Compression of 3-D echocardiographic images using a modified 3-D set-partitioning-in-hierarchical-trees algorithm based on a 3-D wavelet packet transform

Xiyi Hang
California State University, Northridge
Department of Electrical and Computer Engineering
18111 Nordhoff Street
Northridge, California 91330
E-mail: xhang@csun.edu

Neil L. Greenberg
The Cleveland Clinic Foundation
Department of Cardiovascular Medicine
Cleveland, Ohio 44195

Yuan F. Zheng
The Ohio State University
Department of Electrical & Computer Engineering
Columbus, Ohio 43210

James D. Thomas
The Cleveland Clinic Foundation
Department of Cardiovascular Medicine
Cleveland, Ohio 44195

Abstract. *An efficient compression strategy is indispensable for 3-D digital echocardiography, which can provide more accurate diagnostic information than 2-D echocardiography without geometric assumption, but unfortunately requires a huge storage space. We describe a new set partitioning in hierarchical trees (SPIHT) algorithm that is based on the wavelet packet transform and use it to compress 3-D echocardiographic images. The new algorithm originates from the well-known SPIHT algorithm that is based on a wavelet transform but outperforms it for the compression of 3-D echocardiographic images. Experimental results are presented to verify the effectiveness of the new SPIHT algorithm. Results show that the new algorithm achieves a very high compression ratio while retaining good quality 3-D echocardiographic images. © 2006 SPIE and IS&T.* [DOI: 10.1117/1.2194467]

1 Introduction

Real-time three-dimensional (RT3D) echocardiography is becoming feasible due to the advancement of computer and visualization technologies. The heart is a moving 3-D complex structure, optimal examination of which can be achieved with the aid of RT3D echocardiography, enabling more accurate quantification of the left ventricle without geometrical assumptions. The obtained parameter, the ejection fraction, is used in diagnosis of the ischemic and congenital heart diseases. Another important use of RT3D echocardiography is the visualization of valve morphology, effectively assisting the diagnosis of valve disease.

Digital storage and review of the RT3D echocardiographic images plays an important role in its successful application. However, the huge storage requirement and limited networking bandwidth are obstacles to any meaningful use of the method. For example, a 1-s loop (30 frames) of medium-size 3-D images, each with $196 \times 176 \times 196$ voxels and produced by the Philips Live 3D Echo System, which is used in this study, requires nearly 200 Mbytes of storage space. The space required by a large medical center such as the Cleveland Clinic Foundation for storing the results of examinations daily would be more than 1 Tbytes.

Many efforts have been made to reduce the required storage space, and a number of approaches have been developed for the compression of ultrasonic images including echocardiographic images. In general, these approaches can be divided into two categories: standalized and specialized algorithms.

JPEG (Ref. 1), which is a standard for still image compression, was examined in Refs. 2 and 3 for its impact on
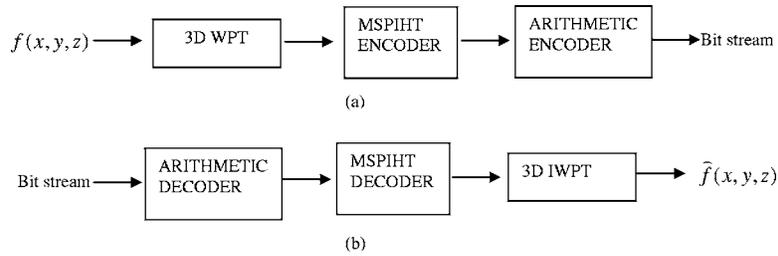
Fig. 1 MSPIHT algorithm diagrams of (a) encoding and (b) decoding.

the clinical interpretation of 2-D digital echocardiography. It was concluded that the diagnostic information is retained at compression ratios as high as 20:1. The JPEG standard has been accepted by the American Society of Echocardiography under the Digital Imaging and Communications in Medicine (DICOM) formatting standard,[4] and is utilized by all major echocardiograph manufacturers. In Ref. 5, the new image compression standard[6] JPEG2000 is applied to the compression of ultrasonic images.

Motion picture compression standards MPEG-1 and MPEG-2 (Ref. 7) were clinically validated at compression ratios up to 100:1 in Refs. 8 and 9, respectively. Furthermore, the new standard MPEG-4 (Ref. 10) was also investigated in Ref. 11. The authors demonstrated the content-based coding capability of MPEG-4 by creating video objects and encoding each of them as a separate bit stream.

As for the specialized compression algorithms, there are three types: segmentation-based, model-based, and transform-based methods. The segmentation-based method aims to preserve the area with important diagnostic information and highly compress the background area. In Ref. 12, a wavelet transform is first applied to the echocardiographic image, and then the information of the region of interest (ROI) is obtained by segmentation. Finally, higher importance is assigned to the segmented anatomical structure by keeping more wavelet coefficients. In Ref. 13, vector quantization with different codebooks is applied to different segmented areas. A larger codebook is created for an important area, and vice versa.

Model-based methods attempt to keep the speckle patterns in the ultrasound image, treating an ultrasound image as the combination of two components: image structure and speckle texture. In Ref. 14, a wavelet-denoising strategy is used to separate the two components. A general-purpose image compression strategy such as JPEG is applied for the compression of the structure component, while the speckle texture component is compressed by means of texture synthesis and first-order spatial statistics. In Ref. 15, the separation of the two components is realized by further decomposing the selected high frequency subbands. The image structure is compressed by a wavelet-based lossy method, and the speckle part is compressed using a lossless method.

A transform-based compression algorithm consists of three steps: transform, quantization, and entropy coding, while wavelet transform is the most often used transform method. The embedded zero-tree wavelet (EZW) algorithm[16] was used in Ref. 17 for the compression of 2-D echocardiographic images. And it was extended to three dimensions in Ref. 18 to compress 3-D echocardiographic images. Set partitioning in hierarchical trees[19] (SPIHT), which is also wavelet-based, along with its 3-D extension has been used for the compression of 2-D echocardiographic image and video.[20,21] In Ref. 22, a space-frequency segmentation (SFS) method[23] was used that builds a balanced wavelet packet tree via a tree-pruning algorithm. It is shown that SFS performs better than SPIHT at the cost of a much longer execution time. Integer-wavelet-transform-based methods are also proposed.[24,25]

Both EZW and SPIHT algorithms, just mentioned, are wavelet zero-tree algorithms that utilize quad-trees to relate wavelet coefficients across scales to take advantage of the decay of the coefficients. SPIHT outperforms EZW with a so-called set-partitioning technique.

In this paper, a new algorithm similar to SPIHT in spirit is proposed for the compression of 3-D echocardiographic images. Two main features characterize the novelty of the algorithm. First, the new SPIHT algorithm is based on a wavelet packet transform, while the original SPIHT is based on a wavelet transform. It is known that wavelet packets can provide a better localization in the high-frequency domain than wavelets, because the Heisenberg box could be narrower with wavelet packets but is usually wider with high-frequency wavelets.[26] Also, highly oscillatory speckle patterns that translate into high-frequency wavelets are important components of ultrasonic images. Second, a new wavelet packet zero-tree is defined that is
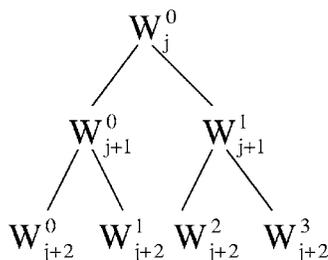
Fig. 2 Binary wavelet packet tree.
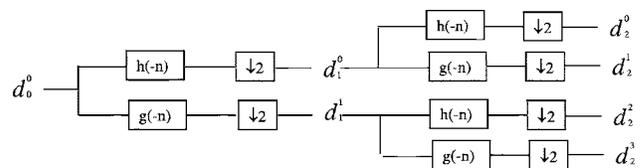
Fig. 3 Full-scale two-level WPT.
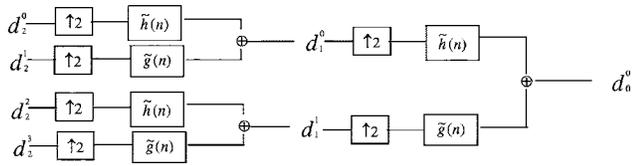
Fig. 4 Full-scale two-level inverse WPT.

(a)

(b)

Fig. 6 (a) Index numbers of 16 subbands in a 2-D two-level WPT and (b) the SOT among these 16 subbands.

totally different from the spatial orientation tree used in SPIHT. We call the new algorithm modified SPIHT (MSPIHT).

Figure 1 shows diagrams of the encoder and decoder procedures of the MSPIHT algorithm. The rest of the paper is organized as follows. Section 2 reviews the wavelet packet transform, Sec. 3 focuses on the MSPIHT algorithm, Sec. 4 gives the coding results and discussion, and the conclusion and future work follow in the final section.

2 Wavelet Packet Transform

The wavelet packet transform (WPT) introduced in Ref. 27 enables a finer and more adjustable analysis of the signal at high frequencies as compared to the wavelet transform. In this section, the theory of the WPT is reviewed. For the benefit of further discussion, see Ref. 28. Denoting the family of the wavelet packet bases as $\{\psi_j^m(t - 2^j n)\}$ of wavelet packet space W_j^m, we have

$$\psi_j^0(t) = 2^{-j/2}\phi(2^{-j}t), \tag{1}$$

where ϕ is the scaling function. Figure 2 shows the binary tree of the wavelet packet space. The recursive equations for the wavelet packet decomposition are

$$\psi_{j+1}^{2m}(t) = \sum_n h(n)\psi_j^m(t - 2^j n), \tag{2}$$

and

$$\psi_{j+1}^{2m+1}(t) = \sum_n g(n)\psi_j^m(t - 2^j n), \tag{3}$$

where $h(n)$ and $g(n)$ are the low-pass scaling filter and the high-pass wavelet filter for analysis, respectively, with

$$g(n) = (-1)^{1-n}h(1-n). \tag{4}$$

Let $f(t) \in W_0^0$ be a continuous signal, the wavelet packet coefficients can be denoted as

Fig. 5 Structure of the SOT in a 2-D two-level wavelet transform.

$$d_j^m = \langle f(t), \psi_j^m(t - 2^j n)\rangle. \tag{5}$$

Equations (2), (3), and (5) lead to the following wavelet packet decomposition in the discrete domain:

$$d_{j+1}^{2m}(n) = d_j^m * \bar{h}(2n) \quad \text{and} \quad d_{j+1}^{2m+1}(n) = d_j^m * \bar{g}(2n), \tag{6}$$

where

$$\bar{h}(n) = h(-n) \quad \text{and} \quad \bar{g}(n) = g(-n). \tag{7}$$

The algorithm in Eq. (6) can be realized by the wavelet packet analysis filter banks, iterating both the high-pass wavelet and the low-pass scaling function branches of the Mallat algorithm tree.[29] Figure 3 shows a full-scale two-level WPT, where $\underline{d}_0^0$ stands for the input discrete-time signal. Similarly, for reconstruction, one has

$$d_j^m(n) = d_{j+1}^{2m} * \tilde{h}(n) + d_{j+1}^{2m+1} * \tilde{g}(n), \tag{8}$$

where $\tilde{h}(n)$ and $\tilde{g}(n)$ are the low-pass scaling filter and the high-pass wavelet filter for synthesis, respectively. The wavelet packet synthesis filter banks can perform the computation of Eq. (8). Figure 4 shows the inverse WPT corresponding to the WPT shown in Fig. 3.

The 3-D WPT can be realized by three 1-D WPTs along *x*, *y*, and *z*, respectively. One concern for the 3-D transform is the order of the 1-D transform. For example, for the compression of video signal based on a 3-D wavelet transform, the order of the 1-D transform does affect the compression effect because the statistical property along the temporal direction is quite different from those along other two spatial directions.[30] For the compression of 3-D echocardiographic images, on the other hand, since the 3-D image has symmetric statistical properties along all the spatial directions, the order does not affect the compression effect. To handle the problem of convolution at the borders of the data field in the wavelet packet filter banks, the symmetric extension method[31] is used, which, for subband coding, has been shown to be superior to other methods such as zero padding and periodic extension. Furthermore, for the

definition of the spatial orientation tree among the wavelet packet coefficients described in Sec. 3, we use the full-scale binary WPT, as shown in Fig. 3. Although the best wavelet packet bases can be obtained by pruning the full binary tree using a binary search algorithm,[27] the corresponding binary wavelet packet tree is variable, which is not suitable for the definition of our fixed wavelet packet zero-tree.

3 MSPIHT Algorithm

The SPIHT algorithm is a very popular image compression scheme. It is a zero-tree-based wavelet codec, exploiting the fact that among the wavelet coefficients, if a point in the higher level of the pyramid is insignificant, its descendants are also insignificant with a high probability. The zero-tree used in SPIHT is called the spatial orientation tree (SOT) and is briefly explained here. Let $O(i,j,k)$ denote the set of coordinates of all offspring of the node (i,j,k), with $0 \le i \le M-1, 0 \le j \le N-1$, and $0 \le k \le P-1$ in a 3-D image with size $M \times N \times P$. The SOT among the wavelet coefficients of an L-level 3-D wavelet transform is built by the recursive definition of $O(i,j,k)$, given next.

IF $(0 \le i \le M/2^L-1)$ and $(0 \le j \le (N/2^L-1)$ and $(0 \le k \le P/2^L-1)$

    IF $(i\%2=0)$ and $(j\%2=0)$ and $(k\%2=0)$

        $O(i,j,k)=\{\};$

    ELSE

        IF $i\%2=0$

                $x_0=i;$

        ELSE

                $x_0=M/2^L+i-1;$

        IF $j\%2=0$

                $y_0=j;$

        ELSE

                $y_0=N/2^L+j-1;$

        IF $k\%2=0$

                $z_0=k;$

        ELSE

                $z_0=P/2^L+k-1;$

        $O(i,j,k)=\{(x_0,y_0,z_0),(x_0+1,y_0,z_0),(x_0,y_0+1,z_0),$
        $(x_0+1,y_0+1,z_0),(x_0,y_0,z_0+1),(x_0+1,y_0,z_0+1),$
        $(x_0,y_0+1,z_0+1),(x_0+1,y_0+1,z_0+1)\};$

ELSE IF $(i \le M/2-1)$ and $(j \le N/2-1)$ and $(k \le P/2-1)$

    $x_0=2*i;$
    $y_0=2*j;$
    $z_0=2*k;$
    $O(i,j,k)=\{(x_0,y_0,z_0),(x_0+1,y_0,z_0),(x_0,y_0+1,z_0),$
    $(x_0+1,y_0+1,z_0),(x_0,y_0,z_0+1),(x_0+1,y_0,z_0+1),$
    $(x_0,y_0+1,z_0+1),(x_0+1,y_0+1,z_0+1)\};$

ELSE

    $O(i,j,k)=\{\}.$

Figure 5 shows the SOT in a 2-D two-level wavelet transform.

The data generated by the SPIHT algorithm are three linked lists: the list of insignificant pixels (LIP), the list of significant pixels (LSP), and the list of insignificant sets (LIS). The lists are operated by three stages of SPIHT:

(a)

(b)
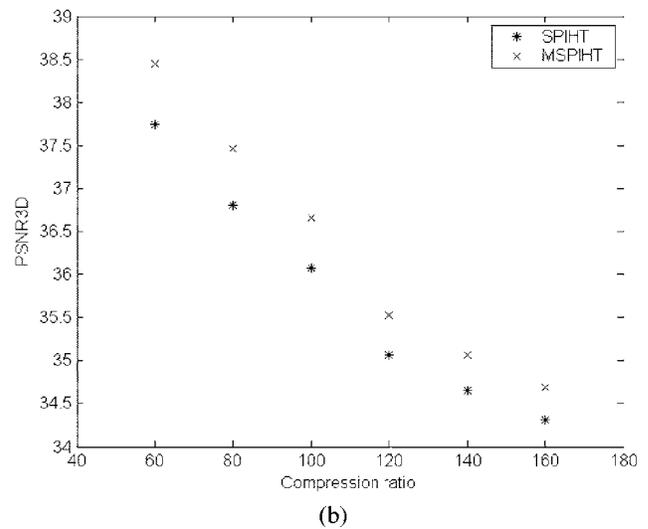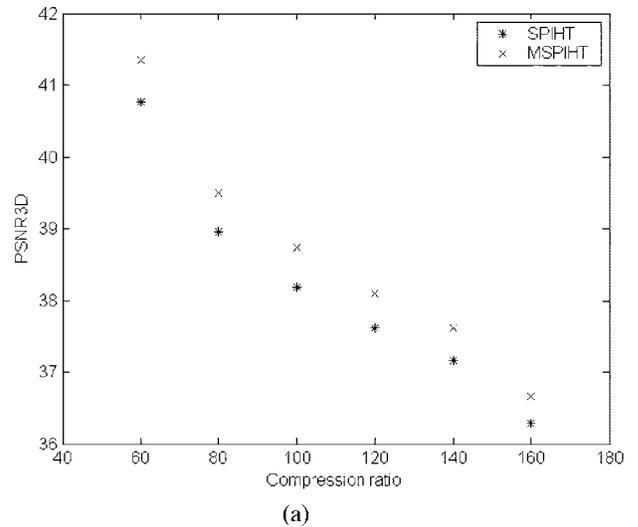
Fig. 7 PSNR3D values under different compression ratios when applying MSPIHT and SPIHT to (a) dataset 1 and (b) dataset 2.

initialization, sorting, and refinement. In the initialization stage, SPIHT sets the LSP as an empty list, adds every point in the highest level to the LIP, and puts only those points with descendants in the highest level to the LIS. At every sorting stage, corresponding to a threshold that is a power of 2 and octavely decreasing, every significant point in an LIP is moved into the LSP, and the sets in the LIS are partitioned into significant points, insignificant points, and smaller insignificant sets, which are moved into LSP, LlP, and LIS, respectively. During this stage, SPIHT also outputs the bits about the significance information and the parent-offspring relation of all the points based on the spatial orientation tree. Following each sorting stage, the refinement stage outputs the bits of significance for those points in LSP at the thresholds that are greater than the current threshold. The preceding process continues until the desired bit rate or image quality is reached. Encoding and decoding are executed along the same path based on the spatial orientation tree. More details can be found in Ref. 19.

Our 3-D MSPIHT is based on a 3-D WPT and has a
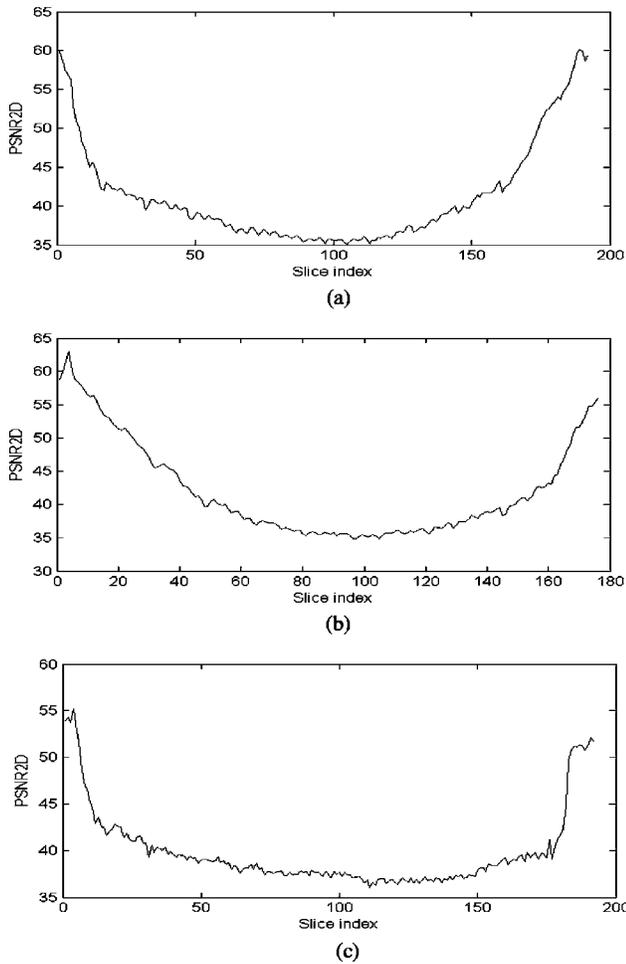
Fig. 8 PSNR2D values of the slice images of the compressed dataset 1 at the compression ratio of 100:1, along (a) *x*, (b) *y*, and (c) *z*, respectively.

Fig. 9 PSNR2D values of the slice images of the compressed dataset 2 at the compression ratio of 100:1, along (a) *x*, (b) *y*, and (c) *z*, respectively.

different SOT among wavelet packet coefficients. The rationale of defining the new SOT is, in spirit, very similar to SPIHT, which utilizes the SOT to define the spatial relationship on the hierarchical pyramid of wavelet subbands. And the new SOT is used to define the spatial relationship on the hierarchical pyramid of full-scale wavelet packet subbands. Let $\{d^I(i,j,k)\}$ be the 3-D subbands obtained by the 3-D *L*-level WPT described in Sec. 2, where *I* is the subband index; $I = 1 \cdots 8^L$; and *i*, *j*, and *k* are coordinates of the 3-D subbands. We denote the offspring of every node $d^I(i,j,k)$ in the new SOT as $O[d^I(i,j,k)]$. The new SOT is defined as

$$
O[d^I(i,j,k)]
$$
$$
= \begin{cases} \{d^b(i,j,k)\}\ b = 2 \cdots 8, & \text{if } I = 1 \\ \{d^{8I-b}(i,j,k)\}\ b = 0 \cdots 7, & \text{else if } 2 \leqslant I \leqslant 8^L - 1 \\ \{\}, & \text{otherwise.} \end{cases}
$$
$$
(9)
$$

Figure 6(a) shows the index numbers of 16 subbands in a 2-D full-scale two-level WPT, and Fig. 6(b) shows the SOT among these 16 subbands. For the 3-D case, obviously, the main difference between the new SOT and the original one lies in the following two facts: (1) in the new SOT, each voxel of the first subband, which has the lowest frequency component, has offspring, while in the original spatial orientation tree one-eighth of the voxels of the first subband do not have offspring; and (2) in the new SOT, all the offspring of one voxel are located in different subbands, while in the original SOT, they are located in the same subband.

The 3-D MSPIHT still retains the same data structure of lists LIP, LIS, and LSP, and the same three stages of processing (initialization, sorting, and refinement) as in SPIHT. However, the new algorithm utilizes the new SOT to traverse the coefficients. According to the definition of the new SOT, every point in the highest level of the 3-D wavelet packet coefficients, which is the first subband, has descendants in the initialization stage. Subsequently, each point of the first subband is put into LIP and LIS, respectively. SPIHT, however, puts only the points at the highest level of the wavelet coefficients with descendants in LIS. And the procedures of sorting and refinement are similar to SPIHT, except that a different SOT is involved during the execution. The MSPIHT encoder is given next along with some new notation.

- $D(i,j,k,I)$: set of all descendents of node (i,j,k,I);
- H: set of all the nodes in the first subband;
- $L(i,j,k,I) = D(i,j,k,I) - O(i,j,k,I)$;
- A node (i,j,k,I) in the LIS is of type A if it belongs to $D(i,j,k,I)$, and of type B if it belongs to $L(i,j,k,I)$;
- $S_n(T) = \begin{cases} 1, & \max_T\{|c(i,k,j,I)|\} \geq 2^n, \\ 0, & otherwise, \end{cases}$

where $c(i,j,k,I)$ is the wavelet packet coefficient at the node (i,j,k,I), and $(i,j,k,I) \in T$.

1. Initialization:
    1.1) Output n=$\lfloor \log_2(\max_{(i,j,k,I)}\{|c(i,j,k,I)|\}) \rfloor$;
    1.2) LSP={ };
    1.3) Add all the nodes in H to LIP;
    1.4) Add all the nodes in H to LIS as type A;
2. Sorting:
    2.1) for each node (i,j,k,I) in the LIP do:
        2.1.1) output $S_n(i,j,k,I)$;
        2.1.2) if $S_n(i,j,k,I)=1$, then
            . delete (i,j,k,I) from LIP;
            . add (i,j,k,I) to LSP;
            . output the sign of $c(i,j,k,I)$;
    2.2) for each node (i,j,k,I) in the LIS do:
        2.2.1) if the node is of type A, then
            . output $S_n(D(i,j,k,I))$;
            . if $S_n(D(i,j,k,I))=1$, then
                . for each node (l,m,n,P) of $O(i,j,k,I)$ do:
                    . output $S_n(l,m,n,P)$;
                    . if $S_n(l,m,n,P)=1$, then
                        . add (l,m,n,P) to LSP;
                        . output the sign of $c(l,m,n,P)$;
                  . else
                        . add (l,m,n,P) to LIP;
            . if $L(i,j,k,I)=\{\}$, then
                . delete (i,j,k,I) from LIS;
            . else
                . delete (i,j,k,I) from LIS;
                . $(i,j,k,I) \rightarrow$LIS, as type B;
        2.2.2) if the node is of type B, then
            . output $S_n(L(i,j,k,I))$;
            . if $S_n(L(i,j,k,I))=1$, then
                . for each node (l,m,n,P) of $O(i,j,k,I)$ do:
                    . add (l,m,n,P) to LIS as type A;
            . delete (i,j,k,I) from LIS;
3. Refinement:
    for each node (i,j,k,I) in the LSP do:
    . if $|c(i,k,j,I)|<2^{n+1}$, then
    . output the n-th MSB of $|c(i,j,k,I)|$;
4. Quantization-step update:
    . n−1→n;
    . go to Step 2.

The MSPIHT decoder executes along the same path as the encoder except that "output" should be changed to "input" and the wavelet packet coefficients are appropriately reconstructed. Finally, arithmetic coding[32] is used for entropy coding. Groups of $2 \times 2 \times 2$ voxels are kept together. Different context models, each with 2^m symbols, $m \in \{1,2,3,4,5,6,7,8\}$, are used to code the information in a group of 8 voxels.

4 Coding Results

The two-dimensional peak SNR (PSNR2D) and the 3-D PSNR (PSNR3D) are used to evaluate the performance of MSPIHT. They are defined as

$$PSNR2D = 10\log_{10}\frac{255^2 M \times N}{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[f(i,j)-\hat{f}(i,j)]^2}, \quad (10)$$

where $f(i,j)$ and $\hat{f}(i,j)$ represent the value of each pixel in the 2-D original image and reconstructed image, respectively, all with size $M \times N$, and

$$PSNR3D = 10\log_{10}\frac{255^2 M \times N \times P}{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{k=0}^{P-1}[f(i,j,k)-\hat{f}(i,j,k)]^2}, \quad (11)$$

where $f(i,j,k)$ and $\hat{f}(i,j,k)$ represent the value of each voxel in the 3-D original image and reconstructed image, respectively, all with size $M \times N \times P$.

We selected the suitable wavelet for the compression of 3-D echocardiographic image by testing a group of wavelets,[33] which are listed as follows:

1.   Daubechies orthogonal wavelets with number p of vanishing moments:

    $p = 1$ ('db2'),

    $p = 2$ ('db2'),

    $p = 3$ ('db3'),

    $p = 4$ ('db4'),

    $p = 5$ ('db5').

2.   Symmlets orthogonal wavelets with number p of vanishing moments:

    $p = 4$ ('sym4'),

    $p = 5$ ('sym5').

3.   Biorthogonal wavelets with number p of vanishing moments for synthesis, and number q of vanishing moments for analysis:

    $p = 1, q = 3$ ('bior1.3'),

    $p = 1, q = 5$ ('bior1.5'),
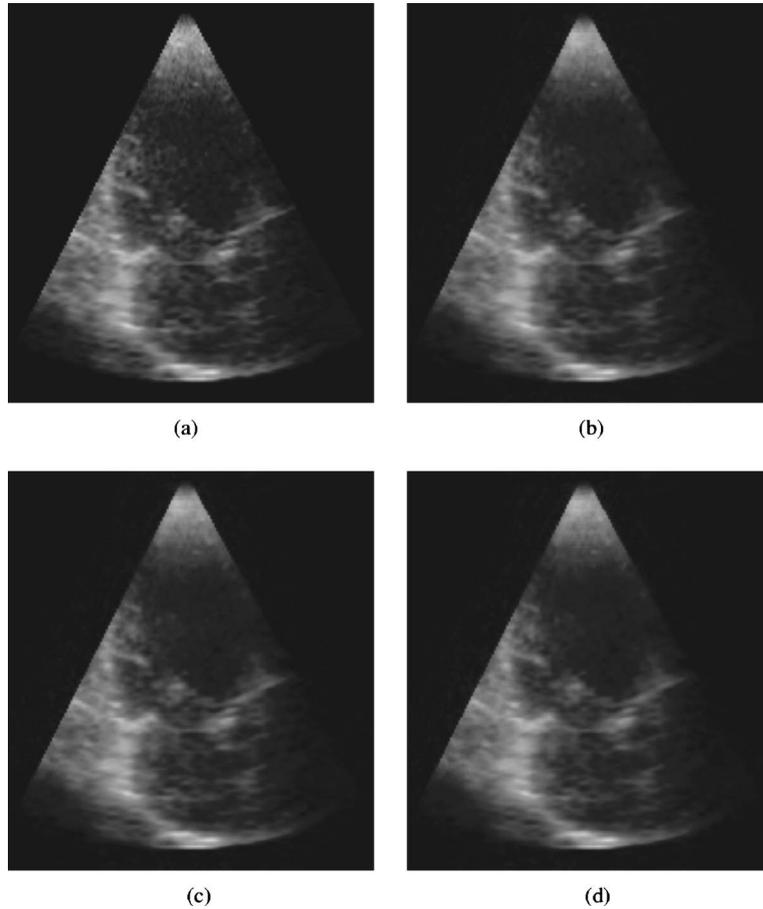
Fig. 10 Slice images at $x=95$ of dataset 1 and the corresponding compressed datasets under differential compression ratios: (a) original, (b) 100:1 (PSNR=35.89 dB), (c) 140:1 (PSNR=34.79 dB), and (d) 160:1 (PSNR=33.82 dB).

$$p = 2, q = 2 \ (\text{'bior2.2'}),$$

$$p = 2, q = 4 \ (\text{'bior2.4'}),$$

$$p = 3, q = 1 \ (\text{'bior3.1'}),$$

$$p = 3, q = 3 \ (\text{'bior3.3'}),$$

$$p = 4, q = 4 \ (\text{'bior4.4'}),$$

$$p = 5, q = 5 \ (\text{'bior5.5'}).$$

4. Reverse biorthogonal wavelets with number p of vanishing moments for synthesis, and number q of vanishing moments for analysis. This type of wavelet is obtained by means of exchanging the analysis part and the synthesis part of biorthogonal wavelets with number p of vanishing moments for synthesis, and number q of vanishing moments for analysis:

$$p = 1, q = 3 \ (\text{'rbior1.3'}),$$

$$p = 1, q = 5 \ (\text{'rbior1.5'}),$$

$$p = 2, q = 2 \ (\text{'rbior2.2'}),$$

$$p = 2, q = 4 \ (\text{'rbior2.4'}),$$

$$p = 3, q = 1 \ (\text{'rbior3.1'}),$$

$$p = 3, q = 3 \ (\text{'rbior3.3'}),$$

$$p = 4, q = 4 \ (\text{'rbior4.4'}),$$

$$p = 5, q = 5 \ (\text{'rbior5.5'}).$$

The mathematical characteristics of the preceding wavelets can be found in Ref. 33.

Our tests are performed on two randomly selected datasets: dataset 1 with a volume size of $192 \times 176 \times 192$ and a voxel size of $1.05 \times 1.05 \times 1.05 \ \text{mm}^3$, and dataset 2 with a volume size of $192 \times 176 \times 160$ and a voxel size of $1.01 \times 1.01 \times 1.01 \ \text{mm}^3$. Both datasets are acquired for the
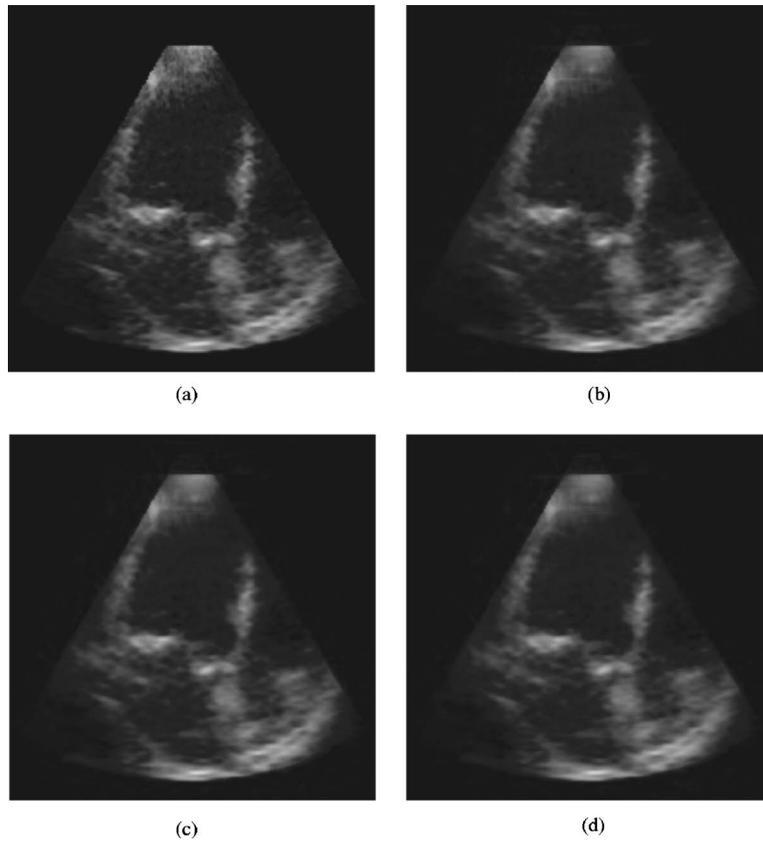
Fig. 11 Slice images at $y=80$ of dataset 1 and the corresponding compressed datasets under differential compression ratios: (a) original, (b) 100:1 (PSNR=35.59 dB), (c) 140:1 (PSNR=34.51 dB), and (d) 160:1 (PSNR=33.64 dB).

purpose of left ventricle quantification. Tables 1–3 show the PSNR3D values of the two datasets at a compression ratio of 80:1, using orthogonal wavelet filters, biorthogonal wavelet filters, and reverse biorthogonal wavelet filters, respectively. Notice that 'sym 5' shows the best performance among the orthogonal wavelet filters, 'bior4.4' among the biorthogonal wavelet filters, and 'rbio1.5' among the reverse biorthogonal wavelet filters, respectively. Further-

more, 'sym5' is superior to both 'bior4.4' and 'rbio1.5'. The other results presented in the remainder of this section are obtained by using 'sym5'.

The 3-D SPIHT has been studied extensively for compression of video signal. And an asymmetric-tree-based SPIHT algorithm has been proposed to take advantage of the asymmetric statistical properties of video signal.[30]

Table 1 The PSNR3D values (in decibels) of two datasets at the compression ratio of 80:1, using orthogonal wavelet filters.

Wavelet Filters	Dataset 1	Dataset 2
'db1'	36.86	34.72
'db2'	38.28	36.23
'db3'	38.45	36.45
'db4'	38.42	36.42
'db5'	38.22	36.10
'sym4'	39.23	37.18
'sym5'	39.49	37.46

Table 2 The PSNR3D values (in decibels) of two datasets at the compression ratio of 80:1, using biorthogonal wavelet filters.

Wavelet Filters	Dataset 1	Dataset 2
'bior1.3'	36.86	34.25
'bior1.5'	35.92	33.90
'bior2.2'	37.86	35.49
'bior2.4'	38.24	35.89
'bior3.l'	30.31	28.05
'bior3.3'	35.04	33.39
'bior4.4'	39.15	37.21
'bior5.5'	37.91	36.16
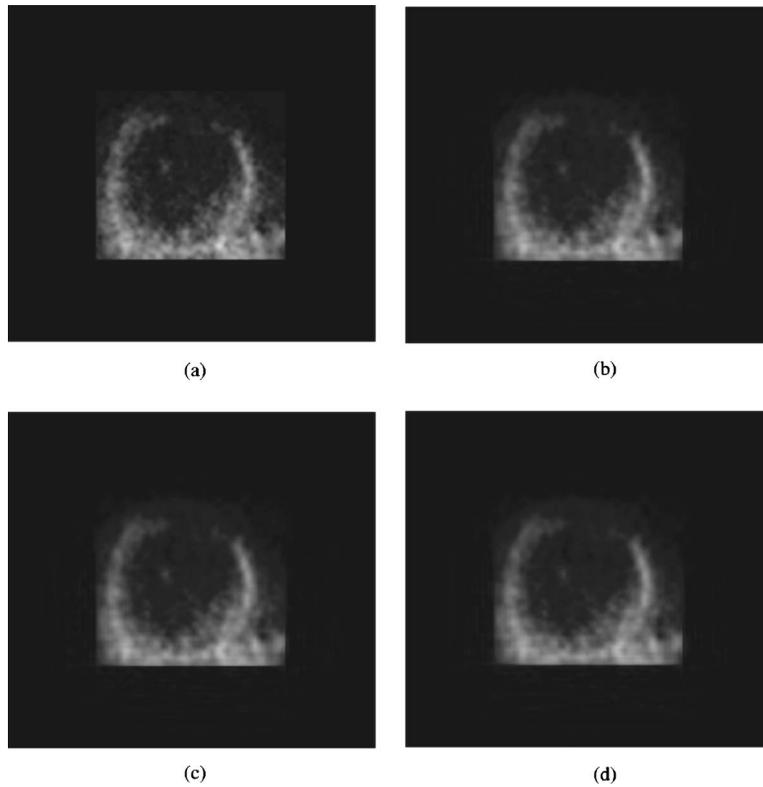
Fig. 12 Slice images at z=85 of dataset 1 and the corresponding compressed datasets under differential compression ratios: (a) original, (b) 100:1 (PSNR=37.41 dB), (c) 140:1 (PSNR=36.33 dB), and (d) 160:1 (PSNR=35.46 dB).
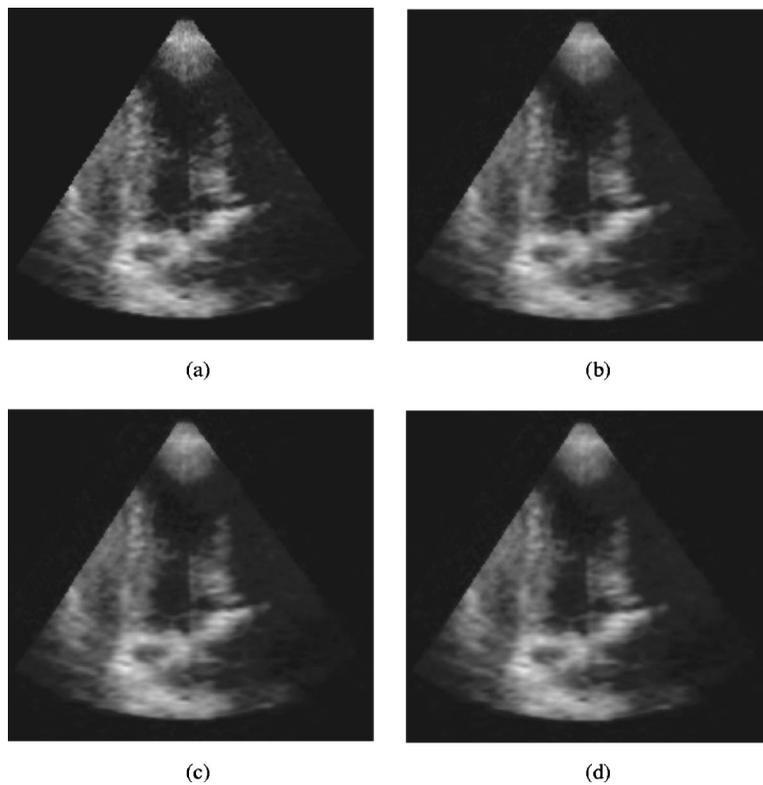
Fig. 13 Slice images at x=95 of dataset 2 and the corresponding compressed datasets under differential compression ratios: (a) original, (b) 100:1 (PSNR=34.83 dB), (c) 140:1 (PSNR=33.21 dB), and (d) 160:1 (PSNR=32.86 dB).

Table 3 The PSNR3D values (in decibels) of two datasets at the compression ratio of 80:1, using reverse biorthogonal wavelet filters.

Wavelet Filters	Dataset 1	Dataset 2
'rbior1.3'	38.89	36.82
'rbior1.5'	39.02	37.07
'rbior2.2'	33.95	32.46
'rbior2.4'	37.09	34.05
'rbior3.1'	9.40	8.57
'rbior3.3'	25.25	23.96
'rbior4.4'	38.90	36.77
'rbior5.5'	38.74	36.38

Since a 3-D echocardiographic image has symmetric statistic properties, we chose a symmetric-tree-based 3-D SPIHT for comparison, i.e., the 3-D extension of the original 2-D SPIHT. Figures 7(a) and 7(b) show the PSNR3D values under different compression ratios when the 3-D MSPIHT and the 3-D SPIHT are applied to datasets 1 and 2, respectively. The 3-D MSPIHT shows better performance than the 3-D SPIHT for both datasets.

Table 4, parts (a) to (c), show the average PSNR2D values of slice images obtained by setting x, y, and z, respectively, to different constants, which are taken from 1 to the number of the spatial resolution of x, y, and z, respectively. When the compression ratio increases, the mean value decreases while the standard deviation increases.

Figures 8(a) to 8(c) show the PSNR2D values of the slice images of compressed dataset 1 at the compression ratio of 100:1, along x, y, and z, respectively. Similar information is shown in Fig. 9 for dataset 2. The PSNR2D values of the slice images in the middle are much smaller than those of the slice images at the beginning and at the end. The reason lies in the fact that most of the content of the slice images at the beginning and at the end is background information.

Figures 10–15 show the slice images $f_1(95, y, z)$, $f_1(x, 80, z)$, $f_1(x, y, 85)$, $f_2(95, y, z)$, $f_2(x, 80, z)$, and $f_2(x, y, 85)$ including their corresponding compressed images under compression ratios of 100:1, 140:1, and 160:1, respectively. And one can see that the MSPIHT approach produces a high compression ratio while maintaining good image quality.

In the preceding we studied the properties of our compression algorithms using PSNR values. For medical image compression, another important concern is how much diagnostic information would be kept after compression. Since both datasets are acquired for the purpose of left ventricle quantification, we measured the left ventricle volume before and after compression to study the impact of compression on left ventricle quantification. To keep the study as objective as possible, a robust geometric deformable model[34] is used to measure the left ventricle volume, instead of using manual tracing method. For dataset 1, the measured volumes are 71.60, 74.05, 74.95, and 74.14 ml, respectively, under the corresponding compression ratios of 1:1, 80:1, 160:1, and 240:1. And for dataset 2, the measured volumes are 108.73, 112.36, 112.36, and 111.57 ml, respectively, under the same compression ratios. The maximal errors caused by compression are 4.66 and 3.34%, respectively, for datasets 1 and 2. The errors caused by the compression are smaller than those caused by manual tracing method,[35] which is used by clinical doctors to measure left ventricle volume. Obviously, compression does not greatly affect the accuracy of left ventricle quantification. Left ven-

Table 4 The average PSNR2D values of slice images of datasets 1 and 2, along (a) x, (b) y, and (c) z, respectively, under different compression ratios.

a						
	60:1	80:1	100:1	120:1	140:1	160:1
Dataset 1	43.64±5.88	42.00±6.34	41.33±6.53	40.71±6.58	40.25±6.62	39.46±6.92
Dataset 2	39.97±4.70	39.02±4.77	38.22±4.78	37.12±4.88	36.68±4.98	36.31±5.03
b						
	60:1	80:1	100:1	120:1	140:1	160:1
Dataset 1	44.79±6.68	43.85±7.79	43.25±8.00	42.66±8.08	42.21±8.13	41.45±8.42
Dataset 2	41.10±5.91	40.18±6.02	39.43±6.12	38.39±6.30	37.97±6.40	37.67±6.54
c						
	60:1	80:1	100:1	120:1	140:1	160:1
Dataset 1	42.37±3.86	40.59±4.06	39.90±4.23	39.27±4.26	38.79±4.27	37.84±4.32
Dataset 2	40.78±4.91	39.84±5.02	39.07±5.06	37.93±5.11	37.54±5.16	37.24±5.28
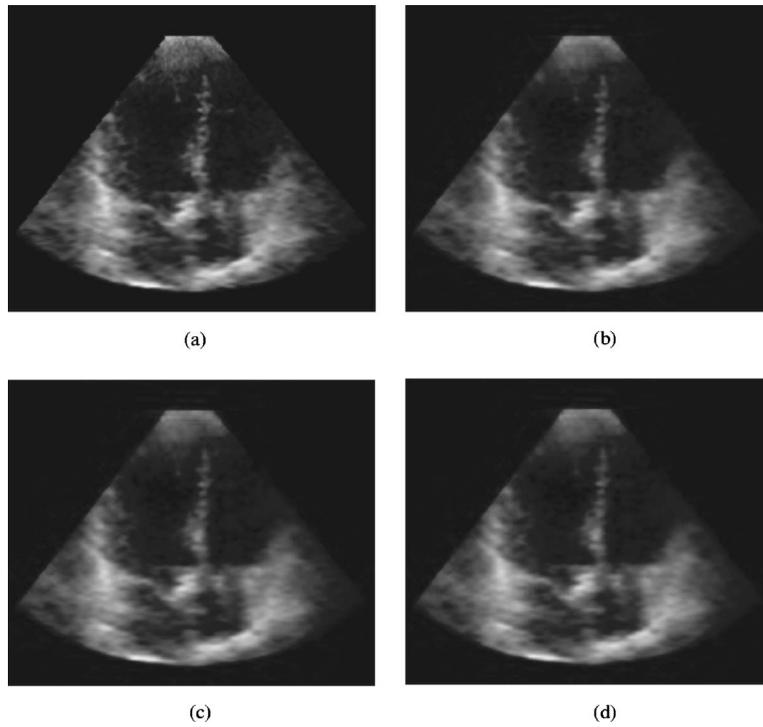
Fig. 14 Slice images at $y=80$ of dataset 2 and the corresponding compressed datasets under differential compression ratios: (a) original, (b) 100:1 (PSNR=33.93 dB), (c) 140:1 (PSNR=32.48 dB), and (d) 160:1 (PSNR=32.16 dB).
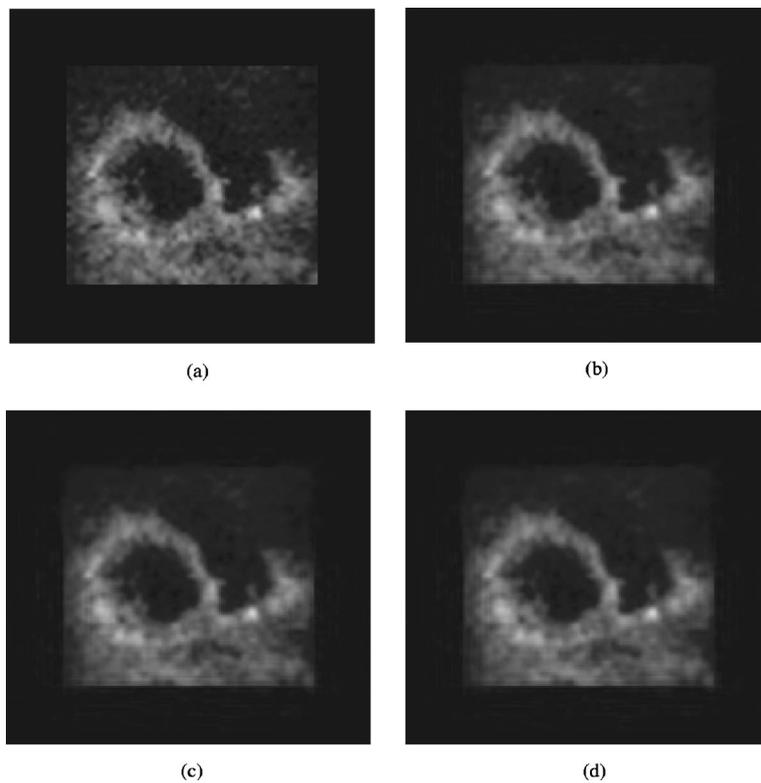
Fig. 15 Slice images at $z=85$ of dataset 2 and the corresponding compressed datasets under differential compression ratios: (a) original, (b) 100:1 (PSNR=34.57 dB), (c) 140:1 (PSNR=33.00 dB), and (d) 160:1 (PSNR=32.55 dB).

tricle volume is an important clinical parameter for cardiac disease diagnosis. The preceding study just shows that under appropriate compression ratio, for example, 80:1, the compression algorithm may still retain the critical information for diagnosis. Further experiments should be performed in the future to study comprehensively the impact of compression on clinical diagnosis.

5 Conclusions

A new SPIHT algorithm based on the WPT was described. The algorithm is very powerful in compressing 3-D echocardiographic images. The experimental results show that the new algorithm can achieve very high compression ratios while still retaining good image quality. Furthermore, the new SPIHT algorithm outperforms the original SPIHT algorithm for the compression of 3-D echocardiographic images. Because medical images share many common features and structures, the new algorithm may be applied to the compression of 3-D medical images of other modalities as well. Future study may incorporate motion compensation into the new algorithm for the compression of 3-D echocardiographic video.

References

1. http://www.jpeg.org/JPEG/
2. T. H. Karson, S. Chandra, A. J. Morehead, W. J. Steward, S. Nissen, and J. D. Thomas, "JPEG compression of digital echocardiographic images: impact on image quality," *J. Am. Soc. Echocardiogr* **8**, 306–318 (1995).
3. T. H. Karson, R. C. Zepp, S. Chandra, A. J. Morehead, and J. D. Thomas, "Digital compression of echocardiograms offers superior image quality to analog storage even with 20:1 digital compression: results of the digital echo record access study," *J. Am. Soc. Echocardiogr* **9**, 769–778 (1996).
4. http://medical.nema.org/
5. J. Puniene, V. Punys, and J. Punys, "Ultrasound and angiographic image compression by cosine and wavelet transforms and its approval in clinical environment," in *Proc. IEEE 2nd Int. Symp. on Image and Signal Processing and Analysis*, pp. 74–79 (2001).
6. http://www.jpeg.org/JPEG2000/
7. http://www.mpeg.org/
8. M. J. Garcia, J. D. Thomas, N. L. Greenberg, J. Sandelski, C. Herrera, C. Mudd, J. Wicks, K. Spencer, A. Neumann, B. Sankpal, and J. Soble, "Comparison of MPEG-1 digital videotape with digitized sVHS videotape for quantitative echocardiographic measurements," *J. Am. Soc. Echocardiogr* **14**, 114–121 (2001).
9. M. L. Main, D. Foltz, M. S. Firstenberg, E. Bobinsky, D. Bailey, B. Frantz, D. Pleva, M. Baldizzi, D. P. Meyers, K. Jones, M. C. Spence, K. Freeman, A. Morehead, and J. D. Thomas, "Real-time transmission of full-motion echocardiography over a high-speed data network: impact of data rate and network quality of service," *J. Am. Soc. Echocardiogr* **13**, 764–770 (2000).
10. http://leonardo.telecomitalialab.com/icjfiles/mpeg-4_si/
11. C. Lau, J. E. Cabal, A. H. Rambhia, H. Avni, and Y. Kim, "MPEG-4 coding of ultrasound sequences," in *Proc. SPIE* **3976**, 573–579 (2000).
12. V. Vlahakis and R. I. Kitney, "Wavelet-based, inhomogeneous, near-lossless compression of ultrasound images of the heart," *Proc. IEEE Computers in Cardiology*, pp. 549–552 (1997).
13. A. Cziho, G. Cazuguel, B. Solaiman, and C. Roux, "Medical image compression using region-of-interest vector quantization," in *Proc. 20th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Vol. 3, pp. 1277–1280 (1998).
14. J. G. do Rosiles and M. J. T. Smith, "Model-based coding of SAR and ultrasound images," in *Proc. 7th IEEE Int. Conf. on Electronics, Circuits and Systems*, Vol. 1, pp. 62–65 (2000).
15. J. Luo, H. Li, and C. Chen, "Ultrasound image coding based on subband decomposition and speckle synthesis," in *Proc. IEEE Conf. on Image Processing*, Vol. 3, pp. 106–109 (1997).
16. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.* **41**, 3445–3462 (1993).
17. J. E. Cabral Jr., D. T. Linker, and Y. Kim, "Compression for pre-scan-converted ultrasound sequences," in *Proc. SPIE* **3335**, 378–387 (1998).
18. L. Zeng, C. P. Jansen, S. Marsch, M. Unser, and P. R. Hunziker, "Four-dimensional wavelet compression of arbitrarily sized echocardiographic data," *IEEE Trans. Med. Imaging* **21**(9), 1179–1187 (2002).
19. A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.* **6**, 243–250 (1996).
20. X. Li, G. Hu, and S. Gao, "Design and implementation of a novel compression method in a tele-ultrasound system," *IEEE Trans. Inf. Technol. Biomed.* **3**(3), 205–213 (1999).
21. N. E. Cabral, D. T. Linker, and Y. Kim, "Ultrasound telemedicine system supporting compression of pre-scan-converted data," in *Proc. SPIE* **3976**, 350–358 (2000).
22. E. Chiu, J. Vaisey, and M. S. Atkins, "Wavelet-based space-frequency compression of ultrasound images," *IEEE Trans. Inf. Technol. Biomed.* **5**(4), 300–310 (2001).
23. C. Herley, Z. Xiong, K. Ramchandran, and M. T. Orchard, "Joint space-frequency segmentation using balanced wavelet packet trees for least-cost image representation," *IEEE Trans. Image Process.* **6**(9), 1213–1230 (1997).
24. Z. Xiong, X. Wu, S. Chen, and J. Hua, "Lossy-to-lossless compression of medical volumetric data using three-dimensional integer wavelet transforms," *IEEE Trans. Med. Imaging* **22**(3), 459–470 (2003).
25. A. A. Kasim, P. Yan, W. S. Lee, and K. Sengupta, "Motion compensated lossy-to-lossless compression of 4-D medical images using integer wavelet transforms," *IEEE Trans. Inf. Technol. Biomed.* **9**(1), 132–138 (2005).
26. R. Coifman and Y. Meyer, "Size properties of wavelet packets," in *Wavelets and their Applications*, M. B. Ruskai *et al.*, Eds. pp. 125–150, Jones and Barlett, Boston (1992).
27. R. R. Coifman and M. V. Wickerhauser, "Entropy based algorithms for best basis selection," *IEEE Trans. Inf. Theory* **38**(2), 713–718 (1992).
28. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego (1999).
29. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (1989).
30. C. He, J. Dong, Y. F. Zheng, and Z. Gao, "Optimal 3D coefficient tree structure for the 3-D wavelet video coding," *IEEE Trans. Circuits Syst. Video Technol.* **13**(10), 961–972 (2003).
31. C. M. Brislawn, "Preservation of subband symmetry in multirate signal coding," *IEEE Trans. Signal Process.* **43**(12), 3046–3050 (1995).
32. I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Commun. ACM* **30**(6), 520–540 (1987).
33. I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia (1992).
34. X. Hang, "Compression and segmentation of three-dimensional echocardiography," Dissertation, The Ohio State University (Summer, 2004).
35. J. X. Qin, M. Jones, T. Shiota, N. L. Greenberg, H. Tsujino, M. S. Firstenberg, P. C. Gupta, A. D. Zetts, Y. Xu, J. P. Sun, L. A. Cardon, J. A. Odabashian, S. D. Flamm, R. D. White, and J. D. Thomas, "Validation of real-time three-dimensional echocardiography for quantifying left ventricular volumes in the presence of a left ventricular aneurysm: in vitro and in vivo studies," *J. Am. Coll. Cardiol.* **35**, 900–907 (2000).

Xiyi Hang received his MS and PhD degrees in biomedical engineering from the Ohio State University, Columbus, in 2002 and 2004, respectively, and his BS degree in biomedical engineering from Zhejiang University, Hangzhou, China, in 1992. Since August 2004 he has been with the Department of Electrical and Computer Engineering, California State University, Northridge. Dr. Hang's research interests are primarily image processing and bioinformatics.

Neil L. Greenberg: Biography and photograph not available.

Yuan F. Zheng received his MS and PhD degrees in electrical engineering from the Ohio State University, Columbus, in 1980 and 1984, respectively. His undergraduate education was received at Tsinghua University, Beijing, China, in 1970. From 1984 to 1989, he was with the Department of Electrical and Computer Engineering, Clemson University, South Carolina and since August 1989 he has been with the Ohio State University, where he is currently a professor and chaired the Department of Electrical and Computer Engineering from 1993 to 2004. Since 2004 Professor Zheng has held a visiting appointment at the Shanghai Jiao Tong University, China. Professor Zheng is an IEEE fellow and his research interests include wavelet transforms for image, video, and audio compression and robotics, including robotics for biological applications, multiple robots coordination, legged walking robots, human-robot coordination, and personal robotics. His research has been supported by the National Science Foundation, the Air Force Research Laboratory, the Office of Naval Research, Department of Energy, DAGSI, ITEC-Ohio, etc. Dr. Zheng was and is on the editorial board of five international journals and he received the Presidential Young Investigator Award from Ronald Reagan in 1986, research awards from the College of Engineering of the Ohio State University in 1993 and 1997, respectively, and with his students the best student award a few times. In 2004, Professor Zheng was appointed to the International Robotics Assessment Panel by the National Science Foundation.

James D. Thomas received his BS degree *summa cum laude* in applied mathematics from Harvard College in 1977. After medical education at Harvard and clinical training at Massachusetts General Hospital and the University of Vermont, he was an assistant professor of medicine at Harvard and Massachusetts General Hospital before assuming his current position at the Cleveland Clinic Foundation in 1992. Dr. Thomas is the Charles and Lorraine Moore Chair of Cardiovascular Imaging in the Department of Cardiovascular Medicine and a professor of medicine and biomedical engineering at the Ohio State University and Case Western University. He is also lead scientist for ultrasound with the National Aeronautics and Space Administration (NASA) and a consultant to the Department of Defense. Dr. Thomas has over 400 peer-reviewed research manuscripts (published or in press), 100 invited articles and book chapters, and over 700 abstract presentations focusing on such areas as valvular heart disease, abnormalities of ventricular filling, and the development of new technology in echocardiography. His editorial board experience includes circulation, circulation research, JACC, JASE, echocardiography, and heart. He is currently on the Cardiovascular Board of Examiners for the American Board of Internal Medicine.