Prioritization of Compressed Data by Tissue Type Using JPEG2000

Karthik Krishnan[a], Michael W. Marcellin[a], Ali Bilgin[a], Mariappan Nadar[b]

[a]Signal Processing and Coding Laboratory, Department of Electrical and Computer Engineering, 1230 E. Speedway Blvd, University of Arizona, Tucson, AZ, USA 85721;
[b]Siemens Corporate Research Inc. Princeton, NJ, USA 08540

ABSTRACT

One of the goals of telemedicine is to enable remote visualization and browsing of medical volumes. Volume data is usually massive and is compressed so as to effectively utilize available network bandwidth. In our scenario, these compressed datasets are stored on a central data server and are transferred progressively to one or more clients over a network. In this paper, we study schemes that enable progressive delivery for visualization of medical volume data using JPEG2000. We then present a scheme for progressive encoding based on scene content, that enables a progression based on tissues or regions of interest in 3D medical imagery. The resulting compressed file is organized such that the tissues of interest appear in earlier segments of the bitstream. Hence a compliant decoder that chooses to stop transmission of data at a given instant would be able to render the tissue of interest with a better visual quality.

Keywords: Telemedicine, JPEG2000, embedded coding, scalable compression, medical volume compression, volume of interest, interactive visualization, DWT.

1. INTRODUCTION

Picture Archiving and Communication Systems (PACS) as a field is gaining popularity. The increasing usage of the internet and the popularity of online medical volume databases, such as those maintained by the National Library of Medicine (NLM) have spawned research dealing with techniques for interactive transmission of compressed data for volume visualization. Typically, data sets are stored and maintained by a database server, so that one or more remote clients can browse the datasets interactively and render them as in Fig. 1. An interactive user may be willing to initially sacrifice some rendering quality or field of view in exchange for real-time performance. For the system to be responsive, the time taken for a rendered image to appear on the screen should be small. The client should also be provided some breadth in terms of interactivity (such as reduced resolution viewing, ability to view a select subsection of the volume, pan, zoom, view select slices, etc.) and a pleasant viewing experience (progressive refinement of the view volume, etc.).

Further author information: (Send correspondence to Ali Bilgin)
Ali Bilgin: E-mail: bilgin@ece.arizona.edu, Phone: +1 520 626 8943, Fax: 1 520 6218076

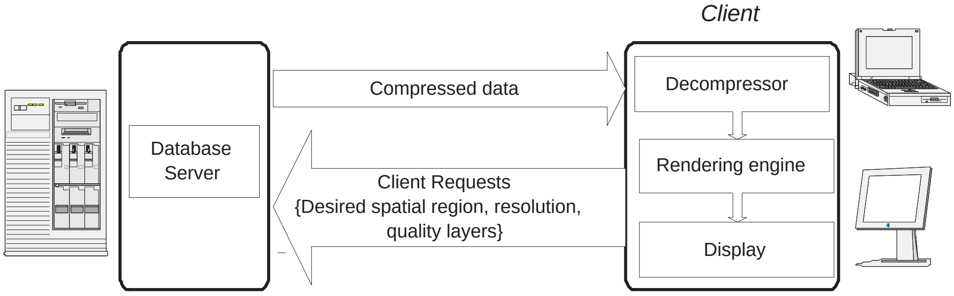


Figure 1. Client-Server model

Medical Imaging 2005: PACS and Imaging Informatics, edited by Osman M. Ratib, Steven C. Horii, Proceedings of SPIE Vol. 5748 (SPIE, Bellingham, WA, 2005)
1605-7422/05/$15 · doi: 10.1117/12.594914

1.1. Compression

Volumetric data sets are massive. For example, the Visible Male data set consists of about 15 gigabytes of volume data. If such data sets need to be transmitted over low-bandwidth networks with varying loads and latency constraints, efficient compression schemes must be employed. Compression may be lossy or lossless. Acceptable rendering quality can generally be obtained at compression ratios of 5-100:1. However for diagnostic purposes it is often required that the medical volume data be compressed losslessly. It would therefore be desirable for the compression scheme to support progressive encoding, so that the volume can be decoded from an initial lossy to, (if required) a final lossless representation.

Reseachers have used a variety of compression schemes for volume compression. JPEG compression was used in[1] . Chiueh et al[2] use a similar scheme, where the volume is divided into bricks which are frequency transformed using Fourier transform, quantized and entropy coded. Rendering is carried out in the frequency domain. However, compositing techniques across bricks generally lead to blocky atrifacts.

Hierarchical compression techniques (wavelet, tree based schemes such as octree, zero trees etc), are ideal for multi-resolution access. A 3D reversible integer wavelet transform combined with Embedded Zero-tree Wavelet (EZW) coding[3] has been used in[4] to achieve lossless compression. 3D-Set Partitioning in Hierarchical Trees (SPIHT)[5] has been employed in[6] and[7] . Lippert et. al[8] use wavelet decomposition for progressive rendering and transmission over networks. A comparison of popular wavelet based 3D coders can be found in[9] .

In our research, we have used the JPEG2000[10] coder. JPEG2000 is the new scalable image compression standard designed to support a variety of applications, including the compression and transmission of medical images. In November 2001, JPEG2000 was selected for inclusion in the DICOM standard for medical image transfer and the DICOM protocol now includes JPEG2000 transfer syntaxes. In[12] , we have used JPEG2000 and JPEG2000 Interactive Protocol (JPIP)[11] to progressively transmit volume data over networks.

2. A BRIEF INTRODUCTION TO JPEG2000

JPEG2000 is a highly scalable image compression standard[14] and offers many progression orders that facilitate interactive access of image segments. Among them are resolution scalability, distortion scalability and spatial random access. It supports both lossy and lossless compression and unlike its DCT based predecessor, JPEG, it allows for an initial lossy transmission to a complete lossless refinement of the data.

JPEG2000 Part I employs the Discrete Wavelet Transform (DWT). Each level of the DWT decomposes its input into four spatial frequency subbands, LL_d, LH_d, HL_d and HH_d. A D_{xy} level DWT decomposes an image into 3 D_{xy}+1 subbands. In our particular example, we are dealing with medical volumes, which are typically acquired as a sequence of axial slices. Depending on the slice thickness, there can be significant correlation in the slice direction. We exploit this correlation by applying a DWT in the z direction as specified in JPEG2000 Part 2[15] prior to the $x - y$ DWT. We refer to the "slices" that result from the z direction transform as "transformed components." These transformed components are further decomposed (by xy DWT) to form subbands as illustrated in Fig. 2.

2.1. Code-Blocks and Coding Passes

Each subband of each transformed component is partitioned into rectangular blocks, known as *code-blocks*. Hence a code-block is a collection of wavelet coefficients. The wavelet coefficients, in their binary representation (sign magnitude) can be thought of as a collection of bit-planes. Binary arithmetic coding of bit-planes is performed independently on each code-block. The quality of reconstruction for a code-block is controlled by the number of bit-planes included in its bitstream. Every bit-plane except the most significant bit-plane is encoded in three different passes, each of which is referred to as a coding pass. Hence a *coding pass* can be understood as a fractional bit-plane. A coding pass represents a code-block's incremental quality contribution to the codestream.

Multi-resolution access is a direct consequence of the wavelet transform. Distortion scalability is due to coding passes. Spatial random access is possible because of the finite extent of the wavelet filters and due to the independent encoding of code-blocks.

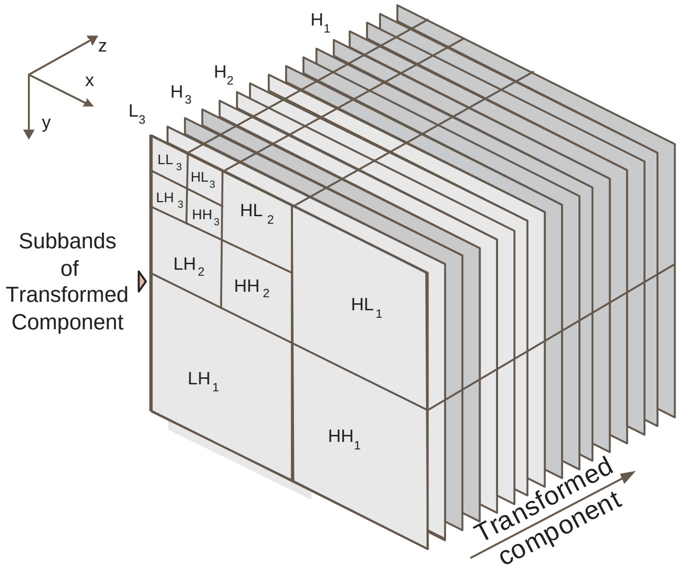


Figure 2. DWT of the volume into subbands.

2.2. Rate Allocation

Every coding pass incrementally contributes to the volume data. This quality increment may be quantified by the MSE (Mean Square Error) improvement, $\Delta\xi$, its inclusion makes to the volume. The typical rate allocation algorithm used in JPEG2000 seeks to maximize the distortion reduction achieved by including end of certain coding passes. Given two coding passes in two different code-blocks, this usually implies including them in the decreasing order of their distortion contributions per unit length, $\Delta\xi/\Delta l$, where Δl is the length of the coding pass. Hence the Distortion-Rate slope of a coding pass, $\Delta\xi/\Delta l$, plays a pivotal role in the order in which coding passes are included in the bitstream.

3. PRIORITIZED ENCODING BY TISSUE TYPE

3.1. Volume Visualization

The fundamental volume visualization algorithms are of two types: *direct volume rendering (DVR)* algorithms and *surface-fitting (SF)* algorithms. DVR includes approaches such as ray-casting[16] and splatting.[17] The disadvantage with DVR is that the entire dataset must be traversed for every image rendered from a viewpoint. Sometimes a low resolution image is quickly created to check the volume and then refined, also called “progressive refinement.”

SF methods fit *iso-surfaces* (planar polygons or surface patches) to constant-value contour surfaces. SF methods are usually faster than DVR methods since they traverse the dataset once, for a given threshold value. New views of the surface can be quickly generated. However, using a new SF threshold is time consuming since the cells must be revisited to extract new surfaces. Moreover SF methods fail to account for data originating from fluid and other materials which may be partially transparent and should be modeled as such.

Data classification involves classifying the data into relevant and irrelevant regions. This involves choosing a scalar threshold value for SF algorithms and opacity values for DVR algorithms. An opacity transfer function is used to expose the relevant segments and to make transparent the uninteresting ones. For data acquired from CT scanners, the pixel intensity generally determines the tissue type. Hence, scalar voxel values play a key role in determining the compressed data of interest.

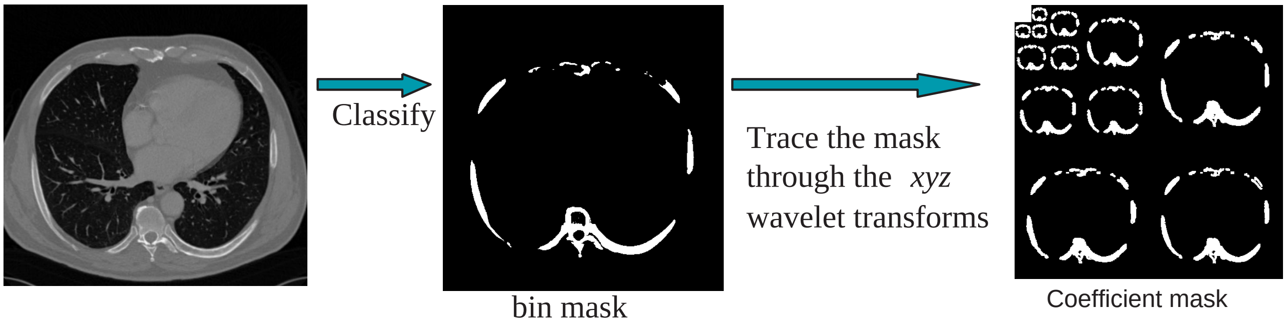


Figure 3. An axial slice, its *bin mask* obtained after classification and the coefficient mask (the set of wavelet coefficients required to reconstruct the bin-mask).

3.2. Identifying Scalar Values in the Code-Stream

We divide the scalar range of the volume data into *bins*, where each bin represents the intensity range to which a particular tissue is confined. To allow for interpolation between voxel vertices and computation of gradients for purposes of rendering, we dilate the bin masks by a few voxels. Bins are generally not uniformly spaced. Each voxel is assigned a particular bin, or may remain unassigned. We then proceed to identify compressed volume data corresponding to each bin.

For each bin, a *bin mask* is generated. The bin mask is the set of voxels that belong to the bin. Due to the finite spatial extent of the DWT, it is possible to trace each voxel to a set of wavelet coefficients using the footprint of the wavelet kernel used to transform the volume data. This can be done recursively from each resolution to the next lower resolution. In this way, each bin mask, which is defined on the volume data grid, is mapped to a mask on the subband grid. This process is illustrated in Fig. 3 for a binary classification of the volume into bone and non-bone.

3.3. Introduction to Region-of-Interest (ROI) Based Schemes

JPEG2000 supports ROI coding that allows certain regions of the image to be encoded with higher fidelity than the rest of the image. ROI coding schemes supported by JPEG2000 are of two kinds: *MAXSHIFT* method as defined in Part 1 of the JPEG2000 standard[18] and the *arbitrary scaling* based method[19] as defined in Part 2 of the standard.[15]

ROI coding is accomplished by shifting the wavelet coefficients that contribute to the foreground to higher bit-planes. This results in the foreground being encoded prior to the background. JPEG2000 Part 2 allows shape information (of the foreground) to be encoded into the compressed codestream as long as the shape is rectangular or elliptical. The shift, ς, controls the relative importance of the foreground and the background. Arbitrary values of ς are supported by JPEG2000 Part 2 as long as the ROI shape is constrained to be rectangular or elliptical. The disadvantage of the "arbitrary scaling" method is that arbitrary shapes are not supported.

Arbitrary shapes of foreground are enabled by the *MAXSHIFT* method, which chooses a shift, ς, such that shape information need not be transmitted to the decoder. This requires ς to be sufficiently large so that the foreground and the background wavelet coefficients can be distinguished based upon the decoded quantization indices alone. This requires

$$\varsigma \geq \max_s \kappa_s^{max} \tag{1}$$

where κ_s^{max} is the maximum number of bitplanes required to represent background wavelet coefficients in sub-band, s.

Hence the encoder and decoder must support a bit-plane precision, $\wp$, given by

$$\wp_{maxshift} = \kappa_s^{max} + \varsigma \tag{2}$$

An obvious disadvantage of the MAXSHIFT method is the bit-plane precision that must be supported, which may be large for 12 bit medical imagery (33 - 34 bit-planes). Most software implementations represent wavelet coefficients using 32 bit integers. Hardware implementations generally provide very limited bit-plane precisions. If lower precision is employed, this would necessitate discarding bit-planes from the background, resulting in loss of information. Hence generally, it is not possible to use ROI coding when lossless compression of medical imagery is desired. If lossy compression is sought, it may be possible to vary the quantization step sizes so as to restrict the number of background bit-planes to a value within the encoder's bit-plane precision. Lossless compression however mandates the use of a quantization step size of 1.

A second disadvantage is the lack of flexibility in obtaining a tradeoff between the foreground and the background quality. Some tradeoff may however be obtained, by including the $LL_{D_{xy}}$ band or a few low resolution subbands in the foreground. This scheme may be employed in our scenario as well.

The above mentioned schemes are easily extended to volume data. The foreground mask needs to be traced through the DWT along the slices prior to being traced through the DWT along the $x - y$ directions.

3.4. The Prioritized Encoding Scheme

Given a bin or set of bins to be emphasised, we first determine its coefficient mask as described earlier. We then identify code-blocks that have a non-zero intersection with the coefficient mask. This forms a code-block mask, comprising the set of code-blocks that contribute to the reconstruction of the bin. The rate allocation algorithm computes length and the change in distortion, $(L,\ \Delta\xi)$ at the end of each coding pass, in each code-block. Prior to entropy coding, we multiply the distortion metrics of the relevant code-blocks by a weight, w_b. Hence relevant code-blocks get included in earlier segments of the bitstream. Hence a bitstream progressively decoded or truncated at a particular point will have larger quality contributions from the bin as opposed to outside the bin. By choosing the weight, w_b appropriately, it is possible to obtain a tradeoff in quality within and outside the bin.

The disadvantage of the preemphasis scheme as opposed to MAXSHIFT is that the mask constructed using code-blocks is a poorer definition of the foreground, since emphasis is given to entire code-blocks rather than individual subband samples. Hence MAXSHIFT based ROI coding is capable of providing greater quality differences between the foreground and the background. However, with an appropriate choice of w_b, preemphasis can be used to obtain a wider range of flexibility between the foreground and background quality. Moreover, preemphasis does not suffer from bit-plane precision effects. Specifically, for 12 bit medical volume data, MAXSHIFT based ROI cannot achieve a lossless representation of the background, while pre-emphasis can.

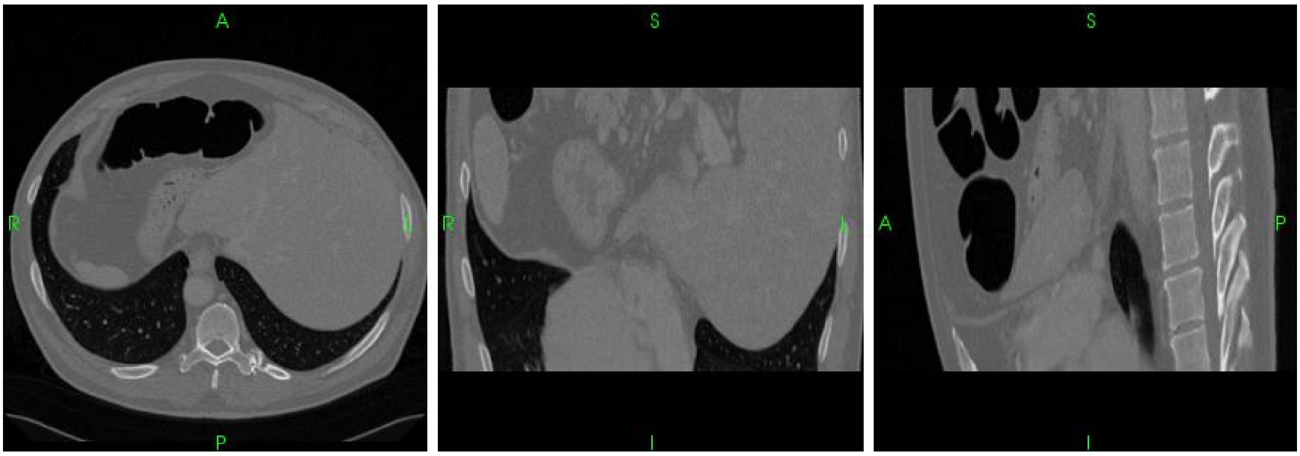
4. RESULTS

We compressed a $512 \times 512 \times 550$ section of a CT scan of the abdomen with the reversible integer (5,3) wavelet using three methods. First, we compressed the volume using a MAXSHIFT ROI scheme. The file could not be reconstructed losslessly. Specifically, the maximum background Peak Signal to Noise Ratio (PSNR) attainable was 25 dB. The minimum value of the shift, ς, needed was 16. We then compressed the volume using the weighting scheme applied to voxels classified as bone. Since the mask is constructed using code-blocks, their size was reduced to 8×8. The lowest resolution, LL_7 subband was in its entirity included in the foreground so that at least a crude estimate of the volume could be decoded at high compression ratios. For purposes of comparison, we also compressed the volume without any prioritization and used a code-block size of 32×32.

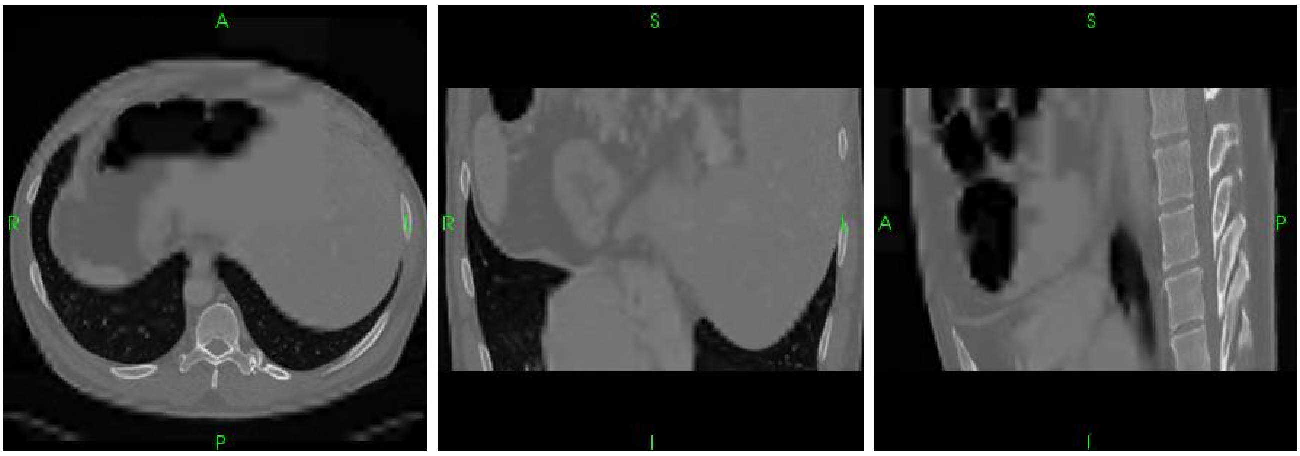
We then decompressed only the initial portions of each file at bit-rates that corresponded to compression ratios varying from 50:1 to 500:1 and compared the resulting lossy outputs. An axial, coronal and a sagittal slice from the decompressed dataset is shown in Fig 4. The PSNR values using each scheme are tabulated in Table 1.

The volumes decompressed at 500:1 were rendered*. In our scenario, we use a step transfer function accompanied by volume rendering with gradient opacity mapping, followed by shading. Linear interpolation is used.

*An encoding scheme which encodes the foreground and background in seperate files may not be a suitable approach since this would necessitate encoding of region information, which often, as seen from Fig. 5 is not well confined. Hence encoding the region information may not be a trivial task.



(a)



(b)

Figure 4. (from top to bottom): An axial, coronal and sagital slice decompressed at a ratio of 20:1 using the two methods: (a) Default compression, (b) Compression applying a pre-emphasis to bone. The reduction in quality in regions away from the bone in the lower figure indicates that more bits have been allocated to the regions in the vicinity of the bone, although it is difficult to perceive the increase in quality in the bone region due to lack of sufficient texture. The value of w_b used was 1024.

Table 1. PSNR (dB) using each of the coding schemes. Two weights w_b=1024 and w_b=8 have been chosen to illustrate the tradeoff that can be obtained by varying the weights.

		Compression Ratio					
		50:1	100:1	200:1	300:1	400:1	500:1
Default	Bone	43.27	40.36	37.35	35.92	34.33	33.04
	Rest of the volume	46.70	44.75	43.08	42.03	41.12	40.43
Pre-emphasized Encoding w=1024	Bone	49.85	44.57	40.63	38.39	37.11	35.59
	Rest of the volume	30.64	30.63	30.61	30.60	30.58	30.57
Pre-emphasized Encoding w=64	Bone	48.89	43.84	40.09	37.82	36.76	35.15
	Rest of the volume	43.76	41.48	39.49	38.10	36.74	36.04

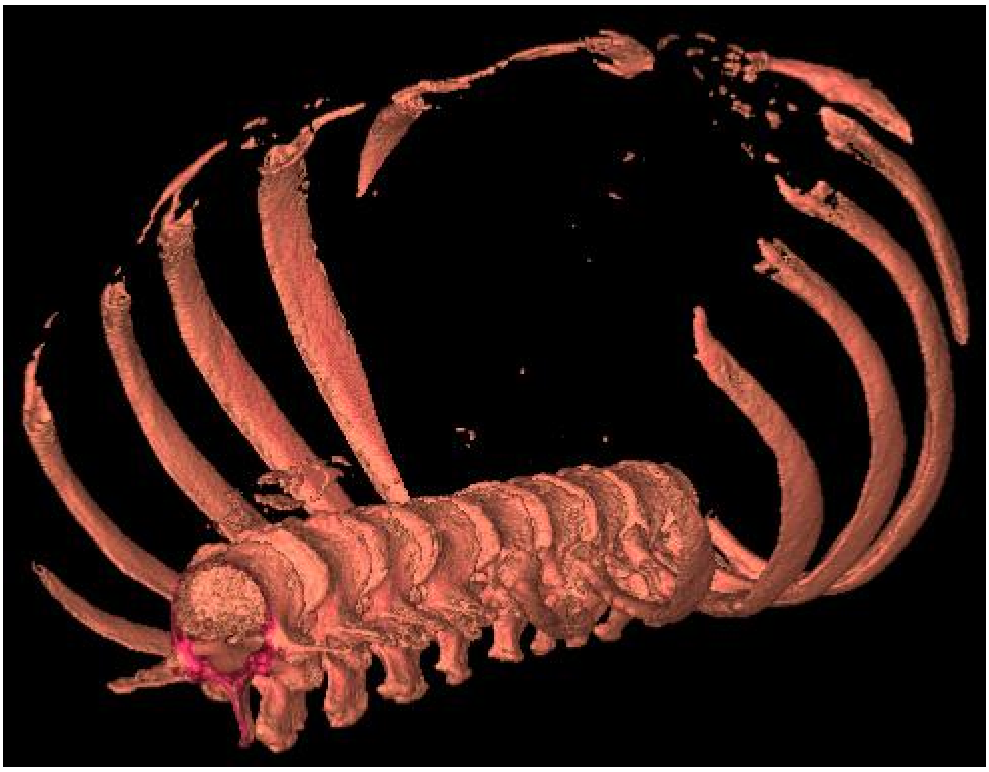


Figure 5. Bone rendered from a 512 × 512 × 550 section of the abdomen. (Uncompressed dataset)

Fig. 5 is rendered from the original dataset. Fig. 6 is rendered from the pre-empahsized file decompressed at 500:1. Fig. 7 is rendered from a file compressed using the default parameters, without any such region based coding.

5. CONCLUSIONS

We have described a scheme that encodes the volume progressively based on scene content using the multi-resolution framework of JPEG2000, which is ideal for transmission over networks for remote visualization. We have examined the shortcomings of ROI based coding supported by JPEG2000 and we have suggested an alternative approach. This approach also makes it possible to obtain a tradeoff in visual quality between the foreground and the background. This scheme is well suited for applications where the intent (such as visualization parameters or the classification process that segments the data as more relevant and less relevant) is known *a priori*; nevertheless, it may not be advisable to discard background data which may later be needed for additional diagonosis. In such scenarios, with this scheme, it is possible to progressively encode volume data based on relevance for purposes of storage or for transmission over networks.

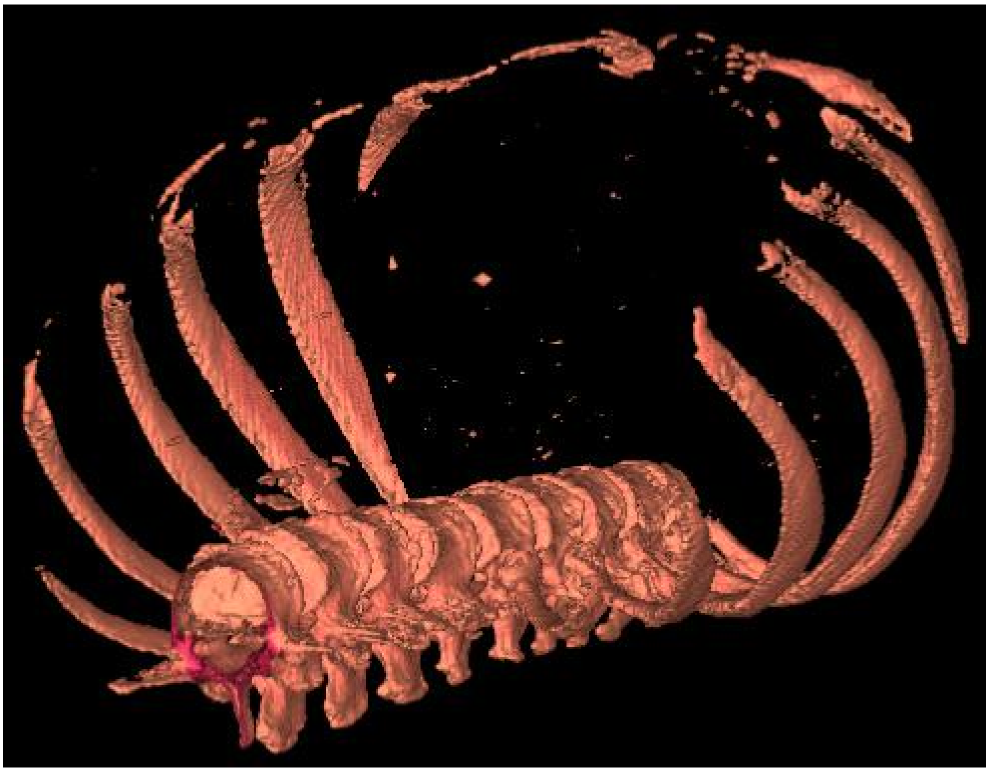


Figure 6. Bone rendered from the preemphasized file decompressed at 500:1. This file was compressed with a preemphasis weight of 1024 applied to bone and to the lowest resolution LL_7 band.

REFERENCES

1. B. Yeo, B. Liu,"Volume rendering of DCT-based compressed 3D scalar data," *IEEE Trans. Visualization and Computer Graphics,* vol. 1, no. 1, pp. 29-43, Mar. 1995.
2. T. Chiueh, C. Yang, T. He, H. Pfister, A. Kaufman,"Integrated volume compression and visualization," *Proc. IEEE Visualization '97,* pp. 329-336, Oct. 1997.
3. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing,* vol. 41, no. 12, pp. 3445-3462, Dec. 1993.
4. A. Bilgin, G. Zweig, M. W. Marcellin, "Three-dimensional compression using integer wavelet transforms," *Applied Optics: Information Processing, Special Issue on Information Theory in Optoelectronic Systems,* vol. 39, no. 11, pp. 1799-1814, Apr. 2000.
5. A. Said, W. A. Pearlman, "A new fast and efficient image codec based on Set Partitioning in Hierarchical Trees," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 6, pp. 243-250, June 1996.
6. Y. Kim, W. A. Pearlman, "Lossless volumetric medical image compression," *Proc. SPIE Conf. Applications of Digital Image Processing XXII,* vol. 3808, pp. 305-312, July 1999.
7. Z. Xiong, X. Wu, D. Yun, W. A. Pearlman, "Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transform," *Proc. IEEE Second Workshop on Multi. Signal Processing,* pp. 553-558, Dec. 1998.
8. L. Lippert, M. H. Gross, C. Kurmann, "Compression domain volume rendering for distributed environments," *Proc. Eurographics '97, COMPUTER GRAPHICS forum,* vol. 14, no. 3, 1997.
9. P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, J. Cornelis, "Wavelet coding of volumetric medical datasets," *IEEE Trans. Medical Imaging,* vol. 22, no. 3, pp. 441-458, Mar. 2003.
10. D. Taubman, M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice.* Massachusetts: Kluwer Academic Publishers, 2002.

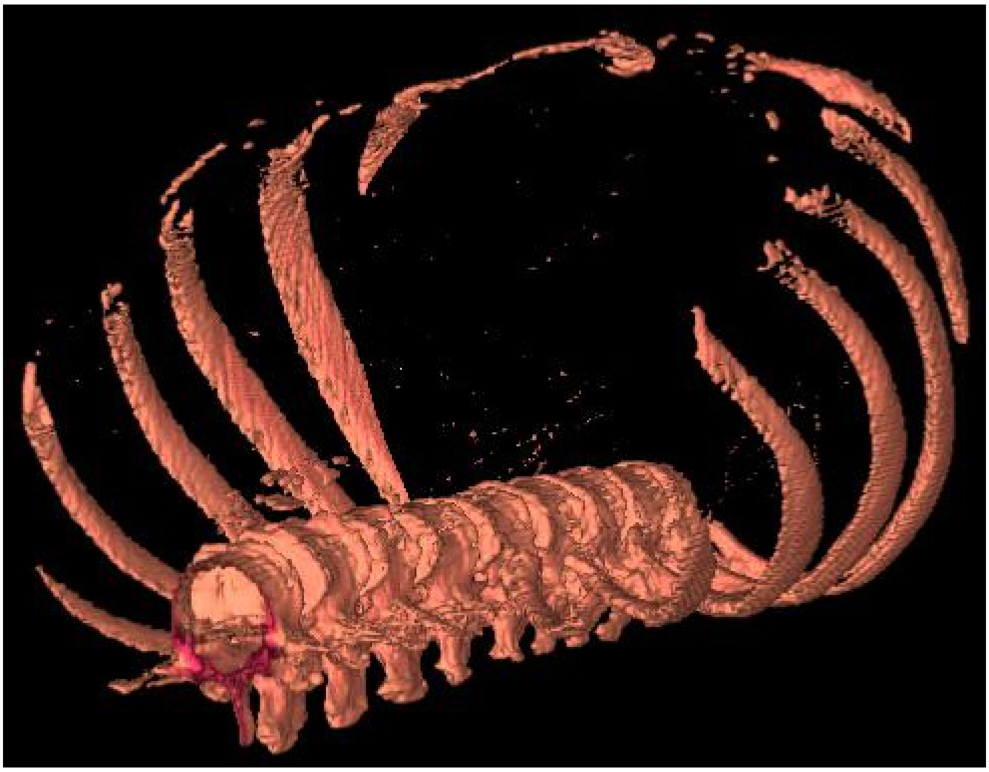


Figure 7. Bone rendered from a volume decompressed at 500:1 from a file compressed without any preemphasis.

11. D. Taubman, R. Prandolini, "Architecture, Philosophy and Performance of JPIP: Internet Protocol Standard for JPEG2000," *International Symposium on Visual Communications and Image Processing*, 2003
12. K. Krishnan, M. W. Marcellin, A. Bilgin, M. S. Nadar, "Compression / Decompression strategies for large volume medical imagery," *Proc. SPIE Medical Imaging: PACS and Imaging Informatics*, vol. 5371, San Diego, California, Feb. 2004.
13. D.Taubman, R. Leung, A. Secker, "Scalable compression of volumetric images," *Proc. IEEE Intl. Conf. Image Processing*, vol. 2, pp. 22-25, Sept. 2002.
14. D. Taubman, "High performance scalable image compression with EBCOT," *Proc. IEEE Int. Conf. Image Proc.*, vol. 3, pp. 344-348, Sept 2002.
15. Information technology - *JPEG2000 image coding system - Part 2.* FDIS, ISO/IEC JTC1/SC29 WG1, 15444-2 Std., July 2002.
16. M. Levoy, "Display of surfaces from volume data," *IEEE Computer Graph. and Applications*, vol. 8, no. 3, pp. 29-37, Mar. 1988.
17. D. Laur, P. Hanrahan, "Hierarchical splatting: A progressive refinement algorithm for volume rendering," *Computer Graphics SIGGRAPH*, vol. 25, no. 4, pp. 285-288, July 1991.
18. Information technology - *JPEG 2000 image coding system - Part 1: Core coding system,* ISO/IEC JTC1/SC29 WG1, 15444-1 Std., July 2002.
19. G. Anastassopoulos, A. Skodras, "JPEG2000 ROI coding in medical imaging applications," *Proc. Second IASTED Intl. Conf. on Visualization, Imaging and Image Processing*, pp. 9-12, Sept. 2002.