LINE BUFFER WORDLENGTH ANALYSIS FOR LINE-BASED 2-D DWT

Chih-Chi Cheng, Chao-Tsung Huang, Jing-Ying Chang, and Liang-Gee Chen

DSP/IC Design Lab, Graduate Institute of Electronics Engineering and
Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
Email: {ccc,cthuang,jychang,lgchen}@video.ee.ntu.edu.tw

ABSTRACT

The on-chip line buffer dominates the total area and power of line-based 2-D DWT. Therefore, the line buffer wordlength has to be carefully designed to maintain the quality level due to the dynamic range growing and the round-off errors. In this paper, a complete analysis methodology is proposed to derive the required wordlength of line buffer given the desired quality level of reconstructed image. The proposed methodology can guarantee to avoid overflow of coefficients, and the difference between predicted and experimental quality level is averagely 0.06 dB in terms of PSNR.

1. INTRODUCTION

2-D DWT is an efficient tool for image and video processing. The line-based implementation [1] can achieve minimum external memory access and has become the mainstream of 2-D DWT VLSI implementation.

Fig. 1 shows a generic scheme for single-level line-based 2-D DWT in column-row order. The line buffer needed in line-based DWT can be decomposed into data buffer and temporal buffer as shown in Fig. 1 [2]. The data buffer can be reduced into only few words of registers [3]. Therefore, the focus of this paper is the wordlength of the temporal buffer.

Temporal buffer is used to buffer the intrinsic register values of every column for column DWT as illustrated in Fig. 2. Therefore, the temporal buffer contains (*Image Width* $\times$ *Number of Intrinsic Registers*) intrinsic register values. This large size makes temporal buffer be the dominant factor of both area and power in 2-D DWT processor [4]. The wordlength of temporal buffer thus has to be carefully designed. However, there are two phenomena that make the wordlength of temporal buffer in multi-level 2-D DWT hard to be determined.

The first phenomenon is the dynamic range growing effect. In multi-level DWT, the coefficients are iteratively filtered for several times. This may lead to variations in signal dynamic range. If overflow occurs, the reconstructed image quality will be severely degraded.

The second phenomenon is the round-off errors which are induced by transferring the floating-point data into data with fixed wordlength. The errors in DWT introduce an upper bound of quality level in image/video system. To control the wordlength such that the reconstructed image can achieve the desired quality level is thus important.

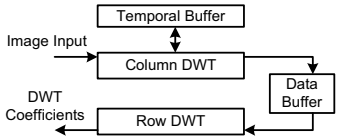

Fig. 1. Line-based scheme for single-level column-row 2-D DWT.

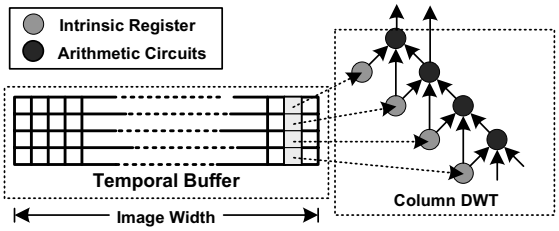

Fig. 2. An example of the temporal buffer scheme. There are 4 words of intrinsic register in lifting-based (9,7) filter, and the temporal buffer contains (4$\times$*Image Width*) intrinsic register values.

Recently there are some reports presented for wordlength designing [5, 6]. Experimental methods in [5] can not guarantee to achieve desired quality level. In [6], the round-off errors of only 1-D DWT are analyzed. Only one error source at each level is considered, and the distortion can not map back to image domain. Moreover, the above works do not take the dynamic range growing effects into consideration.

In this paper, a complete analysis methodology for deriving required line-buffer wordlength in multi-level 2-D DWT is presented. Not only round-off errors but dynamic range growing effects are analyzed. Wordlength derived with the proposed dynamic range analysis methodology can be proved to guarantee the avoidance of overflow. The wordlength required to achieve the desired quality level in reconstructed image can also be derived with the proposed round-off error analysis with a simple distortion model.

This paper is structured as follows. The proposed dynamic range analysis methodology and round-off-error analysis methodology are presented in Section 2 and 3, respectively. The experimental results verifying the proposed method-

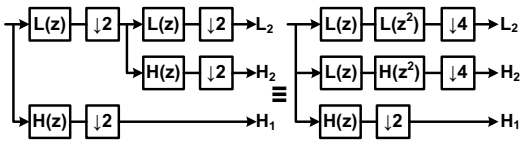

Fig. 3. The Noble identity 1 applied in two-level 1-D DWT.

ologies are presented in Section 4. Finally, Section 5 concludes this work.

2. THE PROPOSED DYNAMIC RANGE ANALYSIS METHODOLOGY

2.1. Dynamic Range Analysis of FIR Filters

Suppose a sequence $x(n)$ with dynamic range within $[-S, S]$ is fed into a FIR filter $H(z) = \sum_{i=-L}^{T} h(i)z^{-i}$, the output is $y(n) = \sum_{i=-L}^{T} h(i)x(n-i)$. Therefore, the maximum possible dynamic range at filter output will be $S \times \sum_{i=-L}^{T} |h(i)|$. The dynamic range gain of a system is defined as the maximum possible value of (*output dynamic range*/*input dynamic range*). For the dynamic range gain G of this filter:

$$G = \sum_{i=-L}^{T} |h(i)| \tag{1}$$

Consider the case of two cascaded FIR filters $H_1(z) = \sum_{i=-L_1}^{T_1} (h_1(i)z^{-i})$ and $H_2(z) = \sum_{i=-L_2}^{T_2} h_2(i)z^{-i}$, these two filters can be merged into an equivalent FIR filter $H_{total}(z)$ with coefficients $h_{total}(n) = \sum_{i=-L_1}^{T_1} h_1(i)h_2(n-i)$. The total dynamic range gain of these two filters G_{total} is:

$$\begin{aligned} G_{total} &= \textstyle\sum_{n=-(L_1+L_2)}^{T_1+T_2} |h_{total}(n)| \\ &= \textstyle\sum_{n=-(L_1+L_2)}^{T_1+T_2} |\sum_{i=-L_1}^{T_1} h_1(i)h_2(n-i)| \end{aligned} \tag{2}$$

2.2. LL-band Dynamic Range Analysis

The coefficients of LL-band are analyzed because the dynamic range of coefficients of $(n-1)$-th level LL band will affect the dynamic range of intrinsic registers in n-th level. Figure 3 shows the Noble identity 1 applied in two-level 1-D DWT. By applying Noble identity, the operations from signal input to each subbands can be equivalent to one filter and one downsampling that will not affect the dynamic range.

For example, if the lowpass filter is $L(z) = \frac{1}{3}z^{-1} + \frac{1}{2} + \frac{1}{3}z$, the dynamic range gain is $|\frac{1}{3}| + |\frac{1}{2}| + |\frac{1}{3}| = \frac{7}{6}$. Taking both column and row directions into consideration, the dynamic range gain of first level LL-band is thus $(\frac{7}{6})^2 = \frac{49}{36}$. As for the second level LL-band, the Noble identity 1 has to be applied. The equivalent filter is $L(z)L(z^2)$, and the corresponding dynamic range gain is 1.361. Therefore, the dynamic range gain in second level LL-band is $1.361^2 = 1.853$. Having the equivalent filter of LL-band at each level, the dynamic range gains of LL-band in all levels can be obtained.

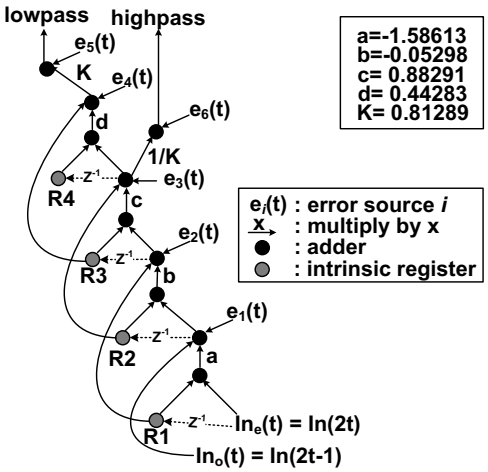

Fig. 4. Hardware architecture of lifting-based (9,7) filter.

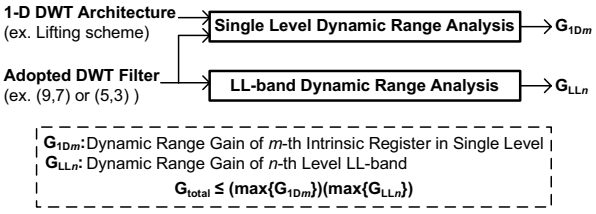

Fig. 5. Analysis flow of the proposed dynamic range analysis methodology.

2.3. Single Level Dynamic Range Analysis

Since DWT is a type of FIR filter, intrinsic register values can be represented as linear combinations of input signal, and this relationship can further be an equivalent FIR filter.

Take lifting-based (9,7) filter in Fig. 4 as an example, there are four intrinsic registers values. Therefore, there are four equivalent filters for column DWT. For example, the equivalent filter for the intrinsic register "*R2*" in Fig. 4 is $a(z^{-1} + z) + 1$. Thus the corresponding filter gain is $2|a| + 1 = 4.17226$. All filter gains from input to first-level intrinsic register values in temporal buffer can thus be obtained.

2.4. Summary of the Proposed Dynamic Range Analysis Methodology

To analyze the dynamic range gains of the intrinsic register values of all levels will be a tedious work. A tight upper bound of dynamic range gain of temporal buffer that can be derived in a much simpler way is thus proposed in this section.

Lemma 1 *Assume two cascaded FIR filters* $H_1(z) = \sum_{i=-L_1}^{T_1} h_1(i)z^{-i}$ *and* $H_2(z) = \sum_{i=-L_2}^{T_2} h_2(i)z^{-i}$ *have filter gains* G_1 *and* G_2*, respectively. The total filter gain* G_{total} *is smaller than or equal to* G_1G_2.

Proof

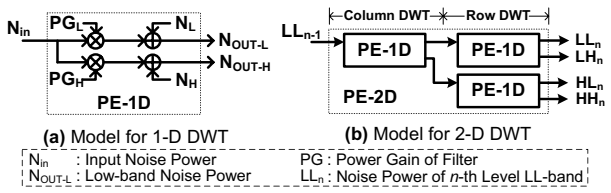

Fig. 6. The proposed noise model of single-level DWT.

From Equ. 2:

$$\begin{aligned} G_{total} &= \sum_{n=-(L_1+L_2)}^{T_1+T_2} |\sum_{i=-L_1}^{T_1} h_1(i)h_2(n-i)| \\ &\le \sum_{n=-(L_1+L_2)}^{T_1+T_2} \sum_{i=-L_1}^{T_1} |h_1(i)h_2(n-i)| \\ &= (\sum_{i=-L_1}^{T_1} |h_1(i)|)(\sum_{j=-L_2}^{T_2} |h_2(j)|) \\ &= G_1 G_2 \end{aligned} \tag{3}$$

Lemma 2 *Let the dynamic range gain from input image to temporal buffer be G_{max}, then $G_{max} \le (\max_m\{G_{1D_m}\})\times(\max_n\{G_{LL_n}\})$, where G_{1D_m} is the dynamic range gain of the m-th intrinsic register value within single level, and G_{LL_n} is the dynamic range gain from input image to the n-th level LL-band.*

Proof

The relationship from the input to one word in intrinsic register value in n-th level can be taken as the cascade of the relationship from input to $(n-1)$-th level LL-band and the relationship from $(n-1)$-th level LL-band to the intrinsic register value. Assume $G_{total_{m,n}}$ be the filter gain from input to the m-th intrinsic register value in n-th level column DWT. From lemma 1:

$$\begin{aligned} G_{max} &= \max_{m,n}\{G_{total_{m,n}}\} \\ &= G_{total_{m1,n1}}(\exists m_1, n_1) \\ &\le G_{1D_{m1}} G_{LL_{n1}} \\ &\le (\max_m\{G_{1D_m}\})(\max_n\{G_{LL_n}\}) \end{aligned} \tag{4}$$

The analysis flow is shown in Fig. 5. G_{LL_n} can be obtained from section 2.2, and G_{1D_m} can be obtained from section 2.3. With these two values, the upper bound of the dynamic range gain from input image to temporal buffer, G_{max}, can be obtained. Therefore, the wordlength needed to prevent overflow can be obtained.

3. THE PROPOSED ROUND-OFF ERROR ANALYSIS METHODOLOGY

In this section, a hierarchical model to estimate round-off errors in reconstructed image is proposed, and it is a function of the number of bits in fractional part of temporal buffer.

3.1. Model of Round-off Operations

The most basic element in the proposed error model is the model of round-off operations. Once a round-off operation is performed, an zero-mean uniformly-distributed and mutually-independent additive error source is introduced, as the $e_1(t)$ to $e_6(t)$ in Fig. 4. If there are n fractional bits to represent a coefficient, for error sources $\{e_i(t) : i = 1, 2, ...\}$:

$$E\{e_i(t_1)e_i(t_2)\} = 0, \qquad t_1 \ne t_2 \tag{5a}$$

$$E\{e_i(t_1)e_j(t_2)\} = 0 \qquad \forall t_1 t_2, i \ne j \tag{5b}$$

$$E\{e_i(t)e_i(t)\} = \sigma^2, \sigma^2 = \int_{-\frac{1}{2^{n+1}}}^{\frac{1}{2^{n+1}}} 2^n x^2\, dx = \frac{2^{-2n}}{12} \tag{5c}$$

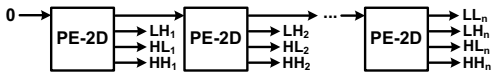

Fig. 7. The noise power model of multi-level 2D DWT.

3.2. Noise Power Model of Single-level 1-D DWT

To analyze the noise power at 1-D DWT output, the noise sources are taken as real signals in the filter architecture. The noise power can be evaluated from the expression of noise sources. For example, consider the contributions of noise source $e_3(t)$ and $e_4(t)$ in Fig. 4 to the lowpass output, the expression of these two noise sources ($e_{LP}(t)$) is:

$$e_{LP}(t) = Kd(e_3(t) + e_3(t-1)) + Ke_4(t) \tag{6}$$

In this case, the estimated noise power $E\{e_{LP}^2(t)\}$ can be derived from the assumptions in Eq. 5:

$$\begin{aligned} E\{e_{LP}^2(t)\} &= E\{(Kd(e_3(t) + e_3(t-1)) + Ke_4(t))^2\} \\ &= K^2d^2E\{e_3^2(t) + e_3^2(t-1)\} + K^2E\{e_4^2(t)\} \\ &= (2K^2d^2 + K^2)\sigma^2 \end{aligned} \tag{7}$$

Thus the noise power induced by 1-D DWT can be calculated in this way. If the input signal is with a known noise power, this error will also be modelled as a noise source that satisfies Eq. 5a and 5b. If the filter is $H(z) = \sum_{i=-L}^{T} h_i z^{-i}$ and the input noise power is N_{in}, the noise power at filter output induced by input noise, N_{out}^{in}, is:

$$\begin{aligned} N_{out}^{in} &= E\{(\sum_{-L}^{T} h_i e_{in}(t-i))^2\} &&= \sum_{-L}^{T} h_i^2 E\{e_{in}^2(t-i)\} \\ &= E\{e_{in}^2(t)\} \times \sum_{-L}^{T} h_i^2 &&= N_{in} \times PG \end{aligned} \tag{8}$$

Where the PG is defined as the power gain of $H(z)$. In the proposed noise model, the input noise will be amplified by PG after the filtering operation.

Figure 6a shows the proposed noise model of single level 1-D DWT. As discussed, the fixed-point hardware will introduce round-off error power (N_L, N_H), and the input noise power will be amplified by noise power gains (PG_L, PG_H).

3.3. Noise Power Model of Multi-level 2-D DWT

Figure 6b further shows the proposed noise model in single level 2-D DWT. It is a cascade of noise models of 1-D DWT.

The proposed noise model shown in Fig. 7 for multi-level 2-D DWT consists of the cascade of the noise models of single level DWT. The input noise power of n-th level is the the LL-band noise power in $(n-1)$-th level. At the first level, the input is the original image, the input noise power is therefore zero.

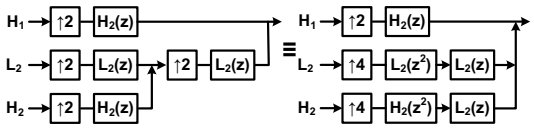

Fig. 8. Noble identity 2 applied in two-level 1-D IDWT.

Table 1. The estimated and the measured quality of reconstructed image with different number of fractional bits.

	Derived Upper Bound	Experimental Maximum Dynamic Range
Lifting-based	731 (10bits)	549 (10bits)
convolution-based	241 (8bits)	233 (8bits)

3.4. Noise Power Analysis in Reconstructed Image

In section 3.3, the noise powers in all sub-bands are calculated. To estimate the corresponding noise power in reconstructed image, noble identity 2 is utilized. As illustrated in Fig. 8, the IDWT can be considered as upsampling, filtering, and addition of each subband. Upsampling by n will make the noise power became $\frac{1}{n}$ times, and the equivalent filter can be modelled as a noise power gain as in section 3.2.

3.5. Summary of Round-off-error analysis

In summary, noise power gains (PG_L, PG_H) can be obtained with the DWT filter type. Noise power induced in 1-D DWT (N_L, N_H) can be obtained by analyzing the 1-D hardware architecture. The model of single level 2-D DWT can be obtained by cascading the 1-D noise models. The multi-level 2-D DWT model can be obtained by cascading the single level DWT models. By feeding zero noise power into the multi-level DWT model, the noise power of each subbands can be obtained. Finally, the noise power expression in reconstructed image can be estimated by calculate the power gain of each sub-band using noble identity 2. This noise power expression is a function of the number of fractional bits of data because the value of σ is defined as in Eq. 5c. The noise power in reconstructed image is the mean-square-error of pixels and can directly be mapped into PSNR. Therefore, the required number of fractional bits in temporal buffer can be obtained from the required quality level in image domain.

4. EXPERIMENTAL RESULTS

Table 1 shows the derived dynamic range upper bound and the experimental maximum dynamic range when random signals are taken as input. Two hardware architectures of DWT, convolutional-based and lifting-based architectures are implemented by Verilog HDL. As shown in table 1, the proposed dynamic range upper bound and the experimental maximum dynamic range yield the same number of required bits in both architectures. Therefore, the proposed upper bound of dynamic range is tight enough in designing the wordlength.

Table 2 shows the estimated and the measured quality of reconstructed image with different number of fractional bits in temporal buffer. As shown in table 2, the average and the maximum prediction error of proposed precision analysis methodology are 0.06 dB and 0.43 dB, respectively. Therefore, the proposed methodology can provide a good reference for designing the number of fractional bits in temporal buffer.

Table 2. The estimated and the measured quality (in PSNR(dB)) of reconstructed images with different number of fractional bits. The average and the maximum prediction error are 0.06 dB and 0.43 dB.

Lifting-based							
Fractional Part	lena	baboon	lake	pepper	average	**Predicted**	**difference**
2 bit	52.21	52.27	52.38	52.35	52.30	**52.41**	**0.11**
3 bit	58.36	58.25	58.00	58.26	58.21	**58.43**	**0.21**
4 bit	64.35	64.43	64.31	64.29	64.34	**64.45**	**0.10**
convolution-based							
Fractional Part	lena	baboon	lake	pepper	average	**Predicted**	**difference**
2 bit	49.43	49.80	49.31	49.55	49.52	**49.66**	**0.14**
3 bit	55.83	55.76	55.80	55.77	55.79	**55.68**	**-0.11**
4 bit	61.75	61.81	61.77	61.79	61.78	**61.70**	**-0.08**

5. CONCLUSION

Line buffer dominates the area and power in line-based 2-D DWT. In this paper, design methodology for the wordlength of line buffer is proposed. The proposed dynamic range analysis methodology can provide a tight upper bound and guarantee to prevent the overflow in line buffer. The proposed round-off error model can predict the image-domain quality level with averagely 0.06 dB difference, and the required fractional wordlength can be obtained from the desired quality level. Both proposed methodologies are simple to be performed.

6. REFERENCES

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

[2] P.-C. Tseng, C.-T. Huang, and L.-G. Chen, "Generic RAM-based architecture for two-dimensional discrete wavelet transform with line-based method," in *Asia-Pacific Conference on Circuits and Systems*, 2002, pp. 363–366.

[3] Mu-Yu Chiu, Kun-Bin Lee, and Chein-Wei Jen, "Optimal data transfer and buffering schemes for JPEG2000 encoder," in *IEEE Workshop on Signal Processing Systems*, 2003, pp. 177–182.

[4] C.-C. Cheng and et al., "Multiple-lifting scheme: Memory-efficient VLSI implementation for line-based 2-D DWT," in *Proceedings of 2005 IEEE International Symposium on Circuits and Systems*, 2005, pp. 5190–5193.

[5] Michael Weeks, "Precision for 2-D discrete wavelet transform processors," in *IEEE Workshop on Signal Processing Systems*, 2000, pp. 80–89.

[6] H. Choi, W. P. Burleson, and D. S. Phatak, "Optimal wordlength assignment for the discrete wavelet transform in VLSI," in *Proceedings of IEEE Workshop on VLSI Signal Processing*, 1993, pp. 325 – 333.