

Digital implementation of linear gammatone filters: Comparison of design methods

Luc Van Immerseel and Stefaan Peeters

Laboratory for Medical Electronics, University of Antwerp, B-2610 Antwerp, Belgium
luc.vanimmerseel@ua.ac.be

Abstract: Digital implementations of linear gammatone filters are regularly part of auditory models and can be used in the sound processing in cochlear implants. How close digital impulse, magnitude, and phase responses match the corresponding properties of the analog gammatone filter is evaluated for one, finite-impulse response filter design, and for five, infinite-impulse response filter designs: base-band impulse invariant transformation, impulse invariant transformation, matched z transformation, bilinear transformation, and mapping of differentials with backward differences. Filter properties and computational cost are compared as a function of the filter's order and center frequency, and sampling frequency. Results show that filters designed with the base-band impulse invariant transformation give the best overall approximation of the analog properties with reasonable computational cost.

© 2003 Acoustical Society of America

PACS numbers: 43.64.Bt, 43.66.Ba, 43.64.Me

Date Received: 22 August 2002 **Date Accepted:** 9 December 2002

Introduction

The gammatone function, introduced by Johannesma [1972], characterizes physiological impulse-response data gathered from primary auditory fibers in the cat. The gammatone auditory filter (GTF), developed by de Boer [1975], describes cochlear impulse responses of the cat's cochlea. The envelope of the impulse response is approximated with the gamma distribution known from statistics, and the fine structure of the impulse response is a sinusoid at the center frequency of the filter. The impulse response therefore equals

$$h(t) = At^{N-1} \exp(-2\pi bt) \cos(2\pi f_c t + \varphi) \quad (t \geq 0, N \geq 1) \quad (1)$$

with A an arbitrary factor that is typically used to normalize the peak magnitude transfer to unity, N the filter order, b a parameter that determines the duration of the impulse response and thus the filter's bandwidth, f_c the center frequency, and φ the phase of the tone. A fourth-order GTF gives the best fit with a wide range of human masking data [de Boer, 1975; Patterson et al., 1992]. Most auditory models use fourth-order GTFs, but lower or higher orders might be used in the sound-processing strategy in cochlear implants to optimize speech reception scores.

The GTF is a good model of the spectral analysis in humans at moderate sound levels where the shape of the auditory filter is reasonably symmetric on a linear frequency scale. Linear gammatone filterbanks are used in models that simulate the signal processing in the (central) auditory system [Patterson et al., 1995; Dau et al., 1997a, b; Shackleton et al., 2000] and can be used in the sound processing in cochlear implants. However, the linear GTF is inadequate at low levels where the active cochlear nonlinearity determines the sharp tuning properties, and at higher levels where the shape of the auditory filter becomes increasingly asymmetric [Glasberg and Moore, 1990]. Therefore, nonlinear (level-dependent) GTFs [Carney, 1993; Pflueger et al., 1997] and gammachirp filters [Irino and Patterson, 2001] have been proposed.

The GTF from Eq. (1) is an analog filter, where as auditory simulations and signal processing in cochlear implants make use of digital implementations. Linear GTFs have been digitally implemented as finite-impulse response (FIR) filters [Evans, 1986; Shackleton et al., 2000] or as infinite-impulse response (IIR) filters [Patterson et al., 1988; Cooke, 1991; Darling, 1991; Slaney, 1993]. The present study evaluates, for different digital filter design methods and

filter parameters, the differences between impulse, magnitude, and phase response properties of the linear analog GTF and of the digital implementation. The computational cost of the digital implementations is also estimated.

Methods

Implementation

The bandwidth of an auditory band-pass filter depends on center frequency and is often described by an equivalent rectangular bandwidth (ERB). For moderate sound levels, the relation between ERB and center frequency f_c in Hz is described by the equation from Glasberg and Moore [1990]. For gammatone filters, this equation is corrected with a factor 1.019 by Patterson et al. [1992], such that bandwidth b (Eq. [1]) of a gammatone filter in this study equals

$$b(f_c) = 1.019 \times 24.7(1 + 4.37f_c/1000). \tag{2}$$

Although this correction is optimized for fourth-order GTFs, it is used in this study independent of filter order, such that changing the filter’s order does not change the filter’s bandwidth.

Phase φ (Eq. [1]) has a neglectable influence on the power spectrum, but determines the shift of the fine structure of the impulse response relative to its envelope. Regardless of its value, the GTF phase response does not correspond to the auditory phase response [Recio and Rhode, 2000]. Therefore, and to simplify the analytical computations, phase φ is set to zero as in most other studies on GTF [Darling, 1991].

The transfer function of the analog GTF is derived using the Laplace transformation of the impulse response (Eq. [1]), as was first done by Slaney [1993], giving

$$H(s) \propto \prod_{n=1}^N H_n(s) = \prod_{n=1}^N \frac{s - s_n}{(s - p_0)(s - p_0^*)}. \tag{3}$$

This means that an N -th order GTF can be implemented as a cascade of N second-order filter sections, all having the same complex conjugate pole pair (p_0, p_0^*) and with one zero each (s_n).

Table 1 gives an overview of the six evaluated implementation methods, the abbreviations that are used throughout the text, and references to the design techniques. FIR filters are designed by rectangular windowing. The window length is set to the duration before the normalized impulse response of the GTF has decreased below a specified truncation level. Truncation levels ranging from 0.01 to 0.0001 are used. For high-order GTFs and small truncation levels, the FIR filters become noisy. Therefore, the FIR filter is always implemented as a cascade of FIR filters, each of them implementing a second-order section of the analog GTF (Eq. [3]) using the truncation criterion.

BBIT filter implementation follows the design rules outlined in Cooke [1991]. Since the filters are low-pass, the effective filter order is only one-half the order of the equivalent band-pass filter. Therefore, no noisy or unstable high-order filters are encountered. To avoid noisy or unstable IIR band-pass filters, they are implemented by applying the appropriate transformations (IIT, MZT, BILT, MDT) to the analog second-order transfer functions (Eq. [3]).

Table 1. Evaluated implementation methods, abbreviations, and references to the design techniques

Implementation method	Abbreviation	Reference
Finite-impulse response filter	FIR	Oppenheim et al. [1999]
Base-band impulse invariant transformation	BBIT	Cooke [1991]
Impulse invariant transformation	IIT	Oppenheim et al. [1999]
Matched z transformation	MZT	Oppenheim et al. [1999]
Bilinear transformation	BILT	Oppenheim et al. [1999]
Mapping of differentials transformation	MDT	Oppenheim et al. [1999]

Evaluation

Analog and digital impulse responses are normalized such that the absolute peak value equals unity. Matched z transformation results in a response lag of $N/2$ sampling periods with N the order of the GTF. This lag is compensated prior to the evaluation. The error is measured as the mean absolute difference between the sampled normalized analog impulse response and the impulse response of the digital implementation. Similar to Cooke [1991], the time range to evaluate the error is limited to the range where the envelope of the analog impulse response is above a specified fraction of its peak value. Envelope fractions ranging from 0.001 to 0.05 are used. Smaller or larger envelope fractions give smaller or larger impulse response errors respectively. The choice, however, does not influence the conclusions of the study. Therefore, only results for envelope fraction equal to 0.01 are presented.

Evaluation of the analog transfer function $H(s)$ (Eq. [3]) for $s = j\omega$ and of the digital transfer function $H(z)$ for $z = e^{j\omega}$ gives the analog and digital magnitude and phase transfer properties respectively. Analog magnitude responses are normalized to make the normalized transfer at the center frequency ω_c equal to 0 dB. Digital magnitude responses are normalized such that the normalized transfer at the effective center frequency equals 0 dB. Because phase shifts over multiples of 2π are not relevant, they are neglected. Magnitude and phase errors are measured as the mean absolute difference between analog and digital properties. The frequency range to compare magnitude and phase responses is restricted to $f_c \pm b(f_c)/2$ (Eq. [2]), which is independent of the filter's order. This frequency range is centered around f_c on a linear scale because the GTF is approximately symmetric on a linear scale. Normalized magnitude responses and phase responses are evaluated at multiples of 1 Hz within this restricted frequency range and thus are compared with the corresponding sampled analog responses. To compare magnitude and phase properties of the BBIIT implementation with those of the band-pass GTF, the frequency axis of the BBIIT filter's properties is shifted upwards by the center frequency f_c .

The differences between impulse, magnitude, and phase responses of the analog gammatone filter and any digital implementation are evaluated as a function of order (1 to 6), center frequency (100 to 5000 Hz), and sampling frequency (11025, 22050, and 44100 Hz). The results section shows only a representative subset: order equal to 2, 4, and 6; center frequency equal to 200, 1000, and 5000 Hz; sampling frequency equal to 22050 Hz. This choice does not influence the conclusions of this study.

The properties of analog and digital filters have been analytically computed with 64 bit floating-point accuracy such that round-off or truncation errors may be neglected in the results.

Results

A FIR filter design with truncation levels below 0.001 is comparable to but never better than an IIT filter design. This should not come as a surprise because, for a low truncation level, both methods essentially aim to preserve the impulse response of the GTF. With a higher truncation level, giving shorter FIR filters, the approximation of the GTF properties, is poorer than with the IIT design.

Fig. 1 shows impulse, magnitude, and phase response errors for all IIR design methods and for a FIR design with truncation level 0.001 only. For each set of parameters, the design method with the smallest error is marked with a bullet. BBIIT gives the best overall approximation of the analog impulse response. Long FIR, IIT, and MZT designs give comparable approximations except for high-order filters with higher center frequencies. BILT and MDT designs clearly result in a poorer impulse response approximation. With errors below 1 dB, long FIR, BBIIT, and IIT designs give excellent approximations of the analog magnitude response for all parameter combinations. MZT is only equally good for small ratios of center frequency to Nyquist frequency. BILT and MDT designs result in larger magnitude errors. BBIIT design always gives a close approximation of the analog phase response. Long FIR, IIT, and MZT designs are only comparable for small ratios of center frequency to Nyquist frequency. BILT and MDT designs give larger phase errors.

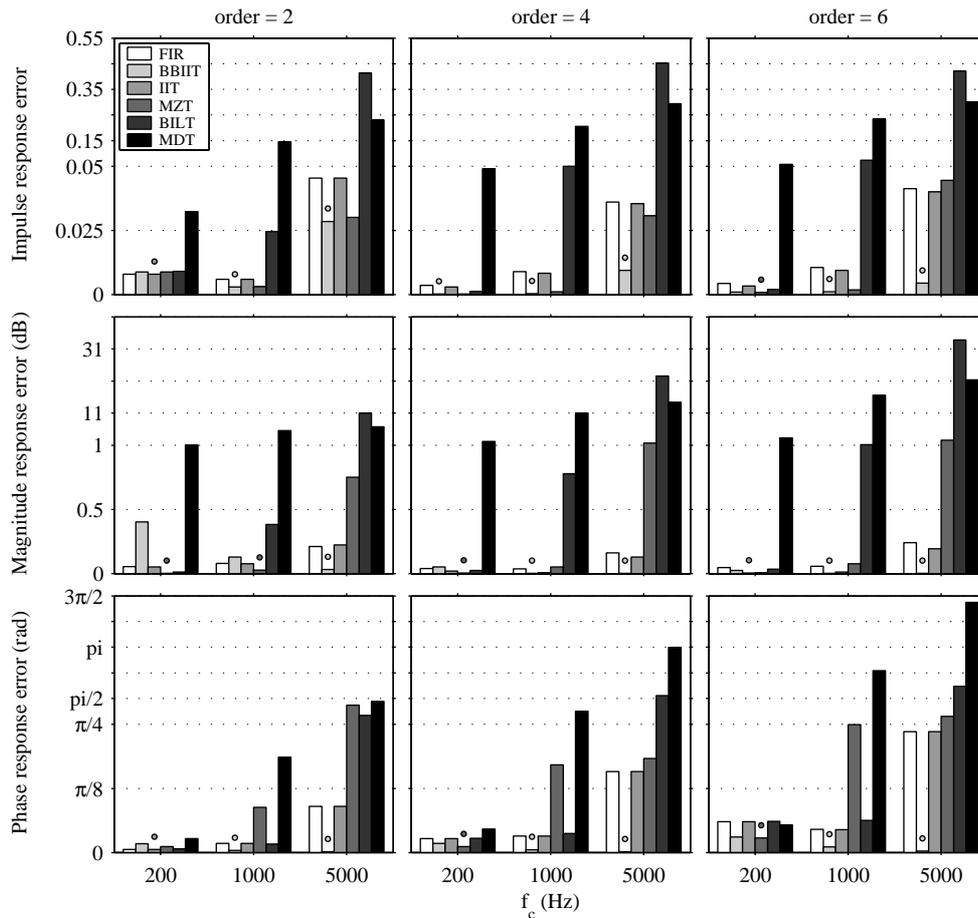


Fig. 1. Errors (mean absolute differences) of the impulse, magnitude, and phase responses. Vertical axes have piecewise linear scales.

Table 2. Computational cost per sample for the evaluated design techniques. N is the order of the GTF.

Design	# multiplications	# additions
BBIIT	$4N + 4$	$2N + 1$
IIT, MZT, BILT, MDT	$5N$	$3N$
FIR ($f_c = 200$ Hz)	$\approx 520N$	$\approx 519N$
FIR ($f_c = 1000$ Hz)	$\approx 180N$	$\approx 179N$
FIR ($f_c = 5000$ Hz)	$\approx 43N$	$\approx 42N$

From these results it follows that BBIIT is the best overall design technique. BBIIT results in an approximation of the analog GTF that is certainly accurate enough, seen the errors in the fit between the original physiological and psychoacoustical data and the GTF [Johannesma, 1972; de Boer, 1975; Patterson et al., 1992]. Long FIR filter, IIT, and MZT designs give comparable approximations for not too high ratios of center frequency to sampling frequency. Short FIR filter, BILT, and MDT implementations result in filters of which the properties deviate more from the analog properties. The observations in this study are in agreement with Cooke [1991] who compared BBIIT, BILT, and MZT designs for low-pass instead of band-pass GTFs.

Table 2 shows the computational cost of the evaluated implementations as a function

of the order N of the GTF. Cost is measured as the number of multiplications and additions per sample. The cost for a FIR design is shown only for truncation level 0.001 and sampling frequency 22050 Hz. The length of a FIR filter and, thus, its computational cost scales linearly with sampling frequency, and a larger truncation level results in shorter filters. The FIR filter cost is a linear fit to the actual cost for orders 1 to 6, with correlation above 0.999. The cost of the IIR filter designs is based on a transposed direct form II implementation and is independent of center and sampling frequencies.

The results show that the cost of the BBIIT implementation is lower than for band-pass IIR filter cascade implementations in case of high-order filters, but higher in case of low-order filters. The FIR filter implementation is much less efficient than the equally good IIT filter.

Discussion

The presented results are valid only for the filter topologies as described in the methods section. Digital implementation of the GTF with IIT, MZT, BILT, or MDT as one high-order filter instead of a cascade of second-order filter sections can result in a much noisier and even unstable filter. The same is true for the FIR filter implementation.

It is important to notice that the impulse response errors as measured in this study are biased. The largest differences between digital and sampled analog impulse responses typically occur at the time of onset. For higher center frequencies, and thus shorter impulse responses, these differences are averaged over a smaller number of samples than for lower center frequencies. When a fixed time range independent of center frequency is used instead, errors for higher center frequencies become relatively smaller, but it does not change the conclusion that BBIIT is the best design technique to approximate the impulse response of the GTF.

Cooke [1991] specifies the frequency range for the evaluation of magnitude and phase responses as the range where the filter's transfer is above an attenuation level, e.g., -60 dB. This gives larger ranges for low filter orders than for higher orders. As a consequence, errors in the tails of low-order filters contribute more to the error than in case of high-order filters. This biases the error because design methods concentrate the mapping accuracy around the most critical points, poles and zeros, where as filter skirts are not approximated equally accurate. In the present study, the frequency range is independent of the order and corresponds, approximately, to attenuation ranges of -20, -40, and -60 dB for second, fourth, and sixth-order filters respectively. Both methods to specify the frequency range, however, indicate that BBIIT is the best overall design technique to approximate magnitude and phase responses.

For sampling frequencies below or above 22050 Hz, errors are higher or lower respectively, mainly due to the different ratio of center frequency to sampling frequency. However, the ranking of implementation techniques is the same as for sampling frequency 22050 Hz.

Impulse response discrepancies might be important to any later processing that uses the fine temporal output of the filter, as some modeling studies do. Patterson [1994] and Patterson et al. [1995] use the base-band design with matched z transformation. Although this design is not evaluated in this study, it can be assumed, based on the results in Fig. 1, that this design will result in a good approximation of the impulse and magnitude responses for not too large ratios of center frequency to Nyquist frequency. Shackleton et al. [2000] use the FIR design and truncate the impulse response after 25.6 ms for all filters. For a center frequency of 200 Hz, this window length corresponds to an effective truncation level around or above 0.01, depending on the GTF order, which results in a design inferior to IIT and with higher computational cost.

It is difficult to assess the importance of differences in phase response because of discrepancies between GTF and auditory phase responses [Recio and Rhode, 2000]. That phase responses are influenced by the design method and by parameters, such as order, center frequency, and sampling frequency remains an important point of consideration for any auditory modeling that is sensitive to (differences in) phase, such as studies of modulation processing [Dau et al., 1997a, b; Derleth and Dau, 2000] or phase effects [Oxenham and Dau, 2001].

The use of a bank of linear GTFs in the sound processing of cochlear implants is a more evident choice than, for instance, a bank of Butterworth filters. At low or high levels,

however, only nonlinear filters can model the auditory properties. If correct phase properties are important, linear GTFs are also not adequate, as was already discussed above. The GTF can also be used as a basic filter in a wavelet filterbank with the advantage of variable time-frequency resolution. Also, then, the digital design technique remains important.

One must be careful to extrapolate the results of this study on linear GTF to the design of nonlinear GTF [Carney, 1993; Pflueger et al., 1997] or gammachirp filters [Irino and Patterson, 2001]. In nonlinear filters the bandwidth is varied with signal level [Carney, 1993], poles are moved dependent on level [Pflueger et al., 1997], or all parameters of the gammachirp filter change with level [Irino and Patterson, 2001]. This results in a much larger set of combinations of filter parameters than for linear filters. However, the claim of greater accuracy with nonlinear GTF models is only justified when the right design technique is used.

Acknowledgments

This work was supported by Advanced Bionics Corporation, Sylmar, CA. The authors thank Malcolm Slaney and an anonymous reviewer for their valuable comments.

References and links

- Johannesma, P. I. M. (1972). "The pre-response stimulus ensemble of neurons in the cochlear nucleus," in *Symposium on Hearing Theory* (IPO, Eindhoven, The Netherlands), pp. 58–69.
- de Boer, E. (1975). "Synthetic whole-nerve action potentials for the cat," *J. Acoust. Soc. Am.*, **58**, 1030–1045.
- Patterson, R. D., Robinson, K., Holdsworth, J. W., McKeown, D., Zhang, C., and Allerhand, M. (1992). "Complex sounds and auditory images," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 429–446.
- Patterson, R. D., Allerhand, M. H., and Giguère, C. (1995). "Time-domain modelling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.*, **98**, 1890–1894.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers," *J. Acoust. Soc. Am.*, **102**, 2892–2905.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," *J. Acoust. Soc. Am.*, **102**, 2906–2919.
- Shackleton, T. M., McAlpine, D., and Palmer, A. R. (2000). "Modelling convergent input to interaural-delay-sensitive inferior colliculus neurones," *Hear. Res.*, **149**, 199–215.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, **47**, 103–108.
- Carney, L. H. (1993). "A model for the responses of low-frequency auditory-nerve fibers in cat," *J. Acoust. Soc. Am.*, **93**, 401–417.
- Pflueger, M., Hoeldrich, R., and Riedler, W. (1997). "Nonlinear all-pole and one-zero gammatone filters," *Acustica*, **83**, 513–519.
- Irino, T., and Patterson, R. D. (2001). "A compressive gammachirp auditory filter for both physiological and psychophysical data," *J. Acoust. Soc. Am.*, **109**, 2008–2022.
- Evans, E. (1986). "Cochlear nerve fibre temporal discharge patterns, cochlear frequency selectivity and the dominant region for pitch," in *Auditory frequency selectivity*, edited by B. C. J. Moore and R. Patterson (Plenum Publishing Corporation), pp. 253–260.
- Patterson, R. D., Holdsworth, J., Nimmo-Smith, I., and Rice, P. (1988). "SVOS Final Report: The auditory filterbank," Tech. Rep. 2341, MRC Applied Psychology Unit, Cambridge.
- Cooke, M. P. (1991). "Modelling auditory processing and organization," Ph.D. diss., University of Sheffield.
- Darling, A. M. (1991). "Properties and implementation of the gammatone filter: a tutorial," in *Speech Hearing and Language, work in progress* (University College London, Department of Phonetics and Linguistics), pp. 43–61.
- Slaney, M. (1993). "An efficient implementation of the Patterson-Holdsworth auditory filter bank," Tech. Rep. 35, Apple Computer, <http://rvl4.ecn.purdue.edu/~malcolm/apple/tr35/>.
- Recio, A., and Rhode, W. S. (2000). "Basilar membrane responses to broadband stimuli," *J. Acoust. Soc. Am.*, **108**, 2281–2298.
- Oppenheim, A. V., Schaffer, R. W., and Buck, J. R. (1999). *Discrete-time signal processing* (Prentice Hall, Englewood Cliffs).
- Patterson, R. D. (1994). "The sound of a sinusoid: Spectral models," *J. Acoust. Soc. Am.*, **96**, 1409–1418.
- Derleth, R. P., and Dau, T. (2000). "On the role of envelope fluctuation processing in spectral masking," *J. Acoust. Soc. Am.*, **108**, 285–295.
- Oxenham, A. J., and Dau, T. (2001). "Reconciling frequency selectivity and phase effects in masking," *J. Acoust. Soc. Am.*, **110**, 1525–1538.