B.Sc. degree in Mathematics from the University of Windsor, Ontario, and the S.M. and Ph.D. degrees in Applied Mathematics from the University of Chicago in 1976, 1978, and 1981, respectively. He has held visiting positions at the Mathematics Research Center, the University of Wisconsin, and at the Institute of Mathematics and its Applications, the University of Minnesota, and is currently Professor of

Mathematics and Computer Science at Purdue University. His research interests include applications of wavelets to data compression, image processing, computer aided geometric design and numerical methods for partial differential equations, as well as adaptive and moving-grid numerical methods for time-dependant partial differential equations.

Wei Qian is currently an Assistant Professor at the Department of Radiology, College of Medicine, University of South Florida and a member in Residence of the H. Lee Moffitt Cancer Center and Research Institute. He has been a Postdoctoral Research Associate at the Department of Electrical Engineering at the University of Notre Dame in 1990/91 and a Postdoctoral Research Associate at the Department of Radiology at the University of South Florida in 1991/92. He is an author of over 30 publications in the area of image processing and medical imaging. Dr. Qian's research interests include image processing, pattern recognition and computer vision as applied to medical imaging and computer assisted diagnosis.