

[レクチャー講演] 聴覚フィルタの測定と定式化について

入野 俊夫

和歌山大学システム工学部 〒640-8510 和歌山市栄谷 930 番地

E-mail: irino@sys.wakayama-u.ac.jp

あらまし 音声や音響信号に対する人間の知覚特性をモデル化するためには、聴覚末梢系の特性を正確に測定し定式化することが第一歩として重要である。本資料では、聴覚フィルタの概念から始め、心理物理学的に帯域幅・振幅周波数特性・圧縮特性を測定する方法について述べる。さらに、その結果を定式化するための聴覚フィルタ関数としてのroexフィルタから、時間応答のあるガンマトーンフィルタやガンマチャープフィルタに至る研究の推移について概説を行う。

キーワード 聴覚フィルタ、聴覚末梢系、周波数特性、同時マスクング、順向性マスクング、roexフィルタ、ガンマトーン、ガンマチャープ

[Invited Lecture] Measurement and formulation of the auditory filter

Toshio IRINO

Faculty of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama, 640-8510, Japan

E-mail: irino@sys.wakayama-u.ac.jp

Abstract The measurement and estimation of frequency selectivity (filter shape) and compression characteristics of the auditory filters are essential to develop perceptual models for individual listeners. This paper introduces an idea of auditory filter and describes psychoacoustic methods for measuring the bandwidth, filter shape, and compression. It also describes the filter functions for fitting the psychoacoustic data and shows the overview of the temporal models from gammatone to gammachirp.

Keywords auditory filter, auditory periphery, frequency characteristics, simultaneous masking, forward masking, roex filter, gammatone, gammachirp

1. はじめに

私たちは、様々な音に囲まれて日常生活を送っている。音声の聞き取りや音源の種類の違いなど、音の解釈は脳で行われている。この前段階として、外界の音響振動を神経発火に変換する感覚系(聴覚末梢系)とその情報を中継する聴覚経路が存在する。感覚系で得られた情報をもとに音に対応する脳内表現が構築され、その表現をもとに外界の解釈が行われる。音の感覚器は脊椎動物で広く確認されており、哺乳類の場合、内耳の蝸牛がその役割を担っている。この蝸牛には神経発火変換機能ばかりでなく、周波数分析機能も付加されている。この周波数分析機能を言い表す概念として「聴覚フィルタ」という用語が用いられている[1, 2, 3]。

蝸牛の研究に関しては、動物実験による生理学的な知見の積み重ねがある。しかし、最も特性を知りたい人間に対して侵襲的な生理実験を行うことはできない

ため、実験心理的な手法が考案されてきた。また、心理実験結果が次の生理実験の基盤となっている場合もある。この意味では、生理実験と心理実験、さらにモデルを対比させながら調べるのが、科学の進歩の仕方を学ぶ上でも最も興味深く勉強になる。しかし、すべてを網羅的に紹介できないため、ここでは聴覚フィルタの特性を測定する心理物理実験と、それを定式化するフィルタ関数/モデルに関して紹介する。

2. 聴覚フィルタ特性の測定

聴覚フィルタを線形と仮定することから特性の推定は始められた。この仮定が厳密でないことは、基底膜振動の生理実験結果からも明らかである。しかし、ある条件下ではおおむね線形性を満足し、妥当と見られる結果が得られるため広く受け入れている。まず、フィルタの帯域幅から振幅周波数特性が測定の対象と

なった。さらに、非線形性の測定も行われるようになった。特に、外界の音圧に対する振動のレベル依存性は、例えば補聴器信号処理等を考える上で重要である。

2.1. 帯域幅の測定：臨界帯域

Fletcher[1]は、聴覚フィルタの「臨界帯域(critical band)」測定を行った。この実験では、帯域雑音(マスキング音)を純音(プローブ音)と同時に再生し、純音のマスキング閾値を測定する。この帯域雑音の周波数帯域幅を増加させていくと、ある幅以上になってもマスキング閾値が変化しなくなる。この臨界点が、特定の周波数成分だけを抽出する聴覚フィルタの帯域幅に相当するという論理である。

さらに、マスキングとプローブの関係を定量化するために、「マスキングのポワースペクトルモデル」を提案している。マスキングと聴覚フィルタのポワースペクトルを $N(f)$ と $W(f)$ 、検出器の効率を定数 K とすると、プローブレベル P_s は以下の式(1)で良く予測できる。

$$P_s = K \int_{-\infty}^{\infty} N(f)W(f)df \quad (1)$$

この式は、同時マスキングの実験結果を良く説明でき、以降行われた聴覚フィルタ特性推定の基本となっている。モデルとして非常に簡単で便利であるが、多くの仮定を伴うため実験条件により成立しない場合もあることを心にとめておく必要がある。

式(1)が十分信頼できるとしても、雑音帯域を拡張させる実験手法には問題点がある。帯域幅を実験中に増減させると、雑音の統計的性質(平均ばかりでなく分散も)がそれに伴い変化してしまい、信号検出理論から計算される閾値との関係が複雑になる。このため、実質的なダイナミックレンジが高々5 dB程度で、フィルタの中心周波数周辺しか推定できないと指摘されている[4]。

2.2. 振幅周波数特性/フィルタ形状の測定

上記の問題点を克服し、フィルタの振幅周波数特性(フィルタ形状)をより広いダイナミックレンジで推定するための手法として、ノッチ雑音法が提案された[2-7]。この手法では、図1に示すように、プローブ音を2つの帯域雑音で挟むように配置する。中心に雑音が無いノッチ領域があるため、ノッチ雑音法と呼ばれている。帯域雑音の幅を一定とすれば、分散が変化することはなく、式(1)をパワーだけの関数としてそのまま使用できる。この時、雑音のポワースペクトル $N(f)$ が N_0 で一定だとすると、

$$P_s = K \cdot \left(\int_{f_{l,min}}^{f_{l,max}} N_0 \cdot W(f)df + \int_{f_{u,min}}^{f_{u,max}} N_0 \cdot W(f)df \right) \quad (2)$$

となる。この雑音帯域を、中心周波数 f_c に対して対称あるいは非対称に配置する。マスキング閾値をノッチ幅の関数として求め、式(2)の予測誤差を最小とするフ

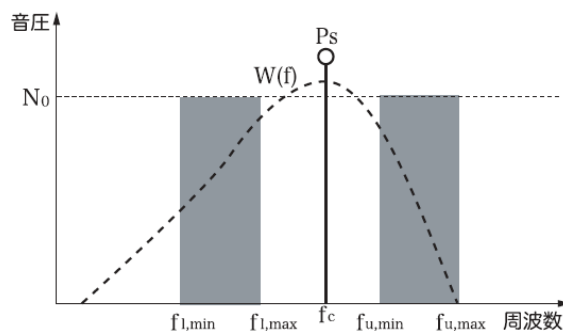


図1 ノッチ雑音法における、プローブ音(P_s)、マスキング雑音(2カ所の灰色部分、 N_0)と聴覚フィルタのポワースペクトル($W(f)$)の関係。横軸周波数、縦軸音圧。

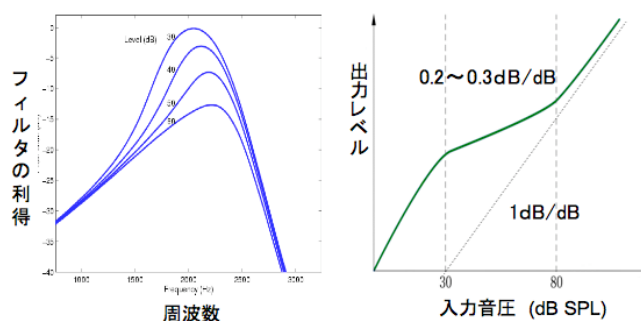


図2 聴覚フィルタ形状のレベル依存性(左)と、フィルタの入出力特性(右)の例。左図)横軸周波数、縦軸フィルタ利得。実線は聴覚フィルタ形状で、上から入力音圧を30, 40, 50, 60 dB SPLとした場合。右図)横軸入力音圧、縦軸出力レベル(dB)。

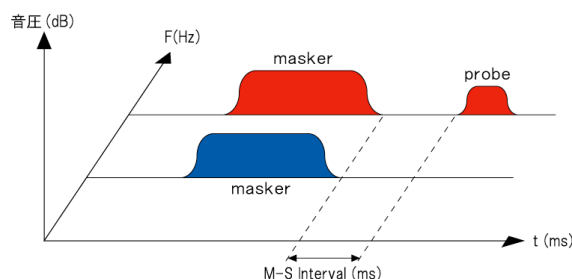


図3 入出力関数を推定するための順向性マスキング実験の実験刺激音の配置。横軸時間、奥行周波数、縦軸音圧。

ィルタ関数 $W(f)$ を推定する。聴覚フィルタは1つだけ存在するわけではなく、連続的に並んでいる。そこで、SNR が最大となるフィルタが聴取時に選ばれと仮定し、離調聴取(off-frequency listening)の影響を取り除くことも推定手法に含まれる[3, 6, 7]。

Patterson は roex (rounded exponential) をフィ

ルタ関数 $W(f)$ の近似式として導入した [6]。この roex フィルタは中心周波数から上下いずれか片側の周波数領域のみで定義される重み付け関数で、聴覚末梢の時間応答をそのままでは表現し得ない。しかし、その分自由度が十分あったため、フィルタ形状が中心周波数に対して非対称であることや音圧による変化も初期の段階で示されている [6, 7]。

このノッチ雑音法と roex フィルタの組み合わせは、実験的にも比較的安定にフィルタ形状を推定できるため、以降幅広く利用されて来た。対象も健聴者ばかりではなく難聴者にも広げられ、指標として特性の把握にも役だっている [2, 8]。

2.3. 位相周波数特性の測定

聴覚フィルタが、インパルス応答によって定義できる線形フィルタとして十分良く近似できるのであれば、位相周波数特性も求めることができるはずである。Oxenham and Dau [9] は、Schroeder phase wave [10] を用いて、位相特性を特定しようと試みた。結果として、伝送線路モデル [11] やガンマチャープフィルタ [12] の位相特性で説明できないと結論づけている。このことは逆に、特定周波数の聴覚フィルタの位相特性を広帯域な刺激音で推定するのが難しいことを示唆する。測定にはさらなる工夫が必要である。ただし、位相特性を心理物理的に直接推定する必要性は高くないと考えられる。後述のように、生理実験から得られているフィルタのインパルス応答を制約として用いれば、振幅周波数特性だけでも十分モデル化ができるからである。

2.4. フィルタ形状の音圧依存性の測定

基底膜振動特性は音圧によって変化するため、対応する聴覚フィルタの形状もレベル依存性があるはずである。前述の聴覚フィルタ推定手法では、マスキング音圧レベルが一定である範囲での線形性を仮定している。マスキング音の固定音圧を順次変更し、その音圧ごとの形状を測定・推定すれば、音圧変化に対する形状変化も見ることができる。

図 2 左に、推定された聴覚フィルタの振幅周波数特性の例を示す。入力音圧が低い場合、フィルタ形状は鋭く、音圧上昇とともになだらかになっている。さらに、音圧レベルの上昇に伴ってフィルタ利得が減少していることもわかる。

この図は、後述の圧縮型ガンマチャープフィルタ [13] を推定に用いた場合の例で、両方の特性が同時に得られている。フィルタ形状の変化に関しては、従来から用いられてきた roex フィルタでも推定でき、報告されている [7]。しかし、利得変化に関しては roex だけでは推定できず、音圧依存の利得を表す関数が必要となる。用いるフィルタ関数によって、式 (2) から推定される内容も異なることがわかる。

2.5. 圧縮特性とその測定

基底膜振動において、外界の音圧が大きくなっても振幅が制限される「圧縮特性」がある [3, 29]。これによりダイナミックレンジの大きい外界の音を、比較的狭い範囲の振動に変換している。これが、まさしく図 2 左の音圧依存の利得変化に対応する。図 2 右に、入力音圧に対する振動レベル（出力レベル）を示す入出力特性の例を示す。入力音圧が 30–80 dB SPL の範囲で入力に対して出力が 0.2–0.3 dB/dB の割合でしか増加しない。これを圧縮特性と呼び、その範囲外では線形 (1dB/1dB の点線と平行) になる。この特性は、図 2 左のフィルタにおいて、パラメータの音圧を横軸に、その最大利得の値を縦軸に取って結ぶと得られる。

この入出力特性は、マスキング音がプローブ音と時間的に重畳せずに先行する、順向性マスキング法を用いて推定される。図 3 にマスキング音とプローブ音の配置を示す。まず、調べたい周波数のプローブ音に対して、同じ周波数のマスキング音による閾値を測定する。次に低い周波数のマスキング音による閾値を測定する。この両者の違いから入出力特性を推定する。

この方法の基本的として、図 2 左の聴覚フィルタ特性を見ると、中心周波数付近は音圧レベルによって大きく利得が変化するのに対して、そこから離れた低い周波数では、利得の変化しない（すなわち線形の）状態であることを用いる。プローブ音とマスキング音が同じ周波数の場合、フィルタ特性に対して同じだけ圧縮特性の影響があるが、マスキング音が低い周波数の場合、マスキング音圧の増加によってフィルタ特性への影響はほとんどない。その違いを見るわけである。

マスキング音の音圧を順次増加させ、それに伴うプローブ閾値の変化を直接測定する、Growth-of-Masking (GOM)/Growth-of-Maskability (GMB) 法が最初提案された [14, 15, 16]。しかし、特性に本来関係ないはずの雑音によって測定値が異なったり、プローブ音増加とともに、基底膜の振動パターン（この包絡線を興奮パターン excitation pattern と呼ぶ）の範囲が広がって測定結果に影響を与えるといった問題点があった。そこで、Nelson ら [16] は、プローブ音を絶対閾値上 10dB としてその興奮パターンを固定し、マスキング音とプローブ音との時間間隔 (M-S Interval) をパラメータとして、プローブ音を丁度マスクするマスキング音圧を測定する方法を提案した [17, 18]。横軸が時間間隔、縦軸にマスキングレベルを取ると、Temporal Masking Curve (TMC) と呼ばれる線がマスキング周波数に対応して 2 本描ける。時間間隔を媒介変数として、横軸を同一周波数のレベル、縦軸に異なる周波数のレベルを取り直すと、ちょうど入出力特性に相当する [18]。時間間隔をマスキングの度合いの調整として利用する巧妙な手法である。

ただし、時間的なマスキング回復特性が異なるマスク一周波数でも同一であるという仮定を置いており、その妥当性も現在議論されている。

2.6. その他の特性の測定

上記は、聴覚フィルタの特性として心理物理的に現在重要と考えられているものである。しかし、基底膜振動に関しては二音抑圧(two-tone suppression)等の非線形現象が知られている。この測定は Houtgast や Duifhuis の古典的な実験がある[19, 20] (詳細略)。1つの聴覚フィルタだけでは説明できないが、聴覚フィルタを並べたフィルタバンクを作る際には考慮に入れる必要がある。さらに、非線形性の代表例として、耳音響放射(otoacoustic emission, OAE)がある。しかし、これをモデル化することは複雑で、簡潔な関数系を導入することが目的の聴覚フィルタの概念から出発するのは、妥当とは思われない。むしろコルチ器を含めた機械的な振動も記述して拡張できる、有限要素法や伝送線路モデル等を用いて定式化すべきであろう。

3. 聴覚フィルタの定式化

ノッチ雑音マスキング実験結果から聴覚フィルタを推定するために、周波数軸上の重み関数である roex(rounded exponential)が用いられてきた[2-8]。これにより得られた成果は重要で影響力が大きい。しかし、本来、基底膜振動に対応する聴覚フィルタを考えるのであれば、時間応答を持つフィルタ関数を定義したい。これにより人間の特性を適切に反映させた基底膜振動のシミュレーションができるようになる。

3.1. ガンマトーンフィルタ

このための第一歩として、Patterson は roex フィルタをさらに近似するための関数として線形のガンマトーンフィルタを用いた[21]。このフィルタは、聴神経発火と入力音との相関を取ることによりインパルス応答を推定する reverse-correlation(revcor)法による測定結果から定式化された[22, 23]。インパルス応答の包絡線がガンマ関数、搬送波が正弦波(pure tone)で近似されることから“gammatone”と呼ばれるようになった。時間関数は、以下のように表される。

$$g_c(t) = at^{n-1} \exp(-2\pi b \text{ERB}(f_c)t) \cos(2\pi f_c t + \phi) \quad (t > 0) \quad (3)$$

ただし、 t :時間、 f_c :中心周波数、 a :振幅、 $\text{ERB}(f_c)$: f_c における矩形帯域幅、 b :係数、 ϕ :位相。中心周波数(ピーク周波数)付近の振幅周波数特性がガンマトーンと roex で似ているため、多数並べてフィルタバンクを構成すれば人間の末梢特性をある程度良く近似できることになる。これは、聴覚系の特性解明ばかりでなく、信号処理に関しても多くの研究者に利用されてきている[24-28]。

3.2. レベル依存性と非対称性の導入

図2に示すように、聴覚フィルタは基本的に周波数特性が中心周波数に対して非対称である。また、図2左で示したとおり、入力音圧レベルが大きいと利得が小さく広帯域で非対称性が大きくなり、レベルが小さくなるに従い利得が大きく帯域も狭まって非対称性も小さくなる。この特性は生理実験的にも確認されている[29, 30]。

これに対して上記のガンマトーンフィルタは基本的に線形で対称な周波数特性しか近似できなかった。そこでガンマトーンフィルタを拡張するモデルが提案されている。Carney[31]は、線形とレベル依存性のある非線形のガンマトーンを並列に接続して、レベル依存の利得や帯域幅を説明するモデルを作った。さらに非対称性の変化も表せるように改良されている[31]。Lyon[33, 34]は、ガンマトーンを IIR フィルタで近似する[26]時に、全極型か零点を1つにすることによって非対称性[30]を導入できることを示した。Meddis ら[35, 36]は、2つのガンマトーンを並列に入れた Dual-Resonance, Non-Linear (DRNL) フィルタを導入し、圧縮特性や抑圧特性を説明している。主に、生理実験データを良く説明するために多数のパラメータを導入してもよしとする立場で研究はすすめられてきた。

3.3. ガンマチャープフィルタ

これらと異なる最適性の観点(後述)から、Iriano and Patterson[12]は、ガンマトーンを拡張したガンマチャープを提案した。包絡線がガンマ関数、搬送波が周波数変化する chirp 波となるため“gammachirp”と呼ぶ。この時間関数は、以下のように表される。

$g_c(t) = at^{n-1} \exp(-2\pi b \text{ERB}(f_c)t) \cos(2\pi f_c t + c \ln t + \phi) \quad (t > 0) \quad (4)$
式(3)との違いは $c \ln t$ 、すなわち時間の自然対数に chirp の度合いを決める係数 c をかけたものが加わっているだけである。この係数を入力音圧の関数とすることによって、非対称でレベル依存性のあるフィルタ形状を定量的に良く近似できることを示した。

生理学的にも、基底膜のインパルス応答の搬送波の周波数は一定ではなく遷移することも観測されており[37]、この定式化自体は良い近似となっている。しかし、このガンマチャープ提案の前後に、周波数遷移の度合いが入力レベルに依存しないことも示された[38-40]。このことは、上記の「解析的」ガンマチャープにレベル依存性を導入する方法に問題があることを示している[40]。

これを受けて、Iriano and Patterson[13]は、生理データと心理物理データを同時に説明できる「圧縮型」ガンマチャープ(cGC, compressive gammachirp)を提案した。これにより、レベル依存の形状と圧縮特性を一度に導入できた。この cGC を多数の被験者の大規模

なノッチ雑音マスキングデータ[42, 43]に適合させ、健聴者の平均的な聴覚フィルタ形状を少数パラメータで表現できることも示した[44, 45]。

さらに、時間的に変化する音に対して動的に特性を変化させる動的圧縮型ガンマチャープフィルタバンク(dcGC, dynamic compressive gammachirp)も提案されている[46]。

3.4. パラメータ数の意味での妥当性

有限要素法や伝送線路モデルの方が基底膜の機械振動を良く近似でき、妥当性が高いように見える。しかし、個々の人間の特性を定量的に近似するためには、実験データに対する適合(fitting)が不可欠である。これゆえに、聴覚フィルタの概念に基づいて少数パラメータ数の roex フィルタが使われてきた。これから発展した圧縮型ガンマチャープは、他のガンマトーンの末裔のフィルタ[31-36]に比べてもパラメータ数が少ないため、容易にデータ適合に使用できる。ただし、計算機のスピードアップに伴い大規模な適合も可能となっているため、圧縮型ガンマチャープが唯一の手段ではない。しかし、モデルがパラメータを多数持つ場合、お互いの整合性に配慮と妥当な結果は得られないと考えられる。これは、ある事を説明するときに必要以上に多くの前提を仮定すべきでないとする「オッカムの剃刀」「けちの原理」にもかなっている。

3.5. 最適性

ガンマトーンは、外界の音を信号処理する上で最適なフィルタなのであろうか。フーリエ変換で得られる時間周波数空間で最小不確定性を持つ関数は、良く知られているようにガウス関数/ガボール関数で、明らかにガンマトーンと関数系が異なる。それでは、聴覚信号処理系は最適である必要がないのであろうか。

この観点の思索から、メルン変換で得られる時間-スケール空間における最小不確定性を持つ関数として、ガンマチャープ[12]は生まれた。これはメルン変換を聴覚系内部に持っているという仮説に基づいている[47]。この変換系は、たとえば音声であれば、声道の寸法の情報と形状の情報を分離して計算することを可能とする。聴覚フィルタバンクはウェーブレット変換近似でき、外界の音の寸法変形に対してスケール軸上でのシフトしか与えない。さらに核関数のガンマチャープは、最小限の不確定性しか与えない。ある意味「透明」なフィルタ系であるので、後段の寸法/形状情報抽出が効果的に実行できる。この仮説について議論はつきないが、実験的に支持する結果も数多く出てきている[48-53]。理論が先行し実験で確認する科学のサイクルに乗った研究となっている。

最近、音符号化の効率性の意味で最適化計算を行うと、ガンマトーンに似たベクトル系列になることが報

告されている[54]。面白い観点ではあるが、少数パラメータの関数系を特定できたわけではないため、実験結果に適合させる目的では利用できない。

4. おわりに

日本はすでに高度高齢化社会に突入している。2013年には4人のうち1人、2035年には3人に1人、2055年に2.5人のうち1人が高齢者となる[55]。多くの人が加齢と共に耳が遠くなり[56]、老人性難聴者の数も増えるであろう。多くの人々が楽しく共生できる社会の実現のためには、解決しなければいけない問題の一つである。聴覚末梢系の劣化が大きな要因であることは間違いないが、それをモデル化するだけでは不十分で、それ以降の知覚情報処理機構の解明も必要である。これらの研究を通して個々人の難聴の模擬をすることや、そのための実験的データ測定 of 最適化・効率化は、今後の補聴器開発に不可欠である。

たとえ今20代でも、50年後には確実に高齢者になる。ぜひ、多くの若手に自分自身の問題として、聴覚研究に参加してもらえれば幸いである。

謝辞 聴覚フィルタを含む聴覚心理/モデル化の研究に関して Cambridge 大学 Patterson 博士から多くの薫陶を受け、共同研究も進めることもできました。心より感謝いたします。

文 献

- [1] Fletcher, H., "Auditory patterns," Rev. Mod. Phys. **12**, 47-61, 1940.
- [2] Patterson, R. D., "Auditory filter shape," J. Acoust. Soc. Am. **55**, 802-809, 1974.
- [3] Moore, B. C. J., "Psychology of Hearing (5th ed)," Academic Press, London, 2003. (注:大串監訳は 3rd ed.)
- [4] Patterson, R. D. and Henning, G. B., "Stimulus variability and auditory filter shape," J. Acoust. Soc. Am., **62**, 649-664, 1977.
- [5] Patterson, R. D., "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640-654, 1976.
- [6] Patterson, R. D. and Nimmo-Smith, I., "Off-frequency listening and auditory-filter asymmetry," J. Acoust. Soc. Am., **67**, pp.229-245, 1980.
- [7] Glasberg, B. R. and Moore, B. C. J. "Derivation of auditory filter shapes from notched noise data", Hearing Research, **47**, pp.103-138, 1990.
- [8] Moore, B.C.J., "Perceptual Consequences of Cochlear Damage," Oxford University Press, Oxford, 1995.
- [9] Oxenham A.J. and Dau, T., "Towards a measure of auditory-filter phase response," J. Acoust. Soc. Am. **110**, pp. 3169-3178, 2001.
- [10] Schroeder, M. R., "Synthesis of low peak-factor signals and binary sequences with low autocorrelation," IEEE Trans. Inf. Theory **16**, 85-89, 1970.
- [11] Strube, H. W., "A computationally efficient basilar-membrane model," Acustica **58**, 207-214, 1985.
- [12] Irino T. and Patterson, R.D., "A time-domain, level-dependent auditory filter: the gammachirp," J. Acoust. Soc. Am., **101** (1), pp.412-419, January 1997.
- [13] Irino T. and Patterson, R.D, "A compressive gammachirp auditory filter for both physiological and psychophysical data," J. Acoust. Soc. Am., **109** (5), pp.2008-2022, 2001.
- [14] Hicks, M. L., and Bacon, S. P. "Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing," J. Acoust. Soc. Am. **105**, 326-338, 1999.
- [15] Oxenham, A. J., and Plack, C. J. "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666-3675, 1997.

- [16] Nelson, D. A., Schroder, A. C., and Wojtczak, M., "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners", *JASA*, 110, pp. 2045-2064, 2001
- [17] Plack, C. J., Daga, V., and Lopez-Poveda, E.A. "Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss", *JASA*, 115, pp.1684-1695,2004.
- [18] 入野,高橋,河原,バターンソン,"圧縮型ガンマチャープ聴覚フィルタによるノッチ雑音データと圧縮特性データへの同時適合," 日本音響学会聴覚研究会資料, H-2009-51, 39(4), pp.283-288, 2009.
- [19] Houtgast, T., "Psychophysical evidence for lateral inhibition in hearing," *J. Acoust. Soc. Am.* 51,1885-1894, 1972.
- [20] Duifhuis, H., "Level effects in psychophysical two-tone suppression," *J. Acoust. Soc. Am.* 67, 914-927, 1980.
- [21] Patterson, R. D., J. Holdsworth, I. Nimmo-Smith and P. Rice,"SVOS Final Report: The Auditory Filterbank," APU report 2341, 1987.
- [22] Johannesma, P. I. M., "The pre-response stimulus ensemble of neurons in the cochlear nucleus," in *Symposium on Hearing Theory (IPO, Eindhoven, The Netherlands, 58-69, 1972.*
- [23] de Boer, E. and de Jongh, H.R."On cochlear encoding: Potentialities and limitations of the reverse-correlation technique," *J. Acoust. Soc. Am.* 63, 115-135, 1978.
- [24] Meddis, R. and M. J. Hewitt,"Virtual pitch and phase sensitivity of a computer model of the auditory periphery: I pitch identification," *J. Acoust. Soc. Am.* 89, 2866-2882, 1991.
- [25] Patterson, R.D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C. and Allerhand M., "Complex sounds and auditory images," in *Auditory physiology and perception, Proceedings of the 9th International Symposium on Hearing, Y. Cazals, L. Demany, K. Horner (eds), Pergamon, Oxford, 429-446, 1992.*
- [26] Slaney, M., "An efficient implementation of the Patterson-Holdsworth auditory filterbank," *Apple Computer Technical Report #35,1993.*
- [27] Cooke, M., "Modelling Auditory Processing and Organisation," Cambridge University Press, 1993.
- [28] Patterson, R. D., Allerhand, M. and Giguère, C., "Time-domain modelling of peripheral auditory processing: a modular architecture and a software platform," *J. Acoust. Soc. Am.* 98, 1890-1894,1995.
- [29] Pickles, J.O., "An Introduction to the Physiology of Hearing," Academic Press, London, 1988.
- [30] Ruggero, M.A., "Responses to sound of the basilar membrane of the mammalian cochlea," *Current Opinion in Neurobiology*, 2, 449-456, 1992.
- [31] Carney, L. H. "A model for the response of low-frequency auditory-nerve fibers in cat," *J. Acoust. Soc. Am.* 93, 401-417, 1993.
- [32] Zhang, X., Heinz, M. G., Bruce, I. C. and Carney, L. H., "A phenomenological model for the responses of auditory-nerve fibres: I. Nonlinear tuning with compression and suppression," *J. Acoust. Soc. Am.* 109, 648-670, 2001.
- [33] Lyon, R.F., "The all-pole gammatone filter and auditory models," *Forum Acusticum '96, Antwerp, Belgium, 1996.*
- [34] Lyon, R.F., "All-pole models of auditory filtering," in *Diversity in Auditory Mechanics, Lewis et al. Eds, World Scientific, Singapore, 1997.*
- [35] Meddis, R., O'Mard, L. P. and Lopez-Poveda, E. A., "A computational algorithm for computing nonlinear auditory frequency selectivity," *J. Acoust. Soc. Am.* 109, 2852-2861,2001.
- [36] Lopez-Poveda, E. A. and Meddis R., "A human nonlinear cochlear filterbank," *J. Acoust. Soc. Am.* 110, 3107-3118, 2001.
- [37] Møller, A.R. and Nilsson, H.G., "Inner ear impulse response and basilar membrane modelling," *Acustica* 41, 258-262,1979.
- [38] de Boer, E. and Nuttall, A. L. (1997). "The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions," *J. Acoust. Soc. Am.*, 101, 3583-3592.
- [39] de Boer, E. and Nuttall, A. L. (2000). "The mechanical waveform of the basilar membrane. III. Intensity effects," *J. Acoust. Soc. Am.* 107, 1497-1507.
- [40] Recio, A.R., Rich, N.C., Narayan, S.S. and Ruggero, M.A. (1998). "Basilar-membrane response to clicks at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* 103, 1972-1989.
- [41] Carney, L. H., Megean, J.M., and Shekhter, I. (1999). "Frequency glides in the impulse responses of auditory-nerve fibers," *J. Acoust. Soc. Am.* 105, 2384-2391.
- [42] Baker, R. J., Rosen, S. & Darling, A. M. (1998) "An efficient characterisation of human auditory filtering across level and frequency that is also physiologically reasonable," in *Psychophysical and physiological advances in hearing: Proceedings of the 11th International Symposium on Hearing, Eds A. Palmer, A. Rees, Q. Summerfield and R. Meddis. Whurr, London, 81-88.*
- [43] Glasberg, B. R. and Moore, B. C. J. (2000). "Frequency selectivity as a function of level and frequency measured with uniformly exciting noise," *J. Acoust. Soc. Am.* 108, 2318-2328.
- [44] Patterson, R.D., Unoki, M. and Irino T., "Extending the domain of center frequencies for the compressive gammachirp auditory filter," *J. Acoust. Soc. Am.*, 114 (3), pp.1529-1542, Sept 2003.
- [45] Unoki, M., Irino, T., Glasberg, B., Moore, B. C. J., and Patterson R. D., "Comparison of the roex and gammachirp filters as representations of the auditory filter," *J. Acoust. Soc. Am.* , 120(3), pp.1474-1492, Sept., 2006.
- [46] Irino T. and Patterson, R.D., "A dynamic compressive gammachirp auditory filterbank" *IEEE Trans. Audio, Speech, and Lang. Process.*, 14(6), pp.2222-2232, 2006.
- [47] Irino T. and Patterson, R.D., "Segregating information about the size and shape of the vocal tract using a time-domain auditory model: The Stabilised Wavelet Mellin Transform," *Speech Communication*, 36 (3-4),pp.181-203, March 2002.
- [48] Smith D.R., Patterson, R.D., Turner D., Kawahara H., and Irino T., "The processing and perception of size information in speech sounds," *J. Acoust. Soc. Am.*, 117(1), pp. 305-318, Jan., 2005.
- [49] Ives, D.T., Smith, D.R.R. and Patterson, R.D., "Discrimination of speaker size from syllable phrases." *J. Acoust. Soc. Am.*, 118, p.3816-3822, 2005.
- [50] Smith, D.R.R. and Patterson, R.D., "The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age." *J. Acoust. Soc. Am.*, 118, p.3177-3186, 2005.
- [51] van Dinther, R. and Patterson, R.D., "Perception of acoustic scale and size in musical instrument sounds." *J. Acoust. Soc. Am.*, 120, p.2158-76, 2006.
- [52] Smith, D.R.R., Walters, T.C. and Patterson, R.D., "Discrimination of speaker sex and size when glottal-pulse rate and vocal-tract length are controlled." *J. Acoust. Soc. Am.*, 122, p.3628-3639, 2006.
- [53] Patterson, R.D., Smith, D.R.R., van Dinther, R. and Walters, T.C., "Size Information in the Production and Perception of Communication Sounds", in *Auditory Perception of Sound Sources*, Yost, W.A., Popper, A.N. and Fay, R.R. editors (Springer Science+Business Media, LLC, New York), 2008.
- [54] Lewicki, M. S., "Efficient coding of natural sounds," *Nature Neuroscience*, 5, 356 - 363, 2002.
- [55] 内閣府 高齢社会白書 <http://www8.cao.go.jp/kourei/index.html>
- [56] ISO7029:1984, ISO7029:2000, http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csn_umber=26314