TECHNICAL REPORT

Improvement of an IIR asymmetric compensation gammachirp filter

Masashi Unoki^{1,2,5}, Toshio Irino^{1,3,6} and Roy D. Patterson⁴

¹ATR Human Information Processing Research Laboratories,
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0288 Japan

²Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Tatsunokuchi-machi, Nomi-gun, Ishikawa, 923-1292 Japan

³CREST, Japan Science and Technology Corporation

⁴CNBH, Department of Physiology, University of Cambridge,
Downing Street, Cambridge, CB2 3EG United Kingdom

⁵Currently with CNBH, Department of Physiology, University of Cambridge

⁶Currently with NTT Communication Science Laboratories

(Received 22 January 2001, Accepted for publication 22 May 2001)

Abstract: An IIR implementation of the gammachirp filter has been proposed to simulate basilar membrane motion efficiently (Irino and Unoki, 1999). A reasonable filter response was provided by a combination of a gammatone filter and an IIR asymmetric compensation (AC) filter. It was noted, probably however, that the rms error was high when the absolute values of the parameters are large, because the coefficients of the IIR-AC filter were selected heuristically. In this report, we show that this is due to the sign inversion of the phase of poles and zeros in the conventional model. We propose a new definition of the IIR-AC filter and we describe a method of systematic determining the optimum coefficients and number of cascade for the second-order filter. This results in a reduction of the error to about 1/3 that produced by the conventional model.

Keywords: Auditory filter, Gammatone, Gammachirp, IIR asymmetric compensation filter, Filter design

PACS number: 43.64.Bt, 43.66.Ba

1. INTRODUCTION

Recently, the gammachirp filter was proposed as an auditory filter to explain both perceptual and physiological data [1–3]. The gammachirp filter is an extension of the popular gammatone filter; it has an additional frequency-modulation term [1]. Irino and Unoki (1999) developed an infinite impulse response (IIR) implementation of the gammachirp filter to simulate dynamic properties that of the cochlear processing, that are difficult with the FIR filters [4]. They demonstrated that the analytic gammachirp filter is reasonably approximated with a combination of an IIR asymmetric compensation (AC) filter and an IIR gammatone filter [5]. The residual error was still fairly large however when the parameters were close to the boundary of the reported range, because the coefficients of the IIR-AC filter were selected heuristically. Moreover, there was no detailed investigation of the error over the wide parameter space that the gammachirp filter uses when fitting perceptual and physiological data. In this report, we

investigate the error and propose a new IIR-AC filter to reduce the error.

2. IIR GAMMACHIRP FILTER

In this section, we describe the definition of the gammachirp filter, the conventional design of the IIR implementation, and its limitation.

2.1. Gammachirp Filter

The amplitude spectrum of the gammachirp [1] is

$$|G_C(f)| = \frac{|a\Gamma(n+jc)|}{|2\pi\sqrt{\{b\text{ERB}(f_r)\}^2 + (f-f_r)^2}|^n} \cdot e^{c\theta}, \quad (1)$$

$$= a_0 \cdot |G_T(f)| \cdot |H_A(f)|,$$

where a is a constant, n and b are the parameters defining the envelope of the gamma distribution, f_r is the asymptotic frequency, $\text{ERB}(f_r)$ is the equivalent rectangular bandwidth of an auditory filter at f_r , and c determines the rate of the frequency modulation or the ‘chirp’ rate. a_0 is a constant, $|G_T(f)|$ is the amplitude spectrum of the

gammatone filter, $H_A(f) = e^{c\theta}$ is the asymmetric function, and $\theta = \arctan[(f - f_r)/\{b\text{ERB}(f_r)\}]$. The filter shape changes with level and it is important to associate the level dependency with the correct parameters [1]. The envelope parameters, n and b , determine the bandwidth of the gammatone filter; the bandwidth increases as b increases when n is fixed. The chirp parameter, c , largely determines the filter asymmetry on a linear frequency axis. c is typically negative in which case, the lower skirt of the frequency response moves upward and the upper skirt moves downward with the absolute value of c . See Ref. [3] for a detailed description with figures.

2.2. Conventional Design

To implement the gammachirp filter as an IIR filter, Irino and Unoki (1999) defined an asymmetric compensation filter $H_C(f)$ to simulate the asymmetric function $H_A(f)$ so that

$$G_C(f) \cong G_{CAC}(f) = a_1 \cdot G_T(f) \cdot H_C(f), \quad (2)$$

where a_1 is a real constant. The IIR asymmetric compensation (AC) filter is defined as follows:

$$H_C(z) = \prod_{k=1}^N H_{Ck}(z), \quad (3)$$

$$H_{Ck}(z) = \frac{(1 - r_k e^{j\phi_k} z^{-1})(1 - r_k e^{-j\phi_k} z^{-1})}{(1 - r_k e^{j\phi_k} z^{-1})(1 - r_k e^{-j\phi_k} z^{-1})}, \quad (4)$$

$$r_k = \exp(-k \cdot p_1 \cdot 2\pi b \text{ERB}(f_r)/f_s), \quad (5)$$

$$\phi_k = 2\pi\{f_r + p_0^{k-1} \cdot p_2 \cdot c \cdot b \text{ERB}(f_r)\}/f_s, \quad (6)$$

$$\varphi_k = 2\pi\{f_r - p_0^{k-1} \cdot p_2 \cdot c \cdot b \text{ERB}(f_r)\}/f_s, \quad (7)$$

where N is the number of cascade, p_0 , p_1 , and p_2 are positive coefficients, and f_s is the sampling frequency. The k -th pole and zero are defined with the absolute value, r_k , and the phases, ϕ_k and φ_k . In addition, another parameter was introduced to normalize the peak amplitude of $|G_{CAC}(f)|$ independently with the parameter variability. That is the frequency $f_{p,k}$ where the amplitude of the k -th second-order filter, $|H_{Ck}(z)|$, is normalized to unity.

$$f_{p,k} = f_r + k \cdot p_3 \cdot c \cdot b \text{ERB}(f_r)/n. \quad (8)$$

Irino and Unoki (1999) selected the coefficients for p_0 , p_1 , p_2 , and p_3 heuristically: the values are listed in columns 2 to 5 of the first row of Table 1 (when $n = 4$ and $N = 4$) [4]. The root-mean-squared (rms) error between $|G_C(f)|$ and $|G_{CAC}(f)|$ is 0.63 dB for 90 sets of parameter combinations $\{ N = 4; n = 4; b = 1.0, 1.35, 1.7; c = -3.0, -2.0, -1.0, 0.0, 1.0; f_r = 250, 500, 1,000, 2,000, 4,000, 8,000 \text{ Hz} \}$ in the range where $|G_C(f)| > -50$ dB.

2.3. Limitation of the Conventional Filter

The coefficients are sufficient in the most cases. However Irino and Patterson (2000) found it necessary to extend the parameter range when using a modified gammachirp with c values greater than 1 [2]. In this case, the average rms error was 1.48 dB for 6,816 sets of parameter combinations $\{ N = 4; n = 4; b = 0.5, 0.6, \dots, 2.0; c = -4.0, -3.9, -3.8, \dots, 3.0; f_r = 250, 500, 1,000, 2,000, 4,000, 8,000 \text{ Hz} \}$; that is, about two times the error for 90 sets. It is, however, possible to reduce the error using a nonlinear least-squared-error method, and this reduces the average rms error to 1.13 dB. The coefficients are listed in columns 2 to 5 in the second row of Table 1. To reduce the error below 1 dB, it is necessary to modify the definition of the AC filter, Eqs. (3) and (8) or the parameter dependency listed in Table 1.

The error distribution for the 6,816 parameter sets that produced on error of 1.48 dB is presented in Fig. 1(a) when $f_r = 2,000$ Hz. A horizontal plane at 1 dB is included to show where the error is above 1 dB. The rms error increases rapidly as the absolute value of c increases, and the rate of increase is greater when b is small. To understand the form of the error surface, we investigated the relationship between the parameter values and the location of the poles and zeros in $H_{Ck}(z)$.

2.4. Source of the Excessive Error

Figure 2(a) shows the absolute values of the pole and zero, r_k , and the magnitude of $|H_{Ck}(0)|$ at frequency zero as a function of the asymmetry parameter c . Figure 2(b) shows the phases of the poles, $\pm\phi$, and zeros, $\pm\varphi$, when $N = 4$, $n = 4$, $b = 1.68$, and $f_r = 2,000$. The absolute

Table 1 Coefficients and errors for the IIR Asymmetric compensation filter.

Condition	p_0	p_1	p_2	p_3	p_4	Error (dB)
Original	2	$1.35 - 0.19 \cdot c $	$0.29 - 0.0040 \cdot c $	$0.23 + 0.0074 \cdot c $	—	1.48
Opti.	2	$1.5363 - 0.2531 \cdot c $	$0.3437 - 0.0099 \cdot c $	$0.2342 + 0.0144 \cdot c $	—	1.13
Proposed		1.7818	0.5689	0.2523		
$N = 4$	2	$\cdot(1 - 0.0791 \cdot b)$ $\cdot(1 - 0.1655 \cdot c)$	$\cdot(1 - 0.1620 \cdot b)$ $\cdot(1 - 0.0857 \cdot c)$	$\cdot(1 - 0.0244 \cdot b)$ $\cdot(1 + 0.0574 \cdot c)$	1.0724	0.68
Proposed		1.7256	0.5874	0.2731		
$N = 5$	2	$\cdot(1 - 0.0809 \cdot b)$ $\cdot(1 - 0.1685 \cdot c)$	$\cdot(1 - 0.1345 \cdot b)$ $\cdot(1 - 0.1153 \cdot c)$	$\cdot(1 + 0.0303 \cdot b)$ $\cdot(1 - 0.0066 \cdot c)$	1.0230	0.44

Fig. 1 Rms error as a function of parameters b and c for (a) the conventional model, and the proposed models when (b) $N = 4$ and (c) $N = 5$. $f_r = 2,000$ Hz. The coefficients for the IIR-AC filter are listed in the first, third, and fourth rows in Table 1, respectively.

value increases monotonically as the parameter c decreases. The poles and zeros are symmetrically located about the radian $2\pi f_r/f_s$, and they diverge monotonically as

Fig. 2 The relationship between the parameter value of c and the location of the poles and zeros. (a) The absolute values of the poles and zeros, r_k , (lines with rectangles) and the magnitude of the second-order filter $|H_{Ck}(0)|$ at frequency zero (lines with triangles). (b) The phase of the poles (lines with crosses) and zeros (lines with circles). The line type distinguishes the number of filters in the cascade. Dotted line: $k = 1$; Dashed and dotted line: $k = 2$; Dashed line: $k = 3$; Solid line: $k = 4$.

parameter c decreases.

The asymmetric compensation filter, $H_C(f)$, is designed to simulate the asymmetric function, $|H_A(f)| = e^{c\theta}$ with a set of cascaded filters, $H_{Ck}(f)$, in which the poles and zeros gradually diverge as k increases. The magnitude of the asymmetric function at zero frequency, $|H_A(0)|$, increases exponentially as the absolute value of parameter c increases. So, it is necessary for the asymmetric compensation filter to have the same property. Figure 2(a) shows the magnitude of the asymmetric compensation filter at zero frequency, $|H_{Ck}(0)|$, and the absolute values of poles and zeros, r_k , as a function of the value of the parameter c . The magnitude $|H_{Ck}(0)|$ increases exponentially as c decreases when k is between 1 and 3. It is not, however, the case when k is 4. The magnitude of $|H_{C4}(0)|$ increases exponentially between c values of 0 and -2 , but saturates around -2.5 and it decreases for c below -3 . It is possible to explain this property using Fig. 2(b). The phase of the pole of the 4-th filter, ϕ_4 , becomes negative when c is less than about -2 , and the pole of the conjugate pair, $-\phi_4$, becomes positive. In this case, the pole of $-\phi_4$ forms a pair with the zero of ϕ_4 in the positive half plane. The frequency response of the filter has its maximum at the pole frequency and it decreases monotonically as frequency decreases below the pole frequency in the positive half plane. As shown in Fig. 2(b), the value of $-\phi_4$ and the corresponding peak frequency increase as c decreases below -2 . As a result, the magnitude response at zero frequency decreases

monotonically as c decreases. The distortion in the 4-th filter affects the composite response of the asymmetric compensation filter and results in the increment of the error at low c values. This is the case for the 3-rd pole, ϕ_3 , when b is small and for the 4-th zero, φ_4 , when c is positive and becomes large.

3. IMPROVEMENT

To overcome the problem in the previous section, we introduced another constraint into the phase variation defined in Eqs. (6) and (7). We also determined how the error depends on p_k , it was a linear function of c previously, as listed in the first row of Table 1.

3.1. Redefinition

We restricted the phase of the pole, ϕ_k , and zero, φ_k , to be within the positive half plain and zero. We also introduced one coefficient, p_4 , to improve the fit as a result of preliminary simulation. The new definition of r_k , ϕ_k , and φ_k is:

$$r_k = \exp(-p_1(p_0/p_4)^{k-1} \cdot 2\pi b \text{ERB}(f_r)/f_s), \quad (9)$$

$$\phi_k = \begin{cases} 2\pi\{f_r + \Delta f_r\}/f_s, & f_r + \Delta f_r \geq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

$$\varphi_k = \begin{cases} 2\pi\{f_r - \Delta f_r\}/f_s, & f_r - \Delta f_r \geq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

$$\Delta f_r = (p_0 \cdot p_4)^{k-1} \cdot p_2 \cdot c \cdot b \text{ERB}(f_r), \quad (12)$$

where p_0 , p_1 , p_2 , and p_4 are positive coefficients. The frequency for normalization using p_3 is the same as in Eq. (8).

We surveyed the dependency of the parameters b and c on the coefficients p_1 , p_2 , and p_3 as:

$$p_m = p_{m,a} \left(1 + \sum_{\beta=1}^{N_\beta} p_{m,b_\beta} \cdot b^\beta \right) \left(1 + \sum_{\gamma=1}^{N_\gamma} p_{m,c_\gamma} \cdot |c|^\gamma \right), \quad (13)$$

where $m = 1, 2$, and 3 ; p_0 is set to a constant 2 as previously; p_4 is set to a variable. The order of the polynomial is controlled by N_β and N_γ . The total number of variable coefficients is $N_t = 3(1 + N_\beta + N_\gamma) + 1$. Now, we can redefine the asymmetric compensation filters in Eqs. (3) and (4) using Eqs. (9) to (12) for parameter relationships defined with Eq. (13). A nonlinear least-squared-error method was applied to minimize the error between the original gammachirp in Eq. (1) and the IIR gammachirp in Eq. (2).

3.2. Determination of the Coefficients

We determined the coefficients that minimize the error for 6,816 sets of parameter combinations as described above. Table 2 shows results when parameters b and c take

Table 2 Relationship between the order of coefficients and the average rms error.

#Condition	Restrict sign inversion	Order for b N_β	Order for $ c $ N_γ	Total number N_t	RMS Error (dB)
0	no	0	1	6	1.13
1	yes	0	1	7	0.86
2	yes	1	0	7	1.84
3	yes	1	1	10	0.68
4	yes	0	2	10	0.86
5	yes	2	0	10	1.83
6	yes	1	2	13	0.68
7	yes	2	1	13	0.68
8	yes	2	2	16	0.66

on order 1 or 2 (columns 3 and 4); the total number of variable coefficients is listed in column 5, and the average rms error in column 6. The second column indicates whether the restriction for the sign inversion of the phase is applied or not. Condition 0 is for the conventional filter using Eqs. (5) to (7); the error is 1.13 dB as shown in the second row of Table 1. Conditions 1 to 8 are for the proposed filter with the restrictions of Eqs. (9) to (12). The total number of variable coefficients basically increases as the condition number increases.

The error in condition 1 is about 0.27 dB less than that of condition 0 mainly because of the restriction on sign inversion. The errors in conditions 2 and 5 are about 1.8 dB which is much greater than those of the other conditions including condition 0. So, the coefficients noted to vary with parameter c . Condition 8 has the minimum rms error, 0.66 dB. It is, however, inefficient relative to other options in the table. The number of the coefficients for condition 8 is 16. The errors for conditions 3, 6, and 7 are all 0.68 dB, i.e., only 0.02-dB worse than for condition 8. Condition 3 has 10 coefficients and so it is most efficient in this set. The errors for conditions 1 and 4 are 0.86 dB which is worse than that for condition 3. The trade off between error and number of coefficients, we suggest that condition 3 provides the best solution. The values for this condition are listed in the third row of Table 1. The error is reduced by 0.45 dB from 1.13 dB to 0.68 dB with this conventional, but optimized, version of the IIR gammachirp.

3.3. Number of Cascade

We have focused to this point on the number of filters in the cascade (N in Eq. (3)) of 4 up to values of consistency with the results from the conventional filter. It would be possible, however, to improve the results further using the new definition. Figure 3 shows the average rms error as a function of the number of filters in the cascade when the filter coefficients were bi-linear functions of the parameters b and c as condition 3 in Table 2; the nonlinear

Fig. 3 Relationship between rms error and the number of cascaded filters, N .

least-mean-squared method was applied to minimize the error for 6,816 sets of parameter combinations. The dashed line shows the rms error for the conventional filter of Eqs. (5) to (7). The error increases for N below 4 and reduces for N above of 5. It appears that the conventional filter cannot decrease the error simply by increasing the length of the cascade. The solid line shows the rms error for the filter defined with Eqs. (9) to (12). The error is always smaller than for the corresponding conventional filter; moreover it decreases monotonically as N increases. The function asymptotes, however, around an N of 5. So, N of 5 is a reasonable solution for minimizing both the error and the number of filters in the cascade. The final values for the coefficients are listed in the final row of Table 1; the average rms error is 0.44 dB, about 1/3 of the error with the original filter (1.13 dB) over a wide parameter space.

Figure 1(b) and (c) show the error distributions when $N = 4$ and $N = 5$ as a function of the parameters b and c , when $f_r = 2,000$ Hz. It is obvious that, when $N = 5$, (Fig. 1(c)) the error is smaller than when $N = 4$, and much smaller than for the conventional cascade (Fig. 1(b)). The error is less than 1 dB for most of the parameter sets. The maximum rms error is 3.52 dB when $b = 0.5$ and $c = -4$. So, it is a sufficient result for our propose for simulation of the gammachirp.

4. SUMMARY

We have shown how the approximation error of the IIR asymmetric compensation (AC) gammachirp filter can be reduced. Firstly, we showed that the reason for the large

errors was the relationship between the parameter values and the locations of the poles and zeros. The rms error was reduced by restricting the IIR AC filter to avoid sign inversion of the phases of poles and zeros. Then, the coefficients were systematically determined as a bi-linear function of the filter parameters b and c . Finally, the number of filters in the cascade was determined from the relationship between the number of cascaded filters and the rms error of the fit. The average rms error was reduced to 0.44 dB for the parameter sets of 6,816, which is about 1/3 of the error for the conventional filter, although the total number of variable coefficients increases from 6 to 10. The error distribution was significantly improved especially when the absolute value of c is large and b is small.

ACKNOWLEDGEMENTS

The authors wish to thank M. Akagi of JAIST and S. Katagiri of ATR-HIP for the arrangements, H. Kawahara of Wakayama University / CREST and P. Zolfaghari of Nagoya University for the discussion. This work was partially supported by Grant-in-Aid for Science Research from the Ministry of Education (Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists) and by CREST (Core Research for Evolutional Science and Technology) of JST (Japan Science and Technology Corporation).

REFERENCES

- [1] T. Irino and R. D. Patterson, "A time-domain level dependent auditory filter: The gammachirp", *J. Acoust. Soc. Am.*, **101**, 412 (1997).
- [2] T. Irino and R. D. Patterson, "A gammachirp framework for auditory filtering: unification of cochlea frequency-glide data and psychoacoustical masking data", *XII-th Int. Symp. on Hearing*, Mierlo (2000).
- [3] T. Irino and R. D. Patterson, "A compressive gammachirp auditory filter for both physiological and psychophysical data", *J. Acoust. Soc. Am.*, **109**, 2008 (2001).
- [4] T. Irino and M. Unoki, "An analysis/synthesis auditory filterbank based on an IIR implementation of the gammachirp", *J. Acoust. Soc. Jpn. (E)*, **20**, 397 (1999).
- [5] M. Slaney, "An efficient implementation of the Patterson-Holdsworth auditory filter bank", *Apple Comput. Tech. Rep.* #35 (1993).