Universitat Autònoma de Barcelona

Departament d'Enginyeria de la Informació i de les Comunicacions

Diagnostically lossless compression strategies for X-ray angiography images

Submitted to Universitat Autònoma de Barcelona
in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in Computer Science

by Zhongwei Xu

Bellaterra, June 2015

Directed by
Dr. Joan Bartrina-Rapesta
and Dr. Joan Serra-Sagristà

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, June 2015

Dr. Joan Bartrina-Rapesta and Dr. Joan Serra-Sagristà
(Supervisors)

Committee:

Dr. Vicente González Ruiz
Dr. Ian Blanes García
Dr. Anton Bardera Reig
Dr. Jordi Serra Ruiz
Dr. Francesc Aulí Llinàs

Abstract

The past several decades have witnessed a major evolution in medical imaging techniques, making medical images become commonplace in healthcare systems and an integral part of a patient medical record. Among the existing medical imaging modalities, X-ray imaging is one of the most popular technologies due to its low cost, high resolution and excellent capability to penetrate deep within tissue. In particular, X-ray angiographies –which use minimally invasive catheterization– and X-ray imaging are widely used to identify irregularities in the vascular system. X-ray angiography images can be classified into two types: general X-ray angiography (GXA) images, which present blood vessels in several body parts like arms, legs, foots, etc.; and coronary angiogram video squences (CAVSs), which only focus on coronary vessel trees for diagnosing cardiovascular diseases. Because of the differences in functions, these two types of images have different features: GXA images normally have high spatial resolutions (the width and height sizes) but low temporal resolution (the number of frames), while CAVSs usually have lower spatial resolutions but higher temporal resolution.

Due to the increasing number of medical studies using X-ray angiography images and the need to store and share them, compression of these images is becoming critical. Lossy compression has the advantage of high data reduction capability, but it is rarely accepted by medical communities because of the modification of data that may affect the diagnosis process. Lossless compression guarantees perfect reconstruction of the medical signal, but results in low compression ratios. Diagnostically lossless compression is becoming the preferred choice, as it provides an optimal trade-off between compression performance and diagnostic accuracy. In diagnostically lossless compression, the clinically relevant data is encoded without any loss while the irrelevant data is encoded with loss. In this scenario, identifying and distinguishing the clinically relevant from the clinically irrelevant data in medical images is the first and usually most important stage in diagnostically lossless compression methods.

In this thesis, two diagnostically lossless compression strategies are developed. The first one is proposed for GXA images. The second one if proposed for CAVSs. For GXA images, the clinically relevant focal area in each frame is first identified; and then a background-suppression approach is employed to increase the data redundancy of the images and hence improve the compression performance. For CAVSs, a frame-identification procedure is implemented to recognise the diagnostically unimportant frames that do not contain visible vessel structures; then, lossy compression is applied to these frames, and lossless compression is applied to the other frames.

Several compression techniques have been investigated for both types of images, including the DICOM-compliant standards JPEG2000, JPEG-LS and H.264/AVC, and the latest advanced video compression standard HEVC. For JPEG2000, multicomponent-transform and progressive lossy-to-lossless coding are also tested. Experimental results suggest that both the focal-area-identification and frame-identification processes are automatic in computation and accurate in clinically relevant data identification. Regarding the compression performance, for GXA images, when compared to the case of coding with no background-suppression, the diagnostically lossless compression method achieves average bit-stream reductions of as much as 34% and improvements on the reconstruction quality of up to 20 dB-SNR for progressive decoding; for CAVSs, the frame-identification followed by selective lossy & lossless compression strategy achieves bit-stream reductions of more than 19% on average as compared to lossless compression.

Contents

Chapter 1

Introduction

1.1 Motivation

Medical images have become commonplace in healthcare systems and an integral part of a patient medical record. It is hard nowadays to imagine how diagnosis could be accomplished without medical images. There exist several medical imaging modalities that allow studying the functionality and anatomy of the human body in a non- or minimally- invasive manner, including X-ray imaging, ultrasonography, magnetic resonance and radionuclide [1]. Since the world-wide growth of the ageing population, the vascular system diseases (e.g., coarctation of the aorta, arteriosclerosis, etc.) are the main risks for human health today [2, 3]. Angiography, which uses catheterization with a particular "contrast agent", was developed and is now one of the most popular existing imaging modalities. Angiography images help improve the diagnosis of vascular system diseases. Among the different angiography-combined modalities, X-ray angiography using organic iodine compounds as the "contrast agent" remains the gold standard for identification of underlying blood vessels, due to its low cost, high resolution and excellent capability to penetrate deep within tissue [4].

X-ray angiography images can be divided into two categories: images aiming only at the coronary vessel trees, called *Coronary Angiogram Video Sequences* (CAVSs); and images aiming at the other parts of the human body like arms, legs, feet, etc., called *General X-ray Angiography* (GXA) images. A huge number of medical studies

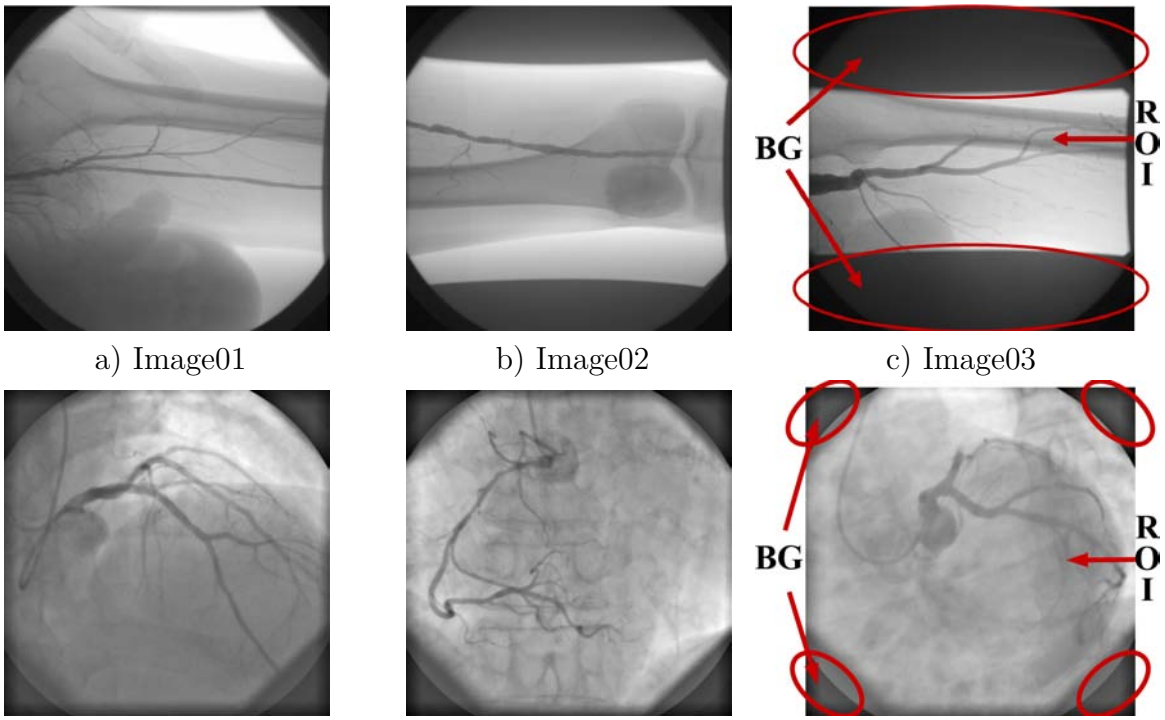
using X-ray angiography are processed in today's world. For instance, the American College of Cardiology National Cardiovascular Data Registry, which collects data from over 80% of the hospitals in the United States, has reported more than 12 million coronary angiography and percutaneous coronary intervention procedures performed from 1998 to 2011 [5], and this figure is only expected to grow. Recent advances in telemedicine require that X-ray angiography images be efficiently transmitted over networks of limited bandwidth. Moreover, a recent trend towards facilitating the general public on-line access to their own medical records has also become of significant interest to major companies and healthcare institutions [6]. Considering the usually large file size of X-ray angiography images, this medical data hence poses heavy demands on storage and transmission resources, which makes the compression of these images become critical.

Regarding compression strategies, lossy compression can achieve a high compression ratio but it is rarely accepted by the medical community [7]; the lossless compression obtains perfect reconstruction for medical data, but it yields low compression ratios. In this scenario, diagnostically lossless compression, which compresses the clinically relevant data without any loss and encodes the irrelevant data with loss, turns out to be an optimal solution dealing with medical data to get high compression performance while maintaining diagnostic accuracy. Regarding compression techniques, those which are compatible with the Digital Imaging and Communications in Medicine (DICOM) standard [8] and its associated Picture Archiving and Communication Systems (PACS) [9] are preferred; they include JPEG2000 [10], JPEG-LS [11] and H.264/AVC [12]. The latest HEVC [13] coding technique is also investigated because of its competitive compression ability.

This thesis started in September 2011 studying the state-of-the-art diagnostically lossless compression strategies for X-ray angiography images. From September 2011 to July 2013, the research focused on developing the compression strategies for GXA images. After this research was successfully finished, our research shifted to the compression of CAVSs, which was implemented during September 2013 and April 2015. Based on the different features of GXA images and CAVSs, we developed two main strategies. The first one is *Background suppression*, which uses image

segmentation to separate the Region Of Interest (ROI) and the clinically irrelevant background areas in each single frame of the GXA image, and suppresses the data in background areas to improve the compression performance. The second one is *Frame identification*, which distinguishes the clinically relevant and irrelevant frames, in order to exploit further the extra data redundancy in the temporal domain of the CAVS.

1.2 X-ray angiography images



a) Image01 b) Image02 c) Image03

d) Image04 e) Image05 f) Image06

Figure 1.1: Sample frames of three different GXA images (row 1) and three CAVSs (row 2).

As previously mentioned, X-ray angiography images can be categorized into two types: GXA images and CAVSs. Both types of images employ the catheterization process with organic iodine compounds as the "contrast agent". The X-ray imaging

projector is used to record the flow of the injected "contrast agent" through the blood vessels over a specific period of time, obtaining a collection of frames. Examples of several frames from different GXA images and CAVSs are presented in Figure 1.1.

Because of the differences in functions, GXA images and CAVSs have several distinguishable features:

(1) *Number of frames*: as coronary vessels are closer to the heart with a fast blood flow speed, CAVSs usually use a high CineRate[1] (i.e., a low FrameTime[2]) to trace the coronay vessel trees over several heart cycles, resulting in a large number of frames. On the other hand, GXA images are normally acquired using a lower CineRate (i.e., a higher FrameTime) with less imaging time because of the slower flow speed of blood, which generates less frames. We present the values of CinRate, FrameTime and the number of frames for our experimental corpus of GXA images and CAVSs in Table 1.1.

(2) *Spatial resolution*: regarding the 2-dimensional spatial resolution of each single frame, GXA images usually have higher width and height sizes than CAVSs, in order to contain a big body region (e.g., one foot). For the tested data in Table 1.1, for instance, the area size of a GXA frame is 4 times bigger than that of a CAVS frame.

(3) *Proportion of background areas*: both GXA images and CAVSs have background (BG) areas that do not contain any medical related information. These BG areas are derived from the geometric masks applied during the imaging process for presentation purposes in order to reject any of the pixels located outside of ROI [14]. As shown in Figure 1.1 and Table 1.1, GXA images normally have higher proportion of BG areas than CAVSs.

Table 1.1 lists the mentioned data of our experimental image corpus. The image corpus includes 60 GXA images and 72 CAVSs. More details of the experimental data are introduced in sections 2.2 and 3.2

[1]DICOM Tag: (0018,0040). Description: Number of frames per second.

[2]DICOM Tag: (0018,1063). Description: Nominal time (in msec) per individual frame. Note that: $FrameTime = \frac{1000\,msec}{CineRate}$.

Table 1.1: Information of GXA images and CAVSs

Image type (Number of images)	CineRate (FrameTime)	Frame Number	Spatial resolution (width × height)	background area in %
GXA Images (60)	4 ∼ 1 (250 msec ∼ 1000 msec)	2 ∼ 29 (in average 7.63)	1024 × 1024	9.80% ∼ 59.47% (in average 28.82%)
CAVSs (72)	15 (66.67 msec)	41 ∼ 151 (in average 79.07)	512 × 512	6.34% ∼ 9.20% (in average 7.34%)

1.3 Medical Image Compression

During the last two decades, because of the rapid development and wide use of digital medical image data, PACS and its associated DICOM standard were developed to facilitate the storage and exchange of medical images and videos within medical centers and hospitals.

Both the DICOM standard and the definition of PACS are briefly introduced below. The image compression techniques and strategies applied on medical images are also introduced here.

1.3.1 DICOM standard and PACS

DICOM

DICOM was first developed by the American College of Radiology (ACR) and National Electrical Manufacturers Association (NEMA) in 1985, in order to unify the output formats of medical images generated by the imaging machines from different manufacturers. The latest version of DICOM is "The DICOM Standard 2015b", which is presently under the management from the Medical Imaging & Technology Alliance - a division of NEMA [8].

DICOM standard now comprises 20 main parts, covering guidelines from file formats definition to network communication protocols. Among these parts, the "DICOM Part 5: Data Structures and Encoding" specifies the compressed formats allowed in the standard, including: JPEG [15], Run Length Encoding [16], JPEG-LS [11], JPEG2000 [10], MPEG-2 [17], and H.264/AVC [18]. JPEG2000, JPEG-LS and H.264/AVC are chosen in our experiments due to their unique advantages, which

are indicated in the following section 1.3.2; moreover, the latest HEVC technique is also applied in our proposals, and is also briefly introduced in section 1.3.2.

PACS

PACS is developed for providing economical storage and efficient access of various medical image records of different modalities and remote locations. A PACS system is primarily comprised of four components: image acquisition modules for obtaining input medical images, data management workstations for displaying and processing these images, storage infrastructures for archiving the images and the related medical reports, and a secured transmission network for connecting another PACS and sharing the medical data [9, 19]. Note that, DICOM is the universal format used in PACS image storage and transfer [19].

1.3.2 Image compression techniques

The global generation of digital multimedia data (including images and videos) has never ceased, which is posing an increasing pressure to the finite worldwide storage and transmission capacity. Hence, various advanced image compression techniques were proposed during the last two decades. Among these techniques, several have been accepted by the DICOM standard, helping medical centers and hospitals accomplish the compression of this legally and diagnostically sensitive data [20]. We would like to introduce the following four image compression techniques and their unique features. All these four techniques are employed in the experiments of this thesis.

JPEG2000

Thanks in part to JPEG [15], a compression technique for still images based on Discrete Cosine Transform (DCT), developed by members from the International Telecommunication Union (ITU) and the International Organization for Standardization (ISO) during the beginning of 1990s, Internet nowadays is full of images. However, JPEG compression misses several features that are required in some professional fields (e.g., the random codestream access and processing feature needed in

telemedicine over networks of limited bandwidth). Thus, JPEG2000 [10], a new coding system for still images with more features was built in the end of 1990s. JPEG2000 is based on discrete wavelet transform (DWT) and uses Embedded Block Coding with Optimized Truncation of the embedded bitstreams (EBCOT) [21] algorithm as the basic encoding engine. Several significant features of JPEG2000 are listed here:

1) Superior low bit-rate performance: compared to the JPEG standard, JPEG2000 achieves a superior reconstruction performance at low bit-rates, which is useful for the image transmission through a network with limited bandwidth.

2) Region-of-Interest Coding: images normally contains some areas more important than the others. The Region-of-Interest Coding feature is developed for this case, where the important areas are encoded with high quality and the rest with lower quality. This feature is widely used in medical image compression for achieving the diagnostically lossless compression.

3) Scalability and bitstream parsing: this feature allows JPEG2000 to decode the image with different resolutions or qualities, which enables the image retrieval without decoding the whole image.

4) Progressive transmission by pixel accuracy and resolution: this feature allows the image to be reconstructed with increasing pixel accuracy or spatial resolution.

For more details of the architecture and features of JPEG2000 standard, please refer to [22] and [23].

Available implementation software for JPEG2000 standard Part 1 (ISO/IEC 15444-1) include: Kakadu [24] and BOI [25]; both are used in the experimental part of the thesis.

JPEG-LS

JPEG-LS [11] uses LOCO-I (LOw COmplexity LOssless COmpression for Images) [26] algorithm as its core encoding engine to achieve lossless and near-lossless compression

for still images. The advantage of LOCO-I algorithm is that it attains compression ratios similar or superior to those state-of-the-art techniques based on arithmetic coding (e.g., JPEG2000 lossless compression), but with a much lower computing complexity level [27].

The LOCO-I algorithm was developed at Hewlett-Packard Laboratories. LOCO-I/JPEG-LS implementation is available in Hewlett-Packard Labs webpage [28].

H.264/AVC

H.264/AVC [18] video coding standard was developed about 10 years ago to replace the even older MPEG-2 video coding standard (also known as H.262). It uses block-based DCT and motion-compensation as the basic processes to encode the video contents [29]. Today H.264/AVC is the most commonly used format for recording, encoding and transferring video data. The following highlighted features are included in H.264/AVC for achieving an enhanced coding efficiency [12]:

1) H.264/AVC supports more selection of motion compensation block sizes and shapes than any previous standard.

2) The motion-compensated prediction signal can be weighted and offset with values specified by the encoder, which makes the encoder more flexible for various compression purposes.

3) Small block-size transform is used in H.264/AVC, which reduces the "ringing" artifacts.

4) The advanced entropy coding method – arithmetic coding – is included in H.264/AVC to get improved compression performance compared to the previous standards.

For more of the advanced features of H.264/AVC standard, please refer to [12]. The reference software of H.264/AVC is available at [30].

HEVC

The appearance of videos with beyond-HD formats (e.g., 4k × 2k or 8k × 4k resolution) and the expectation of growing popularity of these formats have made H.264/AVC less efficient than before. Also, the increasing popularity of mobile devices and tablet PCs that have limited computing capability also calls for a new well-designed video coding standard that supports parallel processing.

HEVC (High Efficiency Video Coding) is a directly succeeding project of H.264/AVC video coding standard [13]. It has been developed to be compatible with all the existing applications of H.264/AVC, but can also address the aforementioned two issues: the increased video resolution and the use of parallel processing architectures. Compared to H.264/AVC, the highlighted features of HEVC include [31]:

1) HEVC supports more block sizes (from 64 × 64 to 8 × 8 pixels), and square or rectangular (non-square) prediction and transform units. The usage of larger block sizes benefits the compression of videos with HD or beyond-HD resolutions, since larger smooth regions are contained in these videos.

2) Several mechanisms are applied in HEVC for supporting parallel encoding and decoding, including tiles and wavefront parallel processing (WPP).

3) More intra-prediction modes are supported in HEVC.

4) More integer transforms are supported, with size from 32 × 32 to 4 × 4 and shapes of square or even non-square.

For detailed comparison between HEVC and H.264/AVC, and the compression performance of HEVC, please refer to [31] and [13].

The reference software of HEVC is available at [32], which includes a detailed software manual and several ready-to-use configuration profiles (e.g., Intra main, Random Access and Random Access main RExt, which are employed in the thesis).

1.3.3 Medical image compression strategies

In the past 5 years, many proposals using various compression strategies were presented for different types of medical images: a segmentation and JPEG2000 compression combined strategy was proposed for X-ray computed tomography (CT) images in [33]; a similar idea was used in [34] for medical ultrasound images, where the particular fan-shaped ROI was first extracted, and the lossy compression with different quantization values was applied to the ROI and non-ROI, respectively; in [35], the authors discussed the influence of noise-filtering process in JPEG2000 compression of X-ray CT images, and a correlation modeling strategy for coding this type of 3-D images was developed in [36]; the lossless compression for DNA microarray images using DICOM-compliant JPEG2000 technique was researched in [37], and distortion metrics for lossy compression of the DNA microarray images were also investigated [38]; in [39] the lossless compression of medical images using HEVC method was analysed, which was then applied to digital pathology images in [40] and [41].

Regarding the compression strategies for X-ray angiography images, several contributions were presented during the last decade:

1) In [42] the angiography image was first split into equal size macroblocks; those blocks containing vessel parts were then detected through standard deviation computing, and a block-based adaptive quantization strategy using H.264/AVC technique was applied to the image.

2) [43] proposed a wavelet-based contourlet transform compression method. ROI was first identified through motion detection approach. The coefficients results from contourlet transform were then combined with the ROI results, gaining more data redundancy by removing the low-level contourlet coefficients belonging to non-ROI areas.

3) [44] introduced a wavelet based ROI compression strategy. The ROI was also defined by motion detection approach, which only gave a coarse vessel region and could miss the tiny but important vessel branches area.

4) Similar to [44], in [45], a segmentation process was also applied to separate ROI and non-ROI areas in angiography images. The ROI areas were compressed without loss but the non-ROI areas were lossy compressed. The method of [45] has the same shortage as [44], i.e., the segmentation of ROI is not accurate, which could wrongly consider some tiny vessels as non-ROI.

The methods of [42] and [43] are lossy compression approaches, which the medical community is usually reluctant to deal with. The proposed methods in [44] and [45], and also the methods of [33] and [34] all belong to diagnostically lossless compression strategies, which have inspired the contributions in this thesis.

1.4 Thesis contribution and organization

The main contributions of this thesis are the two diagnostically lossless compression strategies (i.e., the Background Suppression strategy and Frame Identification strategy) developed for the GXA images and CAVSs, respectively. The former contribution has originated one conference paper [46] and one journal paper (currently under review) [47], and the latter contribution has originated another journal paper (currently under review too) [48].

Chapter 2 introduces the Background Suppression strategy in details. Experimental results using 60 GXA images are also included in Chapter 2 for evaluating not only the ROI segmentation performance, but also the improved compression performance.

In Chapter 3, all stages of the Frame Identification strategy are presented. Sufficient experimental results are also provided to prove the high accuracy of the identification process, and to show the compression benefits of this strategy.

Finally, Chapter 4 concludes this thesis and indicates the future works after this research.

Chapter 2

Background suppression strategy for GXA compression

General X-ray angiography (GXA) images are widely used to identify irregularities in the vascular system. Because of their high spatial and time resolutions and the increasing amount of X-ray angio images generated, coding of these images is becoming essential.

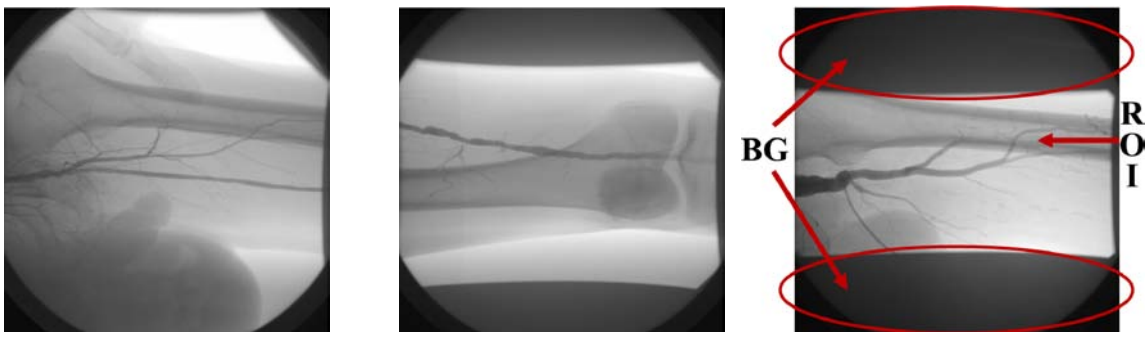
This chapter proposes a diagnostically lossless coding method based on automatic segmentation of the focal area using ray-casting and α-shapes. The diagnostically relevant ROI is first identified by exploiting the inherent symmetrical features of the GXA images. The background is then suppressed and the resulting images are encoded using lossless and progressive lossy-to-lossless coding methods, including JPEG-LS, JPEG2000, H.264/AVC and HEVC. Experimental results on a large set of GXA images suggest that the proposed method correctly identifies the ROI. When compared to the case of coding with no background suppression, the method achieves average bit-stream reductions of as much as 34% and improvements on the reconstruction quality of up to 20 dB-SNR for progressive decoding.

The chapter is organized as follows: our proposed compression strategy is described in Section 2.1. Section 2.2 presents an extensive experimental evaluation for the cases of diagnostically lossless and progressive lossy-to-lossless coding. This section also discusses, in collaboration with physicians from Hospital Fundació Mútua

de Terrassa, Spain, the accuracy of the proposed segmentation technique. Section 2.3 concludes this chapter.

2.1 Proposed coding scheme

The proposed coding method is based on the fact that improvements in coding efficiency may be achieved by exploiting some of the inherent symmetrical features of medical images [49, 50]. For example, in GXA images, there are usually two distinguishable areas: the ROI, depicting skeleton and tissues, and the BG, depicting non-clinically relevant information, as shown in Figure 2.1. Note that in these sample frames, the ROI is located in the center of the image (i.e., the focal area) and the BG features radially symmetrical properties around the ROI. Based on this observation, we focus on exploiting these symmetrical features to attain automatic segmentation and thus increase coding efficiency.



a) Image01 b) Image02 c) Image03

Figure 2.1: Sample frames of three different GXA images.

Our method, as illustrated in Figure 3.2, consists of two main stages, the first stage deals with automatic ROI segmentation, while the second stage focuses on data coding. The automatic ROI segmentation is based on ray-casting and α-shapes, which provide a high level of accuracy with low computational complexity. After segmentation, our method suppresses the BG from the image to increase data redundancy. In the second stage, the method employs lossless or progressive lossy-to-lossless (PLL)

coding on the BG-suppressed image. In the following sections, we describe in more detail these two stages.

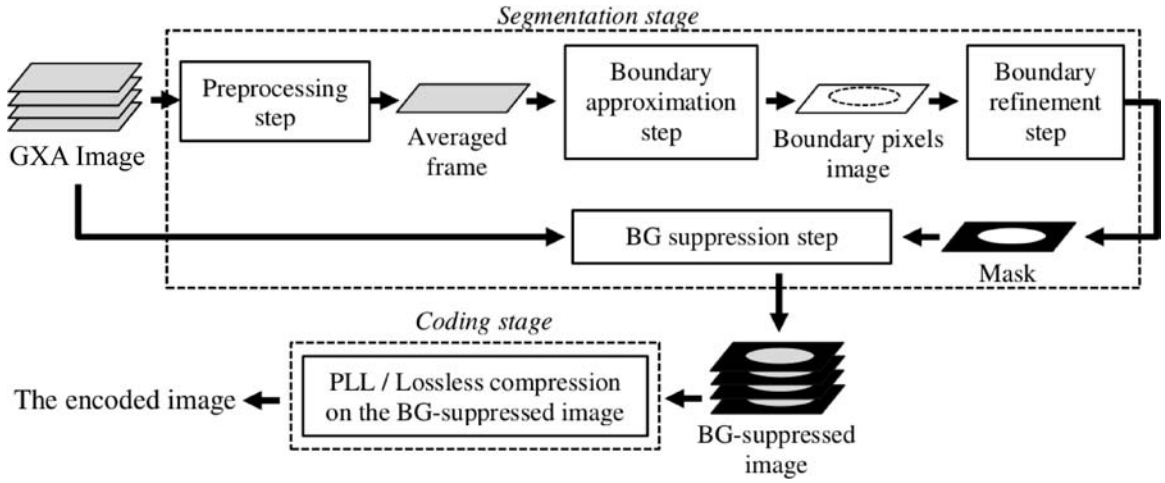


Figure 2.2: Block diagram of the proposed diagnostically lossless coding method.

2.1.1 Segmentation Stage

The segmentation stage comprises four steps: 1) preprocessing; 2) boundary approximation; 3) boundary refinement; and 4) BG suppression.

Preprocessing step

This step reduces the amount of noise in the data and exploits correlations among frames. GXA images contain random noise introduced by unblocked secondary radiation, poor film-developing and handling, or by the digitization process [51], which may affect the segmentation accuracy. To reduce this random noise, several techniques may be employed, such as neighbor average filtering [52], median filtering [53], non-local means [54] and 3D block matching [55]. It is highly desirable, however, to minimize the blurring effect that these common techniques have on the edge information, as edges are a prominent feature for ROI identification. To this end, we employ anisotropic diffusion filtering in each frame as it is capable to efficiently reduce noise while preserving edge information [56].

Since GXA images commonly consist of several frames that are usually highly correlated, a simple averaging operation may be used after noise reduction to generate a single frame that preserves the boundary between the ROI and BG. Here, we employ an averaging operation defined as $I_{avg}(x,y) = \left(\sum_{f=1}^{F} I_f(x,y)\right)/F$, where F is the number of frames in the GXA image, and $I_f(x,y)$ and $I_{avg}(x,y)$ denote the intensity value of the spatial position (x,y) in frame f and the average frame I_{avg}, respectively. By employing this simple averaging operation, we reduce the computational complexity of subsequent steps since segmentation can now be performed on the average frame and the results be used to identify the ROI in each frame.

Boundary approximation step

This second step computes a coarse approximation of the location of the boundary between the ROI and BG on the average frame I_{avg} by employing ray-casting and the image (pixel intensity) profiles computed along a set of rays [57]. Let P denote the center of I_{avg}, $R_n(P,\theta_n)$ denote ray n projected from P towards the periphery at an angle θ_n (see Figure 2.3(a)), and A_n denote the image profile along ray $R_n(P,\theta_n)$ computed using nearest-neighbor interpolation (see Figure 2.3(b)). The image profile A_n provides information about important intensity changes along the ray $R_n(P,\theta_n)$, which may be used to locate the position of the boundary between the ROI and BG. Due to the symmetrical properties of the ROI, important intensity changes along ray $R_n(P,\theta_n)$ usually occur at a very similar Euclidean distance from P as in ray $R_n(P,\theta_n+\pi)$. We call such two rays, $R_n(P,\theta_n)$ and $R_n(P,\theta_n+\pi)$, symmetrical rays. Figure 2.3(a) illustrates this concept by depicting rays $R_1(P,\theta_1)$ and $R_2(P,\theta_2)$, and their symmetrical rays $R_1(P,\theta_1+\pi)$ and $R_2(P,\theta_2+\pi)$. Figure 2.3(b)-(c) plots the corresponding image profiles A_1, $A_{1+\pi}$ and A_2, $A_{2+\pi}$. Note that the first significant intensity changes along the image profile plot, moving from the periphery towards P, usually happen at the boundary between ROI and BG and therefore, may be used to approximate the location of this boundary.

Nevertheless, in many cases, the intensity values of ROI and BG tend to be very similar in the boundary region for some of the rays, making it challenging to approximate the boundary location by simply analyzing the image profile along such rays. To

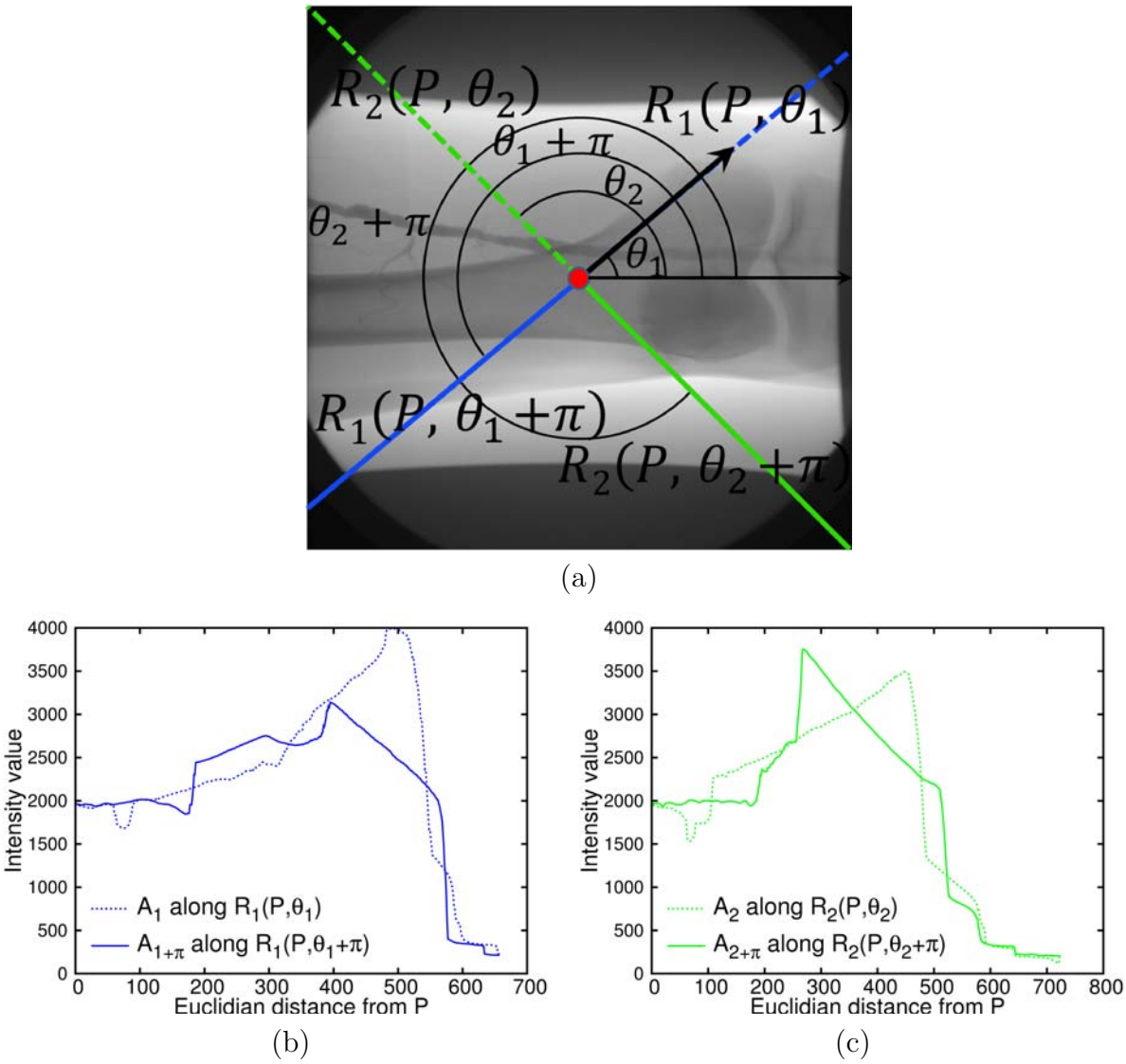


Figure 2.3: (a) Two pairs of symmetrical rays projected from P (red dot): $R_1(P,\theta_1)$, $R_1(P,\theta_1+\pi)$, and $R_2(P,\theta_2)$ and $R_2(P,\theta_2+\pi)$; (b)- (c) the corresponding image profiles.

overcome this, we exploit the radially symmetrical properties of ROIs and estimate the location of the boundary along a challenging ray by using the location of the first significant intensity change along the corresponding symmetrical ray. This idea is illustrated in Figure 2.4. We follow the next procedure and criterion to identify the significant intensity changes and determine the location of the boundary along a pair

of symmetrical rays:

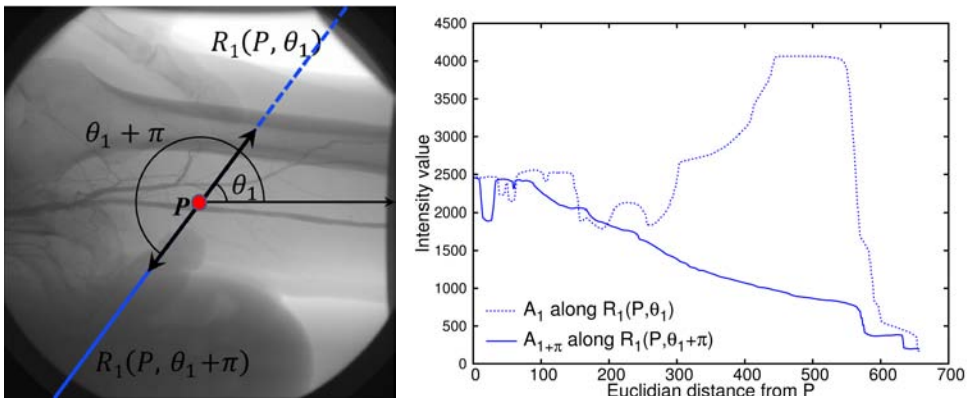


Figure 2.4: Sample case where the boundary is hard to locate in the intensity profile of one ray (solid ray) but it is easy to detect in the corresponding symmetrical ray (dashed ray).

1. For each pair of symmetrical rays, $R_n(P, \theta_n)$ and $R_n(P, \theta_n + \pi)$, we compute the corresponding image profiles, A_n and $A_{n+\pi}$.

2. For each pair A_n and $A_{n+\pi}$, we compute the corresponding intensity-change sets, denoted by C_n and $C_{n+\pi}$, respectively. An intensity-change set stores the largest intensity change within a small window w_1. Figure 2.5 depicts a sample computation of C_n.

3. We compute the maximum value $M = max\{C_n, C_{n+\pi}\}$ and threshold $T = M \times t$, where $0 < t < 1$.

4. In the intensity-change set where M is found, we search for the first element larger than T, denoted by B_n (or $B_{n+\pi}$).

5. We estimate $B_{n+\pi}$ (or B_n) in $C_{n+\pi}$ (or C_n) by searching for the largest element within a window of size w_2 centered in $C_{n+\pi}[B_n]$ (or $C_n[B_{n+\pi}]$) (see Figure 2.6). Positions $A_n[B_n]$ and $A_{n+\pi}[B_{n+\pi}]$ correspond to the position of the boundary along rays $R_n(P, \theta_n)$ and $R_n(P, \theta_n + \pi)$, respectively.

6. We repeat steps 1-5 for all pairs of symmetrical rays.

Section 2.2 reports on the values for parameters t, w_1 and w_2 that result in the best performance for the data set used in the experimental evaluation.

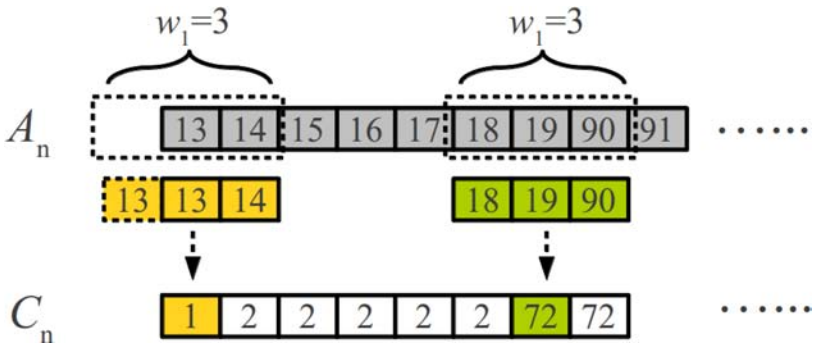


Figure 2.5: Example of an image profile A_n and the corresponding intensity-change set C_n computed with a window of width $w_1 = 3$.

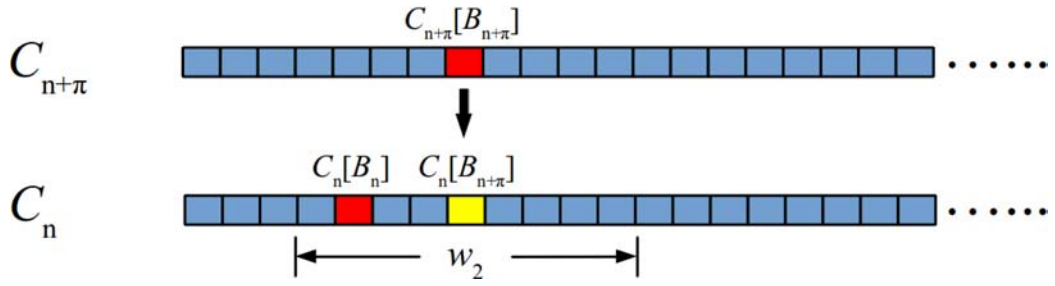


Figure 2.6: Sample estimation of B_n: $B_{n+\pi}$ is used to determine B_n in C_n by searching for the largest element within a window of size w_2 centered in $C_n[B_{n+\pi}]$.

Boundary refinement step

The previous step results in a set of locations that approximates the position of the boundary between the ROI and BG. Figure 2.7(a)-(b) show examples of such locations, depicted as white pixels over a black background, for an average frame I_{avg}. Note that these pixels only provide a coarse approximation of the overall boundary, for instance in the zoom area in Figure 2.7(b) we can see that several pixels are cluttered in a small area and disconnected with the other pixels. In order to refine the

boundary location and compute a closed contour, we link these pixels by employing α-shapes [58]. The objective is to create a closed contour that accurately describes the boundary between ROI and BG, so that this contour can be used to create a binary mask.

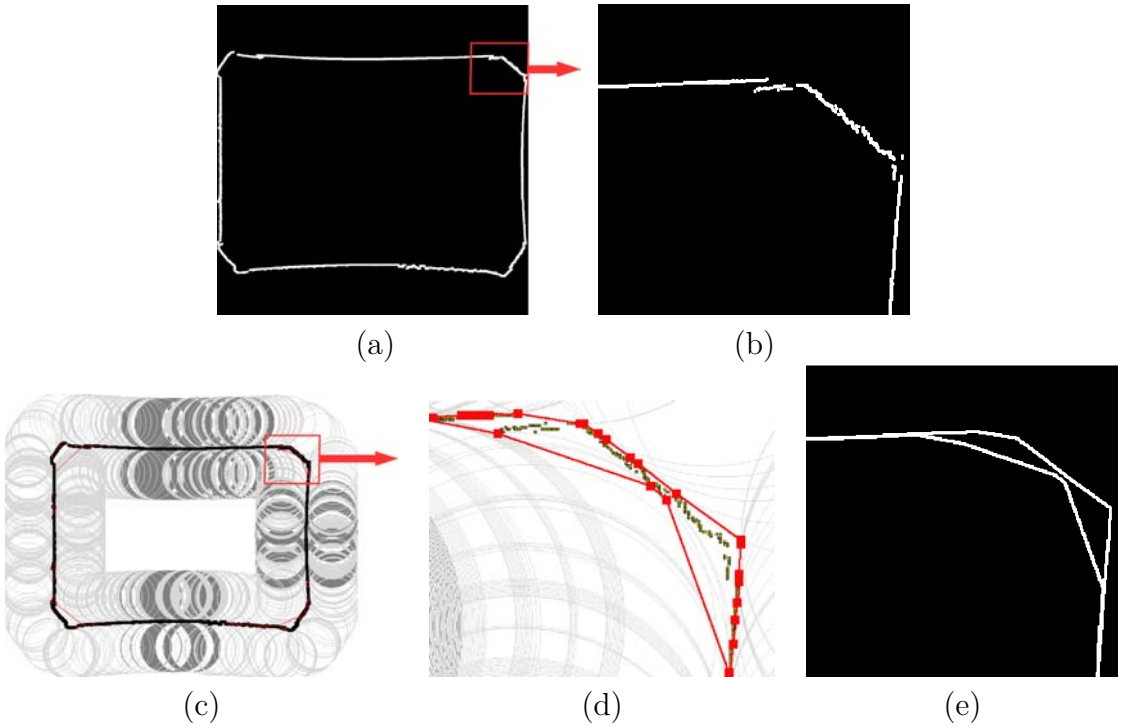


Figure 2.7: (a)-(b) Boundary points for I_{avg} of Image02 depicted in Figure 2.1. (c)-(d) Points (red squares) where an edge can be traced for a closed disk with r = -1/α. (e) Inner and outer contours computed with $\alpha = 0.01$.

Let us define a disk of radius $r = 1/\alpha$, such that if $\alpha > 0$, we obtain a closed disk; if $\alpha = 0$, we obtain a closed half-plane; and if $\alpha < 0$, we obtain the closure of the complement of a closed disk. Let us assume that the set of pixels corresponding to the boundary location forms a set of points on a plane, where the location of each pixel i denotes the location of point P_i in the point set. Based on this assumption, we compute a closed contour as follows:

1. For each point P_i in the point set, we create a vertex V_i.
2. We create an edge between two vertices V_i and V_j whenever there exists a disk

of radius $1/\alpha$ containing the entire point set and which has the property that P_i and P_j lie on the disk boundary.

After employing α-shapes we obtain a closed contour with less distortion than those generated by simple morphological operations like closing or dilation. Figure 2.7(c) shows those points –squared in red in the figure– where an edge can be traced for a closed disk of radius $r = -1/\alpha$, with $\alpha = 0.01$. Note that such small r values may result in additional closed contours inside the set of boundary points (see Figure 2.7(e)). In such cases, we select the outermost contour as the ROI boundary to ensure that the whole ROI is inside the contour.

BG suppression step

After computing the closed contour between the ROI and BG, we compute a binary mask by setting the intensity values of those points inside the contour to 1 (ROI) and those outside the contour to 0 (BG). We then achieve BG suppression by applying a logical AND operation between this mask and each frame of the GXA image, which sets the BG to zero. There is no need to transmit this mask. Mask results and BG-suppressed images are reported in Section 2.2.

2.1.2 Coding Stage

In this work, we focus on four coding techniques, JPEG-LS, JPEG2000, H.264/AVC and HEVC. All of them support lossless coding and provide excellent coding performance. Note that only JPEG-LS, JPEG2000 and H.264/AVC are included in DICOM. We are particularly interested in JPEG2000 as this coding standard offers a richer set of coding features than any other lossless coding method. These features include scalability by resolution and quality and the capability to exploit data redundancies among frames of GXA images through the use of a multi-component transform.

It is important to mention that BG-suppressed frames of GXA images usually contain sharp boundaries between BG and ROI that may generate a large amount of high frequencies responses during the spatial wavelet transform (WT) process

of JPEG2000, penalizing the coding performance. The shape-adaptive version of JPEG2000 (SA-JPEG2000) [59] is designed to overcome this issue. SA-JPEG2000 modifies the spatial WT and bit-plane encoder of JPEG2000 so that only the ROI data is processed, without the need to encode the BG. SA-JPEG2000 also allows for the use of multi-component transforms but requires that the binary mask used to identify the ROI be encoded and transmitted. SA-JPEG2000 may then provide a theoretical optimal coding performance for ROI coding using JPEG2000. In our evaluation results we consider SA-JPEG2000 as the benchmark coding method.

2.2 Experimental results

We performed extensive performance evaluations to verify the accuracy and advantages of our proposed method. In particular, we carried out two different sets of evaluations aimed at assessing: a) the segmentation stage, and b) the coding stage, which includes diagnostically lossless and PLL coding.

Our test data set comprises 60 GXA images of various frames, each frame with a resolution of 1024×1024 pixels of 12 bits of unsigned precision. All the images were routinely acquired at Hospital Mútua de Terrassa, Spain [60], with a Siemens AXIOM-Artis [61] system using organic iodine compounds as the X-ray "contrast agent".

2.2.1 Evaluation of Segmentation accuracy

We first compare our segmentation technique to several edge-based and region-based segmentation methods. In general, edge-based methods use exclusively edge information to identify the ROI, while region-based methods use texture, intensity or statistical features extracted from the image.

In order to detect the boundary using our technique, we cast rays every 0.1 deg (0.0017 rad) and use a window of $w_1 = 10$, $w_2 = 50$ with $t = 0.7$ to compute threshold T. We employ a value of $\alpha = 0.01$ to create disks with radius $r = -1/\alpha$. Note that the ray-casting angle interval and α value are determined based on the size of the

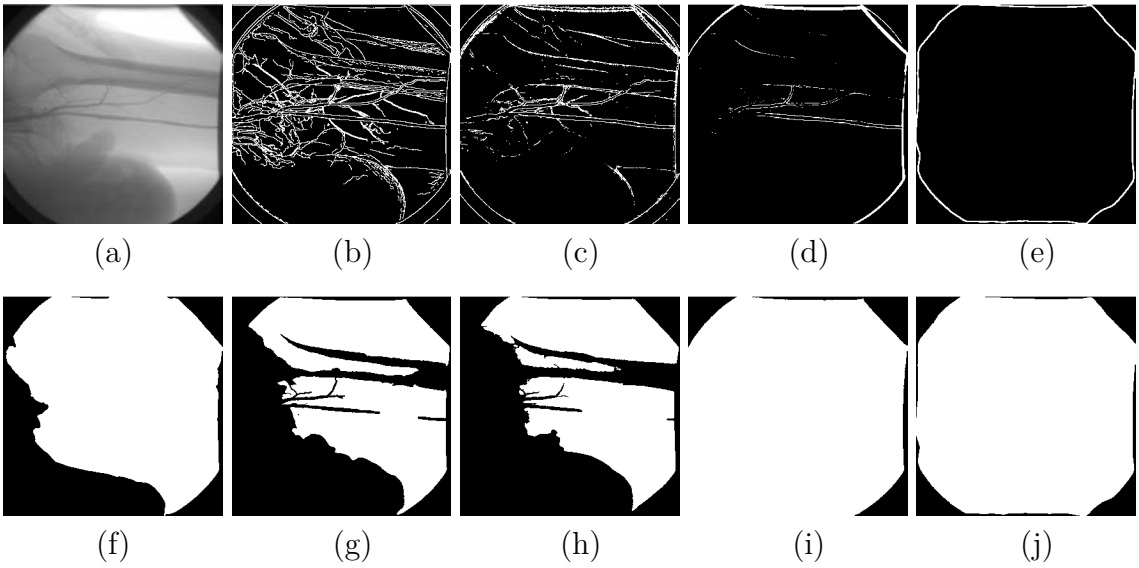


Figure 2.8: (a) Average frame of one GXA image; the corresponding edge detection results by using (b) Canny edge detector, (c) Sobel edge detector, (d) orthogonal projection method, and (e) our proposed boundary detection technique; and the corresponding binary masks computed using (f) Active Contour WE, (g) BC Level Set, (h) Adaptive SRG, (i) MC Watershed and (j) our proposed segmentation technique.

frames and the trade off between boundary refinement accuracy and computational complexity. A small w_1 value is used to find sharp intensity changes, while a large w_2 value is used to increase the probability of locating the symmetrically boundary, as shown in Figure 2.6. Our extensive evaluations indicate that these values result in the best performance for all images in the test data set.

Figure 2.8(a)-(e) show one average frame and its visual segmentation results using several edge-based methods; specifically Canny edge detector [62], Sobel edge detector [63], orthogonal projection [64] and our proposed boundary detection technique. It is important to mention that the average frame in Figure 2.8(a) is one of the most challenging average frames in our test data set. It can be seen that the first three methods detect several edges that are not part of the boundary between the ROI and BG. Moreover, in several cases, they fail to detect all the edges describing the boundary. Our proposed technique detects only the true boundary points, which can be used to generate more accurate binary masks.

Figure 2.8(f)-(j) show the binary masks for the average frame in Figure 2.8(a) computed using state-of-the-art region-based methods. These methods are: Active Contour Without Edges (*Active Contour WE*) [65], Bias Correction Level Set (*BC Level Set*) [66], Adaptive Seeded Region Growing (*Adaptive SRG*) [67], and Marker-Controlled Watershed (*MC Watershed*) [68].

In *Active Contour WE*, the deformation process of the curve does not depend on the gradient of the image as in classical active contour models; instead, it depends on the difference of intensities inside and outside the contour, making these curves less sensitive to noise and the initial curve position. In our experiments, we set the most outside square boundary of I_{avg} as the initial curve. Figure 2.8(f) shows the result of Active contour WE, which fails to correctly detect the ROI boundary in those regions where the intensities between BG and ROI are very similar. *BC Level Set* is a region-based method capable of dealing with intensities inhomogeneities while using the well-known level-set formulation [66, 69] based segmentation process. *Adaptive SRG* combines Otsu's thresholding method [70] and regular SRG [71], avoiding the "trial-and-error" threshold selection of SRG, which is commonly done with human supervision. In our experiments, for BC Level Set the initial curve is the most outside square boundary of I_{avg}, while for Adaptive SRG the initial seeds are a selection of pixels belonging to the four corners of I_{avg}. Figure 2.8(g) and (h) show results for BC Level Set and Adaptive SRG, both methods miss-classify dark bones and tissues areas as being part of the BG. *MC Watershed* is based on watershed transform; it employs predefined background-region marker pixels and foreground-region marker pixels to solve the embedded "over-segmentation" problem of regular watershed methods. Figure 2.8(i) shows the result of MC Watershed. It is important to mention that in our experiments, after an extensive search to define good background-region markers, we were able to segment correctly 27 of the 60 images, which accounts for less than 50% of the images.

For all the methods tested, the corresponding parameters were adjusted according to the values recommended by the authors and according to our evaluations in order to provide the most accurate segmentation results.

With the aim of providing quantitative results, we quantify the segmentation

accuracy of Active Contour WE, BC Level Set, Adaptive SRG, MC Watershed and our proposal by comparing their results to the manual segmentation performed with the help of physicians from Hospital Mútua de Terrassa (Spain), using the following Dice Similarity Coefficient (DSC) [72]:

$$DSC = \frac{2 \times \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (M\_ROI(x,y) \times P\_ROI(x,y))}{\#M\_ROI + \#P\_ROI(x,y)}, \tag{2.1}$$

where M_ROI and P_ROI represent the binary masks detected, respectively, manually and automatically; $\#M\_ROI$ and $\#P\_ROI$ denotes the number of ROI samples in M_ROI and in P_ROI, and X and Y are the number of rows and columns of the image. Note that $DSC \in [0, 1]$, and higher DSC indicates higher similarity between M_ROI and P_ROI, therefore indicates higher segmentation accuracy. The mean and the standard deviation (Std) values of DSC are presented in Table 2.1 for the 60 GXA images. It can be seen from these results that the proposed method has not only the most accurate segmentation results (highest mean DSC), but also the most consistent performance (lowest Std of DSC).

Table 2.1: Segmentation quantitative results. Mean and Std values of the DSC for all 60 images for Active Contour WE, BC Level Set, Adaptive SRG, MC Watershed and the proposed method.

	Active Contour WE	BC Level Set	Adaptive SRG	MC Watershed	Proposed
Mean	0.96	0.87	0.89	0.93	0.99
Std	0.029	0.081	0.12	0.082	0.0025

Figure 2.9 shows the BG-suppressed average frames for twelve different GXA images, with the boundary between ROI and BG enhanced in red. Note that our proposal distinguishes the ROI from the BG with high accuracy. For the rest of the images in the test data set, the results are equivalent.

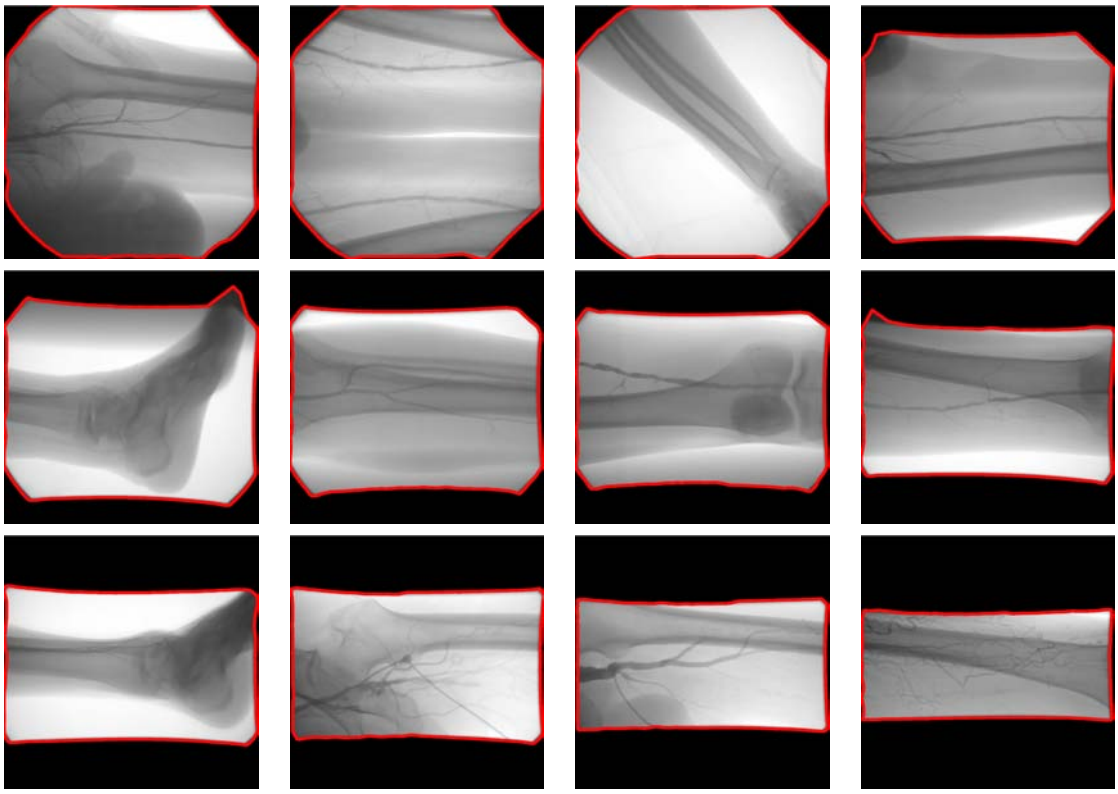


Figure 2.9: Average frames for 12 different GXA images after BG suppression. The boundary (closed contour) between ROI and BG, as detected using α-shapes, is enhanced in red.

2.2.2 Evaluation of Compression performance

We compare several lossless coding methods after applying our segmentation technique to the case of no BG suppression. To better understand the relationship between the amount of BG and the coding performance, the 60 tested images are divided into various subsets according to the amount of BG (in %). Our evaluations include lossless coding and PLL coding. The later is important in interactive telemedicine applications to access and display GXA images over channels of various capacities.

We first compare lossless JPEG2000, JPEG-LS, H.264/AVC and HEVC to the case of coding after BG suppression using our segmentation technique, denoted by BGS-JPEG2000, BGS-JPEG-LS, BGS-H.264/AVC and BGS-HEVC, respectively. In

order to obtain a theoretical optimal rate for JPEG2000, SA-JPEG2000 is applied on the BG-suppressed images. We employ 5 levels of 5/3 reversible spatial WT and codeblocks of size 64x64 for JPEG2000 and SA-JPEG2000, using the BOI [25] software. For JPEG-LS, the reset interval to 64 and the line-interleaved mode for multi-component images are used within the HP implementation of JPEG-LS [28]. For H.264/AVC, the reference software JM 16.2 [73] is used, with FREXt Profile "High 4:4:4" selected for Intra coding and QP and QP Offsets set to 0. For HEVC, the reference software HM 16.2 [74] is used. Three coding modes of HEVC are tested: Intra mode, using the Intra main profile; Random Access (RA) mode, using the Random Access profile and RExt mode, which uses HM 16.2 software with the SCM 3.0 extension and the Random Access main RExt profile. For all these three modes, QP was set to 0 and both TransquantBypassEnableFlag and CUTransquantBypassFlagForce are set to 1, and in RExt mode, CostMode is set to lossless. Note that, in order to comply with the profiles used in H.264/AVC and HEVC, all GXA frames are coded using the colour space YUV 4:4:4 and YUV 4:0:0, respectively.

Table 2.2: Coding performance (in bpp) of JPEG2000, BGS-JPEG2000, SA-JPEG2000, JPEG-LS, BGS-JPEG-LS, H.264/AVC, BGS-H.264/AVC, HEVC (in three modes) and BGS-HEVC (in three modes); bpp savings are reported in % within parenthesis with respect to the case of coding with no BG suppression.

#Images	BG percentage range	JPEG2000	BGS-JPEG2000	SA-JPEG2000	JPEG-LS	BGS-JPEG-LS	H.264/AVC	BGS-H.264/AVC	HEVC Intra	BGS-HEVC Intra	HEVC RA	BGS-HEVC RA	HEVC RExt	BGS-HEVC RExt
23	10% - 20%	6.08	5.39 (11.35%)	5.24 (13.82%)	6.01	5.18 (13.81%)	8.56	7.37 (13.90%)	6.40	5.56 (13.13%)	6.12	5.27 (13.89%)	5.82	5.06 (13.06%)
8	20% - 30%	6.20	4.60 (25.81%)	4.46 (28.06%)	6.17	4.42 (28.36%)	8.90	6.16 (30.79%)	6.51	4.73 (27.34%)	6.29	4.50 (28.46%)	5.95	4.31 (27.56%)
12	30% - 40%	6.24	3.97 (36.38%)	3.84 (38.46%)	6.18	3.78 (38.83%)	9.14	5.19 (43.22%)	6.55	4.08 (37.71%)	6.33	3.86 (39.02%)	5.96	3.68 (38.26%)
13	40% - 50%	6.42	3.57 (44.39%)	3.45 (46.26%)	6.36	3.39 (46.70%)	9.60	4.69 (51.15%)	6.74	3.67 (45.55%)	6.55	3.49 (46.72%)	6.12	3.31 (45.92%)
4	50% - 60%	6.58	2.98 (54.71%)	2.87 (56.38%)	6.51	2.82 (56.68%)	10.35	3.93 (62.03%)	6.87	3.06 (55.46%)	6.71	2.90 (56.78%)	6.21	2.74 (55.88%)
Average		6.23	4.46 (28.41%)	4.31 (30.82%)	6.17	4.27 (30.79%)	9.07	5.96 (34.29%)	6.55	4.58 (30.08%)	6.32	4.34 (31.33%)	5.96	4.15 (30.37%)

Table 2.2 reports the average coding results, in bits per pixel (bpp), for each image subset and for the whole test data set. When no BG suppression is used, the entire image is losslessly encoded. These results indicate that by employing BG suppression the coding performance improves by more than 28%, on average, for all coding methods compared to the case of no BG suppression. H.264/AVC does not achieve as good coding performance as the other coding methods for the GXA images. BGS-HEVC REext attains the best coding performance, followed by BGS-JPEG-LS. BGS-JPEG2000 gets a similar coding performance as BGS-HEVC and BGS-JPEG-LS, while allowing accessing the coded data in a progressive manner. Note that SA-JPEG2000 is, on average, 0.1 bps better than BGS-JPEG2000 even though it requires that the ROI binary mask be encoded and included in the bit-stream. This improvement is mainly due to skipping all of the BG samples during spatial WT and bitplane coding.

As video coding standards H.264/AVC and HEVC are developed with also exploiting the redundancy among frames, we also compare the lossless coding performance when the redundancy among frames is exploited through different multi-component transforms included in Part-2 of JPEG2000 [75], namely Reversible Haar Transform (RHAAR), Reversible Karhunen Loeve Transform (RKLT)[76], 5/3 Reversible Wavelet Transform (RWT) and Differential Pulse Code Modulation (DPCM) [77]. Although JPEG-LS does not include any multi-component transformation, we also introduce the use of a multi-component transform in JPEG-LS to provide a fair comparison. For RHAAR and RWT, the number of decomposition levels along frames is determined by $\min(5, \lfloor \log_2 F \rfloor)$. For RKLT, the side information is encoded with LZMA and included in the final bit-rate. Table 2.3 reports the average coding results for the same image subsets in Table 2.2 when multi-component transforms are employed. It is easy to see that JPEG2000 and JPEG-LS with multi-component transforms get closer or even better coding performance than HEVC, for X-ray angio images. BGS-RKLT-JPEG-LS yields, on average, the best coding performance, closely followed by SA-RKLT-JPEG2000, and both are slightly better than BGS-HEVC REext, while SA-RKLT-JPEG2000 also supports PLL coding.

Table 2.3: Coding performance for multi-component transform using RHAAR, RKLT, RWT and DPCM, followed by JPEG2000 and JPEG-LS. First row reports bitrate (in bpp); second row reports bpp savings (in %) with respect to the case of coding with no BG suppression and no multi-component transform (see Table 2.2)

BG % range	# Frame average	JPEG2000							JPEG-LS				
		BGS-RHAAR	SA-RHAAR	BGS-RKLT	SA-RKLT	BGS-RWT	SA-RWT	BGS-DPCM	SA-DPCM	BGS-RHAAR	BGS-RKLT	BGS-RWT	BGS-DPCM
10 - 20	10	5.05	4.97	4.99	4.92	5.07	4.99	5.25	5.17	4.94	4.89	4.96	5.14
		(16.94%)	(18.26%)	(17.93%)	(19.08%)	(16.61%)	(17.93%)	(13.65%)	(14.97%)	(17.80%)	(18.64%)	(17.47%)	(14.48%)
20 - 30	8	4.36	4.28	4.32	4.24	4.38	4.29	4.51	4.43	4.26	4.23	4.28	4.41
		(29.68%)	(30.97%)	(30.32%)	(31.62%)	(29.35%)	(30.81%)	(27.26%)	(28.55%)	(30.96%)	(31.44%)	(30.63%)	(28.53%)
30 - 40	6	3.74	3.66	3.68	3.61	3.73	3.65	3.82	3.75	3.62	3.57	3.61	3.70
		(40.06%)	(41.35%)	(41.03%)	(42.15%)	(40.22%)	(41.51%)	(38.78%)	(39.90%)	(41.42%)	(42.23%)	(41.59)%	(40.13%)
40 - 50	7	3.41	3.33	3.33	3.26	3.39	3.31	3.45	3.38	3.28	3.22	3.27	3.34
		(46.88%)	(48.13%)	(48.13%)	(49.22%)	(47.20%)	(48.44%)	(46.26%)	(47.35%)	(48.43%)	(49.37%)	(48.58%)	(47.48%)
50 - 60	6	2.80	2.73	2.75	2.68	2.79	2.72	2.84	2.78	2.70	2.66	2.69	2.75
		(57.45%)	(58.51%)	(58.21%)	(59.27%)	(57.60%)	(58.66%)	(56.84%)	(57.75%)	(58.53%)	(59.14%)	(58.68%)	(57.76%)
Average		4.19	4.11	4.13	4.06	4.19	4.11	4.32	4.24	4.08	4.03	4.08	4.20
		(32.74%)	(34.03%)	(33.71%)	(34.83%)	(32.74%)	(34.03%)	(30.66%)	(31.94%)	(33.87%)	(34.68%)	(33.87%)	(31.93%)

For progressive coding, we only compare DICOM-compliant methods that support PLL coding. Figure 2.10 shows the rate-distortion performances for BGS-RKLT-JPEG2000, BGS-RHAAR-JPEG2000, BGS-RWT-JPEG2000, BGS-JPEG2000 and JPEG2000 for three images with various amounts of BG.

The rate-distortion performances are evaluated in terms of the Signal-to-Noise Ratio (SNR), which is defined as $10\log_{10}\frac{\sigma^2}{MSE}$. The mean-squared error (MSE) is computed as $\frac{1}{F}\frac{1}{X}\frac{1}{Y}\sum_f^F\sum_x^X\sum_y^Y(I_f(x,y)-\hat{I}_f(x,y))^2$, where $I_f(x,y)$ and $\hat{I}_f(x,y)$ denote, respectively, the original frame and the recovered frame, and σ^2 denotes the variance of the original image. The distortion gains between JPEG2000 and the BG-suppression strategies vary according to the amount of BG, and are up to 4dB, 10dB and 20dB for images with a BG percentage of 10.40%, 31.31% and 58.97%, respectively. Note that, when multi-component transforms are used, the best results are achieved by BGS-RKLT-JPEG2000, and the distortion gain compared to BGS-JPEG2000 is on average 5dB.

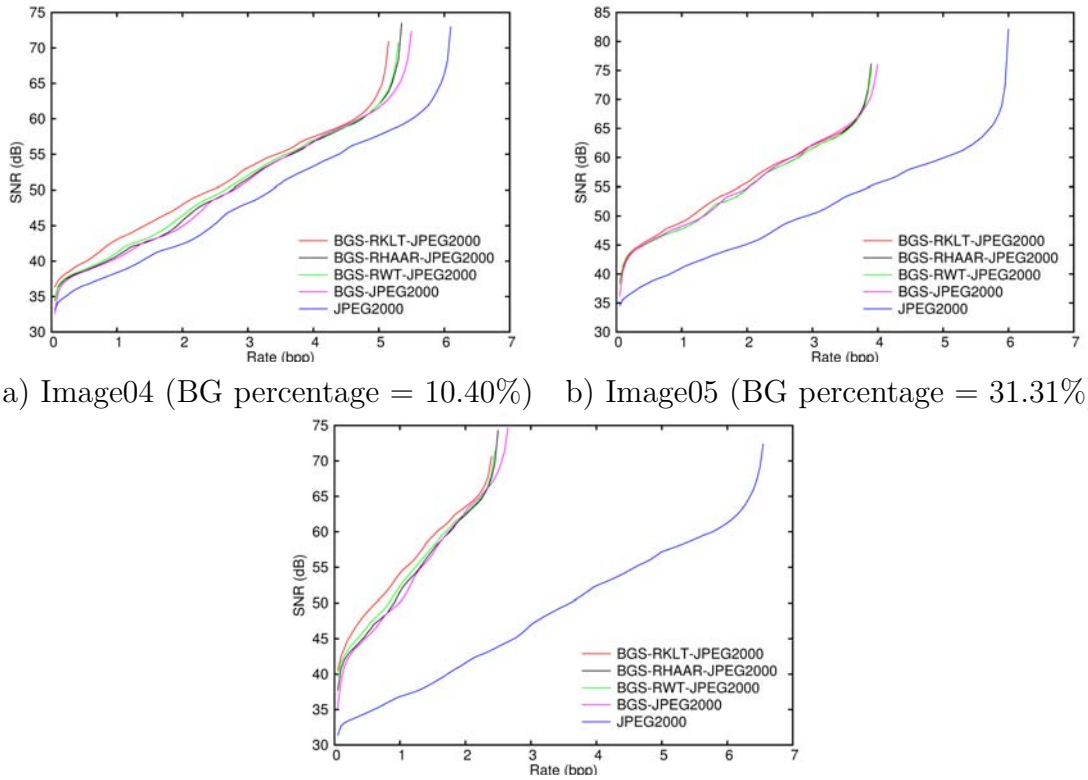
Figure 2.11 depicts a region of two sample frames decoded at 0.01 bpp after JPEG2000 and BGS-RKLT-JPEG2000 PLL coding. It can be observed that the visual quality attained by the latter is better. This is a useful feature that can be exploited in situations where physicians need to access and analyze GXA images in limited bandwidth network environments, e.g., using mobile phones.

To summarize, background suppression helps achieving significant bit-rate savings, and while JPEG-LS with RKLT multi-component is the best lossless coding technique for the tested images, JPEG2000 with RKLT multi-component transform becomes the best alternative when different types of scalability are needed.

2.2.3 Computational complexity

At last, the computational complexity of the proposed strategy is also provided. The segmentation stage has been implemented and run with MATLAB R2012a. The software used in the coding stage is indicated in Section 2.2.2. This experiment ran on an Intel Core i5 CPU 650 (3.20GHz $\times$ 4) with 8Gb.

The segmentation stage of this strategy in average takes 15.68 seconds per frame



a) Image04 (BG percentage = 10.40%) b) Image05 (BG percentage = 31.31%)

c) Image06 (BG percentage = 58.97%)

Figure 2.10: Rate-distortion performance for BGS-RKLT-JPEG2000, BGS-RHAAR-JPEG2000, BGS-RWT-JPEG2000, BGS-JPEG2000 and JPEG2000 for three GXA images with different BG amounts.

for the GXA image, while the coding stage in average takes 0.0639, 0.23, 10.69, 5.31, 35.62 and 37.23 seconds per frame for JPEG2000, JPEG-LS, H.264/AVC, HEVC-Intra, HEVC-RA and HEVC-RA-RExt, respectively. The computing time for the coding stage are computed without using the multi-component transforms. Note that the segmentation stage is usually applied off-line and is not programmed to yield fast real-time execution.

From the results above, the time increment of the segmentation stage amounts 99.59%, 98.55%, 59.46%, 74.70%, 30.57% and 29.64% of the total implementation time when the GXA images are encoded with, respectively, JPEG2000, JPEG-LS, H.264/AVC and HEVC (in three modes).

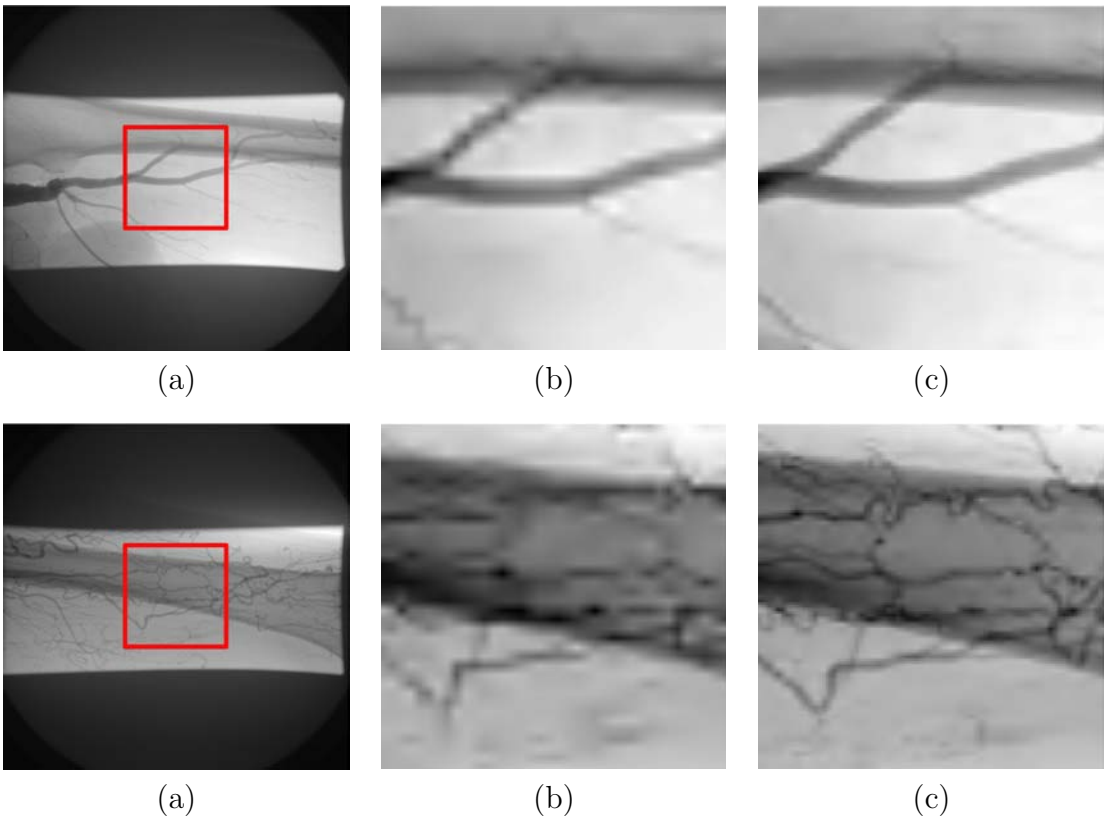


Figure 2.11: (a) Two sample frames decoded at 0.01 bpp after PLL coding using (b) JPEG2000 and (c) BGS-RKLT-JPEG2000. A zoomed-in region of the ROI is presented to show visual differences.

2.3 Chapter Summary

In this chapter, we present a two-staged diagnostically lossless coding method for GXA images. The first stage performs automatic segmentation by employing ray-casting and α-shapes to distinguish the clinically relevant ROI from the BG. The second stage performs lossless or progressive lossy-to-lossless coding on the BG-suppressed images by using JPEG-LS, JPEG2000, H.264/AVC and HEVC.

Experimental results suggest that our segmentation technique identifies the ROI with an average Dice Similarity Coefficient of 0.99 with respect to manual segmentation. When combined with lossless coding methods, our proposed method improves coding performance, on average, by more than 28% and up to 34% compared to the

case of no BG suppression. JPEG-LS technique with multi-component transform has the best coding results, closely followed by JPEG2000 with multi-component transform and HEVC. In addition, evaluations of JPEG2000 with multi-component transform and progressive lossy-to-lossless coding also indicate that, by employing BG suppression, significant improvements on the reconstruction quality of the images may be attained at all bit-rates. Computational complexity evaluation shows that the proposed segmentation process takes at most 99.59% and at least 29.64% of the entire implementation time, depending on the coding techniques used in the coding stage. The increment time from the proposed segmentation process is usually applied off-line, hence it should not affect the whole real-time coding efficiency.

Chapter 3

Frame identification strategy for CAVSs compression

Coronary angiogram video sequences (CAVSs) have become one of the most important instruments for diagnosing cardiovascular diseases. Because of the increasing number of studies using CAVSs and the need to store and share them, the compression of these sequences is becoming critical.

In this chapter, we propose a new diagnostically lossless compression strategy based on frame identification for CAVSs. Our proposal automatically identifies the irrelevant and relevant data in the third dimension of the CAVS, distinguishing the frames according to the "contrast agent" flow motion phases. Once the identification is performed, any compression technique could be applied to the irrelevant and the relevant frames, encoding them, respectively, with and without loss. HEVC [13] and JPEG2000 [10] are the compression techniques evaluated, because JPEG2000 is one of the compression standards included in DICOM, and also because of its competitive coding performance and rich capabilities; and for HEVC, because it is the latest advanced video compression standard.

Experimental results on a large set of CAVSs suggest that the proposed strategy correctly identifies the last irrelevant frame in each CAVS, with an accuracy of 1 frame as compared to manual identification, and achieves bit-stream reductions of more than 19% on average as compared to lossless compression. Assuming that all

the frames in a CAVS must be compressed, the proposed methodology may suppose a computational complexity increment due to the identification process. This increment is proportional to the amount of frames to be processed in the identification stage.

The rest of the chapter is organized as follows: the proposed frame identification and diagnostically lossless compression strategy are described in Section 3.1. Section 3.2 validates our proposal in terms of frame classification and compression performance. Section 3.3 concludes this work.

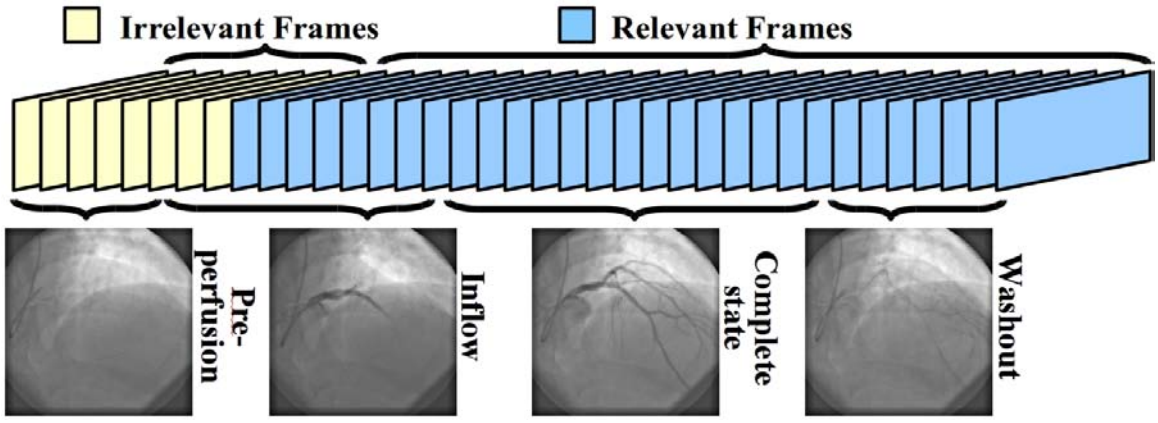


Figure 3.1: A CAVS showing the irrelevant and relevant frames and the 4 phases of the contrast agent flow.

3.1 Proposed Compression Strategy

A CAVS comprises a collection of X-ray projection frames describing the flow of an injected "contrast agent" through the coronary vessel tree over several heart cycles. It is common that CAVS frames are manually labelled pursuant to the four phases of the contrast agent flow motion: *pre-perfusion*, where only the catheter is visible; *inflow*, where the contrast agent is visible in the coronary vessels but it has not yet reached all the vessel branches; *complete state*, where the entire vessel tree is visible; and *washout*, where the perfusion is finished and the contrast agent starts leaving the coronary arteries [78]. In 1997, Prince et al. [79] explored the idea of identifying the

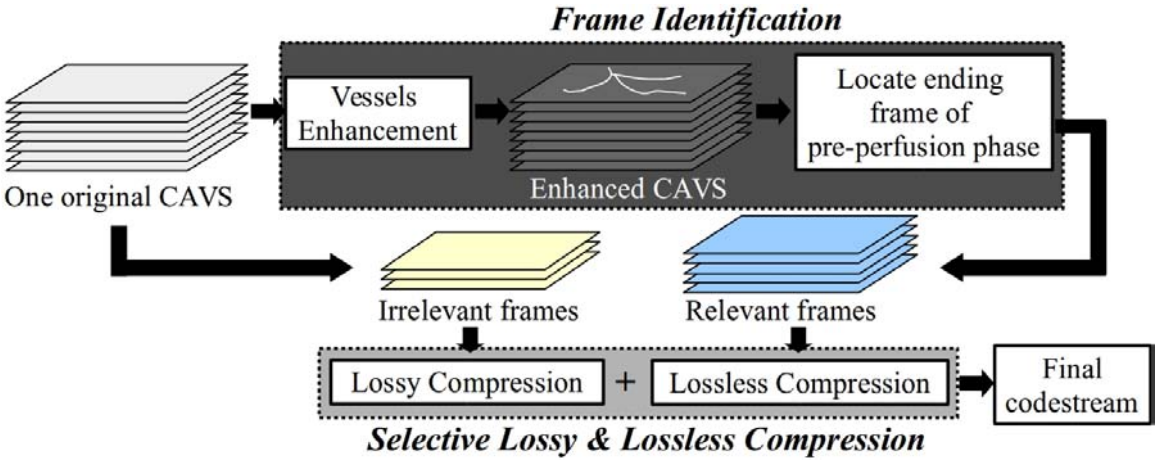


Figure 3.2: Block diagram of the proposed diagnostically lossless compression strategy.

relevant frames during the imaging process of MR angiography evaluating the intensity gradient. Later, a similar idea was investigated for CAVSs in [80], not during the imaging process but during the post-imaging process. In the paper [80], the authors claim that the frames showing the coronary vessel tree (i.e., the frames in the *complete state*) are used by doctors for pre-intervention diagnosis and as a roadmap during the intervention. The pre-perfusion frames are thus not considered diagnostically relevant. The consideration of relevant or irrelevant frames for the diagnosis and to serve as a roadmap during intervention is also undertaken by cardiologists [81]. Figure 3.1 depicts an example of the relation between the four phases and the irrelevant and relevant frames.

The block diagram of the proposed strategy is illustrated in Figure 3.2. The proposal comprises two stages: Frame identification, enhanced in dark grey, and Selective lossy & lossless compression, enhanced in light grey. Both stages are respectively described in detail in Sections 3.1.1 and 3.1.2.

3.1.1 Frame identification

The first stage is aimed to distinguish relevant frames from irrelevant frames. To identify these two types of frames, the last irrelevant frame is used as a delimiter.

This frame can be identified by evaluating the change of vessel structure along all frames. In order to localize this specific frame, this stage includes two steps: 1) *Vessel enhancement* and 2) *Significant vessel structure change detection*.

Vessel enhancement

CAVS frames usually have low contrast and non-uniform tissues background, which makes it difficult to detect the vessels directly. To tackle this issue, some contributions have been presented [82], [80], [83]. [82] uses the eigenvalue analysis of the Hessian matrix and multi-scale filtering approach to enhance the vessels and suppress the background. [80] first employs the morphological top hat filter to equalize the background and then a combination of first- and second-order derivatives to finally enhance the vascular structures. Finally, [83] applies an isotropic undecimated wavelet transform to generate one smooth residual image and several wavelet level images containing the high-frequency data representing details and edges of the original image; then the sum of the second and third wavelet level images is computed as the final vessel enhanced image.

All these three enhancement methods have been tested with our CAVSs. The parameters were adjusted according to the values recommended by the authors and according to our evaluations in order to provide the best enhancement results for the corpora used in this manuscript.

Figure 3.3 depicts an original frame and the same frame after applying the vessel enhancement techniques discussed above. From this figure, it is easy to conclude that [82] attains a higher contrast ratio and a cleaner background than [80] and [83], which makes it simpler to evaluate the changes of vessel structures in each frame. [82] is thus the method used for vessel enhancement in our proposal. Note that, in the following Section 3.2.1, the identification results based on these three enhancement methods are also provided, in order to further prove the advantage of the method [82].

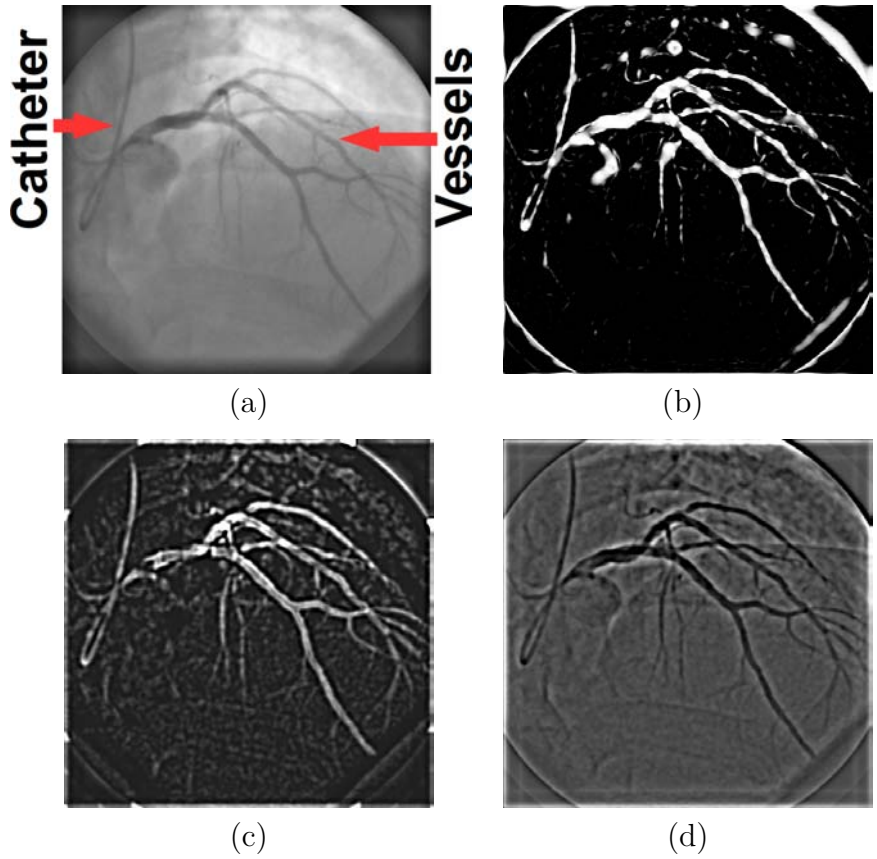


Figure 3.3: Vessel enhancement comparison results for a sample CAVS frame. (a) Original frame, (b) [82], (c) [80], and (d) [83].

Significant vessel structure change detection

After computing the vessel enhancement, the intensity of the sample frames is normalized to the range $[0, 1]$. Pixels belonging to the catheter, the vessels, the muscular tissue and the focal spot edges (those curves close to the corners) have intensities close or equal to 1. Except for the vessels, all the other elements appear in all frames with similar proportions and can be considered constant. Therefore, variations in intensities are attributed to the vessels. This allows tracking the vessel changes, and consequently classifying the frames by using the sum of intensities in each frame. This sum is stored into an array S as follows: $S[n] = \sum_x F_n[x]$, where $F_n[x]$ represents the intensity of pixel x of the nth enhanced frame. Fig. 3.4 plots $S[n]$ for all

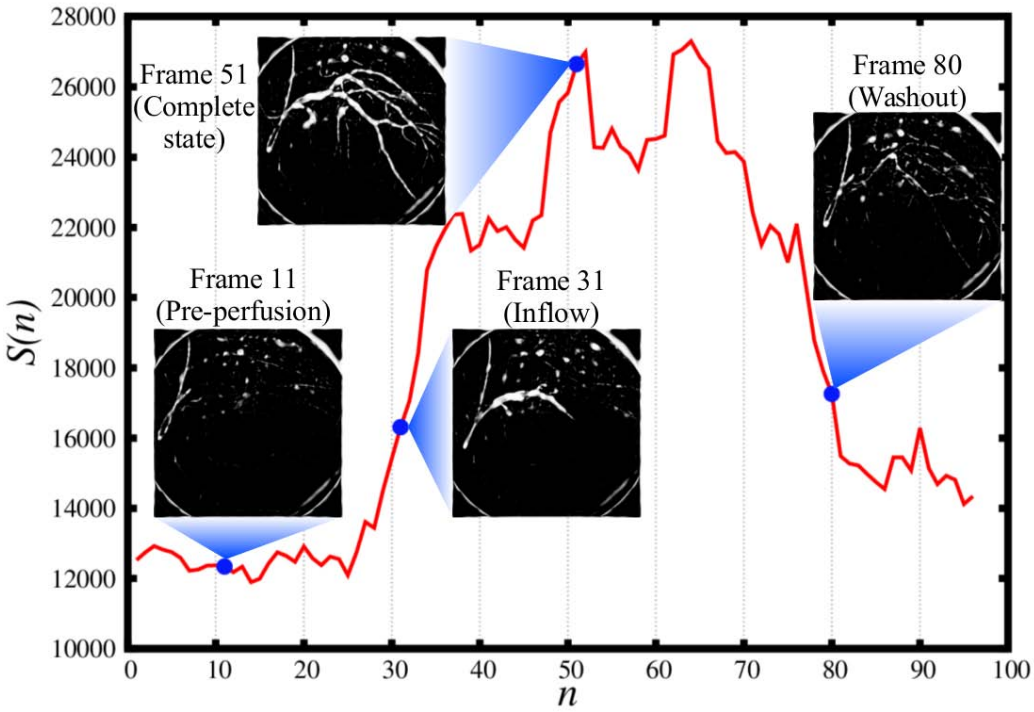


Figure 3.4: Plot of $S[n]$ (sum of pixel intensities) for all enhanced frames of a CAVS. Four enhanced frames corresponding to the four different phases are emphasized in the blue points.

frames of a single CAVS. Different frames corresponding to the different contrast flow motion phases are depicted: frame11 for pre-perfusion, frame31 for inflow, frame51 for complete state, and frame80 for washout. It is worth noting that there exists a relationship between $S[n]$ and the four phases. Let us remind that relevant frames start roughly when the inflow phase occurs, which corresponds to the first sharp rise in the plot of $S[n]$.

The following operations are used to detect the first sharp rise in $S[n]$:

1. An order 5 one-dimensional median filter is first applied to S to remove small irregular changes, yielding S'.

2. An array of rise-increments, denoted as IN, is then computed using a sliding window whose width $\mathcal{W}$ is adapted to the length of S', i.e., $\mathcal{W}$=$round(V/w)$,

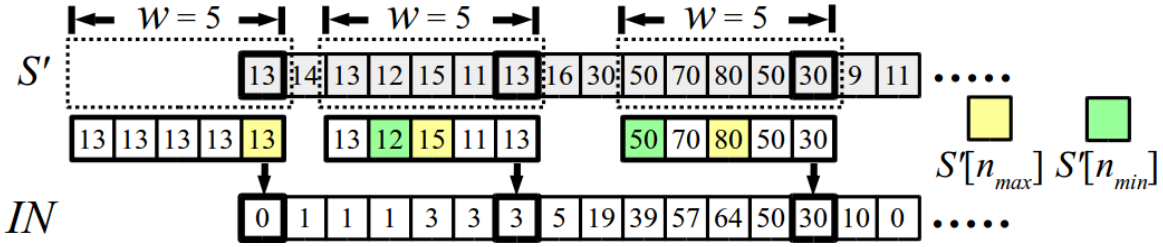


Figure 3.5: Example of an S' and its corresponding intensity-rise array IN computed with a sliding window of width $\mathcal{W} = 5$.

$w > 0$, where V is the length of S'. In each sliding window, the rise-increment of the last frame, $IN[n]$, is computed as follows:

(a) $S'[n_{max}]$, the frame with the largest sum of intensities in that sliding window is found;

(b) $S'[n_{min}]$, the frame with the lowest sum of intensities and to the left of $S'[n_{max}]$ in that sliding window is found;

(c) $IN[n] = S'[n_{max}] - S'[n_{min}]$;

(d) the sliding window slides one position at a time to the right.

3. Let $\max_{IN}$ be the largest rise-increment in the array S' and let threshold T be $T{=}t{\times}\max_{IN}$, with $0{\leq}t{\leq}1$. Then, n' is the position of the first frame with $IN[n'] > T$, which corresponds to the first sharp rise in S'. n' is thus identified as the last irrelevant frame in the CAVS.

The procedure described in step 2 above is graphically illustrated in Fig. 3.5, with $S'[n_{max}]$ and $S'[n_{min}]$ enhanced in yellow and green, for some example sliding windows.

In our proposal, parameters w and t are set by users. The values for these parameters are dealt with in section 3.2.

3.1.2 Selective lossy & lossless compression

As shown in Fig. 3.2, after the frames in CAVSs have been successfully identified, any compression technique could be applied to the irrelevant and relevant frames, coding them with and without loss, respectively. HEVC and JPEG2000 are employed as both yield competitive coding performance for video compression and both support lossy and lossless coding. And it is worth noting that JPEG2000 as this coding standard offers a richer set of coding features than any other lossless coding method, including resolution, quality scalability and compatibility with DICOM.

3.2 Experimental results

Our proposal is evaluated through two different sets of experiments: *A) Frame Identification Accuracy* and *B) Selective Lossy & Lossless Compression Performance*. For both experiments, 72 X-ray CAVSs of various numbers of frames have been used. Each frame has a resolution of 512 × 512 pixels with 12 bits of unsigned precision. All of them were routinely acquired at Hospital Mútua de Terrassa, with a Siemens AXIOM-Artis [61] system using organic iodine compounds as the X-ray "contrast agent". The cardiologists at Hospital Mútua de Terrassa have identified the irrelevant frames manually; the number of irrelevant frames is referred to as n''. Fig. 3.6 depicts the number of total frames (solid green) for each CAVS, the index of the last irrelevant frame (solid blue) identified manually, and the proportion of these pre-perfusion frames (solid yellow) in each CAVS.

3.2.1 Identification error

To appraise the accuracy of the identification stage in detecting the last irrelevant frame, an identification error is computed as E=n'-n'', where n' and n'' are the frame indexes identified automatically and manually, respectively. Proportion of the identification error in each CAVS is also provided as $P = \frac{(n'-n'')\times 100}{\mathrm{N}}$, where N denotes the total number of frames of a CAVS. Fig. 3.7 depicts E and P for all 72 test sequences. The parameters used during the frame identification procedure are set to w=15 and

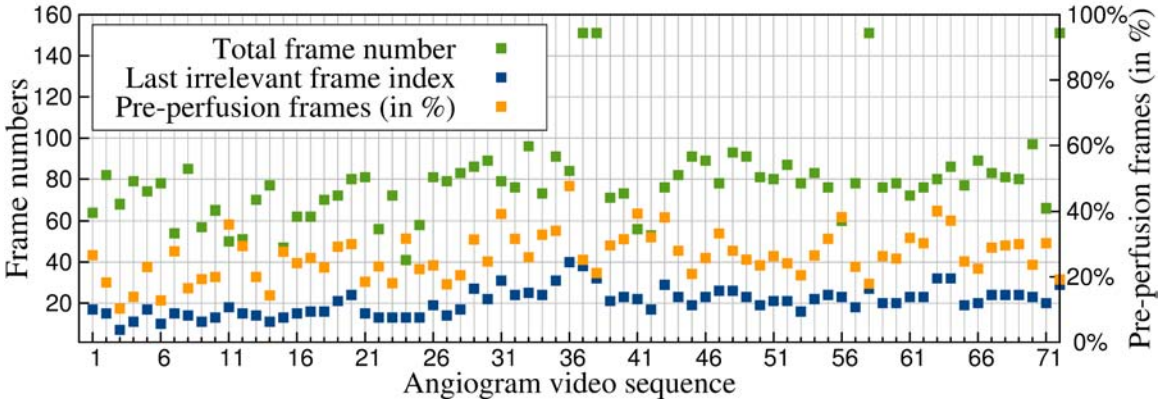


Figure 3.6: The number of frames for the 72 CAVSs used, the index of the last irrelevant frame (n'') obtained by manual frame-classification and the proportion of these pre-perfusion frames in each CAVS.

t=0.5. Note that for most of the CAVSs we tested, changing this parameter setting in a certain range does not alter the accuracy of the frame identification results too much; and the parameters we are using are found to provide the most accurate classification results compared to manual identification. By using the adaptive window size, we achieve better identification results than the case of setting a fix window size (e.g., $\mathcal{W} = 5$, which is the average value of the adaptive window sizes we used). From these results, we can see that our automatic identification procedure identifies the last irrelevant frame with high accuracy. The average of the absolute identification error is below 1 frame, with a maximum error of 4 frames, and the percentage of the absolute identification error is about 1.18%. Notice that when E (or P)is positive, relevant data is encoded with loss; while for negative error values, irrelevant data is encoded losslessly. However, it is worth noting that, according to the cardiologists, these small errors are negligible during the diagnosis, since a small error during the identification is insignificant compared to the number of frames in each CAVS.

The percentage of the identification error with using [80] and [83] are also computed and depicted together with the ones of [82] in Figure 3.8. The parameters $w = 15$ and $t = 0.5$ are used for all the three enhancement methods. From Figure 3.8, we can see that [82] obtains the highest accuracy, followed by [83] and then [80]. That

also proves our selection of using [82] as the enhancement method (even it is the oldest one), as using [80] and [83], the accuracy in "significant vessel structure change" detection to correctly identify the relevant frames is penalized.

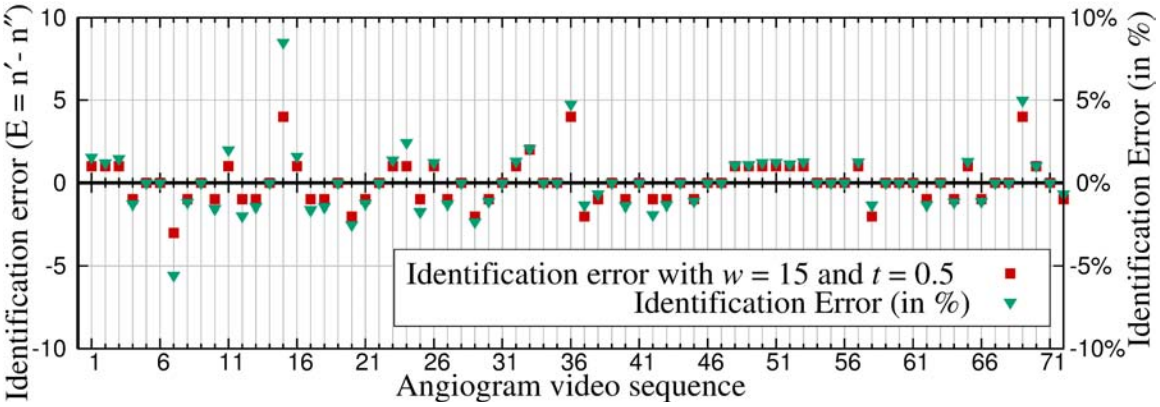


Figure 3.7: Identification error (solid red) of automatic versus manual identification of the first irrelevant frame, and the proportion of this error (solid green) with respect to the frame number in each CAVS.

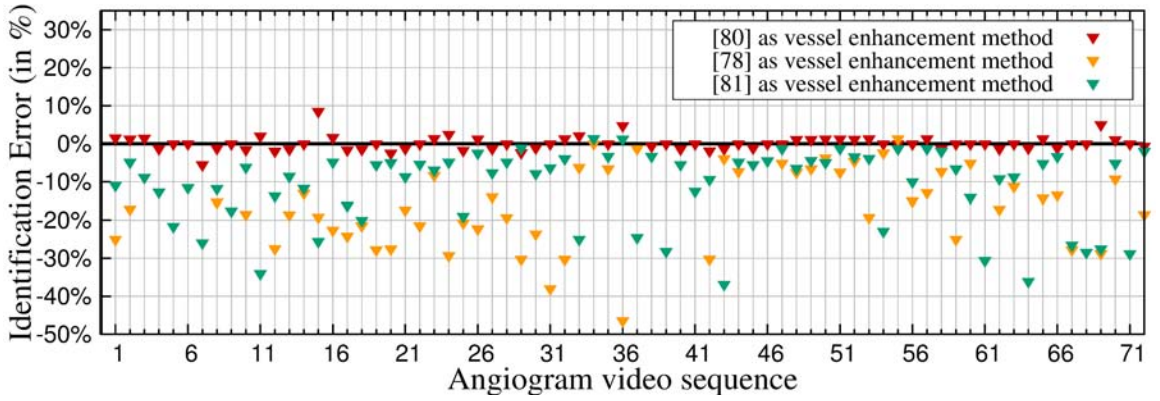


Figure 3.8: Identification error in % using [82], [80] and [83] as enhancement vessel methods.

3.2.2 Compression performance

This second set of experiments evaluates the compression performance of *full-lossless*, that encodes the irrelevant frames without loss; *selective lossy & lossless*, which encodes the irrelevant frames in a lossy mode. In addition we have included results combining [46] (BGS) with full-lossless, and [46] with selective lossy & lossless, named BGS+full-lossless and BGS+selective lossy & lossless, respectively. Note that, in all strategies the relevant frames are all encoded without loss.

The coding techniques compared are the current standards for image and video coding, JPEG2000 and HEVC. For lossless JPEG2000, the parameters are: 5 levels of 5/3 reversible spatial wavelet transform, single quality layer, and codeblocks of size 64x64 with no precincts. For lossy JPEG2000, the parameters are the same except for the use of the 9/7 irreversible spatial wavelet transform and Qstep=1/4096[1]. JPEG2000 experiments have been performed with Kakadu v7.4 [24] implementation. Three coding modes of HEVC are tested: Intra mode (HEVC-Intra), using the Intra main profile; Random Access mode (HEVC-RA), using the Random Access profile and RExt mode (HEVC-RA-RExt). For all these three modes, in lossless compression, QP was set to 0 and both TransquantBypassEnableFlag and CUTransquantBypassFlagForce are set to 1, and in RExt mode, CostMode is set to lossless; in lossy compression, these parameters are set back as default but QP was kept as 0 to achieve the slightest loss. Note that, in order to comply with the profiles used in HEVC, all CAVS frames are coded using the colour space YUV 4:0:0. HEVC experiments have been computed with the reference implementation [74]. All the compression rates are given in bits per sample (bps).

Figure 3.9 depicts the average compression results achieved for full-lossless, selective lossy & lossless, BGS+full-lossless and BGS+selective lossy & lossless, when JPEG2000 and the three different configurations of HEVC are used. For all coding techniques, considering the full-lossless as the reference strategy, the selective lossy & lossless achieves improvements between 0.75 to above 2 bps (which depends on the proportion of pre-perfusion frames), in average, this improvement reaches 1.25 bps;

[1]Qstep is the step size defined during the quantization process in JPEG2000. A large step size translates into more loss.

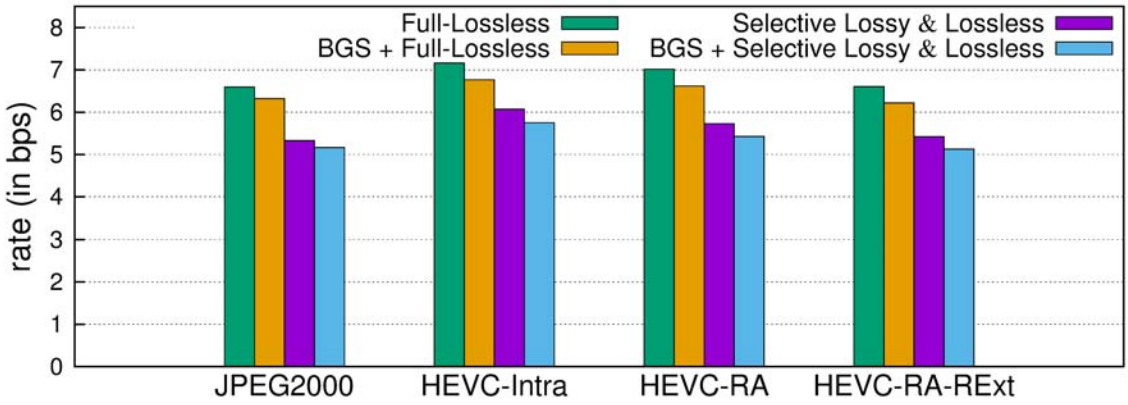


Figure 3.9: Average compression results for the 72 CAVSs, using full-lossless, selective lossy & lossless, BGS+full-lossless, and BGS+selective lossy & lossless compression strategies, with JPEG2000 and HEVC (three modes) techniques.

when BGS is combined with selective lossy & lossless, the results are still improved but with smaller benefits.

In addition, we assess the image quality in terms of Signal-to-Noise Ratio (SNR) and Structural Similarity Index Measure (SSIM) [84], for the irrelevant frames for lossy-lossless when JPEG2000 and HEVC-RA-RExt are employed. Figure 3.10 illustrates the image quality results for JPEG2000 and HEVC-RA-RExt results for all 72 CAVSs. The average SNR of the irrelevant frames is about, 47.80 dB and 50.48 dB, respectively; while for SSIM it is 0.9970 and 0.9982, respectively. The team of cardiologists at Hospital Mútua de Terrassa [81] has visually evaluated the irrelevant decompressed frames compressed with JPEG2000 and HEVC-RA-RExt. In both cases, cardiologists have indicated that visual differences are not appreciable. For instance, Figure 3.11 depicts a zoomed area of an irrelevant frame of an original image, a decompressed frame after employing the lossy-lossless for JPEG2000 and HEVC-RA-RExt [2].

[2]These three CAVS frames in full resolution can be downloaded from `http://www.gici.uab.cat/GiciApps/Frame_Visual_Evaluation.tar.gz` for visual evaluation.

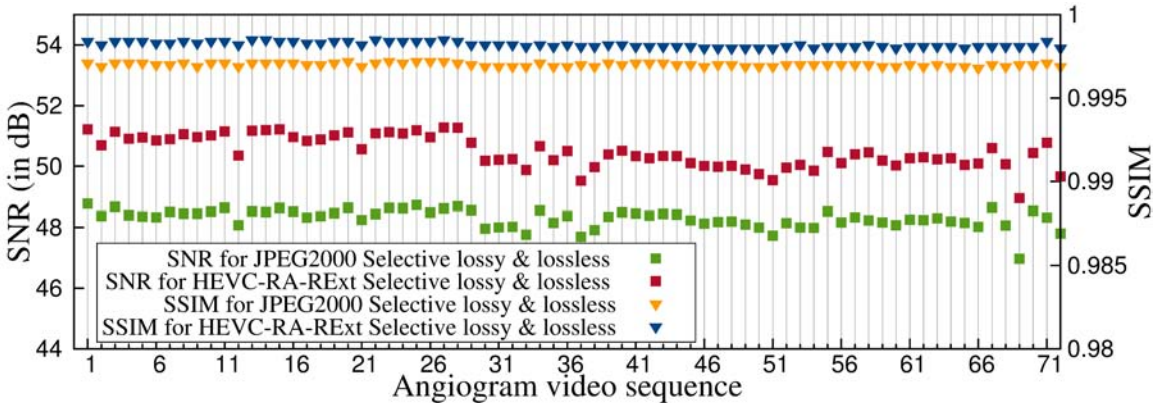


Figure 3.10: Quality of the the irrelevant frames. Results are reported for the 72 CAVSs using lossy-lossless for JPEG2000 and HEVC-RA-RExt.

3.2.3 Computational complexity

Finally, we evaluated the impact of the identification stage in terms of computational time. The software and hardware environments are similar to the ones used in Section 2.2.3, i.e., MATLAB R2012a is used for frame identification; the compression software are indicated in Section 3.2.2; and all of the experiments ran on an Intel Core i5 CPU 650 (3.20GHz $\times$ 4) with 8Gb.

The segmentation stage of the proposal in Chapter 2 takes 0.073 seconds per frame for the CAVS, while the methodology presented here needs 17.58 seconds per frame, about 240 times slower. This increment is mainly produced by the vessel enhancement technique, for which we used [82]. Note that this technique is usually applied off-line and is not programmed to yield fast real-time execution.

The consumed time for the compression stage varies depending on the technique, needing 0.013, 1.76, 10.69 and 11.26 seconds per frame for JPEG2000, HEVC-Intra, HEVC-RA and HEVC-RA-RExt, respectively. All the computational times are given on average for the whole corpus.

In summary, the identification stage in our current proposal would amount to 99.92%, 90.89%, 62.18% and 60.95% of the total execution time when the images are encoded with, respectively, JPEG2000, HEVC-Intra, HEVC-RA and HEVC-RA-

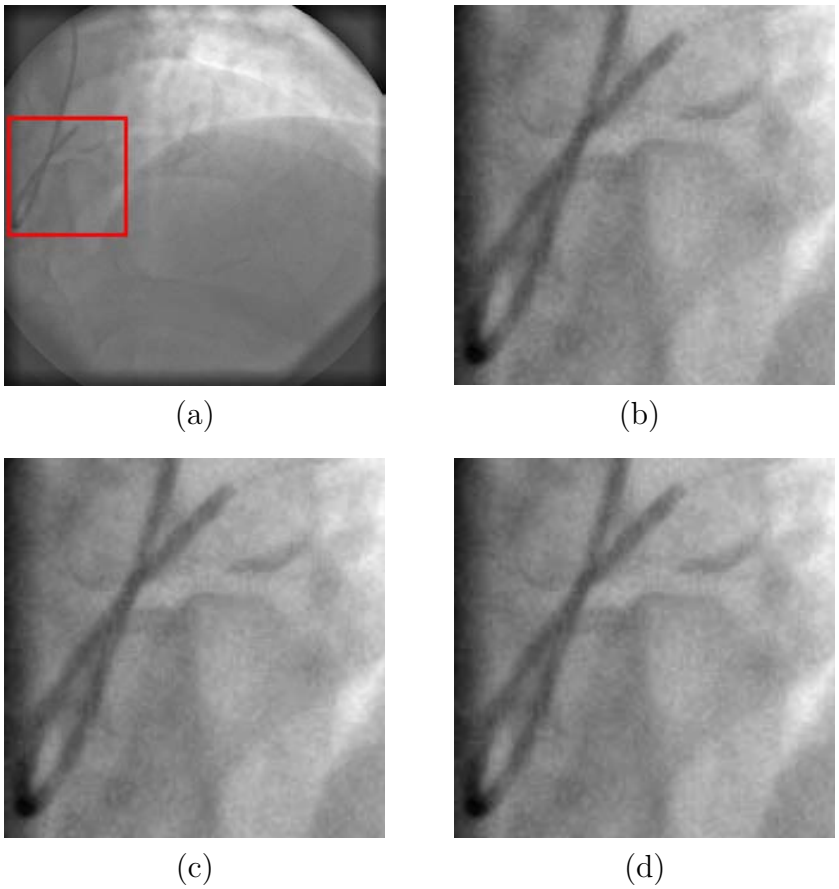


Figure 3.11: (a) One original pre-perfusion frame, (b) a zoomed area from (a), (c)-(d) the zoomed area after selective lossy & lossless compression with JPEG2000 and HEVC-RA-RExt, respectively.

RExt.

3.3 Chapter Summary

Coronary angiogram video sequences (CAVSs) are used in medical centers for Cardiovascular Diseases diagnosis, the number one cause of death globally. These sequences need be stored for medical records and shared for remote telediagnosis. Cardiologists manually identify the clinically irrelevant frames in these CAVSs, which correspond to the initial frames where only the catheter is visible or the contrast agent introduced

into the vessels is imperceptible.

We proposed an automatic method to identify the last irrelevant frame, avoiding the need for the cardiologist manual classification. We then further proposed a diagnostically lossless compression approach, where the irrelevant frames are encoded with loss and the relevant frames without loss. Our proposal employs JPEG2000 and HEVC as the compression techniques. Experimental results suggest that the frame identification correctly distinguishes the diagnostically irrelevant from the relevant frames with high accuracy, and improves the compression performance, on average, by more than 19% and 12% compared to lossless JPEG2000 and HEVC compression, respectively. Computational complexity evaluation indicates the identification stage of the current proposal is slower than the segmentation methodology of Chapter 2. However, this technique is usually applied off-line, thus does not affect the efficiency of the following compression stage. At last, all the experimental results have been supervised and validated by the cardiologist unit of Hospital Mútua de Terrassa.

Chapter 4

Conclusions

4.1 Summary

Medical Image Compression

Because of the global trend of ageing population, vascular system diseases have become the main risk for human health today. Angiography using X-ray imaging and catheterisation with particular "contrast agents" remain the gold standard for the diagnosis of vascular system diseases. For distinct medical examination purposes, there are two main types of X-ray angiography images: Coronary Angiogram Video sequences (CAVSs) and General X-ray Angiography (GXA) images. The former one only targets to detect the irregularities present in coronary vessel trees, while the latter one serves for observing the blood vessel diseases of the other body parts like feet, legs, arms, etc.

Considering the increasing amount of X-ray angiography images generated and the large file sizes of the images (the maximum file size of one single CAVS or GXA image may reach 140MB), the Picture Archiving and Communications Systems (PACS) associated with the Digital Imaging and Communications in Medicine (DICOM) standard are used in hospitals and medical centers to handle this medical data. Compression is an invaluable tool to reduce the data size and hence benefit the storage and transmission of medical images.

Medical image compression methods could be classified into three types: 1) lossy compression methods, which achieve high compression ratios but introduce distortion to the diagnostically sensitive data and thus may affect the subsequent diagnosis accuracy; 2) lossless compression methods, which enable a perfect reconstruction of the decoded image and usually have a lower data-reduction capability than lossy methods; and 3) diagnostically lossless compression methods, developed to avoid the shortages of both lossy and lossless methods, which obtain higher compression ratios than lossless methods while guaranteeing the perfect reconstruction of those regions in the image that are used for diagnostic purposes. In this thesis, we introduce two diagnostically lossless compression methods for the two different main types of X-ray angiography images, GXA images and CAVSs.

Clinically relevant data identification

Diagnostically lossless compression improves the compression performance through the exploitation of data redundancy that exists in clinically irrelevant areas of the images. Lossy compression is applied to the clinically irrelevant areas and lossless compression to the diagnostically relevant areas. Due to this basic concept, the process of distinguishing the clinically relevant and irrelevant data in the images plays a pivotal role in diagnostically lossless compression methods.

Based on the different features of GXA images and CAVSs, we developed two approaches to accomplish the clinically relevant data identification: 1) Background-suppression, which uses segmentation methods to separate the Region of Interest (ROI) and the background (BG) areas in each 2-dimensional GXA frame; and 2) Frame-identification, which recognizes the clinically irrelevant and relevant frames in the third dimension of CAVSs.

Advanced compression techniques

The two proposed diagnostically lossless compression strategies are both comprised of two stages: the clinically relevant data identification stage and the compression stage. These two stages are implemented independently, i.e., different compression

techniques could be employed in the compression stage, without affecting the performance of the previous identification performance.

Several compression formats that are accepted in DICOM are tested in our proposals, including JPEG2000, JPEG-LS and H.264/AVC. The latest video coding standard HEVC, which has been shown to attain important compression gains compared to previous standards, is also applied in the proposals.

Experimental results

For GXA images, after separating the ROI and the background areas, the data in BG areas are suppressed to improve the compression performances; and for CAVSs, after the diagnostically relevant and irrelevant frames are identified, a selective lossy & lossless compression method is used to achieve compression improvements. In both strategies, several compression techniques are employed, including the DICOM-compliant JPEG2000, JPEG-LS, H.264/AVC, and the latest video coding standard HEVC.

Experimental results suggest that both strategies are automatic and accurate in diagnostically relevant data identification (the segmentation process of BG suppression strategy gets an average Dice Similarity Coefficient of 0.99 with respect to manual segmentation; and the percentage of the absolute identification error in frame identification strategy is around 1.18% on average), and efficient in data compression (for GXA images, BG suppression strategy improves the compression performance, on average, more than 28% and up to 34% compared to the case of no BG suppression; and for CAVSs, the frame identification strategy improves the compression performance, on average, by more than 19% and 12% compared to lossless JPEG2000 and HEVC compression, respectively). Both strategies were developed in cooperation with the physicians from Hospital Mútua de Terrassa, who have supervised and validated both the identification and the compression results.

4.2 Future Work

Blood vessels and their neighbourhood areas of X-ray angiography images contain the most important information for diagnosis purpose. The future work may focus on the diagnostically lossless compression strategies based on the vessel segmentation. Following this idea, several works may be done:

1) Search for an accurate vessel segmentation method. One method under our consideration now is [85]. Compared to the Frangi method [82], [85] obtains more precise Hessian eigenvalue analysis in noisy environment and detect smaller and thinner vessels by using a directional filter bank.

2) Develop a hierarchical compression strategy. Based on the vessel segmentation results, a three levels compression strategy may be applied, i.e., background suppression for the most peripheral areas, lossy compression for the tissue areas and lossless compression for the blood vessels and their neighbourhood areas. And regarding the compression techniques, JPEG2000, JPEG-LS, H.264/AVC and HEVC could be employed. For JPEG2000 and JPEG-LS, multi-component transform may also be combined into the compression to further improve the compression performance.

3) Another idea of the hierarchical compression strategy may be background suppression + visual lossless modelling + lossless compression, i.e., different from the method in 2), we could use visual lossless modelling to deal with the tissue areas, which could give a higher visual quality of these areas, compared to directly lossy compression.

4) A distortion metric could be developed to evaluate the compression performance of the hierarchical compression strategy.

Another main work could be the computational complexity optimization. After assessing the computing time of our proposal, the time devoted to the identification stage normally amounts to more than 50% of the whole compression time. One way to reduce the computing time would be to replace the use of MATLAB platform with

other efficient programming languages, e.g., Java and Python. Yet another choice could be running the implementation in GPU, which is under early research in [86].

Appendix A

Acronyms

ACR American College of Radiology

BGS Background Suppression

CAVS Coronary Angiogram Video Squence

CT Computed Tomography

DCT Discrete Cosine Transform

DICOM Digital Imaging and Communications in Medicine

DPCM Differential Pulse Code Modulation

DSC Dice Similarity Coefficient

DWT Discrete Wavelet Transform

EBCOT Embedded Block Coding with Optimized Truncation of the embedded bit-streams

GXA General X-ray Angiography

HEVC High Efficiency Video Coding

ISO International Organization for Standardization

ITU International Telecommunication Union

LOCO-I LOw COmplexity LOssless COmpression for Images

MSE Mean-Squared Error

NEMA National Electrical Manufacturers Association

PACS Picture Archiving and Communications Systems

PLL Progressive Lossy-to-Lossless

RHAAR Reversible Haar

RKLT Reversible Karhunen Loeve Transform

ROI Region of Interest

RWT Reversible Wavelet Transform

SNR Signal-to-Noise Ratio

SSIM Structural Similarity Index Measure

WT Wavelet Transform

Bibliography

[1] H. Kobayashi, M. Ogawa, R. Alford, P. L. Choyke, and Y. Urano, "New strategies for fluorescent probe design in medical diagnostic imaging," *Chemical reviews*, vol. 110, no. 5, pp. 2620–2640, May 2010.

[2] A. Fouras, M. Kitchen, S. Dubsky, R. Lewis, S. Hooper, and K. Hourigan, "The past, present, and future of X-ray technology for in vivo imaging of function and form," *Journal of Applied Physics*, vol. 105, no. 10, p. 102009, May 2009.

[3] K. L. McCance and S. E. Huether, *Pathophysiology: The biologic basis for disease in adults and children (7th)*. 3251 Riverport Lane, St. Louis, Missouri: Mosby Elsevier, 2014, ch. "Alterations of Cardiovascular Function", pp. 1129–1193.

[4] D. R. Coles, M. A. Smail, I. S. Negus, P. Wilde, M. Oberhoff, K. R. Karsch, and A. Baumbach, "Comparison of radiation doses from multislice computed tomography coronary angiography and conventional diagnostic angiography," *Journal of the American College of Cardiology*, vol. 47, no. 9, pp. 1840–1845, May 2006.

[5] I. Moussa, A. Hermann, J. C. Messenger, G. J. Dehmer, W. D. Weaver, J. S. Rumsfeld, and F. A. Masoudi, "The NCDR CathPCI Registry: a US national perspective on care and outcomes for percutaneous coronary intervention," *Heart (British Cardiac Society)*, vol. 99, no. 5, pp. 297–303, Mar 2013.

[6] N. Archer, U. Fevrier-Thomas, C. Lokker, K. A. McKibbon, and S. Straus, "Personal health records: a scoping review," *Journal of the American Medical Informatics Association*, vol. 18, no. 4, pp. 515–522, Jul 2011.

[7] J. Taquet and C. Labit, "Hierarchical Oriented Predictions for Resolution Scalable Lossless and Near-Lossless Compression of CT and MRI Biomedical Images," *Image Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 2641–2652, May 2012.

[8] Digital Image and Communication in Medicine, "DICOM," May 2015. [Online]. Available: http://medical.nema.org/

[9] R. Choplin, "Picture archiving and communication systems: an overview," *Radiographics*, vol. 12, no. 1, pp. 127–129, Jan 1992.

[10] *Information technology - JPEG 2000 image coding system - Part 1: Core coding system*, ISO/IEC Std. 15 444-1, Dec 2000.

[11] *Information technology - Lossless and near-lossless compression of continuous-tone still images - Baseline*, ISO/IEC Std. 14 495-1, Jun 1999.

[12] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, Jul 2003.

[13] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[14] Digital Image and Communication in Medicine, "DICOM Section: C.7.6.11 Display Shutter Module," May 2015. [Online]. Available: http://dicom.nema.org/medical/dicom/current/output/chtml/part03/sect_C.7.6.11.html

[15] *Digital compressing and coding for continuous-tone still images*, ISO/IEC Std. 10 918-1, 1994.

[16] S. Golomb, "Run-length encodings (corresp.)," *Information Theory, IEEE Transactions on*, vol. 12, no. 3, pp. 399–401, Jul 1966.

[17] *Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Video*, ISO/IEC Std. 13 818-2, Nov 1994.

[18] Joint Video Team of ITU-T and ISO/IEC JTC 1, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)," Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, Mar 2003.

[19] G. J. Whitman and T. M. Haygood, *Digital Mammography: A Practical Approach*. Shaftesbury Road, Cambridge: Cambridge University Press, 2012, ch. "PACS, storagy and archiving", pp. 32–45.

[20] R. Spencer, "How far can Lossy compression be utilised throughout digital imaging before diagnostic quality is compromised?" *Shadows: The New Zealand Journal of Medical Radiation Technology*, vol. 55, no. 1, pp. 5–12, Mar 2012.

[21] D. Taubman, "High performance scalable image compression with EBCOT," *Image Processing, IEEE transactions on*, vol. 9, no. 7, pp. 1158–1170, Jul 2000.

[22] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: an overview," *Consumer Electronics, IEEE Transactions on*, vol. 46, no. 4, pp. 1103–1127, Nov 2000.

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

[24] D. Taubman, "Kakadu software," May 2015. [Online]. Available: http://www.kakadusoftware.com/

[25] Group on Interactive Coding of Images, "BOI software," May 2015. [Online]. Available: http://www.gici.uab.es/BOI

[26] M. J. Weinberger, G. Seroussi, and G. Sapiro, "LOCO-I: A low complexity, context-based, lossless image compression algorithm," in *Data Compression Conference, 1996. DCC'96. Proceedings.* IEEE, Mar 1996, pp. 140–149.

[27] ——, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS," *Image Processing, IEEE Transactions on*, vol. 9, no. 8, pp. 1309–1324, Aug 2000.

[28] Hewlett Packard, "HP Labs LOCO-I/JPEG-LS," 2015. [Online]. Available: http://www.hpl.hp.com/research/info_theory/loco

[29] A. Robert, I. Amonou, and B. P esquet Popescu, "Improving H.264 video coding through block oriented transforms," in *Multimedia and Expo, 2008 IEEE International Conference on*, Jun 2008, pp. 705–708.

[30] Joint Video Team (JVT), "Joint model (JM) 18.6, H.264/AVC reference software," May 2015. [Online]. Available: http://iphome.hhi.de/suehring/tml/

[31] M. T. Pourazad, C. Doutre, M. Azimi, and P. Nasiopoulos, "HEVC: the new gold standard for video compression: how does HEVC compare with H.264/AVC?" *Consumer Electronics Magazine, IEEE*, vol. 1, no. 3, pp. 36–46, Jul 2012.

[32] Joint Collaborative Team on Video Coding (JCT-VC), "HEVC reference model (HM) 16.4," May 2015. [Online]. Available: https://hevc.hhi.fraunhofer.de/

[33] 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: improving compressibility by increasing data redundancy outside the body region," *Radiology*, vol. 259, no. 1, pp. 271–277, Apr 2011.

[34] S. M. Hosseini and A.-R. Naghsh-Nilchi, "Medical ultrasound image compression using contextual vector quantization," *Computers in biology and medicine*, vol. 42, no. 7, pp. 743–750, Jul 2012.

[35] J. Muñoz-Gómez, J. Bartrina-Rapesta, M. W. Marcellin, and J. Serra-Sagristà, "Influence of Noise Filtering in Coding Computed Tomography with JPEG2000," in *Data Compression Conference (DCC), 2011.* IEEE, Mar 2011, pp. 413–422.

[36] ——, "Correlation Modeling for Compression of Computed Tomography Images," *Biomedical and Health Informatics, IEEE Journal of*, vol. 17, no. 5, pp. 928–935, Sep 2013.

[37] M. Hernández-Cabronero, J. Munoz-Gómez, I. Blanes, M. W. Marcellin, and J. Serra-Sagristà, "DNA microarray image coding," in *Data Compression Conference (DCC), 2012*. IEEE, Apr 2012, pp. 32–41.

[38] M. Hernandez-Cabronero, V. Sanchez, M. W. Marcellin, and J. Serra-Sagristà, "A distortion metric for the lossy compression of DNA microarray images," in *Data Compression Conference (DCC), 2013*. IEEE, Mar 2013, pp. 171–180.

[39] V. Sanchez and J. Bartrina-Rapesta, "Lossless compression of medical images based on HEVC intra coding," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, May 2014, pp. 6622–6626.

[40] V. Sanchez, F. Auli-Llinas, J. Bartrina-Rapesta, and J. Serra-Sagristà, "HEVC-based lossless compression of Whole Slide pathology images," in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*. IEEE, Dec 2014, pp. 297–301.

[41] V. Sanchez, F. Auli-Llinas, R. Vanam, and J. Bartrina-Rapesta, "Rate Control for Lossless Region of Interest Coding in HEVC Intra-Coding with Applications to Digital Pathology Images," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, Apr 2015, in Press.

[42] A. Ouled Zaid and B. B. Fradj, "Coronary angiogram video compression for remote browsing and archiving applications," *Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 632–641, Dec 2010.

[43] G. Uma Vetri Selvi and R. Nadarajan, "Coronary angiogram video compression using wavelet-based contourlet transform and region-of-interest technique," *Image Processing, IET*, vol. 6, no. 8, pp. 1049–1056, Nov 2012.

[44] D. Gibson, M. Spann, and S. I. Woolley, "A wavelet-based region of interest encoder for the compression of angiogram video sequences," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 8, no. 2, pp. 103–113, Jun 2004.

[45] Z. S. Yazdi, N. Karimi, S. Samavi, and S. Shirani, "Perceptually lossless compression of angiogram sequences," in *Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on.* IEEE, Nov 2010, pp. 1–4.

[46] Z. Xu, J. Bartrina-Rapesta, V. Sanchez, J. Serra-Sagristà, and J. Munoz-Gómez, "Diagnostically lossless compression of x-ray angiography images based on automatic segmentation using ray-casting and α-shapes," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, Sep 2013, pp. 738–742.

[47] Z. Xu, J. Bartrina-Rapesta, I. Blanes, V. Sanchez, J. Serra-Sagristà, M. García-Bach, and J. F. Muñoz, "Diagnostically Lossless Coding of X-ray Angiography Images based on Background Suppression," *Computers & Electrical Engineering*, 2015 (submitted).

[48] Z. Xu, J. Bartrina-Rapesta, J. Serra-Sagristà, V. Sanchez, and J. F. Muñoz, "Diagnostically Lossless Compression of Video Coronary Angiograms based on Frame Identification," *Signal Processing Letters, IEEE*, 2015 (submitted).

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

[50] A. Giachetti, L. Ballerini, E. Trucco, and P. J. Wilson, "The use of radial symmetry to localize retinal landmarks," *Computerized Medical Imaging and Graphics*, vol. 37, no. 5, pp. 369–376, Jul-Sep 2013.

[51] K. I. Jassam, "Removal of random noise from conventional digital X-ray images," in *Proc. International Society for Photogrammetry and Remote Sensing (ISPRS)*, vol. XXIX, no. B5, Aug 1992, pp. 113–118.

[52] R. C. Gonzalez and R. E. Woods, *Digital Image Processing, 3rd Edition*. Upper Saddle River, NJ, USA: Prentice-Hall, 2008, ch. "Intensity Transformations and Spatial Filtering", pp. 126–220.

[53] S. Mitra and J. Kaiser, *Digital Signal Processing Handbook*. John Wiley & Sons, 1993.

[54] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[55] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080–2095, Aug 2007.

[56] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 7, pp. 629–639, Jul 1990.

[57] P. Dodin, J. Martel-Pelletier, J.-P. Pelletier, and F. Abram, "A fully automated human knee 3D MRI bone segmentation using the ray casting technique," *Medical & biological engineering & computing*, vol. 49, no. 12, pp. 1413–1424, Dec 2011.

[58] S. Lou, X. Jiang, and P. J. Scott, "Fast algorithm for morphological filters," in *Journal of Physics: Conference Series*, vol. 311, no. 1. IOP Publishing, 2011, p. 012001.

[59] A. Zabala, J. González-Conejero, J. Serra-Sagristà, and X. Pons, "JPEG2000 encoding of images with NODATA regions for remote sensing applications," *SPIE Journal of Applied Remote Sensing*, vol. 4, no. 1, pp. 041 793–041 793, Jul 2010.

[60] Hospital Mutua de Terrassa, Terrassa, Spain, May 2015. [Online]. Available: http://www.mutuaterrassa.cat/

[61] Siemens, "Siemens medical imaging," May 2015. [Online]. Available: http://www.healthcare.siemens.com/

[62] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 679–698, Nov 1986.

[63] R. C. Gonzalez and R. E. Woods, *Digital Image Processing, 3rd Edition.* Upper Saddle River, NJ, USA: Prentice-Hall, 2008, ch. "Image Segmentation", pp. 711–816.

[64] A. Mejias, S. Romero, and F. Moreno, "A new algorithm to extract the lines and edges through orthogonal projections," *Digital Signal Processing*, vol. 22, no. 1, pp. 147–152, Jan 2012.

[65] T. Chan and L. Vese, "Active contours without edges," *Image Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 266–277, Feb 2001.

[66] C. Li, R. Huang, Z. Ding, J. Gatenby, D. Metaxas, and J. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," *Image Processing, IEEE Transactions on*, vol. 20, no. 7, pp. 2007–2016, Jul 2011.

[67] T. Berber, A. Alpkocak, P. Balci, and O. Dicle, "Breast mass contour segmentation algorithm in digital mammograms," *Computer Methods and Programs in Biomedicine*, vol. 110, no. 2, pp. 150–159, May 2013.

[68] L. Xu and H. Lu, "Automatic morphological measurement of the quantum dots based on marker-controlled watershed algorithm," *NaNotechnology, IEEE Transactions on*, vol. 12, no. 1, pp. 51–56, Jan 2013.

[69] D. X. Ji, S. H. Ong, and K. W. C. Foong, "A level-set based approach for anterior teeth segmentation in cone beam computed tomography images," *Computers in Biology and Medicine*, vol. 50, no. 0, pp. 116–128, Jul 2014.

[70] N. Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, no. 1, pp. 62–66, Jan 1979.

[71] R. Adams and L. Bischof, "Seeded region growing," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 6, pp. 641–647, Jun 1994.

[72] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul 1945.

[73] Joint Video Team (JVT), "Joint model (JM) 16.2, H.264/AVC reference software," May 2015. [Online]. Available: http://iphome.hhi.de/suehring/tml/

[74] Joint Collaborative Team on Video Coding (JCT-VC), "HEVC reference model (HM) 16.2," May 2015. [Online]. Available: https://hevc.hhi.fraunhofer.de/

[75] *Information technology - JPEG 2000 image coding system - Part 2: Extensions*, ISO/IEC Std. 15 444-2, Apr 2004.

[76] K. Karhunen, "Zur spektraltheorie stochastischer prozesse," *Annales Academiae Scientiarum Fennicae*, vol. 37, 1946.

[77] C. C. Cutler, "Differential quantization of communication signals," U.S. Patent 2 605 361, Jul 1952.

[78] A. P. Condurache, T. Aach, A. Kaiser, and P. Radke, "User-defined ROI tracking for estimation of the myocardial blush grade," in *Image Analysis and Interpretation, 2006 IEEE Southwest Symposium on*, 2006, pp. 66–70.

[79] M. R. Prince, T. L. Chenevert, T. K. Foo, F. J. Londy, J. S. Ward, and J. H. Maki, "Contrast-enhanced abdominal MR angiography: optimization of imaging delay time by automating the detection of contrast material arrival in the aorta," *Radiology*, vol. 203, no. 1, pp. 109–114, Apr 1997.

[80] T. Aach, A. P. Condurache, K. Eck, and J. Bredno, "Statistical-model-based identification of complete vessel-tree frames in coronary angiograms," in *Electronic Imaging 2004*. International Society for Optics and Photonics, 2004, pp. 283–294.

[81] Private Communication, Cardiologist Unit, Hospital Mútua de Terrassa, Terrassa, Spain, Sep 2014.

[82] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Interventation–MICCAI'98*, ser. Lecture Notes in Computer Science. Springer, 1998, vol. 1496, pp. 130–137.

[83] P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, "Fast retinal vessel detection and measurement using wavelets and edge location refinement," *PLoS ONE*, vol. 7, no. 3, p. e32435, Mar 2012.

[84] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, Apr 2004.

[85] P. T. Truc, M. A. Khan, Y.-K. Lee, S. Lee, and T.-S. Kim, "Vessel enhancement filter using directional filter bank," *Computer Vision and Image Understanding*, vol. 113, no. 1, pp. 101–112, Jan 2009.

[86] P. Enfedaque, F. Auli-Llinas, and J. C. Moure, "Implementation of the DWT in a GPU through a Register-based Strategy," *Parallel and Distributed Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, Dec 2014.