Auditory Filter Bank Design Using Masking Curves

L. Lin, E. Ambikairajah and W.H. Holmes

School of Electrical Engineering and Telecommunications
The University of New South Wales, UNSW SYDNEY NSW 2052, Australia
ll.lin@ee.unsw.edu.au, Ambi@ee.unsw.edu.au, H.Holmes@unsw.edu.au

Abstract

It is very difficult and costly to experimentally observe the motion of the basilar membrane in a fully functional cochlea with the view to obtaining amplitude response at points along the membrane. This paper presents an inexpensive method of generating psychoacoustic tuning curves from the well-known masking curves in critical band rate. We present a method for designing critical band auditory filters from the tuning curves. It is also known that the auditory filter frequency response becomes broader with increasing input signal levels and becomes narrower with decreasing signal levels. We also propose a method for designing level dependent auditory filters. The proposed filter bank is applicable to various types of signal processing required to model human auditory filtering.

1. Introduction

Considerable research has been done in seeking a model that accurately predicts human neural tuning curves. Earlier models have used transmission line representations to simulate basilar motion[1][4]. Recently, Gammatone filters become very popular as a reasonably accurate alternative for auditory filtering [5][6][7]. A parallel auditory filter bank that consists of Gammatone filters can be easily inverted and hence has applications in auditory-based speech and audio processing [4].

This paper presents a novel method of generating psychoacoustic tuning curves from the well-known masking curves in the critical band scale. We then propose a modelling technique to match the psychoacoustical tuning curves thus resulting in a parallel auditory filter bank with minimum-phase digital filters. Log-magnitude modelling is another technique[3] that can be used to fit the tuning curves, but it is slightly at the expense of extra zeros.

Tuning curves becomes broader with increasing input level and becomes narrower with decreasing signal input levels, hence nonlinear. We use a ‘shell’ compensation filter to model the level-dependence of the tuning curves.

The paper is organized as follows: Section 2 summarizes the method that we have used to generate psychoacoustical tuning curves. Section 3 outlines a method for designing auditory filter bank approximating the psychoacoustical tuning curves. Section 4 gives a solution to the level-dependence problem of auditory filtering.

2. Generation of Psychoacoustical Tuning Curves

Masking is usually described as the sound-pressure level of a test sound necessary to be barely audible in the presence of a masker. Using narrow-band noise of a given centre frequency and bandwidth as maskers and a pure tone as test sound, masking patterns have been obtained by Zwicker and Fastl [8] [9]. The effect of masking produced by narrow-band maskers is level dependent. Fig. 1 shows the masking patterns centred at 1 kHz.

![Figure 1: Masking curves [8]. (The dashed line is threshold in quiet)](image1)

It is known that the shapes of the masking patterns shown in Fig. 1, for different centre frequencies and different levels are very similar when plotted using the critical band rate scale. Masking curve at different centre frequencies can be obtained by simply shifting the available masking curves at \( f_c = 1 \text{kHz} \) (Fig. 2). Masking curves at levels other than \( L_G = 20, 40, 60, 80 \) and 100 dB (Fig. 1) can be generated through interpolation.

![Figure 2: Masking curves obtained by interpolation and then shifting horizontally](image2)
The tuning curves can be obtained from the masking curves by using the following method. The first step is to fix a test tone of a particular frequency at a certain level (Fig. 2). Then the masking curves with different centre frequencies that are just able to mask the testing tone are found and the corresponding levels are noted. Plotting the levels as a function of the centre frequencies provides the tuning curve at that test tone frequency (Fig. 3 and 4).

Figure 3: Tuning curves at test tone level $L_t=20\, \text{dB}$ for different centre frequencies.

Figure 4: Tuning curves at $f_c=1\, \text{kHz}$ for different test tone levels.

Figure 3 shows the tuning curves obtained for the same test tone level ($L_t=20\, \text{dB}$) but at different centre frequencies. The tuning curves obtained in this way are consistent with the measurement on cochlea in the following aspects:

1) The curves are characterized by a narrow but round peak at the centre frequency,
2) The curves are asymmetric with a very steep high-frequency slope and a flat tail extending to the lowest frequency.
3) The amplitude ratio from tip to tail is about 70 dB at higher frequency and decreases as the centre frequency decreases.

Figure 4 shows the tuning curves obtained using the above method, but at various input levels. The curves become broader with increasing signal level (nonlinearity).

The magnitude response plot (Fig. 5) of the membrane can be obtained by vertically reversing and scaling the tuning curves shown in Fig. 3. Similarly, we can obtain magnitude response curves at various input levels as shown in Fig 6. The frequency scales in Fig. 5 and 6 have been converted from critical-band rate to Hertz. We now propose a digital filter model that can approximate the magnitude response with a reasonable accuracy.

3. Auditory Filter Bank Design

The Gammatone filter [5][6][7] is well known for its closeness in modeling the psychoacoustical data. The transfer function of a 4th Gammatone filter has 8 poles and 4 zeros in s-domain [7]. The zeros are real and some of them are positive. Hence we may call it ‘all-phase Gammatone filter’. The all-pole Gammatone filter is obtained by discarding all of the zeros of the transfer function of the all-phase Gammatone filter [5].

Attempt to fit both types of filters to the basilar membrane magnitude response curves has been made. It is shown that the lower-frequency slope of the all-phase Gammatone filter is deeper than that of the magnitude response curves, but the lower-frequency slope of the all-pole Gammatone filter is shallower. This suggests that only one or two zeros placed at lower frequency should be enough to model the lower frequency slope of the tuning curves.

It is also shown that the upper-frequency slopes of these two filters are not steep enough to model the magnitude response curves, but all-pole filter is much closer to the upper-frequency part of the curve. To sharpen the upper-frequency slope, we will use a second order notch filter positioned at a
frequency higher than the centre frequency. To simulate the sharp and round peak of a magnitude response curve, at least 4 pairs of complex poles are required. Hence there will be 8 poles and 3 zeros in our model.

We develop our model directly in z-domain and the model transfer function is written as,

\[
H(z) = \frac{1 - r_0 z^{-1} + (1 - 2r_0 \cos(2\pi f_0 / f_s) z^{-1} + r_0^2 z^{-2})}{(1 - 2r_A \cos(2\pi f_A / f_s) z^{-1} + r_A^2 z^{-2})^4}
\]

(3.1)

The filter has three terms and the function of each term is discussed below:

The term \(1/(1 - 2r_A \cos(2\pi f_A / f_s) z^{-1} + r_A^2 z^{-2})^4\) will produce a narrow peak at the centre frequency \(f_c\). The parameters \(r_A\) and \(f_A\) are calculated by

\[
f_A = \sqrt{f_s^2 + B_w}, \quad r_A = e^{-2\pi f_s / f_s}
\]

(3.2)

where \(f_s\) is the sampling frequency. The bandwidth \(B_w\) in Equation 3.2 is calculated by [8]

\[
B_w = 25 + 7\sqrt{1 + 1.4(f_c / 1000)^2} \times 10^6.
\]

(3.3a)

\[
Z_c = 13\tan^{-1}(0.76 f_c / 1000) + 3.5\tan^{-1}(f_c / 7500)^2
\]

(3.3b)

where \(Z_c\) is the corresponding critical band rate of \(f_c\). The spacing of \(f_c\) is linear in critical band rate.

The term \(1 - r_0 z^{-1}\) ensures that the lower-frequency slope of the filter is similar to that of the amplitude response curve. We choose \(r_0 = 0.955\), which is very close to 1.

\[
f_{\Delta} = 117.5(f_c / 1000)^2 + 1135.5(f_c / 1000) + 277.0
\]

(3.4)

where \(f_c\) should be in Hz. The choice of \(r_0\), \(r_A\) and \(r_A\) will guarantee that the minimum phase filter is stable.

Filters at five different centre frequencies are designed using Equation 3.1 and their frequency response is plotted in Figure 7 together with the corresponding magnitude response curves obtained from the tuning curves at \(L_c = 45\) dB.

4. Level-Dependent Compensation Filter

It is also known that auditory filter shape becomes broader with increasing input signal level or narrower with decreasing signal level. To model this property, we first design a critical-band-rate filterbank that simulates the frequency magnitude response curves at 45 dB input signal level, using the method described in Section 3. We name this filterbank as the reference filterbank and each of the filters as a reference filter. Each reference filter is then cascaded with a ‘shell’ compensation filter. The parameter of the shell filter changes with the signal level so that the overall frequency characteristic is consistent with the level-dependence of the magnitude response curves. The shell filter is defined as

\[
H_s(z) = H_{tL}(z) \cdot H_{tL}(z)
\]

(4.1)

where

\[
H_{tL}(z) = \frac{1 - 2p_1 \cos(2\pi f_{\Delta} / f_s) z^{-1} + p_1^2 z^{-2}}{1 - 2p_2 \cos(2\pi f_{\Delta} / f_s) z^{-1} + p_2^2 z^{-2}}
\]

(4.2a)

\[
H_{tL}(z) = \frac{1 - 2p_1 \cos(2\pi f_{\Delta} / f_s) z^{-1} + p_1^2 z^{-2}}{1 - 2p_2 \cos(2\pi f_{\Delta} / f_s) z^{-1} + p_2^2 z^{-2}}
\]

(4.2b)

The parameters for \(H_{tL}(z)\) are chosen as

\[
p_1 = e^{-2\pi C_1 f_s / f_s},
\]

(4.3a)

\[
f_{\Delta} = (C_2 + C_L) f_s, \quad f_{\Delta} = (C_2 - C_L) f_s
\]

(4.3b)

where \(C_1\) in Equation 4.3a controls the damping ratio of \(H_{tL}(z)\). We choose \(C_1 = 0.22\) which gives a slight overshoot in the magnitude response of \(H_{tL}(z)\).

The value of \(C_2\) in Equation 4.3b will decide the frequency location of the shell filter. We choose \(C_2 = 0.82\), so that the shell filter is centered at a frequency valued about 0.82\(f_c\) and provides a small amount of centre-frequency shift with input level.

The two corner frequencies of the shell filter are placed at \(f_{\Delta} = (C_2 + C_L) f_s\) and \(f_{\Delta} = (C_2 - C_L) f_s\). They change with signal level \(L_s\) through the parameter \(C_L\). The following formula is used to calculate \(C_L\)

\[
C_L = 2.5 \times 10^{-3}(L_s - 45)
\]

(4.4)

The parameters for \(H_s(z)\) are chosen as follows:

\[
p_2 = e^{-2\pi (3C_1) / f_s},
\]

(4.5a)
The choice of these parameters makes \(H_2(z)\) more damped and provides \(H_2(z)\) with two corner frequencies even further away from the center of the shell filter.

The cascade structure of \(H_1(z)\) and \(H_2(z)\) forms a low-pass filter for input levels higher than 45 dB and a high-pass for the other levels.

The frequency response of a group of shell compensation filters designed for \(f_s = 1\) kHz is shown in Figure 8. For \(L_t = 45\) dB (signal level equals the reference level), the frequency response is 0 dB everywhere, which means no change of filter shape for the reference filters. The shell filters act as low-pass and high-pass filters depending on the input level. Their magnitude frequency response looks like “shells” when plotted in a linear frequency scale. The shell filters are always minimum-phase and stable when \(p_1 < 1\) and \(p_2 < 1\).

The frequency response of the cascade of the shell filters with the reference filter is shown in Figure 9, which matches the magnitude response curves in Figure 6.

5. Conclusions

We have provided a novel method for obtaining tuning curves from the masking curves. Also we have shown that the magnitude response of the basilar membrane at various frequency points can be obtained using these tuning curves. An efficient method for designing auditory filters has been proposed. A design procedure for obtaining level dependent auditory filters using shell filters has also been formulated in the paper. Future research will concentrate on inversion of this auditory model and its application to speech and audio processing.

6. References