International Symposium on Intelligent Signal Processing and Communication Systems (**ISPACS**), TA2, B-5, pp.264-267, Dec. 2009, Kanazawa, Japan.

Bit Rate Reduction of Enhancement Layer in Bit-Depth Scalable Coding

Hiroshi KIKUCHI, Wataru OTAKE and Masahiro IWAHASHI
Nagaoka University of Technology, Nagaoka 980-2188 Niigata Japan
E-mail: {hkikuchi, wotake}@tech.nagaokaut.ac.jp
Tel: +81-258-46-6000

Abstract— **In this report, we propose a bit-depth scalable coding system based on the integer wavelet transform. Encoder of the system outputs two layered bit-streams. Decoding the base layer, upper part of the bit planes of each pixel values of an input image signal are displayed. Adding the enhancement layer, all the bit planes are reconstructed without any loss. Since quantization errors in the transformed domain are amplified by the backward transform in the local decoding, we move the quantization to the spatial domain. As a result, the errors are not amplified and the bit rate of the enhancement layer is reduced by 8.4 [%] on average over some image signals.**

I. INTRODUCTION

When a pixel value of an image signal is expressed with 8 bit integer, it is possible to have 256 kinds of intensities. This is called the bit depth (BD). Recently, the BD is extended to 12 [bit] and more for high dynamic range imagery [1]. For this variation of the BD in image systems, the BD scalable coding (SC) has been attracting researchers' attention [2,3].

Based on the discrete cosine transform (DCT) and the motion compensation, the H.264 is extended to the BD scalable video coding (SVC) [4,5]. The system outputs two layered bit-streams. Decoding the base layer, upper part of the bit planes of each pixel of an input image signal are displayed. Adding the enhancement layer, all the bit planes are reconstructed.

In this report, we propose a BD-SC based on the integer discrete wavelet transform (DWT) in the JPEG 2000 [6,7]. Especially, we reduce the bit rate of the enhancement layer by introducing the quantization in the spatial domain [8].

The DCT in the H.264 based BD-SVC is orthonormal. On the contrary, the lifting based DWT in the JPEG 2000 is bi-orthogonal. In this case, the energy gain of the DWT to the white noise is greater than unity. Therefore, the variance of the quantization error is amplified by the backward DWT in the local decoding process.

To avoid this, we perform the quantization before the forward DWT (in the spatial domain). As a result, the variance of the locally decoded signal is reduced. In addition, the rounding error inside the DWT is cancelled. Therefore, the bit rate of the enhancement layer is reduced.

II. BIT DEPTH SCALABLE CODING

A. BD-SC and its Evaluation Criteria

The encoder of the BD-SC in Fig.1 (a) inputs an original signal x_{org} with bit-depth N_{org} [bit]. It outputs the base layer which contains the compressed data of the base layer signal x_{bas} with shortened bit-depth N_{bas} [bit].

Decoding only the base layer by the decoder in Fig.1 (b), the base layer signal is reconstructed for a display with short bit-depth. Adding the enhancement layer data, the original full bit-depth signal x_{org} is displayed without any loss.

In the encoder, the base layer signal x_{bas} is generated from the original signal x_{org} by

$$x_{bas} = F[2^{-\Delta N} x_{org}] \tag{1}$$

where

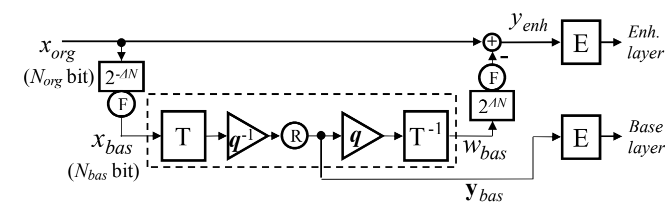
$$F[x] = \lfloor x \rfloor = x - (x \bmod 1)\,,\ \Delta N = N_{org} - N_{bas} > 0\,. \tag{2}$$

After the forward transform T detailed in Fig.2 and the quantization, the band signal vector $\mathbf{y}_{bas}$ is produced by

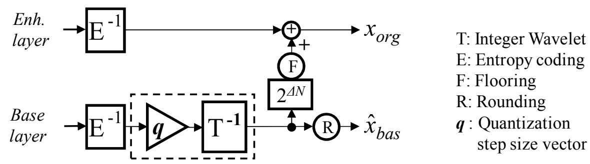
$$\mathbf{y}_{bas} = R[\mathbf{q}^{-1} \cdot \mathrm{T} \circ x_{bas}] \tag{3}$$

where

$$R[x] = \lfloor x + 2^{-1} \rfloor,\ \mathbf{q}^{-1} = diag[q_0^{-1} \quad q_1^{-1}]\,. \tag{4}$$



(a) Encoder



(b) Decoder

Fig.1 The existing bit depth scalable coding (BD-SC).

In this report, we evaluate the system with the following three parameters. The two are the bit rate H_{bas} and H_{enh} of the base layer and the enhancement layer, respectively. The other is the image quality θ_{bas} of the base layer signal $\hat{x}_{bas}$.

B. Bit Rate of the Enhancement Layer

The bit rate we reduce in this report is related to the variance of the signal y_{enh} in Fig.1 (a) by

$$H_{enh} = 2^{-1}\log_2 \gamma_{y_{enh}} \sigma^2_{y_{enh}} \tag{5}$$

where $\gamma_{y_{enh}}$ is determined by the probability density function of y_{enh}. Since the variance of y_{enh} is proportional to that of the locally decoded signal w_{bas} in Fig.1 (a), we focus on analysis on the error in w_{bas}.

Fig.3 indicates an analysis model of the signal processing in the broken line in Fig.1 (a). Two band decomposition case is illustrated as an example. The signal w_{bas} contains two kinds of errors. One is the rounding error e_R which is generated by the rounding of signals into integers inside the transform in Fig.2. Note that the error is accumulated through T, quantization and T $^{-1}$. The other is the quantization error $\mathbf{e}_Q=[e_{q0}\ e_{q1}]$. Note that its variance is multiplied by the L^2 norm of the synthesis filters $\mathbf{G}$. Namely,

$$\left\|\mathbf{G}\circ\mathbf{e}_Q\right\|^2 = 2^{-1}\cdot\left\langle \left\|\mathbf{G}\right\|^2, \left\|\mathbf{e}_Q\right\|^2 \right\rangle \tag{6}$$

where

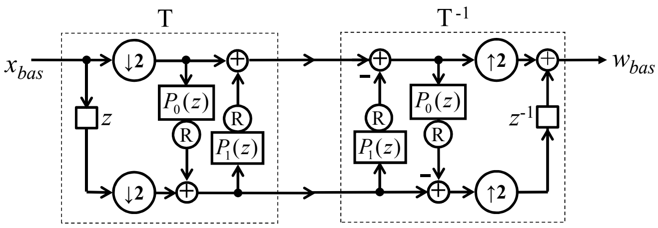


Fig.2 The integer 5/3 DWT [7]. T: forward, T $^{-1}$: backward.

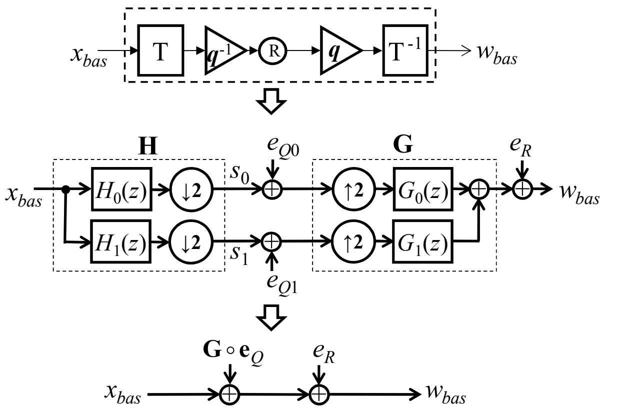


Fig.3 A model for analysis on the existing method.

$$\left\|\mathbf{G}\right\|^2 = \left[\left\|G_0\right\|^2 \quad \left\|G_1\right\|^2\right], \quad \left\|\mathbf{e}_Q\right\|^2 = \left[\sigma^2_{e_{Q0}} \quad \sigma^2_{e_{Q1}}\right]$$

$$\left\|G_n\right\|^2 = \frac{1}{2\pi j}\oint G_n(z)\overline{G_n(z)}z^{-1}dz, \quad n\in\{0,1\}$$

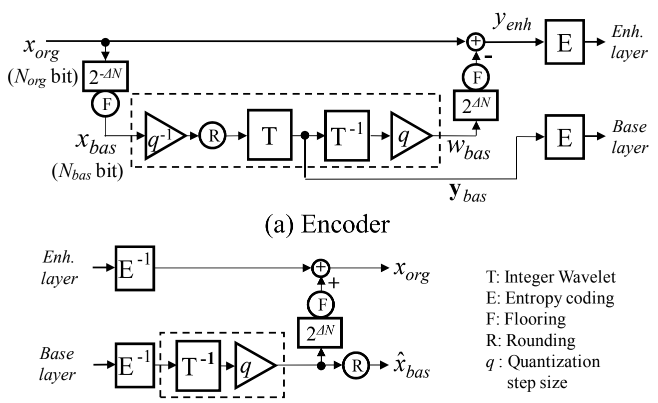
Since the DCT is orthonormal, its L^2 norm is unity. However, that of the bi-orthogonal DWT is greater than unity. Therefore the variance of the quantization error is amplified by the backward DWT in the local decoding process.

As a result, the bit rate of the enhancement layer H_{enh} is increased by replacing the DCT by the bi-orthogonal DWT.

III. Proposed BD-SC based on 5/3 DWT

To reduce the bit rate of the enhancement layer, we introduce the quantization in spatial domain as illustrated in Fig.4. The transform and the quantization in Fig.1 are bartered.

Fig.5 indicates an analysis model of the signal processing in the broken line in Fig.4 (a). In this case, the signal w_{bas} contains only one error. Note that the rounding error e_R inside the transform is cancelled between T and T $^{-1}$ since the transform is lossless. The only one error is the quantization error e_Q. Also note that its variance is not multiplied by the L^2 norm of the synthesis filters $\mathbf{G}$.



(a) Encoder

(b) Decoder

Fig.4 The proposed BD-SC.

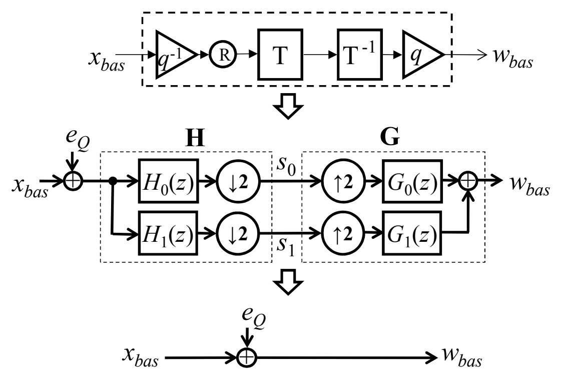


Fig.5 A model for analysis on the proposed method.

In summary, the bit rate of the enhancement layer of the existing method is described by

$$H_{enh} = 2^{-1}\log_2 \gamma\left(\sigma_{e_F}^2 + \left\|\mathbf{G}\circ\mathbf{e}_Q\right\|^2 + \left\|e_R\right\|^2\right) + \Delta N \qquad (7)$$

and that of the proposed method is

$$H_{enh} = 2^{-1}\log_2 \gamma\left(\sigma_{e_F}^2 + \left\|e_Q\right\|^2\right) + \Delta N \qquad (8)$$

where

$$e_F = F\left[2^{-\Delta N} x_{org}\right] - 2^{-\Delta N} x_{org} \ . \qquad (9)$$

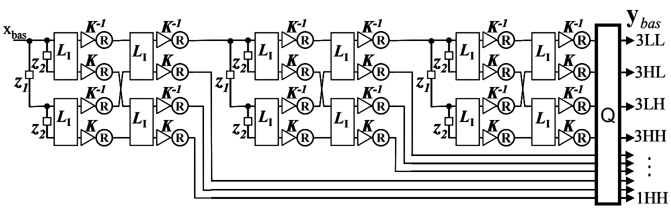
Now, it is obvious that the proposed method reduces the bit rate H_{enh} due to 1) cancellation of the rounding error inside the DWT, and 2) avoidance of amplification of the quantization error by the non-orthonormal synthesis filters.

IV. Proposed BD-SC based on 9/7 DWT

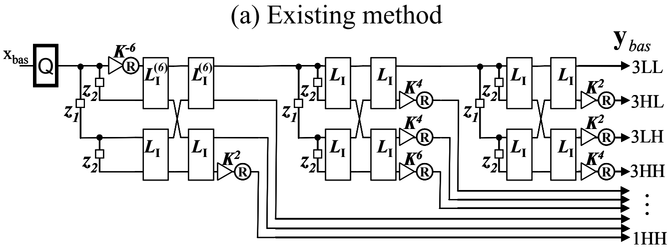
We extend the previously described approach to the integer 9/7 DWT in which all the signals inside the DWT are rounded to integers. It is similar to the integer 5/3 DWT case, however a main difference is that the integer 9/7 DWT is composed of a scaling part $\mathbf{K}$=diag[K^{-1} K] and the lifting filter part L_I.

Fig.6 illustrates an example of the two dimensional three stage octave decomposition. Fig.6 (a) and Fig.6 (b) illustrate the existing method and the proposed method respectively. In Fig.6 (b), the lifting part L_I and the quantization $\mathbf{Q}$ are bartered. As a result, the amplification of the quantization error is avoided and the rounding error is cancelled in all the frequency bands.

In addition, we synthesized all the scaling coefficients K^{-1} in low frequency bands to K^{-6} in the spatial domain. Note that the filter coefficients in the lifting part $L_I^{(6)}$ are adjusted by K^6 so that the transfer function is not changed [10,11].



(a) Existing method



(b) Proposed method

Fig.6 The BD-SC based on the integer 9/7 DWT.

V. Experimental Results

We confirm effectiveness of the proposed method on bit rate reduction using AR(1) model with ρ=0.95 and some image signals. The number of stages in the octave decomposition is set to three. The EBCOT and the Huffman coding are used as the entropy coding in the base layer and the enhancement layer respectively. In the quantization of the existing method, the bit truncation is applied.

A. *Performance of the BD-SC based on 5/3 DWT*

Fig.7 summarizes the bit rate H_{enh} of the enhancement layer versus the image quality θ_{bas} of the base layer signal x_{bas}. We can confirm that the bit rate is reduced by the proposed method. Note that when the step size is one, the signal x_{bas} becomes lossless and there is no difference between the existing method and the proposed method.

Fig.8 summarizes the bit rate H_{bas} of the base layer versus the image quality θ_{bas}. The bit rate is reduced in high bit rate (more than 2 [bpp]) region.

Fig.9 summarizes the result for AR(1) model and some image signals. We can confirm that the bit rate is reduced by 0.38 [bpp] on average. This is equivalent to 8.4 [%] of the bit rate in case of the integer 5/3 DWT based method.

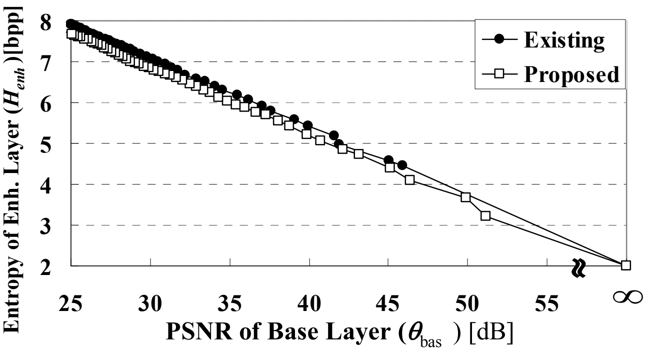


Fig.7 Bit rate of the enhancement layer of the 5/3 DWT based methods. AR(1) model, (N_{org}, N_{bas})=(10, 8) [bit].

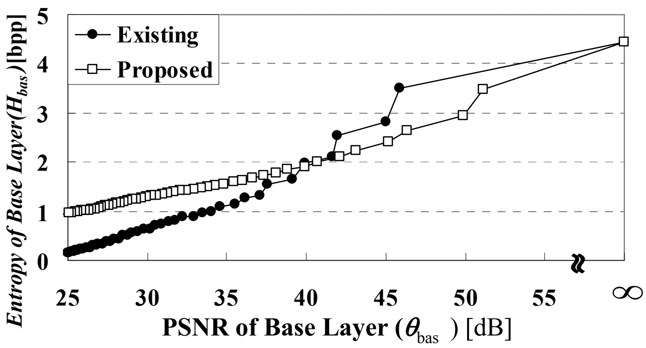


Fig.8 Bit rate of the base layer of the 5/3 DWT based methods. AR(1) model, (N_{org}, N_{bas})=(10, 8) [bit].

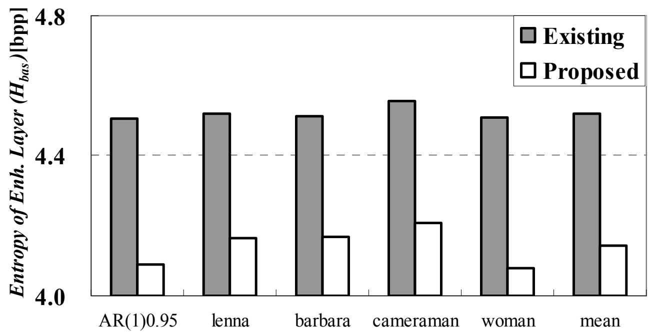


Fig.9 Bit rate of the enhancement layer of the 5/3 DWT based methods for images. (N_{org}, N_{bas})=(8, 6), θ_{bas}=35 [dB].

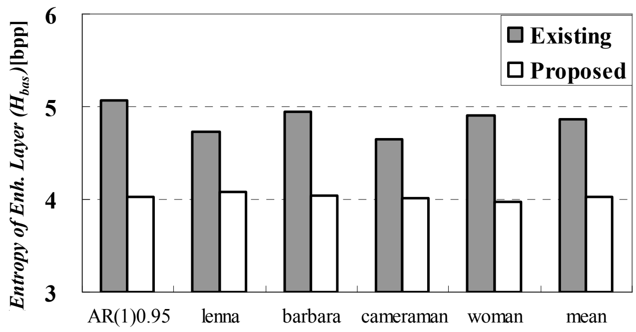


Fig.12 Bit rate of the enhancement layer of the 9/7 DWT based methods for images. (N_{org}, N_{bas})=(8, 6), θ_{bas}=35 [dB].

B. Performance of the BD-SC based on 9/7 DWT

Fig.10 illustrates the bit rate of the enhancement layer of the 9/7 DWT based methods. The bit rate is reduced by 0.5 [bit] on average.This effectiveness is independent of θ_{bas}.

Fig.11 is the bit rate of the base layer. Comparing to the 5/3 DWT based case, the bit rate is much more reduced in higher bit rate region. It is not reduced in the lower bit rate region.

Fig.12 summarizes the bit rate of the enhancement layer for the AR(1) model and some image signals. We can see that the bit rate is reduced by 0.83 [bpp] (17 [%]) on average in case of the integer 9/7 DWT based method.

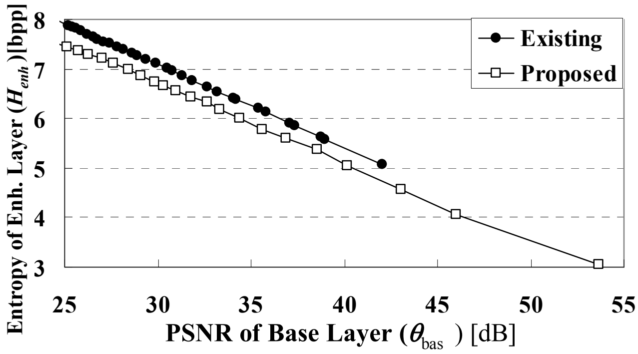


Fig.10 Bit rate of the enhancement layer of the 9/7 DWT based methods for AR(1) model. (N_{org}, N_{bas})=(10, 8) [bit].

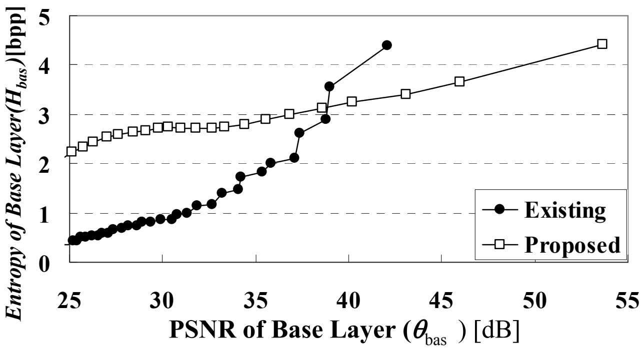


Fig.11 Bit rate of the base layer of the 9/7 DWT based methods for AR(1) model. (N_{org}, N_{bas})=(10, 8) [bit].

VI. CONCLUSIONS

We investigated the bit depth scalable coding based on the integer bi-orthogonal DWT. We reduced the bit rate of the enhancement layer by introducing the quantization in spatial domain. It is confirmed that the entropy rate is reduced by 8.4 [%] and 17 [%] on average by the system based on the 5/3 and the 9/7 integer DWT respectively.

REFERENCES

[1] G. Ward and M. Simmons, "Subband encoding of high dynamic range imagery", Proc. Symposium on Applied perception in graphics and visualization, pp.83-90, 2004.

[2] M. Winken, D. Marpe, H. Schwarz and T. Wiegand, "Bit-depth scalable video coding", IEEE International Conference Image Processing, Vol.1, pp.5-8, Sept. 2007.

[3] S. Liu, W.S. Kim, A. Vetro, "Bit-depth scalable coding for high dynamic range video", SPIE Conf. Visual Communications and Image Processing, TR2007-078, Jan. 2008.

[4] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz, M. Wien, Joint Draft ITU-T Rec. H.264 ISO/IEC 14496-10, "Scalable video coding", July, 2007.

[5] S. Park and K.R. Rao, "Bit-Depth Scalable Video Coding Based on H.264/AVC", IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences, Vol. E91-A, No.6 pp.1541-1544, 2008.

[6] A. Descampe, F.O. Devaux, G. Rouvroy, J.D. Legat, J.J. Quisquater, 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.

[7] ISO/IEC FCD15444-1, "JPEG2000 Image Coding System", March 2000.

[8] M. Iwahashi, Y. Tonomura, S. Chokchaitam, N Kambayashi, "Pre-Post Quantization and Integer Wavelet for Image Compression", IEEE Electronics Letters, vol.39, No.24, pp.1725-1726, 2003.

[9] IEEE Standard 754-1985, IEEE Standard for Binary Floating-Point Arithmetic.

[10] H. Kiya, M. Iwahashi, O. Watanabe, "New Class of Lifting Wavelet Transform for Guaranteeing Losslessness of Specific Signals", IEEE International Conference Acoustics, Speech & Signal Processing, pp.3273-3276, March 2008.

[11] H. Kiya, M. Iwahashi, O. Watanabe, "A New Structure of Lifting Wavelet for Reducing Rounding Error", IEEE Intern'l Symposium Circuits and Systems, pp.2881-2884, May 2008.