3D/2D OBJECT-BASED CODING OF HEAD MRI DATA

G. Menegaz and L. Grewe

Audiovisual Communications Laboratory
Swiss Federal Institute of Technology,
CH–1015 Lausanne
email: Gloria.Menegaz@epfl.ch

ABSTRACT

We propose a coding system featuring 3D encoding/2D decoding object-based functionalities. Any object of any 2D image of the dataset can be recovered at a finely graded up to lossless quality. Compression is improved by exploiting the full correlation among data samples by means of 3D DWT. A swift access to 2D images is obtained by enabling 2D decoding. Given the index of the image of interest along the z axis, only the concerned portion of the bitstream is decoded, at the desired quality. The selective access to data can be improved by splitting the image in regions corresponding to the different objects. Then, a suitable ordering of the encoded information within the bitstream enables random access to any object at the desired rate. This enables a pseudo-lossless regime, where the diagnostically relevant parts of the image are represented without loss, while a lower quality is assumed to be acceptable for the others. Results show that the proposed system is a good compromise between the gain in compression efficiency provided by 3D systems and the fast access to the data of 2D ones.

1. INTRODUCTION

The huge amount of data produced every day in the clinical environment pushes toward the definition of *ad-hoc* compression techniques. Our solution consists in combining 3D encoding/2D decoding capabilities with Region Of Interest (ROI) based functionalities. The agreement of the image processing community on object-based approaches is proved by the fact that JPEG2000 [1] features ROI based capabilities. However, volumetric medical data, which are the subject of our investigation, are out of the scope of the baseline JPEG2000.

The peculiarity of the proposed architecture is the combination of a fully 3D transformation with an encoding policy allowing to access and reconstruct any object of any 2D image of the dataset at the desired bitrate, without reconstructing the entire volume. The subband coefficients that are needed to recover a given object of the 2D image of index $\overline{z}$ are selectively accessed and decoded. This minimizes the decoding time of the object of interest and represent a good trade off between the gain in compression efficiency provided by 3D versus 2D system and a fast access to the data.

At the encoder, the data are first decorrelated by a 3D Discrete Wavelet Transform (DWT) and then coded by the Multidimensional Layered Zero Coding (MLZC) technique [2]. The implementation of the DWT via the lifting steps scheme [3] is particularly advantageous in this framework. First, it provides a very simple way of constructing non-linear wavelet transforms mapping integer-to-integer values [4]. This is very important for medical applications because it enables lossless coding. Second, perfect reconstruction is guaranteed for any kind of signal extension along borders. This greatly simplifies the management of the boundary conditions and facilitates the selection of the coefficients needed to reconstruct a subset of data samples (*e.g.* a 2D image or an object). Third, it is computationally efficient. It can be shown that the lifting steps implementation asymptotically reduces the computational complexity by a factor 4 with respect to the classical filterbank implementation [5]. Finally, the transformation can be implemented in-place, namely by progressively updating the values of the original samples, without allocating auxiliary memory. The 3D transform is followed by bitplane encoding via context adaptive arithmetic coding. Some markers are placed in the codestream to enable the random access to the encoded information. By combining the 3D-DWT with 2D spatial neighborhoods for context adaptive arithmetic coding, the resulting MLZC algorithm features 3D encoding/2D decoding capabilities.

At the decoder, the set of wavelet coefficients necessary to reconstruct an image of index $\overline{z}$ is automatically determined and only the related information is decoded. The Inverse DWT (IDWT) is performed locally, reducing the memory requirements and the computational cost.

The object based functionalities are integrated by independently handling the different objects. Some extra information must be encoded for any object in order to avoid artifacts along the borders in lossy regime. This generates an overheading of the bitstream which slightly degrades the compression efficiency with respect to the case where the volume is encoded as a whole.

This paper is organized as follows. Section 2 gives an overview of the proposed system. Results are discussed in Sec. 3, and conclusions are derived in Sec. 4.

2. 3D/2D OBJECT-BASED MLZC

The 3D/2D MLZC coding system presented in [6] combines the high coding efficiency resulting from the decorrelation properties of the 3D DWT with the swift access to the 2D images of the dataset. We propose here an extension of such a system integrating ROI-based functionalities. The resulting system permits random access to any object of any 2D image of the dataset at the desired up to lossless quality [7].

2.1. Object-based Coding

Medical images usually consists of a region representing the part of the body under investigation (i.e. the heart in a CT or MRI chest scan, the brain in a head scan) on an often noisy background with no diagnostic interest. It seems thus very natural to process such data in a object-based framework: assign high priority to the semantically relevant object, to be represented with up to lossless quality, and lower priority to the background. Artifacts along borders in lossy regime are avoided by ensuring that the boundary samples are processed as if the entire set of transformed coefficients were available. Our system provides decoded images that are exactly the same in the cases: (a) the signal has been encoded/decoded as a whole and (b) each object has has been independently encoded and decoded at a given quality (*e.g.* quantization level). The perfect reconstruction condition is not enough to ensure the absence of artifacts - in terms of discontinuities at borders. Since quantized coefficients are approximations of true values, any signal extension used to reconstruct two adjacent samples belonging to different objects (*e.g.* lying at the opposite sides of a boundary) would generate a discontinuity. To avoid this, the inverse transform must be performed *as if* the whole set of true coefficients were available. The use of the lifting scheme simplifies this task. The idea is to determine which samples are needed at the *input* of the synthesis chain to reconstruct a given sample at its *output*. The key of the proposed solution is to start at the finest resolution ($l = 1$) and select the set of wavelet coefficients which are needed (in each subband) to reconstruct the object in the signal domain (full resolution). Due to the recursiveness of the IDWT, the approximation subband of level $l = 1$ becomes the reference (*critic*) set of samples that must be reconstructed without loss, and so on. By going through all the resolutions and successively iterating the procedure as described for $l = 1, \ldots, L$, the set of concerned wavelet coefficients (*Generalized Projection (GP)*), is selected. We refer to [8, 7] for further details.

In the encoding step, each object is then assigned a portion of the bitstream, which can be independently accessed and decoded. The organization of the encoded information in each bitstream segment ensures a finely graded rate control on the associate object. The random access to any 2D image of the dataset is obtained by the Multidimensional Layered Zero Coding strategy (MLZC). The integration of the object-based functionalities enables the independent decoding of any object of any 2D image of the dataset at the desired bitrate.

2.2. 3D/2D object-based MLZC

The Layer-Per-Layer Progressive (LPL-PROG) operation mode is used for the MLZC coding scheme [6]. This is obtained by adding some markers in order to enable random access to the information of interest in the bitstream and allows direct access to every quantization layer of every subband image [1]. Quality scalability is obtained by successively decoding the quantization layers of the subband images necessary for the reconstruction of the image of interest. Given the index of the image of interest within the dataset (z coordinate in fig. 1), the concerned portions of the bitstream are accessed and decoded. More details can be found in [8, 6]. ¿From a compression point of view, the drawback is the bitstream over-heading due to the additional information needed for data addressing.

Very little modifications are needed to adapt the MLZC system for object-based processing. Each coefficient is encoded only whether it belongs to the generalized projection of the considered object. Figure 2 illustrates the bitstream structure. In the figure, H

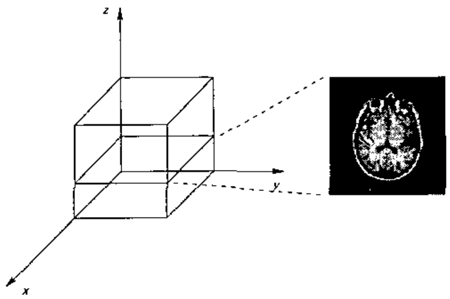

Fig. 1. Volumetric data. We call z the third dimension, and assume that the images are the intersections of the volume with a plan orthogonal to z axis.

is the bitstream header, L_i^j is the i^{th} bitplane, $i = 0, \ldots, I-1$ of the subband image at position j in a given 3D subband. A bitplane coding strategy is followed. The bitplane $\bar{i}$ is encoded for each subband image ($j = 0, \ldots, J-1$, J being the depth of the considered 3D subband) before switching to the next $\bar{i} - 1$ bitplane. Markers separate information concerning the different subband images as well as successive bitplanes.

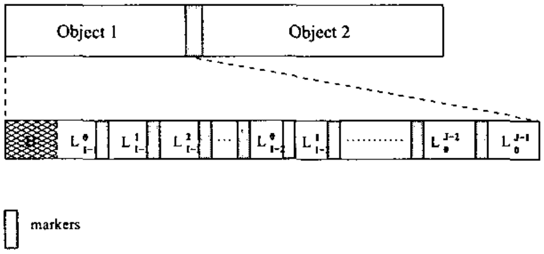

Fig. 2. Structure of the bitstream for the 3D/2D object-based MLZC coding scheme.

3. RESULTS AND DISCUSSION

In the framework of ROI-based coding, semantics is exploited for the allocation of the resources (*e.g.* bit-budget, bandwidth). The efficiency improvement is thus to be intended in the sense of *prioritization* of the information to be transmitted. The impact of the non-relevant information must not be underestimated. This is particularly important when dealing with medical images, where the background often encloses the majority of the voxels. For a typical MRI dataset for example, about the 90% of the voxels are taken by the background. It is thus of prime importance to classify them *a-priori* in order to assign higher priority to the ROI.

The performances of the 3D/2D Object-based MLZC coding system have been compared to those of the JPEG2000 [1] standard in the ROI-based mode. Both the lossless and lossy regimes have been considered. The test set consists of two MRI head scans (saggital view). The first one (MRI) is made of 256 × 256 × 128 voxels. The second (MIR-MRI) [2] consists of 58 images of size

---

[1] Three-dimensional subbands can be considered as collections of so-called *subband images*.

[2] This volume is a courtesy of the Mallinckrodt Institute of Radiology

256 × 256. Since this dataset has also been used as a test set by other authors [9, 10, 11] it is a reference for the evaluation of the compression performances. We assume that the object of interest is the brain, and consider the rest as background. The datasets are presented in fig 3. Particularly, fig. 3(a) and (c) show an image of the MRI and MIR-MRI sets, respectively, while fig. 3(b) and (d) show the corresponding masks selecting the ROI.

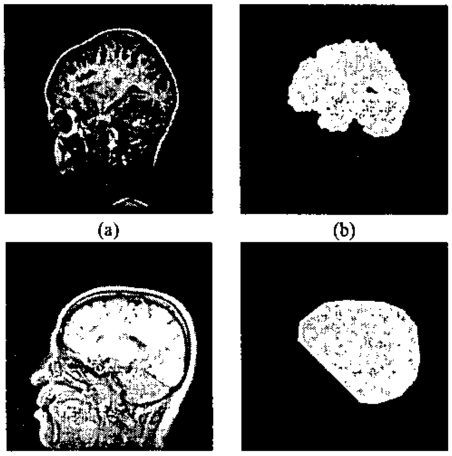

Fig. 3. MR scans of a human head. (a) MRI; (b) MIR-MRI (after histogram equalization); (c) mask for (a); (d) mask for (c).

The performances of the ROI-based mode of JPEG2000 are not completely satisfying for this type of data. In JPEG2000, ROI-based coding is performed by the MAXSHIFT method [12]. The subband coefficients within the ROI mask are shifted up (or, equivalently, those outside the ROI are shifted down) so that the minimum value in the ROI is greater than the maximum value in the background. This splits the bitplanes respectively used for the ROI and the background in two disjoint sets. The rate allocation procedure assigns to each layer of each codeblock (in the different subbands) a coding priority which depends on both the semantics (through the MAXSHIFT method) and the gain in terms of rate/distortion ratio. This establishes the relative order of encoding of the ROI subband coefficients with respect to the background. When applied to the head MRI dataset and with the implementation described in [13], such a procedure tends to move the encoder focus out of the ROI. The ROI and background codeblocks are "mixed up", compromising ROI-based functionalities. What happens is that the background is decoded before the lossless representation of the ROI is achieved. This can be verified by progressively decoding the bitstream up to the rate indicated by the encoder as representative of the ROI. The image in fig. 3(a) was compressed with the mask in 3(b). As the decoding rate increases, the quality of the background increases despite the lossless representation of the ROI is not yet reached. Fig. 4 shows the images for the decoding rates 0.1, 0.5, 1.0 and 1.6 bit/pixel. The last rate is the lossless rate of the ROI as provided by the encoder.

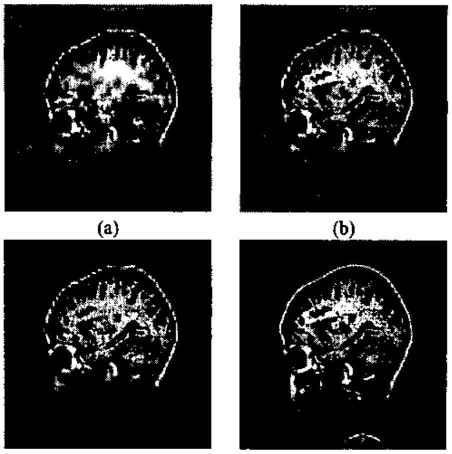

Fig. 4. ROI-based JPEG2000. Images decoded at different rates. (a) 0.1 bpp; (b) 0.5 bpp; (c) 1.0 bpp; (d) Reference rate for lossless ROI.

In order to make the comparison with our system, we independently compressed the ROI and the background with JPEG2000 and compared the respective bitrates to those provided by ROI-based JPEG2000. Results are given in fig. 5. The global lossless rate in the different conditions are shown as a function of the image index. The dash-dot line represents ROI-based mode (JPEG2000 ROI) and the continuous line is for the independent object coding (JPEG2000 IO). The curve represents the sum of the lossless rates concerning the ROI and the background. The corresponding average lossless rates are shown in Table 1. In particular, the OBJ column gives the lossless rate for the object of interest, when it is encoded independently by the different algorithms, while the WHOLE and OBJ+BGND columns provide the lossless rates obtained for the entire volume and the independent coding of the object and the background, respectively. The differences between the compression ratios for the cases WHOLE and OBJ+BGND are

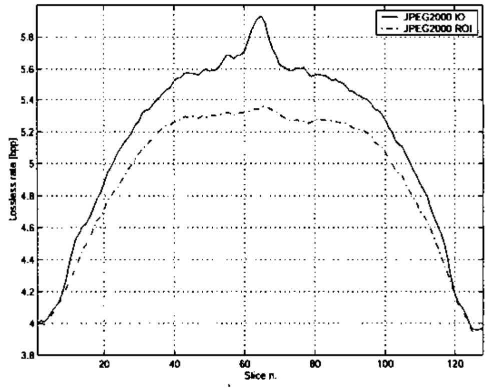

Fig. 5. Lossless performances of JPEG2000 standard on the MRI dataset. The baseline ROI mode (JPEG2000 ROI) is compared to the independent coding of the object of interest and the background (JPEG2000 IO).

(Washington University).

Volume	LR [bpp]	OBJ	WHOLE	OBJ+BGND
MRI	MLZC (030)	0.959	4.666	5.048
	JPEG2000 IO	1.064	4.651	5.130
	JPEG2000 ROI	1.259	4.651	4.910
MIR-MRI	MLZC (030)	0.788	2.310	2.933
	JPEG2000 OI	1.062	2.950	3.300
	JPEG2000 ROI	1.289	2.950	3.267

Table 1. Lossless rates (LR) for MRI and MIR-MRI. The filter is 5×3, $L = 4$. Pure spatial conditioning has been used for the MLZC LPL-PROG mode.

due to two causes. First, the entropy coder performs differently in the two cases because of the different sources. Second, the total number of coefficients to be encoded is larger for OBJ+BGND because of the extra coefficients to be encoded for both the object and the background. The minimum lossless rate for OBJ is obtained by MLZC for both the datasets. For MRI, JPEG2000 ROI outperforms the proposed system for OBJ+BGND reducing the lossless rate of 2.7%. However, in this mode the ROI based functionalities are not completely fulfilled. Conversely, MLZC OBJ+BGND outperforms JPEG2000 IO of about 1.6%, while preserving random access to every object. For the MIR-MRI dataset MLZC OBJ+BGND provides the best performances. The corresponding rate saving is about 11% and 12.5% over JPEG2000 ROI and IO, respectively. The prioritization of the information inherent to separate object processing leads to a significant improvement in coding efficiency when it is possible to relax the lossless constraint in the background region. In this case, the background can be encoded/decoded at a lower resolution.

The performances in lossy and pseudo-lossless regime are still under investigation. However, it is reasonable to expect a trade off between the improvement in coding efficiency provided by the 3D DWT and the need to recover a *set* of subband images for reconstructing a given image of interest.

The major bottleneck of 3D coding systems is computational efficiency. In this work, we did not address the problem of complexity and no optimization has been performed. We leave this subject for future investigation.

4. CONCLUSIONS

We presented a wavelet-based coding system featuring object-based 3D encoding with 2D decoding capabilities. In this way, the improvement in coding efficiency provided by 3D algorithms can be obtained at a lower computational cost. Each object is encoded independently to generate a self-contained segment of the bitstream. The implementation of the DWT via the lifting steps scheme in the non-linear integer version and the embedding of the encoded information allow to reconstruct each object of each image at a progressive up-to-lossless quality. Border artifacts are avoided by encoding some extra-coefficients (for each object). There is clearly a trade off between the improvement in coding efficiency of 3D over 2D systems and the increase in complexity. The complexity issue has not yet been addressed and no optimization has been performed. However, the potential of 3D encoding/2D decoding is very high, especially in the framework of the emerging model-based approach to coding.

5. REFERENCES

[1] ISO/IEC JTC 1/SC 29/WG1, *Information technology – JPEG2000 image coding system, ISO/IEC International Standard 15444-1, ITU Recommendation T.800*, 2000.

[2] G. Menegaz, L. Grewe, and J-Ph. Thiran, "Multirate coding of 3d medical data," in *Proc. of the International Conference on Image Processing (ICIP)*, Vancouver, BC, Canada, Sept 10 - 13 2000, vol. 3, pp. 656 –659.

[3] I. Daubechies and W. Sweldens, "Factoring wavelet transform into lifting steps," *J. Fourier Anal. Appl.*, vol. 41, no. 3, pp. 247–269, 1998.

[4] A.R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," *Appl. Comput. Harmon. Anal.*, vol. 5, no. 3, pp. 332–369, 1998.

[5] J. Reichel, G. Menegaz, M. Nadenau, and M. Kunt, "Integer wavelet transform for embedded lossy to lossless image compression," *IEEE Trans. on Image Processing*, vol. 10, no. 3, pp. 383–392, March 2001.

[6] G. Menegaz and J.-Ph. Thiran, "3d encoding–2d decoding of medical data," *submitted to IEEE Trans. on Medical Imaging*, 2001.

[7] G. Menegaz and J.-Ph. Thiran, "Lossy to lossless object-based coding of 3d mri data," *submitted to IEEE Trans. on Image Processing*, 2001.

[8] G. Menegaz, *Model-based coding of multi-dimensional data with applications to medical imaging*, Ph.D. thesis, Signal Processing Laboratory (LTS), Swiss Federal Institute of Technology (EPFL), May 2000.

[9] A. Bilgin, G. Zweig, and M.V. Marcellin, "Three-dimensional image compression with integer wavelet transform," *Applied Optics*, vol. 39, no. 11, pp. 1799–1814, 2000.

[10] Z. Xiong, X. Wu, D.Y. Yun, and W.A. Pearlman, "Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transform," in *IEEE Second Workshop on Multimedia Signal Processing (Cat.No.98EX175)*, Piscataway, NJ, USA, 1998, pp. 553–558.

[11] Y. Kim and W.A. Pearlman, "Stripe-based sphit lossy compression of volumetric medical images for low-memory usage and uniform reconstruction quality," in *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2000, vol. 4, pp. 2031–2034.

[12] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The jpeg 2000 still image coding system: an overview," *IEEE Trans. on Consumer Electronics*, vol. 46, no. 4, pp. 1103–1127, Nov. 2000.

[13] D. Taubman, "High performance scalable image compression with ebcot," *IEEE Trans. on Image Processing*, vol. 9, no. 7, pp. 1158–1170, July 2000.