Interslice Coding for Medical Three-Dimensional Images Using an Adaptive Mode Selection Technique in Wavelet Transform Domain

Hyun Duk Cho, Jong Hyo Kim, and Jong Beom Ra

In this article the authors propose a novel interslice coding algorithm especially appropriate for medical 3-dimensional (3D) images. The proposed algorithm is based on a video coding algorithm using motion estimation/compensation and transform coding. In the algorithm, warping is adopted for motion compensation. Then, by using adaptive mode selection, an MC residual image and original image are mixed up in the wavelet transform domain for improvement in coding performance. The mixed image is then compressed by the zerotree coding method. It is proven that the adaptive mode selection technique in the wavelet transform domain is very useful for medical 3D image coding. Simulation results show that the proposed scheme provides good performance, regardless of interslice distance, and is prospective for medical 3D image compression.

Copyright © 1999 by W.B. Saunders Company

KEY WORDS: interslice coding, motion compensation, wavelet transform, zerotree coding, adaptive mode selection technique.

THE MOST commonly used digital modalities, such as x-ray computed tomography (XCT), magnetic resonance imaging (MRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT), generate multiple slices in a single examination. Although these multiple slices can be compressed slice by slice, more efficient compression can be achieved by exploiting the correlation between slices. This is because multiple slice images normally are correlated with each other. As a result, various medical 3-dimensional (3D) image compression techniques have been proposed.¹⁻⁵ These techniques can be roughly divided into two classes: one based on a 3D transformation^{1,2} and the other on block motion estimation/compensation techniques, which have been widely used in the video coding area.³⁻⁵ In the remainder of this report, the term *motion*, when applied to medical 3D images, represents interslice correspondences.

Coding methods based on 3D transformations can be considered a straightforward approach for 3D coding, and they have been investigated by many researchers. It has been shown that the coding scheme based on 3D wavelet transform provides a noticeable improvement over 2-dimen-

sional (2D) methods.² But in that study, performances for medical image compression were compared at very low bit rates. In most medical image compressions, relatively high bit rates are required to prevent the risk of misdiagnosis owing to image distortion. It is known that the improvement in the 3D method over the 2D method decreases noticeably as the bit rate increases.¹ Furthermore, 3D transformation methods are sensitive to interslice distances.⁴ In medical imaging, because the interslice spacing can vary with each scanning case, and generally is much greater than that within the slice, it is difficult for 3D transformation methods to provide good coding performances for various imaging modalities and their imaging parameters.

As an alternative, interslice coding methods have been proposed^{4,5} by adopting warping motion compensation (MC) that is being used in video coding. However, these MC methods also are sensitive to interslice distance, and the coding performances have room for improvement with respect to medical images. Therefore, the design of a medical 3D image coder based on interslice coding requires more careful examination of medical 3D image characteristics for improving compression performances.

Because video coding usually requires a lower bit rate, information contained in a motion-compensated residual image can be greatly reduced by coarse quantization. However, for medical images, it must be quantized finely enough to prevent misdiagnosis and it requires higher bit rates. Furthermore, the correlation between slices in medical 3D images is lower than that between frames in video applications. So, even after MC, there is still considerable redundancy left over, and more infor-

From the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Republic of Korea, and the Department of Radiology, College of Medicine, Seoul National University, Seoul, Korea.

Address reprint requests to Jong Beom Ra, Department of Electrical Engineering, KAIST, 373-1 Kusongdong Yusonggu Taejon, Republic of Korea.

Copyright © 1999 by W.B. Saunders Company
0897-1889/99/1204-0004\$10.00/0

mation must be handled by a coder. Therefore, in medical 3D image coding, residual image coding is more important, and a more efficient coder design is needed than in video coding.

In this article, we propose a 3D interslice image coder, which is especially appropriate for medical images in which the interslice distance varies depending on imaging modality and parameters. It overcomes the drawback of the existing coders whose performances rely on interslice distance, and also provides high coding performance. The proposed algorithm adopts warping prediction for MC, and introduces image compression based on adaptive intra/intercoding in the wavelet transform domain.

The interslice redundancy reduction method based on warping is reviewed in the next section. Following that, the adaptive mode selection technique in the wavelet transform domain is introduced, and the zerotree coding method is explained. Afterward, experimental results are provided and our conclusions are summarized.

INTERSLICE REDUNDANCY REDUCTION VIA MC

As mentioned above, the proposed medical 3D image coder is based on a video coder consisting of motion estimation/compensation (interslice decorrelation) and transform coding parts. However, characteristics of interslice variation in medical images are very different from those in a video sequence. The difference between two consecutive frames in a video sequence depends mainly on camera movement or object movement. Consequently, block matching algorithms (BMA), which are proper for a translational motion, are widely used for motion estimation in video coding. However, the interslice variation in medical 3D images results from a change in location of the consecutive slices of a stationary 3D object rather than a temporal change in a video sequence. Accordingly, the translational motion model is not suitable for medical images because most interslice variations consist of a variety of components such as zooming, rotation, new object occurrence, and occlusion.

To overcome the limitation of the translational motion model, a warping prediction method has been presented^{6,7} and applied to medical applications.^{4,5} Herein we also adopt the warping prediction method for motion compensation. The warping prediction is mainly classified into an affine transformation warping or a bilinear transformation

Fig 1. Warping prediction. Pixels in every rectangle block in the k th image are predicted from the corresponding deformed block in the $(k-1)$ th image.

warping. Because we have found that the performance of a bilinear transformation is superior to that of the affine transformation, we adopt bilinear transformation for warping prediction.

In warping prediction, the k th image I_k is considered a set of blocks of equal size, as shown in Fig 1. Then the four vertices, (u_0, v_0) , (u_1, v_1) , (u_2, v_2) , and (u_3, v_3) , in I_{k-1} corresponding to the vertices of a rectangle block in I_k , (m_0, n_0) , (m_1, n_1) , (m_2, n_2) , and (m_3, n_3) , are determined on the basis of the conventional block matching algorithm (see Fig 2). Then, for a given point (m, n) , its corresponding point (u, v) in I_{k-1} can be determined by using the bilinear transformation, ie,

$$\begin{aligned} u &= u_0 + (u_1 - u_0)r_m + (u_2 - u_0)r_n \\ &\quad + (u_3 - u_2 - u_1 + u_0)r_m r_n \\ v &= v_0 + (v_1 - v_0)r_m + (v_2 - v_0)r_n \\ &\quad + (v_3 - v_2 - v_1 + v_0)r_m r_n \end{aligned} \quad (1)$$

where r_m and r_n are given by

$$\begin{aligned} r_m &= (m - m_0)/(m_1 - m_0) \\ r_n &= (n - n_0)/(n_3 - n_0) \end{aligned} \quad (2)$$

Therefore, $I_k(m, n)$ can be predicted from $I_{k-1}(u, v)$, and $I_{k-1}(u, v)$ is determined by the bilinear interpolation with four integer points neighboring (u, v) .

Fig 2. Calculation of the displacement of a pixel, (u, v) , in terms of that of the four vertices, (u_0, v_0) , (u_1, v_1) , (u_2, v_2) , and (u_3, v_3) . Here, the bilinear transformation is used for this calculation.

The difference between an original image, I_k , and its predicted image, $\hat{I}_k$, is used as the motion-compensated residual image.

IMAGE COMPRESSION BASED ON ADAPTIVE MODE SELECTION IN THE WAVELET TRANSFORM DOMAIN

Motivation

In medical 3D image coding, the interslice distance varies according to the purposes of examinations and the kinds of organs examined. Depending on the imaging parameters, the interslice distance can be considerably large, and interslice coding may no longer be effective. Therefore, we need a robust coder that is independent of interslice distance, by increasing the portion of intracoding when images have large slice distances and by increasing the portion of intercoding when images have small slice distances.

This adaptive intra/intercoding in the spatial domain also is widely used in video coding fields such as MPEG, H.261, and H.263. The wavelet transform is performed on the entire image, whereas DCT used in MPEG or H.26x is performed block by block. Therefore, when original blocks and their motion-compensated error blocks are mixed in a single spatial-domain image, in the wavelet transform domain, their block boundaries produce high-frequency components that lower the coding performance severely. However, if we can perform the adaptive intra/intercoding in the transform domain, this problem can be overcome.

It has been found that, in the transform domain, residual images (intercoding) have less information in low frequencies, but original images (intracoding) have less information in high frequencies.^{8,9} Using this observation, the adaptive intra/intercoding technique in the wavelet-transform domain has been suggested.⁹ However, it could not provide noticeable improvement in video coding performance. This is because the relatively coarse quantization in video coding truncates most of the high-frequency coefficients in both original and residual images to zero, thus making the adaptive coding less meaningful. But medical image coding generally requires fine quantization to prevent misdiagnosis. Therefore, high-frequency coefficients after quantization are no longer negligible for residual images, and the adaptive selection of intra/inter mode in high-frequency bands becomes important.

Adaptive Mode Selection in the Wavelet Transform Domain

Before explaining the proposed algorithm in detail, let us consider a “block” in the transform domain as a basic unit for adaptive mode selection. Figure 3 shows the blocks in a wavelet-transformed image. Each block has its own size according to the frequency band to which it belongs. The block size of the highest frequency band is the largest. As the frequency band decreases, block size becomes smaller by four. In this study, the block size of the highest frequency band was fixed at 16×16 . As shown in Fig 3, blocks connected by arrows are located at similar spatial positions in each frequency band. The block at the starting point of the arrow is the “parent block”; the block located at the end point of the arrow is defined as the “child block.” Descendents of a child block are considered “descendent blocks.” The shaded areas in Fig 3 represent blocks that have no descendent blocks. The parent-child relationship will be utilized when the information of intra/intermode is transmitted (see next section). The block and the parent-child relationship correspond to the structure of the spatial orientation tree defined in the SPIHT coder,¹⁰ which is chosen as the image coder in this algorithm.

Figure 4 depicts the adaptive mode selection technique in the wavelet transform domain. First, before selecting the modes, both the original image and its residual one are wavelet transformed, as shown in Fig 4A. In practice, a transformed residual image is obtained by first transforming the original image and its predicted image, then taking the difference between these two transformed im-

Fig 3. Block structure in a wavelet-transformed image and the parent-child relationship of blocks.

Fig 4. Adaptive mode selection technique in the wavelet transform domain; (A) encoding process and (B) decoding process.

Fig 5. Distribution of intra/intermodes in the wavelet transform domain. (A) an XCT chest image and (B) its mode selection result (white, intramode; black, intermode).

Fig 6. Proposed coder using an adaptive mode selection technique; (A) encoder and (B) decoder.

ages. After wavelet transforms are performed, the encoder selects the blocks with the lower energy between blocks of the original image and those of the motion-compensated error image. The resulting image provides minimum energy by mixing the transformed original and residual images properly. Then, zerotree coding is performed on a new mixed image, and its result is transmitted to a decoder along with mode information.

In the decoder, the reconstructed image is obtained from the mixed image, mode information, and motion vectors, as shown in Fig 4B. First, the predicted image is generated using motion vectors, and then is wavelet transformed. In the mixed image, this predicted image is added only to the blocks defined as intermodes by mode information. Then, by inverse-wavelet-transforming the resulting image, the final reconstructed image is obtained. Even if two modes are mixed in the wavelet domain, the image can be reconstructed perfectly in

the decoder because wavelet transformation is a linear operator.

Wavelet Domain Image Compression Using Zerotree Coding

In this section, we will consider coding methods of mode information and mixed images, which are generated by the adaptive intra/intermode selection technique in the wavelet transform domain. Zerotree coding techniques such as EZW¹¹ and SPIHT¹⁰ have attracted much attention in recent years. In the proposed scheme, a modified EZW algorithm is adopted and modified for mode information coding, and the SPIHT coder is adopted for mixed image coding.

*Brief review of zerotree coding.*¹¹ Let us consider a data structure called a *zerotree*. A wavelet coefficient x is said to be *insignificant* with respect to a given threshold T if $|x| < T$ and *significant* otherwise. Given the threshold level T to determine

Table 1. Specifications of Stimulation Images

	XCT				MRI	
	Head	Chest 1	Chest 1'	Chest 2	Head	Head'
Slice distance	3 mm	2 mm	4 mm	7 mm	1.5 mm	3 mm
Pixel resolution	0.85 mm	0.625 mm	0.625 mm	0.625 mm	0.86 mm	0.86 mm
Image size	256 × 256	512 × 512	512 × 512	512 × 512	256 × 256	256 × 256

Fig 7. PSNR versus bit rate curves for various compression methods; (A) XCT Head, (B) XCT Chest 1, (C) XCT Chest 1', (D) XCT Chest 2, (E) MRI Head, (F) MRI Head'.

whether or not a coefficient is significant, a coefficient x is said to be an element of a *zerotree* for threshold T if *all* of its descendents and itself are insignificant with respect to T . An element of a *zerotree* for threshold T is a *zerotree root* if it is not

the descendent of a previously found *zerotree root* for the same threshold. The significance map can be efficiently represented as a string of four symbols. They are (1) *zerotree root*, (2) *isolated zero*, which means that the coefficient is insignificant but has

Fig 8. PE (peak error) versus bit rate curves for various compression methods; (A) XCT Head, (B) XCT Chest 1, (C) XCT Chest 1', (D) XCT Chest 2, (E) MRI Head, (F) MRI Head'.

some significant descendent, (3) positive significant, and (4) negative significant.

In the zerotree coding method, wavelet coefficients are quantized by successive-approximation quatization (SAQ). The SAQ sequentially applies a sequence of thresholds $T_0, \dots, T_{N-1}$ to determine

significance, where the thresholds are chosen so that $T_i = T_{i-1}/2$.

The SAQ is implemented by the iteration of 2 separate passes: the dominant pass and the subordinate pass. During a dominant pass, coefficients that have not yet been found to be significant are

Fig 9. Performance comparison between the proposed coder and the MPEG coder; (A) XCT Head and (B) XCT Chest 1.

compared to threshold T_i to determine their significance. This significance map is then zerotree coded using the four symbols mentioned previously. Each time a coefficient is encoded as significant, its magnitude is appended to the list called the subordinate list.

A dominant pass is followed by a subordinate pass in which all coefficients on the subordinate list are scanned, and the specifications of the magnitudes available to the decoder are refined to an additional bit of precision. For each magnitude on the subordinate list, this refinement can be encoded using a binary alphabet with a "1" symbol indicating that the true value lies in the upper half of the old uncertainty interval, and a "0" symbol indicating the lower half.

The process continues to alternate between dominant passes and subordinate passes where the threshold is halved before each dominant pass. And it stops at the desired rate or distortion.

Mode information coding. Now, let us try to implement zerotree coding for mode information compression. Figure 5B shows the intra/intermode distribution for a typical XCT chest image given in Fig 5A. Here, white blocks represent intrablocks and black blocks represent interblocks. In Fig 5B, note that the interblocks are mainly concentrated on the low-frequency band, and intrablocks appear mainly in the high-frequency band, as mentioned previously. If a block is selected as an intramode in a certain frequency band, its corresponding blocks in higher frequency bands likely will become an intramode. This is similar to the case of embedded zerotree wavelet coding (EZW), which is based on the hypothesis that if a wavelet coefficient in a low-frequency band is zero, all corresponding

wavelet coefficients in higher frequency bands are likely to be zero.¹¹ Therefore, the concept of EZW also can be adopted for our mode selection information coding.

In this algorithm, mode selection information is represented as a string of 3 symbols:

- TRATR (IntraTree Root): Current block itself and all descendents are intramodes
- ITRA (Isolated Intra): Current block is intramode but has some descendents that are intermodes
- INTER (Inter): Current block is intermode

Then the string is entropy coded. Through various experiments we have found that if mode selection information is represented with the 3 symbols above, it can be compressed by a ratio of about 7:1 compared with a fixed rate coding where 1-bit mode selection information is allocated to every block.

Mixed image coding. For mixed image coding, the SPIHT¹⁰ based on zerotree coding (described above) is adopted. In medical 3D images, however, it is important to maintain consistent quality throughout the slices. Therefore, we need a new bitrate control method that allocates bits to each slice so that consistent quality is maintained. If zerotree coding is not supposed to stop in the middle of each pass, the quantization error of wavelet coefficients can be determined by the threshold of the final pass. Therefore, by using the same final threshold in each slice, we can maintain the same quantization error for wavelet coefficients and achieve consistent image quality among all slices.

Then, an initial threshold of the k th slice $T_{0,k}$ can

Fig 10. XCT head image (window level, 1054; window width, 118); (A) original image, (B) reconstructed image, (C) zoomed version of A, and (D) zoomed version of B (T_f , 6; bit rate, 2.40 bpp; PSNR, 64.65 dB; PE, 11). Zoomed versions are magnified 2 times in each direction.

be defined as

$$T_{0,k} = T_f \cdot 2^{p_k-1}, \quad \text{for } k = 1, 2, \dots, M \quad (3)$$

where T_f , p_k , M denote the final threshold, the number of passes for k th slice coding, and the number of slices, respectively. As mentioned above, because consistent image quality is achieved using the same final threshold (T_f) in each slice, the compression ratio can be controlled by adjusting T_f . Then, for a given T_f , $T_{0,k}$ and p_k are to be determined so that equation 3 and the following condition can be satisfied for every wavelet coefficient $W_k(m, n)$

in the k th slice, ie,

$$|W_k(m, n)| < 2T_{0,k},$$

$$\text{for } 1 \leq m \leq N, \quad 1 \leq n \leq N \quad (4)$$

where $N \times N$ is the number of coefficients.

Figure 6 shows our proposed motion-compensated wavelet transform coder. As this figure shows, the number of wavelet transforms (including forward and inverse transforms) is 3 in the encoder and 2 in the decoder. It means that 1 more transform is needed in both the encoder and the decoder of the proposed scheme compared to

Fig 11. XCT chest image (window level, 800; window width, 1596); (A) original image and (B) reconstructed image (T_f , 10; bit rate, 1.76 bpp; PSNR, 60.14 dB; PE, 21).

conventional schemes such as MPEG, H.261, and H.263. The reason for more transforms is that adaptive mode selection is performed not in the spatial domain but in the wavelet transform domain. Although additional transforms are needed, their portion of the entire computational time is relatively small compared with that of motion estimation/compensation and zerotree coding.

EXPERIMENTAL RESULTS AND DISCUSSION

Experimental results are obtained using multiple-slice XCT head images, XCT chest images, and MRI head images. The pixel precision of all images is 12 bits. Table 1 contains the details pertaining to the images used in our experiment. The XCT Chest 1' (or MRI Head') images mentioned in Table 1 are created by selecting every other slice in the XCT Chest 1 (or MRI Head) image set so that its slice distance may be considered 4 mm (or 3 mm). Every image set consists of 30 images, and only the first slice is assigned to the intraslice.

For warping prediction, a block size of 16×16 is used, and the full search block matching algorithm with a search range of 5×5 is used for the initial search of the four vertices (u_i, v_i) in the deformed block (Fig 2). In wavelet transform, a 4-level pyramid is constructed using the 9/7-tap filter.¹²

In this article, distortion is measured by the peak signal to noise ratio (PSNR) and peak error (PE):

$$PSNR = 10 \log_{10} \left(\frac{4095^2}{MSE} \right) \text{ dB} \quad (5)$$

where MSE denotes the mean squared-error between the original and reconstructed images; and

$$PE_k = \max_{m,n} (|I_k(m,n) - \tilde{I}_k(m,n)|), \quad (6)$$

for $1 \leq m \leq N, \quad 1 \leq n \leq N$

where I_k is the k th original image, $\tilde{I}_k$ is the k th reconstructed image, and $N \times N$ is the size of an image.

A common problem in evaluating image quality by PSNR is that small average distortion can result in noticeable visual artifacts if all errors are concentrated to small important regions. Because of this drawback, PE has been introduced.

Figures 7 and 8 illustrate graphs of PSNR versus bit rates and graphs of PE versus bit rates for three coding schemes: (1) the intraslice coding scheme (INTRA), (2) the interslice coding scheme (INTER), and (3) the proposed coding scheme with an adaptive mode selection technique (ADAPT). PE in Fig 8 represents the average value of slice PEs, PE_k . Operating bit rates in this simulation are set high enough so that an original image and a reconstructed image cannot be distinguished.

As explained in section III, the final threshold T_f determines the quantization step. Therefore, the same T_f provides similar coding performance of PSNR or PE. (Every neighboring 3 points in Figs 7 and 8 correspond to the same T_f .)

It may be noted that as the slice gap between images gets larger, the coding performance of

Fig 12. MRI head image (window level, 105; window width, 210); (A) original image, (B) reconstructed image, (C) zoomed version of A, (D) zoomed version of B (T_r , 6; bit rate, 1.01 bpp; $PSNR$, 65.06 dB, PE , 13). Zoomed versions are magnified 2 times in each direction.

interslice coding becomes worse. For example, because the slice gaps of XCT Chest 1' and Chest 2 images, and MRI Head' images are quite large, even the intraslice coding scheme performs better than the interslice scheme (see C-F of Figs 7 and 8). Figure 7 shows that the adaptive mode selection method provides a PSNR improvement of 0.4 to 1.7 dB compared with intraslice coding, and 0.5 to 2.1 dB compared with interslice coding at the same bit rate. The figure also shows that the adaptive method provides a compression ratio improvement of 4.7 to 14.3% compared with intraslice coding, and 4.8% to 17.2% compared with interslice coding at the same PSNR. Figure 8 demonstrates that the

adaptive method increases the compression ratio compared with intraslice coding or interslice coding at the same PE. Therefore, we conclude that the proposed method offers better performance than interslice or intraslice coding schemes, regardless of the size of the interslice gap.

Finally, we compared the proposed method with the well-known MPEG coder. In this simulation, an MPEG coder is modified to handle 12-bit images. Rate control is not applied to the MPEG coder, but the same quantization step is applied to all frames like the zerotree coder adopted in the proposed coder. The group of picture (GOP) size of the MPEG coder is 12, and B frames are not used. As

shown in Fig 9, the proposed method provides drastic improvement over the MPEG coder. Therefore, the proposed method using existing video coding tools such as warping, zerotree coding, and intra/intermode selection in the wavelet-transform domain is meaningful and prospective for medical 3D image coding.

Figures 10, 11, and 12 show coding results of XCT head images with a compression ratio of 5:1, XCT chest images with a ratio of 6.8:1, and MRI head images with a ratio of 11.9:1, respectively. As shown in these figures, it is difficult to distinguish the original image from its reconstructed image, even in the zoomed version.

The encoding time of the proposed algorithm is about 8.4 seconds for an image of 256×256 in an Ultra Sparc I workstation with a CPU of 167 MHz. Most of the time (about 8.1 seconds) is for warping prediction for better motion estimation/compensation. Decoding takes only about 0.4 seconds. Even though the encoding speed is relatively slow, it does not necessarily reduce the usefulness of the proposed scheme because encoding usually is an off-line process in medical applications.

CONCLUSION

Because medical 3D images have low interslice correlation and demand high bit-rate coding compared with video sequences, a more effective image coder is required. In this report, an effective coding method of medical 3D images, which uses an adaptive mode selection technique in the wavelet transform domain, has been presented. The adaptive mode selection technique improves coding performance by mixing residual image coding (intercoding) and original image coding (intracoding) adaptively in a transformed image. This adaptive selection is based on observations that intercoding is profitable in high frequency, that intracoding is profitable in low frequency, and that intercoding is preferred in spatial regions where motion compensation is successful, with intracoding preferred otherwise.

It has been shown that the proposed coding scheme, based on adaptive mode selection, is a generalized and robust compression tool for medical 3D images, which always provides better performance than intraslice or interslice coding, regardless of slice distance.

REFERENCES

1. Wang J, Huang HK: Three-dimensional medical image compression using a wavelet transform with parallel computing. *Proc SPIE Med Imaging* 2431:162-172, 1995
2. Xiong Z, Wu X, Yun DY, et al: Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transform. *Proc SPIE Symp Visual Comm, Image Processing* 3653:327-335, 1999
3. Lee H, Kim Y, Rowberg AH, et al: Statistical distributions of DCT coefficients and their application to an interframe compression algorithm for 3-D medical images. *IEEE Trans Med Imaging* 12:478-485, 1993
4. Orchard MT, Nosratinia A, Rajagopalan R: On interframe coding models for volumetric medical data. *Proc IEEE Int Conf Image Processing* 2:17-20, 1995
5. Nosratinia A, Mohsenian N, Orchard MT, et al: Interframe coding of magnetic resonance images. *IEEE Trans Med Imaging* 15:639-647, 1996
6. Sullivan GJ, Baker RL: Motion compensation for video compression using control grid interpolation. *Proc IEEE IC-ASSP '91* 4:2713-2716, 1991
7. Nakaya Y, Harashima H: Motion compensation based on spatial transformations. *IEEE Trans Circuit Syst Video Technol* 4:339-356, 1994
8. Yamaguchi H: Adaptive DCT coding of video signals. *IEEE Trans Comm* 41:1534-1543, 1993
9. Ito H, Farvardin N: A wavelet video coder using entropy-constrained trellis coded quantization. *Proc IEEE Int Conf Image Processing* 1:598-601, 1995
10. Said A, Pearlman WA: A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans Circuit Syst Video Technol* 6:243-250, 1996
11. Shapiro JM: Embedded image coding using zerotree of wavelet coefficients. *IEEE Trans Signal Processing* 41:3445-3462, 1993
12. Antonini M, Barlaud M, Mathieu P, et al: Image coding using wavelet transform. *IEEE Trans Image Processing* 1:205-220, 1992