OBJECT CODING USING A SHAPE ADAPTIVE WAVELET TRANSFORM WITH SCALABLE WDR METHOD

Bindulal T.S.*, M. R. Kaimal+

*Student member, IEEE, +Member, IEEE
Department of Computer Science, University of Kerala, India

ABSTRACT

In this paper, a shape adaptive wavelet transform coding scheme for arbitrarily shaped visual objects which is also scalable and progressive is proposed. The proposed scheme uses the shape adaptive wavelet transform (SA-DWT) of extracted objects for generating the wavelet coefficients and the wavelet difference method (SWDR) for progressive scalable coding and decoding. The method is developed for the limited bandwidth network where the image quality and data compression are most important. The experiments are performed on the MRI medical images, and video sequence images. The simulation results show that the proposed scheme is slightly better at all bit rates, in PSNR values compared to the other scalable coding schemes like SPIHT etc. Thus, the proposed coding scheme gives a convenient and better object coding scheme with applications in the multimedia image compression and transmission scenario.

***Index Terms*—**object coding, shape adaptive discrete wavelet transform, WDR method, scalable image transmission

1. INTRODUCTION

The wavelet transform based traditional coding techniques are popular in the field of multimedia image compression and transmission. Now a days, the image compression algorithms are focusing on the region of interest of the images so that cost effective and lossless compression can be achieved at the expense of unwanted data. It is often useful to reconstruct the region of interest (ROI) of an image before the background of the image is processed in many applications like web browsing, image database and telemedicine [10].

The coding of arbitrarily shaped objects supporting the two important features viz., namely the rate and resolution scalability is popular in the object based scalable coding scenario [3]. The image can be considered as a composition of different arbitrarily shaped objects and encoding these objects individually rather than considering the whole picture as a rectangular matrix of pixels will be more efficient in situations mentioned above. Recently, Li and Li [9] proposed an important modification of DWT called shape adaptive DWT, (SA-DWT) for enabling the wavelet based coding of arbitrarily shaped objects. The main feature of the SA-DWT is that the number of pixels in the arbitrarily shaped visual objects is identical to the number of coefficients in the transformed wavelet domain. Moreover, the spatial correlation and self similarity across the subbands are well preserved in the SA-DWT.

A lot of research work has been reported in the field of shape adaptive object coding paradigm like ZTE [1] and SPIHT [2, 3] to improve the coding efficiency and for supporting the scalability property. Many coding schemes like the wavelet difference reduction method (WDR) [4] makes use of the run length coding scheme and has been found to be an efficient coding scheme. The performance of WDR algorithm has been improved by various authors [5, 6]. Adaptively scanned WDR (ASWDR) scheme of Walker and Nguan [5] and context-modeling with WDR (CMWDR) method of Yuan and Mandal [6] offer substantial improvements in performance than the powerful SPIHT method. Here, the authors modified the CM-WDR [6] as SWDR [7] to support SNR scalability as well as resolution scalability. The adaptive scalable WDR method proposed by authors [7] is applied along with the shape adaptive wavelet transform method in this paper for enabling more efficient shape adaptive object coding.

Further sections are organized as follows: a brief description of the shape adaptive discrete wavelet transform is presented in the section 2; the proposed scalable WDR (SWDR) algorithm is presented in section 3; the experimental results are discussed in section 4; the conclusion and references are given in section 5 and section 6 respectively.

2. SHAPE ADAPTIVE DISCRETE WAVELET TRANSFORM

Object based coding algorithms are the new approaches in the image coding paradigm, in which the data in the object are transformed with shape adaptive wavelet transform. The shape adaptive wavelet transform is done in accordance with the shape information and it is additionally sent to the decoder for perfect reconstruction [3].

The object coding is useful in the medical images. Generally, the useful clinical information in a medical image is in the central part of the image which needs to be compressed without any loss. The background part does not contain any clinical information and consumes unnecessary bit budget and reduces the performance of the compression scheme. Not only the medical objects, but the visual textures are also coded using object based coding algorithms. The background of the visual objects has lower attention by the viewer and the object with higher attention has to be coded very efficiently as shown in figure 1.

The shape adaptive wavelets transform progresses through each row of the segmented objects [9, 10]. The 1-D transform is applied to each row and then each column to obtain the 2-D transform. Each row intersects the object and forms one or more foreground segments which can be found from the shape information. A lifting scheme with symmetrical boundary extension is applied on each segment independently. The low and high pass bands of horizontal transform are further processed column by column. Each column intersects the object and forms one or more vertical segments. One main feature is that number of wavelet coefficients after SA-DWT is exactly same as the number of pixels in the object.

During the coding of arbitrary shape objects, there are many isolated segments which require special consideration [9, 10]. These segments fall into one of the following cases: 1) even length signal at even position 2) even length signal at odd position 3) odd signal at even position 4) odd signal at odd position [10]. The signal is symmetrically extended around its boundaries before starting the wavelet transform. SADWT is evaluated according to the length of segment as follows:

1) If the length of the isolated segment N is unity, and if the position of the pixel is even, the coefficient is appended to the low frequency band, else the position of the pixel is odd and the coefficient is appended to the high band. The normalization is done by $\sqrt{2}$.
2) If the length of signal N is even, then the isolated segment is transformed using the lifting scheme to produce N/2 high frequency coefficients using even sub sampling (2i+1) and (N/2) low frequency coefficients using (2i) sub sampling, respectively.
3) If the length of signal N is odd, then if the isolated segment starts at an even position, the isolated segment is transformed using the lifting scheme to produce (N/2) high frequency coefficients and (N/2)+1 low frequency coefficients. This done by odd sub sampling (2i) for low frequency coefficients and even sub sampling (2i+1) for high frequency coefficients. If the isolated segment starts at an odd position, the isolated segment is transformed using the lifting scheme to produce (N/2)+1 high frequency coefficients and (N/2) low frequency coefficients. This done by odd sub sampling (2i) for high frequency coefficients and even sub sampling (2i+1) for low frequency coefficients.

After the SA-DWT, the mask information is needed for the decoder to decode the shaped information. The first level row mask is sufficient for generating all column masks and row masks in the further processing. The boundary positions of the masks are coded in lossless and sent to the decoder for perfect reconstruction.

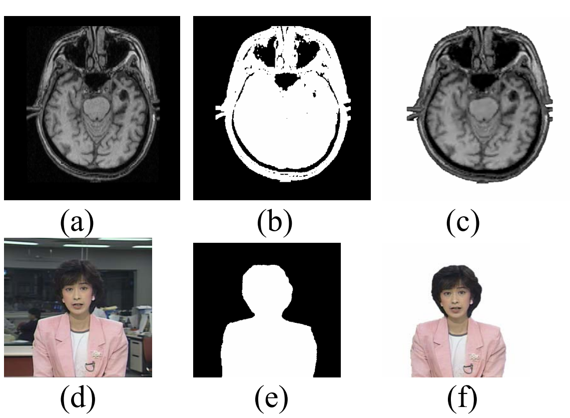


Fig. 1. (a, d) original images, (b, e) generated masks (c, f) extracted objects

3. ADAPTIVE SCALABLE WAVELET DIFFERENCE REDUCTION METHOD

The SWDR [7] is compared with original SPIHT and its scalable version S-SPIHT [3] is used along with the shape adaptive wavelet transform. The difficulty in checking zero tree of SPIHT for object coding can be avoided in the case of SWDR since an adaptive scan path technique is used in the latter method. The region growing procedure [8] will increase the coding efficiency so that the coefficients in the region of interest are coded at the earliest available situation. The generated flexible bit stream will produce different resolution images as per requirement.

The scalability properties are implemented through multiple resolution dependent lists [3]. The scalable WDR (SWDR) coding scheme uses the data structures *RGE* (coefficients that are collected during the adaptive scanning process in Region Growing manner)*, SNS* (Significant Neighbour Sub-array*, SPS* (Significant Parent Sub-array)*, LIP* (List of Insignificant Pixels), *LSP* (List of Significant Pixels), *TPS* (Temporary Set of Significant coefficients) to regroup the wavelet coefficients so as to code efficiently and get good compression results [7].

The transformed image in wavelet domain is partitioned into difference subband groups for implementing the scalability property. For each spatial subband group λ_L , the lists are ordered as RGE_L, SNS_L, SPS_L, LIP_L so that L takes values L_{max}, L_{max}*-1, ---, 1* where L_{max} is the maximum number of spatial resolution level supported by the encoder or decoder. During the processing of wavelet coefficients w_{ij} at pixel location (i, j) from the subband level λ_L , if the coefficients from outside the subband occur, it will be included in the next level of list at *(L-1)* level. Scalable WDR bit stream can be reordered for multi resolution decoding at any desired bit rate. The total number of bits belonging to a particular bit plane is the same for original scheme and its scalable version, but they are re-arranged in accordance with their spatial resolution dependency.

The significant information is generated using the significant test function $\sigma(w_{ij}, t_n)$ at the bit-plane n. The sign of the information is generated using $Sign(w_{ij})$. The encoding process progresses through the mask function $mask(i, j)$ to check whether the wavelet coefficient position is inside the object area or not.

$$\sigma(w_{ij}, t_n) = \begin{cases} 1 : |w_{ij}| \geq t_n \\ 0 : |w_{ij}| < t_n \end{cases}$$

$$Sign(w_{ij}) = \begin{cases} + & : \; w_{ij} \geq 0 \\ - & : \; w_{ij} < 0 \end{cases}$$

$$Mask(i, j) = \begin{cases} 1 & if(i, j)\, inside\, the\, object area \\ 0 & if(i, j) outside the\, object\, area \end{cases}$$

The significant coefficients are collected during the traversing through the wavelet coefficients using a predefined scan path. The restructuring of wavelet coefficients is done by using the function $cluster\,(w_{ij}, t_n)$ for neighboring coefficients around the significant coefficient w_{ij} and the $child(w_{ij}, t_n)$ for child coefficients of significant coefficient w_{ij} .

$cluster\;(w_{ij}, t_n) = \{(m, n)\}$, when

1. $(i-1) \leq m \leq (i+1), (j-1) \leq n \leq (j+1)$

2. $(m,n)\in\lambda_L \,\&\, (i,j)\in\lambda_L$

3. $(m,n)\in\{mask(m,n)=1\}$

$child(w_{ij},t_n)=\{(m,n)\}$, when

1. $(m,n)\in\left\{\begin{matrix}(2i,2j),(2i,2j+1)\\(2i+1,2j),(2i+1,2j+1)\end{matrix}\right\}$

2. $(m,n)\in\lambda_{L-1} \,\&\, (i,j)\in\lambda_L$

3. $(m,n)\in\{mask(m,n)=1\}$

The algorithm is outlined below:

1. Initialization

$$LSP_L=\phi \quad ,TSP_L=\phi \quad \forall L \quad ,1\le L\le L_{\max}$$

$$LIP_L=\begin{cases}\phi \quad , \quad \forall L \quad ,1\le L\le L_{\max} \quad such \quad that\\ RGE_L=\phi, SNS_L=\phi, SPS_L=\phi\end{cases}$$

$$n \quad = \quad \left\lfloor \log_2\left(\max_{(i,j)\in I}|w_{ij}|\right)\right\rfloor$$

$t_{n-1}=2^{n+1}, t_n=t_{n-1}/2$, *L=Lmax;*

2. Sorting pass

If mask (i, j) = 1{

If $LIP_L(\sigma(w_{ij},t_{n-1})=0)$

{If $LIP_L(\sigma(w_{ij},t_n)=1)$

{coding (w_{ij} ,L) ;}}}

If mask (i, j) = 1

{If $\lambda_L\ne\phi$

{If $\sigma(w_{ij},t_{n-1})=0$

{If $\sigma(w_{ij},t_n)=1$

{coding (w_{ij} , *L*);

$RGE_L=cluster\ (w_{ij},t_n)$;

Do {If $RGE_L\ne\phi$

{If $RGE_L(\sigma(w_{ij},t_{n-1}))=0$

{{If $RGE_L(\sigma(w_{ij},t_n))=1$

coding (w_{ij} , *L*);

$RGE_L=cluster\ (w_{ij},t_n)$}}

} while (End (RGE_L)!=True) ;}}}}

Function $coding(w_{ij},L)$

{ Output distance 'd' from previous significant.

Send binary representation of 'd' without leading MSB '1'.

Send sign information of w_{ij} $Sign(w_{ij})$. Add w_{ij} into TPS_L. }

3. Index updating pass:

If $TPS_L\ne\phi$ {

$SNS_L=cluster\ (w_{ij},t_n)\quad ;\forall(i,j)\in TPS_L$

$SPS_{L-1}=child(w_{ij},t_n)\quad ;\quad \forall(i,j)\in TPS_L$ }

$LIP_L=RGE_L+SNS_L+SPS_L$;

4. Refinement Pass:

If $LSP_L\ne\phi$

{ If $LSP_L(\sigma(w_{ij},t_{n-1})=1)$

{Add nth MSB of $LSP_L(w_{ij})$.}}

$LSP_L=LSP_L+TPS_L.;TPS_L=\phi$;

5. *Resolution scale updates:*

Send Header Information;

If (L > 1) {L=L-1; Go to step 2.}

Else{L=Lmax; }

6. *Threshold update:*

$If\ t_n>1\{t_{n-1}=t_n \quad \& \quad t_n=t_n/2$;Go to step 2.}

The algorithm produces four symbols: +, -, 1, 0. These symbols are coded as in CM-WDR algorithm [6] in the sorting pass using 2 bits with 11 for +, 10 for -, 01 for 1 and 00 for 0. The mask is coded by noting the boundary points. The proposed algorithm Scalable WDR coding scheme does not require arithmetic coding.

4. EXPERIMENTAL RESULTS

The proposed scheme SWDR is compared with original SPIHT and its scalable version S-SPIHT with Shape adaptive DWT. The algorithm is simulated on 8 bpp medical images consisting of 10 classes with 100 frames in each class and video objects like announcer etc. The original resolutions of the medical images are (512x512) pixels and video still pictures are 480x512 pixels and are publicly available at http:// www.cipr.rpi.edu / resource /sequences / index .html. The wavelet decomposition is based on the bi-orthogonal 9/7 for announcer object and 5/3 tap wavelet filters for MRI object with symmetric extension at the image boundary [9, 10].

The objects are extracted by hand, but in practice any segmentation algorithm appropriate for the application may be employed. Six levels of wavelet decomposition were first applied to each test image, and then the scalable WDR encoder was set to encode the coefficients from $bitplane_{max}$ to $bitpane_0$ supporting maximum spatial scalability levels as 7.

The bit stream for each spatial resolution at different rates and the fidelity was measured by the peak signal to noise ratio (PSNR) defined as,

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

where MSE is mean squared error between the original and the reconstructed image; ***max*** is the maximum possible magnitude of a pixel inside the image. All the results for SPIHT, scalable SPIHT (S-SPIHT) and scalable WDR (SWDR) were obtained by decoding the binary bit streams without using the arithmetic coding.

The simulation results obtained by the various algorithms and on medical images are given in Table 1, 2. A typical reconstructed image at different resolution levels is shown in figure 2. For full resolution MRI image reconstruction, the performance gain is from 0.34 dB to 0.71 dB for various bit rates. It is observed that the coding performance in PSNR values (in dB) increases when the resolution scale decreases. For resolution level 2, i.e. 256x256, the performance gains of SWDR are from 0.80 dB to 5.20 dB compared to the normal SPIHT and from 0.51 dB to 1.11 dB compared to the scalable SPIHT for various bit rates. Similar experimental results are obtained for the resolution level 3 (128x128).

The experimental results obtained for video still object in YUV format and PSNR values for luminance components (Y) are shown in the table 3, 4. For full resolution object reconstruction, the performance gain is from 0.79 dB to 1.34 dB for various bit rates. For resolution level 2, i.e. 240x256, the performance gains of SWDR are from 3 dB to 12 dB compared to the normal SPIHT and from 1.81 dB to 3.76 dB compared to the scalable SPIHT for various bit. Similar experimental results are obtained for the resolution level 3 (120x128).

Table 1: PSNR values of MRI Object of Full Resolution with size (512x512)

Method	Bit rate (bpp)			
	0.125	0.25	0.5	1.0
SPIHT	31.21	35.84	40.35	47.11
SWDR	31.92	36.43	40.93	47.45

Table 2: PSNR values of MRI Object of Half Resolution with size (256x256)

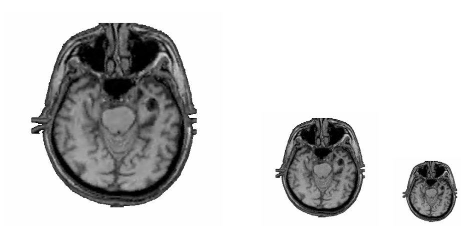
Method	Bit rate (bpp)			
	0.0625	0.125	0.25	0.5
SPIHT	27.56	31.20	36.49	43.63
S-SPIHT	27.76	31.49	37.03	47.72
SWDR	28.58	32.00	38.04	48.83

Table 3: PSNR values of Announcer Object of Full Resolution with size (480X512) (luminance Component Y)

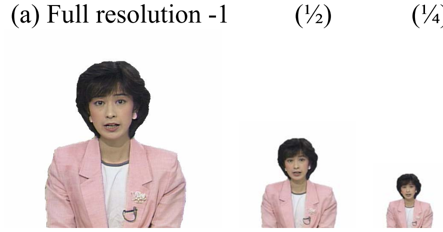
Method	Bit rate (bpp)			
	0.125	0.25	0.5	1.0
SPIHT	36.49	40.77	45.46	51.51
SWDR	37.83	41.66	46.25	52.60

Table 4: PSNR values of Announcer Object of Half Resolution with size (240x256) (luminance Component Y)

Method	Bit rate (bpp)			
	0.0625	0.125	0.25	0.5
SPIHT	33.61	38.14	43.86	50.53
S-SPIHT	34.24	39.70	46.47	58.88
SWDR	36.64	41.51	48.57	62.64



(a) Full resolution -1 (½) (¼)



(a) Full resolution -1 (½) (¼)

Fig. 2. Scalable object reconstruction at bit rate 0.0625

5. CONCLUSION

We propose a scalable WDR coding scheme with shape adaptive wavelet transform that supports spatial and SNR scalability. The flexible bit streams generated by the encoder can be decoded adaptively to get images at any level of spatial resolution. The object based coding using scalable WDR performs much better than the scalable SPIHT and the original SPIHT at any bit rate in terms of scalable properties and has lesser complexity than the zero tree coding technique. The new coding scheme is applied to the multimedia video sequences. The scalability features of new method have interesting perspectives for numerous visual communications applications.

6. REFERENCES

[1] J.M.Shapiro, “Embedded image coding using zero trees of wavelets coefficients”, IEEE Trans. Signal Process. vol. 41, pp. 3445-3462, 1993.

[2] A. Said, W. A. Pearlman, “A new fast and efficient image codec based on set portioning in hierarchical trees”, IEEE Trans. Circuits Syst. Video Techno. vol. 6, no.3, pp. 243-250, 1996.

[3] H. Danyali, A. Mertins, “Flexible, highly scalable, object-based wavelet image compression algorithm for network applications”, IEE proceedings Vis. Image Signal Process., vol 151, No. 6, p.p.- 498-510, 2004.

[4] J. Tian, Jr. R. O. Wells, “Embedded image coding using wavelet difference reduction”, Wavelet image and video compression, P. Topiwala, Ed. Norwell, MA: Kluwer, pp. 289-302, 1998.

[5] James S. Walker, Truong Q. Nguyen, “Lossy image codec based on adaptively scanned wavelet difference reduction”, Optical Engineering 39, pp. 1891-1897, 2000.

[6] Yufei Yuan, Mrinal K. Mandal, “Novel embedded image coding algorithms based on wavelet difference reduction”, In Proc. of IEE, Vol. 152, No. 1, pp. 9 – 19, Feb. 2005.

[7] Bindulal T.S., M.R. Kaimal, “Adaptive Scalable Wavelet Difference Reduction Method for Efficient Medical Image Transmission”, Proc. of IEEE, TENCON, Hong Kong, 2006.

[8] Bindulal T.S., M.R. Kaimal, “A Hybrid Scheme based on Wavelet Transform, SVD and WDR method for medical images”, Proc. of IET International Conference on VIE 2006, India , p.p. 201-206, 2006.

[9] Shipeng Li, Weiping Li, “Shape adaptive Discrete Wavelet Transform for Arbitrarily Shaped Visual Object Coding”, IEEE Trans. On Circuits Syst. Video Technol. Vol. 10, No. 5, pp. 725-43, 2006.

[10] Ping Xu and Shannn Zhu, “A New Method for Arbitrarily Shape ROI Coding Based on ISA-DWT”, Proc. Of IEEE Int. Conference on Control and Automation (ICCA 2005, pp. 1018-21), 2005.