DSPA, 1999, V.1, pp.126-129.

DESIGN OF CASCADE IIR DIGITAL FILTERS USING MINIMUM ADDER MULTIPLIER BLOCKS

Mingazin A.T.
RADIS Ltd
Radio, 12/2, 107005, Moscow, Russia, Tel. 536-83-73, Fax. 267-45-39, e-mail: alexmin@orc.ru

Abstract. For further complexity reduction of fixed-point cascade IIR digital filters using multiplier blocks the minimizing problems of coefficient wordlength, (roundoff noise)-to-signal ratio and adder cost of multiplier blocks should be decided jointly. Two approaches to such decision are proposed. Their efficiency is demonstrated by examples.

1. Introduction

Complexity of fixed-point digital filters has been significantly reduced by applying a multiplier block technique [1,2]. The block contains shift-add elements. Recently, Dempster and Macleod [3] have compared cascade, parallel, direct and wave IIR digital filters using minimum adder multiplier blocks. Their investigation and obtained average results show that the cascade structure is the most efficient structure. Here, we focus on the more detailed design of low-complexity cascade IIR digital filters using multiplier blocks.

For given filter specifications, assuming that the approximation, order, structure of the sections and L_p-norm for scaling are chosen, the minimum number of adders in the multiplier blocks depends on quantized coefficient values. The last, in turn, depend on the discrete solution variant, coefficient wordlength, type of scaling factors, insert method of these factors in the filter transfer function, pole-zero pairing (e.g. for the elliptic approximation) and ordering of the filter sections. On the other hand, all this has an effect on the (roundoff noise)-to-signal ratio, which needs also to be decreased for reduction of the date wordlength inside filter. Thus, the low-complexity design is a difficult problem. For cascade filters using individual multipliers a joint minimizing of the coefficient wordlength and noise-to-signal ratio was discussed in [4]. In this letter two approaches to the joint minimizing of the coefficient wordlength, noise-to-signal ratio and total number of the adders in cascade IIR digital filters using multiplier blocks are proposed. As will be shown on examples such design results to further complexity reduction.

2. Statement of problem

We shall consider L_p-scaled cascade filters composed of second-order transposed direct form I sections. The scaling factors can be of two types, namely equal and not equal to powers of two. It is assumed that they are inserted by change of transfer function numerator coefficients (except the input scaling factor). The scaled transfer function with zeros on the unit circle in z-domain can therefore be written as

$$H(z) = B_{00}\prod_{i=1}^{K}\frac{B_{0i} + B_{1i}z^{-1} + B_{0i}z^{-2}}{1 + A_{1i}z^{-1} + A_{2i}z^{-2}} = B_{00}\prod_{i=1}^{K}H_i(z).$$

The magnitude function |H(z)| with its quantized coefficients should satisfy to given tolerance specifications. We shall define the quantization step as $q=2^{-M}$, where M is the mantissa wordlength of the transfer function coefficients. Reduction of M results in decrease of the minimum number of adders in filter multiplier blocks [2,3].

For our roundoff noise model the noise-to-signal ratio on the filter output (in dB) is

$$N/S = 10\lg\{ \frac{1}{1.5G^2}[1 + \sum_{i=1}^{K}\left\|\prod_{n=i}^{K}H_n\right\|_2^2] \} - 6.02b = R - 6.02b,$$

where G is the filter gain, b is the number of bits (including sign bit) needed to be kept after rounding of date inside the filter.

For given N/S, the reduction of a value R on 6 dB reduces a value of b on 1 bit and in this case the filter complexity is decreased. We shall use the parameter R as the N/S-performance.
The total number of adders in the filter multiplier blocks is

$$\Sigma = \sum_{i=0}^{K} m_i ,$$

where m_i - the number of adders in the i-th block.
The block with m_0 executes one multiplication on the B_{00}. We shall use Σ as the adder cost of multiplier blocks.
Our design aim consists in reduction as much as possible the values of the parameters M, R and Σ at satisfaction of given tolerance specifications.

3. Approaches to problem decision

We shall consider two possible approaches to decision of the statement problem. The first corresponds to the scaling factors equal to powers of two and consists in subsequent minimizing of the coefficient wordlength, adder cost with account of pole-zero pairing, and N/S-performance using ordering of the sections. Clearly, the adder cost minimizing can be executed for a set of tolerable discrete solutions. The variant with the smaller minimum cost should be chosen in this case. The second corresponds to the scaling factors not equal priori to powers of two. In this approach (in our statement task) the quantization problems for both numerator coefficients and scaling factors become undivided [4]. At the beginning a simultaneous minimizing of the coefficient wordlength and N/S-performance using pairing-ordering is executed. Then, the smaller minimum adder cost variant from a set of solutions with acceptable N/S-performances is chosen.
Obviously, for filters with specific zeros (e.g. for Chebyshev or Butterworth filters with multiple zero in the point z = -1 or +1) the pairing should be eliminated from two above approaches and the second of them is actually reduced to the first. The use of these approaches does not exclude a possibility of some tradeoff between the parameters M, R and Σ.

4. Examples

We shall continue our discussion on the design of two L_∞-scaled digital filters, derived from elliptic analog prototypes by the bilinear transformation. The tolerance specifications to them were also used by other authors in relation to the coefficient wordlength reduction [4]. In the design we apply methods from [2,4,5].

4.1Example 1

The low-pass filter specifications are the passband ripple $\Delta a \leq 0.174$ dB, the stopband attenuation $a_0 \geq 60$ dB, the edge frequencies $f_1 = 0.166667$ and $f_2 = 0.188056$. Here and further f_j is normalized in relation to a sampling frequency. The number of sections K=5.
For this filter we shall demonstrate the efficiency of the first proposed approach. In this case the scaling factors (B_{0i}, i=0-5) equal to powers of two. The integer coefficient version of the minimum wordlength solution, obtained by a method based on variation of initial parameters [4], is

$$(35,-11,53)(32,-17,6)(30,-24,-14)(29,-29,-22)(28,-31,-24).$$

This configuration, in which the sequence of ( ) is ($-A_{1i}$,$-A_{2i}$,B_{1i}/B_{0i}), i=1-5, gives the information on the pole-zero pairing and ordering of the sections. The quantization step of coefficients $q=2^{-5}$ and M=5. The real coefficients can be reproduced by multiplication of the integer values by q. Notice, the solution remains tolerable, when B_{11}/B_{01}= 54 or 55, i.e. there are three discrete solution variants. For this configuration the application of the minimum adder cost technique [2] and computation of the N/S-performance give

$$\Sigma = 4+2+3+3+3 = 15 \text{ and } R=18.7 \text{ dB}.$$

The other values B_{11}/B_{01} do not change Σ and effect to R very weakly. The presented configuration corresponds to the one accepted in attention in a heuristic pairing-ordering procedure [5]. The best solution, obtained by using this procedure, has R=17.4 dB. Change of the pole-zero pairing in the above configuration and the application of the technique [2] leads to other minimum adder cost solution. It is

$$(35,-11,-22)(32,-17,-14)(30,-24,-24)(29,-29,6)(28,-31,53),$$
$$\Sigma = 3+2+2+2+4=13, \ R=27.4\text{dB}.$$

The alternate pole-zero pairing, exchange of 53 on 54, technique [2] and ordering of the sections [5] result in following

$$(35,-11,-22)(29,-29,6)(30,-24,54)(32,-17,-24)(28,-31,-14),$$
$$\Sigma = 3+2+3+2+2 = 12, \ R=17.7 \text{ dB}.$$

In this case $B_{00} = 8$, $B_{01} = 16$, $B_{02} = 8$, $B_{03} = 4$, $B_{04} = 16$, $B_{05} = 64$. For this filter the selection of the pole-zero pairing and discrete solution variant results in the adder cost reduction on 20% without deterioration of the N/S-performance. It is interesting, the simplified approach to the design, namely the simple rounding of minimax filter coefficients , their canonic signed-digit representation and the first from above configurations, results in M=9, Σ=10+8+7+8+8=41, R=20.1 dB. Thus, the presented approach reduces the adder cost on 71% and somewhat improves the N/S-performance in comparison with the simplified approach.

4.2 Example 2

The band-pass filter specifications are $\Delta a \leq 0.521$ dB, $a_0 \geq 40$ dB, $f_1 = 0.191667$, $f_2 = 0.205556$, $f_3 = 0.233333$, $f_4 = 0.247222$. The value K=4.

The first design approach (the powers-of-two scaling factors) gives

$$(18,-58,-64)(36,-62,0)(12,-62,-48)(29,-58,24),$$
$$M=6, \ \Sigma = 3+2+2+2 = 9, \ R=18.0 \text{ dB}.$$

In this case $B_{00} = 4$, $B_{01} = 16$, $B_{02} = B_{03} = 32$, $B_{04} = 128$. The use of discrete solution variants and pole-zero pairing has allowed to reduce the minimum adder cost with 12 up to 9 or on 25%.

The second approach (the scaling factors not equal to powers of two in advance) gives a number of the solutions. Two of them are

$$(6)(18,-58,34,-25)(36,-62,19,0)(12,-62,16,7)(29,-58,119,-130),$$
$$M=6, \ \Sigma = 1+4+3+3+4=15, \ R=15.3 \text{ dB};$$

$$(2)(36,-62,17,8)(18,-58,14,-16)(29,-58,40,0)(12,-62,353,-258),$$
$$M=6, \ \Sigma = 0+3+3+3+4 = 13, \ R=23.9 \text{ dB}.$$

Here the sequence of ( ) is (B_{00}), ($-A_{1i}$,$-A_{2i}$,B_{0i},B_{1i}), i=1-4. The first solution corresponds to the minimum of R and second - to the minimum of Σ. The other solutions have intermediate or large values of these parameters. For the obtaining of these results we used the much more number of possible pole-zero pairs than in [5]. Clearly, we can not apply the pole-zero pairing in two presented configurations for additional reduction of Σ, because it will distort the L_∞-scaling. Notice, in each of three obtained solutions one structural adder is eliminated (e.g. the third configuration where B_{13}=0). For this filter the advantage of the first approach over the second consists in reduction of the adder cost by 40% and 30% depending on the presented variants at comparable the N/S-performance.

5. Conclusions

The design of low-complexity fixed-point cascade IIR digital filters using multipler blocks requires to decide three problems: minimizing of the coefficient wordlength, (roundoff noise)-to-signal ratio and adder cost of multiplier blocks. We have proposed two approaches to the joint decision of these problems with account of the type of scaling factors, discrete solution variants, pole-zero pairing and ordering of the sections. Our study shows that such decision results in further filter complexity reduction. The considered approaches can be applied in development of VLSI digital filters and CAD tools for their design.

References

1. Bull D.R. and Horrocks D.H. Primitive operator digital filters. Proc. IEE, pt. G, 1991, v.138, june, pp. 401- 412.
2. Dempster A.G. and Macleod M.D. Use of minimum-adder multiplier block in FIR digital filters. IEEE Trans. on Circuits and Systems-II, 1995, v. 42, N 9, pp. 569-577.

3. Dempster A.G. and Macleod M.D. IIR digital filter design using minimum adder multiplier blocks. IEEE Trans. on Circuits and Systems-II, 1998, v. 45, N 6, pp. 761-763.
4. Mingazin A.T. Synthesis of digital filter transfer functions in discrete coefficient space (a review). Elektronnaya Tekhnika, Ser.10, 1993, N 1,2, pp. 3-35, (rus.).
5. Mingazin A.T. and Zorich A.A. Minimization of roundoff noise in cascade recursive digital filters. Elektronnaya Tekhnika, Ser.10, 1992, N 1,2, pp. 37-43, (rus.).