DISSERTATION

Master in Electrical and Electronic Engineering

Contributions to HEVC Prediction for Medical Image Compression

André Filipe Rodrigues Guarda

Leiria, March 2016

DISSERTATION

Master in Electrical and Electronic Engineering

Contributions to HEVC Prediction for Medical Image Compression

André Filipe Rodrigues Guarda

Master dissertation performed under the guidance of Doctor Nuno Miguel Morais Rodrigues, Professor of Escola Superior de Tecnologia e Gestão of Instituto Politécnico de Leiria and with the co-orientation of Doctor Sérgio Manuel Maciel Faria, Professor of Escola Superior de Tecnologia e Gestão of Instituto Politécnico de Leiria.

Leiria, March 2016

You miss 100% of the shots you don't take.
(Wayne Gretzky)

Acknowledgments

I want to thank my advisers, Dr. Nuno Rodrigues and Dr. Sérgio Faria, for their guidance, and without whom this work would not be possible. Thank you for your patience and for recruiting me to this research group.

I would also like to express my gratitude to Instituto de Telecomunicações and Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria for providing me great working conditions and resources.

To the members of UDICMI and MEDICOMP projects, namely Dr. Luís Cruz and Dr. Pedro Assunção, for their input which allowed to improve this work.

To my colleagues at the Multimedia Signal Processing group from the Leiria delegation of Instituto de Telecomunicações, namely, Filipe Gama, Gilberto Jorge, João Carreira, João Santos, Lino Ferreira, Luís Lucas, Ricardo Monteiro and Sylvain Marcelino, for their friendship. A special thanks to Gilberto Jorge and Ricardo Monteiro for their contributions related to geometrical transformations in this work, to João Carreira and Luís Lucas, who helped me use Linux and understand the HEVC code, and provided tools that made my work easier and more enjoyable, and to João Santos for putting up with me.

To my friends who accompanied me throughout these five and a half years, in particular, David Cruz, Hélder Simões, José Ricardo, Miguel Rasteiro, Paulo Cardoso, Ricardo Santos and Richard Marciano, for allowing me to enjoy both hard-work and leisure moments.

And to my family, for their unconditional support.

This work was a part of the UDICMI project funded by CENTRO-07- ST24-FEDER-002022 of QREN and project Medicomp in the scope of R&D Unit 50008, financed by the applicable financial framework (FCT/MEC through national funds and when applicable co-funded by FEDER – PT2020 partnership agreement).

Abstract

Medical imaging technology and applications are continuously evolving, dealing with images of increasing spatial and temporal resolutions, which allow easier and more accurate medical diagnosis. However, this increase in resolution demands a growing amount of data to be stored and transmitted. Despite the high coding efficiency achieved by the most recent image and video coding standards in lossy compression, they are not well suited for quality-critical medical image compression where either near-lossless or lossless coding is required.

In this dissertation, two different approaches to improve lossless coding of volumetric medical images, such as Magnetic Resonance and Computed Tomography, were studied and implemented using the latest standard High Efficiency Video Encoder (HEVC). In a first approach, the use of geometric transformations to perform inter-slice prediction was investigated.

For the second approach, a pixel-wise prediction technique, based on Least-Squares prediction, that exploits inter-slice redundancy was proposed to extend the current HEVC lossless tools. Experimental results show a bitrate reduction between 45% and 49%, when compared with DICOM recommended encoders, and 13.7% when compared with standard HEVC.

Keywords: HEVC, Lossless Compression, Medical Imaging, Geometric Transformations, Least-Squares Prediction.

Resumo

A tecnologia e as aplicações de imagens médicas estão em constante evolução, lidando com imagens de maiores resoluções espaciais e temporais, o que permite um diagnóstico mais fácil e preciso. No entanto, este aumento de resolução implica uma quantidade crescente de dados que têm que ser armazenados e transmitidos. Apesar da alta eficiência de codificação alcançada pelas mais recentes normas de codificação de imagem e de vídeo, estas não estão bem adaptados para a compressão de imagens médicas, onde é necessária codificação sem perdas ou quase sem perdas.

Nesta tese, foram estudadas e implementadas duas abordagens diferentes para melhorar a codificação sem perdas de imagens médicas volumétricas, como Ressonância Magnética e Tomografia Computorizada, utilizando a mais recente norma *High Efficiency Video Encoder* (HEVC). Na primeira abordagem, foram desenvolvidos alguns métodos que utilizam transformações geométricas para realizar predição entre imagens.

Para a segunda abordagem, foi proposta uma técnica de predição pixel-a-pixel, com base em Mínimos Quadrados, que explora a redundância entre imagens, de modo a estender as atuais ferramentas de compressão sem perdas do HEVC. Os resultados experimentais mostram uma redução da taxa de bits entre 45% e 49%, quando comparado com os codificadores recomendados pela norma DICOM, e 13,7%, quando comparado com a norma HEVC.

Palavras-chave: HEVC, Compressão Sem Perdas, Imagens Médicas, Transformações Geométricas, Predição baseada em Mínimos Quadrados.

Contents

List of Figures

List of Tables

List of Abbreviations

2D	Two-Dimensional
2D-LMP	2D Linearised Median Predictor
3D	Three-Dimensional
3D-DLMP	3D Distances-based Linearised Median Predictor
4D	Four-Dimensional
AAC	Adaptive Arithmetic Coder
AVC	Advanced Video Coding
BMA	Block Matching Algorithm
BMGT	Block Matching with Geometric Transformations
bpp	bits per pixel
CABAC	Context-Adaptive Binary Arithmetic Coding
CALIC	Context based Adaptive Lossless Image Codec
CIPR	Center for Image Processing
CT	Computed Tomography
CTU	Coding Tree Unit
CU	Coding Unit
DCT	Discrete Cosine Transform
DICOM	Digital Imaging and Communications in Medicine
DoF	Degrees of Freedom
DPCM	Differential Pulse-Code Modulation
DTM	Directional Template Matching
EBCOT	Embedded Block Coding with Optimal Truncation
GAP	Gradient Adaptive Predictor
GOP	Group of Pictures
GT	Geometric Transformations
HEVC	High Efficiency Video Coding
JPEG	Joint Photographic Experts Group
LOCO-I	Low Complexity Lossless Compression for Images
LSP	Least-Squares Prediction

MCT	Multi-Component Transform
MED	Median Edge Detector
MFV	Most Frequent Value
MILC	Medical Images Lossless Compression
MMP	Multi-scale Multidimensional Parser
MR	Magnetic Resonance
MRP	Minimum Rate Predictors
MV	Motion Vector
PCM	Pulse Code Modulation
PU	Prediction Unit
R.A.	Random Access
RExt	Range Extension
RLE	Run Length Encoding
SAD	Sum of Absolute Differences
SAP	Sample Adaptive intra-Prediction
SAP-G	Sample-based Angular intra-Prediction with Gradient-based
SAP-ME	Sample-based Angular intra-Prediction with Median and Edge
SWP	Sample-based Weighted Prediction

Chapter 1

Introduction

1.1 Context and Motivation

Recent advances in digital imaging and video, as well as rapid evolution of acquisition and processing systems using increasingly higher resolutions, demand for efficient compression formats for file exchange, storage and visual communications over networks. High Efficiency Video Coding (HEVC) is the most recent standard to fulfil the challenging requirements imposed by new services and applications [1], but in the case of medical imaging systems, with the growing use of telemedicine and information sharing among the medical community, efficient image compression assumes a different meaning due to more challenging requirements [2].

Medical imaging systems are following the same trend of higher spatial and temporal resolutions and bit depths. Therefore, a simple medical examination, e.g. a Magnetic Resonance (MR) or a Computed Tomography (CT), generates a very large amount of data. Typically these data must be stored for several years, with resulting high costs in creating and maintaining medical image record databases. Since the use of lossy compression algorithms can result in the loss of important details in medical images, which can lead to more difficult analysis or even wrong diagnosis, it is required to use lossless or near-lossless compression techniques in this type of applications. The preference for lossless techniques is acknowledged by the Digital Imaging and Communications in Medicine (DICOM) [3], which recommends the use of lossless encoders such as Joint Photographic Experts Group (JPEG) in lossless mode, Run Length Encoding (RLE), TIFF PackBits, JPEG-LS [4], lossless JPEG2000 [5], H.264/Advanced Video Coding (AVC) [6], as well as lossy JPEG, JPEG-LS and JPEG2000 [5] with near-lossless compression ratios for this purpose.

JPEG2000 is a wavelet transform-based image encoder [5], while JPEG-LS uses the Me-

dian Edge Detector (MED) [4] predictor. MED is a simple edge detector that uses three raster-scan order causal pixels to determine if a vertical or horizontal edge exists and then selects one of three predictors accordingly. Other state-of-the-art lossless image encoders are Context based Adaptive Lossless Image Codec (CALIC) [7] and Minimum Rate Predictors (MRP) [8]. CALIC uses a Gradient Adaptive Predictor (GAP), in which six causal pixels are used to estimate the local gradients and determine the prediction value. MRP uses several linear predictors, which are adapted to each image to be encoded.

Volumetric images, such as MR images or CT scans, comprise a set of static images, each one representing a slice of a relevant volume for medical purposes. Although each slice can be encoded independently using a standard image encoder, such option does not exploit the inter-slice redundancy present in volumetric image sets. In order to address this issue, an extension to the JPEG2000 encoder was developed for volumetric images, known as JP3D [9], which uses 3D blocks and a 3D discrete wavelet transform, instead of their Two-Dimensional (2D) counterparts.

Besides developing specific techniques to encode multi-slice medical volumes, it is possible to regard each slice as analogous to a video frame, thus allowing to use state-of-the-art video encoders like HEVC [1]. Despite this type of encoders having been developed and fine tuned for lossy video compression, they can also be used to encode visual data volumes, using the inter-frame prediction tools to exploit the inter-slice correlation.

In this work, an approach to the encoding of these volumetric medical images using geometrical transformations in HEVC was tested. Different methods for motion estimation and compensation between slices were implemented, from image-wise motion estimation, to block matching with geometric transformations, and even a mesh-based technique, in order to improve the prediction and ultimately the compression efficiency. However, this approach did not yield positive results, which led to the focus of this dissertation being shifted into pixel-wise prediction.

Least-Squares Prediction (LSP), introduced in [10], has been successfully used in previous works for image and stereo image compression, due to its ability to represent well the image edges orientations [11–13]. As such, an LSP-based method is proposed to enhance HEVC for lossless coding of volumetric medical images. LSP was modified to exploit inter-slice redundancy in order to improve its prediction efficiency.

1.2 Objectives

The objectives of this work consist firstly in the study of state-of-the-art lossless encoders and encoding techniques for volumetric medical images. Then, the following topics were

developed:

- Implementation of methods for medical images processing to improve their compression;
- Study and development of geometric transformations based techniques;
- Integration of these techniques for motion estimation and compensation in HEVC;
- Performance evaluation of the techniques.

Given the performance of the methods based on geometric transformations, new objectives were proposed for this work:

- Study of pixel-wise prediction methods for lossless encoding;
- Development and implementation of techniques to improve lossless prediction in HEVC;
- Performance evaluation of these techniques, and comparison against state-of-the-art encoders.

1.3 Dissertation Structure

This work is organized as follows: Chapter 2 gives an overview of medical images, describes state-of-the-art lossless encoders used in this work and presents a comparison of their performance. Chapter 3 describes the implementation of the methods for processing the medical images and improve their encoding efficiency, and presents the obtained results. Chapter 4 explains geometric transformations and their applications, details the implemented approaches in HEVC and their respective experimental results. In Chapter 5 the proposed LSP method and its integration in HEVC are described, along with its performance. Finally, in Section 6 some conclusions are drawn.

In Appendix A, contributions associated with this dissertation are presented.

Chapter 2

Medical Image Compression

In this chapter the current state-of-the-art on medical images compression is presented. Firstly, a brief summary on medical imaging is given, followed by a description of the state-of-the-art lossless encoders used in this work. Finally, a review of previous work on medical image compression present in the literature concludes this chapter.

2.1 Medical Imaging

Medical imaging emerged with the increasing understanding of several physical phenomena such as X-rays, γ-rays, ultrasound waves, among others. Later on, the growing use of computers, the development of acquisition procedures and the usage of different physical processes, such as image reconstruction in tomography, allowed a further development of medical imaging systems.

2.1.1 Medical Imaging Types

Each type of medical image has its own characteristics [14] and is used for different purposes. Some can be used to view anatomical structures in the body, showing clearly detailed organs or bones, while others allow to show and trace bodily functions or chemical processes, such as blood flowing [15]. Figure 2.1 shows an example of the human brain acquired with different medical imaging types. Due to their popularity, in this work two of the most common volumetric medical imaging types, Magnetic Resonance (MR) imaging and X-ray Computed Tomography (CT), are used.

CT [16] scanners have been available since mid-1970s and were a revolution in medical imaging. Although CT scans use X-rays, they are able to produce images far more

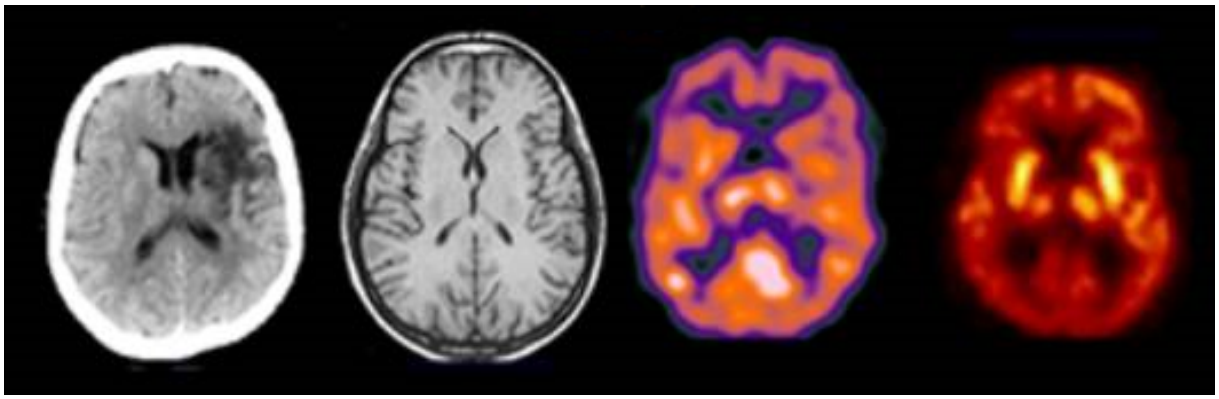


Figure 2.1: Transaxial slice of the human brain acquired with different imaging modalities. From left to right: X-ray Computed Tomography, Magnetic Resonance Imaging, Single-Photon Emission Computed Tomography and Positron Emission Tomography [15].

detailed than conventional X-rays, with the advantage of the scanning speed, since it can be completed in seconds. Despite being capable to show various body parts, such as heart and lungs, the biggest advantage of the CT image is its ability to represent bones, being able to detail even the smallest ones. The use of contrast agents allows the further assessment of blood vessels and organs.

The MR [17] makes use of powerful magnetic fields and radio-frequency pulses, in order to map the internal body structures, with the advantage of this technology not exposing the users to X-rays. MR is mainly used to produce high quality images of soft tissue and anatomic structures, such as grey and white matter in the brain. MR scanners can be used for a wide range of body parts including injuries of the joints, blood vessels, breasts, as well as abdominal and pelvic organs, such as the liver or reproductive organs. Many diseases, such as brain tumours, can be visualized using this type of images because of the high contrast definition, which does not always require contrast agents to produce detailed images of blood vessels.

2.1.2 Tests Dataset

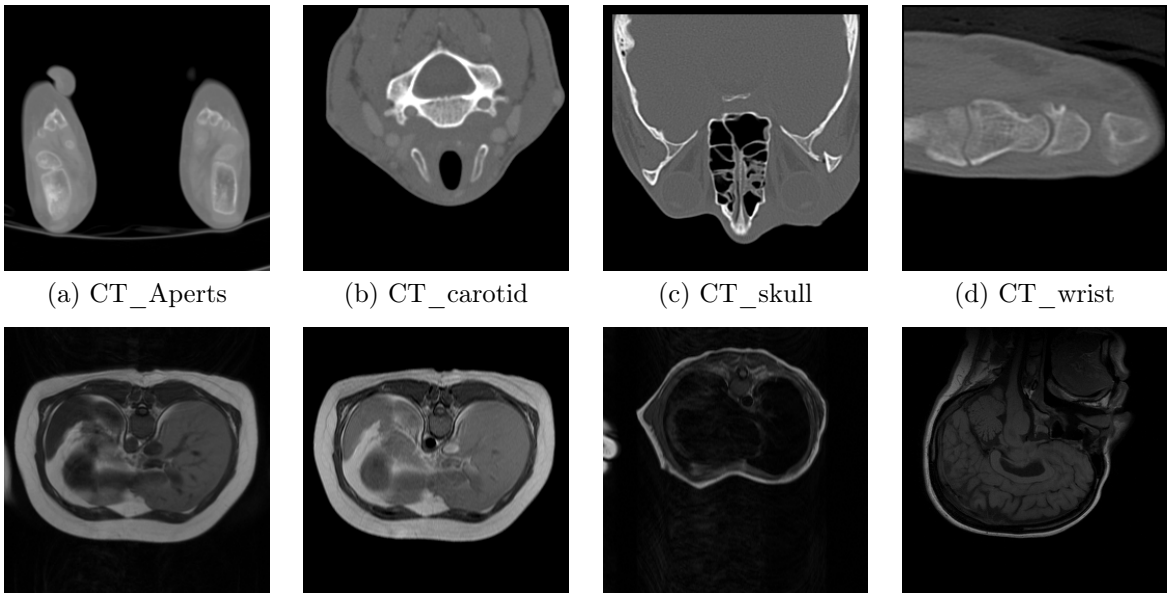
The image dataset used to assess the performance of the proposed predictive scheme is composed of eight volumetric medical images: four CTs and four MRs scans, all available from [18], the image and video repository of the Center for Image Processing (CIPR) of the Rensselaer Polytechnic Institute.

The spatial resolution, bit depth and number of slices of each volume is presented in Table 2.1. These data volumes are sets of spatially adjacent slices, which due to their resolution and number, as well as bit depth, require a large number of bits to be represented uncompressed.

Table 2.1: Medical volumes information.

Sequence	Resolution	Bit Depth	No. Slices
CT_Aperts	256x256	8	97
CT_carotid	256x256	8	74
CT_skull	256x256	8	203
CT_wrist	256x256	8	183
MR_liver_t1	256x256	8	58
MR_liver_t2e1	256x256	8	58
MR_ped_chest	256x256	8	77
MR_sag_head	256x256	8	58

Figure 2.2 shows the midpoint slice of each volume of the dataset. An important characteristic of these slices is that since they represent cuts through anatomical organs which extend into the three dimensions, typically neighbouring pixels in adjacent slices are highly correlated. The encoding methods currently recommended by DICOM, which is presented in the following section, do not exploit this correlation to achieve higher compression ratios. Therefore it is quite opportune to develop new techniques able to take advantage of the inter-slice correlation characteristics of these volumes, to increase the compression efficiency.



(a) CT_Aperts (b) CT_carotid (c) CT_skull (d) CT_wrist

(e) MR_liver_t1 (f) MR_liver_t2e1 (g) MR_ped_chest (h) MR_sag_head

Figure 2.2: Middle slice of each medical volume [18].

2.1.3 Digital Imaging and Communications in Medicine (DICOM)

Digital Imaging and Communications in Medicine (DICOM) is the standard for communication and management of medical imaging information. Among others, DICOM defines the way devices handle medical data, such as transmission or storage, including for these purposes a protocol for network communications and a file format, in which the file of a medical image contains information about the exam and the patient. DICOM recommended encoders for compression of medical images are JPEG, RLE, JPEG-LS, JPEG2000, MPEG2, H.264/AVC and H.264/AVC Stereo [3].

2.2 Lossless Encoders

In this section some of the current state-of-art lossless encoders are described, in order to better understand their compression algorithms.

2.2.1 JPEG2000

JPEG2000 is an image encoding standard [5] based in wavelet transform [19], which aimed to be a more efficient replacement to JPEG. JPEG2000 was designed to be spatially scalable by dividing an image into different sub-bands of frequency, which can be sampled at various resolutions, as can be seen in Figure 2.3.

The block diagram of JPEG2000 algorithm is shown in Figure 2.4. In the first step, each of the image samples is shifted by a DC level, in order for all sample values to be within a dynamic range centred around zero. Then a Multi-Component Transform (MCT) is applied in order to decompose the image in sub-bands, followed by the wavelet transform, with the resulting coefficients being entropically encoded.

The standard allows both lossy and lossless encoding, using a different wavelet transform for each purpose. For lossless encoding, a discrete reversible wavelet transform is used, and the quantization step is bypassed.

Motion JPEG2000

Motion JPEG2000 is an extension of JPEG2000 for video coding, adding support for audio. However, it does not perform inter-slice coding, with every frame being independently encoded by JPEG2000.

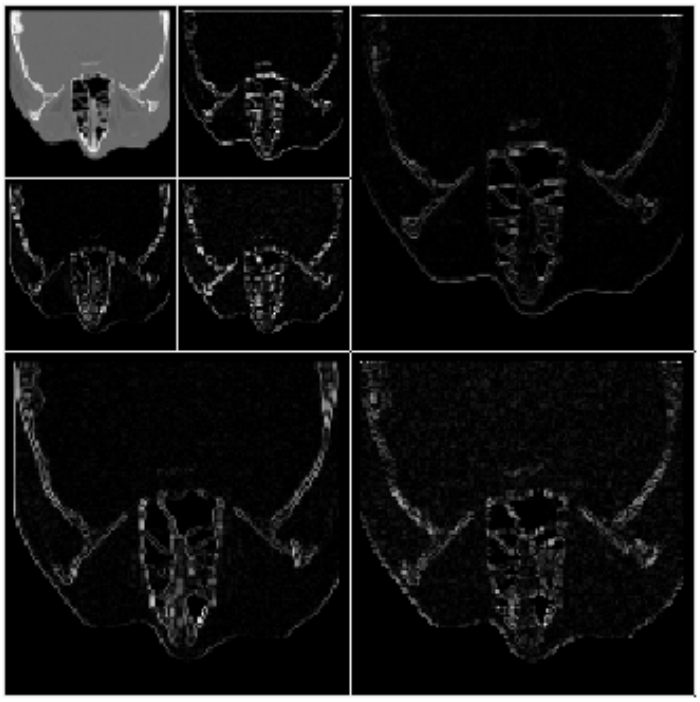


Figure 2.3: Example of a wavelet transform applied to a slice of CT_skull, obtained with MATLAB.

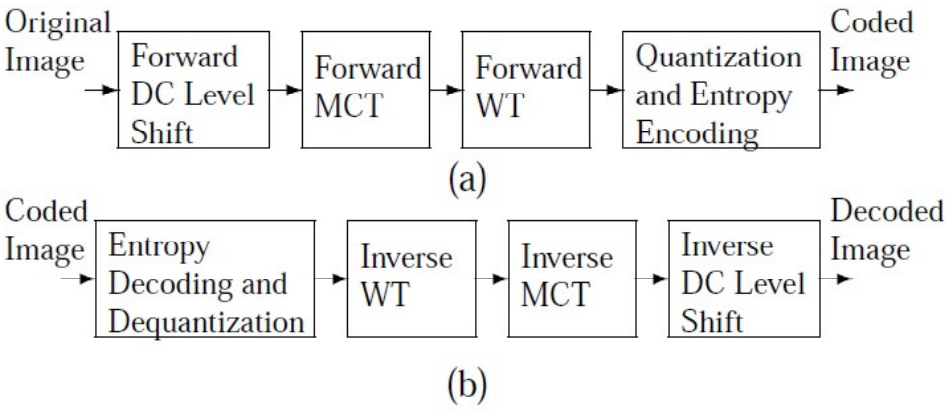


Figure 2.4: Block diagram of JPEG2000 algorithm [19], for the (a) encoder and (b) decoder.

JP3D

JP3D [9], [20] is an extension of JPEG2000 for volumetric images, such as medical images. In this extension, the tilling of the sequence results in 3D blocks, rather than 2D blocks, which are coded independently. Furthermore, a 3D discrete wavelet transform is applied instead of a 2D transform as used in JPEG2000, allowing decompositions along the three axes, and the entropy encoder is also improved for the 3D blocks.

2.2.2 JPEG-LS

JPEG-LS [4] is a standard for lossless and near-lossless compression of continuous-tone images, which is based on the Low Complexity Lossless Compression for Images (LOCO-I) algorithm. The LOCO-I algorithm consists of three steps, Prediction, Context Modelling and Codification, as detailed in the block diagram of JPEG-LS, shown in Figure 2.5.

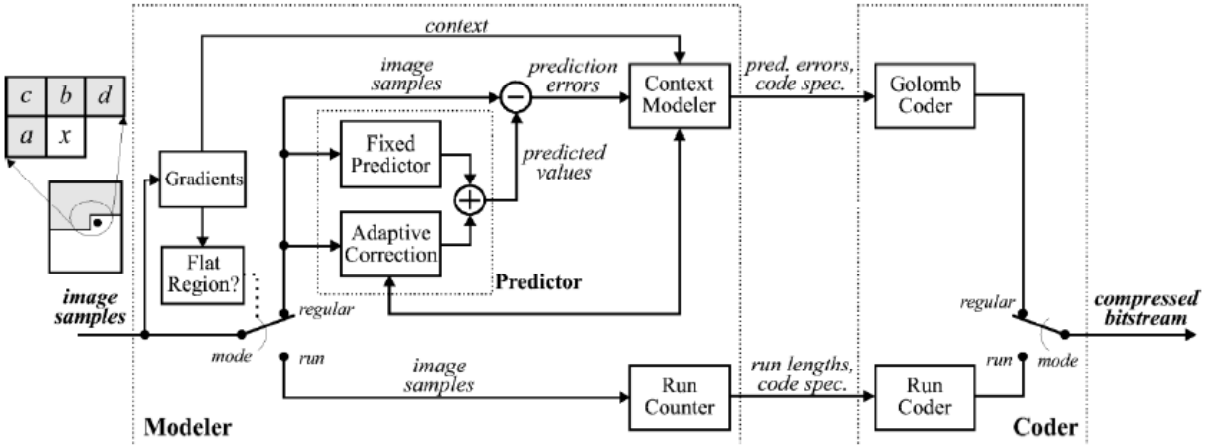


Figure 2.5: JPEG-LS block diagram [4].

The prediction step uses a causal template, presented in Figure 2.6, in which x is the current sample and a, b, c and d represent neighbouring pixels. In order to maintain a low complexity algorithm, a simple edge detection method is applied, referred to as Median Edge Detector (MED), which selects a fixed predictor, as represented by Equation 2.1. This predictor tends to choose pixel b when a vertical edge exists, pixel a when a horizontal edge is present and $a + b - c$ if no edge is detected.

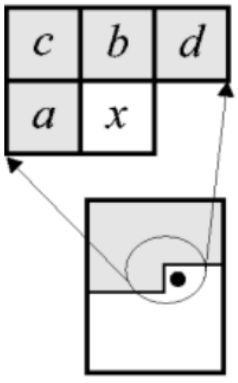


Figure 2.6: Causal template used by the MED predictor in JPEG-LS [4].

$$\hat{x}_{MED} \triangleq \begin{cases} \min(a,b), & \text{if } c \geq \max(a,b) \\ \max(a,b), & \text{if } c \leq \min(a,b) \\ a+b-c, & \text{otherwise.} \end{cases} \tag{2.1}$$

The next step is the context modelling of the prediction error, where LOCO-I calculates the local gradients in order to determine the context model, followed by the coding of the prediction residue using Golomb-type codes, which allow a symbol-by-symbol coding while preserving a low complexity. In uniform regions, a run-length variation is used instead.

2.2.3 Context based Adaptive Lossless Image Codec (CALIC)

Context based Adaptive Lossless Image Codec (CALIC) [7] is a lossless image encoder, whose algorithm is shown by the flow diagram in Figure 2.7. CALIC is composed by two independent modes: binary and continuous tone.

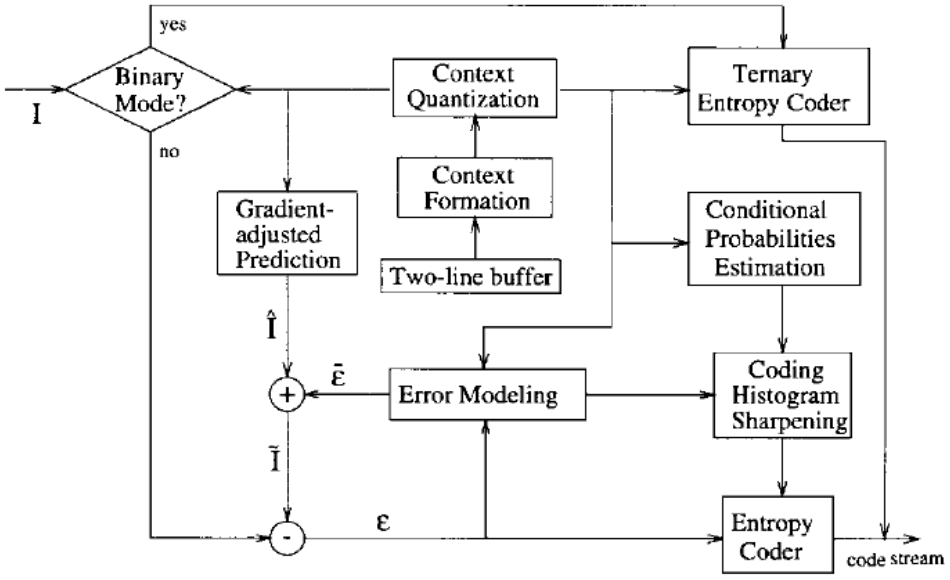


Figure 2.7: Flow diagram of the CALIC algorithm [7].

The binary mode is automatically selected, without explicitly signalling it, when the causal pixel neighbourhood used in the prediction presents only one or two different values. Then, a ternary entropy coder is applied, using one symbol for each different pixel value, and the third one is used as an escape symbol to signal the return to the continuous tone mode.

In the continuous tone mode, the Gradient Adaptive Predictor (GAP) is performed, using the causal template shown in Figure 2.8. GAP determines the horizontal and vertical gradients in order to evaluate the orientation and characteristic of the current edge, using predefined thresholds to select whether it is sharp or weak. Then, the predicted value is calculated in accordance to the edge's attributes, using the corresponding equation from Table 2.2.

After the prediction step, context modelling is performed, and the prediction residue is

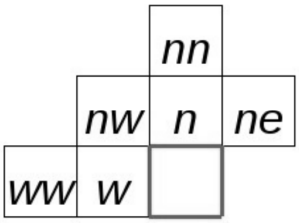


Figure 2.8: Causal pixel neighbourhood used by the GAP in CALIC [7].

Table 2.2: Predicted pixel values for each edge type.

Edge	Horizontal	Vertical
Sharp	w	n
Normal	$\frac{aux+w}{2}$	$\frac{aux+n}{2}$
Weak	$\frac{3aux+w}{4}$	$\frac{3aux+n}{4}$

$$aux = \frac{w+n}{2} + \frac{ne-nw}{4}$$

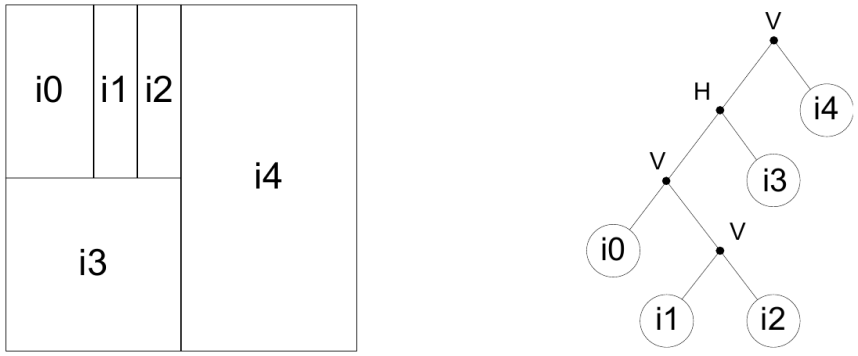
entropically encoded with an adaptive arithmetic coder.

2.2.4 Multi-scale Multidimensional Parser (MMP)

Multi-scale Multidimensional Parser (MMP) [21–23] is a pattern matching compression algorithm, where the block to encode is matched with known or previously used patterns stored in a dictionary, and the dictionary index of the matched block is coded. One of the advantages of MMP is the use of scale transformations in the stored patterns, in order to allow the matching of blocks with different sizes. MMP has been used for a wide range of applications, including still images, video sequences or stereoscopic images, managing to outperform state-of-the-art encoders. More recently, a lossless version of MMP has been proposed in [24].

The MMP algorithm divides the image to encode into non-overlapping sequential 16×16 blocks, which are encoded in raster scan order, with each block processed individually. The blocks are segmented using a flexible partition scheme, generating an optimized segmentation tree, as shown in Figure 2.9.

Instead of performing pattern matching and directly encoding the image blocks, MMP applies intra prediction techniques, which generate residue patterns with lower energy and with higher probability of existing in the dictionary, thus allowing a more efficient encoding. The prediction modes were adapted from H.264/AVC directional modes for



(a) (b)

Figure 2.9: (a) Flexible segmentation of a block, (b) corresponding segmentation tree [23].

4×4 blocks, with the DC mode being replaced by the Most Frequent Value (MFV) mode. There is also an additional mode based on LSP, which is described in detail in Chapter 5.

In order to obtain the optimal segmentation tree, all the prediction modes are tested for every partition, and the resulting residues are coded with MMP. Then, the best prediction mode and partition direction is selected using a rate-distortion optimization function. During the encoding of the image, the used patterns and several variations are added to the dictionary, in order to be reused further in the image.

In the lossless implementation of MMP, the horizontal, vertical and LSP intra prediction modes were enhanced by always using the adjacent pixels in the prediction, instead of using only the ones in the upper or left border of the block. This translates into a Differential Pulse-Code Modulation (DPCM) which allows a more efficient prediction.

3D-MMP

3D-MMP is an extension of the MMP image coder for volumetric images [23, 25]. The first major difference is the use of 3D blocks instead of 2D. Consequently, these blocks can be segmented along three directions, as shown in Figure 2.10, which means that the number of possible block sizes grows considerably, leading to a great increase of the computational complexity. To mitigate this added computational complexity, multiple copies of the patterns in different scales (block sizes) are stored in the dictionary, so as to reduce the number of scale transformations, although this significantly increases the required memory.

In order to take advantage of the inter-slice redundancy present in volumetric images, the intra prediction modes of MMP were adapted for 3D blocks. For LSP, when predicting a pixel, since the previous slice has already been encoded, the causal template includes

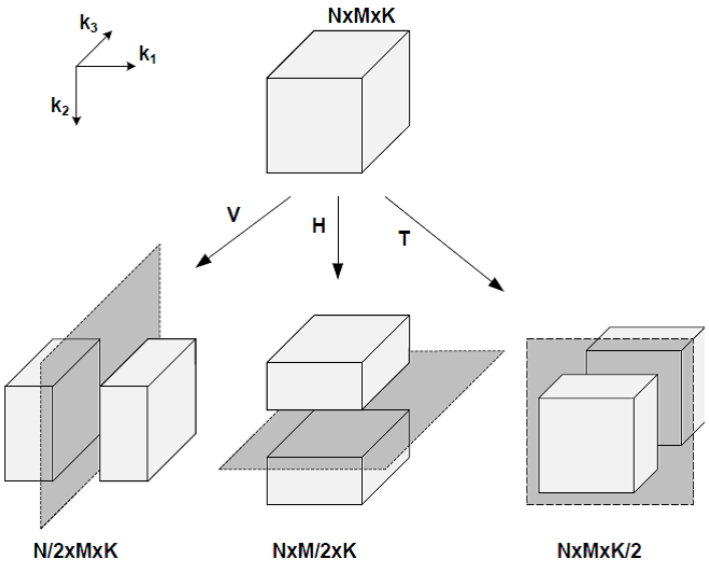


Figure 2.10: Triadic flexible partition [23].

pixels of the previous slice, which can belong to the same or a previous block. The H.264/AVC based modes were also extended, using a causal neighbourhood around the 3D block. Besides these adapted modes, in 3D-MMP a new inter-slice prediction mode based on motion estimation was introduced.

For video coding, an hierarchical architecture is used, in which the 3D blocks are generated by grouping the slices alternately, that is, on the temporal axis, the first group of blocks are composed of just odd slices, while the second group contains only even slices, and so forth. This approach, represented in Figure 2.11, is similar to the I/P and B slices compression used on H.264/AVC and HEVC codecs, with the first group of blocks being I/P blocks, and the second B blocks.

This approach allows B blocks to be encoded after the respective I/P blocks. Thus, LSP prediction can be improved for B blocks, since for each slice both the temporally previous and the next slice have been encoded in the I/P block. This way, when coding a pixel in a B block, the pixels used for LSP prediction are the previously coded neighbours of the same slice, and the pixels in the same region from the temporally previous and next slices coded in the preceding I/P block.

2.2.5 Minimum Rate Predictors (MRP)

Minimum Rate Predictors (MRP) [8] is a lossless encoder that uses linear predictors, which are reconfigured for each image to encode. The image is divided into 32×32 size blocks, which can be partitioned, using a quad-tree, down to 2×2 blocks, with the predictors being used in the resulting blocks. Each block is classified in one of M classes, such that

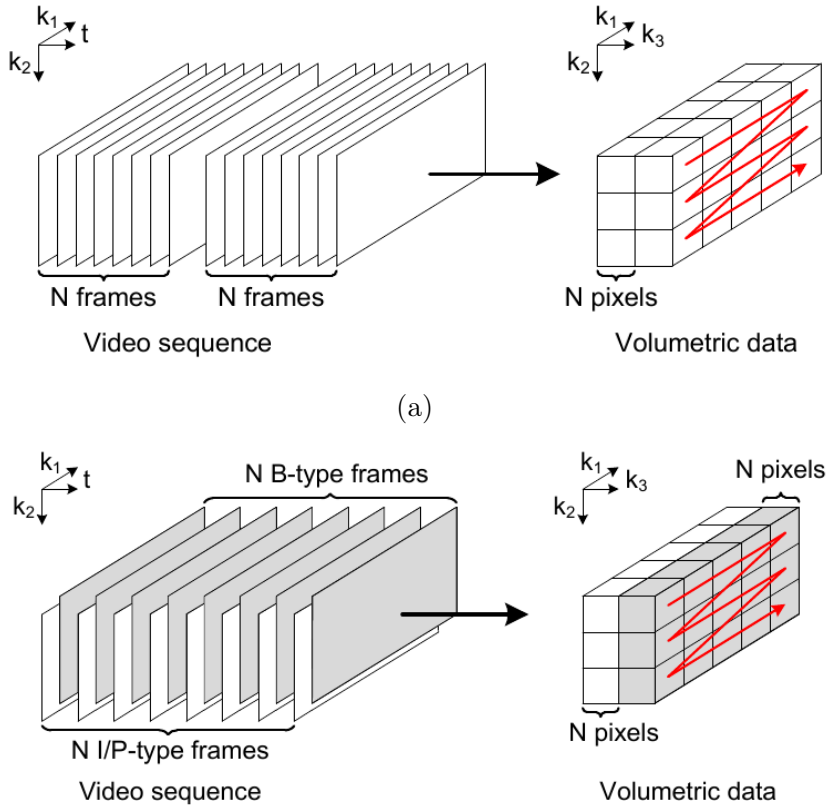


Figure 2.11: Video coding architecture in 3D-MMP: a) Sequential and b) Hierarchical [23].

the linear predictors that are used in each block depends on its class.

For a given pixel, p_0, the value predicted by the m-th predictor ($m = 1, ..., M$) is given by the Equation 2.2:

$$\hat{s}(p_0) = \sum_{k=1}^{K} a_m(k) \cdot s(p_k), \tag{2.2}$$

where K is the prediction order, $s(p_k)$ is the value of the p_k pixel and $a_m(k)$ is the respective prediction coefficient. Figure 2.12 shows the distribution of the reference pixels for a prediction order of $K = 30$. The resulting prediction error is adaptively encoded using context modelling, with the distribution of the prediction errors being estimated by a generalized Gaussian.

Several steps for the optimization of the prediction, the quad-tree and coding information are performed, in order to minimize the cost function for the encoding image. The prediction order K and the number of classes M can be determined from the characteristics of the encoding image, so as to use the predictors which best estimate the image structures.

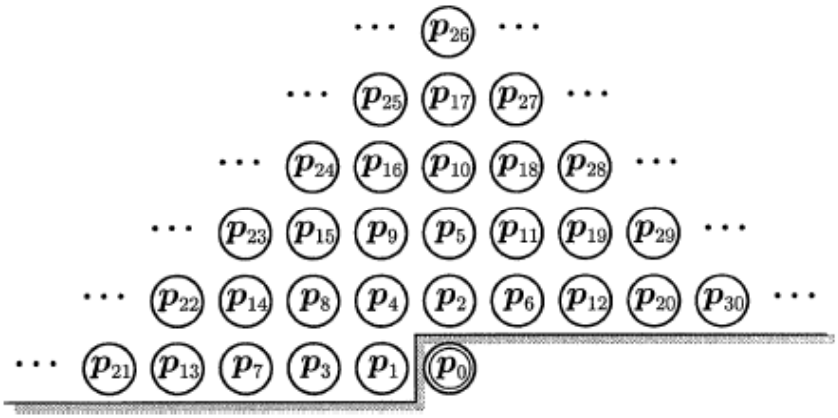


Figure 2.12: Distribution of reference pixels [8].

2.2.6 H.264/Advanced Video Coding (AVC)

H.264/Advanced Video Coding (AVC) is a video coding standard which has a hybrid coding structure [6], [26]. The coding algorithm, whose block diagram is shown in Figure 2.13, includes the following main steps: intra/inter frame prediction, transform, quantization and entropy coding. Each frame is encoded at the block level, thus each frame is divided into macro-blocks of 16×16 pixels, which can be subsequently segmented. After the intra or inter prediction, a transform (Hadamard or Discrete Cosine Transform (DCT)) is applied to the obtained residue, resulting in a block of coefficients, which are later quantized. Finally, entropy coding is used.

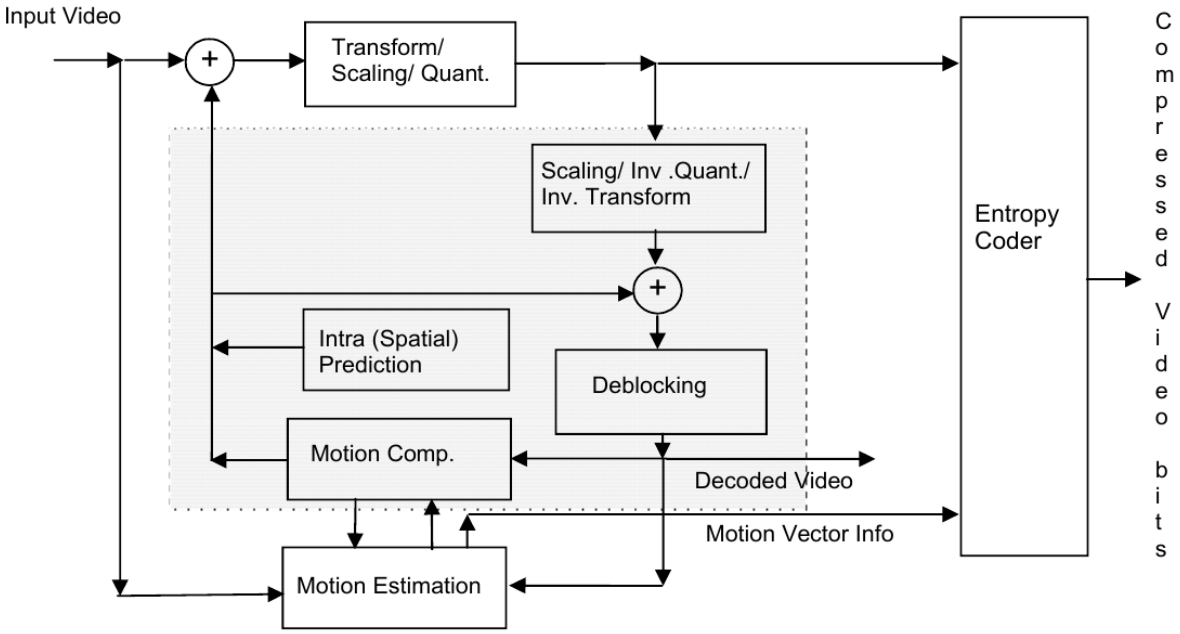


Figure 2.13: Block Diagram of H.264/AVC [6].

The frames of a video sequence are classified in three different types, I, P and B. For

I frames, the prediction used is intra-frame only, that is, spatial prediction. The intra prediction modes of H.264/AVC for 4×4 blocks consist of a DC mode and eight directional modes, which are represented in Figure 2.14, while for 16×16 block only 4 modes are available, horizontal, vertical, DC and planar.

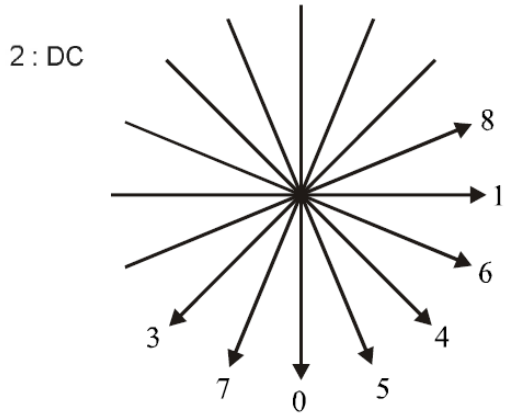


Figure 2.14: H.264/AVC Intra prediction directional modes [26].

As for both P and B frames, besides intra prediction, inter-frame prediction with motion compensation can be used, in which pixels from only previous frames or upcoming frames as well, respectively, can be used to predict the current block. For P frames, only one motion vector for a block is transmitted, while for B frames two can be sent, allowing some form of weighted prediction.

This process is lossy, since by using quantization, the information is lost and can't be recovered in the decoder. Thus, the H.264/AVC standard allows lossless representation of macro-blocks by simply using Pulse Code Modulation (PCM) mode, in which the values of the samples are sent directly without prediction. However, this method is not tuned for efficiency.

Therefore, a more effective lossless coding technique is used: transform-bypass lossless mode. As the name suggests, the transform and the quantization steps are skipped, and only intra prediction and entropy coding are used.

2.2.7 High Efficiency Video Coding (HEVC)

High Efficiency Video Coding (HEVC) [1, 27] is the most recent video coding standard following H.264/AVC, also using a hybrid coding structure. The main difference is the replacement of macro-blocks with Coding Tree Units (CTUs), which can use larger blocks of up to 64×64 pixels, and can also be further partitioned. The use of larger blocks allows the number of intra prediction modes to increase from nine, in the H.264/AVC standard,

to 35, as can be seen in Figure 2.15, and it significantly reduces the bit rate, specially for higher resolution video sequences.

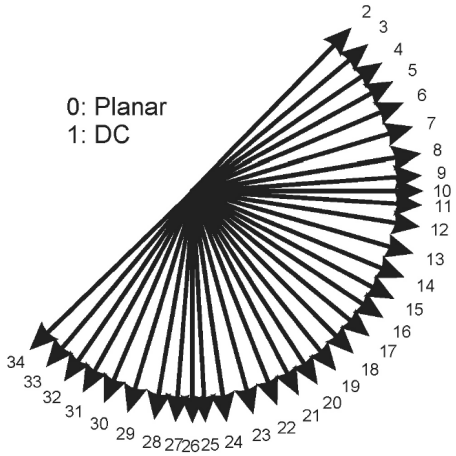


Figure 2.15: HEVC Intra prediction directional modes [1].

Similarly to the H.264/AVC standard, HEVC allows lossless CTU coding in two ways: I_PCM, in which the prediction, transform, quantization, entropy coding and in-loop filtering are not used, and the samples are represented by a fixed number of bits and transmitted directly in the bitstream. And the Lossless mode, in which only the transform, quantization and in-loop filtering are skipped, with the residue resulting from intra and inter prediction being coded entropically, as can be seen in Figure 2.16. The first one can be used to code noise-like signals, and the second one is used to code the remaining signals.

Since the crucial steps of residue transformation and quantization are not used, the compression gain is small and the efficiency of prediction methods becomes of utmost importance for the overall encoding performance. Unlike in H.264/AVC, the lossless mode can be used with all the coding structures: Intra, in which all frames are coded using only intra prediction (I frames), Random Access, using both intra and inter prediction (I, P and B frames), and Low Delay, in which only the closest reference images are used for inter prediction (I and B or P frames).

Following the completion of the first edition of the HEVC standard, several additional coding tools were added as part of the Range Extension (REext) [29]. These extensions include tools to support the encoding of content with higher bit depths and higher chroma-resolution, and a technique for lossless encoding. This technique is based on Differential Pulse-Code Modulation (DPCM) encoding of the residues resulting from intra directional prediction modes. The DPCM encoding of the intra prediction residue, also known as SAP [30], is better suited for lossless encoding since it uses the original pixel values of adjacent neighbours to compute the predicted samples in the current Prediction Unit

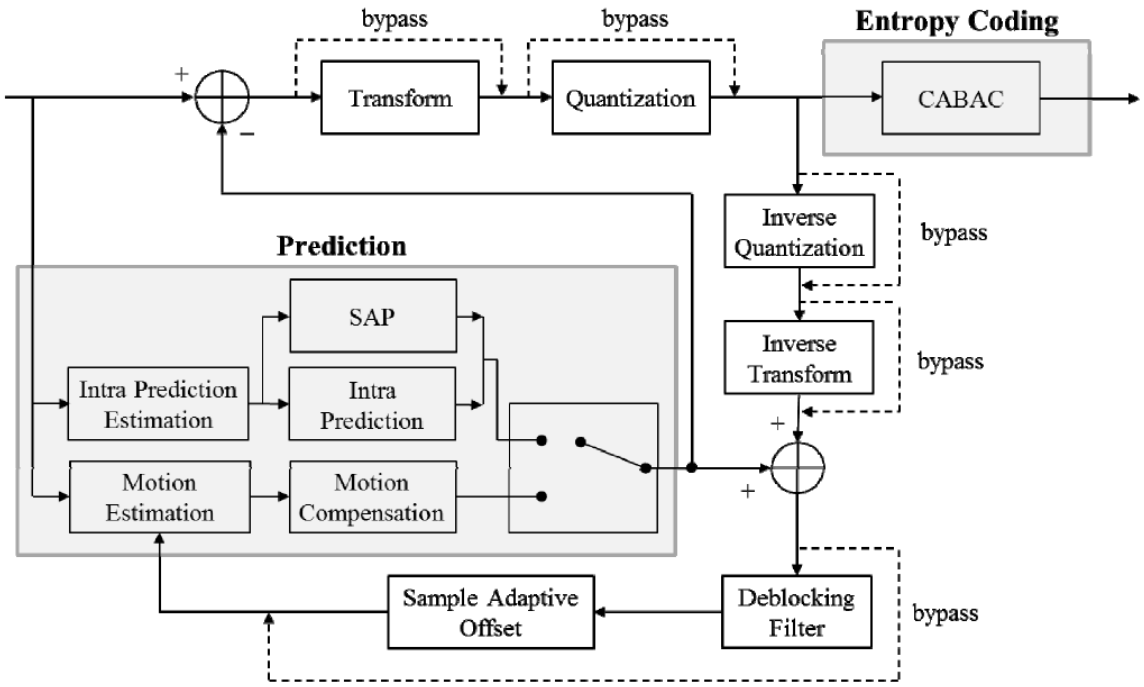


Figure 2.16: Block Diagram of HEVC for lossless coding, with dashed lines [28].

(PU), instead of only those in the edge of the PU.

2.2.8 Performance Comparison

After the study of state-of-the-art lossless encoders, their encoding performance for medical images was compared. Publicly available software implementations of JPEG-LS, JPEG2000, JP3D, CALIC, MRP, H.264/AVC and HEVC were used for these tests.

Firstly, to evaluate the impact of the RExt in HEVC, the test sequences were encoded using HEVC with and without the RExt profile, for both Intra and R.A. coding structures. The encoding configurations of HEVC were mostly left as default, using the latest release of the reference software HM-16.4. To allow lossless coding, the QP was set to 0, and both TransquantBypassEnableFlag and CUTransquantBypassFlagForce were set to 1. For the RExt profile, the parameter CostMode was set to lossless.

Table 2.3 shows the results, in bits per pixel (bpp), obtained in this test. It can be seen that using inter-slice prediction, such as in Random Access, allows a major bitrate reduction, of 39%, in opposed to Intra only, since it exploits the high inter-slice redundancy present in medical image volumes. As for the use of the RExt profile, for both Intra and Random Access coding structures, there is a performance gain of roughly 9%, which is expected due to the SAP technique, devised specifically for lossless encoding. Given these results, from here forward in this work, every test performed with HEVC includes the

Table 2.3: HEVC performance comparison, with and without Range Extension (REext) profile (in bits per pixel (bpp)).

Sequence	HEVC Intra	HEVC REext Intra	HEVC R.A.	HEVC REext R.A.
CT_Aperts	1.289	1.142	0.826	**0.734**
CT_carotid	2.198	1.955	1.587	**1.383**
CT_skull	3.083	2.981	1.905	**1.822**
CT_wrist	2.195	1.919	1.155	**1.024**
MR_liver_t1	3.742	3.399	2.392	**2.118**
MR_liver_t2e1	2.811	2.512	1.726	**1.552**
MR_ped_chest	3.352	3.037	1.699	**1.515**
MR_sag_head	2.732	2.599	1.873	**1.771**
Average	2.675	2.443	1.645	**1.490**

REext profile.

Table 2.4 presents the lossless compression performance comparison of all the tested encoders. For H.264/AVC, the latest release of the reference software JM 18.6 was used. For lossless coding, FREext High 4:4:4 Profile was used for Intra coding, with QP and QP Offsets set to 0. JPEG-LS and CALIC were used with the default configuration for lossless coding. In MRP, the optimization flag was activated, with the remaining options left as default. For JPEG2000, the software implementation OpenJPEG v2.1 is used with the default configuration. This software also includes an implementation of JP3D, which was configured to use a 3D Discrete Wavelet Transform and 3D Embedded Block Coding with Optimal Truncation (EBCOT). As for MMP [24], the fast implementation was used, with the default dictionary size.

Table 2.4: Lossless compression performance comparison of state-of-the-art encoders (in bpp).

Sequence	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
CT_Aperts	1.193	1.142	1.261	1.058	1.178	0.998	0.775	0.941	**0.734**
CT_carotid	2.062	1.955	2.019	1.778	1.977	1.684	**1.374**	1.547	1.383
CT_skull	3.183	2.981	2.991	2.761	2.959	2.628	2.329	2.088	**1.822**
CT_wrist	1.911	1.919	1.757	1.627	1.717	1.550	1.173	1.238	**1.024**
MR_liver_t1	3.489	3.399	3.256	3.160	3.393	3.022	2.582	2.356	**2.118**
MR_liver_t2e1	2.806	2.512	2.572	2.418	2.460	2.269	1.722	1.745	**1.552**
MR_ped_chest	3.080	3.037	3.021	2.937	3.074	2.789	2.337	2.071	**1.515**
MR_sag_head	2.635	2.599	2.905	2.582	2.808	2.519	2.279	2.160	**1.771**
Average	2.545	2.443	2.473	2.290	2.446	2.183	1.821	1.768	**1.490**

For lossless coding, only JP3D and HEVC R.A. take advantage of the redundancy between slices, while all the other encoders only use intra prediction, being image encoders. As expected, encoders which exploit the inter-slice redundancies generally present better results, with HEVC REext R.A. obtaining the smallest bitrate on average, with a consid-

erable difference over JP3D. As for the image encoders, they present a generally similar performance, with the exceptions of CALIC and more notably MRP, which manages to be close to JP3D, even with only intra prediction.

Another test was performed using 3D-MMP, with block size of 16×16×4. Due to the way it is implemented, the number of slices encoded for each sequence must be multiple of four, in order to segment it into complete blocks. Table 2.5 shows the number of encoded slices for each sequence and the results for 3D-MMP. Although the number of encoded slices is not the same for the other encoders, it differs in only one or two slices, which have an almost negligible effect in the results. When comparing to the other encoders, 3D-MMP has a superior performance to most of the image encoders, but falls short of the video and volumetric encoders HEVC and JP3D, standing somewhat behind MRP.

The great disadvantage of 3D-MMP is its unaffordable computational complexity. Given this, and the fact that its results are far from the best, 3D-MMP was not used during the course of this work.

Table 2.5: Lossless encoding results of 3D-MMP (in bpp).

Sequence	No. Slices	3D-MMP
CT_Aperts	96	0.938
CT_carotid	72	1.622
CT_skull	-	-
CT_wrist	-	-
MR_liver_t1	56	2.579
MR_liver_t2e1	56	2.051
MR_ped_chest	76	1.804
MR_sag_head	56	2.303

2.3 Related State-of-the-Art

In this section, a review of previous work relevant to this dissertation is presented, mainly on two topics: new lossless prediction techniques in HEVC, and lossless encoding methods for medical images.

2.3.1 Lossless Coding in HEVC

With the increasing relevance of screen content coding, the need for more efficient lossless prediction techniques increases, which can be aimed at other applications such as medical imaging. Several techniques for lossless prediction have been proposed in addition to SAP to improve the lossless encoding of HEVC RExt.

The Sample-based Weighted Prediction (SWP), proposed in [31], performs a weighted average of the four causal neighbouring pixels immediately surrounding the current pixel, referred to as a, b, c, d in Figure 2.17, to form its prediction. In order to determine the weights attributed to each pixel, the Sum of Absolute Differences (SAD) is calculated between the patch shown in Figure 2.17 and the same patch placed over each of the neighbouring pixels a, b, c, d. After determining the SAD for each of the four pixels, the corresponding weights are obtained through a look-up table. The SWP replaces the planar mode in HEVC, and presents significant reductions of bitrate when encoding intra only (8.5%), and more modest ones for random access coding structure (2.6%).

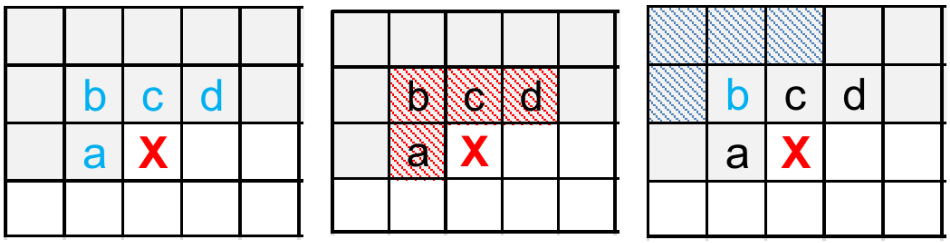


Figure 2.17: Sample-based Weighted Prediction (SWP) [32]; Causal neighbourhood used by SWP (left); causal patch placed over pixel to predict X (centre); patch placed over pixel b.

An alternative to SWP is the Directional Template Matching (DTM) [32], which instead of performing a weighted average of the neighbouring pixels, uses directly the one with the highest similarity, that is, the pixel for which the placed patch presents the lowest SAD, thus achieving better reconstruction of sharp edges. The authors also found that using a combination of SWP and DTM provides better prediction performance, by using DTM as an exception algorithm of SWP, with a bitrate reduction of 1.15 percentage points (pp) over just using DTM. Other possible combination is SWP replacing the planar mode and DTM replacing the DC mode, with a bitrate reduction of 2.15 percentage points over just using DTM.

Another method recently proposed is the Sample-based Angular intra-Prediction with Gradient-based (SAP-G) edge prediction [33], which replaces the planar mode. This technique uses the surrounding pixels to compute the gradients in four directions (0 deg., 45 deg., 90 deg. and 135 deg.), as shown in Figure 2.18, and then selects the adjacent neighbouring pixel in the direction of the smallest gradient as the prediction for the current pixel. In addition to this, SAP-G applies the same concept of SAP to the DC mode, computing the average of immediate neighbours of the current pixel instead of the values of the column to the left and the row above the current block. This technique achieves a slight performance improvement (1.05 percentage points) when compared to SWP+DTM method.

In [34], the Sample-based Angular intra-Prediction with Median and Edge (SAP-ME)

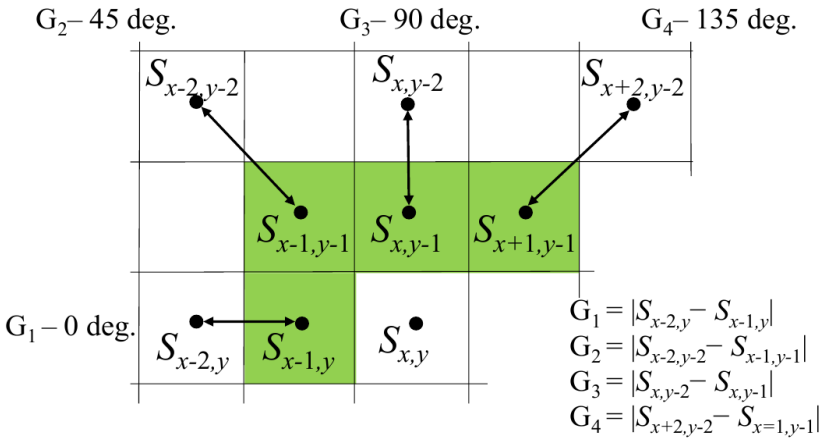


Figure 2.18: Computing of gradients in four directions in SAP-G [33].

predictor was proposed. In order to improve the prediction of sharp edges, the Median Edge Detector (MED) of JPEG-LS [4] is applied, replacing the planar mode, using the samples b, c and d as shown in Figure 2.19, with the predicted value $P_{x,y}$ given by:

$$P_{x,y} = \begin{cases} \min(b,d), & \text{if } c \geq \max(b,d) \\ \max(b,d), & \text{if } c \leq \min(b,d) \\ b+d-c, & \text{otherwise.} \end{cases} \tag{2.3}$$

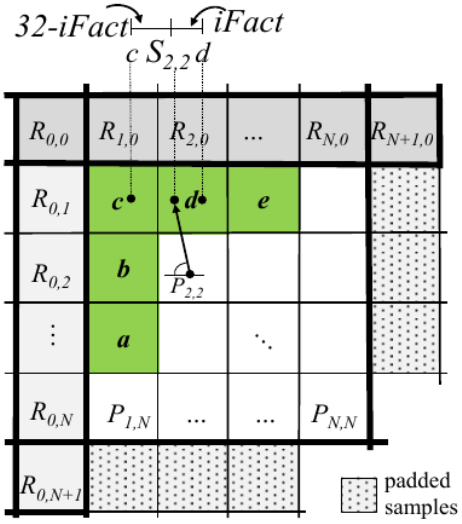


Figure 2.19: Prediction principle of SAP [34], where the prediction value is obtained using bilinear interpolation between the immediate neighbours along the prediction direction.

To improve prediction in smooth regions, SAP-ME also introduces a median predictor to replace the DC mode, in which the predicted pixel assumes the value of the median

between the samples a, b, c, d and e, from Figure 2.19, given by Equation 2.4:

$$P_{x,y} = \text{median}\{a, b, c, d, e\}. \tag{2.4}$$

SAP-ME is used with SAP as well, and presents an improvement of 1.74 percentage points over the previously described SWP+DTM. All the previously described techniques show that pixel-wise prediction is efficient for lossless compression.

2.3.2 Medical Image Encoding

The growth in the use of medical images led to the need for more efficient compression techniques. More recently, the increasing use of three dimensional medical images, such as CTs and MRs, and even Four-Dimensional (4D) ones, such as functional magnetic resonances, consisting of a sequence of 3D volumes obtained over time, resulted in the development and proposal of several novel methods for their compression.

In [35] a lossless compression technique for 4D medical images using H.264/AVC was proposed. H.264/AVC was modified to just bypass the quantization step, in order to maintain the use of inter prediction. Two methods were proposed, as shown in Figure 2.20, using different coding orders. H.264-VOL encodes each volume separately, with each slice of the volume corresponding to a frame, thus allowing inter-slice prediction between slices of the same volume. As for H.264-TIME, it encodes along the fourth dimension (time), that is, for all the volumes, the slices in the same position are grouped, with each group being encoded independently.

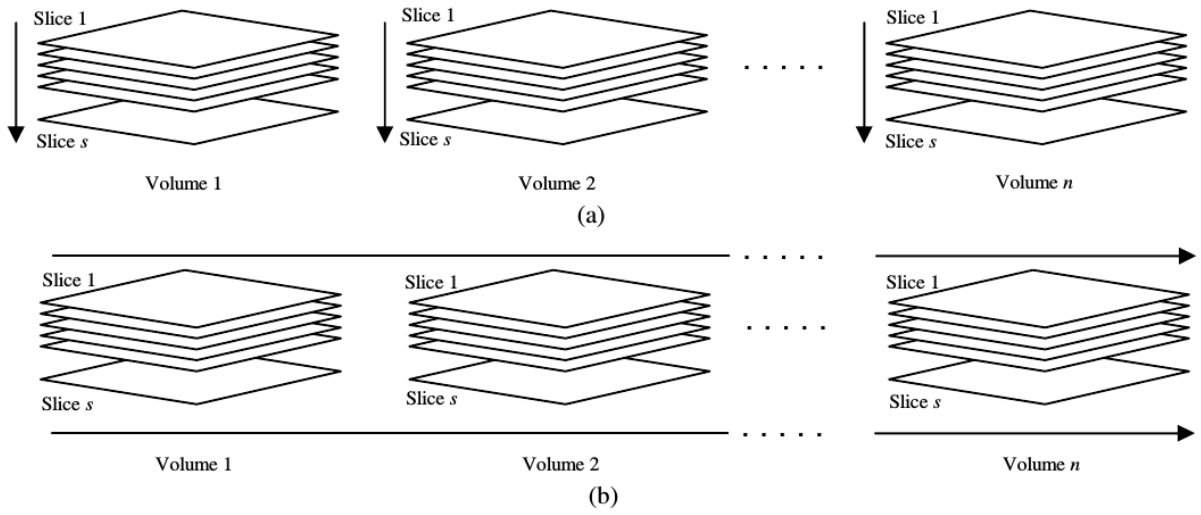


Figure 2.20: Encoding order of 4D medical images using the proposed methods: (a) H.264-VOL and (b) H.264-TIME [35].

Comparing to DICOM recommended encoders, both approaches present significant improvements, however, H.264-TIME achieves a much higher compression ratio, with an increase of almost 70% over previous techniques. This is explained since 4D medical images present the changes of a 3D volume over time, and thus the redundancy between the same slices of the volumes is more significant.

In [36], the same authors propose an improved technique which combines the two methods, as shown in Figure 2.21. Firstly, the motion compensation of H.264-VOL ($MC1$) is applied, generating residual slices and the respective motion vectors (MV_V). Then, the motion compensation of H.264-TIME ($MC2$) is applied to the residual slices, with the resulting residue and motion vectors (MV_T) being encoded with Context-Adaptive Binary Arithmetic Coding (CABAC). The motion vectors MV_V are differentially coded with a proposed method and then also encoded with CABAC, incorporating the bitstream. The proposed technique greatly improves, up to three times, the compression ratio of state-of-the-art compression techniques.

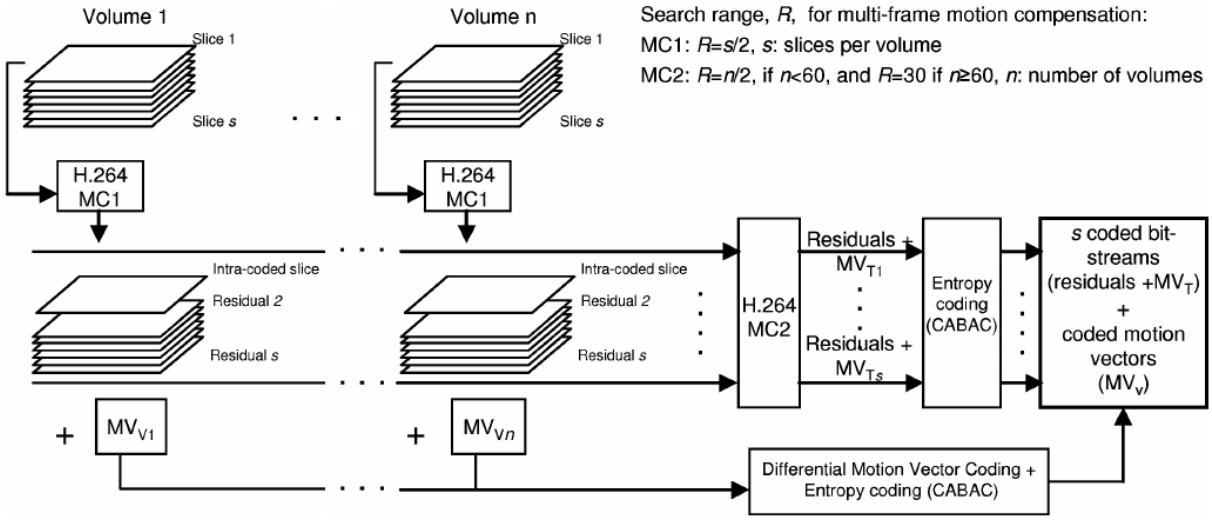


Figure 2.21: Block diagram of the proposed technique with H.264/AVC [36].

A new technique for compression of 3D volumetric medical images using linear prediction was proposed in [37]. The corresponding block diagram is shown in Figure 2.22. The Medical Images Lossless Compression (MILC) algorithm performs the prediction on a pixel-by-pixel basis, using one of two prediction contexts, depending on whether it is predicting the first slice or not. If it is the first slice, a 2D Linearised Median Predictor (2D-LMP) is used. Similar to the MED predictor of JPEG-LS [4], the three neighbours A, B and C of the current pixel X are used to compute the predicted value $\hat{X}$, by performing the average of the three possible predictors of MED, resulting in:

$$\hat{X} = \frac{2 \times (A + B)}{3} - \frac{C}{3}. \tag{2.5}$$

For the remaining slices, since there is at least one previously encoded slice, the 3D Distances-based Linearised Median Predictor (3D-DLMP) is used. In addition to the three neighbouring pixels in the same slice, four collocated pixels A', B', C' and X' from the previous slice are used for the prediction. Firstly, the differences between the collocated pixels values are computed by $\delta_A = A - A', \delta_B = B - B'$ and $\delta_C = C - C'$. Then, the optimized approximated difference, or distance, δ is given by:

$$\delta = \frac{2 \times (\delta_A + \delta_B)}{3} - \frac{\delta_C}{3}, \tag{2.6}$$

similarly to the predictor 2D-LMP. Finally, the predicted value $\hat{X}$ is computed by adding the determined distance δ to the pixel collocated to the current pixel to predict, given by $\hat{X} = X' + \delta$. After obtaining the prediction values, the prediction error is then modelled and coded. The experimental results show that MILC presents a better performance than other state-of-the-art lossless encoders, including CALIC and the DICOM recommended JPEG-LS, reaching a bitrate reduction of 27% and 28%, respectively.

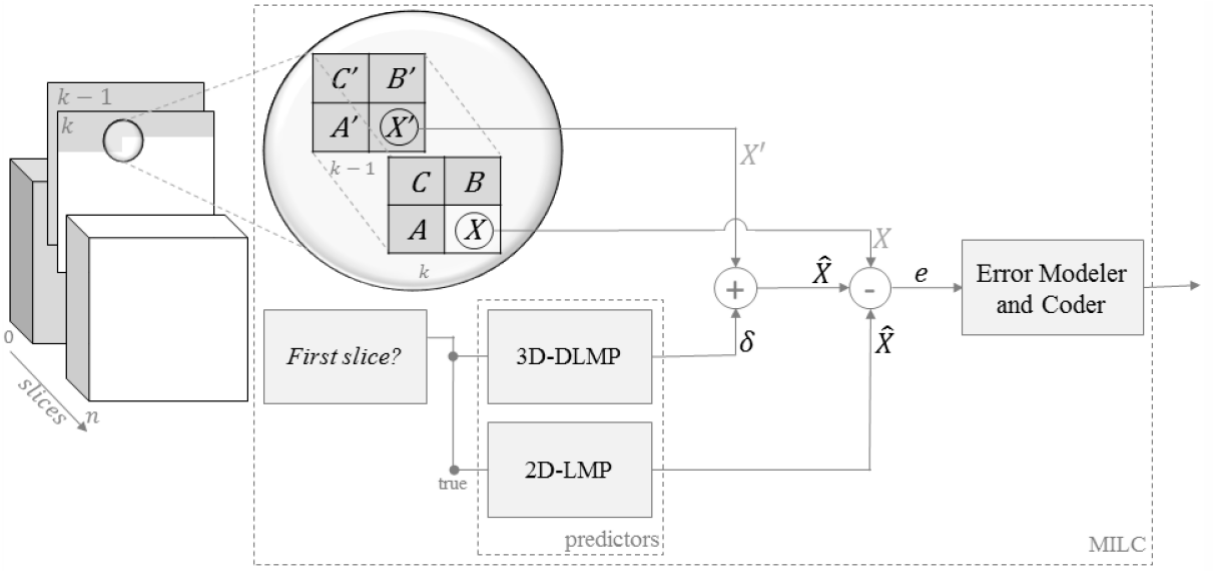


Figure 2.22: Block diagram of the MILC algorithm [37].

A new method for compression of 3D medical images based on their symmetries was proposed in [38] and improved in [39]. As can be seen in the block diagram of Figure 2.23, firstly, a 2D integer wavelet transform is applied to each slice of the volume. Then, an intra-band prediction method based on the symmetry of the sub-band is used, followed by an inter-slice DPCM prediction applied to the resulting residue. The final residue and the prediction parameters are then entropically encoded.

After applying the wavelet transform, in each sub-band the global symmetry axis is determined, selecting one of four directions: horizontal, vertical and diagonal up or down. The sub-band is divided into two areas, according to the axis, and partitioned into 16×16

blocks. To predict each block, a spatial transformation is applied to previously encoded blocks, and the one which generates the residue with lowest energy is selected. Eight different spatial transformations are tested, including flips and rotations along different directions, which allows to take advantage of local symmetries. Since the blocks are processed by alternating between both areas of the global symmetry axis, this method also takes advantage of the global symmetry.

Once the intra-band prediction residue is obtained, the inter-slice DPCM method is applied. The residue is divided into 3D blocks, and for each block, DPCM is applied along five directions, with the one generating the lowest energy error being selected. The proposed method achieves a significant compression gain over DICOM recommended encoders, with an average improvement of 17% over JPEG2000 and H.264/AVC intra coding.

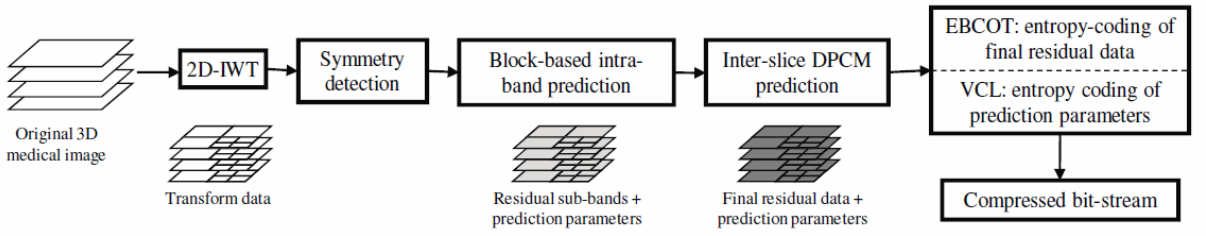


Figure 2.23: Block diagram of the symmetry-based compression technique [39].

Chapter 3

Medical Images Processing

As described in the previous chapter, generally, lossless encoders are not able to exploit the inter-slice redundancy that is present in typical MR and CT images, that are the focus of this work. Thus, in this chapter, several lossless processing techniques are applied to medical image sequences, prior to their encoding. The aim of these techniques is to take advantage of this inter-slice redundancy and improve the compression efficiency.

The following sections describe each technique and present their respective results. Given the characteristic of these techniques, in order to carry out a fair comparison, in this chapter all the tests were performed using a number of slices multiple of 16 in all sequences, as shown in Table 3.1. The performance of all the encoders described in Chapter 2 for these test conditions is presented in Table 3.2.

It can be seen that there are no significant differences over the results with all of the slices, shown in Chapter 2. These results will serve as base of comparison to all the tested approaches.

Table 3.1: Medical volumes with number of slices multiple of 16.

Sequence	No. Slices
CT_Aperts	96
CT_carotid	64
CT_skull	192
CT_wrist	176

Sequence	No. Slices
MR_liver_t1	48
MR_liver_t2e1	48
MR_ped_chest	64
MR_sag_head	48

Table 3.2: Lossless coding performance comparison for the original sequences with a number of slices multiple of 16 (in bpp).

Sequence	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
CT_Aperts	1.200	1.142	1.267	1.064	1.184	1.004	0.776	0.945	**0.734**
CT_carotid	2.017	1.955	1.981	1.739	1.936	1.647	**1.347**	1.507	1.383
CT_skull	3.284	2.981	3.079	2.847	3.058	2.713	2.411	2.143	**1.822**
CT_wrist	1.939	1.919	1.782	1.653	1.743	1.574	1.190	1.261	**1.024**
MR_liver_t1	3.493	3.399	3.257	3.158	3.383	3.019	2.566	2.409	**2.118**
MR_liver_t2e1	2.754	2.512	2.530	2.369	2.416	2.223	1.688	1.771	**1.552**
MR_ped_chest	3.066	3.037	3.015	2.928	3.061	2.780	2.327	2.055	**1.515**
MR_sag_head	2.639	2.599	2.908	2.585	2.810	2.522	2.281	2.190	**1.771**
Average	2.549	2.443	2.477	2.293	2.449	2.185	1.823	1.785	**1.490**

3.1 Concatenated Slices

Despite the main focus of this dissertation being the improvement of HEVC for lossless encoding, a test was performed in order to improve the performance of the lossless encoders. For instance, MMP is a dictionary-based image encoder in which the previously encoded blocks are stored in a dictionary as patterns, used to predict the following blocks.

As such, by concatenating all slices of a sequence side-by-side into a single image, the encoded blocks of previous slices can be used to predict the blocks in the current slice, allowing some form of inter-slice prediction to be used in the image encoders, particularly in MMP.

The achieved results for this approach can be seen in Table 3.3 where all encoders present bitrate reductions, albeit minimal in most cases, with MMP presenting the most significant one as expected, with a reduction of 8.5%, followed by CALIC with 4.6%. Given the large number of slices for the sequences CT_skull and CT_wrist, their respective concatenated images present a very high horizontal resolution, which is not supported in H.264/AVC and JPEG-LS. As such, the results presented in Table 3.3 for these encoders and sequences are the same from encoding the original sequences in Table 3.2, in order to calculate the average and facilitate the comparison.

Table 3.3: Encoding results of the concatenated slices (in bpp).

Sequence	H.264	HEVC RExt Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP
CT_Aperts	1.177	1.136	1.246	1.011	1.017	0.894	**0.736**
CT_carotid	2.003	1.949	1.955	1.698	1.757	1.548	**1.276**
CT_skull	3.284	2.973	3.051	2.847	2.882	2.639	**2.355**
CT_wrist	1.939	1.911	1.757	1.653	1.605	1.492	**1.139**
MR_liver_t1	3.472	3.393	3.233	3.106	3.146	2.903	**2.496**
MR_liver_t2e1	2.742	2.504	2.506	2.316	2.197	2.107	**1.571**
MR_ped_chest	3.040	3.029	3.000	2.873	2.781	2.663	**2.264**
MR_sag_head	2.618	2.592	2.883	2.534	2.546	2.434	**2.243**
Average	2.534	2.436	2.454	2.255	2.241	2.085	**1.760**

3.2 Directional Approaches

Medical images such as CTs and MRs consist of volumes representing a Three-Dimensional (3D) object. These volumes can be observed along different directions, creating slices with varying features, as shown in Figure 3.1. Slices designated by (A) are originally generated in the Z axis, forming a volume from which the slices along X (B) and Y (C) axes are extracted.

The number of slices in the Z axis becomes the width of the slices in the X and Y axes. As some encoders require image dimensions to be multiple of 16, all the tests were performed using a number of slices multiple of 16, as previously stated.

By encoding the volumes along these other directions, different redundancies are exploited by the intra encoders. Tables 3.4 and 3.5 show the compression results for the medical volumes tested along different directions, X and Y, respectively. It can be seen that the image/intra encoders present considerable bitrate reductions, ranging from 14 to 25%

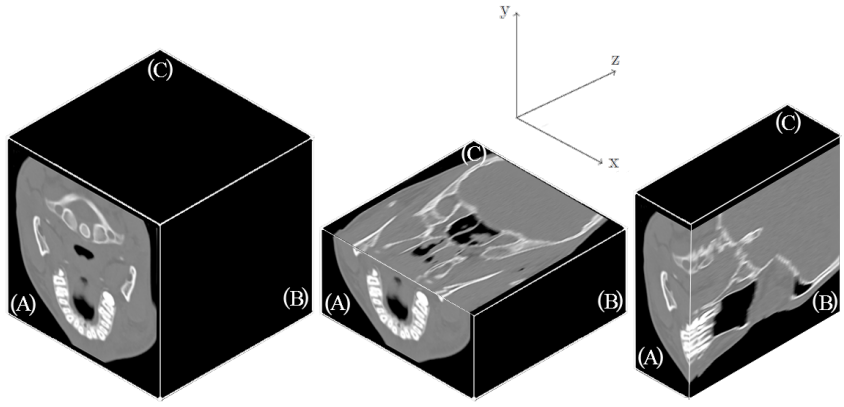


Figure 3.1: Example of slices from a 3D volume along different directions.

over the encoding in the Z axis, with the slices along X showing a slight advantage. As for video or volumetric encoders, such as HEVC and JP3D, respectively, a significant increase of the bitrate is verified, being around 10% for HEVC and 30% for JP3D. With this improvement, MRP surpasses HEVC and becomes the most efficient encoder for these medical volumes.

This can be explained by the different redundancy levels in slices of each axis. For example, the high redundancy between slices in the Z axis, which is not exploited by the image encoders, is now represented in the X and Y slices, and can thus be an advantage for these encoders. However, for HEVC and JP3D, the inter-slice prediction techniques are the basis for their advantageous performance, while the intra-slice prediction is not as effective, hence obtaining better results when there is more redundancy between slices, as in the Z axis.

Table 3.4: Compression results for slices in the X axis (in bpp).

Sequence	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
CT_Aperts	1.074	1.008	1.134	0.937	1.125	0.888	**0.676**	1.630	0.771
CT_carotid	1.777	1.701	1.723	1.459	1.735	1.458	**1.150**	1.910	1.424
CT_skull	2.493	2.314	2.227	2.127	2.352	1.970	**1.599**	2.850	2.038
CT_wrist	1.651	1.602	1.483	1.357	1.460	1.289	**0.962**	1.734	1.136
MR_liver_t1	3.261	3.106	2.717	2.557	3.128	2.583	**2.013**	3.161	2.317
MR_liver_t2e1	2.486	2.289	2.059	1.945	2.328	1.944	**1.393**	2.675	1.889
MR_ped_chest	2.140	2.003	1.966	1.747	2.071	1.712	**1.322**	2.275	1.713
MR_sag_head	2.417	2.325	2.358	2.059	2.498	2.081	**1.687**	2.580	2.033
Average	2.162	2.043	1.958	1.773	2.087	1.741	**1.350**	2.352	1.665

Table 3.5: Compression results for slices in the Y axis (in bpp).

Sequence	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
CT_Aperts	1.090	1.037	1.157	0.943	1.113	0.873	**0.691**	1.572	0.757
CT_carotid	1.792	1.720	1.763	1.530	1.842	1.518	**1.231**	1.864	1.391
CT_skull	2.517	2.339	2.285	2.198	2.411	2.004	**1.639**	2.776	2.006
CT_wrist	1.513	1.483	1.259	1.154	1.276	1.112	**0.821**	1.912	1.209
MR_liver_t1	3.243	3.088	2.705	2.510	3.084	2.535	**2.052**	3.143	2.335
MR_liver_t2e1	2.568	2.319	2.144	2.053	2.430	2.005	**1.642**	2.393	1.737
MR_ped_chest	2.198	2.030	2.103	1.844	2.198	1.848	**1.460**	2.175	1.670
MR_sag_head	2.422	2.335	2.366	2.068	2.522	2.085	**1.682**	2.590	2.036
Average	2.168	2.044	1.972	1.788	2.110	1.747	**1.402**	2.303	1.643

3.3 Inter-Slice Prediction

Another approach on the inter-slice redundancy of medical volumes involved the use of inter-slice prediction techniques, with the resulting residue being encoded instead of the original sequences. Firstly, a simple zero-order inter-slice predictor was tested. Due to its promising results, the more complex prediction scheme of HEVC was also tested.

3.3.1 Zero-Order Inter-Slice Predictor

In order to capitalise on the redundancy, a zero-order inter-slice predictor was used, which is translated into a pixel-wise difference between adjacent slices, resulting in a residue that will then be encoded.

This technique is implicit to an extent. Although it is a simple technique that is applied to all the pixels, there may be additional information to transmit to the decoder. That is, when performing a difference operation between two eight bit pixels, the result may have to be represented by a nine bit value, in order to be lossless. In such cases, the pixel value in the residue is truncated to eight bits and the additional information for that pixel is transmitted to the decoder. Given the very low occurrence of this issue, its impact in the final results is insignificant.

Table 3.6 shows the compression results for this technique, implemented along the Z axis. When compared to encoding the original sequences in the same axis, in Table 3.2, the intra/image encoders present very significant gains, with bitrate reductions between 28 and 36%. The encoders with inter-slice prediction, JP3D and HEVC REext Random Access (R.A.), also present important improvements, although not as accentuated, of approximately 9 and 14%, respectively. With this method, HEVC REext R.A. achieves the best encoding results, on average, followed closely by the MRP image encoder.

Besides the initial Z axis, this technique was also implemented along X and Y axes, whose results are given in Tables 3.7 and 3.8, respectively. Unlike in the Z axis, when compared to their counterparts, this method exhibits much lower gains, being marginal for most encoders, or even losses, as is the case for MRP, JP3D and HEVC REext R.A.. Thus, encoding the difference residue along the Z axis still yields a large advantage.

Given these results, another test was performed, in which the difference residue obtained along the Z axis is encoded along X and Y directions, achieving the results presented in Tables 3.9 and 3.10. Even though it shows improvements compared to the residue obtained along X and Y axes, there is still a considerable distance to the encoding results of the residue along the Z axis.

Table 3.6: Compression results of the pixel-wise difference residue along the original Z axis (in bpp).

Sequence	H.264	HEVC REXt Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REXt R. A.
CT_Aperts	0.791	0.776	0.933	0.790	0.885	0.771	**0.630**	0.890	0.679
CT_carotid	1.459	1.442	1.546	1.354	1.500	1.318	**1.115**	1.380	1.243
CT_skull	2.071	2.045	2.173	2.011	2.141	1.935	1.696	1.924	**1.488**
CT_wrist	1.115	1.117	1.223	1.049	1.145	1.064	**0.862**	1.171	0.879
MR_liver_t1	2.216	2.211	2.287	2.098	2.256	2.049	**1.790**	2.243	1.897
MR_liver_t2e1	1.781	1.774	1.815	1.681	1.714	1.597	1.307	1.632	**1.245**
MR_ped_chest	1.576	1.580	1.761	1.577	1.696	1.536	1.324	1.785	**1.292**
MR_sag_head	1.983	1.979	2.177	1.995	2.216	1.975	1.789	2.046	**1.538**
Average	1.624	1.615	1.739	1.569	1.694	1.531	1.314	1.634	**1.282**

Table 3.7: Compression results of the pixel-wise difference residue along the X axis (in bpp).

Sequence	H.264	HEVC REXt Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REXt R. A.
CT_Aperts	0.914	0.887	1.053	0.869	1.060	0.872	**0.713**	1.497	0.823
CT_carotid	1.608	1.580	1.637	1.399	1.678	1.417	**1.194**	1.935	1.506
CT_skull	2.330	2.294	2.218	2.157	2.353	2.012	**1.729**	2.909	2.205
CT_wrist	1.171	1.160	1.276	1.094	1.278	1.125	**0.907**	1.601	1.116
MR_liver_t1	2.631	2.522	2.543	2.323	2.832	2.361	**2.031**	3.041	2.409
MR_liver_t2e1	2.128	2.047	1.957	1.834	2.135	1.811	**1.411**	2.610	1.989
MR_ped_chest	2.053	1.924	2.131	1.867	2.236	1.893	**1.590**	2.550	1.901
MR_sag_head	2.407	2.327	2.540	2.260	2.706	2.288	**1.946**	2.757	2.288
Average	1.905	1.843	1.920	1.725	2.035	1.722	**1.440**	2.363	1.780

Table 3.8: Compression results of the pixel-wise difference residue along the Y axis (in bpp).

Sequence	H.264	HEVC REXt Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REXt R. A.
CT_Aperts	0.908	0.888	1.059	0.868	1.021	0.845	**0.714**	1.419	0.826
CT_carotid	1.541	1.513	1.622	1.395	1.677	1.401	**1.194**	1.858	1.424
CT_skull	2.252	2.215	2.201	2.127	2.323	1.976	**1.715**	2.794	2.127
CT_wrist	1.325	1.321	1.267	1.129	1.299	1.124	**0.921**	1.879	1.268
MR_liver_t1	2.641	2.537	2.563	2.321	2.826	2.355	**2.064**	3.015	2.395
MR_liver_t2e1	1.973	1.903	1.919	1.795	2.107	1.752	**1.527**	2.213	1.695
MR_ped_chest	1.930	1.796	2.122	1.820	2.185	1.872	**1.591**	2.348	1.748
MR_sag_head	2.407	2.324	2.540	2.265	2.693	2.286	**1.939**	2.757	2.288
Average	1.872	1.812	1.912	1.715	2.016	1.701	**1.458**	2.285	1.721

Table 3.9: Compression results, along the X axis, encoding the pixel-wise difference residue obtained in the Z axis (in bpp).

Sequence	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
CT_Aperts	0.901	0.878	1.127	0.910	1.077	0.948	0.742	1.689	**0.706**
CT_carotid	1.523	1.504	1.606	1.429	1.702	1.476	**1.223**	1.678	1.300
CT_skull	1.958	1.967	2.112	1.784	2.074	1.853	**1.659**	2.567	1.784
CT_wrist	1.220	1.233	1.394	1.152	1.282	1.203	0.994	1.687	**0.987**
MR_liver_t1	2.458	2.474	2.660	2.266	2.810	2.406	2.090	3.058	**2.055**
MR_liver_t2e1	1.793	1.803	2.027	1.618	2.019	1.703	**1.421**	2.681	1.575
MR_ped_chest	1.603	1.609	1.921	1.571	1.781	1.597	**1.374**	1.993	1.481
MR_sag_head	2.063	2.074	2.349	1.878	2.357	2.094	**1.816**	2.793	1.860
Average	1.690	1.693	1.900	1.576	1.888	1.660	**1.415**	2.268	1.469

Table 3.10: Compression results, along the Y axis, encoding the pixel-wise difference residue obtained in the Z axis (in bpp).

Sequence	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
CT_Aperts	0.918	0.910	1.169	0.923	1.072	0.938	0.765	1.631	**0.702**
CT_carotid	1.568	1.568	1.631	1.494	1.783	1.523	**1.295**	1.642	1.301
CT_skull	1.978	1.986	2.156	1.814	2.118	1.870	**1.704**	2.524	1.772
CT_wrist	1.094	1.106	1.232	1.052	1.145	1.009	**0.860**	1.781	0.990
MR_liver_t1	2.443	2.459	2.647	2.248	2.801	2.358	**1.670**	3.017	2.048
MR_liver_t2e1	1.850	1.863	2.089	1.688	2.091	1.785	**1.295**	2.373	1.503
MR_ped_chest	1.673	1.666	2.006	1.646	1.876	1.726	1.489	1.957	**1.475**
MR_sag_head	2.077	2.083	2.352	1.884	2.360	2.096	**1.814**	2.789	1.861
Average	1.700	1.705	1.910	1.594	1.906	1.663	**1.362**	2.214	1.457

3.3.2 HEVC Prediction

Since performing a simple zero-order inter-slice prediction and encoding its residue produced a substantial reduction of the bitrate, then a more complex and efficient prediction technique could further improve the results. HEVC REext R.A. has proven to accomplish a very efficient prediction, both intra and inter slice, as it yields the best encoding performance of volumetric medical images, as shown in Table 3.2. As such, the residue resulting from the HEVC prediction was extracted, and subsequently compressed by all the tested encoders.

Unlike the zero-order prediction, the HEVC prediction is explicit, that is, it is necessary to transmit to the decoder the information signalling the used prediction mode for each block, including the angular mode in intra prediction, the motion vector for inter prediction, or merge and skip flags, among other data. The cost of this extra information, presented for each sequence in Table 3.11, is then added to the compression results.

Table 3.11: Prediction signalling cost of HEVC REext R.A. (in bpp).

Sequence	HEVC REext R.A. Prediction Cost
CT_Aperts	0.043
CT_carotid	0.049
CT_skull	0.041
CT_wrist	0.025
MR_liver_t1	0.053
MR_liver_t2e1	0.036
MR_ped_chest	0.037
MR_sag_head	0.056
Average	0.043

Table 3.12 shows the encoding results of the extracted residue, with the prediction cost already added. It can be seen that for all of the intra/image encoders there is a great reduction of the bitrate when compared to the original sequences encoding, exceeding the performance of the zero-order predictor for almost all, with the exception of MRP which becomes slightly worse with the addition of the prediction cost. JP3D is the only encoder which presents a bitrate increase, while HEVC REext R.A. presents a rather marginal gain over encoding the original sequences, but it is still distant from the zero-order predictor results.

Table 3.12: Compression results for the HEVC RExt R.A. residue (in bpp).

Sequence	H.264	HEVC RExt Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC RExt R. A.
CT_Aperts	0.762	0.762	0.912	0.797	0.874	0.785	**0.679**	1.016	0.737
CT_carotid	1.416	1.425	1.526	1.377	1.508	1.356	**1.207**	1.646	1.391
CT_skull	1.881	1.851	1.998	1.854	1.923	1.765	**1.617**	2.384	1.773
CT_wrist	1.043	1.046	1.176	1.044	1.070	1.019	**0.881**	1.478	1.012
MR_liver_t1	2.137	2.140	2.267	2.100	2.209	2.014	**1.856**	2.755	2.082
MR_liver_t2e1	1.589	1.574	1.674	1.544	1.586	1.466	**1.301**	2.037	1.491
MR_ped_chest	1.537	1.543	1.745	1.621	1.622	1.525	**1.375**	2.228	1.507
MR_sag_head	1.807	1.804	2.027	1.845	1.965	1.808	**1.666**	2.381	1.747
Average	1.521	1.518	1.666	1.523	1.595	1.467	**1.323**	1.991	1.468

3.4 Summary

In this chapter, a few techniques were applied to medical image volumes, in order to take advantage of the inter-slice redundancy and improve the compression efficiency. Table 3.13 displays the average compression results for each tested approach, in bits-per-pixel, and Table 3.14 shows the difference in bitrate of each approach, in comparison to the original (Z axis), in percentage, meaning that a decrease in the bitrate is represented by a negative number.

It can be observed that while encoding the sequences along different directions (X and Y) improves the performance of intra encoders, the use of inter-slice prediction techniques presents the best results. Encoding the HEVC residue obtains the largest bitrate reduction for most encoders, even when considering the additional signalling information. As for the remaining encoders, MRP, JP3D and HEVC RExt R.A., the zero-order inter-slice predictor (diff res) in the original Z axis generates the best results. Overall, HEVC RExt R.A. has the best performance, with the best result for the difference residue, followed closely by MRP.

These tests allowed to assess the importance of inter-slice prediction for volumetric medical image encoding. It is clear that there still is margin for improving the prediction of HEVC RExt for this type of images. As such, the remaining of this dissertation is focused on the development of more efficient inter-slice prediction techniques in HEVC, using two distinct approaches: Chapter 4 focuses on motion estimation using Geometric Transformations (GT), while in Chapter 5 a pixel-wise prediction technique based in Least-Squares Prediction (LSP) is implemented.

Table 3.13: Summary of the results for each approach (in bpp).

Approach	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
Z	2.549	2.443	2.477	2.293	2.449	2.185	1.823	1.785	1.490
cat	2.534	2.436	2.454	2.255	2.241	2.085	1.760	-	-
X	2.162	2.043	1.958	1.773	2.087	1.741	1.350	2.352	1.665
Y	2.168	2.044	1.972	1.788	2.110	1.747	1.402	2.303	1.643
resZ	1.624	1.615	1.739	1.569	1.694	1.531	**1.314**	**1.634**	**1.282**
resX	1.905	1.843	1.920	1.725	2.035	1.722	1.440	2.363	1.780
resY	1.872	1.812	1.912	1.715	2.016	1.701	1.458	2.285	1.721
resZinX	1.690	1.693	1.900	1.576	1.888	1.660	1.415	2.268	1.469
resZinY	1.700	1.705	1.910	1.594	1.906	1.663	1.362	2.214	1.457
resHEVCinZ	**1.521**	**1.518**	**1.666**	**1.523**	**1.595**	**1.467**	1.323	1.991	1.468

Table 3.14: Percentage of bitrate difference in comparison to the original sequences encoding (Z). Negative means a reduction of the bitrate.

Approach	H.264	HEVC REext Intra	JPEG 2000	JPEG LS	MMP	CALIC	MRP	JP3D	HEVC REext R. A.
concatenated	-0.6	-0.3	-1.0	-1.7	-8.5	-4.6	-3.5	-	-
X	-15.2	-16.4	-21.0	-22.7	-14.8	-20.3	-25.9	31.8	11.7
Y	-15.0	-16.3	-20.4	-22.0	-13.9	-20.0	-23.1	29.0	10.2
diff res Z	-36.3	-33.9	-29.8	-31.6	-30.8	-30.0	**-27.9**	**-8.5**	**-13.9**
diff res X	-25.3	-24.6	-22.5	-24.7	-16.9	-21.2	-21.0	32.4	19.5
diff res Y	-26.6	-25.8	-22.8	-25.2	-17.7	-22.1	-20.0	28.0	15.5
diff res Z in X	-33.7	-30.7	-23.3	-31.3	-22.9	-24.0	-22.4	27.1	-1.4
diff res Z in Y	-33.3	-30.2	-22.9	-30.5	-22.2	-23.9	-25.3	24.0	-2.2
HEVC res	**-40.3**	**-37.9**	**-32.8**	**-33.6**	**-34.9**	**-32.9**	-27.5	11.5	-1.5

Chapter 4

Geometric Transformations (GT)

As seen in the previous chapter, performing inter-slice prediction greatly improves the compression efficiency of CTs and MRs, due to its high redundancy between slices. The features of these volumetric images consist of tissues and bones, with changes between slices that may consist on expansions, contractions, rotations or translations, which can be described by geometric transformations.

In this chapter, different approaches to the lossless compression of volumetric medical images using geometric transformations in HEVC are described.

4.1 State-of-the-Art

Geometric Transformations (GT) specify the mapping of a coordinate system into another [40–42]. A point of coordinates $[u, v]$ is corresponded to the output coordinates $[x, y]$ through a function, also known as forward mapping, and can be backward transformed to the original coordinates using an inverse function, also known as inverse mapping, as can be seen in Figure 4.1. The forward and inverse mapping functions can be expressed, respectively, by:

$$[x, y] = [X(u, v), Y(u, v)] \tag{4.1}$$

and

$$[u, v] = [U(x, y), V(x, y)]. \tag{4.2}$$

There are several different types of geometric transformations, some more complex than others, namely affine, perspective, bilinear, etc. Given the purpose of these approaches being motion estimation with geometric transformations, and with the necessity of transmitting its coefficients, the affine, projective and bilinear transformations were consid-

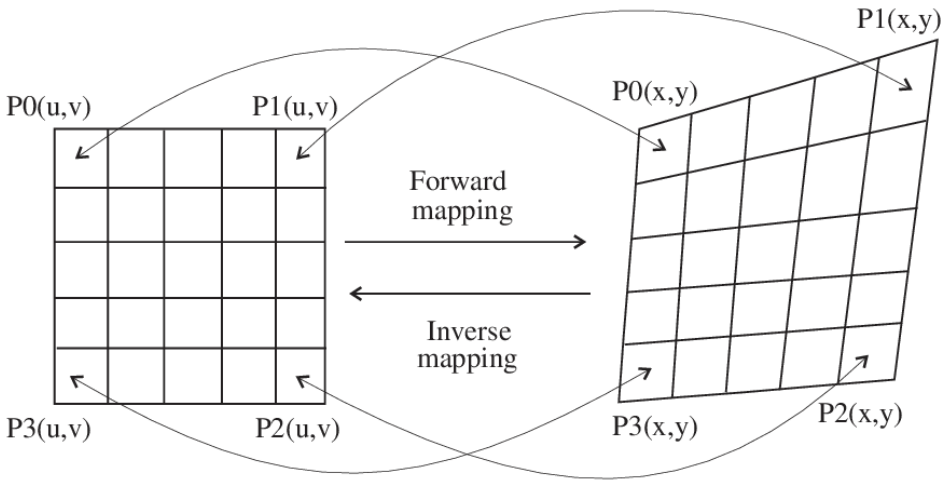


Figure 4.1: Example of forward and inverse geometric mappings of an image [40].

ered. Being simple transformations, these are some of the most used for motion estimation [40, 41], and are described in the following sections.

4.1.1 Projective Transformation

The projective transformation can be defined by a 3×3 matrix, applied to homogeneous coordinate systems. In order to guarantee that a two dimensional coordinate system $[x, y]$ is represented as homogeneous, a third element h is added. The homogeneous form is then represented as $[x', y', h]$, where $x = x'/h$, $y = y'/h$ and $h \neq 0$. Considering that only two dimensional plane transformations are relevant for this work, it can be considered that $h = 1$. Thus, the forward mapping of the point $[u, v]$ into $[x, y]$ is given by:

$$[x, y, 1] = [u, v, 1] \times T = [u, v, 1] \times \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}. \tag{4.3}$$

The matrix T can be divided into four sub-matrices, as shown in Equation 4.4, where each one is responsible for different operations.

$$T = \begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \\ \begin{bmatrix} a_{31} & a_{32} \end{bmatrix} & a_{33} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & a_{33} \end{bmatrix} \tag{4.4}$$

T_{11} allows transformations such as rotation, scaling or shearing, while the sub-matrix T_{12} specifies a projective transformation. T_{21} defines translations in x and y directions, and

finally a_{33} is a scaling factor. Considering that $h_{33} = 1$, the forward mapping functions are defined as:

$$x = a_{11}u + a_{21}v + a_{31} - a_{13}xu - a_{23}xv \tag{4.5}$$

and

$$y = a_{12}u + a_{22}v + a_{32} - a_{13}yu - a_{23}yv. \tag{4.6}$$

Having eight coefficients, that is eight Degrees of Freedom (DoF), they can be determined with four points in a 2D coordinate system and their correspondent points in the desired transformation coordinate system. Therefore, the projective transformation allows a quadrilateral to be transformed into another quadrilateral.

4.1.2 Affine Transformation

The affine transformation is a particular case of the projective transformation, where $a_{13} = a_{23} = 0$ and $a_{33} = 1$, and thus is represented by:

$$[x, y, 1] = [u, v, 1] \times T = [u, v, 1] \times \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix}. \tag{4.7}$$

From Equations 4.5 and 4.6, the forward mapping functions of the affine transformation are then defined as:

$$x = a_{11}u + a_{21}v + a_{31} \tag{4.8}$$

and

$$y = a_{12}u + a_{22}v + a_{32}. \tag{4.9}$$

With only six coefficients or DoF, a triangle can be transformed into another triangle, but it is not possible to transform a quadrilateral into an arbitrary quadrilateral. However, it can perform the mapping of a rectangle into a parallelogram.

4.1.3 Bilinear Transformation

The bilinear transformation allows a more complex mapping, such as a quadrilateral transformed into a non-planar quadrilateral. The forward mapping functions take the

form:

$$[x, y] = [uv, u, v, 1] \times \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix}, \tag{4.10}$$

resulting in the Equations:

$$x = a_{11}uv + a_{21}u + a_{31}v + a_{41} \tag{4.11}$$

and

$$y = a_{12}uv + a_{22}u + a_{32}v + a_{42}. \tag{4.12}$$

Similarly to the projective, the bilinear transformation has eight DoF, and as such requires four points in order to determine the coefficients.

When performing a geometric transformation in an image, the transformed coordinates may not be integer values, which results in some pixels not corresponding to another pixels, leading to "holes" or overlapping pixels. To solve this problem, interpolation is used, which can be a simple nearest neighbour approach, using the value of the nearest pixel, or the bilinear interpolation, where a more robust weighted average of the values of pixels in the neighbourhood is used instead.

4.2 Motion Compensation using Geometric Transformations

The Block Matching Algorithm (BMA) is the preferred method for motion estimation and compensation for video coding standards, such as H.264/AVC and HEVC. It consists of matching the current block to be encoded to the block in the reference picture which minimizes the prediction error. The BMA is essentially a geometric transformation with two Degrees of Freedom (DoF), representing only translation movements.

However, the motion between slices in medical images can be much more complex, including expansions, contractions, rotations, and others. These types of motion can be achieved using geometric transformations with six or eight DoF, such as the affine, projective or bilinear transformations, described in the previous section. In this section, three different approaches using these geometric transformations in HEVC are described.

4.2.1 Image-wise Motion Estimation

As seen in Chapter 3, applying a simple pixel-wise difference between slices and encoding the residue produces the best results. So, by using a more complex image-wise motion estimation it is possible to minimize the energy of the residue, resulting in a more efficient compression.

In this method, geometric transformations are applied to the whole image. The corners of the previous/reference slice are moved independently within a search range, generating different geometric transformations [42]. The resulting transformed slice is compared to the current slice and the geometric transformation which minimizes the prediction error is selected. The residue is then coded with the HEVC REext, and the new coordinates for each slice are transmitted in the form of motion vectors, in order to obtain the coefficients of the transformation in the decoder. For this test, affine, projective and bilinear transformations were used, and the residue was encoded with reference software HM-14 + REext-7.2.

Since a single geometric transformation for a whole slice may not be representative, a test was performed segmenting a slice into different regions, according to the pixel values. Then, a geometric transformation is applied for each region, minimizing the prediction error, and the residue is generated. For the segmentation, the K-Means [43] algorithm is used, with the number of regions to detect being selected beforehand. For this test, only bilinear transformations were used with three, four and five segmented regions.

A similar approach has been proposed in [42] for disparity estimation in multi-view video encoding with the standard HEVC.

Table 4.1: Comparison of the motion estimation results with geometric transformations (in bpp).

Sequence	Orig	Diff Res	GT Affine	GT Projec	GT Bilinear	GT 3 Reg Bilinear	GT 4 Reg Bilinear	GT 5 Reg Bilinear
CT_Aperts	0.731	**0.693**	0.697	0.709	0.705	0.722	0.728	0.741
CT_carotid	1.433	**1.314**	1.318	1.318	1.318	1.362	1.369	1.368
CT_skull	1.776	**1.514**	1.518	1.518	1.522	1.527	1.526	1.525
CT_wrist	1.009	**0.903**	0.907	0.907	0.935	0.958	0.970	0.965
MR_liver_t1	2.048	**1.867**	1.871	1.872	1.872	1.867	1.867	1.867
MR_liver_t2e1	1.512	**1.247**	1.251	1.251	1.251	1.247	1.247	1.247
MR_ped_chest	1.545	**1.366**	1.369	1.369	1.369	1.366	1.369	1.368
MR_sag_head	1.736	**1.523**	1.527	1.610	1.527	1.545	1.539	1.541
Average	1.474	**1.303**	1.307	1.319	1.312	1.324	1.327	1.328

Table 4.1 shows the results of the encoded residues generated in these tests, in comparison to the encoding of the original sequences and the residue from the pixel-wise difference,

using the same version of HEVC. Contrary to what was expected, the segmentation of the slice in regions presented slightly worse results than applying a geometric transformation to the whole slice, with better results obtained when using less regions. For the first test, the affine transformation provides better results than bilinear and projective, respectively. However, when comparing to the pixel-wise difference, it still presents a slight increase in the bitrate. This increase becomes much more significant after adding the extra information of the geometric transformations, which is not presented in these results.

4.2.2 Block Matching with Geometric Transformations (BMGT)

Given the previous results, a more complete approach was implemented, by complementing the BMA of HEVC with the Block Matching with Geometric Transformations (BMGT) [40]. BMGT allows the four corners of the block to be moved independently, inside a search window. Every combination of the four corners' coordinates results in a geometric transformation. The geometric transformation which minimizes the error is then selected, and the corresponding four motion vectors (one for each corner of the block) are transmitted. BMGT has been successfully applied to video [40], stereoscopic images and video [41], and multi-view video sequences [42]. An example of BMGT is shown in Figure 4.2.

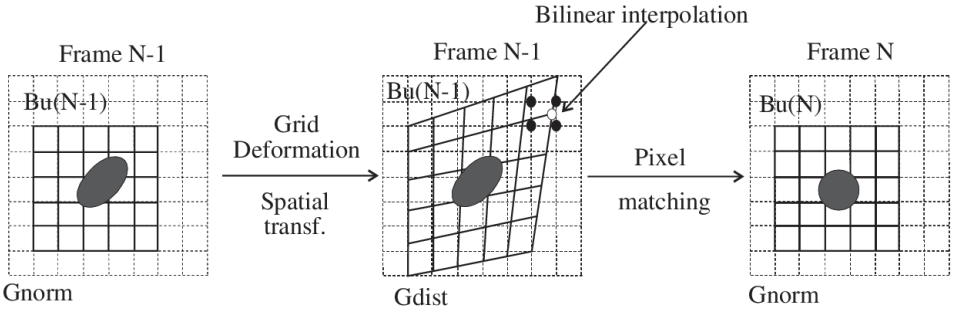


Figure 4.2: Exemplification of Block Matching with Geometric Transformations (BMGT) [40].

Comparing BMGT to BMA, in the former the four corners are moved independently, resulting in an increase in computational complexity. It is also noticeable that it is necessary to transmit four motion vectors instead of only one, since in BMA the 4 corners have the same motion vector. On the other hand, since the BMA is a particular case of the BMGT, the prediction error of the BMGT is potentially inferior, or in the worst case scenario the same as BMA.

With this in mind, the BMGT was implemented in the HEVC reference software HM-14 + RExt-7.2. Due to the way it is implemented, in order to allow the use of BMGT, the

half and quarter pixel prediction was disabled. Then, for motion estimation, the BMA is applied as usual, with the default search window. Obtained the coordinates of the block with minimal error, the BMGT is applied, using a search window of 9×9 pixels centred in each corner of the block with those coordinates. Each combination of the four corners specifies a bilinear transformation, which is then applied to the whole block. This way, the best transformation is selected to minimize the prediction error. After having the 4 motion vectors, only one is encoded in the bitstream, as it would be with BMA, with the remaining three being encoded with an Adaptive Arithmetic Coder (AAC) outside of HEVC.

Table 4.2 shows the results in bpp for the original HEVC REext Random Access (R.A.), for HEVC REext with quarter and half pixel prediction disabled, and for the implemented BMGT. The last column shows the extra information necessary for the BMGT (the three additional motion vectors) encoded with an AAC. By disabling quarter and half pixel prediction, the compression ratio slightly reduces, as it would be expected. However, the implementation of BMGT following BMA was expected to improve the compression ratio, and instead there is a slight decrease. On top of that, it is necessary to add the extra information, decreasing the effectiveness of BMGT.

Table 4.2: Comparison of the results of the implemented BMGT in HEVC REext (in bpp).

Sequence	Original	No Q/H Prediction	BMA + BMGT	MV's coded w/ AAC
CT_Aperts	**0.731**	0.743	0.745	0.010
CT_carotid	**1.433**	1.451	1.455	0.011
CT_skull	**1.776**	1.788	1.790	0.006
CT_wrist	**1.009**	1.021	1.022	0.003
MR_liver_t1	**2.048**	2.061	2.062	0.003
MR_liver_t2e1	**1.512**	1.521	1.523	0.007
MR_ped_chest	**1.545**	1.563	1.564	0.003
MR_sag_head	**1.736**	1.753	1.754	0.003
Average	**1.474**	1.488	1.489	0.006

4.2.3 Mesh-based Motion Estimation

Following the previous methods, a mesh-based technique was implemented. Similarly to the first method, after applying geometric transformations as motion estimation outside of the encoder, the prediction residue is obtained and encoded with HEVC REext, using the reference software HM-16.2.

This method consists of generating a mesh of triangles defined by points in the slice [44]. These points that characterise the mesh can be manually defined or obtained from the information of the image (edges). The points are then independently moved along a search

window, resulting in the triangles undergoing affine geometric transformations. Once the coordinates of the points that minimize the prediction error are determined, the residue is acquired. A similar mesh-based method has been successfully applied to stereo video encoding in [44].

With this method, a more efficient motion estimation can be achieved since it allows for a expansion of a certain area of the image while some other area could be contracted, which is a common feature of medical images. Also, this method does not produce blocking effect, since it is a continuous method, that is, all triangles in the mesh are contiguous, and when one is transformed, all the adjacent triangles are moved accordingly.

Several tests were conducted with different search ranges and number of points, with some being in fixed positions and others placed in the edges/contours of the image, as can be seen in Figure 4.3. Figures (a), (b) and (c) show a fixed grid with 9, 17 and 25 points, respectively. In figure (d) is represented a fixed grid with 9 points + 16 dependent on the edges in the slice, and (e) is a grid with 17 + 32 dependent. As for figures (f) and (g), it is used 1 fixed point in the centre + all the points dependent on the detected edges in the periphery of the slice with a minimal distance of 10 pixels and 5 pixels, respectively. Finally, figures (h) and (i) show meshes with all the determined points dependent on all edges with minimum distance of 10 and 5 pixels, respectively. As the number of points increases, the mesh represents the image with more accuracy, being expected that the motion estimation results in a smaller prediction error, and thus a better coding efficiency. However, the computational complexity and the additional information (motion vectors) increases.

Table 4.3: Comparison of the results of the mesh-based motion estimation (in bpp).

			Mesh Residue				
Sequence	Orig	Diff Res	9 Fix	17 Fix	9 Fix + 16 Dep	Periphery Dist = 10	Edges Dist = 10
MR_liver_t2e1	1.509	**1.228**	1.231	1.371	1.245	1.594	1.718

Table 4.3 shows the results for some of the tests performed with mesh-based motion estimation, in comparison to the coding of the original sequence and the pixel-wise difference residue (no motion estimation). Due to the large amount of tests that were performed, it was used only one test sequence. It can be seen that the best result was achieved for the test with fewer points: only 9 fixed points. Even though this method performs better than the coding of the original sequence, it still does not surpass the coding of the pixel-wise difference residue. Despite what was expected, by increasing the number of points in the mesh or by using points more relevant to the slice, as can be seen in Figure 4.3, the bitrate increases, even when the additional rate for the motion vectors is not taken into account.

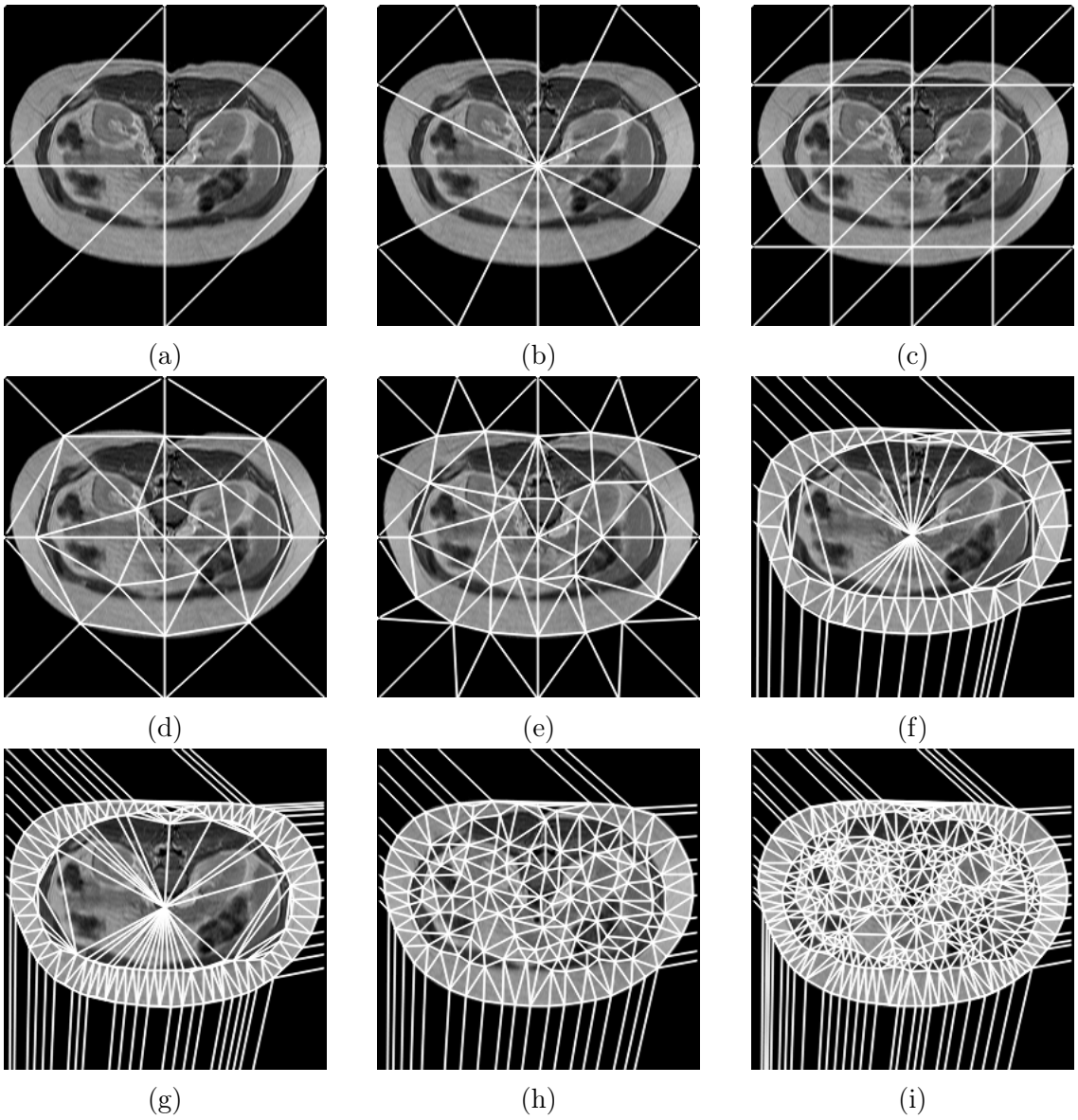


Figure 4.3: Examples of meshes with different number of points, both fixed and edges dependent.

Chapter 5

Least-Squares Prediction (LSP)

In the previous chapter, an approach to improve the lossless encoding of medical volumes using Geometric Transformations (GT) for motion estimation in HEVC was tested, which did not produce satisfactory results. Given the importance of prediction in lossless compression with HEVC, in this chapter, a new approach based on Least-Squares Prediction (LSP) was developed.

When HEVC is used for lossless compression, the transform and quantization steps are not applied to the prediction residue, which is directly encoded using CABAC [1]. This means that it is fundamental to perform an accurate prediction, in order to produce a residue with low energy, thus resulting in a high compression efficiency.

State-of-the-art lossless image encoders, such as JPEG-LS, CALIC and MRP, are characterized by their prediction schemes, which are performed on a pixel-by-pixel basis. As seen in Chapter 2, JPEG-LS and CALIC predictors, MED and GAP, respectively, perform an edge detection and then determine the prediction value accordingly. As for MRP, a linear prediction, similar to LSP, is used. LSP is a more robust alternative to edge detection and orientation estimation [10].

5.1 Description of LSP

The Least-Squares Prediction (LSP) method aims to find the prediction $\hat{X}(n_0)$ of a current pixel $X(n_0)$ by adaptively linear filtering a selected set of N neighbouring pixels, $X(n_i)$, which form the filter support. The optimal filter coefficients a_i are estimated using a previously encoded region, commonly referred to as a training window, of size $M = 2T(T+1)$ pixels, as shown in Figure 5.1. The predicted value of the current pixel is given

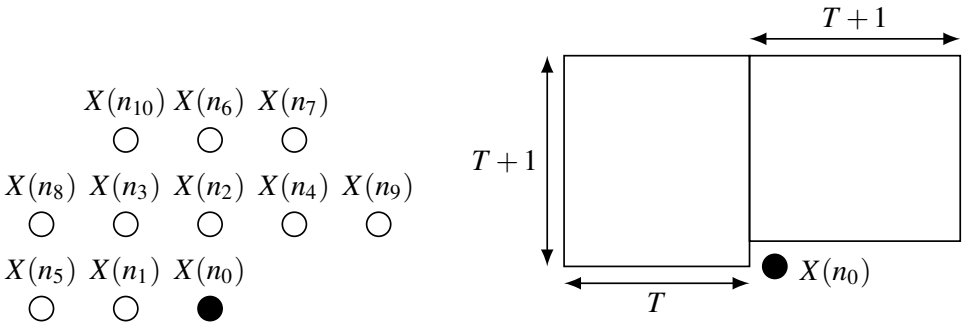


Figure 5.1: Filter support and training window.

by:

$$\hat{X}(n_0) = \sum_{i=1}^{N} a_i X(n_i), \tag{5.1}$$

where N is the filter order. The selected value of T must satisfy the condition $N << M$.

Considering the training window as an $M \times 1$ column vector $\vec{y} = [X(n_1) \cdots X(n_M)]^T$, by applying the filter support to every pixel of the training window, we form the $C_{M \times N}$ matrix:

$$C = \begin{bmatrix} X(n_{1,1}) & \cdots & X(n_{1,N}) \\ \vdots & \ddots & \vdots \\ X(n_{M,1}) & \cdots & X(n_{M,N}) \end{bmatrix}. \tag{5.2}$$

The prediction coefficients are determined using Least Squares optimization, that is, the optimal solution is the one that minimizes the sum of the squared errors for all equations, represented by:

$$\min\{\|\vec{y} - C\vec{a}\|^2\}. \tag{5.3}$$

A well-known closed form solution to this problem [10] is given by:

$$\vec{a} = (C^T C)^{-1}(C^T \vec{y}). \tag{5.4}$$

5.2 State-of-the-Art

Least Squares-based prediction techniques have been successfully used in different approaches [10–13]. By using a training window with causal neighbouring pixels, LSP can accurately predict the pixels in the current prediction unit, even in cases where traditional prediction methods fail, as shown in [10]. For lossless compression of natural images, due to its ability to correctly predict arbitrarily oriented edges, LSP produces better results

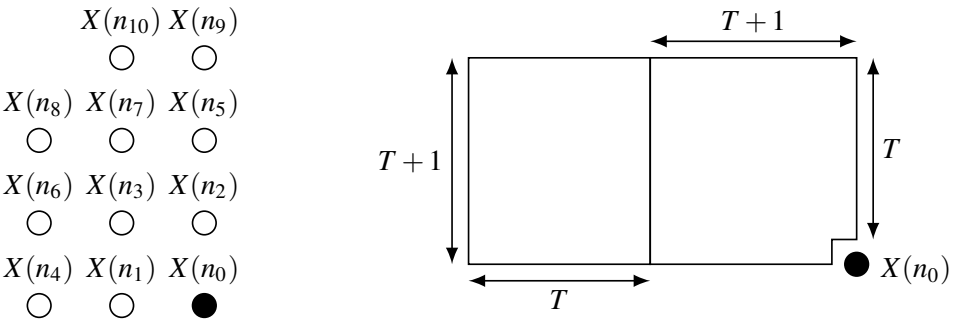


Figure 5.2: Causal filter support and training window as proposed in [11].

than state-of-the-art predictors, such as MED and GAP, used by JPEG-LS and CALIC, respectively.

LSP may also be used in block based encoders, such as HEVC. However, when the current pixel to be predicted belongs to the pixel columns at the right extreme of the current PU, part of the filter support belongs to a block not yet encoded. In those cases, a causal filter support can be used instead, as shown in Figure 5.2. This adaptation of the training window for the current pixel position, proposed by Danillo et al. [11], resulted in better prediction coefficients, improving the LSP capability for block prediction.

This method was also used for stereo image encoding, by extending the filter support to both views, after disparity compensation with template matching [12]. In this case, the pixels located below and to the right of the current pixel in the other view, which has already been encoded, can be used for the filter support.

In order to improve the temporal prediction, LSP was proposed for video encoding in [13], to replace the much used motion estimation and compensation based on block-matching. Due to its edge prediction ability, LSP can also be used to determine the prediction of the current pixel from the previous frame, by extending the filter support. By using an adaptive filter support, dependent on the type of motion, LSP can be more effective, with the additional benefit of not needing to transmit overhead.

5.3 Implementation in HEVC

To improve the lossless compression efficiency of HEVC, the LSP based mode was developed for HEVC Range Extension (main-REext) profile, including some improvements for both intra and inter prediction, which are described in the following subsections. By replacing one of the less used intra-prediction modes (Intra direction mode 27), the HEVC algorithm modifications are minimal. Although the LSP mode is applied to the whole PU, Least Squares optimization is performed on a pixel-by-pixel basis, that is, the predictor

coefficients are estimated for each pixel of the PU, individually. Since this operation uses an implicit method, which relies only on the causal pixels, it can be performed in both the encoder and decoder, without the need of transmitting additional information.

5.3.1 Intra Improvements to LSP

Since LSP is applied for block prediction, the causal filter support and training window adaptation, shown in Figure 5.2 and proposed in [11], were implemented. When the LSP mode is used, every pixel of the PU is predicted individually. For the pixels whose filter support contains pixels that belong to a block not yet encoded, which occurs in the right edge of the PU, this causal filter support is used, while for the remaining pixels is applied the regular one. Similarly to the intra directional prediction modes, for the prediction units at the left and upper limits of the image, since there are not enough causal pixels, the LSP mode is not applied.

In lossy applications, the decoder does not always have the original pixel values, only an approximation. So, when encoding a pixel further in the PU, the filter support uses the predicted values of the causal pixels of the same PU, instead of the original pixel values. By doing this, if the predicted value of a pixel has a significant difference to the original value, this error may be propagated to the following pixels, leading to a high prediction error, and resulting in the LSP mode not being selected. Thus, in order to improve the efficiency of LSP for lossless coding, the original pixel values are used instead for every pixel of the PU, as they will also be available in the decoder due to being lossless compression.

Among the characteristics of medical images such as CTs and MRs are the smooth regions and sharp edges. Due to these regions, when LSP is applied, the matrix C, given by Equation 5.2, can feature many repeated or linearly dependent rows. This implies that, in practice, the condition $N << M$ is not verified. As such, over-fitting may occur, that is, the filter coefficients become overly adapted to the training region, such that they describe small variations that could be attributed to random errors or noise, instead of their subjacent relationship [45], which is the edge to predict. In order to minimize this issue and improve the prediction efficiency of LSP, Tikhonov regularization [46, 47] is used. To this extent a regularization term is added to the minimization problem:

$$\min\{\|\vec{y} - C\vec{a}\|^2 + \|\Gamma\vec{a}\|^2\}, \tag{5.5}$$

and the regularized solution is computed according to

$$\vec{a} = (C^T C + \Gamma^T \Gamma)^{-1}(C^T \vec{y}), \tag{5.6}$$

where Γ is the Tikhonov matrix, defined as $\Gamma = \alpha I$, with I being the identity matrix and α the regularization parameter. The Tikhonov regularization improves stability of the method by giving preference to solutions with smaller norms. A set of experimental tests were performed in order to determine this parameter. These tests show that small variations of α do not have a significant impact in the results, with $0.1 \leq \alpha \leq 5$ producing the best results.

5.3.2 Inter-Slice Improvements to LSP

Due to the characteristics of the volumetric images such as MRs and CTs, in which human organs and tissues are represented, adjacent slices are highly correlated. By expanding the LSP support to include pixels from the previous and the following slices, the prediction accuracy can be improved. This extended support is used only in non-Intra slices, with the previous slice being the first one from Reference_List_0, and the following slice being the first one from Reference_List_1 of HEVC. This approach adapts LSP prediction for volumetric datasets by exploiting the reference frame structure of HEVC, leading to an increased overall coding efficiency.

While for every PU it is necessary to resort to the causal filter support and training window in the current slice, since the PU to the right has not yet been encoded, the reference slices are fully available, and as such the extended filter support can still be used as normal. However, if the pixel to predict belongs to a right or bottom border, not all the pixels from the filter support exist, thus it is used a causal filter support instead.

5.3.3 Experimental Results

The proposed lossless prediction method was implemented into the HEVC reference software and its performance was evaluated using several volumetric medical image sets. The latest release of the HEVC reference software HM 16.4 was used, with the Random Access Main profile and the RExt profile. To guarantee lossless coding, QP was set to 0, and both TransquantBypassEnableFlag and CUTransquantBypassFlagForce were set to 1. For the RExt profile, the parameter CostMode was set to lossless. Default values were used for the remaining configuration parameters. For all the other encoders, the default configurations for lossless encoding were used.

The proposed method LSP mode uses a filter support which was empirically determined after testing several geometries for various images. The best results were achieved for a triangular template with 12 pixels in the current slice, and a diamond shaped template with five pixels centred in the pixel co-located to the pixel to predict in the previous

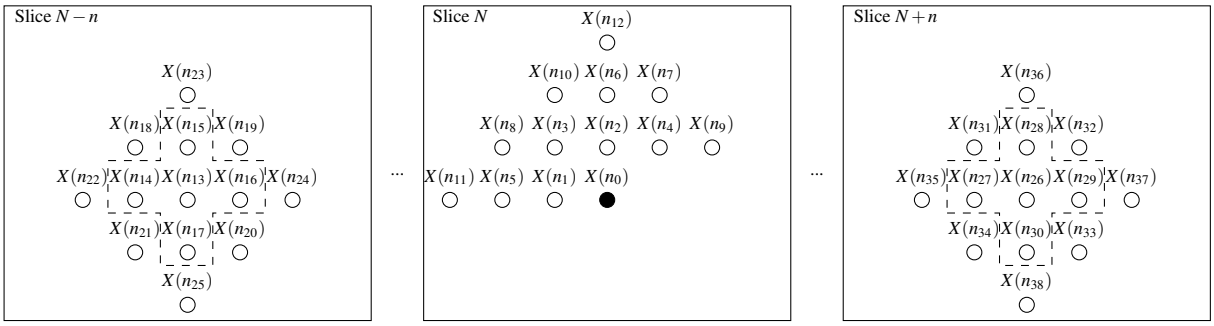


Figure 5.3: LSP filter support extended for volumetric images using the 5 pixels within the dashed line in each additional slice.

and following slices, as depicted in Figure 5.3 by the dashed line. Using more than five pixels did not increase the prediction performance, but presented higher computational complexity, therefore the five pixel region was adopted as the best choice. The training window size was set to $T = 7$.

Given the Group of Pictures (GOP) structure of HEVC, the reference slices are not always at the same distance. In order to assess the influence of this distance in the number of coefficients of the extended filter support, some tests were performed by varying the filter order in the reference slices in accordance to their distance to the current slice. The results were very similar, but using five fixed coefficients in each reference slice still provide slightly better results.

The encoding performance of the proposed method is compared to DICOM [3] recommended encoders, JPEG2000 [5] and JPEG-LS [4], the extension for volumetric images JP3D [9], the state-of-the-art lossless image encoders CALIC [7] and MRP [8], as well as the H.264/AVC [6] and a standard version of HEVC [1], with and without Range Extensions [29]. The compression results are presented in Table 5.1, in bits per pixel (bpp).

The first column shows that H.264/AVC presents the worst performance in lossless compression of volumetric medical images. This is justified since this video encoder was primarily designed for lossy video compression, and for lossless encoding only the Intra profile is available, which further limits its performance. The next four columns show the performance of the lossless image encoders: JPEG2000, JPEG-LS, CALIC and MRP. JPEG2000 has the worst performance, which can be explained by the use of a wavelet transform instead of a pixel-wise predictor, such as the other three. It can also be seen that the more advanced the prediction method, the better the results obtained, with MRP outperforming both CALIC (GAP) and JPEG-LS (MED). The next column presents the results for JP3D, which surpass those of image encoders, as expected, since it exploits inter-slice redundancy. However, when compared to MRP, the compression gain is not

Table 5.1: Lossless performance comparison of the proposed method with HEVC and other state-of-the-art encoders (in bpp).

Sequence	H.264 Intra	JPEG 2000	JPEG LS	CALIC	MRP	JP3D	HEVC	HEVC REext	Proposed
CT_Aperts	1.193	1.261	1.058	0.998	0.775	0.941	0.826	0.728	**0.620**
CT_carotid	2.062	2.019	1.778	1.684	1.374	1.547	1.587	1.425	**1.204**
CT_skull	3.183	2.991	2.761	2.628	2.329	2.088	1.905	1.766	**1.562**
CT_wrist	1.911	1.757	1.627	1.550	1.173	1.238	1.155	1.002	**0.828**
MR_liver_t1	3.489	3.256	3.160	3.022	2.582	2.356	2.392	2.052	**1.764**
MR_liver_t2e1	2.806	2.572	2.418	2.269	1.722	1.745	1.726	1.510	**1.196**
MR_ped_chest	3.080	3.021	2.937	2.789	2.337	2.071	1.699	1.534	**1.367**
MR_sag_head	2.635	2.905	2.582	2.519	2.279	2.160	1.873	1.748	**1.617**
Average	2.545	2.473	2.290	2.183	1.821	1.768	1.645	1.471	**1.270 (-13.7%)**

Table 5.2: Breakdown of the incremental performance of the proposed method and the improvements to LSP.

	HEVC REext	LSP	LSP+Original Pixels	LSP+Extended Filter Support	LSP+Tikhonov Regularization
Average	1.471	1.469 (-0.1%)	1.396 (-5.1%)	1.277 (-13.2%)	**1.270 (-13.7%)**

very relevant, thanks to the effectiveness of MRP prediction techniques.

The performance of HEVC and HEVC with Range Extension can also be compared in Table 5.1. It can be seen that using the Range Extension profile for lossless coding results in an expected bitrate reduction. The use of the SAP lossless encoding technique led to a reduction of the bitrate, on average 10.6%. The last column shows that the proposed method achieves a significant reduction of bitrate, 13.7% on average, in comparison to the REext profile of HEVC. This major improvement of the encoding efficiency is mainly due to the inter-slice prediction efficiency of the proposed method, since it exploits the redundancy of the volumetric medical images. When compared to DICOM recommended encoders, as well as state-of-the-art lossless encoders, HEVC with the proposed method presents a much higher encoding efficiency, decreasing the bitrate by 44.5% and 30.3%, when compared to JPEG-LS and MRP, respectively.

To better evaluate the impact of each of the techniques used in the proposed method, we performed some tests by enabling each of the improvements to LSP incrementally. The results of these experiments, presented in Table 5.2, show that when the LSP is introduced in HEVC, using a straightforward implementation in HEVC with REext profile only a marginal gain is obtained (0.1%). However, when adapting LSP for lossless coding using the original pixel values for the filter support this gain becomes much more significant (5.1%). The coding efficiency has major improvements after extending the LSP filter support to a previous and a following slices (bitrate reduction of 13.2%). Finally, using

Tikhonov regularization further increases the bitrate savings to 13.7%, when comparing to the unmodified HEVC with REext profile.

5.3.4 Computational Complexity

In the current implementation of the proposed method in HEVC, the computational complexity was not taken into account, as it was not a primary objective of this work. In its current form and implementation, the encoding time using the proposed method for all eight sequences was 22.4 times the encoding time using the original HEVC with the REext profile. However there is room for algorithmic optimization and improvement of the proposed method and it is foreseeable that its computational complexity can be reduced significantly.

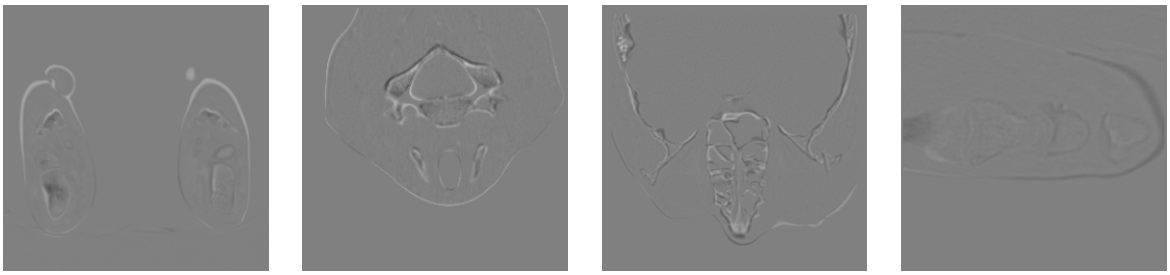
5.4 Inter Slice Predictor

As seen before, the high redundancy between slices present in this type of medical images can be exploited to improve the compression efficiency. In Chapter 3, to take advantage of this similarity, some techniques were tested outside of the encoders, and it was observed that using a zero-order inter-slice predictor was possible to obtain better results.

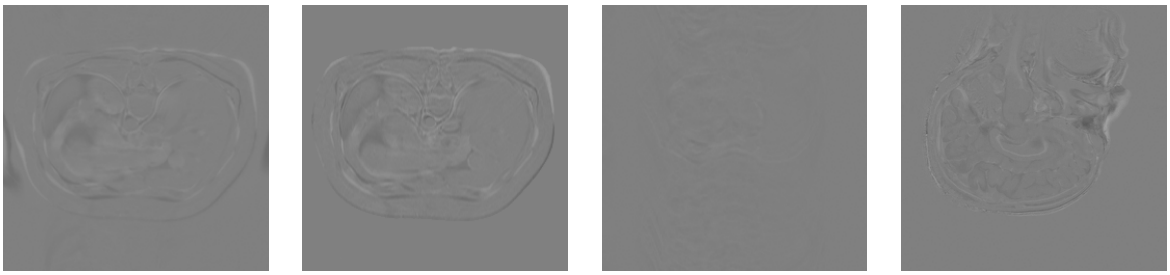
The result of implementing this predictor is a slice-wise difference, between co-located pixels in adjacent slices, which corresponds to a residue to be encoded. Figure 5.4 shows a slice in the middle of the resulting residue for each volumetric image. It can be seen that the resulting residue contains high-frequency regions, mainly along the organs edges. Given its strength in predicting arbitrarily oriented edges, the proposed LSP can introduce further improvements to the compression of this residue, instead of the original medical images.

5.4.1 Experimental Results

Table 5.3 shows the results of encoding the residue resulting from the slice-wise difference for all the encoders, in comparison to the proposed method. When comparing to the previous results from Table 5.1, it can be seen that major improvements for all the encoders were obtained, with the exception of JP3D, which presents a small gain. JPEG2000 and JP3D present the worse results, followed by H.264/AVC. JPEG-LS and CALIC present similar results, with CALIC having a small advantage. MRP brings a large improvement, surpassing HEVC, and getting performances close to those of HEVC REext. This can be



(a) CT_Aperts (b) CT_carotid (c) CT_skull (d) CT_wrist



(e) MR_liver_t1 (f) MR_liver_t2e1 (g) MR_ped_chest (h) MR_sag_head

Figure 5.4: Middle slice of the inter-slice pixel-wise difference for each medical volume [18].

Table 5.3: Lossless performance comparison of the proposed method with state-of-the-art encoders (in bpp), for the residue resulting from the slice-wise difference.

Sequence	H.264 Intra	JPEG 2000	JPEG LS	CALIC	MRP	JP3D	HEVC	HEVC REext	Proposed
CT_Aperts	0.798	0.938	0.794	0.774	0.632	0.942	0.722	0.673	**0.605**
CT_carotid	1.511	1.592	1.396	1.355	1.147	1.472	1.366	1.272	**1.144**
CT_skull	2.033	2.137	1.974	1.899	1.660	1.984	1.511	1.445	**1.348**
CT_wrist	1.112	1.212	1.043	1.054	0.851	1.218	0.905	0.861	**0.784**
MR_liver_t1	2.189	2.255	2.070	2.020	1.765	2.283	1.986	1.852	**1.663**
MR_liver_t2e1	1.789	1.818	1.685	1.604	1.310	1.693	1.328	1.228	**1.059**
MR_ped_chest	1.639	1.815	1.631	1.586	1.365	1.916	1.346	1.302	**1.237**
MR_sag_head	1.988	2.185	2.001	1.979	1.803	2.113	1.580	1.510	**1.420**
Average	1.632	1.744	1.574	1.534	1.317	1.703	1.343	1.268	**1.157 (-8.8%)**

explained since the prediction technique performed by MRP is similar to LSP, and thus takes advantage of its edge directed property to better encode the residue. Nevertheless, the proposed LSP based mode still has the best performance, reaching a bitrate reduction of 8.8%, when compared to the second best (HEVC REext).

5.5 Other Approaches

Following the good results obtained using LSP with the filter support extended to previous and following slices, a few techniques were implemented aiming to improve the current LSP.

5.5.1 LSP with Motion Vectors

When using LSP with the extended filter support, pixels in the previous and following frames are centred in the pixel collocated to the one being encoded in the current frame. However, this does not take into account the movement of the organs/tissues between slices, that is, the region being encoded in the current frame might have an offset, by many pixels, relative to other frames. Thus, in some cases, the extended filter support may not improve the prediction.

In order to minimize this issue, a modification to LSP involving Motion Vectors (MVs) was tested. The filter support in the reference slices is moved in according to the respective motion vector, so as to place it in the more relevant region. The used motion vectors are determined by HEVC motion estimation, which is performed prior to the intra prediction modes, including the proposed LSP, when selecting the best prediction mode for the current block to be encoded (Coding Unit (CU)).

The downside of this approach is the need to transmit the motion vectors. To determine if motion vectors are used or not, the LSP mode is tested both with and without MVs, and the mode with lowest rate-distortion cost is selected.

The results for this approach are shown in Table 5.4, in comparison to the proposed LSP presented in Sections 5.3 and 5.4. Although the introduction of motion vectors in LSP does improve the prediction, albeit rather minimally, when their required cost is added the encoding efficiency diminishes, both for the encoding of the original sequences and of the difference residue. As a result, the use of MVs in LSP was not researched further.

Table 5.4: Comparison of the performance of the proposed LSP with and without motion vectors (in bpp).

	Original Sequences		Slice-wise Difference Residue	
	LSP	LSP+MVs -> with Cost	LSP	LSP+MVs -> with Cost
CT_Aperts	**0.620**	0.618 -> 0.620	**0.605**	0.603 -> 0.605
CT_carotid	**1.204**	1.201 -> 1.215	**1.144**	1.141 -> 1.155
CT_skull	**1.562**	1.560 -> 1.566	**1.348**	1.346 -> 1.352
CT_wrist	**0.828**	0.827 -> 0.832	**0.784**	0.783 -> 0.786
MR_liver_t1	**1.764**	1.763 -> 1.765	**1.663**	1.662 -> 1.663
MR_liver_t2e1	**1.196**	1.192 -> 1.197	**1.059**	1.058 -> 1.059
MR_ped_chest	**1.367**	1.366 -> 1.367	**1.237**	1.237 -> 1.238
MR_sag_head	1.617	1.616 -> **1.616**	1.420	1.419 -> **1.419**
Average	**1.270**	1.268 -> 1.272	**1.157**	1.156 -> 1.160

5.5.2 Sparse-LSP

Traditional LSP typically uses a small filter context region. However, in some cases it might not contain the most relevant pixels to the one being predicted.

Sparse prediction methods aim to address this issue by searching in a large causal region for the few best predictors. Usually these methods perform well for images with similar or repeated patterns and regions.

In [48] a sparse implementation of Least-Squares Prediction (LSP) was proposed. As shown in Figure 5.5, contrary to the traditional LSP, the Sparse-LSP has a large filter context region and a fixed training window for the whole block. The sparse condition is enforced by a k-Nearest Neighbours (k-NN) technique, which selects a few positions for non-null coefficients of the filter support. The k-NN algorithm resorts to template matching of the training window in the filter context region to determine the locations which are most correlated to the current block.

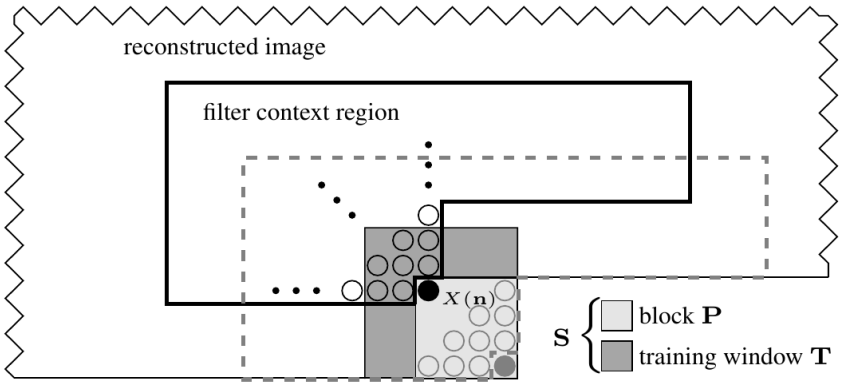


Figure 5.5: Filter support region and training window of Sparse-LSP as proposed in [48].

The non-null coefficients are then estimated similarly to LSP, as described in Section 5.1. The main difference is that they are determined once for all the pixels of the block.

In order to evaluate the performance of Sparse-LSP in comparison to the proposed LSP, it was implemented in HEVC replacing the proposed LSP. The same improvements were introduced in Sparse-LSP, namely the use of the original pixel values instead of the predicted ones, as well as the extension of the filter support to a previous and a following slice.

The extension to the filter support was performed by adding a filter context region all around the collocated block in both a previous and following slices. The k-NN technique performs the template matching for the three regions (in previous, current and following slices) and selects the k best positions. Several tests were performed in order to determine the sparsity condition parameter k, with the best results being obtained for k=10.

Table 5.5 shows the results for this approach in comparison to the proposed LSP presented in Sections 5.3 and 5.4. It can be seen that although the implementation of Sparse-LSP does improve the standard HEVC with Range Extension, the proposed LSP still exhibits a significant advantage. This can be explained by there not being enough repeated patterns in these medical image sequences, and even though there are similar structures, since they still present differences, the Sparse-LSP presents reduced improvements. As such, the use of a sparse implementation of LSP was dismissed.

Table 5.5: Comparison of the performance of the proposed LSP with the implemented Sparse-LSP (in bpp).

	Original Sequences		Difference Residue	
	LSP	Sparse-LSP	LSP	Sparse-LSP
CT_Aperts	**0.620**	0.702	**0.605**	0.655
CT_carotid	**1.204**	1.373	**1.144**	1.248
CT_skull	**1.562**	1.707	**1.348**	1.425
CT_wrist	**0.828**	0.935	**0.784**	0.832
MR_liver_t1	**1.764**	1.967	**1.663**	1.784
MR_liver_t2e1	**1.196**	1.409	**1.059**	1.177
MR_ped_chest	**1.367**	1.467	**1.237**	1.276
MR_sag_head	**1.617**	1.709	**1.420**	1.495
Average	**1.270**	1.409	**1.157**	1.236

Chapter 6

Conclusions

The exponential growth of medical imaging data induces the need to develop efficient compression methods. However, this compression is desired to be lossless so that no relevant information for diagnosis is lost. DICOM, the standard that defines, among others, the way medical information is stored, recommends the use of lossless encoders such as JPEG-LS and JPEG2000. However, for volumetric images, such as MR images or CT scans, these encoders are not the most efficient since they do not exploit the inter-slice redundancy present in these volumes. More recently, JP3D has been developed for volumetric images, and by exploiting the redundancy across the three dimensions, presents a much superior performance over the remaining intra encoders.

The newest video encoding standard HEVC, unlike its predecessor H.264, allows lossless coding using inter-frame prediction, or motion compensation, and can thus be applied to the volumetric medical images, resulting in a great compression efficiency. However, since it was not refined for lossless encoding, there is still room for improvement.

In order to investigate the importance of inter-slice prediction in medical volumes, and to evaluate the best way to exploit this redundancy, several processing techniques were applied to medical images. It was verified that performing inter-slice prediction prior to the encoding resulted in bitrate reductions of around 30% for all intra encoders, and nearly 14% for HEVC, which presents the best results.

MR images and CT scans consist on a set of slices representing a certain anatomical structure. Due to their characteristics, the differences between adjacent slices can correspond to rotations, translations or expansions, which can be represented by geometric transformations. As such, different approaches using geometric transformations were developed in order to improve the inter-slice prediction of HEVC.

Given the great performance of a simple zero-order inter-slice predictor, motion estima-

tion using geometric transformations was applied to the full slices, prior to encoding, however, the compression efficiency was not improved. Following this approach, a more complex one was implemented, block matching with geometric transformations. The current motion estimation of HEVC is a particular case of BMGT, in which only a translation is applied. With this in mind, BMGT was added to HEVC, but did not yield positive results. In the third approach, the geometric transformations were applied using a mesh-based motion estimation, but despite several configurations being tested, none of them produced improvements.

Thus, the focus of this work was aimed at pixel-wise prediction in HEVC. The SAP technique introduced in the Range Extension profile proved its efficiency, leading to several new proposed techniques. The Least-Squares Prediction (LSP) has been widely used in image, stereo and video encoding, due to its edge prediction ability.

In this work, an LSP-based prediction mode was implemented in HEVC for lossless compression of volumetric medical images. This proposed mode improved LSP for lossless encoding by using the original pixels, instead of their predicted values, to predict the following ones within the coding block. This method has also been extended to exploit inter-slice redundancy using the nearest previously encoded images for LSP prediction, thus increasing its efficiency for the compression of volumetric medical images. To increase the robustness of the linear filter parameters estimation, Tikhonov regularization has been included in the algorithm.

Experimental results show that the proposed HEVC prediction mode achieves considerably higher lossless compression efficiency for encoding volumetric medical images, achieving an average bitrate reduction of approximately 13.7% over standard HEVC, and between 45% and 49% over DICOM recommended coding methods.

In order to improve the proposed method, motion estimation was used to determine the filter support in the neighbouring slices. However, the performance gain was not sufficient to cover the motion vectors cost. Another approach was tested with LSP using a sparse implementation, which did not surpass the compression performance of the proposed technique.

The LSP method proposed in this dissertation allows an improvement of the HEVC standard for lossless compression, becoming the most efficient state-of-the-art encoder of volumetric medical images.

Bibliography

[1] 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, 2012. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6316136

[2] J. Taquet and C. Labit, "Hierarchical oriented predictions for resolution scalable lossless and near-lossless compression of ct and mri biomedical images," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2641–2652, May 2012.

[3] *Digital Image and Communications in Medicine (DICOM)*. Strategic document, 2014.

[4] M. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1309–1324, 2000. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=855427

[5] "Information technology: JPEG 2000 image coding system, part 1: Core coding system," *document 15444-1, ISO/IEC, 2000*, 2000.

[6] G. J. Sullivan, P. N. Topiwala, and A. Luthra, "The H.264/AVC advanced video coding standard: overview and introduction to the fidelity range extensions," in *Applications of Digital Image Processing XXVII*, vol. 5558 of Proceedings Of SPIE. SPIE, November 2004, pp. 454–474. [Online]. Available: http://dx.doi.org/10.1117/12.564457

[7] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Transactions on Communications*, vol. 45, no. 4, pp. 437–444, Apr 1997.

[8] I. Matsuda, Y. Umezu, N. Ozaki, J. Maeda, and S. Itoh, "A lossless coding scheme using adaptive predictors and arithmetic code optimized for each image," *Systems and Computers in Japan*, vol. 38, no. 4, pp. 1–11, 2007. [Online]. Available: http://dx.doi.org/10.1002/scj.20630

[9] P. Schelkens, A. Munteanu, A. Tzannes, and C. Brislawn, "Jpeg2000. part 10. volumetric data encoding," in *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*, May 2006, pp. 4 pp.–3877.

[10] X. Li and M. Orchard, "Edge-directed prediction for lossless compression of natural images," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 813–817, 2001. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=923277

[11] D. Graziosi, N. Rodrigues, E. da Silva, S. de Faria, and M. de Carvalho, "Improving multiscale recurrent pattern image coding with least-squares prediction mode," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009, pp. 2813–2816. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5414219

[12] L. Lucas, N. Rodrigues, E. da Silva, and S. de Faria, "Adaptive least squares prediction for stereo image coding," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, 2011, pp. 2013–2016. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6115872

[13] X. Li, "Least-square prediction for backward adaptive video coding," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 126–126, 2006.

[14] C. Cavaro-Menard, L. Zhang, and P. Le Callet, "Diagnostic quality assessment of medical images: Challenges and trends," in *Visual Information Processing (EUVIP), 2010 2nd European Workshop on*, 2010, pp. 277–284. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5699147

[15] H. Zaidi, "Monte carlo simulation studies of scatter correction in positron emission tomography," Ph.D. dissertation, Faculty of Science of the University of Geneva, 2000.

[16] Siemens Healthcare, "Computed tomography," March 2016. [Online]. Available: http://www.medicalradiation.com/types-of-medical-imaging/imaging-using-x-rays/computed-tomography-ct/

[17] Siemens Healthcare, "Magnetic resonance imaging," March 2016. [Online]. Available: http://www.medicalradiation.com/types-of-medical-imaging/other-types-of-medical-imaging/magnetic-resonance-imaging/

[18] C. for Image Processing Research, "Medical images datasets," March 2016. [Online]. Available: http://www.cipr.rpi.edu/resource/sequences/sequence01.html

[19] M. Adams and R. Ward, "Wavelet transforms in the JPEG-2000 standard," in *Communications, Computers and signal Processing, 2001. PACRIM. 2001*

IEEE Pacific Rim Conference on, vol. 1, 2001, pp. 160–163. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=953547

[20] T. Bruylants, A. Munteanu, A. Alecu, R. Deklerck, and P. Schelkens, "Volumetric image compression with JPEG2000," in *SPIE The International Society for Optical Engineering*, 2007.

[21] M. B. de Carvalho, E. A. Da Silva, and W. A. Finamore, "Multidimensional signal compression using multiscale recurrent patterns," *Signal processing*, vol. 82, no. 11, pp. 1559–1580, 2002.

[22] N. Francisco, N. Rodrigues, E. da Silva, M. de Carvalho, and S. de Faria, "Efficient recurrent pattern matching video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 8, pp. 1161–1173, 2012. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6193166

[23] N. C. Francisco, N. M. M. Rodrigues, and E. A. B. da Silva, "Contribuições à codificação eficiente de imagem e vídeo utilizando recorrência de padrões multiescala," Ph.D. dissertation, Universidade Federal do Rio de Janeiro, 2012. [Online]. Available: http://www02.smt.ufrj.br/~eduardo/teses/Tese_doutoramento_nelson_francisco.pdf

[24] D. Graziosi, N. Rodrigues, E. da Silva, M. de Carvalho, and S. Maciel de Faria, "Lossy and lossless image encoding using multi-scale recurrent pattern matching," *IET Image Processing*, vol. 7, no. 6, pp. 556–566, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6616264

[25] N. Francisco, N. Rodrigues, E. da Silva, M. de Carvalho, and S. de Faria, "Video compression using 3D multiscale recurrent patterns," in *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*, 2013, pp. 1412–1415. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6572120

[26] ITU-T and I. 1, *Advanced Video Coding for Generic Audiovisual Services*. ITU-T Rec. H.264 & ISO/IEC 14496-10, 2014.

[27] 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, 2012.

[28] J.-A. Choi and Y.-S. Ho, "Improved residual data coding for high efficiency video coding lossless extension," in *Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 2013 IEEE RIVF International Conference on*, Nov 2013, pp. 18–21.

[29] G. Sullivan, J. Boyce, Y. Chen, J.-R. Ohm, C. Segall, and A. Vetro, "Standardized extensions of high efficiency video coding (HEVC)," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 1001–1016, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6630053

[30] M. Zhou, W. Gao, M. Jiang, and H. Yu, "HEVC lossless coding and improvements," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1839–1843, 2012. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6317159

[31] E. Wige, G. Yammine, P. Amon, A. Hutter, and A. Kaup, "Pixel-based averaging predictor for HEVC lossless coding," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, 2013, pp. 1806–1810. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6738372

[32] E. Wige, G. Yammine, P. Amon, A. Hutter, and A. Kaup, "Sample-based weighted prediction with directional template matching for HEVC lossless coding," in *Picture Coding Symposium (PCS), 2013*, 2013, pp. 305–308. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6737744

[33] V. Sanchez, "Sample-based edge prediction based on gradients for lossless screen content coding in HEVC," in *Picture Coding Symposium (PCS), 2015*, 2015, pp. 134–138. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7170062

[34] V. Sanchez, "Lossless screen content coding in HEVC based on sample-wise median and edge prediction," in *Image Processing (ICIP), 2015 IEEE International Conference on*, 2015, pp. 4604–4608. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7351679

[35] V. Sanchez, P. Nasiopoulos, and R. Abugharbieh, "Lossless compression of 4d medical images using h.264/avc," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 2, May 2006, pp. II–II.

[36] V. Sanchez, P. Nasiopoulos, and R. Abugharbieh, "Efficient lossless compression of 4-d medical images based on the advanced video coding scheme," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 4, pp. 442–446, July 2008.

[37] R. Pizzolante and B. Carpentieri, "Lossless, low-complexity, compression of three-dimensional volumetric medical images via linear prediction," in *Digital Signal Processing (DSP), 2013 18th International Conference on*, July 2013, pp. 1–6.

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

[39] V. Sanchez, R. Abugharbieh, and P. Nasiopoulos, "3d scalable lossless compression of medical images based on global and local symmetries," in *IEEE International Conference on Image Processing (ICIP)*, Nov 2009, pp. 2525–2528.

[40] S. M. M. de Faria, "Very low bit rate video coding using geometric transform motion compensation," Ph.D. dissertation, University of Essex, 1996.

[41] N. M. M. Rodrigues, "Codificação de vídeo usando transformações geométricas e de luminância," Master's thesis, FCT, Universidade de Coimbra, 1997.

[42] R. J. S. Monteiro, "Disparity compensation using geometric transforms," Master's thesis, ESTG, Instituto Politécnico de Leiria, 2014.

[43] G. A. F. Seber, *Multivariate Observations*. JOHN WILEY & SONS INC, 2004. [Online]. Available: http://www.ebook.de/de/product/6972794/g_a_f_seber_multivariate_observations.html

[44] G. Jorge, "Codificação compatível de vídeo 3d com o algoritmo hevc," Master's thesis, ESTG, Instituto Politécnico de Leiria, 2014.

[45] G. G. Irina Rish, *Sparse Modeling*. Taylor & Francis Inc, 2010. [Online]. Available: http://www.ebook.de/de/product/10939778/irina_rish_genady_grabarnik_sparse_modeling.html

[46] A.-I. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill Posed Problems (Scripta series in mathematics)*. Vh Winston, 1977.

[47] A. N. Tikhonov, A. Goncharsky, V. Stepanov, and A. G. Yagola, *Numerical methods for the solution of ill-posed problems*. Springer Science & Business Media, 2013, vol. 328.

[48] L. F. R. Lucas, N. M. M. Rodrigues, C. L. Pagliari, E. A. B. da Silva, and S. M. M. de Faria, "Sparse least-squares prediction for intra image coding," in *Image Processing (ICIP), 2015 IEEE International Conference on*, Sept 2015, pp. 1115–1119.

Appendix A

Contributions

A.1 Conferences

J. M. Santos, A. F. R. Guarda, N. M. M. Rodrigues and S. M. M. Faria, "Contributions to lossless coding of medical images using minimum rate predictors," *Image Processing (ICIP), 2015 IEEE International Conference on*, Quebec City, QC, 2015, pp. 2935-2939. [Online]. Available: `http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7351340`

A. F. R. Guarda, J. M. Santos, L. A. S. Cruz, P. A. A. Assunção, N. M. M. Rodrigues and S. M. M. Faria, "Orientation Dependent Inter-Slice Prediction For Volumetric Medical Images," *Conference on Telecommunications (Conftele) 2015*, Aveiro, Portugal, 2015.

A.2 Journals

A.2.1 Submitted

A. F. R. Guarda, J. M. Santos, L. A. S. Cruz, P. A. A. Assunção, N. M. M. Rodrigues and S. M. M. Faria, "A Method to Improve HEVC Lossless Coding of Volumetric Medical Images," *Signal Processing: Image Communication - Special Issue on Medical Image Communication, Computing and Security.*