A Study on Noise Reduction of Non-Separable Wavelet Transform with Lifting Structure

by

12700985

TEERAPONG ORACHON

Doctoral Dissertation

Submitted to the Faculty of the Graduate School of

Nagaoka University of Technology

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF INFORMATION SCIENCE AND CONTROL ENGINEERING

Supervisor

Professor Masahiro Iwahashi

NAGAOKA UNIVERSITY OF TECHNOLOGY

Nagaoka, Niigata, Japan

Doctoral Dissertation

A Study on Noise Reduction of Non-Separable Wavelet Transform with Lifting Structure

Author: 12700985 Teerapong Orachon

Supervisor: Prof. Dr. Masahiro Iwahashi

Date: 14th June 2016

Abstract

JPEG 2000, a compression standard based on the Discrete Wavelet Transformation (DWT) with entropy coding by a one-dimensional signal processing vertically and horizontally (a separation type structure), offers many novel features than its conventional standard JPEG especially for high-resolution multimedia. The implementation in second generation wavelet transform is composed of the lifting steps which are computed in cascaded style. Hence, the main bottleneck in the scheme is the delay time taken for completion of each step. Presently, several non-separable factorization researches have been proposed as these can be utilized in minimizing the lifting scheme of the WT. In addition, the compatibility of the structure for both lossless and lossy image compression is restricted to three-dimensional data or higher. On the other hand, there is the particular issue related to noise within the lifting steps or on reduction of them is still an on-going interest. This dissertation firstly proposed non-separable lifting structure for three-dimensional lossless and lossy compression. Afterwards, the noise in the lifting structure was analyzed and, the precise noise minimization methods are proposed.

As for the first proposal, this dissertation introduced 5/3 non-separable lifting structure for 3D lossless compression. By applying the two basic

properties, the scheme reduced the lifting steps in the existing, separable lifting structure to non-separable lifting structure. Based on the evaluation result, the proposed non-separable method has higher PSNR in both frequency and the spatial domain. Also, the approach has better coding performance in both lossless and lossy coding within the same bit-rate.

Subsequently, 9/7 non-separable lifting structure for 3D lossy compression with minimum rounding noise was introduced. Comparing in lifting structure, non-separable 3D has minimum lifting step. However, rounding noise in non-separable 3D is larger than separable 3D. To clarify the minimum rounding noise structure, 6 rules were introduced for identifying non-separable candidates with minimum rounding noise. After the observation of rounding noise in non-separable candidate and comparison between separable 3D and non-separable 3D, the results showed that rounding noise of non-separable 2D structure is minimum.

However, the reduction of lifting steps together with rounding operators in 9/7 non-separable structure caused the amplification in the rounding noise. In that case, the research further focuses on the rounding noise by introducing the scaling method which could also improve the reconstructed image quality.

Finally, for the moderating of the word length of the coefficient filter in the lifting structure which could increase the hardware complexity, the truncation method to the word length of the coefficient applied in the lifting structure is initiated. As a consequence, this leads to the truncation noise affecting the PSNR of the reconstructed image. To tackle the particular issue, an optimization procedure for the word length of the each coefficient was lastly proposed.

Table of Contents

List of Figure

List of Tables

Chapter 1

1. Introduction

1.1. Background of Image Compression

Presently, by cause of the widely usage of digital information or digital data for daily life, the concern for data storage and data bandwidth is raising for storage and transmission of the voluminous data. This much, the compression algorithm plays an important role in compression reduction of the desire information. On the other hand, in the age of today's electronic world, the digital imaging such as digital photography, medical image processing and satellite imaging plays a vital role and, many image compression algorithms are emerged and, they become one of the attracting fields in the research area.

There has been two general categories in image compression algorithms [1], lossy and lossless compression. Lossy compression algorithm has high compression ratios trade off with the perception loss, the distortion that human eye can accept. Differently, in lossless compression algorithms decoded image has the same quality as an original image without any loss although there is a low compression ratio. In that sense, lossy compression algorithms are well suited for the applications that can accept the perception loss such as multimedia data, streaming media and online telephony. Comparatively, lossless compression is required for the real time application that requires high quality such as medical image and satellite image. The

system flow of the lossy and lossless compression system can be described as shown in Figure 1.

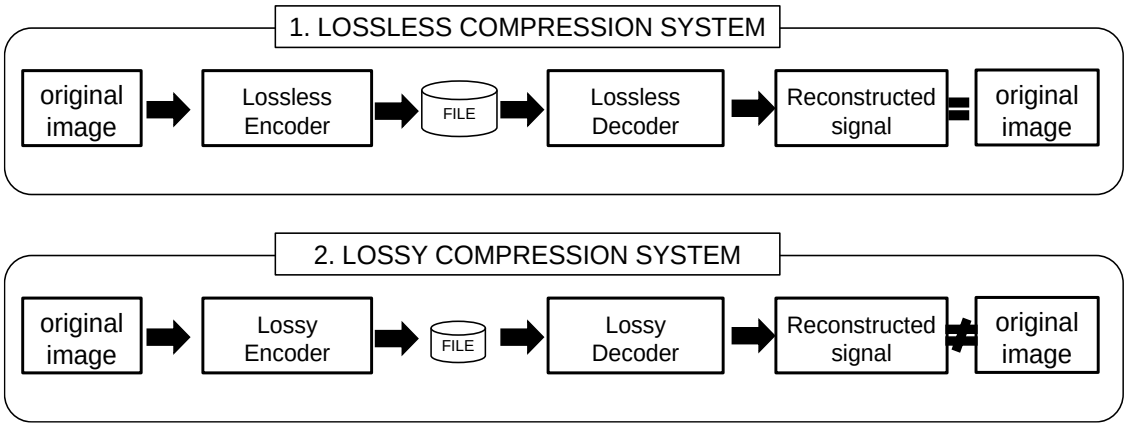


Figure. 1 Lossless and Lossy compression system.

Recently, JPEG 2000 [2] international standard become the extensively used standard compression which is suitable for compression of images the in specific fields such as medical image, digital cinema and high resolution image. In the first international compression standard for continuous-tone still images, the JPEG Standard, the DCT was used as a basic transformation tool. Afterwards, as digital imagery equipment became more widely used, its extension standard JPEG 2000 which offers many novel features than that of it, has led. The standard is based on Discrete Wavelet Transformation (DWT) with arithmetic entropy coding which can be achieved by a one-dimensional signal processing vertically and horizontally (a separation n type structure) based on the line memory. And, each process is implemented by cascading several lifting calculations**.** The wavelet transform is used especially to remove redundancy of information before encoding. Comparison to old

standard JPEG with DCT (Discrete Cosine Transform), JPEG2000 is better than JPEG in image quality and compression ratio as show in Figure 2.

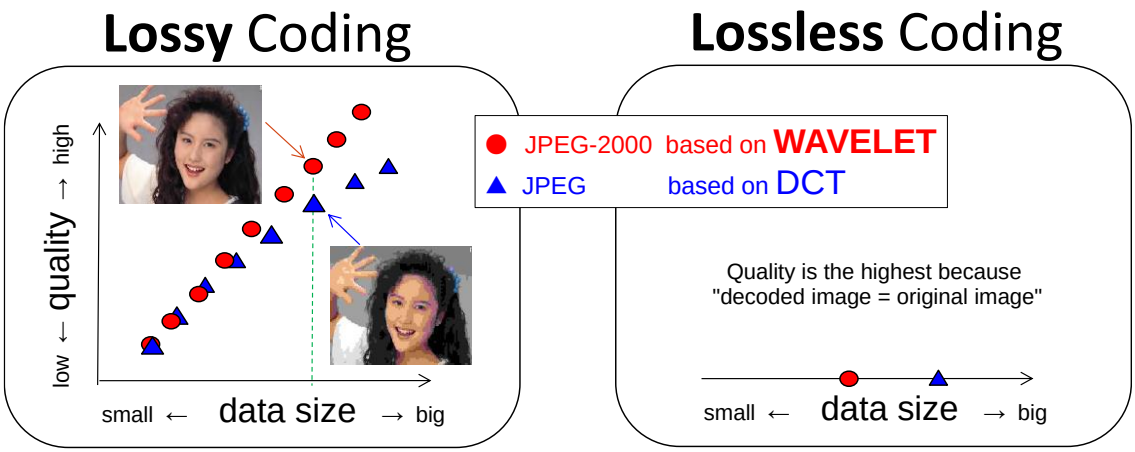


Figure. 2 Comparison between JPEG and JPEG 2000 in image quality and compression ratio.

An example of the lifting structure in the wavelet transform that compatible with JPEG2000 is illustrated in Figure 3. Due to the direction of calculation in each step is separate, so, this lifting structure can be named as “separable”.

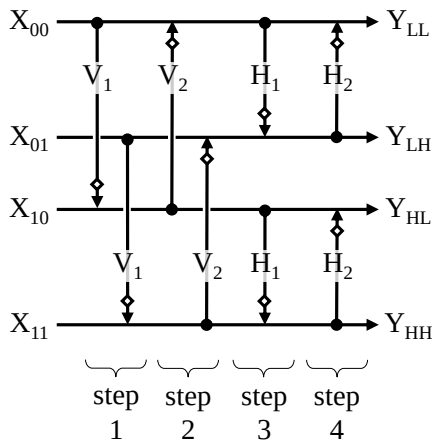


Figure. 3 Lifting structure for 2D data (image) or separable.

Separable has many lifting step which has rounding operators which are used in the process of transforming the floating point numbers to integer numbers, whereas the rounding errors occur in every rounding operator. For that reason, many studies try to focus on the reduction of rounding error where one of them is Non-separable approach. Several evidences studied about “Non-separable” for minimizing rounding errors are based on the lifting steps compensation. Figure 4 shows an example of non-separable lifting structure for image compression.

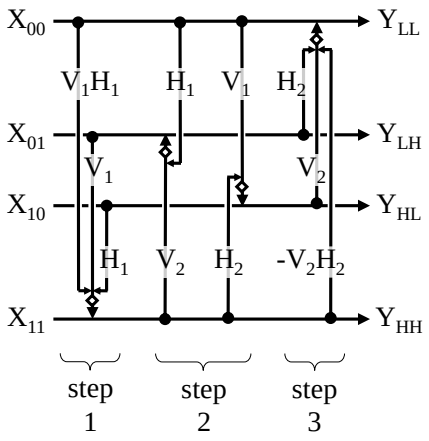


Figure. 4 Non-separable lifting structure for image compression.

For image compression, 2D non-separable lifting structure can be used as shown in Figure 4. But there is some limitation for this lifting. It can use only in 2D data such as image. Non-separable for 3D data has not been much researched yet. And also noise in lifting structure such as rounding noise and coefficient noise has not been clarified yet.

In this dissertation, we focus on the non-separable lifting structure of the wavelet transform. And our proposal exhibited the non-separable lifting structure for 3D data

that compatible with JPEG2000 for lossless and lossy compression respectively. And the deep analysis is about rounding noise and coefficient noise and finally the algorithm for the reduction of that of noise is proposed.

1.2. Overview of Research Topic

In this dissertation, we presented four proposals for solve the issues that we mentioned above as,

- Non-separable 3D integer wavelet transform for lossless data compression
- Integer Implementation of 3D Wavelet Transform for Lossy Data Compression
- Channel Scaling for Integer Implementation of Minimum Lifting 2D Wavelet Transform
- Word Length Allocation for Multiplier Coefficients of Minimum Lifting Non-Separable 2D Wavelet

1.2.1. Overview of Non-separable 3D integer wavelet transform for lossless data compression

The 5/3 non-separable lifting structure for 3D lossless compression was proposed by two basic properties which reduced lifting step of existing, separable lifting structure to non-separable lifting structure. For the evaluation of the noise in lifting structure, the proposed, Non-separable has maximum PSNR in both frequency domain

and pixel domain. Also, non-separable has gain the better coding performance when compared with the same bitrate in lossless and lossy coding.

1.2.2. Overview of Integer Implementation of 3D wavelet transform for lossy data compression

The 9/7 non-separable lifting structure for 3D lossy compression with minimum rounding noise was proposed in this dissertation. From separable 3D, the previous proposed method, non-separable 3D has minimum in lifting step. However, the rounding noise in Non-separable 3D is larger than separable 3D. To clarify the better structure than non-separable 3D, 6 rules were created to find the non-separable candidates with minimum rounding noise. The results showed that rounding noise in non-separable 2D structure is minimum after observation.

1.2.3. Overview of channel scaling for integer implementation of minimum lifting 2D wavelet transform

In the previous our proposed method, reduction of the lifting steps and rounding operators of the 9/7 non-separable lifting structure is introduced however rounding noise was amplified. For that reason, scaling method was proposed not only to reduce rounding noise but also to improve the quality of the decoded images.

1.2.4. Overview of word length allocation for multiplier coefficients of minimum lifting non-separable 2d wavelet

Based on the consideration of the design complexity in hardware perceptive, the word length of coefficient filter in lifting structure is also a concerned and such, this needs to be truncated. As a result, the truncation noise has occurred so that the coefficients have a different effect to PSNR of the reconstructed image. In consequence, this method was proposed for how to optimize word length of each coefficient while maintain PSNR of reconstructed image.

1.3. Outline of the Dissertation

There are 7 chapters in this dissertation which can be divided as shown in Figure 5.

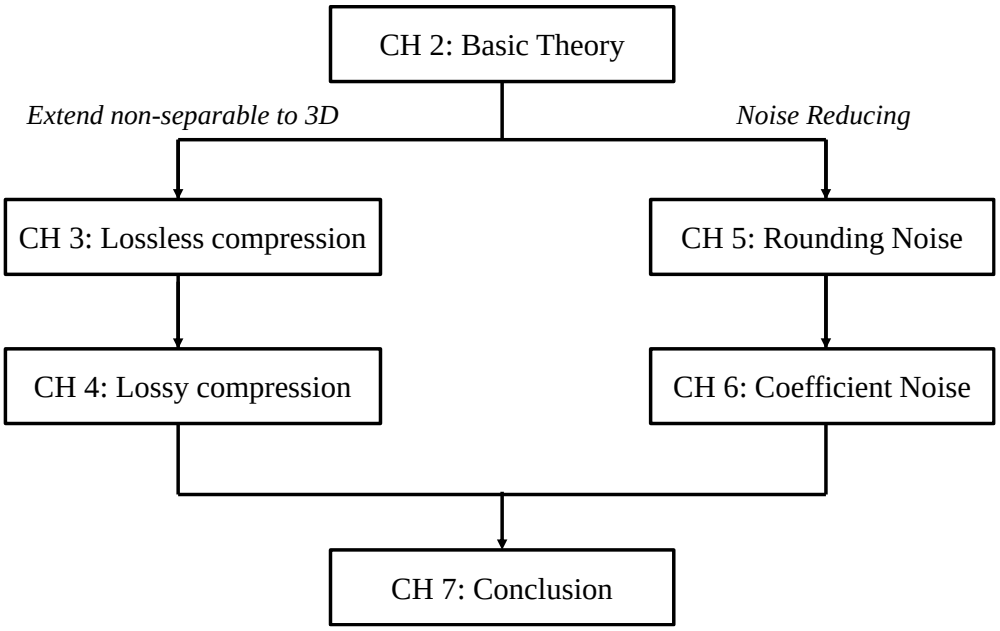


Figure. 5 Outline of this Dissertation

Chapter 2

2. Basic theory

2.1. JPEG2000

The development of data compression algorithms is occurred in a variety of approaches for tackling the issues of voluminous data transmission and storage. From the view of 2D signal image encoding, JPEG, JPEG XR and JPEG2000 [2] standards are widely used today. Again for the compatibility with 3D signal, for instance, video, the encoder standards, such as, H.261, MPEG2, H.264 and HEVC, are considered. In this dissertation, our analyzed and proposed methods are based on the examination of medical images where there is the higher resolution is voluntary as well as the compression practices. The most appropriate encoder in such cases is JPEG2000 as shown in Figure 6.

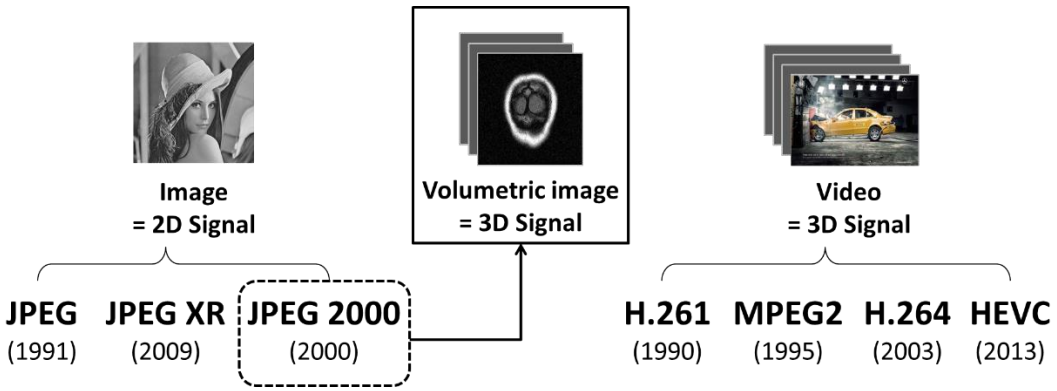


Figure. 6 The signal and encoder compatibility.

JPEG2000 is an image compression standard and coding system which is used for encoding the images in some specific areas such as digital cinema, geographic

information systems and medical imaging. The standard was developed by the JPEG (Joint Photographic Experts Group) committee in 2000 as the next generation of the JPEG standard that created in 1992.The wavelet transform is adopted in this standard for reduction of redundancy information data before encoding. For lossless mode, the 5/3 lifting structure is used to applied and the 9/7 lifting structure for the lossy mode.

2.2. Wavelet transform

According to the concept of data compression in wavelet transform [3], data need to be divided into 2 categories as low frequency and high frequency. At low frequency, the variance of data is large yet in high frequency, the variance of that of data is small which cause the separate inputs to the transform as shown in Figure 7.

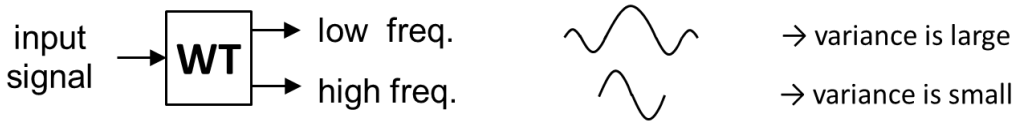


Figure. 7 Compressing input data by wavelet transform.

While comparing the DCT (discrete cosine transform) [4] which used in prior JPEG standard, wavelet transform in JPEG2000 more flexible and better in compression ratio (lossless compression) and reconstruct image quality (lossy compression). This is by because of the fact that the longer basis of wavelet is done in low frequency and short in high frequency while basis of DCT are the same length as shown in Figure 8. Due to the flexible property of representing image, the wavelet

transform in JPEG 2000 is well suited for much image application areas for the reason that the image data is mixed with low and high frequency.

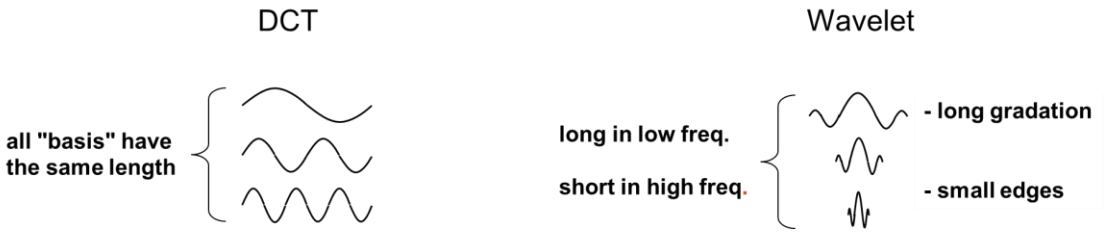


Figure. 8 Comparison between DCT and Wavelet basis.

Wavelet can decompose image data by using high pass filter h[n] and low pass filter g[n]. The output from high pass and low pass filter are written in detail and approximation coefficients, respectively as show in Figure 9.

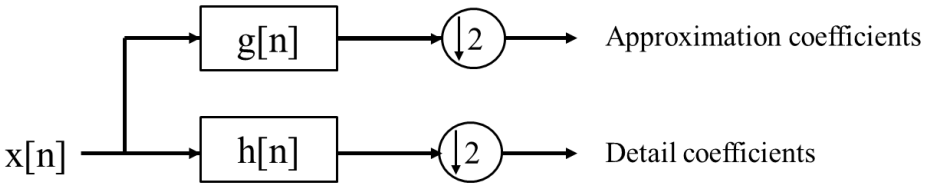


Figure. 9 Block diagram of filter analysis.

From Figure 9, it can be clearly seen that the output sample will become double because of the signal input n sample process with low pass g[n] and high pass h[n] filter. By applying down sampling by 2 to tackle this problem, the output can be express as equation (1)

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k]$$
$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \quad (1)$$

2.3. Lifting structure and integer implementation

For the purpose of studying the wavelet transform in simplest form for better modification [4], we focus on the lifting scheme that utilized in designing wavelet transform as show in Figure. 10.

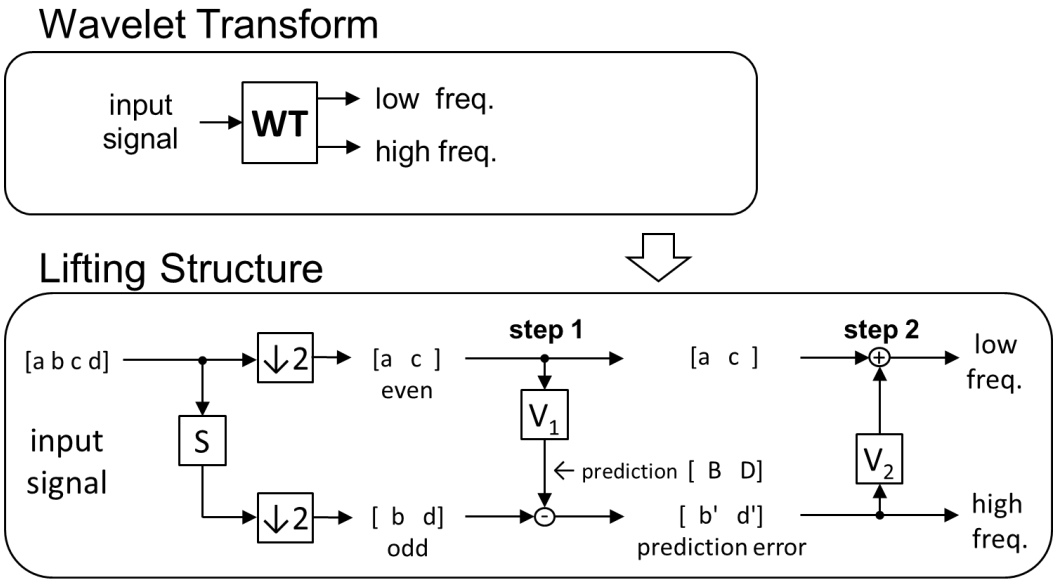


Figure. 10 Wavelet transform in general form and lifting form.

Figure 10 showed the general form and lifting scheme with 2 channels whereas the input signal is divided into low and high frequency. As the input value is image which has a value between 0-255 and it can be treated as an integer value for the encoding part. However as long as the coefficient of every wavelet filter includes the values that are not integer so that the output values of wavelet transform become non-integer result.

For the aforementioned case, the rounding operator is introduced as shown in Figure 11 to this lifting structure in every lifting step for the purpose of transforming

the floating point to integer so that the output signals become integer values. In that case, the perfect reconstruction of lifting structure properties still remains.

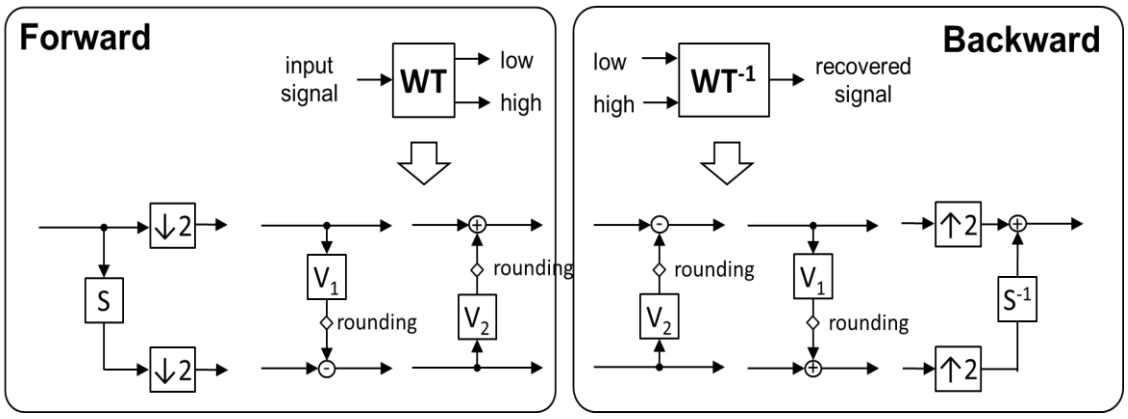


Figure. 11 Rounding operator applied to lifting structure.

However, the rounding noise occurred from the rounding operators in forward transform can be definitely canceled again by the rounding noise in backward transform as shown in Figure 12. Then the output signal which came out from backward transform was same as input signal before coming to forward transform. This was called “perfect reconstruction”

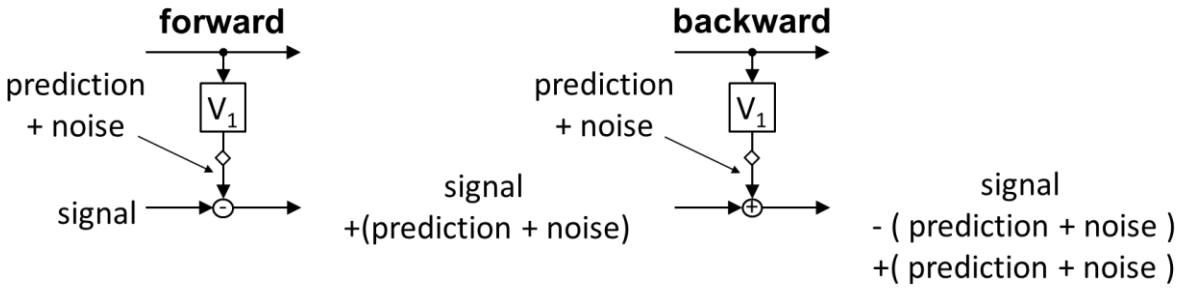
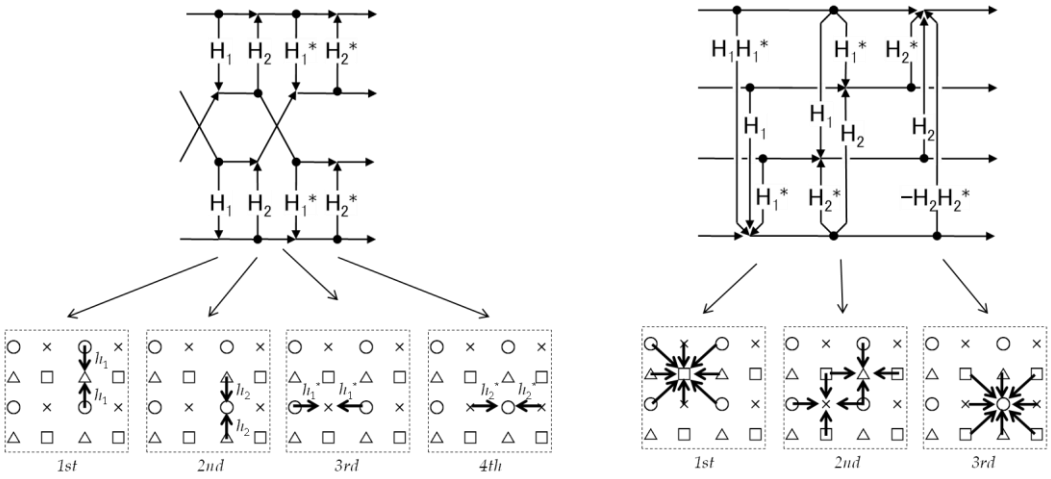


Figure. 12 Noise is canceled in backward transform.

2.4. Separable and Non-Separable



(a) (b)

Figure. 13 Comparison in filter direction between separable (a) and non-separable (b).

Separable and Non-separable are the terms that use to distinguish filters in the lifting structure of wavelet transform. Separable means the filter in lifting structure process one direction per each lifting structure while non-separable filter process two directions per each lifting structure. Examples of 2D separable and 2D non-separable filter are shown in Figure 13.

Figure 13 (a), perform the lifting steps process from neighboring pixels one direction per one lifting steps. Figure 13 (b), perform every direction has process together in one time per lifting steps. And non-separable has reduce lifting step from 4 (100%) to 3 (75%).

2.5. Derivation process

The derivation from separable structure to non-separable can be done by using two basic properties for derivation process, properties I (P_I) and properties II (P_{II}). P_I and P_{II} are expressed as shown in equation (2) and (3), respectively.

$$P_I : \quad \mathbf{Y} = \mathbf{BA} \cdot \mathbf{X} = \mathbf{AC}_0\mathbf{B} \cdot \mathbf{X} \tag{2}$$

$$P_{II} : \quad \mathbf{Y} = \mathbf{AB} \cdot \mathbf{X} = \mathbf{BC}_1\mathbf{A} \cdot \mathbf{X} \tag{3}$$

for

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ A & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathbf{C}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +AB & 0 & 1 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix},. \tag{4}$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & B & 1 \end{bmatrix}, \ \mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -AB & 0 & 1 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \end{bmatrix}. \tag{5}$$

P_I and P_{II} from equation (2) and (3) can be described in lifting structure as shown in Figure 14.

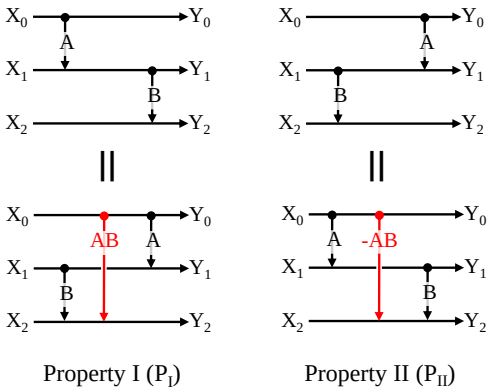


Figure. 14 Basic properties for modification.

2.6. Rounding operator and its errors

Handling the rounding operators and their characteristics plays important role in concern with the wavelet transform. Which exists next to coefficient inside the lifting steps for rounding floating value to integer value with which the rounding errors occurred from each of these rounding operators.

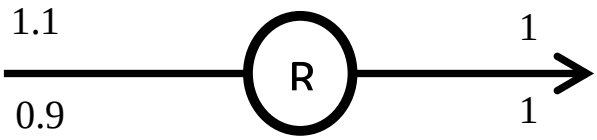
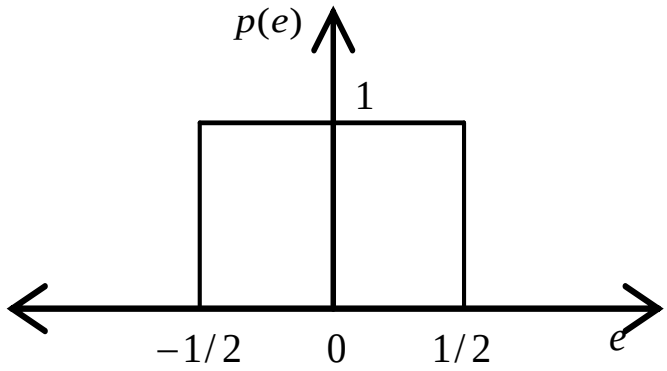


Figure. 15 Rounding process.

Figure 15 shows example of the rounding error from within the range -0.1 and 0.1. Rounding errors can be defined by following equation. The probability density function (PDF) of the rounding error is, as shown in Figure 16.



$$p(e) = \begin{cases} 1, 0.5 \le e \le 0.5 \\ 0, otherwise \end{cases}$$

Figure. 16 PDF of rounding error.

e stands for the rounding error and *p(e)* means PDF of rounding error. Figure 16 shows the values that average of rounding error is 0 and variance of rounding error is 1/12. Variance of rounding error is calculated by equation (6) [5]. $V_e[R]$ is variance of rounding error value.

$$V_e[R] = \int_{-\infty}^{\infty} p(e) \cdot e^2 de \tag{6}$$

2.7. Lagrange multiplier

Lagrange multiplier [6] is used for finding the minimum or maximum of a multivariate function $f(x_1, x_2, ..., x_n)$ subject to the constraint $g(x_1, x_2, ..., x_n) = 0$, where *f* and *g* are continuous function. And $\nabla g \neq 0$ at any point on the curve $g(x_1, x_2, ..., x_n) = 0$ (where ∇ is the gradient).

For the minimum or maximum of *f* on *g*, the gradient of *f* must line up with the gradient of *g*. If the two gradients (*f* and *g*) are in the same direction, then it can be express in equation (7)

$$\nabla f = -\lambda \nabla g \tag{7}$$

The two vectors are equal, so all of their components are similar, giving

$$\frac{\partial f}{\partial x_k} + \lambda \frac{\partial g}{\partial x_k} = 0 \tag{8}$$

Where k = 1,…,n, and the constant λ is called the Lagrange multiplier.

2.8. Basic measurement tools

In this section, evaluation tools and how to comparison between the proposal method and the existing method are explained as shown below.

2.8.1. The-peak-signal-to-noise-ratio (PSNR)

PSNR is a tool which is widely used to measure image reconstructed quality. PSNR is the ratio between the maximum power of a signal and the power of corrupting noise that affects to image. Due to the range is very widespread, so, PSNR is expressed in logarithm term. The power of corrupting noise can be express as mean square error (MSE) in equation (9).

$$MSE = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} [I(i, j) - O(i, j)]^2 \tag{9}$$

I is input image and O is output image, where “i” and “j” is position of image pixel in vertical and horizontal, respectively. The PSNR is determined as

$$PSNR = 10\log_{10}(\frac{Max(I)^2}{MSE}) \text{ [dB]} \quad (10)$$

Max(I) in equation (10) is the maximum pixel value of input image. For example if input image are 8 bit gray scale image then Max(I) is 255.

2.8.2. Rounding Noise

In image compression, lifting scheme of the wavelet transform is for the purpose of separating image input to low and high frequency bands. Lifting structure contains lifting steps and in the each of lifting step, there has been rounding operators which produce rounding errors that effected to the output. The rounding noise in image compression can be evaluated in 2 ways, rounding noise in frequency domain and rounding noise in pixel.

2.8.2.1 Rounding noise in frequency domain

Rounding noise in frequency domain can be illustrated as shown in Figure 17

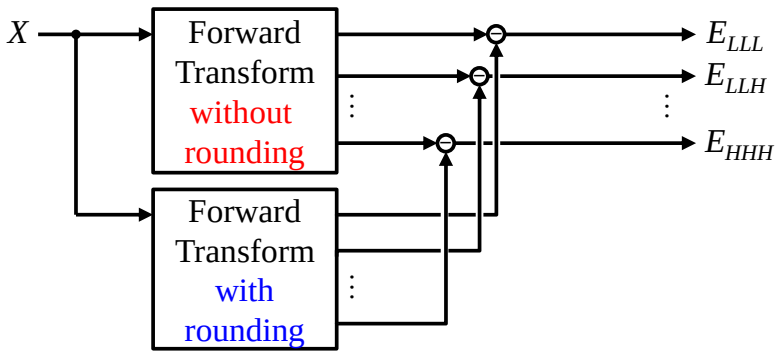


Figure. 17 Rounding noise in frequency domain.

For example, rounding noise in channel LLL can be express as

$$E_{LLL} = FT_{(O)LLL} - FT_{(R)LLL} \tag{11}$$

Where E_{LLL} is rounding noise of channel LLL and $FT_{(O)LLL}$ and $FT_{(R)LLL}$ is forward transform in channel LLL without rounding and with rounding respectively.

2.8.2.2 Rounding noise in pixel domain

In Figure 18, rounding noise in pixel domain can be described as below;

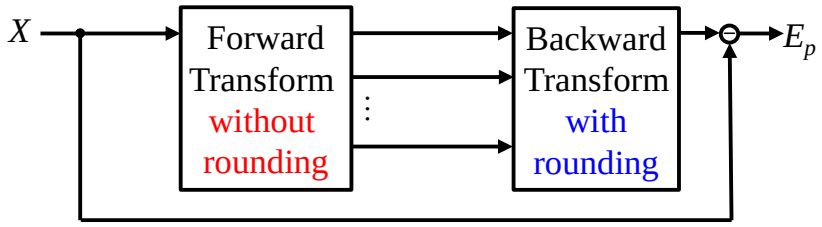


Figure. 18 Rounding noise in pixel domain.

which can be expressed as

$$E = \sum_{j=1}^{m} \sum_{i=1}^{n} [FT_{(O)}(i, j) - BT_{(R)}(i, j)] \tag{12}$$

Where E_p is rounding noise in pixel domain, $FT_{(o)}$ is forward transform without rounding and $FT_{(o)}$ is backward transform with rounding. “i” and “j” is position of image pixel in vertical and horizontal, respectively.

2.8.3. Entropy rate

In the image compression, entropy rate is use for measurement of the randomness [7] which is used to represent symbol of input image. Entropy rate can be defined as minimum bit rate per each symbol in input image which can be express as

$$H = \sum_{i=1}^{n} - p_i \log_2 p_i \tag{13}$$

Where H represents the entropy of grayscale image, P_i is the probability of pixel value at indexes i.

2.8.4. Sum of Power of Two (SPT)

SPT [8] is used to represent decimal number to binary number which is the very basic for digital communication. For example, SPT can show in equation (14) as shown below.

$$(1.101)_2 = \underbrace{1 \cdot 2^0 + 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3}}_{\text{SPT}} \tag{14}$$

1.101 is binary number and term in the right is SPT. In this dissertation, we use the format SPT to represent complexity of coefficient.

2.8.5. Rate distortion curve

The rate distortion curve is the measurement matrix for the comparison between PSNR and bit rate of the reconstructed image for evaluating the coding

performance. For instance, Figure 19 shows an example of rate distortion curve of Separable (Sep) and Non separable (Nsp) at which bit rate higher than 2.1 bpp that coding performance of Nsp is better than Sep.

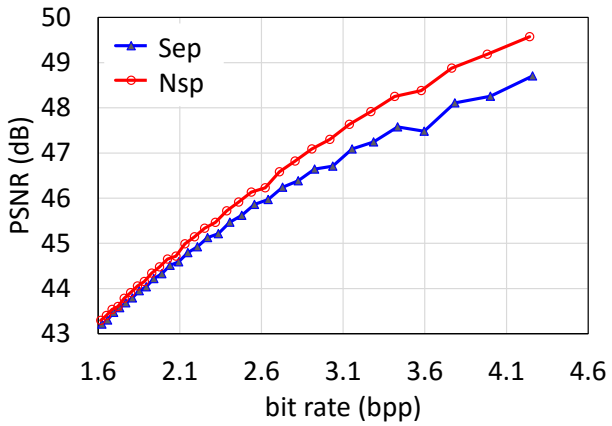


Figure. 19 Example of rate distortion curve.

2.9. Three dimensions input data

Because of the main of this dissertation discussed about 3D wavelet lifting structure. Therefore, 3D input data was used for evaluating the proposal method performance and efficiency. The 3 input data were shown as list below.

- **MRI (Magnetic Resonance Image) data**, resolution 256 x 256, total 8 slices.

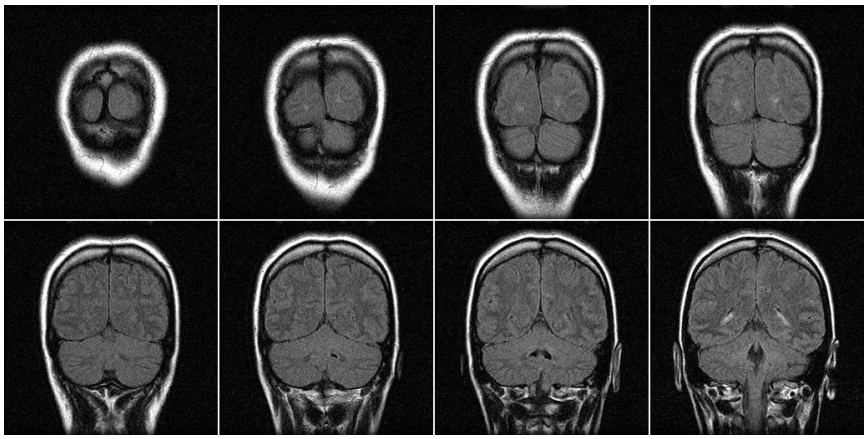


Figure. 20 MRI data input.

- **Random data**, resolution 256 x 256, total 8 slices.

In this dissertation, the stochastic process that we utilize is normal distribution for random numbers. In Figure. 21, the instance of the random data can be expressed as

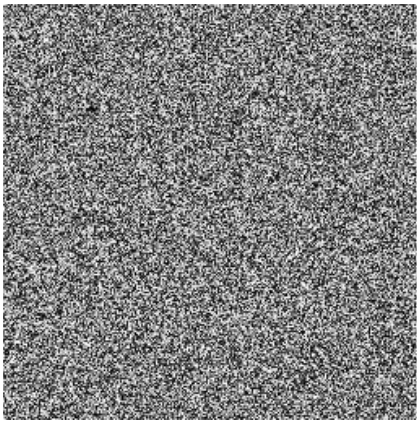


Figure. 21 Example one slide of random data input.

- **AR model data**, resolution 256 x 256, total 8 slices.

The data distribution for the random data input is in uniform distribution still practical image data distribution is commonly in normal or Gaussian distribution.

For that case, the AR model is applied for the transform process of the uniform to normal distribution in the image data. Figure 22 shows Example of AR model data where the model can be expressed as in equation (15).

$$H(z) = \frac{W(z)}{1 - \rho z^{-1}} \tag{15}$$

Where W(z) is white noise or random signal, H(z) is AR model signal and ρ is parameter for control distribution if it is close to 1 then we got normal distribution signal in the output.

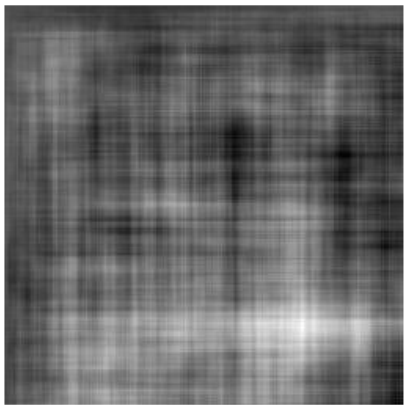


Figure. 22 Example one slide of AR model data input

Chapter 3

3. Non-separable 3D integer wavelet transform for lossless data compression

3.1. Motivation of this proposal

Recently, the image affection has been dramatically increased with respect to pixel resolution, frame rate and dynamic range of pixel values. So that the amount of image data increases as the quality rises. As a result, the image compression techniques have been playing an essential role for storage and communications of digital media data.

The two-dimensional (2D) transform like discrete cosine transform (DCT) and the wavelet transform are mainly stand for still images. Latterly, numerous types of three-dimensional (3D) transforms have been examined for video, hyper spectral images, integral images and medical volumetric data [9-12]. In this chapter, we are dealing with a '3D' wavelet transform together with reduction of the rounding noise within the transform for high exploitation 'lossless' data compression.

Still, there is an extensive development of a category of 'separable'2D wavelet for various applications and, it is also adopted by the JPEG 2000 international standard [13,14] which has been applied in digital cinema [15,16]. On account of the transfer function of the class of transform is constructed with a product of the horizontal 1D transfer function and the vertical one, it can derive legacy of previously designed 1D

structure suitable for hardware implementations [17-19]. Besides, it might have the regularity and low sensitivity to various noises [20, 21]. Nonetheless, most of them are designed for ‘lossy’ coding of images for practically reasonable data compression rate.

While considering the ‘lossless’ coding of images, it is essential that transformed values need to be integers for entropy coding, and also for the case of inverse transform for reconstructing the original integer pixel values without any loss. And there have been an implementation of this type of transform in the lifting structure in the wavelet transform with rounding operations as the integer and, extensions have been investigated based on the international standard [22, 23]. For the reversible color transform and the integer DCT, the radar network has also been utilized [24-30].

Absolutely, those are ‘lossless’ where a set of forward and backward transform assured lossless reconstruction of the initial values of the pixel in integer. Still there is an occurrence of rounding noise’ subsequently the forward transform. For reduction of those noise occurs at each of the lifting steps, a ‘non-separable’ structure was proposed for the condensation of the total number of lifting steps along with the rounding operations [31].

Almost all of these have been initially increase the precision of the prediction in adaptation of the local context of neighboring pixels [32-34].Based on sparse criteria, the other general designing problem has been investigated [35]. Again, there has been a utilization of directionality in a generalized poly-phase representation [36, 37].

Nearly all of these methods were addressed on designing adaptive high pass filters of the wavelet transform.

In [38-41], a new form of non-separable 2D structure has been reported with which transfer function can be expressed as a product of 1D function so that its transform is ‘compatible’ to the separable transform. After all, the implementation of the aforementioned scheme is not a cascade of 1D signal processing in the 1D structure which requires multi-dimensional memory accessing. Yet the mechanism has some drawbacks relating to the unadapting to local context of pixels, it has advantage over reduction of the total number of lifting steps together with the rounding operations.

In this chapter, we examined a ‘3D’ integer wavelet transform for ‘lossless’ coding of 3D signals with the minimization of the rounding noise which is the deviation from the previous studies. A non-separable 3D lifting structure was proposed in the case of reducing the total number of lifting steps [42] without regard to the huge computational complexity for the requirement of 3D memory accessing. On account of video signal processing, it requires inter-frame memory accessing which costs huge memory space. Accordingly, we discuss the non-separable 2D structures to construct the 3D integer wavelet transform as a newly proposed method as in the chapter.

Further, investigation on the performance of the 3D integer wavelet transform was limited to noise variance in frequency domain only along with the limitation of the input signal to only one MRI data set [42]. Regarding this, we first try to clarify variance of the noise in frequency domain and those in pixel domain then we extend our scope on the rate distortion curve in lossy coding mode and the entropy rate in

lossless coding mode by means of computational time of the transforms, and conceive the feature comparison with other methods. Apart from the MRI data set, various kinds of input data are tested in this division.

The proposed wavelet transform discussed here has a merit that its output signals, apart from the rounding noise, are exactly the same as a conventional transform .Based on this compatibility, it becomes possible to utilize legacy of previously designed 1D wavelet transforms. The total amount of the rounding noise is significantly reduced additionally and contributes to increase coding performance of data compression system based on the 3D wavelet transform.

This chapter is organized as in two main sections; section 2 describes the existing method and section 3 for the proposed method from the architectural background to experimental result.

3.2. Existing Method

3.2.1. One-Dimensional (1D) Wavelet Transform

One-dimensional (1D) integer wavelet transform for lossless coding of discrete 1D signal is illustrated in Figure 23. The input signal X is fed into the forward transform at the encoding side and then outputs low frequency band signal Y_L and high frequency band signal Y_H. Each of them is coded with an entropy encoder to generate a bit stream for storage and communications and again the band signals are decoded and inversely transformed at the decoding side.

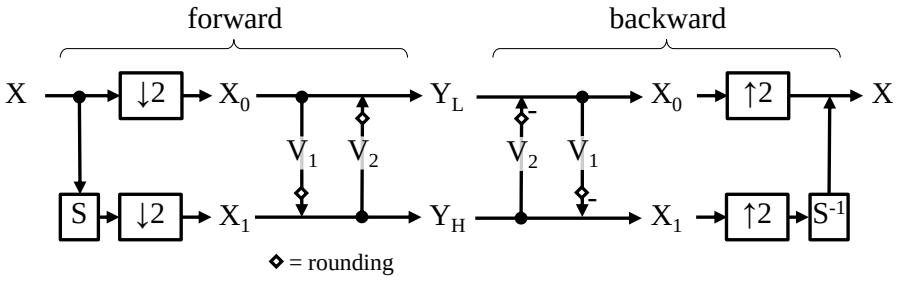


Figure. 23 One-dimensional (1D) integer wavelet transform.

The arrangement can be referred to as the ‘integer’ wavelet transform for ‘lossless’ coding since the original signal *X* can be reconstructed without any loss for the reason that all the signal values inside the transform are rounded to integers, and the rounding noise generated by the rounding operation is cross out at the output of the backward transform.

Precisely, two half-length sequences X_0 and $X_1$are resulted from spilling the original signal *X* given in integer values with length as

$$\begin{cases} x_0(m) = x(2m) \\ x_1(m) = x(2m+1) \end{cases} \tag{16}$$

For*m*=0, 1,⋯, *M*-1 where

$$\begin{cases} X = [\, x(0) \quad x(1) \quad \cdots \quad x(N-1)] \\ X_0 = [x_0(0) \quad x_0(1) \quad \cdots \quad x_0(M-1)] \\ X_1 = [\, x_1(0) \quad x_1(1) \quad \cdots \quad x_1(M-1)] \end{cases} \tag{17}$$

And *M*=*N*/2. In the 1st lifting step, X_1 is predicted from X_0 as

$$y_H(m) = x_1(m) + R[V_1 * x_0(m)] \tag{18}$$

With which R[x] denotes an integer rounded from the original value x, and $*$ denotes convolution which can be defined as

$$V_1 * x_0(m) = h_1\left(x_0(m) + x_0(m+1)\right) \tag{19}$$

for h_1 is a filter coefficient. At 2nd lifting step, X_0 is updated from Y_H as

$$y_L(m) = x_0(m) + R[V_2 * y_H(m)] \tag{20}$$

Where the convolution$V_2*$is defined as

$$V_2 * y_H(m) = h_2\left(y_H(m) + y_H(m-1)\right) \tag{21}$$

where h_2 is a filter coefficient. As the rounding operations $R[\cdot]$ are applied, all the signal values are expressed as integers inside this integer wavelet transform.

In the previous researches, filter coefficients within the lifting wavelet transforms, have been carefully determined. For instance, a set (h_1,h_2)=(-1/2,1/4) is utilized in JPEG 2000 international standard [13,14] where as a low pass filter of the forward transform has unity gain for a constant signal, and zero gain for an alternating signal. Those properties are indicated as

$$\begin{aligned} y_L(m) &= x_0(m) + h_2\left(y_H(m) + y_H(m-1)\right) \\ &= \begin{cases} \text{constant} & for \quad x(n) = \text{constant} \\ 0 & for \quad x(n) = (-1)^n \end{cases} \end{aligned} \tag{22}$$

and a high pass filter has

$$
\begin{aligned}
y_H(m) &= x_1(m) + h_1\left(x_0(m) + x_0(m+1)\right) \\
&= x(2m+1) + h_1\left(x(2m) + x(2m+2)\right) \\
&= 0 \quad for \quad x(n) = \text{constant.}
\end{aligned}
\tag{23}
$$

It is enticing to apply the positive properties like (22) and (23), and other properties such as the vanishing moment in designing a multi-dimensional transform.

3.2.2. Two-Dimensional (2D) Wavelet Transform

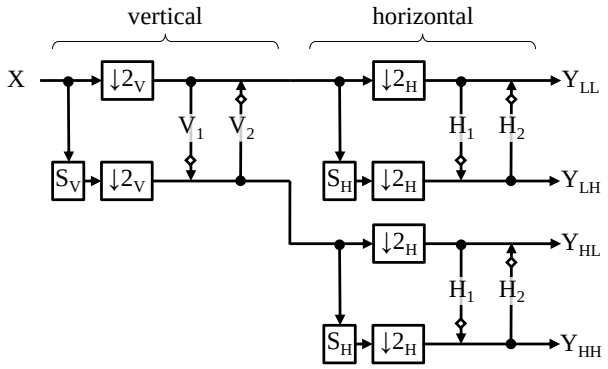


Figure. 24 Two-Dimensional (2D) Wavelet Transform.

A two-dimensional (2D) integer wavelet transform has been illustrated in Figure. 24. Such transform applies the 1D transform vertically to the 2D input image *X* and afterwards, the same 1D transform horizontally. Four frequency band signals Y_{LL}, Y_{LH}, Y_{HL} and Y_{HH} are generated as resultant factors. As it is a product of 1D transforms, it is referred to as the ‘separable’ transform.

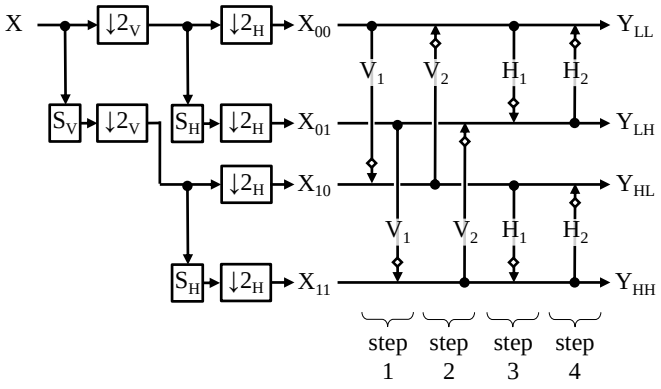


Figure. 25 Separable 2D structure for 2D wavelet transform.

In Figure 25, there is an illustration of the other expression of Figure 24 which outputs the same band signals that of Figure 24. In this, the 2D input signal $x(n_1,n_2)$ is split into 4 groups as

$$\begin{cases} x_{00}(m_1,m_2) = x(2m_1 \quad , 2m_2 \quad) \\ x_{10}(m_1,m_2) = x(2m_1+1, 2m_2 \quad) \\ x_{01}(m_1,m_2) = x(2m_1 \quad , 2m_2+1) \\ x_{11}(m_1,m_2) = x(2m_1+1, 2m_2+1) \end{cases} \tag{24}$$

wherem_1=0, 1,⋯, $M1$-1and m_2=0,1,⋯,M_2-1 for M_1=N_1/2 and M_2=N_2/2. In the 1st lifting step, X_{10} and X_{11} are predicted from X_{00} and X_{01} respectively as

$$\begin{cases} x_{10}^{(1)}(m_1,m_2) = x_{10}(m_1,m_2) + R[V_1 * x_{00}(m_1,m_2)] \\ x_{11}^{(1)}(m_1,m_2) = x_{11}(m_1,m_2) + R[V_1 * x_{01}(m_1,m_2)] \end{cases} \tag{25}$$

where

$$\begin{cases} V_1 * x_{00}(m_1,m_2) = h_1\big(x_{00}(m_1,m_2) + x_{00}(m_1+1,m_2)\big) \\ V_1 * x_{01}(m_1,m_2) = h_1\big(x_{01}(m_1,m_2) + x_{01}(m_1+1,m_2)\big). \end{cases} \tag{26}$$

Along with the variable $m_1$the convolution is performed vertically. Likewise, in the 2nd lifting step, X_{00} and X_{01} are updated from X_{10} and X_{11} respectively as

$$\begin{cases} x_{00}^{(2)}(m_1,m_2) = x_{00}(m_1,m_2) + R[V_2 * x_{10}^{(1)}(m_1,m_2)] \\ x_{01}^{(2)}(m_1,m_2) = x_{01}(m_1,m_2) + R[V_2 * x_{11}^{(1)}(m_1,m_2)] \end{cases} \tag{27}$$

where

$$\begin{cases} V_2 * x_{10}^{(1)}(m_1,m_2) = h_2\left(x_{10}^{(1)}(m_1,m_2) + x_{10}^{(1)}(m_1-1,m_2)\right) \\ V_2 * x_{11}^{(1)}(m_1,m_2) = h_2\left(x_{11}^{(1)}(m_1,m_2) + x_{11}^{(1)}(m_1-1,m_2)\right). \end{cases} \tag{28}$$

Also, the process is repeated horizontally along with the variable m_2 in the 3rd and 4th lifting steps, For example, the horizontal convolutions are given as

$$\begin{cases} H_1 * x_{00}^{(2)}(m_1,m_2) = h_1\left(x_{00}^{(2)}(m_1,m_2) + x_{00}^{(2)}(m_1,m_2+1)\right) \\ H_2 * x_{01}^{(3)}(m_1,m_2) = h_2\left(x_{01}^{(3)}(m_1,m_2) + x_{01}^{(3)}(m_1,m_2-1)\right). \end{cases} \tag{29}$$

It can be clearly seen that this kind of ‘separable’ 2D lifting structure has a total of 8 rounding operations since each of 8 lifting step has one rounding operation. They are the source of the rounding noise to be reduced by introducing a non-separable structure discussed detailed in the chapter.

3.2.3. Non-separable 2D Structure

Figure 26 demonstrates a ‘non-separable’ 2D lifting structure. Unlike the separable architecture, it has multi-input single-output lifting steps. Besides, it has less lifting steps comparing to the separable one. As a result, there have been less rounding

operations as well as rounding noise. While comparing with the Figure 25, the overall lifting steps can be compensated from 4 to 3 (75 %), and the total number of rounding operations is reduced from 8 to 4 (50 %) in Figure 26. Additionally, the total amount of rounding noise [41] is minimized. Note that this transform has the properties of the 1D structure in (22) and (23) since its transfer function is the same as that of the 'separable' transform.

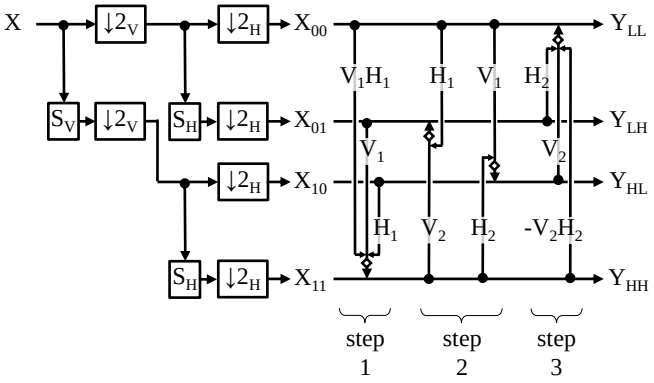


Figure. 26 Non-separable 2D structure for 2D wavelet transform.

Particularly, from multi-inputs X_{00}, X_{01} and X_{10}, the 1st lifting step predicts X_{11} as

$$\begin{aligned} y_{HH}(\mathbf{m}) = x_{11}(\mathbf{m}) + R[\, V_1 * H_1 * x_{00}(\mathbf{m}) \\ + V_1 * x_{01}(\mathbf{m}) + H_1 * x_{10}(\mathbf{m})\,] \end{aligned} \tag{30}$$

for (m)=(m_1, m_2). In this prediction, a diagonal prediction

$$\begin{aligned} V_1 * H_1 * x_{00}(\mathbf{m}) = h_1 h_1 \big(x_{00}(m_1, m_2) + x_{00}(m_1 + 1, m_2) \\ + x_{00}(m_1, m_2 + 1) + x_{00}(m_1 + 1, m_2 + 1) \big) \end{aligned} \tag{31}$$

is utilized. In the 2nd lifting step, two predictions

$$y_{LH}(\mathbf{m}) = x_{01}(\mathbf{m}) + R[\, H_1 * x_{00}(\mathbf{m}) + V_2 * y_{HH}(\mathbf{m})\,] \tag{32}$$

and

$$y_{HL}(\mathbf{m}) = x_{10}(\mathbf{m}) + R[\, V_1 * x_{00}(\mathbf{m}) + H_2 * y_{HH}(\mathbf{m}) \,] \tag{33}$$

are applied whereas the two steps can be executed simultaneously with a parallel processor as there is no need to waiting for the calculation results. Finally, the 3rd lifting step completes the transform as

$$\begin{aligned} y_{LL}(\mathbf{m}) = x_{00}(\mathbf{m}) + R[\, & H_2 * y_{LH}(\mathbf{m}) \\ & + V_2 * y_{HL}(\mathbf{m}) - V_2 * H_2 * y_{HH}(\mathbf{m}) \,]. \end{aligned} \tag{34}$$

Correspondingly, the same band signals generated from this non-separable structure as the separable 2D transform apart from the rounding noise which can be referred to as ‘compatibility’. In the next section, this theory is extended to 3D case.

3.2.4. Separable 3D Structure

In, figure 27 a ‘separable’3D lifting structure is illustrated. Single-input single-output lifting steps are applied to generate 8 frequency band signals group where they are resulting from the splitting of a 3D input signal *X*. This structure has 6 lifting steps and 24 rounding operations which are reduced by the proposed method in this chapter.

In detail, the 3D input signal $x(n_1,n_2,n_3)$ is split into 8groups as

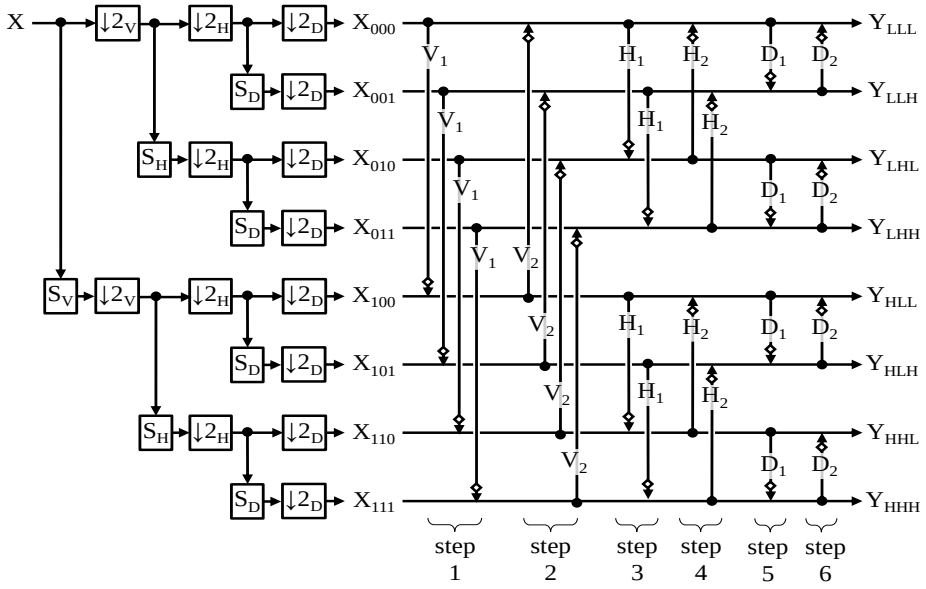


Figure. 27 Separable 3D structure ‘Sep3D’ for 3D wavelet transform.

$$\begin{cases} x_{000}(\mathbf{m}) = x(2m_1 \quad , 2m_2 \quad , 2m_3 \quad) \\ x_{001}(\mathbf{m}) = x(2m_1 \quad , 2m_2 \quad , 2m_3+1) \\ x_{010}(\mathbf{m}) = x(2m_1 \quad , 2m_2+1, 2m_3 \quad) \\ x_{011}(\mathbf{m}) = x(2m_1 \quad , 2m_2+1, 2m_3+1) \\ x_{100}(\mathbf{m}) = x(2m_1+1, 2m_2 \quad , 2m_3 \quad) \\ x_{101}(\mathbf{m}) = x(2m_1+1, 2m_2 \quad , 2m_3+1) \\ x_{110}(\mathbf{m}) = x(2m_1+1, 2m_2+1, 2m_3 \quad) \\ x_{111}(\mathbf{m}) = x(2m_1+1, 2m_2+1, 2m_3+1) \end{cases} \tag{35}$$

for (m)=(m_1,m_2,m_3).In the 1st lifting step, the vertical prediction is applied as

$$\begin{bmatrix} x_{100}^{(1)}(\mathbf{m}) \\ x_{101}^{(1)}(\mathbf{m}) \\ x_{110}^{(1)}(\mathbf{m}) \\ x_{111}^{(1)}(\mathbf{m}) \end{bmatrix} = \begin{bmatrix} x_{100}(\mathbf{m}) \\ x_{101}(\mathbf{m}) \\ x_{110}(\mathbf{m}) \\ x_{111}(\mathbf{m}) \end{bmatrix} + \begin{bmatrix} R[V_1 * x_{000}(\mathbf{m})] \\ R[V_1 * x_{001}(\mathbf{m})] \\ R[V_1 * x_{010}(\mathbf{m})] \\ R[V_1 * x_{011}(\mathbf{m})] \end{bmatrix} \tag{36}$$

where ‘V_1*’ denotes the vertical convolution along with the variable m_1 similar to (26). After the updating

$$\begin{bmatrix} x_{000}^{(2)}(\mathbf{m}) \\ x_{001}^{(2)}(\mathbf{m}) \\ x_{010}^{(2)}(\mathbf{m}) \\ x_{011}^{(2)}(\mathbf{m}) \end{bmatrix} = \begin{bmatrix} x_{000}(\mathbf{m}) \\ x_{001}(\mathbf{m}) \\ x_{010}(\mathbf{m}) \\ x_{011}(\mathbf{m}) \end{bmatrix} + \begin{bmatrix} R[V_2 * x_{100}^{(1)}(\mathbf{m})] \\ R[V_2 * x_{101}^{(1)}(\mathbf{m})] \\ R[V_2 * x_{110}^{(1)}(\mathbf{m})] \\ R[V_2 * x_{111}^{(1)}(\mathbf{m})] \end{bmatrix} \tag{37}$$

in the 2nd lifting step, the horizontal prediction

$$\begin{bmatrix} x_{010}^{(3)}(\mathbf{m}) \\ x_{011}^{(3)}(\mathbf{m}) \\ x_{000}^{(4)}(\mathbf{m}) \\ x_{001}^{(4)}(\mathbf{m}) \end{bmatrix} = \begin{bmatrix} x_{010}^{(2)}(\mathbf{m}) \\ x_{011}^{(2)}(\mathbf{m}) \\ x_{000}^{(2)}(\mathbf{m}) \\ x_{001}^{(2)}(\mathbf{m}) \end{bmatrix} + \begin{bmatrix} R[H_1 * x_{000}^{(2)}(\mathbf{m})] \\ R[H_1 * x_{001}^{(2)}(\mathbf{m})] \\ R[H_2 * x_{010}^{(3)}(\mathbf{m})] \\ R[H_2 * x_{011}^{(3)}(\mathbf{m})] \end{bmatrix} \tag{38}$$

is applied at‘H_1*’ and ‘H_2*’ denote the horizontal convolution along with the variable m_2 similar to (29). In the 5th and 6th lifting steps, the convolution ‘D_1*’ and ‘D_2*’ are performed along with the variable m_3.

As per this separable structure, the *N*th lifting step must wait for calculation results of its previous *N*-1th lifting step which causes delay from the input to the output steps. First the lifting steps are reduced by the initiating of the non-separable structures composed of multi-dimensional memory accessing.

3.3. Proposed Method

3.3.1. Symmetrically derive to Non-separable 3D Structure

Based on the two basic properties which is described in chapter 2, the 5/3 separable 3D structure in figure 27 can be symmetrically derived to minimum lifting structure or non-separable 3D structure in figure 39 as follow:

First, lifting in the same direction (figure 28) were moved together in the nearby position (figure 29)

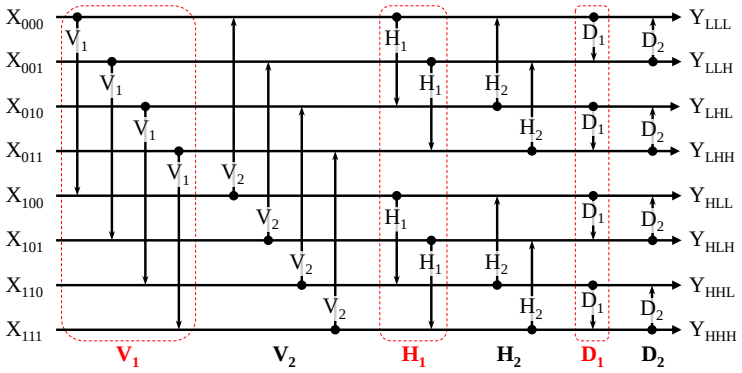


Figure. 28 Before moving V_1, H_1 and D_1.

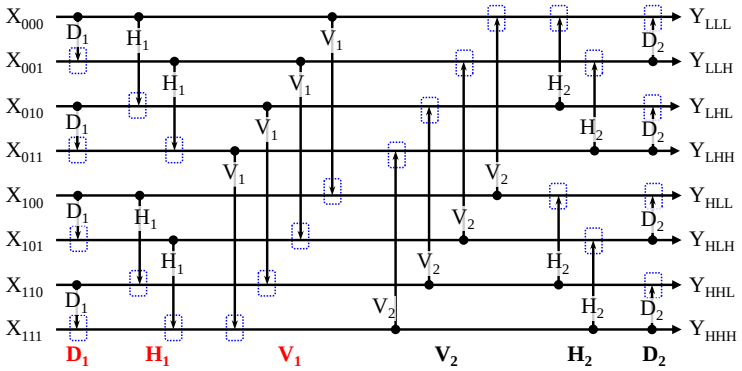


Figure. 29 After moved V_1, H_1 and D_1 (24 rounding operators)

Second, by using the basic properties P_I and P_{II}, and we move forward the lifting H_1 V_1 and H_2 D_2 in figure 30 and then backward. Finally it can be expressed lifting scheme which is shown in figure 31.

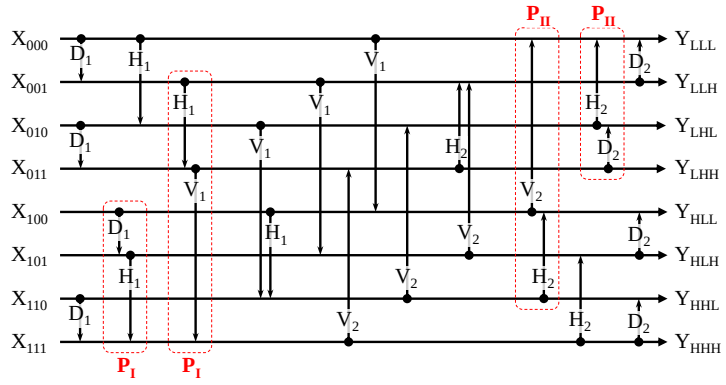


Figure. 30 Before move in second step.

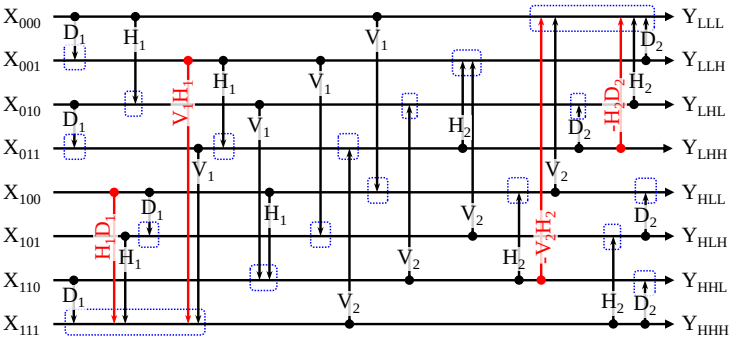


Figure. 31 After move in second step (18 rounding operators)

Third, shifting V_1H_1, V_1 and the other V_1 backward based on P_I and V_2 another V_2 and $-V_2H_2$ forward by P_{II}.

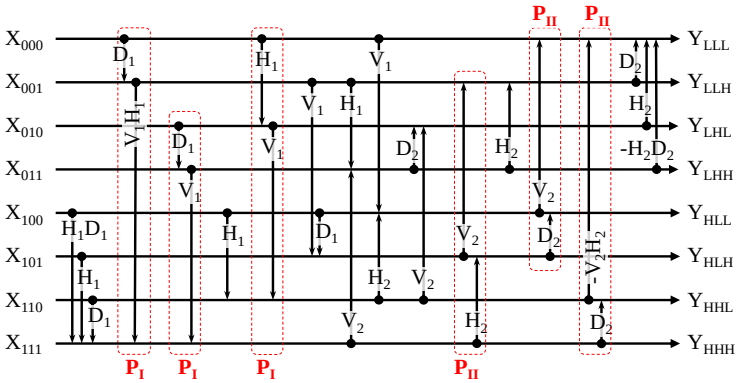


Figure. 32 Before move in third step.

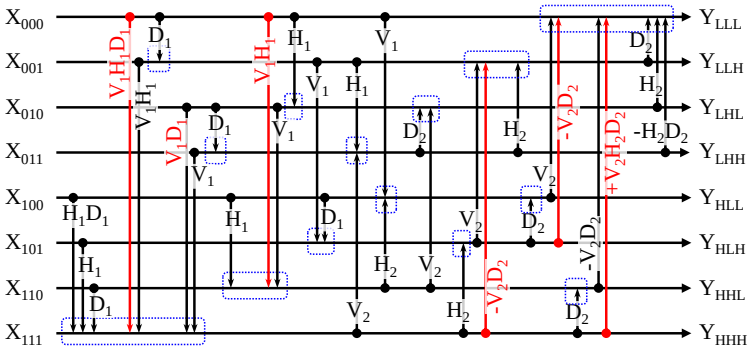


Figure. 33 After move in third step (14 rounding operators)

In the fourth stage, shift D_1 forward by P_I, D_2 backward by P_{II}.

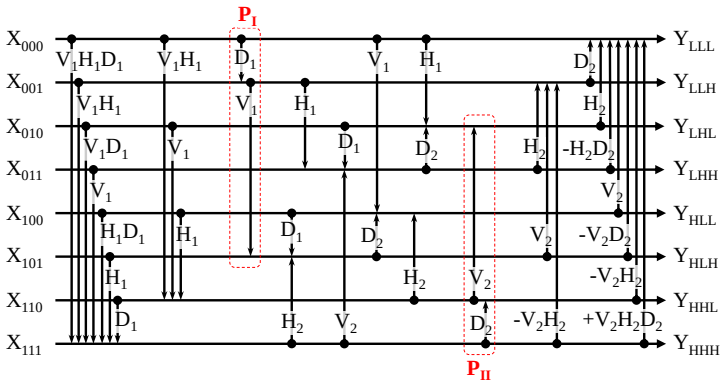


Figure. 34 Before moving in forth step.

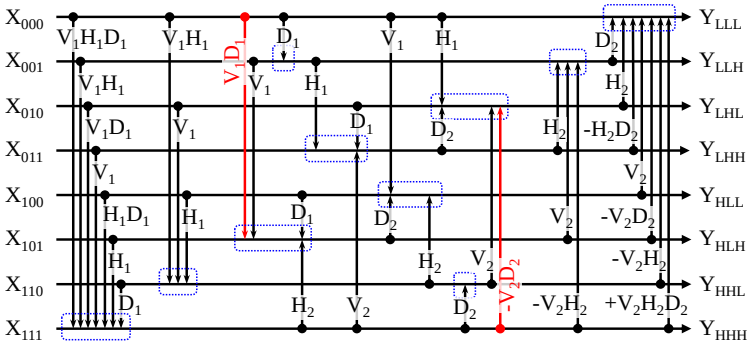


Figure. 35 After move in forth step (10 rounding operators).

Fifth, shift D_1 forward again by P_I after that D_2 backward by P_{II}.

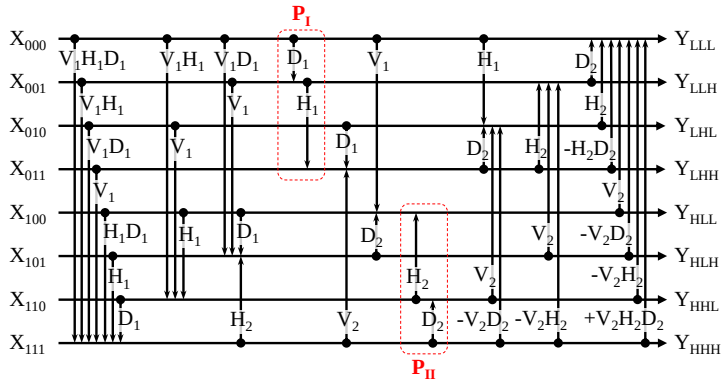


Figure. 36 Before moving in fifth step

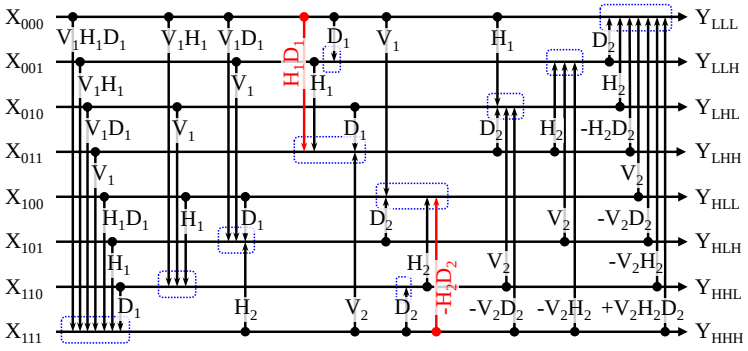


Figure. 37 After moving in fifth step (10 rounding operators)

Sixth, brings D_1 to the lifting H_2 V_2 and $-V_2H_2$ and then bring D_2 to V_1, H_1 and V_1H_1. Eventually, we can get the non-separable lifting structure in Figure. 39.

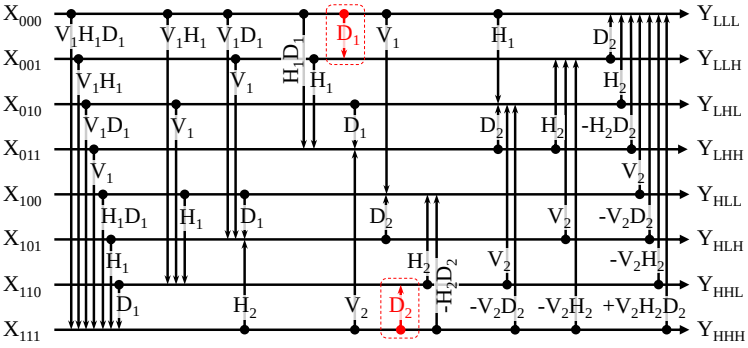


Figure. 38 D_1 and D_2 before moved.

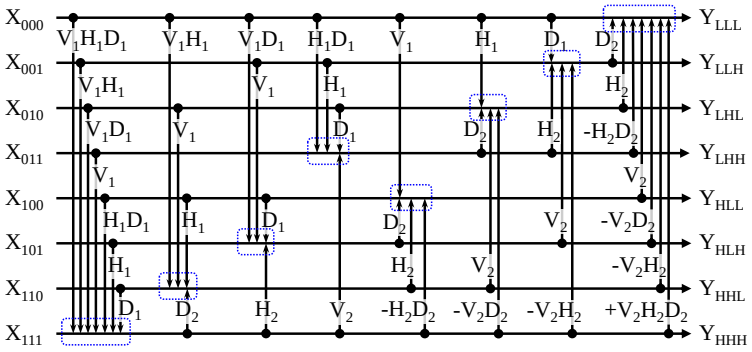


Figure. 39 After moving D_1 and D_2 we got Non separable structure (8 rounding operators).

3.3.2. Non-separable 3D Structure

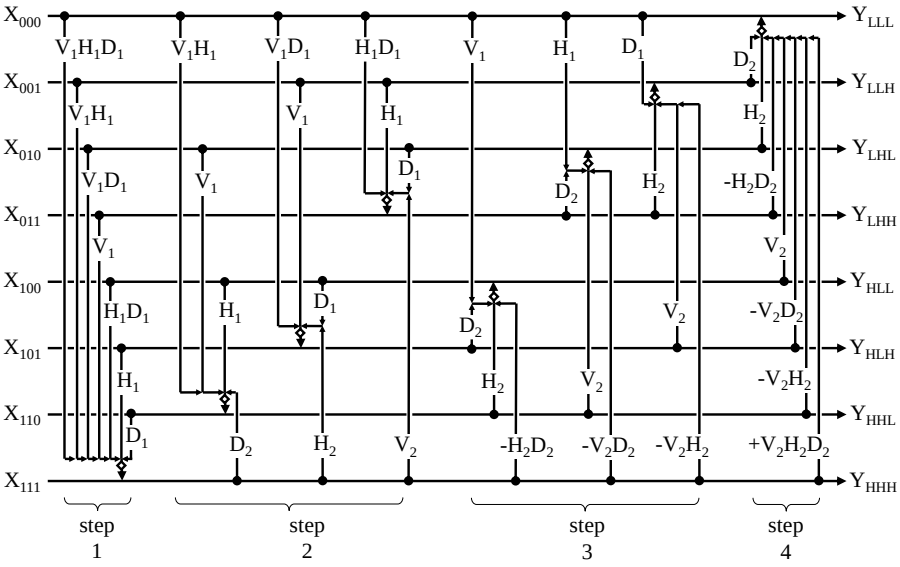


Figure. 40 Non-separable 3D structure ‘Ns3D’ for 3D wavelet transform.

Figure 40 illustrates the ‘non-separable’3D lifting structure in [42]. In this structure, multi-input single-output lifting steps are introduced. It means 3D memory accessing. Comparing to the existing separable structure, the total number of lifting steps is reduced from 6 to 4 (66.7 %). Therefore it is expected to reduce the delay due

to waiting for results of the previous lifting step utilizing parallel signal processing. The total number of rounding operations is also decreased from 24 to 8 (33.3 %). We experimentally demonstrate that it contributes to reducing total amount of the rounding noise in 5.

In detail, the 1st stage predicts X_{111} from all of the remaining groups as

$$\begin{aligned} x_{111}^{(1)}(\mathbf{m}) = x_{111}(\mathbf{m}) + R[&V_1 * H_1 * D_1 * x_{000}(\mathbf{m}) \\ &+V_1 * H_1 * x_{001}(\mathbf{m}) \\ &+V_1 * D_1 * x_{010}(\mathbf{m}) \\ &+V_1 * x_{011}(\mathbf{m}) \\ &+H_1 * D_1 * x_{100}(\mathbf{m}) \\ &+H_1 * x_{101}(\mathbf{m}) \\ &+D_1 * x_{110}(\mathbf{m})]. \end{aligned} \tag{39}$$

where there is a convolution with multi-dimensional memory accessing is included. The convolution '$V_1 * H_1 * D_1 *$' is defined for instance as

$$\begin{aligned} V_1 * H_1 * D_1 * x_{000}(\mathbf{m}) = h_1^3 \{ & \\ x_{000}(m_1, m_2, m_3) + x_{000}(m_1^+, m_2, m_3) & \\ + x_{000}(m_1, m_2^+, m_3) + x_{000}(m_1^+, m_2^+, m_3) & \\ + x_{000}(m_1, m_2, m_3^+) + x_{000}(m_1^+, m_2, m_3^+) & \\ + x_{000}(m_1, m_2^+, m_3^+) + x_{000}(m_1^+, m_2^+, m_3^+) \} & \end{aligned} \tag{40}$$

where

$$m_1^+ = m_1 + 1, \quad m_2^+ = m_2 + 1, \quad m_3^+ = m_3 + 1. \tag{41}$$

Then there are three predictions in the 2nd lifting step,

$$x_{110}^{(2)}(\mathbf{m}) = x_{110}(\mathbf{m}) + R[V_1 * H_1 * x_{000}(\mathbf{m}) + V_1 * x_{010}(\mathbf{m}) + H_1 * x_{100}(\mathbf{m}) + D_2 * x_{111}^{(1)}(\mathbf{m})] \quad (42)$$

$$x_{101}^{(2)}(\mathbf{m}) = x_{101}(\mathbf{m}) + R[V_1 * D_1 * x_{000}(\mathbf{m}) + V_1 * x_{001}(\mathbf{m}) + D_1 * x_{100}(\mathbf{m}) + H_2 * x_{111}^{(1)}(\mathbf{m})] \quad (43)$$

$$x_{011}^{(2)}(\mathbf{m}) = x_{011}(\mathbf{m}) + R[H_1 * D_1 * x_{000}(\mathbf{m}) + H_1 * x_{001}(\mathbf{m}) + D_1 * x_{010}(\mathbf{m}) + V_2 * x_{111}^{(1)}(\mathbf{m})] \quad (44)$$

Each of these steps can be done simultaneously on a parallel signal processing platform for there is no need to for calculation results from the particular step. Updating in the 3rd and 4th lifting steps is performed as the same manner.

It can be observed that there is no difference between the separable structure in Figure 27 and the non-separable structure in figure 40 in respect of signals. We can denote that both these methods are expressed with the same transfer function which is a product of 1D function. Based on that particular fact, the transform to be implemented is ‘separable ’yet in respect of noise, those are different. The structure in figure 40 is expected to have less rounding noise as it has less rounding operations.

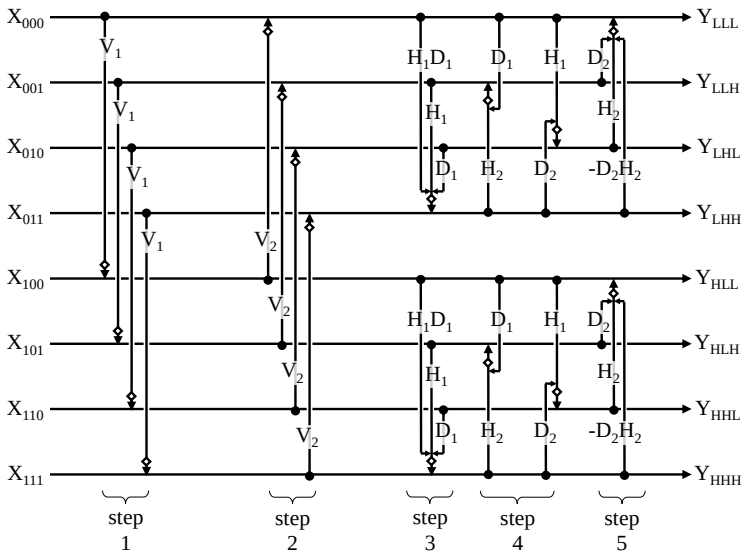


Figure. 41 Non-separable 2D structure ‘Ns2D(1)’ for 3D wavelet transform

A newly introduced structure ‘Ns2D (1)’ for the 3D lifting wavelet transform is illustrated Figure 41. In this, although the 3rd, 4th, 5th 6th lifting steps are implemented as the non-separable ‘2D’ structure, the 1st and the 2nd lifting steps are the same as those of the ‘separable’ structure in Figure 27. That is in contribution for the reduction of hardware complexity comparing to ‘Ns3D’ in Figure 40 for it does not utilize the 3D memory accessing.

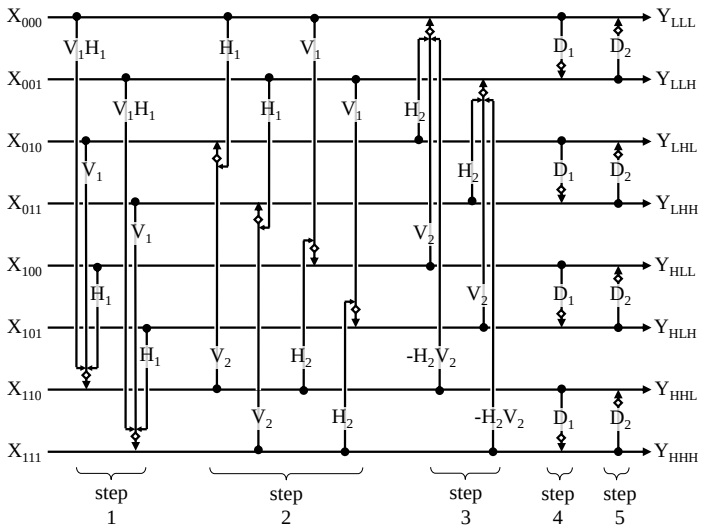


Figure. 42 Non-separable 2D structure ‘Ns2D(2)’ for 3D wavelet transform.

Figure 42 illustrates another variation ‘Ns2D (2)’ for the 3D transform. Similar to the scheme in Figure 41, yet the 4th and the 5th lifting steps are the same as those of the ‘separable’ structure in figure 27, the 1st, the 2nd and the 3rd lifting steps are implemented as the non-separable ‘2D’ structure.

In brief, the structures discussed in this chapter are expressed as

$$\begin{cases} Sep3D & V_1V_2H_1H_2D_1D_2 & (Fig.5) \\ Ns3D & [V_1V_2H_1H_2D_1D_2] & (Fig.6) \\ Ns2D(1) & V_1V_2(H_1H_2D_1D_2) & (Fig.7) \\ Ns2D(2) & (V_1V_2H_1H_2)D_1D_2 & (Fig.8) \end{cases} \tag{45}$$

where [] in ‘Ns3D’ and () in ‘Ns2D’ denote the non-separable 3D structure and the non-separable 2D structure.

3.3.3. Comparison of Structures

The summary the variations of the structure of the 3D integer wavelet transform is described in Table 1. In the newly introduced structures ‘Ns2D(1) and ‘Ns2D(2)’ there is no utilization of inter-frame memory accessing which requires huge memory space which is different from the ‘Ns3D’ and in place of the total number of lifting steps is increased comparing to ‘Ns3D’. The number of rounding operations is in the middle between ‘Sep3D’ and ‘Ns3D’. Here, we compare the variations in (42) for various input signals from various aspects in 5.

Table. 1 Comparison of the structures**.**

	Sep3D	**Ns3D**	**Ns2D(1)**	**Ns2D(2)**
Rounding operations	24	8	16	16
Lifting steps	6	4	5	5
Memory accessing	1D	3D	2D	2D

3.4. Experimental Result

A set of MRI volumetric data in Figure 43 provided by MATLAB is tested in the following experiments while each of frame has 256×256 pixels and each pixel is expressed with an 8 bit depth integer. The results of applying the 3D wavelet transform to the 8 frames of the MRI data are illustrated in figure 44. In that, each frequency band signals are normalized to the range of [0,255] for display purpose in this figure

In this proposed method, there is no limitation to the input signal to the MRI data like [42], and a random 3D input signal is included. Moreover, the 3D auto-regressive model can be demonstrated as

$$\begin{cases} x\ (n_1,n_2,n_3) = x'\ (n_1,n_2,n_3) + \rho \cdot x\ (n_1,n_2,n_3-1) \\ x'\ (n_1,n_2,n_3) = x''(n_1,n_2,n_3) + \rho \cdot x'\ (n_1,n_2-1,n_3) \\ x''(n_1,n_2,n_3) = w\ (n_1,n_2,n_3) + \rho \cdot x''(n_1-1,n_2,n_3) \end{cases} \tag{46}$$

is included into our experiments. Note that ρ is set to 0.9 in the experiments in this chapter.

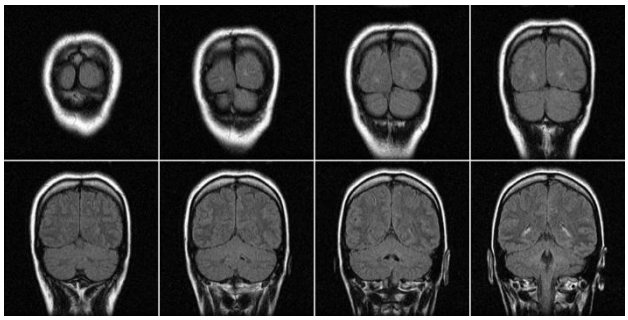


Figure. 43 Tested data set ‘MRI’.

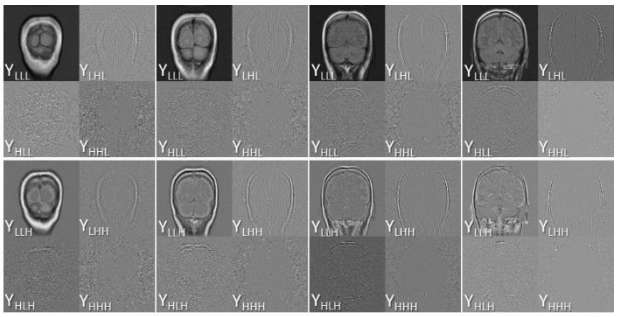
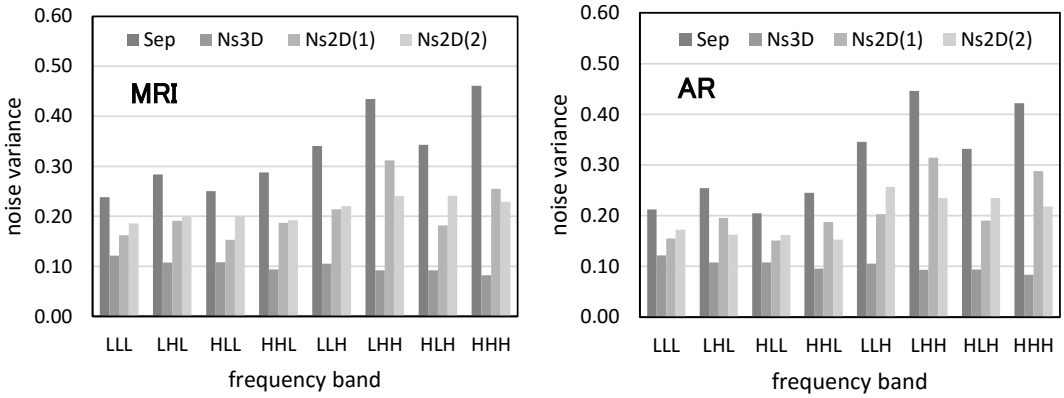


Figure. 44 Results of the 3D wavelet transform.

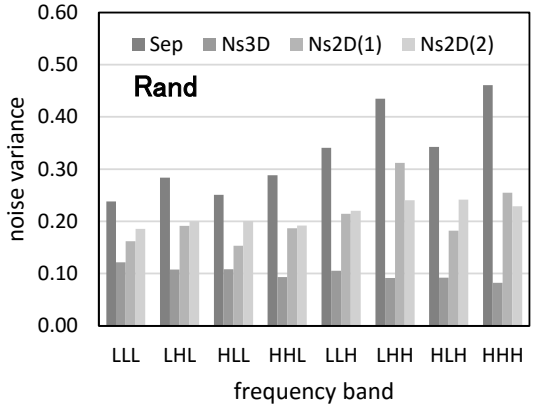
3.4.1. Evaluation of Rounding Noise

First of all, the methods summarized in table 1 are compared in respect of the rounding noise. Figure 45 (a), (b) and (c) show the variance of the rounding noise in each frequency bands for MRI data, AR model, and random input, respectively. In all of these cases, variance of the noise in ‘HHH’ band of ‘Ns3D’ is 0.08 due to the fact that the noise is not amplified. As the value of the rounding noise is within the range of [-0.5, 0.5], its variance is 1/12=0.08 just after each of the rounding operation.



(a)

(b)



(c)

Figure. 45 Rounding noise in each frequency band.

Vice versa, the noise variance in ‘LLL’ band of ‘Ns3D’ is greater than 0.08 as a result of multiple noise amplified though convolutions are summed up in Y_{LLL}. In ‘LHH’ band and ‘HHH’ band of ‘Sep’, the highest noise variance is observed. The variance of the noise is amplified approximately 0.45/0.08=5.6 times in those frequency bands in the separable structure in Figure 27.

The averaged variance over all frequency bands in peak-signal to noise ratio (PSNR) in figure 46 defined as

$$PSNR = 10\log_{10}\frac{255^2}{\sigma^2} \quad (dB) \tag{47}$$

Where σ^2 stands for variance of the rounding error and also observing that ‘Ns3D’ and ‘Ns2D’ increase PSNR by 5 (dB) and 2 (dB) respectively comparing to ‘Sep’.

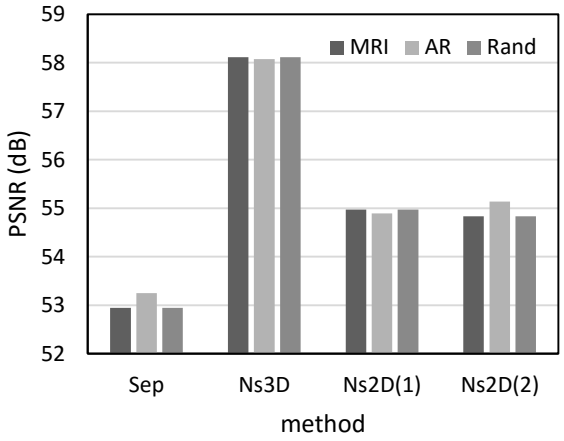


Figure. 46 Rounding noise in frequency domain.

The variance of the rounding noise measured in pixel domain is described Figure 47. Without the rounding operation, the input signal is transformed and unlikely transformed backward with the rounding operations. Similarly to Figure 47, ‘Ns3D’ is the best and ‘Ns2D’ is the second. There is no significant difference among input signals.

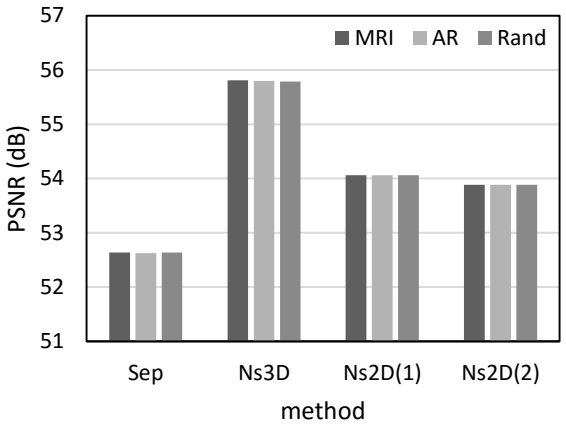


Figure. 47 Rounding noise in pixel domain.

3.4.2. Evaluation of Coding Performance

As per the second evaluation, the entropy rate is examined for the evaluation of the lossless coding performance. The comparison of the structures in terms of lossless data compression rate for various bit depth of the input signal 'MRI' can be seen in figure 45. The compression rate is defined as

$$\eta = \frac{H_{Str}}{H_{Sep}} \times 100 \quad (\%) \tag{48}$$

for $Str \in \{\text{Ns3D, Ns2D(1), Ns2D(2)}\}$ where H denotes the entropy rate in bit per pixel (bpp) averaged over all frequency band signals which approximates the average code length of the compressed data volume of the transformed signal.

As to the Figure 48 we observed that there is no significant difference between the three methods for the original 8 bit depth signal. Although the bit depth of the input signal is reduced to 4 (bit), the ratio of the entropy rate is decreased to 91.9 (%) and

96.2 (%) by ‘Ns3D’ and ‘Ns2D(1)’ respectively. It might be noted that the bit depth of pixel values in the range of [0, 2^B-1] is defined as B (bit). To summarize, performance in lossless coding mode is improved by 8.1 (%) and 3.8 (%) by ‘Ns3D’ and ‘Ns2D(1)’ respectively for 4 bit depth ‘MRI’ volumetric data.

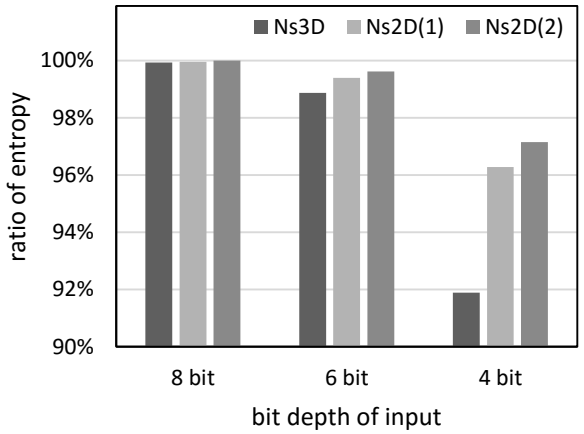


Figure. 48 Performance in lossless coding mode.

In Figure 49, rate-distortion curve of ‘Sep’ and that of ‘Ns3D’ is indicated where the horizontal axis indicates the entropy rate and the vertical axis PSNR of the reconstructed signal, respectively. The quantization is originated just after the forward transform and these curves indicate performance of the methods in lossy coding mode while observing that ‘Ns3D’ is higher than ‘Sep’ by approximately 1.5 (dB) at 6.5 (bpp). It can be summarized that superiority of the non-separable structure over the conventional separable structure was observed.

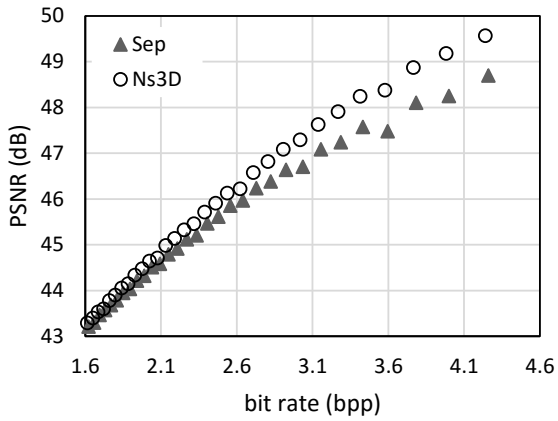


Figure. 49 Performance in lossy coding mode.

3.4.3. Evaluation of Computational Time

From the view point of computational time of the structures for the 3D transform was investigated as the third evaluation. The computational time summary of 'Sep3D', 'Ns3D', 'Ns2D (1)' and 'Ns2D(2)' is described in Table 2 . The performance of these algorithms were executed with MATLAB program on Dell Inspiron 580 PC, Intel Core i3-550 Processor 3.2 GHz, 4GB RAM on Windows 7.

As stated in the table 2, it was observed that 'Ns3D' is the fastest. Both of 'Ns2D(1)' and 'Ns2D(2)' are almost the same, and the second in computational speed which are still faster than 'Sep3D' by 19 (ms). From these results, it can be confirmed that less lifting steps mean faster computation of the transform.

Table. 2 Comparison of computational time.

	Sep3D	**Ns3D**	**Ns2D(1)**	**Ns2D(2)**
Computational time	120.7 ms	97.9 ms	101.3 ms	101.4 ms

3.5. Summary and Discussion on this Proposal

This chapter discussed the 3D integer lifting wavelet transform with reduced amount of rounding noise. In the chapter, we introduce Non-separable multi-dimensional lifting structures for the purpose of the minimization of the total number of lifting steps as well as the rounding operations. And the proposed method is experimentally confirmed that the non-separable 3D structure increases PSNR by 5 (dB) in frequency for MRI, AR model and random input signals. Also for the case the non-separable 2D structure, it increases PSNR by 2 (dB) without using inter-frame memory accessing. These improvements in PSNR measures were also observed in rate-distortion curves at high bit rates in lossy coding mode. In lossless coding mode, data compression performance is improved for low bit depth input images. Computational time was also improved by the method outline detailed in the chapter.

Chapter 4

4. Integer Implementation of 3D Wavelet Transform for Lossy Data Compression

4.1. Motivation of this proposal

As mentioned in the previous chapter, by virtue of that the wavelet transform [13, 14] is utilized in JPEG 2000 international standard and implementing upon digital cinema video signals compression [15, 16] wavelet transforms have been attracting researchers' attention. As for the previous studies, there have been addressing on two main areas, one for the (5/3) type transform for lossless coding of images and the other (9/7) type transform for lossy coding where the former one is based on the lifting structure. Based on the symmetric property of the forward and backward transform that makes the rounding noise of integer within the transform to disappear, it is possible to decode the original integer input signal without any loss. For that reason, (5/3) type transform has been used for lossless coding of images. But it also becomes possible to perform lossy coding in that transform by originating quantization between the forward transform and entropy coding. However its performance is inferior to the (9/7) type transform in lossy coding mode.

This chapter concentrated on the lossy coding of 3D signals and, discussed on the analysis of the (9/7) type transform based on quadruple lifting steps implementation with two aspects. One is delay from input to output in such case a particular lifting

step could be processed only after getting the calculation result of its previous lifting step, there is a long delay among many sequential lifting steps from input to output. And the next aspect is regarding with the integer implementation of the transform for the signal values inside the transform are assumed to be real numbers as an ideal case. Per contra, those are rounded into finite word length rational numbers. Such, the shorter the word length, the less computational load and the more rounding noise might be resulted. The trade-off made the contribution upon memory space reduction that challenges with the rounding noise [44] which is independent of coding efficiency. However, the quantization noise is closely related to performance in lossy coding.

Relating to the first aspect, studies were made upon a non-separable 2D structure introduced to (5/3) type 2D transform [31,45] which was then extended to be adaptive to local property of pixels to increase precision of the prediction [33-35]. The extended researches ware based on minimization of rounding noise in integer implementation of the (5/3) type 2D transform and for the 3D case in [42]. Yet from the (9/7) type transform perspective, it is not the case. Despite the introduction of non-separable structures surely reduces total number of lifting steps [36-37], total amount of the rounding noise was not always reduced [40,41] and the issues can still find in 3D case (existing method) [46].

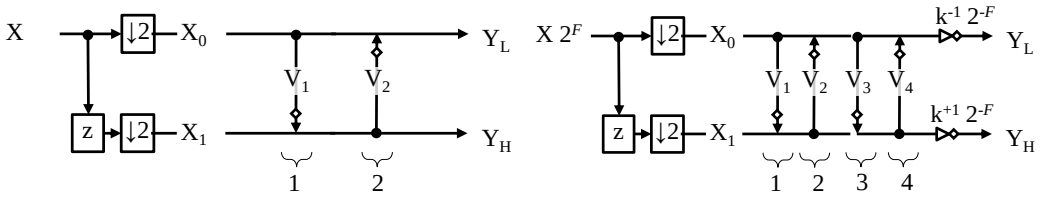
Non-separable 2D structure was introduced instead of the 3D structure in the existing method, whereas the order of the lifting steps in the original separable 3D transform is preserved. It was investigated by observing all possibilities combination which was determined by 6 rules. Accordingly, the total amount of the rounding noise

is reduced, although it raises the total number of the rounding operations. Here we experimentally confirmed significant reduction of that total amount of the rounding noise observed in pixel values of the decoded image. Upper bound of quality of decoded images in lossy coding mode is also improved.

4.2. Existing Method

4.2.1. Two Types of Wavelet Transform

A forward transform of (5/3) type wavelet transform developed for ‘lossless’ coding of 1D signals for JPEG 2000 standard [13] is illustrated in figure 50 (a) which compose of two lifting steps. In figure 50 (b) , a (9/7) type transform developed for ‘lossy’ coding is describe with which two more lifting steps and scaling with a constant *k* are added.



(a) (5/3) type for lossless coding (b) (9/7) type for lossy coding

Figure. 50 Two types of wavelet transform.

The (5/3) type transform decomposes a 1D input signal *X* into two frequency band signals Y_L and Y_H with two lifting steps where 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

$$X_c(z) = \downarrow 2[z^c X(z)], \quad c \in \{0,1\} \tag{49}$$

for

$$\downarrow 2[X(z)] = \frac{1}{Q}\sum_{p=0}^{Q-1} X(z^{1/Q} \cdot W_Q^p), \quad W_Q = e^{j2\pi/Q} \tag{50}$$

Where Q=2 and

$$X(z) = \sum_{n=0}^{N-1} x(n) z^{-n} . \tag{51}$$

Then, the 1st lifting step is applied as

$$X_1^{(1)}(z) = X_1(z) + R[V_1(z) X_0(z)] \tag{52}$$

and the 2nd lifting is applied as

$$X_0^{(2)}(z) = X_0(z) + R[V_2(z) X_1^{(1)}(z)] \tag{53}$$

where $V_1(z)$ and $V_2(z)$ are filters given as

$$\begin{bmatrix} V_1(z) \\ V_2(z) \end{bmatrix} = \begin{bmatrix} \alpha \cdot (1 + z^{+1}) \\ \beta \cdot (1 + z^{-1}) \end{bmatrix} \tag{54}$$

for (α, β)=(-1/2, 1/4) in the (5/3) type transform in JPEG 2000. Finally the frequency band signals are generated as

$$\begin{bmatrix} Y_L(z) \\ Y_H(z) \end{bmatrix} = \begin{bmatrix} X_0^{(2)} \\ X_1^{(1)} \end{bmatrix} \tag{55}$$

for

$$Y_b(z) = \sum_{m=0}^{M-1} y_b(m) z^{-m}, \quad b \in \{L, H\} \tag{56}$$

R [ ] in (52) and (53) is the rounding operator which truncates a pixel value in real number to an integer which becomes source of the rounding noise to be reduced in this chapter.

4.2.2. Integer Implementation of the Transform

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

$$X_1^{(3)}(z) = X_1^{(1)}(z) + R[V_3(z) X_0^{(2)}(z)] \tag{57}$$

and the 4-th lifting step

$$X_0^{(4)}(z) = X_0^{(2)}(z) + R[V_4(z) X_1^{(3)}(z)] \tag{58}$$

are added where $V_3(z)$ and $V_4(z)$ are given as

$$\begin{bmatrix} V_3(z) \\ V_4(z) \end{bmatrix} = \begin{bmatrix} \gamma \cdot (1 + z^{+1}) \\ \delta \cdot (1 + z^{-1}) \end{bmatrix} \tag{59}$$

where

$$\begin{cases} \alpha = -1.586134342059924, \\ \beta = -0.052980118572961, \\ \gamma = 0.882911075530934, \\ \delta = 0.443506852043971, \\ k = 1.230174104914001 \end{cases} \tag{60}$$

in the (9/7) type transform in JPEG 2000. Finally the frequency band signals are generated with scaling as

$$\begin{bmatrix} Y_L(z) \\ Y_H(z) \end{bmatrix} = \begin{bmatrix} R[k^{-1}2^{-F} X_0^{(4)}] \\ R[k^{+1}2^{-F} X_1^{(3)}] \end{bmatrix}. \tag{61}$$

Note that the input signal is scaled with 2^F beforehand as illustrated in Figure 50 (b). In the integer implementation, F is set to a positive number as in integer implementation. The smaller the F is, the shorter the bit depth of signals inside the transform which then contributes to reduce hardware complexity implementation of the transform in general.

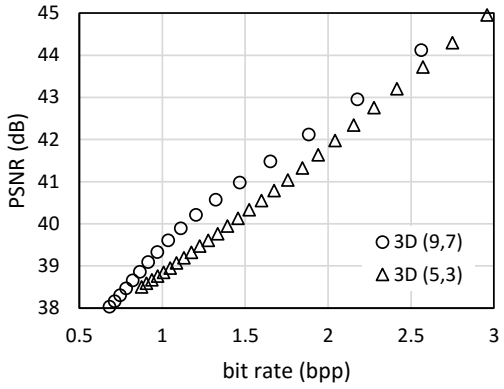
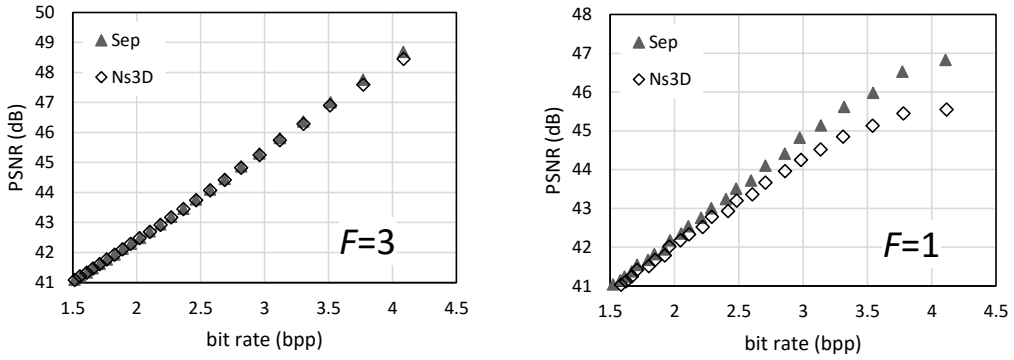


Figure. 51 Lossy coding performance.

The comparison of the (5/3) type 3D transform in [25] and the (9/7) type 3D transform (F=8) in [32] in respect of 'lossy' coding performance is demonstrated in Figure 51. The horizontal axis denotes the bit rate (= code length per pixel). The frequency band signals are quantized with the optimum bit allocation, and entropy coded according to JPEG 2000. The lower the bit rate is, the better the data compression ratio becomes. The vertical axis denotes the peak-signal to noise ratio (PSNR). Larger value means higher quality of the decoded image. As it is well known, it is confirmed that the (9/7) type is better than the (5/3) type in 'lossy' coding of a 3D signal. Therefore this paper focuses on improving the (9/7) type transform for lossy compression of 3D signals.

4.2.3. Problem in Integer Implementation

The two structures comparison for the (9/7) type 3D wavelet transform at F=3 is illustrated in Figure 52 (a) whereas one of these is the separable 3D structure (Sep) in which the 1D transform in Figure 50 (b) is applied three times ,in horizontal direction, in vertical direction, and in inter-frames. The other is the non-separable 3D structure (Ns3D) reported in [32]. It has advantage that it is composed of the minimum number of lifting steps. However it has a problem described as below.



(a) F=3 (b) F=1

Figure. 52 Image quality degradation in integer implementation of the (9/7) type wavelet transform.

Figure 52 (b) compares the two structures at F=1. Here, the shorter the setting of the value of F, the less the hardware complexity. Yet, it can be concluded from the figure, PSNR is lowered which is a proof that quality of the decoded image is degraded by the rounding noise. Particularly, the quantization noise generated in the lossy encoder after the forward transform contributes in maximization of the compression ratio still the rounding noise generated by the rounding operator does not. And this factor defines the upper bound of PSNR in the rate distortion curve as illustrated in Figure 52.

4.2.4. Separable 3D Structure (Sep)

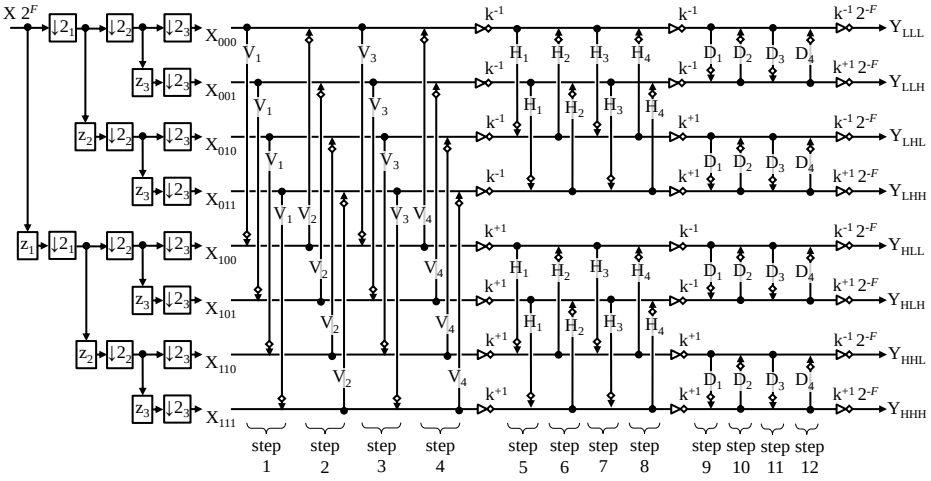


Figure. 53 Separable 3D structure ‘Sep’ for (9/7) type 3D transform.

Figure 53 illustrates the (9/7) type separable 3D wavelet transform (Sep). By applying the 1D structure three times, 12 lifting steps have been result which becomes bottleneck for fast implementation of the transform due to the aforementioned delay with the lifting steps. Signal processing of Sep is detailed as below. Denoting a 3D input signal as

$$X(z_1, z_2, z_3) = \sum_{n_3=0}^{N_3-1} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} x(n_1, n_2, n_3) z_1^{-n_1} z_2^{-n_2} z_3^{-n_3} \tag{62}$$

the input signal is classified into 8 groups as

$$\begin{cases} X_{c_1c_2c_3}(z_1, z_2, z_3) = \\ \downarrow 2_3[\downarrow 2_2[\downarrow 2_1[z_3^{c_3} z_2^{c_2} z_1^{c_1} X(z_1, z_2, z_3) \cdot 2^F]]], \\ c_1, c_2, c_3 \in \{0,1\} \end{cases} \tag{63}$$

for

$$\begin{cases} \downarrow 2_1[X(z_1,z_2,z_3)] = \dfrac{1}{Q}\displaystyle\sum_{p=0}^{Q-1} X(z_1^{1/Q}\cdot W_Q^p, z_2, z_3) \\ \downarrow 2_2[X(z_1,z_2,z_3)] = \dfrac{1}{Q}\displaystyle\sum_{p=0}^{Q-1} X(z_1, z_2^{1/Q}\cdot W_Q^p, z_3) \\ \downarrow 2_3[X(z_1,z_2,z_3)] = \dfrac{1}{Q}\displaystyle\sum_{p=0}^{Q-1} X(z_1, z_2, z_3^{1/Q}\cdot W_Q^p) \end{cases} \tag{64}$$

Where $Q = 2$. Next, the filters

$$\begin{bmatrix} V_1(z_1) & V_3(z_1) \\ V_2(z_1) & V_4(z_1) \end{bmatrix} = \begin{bmatrix} \alpha\cdot(1+z_1^{+1}) & \gamma\cdot(1+z_1^{+1}) \\ \beta\cdot(1+z_1^{-1}) & \delta\cdot(1+z_1^{-1}) \end{bmatrix} \tag{65}$$

$$\begin{bmatrix} H_1(z_2) & H_3(z_2) \\ H_2(z_2) & H_4(z_2) \end{bmatrix} = \begin{bmatrix} \alpha\cdot(1+z_2^{+1}) & \gamma\cdot(1+z_2^{+1}) \\ \beta\cdot(1+z_2^{-1}) & \delta\cdot(1+z_2^{-1}) \end{bmatrix} \tag{66}$$

$$\begin{bmatrix} D_1(z_3) & D_3(z_3) \\ D_2(z_3) & D_4(z_3) \end{bmatrix} = \begin{bmatrix} \alpha\cdot(1+z_3^{+1}) & \gamma\cdot(1+z_3^{+1}) \\ \beta\cdot(1+z_3^{-1}) & \delta\cdot(1+z_3^{-1}) \end{bmatrix} \tag{67}$$

are applied in the lifting steps. Finally, the 8 frequency band signals are obtained as illustrated in the figure 53.

4.2.5. Non-separable 3D Structure (Ns3D)

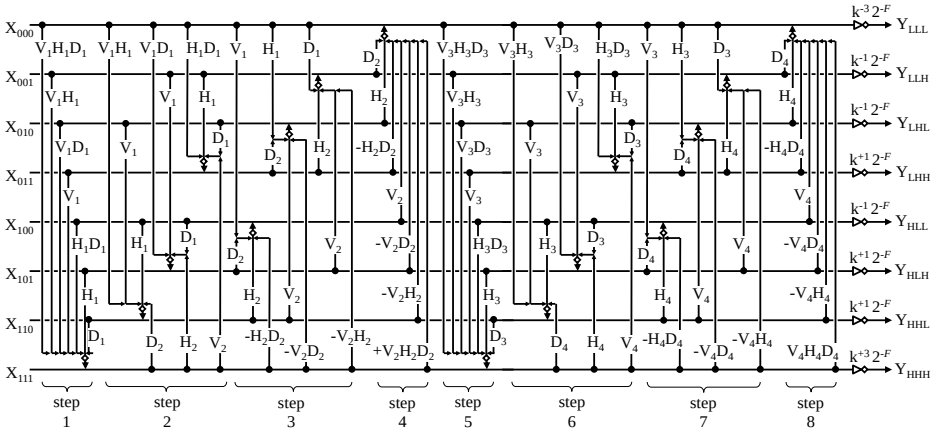


Figure. 54 Non-separable 3D structure ‘Ns3D’ for 3D transform (existing method).

The illustration of the existing method (Ns3D) reported in [32] in Figure 54 which has 8 lifting steps. As Ns3D uses non-separable structure in each lifting step, the total number of lifting steps is decreased from 12 to 8 (66.7 %) comparing to Sep. For example, the 1st lifting steps is described as

$$\begin{aligned} X_{111}^{(1)}(\mathbf{z}) = X_{111}(\mathbf{z}) + R[&V_1H_1D_1X_{000}(\mathbf{z}) \\ &+ V_1H_1X_{001}(\mathbf{z}) + V_1D_1X_{010}(\mathbf{z}) + V_1X_{011}(\mathbf{z}) \\ &+ H_1D_1X_{100}(\mathbf{z}) + H_1X_{101}(\mathbf{z}) + D_1X_{110}(\mathbf{z})] \end{aligned} \tag{68}$$

where $\mathbf{z}=(z_1, z_2, z_3)$ which requires multi-dimensional memory accessing such for instance, $V_1H_1D_1$ and V_1H_1 require 3D and 2D memory accessing, respectively. However, there is a degradation of image quality in its integer implementation which is alleviated by the proposed method as explained in the chapter.

4.3. Proposed Method

In the proposed method, instead of the non-separable 3D structure in the existing method, another structure with minimum rounding noise when compare with non-separable 3D and separable 3D need to be found.

4.3.1. Approach to find the best structure

The separable 3D (9/7) has many coefficients which can be permutated 12! combinations or 479,001,600 combinations as shown in Figure 55 which is really a numerous amount of numbers. Moreover, from every individual combination, there might be a variety of possibilities to become the candidates of Ns3D and Ns2D structures. For the purpose of observing the candidates with minimum rounding noise, the rules to exclude unnecessary combination are determined. In this dissertation, 6 rules are introduced to find minimum rounding noise structure.

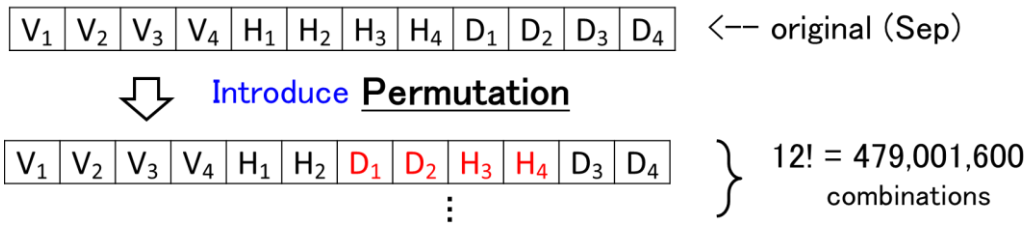


Figure. 55 Permutation of separable 3D (9/7).

- **Rule 1**

Permutation of $V_i H_j$, $V_i D_j$ and $H_i D_j$ for $i \in \{1, 2, 3, 4\}$ are excluded as shown in the example, figure 56. Because they have the same rounding noise

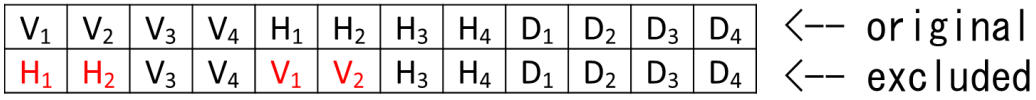


Figure. 56 Example case for the rule 1.

- **Rule 2**

Permutation of V_i H_j, H_i H_j and D_i D_j for i < j are excluded as shown in the example, figure 57. Because they are different from 9/7 wavelet.

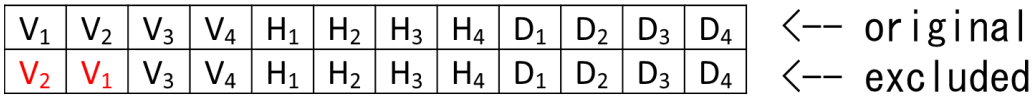


Figure. 57 Example case for the rule 2.

- **Rule 3**

V_i V_j are coupled as V_{ij} for ij $\in$ {12} or {34}. As well as H_i H_j and D_i D_j (As shown in the example, figure 58). Because they can't be separated.

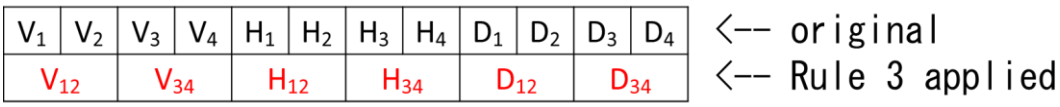


Figure. 58 Example case for the rule 3.

- **Rule 4**

V_{ij} H_{pq} can be implemented with Ns2D as (V_{ij} H_{pg}) for ij,pq $\in$ {12, 34} as shown in example, figure 59. V_{ij} H_{pq} and H_{pq} V_{ij} can be implemented to the same Ns2D, also in the case of V_{ij} D_{pq} and H_{ij} D_{pq}.

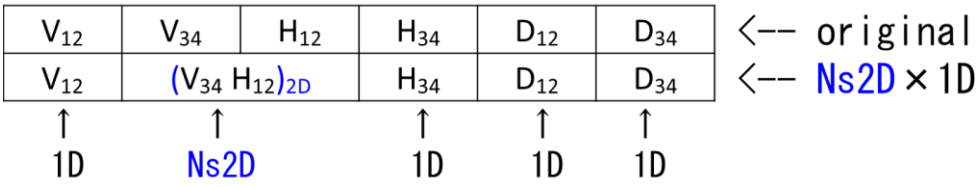


Figure. 59 Example case after implemented with Ns2D.

- **Rule 5**

V_{ij} H_{pq} D_{st} can be implemented with Ns3D as [V_{ij} H_{pg} D_{st}] for ij,pq,st $\in$ {12, 34} as shown in example, figure 60. Also, the same Ns3D [V_{ij} H_{pg} D_{st}] can be applied to V_{ij} D_{st} H_{pq}, $H_{pq}V_{ij}$ D_{st}, H_{pq} D_{st} V_{ij}, D_{st} H_{pq} V_{ij} and D_{st} V_{ij} H_{pq}.

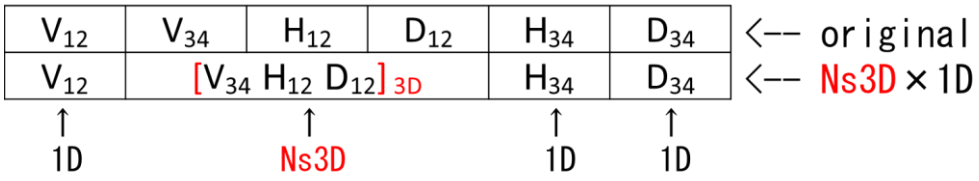


Figure. 60 Example case after implemented with Ns3D.

- **Rule 6**

In case of the original order (figure 55) is permutated, it was experimentally found that the structure has large rounding noise. Because the rounding noise is amplified by large coefficient values. Therefore, the original order must be preserved.

After applied all these aforementioned rules, there exist only 4 candidates left as shown in Figure 61. Observing rounding noise from these 4 candidates, it was found that the structure in candidate 4 has the minimum rounding noise.

						Sep
V_{12}	V_{34}	H_{12}	H_{34}	D_{12}	D_{34}	<-- candidate 1 (original)
V_{12}	$(V_{34}\ H_{12})_{2D}$		H_{34}	D_{12}	D_{34}	<-- candidate 2
V_{12}	V_{34}	H_{12}	$(H_{34}\ D_{12})_{2D}$		D_{34}	<-- candidate 3
V_{12}	$(V_{34}\ H_{12})_{2D}$		$(H_{34}\ D_{12})_{2D}$		D_{34}	<-- candidate 4

Figure. 61 Four candidates for investigating the best structure.

4.3.2. Non-separable 2D structure (Ns2D).

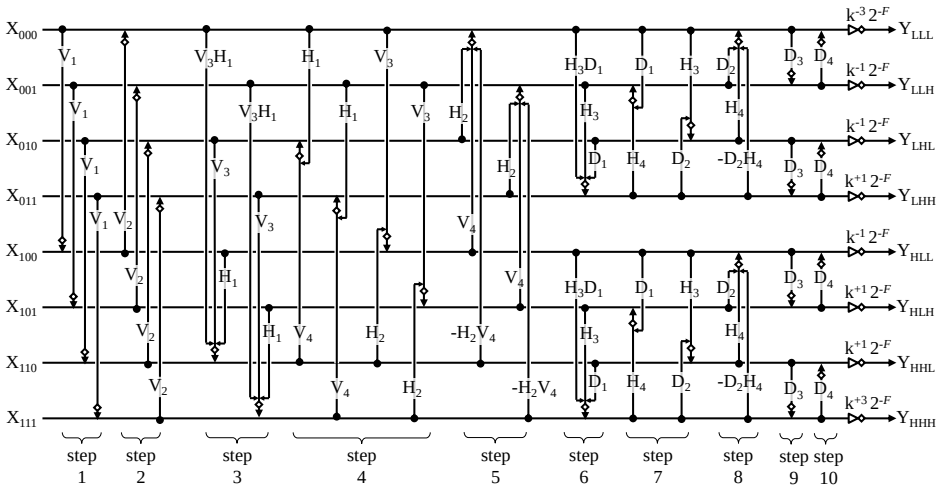


Figure. 62 Combination with non-separable 2D structure ‘Ns2D’ for 3D transform (proposed method).

From section 4.3.1, non-separable 2D structure ‘Ns2D’ was proposed. Figure 62 illustrates the proposed method, composed of non-separable 2D structures in combination with 1D separable structure, differed from the existing method Ns3D. In

which, yet the number of lifting steps is increased comparing to Ns3D but it is not exceeded to the Sep. Nonetheless, total amount of the rounding noise is reduced

The 1st and the 2nd lifting steps are 1D structures can be expressed as

$$\begin{cases} X^{(1)}_{1c_2c_3}(\mathbf{z}) = X_{1c_2c_3}(\mathbf{z}) + R[V_1 X_{0c_2c_3}(\mathbf{z})] \\ X^{(2)}_{0c_2c_3}(\mathbf{z}) = X_{0c_2c_3}(\mathbf{z}) + R[V_1 X^{(1)}_{1c_2c_3}(\mathbf{z})] \end{cases} \tag{69}$$

Where c_2, $c_3 \in \{0,1\}$. The 3rd, 4th and 5th lifting steps are non-separable 2D structures. For example, the 1st step is expressed as

$$\begin{aligned} X^{(2)}_{11c_3}(\mathbf{z}) = X^{(1)}_{11c_3}(\mathbf{z}) + R[V_3 V_1 X^{(2)}_{00c_3}(\mathbf{z}) \\ + V_3 X^{(2)}_{01c_3}(\mathbf{z}) + H_1 X^{(1)}_{10c_3}(\mathbf{z})] \end{aligned} \tag{70}$$

Where $c_3 \in \{0,1\}$. A set of the 6th, 7th and 8th lifting steps is a non-separable 2D structure. The final set of the 9th and 10th lifting steps is the 1D structure again. Thus, the proposed method is a combination of non-separable 2D structures and a product of 1D structures (= separable structure).

4.3.3. Comparison of the Structures

Table. 3 Comparison of the methods

(9/7) type	Sep	Ns3D	Ns2D
rounding operations	72	24	40
lifting steps	12	8	10
memory accessing	1D	3D	2D

The comparison of the three structures: Sep, Ns3D and Ns2D is summarized in the Table 3 where total number of the lifting steps of the proposed Ns2D is increased from 8 to 10 comparing to the existing Ns3D. However, although the total number of the rounding operations is still fewer than Sep, it is raised than the existing method so that no merit in reducing the rounding noise. However the proposed method surely reduces total amount of the rounding noise as confirmed in 4.4. This is considered to be due to the difference between Ns3D and Ns2D described below.

Sep illustrated in figure 53, is composed of the lifting steps V_1, V_2, $\cdots$,D_4 is expressed as

$$Sep \quad V_1V_2V_3V_4H_1H_2H_3H_4D_1D_2D_3D_4 . \tag{71}$$

Firstly, order of the lifting steps in the existing method is changed as

$$Sep' \quad V_1V_2H_1H_2D_1D_2V_3V_4H_3H_4D_3D_4 , \tag{72}$$

And, a part of it is implemented in the non-separable 3D structure as

$$Ns3D \quad [V_1V_2H_1H_2D_1D_2][V_3V_4H_3H_4D_3D_4] \tag{73}$$

Where the [ ] part denotes the non-separable 3D structure. Differ from the existing method in (72), the proposed method can be expressed as

$$Ns2D \quad V_1V_2(V_3V_4H_1H_2)(H_3H_4D_1D_2)D_3D_4 \tag{74}$$

Does not change the order of (71) where the ( ) part denotes the non-separable 2D structure.

4.4. Experimental Result

4.4.1. Evaluation of rounding noise

A set of MRI volumetric image tested for the method in this chapter is illustrated Figure 43 which compose of 8 frames, 256×256 pixels for each particular frame whereas the individual pixel has 8 bit depth integer value; $x(n) \in [0,255]$.

In, figure 63 we summarizes variance of the rounding noise observed in each frequency band signals as F=0. In all bands, although the existing Ns3D is the biggest (=worst), the proposed Ns2D is apparently better than Ns3D. Figure 64 summarized the variance averaged over all frequency bands. An auto-regressive (AR) model and a random signal are added as the input signal. It can be clearly detected that the proposed method is the best in respect of variance of the rounding noise in frequency domain.

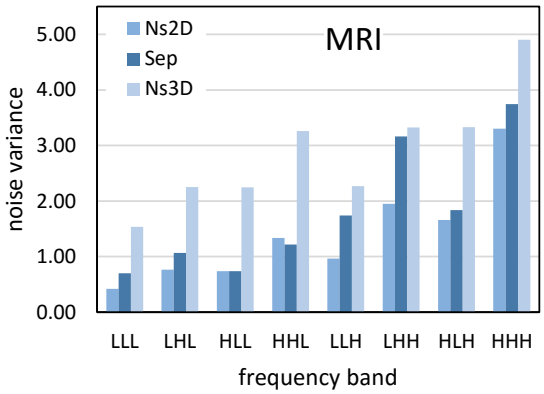


Figure. 63 Rounding noise in each frequency band.

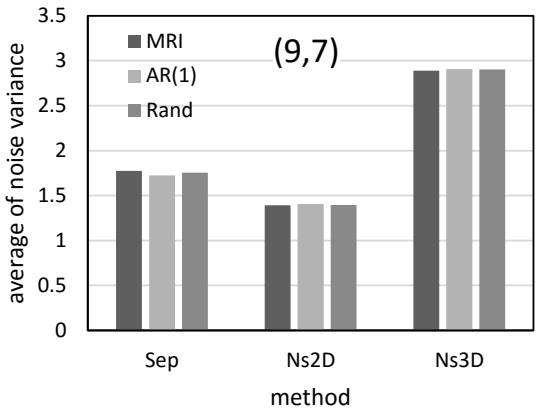


Figure. 64 Rounding noise averaged over all frequency bands.

For the variance of the rounding noise in pixel domain, we summarizes in Figure 65. There is no rounding operation is applied in the forward transform as well as the quantization for lossy coding. In the backward transform, signals are rounded at F=0. In this domain, difference between the decoded signal and the input signal is defined also. Its variance is measured with the PSNR defined as

$$PSNR = 20\log_{10} 255 - 10\log_{10} \sigma_{Rnd}^2 \tag{75}$$

Where σ_{Rnd} denotes the standard deviation of the rounding noise in pixel domain. Based on the figure 65, it can be confirmed the proposed method significantly reduces total amount of the rounding noise, and has the best quality of the decoded signal.

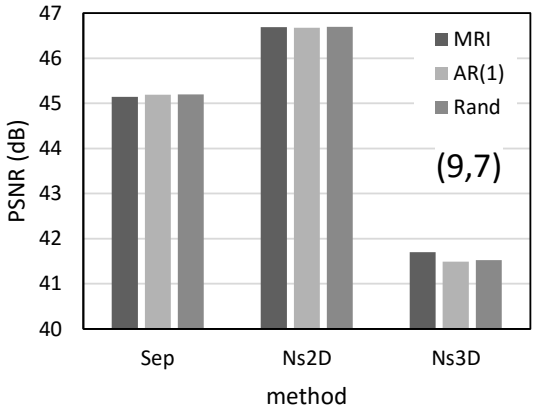


Figure. 65 Rounding noise in pixel domain.

4.4.2. Evaluation of Lossy Coding Performance

Figure 66 illustrates the rate-distortion curves which compare performance of the methods in lossy coding mode. In Figure 66 (a) at F=0, the existing Ns3D is the worst at high bit rate. Obviously, the proposed Ns2D is observed to be the best. Whereas the degradation of the PSNR is related to the rounding noise in pixel domain in Figure 65 that determines the upper bound of the PSNR in figure 66. It can be modelled as

$$PSNR = 20\log_{10} 255 - 10\log_{10}(\sigma_{Rnd}^2 + \sigma_{Qnt}^2) \tag{76}$$

Where σ_{Rnd} and σ_{Qnt} denotes the standard deviation of the rounding noise and the quantization noise, respectively.

As for the case $\sigma_{Qnt} >> \sigma_{Rnd}$ holds, it can be clearly observed that the quantization noise is related to the bit rate Bit as

$$\frac{\Delta PSNR}{\Delta Bit} = \frac{-10\log_{10}\Delta\sigma_{Qnt}^2}{\Delta Bit} \propto 6.02 \tag{77}$$

where Δ denotes deviation of the value. Particularly, one bit increase of *Bit* provides 6.02 (dB) increase of the PSNR which from decrease of standard deviation of the quantization noise σ_{Qnt}. In such case, it can be concluded that degradation of quantization noise raises PSNR whereas the total amount of the rounding noise is independent of PSNR. In case of $\sigma_{Qnt} << \sigma_{Rnd}$ holds, (76) becomes

$$PSNR \approx 20\log_{10} 255 - 10\log_{10} \sigma_{Rnd}^2 \tag{78}$$

which determines the upper bound of PSNR in the rate distortion curves.

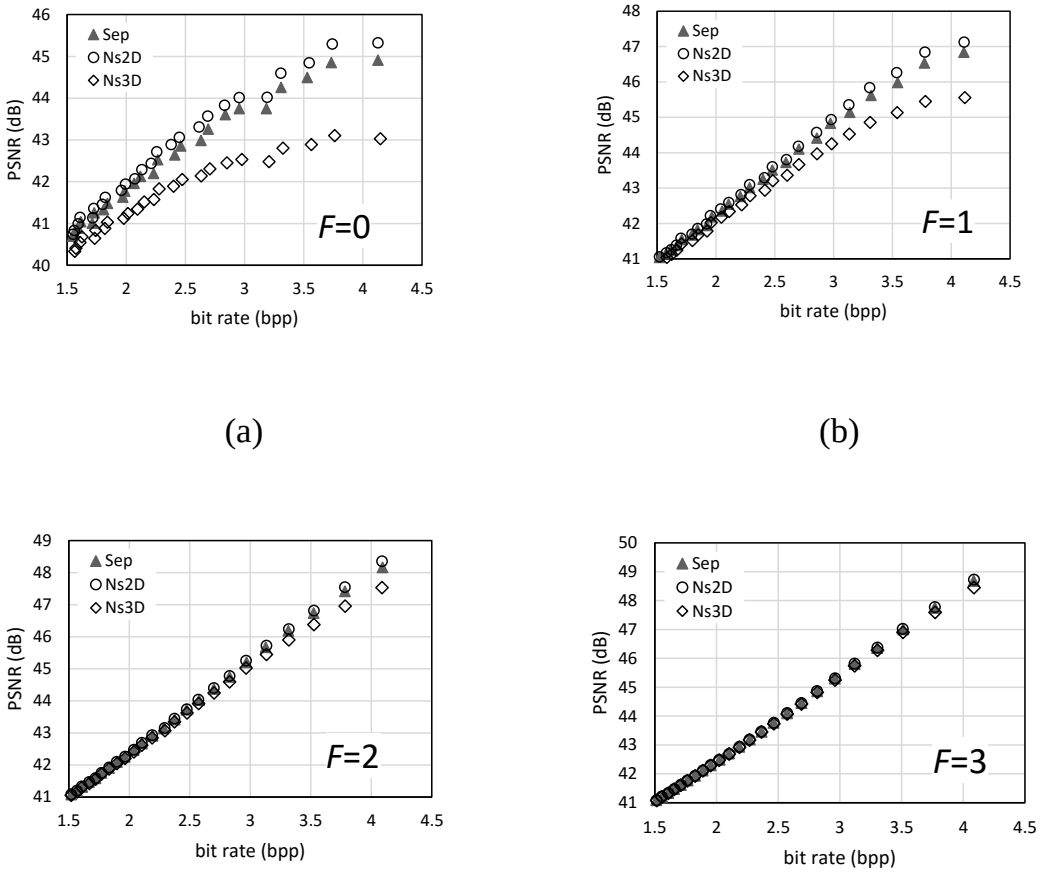


Figure. 66 Performance in lossy coding mode.

The relation between the bit depth F and the standard deviation of the rounding noise σ_{Rnd} as an another aspect can be modelled as

$$\sigma_{Rnd}^2 \propto 2^{-F} . \tag{79}$$

For that reason, as long as larger F is used, the upper bound increases which is confirmed according to Figure 66 (b), (c) and (d). At F=3, there is almost no difference among the existing method and the method discussed in this chapter as indicated in Figure 66 (d). Also, it can be clearly seen as the proposed method is superior to the existing method in respect of PSNR at high bit rate in short length integer implementation of the transform.

4.5. Summary

In this chapter, non-separable 2D structures, which was found by 6 rules, was introduced. Comparing to the 3D structures existing method, total number of lifting steps was increased. However, total amount of rounding noise due to integer expression of signal values inside the transform (integer implementation) was minimized. It was confirmed that it contributes to increasing upper bound (quality of decoded images) in the rate distortion curve in lossy coding mode.

Chapter 5

5. Channel Scaling for Integer Implementation of Minimum Lifting 2D Wavelet Transform

5.1. Motivation of this proposal

In the integer implementation within the wavelet transform, signal values inside the transform are rounded to integers so that shorter word length of the integer contributes to smaller memory space and faster calculation [48-53] which is beneficial for hardware implementation of WT. However it generates noise due to the rounding (rounding noise).

Unlike the quantization noise, the rounding noise is independent of the bit rate (compression ratio). Therefore the rounding noise appears as the upper bound of the rate distortion curve (the bit rate versus the peak signal to noise ratio) in high bit rate lossy coding of images. In this chapter, we focus on the reduction the rounding noise of the minimum lifting WT by introducing the channel scaling.

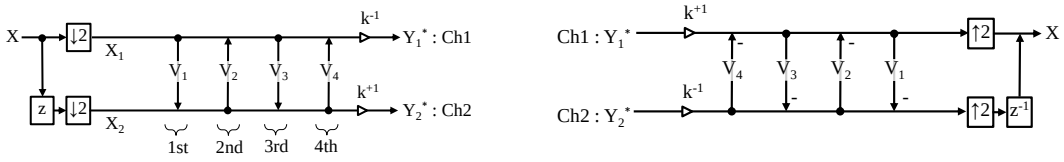
More precisely, we focus on the reduction of the rounding noise of the minimum lifting WT based on the 9/7 WT for 2D case in integer implementation by introducing the channel scaling. There have been four channels in 2D WT and, in all channels; the maximum absolute value (MAV) of signals inside the transform is limited by word length of signals in integer implementation. Based on the fact that MAV in each

channel is not the same, different scaling in each channel maintaining the original signal value at output of the transform is introduced. The scaling parameters are defined for the case that MAV of all channels become the same. This expressed that the dynamic range given by the word length is fully utilized in each channel. Accordingly, the rounding noise is reduced without altering the signal at output of WT.

5.2. Existing Method

A problem of the minimum lifting WT in the integer implementation is addressed based on the explanation of the Integer implementation of WT.

5.2.1. Lifting Wavelet and its Integer Implementation



(a) (b)

Figure. 67 Lifting wavelet transform for 1D signal.

In figure 67 (a), a forward transform with the 9/7 WT for a 1D input signal sequence is illustrated:

$$X(z) = \sum_{n=0}^{N-1} x(n) z^{-n}, \quad x(n) \in [-2^{D-1}, 2^{D-1}) \tag{80}$$

where $x(n)$ denotes a D bit depth signal value at location n. The decomposition of the transform $X(z)$ into two channels is mentioned as

$$\begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} = \downarrow 2\left[\begin{bmatrix} 1 \\ z \end{bmatrix} X(z)\right] = \begin{bmatrix} \downarrow 2[X(z)] \\ \downarrow 2[zX(z)] \end{bmatrix} \tag{81}$$

where

$$\downarrow 2[X(z)] = \left\{X(z^{1/2}) + X(-z^{1/2})\right\} \cdot 2^{-1}. \tag{82}$$

$X_2(z)$ in channel 2 (Ch2) is predicted from $X_1(z)$ in channel 1 (Ch1) with an FIR filter $V_1(z)$ for the first lifting step. Then being in four lifting steps, the transform outputs two frequency band signals $Y_1^*(z)$ and $Y_2^*(z)$ as

$$\begin{bmatrix} Y_1^*(z) \\ Y_2^*(z) \end{bmatrix} = \begin{bmatrix} k^{-1} & 0 \\ 0 & k^{+1} \end{bmatrix} \begin{bmatrix} 1 & V_4(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ V_3(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & V_2(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ V_1(z) & 1 \end{bmatrix} \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} \tag{83}$$

where

$$\begin{bmatrix} V_p(z) \\ V_q(z) \end{bmatrix} = \begin{bmatrix} h_p(1+z^{+1}) \\ h_q(1+z^{-1}) \end{bmatrix}, \quad p = 1,3, \quad q = 2,4 \tag{84}$$

and

$$\begin{cases} h_1 = -1.5861\cdots, & h_2 = -0.0530\cdots \\ h_3 = +0.8829\cdots, & h_4 = +0.4435\cdots \\ k = +1.23017\cdots. \end{cases} \tag{85}$$

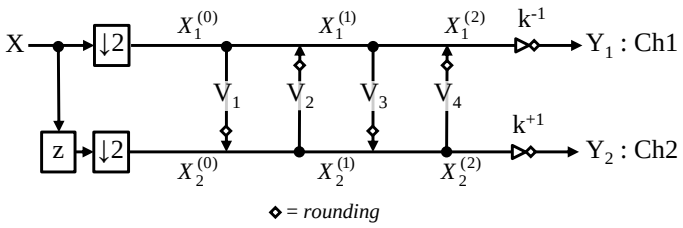


Figure. 68 Integer implementations of the lifting wavelet transform.

These band signals are quantized and encoded with an entropy encoder such as EBCOT [13] to produce a bit stream (compressed data) in a lossy coding. For the input signal reconstruction, the backward transform illustrated in Figure 68 (b) is applied immediately after the bit stream is decoded with an entropy decoder and processed with inverse of the quantization. In such case, the reconstructed signal contains noise due to the quantization (quantization noise) so that the more the volume of noise, the less the amount of data Figure 69 illustrates an integer implementation of the transform in Figure 68 (a). After each filter, a rounding operation is introduced. In the 1st lifting step, Ch2 is predicted from Ch1 as

$$X_2^{(1)}(z) = X_2^{(0)}(z) + R[V_1(z)X_1^{(0)}(z)] \tag{86}$$

where a rounding operator $R[x]$ rounds a rational number x to the nearest integer. Similarly, Ch1 is updated from Ch2 as

$$X_1^{(1)}(z) = X_1^{(0)}(z) + R[V_2(z)X_2^{(1)}(z)]. \tag{87}$$

After two more lifting steps:

$$X_2^{(2)}(z) = X_2^{(1)}(z) + R[V_3(z)X_1^{(1)}(z)] \tag{88}$$

and

$$X_1^{(2)}(z) = X_1^{(1)}(z) + R[V_4(z)X_2^{(2)}(z)], \tag{89}$$

two frequency band signals are generated as

$$\begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} = \begin{bmatrix} R[k^{-1} \cdot X_1^{(2)}(z)] \\ R[k^{+1} \cdot X_2^{(2)}(z)] \end{bmatrix}. \tag{90}$$

Again, an integer implementation case where the bit depth of signals inside the transform is limited to B (bit) is considered in this chapter which is then described as

$$\max\left\{ |X_c^{(s)}(z)| \;\middle|\; \forall s, \forall c \right\} < 2^{B-1} \tag{91}$$

where the maximum absolute value (MAV) of signal values of the signal sequence $X_c^{(s)}(z)$ in the left hand side. Yet the smaller values B causes the lower numerical complexity, the band signals in the integer implementation contain noise due to the rounding operation (rounding noise) even though there is no quantization noise.

5.2.2. Standard Lifting Wavelet for 2D Signal

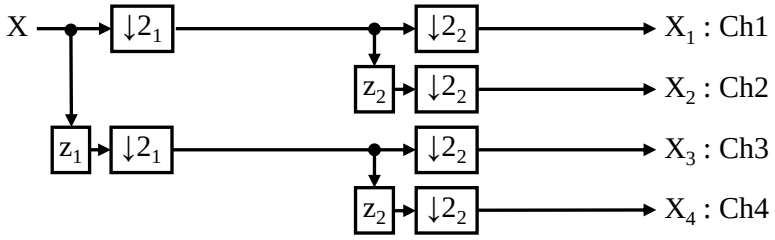


Figure. 69 Channel decomposition of a 2D signal.

For a 2D input signal $X(\mathbf{z})$, the transform splits the input signal into four channels as illustrated in Figure 69. It is denoted as

$$\begin{bmatrix} X_1(\mathbf{z}) \\ X_2(\mathbf{z}) \\ X_3(\mathbf{z}) \\ X_4(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \downarrow 2_2 \left[\begin{bmatrix} 1 \\ z_2 \end{bmatrix} W_1(\mathbf{z}) \right] \\ \downarrow 2_2 \left[\begin{bmatrix} 1 \\ z_2 \end{bmatrix} W_2(\mathbf{z}) \right] \end{bmatrix} \tag{92}$$

where

$$\begin{bmatrix} W_1(\mathbf{z}) \\ W_2(\mathbf{z}) \end{bmatrix} = \downarrow 2_1 \left[\begin{bmatrix} 1 \\ z_1 \end{bmatrix} X(\mathbf{z}) \right] \tag{93}$$

and

$$\begin{bmatrix} \downarrow 2_1[X(\mathbf{z})] \\ \downarrow 2_2[X(\mathbf{z})] \end{bmatrix} = \begin{bmatrix} X(z_1^{1/2}, z_2) + X(-z_1^{1/2}, z_2) \\ X(z_1, z_2^{1/2}) + X(z_1, -z_2^{1/2}) \end{bmatrix} \cdot 2^{-1} \tag{94}$$

for

$$X(\mathbf{z}) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x(\mathbf{n}) z_1^{-n_1} z_2^{-n_2} \tag{95}$$

where $\mathbf{z}=(z_1,z_2)$ and $\mathbf{n}=(n_1,n_2)$.

By applying the 1st, 2nd, 3rd and 4th lifting steps vertically with

$$\begin{bmatrix} V_p(\mathbf{z}) \\ V_q(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_p(1 + z_1^{+1}) \\ h_q(1 + z_1^{+1}) \end{bmatrix}, \quad p = 1,3, \quad q = 2,4\,, \tag{96}$$

and the 5th, 6th, 7th and 8th lifting steps horizontally with

$$\begin{bmatrix} H_p(\mathbf{z}) \\ H_q(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_p(1+z_2^{+1}) \\ h_q(1+z_2^{+1}) \end{bmatrix}, \quad p=1,3, \quad q=2,4\,, \tag{97}$$

to the channel signals in (92), the transform outputs four frequency band signals $Y_1(\mathbf{z})$, $Y_2(\mathbf{z})$, $Y_3(\mathbf{z})$ and $Y_4(\mathbf{z})$ as illustrated in Figure 70 which is referred to as Separable structure (Sep).

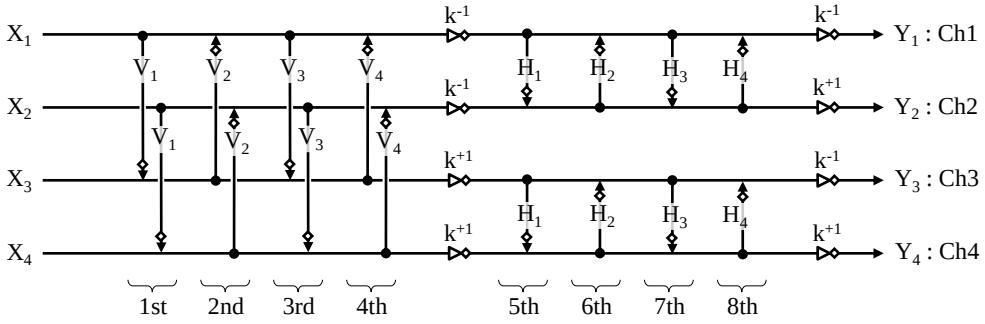


Figure. 70 The standard wavelet transform based on the separable 2D structure (Sep).

5.2.3. Minimum Lifting Wavelet and its Problem

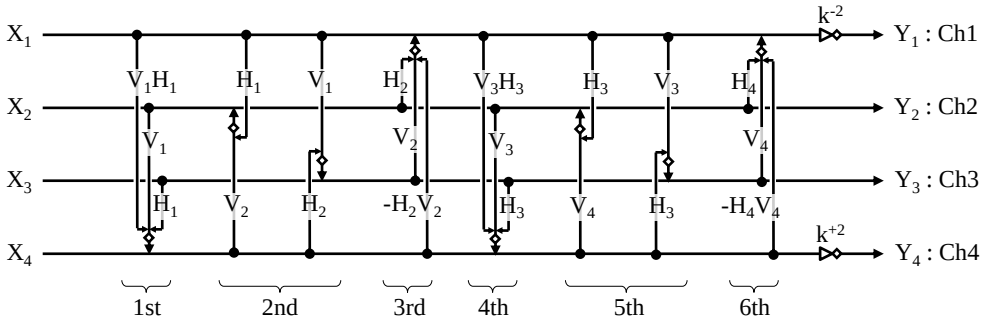


Figure. 71 The minimum lifting wavelet transform based on the non-separable 2D structure (Nsp).

The minimum lifting steps in WT based on the non-separable structure (Nsp) has reduced number of lifting steps in contrast to Sep. In Figure 71, Nsp with 6 lifting steps is illustrated by the introduction of 2D memory accessing. The signal in Ch4 is produced as (for instance);

$$X_4^{(1)}(\mathbf{z}) = X_4(\mathbf{z}) + R[P_4^{(1)}(\mathbf{z})] \tag{98}$$

for

$$\begin{aligned} P_4^{(1)}(\mathbf{z}) = &\left[V_1(\mathbf{z})H_1(\mathbf{z}) \quad V_1(\mathbf{z}) \quad H_1(\mathbf{z})\right] \\ &\cdot \left[X_1(\mathbf{z}) \quad X_2(\mathbf{z}) \quad X_3(\mathbf{z})\right]^T \end{aligned} \tag{99}$$

in the 1st lifting step. In this step, a 2D filtering with 2D memory accessing $V_1(\mathbf{z})H_1(\mathbf{z})$ is used. In the 2nd lifting step, the calculation of Ch2 and Ch3:

$$\begin{bmatrix} X_2^{(1)}(\mathbf{z}) \\ X_3^{(1)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} X_2(\mathbf{z}) + R[P_2^{(1)}(\mathbf{z})] \\ X_3(\mathbf{z}) + R[P_3^{(1)}(\mathbf{z})] \end{bmatrix} \tag{100}$$

for

$$\begin{bmatrix} P_2^{(1)}(\mathbf{z}) \\ P_3^{(1)}(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} H_1(\mathbf{z}) & V_2(\mathbf{z}) \\ V_1(\mathbf{z}) & H_2(\mathbf{z}) \end{bmatrix} \begin{bmatrix} X_1(\mathbf{z}) \\ X_4^{(1)}(\mathbf{z}) \end{bmatrix} \tag{101}$$

can be done simultaneously without necessary to wait for calculation results each other.

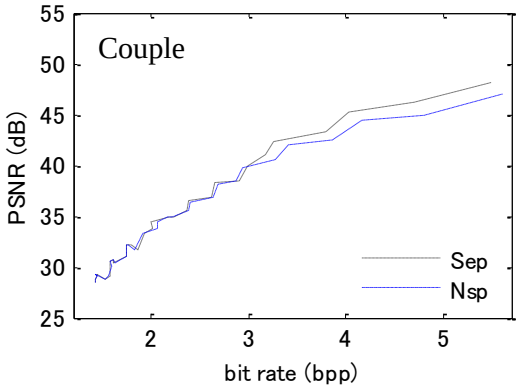


Figure. 72 Problem of the existing method in lossy coding.

Nonetheless, there is a problem remaining in the existing method in lossy coding. As indicated by the rate distortion curve in Figure 72, the horizontal axis for the average code length (data volume in bit per pixel) and the vertical axis the peak-signal to noise ratio (PSNR) as follows;

$$PSNR = 10\log_{10}\frac{255^2}{V[E(\mathbf{z})]} \quad (dB) \tag{102}$$

where $V[E(\mathbf{z})]$ stands for the variance of the noise $E(\mathbf{z})$. As illustrated in the figure 72, PSNR is saturated in high bit rates especially in Nsp.

The root cause of the issue is explained as below. The difference between the original input signal $X(\mathbf{z})$ and the reconstructed signal (output of the backward transform) contains two noises uncorrelated each other. One is the quantization noise $N_Q(\mathbf{z})$ and the other is the rounding noise $N_R(\mathbf{z})$. Namely variance of $E(\mathbf{z})$ becomes

$$\begin{aligned} V[E(\mathbf{z})] &= V[N_Q(\mathbf{z})] + V[N_R(\mathbf{z})] \\ &\approx \begin{cases} V[N_Q(\mathbf{z})] & case\ A \\ V[N_R(\mathbf{z})] & case\ B \end{cases} \end{aligned} \tag{103}$$

where the quantization noise dominates over the rounding noise in *case A* and vice versa in *case B*. Such, the logarithm of the variance of the quantization noise is proportional to the bit rate. On the other hand, the variance of the rounding noise is independent of the bit rate so that upper bound of PSNR is determined by the rounding noise. Namely,

$$PSNR \propto \begin{cases} T & case\ A \\ U & case\ B \end{cases} \tag{104}$$

where

$$T \propto -\log V[N_Q(\mathbf{z})], \quad U \propto \log V[N_R(\mathbf{z})]. \tag{105}$$

whereas T denotes the bit rate and U is a constant which is independent of T.

5.3. Proposed Method

The primitive concept of the channel scaling explained for 1D case is applied to the non-separable 2D structure as the proposed method.

5.3.1. Channel Scaling in the 1D Transform

In figure 73, a new integer implementation of the 1D forward transform is described while the scaling parameters s_1 and s_2 are introduced comparing to figure 68. And the ratios of these are expressed as;

$$[s_{12} \quad s_{21}] = [s_2 / s_1 \quad s_1 / s_2] \tag{106}$$

This also introduced in each lifting step for having the same signal as those in Figure 67 (a) except the rounding noise. Particularly, the filters in Figure 68 are replaced as

$$\begin{bmatrix} V_1(\mathbf{z}) & V_2(\mathbf{z}) & V_3(\mathbf{z}) & V_4(\mathbf{z}) \end{bmatrix} \rightarrow \begin{bmatrix} s_{12}V_1(\mathbf{z}) & s_{21}V_2(\mathbf{z}) & s_{12}V_3(\mathbf{z}) & s_{21}V_4(\mathbf{z}) \end{bmatrix} \tag{107}$$

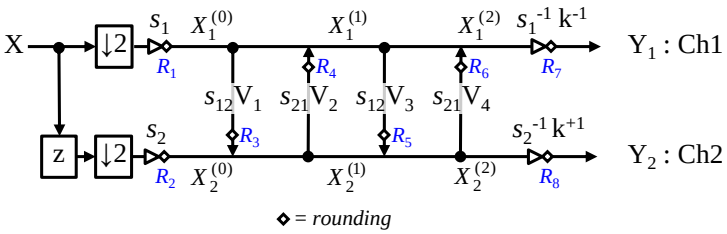


Figure. 73 Channel scaling (Scl) in the 1D transforms.

in figure 73. In this paper, the scaling parameter s_c in channel $c \in \{1,2\}$ is determined so that the dynamic range of signal values in each channel is fully utilized.

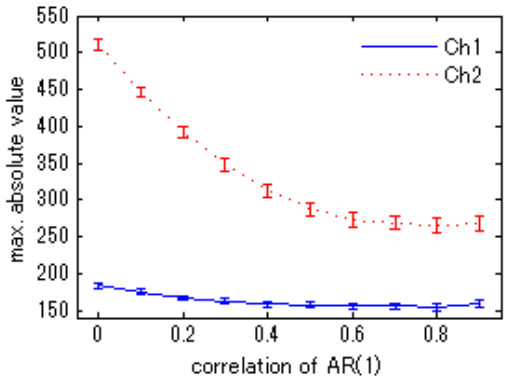


Figure. 74 The maximum absolute value in each channel.

The maximum absolute value (MAV) in channel $c \in \{1,2\}$ defined in instance as

$$m_c = \max\left\{ | X_c^{(s)}(z) | \;\middle|\; \forall s \right\} \tag{108}$$

are different in general. Figure 74 illustrates m_1 and m_2 of D=8 bit depth AR(1) model signals with different correlation coefficient ρ .For example, (m_1, m_2) = (157.86, 272.08) in average for 64 kind of signals with ρ = 0.9. Here, we prepare the bit depth B which satisfies (91) is prepared for the integer implementation of the transform so that the dynamic range of Ch2 is fully utilized under this bit depth yet that of Ch1 is not. For that case, we determine the scaling parameters as

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} \max\{ m_1, m_2 \} \cdot m_1^{-1} \\ \max\{ m_1, m_2 \} \cdot m_2^{-1} \end{bmatrix} \tag{109}$$

so that MAV m_1' and m_2' after the channel scaling become almost the same under the bit depth constraint:

$$m_1' \approx m_2' < 2^{B-1} \tag{110}$$

for

$$\begin{bmatrix} m_1' & m_2' \end{bmatrix} \approx \begin{bmatrix} m_1 \cdot s_1 & m_2 \cdot s_2 \end{bmatrix}. \tag{111}$$

5.3.2. Proposed Transcoding System II

In the following, the relationship between the scaling parameters (s_1, s_2) and the variance of the rounding noise in frequency domain is analyzed. By characterizing the

rounding noise as the difference between the frequency band signals in figure 67 (a) and those in figure 73 as

$$\begin{bmatrix} E_1(\mathbf{z}) \\ E_2(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} Y_1(\mathbf{z}) \\ Y_2(\mathbf{z}) \end{bmatrix} - \begin{bmatrix} Y_1^*(\mathbf{z}) \\ Y_2^*(\mathbf{z}) \end{bmatrix}, \tag{112}$$

Which is described with the rounding noise in each lifting steps denoted as R_i , $i \in \{1,2,\cdots,8\}$ as

$$\begin{aligned} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} &= \begin{bmatrix} R_7 \\ R_8 \end{bmatrix} + \begin{bmatrix} s_1^{-1}k^{-1} & 0 \\ 0 & s_2^{-1}k^{+1} \end{bmatrix} \\ &\cdot \left(\begin{bmatrix} R_6 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & s_{21}V_4 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 0 \\ R_5 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ s_{12}V_3 & 1 \end{bmatrix} \right. \right. \\ &\cdot \left. \left(\begin{bmatrix} R_4 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & s_{21}V_2 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 0 \\ R_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ s_{12}V_1 & 1 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \right) \right) \right) \end{aligned} \tag{113}$$

where (**z**) are omitted for simplification of notation. Substituting (106) into (113), it is expressed as

$$\begin{aligned} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} R_7 \\ R_8 \end{bmatrix} &+ \mathbf{K}\left(\mathbf{G}_6 R_6 + \mathbf{G}_4 R_4 + \mathbf{G}_1 R_1\right) s_1^{-1} \\ &+ \mathbf{K}\left(\mathbf{G}_5 R_5 + \mathbf{G}_3 R_3 + \mathbf{G}_2 R_2\right) s_2^{-1} \end{aligned} \tag{114}$$

where

$$\mathbf{K} = \begin{bmatrix} k^{-1} & 0 \\ 0 & k^{+1} \end{bmatrix}, \quad \mathbf{G}_6 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{G}_4 = \begin{bmatrix} 1+V_3V_4 \\ V_3 \end{bmatrix},$$
$$\mathbf{G}_1 = \begin{bmatrix} (1+V_3V_4)(1+V_1V_2)+V_1V_4 \\ (1+V_1V_2)V_3+V_1 \end{bmatrix}, \quad \mathbf{G}_5 = \begin{bmatrix} V_4 \\ 1 \end{bmatrix},$$
$$\mathbf{G}_3 = \begin{bmatrix} (1+V_3V_4)V_2+V_4 \\ V_2V_3+1 \end{bmatrix}, \quad \mathbf{G}_2 = \begin{bmatrix} (1+V_3V_4)V_2+V_4 \\ V_2V_3+1 \end{bmatrix}$$

with scaling parameters (s_1, s_2). Defining the total variance of the rounding noise in (114) as

$$I(s_1, s_2) = \frac{V[E_1]+V[E_2]}{2}, \tag{115}$$

it becomes

$$I(s_1, s_2) = \alpha_0 + \alpha_1 \cdot s_1^{-2} + \alpha_2 \cdot s_2^{-2} \tag{116}$$

under the assumptions that all the rounding noise have the same variance:

$$V[R_i] = 1/12 \quad for \ \forall i \tag{117}$$

and those are uncorrelated each other:

$$V[R_i R_j] = 0 \quad for \ i \neq j\,. \tag{118}$$

Consequently, the total variance of the rounding noise to be reduced is expressed as a function of the scaling parameters s_1 and s_2.

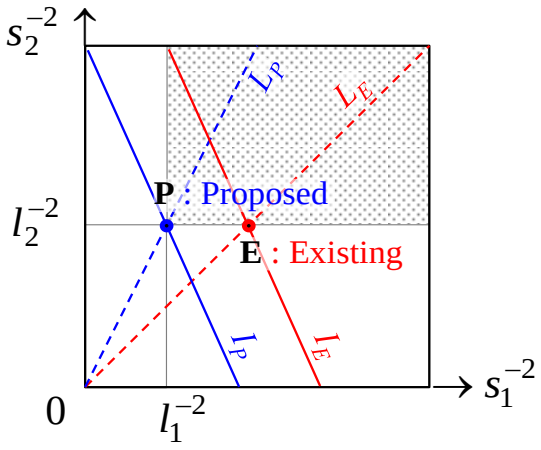


Figure. 75 Channel scaling (Scl) and rounding noise.

Figure 75 indicates contours of the total variance I in (115) with solid lines on (s_1^{-2}, s_2^{-2}) plane. From (110) and (111), the bit depth constraint is expressed as

$$m_1 \cdot s_1 < 2^{B-1} \quad \wedge \quad m_2 \cdot s_2 < 2^{B-1} \tag{119}$$

which means the region

$$s_1^{-2} > l_1^{-2} \quad \wedge \quad s_2^{-2} > l_2^{-2} \tag{120}$$

where

$$[l_1 \quad l_2] = [m_1^{-1} \quad m_2^{-1}] \cdot 2^{B-1} \tag{121}$$

on the plane (shadowed region in figure 75). From (109), the scaling parameters of the proposed method are on the line:

$$L_P : \quad [m_1 \quad -m_2] \cdot [s_1 \quad s_2]^T = 0 \tag{122}$$

and therefore the proposed method sets the parameters on the point **P** in figure 75. In contrast, the existing method sets on the line:

$$L_E: \quad [1 \quad -1]\cdot[s_1 \quad s_2]^T = 0\,. \tag{123}$$

The existing method sets on the point **E** in the figure results,

$$I_E(s_1, s_2) \geq I_P(s_1, s_2) \tag{124}$$

Holds that the proposed method decreases the total variance of the rounding noise with the scaling parameters in (114).

5.3.3. Channel Scaling in the 2D Transform

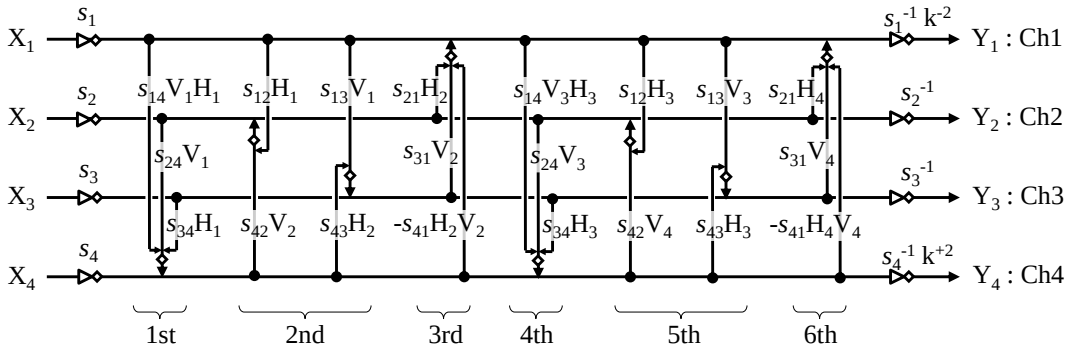


Figure. 76 The minimum lifting wavelet transform with the channel scaling (Nsp+Scl).

The proposed method (Nsp+Scl) for 2D signals is shown in Figure 76. The channel scaling described in Figure 71 has four scaling parameters

$$\mathbf{s} = [s_1 \quad s_2 \quad s_3 \quad s_4] \tag{125}$$

and their ratio

$$s_{pq} = s_q / s_p \quad for \quad p,q \in \{1,2,3,4\} \tag{126}$$

as illustrated in the figure 76 The scaling parameters in the proposed method are set to

$$s_c = \max\{ m_1, m_2, m_3, m_4 \} \cdot m_c^{-1} \tag{127}$$

for $c \in \{1,2,3,4\}$. In contrast, the existing method sets the parameters to

$$s_c = 1 \quad for \quad c \in \{1,2,3,4\} . \tag{128}$$

Based on the fact that the proposed method is expected to reduce the total variance:

$$I(\mathbf{s}) = \frac{1}{4}\sum_{c=1}^{4} V[E_c] \quad = \alpha_0 + \sum_{c=1}^{4} \alpha_c \cdot s_c^{-2} \tag{129}$$

of the rounding noise in frequency band signals for raising the upper bound U in (104) in lossy coding of images.

5.4. Experimental Result

In this section, effect of the channel scaling is examined. After determining the scaling parameters, the total variance of the rounding noise is evaluated together with the upper bound in lossy coding.

5.4.1. Parameter Setting in Channel Scaling

First of all, the maximum absolute value of signals inside the transform in the existing method (Nsp) in figure 71 is investigated. The bit depth defined as

$$B_c = 1 + \log_2 \max\left\{ |X_c^{(s)}(z)| \;\middle|\; \forall s \right\} \tag{130}$$

for each channel $c \in \{1,2,3,4\}$ is measured for 10 input images {1.Couple, 2.Boat, 3.Lena, 4.Cameraman, 5.Girl, 6.Airplane, 7.Woman, 8.Building, 9.Barbara, 10.Text} with 8 bit depth each. B_c averaged over the input images is summarizes in figure 77. The longest bit depth is in Ch4 which has $B_4 = 10.3 \pm 0.2$ (bit) where 0.2 denotes the standard deviation. Therefore the transform must have the bit depth B in (87) which is longer than B_4. The shortest bit depth is $B_1 = 8.6 \pm 0.2$ (bit) in Ch1. Therefore the dynamic range provided by B is not fully utilized in Ch1, Ch2 and Ch3 in the existing method.

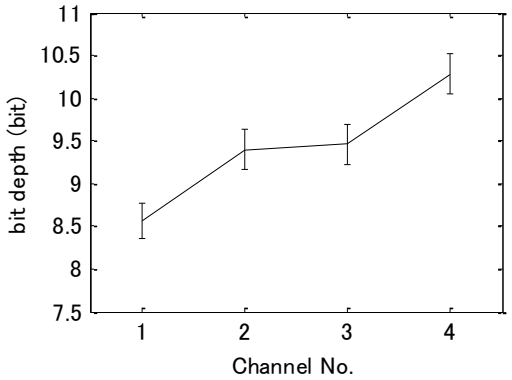


Figure. 77 Bit depth of signal values in each channel.

For absolutely applying the dynamic range of the transform, the scaling parameters are determined according to (126). Figure 78 summarizes the parameters averaged

over the input images. For example, s_4 =1.0 ± 0.0 in Ch4. The largest value was observed to be s_1 =3.33 ± 0.35 in Ch1. Of all the channels have almost the same bit depth by introducing these scaling parameters in our proposed method in this chapter.

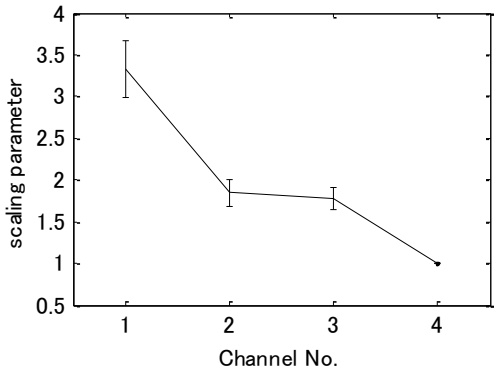


Figure. 78 Scaling parameters in each channel.

5.4.2. Effect of Channel Scaling for Various Images

The lowest bit depth required for the integer implementation of the transform illustrated in Figure 79 is defined as

$$B_{\min} = 1 + \log_2 \max\left\{ | X_c^{(s)}(z) | \;\middle|\; \forall s, \forall c \right\} \tag{131}$$

for all the channels and measured for each of the input images. It can be confirmed that the proposed method (Nsp+Scl) has the same $B_{\min}$ as the existing method (Nsp). Note that B_1, B_2, B_3, B_4 and $B_{\min}$ have almost the same value in the proposed method and therefore the dynamic range is fully utilized in all the channels.

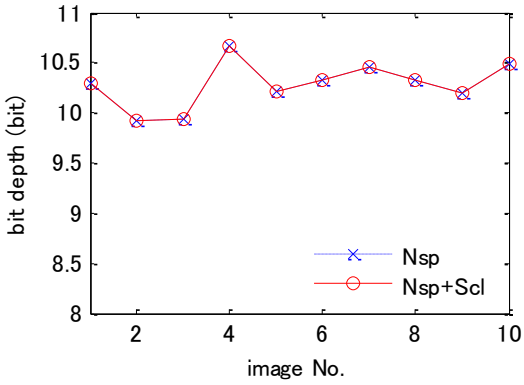


Figure. 79 Bit depth necessary for integer implementation.

The total variance $I(\mathbf{s})$ in (116) is summarized in figure 79 is converted to PSNR as

$$PSNR = 10\log_{10}\frac{255^2}{I(\mathbf{s})} \quad (dB) \tag{132}$$

as the same in (102). PSNR for the existing method (Nsp), was observed to be 48.24 ± 0.11 (dB) in average while that of the proposed method (Nsp+Scl) is improved to 52.27 ± 0.33 (dB). We can observe that the channel scaling reduces variance of the rounding noise by 4.03 (dB) in average.

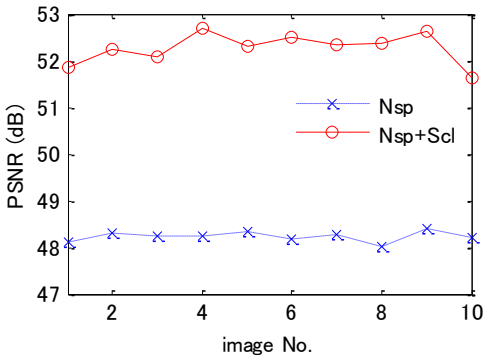


Figure. 80 Variance of the rounding noise measured in PSNR.

5.4.3. Effect of Channel Scaling on Lossy Coding

The effect of the channel scaling over lossy coding of images is examined lastly. The forward transform in the integer implementation in figure 70, in Figure 71 and in Figure 76, are utilized for the encoding of the input images respectively. Then, the backward transform without rounding operations is applied in decoding process, the variance of the difference between the decoded image and the original image is measured in PSNR so that the measured value indicates the total variance (in pixel domain) of the rounding noise generated in the forward transform.

The rate distortion curves are summarized in figure 81. The vertical axis represents the PSNR described above and the horizontal the compressed data volume measured in the bit rate (bpp: bit per pixel) consequently. In lossy coding mode, there is no difference between the methods in low bit rate remarked as '*case A*' in (103).On the contrary, for the high bit rate lossy coding mode, '*case B*' in (103), PSNR is improved by the channel scaling. In case of 'Couple' input image for example, the existing method (Nsp) has approximately 2 (dB) lower PSNR at 5.5 (bpp) comparing to the standard wavelet transform (Sep). However, it was observed that the channel scaling increases PSNR by approximately 5 (dB) in the proposed method (Nsp+Scl).

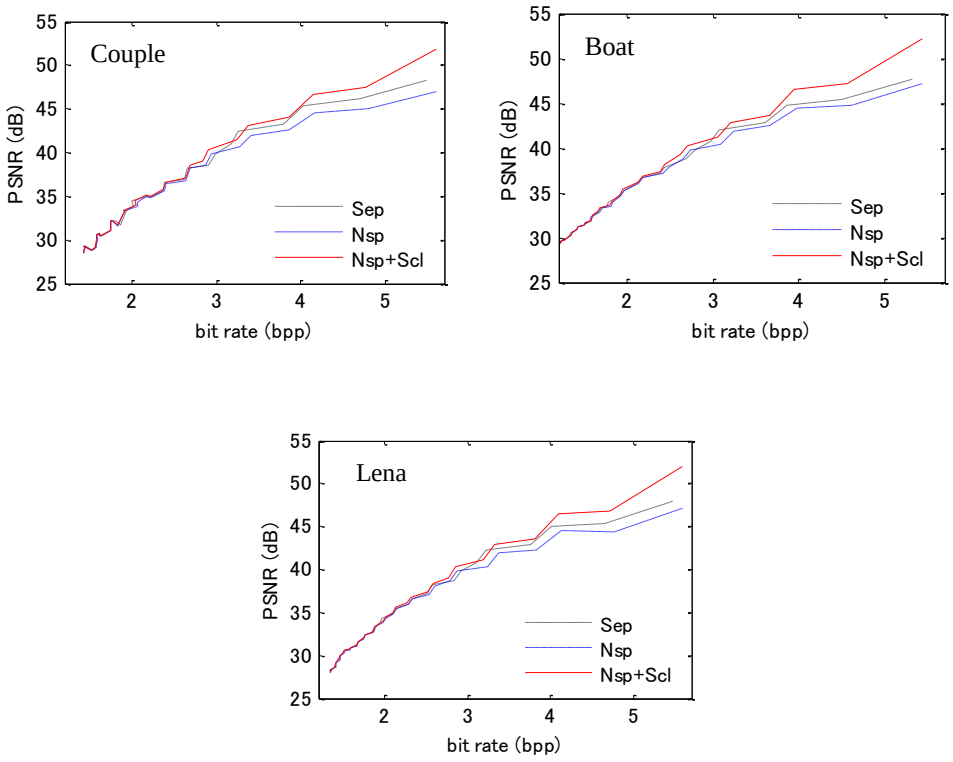


Figure. 81 Effect of channel scaling on lossy coding.

5.5. Summary and Discussion on this Proposal

The channel scaling which was designed for fully utilization of the dynamic range of signal values at each channel of the transform in integer implementation was introduced to the minimum lifting WT based on the non-separable 2D structure. The fact that the variance of the rounding noise was decreased at output of the transform is confirmed in such a case where the rounding noise reduction contributes to improve decoded image quality in high quality (= high bit rate) lossy coding of images.

Chapter 6

6. Word Length Allocation for Multiplier Coefficients of Minimum Lifting Non-Separable 2D Wavelet

6.1. Motivation of this proposal

When the era of that the lifting wavelet transform was adopted by the JPEG 2000 international standard [13], various implementation issues have been discussed on this transform [15,19,48]. In this standard, there are two types of filters 5/3 filter for the lossless coding and 9/7 the lossy coding of images, [19] where each of them are composed of cascaded lifting steps.

According to the cascaded scheme, there is a delay time for waiting the result from the previous lifting step. In that case, the more the lifting steps exist, the more the delay arises. Such, the decrement in the number of lifting steps can minimize the latency of the transform in a parallel processing platform.

However, considering the case of the minimum lifting step structure, the output signal is distorted due to truncation of multiplier coefficient values. As regards, a few coefficients are extremely sensitive to the truncation in this structure while some other coefficients are relatively tolerant. For this reason, a special treatment to each

coefficient is necessary to compensate the implementation cost in the minimum lifting structure.

This chapter deals with the word length of coefficients as reference to an implementation cost [60] which is closely related to the sum-of-power-of-two (SPT) format [61-65]. [62,63] reports designing multiplier-less adaptive filters to be beneficial for. Different number of SPT terms are allocated to each coefficient value under a given total number of SPT terms to reduce implementation cost in [64].

Inspired by [64], we proposes an optimization procedure for dictating the tolerable (the maximum) truncation error (tolerance) for each coefficient of the minimum lifting structure. In this method, the noise gain (sensitivity) of each coefficient is taken into account for an input signal with colored spectrum. In comparison with the tabu-search method in [65], the propose method is simple and stable with satisfactory performance for this case.

In the experiments and discussions, we examined the exploitation of the proposed scheme in lossy coding. The tolerance is assigned to each coefficient so that coding performance of the transform is maintained. It can be well proved that the proposed method can reduce not only the word length cost but also the total number of SPT terms maintaining the same lossy coding performance.

6.2. Existing Method

Nsp 2D minimum lifting structure and its multiplier coefficients to be expressed in SPT format are summarized.

6.2.1. Minimum Lifting Structure

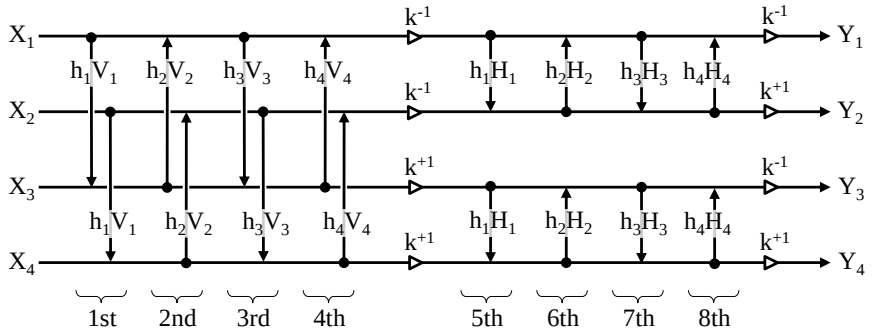


Figure. 82 The separable (Sep) 2D structure of the 9/7 forward wavelet transform in JPEG 2000 standard. This structure has 8 lifting steps.

In Figure 82, the Sep 2D structure of the forward 9/7 wavelet transform is illustrated. Pixel values $x(n_1,n_2)$ at location n_1, $n_2$of the input image is divided into 4 groups as

$$\begin{cases} x_1(\mathbf{m}) = x(2m_1, 2m_2), & x_2(\mathbf{m}) = x(2m_1, 2m_2 + 1), \\ x_3(\mathbf{m}) = x(2m_1 + 1, 2m_2), & x_4(\mathbf{m}) = x(2m_1 + 1, 2m_2 + 1), \end{cases} \tag{133}$$

for $\mathbf{m}=(m_1, m_2)$. For the first step, x_3 is predicted from x_1 as

$$X_3^{(1)}(\mathbf{z}) = X_3(\mathbf{z}) + h_1 V_1(\mathbf{z}) X_1(\mathbf{z}) \tag{134}$$

where

$$X_b(\mathbf{z}) = \sum\nolimits_{\forall m_1, m_2} x_b(m_1, m_2) z_1^{-m_1} z_2^{-m_2} \tag{135}$$

for $c \in \{1,2,3,4\}$, $\mathbf{z}=(z_1,z_2)$ and $V_1(\mathbf{z})=(1+z_1)$. x_4 is predicted from x_2 in the same manner concurrently. In (134), h_1 denotes a multiplier coefficient. Updating in the second step can start after having the result of the first step, As illustrated in figure 82, it has 8 lifting steps in total where

$$\begin{cases} V_1(\mathbf{z}) = V_3(\mathbf{z}) = 1 + z_1, & V_2(\mathbf{z}) = V_4(\mathbf{z}) = 1 + z_1^{-1}, \\ H_1(\mathbf{z}) = H_3(\mathbf{z}) = 1 + z_2, & H_2(\mathbf{z}) = H_4(\mathbf{z}) = 1 + z_2^{-1}, \end{cases} \tag{136}$$

and coefficient values are defined as listed in Table 4 [5].

Table. 4 Coefficient values of the 2D wavelet transform with 9/7 filters.

c	coefficient		c	coefficient	
1	h_1	−1.586134342059924	6	h_{22}	0.0028068929640050
2	h_2	−0.052980118572961	7	h_{33}	0.7795319672951906
3	h_3	0.882911075530934	8	h_{44}	0.1966983278099528
4	h_4	0.443506852043971	9	k^{+2}	1.5133283284009633
5	h_{11}	2.515822151061868	10	k^{-2}	0.6607951369394080

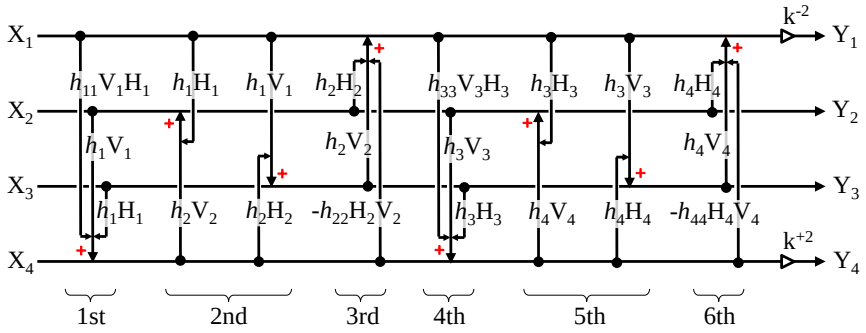


Figure. 83 The non-separable (Nsp) 2D minimum-lifting structure of the forward transform. The number of lifting steps is reduced from 8 to 6.

Figure 83 demonstrates the Nsp 2D structure of the forward transform [39,40]. Contrary to the structure in Figure 82, x_4 is predicted from x_1, x_2 andx_3 in the first step. Then in the next step, x_2 and x_3 are updated from x_1 and $x_4$simultaneously. This minimum lifting structure has total of 6 steps which is reduced from 8 to 6 (75 %). figure 83 illustrates the backward transform in this structure. The forwardly transformed signals Y_1, Y_2, Y_3 and Y_4 are backwardly transformed to reconstruct the original input signals (from right to the left in figure 84). In this chapter, we addresses

on the expressing coefficients of this backward transform in SPT format for reducing computational complexity of the decoder.

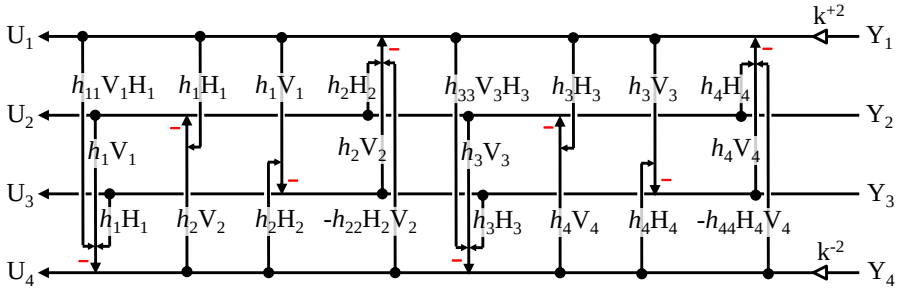


Figure. 84 The backward transform of the non-separable (Nsp) structure. In this paper, coefficients are expressed in the Sum-of-Power-of-Two format.

6.2.2. Truncation of coefficient

In minimization of the computational complexity, coefficient values are first truncated and, then expressed with short word length binary or the SPT format. When an original coefficient value h_c is truncated as (for instance)

$$h'_c = \left\lfloor h_c \cdot 2^{+W_c} \right\rfloor \cdot 2^{-W_c}, \tag{137}$$

The truncated coefficient h'_c in binary expression has W_c [bit] in its fraction part. Hence, the truncation error takes values in the range of

$$\Delta h_c \quad = h_c - h'_c \quad \in [0, 2^{-W_c}) . \tag{138}$$

Accordingly, the shorter the word length W_c is, the larger the truncation error Δh_c becomes. In this chapter, we use the word length as the implementation cost. h_c is

expressed in the SPT format by giving the word length or the tolerable truncation error,. For example, h_1= -(1.100101100000$\cdots$)$_2$is truncated to h'_1= -(1.1001)$_2$ at W_c=4 and 5, respectively (the same result). In SPT format, it is expressed with 3 terms as

$$h'_1 = -(2^0 + 2^{-1} + 2^{-4}) . \tag{139}$$

From [64], it can be regarded as the less number of terms is preferable for low computational cost.

6.3. Proposed Method

In the proposed method, different word length (or number of SPT terms) are assigned to each coefficient and so the word length cost can be lessen under the same coding performance.

6.3.1. Effect of the Truncation on signal

We examine the effect of truncating each coefficient value on the reconstructed signal (output signal of the backward transform) in Figure 69 first. Here, the effect is measured as

$$I_c = \left(\frac{1}{\#\mathbf{M}} \sum\nolimits_{\forall \mathbf{m}} \sum\nolimits_{b=1}^{4} \left\{ u_b(\mathbf{m} \mid c) - x_b(\mathbf{m}) \right\}^2 \right)^{1/2} \tag{140}$$

for $c \in \{1,2,\cdots,N\}$ where #**M** stands for the total number of pixels which indicates the standard deviation (SD) of the noise due to the truncation. In (8), $u_b(\mathbf{m}|c)$ denotes

pixel values of the output signals U_b in Figure 69 in which only the coefficient number c is truncated. When all the coefficients are truncated,

$$I_{\forall c} = \left(\frac{1}{\#\mathbf{M}} \sum\nolimits_{\forall \mathbf{m}} \sum\nolimits_{b=1}^{4} \left\{ u_b(\mathbf{m} \mid \forall c) - x_b(\mathbf{m}) \right\}^2 \right)^{1/2} \tag{141}$$

is used as a substitution of (140). Relation between (140) and (141) is mentioned as

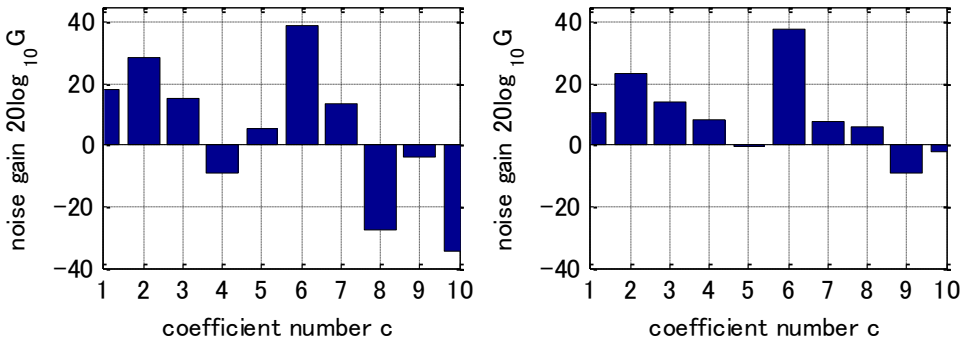
$$I_{\forall c}^2 = \sum\nolimits_{c=1}^{N} I_c^2 + \beta \tag{142}$$

where N for the total number of coefficients. When the noise $u_b(\mathbf{m}|c)$ are uncorrelated each other, β becomes zero.

Second, noise gain $G(c)$ of the coefficient number c is defined as

$$G(c) = I_c \cdot \sigma_x^{-1} \cdot \Delta h_c^{-1} \tag{143}$$

where σ_x denotes SD of pixel values of the input image. In Figure 70, the measured noise gain for a 2D AR(1) model signal with correlation ρ is illustrated. It is observed that the coefficient number 6 has the largest value of G, whereas the number 10 has the least effect for ρ=0.99 case.



(a) ρ=0.99 (b) ρ=0.00

Figure. 85 The noise gain G(c) of the coefficient number c in log scale. The coefficient number 6 has the largest effect on the output signal.

6.3.2. Optimum Word Length Assignment

In the proposed scheme, the tolerance Δh_c in (143) is optically allocated for each coefficient. Then the word length cost can be determined as

$$J = \frac{1}{N} \sum_{c=1}^{N} W_c \ \text{[bit]} \tag{144}$$

where the word length of a coefficient number c is denoted as W_c. The optimization issue discussed in this chapter is described as

$$\min J \quad s.t. \quad I_{\forall c}^2 = \varepsilon \quad \wedge \quad \max\{W_c\} = W \tag{145}$$

for a given word length W. The parameter ε is set to the variance in (142) of the existing method, so that the addressed scheme does not change the lossy coding performance of the existing method. From (142), (143), (144) and $-\log_2(\Delta h_c) > W_c$ from (138), the problem in (145) is expressed with the tolerance Δh_c as

$$\Delta\hat{h}_c = \arg\max_{\Delta h_c} \log_2\left(\prod_{c=1}^{N} \Delta h_c\right)^{1/N} + \lambda_1\left(\sum_{c=1}^{N} G^2(c)\sigma_x^2\Delta h_c^2 + \beta - \varepsilon\right) + \lambda_2\left(\min\{\Delta h_c\} - 2^{-W}\right) \quad (146)$$

where λ_1 and λ_2 the Lagrange multipliers. To alleviate the particular issue, assign the tolerance for each coefficient can be assigned as

$$\Delta h_c = \left(\frac{\max\{G(c)\}}{G(c)}\right)^{\alpha} \cdot 2^{-W}, \quad \alpha \in [0,1] \quad (147)$$

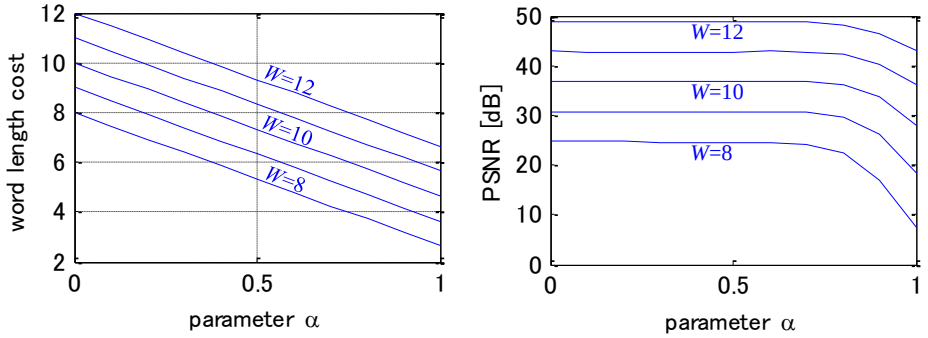
where α is a hyper parameter to be determined experimentally where as α=0 refers to the existing method.

6.4. Experimental Result

The minimization effect of the proposed method upon the word length cost is experimentally confirmed under the same coding performance of the existing method.

6.4.1. Parameter Setting in Channel Scaling

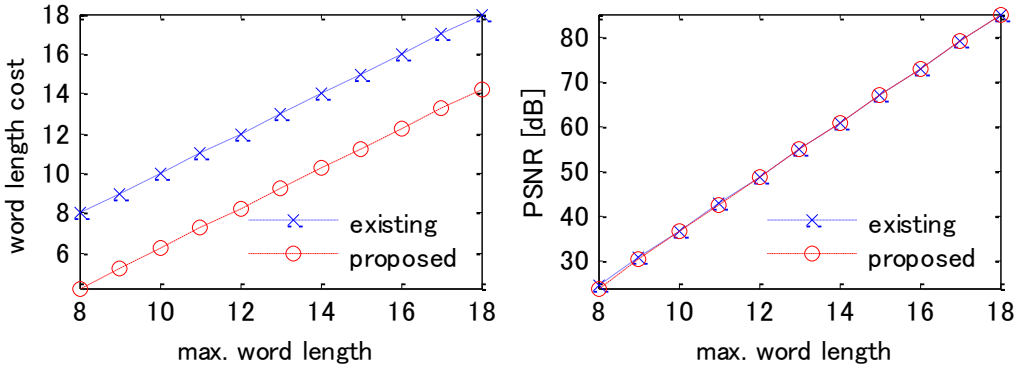
Yet our solution in (147) includes the hyper parameter α, it is simple and stable comparing to a numerical optimization procedure. Figure 86 (a) indicates that the word length cost J decreases as α becomes close to 1 despite that the peak signal to noise ratio:



(a) (b)

Figure. 86 The parameter α>0is determined so that the proposed method has the least length cost under the same PSNR as the existing method with α=0.

$$PSNR = 10(\log_{10} 255^2 - \log_{10} I_{\forall c}^2) \quad [dB] \tag{148}$$



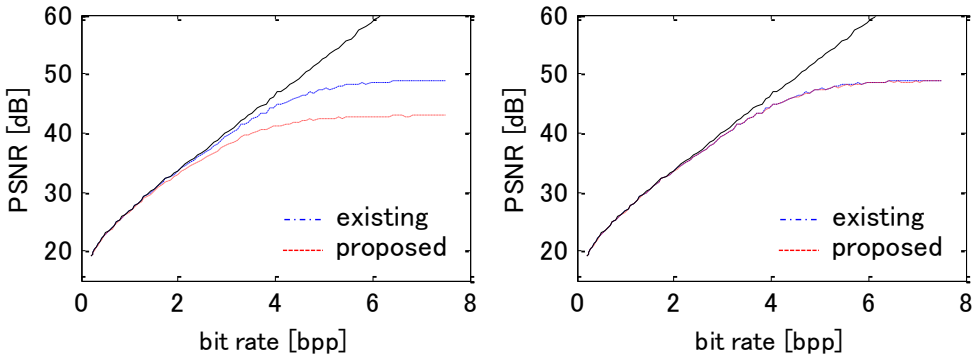
(a) (b)

Figure. 87 The proposed method with α=0.7 has less word length cost under the same PSNR comparing to the existing method with α=0.

Decreases for αgreater than approximately 0.7as indicated in figure 86 (b). Such it is suggested to set the value of α=0.7. The indication of the proposed method decreases the word length cost by approximately 4 [bit] for a given word length W in

Figure 87 (a).Again for case of the proposed method has the same PSNR as the existing method, it is describe in Figure 87 (b).

Figure 88 mentions the rate-distortion curves in lossy coding at W=12. It could be examined that the proposed method keeps coding performance of the existing method under the lower word length cost. Note that the PSNR ceiling (approximately 49 [dB] in Figure 88 (b) is increased by 6 [dB] per 1 [bit] increase of W. It's up to the capacity of the implementation platform.



(a) α = 1.0, ρ = 0.9 (b) α = 0.7, ρ = 0.9

Figure. 88 The proposed method with α = 0.7 has the same performance in the rate-distortion curve as the existing method.

6.4.2. Expression of Coefficients in SPT format

The parameters are set to W=12 and α=0.7 for instance in the following experiments. The assigned word length W_c for each coefficient is summarized in Table 5. Each of them satisfies $W_c < -\log_2(\Delta h_c)$ where the tolerance Δh_c is derived by the assignment in (147). For example, the tolerance of the existing method is $\Delta h_c = 2^{-12}$ for all coefficients. Since $h'_1 = -(1.100101100000)_2$ in 12 [bit] and $h'_1 = -(1.1001011)_2$ in 7

[bit] are the same, W_c is counted as 7 in the Table. It can be proved that the word length cost is reduced to 71.8 [%] by the proposed method. Under these tolerances, the coefficient values are expressed in SPT format. Table 6 summarize the results where Table 7 the total number of the SPT terms which is reduced to 74.5 [%] by the proposed method. It can be regard that the number of terms can be decreased by setting smaller value of the word length W at the cost of lower ceiling of PSNR (illustrated in figure 88).

Table. 5 The word length W_c of each coefficient (example for W=12).

	$-h_1$	$-h_2$	h_3	h_4	h_{11}	h_{22}
Existing	7	12	7	9	6	12
Proposed	7	9	7	4	6	12
	h_{33}	h_{44}	k^{+2}	k^{-2}	total	
Existing	9	12	11	11	96	100 %
Proposed	9	4	7	3	68	71.8%

Table. 6 Coefficients in the SPT format (example for W=12).

	Existing method	Proposed Method
$-h_1$	$2^{0}+2^{-1}+2^{-4}+2^{-6}+2^{-7}$	$2^{0}+2^{-1}+2^{-4}+2^{-6}+2^{-7}$
$-h_2$	$2^{-5}+2^{-6}+2^{-8}+2^{-9}+2^{-12}$	$2^{-5}+2^{-6}+2^{-8}+2^{-9}$
h_3	$2^{-1}+2^{-2}+2^{-3}+2^{-7}$	$2^{-1}+2^{-2}+2^{-3}+2^{-7}$
h_4	$2^{-2}+2^{-3}+2^{-4}+2^{-8}+2^{-9}$	$2^{-2}+2^{-3}+2^{-4}$
h_{11}	$2^{+1}+2^{-1}+2^{-6}$	$2^{1}+2^{-1}+2^{-6}$
h_{22}	$2^{-9}+2^{-11}+2^{-12}$	$2^{-9}+2^{-11}+2^{-12}$
h_{33}	$2^{-1}+2^{-2}+2^{-6}+2^{-7}+2^{-8}+2^{-9}$	$2^{-1}+2^{-2}+2^{-6}+2^{-7}+2^{-8}+2^{-9}$
h_{44}	$2^{-3}+2^{-4}+2^{-7}+2^{-10}+2^{-12}$	$2^{-3}+2^{-4}$
k^{+2}	$2^{0}+2^{-1}+2^{-7}+2^{-8}+2^{-10}+2^{-11}$	$2^{0}+2^{-1}+2^{-7}$
k^{-2}	$2^{-1}+2^{-3}+2^{-5}+2^{-8}+2^{-11}$	$2^{-1}+2^{-3}$

Table. 7 The total number of SPT terms (example for W=12).

	$-h_1$	$-h_2$	h_3	h_4	h_{11}	h_{22}
Existing	5	5	4	5	3	3
Proposed	5	4	4	3	3	3
	h_{33}	h_{44}	k^{+2}	k^{-2}	total	
Existing	6	5	6	5	47	100 %
Proposed	6	2	3	2	35	74.5%

6.5. Summary and Discussion on this Proposal

A simple and stable allocation method of the word length of coefficients of the minimum lifting 2D wavelet transform was proposed and discussed in this chapter. Also, the fact that the proposed method reduces the word length cost keeping the same PSNR as the existing method in lossy coding was experimentally confirmed. For 12 bit maximum word length case, an allocation example was indicated in SPT format. Since the analysis discussed in this is limited to the truncation of coefficients, the rounding of signal values should be also considered as described in [58].

Chapter 7

7. Conclusion of the Dissertation

In this dissertation, we provided 4 proposals as; (1-2) Chapter 3 and 4 for non-separable lifting structure compatibility with 3D data for lossless and lossy compression, (3) Chapter 5 for the rounding noise was reduced by channel scaling method, and (4) Chapter 6, Coefficient noise which was reduced by word length optimization.

In (1) Chapter 3, our findings showed that the 3D non-separable lifting structure for lossless coding is better than the existing separable lifting structure in all ways. The rounding noise in frequency domain and pixel domain was also improved. Moreover, in coding performance, at low bit rate, our proposal is better than existing obviously. Therefore, the lossless 3D non-separable is the best choice for the minimum rounding noise and good coding performance.

In (2) Chapter 4, the non-separable 2D structures for 3D lossy compression, which was found by the 6 rules, was proposed. Comparing in lifting structure to the existing, the total number of lifting steps was increased. However, the rounding noise that caused by integer implementation was minimized and the lossy coding performance was also increased when compared to the other lifting structure.

In (3) Chapter 5, scaling method was proposed for improvement of rounding error of 2D non-separable lifting structure for lossy compression. After applying scaling parameter to lifting structure, we found that the rounding error was improved while the maximum of each channel does not exceed hardware constraint.

In the last, (4) Chapter 6, Word length optimization can reduce the coefficient noise. When $\alpha = 0.7$, word length cost was reduced while a little bit drop PSNR (reconstructed image quality). Worse, when $\alpha = 1.0$, word length cost was reduced more than $\alpha = 0.7$ and more coefficient noise occurred and PSNR was done to drop from $\alpha = 0.7$. Therefore, the parameter $\alpha = 0.7$ is suitable for reduction of the word length while it can maintain PSNR for lossy coding.

The conclusion above can be summarized for advantages and disadvantages as shown in Table 8

Table. 8 Advantage and disadvantage of each proposed method

	Chapter 3	**Chapter 4**	**Chapter 5**	**Chapter 6**
Proposed method	5/3 Non-separable 3D lifting structure	9/7 Non-separable 2D lifting structure	Channel Scaling	Word length optimization
Advantage	Minimum lifting steps and rounding errors	Minimum rounding errors	Rounding errors reduce	Word length cost reduce
Disadvantage	3D memory accessing (Complexity Hardware)	2D memory accessing (Complexity Hardware)	Maximum number in each channel increase	Reconstruction quality a little bit drop

References

[1] F. Sheng, A. Bilgin, P. J. Sementilli, M. W. Marcelling, “Lossy and lossless image compression using reversible integer wavelet transforms,” International Conference on Image Processing, Vol. 3, pp. 876-880, 1998.

[2] B. Shi, L. Liu, C. Xu, “Comparison between JPEG2000 and H.264 for digital cinema,” IEEE International conference on Multimedia and Expo, pp. 725-728, 2008.

[3] M. Unser, T. Blu, “Mathematical properties of the JPEG2000 wavelet filters,” IEEE Transactions on Image Processing, Vol. 12, pp. 1080-1090, 2003.

[4] W. Sweldens, “The lifting scheme: A new philosophy in biorthogonal wavelet constructions”, Wavelet App. in Signal and Image Processing III, pp. 68-79, Proc. SPIE 2569, 1995.

[5] H. G. Feichtinger, “Parseval's relationship for nonuniform samples of signals with several variables,” IEEE Transactions on Signal Processing, Vol. 40, pp. 1262-1263, 1992.

[6] P. Zhao, Y. Liu, J. Liu, R. Yao, S. Ci and H. Tang, “Low-complexity content-adaptive Lagrange multiplier decision for SSIM-based RD-optimized video coding,” IEEE International Symposium on Circuits and Systems, pp. 485-488, 2013.

[7] I. Chattopadhyay and H. Lipson, “Computing entropy rate of symbol sources & a distribution-free limit theorem,” Information Sciences and Systems, pp. 1-6, 2014.

[8] C. Damian and E. Lunca, “A low area FIR filter for FPGA implementation,” Telecommunications and Signal Processing, pp. 521-524, 2011.

[9] Chao He, Dong, J., Zheng, Y. F., Zhigang Gao: Optimal 3-D coefficient tree structure for 3-D wavelet video coding, IEEE Trans. Circuits and Systems for Video Technology, vol.13, Issue 10, pp.961-972 (2003)

[10] Aggoun, A.: Compression of 3D integral images using 3D wavelet transform,

IEEE Journal of Display Technology, vol.7, Issue 11, pp.586-592 (2011)

[11] B. Penna, T. Tillo, E. Magli, G. Olmo: Progressive 3-D coding of hyperspectral images based on JPEG 2000, IEEE Geoscience, Remote Sensing Letters, vol.3, issue 1, pp.125-129 (2006)

[12] 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, 22 (3), pp.459-70 (2003)

[13] ISO / IEC FCD 15444-1, Joint Photographic Experts Group: JPEG 2000 image coding system, (March 2000)

[14] Skodras, A., Christopoulos, C., Ebrahimi, T.: The JPEG 2000 still image compression standard. IEEE Signal ProcessingMagazine, 18, pp.36-58 (2001)

[15] A. Descampe, F. Devaux, G. Rouvroy, J. D. Legat, J. J. Quisquater and B. Macq: A flexible hardware JPEG 2000 decoder for digital cinema, IEEE Trans. Circuits and Systems for Video Technology, vol. 16, issue 11, pp.1397-1410 (Nov. 2006)

[16] Kaneko, K., Ohta, N.: 4K applications beyond digital cinema. In: Proc. Int. Conf. Virtual Syst. Multimedia, pp. 133-136 (2010)

[17] C. Chrysafis, A. Ortega: Line-based, reduced memory, wavelet image compression, IEEE Trans. Image Processing, vol.9, no.3, pp.378-389 (March 2000)

[18] 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 (April 2009)

[19] Bing-Fei Wu, Chung-Fu Lin: A high-performance and memory-efficient pipeline architecture for the 5/3 and 9/7 discrete wavelet transform of JPEG 2000 codec, IEEE Trans. Circuits and Systems for Video Technology, vol.15, no.12, pp.1615-1628 (Dec. 2005)

[20] Siqi Li: The regularity of wavelet transform with the high order vanishing moments, Proc. IEEE International Conf. Measurement, Information and Control

(ICMIC), vol.2, pp.1358-1361 (2013)

[21] Vetterli, M., Herley, C.: Wavelets and filter banks: theory and design," IEEE Trans. Signal Processing, vol.40, Issue 9, pp.2207-2232 (1992)

[22] Zhang, Y., Bull, D.R., Reinhard, E.: Perceptually lossless high dynamic range image compression with JPEG 2000, Proc. IEEE International Conference on Image Processing, pp.1057-1060 (2012)

[23] Shih, Y.S., Zhang, W.C., Sheng, H., et.al.: Bio-inspired JPEG XR CODEC design for lossless HDR biomedical image, Proc. International Computer Symposium, pp. 148-153 (2010)

[24] N. Zhang, X. Wu: Lossless compression of color mosaic images, IEEE Trans. on Image Processing, vol.15, no.6, pp.1379-1388 (June 2006)

[25] PengweiHao,Qingyun Shi: Reversible Integer KLT for progressive-to-lossless compression of multiple component images, IEEE International Conf. on Image Processing (ICIP), vol. 1, pp. 633-636 (Sept. 2003)

[26] M. Iwahashi, M. Ogawa, H. Kiya: Avoidance of singular point in integer orthonormal transform for lossless coding, IEEE Trans. on Signal Processing, vol.60, no.5, pp.2648-2653 (May 2012)

[27] S.Poomrittigul, M. Ogawa, M.Iwahashi, H.Kiya: Reversible color transform for Bayer color filter array images, APSIPA Transactions on Signal and Information Processing, vol.2, pp.1-10 (Sept. 2013)

[28] H. S. Malvar, G. J. Sullivan, S. Srinivasan: Lifting-based reversible color transformations for image compression, SPIE vol. 7073 (2008)

[29] V. Britanak, P. Yip and K. R. Rao: Discrete cosine and sine transform, general properties, fast algorithm and integer approximations, Academic Press (2007)

[30] A. Benazza-Benyahia, J. C. Pesquet and M. Hamdi: Vector-lifting schemes for lossless coding and progressive archival of multispectral images, IEEE Trans. on Geoscience & Remote Sensing, vol.40, issue 9, pp.2011-2024 (Sep. 2002)

[31] S. Chokchaitam, M. Iwahashi: Lossless lossy image compression based on non-separable two-dimensional L-SSKF, IEEE International Symposium, Circuits and Systems (ISCAS), pp.421-424 (May 2002)

[32] D. S. Taubman: Adaptive, non-separable lifting transforms for image compression, IEEE International Conference on Image Processing (ICIP), vol.3, pp.772-776 (1999)

[33] S. Fukuma, M. Iwahashi and N. Kambayashi: Adaptive multi-channel prediction for lossless scalable coding, IEEE International Symposium on Circuits and Systems (ISCAS), no.IV, pp. 467-470 (May 1999)

[34] M. Kaaniche, J. C. Pesquet, A. B. Benyahia, B. P. Popescu: Two-dimensional non separable adaptive lifting scheme for still and stereo image coding, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp.1298-1301 (March 2010)

[35] M. Kaaniche, B. P. Popescu, A. B. Benyahia, J.C. Pesquet: Adaptive lifting scheme with sparse criteria for image coding, EURASIP Journal on Advances in Signal Processing: Special Issue on New Image and Video Representations Based on Sparsity, vol. 2012, pp.1-22 (Jan.2012)

[36] T. Yoshida, T. Suzuki, S. Kyochi, M. Ikehara: Two dimensional non-separable adaptive directional lifting structure of discrete wavelet transform, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.1529-1532 (May 2011)

[37] T. Yoshida, T. Suzuki, S. Kyochi, M. Ikehara: Two dimensional non-separable adaptive directional lifting structure of discrete wavelet transform, IEICE Trans. Fundamentals, vol.E94-A, no.10, pp.1920-1927 (Oct. 2011)

[38] M. Iwahashi, H. Kiya: Non separable 2D factorization of separable 2D DWT for lossless image coding, IEEE International Conference Image Processing (ICIP), pp.17-20 (Nov. 2009)

[39] M. Iwahashi, H. Kiya: A New lifting structure of non separable 2D DWT with

compatibility to JPEG 2000, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), IVMSP, P9.7, pp.1306-1309 (March 2010)

[40] T. Strutz, I. Rennert, Two-dimensional integer wavelet transform with reduced influence of rounding operations, EURASIP Journal on Advances in Signal Processing, vol.2012, 2012:75, ISSN:1687-6180 (April 2012)

[41] M. Iwahashi, H. Kiya: Discrete wavelet transforms: Non separable two dimensional discrete wavelet transform for image signals, ISBN 980-953-307-580-3, InTech (2013)

[42] M. Iwahashi, T. Orachon, H. Kiya: Three dimensional discrete wavelet transform with deduced number of lifting steps, IEEE International Conference on Image Processing (ICIP), no.WA.L4, pp.1651-1654 (Sept. 2013)

[35] Aggoun, A.: Compression of 3D integral images using 3D wavelet transform, IEEE Journal of Display Technology, vol.7, Issue 11, pp.586-592 (2011)

[36] M. Iwahashi, H.Kiya, Discrete wavelet transforms, InTech, ISBN 978-953-307-313-2, Chapter 14, Condition on Word Length of Signals and Coefficients for DC Lossless Property, pp.231-254, Sept. 2011.

[37] S. Chokchaitam, M. Iwahashi, Lossless / lossy image compression based on non-separable two- dimensional LWT", Proc. International Conference on Circuits / Systems Computers and Communications (ITC- CSCC 2002), pp. 912-915, (July 2002)

[38] M. Iwahashi, T. Orachon, H. Kiya, Non separable 3D lifting structure compatible with separable quadruple lifting DWT, Asia-Pacific Signal and Information Processing Association 2013 Annual Summit and Conference (APSIPA), OS.26, IVM.11, no.4, pp.1-4 (Oct. 2013)

[39] W. Jiang and A. Ortega, “Lifting Factorization-Based Discrete Wavelet Transform Architecture Design,” IEEE Trans. Circuits and Systems for Video Technology, vol.11, no.5, pp.651-657, May 2001.

[40] M. Grangetto, et.al., "Optimization and Implementation of the Integer Wavelet

Transform for Image Coding," IEEE Trans. Image Processing, vol.11, Issue 6, pp. 596-604, June 2002.

[41] A. M. Reza, Lian Zhu, "Analysis of Error in the Fixed-point Implementation of Two-dimensional Discrete Wavelet Transforms," IEEE Trans. circuits and systems, fundamental theory and applications, vol.52, issue 3, pp.641-655, March 2005.

[42] M. D. Adams and F. Kossentini, "Reversible Integer-to-integer Wavelet Transforms for Image Compression: Performance Evaluation and Analysis," IEEE Trans. on Image Processing, 9 (6), pp. 1010-1024, 2000.

[43] M. Iwahashi, O. Nishida, S. Chokchaitam, N. Kambayashi, "Optimum Word Length Allocation for Multipliers of Integer DCT," Proc. IEEE International Symposium on Circuits and Systems (ISCAS), Vol. 2, II, pp.400 -403, May 2003.

[44] S. Chokchaitam, M. Iwahashi, N. Kambayashi, "Optimum Word Length Allocation of Integer DCT and Its Error Analysis," Signal Processing: Image Communication, Vol. 19, Issue 6, pp.46 5-478, July 2004.

[45] S. Chokchaitam, M. Iwahashi, S. Jitapunkul, "A New Unified Lossless/Lossy Image Compression Based on A New Integer DCT," IEICE Trans. on Information and Systems, Vol. E88-D, Issue 7, pp.1598 -1606, July 2005.

[46] M. Iwahashi, “Four Band Decomposition Module with Minimum Rounding Operations,” IET Electronics letters, vol.43, no.6, pp.333-335, March 2007.

[47] T. Orachon, S. Poomrittigul, T. Yoshida, M. Iwahashi, S. Chokchaitam, "Non-separable 3D Integer Wavelet Transform for Lossless Data Compression," Science Journal of Circuits, Systems and Signal Processing, 3(6), pp.35-46, Jan.20, 2015.

[48] S. Poomrittigul, M. Iwahashi, H. Kiya, "Reduction of Lifting Steps of Non Separable 2D Quadruple Lifting DWT Compatible with Separable 2D DWT," IEICE Trans. Vol.E97-A, No.7,pp.1492-1499, Jul. 2014.

[49] M. Iwahashi, T. Orachon, H. Kiya, “Non Separable 3D Lifting Structure Compatible with Separable Quadruple Lifting DWT”, Asia-Pacific Signal and Information Processing Association 2013 Annual Summit and Conference

(APSIPA), OS.26, IVM.11, no.4, pp.1-4, Oct. 2013.

[50] T. Suzuki, M. Ikehara, and T. Nguyen, “Generalized Block-lifting Factorization of M-channel Biorthogonal Filter Banks for Lossy-to Lossless Image Coding,” IEEE Trans. Image Process., vol.21, no.7, pp.3220–3228, 2012.

[51] T. Suzuki and H. Kudo, "Two-dimensional Non-separable Block Lifting-based M-channel Biorthogonal Filter Banks," Proc. Signal Processing Conference (EUSIPCO), pp. 291- 295, Sept. 2014.

[52] V. Sandhiya, S. Karthick, M. Valarmathy, "A survey of new reconfigurable architectures for implementing FIR filters with low complexity," Int'l Conf. Computer Communication, Informatics (ICCCI), pp.1-9, 2014.

[53] S. C. Chan, P. M. Yiu, "A multiplier-less 1-D and 2-D fast Fourier transform-like transformation using sum-of-powers-of-two (SOPOT) coefficients," IEEE Int'l Symposium Circuits, Systems (ISCAS), vol.4, pp.755 -758, 2002.

[54] S. Choudhary, P. Mukherjee, M. Chakraborty, "A SPT treatment to the bit serial realization of the sign-LMS based adaptive filter," IEEE Int'l Symposium, Circuits, Systems (ISCAS), pp.2678 -2681, 2010.

[55] A. Chandra, S. Chattopadhyay, "Efficient encoding of powers-of-two coefficients through minimum index floating point representation (MIFPR)," Int'l Conf. Control, Instrumentation, Energy, Communication (CIEC), pp.650 -653, 2014.

[56] Y. C. Lim, R.i Yang, Li Dongning, J. Song, "Signed power-of-two term allocation scheme for the design of digital filters," IEEE Trans. Circuits, Systems II, Vol.46, Issue 5, pp.577 -584,1999.

[57] S. Traferro, A. Uncini, "Power-of-two adaptive filters using tabu search," IEEE Trans. Circuits, Systems II, Vol. 47, Issue 6. pp.566 -569, 2000.

[58] M. Iwahashi, H. Kiya, "Word Length Condition for DC Lossless DWT", Asia-Pacific Signal and Information Processing Association 2009 Annual Summit and Conference (APSIPA), no.TA-P2-6, pp.469-472, Oct.2009.

List of Publications

(a) Journal papers

[1] T. Orachon, S. Poomrittigul, T. Yoshida, M. Iwahashi, S. Chokchaitam, "Non-separable 3D Integer Wavelet Transform for Lossless Data Compression," Science Journal of Circuits, Systems and Signal Processing (SJCSSP), 3(6), pp.35-46, Jan., 2015.

[2] T. Orachon, S. Poomrittigul, T. Yoshida, M. Iwahashi, S. Chokchaitam, "Integer Implementation of 3D Wavelet Transform for Lossy Data Compression," International Journal of Emerging Technology and Advanced Engineering (IJETAE), Vol. 5, Issue 1, pp.17-25, Jan. 2015.

[3] T. Orachon, T. Yoshida, S. Chokchaitam, M. Iwahashi, H. Kiya, “Channel Scaling for Integer Implementation of Minimum Lifting 2D Wavelet Transform,” Institute of Electronics, Information and Communication Engineers (IEICE) Trans. on Fundamentals, Vol.E99-A, No.7, July 2016.

[4] T. Orachon, T. Yoshida, S. Chokchaitam, M. Iwahashi, H. Kiya, “Word Length Allocation for Multiplier Coefficients of Minimum Lifting Non-Separable 2D Wavelet Transform,” IEEJ Transactions on Electronics, Information and Systems vol. 136, no. 9, Sept. 2016.

(b) Conference papers

[1] M. Iwahashi, T. Orachon, H. Kiya, “Three Dimensional Discrete Wavelet Transform with Deduced Number of Lifting Steps”, Proc. IEEE International Conference on Image Processing (ICIP), no.WA.L4, pp.1651-1654, Sept. 2013.

[2] T. Orachon, M. Iwahashi, "Compression of Medical Volumetric Data using Lossless Non-separable Discrete Wavelet Transform”, Proc. ICEEHE, pp.73, Dec., 2013.

[3] M. Iwahashi, T. Orachon, H. Kiya, “Non Separable 3D Lifting Structure Compatible with Separable Quadruple Lifting DWT”, Asia-Pacific Signal and Information Processing Association 2013 Annual Summit and Conference (APSIPA), OS.26, IVM.11, no.4, pp.1-4, Oct. 2013.

[4] M. Iwahashi, T. Orachon, T. Moe Aye, T. Yoshida and H. Kiya,"Optimum SPT Allocation for Multipliers of Minimum Lifting 2D Wavelet Transform", IEEE International Symposium on circuits and systems (ISCAS), May. 2016.

Acknowledgement

I would like to express the deepest appreciation to my supervisor, Professor Iwahashi Masahiro for his evaluation, inestimable suggestion, useful critiques, valuable guidance, and supervision thoroughly my doctoral research.

My grateful thanks are also extended to Japanese government and Pathumwan Institute of Technology, Bangkok, Thailand, for Mobukagakusho by financial supported my research in Nagaoka University of Technology since September 2012.

I also would like to acknowledge the help provided by Professor Nakagawa Kenji, Assistant Professor Taichi Yoshida and all Dissertation Committees for their patience, enthusiasm, motivation, and immense knowledge. I would not have been able to succeed this dissertation without their guidance, which has helped me throughout my study. In addition, I am particularly delightful to represent my gratitude to Dr. Suvit Poomrittikul, Miss Fairoza Amira Binti Hamzah, Mr. Kentaro Yamaura, and Mr. Akira kondo for their devotion on mentoring me how to perform several laboratory procedures and finish study here.

I also would like to express my sincere gratitude to Miss Tin Moe Aye and Miss Suteera Vibulyaseck for their encouragement and generous English editing on the dissertation. This dissertation would not have been accomplished without their advice and invaluable help. Finally, I would like to give a thankful regard to my parents and friends for their love, support, and enjoyable moments throughout the research.