カルマンフィルタを用いた DPOAE の時間変動の測定

原田 竜彦, 神崎 仁

国際医療福祉大学附属熱海病院耳鼻咽喉科

要旨：カルマンフィルタを用いて、歪成分耳音響放射（DPOAE）の時間変動を観察するシステムを作成し、刺激音圧を変動させた際および刺激音圧は一定として音圧変動する抑圧音を同時負荷した際の DPOAE の音圧変動を正常聴力被験者 2 名にて観察した。第 1 の実験では 2 つの刺激音圧を同時に振幅変調し、2 名の被験者ともに DPOAE 2f₁-f₂成分が刺激音圧の変動に約 40 から 50 ミリ秒程度遅れて周期的に変動する様子が確認できた。第 2 の実験では刺激音は振幅変調せず音圧を一定とし、振幅変調を行った第 3 音で同時に刺激し、DPOAE に対する抑圧を観察した。2 名の被験者ともに第 3 音の音圧ピークからやはり 40 から 50 ミリ秒程度遅れて DPOAE 2f₁-f₂成分が最も抑圧されることが確認できた。

これらの結果からカルマンフィルタは DPOAE の時間変動を観察する手法として有用であることが確認できた。DPOAE の時間変動の観察法はこれまでも複数の方法が提案されているが、それらと比べてカルマンフィルタは測定と同時に周波数成分の結果が表示できるリアルタイム性ならびに複数の周波数成分を同時に算出できる点で優れている。この方法を用いて今後、語音など時間的に変動する音響を聴取する際の蝸牛の働きを明らかにすることが期待できる。

－キーワード－

DPOAE, カルマンフィルタ, 振幅変調, 抑圧

1. はじめに

歪成分耳音響放射（DPOAE）は非侵襲的かつ周波数特異的な蝸牛機能評価が可能な測定法であり、基礎研究および臨床検査に広く活用されている。しかし DPOAE は一定時間の連続した音響記録から周波数解析を行う必要があることから、短時間での変動が観察し難いことが欠点として挙げられる。音声をはじめとするわれわれが日常聴取している音響の多くは短時間で変動するものであり、このような音響に対する蝸牛の活動の評価に DPOAE が活用できないことは研究ならびに臨床検査において大きな支障となっている。これまでも DPOAE 時間変動の観察の試みは、DPOAE の発生潜時の測定¹⁾や低周波音負荷時の影響の評価²⁾ならびに DPOAE

の発生起源の探索³⁾を目的とした研究において試みられているが、煩雑な測定手法であったことから広く用いられるには至っていない。

一方、近年カナダの Vivosonic 社によりカルマンフィルタを用いた DPOAE 測定システムが開発され、2005 年に James らはこのシステムを用いてヒトならびに実験動物を対象に対側音響負荷時の微細な DPOAE 変動を測定した結果を報告した⁴⁾。カルマンフィルタとは、1960 年代初頭にカルマンらにより報告された手法⁵⁾で、雑音を含んで観測される値から時間的な変化を伴う対象の状態を推定する手法である。現代のシステム制御理論の基礎となっており、アポロ計画でのロケットの位置推定への活用にはじまり、現在では情報工学や土木工学、医学、社会・経済学など広い分野にも応用されている⁶⁾。

James らの報告では対側音響負荷時の DPOAE の変動の測定に使用した結果、音響負荷時の抑圧を捉えている様子が明らかにされているが、測定手法の詳細については記載がない。筆者は、このカルマンフィルタを用いる方法が DPOAE 測定の新しい手法として今後広く用いられる可能性があると考え、カルマンフィルタに関する他領域の文献^{7,8)}を参考に測定システムを自作し、刺激音圧の変動時ならびに抑圧音の同時負荷時の DPOAE の時間変動を観察した。

2. 理論的背景とパラメータ設定

カルマンフィルタとは、測定対象とする状態（通常、複数のパラメータで表される）が既知の初期値から一定の規則で経時的に変化してゆく系に対し、規則性を乱す因子（外乱）が加わり、さらに観察の際の雑音も加わった結果がとらえられることを想定し、この結果から逆に真の状態を推定して行く手法である。このようなモデルは状態空間モデル(図1)と呼ばれる。雑音中の一定周波数の信号の検出に適用する場合は、状態空間モデルにおける状態 x は、信号の振幅と位相による表示ではなく複素数の実数部と虚数部で表示された 2 次元のベクトルで表す。

図1 状態空間モデル

数値パラメータで表わされる状態 x が時間とともに変化してゆき、それぞれの時刻における状態が反映された結果として観測値 y が観察されるというモデルである。 x は数値群としてベクトル（列ベクトル）で表される。時刻 $t-1$ における状態 x_{t-1} から次の時刻 t における状態 x_t への変化は、規則的な変化とそれに加わる不規則な変動（外乱）に分けて考え、規則的な変化は x_{t-1} から Fx_{t-1} へとベクトルを変換する行列として表わすことができる。この F を状態遷移行列と呼ぶ。この Fx_{t-1} に外乱が加わったものが x_t となる。状態 x と観測値 y の関係も、規則的な関係を観測行列 H として表し、時刻 t における状態 x_t が H により Hx_t に変換されこれに観測時の雑音を加わったものが観測値 y_t となる。

また、状態遷移行列 F は時間の経過により x がいかに変化するかを表すもので、一定周波数の信号であれば経過時間分だけ位相が変化する（周波数 f に対する角速度 $2\pi f$ と経過時間との積だけ増加）ため、 F は x のベクトルをこの位相変化の相当する角度だけ回転させる操作を表す行列となる。これは三角関数の公式に基づき、経過時間を t とすると 1 行目は $(\cos 2\pi ft \sin 2\pi ft)$ 2 行目は $(-\sin 2\pi ft \cos 2\pi ft)$ の 2 行 2 列の行列となる。 x のうち実際に観察されるのは実数部であるから観測行列 H は x との積が実数部の値になるように $(1 \ 0)$ の行ベクトルとなる。 x の次元とそれに伴う F や H の行と列をさらに増やせば、複数周波数からなる信号の推定にも使用できる。DPOAE 測定に応用する場合は、観測値 y は測定された外耳道内音圧、状態 x は f_1 と f_2 および $2f_1-f_2$ の 3 つの正弦波信号となる。したがって x は 6 次元（各成分が実数部と虚数部で 2 次元ずつ）のベクトルであらわすことができる。 $2f_1-f_2$ 成分以外の歪成分も対象とする場合はさらに次元を増やせばよい。状態遷移行列 F はそれぞれの周波数に対する角速度 $(2\pi f_1, 2\pi f_2, 2\pi (2f_1-f_2))$ とサンプリング時間の積から計算する。

カルマンフィルタを用いて y から x を推定する

図2 カルマンゲインを用いて x_t の推定を行うためのブロック図

直前の時刻 $t-1$ における状態推定値が x_{t-1} として得られている場合、この x_{t-1} から状態遷移行列 F をもとに時刻 t の仮の推定値 $x_{t/t-1}$ を Fx_{t-1} として計算し、さらにこの値に基づき時刻 t における観測値の推定値 $\hat{y}_t$ を $Hx_{t/t-1}$ として計算する。時刻 t における実際の観測値 y_t と推定値 $\hat{y}_t$ を比較し、この差 $(y_t - \hat{y}_t)$ とカルマンゲイン K の積、 $K(y_t - \hat{y}_t)$ を x の仮の推定値 $x_{t/t-1}$ に加えることで仮の推定値を補正し、時刻 t における x の推定値 x_t を計算する。カルマンゲイン K は、原法では時刻 $t-1$ で用いた K をアルゴリズムにより時刻 t における K を計算し使用するが、本研究では定常カルマンゲインとして同一の K を使用している。

際には、図2のように直前の x の推定値から予想される y の値を計算し実際の観測値との誤差を求め、これにカルマンゲイン (K) と名づけられる値を掛けたもので x の推定値を補正するという計算を逐次繰り返す。このカルマンゲイン (K) の算出がカルマンフィルタの最も重要な部分であるが、アルゴリズムとしても確立されており具体的には成書^{6,7)}を参照していただきたい。一方、この計算を行うには逐次得られる y の値に加え、 x の初期値が必要である(カルマンゲインの計算にも x の初期値は必要である)。DPOAEに応用する場合初期値を知ることにはできないが、このような場合のために定常カルマンフィルタという方法があり、本研究ではこれを適用した。状態 x を推定してゆく操作を繰り返してゆくと、 x の初期値にかかわらずカルマンゲイン (K) が一定の値に収束してゆく性質があり、この収束値(定常カルマンゲイン)は y の値を使用せずに他のパラメータのみであらかじめ計算できる⁷⁾。この定常カルマンゲインを K として使用するのが定常カルマンフィルタの考えで、本研究ではこれを用いた。定常カルマンゲインを使用することにはこの他に、測定開始前にあらかじめ K を計算しておくことができるため計算時間が短縮できリアルタイムで結果を表示できるメリットもある。定常カルマンゲインの算出法には倍加アルゴリズム⁷⁾を使用した。なお、この後に記載する実験のうち、実験1は f_1 , f_2 , $2f_1-f_2$ 成分を測定対象としたために、 x は6つの要素で構成される6次元ベクトル、実験2ではこれに第3の刺激音の周波数成分 f_3 も測定対象に加わるため8次元ベクトルとなり、状態遷移行列 F も実験1で6行6列、実験2では8行8列、観測行列 H は実験1で1行6列(6次元の行ベクトルと同じ)、実験2では1行8列(8次元の行ベクトルと同じ)となっている。

3. 測定方法

測定には、DPOAE測定プローブとしてEtymotic Research社のER-10B、刺激に用いる正弦波信号の発生とマイクロホンの記録信号のデジタル化にTucker Davis社のシグナルプロセッサRM-1、シグナルプロセッサのコントロールおよび記録されたデジタル波形の処理にパーソナルコンピュータ

(SONY VAIO PCG-FX99V/BP)を使用した。測定用プログラムはナショナルインスツルメンツ社のLabVIEW version6.1を使用して自作した。プログラムは、シグナルプロセッサへのコントロールにより2つの刺激音の周波数(f_1 と f_2)ならびにこれらの音圧(L_1 と L_2)ならびに測定設定を行う部分、刺激開始のトリガーをかけシグナルプロセッサから記録データの取り込みを行う部分、取り込まれたデータに対し高速フーリエ変換もしくはカルマンフィルタによる処理を行う部分から構成した。最後の処理部分はソフトウェア的に切り替え可能とし、高速フーリエ変換モードとして出力音圧の校正を行ったのちカルマンフィルタモードとして時間波形記録を行う形とした。

具体的な測定としては、DPOAEが時間的変動する状態として、ふたつの刺激音を同時に振幅変調した場合(実験1)と刺激音圧は一定として第3音をsuppressor toneとして同時刺激し、第3音を振幅変調した場合(実験2)のふたつを行った。実験1と2ともに $f_1=3200\text{Hz}$ $f_2=4000\text{Hz}$ としsuppressor toneの周波数(f_3)= 5000Hz , L_1 (f_1 刺激の音圧)= 70dBSPL , L_2 (f_2 刺激の音圧)= 60dBSPL , suppressor toneの音圧(L_3)= 70dBSPL とした。実験1における刺激音および実験2におけるsuppressor toneの振幅変調周期はともに 5Hz , 記録時間は1秒とした。デジタル化におけるサンプリング周波数は 48kHz とした。測定対象は聴力正常な被験者2名(41歳男性と38歳女性)の一側耳とした。

予備実験において、 $2f_1-f_2$ 成分波形においてより音圧の大きい f_1 および f_2 成分の変動の影響による波形のゆらぎが見られたため、上記に加えシグナルプロセッサRM-1において $2f_1-f_2$ 周波数を中心周波数と帯域幅 5Hz とした1段双2次バンドパスフィルタによる、デジタルフィルタリングを前処理として行った。

4. 結 果

実験1の測定結果を図3と図4に示す。図3では音圧に表示にデシベルによる対数表示ではなく 0dBSPL を1とした場合の線形表示(0dB の何倍であるかで表す)とし、図4では音圧を対数スケールであるデシベルによる表示とした。線形表示では

図3 被験者1 (左) および被験者2 (右) における実験1の線形表示した測定結果 (線形表示)
 $f_1 = 4000\text{Hz}$, $f_2 = 4800\text{Hz}$, $L_1 = 70\text{dB SPL}$, $L_2 = 60\text{dB SPL}$ の刺激音を $5\text{Hz} 100\%$ の振幅変調した場合の f_1 , f_2 および DPOAE である $2f_1 - f_2$ の各成分の音圧を 0dB SPL を1としてこれの何倍であるかを表す線形スケールにて表示した。 f_1 および f_2 成分の波形が 5Hz の周期で正弦波的振幅変調となっていることが確認できる。

図4 被験者1 (左) および被験者2 (右) における実験1の測定結果 (デシベル表示)
 図3に示したものと同一の測定結果を音圧をデシベル表示にして示す。図4上の図では上に f_1 および f_2 成分は図3では正弦波波形であったものが上に凸の曲線状を呈している。ピークでの音圧は前処理でのデジタルフィルタリングの効果により、 f_1 成分が $24\text{--}26\text{dB}$, f_2 成分が $9\text{--}10\text{dB}$ で、 $2f_1 - f_2$ 成分とはほぼ同等の振幅レベルとなっている。 $2f_1 - f_2$ 成分は、 f_1 成分および f_2 成分の変動より 40 から 50 ミリ秒程度遅れて周期的に変動している。測定開始から 15 から 20 ミリ秒程度の間は波形は不安定である。

図5 被験者1 (左) および被験者2 (右) における実験2の測定結果
 $f_1 = 4000\text{Hz}$, $f_2 = 4800\text{Hz}$, $L_1 = 70\text{dB SPL}$, $L_2 = 60\text{dB SPL}$ で音圧一定とした刺激音とこれを抑圧する 5000Hz , 70dB SPL で $5\text{Hz} 100\%$ の振幅変調を加えた第3音を suppressor tone として同時刺激した際の f_1 , f_2 , f_3 (第3音の音圧) および DPOAE $2f_1 - f_2$ 周波数成分の時間変動を示す。 f_1 および f_2 成分は一定で、 f_3 成分は 5Hz の周期で変動し、 $2f_1 - f_2$ 成分は f_3 成分のピークから 40 から 50 ミリ秒程度遅れて最大に抑圧されている。実験1と同様に測定開始から概ね 20 ミリ秒程度の間は波形が不安定である。

刺激音である f_1 および f_2 成分の波形は 5 Hz の周期で正弦波的振幅変調となっていることが確認できた。一方、音圧をデシベル表示した図 4 では f_1 および f_2 成分は上に凸の曲線状を呈した。前処理でのデジタルフィルタリングの効果により、ピーク時の音圧は 2 被験者とも f_1 が 24-26 dB, f_2 が 9-10 dB で、刺激音の設定音圧 (L_1 と L_2) より小さく $2f_1-f_2$ 成分とはほぼ同等の振幅レベルとなっている。DPOAE の主たる成分である $2f_1-f_2$ 周波数成分は、ふたりの被験者で音圧変動の大きさに差はあるが、ともに f_1 成分および f_2 成分の変動より 40 から 50 ミリ秒程度遅れて周期的に変動した。測定開始から 15 から 20 ミリ秒程度の間の波形は不安定であった。

次に実験 2 の結果を図 5 に示す。実験 1 と同様に測定開始から概ね 20 ミリ秒を過ぎると波形は安定し、刺激音圧は一定であるため f_1 および f_2 成分は一定で実験 1 におけるピークレベルとほぼ同一であった。suppressor tone の周波数である f_3 成分は 5 Hz の周期で変動し、これに伴って $2f_1-f_2$ 成分は f_3 成分のピークから 40 から 50 ミリ秒程度遅れて最大に抑圧された。

5. 考 察

実験 1 の測定結果では、カルマンフィルタで得られた結果において刺激音圧変動のタイミングに一致して記録の刺激音の周波数成分の変動が記録され、これに一定の時間遅れで $2f_1-f_2$ DPOAE 成分の変動が記録されており、刺激音圧の変動に対する DPOAE の生理的変動を観察できていると考えられる。また実験 2 においても、suppressor tone の音圧変動から一定の時間遅れをもって $2f_1-f_2$ 周波数成分が変動しており、やはり DPOAE の生理的変動を観察できていると考えられる。また、この際に f_1 および f_2 周波数成分には変動は見られず、suppressor tone および DPOAE の変動の影響は受けていないこともわかる。これらの結果から、カルマンフィルタは DPOAE の時間変動を観察する目的に有用な技法であると言える。

DPOAE の時間変動の観察の手法としてこれまでに報告されている方法としては、Whitehead らによる位相の異なる刺激の加算により刺激音を消去し $2f_1-f_2$ 成分の時間波形を得る方法¹⁾、Talmadge らに

よる least-square fit (LSF) 法³⁾、Scholz らによる短時間フーリエ変換を用いた方法²⁾、Kim らによるヘテロダイン法⁹⁾がある。Whitehead らの方法は、 $2f_1-f_2$ DPOAE の位相が 2 つの刺激音の位相 ϕ_1 と ϕ_2 に対し $2\phi_1-\phi_2$ となることから ϕ_1 を 90 度ずつ ϕ_2 を 180 度ずつ変えた 8 つの組み合わせで刺激し加算することにより刺激音を消去し $2f_1-f_2$ 成分のみを残すというものである。Talmadge らの LSF 法は短く区切った時間ごとに $2f_1-f_2$ 成分について線形予測分析を行い誤差が最小となる値から振幅を推定するものである。線形予測分析 (linear predictive coding) とは、カルマンフィルタにも活用されているが、それ以前の測定値の線形結合として次の値を推定して行く手法である。短時間フーリエ変換は時間波形に窓関数をかけて対象とする時間を限定し高速フーリエ変換 (FFT) を行う手法である。Kim らのヘテロダイン法は時間波形を FFT の後 $2f_1-f_2$ 周囲の周波数成分のみを取り出し逆 FFT を行い、得られた解析信号から各時間における瞬時振幅を計算するものである。

これらの過去に報告された方法と比較し、本研究で用いたカルマンフィルタを応用した方法の優れた点としては、測定値が得られるごとに振幅データが得られる (オンラインとも言われる) ためリアルタイムに結果が得られる点と、希望する複数の周波数成分の変動が同時に検出できる点が挙げられる。Whitehead らの方法および Kim らのヘテロダイン法は測定時間全長にわたりデータを収集してから計算を行う方法 (オフラインと呼ばれる) であり、短時間フーリエ変換と LSF 法は逐次計算する方法ではあるが一定量のデータがまとまってから計算する形であり、リアルタイム性という点では劣る。また、本研究においては測定される主要な音響成分である f_1 , f_2 , $2f_1-f_2$ ならびに suppressor tone の成分を同時に検出できたが、Whitehead らの方法や解析信号を用いる Kim らの方法では原理的に $2f_1-f_2$ 成分しか検出できない。短時間フーリエ変換ではすべての周波数成分の検出が可能であるが、検出する必要のない成分まで計算するため効率が悪い。一方、Talmadge らの LSF 法では複数周波数の検出は意図されていないが、線形予測分析を使用しており複数周波数が検出となるように改変することも可能で

ある。

カルマンフィルタを DPOAE の時間変動測定への活用する具体的方法については、本研究で示した以外の方法も考えられる。カルマンの原法では初期値は既知であり、ここから順次カルマンゲインと各パラメータを計算してゆくが、初期値が未知の場合仮の初期値で一定の時間記録分を順次計算し、最終値の時点の推定値から逆行性に初期値まで戻って推定しなおすという方法（固定区間スムージング）がある⁷⁾。この方法は測定開始時～正確な推定値が得られるというメリットがあるが、リアルタイムの計算はできず計算量も大幅に増加するというデメリットがある。一方、定常カルマンゲインを使用せず、初期値を任意の値として原法のとおりに逐次カルマンゲインを計算するという方法も考えられるが、予備実験で両者を比較したところむしろ定常カルマンフィルタを使用した場合の方が、値がよりすみやかに安定した。初期値が不確定な値であるためこれを元にカルマンゲインを逐次計算しても、安定しにくいと考えられる。これらを踏まえると、本研究で使用方法はカルマンフィルタの活用法としては適当なものであると考えられる。

一方、この方法の問題点は測定開始直後の計算結果が不安定で誤差が大きいことである。これは、上に考察したようにカルマンフィルタの初期値に不確定な値を使用しているためやむを得ないものではあるが、立ち上がり時の DPOAE 時間波形の挙動が観察できないことは検査としてデメリットである。本研究では振幅変調刺激での測定を行ったが、その際に測定開始後一定時間が経過した後では振幅変調における波形の立ち上がり時に刺激音および DPOAE とともに正確に立ち上がりの様子が検出できた。さらに、追加実験として振幅変調の途中で刺激休止時間において断続的な刺激することを行ったが、やはり立ち上がり波形は検出できた。したがって、カルマンフィルタの値が安定するまでの時間にダミーの刺激を行い、刺激音の次の立ち上がり時からを本測定とすることにすれば測定開始時の誤差の影響は回避できると考える。

DPOAE についてはこれまでにさまざまな研究が行われ、測定手法としてはすでに成熟した感もあったが、リアルタイムで DPOAE の時間変動を観察

できるようになったことで、これまで研究対象にできなかった語音をはじめとする過渡的な音響に対する蝸牛の活動を明らかにするなど新たな展開が期待できると考える。

Measurement of distortion product otoacoustic emission fluctuation using a Kalman filtering approach

Tatsuhiko Harada, Jin Kanzaki

Department of Otolaryngology, International University of Health and Welfare Atami Hospital

A new measurement system for detecting fluctuations in distortion product otoacoustic emission (DPOAE) was developed using a Kalman filtering approach. With this system, DPOAE fluctuations in response to changing stimulus levels and to changing suppressor tone levels were measured in two normal hearing subjects. In the first experiment, the stimulus tones were 4000 Hz · 70 dB SPL and 4800 Hz · 60 dB SPL with a 5 Hz 100% amplitude modulation. In both subjects, the 2f₁-f₂ component of the DPOAE periodically changed in keeping with the cycle of stimulus levels but with a 40-50 millisecond delay. In the second experiment, the levels of the stimulus tones were kept constant, but a third tone of 5000 Hz · 70 dB SPL with a 5 Hz 100% amplitude modulation was added to the stimulus tones as a DPOAE suppressor. In both subjects, the 2f₁-f₂ component of the DPOAE was maximally suppressed from the peaks of the suppressor tone after a delay of about 40-50 milliseconds. These results indicate that a Kalman filtering approach is useful for detecting fluctuations in DPOAE. This approach has advantages over previously reported approaches in that it enables real-time processing and parallel processing to multiple frequency components. This approach will contribute to the elucidation of cochlear function in response to time, variable

tones, such as speech sounds.

参考文献

- 1) Whitehead ML, Stagner BB, Martin GK, et al : Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency, *J Acoust Soc Am* **100** : 1663-79, 1996
- 2) Scholz G, Hirschfelder A, Marquardt T, et al : Low-frequency modulation of the 2f1-f2 distortion product otoacoustic emissions in the human ear, *Hear Res* **130** : 189-96, 1999
- 3) Talmadge CL, Long GR, Tubis A, et al : Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emission, *J Acoust Soc Am* **105** : 275-292, 1999
- 4) James AL, Harrison RV, Pienkowski M, et al : Dynamics of real time DPOAE contralateral suppression in chinchillas and humans, *Int J Audiol* **44** : 118-29, 2005
- 5) Kalman RE : A New Approach to Linear Filtering and Prediction Problems. *Trans ASME, J Basic Eng* : Vol. 82D, 34-45, 1960
- 6) 西山 清 : 最適フィルタリング。培風館, 東京, 2001
- 7) 片山 徹 : 新版応用カルマンフィルタ。朝倉書店, 東京, 2001
- 8) Bimead RR, Tsoi AU, Parke PJ : A Kalman filtering approach to short-time Fourier analysis, *IEEE Trans Acoust, Speech, Signal Processing* : ASSP-34, 1493-1501, 1986
- 9) Kim DO, Dohn PA, Neely ST, et al : Adaptation of distortion product otoacoustic emission in humans, *J Assoc Res Otolaryngol* **2** : 31-40, 2001

(原稿受付 平成18. 6. 20)

別冊請求先 〒413-0012 静岡県熱海市東海岸町
13-1
国際医療福祉大学附属熱海病院耳鼻咽喉科

原田 竜彦

Reprint request :

Tatsuhiko Harada
Department of Otolaryngology, International University of Health and Welfare Atami Hospital, 13-1 Higashi-Kaigan-cho Atami, Shizuoka, 413-0012, Japan