論文 / 著書情報
Article / Book Information

Title	Analysis and Minimization of Output Errors of 2-D Non-separable FIR Digital Filters with Finite Precision Internal Signals
Authors	MITSUHIKO YAGYU, AKINORI NISHIHARA, NOBUO FUJII
出典 / Citation	IEICE Trans. Fundamentals., Vol. E80-A, No. 8, pp. 1391-1402
発行日 / Pub. date	1997,
URL	http://search.ieice.org/
権利情報 / Copyright	本著作物の著作権は電子情報通信学会に帰属します。 Copyright (c) 1997 Institute of Electronics, Information and Communication Engineers.

PAPER Special Section on Digital Signal Processing

Analysis and Minimization of Output Errors of 2-D Non-separable FIR Digital Filters with Finite Precision Internal Signals

Mitsuhiko YAGYU†, *Student Member*, Akinori NISHIHARA†, *and* Nobuo FUJII†, *Members*

SUMMARY This paper presents a method to analyze and minimize output errors of 2-D non-separable FIR filters with finite wordlength. Finiteness in the wordlength causes output errors, which can be analyzed in the frequency domain when the statistics of input signals are known. The output errors can be minimized by optimizing responses corresponding to all levels of input impulses. A new ROM-based filter structure is proposed in which the optimized impulse responses are stored in the ROM. The output signals are generated by superposing the impulse responses corresponding to the input levels. Many simulation results confirm that the output signals of the proposed filters have far less errors compared to conventional filters. The hardware size of the ROM-based filters is estimated and compared with that of conventional structures. The proposed structures are more effective than the conventional ones especially when the signal wordlength is short.

key words: *2-D non-separable FIR digital filters, finite wordlength, ROM-based filter*

1. Introduction

Two dimensional linear phase FIR digital filters are used in many applications such as image signal processing. In practical implementations, filters must have a finite wordlength of both coefficients and internal signals. Naturally, the output signals of the filters have errors due to those finiteness. If the filters were implemented as cascade forms, the finite wordlength errors could be reduced [1]. It is, however, difficult to implement 2-D non-separable FIR filters as cascade forms. Many methods have been proposed to design optimal 2-D linear phase non-separable FIR digital filters in the Chebyshev sense, when the wordlength of coefficients are given as finite numbers [2]–[5]. The finite wordlength effects of internal signals are not considered in those design procedures, and are usually dealt with as additive roundoff noise. The random roundoff noise model is merely a rough approximation, and is not suitable for accurate evaluation of the errors.

Firstly, this paper presents a method to analyze the output errors due to the finite wordlength. Since the wordlength of internal signals is finite, the products in the filters are rounded off and thus nonlinear operations are introduced. Then the output errors are most properly analyzed by their responses corresponding to all levels of input impulses rather than frequency responses. Naturally, the deteriorations in the responses are dependent on the levels of the input impulses. Mean squared output error spectra (MSOES) and maximum output error spectra (MOES) for random inputs are used as criteria.

Next, the output errors generated by the responses are minimized in the frequency domain by using the mixed integer linear programming (MILP) for every possible input level. To realize the minimal error filters, the optimized responses corresponding to every input level are stored in a ROM and the output signals are generated by superposing the stored responses which correspond to the input signal. Many examples show that the proposed filters are superior to the conventional ones.

Finally, the hardware size of the proposed filters is estimated and compared with that of conventional filters. Several results show that the proposed structures are more effective than the conventional ones especially when the signal wordlength is short.

Manuscript received December 12, 1996.
Manuscript revised March 4, 1997.

†The authors are with the Faculty of Engineering, Tokyo Institute of Technology, Tokyo, 152 Japan.

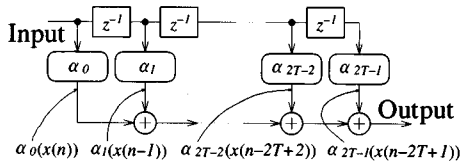
2. Generation of Output Signals

2.1 Internal Signal Rounding

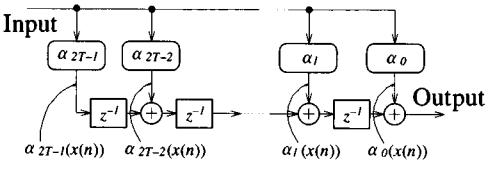
It is assumed that input, output and all internal signals in a 2-D linear phase FIR digital filter are represented as fixed point binary numbers of wordlength l. Negative numbers are expressed as two's complement. In filter operations, input signals are multiplied with filter coefficients. The products of two binary numbers of wordlength l are binary numbers of wordlength $2l-1$, if the products are computed in full resolution. The products are then rounded off to wordlength l. This rounding operation is a source of a nonlinear error.

2.2 Superpositions of Responses

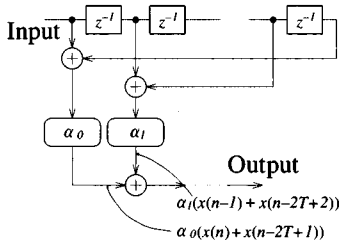
Consider at this moment a 1-D FIR filter having $2T$ taps as an example to make a discussion simple. Figures 1 (a) and (b) show the direct and the transposed forms of the filter which has the roundoff operations. In those figures, the operation block $\alpha_m(\cdot)$ for $0 \leqq m \leqq$



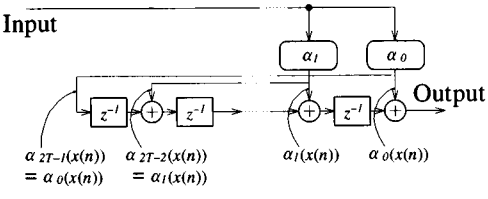
(a) A direct form.



(b) A transposed form.



(c) An efficient structure of a direct form.



(d) An efficient structure of a transposed form.

Fig. 1 Direct and transposed structures.

$2T-1$ indicates the multiplication by the m-th filter coefficient and the roundoff operation. Now $h(a,m)$ is defined as the rounded products of a multiplicand a and the m-th coefficient. Then $h(a,m)$ is written as

$$h(a,m) = \alpha_m(a). \tag{1}$$

In Figs. 1 (a) and (b), the output signal $y(n)$ is written as

$$y(n) = \sum_{m=0}^{2T-1} h(x(n-m),m) \tag{2}$$

$$= \sum_{k=n}^{n-2T+1} h(x(k),n-k). \tag{3}$$

Now $h(a,m)$ for $m<0$ or $m>2T-1$ can be assumed to be zero. Then $y(n)$ is written as

$$y(n) = \sum_{k=-\infty}^{\infty} h(x(k),n-k). \tag{4}$$

Since any input signals are trains of impulses of different levels, $h(x(m),n-m)$ is regarded as a response corresponding to an input impulse $x(m)$. Then the output signal $y(n)$ is obtained as the superpositions of the shifted responses corresponding to the impulses of input signals.

If the filter coefficients have symmetry, the filter can efficiently be implemented as shown in Figs. 1 (c) and (d). In Fig. 1 (d), the output signal is obtained as Eq. (4). However in Fig. 1 (c), the output signal is not the superpositions of the shifted responses, because the function $\alpha_m(\cdot)$ is nonlinear such as

$$\alpha_m(a+b) \neq \alpha_m(a) + \alpha_m(b). \tag{5}$$

Hereafter, the generations of output signals of 2-D non-separable FIR filters are considered. Let the input image signal be $x(n_1,n_2)$ and its size be $N_1 \times N_2$. The region where an image signal is defined is referred to as Φ. We define a vector $\boldsymbol{n}$ as $[n_1,n_2]^T$, $\boldsymbol{m}$ as $[m_1,m_2]^T$ and $\boldsymbol{\omega}$ as $[\omega_1,\omega_2]^T$. Now let the tap size of the 2-D filter be $T_1 \times T_2$. The case of odd T_1 and T_2 is analyzed as an example, but other cases can similarly be treated. Consider the input impulse at the point $\boldsymbol{n}$ and thus having the value $x(\boldsymbol{n})$. The corresponding response is expressed as $h(x(\boldsymbol{n}),\boldsymbol{m}-\boldsymbol{n})$ for $n_1 \leqq m_1+n_1 \leqq T_1-1+n_1$ and $n_2 \leqq m_2+n_2 \leqq T_2-1+n_2$. If the 2-D non-separable FIR filter is implemented as shown in Figs. 1 (a), (b) or (d), the output image signal $y(\boldsymbol{m})$ is obtained as

$$y(\boldsymbol{m}) = \sum_{\boldsymbol{n}\in\Phi} h(x(\boldsymbol{n}),\boldsymbol{m}-\boldsymbol{n}). \tag{6}$$

3. Output Error Analysis

3.1 Output Signal Estimation Sequence

The output signals are generated by superposing the responses corresponding to the levels of impulses of input signals as shown in Eq. (6). Naturally, those responses have errors due to the finite number of taps and wordlength. In this section, the effects of the errors on output signals are analyzed.

Firstly, we define an ideal 2-D linear phase FIR filter as F_d, and its frequency response $H_{ideal}(e^{j\omega_1},e^{j\omega_2})$ as

$$H_{ideal}(e^{j\omega_1},e^{j\omega_2}) = H_d(\boldsymbol{\omega})e^{-j\theta_T} \tag{7}$$

where $\theta_T = (T_1-1)\omega_1/2+(T_2-1)\omega_2/2$. The real function $H_d(\boldsymbol{\omega})$ has 1.0 in whole passband and 0.0 in whole stopband. The discrete-time Fourier transform (DTFT) of the ideal output of F_d can be written as

$$Y_d(e^{j\omega_1},e^{j\omega_2}) = \sum_{\boldsymbol{n}\in\Phi} x(\boldsymbol{n})H_d(\boldsymbol{\omega})e^{-j\boldsymbol{n}^T\boldsymbol{\omega}}e^{-j\theta_T}. \tag{8}$$

Two-D linear phase FIR filters have symmetric coefficients so that the response $h(x(\boldsymbol{n}),\boldsymbol{m})$ has the symmetry. Then the frequency-domain response is defined by the DTFT as

$$H(x(\boldsymbol{n}),\boldsymbol{\omega})e^{-j\theta_T} = \sum_{\boldsymbol{m}} h(x(\boldsymbol{n}),\boldsymbol{m})e^{-j\boldsymbol{m}^T\boldsymbol{\omega}} \tag{9}$$

where $H(x(\boldsymbol{n}),\boldsymbol{\omega})$ is a real function. By using Eqs. (6) and (9), the DTFT of the output signal is written as

$$Y(e^{j\omega_1},e^{j\omega_2}) = \sum_{\boldsymbol{n}\in\Phi} H(x(\boldsymbol{n}),\boldsymbol{\omega})e^{-j\boldsymbol{n}^T\boldsymbol{\omega}}e^{-j\theta_T}. \tag{10}$$

The error in the output signal is given as

$$R_o(e^{j\omega_1},e^{j\omega_2}) = Y(e^{j\omega_1},e^{j\omega_2}) - Y_d(e^{j\omega_1},e^{j\omega_2}). \tag{11}$$

Then we define an error included in $H(x(\boldsymbol{n}),\boldsymbol{\omega})$ as $R(x(\boldsymbol{n}),\boldsymbol{\omega})$, which can be written as

$$R(x(\boldsymbol{n}),\boldsymbol{\omega}) = H(x(\boldsymbol{n}),\boldsymbol{\omega}) - x(\boldsymbol{n})H_d(\boldsymbol{\omega}). \tag{12}$$

By using (8), (10), (11) and (12), the output error $R_o(e^{j\omega_1},e^{j\omega_2})$ can be written as

$$R_o(e^{j\omega_1},e^{j\omega_2}) = \sum_{\boldsymbol{n}\in\Phi} R(x(\boldsymbol{n}),\boldsymbol{\omega})e^{-j\boldsymbol{n}^T\boldsymbol{\omega}}e^{-j\theta_T}. \tag{13}$$

$R(x(\boldsymbol{n}),\boldsymbol{\omega})$ is a real function of $x(\boldsymbol{n})$ and $\boldsymbol{\omega}$. So if $\boldsymbol{\omega}$ is fixed, $R(x(\boldsymbol{n}),\boldsymbol{\omega})$, $n_1 = 0,1,\ldots$, $n_2 = 0,1,\ldots$ is a real sequence. We call this real sequence as an output signal estimation sequence (OSES). Equation (13) is the DTFT of the OSES. Then an input signal $x(\boldsymbol{n})$ determines an OSES $R(x(\boldsymbol{n}),\boldsymbol{\omega})$, $n_1 = 0,1,\ldots$, $n_2 = 0,1,\ldots$ corresponding to each frequency. The output error at $\boldsymbol{\omega}_a$ can be referred to as the value of the DTFT at $\boldsymbol{\omega}_a$ of the OSES corresponding to the frequency $\boldsymbol{\omega}_a$.

In Eq. (13), the DTFTs of the OSES, which are defined in the region Φ, are evaluated. If the region Φ is not so wide, those DTFTs may not be properly evaluated. The output error at $\boldsymbol{\omega}_n = [\omega_{n1},\omega_{n2}]^T$ is referred to as the inner product of the OSES corresponding to $\boldsymbol{\omega}_n$ and the complex sinusoid $e^{-j\boldsymbol{n}^T\boldsymbol{\omega}_n}$. So in the case of small size image, the output error is evaluated as

$$\tilde{R}(e^{j\omega_{n1}},e^{j\omega_{n2}}) = \sum_{\boldsymbol{n}\in\Phi} R(x(\boldsymbol{n}),\boldsymbol{\omega}_{\boldsymbol{n}})e^{-j\boldsymbol{n}^T\boldsymbol{\omega}_n}e^{-j\theta_T} - \sum_{\boldsymbol{n}\in\Phi}\frac{R(x(\boldsymbol{n}),\boldsymbol{\omega}_{\boldsymbol{n}})}{N_1N_2}\sum_{\boldsymbol{n}\in\Phi} e^{-j\boldsymbol{n}^T\boldsymbol{\omega}_n}e^{-j\theta_T}, \tag{14}$$

where $\boldsymbol{\omega}_n \neq \mathbf{o}$. For the evaluation of the output error at DC, Eq. (13) is used. When the size of Φ is not so large, the second term in the righthand side in Eq. (14) has large spectrum at frequencies $[\omega_{n1},0]$ and $[0,\omega_{n2}]$. That term becomes zero with the growth of the size of Φ at those frequencies. Accordingly, the removal of that term can somewhat neutralize the effects which are caused by the small size of Φ.

3.2 Mean Squared Output Error Spectrum (MSOES)

In this section, we show that the MSOES can be analyzed, even if the given input signals are not deterministic. Let the wordlength of the input signals be l and all the levels of input impulses be x_i, $i = 0,\ldots,L-1$ where $L = 2^l$. The OSES corresponding to a frequency $\boldsymbol{\omega}_a$ is $R(x(\boldsymbol{n}),\boldsymbol{\omega}_a)$ corresponding to an input signal $x(\boldsymbol{n})$ and then it is trains of elements in a set Ψ_{ω_a} given by

$$\Psi_{\omega_a} = \{R(x_i,\boldsymbol{\omega}_a) \mid i = 0,\ldots,L-1\}. \tag{15}$$

Now we assume that the stochastic input signals are stationarily independent process and that their probability density function is defined as $p(x_i)$, $i = 0,\ldots,L-1$. Then the OSESs become stationarily independent and thus their probability density function $q(R(x_i,\boldsymbol{\omega}_a))$ is obtained as

$$q(R(x_i,\boldsymbol{\omega}_a)) = p(x_i),\ i = 0,\ldots,L-1. \tag{16}$$

The mean power spectral density function of the OSESs corresponding to the frequency $\boldsymbol{\omega}_a$ is defined as $E[\mathcal{S}(e^{j\omega})]$ and given in Eq. (A·9). Then the MSOES at frequency $\boldsymbol{\omega}_a$ is obtained as $E[\mathcal{S}(e^{j\omega_a})]$. The second term in the right hand side in Eq. (A·9) becomes equivalent to the delta function and very large at DC, when N_1 and N_2 are infinite. So in this paper, the MSOES $E[\mathcal{P}(\boldsymbol{\omega})]$ is written as

$$E[\mathcal{P}(\boldsymbol{\omega})] \approx \begin{cases} R_{mse}(\mathbf{o}) + (N_1N_2-1)R_{me}^2(\mathbf{o}), & \boldsymbol{\omega}=\mathbf{o} \\ R_{mse}(\boldsymbol{\omega}) - R_{me}^2(\boldsymbol{\omega}) & \text{otherwise} \end{cases} \tag{17}$$

where

$$R_{me}(\boldsymbol{\omega}) = \sum_{i=0}^{L-1} R(x_i,\boldsymbol{\omega})p(x_i), \tag{18}$$

$$R_{mse}(\boldsymbol{\omega}) = \sum_{i=0}^{L-1} R^2(x_i,\boldsymbol{\omega})p(x_i). \tag{19}$$

If $R_{me}(\mathbf{o}) \neq 0$, the MSOES is very large at DC. In other frequencies, the MSOES is given as the variance of the error responses $R(x_i,\boldsymbol{\omega})$, $i = 0,\ldots,L-1$.

The spectra of the uncorrelated input signals $x(\boldsymbol{n})$ are constant. Its constant value is equal to the variance of the input signals. In this paper, the MSOES at DC and the peak of the MSOES are normalized with the mean spectrum of the input signals and also evaluated as follows.

$$\begin{cases} \dfrac{R_{me}^2(\boldsymbol{n})}{E^2[x(\boldsymbol{n})]} & \boldsymbol{\omega} = \mathbf{o} \\ \max\limits_{\substack{\boldsymbol{\omega} \\ \boldsymbol{\omega}\neq\mathbf{o}}} \dfrac{R_{mse}(\boldsymbol{\omega}) - R_{me}^2(\boldsymbol{\omega})}{V[x(\boldsymbol{n})]} & \text{otherwise} \end{cases} \tag{20}$$

where

$$E[x(\boldsymbol{n})] = \sum_{i=0}^{L-1} x_i p(x_i), \tag{21}$$

$$V[x(\boldsymbol{n})] = \sum_{i=0}^{L-1} x_i^2 p(x_i) - E^2[x(\boldsymbol{n})]. \tag{22}$$

3.3 Maximum Output Error Spectrum (MOES)

In this section, we present a method to obtain the maximum output error spectrum (MOES) at each frequency. An input signal determines an OSES corresponding to a frequency ω_a. Conversely, an arbitrary OSES corresponding to a frequency ω_a necessarily corresponds to at least a certain input signal. The MOES at ω_a is equivalent to the maximum spectrum at ω_a in all spectra of all OSESs corresponding to ω_a. Then the MOES $S(\omega_a)$ can be written as

$$S(\omega_a) = \lim_{\substack{N_1,N_2 \\ \to\infty}} \max_{\substack{R(x(\boldsymbol{n}),\omega_a) \\ \in\Psi_{\omega_a}}} \left| \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \frac{R(x(\boldsymbol{n}),\omega_a) e^{-j\boldsymbol{n}^T\omega_a}}{N_1 N_2} \right|. \tag{23}$$

Now let the maximum and the minimum elements in the set Ψ_{ω_a} be $R_{max}(\omega_a)$ and $R_{min}(\omega_a)$, respectively. Further, $R_{dis}(\omega_a)$ and $R_{mid}(\omega_a)$ are defined as follows.

$$R_{dis}(\omega_a) = \frac{R_{max}(\omega_a) - R_{min}(\omega_a)}{2}, \tag{24}$$

$$R_{mid}(\omega_a) = \frac{R_{max}(\omega_a) + R_{min}(\omega_a)}{2}. \tag{25}$$

As shown in Eqs. (A· 19) and (A· 20) in Appendix B, the MOES $S(\omega_a)$ can be obtained as

$$S(\omega_a) = R_{dis}(\omega_a) w(\omega_a). \tag{26}$$

$R_{dis}(\omega_a)$ is determined by the dynamic range of the OSES $R(x(\boldsymbol{n}),\omega_a)$. From Eq. (26), the MOES $S(\omega_a)$ is given as a product by $R_{dis}(\omega_a)$ and the inherent weight $w(\omega_a)$. In the same way, if input signals satisfy $x_{min} \leqq x(\boldsymbol{n}) \leqq x_{max}$, the maximum spectrum at ω_a in all input signal spectra can be obtained as a product of $(x_{max} - x_{min})/2$ and the weight. To cancel the weight, which is intractable, the ratio of the MOES and the maximum input spectrum is introduced. Then to evaluate the maximum output errors of filters, we introduce the following normalized worst-case ratio as

$$\max_{0<\omega_a\leqq\pi} \left\{ \frac{2R_{dis}(\omega_a)}{x_{max}-x_{min}}, \frac{\max\{|R_{max}(\mathbf{o})|, |R_{min}(\mathbf{o})|\}}{\max\{|x_{max}|, |x_{min}|\}} \right\}. \tag{27}$$

3.4 Ideal Error Responses

In the above sections, the MSOES and the MOES have been analyzed. In this section, we consider the error responses corresponding to all the input levels such that the MSOES and the MOES are zero.

Firstly the MSOES and the MOES at DC are considered. From Eqs. (17) and (27), all the error responses $R(x_i, \mathbf{o})$ must have zero so that the MSOES and the MOES at DC are zero. Next in the case of other frequencies, the condition for the error responses to satisfy is considered. In Eq. (17), the MSOES at those frequencies is given as the variance of all the error responses. Accordingly, if and if only all the error responses are equal to their mean error response, the MSOES also becomes zero. In that case, the MOES becomes zero. Then if all the error responses satisfy

$$R(x_i,\omega) = \sum_{k=0}^{L-1} R(x_k,\omega) p(x_k), \quad i = 0,\ldots,L-1 \tag{28}$$

and

$$R(x_i, \mathbf{o}) = 0, i = 0,\ldots,L-1, \tag{29}$$

the MSOES and the MOES are zero in whole frequency range. We call those error responses *ideal* error responses. If the responses $H(x_i,\omega)$ have the *ideal* error responses, the MOES is zero in whose frequency. This means that an output error corresponding to an arbitrary input signal is zero.

4. New ROM-Based Filters

4.1 Implementation

We propose a filter structure having a ROM as shown in Fig. 2. Figure 2 shows a 1-D filter structure, but 2-D structures can be implemented in the same way. In the design of the ROM-based filters, the responses corresponding to all the input levels are optimized to minimize the output errors and stored in the ROM. Those responses are successively referred. Then by superposing the referred responses, the output signals are computed in the ROM-based filters.

4.2 Optimization of All the Responses

From Eq. (17), the MSOES is given as the variance of the error responses. Then the MSOES at frequencies but DC can be written as

$$E[\mathcal{P}(\omega)] = \sum_{i=0}^{L-1} p(x_i) \{R(x_i,\omega) - R_{me}(\omega)\}^2. \tag{30}$$

Only by optimizing all the error responses simultaneously, an optimum solution can be obtained. It is, however, difficult to carry out that optimization due to its

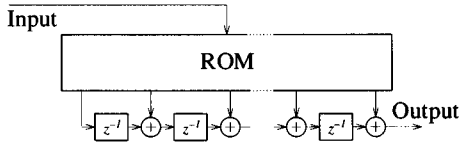


Fig. 2 A ROM-based structure.

enormous computing cost. Accordingly, we propose a method that $||R(x_i,\omega)-R_{me}(\omega)||$, $i=0,\cdots,L-1$ are minimized under the condition $R(x_i,\mathbf{o})=0$, iteratively, while the mean error response $R_{me}(\omega)$ is updated. By using that method, the MSOES shown in Eq. (17) can be minimized.

In the spatial domain, the responses $h(x_i,\boldsymbol{n})$ of linear phase filters have the symmetry. Then $H(x_i,\omega)$ can be written as

$$H(x_i,\omega)=A(\omega)\boldsymbol{h}_{x_i} \tag{31}$$

where $\boldsymbol{A}(\omega)$ is a vector whose elements are trigonometric functions and $\boldsymbol{h}_{x_i}$ is a vector whose elements are the independent coefficients of $h(x_i,\boldsymbol{n})$ [6]. We propose the following algorithm, which is referred to as the algorithm 1, to obtain the responses corresponding to all the input levels.

1. Let elements in a set Λ be probability densities $p(x_i)$, $i=0,\ldots,L-1$, $R^*_{me}(\omega):=0$ and $s:=0$.
2. If the set Λ is empty, then stop.
3. Choose $p(x_k)$ which is the largest value in all elements in the set Λ. Exclude the element $p(x_k)$ from the set Λ.
4. The following mixed integer linear programming (MILP) problem is solved.

$$\begin{aligned} &\textit{Minimize} && ||\boldsymbol{A}(\omega)\boldsymbol{h}_{x_k}-x_kH_d(\omega)-R^*_{me}(\omega)||_\infty,\\ &\textit{subject to} && \boldsymbol{A}(\mathbf{o})\boldsymbol{h}_{x_k}=x_kH_d(\mathbf{o}) \text{ and each element of } \boldsymbol{h}_{x_k} \text{ has wordlength } l. \end{aligned}$$

5. By using the obtained $\boldsymbol{h}_{x_k}$, $R^*_{me}(\omega)$ is updated by

$$R^*_{me}(\omega) := \frac{sR^*_{me}(\omega)+p(x_k)\{\boldsymbol{A}(\omega)\boldsymbol{h}_{x_k}-x_kH_d(\omega)\}}{s+p(x_k)}. \tag{32}$$

6. Let $s:=s+p(x_k)$ and go to Step 2.

In the practical image signal processing, the probability density functions $p(x_i)$ of image signals are not known *a priori*. To prepare for various input signals, the probability density function used in the algorithm 1 is given as a uniform distribution. Then the probability density function $p(x_i)$ can be written as

$$p(x_i)=\frac{1}{L},\quad i=0,\ldots,L-1. \tag{33}$$

Let x_{me} be a mean value of input signals $x(\boldsymbol{n})$. In Step 3, the largest values of the probability densities are successively chosen. If the probability density function $p(x_i)$ is the uniform distribution, those largest values can not be determined uniquely. Therefore in that case, we modify Step 3 as follows.

3. Choose x_k which is the nearest value to x_{me} in all elements in the set Λ. Exclude the element x_k from the set Λ.

If a probability density function of input signals is known *a priori*, the original procedure is used as Step 3. Then better solutions can be obtained.

The MILP problem can be solved by using the branch and bound algorithm. $R^*_{me}(\omega)$ is the provisional mean error response during the optimization of the error responses. In Step 4, the difference between the error response $H(x_k,\omega)-x_kH_d(\omega)$ and $R^*_{me}(\omega)$ is minimized. The algorithm 1 minimizes not the error responses but the MSOES and the MOES. The responses $H(x_i,\omega)$ $i=0,\ldots,L-1$ are optimized so as to have similar error responses. Accordingly, $H(x_k,\omega)$ optimized by using the algorithm 1 may have large deviations from $x_kH_d(\omega)$.

To clarify effectiveness of the algorithm 1, consider another minimization subject to $R^*_{me}(\omega)=0$ in Step 4. This modified algorithm is called the algorithm 2. The algorithm 2 minimizes the error response $||H(x_i,\omega)-x_iH_d(\omega)||_\infty$, where $i=0,\ldots,L-1$. Accordingly in this algorithm, the deviations of $H(x_i,\omega)$, $i=0,\ldots,L-1$ from $x_iH_d(\omega)$ are minimized. Thus the error shown in Eq. (17) is not minimized in this algorithm.

5. Design Examples

5.1 Specifications

The ROM-based filters obtained by using the algorithm 1 are compared with the conventional filters where the products are rounded off. In this paper, we give the probability density function of input signals as the uniform distribution shown in Eq. (33). Filter specifications are as follows.

	tap size	:	9×9
	wordlength	:	6, 8
Type I	passband	:	$\mid\omega 1\mid+\mid\omega 2\mid\leqq 0.2\pi$
	stopband	:	$\mid\omega 1\mid+\mid\omega 2\mid\geqq 0.6\pi$
Type II	passband	:	$\mid\omega 1\mid+\mid\omega 2\mid\leqq 0.04\pi$
	stopband	:	$\mid\omega 1\mid+\mid\omega 2\mid\geqq 0.4\pi$

The coefficients of conventional filters are designed by using MILP [2], [7]. These coefficients are designed under the condition that the errors at DC are strictly zero, because the power spectra of image signals are usually very high at DC. The conventional filters have the roundoff operations. To ensure an accuracy in binary arithmetic, we assume that multiplications are carried out in full resolution and then only the outputs of multipliers are rounded off. To calculate PSNRs of output images of the conventional and the ROM-based filters, we need a reference which is regarded as the output images of the ideal filters. For this purpose, a 25×25 tap

FIR filter is designed by using linear programming [6], which has −67.7 dB Chebyshev error under the DC response condition.

5.2 Detailed Comparisons in a Specification

Firstly in the specification wordlength 6 and Type I, the ROM-based filters designed by the algorithm 2 also are compared. To analyze the MSOES of the designed filters, probability density functions of input signals are needed. We use four types of probability density functions shown in Figs. 3(a), 4(a), 5(a) and 6(a) and referred to as PDF1, 2, 3 and 4. PDF1 is the sample probability density function that used in the algorithm 1. Using Eq. (20), we obtain the MSOES of the designed filters normalized with the power spectrum of the input signals. Figures 3, 4, 5 and 6 show the un-normalized MSOES of the ROM-based filters designed by the algorithm 1 and 2 and the conventional filter. Table 1 shows the normalized peaks of MSOES of the filters, which are defined in Eq. (20). From this table, the normalized MSOES of the conventional filter at DC are small but nonzero. In such cases, the un-normalized MSOES at DC defined in Eq. (17) become very large, when the size of input image signals is larger. The ROM-based filters generate no error at DC, but the conventional filter generates very large un-normalized MSOES at DC. Responses designed by the algorithm 1 have similar error responses. Then the MSOES of the ROM-based filter designed by the algorithm 1 are not so much dependent on the probability density functions of input signals. Error responses designed by the algorithm 2 are smaller but not similar to one another. So the MSOES of the ROM-based filter designed by the algorithm 2 and the conventional filter are strongly dependent on the probability density functions of input signals and comparatively large.

Next we compare practical output images of the ROM-based and the conventional filters and confirm the effectiveness of the ROM-based filters. The stan-

Table 1 Peak value of MSOES [dB].

(a) PDF1.

	Algorithm 1	Algorithm 2	Conventional
$\omega \neq 0$	−12.9	−11.8	−8.0
$\omega = 0$	$-\infty$	$-\infty$	−32.7

(b) PDF 2.

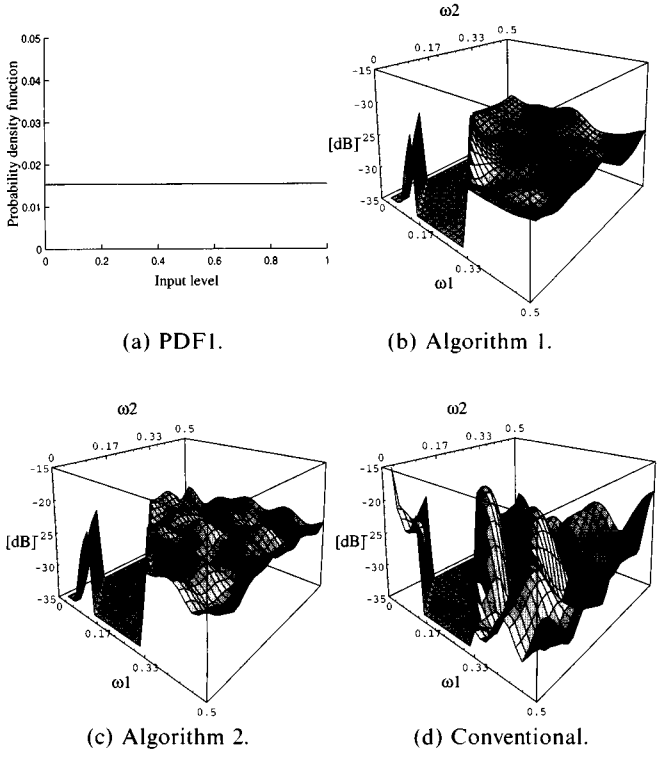
	Algorithm 1	Algorithm 2	Conventional
$\omega \neq 0$	−7.10	−3.81	−4.90
$\omega = 0$	$-\infty$	$-\infty$	−31.4

(c) PDF 3.

	Algorithm 1	Algorithm 2	Conventional
$\omega \neq 0$	−9.38	−7.21	−0.46
$\omega = 0$	$-\infty$	$-\infty$	−27.8

(d) PDF 4.

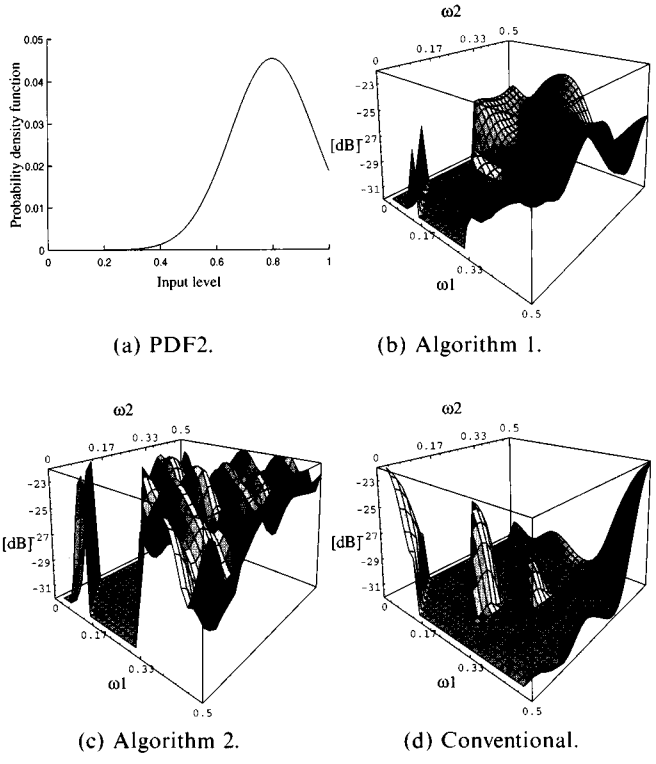
	Algorithm 1	Algorithm 2	Conventional
$\omega \neq 0$	−7.62	−7.78	−4.97
$\omega = 0$	$-\infty$	$-\infty$	−15.13



(a) PDF1. (b) Algorithm 1.

(c) Algorithm 2. (d) Conventional.

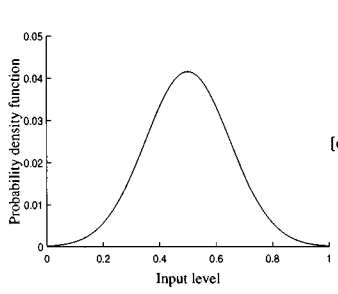
Fig. 3 MSOES of designed filters — PDF1.



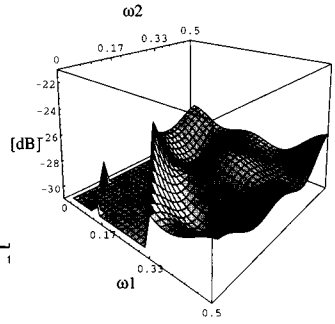
(a) PDF2. (b) Algorithm 1.

(c) Algorithm 2. (d) Conventional.

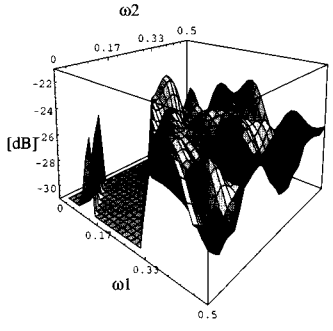
Fig. 4 MSOES of designed filters — PDF2.



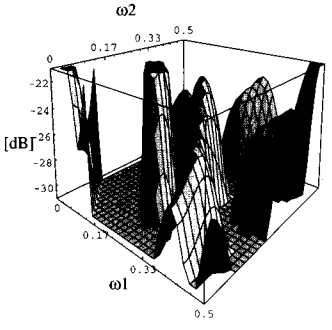
(a) PDF3.



(b) Algorithm 1.

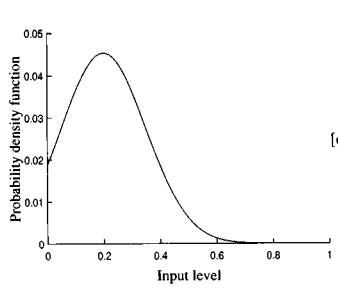


(c) Algorithm 2.

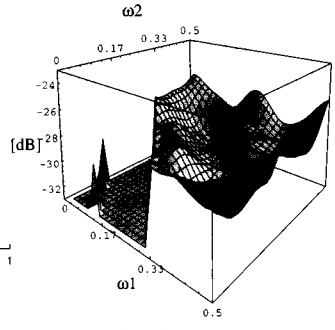


(d) Conventional.

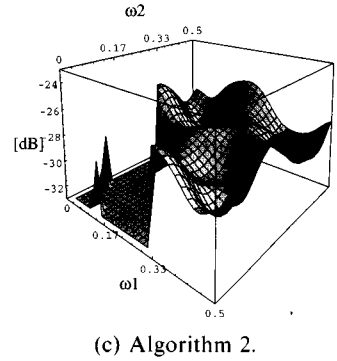
Fig. 5 MSOES of designed filters — PDF3.



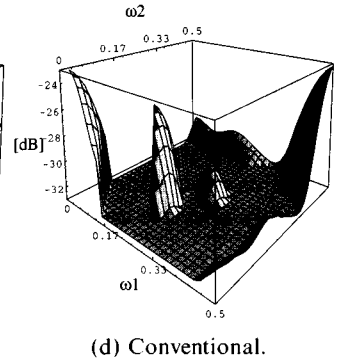
(a) PDF4.



(b) Algorithm 1.



(c) Algorithm 2.



(d) Conventional.

Fig. 6 MSOES of designed filters — PDF4.

(a) Input image Lenna, 6 bit/pixel.

(b) Ideal output image.

Fig. 7 Input and ideal output images.

dard image Lenna is quantized to 6 bits. The quantized image is filtered by using the ROM-based, the conventional and the ideal filters. Figures 7 (a) and (b) show the input and the ideal output images, respectively. The ideal output images are created by using the ideal filter. Figures 8 (a), (b) and (c) show the output images of the ROM-based filters designed by the algorithm 1 and 2 and the conventional filter, respectively. The binary arithmetic is carried out in the ROM-based and the conventional filters, but the real arithmetic in the ideal filter. In Fig. 8, the output image of the conventional filter has false contours, but they are not observed in the images by the ROM-based filters. It can be seen that the output image of the ROM-based filter designed by the algorithm 2 has larger error spectrum in high frequencies. The output images of the conventional filters have very large errors at low frequencies. Since the ROM-based filters designed by the algorithm 2 do not generate any errors at DC, those filters are more effective than the conventional filters. The algorithm 2, however, does not minimizes the output error spectrum in other frequencies. The algorithm 1 minimizes the output error

(a) Algorithm 1, PSNR = 26.0 [dB].

(b) Algorithm 2, PSNR = 22.1 [dB].

(c) Conventional, PSNR = 15.2 [dB].

Fig. 8 Output images obtained by ROM-based and conventional filters.

spectrum in whole frequency range under the condition of no error at DC. Accordingly, the output images of the ROM-based filters designed by the algorithm 1 have far smaller errors in whole frequency range than those designed by the algorithm 2.

Table 2 Peak value of MSOES [dB].

		Type I	
		Conventional	Algorithm 1
$l = 6$	$\omega \neq 0$	−8.0	−12.9
	$\omega = 0$	−32.7	$-\infty$
$l = 8$	$\omega \neq 0$	−20.7	−22.0
	$\omega = 0$	−55.4	$-\infty$
		Type II	
		Conventional	Algorithm 1
$l = 6$	$\omega \neq 0$	−1.58	−12.1
	$\omega = 0$	−36.3	$-\infty$
$l = 8$	$\omega \neq 0$	−17.7	−21.6
	$\omega = 0$	−54.2	$-\infty$

Table 3 Peak value of MOES [dB].

	Type I	
	Conventional	Algorithm 1
$l = 6$	−8.52	−13.0
$l = 8$	−21.3	−22.2
	Type II	
	Conventional	Algorithm 1
$l = 6$	−7.18	−12.6
$l = 8$	−19.2	−22.2

5.3 Comparisons in All the Given Specifications

In the above section, it has been shown that the algorithm 1 is more efficient than the algorithm 2. So in this section, filters are designed by using the algorithm 1 and the conventional method to meet the specifications Type I, II and wordlength 6, 8 and then compared. Tables 2 and 3 show the normalized MSOES and the normalized MOES of the designed filters, respectively. From these Tables, the normalized MSOES at DC of the conventional filters are small but nonzero. Then the un-normalized MSOES at DC become very large with the growth of the size of input image signals. The normalized MSOES of the ROM-based filters designed by using the algorithm 1 are strictly zero at DC. In this case, the un-normalized MSOES at DC are independent of the size of input image signals and are strictly zero. In other frequencies, the ROM-based filters designed by using the algorithm 1 have far smaller unnormalized MSOES than the conventional filters. Next many standard images shown in Table 4 are quantized to 6 and 8 bits and then filtered by using the designed filters. Table 4 shows peak values of output errors shown in Eq. (14) and PSNR. From Table 4, the ROM-based filters designed by using the algorithm 1 generate far smaller errors than the conventional filters. When the wordlength is short, the ROM-based filters are more especially effective compared to the conventional filters.

5.4 Comparison of Hardware Size

In this section, the hardware size of the ROM-based filters and the conventional filters are compared. The ROM-based filters have a ROM which has an input and

Table 4 Peaks of output error spectra.

(a) Lenna.

		Type I		Type II	
		Conventional	Algorithm 1	Conventional	Algorithm 1
$l=6$	$\omega \neq 0$	9.88×10^{0}	1.67×10^{-2}	1.31×10^{2}	4.10×10^{-2}
	$\omega = 0$	7.65×10^{1}	0.0	1.29×10^{2}	0.0
	PSNR[dB]	15.2	26.0	5.33	24.5
$l=8$	$\omega \neq 0$	2.24×10^{-1}	2.38×10^{-3}	9.40×10^{-1}	2.69×10^{-3}
	$\omega = 0$	2.94×10^{-1}	0.0	6.22×10^{-2}	0.0
	PSNR[dB]	30.5	33.0	25.5	32.1

(b) Swiss Mountain.

		Type I		Type II	
		Conventional	Algorithm 1	Conventional	Algorithm 1
$l=6$	$\omega \neq 0$	1.74×10^{0}	2.21×10^{-2}	2.75×10^{1}	4.23×10^{-2}
	$\omega = 0$	6.18×10^{1}	0.0	2.63×10^{0}	0.0
	PSNR[dB]	17.5	22.9	8.53	21.5
$l=8$	$\omega \neq 0$	4.10×10^{-2}	3.31×10^{-3}	2.64×10^{-1}	5.42×10^{-3}
	$\omega = 0$	1.09×10^{0}	0.0	6.06×10^{0}	0.0
	PSNR[dB]	30.6	30.7	27.4	30.3

(c) Girl.

		Type I		Type II	
		Conventional	Algorithm 1	Conventional	Algorithm 1
$l=6$	$\omega \neq 0$	1.45×10^{0}	1.31×10^{-2}	1.46×10^{1}	4.33×10^{-2}
	$\omega = 0$	9.33×10^{1}	0.0	1.54×10^{3}	0.0
	PSNR[dB]	14.2	21.1	3.97	19.6
$l=8$	$\omega \neq 0$	2.48×10^{-2}	1.96×10^{-3}	3.91×10^{-1}	2.29×10^{-3}
	$\omega = 0$	5.63×10^{0}	0.0	1.22×10^{-1}	0.0
	PSNR[dB]	27.2	28.6	23.1	27.3

(d) Lady.

		Type I		Type II	
		Conventional	Algorithm 1	Conventional	Algorithm 1
$l=6$	$\omega \neq 0$	3.53×10^{0}	1.42×10^{-2}	3.17×10^{1}	2.43×10^{-2}
	$\omega = 0$	1.49×10^{1}	0.0	5.75×10^{2}	0.0
	PSNR[dB]	16.8	26.2	6.65	24.3
$l=8$	$\omega \neq 0$	1.29×10^{-1}	3.54×10^{-3}	8.42×10^{-2}	1.89×10^{-3}
	$\omega = 0$	1.99×10^{-1}	0.0	1.71×10^{0}	0.0
	PSNR[dB]	32.3	33.2	28.1	32.2

(e) Moon.

		Type I		Type II	
		Conventional	Algorithm 1	Conventional	Algorithm 1
$l=6$	$\omega \neq 0$	1.61×10^{0}	8.97×10^{-3}	3.68×10^{1}	2.97×10^{-2}
	$\omega = 0$	1.49×10^{1}	0.0	3.74×10^{2}	0.0
	PSNR[dB]	15.4	26.1	5.11	24.6
$l=8$	$\omega \neq 0$	4.91×10^{-2}	4.69×10^{-3}	2.97×10^{-1}	1.55×10^{-3}
	$\omega = 0$	5.24×10^{-2}	0.0	2.40×10^{0}	0.0
	PSNR[dB]	30.2	33.5	25.6	32.3

several outputs, adders and delays as shown in Fig. 2. We assume that the ROM is realized as a mask-ROM. A ROM structure, which is shown in [8], is exemplified in Fig. 9 for a case of a wordlength 4. As shown in Fig. 9, a decoder of a ROM is realized as a matrix of circuit elements called the AND plane. This AND plane decodes inputs of a ROM-based filter so that corresponding data words may be selected. The output of the AND plane leaves at right angles to its input and run horizontally through other matrices called the OR planes. The outputs of the OR planes are the outputs of the ROM shown in Fig. 2.

In the implementations of the conventional filters, filter coefficients are represented as canonic signed digit (CSD) [2], [7]. Then multipliers are implemented by using adders and subtractors [9]. The conventional filters have the roundoff operations where $2l-1$ bit products are rounded off to l bit numbers. To carry out such roundoff operations, only the $l+1$ bits of outputs of multipliers are needed. Since the unnecessary $l-2$ bits may not be computed, the hardware size of multipliers in the conventional filters can be reduced. Then the conventional filters have logic circuits to round off only $l+1$ bit numbers.

The ROM-based filters have the same number of delays used in the conventional filters. Accordingly, we compare the number of transistors used in the ROM, adders, subtractors and the roundoff circuits. To estimate the number of transistors used in those circuit blocks, we assume that a full adder, a half adder and the

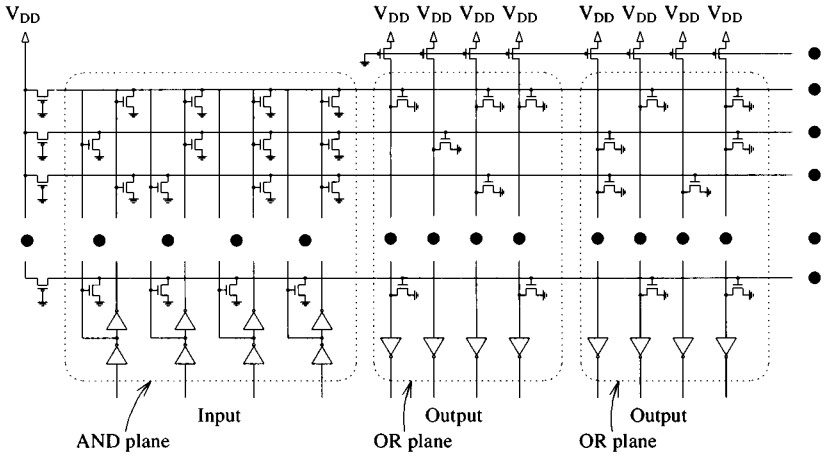
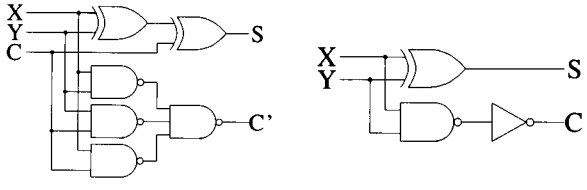
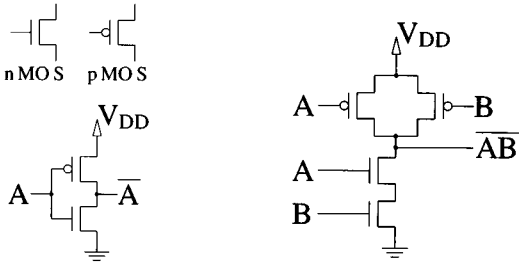


Fig. 9 A ROM structure.

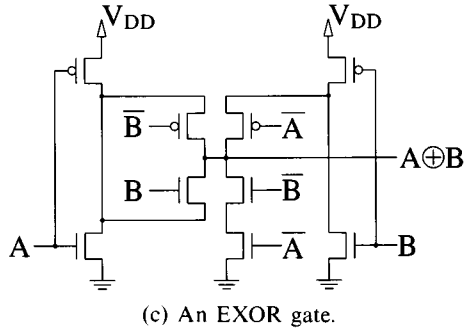


(a) A full adder. (b) A half adder.

Fig. 10 Adder structures.



(a) An inverter gate. (b) A NAND gate.



(c) An EXOR gate.

Fig. 11 Transistor diagrams of gates.

ROM are realized as CMOS circuits shown in Figs. 9, 10 and 11. The roundoff operation is performed with an LSB full adder by feeding a carry input.

Table 5 shows the number of transistors used in the conventional and the ROM-based filters which are designed by using the algorithm 1. In Table 5, there is a case that the ROM-based filters can be implemented by using 60% of transistors used in the conventional filter. In general when the wordlength is short, the ROM-based filters have the effectiveness. In other cases, the numbers of transistors used in the ROM-based filters are larger than those of the conventional filters, but the differences between those numbers of both filters are less than 10%. The ROM-base filters having long wordlength could not efficiently be implemented.

Table 5 Number of transistors used in designed filters.

	Type I	
	Conventional	ROM-based
$l = 6$	17946	10215
$l = 8$	30708	32818
	Type II	
	Conventional	ROM-based
$l = 6$	15120	15290
$l = 8$	31914	30823

6. Conclusion

This paper has presented a method to analyze and minimize output errors of 2-D FIR digital filters with finite precision internal signals. Such filters generate nonlinear output errors. In the proposed method, responses corresponding to input impulses of all possible levels have been considered so that the output errors have been accurately analyzed. The responses corresponding to all input impulses are optimized to minimize the analyzed output errors. We have proposed ROM-base filters where the optimized responses are stored. The output signals are generated by superposing the impulse responses corresponding to the input impulses. In many design examples, we have confirmed the superiority of the proposed filters to the conventional filters, where the roundoff operations are carried out. To estimate the hardware size of the designed filters, the numbers

of transistors used in the ROM-based and the conventional filters are compared. We have found that there is a case that the ROM-based filters can be implemented by using 60% of those used in the conventional filter in many design examples. In other cases, the numbers of transistors used in the ROM-based filters have been larger than those of the conventional filters, but the differences between those numbers of both filters have been less than 10% of those numbers.

Error spectrum shaping techniques [10], [11] can be applied to the ROM-based filters. Then the output errors could be reduced.

References

[1] L.R. Rabiner and B.Gold, "Theory and application of digital signal processing," Prentice-Hall, 1975.
[2] P. Siohan and A. Benslimane, "Finite precision design of optimal linear phase 2-D FIR digital filters," IEEE Trans. Circuits & Syst., vol.36, no.1, pp.11–22, Jan. 1989.
[3] N. Benvenuto, M. Marchesi, and A. Uncini, "Applications of simulated annealing for the design of special digital filters," IEEE Trans. Signal Process., vol.40, no.2, pp.323–332, Feb. 1992.
[4] J. Radecki, J. Konrad, and E. Dubois, "Design of multidimensional finite-wordlength FIR and IIR filters by simulated annealing," IEEE Trans. Circuits & Syst., vol.42, no.6, pp.424–431, Jan. 1995.
[5] M. Park and W. Song, "A new design method of 2-d linear phase FIR filters with finite-precision coefficients," IEEE Trans. Circuits & Syst., vol.41, no.7, pp.478–482, July 1994.
[6] C. Charalambous, "The performance of an algorithm for minimax design of two-dimensional linear phase FIR digital filters," IEEE Trans. Circuits & Syst., vol.CAS-32, no.10, pp.1016–1028, Oct. 1985.
[7] Y.C. Lim, "Design of discrete-coefficient-value linear phase FIR filters with optimum normalized peak ripple magnitude," IEEE Trans. Circuits & Syst., vol.37, no.12, pp.1480–1486, Dec. 1990.
[8] C. Mead and L. Conway, "Introduction to VLSI systems," Addison-Wesley Series in Computer Science, 1980.
[9] R. Jain, P.T. Yang, and T. Yoshino, "FIRGEN: A computer-aided design system for high performance FIR filter integrated circuits," IEEE Trans. Signal Process., vol.39, no.7, pp.1655–1668, July 1991.
[10] M. Iwahashi, T. Johba, and N. Kambayashi, "Finite word length expression of FIR digital filter coefficients using delta-sigma modulation," IEICE Technical Report, DSP95-129, Dec. 1995.
[11] T. Johba, A. Oguma, M. Iwahashi, and N. Kambayashi, "Finite word length expression of FIR digital filter coefficients using error spectrum shaping, and the reduction of operational error," IEICE Technical Report, DSP95-186, March 1996.

Appendix A: The Mean Power Spectral Density Function of OSESs

In this section, the OSES $R(x(\boldsymbol{n}), \boldsymbol{\omega_a})$ is called $v(\boldsymbol{n})$, briefly. The stochastic OSESs $R(x(\boldsymbol{n}), \boldsymbol{\omega_a})$ are stationarily independent process and have the probability density function $q(R(x_i, \boldsymbol{\omega_a}))$ $i = 0, \ldots, L-1$. By using Eqs. (15) and (16), the mean and the variance of OSESs can be obtained as

$$E[v(\boldsymbol{n})] = \sum_{i=0}^{L-1} R(x_i, \boldsymbol{\omega_a}) p(x_i) \tag{A·1}$$

and

$$V[v(\boldsymbol{n})] = \sum_{i=0}^{L-1} R^2(x_i, \boldsymbol{\omega_a}) p(x_i) - E^2[v(\boldsymbol{n})]. \tag{A·2}$$

Now we define another sequence $\hat{v}(\boldsymbol{n})$ satisfying

$$\hat{v}(\boldsymbol{n}) \triangleq v(\boldsymbol{n}) - E[v(\boldsymbol{n})]. \tag{A·3}$$

Then $\hat{v}(\boldsymbol{n})$ also satisfies

$$E[\hat{v}(\boldsymbol{n})] = 0, \tag{A·4}$$

$$E[\hat{v}^2(\boldsymbol{n})] = E[v^2(\boldsymbol{n})] - E^2[v(\boldsymbol{n})] = V[v(\boldsymbol{n})] \tag{A·5}$$

and

$$E[\hat{v}(\boldsymbol{n})\hat{v}(\boldsymbol{n} - \boldsymbol{m})] = 0 \tag{A·6}$$

where $\boldsymbol{m} \neq \mathbf{o}$. Now let $E[\mathcal{S}(e^{j\omega})]$ be the mean power spectral density function of $v(\boldsymbol{n})$ and $E[\hat{\mathcal{S}}(e^{j\omega})]$ be that of $\hat{v}(\boldsymbol{n})$, respectively. By using (A·6), $E[\hat{\mathcal{S}}(e^{j\omega})]$ is obtained as

$$E[\hat{\mathcal{S}}(e^{j\omega})] = E[\hat{v}^2(\boldsymbol{n})]. \tag{A·7}$$

By using $\hat{v}(\boldsymbol{n})$, $E[\mathcal{S}(e^{j\omega})]$ can be written as

$$E[\mathcal{S}(e^{j\omega})] = E[\hat{\mathcal{S}}(e^{j\omega})] + \frac{E^2[v(\boldsymbol{n})]}{N_1 N_2} \left| \sum_{\boldsymbol{n} \in \Phi} e^{-j\boldsymbol{n}^T \omega} \right|^2 + \frac{2E[v(\boldsymbol{n})]}{N_1 N_2} \sum_{\boldsymbol{n} \in \Phi} E[\hat{v}(\boldsymbol{n})] \sum_{\boldsymbol{m} \in \Phi} \cos[\boldsymbol{n} - \boldsymbol{m}]^T \boldsymbol{\omega}. \tag{A·8}$$

By substituting Eqs. (A·4), (A·5) and (A·7), Eq. (A·8) can be rewritten as

$$E[\mathcal{S}(e^{j\omega})] = V[v(\boldsymbol{n})] + \frac{E^2[v(\boldsymbol{n})]}{N_1 N_2} \left| \sum_{\boldsymbol{n} \in \Phi} e^{-j\boldsymbol{n}^T \omega} \right|^2. \tag{A·9}$$

Appendix B: The Maximum Spectrum at a Frequency of OSESs

In this section, we call the OSESs $R(x(\boldsymbol{n}), \boldsymbol{\omega_a})$ as $v(\boldsymbol{n})$, briefly. The limit value of the maximum shown in Eq. (23) is obtained hereafter. By using additional variable α, the complex maximization problem can be rewritten as follows.

$$\max_{v(\boldsymbol{n}) \in \Psi_{\omega_a}} \left| \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \frac{v(\boldsymbol{n}) e^{-j\boldsymbol{n}^T \omega_a}}{N_1 N_2} \right|$$

$$= \max_{\substack{v(\boldsymbol{n})\in\Psi_{\omega_a} \\ 0\leqq\alpha\leqq 2\pi}} \mathrm{Re}\left[\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\frac{v(\boldsymbol{n})e^{-j(\boldsymbol{n}^T\omega_a+\alpha)}}{N_1N_2}\right] \tag{A·10}$$

$$= \max_{\substack{v(\boldsymbol{n})\in\Psi_{\omega_a} \\ 0\leqq\alpha\leqq 2\pi}} \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\frac{v(\boldsymbol{n})\cos(\boldsymbol{n}^T\omega_a+\alpha)}{N_1N_2}. \tag{A·11}$$

In the case of $\omega_a = \mathbf{o}$, Eq. (A·10) can easily be obtained by assigning $v(\boldsymbol{n}) := \max\{|R_{min}(\omega_a)|, |R_{max}(\omega_a)|\}$. So let $\omega_a \neq \mathbf{o}$ hereafter. Consider the following assignment of $v(\boldsymbol{n})$.

$$v(\boldsymbol{n}) := \begin{cases} R_{max}(\omega_a) & \cos(\boldsymbol{n}^T\omega_a+\alpha) \geqq 0 \\ R_{min}(\omega_a) & \cos(\boldsymbol{n}^T\omega_a+\alpha) < 0 \end{cases}. \tag{A·12}$$

Then Eq. (A·11) is maximum for an arbitrary α and thus has only one variable α. By using Eqs. (A·11) and (A·12), Eq. (23) can be written as

$$S(\omega_a) = R_{dis}(\omega_a)\lim_{N_1,N_2\to\infty}\max_{\alpha} f_{N_1N_2\omega_a}(\alpha) + R_{mid}(\omega_a)\lim_{N_1,N_2\to\infty}\max_{\alpha} g_{N_1N_2\omega_a}(\alpha), \tag{A·13}$$

where

$$f_{N_1N_2\omega_a}(\alpha) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\frac{|\cos(\boldsymbol{n}^T\omega_a+\alpha)|}{N_1N_2} \tag{A·14}$$

and

$$g_{N_1N_2\omega_a}(\alpha) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\frac{\cos(\boldsymbol{n}^T\omega_a+\alpha)}{N_1N_2}. \tag{A·15}$$

Then $g_{N_1N_2\omega_a}(\alpha)$ defined in Eq. (A·15) uniformly converges to

$$\lim_{N_1,N_2\to\infty} g_{N_1N_2\omega_a}(\alpha) = 0. \tag{A·16}$$

From Eq. (A·16), the second term in Eq. (A·13) becomes

$$\lim_{N_1,N_2\to\infty}\max_{\alpha} g_{N_1N_2\omega_a}(\alpha) = \max_{\alpha}\lim_{N_1,N_2\to\infty} g_{N_1N_2\omega_a}(\alpha) \tag{A·17}$$

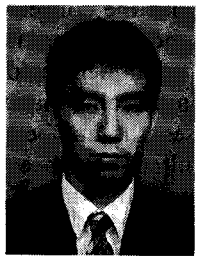
$$= 0. \tag{A·18}$$

By substituting Eq. (A·18) into Eq. (A·13), Eq. (23) can be written as

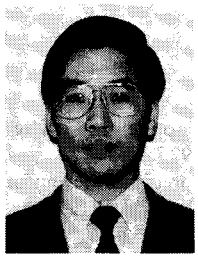
$$S(\omega_a) = R_{dis}(\omega_a)w(\omega_a) \tag{A·19}$$

where

$$w(\omega_a) = \lim_{N_1,N_2\to\infty}\max_{\alpha} f_{N_1N_2\omega_a}(\alpha). \tag{A·20}$$



Mitsuhiko Yagyu was born in Osaka, Japan, on December 3, 1969. He received the B.S. and M.E. degrees from Tokyo Institute of Technology, Tokyo, Japan, in 1993 and 1995 respectively. He is currently working toward the Ph.D. degree in the Graduate School of Science and Engineering, Tokyo Institute of Technology. His main research interest is in digital filter design.



Akinori Nishihara was born in Fukuoka, Japan, on February 26, 1951. He received the B.E., M.E. and Dr. Eng. degrees in electronics from Tokyo Institute of Technology in 1973, 1975 and 1978, respectively. Since 1978 he has been with Tokyo Institute of Technology, where he is now Professor of the Center for Research and Development of Educational Technology. His main research interests are in filter design, 1D and multi-D signal processing, and educational technology. From 1990 to 1994 he served as an Associate Editor of the IEICE Trans. Fundamentals. During 1995–1996 he was Student Activities Committee Chair, IEEE Region 10 (Asia Pacific Region). He is now serving as an Associate Editor of the IEEE Transactions on Circuits and Systems II. Dr. Nishihara is a member of IEEE, EURASIP, ECS and JET.

Nobuo Fujii received B.E. degree from Keio University, Yokohama, Japan, and M.E. and Doctor of Engineering degrees from Tokyo Institute of Technology, Tokyo, Japan, in 1966, 1968, and 1971, respectively. Since 1971, he has been with the Faculty of Engineering, Tokyo Institute of Technology where he is now a professor in the Department of Physical Electronics. From 1984 to 1985, he was a visiting scholar at the University of California, Santa Barbara. From 1990 to 1992, he served as an editor of the Transaction of the Institute of Electronics, Information, and Communication Engineers and is now one of the chief editors of the International Journal of Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers. He is the chairman of the technical group of electronic circuits of IEE Japan and the chairman of the Circuits and Systems Society of IEEE Tokyo Chapter. His main interest lies in the fields of active networks, analog integrated circuits, and analog signal processing. He is the recipient of the Best Paper Award of the Institute of Electrical and Communication Engineers of Japan. He is the author of more than 10 books. Dr. Fujii is a member of the Institute of Electrical and Electronics Engineers, the Institute of Electrical Engineers of Japan.