Channel Scaling for Rounding Noise Reduction in Minimum Lifting 3D Wavelet Transform

Fairoza Amira Binti Hamzah*, Taichi Yoshida*, Masahiro Iwahashi* and Hitoshi Kiya†
* Nagaoka University of Technology, Nagaoka, Japan
† Tokyo Metropolitan University, Tokyo, Japan

Abstract— **An integer transform is used in lossless-lossy coding since it can reconstruct an input signal without any loss at output of the backward transform. Recently, its number of lifting steps is reduced as well as delay from input to output introducing multi-dimensional memory accessing. However it has a problem that quality of the reconstructed signal in lossy coding has its upper bound in the rate distortion curve. This is because the noise generated by rounding operations in each lifting step inside the integer transform does not contribute to data compression. This paper tries to reduce the rounding noise observed at output of the integer transform introducing channel scaling inside the transform. As a result of experiments, it was observed that the proposed method improves quality of the decoded signal in lossy coding mode.**

I. INTRODUCTION

This paper deals with a picture quality degradation problem due to integer implementation of a three dimensional (3D) lifting wavelet transform (WT). In an integer implementation, signal values inside the transform are rounded to integers. Shorter word length of the integer contributes to smaller memory space and faster calculation [1,2]. However it generates noise due to the rounding (rounding noise).

In the lifting WT, the rounding noise does not appear at output of the backward transform as far as the output signal values of the forward WT are fed into the backward WT without any alteration. Therefore it is applied to lossless coding of signals. Lossless coding based on the 5/3 WT in the JPEG 2000 is an example for this case [3]. However the rounding noise appears at output of the backward WT in lossy coding in which the quantization and its inverse are inserted between the forward WT and the backward WT. Lossy coding based on the 9/7 WT is an example for this case [3].

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. This paper tries to reduce the rounding noise of the minimum lifting WT by introducing the channel scaling to increase the upper bound.

Recently, based on non-separable structure [4,5], the minimum lifting WT which has less number of lifting steps in cascade comparing to the conventional separable structure has been reported [6-11]. It contributes to reduce delay from input to output under a parallel signal processing architecture. This is because a lifting step must wait for a calculation result of the previous lifting step.

The minimum lifting WT based on the 5/3 WT was been reported for 2D signals in [6]. It was extended to 3D signals in [7]. In both cases, the total number of lifting steps was decreased as well as the variance of the rounding noise. Similarly, the minimum lifting 2D WT based on the 9/7 WT was reported in [8-10]. It was extended to 3D case in [11]. However, unlike the 5/3 WT, the variance of the rounding noise is increased in the 9/7 WT even though the number of lifting steps is decreased.

This paper reduces the rounding noise of the minimum lifting WT based on the 9/7 WT for 3D case by introducing channel scaling. The 3D WT has eight channels. In all channels, the maximum absolute value (MAV) of signals inside the transform is limited by word length of signals in its integer implementation. Using a fact that MAV in each channel is not the same, we introduce different scaling in each channel. The channel scaling is designed so that 1) each channel has almost the same MAV inside the transform, 2) output signal values of the transform are not changed, and 3) rounding noises generated inside the transform are reduced at output of the transform.

II. MINIMUM LIFTING WAVELET

Figure 1 illustrates the forward transform of the 3D WT based on the "separable" structure. The input signal

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

where $(\mathbf{z})=(z_1, z_2, z_3)$ and $(\mathbf{n})=(n_1, n_2, n_3)$ is decomposed into 8 subgroups $X_{000}, X_{001}, \cdots, X_{111}$ depending on parity of the pixel location $(\mathbf{n})$. In its integer implementation, signal values are multiplied with 2^F and rounded to integers. In the 1st lifting step, X_{100} in channel (Ch) 5 is predicted from X_{000} in Ch1 with a filter V_1. Note that signal values of the prediction are rounded to integers. Similarly, Ch6, Ch7 and Ch8 are predicted from Ch2, Ch3 and Ch4, respectively. There predictions are performed simultaneously under a parallel processing platform. However the updating in the 2nd lifting must wait for calculation results of the 1st lifting step. It means that this structure takes long delay from input to output since it has 12 lifting steps.

Figure 2 illustrates the forward transform of the 3D WT based on the “non-separable” structure [11]. In this structure, the total number of the lifting steps is reduced from 12 to 8. It contributes to decrease the delay. If there is no rounding, both of Fig. 1 and Fig. 2 have the same output for the same input. However, if F is very small, they have different amount of the rounding noise at their output. Unfortunately, Fig. 2 has greater amount of the rounding noise even though the number of lifting steps is decreased comparing to Fig. 1. This paper tries to reduce the variance of the rounding noise at output of the transform by introducing the channel scaling explained in the next section.

Note that the filters V_1, V_2, V_3, V_4 are defined as

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

Similarly, other filters are defined as

$$\begin{bmatrix} H_p(\mathbf{z}) \\ H_q(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_p(1+z_2^{+1}) \\ h_q(1+z_2^{+1}) \end{bmatrix}, \quad \begin{bmatrix} D_p(\mathbf{z}) \\ D_q(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} h_p(1+z_3^{+1}) \\ h_q(1+z_3^{+1}) \end{bmatrix}, \tag{3}$$

where the filter coefficients h_p and h_q are defined in the JPEG 2000 as the 9/7 WT as well as the coefficient k in the figure.

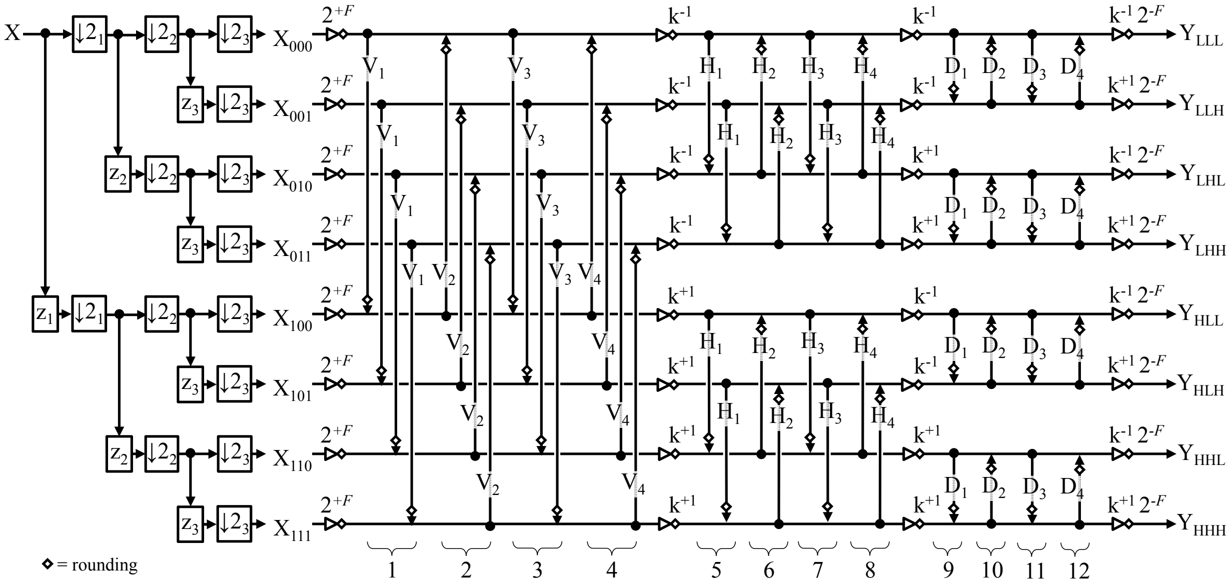

Fig. 1 Separable 3D structure

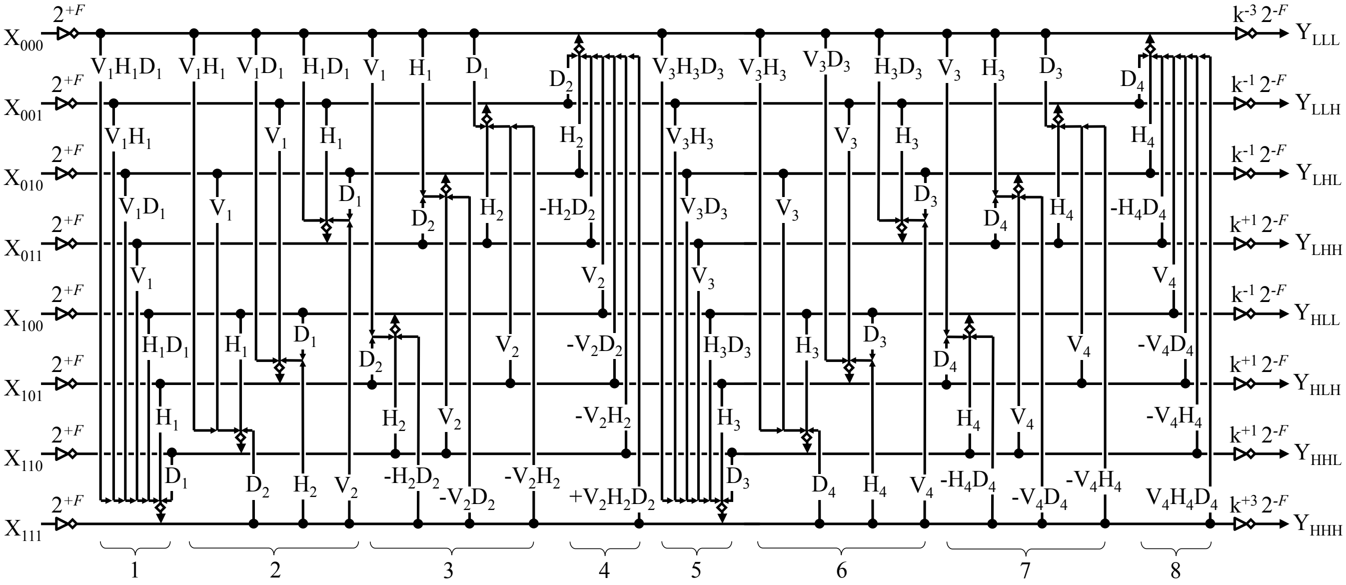

Fig. 2 Non-separable 3D structure without channel scaling (existing method).

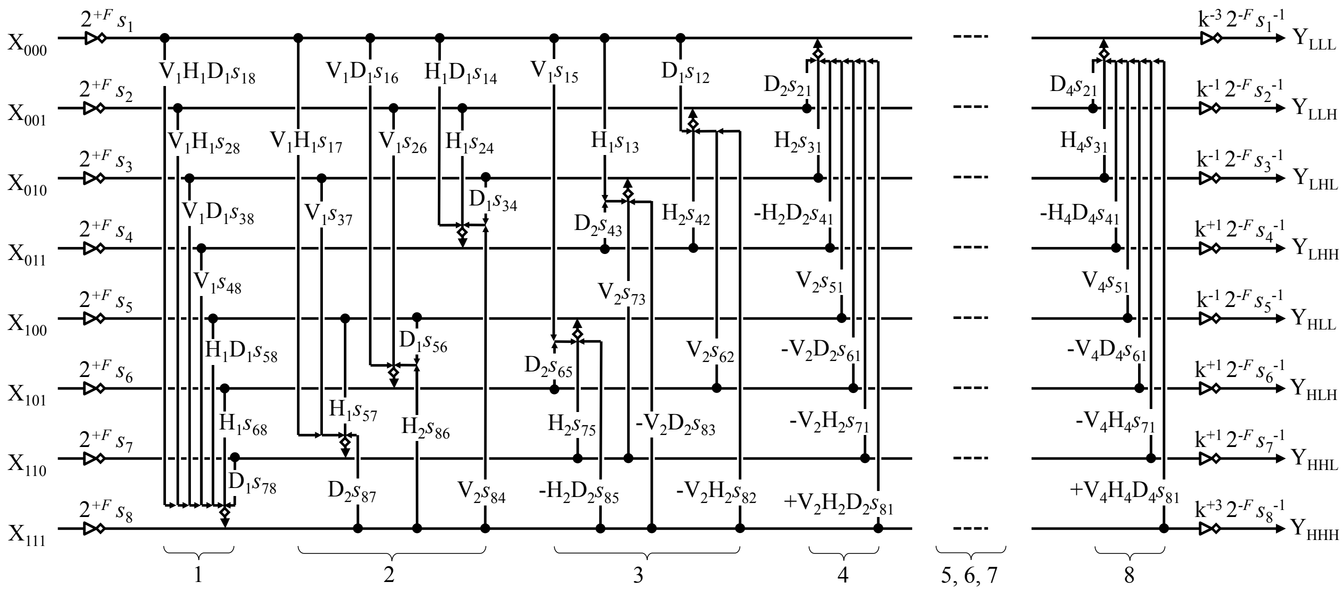

Fig. 3 Non-separable 3D structure with channel scaling (proposed method).

III. PROPOSED METHOD

Figure 3 illustrates the proposed method. It has the minimum number of the lifting steps as the same as the existing method in Fig. 2. Unlike the existing method, the scaling parameters $s_1, s_2, \cdots, s_8$ are introduced so that each channel has almost the same maximum absolute value (MAV). If there is no such scaling, each channel has different MAV. It means that the dynamic range for signal values of the integer implementation is not fully utilized inside the transform. Fig. 3 also has other parameters

$$s_{pq} = s_q / s_p \quad for \quad p,q \in \{1,2,\cdots,8\}. \tag{4}$$

Those are introduced so that output signal values of the transform becomes the same as the existing method if there is no rounding noise.

In the proposed method, signal values are multiplied with different parameters s_p in each channel before the 1st lifting step. In the inter channel prediction and up-dating, scales are balanced with other parameters s_{pq}. Finally, both of the signal value and the rounding noise are divided with the parameters s_p in each channel. As a result, the rounding noise generated inside the transform is reduced at output of the transform without changing the signal values.

In detail, MAV m_p in channel $p \in \{1,2,\cdots,8\}$ of the existing method are measured for an input signal. Next, the parameters s_p are set as

$$s_p = \max\{m_1, m_2, \cdots, m_8\} \cdot m_p^{-1}. \tag{5}$$

As a result, the parameters of this channel scaling becomes greater than or equal to one. In contrast, the existing method sets all parameters to one.

IV. EXPERIMENTAL RESULTS

In the following experiments, a 3D AR(1) model was used as the input signal to the forward WT. Its auto-correlation coefficient was set to $\rho = 0.9$. The range of signal values is set to [-128, 127], namely 8 bit depth.

Figure 4 summarizes MAV in each channel at F=0. Before the channel scaling (scl_OFF), each channel had different MAV. After, the channel scaling (scl_ON), all channels have almost the same MAV (=1024). It means that the dynamic range [-1024, 1023] of signal values at each channel in 11 bit-depth integer implementation is fully utilized.

Figure 5 summarizes the variance of the rounding noise measured at output of the forward WT in each channel at F=0. For example, the variance is reduced from 6.1 to 1.62 in channel 8. In average, it is reduced from 3.13 to 0.56. it was observed that the rounding noise is reduced to 17.9 (%) at output of the forward WT by the channel scaling.

Figure 6 illustrates the rate-distortion curves. The horizontal axis and the vertical axis indicates the bit rate (= compressed data volume) and the peak signal to noise ratio (PSNR), respectively. Both of the forward and the backward transforms are implemented as integer transforms at F=0. After the forward transform, the quantization and the entropy coding are embedded to compress its data volume. It was observed that quality of decoded signals measured with PSNR is improved from 43.0 (dB) to 46.6 (dB) at 4.14 (bpp). Quality of decoded signals in high quality (= high bit rate) lossy coding is improved by the proposed method.

Figure 7 indicates the bit depth = $\log_2$(MAV) +1 +F (bit) versus PSNR of the output signal from the backward WT at F=8 and the forward WT at F=0. There is no quantization between them. It was observed that the proposed method reduces the rounding noise by 4 (dB).

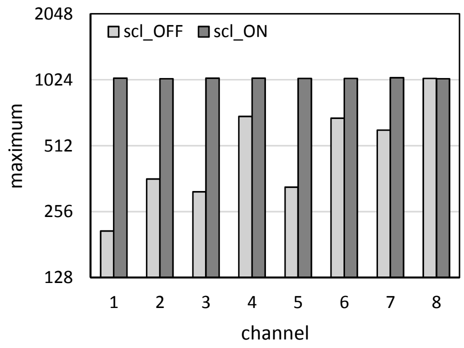

Fig. 4 Maximum absolute value (MAV) in each channel.

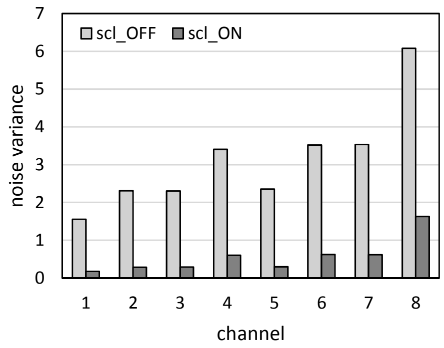

Fig. 5 Variance of the rounding noise in frequency domain.

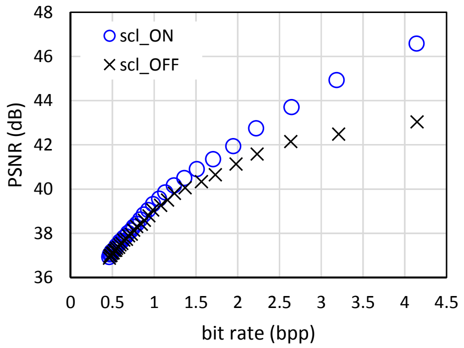

Fig. 6 The rate-distortion curves in lossy coding mode.

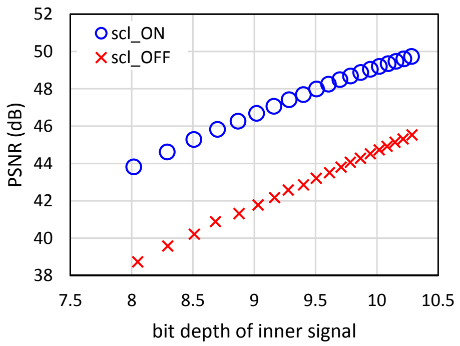

Fig. 7 Quality of decoded signals versus word length of integers.

V. CONCLUSIONS

The channel scaling was introduced to the minimum lifting WT based on the non-separable 3D structure. It was designed so that the dynamic range of signal values at each channel of the transform in integer implementation is fully utilized. It was confirmed that the variance of the rounding noise was decreased at output of the transform. Analysis on rounding should be extended to multiplier coefficients as in [12], since it is limited to signal values in this paper.

This work was supported by JSPS KAKENHI Grant Number 26289117.

REFERENCES

[1] 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.

[2] 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.

[3] ISO / IEC FCD 15444-1, Joint Photographic Experts Group, “JPEG 2000 Image Coding System,” March 2000.

[4] M. Kaaniche, J. C. Pesquet, A. B. Benyahia and 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.

[5] S. Muramatsu, T. Kobayashi, M. Hiki, H. Kikuchi, “Boundary Operation of 2-D Non-separable Linear-phase Paraunitary Filter Banks,” IEEE Trans. on Image Processing, DOI 10.1109/ TIP. 2011.2181527, Apr. 2012.

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

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

[8] T. Strutz and 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.

[9] M. Iwahashi, and H. Kiya, “Non Separable 2D Factorization of Separable 2D DWT for Lossless Image Coding,” IEEE International Conference on Image Processing (ICIP), pp. 17-20, Nov. 2009.

[10] 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.

[11] 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.

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