Implementing a GammaTone Filter Bank*

John Holdsworth
Cambridge Electronic Design
Science Park
Milton Road
Cambridge

Ian Nimmo-Smith
MRC Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF

Roy Patterson
MRC Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF

Peter Rice
Cambridge Electronic Design
Science Park
Milton Road
Cambridge

26th February 1988

Introduction

The purpose of this technical note is to provide a self-contained summary of properties of the GammaTone filter and of the manner in which it can be effectively implemented through a cascade of identical recursive digital filters.

- Section 1 defines the GammaTone filter as an IIR filter in the time domain and describes its provenance and some of its elementary properties.
- Section 2 examines the behaviour of the GammaTone filter in the frequency domain, and shows that its form can lead to an (approximate) representation as a cascade of identical first order filters.
- Section 3 continues the analysis of Section 2 to provide a way of calculating the parameters needed for a GammaTone filter to have a specified equivalent rectangular bandwidth.
- Section 4 discusses the issue of phase compensation for the GammaTone filter.
- Section 5 describes the way in which these features have been exploited to achieve an efficient digital implementation of the GammaTone filter on a general purpose computer.

1 The GammaTone filter in the time domain

Prompted by de Boer and Kuyper (1968), the GammaTone filter was introduced by Johannsma (1972) to describe the shape of the impulse response function of the auditory system as estimated by the reverse correlation function of neural firing times. This was subsequently developed by de Boer and de Jongh (1978) and de Boer and Kruidenier (1988). For further details of the comparison between GammaTone filters, rounded-exponential filter shapes and experimental evaluation of the shape of the human auditory filter, see Patterson *et al.* (1987).

The GammaTone filter is defined in the time domain (impulse response function) as

$$gt(t) \propto t^{n-1} \exp(-2\pi bt) \cos(2\pi f_0 t + \phi) \qquad (t \geq 0) \tag{1}$$

*Annex C of the SVOS Final Report (Part A: The Auditory Filter Bank)

It is thus a causal filter with an infinite response time. The form of this function is that of an amplitude modulated carrier tone of frequency f_0Hz, with an envelope proportional to $t^{n-1}\exp(-2\pi bt)$, which is the familiar Gamma distribution from statistics. These two components give rise to the name GammaTone (de Boer and de Jongh, 1978).

The parameters of the GammaTone filter are n, the order, which (for fixed b) controls the relative shape of the envelope, becoming less skewed as n increases; b (in Hz) which (for fixed n) controls the duration of the impulse response function, increased b leading to shorter duration; f_0 (in Hz) which determines the frequency of the carrier; and ϕ (in radians), the carrier phase, which determines the relative position of the fine structure of the carrier to the envelope. All four parameters have corresponding effects on the frequency domain characteristics of the GammaTone filter.

2 The GammaTone filter in the frequency domain

If $GT(f)$ represents the GammaTone filter in the frequency domain (frequency response function) then

$$GT(f) \propto [1 + j(f - f_0)/b]^{-n} + [1 + j(f + f_0)/b]^{-n} \qquad (-\infty < f < \infty) \qquad (2)$$

This can be derived either directly from the time domain definition above, by application of the Fourier transform, or by observing that the product in the time domain of a Gamma function and a cosine function will correspond to the convolution in the frequency domain of the Fourier transform $(1 + jf/b)^{-n}$ of the Gamma function with a two-point distribution at $\pm f_0$. Here for simplicity we have set $\phi = 0$, as it has no important effect on the frequency domain characteristics of the filter.

The rôle of the parameters in the frequency domain is apparent from the above formula. f_0 is the centre frequency of the filter; for fixed order n, b acts as a scaling parameter such that the bandwidth of the filter increases with b; the order parameter n controls the overall shape of the filter. For fixed b, the bandwidth decreases as n increases. $GT(f)$ is approximately symmetric about f_0 on a linear frequency scale.

As stated by de Boer and Krudenier (1988), the second term in Eqn.(2) can be ignored when f_0/b is sufficiently large — which is always the case when modelling the human auditory filter — yielding an approximate frequency response function

$$GT(f) \approx [1 + j(f - f_0)/b]^{-n} \qquad (0 < f < \infty) \qquad (3)$$

For the GammaTone filters required in modelling the human auditory system ($4 < f_0/b < 8$), the second term in Eqn.(2) is typically 25 dB below the level of the first term, even at two bandwidths below the centre frequency.

Because of the n^{th} power form of Eqn.(3), the GammaTone filter of order n can be approximated by a cascade of n identical GammaTone filters of order 1 with frequency response function

$$[1 + j(f - f_0)/b]^{-1} \qquad (4)$$

each of which has the form of a frequency-shifted low-pass filter — a first-order filter which has a well-known recursive digital implementation.

The next section describes how to calculate the bandwidth of a GammaTone filter. Section 4 describes the way in which these cascade and recursive properties can be exploited to create efficient digital approximations to the GammaTone filter.

3 The equivalent rectangular bandwidth of the GammaTone filter

When f_0/b is large, the equivalent rectangular bandwidth of the GammaTone filter is proportional to b, and the proportionality constant a_n only depends on n.

We can calculate the equivalent rectangular bandwidth of the GammaTone filter by an application of Parseval's Theorem. If we assume that $|GT(f)|^2$ is symmetric in f, and that its maximum value occurs when $f = f_0$ then

$$\begin{aligned} \int_{-\infty}^{+\infty} gt(t)^2\, dt &= \int_{-\infty}^{+\infty} |GT(f)|^2\, df \\ &= 2|GT(f_0)|^2 ERB[GT] \end{aligned} \tag{5}$$

where $ERB[GT]$ is the equivalent rectangular bandwidth of the GammaTone filter.

Applying Eqn.(5) to the GammaTone filter leads to a tractable integral, yielding the following good approximation for the constant a_n:

$$a_n = \frac{\pi(2n-2)!\, 2^{-(2n-2)}}{(n-1)!^2} \tag{6}$$

The 3 dB bandwidth of a GammaTone filter of order n is $c_n b$ Hz where

$$c_n = 2\sqrt{2^{1/n} - 1} \tag{7}$$

Table 1 gives values of a_n and c_n which are accurate both for the 'true' GammaTone filter and to its approximation described above.

Table 1

n	a_n	a_n^{-1}	c_n	c_n^{-1}
1	3.141	0.318	2.000	0.500
2	1.570	0.637	1.288	0.777
3	1.178	0.849	1.020	0.980
4	0.982	1.019	0.870	1.149
5	0.889	1.164	0.772	1.296
6	0.773	1.293	0.700	1.429
7	0.709	1.411	0.646	1.550
8	0.658	1.520	0.602	1.662
9	0.617	1.621	0.566	1.767

Moore and Glasberg (1983) give the following formula for the equivalent rectangular bandwidth (in Hz) of the human auditory filter shape of centre frequency f_0 Hz.

$$ERB[f_0] = 6.23 \times 10^{-6} f_0^2 + 93.39 \times 10^{-3} f_0 + 28.52 \qquad (100 < f_0 < 10000) \tag{8}$$

Thus to match a GammaTone filter of order 4 and centre frequency 1000 Hz, the value of b is $a_n^{-1} ERB[1000] = 1.019 \times 128.14 = 130.57$ Hz. The 3 dBbandwidth of this filter is $c_n b = 0.870 \times 130.57 = 113.59$ Hz.

4 Phase compensation

When displaying the output of a bank of filters graphically, it is often useful to align the peaks of the envelopes of the impulse response functions and to align the peak of the fine structure with the peak of each envelope.

First, the peaks of the envelopes can be aligned by introducing a lead $t_c = (n-1)/2\pi b$ to the output of the filter. Second, a peak in the fine structure can be shifted to the peak of the envelope by introducing a phase correction $\phi_c = -2\pi f_0 t_c$.

Formally this leads to the (non-causal) filter

$$\tilde{gt}(t) \propto (t + t_c)^{n-1} \exp[-2\pi b(t + t_c)] \cos(2\pi f_0 t)(t \geq -t_c)(1')$$

which aligns the peak impulse response of all filters at $t = 0$.

5 Digital approximation by a cascade of recursive filters

In Section 2 it was shown how a GammaTone filter of (integral) order n could be approximated by a cascade of n GammaTone filters of order 1. The digital implementation of the frequency-shifted low-pass filter with frequency response function given by Eqn.(4) is achieved by the following algorithm.

Let the array $x[k]$ contain the input waveform, sampled at intervals of Δt. In order to avoid aliasing problems when the passband of the filter is close to the Nyquist frequency of the raw waveform ($f_0 > (4\Delta t)^{-1}$), we have used data-doubling (artificial over-sampling) before filtering, and data-halving (down-sampling) afterwards. The value of Δt in the algorithm is correspondingly halved.

- Start by frequency shifting the array $x[k]$ by an amount $-f_0$ Hz to produce the complex array $z[k]$.

$$z[k] = e^{-2\pi j f_0 k \Delta t} x[k] \tag{9}$$

- Next the array $z[k]$ is passed through a first-order recursive filter with output $w[k]$. When the order n of the required GammaTone is greater than 1, the output array is re-input to this recursive filter a further $n - 1$ times.

$$w[k] = w[k-1] + (1 - e^{-2\pi b \Delta t})(z[k-1] - w[k-1]) \tag{10}$$

- Finally a frequency shifting by $+f_0$ Hz is applied and the real part is taken to produce an output array $y[k]$.

$$y[k] = \Re(e^{+2\pi j f_0 k \Delta t} w[k]) \tag{11}$$

The form of digital approximation used in Eqn.(10) for the low-pass filter stage is that of 'impulse invariant design' (Oppenheim and Schafer, 1975, §5.1.1). As there is an implicit time delay of $\Delta t/2$ between the input x and the output y in this scheme, an appropriate adjustment needs to be made in implementing a cascade of such filters to get correct temporal alignment.

Informal testing indicates that a single GammaTone filter of order 4 implemented in this way can filter data sampled at 10 kHz in about real time on a Sun 3 system and about twice real time on a VAX system.

References

[1] E. de Boer and P. Kuyper, (1968) "Triggered correlation," *I.E.E.E. Trans. Bio. Med. Eng.*, **BME15-3**, 169-179

[2] E. de Boer and H. R. de Jongh, (1978) "On cochlear encoding: potentialities and limitations of the reverse-correlation technique," *J. Acoust. Soc. Am.*, **63**, 115–135.

[3] E. de Boer and C. Kruidenier, (1988) "On ringing limits of the auditory periphery," *(Manuscript in preparation).*

[4] P. I. M. Johannesma, (1972) "The pre-response stimulus ensemble of neurons in the cochlear nucleus," *Symposium on Hearing Theory,* 22–23 June, 1972, IPO, Eindhoven

[5] B. C. J. Moore and B. R. Glasberg, (1983) "Suggested formulae for calculating auditory filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.,* **74**, 750–753.

[6] A. V. Oppenheim and R. W. Schafer, (1975) *Digital Signal Processing,* Prentice-Hall International Inc., Englewood Cliffs, N.J.

[7] R. Patterson, I. Nimmo-Smith, J. Holdsworth and P. Rice, (1987) "An efficient auditory filterbank based on the gammatone function," *Paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE,* December 14-15, 1987