LETTER

Methods for Avoiding the Checkerboard Distortion Caused by Finite Word Length Error in Multirate System

Hiroaki IWAI[†], *Student Member*, Masahiro IWAHASHI[††], *Member*, *and* Hitoshi KIYA[†a)], *Fellow*

SUMMARY We propose two methods for avoiding the checkerboard distortion which is caused by finite word length error. The first method derives the bit length of filter coefficients required for avoiding the checkerboard distortion under a certain word length. In the second method, the checkerboard distortion can be avoided by using the cascade structure which consists of zero-hold kernel and a time-invariant filter factorized from a filter with structure for avoiding the checkerboard distortion under linear systems. It is demonstrated by simulations that we can avoid the checkerboard distortion by using these proposed methods.

key words: *checkerboard distortion, multirate system, finite word length error, cascade structure*

1. Introduction

Multirate signal processing techniques are applied to many areas such as filter banks, analog-to-digital, digital-to-analog conversion, image resolution conversion, and so on [1], [2]. The multirate system to perform multirate signal processing generally consists of up-sampler, a digital filter, and down-sampler, since these techniques are implemented based on sampling rate conversion. It is known that the perceptible artifacts called checkerboard distortion occurs owing to periodically time varying process of the interpolator in multirate system [4], [5]. The condition for avoiding this checkerboard distortion has been proven [4], [5]. However, this condition cannot be necessarily approved in practical systems because infinite word length is assumed in this condition. In other words, this condition is not guaranteed because of error caused by finite word length. Then there is a possibility that the checkerboard distortion occurs.

In this paper, two methods for avoiding the checkerboard distortion under finite word length are proposed. The first method derives the bit length of filter coefficients required for avoiding the checkerboard distortion under a certain word length. In the second method, we construct the cascade structure which consists of zero-hold kernel and a time-invariant filter factorized from a filter with structure for avoiding the checkerboard distortion under linear systems, and propose that the checkerboard distortion can be avoided by using this cascade structure. The cascade structure used in the second method does not cause the checkerboard distortion under any word length of coefficients. On the other hand, the first method is needed, when the cascade structure cannot be applied for some reasons.

It is demonstrated by experiments of image resolution conversion in Sect. 4 that we can avoid the checkerboard distortion by using these proposed methods.

Manuscript received June 18, 2009.
Manuscript revised November 6, 2009.
[†]The authors are with the Department of Information and Communication Systems Engineering, Tokyo Metropolitan University, Hino-shi, 191-0065 Japan.
[††]The author is with the Department of Electrical Engineering, Nagaoka University of Technology, Nagaoka-shi, 940-2188 Japan.
a) E-mail: kiya@sd.tmu.ac.jp
DOI: 10.1587/transfun.E93.A.631

2. Checkerboard Distortion

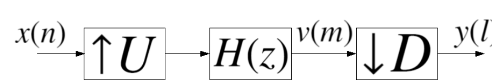
In this section, firstly, sampling rate conversion is described. Secondly, we explain about the cause of the checkerboard distortion. Moreover, the condition for avoiding the checkerboard distortion under linear systems is described.

2.1 Sampling Rate Conversion

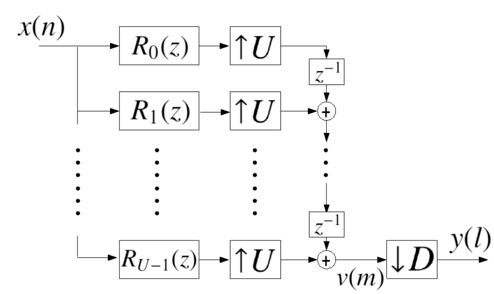
A composition of a sampling rate converter is shown in Fig. 1(a), where U and D are mutually prime positive integers, $H(z)$ is a digital filter, $\uparrow U$ is up-sampler with the factor U, and $\downarrow D$ is down-sampler with the factor D. The sampling rate of the output signal obtained of this converter is U/D times compared to that of input signal. $H(z)$ is the transfer function of an FIR filter given as

$$H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n}, \tag{1}$$

where $h(n)$ denotes the impulse response of $H(z)$.



(a)General structure



(b)Polyphase structure

Fig. 1 Sampling rate converter.

Figure 1(a) can be equivalently represented as polyphase structure shown in Fig. 1(b) [3]–[5]. $H(z)$ can be represented as

$$H(z) = \sum_{i=0}^{U-1} R_i(z^U) z^{-U-1-i}, \tag{2}$$

where

$$R_i(z) = \sum_{n=-\infty}^{\infty} h(Un - i - 1 + U) z^{-n}. \tag{3}$$

$R_i(z)$ are the polyphase components of $H(z)$.

2.2 Definition of Checkerboard Distortion

It is known that the checkerboard distortion occurs in multirate systems due to periodically time varying processing of the interpolator which consists of an up-sampler and a digital filter [4], [5].

In Fig. 1(b), if the input signal $x(n)$ is the unit step signal $u(n)$, and besides, n is large enough, the value of interpolated signal $v(m)$ becomes steady state given as

$$s(m) = \begin{cases} R_0(1), & (m = Un + (U-1)), \\ R_1(1), & (m = Un + (U-2)), \\ \vdots & \\ R_{U-1}(1), & (m = Un), \end{cases} \tag{4}$$

where $R_i(1)$ denote the DC gain of ith polyphase component $R_i(z)$. Consequently, the value of output signal $y(l)$ is also steady state given as

$$y(l) = s(Dm). \tag{5}$$

Because each $R_i(1)$ is generally different, $s(m)$ is not constant and has the period U. In consequence, $y(l)$ also has the period U. The periodic artifact caused by this periodic step response is called the checkerboard distortion [4], [5].

2.3 Conditions for Avoiding the Checkerboard Distortion under Linear Systems

The condition for avoiding the checkerboard distortion under linear systems has been proven [4], [5].

Theorem: The necessary and sufficient condition for avoiding the checkerboard distortion is that the DC gain of each polyphase component is equivalent to a constant. Namely,

$$R_0(1) = R_1(1) = \cdots = R_{U-1}(1) = G. \tag{6}$$

We note that when $H(z)$ can be factorized as Eq. (7), Eq. (6) is satisfied.

$$H(z) = H_0(z)P(z). \tag{7}$$

where

$$H_0(z) = \sum_{i=0}^{U-1} z^{-i}, \tag{8}$$

$H_0(z)$ denotes the interpolation kernel of zero-order hold. A condition for satisfying Eq. (7) is that $H(z)$ has zero values at z shown in Eq. (9).

$$z = e^{j\omega}\,|_{\omega=\frac{2\pi m}{U}}, m = 1, 2, \cdots, U-1, \tag{9}$$

3. Methods for Avoiding the Checkerboard Distortion under Finite Word Length

The necessary and sufficient condition described in Sect. 2.3 premises that filter coefficients do not have finite word length error caused by quantization. Therefore, in a practical system, a method unaffected by finite word length error is needed to build the multirate system without the checkerboard distortion. In this section, firstly, a quantization used in this paper is described. Secondly, we consider the checkerboard distortion under finite word length, and propose two methods for avoiding it.

3.1 Quantization Used in This Paper

In this paper, we assume that signal values are expressed in fixed point number format defined as

$$x = s\sum_{p=-F}^{I-1} b_p 2^p, \quad b_p \in \{0,1\}, \quad s \in \{-1, 1\},$$
$$I \geq 0, F \geq 0, \qquad I, F \in \mathbf{Z}, \tag{10}$$

where I is the number of assigned bits to the integer part without a sign bit, F is the number of assigned bits to the fractional part, and s is the sign bit. The sign bit is assigned to most significant bit. We define quantization of a real number x as Eq. (11).

$$\begin{cases} O[x] = \lfloor x' \rfloor = x' - (x' \mod 1), \ 0 \leq x' \mod 1 < 1, \\ x' = x + 2^{-1}. \end{cases} \tag{11}$$

In addition, after this, we normalize a real number W such as an input signal or filter coefficients by bit shifting, and proceed to a discussion on the assumption that W_{nor} is more than -1 but less than 1, where W_{nor} denotes normalized W. When the number of assigned bits to the integer part with a sign bit is I bits, the normalization process and its inverse process are expressed as Fig. 2. In case of a real number is quantized at V bits with a sign bit, Eq. (12) is employed.

$$Q_V[W_{nor}] = O[W_{nor}2^{(V-1)}]2^{-(V-1)}. \tag{12}$$

3.2 The Checkerboard Distortion Caused by Finite Word Length Error

In practical systems, the digital filter $H(z)$ must be quantized

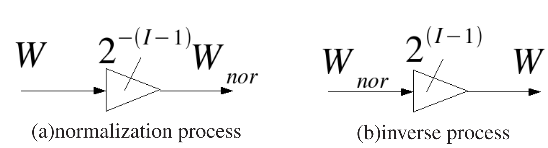


Fig. 2 The normalization process and its inverse process.

because of constraint of finite word length. The polyphase structure of this quantized filter is defined as

$$R'_i(z) = \sum_{n=-\infty}^{\infty} h'(Un - i - 1 + U)z^{-n}, \tag{13}$$

where

$$h'(Un - i - 1 + U) = Q_B[h(Un - i - 1 + U)], \tag{14}$$

and B denotes the number of assigned bits to filter coefficients. Even if $H(z)$ is designed under the condition of Eq. (6), there is a possibility that the value of step response has periodicity because the DC gain of each polyphase components $R'_i(z)$ is respectively not equivalent owing to the error caused by finite word length. As a result, the checkerboard distortion occurs.

3.3 Proposed Methods for Avoiding the Checkerboard Distortion

method 1: calculating the bit length of filter coefficients required for avoiding the checkerboard distortion

In general, the output values of a sampling rate converter are quantized at bit length of output signal. Thus, we can consider that the checkerboard distortion does not occur if the value of conclusive step response obtained after this quantization is constant, even if finite word length error exists. From here onwards, we can derive a conditional equation for the checkerboard distortion by using the relationship between filter coefficients, input value, and output value.

When bit length of the input signal is L bits, conceivable input step signals are given as

$$x_K(n) = K2^{-(L-1)}u(n),$$
$$K = -2^{(L-1)} + 1, -2^{(L-1)} + 2, \cdots, 2^{(L-1)} - 1, \tag{15}$$

In Sect. 2.2, we described the checkerboard distortion as periodical oscillation of the unit step response. However, it is not enough to deal with unit gain when we cannot suppose linearity property. Therefore, after this, we consider it as periodicity of a step response. When the input signals are the step signals $x_K(n)$, the output value of interpolator becomes steady state given as

$$s_K(m) = \begin{cases} Q_V[R'_0(1)x_K(1)], \\ \qquad (m = Un + (U-1)), \\ Q_V[R'_2(1)x_K(1)], \\ \qquad (m = Un + (U-2)), \\ \vdots \\ Q_V[R'_{U-1}(1)x_K(1)], \\ \qquad (m = Un), \end{cases} \tag{16}$$

where V denotes the bit length of output signal. Thus, the condition for avoiding the checkerboard distortion can be defined as

$$g'_0 = g'_1 = \cdots = g'_{U-1}, \tag{17}$$

where

$$g'_i = Q_V[f'_i], \tag{18}$$

and

$$f'_i = R'_i(1)x_K(1). \tag{19}$$

If Eq. (20) is satisfied, Eq. (17) would be approved (from appendix).

$$-2^{-V} < g'_{i'} - f'_{i'+1} \leq 2^{-V}, i' = 0, 1, \cdots, U-2. \tag{20}$$

Denoting the difference between $R'_{i'+1}(1)$ and $R'_{i'}(1)$ as

$$e_{i'} = R'_{i'+1}(1) - R'_{i'}(1), \tag{21}$$

the equation obtained from Eq. (21) and Eq. (19) can be written as

$$f'_{i'+1} = f'_{i'} + e_{i'}x_K(1). \tag{22}$$

The difference between $Q_V[f'_{i'}]$ and $f'_{i'}$ is written as

$$\Delta f'_{i'} = Q_V[f'_{i'}] - f'_{i'}, \tag{23}$$

additionally, Eq. (20) can be expressed as

$$-2^{-V} < \Delta f'_{i'} - e_{i'}x_K(1) \leq 2^{-V}. \tag{24}$$

Finally, Eq. (24) can be transformed as

$$\begin{cases} \frac{(-2^{-V}+\Delta f'_{i'})}{x_K(1)} \leq e_{i'} < \frac{(2^{-V}+\Delta f'_{i'})}{x_K(1)}, \\ \qquad (x_K(1) > 0), \\ \frac{(2^{-V}+\Delta f'_{i'})}{x_K(1)} < e_{i'} \leq \frac{(-2^{-V}+\Delta f'_{i'})}{x_K(1)}, \\ \qquad (x_K(1) < 0). \end{cases} \tag{25}$$

The checkerboard distortion does not occur if Eq. (25) are satisfied.

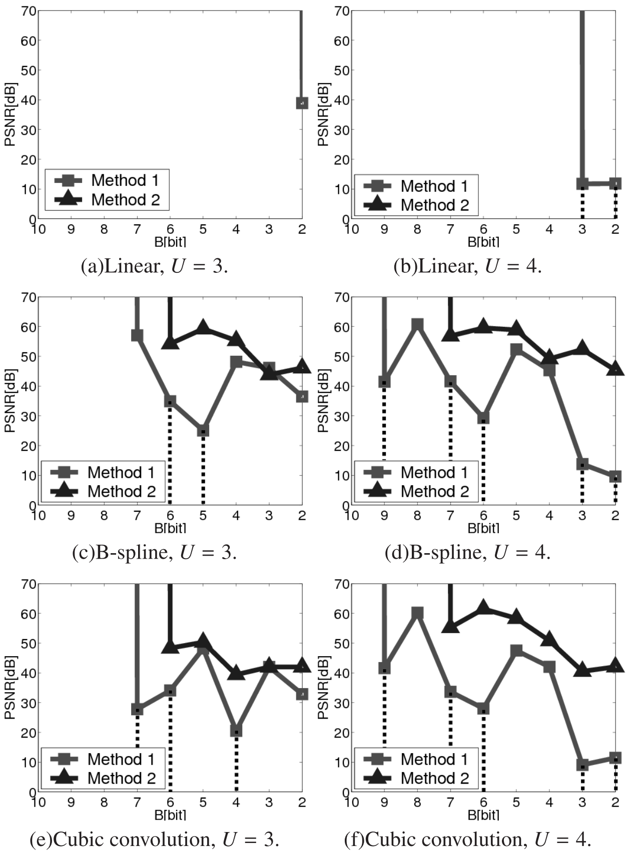
method 2: using the cascade structure with zero hold kernel

In method 1, it depends on not only the bit length of the filter but also input value and filter coefficients whether Eq. (25) is satisfied or not. Alternatively, we propose another method for avoiding the checkerboard distortion, which does not depend on such situations. A digital filter designed from Theorem in Sect. 2.2 can be factorized as shown by Eq. (7). Thus, the interpolator can consists of zero-order hold and a time-invariant filter $P(z)$. The checkerboard distortion can be avoided even though the time-invariant filter is quantized, because zero-order hold is unaffected by quantization, in addition, time-invariance of $P(z)$ is retained. Note that the filtering of $H_0(z)$ must be implemented before $P(z)$.

It is advisable to use the method 2, if the zero-order hold and time-invariant filter can be separately implemented. However, if this implementation cannot be applied, the use of method 1 is required.

4. Simulation

We demonstrate that the checkerboard distortion caused by



(a)Linear, $U = 3$. (b)Linear, $U = 4$.

(c)B-spline, $U = 3$. (d)B-spline, $U = 4$.

(e)Cubic convolution, $U = 3$. (f)Cubic convolution, $U = 4$.

Fig. 3 The results of interpolation.

finite word length error affects the image quality by image resolution conversion. We confirm that the checkerboard distortion could be avoided by using the proposed methods.

The image quality is evaluated by calculating PSNR between the ideal condition and the results of interpolation obtained by using method 1 and method 2. The ideal condition is the results obtained by using the filter with long enough word length. The bit length of input signal L and bit length of output signal V are respectively 8, and the test image is "Lena." The bit length of filter coefficients B are changed from 10 to 2. Linear, B-spline, and cubic convolution are used for interpolation function. The factor of upsampler U is two types of 3 and 4. The factor of downsampler D is 1. The filter of linear interpolation can be represented as

$$H(z) = \frac{1}{U} Q(z). \tag{26}$$

Thus, the multiplication by $1/U$ shown in Eq. (26) is implemented behind interpolator. Similarly, since B-spline and cubic convolution can be factorized $H(z)$ into $1/(6U^3)$ and the other, $1/(6U^3)$ is implemented at the last. Furthermore, if the DC gain of quantized filter is not 1, gain adjustment is implemented.

A results of interpolation are shown in Fig. 3. The dashed line is drawn on the word length that the checkerboard distortion occurred. In method 1, when the checkerboard distortion occurred, we can recognize that PSNR decreases significantly from Fig. 3. By contrast, in method 2, because the checkerboard distortion don't occur, high PSNR is generally obtained. Figure 4(a) shows the image with checkerboard distortion, which was interpolated threefold using method 1. Cubic convolution was used for interpolation function, additionally $B = 6$. In contrast, Fig. 4(b) is similarly the result interpolated using method 2. This image do not have the checkerboard distortion. From these results, we could say that if the method 2 can be implemented and employed, we can avoid the checkerboard distortion and may obtain nearly ideal image.

(a)method 1 (b)method 2

Fig. 4 A part of interpolation results. Cubic convolution, $B = 6$.

5. Conclusion

The checkerboard distortion caused by finite word length error was discussed. The two methods for avoiding the checkerboard distortion under finite word length were proposed. In the first method, the bit length of filter coefficients required for avoiding the checkerboard distortion was derived. In the second method, it was proposed that the checkerboard distortion can be avoided by using the cascade structure which consists of zero-hold kernel and a time-invariant filter factorized from the filter with structure for avoiding the checkerboard distortion under linear systems. It was presented by image resolution conversion that we can avoid the checkerboard distortion by using the proposed methods.

References

[1] A.V. Oppenheim, R.W. Schafer, and J.R. Buck, Discrete-Time Signal Processing, Pearson Education, 1999.

[2] F. Harris, Multirate Signal Processing for Communication Systems, Prentice Hall, 2004.

[3] S. Muramatsu, A. Yamada, and H. Kiya, "A design method of multidimensional linear-phase paraunitary filter banks with a lattice structure," IEEE Trans. Signal Process., vol.47, no.3, pp.690–700, March 1999.

[4] T. Tamura, M. Kato, T. Yoshida, and A. Nishihara, "Design of checkerboard-distortion-free multidimensional multirate filters," IEICE Trans. Fundamentals, vol.E81-A, no.8, pp.1598–1606, Aug. 1998.

[5] Y. Harada, S. Muramatsu, and H. Kiya, "Multidimensional multirate filter and filter bank without checkerboard effect," IEICE Trans. Fundamentals, vol.E81-A, no.8, pp.1607–1615, Aug. 1998.

Appendix: The Proof that Eq. (17) is Satisfied, if Eq. (20) Holds

Substituting Eq. (18) into Eq. (12), we obtain the following.

$$\begin{aligned} g'_{i'+1} &= Q_V[f'_{i'+1}] \\ &= O[f'_{i'+1}2^{(V-1)}]2^{-(V-1)}. \end{aligned} \qquad \text{(A· 1)}$$

Using Eq. (11), Eq. (A· 1) is rewritten as

$$\begin{aligned} g'_{i'+1} &= (f'_{i'+1}2^{(V-1)} + 2^{-1})2^{-(V-1)} - t_{i'+1} \\ &= f'_{i'+1} + 2^{-V} - t_{i'+1}, \end{aligned} \qquad \text{(A· 2)}$$

where

$$t_{i'+1} = \{(f'_{i'+1}2^{(V-1)} + 2^{-1}) \mod 1\}2^{-(V-1)}. \qquad \text{(A· 3)}$$

Then, from Eq. (A· 2), we obtain

$$f'_{i'+1} = g'_{i'+1} - 2^{-V} + t_{i'+1}. \qquad \text{(A· 4)}$$

Substituting Eq. (A· 4) into Eq. (20), we obtain

$$-2^{-(V-1)} + t_{i'+1} < g'_{i'} - g'_{i'+1} \leq t_{i'+1}. \qquad \text{(A· 5)}$$

Note that $g'_{i'} - g'_{i'+1}$ is given as

$$g'_{i'} - g'_{i'+1} = \{2^{-(V-1)}a \mid a \in \mathbf{Z},\ -2^V + 2 \leq a \leq 2^V - 2\}, \qquad \text{(A· 6)}$$

from Eq. (18). Thus, since $0 \leq t_{i'+1} < 2^{-(V-1)}$, the range of Eq. (A· 5) is given as

$$-2^{-(V-1)} \leq -2^{-(V-1)} + t_{i'+1} < g'_{i'} - g'_{i'+1} \leq t_{i'+1} < 2^{-(V-1)}. \qquad \text{(A· 7)}$$

When Eq. (A· 5) is satisfied, in other words, when Eq. (20) is satisfied, we obtain $g'_{i'} - g'_{i'+1} = 0$ from Eq. (A· 7) and Eq. (A· 6). Consequently, Eq. (17) is satisfied under Eq. (20).