ガンマチャープフィルタバンクにおける非対称性の制御方法 *

◎ 鵜木　祐史 (北陸先端大/ATR 人間情報研)　　入野　俊夫 (ATR 人間情報研)

1. はじめに

時間–スケール表現における最小不確定性を満たす関数をインパルス応答として持つガンマチャープフィルタが聴覚フィルタとして良い特性を持つことが示された [1]。しかし、このフィルタを用いて時変系の聴覚末梢系モデルを構築するにあたり、フィルタの時間応答の計算効率と音圧に依存した非対称性パラメータの制御方法が問題となった。前者については、IIR の非対称性補償フィルタをガンマトーンフィルタと組み合わせることにより、計算効率の高いガンマチャープフィルタを精度良く実現することが可能となった [2]。本稿では、後者の改善として、音圧推定の方法と非対称性の制御方法を提案し、時変系の聴覚末梢系モデルとしてガンマチャープフィルタバンクの構成法を示す。

2. 非対称性補償形ガンマチャープフィルタ

ガンマチャープフィルタの振幅周波数特性は

$$|G_c(f)| = \frac{a\Gamma(n+jc)}{|2\pi b\mathrm{ERB}(f_r)+j2\pi(f-f_r)|^2}\cdot e^{c\theta} \quad (1)$$

$$\theta = \arctan\{(f-f_r)/b\mathrm{ERB}(f_r)\} \quad (2)$$

と表される [1]。ここで、a は振幅、n と b はガンマ分布関数のパラメータ、f_r は最終的に収束する近似周波数、$\mathrm{ERB}(f_r)$ は等価矩形帯域幅、c は周波数変調項の係数で非対称性の度合を決めるパラメータを表す。また、振幅が最大になる周波数 f_{peak} は

$$f_{\mathrm{peak}} = f_r + c\cdot b\mathrm{ERB}(f_r)/n \quad (3)$$

となる。振幅を正規化して考えると式 (1) の第 1 項はガンマトーンの振幅周波数特性 $|G_T(f)|$ であり、残りの $e^{c\theta}$ が式 (3) の周波数シフトを伴った非対称性を決定している。これより、

$$|G_c(f)| = |G_T(f)|\cdot|G_A(f)| \quad (4)$$

$$|G_A(f)| = e^{c\theta} = e^{c\cdot\arctan\{(f-f_r)/b\mathrm{ERB}(f_r)\}} \quad (5)$$

と書き表すことができる。ここで、$|G_T(f)|$ については IIR フィルタで実現でき [3]、$|G_A(f)|$ については少ないフィルタ係数 (2 次の IIR を 4 段で縦続接続) で近似した非対称性補償フィルタ ($|H_C(f)|$) で実現できる [2]。この $|G_c(f)| = |G_T(f)|\cdot|H_C(f)|$ で構成されたフィルタを非対称性補償形ガンマチャープフィルタと呼び、式 (1) の特性を少ないパラメータで精度良く実現できる [2]。

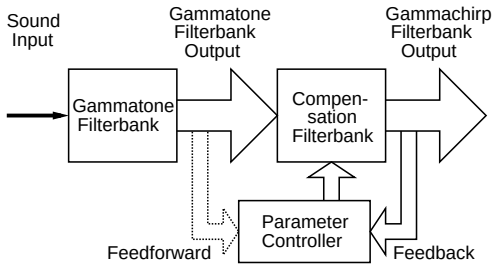


図 1 . ガンマチャープフィルタバンクの構成

3. ガンマチャープフィルタバンク

ガンマチャープフィルタバンクを構築するためには、フィルタ形状の非対称性を決定するパラメータ制御の問題を解決する必要がある。ここでは、フィルタ出力から、非対称性の度合を決める入力信号の音圧レベルを推定する方法を考える。

3.1 音圧推定と非対称性の制御方法

図 1にガンマチャープフィルタバンクのブロック図を示す。これは、ガンマトーンフィルタバンク、非対称性補償フィルタバンク、パラメータ制御系 (音圧推定とパラメータ決定) で構成される。

入力信号は、ガンマトーンフィルタバンクに入力され、この出力は非対称性補償フィルタバンクに入力される。次に、パラメータ制御系ではフィルタ出力から音圧を推定し、その結果からフィルタ形状の非対称性を決定するパラメータ c を決定する。このパラメータ c は非対称性補償フィルタバンクの各フィルタ形状を決定し、このフィルタ出力がガンマチャープフィルタバンクの出力となる。本稿では、この非対称性の制御方法として、ガンマトーンフィルタバンク出力を利用するフィードフォワード (FF) 形とガンマチャープフィルタバンク出力を利用するフィードバック (FB) 形を考える。

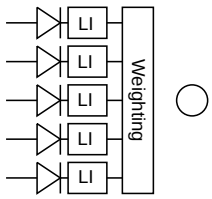
1 チャンネル分のパラメータ制御回路を図 2に示す。パラメータ制御系では、FF/FB 形に対応したフィルタバンク出力を半波整流し、漏洩積分器 (LI) を通過した信号から平均振幅を計算する。本稿では LI の時定数を 30ms とするが、フィードバック系 (FB 形) の安定性に問題は生じなかった。次に、処理対象となるチャンネルを中心に隣接するチャンネルの平均振幅は荷重平均され、そのチャンネルでの音圧推定のための平均活性度となる。ここで、この荷重関数は自由に選択でき、本稿では、 3 つの窓関数 ((i) 3ERB 幅の Hanning 窓、(ii) $c=-1$、(iii) $c=-2$

∗ A method for controlling the asymmetric parameters in the gammachirpfilterbank

By Masashi Unoki (JAIST/ATR HIP) and Toshio Irino (ATR HIP)

表 **1** . 係数 q と音圧推定誤差

control type	weighting	rms error (dB)	q
Feed Forward (FF)	Hanning	1.7	14.0
	$c=-1$	1.9	14.5
	$c=-2$	2.1	15.0
Feed Back (FB)	Hanning	2.3	16.9
	$c=-1$	3.7	16.6
	$c=-2$	4.2	16.8



に固定したガンマチャープ) を用いる。また、k 番目のチャンネルにおける推定音圧 $\hat{P}_{sk}$ は、

$$\hat{P}_{sk} = 20\log_{10}(q \cdot a_{sk}) \tag{6}$$

で決定する。ここで、係数 q は定数、a_{sk} は平均活性度である。最後に、非対称性のパラメータ推定値 $\hat{c}_k$ は、文献 [1] のレベル依存性の式から、$\hat{c}_{sk} = 3.38 - 0.107\hat{P}_{sk}$ で求める。これは k 番目の非対称性補償フィルタのパラメータとして用いられる。

3.2 係数 q の決定と推定誤差の評価

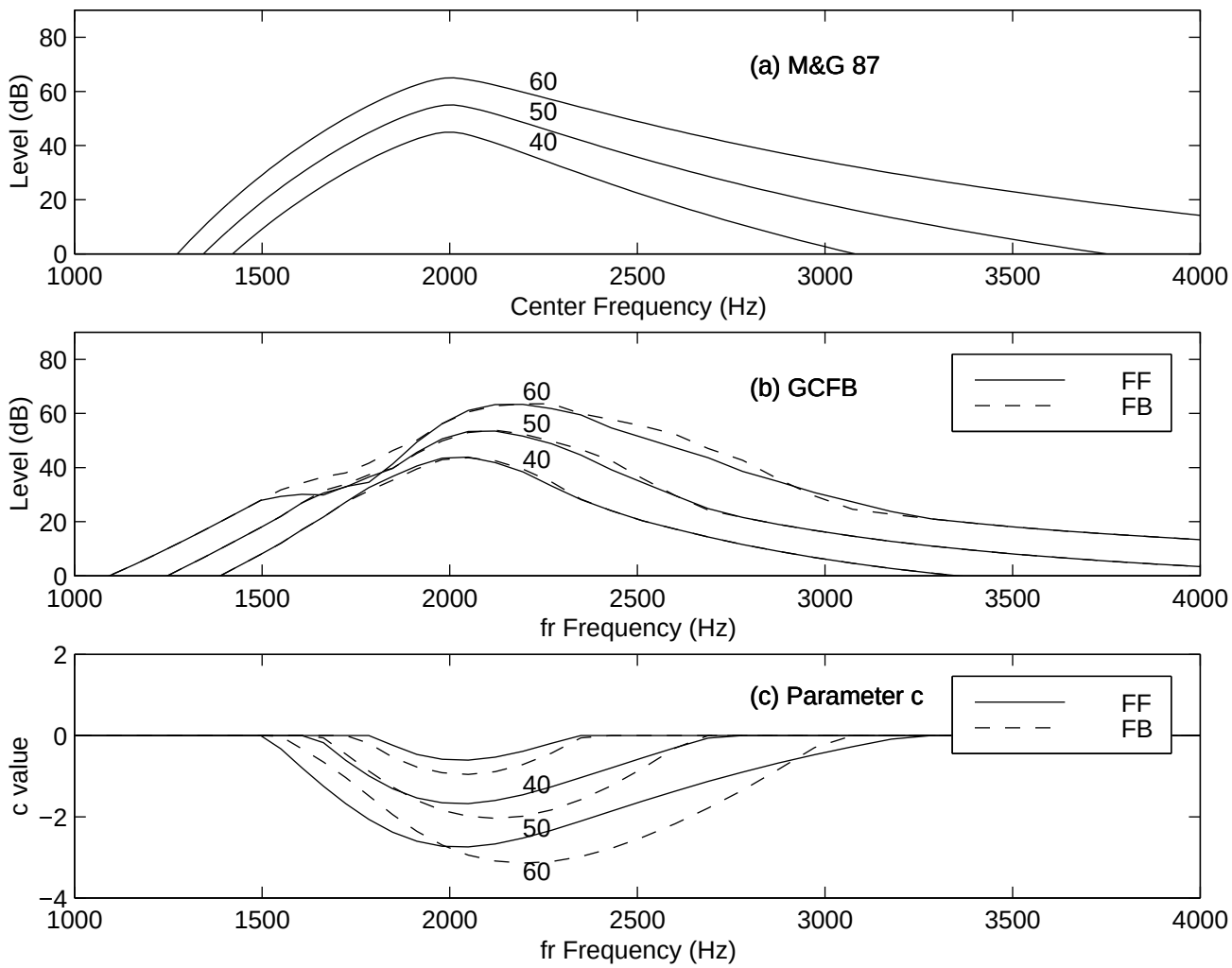
入力信号をガンマチャープフィルタのパラメータ決定で用いたノッチ雑音によるマスキングデータ (周波数 2 kHz、78 個のプローブ音と雑音の組み合わせ) とし、そのときに得られる推定音圧 $\hat{P}_s$ と入力したプローブ音圧 P_s の誤差を最小とするような係数 q を決定する。ここで、非対称なノッチ雑音の場合にも離聴聴取が起きて、SNR が最も大きいチャンネルでのフィルタ形状を決定できると仮定する。そのため、FF 形では離聴聴取に最も近いピーク周波数を持つガンマトーンフィルタのチャンネルの出力レベルが、FB 形では離聴聴取に最も近いピーク周波数を持つ過去のガンマチャープフィルタのチャンネルの出力レベルが、このチャンネルに対応する原時刻の非対称性補償フィルタの形状を決定していると仮定する。

本稿では、パラメータ制御のタイプ (FF/FB 形) と 3 つの荷重関数について、それぞれ音圧推定値の平均二乗誤差が最小となる係数 q を求めた。この結果を表 1に示す。表 1の結果から、いずれの荷重関数を用いても、q はほとんど同じ値で誤差最小となり、FF/FB 形でも概ね近い値となった。

3.3 エキサイテーションパターンの比較

ここでは、フィルタバンクの評価として、2 kHz の純音を入力した場合のエキサイテーションパターン (以後、本稿では EP と略記する) の比較を行なった。フィルタバンクの設計条件は、$n=4$、$b=1.019$、荷重関数を Hanning、周波数範囲を 200 ~ 4000 Hz、チャンネル数を 75 (ERB で等間隔)、FF のときに $q=14.0$、FB のときに $q=16.9$ とし、c の値は 0 から -3.5 の間に制限した。

図 3(a) に Moore and Glasberg(1987) の方法で算出した roex フィルタによる EP[4] を、同図 (b) にガンマチャープフィルタバンクの出力活性度を平均し


(a) M&G 87
60
50
40
Level (dB)
Center Frequency (Hz)
(b) GCFB
FF
FB
60
50
40
Level (dB)
fr Frequency (Hz)
(c) Parameter c
FF
FB
40
50
60
c value
fr Frequency (Hz)