



はじめての聴覚フィルタ*

入野 俊夫 (和歌山大学システム工学部)**

43.66.Ba

1. はじめに

私たちは、音を聞いたときに「高い音」や「低い音」といった印象を持つことがあります。これは、音の高さ「ピッチ」を知覚する機能が生体に備わっていることを意味します。この機能のおかげで、音声を上手に聞き取り、音楽を楽しむことができます。ピッチ知覚には、内耳の基底膜振動による周波数分析と、時間的な周期性分析の両方が関与しています。ここでの周波数分析法は「聴覚フィルタ」(2章参照)で近似することができます。これはよく知られているフーリエ変換 [1] を用いた分析法とは異なる特性を持っています。

この聴覚フィルタは、最近の圧縮オーディオ [2] や音質客観測定法 [3] などに入っている「知覚モデル」の重要な要素となっています。これらで用いられているのは、臨界帯域の概念 (3章参照) [4] から出された、健聴者の平均的な周波数選択性を反映させた線形フィルタです。ところが、実際の聴覚フィルタは、外界の音圧や音環境によって変化する非線形フィルタとなっています。更に、個人ごとに異なり、特に高齢者や難聴者の場合は標準的な特性から大きくばらつきます。今後の高度高齢化社会 [5] における音響処理を考える上では、個々人の特性に合わせたモデル化が必要となります。そのためには、心理物理的な非侵襲実験により「その人」のデータを収集し、聴覚フィルタ特性を推定することも必要となるでしょう。

そこで本解説では、聴覚フィルタ特性推定の具体的な方法を体験してもらいます。これを通して、聴覚フィルタとは何かを理解しましょう。まず2章で「聴覚フィルタ」の基本概念と特性を述べます [6]。

3章では、2種類の心理物理実験のデモを体験してもらいます。4章において「自分の聴覚フィルタ」を推定してみましょう。ここでは、Mathworks社製のMATLAB上で動くデモソフト ([7] からダウンロード可) を利用して順次進めます^(注1)。なお、本解説で提供するデモソフトは、その原理及び概念を理解するための教材であり、厳密な測定結果を提供するものではないことをご理解の上ご使用下さい^(注2)。

2. 聴覚フィルタとは

2.1 聴覚末梢系の周波数分析

外界の音は、鼓膜を振動させ、耳小骨を通して、内耳の蝸牛に伝達されます。ここで、基底膜を機械的に振動させます。基底膜の上には、コルチ器で保持された聴神経があり、機械振動が神経発火情報に変換されます。この神経発火情報は、聴覚経路を經由して大脳皮質に伝達され、更に様々な処理が行われることとなります [8, 9]。ここで、入力される音の周波数によって、基底膜の振動が最大となる場所が変わります。耳小骨に接続する入口側が高い周波数を担当し、逆の頂上側が低い周波数を担当するようになっています^(注3)。これは、場所ごとに中心周波数と帯域幅が異なるフィルタが多数並んでいる形で定式化できます。この各々のフィルタのことを「聴覚フィルタ」と呼びます。また、それを周波数順に多数並べた1セットを、聴覚フィルタバンクと呼びます。

2.2 デモソフトの実行

それでは、この聴覚フィルタの特性がどのようなものか、MATLABを立ち上げてデモソフトを走らせてみましょう。実行コマンドはDemoAuditoryFilterです。すぐに動かない場合は、読者の使用環境に合わせて、プログラムDemoAuditoryFilter.mを編集して下さい^(注1)。正常に実行できると、デモ番号が選択できるよう

* An introduction to auditory filter.

** Toshio Irino (Faculty of Systems Engineering, Wakayama University, Wakayama, 640-8510)
e-mail: irino@sys.wakayama-u.ac.jp
URL: <http://www.wakayama-u.ac.jp/~irino/>

になります。ここでは聴覚フィルタの基礎を学ぶため1を入力します。結果として、四つのグラフが出力されると思います。

```
>> DemoAuditoryFilter ◀ [注: return 入力]
デモ :
  1) 聴覚フィルタの基礎
  2) 臨界帯域幅
  3) ノッチ雑音マスキング法
  4) 聴覚フィルタ形状推定
デモ番号の選択 >> 1

— デモ: 1) 聴覚フィルタの基礎 —
Figure 1: Frequency response [注: 図-1]
Figure 10: ERB_N [注: 本文中図無し]
Figure 2: Filter level dependency [注: 図-2]
Figure 3: Input-output function [注: 図-3]
```

2.3 振幅周波数特性

図-1は、フィルタの振幅周波数特性で、横軸は周波数、縦軸はフィルタの利得をdB表示しています。ただし横軸は、周波数の値の並びを見ても分かるように、線形軸ではなく対数間隔に近い軸となっています。これにより、六つの中心周波数を持つフィルタが、周波数に関わらず同じような形状を持つことが分かります。なお、この振幅周波数特性のことを、心理物理では「フィルタ形状」と呼びならわしています。

ここで、中心周波数と帯域幅の間には、以下の関係があることが測定から知られています [9]。

$$ERB_N = 24.7(4.37F/1000 + 1) \quad (1)$$

ただし、 ERB_N : 健聴者の測定から得られた「矩形等価帯域幅」(Hz)、 F : フィルタの中心周波数 (Hz)。この関係は、MATLABのFigure 10 (本文中図略) に示されるように、中心周波数が高くなるにつれて帯域幅は広がります。この周波数によって変わる帯域幅のフィルタが等間隔に並ぶように ERB_N 番号が定義できます。

$$ERB_N \text{ number} = 21.4 \log_{10}(4.37F/1000 + 1)$$

図-1の横軸は、この ERB_N 番号が等間隔になる周波数軸となっています。

2.4 音圧依存性と入出力特性

実際の聴覚フィルタは、外界の音環境や音圧に依存して、その特性が変化する非線形の時変フィルタです。この点で、信号処理の講義の最初に習う線形時不変フィルタや短時間フーリエ変換とは異なり、インパルス応答だけで特性のすべてを語ることはできません。

図-2に、音圧に対するフィルタ形状・利得変化の様子を示します。30 dBで中心周波数における利得が0 dBと最も大きく、音圧上昇と共に利得が減少して、90 dBでは利得が約-35 dBになっていることが分かります。また、同時に、フィルタの帯域幅も音圧上昇と共に徐々に広がる傾向があることが分かります。この意味では、帯域幅を示す式(1)は、健聴者を用いて中程度の音圧で測定した場合の一つの目安として見る必要があります。

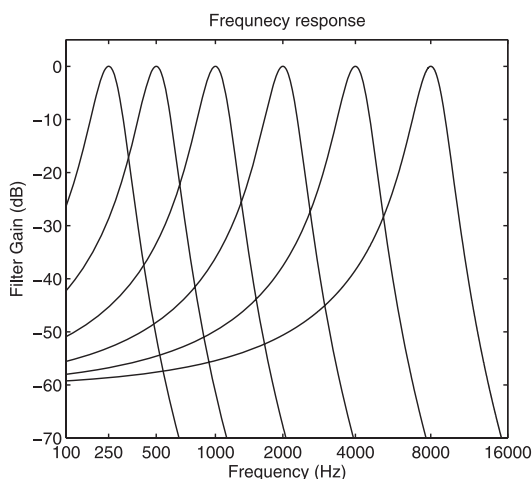


図-1 聴覚フィルタの周波数特性の例
振幅が最大となる周波数が250, 500, 1,000, 2,000, 4,000, 8,000 Hzの6個の聴覚フィルタを示している。

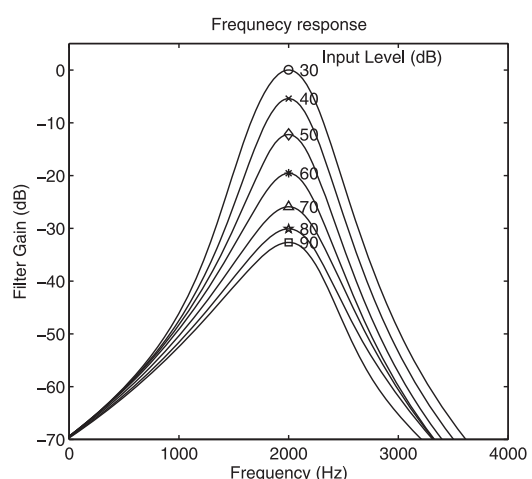


図-2 聴覚フィルタの音圧依存性
音圧が低いとフィルタの利得が大きく、音圧上昇とともに利得が減少する。

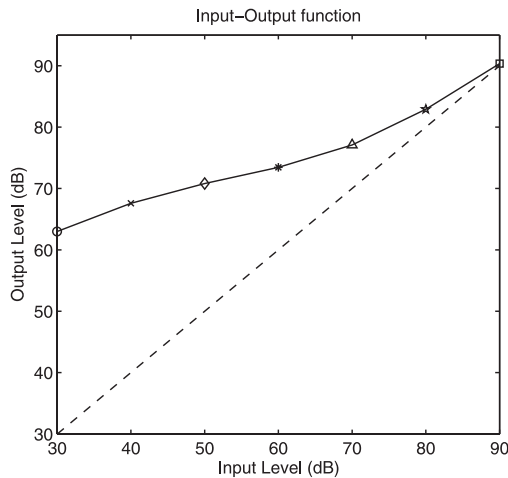


図-3 聴覚フィルタの入出力特性例

図-2の入力音圧に対する最大利得の値を用いて、対応する出力音圧を計算し、入力90 dBに対して出力90 dBとなるようにレベルを調整した後に図示した。図-2と図-3で図記号(○×△等)は対応している。

図-2の特性から計算される、フィルタの入力音圧に対する出力音圧の関係を、図-3に示します。破線で示された増加率1 dB/dB (1:1)の線形関係に比べて、曲線の傾きがゆるやかです。健聴者の場合、中音圧域で約0.2~0.3 dB/dBの増加率になると推定されています。これに対し、低音圧域や高音圧域においては、線形関係(1 dB/dB)に近くなります。このように、入力増加に対して、出力増加の割合が小さいために、心理物理実験的には「圧縮特性」と呼びならわされています。

以上が聴覚フィルタの主な非線形特性です。この他にも、二音抑圧や耳音響放射などの蝸牛における非線形現象が知られています[8]。

3. 聴覚フィルタ特性の心理物理的測定

前章で述べた特性を持つ聴覚フィルタは、マスキング閾値を計る心理物理実験から推定できます。ここでは、まずこのマスキング現象の概念を述べた上で、聴覚フィルタの概念の先駆けとなったFletcherの臨界帯域幅測定法[4]を紹介します。更に、その問題点を克服して現在の標準手法となっているPattersonのノッチ雑音マスキング法[10]を紹介します。

3.1 マスキング現象と心理物理実験

例えば、駅や道路際で電話をかけているときに、電車や車が近づくと聞き取りにくくなることは、誰でも経験したことがあるでしょう。大きい音の

存在によって小さい音が聞こえなくなる現象がマスキングです。

心理物理実験においては、様々な検出音(プローブ音)が、どの程度の音圧ならば検出できるかという測定(閾値測定/弁別域測定)から様々な特性の推定が行われています。例えば、純音(sin波)の検知を考えましょう。同じ周波数帯域に雑音が同時に提示される場合には、それが無い場合に比べて、純音の音圧を高めないと検出できなくなります。この増加分の音圧量がマスキング量に相当します。このマスキング現象を用いて、聴覚フィルタが推定できます。

3.2 Fletcherによる臨界帯域幅測定

Fletcherは、「臨界帯域(critical band)」の測定を行い、人間の周波数分析特性もフィルタとして定式化できることを示しました[4]。実験では、純音の検出音(プローブ音)と、様々な帯域幅の雑音(マスカ音)を同時に再生します。そこで、雑音の帯域幅変化に対する純音のマスキング量変化を調べます。それでは、DemoAuditoryFilterのデモ番号2を選択して臨界帯域幅を測定してみましょう([11]に準拠しています)。

```
>> DemoAuditoryFilter ◀
...
デモ番号の選択 >> 2 ◀

== デモ:2) 臨界帯域幅 ==
5 dB ずつ減衰するプローブ音系列を 2 回再生します。
何個聞こえるか教えてください。
リターンで再生開始 >> ◀
再生中... (Figure 11 にスペクトル表示中)
聞こえた数 >> 15 ◀ [注:自分の回答を入力]

次に帯域雑音を重畳します。
帯域雑音の種類ごとに、何個聞こえるか教えてください。
帯域幅: 4000 (Hz)
リターンで再生開始 >>
再生中... (Figure 11 にスペクトル表示中)
聞こえた数 >> 8 ◀

帯域幅: 1000 (Hz) [注:以下繰り返し]
...
Figure 4: Result [注:図-4]
```

このデモ音再生時には、MATLABのFigure 11に、提示音の一部分のスペクトルが表示されてい

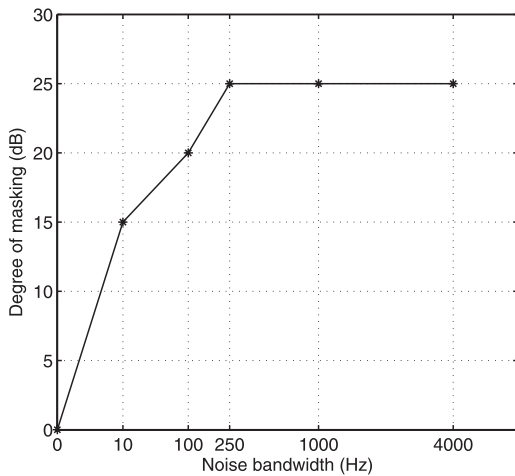


図-4 雑音帯域幅に対するマスキングの度合いの例
250Hz以上ではマスキング量が変化しない。

ますので、その変化も見てください。この実験においては、プローブ音の変化は5dBステップと粗いですが、聞こえた数を答えることにより、おおよそのマスキング量を推定することができます。聞こえる数が多いことは検知できるプローブ音圧が小さく、マスキング量が小さいことに相当します。図-4に、最後まで回答した後に表示されるグラフの例を示します。横軸は雑音の帯域幅 (Hz)、縦軸はマスキング量 (dB) となっています。この結果を見ると、雑音帯域幅が狭い場合と広い場合でプローブ音圧が異なることがわかります。また、雑音帯域 250 Hz 以上ではプローブ音圧が一定、すなわちマスキング量が一定であることがわかります。つまり、その帯域幅を超えた部分の雑音は聴覚フィルタにより除去されていて、プローブ音を検出することに影響しないと考えられます。Fletcher は、この丁度境界となっているところを聴覚フィルタの「臨界帯域幅」と呼びました。

3.3 マスキングのパワースペクトルモデル

Fletcher は、マスカレベルと丁度検出できるプローブレベルの関係も定式化しました [4]。マスカ雑音のパワースペクトルを $N(f)$ 、聴覚フィルタのパワースペクトルを $W(f)$ 、聴覚内部にあると仮定する検出器の効率を定数 K とすると、プローブレベル P_s が以下の式 (2) でよく予測できると提案したのです。

$$P_s = K \int_{-\infty}^{\infty} N(f)W(f)df \quad (2)$$

これは、「マスキングのパワースペクトルモデル」

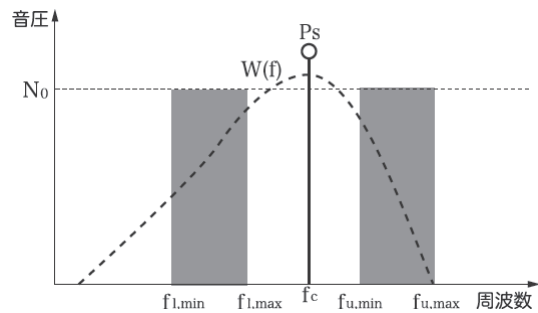


図-5 ノッチ雑音マスキング法における、プローブ音 (P_s)、マスカ雑音 (2か所の灰色部分, N_0) と聴覚フィルタのパワースペクトル ($W(f)$) の関係
横軸周波数, 縦軸音圧。

と呼ばれ、フィルタ関数が既知であれば、同時マスキングの実験結果をよく説明できます。プローブレベル P_s が測定から分かれば、逆にフィルタ関数 $W(f)$ を推定することもできます。ここで、 $N(f)$ は実験で用いて既知なので、フィルタ関数 $W(f)$ を仮定し、式 (2) に代入して得られるプローブレベルの予測値 \hat{P}_s と測定値 P_s の誤差が最小となるように、フィルタ関数パラメータを振って最適化計算を行います。

マスキングのパワースペクトルモデルの利用は、次節で述べるノッチ雑音マスキング法も含め聴覚フィルタ特性推定における基本手法となっています。非常に簡単で便利ですが、大前提となる線形性の仮定等、多くの仮定を伴います。例えば、短時間の過渡音は、式 (2) が前提としている長時間パワースペクトルで表すことができません。従って、実験条件によっては成立しない場合もあることを心にとめておく必要があります。

3.4 ノッチ雑音マスキング法

3.2 節の臨界帯域幅測定法よりも、フィルタ形状を広いダイナミックレンジで精度良く推定するための手法として、「ノッチ雑音マスキング法」が Patterson によって提案されました [10]。この手法では、図-5 に示すように、二つの帯域雑音で挟まれた無音のノッチ領域にプローブ音を置き、条件ごとのマスキング量を測定します。その上で、式 (2) のマスキングのパワースペクトルモデルをそのまま適用して聴覚フィルタ形状を推定します。この時、雑音帯域が二つあるので、二つの積分の和になります。また、パワースペクトル $N(f)$ が一定レベル N_0 の場合、 N_0 を積分の外にくくり出すこともできます。この手法の利点は、中心周波

数 f_c の上下に雑音帯域が分かれているため、 f_c に対して非対称にも雑音を配置できることにあります。これにより、フィルタ形状の非対称性も推定できるようになりました。

それでは、DemoAuditoryFilter のデモ番号 3 を選択して、ノッチ雑音マスキング実験デモを行ってみましょう。

```
>> DemoAuditoryFilter ◀
...
デモ番号の選択 >> 3 ◀
— デモ：3) ノッチ雑音マスキング法 —
5 dB ずつ減衰するプローブ音系列を 2 回再生します。
何個聞こえるか教えてください。
リターンで再生開始 >>
再生中...
聞こえた数 >> 15 ◀

次にノッチ雑音を重畳します。
ノッチ雑音の種類ごとに、何個聞こえるか教えてください。
ノッチ幅： 0 (Hz)
リターンで再生開始 >> ◀
再生中... (Figure 11 にスペクトル表示中)
聞こえた数 >> 5 ◀

ノッチ幅： 400 (Hz) [注：以下繰り返し]
...
Figure 6: Result [注：図-6 の実線部分のみ]
```

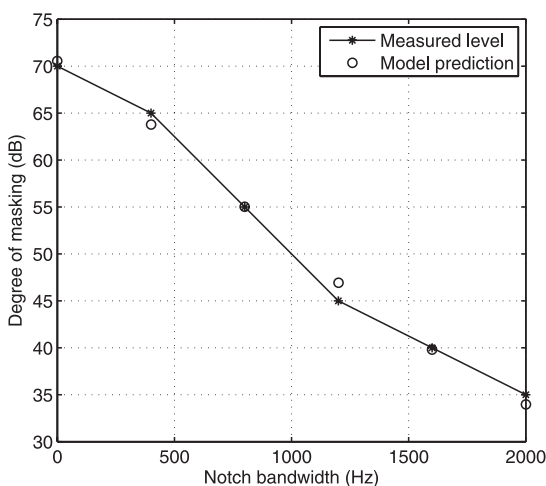


図-6 ノッチ雑音法で得られたノッチ帯域幅に対するマスキングの度合いの例 (*).
○は、モデルから予測される値。

図-6 に、結果の一例を示します。なお、このデモでは、スペクトルを見ても分かるように、中心周波数に対して対称に配置した帯域雑音だけを用いました。横軸は、中心からノッチの端までの周波数差 (全体のノッチ幅の半分) (Hz), 縦軸は、マスキングの度合い (dB) を示します。これで、式 (2) の中にある丁度検出できるプローブ音の音圧 P_s に相当するものが測定できました。後は式 (2) を用いて、最適化計算 (3.3 節) を行えば、パワースペクトル $W(f)$ が推定できます。

4. 聴覚フィルタ形状の推定

4.1 フィルタ関数

Patterson はフィルタ関数 $W(f)$ の近似式として roex (rounded exponential) フィルタを導入しました。この roex フィルタは、フィルタ形状の中心周波数に対する非対称や音圧による変化もよくモデル化できたため、今まで広く使われてきました [9]。しかし、周波数軸における重み関数で、インパルス応答を持たないため、聴覚末梢系における時間的なフィルタを表現できません。そこで、この roex を更に近似するものとして、インパルス応答のあるガンマトーン (gammatone) フィルタが導入されました [6]。

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

ただし、 t : 時間 ($t > 0$), f_c : 中心周波数, a : 振幅, $\text{ERB}_N(f_c)$: f_c における矩形帯域幅, b : 係数, ϕ : 位相。しかし、これは中心周波数に対して対称なフィルタなので、roex の非対称性をそのままでは表現できません。そこで、非対称性も導入できるガンマチャープ (gammachirp) フィルタが提案されました [12]。

$$g_c(t) = at^{n-1} \exp(-2\pi b \text{ERB}_N(f_r) t) \cdot \cos(2\pi f_r t + c \ln t + \phi) \quad (4)$$

ただし、 t : 時間 ($t > 0$), f_r : 漸近周波数, c : 周波数の変化 (chirp) の係数。式 (3) との違いは $c \ln t$ だけで、係数 $c = 0$ とすれば同じ式になります。これを使うと非対称のノッチ雑音マスキング法のデータにも、よく適合させられることが示されています。導入に至る歴史的経緯やその後の動向は文献 [6, 13] を参照して下さい。

4.2 自分の聴覚フィルタ形状

それでは、DemoAuditoryFilter のデモ番号 4 を選択して、デモ番号 3 で得られたノッチ雑音マスキングデータから、「自分の聴覚フィルタ形状」を推定してみましょう。

```
>> DemoAuditoryFilter ◀
...
デモ番号の選択 >> 4 ◀

== デモ：4) 聴覚フィルタ形状推定 ==
パラメータ値：b = 1.59, K = 4.8
推定誤差：RMS error = 1.05 (dB)
Figure 7: Estimated filter shape [注：図-7]
Figure 6: Result and Prediction [注：図-6]
Figure 12: Impulse response of GT and GC
[注：本文中図無し]
Figure 13: Fourier spectrum of GT and GC
[注：本文中図無し]
```

図-6の実験結果から、 $W(f)$ をガンマトーンフィルタ (式 (3)) の振幅周波数特性の 2 乗とした場合の、フィルタ形状を図-7 に実線で示します。実験で用いたのは、中心周波数に対し対称のノッチ雑音だけだったので、対称のフィルタ形状しか求まりません。デモソフトのプログラム中ではガンマチャープフィルタを用いていますが、実際は $c = 0$ としてガンマトーンフィルタに設定して計算しています。図-7 の破線は、式 (1) の ERB_N と同じ帯

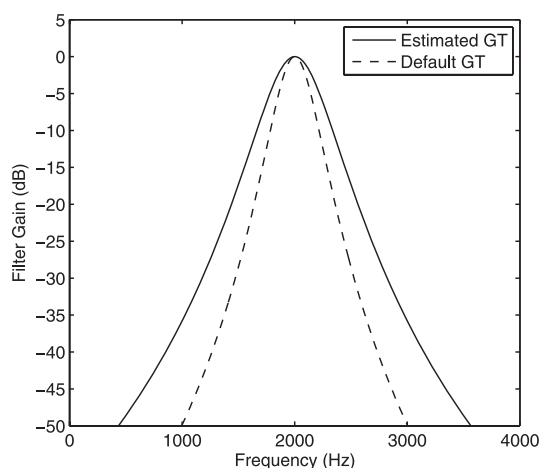


図-7 図-6 のデータから求まったガンマトーンフィルタの振幅周波数特性 (実線)
破線は、式 (1) の ERB_N と帯域幅を等しくしたガンマトーンフィルタ。

域幅になる係数 ($b = 1.019$) を持つガンマトーンフィルタ (MATLAB Figure 12, 13, 破線) の特性です。これに比べ実線の推定されたガンマトーンフィルタの方が、帯域幅が広いことがわかります。もちろん、実験条件や個人によってこれは変わります。各自で自分の聴覚フィルタ形状を算出し、どの程度異なるか比べてみて下さい(注2)。

また、式 (2) で計算された予測レベルの収束値が、図-6 の○印で示されています。測定値との RMS 誤差は、高々 1 dB 強で、5 dB ステップという比較的粗い間隔の音系列を用いた実験としては良い精度程度で求まっています。これに対して、図-4 の臨界帯域幅デモの結果にガンマトーンフィルタを直接当てはめると、誤差はもっと大きくなりそうです。

4.3 聴覚フィルタの実際の測定と推定

以上で、聴覚フィルタ形状の推定法を、デモソフトを用いて説明してきました。実際の測定では、GUI (Graphic User Interface) や反応ボタンを用いた 2 区間 2 肢強制選択法と、刺激音圧の変形上下法を用いて行われます。また、マスクの音圧を段階的に変えて実験を繰り返すことにより、図-2 のような音圧依存のフィルタ形状を出します。更に、中心周波数に対し雑音を非対称に配置する刺激音条件も実験に取り入れて、フィルタの非対称性も推定することが行われます(注4)[6,9]。非対称性をよく表現できるガンマチャープフィルタ (MATLAB Figure 12, 13 の実線) はそのために用いられています。更に、図-2, 図-3 のような圧縮特性も導入した圧縮型ガンマチャープフィルタを用いて、大規模な健聴者のデータから聴覚フィルタ形状が推定されています [13]。そこでは、式 (1) の ERB_N の値よりも帯域幅が広めでレベル依存性があることが示されています。更に、難聴者の場合は、健聴者よりも広い帯域幅を持つように推定され、個人ごとのばらつきが大きくなる傾向にあります [9]。

聴覚フィルタには図-3 のような圧縮特性もあることも忘れてはいけません [9]。この圧縮特性に関して、順向性マスキングを用いて直接推定する手法が提案されています [14]。しかし、まだ精度良く安定に求められる状況にはなっておらず、最前線で議論されています。「聴覚フィルタ」は、決して古典的な話ではないのです。

5. おわりに

日本は高度高齢化社会にすでに突入しています [5]。2055 年には人口の 2.5 人に 1 人が高齢者になると予測されています。たとえ今 20 歳だとしても、45 年後には確実にその仲間に入り、もしかすると耳も遠くなっているかもしれません。このような時代であるからこそ、補聴器を含む音響機器の設計には、老若男女を問わないユニバーサルデザインの考え方が必要でしょう。そのためには、人間の特性を機器に反映させるための基本的な心理物理実験は不可欠となります。心理物理実験の成果や手法を欧米からの輸入に頼るばかりでなく、自前で実験を行うことこそが、真に必要とされる機器の開発につながります。このためにも、最新の動向を含めた聴覚研究の背景知識を得ることは重要で、最近出版された聴覚のハンドブック [15] を熟読することをお勧めします。本解説が、特に若手にとって、心理物理実験への第一歩となることを切望して、まとめとしたいと思います。

注 釈

注 1) [7] からダウンロードできるのは、MATLAB の m-file を集めて圧縮した DemoAuditoryFilter.tar.gz です。MATLAB_R2010a, signal processing toolbox 付を用いて作成/動作確認されています。以前のバージョンでは日本語が文字化けする可能性があります。その時は DemoAuditoryFilter.m の中の変数 SwEnglish を変更下さい。また変数 DirWork も適切に設定する必要があります。MATLAB と同等な機能がある GNU octave においては、音を出力する sound 関数がないため動きません。sound 関数を web で探し出し、適切にインストールできると動くかもしれません。また、Mac 用 MATLAB のあるバージョンでは、sound 関数出力波形自体が歪んでいる場合もありました。ここでは、ソフト起因の歪みがない条件にしたつもりですが、保証できません。

注 2) ここでのデモは、出力音響系や音圧の設定を規定していないので「実験」や「測定」とは呼びません。従って、結果は目安にすぎず、一喜一憂できる性質のものではありません。再生はスピーカでもヘッドホンでもかまいませんので、うるさくない音圧に設定して実施して下さい。このデモソフトを使用した結果生じた損害や損失等は一切補償いたしません。また、このデモを使って他人の特性を計ってはいけません。研究倫理審査を通さない限り、人間や動物を対象とした「実験」を行ってはいけません。

更に、正確な心理物理実験を実施したい場合は、注意深くプログラム作成や音響系の設定をするのは当然として、それ以上に実験に関する深い know how が必要です。しかし、これに関しては、いくら注意深く文献を読んでもよく分からないことも事実です。研究実績のある研究者の指導をおおぐことが、一番の早道です。

注 3) 蝸牛における周波数分析の研究に関しては、動物生

理実験等による膨大な知見の積み重ねがあります [8]。また、機械振動なので、材料定数やその動的変化が分かれば、波動方程式を数値演算的に解くことである程度近似できます。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実施できないため、それらの情報を得ることは不可能です。更には、個人ごとの特性を知りたいければ、その個人で測定できなければ、意味がありません。そこで、心理物理実験による推定手法が開発され、利用されています。

注 4) 特に非対称ノッチ雑音を用いた場合、多数あるフィルタのうちで、SNR が最大となるフィルタが聴取時に選ばれればと仮定して推定します。これは離調聴取 (off-frequency listening) と呼ばれます。また、安定な推定を行うため、フィルタ形状の制約を入れる等、様々なテクニックが実際には使われています。

文 献

初学者むけに日本語の二次資料もなるべく入れました。原著論文を孫引きして、ぜひ読んで下さい。

- [1] 河原一彦, “はじめての周波数分析,” 音響学会誌, 65, 20-25 (2009).
- [2] T. Painter, and A. Spanias, “Perceptual coding of digital audio,” *Proc. IEEE*, 88, 451-513 (2000).
- [3] ITU-T Recommendation P.862, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs” (2001).
- [4] H. Fletcher, “Auditory patterns,” *Rev. Mod. Phys.*, 12, 47-61 (1940).
- [5] 内閣府, 高齢社会白書, <http://www8.cao.go.jp/kourei/index.html>
- [6] 入野俊夫, “聴覚フィルタの測定と定式化について,” 音響学会聴覚研資, H-2009-73, 39, 413-418 (2009).
- [7] <http://researchmap.jp/irino/>資料公開/
- [8] J.O. Pickles, *An Introduction to the Physiology of Hearing*, 3rd ed. (Academic Press, London, 2008); 旧版の邦訳: 谷口郁雄監訳, ピクルス聴覚生理学 (二瓶社, 大阪, 1995).
- [9] B.C.J. Moore, *Psychology of Hearing*, 5th ed. (Academic Press, London, 2003); 3rd ed. の邦訳: 大串健吾監訳, 聴覚心理概論 (誠心書房, 東京, 1994).
- [10] R.D. Patterson, “Auditory filter shape,” *J. Acoust. Soc. Am.*, 55, 802-809 (1974).
- [11] “Demonstration 2. Critical Bands by Masking,” in A.J.M. Houtsma, T.D. Rossing and W.M. Wagenaars, “Auditory Demonstrations on Compact Disk” (1987). 米国音響学会から入手可 (<http://asa.aip.org/discs.html>)
- [12] T. Irino and R.D. Patterson, “A time-domain, level-dependent auditory filter: The gammachirp,” *J. Acoust. Soc. Am.*, 101, pp. 412-419 (1997).
- [13] R.D. Patterson, M. Unoki and T. Irino, “Extending the domain of center frequencies for the compressive gammachirp auditory filter,” *J. Acoust. Soc. Am.*, 114, 1529-1542 (2003).
- [14] D.A. Nelson, A.C. Schroder and M. Wojtczak, “A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners,” *J. Acoust. Soc. Am.*, 110, 2045-2064 (2001).
- [15] C. Plack, Eds., *The Oxford Handbook of Auditory Science: Hearing* (Oxford University Press, New York, 2010).