Region of interest coding in volumetric images with shape-adaptive wavelet transform

Ikuro Ueno[a], William A. Pearlman[b]

[a]Mitsubishi Electric Corp., 5-1-1 Ofuna Kamakura, Kanagawa 247-8501 JAPAN

[b]Rensselaer Polytechnic Institute, 110 8th street, Troy, NY 12180

ABSTRACT

We evaluated some arbitrary-shape ROI (Region of Interest) coding techniques for three-dimensional volumetric images, and through the observation we propose a flexible ROI coding with efficient compression performance. In arbitrary-shape ROI coding, the object in an image is coded with higher fidelity than the rest of the image, together with the shape information, which indicates the region of the object. In our proposed method (named as SA-ROI), in which shape-adaptive wavelet transform and scaling-based ROI coding are incorporated, the samples within the object are transformed with three-dimensional shape-adaptive wavelet transform according to its shape-information. If necessary, the background is also transformed by shape-adaptive wavelet transform independently. Then the samples within the object are scaled up by a certain number of bit-shifts, and encoded from the MSB (Most Significant Bit) plane by plane so that coefficients within the object are encoded earlier and have higher fidelity than the background. Compared with the scaling-based ROI coding using ordinary wavelet transform, which has almost the same ROI coding functionalities, SA-ROI outperforms by 5% in lossless compression ratio. Also in lossy coding, except at very low coding rates, SA-ROI has better compression performance.

Keywords: Wavelet transform, region of interest, shape-adaptive, volumetric image, three-dimension, object-based, ROI, bit-plane

1. INTRODUCTION

In many medical volumetric images, an object is placed in the center of the image and the background does not contain any important information as an image in Figure 1. For applications dealing with such volumetric images, if the object can be extracted before encoding, arbitrary-shape ROI (Region of Interest) coding, i.e., coding the object with higher fidelity than the rest of the image together with the shape information which indicates the region of the object, is beneficial for efficient coding of the object. The shape information itself can make the progressive coding more flexible, because a rough sketch of a volumetric image generated from the shape information would be useful for fast browsing as shown in Figure 2. Considering these advantages, we evaluated some coding techniques that can realize arbitrary-shape ROI coding, and propose a flexible ROI coding with efficient compression performance.

In the experiments in this document, 3D-SPIHT[1] is used as entropy coder, because it provides excellent compression performance with low complexity by encoding bit-planes of wavelet coefficients. Since entropy coding is incidental to ROI coding, other coding methods can be applied if they are bit-plane coding using wavelet transform. It should be also noted that extraction of an object is not examined, because it is regarded as being out of scope of this document.

2. CONVENTIONAL ROI CODING

Many of conventional shape-oriented coding techniques for two dimensional motion/still pictures can be extended to three-dimensional volumetric image coding. We compared some of them that are applicable to arbitrary-shape ROI coding in the following and summarized their features in Table 1.

2.1 JPEG2000 MAX-SHIFT ROI CODING

In the max-shift ROI coding adopted in JPEG2000 part-1 [5], which we call MS-ROI hereafter in this document, an entire volumetric image is transformed and only the coefficients associated with the ROI are scaled up through a given number of bit-shifts, where the number of bit-shifts, which is called scaling value *s*, is given by the largest number of non-empty magnitude bit-planes of the coefficients. A conceptual figure of max-shift ROI coding is shown in Figure 3(a). The bit-planes of coefficients are encoded plane by plane to let the ROI have higher fidelity than the rest of the image. The same concept can be applied to coefficients produced with three-dimensional wavelet transform. Note that not only the coefficients within the ROI but also coefficients surrounding the ROI that affect the image samples within the ROI need to be encoded to realize lossless coding of the ROI. In Figure 3, the additional coefficients are marked with dark hatch. Figure 4 shows an example of one-dimensional image samples and corresponding high-pass and low-pass subbands using the 5-3 wavelet filter [5].

$$x_{2n} = l_n - \lfloor (h_{n-1} + h_n + 2) / 4 \rfloor$$
$$x_{2n+1} = h_n + \lfloor (l_n + l_{n+1}) / 2 \rfloor$$

Considering the inverse wavelet transform described above, it can be seen that coefficients l_5, h_1 and h_5 (marked with dark hatch) as well as the coefficients within the object are necessary to reconstruct samples from x_4 to x_9, where x_i, l_i and h_i represent values of an image sample, a low-pass subband coefficient and a high-pass subband coefficient respectively. The additional coefficients depend on the filter length and the wavelet decomposition depth.

One of the advantages of this method is that it does not need to transmit the shape information as additional information and just send the scaling value s, because the decoder can identify coefficients scaled up just by comparing each coefficient with a threshold 2^s. However with code stream associated with the object (most significant s bit-planes) the object cannot be exactly decoded, since the decoder cannot distinguish coefficients within the object from coefficients surrounding the object. Figure 6(a) shows an example of a decoded volumetric image, in which samples surrounding the object are affected by the object and cause degradation of image quality.

2.2 JPEG2000 SCALING-BASED ROI CODING

In the scaling-based ROI coding adopted in JPEG2000 part-2 [6], which we call SB-ROI hereafter, an entire volumetric image is transformed, and the coefficients associated with the ROI (within and around the ROI) are scaled up by a certain number of bit-shifts as shown in Figure 3(b). Then the bit-planes of coefficients are encoded plane by plane. The difference of image quality between the ROI and non-ROI can be controlled by specifying the scaling value. Although JPEG2000 part-2 specifies scaling-based ROI coding only for rectangular or elliptic areas of a two-dimensional image, the concept of scaling-based ROI coding can be easily extended to arbitrary-shape ROI coding for volumetric imagery. In the scaling-based ROI coding, shape information has to be transmitted to the decoder unlike the max-shift ROI coding. Therefore, in scaling-based ROI coding, the object can be exactly decoded by discarding all of the background, but looking at the background near the object, the additional coefficients still might cause unwanted effect at an early stage of progressive coding.

2.3 OBJECT-BASED CODING WITH SHAPE-ADAPTIVE WAVELET TRANSFORM

Object-based approaches for video coding [2] [3], which we call SA-DWT, are being studied as a new video coding paradigm, in which only the samples within an object are transformed with shape-adaptive wavelet transform [4] according to the shape information additionally sent to the decoder and the coefficients are encoded. (Usually the background is not considered to be encoded.) The length of each one-dimensional segment to be transformed varies from segment to segment. Figure 7 shows an example of shape-adaptive wavelet transform, in which only the hatched samples are transformed and both edges of the segment are extended by the symmetric extension.

In shape-adaptive wavelet transform, the number of coefficients associated with an object is identical to the number of image samples within the object. Compared with SB-ROI, fewer samples are necessary to encode the object, so more efficient coding of the object can be expected. However flexible ROI coding such as user-driven ROI coding on interactive applications is difficult to realize. In user-driven ROI coding, the user modifies or specifies an ROI in the middle of the coding. In case of SB-ROI, it can be easily realized by specifying/modifying the ROI of the coefficients to be shifted as shown in Figure 5. In SA-DWT, however, the wavelet transform needs to be performed again when the ROI is specified/modified, and it would cause much increase of computational cost to efficiently use code data having been transmitted until that time.

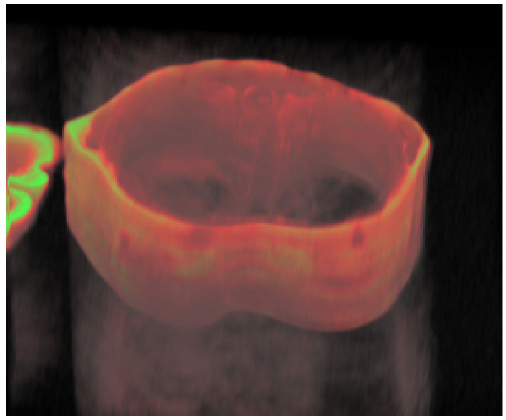


Figure 1 MRchest (original).

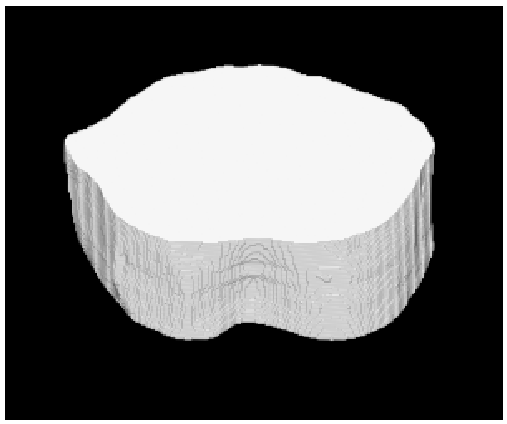
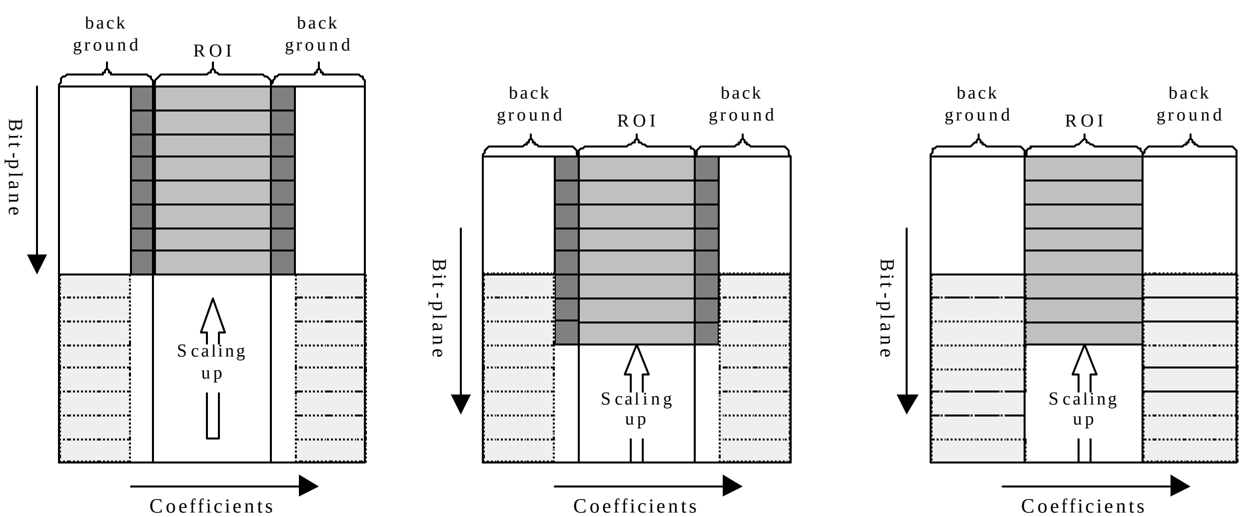


Figure 2 Rough sketch of “MRchest” generated from its shape information.



(a) MS-ROI (b) SB-ROI (c) SA-ROI

Figure 3 coefficents bit-shifting in ROI coding.

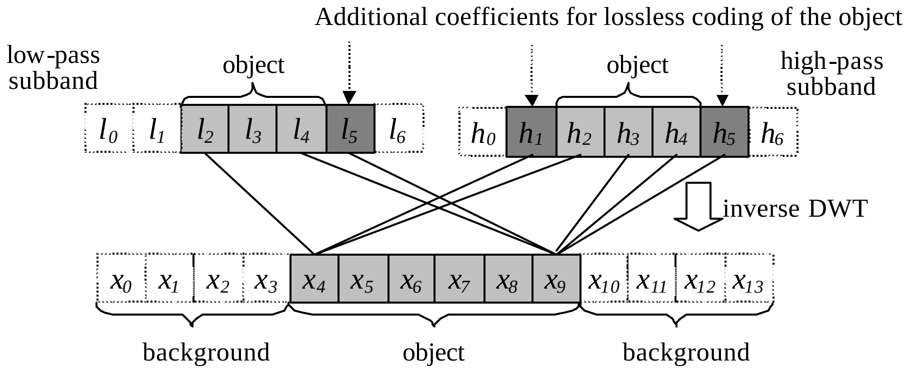


Figure 4 The inverse wavelet transform with the 5-3 filter.

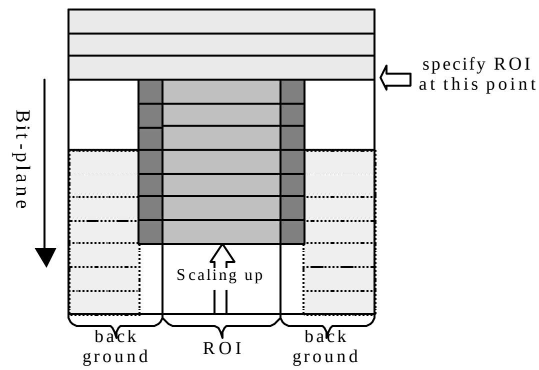
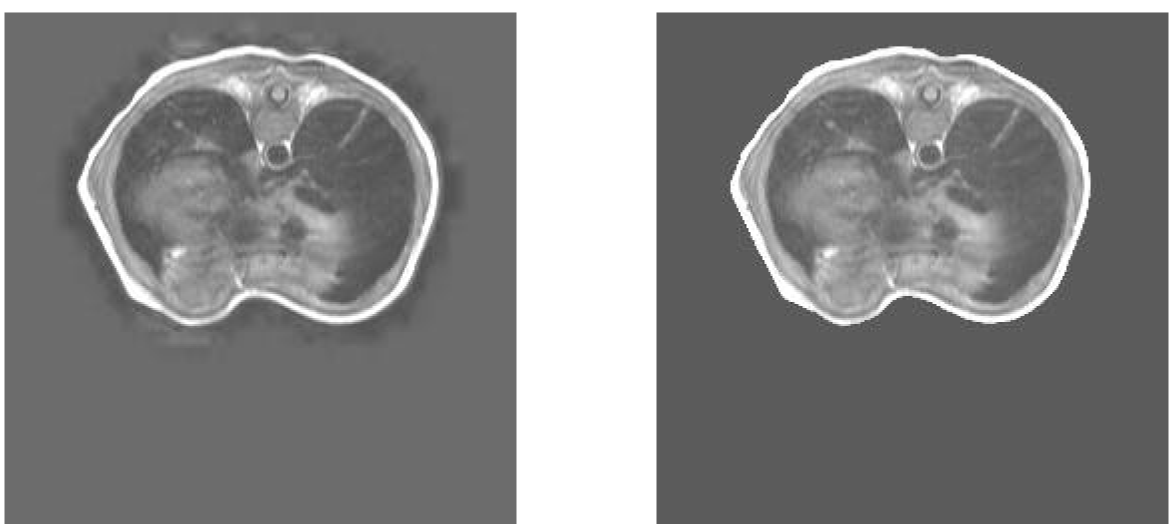


Figure 5 User-driven ROI coding with SB-ROI.



(a) MS-ROI (b) SA-ROI

Figure 6 Reproduce volumetric images (“MRchest” slice 9, 0.50 bits/sample)

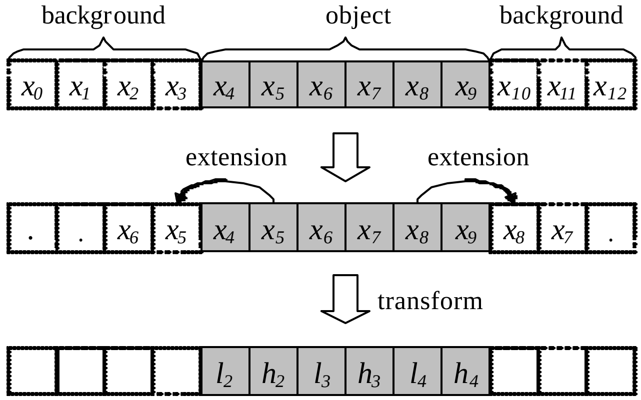


Figure 7 Shape adaptive wavelet transform with symmetric extension.

Table 1 Comparison of arbitrary-shape ROI coding

		MS-ROI	SB-ROI	SA-DWT	SA-ROI
Factors concerning compression performance	Number of coefficients needed to decode the object.	more	more	fewer	fewer
	Number of coefficients needed to decode the entire image.	same	same	---	same
	Shape information transmission	unnecessary	necessary	necessary	necessary
ROI coding functionality	User-driven ROI coding	partly possible	possible	difficult	possible
	Decoding exactly the object.	impossible	possible	possible	possible

3. ROI CODING WITH SHAPE-ADAPTIVE WAVELET TRANSFORM

We propose an ROI coding technique combining shape-adaptive wavelet transform and scaling-based ROI, which we call SA-ROI. In this method, the samples within the object are transformed with shape-adaptive wavelet transform according to the shape-information. If necessary, the background is also transformed by shape-adaptive wavelet transform independently. Then the samples within the object are scaled up by a certain number of bit-shifts, and encoded plane by plane as shown in Figure 3(c). In this case, the number of coefficients to be encoded does not change from the number of image samples within the object, and by scaling-up the coefficients of the object, the difference of image quality between the object and background can be controlled. If all of the samples (both of the object and background) are transformed with shape-adaptive wavelet transform, an ROI can be specified independently of the object/background, and user-driven ROI coding can also be realized as in the same manner as SB-ROI.

Computational cost of shape-adaptive wavelet transform is generally higher than conventional transform, because the length of each segment is not constant. However, if the object occupies only a small portion of the image and the background is not considered, the total computational cost could be reduced.

4. CODING PERFORMANCE

Comparison of the ROI coding techniques described above is summarized in Table 1. In terms of ROI coding functionality, SA-ROI and SB-ROI are superior to the others, since these two can decode the object exactly and user-driven ROI coding can be easily realized. Also they can flexibly regulate the difference of the image quality between the background and the object by the scaling value. Figure 10 illustrates PSNR in the object together with the PSNR in the background in the case of SA-ROI with scaling value 3 and 6. MS-ROI has the advantage of simplicity, and user-driven ROI coding is also possible but it is not so flexible as that of scaling-based ROI coding. SA-DWT is not appropriate if flexible ROI coding such as user-driven ROI coding is required.

We compared the compression performance of SA-ROI and SB-ROI, because they have more flexible functionalities than the others. A common block diagram of both is shown in Figure 9. Wavelet transform and scaling-up of coefficients are different between these two. Three-dimensional shape-adaptive wavelet transform is used in SA-ROI, while the entire image samples are transformed with conventional three-dimensional wavelet transform in SB-ROI. Scaling-up of coefficients is shown in Figure 3(b) and (c). Before coding, the object is extracted and shape information is produced. The shape information, which indicates whether each sample belongs to the object or background with its value 0 or 1, is encoded losslessly by using context-based binary arithmetic coding. Various techniques to encode the shape of an object are available, but here we adopt a simple method that can utilize the arithmetic coder used in the coding of the coefficients in SPIHT. Ten bits near a target bit are used as a context; five bits from the current slice and five bits from the previous slice, i.e., totally 2^{10} contexts. As entropy coder, 3D-SPIHT[1], which is extended to three-dimension from SPIHT[7], is used. In this experiment, a volumetric image is divided into several groups of slices (GOS) and each GOS is encoded independently to save memory consumption. Each GOS, which consists of 16 splices, is decomposed with wavelet transform for all special orientations and balanced three-dimensional zero-trees of each binary bit-plane of the coefficients are constructed as in Figure 8. Other experimental conditions are as follows:

DWT filters:	5-3 reversible filter, 9-7 irreversible filter [5].
DWT decomposition:	3-level dyadic decomposition.
Volumetric image:	MRchest, 256x256x64, 8 bits/sample.

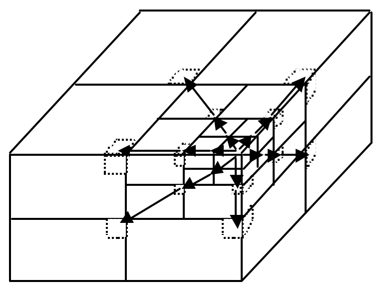


Figure 8 Three-dimensional wavelet decomposition and zero-trees of a GOS (Three-level decomposition).

In Table 2, the numbers of coefficients necessary to decode the object are shown. The numbers in SB-ROI depend on the filter length, while the numbers in SA-ROI are constant regardless of the filters. The differences of the numbers between these two are 17.6% for the 9-7 filter and 7.5% for the 5-3 filter. As the filter length or the decomposition level increases, the difference would be larger.

Table 3 shows the coding rates for lossless coding of the object (Coding of the background is not considered). The code size by SA-ROI is smaller by 4.9% in case of the reversible 5-3 filter. Note that both coding rates include the additional information for shape coding, which is 0.0347 bits/sample or about 4% of the total lossless code size.

The lossy coding performances using the irreversible 9-7 filter are illustrated in Figure 11. In Figure 11, the scaling value is the largest number of magnitude bit-planes of coefficients, and PSNR in the object are compared between SA-ROI and SB-ROI. Except at very low coding rates, SA-ROI has better compression performance than SB-ROI.

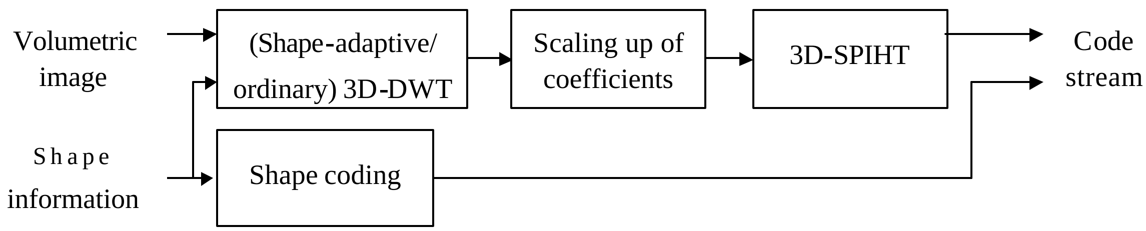


Figure 9 Block diagram.

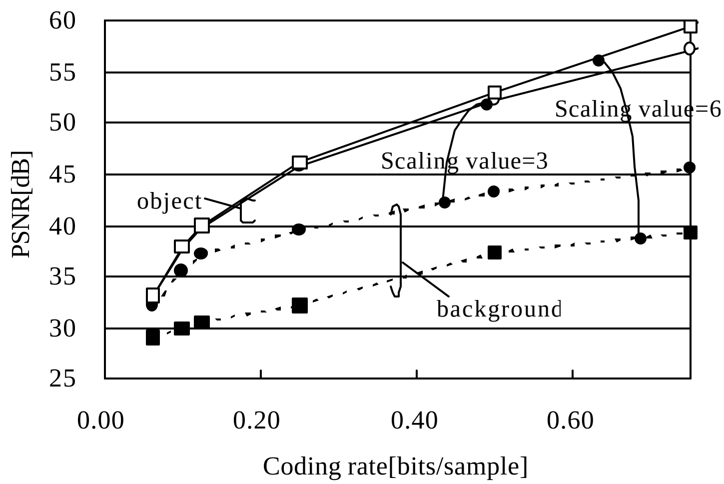


Figure 10 PSNR by SA-ROI(Irreversible 9-7 filter, scaling value=3, 6).

Table 2 The number of coefficients necessary to decode the object.

	SB-ROI	SA-ROI	difference(%)
9-7 filer	1433039	1218283	17.6
5-3 filter	1309518	1218283	7.5

Table 3 Coding rates for lossless coding of the object [bits/sample] (Reversible 5-3 filter)

	SB-ROI	SA-ROI	difference(%)
Coding rate [bits/sample]	0.872	0.831	4.9

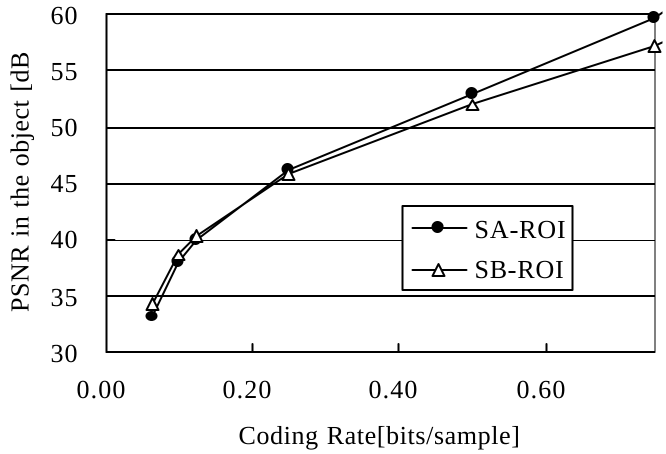


Figure 11 PSNR in the object by SA-ROI and SB-ROI (Irreversible 9-7 filter).

5. CONCLUSION

We evaluated some arbitrary-shape ROI coding techniques for volumetric images and proposed SA-ROI (scaling-based ROI coding with shape-adaptive wavelet transform). We showed that SB-ROI (scaling-based ROI coding adopted in JPEG2000) and SA-ROI have superior ROI coding functionalities than the others, and SA-ROI is better than SB-ROI in terms of compression performance of the object. In applications where the object in an image is the main target to be coded and the object can be extracted before encoding, our proposed method can be a useful solution for functional and efficient coding.

6. REFERENCES

[1] B-J. Kim and W.A. Pearlman, "An Embedded Wavelet Video Coder Using Three-Dimensional Set Partitioning in Hierarchical Trees," *IEEE Data Compression Conference*, pp. 251-260, March 1997.

[2] G.Minami, Z.Xiong, A.Wang and S.Mehrotra, "3-D wavelet coding of video with arbitrary regions of support," *IEEE Trans. Circuits Syst. Video Technol*., vol. 11, No.9, pp.1063-1068, 2001.

[3] Z. Lu and W. A. Pearlman, "Wavelet Video Coding of Video Object by Object-Based SPECK Algorithm," *Picture Coding Symposium (PCS-2001)*, pp. 413-416, April, 2001.

[4] S.Li and W.Li, "Shape-adaptive discrete wavelet transform for arbitrary shaped visual object coding," *IEEE Trans. Circuits Syst. Video Technol*., vol. 10, pp.725-743, 2000.

[5] ISO/IEC 15444-1, JPEG2000 Image Coding System.

[6] ISO/IEC 15444-2, JPEG2000 Image Coding System:Extensions.

[7] 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.*, vol. 6., pp. 243--250, June 1996.