DIAGNOSTIC COMPRESSION OF BIOMEDICAL VOLUMES

Alberto Signoroni, Riccardo Leonardi
Signals and Communications Lab., Dept. of Electronic for Automation,
University of Brescia,
via Branze 38, I-25123 Brescia, ITALY
Tel: +39030 3715432; fax: +39030 380014
e-mail: signoron,leon@ing.unibs.it

ABSTRACT

In this work we deal with lossy compression of biomedical volumes. By force of circumstances, diagnostic compression is bound to a subjective judgment. However, with respect to the algorithms, there is a need to shape the coding methodology so as to highlight beyond compression three important factors: the medical data, the specific usage and the particular end-user. Biomedical volumes may have very different characteristics which derive from imaging modality, resolution and voxel aspect ratio. Moreover, volumes are usually viewed slice by slice on a lightbox, according to different cutting direction (typically one of the three voxel axes). We will see why and how these aspects impact on the choice of the coding algorithm and on a possible extension of 2D well known algorithms to more efficient 3D versions. Crosscorrelation between reconstruction error and signal is a key aspect to keep into account; we suggest to apply a non uniform quantization to wavelet coefficients in order to reduce slice PSNR variation. Once a good "neutral" coding for a certain volume is obtained, non uniform quantization can also be made space variant in order to reach more objective quality on Volumes of Diagnostic Interest (VoDI), which in turns can determine the diagnostic quality of the entire data set.

1 INTRODUCTION

In the field of biomedical signal compression there is a strong requirement of errorless reconstruction. This derives from different and apparent obvious motivations, most importantly to overcome legal and diagnostic issues. However, even in its most effective implementations, lossless compression cannot overcome a data dependent [1] compression factor which remain too low for efficient storage and transmission in a variety of PACS [2] application contexts. By the way, diagnostic coding has been introduced in the case of image coding [3] and it is based on the *diagnostic quality* definition: "a lossy reconstructed biomedical volume can be accepted from a diagnostic quality point of view, if a physician with same qualification level can establish the same diagnosis on the reconstructed volume (with all coding information enclosed) with respect to the original one".

In this work, a wavelet based approach is considered in the case of volumetric compression, using MRI and CT anatomical data (Sec.2). Volumetric compressed material may exhibit peculiar artifacts depending on the visualization strategy mainly because of signal to error crosscorrelation. This must be taken into account to improve on the coding scheme. In Sec.3 an analysis of these sorts of artefacts is carried out. Let us stress that in the present JPEG2000 standardization process, even if biomedical imaging is considered, 3D coding is not directly addressed and the coding strategy does not take into account the implications of the above concern. Moreover, video coding implements 2D+t predictive motion oriented coding schemes with paradigms that we cannot consider adequate for static volumes, given the lack of moving objects.

2 3D WAVELET CODING

Zerotree-wavelet coding [4, 5], with all its variants, currently represents one of the the most efficient ways for progressive image compression, with the possibility to embed lossy and lossless coding in a unique multiresolution framework. Progressive coding is essential for communicating and archiving images and volumes. This is due to the possibility to refine, even in a spatially localized fashion, or to erode the compressed bit-stream without having to recode the entire data. CT and MRI volumetric scans produce a set of samples (voxels) on a 3D grid (usually regular but not necessarily isotropic) and quantize them on a gray level scale with fine granularity (usually 12bit/voxel stored using a 16 bit short integer) but with a relatively low SNR [1]. The potential of a 3D extension of zerotree schemes has already been presented by some authors[6, 7]. In order to have a perfect reconstruction, it is necessary to perform an integer-to-integer wavelet transform (WT), but in practice it is useless to spend bit to reconstruct the signal below its source noise variance. Integer WT may be used to lower the computational cost, but floating point representations are better suited for coding gain optimality as well as for coefficient manipulation before and

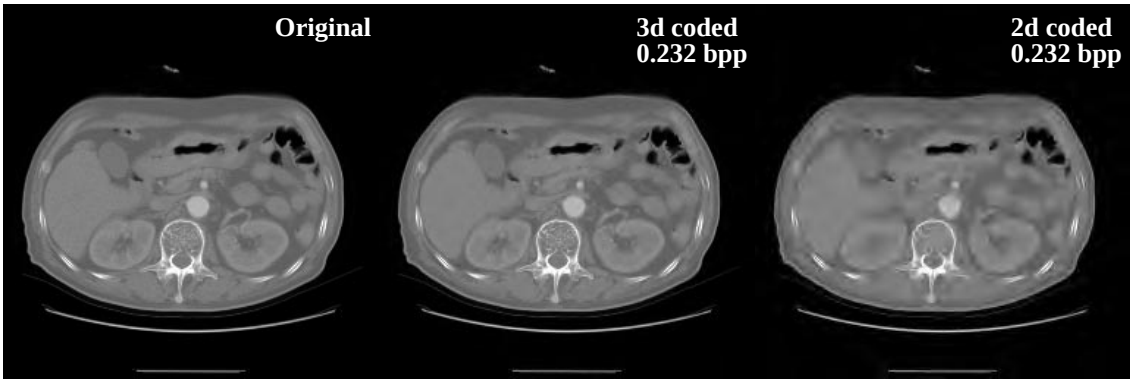

Figure 1: Different results for 3D (PSNR=37dB) and repeated 2D(z) coding (32.5dB) at the same CR=52.

after quantization. For example a product mask may be applied to the wavelet coefficients to highlight Regions or Volumes of Diagnostic Interest (RoDI [3], VoDI) or to perform a non-uniform quantization of subbands.

2.1 The Voxel anisotropy problem

For the greater part of imaging modalities, the spectral characteristics of each slice follow the typical $1/f$ model for natural images well, whereas a good match with the $1/f$ model along the third dimension strongly depends on the voxel resolution anisotropy with respect to the isotropic cube. In other words, some energy compaction deficiencies of the WT along the z dimension may occur due to the typical lower resolution in the slicing direction. This may compromise the zerotree algorithm performance due to a breakdown of the $1/f$ spectral characteristics. We will highlight, giving an example, the dependencies of 3D wavelet coding with respect to the image source and z-resolution. As a reference we compare 3D coding performance over that of a set of repeated 2D ones.

2.2 3D coding gain over repeated 2D

In Tab.1, a comparison of coding performance is shown between two different imaging data sets, i.e. a CT-abdomen and a MRI-brain (256x256)x128 slice sets, which has been subsampled to obtain different z-resolution datasets (128 and 64 slices). The initial data sets have been produced with anisotropic voxels of respectively 2x2x4mm and 0.78x0.78x1.17mm spatial resolution. As for the decomposition, we use the set of biorthogonal filters proposed in [4]. The 3D separable WT decomposition is performed with 5 levels for 128 slices dataset and with 4 levels for the 64 resliced ones. It is reasonable to consider only a separable WT implementation for computational efficiency, and different basis along each dimension, in order to better deal with anisotropy of the voxel characteristics. As coding algorithm we perform SPIHT [5] coding repeated on each slice and a 3D version of SPIHT. The total bitstream lengths generated at various target PSNR distortion are used to compare the resulting compression ratios (CR=12bpp/stream-bpp). Here the PSNR value of 45dB represents an absolute visually lossless quality, even if reconstructed images are magnified and compared at high zoom factors. As we can notice, the CR %gain of the 3D algorithm over the 2D(z) algorithm is more important for the CT data set. On the other hand, for the subsampled 64 slice MRI there is only a little gain in using the 3D coding approach. Usually, MRI data are more detailed and noisy with respect to CT ones, thus the WT does not allow a good decorrelation while high predictivity of coefficients (by zerotree) in the z direction is weakened. In Fig.1, we show an example of 2D(z) and 3D coding results for the slice 40 of CT data set. For low z resolution datasets it is sometimes better to use shorter filters in the z direction, such as the haar basis [6]. An alternative is to change the wavelet representation in order to design a predictive coding schemes along z. This is an open research area, because it is important, given the imaging modality and the voxel aspect ratio and resolution, to find an optimal wavelet representation in terms of coding efficiency. Moreover, there are other issues that must be taken into account in order to envisage a "diagnostic compression".

3 VOLUMETRIC CODING ERRORS

It is well known that the reconstruction error autocorrelation and the error to signal cross-correlation depends on the decomposition structure, the quantization model and the reconstruction filter shape. A quantitative analysis of the phenomena is out of the scope of this short contribution, but we want to summarize some derivations of this type of analysis in a qualitative manner.

A 2D(z) coding strategy has the advantage of allowing a precise PSNR control over each single slice, as depicted in the above experiment and shown in Fig.3(a); in the 3D case this is not guaranteed due to the quantization process used in the zerotree coding, and to the spreading of quantization error in a localized fashion (ringing) by the inverse WT. Nevertheless, this advantage is not real, as the radiologist does not necessarily

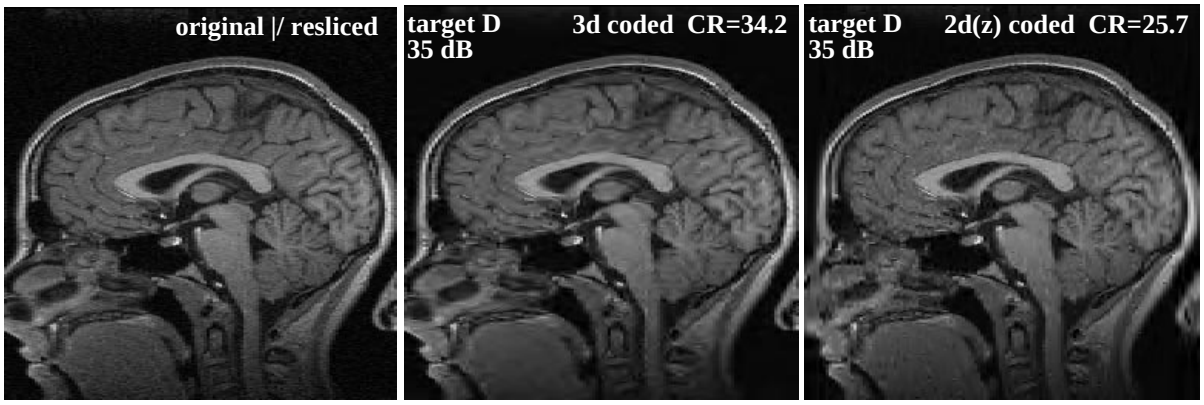

Figure 2: MRI slices originated by a volume reslicing on y. For both the coding modalities a target distortion PSNR of 35 dB was used, but some streaking artifacts, due to error z-uncorrelation, are evident on the right most image.

Data	**CT**		**CT**		**MRI**		**MRI**	
	128		**64**		**128**		**64**	
	2D	*3D/g*	*2D*	*3D/g*	*2D*	*3D/g*	*2D*	*3D/g*
D(dB)	**CR**	–>						
45	7.8	14.3	7.5	9.3	5.5	5.7	5.5	5.5
		83%		24%		4%		0%
40	14.1	30.7	13.7	18.4	9.3	9.6	9.2	9.1
		118%		34%		3%		-1%
37	20.8	53.1	19.9	29.3	16.2	18.2	16.0	16.1
		155%		47%		12%		0.6%
35	28.5	74.4	26.9	43.9	25.8	31.9	25.2	26.0
		161%		63%		24%		3.2%

Table 1: Comparison between 2D(z) and 3D coding on different data sets. For each dataset and for each target PSNR distortion D (dB), repeated 2D and 3D coding are performed, and the compression ratio (CR) and CR % gain (g) of 3D over 2D(z) is quoted.

consider only the z slicing direction for diagnostic purposes. He/she can use any cutting plane, for example following the x or y axes. Fig.3(b) shows the 2D PSNR vs y-slice number in an interval taken from the same dataset of Fig.3(a). From a perceptual point of view, the artifacts caused by 3D coding are always defocusing and eventually ringing (depending on the CR, the cut direction vs voxel anisotropy, the zooming factor...); in other words, these artefacts are correlated with the signal characteristics. In the 2D case, instead, the quantization may introduce objectionable errors along the z dimension, because the reconstructed signal lacks correlation in such direction. This determines uncontrollable mismatch which, at medium or low bit rates, may alter the anatomical structures in the data set, as can be seen in Fig.2.

3.1 PSNR constancy

Fig.3(a,b) show a regular oscillation of the per-slice PSNR curve. The amplitude and frequency of the oscillation is shaped by the quantization inverse WT error of the first level wavelet coefficients (higher freq.). Despite the localized PSNR control on z slice for 2D(z) coding, the PSNR oscillation with respect to other slicing axes is nearly the same for both coders(Fig.3(b)). This superposition is due to the intrinsic separable nature of embedded wavelet coding which is robust with respect to PSNR control in one direction. Being the slices at adjacent position quite similar one to the other, the PSNR variation may be perceived on the screen lightbox even at medium-low rates, causing objectionable artifacts and lowering the reliability of the coding. It is important to find some methodology to lower the amplitude of the oscillation. We propose to operate a non uniform quantization on the first stages of wavelet coefficients.

3.2 Non-uniform quantization

In Fig.3(c) three curves are shown representing the effect of two WT coefficient level weighting. Multiplicative weight are applied prior to SPIHT quantization and taken off prior to inverse WT. The measurements are performed at the same CR. It is easy to see the immediate benefit in terms of PSNR oscillations, while the global PSNR performance get slightly worse, as it can be expected. Results on z slicing directions reflects what happens in the other direction as well. The oscillation standard deviation in the original case is $\pm$0.4dB, while weighting the first WT level by a factor of 2, the σ drop to $\pm$0.2dB, and with 1st WT level weight, WTlw=4 and 2nd WTlw=2, $\sigma = \pm$0.13dB. This kind of *masking* is in accordance with the results of perceptual studies aiming to obtain a good match between subband weighting and HVS characteristics. A combined study should give optimal results. Thus, even if global PSNR worsens this does not imply a worse image quality. It is also interesting to note how the PSNR oscillation pattern changes if we reduce quantization errors for example in the first WT level. In Fig.3(d) we apply two masks: M1 consists of 1st WTlw=2, and M2 is 1st WTlw=4. Each coding is stopped when a 10 bit-plane quantization refine-

ment is performed on each WT coefficient. With M2 the 1st level coefficients are best represented (12 bit-planes WT precision) but the PSNR oscillation pattern is now dominated from higher level coefficients. It is interesting to see how with M1 we take advantage of a favorable intermediate situation between the unweighted coding pattern ($\sigma = \pm 0.30$dB) and what corresponds to M2 ($\sigma = \pm 0.50$dB). With M1 we obtain a PSNR $\sigma = \pm 0.17$dB.

4 CONCLUSION

In this presentation we have seen that 2D(z) coding is not well suited for compression of biomedical volumes because it is less efficient with respect to 3D coding and it allows a distortion control only along z while presenting PSNR inter-slice oscillations as well as potential objectionable artifact considering other slicing directions. We have proposed 3D SPIHT coding with non uniform quantization of subbands in order to reduce PSNR oscillations for every slicing directions. Good results have been obtained without impairing subjective quality (which can become even better from a perceptual point of view). All this, with a good wavelet decomposition, must be done on a given image modality, voxel aspect ratio and resolution, to guarantee an efficient and reliable compression of biomedical volumes that can be the basis for diagnostic coding.

References

[1] B. Aiazzi, L. Alparone, S. Baronti, "Virtually-lossless compression of medical images through classified prediction and context-based arithmetic coding" VCIP'99, SPIE vol.3653, 1033-1040, San José, California, Jan.1999.

[2] H.K. Huang, *PACS Picture Archiving and Communication Systems in Biomedical Imaging*, VCH Publishers, 1996.

[3] A. Signoroni, R. Leonardi, "Progressive medical image compression using a diagnostic quality measure on regions-of-interest" EUSIPCO'98, 2325-2328, Rodi (GR), Sep.1998.

[4] J.M. Shapiro, "Embedded image coding using zerotrees of wavelets coefficients", IEEE Trans. on Signal Processing, 41(12), 3445-3462, Dec. 1993.

[5] A. Said, W. Pearlman, "A new, fast and efficient image codec based on set partitionning in hierarchical trees", IEEE Trans. Circuit Syst. Video Technol., vol.6, 243-250, June 1996.

[6] J. Wang, H.K. Huang, "Medical Image Compression by using Three-Dimensional Wavelet Transformation", IEEE Trans. on Medical Imaging, 15(4), 547-554, Aug.1996.

[7] A. Bilgin, G. Zweig, M.W. Marcellin, "Three-Dimensional Image Compression using Integer Wavelet Transforms" , Applied Optics: Information Processing, 39(11), 1799/1814, Apr. 2000

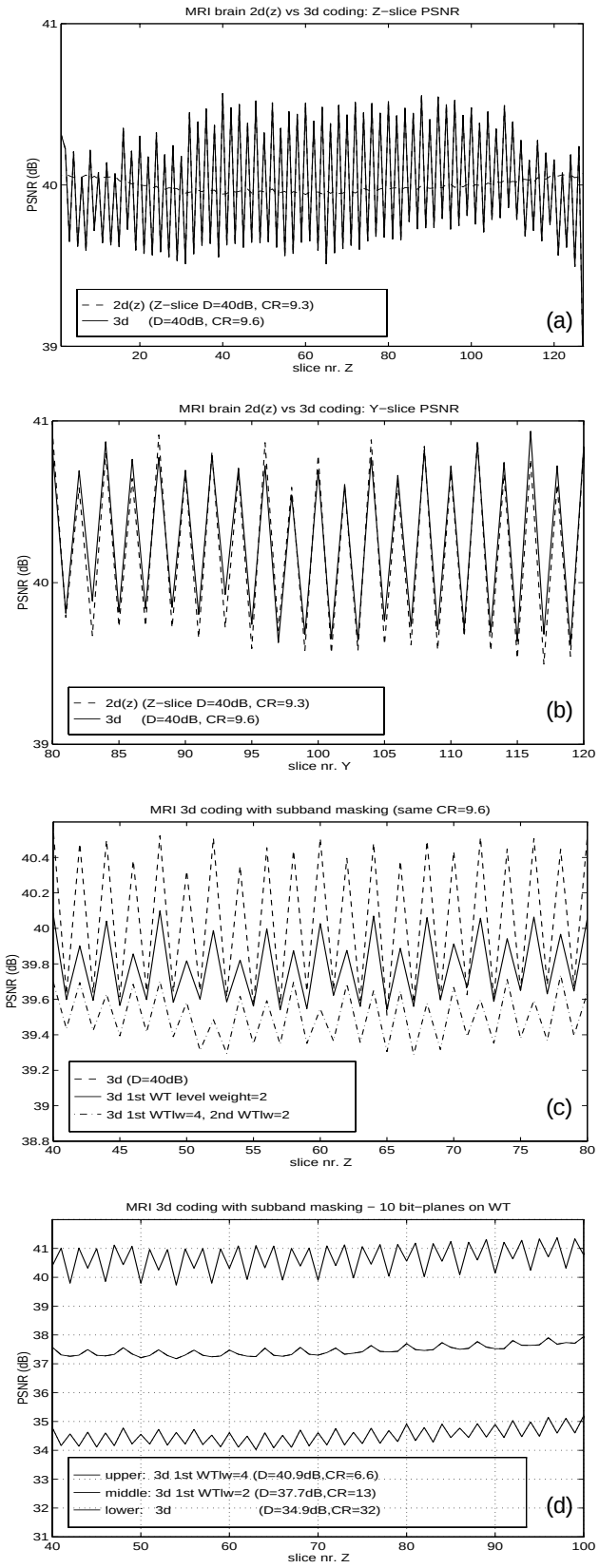

Figure 3: PSNR performance comparison between 2D(z) and 3D coding (a), (b) and different weighting of wavelet coefficients (c), (d) in order to obtain lower PSNR oscillations.