Pre-print of: "Correlation modeling for compression of computer tomography images" / J. Muñoz-Gómez, et al. in EEE Journal of Biomedical and Health Informatics (IEEE), Vol. 17, Issue 5 (Set. 2013) , p. 928-935. DOI 10.1109/JBHI.2013.2264595

Correlation Modeling for Compression of Computed Tomography Images

Juan Muñoz-Gómez, Joan Bartrina-Rapesta, Michael W. Marcellin, *Fellow, IEEE*, and Joan Serra-Sagristà, *Senior Member, IEEE*

***Abstract*—Computed Tomography (CT) is a noninvasive medical test obtained via a series of X-ray exposures resulting in 3D images that aid medical diagnosis. Previous approaches for coding such 3D images propose to employ multi-component transforms to exploit correlation among CT slices, but these approaches do not always improve coding performance with respect to a simpler slice-by-slice coding approach. In this work, we propose a novel analysis which accurately predicts when the use of a multi-component transform is profitable. This analysis models the correlation coefficient r based on image acquisition parameters readily available at acquisition time. Extensive experimental results from multiple image sensors suggest that multi-component transforms are appropriate for images with correlation coefficient r in excess of 0.87.**

***Index Terms*—Computed Tomography Image Compression, Correlation modeling, Multi-component transforms, JPEG2000 coding standard, DICOM protocol**

I. Introduction

Human body medical imaging is often used for clinical diagnosis. One of the medical imaging modalities that is more commonly used is Computed Tomography (CT), which combines special X-ray equipment with sophisticated software to produce three-dimensional (3D) images, each consisting of a set of image slices. These images of the inside of the human body show organs, bones, soft tissue and blood vessels with greater clarity than standard X-rays, allowing radiologists to more easily diagnose problems such as cancer, infectious diseases, appendicitis, cardiovascular diseases, trauma and musculoskeletal disorders [1].

The use of CT imagery has increased rapidly. In 2007, it was estimated that more than 62 million CT scans were obtained per year in the United States [2]. To manage these data, medical centers use Picture Archiving and Communications Systems (PACS) [3] to store, retrieve, distribute, and display medical images. PACS are commonly constituted of large computer networks, servers, and workstations [4], [5]. The Digital Imaging and Communications in Medicine (DICOM) standard [6] specifies the format used to store and distribute images in PACS. Due to the number of images managed, data compression plays a key role in DICOM.

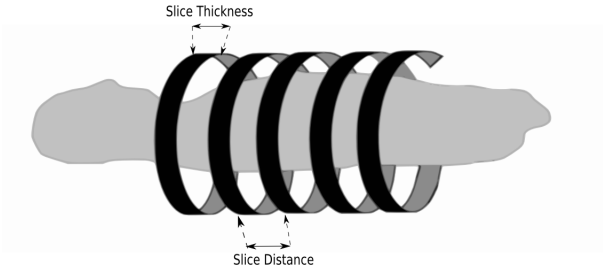


Figure 1: Graphical representation of slice thickness and slice distance during a CT scan.

During the CT scanning process two main parameters can be manipulated by the radiologist to capture the desired information, *slice thickness* and *slice distance*. Slice thickness is defined as the width (in mm) of the region in the human body represented by each slice. Its value can be selected according to clinical requirements and commonly lies between 1 mm and 10 mm. In general, a larger slice thickness results in poorer contrast resolution in the image. On the other hand if the slice thickness is small (e.g., 0.75-2 mm), higher radiation doses are required to achieve a high quality image [7]. Slice distance is defined as the distance (in mm) between two adjacent slices. Similar to slice thickness, common slice distances lie between 0 mm and 10 mm. It is possible to choose the slice distance to be less than the slice thickness. Figure 1 provides a graphical representation of these two concepts.

Regarding CT image acquisition parameters, Siegel et al. [8] presented an empirical study of the effects of slice thickness in CT coding, concluding that thinner CT slices are less compressible than thicker slices when 2D coding is employed, and recommended the use of a 3D coder to obtain higher compression ratios. Such 3D coding exploits the fact that CT images can have a significant amount of redundancy among slices, which can be exploited through multi-component transforms to improve coding performance. Under certain assumptions, the potential for such improvement can be characterized via the correlation coefficient r [9].

Compression of medical imagery is an active topic of research [10]. Recently published work on this topic includes Schelkens et al. [11], which presented an extensive review of 3D wavelet coders, and proposed three different coding methods which attempted to exploit correlation among components. Xiong et al. [12] proposed a 3D modification of set

This work has been partially supported by European Union, Spanish Government (MINECO), FEDER, Catalan Government, and Universitat Autònoma de Barcelona, under Grants FP7-PEOPLE-2009-IIF, FP7-250420, TIN2009-14426-C02-01, TIN2012-38102-C03-03, 2009-SGR-1224 and UAB-BI3INT2006-08.

J.Muñoz-Gómez, J.Bartrina-Rapesta, and J.Serra-Sagristà are with the Department of Information and Communications Engineering, Universitat Autònoma de Barcelona, E-08193 Cerdanyola del Vallès, Barcelona, Spain (e-mail: jmunoz@deic.uab.cat, joan.bartrina@uab.cat, joan.serra@uab.cat).

M.W. Marcellin is with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721, USA (e-mail: marcellin@ece.arizona.edu).

partitioning in hierarchical trees (SPIHT) [13] and of embedded subband coding with optimal truncation [14]. Agarwal et al. [15] presented a fast JPEG2000 decoder and discussed its usefulness in medical image coding. In 2009, Miaou et al. [16] developed a lossless coding scheme, using JPEG-LS and an interframe coding stage, which outperforms JPEG2000 and JPEG-LS for lossless coding. Sanchez et al. [17] have exploited image symmetries to predict the value of wavelet coefficients on a block-by-block basis. On the other hand, methods based on Region Of Interest (ROI) techniques, aimed to encode only the biological area of the image –or the relevant area detected by computer-aided diagnosis procedures– have been proposed. Penedo et al. [18] presented object-based extensions for the set partitioning in hierarchical trees and the set partitioning embedded block coder algorithms for digital mammography, and explored the effects of lossy compression for detecting microcalcifications in digital mammography [19]. Sanchez et al. [20] proposed a 3D scalable compression method for medical images with optimized volume of interest coding. More recently, Bartrina et al. [21] introduced an ROI coding method for digital mammography based on component priority. Kim et al. [22] presented a preprocessing method for CT images that replaces the pixels of the non-body region by a constant value, maximizing the data redundancy. A similar approach had previously been employed in the framework of remote sensing scenarios [23]. Kassim et al. [24] proposed a 4D image coding scheme combining a 3D wavelet transform, 3D motion compensation and a 3D extension of SPIHT. Sanchez et al. [25] presented a 4D image coding method based on H.264/AVC and a modification of context-adaptive binary arithmetic coding that takes into account the probability distribution of the residual and motion vector data.

In the next sections we show that multi-component transforms improve compression performance significantly when the correlation among slices is sufficiently high. However, precomputing image correlation is a computationally demanding task. In this paper we propose a new correlation model specifically designed for CT images. Our novel contribution employs CT image acquisition parameters to model the correlation among slices. Results indicate that the proposed method accurately models the correlation among slices.

The manuscript is structured as follows: Section II introduces the employed image corpus, the metrics used to evaluate our proposal, and a short review of the JPEG2000 standard. Section III describes our correlation model. Section IV provides experimental results. Section V closes the paper with discussion and conclusions.

II. Materials and Methods

A. Image Corpus

The images employed in this work were acquired with four different CT scanners: Siemens Sensation 16, Siemens Somatom Plus 4, General Electric LightSpeed 16, and Philips Brilliance 40. Images from the first sensor were provided by Parc Taulí Health Corporation [26], while images from the second, third and fourth sensors were obtained from The Cancer Imaging Archive [27]. All images have a bit-depth of 12 bits per pixel per slice (bppps) with sign, but are stored using 16 bppps. The corpus contains 100 3D images. Different acquisition parameters –selected by the radiologist for the purpose of specific examinations– are considered.

Table I summarizes the corpus characteristics. The first column indicates the sensor used to acquire the imagery. The second column provides image names, which end with an integer suffix to differentiate between multiple 3D images having the same acquisition characteristics. The third column gives the number of 3D images with the same image characteristics. The fourth and fifth columns give, respectively, slice thickness and slice distance, while pixel spacing within a slice is provided in column six. The last column reports the number of slices, N_z, in each 3D image, which is given as a range, since images with the same characteristics may have a different numbers of slices. In every case, the slice size is 512 by 512 pixels.

B. Coding Performance Metrics

The performance of a coding system is established as a trade-off between the rate achieved by the coding process, and the quality of the recovered image after decoding. To evaluate performance, two quality metrics are used in this work: Signal-to-Noise Ratio (SNR) and High Dynamic Range - Visual Difference Predictor (HDR-VDP):

- One of the most common metrics to evaluate reconstruction quality is *Signal-to-Noise Ratio (SNR)*, defined as

$$\text{SNR} = 10\log_{10}\frac{\sigma^2}{\text{MSE}}\ \text{(dB)},$$

where

$$\text{MSE} = \frac{1}{N_z}\frac{1}{N_x}\frac{1}{N_y}\sum_{k=1}^{N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}(I_{ijk}-\hat{I}_{ijk})^2.$$

I_{ijk} and $\hat{I}_{ijk}$ denote, respectively, the values of the original samples and decompressed samples at position ijk corresponding to the horizontal, vertical and slice axes. σ^2 denotes the variance of the original image. N_x and N_y are the number of pixels in a row and in a column, respectively. Higher SNR represents better quality of the decompressed image.

- HDR-VDP is a perceptual metric, suitable for medical applications [28]. This metric returns a probability-of-detection map where each value in the map indicates the probability that a human observer would detect a difference in the pixel at the corresponding location in the compressed image. This probability-of-detection map is summarized by a single value as

$$\text{HDR-VDP} = \Big(\sum_i\sum_j p(i,j)^\beta\Big)^{\frac{1}{\beta}},$$

where $p(i,j)$ is the probability of detection for pixel (i,j), and $\beta = 2.4$ [28]. On a dB scale,

$$\text{HDR-VDP} = 20\log_{10}\frac{\text{HDR-VDP}_{max}}{\text{HDR-VDP}}\ \text{(dB)},$$

Table I: Image Corpus Characteristics.

Sensor	Image Names	# Images	Slice Thickness (mm)	Slice Distance (mm)	Pixel Spacing (mm,mm)	# slices (N_z)
Sensation 16	SS16-T1-D075_{1..7}	7	1	0.75	0.78,0.78	[337,637]
	SS16-T2-D1_{1..5}	5	2	1	0.75,0.75	[399,935]
	SS16-T2-D2_{1..3}	3	2	2	0.75,0.75	[105,180]
	SS16-T7-D5_{1..7}	7	7	5	0.75,0.75	[77,85]
	SS16-T5-D5_{1..22}	22	5	5	0.66,0.66	[53,110]
	SS16-T1-D10_{1..6}	6	1	10	0.66,0.66	[28,32]
LightSpeed 16	LS16-T125-D125_{1..7}	7	1.25	1.25	0.78,0.78	[241,261]
	LS16-T25-D25_{1..16}	16	2.5	2.5	0.82,0.82	[116,209]
Somatom Plus 4	SP4-T5-D5_{1..10}	10	5	5	0.65,0.65	[48,71]
Brilliance 40	B40-T1-D08_{1..17}	17	1	0.8	0.83,0.83	[48,71]

where HDR-VDP$_{max}$ is the maximum value that HDR-VDP can have, which would occur if all values in the probability-of-detection map were 1. In their paper, the authors conclude that an image recovered at 25.8 dB or above is visually lossless.

C. JPEG2000

JPEG2000 is a powerful image compression standard [29] that provides advanced features for imaging applications and has been included in DICOM since November 2001 [30]. Based on a wavelet coding scheme, it is composed of a two-tiered coding system: tier-1 carries out bitplane-by-bitplane arithmetic entropy encoding while tier-2 organizes the code-stream. JPEG2000 achieves high compression ratios in lossy, lossless and progressive lossy-to-lossless regimes, supports more than 16-bits of signed or unsigned data, includes tools for interactive transmission [31], and provides some interesting capabilities for 3D image coding, such as support for multi-component transforms [32], aimed to exploit redundancy among image components, commonly increasing compression performance when applied. It is worth noting that we use the JPEG2000 standard language that refers to "components" and "multi-component transforms." In the context of CT imagery, these can be understood as "slices" and "multi-slice transforms." In particular, if z refers to the slice dimension, with x and y being the spatial dimensions within a slice, then a multi-component transform is applied in the z dimension.

An important feature provided by JPEG2000 is scalability in terms of spatial location, resolution, component, and quality. Spatial scalability provides access to different spatial regions of an image. Resolution scalability allows one to obtain images in different resolutions or sizes. Quality scalability permits access to image data corresponding to different compression ratios or bitrates. Finally, component scalability is the ability to retrieve a set of selected components (or slices) of the image. All JPEG2000 scalabilities can be exercised without needing to decode the full code-stream. Spatial location scalability, quality scalability and resolution scalability are not affected by multi-component transforms. However, component scalability is.

All experiments presented in this paper were performed with Kakadu v6.4.1 [33]. In the (x, y) dimensions the reversible 5/3 wavelet transform (RWT) was used with 5 decomposition levels. In the third (z) dimension, either the RWT or the reversible HAAR transform (RHAAR) was used with a number of transform levels that depends on the total number of slices of each image. Specifically, the number of wavelet transform levels chosen for the z dimension was $\min\{5, \log_2 N_z\}$. Five levels of wavelet transform is typical in the literature. More than five levels does not generally provide additional increases in compression performance. Additionally, each level of wavelet transform reduces the number of "low band slices" by a factor of 2. Thus, $\log_2 N_z$ is a practical upper bound on the number of wavelet transform levels in the z dimension. A code-block size of 64×64 was used throughout. With respect to rate allocation, multi-component post compression rate distortion optimization was employed because it yields the best progressive lossy-to-lossless coding performance [34].

III. Correlation Modelling for Multi-Component Transform Selection

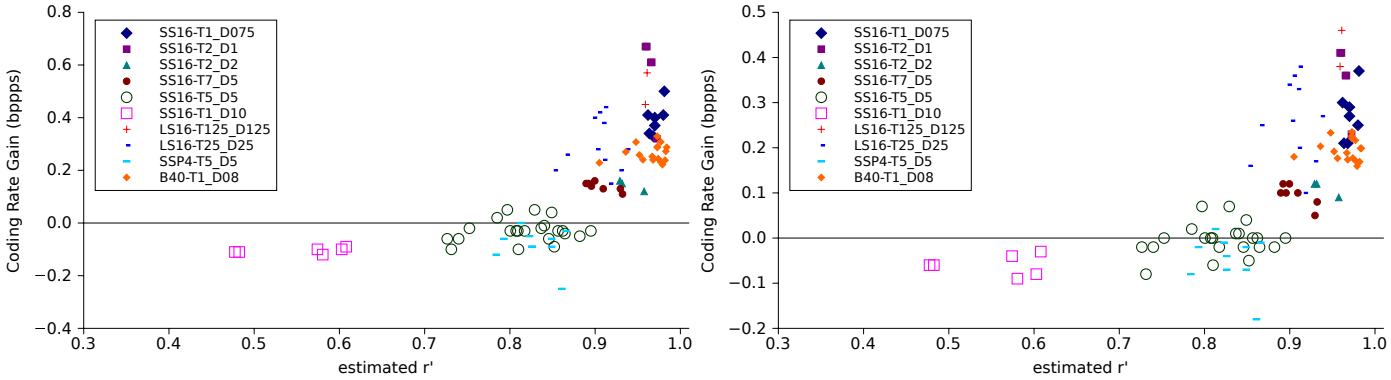
The work proposed in this section is based on the fact that CT images can have a significant amount of redundancy among slices, which may be exploited via multi-component transforms to improve coding performance. This performance improvement can be characterized by the correlation coefficient [9]. The correlation among slices varies significantly, depending on the two scanning parameters used to acquire an image: the slice thickness and slice distance.

Given two random variables A and B, their correlation coefficient is given by

$$r_{A,B} = \frac{E[(A - \bar{A})(B - \bar{B})]}{\sigma_A \sigma_B}, \tag{1}$$

where $E[\bullet]$ indicates the expectation or probabilistic average, $\bar{A} = E[A]$ is the mean of A, and $\sigma_A^2 = E[(A - \bar{A})^2]$ is the variance of A. Similarity $\bar{B}$ and σ_B^2 are the mean and variance of B. The correlation coefficient between two consecutive CT slices k and $k + 1$ can be estimated by

$$r'_{k,k+1} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{(x_{ijk} - \bar{x}_k)(x_{ijk+1} - \bar{x}_{k+1})}{\sigma_k \sigma_{k+1}}, \tag{2}$$



(a) RWT+JPEG2000 (b) RHAAR+JPEG2000

Figure 2: Coding rate gain vs estimated r'. (a) and (b) respectively depict the coding rate gain for RWT and RHAAR.

where x_{ijk} denotes the pixel at column i and row j of slice k, and $\bar{x}_k$ and σ_k respectively denote the sample pixel mean and standard deviation of slice k. The average correlation coefficient between consecutive slices of an image is estimated as

$$r' = \frac{1}{N_z - 1} \sum_{k=1}^{N_z - 1} r'_{k,k+1}. \tag{3}$$

Usually, "non-biological areas" in a 3D image do not change from slice to slice; however, these areas substantially influence the computation of r', bringing it artificially close to 1. To avoid this effect, r' is estimated using only a 170×170 pixel square window centered in the slices, which corresponds roughly to the biological area in the slices.

To evaluate the relationship between r' and the benefit of multi-component transforms in terms of lossless coding performance, Figure 2 depicts the difference in lossless coding rate between JPEG2000 with and without a multi-component transform as a function of r'. Two multi-component transforms are explored: the 5/3 RWT and the RHAAR transform. In particular, the figure depicts the bit-rate obtained by JPEG2000 (without multi-component transform) minus the bit-rate obtained by JPEG2000 with a multi-component transform (RWT+JPEG2000 or RHAAR+JPEG2000). We refer to this quantity as the coding rate gain, where positive values indicate improvement for multi-component transforms. For the images used in this manuscript, the results of Figure 2 suggest that r' is a good indicator of when a multi-component transform can improve coding performance. Roughly, a multi-component transform should be applied among slices when r' is greater than or equal to 0.87. Unfortunately, the computation of r' is quite computationally and memory intensive. This issue is addressed below.

A. Correlation Modeling Based on CT Image Acquisition Parameters

As mentioned in the introduction, Siegel et al. investigated the performance of JPEG2000 as a function of slice thickness. That work was empirical in nature and did not explore the role of slice distance. In this section, we propose a theoretical model for the correlation among components, denoted by r, as a function of slice thickness T and slice distance D. This model provides a basis for explaining compression performance in terms of these two parameters and is used to determine when a multi-component transform will be profitable.

In the previous section, the pixel at spatial location ij of slice k was denoted by x_{ijk}. In what follows, we consider a sequence of pixels, indexed by k, obtained by fixing a spatial location ij. To reduce notational clutter, we drop the explicit dependence on ij and write $x(k)$. We assume that the pixel $x(k)$ can be modeled as arising from the integration of some underlying continuous signal $y(z)$ over the extent corresponding to a slice thickness. For computational purposes, we discretize y with a sample distance significantly smaller than both T and D, and replace the integration of y by a sum. Hereafter, this sample distance is fixed at 0.0625 mm. The number of samples of y that correspond to one slice thickness is then $L = T/0.0625$. Similarly, the number of samples corresponding to the slice distance is $M = D/0.0625$. For example, when $T = 1$ mm, each $x(k)$ is modeled as a sum of $L = 16$ consecutive samples of y. Figure 3 depicts this example for $D = 0.75$ mm, 1 mm and 1.5 mm, resulting in $M = 12$, 16 and 24, respectively.

The k^{th} pixel value $x(k)$ can then be written as

$$x(k) = \sum_{l=1}^{L} a(l) y(kM - l). \tag{4}$$

The constants $a(l)$ are included in (4) for two reasons. First, they allow for the possibility of generalizing the expression to a weighted sum. Second, they facilitate the observation that (4) corresponds to a filtering (or convolution) operation. Specifically, $x(k)$ is a subsampled version of

$$w(n) = a(n) * y(n). \tag{5}$$

That is,

$$x(k) = w(kM), \tag{6}$$

where

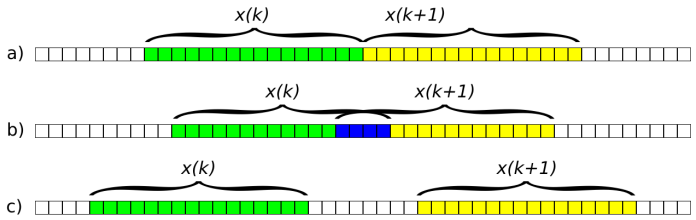


Figure 3: Samples of y used to compute $x(k)$ for three choices of D. In each case, $T = 1$ mm, a) $D = 1$ mm, b) $D = 0.75$ mm, and c) $D = 1.5$ mm.

$$w(n) = \sum_{l=1}^{L} a(l)y(n-l). \tag{7}$$

We now assume that the samples $y(n)$ arise from a simple auto-regressive random process

$$Y(n) = bY(n-1) + \Theta(n), \tag{8}$$

where $b \in$ (0,1) is a constant and $\Theta(n)$ is a stationary white Gaussian random process. The corresponding random processes for $w(n)$ and $x(k)$ are denoted by $W(n)$ and $X(k)$. The autocovariance function of $Y(n)$ is

$$C_Y(j) = E[(Y(n) - \bar{Y})(Y(n+j) - \bar{Y})], \tag{9}$$

where $\bar{Y} = E[Y(n)] = E[Y(n+j)]$, regardless of n. It is then easily shown that for the specific choice of (8),

$$C_Y(j) = \sigma_Y^2 b^{|j|}. \tag{10}$$

From (5), it follows that the autocovariance function of $W(n)$ is

$$C_W(j) = C_Y(j) * a(j) * a(-j). \tag{11}$$

From (6), we then have

$$C_X(k) = C_W(kM). \tag{12}$$

Thus, for given values of T, D, b, and $a(j)$, $j = 1, 2, 3, ..., L$, it is straightforward to compute $C_X(k)$ via (10), (11), and (12). Consistent with our simple integration model, $a(j) = 1$ in all discussions that follow, but other choices pose no complications. Finally, it follows that, given values for T, D and b, the correlation coefficient between two pixels $X(k)$ and $X(k+1)$ at the same location (i, j) in two consecutive slices is modeled by

$$r = \frac{E[(X(k) - \bar{X})(X(k+1) - \bar{X})]}{\sigma_X^2} = \frac{C_X(1)}{\sigma_X^2}. \tag{13}$$

We have used 24 images from the corpus of Table I to find a suitable value of b by minimizing the least squared error between r as computed by (13) and r' as computed via (3). Images with a variety of values of D and T were used in this process to obtain a *single* value of $b = 0.9962$. The results of Table II are provided to assess the performance of our model. In particular, each row of Table II corresponds to data from a collection of images having the same acquisition parameters T and D. For each row, one fixed value of r is reported. This value is computed via (13) using $b = 0.9962$ together with the values of T and D indicated by the image name. Additionally, a separate value of r' is computed for each image via (3). The average of these values is reported as $\bar{r}'$ in Table II. Finally, the mean error and variance of the error between r and r' is reported for each image set. We note that the 76 images used to obtain Table II are from the corpus of Table I but are different from the 24 images used to calculate b. As can be seen in Table II, the modeled values for r agree closely with the estimated values r'. Figure 4 depicts the modeled value of r as a function of T and D, where the color scale represents the different correlation values, as indicated on the right side of the figure. We can see that the correlation decreases when the

Table II: Modeled r and estimated $\bar{r}'$, together with mean error and standard deviation of the difference between r and r' for images with the same acquisition parameters T and D.

Images	r	$\bar{r}'$	Mean Error	Std. Deviation
SS16-T1-D075	0.979	0.971	0.0083	0.006
SS16-T2-D1	0.978	0.964	0.0133	0.0052
SS16-T2-D2	0.926	0.946	0.0132	0.0157
SS16-T7-D5	0.871	0.907	0.0359	0.0176
SS16-T5-D5	0.828	0.818	0.0103	0.0295
SS16-T1-D10	0.554	0.554	0.0001	0.0232
LS16-T125-D125	0.954	0.955	0.0007	0.014
LS16-T25-D25	0.908	0.906	0.0016	0.0153
SP4-T5-D5	0.828	0.827	0.0011	0.0152
B40-T1-D08	0.978	0.965	0.0129	0.0191

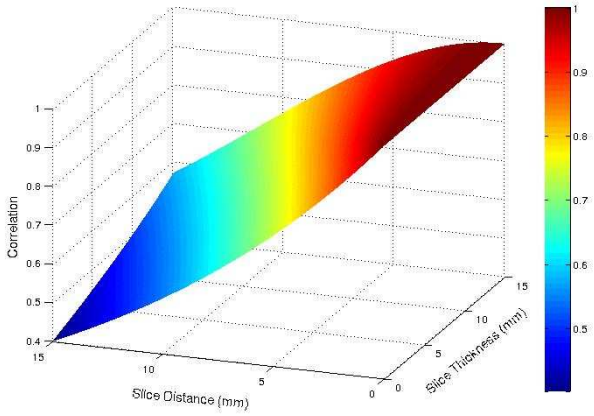


Figure 4: Modeled r.

slice distance D is increased. On the other hand, correlation increases as a function of slice thickness T.

It is worth noting the significant difference in complexity between estimating the correlation coefficient directly as r' vs. computing the modeled value r. It is evident that the computation of r' via (3) requires several calculations per pixel multiplied by $N_x \times N_y \times N_z$ pixels per 3D image. On the other hand, the complexity of the proposed method is constant, independent of the dimensions of the image.

IV. Experimental Results

Extensive experiments have been carried out to evaluate our correlation model. In particular, we have carried out: A) a lossless coding performance evaluation, B) a rate-distortion evaluation, and C) a component scalability evaluation. A) and B) aim to analyze the compression performance of multi-component transforms (RWT and RHAAR) on images with different acquisition parameters, and C) assesses the rate-distortion performance when a subset of components are decoded from a code-stream.

A. Lossless Coding Performance

In these experiments, we compare the lossless coding performance of JPEG2000 with two different multi-component transforms (RWT+JPEG2000 and RHAAR+JPEG2000) with that of JPEG2000 (without any multi-component transform) as a function of the modeled correlation coefficient r. Figure 5 is equivalent to Figure 2 but depicts r rather than r'. The same conclusion is apparent: performing a multi-component transform is profitable when r exceeds 0.87.

B. Rate-Distortion Evaluation

In this section, the rate-distortion performance of JPEG2000 with and without the two multi-component transforms is evaluated in terms of SNR and HDR-VDP. Figure 6 shows the rate-distortion performance in terms of SNR for three images from two different sensors with various acquisition parameters. As was the case for lossless compression, results suggest that for images with $r > 0.87$ the multi-component transforms improve the rate-distortion coding performance, while for images with low r, the rate-distortion performance of JPEG2000 without a multi-component transform is superior.

Figure 7 depicts the rate-distortion performance in terms of HDR-VDP. The horizontal black-dashed line identifies the visually lossless threshold determined by the authors of HDR-VDP. Results indicate that for images with $r > 0.87$, RWT+JPEG2000 reaches visually lossless performance at a lower rate than JPEG2000. However, for images with low r, JPEG2000 (without multi-component transform) provides visually lossless performance at a lower rate. Results are similar for other images in the corpus.

C. Component Scalability

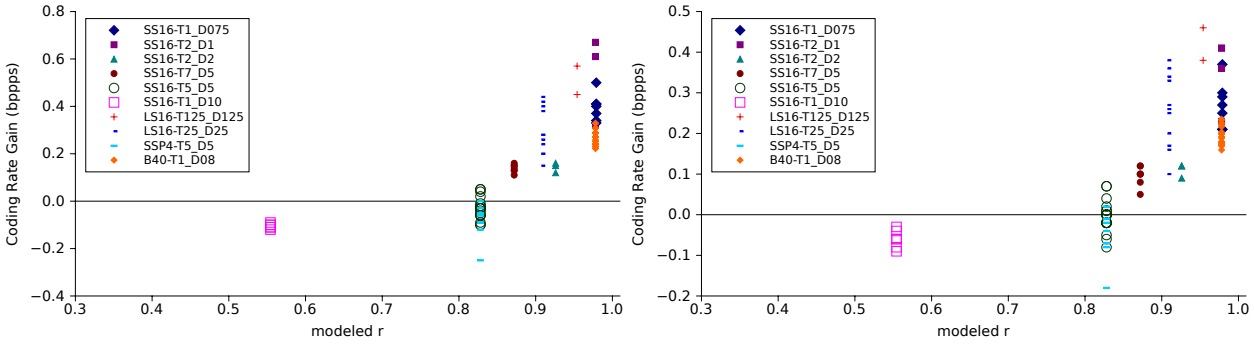
As mentioned in the introduction, component scalability is negatively impacted when multi-component transforms are employed. To explore this effect, we consider decoding a subset of N slices of interest. Due to the non-zero length impulse response of the filters employed in the inverse transform, K (multi-component) transformed slices are involved in the reconstruction of the N slices of interest, where $K > N$. The number of transformed slices K needed varies depending on the slice axis transform (RHAAR or RWT) and the number of transform levels used. Thus, even though a multi-component transform may improve the compression performance for an entire image, it may cause more data to be read and decompressed when only a subset of slices is desired. Accordingly, the aim of the following experiment is to evaluate the component scalability of the proposed coding scheme. To assess this, we have analyzed the number of bytes needed to decode a set of consecutive slices from the center of an image. Since the multi-component transform only provides a gain for images with $r > 0.87$, only such images are considered below.

Figure 8 shows the amount of data decoded (in MB) for the three tested coding approaches as a function of the number of slices decoded N. Note that, for small N, RWT+JPEG2000 and RHAAR+JPEG2000 result in more data being decoded, corresponding to a deterioration in performance with respect to JPEG2000. However, as N grows, the trend reverses, corresponding to an improved compression performance. The number of slices needed to achieve a positive gain for RWT+JPEG2000 and RHAAR+JPEG2000 is larger for images with lower correlation among slices, owing to the lower performance improvement achieved by the multi-component transforms. For small N, RHAAR outperforms RWT due to the fact that the RHAAR filters have shorter lengths than those of the RWT. However, as the number of retrieved slices is increased, RWT eventually produces better coding performance.

V. Conclusions

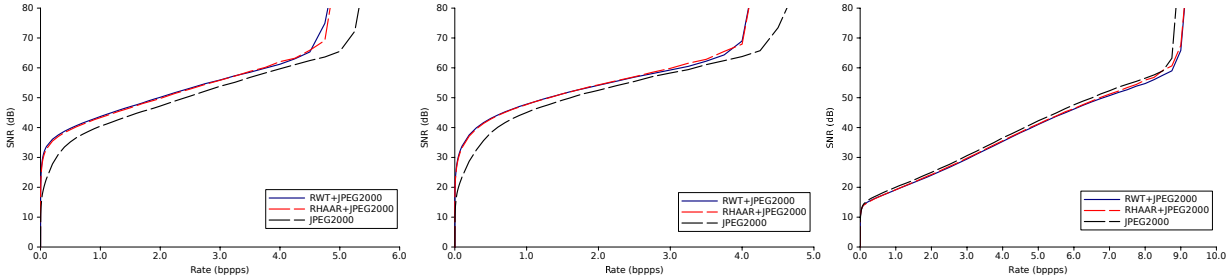
Given the extensive use of Computed Tomography and the huge volume of data, CT image coding is a relevant topic for practical medical scenarios and research. This manuscript proposes a new correlation modeling specifically designed for CT images. Our model is aimed to determine whether a multi-component transform helps improve the coding performance, both for lossless and for progressive lossy-to-lossless cases. This model employs CT image acquisition parameters to model the correlation among slices, without the computationally demanding step of precomputing image correlation. A study of the influence of correlation in 3D coding performance is carried out, which shows, for the evaluated corpus, that for images with $r > 0.87$, the RWT and RHAAR along the z dimension can provide significant coding gain.

Experimental results indicate that when the multi-component transform is profitable, RWT+JPEG2000 yields the best coding performance in terms of SNR, HDR-VDP and lossless bitrate, always outperforming RHAAR. On the other hand, when a specific subset of components needs to be retrieved, JPEG2000 or RHAAR+JPEG2000 can sometimes yield better rate-distortion performance, depending on the value of r and on the number of slices decoded.



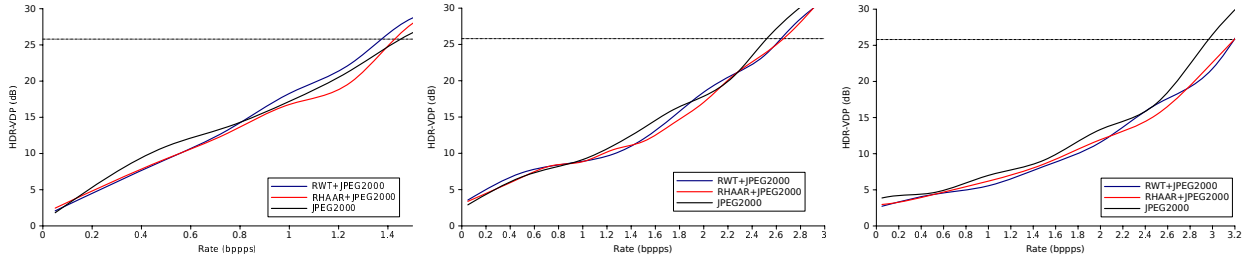
(a) RWT+JPEG2000 (b) RHAAR+JPEG2000

Figure 5: Coding rate gain vs modeled r. (a) and (b) respectively depict the coding gain between RWT+JPEG2000 and RHAAR+JPEG2000, with respect to JPEG2000.



(a) LS16-T125-D125_2 ($r = 0.961$) (b) LS16-T25-D25_2 ($r = 0.909$) (c) SS16-T1-D10_5 ($r = 0.554$)

Figure 6: Rate-distortion performance for RWT+JPEG2000, RHAAR+JPEG2000, and JPEG2000 for different images: (a) LS16-T125-D125_2, (b) LS16-T25-D25_2, (c) SS16-T1-D10_5.



(a) LS16-T125-D125_1 ($r = 0.954$) (b) SP4-T5-D5_1 ($r = 0.828$) (c) SS16-T1-D10_5 ($r = 0.554$)

Figure 7: HDR-VDP performance for RWT+JPEG2000, RHAAR+JPEG2000, and JPEG2000 for different images: (a) LS16-T125-D125_1, (b) SP4-T5-D5_1, (c) SS16-T1-D10_5.

REFERENCES

[1] American Cancer Society, "ACS," Feb 2013. [Online]. Available: http://www.cancer.org

[2] D. J. Brenner and E. J. Hall, "Computed tomography - an increasing source of radiation exposure," *New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.

[3] R. Choplin, "Picture Archiving and Communication Systems: an overview." *Radiographics*, vol. 12, pp. 127–129, Jan 1992.

[4] X. Cao and H. Huang, "Current status and future advances of digital radiography and PACS," *IEEE Eng. in Medicine and Biology Magazine*, vol. 19, no. 5, pp. 80–88, Sep 2000.

[5] D. J. Foran, P. P. Meer, T. Papathomas, and I. Marsic, "Compression guidelines for diagnostic telepathology," *IEEE Trans. on Information Technology in Biomedicine*, vol. 1, pp. 55–59, 1997.

[6] Digital Image and Communication in Medicine, "DICOM," Jul 2013. [Online]. Available: http://medical.nema.org/

[7] International Atomic Energy Agency, "Radiation protection of patients," Feb 2013. [Online]. Available: https://rpop.iaea.org

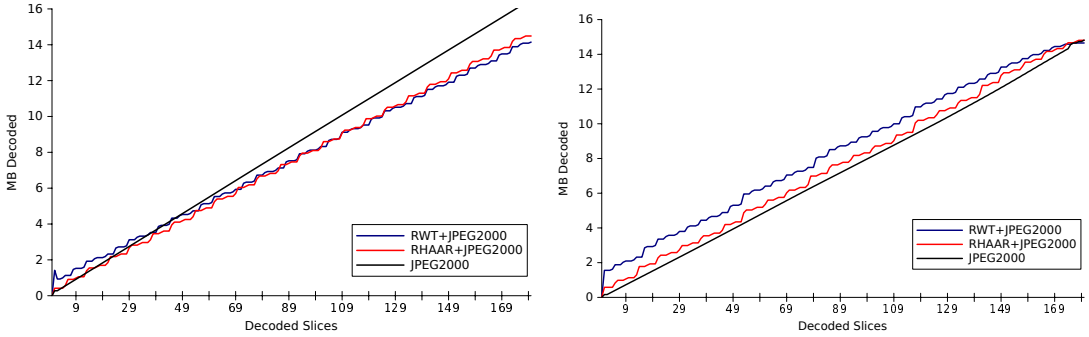
[8] E. Siegel, K. Siddiqui, J. Johnson, O. Crave, Z. Wu, J. Dagher, A. Bilgin, M. Marcellin, M. Nadar, and B. Reiner, "Compression of multi-slice CT: 2D vs. 3D JPEG2000 and effects of slice thickness," in *Medical Imaging 2005: PACS and Imaging Informatics*, ser. Proceedings of SPIE Medical Imaging, vol. 5748, 2005, pp. 162–170.

[9] D. Taubman and M. W. Marcellin, *JPEG2000: Image compression fundamentals, standards and practice. Chapter 3: Rate-Distortion Theory.* Norwell, Massachusetts 02061 USA: Kluwer Academic Publishers, 2002.

[10] J. Serra-Sagristà and A. Bilgin, "Special issue on image compression technologies for medical applications." *IEEE COMSOC MMTC E-letter*, vol. July, pp. 5–37, 2011.

[11] P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, and J. Cornelis, "Wavelet coding of volumetric medical datasets," *IEEE Trans. on Medical Imaging*, vol. 22, no. 3, pp. 441–458, Mar 2003.

[12] Z. Xiong, X. Wu, S. Cheng, and J. Hua, "Lossy-to-lossless compression of medical volumetric data using three-dimensional integer wavelet



(a) SS16-T1-D075_2 ($r = 0.979$) (b) SS16-T2-D2_2 ($r = 0.926$)

Figure 8: Data decoded (in MB) for JPEG2000, RWT+JPEG2000 and RHAAR+JPEG2000 as a function of N, where N is the number of slices decoded from the center of the image.

transforms," *IEEE Trans. on Medical Imaging*, vol. 22, no. 3, pp. 459 –470, Mar 2003.

[13] B.-J. Kim, Z. Xiong, and W. A. Pearlman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE Trans. on Circuits and Systems for Video Technol.*, vol. 10, pp. 1365–1374, Dec 2000.

[14] J. Xu, Z. Xiong, S. Li, and Y. Zhang, "3-D embedded subband coding with optimal truncation (3-D ESCOT)," *J. Appl. Comput. Harmon. Anal.*, vol. 10, p. 290315, May 2001.

[15] A. Agarwal, A. H. Rowberg, and Y. Kim, "Fast JPEG 2000 decoder and its use in medical imaging," *IEEE Trans. on Information Technology in Biomedicine*, vol. 7, pp. 184–190, 2003.

[16] S.-G. Miaou, F.-S. Ke, and S.-C. Chen, "A lossless compression method for medical image sequences using JPEG-LS and interframe coding," *IEEE Trans. on Information Technology in Biomedicine*, vol. 13, pp. 818–821, 2009.

[17] V. Sanchez, R. Abugharbieh, and P. Nasiopoulos, "Symmetry-based scalable lossless compression of 3D medical image data," *IEEE Trans. on Medical Imaging*, vol. 28, no. 7, pp. 1062–1072, Jul 2009.

[18] M. Penedo, W. Pearlman, P. Tahoces, M. Souto, and J. Vidal, "Region-based wavelet coding methods for digital mammography," *IEEE Trans. on Medical Imaging*, vol. 22, no. 10, pp. 1288–1296, Oct 2003.

[19] M. Penedo, M. Lado, P. Tahoces, M. Souto, and J. Vidal, "Effects of JPEG2000 data compression on an automated system for detecting clustered microcalcifications in digital mammograms," *IEEE Trans. on Information Technology in Biomedicine*, vol. 10, pp. 354–361, 2006.

[20] V. Sanchez, R. Abugharbieh, and P. Nasiopoulos, "3-D scalable medical image compression with optimized volume of interest coding," *IEEE Trans. on Medical Imaging*, vol. 29, no. 10, pp. 1808–1820, Oct 2010.

[21] J. Bartrina-Rapesta, J. Serra-Sagristà, and F. Aulí-Llinàs, "JPEG2000 ROI coding through component priority for digital mammography," *Elsevier Computer Vision and Image Understanding*, vol. 115, no. 1, pp. 59–68, Jan 2011.

[22] K. J. Kim, K. H. Lee, B. Kim, T. Richter, I. D. Yun, S. U. Lee, K. T. Bae, and H. Shim, "JPEG2000 2D and 3D reversible compressions of thin-section chest CT images," *Radiology*, vol. 259, no. 1, pp. 271–277, 2011.

[23] J. Gonzalez-Conejero, J. Bartrina-Rapesta, and J. Serra-Sagristà, "JPEG2000 encoding of remote sensing multispectral images with no-data regions," *IEEE Signal Processing Letters*, vol. 7, no. 2, pp. 251–255, Apr 2010.

[24] A. A. Kassim, P. Yan, W. S. Lee, and K. Sengupta, "Motion compensated lossy-to-lossless compression of 4-D medical images using integer wavelet transforms," *IEEE Trans. on Information Technology in Biomedicine*, vol. 9, pp. 132–138, 2005.

[25] V. Sanchez, P. Nasiopoulos, and R. Abugharbieh, "Novel lossless fMRI image compression based on motion compensation and customized entropy coding," *IEEE Trans. on Information Technology in Biomedicine*, vol. 13, pp. 645–655, 2009.

[26] Corporacio Sanitaria Parc Tauli, Sabadell, Spain, Feb 2013. [Online]. Available: http://www.cspt.es

[27] National Cancer Institute, "The cancer imaging archive," Feb 2013. [Online]. Available: http://http://www.cancerimagingarchive.net/

[28] K. J. Kim, B. Kim, R. Mantiuk, T. Richter, H. Lee, and K. H. L. Heung-Sik Kang, Jinwook Seo, "A comparison of three image fidelity metrics of different computational principles for JPEG2000 compressed abdomen CT images," *IEEE Trans. on Medical Imaging*, vol. 29, no. 8, pp. 1496–1503, 2010.

[29] "Information technology - JPEG2000 image coding system - Part 1: Core coding system," ISO/IEC, Dec. 2000.

[30] D. Clunie, "DICOM supplement 61: JPEG2000 transfer syntaxes," Nov 2001. [Online]. Available: ftp://medical.nema.org/medical/dicom/final/sup61_ft.pdf

[31] "Information technology - JPEG 2000 image coding system - Part 9: Interactivity tools, APIs and protocols," ISO/IEC, Dec. 2005.

[32] "Information technology - JPEG 2000 image coding system - Part 2: Extensions," ISO/IEC, May 2004.

[33] D. Taubman, "Kakadu software," 2013. [Online]. Available: {http://www.kakadusoftware.com/}

[34] J. E. Fowler and J. T. Rucker, *Hyperspectral data exploitation: Theory and applications. Chapter 14: 3D wavelet-based compression of hyperspectral imager*. Hoboken, NJ, USA: John Wiley & Sons Inc., 2007, pp. 379–407.

Juan Muñoz-Gómez received the B.E. and M.S. degrees in computer science from the Universitat Autònoma de Barcelona, Spain, in 2008 and 2009, respectively. He joined the Group on Interactive Coding of Images at the Universitat Autònoma de Barcelona in July of 2008 where he is currently working towards his Ph.D. degree. His research interests includes a wide range of image coding and signal processing topics, including highly scalable image and video coding systems, interactive image and video transmission, statistical modeling and medical image coding.

Joan Bartrina-Rapesta (S'06-M'10) received the B.Sc., B.E., M.S., and Ph.D. degrees in computer science from the Universitat Autònoma de Barcelona, Spain, in 2002, 2004, 2006, and 2009, respectively. He was awarded with a doctoral fellowship from Universitat Autònoma de Barcelona. He has collaborated in the development of BOI, a JPEG2000 Part 1 implementation. His research interests includes a wide range of image coding topics, including highly scalable image and video coding systems, region-of-interest coding, rate-distortion optimization techniques, distortion estimation, interactive image and video transmission, and medical image coding.

Michael W. Marcellin (S'81-M'87-SM'93-F'02) received the B.S. in Electrical Engineering from San Diego State University in 1983, and the MS and PhD in Electrical Engineering from Texas A&M University in 1985 and 1987, respectively. Since 1988, Dr. Marcellin has been with the University of Arizona, where he currently holds the title of Regents' Professor, and is the International Foundation for Telemetering Chaired Professor. His research interests include digital communication and data storage systems, data compression, and signal processing. He has authored or coauthored more than two hundred publications in these areas. He has received numerous honors, including six teaching awards.

Joan Serra-Sagristà (S'97-M'05-SM'11) received the Ph.D. degree in computer science from Universitat Autònoma de Barcelona (UAB), Spain, in 1999. He is currently an Associate Professor at Department of Information and Communications Engineering, UAB. From September 1997 to December 1998, he was at University of Bonn, Germany, funded by DAAD. His current research interests focus on data compression, with special attention to image coding for remote sensing and telemedicine applications. He serves as Associate Editor of IEEE Trans. on Image Processing. He has co-authored over one hundred publications. He was the recipient of the Spanish Intensification Young Investigator Award in 2006.