長岡技術科学大学
Nagaoka University of Technology

Four-Dimensional Medical Image Compression by Using Non-Separable Integer Wavelet Transform

by

12316689

Fairoza Amira Binti Hamzah

Ph.D Dissertation

Submitted to the Graduate School of Engineering
Nagaoka University of Technology
in partial fulfilment of the requirements
for the Degree of

Ph.D in INFORMATION SCIENCE AND CONTROL ENGINEERING

Supervisor

Professor Masahiro Iwahashi

NAGAOKA UNIVERSITY OF TECHNOLOGY

Nagaoka, Niigata, Japan

Ph.D Dissertation

Four-Dimensional Medical Image Compression by Using Non-Separable Integer Wavelet Transform

Author: Fairoza Amira Binti Hamzah (Student ID: 12316689)

Supervisor: Professor Masahiro Iwahashi

Date: 28th January 2019

Abstract

The four-dimensional (4D) images such as 4D magnetic resonance image (MRI), 4D computed tomography (CT) and other types of 4D medical images have high slice resolution up to 1024×1024 or more at increasing pixel bit-depth. The limitations in storage space and transmission bandwidth has triggered a huge investigation on multidimensional and dynamic image compression to efficiently store and transmit the 4D images.

Since few decades ago, Discrete Cosine Transform (DCT) based digital image signal compression had been adopted as the JPEG international standard. Later, Discrete Wavelet Transform (DWT) has replaced the DCT and its being applied in medical image compression. JPEG 2000, the international standardization of DWT is using separable lifting structure where the multidimensional image signal is transformed separately in its horizontal and vertical direction. Besides that, each process is realized by cascading in lifting calculation. However, the necessity of waiting for previous step before calculating to the next step will make the overall delay time become longer. The delay time between input and output of WT is reduced as the proposed method reduces its lifting steps. Since the lifting step contains a rounding operation, variance of the rounding noise generated due to the rounding operation inside the transform is reduced.

Unlike the conventional separable structure, the proposed non-separable structure reduces the rounding noise inside the transform, which will lead to the increasing of coding performance. The proposed wavelet transform has a merit that its output signal, apart from the rounding noise, is the same as the conventional separable structure which is a cascade of 1D

structure. As a result of experiments in the first proposal, it was observed that the proposed method reduces the rounding noise as well as increases the performance of data compression of various 4D input signals.

The JPEG 2000 restricts the user's choice to two wavelet transforms—Daubechies 9/7 for lossy compression and the 5/3 LeGall wavelet, which has rational coefficients for reversible or lossless compression. The non-separable 3D and 4D structure in the 5/3-type transform for lossless coding reduces rounding noise, but it increases in the 9/7-type transform for lossy coding in the structure. A combination of 2D and 3D non-separable structures for 4D integer WT has been proposed to solve this problem, but it is found that the original filter arrangements need to be preserved to reduce rounding noise.

Therefore, in the second proposal, a non-separable 2D structure for the integer implementation of a 4D quadruple lifting WT with a 9/7 filter is proposed. The proposed wavelet transform has the same output signal as the conventional separable structure except for the rounding noise. As the order of the original lifting scheme is preserved, rounding noise in pixels of the decoded image can be significantly reduced, and the upper bounds of quality and lossy decoded 4D medical images can be improved.

Furthermore, the proposed methods can be extended in various ways to increase the coding performance for such multidimensional images, such as by implementing the region-of-interest (ROI) coding to utilize both lossless and lossy coding efficiently. Thus, an ROI coding is introduced in the third proposal for the non-separable lifting structure of 4D integer WT to compress the images by preserving the important part of the image only. It was observed that the proposed method increased the coding performance for 4D images.

Moreover, to further improve the performance of lossless coding, the adaptive directional lifting structure is introduced in the fourth proposal. This structure uses different lifting directions and optimally applies to the non-separable three-dimensional (3D) integer WT according to their own property. The adaptive structure has the advantage over the non-adaptive by further compacting the energy of low-frequency band signals and thus will increase its coding performance. It is proved that the proposed method in this proposal increase the lossless coding performance of various signals such as the video, medical image and light field image.

Acknowledgement

In the name of Allah, the most gracious, the most merciful, all the thanks go to the almighty Him for giving me strength and ability to finish this dissertation. It is all out of his willingness and endowments upon us. May Allah reward us for the time and effort spent in producing this work. I ask for forgiveness if there are any pitfalls found here and for prayer if it is beneficial.

The work in this dissertation would not have been possible without the support of many staffs, organizations and individuals. My words are not sufficiently eloquent to express my gratitude feelings towards them. First, I would like to express my deep gratitude to Malaysia government, specifically the MARA Education Foundation for providing me the scholarship to pursue Ph.D of Engineering study here.

The deepest gratitude to Professor Masahiro Iwahashi, my supervisor, for continuously supporting and supervising me throughout these whole years of study in Nagaoka University of Technology (NUT). Words cannot express my gratitude for his kindness, understanding, support and deep guidance. I would like to express my appreciation to Professor Kenji Nakagawa, Professor Masato Aketagawa, Professor Takashi Yukawa and Associate Professor Yasunori Sugita, panels of judging committee for my Ph.D degree, for their useful comments in helping me modifying this thesis.

I would also like to express my appreciation to my research internship supervisor, Associate Professor Victor Sanchez of University of Warwick, UK for his supportive guidance during the internship study in his research laboratory. Thanks to Assistant Professor Taichi Yoshida for his guidance as one of the co-authors of my research articles.

Special thanks to Nagaoka University of Technology, especially to International Student Affairs Division for taking care of me in here. My appreciation to my previous scholarship during undergraduate and postgraduate (Master of Engineering) study, MARA for providing me the opportunity to study abroad in Japan since my Bachelor's Degree. I am greatly thankful to all lecturers in Japanese Associate Degree (JAD) for advising me to study in NUT and for your helpful guidance in Malaysia.

My appreciation and thanks to my colleagues in Image Information System Laboratory, NUT, Dr. Teerapong Orachon, Dr. Suvit Poomrittigul, Dr. Khomdet Phapatanaburi, Mr. Zeyan Oo, Mr. Kentarou Yamaura and others for lots of discussions and kind support during my study here. I would also like to express my gratitude to my Malaysian, foreign and Japanese friends in Japan, Malaysia and other parts of the world, for your continuous supports and encouraging words to complete my Ph.D degree.

My mother and family supported me with every possible way. They gave me more than I deserve, all over my life. I respect them, love them, and owe them my success and all what I achieve in my life. Dear my lord, bestow on them your mercy as they did bring me up when I was young.

Last but not least, a big thanks to Malaysia and Japan for giving me a chance to study here. I do respect and love Japan as well as its people. I wish I can help my country to reach the same level of perfection, to be a well-developed country in the next few years, *Insha'Allah*.

Fairoza Amira Binti Hamzah

22nd January 2019

Table of Contents

Abstract I

Acknowledgement III

List of Figures X

List of Tables XIII

Abbreviations List XIV

Chapter 1 Introduction 1

1.1 Research Motivation 1

1.2 Research Objectives 3

1.3 Research Methodology 4

1.4 Dissertation Overview 5

Chapter 2 Research Background: Overview on Image Compression Standards 6

2.1 Introduction to Medical Images 6

2.2 The Importance of Medical Image Compression 7

2.3 Image Compression Scheme 8

2.3.1 Transform / Prediction 8

2.3.2 Quantization 11

2.3.3 Entropy Coding 12

2.4 Type of Image Compression 14

2.4.1 Lossless compression 14

2.4.2 Lossy compression 14

2.5 Image Compression Standards 16

2.5.1 JPEG 16

2.5.2 JPEG 2000 16

2.5.3 Others 17

Chapter 3 Research Background: Introduction to Wavelet Transform 18

3.1 Wavelet Transform 18

3.2 Integer Transform 19

3.2.1 Lifting Scheme 19

3.2.2 Factorization 20

3.2.3 Rounding Noise 20

3.2.3 Latency 22

3.3 Integer Wavelet Transform 23

3.3.1 One-Dimensional 23

3.3.2 Two-Dimensional (Non-Separable) 25

3.3.3 Three-Dimensional (Non-Separable) 26

Chapter 4 Four-Dimensional Non-Separable Integer Wavelet Transform with Double Lifting Structure for Lossless Compression 28

4.1 Proposal Motivation 28

4.2 Existing Method 31

4.2.1 Four-Dimensional Separable Integer Wavelet Transform for 5/3 filter 31

4.3 Proposed Method....35
4.3.1 Four-Dimensional Non-Separable Integer Wavelet Transform for 5/3 filter....35
4.3.2 Derivation of the Structure....38
4.3.2.1 Basic Properties for Modification....38
4.3.2.2 Derivation Process....39
4.3.3 Lifting Steps and Latency....41
4.4 Experimental Results....43
4.4.1 Evaluation on Rounding Noise....44
4.4.2 Evaluation on Coding Performance....47
4.5 Conclusion of this Chapter....51
Chapter 5 Four-Dimensional Non-Separable Integer Wavelet Transform with Quadruple Lifting Structure for Lossy Compression....52
5.1 Proposal Motivation....52
5.2 Existing Methods....54
5.2.1 Separable 4D Integer Wavelet Transform for 9/7 filter (Existing I)....54
5.2.2 Non-Separable 3D Integer Wavelet Transform for 4D 9/7 filter (Existing II)....56
5.2.3 Non-Separable 4D Integer Wavelet Transform for 9/7 filter (Existing III)....58
5.3 Proposed Methods....60
5.3.1 Designing a new lifting structure for 4D 9/7 filter....60
5.3.2 Non-separable 2D and 3D structure for 4D 9/7 filter (Proposed I)....62
5.3.3 Non-separable 2D structure for 4D 9/7 filter (Proposed II)....62
5.3.4 Comparison of the structures....63

5.4 Experimental Results....65
5.4.1 Evaluation on Rounding Noise....66
5.4.2 Evaluation on Coding Performance....70
5.5 Conclusion of this Chapter....73
Chapter 6 Further Proposals on the Non-Separable Integer Wavelet Transform....74
6.1 Region-of-Interest (ROI) Coding for 4D Image Compression....74
6.1.1 Proposal Motivation....74
6.1.2 Existing Methods....76
6.1.2.1 Separable 4D Integer Wavelet Transform without ROI....76
6.1.2.2 Non-Separable 4D Integer Wavelet Transform without ROI....76
6.1.2.3 Separable 4D Integer Wavelet Transform with ROI....76
6.1.3 Proposed Methods....77
6.1.3.1 Non-Separable 4D Integer Wavelet Transform with ROI....77
6.1.4 Experimental Results....78
6.1.4.1 Evaluation on Rounding Noise....78
6.1.4.2 Evaluation on Coding Performance....80
6.1.5 Conclusion of this Section....82
6.2 Adaptive Directional Lifting for 3D Image Compression....83
6.2.1 Proposal Motivation....83
6.2.2 Existing Method....85
6.2.2.1 Non-Adaptive Non-Separable 3D Integer Wavelet Transform....85

6.2.3 Proposed Method 87

6.2.3.1 Adaptive Non-Separable 3D Integer Wavelet Transform 87

6.2.4 Experimental Results 91

6.2.5 Conclusion of this Section 97

6.3 Conclusion of this Chapter 98

Chapter 7 Conclusion of the Dissertation 99

7.1 Conclusion and Discussions 99

7.2 Future Works 101

References 102

Work Achievements 112

List of Peer-Reviewed Research Articles 112

List of Peer-Reviewed Conference Proceedings 112

Others 112

List of Figures

Fig. 1 The evolution of medical image's dimension.....6

Fig. 2 Summary of the compression scheme 8

Fig. 3 The DCT 8 x 8 basis function, where the most upper left of the block is considered as the low frequency part and the lower right of the block as the high frequency part of the image [18].....10

Fig. 4 Example of a 4-4 connected combinatorial graph Laplacian 11

Fig. 5 Schematic diagram for lossless compression 14

Fig. 6 Schematic diagram for lossy compression 15

Fig. 7 Summary of current image compression standard 17

Fig. 8 Comparison between DCT and DWT basis 18

Fig. 9 Lifting structure of 1D DWT.....20

Fig. 10 Factorization of a rotation matrix to lifting structure 20

Fig. 11 Rounding operation and its equivalent expression.....21

Fig. 12 Probability Density Function (PDF) for an amplitude of the additive noise.....21

Fig. 13 Rounding noise between integer and real number.....22

Fig. 14 Example of latency in DWT 22

Fig. 15 The 1D integer WT for 5/3-type transform (double lifting structure).....23

Fig. 16 The integer WT of the 9/7-type transform for lossy coding (quadruple lifting structure) 23

Fig. 17 Two-dimensional integer wavelet transform 26

Fig. 18 Three-dimensional non-separable integer wavelet transform 27

Fig. 19 Decomposition of a 4D signal 33

Fig. 20 Separable structure for 4D integer WT 34

Fig. 21 Non-separable structure for 4D integer WT 37

Fig. 22 Basic properties for modification 38

Fig. 23 Derivation process for non-separable 2D 39

Fig. 24 Example derivation of 4D Non-Separable for the 1st lifting step 40

Fig. 25 Separable and non-separable 2D structures 41

Fig. 26 Implementation examples in parallel processing platform 42

Fig. 27 Tested data 44

Fig. 28 Rounding noise in each frequency band 46

Fig. 29 Coding performance in lossless mode 47

Fig. 30 Coding performance in lossy mode. 49

Fig. 31 Decomposition of a 4D signal 56

Fig. 32 Separable 4D structure for 9/7-type of transform (Existing I) 56

Fig. 33 Non-separable 3D structure for 9/7-type of transform (Existing II) 58

Fig. 34 Non-separable 4D structure for quadruple lifting (Existing III) 59

Fig. 35 Non-separable 2D combined with 3D structure for 9/7-type of transform (Proposed I) 62

Fig. 36 Non-separable 2D structure for 9/7-type of transform (Proposed II) 63

Fig. 37 Tested data 66

Fig. 38 Variance of rounding noise for all data in each frequency band 68

Fig. 39 Average variance of rounding noise in each frequency band 68

Fig. 40 Effect of one rounding operator for fMRI(II) data 69

Fig. 41 Coding performance in lossy mode for CT image 70

Fig. 42 Coding performance in lossy mode for fMRI(I), fMRI(II), MRI, AR and US data.... 72

Fig. 43 ROI coding for non-separable double lifting structure of integer WT (Proposed) 77

Fig. 44 Tested data 78

Fig. 45 Variance of rounding noise in each frequency band 79

Fig. 46 Coding performance for the 4D MRI Image 80

Fig. 47 Transform classification according to M_1 87

Fig. 48 Directions of $M_{i,j}$ (i = 0,1,2 and j = 0,1,2,3,4). 89

Fig. 49 Prediction of depth axis between two planes based on depth direction in the matrices. 89

Fig. 50 A part of test videos, medical slice data and light field data. 92

Fig. 51 Test data (Light field image=Tarot card) 95

Fig. 52 Graph of the relationship between the bit per pixel of compressed data and its side information of a medical data. 96

List of Tables

Table 1 Size of uncompressed medical image 7

Table 2 Comparison of the structures 41

Table 3 Type of data used in the experiments 43

Table 4 Comparison of the methods 64

Table 5 Type of data used in the experiments 65

Table 6 Lossless coding resuts in bpp 92

Table 7 Processing time for each method 94

Table 8 Bpp of compressed data and their overhead 96

Abbreviations List

2D	Two-Dimensional
3D	Three-Dimensional
4D	Four-Dimensional
ADL	Adaptive Directional Lifting
APSIPA	Asia-Pacific Signal Processing Association Annual Summit and Conference
AR	Auto-regressive
BPG	Better Portable Graphics
Bpp	Bits per pixel
CDF	Cohen-Daubechies-Feauveau
CT	Computed Tomography
DCT	Discrete Cosine Transform
DICOM	Digital Imaging and Communications in Medicine
DWT	Discrete Wavelet Transform
EBCOT	Embedded Block Coding with Optimized Truncation
EURASIP	European Association for Signal Processing
EZWIP	Embedded Zerotree Wavelet Intra-band Partitioning
fMRI	Functional Magnetic Resonance Image
GBT	Graph-based Transform
HEVC	High Efficiency Video Coding
HHHH	High-frequency band signal
IEEE	Institute of Electrical and Electronics Engineers
IEICE	Institute of Electronics, Information and Communication Engineers
JPEG	Joint Photographic Experts Group
KLT	Karhunen-Loève Transform
LLLL	Low-frequency band signal
MRI	Magnetic Resonance Image
NSP	Non-Separable
PACS	Picture Archiving Communication System

PCA	Principle Component Analysis
PDF	Probability Density Function
PET	Positron Emission Tomography
PROP	Proposed method
PSNR	Peak signal-to-noise ratio
RGB	Red-Green-Blue
ROI	Region-of-Interest
SEP	Separable
SPECT	Single-Photon Emission Computed Tomography
SPIHT	Set Partitioning in Hierarchical Trees
US	Ultrasound
WT	Wavelet Transform
YCbCr	Luma, Blue Chroma, Red Chroma

Chapter 1 Introduction

1.1 Research Motivation

Prompt advances in multidimensional image data have caused the development study of the most suitable compression method for them to be very crucial. As the medical imaging industry progresses toward a filmless environment, the amount of digital data that need to be managed presents a significant challenge. 4D images are gradually being collected and used in many clinical and research applications, such as 4D Magnetic Resonance Image (MRI), 4D computed tomography (CT), 4D ultrasound, and 4D functional MRI (fMRI). 4D medical images have had an ample influence on the diagnosis of diseases and surgical planning [1]. Four-dimensional medical images typically represent sequences of 3D image. The image slice resolution of 512x512 are considered to be minimum standard, but, nowadays, more recent scanning systems are able to output image slices with spatial resolutions of 1024x1024 or more at increasing pixel bit-depths [2]. The limitations in storage space and transmission bandwidth, and the growing size of medical image data sets on the other, have pushed the design of ad-hoc tools for their handling. The increasing demand for efficiently storing and transmitting digital medical data set has triggered a huge investigation of multidimensional and dynamic image compression. Thus, several studies have examined the compression of 4D medical images, such as [3, 4, 5].

Since every single bit of information in medical image is particularly important, therefore, a scalable lossless coding is necessary. The increasing of internet speed from 3G to 4G, then in the near future, 5G network, from 10MBps to 1GBps in fibre internet does not limit the necessity of study regarding data compression. This is due to the limited bandwidth of each network coverage and its quota per person. For that reason, data compression will be the biggest help to provide the highest internet speed to their users, especially to countries with slow internet speed.

The main goal of designing any system for communication or storage is first, the performance should be as good as possible, and second, the system should be as simple as possible. The long history of image compression begins with the well-known JPEG (Joint Photographic Expert Groups) compression scheme which use the ‘.jpg’ and ‘.jpeg’ image

format. The JPEG has released various version of image compression algorithm such as the JPEG-LS and JPEG 2000 for lossless image compression [6]. There are several important steps in image compression scheme which includes the transform or prediction part, quantization part and entropy coding part. Transform is the most important part as it is needed to classify the important and unimportant details of the image thus reducing the redundancy of the image.

The JPEG is adopting Discrete Cosine Transform (DCT) as its transform coding, while JPEG 2000 use Discrete Wavelet Transform (DWT) to transform the image. Since DWT has several better properties than DCT to transform the image efficiently, this dissertation will design the DWT based integer transform for 4D image. Further explanation regarding this method will be thoroughly describe in this dissertation.

1.2 Research Objectives

In this dissertation, we address different aspects of 4D medical image compression for efficient storage and transmission of the 4D data. The main goal of image compression is to reduce the energy in the transform and thus to attain a higher compression performance. This research has shown that the wavelet transform is an efficient way to pack the energy of 4D medical images in a few coefficients. The objectives of this research include the following:

- Design a 4D integer-based wavelet transform to accommodate the 4D medical image compression in both lossless and lossy coding mode.

- Reduce the rounding noise inside the non-separable integer wavelet transform to compact more energy in low frequency part, thus will increase the compression ratio of the 4D medical image.

- Reduce the latency in 4D image compression system by introducing the non-separable structure, which will reduce the lifting steps inside the transform.

- Exploiting the energy compaction property of discrete wavelet transform to compress the 4D medical images by further compacting the energy in the detailed component of the transformed image, which is the low frequency part.

- Efficiently utilize integer transform by using region-of-interest (ROI) coding to compress the 4D medical images efficiently.

- Adaptively transform the 3D images to compact more energy in the low frequency band signal.

1.3 Research Methodology

The applied methodology started by thoroughly investigating and studying the theoretical concepts of data compression and its digital design approaches. The theoretical study was followed by real development through performing the transform method on the system. The results were used to practically demonstrate the design applicability to the design concepts.

In order to fulfil the main goal of this study as stated in Section 1.2, the most scalable and appropriate method is necessary to be designed. This study is focusing on designing a suitable transform or predictive method for multidimensional data coding. This is because the transformed coefficients are important to determine its coding efficiency. The non-separable lifting structure for integer wavelet transform yield better coding efficiency as it can reduce the rounding noise and compact the energy in low-frequency band signals. Lifting is like a ladder structure to predict neighbouring pixels. It first proposed in [7] to be applied in DWT.

Latest development of image coding method, which is the Better Portable Graphics (BPG) [8] applying the DCT as it's based. However, the BPG has its limitation, which it can efficiently be applied in lossy image coding only. Joints Photographic Experts Group (JPEG) 2000 which used DWT as it's based can be applied in lossless video compression, but in separable condition, which will be increasing its computing times, thus the data cannot be compress in the best manner. Moreover, the DWT has an advantage over DCT, which it will not produce any blocking artefact when the data is decompressed back again. By studying on implementing integer WT for 4D image compression, the most scalable image compression scheme for these images is to be proposed.

1.4 Dissertation Overview

This dissertation consists of seven main chapters, which are:

- Chapter 1 is the introduction of this dissertation which clarify our research motivation that includes the significance of 4D medical image compression. The objectives and detailed problems definition are investigated here.
- Chapter 2 presents research background which is the overview on image compression standards. The image compression scheme and standards are explained in detail here.
- Chapter 3 explains the second part of this research background which the introduction of wavelet transform. The design of integer-based wavelet transform and its necessity are presented here.
- Chapter 4 highlights the first proposal of this dissertation which is to construct the 4D non-separable integer wavelet transform for lossless coding. The derivation of this transform and details experiment to proof the advantage of the first proposal is clarified in this chapter.
- Chapter 5 introduces the second proposal, which is to design the 4D non-separable integer wavelet transform for lossy coding. Various lifting structure for lossy coding are designed to find the best lifting structure with the best coding performance.
- Chapter 6 further introduces third and fourth proposal which will utilize the non-separable integer wavelet transform for 3D and 4D image compression. This chapter consists of two sections, which are the ROI coding for 4D image and adaptive lifting structure for 3D image.
- Chapter 7 will conclude the thesis and presents its future work.

Chapter 2 Research Background: Overview on Image Compression Standards

2.1 Introduction to Medical Images

The origin of medical images began around early 20th century, after the discovery of the x-ray. The 2D digital medical images comprises of two spatial dimensions which is the *x*-dimension and *y*-dimension. The 3D medical images are the stack of the 2D slices which will form the *z*-dimension. Meanwhile the 4D medical images are the sequence of the 3D images with another temporal, *t*-dimension. The summary of the evolution of medical image's dimension is summarized as in Fig. 1.

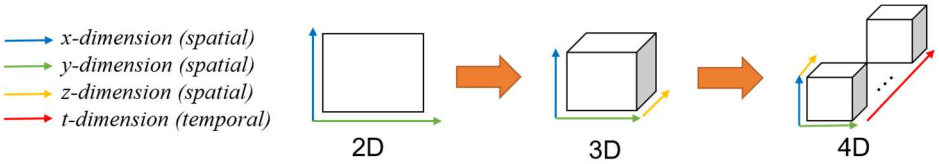


Fig. 1 The evolution of medical image's dimension.

The imaging modalities nowadays are totally dependent on computing capabilities such as the MRI (Magnetic Resonance Imaging) and CT (Computed Tomography). The 4D medical images includes volumetric CT, MRI, positron emission tomography (PET), single-photon emission computed tomography (SPECT) and ultrasound (US) imaging [9]. Recent study on reconstruction of these medical images using 4D structured mesh models has been done in [10].

These medical images are managed digitally by radiology Picture Archiving and Communications System (PACS), which is invented to make it easier to store and access the vast volume of medical images [11]. Image compression used in PACS is Digital Imaging and Communications in Medicine (DICOM)-compatible JPEG format which use the DCT as its transform. The DICOM standard is globally used to store, exchange and transmit medical images.

2.2 The Importance of Medical Image Compression

Nowadays, the medical images, besides being stored in PACS, researchers have uploaded those images in their website for online sharing among medical experts. However, these 4D medical images have huge number of pixels which led to large uncompressed file size as shown in Table 1.

Table 1 Size of uncompressed medical image

Type of Image	Image Dimension (x,y,z,t)	Uncompressed File Size
CT	512,512,100,20	1 GB
MRI	512,512,100	50 MB
fMRI	256,256,20,25	250 MB

The major advances of communication system such as the internet speed enhancement by optical fibre communication networks and 5G wireless network will not limit the necessity of medical image compression. This is because such high volumes medical data can make any transmission media, including optical fibre to be quickly overwhelmed. Furthermore, research on building better medical imaging modalities which will increase the data dimensionality and uncompressed file size is still an ongoing research, perhaps, in the near future, medical experts will see these images in a virtual or mixed reality environment, rather than in a computer software. Thus, image compression research is always needed to store and transmit these images in the best manner.

Image compression is done to reduce irrelevance and redundancy of the image data in order to be able to store or transmit them in the most efficient form. The large uncompressed file size of the data above can be reduced to various compression ratio if it is compressed in lossy mode. But as the medical images contain important information, most of the medical experts prefer the lossless compression more than the lossy compression. Lossless compression has the disadvantage that the compression ratio will be lower than the lossy compression, but the decompressed image will be exactly the same as the original image.

2.3 Image Compression Scheme

Most image compression scheme comprised of the three main components, which are the decorrelation algorithm which is the transform or prediction part, the quantization part and the entropy encoding part. The image compression scheme is summarized as in Fig. 2.

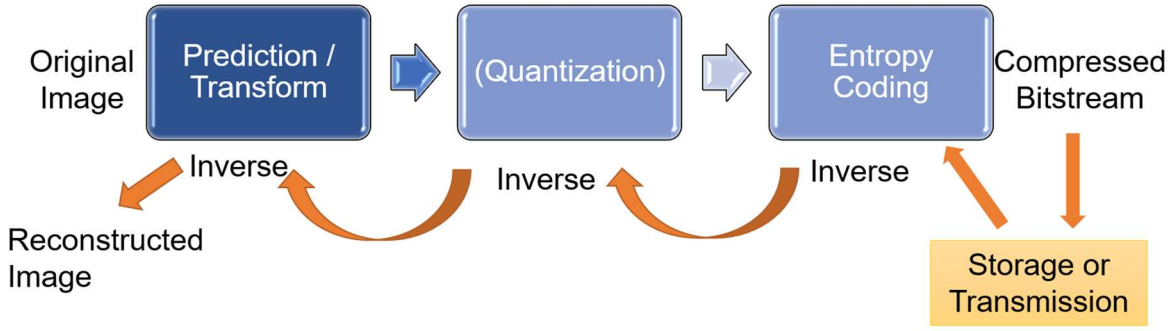


Fig. 2 Summary of the compression scheme

The quantization is done only for lossy compression scheme.

2.3.1 Transform / Prediction

The focus of this dissertation is the transform part which is the crucial part in image compression scheme. Transform coding is needed to decorrelate the data and discarded the unnecessary information. Data is compressible because it is correlated. The transform coding will produce transform coefficients which are statistically independent. It also needs to perform energy compaction, where the small number of coefficients should represent most of the energy in an image. The basis functions should be independent otherwise it is necessary to transmit the basis to the decoder. The transform is also preferred to be fast implemented.

For instance, there is a block of consecutive samples of a stationary random process and this block is needed to be coded efficiently with a specified number of bits. Let **X** denote the sample vector

$$\mathbf{X} = (X_1, X_2, \cdots, X_k)^t, \qquad (2.1)$$

These samples will have significant correlation and separate quantization and it is inefficient to encode them when the quota of bits is small. The input vector, **X** can be linearly transformed

to obtain a new vector, **Y**, also with k components, which are known as the transform coefficients where these coefficients are less correlated than the original samples. The information is more compact as it is being concentrated in only a few of the transform coefficients [12].

There are various transforms coding since the beginning of the image compression history, such as the Karhunen-Loève Transform (KLT) which is the optimal transform among the others. The KLT can be used within a block-based coding transform context depends on the covariance of the source [13]. It consists of the eigenvectors of the autocorrelation matrix. Even though it can completely decorrelates the signal in the transform domain, its signal dependency and lack of fast transform algorithm usually makes it an impractical tool in signal processing [14]. In addition, it is also necessary to transmit the KLT basis to the decoder to reconstruct the image back. KLT is also known as the Principle Component Analysis (PCA) transform.

The JPEG and BPG image compression standards use the DCT to decorrelate the image samples and generate a set of transform coefficients [15]. The two-dimensional transforms are computed separately by applying them in the horizontal and vertical directions. The elements of the core transform matrices were derived by approximating scaled DCT basis functions. The N transform coefficients v_i of an N-point 1D DCT applied to the input samples u_i can be expressed as

$$v_i = \sum_{j=0}^{n-1} u_j c_{ij}, \qquad (2.2)$$

where $i = 0, \dots, N-1$. Elements c_{ij} of the DCT matrix C are defined as

$$c_{ij} = \frac{P}{\sqrt{N}} cos\left[\frac{\pi}{N}\left(j + \frac{1}{2}\right) i\right], \quad (2.3)$$

where $i, j = 0, \dots, N-1$ and where P is equal to 1 and $\sqrt{2}$ for $i = 0$ and $i > 0$, respectively. The basis vectors $\mathbf{c}_i$ of the DCT are defined as $\mathbf{c}_i = [c_{i0}, \dots, c_{i(N-1)}]^T$ where $i = 0, \dots, N-1$ [16]. The DCT has several properties that are considered useful for both compression efficiency and for efficient implementation such as the basis vectors are orthoginal and have equal norm. They have been shown to provide good energy compaction which is also necessary for compression efficiency [17]. However, the DCT has disadvantages that it has blocking artifact when decompressing the image and it is a fixed block linear transform which is not much applicable with different statistical properties of blocks.

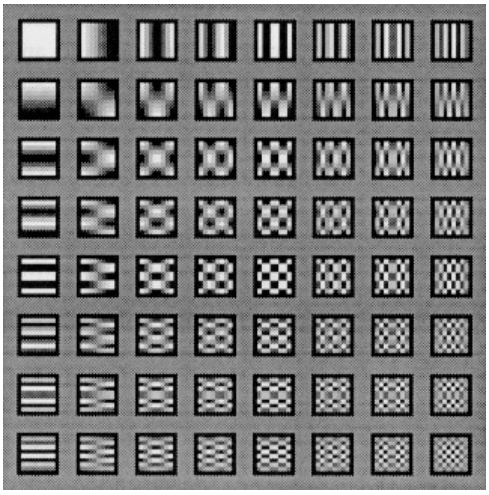


Fig. 3 The DCT 8 x 8 basis function, where the most upper left of the block is considered as the low frequency part and the lower right of the block as the high frequency part of the image *[18]*.

Graph-based signal processing has recently gained interest for compression by employing Graph-based Transforms (GBTs). Important works may be found for depth-map coding [19] , [20] and [21]; video coding [22] [23]; image coding [24] [25]; and for two-dimensional (2D) piecewise smooth signals [26]. GBTs have also been applied to decoding of JPEG images such as in [27] [28].

In graph signal processing [29], signals are supported on an undirected, weighted and connected graph, $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{W}\}$, where signal values are attached to nodes of the graph ($\mathcal{V}$) , and its links ($\mathcal{E}$) capture inter-sample relations among signal's samples. The adjacency matrix, **W**, represents the graph's links weights. The GBT is defined by using its combinatorial Laplacian,

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \quad (2.4)$$

where **D** is the diagonal degree matrix. The other type of graph which is the generalized graph Laplacian of a weighted graph $\mathcal{G}$ is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{W} + \mathbf{V}, \quad (2.5)$$

where **W**, **D**, and **V** are the adjacency matrix, the degree matrix and the self-loop matrix of $\mathcal{G}$, respectively. To find the GBT associated with graph $\mathcal{G}$, the eigen-decomposition of graph Laplacian is done, that is

$$\mathbf{L} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{t}, \quad (2.6)$$

where the columns of **T** are the basis vectors of the corresponding GBT. An example of the Laplacian graph is shown in Fig. 4. The $v_1, \cdots, v_4$ variables are the vertices of the graph which represent the pixel values of the image and $e_1, \cdots, e_4$ are the edges of the graph which denotes the pixel similarity in the image. Then, by decomposing the image with the eigenvalues from the graph created, the image blocks can be transformed according to their statistical properties.

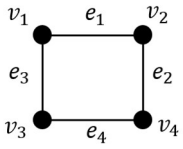


Fig. 4 Example of a 4-4 connected combinatorial graph Laplacian

GBTs have the advantage that they can adapt to the different statistical properties of the blocks to be transformed. Unfortunately, GBT is the same with KLT, where the basis functions are eigenvectors of the covariance matrix of the input signal, thus, it is necessary to transfer the transform basis as side information to the decoder which will increase the computational complexity of the compression scheme.

The DWT, on the other hand, can represent a signal as a superposition of a wavelet basis that is discretely sampled [30], and the coefficients of the basis can then be used to reconstruct the original signal. The DWT has some features that make it suitable for scalable compression, for examples, representation of an image at a different resolution, compacting most of the energy in a few wavelet coefficients and no blocking artifact produces as the image is wholly analyzed and reconstructed. The DWT is localized in both time and frequency where it is obtained by using scaling and translation of a scaling and wavelet function, as shown in equation (2.7).

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}), \quad (2.7),$$

where *a* is the scaling parameter and *b* is the shifting parameter. DWT is realized by using subband coding in JPEG 2000.

2.3.2 Quantization

Quantization is the process of mapping input values from a large set (often a continuous set) to output values in a (countable) smaller set, often with a finite number of elements. The quantization process is only done in lossy compression technique to reduce the number of discrete symbols to make it more compressible. Scalar quantization is an irreversible step and

cannot be used in lossless image compression. In order to proceed with quantization, a transform is necessary as they provide coefficients for the scalar quantization. JPEG 2000 use a dead-zone uniform scalar quantizer to coefficients from the wavelet transform of the image [31].

For a given wavelet coefficient *y*, the quantizer produces a signed integer q given by

$$q = Q(y), \quad (2.8)$$

The quantization index q indicates the interval which *y* lies.

2.3.3 Entropy Coding

Entropy coding removes redundancy from the representation of data. The entropy rate [32] [33] indicates the approximate amount of information per source symbols in average. And is defined as

$$Entropy\ (H) = -\sum_{i=1}^{n} p(s_i) \log_2 p(s_i), \quad (2.9)$$

where $p(s)$ is a probability of symbol (element of data set) of data and $p(s_i)$ is a probability of each symbol. Entropy rate interpreted as a minimum bit rate required in representing the equation. Entropy rate in image processing used bit-per-pixel (bpp) as its unit.

The most popular variable-rate entropy coding techniques are Huffman coding and arithmetic coding. The JPEG 2000 employs Embedded Block Coding with Optimized Truncation (EBCOT) for its entropy coding method [34]. For each code-block EBCOT generates a separate embedded bitstream where a rate distortion optimization is done to find how many bits in each code-block's embedded bit stream should be kept resulting in the highest quality image at any given bit rate. However, this dissertation uses the Embedded Zerotree Wavelet Intra-band Partitioning (EZWIP) [35] which combines the EZW coding and Set Partitioning in Hierarchical Trees (SPIHT) codec. The EZWIP provides simpler implementation compared to SPIHT and EBCOT.

The EZW allows complete embedded bit representation. SPIHT is the improvement of EZW which uses spatial orientation tree structure and bit-plane coding method which can achieve embedded code stream of high compression efficiency.

2.4 Type of Image Compression

2.4.1 Lossless compression

Image compression process starts with predictive or transform coding as shown in Fig. 5, in this stage, the data is transformed into its low pass and high pass coefficients. It is a first step to transform the original image signal into a new data set of signals where it is done to remove the redundant information in image signal without any distortion.

Then, the transformed data will undergo the entropy coding, where the transformed data will be encoded, to form the final bit stream or symbol data file for transmission or storage. This step will affect the reduction of total bit rate to represent the transform samples of image data. Later, the compressed data will be decoded, followed by final steps of data compression, which is the inverse transform of predictive coding or transform method. The above process is applied in lossless image coding, the reversible coding method to ensure there will be no information loss during the process. This method is usually being applied in medical image compression and hyperspectral image, where every single bit information is mostly essential.

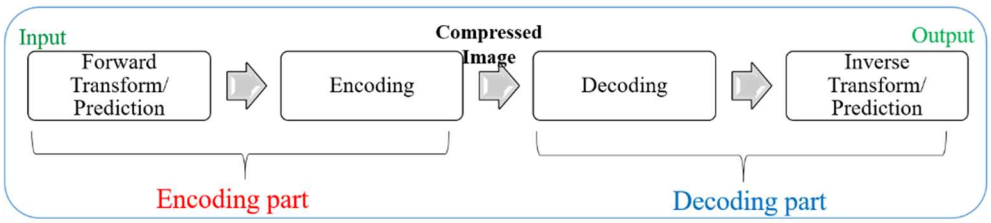


Fig. 5 Schematic diagram for lossless compression

2.4.2 Lossy compression

Lossy image coding happened by quantizing the transformed data before it is being encoded as shown in Fig. 6. This lossy compression method is usually applied in television broadcasting, web data uploading etc., because they do not require the exact same reconstructed output data.

Most of the applications in the image processing field requires a good compression ratio of a compressed image [36]. Compression ratio is a ratio between compressed data with input data. The higher the compression ratio, the lower the file size and its storage and transmission capacity. Besides compression ratio, computational cost is also an important factor in this field. The lower the computational cost, the lower the transmission time, thus resulting to a better transmission system. Not only that, the data quality after it is decompressed is also important

for certain fields such as the medical data, satellite data and others. The higher the quality of data, the better the compression system. In conclusion, there are three main aspects to be considered in order to build the most efficient compression system, which are; the compression ratio, the computational cost and the quality of decompressed data.

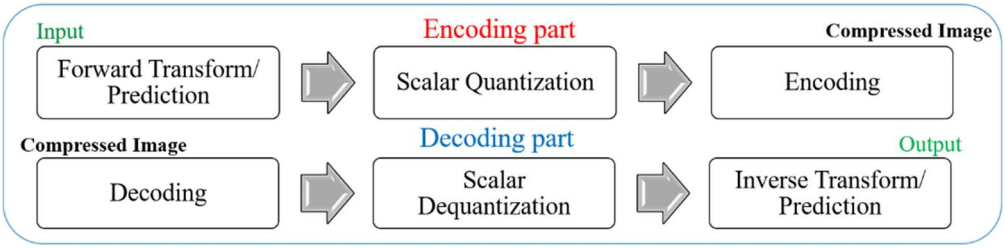


Fig. 6 Schematic diagram for lossy compression

The quality of the decompressed image is often measured by using the peak signal-to-noise ratio (PSNR), where it is measured by using decibel value [dB]. The higher the dB value, the higher the quality of the decompressed image. Meanwhile for image compressed by using lossless method, the PSNR value will be infinity as the decompressed image will be reconstructed exactly the same as the original image. The PSNR value is measured as in equation (2.10).

$$PSNR = 10\,log_{10}\left(\frac{S_p^2}{MSE}\right), \qquad (2.10)$$

where S refers to the bit-depth of the image, and MSE is the mean-squared-error, which is the different between the decompressed and original image.

2.5 Image Compression Standards

2.5.1 JPEG

JPEG is usually used for lossy compression of digital images. The degree of compression is adjustable which allows a trade-off choice between storage size and image quality. The colour red-green-blue (RGB) image is first transformed into YCbCr colour space which yields three monochromatic images that are processed separately. They are divided into small blocks such as the 8×8 pixels and each block are transformed into spatial frequency domain by means of a two-dimensional DCT. The transformed blocks contain low spatial frequency components in the upper left corner while higher frequency components in the lower right corner. The DCT concentrates most signal energy in the low spatial frequency analysis [6]. After transforming the input data, the transformed coefficients are quantized to reduce more information. Later, they are sequenced and packed into the output bitstream.

2.5.2 JPEG 2000

JPEG 2000 is an image compression standard and coding system, which was created by the JPEG committee in 2000 to supersede their original DCT-based JPEG standard with a newly designed, wavelet-based method. The JPEG 2000 provides both lossless and lossy compression in a single compression architecture. Lossless compression is provided using reversible integer wavelet transform in JPEG 2000, which is a rounded version of the biorthogonal Cohen-Daubechies-Feauveau (CDF) 5/3 wavelet transform [37]. The CDF 5/3 wavelet uses only integer coefficients which the output noise does not require rounding (quantization) and therefore will not introduce any quantization noise. Lossy compression can be done by using irreversible biorthogonal CDF 9/7 wavelet [38] because it introduces quantization noise that depends on the precision of the decoder.

Compared to the previous JPEG standard, JPEG 2000 delivers a typical compression gain in the range of 20%, depending on the image characteristics. Higher-resolution images tend to benefit more, where JPEG-2000's spatial-redundancy prediction can contribute more to the compression process. In very low-bitrate applications, studies have shown JPEG 2000 to be outperformed by the intra-frame coding mode of H.264 [39]. Good applications for JPEG 2000 are large images, images with low-contrast edges — e.g., medical images. The JPEG 2000 also does not produce blocking artifacts for the reconstructed image compared to the JPEG.

2.5.3 Others

Besides JPEG and JPEG 2000, there are various image compression standard such as the latest BPG, JPEG-LS, JPEG-XR etc. The JPEG-LS was designed for effective lossless and near-lossless compression of continuous-tone still images. The JPEG XR supports a wide range of colour encoding formats which combines the optimized image quality and compression efficiency together with low-complexity encoding and decoding. The latest JPEG PLENO [40] is designed for light field image coding and virtual reality images. Meanwhile the new BPG image compression standard is designed based on the intra-prediction of High Efficiency Video Coding (HEVC) applies to still image. The BPG has the same disadvantage with JPEG, as the reconstruction image may produce blocking artefact due to the block splitting at the first stage of the compression. The clear comparison between all the compression standards are shown in Fig. 7.

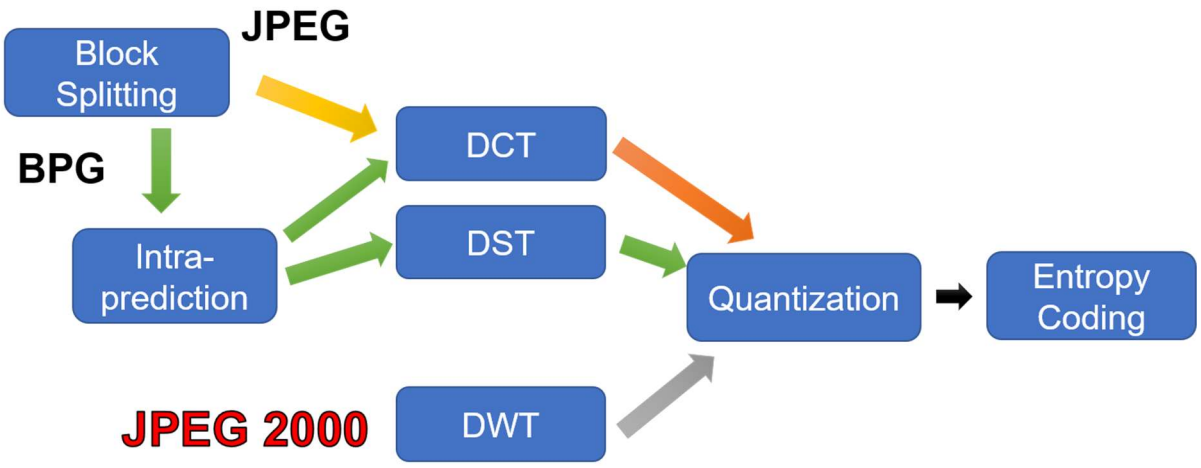


Fig. 7 Summary of current image compression standard

Chapter 3 Research Background: Introduction to Wavelet Transform

3.1 Wavelet Transform

Wavelets are mathematical functions that cut up data into different frequency components and study each component with a resolution matched to its scale. It will decompose a signal into a set of basis functions which are then called as wavelets. DWT transforms a discrete time signal to a discrete wavelet representation.

The advantage of DWT over DCT is that the DWT are well localized in both time and frequency domain and have different length of basis, whereas the DCT is only localized in frequency domain and have the same length of basis [41]. Therefore, the DWT's basis can easily be classified to low frequency and high frequency components and making it easier to compress the images compared to the DCT, as shown in Fig. 8.

Furthermore, the DWT does not need to divide image into blocks like the DCT, as it can analyse the whole image. The DWT is then applied to the sub images, and thus will not produce any blocking artifacts like DCT, KLT etc.

(a) DCT basis (b) DWT basis

Fig. 8 Comparison between DCT and DWT basis

3.2 Integer Transform

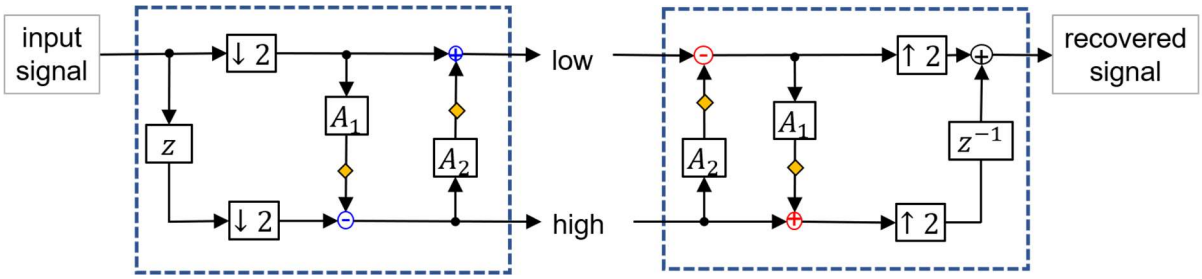
Integer transform [42] is vital for lossless image compression. The general input image values are all in integer numbers and may change into non-integer values during the compression process. This non-integer values can cause some loss of image information when reconstructing the image. Due to the importance of integer-to-integers transform in lossless image compression, this dissertation presents an approach to build integer transform based on the idea of factorizing into lifting steps by rounding the filter coefficients. Note that the factoring parameters are almost real number in factorization process. Thus, the integer transform is a necessary method to make a reversible transform in lossless compression system.

3.2.1 Lifting Scheme

The basic idea behind lifting is that it provides a simple relationship between all multiresolution analyses that share the same low-pass filter or high pass filter. Lifting allows for an in-place implementation for the fast wavelet transform, a feature similar to the fast Fourier transform. The signals can be transformed from integer to integer with rounding operators in each lifting step.

Lifting also allows the construction of wavelets without making the use of the Fourier transform and every transform built is immediately invertible, where the inverse transform has exactly the same computational complexity as the forward transform. Furthermore, lifting allows for adaptive wavelet transforms, where one can start the analysis of a function from the coarsest levels and then build the finer levels by refining only in the area of interest [43].

Fig. 9 shows basic lifting structure of DWT. Fig. 9 (a) shows the forward transform of DWT, and Fig. 9 (b) shows the inverse transform of DWT. x_0 and x_1 is the input data, y_0 and y_1 are transformed data, y_0 and y_1 are the low pass data and high pass data, respectively. P and U are coefficients of filter bank, ○ and ◇ represents the addition or subtraction operation, and rounding operator, respectively.



(a) Forward transform (b) Inverse transform

Fig. 9 Lifting structure of 1D DWT

It is clearly shown in Fig. 9 that the above lifting structure will produce the same recovered signal comparing with the input signal. The inverse transform is done by inversing the arithmetic operation from the forward transform to avoid any reconstruction loss.

3.2.2 Factorization

Polyphase matrix of wavelet or subband filters are factorized into basic matrices as in [7]. The factorization simplifies matrices to the lifting structure in the biorthogonal case. In a given rotation matrix is factorized as shown in Fig. 10 by

$$\begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix} = \begin{bmatrix}1 & 0 \\ f_3 & 1\end{bmatrix}\begin{bmatrix}1 & f_2 \\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 \\ f_1 & 1\end{bmatrix}, \qquad (3.1)$$

where

$$[f_1 \quad f_2 \quad f_3] = \left[\tan\frac{\theta}{2} \quad -\sin\theta \quad \tan\frac{\theta}{2}\right], \qquad (3.2)$$

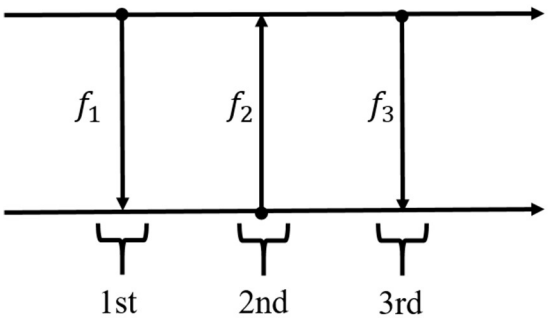


Fig. 10 Factorization of a rotation matrix to lifting structure

3.2.3 Rounding Noise

As fewer rounding operators do not imply fewer rounding errors in total [44] [45], the

total rounding noise in the output frequency band signal was investigated. The rounding operation is a vital part of lossless compression.

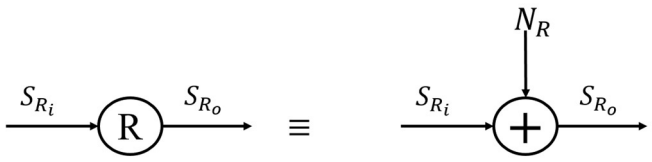


Fig. 11 Rounding operation and its equivalent expression

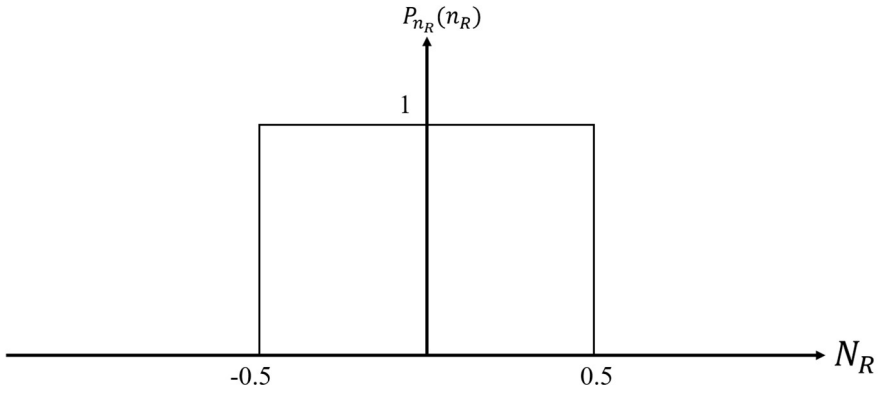


Fig. 12 Probability Density Function (PDF) for an amplitude of the additive noise

A non-linear operation transforms a floating-point signal into an integer signal. An equivalent expression to the rounding operation is shown in Fig. 11. A non-linear equation with additive noise is defined as

$$S_{R_o} = S_{R_i}(z) + N_R(z), \qquad (3.3)$$

where S_{R_i}, S_{R_o}, and N_R denote the input signal, the output signal, and the additive noise of the rounding operation, respectively. As the correlation between each of the of the errors and the signals was zero (based on statistical independence), the variance of output signal was calculated from

$$\sigma^2_{S_{R_o}} = \sigma^2_{S_{R_i}} + \sigma^2_{N_R}, \quad (3.4)$$

where $\sigma^2_{S_{R_i}}$, $\sigma^2_{S_{R_o}}$, and $\sigma^2_{N_R}$ refer to the variance of the input signal, the output signal, and the additive noise of the rounding operators, respectively. As the power of the probability density function (PDF) of additive noise is approximately flat, as shown in Fig. 12, the variance of the additive noise for the rounding operations was calculated as

$$\sigma_{N_R}^2 = \int_{-0.5}^{0.5} x^2 dx = \frac{1}{12}, \qquad (3.5)$$

In this dissertation, rounding noise inside the circuit was measured to observe the accumulative error in it due to the rounding operators. Fig. 13 illustrates that rounding noise was examined by using the error between the integer signal and the real number signal. Rounding noise is defined by Equation (3.6).

$$error = y - \hat{y}, \qquad (3.6)$$

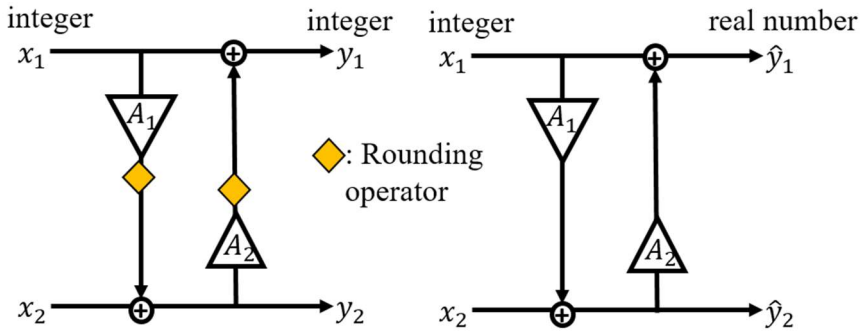


Fig. 13 Rounding noise between integer and real number

3.2.3 Latency

Latency in DWT is the delay of calculation results of previous lifting step for the input of the next step [46]. Fig. 14 shows the example of latency of 2D DWT with 4 lifting steps. There is a delay occurred in every lifting step. Further explanation regarding lifting steps and latency is in Section 4.3.3.

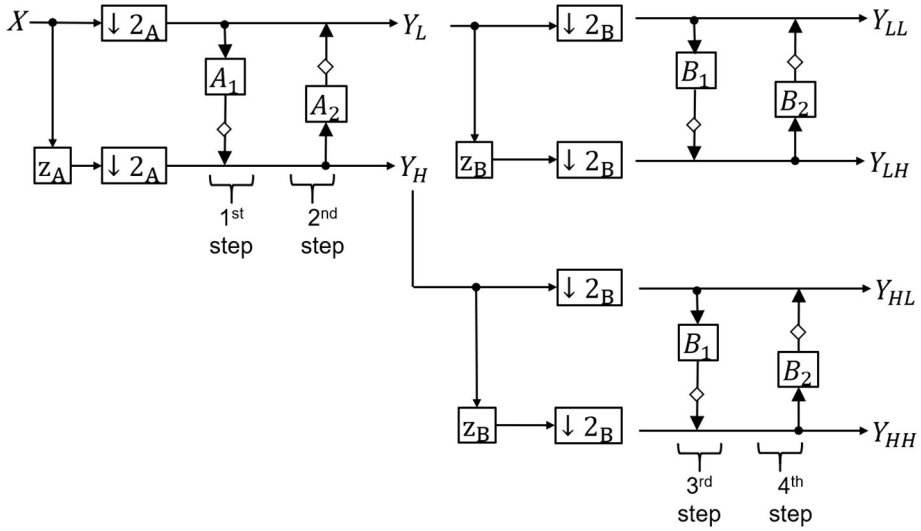
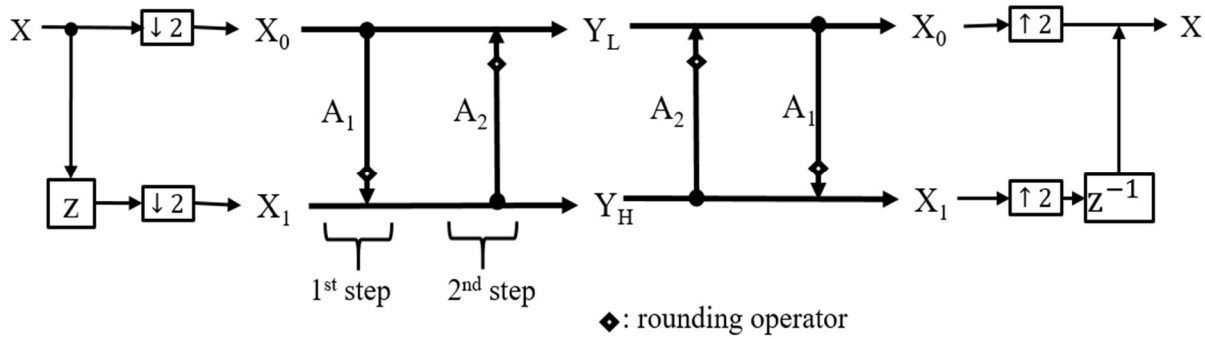


Fig. 14 Example of latency in DWT

3.3 Integer Wavelet Transform

3.3.1 One-Dimensional

Fig. 15 shows the forward and backward transforms of the integer WT developed for the 5/3-type transform of the lossless coding of a discrete 1D signal in JPEG 2000 [47]. Fig. 15(a) shows the forward transform of the integer WT and Fig. 15(b) shows its backward transform. It is composed of two lifting steps. The input signal X is down-sampled and fed into the forward transform, which is transformed into a low-frequency band signal, Y_L, and a high-frequency band signal Y_H. A_1 and A_2 are the coefficients of the filter bank, and Y_L and Y_H are coded with an entropy encoder to generate a bit stream for storage and communication. The band signals are then decoded and inversely transformed to obtain the reconstructed signal, X.



(a) Forward transform (b) Backward transform

Fig. 15 The 1D integer WT for 5/3-type transform (double lifting structure)

Fig. 16 shows a 9/7-type transform developed for lossy coding. Two more lifting steps and scaling with a constant k are added in this type.

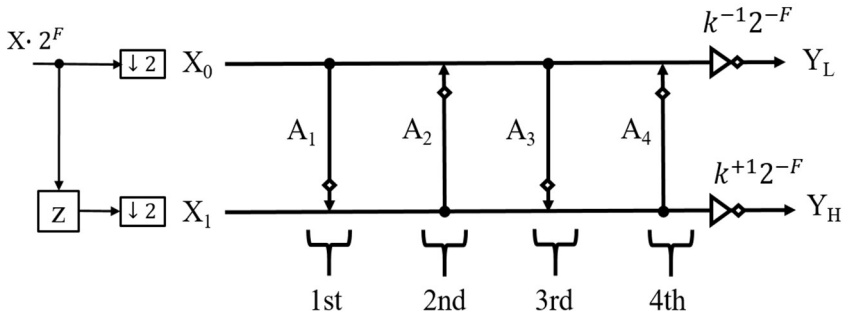


Fig. 16 The integer WT of the 9/7-type transform for lossy coding (quadruple lifting structure)

In detail, the input signal $x(n)$, n= 0,1,⋯, N-1 is divided into two groups $x_0(m)$ and $x_1(m)$, m= 0,1,⋯, M-1, M=N/2. It is expressed with the z transform as

$$X_c(z) = \downarrow 2[z^c X(z)], \quad c \in \{0,1\}, \tag{3.7}$$

for

$$\downarrow 2[X(z)] = \frac{1}{Q}\sum_{p=0}^{Q-1} X\left(z^{\frac{1}{Q}} W_Q^p\right), \quad W_Q = e^{\frac{j2\pi}{Q}}, \tag{3.8}$$

where Q=2 and

$$X(z) = \sum_{n=0}^{N-1} x(n) z^{-n}, \tag{3.9}$$

Secondly, the 1st lifting step is applied as

$$X_1^{(1)}(z) = X_1(z) + R[A_1(z)X_0(z)], \tag{3.10}$$

and the 2nd lifting is applied as

$$X_0^{(2)}(z) = X_0(z) + R[A_2(z)X_1^{(1)}(z)], \tag{3.11}$$

where $A_1(z)$ and $A_2(z)$ are filters given as

$$\begin{bmatrix} A_1(z) \\ A_2(z) \end{bmatrix} = \begin{bmatrix} h_1(1+z^{-1}) \\ h_2(1+z^{-1}) \end{bmatrix}, \tag{3.12}$$

Finally, the frequency band signals are generated as

$$\begin{bmatrix} Y_L(z) \\ Y_H(z) \end{bmatrix} = \begin{bmatrix} X_0^{(2)} \\ X_1^{(1)} \end{bmatrix}, \tag{3.13}$$

for

$$Y_b(z) = \sum_{m=0}^{M-1} y_b(m) z^{-m}, \quad b \in \{L, H\}, \tag{3.14}$$

Note that R[ ] denotes the rounding operator that truncates a pixel value in real numbers to an integer. Calculation in a lifting step starts after the calculation results of the previous have been obtained. The greater the number of lifting steps, the higher the latency (or delay). Therefore, we reduce the total numbers of lifting steps and rounding operators in the 4D integer WT in the 5/3-type and 9/7-type transform. Note that this dissertation focuses on reducing rounding

noise in the transform to increase the coding performance for 4D data in lossless and lossy mode.

In 5/3 type transform, the coefficient values are given as

$$\begin{bmatrix} h_1 & h_3 & k^{-1} \\ h_2 & h_4 & k \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & 0 & 1 \\ \frac{1}{4} & 0 & 1 \end{bmatrix}, \qquad (3.15)$$

Lossless reconstruction can be guaranteed with the scaling factors, k^{-1} and k are 1.

The 9/7 type transform has two more lifting steps and scaling. Namely, the 3rd lifting step

$$X_1^{(3)}(z) = X_1^{(1)}(z) + R[A_3(z)X_0^{(2)}(z)], \qquad (3.16)$$

and the 4th lifting is applied as

$$X_0^{(4)}(z) = X_0^{(2)}(z) + R[A_4(z)X_1^{(3)}(z)], \qquad (3.17)$$

where $A_3(z)$ and $A_4(z)$ are filters given as

$$\begin{bmatrix} A_3(z) \\ A_4(z) \end{bmatrix} = \begin{bmatrix} h_3(1+z^{-1}) \\ h_4(1+z^{-1}) \end{bmatrix}, \qquad (3.18)$$

In 9/7 type transform, the coefficient values are given as

$$\begin{cases} h_1 = -1.586134342059924 \\ h_2 = -0.052980118572961 \\ h_3 = 0.882911075530934 \\ h_4 = 0.443506852043971 \\ k = 1.230174104914001 \end{cases}, \qquad (3.19)$$

Finally, the frequency band signals are generated with scaling as

$$\begin{bmatrix} Y_L(z) \\ Y_H(z) \end{bmatrix} = \begin{bmatrix} R[k^{-1}2^{-F}X_0^{(4)}] \\ R[k^{+1}2^{-F}X_1^{(3)}] \end{bmatrix}, \qquad (3.20)$$

Note that the input signal is scaled with 2^F beforehand as shown in Fig. 16. In the integer implementation, F is set as a positive number. The smaller the F is, the shorter the bit depth of signals inside the transform will be.

3.3.2 Two-Dimensional (Non-Separable)

Fig. 17 illustrates the two-dimensional non-separable double lifting structure designed for 5/3-type filter. The non-separable structure is derived from the separable structure, which is the two one-dimensional double lifting structure with horizontal and vertical dimension in it.

The derivation process is explained in Section 4.3.2. Compared to the separable structure, the non-separable structure only has three lifting steps with four rounding operators. This structure is first proposed in [48].

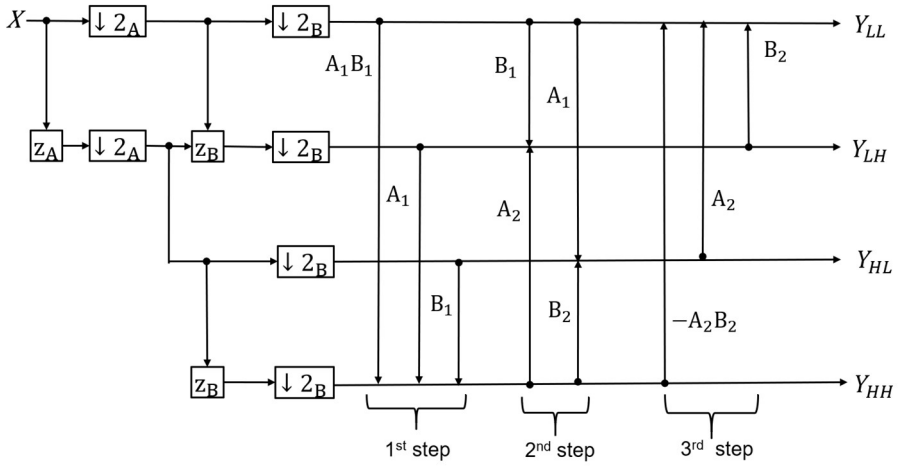


Fig. 17 Two-dimensional integer wavelet transform

3.3.3 Three-Dimensional (Non-Separable)

Comparing to Sep, the minimum lifting WT based on the non-separable structure has reduced number of lifting steps. As illustrated in Fig. 18, non-separable has 6 lifting steps. This is due to introduction of 3D memory accessing. For example, the signal in Y_{HHH} is produced as

$$X_{111}^{(1)}(\mathbf{z}) = X_{111}(\mathbf{z}) + R\left[P_{HHH}^{(1)}(\mathbf{z})\right] \qquad (3.21)$$

for

$$P_{HHH}^{(1)} = [V_1(\mathbf{z})H_1(\mathbf{z})D_1(\mathbf{z}) \quad V_1(\mathbf{z})H_1(\mathbf{z}) \quad V_1(\mathbf{z})D_1(\mathbf{z}) \quad V_1(\mathbf{z}) \quad H_1(\mathbf{z})D_1(\mathbf{z}) \quad H_1(\mathbf{z}) \quad D_1(\mathbf{z})] \cdot [X_{000} \quad X_{001} \quad X_{010} \quad X_{011} \quad X_{100} \quad X_{101} \quad X_{110}]^T \qquad (3.22)$$

in the 1st lifting step. In this step, a 3D filtering with 3D memory accessing $V_1(\mathbf{z})H_1(\mathbf{z})D_1(\mathbf{z})$ is used. In the 2nd lifting step, the calculation of Y_{HHL}, Y_{HLH} and Y_{LHH} :

$$\begin{bmatrix} X_{110}^{(1)}(\mathbf{z}) \\ X_{101}^{(1)}(\mathbf{z}) \\ X_{011}^{(1)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} X_{110}(\mathbf{z}) + R\left[P_{HHL}^{(1)}(\mathbf{z})\right] \\ X_{101}(\mathbf{z}) + R\left[P_{HLH}^{(1)}(\mathbf{z})\right] \\ X_{011}(\mathbf{z}) + R\left[P_{LHH}^{(1)}(\mathbf{z})\right] \end{bmatrix} \qquad (3.23)$$

for

$$\begin{bmatrix} P_{HHL}^{(1)}(\mathbf{z}) \\ P_{HLH}^{(1)}(\mathbf{z}) \\ P_{LHH}^{(1)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} V_1(\mathbf{z})H_1(\mathbf{z}) & 0 & V_1(\mathbf{z}) & H_1(\mathbf{z}) & D_2(\mathbf{z}) \\ V_1(\mathbf{z})D_1(\mathbf{z}) & V_1(\mathbf{z}) & 0 & D_1(\mathbf{z}) & H_2(\mathbf{z}) \\ H_1(\mathbf{z})D_1(\mathbf{z}) & H_1(\mathbf{z}) & D_1(\mathbf{z}) & 0 & V_2(\mathbf{z}) \end{bmatrix} \begin{bmatrix} X_{000}(\mathbf{z}) \\ X_{001}(\mathbf{z}) \\ X_{010}(\mathbf{z}) \\ X_{100}(\mathbf{z}) \\ X_{111}(\mathbf{z}) \end{bmatrix} \quad (3.24)$$

where R[] denotes the rounding operation on a signal value. Similarly, prediction of $X_{100}, X_{010}, X_{001}$ and X_{000} are also independent. As a result, the total number of the lifting steps and rounding operators of the non-separable structure in Fig. 18 is decreased from 6 to 4 (66.7 %) and from 24 to 8 (33.3%), respectively, comparing to the separable structure.

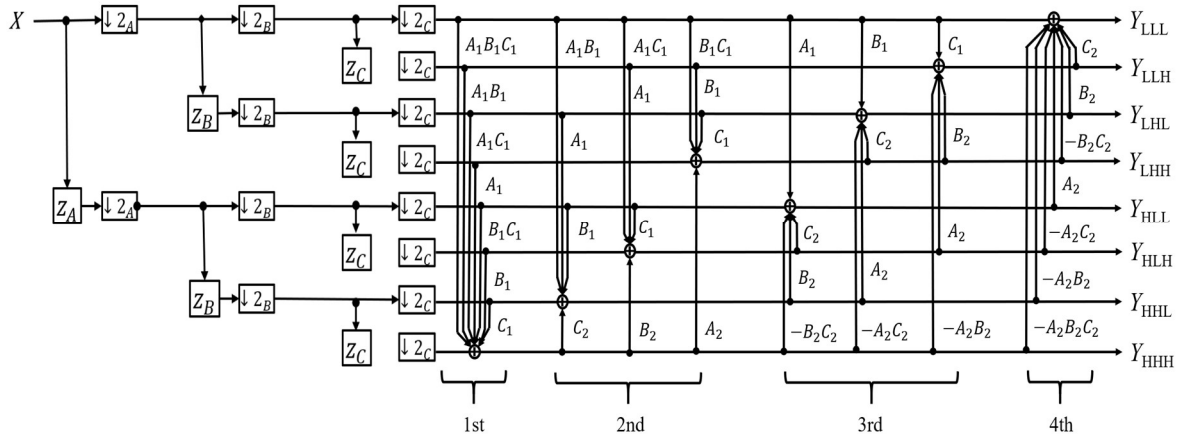


Fig. 18 Three-dimensional non-separable integer wavelet transform

Chapter 4 Four-Dimensional Non-Separable Integer Wavelet Transform with Double Lifting Structure for Lossless Compression

(This chapter is published in EURASIP Signal Processing Image Communication, Vol. 58, pp. 123-133, October 2017)

4.1 Proposal Motivation

The two-major type of image compression are lossless and lossy. Lossless techniques allow an exact reconstruction of the original image. Lossy techniques can reduce images by large ratios but do not perfectly reproduce the original image. Since every single bit of information in medical image is significantly essential, therefore a scalable lossless coding is necessary.

For still images, a two-dimensional (2D) transform such as the DCT and the DWT have been applied. Currently, various types of 3D transforms have been investigated for video, hyper spectral images, integral images and medical volumetric data [49, 50, 51, 52, 53, 54]. These 3D transforms have been recently extended to 4D transforms for the 4D medical image compression [4, 5, 55, 56]. This paper deals with a 4D WT with reduced amount of rounding noise inside the transform for high performance lossless data compression.

Over the past few decades, several constructions of compactly supported wavelet originating both from mathematical analysis and the signal processing community are designed. JPEG 2000, which has been approved as an international standard for the compression of still digital pictures, is currently wavelet-based unlike JPEG which use DCT [57]. There are two types of wavelet transform (WT) in JPEG 2000, which are the Daubechies (9,7) for lossy compression [38] and the (5,3) LeGall wavelet [37], which has rational coefficients, for the reversible or lossless compression. Note that the JPEG 2000 also supports arbitrary transform kernels. It also specifies that these should be executed using the lifting scheme [43].

A class of 'separable' 2D wavelet transform has been broadly developed for various applications. Since its transfer function is composed of a product of the horizontal 1D transfer function and the vertical one in spatial dimension, it can inherit legacy of previously designed 1D structure suitable for hardware implementations [58, 59]. It can also have the regularity and low sensitivity to various noises [60]. However, most of them are designed for lossy coding of images for practically reasonable data compression rate.

It is necessary to have transformed values as integers for entropy coding and inversely transform them to reconstruct the original integer pixel values without any loss for lossless coding of images. The lossless coding has been implemented in the lifting structure with rounding operations as the integer WT and extended based on the international standard, JPEG 2000 [61]. A set of forward and backward transform ensures lossless reconstruction of the original integer pixel values. However, there are rounding noise after the forward transform which will affect the efficiency of lossless coding structure. A non-separable structure was introduced [48] [62] reduce those noises from the rounding operations applied at each lifting steps as well as decreasing the total number of lifting steps.

The non-separable structures have been primarily introduced to increase the accuracy of prediction by adapting to local context of neighbouring pixels [63, 64]. Recently, directionality has been utilized in a generalized poly-phase representation [54, 65, 66]. They have been aiming on designing the adaptive high pass filters of the wavelet transforms.

A new class of non-separable 2D structure has been reported in [67, 44, 45]. The transfer function can be expressed as a product of two 1D functions. The transform based on this structure is compatible to the separable transform. The non-separable structure is not a cascade of 1D signal processing in 1D structure, but it requires the multidimensional memory accessing. This structure has an advantage that it will decrease the total number of lifting steps and the rounding operations inside it.

The non-separable 3D structure is also recently proposed in [62]. Unlike those previous studies on the 2D and 3D case, this chapter proposes a 4D integer WT for lossless coding of 4D signals with reduced amount of the rounding noise. The non-separable 4D structure has the advantage of rounding noise reduction inside the transform that will increase its coding performance. The existing studies of 4D image compression based on wavelet transform is using the separable structure [5] [56], hence, the non-separable 4D structure is proposed to increase the coding performance of 4D image compression. Therefore, in this proposal, the

variance of the noise in frequency domain, and the coding performance in lossless and lossy coding mode are clarified.

The proposed WT has a benefit that its output signals, apart from the rounding noise, is exactly the same as an existing transform whose transfer function is expressed as a product of 1D transfer functions. Moreover, the total amount of rounding noise is significantly reduced. It contributes to increase coding performance of data compression system based on 4D WT.

4.2 Existing Method

4.2.1 Four-Dimensional Separable Integer Wavelet Transform for 5/3 filter

Fig. 20 illustrates a 'separable' 4D structure. In JPEG 2000 standard, the 1D processing in Fig. 15 is applied to a 4D signal in *x*, *y*, z and *t* dimension, where *x* and *y* denotes the two spatial dimensions within a slice, the variable *z* denotes the third spatial dimension within a volume and *t* denotes the fourth temporal dimension. This structure has 8 lifting steps and 64 rounding operators for 2^4 input voxels. Thus, the rounding noise inside it is high. To cope with this problem, the minimum lifting WT was introduced [48, 44]. It is summarized as below.

For a 4D input signal $X(\mathbf{z})$, the transform splits the input signal into sixteen channels as illustrated in Fig. 19. It is denoted as

$$\begin{bmatrix} X_{0000}(\mathbf{z}) \\ X_{0010}(\mathbf{z}) \\ X_{0100}(\mathbf{z}) \\ X_{0011}(\mathbf{z}) \\ \vdots \\ X_{1110}(\mathbf{z}) \\ X_{1111}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_D \left[\begin{bmatrix} 1 \\ z_D \end{bmatrix} W_1(\mathbf{z}) \right] \\ \downarrow 2_D \left[\begin{bmatrix} 1 \\ z_D \end{bmatrix} W_2(\mathbf{z}) \right] \\ \vdots \\ \downarrow 2_D \left[\begin{bmatrix} 1 \\ z_D \end{bmatrix} W_8(\mathbf{z}) \right] \end{bmatrix}, \qquad (4.1)$$

where

$$\begin{bmatrix} W_1(\mathbf{z}) \\ W_2(\mathbf{z}) \\ W_3(\mathbf{z}) \\ W_4(\mathbf{z}) \\ W_5(\mathbf{z}) \\ W_6(\mathbf{z}) \\ W_7(\mathbf{z}) \\ W_8(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_1(\mathbf{z}) \right] \\ \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_2(\mathbf{z}) \right] \\ \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_3(\mathbf{z}) \right] \\ \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_4(\mathbf{z}) \right] \end{bmatrix}, \qquad (4.2)$$

$$\begin{bmatrix} V_1(\mathbf{z}) \\ V_2(\mathbf{z}) \\ V_3(\mathbf{z}) \\ V_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_B \left[\begin{bmatrix} 1 \\ z_B \end{bmatrix} P_1(\mathbf{z}) \right] \\ \downarrow 2_B \left[\begin{bmatrix} 1 \\ z_B \end{bmatrix} P_2(\mathbf{z}) \right] \end{bmatrix}$$

$$\begin{bmatrix} P_1(\mathbf{z}) \\ P_2(\mathbf{z}) \end{bmatrix} = \downarrow 2_A \left[\begin{bmatrix} 1 \\ z_A \end{bmatrix} X(\mathbf{z}) \right]$$

and

$$\begin{bmatrix}\downarrow 2_A[X(\mathbf{z})]\\ \downarrow 2_B[X(\mathbf{z})]\\ \downarrow 2_C[X(\mathbf{z})]\\ \downarrow 2_D[X(\mathbf{z})]\end{bmatrix}=\begin{bmatrix}X\left(z_A^{1/2},z_B,z_C,z_D\right)+X\left(-z_A^{1/2},z_B,z_C,z_D\right)\\ X\left(z_A,z_B^{1/2},z_C,z_D\right)+X\left(z_A,-z_B^{1/2},z_C,z_D\right)\\ X\left(z_A,z_B,z_C^{1/2},z_D\right)+X(z_A,z_B,-z_C^{1/2},z_D)\\ X\left(z_A,z_B,z_C,z_D^{1/2}\right)+X(z_A,z_B,z_C,-z_D^{1/2})\end{bmatrix}\cdot 2^{-1}\ , \qquad (4.3)$$

for

$$X(\mathbf{z})=\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\sum_{n_3=0}^{N_3-1}\sum_{n_4=0}^{N_4-1}X(\mathbf{n})z_A^{-n_1}z_B^{-n_2}z_C^{-n_4}z_D^{-n_4}, \qquad (4.4)$$

where $\mathbf{z}=(z_A, z_B, z_C, z_D)$ and $\mathbf{n}=(n_1, n_2, n_3, n_4)$.

In JPEG 2000 standard, applying the 1st and 2nd lifting steps vertically with

$$\begin{bmatrix}A_1(\mathbf{z})\\ A_2(\mathbf{z})\end{bmatrix}=\begin{bmatrix}h_1(1+z_A^{+1})\\ h_2(1+z_A^{-1})\end{bmatrix}, \qquad (4.5)$$

and the 3rd and 4th lifting steps horizontally with

$$\begin{bmatrix}B_1(\mathbf{z})\\ B_2(\mathbf{z})\end{bmatrix}=\begin{bmatrix}h_1(1+z_B^{+1})\\ h_2(1+z_B^{-1})\end{bmatrix}, \qquad (4.6)$$

and the 5th and 6th lifting steps in slice-wise with

$$\begin{bmatrix}C_1(\mathbf{z})\\ C_2(\mathbf{z})\end{bmatrix}=\begin{bmatrix}h_1(1+z_C^{+1})\\ h_2(1+z_C^{-1})\end{bmatrix}, \qquad (4.7)$$

and the 7th and 8th lifting steps temporally with

$$\begin{bmatrix}D_1(\mathbf{z})\\ D_2(\mathbf{z})\end{bmatrix}=\begin{bmatrix}h_1(1+z_D^{+1})\\ h_2(1+z_D^{-1})\end{bmatrix}, \qquad (4.8)$$

to the channel signals in equation (3.13), the transform outputs sixteen frequency band signals $Y_{LLLL}(\mathbf{z})$, $Y_{LLLH}(\mathbf{z})$, $Y_{LLHL}(\mathbf{z})$, $Y_{LLHH}(\mathbf{z})$, $Y_{LHLL}(\mathbf{z})$, $Y_{LHLH}(\mathbf{z})$, $Y_{LHHL}(\mathbf{z})$, $Y_{LHHH}(\mathbf{z})$, $Y_{HLLL}(\mathbf{z})$, $Y_{HLLH}(\mathbf{z})$, $Y_{HLHL}(\mathbf{z})$, $Y_{HLHH}(\mathbf{z})$, $Y_{HHLL}(\mathbf{z})$, $Y_{HHLH}(\mathbf{z})$, $Y_{HHHL}(\mathbf{z})$, and $Y_{HHHH}(\mathbf{z})$ as illustrated in Fig. 18. It is referred to as separable structure. The separable structure has lots of rounding operations, which will increase the rounding noise inside the transform. Therefore, the non-separable 4D structure proposed in Section 4.3.1 is based on two basic properties of modification clarified in Section 4.3.2 to overcome the problems stated above.

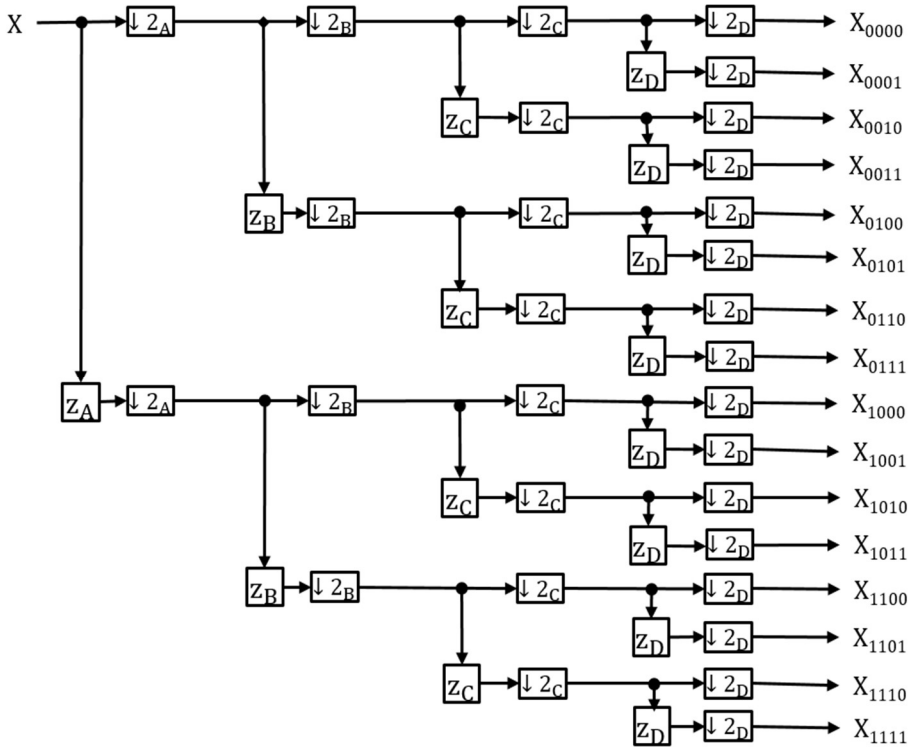


Fig. 19 Decomposition of a 4D signal

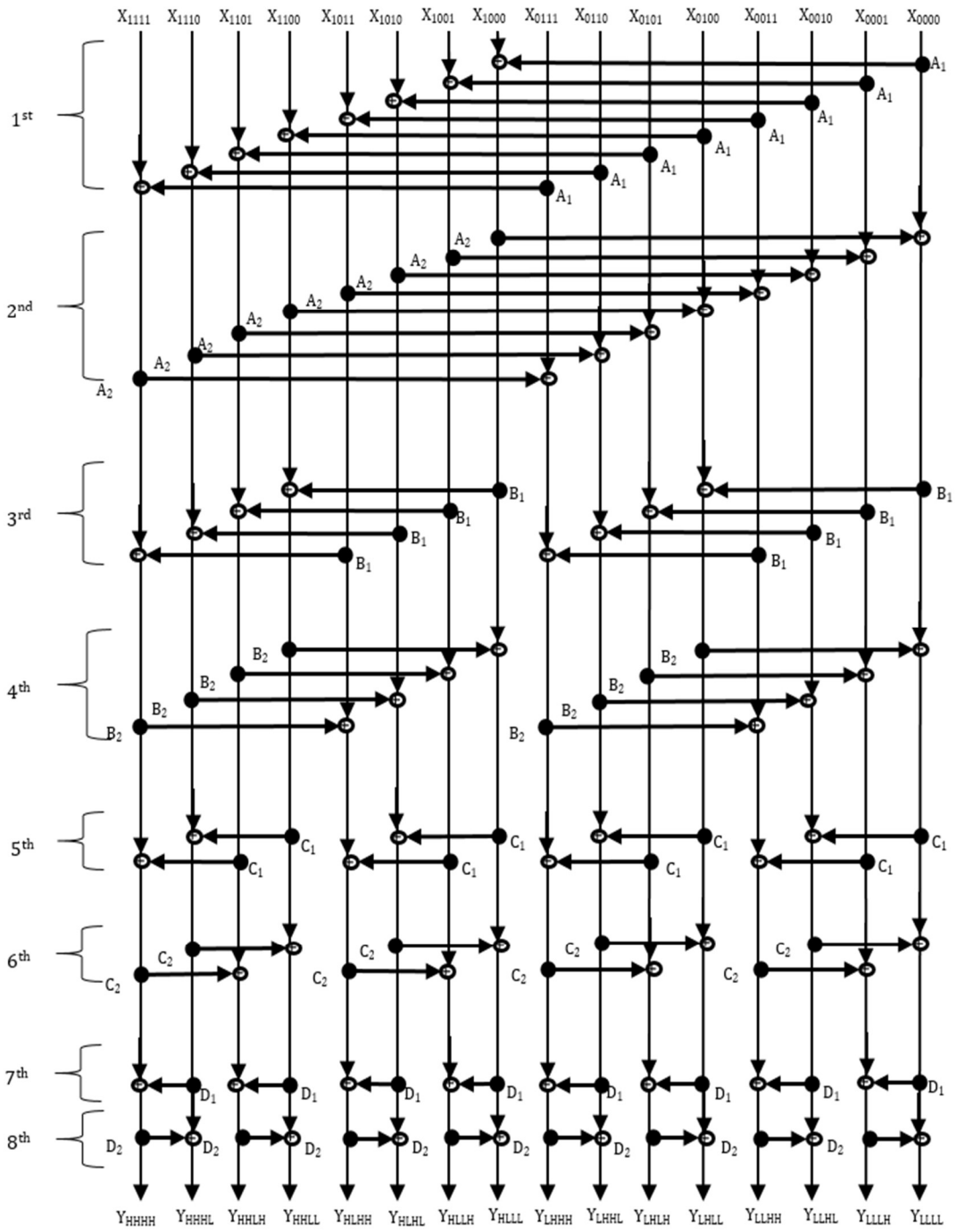


Fig. 20 Separable structure for 4D integer WT

4.3 Proposed Method

4.3.1 Four-Dimensional Non-Separable Integer Wavelet Transform for 5/3 filter

Fig. 21 illustrates the non-separable 4D lifting structure. In this structure, multi-input single-output lifting steps are introduced. Comparing to the existing separable structure, the total number of lifting steps is reduced from 8 to 5 (62.5%). The total number of rounding operations is also decreased from 64 to 16 (25%). The total amount of rounding noise is reduced as shown in Section 4.4.1. The derivation process of this structure is shown in Section 4.3.2.

For example, the signal in Y_{HHHH} is produced as

$$X_{1111}^{(1)}(\mathbf{z}) = X_{1111}(\mathbf{z}) + R\left[P_{HHHH}^{(1)}(\mathbf{z})\right], \qquad (4.9)$$

for

$$P_{HHHH}^{(1)} = \begin{bmatrix} A_1(\mathbf{z})B_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z}) \\ A_1(\mathbf{z})B_1(\mathbf{z})C_1(\mathbf{z}) \\ A_1(\mathbf{z})B_1(\mathbf{z})D_1(\mathbf{z}) \\ A_1(\mathbf{z})B_1(\mathbf{z}) \\ A_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z}) \\ A_1(\mathbf{z})C_1(\mathbf{z}) \\ A_1(\mathbf{z})D_1(\mathbf{z}) \\ A_1(\mathbf{z}) \\ B_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z}) \\ B_1(\mathbf{z})C_1(\mathbf{z}) \\ B_1(\mathbf{z})D_1(\mathbf{z}) \\ B_1(\mathbf{z}) \\ C_1(\mathbf{z})D_1(\mathbf{z}) \\ C_1(\mathbf{z}) \\ D_1(\mathbf{z}) \end{bmatrix}^T \cdot \begin{bmatrix} X_{0000} \\ X_{0001} \\ X_{0010} \\ X_{0011} \\ X_{0100} \\ X_{0101} \\ X_{0110} \\ X_{0111} \\ X_{1000} \\ X_{1001} \\ X_{1010} \\ X_{1011} \\ X_{1100} \\ X_{1101} \\ X_{1110} \end{bmatrix}, \qquad (4.10)$$

in the 1st lifting step. In this step, a 4D filtering with 4D memory accessing $A_1(\mathbf{z})B_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z})$ is used. In the 2nd lifting step, the calculation of Y_{HHHL}, Y_{HHLH}, Y_{HLHH} and Y_{LHHH} is shown in (4.11). It can be done simultaneously on a parallel signal processing platform since it is not necessary to wait for calculation results of each other.

$$\begin{bmatrix} X_{1110}^{(1)}(\mathbf{z}) \\ X_{1101}^{(1)}(\mathbf{z}) \\ X_{1011}^{(1)}(\mathbf{z}) \\ X_{0111}^{(1)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} X_{1110}(\mathbf{z}) + R\left[P_{HHHL}^{(1)}(\mathbf{z})\right] \\ X_{1101}(\mathbf{z}) + R\left[P_{HHLH}^{(1)}(\mathbf{z})\right] \\ X_{1011}(\mathbf{z}) + R\left[P_{HLHH}^{(1)}(\mathbf{z})\right] \\ X_{0111}(\mathbf{z}) + R\left[P_{LHHH}^{(1)}(\mathbf{z})\right] \end{bmatrix}, \qquad (4.11)$$

for

$$\mathbf{P} = \mathbf{H} \times \mathbf{X}, \quad (4.12)$$

where,

$$\mathbf{P} = \begin{bmatrix} P_{HHHL}^{(1)}(\mathbf{z}) \\ P_{HHLH}^{(1)}(\mathbf{z}) \\ P_{HLHH}^{(1)}(\mathbf{z}) \\ P_{LHHH}^{(1)}(\mathbf{z}) \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} A_1(\mathbf{z})B_1(\mathbf{z})C_1(\mathbf{z}) & A_1(\mathbf{z})B_1(\mathbf{z})D_1(\mathbf{z}) & A_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z}) & B_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z}) \\ 0 & A_1(\mathbf{z})B_1(\mathbf{z}) & A_1(\mathbf{z})C_1(\mathbf{z}) & B_1(\mathbf{z})C_1(\mathbf{z}) \\ A_1(\mathbf{z})B_1(\mathbf{z}) & 0 & A_1(\mathbf{z})D_1(\mathbf{z}) & B_1(\mathbf{z})D_1(\mathbf{z}) \\ 0 & 0 & A_1(\mathbf{z}) & B_1(\mathbf{z}) \\ A_1(\mathbf{z})C_1(\mathbf{z}) & A_1(\mathbf{z})D_1(\mathbf{z}) & 0 & C_1(\mathbf{z})D_1(\mathbf{z}) \\ 0 & A_1(\mathbf{z}) & 0 & C_1(\mathbf{z}) \\ A_1(\mathbf{z}) & 0 & 0 & D_1(\mathbf{z}) \\ B_1(\mathbf{z})C_1(\mathbf{z}) & B_1(\mathbf{z})D_1(\mathbf{z}) & C_1(\mathbf{z})D_1(\mathbf{z}) & 0 \\ 0 & B_1(\mathbf{z}) & C_1(\mathbf{z}) & 0 \\ B_1(\mathbf{z}) & 0 & D_1(\mathbf{z}) & 0 \\ C_1(\mathbf{z}) & D_1(\mathbf{z}) & 0 & 0 \\ D_2(\mathbf{z}) & C_2(\mathbf{z}) & B_2(\mathbf{z}) & A_2(\mathbf{z}) \end{bmatrix}^T$$

$$\mathbf{X} = \begin{bmatrix} X_{0000}(\mathbf{z}) \\ X_{0001}(\mathbf{z}) \\ X_{0010}(\mathbf{z}) \\ X_{0011}(\mathbf{z}) \\ X_{0100}(\mathbf{z}) \\ X_{0101}(\mathbf{z}) \\ X_{0110}(\mathbf{z}) \\ X_{1000}(\mathbf{z}) \\ X_{1001}(\mathbf{z}) \\ X_{1010}(\mathbf{z}) \\ X_{1100}(\mathbf{z}) \\ X_{1111}(\mathbf{z}) \end{bmatrix}$$

and $R[]$ denotes the rounding operation on a signal value. Similarly, prediction of $X_{1100}, X_{1010}, X_{1001}, X_{0110}, X_{0101}, X_{0011}, X_{1000}, X_{0100}, X_{0010}, X_{0001}$ and X_{0000} are also independent. The updating in the 3rd, 4th and 5th lifting steps are also performed.

Note that there is no difference between separable structure in Fig. 20 and the non-separable structure in Fig. 21 in respects of signals. It means that both of them are expressed

with the same transfer function which is a product of 1D function. In this sense, the transform to be implemented is separable. However, in case of noise, those are different. The structure in Fig. 21 is expected to have less rounding noise because it has fewer rounding operations. The rounding noise is investigated in Section 4.4.1.

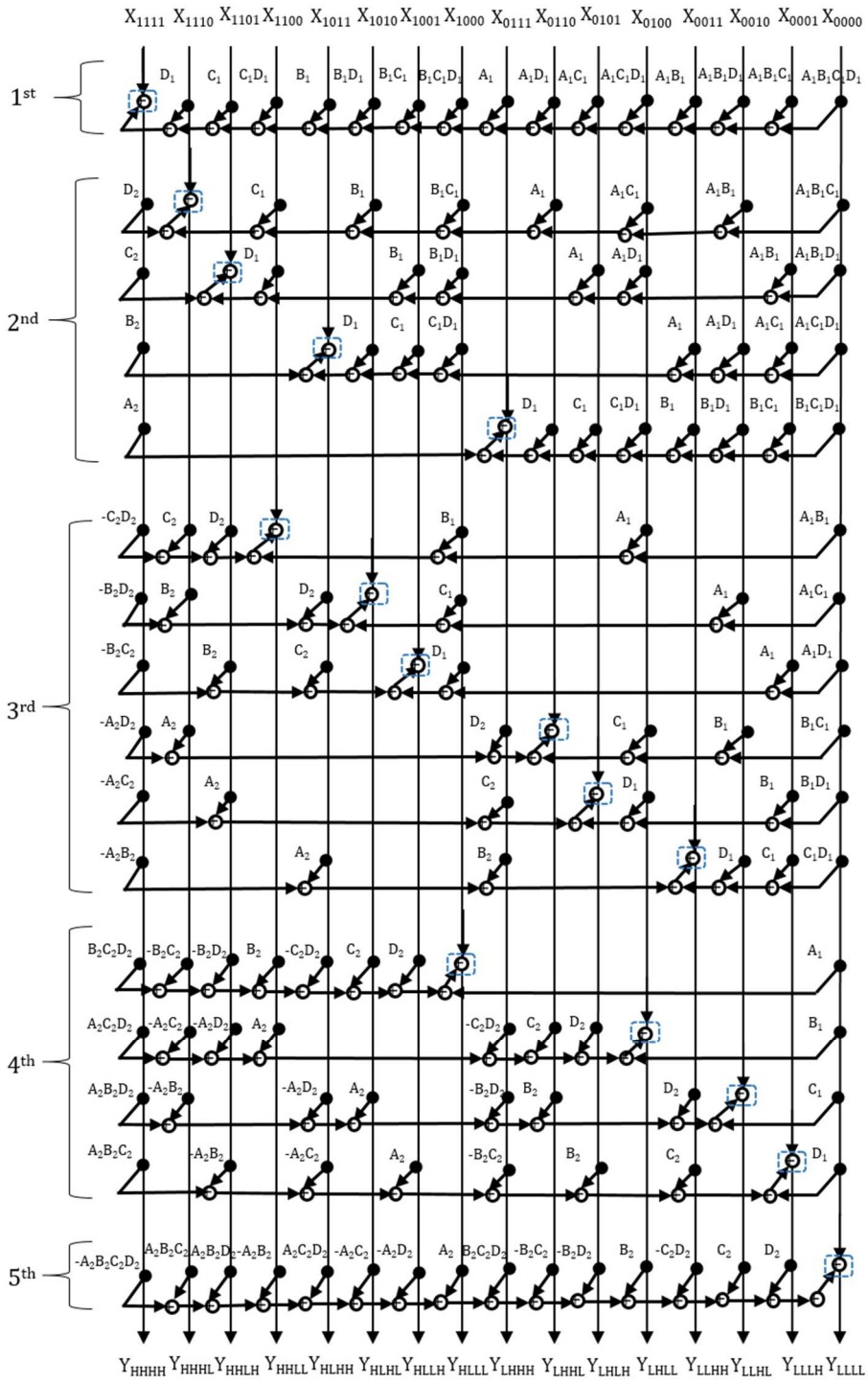


Fig. 21 Non-separable structure for 4D integer WT

4.3.2 Derivation of the Structure

4.3.2.1 Basic Properties for Modification

The derivation the non-separable structure in Fig. 21 from the separable structure in Fig. 20 is shown here. We derived it using two basic properties illustrated in Fig. 22, as shown in [46]. These are denoted as

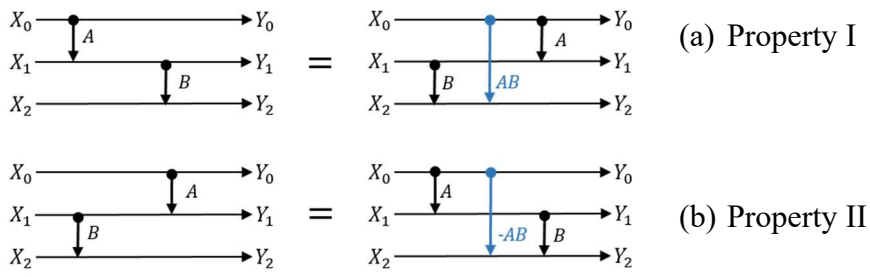
$$\mathrm{P_I}: \quad \mathbf{Y} = \mathbf{BA} \cdot \mathbf{X} = \mathbf{AC}_0\mathbf{B} \cdot \mathbf{X}, \qquad (4.13)$$

$$\mathrm{P_{II}}: \quad \mathbf{Y} = \mathbf{AB} \cdot \mathbf{X} = \mathbf{BC}_1\mathbf{A} \cdot \mathbf{X}, \qquad (4.14)$$

for

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ A & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{C}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +AB & 0 & 1 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix}, \qquad (4.15)$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & B & 1 \end{bmatrix}, \mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -AB & 0 & 1 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \end{bmatrix}, \qquad (4.16)$$



(a) Property I

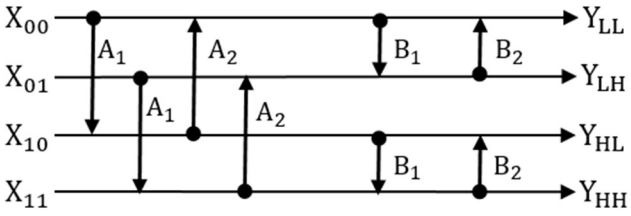
(b) Property II

Fig. 22 Basic properties for modification

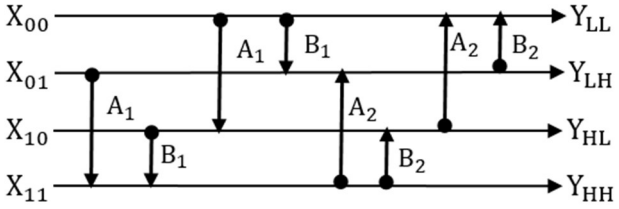
By utilizing both properties above, the derivation process of the non-separable structure for 4D WT is clarified. The basic properties for modification as shown in Fig. 22(a) and (b) have two different lifting procedures, but they are equivalent to each other as shown in equation 4.15 and 4.16. By adding property I and property II in the conventional separable structure, the lifting can be grouped together, thus reducing the number of lifting steps and rounding operations inside the transform. Therefore, the non-separable structure can be created systematically to reduce the rounding noise inside it as well as increasing the coding performance of it.

4.3.2.2 Derivation Process

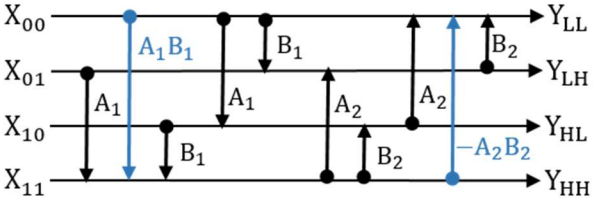
The derivation process of the non-separable 2D is shown in Fig. 23 based on [46]. Note that the same derivation process is done to get the non-separable 4D structure as shown in Fig. 21.



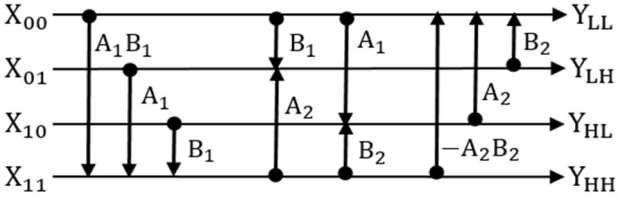
(a) Separable 2D structure



(b) Rearranging the lifting structure



(c) Applying property I and property II



(d) Non-separable 2D structure

Fig. 23 Derivation process for non-separable 2D

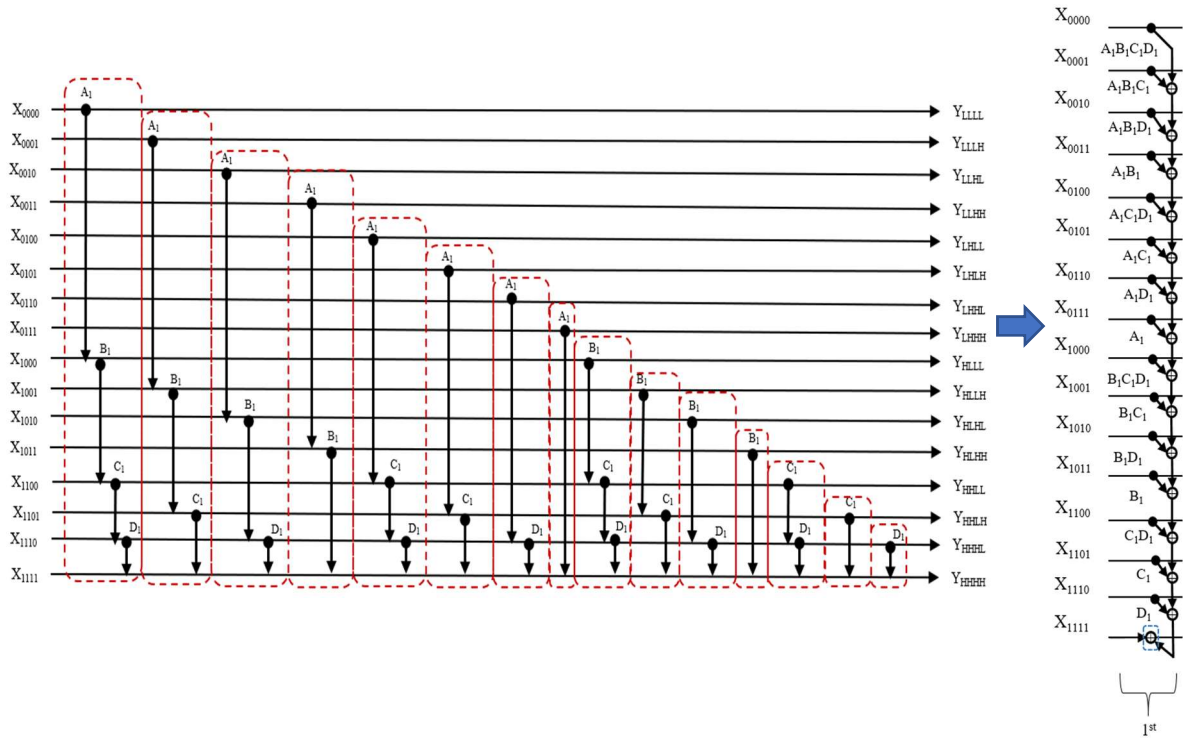
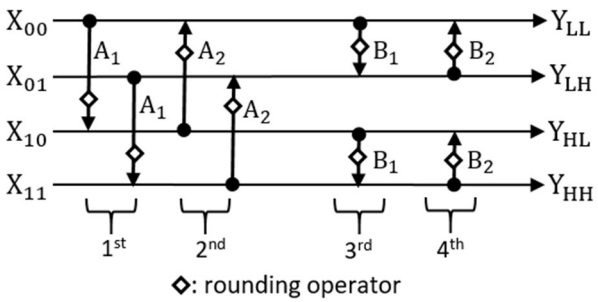


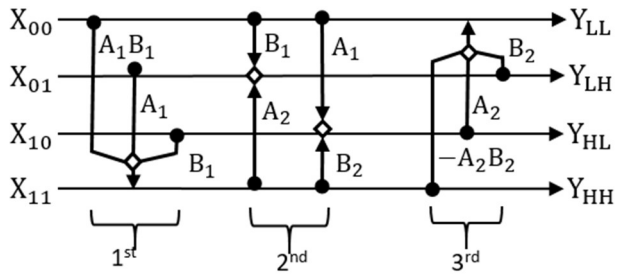
Fig. 24 Example derivation of 4D Non-Separable for the 1st lifting step

Fig. 24 shows an example of the 4D Non-Separable derivation for the 1st lifting step. All the filters, $A_1 \sim D_2$ are rearranged and modification properties as in Fig. 22 is applied to the structure. The number of lifting filters should be the same between the non-separable and separable structure.

Table 2 explains the variations of the structure of the 4D integer WT. In this paper, we compare the structures for five input signals from various aspects in Section 4.4. A clear comparison on number of rounding operators and lifting steps between separable and non-separable structure is explained in Fig. 25.



(a) Separable 2D structure with 8 rounding operators and 4 lifting steps



(b) Non-separable 2D structure with 4 rounding operators and 3 lifting steps

Fig. 25 Separable and non-separable 2D structures

Table 2 Comparison of the structures

	Separable 4D	**Non-separable 4D**
Rounding operations	64	16
Lifting steps	8	5
Memory accessing	1D	4D

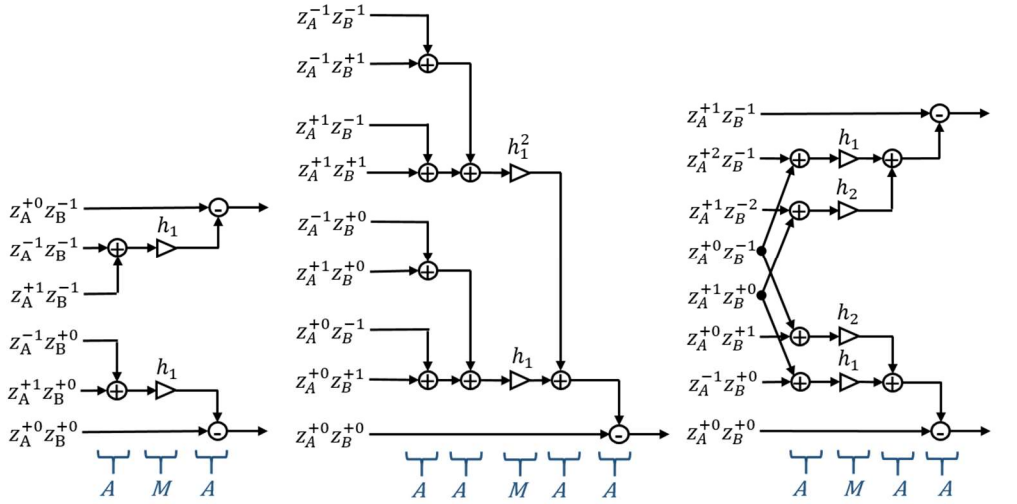
4.3.3 Lifting Steps and Latency

The non-separable structure can reduce rounding operations and lifting steps inside the transform. The fewer the number of lifting steps exist, the lower the overall latency of the transform as shown in [46]. Fig. 26(a) illustrates the implementation example for the 1st lifting step in separable 2D structure as in Fig. 25(a). In this structure, each adder is performed one-by-one in a parallel processor taking the latency denoted as '*A*'. Similarly, denoting the latency of multipliers as '*M*', it takes '*M+2A'* in total. Compared to the non-separable 2D structure in Fig. 25(b), four adders are simultaneously performed in a parallel processor for the 1st and 3rd lifting steps as shown in Fig. 26(b), which will take *'M+4A'* each of them. The 2nd lifting step shown in Fig. 26(c) is implemented as in Fig. 25(b), which will have *'M+3A'*. As a result, the separable 2D structure and the non-separable 2D structure takes *'4M+8A'* and *'3M+11A'*, respectively. Thus, the latency ratio is defined as

$$L = \frac{3\eta+11}{4\eta+8},\ \eta = \frac{M}{A}, \quad (4.17)$$

Latency due to the number or lifting steps is reduced if η>3. Note that the above estimation is valid only for examples in Fig. 26. As the non-separable 4D is designed according to the same derivation process of non-separable 2D, the implementation examples in Fig. 26 are also

applicable to the non-separable 4D structure. The proposed non-separable 4D structure will have lower latency compared to the separable 4D structure.



(a) 1st in Fig. 25(a) (b) 1st in Fig. 25(b) (c) 2nd in Fig. 25(b)

Fig. 26 Implementation examples in parallel processing platform

4.4 Experimental Results

In the following experiments, there are five types of data used to evaluate the rounding noise and their coding performance as shown in Table 3.

Table 3 Type of data used in the experiments

Type of data	**Size of data in $(\mathbf{x} \times \mathbf{y} \times \mathbf{z} \times \mathbf{t})$ sequence**	**Bit depth**	**Referred here as**
4D functional Magnetic Resonance Image	$56 \times 88 \times 32 \times 16$	16	fMRI (I)
4D Computed Tomography	$256 \times 256 \times 16 \times 16$	12	CT
4D functional Magnetic Resonance Image	$40 \times 64 \times 64 \times 16$	12	fMRI (II)
4D Magnetic Resonance Image	$52 \times 224 \times 64 \times 16$	8	MRI
4D Auto-regressive Model	$256 \times 256 \times 32 \times 16$	8	AR

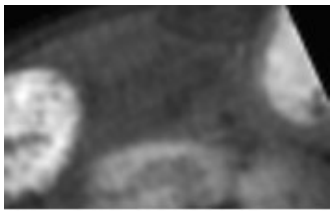
Note that the MRI data is retrieved from [68] and fMRI(II) data is retrieved from [69] . MRI data represent highly correlated data at consecutive time points with limited motion in the temporal dimension *t* and structural changes in the spatial dimension *z*.

Each frequency band signals are normalized to the range [0,255] for display purpose in this figure as shown in Fig. 27 (a) – (e). In this paper, the variance of the noise in frequency domain, and the coding performance in lossless and lossy coding mode are investigated.

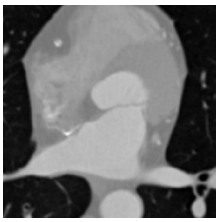
The 4D auto-regressive (AR) model included in our experiments is expressed as

$$\begin{cases} x^{(1)}(n_1,n_2,n_3,n_4) = x(n_1,n_2,n_3,n_4) + \rho \cdot x^{(1)}(n_1-1,n_2,n_3,n_4) \\ x^{(2)}(n_1,n_2,n_3,n_4) = x^{(1)}(n_1,n_2,n_3,n_4) + \rho \cdot x^{(2)}(n_1,n_2-1,n_3,n_4) \\ x^{(3)}(n_1,n_2,n_3,n_4) = x^{(2)}(n_1,n_2,n_3,n_4) + \rho \cdot x^{(3)}(n_1,n_2,n_3-1,n_4) \\ x^{(4)}(n_1,n_2,n_3,n_4) = x^{(3)}(n_1,n_2,n_3,n_4) + \rho \cdot x^{(4)}(n_1,n_2,n_3,n_4-1) \end{cases}, \quad (4.18)$$

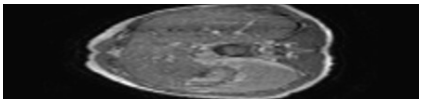
Note that ρ is set as 0.9 in the experiments in this paper.



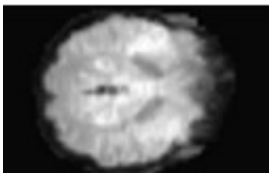
(a) fMRI (I)



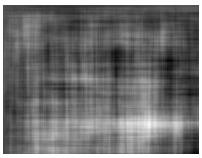
(b) CT



(c) MRI



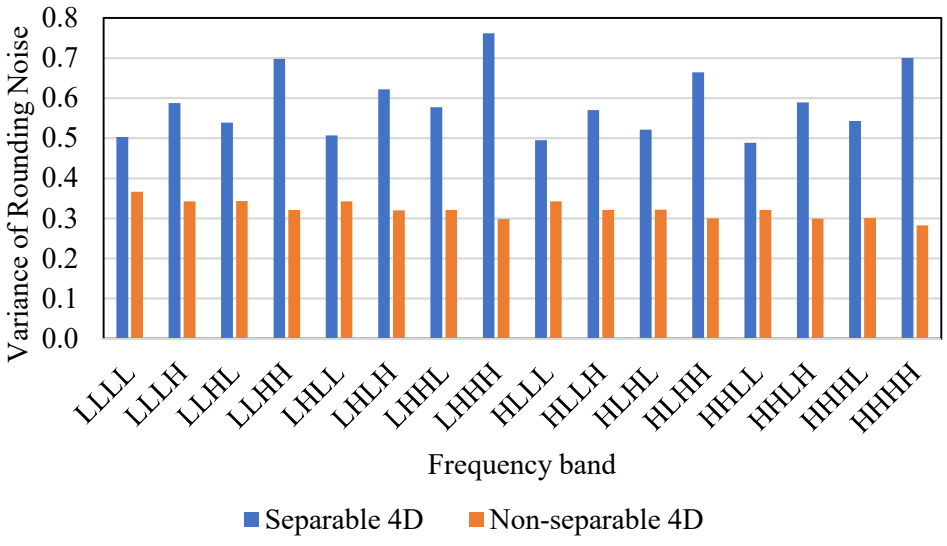
(d) fMRI (II)



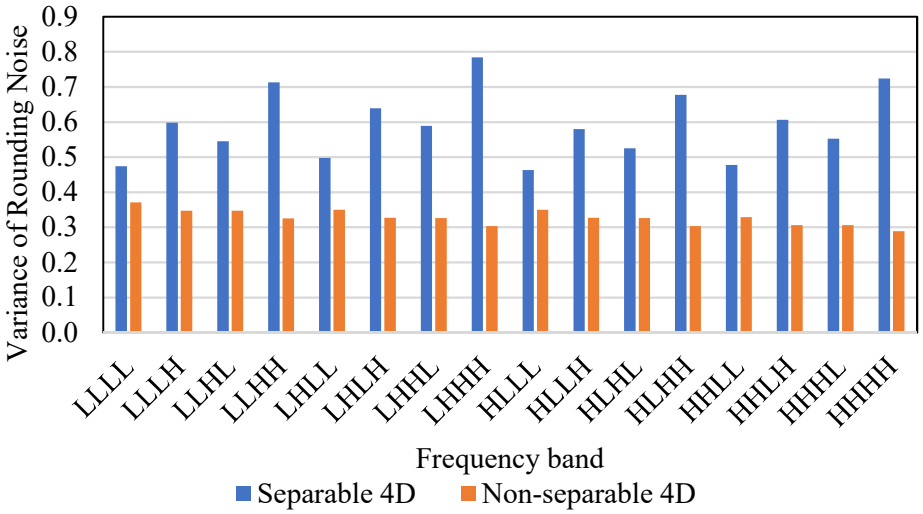
(e) AR

Fig. 27 Tested data

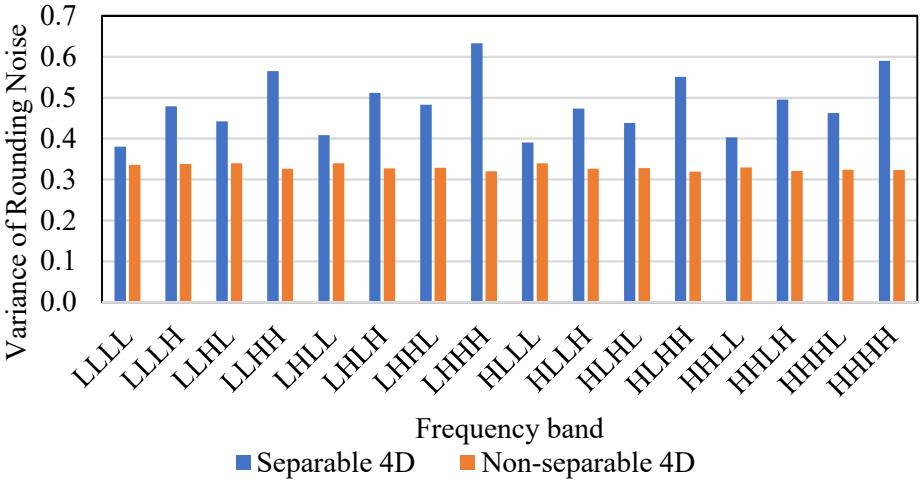
4.4.1 Evaluation on Rounding Noise



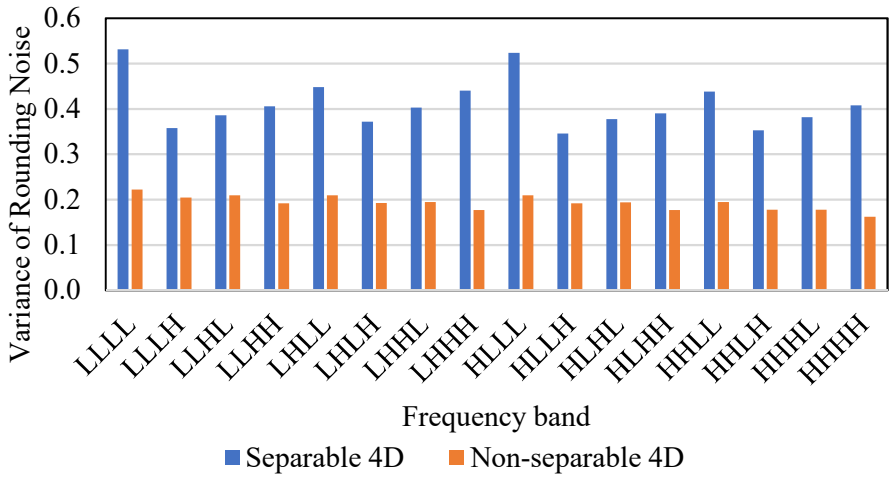
(a) fMRI (I)



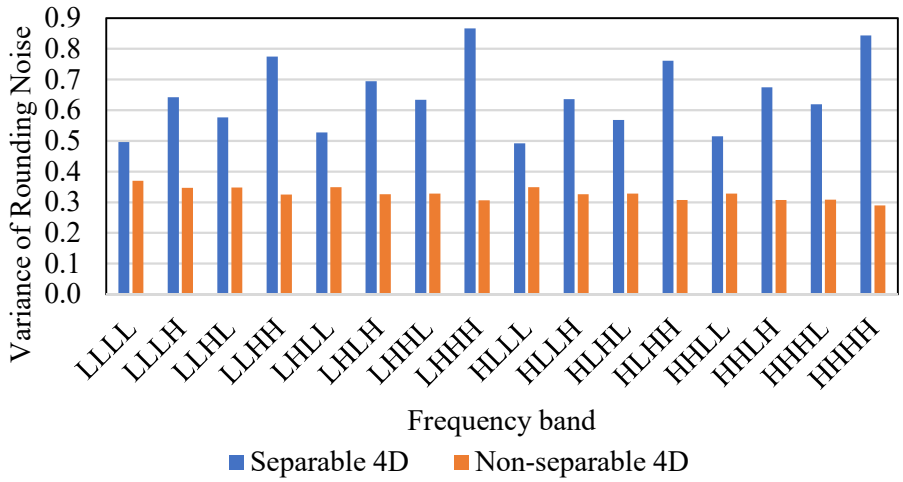
(b) CT



(c) MRI



(d) fMRI (II)

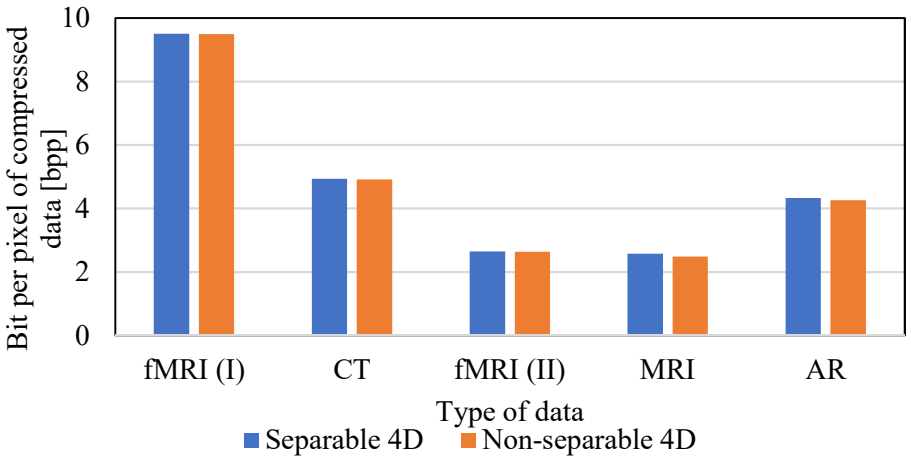


(e) AR

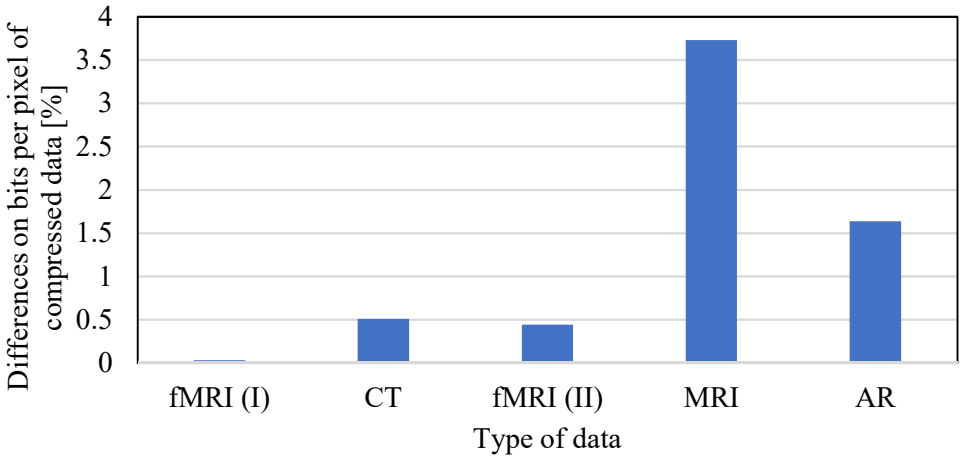
Fig. 28 Rounding noise in each frequency band

Since fewer rounding operators does not always imply fewer rounding errors in total [44, 45], total amount of rounding errors contained in the output frequency band signals is investigated. The difference between output from WT without rounding and that with rounding is defined as the error. Fig. 28 (a) - (e) indicate variance of the rounding error in each frequency bands for fMRI(I), CT, fMRI(II), MRI and AR data respectively. The average amount of rounding errors for all data are reduced to 50.12[%], 45.30[%], 52.85[%], 32.85[%] and 49.94[%] respectively. It is observed that variance of rounding noise in the separable 4D structure is distributed randomly in all type of data. In all cases, variance of the rounding error in 'HHHH' band of non-separable is the lowest. On the contrary, variance of rounding error in 'LLLL' band of non-separable is the highest. This is because multiple noise amplified through convolutions are summed up in low frequency band signal. By using the non-separable 4D integer WT structure, the energy is further compacted to the low frequency band signal, thus reducing the entropy rate of the compressed image and will eventually increase its coding performance. The higher the variance of rounding noise in low frequency band signal, the lower the entropy rate of the compressed image. The coding performance is evaluated in Section 4.4.2.

4.4.2 Evaluation on Coding Performance



(a) Lossless coding performance of all data

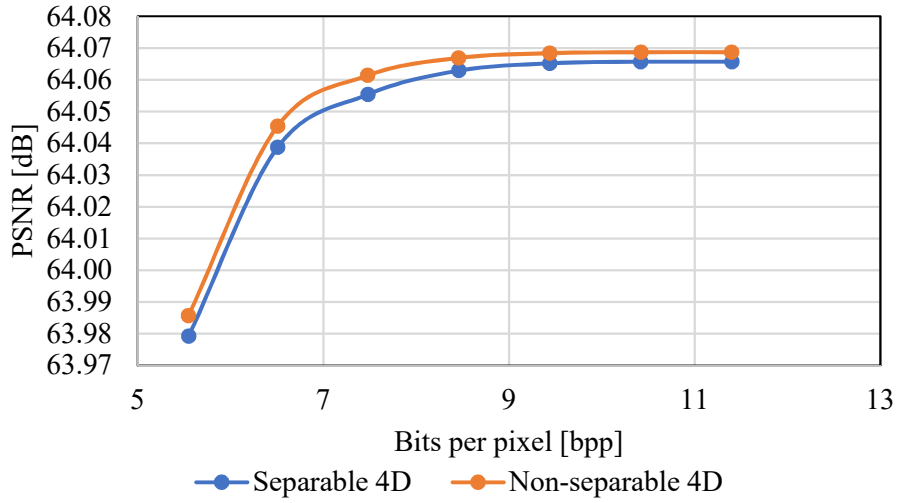


(b) Differences of lossless coding performance between the separable 4D and non-separable 4D structure for all data

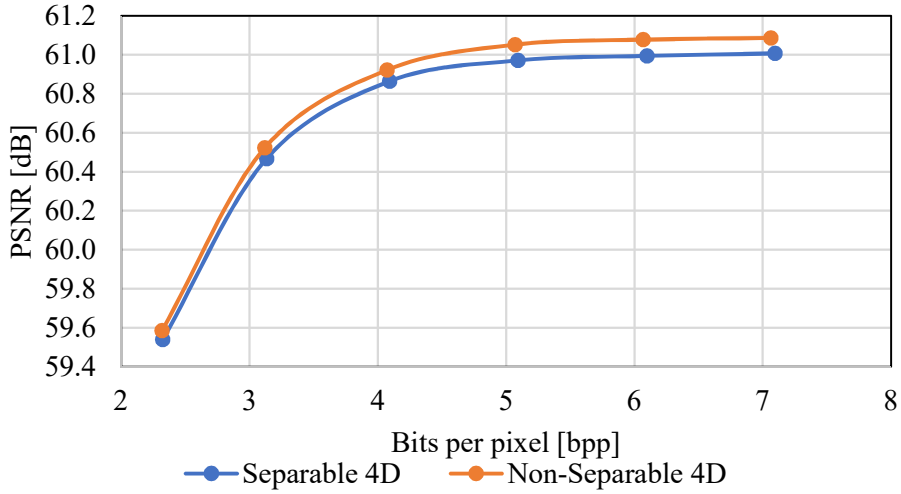
Fig. 29 Coding performance in lossless mode

Next, we investigate the entropy rate of the compressed data to evaluate lossless coding performance. Fig. 29(a) compares the separable and non-separable structure in respect of lossless data compression rate for fMRI(I), CT, fMRI(II), MRI and AR with various original bit depth. The transformed data is encoded by using the Embedded Zerotree Wavelet based on Intraband Partitioning (EZWIP) [70] encoder developed in three dimensional (3D), where each temporal subband are independently coded. The bit rate of compressed data is reduced by 0.04[%], 0.51[%], 0.44[%], 3.73[%], and 1.64[%] for fMRI(I), CT, fMRI(II), MRI and AR data respectively, as shown in Fig. 29(b). The MRI and AR data which has 8-bit depth, is

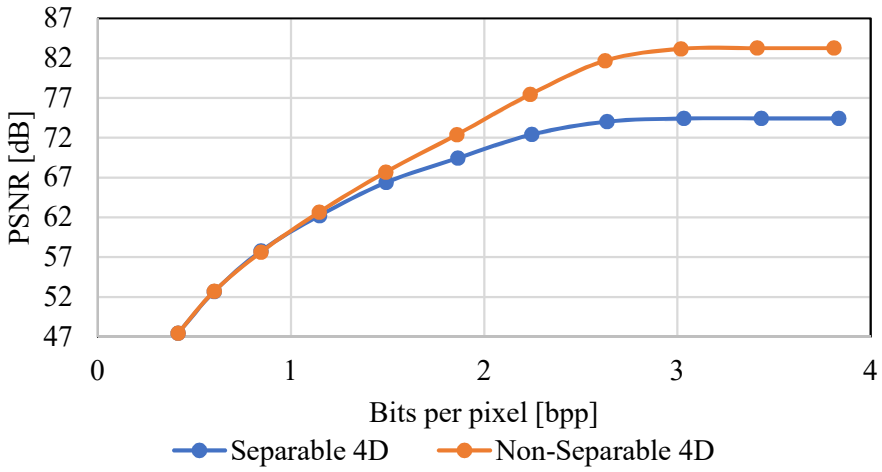
significantly reduced in this experiment, thus proving that the proposed structure can be applied for lossless compression of 4D image.



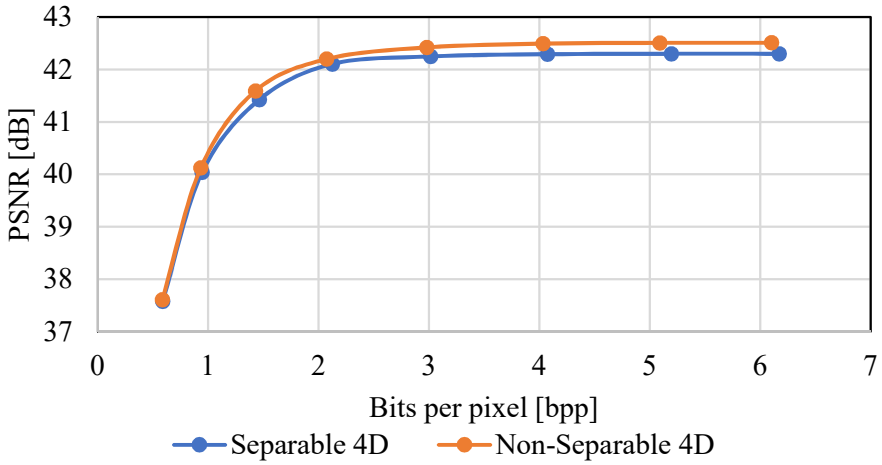
(a) fMRI (I)



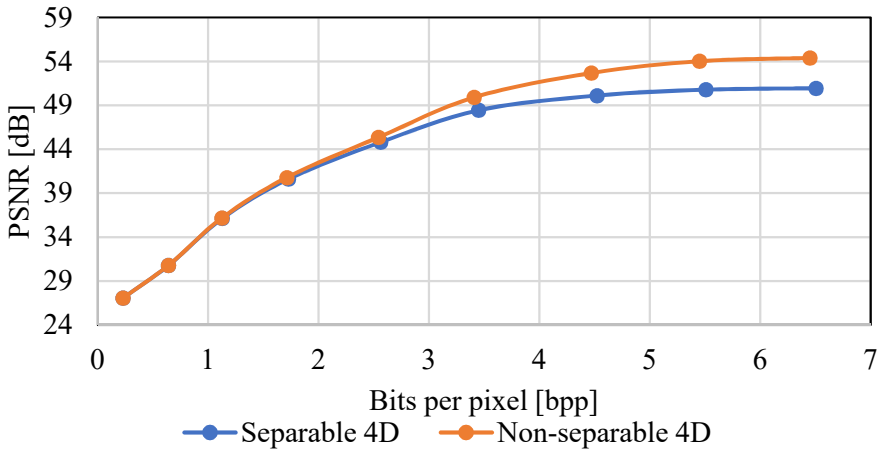
(b) CT



(c) fMRI (II)



(d) MRI



(e) AR

Fig. 30 Coding performance in lossy mode.

Fig. 30 (a) – (e) indicates rate-distortion curve of separable and non-separable structure of 4D WT tested in fMRI(I), CT, fMRI(II), MRI and AR data. The horizontal axis and the vertical axis represent the entropy rate measured in bits per pixel [bpp] and peak signal noise-to ratio (PSNR) of the reconstructed signal, respectively. These curves indicate performance of the methods in lossy coding mode. It was observed that non-separable structure performed better than the separable structure. Under the same bit rate in each data, the PSNR is increased by 0.007[dB], 0.08[dB], 8.82[dB], 0.21[dB] and 3.24[dB] for the fMRI(I), CT, fMRI(II), MRI and AR data, respectively. Compared to the lossless coding performance, the non-separable 4D performed better in lossy coding. This is because, there are both quantization noise and rounding noise inside the compression system. The quantization noise for separable and non-separable 4D is the same, but, the rounding noise in the non-separable is lower than in the separable, thus, increasing the coding performance in lossy mode. Unlike the lossless coding mode, there is only rounding noise inside the system which will only give significant effect to the difference of performance. In conclusion, the non-separable structure outperformed the conventional separable structure in these experiments.

4.5 Conclusion of this Chapter

A non-separable 4D integer WT compatible with the existing separable 4D double lifting integer WT is proposed. Non-separable multidimensional lifting structure is introduced to decrease the total number of lifting steps as well as the rounding operators. By introducing 4D direct memory accessing, the total number of lifting steps is reduced by 37.5[%]. The total number of rounding operators is also reduced by 75[%], and further reducing the rounding noise by approximately 50[%], thus increasing its coding performance. It is advantageous to implement our integer WT which has compatibility with a WT module that has been broadly used as an international standard nowadays.

Chapter 5 Four-Dimensional Non-Separable Integer Wavelet Transform with Quadruple Lifting Structure for Lossy Compression

(This chapter is published in EURASIP Journal on Image and Video Processing, Vol. 36, pp. 1-18, May 2018 and presented in IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Mar 2017)

5.1 Proposal Motivation

Various types of wavelet transforms have been reported to compress the 4D images [56], 4D hyperspectral images [55], 4D medical volumetric data [4, 5], 4D light field data [71], and 4D color images [72]. However, most of them use the separable 4D WT that contains a large amount of rounding noise. Therefore, a non-separable 3D integer WT [62] and 4D integer WT in Chapter 4 was proposed to overcome its limitations. Unfortunately, this was limited to a double-lifting integer WT with a 5/3 filter especially applied for lossless coding. A non-separable quadruple 3D WT with a 9/7 filter was subsequently proposed in [73] for lossy coding. Nevertheless, unlike in the double-lifting WT, the variance of rounding noise increased in the quadruple lifting WT even though the number of lifting steps decreased. Rounding noise in the transform can reduce the efficiency of the lossy coding structure. This research is the first to use a non-separable 4D quadruple lifting integer WT with the aim of reducing rounding noise inside the transform as well as improving its coding performance. Note that a part of this chapter was presented in [74].

This chapter focuses on lossy coding of 4D signals using the 9/7-type transform based on the quadruple lifting steps, and a reduction in rounding noise in the integer implementation of the transform is affected. As a lifting step needs to wait for the results of calculations from the previous lifting step, many sequential lifting steps incur a long delay between input and output. The real numbers assumed as signal values inside the transform are rounded into finite-length rational numbers. Shorter lengths imply lower computational load but more rounding noise. The space needed for memory storage can be reduced in a trade-off with rounding noise [75].

A non-separable 2D quadruple lifting structure for 4D input signals to deal with the problem of degradation in image quality due to integer implementation is proposed in this chapter. It has the advantage that its output signals, apart from rounding noise, are identical to those of a conventional transform the transfer function of which is a product of 16 1D transfer functions. Unlike the prevalent 3D and 4D quadruple lifting structure, the order of lifting steps in the original separable 4D transform is preserved. Thus, the total rounding noise is reduced even though the total number of rounding operators remain the same as in prevalent methods. Experiments confirmed that the total rounding noise observed in each frequency band of the decoded images was significantly smaller. The upper bounds of the quality-decoded images in lossy coding mode also improved.

5.2 Existing Methods

5.2.1 Separable 4D Integer Wavelet Transform for 9/7 filter (Existing I)

Fig. 32 shows the 9/7-type separable 4D integer WT. In the JPEG 2000 standard, the 1D processing shown in Fig. 16 is applied to a 4D signal along the *x*, *y*, *z,* and *t* dimensions, where *x* and *y* denote two spatial dimensions within a slice, *z* denotes the third spatial dimension within a volume, and *t* denotes the fourth, temporal, dimension. However, the separable 4D structure increases the number of rounding operators in the transform. This structure has 192 rounding operators.

For a 4D input signal $X(\mathbf{z})$, the transform splits the input signal into 16 channels, X_{0000}, X_{0001}, X_{0010}, X_{0011}, X_{0100}, X_{0101}, X_{0110}, X_{0111}, X_{1000}, X_{1001}, X_{1010}, X_{1011}, X_{1100}, X_{1101}, X_{1110}, and X_{1111} as shown in Fig. 31. It is denoted as

$$\begin{bmatrix} X_{0000}(\mathbf{z}) \\ X_{0010}(\mathbf{z}) \\ X_{0100}(\mathbf{z}) \\ X_{0011}(\mathbf{z}) \\ \vdots \\ X_{1110}(\mathbf{z}) \\ X_{1111}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_D \left[\begin{bmatrix} 1 \\ z_D \end{bmatrix} W_1(\mathbf{z}) \right] \\ \downarrow 2_D \left[\begin{bmatrix} 1 \\ z_D \end{bmatrix} W_2(\mathbf{z}) \right] \\ \vdots \\ \downarrow 2_D \left[\begin{bmatrix} 1 \\ z_D \end{bmatrix} W_8(\mathbf{z}) \right] \end{bmatrix}, \quad (5.1)$$

where

$$\begin{bmatrix} W_1(\mathbf{z}) \\ W_2(\mathbf{z}) \\ W_3(\mathbf{z}) \\ W_4(\mathbf{z}) \\ W_5(\mathbf{z}) \\ W_6(\mathbf{z}) \\ W_7(\mathbf{z}) \\ W_8(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_1(\mathbf{z}) \right] \\ \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_2(\mathbf{z}) \right] \\ \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_3(\mathbf{z}) \right] \\ \downarrow 2_C \left[\begin{bmatrix} 1 \\ z_C \end{bmatrix} V_4(\mathbf{z}) \right] \end{bmatrix} \quad , (5.2)$$

$$\begin{bmatrix} V_1(\mathbf{z}) \\ V_2(\mathbf{z}) \\ V_3(\mathbf{z}) \\ V_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_B \left[\begin{bmatrix} 1 \\ z_B \end{bmatrix} P_1(\mathbf{z}) \right] \\ \downarrow 2_B \left[\begin{bmatrix} 1 \\ z_B \end{bmatrix} P_2(\mathbf{z}) \right] \end{bmatrix}$$

$$\begin{bmatrix} P_1(\mathbf{z}) \\ P_2(\mathbf{z}) \end{bmatrix} = \downarrow 2_A \left[\begin{bmatrix} 1 \\ z_A \end{bmatrix} X(\mathbf{z}) \right]$$

and

$$\begin{bmatrix} \downarrow 2_A[X(\mathbf{z})] \\ \downarrow 2_B[X(\mathbf{z})] \\ \downarrow 2_C[X(\mathbf{z})] \\ \downarrow 2_D[X(\mathbf{z})] \end{bmatrix} = \begin{bmatrix} \frac{1}{Q}\sum_{p=0}^{Q-1} X\left(z_A^{1/Q}\cdot W_Q^p, z_B, z_C, z_D\right) \\ \frac{1}{Q}\sum_{p=0}^{Q-1} X\left(z_A, z_B^{1/Q}\cdot W_Q^p, z_C, z_D\right) \\ \frac{1}{Q}\sum_{p=0}^{Q-1} X\left(z_A, z_B, z_C^{1/Q}\cdot W_Q^p, z_D\right) \\ \frac{1}{Q}\sum_{p=0}^{Q-1} X\left(z_A, z_B, z_C, z_D^{1/Q}\cdot W_Q^p\right) \end{bmatrix} \cdot 2^F \quad , \qquad (5.3)$$

for

$$X(\mathbf{z}) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\sum_{n_3=0}^{N_3-1}\sum_{n_4=0}^{N_4-1} X(\mathbf{n}) z_A^{-n_1} z_B^{-n_2} z_C^{-n_4} z_D^{-n_4}, \qquad (5.4)$$

where $\mathbf{z}=(z_A, z_B, z_C, z_D)$ and $\mathbf{n}=(n_1, n_2, n_3, n_4)$.

In JPEG 2000 standard, applying the 1st, 2nd, 3rd and 4th lifting steps in the spatial dimension, *x* with

$$\begin{bmatrix} A_1(\mathbf{z}) & A_3(\mathbf{z}) \\ A_2(\mathbf{z}) & A_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_1(1+z_A^{+1}) & h_3(1+z_A^{+1}) \\ h_2(1+z_A^{-1}) & h_4(1+z_A^{-1}) \end{bmatrix}, \qquad (5.5)$$

and the 5th, 6th, 7th and 8th lifting steps in the spatial dimension, *y* with

$$\begin{bmatrix} B_1(\mathbf{z}) & B_3(\mathbf{z}) \\ B_2(\mathbf{z}) & B_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_1(1+z_B^{+1}) & h_3(1+z_B^{+1}) \\ h_2(1+z_B^{-1}) & h_4(1+z_B^{-1}) \end{bmatrix}, \qquad (5.6)$$

and the 9th, 10th, 11th and 12th lifting steps in the spatial dimension, *z* with

$$\begin{bmatrix} C_1(\mathbf{z}) & C_3(\mathbf{z}) \\ C_2(\mathbf{z}) & C_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_1(1+z_C^{+1}) & h_3(1+z_C^{+1}) \\ h_2(1+z_C^{-1}) & h_4(1+z_C^{-1}) \end{bmatrix}, \qquad (5.7)$$

and the 13th, 14th, 15th and 16th lifting steps in the temporal dimension, *t* with

$$\begin{bmatrix} D_1(\mathbf{z}) & D_3(\mathbf{z}) \\ D_2(\mathbf{z}) & D_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_1(1+z_D^{+1}) & h_3(1+z_D^{+1}) \\ h_2(1+z_D^{-1}) & h_4(1+z_D^{-1}) \end{bmatrix}, \qquad (5.8)$$

to the channel signals in (3.13), the transform outputs sixteen frequency band signals $Y_{LLLL}(\mathbf{z})$, $Y_{LLLH}(\mathbf{z})$, $Y_{LLHL}(\mathbf{z})$, $Y_{LLHH}(\mathbf{z})$, $Y_{LHLL}(\mathbf{z})$, $Y_{LHLH}(\mathbf{z})$, $Y_{LHHL}(\mathbf{z})$, $Y_{LHHH}(\mathbf{z})$, $Y_{HLLL}(\mathbf{z})$, $Y_{HLLH}(\mathbf{z})$, $Y_{HLHL}(\mathbf{z})$, $Y_{HLHH}(\mathbf{z})$, $Y_{HHLL}(\mathbf{z})$, $Y_{HHLH}(\mathbf{z})$, $Y_{HHHL}(\mathbf{z})$, and $Y_{HHHH}(\mathbf{z})$ as illustrated in Fig. 32. This is referred to as a separable structure. As it has a large number of rounding operators, there is a large volume of rounding noise in the transform. A non-separable 3D structure was thus proposed in [62] and a non-separable 4D structure is proposed in Chapter 4. However, when used for a 4D signal of quadruple lifting structure, the rounding noise in it increases compared with that in a separable 4D structure. Thus, its coding performance is significantly affected by the rounding noise generated inside it.

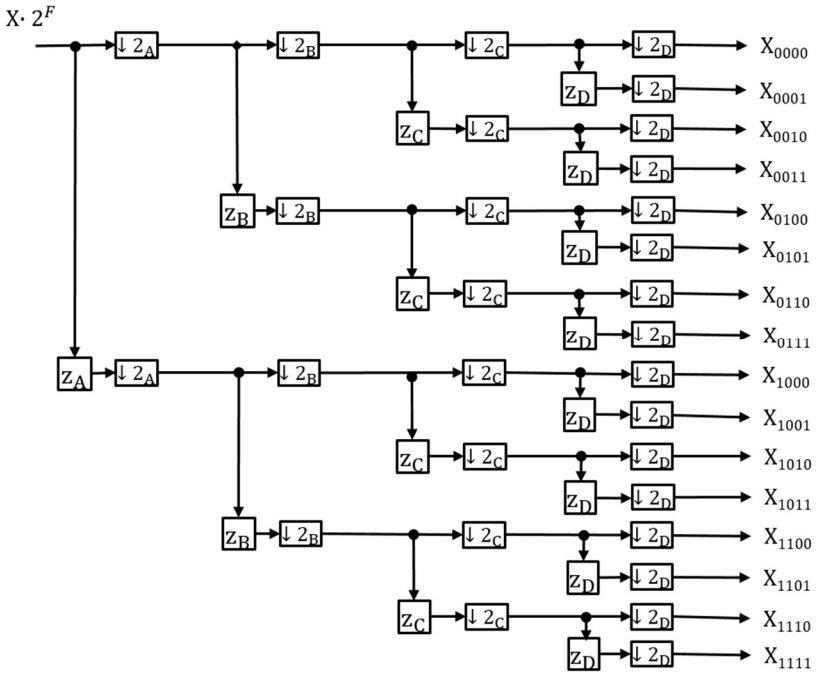


Fig. 31 Decomposition of a 4D signal

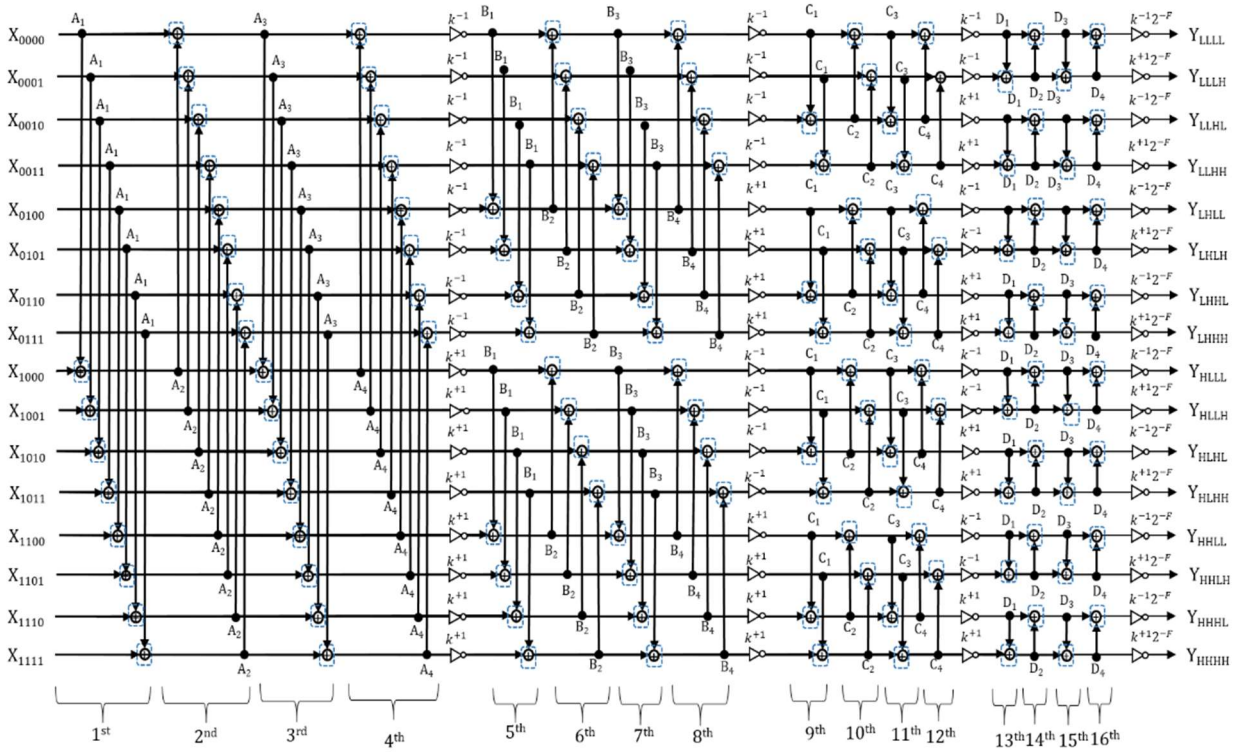


Fig. 32 Separable 4D structure for 9/7-type of transform (Existing I)

5.2.2 Non-Separable 3D Integer Wavelet Transform for 4D 9/7 filter (Existing II)

Fig. 33 shows the non-separable 3D structure of integer WT for a 4D input signal designed in the 9/7-type transform based on the structure proposed in [73]**.** In the first to the

fourth lifting steps, the 4D input signal, once it is decomposed into 16 channels, is applied to the spatial dimension *x* as in Equation (5.5).

$$\begin{bmatrix} X_{0000}^{(B)}(\mathbf{z}) \\ X_{0001}^{(B)}(\mathbf{z}) \\ \vdots \\ X_{1111}^{(B)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} R\left[k^{-1}X_{0000}^{(A)}(\mathbf{z})\right] \\ R\left[k^{-1}X_{0001}^{(A)}(\mathbf{z})\right] \\ \vdots \\ R\left[k^{+1}X_{1111}^{(A)}(\mathbf{z})\right] \end{bmatrix}, \qquad (5.9)$$

Then, from the fifth to the 12th lifting steps, the signals are transformed simultaneously in spatial dimensions *y* and *z*, and temporal dimension *t* using the non-separable 3D structure. For instance, the signal in Y_{LHHH} is produced as

$$X_{0111}^{(D)}(\mathbf{z}) = X_{0111}^{(B)}(\mathbf{z}) + R\left[k^{+3}2^{-F}P_{LHHH}^{(D)}(\mathbf{z})\right] \qquad (5.10)$$

for

$$P_{LHHH}^{(D)} = \begin{bmatrix} B_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z}) \\ B_1(\mathbf{z})C_1(\mathbf{z}) \\ B_1(\mathbf{z})D_1(\mathbf{z}) \\ B_1(\mathbf{z}) \\ C_1(\mathbf{z})D_1(\mathbf{z}) \\ C_1(\mathbf{z}) \\ D_1(\mathbf{z}) \end{bmatrix}^T \cdot \begin{bmatrix} X_{0000}^{(B)}(\mathbf{z}) \\ X_{0001}^{(B)}(\mathbf{z}) \\ X_{0010}^{(B)}(\mathbf{z}) \\ X_{0011}^{(B)}(\mathbf{z}) \\ X_{0100}^{(B)}(\mathbf{z}) \\ X_{0101}^{(B)}(\mathbf{z}) \\ X_{0110}^{(B)}(\mathbf{z}) \end{bmatrix} + \begin{bmatrix} B_3(\mathbf{z})C_3(\mathbf{z})D_3(\mathbf{z}) \\ B_3(\mathbf{z})C_3(\mathbf{z}) \\ B_3(\mathbf{z})D_3(\mathbf{z}) \\ B_3(\mathbf{z}) \\ C_3(\mathbf{z})D_3(\mathbf{z}) \\ C_3(\mathbf{z}) \\ D_3(\mathbf{z}) \end{bmatrix}^T \cdot \begin{bmatrix} X_{0000}^{(B)}(\mathbf{z}) \\ X_{0001}^{(B)}(\mathbf{z}) \\ X_{0010}^{(B)}(\mathbf{z}) \\ X_{0011}^{(B)}(\mathbf{z}) \\ X_{0100}^{(B)}(\mathbf{z}) \\ X_{0101}^{(B)}(\mathbf{z}) \\ X_{0110}^{(B)}(\mathbf{z}) \end{bmatrix}, \qquad (5.11)$$

in the 5th lifting step. In this step, a 3D filtering with 3D memory accessing $B_1(\mathbf{z})C_1(\mathbf{z})D_1(\mathbf{z})$ is used. In the 6th lifting step, the calculation of Y_{LHHL}, Y_{LHLH} and Y_{LLHH} :

$$\begin{bmatrix} X_{0110}^{(D)}(\mathbf{z}) \\ X_{0101}^{(D)}(\mathbf{z}) \\ X_{0011}^{(D)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} X_{0110}^{(B)}(\mathbf{z}) + R\left[k^{+1}2^{-F}P_{LHHL}^{(D)}(\mathbf{z})\right] \\ X_{0101}^{(B)}(\mathbf{z}) + R\left[k^{+1}2^{-F}P_{LHLH}^{(D)}(\mathbf{z})\right] \\ X_{0011}^{(B)}(\mathbf{z}) + R\left[k^{+1}2^{-F}P_{LLHH}^{(D)}(\mathbf{z})\right] \end{bmatrix}, \qquad (5.12)$$

for

$$\begin{bmatrix} P_{LHHL}^{(D)\prime}(\mathbf{z}) \\ P_{LHLH}^{(D)\prime}(\mathbf{z}) \\ P_{LLHH}^{(D)\prime}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} B_1(\mathbf{z})C_1(\mathbf{z}) & 0 & B_1(\mathbf{z}) & C_1(\mathbf{z}) & D_2(\mathbf{z}) \\ B_1(\mathbf{z})D_1(\mathbf{z}) & B_1(\mathbf{z}) & 0 & D_1(\mathbf{z}) & C_2(\mathbf{z}) \\ C_1(\mathbf{z})D_1(\mathbf{z}) & C_1(\mathbf{z}) & D_1(\mathbf{z}) & 0 & B_2(\mathbf{z}) \end{bmatrix} \begin{bmatrix} X_{0000}^{(B)}(\mathbf{z}) \\ X_{0001}^{(B)}(\mathbf{z}) \\ X_{0010}^{(B)}(\mathbf{z}) \\ X_{0100}^{(B)}(\mathbf{z}) \\ X_{0111}^{(B)}(\mathbf{z}) \end{bmatrix}, \qquad (5.13)$$

$$\begin{bmatrix} P_{LHHL}^{(D)\prime\prime}(\mathbf{z}) \\ P_{LHLH}^{(D)\prime\prime}(\mathbf{z}) \\ P_{LLHH}^{(D)\prime\prime}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} B_3(\mathbf{z})C_3(\mathbf{z}) & 0 & B_3(\mathbf{z}) & C_3(\mathbf{z}) & D_4(\mathbf{z}) \\ B_3(\mathbf{z})D_3(\mathbf{z}) & B_3(\mathbf{z}) & 0 & D_3(\mathbf{z}) & C_4(\mathbf{z}) \\ C_3(\mathbf{z})D_3(\mathbf{z}) & C_3(\mathbf{z}) & D_3(\mathbf{z}) & 0 & B_4(\mathbf{z}) \end{bmatrix} \begin{bmatrix} X_{0000}^{(B)}(\mathbf{z}) \\ X_{0001}^{(B)}(\mathbf{z}) \\ X_{0010}^{(B)}(\mathbf{z}) \\ X_{0100}^{(B)}(\mathbf{z}) \\ X_{0111}^{(B)}(\mathbf{z}) \end{bmatrix}, \quad (5.14)$$

$$\begin{cases} P_{LHHL}^{(D)}(\mathbf{z}) = P_{LHHL}^{(D)\prime}(\mathbf{z}) + P_{LHHL}^{(D)\prime\prime}(\mathbf{z}) \\ P_{LHLH}^{(D)}(\mathbf{z}) = P_{LHLH}^{(D)\prime}(\mathbf{z}) + P_{LHLH}^{(D)\prime\prime}(\mathbf{z}), \\ P_{LLHH}^{(D)}(\mathbf{z}) = P_{LLHH}^{(D)\prime}(\mathbf{z}) + P_{LLHH}^{(D)\prime\prime}(\mathbf{z}) \end{cases} \quad (5.15)$$

where $R[]$ denotes the rounding operation on a signal value. Similarly, prediction of $X_{1111}, X_{1110}, X_{1101}, X_{1100}, X_{1011}, X_{1010}, X_{1001}, X_{1000}, X_{0100}, X_{0010}, X_{0001}$ and X_{0000} are also independent. The total numbers of lifting steps and rounding operators in the non-separable structure were hence reduced from 16 to 12 and 192 to 96, respectively, comparing with the separable structure in Fig. 32.

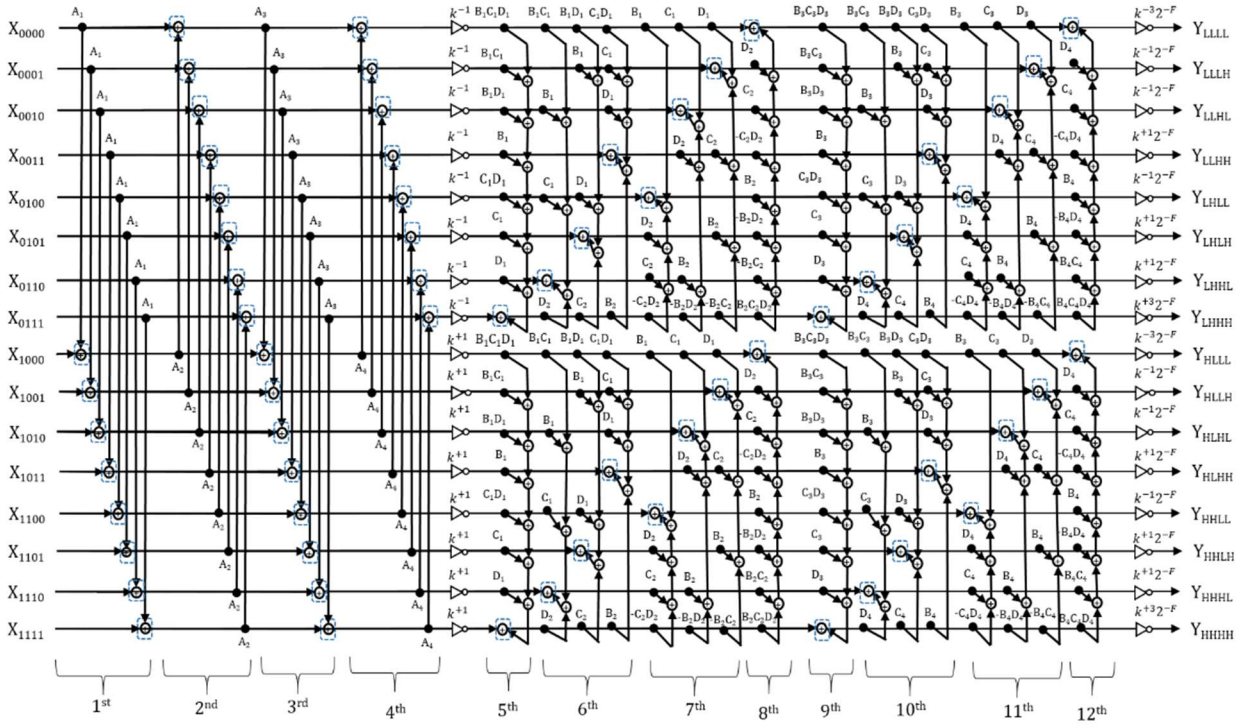


Fig. 33 Non-separable 3D structure for 9/7-type of transform (Existing II)

5.2.3 Non-Separable 4D Integer Wavelet Transform for 9/7 filter (Existing III)

Fig. 34 shows the non-separable 4D structure for 9/7-type transform. The non-separable quadruple lifting structure is the combination of two double lifting structure as in Fig. 21 in Section 4.3.1. The different between those structures are the addition of another two filters as the 1D DWT 9/7-type transform of Fig. 16 in Section 3.3.1. The derivation of this structure is

also the same as in Section 4.3.2. The total numbers of lifting steps and rounding operators in the non-separable structure were hence reduced from 16 to 10 and 192 to 48, respectively, comparing with the separable structure in Fig. 32 and non-separable 3D structure in Fig. 33. However, the quality of the decoded image was degraded by the rounding noise inside the transform in its integer implementation. The proposed methods that solved this problem is explained in Section 5.3.

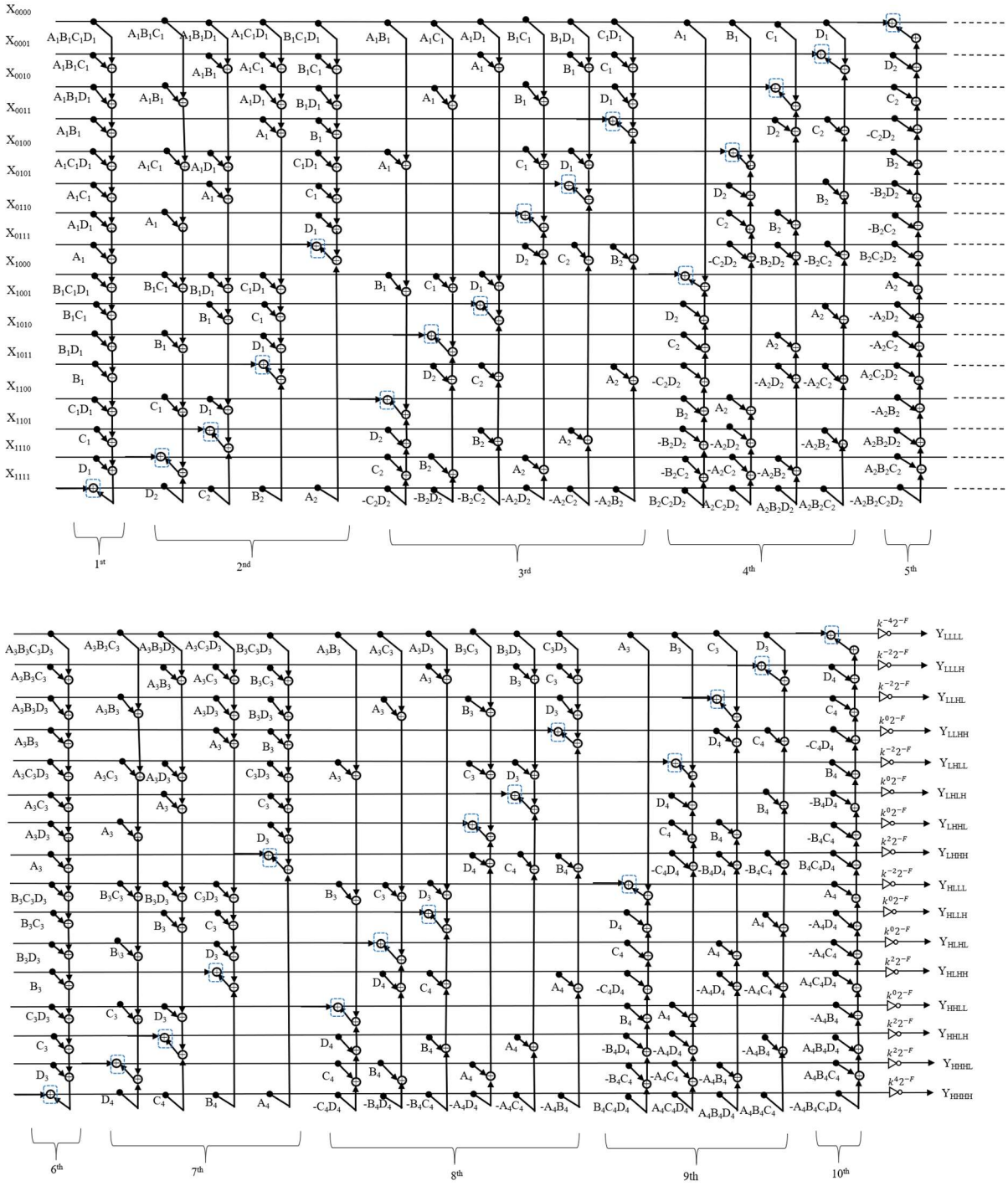


Fig. 34 Non-separable 4D structure for quadruple lifting (Existing III)

5.3 Proposed Methods

5.3.1 Designing a new lifting structure for 4D 9/7 filter

The non-separable structure is believed can help reduce rounding noise inside 4D quadruple DWT. The filter can be rearranged and recombined from to form the non-separable structure. There are 24 possible filter combinations for 2D structure, 40320 for 3D structure and 2.092×10^{13} for 4D structures. Therefore, it is necessary to find the best possible combinations for this filter in order to design a non-separable structure with the lowest rounding noise.

There are several rules needed to abide in order to combine and rearrange the filter for the JPEG 2000 DWT. Below are the rules for the rearrangement of the filter coefficients:

1. Permutations (A_i, B_i),(B_i, C_i), (A_i, C_i), (A_i, D_i), (B_i, D_i) and (C_i, D_i) for $i\in\{1,2,3,4\}$ are excluded because they will have the same rounding noise.
   For example, changing the $A_1A_2A_3A_4B_1B_2B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$ to $B_1B_2A_3A_4A_1A_2B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$ will yield the same amount of rounding noise.
2. Permutations (A_i, A_j), (B_i, B_j), $(C_i, C_j)$and (D_i, D_j) for $i<j$ are excluded because they are different from the 9/7-type DWT.
   For example, turning the $A_1A_2A_3A_4B_1B_2B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$ into $A_2A_1A_3A_4B_1B_2B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$ is prohibited.
3. A_i and A_j are coupled as A_{ij} for $ij = \{12\} or \{34\}$. Similarly, $B_i \; and \; B_j$, --> B_{ij}, $C_i \; and \; C_j$, --> C_{ij}, $D_i \; and \; D_j$, --> D_{ij}.
   For example, $A_1A_2A_3A_4B_1B_2B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$ will be coupled as $A_{12}\ A_{34}\ B_{12}\ B_{34}\ C_{12}\ C_{34}\ D_{12}\ D_{34}$.

By following the above rules of filter combinations and arrangements, we can form the non-separable 2D,

$$A_1A_2(A_3A_4B_1B_2)_{2D}B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$$

non-separable 3D,

$$A_1A_2(A_3A_4B_1B_2C_1C_2)_{3D}B_3B_4C_3C_4D_1D_2D_3D_4$$

and non-separable 4D,

$$A_1A_2(A_3A_4B_1B_2C_1C_2D_1D_2)_{4D}B_3B_4C_3C_4D_3D_4$$

Thus, the possible combinations to lower the rounding noise of 4D DWT for 9/7-type filter can be lowered down. As a result, in this proposal, we tried to combine and rearrange the filter to form the 2D and 3D non-separable for this 4D structure. Then, we tried to maintain the original filter arrangement by combining them with only using the 2D non-separable structure. The details of both structures are explained in Section 5.3.2 and Section 5.3.3. From above structures, we experimentally investigated which structures can reduce the rounding noise inside the 4D quadruple lifting of DWT as shown in Section 5.4.

5.3.2 Non-separable 2D and 3D structure for 4D 9/7 filter (Proposed I)

To solve the problems of higher rounding noise and degraded quality of lossy coding problems, the non-separable structure, which combines both 2D and 3D structures, is proposed. The total number of rounding operators is thus further reduced from 96 to 72. Fig. 35 illustrates the structure of the Proposed I method. Once the 4D input signal is decomposed into 16 channels, the first, second, and third lifting steps are cascaded by using the non-separable 2D structure, followed by the non-separable 3D structure and the separable 1D structure. For example, the first lifting step is expressed as

$$X^{(1)}_{11c_1c_2}(\mathbf{z}) = X_{11c_1c_2}(\mathbf{z}) + R\big[A_1B_1X_{00c_1c_2}(\mathbf{z}) + A_1X_{01c_1c_2}(\mathbf{z}) + B_1X_{10c_1c_2}(\mathbf{z})\big], \quad \text{where } c_1, c_2 \in \{0,1\}, \tag{5.16}$$

Even though this proposed method has fewer rounding operators, the rounding noise in it is still higher than in Existing I method, the separable 4D structure. Therefore, it is necessary to maintain the original structure so that rounding noise in the transform is lower.

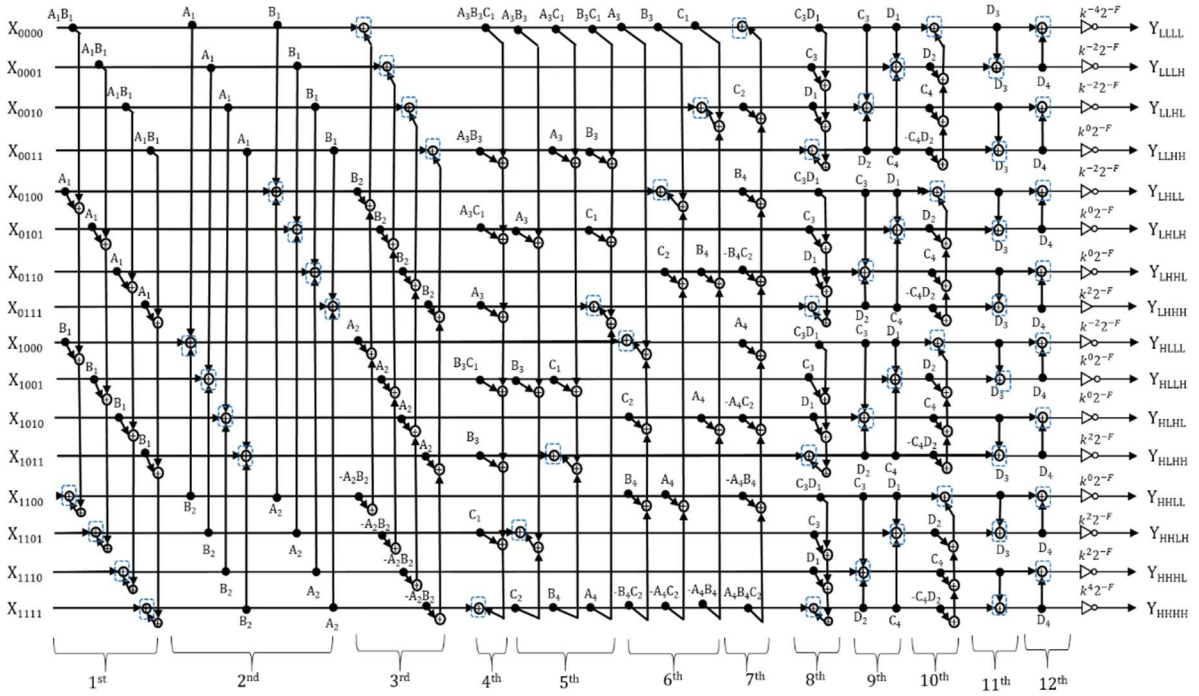


Fig. 35 Non-separable 2D combined with 3D structure for 9/7-type of transform (Proposed I)

5.3.3 Non-separable 2D structure for 4D 9/7 filter (Proposed II)

To lower the rounding noise inside the transform, the non-separable 2D structure for 4D input signals is proposed. Fig. 36 shows the Proposed II method. Unlike the Existing I, Existing II, Existing III and Proposed I methods, it consists of only a non-separable 2D

structure. The original order of lifting steps for each dimension is also maintained in this structure. The total number of lifting steps, however, is larger than in Existing II and Existing III but smaller than in Existing I. The total number of rounding operators is smaller than in Existing I, and the rounding noise is lower, as confirmed in Section 5.4.

The first and second lifting steps involve 1D structures expressed as

$$\begin{cases} X^{(1)}_{1c_1c_2c_3}(\mathbf{z}) = X_{1c_1c_2c_3}(\mathbf{z}) + R[A_1 X_{0c_1c_2c_3}(\mathbf{z})] \\ X^{(1)}_{0c_1c_2c_3}(\mathbf{z}) = X_{0c_1c_2c_3}(\mathbf{z}) + R[A_2 X_{1c_2c_3}(\mathbf{z})] \end{cases}, \quad where\ c_1, c_2, c_3 \in \{0,1\}, \qquad (5.17)$$

The third, fourth, and fifth lifting steps consist of non-separable 2D structures, and the same goes for the 5th to the 11th steps. Finally, the 12th and 13th lifting steps involve the separable 1D structure. For example, the third lifting step is expressed as

$$X^{(2)}_{11\ {}_1c_2}(\mathbf{z}) = X^{(1)}_{11c_1c_2}(\mathbf{z}) + R\Big[A_3B_1X^{(1)}_{00c_1c_2}(\mathbf{z}) + A_3X^{(1)}_{01\ {}_1c_2}(\mathbf{z}) + B_1X^{(1)}_{10c_1c_2}(\mathbf{z})\Big], \text{ where } c_1, c_2 \in \{0,1\}, \qquad (5.18)$$

Thus, Proposed II is a combination of non-separable 2D structures and separable 1D structures.

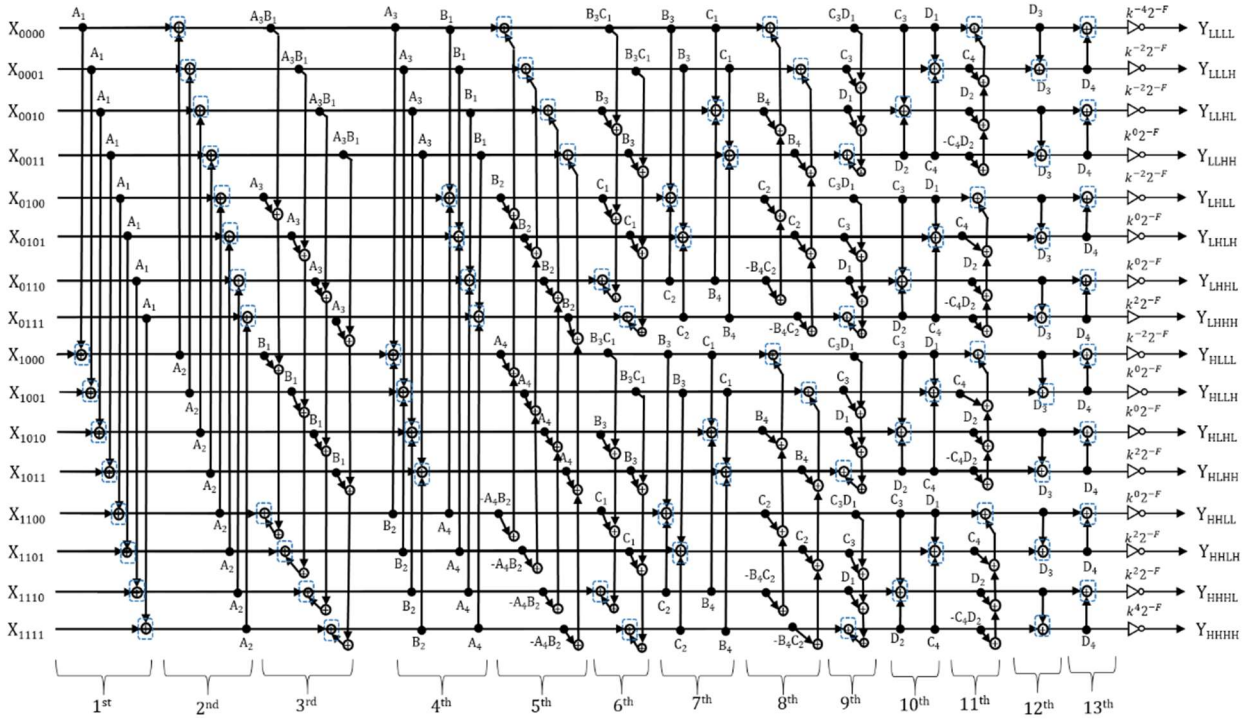


Fig. 36 Non-separable 2D structure for 9/7-type of transform (Proposed II)

5.3.4 Comparison of the structures

Table 4 compares the four structures: separable 4D, non-separable 3D, non-separable 4D, non-separable 2D and 3D, and non-separable 2D. As summarized in the table, the total number of lifting steps in the proposed non-separable 2D structure increases from 10 to 13

comparing with the existing non-separable 4D structure. However, the total number of rounding operators remains the same. Note that both are still smaller in number than the existing separable 4D structure. A different structure obtained by combining the 2D and 3D structures is first proposed to further reduce the number of rounding operators; but to reduce rounding noise, it is necessary to maintain the original, separable 4D structure. Therefore, the non-separable 2D structure is proposed based on this original structure.

As shown in Fig. 4, the Existing I method is composed of lifting steps A_1, A_2, …., D_4, and is expressed as

Separable 4D $A_1A_2A_3A_4B_1B_2B_3B_4C_1C_2C_3C_4D_1D_2D_3D_4$ (5.19)

In Existing I, the order of lifting steps for spatial dimension *x* remains the same, but those of the steps for dimensions *y* and *z*, and temporal dimension *t,* change as

Separable 3D' $B_1B_2C_1C_2D_1D_2B_3B_4C_3C_4D_3D_4$ (5.20)

Part of this is implemented in the non-separable 3D structure (Existing II).

Non-separable 3D $A_1A_2A_3A_4(B_1B_2C_1C_2D_1D_2)_{3D}(B_3B_4C_3C_4D_3D_4)_{3D}$ (5.21)

The non-separable 4D structure (Existing III) is expressed as

Non-separable 4D $(A_1A_2B_1B_2C_1C_2D_1D_2)_{4D}(A_3A_4B_3B_4C_3C_4D_3D_4)_{4D}$ (5.22)

Unlike Existing II, the Proposed I structure is expressed as

Non-separable 2D & 3D $(A_1A_2B_1B_2)_{2D}(A_3A_4B_3B_4C_1C_2)_{3D}(C_3C_4D_1D_2)_{2D}D_3D_4$ (5.23)

To maintain the original structure of Existing I, Proposed II is expressed as

Non-separable 2D $A_1A_2(A_3A_4B_1B_2)_{2D}(B_3B_4C_1C_2)_{2D}(C_3C_4D_1D_2)_{2D}D_3D_4$ (5.24)

Note that the brackets refer to the non-separable structure.

Table 4 Comparison of the methods

Structure	**Lifting steps**	**Rounding Operators**	**Memory Accessing**
Separable 4D (Existing I)	16	192	1D
Non-separable 3D (Existing II)	12	96	1D & 3D
Non-separable 4D (Existing III)	10	48	4D
Non-separable 2D & 3D (Proposed I)	12	72	1D, 2D & 3D
Non-separable 2D (Proposed II)	13	96	1D & 2D

5.4 Experimental Results

In the following experiments, six types of data were used to evaluate the rounding noise and coding performance, as shown in Table 5 and Fig. 37.

Table 5 Type of data used in the experiments

Type of data	**Size of data in ($\mathbf{x} \times \mathbf{y} \times \mathbf{z} \times \mathbf{t}$) sequence**	**Bit depth**	**Referred here as**
4D functional Magnetic Resonance Image	$56 \times 88 \times 32 \times 16$	16	fMRI (I)
4D Computed Tomography Image	$256 \times 256 \times 16 \times 16$	12	CT
4D functional Magnetic Resonance Image	$40 \times 64 \times 64 \times 16$	12	fMRI (II)
4D Magnetic Resonance Image	$52 \times 224 \times 64 \times 16$	8	MRI
4D Auto-regressive Model	$256 \times 256 \times 32 \times 16$	8	AR
4D Ultrasound Image	$532 \times 416 \times 16 \times 16$	8	US

Note that the MRI, fMRI(II), and US data were retrieved from [68], [69], and [76], respectively. MRI data represent highly correlated data at consecutive time points with limited motion in the temporal dimension *t* and structural changes in the spatial dimension *z*.

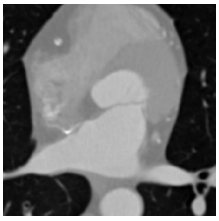
Each image is normalized to the range [0, 255] for display purposes as shown in Fig. 33 (a)–(f). In this paper, the variance in noise in frequency domain, and coding performance in lossless and lossy coding modes were investigated.

The 4D auto-regressive (AR) model used in our experiments can be expressed as

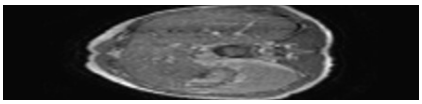
$$\begin{cases} x^{(1)}(n_1,n_2,n_3,n_4) = x(n_1,n_2,n_3,n_4) + \rho \cdot x^{(1)}(n_1-1,n_2,n_3,n_4) \\ x^{(2)}(n_1,n_2,n_3,n_4) = x^{(1)}(n_1,n_2,n_3,n_4) + \rho \cdot x^{(2)}(n_1,n_2-1,n_3,n_4) \\ x^{(3)}(n_1,n_2,n_3,n_4) = x^{(2)}(n_1,n_2,n_3,n_4) + \rho \cdot x^{(3)}(n_1,n_2,n_3-1,n_4) \\ x^{(4)}(n_1,n_2,n_3,n_4) = x^{(3)}(n_1,n_2,n_3,n_4) + \rho \cdot x^{(4)}(n_1,n_2,n_3,n_4-1) \end{cases}, \quad (5.25)$$

Note that ρ was set to 0.9 in the experiments in this paper, as the typical values of ρ for natural images are between 0.9 and 0.98 [77].

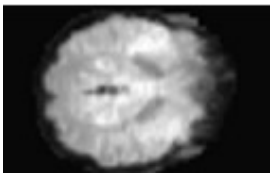
(a) fMRI (I)



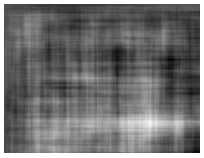
(b) CT



(c) MRI



(d)fMRI (II)

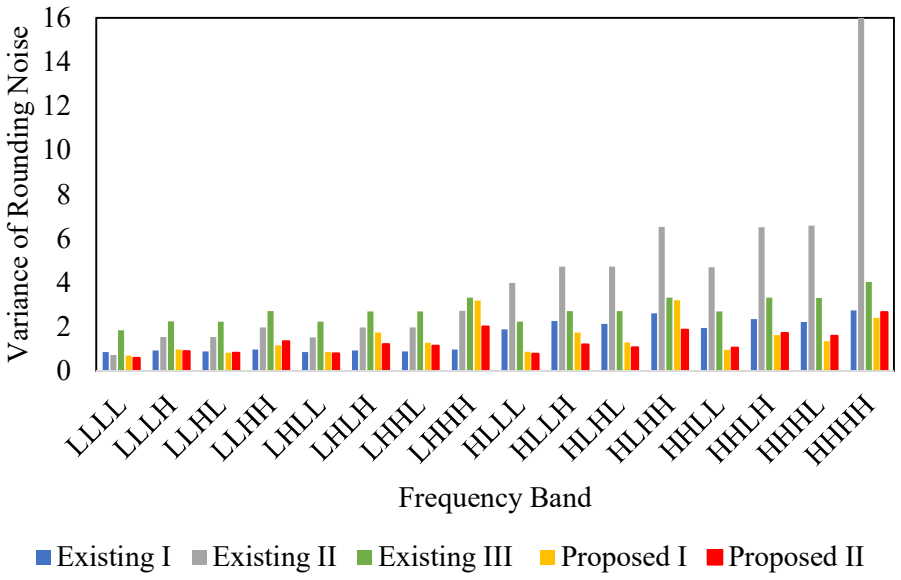


(e) AR

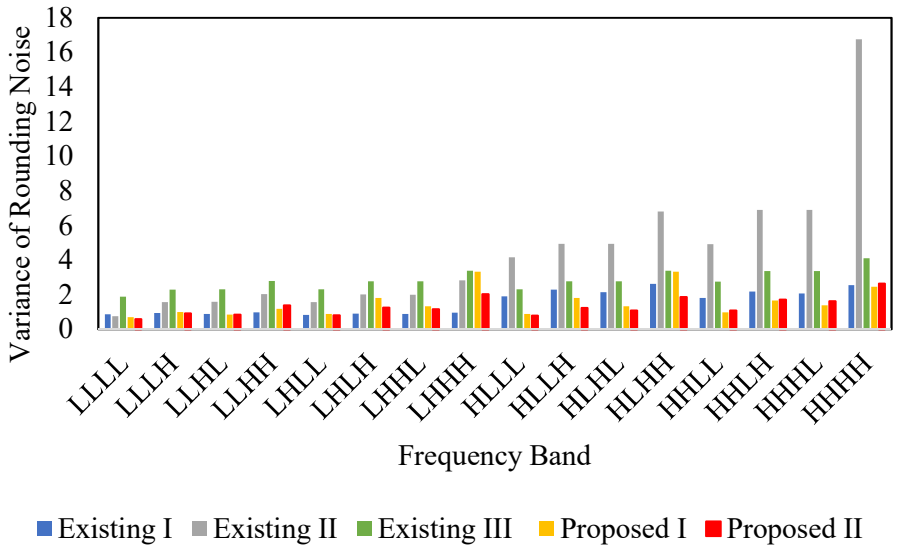
(f) US

Fig. 37 Tested data

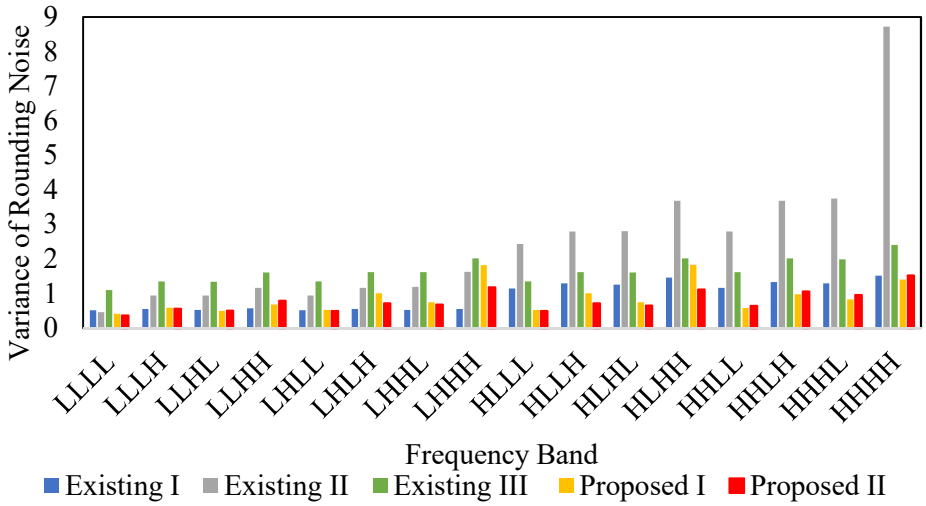
5.4.1 Evaluation on Rounding Noise



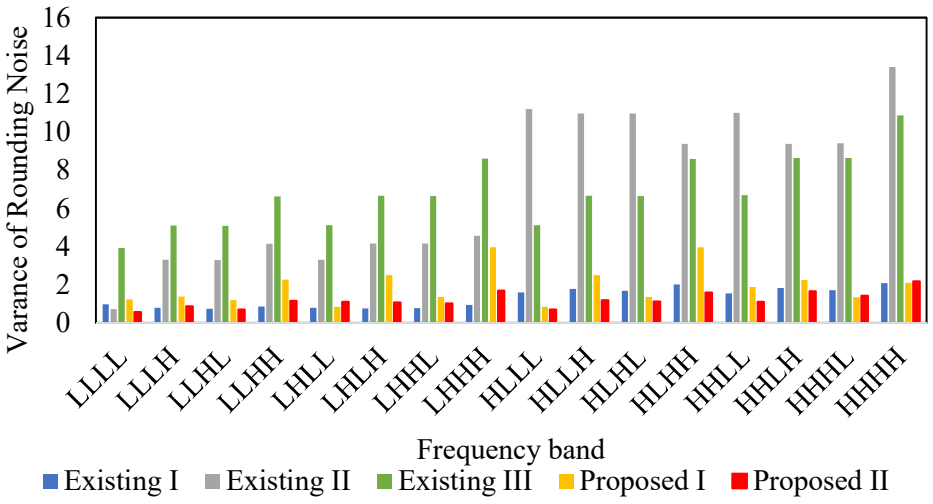
(a) fMRI(I)



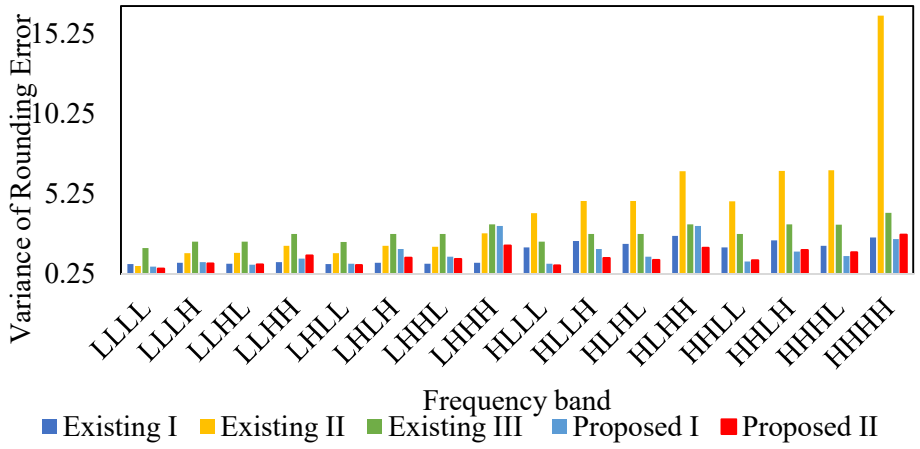
(b) CT



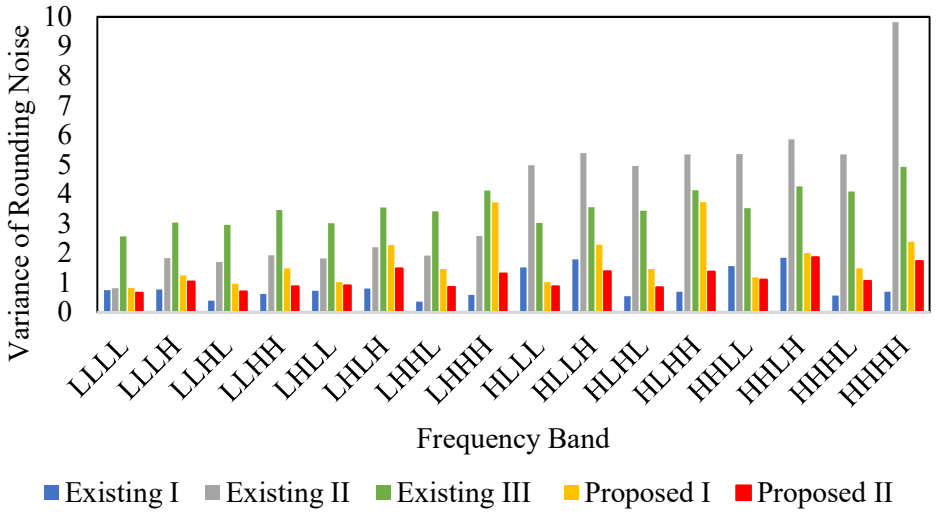
(c) fMRI(II)



(d) MRI



(e) AR



(f) US

Fig. 38 Variance of rounding noise for all data in each frequency band

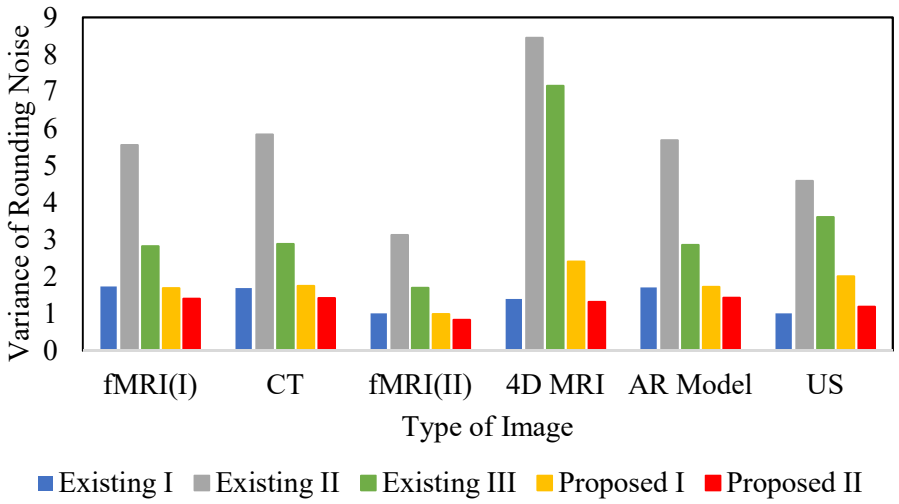


Fig. 39 Average variance of rounding noise in each frequency band

Figs. 38 (a)–(f) indicate the variance of the rounding error in each frequency band for fMRI(I), CT, fMRI(II), MRI, AR, and US data respectively. The average variance of rounding errors for all data was reduced to 19.11%, 16.08%, 16.81%, 5.63%, and 16.09%, respectively, except for US data, which was increased by 15.47%, between Existing I structure and Proposed II structure, as shown in Fig. 39. Although the rounding noise in the LLLL frequency band for Proposed II structure for the US data was lower than that in Existing I, the coding performance on the US data for Proposed II method improved as shown in Fig. 38 (f). The HHHH frequency band of the Existing II structure was the highest and resulted in a degradation in coding performance. The energy of the frequency band should be compacted to a low-frequency band signal to reduce entropy in the compressed image and enhance coding performance. The higher the variance of rounding noise in the low-frequency band signal, the lower the entropy of the compressed image. However, all structures had the highest compacted energy in the high-frequency band signals. Therefore, we conclude that the lowest variance in rounding noise among all frequency band signals yields the best coding performance.

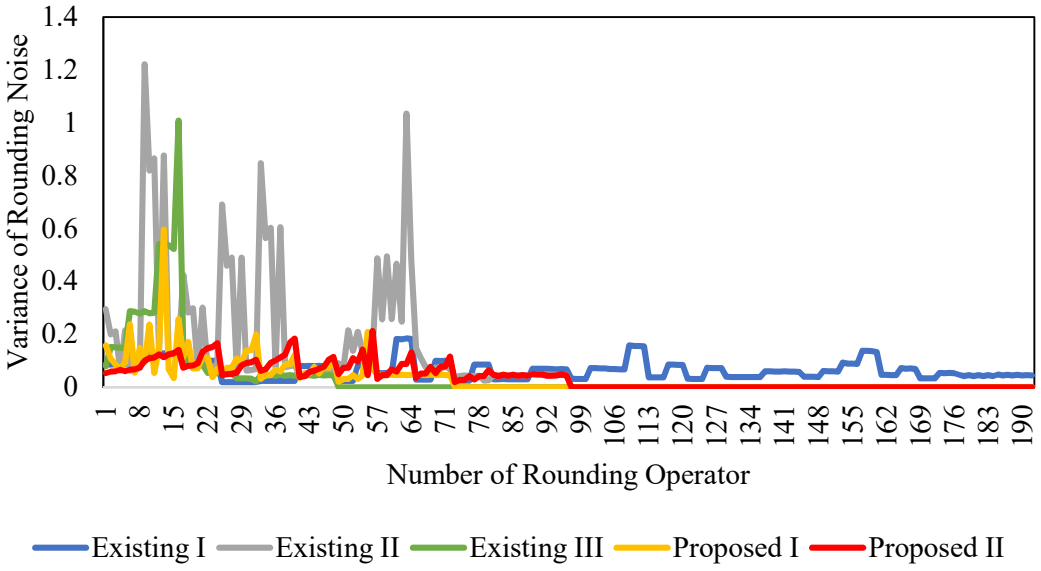


Fig. 40 Effect of one rounding operator for fMRI(II) data

The reason for why the total number of rounding operators did not influence total rounding noise is investigated in Fig. 40, which shows the variance of rounding noise in the frequency domain. In an experiment, only one rounding operator in the forward transform was activated. The horizontal axis denotes the activated rounding operator, numbered from top to bottom, from left to right in Figs. 32 to 36. As summarized in Table 4, the Existing I, Existing II, Existing II, Proposed I, and Proposed II structures had 192, 96, 48, 72, and 96 rounding

operations, respectively. The total rounding noise in Proposed II decreased even though the number of rounding operations, the source of the noise, remained the same as in the Existing II structure. This is because of the relatively large variance in error in the structure. The variance in rounding noise was the highest in the 63rd rounding operator at 0.1844, the 9th rounding operator at 1.2217, the 16th rounding operator at 1.0084, the 14th rounding operator at 0.5964 and the 56th rounding operator at 0.2127 for Existing I, Existing II, Existing III, Proposed I, and Proposed II, respectively. However, the average rounding noise in Proposed II relative to Existing I decreased by almost half from 0.0646 to 0.0382. It can be concluded that the more frequent and the greater the extent to which noise is amplified in the lifting structure, the higher the rounding noise inside it. As the rounding noise inside the transform influences coding performance, it is investigated in Section 5.4.2.

5.4.2 Evaluation on Coding Performance

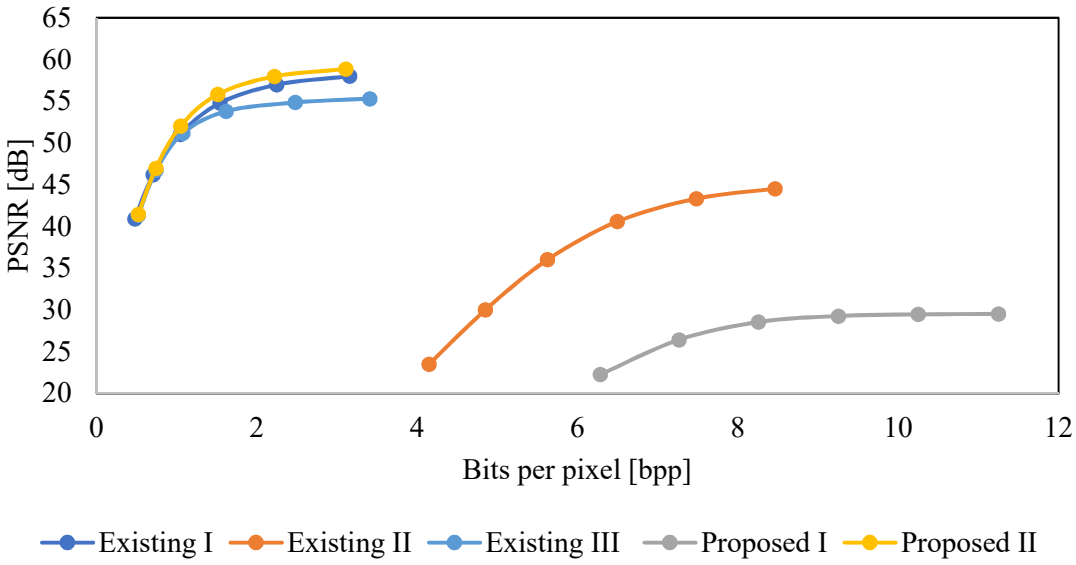
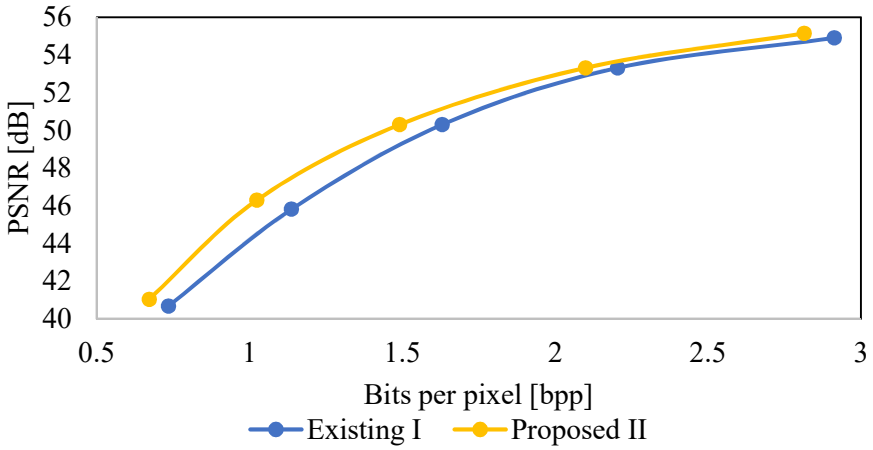


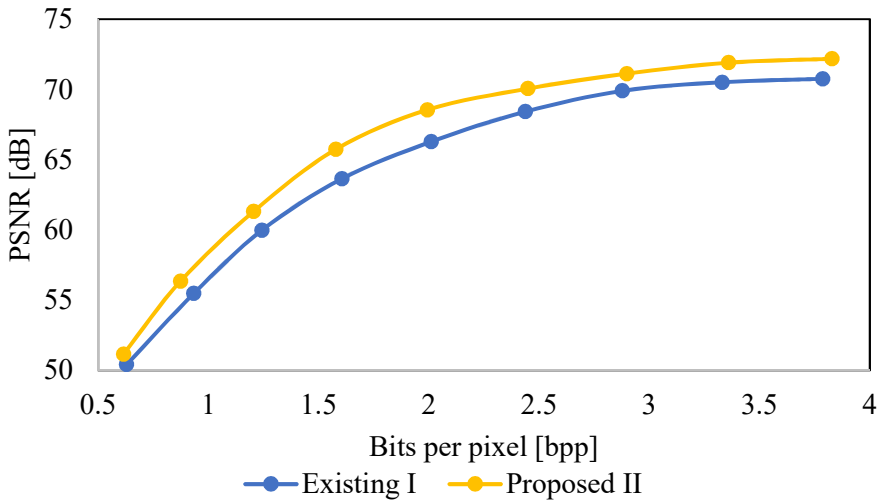
Fig. 41 Coding performance in lossy mode for CT image

Fig. 41 illustrates the rate distortion curve, which compares the performance of the methods in the lossy coding mode for CT images. The horizontal and vertical axes represent the entropy rate measured in bits per pixel [bpp] and the peak signal-to-noise ratio (PSNR) of the reconstructed signal, respectively. It shows that the Proposed II structure outperformed the others at the same bitrate. The quality of the reconstructed signal in Proposed II increased by 0.18 dB, 16.11 dB, 3.95 dB and 33.00 dB over Existing I, Existing II, Existing III and Proposed I, respectively. As the quality of the reconstructed image deteriorated considerably if it was

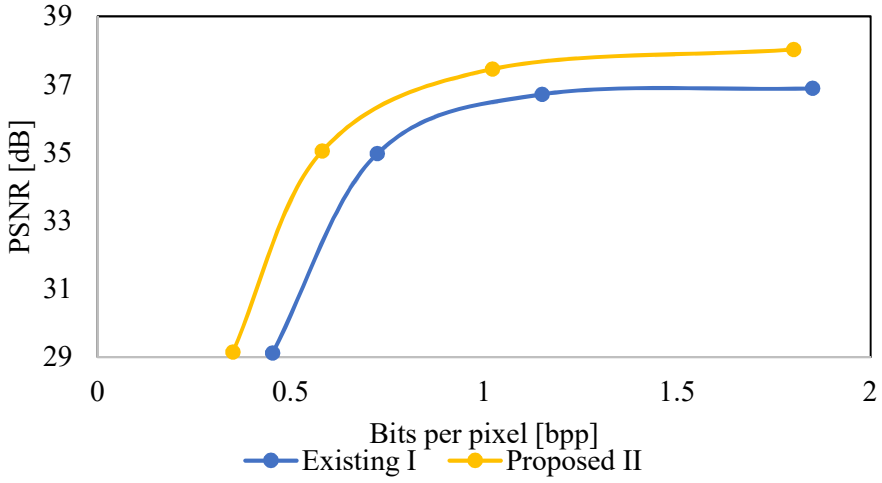
transformed using Existing II and Proposed I, Figs. 42(a)–(e) show coding performance results when images were transformed using Existing I and Proposed II only.



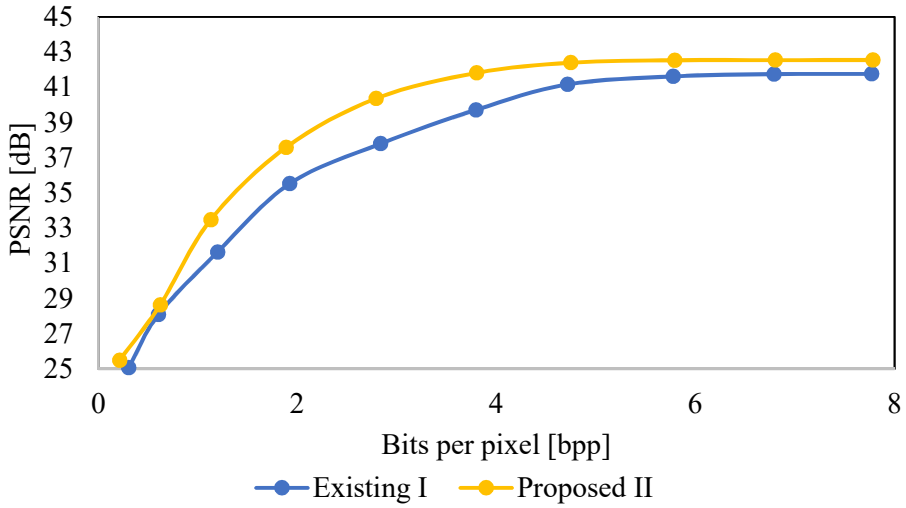
(a) fMRI(I)



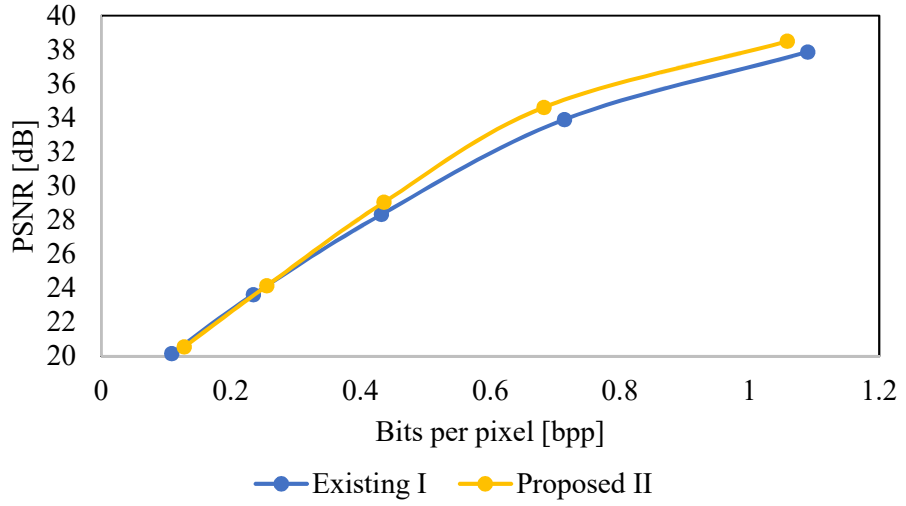
(b) fMRI(II)



(c) MRI



(d) AR



(e) US

Fig. 42 Coding performance in lossy mode for fMRI(I), fMRI(II), MRI, AR and US data

Figs. 42 (a)–(e) show the coding performance of Existing I and Proposed II on fMRI(I), fMRI(II), MRI, AR, and US data, respectively. Under the same bitrate, the quality of the reconstructed signal for Proposed II increased by 0.40 dB, 2.10 dB, 0.27 dB, 2.57 dB, and 0.72 dB for fMRI(I), fMRI(II), MRI, AR, and US data, respectively. As both quantization noise and rounding noise were present in the lossy compression system, even though the former for all structures was the same, rounding noise in the Proposed II structure was the lowest. Thus, the coding performance of this structure was the best. In conclusion, Proposed II structure outperformed all other structures in the experiments.

5.5 Conclusion of this Chapter

This chapter proposed a non-separable 2D structure for 4D input signals in lieu of non-separable 3D structures. The total number of rounding operators was reduced by half compared with the prevalent separable 4D structure. However, rounding noise owing to the integer implementation of signal values inside the transform increased in the non-separable 3D structure, due to a change in its original lifting scheme. Therefore, by maintaining the original scheme, the proposed non-separable 2D structure reduced total rounding noise in it and enhanced the quality of the reconstructed signal in lossy coding. Furthermore, the number of lifting steps, a reduction in which reduces the latency of the overall transform, were also reduced by 18.75% between the proposed non-separable 2D structure and the conventional separable 4D structure on the quadruple 4D integer WT. Our integer WT has the advantage of compatibility with the conventional integer WT. It also enhances compression performance on 4D images, such as medical images.

Chapter 6 Further Proposals on the Non-Separable Integer Wavelet Transform

6.1 Region-of-Interest (ROI) Coding for 4D Image Compression

(This section is presented in Asia-Pacific Signal Processing Association Annual Summit and Conference (APSIPA ASC), Dec. 2017)

6.1.1 Proposal Motivation

The integer wavelet transform has been extended to 4D transforms for the 4D image compression [5, 55]. However, the transform is composed of separable structure based on JPEG 2000 where the data is transformed separately in spatial and temporal dimensions.

A new class of non-separable 2D structure has been reported in [45, 44]. The transfer function can be expressed as a product of 1D function. The transform based on this structure is compatible to the separable transform. The non-separable structure is not a cascade of 1D signal processing in 1D structure, but it requires the multidimensional memory accessing. The non-separable three-dimensional (3D) structure is also proposed in [62]. This structure has an advantage that it will decrease the total number of lifting steps and the rounding operations inside it.

Unlike those previous studies on the 2D and 3D case, this proposal proposes a non-separable double lifting structure of 5/3-type filter for 4D integer WT with reduced amount of the rounding noise. It contributes in increasing the coding performance of 4D data compression system. But, as the lossless coding only includes rounding noise in the system compared to the lossy coding which has both quantization and rounding noise, the coding performance of the double non-separable lifting integer WT is found to be just slightly increased. The performance of lossy coding which use the non-separable quadruple lifting structure of 9/7-type filter for 4D integer WT in [74] is increased too, but, as it is important to preserve the information in medical image, the lossy coding is not preferable to compress this type of data.

Therefore, a region of interest (ROI) coding is introduced to utilize both lossy and lossless coding of the non-separable lifting structure 4D integer WT to compress the 4D image efficiently by preserving the important part of the image only. Several studies of ROI coding

for 2D image are done in [78] and 3D image in [79]. To the best of our knowledge, this proposal is the first to propose the ROI coding for 4D image. The variance of the noise in frequency domain and the coding performance of a 4D image are clarified in this proposal.

6.1.2 Existing Methods

6.1.2.1 Separable 4D Integer Wavelet Transform without ROI

The separable 4D structure used as existing method in this proposal is the same as in Section 4.2.1. This structure has 8 lifting steps and 64 rounding operators, which will increase the rounding noise inside it. Therefore, a non-separable double lifting structure is proposed to overcome this problem. The increasing of rounding noise inside the transform will affect the coding performance for 4D image. Please refer to Section 4.2.1 for detail explanation regarding this structure.

6.1.2.2 Non-Separable 4D Integer Wavelet Transform without ROI

The non-separable 4D structure used in this section is the same with the structure in Section 4.3.1. This structure has 5 liftings steps with 16 rounding operators, which will result on reduction of rounding noise inside the transform. Further explanation on this method can be referred in Section 4.3.1.

6.1.2.3 Separable 4D Integer Wavelet Transform with ROI

The separable 4D structure implemented in this Section is according to Fig. 43, where the non-separable 4D structure is replaced with the separable 4D structure for its transform method.

6.1.3 Proposed Methods

6.1.3.1 Non-Separable 4D Integer Wavelet Transform with ROI

In order to further improve compression performance of four-dimensional image, the ROI based on non-separable double lifting 4D integer WT is proposed as shown in Fig. 43. First, the ROI of image is predetermined manually, and the image is split into the ROI part and the non-ROI part. Then, the ROI part will be compressed in lossless mode to preserve the part's essential bits of information, whereas the non-ROI part will be compressed in lossy mode as it is assumed that this part's information is permitted to be reconstructed with accepted perception loss. Then, the image is reconstructed by combining both the reconstructed ROI and non-ROI part as shown in the equation (6.1).

$$I' = I'_1 + I'_2, \qquad (6.1)$$

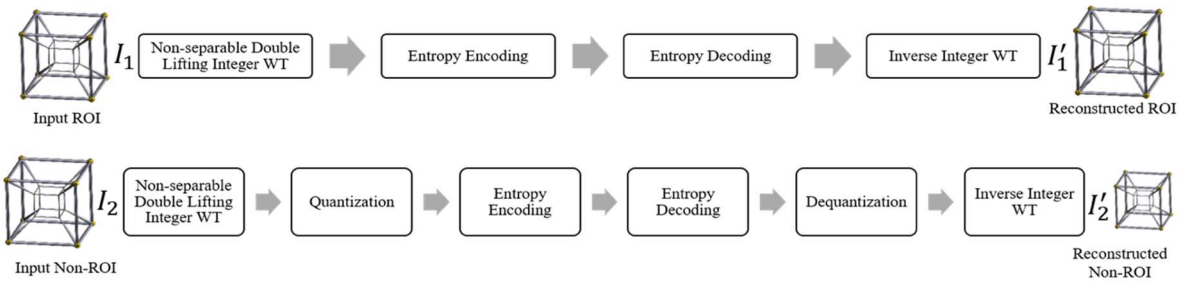
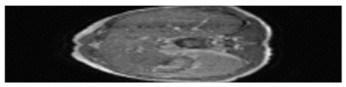


Fig. 43 ROI coding for non-separable double lifting structure of integer WT (Proposed)

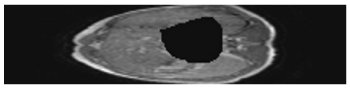
6.1.4 Experimental Results

In the following experiments, a set of grayscales 4D MRI data with 8-bit depth provided in [68] was tested as shown in Fig. 44. The size of the image is 50×224×16×16 pixels, which is expressed in x × y × z × t sequence. Note that each image was normalized to the range [0,255] for display purpose in Fig. 44. In this proposal, the variance of the noise in frequency domain and the coding performance of the image was investigated.



(a) Tested image

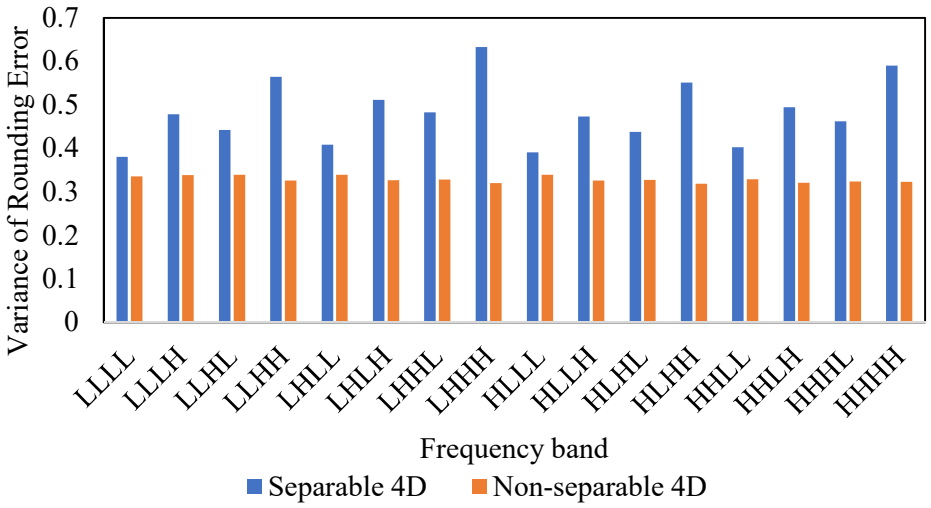
(b) ROI part of the image



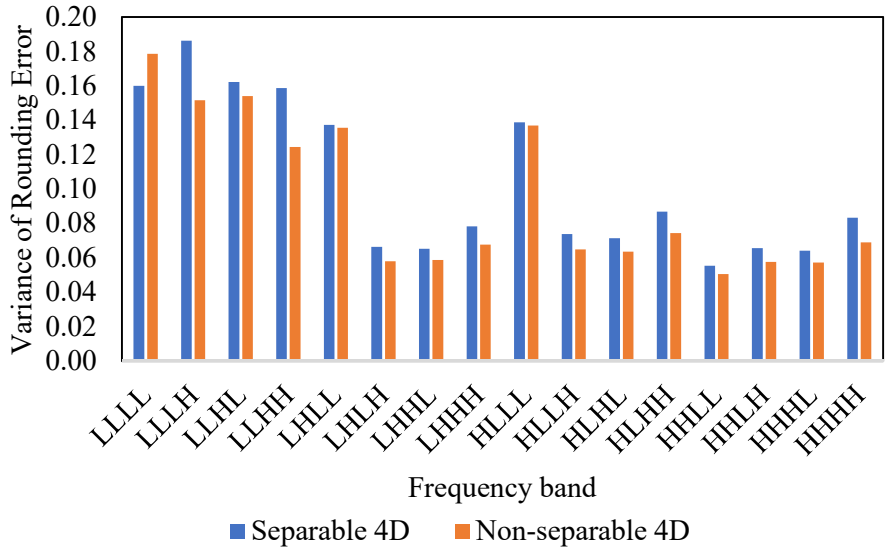
(c) Non-ROI part of the image

Fig. 44 Tested data

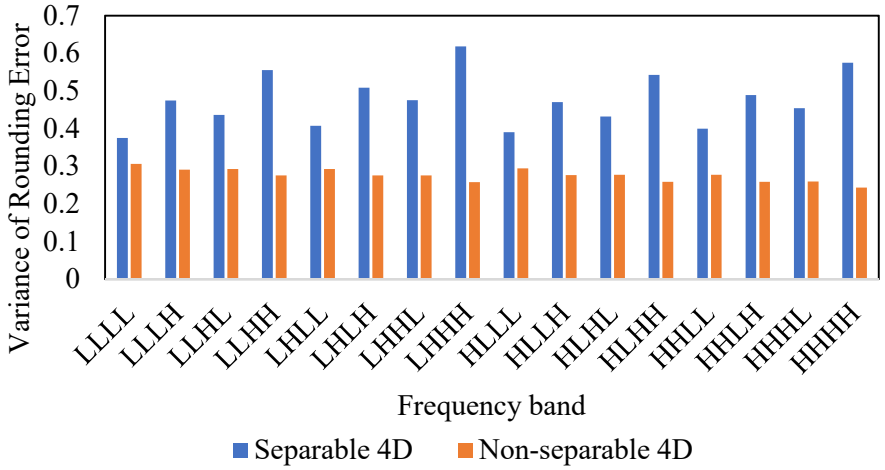
6.1.4.1 Evaluation on Rounding Noise



(a) Tested image



(b) ROI part of the image



(c) Non-ROI part of the image

Fig. 45 Variance of rounding noise in each frequency band

Since fewer rounding operators does not always imply fewer rounding errors in total [45], total amount of rounding errors contained in the output frequency band signals was investigated. The difference between output from wavelet transform without rounding and that with rounding is defined as the error. Fig. 45 (a) - (c) indicate variance of the rounding error in each frequency bands for tested image, the ROI part of the image and the non-ROI part of the image, respectively. The average amount of rounding errors for the tested image, the ROI part of the image, and the non-ROI part of the image were reduced to 32.85[%], 9.11[%] and 43.24[%] respectively. It was found that variance of rounding noise in the separable 4D structure is distributed randomly in all type of data. In most cases, variance of the rounding error in 'HHHH'

band of non-separable was the lowest. On the contrary, variance of rounding error in 'LLLL' band of non-separable was the highest. By using the non-separable 4D integer WT structure, the rounding noise is reduced, and the energy is further compacted to the low frequency band signal, thus reducing the entropy rate of the compressed image, which will eventually increase its coding performance.

6.1.4.2 Evaluation on Coding Performance

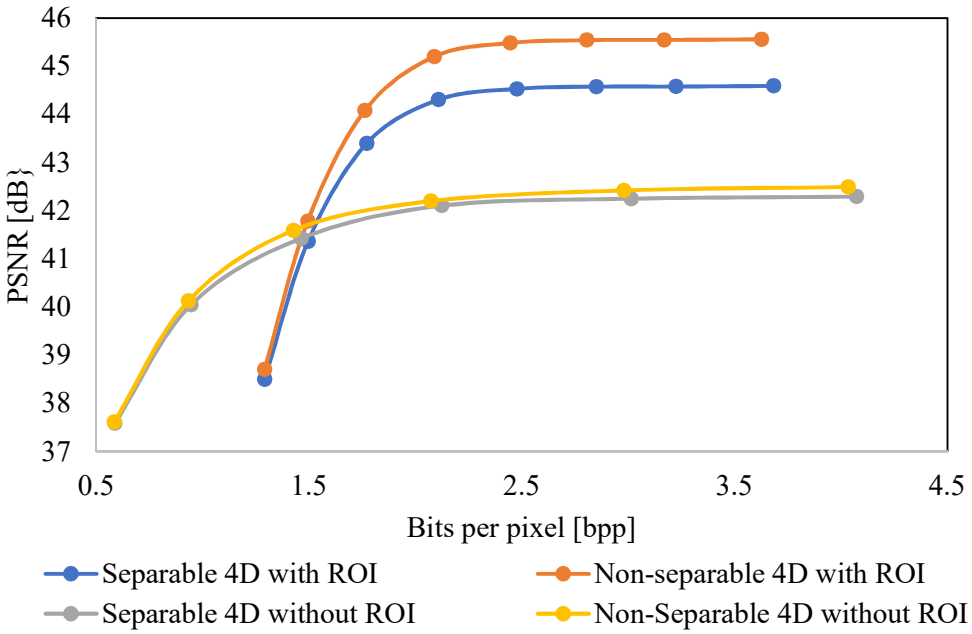


Fig. 46 Coding performance for the 4D MRI Image

Fig. 46 indicates the rate-distortion curve of separable and non-separable structure of 4D WT with and without ROI tested in 4D MRI data. The horizontal axis and the vertical axis represent the entropy rate measured in bits per pixel [bpp] and peak signal noise-to-ratio (PSNR) of the reconstructed signal which represents the quality of reconstructed image, respectively. The transformed data is encoded by using the Embedded Zerotree Wavelet based on Intraband Partitioning (EZW-IP) [35] encoder developed in 3D, where each temporal subband are independently coded. It was observed that non-separable 4D with ROI performed best compared to others. Under the same bit rate, the PSNR was increased by 3.30[dB], 3.13[dB] and 0.97[dB] when the non-separable 4D with ROI is compared with the separable 4D without ROI, non-separable 4D without ROI and separable 4D with ROI. Compared to the lossless coding, the non-separable 4D performed better in lossy coding. This is because, there are both quantization noise and rounding noise inside the compression system. The quantization noise for the separable and the non-separable 4D is the same, but, the rounding

noise in the non-separable is lower than in the separable, thus, increasing the coding performance in lossy mode. Unlike the lossless coding mode, there is only rounding noise inside the system which will only give significant influence on the difference of performance. Therefore, by implementing both lossy and lossless coding in non-separable with ROI, a high coding performance can be achieved.

6.1.5 Conclusion of this Section

A non-separable 4D integer WT compatible with the existing separable 4D double lifting integer WT is proposed. Non-separable multidimensional lifting structure is introduced to decrease the rounding operators by 75[%] and further reducing the rounding noise by approximately 40[%]. It is beneficial to implement our integer WT with ROI which has compatibility with the conventional integer WT that has been broadly used as an international standard nowadays, since the important information can be well preserved while reducing the rounding noise and increasing its coding performance.

6.2 Adaptive Directional Lifting for 3D Image Compression

(This section is published in IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E99-A, No. 5, pp. 892-899)

6.2.1 Proposal Motivation

The existing DWT lifting scheme can only be applied in horizontal and vertical directions which is not suitable to be used in images with lots of directional high frequency components, such as edges and lines. The existing DWT will further refer as non-adaptive DWT in this proposal. The non-adaptive DWT if applied in such images will increase the energy of high pass subband, thus further decreasing its compression ratio and its coding efficiency.

To overcome these problems, lots of new techniques, for instance, curvelets [80], contourlets [81], bandelets [82] [83], wedgelets [84], directionlets [85] [86] and interpolation based direction-adaptive lifting [87] have been proposed. However, all of them have weaknesses in high computing times and complex filter design and they are not usually used in image compression. The adaptive discrete wavelet transform is first proposed by D. Taubman et al [88] by utilizing orientation of local image features to avoid the highly objectionable Gibbs-like phenomena observed at reconstructed image edges with conventional subband schemes at low bit rates. The image is partitioned into blocks then the blocks are sheared through a reversible resampling filter, and conventional 2D DWT is applied to the sheared block thus provides fading moments along the edges.

A decade later, various adaptive directional prediction methods based on the 3D DWT had been proposed in [89] [90] [91] [92]. They are realized by applying the 1D adaptive directional lifting [93] [94] [95] to the separable 3D DWT. They are adaptively changing the transform direction for local regions of the input 3D data to decrease their entropy. Since the entropy coding is generally used, the adaptive directional prediction improves the coding performance. Because of the separable structure, the conventional methods are naturally incompatible for lossless coding.

Hence, to further improve the performance of lossless coding, the adaptive directional lifting (ADL) structure based on non-separable 3D DWT is proposed. The proposed structure

is realized by introducing an adaptive directional framework into a 3D direct non-separable lifting structure based on DWT.

6.2.2 Existing Method

6.2.2.1 Non-Adaptive Non-Separable 3D Integer Wavelet Transform

The polyphase representation of one dimensional (1D) lifting structure of DWT is expressed in z domain as [96, 7, 97]

$$\begin{bmatrix} H_L(z) \\ H_H(z) \end{bmatrix} = \begin{bmatrix} s & 0 \\ 0 & \frac{1}{s} \end{bmatrix} \prod_{i=N-1}^{0} \left\{ \begin{bmatrix} 1 & U_i(z^2) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ P_i(z^2) & 1 \end{bmatrix} \right\} \times \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} = \mathbf{E}(z^2) \times \mathbf{d}(z), \quad (6.2)$$

where $H_L(z)$ and $H_H(z)$ are lowpass and highpass filters, respectively, N is a parameter which determines the length of filters, s, $U_i(z)$ and $P_i(z)$ denote scaling, update and prediction parameters, respectively. E(z) means the type-I polyphase matrix of the analysis bank of 1D DWT and d(z) is called a delay vector. In the case of 5/3 DWT, the parameters are defined as $N = 1$, $s = 1$, $U_0(z) = (1 + z)/4$, and $P_0(z) = -(1 + z^{-1})/2$, respectively. On the other hand, the parameters are defined in the case of 9/7 DWT as $N = 2$, $s = 1.1496$, $U_0(z) = -0.0530(1+z)$, $U_1(z) = 0.4435(1+z)$, $P_0(z) = -1.5861(1 + z^{-1})$, and $P_1(z) = 0.8829(1 + z^{-1})$, respectively.

According to the separable implementation of the above 1D DWT, the polyphase representation of the lifting structure of the separable 3D DWT based on 5/3 DWT is derived as

$$\mathbf{H}(z) = \begin{bmatrix} \mathbf{E}(z_2^2) & 0 & 0 & 0 \\ 0 & \mathbf{E}(z_2^2) & 0 & 0 \\ 0 & 0 & \mathbf{E}(z_2^2) & 0 \\ 0 & 0 & 0 & \mathbf{E}(z_2^2) \end{bmatrix} \times \begin{bmatrix} \mathbf{E}(z_1^2) & 0 \\ 0 & \mathbf{E}(z_1^2) \end{bmatrix} \mathbf{E}(z_2^2) \times \mathbf{d}_0(\mathbf{z}) \quad , \quad (6.3)$$

where $\mathrm{z} = [z_0, z_1, z_2]^T$, z_0, z_1 and z_2 means z components along the height, width and depth direction of 3D data, respectively, its sampling matrix M_0 is defined as

$$\mathbf{M}_0 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad (6.4)$$

and its delay vector $\mathbf{d}_0(\mathbf{z})$ is

$$\mathbf{d}_0^T(\mathbf{z}) = \left[1, \frac{1}{z_2}, \frac{1}{z_1}, \frac{1}{z_1 z_2}, \frac{1}{z_0}, \frac{1}{z_0 z_2}, \frac{1}{z_0 z_1}, \frac{1}{z_0 z_1 z_2}\right], \quad (6.5)$$

.

In [62], we propose the non-adaptive lifting structure of non-separable 3D DWT. The structure based on 1D 5/3 DWT is shown in Fig. 15, and the detail explanation of it can be referred to Section 3.3.3. Note that the number of rounding operators in each lifting step is reduced from 24 to 8, and then the rounding error is also reduced. However, this structure can only be utilized in horizontal and vertical directions, which is not suitable to be used in images with lots of directional high frequency components, the edges and lines, for instance, the light field image. Hence, in order to reduce its coding efficiency, the irrelevance data the image which is the high frequency band signals also needs to be reduced. Therefore, the adaptive directional lifting structure based on non-separable 3D DWT in [62] is proposed to overcome this problem.

6.2.3 Proposed Method

6.2.3.1 Adaptive Non-Separable 3D Integer Wavelet Transform

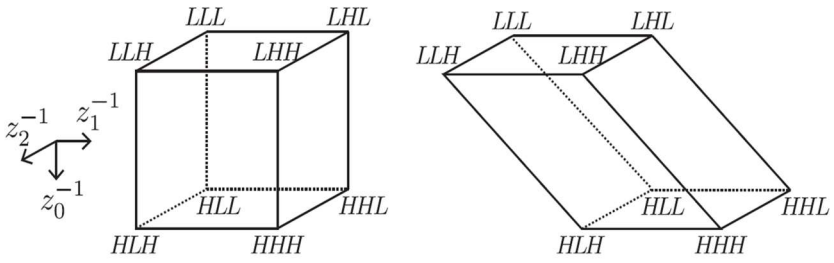
In this proposal, we propose an adaptive directional prediction based on the non-separable 3D DWT in Fig. 18. The proposed method divides the input 3D data into $N \times N \times N$ small 3D cubes and adaptively applies the non-separable 3D DWT to them along optimal directions according to their own property. The optimal direction means that the entropy of transformed coefficients along it is minimum in various directions. For example, if a cube is consisted of images with same diagonal stripe textures, we apply the non-separable 3D DWT to it along the diagonal direction in the height-width plane and the normal direction in the depth axis.

Similar to [66], to adaptively change the transform direction, we switch the sampling matrix for each cube. Of course, the delay vector is also changed according to the sampling matrix. Fig. 47 shows the transform classifications for $\mathbf{M}_0$ in (6.4) and $\mathbf{M}_1$ as

$$\mathbf{M}_1 = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad (6.6)$$

The figures claim that the transform direction is changed with switching the sampling matrix, and the adaptive directional prediction is realized by the adaptive switch method.

The candidates of the sampling matrix $\mathbf{M}_i$ can be arbitrarily determined under the restriction as $\det(\mathbf{M}_i) = 8$, where $\det(\mathbf{M}_i)$ denotes the determinant of $\mathbf{M}_i$. If the proposed method is applied for coding, the sampling matrices of each cube are stored as the side-information. Therefore, we should consider the trade-off between the number of sampling matrix candidates and the budget of the side-information.



(a) $\mathbf{M}_0$ (b) $\mathbf{M}_1$

Fig. 47 Transform classification according to M_1

The decision of finding suitable candidates influences the performance of image coding. Moreover, by increasing the candidates, the computational complexity of the proposed method is increased and one of particular processes in boundary of each cube is exponentially increased. The proposed method with many candidates becomes unpractical. In this proposal, we limit the candidates to 15 directions to reduce data volume of the side information. The direction for height-width directions are the horizontal-vertical, horizontal-right diagonal and horizontal-right diagonal directions. Whereas the direction for depth direction are the normal, top, bottom, left and right directions. Namely, the candidates are restrict to its minimum in $\{\mathbf{M}_{i,j}\}$ ($i = 0,1,2$ and $j = 0,1,2,3,4$) as

$$\mathbf{M}_{i,j} = \mathbf{M}_i + \mathbf{T}_j \qquad (6.7)$$

where

$$\mathbf{M}_2 = \begin{bmatrix} 2 & 0 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad (6.8)$$

$\mathbf{T}_0$ is a matrix whose elements are 0,

$$\mathbf{T}_1 = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{T}_2 = \begin{bmatrix} 0 & 0 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{T}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{T}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix}, \qquad (6.9)$$

In this case, the side-information is encoded in up to a 4 bit-per cube, i.e., $4/N^3$ bit-per-pixel (bpp). This is a reasonable value for lossless coding. In the above restriction, we can separately treat the transform direction in the height-width plane and the depth axis as shown in Fig. 48. In detail, prediction of depth axis between two planes based on depth direction in the matrices in (6.9) is illustrated in Fig. 49. It is reasonable to implement the proposed method, because it leads to the decrease of the variety of cube boundary processes and then reduce the hardware cost of the additive processes.

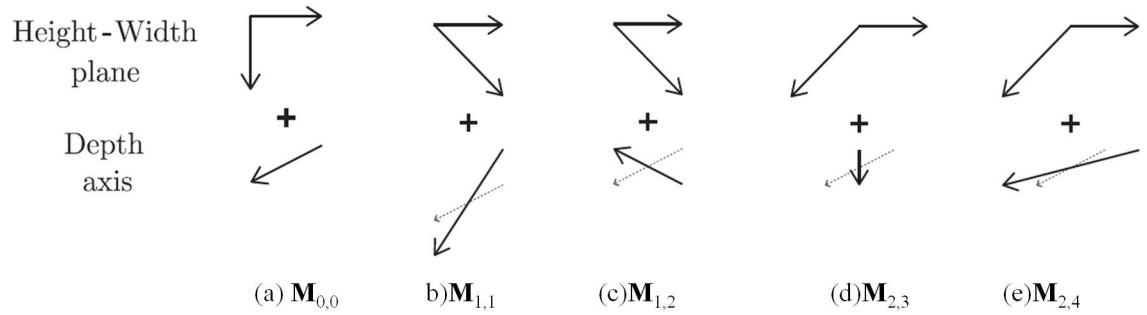


Fig. 48 Directions of $M_{i,j}$ (i = 0,1,2 and j = 0,1,2,3,4).

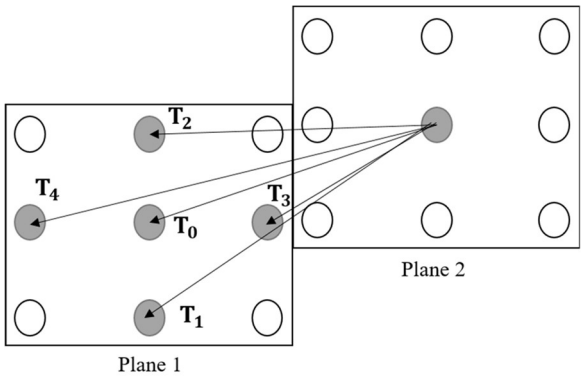


Fig. 49 Prediction of depth axis between two planes based on depth direction in the matrices.

For image coding, the optimal direction should maximally decrease the entropy of transformed image. It is well known that the entropy rate decreases as the energy of the signal decreases [96]. Therefore, we define the optimal direction according to the absolute mean of HHH subband coefficients for simplicity in this paper, similar to [32]. For all the sampling matrix $\mathbf{M}_i$, the HHH subband coefficients are described as h_i. The optimal direction $Dir(n)$ of the n-th cube is chosen such that it minimizes the energy of $h_d(n)$,

$$Dir\ (n) = \underset{i}{\operatorname{argmin}}|h_i(n)|, \qquad (6.10)$$

To reduce the budget of the side information, we divide the 3D data according to the octree decomposition. The decomposition procedure is explained as follows: The full quad tree T is constructed by applying the quad tree decomposition to the image until reaching the predefined block size, i.e., $N \times N \times N$. B, $D(B)$ and $R(B)$ are defined as a subtree, distortion and rate of bits. The most suitable subtree B^* with optimal directions is provided by minimizing the cost function defined as

$$B^* = \min_B J(B) = \min_B (D(B) + \lambda R(B)) \quad (6.11)$$

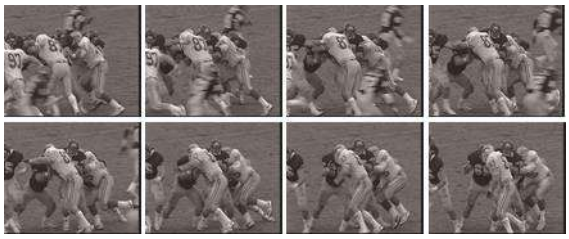
where

$$\begin{cases} D(B) = \sum_{\tau \epsilon B} \sum_{n} |h_{\tau}(n)| \\ R(B) = \sum_{\tau \epsilon B} r_{C}(\tau) + \sum_{v \epsilon B} r_{T}(v) \end{cases} \quad (6.12)$$

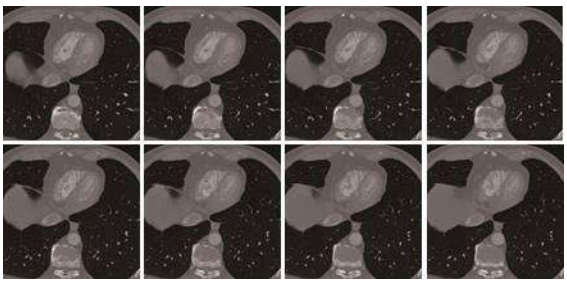
6.2.4 Experimental Results

To compare the proposed method with the separable and non-separable 3D DWT based on 5/3 DWT in Section 6.2.2, we applied them to lossless coding of 3D signals. The conventional methods are used with 5 decomposition levels, whereas the proposed method is performed first 1-level followed by the 4-level of the non-separable 3D DWT. The minimum cube size N is determined as $N = 16$. The rounding operator is introduced into each lifting step to achieve integer-to-integer transform. As test data, we used 5 videos, 5 medical slice data, and 5 images of light field data with size 288×352×8 , 512×512×8 , and 640×768×8 respectively, where $L \times M \times N$ shows the height×width×frames for video, and height×width×depth for medical data and light field data, as shown in Fig. 46. 'Foreman' video is used for small motion video, 'Football' video is used as a large motion video, 'Body' is used as a medical image data, and 'Treasure box' is used as a light field data as their test data. All data are in gray-scale with 8 bits per each pixel. As the coding method, the embedded zerotree wavelet based on intra-band partitioning (EZW-IP) [70] is developed for 3D signals. The side information is encoded by the arithmetic coding algorithm [98].

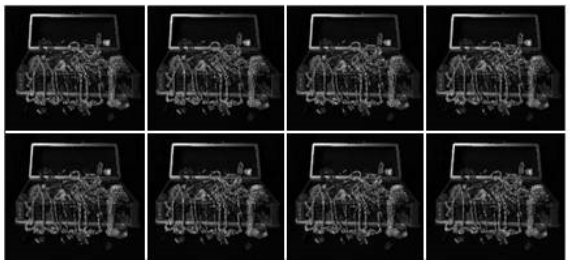
(a) Foreman



(b) Football



(c) Body



(d) Treasure box

Fig. 50 A part of test videos, medical slice data and light field data.

Table 6 shows the lossless coding results in bpp and their averages, where 'SEP', 'NSP' and 'PROP' mean the separable 3D DWT, the non-separable 3D DWT and the proposed method, respectively. The non-separable methods outperform the separable one due to a smaller number of rounding operators. The proposed method is the most effective in light field data, followed by videos, but slightly effective for medical slice data.

This is due to the characteristics of light field data which has clearer and sharper edges and has smaller changes in *z*-direction compared to videos and volumetric medical data. Here, the *z*-direction means the time direction (between frames) for video data. The improvement of the proposed method for videos with small motions is larger than ones with large motions from the foreman and football. Comparing the small motion video and large motion video, the small motion between frames in the small motion video making this method becoming more efficient to be used compared to the large motion one. However, the medical data which contains smooth edge and large difference between its slices has made the proposed method become slightly effective compared to the conventional ones. In conclusion, the sharper the edge, the smaller the difference between frames or slices in an image data, the better the performance of proposed method compared to both conventional methods.

Table 6 Lossless coding resuts in bpp

Type of Data	Name of Data	SEP	NSP	PROP
Large Motion Video	Football_A	4.7141	4.6713	**4.6540**
	Football_B	4.1775	**4.1129**	**4.1129**
Small Motion Video	Foreman_A	4.3516	4.3060	**4.2690**

	Foreman_B	4.2354	4.1906	**4.1712**
	Foreman_C	2.8968	**2.7043**	**2.7043**
Average		4.0751	3.9970	**3.9823**
Differences				-0.0147
Medical Data	Body_A	4.0450	4.0025	**3.9982**
	Body_B	4.2025	4.1564	**4.1495**
	Body_C	4.5022	4.4519	**4.4451**
	Body_D	4.5569	**4.5228**	4.5243
	Body_E	4.4970	4.4615	**4.4571**
Average		4.3607	4.3190	**4.3148**
Differences				-0.0042
Light Field Data	Robot	2.7503	2.6136	**2.6040**
	Tarot Cards	3.6458	3.5535	**3.4864**
	Treasure Box	2.7129	2.5741	**2.4899**
	Lego	3.1211	3.0096	**2.9742**
	Self	2.5171	2.3488	**2.3463**
Average		2.9494	2.8199	**2.7802**
Differences				-0.0398
Overall Average		3.7951	3.7120	**3.6924**

As for the computational cost of the proposed method which referred to processing time, the average processing time is approximately 15 times higher compared to the existing methods, where lots of time is needed to choose the optimal directions. The processing time is tested with MATLAB R2013a, Intel core i5-4590 CPU 3.30GHz, 64-bit Windows 10 Pro and 8GB RAM.

Table 7 Processing time for each method

		Processing Time [s]		
Type of Data	**Name of Data**	**SEP**	**NSP**	**ADL**
Large Motion Video	Football_A	0.069	0.126	16.631
	Football_B	0.110	0.130	14.411
Small Motion Video	Foreman_A	0.090	0.129	15.535
	Foreman_B	0.083	0.122	14.712
	Foreman_C	0.078	0.100	14.613
	Average	**0.094**	**0.127**	**14.886**
Medical Data	Body_A	0.184	0.400	93.347
	Body_B	0.210	0.400	96.721
	Body_C	0.229	0.348	98.993
	Body_D	0.223	0.353	97.168
	Body_E	0.185	0.348	97.795
	Average	**0.206**	**0.370**	**96.805**
Light Field Data	Robot	0.347	0.775	359.171
	Tarotcards	0.361	0.792	347.580

	Treasure Box	0.370	0.748	340.213
	Lego	0.358	0.737	348.905
	Self	0.375	0.686	369.697
Average		**0.362**	**0.748**	**353.113**
Overall avarage		**0.240**	**0.459**	**176.481**

The computational cost mentioned here is the processing time of the existing and proposed methods in the MATLAB. Since the increase of filter directions in the proposed method compared to the existing method, which is from 1 to 15, thus, making the processing time increase along with it. However, by making the non-separable 3D DWT as a base to apply the adaptive 3D DWT, the adaptive 3D DWT has the same latency with the non-separable 3D DWT, which is lower than the separable 3D DWT. Thus, in memory 3D accessing, the adaptive 3D DWT still outperform the conventional method, separable 3D DWT.

The graph in Fig. 52 shows the relationship between the overhead and its entropy rate. A medical data, 'Body_B' is used in this experiment. Three types of information bit are used in this experiment, which is 2[bit], 4[bit] and 16[bit] of transformed method. The 2[bit], 4[bit] and 16[bit] refers to the 1, 3 and 15 directions of adaptive 3D DWT, respectively. The original data is 8[bpp]. The data used in this experiment is 'foreman' as a small motion video as shown in Fig. 50(a), 'body' as a medical data as in Fig. 50(c) and 'tarot cards' as a light field data as in Fig. 51, with size 288×352×8, 512×512×8, and 640×768×8 respectively, where $L \times M \times N$ shows the height×width×frames for video, and height×width×depth for medical data and light field data. Results are shown in Table 8 and Fig. 52. It can be concluded that the higher the image information, the lower the bit per pixel of the compressed data, and the higher the bit per pixel of side information.

Fig. 51 Test data (Light field image=Tarot card)

Table 8 Bpp of compressed data and their overhead

	Bit per Pixel of Compressed Data [bpp]			**Bit per Pixel of Side Information [bpp]**		
Type of Data	2 [bit]	4 [bit]	16 [bit]	2 [bit]	4 [bit]	16 [bit]
Video	4.1906	4.1748	4.1712	0.0000	0.0010	0.0020
Medical Image	4.1564	4.1512	4.1495	0.0000	0.0002	0.0015
Light Field Data	3.5535	3.5256	3.4864	0.0000	0.0010	0.0024

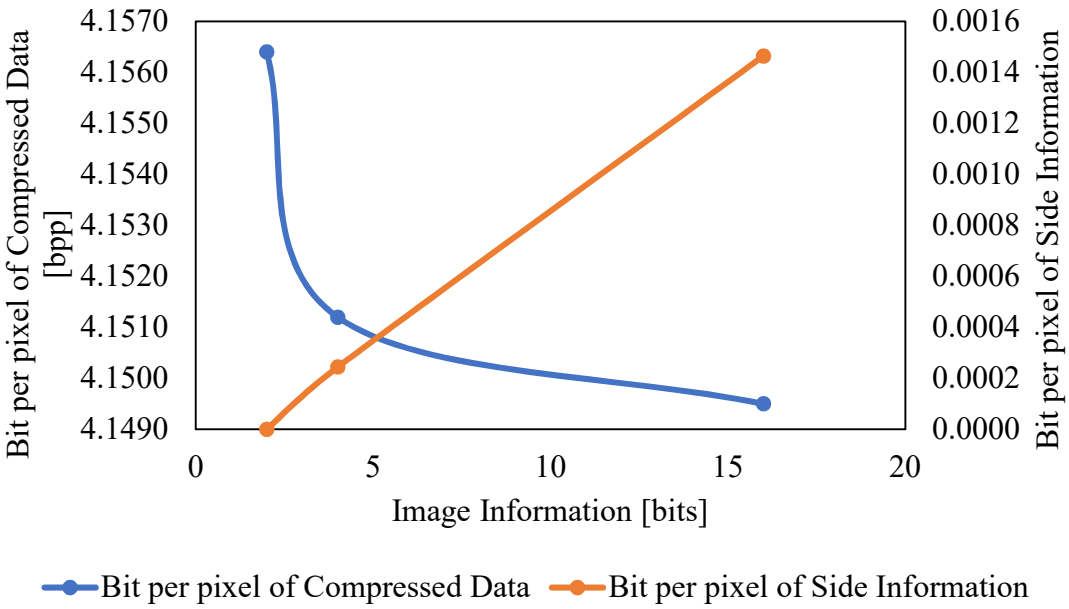


Fig. 52 Graph of the relationship between the bit per pixel of compressed data and its side information of a medical data.

6.2.5 Conclusion of this Section

In this chapter, we propose the directionally adaptive prediction of the non-separable 3D DWT. We adaptively switch its sampling matrix for each local region of input 3D data, based on the non-separable 3D DWT, to realize the adaptive method. Through simulations of lossless coding compared with the separable and non-separable 3D DWT, it is proven that the proposed method is effective for high resolution volumetric data compression, specifically in compressing light field data. In video, the proposed method is better effective for small motion videos than large motion ones.

6.3 Conclusion of this Chapter

This chapter presents further proposals for non-separable double lifting structure designed for 5/3-type integer wavelet transform. Section 6.1 explained on the region-of-interest (ROI) coding for 4D non-separable integer wavelet transform. The main idea of this proposal is to efficiently utilize the lossless and lossy coding of the double lifting structure. The ROI coding is important for medical images where the information preservation and increment of the compression performance are the important aspects to compress these images. It was found that the non-separable 4D integer wavelet transform with ROI coding is the best among other prevalent methods.

Section 6.2 covers on the forth proposal of the non-separable integer wavelet transform for 3D images. This proposal explained on the adaptive directional lifting structure for 3D non-separable integer wavelet transform for 5/3-type filter. This structure is designed to further compacting the energy of the image in low-frequency band signals thus increasing the coding performance of 3D images. The adaptive structure is more suitable to be applied in light field image with sharp edges and slow movement between the slice of the images. Although the structure is more complex than the non-adaptive, but it is reasonable to use this structure to further increase compression ratio for the 3D images.

Chapter 7 Conclusion of the Dissertation

7.1 Conclusion and Discussions

Chapter 1 introduces the structure of this dissertation and the main objective of this research. As the medical images have evolved from the 2D, 3D and 4D, such images will have huge uncompressed files size that will lead to inefficient storage and transmission. Therefore, it is necessary to research on the best compression method for them. The objective of this dissertation is to propose the non-separable integer wavelet transform for 4D image compression in both lossless and lossy mode.

Chapter 2 explains on the background of this research. Chapter 2 mainly covers on the introduction on image compression standards. The significance of compressing medical images is explained in this chapter. The image compression scheme and its type are presented here. Then, current international standards for image compression are introduced. This research is adopting the JPEG 2000 image compression standard that use wavelet transform to transform the image. This dissertation is also proposing the most suitable method for both lossless and lossy compression of 4D medical images.

Next, in Chapter 3, the second part of background for this research is presented. This chapter introduces the properties of wavelet transform. Wavelet transform is found to be more scalable than the DCT as it can compress image easily due to its basis function. Then, the necessity and basic introduction of integer transform are explained in this chapter. The integer transform is implemented by using lifting structure, where the products of the input signal and the filter coefficients are rounded to ensure the perfect reconstruction of the structure. Lifting structure is an important part to implement a lossless compression. Rounding noise existed in the transform is also explained here. Besides that, current existing methods of the lifting structure, which are the separable structure, 2D and 3D non-separable structure is introduced in this chapter.

Chapter 4 presents the first proposal of this dissertation which is the non-separable 4D integer wavelet transform for 5/3-type filter. This lifting structure is known as double lifting structure because it has only two filter coefficients. It is proved in this proposal that the 4D

non-separable structure has fewer rounding noise due to the fewer rounding operators inside the transform compared to the separable 4D structure. The fewer the rounding noise, the more the energy can be compacted in the low frequency band signal, and the higher the coding performance for 4D image. This proposal is tested in 5 types of input signal (4 of them are different types of medical images) and experimented in both lossless and lossy mode.

In Chapter 5, the second proposal for this dissertation is explained. The second proposal is the continuance of the first proposal. As JPEG 2000 presented two type of filters for lossless and lossy coding, which are the 5/3-type transform (lossless) and 9/7-type transform (lossy). The 9/7-type transform use a lifting structure known as the quadruple type. Although the non-separable 4D double lifting structure has lower rounding noise compared to the separable 4D double lifting structure, but the non-separable quadruple lifting structure has higher rounding noise despite of the reduction of the rounding operators inside the transform. This is due to the noise amplification inside the transform. Therefore, various structures are introduced to reduce the rounding noise. As a result of the experiments, it was found that it is necessary to maintain the arrangement of filter coefficients inside the transform in order to reduce the rounding noise. This experiment is done with 6 type of 4D input signals (5 of them are different type of medical images) and tested in the lossy mode.

Chapter 6 presents the third and fourth proposal of this dissertation. The third proposal is the extension of the first proposal. The non-separable structure is tested in the region-of-interest coding mode. The ROI coding is proposed to efficiently utilize both lossless and lossy coding with the aim to increase the coding performance of 4D medical images while preserving the important information in it. It is important for medical image compression as every loss of information may affect other's life. Meanwhile the fourth proposal clarifies on how to further compact the energy in low frequency band signal in order to increase the coding performance of 3D images. The adaptive lifting structure is designed in the fourth proposal, where the best lifting direction is chosen to get the most energy in the low frequency band signal. Although the computational complexity for this structure is higher than the non-adaptive structure, it is experimentally proven that the images transformed by adaptive structure yields higher compression ratio compared to the prevalent ones.

Finally, this dissertation is concluded in Chapter 7 and future works are discussed here.

7.2 Future Works

Although wavelet transform has an age of almost 20 years of research but is still worth to study this transform due to its scalable properties for image analysis and compression. The 4D image compression is still a hot research topic as there are various multidimensional images being reconstructed nowadays, for example the light field image and the mixed reality image. The most important thing is that in order to realize the lossless compression, the integer transform is the only key for that.

Besides wavelet transform, the integer transform can be realized in DCT, KLT, DST and even the latest transform coding, which is the GBT. However, to construct a perfect reconstruction for those integer transforms remains a challenge to researchers. In the future, it is hoped that this integer transform can be applied in those transforms and to other types of multidimensional data.

Other than that, the advanced of deep learning has opened the opportunity among researchers to implement the traditional digital signal processing with them. There are various studies have been done on combining wavelet transform with some machine learning architectures in the computer vision field. Thus, this research can still be a part of the ongoing works in the artificial intelligence field for image processing, specifically on image compression.

References

[1] A. Nait-Ali, and C. Cavaro-Menard, Compression of Biomedical Images and Signals, USA: John Wiley & Sons, Inc., 2008.

[2] H. K. Huang, PACS and Imaging Informatics: Basic Principles and Applications, USA: Wiley, 2010.

[3] R. Rajeswari and R. Rajesh, "Efficient Compression of 4D fMRI Images Using Bandelet Transform and Fuzzy Thresholding," in *World Congress on Nature & Biologically Inspired Computing*, India, 2009.

[4] V. Sanchez, P. Nasiopoulos, R. Abughrabieh, "Novel Lossless fMRI Image Compression based on Motion Compensation and Customized Entropy Coding," *IEEE Trans. on Information Technology in Biomedicine,* vol. 13, no. 4, pp. 645-655, 2009.

[5] H. G. Lalgudi, A. Bilgin, M. W. Marcellin, A. Tabesh, M. S. nadar, T. P. Trouard, "Four-Dimensional Compression of fMRI using JPEG 2000," *Proc. of SPIE, Medical Imaging: Image Processing,* vol. 5747, pp. 1028-1037, 2005.

[6] G. Hudson, A. Leger, B. Niss, I. Sebestyen, "JPEG at 25: Still Going Strong," *IEEE Multimedia,* vol. 24, no. 2, 2017.

[7] I. Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Steps," *Journal of Fourier Analysis and Applications,* vol. 3, no. 3, pp. 247-269, 1998.

[8] U. Albalawi, S. P. Mohanty, E. Kougianos, "A Hardware Architecture for Better Portable Graphics (BPG) Compression Encoder," in *IEEE Int. Symposium on Nanoelectric and Information Systems*, India, 2015.

[9] Li G, Citrin D, Camphausen K, Mueller B, Burman C, Mychalczak B, Miller RW, Song Y., "Advances in 4D medical imaging and 4D radiation therapy," *Technol Cancer Res Treat.,* vol. 7, no. 1, pp. 67-81, 2008.

[10] F. Cane, B. Verhegghe, M. D. Beule et al, "From 4D Medical Images (CT, MRI, Ultrasound) to 4D Structured Mesh Models for Patient-Specific Flow Simulations in the Human Heart," in *16th National Day on Biomedical Engineering & IEEE EMBS Benelux Chapter*, Brussels, 2017.

[11] Marc J. Gollub, "Picture Archiving and Communication Systems (PACS)," *Encyclopedia of Gastroenterology,* pp. 180-188, 2004.

[12] A. Gersho, R. M. Gray, Vector Quantization and Signal Compression, Massachusetts, USA: Kluwer Academic Publishers, 1992.

[13] V. K. Goyal, "Theoretical Foundations of Transform Coding," *IEEE Signal Processing Magazine,* vol. 01, pp. 9-21, 2001.

[14] V. Britanak, P. Yip and K. R. Rao, Discrete Cosine and Sine Transforms, Great Britain: Elsevier Ltd., 2007.

[15] G. K. Wallace, "The JPEG Still Picture Compression Standard," *IEEE Trans. on Consumer Electronics,* vol. 38, no. 1, pp. xviii-xxxiv, 1992.

[16] M. Budagavi, A. Fuldseth, G. Bjontegaard, "HEVC Transform and Quantization," in *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Switzerland, Springer International Publishing, 2014, pp. 141-169.

[17] K. R. Rao and P. Yip, Discrete Cosine Transform: Algorithms, Advantages, Applications, San Diego, CA, USA: Academic Press Professional, 1990.

[18] D. Marshall, "The Discrete Cosine Transform," 10 Apr 2001. [Online]. Available: http://users.cs.cf.ac.uk/Dave.Marshall/Multimedia/node231.html. [Accessed 09 Jan 2019].

[19] G. Shen, W.-S. Kim, A. Ortega, J. Lee, H. C. Wey, "Edge-adaptive Transforms for Efficient Depth Map Coding," in *Proc. of Picture Coding Symposium*, Nagoya, Japan, 2010.

[20] W.-S. Kim, S. K. Narang and A. Ortega, "Graph Based Transforms for Depth Video Coding," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Kyoto, Japan, 2012.

[21] Su Weng-Tai, Cheung Gene, Lin Chia-Wen, "Graph Fourier Transform with Negative Edges for Depth Image Coding," in *IEEE Int. Conf. on Image Processing*, 2017.

[22] H. E. Egilmez, Y.-H. Chao, A. Ortega, B. Lee and S. Yea, "GBST: Separable Transforms based on Line Graphs for Predictive Video Coding," in *Proc. of IEEE Int. Conf. on Image Processing*, 2016.

[23] H. E. Egilmez, A. Said, Y.-H. Chao and A. Ortega, "Graph-based Transforms for Inter Predicted Video Coding," in *Proc. of IEEE Int. Conf. on Image Processing*, 2015.

[24] S. I. Young, A. T. Nama, R. K. Mathew and D. Taubman, "Optimized Decoding of JPEG Images Based on Generalized Graph Laplacians," in *Proc. of Picture Coding Symposium*, 2016.

[25] F. Verdoja and M. Grangetto, "Directional Graph Weight Prediction for Image Compression," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, New Orleans, USA, 2017.

[26] D. Zhang and J. Liang, "Graph-based Transform for 2D Piecewise Smooth Signals with Random Discontinuity Locations," *IEEE Trans. on Image Processing,* vol. 26, no. 4, pp. 1679-1693, 2017.

[27] Xianming Liu, Gene Cheung, Xiaolin Wu, Debin Zhao, "Random Walk Graph Laplacian-Based Smoothness Prior for Soft Decoding of JPEG Images," *IEEE Transactions on Image Processing ,* vol. 26, no. 2, pp. 509-524, 2017.

[28] Wei Hu ; Gene Cheung ; Masato Kazui, "Graph-based Dequantization of Block-Compressed Piecewise Smooth Images," *IEEE Signal Processing Letters,* vol. 23, no. 2, pp. 242-246, 2016.

[29] D. I. Shuman, S. K. Narang, P. Frossards, A. Ortega, P. Vandergheynst, "The Emerging Field on Signal Processing on Graphs," *IEEE Signal Processing Magazine,* vol. 30, no. 3, pp. 83-98, May 2013.

[30] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.,* vol. 41, pp. 909-996, 1998.

[31] M. W. Marcellin, M. A. Lepley et al, "An overview of quantization in JPEG 2000," *Signal Processing: Image Communication,* vol. 17, pp. 73-84, 2002.

[32] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Prentice Hall, 2004.

[33] D. Taubman and M. Marcellin, JPEG2000: Image Compression Fundamentals, Standards and Practice, Springer, 2002.

[34] D. Taubman, "High Performance Scalable Image Compression with EBCOT," *IEEE Trans. on Image Processing,* vol. 9, no. 7, pp. 1158-1170, 2000.

[35] Z. Liu, L. J. Karam, "An efficient embedded zerowavelet image codec based on intraband partitioning," in *IEEE Int. Conf. Image Processing*, Vancouver, Canada, Sept. 2000.

[36] T. Berger, "Optimum Quantizers and Permutation Codes," *IEEE Transactions on Information Theory,* vol. 18, no. 6, pp. 759-765, 1972.

[37] D. LeGall and A. Tabatai, "Subband Coding of Digital Images using Symmetric Short Kernel Filters and Arithmetic Coding Techniques," in *Int. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New York, 1988.

[38] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image Coding using Wavelet Transform," *IEEE Trans. on Image Processing,* vol. 1, pp. 205-220, Apr. 1992.

[39] T. Halbach, "Performance Comparison: H.26L Intra Coding vs. JPEG 2000," in *JVT-D039*, July 2000.

[40] T. Ebrahimi, S. Foessel et al, "JPEG Pleno: Toward an Efficient Representation of Visual Reality," *IEEE Multimedia,* vol. 23, no. 4, pp. 14-20, 2016.

[41] L. A. Barford, R. Shane Fazzio, D. R. Smith, "An Introduction to Wavelets," Hewlett Packard Technical Report, 1992.

[42] P. Schelkens, A. Skodras, T. Ebrahimi, The JPEG 2000 Suite, A John Wiley and Sons, Ltd., Publication, 2009.

[43] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Appl. Comput. Harmon. Anal.,* vol. 3, no. 2, pp. 186-200, 1996.

[44] T. Strutz, I. Rennert, "Two-dimensional integer wavelet transform with reduced influence of rounding operations," *EURASIP Journal on Advances in Signal Processing,* vol. 75, 2012.

[45] M. Iwahashi, H. Kiya, "Discrete wavelet transforms: Non separable two dimensional discrete wavelet transform for image signals," in *Discrete Wavelet Transforms - A Comnpendium of New Approaches and Recent Applications*, InTech, 2013.

[46] S. Poomrittigul, M. Iwahashi, and H. Kiya, "Reduction of Lifting Steps of Non Separable 2D Quadruple Lifting DWT Compatible with Separable 2D DWT," *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences ,* Vols. E97-A, no. 7, pp. 1492-1499, July 2014.

[47] "Joint Photographic Experts Group: JPEG 2000 image coding system". Patent ISO / IEC FCD 15444-1, March 2000.

[48] M. Iwahashi and H. Kiya, "A New Liftings Structure of Non Separable 2D DWT with Compatibility to JPEG2000," in *IEEE ICASSP*, Texas, USA, March 2010.

[49] Chao He, J. Dong, Y. F. Zheng, Z. Gao, "Optimal 3-D coefficient tree structure for 3D wavelet coding," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 13, no. 10, pp. 961-972, 2003.

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

[51] A. Aggoun, "Compression of 3D integral images by using 3D wavelet transform," *IEEE Journal of Display Technology,* vol. 7, no. 11, pp. 586-592, 2011.

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

[53] B. Penna, T. Tillo, E. Magli, G. Olmo, "Progressive 3D coding of hyperspectral images based on JPEG 2000," *IEEE Geoscience Remote Sensing Letters,* vol. 3, no. 1, pp. 125-129, 2006.

[54] T. Bruylants, A. Munteanu, P. Schelkens, "Wavelet based volumetric medical image compression," *Signal Processing: Image Communication,* vol. 31, pp. 112-133, 2015.

[55] J. M. Gomez, J. B. Rapesta, I. Blanes, L. J. Rodriguez, F. A. Llinas, J. S. Sagrista, "4D remote sensing image coding with JPEG2000," *Satellite Data Compression, Communications and Processing VI,* vol. 7810, 2010.

[56] Y. Wang and H. Hamza, "4D Geometry Compression based on Lifting Wavelet Transform," in *Industrial Engineering Research Conference*, Florida, 2006.

[57] M. Unser and T. Blu, "Mathematical properties of the JPEG 2000 wavelet filters," *IEEE Trans. Image Processing,* vol. 12, no. 9, pp. 1080-1090, 2003.

[58] C.Chrysafis, A. Ortega, "Line-based, reduced memory, wavelet image compression," *IEEE Trans. on Image Processing,* vol. 9, no. 3, pp. 378-389, 2000.

[59] G. Shi, W. Liu, Li Zhang and Fu Li, "An efficient folded architecture for lifting-based discrete wavelet transform," *IEEE Trans. Circuits, Systems II express briefs,* vol. 56, no. 4, pp. 290-294, 2009.

[60] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing,* vol. 40, no. 9, pp. 2207-2232, 1992.

[61] Y. Zhang, D. R. Bull, E. Reinhard, "Perceptually lossless high dynamic range image compression with JPEG 2000," in *IEEE International Conference on Image Processing*, 2012.

[62] M. Iwahashi, T. Orachon and H. Kiya, "Three dimensional discrete wavelet transform with deduced number of lifting steps," in *IEEE International Conference on Image Processing*, 2013.

[63] D. S. Taubman, "Adaptive, non-separable lifting transforms for image compression," in *IEEE International Conference on Image Processing*, 1999.

[64] S. Fukuma, M. Iwahashi and N. Kambayashi, "Adaptive multi-channel prediction for lossless scalable coding," in *IEEE International Symposium on Circuits and Systems*, 1999.

[65] F. A. B. Hamzah, T. Yoshida, M. Iwahashi and H. Kiya, "Adaptive directional lifting structure of three dimensional non-separable discrete wavelet transform for high resolution volumetric data compression," *IEICE Trans. Fundamentals,* Vols. E99-A, no. 5, pp. 892-899, 2016.

[66] T. Yoshida, T. Suzuki, S. Kiyochi and M. Ikehara, "Two Dimensional Non-separable Adaptive Directional Lifting Structure of Discrete Wavelet Transform," *IEICE Trans. Fundamentals,* Vols. E94-A, no. 10, pp. 1920-1927, Oct. 2011.

[67] M. Iwahashi, H. Kiya, "Non separable 2D factorization of separable 2D DWT for lossless image coding," in *IEEE International Conference Image Processing*, 2009.

[68] D. Boye, et. al, "Population based Modeling of Respiratory Lung Motion and Prediction from Partial Information," in *Proc. SPIE 8669, Medical Imaging 2013: Image Processing*, 2013.

[69] Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L. & Pietrini, P. , " Distributed and overlapping representations of faces and objects in ventral temporal cortex," *Science,* pp. 2425-2430, 2001.

[70] Z. Liu and L. J. Karam, "An Efficient Embedded Zerowavelet Image Codec based on Intraband Partitioning," in *IEEE Int. Conf. Image Processing*, Vancouver, Canada, Sept. 2000.

[71] C.-L. Kuo, Y.-Y. Lin and Y.-C. Lu, "Analysis and Implementation of Discrete Wavelet Transform for Compressing Four-Dimensional Light Field Data," in *IEEE Int. SOC Conference*, Germany, 2013.

[72] A. Sang, T. Sun , H. Chen and H. Feng, "A 4D nth-order Walsh Orthogonal Transform Algorithm Used for Color Image Coding," in *Int. Conf. on Image Analysis and Signal Processing*, China, 2010.

[73] M. Iwahashi, T. Orachon, H. Kiya, "Non separable 3D lifting structure compatible with separable quadruple lifting DWT," in *Proc. of Asia Pacific Signal and Information Processing (APSIPA) Annual Summit and Conference (ASC)*, 2013.

[74] F. A. Binti Hamzah, T. Yoshida and M. Iwahashi, "Non-Separable Quaruple Lifting Structure for Four-Dimensional Integer Wavelet Transform with Reduced Rounding Noise," in *Proc. of IEEE ICASSP*, New Orleans, USA, Mar 2017.

[75] M. Iwahashi and H. Kiya, "Condition on Word Length of Signals and Coefficients for DC Lossless Property," in *Discrete Wavelet Transforms*, InTech, ISBN 978-953-307-313-2, 2011, pp. 231-254.

[76] C. Cortes, L. Kabongo, I. Macia, O. E. Ruiz and J. Florez, "Ultrasound Image Dataset for Image Analysis Algorithms Evaluation," *Innovation in Medicine and Healthcare 2015. Smart Innovation, Systems and Technologies,* vol. 45, pp. 447-457, 2015.

[77] J.-R. Ohm, "Linear Systems and Transforms," in *Multimedia Communication Technology: Representation, Transmission and Identification of Multimedia Signals*, New York, Springer-Verlag, 2004, pp. 96-104.

[78] J. Batrina-Rapesta, J. Serra-Sagrista, F. Auli-Llinas, "JPEG 2000 ROI coding through component priority for digital mammography," *Computer Vision and Image Understanding,* vol. 115, pp. 59-68, 2011.

[79] V. Sanchez, "Joint Source/Channel Coding for Prioritized Wireless Transmission of Multiple 3-D Regions of Interest in 3-D Medical Imaging Data," *IEEE Trans. on Biomedical Engineering,* 2013.

[80] E. J. Cands and D. L. Donoho, "Curvelets a Surprisingly Effective Nonadaptive Representation for Objects with Edges," *Curve and Surface Fitting: Saint-Molo. Nashville, TN: Univ. Press,* pp. 105-120, 1999.

[81] M. N. Do and M. Vetterli, "The Contourlet Transform: An Efficient Directional Multiresolution Image Representation," *IEEE Trans. Image Process.,* vol. 14, no. 3, pp. 2091-2106, Dec. 2005.

[82] E. L. Pennec and S. Mallat, "Sparse Geometric Image Representation with Bandelets," *IEEE Trans. Image Process.,* vol. 14, no. 4, pp. 423-438, Apr. 2005.

[83] G. Peyre and S. Mallat, "Surface Compression with Geometric Bandelets," *ACM Trans. Graphics,* vol. 24, no. 3, pp. 601-608, 2005.

[84] D. L. Donoho, "Wedgelets: Nearly Minimax Estimation of Edges," *Ann. Statist.,* vol. 27, no. 3, pp. 859-897, 1999.

[85] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Directionlets: Anisotropic Multi-directional Representation with Separable Filtering," *IEEE Trans. Image Process.,* vol. 15, no. 7, pp. 1916-1933, Jul. 2006.

[86] D. Jayachandra and A. Makur, "Directionlets using In-phase Lifting for Image Representation," *IEEE Trans. Image Process.,* vol. 23, no. 1, pp. 240-249, Jan. 2014.

[87] Zhijun Fang, Naixue Xiong, Yang, L.T., Xingming Sun, Yan Yang, "Interpolation-based Direction-adaptive Lifting DWT and Modified SPIHT for Image Compression in Multimedia Communications,," *IEEE System Journal,* vol. 5, no. 4, pp. 484-493, Nov. 2011.

[88] D. Taubman and A. Zakhor, "Orientation Adaptive Subband Coding of Images," *IEEE Trans. Image Process.,* vol. 3, no. 4, pp. 421-437, Jul. 1994.

[89] J. R. Ohm, M. van der Schaar, and J. W. Wood, "Interframe Wavelet Coding–Motion Picture Representation for Universal Scalability," *Signal Process.-Image Commun.,* vol. 19, no. 9, pp. 877-908, 2004.

[90] Y. Liu, F. Wu, and K. N. Ngan, "3-D Object-based Scalable Wavelet Video Coding with Boundary Effect Suppression," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 18, no. 7, pp. 888-899, July 2008.

[91] Y. Liu, K. N. Ngan, and F. Wu, "3-D Shape-adaptive Directional Wavelet Transform for Object-based Scalable Video Coding," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 17, no. 5, pp. 639-644, May 2007.

[92] A. Das, A. Hazra, and S. Banerjee, "An Efficient Architecture for 3-D Discrete Wavelet Transform," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 20, no. 2, pp. 286-296, Feb. 2010.

[93] O. N. Gerek and A. E. C¸etin, "A 2-D Orientation-adaptive Prediction Filter in Lifting Structures for Image Coding," *IEEE Trans. Image Process.,* vol. 15, no. 1, pp. 106-111, Jan. 2006.

[94] W. Ding, F. Wu, X. Wu, S. Li, and H. Li, "Adaptive Directional Lifting based Wavelet Transform for Image Coding," *IEEE Trans. Image Process.,* vol. 16, no. 2, pp. 416-427, Feb. 2007.

[95] C.L. Chang and B. Girod, "Direction-Adaptive Discrete Wavelet Transform for Image Compression," *IEEE Trans. Image Processing,* vol. 16, no. 5, pp. 1289-1302, May 2007.

[96] P. P. Vaidyanathan, Multirate Systems and Filter Banks, Englewood Cliffs, NJ: Prentice Hall, 1993.

[97] S. Mallat, A Wavelet Tour of Signal Processing, New York: Academic, 1998.

[98] P. G. Howard and J. S. Vitter, "Arithmetic Coding for Data Compression," in *IEEE*, Jun 1994.

Work Achievements

List of Peer-Reviewed Research Articles

1. **F. A. Binti Hamzah**, T. Yoshida, M. Iwahashi, H. Kiya, "Adaptive Directional Lifting Structure of the Three-Dimensional Discrete Wavelet Transform for High Resolution Volumetric Data," *IEICE Transactions on Fundamentals*, Vol.E99-A, No.5, pp. 892-899, May. 2016.

2. **F. A. Binti Hamzah,** T. Yoshida, M. Iwahashi, "Non-separable Four Dimensional Integer Wavelet Transform with Reduced Rounding Noise," *EURASIP Signal Processing Image Communication*, Vol. 58, pp. 123-133, Oct. 2017.

3. **F. A. Binti Hamzah**, S. Minewaki, T. Yoshida, M. Iwahashi, "Reduction of Rounding Noise and Liftings Steps in Non-Separable Four-Dimensional Quadruple Lifting Integer Wavelet Transform," *EURASIP Journal on Image and Video Processing*, Vol. 36, pp. 1-18, May 2018.

List of Peer-Reviewed Conference Proceedings

1. **F. A. Binti Hamzah,** T. Yoshida, M. Iwahashi, "Non-Separable Quadruple Lifting Structure for Four-Dimensional Integer Wavelet Transform with Reduced Rounding Noise," *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Mar 2017.

2. **F. A. Binti Hamzah**, T. Yoshida, M. Iwahashi, "Four-Dimensional Image Compression with Region of Interest Based on Non-separable Double Lifting Integer Wavelet Transform," *Proc. of Asia Pacific Signal Processing Annual Summit and Conference (APSIPA ASC)*, Dec 2017.

Others

1. **F. A. Binti Hamzah**, "The Non-Separable JPEG 2000 for Better 4D Image Compression," *Top-10 finalist of Three-Minute Thesis Competition, European Signal Processing Conference (EUSIPCO*), Rome, Italy, Sep 2018.