FINITE WORD-LENGTH EFFECTS IN DIGITAL STATE-SPACE FILTERS

Vratislav DAVÍDEK[1], Marcela ANTOŠOVÁ[1], Bohumil PŠENIČKA[2]
[1] Dept. of Circuit Theory
Czech Technical University in Prague
Technická 2, 16627 Praha 6, Czech Republic
[2] Dept. of Telecommunication
Universidad National Autonoma de Mexico

Abstract

The state-space description of digital filters involves except the relationship between input and output signals an additional set of state variables. The state-space structures of digital filters have many positive properties compared with direct canonical structures. The main advantage of digital filter structures developed using state-space technique is a smaller sensitivity to quantization effects by fixed-point implementation. In our presentation, the emphasis is on the analysis of coefficient quantization and on existence of zero-input limit cycles in state-space digital filters. The comparison with direct form II structure is presented.

Keywords

digital filter, state-space filter, quantization, limit cycle, finite word-length

1. Introduction

A lower sensitivity to roundoff effects of state-space digital filters (SDF) by fixed-point implementation is a main reason for using special filter structures. In this paper there are analysed sensitivity to coefficient quantization and zero-input limit cycles in the second order digital state-space filter in comparison to the second order direct canonical structure. The existence of limit cycles in recursive digital filters is the object of interest of some publications as [1] - [3], [5], [6] etc. For the limit cycles existence play a deciding role the feed-back coefficients of a system matrix **A**, similarly as denominator coefficients of a system function in the case of canonical direct form II (CDF). The zero-input limit cycles existing in absolute stable digital filter due to an accumulation of quantization errors by multiplying in recursive section of a filter.

In contribution there will be obtained relationships between coefficients of second-order state and canonical structures for the same frequency response. This relationships give a possibility to transform both structures mutually and to compare their roundoff effects.

2. Digital State-Space Filters

The N-th order single-input/single-output SDF can be described by two equations, by the state equation and by the output equation:

$$\begin{aligned} \mathbf{u}[n+1] &= \mathbf{A}\mathbf{u}[n] + \mathbf{B}x[n] \\ y[n] &= \mathbf{C}\mathbf{u}[n] + \mathbf{D}x[n] \end{aligned} \tag{1}$$

where $\mathbf{u}[n]$ is the N-dimensional vector of state variables, $x[n]$ is an input sequence, $y[n]$ is an output sequence and the state-matrixes **A**, **B**, **C** and **D** contain the filter coefficients. The matrix **A** represents a system matrix of the dimension N x N. The associated system function H(z) is given by

$$H(z) = \mathbf{D} + \mathbf{C}\cdot(z\cdot\mathbf{I} - \mathbf{A})^{-1}\cdot\mathbf{B} \tag{2}$$

where **I** is the identity matrix.

2.1 Second order state-space structure

The recursive second-order SDF can be described by the state matrixes following from the Eq. (1)

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},\ \mathbf{C} = \begin{bmatrix} c_1 & c_2 \end{bmatrix},\ \mathbf{D} = d \tag{3}$$

The equations (1) for the second-order SDF can be written in the form

$$\begin{bmatrix} u_1[n+1] \\ u_2[n+2] \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1[n] \\ u_2[n] \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} . x[n]$$

$$y(n) = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \cdot \begin{bmatrix} u_1[n] \\ u_2[n] \end{bmatrix} + d.x[n]. \tag{4}$$

The second-order SDF structure derived from the state-equations (4) is shown in Fig. 1a). The common system function of the second-order CDF is given by

$$H(z) = \frac{B_0 + B_1 z^{-1} + B_2 z^{-2}}{1 + A_1 z^{-1} + A_2 z^{-2}} \tag{5}$$

or equivalently

$$H(z) = B_0 + \frac{(B_1 - B_0 A_1) z^{-1} + (B_1 - B_0 A_2) z^{-2}}{1 + A_1 z^{-1} + A_2 z^{-2}} \qquad (6)$$

The CDF structure is shown in Fig. 1 b).

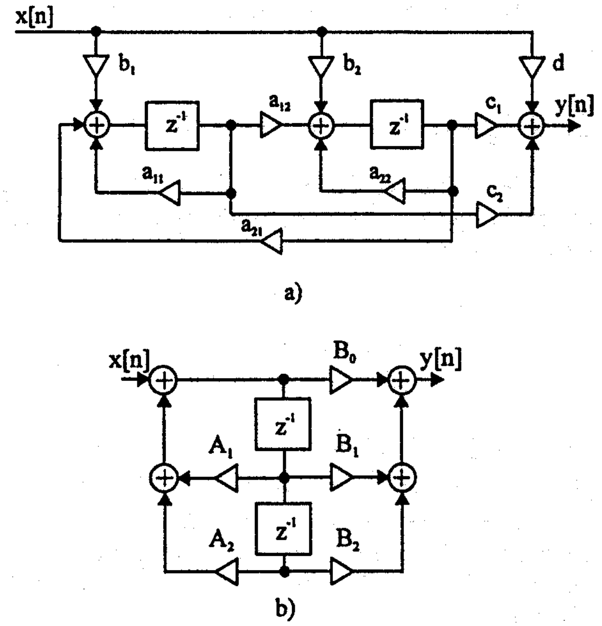


Figure 1: Second-order digital filters a) state-space filter b) direct form II.

Substituting matrices (3) in the Eq. (2) can imply the system function H(z) of SDF in the form

$$H(z) = d + \frac{\alpha_1 z^{-1} + \alpha_2 z^{-2}}{1 + \beta_1 z^{-1} + \beta_2 z^{-2}} \qquad (7)$$

where constants α and β are expressed as

$$\begin{aligned} \alpha_1 &= b_1 c_1 + b_2 c_2 \\ \alpha_2 &= b_1 c_2 a_{21} + b_2 c_1 a_{12} - b_1 c_1 a_{22} - b_2 c_2 a_{11} \qquad (8) \\ \beta_1 &= -(a_{11} + a_{22}) = -tr\,\mathbf{A} \\ \beta_2 &= (a_{11} a_{22} - a_{12} a_{21}) = \det \mathbf{A} \end{aligned}$$

Symbols $tr\,\mathbf{A}$ and $\det \mathbf{A}$ denote the trace and the determinant of a system matrix $\mathbf{A}$, respectively.

Comparing H(z) in equations (5) and (7) we get five equations for the computation of nine state-space filter coefficients. The next necessary equations follow for example from the relations for sensitivities of zeros and poles of transfer function to the filter coefficients, as it was derived in [4] and [8]. As equations (2) and (6) are expected to be equal the following equations have to hold for the system function of the second order

$$\begin{aligned} B_0 &= d \\ B_1 &= b_1 c_1 + b_2 c_2 - \det \mathbf{A} \\ B_2 &= d \cdot \det \mathbf{A} + b_1 c_2 a_{21} + b_2 c_1 a_{12} - b_1 c_1 a_{22} - b_2 c_2 a_{11} \qquad (9) \\ A_1 &= -tr\,\mathbf{A} \\ A_2 &= \det \mathbf{A}\,. \end{aligned}$$

The obtained relations between coefficients of state-space filter (7) and coefficients of direct canonical structure (5) are given by the following equations:

$$a_{11} = a_{22} = -\frac{A_1}{2}, \qquad a_{12} = -a_{21} = \sqrt{A_2 - \frac{A_1^2}{4}}$$

$$c_2 = -b_2 = \sqrt{\frac{\sqrt{\alpha^2 - 4\varepsilon} - \alpha}{2}}, \qquad (10)$$

$$c_1 = b_1 = -\sqrt{\alpha + c_2^2}\,, \qquad d = B_0\,,$$

where

$$\alpha = B_1 - B_0 A_1, \quad \beta = B_2 - B_0 A_2, \quad \varepsilon = \frac{(\beta + \alpha \cdot a_{22})^2}{4 a_{12}^2}.$$

3. Finite word-length effects

Digital filter development methods it is assume that the filter coefficients and signal variables have infinite precision. However, when implemented in special purpose hardware form, the system parameters and signal variables can take only limited word length. This discrete values are specified by the length of registers provided to store the coefficients and signal values. Quantization process in digital filters results in various sources of errors. Two of this sources will be mentioned in next paragraphs, coefficient quantization and quantization of arithmetic operations.

3.1 Quantization of filter coefficients

When implemented on a digital hardware, the filter coefficients can assume only certain discrete values. The poles and zeros of the system function will, in general, be different from the desired poles and zeros. Therefore, the actual frequency response may be quite different from the desired frequency response.

The possibility of realisation of band-pass IIR digital filter by a cascade structure of second-order SDF giving a lower sensitivity of the transfer function to coefficient quantization compared with the second-order CDF structure, will be now demonstrated. Fig. 2 presents magnitude responses of an elliptic band-pass four-order filters for six-bit coefficient length. It is obvious from Fig. 2, that the frequency response of SDF (dotted line) much better approximates the full precision response (solid line) than CDF frequency response (dached line).

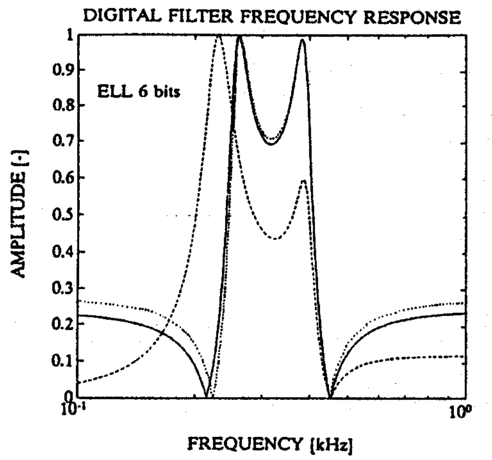


Fig. 2: Frequency responses of four-order band pass DF's

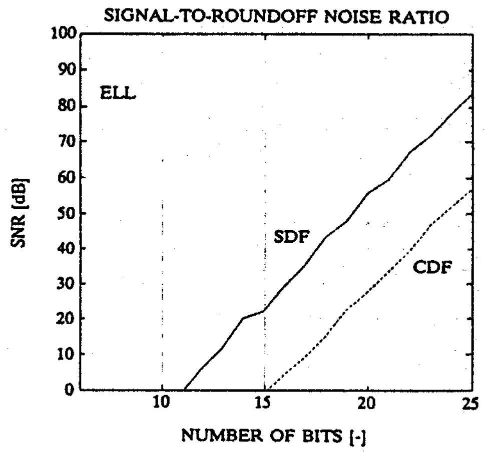


Fig. 3. SNR for different word length for SDF and CDF structures.

The designed DF's have been simulated for some particular coefficient length and for fixed-point arithmetic using a computer simulation program. As a further measure of the finite word-length effects the SNR in [dB] has been used in the following formula

$$\mathrm{SNR} = 10\log\frac{\sum_{n=0}^{M} y^2[n]}{\sum_{n=0}^{M}\left(y_Q[n]-y[n]\right)^2} \qquad (11)$$

where $y[n]$ is the output sequence of the analysed DF with particular coefficient length from 6 do 25 bits and $y_Q[n]$ is the quantized output sequence simulated in the fixed-point arithmetic using rounding and two's complement representation.

3.2 Zero-input Limit Cycles Analysis

In recursive digital filters, the finite precision of fixed-point arithmetic (e.g., quantization of the output of multiplication and/or additions) causes limit cycle oscillations at the filter output. The period and the maximum amplitude of limit cycle depends on the filter coefficients. A second order recursive filter section can generate a variety of limit cycles if rounding arithmetic is used at the multiplier outputs.

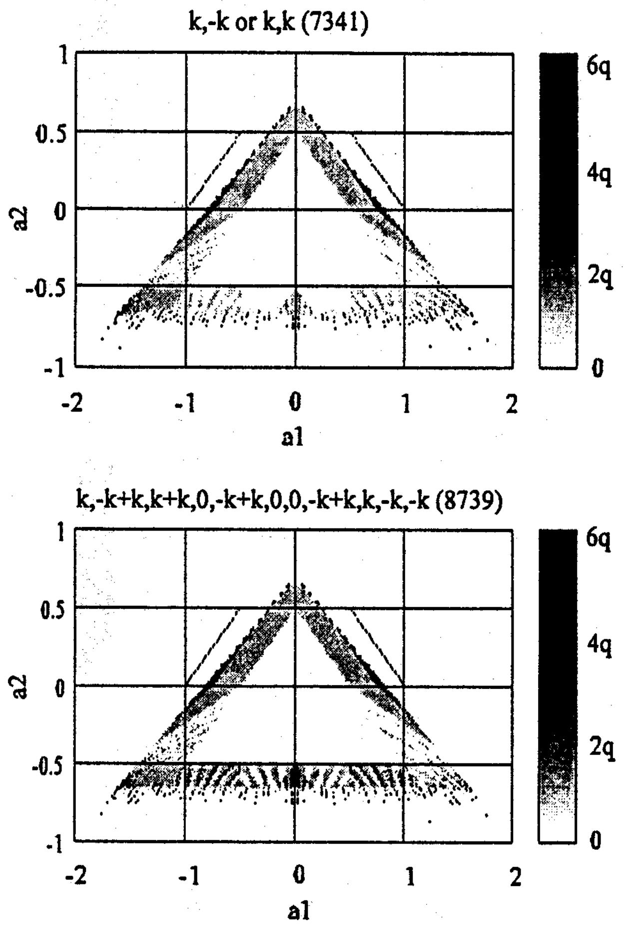
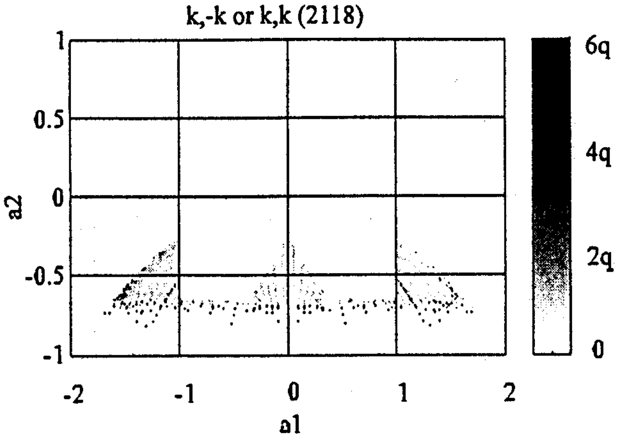


Fig. 4 Area of limit cycles existence in direct form II DF.

It is known that the linear version of direct form realisation of a second order digital filter is asymptotically stable when the coefficients A_1 and A_2 lie inside the triangular spaced in region $|A_2| < 1$ and $|A_1| < 1+A_2$. Let quantize booth of coefficients with the step of 0.01. At the output of CDF structure there are to observe limit cycles with different amplitude and period. In Fig. 4 is shown the region of quantized coefficients for which exist limit cycles with period length $L=1$ and $L=2$ with amplitude of K. The amplitude of limit cycles K is given in multiples of the quantization step q. The period $L=1$ corresponds to output sequence {K,K} and $L=2$ *to* sequence {K,-K} respectively. In Figures 4 and 5 are illustrated the founded amplitudes of limit cycles. For combinations of filter coefficients $a1$ and $a2$ is the amplitude of existing limit cycles represented by the scale of tints between white and black colours. Darker tints of grey describe an existence of LC with a larger amplitude. The results of LC analysis is shown in Fig. 4 and 5. It was found about 7341of zero-input limit cycles of period K,K or K,-K for CDF and only 2118 limit cycles K,K or K,-K for SDF. Another founded LC are presented in next parts of Fig. 4 and 5.



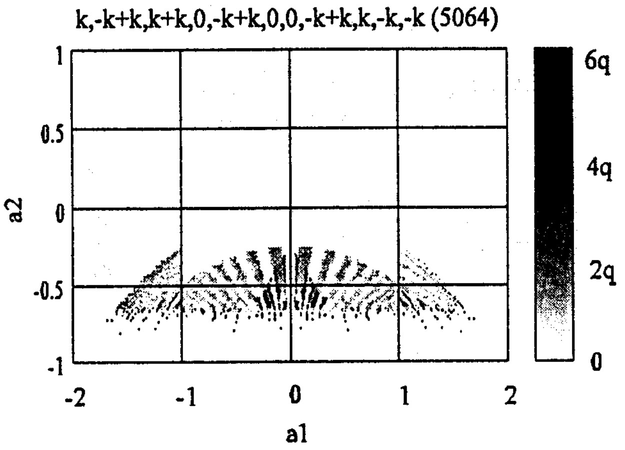


Fig. 5 Area of limit cycles existence in state-space DF.

4. Conclusion

The state-space DF's are special structures of digital filters with lower sensitivity to roundoff effects by fixed-point implementation in comparison to canonical direct form II. Based on the analysis shown in Figures 2, 4 and 5 we can make several comments. The SDF's have generally a lower sensitivity to coefficient quantization in comparison to CDF's. The probability of occurrence of zero-input limit cycles by rounding is lower as in the case of direct canonical form. Unfortunately, if the limit cycle occurs its amplitude can be higher (two times maximally), compared to direct form realisation. Digital filters with minimum norm of system matrix **A** have asymptotically stable realisation for overflow oscillations and for magnitude truncation limit cycles. The disadvantage of SDF's is a higher number of multipliers, 9 coefficients are necessary against 5 coefficients of CDF structures.

References

[1] JACKSON, L. B.-HILL, N. M.: Roundoff-noise Analysis for Fixed-point Digital Filtres Realized in Cascade or Parallel Form, IEEE Trans. Audio Electroacoust.}, vol. AU-18, June 1970.

[2] LIU, B.: Effect of finite word lenght on the accuracy of digital filters - a review, IEEE Trans. Circuit Theory, vol. CT-18, Nov. 1971.

[3] BARNES, C. W.-FAM, A. T. : Minimum norm recursive digital filters that are free of limit cycles, IEEE Trans. on Circuits and Systems, vol. CAS-24, no. 10, October 1977.

[4] PŠENIČKA, B.-ZADÁK, J. -DAVÍDEK, V. : Design of State Digital Filters, Acta Polytechnica 14(III, 3), 1991.

[5] LECLERC, J. L.-BAUER, H. P. : New Criteria for Absence of Limit Cycles in State-Space Digital Filter with Combinations of Quantization and Overflow Nonlinearities, IEEE International Symposium on Circuits and Systems, San Diego, 1992.

[6] BOSE, T.-BROWN, D. P. : Limit cycles due to roundoff in state-space digital filters, IEEE Trans. On Acoustic, Speech and Signal Processing, vol. 38, no. 8, August 1990

[7] UGALDE, F.G. - PŠENIČKA, B. - DAVÍDEK V. : On the Design of the State Digital Filters, ICSPAT'95 Boston, October 1995.

[8] PŠENIČKA, B.-UGALDE, F. G.-SAVAGE J.-DAVÍDEK V.: Design of State Digital Filters, IEEE Trans. On Signal Processing, vol. 46, no. 9, September 1998

[9] TIHELKA, J.: Quantization Effects Analysis in IIR Digital Filters, (in Czech) Student project, Faculty of Electrical Engineering of CTU, 1997

About authors

Vratislav DAVÍDEK was born in Prague in 1942. He received (M.Sc.) equiv. degree in control technique in 1964 and (Ph.D.) equiv. degree in Radio Electronics in 1977 at the TU Prague. Since 1993 he is an Associated Professor, currently he is the head of the Department of Circuit Theory at the Faculty of Electrical Engineering of TU in Prague.

Marcela ANTOŠOVÁ was born in Prague in 1965. She received (M.Sc.) and (Ph.D.) equivalent degrees both in Telecommunication in 1988 and 1992 respectively at the TU Prague. Since 1988 she is an Assistant Professor at the Department of Circuit Theory at the Faculty of Electrical Engineering of TU in Prague.

Bohumil PŠENIČKA was born in Prague in 1933. He received (M.Sc.) and (Ph.D.) equivalent degrees both in Telecommunication in 1968 and 1972 respectively at the TU Prague. In 1969-1990 he was an Assistant Professor and in 1990-1993 he vas an Associated Professor at the Department of Circuit Theory at the Faculty of Electrical Engineering of TU in Prague. Currently he is a Professor at the University in Mexico City, Department of Telecommunication.