

## AN ANALYSIS/SYNTHESIS AUDITORY FILTERBANK BASED ON AN IIR GAMMACHIRP FILTER

Toshio Irino<sup>1</sup> and Masashi Unoki<sup>1,2</sup>

<sup>1</sup> ATR Human Information Processing Research Labs.

2-2 Hikaridai Seika-cho Soraku-gun Kyoto, 619-02, JAPAN

<sup>2</sup> Japan Advanced Institute of Science and Technology

1-1 Asahidai Tatsunokuchi Nomi Ishikawa, 923-1292, JAPAN

### 1. Introduction

A number of auditory models have been developed for telecommunications systems that incorporate human auditory characteristics. Recent attempts do a good job, based on physiological comparisons, simulating the peripheral auditory system (for a review, see [4]). But, unfortunately, none of these models have had as much success in speech recognition systems as linear predictive analysis and the Fourier transforms. There are a number of reasons for this dilemma. An obvious problem is that realistic, non-linear auditory models require complex calculation that preclude real-time processing. However, this problem should be resolved in the near future by fast digital signal processors. Another factor is that these models do not facilitate proper signal resynthesis procedure, which are straight forward for both linear predictive analysis and Fourier analysis because of their linearity.

Linear auditory filterbanks or wavelet transforms have been used for signal resynthesis [2][24], but they are unable to account for the dynamic characteristics of basilar membrane motion. Iterative procedures to reconstruct signals from cochleagrams (i.e., short-time averaged amplitude responses of basilar membrane motion without phase information) [6][23] are applicable to such non-linear filterbanks, but can not guarantee the precision of the resynthesis because of local minima. Thus it is desirable to develop an adaptive auditory filterbank that also provides a sound resynthesis procedure resulting in no perceptual distortion. This paper shows that such an adaptive, analysis/synthesis filterbank is possible through the implementation of a new “gammachirp” auditory filter [9].

The gammachirp function was analytically derived to have minimal uncertainty in a joint time-scale representation [1][7][8]. The gammachirp auditory filter is an extension of the popular gammatone filter (for a review, see [17]); it has an additional frequency-modulation term to produce an asymmetric amplitude spectrum. When the degree of asymmetry is associated with the stimulus level, the gammachirp filter provides an excellent fit to 12 sets of notched-noise masking data from three different studies [9]. The gammachirp has a much simpler impulse response than recent physiological models on cochlear mechanics [4], which do provide a good fit to human masking data. Moreover, the chirp term in the gammachirp is consistent with physiological observations on frequency-modulations or frequency “glides” in measurements of the mechanical responses of the basilar membrane [3][15][20].

The gammachirp filter has been implemented as a finite impulse response (FIR) filter because the gammachirp is defined as a time-domain function. Including this filter in an

auditory filterbank, however, poses problems. For simulations of the dynamic characteristics of the cochlea the filter coefficients have to be recalculated and then convolved with the signal on a moment-by-moment basis. Unfortunately, the large number of FIR coefficients, especially at low frequencies, precludes fast implementations. Moreover, the simulation becomes unrealistic if the filter output is not calculated simultaneously with the update of the filter coefficients. The calculation of the filter output and the update of the filter coefficients need to be performed simultaneously. Therefore, the gammachirp filter should be implemented with a small number of filter coefficients using an infinite impulse response (IIR) filter [10][11].

IIR implementations of modified gammatone filters have been developed to introduce asymmetry into auditory filter shapes, i.e., the All-Pole Gammatone Filter (APGF) or One-Zero Gammatone Filter (OZGF) [13][18][22]. The shapes of these filters, however, depend on the sampling rate of the system [10] and have not been directly fitted to psychoacoustical masking data. Moreover, it seems difficult to resynthesize signals from their output representations without uncontrollable errors since they did not provide a well-defined synthesis scheme. These issues are the main topics of this paper.

Section 2 describes an IIR implementation of the gammachirp. Subsection 2.1 shows the definition and the Fourier transform of the gammachirp decomposed into a gammatone and an asymmetric function. Subsection 2.2 explains the characteristics of the asymmetric function. Subsection 2.3 shows that the asymmetric function can be implemented by an IIR “asymmetric compensation filter”. Subsection 2.4 shows the approximation error in the amplitude spectrum between the original gammachirp filter and the combination of a gammatone and an IIR asymmetric compensation filter. Subsection 2.5 shows the stability of the inverse filter of the IIR filter, which enables signal resynthesis using the procedure described in subsection 3.3. Section 3 shows an implementation of the gammachirp filterbank. Subsection 3.1 shows an example of an adaptive analysis filterbank controlled by the sound pressure level estimated at the output of the filterbank. Subsection 3.2 shows another example of a filterbank based on physiological constraints. Finally, subsection 3.3 describes an adaptive, analysis/synthesis auditory filterbank that has never been accomplished by conventional auditory models simulating basilar membrane motion.

## 2. Implementation of the Gammachirp Filters

### 2.1 Definition and Fourier Transform of the Gammachirp

The complex impulse response of the gammachirp [7][8][9] is given as

$$g_c(t) = at^{n-1} \exp(-2\pi b \text{ERB}(f_r)t) \exp(j2\pi f_r t + jc \ln t + j\phi), \quad (1)$$

where time  $t > 0$ ,  $a$  is the amplitude,  $n$  and  $b$  are parameters defining the envelope of the gamma distribution,  $f_r$  is the asymptotic frequency,  $c$  is a parameter for the frequency modulation or the chirp rate,  $\phi$  is the initial phase,  $\ln t$  is a natural logarithm of time, and  $\text{ERB}(f_r)$  is the equivalent rectangular bandwidth of the auditory filter at  $f_r$ . At moderate levels,  $\text{ERB}(f_r) = 24.7 + 0.108f_r$  in Hz [5]. When  $c=0$ , the chirp term,  $c \ln t$ , vanishes and this equation represents the complex impulse response of the gammatone that has the envelope of a gamma distribution function and its carrier is a sinusoid at frequency  $f_r$  [17]. Accordingly, the gammachirp is an extension of the gammatone with a frequency modulation term.

The Fourier transform of the gammachirp in Eq. (1) is derived as follows.

$$\begin{aligned}
G_C(f) &= \frac{a\Gamma(n+jc)e^{j\phi}}{\{2\pi b\text{ERB}(f_r) + j2\pi(f-f_r)\}^{n+jc}} \\
&= \frac{\bar{a}}{\left\{2\pi\sqrt{\bar{b}^2 + (f-f_r)^2} \cdot e^{j\theta}\right\}^{n+jc}} \\
&= \bar{a} \cdot \frac{1}{2\pi\left\{\sqrt{\bar{b}^2 + (f-f_r)^2}\right\}^n \cdot e^{jn\theta}} \cdot \frac{1}{\left\{2\pi\sqrt{\bar{b}^2 + (f-f_r)^2}\right\}^{jc} \cdot e^{-c\theta}}, \quad (2)
\end{aligned}$$

$$\theta = \arctan \frac{f-f_r}{\bar{b}} \quad (3)$$

where  $\bar{a} = a\Gamma(n+jc)e^{j\phi}$  and  $\bar{b} = b\text{ERB}(f_r)$ . The first term  $\bar{a}$  is a constant. The second term is known as the Fourier spectrum of the gammatone,  $G_T(f)$ . The third term represents an asymmetric function,  $H_A(f)$ , that is described in more detail in the next subsection. If we normalize the amplitude, the frequency response of the gammachirp can be represented as

$$G_C(f) = G_T(f) \cdot H_A(f). \quad (4)$$

The amplitude spectrum is

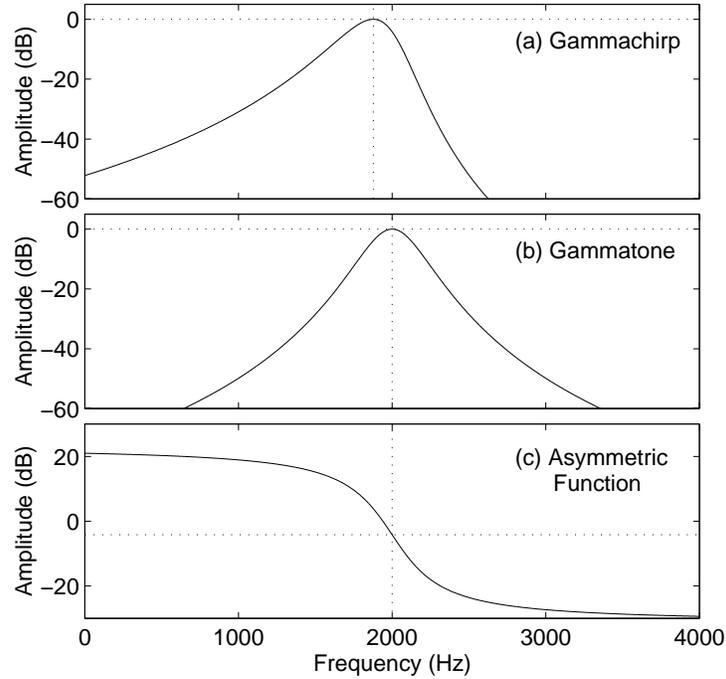
$$|G_C(f)| = |G_T(f)| \cdot |H_A(f)| = \frac{\bar{a}}{\left\{2\pi\sqrt{\bar{b}^2 + (f-f_r)^2}\right\}^n} \cdot e^{c\theta}. \quad (5)$$

Obviously, when  $c=0$ ,  $|H_A(f)|$  ( $=e^{c\theta}$ ) becomes unity and Eq. (5) represents the amplitude spectrum of the gammatone,  $|G_T(f)|$ . Figure 1 shows the amplitude spectra of (a) a gammachirp filter  $|G_C(f)|$ , (b) a gammatone filter  $|G_T(f)|$ , and (c) an asymmetric function  $|H_A(f)|$  with the chirp parameter  $c=-2$ . The amplitude of  $|H_A(f)|$  is biased by about -4 dB to normalize the peak of  $|G_C(f)|$  to 0 dB. Since the amplitude spectrum of the gammatone filter  $|G_T(f)|$  is almost symmetric on a linear-frequency axis, the asymmetric function  $|H_A(f)|$  introduces spectral asymmetry and a peak frequency shift into the gammachirp  $|G_C(f)|$ .

The peak frequency  $f_p$  in the amplitude spectrum is obtained analytically by setting the derivative of Eq. (4) to zero and solving the equation for the peak frequency. The result is

$$f_p = f_r + \frac{c \cdot \bar{b}}{n} = f_r + \frac{c \cdot b\text{ERB}(f_r)}{n}. \quad (6)$$

Therefore, the size of the peak shift is proportional to the chirp parameter  $c$  and the ratio of the envelope parameter  $b\text{ERB}(f_r)$  to  $n$ .



**Figure 1** Amplitude spectra of (a) a gammachirp filter  $|G_C(f)|$ , (b) a gammatone filter  $|G_T(f)|$ , and (c) an asymmetric function  $|H_A(f)|$ , where  $n=4$ ,  $b=1.019$ ,  $c=-2$ , and  $f_r=2000$  Hz.

## 2.2 Characteristics of the gammachirp and the asymmetric function

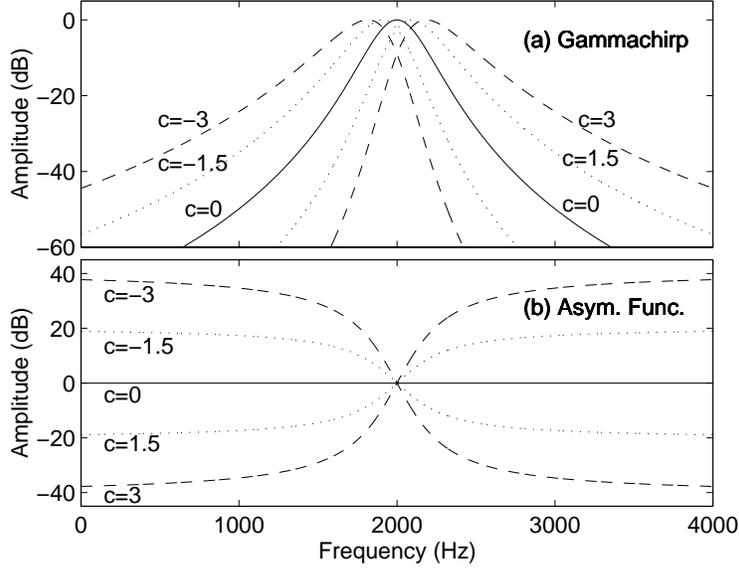
To describe the spectral characteristics of the gammachirp and the asymmetric function precisely Eq. (4) is rewritten in a form that explicitly uses the relevant parameters, that is,

$$G_C(f; n, b, c, f_r) = G_T(f; n, b, f_r) \cdot H_A(f; b, c, f_r). \quad (7)$$

The asymmetric function uses parameters  $b$ ,  $c$ , and  $f_r$ , whereas the gammatone uses parameters  $n$ ,  $b$ , and  $f_r$ .

Figure 2 shows the amplitude spectra of (a) the gammachirp  $G_C(f; n, b, c, f_r)$  and (b) the asymmetric function  $H_A(f; b, c, f_r)$  when the values of the chirp parameter  $c$  are integers between -3 and 3. Several characteristics are derived from this figure and the equations described above.

(a) Figure 2(a) shows that the filter slope of a gammachirp below the peak frequency is shallower than the slope above it when the parameter  $c$  is negative. The situation is the reverse when the parameter  $c$  is positive. The filter shape is symmetric when  $c$  is zero because the chirp term is removed and the resulting function is identical to the standard gammatone function.



**Figure 2** Amplitude spectra of (a) a gammachirp filter  $|G_c(f)|$  and (b) an asymmetric compensation filter  $|H_A(f)|$  as a function of the chirp parameter  $c$  where  $n=4$ ,  $b=1.019$ , and  $f_r=2000$  Hz. The amplitude is normalized to 0 dB at the peak frequency in panel (a) and at  $f_r$  in panel (b).

(b) The asymmetric function  $H_A(f; b, c, f_r)$  in Fig. 2(b) is an all-pass filter when  $c=0$ . Using Eq. (2),

$$H_A(f; b, 0, f_r) = 1. \quad (8)$$

$H_A(f; b, c, f_r)$  is a high-pass filter when  $c > 0$ , and a low-pass filter when  $c < 0$ . The slope and the range of the amplitude increase when the absolute value of  $c$  increases. The filter shapes of the gammachirp in Fig. 2(a) reflect these characteristics.

(c)  $H_A(f; b, c, f_r)$  changes monotonically in frequency. Neither a peak nor a dip ever occurs in this function.

(d) For an arbitrary frequency  $f_i$ , the asymmetric function is restricted by

$$|H_A(f_r - f_i; b, c, f_r)| = |H_A(f_r + f_i; b, c, f_r)|^{-1}. \quad (9)$$

(e) With Eq. (2), the asymmetric function satisfies:

$$H_A(f; b, c, f_r) = H_A(f; b, -c, f_r)^{-1}. \quad (10)$$

(f) For arbitrary chirp parameters  $c_1$  and  $c_2$ , the asymmetric function is multiplicative with respect to  $c$ :

$$H_A(f; b, c_1 + c_2, f_r) = H_A(f; b, c_1, f_r) \cdot H_A(f; b, c_2, f_r). \quad (11)$$

(g) Using Eqs. (7), (10), and (11),

$$\begin{aligned} G_C(f; n, b, c, f_r) &= G_T(f; n, b, f_r) \cdot H_A(f; b, c, f_r) \\ &= G_T(f; n, b, f_r) \cdot H_A(f; b, c_1 + c_2, f_r) \cdot H_A(f; b, -c_2, f_r) \\ &= G_C(f; n, b, c_1 + c_2, f_r) \cdot H_A(f; b, -c_2, f_r) \end{aligned} \quad (12)$$

Equation (12) states that a gammachirp with an arbitrary chirp parameter  $c_1$  is a product of a gammachirp with a different chirp parameter  $c_1 + c_2$ , and an asymmetric function having the difference between them  $-c_2$ . This is because the asymmetric function  $H_A(f; b, c, f_r)$  is an exponential function of the parameter  $c$ .

These characteristics are necessary conditions for designing the approximation filter in the next subsection, and they act as a guide for establishing an analysis/synthesis filterbank in Section 3.

### 2.3 Asymmetric compensation filter

As shown by Eq. (4), a gammachirp filter can be implemented by cascading a gammatone filter and an asymmetric filter. Since efficient implementations of the gammatone are already known [17][22], this section concentrates on an approximation filter for the asymmetric function described in the previous section. It is necessary to design a filter satisfying the conditions (a) through (g) in the previous section. As a first step, a filter satisfying condition (d) is considered because its characteristic seem the most relevant for filter design purposes.

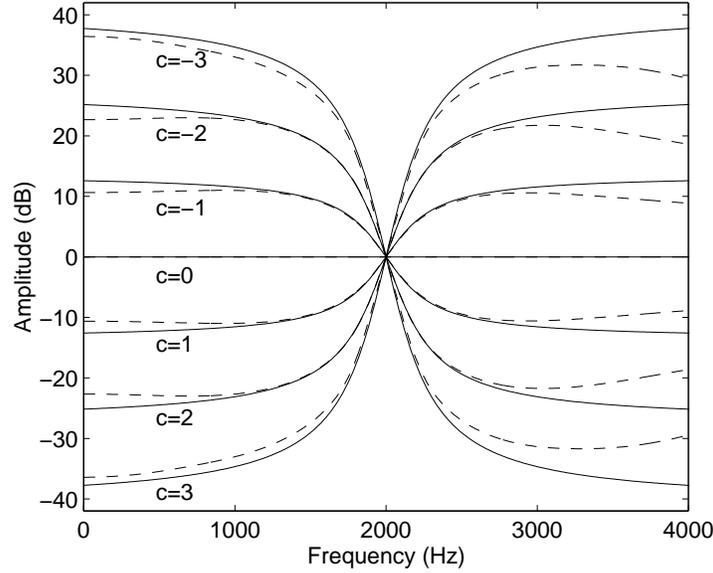
FIR filters cannot satisfy Eq. (9) in the strict sense since they only have zeros and no poles. They can, however, satisfy Eq. (9) approximately if a linear-phase FIR filter designed with the Remez algorithm is employed. Unfortunately, this is ineffective since the number of coefficients is comparable to that of the original FIR gammachirp and, moreover, the coefficients seem to require a table indexed with parameters  $b$ ,  $c$ , and  $f_r$ . The well-known IIR Butterworth and Chebyshev filters cannot satisfy Eq. (9) either. Consequently, a new IIR filter has to be designed that is an explicit function of these parameters to satisfy this condition.

IIR filters satisfying Eq. (9) have the same numbers of poles and zeros symmetrically located at  $f_r + \Delta f$  and  $f_r - \Delta f$  for a design frequency  $\Delta f$ . This makes the magnitude,  $r$ , of the corresponding poles and zeros equal. In addition these magnitude must be less than unity for the IIR filters to be stable; this is known as the minimum phase condition [16]. Since the bandwidth gets narrower when  $r$  gets closer to unity,  $r$  is negatively correlated with the bandwidth parameter  $b\text{ERB}(f_r)$ . Condition (b) implies that  $\Delta f$  is proportional to  $c$  and is positively correlated with  $b\text{ERB}(f_r)$ . A cascaded second-order digital filter satisfying these properties is

$$H_C(z) = \prod_{k=1}^N H_{Ck}(z) \quad (13)$$

$$H_{Ck}(z) = \frac{(1 - r_k e^{j\phi_k} z^{-1})(1 - r_k e^{-j\phi_k} z^{-1})}{(1 - r_k e^{j\phi_k} z^{-1})(1 - r_k e^{-j\phi_k} z^{-1})}, \quad (14)$$

$$r_k = \exp\{-k \cdot p_1 \cdot 2\pi b\text{ERB}(f_r)/f_s\} \quad (15)$$



**Figure 3** Amplitude spectra of asymmetric functions  $|H_A(f)|$  (solid lines) and asymmetric compensation filters  $|H_C(f)|$  (dashed lines) where  $n=4$ ,  $b=1.019$ ,  $c$  is an integer between -3 and 3, and  $f_r=2000$  Hz. The amplitude is normalized to 0 dB at  $f_r$ .

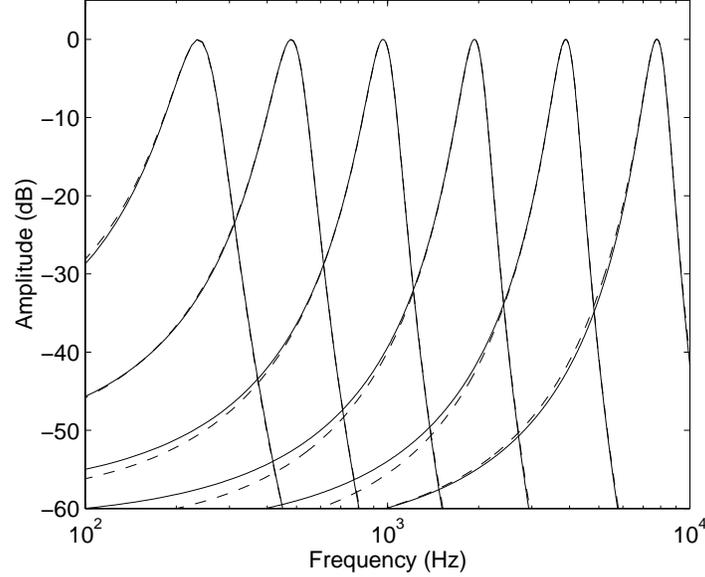
$$\phi_k = 2\pi \{ f_r + p_0^{k-1} \cdot p_2 \cdot c \cdot 2\pi b \text{ERB}(f_r) \} / f_s \quad (16)$$

$$\varphi_k = 2\pi \{ f_r - p_0^{k-1} \cdot p_2 \cdot c \cdot 2\pi b \text{ERB}(f_r) \} / f_s \quad (17)$$

where  $p_0$ ,  $p_1$ , and  $p_2$  are positive coefficients and  $f_s$  is the sampling rate. The reason for cascading filters with gradually located poles and zeros is to satisfy condition (c) approximately. This filter is referred to as an “asymmetric compensation (AC)” filter.

Figure 3 shows the amplitude spectra of this digital filter  $H_C(f)$  (dashed lines) and the asymmetric function  $H_A(f)$  (solid lines) in Eq. (5) as a function of the chirp parameter  $c$ . There were four cascaded filters, the amplitude was normalized at frequency  $f_r$ , and the values of  $p_0$ ,  $p_1$ , and  $p_2$  were set as described in the next subsection. The dashed lines are very close to the solid lines when the frequency is less than 3000 Hz. Above 3000 Hz, however, the disparity gets larger. This, however, does not cause serious errors, because the asymmetric compensation filter is always accompanied by the gammatone filter, which is a band-pass filter.

The results will show that four cascaded second-order filters provide a reasonable fit when the parameter  $b$  is equal to or greater than unity and the chirp parameter  $c$  is between -3 and 1. In this case, there are a total of 16 poles and zeros. Although it is possible to improve the fitting by increasing the number of cascaded filters, a reasonable number can be deter-



**Figure 4** Amplitude spectra of original FIR gammachirp filters  $|G_c(f)|$  (solid lines) and asymmetric compensation (AC) gammachirp filters  $|G_{CAC}(f)|$  (dashed lines) where  $n=4$ ,  $b=1.019$ ,  $c=-1$ , and the values for  $f_r$  are 250, 500, 1000, 2000, 4000, and 8000 Hz.

mined by considering the trade-off between the number of coefficients and the degree of fitting.

#### 2.4 Asymmetric compensation gammachirp

The asymmetric compensation filter cascaded with the gammatone filter approximates the gammachirp filter. The amplitude spectrum of this filter is found by replacing  $H_A(f)$  with  $H_C(f)$  in Eq. (6),

$$|G_{CAC}(f)| = |G_T(f)| \cdot |H_C(f)|. \quad (18)$$

This filter  $|G_{CAC}(f)|$  is referred to as an “Asymmetric Compensation—gammachirp” or “AC—gammachirp” filter until the end of Section 2, so as to distinguish it from the original gammachirp defined by Eq. (1).

##### 2.4.1 Comparison in the amplitude spectrum

Figure 4 shows the amplitude spectra of the gammachirp  $|G_C(f)|$  in Eq. (5) (solid lines), the AC-gammachirp  $|G_{CAC}(f)|$  in Eq. (14) (dashed lines), and the gammatone  $|G_T(f)|$ . The amplitude  $|G_{CAC}(f)|$  has been normalized properly to improve the fit. The frequency for normalizing the amplitude of each second-order filter is closely related to the peak shift in Eq. (6) and is set with a coefficient,  $p_s$ , for the  $k$ -th filter,

$$f = f_r + k \cdot p_3 \cdot c \cdot b\text{ERB}(f_r)/n. \quad (19)$$

The coefficients  $p_0, p_1, p_2$ , and  $p_3$  are set heuristically as

$$p_0 = 2, \quad (20)$$

$$p_1 = 1.35 - 0.19 |c|, \quad (21)$$

$$p_2 = 0.29 - 0.0040 |c|, \quad (22)$$

$$p_3 = 0.23 + 0.0072 |c|. \quad (23)$$

The root-mean-squared (rms) error between the original gammachirp filter and the AC-gammachirp filter is less than 0.41 in Fig. 4 over the range where  $|G_C(f)| > -50\text{dB}$ . The average rms error is only 0.63 dB for 90 sets of parameter combinations  $\{n = 4; b = 1.0, 1.35, \text{ and } 1.7; c = 1, 0, -1, -2, \text{ and } -3; f_r = 250, 500, 1000, 2000, 4000, \text{ and } 8000 \text{ (Hz)}\}$ . The rms error exceeds 2 dB only for three sets when  $f_r = 8000 \text{ Hz}$  and  $c = -3$ .

The fit improved only slightly when the coefficients in Eqs. (21), (22), and (23) were optimized using an iterative least squared-error method. It is possible to improve the fit by changing the locations of the poles and zeros defined in Eqs. (15), (16), and (17), but this is beyond the scope of this paper.

#### 2.4.2 Comparison of the impulse response and the phase spectrum

Figure 5(a) shows an example of the impulse response of the gammachirp defined in Eq. (1) (solid line) and the AC-gammachirp obtained from Eq. (18) (dashed line). The difference in the impulse responses between the original gammachirp and the AC-gammachirp is about -50 dB in the rms amplitude and, therefore, is almost negligible. Their phase spectra, shown in Fig. 5(b), are very close to each other. Therefore, the AC-gammachirp provides an excellent approximation to the original gammachirp in terms of its phase characteristics, i.e.,

$$G_C(f) \cong G_{CAC}(f) = G_T(f) \cdot H_C(f) \quad (24)$$

and also in the time domain,

$$g_c(t) \cong g_{CAC}(t) = g_T(t) * h_c(t), \quad (25)$$

where \* denotes convolution.

#### 2.4.3 Similarity to the asymmetric function

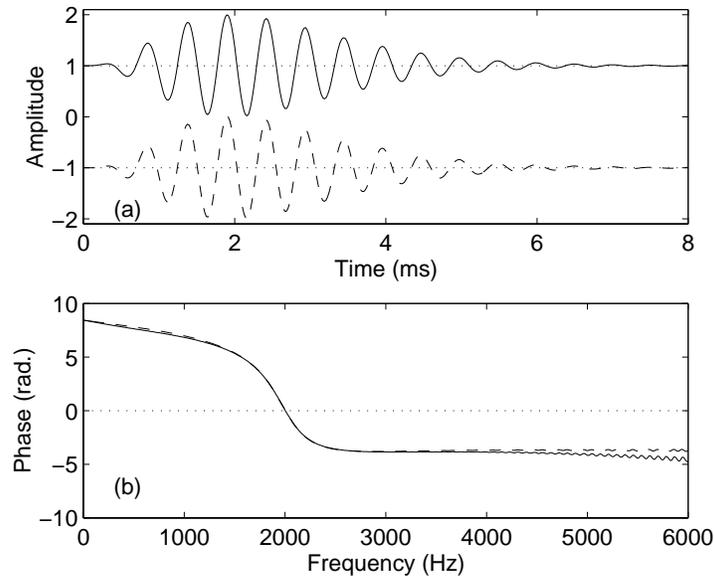
The asymmetric compensation filter  $H_C(z)$  defined in Eqs. (13) and (14) can strictly satisfy Eqs. (8), (9), and (10) and condition (b), and approximately satisfy Eq. (11) and condition (c). The reasons are as follows. For conditions (b), (c) and Eq. (11), the correspondence is obvious from Fig. 2. For Eq. (8), when  $c$  is 0, the locations of the poles and zeros of Eqs. (16) and (17) are the same and, then, Eqs. (13) and (14) become unity. For Eq. (9), since  $\text{ERB}(f_r)$  is a linear function of  $f_r$ , changing  $f_r + f_l$  to  $f_r - f_l$  simply replaces the poles and zeros in Eqs. (16) and (17). For Eq. (10), changing the sign of  $c$  replaces the poles and zeros in Eqs. (16) and (17) and it is possible to derive a stable inverse filter since the asymmetric compen-

sation filter satisfies the minimum phase condition. The inverse filter is always stable even if the parameter values are time varying. Accordingly, it is possible to cancel the forward filter with the inverse filter. Then, the total response of the combination is a unit impulse. This feature leads to an analysis/synthesis filterbank (described in subsection 3.3).

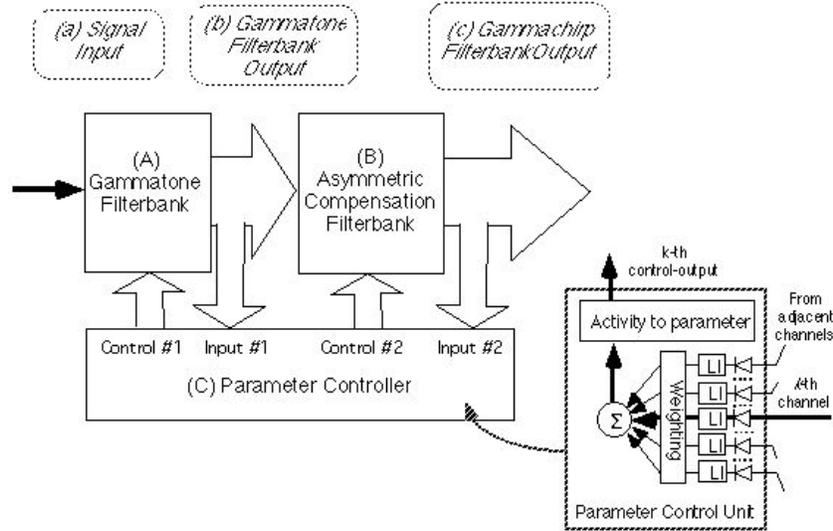
Since the IIR asymmetric compensation filter has few coefficients, fast level-dependent, adaptive auditory filtering can be performed by a combination of the compensation filter with a fast implementation of the gammatone [17][22].

### 3. GAMMACHIRP FILTERBANK

This section describes an adaptive, analysis/synthesis gammachirp filterbank. Since the auditory filter shape is level-dependent [5][9][14], it is necessary to estimate the sound pressure level of incoming signals. Subsection 3.1 shows an example of the analysis filterbank with a level estimation mechanism. Subsection 3.2 shows another type of filterbank structure using physiological constraints. Although no specific structure or parameter set has been determined yet, these examples are sufficient for presenting the most important issue in this paper. Subsection 3.3 shows the general structure of an adaptive, analysis/synthesis gammachirp filterbank. This chapter shows that sound resynthesis is always possible independent of the method of parameter control. The analysis/synthesis errors are shown to be time-invariant and small enough to avoid perceptual distortions, even when listening to synthetic sounds.



**Figure 5** (a) Impulse responses and (b) phase spectra of an FIR gammachirp filter (Eq. (1)) (solid lines) and an asymmetric compensation (AC) gammachirp filter (dashed lines). The parameters are  $n=4$ ,  $b=1.019$ ,  $c=-1$ , and  $f_p=2000$  Hz.



**Figure 6** Block diagram of a level-dependent gammachirp filterbank.

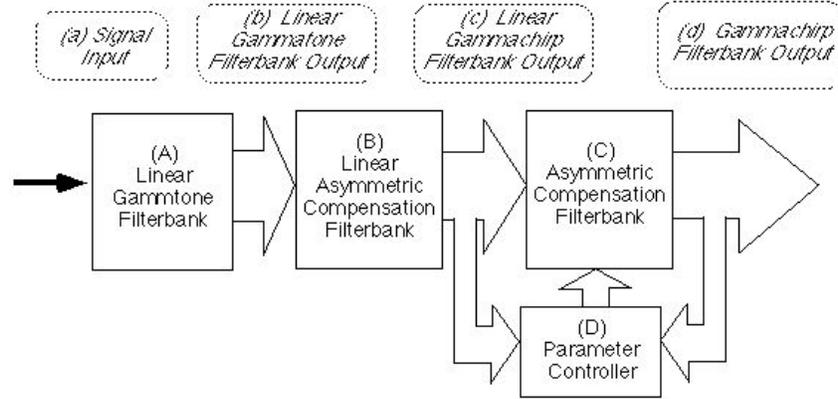
### 3.1 Implementation with level estimation

Figure 6 shows an example of a gammachirp filterbank that consists of a gammatone filterbank, a bank of asymmetric compensation filters, and a parameter controller. The sound pressure level of incoming signals is estimated in the parameter controller using the output of the gammatone filterbank and the asymmetric compensation filterbank. The parameter controller is a bank of parameter control units as shown in the right-bottom block. When considering the  $k$ -th channel, the input signal to this block is first rectified, and is then put into a leaky integrator (LI) for smoothing. Any value for the time constant is possible in the following simulation as long as the feedback system is stable; 30 ms was used for the error estimation in subsection 3.3. A weighting function, i.e., a Hamming window of 3 ERB width across the filter channels, is applied to the LI output of the  $k$ -th and adjacent channels, which are summed together to obtain the activity  $a_{ak}$  for the  $k$ -th channel. The estimated sound pressure level  $P_s$  in decibels is calculated using a straightforward equation,

$$P_s = 20 \log(q \cdot a_{ak}) \quad (26)$$

where  $q$  is a constant. The estimated sound pressure level controls the parameters of the gammachirp filterbank.

It has been demonstrated that the constant  $q$  can be determined using psychoacoustical masking data [10][11]. It does not, however, sufficiently describe the procedure and the result in this paper since they largely depend on the filterbank structure and the parameter controller. In those simulations, however, the individual filters were all set on the basis of the auditory filter shape at a probe frequency of 2000 Hz [9]. Obviously, it is necessary to use the outputs of several adjacent auditory filters. The formulation with the leaky integrator and



**Figure 7** Block diagram of a gammachirp filterbank based on physiological constraints.

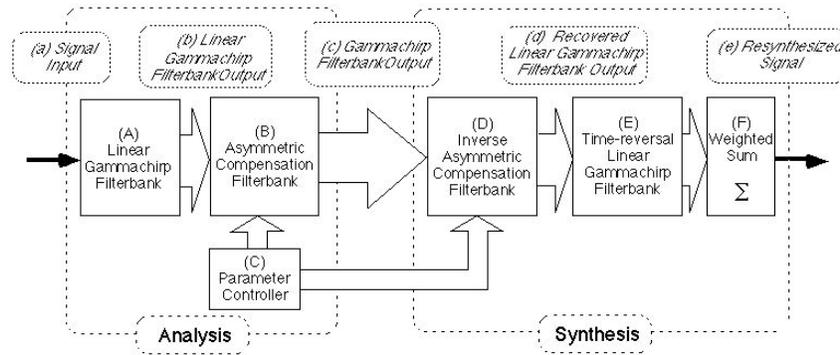
Eq. (26) is restricted to an initial approximation of the parameter control since it does not include fast compression [21] and the compression function is physiologically realistic [4].

Instead of determining the parameter values for this model, we use a basic structure to establish a synthesis procedure (described in subsection 3.3) that is independent of the method of parameter control. For the filterbank in Fig. 6, it is possible to perform signal resynthesis, provided only the chirp parameter  $c$  varies with the level in the level-dependent auditory filters. Then, control #1 (Fig. 6 (c)) is unnecessary because the gammatone filter is not a function of the chirp parameter  $c$ . The next subsection shows another filterbank structure based on physiological observations.

### 3.2 Another filterbank structure

Let us introduce physiological knowledge into the filterbank structure. When the sound pressure level is sufficiently high, the cochlear filter has a broad bandwidth and behaves like a passive and linear filter. As the signal level decreases, the filter gain increases and the bandwidth becomes narrower because of the active processes [19]. This suggests a physiologically plausible auditory filter is a combination of a linear, broadband filter and a nonlinear, level-dependent filter that sharpens the filter shape. Recent observations have shown that the frequency modulation or “glide” persists even post-mortem or after high sound pressure levels [20]. Accordingly, the linear filter can be simulated with a broadband gammachirp filter. As shown in Eq. (12), a gammachirp filter with an arbitrary chirp parameter  $c$  can be produced with a combination of another gammachirp filter and an asymmetric function. Therefore, the second filter can be simulated by a level-dependent asymmetric compensation filter as long as the total filter response can be simulated with the gammachirp.

Accordingly, a candidate filterbank structure is proposed in Fig. 7. It consists of a linear gammatone filterbank, a linear asymmetric compensation filterbank, and an adaptive asymmetric compensation filterbank controlled by a parameter controller. The output of the linear asymmetric compensation filterbank is equivalent to the output of a linear gammachirp filterbank (c). This output is fed into the asymmetric compensation filterbank to obtain the total output (d). The parameter controller is similar to that described in subsection 3.1. However, before determining the structure and the parameters it is necessary to wait for results fitting



**Figure 8** Block diagram of an adaptive, analysis/synthesis gammachirp filterbank.

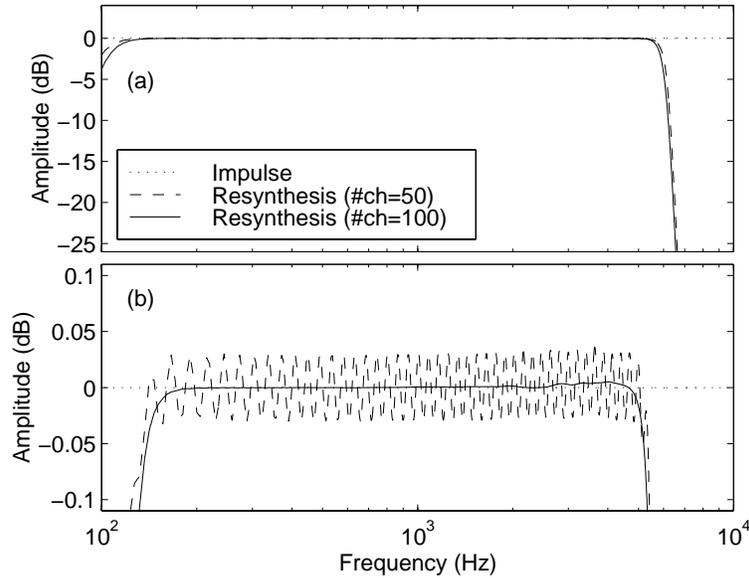
the gammachirp to psychoacoustical masking data across the full range of center frequencies. The structure is based on a combination of a linear filterbank with bandpass filters and a non-linear asymmetric compensation filterbank. For signal processing applications, this filterbank structure has a very important feature that has never been accomplished by conventional auditory filterbanks as described in the next subsection.

### 3.3 Analysis/synthesis filterbank

One of the most important features of the gammachirp filterbank is its ability to establish an analysis/synthesis system as shown in Fig. 8. Moreover, this feature is valid for any kind of parameter controller. Initially, a signal (a) is filtered by a linear, passive gammachirp filterbank (A). When the chirp parameter  $c$  is set to zero for all channels, this is a gammatone filterbank. The output of the linear filterbank (b) is converted into the output of the gammachirp filterbank (c) using a bank of active asymmetric compensation filters (B). Subsection 2.4 shows how to make a bank of inverse asymmetric compensation filters (D). The output of the adaptive gammachirp filterbank (c) is then converted into a representation (d) that is strictly the same as the output of the linear gammachirp filterbank (b) when using the same parameter set produced by the parameter controller (C) at each moment in time. The filterbank output is, then, equalized in phase using the time-reversal gammachirp filterbank (E); this is identical to the linear filterbank (A) except that the impulse response of each filter is reversed in time. Finally, the output after this phase equalization is summed with a weighting function to reproduce the signal.

A combination of the linear analysis filterbank (A), the linear synthesis filterbank (E), and the weighted sum (F) is almost equivalent to a linear, wavelet, analysis/synthesis procedure [2]. Since the combination of the asymmetric compensation filterbank (B) and its inverse filterbank (D) produces unit impulses for all channels, the error between the original and synthetic signals is strictly determined by this linear analysis/synthesis filterbank.

Figure 9 shows an example of analysis/synthesis frequency characteristics for an adaptive gammachirp filterbank with equally-spaced filters for ERB rates between 100 and 6000 Hz using a gammatone filterbank in (A) and (E), i.e., the gammachirp filterbank when  $c=0$  for all channels. Figure 9(b) shows the same graph with a magnified ordinate scale. The maximum error is less than 0.01 dB with 100 channels and is only about 0.03 dB even with



**Figure 9** Frequency responses of the analysis/synthesis gammachirp filterbank shown in Fig. 8 when the frequency range of the filterbank is between 100 and 6000 Hz and the number of channels is 50 (dashed lines) or 100 (solid lines). Panel (b) is the magnified ordinate of panel (a).

50 channels. It appears that about 100 channels are sufficient to minimize the errors. Moreover, the errors are completely independent of parameter control. Consequently, the gammachirp filterbank is able to perform signal resynthesis without producing any undesirable distortion.

The discussion above guarantees the minimum distortion of the analysis/synthesis filterbank system. This filterbank is applicable to various applications when inserting a modification block between the asymmetric compensation filterbank (B) and its inverse filterbank (C). For example, it is possible to construct a noise-suppression filterbank which does not produce any musical noise that is perceptually undesirable [12].

#### 4. SUMMARY

This paper presents an adaptive, analysis/synthesis auditory filterbank using the gammachirp. Initially, the gammachirp function is analyzed to find characteristics for effective digital filter simulation. The gammachirp filter is shown to be excellently approximated by the combination of a gammatone filter and an IIR asymmetric compensation filter. The new implementation reduces the computational cost for adaptive filtering because both filters can be implemented with only a few filter coefficients. The inverse filter of the asymmetric compensation filter is shown to be stable. Then, two examples of gammachirp filterbanks are presented, each is a combination of a linear gammachirp filterbank and a bank of adaptive, nonlinear asymmetric compensation filters, controlled by the signal-level estimation mechanism. A synthesis procedure for such analysis filterbanks is proposed to accomplish signal

resynthesis with a guaranteed precision and no undesirable distortion. This feature has never been accomplished with conventional auditory filterbanks. The adaptive, analysis/synthesis gammachirp filterbank is usable in various signal processing applications requiring the modeling of human auditory filtering.

#### ACKNOWLEDGMENTS

The authors wish to thank Roy D. Patterson of Cambridge Univ., for his continuous advice, Minoru Tsuzaki and Hani Yehia of ATR-HIP and Malcolm Slaney of IBM for their valuable comments. A part of this work was performed while the second author was a visiting student at ATR-HIP. The authors also wish to thank Masato Akagi of JAIST, and Yoh'ichi Tohkura, Hideki Kawahara, and Shigeru Katagiri of ATR-HIP for the arrangements. This work was partially supported by CREST (Core Research for Evolutional Science and Technology) of the Japan Science and Technology Corporation (JST).

#### REFERENCES

- [1] Cohen, L. "The scale transform." *IEEE Trans. Signal Processing*, 41, 3275–3292, 1993.
- [2] Combes, J.M., Grossmann, A. and Tchamitchian, Ph. Eds., *Wavelets*, Springer-Verlag, Berlin, 1989.
- [3] de Boer, E. and Nuttall, A. L. "The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions." *J. Acoust. Soc. Am.*, 101, 3583–3592, 1997.
- [4] Giguère, C. and Woodland, P.C. "A computational model of the auditory periphery for speech and hearing research. I. Ascending path." *J. Acoust. Soc. Am.*, 95, 331–342, 1994.
- [5] Glasberg, B. R. and Moore, B. C. J. "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, 47, 103–138, 1990.
- [6] Irino, T. and Kawahara, H. "Signal reconstruction from modified auditory wavelet transform." *IEEE Trans. Signal Processing*, 41, 3549–3554, 1993.
- [7] Irino, T. "An optimal auditory filter." in *IEEE Signal Processing Society*, 1995 Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 1995.
- [8] Irino, T. "A 'gammachirp' function as an optimal auditory filter with the Mellin transform." *IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP-98)*, 981–984, Atlanta, GA, 1996.
- [9] Irino, T. and Patterson, R.D. "A time-domain, level-dependent auditory filter: The gammachirp", *J. Acoust. Soc. Am.*, 101, 412–419, 1997.
- [10] Irino, T. and Unoki, M. "An efficient implementation of the gammachirp filter and its filterbank design." *ATR Technical Report*, TR-H-225, 1997.
- [11] Irino, T. and Unoki, M. "A time-varying, analysis/synthesis auditory filterbank using the gammachirp." *IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP-98)*, Seattle WA, 1998.
- [12] Irino, T. "Noise suppression using a time-varying, analysis/synthesis gammachirp filterbank." *IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP-98)*, Phoenix, AZ, 1999.
- [13] Lyon, R. F. "The all-pole gammatone filter and auditory models." in *Forum Acusticum '96*, Antwerp, Belgium, 1996.
- [14] Lutfi, R.A. and Patterson, R.D. "On the growth of masking asymmetry with stimulus intensity," *J. Acoust. Soc. Am.* 76, 739–745, 1984.
- [15] Møller, A.R. and Nilsson, H.G. "Inner ear impulse response and basilar membrane modelling." *Acustica*, 41, 258–262, 1979.
- [16] Oppenheim, A.V. and Schaffer, R.W., *Digital Signal Processing*. Prentice-Hall, New Jersey, 1975.
- [17] Patterson, R. D., Allerhand, M. and Giguère, C. "Time-domain modelling of peripheral auditory processing: a modular architecture and a software platform." *J. Acoust. Soc. Am.*, 98, 1890–1894, 1995.
- [18] Plüeger, M., Hoeldrich, R. and Reidler, W. "Nonlinear All-Pole and One-Zero Gammatone Filters," *Acta Acustica*, 84, 513–519, 1998.
- [19] Pickles, J.O., *An Introduction to the Physiology of Hearing*. Academic Press, London, 1988.

- [20] Recio, A.R., Rich, N.C., Narayan, S.S. and Ruggero, M.A. "Basilar-membrane response to clicks at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.*, 103, 1972–1989, 1998.
- [21] Robles, L., Rhode, W.S. and Geisler, C.D. "Transient response of the basilar membrane measured in squirrel monkeys using the Moessbauer effect," *J. Acoust. Soc. Am.*, 59, 926–939, 1976.
- [22] Slaney, M. "An efficient implementation of the Patterson-Holdsworth auditory filter bank." *Apple Computer Technical Report #35*, 1993.
- [23] Slaney, M. "Pattern Playback from 1950 to 1995," *IEEE Conf. Syst. Man., Cybernetics*, Vancouver, Canada, 1995.
- [24] Yang, X, Wang, K. and Shamma, S. A. "Auditory representations of acoustic signals." *IEEE Trans. Information Theory*, 38, 824–839, 1992.