Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

FrC10.5

UNBALANCED 3-D TREE STRUCTURE FOR REGION-BASED CODING OF VOLUMETRIC MEDICAL IMAGES

Kaibin Wang and Aysegul Cuhadar
Carleton University, Department of Systems and Computer Engineering, Ottawa, ON, Canada

Abstract - **In this paper, we propose an unbalanced three-dimensional (3-D) coefficient tree structure for 3-D region-of-interest (ROI) coding of medical volumetric data. We compare the proposed coding scheme against 3-D SPIHT [5] and a recent technique that employs optimal tree construction [3] shown to improve the performance of the 3-D SPIHT algorithm. For a MR volumetric dataset, at 0.1bpp, the proposed region-based coding scheme outperforms [3] and [5] by about 0.5 dB. The improvement is more significant at higher bitrates, up to 1dB compared to [3] and [5]. In addition, the proposed tree structure provides a more general multiregion multiquality coding framework rather than ROI/non-ROI coding.**

I. INTRODUCTION

Object-based coding is one of the most important features provided by MPEG-4 and JPEG-2000. It allows imposing heterogeneous (or region-dependent) fidelity constraints rather than encoding the entire image with a single fidelity constraint. For applications where images contain a number of objects that can be encoded at different bit rates, such as compression of medical image data, such a scheme offers better utilization of available bit rate, since high fidelity shall be maintained only for relatively small regions rather than for the entire image. Region-based extensions of wavelet coefficient trees - embedded zerotree wavelet (EZW) and SPIHT - are successfully employed for coding arbitrary-shape regions [1], [2]. These algorithms use shape-adaptive discrete wavelet transform (SA-DWT) that retains most of the features of the conventional DWT, including multiresolution property and locality. In JPEG2000, instead of using SA-DWT, conventional DWT is used. The coefficients required for reconstruction of various regions are differentiated in the transform domain. Prior to bit plane coding, ROI coefficients are scaled up by a number of power of 2. Although this mechanism provides interactive encoding utility, it is inferior to the other schemes that employ SA-DWT in terms of rate-distortion (R-D) performance. In light of this, we address 3-D wavelet coding of volumetric medical data using SA-DWT. We introduce a new 3-D coefficient tree structure, namely the unbalanced tree, so that multiple regions of interest can be defined and encoded at different bit rates. The proposed coding scheme is transparent to its conventional version (i.e. 3-D SPIHT). In other words, encoding each region independently with the proposed scheme does not introduce annoying artifacts at the boundaries of the selected arbitrary-shape regions.

II. 3-D WAVELET CODING

Success of the two-dimensional (2-D) embedded image coding algorithms including SPIHT, EZW and Embedded Block Coding with Optimal Truncation (EBCOT), led the researchers to extend these compression schemes to three dimensions for compression of video and volumetric image data. The main goal is to exploit the interslice/temporal correlation using 3-D wavelet transforms to remove the interslice/temporal redundancy. In the case of video data, the motion compensation which is not easily scalable and involves an expensive searching procedure is replaced by wavelet transform applied along the temporal dimension as in [3]. 3-D version of EZW has been proposed for compression of volumetric medical images in [4]. The authors obtained above 20% decrease, in average, in compressed file size for representation of volumetric medical images compared with 2-D SPIHT for lossless compression. 3-D version of the original SPIHT algorithm using dyadic and packet wavelet transforms have been used for video coding in [5], [2], and also employed for volumetric image compression in [6]. In [3], the authors propose a modified 3-D coefficient tree structure, namely optimal 3-D coefficient tree (asymmetric), and obtain better coding performance than original 3-D SPIHT algorithm that employs a symmetric coefficient tree structure. 3-D Embedded Subband Coding with Optimal Truncation (ESCOT) algorithm extended from EBCOT algorithm (used in JPEG2000) is used in volumetric medical data compression [7].

III. 3-D ROI CODING

By allocating more of the total bit rate for region(s) of interest (ROI) and less bit rate for the remaining regions, i.e. the background, it is possible to achieve high compression ratios. Region-based image coding schemes using heterogeneous (multiple) quality constraints are especially attractive because they not only can well preserve the diagnostic features in region(s) of interest, but also meet the requirements of less storage and shorter transmission time for medical imaging applications and video transmission.

In [8], regular cuboids are used for specifying the ROIs in volumetric image coding. Several different types of volume cropping for the specification of a 3-D ROI's are defined, such as sub-volume, fence, inverted fence, cross and inverted cross. 3-D integer wavelet transform is applied on computed tomography (CT) and magnetic resonance (MR) image

1-4244-0033-3/06/$20.00 ©2006 IEEE.

datasets and the coefficients belonging to the ROIs are scaled up. Subsequently, original 3-D SPIHT coding algorithm is used to encode the resulting subbands. In this work, we adopt arbitrary-shape representations for the ROIs, since it is more general and no information from the background region is required in the representation and coding of the ROIs. Note that, in conventional embedded image coders, the spatio-temporal orientation trees are heterogeneous when applied to region-based coding, *i.e.* the coefficients in a particular tree may belong to different ROIs. In this paper, we address this problem by introducing a new spatio-temporal orientation tree, namely unbalanced tree structure, instead of conventional spatio-temporal orientation tree used in SPIHT algorithm for coding of volumetric images.

IV. UNBALANCED 3-D TREE STRUCTURE FOR REGION-BASED CODING

Previously, a 2-D unbalanced tree structure is introduced to construct the coefficient trees for ROI coding [9], [10]. Labels are used to differentiate ROI coefficients from background coefficients. Then, unbalanced orientation trees are constructed according to the label information. If the label of a coefficient in a tree does not agree with the label of its parent coefficient, parent coefficient is said to be pseudo-parent, and in this case, the coefficient is disconnected from its pseudo-parent and a new (real) parent coefficient is sought for this coefficient. The real parent is defined as one of the eight-connected neighbors of the pseudo-parent coefficient and has the same label as the corresponding coefficient. If such a parent coefficient is found, the coefficient is connected to its real parent coefficient; otherwise the coefficient itself is used as a zerotree root upon LIS (list of insignificant sets) initialization.

Similar to 2-D case, after the computation of 3-D SA-DWT the spatio-temporal orientation trees are constructed. These 3-D coefficient trees are classified into three categories: 1) All nodes in the tree are inside the wavelet domain ROIs and they belong to the ROIs; 2) All nodes in the tree are outside the ROIs and contain no information about ROI; 3) Some nodes in the tree are outside the ROIs, and some inside the ROIs. Thus, parts of the coefficient trees contain useful information to improve the R-D performance of the coding algorithm. First two types of trees are easily treated just by keeping the first and skipping the second as in conventional SPIHT method. For the 3rd type of coefficient tree, if a coefficient in the tree is outside the ROIs and all of its descendant coefficients are outside the ROIs as well, the branch from this coefficient is pruned from the ROI coefficient tree. If a coefficient in the tree is inside the ROIs and its parent coefficient is outside the ROIs, called *pseudo-parent*, the coefficient is disconnected from its pseudo-parent. A new parent, which is inside the ROIs, is searched within the subband of its parent. The searching process is performed from the nearest coefficients to its farthest coefficients of its pseudo-parent. If new parent is found, called the real-parent, that coefficient is connected to its real-parent. After reconnection, there are only two kinds of coefficient trees left. One is with all coefficients in the tree inside the ROIs, the other with all coefficients in the tree outside ROIs. The resulting coefficient trees are not balanced any longer. A data structure, a list named LNP, is used to collect the isolated coefficients for which no parents in the lowest subband are found. The algorithm to construct unbalanced coefficient trees is described below:

```
For each coefficient (i, j, k) in the 3-D subband pyramid do:
    1.  If (i, j, k) is outside ROI, then
    1.1 If all descendants of (i, j, k) are outside ROI, then
          prune the (i, j, k) branch.
    2.  Else if (i, j, k) is inside ROI, then
    2.1 If parent of (i, j, k) is outside ROI, then disconnect
          (i, j, k) from its parent and look for a real parent
          within the neighbors of its parent.
    2.2 If a real parent is found, then connect (i, j, k) to its
          real parent else,
    2.3 Add (i, j, k) to the no parent list.
```

The proposed algorithm places the coefficients both in the lowest subband and inside ROIs into LIP (list of insignificant pixels) or LIS, and the coefficients in LNP are also appended to LIP or LIS. These coefficients, which are in the lowest subband but outside ROIs, are skipped when the SPIHT coding algorithm is initialized. The sorting pass is modified using unbalanced trees structure to get the descendents of a coefficient and refinement pass in our algorithm is the same as original SPIHT. With this method, coefficients not belonging to the ROIs are not included in the corresponding trees through the sorting pass procedure in the SPIHT algorithm.

V. EXPERIMENTAL RESULTS

Rate-distortion results are first presented for a MR brain volumetric dataset (256x192x60 at 8bits) obtained from the Ottawa General Hospital. In order to perform wavelet decomposition, the last slice of the MR sequence is duplicated four times and copied slices are padded at the end of the sequence to make the number of whole sequence power of 2, (*i.e.* 64). A ROI mask is created by generating a contour of the part of the brain. One of the slices from the original sequence with the segmented ROI are shown in Fig.1. It should be noted that the mask changes between the slices so that it fits the appropriate contour of the part of the brain in each slice. An issue involved with the 3-D wavelet coding is the sequence of the wavelet transform. For still images, the dyadic 2-D wavelet transform which performs one-dimensional wavelet transform along the two spatial dimensions alternately, is used. For 3-D data, with the asymmetric coefficient structure the wavelet packet transform is found to perform better than the dyadic sequence [3].

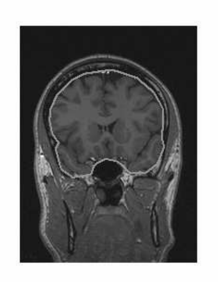

Fig.1. Pre-segmented slice number 10

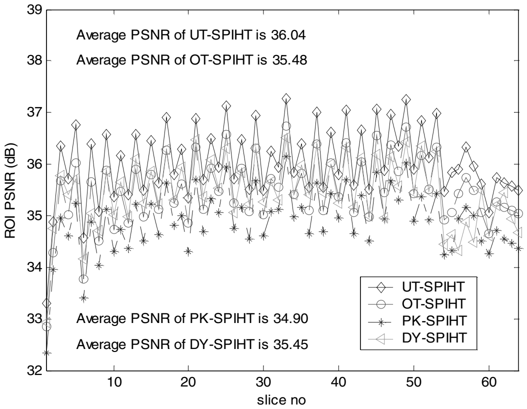

Fig.2. Comparison of ROI PSNRs at bitrate of 0.1 bpp

Fig.2 shows the R-D performance of the proposed algorithm (UT-SPIHT) at total bitrate of 0.1 bpp using 9/7 wavelet with 5 transform levels for all dimensions using packet decomposition. For reference, results obtained by other three algorithms, optimal tree (OT-SPIHT) [3], and original 3-D SPIHT using wavelet packet decomposition (PK-SPIHT) and dyadic decomposition (DY-SPIHT) [12] are also included. As can be seen the PSNR of ROI in each slice using UT-SPIHT algorithm is better than that of other algorithms with an improvement up to 1.14 dB. Using the same setting, this time at bit rate of 0.4 bpp, the average PSNR obtained with UT-SPIHT is 43.29 dB whereas OT-SPIHT has an average PSNR of 42.38 dB. The average PSNR obtained with DY-SPIHT (42.30dB) is close to that of OT-SPIHT at this bit rate. PK-SPIHT has an average PSNR of 41.97 dB.

The average ROI PSNRs for the four algorithms using the same setting as in Fig. 2, but this time at bit rates changing from 0.1 to 1.0 bpp are shown in Fig. 3. On average, UT-SPIHT algorithm outperforms the OT-SPIHT algorithm by 0.5 dB, DY-SPIHT algorithm by 1.0 dB, and PK-SPIHT algorithm by 1.1 dB.

We also evaluate the R-D performance of the proposed technique on other volumetric datasets. For CT_skull (256x256x203) dataset [12], at total bit rate of 0.1 bpp the average ROI PSNR is 33.16 dB (42.85 dB at 0.4 bpp) with the proposed technique while with the OT-SPIHT it is 32.47 dB (41.55 dB at 0.4 bpp). Average PSNR obtained with DY-SPIHT is 33.03 dB (41.83 dB at 0.4 bpp) and 32.03 dB (40.42 dB at 0.4 bpp) with PK-SPIHT. For the third dataset, MR_liver_t1 (256x256x58) [12], we observe similar trends; the proposed technique still performs better than the others. At total bitrate of 0.1bpp, the average ROI PSNR is 43.09 dB (56.17dB at 0.4 bpp) with the proposed technique whereas it is 42.31 dB (54.64 dB at 0.4 bpp) with OT-SPIHT. For this dataset, average PSNR obtained with DY-SPIHT is 41.98 dB (51.84 dB at 0.4 bpp) and 41.07 dB (53.22 dB) with PK-SPIHT.

A significant feature of the unbalanced tree structure is that it can be used to implement a multiple-region multiple-quality coding algorithm (MRMQ-SPIHT) to support compression of multiple ROIs at different qualities. Since other 3-D coding algorithms do not support encoding of multiple ROIs, we compare the R-D performance of the MRMQ-SPIHT algorithm with the dyadic 3-D SPIHT in this experiment. First, we consider that the image sequence is partitioned into two parts, the region that encloses the brain and the background. 3-D SPIHT is used to compress the entire image sequence and MRMQ-SPIHT algorithm is used to compress the ROI and non-ROI at different bitrates (quality). In this experiment, 90% of the bit budget is used to code the ROI and the remaining 10% is used for coding the background. Fig.4 shows the R-D performance of the two coding algorithms. As expected the ROI PSNR of MRMQ-SPIHT algorithm is better than that of original 3-D SPIHT algorithm and it quickly becomes nearly lossless after 0.8bpp. Of course, the background PSNR of the proposed algorithm is lower than that of conventional 3-D SPIHT, *e.g.* 30.21 dB with MRMQ-SPIHT and 39.62 dB with 3-D SPIHT at rate 1.0 bpp. Assuming that diagnostically significant information is included within the ROI, degradation in the remaining part of the image slices can be tolerated.

Fig.5 shows one of the reconstructed image slices using MRMQ-SPIHT algorithm (left) by bit allocation among the ROI and the background at a total bit rate of 0.2 bpp. For comparison, reconstructed image slices by 3-D SPIHT (right) at the same bit rate are also shown in Fig.5. Compared to the previous experiment, here we obtain a background reconstructed at a certain quality, which may be more preferable depending on the application. Comparing the two reconstructed slices in Fig.5, the ROI quality of reconstructed image slices by our proposed algorithm is much better than that of using SPIHT, while the background is degraded as only 10% of the total bit rate has been allocated for encoding the coefficients belonging to the background. At rate 0.2bpp, the average PSNR in the ROI is 38.88 dB with corresponding compression ratio of 40:1 using the MRMQ-SPIHT algorithm. If we wanted to achieve similar PSNR in ROI using 3-D SPIHT, we would have to encode this image sequence at a rate of 0.9 bpp. This corresponds to a compression ratio of only 9:1. The compressed data file using MRMQ-SPIHT method at rate 0.2bpp is 4.5 times smaller than that of 3-D SPIHT.

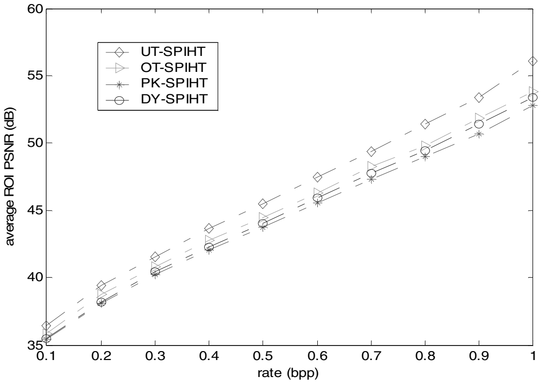

Fig. 3. Average ROI PSNR for different coding schemes.

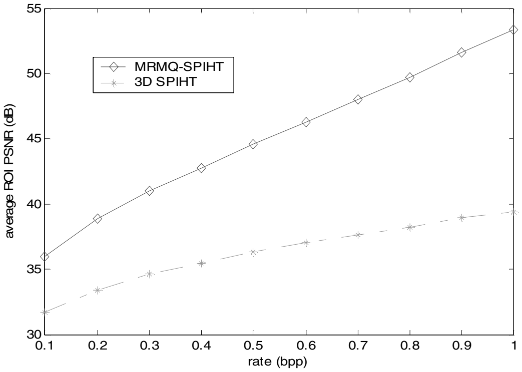

Fig. 4. Average ROI PSNR with 3-D SPIHT and the proposed coding scheme (MRMQ-SPIHT).

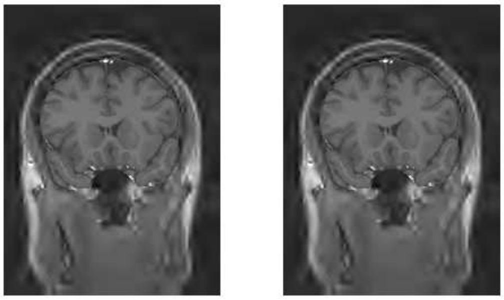

Fig. 5. Reconstructed image slice#10 using MRMQ-SPIHT algorithm (left) and 3-D SPIHT (right) at total bitrate of 0.2 bpp.

VI. CONCLUSION

In this paper we propose a 3-D unbalanced tree structure for region-based coding of volumetric datasets. With the proposed scheme, it is possible to compress multiple regions at various quality levels. This can be applied in medical imaging to offer better utilization of available bit rate, since high fidelity shall be maintained only for relatively small regions of diagnostic relevance rather than for the entire volume.

REFERENCES

[1] S. Li and W. Li, “Shape-Adaptive Discrete Wavelet Transforms for Arbitrarily Shaped Visual Object Coding”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 5, pp. 725-743, August 2000.

[2] G. Minami, Z. Xiong, A. Wang, and S. Mehrotra, “3-D Wavelet Coding of Video with Arbitrary Regions of Support”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 9, pp. 1063-1068, September 2001

[3] C. He, J. Dong, Y. F. Zheng, Z. Gao, “Optimal 3-D Coefficient Tree Structure for 3-D Wavelet Video Coding”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 10, pp. 961-972, October 2003

[4] Ali Bilgin, G. Zweig, and M. W. Marcellin, “Three-dimensional image compression with integer wavelet transforms”, *Applied Optics*, vol. 39, pp. 1799-1814, April 2000

[5] B. Kim, Z. Xiong, and W. A. Pearlman, “Low Bit-rate Scalable Video Coding with 3-D SPIHT”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp. 1374-1387, December 2000

[6] G. Ginesu, D. Giusto, W. A. Pearlman, “Lossy to Lossless SPIHT-based Volumetric Image Compression”, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 693-696, May 17-21, 2004

[7] Z. Xiong, X.Wu, S. Cheng and J. Hua, “Lossy-to-Lossless Compression of Medical Volumetric Data Using Three dimensional Integer Wavelet Transforms,” *IEEE Trans. on Medical Imaging*, vol. 22, no. 3, pp. 459-470, March 2003

[8] D. Agrafiotis, D. R. Bull and N. Canagarajah, “Region of Interest Coding of Volumetric Medical Images”, *IEEE International Conference on Image Processing*, vol. 3, pp. 217-220, September 2003

[9] A. Cuhadar and S. Tasdoken, “Multiple, arbitrary shape ROI coding with zerotree based wavelet coders”, *IEEE International Conference on Acoustics, Speech, and Signal Processing,* vol. III, pp. 257-260, April 6-10, 2003

[10] S. Tasdoken and A. Cuhadar, “Quadtree-based multiregion multiquality image coding”, *IEEE Signal Processing Letters*, Vol. 11, No. 2, pp. 101-103, February 2004

[11] J.E. Fowler, “QccPack: An open-source software library for quantization, compression and coding,” *SPIE Applications of Digital Image Processing XXIII*, vol.4115, pp. 294-301, 2000

[12] Rensselaer Polytechnic Institute, Center for Image Processing Research
http://www.cipr.rpi.edu/resource/sequences/index.html