COST-EFFICIENT MULTIPLIER-LESS FIR FILTER STRUCTURE BASED ON MODIFIED DECOR TRANSFORMATION

I-Hsien Lee, Cheng-Shing Wu, and An-Yeu Wu[*]

Dept. of Electrical Engineering, National Central University,
Chung-Li, 320, Taiwan, R.O.C.
[*]Dept. of Electrical Engineering, National Taiwan University,
Taipei, 106, Taiwan, R.O.C.

ABSTRACT

In this paper, we propose a new design approach to implement FIR filter using CSD multipliers based on *Modified Decorrelating transformation (MDECOR)*. The direct CSD approach will introduce serious quantization errors since the distribution of CSD numbers is very non-uniform. The proposed MDECOR tranformation provides a systematic solution to reduce the dynamic range effectively. By combining the proposed MDECOR transformation followed by CSD quantization, we can avoid the aforementioned quantization problem. As a result, we do not need to employ additional non-zero bits to compensate for the distortion caused by direct CSD quantization, which helps to save the number of adders in VLSI implementations. Furthermore, the MDECOR transformation offers more design of freedom in the filter design. It can achieve high-precision performance under the same hardware complexity as the direct CSD approach. Our simulation results show that we can save 20% number of adders compared with the direct CSD approach.

1. INTRODUCTION

FIR filter is one of the key functional blocks in many digital signal processing (DSP) applications. The hardware complexity of the FIR filter can be very high when the filter is implemented with array multipliers. Many researches are devoted to reduce the hardware complexity of the FIR filter. Among them, the *Canonical Sign Digit* (CSD) approach is one of the most popular approaches [1][2]. In CSD approach, each coefficient is represented/quantized as the sum of *Signed Power-of-Two (SPT)* terms. For example, -0.37521 can be re-formulated as $0.\bar{1}01$ for 2 non-zero digits. By doing so, the implementation of multipliers can be accomplished with only shift-and-add operations. Hence, the cost can be reduced significantly in ASIC designs.

However, the major problem of CSD approach is that the distribution of CSD coefficients is highly non-uniform, which limits the precision performance of the CSD multipliers. In Fig. 1, we demonstrate the distribution of 2-nonzero-digit CSD numbers between 0.0 and 1.0 for wordlength 6, 7, and 8, respectively. As we can see, the gaps of the CSD distribution cannot be reduced even if the wordlength of the CSD numbers increases. To achieve higher precision performance (reducing the gaps), we have to employ more non-zero digits. However, increasing the non-zero digit has the effect of increasing the number of adders in each filter tap.

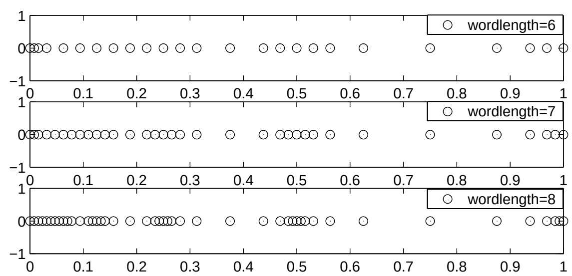

Figure 1: The distribution of 2-nonzero-digit CSD numbers between 0.0 and 1.0 (wordlength = 6,7,8). Each circle denotes the location/value that can be represented by the CSD number.

In this paper, we propose a new CSD approach in FIR filter implementation based on *reduction of dynamic range*. This can be achieved by modifying the *Differential Coefficient Method (DCM)* [3] and *Decorrelating transformation (DECOR)* [4]. That is, we *transform* the original coefficients into a new set of coefficients with much smaller dynamic range than the original ones. Then, CSD quantization process is applied to these transformed coefficients.

In addition, to prevent from the stability problem, we also make some modifications on the DECOR transformation, called *Modified DECOR (MDECOR)* transformation. It introduces more degrees of freedom in the filter designs, which helps to further reduce the dynamic range of coefficients. As will shown in this paper that by combining the proposed MDECOR followed by CSD quantization, we can save 20% number of adders compared with direct CSD approach.

2. REVIEW OF DCM AND DECOR TRANSFORMATION

2.1. Time-domain Representation

Mathematically, an N-tap FIR filter performs the following convolution

$$y(n) = \sum_{k=0}^{N-1} h_k x(n-k), \tag{1}$$

where h_k is the k-th coefficient of the FIR filter; $x(n)$ and $y(n)$ denote the input and the output signals at time instance n, respec-

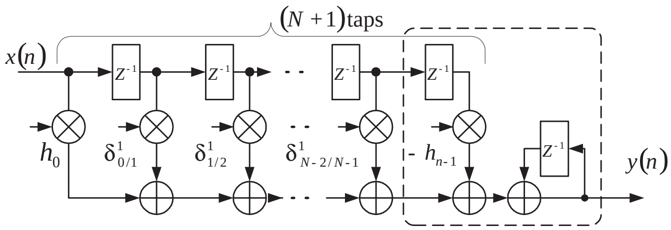

Figure 2: FIR filter structure based on 1^{st}-order DCM.

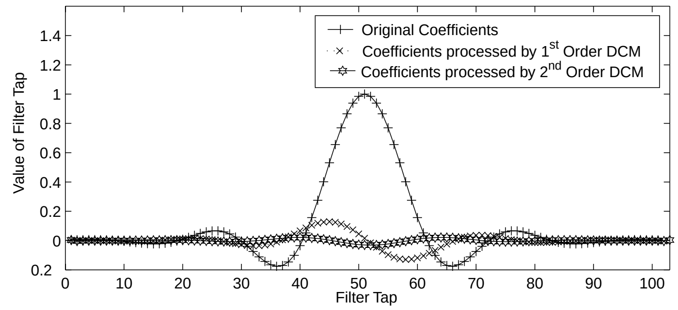

Figure 3: Dynamic range reduction after applying DCM.

tively. In 1^{st}-order DCM, Eq. (1) can be reformulated as [3]

$$\begin{aligned} y(n) &= \sum_{k=0}^{N-1} \left[(h_k - h_{k-1}) + h_{k-1} \right] x(n-k) \\ &= h_0 x(n) + \sum_{k=1}^{N-1} \left[\delta_k^1 x(n-k) \right] + \\ &\qquad y(n-1) - h_{N-1} x(n-N), \quad (2) \end{aligned}$$

where

$$\delta_k^1 \triangleq h_k - h_{k-1}. \quad (3)$$

The "1^{st}-order" operation denotes that we take difference between the contiguous coefficients *only once*. The corresponding structure of the DCM-based FIR filter is depicted in Fig. 2. As we can see from Fig. 2, the extra cost of the 1^{st}-order DCM is *one additional tap and one accumulator* (circled by the dotted line). For m^{th}-order DCM, the coefficients are generated by taking the difference of the $(m-1^{th})$-order DCM coefficients as

$$\delta_k^m \triangleq \delta_k^{m-1} - \delta_{k-1}^{m-1}. \quad (4)$$

To see the effectiveness of the DCM, we apply the 1^{st} and 2^{nd} DCM on the coefficients of a 101-tap low-pass FIR filter with cutoffs frequency of $\pm 0.0945 f_s/2$. The original coefficients as well as the coefficients after 1^{st} and 2^{nd} order DCM are shown in Fig. 3. For the 1^{st}-order DCM, we find the magnitude of coefficients is reduced significantly; all the values of processed coefficients are suppressed within 20% compared with the original coefficients. Similarly, the coefficients of the 2^{nd}-order DCM are further reduced within 10% of the original dynamic range. The reduction of dynamic range of FIR filter imply that the wordlength can be reduced. Hence, a lower cost of hardware complexity can be achieved when we apply the DCM to FIR filter implementation.

2.2. z-domain Representation

From z domain point of view, we can represent the 1^{st}-order DCM of the FIR filter as

$$H'(z) = \frac{H(z)\left(1 - z^{-1}\right)}{(1 - z^{-1})}, \quad (5)$$

where $H(z)$ denotes the original FIR filter response. In [4], Eq. (5) is generalized to the DECOR, in which the transfer function is rewritten as

$$H'(z) = \frac{H(z)\left(1 + \alpha z^{-\beta}\right)^m}{(1 + \alpha z^{-\beta})^m}. \quad (6)$$

Filter Type	α	β	$f(z)$
Low-pass	-1	1	$\left(1 - z^{-1}\right)^m$
High-pass	1	1	$\left(1 + z^{-1}\right)^m$
Band-pass (ω_c: center freq.)	1	π/ω_c	$\left(1 - z^{-\pi/\omega_c}\right)^m$
Band-stop	-1	2	$\left(1 - z^{-2}\right)^m$

Table 1: Setting of α and β for different types of FIR filters in DECOR transform.

The parameters of α and β are chosen depending on the filter type, and m denotes the order (iteration numbers of coefficient operation) of DECOR as listed in Table 1. Note that, when $m = \alpha = 1$, DECOR is reduced to 1^{st}-order DCM of Eq. (5).

3. THE PROPOSED MDECOR-BASED CSD QUANTIZATION PROCESS

In general, in the CSD-based FIR filter design, the coefficients are quantized to the *nearest* CSD numbers. Recall in Fig. 1 that the distribution of CSD numbers is denser for smaller values, and the distribution is sparser for larger values. As one can expect, *the quantization error can be higher for those coefficients in the sparse region of the CSD distribution*. Consequently, in [5], the authors suggested that we can add one more non-zero digit in CSD representation when the magnitudes of FIR filter coefficients exceed 0.5. In fact, this approach can increase the distribution density of CSD number, resulting in smaller quantization errors. However, the reduction of quantization error is gained at the expense of hardware complexity. More adders are required in FIR filter implementation due to the increased non-zero digits. Moreover, since more number of adders of each filter tap is employed, the critical path is also increased correspondingly, which will degrade the operation speed of the filter.

Instead of employing additional non-zero digits, in this paper, we propose a cost-efficient solution to enhance the precision performance. Our idea is to *move* those large-valued coefficients into the region with dense CSD distribution. By doing this, we can avoid the serious quantization errors that are resulted from the *large gap* of the highly non-uniform CSD distribution. As a consequence, we can improve the overall precision performance of FIR filter in practical fixed-point implementation without increasing the hardware complexity.

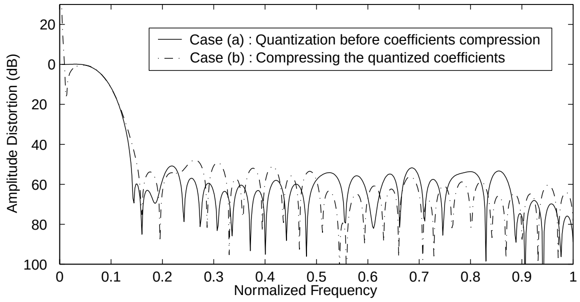

Figure 4: Frequency response of (a) Quantization followed by coefficients compression, (b) Quantization of the compressed coefficients.

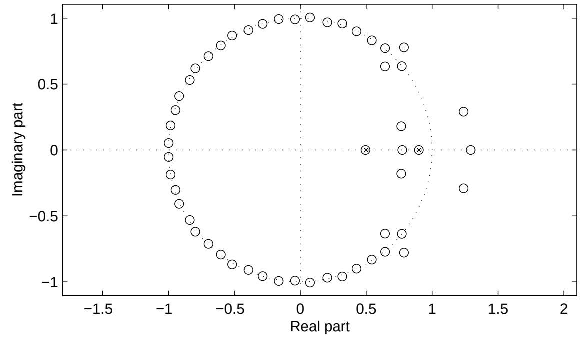

Figure 5: Location of pole-zero pairs of 2^{nd}-order MDECOR while $|\alpha_1| = 0.9$ and $|\alpha_2| = 0.5$.

3.1. CSD Problems in DECOR Transformation

From Eq. (6), we can see that the DECOR transformation is equivalent to inserting m pole-zero pairs at $z = \alpha e^{-j\beta}$ on the z-plane. In Table 1, all α's are set to 1 or -1, *i.e.*, all the inserted poles and zeros lie on the unit circle. In fact, to retain the stability of the IIR filter, the inserted poles must be exactly cancelled with the inserted zeros.

To avoid the aforementioned stability problem, in applying CSD to the DECOR transformation, we perform the following procedures:

1. Quantize the filter coefficients into CSD numbers.
2. Reduce the dynamic range of filter coefficients into a smaller set by using DECOR transformation.

That is, we *first quantize $H(z)$ as a quanitzed $\tilde{H}(z)$, followed by the multiplication of* $\left(1 + \alpha z^{-\beta}\right)^m$.

The resulting frequency response of the filter is shown in Case (a) of Fig. 4. It can still maintain the filter response since the zero-pole pairs will not be drifted from the original location. On the contrary, quantizing the compressed coefficients after the DECOR transformation leads to the imperfect cancellation of inserted pole-zero pairs. Therefore, frequency response of the filter will be changed as shown in Case (b).

3.2. Modified DECOR (MDECOR) Transformation

As one can expect, the aforementioned DCM and DECOR transformation provide a good solution in *transforming* the original coefficients into a new coefficients set with smaller values. However, as will be described later that some basic operations of DCM and DECOR make it difficult to apply DCM and DECOR to CSD process directly. To facilitate the proposed operation, we make two modifications on DECOR transformation, as described below:

A. *Interchange the Quantization Process*

In the DCM and DECOR approaches, we quantize the original coefficients before compressing the quantized coefficients into smaller values. Combined with CSD quantization, these coefficients to be quantized cannot avoid the large gap in CSD distribution. On the contrary, in our approach, the operation of moving those large-valued coefficients into the dense region must be processed *before* CSD quantization.

Moreover, it is hard to assign the number of adders precisely in FIR filter implementation based on DECOR-based CSD process. In fact, it may be easy to assign number of adders in each tap while we quantize filter coefficients into CSD numbers, however, the number of adders may be increased after we take the operation of CSD quantized coefficient by following Eqs. (3) and (4). As a result, *DCM and DECOR transformation cannot be applied to CSD implementation directly*. To facilitate the proposed CSD operation, we interchange operation of the coefficient quantization and dynamic range compressing process, as described below.

1. Compress the dynamic range of filter coefficients into a smaller set.
2. Quantize these pre-processed coefficients, which are generated by step 1, into CSD numbers.

B. *A Generalized Value of α_i*

In high-order DECOR transformation, all α are determined as the same value (see Eq.(6)). In our proposed scheme, we relax the operation of DECOR by modifying the transfer function in Eq. (6). The transfer function of the m^{th}-order *Modified DECOR (MDECOR)* is represented as

$$H'(z) = H(z)\frac{\prod_{i=1}^{m-1}\left(1 + \alpha_i z^{-\beta}\right)^m}{\prod_{i=1}^{m-1}\left(1 + \alpha_i z^{-\beta}\right)^m}. \quad (7)$$

where α_i has the constraint

$$-1 \leq \alpha_i \leq 1. \quad (8)$$

Note that the value of β and sign of α_i are still determined depending on filter types listed in Table 1. However, the value of α_i can be chosen as an arbitrary value with absolute value less than 1. By doing this, the poles in Eq. (7) will not cause serious distortion even the zeros drift away from the original location.

Furthermore, in the m^{th}-order MDECOR transformation different values of α_i can be assigned to optimize filter performance, as shown in Fig. 5. This phenomenon may result in smaller dynamic range of coefficients, implying that we may use fewer adders to implement FIR filter for the same design specification. Under the same hardware complexity, the proposed operation increases the degree of freedom in FIR filter design; hence a better *Signal-to-Quantize Noise Ratio (SQNR)* performance can be expected.

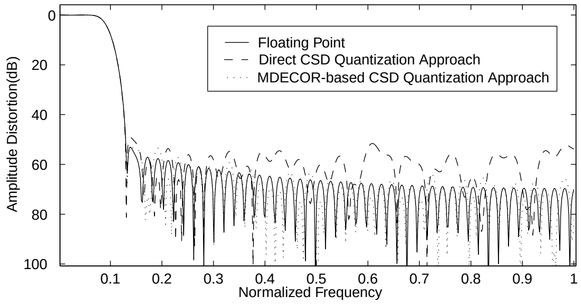

Figure 6: Frequency response of MDECOR-based CSD approach and direct CSD approach.

4. SIMULATION RESULTS AND PERFORMANCE COMPARISON

Fig. 6 shows the frequency responses of a narrow-band low-pass FIR filter using CSD multipliers and the specification is mentioned in section 2. All CSD multipliers employ 3 non-zero digits and 12 bits coefficient wordlength. As one can see, direct CSD quantization of coefficients results in serious distortion in frequency response. On the contrary, the difference between the proposed MDECOR-based CSD approach and ideal frequency response is much smaller.

Fig. 7 shows the SQNR values for different coefficient wordlength. Note that in this simulation, we also consider the truncation error caused by shifters of CSD multiplier. While direct CSD quantization and the proposed MDECOR-based CSD quantization process are all employed 3 non-zero digits in their CSD representation. From Fig. 7, we can make the following observations:

1. SQNR of the proposed method is about 10 dB higher than the direct CSD quantization approach.
2. The wordlength of direct CSD approach reaches its saturation value, which is smaller than the saturation wordlength of the proposed scheme.
3. The topmost curve in Fig. 7 represents the precision performance of the direct CSD approach with 4 non-zero digits. It can achieve a higher SQNR value but the complexity is higher than the proposed approach under the same specification. As a result, if the performance of direct CSD approach cannot satisfy the specification under a specified number of non-zero digits. We can try to pre-process filter coefficients by using the proposed MDECOR-based CSD approach to achieve the SQNR requirement, instead of using more non-zero digits in practical implementations.

Fig. 8 shows the SQNR values of the FIR filter with 1^{st}-order and 2^{nd}-order MDECOR transformation under the same hardware complexity. As we expect, 2^{nd}-order MDECOR transformation has a better SQNR performance. The reason is that we have more design freedom to search for the optimized value of α_i in fixed-point filter design.

From the above simulation results, we can conclude that when the processed coefficients are used, FIR filter realization with CSD multiplier evinces a better performance. That is, we can use fewer adders to realize FIR filter with CSD multipliers.

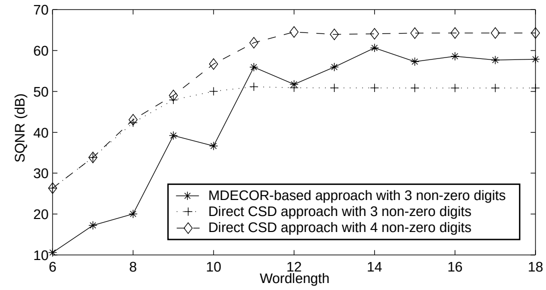

Figure 7: SQNR results of MDECOR-based implementation (3 non-zero digits) and direct CSD approach (3 and 4 non-zero digits).

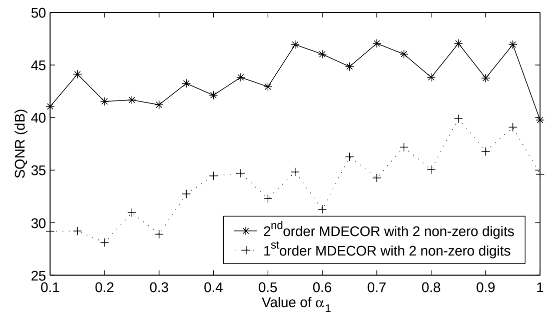

Figure 8: SQNR results of the coefficients by applying 1^{st}-order and 2^{nd}-order MDECOR.

5. CONCLUSIONS

In this paper, we introduced a MDECOR transformation to reduce the serious quantization error caused by non-uniform distribution of CSD numbers. The MDECOR compresses the magnitude of coefficients before quantizing them. Suppose that we want to implement a FIR filter by CSD multipliers but the performance cannot satisfy the specification under limited numbers of non-zero digits. We can employ the proposed MDECOR to achieve the target precision performance, instead of using more adders in practical implementation. Hence, a lower cost can be achieved in designing the multiplier-less FIR filters.

References

[1] K. Hwang, *Computer Arithmetic, Principle, Aritecture, and Design*. New York: Wiley, 1979.

[2] Y. C. Lim and S. R. Parker, “Fir filter design over a discrete powers-of-two coefficients,” *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 31, pp. 583–591, June 1983.

[3] N. Sankarayya, K. Roy, and D. Bhattacharya, “Algorithms for low power and high speed FIR filter realization using differential coefficients,” *IEEE Trans. Circuits Syst.*, vol. 44, pp. 488–497, June 1997.

[4] S. Ramprasad, N. R.Shanbhag, and I. N. Hajj, “Decorrelating (decor) transformations for low-power digital filters,” *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 776–788, June 1999.

[5] H. Samueli, “An improved search algorithm for the design of multiplierless fir filters with powers-of-two coefficients,” *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1044–1047, 1989.