Universitat Autònoma de Barcelona

Departament d'Enginyeria de la Informació i de les Comunicacions

Contributions to Computed Tomography Image Coding for JPEG2000

by Juan Muñoz Gómez

Bellaterra, September 2013

Supervisor:

Dr. Joan Bartrina Rapesta

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Bellaterra, September 2013

Dr. Joan Bartrina Rapesta
(Principal Adviser)

Committee:

    Dr. Víctor Francisco Sánchez Silva

    Dr. Vicente González Ruiz

    Dr. Ian Blanes García

    Dr. Javier Melenchón Maldonado (substitute)

    Dr. Javier Ruiz Hidalgo (substitute)

A todas esas personas que
siempre son capaces de
hacerte sonreír

Abstract

Nowadays, thanks to the advances in medical science, there exist many different medical imaging techniques aimed at seeking to reveal, diagnose, or examine a disease. Many of these techniques produce very large amounts of data, especially from Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) modalities. To manage these data, medical centers use PACS and the DICOM standard to store, retrieve, distribute, and display medical images. As a result of the high cost of storage and transmission of medical digital images, data compression plays a key role.

JPEG2000 is the state-of-the-art of image compression for the storage and transmission of medical images. It is the latest coding system included in DICOM and it provides some interesting capabilities for medical image coding. JPEG2000 enables the use of use of windows of interest, access different resolutions sizes of the image or decode an specific region of the image.

This thesis deals with three different problems detected in CT image coding. The first coding problem is the noise that CT have. These noise is produced by the use of low radiation dose during the scan and it produces a low quality images and penalizes the coding performance. The use of different noise filters, enhance the quality and also increase the coding performance. The second question addressed in this dissertation is the use of multi-component transforms in Computed Tomography image coding. Depending on the correlation among the slices of a Computed Tomography, the coding performance of these transforms can vary even decrease with respect to JPEG2000. Finally, the last contribution deals with the diagnostically lossless coding paradigm, and it is proposed a new segmentation method. Through the use of segmentation methods to detect the biological area and to discard the non-biological area, JPEG2000 can achieve improvements of more than 2bpp.

Hoy en día, gracias a los avances en la ciencia médica, existen diversas técnicas de imágenes médicas destinadas a tratar de revelar, diagnosticar o examinar una enfermedad. Muchas de estas técnicas producen grandes cantidades de datos, en especial las modalidades de tomografía computarizada (CT), imágenes por resonancia magnética (MRI) y tomografía por emisión de positrones (PET). Para gestionar estos datos, los centros médicos utilizan PACS y el estándar DICOM para almacenar, recuperar, distribuir y visualizar imágenes médicas. Como resultado del alto coste de almacenamiento y transmisión de imágenes médicas digitales, la compresión de datos juega un papel clave.

JPEG2000 es el estado del arte en técnicas de compresión de imágenes para el almacenamiento y transmisión de imágenes médicas. Es el más reciente sistema de codificación incluido en DICOM y proporciona algunas características que son interesantes para la codificación de estas imágenes. JPEG2000 permite el uso de ventanas de interés, acceso a diferentes tamaños de la imagen o decodificar una región específica de ella.

Esta tesis aborda tres problemas diferentes detectados en la codificación de CT. El primer problema de la codificación de estas imágenes, es el ruido que tienen. Este ruido es producido por el uso de unas dosis bajas de radiación durante la exploración, lo cual produce imágenes de baja calidad y penaliza el rendimiento de la codificación. El uso de diferentes filtros de ruido, hace mejorar la calidad y también aumentar el rendimiento de codificación. La segunda cuestión que se aborda en esta tesis, es el uso de transformaciones multicomponente en la codificación de las CT. Dependiendo de la correlación entre las diferentes imágenes que forman una CT, el rendimiento en la codificación usando estas transformaciones puede variar, incluso disminuir con respecto a JPEG2000. Finalmente, la última contribución de esta tesis trata sobre el paradigma de la codificación diagnóstica sin pérdida, y propone un nuevo método de segmentación. A través de la utilización de métodos de segmentación, para detectar el área biológica y descartar la zona no-biológica, JPEG2000 puede lograr mejoras de rendimiento de más de 2bpp.

Avui dia, gràcies als avanços en la ciència mèdica, existeixen diverses tècniques d'imatges mèdiques destinades a tractar de revelar, diagnosticar o examinar una malaltia. Moltes d'aquestes tècniques produeixen grans quantitats de dades, especialment les modalitats de tomografia computada (CT), imatges per ressonància magnètica (MRI) i tomografia per emissió de positrons (PET). Per gestionar aquestes dades, els centres mèdics utilitzen PACS i l'estàndard DICOM per emmagatzemar, recuperar, distribuir i visualitzar imatges mèdiques. Com a resultat de l'alt cost d'emmagatzematge i transmissió d'imatges mèdiques digitals, la compressió de dades juga un paper clau.

JPEG2000 és l'estat de l'art en tècniques de compressió d'imatges per a l'emmagatzematge i transmissió d'imatges mèdiques. És el més recent sistema de codificació inclòs en DICOM i proporciona algunes característiques que són interessants per a la codificació d'aquestes imatges. JPEG2000 permet l'ús de finestres d'interès, accés a diferents grandàries de la imatge o la decodificació una regió específica d'ella.

Aquesta tesi aborda tres problemes diferents detectats en la codificació de CT. El primer problema de la codificació d'aquestes imatges, és el soroll que tenen. Aquest soroll és produït per l'ús d'unes dosis baixes de radiació durant l'exploració, produint imatges de baixa qualitat i penalitzant el rendiment de la codificació. L'ús de diferents filtres de soroll, fa millorar la qualitat i també augmentar el rendiment de codificació. La segona qüestió que s'aborda en aquesta tesi, és l'ús de transformacions multi-component en la codificació de les CT. Depenent de la correlació entre les diferents imatges que formen una CT, el rendiment en la codificació usant aquestes transformacions pot variar, fins i tot disminuir pel que fa a JPEG2000. Finalment, l'última contribució d'aquesta tesi tracta sobre el paradigma de la codificació diagnòstica sense pèrdua, i proposa un nou mètode

de segmentació. A través de la utilització de mètodes de segmentació, per detectar l'àrea biològica i descartar la zona no-biològica, JPEG2000 pot aconseguir millores de rendiment de més de 2 bpp.

Acknowledgements

The acknowledgements are written in my mother tongue.

Hay mucha gente a la que quiero agradecer que esta tesis haya podido ser finalizada, pero por si me olvido de alguien, gracias a todos.

Hacer una tesis no es un camino fácil, y menos es dirigirla. Por eso, la primera persona a la que quiero agradecer esta tesis es a mi director y amigo Joan Bartrina. Se que no ha sido un camino fácil, y que han habido unas cuantas discusiones, pero el camino ha merecido la pena y hemos llegado al final.

Como no, quiero dar las gracias a todo GICI, que no solo son buenos en su trabajo sino que también te hacen sentir como en familia. Hemos tenido momentos de mucho trabajo, momentos muy divertidos y baños en una piscina rodeado de nieve tomando agua, siempre agua. En particular, quiero agradecer a Joan Serra la oportunidad de poder realizar este doctorado en este maravilloso mundo de las imágenes, y todo lo que me ha enseñado y transmitido desde su experiencia como investigador. I want to specially thank to Prof. Michael Marcellin for their advices and their signal processing lessons encouraging me in this topic.

Continuando con los agradecimientos, quiero agradecer a todos los miembros del dEIC esos buenos momentos que hemos pasado juntos como las *excursions de departament*, el dEIC football team, y esas sesiones de 2 horas continuas de Sergi haciendo chistes en el *menjador*. Quiero agradecer y animar a los *becaris petits* para que en breve sean ellos los que están escribiendo los agradecimientos de sus respectivas tesis. Y como no, a todos vosotros Lorena, Jaume, Bernat, Lau, Abraham, Ian, Carlos por todos esos momentos que nos hemos reído tantísimo juntos ya sea de BBQ en

Can Pernas, en el Osaka o en cualquiera de los despachos del departament donde pasábamos tan buenos ratos.

También quiero agradecer a Fernando, Lino y Jorge por todos esos momentos de *consejos primales* el los cuales tantas cosas hemos solucionado.

Finalmente, y no menos importante, quiero agradecer a mi familia, en especial a mi madre y a mi hermana, por el apoyo, cariño e interés que me han transmitido durante este tiempo.

A todos vosotros,

Gracias

Preface

In the last decades, thanks to the advances in Physical science and Medicine, there exist many different medical image techniques aimed at diagnosing or addressing certain diseases. Most of these techniques produce a large amount of data sets which need to be stored and transmitted. One may wonder whether modern compression systems could indeed improve the cost needed to access and transmit such data. As this thesis will later show, they do.

This thesis has been developed in the Group on Interactive Coding of Images (GICI). The work began in September 2008 with the study of the state-of-the-art of still image compression techniques. Our first step was to try different state-of-the-art coding techniques in Computed Tomography. This helped to evaluate and understand the compression performance in Computed Tomography. Throughout the evaluation of different coding techniques, several ideas were tested to improve the coding performance without penalizing the the medical diagnostic.

This dissertation shows much of the work that the student has done during the PhD degree in the Departament d'Enginyeria de la Informació i les Comunicacions at the Universitat Autònoma de Barcelona. This dissertation proposes three different contributions to improve the coding performance of Computed Tomography Images. In addition, this dissertation includes in an introduction to medical imaging and a image coding chapter for a better understanding of the thesis.

x

Contents

Chapter 1

Introduction

This chapter motivates the use of digital images in medical scenarios and some of the problems tackled in this thesis. It briefly reviews the most relevant types of medical image, doing a special attention in Computed Tomography, which is the image modality used in this thesis. Furthermore, this chapter indicates how this thesis is organized.

1.1   Motivation

The use of medical imaging has increased rapidly, being fundamental in the medical community. Computed Tomographies are one of the most used imagery. In 2007, it was estimated that more than 62 million CT scans were obtained per year in the United States [1]. To manage these data, medical centers use Picture Archiving and Communications Systems (PACS) [2] to store, retrieve, distribute, and display medical images. PACS are commonly constituted of large computer networks, servers, and workstations [3, 4]. The Digital Imaging and Communications in Medicine (DICOM) standard [5] specifies the format used to store and distribute images in PACS. Because of the large number of images managed, data compression plays a key role in DICOM.

There are several image compression standards compatible with DICOM [6], such as JPEG [7], JPEG-LS [8] and JPEG2000 [9]. JPEG2000 is the latest coding sys-

tem included in DICOM, and it provides some interesting capabilities, explained in Chapter 2, which permit the use of windows of interest, access different resolutions sizes of the image or decode an specific region of the image. These features makes JPEG2000 a good choice for medical community.

However, not only the storage is a problem in the medical centers. Thanks to the advances in information and communications engineering, the use of telemedicine services has increased, and the medical diagnosis applications on smartphones and tablets are a reality [10]. The advances in communications and the growth of new technologies have enabled the development new standards that could achieve a high speed ratios. This high speed ratios could be achieved in certain conditions, but in practice they cannot be achieved [11]. Using JPEG2000, the transmission neither the storage of the whole image is not needed because it is possible to retrieve and transmit only the desired parts.

1.2 Medical Imaging

Medical Imaging is a set of techniques used to create images of the human body for different medical purposes, such as reveal, diagnose or examine a disease. These techniques require the application of some form of energy into the human body. Depending on the energy intensity applied different type of medical images are obtained, such as Magnetic Resonance Imaging, Endoscopy, Radiography, Ultrasound, Nuclear Imaging, and Computed Tomography. All these imaging techniques are briefly described below. Figure 1.1 shows a distribution of the medical imaging techniques over the electromagnetic spectrum.

- **Endoscopy** was introduced by Philip Bozzini in 1806. He introduced the use of a light conductor for the examinations of the canals and cavities of the human body. This procedure lets the physician look inside human body using a camera. The most common medical scenarios where they are used are surgery and gastroenterology.

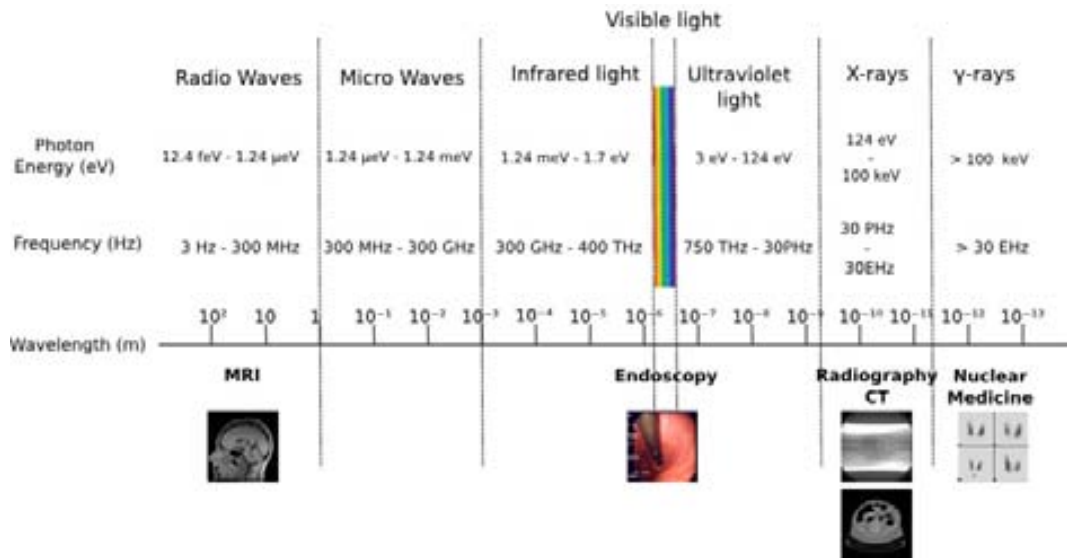- **X-ray imaging** was discovered by Wilhelm Conrad Röntgen in 1895. It is

Figure 1.1: Distribution of the medical imaging techniques over the light spectrum.
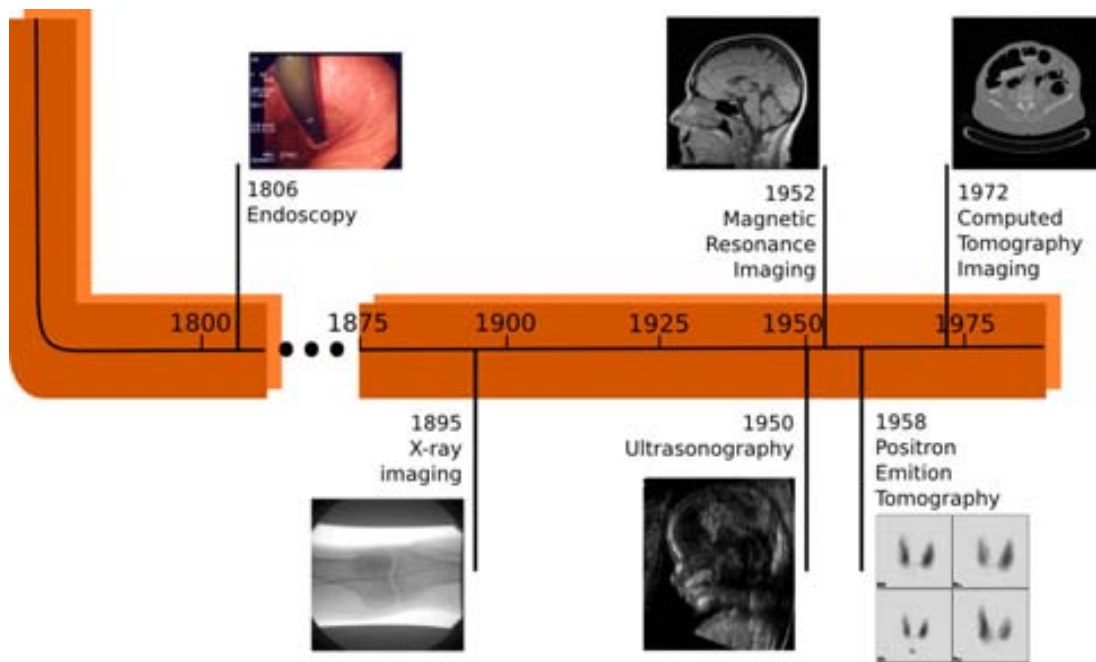

Figure 1.2: Timeline of the Medical Imaging techniques.

a technique that permits capture images of inside of the human body. The technique consists on projecting an amount of X-rays beams toward an specific area of the human body. Depending on the density tissue part of the projected beam is absorbed, and the remnant beam is captured by a detector with different intensities in the X-ray image. The problem is that during the acquisition process the patient is exposed to ionize radiation, which at high expositions increase the risk of death from cancer [12].

- **Ultrasonography** is a technique that was developed by John J. Wild and his colleagues in the 1950s. It is a diagnostic imaging technique based on ultra-sounds. It has several advantages, which make it ideal in numerous situations, in particular in those studies that the physicians want to evaluate the function of moving structures in real-time.

- The first **Magnetic Resonance Imaging** (MRI) scanning device was produced in 1952 by Herman Car. MRI is a three dimensional (3D) medical imaging technique that visualize internal structures of the body in more detail that X-Ray imaging. MRI is especially useful in imaging the brain, muscles and the heart.

- **Nuclear Medicine** was introduced in 1958 by Hal Anger. It consists with the application of radioactive substances in the diagnosis and treatment of disease. These substances are administered to the patient to localize organs or specific cells. This permits to create images with the extent of a disease process in the body, based on the cellular function and physiology, instead of the physical changes in the anatomy. As a difference with X-rays, where the images are acquired from outside to inside; nuclear medicine captures the radiation from inside to outside of the body.

- **Computed Tomography** (CT) produces a 3D image of the structures of a determined volume inside of the body through the acquisition of several X-ray images. The first commercial CT scanner was invented by Sir Godfrey Hounsfield in 1972 at Atkinson Morley Hospital in Wimbledon. The next section
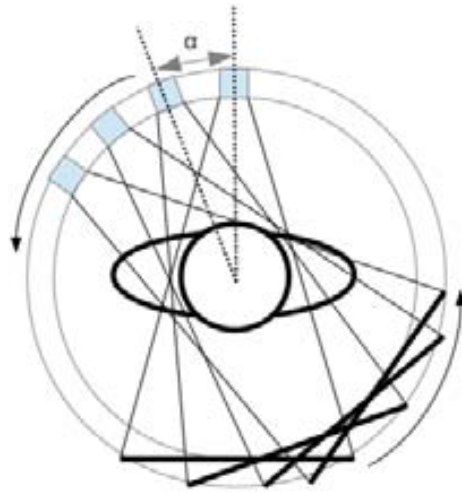
Figure 1.3: Example of the different line attenuation measurements which produce a CT.

includes more details of this imaging technique.

1.2.1 Computed Tomography

CT combines special X-ray beam generator equipment with sophisticated software to produce 3D images of the inside of the human body; showing organs, bones, soft tissue and blood vessels with greater clarity than standard X-rays. Hence, CT images permits to radiologists diagnose several diseases or disorders such as cancer, infectious diseases, appendicitis, cardiovascular diseases, trauma and musculoskeletal disorders [12].

One of the most important advantage of the CT is the high-contrast resolution, which allows to differentiate substances that have small differences in density. Other advantage is the perspective visualization, physicians can view CT images in the axial, coronal, or sagittal planes, depending on the diagnostic procedure. Figure 1.4 shows a graphical example of these three views.

CT devices capture a large set of X-ray images, called 2D projections, yielding line attenuation measurements for all possible angles and for a determined distance from the center. Figure 1.3 shows schematically the acquisition procedure of the 2D
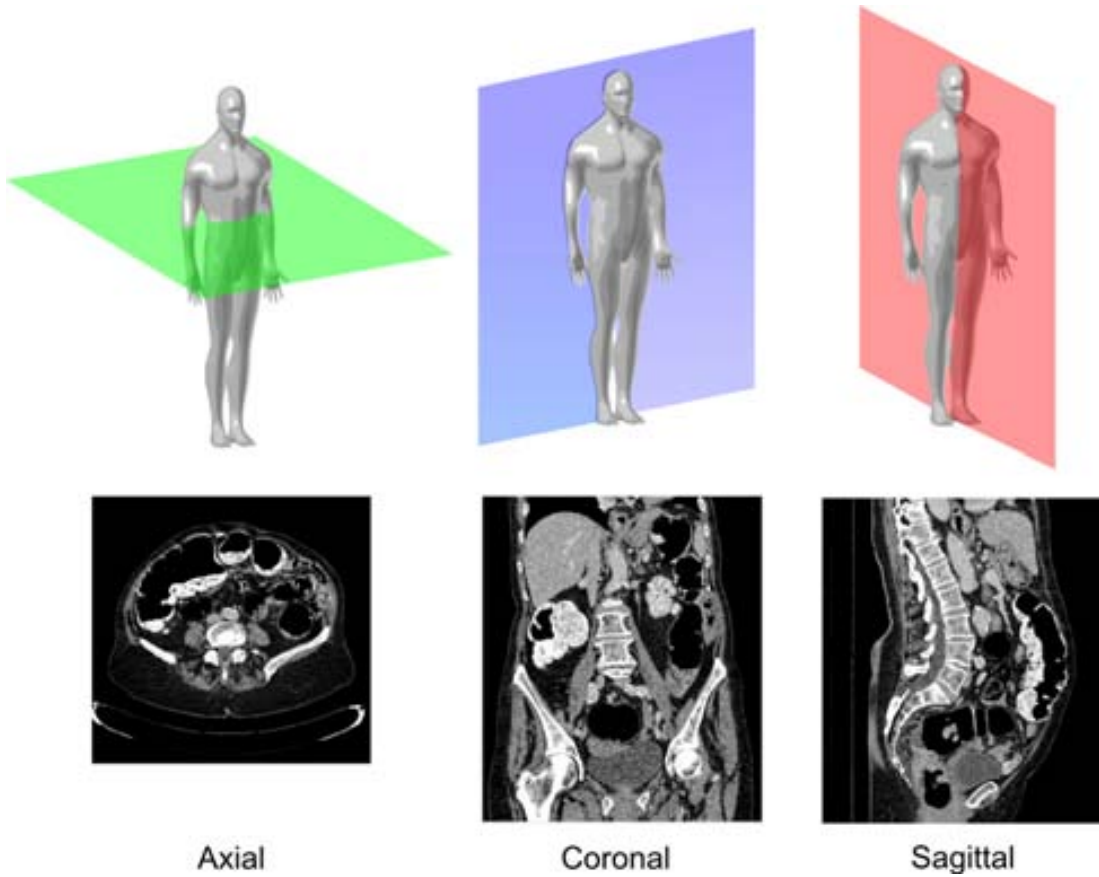
Figure 1.4: Axial, Coronal and Sagittal views.

projections. After the acquisition process, the 2D projections are reconstructed to a 3D image through high computational algorithms [13].

To facilitate the interpretation of CT images by specialists, CT values are given in Hounsfield Units (HU) [14], which allows the identification of biological substances according to different value ranges. A CT value is defined as

$$\text{CT value} = 1000 \times \frac{\mu - \mu_{H_2O}}{\mu_{H_2O}}$$

where μ is the linear attenuation coefficient of the evaluated substance and μ_{H_2O} is the linear attenuation coefficient of water. Figure 1.5 shows the HU scale and the represented biological substances.
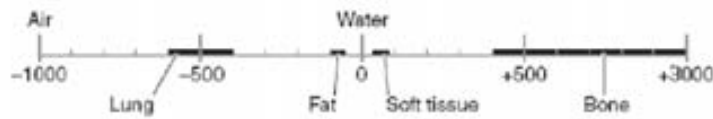
Figure 1.5: The Hounsfield Units scale permits to identify the substances.

As X-ray images, CT images radiates the patient , which means that the dose and the number of CT scans performed increases the probability of inducing cancer, teratogenesis, cognitive decline or heart diseases. This does not means that CT are unsafe medical imaging technique but the effects of continuous ionizing radiation can be harmful and potentially lethal [12].

1.3   Contributions and Thesis organization

This thesis is organized as follows: Chapter 2 reviews the medical image coding literature and gives the main concepts an techniques of image coding. Chapter 3, 4 and 5 can be read independently since each of them corresponds to a scientific contribution presented in this thesis.

A image coding expert would probably skip Chapter 2.

Chapter 3 introduces the first scientific contribution of this thesis, which presents the use of a noise filter over CT to improve the coding performance of the JPEG2000. These contribution includes an explanation of the noise problem in CT images, an study of the state of the art of different noise filtering techniques and its fields of application, and finally the impact of the noise filter in Computed Tomography image coding. These studies include the coding performance in 2D transforms and also over the different evaluated multi-component transforms. Detailed analysis and results are presented and discussed in Chapter 3 of this document. This research was published in

Juan Muñoz-Gómez, Joan Bartrina-Rapesta, Michael W. Marcellin, and Joan Serra-Sagristà, "Influence of Noise Filtering in Coding Computed Tomography

with JPEG2000," In Proceedings of the IEEE Data Compression Conference, March 2011, pp 413-422. Digital Object Identifier 10.1109/DCC.2011.48.

Chapter 4 presents the second scientific contribution of this thesis, which is focused on the evaluation of the multi-component transforms and its relation with the correlation among the slices of the CT. Finally, a statistical model that estimates the correlation between slices using the acquisition parameters is developed. This work has been published in

Juan Muñoz-Gómez, Joan Bartrina-Rapesta, Michael W. Marcellin, and Joan Serra-Sagristà, "Correlation Modeling for Compression of Computed Tomography Images", IEEE Journal of Biomedical and Health Informatics, May 2013, in Press. Digital Object Identifier 10.1109/JBHI.2013.2264595

Chapter 5 introduces the last contribution of this thesis, a segmentation method for CT. Using segmentation methods it is possible to select the biological area and remove everything that is not relevant for medical diagnosis. Besides this, the fact that such non-relevant information be removed, permits a considerable improvements in the coding performance.

Finally, Chapter 6 contains a brief summary of this work, draws conclusions and exposes the future work.

A list of all contributions produced during this PhD. in collaboration with other member of the group is detailed in Appendix A. The image corpus used to perform the experiments are described in the Appendix B.

Chapter 2

Image Coding

This chapter presents a brief introduction to image coding, the fundamental concepts and techniques within the core coding system of JPEG2000, the state of the art of medical image coding and finally the metrics used to evaluate the experiments of this thesis.

2.1 Introduction

Since the Information Theory was formulated by C.E. Shannon in 1948 [15], relevant works based on this topic began to appear. Data compression began to be a fundamental research area due to the increase of amount of data that a computer could generate.

Data compression research is aimed to improve the efficiency of such methods in terms of storage size and transmission. Several methods appeared on this field since Shannon developed its theory. Huffman in 1952 [16], Golomb-Rice in 1966 [17], and arithmetic coding in 1979 [18] are methods that encode information using the number of occurrences of a symbol to avoid redundancy and then produce a shorter message.

For a more efficient compression, it is useful to know the data previously. The images usually have a nature that make the compression more efficient, given that have much spatial redundancy. Therefore, before applying any coding method to the data, mathematical operations can be performed to reduce data redundancy and thus

achieve greater compression performance.

In 1974 an important step in the field of the image compression was done with the introduction of the discrete cosine transform (DCT) [19], which compacts the energy of the image in the transformed domain concentrating most of the information in a few coefficients. In 1986, the Joint Photographic Experts Group (JPEG) [20] was founded with the aim at creating standards for image compression. In 1991, they published the first draft of the JPEG (ISO94). This coding system has three basic stages: DCT transform, a scalar quantization, and a Huffman encoder. In 1992, M.Antonini et al. [21] presented an article, which proposed the use of Wavelet transform which compact the information more efficiently than DCT. With the apparition of the wavelet transform, new image compression algorithms were developed, such as Embedded ZeroTree Wavelet [22], Index Coding [23], Set Partitioning In Hierarchical Trees [24]. They take the advantage of the inherent similarities across the subbands in a wavelet decomposition of an image. All of these methods revolutionized the image compression topic because all of them introduced a new concept: scalability by quality which means that if we stop at some point the transmission of a coded image, any of these algorithms could decompress the image at a quality determined by the number of bits received.

2.2   JPEG2000 Coding Standard

JPEG2000 is a powerful image compression standard developed by JPEG committee in 2000. Based on a wavelet coding scheme, JPEG2000 achieves high compression ratios in lossy, lossless and progressive lossy-to-lossless compression, supports more than 16-bits of signed or unsigned data, includes tools for interactive transmission [25], and provides some interesting capabilities for 3D image coding, such as support for multi-component transforms [26], aimed to exploit redundancy between slices [1]. An

---

[1]It is worth noting that the JPEG2000 standard uses 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", respectively. Hence, in this thesis the term slice is used to refer to a component. 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.

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 us 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 codestream.

The JPEG2000 image coding system consists of 14 parts:

- **Part 1 Core compression system:** It includes the syntax of the JPEG2000 codestream and the necessary steps involved for coding and decoding JPEG2000 images.

- **Part 2 Extensions of the compression system:** It defines various extensions to the Part 1, like, for instance, different wavelet filters, multi-component transforms, quantization alternatives, an alternative way of ROI coding, a new file format named JPX, etc.

- **Part 3 Motion JPEG2000 (MJP2):** It defines a file format called MJ2 (or MJP2) for motion sequences of JPEG2000 images

- **Part 4 Conformance testing:** It specifies test procedures for both encoding and decoding processes, including the definition of a set of decoder compliance classes.

- **Part 5 Reference Software:** Two source code packages that implement the JPEG2000 Part 1 are presented. The two coders were developed alongside Part 1 and were used to check interoperability. One is called JASPER [27], which is written in C, and the other one is JJ2000 [28], which is written in Java.

- **Part 6 Compound image file format:** It is used to store multi-page documents with many objects per page, using many other coding technologies as well.

- **Part 7 This part was abandoned**

- **Part 8 Secure JPEG2000 (JPSEC):** It defines content protection, technology protection, and it allows applications to generate, consume, and exchange JPEG2000 secured bitstreams.

- **Part 9 Interactive transfer protocol (JPIP):** It presents standard protocols supporting interactive transmission of JPEG2000 imagery.

- **Part 10 Volumetric JPEG2000 (JP3D):** It is concerned with the coding of three-dimensional and floating-point data sets, extending JPEG2000 from planar to volumetric images.

- **Part 11 Wireless JPEG2000 (JPWL):** It includes tools and methods to achieve the efficient transmission of JPEG2000 imagery over an error-prone wireless network.

- **Part 12 ISO base media file format:** It provides the definition of the ISO media file, providing an extensible format which facilitates interchange, management and editing.

- **Part 13 An entry level JPEG2000 encoder:** It defines a normative entry level JPEG2000 encoder providing one or more optional complete encoding paths that use various features defined in ISO/IEC 15444.

- **Part 14 XML representation and reference:** It specifies an XML document, referred to as JPXML, which is designed primarily for representing JPEG2000 file format and marker segments in codestream, and referring method of internal data in a JPEG 2000 image.

2.2.1   Core coding system

The core coding system of JPEG2000, depicted in Figure 2.1, is constituted by four main stages: sample data transformations, sample data coding, rate-distortion optimization, and codestream reorganization. The two first stages correspond to the
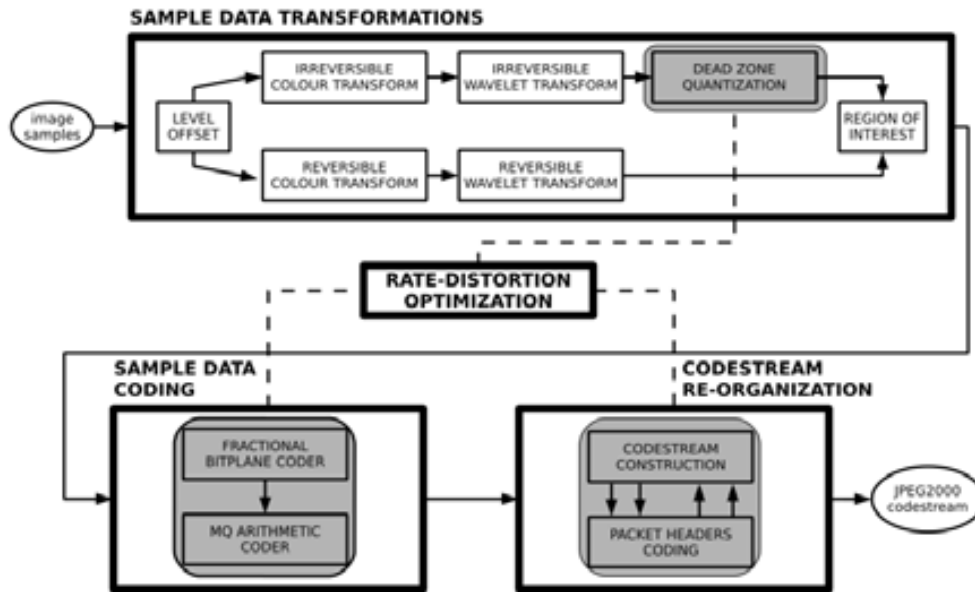
Figure 2.1: Scheme of the Part 1 of the JPEG2000 coding standard.

proper coding stage, while the codestream re-organization and the rate-distortion optimization are the responsible of the structure of the codestream depending on the required scalability.

Sample data transformations

Sample data transformations stage prepares the image to be coded more efficiently in the next stage. This stage could be applied in two different paths, reversible (for lossless compression) and irreversible (for lossy compression).

The first step of this stage is a level offset. Each unsigned sample values of the original image are level shifted (DC offset) by subtracting a fixed value of 2^{B-1} where B is the number of bits of the original image. This operation simplifies certain implementation issues, such as numerical overflow, and has no effect on the coding efficiency [29].

Next is the colour transform which is only applied in case that the image has the RGB channels. This transform takes the original RGB space and transforms
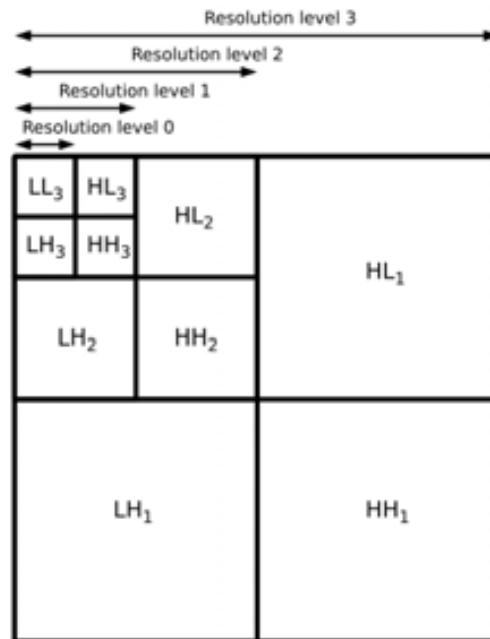
Figure 2.2: Dyadic image decomposition of the DWT.

it in a different color space. In the irreversible path the original components are transformed to the Y, C_b, C_r representation by a linear transform, where Y stands for the luminance and C_b and C_r for the blue and red chrominance respectively. For the reversible path, a non-linear transform and is used to transform the RGB components in a similar color space as Y, C_b, C_r.

The third step is the Wavelet Transform. The 9/7 Discrete Wavelet Transform (DWT) is applied for irreversible path, and the 5/3 Reversible Wavelet Transform (RWT) is applied in the reversible path. Although the filter applied to both cases are different, these transforms generates the same compact representation of the image. The transform is performed in two dimensions and then it produces four subbands, which contain the Low-Low (LL), High-Low (HL),Low-High (LH) and High-High (HH) frequencies in the horizontal and vertical directions respectively. Figure 2.2 shows a graphical representation of this structure. The LL subband can be decomposed successively into four smaller subbands, increasing the number of wavelet levels and obtaining the usual dyadic image decomposition.

The last step for the irreversible path is the dead zone quantization. This operation convert the floating point coefficients provided by the DWT to an integer, which will be encoded. Since lossless compression needs a reversible process, quantization is not used when compressing in lossless mode.

Sample data coding

Once the image is transformed, the next stage is the Sample data coding. This stage uses the Embedded Block Coding with Optimized Truncation [30] (EBCOT) paradigm, where each subband is divided in a small data structures called code-blocks, where each code-block is encoded independently. The size of the code-block are specified by the encoder. The restrictions of the code-block size are two: it must be and integer power of two, and the total number of coefficients in a code-block can not exceed 4096.

Each bitplane is coded by three different coding passes which provides a set of large potential truncation points, each one at the end of each coding pass. Once the code-block is coded by the bitplane coder, the arithmetic coder called MQ-coder is performed. The MQ-coder is aimed of exploiting the spatial redundancy of the coefficients within the code-block using contextual information of the coefficients. This contextual is utilized in the MQ-coder to adjust the probabilities of the incoming symbols. Finally, the MQ-coder produces an embedded codestream for each code-block.

Rate-distortion organization

The rate-distortion optimization stage is the responsible to manage the bitrate and the distortion of the final codestream. It maximizes the quality of the final codestream given a certain bitrate. The criterion for optimally can be based on mean squared error (MSE) between the original and reconstructed image, visual distortion, or any other metric that determine differences with the original image. Most of the JPEG2000 implementations employ the Post Compression Rate-Distortion optimization (PCRD) [30] method to conduct this optimization process, which is based on the

MSE.

Codestream reorganization

The final stage of the JPEG2000 pipeline is the codestream re-organization. It includes the headers and auxiliary data needed to properly identify the content of the codestream, and organizes it in containers that encapsulate and sort the bitstream segments using one or several progression orders. In JPEG2000 there are four different progression orders.

- Layer–resolution–component–position (LRCP)

- Resolution–layer–component–position (RLCP)

- Resolution–position–component–layer (RPCL)

- Position–component–resolution–layer (PCRL)

2.2.2   Multi-component transforms

Multi-component transforms are defined in the Part 2 of JPEG2000, they are designed to exploit the inter-component redundancy that images with a large number of components/bands have. The use of these transforms may improve the coding performance of JPEG2000 coder if there exist a sufficient redundancy among components. The transform is applied to each pixel in the image, in the 3rd direction. After applying multi-component transformation, each frame is compressed independently using the Part 1 of the JPEG 2000 compression standard. It is important to remark that Part 2 of JPEG2000 is defined in a way that more than one transform can be applied iteratively, and diverse transforms can be applied independently to different component sets or component collections.

In the Part 2 of the JPEG2000 standard there are defined 2 different types of multi-component transforms:

- Array Based Transforms: Based on a matrix multiplication

$$Y = A_t I,$$

where Y is the resulting image of the the application of the transform A_t to image I. Examples of this transform are: Part 1 color transforms, Karhunen Loève Transforms (KLT), etc.

- Wavelet transforms: There are several wavelet transforms in the literature that can be implemented in Part 2. For example, the 9/7 DWT, the 5/3 RWT, and the Reversible Haar wavelet.

2.3 Medical Image Coding Literature

Nowadays, medical imaging is prominent in data compression research, as witnessed by a recent special issue in image compression for medical applications [31]. Regarding research presented in the last decade, Schelkens et al. [32] presented an extensive review of 3D wavelet coders, and proposed three different coding methods. Two of these were based on the wavelet transform, while one was based on the discrete cosine transform. Xiong et al. [33] proposed a 3D modification of set partitioning in hierarchical trees [34] and of embedded subband coding with optimal truncation [35] to obtain better coding performance than the then state of the art 3D integer wavelet coding. Other authors have exploited image symmetries, as in [36], with a coding scheme that predicts the value of wavelet coefficients on a block-by-block basis. On the other hand, diagnostically lossless coding methods, which are aimed at encoding only the biological area of the image, or just the relevant area detected by computer-aided diagnosis procedures. This methods guarantees the perfect reconstruction of only those Regions of Interest (ROIs) in the image that are used for diagnostic purposes. An important number of diagnostically lossless coding methods are reported in the literature. Penedo et al. [37] presented object-based extensions for the set partitioning in hierarchical trees and the set partitioning embedded block coder algorithms, both of which were tested with digital mammography. Sanchez et al. [38], proposed a 3D

scalable compression method for medical images with optimized volume of interest coding. More recently, Bartrina et al. [39] introduced an ROI coding method for digital mammography based on component priority. Other authors developed lossless coding methods which rely on segmentation to identify the ROI, followed by lossless coding of the segmented area. In 2011, Kim et al. [40] presented a preprocessing method for CT images that replaces the pixels of the non-body region by a constant value. Finally, Xu et al. [41] developed a diagnostically lossless method which includes a segmentation method using ray-casting and α-shapes followed by a lossless coding compressor compliant with DICOM. All of these methods are aimed at maximizing the data redundancy in the original image, which is similar to an approach employed in the framework of remote sensing scenarios [42].

2.4   Image Coding Metrics

The purpose of image compression is to represent the image I with a string of bits, called encoded bitstream or codestream, which is denoted as L. The objective is to keep the length of L as small as possible, with a minimum loss of information from the original image (lossy compression) or without a loss of information (lossless compression). Below are defined most of the relevant and used image metrics in this thesis.

Compression Ratio (CR)

The CR is one of the most common measures that are used to express the efficiency of a compression method. It is defined as

$$\text{CR} = \frac{size(L)}{size(I)}.$$

The CR could be showed as a number between 0 and 1 which means that the images is compressed. It is also shown as a fraction where the numerator is 1 and the denominator is the number of times that the original image is bigger than the coded

image. For example if we have a CR of 0.25 means that the original image is 4 times bigger than the original and can be expressed as 1 : 4.

Bitrate

The bitrate is the average number of bits that we need to store each pixel of the compressed image. It is estimated as

$$\text{Bitrate} = \frac{size(L) \times 8}{size(I)}.$$

and its units are bits per pixel (bpp).

The lossy performance of a coding system is established as a trade-off between the rate achieved by the coding process, and the distortion of the recovered image after decoding $\hat{I}$. In this thesis two different metrics are used to evaluate the lossy performance.

Signal-to-Noise-Ratio

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

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

where

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

and

$$\sigma^2 = \frac{1}{N_x} \frac{1}{N_y} \sum_{j=1}^{N_y} \sum_{i=1}^{N_x} (I_{ji} - \bar{I})^2.$$

I_{ji} and $\hat{I}_{ji}$ denote, respectively, the values of the original samples and decompressed samples at position ji corresponding to the vertical and horizontal axis, and $\bar{I}$ the mean of the image I. N_x and N_y are the number of pixels in a column and in a row, respectively. Furthermore, if the image quality is calculated for volumetrical images, the *SNR* is defined as,

$$SNR = \frac{1}{N_z} \sum_{k=1}^{N_z} (SNR(I_k, \hat{I}_k)).$$

Structural SIMilarity (SSIM)

The SSIM [43] is based on the assumption that the Human Visual System is highly adapted to extract structural information from visual scenes, and therefore, that a measurement of structural information change can provide a good approximation of perceptual image fidelity. Values of SSIM are between -1 and 1. A value of 1 is only reachable in the case of two identical sets of data.

$$\text{SSIM}(I, \hat{I}) = \frac{(2\mu_I \mu_{\hat{I}} + c_1)(2\sigma_{I\hat{I}} + c_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + c_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + c_2)},$$

with μ_I is the average of I; $\mu_{\hat{I}}$ the average of $\hat{I}$; σ_I^2 is the variance of I; $\sigma_{\hat{I}}^2$ is the variance of $\hat{I}$; $\sigma_{I\hat{I}}$ is the covariance of I and $\hat{I}$; and $c_1 = (k_1 L)^2$ and $c_2 = (k_2 L)^2$ where L is the dynamic range of the image I and $k_1 = 0.01$ and $k_2 = 0.03$. As *SNR*, if the similarity is calculated for volumetrical images, the *SSIM* is defined as,

$$SSIM = \frac{1}{N_z} \sum_{k=1}^{N_z} (SSIM(I_k, \hat{I}_k)).$$

Chapter 3

Noise Filtering in CT coding

In this chapter, an evaluation of the noise filtering effect in conjunction with different decorrelator strategies, and a new encoding scheme for CT images through the JPEG2000 framework is presented.

3.1 Introduction

As it is explained in the Chapter 1, X-rays are a form of ionizing radiation, and consequently, an exposure to this radiation can be a health hazard. Since an additional irradiation increases the risk of getting cancer by 0.6-1.8% for a 75 year old person [44], clinical research is aimed to reduce the radiation dose in each test [45]. The radiation reduction introduces noise generating CT images with a poorer quality than those obtained with higher radiation doses, which complicates the medical diagnosis.

Noise filtering techniques are essential to improve the quality of CT images acquired at low radiation dose [45]. In the literature there exist many noise filtering techniques, aimed to enhance the biological structures to ease the medical diagnosis [46, 47, 45]. Figure 3.2 depicts a component of a CT image, obtained at low radiation dose, before and after noise filtering (The reader is invited to increase the pdf reader zoom level for a clearer illustration).

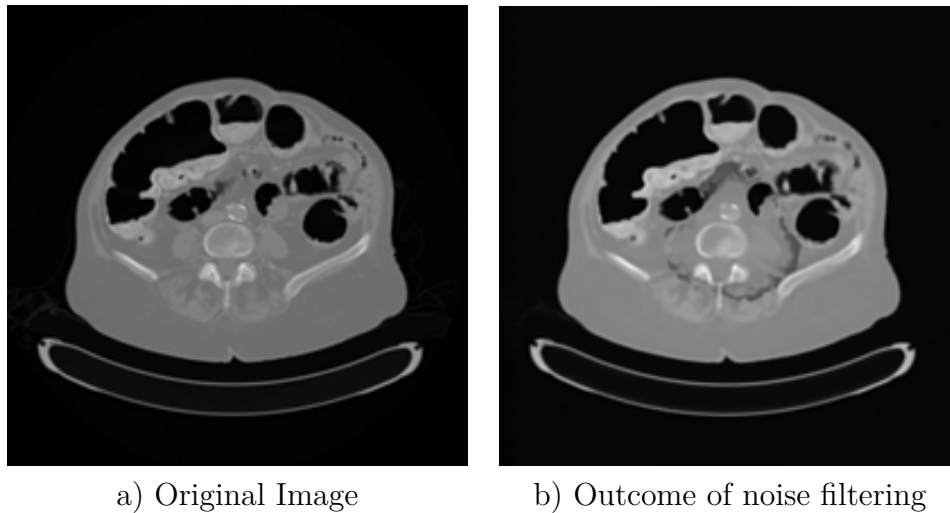| a) Original Image | b) Outcome of noise filtering |

Figure 3.1: A component of a CT image, a) obtained at low radiation dose, and b) after noise filtering.

3.1.1   Review of Noise Filtering in CTs

Aimed to improve the image quality, while maintaining the contrast, different filtering methods exist in the literature. There are two main strategies in noise filtering depending on when the filter is applied: pre-reconstruction, performed during the raw data scanning; and post-reconstruction, applied when the image has already been formed. Although pre-reconstruction techniques work especially well [47], the raw data is often unavailable because manufacturers don't provide access to this data.

Regarding post-reconstruction techniques, Regularized Perona-Malik (RPM) [48] was presented in 1992, obtaining competitive results for enhancing vessels with small diameter. Subsequently, Edge-Enhancing Diffusion (EED) and Coherence-Enhancing Diffusion (CED) were presented in [46], where EED preserves structures while filtering noise from homogeneous areas, and CED maintains small spherical structures while filtering tubular forms. In 2002, RPM and EED were evaluated on 3D angiography [49], suggesting that EED is more effective than RPM for high-resolution reconstructions. Recently, Mendrik et al. [50] have presented Hybrid Diffusion with Continuous Switch (HDCS), which exploits the properties of EED and CED by applying a continuous combination of these two filters to obtain suitable noise filtering of
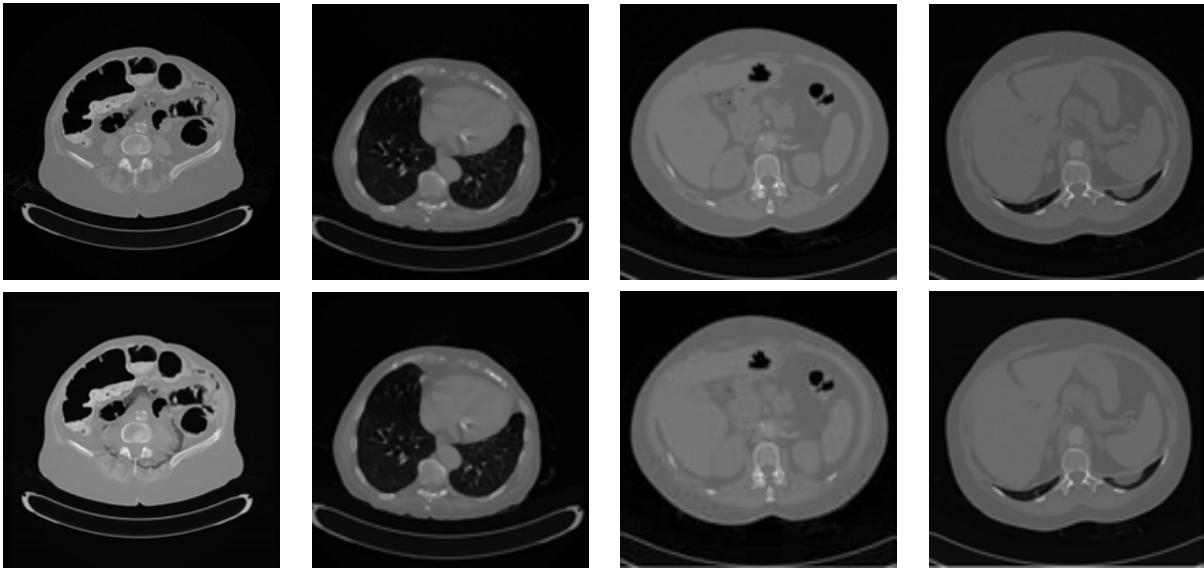
Figure 3.2: Different slices of CT images captured at low radiation dose: top row) original; bottom row) after noise filtering with HDCS. The reader is invited to zoom in to view the details of these images.

CT images; it is considered beneficial for improving detail visibility, surface smoothness, and overall impression of gated scans. Figure 3.2 depicts slices of different CT images, obtained at low radiation dose, containing significant noise. Also shown are the same slices after noise filtering using HDCS. It is worth noting that HDCS is one of the latest filters presented in the literature, improves over previous methods applied to CT images, and has been evaluated and approved by physicians. For these reasons, this filter is used in all discussions below.

3.2   Noise Filtering in Computed Tomography image coding

3.2.1   Influence of Noise Filtering in JPEG2000

In addition to improving image quality as discussed above, noise filtering also improves coding performance. The noise reduction results in higher correlation between

pixels, making the wavelet transform and subsequent entropy coding more efficient. The main objective of the wavelet transform is to concentrate information into a few transform coefficients. Consequently, the number of "small" coefficients (or, ideally, zero valued coefficients) should be as high as possible to improve the coding performance. Table 3.1 shows, for the first slice of several images, the number of zero coefficients after 5 levels of the 2D RWT is performed. Also reported is the bitrate obtained when JPEG2000 is used to compress the slices losslessly. These results are shown when the Noise Filter is not applied (JPEG2000) and when the Noise Filter is applied (NF+JPEG2000). Note that when NF is applied, the number of zero coefficients increases significantly, influencing the achievable lossless bitrate. Specifically, the lowest number of zero coefficients corresponds to the highest bitrate, and the highest number of zero coefficients corresponds to the lowest bitrate.

Table 3.1: Number of zero valued wavelet coefficients and bitrate for lossless compression of the first slice of several images using JPEG2000 and NF+JPEG2000.

Image	#Zeros		bitrate (bpp)	
	JPEG2000	NF+JPEG2000	JPEG2000	NF+JPEG2000
SS16-T1_D075_2	20,281	115,003	5.88	2.80
SS16-T2_D1_1	11,805	99,057	6.46	3.47
SS16-T2_D2_1	16,697	139,689	5.82	2.45
SS16-T7_D5_1	29,845	122,373	5.04	2.58

Furthermore, not only the performance of the wavelet transforms influences in the lossless bitrate, also in the progressive lossy-to-lossless performance. To evaluate it, the rate-distortion performance of JPEG2000 with and without NF is considered. Figure 3.3 shows the rate-distortion performance in SNR for JPEG2000 and NF+JPEG2000 for eight different images. The JPEG2000 (without NF) strategy compresses the original 3D image without employing a multi-component transform. Then, SNR results are then reported for the decoded image with respect to the original. However, the NF+JPEG2000 strategy compresses the noise filtered 3D image also without multi-component transform. In this case, SNR results then compare the decoded image with respect to the filtered image. Results suggest that the rate-distortion coding performance is improved when the NF stage is applied, obtaining

on average, more than 15 dB of benefit with respect to the no-NF strategy.

Summarizing, the noise filter stage reduces the image noise, improving the detail visibility of the image and providing a substantial improvement in coding performance.

3.2.2 Multi-component Transform Analysis

In this section, the effect of the noise filtering before different multi-component transforms are evaluated. Notation is as follows: NF stands for the application of the Noise Filtering, while no-NF means dealing with the original image. Four different multi-component transforms strategies are used: 1) no multi-component transform, only the 5/3 Reversible Wavelet Transform (RWT) is applied in the spatial domain (denoted as "2D-RWT"); 2) multi-component RWT followed by a 2D-RWT (denoted as "RWT+2D-RWT"); 3) the Reversible Karhunen-Loêve transform [51] (RKLT) followed by the 2D-RWT (denoted as "RKLT+2D-RWT"); and 4) the Reversible Haar transform (RHAAR) followed by the 2D-RWT (denoted as "RHAAR+2D-RWT"). For "RWT+2D-RWT" and "RHAAR+2D-RWT", the number of wavelet levels are applied taking into account the number of slices of each image. For the spatial "2D-RWT", 5 decomposition levels have been applied.

Table 3.2 shows the first-order entropy in bits per pixel (bpp) for the no-NF and the NF images using the four decorrelation strategies.

From these results we can see that applying Noise Filtering to the original images reduces first-order entropy only slightly. Even so, and for all investigated transforms, the application of the Noise Filtering can provide significant benefits. As for the decorrelation strategies, once the NF has been applied, RKLT provides a reduction in first-order entropy of about 28%, while the other strategies yield a first-order entropy reduction close to 60%. Due to the differences in entropy reduction between RKLT, RWT and RHAAR transforms, RKLT is not a valid multi-component transform to be applied.

Table 3.2: Entropy in bpp of the images using different decorrelation strategies: 2D-RWT, RWT+2D-RWT, RKLT+2D-RWT, and RHAAR+2D-RWT.

Image		Original	2D-RWT	RWT + 2D-RWT	RKLT + 2D-RWT	RHAAR + 2D-RWT
SS16-T1-D075_1	no-NF	8.68	4.50	4.47	4.77	4.34
	NF	8.12	2.80	2.66	3.71	2.62
SS16-T1-D075_2	no-NF	8.76	4.38	4.39	4.67	4.24
	NF	8.13	2.41	2.27	3.47	2.22
SS16-T1-D075_3	no-NF	9.10	4.39	4.33	4.67	4.20
	NF	8.42	2.32	2.39	3.64	2.31
SS16-T1-D075_4	no-NF	9.45	5.59	5.58	5.79	5.45
	NF	8.54	2.58	2.67	3.95	2.59
SS16-T2-D1_2	no-NF	8.97	4.45	4.38	4.66	4.27
	NF	8.24	2.18	2.08	3.25	2.02
SS16-T2-D1_3	no-NF	8.49	4.46	4.39	4.66	4.28
	NF	7.47	2.20	2.11	3.23	2.04
SS16-T2-D2_1	no-NF	9.17	2.34	4.13	4.42	3.95
	NF	8.81	2.34	2.45	3.47	2.30
SS16-T2-D2_3	no-NF	8.49	4.30	4.13	4.51	4.29
	NF	7.49	1.91	2.03	2.89	1.90
SS16-T7-D5_3	no-NF	8.55	3.94	3.93	4.31	3.74
	NF	8.07	2.26	2.40	3.36	2.22
SS16-T7-D5_4	no-NF	8.55	4.00	3.98	4.32	3.81
	NF	8.14	2.30	2.44	3.35	2.28

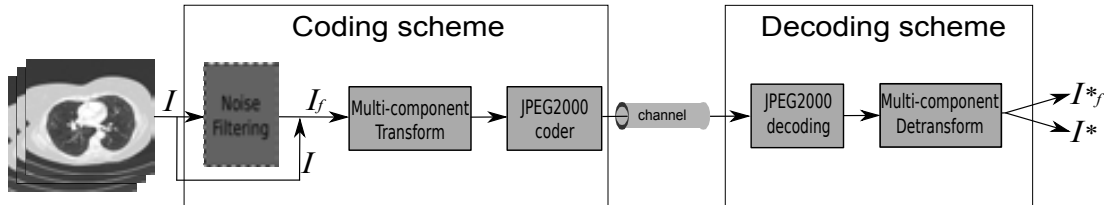### 3.2.3   Noise-Filtering Multi-Component JPEG2000

Figure 3.4: Proposed Noise-Filtering Multi-Component JPEG2000 coder.

Considering the entropy reduction increase of the decorrelation strategies after a Noise Filter, "Noise-Filtering Multi-Component JPEG2000" coding scheme is presented. The coding scheme, depicted in Figure 3.4, has three main stages: the Noise Filter

(NF) stage, which is depicted in dark gray, the multi-component transform stage, and the JPEG2000 coder stage, both depicted in light gray.

The NF stage is responsible for the application of the noise filter. The NF stage takes one original CT input image, I, and produces an output image denoted by I_f. The noise in this stage is discarded because it does not provide relevant information to the medical diagnosis [50]. The Multi-component transform stage is responsible for applying RWT or RHAAR over slices. The JPEG2000 stage generates a compliant JPEG2000 code-stream. The coder stage is commonly constituted by four coding sub-stages: sample data transformation, sample data coding, rate-distortion optimization, and code-stream reorganization. In our proposal scheme, the application of noise filtering is optional: when it is applied, the decoded image is refereed as I_f^*; otherwise, it is denoted as I^*.

3.3 Experimental results

In this section, the coding performance of the proposed coding scheme is evaluated. Three different sets of experimental results are provided: A) perceptual results, intended to analyze the perceptual similarity of the decoded denoised image (I_f^*) at specific bitrates with respect to the original image (I); B) rate-distortion results, to calculate the coding gain in a progressive lossy-to-lossless scenario; and, C) lossless results, to analyze the size of the final code-streams.

All the experiments have been performed with the Kakadu [52] software where the 2-D coding parameters are: 5/3 Reversible WT kernel with 5 decomposition levels, code-block size 64x64. In the third dimension, different transforms such as RWT, and RHAAR have been evaluated. For the RWT and the RHAAR transform, the number of wavelet levels are applied taking into account the number of slices of each image. The rate allocation strategy used in all the experiments is the multi-component PCRD [53].

3.3.1   Perceptual Results

The SSIM metric has been computed using the SSIM implementation provided by its authors [43]. In all the experiments, the SSIM metric is obtained comparing I with I^* and I_f^* at different bitrates. Figure 3.5 depicts the rate-similarity for different images. Results are reported for JPEG2000, NF+JPEG2000, RWT+JPEG2000, NF+RWT+JPEG2000, RHAAR+JPEG2000, and NF+RHAAR+JPEG2000. Results indicate that the NF strategy obtains very competitive results in terms of structural similarity, taking values of SSIM around 0.9998. These results are obtained because the HDCS filter is very competitive in preserving structures and noise filtering in homogeneous areas.

3.3.2   Rate-Distortion Results

These experiments evaluate the use of the NF strategy in terms of rate-distortion performance. The images are encoded at different target bitrates, decoded, and the image quality is assessed.

The experiments exhibit the effect of the NF strategy and the no-NF strategy, for different decorrelation techniques applied in the third dimension. Figure 3.6 shows the rate-distortion performance for JPEG2000, RWT+JPEG2000 and RHAAR+JPEG2000 when the no-NF and NF strategy are used. Results suggest that coding performance is improved for all decorrelation strategies when the NF stage is applied. The best performance is achieved for NF+RWT+JPEG2000, which yields an improvement of more than 10 dB at 1.0 bpp as compared to RWT+JPEG2000, the best approach when noise filtering is not applied.

Figure 3.7 shows the rate-distortion between NF+JPEG2000 and NF+RWT+JPEG2000. Results indicate that for most of the cases the performance of the NF+RWT+JPEG2000 outperforms the NF+JPEG2000, but in some cases the performance of NF+RWT+JPEG2000 are below the NF+JPEG2000 strategy. In the next chapter the theoretical analysis of this phenomenon is explained.

3.3.3   Lossless Results

Table 3.3 shows the code-stream size in bpp, for all decorrelation strategies evaluated: JPEG2000, RWT+JPEG2000, and RHAAR+JPEG2000, with and without the NF stage. Results suggest that NF provides a considerable benefit in compression ratio for all the images. Results suggest that for all decorrelation strategies with NF have better coding performance than no-NF approaches. In addition, RWT in the third dimension yield better coding performance than RHAAR for all the images evaluated.

SS16-T1-D075_1   SS16-T7-D5_2

SS16-T2-D1_2   SS16-T2-D1_4

SS16-T2-D2_2   SS16-T2-D2_3

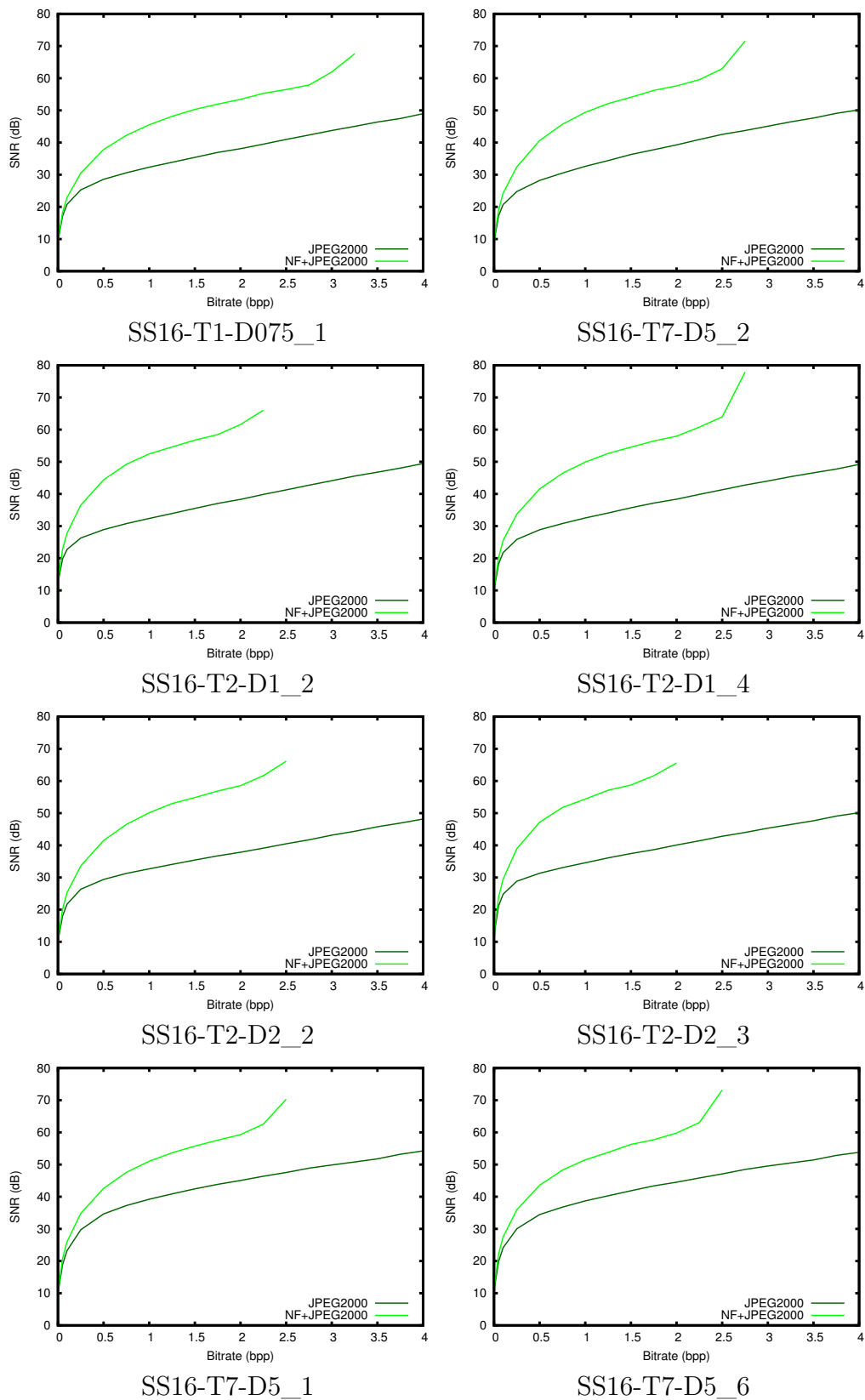SS16-T7-D5_1   SS16-T7-D5_6
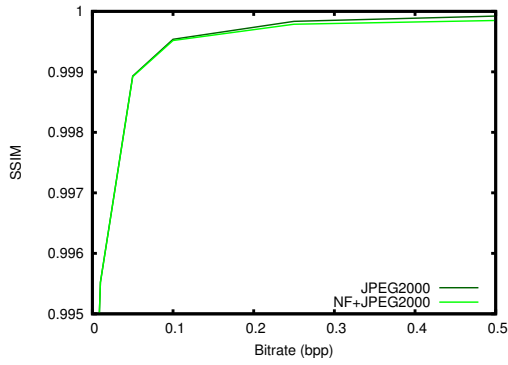
Figure 3.3: Rate-distortion performance of JPEG2000 and NF+JPEG2000 for different images.
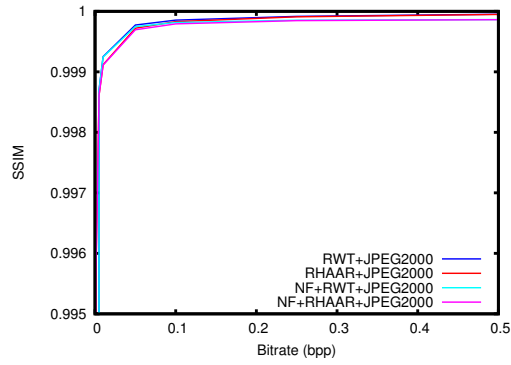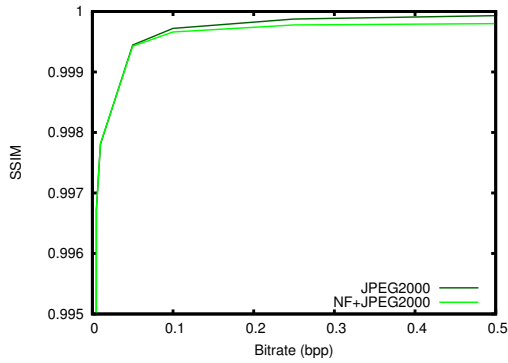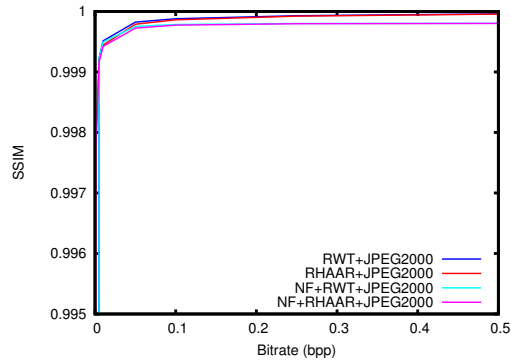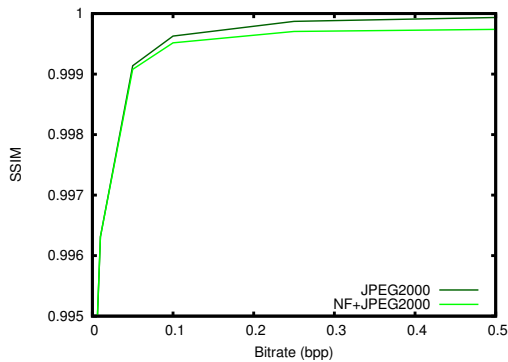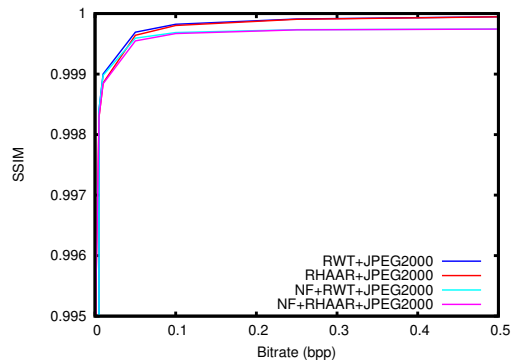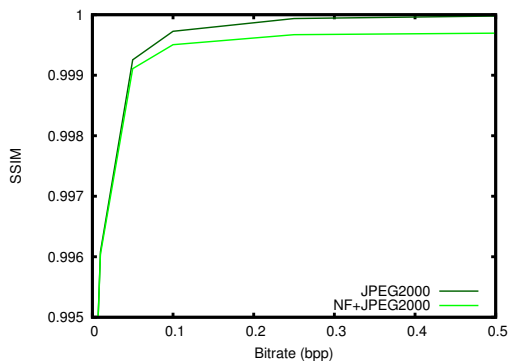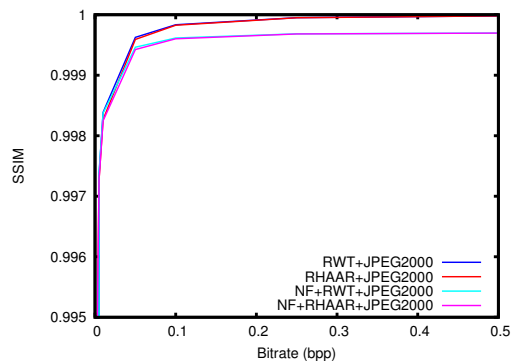
Figure 3.5: Rate-similarity for NF and no-NF strategy. a) NF and no-NF strategy for JPEG2000. (b) NF and no-NF strategy for RWT and RHAAR.

Figure 3.6: Rate-Distortion performance for no-NF strategy and NF strategy. For both strategies, results reported RWT+JPEG2000 and RHAAR+JPEG2000.

SS16-T1-D075_1

SS16-T7-D5_2

SS16-T2-D1_2

SS16-T2-D1_4

SS16-T2-D2_2
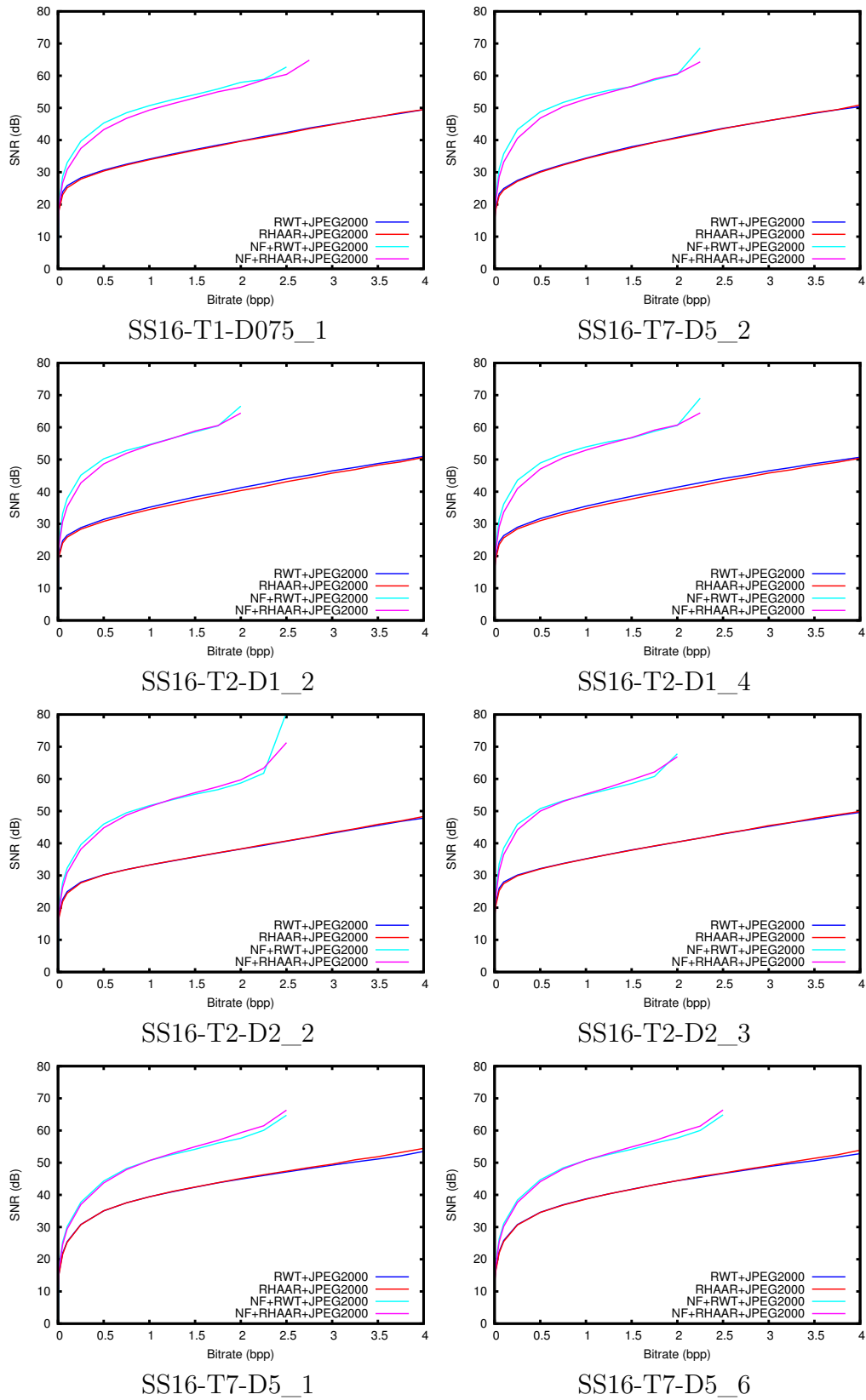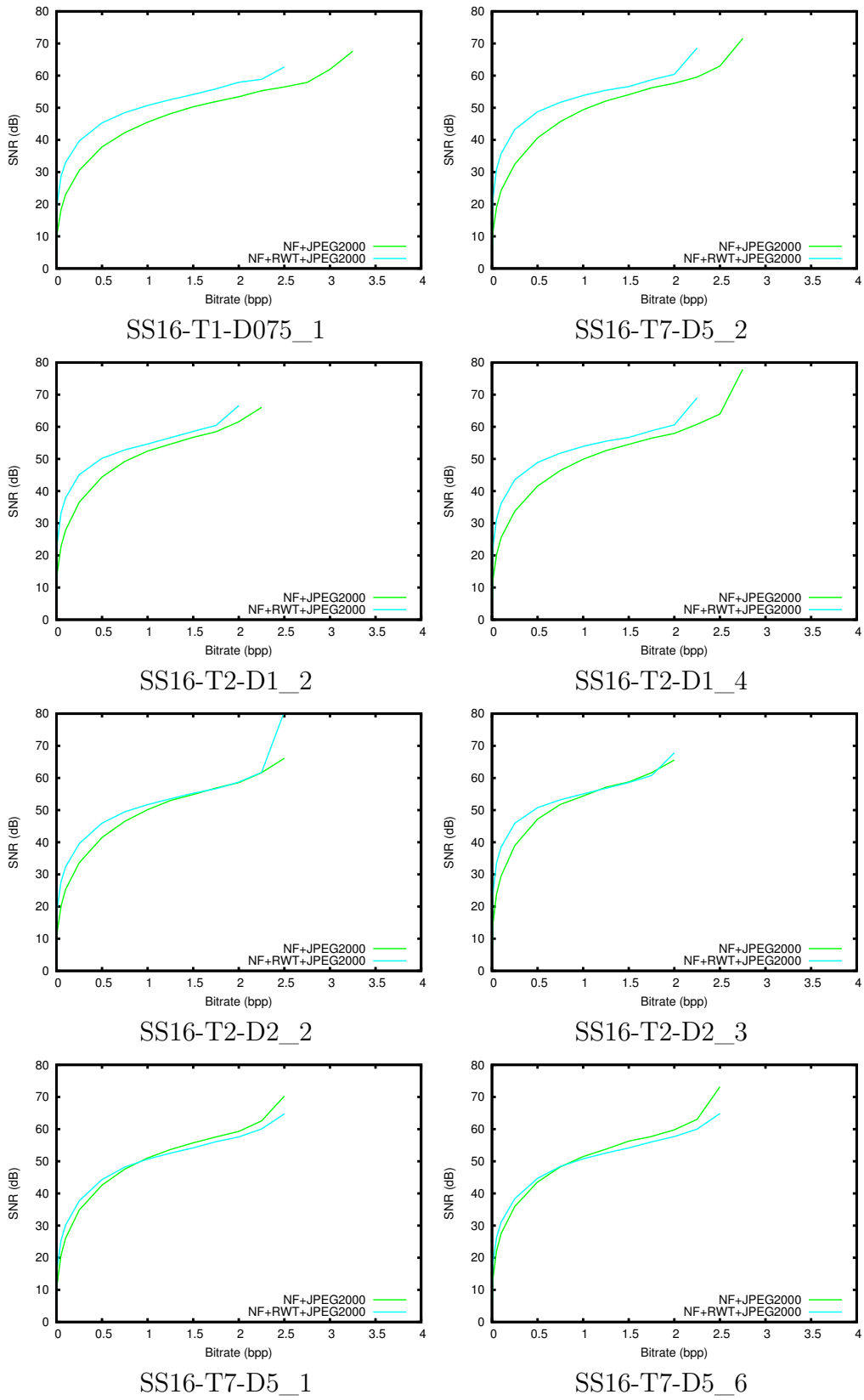
SS16-T2-D2_3

SS16-T7-D5_1

SS16-T7-D5_6

Figure 3.7: Rate-Distortion performance comparison between NF+JPEG2000 and NF+RWT+JPEG2000 for different images

Table 3.3: Lossless coding performance (in bpp). From column 3 to 5, the evaluated transforms are depicted: JPEG2000, RWT+JPEG2000, and RHAAR+JPEG2000.

Image Name	Strategy	JPEG2000	RWT + JPEG2000	RHAAR + JPEG2000
SS16-T1_D075_1	no-NF	6.20	5.80	5.91
	NF	3.37	**2.68**	2.92
SS16-T1_D075_2	no-NF	5.88	5.51	5.61
	NF	2.80	**2.28**	2.46
SS16-T1_D075_3	no-NF	5.86	5.45	5.61
	NF	2.77	**2.40**	2.56
SS16-T1_D075_4	no-NF	8.37	8.03	8.16
	NF	3.16	**2.73**	2.94
SS16-T1_D075_5	no-NF	5.22	4.89	5.01
	NF	2.62	**2.32**	2.45
SS16-T1_D075_6	no-NF	6.04	5.63	5.74
	NF	2.78	**2.24**	2.39
SS16-T1_D075_7	no-NF	7.45	6.95	7.08
	NF	2.63	**2.28**	2.40
Average	no-NF	6.43	6.04	6.16
	NF	2.87	**2.42**	2.58
SS16-T2_D1_1	no-NF	5.04	4.72	4.81
	NF	2.08	**1.82**	1.87
SS16-T2_D1_2	no-NF	5.99	5.38	5.63
	NF	2.43	**2.1**	2.21
SS16-T2_D1_3	no-NF	6.01	5.40	5.65
	NF	2.45	**2.12**	2.23
SS16-T2_D1_4	no-NF	6.22	5.55	5.81
	NF	2.76	**2.31**	2.45
SS16-T2_D1_5	no-NF	6.18	5.51	5.77
	NF	2.79	**2.33**	2.48
Average	no-NF	5.88	5.31	5.77
	NF	2.50	**2.13**	2.24
SS16-T2_D2_1	no-NF	5.38	5.26	5.29
	NF	2.60	**2.56**	2.59
SS16-T2_D2_2	no-NF	6.34	6.19	6.22
	NF	2.68	**2.50**	2.55
SS16-T2_D2_3	no-NF	5.91	5.75	5.79
	NF	2.21	**2.08**	2.21
Average	no-NF	5.88	5.73	5.76
	NF	2.49	**2.38**	2.45
SS16-T7_D5_1	no-NF	5.04	4.91	4.94
	NF	**2.58**	2.59	2.58
SS16-T7_D5_2	no-NF	5.11	4.95	4.99
	NF	2.50	**2.47**	2.49
SS16-T7_D5_3	no-NF	5.04	4.91	4.99
	NF	2.46	**2.44**	2.44
SS16-T7_D5_4	no-NF	5.16	5.01	5.06
	NF	2.51	**2.48**	2.50
SS16-T7_D5_5	no-NF	5.24	5.09	5.12
	NF	2.55	**2.51**	2.53
SS16-T7_D5_6	no-NF	5.11	5.00	5.03
	NF	**2.54**	2.58	2.58
SS16-T7_D5_7	no-NF	5.21	5.07	5.11
	NF	2.57	**2.56**	2.57
Average	no-NF	5.13	4.99	5.03
	NF	2.53	**2.52**	2.53

Chapter 4

Correlation Modeling for Compression of Computed Tomography Images

In this chapter it is presented an study of the performance of the multi-component transforms taking into account the correlation among the slices of the images. In addition, it is introduced a model specifically designed for CT images which takes into account the acquisition parameters to model the correlation among slices.

4.1 Introduction

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 [54]. 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
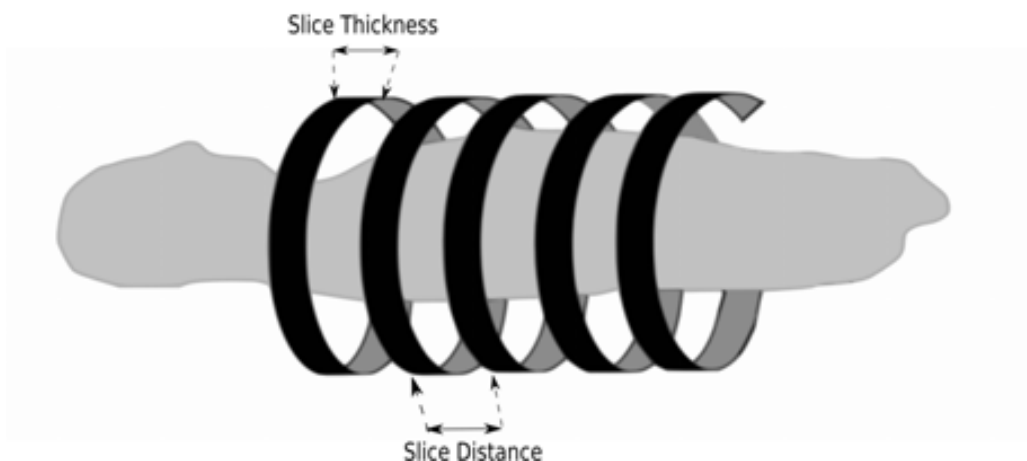
Figure 4.1: Graphical representation of slice thickness and slice distance during a CT scan.

mm. It is possible to choose the slice distance to be less than the slice thickness. Figure 4.1 provides a graphical representation of these two concepts.

Regarding CT image acquisition parameters, Siegel et al. [55] 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 [56].

4.2 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 [56]. 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

$$r_{A,B} = \frac{E[(A - \bar{A})(B - \bar{B})]}{\sigma_A \sigma_B}, \tag{4.1}$$

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

$$r'_{k,k+1} = \frac{1}{N_y N_x} \sum_{j=1}^{N_y} \sum_{i=1}^{N_x} \frac{(I_{kji} - \bar{I}_k)(I_{k+1ji} - \bar{I}_{k+1})}{\sigma_k \sigma_{k+1}}, \tag{4.2}$$

where I_{kji} denotes the pixel at column i and row j of slice k, and $\bar{I}_k$ 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

$$r' = \frac{1}{N_z - 1} \sum_{k=1}^{N_z - 1} r'_{k,k+1}. \tag{4.3}$$

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 4.2 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 bitrate obtained by JPEG2000 (without multi-component transform) minus the bitrate obtained by JPEG2000 with a multi-component transform (RWT+JPEG2000 or RHAAR+JPEG2000).

RWT+JPEG2000

RHAAR+JPEG2000
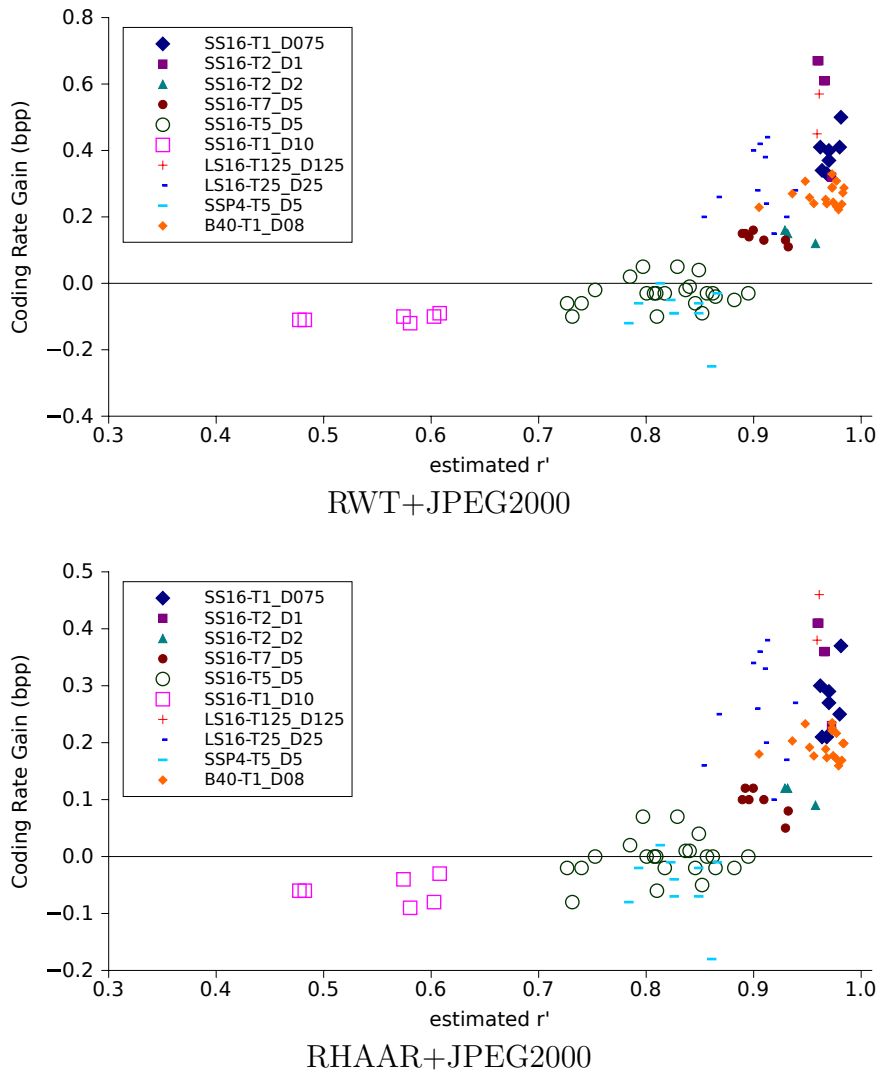
Figure 4.2: Coding rate gain vs estimated r'. (a) and (b) respectively depict the coding rate gain for RWT and RHAAR.

These differences are referred as the coding rate gain, where positive values indicate improvement for multi-component transforms, and the negative values indicate that multi-component transforms have lower coding performance than JPEG2000. For the images used in this manuscript, the results of Figure 4.2 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.

4.3 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, a theoretical model for the correlation among slices is proposed, 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 ji of slice k was denoted by I_{kji}. In what follows, we consider a sequence of pixels, indexed by k, obtained by fixing a spatial location ji. To reduce notational clutter, we drop the explicit dependence on ji 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, y is discretized 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 4.3 depicts this example for $D = 0.75$ mm, 1 mm and 1.5 mm, resulting in $M = 12$, 16 and 24, respectively.

The kth pixel value $x(k)$ can then be written as

$$x(k) = \sum_{l=1}^{L} a(l)y(kM - l). \tag{4.4}$$
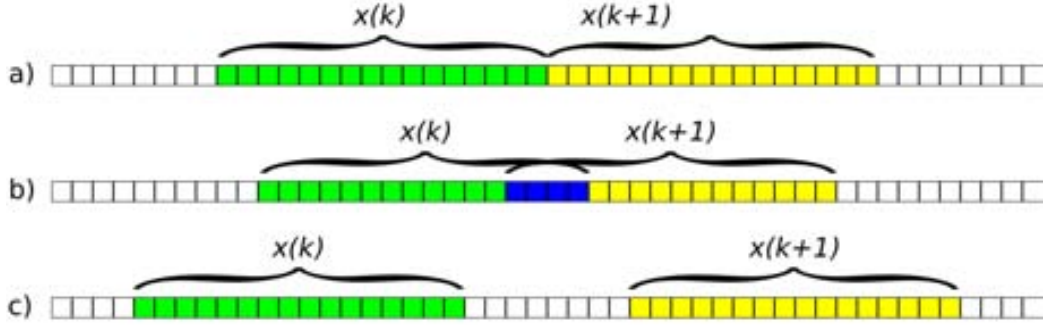
Figure 4.3: Samples of y used to compute $x(k)$ for three choices of D. In each case, $T = 1$ mm, a) $D = 1$ mm, b) $D = 0.75$ mm, and c) $D = 1.5$ mm.

The constants $a(l)$ are included in (4.4) for two reasons. First, they allow for the possibility of generalizing the expression to a weighted sum. Second, they facilitate the observation that (4.4) corresponds to a filtering (or convolution) operation. Specifically, $x(k)$ is a subsampled version of

$$w(n) = a(n) * y(n). \tag{4.5}$$

That is,

$$x(k) = w(kM), \tag{4.6}$$

where

$$w(n) = \sum_{l=1}^{L} a(l)y(n - l). \tag{4.7}$$

Now, the assumption that the samples $y(n)$ arise from a simple auto-regressive random process is taken

$$Y(n) = bY(n - 1) + \Theta(n), \tag{4.8}$$

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

$$C_Y(j) = E[(Y(n) - \bar{Y})(Y(n+j) - \bar{Y})], \tag{4.9}$$

where $\bar{Y} = E[Y(n)] = E[Y(n+j)]$, regardless of n. It is then easily shown that for the specific choice of (4.8),

$$C_Y(j) = \sigma_Y^2 b^{|j|}. \tag{4.10}$$

From (4.5), it follows that the autocovariance function of $W(n)$ is

$$C_W(j) = C_Y(j) * a(j) * a(-j). \tag{4.11}$$

From (4.6), we then have

$$C_X(k) = C_W(kM). \tag{4.12}$$

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 (4.10), (4.11), and (4.12). Consistent with this 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

$$r = \frac{E[(X(k) - \bar{X})(X(k+1) - \bar{X})]}{\sigma_X^2} = \frac{C_X(1)}{\sigma_X^2}. \tag{4.13}$$

Different 24 images from the corpus of Table B.4 are used to find a suitable value of b by minimizing the least squared error between r as computed by (4.13) and r' as computed via (4.3). 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 4.1 are provided to assess the performance of our model. In particular, each row of Table 4.1 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 (4.13) 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 (4.3). The average

Table 4.1: 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.

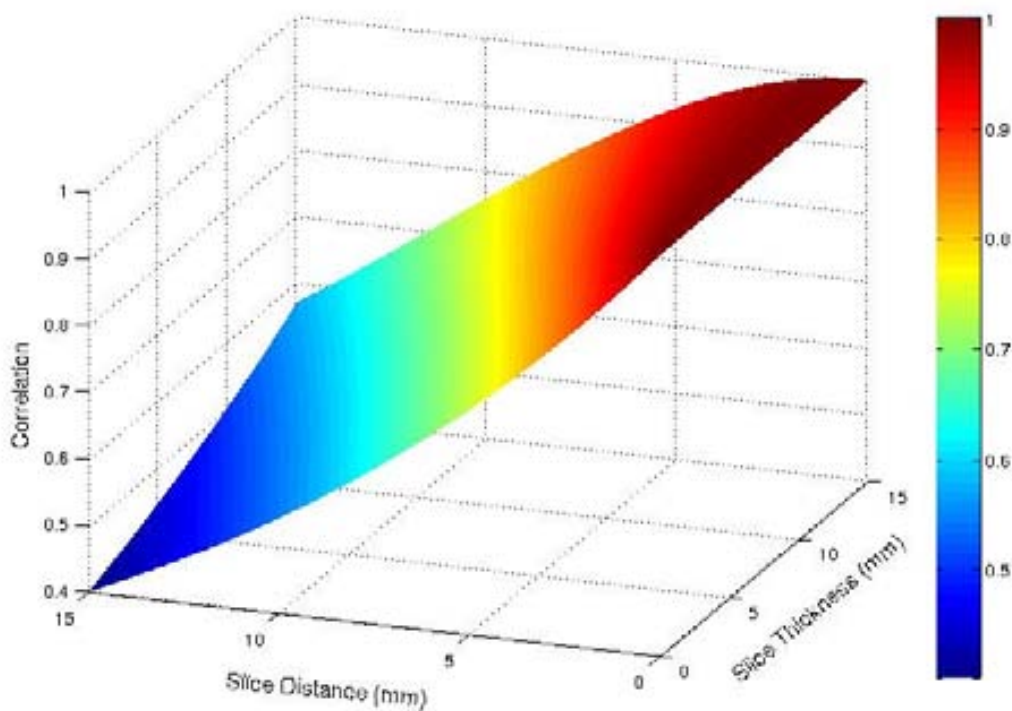| Images	r	$\bar{r}'$	Mean Error	Std. Deviation
SS16-T1-D075	0.979	0.971	0.0083	0.006
SS16-T2-D1	0.978	0.964	0.0133	0.0052
SS16-T2-D2	0.926	0.946	0.0132	0.0157
SS16-T7-D5	0.871	0.907	0.0359	0.0176
SS16-T5-D5	0.828	0.818	0.0103	0.0295
SS16-T1-D10	0.554	0.554	0.0001	0.0232
LS16-T125-D125	0.954	0.955	0.0007	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 4.4: Modeled r.

of these values is reported as $\bar{r}'$ in Table 4.1. Finally, the mean error and variance of the error between r and r' is reported for each image set. Note that the 76 images used to obtain Table 4.1 are from the corpus of Table B.4 but are different from the 24 images used to calculate b. As can be seen in Table 4.1, the modeled values for r agree closely with the estimated values r'. Figure 4.4 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. Figure shows that the correlation decreases when the 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 (4.3) 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.

4.4 Experimental Results

Extensive experiments have been carried out to evaluate our correlation model. In particular, we have carried out: 1) a lossless coding performance evaluation, 2) a rate-distortion evaluation, and 3) a component scalability evaluation. 1) and 2) aim to analyze the compression performance of multi-component transforms (RWT and RHAAR) on images with different acquisition parameters, and 3) assesses the rate-distortion performance when a subset of components are decoded from a code-stream.

4.4.1 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 4.5 is equivalent to Figure 4.2 but depicts r rather than r'. The same conclusion is apparent: performing a multi-component

RWT+JPEG2000

RHAAR+JPEG2000

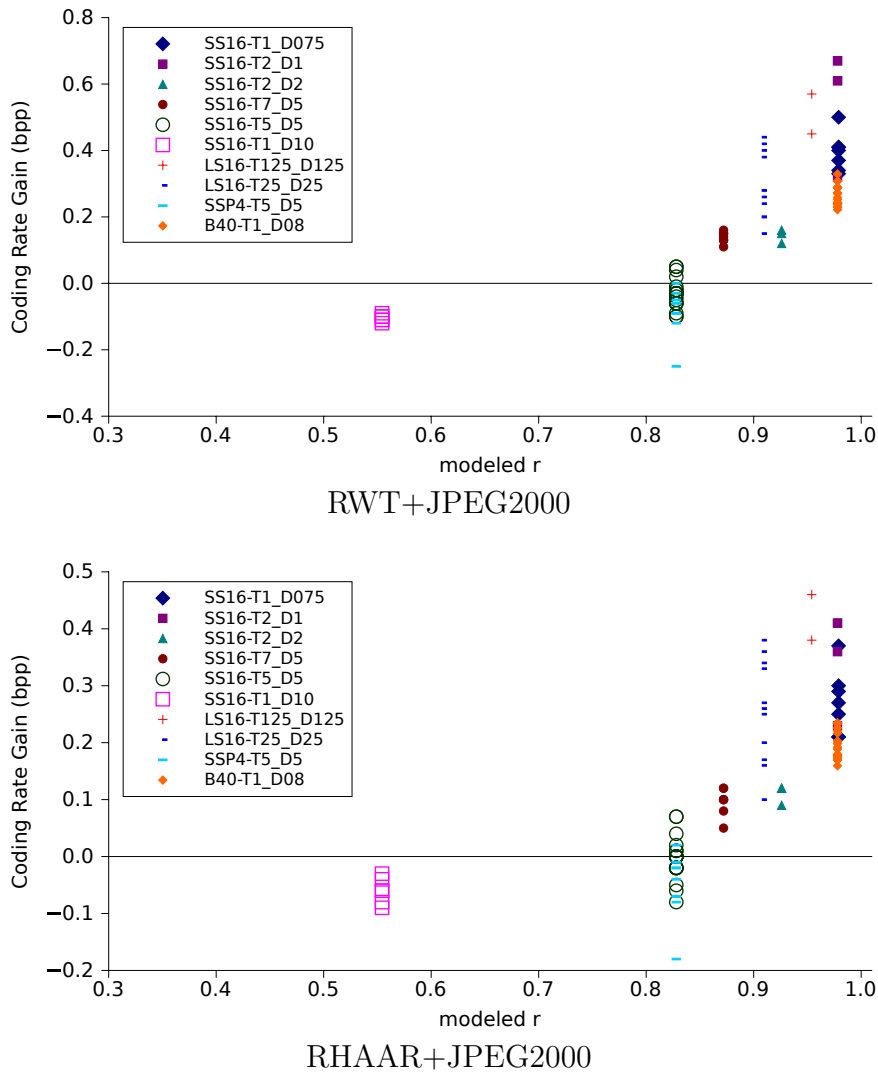Figure 4.5: Coding rate gain vs modeled r. (a) and (b) respectively depict the coding gain between RWT+JPEG2000 and RHAAR+JPEG2000, with respect to JPEG2000.

transform is profitable when r exceeds 0.87.

4.4.2 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. Figure 4.6 shows the

rate-distortion performance in terms of SNR for eight images from 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.

4.4.3 Component Scalability

Component scalability is negatively impacted when multi-component transforms are employed. To explore this effect, a subset of N slices of interest are considered to decode. 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, the number of bytes needed to decode a set of consecutive slices from the center of an image are analyzed.

Figure 4.7 shows the amount of data decoded (in MB) for the three tested coding approaches as a function of the number of slices decoded N. On the one hand, when $r > 0.87$, 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. On the other hand, when $r < 0.87$, the coding performance of JPEG2000 is always better than strategies using RHAAR or RWT.

LS16-T25-D25_2 ($r = 0.908$)

LS16-T125-D125_2 ($r = 0.954$)

SS16-T075-D1_1 ($r = 0.979$)

SS16-T2-D1_2 ($r = 0.978$)

SS16-T2-D2_2 ($r = 0.926$)

SP4-T5-D5_6 ($r = 0.828$)

SS16-T5-D5_8 ($r = 0.828$)
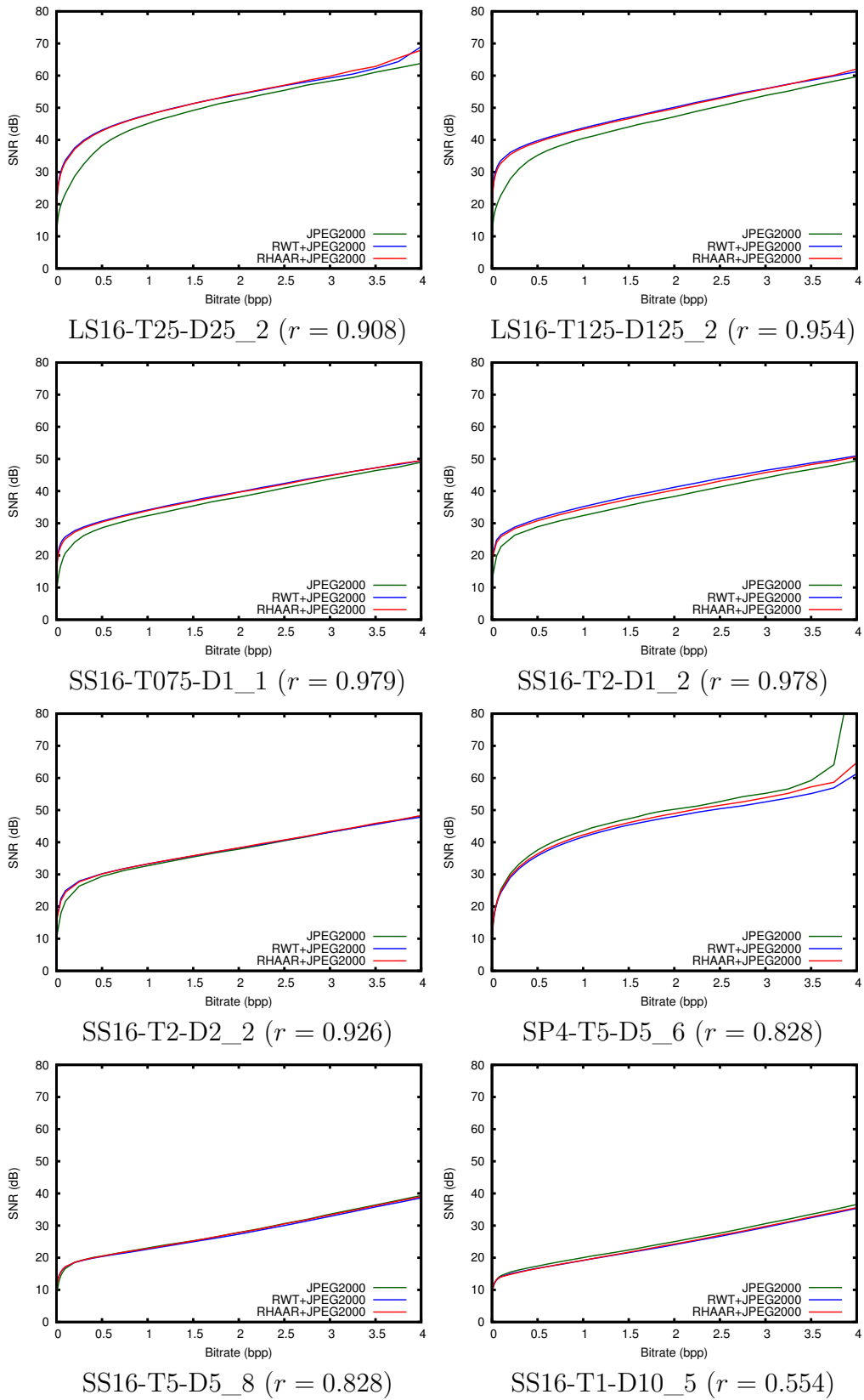
SS16-T1-D10_5 ($r = 0.554$)

Figure 4.6: Rate-distortion performance for RWT+JPEG2000, RHAAR+JPEG2000, and JPEG2000.
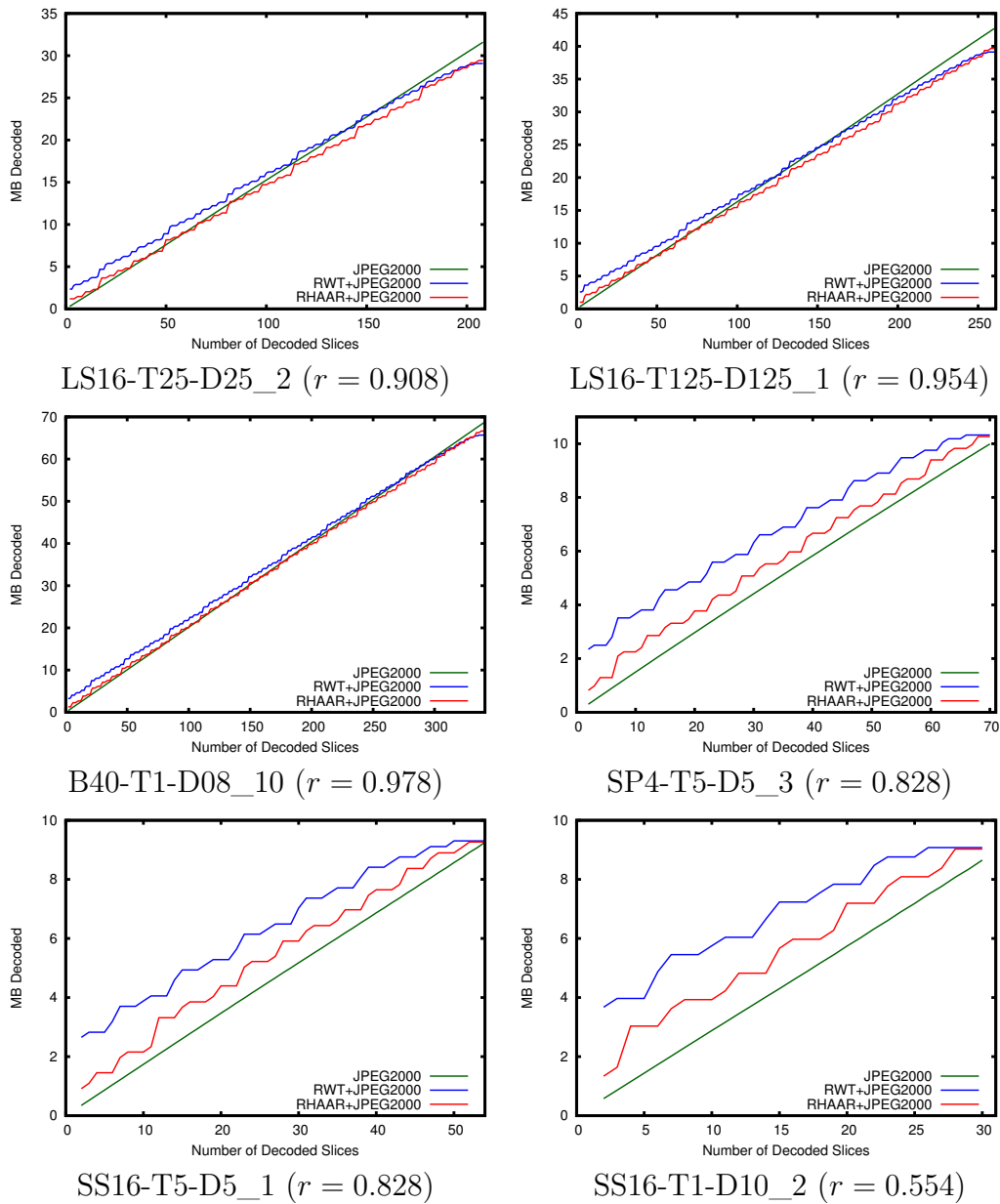
Figure 4.7:   Data decoded (in MB) for JPEG2000, RWT+JPEG2000 and
RHAAR+JPEG2000 as a function of N, where N is the number of slices decoded
from the center of the image.

Chapter 5

CT Segmentation Coding

This chapter presents a efficient thresholding-based segmentation method that removes the non-relevant information of the image and increases the efficiency of the coding system.

5.1   Introduction

In a CT image, the intensity of each pixel represents the attenuation value of the X-ray emitted through the body. To facilitate the interpretation of the specialist, the intensity values are given in Hounsfield units (HU). However, the images are not stored in these units, and they need to be converted to HU through the following linear transform:

$$I_{k,j,i} = m \cdot s_{k,j,i} + b,$$

where m is the slope, b the intercept, and $s_{k,j,i}$ denotes the intensity value of the pixel in the 3D image. The spatial position of the pixel in the 3D image is denoted with the subindex k, j, i, where k represents the slice, j the row, and i the column. The values b and m are set by the manufacturer depending on the characteristics of the scanning device [57]. Recalling from the Introduction of this thesis, the HU permits to identify

biological substances according different ranges. For example values around -1000, in HU, represent air and values around 3000 correspond high density bones.

Commonly, in medical diagnosis through CT images the physician specifies the range of HU values, to identify determined tissues or biological substances, to be displayed. This range is specified with two parameters, window level and window width. The window level represents the center of the visualized range, whereas the window width covers the range of intensities to display. Figure 5.1 shows an area of a CT image with two different configurations of window levels and ranges. Note that depending on the window configuration used, the noise is visually appreciable on different substances. For example, in Figure 5.1 (a) the noise artifacts are appreciable in those air pixels that are located out of the body area, which penalize significantly the coding performance [58] and are considered non-relevant data for medical purposes. This affirmation is supported by the radiology team of the Hospital Mutua de Terrassa [59].

As it is explained in the Chapter 2, the wavelet transforms are employed in many state-of-the-art image coding schemes, including JPEG2000. The proposed segmentation method is applied in the LL subband of each slice to take the advantage of the reduced size of the image and then it reduces the computational cost of the segmentation methods applied in the original image. The proposed method is based on thresholding to take advantage of the HU scale. The main insight behind this contribution is to exploit the multi-resolution structure of the wavelet transform to segment the biological area of a CT image in the LL subband *during* the coding process. Although, wavelet transforms are used in several coding systems, the proposed method is implemented in the JPEG2000 framework due to its inclusion in DICOM and its widespread used in medical scenarios.

5.1.1   Review of segmentation techniques in medical imaging

In the literature, researchers presents different diagnostically lossless coding methods which includes different segmentation algorithms to detect the biological areas of the images and then erase the non-relevant areas [60, 61, 62, 63, 40, 64, 41]. Although
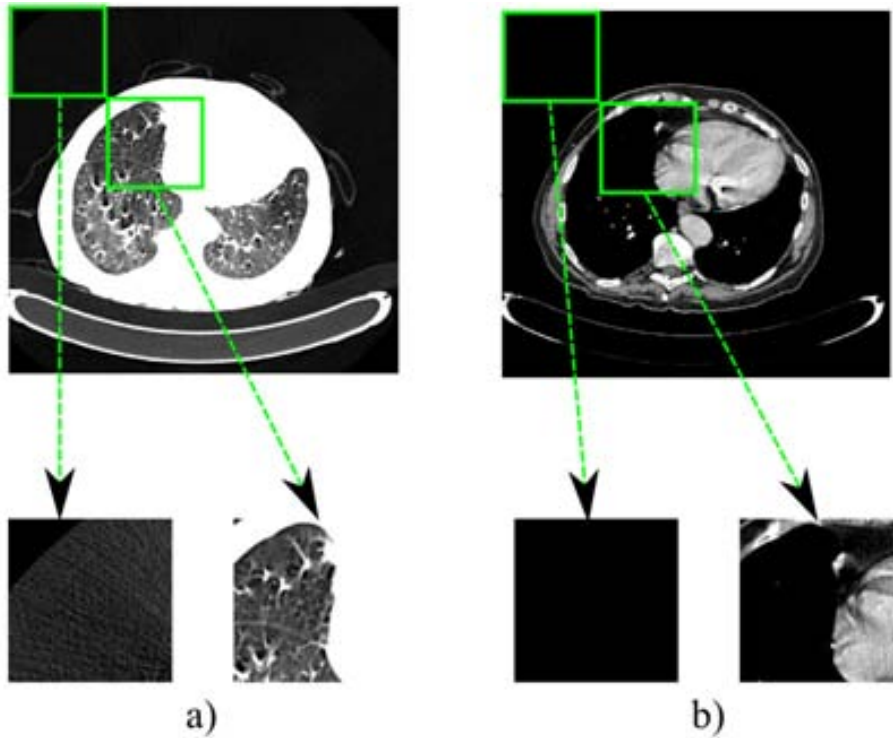
Figure 5.1: Slice of a CT image with two zoomed-in areas, corresponding to areas that contain air and biological material. (a) employs a window level of -812, and a window width of 424, whereas (b) employs a window level of 40, and a window width of 350.

these methods achieve competitive results in lossless coding performance (improvements around 2 bpp), they present different drawbacks. Some of the methods used in these coding schemes are designed for specific substances or images [65, 64, 41], or are focused on to provide substance scalabilities in the codestream [63] are computational expensive [61] because of the use of morphological operations, or use many computational resources due to the image dimensions [40].

5.2 Segmentation Method

Typically, in CT images there are three difference areas: the human body, the stretcher, and the air. The HU values of the stretcher lie on the range of the relevant information, which causes the impossibility of detecting the stretcher from the
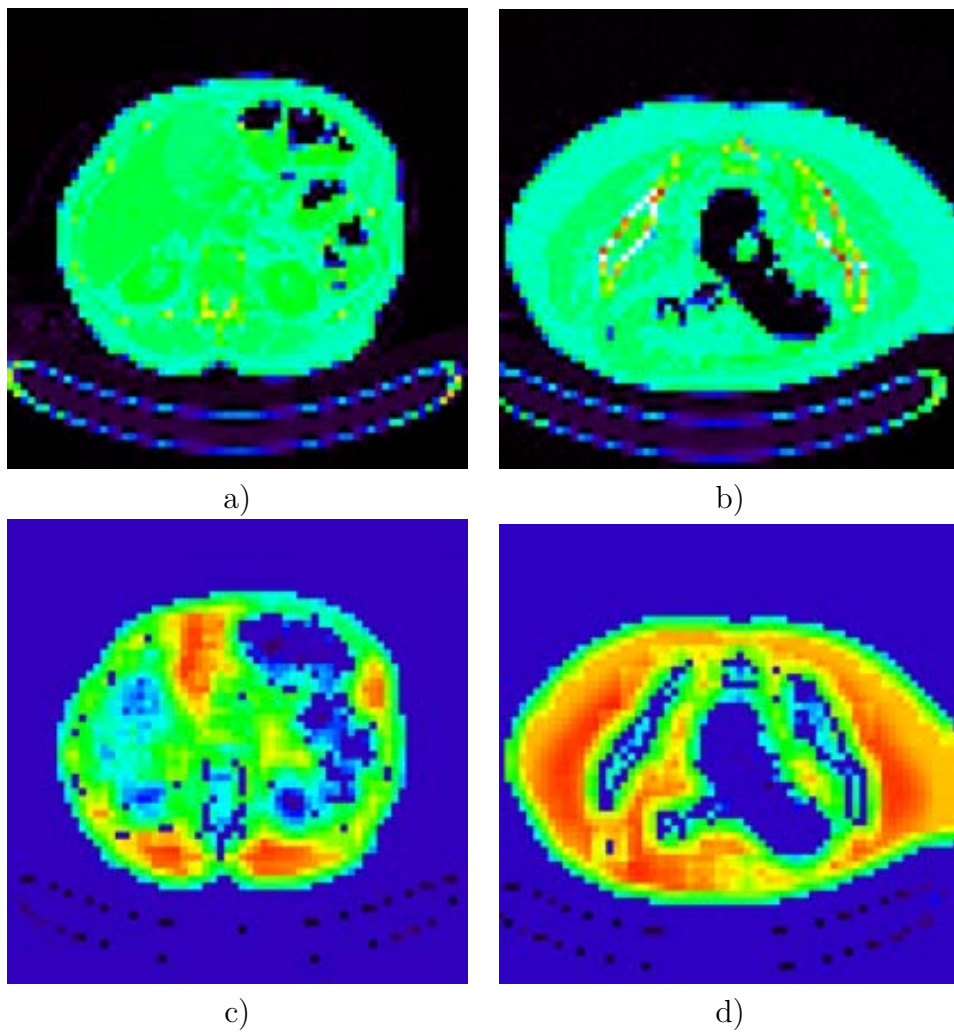
Figure 5.2: Two examples of LL subband (a and b) and its I_g (c and d).

CT image using segmentation methods based on thresholding. Figure 5.2 a) and b) depicts two examples of a LL subband of two different slices of a different CT images. Note that, there are an important amount of pixels with the same intensity value that corresponds to the stretcher and the human body.

To make the value ranges of the stretcher and the body are less overlapped a new configuration of window level and range has been set up the corpora. To set up correctly this configuration, the histogram of those samples belonging to the body
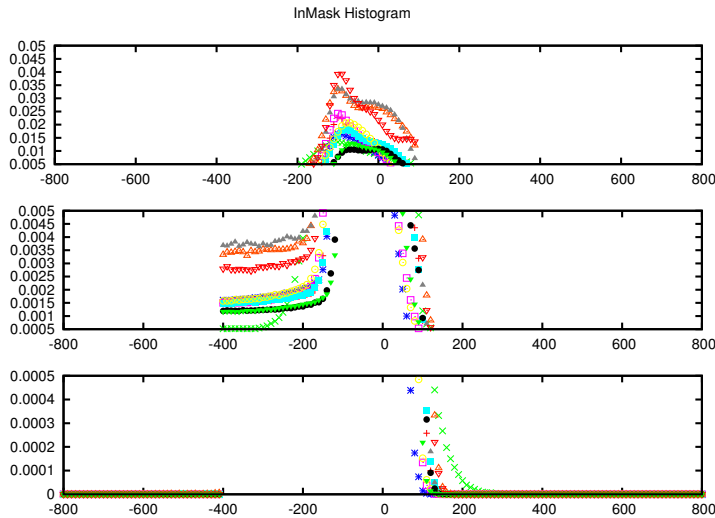
Figure 5.3: Histogram (percentage of apparition) of the contour of the human body of the original images.

contour with a thick of 15 pixels is computed using the algorithm presented in [40]. Figure 5.3 depicts the histogram of the body contour from five 3D images of different sensors and acquisition parameters. From this picture it is set up a new window level and range configuration R experimentally for our corpora as $R = [-400, 100]$. It is important to remark that, although there are pixels inside the human body not belonging to R, they are not needed to perform the correct segmentation. Once it is set the range in the original domain, the next step is to adjust this range to the wavelet domain. Due to the energy gain factor [9] of the 5/3 Wavelet Transform, the new range in the LL is $R_{LL} = [-500, 200]$. Figure 5.4 depicts the histogram of the body contour in the LL subband for the same five 3D images.

Once the R_{LL} configuration is used, there are still few LL coefficients that using a thresholding technique is not possible to differentiate correctly. An example of this effect is shown in Figure 5.3 (c) and (d), where purple data identifies stretcher and body region.

Given the property that in a natural image the pixel are highly correlated, it is very common the use of different filtering techniques to detect edges. Using a smoothing filter, it is possible to sparse the data, making it more distinguishable
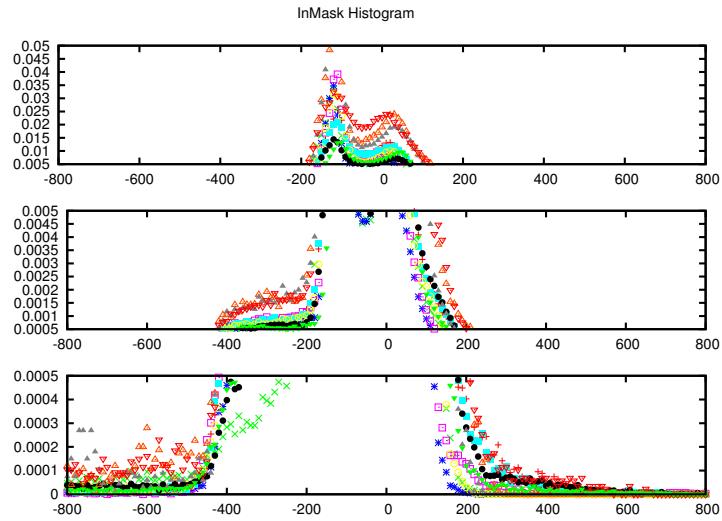
Figure 5.4: Histogram (percentage of apparition) of the contour of the human body of the LL subband of the images.
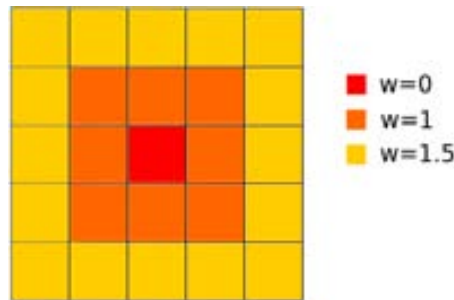
Figure 5.5: 5×5 kernel for smoothing filter.

through thresholding. A 5×5 (K_{5x5}) kernel is defined following two specifications: 1) Not only the first level of adjacency should be evaluated, also a second level because of the LL coefficients are highly correlated; and 2) w_0 is set to 0 because no matter intensity coefficient evaluated, due to this filter evaluates the neighbourhood. w_1 and w_2 have weights 1 and 1.5 because as far from the center coefficient less correlation have the coefficients among themselves. Figure 1.5 depicts graphically the 5x5 kernel defined. To compute the new image I_g, K_{5x5} is applied to all pixels that lies in the R_{LL} range. If the evaluated pixels does not belong to R_{LL}, the pixel is set to -2000
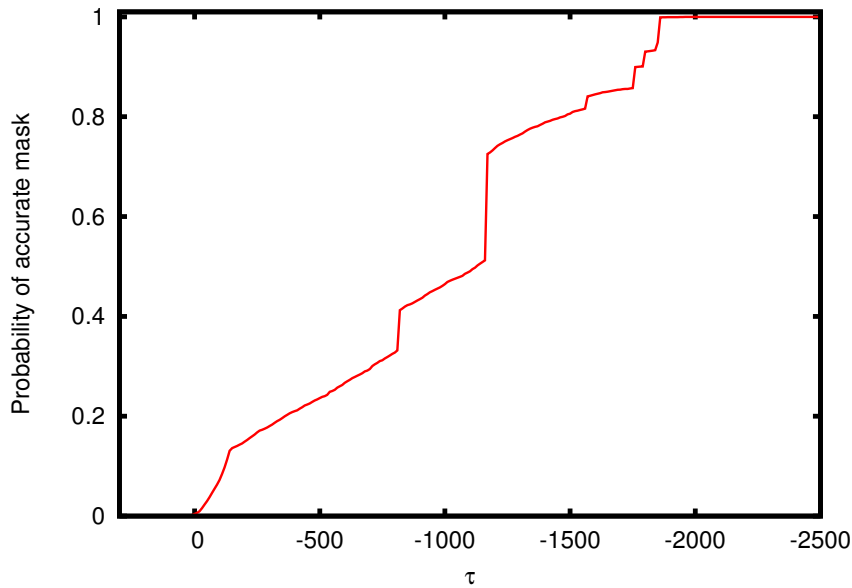
Figure 5.6: Probability p of to determine a correct mask given a determined threshold τ.

value, which is a lower value than the minimum value in the LL subband. Using this penalization it is obtained a bigger difference between pixels that belongs to human body and pixels outside the human body. Figure 5.2 c) and d) shows two different examples of I_g.

Once the image I_g is computed, it is needed to determine the threshold τ which divides the images into the two subsets. The first subset is the one that holds that $I_g \geq \tau$, which corresponds to the pixels of the human body. The second subset, which holds that $I_g < \tau$, belongs to the areas outside the human body. To determine τ, the following metodology is used: 1) compute the mask of the original image with [40] and then the LL subband of the resulting mask is computed; 2) compute the different masks of the I_g using all possible values of τ; and finally 3) compute the probability p of to determine a correct mask for each τ.

Figure 5.6 shows the probability p given a τ value. This value is calculated with the average of p of 2000 different I_g of different images. From this figure it can be determined that using a τ value of -2000 the probability of getting an accurate mask is 1.
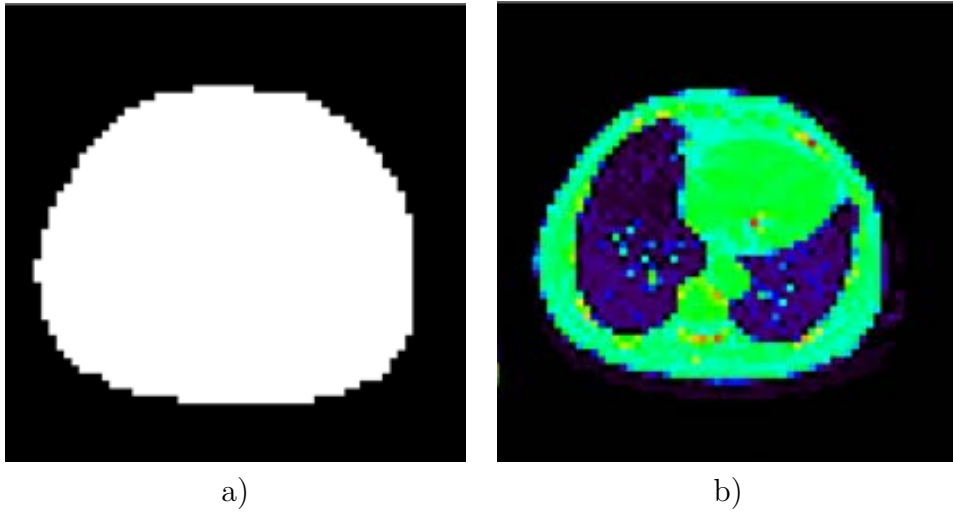
Figure 5.7: Binary mask (a) and LL subband after mask application (b).

After the mask generation, the expansion of the mask is performed to recover accurately the coefficients inside the mask. Using a 5/3 wavelet transform, the number of coefficients needed to undo the wavelet transform is 3 for the even subbands and 5 for the odd subbands [66].

The last step of the algorithm is the to fill the holes using a raster mode and to propagate the mask to the other subbands following the pyramidal relation of the wavelet transform [21]. Figure 5.7 a) shows the resulting mask in the LL subband and Figure 5.7 b) depicts the resulting segmented LL subband.

Summarizing, this method is divided in three main steps for each slice: 1) Generate a new image I_g applying K_{5x5} filter in LL subband of each slice; 2) apply the threshold τ to the image I_g and generate the binary mask; and 3) fill the mask and propagate to other subbands.

5.3 Experimental results

5.3.1 Segmentation Accuracy

The accuracy of the proposed segmentation method is obtained comparing the proposed mask with the reference mask. The accuracy is defined as

$$A = 100 \times \frac{\#R_{RoI} - \sum_{j=y_0}^{y_n} \sum_{i=x_0}^{x_n} |R_{RoI}(j,i) - P_{RoI}(j,i)|}{\#R_{RoI}},$$

where R_{RoI} and P_{RoI} are, respectively, the reference binary mask and the evaluated mask, $\#R_{RoI}$ denotes the number or RoI samples in R_{RoI} , and $(Y, X) = (y_0, x_0), ..., (y_n, x_n)$ are the set of pixel positions that defines R_{RoI}. In addition, the segmentation overhead that produces the proposed method is also reported. The overhead is defined as

$$O = 100 \times \frac{\#R_{RoI} + \sum_{j=y_0'}^{y_n'} \sum_{i=x_0'}^{x_n'} |R_{RoI}(j,i) - P_{RoI}(j,i)|}{\#R_{RoI}},$$

where R_{RoI} and P_{RoI} are, respectively, the reference binary mask and the evaluated mask, $\#R_{RoI}$ denotes the number or RoI samples in R_{RoI} , and $(Y', X') = (y_0', x_0'), ..., (y_n', x_n')$ are the set of pixel positions that defines P_{RoI}.

Table 5.1 shows the accuracy of the proposed method over different images of the corpus.

5.3.2 Coding Performance

To illustrate that air pixels greatly penalize the coding performance of JPEG2000, the impact produced by air pixels on the coding process of JPEG2000 is evaluated. All the experiments are obtained using 3 level of 5/3 wavelet transform and code-block sizes of 64x64. Figure 5.9 reports the number of bytes that are emitted by the bitplane coding engine at the end of each bitplane, for a single code-block of one wavelet subband of two CT images. The analyzed code-blocks lie in the non-biological region of the image, being most of the samples air pixels in the original domain. The Figure reports results when the image is coded with JPEG2000 without any segmentation stage, and with JPEG2000 with our segmentation proposal. In addition, Figure 5.8

Image	A	O	Image	A	O
SP4_T5-D5_1	100	19	SS16_T1-D075_6	100	28
SP4_T5-D5_2	100	30	SS16_T1-D075_7	100	21
SP4_T5-D5_3	100	25	SS16_T2-D1_2	100	22
SP4_T5-D5_4	100	25	SS16_T2-D1_3	100	22
SP4_T5-D5_5	100	30	SS16_T2-D1_4	100	22
SP4_T5-D5_6	100	16	SS16_T2-D1_5	100	22
SP4_T5-D5_7	100	15	SS16_T2-D2_1	99.99	25
SP4_T5-D5_8	100	16	SS16_T2-D2_2	100	25
SP4_T5-D5_9	100	17	SS16_T2-D2_3	100	25
SP4_T5-D5_10	100	19	SS16_T5-D5_11	100	18
SS16_T1-D075_1	100	28	SS16_T5-D5_12	100	19
SS16_T1-D075_2	100	25	SS16_T5-D5_1	100	20
SS16_T1-D075_3	100	17	SS16_T7-D5_1	99.99	25
SS16_T1-D075_4	100	18	SS16_T7-D5_2	100	26
SS16_T1-D075_5	100	16	SS16_T7-D5_3	100	24

Table 5.1: Lossless coding performance of JPEG2000 and JPEG2000 SEG with the percentage of non-biological area in the image.

SS16-T1__D075__2   SS16-T1__D075__4
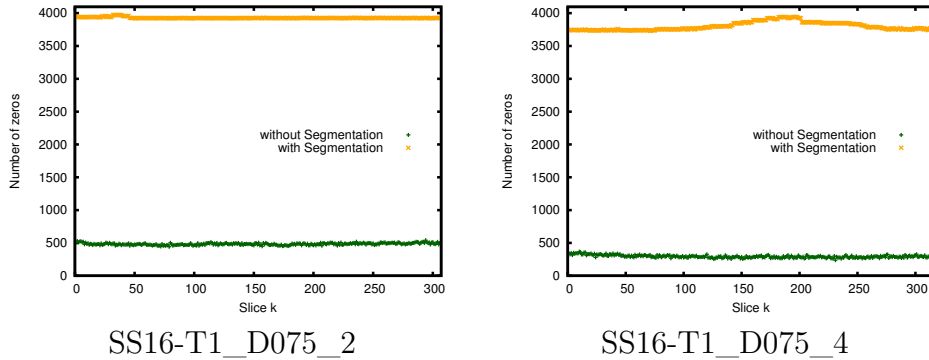
Figure 5.8: Number of samples zero with and without a segmentation method of one code-block of two images. The code blocks are in the low-horitzontal high-vertical frequencies subband of the first decomposition level (HL_1) of a CT image.

shows the number of zeros that our proposal generates into the same code-block in front of the number of zeros the produced by the wavelet transform.

The segmentation process reduces significantly the amount of bytes emitted by the bitplane encoder, due to of the suppression of the noise on those code-blocks that not contain biological data. These results indicate that the imaging artifacts produced

SS16-T1__D075__2   SS16-T1__D075__4
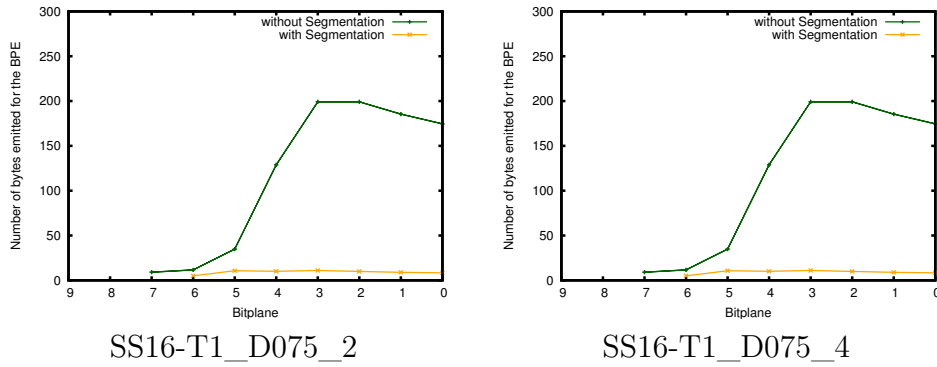
Figure 5.9: Bytes emitted by the bitplane coding engine of JPEG2000 with and without a segmentation method of one code-block of two images. The code blocks are in the low-horitzontal high-vertical frequencies subband of the first decomposition level (HL_1) of a CT image.

by air pixels strongly affect the performance of the JPEG2000 coding system.

Table 5.2 shows the coding performance evaluation of JPEG2000 and JPEG2000 with segmentation (JPEG2000 SEG). Results show that as higher percentage of non-biological area as better coding performance are achieved by JPEG2000 SEG.

Image	JPEG2000	JPEG2000 SEG	% BG	%Coding Gain
SP4_T5-D5_1	7.77	5.56	21	28
SP4_T5-D5_2	4.47	3.46	32	22
SP4_T5-D5_3	4.47	3.46	31	22
SP4_T5-D5_4	7.77	5.56	31	28
SP4_T5-D5_5	7.11	4.67	32	34
SP4_T5-D5_6	3.94	2.89	39	27
SP4_T5-D5_7	4.74	3.71	39	21
SP4_T5-D5_8	8.01	6.06	26	24
SP4_T5-D5_9	4.64	3.90	26	15
SP4_T5-D5_10	8.01	6.48	21	20
SS16_T1-D075_1	6.20	3.99	42	35
SS16_T1-D075_2	5.88	3.72	43	37
SS16_T1-D075_3	5.86	4.25	32	27
SS16_T1-D075_4	8.37	5.94	32	29
SS16_T1-D075_5	5.22	4.02	29	24
SS16_T1-D075_6	6.04	3.98	40	34
SS16_T1-D075_7	7.45	3.75	37	49
SS16_T2-D1_2	5.99	3.89	44	35
SS16_T2-D1_3	6.01	3.89	44	35
SS16_T2-D1_4	6.22	4.28	35	31
SS16_T2-D1_5	6.18	4.31	35	30
SS16_T2-D2_1	5.38	3.61	38	32
SS16_T2-D2_2	6.34	3.88	45	39
SS16_T2-D2_3	5.91	3.86	41	35
SS16_T5-D5_11	5.44	4.18	29	24
SS16_T5-D5_12	8.15	5.80	32	29
SS16_T5-D5_1	8.05	4.39	22	45
SS16_T7-D5_1	5.04	3.03	47	40
SS16_T7-D5_2	5.11	3.04	49	40
SS16_T7-D5_3	5.04	3.02	47	40

Table 5.2: Lossless coding performance of JPEG2000 and JPEG2000 SEG with the percentage of non-biological area in the image and the coding gain of the proposed method.

Chapter 6

Conclusions

6.1 Summary

Computed Tomography is one of the most used medical imaging technique. It is a noninvasive medical test obtained via a series of X-ray exposures, and generates 3D images that aid medical diagnosis. These images, of the inside of the human body, show organs, bones, soft tissue and blood vessels with greater clarity than standard X-rays, permitting radiologists to more easily diagnose problems such as cancer, infectious diseases, appendicitis, cardiovascular diseases, trauma and musculoskeletal disorders.

Taking into account the number of CT scans that are performed in a medical center every day, and the typical sizes of a CT image (around 150MB), medical centers use Picture Archiving and Communications Systems (PACS) [2] to manage all of this data. The Digital Imaging and Communications in Medicine (DICOM) standard [5] specifies the format used to store and distribute images in PACS. Due to the high cost of storage and transmission of medical digital images, data compression plays a key role.

Noise in Computed Tomography (CT) images appears because of the need of performing a CT with a low radiation dose, aimed to reduce the risk of patients getting cancer. Noise filtering is an important topic in the medical research field, since noise can obscure lesions otherwise visible on images obtained with higher dose parameters.

We have seen that noise not only complicates the medical diagnosis, but also influences the image coding process, producing low compression ratios. As a consequence, we introduce a noise-filtering multi-component JPEG2000 coding scheme.

The proposal includes a NF stage, aimed to reduce the noise of the CTs. The filtered noise is discarded since it does not supply relevant information to the medical diagnosis. The inclusion of a NF stage does not affect JPEG2000 compliance, thus easing its use in medical centers for image sharing. Experimental results indicate that suppressing the noise of the CTs through a NF stage does not penalize the visual perception of the image, additionally, the noise extraction permits a significant coding performance gain. The strategy based on NF+RWT+JPEG2000 produces the best coding performance, both for lossless and for progressive lossy-to-lossless coding. The benefits of this coding scheme are: to encode filtered images obtained by low radiation dose in CT acquisition, to obtain better image quality and coding performance, and to be compliant with JPEG2000 for use in DICOM.

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, 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.

The use of segmentation methods over CT images improve the coding performance significantly. Some of the methods presented in the literature are designed for specific substances or images, or are computational expensive. A new method for biological segmentation in the LL subband of the wavelet transform is presented.

This method obtain the same accuracy as methods performed in the original domain but with a considerable reduction of computational cost because of the reduced size of the image. However, they produce an overhead because for some images the threshold $\tau = -2000$ is so high and to recover correctly the segmented area, a mask expansion in needed. Moreover, the coding performance improvement are highly correlated with the percentage of background suppressed.

With this thesis, it is enhanced that the use of generic coding schemes are not sufficient to encode CT. Given the nature of the CT, the use of different preprocessing techniques are needed to obtain a competitive results in coding performance, and also to reduce the transmission times and costs.

6.2 Future research

The research presented in this thesis has been focused on Computed Tomography image coding within the JPEG2000 framework. Following this line of work, there are several directions in which the research can be extended.

The first future line of research, which is immediate, is the application of the segmentation method to other medical images, such as X-ray Angiography, MRI, and PET. As it is exposed in [41], the use of segmentation methods over X-ray Angiography images, improves the coding performance. In addition, there are studies that affirm the same idea to MRI [60, 61].

The second future line of research is the use of correlation modeling to other medical imaging techniques such as, MRI or PET. Note that, the direct application of this model is not possible to these other test, because there are differences in the acquisition. However, they also are a 3D images and have similarities between the slices. In this way, the study of the different acquisition processes, and the parameters involved in them should be studied.

The third future line of research is involved with the kernel of the filter used in the segmentation process. In these thesis, it is proposed a kernel that sets the weigh of the coefficients depending on the distance to the center coefficient. The study of different kernels, not only taking into account the distance, also the direction or the contribution on the inverse transform could be interesting.

Finally, following the research presented in [63] that propose the use of the compression of CT images for different substance, to add substance scalability to JPEG2000 could be an interesting topic for the medical community. A first approach of this research could be developed using the proposed method for multiple ROI in [39], where each ROI could be a different substance of the CT image. Note that with this approach, the size of the final codestream could be greater than the typical approach, but in terms of transmission and decoding, the results could be significantly improved.

Appendix A

List of all produced publications

- <u>Juan Muñoz-Gómez</u>, Joan Bartrina-Rapesta, Michael W. Marcellin, and Joan Serra-Sagristà. Correlation Modeling for Compression of Computed Tomography Images. IEEE Journal of Biomedical and Health Informatics. May 2013. Digital Object Identifier 10.1109/JBHI.2013.2264595

- Zhongwei Xu, Joan Bartrina-Rapesta, Victor Sanchez, Joan Serra-Sagristà, and <u>Juan Muñoz-Gómez</u>. Diagnostically Lossless compression of X-ray angiography images based on automatic segmentation using ray-casting and α-shapes. In Proceedings of the IEEE International Conference on Image Processing, Sept 2013.

- <u>Juan Muñoz-Gómez</u>, Joan Bartrina-Rapesta, Francesc Aulí-Llinas, and Joan Serra-Sagristà.Computed Tomography Image Coding through Air Filtering in the Wavelet Domain. In Proceedings of the IEEE Data Compression Conference. Apr. 2013.

- Zhongwei Xu, Joan Bartrina-Rapesta, Victor Sanchez, and Joan Serra-Sagristà, <u>Juan Muñoz-Gómez</u>. Diagnostically Lossless Compression of X-ray Angiographic Images through Background Suppression. In Proceedings of the IEEE Data Compression Conference, Apr. 2013.

- Miguel Hernández-Cabronero, Francesc Aulí-Llinàs, Joan Bartrina-Rapesta, Ian Blanes, Leandro Jiménez-Rodríguez, Michael W. Marcellin, <u>Juan Muñoz-Gómez</u>, Victor Sanchez, Joan Serra-Sagristà, Zhongwei Xu. Multicomponent compression of DNA microarray images. In Proceedings of the CEDI Workshop on Multimedia Data Coding and Transmission (WMDCT). Sep. 2012

- Joan Bartrina-Rapesta, Marc Navarro, <u>Juan Muñoz-Gómez</u>, Michael W. Marcellin, Jesús Ruberte, and Joan Serra-Sagristà. MicroCT Image Coding Based on Air Filtering. In Proceedings of the IEEE Data Compression Conference. April 2012.

- Miguel Hernández-Cabronero, <u>Juan Muñoz-Gómez</u>, Ian Blanes, Michael W. Marcellin, and Joan Serra-Sagristà. DNA microarray image coding. In Proceedings of the IEEE Data Compression Conference. April 2012.

- <u>Juan Muñoz-Gómez</u>, Joan Bartrina-Rapesta, Michael W. Marcellin, and Joan Serra-Sagristà. Influence of Noise Filtering in Coding Computed Tomography with JPEG2000. In Proceedings of the IEEE Data Compression Conference. March 2011.

- <u>Juan Muñoz-Gómez</u>, J. Bartrina-Rapesta, J. Serra-Sagristà, F. Aulí-Llinàs, I. Blanes, and L. Jiménez-Rodríguez. 4D remote sensing image coding with JPEG2000. SPIE International Conference on Satellite Data Compression, Communication and Processing. Society of Photo-optical Instrumentation Engineers (SPIE). August 2010.

- L. Jiménez-Rodríguez, F. Aulí-Llinàs, <u>Juan Muñoz-Gómez</u>, J. Bartrina-Rapesta, I. Blanes, and J. Serra-Sagristà. Rate allocation method for the fast transmission of pre-encoded meteorological data over JPIP. SPIE International Conference on Satellite Data Compression, Communication and Processing. Society of Photo-optical Instrumentation Engineers (SPIE). August 2010.

- Ian Blanes, Joan Serra-Sagristà, Francesc Aulí-Llinàs, Joan Bartrina-Rapesta, Leandro Jiménez-Rodríguez, <u>Juan Muñoz-Gómez</u>. Side information coding for the Pairwise Orthogonal Transform. in Proceedings of Workshop on Multimedia Data Coding and Transmission (WMDCT). Sep.2010.

- J. Bartrina-Rapesta, J. Serra-Sagristà, F. Aulí-Llinàs, and <u>Juan Muñoz-Gómez</u>. JPEG2000 ROI coding method with perfect fine-grain accuracy and lossless recovery. In Proceedings of the IEEE Asilomar Conference on Signals, Systems, and Computers. Nov. 2009 (Invited Paper)

Appendix B

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 [67], while images from the second, third and fourth sensors were obtained from The Cancer Imaging Archive [68].

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.

The following tables summarize the corpus characteristics. The first column provides image names, which end with an integer suffix to differentiate between multiple 3D images having the same acquisition characteristics. The second column gives the number of 3D images with the same image characteristics. The third and fourth columns give, respectively, slice thickness and slice distance, while pixel spacing within a slice is provided in column five. 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.

Table B.1: Siemens Sensation 16.

Image Names	# Images	Slice Thickness (mm)	Slice Distance (mm)	Pixel Spacing (mm,mm)	# slices (N_z)
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]
S16-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]

Table B.2: Siemens Somatom Plus 4.

Image Names	# Images	Slice Thickness (mm)	Slice Distance (mm)	Pixel Spacing (mm,mm)	# slices (N_z)
SP4-T5-D5_{1..10}	10	5	5	0.65,0.65	[48,71]

Table B.3: General Electric LightSpeed 16.

Image Names	# Images	Slice Thickness (mm)	Slice Distance (mm)	Pixel Spacing (mm,mm)	# slices (N_z)
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]

Table B.4: Philips Brilliance 40.

Image Names	# Images	Slice Thickness (mm)	Slice Distance (mm)	Pixel Spacing (mm,mm)	# slices (N_z)
B40-T1-D08_{1..17}	17	1	0.8	0.83,0.83	[48,71]

Bibliography

[1] 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.

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

[3] 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.

[4] 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.

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

[6] ——, "DICOM Part 5: Data Structures and Encoding," Sept 2013. [Online]. Available: http://medical.nema.org/dicom/2007/07_05pu.pdf

[7] Joint Photographic Experts Group, "JPEG Standard ISO/IEC ISO 10918-1 ITU-T Recommendation T.81," 2013. [Online]. Available: http://www.jpeg.org/jpeg/index.html

[8] ——, "JPEG Standard ISO/IEC ISO 10918-1 ITU-T Recommendation T.87," 2013. [Online]. Available: http://www.jpeg.org/jpeg/jpegls.html

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

[10] OsiriX, Advanced Imaging, "OsiriX HD," Aug 2013. [Online]. Available: https://itunes.apple.com/app/osirix-hd/id419227089

[11] International Telecomunication Union (ITU), "Cellular standards for 3G." [Online]. Available: "http://www.itu.int/osg/spu/imt-2000/technology.html"

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

[13] J. T. Bushberg, J. Seibert, E. M. L. Jr., and J. M. Boone, *The essential physics of Medical Imaging.* Philadelphia, Pennsylvania 19106 USA: Lippincott Williams and Wilkins, 2002.

[14] G. N. Hounsfield, "Computed medical imaging," *Medical Physics*, vol. 7, no. 4, pp. 283–290, 1980.

[15] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 623–656, 1948.

[16] D. Huffman, "A method for the construction of minimum redundancy codes," *Proceedings of the Institute of Radio Engineers*, vol. 40, p. 1098–1101, 1952.

[17] S. Golomb, "Run-length encodings," *IEEE Transactions on Information Theory*, vol. 12, p. 399–401, 1966.

[18] J. Rissanen and J. Langdon, G.G., "Arithmetic coding," *IBM Journal of Research and Development*, vol. 23, no. 2, pp. 149–162, 1979.

[19] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. on Computers*, vol. C-23, no. 1, pp. 90–93, Jan 1974.

[20] Joint Photographic Experts Group, "JPEG," Aug 2013. [Online]. Available: http://www.jpeg.org

[21] M. Antonini, M. Barlaud, P. Mathieu, , and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, April 1992.

[22] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, p. 3445–3462, Dec 1993.

[23] J. Tian and R. W. Jr., "A lossy image codec based on index coding," *In Proceedings of the IEEE Data Compression Conference (DCC)*, p. 456, Mar 1996.

[24] A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, p. 243–250, Jun 1996.

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

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

[27] M. D. Adams, "The JasPer Project ," 2013. [Online]. Available: http://www.ece.uvic.ca/~frodo/jasper/

[28] JJ2000, "A pure Java JPEG 2000 image codec ," 2013. [Online]. Available: http://code.google.com/p/jj2000/

[29] M. Rabbani and R. Joshi, "An overview of the jpeg2000 still image compression standard," *Signal Processing: Image Communication*, vol. 17, p. 3–48, 2002.

[30] D. S. Taubman, "High performance scalable image compression with ebcot," *IEEE Transactions on Image Processing*, vol. 9, no. 7, p. 1158–1170, Jul 2000.

[31] J. Joan 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.

[32] 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.

[33] Z. Xiong, X. Wu, S. Cheng, and J. Hua, "Lossy-to-lossless compression of medical volumetric data using three-dimensional integer wavelet transforms," *IEEE Trans. on Medical Imaging*, vol. 22, no. 3, pp. 459 –470, Mar 2003.

[34] 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.

[35] 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. 290–315, May 2001.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] Z. Xu, J. Bartrina-Rapesta, V. Sanchez, J. Serra-Sagristà, and J. Muñoz-Gomez, "Diagnostically lossless compression of X-ray angiography images based on automatic segmentation using ray-casting and α-shapes," *IEEE International Conference on Image Processing*, Sept 2013.

[42] 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.

[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "The SSIM index for image quality assessment," 2010. [Online]. Available: http://www.ece.uwaterloo.ca/~z70wang/research/ssim/

[44] A. Gonzalez and S. Darby, "Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries," *The Lancet*, vol. 363, pp. 345–351, 2004.

[45] M. Kalra, M. Maher, T. Toth, M. Hamberg, M. Blake, J.-A. Shepard, and S. aini., "Strategies for CT radiation dose optimization," *Radiology*, vol. 230, no. 3, pp. 619–628, 2004.

[46] J. Weickert, "A review of nonlinear diffusion filtering," *Scale-Space Theory in Computer Vision, Lecture Notes in Computer Science*, vol. 1252/1997, pp. 1–28, 1997.

[47] M. Kachelriess, O. Watzke, and W. A. Kalender, "Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT," *AAPM Medical Physics*, vol. 28, no. 4, pp. 475–490, 2001.

[48] F. Catte, P. Lions, J. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM Journal on Numerical Analisys.*, vol. 29, no. 1, pp. 182–193, 1992.

[49] E. Meijering, W. Niessen, and M. Viergever, "Diffusion-enhanced visualization and quantification of vascular anomalies in three-dimensional rotational angiography: Results of an in-vitro evaluation," *ELSEVIER Medical Image Analysis*, vol. 6, no. 3, pp. 215–233, 2002.

[50] A. Mendrik, E.-J. Vonken, A. Rutten, M. Viergever, and B. van Ginneken, "Noise reduction in computed tomography scans using 3-D anisotropic hybrid diffusion with continuous switch," *IEEE Trans. on Medical Imaging*, vol. 28, no. 10, pp. 1585 –1594, Oct 2009.

[51] P. Hao and Q. Shi, "Reversible integer KLT for progressive-to-lossless compression of multiple component images," *ICIP 2003. Proceedings of 2003 International Conference on Image Processing*, vol. 1, pp. 633–636, 2003.

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

[53] 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.

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

[55] 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.

[56] 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.

[57] N. E. M. Association, "DICOM Part 3: Information Object Definitions - A.1.2.9," 2004. [Online]. Available: http://medical.nema.org/dicom/2004/04_03PU3.PDF

[58] K. J. Kim, K. H. Lee, B. Kim, T. Richter, I. D. Yun, S. U. Lee, K. T. Bae, and H. Shim, "Improved Compressibility in JPEG2000 2D and 3D Reversible

Compressions of Thin section Chest CT Images by Increasing the Data Redundancy outside the Body Region," *IEEE COMSOC MMTC: Special Issue on Image Compression Technologies for Medical Applications*, vol. 6, no. 7, pp. 22–31, 2001.

[59] Hospital Mutua de Terrassa, "Terrassa, spain," Sept 2013. [Online]. Available: http://www.mutuaterrassa.cat/

[60] L. Shen and R. Rangayyan, "A segmentation-based lossless image coding method for high-resolution medical image compression," *IEEE Trans. on Medical Imaging*, vol. 16, no. 3, pp. 301–307, 1997.

[61] X. Bai, J. S. Jin, and D. Feng, "Segmentation-based multilayer diagnosis lossless medical image compression," in *Proceedings of the Pan-Sydney area workshop on Visual information processing.* Australian Computer Society, Inc., 2004, pp. 9–14.

[62] M. J. Zukoski, T. Boult, and T. Iyriboz, "A novel approach to medical image compression," *Int. J. Bioinformatics Res. Appl.*, vol. 2, no. 1, pp. 89–103, 2006.

[63] D. Spelic, D. Mongus, and B. Zalik, "Lossless compression of segmented ct medical images according to the hounsfield scale," in *Digital Image Computing Techniques and Applications (DICTA), 2011 International Conference on*, 2011, pp. 297–301.

[64] K.-U. Ilona, E. R. Vrscay, Z. Wang, C. Cavaro-Menard, D. Koff, B. Wallace, and B. Obara, "The impact of skull bone intensity on the quality of compressed ct neuro images," *Proc. SPIE*, vol. 8319, no. 1, 2012.

[65] S. Gokturk, C. Tomasi, B. Girod, and C. Beaulieu, "Medical image compression based on region of interest, with application to colon ct images," in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 3, 2001, pp. 2453–2456.

[66] J. Bartrina-Rapesta, *PhD. Thesis: Region of Interet Coding Methods for JPEG2000.* Universitat Autònoma de Barcelona, 2008.

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

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