On The Design of Short Wordlength FIR Filters

Miljan Petrović, *Faculty of Electronic Engineering in Nis*, Branislav Petrović, *Faculty of Electronic Engineering in Nis*

***Abstract*— This paper evaluates the design method of short wordlength FIR (finite impulse response) filters. The design is based on the change of basis of polynomial space, which means that a filter is designed as a weighted parallel structure consisting of FIR filters with only 1s and −1s as coefficients of the transfer function, i.e. only adders/subtractors and unit delay lines in hardware implementation. The structure allowed for a higher minimal attenuation in the stopband in 87.8% of designed filters. Negligible errors in the cutoff frequency and passband attenuation were noticed. Also, this paper broadens the existing design method to high pass, bandpass and bandstop filters. Finally, IIR (infinite impulse response) filters were discussed as a possible generalization of the design procedure.**

***Index Terms*— filter design; filter sensitivity; FIR; short wordlength; stopband attenuation.**

I. Introduction

FILTER design is often constrained by the wordlength of the digital system encompassing the filter. Short wordlength is a necessity in microprocessor and FPGA (field programmable gate array) filter implementations in order to achieve sufficient computational speed, especially in real time application [1]-[2]. Also, low number of bits for number representation can lead to power efficient systems [3]-[5].

Reference [6] presents the realization structure for FIR (finite impulse response) filters, that can be used for short wordlength filter design. It was shown that, comparing to direct form realization, negative quantization effects are reduced, i.e higher minimal attenuation in the stopband is achieved. In other words, the proposed structure exhibits lower filter sensitivity (dependence of the transfer function to its coefficients) [7]-[8]. This paper reexamines low pass filter design from [6]. Further, it introduces the high pass transform matrix and expands the design method to high pass, bandpass and bandstop filters.

The research presented in this paper is structured as follows. First, the design method is summarized. Transform matrices and used polynomial bases are exhibited and discussed. Then, low pass and high pass filters with cutoff frequencies in the range from 0.1π to 0.9π are designed with short wordlength in the direct form and the proposed structure. The consequent comparison shows differences and advantages of the design method in [6]. Quantization errors of the passband and stopband attenuation, and location of the cutoff frequency are assessed.

Miljan Petrović – Faculty of Electronic Engineering in Nis, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia (e-mail: miljan.petrovic@elfak.ni.ac.rs).

Branislav Petrović – Faculty of Electronic Engineering in Nis, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia (e-mail: branislav.petrovic@elfak.ni.ac.rs).

II. Design for Short Wordlength

The inspected design method for short wordlength FIR filters is presented with detailed mathematical explanation in [6]. It is based on designing a FIR filter using a well known method (set of coefficients a_i, $i=0, 1,\ldots, n$), and finding the set of coefficients b_i, $i=0, 1,\ldots, n$ using the formula

$$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix} = M_n^{-1} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \qquad (1)$$

where M_n^{-1} is the inverse matrix of M_n, and M_n consists of columns equal to vectors of coefficients of base polynomials, i.e. base filters. In order to derive a low pass structure for short wordlength design shown in Fig. 1, base filters should be connected in parallel with weight coefficients b_i, $i=0, 1,\ldots, n$. The structure needed for implementation of high pass filters would differ from the structure in Fig. 1. Subtractors should be put instead of certain adders. Finally, quantization is applied to the set of coefficients.

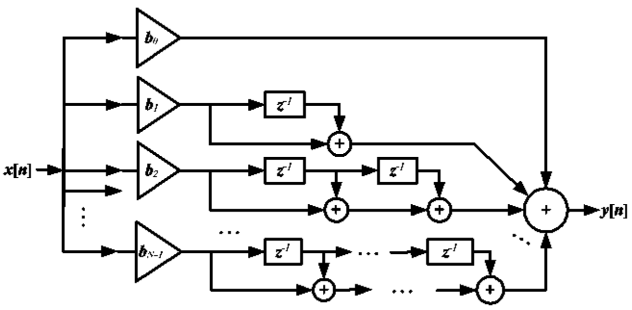

Fig. 1. Filter realization structure based on polynomial basis transform [6].

Due to the fact that transforming the coefficients can lead to a filter which has magnitude response higher than 0 dB in some points, one must perform normalization. Before the quantization step in the design process, coefficients are normalized with the maximal value of the magnitude response. In [6], a different normalization method was implemented (base vectors normalization using Euclidean distance), which produced different results and evaluation of the design procedure. This paper deals with improvement of the normalization step and presents new results.

Reference [6] only explores low pass filter design. The key difference between low pass and high pass filters is the transform matrix M_n and the set of base polynomials.

Proceedings of 4th International Conference on Electrical, Electronics and Computing Engineering, IcETRAN 2017, Kladovo, Serbia, June 05-08, ISBN 978-86-7466-692-0 pp. ELI1.4.1-5

A. *Low pass filters*

Transforming the coefficients while preserving the frequency response of the filter is based on the mathematical expression of *z*-transform of the transfer function $H(z)$:

$$\begin{aligned} H(z) &= a_0 \cdot 1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} + \ldots + a_n \cdot z^{-n} \\ &= b_0 \cdot 1 + b_1 \cdot (1 + z^{-1}) + b_2 \cdot (1 + z^{-1} + z^{-2}) + \\ &\quad + \ldots + b_n \cdot (1 + z^{-1} + z^{-2} + \ldots + z^{-n}), \end{aligned} \tag{2}$$

where n is the order of the filter, and the rest of the notation is as in (1). Equation (2) stands for low pass filters. It is clear that base polynomials, which are being multiplied by coefficients b_i, $i=0, 1,\ldots, n$, are i^{th} order polynomials with all coefficients equal to 1. This defines the matrix M_n, whose columns contain these coefficients. E.g, transform matrix for low pass filter of 3rd order is a 4 by 4 matrix

$$M_3^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

It can be seen that M_n has ones on and above the main diagonal and has zeros below it. Its inverse matrix has ones on the main diagonal, -1 on the diagonal above it, and all other entries equal to 0.

The fact that matrix produces great results only in the case of low pass filters is explained in Fig. 2. Magnitude responses of base polynomials of orders from 1 to 4 are shown. It is clear that all these base filters are low pass. Their linear combination can be a high pass filter too, because the set of polynomials is an orthogonal basis. However, it is this low pass character of the structure that renders it less sensitive to quantization effects only in the case of the original low pass filter design. This means that utilization of low pass base filters for the construction of a low pass filter provides a more feasible design. Similarly, it will be shown that utilization of high pass base filters provides a more feasible design of high pass filters.

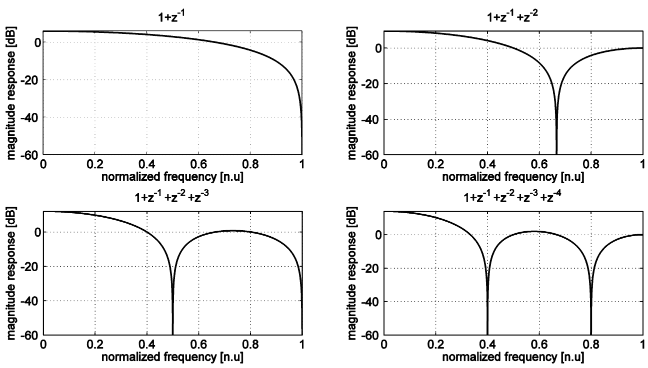

Fig. 2. Magnitude responses of base polynomial filters for low pass filter design.

B. *High pass filters*

Similarly to low pass design, high pass filters are based on the expression:

$$\begin{aligned} H(z) &= a_0 \cdot 1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} + \ldots + a_n \cdot z^{-n} \\ &= b_0 \cdot 1 + b_1 \cdot (1 - z^{-1}) + b_2 \cdot (1 - z^{-1} + z^{-2}) + \\ &\quad + \ldots + b_n \cdot (1 - z^{-1} + z^{-2} - \ldots + (-1)^n \cdot z^{-n}), \end{aligned} \tag{4}$$

where the notation is the same as in (2). The difference between (2) and (4) is in the base polynomials. High pass base polynomials consist of 1s and -1s alternately appearing in the vectors of coefficients. This means that the transform matrix in the case of a 3rd order filter would be

$$M_3^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{5}$$

It is observed that the inverse matrix used for calculating the filter coefficients according to (1), consists of 1s and -1 alternately appearing on the main diagonal, and also on the diagonal above it, starting from 1. All other matrix entries are equal to 0.

Filters designed with this transform matrix exhibit great performance only in the case of high pass filters. This is illustrated in Fig. 3, which shows magnitude responses of base polynomials of orders from 1 to 4. It is obvious that these polynomials are high pass filters, hence, the explanation is as in the case of low pass filters (Fig. 2).

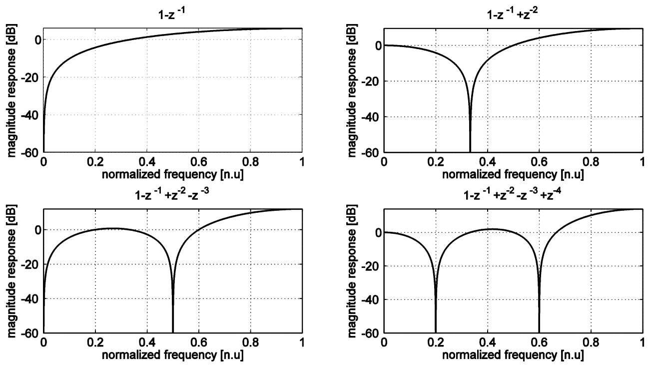

Fig. 3. Magnitude responses of base polynomial filters for high pass filter design.

C. *Banspass and bandstop filters*

Bandpass and bandstop filters can be designed as cascade connection of one low pass and one high pass filter with appropriate cutoff frequencies. Accordingly, the presented design for short wordlength can be applied to these types of filters too.

However, although stopband improvement is expected to be even higher and cutoff frequency error negligible, one must pay attention to the oscillations in the passband. The bandpass (bandstop) filter is considered valid if both incorporated low pass and high pass filters have low enough error. Namely, if the sum of errors of low pass and high pass filter exceeds 3dB (or any other set value of the constraint) in any point, then the passband of the filter is significantly degraded.

III. Examples and Evaluation

Evaluation of the design for short wordlength is performed on low pass and high pass FIR filters with cutoff frequencies set equidistantly in the range from 0.1π to 0.9π with the distance of 0.02π. All filters are of order 54, which is chosen to be high enough to get designed cutoff frequency ω_c close to the desired ω_{des} (frequency set as an argument for the function for Hamming window based FIR design [9]). This can be seen in upper plots in Figs. 8 and 9, by noticing that slopes of functions $\omega_c=f\ (\omega_{des})$ are approximately equal to 1.

Each designed filter is presented in three forms: preliminary design with “infinite” precision wordlength (maximum available computer precision), design with coefficients quantized to format Q1.9 (1 bit for the sign and 9 for the fractional part of the number), and transformed design (as in Fig. 1) with coefficients quantized to Q1.9. Further decrease of the wordlength may lead to rounding all coefficients to zeros in the quantization step, thus completely diminishing the design.

Example of frequency response (magnitude response in upper subplot and phase response in lower subplot) of a low pass filter (ω_{des}=0.32π) is presented in Fig. 4. Similar example for a high pass filter (ω_{des}=0.28π) is shown in Fig. 5. Black solid lines denote responses of infinite precision filters, blue dotted lines responses of quantized filters, and red dashed lines responses of transformed and quantized filters.

Magnitude responses suggest higher minimal attenuation in the stopband, and negligible degradation of the passband attenuation and cutoff frequency due to quantization in the case of transformed design. The phase response remains a linear function in the full range of normalized frequency.

Zoomed subplots show passbands of filters. It can be seen that quantization produces a small error in the passband attenuation. This error is slightly higher in the case of the transformed design. However, for many of the designed filters, it is below 1dB, which can be considered negligible for some applications. Mean values of differences between attenuation in the passband of the infinite precision filter and the others, are presented in Fig. 6 (for low pass filters) and Fig. 7 (for high pass) filters.

It can be seen that the mean error functions for low and high pass filters are symmetric with the line ω_{des}=0.5. This means that the mean degradation of the passband attenuation depends only on the width of the passband. The same error is produced for the low pass filter with the cutoff frequency ω_{des} and the high pass filter with $1-\omega_{des}$. This is due to the analogy present in the window based design approach for low pass and high pass filters [9].

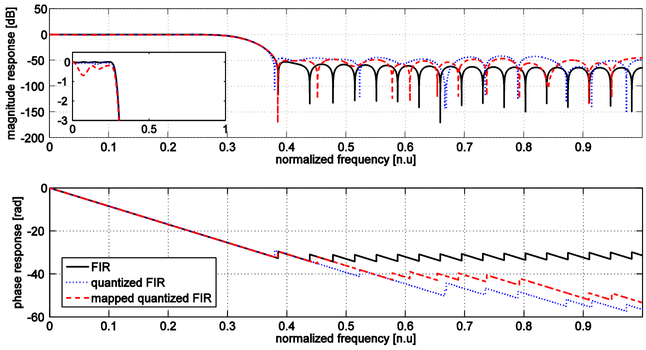

Fig. 4. Magnitude and phase responses of a low pass filter with ω_{des}=0.32π, showing desired characteristics (black solid), direct form realization with quantized coefficients (blue dotted), and transformed realization with quantized coefficients (red dashed).

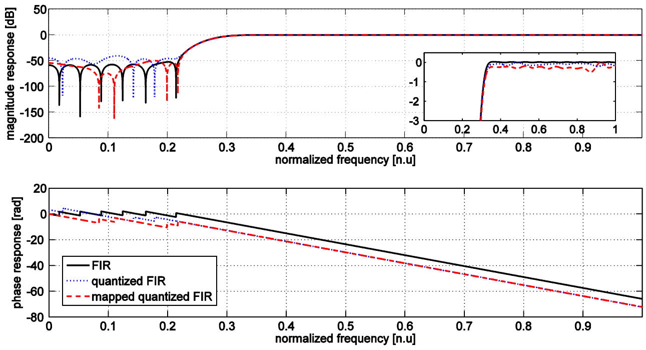

Fig. 5. Magnitude and phase responses of a high pass filter with ω_{des}=0.28π, showing desired characteristics (black solid), direct form realization with quantized coefficients (blue dotted), and transformed realization with quantized coefficients (red dashed).

Further, it is obvious that the functions of mean degradation in Figs. 6 and 7 have higher mean and standard deviation in the case of transformed filter. However, the function of mean degradation is bounded above by 1.4dB.

The main advantage of the design approach presented in [6] is the fact that, in the case of quantization, higher minimal attenuation in the stopband is achieved with the basis transform. This fact is evaluated in Figs. 8 and 9.

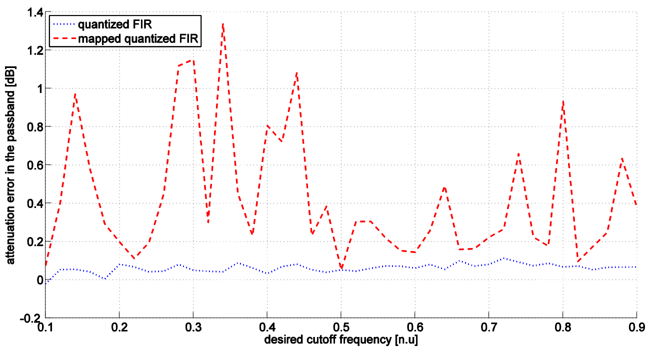

Fig. 6. Mean value of degradation of passband attenuation in low pass filters due to quantization, in the case of direct form realization (blue dotted), and transformed realization (red dashed).

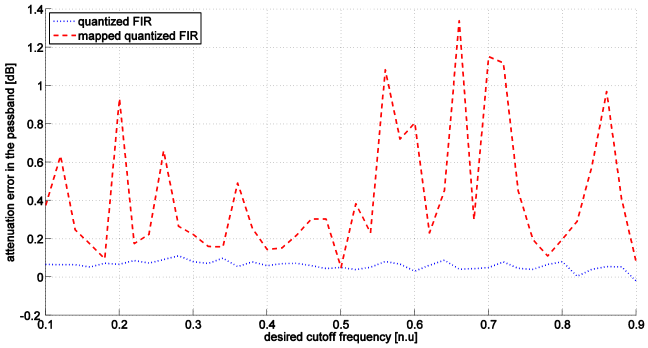

Fig. 7. Mean value of degradation of passband attenuation in high pass filters due to quantization, in the case of direct form realization (blue dotted), and transformed realization (red dashed).

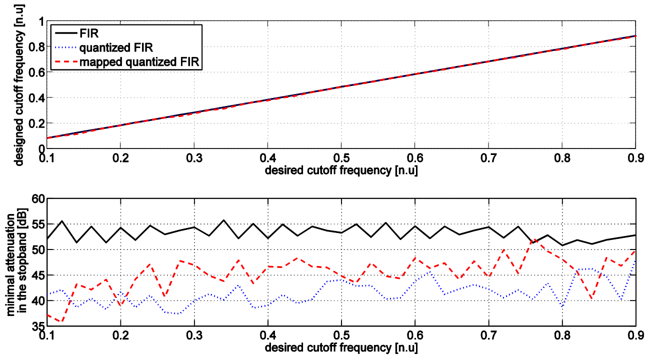

Fig. 8. Designed cutoff frequencies (upper subplot) and minimal attenuations in the stopband (lower subplot) of low pass filters, in the case of desired filter (black solid), direct form realization with quantized coefficients (blue dotted), and transformed realization with quantized coefficients (red dashed).

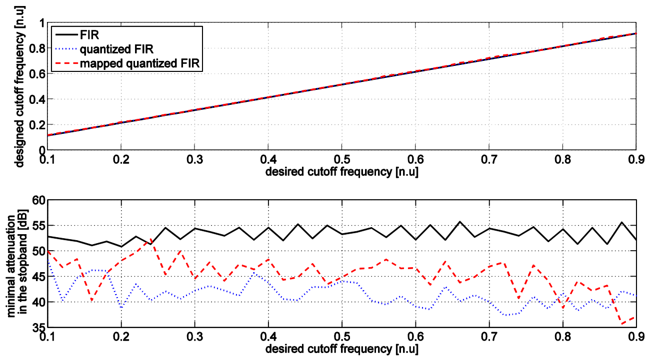

Fig. 9. Designed cutoff frequencies (upper subplot) and minimal attenuations in the stopband (lower subplot) of high pass filters, in the case of desired filter (black solid), direct form realization with quantized coefficients (blue dotted), and transformed realization with quantized coefficients (red dashed).

For each filter and its desired cutoff frequency, minimal attenuation in the stopband (lower subplot) and designed cutoff frequency (upper subplot) are plotted in Figs. 8. and 9. Only several filters show higher minimal attenuation without basis transform, and they all have ω_{des} between 0.1π and 0.2π, or above 0.8π. Better performance is observed for filters with the cutoff frequency in the middle of the explored frequency range if the basis transform is applied. Namely, in the case of 87.8% of all designed filters, higher minimal attenuation in the stopband is observed for the transformed realizations.

Table I shows mean and standard deviation of the design errors in the case of low pass filters. The included errors are the change of the stopband attenuation and the shift of the cutoff frequency.

It can be observed that the mean attenuation error is significantly lower in the case of mapped quantized FIR, which means that minimal attenuation is higher. Further, oscillations and slightly higher maximal attenuations in the passband are compensated with significantly lower mean value of cutoff frequency error (with order of magnitude 10^{-5}).

Since the same design approach has been implemented for both low pass and high pass filters, it was shown that the design errors depend only on the passband width. This means that the mean and standard deviation of the error of the high pass filters are the same as in Table I. The only difference is that the cutoff frequency error has a different sign, suggesting that for low pass filters, quantization leads to a decrease of the cutoff frequency and for high pass filters to an increase. However, for filters of the same passband width, the value of the decrease and the increase are the same, and are negligible.

TABLE I
ERRORS IN FILTER DESIGN (MEAN ± STANDARD DEVIATION)

	Quantized FIR	Mapped quantized FIR
Stopband attenuation error [dB]	11.6826 ± 2.907	7.8103 ± 3.9416
Cutoff frequency error [n.u]	$0.4280\cdot10^{-3} \pm 0.3592\cdot10^{-3}$	$0.003\cdot10^{-3} \pm 0.0029\cdot10^{-3}$

IV. CONCLUSION AND FUTURE WORK

Filter sensitivity and quantization effects are being examined for years. Both solutions in the form of filter structures [10]-[11] and numerical optimizations [12]-[13] have been presented. This paper has provided insight into why the design procedure [6] exhibits great performance. It showed that the design produces higher stopband attenuation with insignificant errors in the passband for a wide range of cutoff frequencies. However, only Hamming window based FIR filters have been examined. Other design methods as the first step of the design for short wordlength should be analyzed in further research.

Infinite impulse response (IIR) filters have also been explored in the context of short wordlength [14]. Analogous to the design procedure in [6], IIR filters may be designed as weighted parallel connections of base transfer functions, whose numerator and denominator polynomials contain only 1s and –1s in order to avoid an increase of the number of multipliers. According to the analysis presented in this paper, those transfer function should be either low pass or high pass filters. Further, calculating the final set of coefficients cannot be a simple matrix-vector multiplication. A more complex computational procedure must be implemented, taking in consideration the right set of base transfer functions.

REFERENCES

[1] P. D. Fiore, L. Lee, "Closed-form and real-time wordlength adaptation," *IEEE Int. Conf. on Acoustics, Speech, and Signal Process.*, Mar., 1999.

[2] G. A. Constantinides, "Perturbation analysis for word-length optimization," 11th Annual *IEEE Symp. on Field-programmable Custom Comp. Mach.*, Apr., 2003.

[3] K. Han, B. L. Evans, E. E. Swartzlander, "Data wordlength reduction for low-power signal processing software," *IEEE Workshop on Signal Process. Syst.*, pp. *343-348*, Oct., 2004.

[4] K. Han, B. L. Evans, "Wordlength optimization with complexity-and-distortion measure and its application to broadband wireless demodulator design," *Proceed. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Process.*, pp. *37-40*, May, 2004.

[5] J. T. Ludwig, S. H. Nawab, A. P. Chandrakasan, "Low-power digital filtering using approximate processing," *IEEE J. of Solid-State Circ.*, vol. *31*, no. *3*, pp. *395-400*, Mar., 1996.

[6] M. Petrovic, B. Petrovic, "A network for low-pass FIR filters realization with short wordlength," *Proceedings of 3rd International Conference on Electrical, Electronic and Computing Engineering IcETRAN*, Zlatibor, Serbia, pp. ELI1.5.1-5, June, 2016.

[7] T. Saramaki, T.–H. Yu, S. Mitra, "Very low sensitivity realization of IIR digital filters using a cascade of complex all-pass structures," *IEEE Trans. Circ. Syst.*, vol. *34*, no. *8*, pp. *876-886*, Jan., 1987.

[8] P. P. Vaidyanathan, S. K. Mitra, "Very Low Sensitivity FIR Filter Implementation Using "Structural Passivity" Concept," *IEEE Trans. Circ. Syst.*, vol. *CAS-32*, no. *4*, pp. *360-364*, Apr., 1985.

[9] Digital Signal Processing Committee of the IEEE Acoustics, Speech, and Signal Processing Society, eds., *Programs for Digital Signal Processing*, New York, USA: IEEE Press, 1979.

[10] T. Yoshikawa, Y. Watanabe, A. Khare, Y. Wang, "Low sensitivity realization of FIR filters with 2-stage cascade structure," *Proceed. of the IEEE Int. Symp. on Industrial Electron.*, pp. *454-455*, May, 1992

[11] C. A. Corral, C. S. Lindquist, "Sensitivity of classical filter transfer functions," *Proceed. of the 44th IEEE Midwest Symp. on Circ. And Syst.*, pp. *45-48*, Aug., 2001.

[12] F. Brglez, "Digital filter design with short word-length coefficients," *IEEE Trans. Circ. Syst.*, vol. *25*, no. *12*, pp. *1044-1050*, Dec., 1978.

[13] D. Kodek, "An algorithm for the design of optimal finite word-length FIR digital filters," *IEEE Int. Conf. on Acoustics, Speech, and Signal Process.*, Apr., 1980.

[14] K. Hon-Keung, "Erratum: Method for designing recursive digital filters with short wordlength," *Electronic Letters*, vol. *15*, no. *12*, pp. *361*, June, 1979.

[15] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, *1*st ed. Philadelphia, USA: SIAM: Society for Industrial and Applied Mathematics, 2001.

[16] J. G. Proakis, D. G. Manolakis, *Introduction to Digital Signal Processing*, *1*st ed. New York, USA: Macmillan Publish. Co., 1988.