

Title	ガンマチャープフィルタバンクの構築
Author(s)	鷓木, 祐史; 入野, 俊夫; 下平, 博
Citation	Research report (School of Information Science, Japan Advanced Institute of Science and Technology), IS-RR-98-0012P: 1-11
Issue Date	1998-03-27
Type	Technical Report
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/8383">http://hdl.handle.net/10119/8383</a>
Rights	
Description	リサーチレポート (北陸先端科学技術大学院大学情報科学研究科)

# ガンマチャープフィルタバンクの構築

鵜木 祐史 入野 俊夫 下平 博

1998年3月27日

IS-RR-98-0012P

北陸先端科学技術大学院大学

情報科学研究科

〒923-12 石川県能美郡辰口町旭台 1-1

unoki@jaist.ac.jp, irino@hip.atr.co.jp, sim@jaist.ac.jp

©M. Unoki, T. Irino and H. Shimodaira, 1998

ISSN 0918-7553

# ガンマチャープフィルタバンクの構築

鵜木 祐史<sup>†</sup> 入野 俊夫<sup>‡</sup> 下平 博<sup>†</sup>

<sup>†</sup> 北陸先端科学技術大学院大学 情報科学研究科

<sup>‡</sup> ATR 人間情報通信研究所 第一研究室

## 1 はじめに

最近、聴覚フィルタモデルとして広く使われている線形のガンマトーンフィルタ [1] に、周波数変調項を加えたガンマチャープフィルタが提案された [2, 3, 4]。このフィルタのインパルス応答を表す“ガンマチャープ関数”は、時間-スケール表現における最小不確定性の意味で最適な関数として理論的に導出されたものである。このフィルタは roex フィルタ [5] と異なり、時間応答で定義され、かつ周波数領域で非対称性をもっている。この非対称性の度合を信号の音圧レベルに依存させることにより心理物理的なデータに適合できるようになり、三つの異なる研究機関の 12 組のノッチ雑音法によるマスキングデータに良く適合させられた [6, 7]。この適合は、線形で対称なガンマトーンフィルタでは不可能であり、パラメータ数の多い基底膜振動解析モデルなどでは行なわれていなかった。また最近、基底膜振動のインパルス応答の瞬時周波数変化が報告されている [8] が、ガンマチャープの搬送波の特性はこれと整合性がある。

このようにガンマチャープフィルタは優れた特性を持っているため、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築が望まれている。このフィルタバンクを構築するにあたり、実際の聴覚末梢系で起こっていると考えられる音圧レベル依存の時変系のフィルタ処理を高い計算効率で実現させる必要がある。ガンマトーンフィルタバンクのように線形時不変のフィルタ処理であれば、比較的計算効率の高いフィルタバンク処理を行なうことができる。しかし、ガンマチャープフィルタバンクを構築する場合には、ガンマチャープフィルタが有限長インパルス応答 (FIR) 形でしか実現できていないため、フィルタバンク処理の計算効率の問題が生じる。仮に、このフィルタバンク処理を FIR フィルタで近似させた場合、各サンプル点で音圧レ

ベルに対応するインパルス応答を求め、次の時刻の出力値を計算する形となるため、計算量が膨大になるだけでなく、インパルス応答長に対し速い変化には追従できないおそれが生じる。また、時変系の処理としてフィルタ形状の非対称性が入力信号の音圧レベルに依存して時々刻々と動的に処理されなければならないが、この音圧推定をどのように推定するかという問題も生じる。

以上のことから、ガンマチャープフィルタバンクを構築するにあたり、フィルタの時間応答の計算効率とフィルタ形状の非対称性パラメータの制御方法が問題となる。

前者の問題点に対する一つの解決法として、全極形ガンマトーンフィルタ (APGF, All-pole gammatone filter) や一零形ガンマトーンフィルタ (OZGF, One-zero gammatone filter) [9, 10] を用いることが考えられる。これらは IIR フィルタで実現したガンマトーンフィルタの極の個数を変えることにより周波数非対称性を与えたものである。しかし、まだ実際のノッチ雑音法によるマスキングデータに直接適合されておらず、またサンプリング周波数や設計する周波数に特性が依存するため、制御パラメータの設定の議論は部分的にしか報告されていない [11]。そこで、本稿では、ガンマチャープフィルタを無限インパルス応答 (IIR) フィルタを用いて少ないパラメータで近似実現することで、計算効率の問題点の改善を計る。

一方、後者の問題点に対しては、まだ明確な解法は報告されていないが、入力信号の音圧をフィルタバンクのどこかの段階で推定することで解決できるものと考えられる。この音圧の推定方法については、あらかじめフーリエ変換で音圧を決定するフィルタの入力側説 [12] と非対称性をいれた roex フィルタの出力で決定する出力側説 [14] が報告されている [19]。これら二つの説に対し、聴覚フィルタの

信号処理の観点からは後者が、処理の計算量の観点では前者が望ましいと考えられる。そこで、本稿では、出力側説を支持するが、いずれの処理形式にも対応できるようにフィードフォワード/フィードバック形式の二種類の音圧推定の制御方法を提案する。

以上、フィルタバンクの構築にあたる上記二つの問題点に対して、一つの解決方法を提案し、サンプリング周波数や設計する周波数に依存せず、基本的にフィルタの帯域幅を決定する二つパラメータと非対称性を決定する一つのパラメータのみで特性を特定できるガンマチャープフィルタバンクの構成法を示す。はじめに、第2節では、フィルタの時間応答の計算効率の問題点に関して、ガンマチャープフィルタがガンマトーンフィルタとIIRの非対称性補償フィルタの組み合わせで精度良く実現できることを示す。次に、第3節では、音圧に依存した非対称性パラメータの制御方法の問題点に関して、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。次に、第4節では、ガンマチャープフィルタバンクの評価として、エキサイテーションパターンの比較を示す。最後に、第5節では、本フィルタバンクの信号処理への応用の一例を示す。

## 2 非対称性補償形ガンマチャープフィルタ

### 2.1 ガンマチャープフィルタ

ガンマトーンフィルタの複素表現のインパルス応答は、

$$g_t(t) = at^{n-1} \exp(-2\pi b \text{ERB}(f_c)t) \times \exp(j2\pi f_c t + \phi) \quad (1)$$

で定義されている。ここで、時間  $t > 0$ 、 $a$  は振幅、 $n$  と  $b$  はガンマ分布関数のパラメータ、 $f_c$  は搬送波周波数、 $\phi$  は初期位相を表す。また、 $\text{ERB}(f_c)$  は等価矩形帯域幅であり、 $\text{ERB}(f_c) = 24.7 + 0.108f_c$  Hz で表される [16]。

これに対し、ガンマチャープフィルタの複素表現のインパルス応答は、

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

で定義されている [6, 7]。ここで、時間  $t > 0$ 、 $f_r$  は最終的に収束する漸近周波数、 $c$  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、 $\ln t$  は自然対数を表す。ガンマトーンフィルタとの違いは  $jc \ln t$  の項であり、 $c = 0$  の時、ガンマチャープフィルタはガンマトーンフィルタと完全に一致する。この関数は瞬時周波数が周波数の分数で変化するFMの搬送波（チャープ信号）とガンマ分布関数の包絡線を持つことからガンマチャープと名付けられた [2]。ノッチ雑音マスキングのデータへの適合の結果、 $c$  が音圧レベルに依存するパラメータ、 $n$  と  $b$  が定数パラメータになる [2, 4]。本稿では、特に断らない限りこの定数パラメータを代表的な値  $n = 4$  と  $b = 1.68$  として使うことにする。

### 2.2 振幅周波数特性

ガンマトーンフィルタの振幅周波数特性は

$$|G_T(f)| = \frac{a\Gamma(n)}{|2\pi b \text{ERB}(f_c) + j2\pi(f - f_c)|^2} \quad (3)$$

と表される。また、ガンマチャープフィルタの振幅周波数特性は

$$|G_c(f)| = \frac{a\Gamma(n + jc)}{|2\pi b \text{ERB}(f_r) + j2\pi(f - f_r)|^2} \cdot e^{c\theta} \quad (4)$$

$$\theta = \arctan\{(f - f_r)/b \text{ERB}(f_r)\} \quad (5)$$

と表される [6]。ここで、振幅が最大になる周波数  $f_{\text{peak}}$  は

$$f_{\text{peak}} = f_r + c \cdot b \text{ERB}(f_r)/n \quad (6)$$

となる。図1に、ガンマトーンフィルタの振幅周波数特性（式(3)）と音圧レベルを60 dB SPLとしたガンマチャープフィルタの振幅周波数特性（式(4)）を示す。ここで、振幅を正規化して考えると式(4)の第1項は式(3)の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(6)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、

$$|G_c(f)| = |G_T(f)| \cdot |H_A(f)| \quad (7)$$

$$|H_A(f)| = e^{c\theta} = e^{c \cdot \arctan\{(f - f_r)/b \text{ERB}(f_r)\}} \quad (8)$$

と書き表すことができる。この結果を振幅周波数特性図で考えると、図1の特性が図2に示すように各振幅周波数特性の積になることが容易に理解できる。従って、 $n$  と  $b$  は定数であることから、ガン

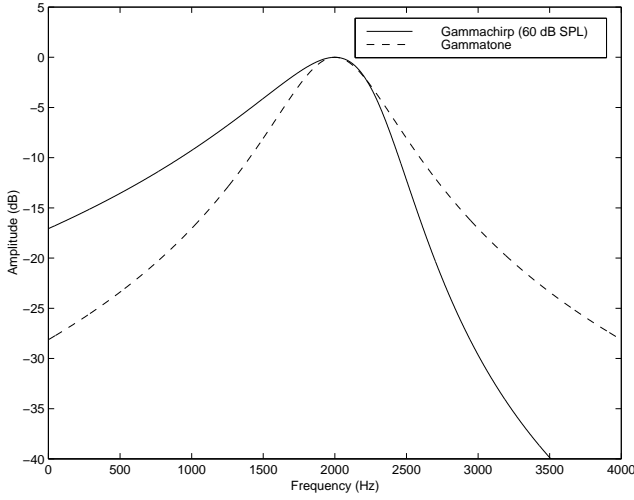


図 1: ガンマチャープフィルタ形状とガンマトーンフィルタ形状 ( $n = 4$ ,  $b = 1.68$ , 音圧レベル 60 dB SPL)

マチャープフィルタの振幅特性は、時不変のガンマトーンフィルタの振幅特性に、レベル依存する非対称性フィルタ  $H_A(f)$  を縦続接続した形となる。ここで、ガンマトーンフィルタは IIR フィルタで実現できることが報告されている [9] ため、残りの  $H_A(f)$  を少ないフィルタ係数で近似できれば IIR 形のガンマチャープフィルタを実現できる。

$f_r = 2000$  Hz の場合の  $|H_A(f)|$  の特性 (式 (8)) を図 2 に示す。この結果から、次のような特徴が言える。

- (a)  $|H_A(f)|$  の特性は、 $c = 0$  のとき全通過形、 $c > 0$  のとき高域通過形、 $c < 0$  のとき低域通過形となる。また、 $c$  の絶対値が大きくなるほど振幅の範囲が広がり  $f_r$  での傾きが大きくなる。
- (b) 任意の周波数  $f_r$  に対し次式が成立する。

$$|H_A(f_r + f_a)| = |H_A(f_r - f_a)|^{-1} \quad (9)$$

- (c) 単調変化で、凸 (ピーク) や凹 (ディップ) がない。

この特徴は、近似フィルタが満たすべき条件でもある [17]。

### 2.3 非対称性補償フィルタの実現

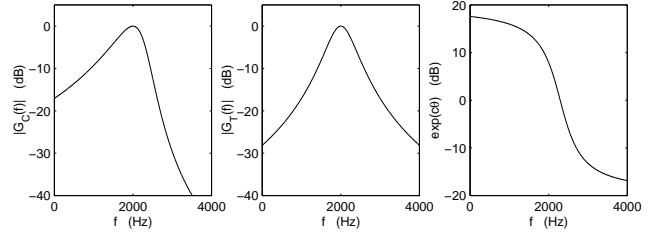


図 2: ガンマチャープフィルタの振幅周波数特性 : (左) ガンマチャープフィルタの振幅周波数特性  $|G_C(f)|$ 、(中央) ガンマトーンフィルタの振幅周波数特性  $|G_T(f)|$ 、(右) 周波数シフトの振幅周波数特性  $|H_A(f)|$

$|H_A(f)|$  を近似するフィルタで、まず条件 (b) を満たすことを考える。まず、FIR フィルタは零点しか持たないので、式 (9) を厳密には満たせない。また、Remze アルゴリズム [18] で近似度高く設計しようとすると、元のガンマチャープフィルタのインパルス応答長と同等な係数長になり、ここでの目的に合わない。これに対し、IIR フィルタで、極と零点の数が等しく、それらが適当な  $\Delta f$  に対し、 $f_r + \Delta f$  と  $f_r - \Delta f$  に相当する位置に対称的に置かれていれば満たすことができる。このことから極・零点の大きさ  $r$  は等しく、発散しないために  $|r| \leq 1$  としなければならない。また、 $r$  はフィルタの  $Q$  に関係するため、帯域幅  $b\text{ERB}(f_r)$  と負の相関を持つこと、 $\Delta f$  は条件 (a) より非対称性の  $c$  に比例して帯域幅  $b\text{ERB}(f_r)$  にも相関があることが予想される。このような特性を、デジタルフィルタで直接実現することを考えるため、縦続接続された 2 次の IIR デジタルフィルタ  $H_C(z)$  を次のように定義する。

$$H_C(z) = \prod_k H_{Ck}(z) \quad (10)$$

$$H_{Ck}(z) = \frac{(1 - r_k e^{j\varphi_k} z^{-1})(1 - r_k e^{-j\varphi_k} z^{-1})}{(1 - r_k e^{j\phi_k} z^{-1})(1 - r_k e^{-j\phi_k} z^{-1})} \quad (11)$$

$$r_k = \exp\{-k \cdot p_1 \cdot 2\pi b\text{ERB}(f_r)/f_s\} \quad (12)$$

$$\phi_k = \frac{2\pi(f_r + 2^{k-1} \cdot p_2 \cdot c \cdot b\text{ERB}(f_r))}{f_s} \quad (13)$$

$$\varphi_k = \frac{2\pi(f_r - 2^{k-1} \cdot p_2 \cdot c \cdot b\text{ERB}(f_r))}{f_s} \quad (14)$$

但し、 $p_1$  と  $p_2$  は正係数、 $f_s$  はサンプリング周波数である。ここでパラメータの値を変化させたフィル

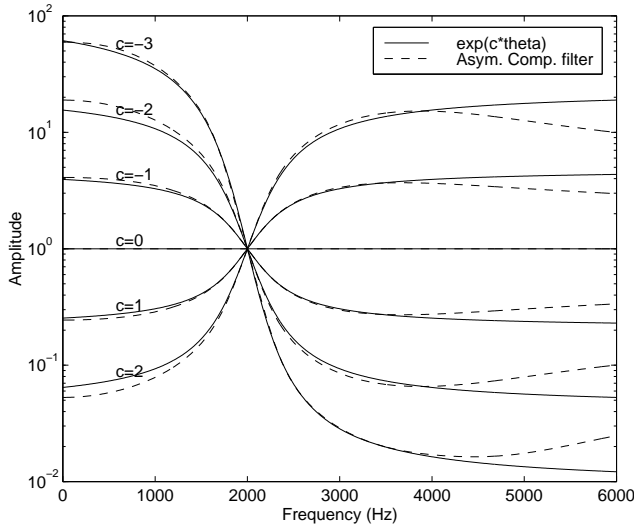


図 3:  $|H_A(f)| = e^{c\theta}$  の振幅スペクトルと非対称性補償フィルタ  $|H_C(f)|$  ( $n = 4, b = 1.68, f_r = 2000$  Hz, 振幅は  $f = f_r$  で正規化)

を縦続接続する理由は、条件 (a) での周波数  $f_r$  での傾きを十分に保ちながら、ピークやディップを打ち消し、条件 (c) を満たすことにある。

縦続接続数を 4 段とした場合の振幅周波数特性  $|H_C(f)|$  を、図 3 に破線で示す。但し、振幅特性は  $f = f_r$  での値で正規化されている。周波数  $f_r$  の 2 倍 (4000 Hz) 以下では近似すべき  $|H_A(f)|$  の特性と非常によく一致し、それ以上では誤差が大きくなる。縦続接続数を増やすことにより、さらに近似誤差の改善に期待できるが、パラメータ数とのトレードオフで決めることになるため、ここでは 4 次以上の多段縦続接続は考慮しない。次節では、ガンマチャープフィルタを構成するために、この 2 次フィルタの 4 段縦続接続で十分近似度が高いことを示す。この極と零、各々 8 係数の IIR フィルタを非対称性補償フィルタ (Asymmetric Compensation Filter) と呼ぶことにする [17]。

## 2.4 振幅周波数特性の近似度

この非対称性補償フィルタにガンマトーンフィルタを縦続接続すれば、ガンマチャープフィルタを近似できる。振幅特性は、式 (7) の非対称性項  $|H_A(f)|$  を非対称性補償フィルタの振幅特性  $|H_C(f)|$  で置き換え、

$$|G_{CAC}(f)| = |G_T(f)| \cdot |H_C(f)| \quad (15)$$

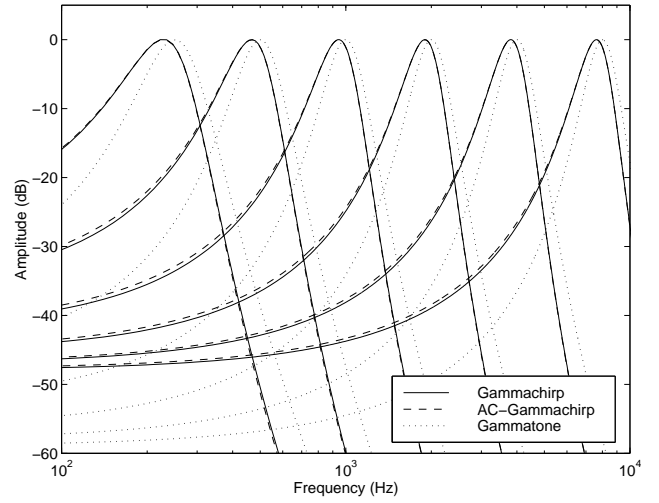


図 4: オリジナルのガンマチャープフィルタ (実線)、非対称性補償形ガンマチャープフィルタ (破線)、ガンマトーンフィルタ (点線) の振幅スペクトル ( $n = 4, b = 1.68, c = -1, f_r = 250, 500, 1000, 2000, 4000, 8000$  Hz)

で表される。以下は簡単のため、このような構成のガンマチャープフィルタを非対称性補償形ガンマチャープフィルタ (Asymmetric Compensation Gammachirp filter, AC-Gammachirp filter) と呼び、式 (2) の時間応答で定義されるガンマチャープフィルタとは区別する [17]。

図 4 に、例としてこの振幅周波数特性  $|G_{CAC}(f)|$  を破線で、式 (4) のガンマチャープフィルタの振幅周波数特性  $|G_c(f)|$  を実線で示す。また、参考のために、式 (3) のガンマトーンフィルタの振幅周波数特性  $|G_T(f)|$  を図 3 に点線で示す。ここで、補償フィルタをかける前後でピーク周波数での振幅を同じにするため、振幅特性  $|H_C(f)|$  を式 (6) 中のピークシフト量に比例した周波数での値で正規化する。この周波数は、第  $k$  番目の 2 次フィルタで、

$$f = f_r + k \cdot p_3 \cdot c \cdot b \text{ERB}(f_r)/n \quad (16)$$

で与えられる。但し、 $p_3$  は定数である。また、 $c = -1$  以外の条件も勘定して式 (11)、式 (12)、式 (16) 中の定数  $p_1, p_2, p_3$  を発見的に

$$p_1 = 1.35 - 0.19 \cdot |c| \quad (17)$$

$$p_2 = 0.29 - 0.0040 \cdot |c| \quad (18)$$

$$p_3 = 0.23 + 0.0072 \cdot |c| \quad (19)$$

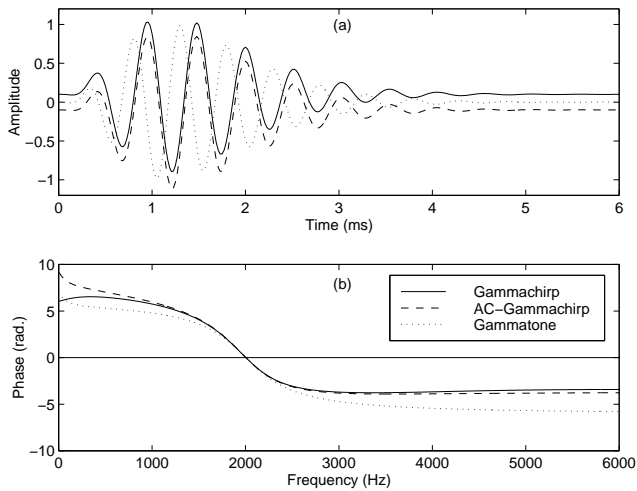


図 5: インパルス応答と位相スペクトル：(実線) オリジナルのガンマチャープフィルタ、(破線) 非対称性補償形ガンマチャープフィルタ、(点線) ガンマトーンフィルタ ( $n = 4$ ,  $b = 1.68$ ,  $c = -1$ ,  $f_r = 2000$  Hz)

と置いた。

図 4 で、 $|G_c(f)| > -50$  dB の範囲での最大実効誤差は  $f_r = 500$  Hz の時で 0.41 dB である。また、 $b$  を 3 セット (1.0, 1.35, 1.7)、 $c$  を 5 セット (1.0, 0, -1.0, -2.0, -3.0)、周波数  $f_r$  を 6 セット (250, 500, 1000, 2000, 4000, 8000 Hz) とし、組み合わせた 90 個のパラメータセットに対し、 $|H_c(f)| > -50$  dB の範囲での平均実効誤差は約 0.63 dB であり、実効誤差が 2 dB を超えるものは  $f_r$  が 8000 Hz で  $c$  が -3.0 の 3 組のみであった。このように、非対称性補償形ガンマチャープフィルタでももとのガンマチャープフィルタを精度良く近似することができる。

更に、式 (17)~(19) の定数の取り方を最小自乗の意味で最適化を試みたが、あまり改善が見られなかった。但し、式 (11)、式 (12)、式 (16) の第  $k$  番目の極や零点を決める係数の取り方を変えれば更に改善できる可能性はある。

## 2.5 時間応答波形と周波數位相

図 5(a) に、非対称性補償形ガンマチャープフィルタの時間応答波形 (破線) と式 (2) のガンマチャープフィルタ (実線) の時間応答波形を示す。この図

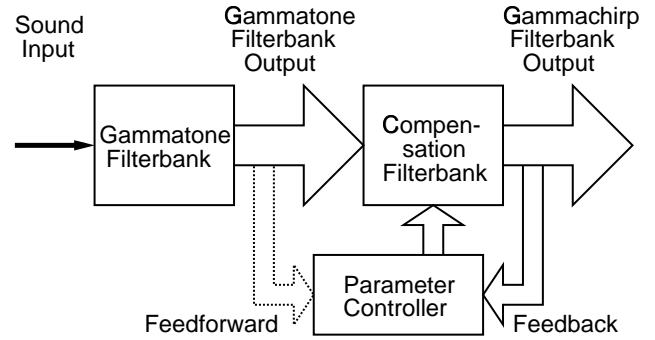


図 6: ガンマチャープフィルタバンクの構成

から、 $c = -1$  で応答波形はほとんど違いがないことがわかる。この図での差は、実効振幅で 0.0038 (最大振幅に対して約 -48 dB) 程度である。ここで、初段のガンマトーンフィルタの初期位相が両者が一致するように合わせ込んでいる。一致する時の初期位相は  $c$  の値によって線形に変化する。式 (2) で適当な定数  $t'$  に対し、

$$jc \ln t = jc \ln(t/t') + jc \ln t' \quad (20)$$

が成立するため、もともとのガンマチャープフィルタでも同様に初期位相は  $c$  の一次関数となり、両者の違いはこの定数  $t'$  の値の違いだけとなる。

図 5(b) に、非対称性補償形ガンマチャープフィルタの位相周波数応答 (破線) とガンマチャープフィルタの位相周波数応答 (実線) を示す。 $f = f_r$  の位相を零として表示している。振幅周波数応答の場合と同様、中心周波数  $f_r$  の上下 2 ERB (= 480 Hz) 以内はもちろん、周波数零に近い場所以外では概ね一致している。

## 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、非対称性の度合い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 6 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

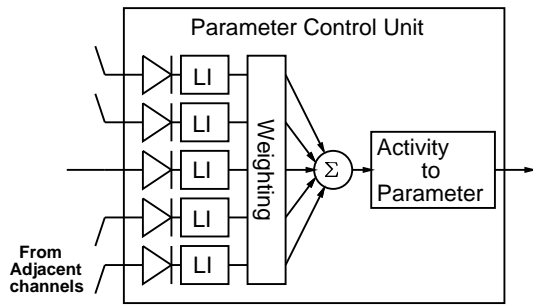


図 7: パラメータ制御回路

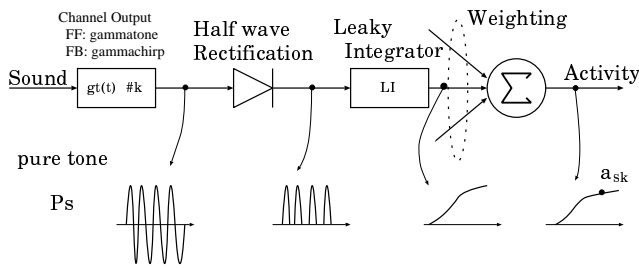


図 8: パラメータ制御回路における信号処理の流れ

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。

信号音は、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。

非対称性の制御方法としては、ガンマトーンフィルタバンク出力/ガンマチャープフィルタバンク出力を利用すフィードフォワード形/フィードバック形や、両者を適当な割合で利用するハイブリット形などが考えられる。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) 形の二種類の場合について検討する。

1 チャンネル分のパラメータ制御回路を図 7 に示す。また、図 7 における細部の信号処理の流れを図

8 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: Leaky Integrator) 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 30 ms 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化度  $a_{sk}$  となる。ここで、この荷重関数は自由に選択でき、本稿では、三つの窓関数

(i) 3 ERB 幅の Hanning 窓

(ii)  $c = -1$  に固定したガンマチャープフィルタ

(iii)  $c = -2$  に固定したガンマチャープフィルタ

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、

$$\hat{P}_{sk} = 20 \log_{10}(q \cdot a_{sk}) \quad (21)$$

で決定する。ここで、係数  $q$  は定数、 $a_{sk}$  は平均活性化度である。最後に、非対称性のパラメータ推定値  $\hat{c}_k$  は、文献 [7] のレベル依存性の式から、

$$\hat{c}_{sk} = 3.38 - 0.107 \hat{P}_{sk} \quad (22)$$

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

聴覚フィルタ形状はマスキングのパワースpektralモデル [15] を基にして心理物理的に決定できる。入野らによるガンマチャープフィルタの代表的なパラメータ [7] は、周波数 2000 Hz のプローブ音を用いたノッチ雑音によるマスキングデータ [14] に適合させて決定した。ここでは、同様のガンマチャープフィルタのパラメータ値を用いて係数  $q$  を決定するために、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, 500 ms, 78 個のプローブ音と雑音の組み合わせ) とする。この 78 個の入力信号に対し、ガンマチャープフィルタバンクで得られる推定音圧  $\hat{P}_s$  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。

聴覚フィルタの形状の場合、非対称なノッチ雑音の場合にも離聴聴取 (off-frequency listening) が起



表 1: 係数  $q$  と音圧推定誤差

control type	weighting	rms error (dB)	$q$
Feed Back (FB)	Hanning	1.9	6.8
	$c = -1$	3.4	5.0
	$c = -2$	3.7	4.8
Feed Forward (FF)	Hanning	2.5	7.4
	$c = -1$	3.6	5.6
	$c = -2$	3.7	5.7

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 [5] と仮定している。この最大 SNR を与えるフィルタの  $f_r$  や  $f_{\text{peak}}$  はこの適合プログラムで決定できる。FF 形では最大 SNR フィルタの  $f_r$  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 形では最大 SNR フィルタの  $f_{\text{peak}}$  に最も近いピーク周波数を持つ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) と三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。ここで、ガンマチャープフィルタバンクの設計条件は、 $n = 4$ ,  $b = 1.68$ , チャンネル数を 31 (2000 Hz を中心に 0.4 ERB 間隔に 31 個のフィルタが存在) とした。また、非線形処理のため、音圧レベル 60 dB SPL の正弦波信号を計算機上では振幅値  $200\sqrt{2}$  (実効値 200) に対応づけることにした。また、ガンマチャープフィルタの適合のときと同様に、刺激音を等ラウドネス曲線 (ELC) の振幅特性を近似した直線位相 FIR フィルタに通し、フィルタバンク出力の振幅が十分に安定した後の平均活性度 (80~150 ms) から式 (21) に基づいて音圧レベルを推定した。

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、係数  $q$  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。また、推定誤差は、Hanning 窓、 $c = -1$ 、 $c = -2$  に固定したガンマチャープフィルタの順に小さい結果となった。

### 3.3 音圧推定誤差に関する考察

表 1 の結果、誤差が最小となる係数  $q$  を求めることができたが、この値がチャンネル数や音圧の定義値の変化に影響を受ける可能性も残っている。フィ

表 2: チャンネル数の変化による係数  $q$  と音圧推定誤差への影響

Type	チャンネル数	フィルタ間隔	rms error (dB)	$q$
FB	61	0.2 ERB	1.7	7.2
	41	0.3 ERB	1.8	7.1
	31	0.4 ERB	1.9	6.8
	21	0.6 ERB	2.0	6.4
	11	1.2 ERB	2.4	6.0
FF	61	0.2 ERB	2.2	7.9
	41	0.3 ERB	2.4	7.6
	31	0.4 ERB	2.5	7.4
	21	0.6 ERB	2.7	7.3
	11	1.2 ERB	3.6	6.8

ルタバンクの設計条件は任意に変更可能であるため、この設計変更に伴う影響が係数  $q$  に現れることは、フィルタバンクの構造上問題となる。

そこで、はじめにフィルタバンクのチャンネル数を変動させ、前節と同様の方法で係数  $q$  を決定することで、チャンネル数による係数  $q$  への影響を調べる。ここでは、荷重関数を Hanning 窓のみとし、チャンネル間隔を 0.2, 0.3, 0.4, 0.6, 1.2 ERB としたときの音圧推定誤差が最小となるときの係数  $q$  を求めた。この結果を表 2 に示す。表 2 の結果から、係数  $q$  を決定した音圧推定誤差の最小値とチャンネル間隔との間には、トレードオフの関係があることがわかる。つまり、チャンネル間隔を密にすると推定誤差が小さくなり、チャンネル間隔を疎にすると推定誤差が大きくなるという傾向がある。しかし、この推定誤差の変動はさほど大きいものではないため、精密な係数推定を考えない限り、このチャンネル数の変動による係数  $q$  への影響を無視できるものと思われる。

次に、音圧レベルの定義値の変化による係数  $q$  への影響を調べる。本稿では、音圧レベルが 60 dB SPL の正弦波信号を計算機上では振幅値  $200\sqrt{2}$  (実効値 200) と定義している (40 dB SPL の正弦波信号の振幅値は  $20\sqrt{2}$ 、50 dB SPL の正弦波信号の振幅値は  $63.2\sqrt{2}$  に対応)。そこで、計算機上で利用する音圧レベルを

- (1) 60 dB SPL の正弦波信号の振幅値を  $100\sqrt{2}$ 、40 dB SPL の正弦波信号の振幅値を  $10\sqrt{2}$
- (2) 60 dB SPL の正弦波信号の振幅値を  $400\sqrt{2}$ 、40 dB SPL の正弦波信号の振幅値を  $40\sqrt{2}$

表 3: 係数  $q$  と音圧推定誤差 (条件 (1))

control type	weighting	rms error (dB)	$q$
Feed Back (FB)	Hanning	1.9	13.5
	$c = -1$	3.4	9.8
	$c = -2$	3.7	9.5
Feed Forward (FF)	Hanning	2.5	15.0
	$c = -1$	3.6	11.0
	$c = -2$	3.7	11.5

表 4: 係数  $q$  と音圧推定誤差 (条件 (2))

control type	weighting	rms error (dB)	$q$
Feed Back (FB)	Hanning	1.9	3.3
	$c = -1$	3.4	2.4
	$c = -2$	3.7	2.4
Feed Forward (FF)	Hanning	2.5	3.8
	$c = -1$	3.6	2.8
	$c = -2$	3.7	2.8

としたときの、音圧推定誤差が最小となる係数  $q$  を求めた。この結果をそれぞれ表 3,4に示す。これらの結果、音圧レベルの定義値の変化と係数  $q$  の変化の間には反比例の関係があることがわかる。具体的には、条件 (1) のとき、音圧レベルの定義値が 1/2 倍になったのに対し、得られた係数  $q$  は約 2 倍に増加した。また、条件 (2) のとき、音圧レベルの定義値が 2 倍になったのに対し、得られた係数  $q$  は約半分に減少した。この結果、音圧レベルの定義値の変化による係数  $q$  への影響は、単に反比例の関係に過ぎず、定義値の変化から逆に係数  $q$  の推定が容易になることがわかる。

以上の考察の結果、フィルタバンクのチャンネル数や音圧レベルの定義値の変化による係数  $q$  への問題となる影響はなく、任意のフィルタバンク設計に影響を受けずに係数  $q$  のみで非対称性を制御することができる。

## 4 フィルタバンクの評価

ガンマチャープフィルタバンクの評価には、エキサイテーションパターンの比較を用いる。本稿では、純音と調波複合音の二種類の入力信号に対して、Mooreら [13] の方法で算出した roex フィルタによるエキサイテーションパターンとガンマチャープフィルタバンク出力のエキサイテーションパター

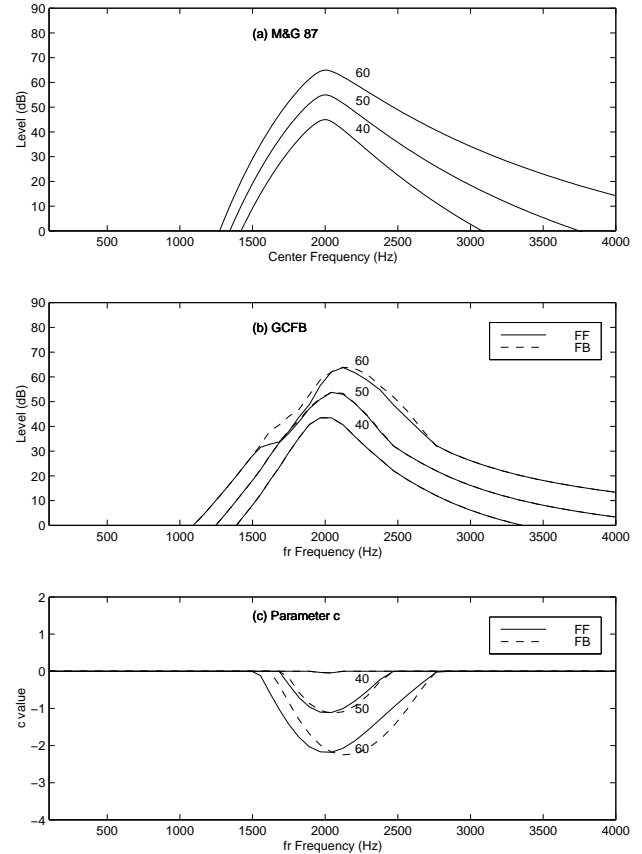


図 9: 純音に対するエキサイテーションパターン: (a)Moore らによるエキサイテーションパターン、(b)ガンマチャープフィルタバンクによるエキサイテーションパターン、(c)非対称性パラメータ  $c$  の特性

ン (平均活性度) を比較する。

### 4.1 純音に対するエキサイテーションパターンの比較

ここでは、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  $q = 7.4$ 、FB のときに  $q = 6.8$  とし、 $c$  の値は 0 から  $-3.5$  の間に制限した。

図 9(a) に Moore and Glasberg(1987) の方法で算出した roex フィルタによるエキサイテーションパターン [13] 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たエキサイテーショ

ンパターンを示す。図 9(b) は同図 (a) と完全に一致しないが、Moore らの結果に類似した結果が得られた。エキサイテーションパターンのピーク周波数は、レベルとともに上昇しているように見えるが、この現象は横軸が  $f_r$  であり、 $c$  によって  $f_{\text{peak}}$  がずれることに起因している。また、図 9(b) の結果から次のことがわかる。

1. 約 1500 Hz 以下と 3500 Hz 以上では活性度の増加度 (growth) は入力音圧に比例する。
2. 2000 Hz 付近のピーク周波数以下では、増加度が 40 dB 以下で比例、それ以上で抑制に変わる。
3. 2000 Hz 以上では、増加度が比例から拡大に変わる。

これは、図 9(b) の 2000 Hz 以下で増加量が比例し、それ以上では拡大的な特性と異なる特性である。

#### 4.2 複合音に対するエキサイテーションパターンの比較

ここでは、基本周波数が 200 Hz、高調波の次数が 12 次の調波複合音を入力した場合のエキサイテーションパターンの比較を行なった。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 100~4000 Hz、チャンネル数を 75 (ERB で等間隔)、FF のときに  $q = 7.4$ 、FB のときに  $q = 6.8$  とし、 $c$  の値は 0 から -3.5 の間に制限した。

図 10(a) に Moore and Glasberg(1987) の方法で算出した roex フィルタによるエキサイテーションパターン [13] 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たエキサイテーションパターンを示す。図 10(b) は同図 (a) と完全に一致しないが、Moore らの結果に類似した結果が得られた。また、エキサイテーションパターンの山に対し、先の結果と同様、高域側に活性度の増加が見られる。

以上、二種類の信号に対するエキサイテーションパターンの比較を行なったが、いずれも Moore らの結果に類似した結果が得られた。この結果、静的なエキサイテーションパターンに限らず、本フィルタバンクでは正しくパラメータ推定とその制御ができれば、人間の聴覚の動特性をも考慮した信号処理が可能になると思われる。

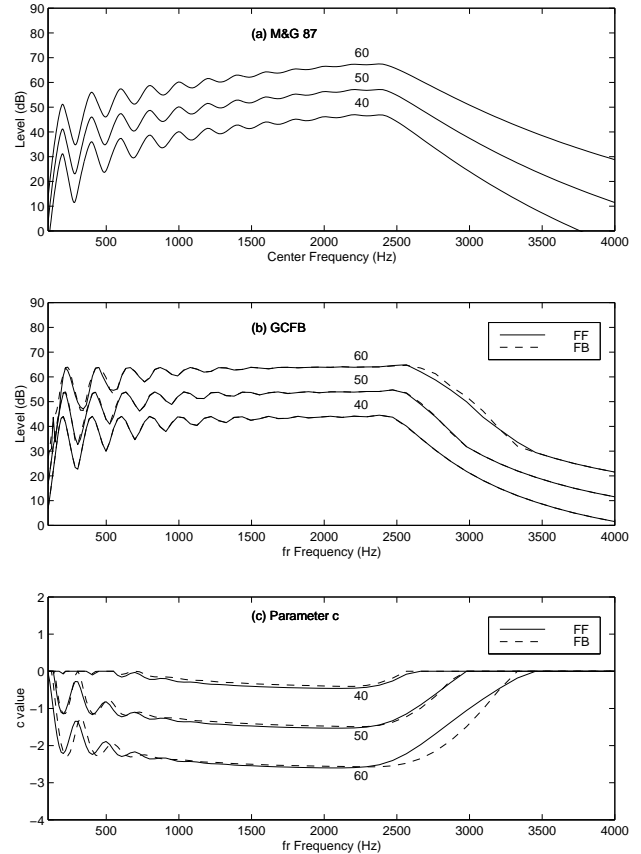


図 10: 調波複合音に対するエキサイテーションパターン: (a)Moore らによるエキサイテーションパターン、(b) ガンマチャープフィルタバンクによるエキサイテーションパターン、(c) 非対称性パラメータ  $c$  の特性

## 5 ガンマチャープフィルタバンクの工学的応用の一例

ここでの構成のガンマチャープフィルタバンクの最も重要な特性の一つは、図 11に示すような分析合成系を構成できることである [21]。入力された信号は、線形のガンマトーンフィルタバンクで周波数分析される。これが制御回路からのパラメータ値に基づいた非対称補償フィルタバンクに入力されてガンマチャープフィルタバンクの出力となる。ここまでは、図 11と同じである。この出力を何も変形操作を行わずにそのまま各チャンネルごとに、安定性が保証されている非対称性保証フィルタの逆フィルタに通せば、もとのガンマトーンフィルタバンク出力を数値演算精度で復元できる。この時の逆フィルタには分析時の順フィルタと同じパラメータを使

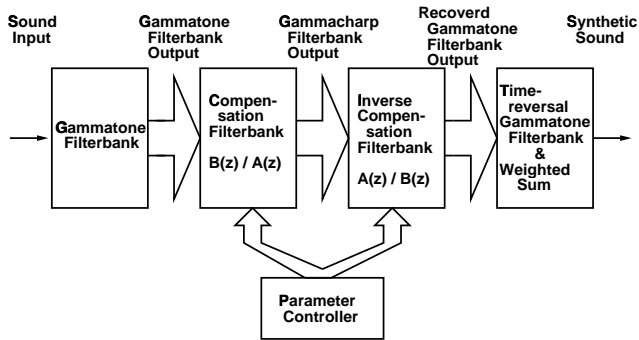


図 11: ガンマチャープ分析合成系のブロックダイアグラム

う必要がある。その後、wavelet フィルタバンクで行なわれているように [20]、ガンマトーンフィルタバンク出力を各チャンネルごとにガンマトーンの時間反転フィルタをかけて位相等化を行い、重付け和を取れば信号を再合成できる。全体の分析合成誤差は、この線形のガンマトーンフィルタバンク分析合成系の誤差となる。すなわち、このことは人間の聴覚末梢系の特性を良く近似している時変系の聴覚フィルタバンクにおいて、精度が保証された形で合成系を構成できることを示している。更に、ガンマチャープフィルタバンク出力を変形操作することによって、振幅・位相情報が欠落しない限りにおいて、変形された音がこの系だけで合成できることも示している。

## 6 まとめ

本稿では、ガンマトーンフィルタと IIR フィルタの非対称補償フィルタを組み合わせることにより、従来 FIR フィルタでしか構成できなかったガンマチャープフィルタの特性を精度良く実現できることを示した。また、時変系の聴覚末梢系を模擬するために、音圧推定と非対称性パラメータの制御法を提案し、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、パラメータ制御回路からなるガンマチャープフィルタバンクの構成法を示した。従来の聴覚モデルでは達成できなかった合成系を実現できることは、本方式の最も重要な特徴であり、人間の聴覚特性を考慮した周波数分析が必要となる様々な信号処理に応用可能である。

また、生理学的な測定で、聴覚末梢系の基底膜の動特性について知見が得られれば、伝送線路モデル

よりパラメータ数が少ない分、より明確な形で与えて模擬することが可能になった。聴覚末梢系自体を考察する際にも重要なモデルと考えられる。尚、動的パラメータの推定は今後の課題であり、よい心理物理実験の考案が待ち望まれる。

## 参考文献

- [1] Patterson, R. D., Allerhand, M., and Giguere, C., "Time-domain modelling of peripheral auditory processing: a modular architecture and a software platform," J Acoust. Soc. Am., 98, pp. 1890-1894, 1995.
- [2] 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.
- [3] 入野 俊夫, "聴覚末梢系の計算理論," 信学技報, SP95-40, July 1995.
- [4] Irino, T., "A 'gammachirp' function as an optimal auditory filter with the Mellin transform," in IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP-96), 2, pp. 981-984, Atlanta, GA., 1996.
- [5] Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R., "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am., 72, 17880-1803, 1982.
- [6] 入野 俊夫, Roy D. Patterson, "レベル依存聴覚フィルタとしてのガンマチャープ," 音響学会聴覚研資, H-96-73, 1997.
- [7] T. Irino and Patterson R. D., "A time-domain level dependent auditory filter: The gammachirp," J. Acoust. Soc. Am. 101, pp. 412-419, 1997.
- [8] de Boer, E. and Nuttall, A. L., "The mechanical waveform of the vasillar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions," J. Acoust. Soc. Am. 101, pp. 3583-3592, 1997.
- [9] Slaney M., "An efficient implementation of the Patterson-Holdsworth auditory filter bank," Apple Computer Technical Report #35.
- [10] Lyon, R. F., "The all-pole gammatone filter and auditory models," in Forum Acusticum'96, Antwerp, Belgium.
- [11] Pflueger, M. Hoeldrich, R. and Reidler, W., Personal communication, 1997.

- [12] Moore, B. C. J., Peters, R. W., and Glasberg, B. R., "Auditory filter shapes at low center frequencies," *J. Acoust. Soc. Am.*, 88, pp. 132-140, 1990.
- [13] Moore B. C. J. and Glasberg B. R., "Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns," *Hear. Res.*, 28, 209-225, 1987.
- [14] Rosen, S. and Baker, R. J., "Characterizing auditory filter nonlinearity," *Hear. Res.*, 73, 231-243.
- [15] Fletcher, H., "Auditory patterns," *Rev. Mod. Phys.*, 12, 47-65.
- [16] Glasberg, B. R. and Moore, B. C. J., "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, 47, pp. 103-138, 1990.
- [17] 入野 俊夫, 鶴木 祐史, "ガンマチャープフィルタとフィルタバンクの効率的な構成," 音響学会聴覚研資, H-97-69, Oct. 1997.
- [18] McClellan, J. H., Parks, T. W. and Rabiner, L. R., "FIR linear phase filter design program," in "Programs for Digital Signal Processing" Ed. IEEE, IEEE press, ch. 5.1, 1979.
- [19] Moore, B. C. J., "Frequency analysis and masking," in *Hearing*, edited by B. C. J. Moore, Academic Press, London, 1995.
- [20] Combes, J.M., Grossmann, A., and Tchamitchian, Ph. Esd., "Wavelets," Springer-Verlag, Berlin, 1989.
- [21] Irino, T. and Unoki, M., "A TIME-VARYING, ANALYSIS/SYNTHESIS AUDITORY FILTER-BANK USING THE GAMMACHIRP," in *IEEE Int. Conf. on Acoustics, Speech and Signal processing (ICASSP'98)*, 1998.