Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'98), vol. 3, pp. 1253–1256, Seattle, Washington, May 12–15, 1998

ENERGY-BASED EFFECTIVE LENGTH OF THE IMPULSE RESPONSE OF A RECURSIVE FILTER

Timo I. Laakso[1] and Vesa Välimäki[2]

[1]Helsinki University of Technology, Laboratory of Telecommunications Technology
P.O. Box 3000, FIN-02015 HUT, Finland
E-mail: Timo.Laakso@hut.fi

[2]Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing
P.O. Box 3000, FIN-02015 HUT, Finland
E-mail: Vesa.Valimaki@hut.fi, URL: http://www.acoustics.hut.fi/~vpv/

ABSTRACT

A measure for the effective length of the impulse response of a stable recursive digital filter based on accumulated energy is proposed. A general definition and a simple algorithm for its evaluation are introduced, and closed-form expressions are derived for first-order IIR filters. The effect of zeros on the effective length is analyzed. An upper bound for the effective length of higher-order filters is derived using results for low-order filters. The new measure finds applications in several fields of digital signal processing, including estimation of the extent of attack transients for filters with dynamically varying inputs, elimination of transients in variable recursive filters, and design and implementation of linear-phase IIR systems.

1. INTRODUCTION

The impulse response of a stable recursive digital filter is infinitely long in principle, but due to exponential decay it eventually sinks below the quantization step or the noise in the system. Thus, in practice the impulse response of a stable recursive filter can be regarded as finite. A measure for the effective length of the impulse response of an IIR filter is needed in several applications, e.g., in estimation of the effective length of the attack transient of a recursive filter [2].

When changing the coefficients of a recursive filter, transients will occur. These transients depend on the filter input, but an impulse-response-based measure can be used to characterize them. A special case of this problem is encountered when the transients are eliminated using a novel technique by updating the state variables of the filter [10], [11]. The transient can be canceled within desired accuracy, but this accuracy depends on the effective length of the impulse response of the filter after the change of coefficients.

Still another application for the effective length of an infinite impulse response is a realization technique for linear-phase IIR filters based on cascading a minimum-phase IIR filter $H(z)$ and its maximum-phase (unstable or noncausal) counterpart $H(z^{-1})$ [4], [1], [8]. The filtering is based on processing the input signal in finite-length blocks of L samples. The basic constraint is to choose L so that the impulse response of $H(z)$ has decayed to a small enough level. On the other hand, block length L should be chosen as small as possible to minimize latency. Although L is an essential system parameter, techniques to determine its value are rather heuristic and do not attempt to find an optimal value. In [4] it was suggested that the filter be implemented in parallel form employing second-order filter sections and using a rough time-constant-based measure for the length of the impulse response of each section. An upper bound for estimating the resulting errors for a given L was derived in [1] and [8] but no explicit measure for determining L was given.

Previously, three different amplitude-based methods have been used for measuring the effective length of an infinitely long but decaying impulse response. 1) In [7], a general duration d of a signal was defined. The discrete-time version of the expression is

$$d^2 = \frac{1}{E}\sum_{n=-\infty}^{\infty} n^2 |x(n)|^2 \quad \text{with} \quad E = \sum_{n=-\infty}^{\infty} |x(n)|^2 \qquad (1)$$

where E is the total energy of the signal. 2) A traditional technique is based on the concept of a time constant. Typically, the time constant of the pole with the largest radius r_{max} is used for estimating the decay rate of the impulse response and an amplitude threshold is chosen to determine the effective length [6]. Smith has proposed to approximate this time constant as $1/(1 - r_{max})$ which is obtained by truncating the Taylor series of the exact equation [9], [11]. Based on merely one pole of the system, this measure is easy to use but gives a crude estimate for the effective length. 3) Furthermore, an amplitude threshold can be set and the effective length be determined as the sample index where the impulse response ultimately goes below this threshold [10]. In principle, this technique gives a better approximation. The drawbacks are the lack of analytical methods and the complication of the measure when the impulse response does not decay monotonically.

From the above it is apparent that several ways to measure the effective length of infinite impulse responses have been suggested but none of them seems to have gained wide acceptance. This paper introduces a meaningful yet simple and practical definition. *We define the effective length of the impulse response of a general recursive filter based on the accumulated percentage of the total energy*. This concept has several advantages: 1) the energy of an additive disturbance is a natural measure in many applications, 2) the total energy of a given filter is easy to determine either in the time or in the frequency domain, thanks to Parseval's theorem, and 3) the measure is parametric and thus flexible.

2. EFFECTIVE LENGTH OF A GENERAL RECURSIVE FILTER

2.1 Definitions

Consider an Nth-order recursive filter with transfer function

$$H(z)=\frac{B(z)}{A(z)}=\frac{b_0+b_1z^{-1}+\ldots+b_N z^{-N}}{1+a_1z^{-1}+\ldots+a_N z^{-N}} \quad (2)$$

where filter coefficients a_k and b_k are real-valued (k = 0, 1, ..., N). Assuming a stable and causal implementation, the recursive filter (2) can also be described via an equivalent difference equation as

$$y(n)=\sum_{k=0}^{N} b_k x(n-k)-\sum_{m=1}^{N} a_m y(n-m), \quad \text{for } n\geq 0 \quad (3)$$

where $x(n)$ and $y(n)$ are the input and output of the filter, respectively. When the input signal is a unit impulse $x(n) = \delta(n)$, which equals unity at n = 0 and zero elsewhere, the output $y(n) = h(n)$ is the impulse response of the filter.

The *total energy* of the causal impulse response $h(n)$ is defined as

$$E=\sum_{n=0}^{\infty} h^2(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|H(e^{j\omega})\right|^2 d\omega = \frac{1}{2\pi j}\oint H(z)H(z^{-1})z^{-1}dz \quad (4)$$

where the frequency-domain expression follows from the Parseval relation. The determination of the integral in the z-domain has been addressed in [3], for example.

We define the *energy-based effective length (EL)* as the smallest nonnegative integer time index N_P by which at least P% of the total energy of the impulse response has arrived. The corresponding *accumulated energy* $E_A(N_P)$ can be expressed as

$$E_A(N_P)=\sum_{n=0}^{N_P} h^2(n)\geq E_P=\frac{P}{100}E \quad (5)$$

Hence, we always require $E_A(N_P) \geq E_P$ since the effective length N_P must be an integer. Note that this differs slightly from the usual definition of length of the corresponding FIR filter: the truncated part contains N_P + 1 samples but the effective length (5) is one less, N_P. The energy-based length (for any percentage) of a filter with a unit impulse as the impulse response is thus zero, and that of a two-point averager is unity, which is in accordance with common sense.

2.2 General Algorithm

The most straightforward way to compute the impulse response of a given causal and stable recursive filter is to use the difference equation (3). When the total energy E is precomputed, the corresponding accumulated energy $E_A(N_P) \geq E_P$ for the chosen percentage P can be determined recursively via the algorithm presented in Table 1. This simple algorithm can be used for many recursive filters. However, for narrowband filters the length can be hundreds of samples. For low-order all-pole filters more practical closed-form expressions can be derived.

Table 1. Algorithm for computing the effective length of a general recursive filter.

Step 0: Compute E and E_P for the chosen P. Initialize: n = 0, $x(n)=\delta(n)$, $h(-1)=h(-2)=\ldots=h(-N)=0$, $E_A(-1)=0$
Step 1: $h(n)=\sum_{k=0}^{N} b_k x(n-k)-\sum_{m=1}^{N} a_m h(n-m)$
Step 2: $E_A(n)=E_A(n-1)+\lvert h(n)\rvert^2$
Step 3: If $E_A(N_P) \geq E_P = PE/100$, then $N_P = n$ and stop; else $n = n + 1$ and go to Step 1.

3. LOW-ORDER ALL-POLE FILTERS

3.1 First-Order All-Pole Filter

Consider a first-order all-pole filter with the transfer function

$$H(z)=1/(1-az^{-1}) \quad (6)$$

where a is real-valued and the pole radius $|a| = r < 1$ for stability. Its causal impulse response is simply $h(n) = a^n$ for nonnegative n. Accumulated energy $E_A(N_P)$ can be expressed as

$$E_A(N_P)=\sum_{n=0}^{N_P}\left(a^2\right)^n=\left[1-\left(r^2\right)^{N_P+1}\right]\Big/\left(1-r^2\right) \quad (7)$$

from which the total energy is also obtained as a limit ($N_P \to \infty$) as $E = 1/(1 - r^2)$. The requirement (5) now becomes

$$E_A(N_P)\geq\frac{P}{100}\frac{1}{1-r^2}=\frac{P}{100}E \quad (8)$$

and the EL can be solved as

$$N_P=\left\lceil\frac{\log(1-P/100)}{\log(r^2)}-1\right\rceil \quad (9)$$

where the logarithm can have any (positive) base and $\lceil\cdot\rceil$ denotes the ceiling operation (i.e., rounding upwards). Note that quantization is necessary because N_P must be an integer.

Figure 1 presents the EL N_P for P = 90%, 95%, and 99% as a function of pole radius r computed according to (9). These curves show the expected phenomenon that the EL of the impulse response increases rapidly as pole radius r approaches the value 1. Furthermore, it is seen that the EL is fairly insensitive to the percentage value so that the lengths corresponding to 90%-99% energy do not differ much except for when pole radius r is larger than 0.9.

3.2 Second-Order All-Pole Filters

Similar derivations can be conducted for second-order all-pole filters. Three different cases have to be elaborated separately: a complex-conjugate pair, a double real pole, and two distinct real poles. The derivations are more involved than in the first-order case. Furthermore, exact closed-form formulas cannot be derived, but simplified approximations or upper and lower bounds can be arrived at for the complex-conjugate case. For the other two cases, it is only possible to derive closed-form formulas for accumulated energy $E_A(N_P)$ and total energy E. Unfortunately,

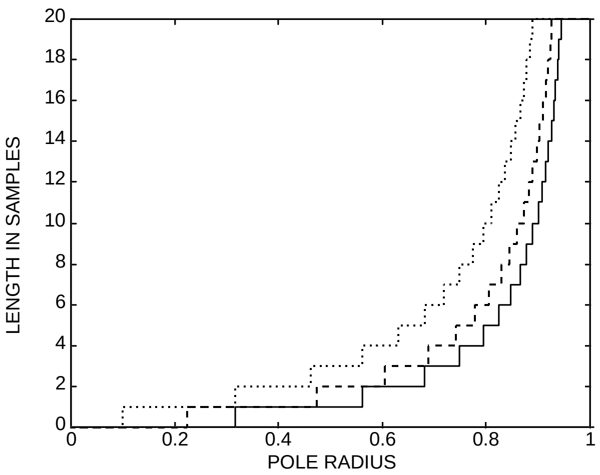

Figure 1. The effective length of a first-order all-pole filter for P = 90% (solid line), P = 95% (dashed line), and P = 99% (dotted line) as a function of pole radius r.

these do not lend to an easy closed-form solution for N_P, but they can be used to efficiently search for minimum N_P by successive evaluations. Using binary search, about $\log_2(N_P)$ evaluations are needed, as compared to N_P steps of the algorithm of Table 1. For example, if we can assume that the EL is at most 256, only 8 evaluations of $E_A(N_P)$ and E are required. The derivations are omitted due to space limitations. Details are available in a long version of this work [5].

4. ON THE EFFECT OF ZEROS

The above results consider all-pole filters only. In this section we show how the zeros affect the EL of recursive filters' impulse response. A general first-order filter is studied in detail after which general conclusions are drawn for higher-order filters.

4.1. General First-Order IIR Filter

Let us consider a first-order IIR filter with transfer function

$$H(z) = c(1 - bz^{-1}) \big/ (1 - az^{-1}) \tag{10}$$

where a, b, and c are real-valued and $|a| < 1$. The impulse response is now

$$h(n) = \begin{cases} 0, & n < 0 \\ c, & n = 0 \\ c(a-b)a^{n-1}, & n \geq 1. \end{cases} \tag{11}$$

The accumulated energy $E_A(N_P)$ is (for $N_P > 0$)

$$\begin{aligned} E_A(N_P) &= c^2\left[1 + \sum_{n=1}^{N_P} (a-b)^2 a^{2(n-1)}\right] \\ &= c^2 + c^2(a-b)^2(1 - a^{2N_P}) \big/ (1 - a^2) \end{aligned} \tag{12}$$

from which the total energy is obtained as a limit (as $N_P \to \infty$)

$$E = c^2(1 - 2ab + b^2) \big/ (1 - a^2) \tag{13}$$

The EL can now be solved as

$$N_P = \left\lceil \frac{\log(1 - P/100) + \log[L(a,b)]}{\log(a^2)} - 1 \right\rceil \tag{14}$$

where $L(a,b) = (1 - 2ab + b^2) \big/ (1 - b/a)^2$.

It is seen that (14) is the same as (9) except for a additive new term $\log[L(a,b)]$. Since $\log(a^2) < 0$, this term increases the length of the impulse response when $L(a,b)$ is smaller than unity, which happens when $|b - a| > |a|$. In the limit the additional term goes asymptotically towards the minimum value $\log[L(a,b)] \to \log(a^2)$ when $|b| \to \infty$, which means that the impulse response is lengthened by one sample at most. In this case the numerator approximates a unit delay, i.e., $1 - bz^{-1} \approx bz^{-1}$.

On the other hand, the impulse response is shorter than (or equal to) that without the zero when $|b - a| < |a|$. For zeros close enough to the pole, the EL is suppressed down to zero. When $b = a$, the zero exactly cancels the pole and the impulse response reduces to a unit impulse.

4.2 *N* Zeros

The conclusions for the first-order filter can readily be generalized for higher-order filters. Consider a general recursive transfer function $H(z) = B(z)/A(z)$ with the numerator $B(z)$ of order M_B. Assuming a fixed denominator, the longest possible impulse response corresponds to a delay of M_B units (one per each zero) and it is attained when the highest-order coefficient b_{M_B} of $B(z) = b_0 + b_1 z^{-1} + \ldots + b_{M_B} z^{-M_B}$ is large enough compared to the others. The EL thus has an upper bound

$$N_P\{H(z)\} \leq N_P\left\{\frac{1}{A(z)}\right\} + M_B \tag{15}$$

The smallest possible EL for the high-order filter is zero which naturally occurs due to (approximate) cancellation of all of the poles by corresponding zeros. This result is used in the next section to obtain a general bound for high-order filters.

5. HIGH-ORDER RECURSIVE FILTERS

Analytical treatment of higher-order filters soon becomes cumbersome. Instead of trying to derive complicated formulas of questionable utility, approximate upper bounds are derived. Let us focus on the case of effective length for a relatively large P (90...99.99%) so that most of the energy has arrived by time index N_P and we can neglect the tail of the impulse response. We define the length-N_P truncated impulse response as

$$h_{TR}(n) = \begin{cases} h(n), & \text{for } n = 0,\ 1,\ \ldots,\ N_P \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

As the truncated impulse response is genuinely finite-length, we can obtain a simple approximative limit for the length of the convolution of two impulse responses $h_1(n)$ and $h_2(n)$ with effective lengths N_{P1} and N_{P2} as

$$N_P\{h_1(n) * h_2(n)\} \approx N_P\{h_{1TR}(n) * h_{2TR}(n)\} \leq N_{P1} + N_{P2} \tag{17}$$

This follows because the length of the convolution of two sequences of lengths (N_{P1} + 1) and (N_{P2} + 1) is equal to $N_P + 1 = (N_{P1} + 1) + (N_{P2} + 1) - 1 = N_{P1} + N_{P2} + 1$, or $N_P = N_{P1} + N_{P2}$ (remember that the *effective* length is one shorter than the number of coefficients!). Applying this result for many convolutions we can express a formula for a filter consisting of K subsections:

$$N_P\{h_1(n)*h_2(n)*\ldots*h_K(n)\} \le N_{P1} + N_{P2} + \ldots + N_{PK} \quad (18)$$

Let us then consider a transfer function where poles are divided into at most second-order real-coefficient sections as follows:

$$H(z) = B(z)/A(z) = B(z)\Big/\prod_{k=1}^{K_A} A_k(z) \quad (19)$$

where the numerator $B(z)$ is of order M_B, and K_A denotes the number of sections in the denominator. Combining (18) with (15), we obtain an approximative upper bound for the EL as

$$N_P\{H(z)\} = N_P\left\{B(z)\Big/\prod_{k=1}^{K_A} A_k(z)\right\} \le M_B + \sum_{k=1}^{K_A} N_P\left\{\frac{1}{A_k(z)}\right\} \quad (20)$$

This is a general-purpose result which can be applied to any kind of stable filters when the factorization to first or second-order real-coefficient sections is available. Note that the obtained estimate for the EL is an approximate upper bound and it may be pessimistic for filters with poles and zeros close to each other.

6. APPLICATION EXAMPLE

Let us then consider a real-life example where the estimation of the length of the impulse response of the IIR filter is crucial. As discussed in the Introduction, linear-phase IIR filters can be implemented by cascading a minimum-phase IIR filter $H(z)$ and its maximum-phase counterpart $H(z^{-1})$. For this the effective length of $H(z)$ must be determined. In [4], Kormylo and Jain designed a third-order elliptic lowpass filter for the processing of a noisy ECG signal. The filter specifications were: passband ripple A_p = 0.05 dB, passband cutoff frequency $\omega_p = 0.175\pi$ (or 35 Hz for 400 Hz sampling frequency), and stopband attenuation A_p = 16 dB. For the cascaded linear-phase system the ripple values are of course doubled, i.e., the composite stopband attenuation is 32 dB.

For block implementation, an estimate for the length of the impulse response of the elliptic filter is required. In [4] it was suggested (apparently heuristically) that the length of four times the time constant τ of the pole with the largest radius should be used, which yields the length estimate of 24.25 sample intervals (using Smith's approximation, i.e., time constant $\tau = 1/(1 - r_{max})$ —in [4] no figures were given). The desired 32 dB stopband attenuation suggests that at most $10^{-3.2}$ = 0.00063096 or 0.063% of the impulse response energy can be lost in the truncation, which corresponds to P = 99.937%. This yields an energy-based EL (exact, using the algorithm of Table 1) of N_P = 21 samples, which is not far from the 4τ estimate.

In [8], Powell and Chau employed a seventh-order elliptic lowpass filter with the passband ripple A_p = 0.005 dB, passband cutoff frequency $\omega_p = 0.65\pi$ and stopband attenuation A_p = 35 dB. Requiring that a bound for the maximum amplitude of transient errors be 70 dB below the signal level, it was derived in [8] that the block length of 200 samples is necessary. By requiring the residual energy of the impulse response to be below 70 dB, i.e., $P = 100\% \times (1 - 10^{-7})$ = 99.99999%, results in the exact EL of N_P = 160 samples. Hence, assuming that the energy-based criterion is suitable for the application, 20% savings in the processing delay can be achieved by using the proposed EL of the impulse response.

7. CONCLUSIONS

A new approach for determining the effective length (EL) of the impulse response of a recursive filter based on the accumulated energy was proposed. The energy-based measure is argued to be better suited for many signal processing problems than former techniques that focus on the amplitude of the impulse response or the time constant of the system. Alongside a simple recursive algorithm to determine the EL for any stable IIR filter, closed-form formulas were derived for first-order all-pole and pole-zero filters. The effect of zeros was studied in a general case, and an approximate upper bound was derived for estimating the EL for higher-order filters using formulas for low-order filters. The results of this paper find applications in several fundamental and advanced signal processing problems. An example of the application of the new measure to the design of the block length in linear-phase IIR filtering was presented.

8. ACKNOWLEDGMENTS

The authors are grateful to Dr. Jonathan Mackenzie and Dr. Tony Wicks for helpful discussions.

9. REFERENCES

[1] R. Czarnach, "Recursive processing by noncausal digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, no. 3, pp. 363–370, June 1982.

[2] R. W. Hamming, *Digital Filters. Second Edition.* Englewood Cliffs, NJ: Prentice-Hall, 1983, pp. 244–245.

[3] E. I. Jury, *Theory and Application of the z-Transform Method*. New York, NY: Wiley, 1964, p. 298.

[4] J. Kormylo and V. K. Jain, "Two-pass recursive digital filter with zero phase shift," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 22, pp. 384–387, Oct. 1974.

[5] T. I. Laakso and V. Välimäki, "Energy-based effective length of the impulse response of a recursive filter," unpublished manuscript, Mar. 1997.

[6] S. J. Orfanidis, *Introduction to Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[7] A. Papoulis, *Signal Analysis*. New York, NY: McGraw-Hill, 1977, p. 291.

[8] S. R. Powell and P. M. Chau, "A technique for realizing linear phase IIR filters," *IEEE Trans. Signal Processing*, vol. 39, no. 11, pp. 2425–2435, Nov. 1991.

[9] J. O. Smith, discussion with V. Välimäki, Nov. 1995.

[10] V. Välimäki, T. I. Laakso, and J. Mackenzie, "Elimination of transients in time-varying allpass fractional delay filters with application to digital waveguide modeling," in *Proc. Int. Computer Music Conf.*, Banff, AB, Canada, pp. 327–334, Sept. 1995.

[11] V. Välimäki, *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*. Dr. Tech. thesis. Espoo, Finland: Helsinki Univ. of Tech., Dec. 1995.