

寄贈	
梅本敏孝氏	平成 年 月 日

14
1270
1993
①

適応フィルタによる短時間スペクトル分析システム

および

小振幅周期性雑音除去システムの構築とその応用

1994年3月

梅本 敏孝

目次

序論	4
I 適応アルゴリズムによる短時間スペクトル分析法とその応用	6
1 適応アルゴリズムによる短時間スペクトル分析法	7
1.1 はじめに	7
1.2 すべての周波数成分に一つの適応アルゴリズムを適用する方法	7
1.2.1 適応アルゴリズムの導出	7
1.2.2 収束条件	14
1.2.3 周波数分析に必要なデータ数	18
1.3 各周波数成分ごとに適応アルゴリズムを適用する方法	19
1.3.1 適応アルゴリズムの導出	19
1.3.2 収束条件	22
1.3.3 適応離散フーリエ変換法の性能	26
1.4 二つのアルゴリズムの比較	31
1.4.1 分析精度	31
1.4.2 信号に対する追従性	33
1.5 結論	38
2 採譜処理への応用	39
2.1 採譜とは	39
2.2 採譜処理のための周波数分析法	41
2.3 採譜処理の過程	41

2.3.1	音価の決定	43
2.3.2	音階の同定法	44
2.3.3	拍子の決定	46
2.3.4	調子の決定	50
2.3.5	楽譜シンボルへの変換	52
2.3.6	楽譜の表示	53
2.4	実験	55
2.4.1	実験環境	55
2.4.2	実験条件	55
2.4.3	音階の同定精度	57
2.4.4	計算時間の比較	58
2.5	結論	62
3	音声の分析への応用	63
3.1	はじめに	63
3.2	周波数分析	63
3.2.1	周波数分析器	64
3.2.2	スケーリング	65
3.2.3	平滑化	66
3.3	視察による特徴抽出	68
3.3.1	セグメンテーション	68
3.3.2	フォルマント抽出	68
3.3.3	フォルマント移動量の算出	70
3.3.4	/m/, /n/ が混在している範囲の音声の検討	70
3.3.5	特徴のまとめ	71
3.4	結論	72
4	まとめ	75

II 適応フィルタによる小振幅周期性雑音除去システムの構築とその応用 **77**

1	適応フィルタ	78
1.1	はじめに	78
1.2	適応性とは	78
1.3	基本的な LMS アルゴリズムを用いた適応システム	79
1.4	LMS アルゴリズム	79
1.4.1	LMS アルゴリズムの導出	81
1.4.2	収束条件	82
1.4.3	時定数	86
1.4.4	過剰誤差	87
2	周期性雑音除去システム	88
2.1	はじめに	88
2.2	1 チャンネル適応雑音除去システムとその問題点	89
2.3	ϵ -分離非線形デジタルフィルタの原理	90
2.4	小振幅周期性雑音除去システム	92
2.5	シミュレーションと実験	94
2.5.1	シミュレーション	94
2.5.2	実験	95
3	結論	102
	付録	103
A	複素スペクトルを用いた周波数補正法	103
A.1	連続フーリエ変換の位相特性	103
A.2	離散フーリエ変換の位相特性	105
A.3	ピーク周波数同定の方法	106
	参考文献	108
	謝辞	111

序論

適応性を持つシステムが種々の分野で着目されている。これは、LSI 技術進歩により性能のよい A/D 変換器、D/A 変換器、メモリ、マイクロコンピュータ、乗算器、除算器、加算器などが安価にかつ簡単に入手できるようになり、コスト的にも比較的容易に適応デジタルフィルタが実現できるようになってきたことや、従来の技術では困難であったことが適応信号処理技術の導入によって可能となったことなどが理由として考えられる。例えば、適応フィルタを用いた自動等化器やエコーキャンセラは多くの研究がなされており、実用化もされている。また、適応フィルタは適応予測符号化、ゴーストキャンセラ、ノイズキャンセラ、ハウリングキャンセラなどとしても多くの研究がなされている。しかし、周波数分析に関する研究や非線形フィルタを組み合わせた雑音除去システムに関する研究はあまり行なわれていない。これは、周波数分析の分野では高速フーリエ変換法という高速でかつ、安定した周波数分析法があるからである。しかし、高速フーリエ変換法にも種々の欠点がある。例えば、求まる周波数成分の間隔が一定であることや求まる周波数成分が観測窓内の信号の平均的なものであることが挙げられる。このため、採譜などに応用するには周波数の補正という問題や信号に対する追従性という問題が生ずる。また、非線形フィルタを組み合わせた雑音除去システムに関する分野においては、非線形フィルタの理論があまり確立されておらず、その応用に関する研究もあまり行なわれていないことが理由として考えられる。

そこで、本論文では適応フィルタの応用としてこれら二つの分野に注目する。周波数分析に関する分野では適応フィルタを用いた新しい周波数分析法を提案し、高速フーリエ変換法の上述のような欠点を克服できることを示す。また、非線形フィルタと組み合わせた雑音除去システムに関する分野では、 ϵ -分離非線形デジタルフィルタを組み合わせた適応雑音除去システムを提案

し、画像信号中に加わった周期性雑音の除去に有効であることを示す。

本論文は2部構成となっている。

第1部では、取り扱う信号を帯域制限周期波形に限定し、すべての周波数成分に対して一つの適応アルゴリズムを適用した場合と各周波数成分ごとに適応アルゴリズムを適用した場合について短時間スペクトル分析法を提案し、その収束条件を係数および学習曲線の収束条件から求め、本アルゴリズムの安定性を証明し、また、各アルゴリズムについてその性能を周波数分析に必要なデータ数、周波数分解能、分析精度などによって求め、各アルゴリズムの利点をシミュレーションによって比較する。また、採譜処理と音声分析における周波数分析法の二つの例に応用することによって提案した適応アルゴリズムによる短時間スペクトル分析法の有効性を示す。

第2部では、適応フィルタに非線形フィルタの一種である ϵ -分離非線形デジタルフィルタを組み合わせた雑音除去システムを提案し、従来の1チャンネル適応雑音除去システムでは、雑音除去が困難であった画像信号に加わる電源からのハム、隣接チャンネルからの妨害、色副搬送波と音声搬送波とのビート妨害、他の電波源からの妨害などによって生じる小振幅の周期性雑音の除去に有効であることをシミュレーションや実験によって示す。

また、付録として採譜処理における周波数分析法で提案した適応アルゴリズムによる短時間スペクトル分析法と比較するために用いた複素スペクトルを用いた周波数補正法について簡単に述べる。

第 I 部

適応アルゴリズムによる短時間スペクトル分 析法とその応用

第 1 章

適応アルゴリズムによる短時間スペクトル分析法

1.1 はじめに

高速フーリエ変換法の欠点は、求まる周波数成分の間隔が一定であることや求まる周波数成分が観測窓内の信号の平均的なものであることなどが挙げられる。一方、採譜などへの応用を考えた場合、周波数軸上の分析間隔も自由に選択できる短時間スペクトル分析が必要となる。そこで、本論文では取り扱う信号を帯域制限周期波形に限定し、すべての周波数成分に対して一つの適応アルゴリズムを適用した場合と各周波数成分ごとに適応アルゴリズムを適用した場合の周波数分析法を提案し、その収束条件を求めるとともに、その性能を検討する。

1.2 すべての周波数成分に一つの適応アルゴリズムを適用する方法

1.2.1 適応アルゴリズムの導出

連続時間関数 $g(t)$ は、連続フーリエ係数 $G(f)$ を用いればつぎのように表すことができる。

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df \\ &= \int_{-\infty}^0 G(f) \exp(j2\pi ft) df + \int_0^{\infty} G(f) \exp(j2\pi ft) df \\ &= - \int_{\infty}^0 G(-f) \exp(-j2\pi ft) df + \int_0^{\infty} G(f) \exp(j2\pi ft) df \end{aligned}$$

$$= \int_0^{\infty} \{G(f) \exp(j2\pi ft) + G(-f) \exp(-j2\pi ft)\} df \quad (1.1)$$

ここで、 $g(t)$ が $f < f_1$ と $f > f_L$ ($f_1 < f_L$) の区間でスペクトル成分を持たない帯域制限周期波形であるとするれば、 $f < f_1$ と $f > f_L$ の区間で $G(f) = 0$ となるので $G(f) \exp(j2\pi ft) = 0$ となり、式(1.1)は次式のように表すことができる。

$$g(t) = \int_{f_1}^{f_L} \{G(f) \exp(j2\pi ft) + G(-f) \exp(-j2\pi ft)\} df \quad (1.2)$$

しかし、連続系で計算することは不可能であるので離散系を用いて近似すれば、式(1.2)の積分式は和の形で近似することができる。連続時間関数 $g(t)$ が実数ならばフーリエ係数 $G(f)$ は複素共役になるので、式(1.2)はつぎのように近似することができる。

$$\bar{g}(t) = \sum_{i=1}^L \{G_i \exp(j2\pi f_i t) + \bar{G}_i \exp(-j2\pi f_i t)\} \quad (1.3)$$

式(1.3)は $\bar{g}(t)$ のサンプリングデータ $\hat{g}(k)$ (正確には $\hat{g}(k\Delta T)$: ΔT はサンプリング周期) に対しても成り立つのでつぎのように表すことができる。

$$\hat{g}(k) = \sum_{i=1}^L \{G_i \exp(j2\pi f_i k\Delta T) + \bar{G}_i \exp(-j2\pi f_i k\Delta T)\} \quad (1.4)$$

ここで、フーリエ係数 G と入力ベクトル $X(k)$ をつぎのように定義する。

$$G = [G_1, \dots, G_i, \dots, G_L]^T \quad (1.5)$$

$$X(k) = [\exp(j2\pi f_1 k\Delta T), \dots, \exp(j2\pi f_i k\Delta T), \dots, \exp(j2\pi f_L k\Delta T)]^T \quad (1.6)$$

これらを用いると、式(1.4)は、

$$\hat{g}(k) = X^T(k)G + \bar{X}^T(k)\bar{G} = G^T X(k) + \bar{G}^T \bar{X}(k) \quad (1.7)$$

と表すことができる。ここで入力ベクトル $X(k)$ とフーリエ係数 G はつぎに表すような複素数である。

$$G = G_R + jG_J \quad (1.8)$$

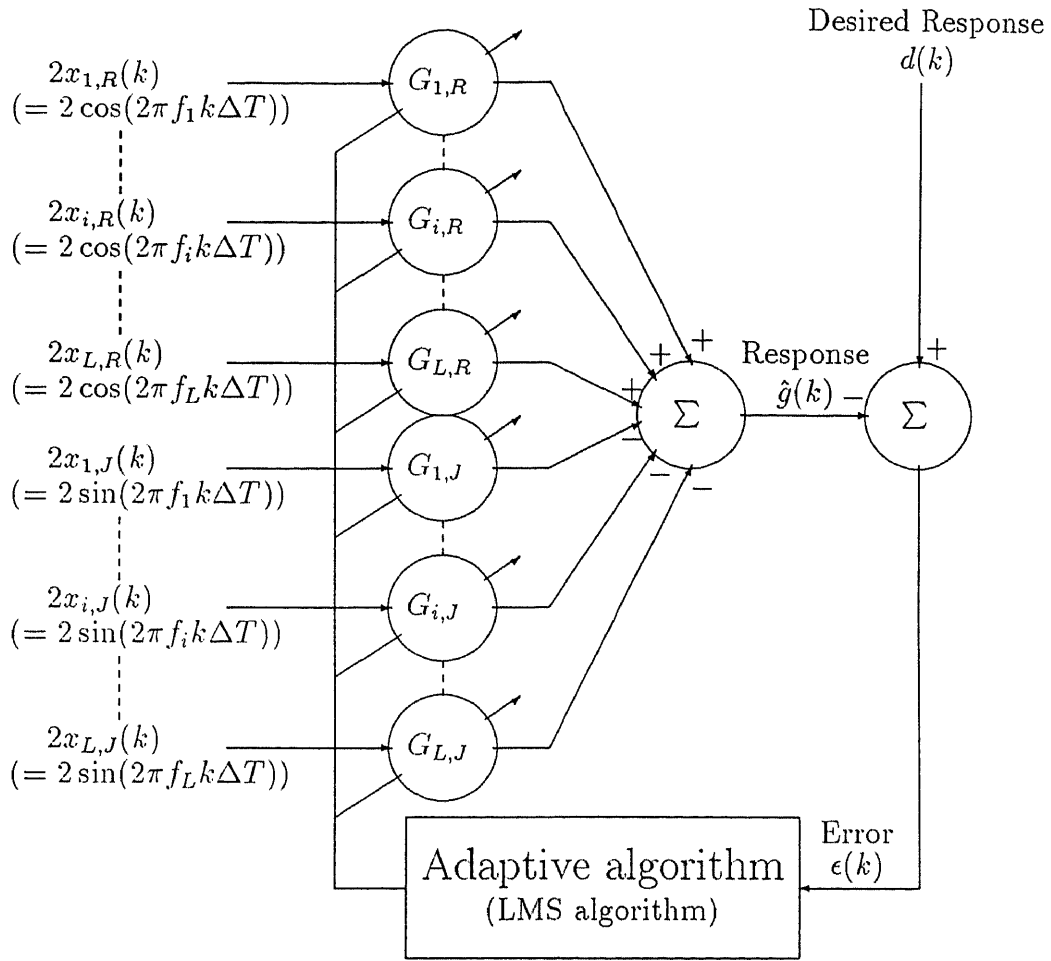


図 1.1: 適応アルゴリズムによる周波数分析システム

$$\mathbf{X}(k) = \mathbf{X}_R(k) + j\mathbf{X}_J(k) \tag{1.9}$$

このように、上述の $\mathbf{X}(k)$ を入力ベクトル、フーリエ係数 G を重みベクトルと考えれば、一種のフィルタと考えることができ、図 1.1 に示すような適応フィルタを構成すれば、周波数分析が可能であると考えられる。

図 1.1 において、フーリエ変換を行ないたい信号を $d(k)$ とすれば、瞬時誤差 $\epsilon(k)$ は、 $d(k)$ とフィルタ出力 $\hat{g}(k)$ の差からつぎのように表すことができる。

$$\epsilon(k) = d(k) - \hat{g}(k)$$

$$\begin{aligned}
&= d(k) - \{G^T X(k) + \bar{G}^T \bar{X}(k)\} \\
&= d(k) - 2G_R^T X_R(k) + 2G_J^T X_J(k) \quad (1.10)
\end{aligned}$$

これより、2乗平均誤差 $E[\epsilon^2(k)]$ はつぎのようになる。

$$\begin{aligned}
\text{MSE} &= E[\epsilon^2(k)] \\
&= E[\{d(k) - 2G_R^T X_R(k) + 2G_J^T X_J(k)\}^2] \\
&= E[d^2(k)] - 4E[d(k)X_R^T(k)]G_R \\
&\quad + 4E[d(k)X_J^T(k)]G_J \\
&\quad + 4G_R^T E[X_R(k)X_R^T(k)]G_R \\
&\quad + 4G_J^T E[X_R(k)X_R^T(k)]G_J \\
&\quad - 8G_R^T E[X_R(k)X_J^T(k)]G_J \quad (1.11)
\end{aligned}$$

式(1.11)は G_R および G_J についての2次関数であり、最小値を有する。 G_R および G_J について偏微分すれば、つぎのようになる。

$$\begin{aligned}
\nabla_R &= \frac{\partial \text{MSE}}{\partial G_R} \\
&= -4E[d(k)X_R(k)] \\
&\quad + 8E[X_R(k)X_R^T(k)]G_R \\
&\quad - 8E[X_R(k)X_J^T(k)]G_J \quad (1.12)
\end{aligned}$$

$$\begin{aligned}
\nabla_J &= \frac{\partial \text{MSE}}{\partial G_J} \\
&= 4E[d(k)X_J(k)] \\
&\quad + 8E[X_J(k)X_J^T(k)]G_J \\
&\quad - 8E[X_R(k)X_J^T(k)]G_R \quad (1.13)
\end{aligned}$$

ここで、 $E[X_R(k)X_R^T(k)]$ 、 $E[X_J(k)X_J^T(k)]$ および $E[X_R(k)X_J^T(k)]$ はつぎのようになる。

$$E[X_R(k)X_R^T(k)]$$

$$= \frac{1}{2} \mathbf{I} \begin{bmatrix} \sin(2\pi f_1 k \Delta T) \sin(2\pi f_L k \Delta T) \\ \vdots \\ \vdots \\ \vdots \\ \sin^2(2\pi f_L k \Delta T) \end{bmatrix} \quad (1.15)$$

$$\begin{aligned} & \mathbf{E}[\mathbf{X}_R(k) \mathbf{X}_J^T(k)] \\ &= \mathbf{E} \left[\begin{pmatrix} \cos(2\pi f_1 k \Delta T) \\ \vdots \\ \cos(2\pi f_i k \Delta T) \\ \vdots \\ \cos(2\pi f_L k \Delta T) \end{pmatrix} (\sin(2\pi f_1 k \Delta T), \dots, \sin(2\pi f_i k \Delta T), \dots, \sin(2\pi f_L k \Delta T)) \right] \\ &= \mathbf{E} \begin{bmatrix} \cos(2\pi f_1 k \Delta T) \sin(2\pi f_1 k \Delta T) & \dots & \cos(2\pi f_1 k \Delta T) \sin(2\pi f_i k \Delta T) & \dots \\ \vdots & \dots & \dots & \dots \\ \vdots & \dots & \cos(2\pi f_i k \Delta T) \sin(2\pi f_i k \Delta T) & \dots \\ \vdots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ & & \cos(2\pi f_1 k \Delta T) \sin(2\pi f_L k \Delta T) \\ & & \vdots \\ & & \vdots \\ & & \vdots \\ & & \cos(2\pi f_L k \Delta T) \sin(2\pi f_L k \Delta T) \end{bmatrix} \\ &= 0 \end{aligned} \quad (1.16)$$

以上のことから最適解の実部 G_R^* と虚部 G_J^* はつぎのようになる。

$$G_R^* = \mathbf{E}[d(k) \mathbf{X}_R(k)] \quad (1.17)$$

$$G_J^* = -\mathbf{E}[d(k) \mathbf{X}_J(k)] \quad (1.18)$$

ここで、最適解 G^* は $G^* = G_R^* + jG_J^*$ であるので式 (1.17), (1.18) はつぎのようにまとめることができる。

$$G^* = E[d(k)\overline{X}(k)] \quad (1.19)$$

この式は離散フーリエ変換の式であり、図 1.1 で示されるシステムによる最適解は離散フーリエ係数となる。しかし、この方法を用いてスペクトルを直接的に求めるには、入力の複素共役 $\overline{X}(k)$ の時間関数列とフーリエ変換を行ないたい信号 $d(k)$ との相互相関ベクトル $E[d(k)\overline{X}(k)]$ を求める必要があり、そのため計算量が多くなり、あまり効果的な方法とはいえない。そこで、適応アルゴリズムを用いることにする。この方法ではフーリエ係数 G を指定した初期値から出発して、フーリエ変換を行ないたい信号 $d(k)$ と入力 $X(k)$ の新しいサンプル値を受け取る時刻ごとに更新していく。そして、この適応は係数が最適値に十分近づくまで続けられる。通常、適応アルゴリズムとして用いられるのは最急降下法である。しかし、最急降下法では 2 乗平均誤差から各繰り返しごとの勾配ベクトル $\nabla(k)$ を正確に測定しなければならず、実際に行なうことは多くの計算を必要とし、困難であるので、勾配ベクトル $\nabla(k)$ を有限個の入力データサンプルから推定しなければならない。そこで、本節ではその一つの推定法として LMS アルゴリズムを用いることにする。この方法の利点は、2 乗平均誤差の推定量として瞬時 2 乗誤差そのものを用いて勾配ベクトルの推定量 $\widehat{\nabla}(k)$ を求めるので、相関関数の測定を必要としないことである。勾配ベクトルの推定量 $\widehat{\nabla}_R(k)$ および $\widehat{\nabla}_J(k)$ はつぎのようになる。

$$\widehat{\nabla}_R(k) = -4\epsilon(k)X_R(k) \quad (1.20)$$

$$\widehat{\nabla}_J(k) = 4\epsilon(k)X_J(k) \quad (1.21)$$

LMS アルゴリズムにおいては、それぞれの負の勾配ベクトルに沿って重みベクトルの実部と虚部を変化させるのでつぎのように定式化することができる。

$$G_R(k+1) = G_R(k) + \mu(-\widehat{\nabla}_R(k)) \quad (1.22)$$

$$G_J(k+1) = G_J(k) + \mu(\widehat{\nabla}_J(k)) \quad (1.23)$$

ここで μ はステップサイズパラメータと呼ばれる正定値である。これらの式に式 (1.20), (1.21) を代入すると式 (1.22), (1.23) はつぎのようになる。

$$G_R(k+1) = G_R(k) + 4\mu\epsilon(k)X_R(k) \quad (1.24)$$

$$\mathbf{G}_J(k+1) = \mathbf{G}_J(k) - 4\mu\epsilon(k)\mathbf{X}_J(k) \quad (1.25)$$

係数ベクトル $\mathbf{G}(k)$ は $\mathbf{G}(k) = \mathbf{G}_R(k) + j\mathbf{G}_J(k)$ であるので、式(1.24)、(1.25)はつぎのようにまとめることができる。

$$\mathbf{G}(k+1) = \mathbf{G}(k) + 4\mu\epsilon(k)\overline{\mathbf{X}}(k) \quad (1.26)$$

また、瞬時誤差 $\epsilon(k)$ はつぎのような式で与えられる。

$$\epsilon(k) = d(k) - \{\mathbf{G}^T(k)\mathbf{X}(k) + \overline{\mathbf{G}}^T(k)\overline{\mathbf{X}}(k)\} \quad (1.27)$$

周波数分析を行なうための LMS アルゴリズムはこれら簡単な二つの式で記述することができる。

1.2.2 収束条件

係数の収束条件

LMS アルゴリズムは厳密には非定常確率過程として取り扱う必要がある。しかし、解析を簡単化するため、係数ベクトル $\mathbf{G}(k)$ 、入力ベクトル $\mathbf{X}(k)$ やフーリエ変換を行ないたい信号 $d(k)$ が互いに統計的に独立であると仮定する。まず、誤差信号 $\epsilon(k)$ はつぎのように表すことができる。

$$\begin{aligned} \epsilon(k) &= d(k) - \{\mathbf{G}^T(k)\mathbf{X}(k) + \overline{\mathbf{G}}^T(k)\overline{\mathbf{X}}(k)\} \\ &= d(k) - 2\mathbf{G}_R^T(k)\mathbf{X}_R(k) + 2\mathbf{G}_J^T(k)\mathbf{X}_J(k) \end{aligned} \quad (1.28)$$

この式を式(1.24)に代入するとつぎのようになる。

$$\begin{aligned} \mathbf{G}_R(k+1) &= \mathbf{G}_R(k) + 4\mu\mathbf{X}_R(k)\{d(k) - 2\mathbf{G}_R^T(k)\mathbf{X}_R(k) \\ &\quad + 2\mathbf{G}_J^T(k)\mathbf{X}_J(k)\} \\ &= \{\mathbf{I} - 8\mu\mathbf{X}_R(k)\mathbf{X}_R^T(k)\}\mathbf{G}_R(k) + 4\mu d(k)\mathbf{X}_R(k) \\ &\quad + 8\mu\mathbf{X}_R(k)\mathbf{X}_J^T(k)\mathbf{G}_J(k) \end{aligned} \quad (1.29)$$

同様に式(1.25)に代入するとつぎのようになる。

$$\begin{aligned} \mathbf{G}_J(k+1) &= \{\mathbf{I} - 8\mu\mathbf{X}_J(k)\mathbf{X}_J^T(k)\}\mathbf{G}_J(k) - 4\mu d(k)\mathbf{X}_J(k) \\ &\quad + 8\mu\mathbf{X}_R(k)\mathbf{X}_J^T(k)\mathbf{G}_R(k) \end{aligned} \quad (1.30)$$

ここで、複素係数誤差ベクトル $C(k) = C_R(k) + jC_J(k)$ をつぎのように定義する。

$$C(k) = G(k) - G^* \quad (1.31)$$

k が無限大になる時、複素係数誤差ベクトル $C(k)$ が零ベクトルになれば、係数が収束したことを意味する。式 (1.31) の実部を式 (1.29) に代入するとつぎのようになる。

$$\begin{aligned} G_R(k+1) &= \{I - 8\mu X_R(k)X_R^T(k)\}\{C_R(k) + G_R^*\} + 4\mu d(k)X_R(k) \\ &\quad + 8\mu X_R(k)X_J^T(k)G_J(k) \\ &= \{I - 8\mu X_R(k)X_R^T(k)\}C_R(k) + G_R^* \\ &\quad + 4\mu\{d(k)X_R(k) - 2X_R(k)X_R^T(k)G_R^*\} \\ &\quad + 8\mu X_R(k)X_J^T(k)G_J(k) \end{aligned} \quad (1.32)$$

同様に、式 (1.31) の虚部を式 (1.30) に代入するとつぎのようになる。

$$\begin{aligned} G_J(k+1) &= \{I - 8\mu X_J(k)X_J^T(k)\}C_J(k) + G_J^* \\ &\quad - 4\mu\{d(k)X_J(k) + 2X_J(k)X_J^T(k)G_J^*\} \\ &\quad + 8\mu X_R(k)X_J^T(k)G_R(k) \end{aligned} \quad (1.33)$$

ここで、最適解の実部 G_R^* を左辺に移項すると、 $G_R(k+1) - G_R^*$ は係数誤差ベクトルの更新値となり、式 (1.32) はつぎのようになる。

$$\begin{aligned} C_R(k+1) &= \{I - 8\mu X_R(k)X_R^T(k)\}C_R(k) \\ &\quad + 4\mu\{d(k)X_R(k) - 2X_R(k)X_R^T(k)G_R^*\} \\ &\quad + 8\mu X_R(k)X_J^T(k)G_J(k) \end{aligned} \quad (1.34)$$

同様に、式 (1.33) もつぎのようになる。

$$\begin{aligned} C_J(k+1) &= \{I - 8\mu X_J(k)X_J^T(k)\}C_J(k) \\ &\quad - 4\mu\{d(k)X_J(k) + 2X_J(k)X_J^T(k)G_J^*\} \\ &\quad + 8\mu X_R(k)X_J^T(k)G_R(k) \end{aligned} \quad (1.35)$$

しかし、LMS アルゴリズムは雑音を含む勾配ベクトルの推定量によっているためその収束条件を上式そのものから求めることは不可能である。一般に LMS

アルゴリズムにおいては集合平均の動的振舞いが議論されるのでここでも、平均係数誤差ベクトル $E[C(k)]$ の収束条件を求めることにする。係数ベクトル $G(k)$ 、入力ベクトル $X(k)$ やフーリエ変換を行ないたい信号 $d(k)$ が互いに統計的に独立であると仮定しているので、係数誤差ベクトル $C(k)$ も入力ベクトル $X(k)$ やフーリエ変換を行ないたい信号 $d(k)$ と統計的に独立である。そこで式 (1.34), (1.35) の両辺の期待値をとれば、つぎのような二つの関係が得られる。

$$\begin{aligned} E[C_R(k+1)] &= E[\mathbf{I} - 8\mu \mathbf{X}_R(k) \mathbf{X}_R^T(k)] E[C_R(k)] \\ &\quad + 4\mu E[d(k) \mathbf{X}_R(k)] - 8\mu E[\mathbf{X}_R(k) \mathbf{X}_R^T(k)] \mathbf{G}_R^* \\ &\quad + 8\mu E[\mathbf{X}_R(k) \mathbf{X}_J^T(k)] E[\mathbf{G}_J(k)] \end{aligned} \quad (1.36)$$

$$\begin{aligned} E[C_J(k+1)] &= E[\{\mathbf{I} - 8\mu \mathbf{X}_J(k) \mathbf{X}_J^T(k)\} E[C_J(k)] \\ &\quad - 4\mu E[d(k) \mathbf{X}_J(k)] + 8\mu E[\mathbf{X}_J(k) \mathbf{X}_J^T(k)] \mathbf{G}_J^* \\ &\quad + 8\mu E[\mathbf{X}_R(k) \mathbf{X}_J^T(k)] E[\mathbf{G}_R(k)] \end{aligned} \quad (1.37)$$

式 (1.14), (1.15), (1.16), (1.17), (1.18) から式 (1.36), (1.37) はつぎのような等比級数になる。

$$E[C_R(k+1)] = (1 - 4\mu) E[C_R(k)] \quad (1.38)$$

$$E[C_J(k+1)] = (1 - 4\mu) E[C_J(k)] \quad (1.39)$$

ここで、係数誤差ベクトル $C(k)$ は $C(k) = C_R(k) + j C_J(k)$ であるので、式 (1.38), (1.39) はつぎのようになる。

$$E[C(k+1)] = (1 - 4\mu) E[C(k)] \quad (1.40)$$

k が無限大に近づく時、ステップパラメータ μ がつぎの条件を満たせば、LMS アルゴリズムは収束する。つまり、平均係数誤差ベクトル $E[C(k)]$ は零に近づくので係数ベクトルも最適値に近づき、スペクトル分析が可能となる。

$$0 < |1 - 4\mu| < 1 \quad (1.41)$$

ステップサイズパラメータ μ からみれば、つぎの条件を満たせば LMS アルゴリズムは収束するので、スペクトル分析が可能となる。

$$0 < \mu < \frac{1}{4} \quad (1.42)$$

学習曲線の収束条件

周波数分析が可能になるためには係数ベクトルが収束するとともに、学習曲線も収束しなくてはならない。そこで次に学習曲線が収束するための条件を求める。前に述べたように LMS アルゴリズムは雑音を含んだ勾配ベクトルによっているため、ここでも 2 乗平均誤差 $E[\epsilon^2(k)]$ の動的振舞いから学習曲線が収束するための条件を求めることにする。2 乗平均誤差 $E[\epsilon^2(k)]$ を ξ とおくと最小 2 乗平均誤差 ξ_{min} はつぎのようになる。

$$\begin{aligned}\xi_{min} &= E[d^2(k)] - 4E[d(k)X_R^T(k)]G_R^* \\ &\quad + 4E[d(k)X_J^T(k)]G_J^* + 2G_R^{*T}G_R^* + 2G_J^{*T}G_J^*\end{aligned}\quad (1.43)$$

式(1.17), (1.18)を代入すれば、上式はつぎのようになる。

$$\begin{aligned}\xi_{min} &= E[d^2(k)] - 2G_R^{*T}G_R^* - 2G_J^{*T}G_J^* \\ &= E[d^2(k)] - 2G^{*T}\overline{G}^*\end{aligned}\quad (1.44)$$

パーセバルの定理より上式は、零となる。つぎに 2 乗平均誤差 ξ_k について求めていく。

$$\begin{aligned}\xi_k &= E[d^2(k)] - 4E[d(k)X_R^T(k)]E[G_R(k)] + 4E[d(k)X_J^T(k)]E[G_J(k)] \\ &\quad + 2E[G_R^T(k)G_R(k) + G_J^T(k)G_J(k)] \\ &= E[d^2(k)] - 4G_R^{*T}E[G_R(k)] - 4G_J^{*T}E[G_J(k)] \\ &\quad + 2E[G_R^T(k)G_R(k) + G_J^T(k)G_J(k)] \\ &= \xi_{min} + 2E[C_R^T(k)C_R(k) + C_J^T(k)C_J(k)] \\ &= \xi_{min} + 2E[C^T(k)\overline{C}(k)] \\ &= 2E[C^T(k)\overline{C}(k)]\end{aligned}\quad (1.45)$$

上式から 2 乗平均誤差 ξ_k は $2E[C^T(k)C(k)]$ になる。つまり、学習曲線が収束するためにはつぎのような条件を満たせばよいということになる。

$$E[C^T(k+1)\overline{C}(k+1)] - E[C^T(k)\overline{C}(k)] < 0 \quad (1.46)$$

この式に式(1.26)を代入するとつぎのようになる。

$$E[C^T(k+1)\overline{C}(k+1)] - E[C^T(k)\overline{C}(k)]$$

$$\begin{aligned}
&= E[\{\mathbf{G}(k) + 4\mu\epsilon(k)\bar{\mathbf{X}}(k) - \mathbf{G}^*\}^T \{\bar{\mathbf{G}}(k) + 4\mu\epsilon(k)\mathbf{X}(k) - \bar{\mathbf{G}}^*\} \\
&\quad - \{\mathbf{G}(k) - \mathbf{G}^*\}^T \{\bar{\mathbf{G}}(k) - \bar{\mathbf{G}}^*\}] \\
&= E[\{\mathbf{G}(k) - \mathbf{G}^*\}^T 4\mu\epsilon(k)\mathbf{X}(k) + \{\bar{\mathbf{G}}(k) - \bar{\mathbf{G}}^*\}^T 4\mu\epsilon(k)\bar{\mathbf{X}}(k) \\
&\quad + 16\mu^2\epsilon^2(k)\bar{\mathbf{X}}^T(k)\mathbf{X}(k)] \\
&= E[-4\mu\epsilon^2(k) + 16\mu^2\epsilon^2(k)\bar{\mathbf{X}}^T(k)\mathbf{X}(k)] \\
&= -4\mu E[\epsilon^2(k)\{1 - 4\mu\bar{\mathbf{X}}^T(k)\mathbf{X}(k)\}] \\
&= -4\mu(1 - 4\mu L)E[\epsilon^2(k)] < 0
\end{aligned} \tag{1.47}$$

ここで、 L は係数ベクトルの要素の数である。このことから、つぎのような関係を満たせば、LMS アルゴリズムは収束し、2乗平均誤差 ξ_k が零に近づき、学習曲線は収束するのでスペクトル分析が可能となる。

$$1 - 4\mu L < 0 \tag{1.48}$$

ステップサイズパラメータ μ からみれば、つぎのような関係を満たせば、LMS アルゴリズムは収束し、スペクトル分析が可能となる。

$$0 < \mu < \frac{1}{4L} \tag{1.49}$$

収束条件

以上のことからステップサイズパラメータ μ がつぎの条件を満たせば、LMS アルゴリズムは収束し、周波数分析を行なうことができる。

$$0 < \mu < \frac{1}{4L} \tag{1.50}$$

1.2.3 周波数分析に必要なデータ数

LMS アルゴリズムを用いているのでスペクトルを求めるに必要なデータ数はシステムがどの程度で収束するか、つまり学習曲線の時定数から求めることができる。2乗平均誤差 $E[\epsilon^2(k)]$ は式(1.45)で与えられ、式(1.40)を代入するとつぎのようになる。

$$\begin{aligned}
\xi_{i,k} &= 2E[\mathbf{C}^T(k)\bar{\mathbf{C}}(k)] \\
&= 2(1 - 4\mu_i)^{2k}E[\mathbf{C}^T(0)\bar{\mathbf{C}}(0)]
\end{aligned} \tag{1.51}$$

この式は項比 $(1 - 4\mu_i)^2$ の等比数列であり、指数関数で近似できる。時定数 τ はサンプリング周期を ΔT とすればつぎのようになる。

$$(1 - 4\mu)^2 = \exp\left(-\frac{\Delta T}{\tau}\right) = 1 - \frac{\Delta T}{\tau} + \left\{\frac{\Delta T}{2\tau}\right\}^2 - \dots \quad (1.52)$$

ここで、適応過程があまり速くない場合、つまり、ステップサイズパラメータ μ が小さい場合を考えれば、式 (1.52) はつぎのように近似することができる。

$$1 - 8\mu \simeq 1 - \frac{\Delta T}{\tau} \quad (1.53)$$

この式から時定数 τ はつぎのようになる。

$$\tau = \frac{\Delta T}{8\mu} \quad (1.54)$$

つまり、1.2.1 節で提案した周波数分析法はステップサイズパラメータ μ によって周波数分析に必要なフレームの長さを決めることができる。

1.3 各周波数成分ごとに適応アルゴリズムを適用する方法

1.3.1 適応アルゴリズムの導出

G_i を周波数成分 f_i におけるフーリエ係数とし、 $X_i(k)$ をつぎのように定義する。

$$X_i(k) = \exp(j2\pi f_i k \Delta T) \quad (1.55)$$

ここで、 ΔT はサンプリング周期である。また、 G_i と $X_i(k)$ は複素数であり、つぎのように表す。

$$G_i = G_{i,R} + jG_{i,J} \quad (1.56)$$

$$X_i(k) = X_{i,R}(k) + jX_{i,J}(k) \quad (1.57)$$

つぎに、実部のみからなる信号に対するフーリエ係数が複素共役になることを考慮し、図 1.2 のようなシステムを考える。 $d(k)$ を離散フーリエ変換を行ないたい信号とすると i 番目の出力 $g_i(k)$ との差はつぎのようになる。

$$\begin{aligned} \epsilon_i(k) &= d(k) - g_i(k) \\ &= d(k) - \{G_i X_i(k) + \overline{G_i} \overline{X_i(k)}\} \\ &= d(k) - 2G_{i,R} X_{i,R}(k) + 2G_{i,J} X_{i,J}(k) \end{aligned} \quad (1.58)$$

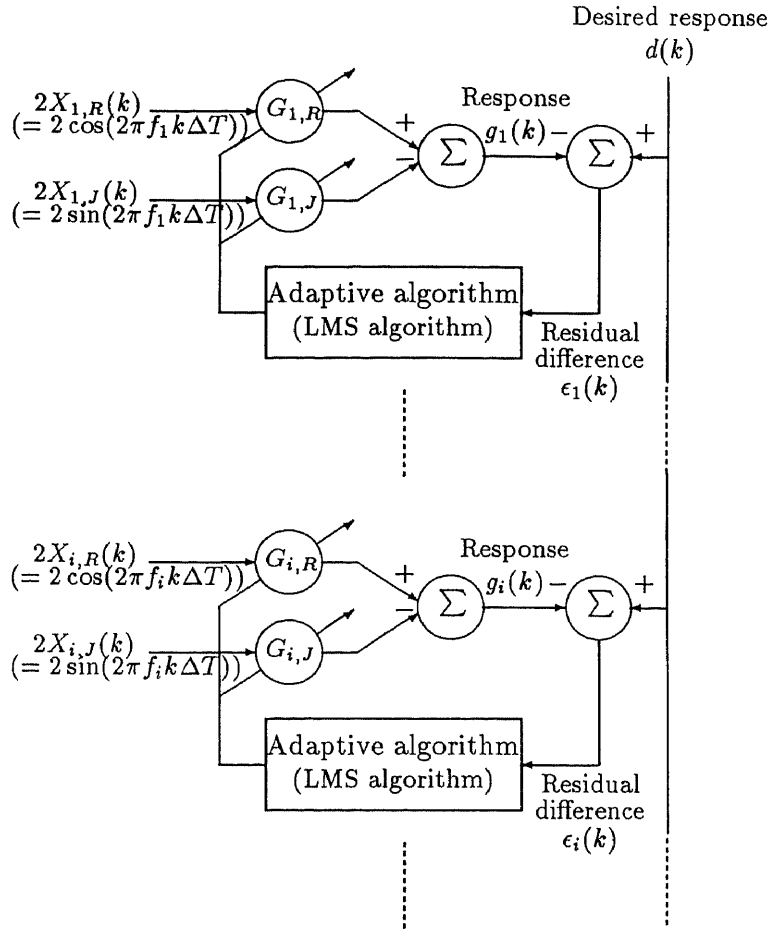


図 1.2: 適応アルゴリズムによる周波数分析システム

よって差の 2 乗平均 $E[\epsilon_i^2(k)]$ はつぎのようになる.

$$\begin{aligned}
 \text{MSE} &= E[\epsilon_i^2(k)] \\
 &= E[\{d(k) - 2G_{i,R}X_{i,R}(k) + 2G_{i,J}X_{i,J}(k)\}^2] \\
 &= E[d^2(k) - 4G_{i,R}E[d(k)X_{i,R}(k)] + 4G_{i,J}E[d(k)X_{i,J}(k)] \\
 &\quad - 8G_{i,R}G_{i,J}E[X_{i,R}(k)X_{i,J}(k)] + 2(G_{i,R}^2 + G_{i,J}^2)] \quad (1.59)
 \end{aligned}$$

式 (1.59) は $G_{i,J}$ および $G_{i,R}$ についての 2 次式であり, 最小値を有する. $G_{i,R}$

および $G_{i,J}$ で偏微分すれば、つぎのようになる。

$$\begin{aligned}\nabla_{i,R} &= \frac{\partial MSE}{\partial G_{i,R}} \\ &= -4E[d(k)X_{i,R}(k)] - 8G_{i,J}E[X_{i,R}(k)X_{i,J}(k)] + 4G_{i,R}\end{aligned}\quad (1.60)$$

$$\begin{aligned}\nabla_{i,J} &= \frac{\partial MSE}{\partial G_{i,J}} \\ &= 4E[d(k)X_{i,J}(k)] - 8G_{i,R}E[X_{i,R}(k)X_{i,J}(k)] + 4G_{i,J}\end{aligned}\quad (1.61)$$

ここで、 $E[X_{i,R}(k)X_{i,J}(k)] = 0$ であるので、最適解の実部 $G_{i,R}^*$ と虚部 $G_{i,J}^*$ はつぎのようになる。

$$G_{i,R}^* = E[d(k)X_{i,R}(k)] \quad (1.62)$$

$$G_{i,J}^* = -E[d(k)X_{i,J}(k)] \quad (1.63)$$

ここで、最適解 G_i^* は $G_i^* = G_{i,R}^* + jG_{i,J}^*$ であるので式 (1.62), (1.63) はつぎのようにまとめることができ、離散フーリエ変換の式と等価になる。

$$G_i^* = E[d(k)\bar{X}_i(k)] \quad (1.64)$$

このように図 1.2 で示されるシステムによる最適解は 1.2 節の場合と同様に離散フーリエ係数となる。しかし、1.2 節でも述べたようにこの方法を用いてスペクトルを直接的に求めるのは、あまり効果的な方法ではない。そこで、本節でも適応アルゴリズムの一つである LMS アルゴリズムを用いることにする。勾配の推定量 $\hat{\nabla}_{i,R}(k)$ および $\hat{\nabla}_{i,J}(k)$ はつぎのようになる。

$$\hat{\nabla}_{i,R}(k) = -4\epsilon_i(k)X_{i,R}(k) \quad (1.65)$$

$$\hat{\nabla}_{i,J}(k) = 4\epsilon_i(k)X_{i,J}(k) \quad (1.66)$$

このことから周波数分析を行なうための LMS アルゴリズムはつぎのように定式化することができる。

$$G_{i,R}(k+1) = G_{i,R}(k) + 4\mu_i\epsilon_i(k)X_{i,R}(k) \quad (1.67)$$

$$G_{i,J}(k+1) = G_{i,J}(k) - 4\mu_i \epsilon_i(k) X_{i,J}(k) \quad (1.68)$$

ここで μ_i は i 番目のステップサイズパラメータ (以下, 単にステップサイズパラメータと呼ぶ) と呼ばれる正定値であり, 1.2 節の方法とは異なり, 各周波数成分ごとに変えることができる. 係数ベクトル $G_i(k)$ は $G_i(k) = G_{i,R}(k) + jG_{i,J}(k)$ であるので, 式 (1.67), (1.68) はつぎのようにまとめることができる.

$$G_i(k+1) = G_i(k) + 4\mu_i \epsilon_i(k) \bar{X}_i(k) \quad (1.69)$$

また, 瞬時誤差 $\epsilon_i(k)$ はつぎのような式で与えられる.

$$\begin{aligned} \epsilon_i(k) &= d(k) - \{G_i(k)X_i(k) + \bar{G}_i(k)\bar{X}_i(k)\} \\ &= d(k) - 2G_{i,R}(k)X_{i,R}(k) + 2G_{i,J}(k)X_{i,J}(k) \end{aligned} \quad (1.70)$$

周波数分析を行なうための LMS アルゴリズムはこれら簡単な二つの式で記述することができる.

1.3.2 収束条件

係数の収束条件

1.2 節で述べたように LMS アルゴリズムは非定常確率過程として取り扱う必要があるが, 本節でも係数 $G_i(k)$, 入力 $X_i(k)$ やフーリエ変換を行ないたい信号 $d(k)$ が互いに統計的に独立であると仮定し, 解析する. 式 (1.67), (1.68) に式 (1.70) を代入すると係数はつぎのようになる.

$$\begin{aligned} G_{i,R}(k+1) &= G_{i,R}(k) + 4\mu_i X_{i,R}(k) \{d(k) - 2G_{i,R}X_{i,R}(k) + 2G_{i,J}X_{i,J}(k)\} \\ &= \{1 - 8\mu_i X_{i,R}^2(k)\} G_{i,R}(k) + 4\mu_i d(k) X_{i,R}(k) \\ &\quad + 8\mu_i X_{i,R}(k) X_{i,J}(k) G_{i,J}(k) \end{aligned} \quad (1.71)$$

$$\begin{aligned} G_{i,J}(k+1) &= G_{i,J}(k) + 4\mu_i X_{i,J}(k) \{d(k) - 2G_{i,R}X_{i,R}(k) + 2G_{i,J}X_{i,J}(k)\} \\ &= \{1 - 8\mu_i X_{i,J}^2(k)\} G_{i,J}(k) + 4\mu_i d(k) X_{i,J}(k) \\ &\quad + 8\mu_i X_{i,R}(k) X_{i,J}(k) G_{i,R}(k) \end{aligned} \quad (1.72)$$

ここで, 係数とその最適値との差をつぎのような式で定義する.

$$C_i(k) = G_i(k) - G_i^* \quad (1.73)$$

ここで、 $C_i(k)$ は $C_i(k) = C_{i,R}(k) + jC_{i,J}(k)$ なる複素数である。もし、 k が無限大になる時、 $C_i(k)$ が零になれば、係数が収束したことを意味する。そこで、式(1.71)、(1.72)に式(1.73)を代入するとつぎのようになる。

$$\begin{aligned} G_{i,R}(k+1) &= \{1 - 8\mu_i X_{i,R}^2(k)\}C_{i,R}(k) + G_{i,R}^* \\ &\quad + 4\mu_i\{d(k)X_{i,R}(k) - 2X_{i,R}^2(k)G_{i,R}^*\} \\ &\quad + 8\mu_i X_{i,R}(k)X_{i,J}(k)G_{i,J}(k) \end{aligned} \quad (1.74)$$

$$\begin{aligned} G_{i,J}(k+1) &= \{1 - 8\mu_i X_{i,J}^2(k)\}C_{i,J}(k) + G_{i,J}^* \\ &\quad + 4\mu_i\{d(k)X_{i,J}(k) - 2X_{i,J}^2(k)G_{i,J}^*\} \\ &\quad + 8\mu_i X_{i,R}(k)X_{i,J}(k)G_{i,R}(k) \end{aligned} \quad (1.75)$$

ここで、最適解の実部 $G_{i,R}^*$ と虚部 $G_{i,J}^*$ を左辺に移項すると、 $G_{i,R}(k+1) - G_{i,R}^*$ および $G_{i,J}(k+1) - G_{i,J}^*$ は係数と係数の最適値との差の更新値となり、つぎのように表すことができる。

$$\begin{aligned} C_{i,R}(k+1) &= \{1 - 8\mu_i X_{i,R}^2(k)\}C_{i,R}(k) \\ &\quad + 4\mu_i\{d(k)X_{i,R}(k) - 2X_{i,R}^2(k)G_{i,R}^*\} \\ &\quad + 8\mu_i X_{i,R}(k)X_{i,J}(k)G_{i,J}(k) \end{aligned} \quad (1.76)$$

$$\begin{aligned} C_{i,J}(k+1) &= \{1 - 8\mu_i X_{i,J}^2(k)\}C_{i,J}(k) \\ &\quad + 4\mu_i\{d(k)X_{i,J}(k) - 2X_{i,J}^2(k)G_{i,J}^*\} \\ &\quad + 8\mu_i X_{i,R}(k)X_{i,J}(k)G_{i,R}(k) \end{aligned} \quad (1.77)$$

しかし、1.2節でも述べたようにLMSアルゴリズムは雑音を含む勾配の推定量によっているためその収束条件を求めることは不可能であるので本節でも集合平均の動的振舞いを議論することにする。1.2節と同様に $C_i(k)$ も入力ベクトル $X_i(k)$ やフーリエ変換を行ないたい信号 $d(k)$ と統計的に独立であると考えられる。そこで式(1.76)、(1.77)の両辺の期待値をとれば、つぎのような二つの関係が得られる。

$$\begin{aligned} E[C_{i,R}(k+1)] &= E[1 - 8\mu_i X_{i,R}^2(k)]E[C_{i,R}(k)] \\ &\quad + 4\mu_i E[d(k)X_{i,R}(k)] - 8\mu_i E[X_{i,R}^2(k)]G_{i,R}^* \\ &\quad + 8\mu_i E[X_{i,R}(k)X_{i,J}(k)]E[G_{i,J}(k)] \end{aligned} \quad (1.78)$$

$$\begin{aligned}
E[C_{i,J}(k+1)] &= E[1 - 8\mu_i X_{i,J}^2(k)]E[C_{i,J}(k)] \\
&\quad + 4\mu_i E[d(k)X_{i,J}(k)] - 8\mu_i E[X_{i,J}^2(k)]G_{i,J}^* \\
&\quad + 8\mu_i E[X_{i,R}(k)X_{i,J}(k)]E[G_{i,R}(k)] \quad (1.79)
\end{aligned}$$

$E[X_{i,R}^2(k)] = \frac{1}{2}$, $E[X_{i,J}^2(k)] = \frac{1}{2}$, $E[X_{i,R}(k)X_{i,J}(k)] = 0$ であるので上式はつぎのような等比数列にまとめることができる.

$$\begin{aligned}
E[C_i(k+1)] &= (1 - 4\mu_i)E[C_i(k)] \\
&= (1 - 4\mu_i)^{k+1}E[C_i(0)] \quad (1.80)
\end{aligned}$$

k が無限大に近づく時, ステップパラメータ μ_i がつぎの条件を満たせば, LMS アルゴリズムは収束する. つまり, $E[C_i(k)]$ は零に近づくので係数ベクトルも最適値に近づき, スペクトル分析が可能となる.

$$0 < |1 - 4\mu_i| < 1 \quad (1.81)$$

ステップサイズパラメータ μ_i からみれば, つぎの条件を満たせば LMS アルゴリズムは収束するので, スペクトル分析が可能となる.

$$0 < \mu_i < \frac{1}{4} \quad (1.82)$$

学習曲線の収束条件

周波数分析が可能になるためには係数が収束するとともに, 学習曲線も収束しなくてはならない. そこでつぎに学習曲線が収束するための条件を求める. 1.2 節で述べたように LMS アルゴリズムは雑音を含んだ勾配によっているため, ここでも差の 2 乗平均 $E[e_i^2(k)]$ の動的振舞いから学習曲線が収束するための条件を求めることにする. まず, 初めに差の 2 乗平均の最小値を求める. 差の 2 乗平均 $E[e_i^2(k)]$ を ξ_i とおくと差の 2 乗平均の最小値 $\xi_{i,min}$ はつぎのようになる.

$$\begin{aligned}
\xi_{i,min} &= E[\{d(k) - 2G_{i,R}^*X_{i,R}(k) + 2G_{i,J}^*X_{i,J}(k)\}^2] \\
&= E[d^2(k)] - 4E[d(k)X_{i,R}(k)]G_{i,R} \\
&\quad + 4E[d(k)X_{i,J}(k)]G_{i,J} + 2\{G_{i,R}^{*2} + G_{i,J}^{*2}\} \\
&= E[d(k)^2] - 2\{G_{i,R}^{*2} + G_{i,J}^{*2}\} \\
&= E[d(k)^2] - 2G_i^* \overline{G_i^*} \quad (1.83)
\end{aligned}$$

1.2節の方法ではパーセバルの定理が成り立ち、差の2乗平均の最小値 $\xi_{i,min}$ は零になるが、この方法では零にはならず、収束しても (1.83) 式で与えられる差が残る。この式を用いて差の2乗平均 $\xi_{i,k}$ を求めるとつぎのようになる。

$$\begin{aligned}
\xi_{i,k} &= E[\{d(k) - 2G_{i,R}(k)X_{i,R}(k) + 2G_{i,J}(k)X_{i,J}(k)\}^2] \\
&= E[d(k)^2] - 4E[\{G_{i,R}^*G_{i,R}(k) + G_{i,J}^*G_{i,J}(k)\} + 2\{G_{i,R}^2(k) + G_{i,J}^2(k)\}] \\
&= \xi_{i,min} + 2E[G_i^*\bar{G}_i^* - 2G_i^*\bar{G}_i(k) - 2G_i(k)\bar{G}_i^* + 2G_i(k)\bar{G}_i(k)] \\
&= \xi_{i,min} + 2E[C_i(k)\bar{C}_i(k)] \tag{1.84}
\end{aligned}$$

このことから学習曲線が収束するためにはつぎのような条件を満たせばよいということになる。

$$E[C_i(k+1)\bar{C}_i(k+1)] - E[C_i(k)\bar{C}_i(k)] < 0 \tag{1.85}$$

この式に式 (1.69), (1.73) を代入するとつぎのようになる。

$$\begin{aligned}
&E[C_i(k+1)\bar{C}_i(k+1)] - E[C_i(k)\bar{C}_i(k)] \\
&= E[\{G_i(k) + 4\mu_i\epsilon_i(k)\bar{X}_i(k) - G_i^*\}\{\bar{G}_i(k) + 4\mu_i\epsilon_i(k)X_i(k) - \bar{G}_i^*\}] \\
&\quad - E[\{G_i(k) - G_i^*\}\{\bar{G}_i(k) - \bar{G}_i^*\}] \\
&= E[\{G_i(k) - G_i^*\}4\mu_i\epsilon_i(k)X_i(k) + \{\bar{G}_i(k) - \bar{G}_i^*\}4\mu_i\epsilon_i(k)\bar{X}_i(k) \\
&\quad + 16\mu_i\epsilon_i^2(k)X_i(k)\bar{X}_i(k)] \\
&= -4\mu_i(1 - 4\mu_i)E[\epsilon_i^2(k)] < 0 \tag{1.86}
\end{aligned}$$

このことから、つぎのような関係を満たせば、LMS アルゴリズムは収束し、2乗平均誤差 $\xi_{i,k}$ が $\xi_{i,min}$ に近づき、学習曲線は収束するのでスペクトル分析が可能となる。

$$1 - 4\mu_i < 0 \tag{1.87}$$

ステップサイズパラメータ μ_i からみれば、つぎのような関係を満たせば、LMS アルゴリズムは収束し、スペクトル分析が可能となる。

$$0 < \mu_i < \frac{1}{4} \tag{1.88}$$

収束条件

式(1.82), (1.88)よりステップサイズパラメータ μ_i が次の条件を満たせば, LMS アルゴリズムは安定となり, 周波数分析を行なうことができる.

$$0 < \mu_i < \frac{1}{4} \quad (1.89)$$

1.3.3 適応離散フーリエ変換法の性能

本節では, 1.3.1 節で提案した周波数分析法の性能をスペクトルを求めるのに必要なデータ数, 周波数分解能, 分析精度から評価する.

スペクトルを求めるのに必要なデータ数

LMS アルゴリズムを用いているのでスペクトルを求めるに必要なデータ数というのはシステムがどの程度で収束するか, つまり学習曲線の時定数から求めることができる. 差の2乗平均 $E[\epsilon_i^2(k)]$ は式(1.84)で与えられ, 式(1.80)を代入するとつぎのようになる.

$$\begin{aligned} \xi_{i,k} &= \xi_{i,min} + 2E[C_i(k)\bar{C}_i(k)] \\ &= \xi_{i,min} + 2(1 - 4\mu_i)^{2k}E[C_i(0)\bar{C}_i(0)] \end{aligned} \quad (1.90)$$

この式は項比 $(1 - 4\mu_i)^2$ の等比数列であり, 指数関数で近似できる. 時定数 τ_i はサンプリング周期を ΔT とすればつぎのようになる.

$$(1 - 4\mu_i)^2 = \exp\left(-\frac{\Delta T}{\tau_i}\right) = 1 - \frac{\Delta T}{\tau_i} + \left\{\frac{\Delta T}{2!\tau_i}\right\}^2 - \dots \quad (1.91)$$

ここで, 適応過程があまり速くない場合, つまり, ステップサイズパラメータ μ_i が小さい場合を考えれば, 式(1.91)はつぎのように近似することができる.

$$1 - 8\mu_i \simeq 1 - \frac{\Delta T}{\tau_i} \quad (1.92)$$

この式から時定数 τ_i はつぎのようになる.

$$\tau_i = \frac{\Delta T}{8\mu_i} \quad (1.93)$$

つまり, 1.3.1 節で提案した周波数分析法はステップサイズパラメータ μ_i によって周波数分析に必要なフレームの長さを決めることができる.

$$(1.95)$$

周波数分解能を周波数分析を行なっている周波数帯の中心周波数 f_c とそのスペクトル値が $\frac{1}{\sqrt{2}}$ になる周波数成分 f_1, f_2 の間の幅 $f_2 - f_1$ の比から評価する.

$$A = \frac{f_c}{f_2 - f_1} \quad (1.96)$$

一方, この方法は式 (1.95) からわかるように入力 $d(k)$ に $\bar{X}(k) = \exp(j2\pi f_i k \Delta T)$ を掛けているので, $d(k)$ の周波数成分を $-f_i$ だけシフトしていることと同値であると考えることができる. このことから, 式 (1.95) から求まる伝達関数の共振周波数 f_r や伝達関数 $H(f)$ の大きさが $\frac{|H(f_r)|}{\sqrt{2}}$ になる周波数成分 f'_1, f'_2 と f_c, f_1, f_2 との間にはつぎのような関係が成り立つと考えられる.

$$f_c = f_r + f_i \quad (1.97)$$

$$f_1 = f'_1 + f_i \quad (1.98)$$

$$f_2 = f'_2 + f_i \quad (1.99)$$

このことから, 式 (1.96) はつぎのようになる.

$$A = \frac{f_r + f_i}{f'_2 - f'_1} \quad (1.100)$$

ここで, 式 (1.94) の伝達関数を考えれば, つぎのようになる.

$$H(s) = \frac{4\mu_i z}{\{z - (1 - 4\mu_i)\}\{z - \exp(j2\pi f_i \Delta T)\}} \Big|_{z=\exp(s\Delta T)} \quad (1.101)$$

ΔT が十分小さい場合, $z = \exp(s\Delta T)$ はつぎのように近似することができる.

$$\begin{aligned} z &= \exp(s\Delta T) = 1 + s\Delta T + \frac{1}{2!}(s\Delta T)^2 + \dots \\ &\simeq 1 + s\Delta T \end{aligned} \quad (1.102)$$

よって, 式 (1.101) はつぎのように近似することができる.

$$H(s) = \frac{4\mu_i f_s}{(s\Delta T + 4\mu_i)(s + j2\pi f_i)} \quad (1.103)$$

ここで、 f_s はサンプリング周波数である。 μ_i が十分小さいので式 (1.103) の共振周波数 f_r は 0Hz 付近に存在すると考えられる。 0Hz 付近で $\frac{|H(f_r)|}{\sqrt{2}}$ になる周波数成分 f'_1 , f'_2 を求めると周波数分解能 A はつぎのようになる。

$$A \simeq \frac{\pi f_i}{4\mu_i f_s} \quad (1.104)$$

つまり、1.3.1 節で提案した周波数分析法はステップサイズパラメータ μ_i によって周波数分解能を決めることができる。

分析精度

すべての周波数成分に対して一つの LMS アルゴリズムを適用する場合のように差の 2 乗平均の最小値が零になるとはかぎらない。 また、LMS アルゴリズムを用いているため、勾配量の推定値には雑音が含まれ、多数回の適応を繰り返しても集合平均的に収束するだけで係数は最適解の近傍でゆらぎを生じ、差の 2 乗平均はその最小値より大きくなってしまふ。 このため、各周波数成分がどの程度の精度で求まるかが問題となる。 そこで、本論文では、分析精度を [係数とその最適解の差の分散]/[真のスペクトルパワー] で評価することにする。 まず、係数とその最適解の差の分散について考える。 係数とその最適解の差の分散は Widrow らによって導出されている。 この手順に従って、係数とその最適解の差の分散の導出を行なう。 まず、LMS アルゴリズムにおける勾配量の推定値 $\hat{\nabla}(k)$ は、真の勾配量 $\nabla(k)$ にノイズ $N(k)$ を含んだものであると考えられ、つぎのように表すことができる。

$$\hat{\nabla}(k) = \nabla(k) + N(k) \quad (1.105)$$

係数とその最適解に十分収束した場合を考えれば、真の勾配量 $\nabla(k)$ は勾配量の推定値 $\hat{\nabla}(k)$ に含まれるノイズ $N(k)$ に比べ十分に小さくなるので、上式はつぎのようになる。

$$\hat{\nabla}(k) = N(k) = -4\mu_i \bar{X}_i(k) \quad (1.106)$$

このことから、勾配量の推定値 $\hat{\nabla}(k)$ に含まれるノイズ $N(k)$ の分散はつぎのようになる。

$$E[|N(k)|^2] = E[N(k)\bar{N}(k)]$$

$$\begin{aligned}
&= 16E[\epsilon^2(k)]E[X(k)\bar{X}(k)] \\
&= 16E[\epsilon^2(k)] \\
&\simeq 16\xi_{i,min}
\end{aligned} \tag{1.107}$$

つぎに係数とその最適解の差の分散を考える。まず、式 (1.69) の両辺から最適解 G_i^* を引くとつぎのようになる。

$$C_i(k+1) = C_i(k) - \mu_i \nabla(k) - \mu_i N(k) \tag{1.108}$$

$\nabla(k)$ はつぎのようになる。

$$\begin{aligned}
\nabla(k) &= \frac{\partial E[\epsilon^2(k)]}{\partial G_i(k)} \\
&= 4\{G_i(k) - E[d(k)\bar{X}_i(k)]\} \\
&= 4\{G_i(k) - G_i^*\} \\
&= 4C_i(k)
\end{aligned} \tag{1.109}$$

よって、式 (1.108) はつぎのようになる。

$$C_i(k+1) = (1 - 4\mu_i)C_i(k) - \mu_i N(k) \tag{1.110}$$

このことから、係数とその最適解の差 $C_i(k)$ と勾配量の推定値 $\hat{\nabla}(k)$ に含まれるノイズ $N(k)$ が統計的に独立であるとすれば、つぎのようになる。

$$E[|C(k+1)|^2] = (1 - 4\mu_i)^2 E[|C(k)|^2] + \mu_i^2 E[|N(k)|^2] \tag{1.111}$$

ステップサイズパラメータ μ_i が十分小さく、係数とその最適解に十分収束した場合、 $E[|C(k+1)|^2] \simeq (1 - 4\mu_i)^2 E[|C(k)|^2]$ と考えることができるので、上式はつぎのようになる。

$$\begin{aligned}
E[|C(k)|^2] &= \frac{\mu_i}{8(1 - 2\mu_i)} E[|N(k)|^2] \\
&\simeq 2\mu_i \xi_{i,min} \\
&= 2\mu_i \{E[d^2(k)] - |G^*|^2\}
\end{aligned} \tag{1.112}$$

このことから、真のスペクトルパワーによって正規化された係数とその最適解の差の分散 B はつぎのようになる。

$$B = 2\mu_i \left\{ \frac{E[d^2(k)]}{|G^*|^2} - 1 \right\} \tag{1.113}$$

つまり、1.3.1節で提案した周波数分析法の分析精度はステップサイズパラメータに比例し、真のスペクトルパワーの2乗に反比例する。このことから、真のスペクトルパワーが小さい場合でもステップサイズパラメータを小さな値にすれば、精度良い分析が可能であると考えられる。

1.4 二つのアルゴリズムの比較

1.4.1 分析精度

すべての周波数成分に一つの適応アルゴリズムを適用した方法では信号が帯域制限周期波形であれば、パーセバルの定理が成り立ち差の2乗平均の最小値 $\xi_{i,min}$ は0になるが、各周波数成分ごとに適応アルゴリズムを適用した方法ではそのようになるとは限らない。また、LMSアルゴリズムを用いているため、勾配ベクトルの推定値には雑音が含まれ、多数回の適応を繰り返しても集合平均的に収束するだけで係数は最適解の近傍でゆらぎを生じ、差の2乗平均 $\xi_{i,k}$ は差の2乗平均の最小値 $\xi_{i,min}$ より大きくなってしまふ。この誤差量は前節からわかるように差の2乗平均の最小値 $\xi_{i,min}$ とステップサイズパラメータ μ_i の積に比例する。これらのことから、信号に対する追従性を考慮しなければ、すべての周波数成分に一つの適応アルゴリズムを適用した方法の方が精度良く周波数分析が行なえると考えられる。このことを示すために、定常的な2つの信号について信号のパワー比とステップサイズパラメータ μ の値を変え、スペクトル分析(分析間隔 $\Delta f = \frac{f_s}{64}$, f_s : サンプル周波数)を行なった。用いた信号はサンプル周波数の4分の1の正弦波とサンプル周波数の8分の1の正弦波とした。図1.4, 図1.5にサンプル周波数の4分の1の正弦波の分析精度を示す。ここで、 A は入力信号のパワー比、 B は出力の分析精度を表し、つぎに示すような式を用いて計算した。

$$A = \sqrt{\frac{E[s_1^2(k)]}{E[s_2^2(k)]}} \quad (1.114)$$

$$B = \frac{|G_o|^2 - |G^*|^2}{|G^*|^2} \quad (1.115)$$

s_2 はサンプル周波数の4分の1の正弦波のパワーを、 s_1 はサンプル周波数の8分の1の正弦波のパワーを、

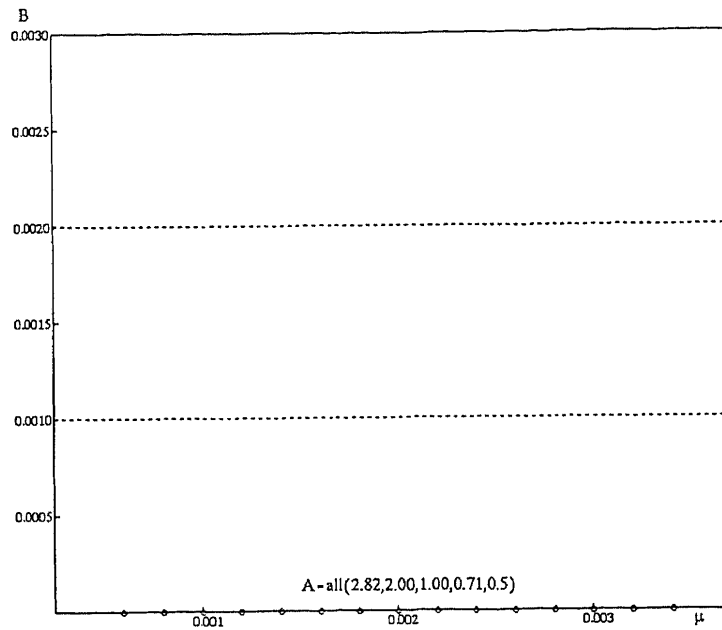


図 1.4: 図 1.1 を用いた場合のステップサイズパラメータと出力の分散との関係

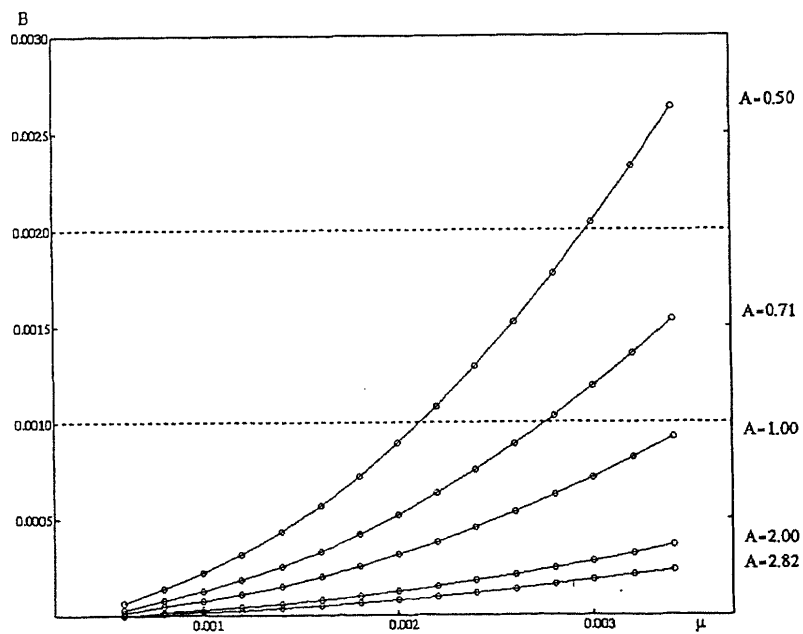


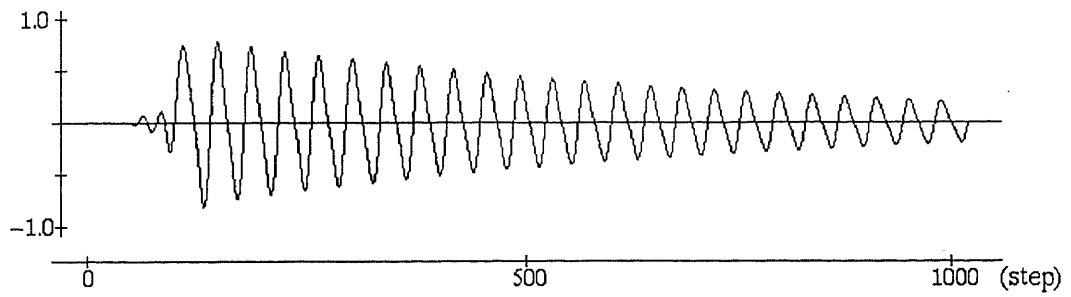
図 1.5: 図 1.2 を用いた場合のステップサイズパラメータと出力の分散との関係

周波数の8分の1の正弦波のパワーを、 G_0 は出力から得られるサンプリング周波数の8分の1の正弦波に対するフーリエ係数を、 G^* はその最適値を表す。これらの結果から時間変化の少ない信号に対してはすべての周波数成分の一つの適応アルゴリズムを適用した方法の方が分析精度が良いことがわかる。

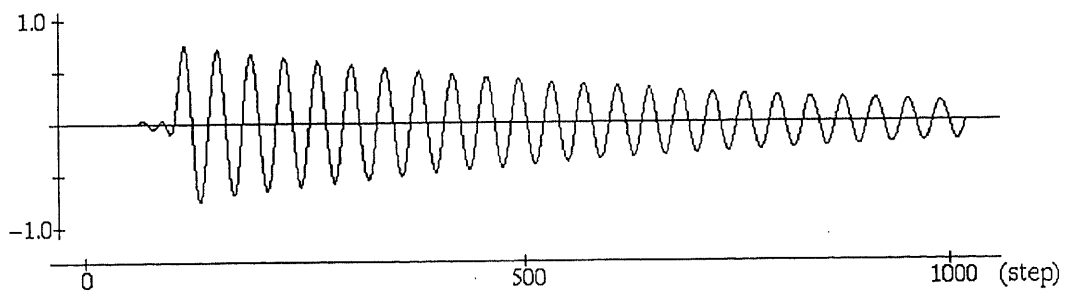
1.4.2 信号に対する追従性

すべての周波数成分の一つの適応アルゴリズムを適用する方法(図1.1)の応用として、採譜のように信号は帯域制限周期波形であるが、どの周波数成分が含まれているかがわからない場合と楽器の構成音の分析の場合のように含まれている周波数成分がわかっている場合の二つが考えられる。そこで、これら二つの場合と、図1.2に示したシステムを用いて各周波数成分ごとに適応アルゴリズムを適用した方法の三つについてのモデルとして、図1.6(b)(c)(d)に示されるような3つの周波数成分からなる合成波形(図1.6(a))のスペクトルの時間的な変化を求め、信号に対する追従性の比較を行なった。図1.7は採譜への応用のようにどの周波数成分が含まれているかがわからないと仮定し、第2オクターブから第5オクターブまでの48点について、図1.1を用いて周波数分析を行なった結果である。図1.8は楽器の構成音の分析の場合のように含まれている周波数成分がわかっていると仮定し、三つの周波数成分について、図1.1を用いて周波数分析を行なった結果である。また、図1.9は図1.2に示したシステムを用いて、48点の周波数成分を別々に周波数分析した場合の結果である。

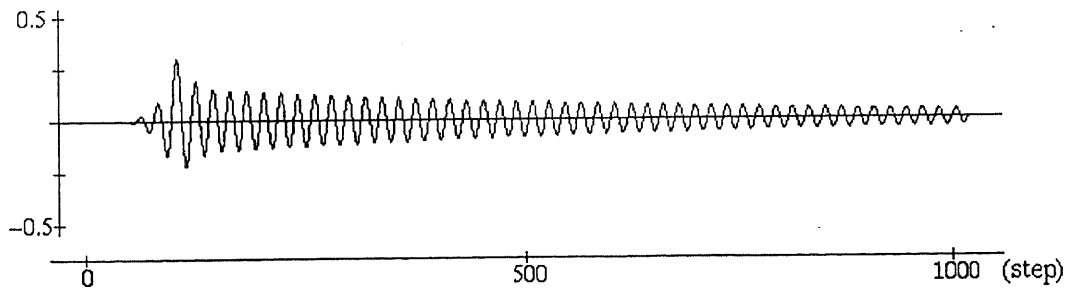
これらの結果からわかるように構成音がわかっている場合には図1.1に示されるようなすべての周波数成分の一つの適応アルゴリズムを適用した場合が良く、採譜への応用などのように信号は帯域制限周期波形であるが、どの周波数成分が含まれているかがわからない場合には、図1.1に示したシステムを用いて音に含まれると考えられるすべての周波数成分の一つの適応アルゴリズムを適用するより、図1.2に示したシステムを用いて、音に含まれると考えられる周波数成分に別々に適応アルゴリズムを適用した方が追従性が良いと考えられる。



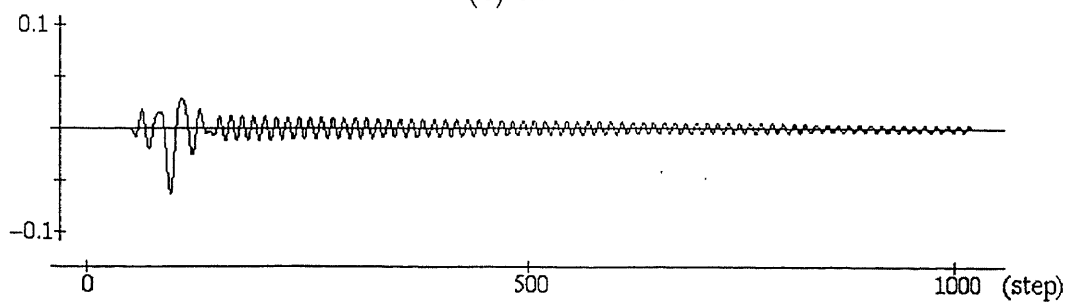
(a) 合成波



(b) 261.63Hz

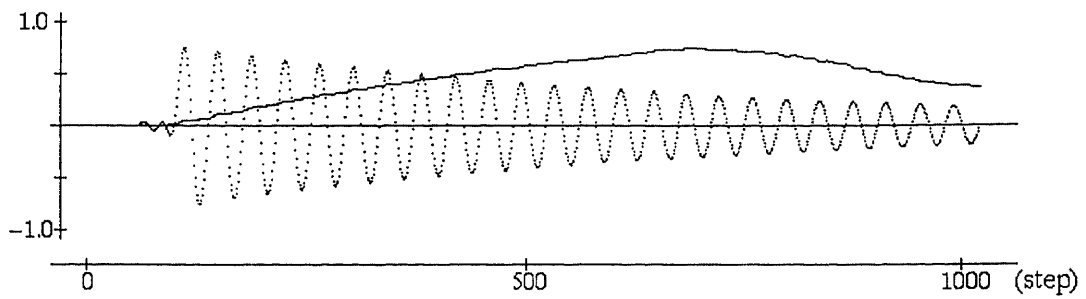


(c) 523.25Hz

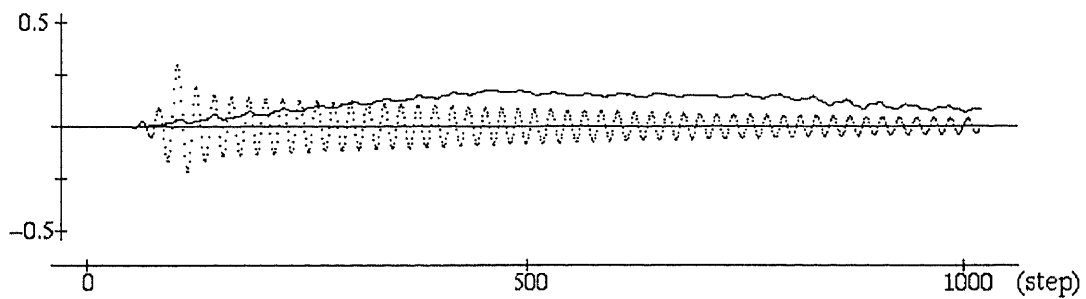


(d) 1046.50Hz

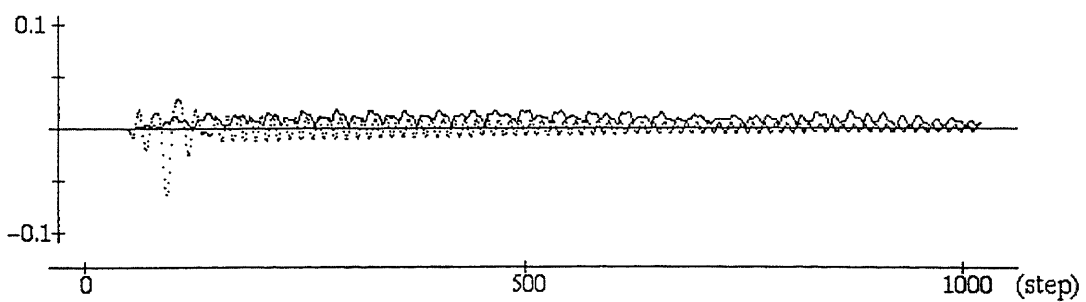
図 1.6: シミュレーションのための合成波形と各周波数成分の波形



(a) 261.63Hz

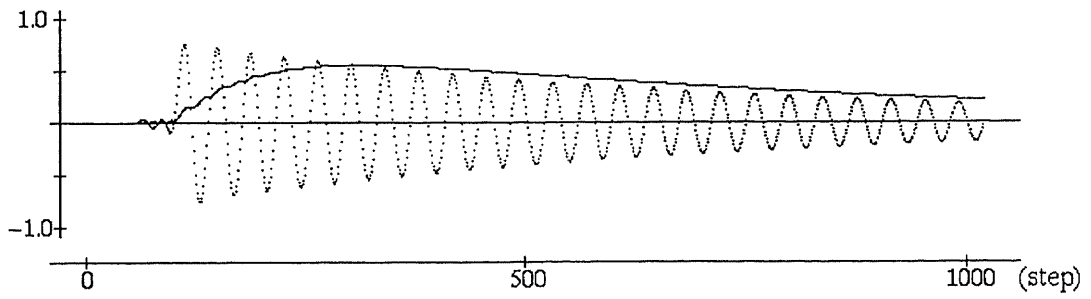


(b) 523.25Hz

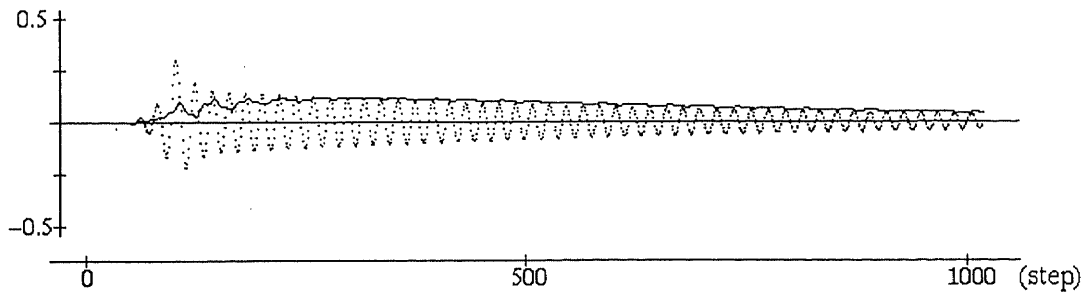


(c) 1046.50Hz

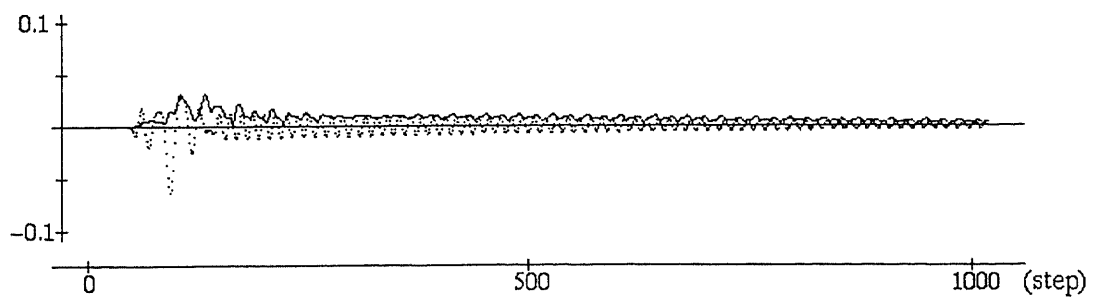
図 1.7: 図 1.1 を用いた場合の結果 (L=48)



(a) 261.63Hz

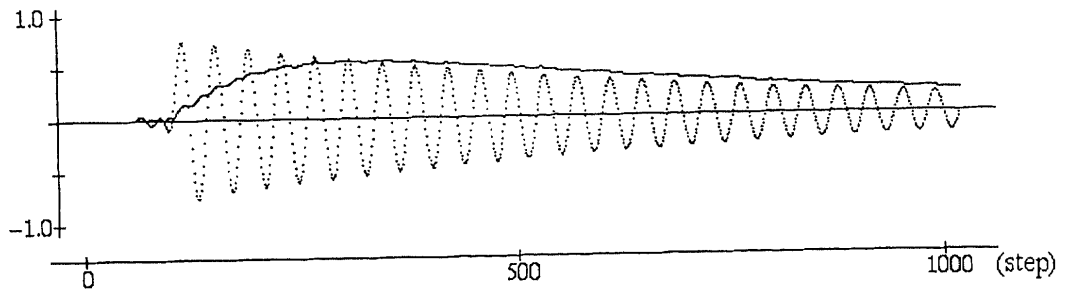


(b) 523.25Hz

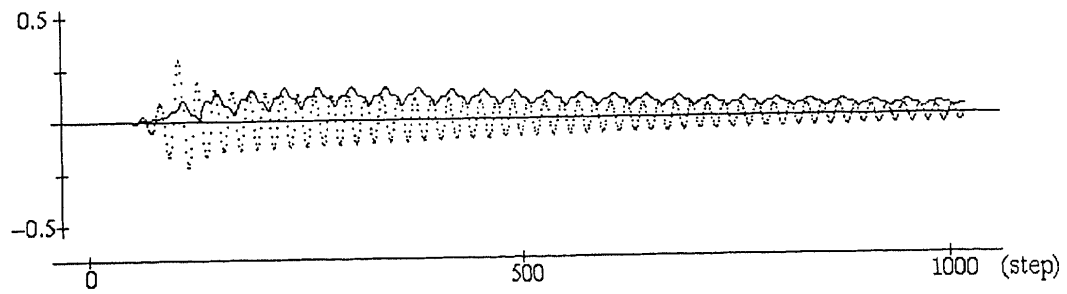


(c) 1046.50Hz

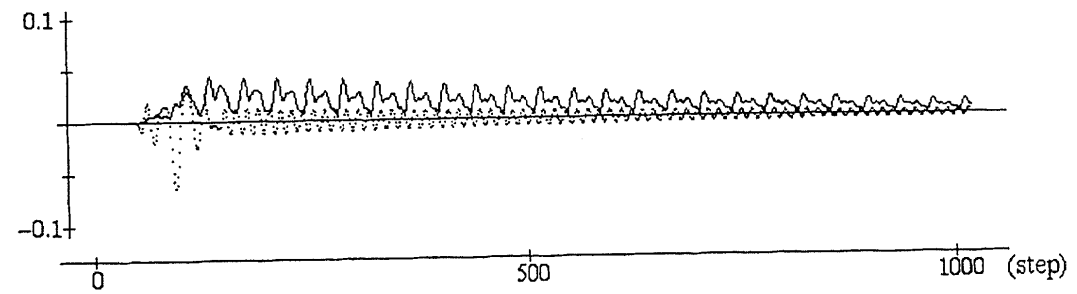
図 1.8: 図 1.1 を用いた場合の結果 ($L=3$)



(a) 261.63Hz



(b) 523.25Hz



(c) 1046.50Hz

図 1.9: 図 1.2 を用いた場合の結果

1.5 結論

本章では、取り扱う信号を帯域制限周期波形に限定し、すべての周波数成分に対して一つの適応アルゴリズムを適用する場合と各周波数成分ごとに適応アルゴリズムを適用する場合の周波数分析法を提案し、その収束条件を求めるとともに、その性能の検討を行ない、二つの周波数分析法を分析精度と信号に対する追従性で比較を行なった。これら二つの周波数分析法は、任意の分析間隔・フレーム長でのスペクトル分析を行なうこともでき、採譜、音声分析や楽器の分析などに適した周波数分析法であると考えられるが分析精度と信号に対する追従性の比較からわかるようにすべての周波数成分に対して一つの適応アルゴリズムを適用する場合は時間変化の少ない信号を精度良く分析する場合や楽器の構成音进行分析する場合のように信号を構成する周波数成分がわかっている場合に有効な周波数分析法である。一方、各周波数成分ごとに適応アルゴリズムを適用する場合は、構成音の周波数成分がわからない場合では信号に対する追従性が、すべての周波数成分に対して一つの適応アルゴリズムを適用する場合に比べ良いので音声分析などに有効であると思われる。

第 2 章

採譜処理への応用

2.1 採譜とは

人間は人間の話す言葉を認識するのと同様に，楽器演奏などの音から楽譜を作ることができる．このことを採譜と言う．採譜処理は大変難しく，長年の経験を踏まえてきた専門家でさえも，日頃，聞き慣れている西洋音楽を正しく楽譜に再現することはなかなかできない．まして民族音楽となると，ほとんど不可能に近いと考えられる．

一方，コンピュータを用いた情報処理の分野においては，人間の話す言葉を認識する技術は，ある程度の進歩をとげ，種々の分野で実用化されている．これと同様に音楽もコンピュータで認識ができれば，その応用分野は色々ある．以下にその例を挙げる．

民族音楽等の記録

この地球上には，多くの民族が住んでおり，固有の音楽を持っている．しかし，これらの音楽情報の多くは，音響データとしてしか残っていない．民族学的観点からこのような音楽を研究するためには，音響データそのものを記録するとともに楽譜を作成する必要がある．このことが採譜の応用として考えられる．

浄写あるいは写符

楽譜をきれいに書くことはなかなか面倒なことであり、多くの時間を必要とする。録音テープから楽譜ができれば、作曲家にとって、浄写の手間が省け、作曲に専念することができる。このことが採譜の応用として考えられる。

演奏技術の解析や名演奏家の再現

最近、パーソナル・コンピュータのなどで簡単に楽器を演奏（正確に言えば、制御）することができるようになってきた。しかし、このように演奏された曲には感情がないと一般に言われる。これは、楽譜には演奏のための情報が100%含まれているわけではないことから生じる。楽譜を絵に例えるとスケッチであり、色が無い。それに色づけするのが演奏家の才能である。そこで、コンピュータで周波数分析などを行なうことによって得られた物理的データを五線譜上に表した奏譜を作成し、それを色々な角度から解析することが必要になってきた。このことが採譜の応用として考えられる。

音楽教育

これまでの演奏技術の習得はすべて耳で行なってきた。つまり、生徒は先生などの演奏を耳で聞いてそれを覚える。また、先生は生徒が間違えたところを再演したり、口で教えてきた。これらのことはともすれば主観的になり易い。しかし、奏譜によって視覚によっても教えることが可能となれば、音楽教育を客観的に行なうことができる。このことが採譜の応用として考えられる。

このように種々の応用が考えられる採譜は本来、音楽知識、経験に基づいた知的作業である。しかし、コンピュータを用いて採譜を行なう場合、その過程は大きく二つに分けて考えることができる。その一つは音響データから奏譜を作成する過程であり、もう一つは音響データから作成された奏譜から楽譜を作成する過程である。上述のように奏譜とは音という物理的データを五線譜上に表したものであり、楽譜とは奏譜を楽譜表記規則に基づいて書き直したものである。つまり、この二つの処理過程は全く異なったものとなる。前者は主に物

理的な情報によって変換されるものであり、音の音階を同定するのに有効な周波数分析法が必要とされる。しかし、後者の処理はそれとは異なり知識工学の分野であり、豊富な音楽知識が必要とされる。

2.2 採譜処理のための周波数分析法

上述のように、コンピュータを用いて採譜処理を行なう過程で最も重要な処理の一つに音の音階を同定することが考えられる。従来から一般的に用いられてきた高速フーリエ変換法 (FFT) では、周波数軸上の分析間隔は一定であり、何らかの方法を用いて周波数を補正することで基本周波数を同定する必要がある。そこで、高速フーリエ変換を採譜に応用する場合によく用いられるのが複素スペクトル面での内挿を用いた周波数補正法である。しかし、この内挿が有効に作用するためには隣合う周波数サンプル点間に存在する周波数成分は一つである必要性があり、求める音の周波数成分が低い場合にはそのような状態になるように、高速フーリエ変換の点数を長くしなくてはならず、リズムの速い曲の場合、音の抽出には不向きである。一方、本論文で提案した適応アルゴリズムによる周波数分析法は分析する周波数成分を自由に選ぶことができ、分析に必要なデータの数もステップサイズパラメータで決めることができるので、このような採譜処理に有効であると考えられる。このことから、この周波数分析法の採譜処理への応用を試みる。

2.3 採譜処理の過程

本節では採譜処理の過程について述べる。採譜処理における周波数分析法として前章で述べた適応アルゴリズムによる周波数分析法を用いることにする。採譜の目的は種々の音響データを楽譜の形で保存することである。この採譜処理全体の流れを図 2.1 に示す。上述のように音響データを楽譜に変換する過程は大きく分けて二つからなる。第一の過程は、音響データを奏譜に変換する過程である。この過程はさらに二つの過程に細分化することができる。その一つの過程は音が持続する時間、つまり音価を抽出する過程であり、もう一つの過程は音の音階を同定する過程である。そして、第二の過程は前述の過程から求めた奏譜から楽譜を作成する過程であり、豊富な音楽知識を必要とする知識

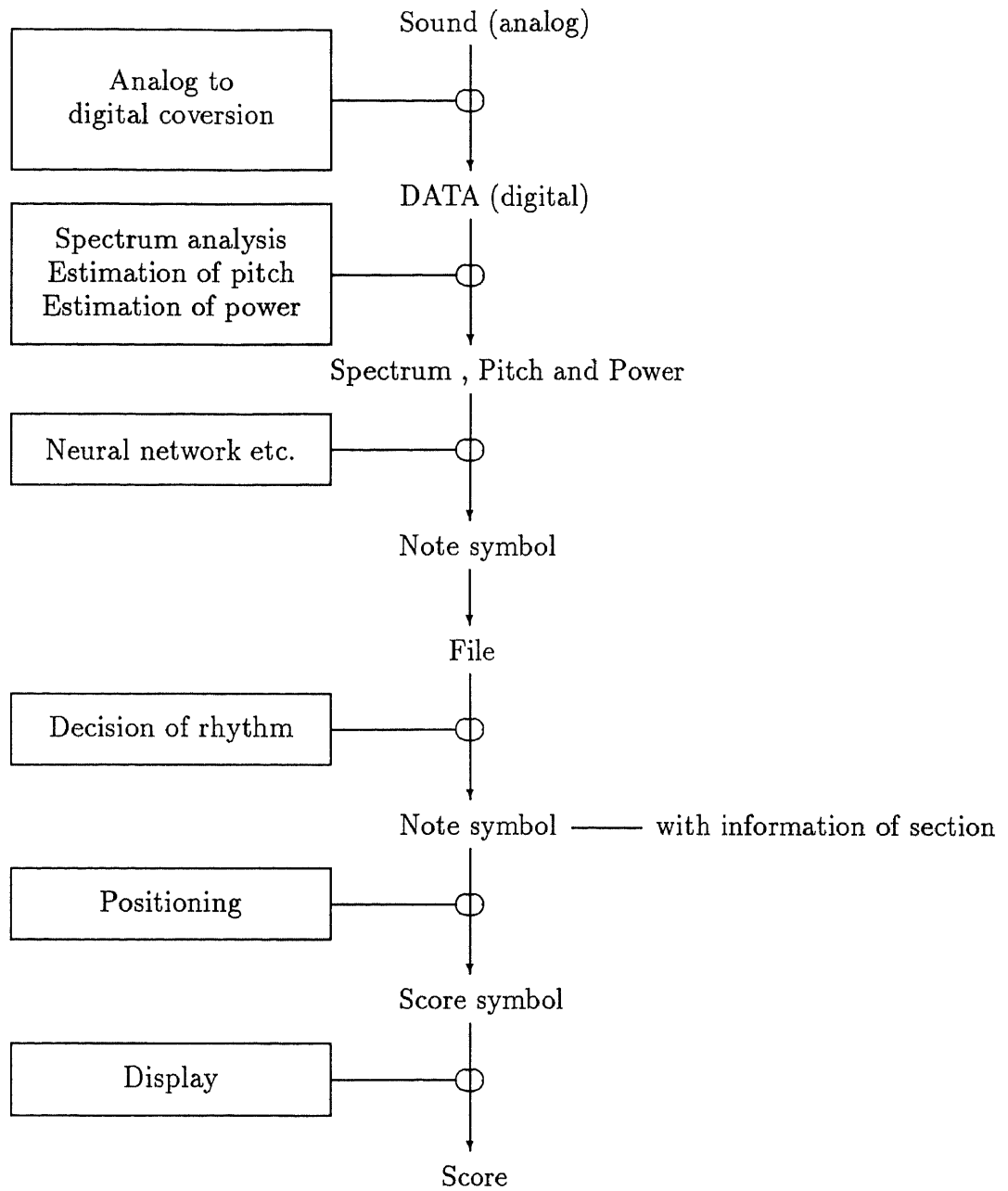


図 2.1: 採譜システム

工学の分野である。

本論文での主目的は、本論文で提案した周波数分析法が音階の同定に有利な方法であることを示すことである。このことから、特に音階の同定法について詳しく説明し、合わせて他の採譜処理の過程についても説明を行なう。

2.3.1 音価の決定

音価とは、音の持続する時間であり、音の立ち上がりを調べることによって決定することができる。本論文においては振幅の変化を調べるのに有利な表現である短時間エネルギーを用いて音価の決定を行なう。短時間エネルギーはつぎの式で定義される。

$$E_n = \sum_{m=-\infty}^{m=\infty} [x(m)w(n-m)]^2 \quad (2.1)$$

これはつぎのように書き直すことができる。

$$E_n = \sum_{m=-\infty}^{m=\infty} x^2(m)h(n-m) \quad (2.2)$$

ここで、 $h(n)$ はつぎの式で与えられる。

$$h(n) = w^2(n) \quad (2.3)$$

このように考えると、式(2.1)は図2.2のように考えることができる。すなわち、信号 $x^2(n)$ は式(2.3)で与えられるインパルス応答を持つ線形フィルタでフィルタリングされる。ここで、 $h(n)$ は一般的に窓関数と呼ばれる。本章ではハミング窓を用いる。ハミング窓はつぎのように定義される。

$$h(n) = \begin{cases} 0.54 - 0.46 \cos(2\pi \frac{n}{N-1}) & 0 < n < N-1 \\ 0 & n < 0 \quad n > N-1 \end{cases} \quad (2.4)$$

図2.3に $N = 101$ として求めた結果を示す。このようにして求められた結果を微分することによって音の立ち上がりを求め、音価を決定し、音価のデータ数とその音価内の最大パワーをファイルに格納する。

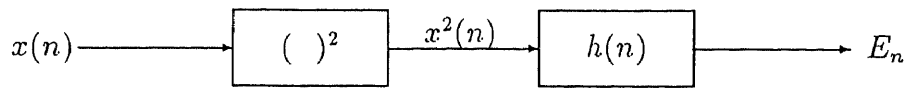


図 2.2: 短時間エネルギーの計算

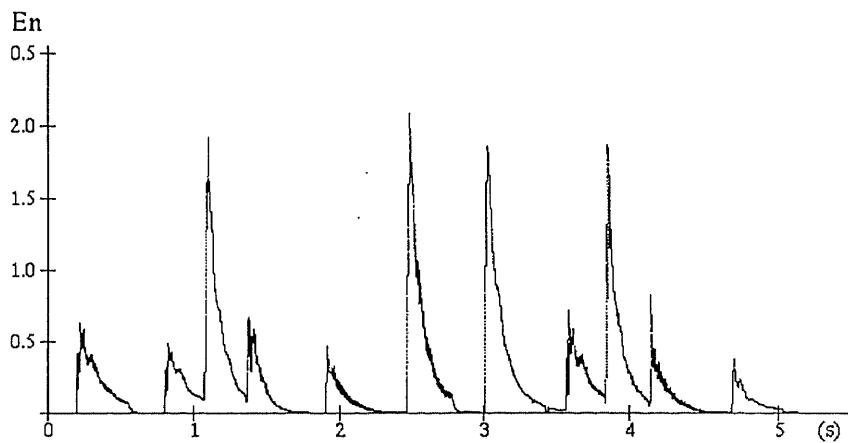


図 2.3: 音価を決めるための短時間エネルギー

2.3.2 音階の同定法

音階とは

採譜における周波数分析は、各音素の基本周波数がどの音階に属するかを同定することである。音楽で用いられる音階には、オクターブをどのように分割するかによって純正音階、ピタゴラス音階、平均律音階などがあるが、工学の分野では最も数学的に取り扱いの容易な1オクターブを対数的に12等分した等分平均律音階を用いることが多い。

本論文では、国際標準周波数(A3音を440Hzとする)による平均律音階を用いて音高の抽出を行なう。

表 2.1: 国際標準周波数 (A3 音を 440Hz とする) に基づいた平均律音階

Note	Number of octave							
	0	1	2	3	4	5	6	7
C	16.35	32.70	65.41	130.81	261.62	523.25	1046.50	2092.99
C \sharp	17.32	34.65	69.30	138.59	277.18	554.36	1108.72	2217.45
D	18.35	36.71	73.42	146.83	293.66	587.33	1174.65	2349.30
D \sharp	19.45	38.89	77.78	155.56	311.13	622.25	1244.50	2489.00
E	20.60	41.20	82.41	164.81	329.63	659.25	1318.50	2637.00
F	21.83	43.65	87.31	174.61	349.23	689.45	1396.90	2793.81
F \sharp	23.12	46.25	92.50	185.00	369.99	739.98	1479.97	2959.93
G	24.50	49.00	98.00	196.00	391.99	783.99	1567.97	3135.94
G \sharp	25.96	51.91	103.83	207.65	414.30	830.60	1661.21	3322.41
A	27.50	55.00	110.00	220.00	440.00	880.00	1759.99	3519.97
A \sharp	29.14	58.27	116.54	233.08	466.16	932.32	1864.64	3729.28
B	30.87	61.74	123.47	246.94	493.88	987.76	1975.52	3951.04

周波数分析

音素の基本周波数を求めることが目的であるので、1kHz 以下に帯域制限されたデジタルデータを以下で示すような周波数分析法を用いて音階の同定を行なう。高速フーリエ変換を用いる場合(付録 A 参照)は観測窓を 50 点ずつシフトさせ、1kHz までの周波数帯について周波数補正し、ピーク周波数を求め、各音階に当てはめる。適応アルゴリズムによる短時間スペクトル分析法を用いて行なう場合は、第 2 オクターブから第 4 オクターブまでの 4 オクターブ、48 点の周波数成分について周波数分析を行なう。これらの周波数分析法から得られたパワースペクトルマップを図 2.4 に示す。

ノイズの除去とビブラートの吸収

上述の周波数分析で求めた各音階の周波数成分は図 2.4 で示されるようにインパルス状のノイズやビブラートを含んでいる。音高を精度良く抽出するためにはこのようなインパルス状のノイズやビブラートを抑制する必要性が

ある。前節の方法で求めた音価単位ごとに求めた周波数成分を積分し、隣合う音階を比較し、また、ピークとして出現した回数を比較することによってインパルス状のノイズの除去とビブラート吸収を行ない、ピークを判断する。図2.4のデータから得られた結果を図2.5に示す。

高調波成分の吸収

楽器が奏でる音は基本波成分と高調波成分からなる。また、その基本波成分と高調波成分のパワーの比は楽器の種類や周波数によって異なる。このことから、あらかじめすべての周波数成分に対して基本波成分と高調波成分のパワーの比のテーブルを用意しておき、上述の処理によって得られたピーク周波数成分に対して、図2.6に示すようなアルゴリズムを繰り返すことによって、各音階のパワーを求める。図2.5のデータから得られた結果を図2.7に示す。

音階の同定

音階の同定は上述までの処理によって得られたピーク周波数成分に対して楽器の種類に応じたノイズレベルまで行なう。図2.7のデータから得られた結果を図2.8に示す。図2.8に示すように抽出されたデータはつぎのような2つの種類に分けて考える。一つ目はその音価内にその周波数の音の立ち上がりがある場合であり、2という数字で表されている。二つ目は前音価からの継続する音である場合であり、1という数字で表されている。また、0はその音階の音が存在しないことを示す。

2.3.3 拍子の決定

拍子

音楽は一定の時間内に刻まれる拍に対して、周期的に強拍部をつくる。このことを拍子という。例えば、2拍ごとに強拍部をつくるならば2拍子であり、3拍目ごとに強拍部をつくるならば3拍子である。つまり音楽を聞けば、周期的に強拍部の位置を感じることができ、そのことから2拍子であるとか3拍子であるとかを判断することができる。この周期的な強拍部を利用して拍子の

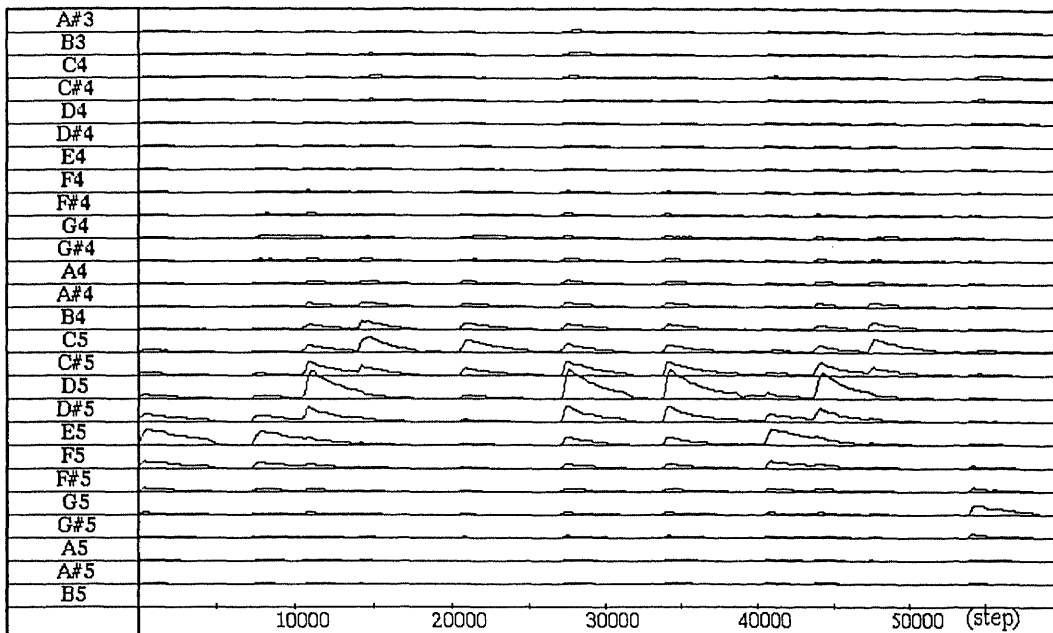


図 2.4: パワースペクトルマップの例

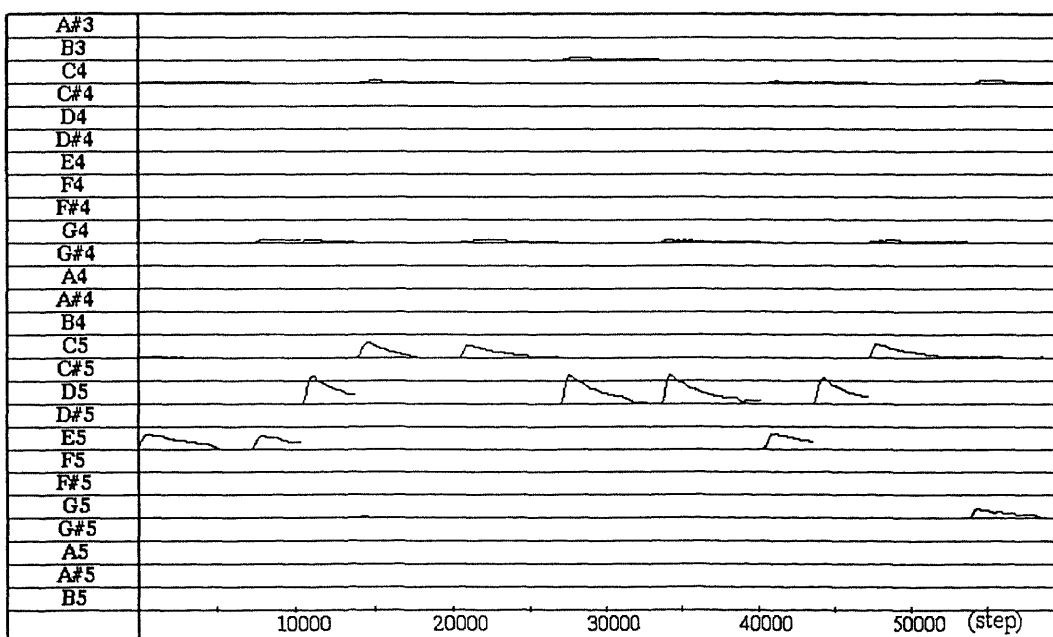


図 2.5: 図 2.4 のパワースペクトルに対してインパルス状の雑音を除去し、ビブラートを抑制した結果の例

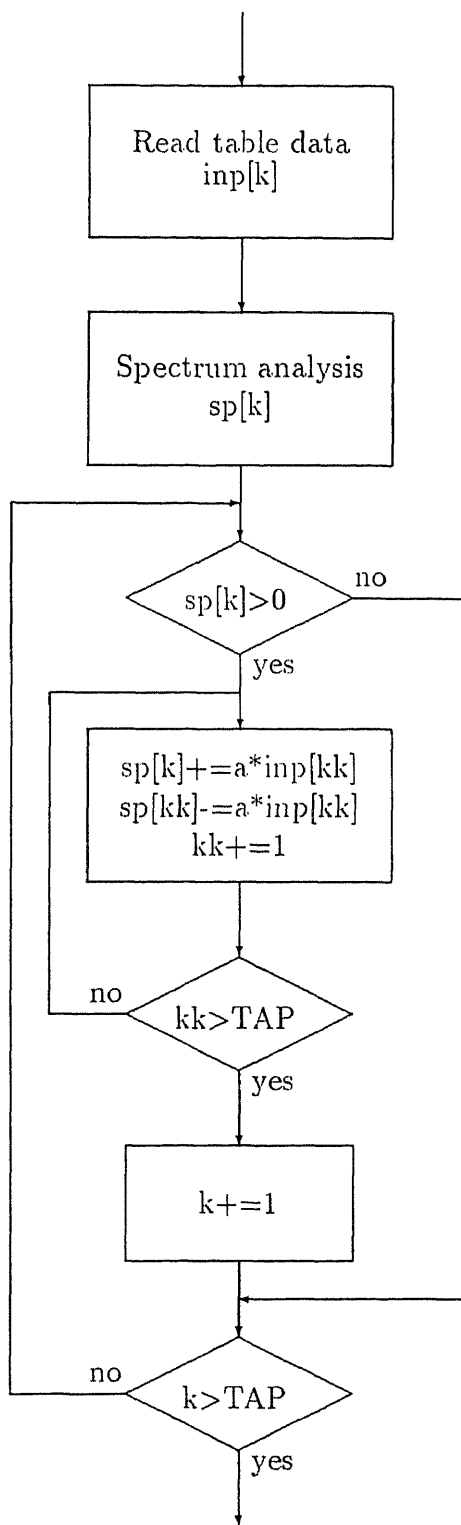


図 2.6: 高調波成分を除去するためのアルゴリズム

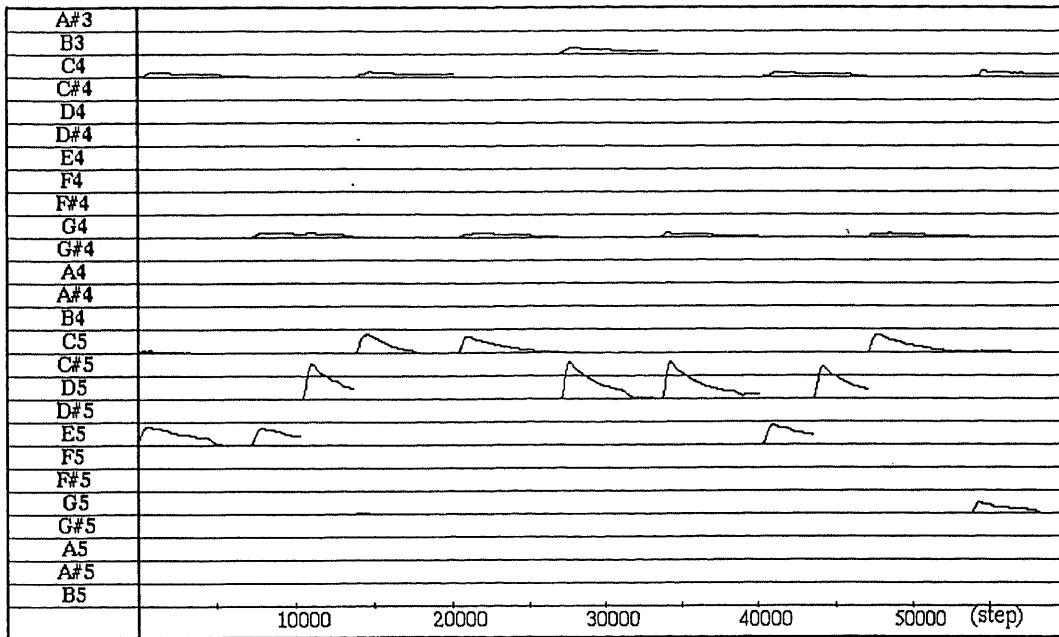


図 2.7: 図 2.5 のパワースペクトルに対して高調波成分を除去した結果の例

A3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A#3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B3	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0	0	0	0
C4	2	0	0	2	0	0	0	2	0	0	2	0	0	2	0	0	0	0	2
C#4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D#4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F#4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G4	0	2	0	0	2	0	2	0	2	0	2	0	0	2	0	0	2	0	2
G#4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A#4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	2	2	0	0	0	2	0	0	0	0	0	0	2	0	0	2
C#5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D5	0	0	2	0	0	2	2	0	2	0	0	0	0	0	2	0	2	0	0
D#5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E5	2	2	0	0	0	0	0	2	0	0	0	0	0	2	2	0	0	2	0
F5	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
F#5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G5	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0
G#5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A#5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

図 2.8: 図 2.7 から求めた音階の同定結果

決定を行なうことにする。

強起と弱起

拍子を決定するのに問題になるのが曲の第 1 拍が強拍部なのか弱拍部なのかである。つまり、音楽には第 1 拍の強拍部から始まるものと弱拍部から始まるものがあり、前者を強起といい、後者を弱起という。弱起の曲は最初と最後の拍数を合わせて 1 小節分になるので、その小節は不完全小節となる。

拍子の決定

この周期的な強拍部を利用し、ニューラルネットワークを用いて拍子の決定を行なうことにした。ニューラルネットワークを用いる場合、学習が必要である。本論文では 8 分音符を用いた標準的なリズムパターンを作成し学習を行なった。このため、拍子を決定するには基準拍を決定し、決定された基準拍から 8 分音符の長さ (データ数) を決めなくてはならない。しかし、人が演奏した曲の場合、同じ音符の音でも音価のデータ数が微妙に異なるし、上述のように曲には、強起の曲と弱起の曲がある。このことから、各音価のデータ数を 100 の位で四捨五入し、1000 ごとに出現頻度を調べ、最も多く出現したものを基準拍の音価のデータ数とすることにした。このようにして得られた基準拍から 8 分音符のデータ数を推測し、各音価を種々の音符に当てはめ、その音価の強さをニューラルネットワークの入力として可能性を確かめる。この時の入力信号の先頭は、始めの 10 個の音価のうちパワーが最も大きいものとする。拍子はこの過程を繰り返し、最も可能性の高いものとする。

2.3.4 調子の決定

音階と調子関係

音楽の旋律は色々な音で構成されているが、それらの音は低い音から高い音への音列に配置することができる。その音の配列を音階という。音階はオクターブを構成する音の数で 5 音音階、7 音音階、12 音音階などに分類することができる。音階の同定で用いた音階は、12 音音階の一つである平均律音階である。一方、一般に譜表表示では 7 音音階が用いられる。12 音音階を

7音音階で表すためには長音階短音階合わせて24種類の調が必要となる。つまり、楽譜を楽譜に変換するには周波数分析で求めた音から調子を決め7音音階で表さなくてはならない。この節では調子の決定のために重要な主要音と転調について説明を行ない、つぎに調子を決定する方法についての説明を行なう。

主要音

上述のように譜表表示には7音音階が用いられる。第1音から順に主音、上主音、上中音、下屬音、屬音、下中音、導音と呼ばれる。主音は音階の基本音で調の出発音であり、終止音でもある。調名はこの音名をもって表す。屬音は音階において主音について重要な音で、調性を安定させる機能を有する音である。下屬音は屬音について重要な音である。導音は主音を導くための音である。これら四つの音は重要な音であり、調子の決定に重要な役割を果たすものである。

転調

音楽においては最初の調のまま曲が終ることも多いが、途中で他の調へ移ることもしばしばある。このように途中で調が移ることを転調という。前の調を原調というのに対して後の調のことを新調という。

転調には曲の一部が一時的に転調するものと、かなり長い部分が転調するものとの2種類がある。前者は、一般的に臨時記号によって転調を表し、後者は原調と新調のさかいに複縦線をつけ、新調の調号を記し転調を明確に表す。

調子の決定

一般に曲の調子はメロディを構成する音の遷移、伴奏和音のコードや終止音の位置などによって決定することができるがメロディと伴奏の分離や最後の音が終止音であるかの判断は難しくこれらによって調子を決定することは困難である。また、主音、下屬音、屬音および導音などの主要音の出現頻度や強拍の音高ヒストグラムから調子を決定することもできるが、曲全体でこのような方法を用いて調子を決定すれば転調に対して対処ができない。しかし、調子決定の目的は譜表表示に用いられる臨時記号を少なくすることであると考えること

ができる。このことから、ある短い小節単位ごとに臨時記号を最も少なくできる調子をその小節単位の調子として決定することにする。

2.3.5 楽譜シンボルへの変換

音素シンボルを楽譜シンボルに変換する流れを図2.9に示す。以下にその説明をする。

音の各譜表への割付け

楽譜は大譜表と呼ばれるト音譜表とヘ音譜表とで表されることが多い。ト音譜表はト音記号で表され、高音部を記すものである。ヘ音譜表はヘ音記号で表され、低音部を記すものである。つまり和音、連符などを表す場合、ト音譜表、ヘ音譜表ごとに認識した方が有利である。このことから本研究ではB3以下を低音部と考え、ヘ音譜表で表し、C4以上を高音部と考え、ト音譜表で表すこととする。

小節の認識

臨時記号の機能が働くのは1小節単位であり、また連符の認識も1小節単位で考えなくてはならない。このことから本論文では得られた音をあらかじめ決められた拍子に従って、小節ごとに音を分けてから種々の処理を行なうことにする。

音の7音音階への割付けと臨時記号の認識

周波数分析で求めた音素シンボルは平均律音階と呼ばれる12音音階である。しかし、譜表は7音音階で表されるのが一般的である。このため前節の方法で調子を決め、12音音階の音素シンボルを7音音階の楽譜シンボルに変換する。しかし、前節で述べたように曲には一時的に転調する場合もあるし、決められた調子では表すことができない音階の音が存在する場合もある。この場合、一般的に臨時記号を用いることになっている。このことから各小節ごとに決められた調子で表すことのできない音素シンボルを認識した場合臨時記号によって表す。

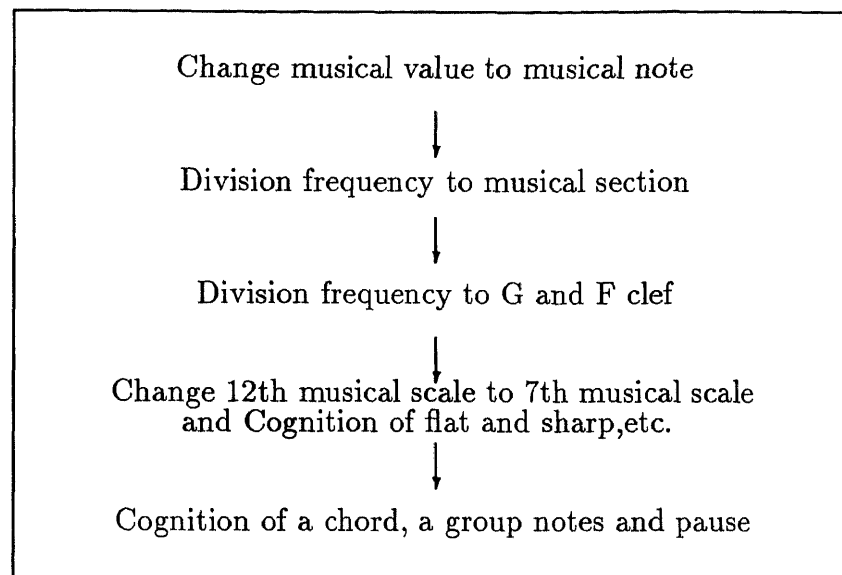


図 2.9: 音素シンボルを楽譜シンボルに変換するためのアルゴリズム

和音，連符の認識

楽譜においては和音，連符などは一つのシンボルとしてとらえた方が自然である。このことから和音に関しては各音部ごとに同じ時刻で鳴り始めて，同じ時刻で鳴り終るものを和音として認識する。連符に関しては各小節において8分音符以下の長さの音価が連続してあれば連符として認識する。

休符の認識

ある音が鳴り終る時間に鳴り始める音が存在しない場合休符を用いて表さなければならない。このことからある音価に存在する音がなければ休符として認識することにする。

2.3.6 楽譜の表示

楽譜シンボルは，楽譜上に表示される図形に対応するものである。楽譜として表示するためには前節で求めた楽譜シンボルに対応する図形を Font として用意しておかなくてはならないし，それを表示する位置も決めなくてはならない。以下にその方法について説明を行なう。本研究で用いているコンピュー



図 2.10: 採譜結果の例

タは SUN4 であり、図形表示は X Window システムを用いている。X Window で表示できる図形はすべて BitmapFont と呼ばれるものである。この BitmapFont は bitmap プログラムで作成することもできるし、線や文字のようにあらかじめ用意されているものもある。しかし、楽譜で用いられる音符などの楽譜シンボルはほとんど用意されていない。このため、あらかじめ使用されると考えられる楽譜シンボルを bitmap プログラムで作成し、準備しておかなくてはならない。このようにして準備され、必要に応じて呼び出された楽譜シンボルは、相対的な位置関係を示すものであるが、これまでの時点で絶対位置の情報を有していない。音符の長さに応じた座標を決定し、楽譜シンボルの属性として絶対位置の情報を与え、楽譜上に表す。このようにして得られた結果を図 2.10 に示す。

2.4 実験

本論文で提案した適応アルゴリズムによる短時間スペクトル分析法は、任意の分析間隔・フレーム長でのスペクトル分析を行なうことができるので、採譜への応用が可能であると考えられる。そこで、従来の方法(付録Aの方法)と音階の同定精度と計算時間の比較を行なうことによってそのことを示す。

2.4.1 実験環境

本研究で用いたシステムを図2.11に示す。メインコンピュータはSUN4/370およびSUN4/60を用いている。SONY PCX-300はDATインターフェイス用のコンピュータとして用いている。また、音響データはSUN4/370上のADA VMEシステムおよびSONY PCX-300上のDATインターフェイスボード(IS-3680)から入力することができる。

2.4.2 実験条件

採譜処理の場合、各音素の基本周波数を求めることが目的であるので、サンプリング周波数を高くする必要性はあまりない。そこでDATを用いた場合は録音したデジタルデータを12kHzにダウンサンプリングし、コンピュータに取り込む。ADA VMEシステム場合はサンプリング周波数を10kHzとし、コンピュータに取り込む。このようにして得られたデジタルデータより種々の周波数分析法を用いて音階の同定を行なう。高速フーリエ変換を用いる場合は観測窓を50点ずつシフトさせ、1kHzまでの周波数帯を周波数補正し、ピーク周波数を求め、各音階に当てはめる。計算時間を比較する場合は、観測窓の長さを512点、1024点および2048点とし、音階の同定精度を比較する場合は、観測窓の長さを1024点とした。また、適応アルゴリズムによる短時間スペクトル分析法を用いた周波数分析法を用いて周波数分析を行なう場合には、第2オクターブから第5オクターブまでの4オクターブ、48個の周波数成分で周波数分析を行ない、音階の同定を行なう。

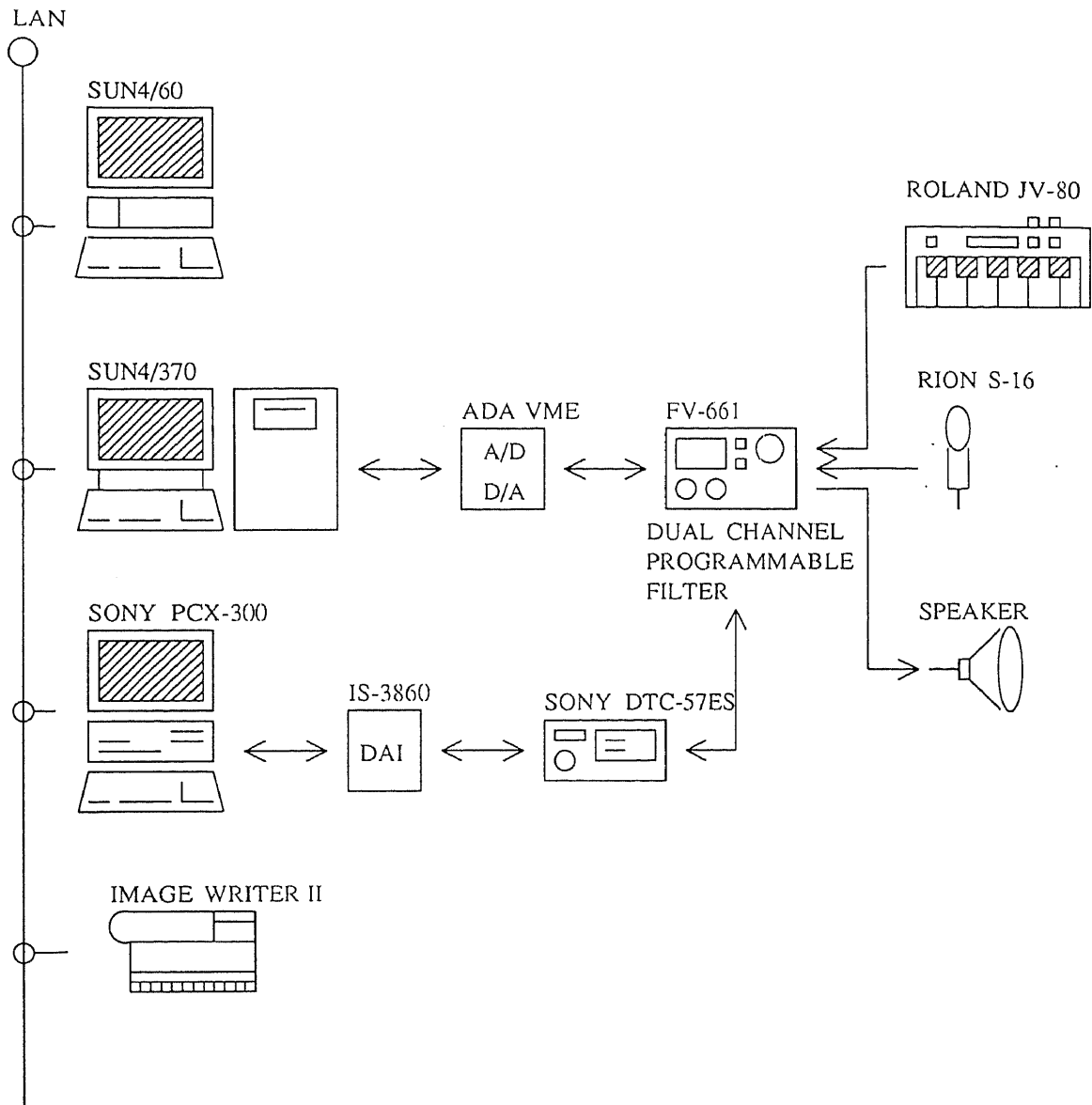


図 2.11: 実験環境

表 2.2: 音階の同定精度の結果

Music	Adaptive method (Fig.1.1)	Adaptive method (Fig.1.2)	1024 points FFT method
Musundehiraite	90.7%	90.7%	90.0%
Izumi no hotori	91.5%	92.0%	79.9%
Tanigawa no nagare	95.1%	97.3%	80.0%
Ogawa no suisya	92.4%	94.5%	87.7%
Kari	90.4%	92.3%	89.7%
Daisukina warutu	90.5%	95.2%	86.8%
Yogisya	91.0%	93.4%	81.0%
Merisan no hituji	92.6%	92.8%	85.3%
Yama no ongakuka	88.3%	92.0%	76.7%
Average	91.3%	93.7%	83.7%

2.4.3 音階の同定精度

上述の条件でピアノ曲について音階の同定を行なった結果を表 2.2 に示す。その結果、各周波数成分ごとに適応アルゴリズムを適用した周波数分析法が最も精度良く音階を同定することができた。つまり、音階を精度良く同定するには周波数分析精度を向上させることよりも信号の変化に対する追従性を向上させなくてはならないことを意味している。高速フーリエ変換では観測窓内の信号の平均的な周波数成分しか求められないので信号に対する追従性に問題が生じるし、すべての周波数成分に対して一つの適応アルゴリズムを適用した場合でも分析しなくてはならない周波数の数 L を小さくできないので信号に対する追従性が劣るからである。

図 1.1の方法でも 94% 程度である理由としてつぎのことが考えられる。和音を扱う場合、同定精度を向上させるためには、ピーク周波数が実際に演奏された音の基本波成分なのか、他の音の高調波成分であるのかを決めなくてはならない。そこで、基本波成分と高調波成分のパワーの比のテーブルを用意しておき、基本波成分のパワーをその基本波成分が持つ高調波成分のパワーを加えたものとして高調波成分の除去を行ない、音階を同定することが有効な方法の一つであると考えられる。しかし、基本波成分と高調波成分のパワーの比は変動し、変動の仕方は楽器の種類や基本周波数によって異なるので、より音階の同定精度を向上させるにはその変動の仕方を記憶しなくてはならない。しかし、実際のシステムを考えれば、メモリの制約から変動する基本波成分と高調波成分とのパワーの比を平均化したテーブルを用いなくてはならないことである。

2.4.4 計算時間の比較

計算時間の短縮化の方法

本論文で提案した適応フィルタを用いた周波数分析法は逐次計算を必要とするので、一般的に高速フーリエ変換に比べ計算時間がかかるという問題がある。これは、係数を更新する式において毎回、入力 $X(k)$ を C 言語による倍精度の算術関数によって求めていたためである。この問題を解決するため本章では入力 $X(k)$ を高速に求める方法としてつぎの二つの方法を考えた。

テーブル読みだし法 この方法は、図 2.12のように 1 周期の正弦波形を $N(=2^n)$ 等分し、テーブルを作成し、その番地を次で示す式によって呼び出すことによって $\sin \omega t$ の値とする方法である。

$$M = \text{int}(N \cdot f / f_s \cdot t) \quad (2.5)$$

$$M = M(\text{AND})N - 1 \quad (2.6)$$

ここで M は呼び出すアドレス、 f は求めたい周波数、 f_s はサンプリング周波数、 T は時刻である。ただし、この方法では図 2.13および図 2.14のような誤差が生じる。その誤差は $N = 512$ では -32dB 、 $N = 1024$ では -35dB 程度となる。

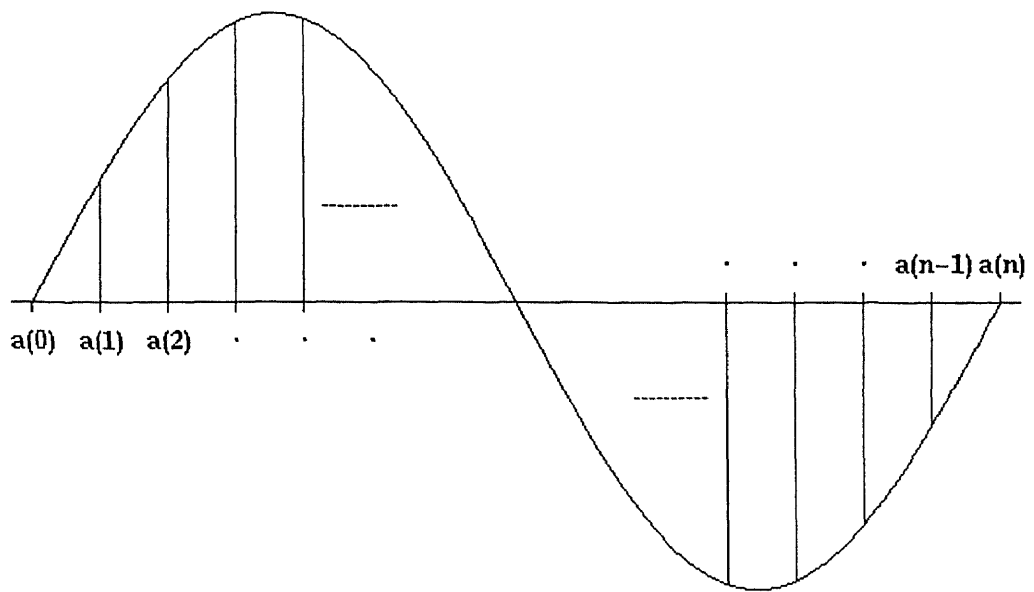
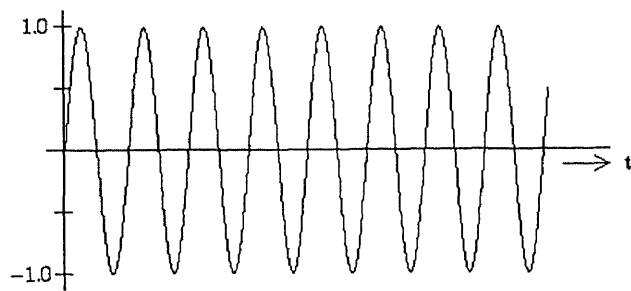
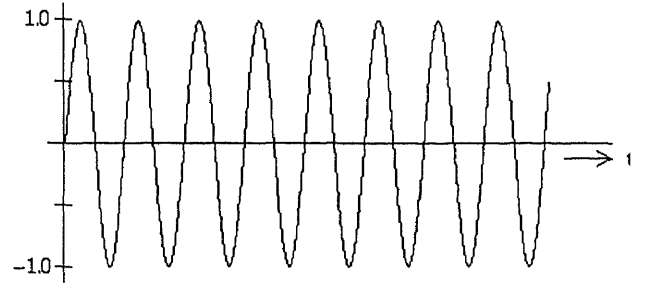


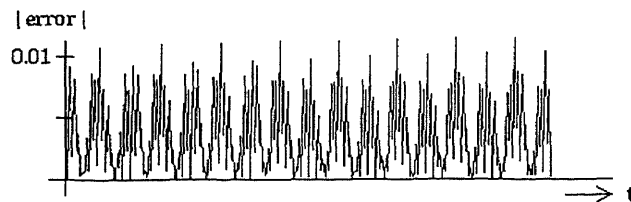
図 2.12: テーブル読みだし法のためのテーブルの作成



(a) C 言語による正弦波波形

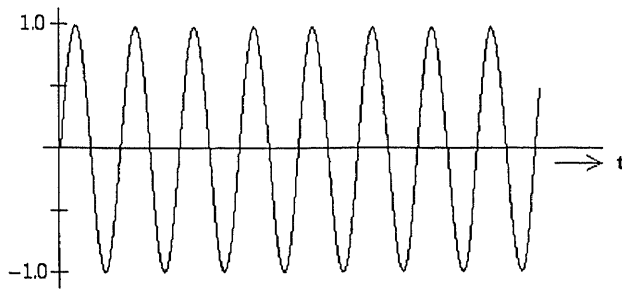


(b) テーブル読みだし法による正弦波波形

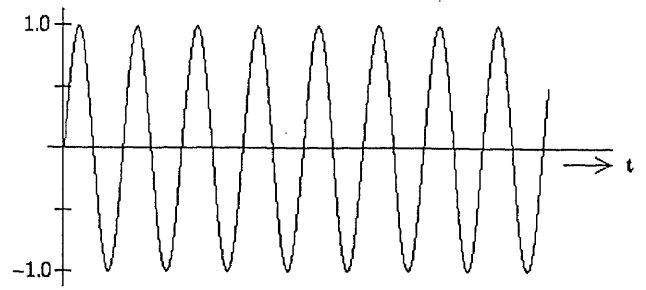


(c) 誤差

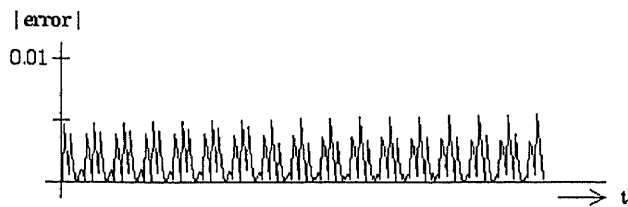
図 2.13: 512 分割した場合のテーブル読みだし法による正弦波波形と誤差



(a) C 言語による正弦波波形



(b) テーブル読みだし法による正弦波波形



(c) 誤差

図 2.14: 1024 分割した場合のテーブル読みだし法による正弦波波形と誤差

三角関数の加法定理を用いる方法 フーリエ変換を行ないたい信号 $d(k)$ はサンプリング周期ごとの離散データであるので入力 $X(k)$ もサンプリング周期ごとの離散データでよい。つまり、 $\exp j\omega t$ の値は、サンプリング周期ごとの値である。このことから次に示す加法定理を用いて $\sin \omega t$ および $\cos \omega t$ を更新し、その和から $\exp j\omega t$ を求めていく方法が考えられる。

$$\begin{aligned}\sin\{\omega(k+1)T\} &= \sin(\omega kT + \omega T) \\ &= \sin(\omega kT) \cdot \cos(\omega T) + \cos(\omega kT) \cdot \sin(\omega T) \quad (2.7)\end{aligned}$$

$$\begin{aligned}\cos\{\omega(k+1)T\} &= \cos(\omega kT + \omega T) \\ &= \cos(\omega kT) \cdot \cos(\omega T) - \sin(\omega kT) \cdot \sin(\omega T) \quad (2.8)\end{aligned}$$

つまり、あらかじめ式 (2.7), (2.8) の $\sin(\omega T)$, $\cos(\omega T)$ (T : サンプリング周期) の値をテーブルにしておく。ここでは、 $\sin(\omega T)$, $\cos(\omega T)$ をそれぞれ sw

表 2.3: 計算時間の比較

		Calculation time
Adaptive method	Direct method	11'20"
	Addition theorem method	4'18"
	Read table method	5'55"
FFT method	512 points	3'39"
	1024 points	6'32"
	2048 points	12'32"

および cw とする。また、 $\sin(0)$ は 0.0 であり、 $\cos(0)$ は 1.0 であるので初期値をつぎのようにおく。

$$s(0) = 0.0 \quad (2.9)$$

$$c(0) = 1.0 \quad (2.10)$$

よって $\sin(\omega kT)$ および $\cos(\omega kT)$ の値はつぎの式から次々に計算でき、その和から $\exp j\omega kT$ の値を求めることができる。

$$s(k+1) = s(k) \cdot cw + c(k) \cdot sw \quad (2.11)$$

$$c(k+1) = c(k) \cdot cw - s(k) \cdot sw \quad (2.12)$$

この方法では丸め誤差は生じるがテーブル読みだし法のような誤差は生じないので $\exp j\omega kT$ の計算精度は良いと考えられる。

実験結果

これら二つの方法を用いて、2.8.2節の条件で130000点のデータに対して行なった計算時間の比較を表2.3に示す。ここで用いたコンピュータはSUN4/60である。音高の抽出を行なう場合、分析すべき周波数帯は限定できるので表2.3からわかるように加法定理を用いて入力 $\exp j\omega t$ を求めれば、本論文で提案した適応アルゴリズムによる短時間スペクトル分析法は高速フーリエ変換による複素スペクトル面での内挿法と同程度の計算速度が得られることがわかった。

2.5 結論

本章では適応アルゴリズムによる短時間スペクトル分析法の採譜処理への応用を試みた。音階を同定するための周波数分析法については、従来高速フーリエ変換による複素スペクトル面での内挿法が最も効率の良く精度の良い方法である言われていた。しかし、この内挿が有効に作用するためには隣合う周波数サンプル点間に存在する周波数成分は一つである必要性があり、このため、音素の基本周波数が低くかつパワーが高調波のパワーに比較して小さい場合にはそのような状態になるように高速フーリエ変換の点数を長くしなくてはならず、観測窓内の信号の平均的な周波数成分しか求められないためどうしても音の立ち上がりや鳴り終りが精度良く求められないという欠点があった。一方、本論文で提案した周波数分析法では音の持続時間とステップサイズパラメータとの関係がはっきりしており、ステップサイズパラメータを適切に選定することによって信号の時間的变化に対して追従性が良く、そのため他の周波数分析法に比べ音の立ち上がりや鳴り終りが精度良く求められると考えられる。また、計算速度の面でも加法定理を用いて $\exp j\omega kT$ を計算すれば、高速フーリエ変換を用いた複素スペクトル面での周波数補正と同等の計算速度が得られる。これらのことから、本論文で提案した適応アルゴリズムによる短時間スペクトル分析法は採譜処理に適していると考えられる。

また、採譜システムの構築法についても2.3節において述べたような音楽的知識を用いて有効な採譜システムが構成できることを示した。

第 3 章

音声の分析への応用

3.1 はじめに

不特定話者，大語彙音声認識システムの実現は重要な課題である．このようなシステムの構成としては，単語単位でパターン・マッチングを行なう方法よりも，音素単位の認識を行ない，単語辞書と音素列のマッチングを行なう方式の方が，単語選択の可能性があり，単語辞書の追加・変更が容易であるなど有利な点が多い．しかし，ここで問題となるのが音韻レベルでの認識率の低さである．この問題を克服するために基本的な音節の特徴抽出の実験が多くなされている．これらの研究の中で最も精度良く特徴抽出を行なえる方法として人間が聴いて明確に識別できる音声の音韻を表す特徴はスペクトル包絡上に存在するとの考え方に基づいた高分解能フィルタ・バンクによるスペクトル包絡線データを観察する方法がある．

一方，本論文で提案した適応アルゴリズムによる短時間スペクトル分析法は周波数軸上の分析間隔を自由に選択できるので，このような音声の特徴抽出に有利であると考えられる．そこで，本論文では，特に特徴が微妙であると言われている /me/, /ne/ について成人男性 9 人が 3 回ずつ発声した 27 単語でその特徴を観察することによってこの二つの音節を区別することを試みる．

3.2 周波数分析

周波数分析システムは，図 3.1 に示すように二つの周波数分析器 (分析器 I・II) とスケーリング，平滑化からなる．

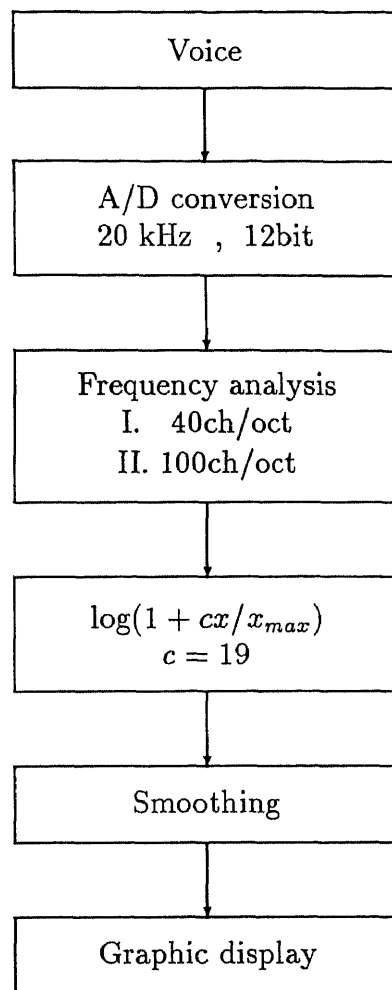


図 3.1: 周波数分析システム

3.2.1 周波数分析器

分析器 I・II は図 1.2 で構成されており、分析器 I はスペクトル包絡線全体の形状を観測するためのものであり、分析器 II はフォルマントなどの移動を詳しく観測するためのものである。分析器 I は 40ch/oct, 156Hz から 5kHz の 5 オクターブ、全 200ch であり、各チャンネルの分析周波数はずぎのようになる。

$$f_i = 0.0065 f_s \cdot 2^{\frac{(i-1)}{40}} \quad (i = 1, 2, \dots, 200) \quad (3.1)$$

また、尖鋭度が高くなりすぎずスペクトル包絡線の観測を容易にするために、ステップサイズパラメータはつぎに示すようなものを用いた。

$$\mu_i = 0.0008 + \frac{i}{38500} \quad (i = 1, 2, \dots, 200) \quad (3.2)$$

このステップサイズパラメータを用いれば、周波数分解能は第1チャンネル(分析周波数156Hz)で7.4, 第100チャンネル(分析周波数852Hz)で9.8, 第200チャンネル(分析周波数4820Hz)で32.8とゆるやかに増加させることができるので、高周波数帯でも尖鋭度はあまり高くない。同様に、分析器IIは100ch/oct, 1.25kHzから5kHzの3オクターブ, 全300chであり、各チャンネルの分析周波数はつぎのようになる。

$$f_i = 0.0065 f_s \cdot 2^{\frac{(i+149)}{100}} \quad (i = 1, 2, \dots, 300) \quad (3.3)$$

また、ステップサイズパラメータはつぎに示すようなものを用いた。

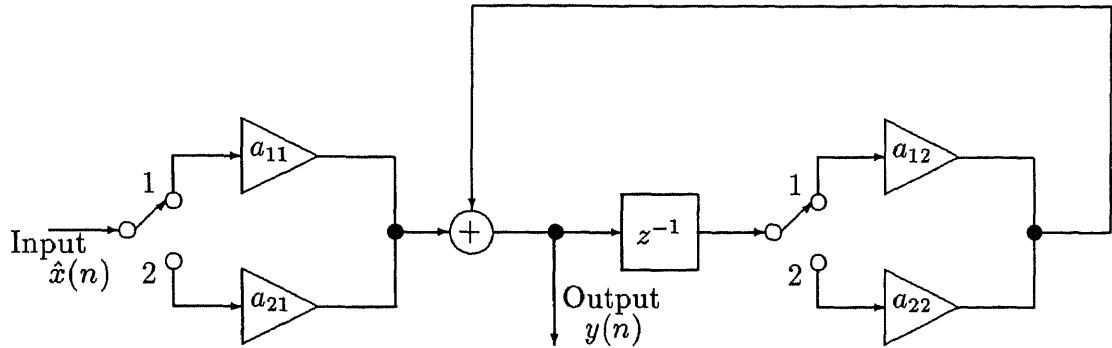
$$\mu_i = 0.0008 + \frac{i + 149}{77000} \quad (i = 1, 2, \dots, 300) \quad (3.4)$$

このことによって分析器Iと同様に周波数分解能は第1チャンネル(分析周波数1.25kHz)で17.9, 第150チャンネル(分析周波数2.5kHz)で20.9, 第300チャンネル(分析周波数4820Hz)で29.6とゆるやかに増加させることができるので、高周波数帯でも尖鋭度はあまり高くなりすぎず、スペクトル包絡線の観測に有利であると考えられる。

3.2.2 スケーリング

スペクトル包絡線の振幅軸は一般的にデシベルが用いられる。このことによって子音部の低レベル部分が強調され、線形軸を用いるよりもスペクトル包絡線の観測が容易になる。しかし、子音部から母音部へのわたり部では第2フォルマント(F_2)から第4フォルマント(F_4)のパワーの立ち上がりが小さくなり、子音部から母音部へのわたりの開始点の抽出が困難になる。そこで、デシベルと線形の間中間的な特性を得るためにつぎのような変換を行なうことにした。

$$\hat{x}(i, k) = \log\left(1 + c \frac{x(i, k)}{x_{max}}\right) \quad (3.5)$$



$$\begin{aligned}
 1. \hat{x}(n) \geq y(n-1) \quad & a_{i1} = \frac{T}{T + \tau_u} \quad a_{i2} = \frac{\tau_u}{T + \tau_u} \\
 2. \hat{x}(n) < y(n-1) \quad & a_{i1} = \frac{T}{T + \tau_d} \quad a_{i2} = \frac{\tau_d}{T + \tau_d} \\
 & y(n) = a_{i1}\hat{x}(n) + a_{i2}y(n-1) \quad (i = 1, 2)
 \end{aligned}$$

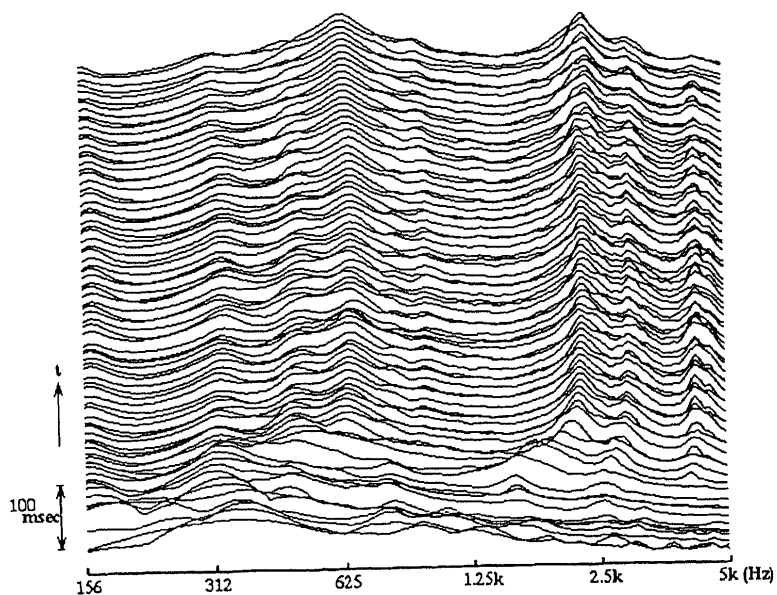
図 3.2: 平滑化のための 1 次遅れ系

ここで、 $x(i, k)$ は分析器の出力、 x_{max} は $x(i, k)$ の最大値、 c は定数 (=19, 視察により決定した) である。

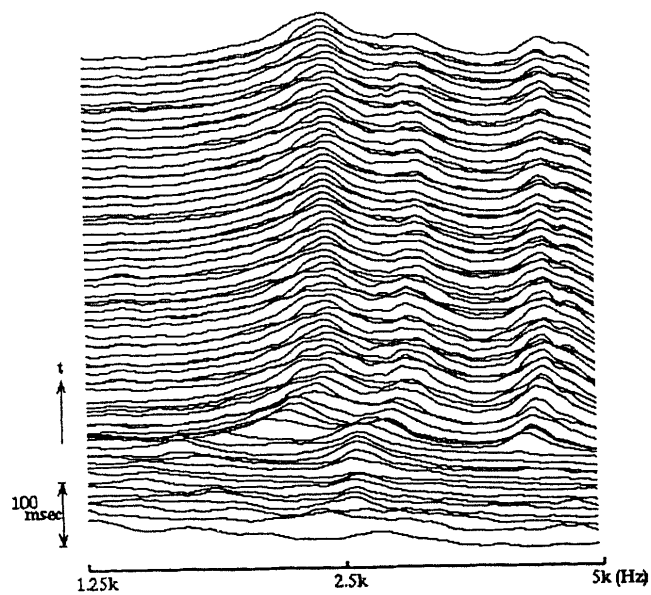
3.2.3 平滑化

対数変換の後、ピッチによるパワーの変動を減少させ、観測を容易にするとともにピークの立ち上がりを重視するという条件を満たすために図 3.2 に示すような立ち上がり時定数 τ_u が 1.6ms と小さく、立ち下がり時定数 τ_d が 6.4ms と大きいような 1 次遅れ系を用いて平滑化を行なった。

このシステムによって作成されたスペクトル包絡線を図 3.3 に示す。



(a) 分析器 I を用いた場合



(b) 分析器 II を用いた場合

図 3.3: スペクトル包絡線の例

3.3 視察による特徴抽出

特徴抽出は、セグメンテーション、フォルマント (F_2 , F_3 , F_4) 抽出、フォルマントの移動量の算出、わたり部でのフォルマントのパワーの立ち上がり方の観測によって行なう。

3.3.1 セグメンテーション

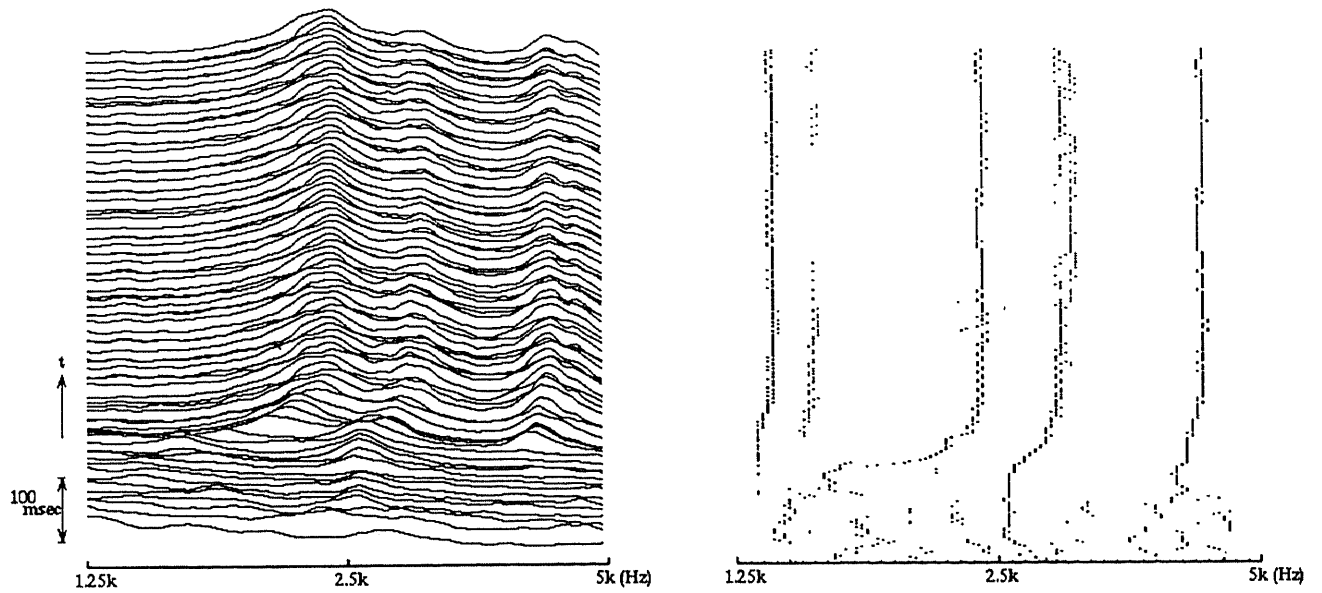
第2フォルマント F_2 から第4フォルマント F_4 のパワーの立ち上がりの観測が容易な分析器Iによるスペクトル包絡線を用いて行なう。

鼻子音のスペクトル包絡線は、子音部 (ハズ)、わたり部、母音部の三つの部分に分割することができ、図3.3(a)からわかるように、わたり部の始点ではフォルマント (F_2 , F_3 , F_4) のパワーが急激に増加したり、フォルマントの急激な移動が生じたりする。このことから、これらの点をわたり部の開始点とし、その手前30ms程度の区間を子音部、わたり部の開始点からフォルマントの位置、形状が比較的定常になるまでの区間をわたり部、それ以後を母音部とする。

3.3.2 フォルマント抽出

前節で述べた三つの部分に対して、分析器IIによるスペクトル包絡線とそのピーク出現マップ (図3.4) を用いて第2フォルマントから第4フォルマントの位置の抽出を行なう。

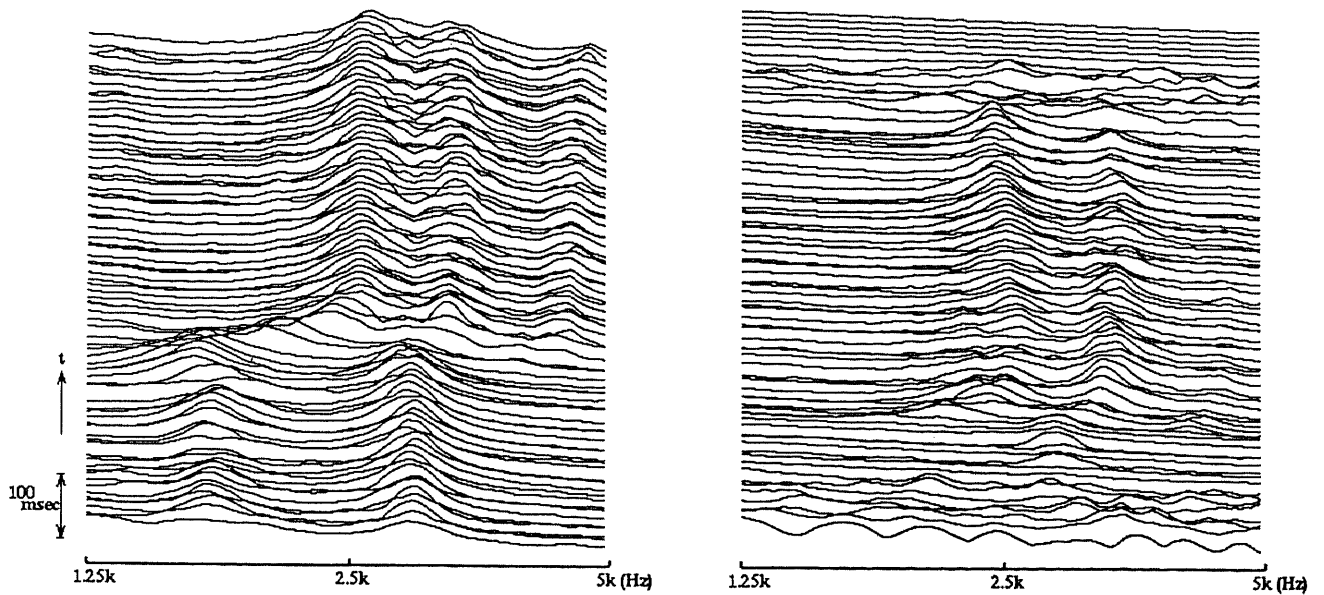
ここで問題となるのが一つのフォルマントが一つのピークにならない場合であり、このような場合にはスペクトル包絡線とそのピーク出現マップから一つのフォルマントであることを視察によって決定し、フォルマントの位置はそれらピークの間の中間の位置として抽出する。さらに、母音部の持続時間が長いので時間軸方向のフォルマントの位置の移動も考慮するために上述のように求められたピーク的位置を平均することによってフォルマントの位置を決定する。また、わたり部ではわたり開始点直後の約1ピッチ分のスペクトル包絡線からフォルマントの位置を決定する。そして、子音部ではわたり開始点直前のピークに対して、第2フォルマント F_2 の抽出を行なうか否かを分析器Iのスペクトル包絡線を用いて判断し、抽出を行なうと判断した場合のみ分析器IIによ



(a) スペクトル包絡線

(b) (a) から求めたピーク出現マップ

図 3.4: スペクトル包絡線とそのピーク出現マップの例



(a) ピークを抽出しない場合

(b) ピークを抽出する場合

図 3.5: スペクトル包絡線の例

るスペクトル包絡線を用いてフォルマントの位置を抽出する。第3フォルマントと第4フォルマントに関しては上述のような判断はせずに無条件でフォルマントの位置を抽出する。第2フォルマントの抽出を行なうか否かは、つぎのように判断する。第2フォルマントの大きさが第3フォルマントに比較して非常に小さい場合(図3.5(a))には抽出しない。ただし、わたり開始点に向かって大きくなるピークや母音部から子音部に向かってピークトラッキングを行なって自然に接続するようなピークの場合(図3.5(b))は抽出する。

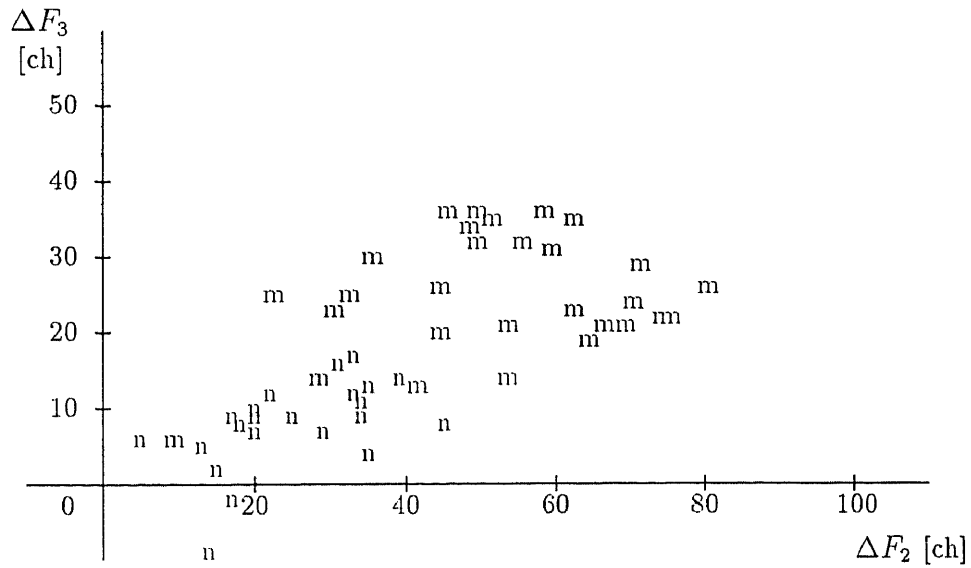
3.3.3 フォルマント移動量の算出

移動量の算出は、わたり部と子音部のフォルマント位置の内、チャンネル数の低い(周波数の低い)方と母音部のフォルマント位置の差とし、子音部またはわたり部から母音部にかけてチャンネルが高くなる(周波数の高くなる)向きを正の値とする。このようにして算出した移動量による $\Delta F_2 - \Delta F_3$ 、 $\Delta F_2 - \Delta F_4$ の分布を図3.6に示す。また、総移動量 $\Delta F (= \Delta F_2 + \Delta F_3 + \Delta F_4)$ のヒストグラムを図3.7に示す。ここで各移動量の重みを1としたのは図3.6からわかるように傾きが-1の直線でほぼ/m/, /n/を分割することができるからである。

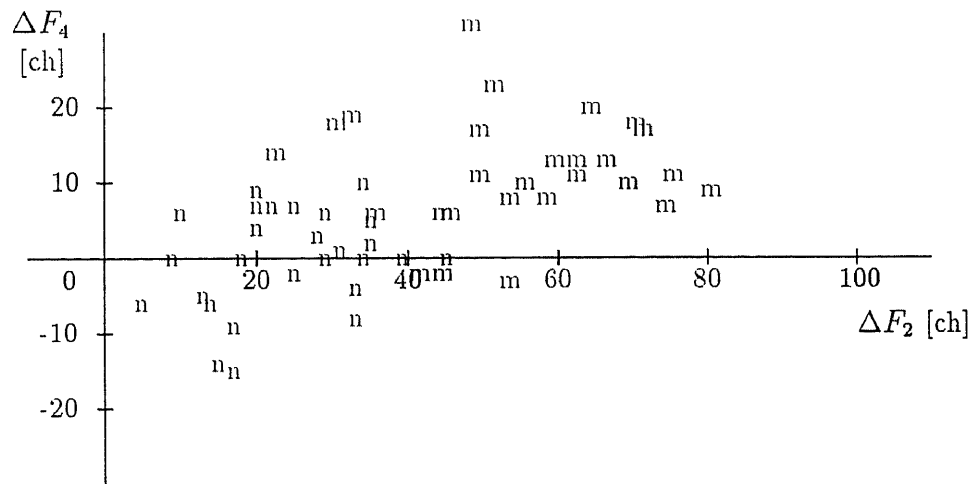
図3.7からわかるように/m/, /n/を区別する境界は55chから60chの間にあると考えられる。この前後10chの間に/m/, /n/が混在しているので $45 \leq \Delta F \leq 65$ chの20ch間の包絡線についてさらに詳しく観察を行なうことにする。

3.3.4 /m/, /n/が混在している範囲の音声の検討

$45 \leq \Delta F \leq 65$ ch間の15個の音声に対して分析器Iによってわたり部でのフォルマントのパワーの移り変わりを観察してみた。その結果、図3.8でわかるように/m/は急激なパワーの立ち上がりがあり、/n/にはそのような特徴が見られなかった。このことによってフォルマント移動量の算出では区別し難い音声も区別できると考えられる。



(a) $\Delta F_2 - \Delta F_3$ 分布



(b) $\Delta F_2 - \Delta F_4$ 分布

図 3.6: フォルマント位置の移動量

3.3.5 特徴のまとめ

特徴 I

子音部から母音部にかけて第2, 3, 4 フォルマントの移動量の総和が /m/ では大きく, /n/ では小さい. その境界は 100ch/oct の分析器を用いた場合で

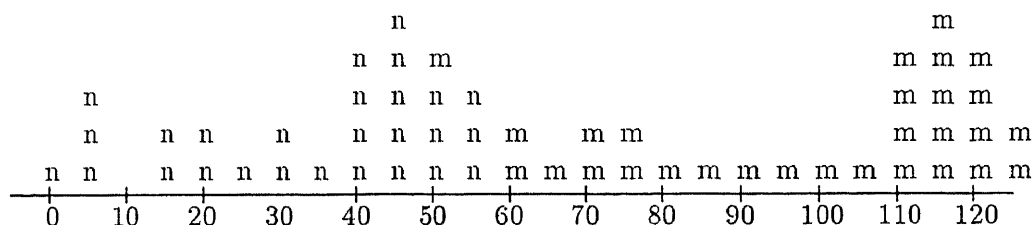


図 3.7: フォルマント位置の総移動量のヒストグラム

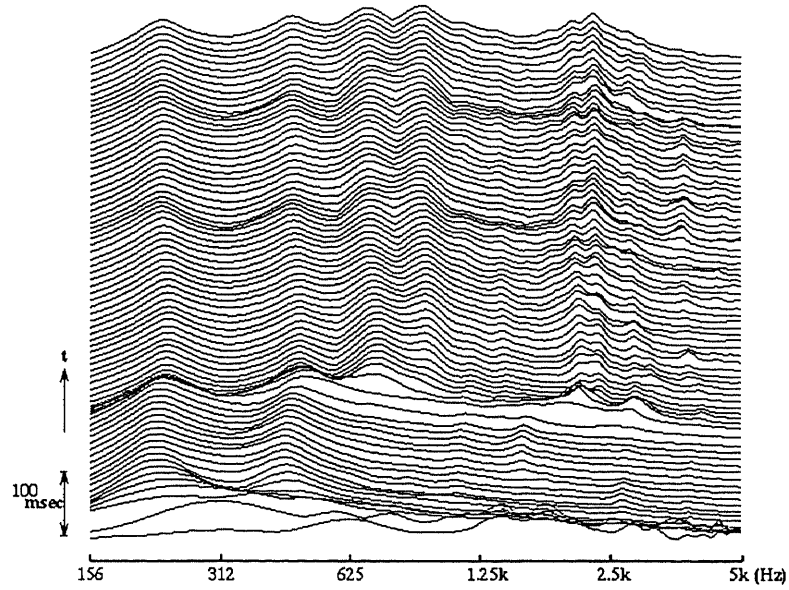
は 55ch から 60ch の間にあると考えられる.

特徴 II

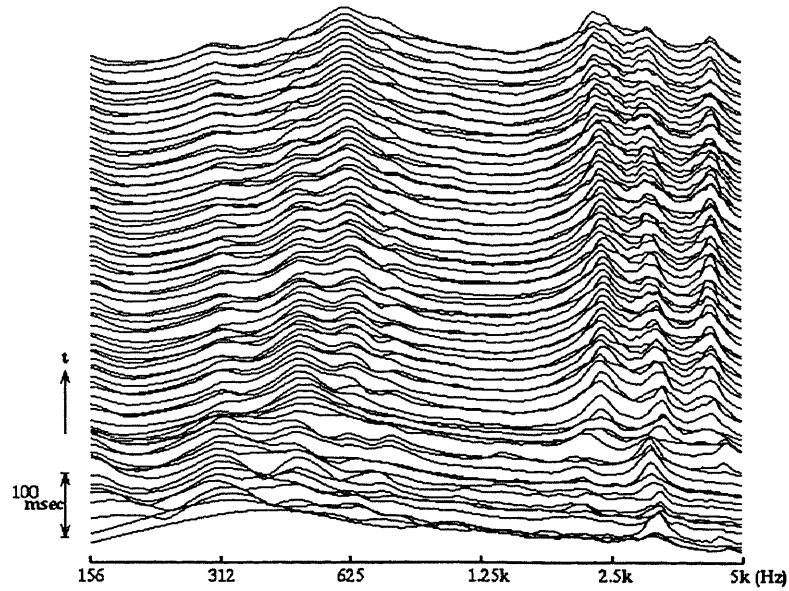
子音部から母音部へのわたりの部分における第2, 3, 4 フォルマントのパワーの立ち上がり方は /m/ では急であり, /n/ では緩やかである. これらの特徴を同時に用いることによって 54 個の音声の内 53 個の音声を正しく識別でき, 98% の識別率が得られた. このことは特徴 I, II が比較的簡単な特徴でありながら非常に有効であることを示している.

3.4 結論

100ch/oct の高分解能周波数分析の後, 対数スケーリング, ピーク立ち上がりを重視した平滑化を行なうことによって得られたスペクトル包絡線を視察することによって, /me/, /ne/ の話者によらない特徴を抽出することができた. この特徴は第2, 3, 4 フォルマントの子音部から母音部にかけての移動量と, 子音部から母音部へのわたりの部分での第2, 3, 4 フォルマントのパワーの立ち上がりの緩急という簡単なものである. この特徴によって 98% の識別率が得られた. このことはこれら二つの特徴の有効性を示しているものである. また, このわたり部のフォルマントの変化は微妙なものであり, 高い分解能と短い時間でスペクトル分析を必要とするので, 本論文で提案した適応



(a) /m/ のスペクトル包絡線



(b) /n/ のスペクトル包絡線

図 3.8: スペクトル包絡線

アルゴリズムによる短時間スペクトル分析法は有効な周波数分析法であると考えられる。本論文では、適応アルゴリズムによる短時間スペクトル分析法が音声分析に有効な周波数分析法であることを示すことに重点を置いたため、コンピュータによるこれら音節の自動認識システムの作成までには至らなかった。これからの課題はより多くのデータによってこれらの特徴の有効性をさらに確かめるとともにコンピュータによるこれら音節の自動認識システムの作成である。

第 4 章

まとめ

本論文では、取り扱う信号を帯域制限周期波形に限定し、すべての周波数成分に対して一つの適応アルゴリズムを適用した方法と各周波数成分ごとに適応アルゴリズムを適用した方法の周波数分析法を提案し、その収束条件を求めるとともに、その性能の検討を行なった。これら二つの周波数分析法では、任意の分析間隔・フレーム長でのスペクトル分析を行なうことができる。分析精度と信号に対する追従性などからすべての周波数成分に対して一つの適応アルゴリズムを適用した方法は時間変化の少ない信号を精度良く分析する場合や楽器の構成音を分析する場合のように信号を構成する周波数成分がわかっている場合に有効な周波数分析法である。また、各周波数成分ごとに適応アルゴリズムを適用した方法は、信号は帯域制限周期波形であるが、どの周波数成分が含まれているかがわからない場合や音を構成している周波数成分がわからない場合には音を構成していると考えられるすべての周波数成分に対して一つの適応アルゴリズムを適用した方法に比べ、信号に対する追従性が良いので採譜や音声分析などに有効であることがシミュレーションから示された。また、採譜や音声分析への応用実験から確かめた。その結果、採譜処理における周波数分析に用いた場合、二つの周波数分析法とも、従来、最も効率的で精度が良いと言われてきた高速フーリエ変換による複素スペクトル面での内挿法より良い精度で音階の同定が行なえることがわかった。また、これら二つの周波数分析法では各周波数成分ごとに適応アルゴリズムを適用した方が良い結果が得られた。これは、採譜処理における周波数分析においては信号に対する追従性が分析精度以上に必要であることを意味していると考えられる。また、音声分析、特に鼻子

音の特徴抽出では、100ch/oct 程度の高分解能周波数分析が必要であり、適応アルゴリズムによる短時間スペクトル分析法は有効な周波数分析法であると考えられる。

第 II 部

適応フィルタによる小振幅周期性雑音除去システム の構築とその応用

第 1 章

適応フィルタ

1.1 はじめに

本章では、本論文で用いる適応フィルタについての説明を行ない、その収束条件の求め方や適応フィルタの持つ特徴について述べる。

1.2 適応性とは

ウィナーフィルタは信号と雑音の統計的性質が既知であり、時間的に変化しないような定常過程の場合に有効である。しかし、現実の場面を考えれば、このような先験的な情報が十分に得られない場合や信号や雑音が非定常な場合などが多い。このような場合には、適応フィルタを用いるのが良い。適応フィルタはある意味で自己設計する装置であり、一組の調節可能なパラメータを持っていて、扱っている信号の推定した統計的性質に基づいてそのパラメータ値が自動的にコントロールされるフィルタである。この適応フィルタの理論は最適フィルタの設計理論と密接な関係がある。しかし、最適フィルタの設計理論が、扱っている信号の性質についての知識を用い、2乗平均誤差を最小にすることによってフィルタの最適値を求めることであるのに対して、適応フィルタの理論は扱っている信号の性質が十分に得られない場合においてもフィルタのパラメータを調節するアルゴリズムを決め、このアルゴリズムを用いて多くの適応を繰り返すことによってフィルタの特性を最適フィルタのそれに近づけようとするものである。適応アルゴリズムは、非繰り返し型のアルゴリズムと繰り返し型のアルゴリズムに分けられる。本論文では繰り返し型のアルゴリズム

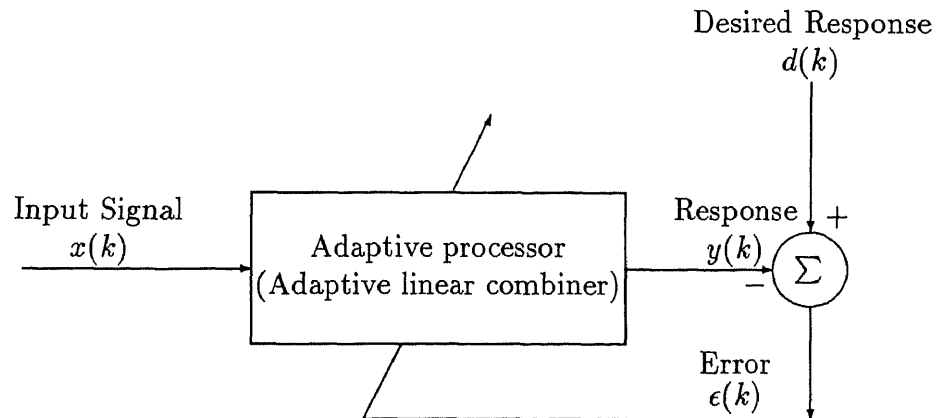


図 1.1: 基本的な適応システム

ムの一つである最小 2 乗平均 (least mean square) アルゴリズム (以下, LMS アルゴリズムと呼ぶ) を用いる.

1.3 基本的な LMS アルゴリズムを用いた適応システム

一般的に用いられるデジタルフィルタは有限長インパルス応答 (FIR) フィルタか無限長インパルス応答 (IIR) フィルタかのいずれかである. 本論文では構造が簡単でかつ安定である FIR フィルタを用いた適応フィルタを取り扱う. 図 1.1 に一般的な適応システムを, 図 1.2 にその内部プロセッサを示す. 本論文で扱う適応フィルタは可変なタップ重みを有する FIR フィルタとそれを制御する LMS アルゴリズムから構成される.

1.4 LMS アルゴリズム

図 1.2 において適応フィルタの出力は, 入力 $x(k)$ と係数 w_i の畳み込みで与えられるのでつぎのような式になる.

$$y(k) = \sum_{i=0}^{L-1} w_i x(k-i) \quad (1.1)$$

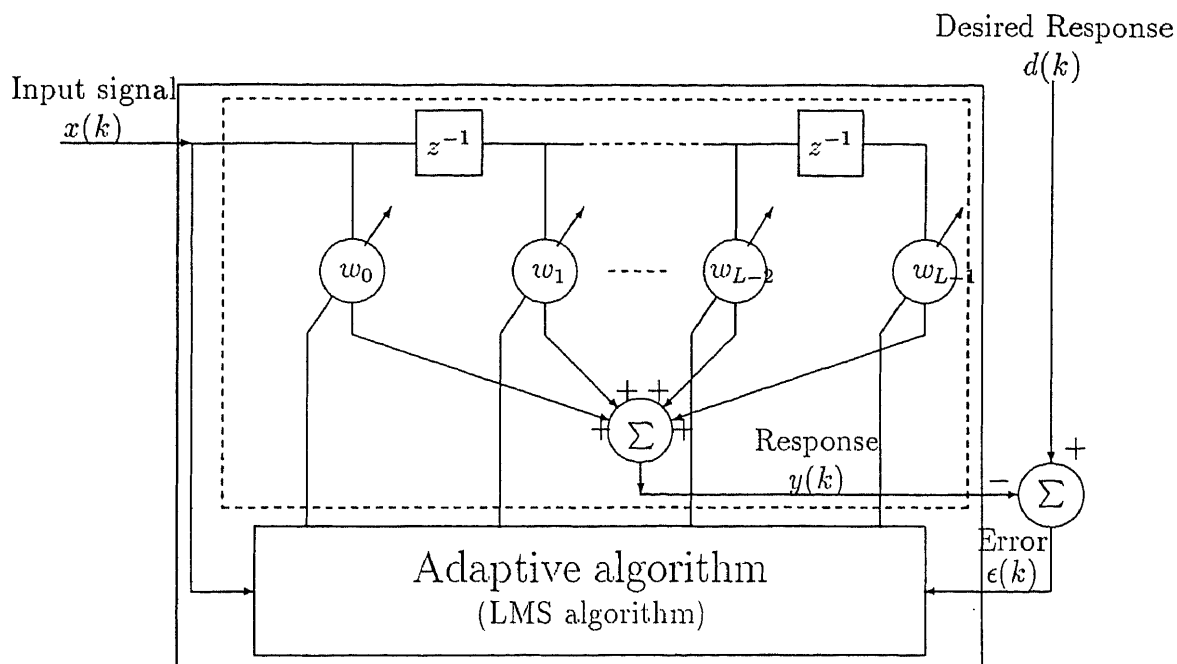


図 1.2: 基本的な適応システムの内部プロセッサ

ここで、 L は係数ベクトルの数である。係数ベクトル \mathbf{W} および入力ベクトル $\mathbf{X}(k)$ をつぎのように定義する。

$$\mathbf{W} = [w_0, \dots, w_i, \dots, w_{L-1}]^T \quad (1.2)$$

$$\mathbf{X}(k) = [x(k), \dots, x(k-i), \dots, x(k-L+1)]^T \quad (1.3)$$

これらの式を用いれば、式 (1.1) はつぎのようになる。

$$y(k) = \mathbf{X}^T(k)\mathbf{W} = \mathbf{W}^T\mathbf{X}(k) \quad (1.4)$$

瞬時誤差 $\epsilon(k)$ は、希望応答とフィルタ出力との差からつぎのように表すことができる。

$$\epsilon(k) = d(k) - y(k) = d(k) - \mathbf{W}^T\mathbf{X}(k) \quad (1.5)$$

ここで、瞬時誤差 $\epsilon(k)$ 、希望応答 $d(k)$ 、入力ベクトル $\mathbf{X}(k)$ が定常であると仮定すれば、2乗平均誤差 $E[\epsilon^2(k)]$ はつぎのようになる。

$$\begin{aligned} \text{MSE} &= E[\epsilon^2(k)] \\ &= E[\{d(k) - \mathbf{W}^T\mathbf{X}(k)\}^2] \end{aligned}$$

$$= E[d^2(k)] - 2E[d(k)\mathbf{X}^T(k)]\mathbf{W} + \mathbf{W}^T E[\mathbf{X}(k)\mathbf{X}^T(k)]\mathbf{W} \quad (1.6)$$

ここで、 $E[\cdot]$ は期待値を示す。つぎに入力相関行列 \mathbf{R} と、希望応答と入力の相互相関ベクトル \mathbf{p} をつぎのように定義する。

$$\mathbf{R} = E[\mathbf{X}(k)\mathbf{X}^T(k)] \quad (1.7)$$

$$\mathbf{p} = E[d(k)\mathbf{X}(k)] \quad (1.8)$$

これらの式を用いて2乗平均誤差を表すとつぎのようになる。

$$\text{MSE} = E[d^2(k)] - 2\mathbf{p}^T\mathbf{W} + \mathbf{W}^T\mathbf{R}\mathbf{W} \quad (1.9)$$

この式は明らかに係数ベクトル \mathbf{W} の2次関数であり、最小値を有する。係数ベクトル \mathbf{W} で偏微分すれば、つぎのようになる。

$$\begin{aligned} \nabla &= \frac{\partial \text{MSE}}{\partial \mathbf{W}} \\ &= -2\mathbf{p} + 2\mathbf{R}\mathbf{W} \end{aligned} \quad (1.10)$$

以上のことから係数ベクトルの最適解 \mathbf{W}^* はつぎのようになる。

$$\mathbf{W}^* = \mathbf{R}^{-1}\mathbf{p} \quad (1.11)$$

この式は Wiener-Hoff 方程式の離散表現に等しい。

1.4.1 LMS アルゴリズムの導出

LMS アルゴリズムは Widrow と Hoff によって最初に導かれた。このアルゴリズムは、その簡単さと実現の容易さから多くの適応信号処理に応用されている。2乗平均誤差の推定量として瞬時2乗誤差 $e^2(k)$ そのものを用いて勾配ベクトル $\nabla(k)$ の推定量 $\hat{\nabla}(k)$ を求める。つまり、このアルゴリズムの利点は相関関数の測定を必要としないことである。勾配ベクトルの推定量 $\hat{\nabla}(k)$ はつぎのようになる。

$$\begin{aligned} \hat{\nabla}(k) &= \frac{\partial e^2(k)}{\partial \mathbf{W}(k)} \\ &= -2e(k)\mathbf{X}(k) \end{aligned} \quad (1.12)$$

つぎに LMS アルゴリズムを定式化する。このアルゴリズムでは、係数ベクトルを推定勾配ベクトルの向きにつぎのように変化させる。

$$\begin{aligned} \mathbf{W}(k+1) &= \mathbf{W}(k) + \mu\{-\widehat{\nabla}(k)\} \\ &= \mathbf{W}(k) + 2\mu\epsilon(k)\mathbf{X}(k) \end{aligned} \quad (1.13)$$

この式は、入力ベクトルと誤差信号の積に比例する量を古い推定量に加えれば、更新された係数ベクトルの推定量が得られることを示している。この式は適応過程を構成している。この時、瞬時誤差 $\epsilon(k)$ はつぎに示すような式になり、入力を $\mathbf{X}(k)$ および $d(k)$ 、出力を $\epsilon(k)$ としたフィルタリング過程を構成している。

$$\epsilon(k) = d(k) - y(k) = d(k) - \mathbf{W}^T(k)\mathbf{X}(k) \quad (1.14)$$

これら簡単な二つの式で、LMS アルゴリズムは完全に記述することができる。しかも、勾配ベクトルの推定量 $\widehat{\nabla}(k)$ の分散が大きいのので一見 LMS アルゴリズムは良い特性が得られないように思われるが再帰アルゴリズムであるため、適応過程の繰り返しによってこの粗い推定が効果的に平均化され最適解に収束することが証明されている。

1.4.2 収束条件

係数の収束条件

LMS アルゴリズムはその動作からわかるように、フィルタ係数にもランダムさが伝搬してしまふ。従って、厳密には非定常確率過程として取り扱う必要がある。しかし、解析を簡単化するため係数ベクトル $\mathbf{W}(k)$ 、入力ベクトル $\mathbf{X}(k)$ や希望応答 $d(k)$ が互いに統計的に独立であると仮定する。式 (1.13) に式 (1.14) を代入し、瞬時誤差 $\epsilon(k)$ を消去するとつぎのようになる。

$$\begin{aligned} \mathbf{W}(k+1) &= \mathbf{W}(k) + 2\mu\mathbf{X}(k)\{d(k) - \mathbf{W}^T(k)\mathbf{X}(k)\} \\ &= \{\mathbf{I} - 2\mu\mathbf{X}(k)\mathbf{X}^T(k)\}\mathbf{W}(k) + 2\mu d(k)\mathbf{X}(k) \end{aligned} \quad (1.15)$$

ここで、 \mathbf{I} は単位行列である。係数誤差ベクトル $\mathbf{C}(k)$ をつぎのように定義する。 k が無限大になる時に係数誤差ベクトル $\mathbf{C}(k)$ が零ベクトルになれば、係

数ベクトルが最適解に収束したことを意味する.

$$C(k) = W(k) - W^* \quad (1.16)$$

この式を式(1.15)に代入し, 係数ベクトル $W(k)$ を消去すればつぎのようになる.

$$\begin{aligned} W(k+1) &= \{I - 2\mu X(k)X^T(k)\}\{C(k) + W^*\} + 2\mu d(k)X(k) \\ &= \{I - 2\mu X(k)X^T(k)\}C(k) + W^* \\ &\quad + 2\mu\{d(k)X(k) - X(k)X^T(k)W^*\} \end{aligned} \quad (1.17)$$

ここで, 最適解 W^* を左辺に移項すると, $W(k+1) - W^*$ は係数誤差ベクトルの更新値となり, 式(1.17)はつぎのようになる.

$$\begin{aligned} C(k+1) &= \{I - 2\mu X(k)X^T(k)\}C(k) \\ &\quad + 2\mu\{d(k)X(k) - X(k)X^T(k)W^*\} \end{aligned} \quad (1.18)$$

しかし, LMS アルゴリズムは雑音を含む勾配ベクトルの推定量によっているのでその収束条件を求めることは困難である. そこで, LMS アルゴリズムにおいては集合平均の動的振舞いが議論される. このことから本論文でも, 平均係数誤差ベクトル $E[C(k)]$ の収束条件を求めることにする. 係数ベクトル $W(k)$, 入力ベクトル $X(k)$ や希望応答 $d(k)$ が互いに統計的に独立であると仮定しているので式(1.18)はつぎのようになる.

$$\begin{aligned} E[C(k+1)] &= E\{\{I - 2\mu X(k)X^T(k)\}\}E[C(k)] \\ &\quad + 2\mu\{E[d(k)X(k)] - E[X(k)X^T(k)]W^*\} \\ &= (I - \mu R)E[C(k)] + 2\mu(p - RW^*) \end{aligned} \quad (1.19)$$

ここで, Wiener-Hoff の関係式 $p = RW^*$ より上式はつぎのようになる.

$$E[C(k+1)] = (I - 2\mu R)E[C(k)] \quad (1.20)$$

つぎに, 相関行列 R をその固有値と固有ベクトルによって対角化を行なう. 相関行列 R はその固有値と固有ベクトルを用いれば, つぎのように表すことができる.

$$Q^T R Q = \Lambda \quad (1.21)$$

ただし、対角行列 \mathbf{A} は相関行列 \mathbf{R} の固有値からなり、ユニタリ行列 \mathbf{Q} の列は固有ベクトルである。このようにすれば、式(1.20)はつぎのように対角化できる。

$$\begin{aligned} E[\mathbf{V}(k+1)] &= (\mathbf{I} - 2\mu\mathbf{A})E[\mathbf{V}(k)] \\ &= (\mathbf{I} - 2\mu\mathbf{A})^k E[\mathbf{V}(0)] \end{aligned} \quad (1.22)$$

ここで、 $\mathbf{V}(k+1)$ はつぎのような関係を満たす。

$$\mathbf{V}(k) = \mathbf{Q}^T \mathbf{C}(k) \quad (1.23)$$

式(1.22)は等比級数であるので、すべてのモードに対してつぎのような関係を満たせば LMS アルゴリズムは収束すると考えられる。

$$0 < |1 - 2\mu\lambda_i| < 1 \quad \text{for all } i \quad (1.24)$$

ここで、 λ_i は相関行列 \mathbf{R} の i 番目の固有値である。ステップサイズパラメータ μ からみれば、つぎのような関係を満たせば LMS アルゴリズムは収束すると考えられる。

$$0 < \mu < \frac{1}{\lambda_{max}} \quad (1.25)$$

ただし、 λ_{max} は相関行列 \mathbf{R} の最大固有値である。

学習曲線の収束条件

LMS アルゴリズムが収束することを証明するためには、係数ベクトルが収束するとともに、学習曲線も収束しなくてはならない。そこで次に学習曲線が収束するための条件を求める。前に述べたように LMS アルゴリズムは雑音を含んだ勾配ベクトルによっているため、ここでも 2 乗平均誤差 $E[\epsilon^2(k)]$ の動的振舞いから学習曲線が収束するための条件を求めることにする。2 乗平均誤差 $E[\epsilon^2(k)]$ を ξ とおくと最小 2 乗平均誤差 ξ_{min} はつぎのようになる。

$$\begin{aligned} \xi_{min} &= E[\{d(k) - \mathbf{W}^{*T} \mathbf{X}(k)\}^2] \\ &= E[d^2(k)] - 2\mathbf{W}^{*T} E[d(k)\mathbf{X}^T(k)] + \mathbf{W}^{*T} E[\mathbf{X}(k)\mathbf{X}^T(k)] \mathbf{W}^* \\ &= E[d^2(k)] - 2\mathbf{W}^{*T} \mathbf{p} + \mathbf{W}^{*T} \mathbf{R} \mathbf{W}^* \end{aligned} \quad (1.26)$$

Wiener-Hoff の関係式より最小 2 乗平均誤差 ξ_{min} は次式で与えられる.

$$\xi_{min} = E[d^2(k)] - \mathbf{W}^{*T} \mathbf{R} \mathbf{W}^* \quad (1.27)$$

つぎに 2 乗平均誤差 ξ_k を求める. 2 乗平均誤差 ξ_k は最小 2 乗平均誤差 ξ_{min} を用いて表すとつぎのようになる.

$$\begin{aligned} \xi_k &= E[\{d(k) - \mathbf{W}^T(k) \mathbf{X}(k)\}^2] \\ &= E[d^2(k)] - 2E[\mathbf{W}^T(k)]E[d(k) \mathbf{X}^T(k)] \\ &\quad + E[\mathbf{W}^T(k)]E[\mathbf{X}(k) \mathbf{X}^T(k)]E[\mathbf{W}(k)] \\ &= \xi_{min} + E[\{\mathbf{W}^* - \mathbf{W}(k)\}^T] \mathbf{R} E[\{\mathbf{W}^* - \mathbf{W}(k)\}] \\ &= \xi_{min} + E[\mathbf{C}^T(k) \mathbf{C}(k)] \text{tr}[\mathbf{R}] \end{aligned} \quad (1.28)$$

ここで, $\text{tr}[\mathbf{R}]$ は \mathbf{R} のトレースを示し, $\text{tr}[\mathbf{R}] = E[\mathbf{X}^T(k) \mathbf{X}(k)]$ である. このことから学習曲線が収束するためにはつぎのような条件を満たせばよいということになる.

$$E[\mathbf{C}^T(k+1) \mathbf{C}(k+1)] - E[\mathbf{C}^T(k) \mathbf{C}(k)] < 0 \quad (1.29)$$

この式に式 (1.13) を代入するとつぎのようになる.

$$\begin{aligned} &E[\mathbf{C}^T(k+1) \mathbf{C}(k+1)] - E[\mathbf{C}^T(k) \mathbf{C}(k)] \\ &= E[\{\mathbf{W}(k) + 2\mu\epsilon(k) \mathbf{X}(k) - \mathbf{W}^*\}^T \{G(k) + 2\mu\epsilon(k) \mathbf{X}(k) - \mathbf{W}^*\} \\ &\quad - \{\mathbf{W}(k) - \mathbf{W}^*\}^T \{\overline{\mathbf{W}}(k) - \overline{\mathbf{W}}^*\}] \\ &= E[\{\mathbf{W}(k) - \mathbf{W}^*\}^T 4\mu\epsilon(k) \mathbf{X}(k) \\ &\quad + 4\mu^2\epsilon^2(k) \mathbf{X}^T(k) \mathbf{X}(k)] \\ &= -4\mu\{1 - \mu \text{tr}[\mathbf{R}]\} E[\epsilon^2(k)] < 0 \end{aligned} \quad (1.30)$$

このことから, つぎのような関係を満たせば, LMS アルゴリズムは収束すると考えられる.

$$1 - \mu \text{tr}[\mathbf{R}] < 0 \quad (1.31)$$

ステップサイズパラメータ μ からみれば, つぎのような関係を満たせば LMS アルゴリズムは収束すると考えられる.

$$0 < \mu < \frac{1}{\sum_{i=0}^{L-1} \lambda_i} = \frac{1}{\text{tr}[\mathbf{R}]} \quad (1.32)$$

1.4.3 時定数

係数の収束に関する時定数

式(1.22)より係数誤差ベクトルの*i*番目の式はつぎのような等比級数となる.

$$E[v_i(k+1)] = (1 - 2\mu\lambda_i)^k E[v_i(0)] \quad k = 0, 1, \dots, L-1 \quad (1.33)$$

その項比 r はつぎのようになる.

$$r_i = 1 - 2\mu\lambda_i \quad k = 0, 1, \dots, L-1 \quad (1.34)$$

ここで, 一線り返し周期を時間の単位にとり, 指数包絡線の時定数 τ_i を次式のように選べば, 包絡線を等比級数に合わせることができる.

$$r_i = \exp\left(-\frac{1}{\tau_i}\right) = 1 - \frac{1}{\tau_i} + \frac{1}{2!\tau_i^2} - \dots \quad (1.35)$$

ここで, 適応過程がゆっくりした場合つまり, 時定数 τ_i が大きい場合を考えることにする. この場合には式(1.34), (1.35)よりつぎのような関係が成り立つ.

$$1 - 2\mu\lambda_i = 1 - \frac{1}{\tau_i} \quad k = 0, 1, \dots, L-1 \quad (1.36)$$

よって, 時定数 τ_i はつぎのようになる.

$$\tau_i = \frac{1}{2\mu\lambda_i} \quad k = 0, 1, \dots, L-1 \quad (1.37)$$

学習曲線の収束に関する時定数

式(1.28)より2乗平均誤差はつぎのように与えられる.

$$\xi_k = \xi_{min} + E[C^{*T}]RE[C^*] \quad (1.38)$$

この式を式(1.21), (1.23)を用いて対角化するとつぎのようになる.

$$\xi_k = \xi_{min} + E[V^{*T}]AE[V^*] \quad (1.39)$$

行列 A は対角化行列であるから, $E[V^{*T}]AE[V^*]$ は相互積の項を含まないのでつぎのように表すことができる.

$$\xi_k = \xi_{min} + \sum_{i=0}^{L-1} \lambda_i E[v_i^2(k)] \quad k = 0, 1, \dots, L-1 \quad (1.40)$$

この式に式 (1.33) を代入するとつぎのような等比級数になる.

$$\xi_k = \xi_{min} + \sum_{i=0}^{L-1} (1 - 2\mu\lambda_i)^{2k} \lambda_i E[v_i^2(0)] \quad k = 0, 1, \dots, L-1 \quad (1.41)$$

係数の収束に関する時定数と同様な手順を用いれば, 時定数 τ_i はつぎのようになる.

$$\tau_i = \frac{1}{4\mu\lambda_i} \quad k = 0, 1, \dots, L-1 \quad (1.42)$$

1.4.4 過剰誤差

LMS アルゴリズムでは, 先にも述べたように雑音を含んだ勾配ベクトルの推定値 $\widehat{\nabla}(k)$ を用いているために多数回の適応を繰り返しても集合平均的に収束するだけで係数ベクトル $\mathbf{W}(k)$ は最適解 \mathbf{W}^* の近傍でゆらぎを生じる. そのため, 最終的な 2 乗平均誤差 ξ_k は最小 2 乗平均誤差 ξ_{min} より大きくなる. この誤差量を過剰 2 乗誤差 excess MSE と定義すれば, excess MSE はつぎのようになる.

$$\begin{aligned} \text{excessMSE} &= E[\xi_k - \xi_{min}] \quad (\text{after convergence}) \\ &= E[\mathbf{V}^T(k)] \Delta E[\overline{\mathbf{V}}(k)] \\ &\simeq \mu \xi_{min} \text{tr}[\mathbf{R}] \end{aligned} \quad (1.43)$$

このように LMS アルゴリズムではステップサイズパラメータ μ を大きく取れば, 時定数 τ_i は小さくなり収束は速いが収束後の分散は大きくなる. 逆にステップサイズパラメータ μ を小さく取れば, 時定数 τ_i は大きくなり収束が遅くなるが収束後の分散は小さくなることがわかる.

第 2 章

周期性雑音除去システム

2.1 はじめに

テレビジョンの画質を改善するには伝送路で加わるゴーストや雑音を除去しなくてはならない。これらのうち、周期性雑音は電源からのハム、隣接チャンネルからの妨害、色副搬送波と音声搬送波とのビート妨害、他の電波源からの妨害などによって生じ、線スペクトルに近い周波数スペクトルをもち、周波数に対応した画像としてしま状の妨害となる。このことから周期性雑音の除去はテレビジョンの画質を改善するには重要なものであると考えられるがゴーストキャンセラに関する研究や非周期性雑音の除去に関する研究に比べあまり研究がなされていない。

従来、このような周期性雑音の除去にはウィナーフィルタが用いられてきたが、この方法では原画像を必要とし、テレビのように原画像が得られず、実時間処理が必要な場合を考えると有効な手段とは言えない。そこで、適応雑音除去システムを用いることを試みる。適応雑音除去システムを用いる利点は、信号に対する追従性などから実時間処理に有利であることと、雑音源の情報が利用できない場合においても雑音が周期性雑音であれば 1 チャンネル適応雑音除去システムによって雑音除去が行なえることが Widrow らによって証明されていることである。しかし、テレビ信号のように雑音が小振幅であり、信号が急峻な変化をともなうような場合には有効に動作しないことがある。そこで、本論文ではこのような適応雑音除去システムに原島らによって提案された信号を大振幅成分と小振幅成分とに効果的に分離することができる ϵ -分離非線形

デジタルフィルタを組み込んだ雑音除去システムを提案し、上述のような場合においても有効に雑音除去が行なえることをシミュレーションなどによって示す。

2.2 1 チャネル適応雑音除去システムとその問題点

適応雑音除去システムが適切に動作するのは雑音 $n(k)$ と非常に相関が強く信号 $s(k)$ とは無相関である雑音 $n'(k)$ が得られる場合と雑音が周期性雑音である場合である。テレビ信号などを考える場合にはこのような雑音源の情報が利用できないことがほとんどであるが、雑音が周期性雑音である場合もある。このような場合、図 2.1 で示されるような 1 チャネル適応雑音除去システムを構成すればうまく動作することが Widrow らによって報告されている。これは信号間に相関がなくなる程度の遅延を与えても雑音が周期性なら雑音間に相関がなくなるためである。このことから図 2.1 において信号 $s(k)$ と $s(k-d)$ との間に相関がなくなる程度の遅延 d を与えシステムが収束した場合を考えればつぎのような関係を満たしていると考えられる。

$$n(k) = \sum_{i=0}^{L-1} w(i)n(k+i-d) \quad (2.1)$$

ここで $w(i)$ はフィルタの係数であり、 $L-1$ はタップ数である。この場合の適応フィルタの出力 $\hat{n}(k)$ はつぎのようになる。

$$\hat{n}(k) = \sum_{i=0}^{L-1} w(i)\{s(k+i-d) + n(k+i-d)\} \quad (2.2)$$

このことからシステム全体の出力 $\hat{s}(k)$ はつぎのようになる。

$$\begin{aligned} \hat{s}(k) &= e(k) = x(k) - \sum_{i=0}^{L-1} w(i)x(k+i-d) \\ &= s(k) + n(k) - \sum_{i=0}^{L-1} w(i)\{s(k+i-d) + n(k+i-d)\} \\ &= s(k) - \sum_{i=0}^{L-1} w(i)s(k+i-d) \end{aligned} \quad (2.3)$$

つまり、このシステムが有効に動作するのは、つぎのような関係が成り立つ場合である。

$$|s(k)| \gg \left| \sum_{i=0}^{L-1} w(i)s(k+i-d) \right| \quad (2.4)$$

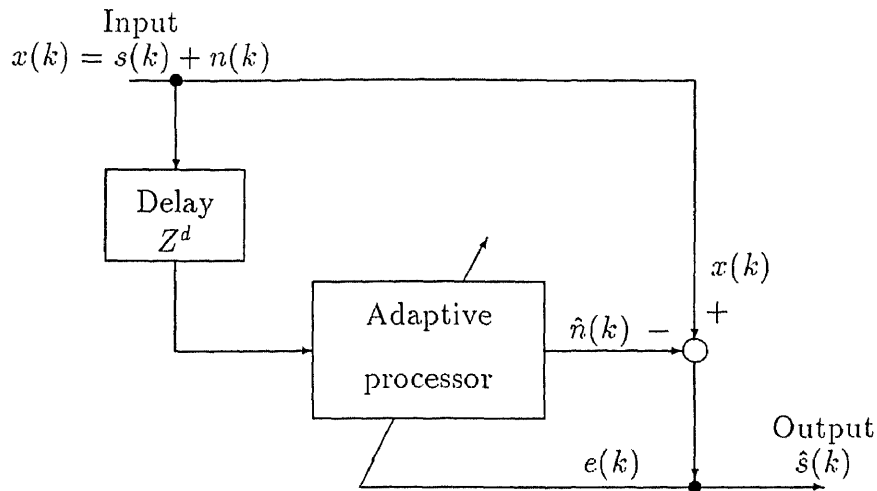


図 2.1: 1チャンネル適応雑音除去システム

テレビなどのように画像信号が伝送される間に周期性雑音加わるような場合を考えれば、一般に画像信号 $s(k)$ はエッジなどによる大振幅成分 $s_l(k)$ と小振幅成分 $s_s(k)$ からなり、伝送する間に加わる周期性雑音 $n(k)$ は小振幅雑音であると考えることができる。このような場合、仮に式 (2.1) の関係を満たすようにフィルタ係数が収束したとしても大振幅信号成分 $s_l(k)$ のために式 (2.4) のような関係を満たすとは限らず、 $\sum_{i=0}^{L-1} w(i)s(k+i-d)$ の値が大きくなり、ゴースト状の雑音となることがある。そこで、本研究では大振幅信号成分 $s_l(k)$ の影響を除去するために ϵ -分離非線形デジタルフィルタを組み込むことを考える。

2.3 ϵ -分離非線形デジタルフィルタの原理

ϵ -分離非線形デジタルフィルタを組み込んだ雑音除去システムについて述べる前に ϵ -分離非線形デジタルフィルタについて簡単に説明し、信号を大振幅成分と小振幅成分とに分離できることを示す。 ϵ -分離非線形デジタルフィルタは原島らによって提案されたデジタルフィルタであり、線形デジタルフィルタを非線形化することによって導くことができる。まず、非巡回

形線形デジタルフィルタは、一般につぎのように表される。

$$y(k) = \sum_{i=-N}^M a(i)x(k-i) \quad (2.5)$$

ここで、 $x(k)$ は入力、 $y(k)$ は出力、 $a(i)$ はフィルタ係数であり、フィルタの直流伝達特性を1とするためにつぎの条件を満たすと考える。

$$\sum_{i=-N}^M a(i) = 1 \quad (2.6)$$

式(2.6)を用いると、式(2.5)はつぎのように変形される。

$$\begin{aligned} y(k) &= \sum_{i=-N}^M a(i)[x(k) - \{x(k) - x(k-i)\}] \\ &= x(k) \sum_{i=-N}^M a(i) - \sum_{i=-N}^M a(i)\{x(k) - x(k-i)\} \\ &= x(k) - \sum_{i=-N}^M a(i)\{x(k) - x(k-i)\} \end{aligned} \quad (2.7)$$

ここで、図2.2のような非線形関数 $F[x]$ を考える。この関数についてつぎの式が成り立つ。

$$|F[x]| \leq \epsilon_0 ; \quad -\infty \leq x \leq \infty \quad (2.8)$$

この式を式(2.7)に導入し、次式で表されるフィルタを考える。

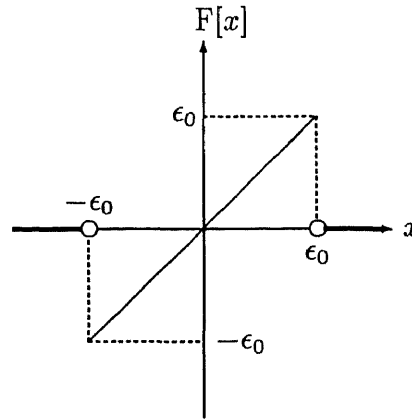
$$y(k) = x(k) - \sum_{i=-N}^M a(i)F[x(k) - x(k-i)] \quad (2.9)$$

この非線形フィルタにおいて、入出力の差 $u(k) = x(k) - y(k)$ はつぎのようになる。

$$\begin{aligned} |u(k)| &= |x(k) - y(k)| = \left| \sum_{i=-N}^M a(i)F[x(k) - x(k-i)] \right| \\ &\leq \sum_{i=-N}^M |a(i)| |F[x(k) - x(k-i)]| \leq \epsilon_0 \sum_{i=-N}^M |a(i)| \end{aligned} \quad (2.10)$$

ここで、 ϵ をつぎのように定義する。

$$\epsilon_0 \sum_{i=-N}^M |a(i)| \equiv \epsilon \quad (2.11)$$

図 2.2: 非線形関数 $F[x]$ の例

このように定義すれば，式 (2.9) で表されるフィルタは入出力信号の差を ϵ 以下に制限しながら信号を平滑化するものであると言える．つまり，このフィルタは急峻に変化する信号成分をそこなく小振幅の成分を取り除くことができるものである．

逆に $u(k)$ を出力するフィルタを考えれば， $u(k)$ はつぎのように表すことができる．

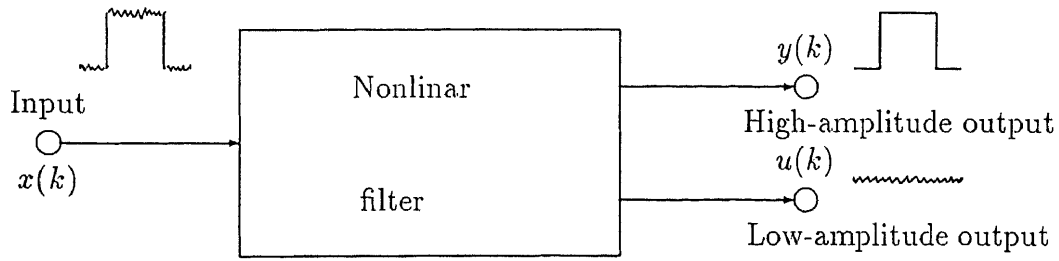
$$u(k) = - \sum_{i=-N}^M a(i) F[x(k) - x(k-i)] \quad (2.12)$$

このフィルタは， $y(k)$ を出力するフィルタとは逆に急峻に変化する大振幅成分を除いた小振幅成分だけを出力するフィルタであると考えることができる．

このように ϵ -分離非線形デジタルフィルタは大振幅信号成分 $y(k)$ と小振幅信号成分 $u(k)$ とを分離できるフィルタと考えることができる (図 2.3)．

2.4 小振幅周期性雑音除去システム

本論文で提案する ϵ -分離非線形デジタルフィルタを組み込んだ雑音除去システムを図 2.4 に示す． ϵ -分離非線形デジタルフィルタは前節で述べたように画像信号 $s(k)$ 中のエッジなどによる大振幅成分 $s_l(k)$ とそれら以外的小振幅成分 $s_s(k)$ および小振幅の雑音 $n(k)$ とを分離することができる． ϵ -分離非線形デジタルフィルタの出力 $u(k)$ は次式で示すような小振幅成分 $s_s(k)$ お

図 2.3: ϵ -分離非線形デジタルフィルタ

よび雑音 $n(k)$ の和であると考えられる。

$$u(k) = s_s(k) + n(k) \quad (2.13)$$

このように ϵ -分離非線形デジタルフィルタの出力 $u(k)$ には信号 $s(k)$ の小振幅成分 $s_s(k)$ が含まれているが雑音 $n(k)$ が周期性であるので図 2.1 で示した 1 チャンネル適応雑音除去システムの場合と同様に信号 $s(k)$ とその小振幅成分 $s_s(k)$ との間に相関がなくなる程度の遅延 d を与えても雑音間の相関はなくなり、式 (2.1) の関係を満たすようにフィルタ係数が収束すると考えることができ、その場合のシステム全体の出力 $\hat{s}(k)$ はつぎのようになると考えられる。

$$\begin{aligned} \hat{s}(k) &= e(k) = x(k) - \sum_{i=0}^{L-1} w(i)u(k+i-d) \\ &= s(k) + n(k) - \sum_{i=0}^{L-1} w(i)\{s_s(k+i-d) + n(k+i-d)\} \\ &= s(k) - \sum_{i=0}^{L-1} w(i)s_s(k+i-d) \end{aligned} \quad (2.14)$$

ここで、 $s_s(k)$ は信号中の小振幅成分であり、一般的につぎのような関係が成り立つ

$$|s(k)| \gg |s_s(k)| \quad (2.15)$$

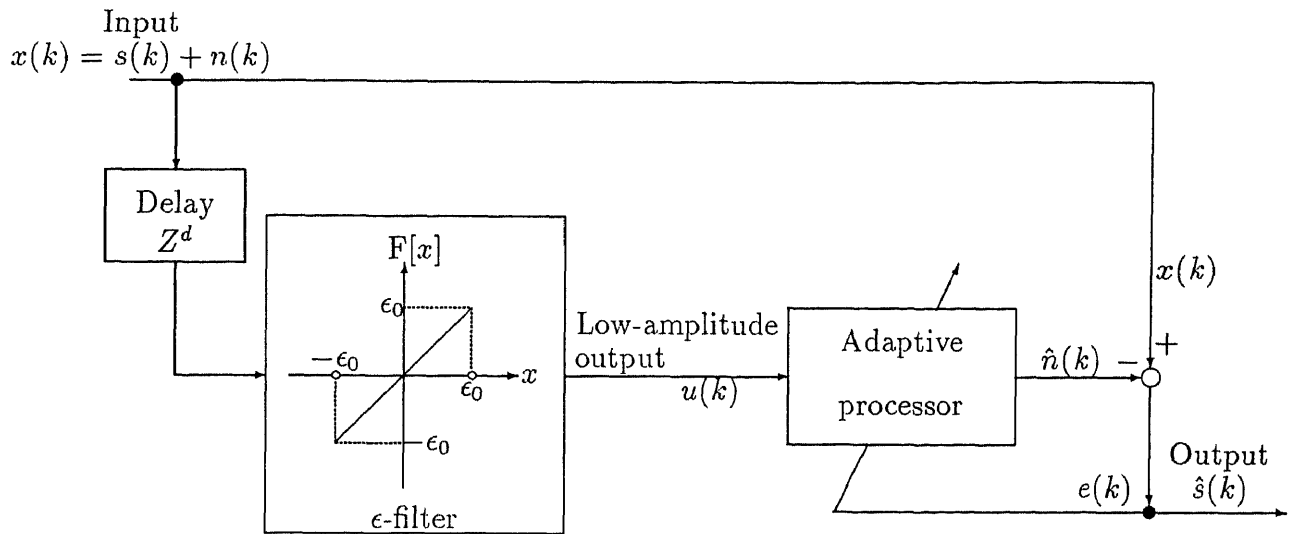


図 2.4: 小振幅周期性雑音除去システム

よって、つぎの関係を満たすと考えられる。

$$\left| \sum_{i=0}^{L-1} w(i)s(k+i-d) \right| \gg \left| \sum_{i=0}^{L-1} w(i)s_s(k+i-d) \right| \quad (2.16)$$

このことから、 ϵ -分離非線形デジタルフィルタを組み込んだ雑音除去システム (図 2.4) は従来の 1 チャンネル適応雑音除去システム (図 2.1) で問題となっていた画像信号中の大振幅成分の影響を受けないと考えることができる。

2.5 シミュレーションと実験

本研究ではテレビ信号などが伝送する間に加わる周期性雑音の除去を最終目標と考えているので原画像が得られないとし、伝送してきた信号を大振幅信号を含む 1 次元信号列として考える。本章では、シミュレーションおよび実際の画像を用いた実験によって本論文で提案した方法の有効性を示す。

2.5.1 シミュレーション

適応フィルタのタップ数 L を 15, ステップサイズパラメータ μ を 0.01, 遅延数 d を 30 とし, ϵ -分離非線形デジタルフィルタのタップ数 L を 7, ϵ_0

を 0.06 として雑音除去システムを構成し、以下に示すような 4 つのタイプの信号に振幅が 0.02 及び 0.04, 周波数がサンプリング周波数の 10 分の 1 の正弦波雑音加わったものに対して雑音の除去を行なった.

例 1 : 階段状の信号列 (1 次元) (図 2.5(a))

例 2 : 直線状の信号列 (1 次元) (図 2.5(b))

例 3 : 種々の性質が混在した信号列 (1 次元) (図 2.5(c))

例 4 : 2 次元信号 (図 2.6)

振幅が 0.04 の場合の雑音除去の結果を図 2.7 から図 2.10 に, 例 1 の場合における適応フィルタの係数ベクトルの収束状況を図 2.11 に示す. また, 次式で示すような入力信号または出力信号と原信号との差の 2 乗平均を原信号の 2 乗平均で規格化したものによって雑音除去の効果を比較し, 表 2.1 にまとめた.

$$\sigma_i = \sqrt{\frac{E[\{x(k) - s(k)\}^2]}{E[s^2(k)]}} \quad (2.17)$$

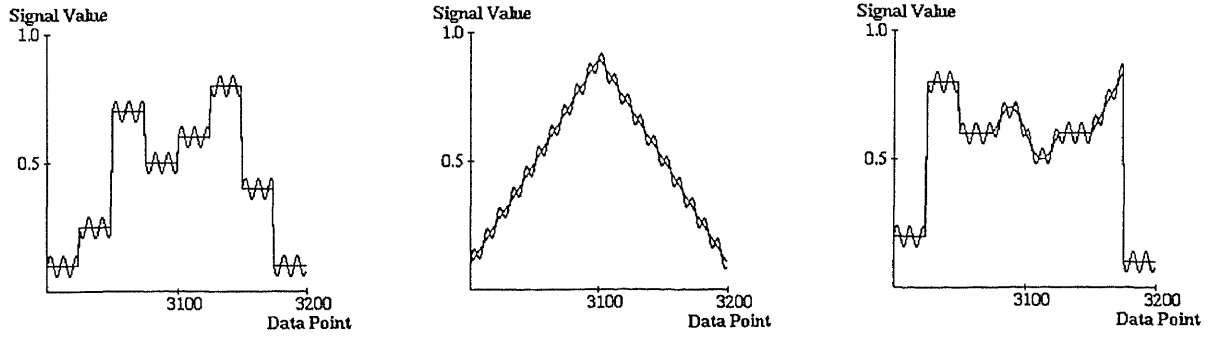
$$\sigma_o = \sqrt{\frac{E[\{\hat{s}(k) - s(k)\}^2]}{E[s^2(k)]}} \quad (2.18)$$

ここで, $s(k)$ は原信号を, $x(k)$ は入力信号を, $\hat{s}(k)$ は出力信号を表す.

従来の 1 チャネル適応雑音除去システムでは, 信号の大振幅特性のために式 (2.4) のような関係が成り立たず, 式 (2.3) からわかるように出力は $\sum_{i=0}^{L-1} w(i) s(k+i-d)$ の影響を受ける. このことが, 表 2.1 において σ_o の値が σ_i の値より大きくなったことや図 2.11(b) のように適応フィルタの係数ベクトルが適切に収束せず, 図 2.7, 図 2.8 において加えた周期性雑音も除去されなかったことの原因として考えられる. しかし, ϵ -分離非線形デジタルフィルタを組み込むことによってこのような信号の大振幅特性の影響を受けなくなり, 雑音が除去できるようになったことがわかる.

2.5.2 実験

つぎに実際の画像信号に対する効果を確かめるために適応フィルタのタップ数 L を 29, ステップサイズパラメータ μ を 0.02, 遅延数 d を 58 とし, ϵ -分離非線形デジタルフィルタのタップ数 L を 7, ϵ_0 を 0.08 として雑音除去システムを構成し, 以下に示す画像信号 (振幅が 0.04, 周波数がサンプリング



(a) 例 1

(b) 例 2

(c) 例 3

図 2.5: 振幅 0.04 の正弦波状雑音を含んだ信号

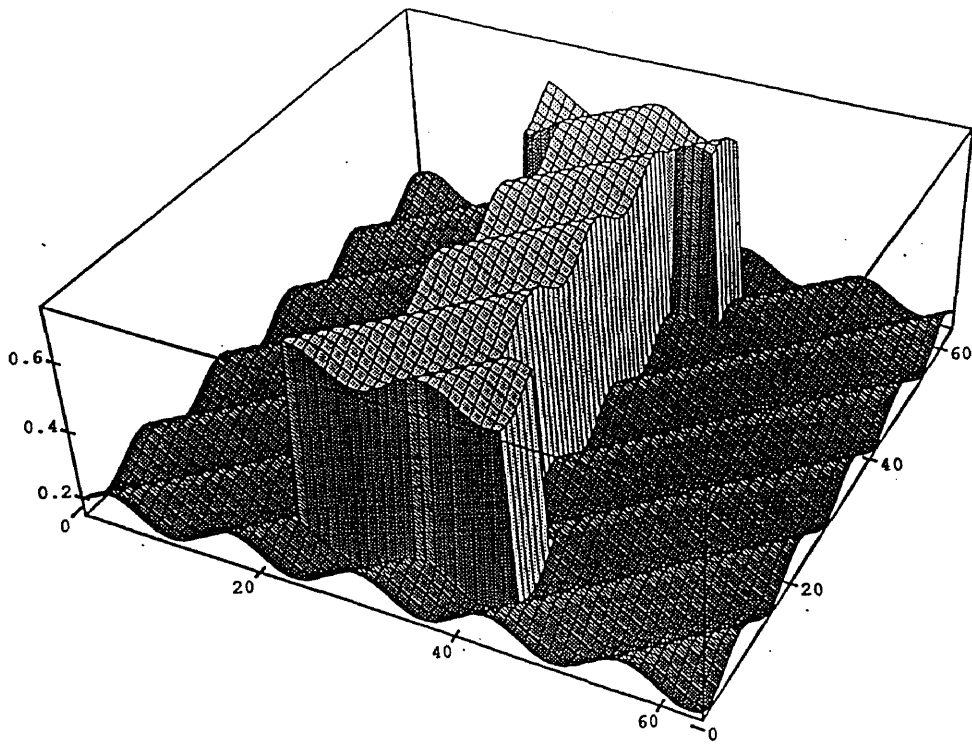


図 2.6: 振幅 0.04 の正弦波状雑音を含んだ信号 (例 4)

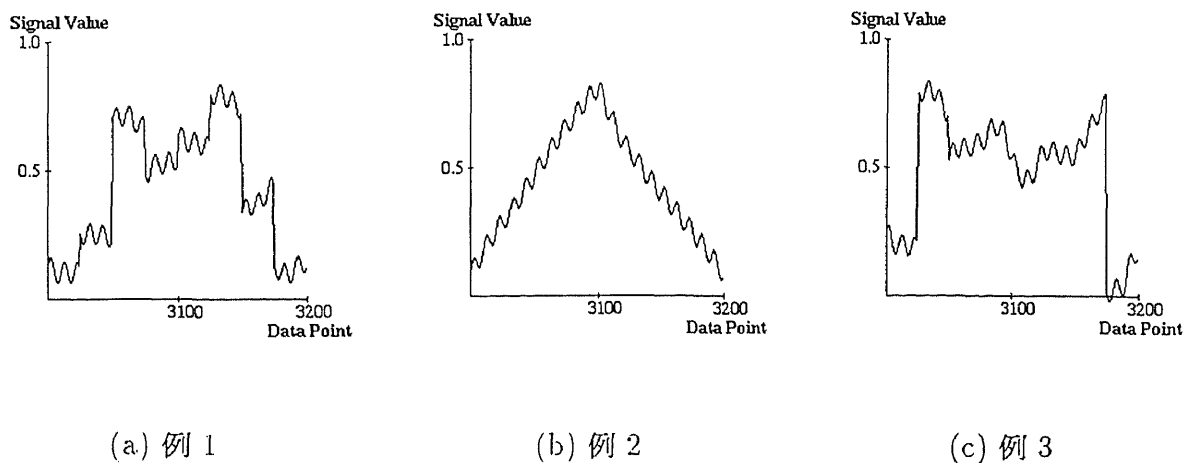


図 2.7: 図 2.1 を用いた場合のシミュレーション結果

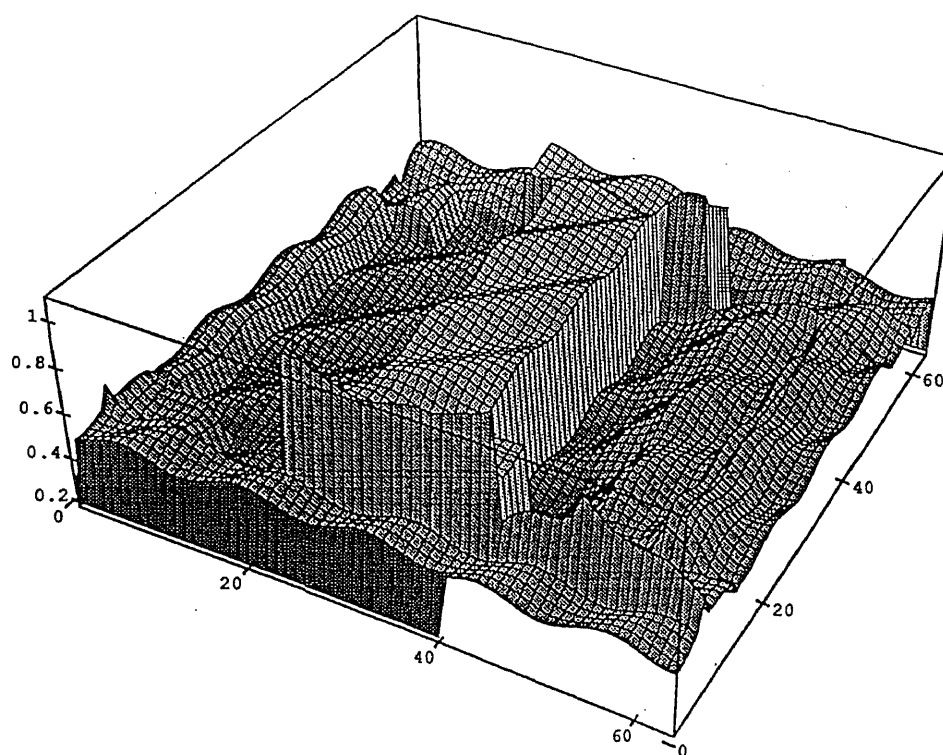


図 2.8: 図 2.1 を用いた場合のシミュレーション結果 (例 4)

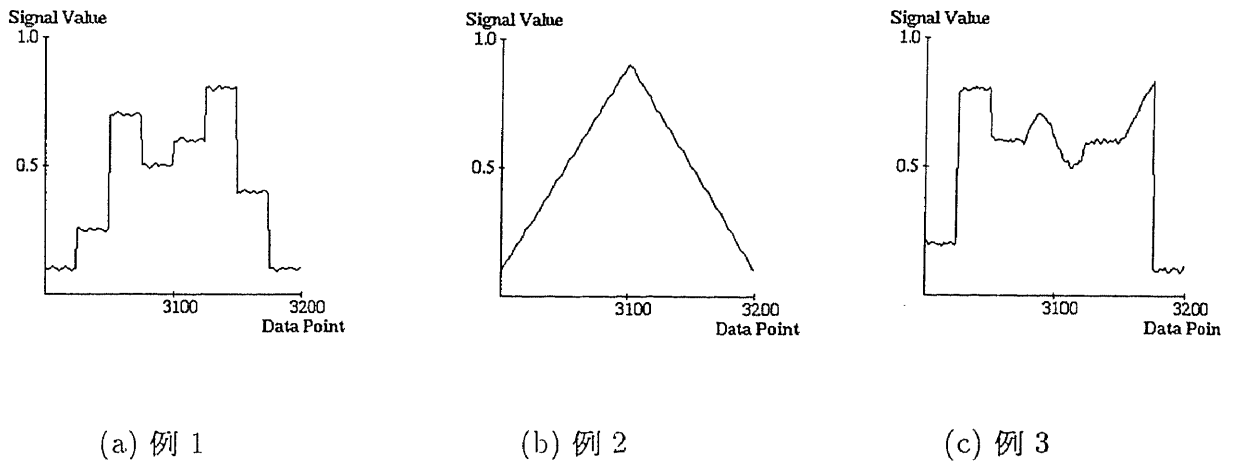


図 2.9: 図 2.4 を用いた場合のシミュレーション結果

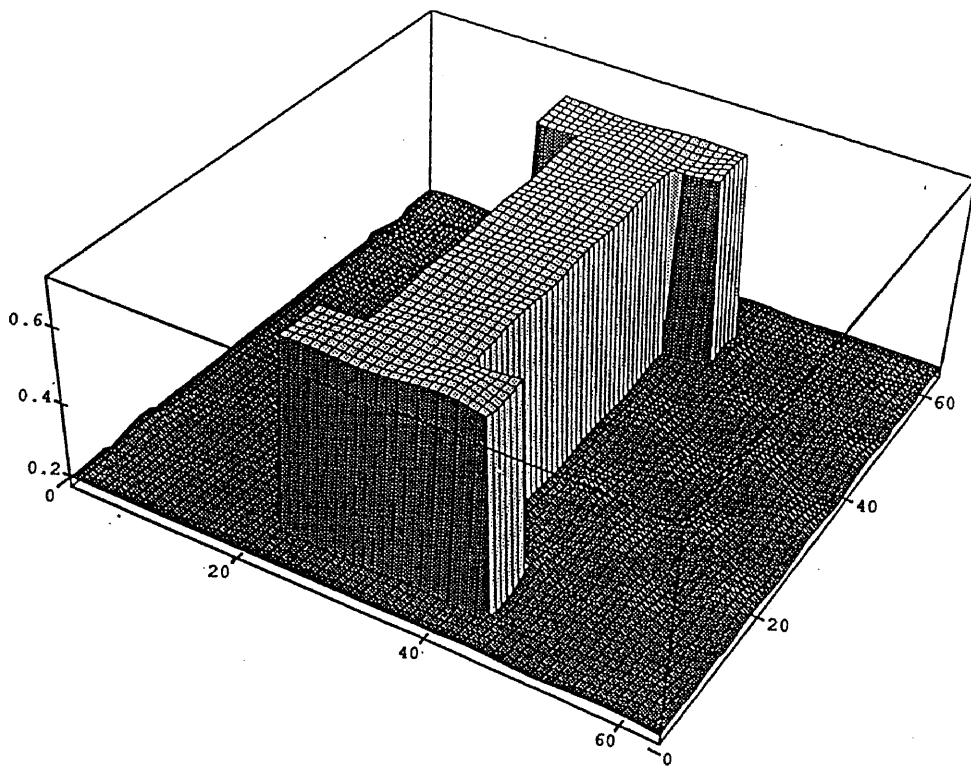
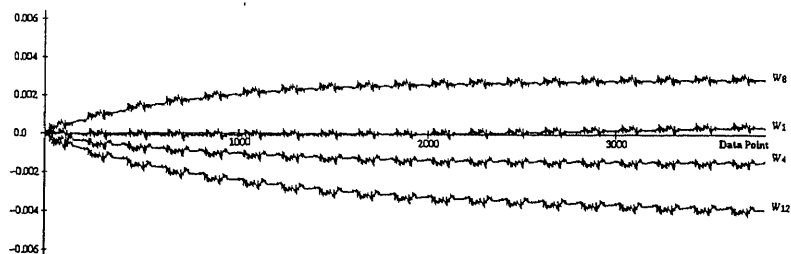
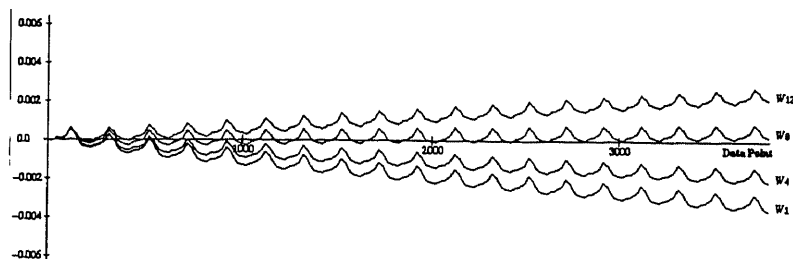


図 2.10: 図 2.4 を用いた場合のシミュレーション結果 (例 4)



(a) 図 2.4を用いた場合



(b) 図 2.1を用いた場合

図 2.11: 係数ベクトル (w_1, w_4, w_8, w_{12}) の収束状況 (例 1)

表 2.1: 雑音除去の結果

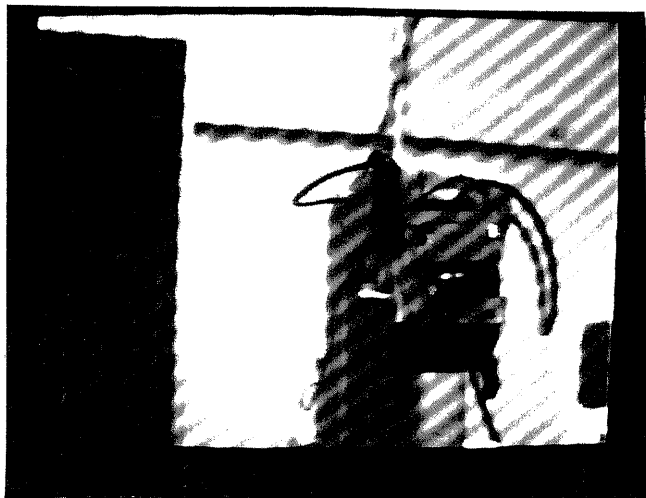
Example	Amplitude	Proposed method (Fig.2.4)	Conventional method (Fig.2.1)
1	Amp.=0.02 ($\sigma_i=0.028$)	$\sigma_o=0.0045$	$\sigma_o=0.052$
	Amp.=0.04 ($\sigma_i=0.055$)	$\sigma_o=0.0091$	$\sigma_o=0.071$
2	Amp.=0.02 ($\sigma_i=0.025$)	$\sigma_o=0.010$	$\sigma_o=0.13$
	Amp.=0.04 ($\sigma_i=0.050$)	$\sigma_o=0.0059$	$\sigma_o=0.14$
3	Amp.=0.02 ($\sigma_i=0.024$)	$\sigma_o=0.0081$	$\sigma_o=0.080$
	Amp.=0.04 ($\sigma_i=0.048$)	$\sigma_o=0.011$	$\sigma_o=0.091$
4	Amp.=0.02 ($\sigma_i=0.039$)	$\sigma_o=0.012$	$\sigma_o=0.60$
	Amp.=0.04 ($\sigma_i=0.078$)	$\sigma_o=0.014$	$\sigma_o=0.61$

周波数の 13 分の 1 の正弦波雑音を加わったもの) に対して雑音の除去を行なった。

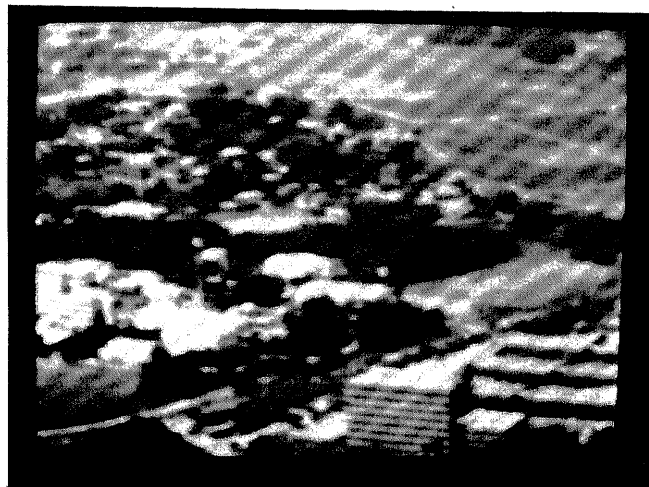
例 1 : 背景に物体が置かれた場合のように急峻な信号の変化が存在するような画像信号 (図 2.12(a))

例 2 : 風景のように信号の変化が繰り返して存在するような画像信号 (図 2.12(b))

その結果を図 2.13, 図 2.14 に示す。シミュレーションの場合と同じ理由によって、従来の 1 チャネル適応雑音除去システムでは図 2.13(a) のようにゴースト状の雑音を生じるが、 ϵ -分離非線形デジタルフィルタを組み込んだ雑音除去システムではそのような症状はなく、雑音が除去されていることがわかる。

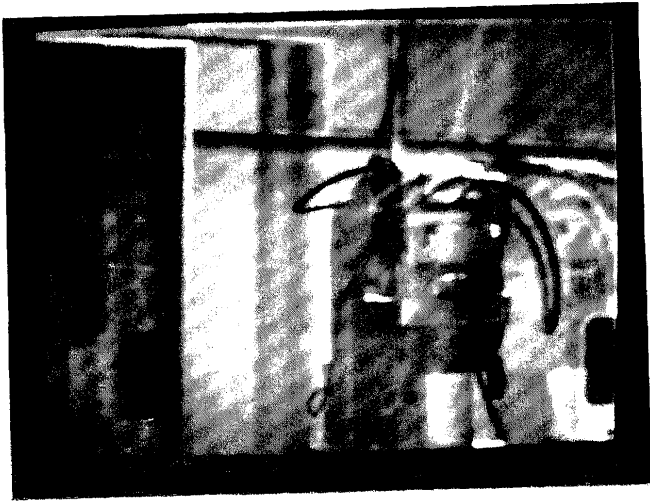


(a) 例 1

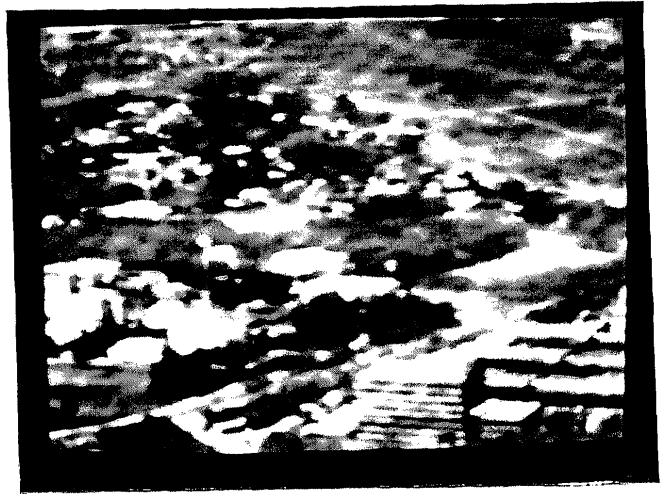


(b) 例 2

図 2.12: 振幅 0.04 の正弦波状雑音を含んだ信号

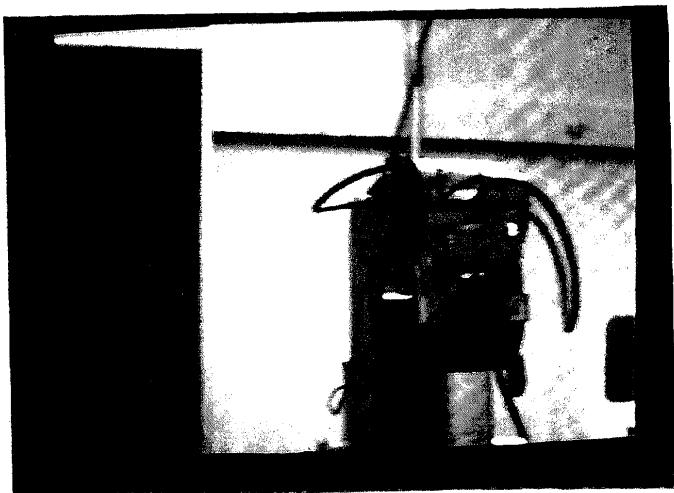


(a) 例 1



(b) 例 2

図 2.13: 図 2.1 を用いた場合の実験結果



(a) 例 1



(b) 例 2

図 2.14: 図 2.4 を用いた場合の実験結果

第 3 章

結論

本論文では従来の 1 チャネル適応雑音除去システムに ϵ -分離非線形デジタルフィルタを組み込んだ雑音除去システムを提案した。そのことによって、従来の 1 チャネル適応雑音除去システム (図 2.1) ではエッジなどのような信号が急峻な変化をする場合には有効に小振幅の周期性雑音が除去できずに、かえって図 2.13(a) のようにゴースト状の雑音を生ずるという問題点を解決することができた。これは、 ϵ -分離非線形デジタルフィルタを組み込むことにより従来の方法で問題となった信号中の大振幅成分をうまく分離することができるからである。この方法は、適応ゴーストキャンセラなどにすでに用いられている適応デジタルフィルタに ϵ -分離非線形デジタルフィルタを加えただけであるので実時間処理が可能であると考えられる。本研究で用いた適応フィルタのタップ数、ステップサイズパラメータの値、遅延数や ϵ -分離非線形デジタルフィルタのタップ数、 ϵ_0 の値などは経験的に値を選んだものであり、これらの値の決定法はこれからの課題である。

付録 A

複素スペクトルを用いた周波数補正法

A.1 連続フーリエ変換の位相特性

複素スペクトルから周波数成分を精度よく抽出することを目的にフーリエ変換の位相特性について考える。今、周波数 f 、観測区間 T 、区間中央での位相 ψ なる余弦波を $x(t)$ とすると

$$\begin{aligned}x(t) &= A \cos\left\{\frac{2\pi f(t - T/2)}{T} + \psi\right\} \\&= \frac{A}{2}\left[\exp\left\{j\left(\frac{2\pi f(t - T/2)}{T} - \pi f + \psi\right)\right\}\right. \\&\quad \left. + \exp\left\{-j\left(\frac{2\pi f(t - T/2)}{T} - \pi f + \psi\right)\right\}\right] \quad (\text{A.1})\end{aligned}$$

と表される。式 (A.1) をフーリエ変換すると $\exp\{j(\cdot)\}$ と $\exp\{-j(\cdot)\}$ の二つの成分に応じて、二つのスペクトルピークが生じる。実波形 $x(t)$ はこの二つの成分の逆変換の重ね合わせで表現することができるので、ここでは単一周波数として $\exp\{j(\cdot)\}$ の成分だけを考えることにする。

$$y(t) = A \exp\left\{j\left(\frac{2\pi ft}{T} - \pi f + \psi\right)\right\} \quad (\text{A.2})$$

式 (A.2) をフーリエ変換すると、複素スペクトルはつぎのようになる。

$$\begin{aligned}z_m &= \frac{A}{T} \int_0^T \exp\left\{j\left(\frac{2\pi ft}{T} - \pi f + \psi\right)\right\} \cdot \exp\left\{-j\frac{2\pi mt}{T}\right\} dt \\&= \frac{A \sin(\pi f)}{\pi(f - m)} \exp(j\psi) \quad (\text{A.3})\end{aligned}$$

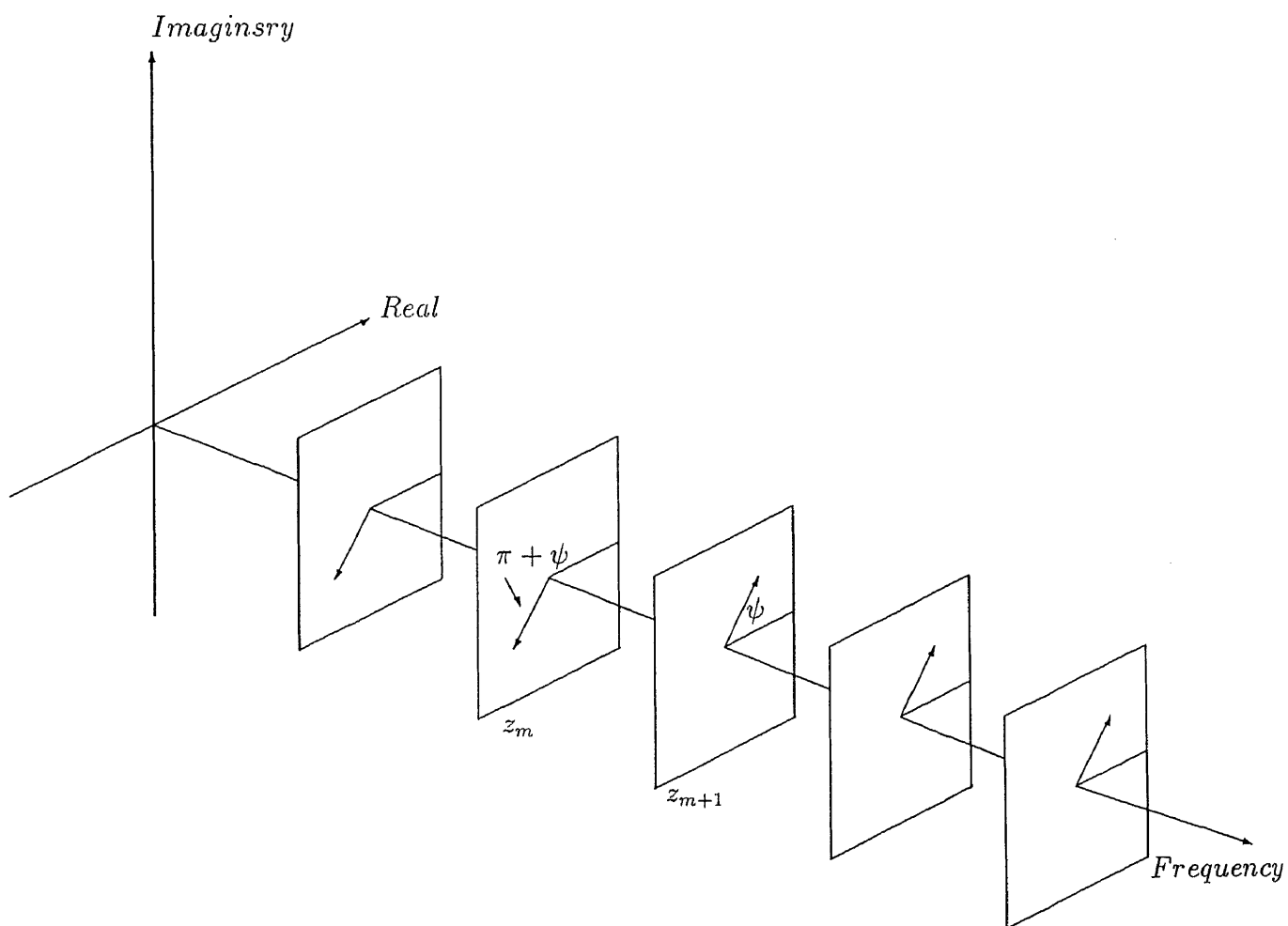


図 A.1: ピーク周波数成分付近のスペクトル成分のベクトル表示

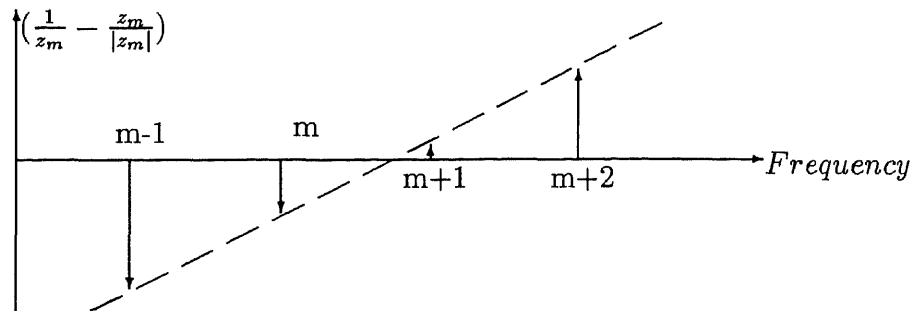


図 A.2: ピーク周波数成分付近のスペクトルの逆数表示

ここで、 $m = [f]$ ($[\cdot]$ はガウス記号) の前後を図示すると、 z_m は図 A.1 のように、方向 ψ の直線上に乗っており、 $m < f$ と $m > f$ に応じて、その向きを反転する。また、 z_m の方向 ψ に対する成分の逆数 ω_m はつぎのようになる。

$$\omega_m = \frac{\pi(f - m)}{A \sin(\pi f)} \quad (\text{A.4})$$

これは図 A.2 のように、 m に関する 1 次関数となる。この直線と横軸との交点から周波数 f を求めることができる。以上の式 (A.3)、(A.4) から、単一周波数から求められた複素スペクトルにはつぎのような性質があることがわかる。

- 1) すべての成分のベクトルの方向は、等しく、位相 ψ で決まる。
- 2) ベクトルの ψ 方向の成分の逆数は、 m についての 1 次関数で表される。
- 3) その 1 次関数の直線が周波数軸と交わる点はその周波数を与える。

A.2 離散フーリエ変換の位相特性

次に、前節の三つの性質を利用して、周期波形に含まれる成分の周波数を求める方法について考える。一般に周波数分析として用いられるのは高速フーリエ変換である。そこで、連続波形 $x(t)$ の積分によって求められた式 (A.3)、(A.4) を離散データについて検討しておく。式 (A.2) を離散データで表すとつぎのようになる。

$$y_k = A \exp\left\{j\left(\frac{2\pi f k}{N} - \pi f + \psi\right)\right\} \quad (\text{A.5})$$

この式 (A.5) に離散フーリエ変換を適用するとつぎのようになる。

$$\begin{aligned} z_m &= \frac{1}{N} \sum_{k=0}^{N-1} y_k \exp\left\{-j\frac{2\pi k m}{N}\right\} \\ &= \frac{A}{N} \exp\{j(\psi - \pi f)\} \sum_{k=0}^{N-1} \exp\left\{j\frac{2\pi k(f-m)}{N}\right\} \\ &= \frac{A}{N} \frac{\sin(\pi f)}{\sin \frac{\pi(f-m)}{N}} \exp(j\psi) \exp\left\{-j\frac{\pi(f-m)}{N}\right\} \end{aligned} \quad (\text{A.6})$$

ただし, $m = 0, 1, 2, \dots, N-1$ である. 式 (A.6) から離散フーリエ変換によって得られる複素スペクトル z_m は $\exp\left\{-j\frac{\pi(f-m)}{N}\right\}$ のために, m とともにベクトルの方向が変化する. N が大きい場合は z_m は $\exp\left\{-j\frac{\pi(f-m)}{N}\right\}$ は無視することができ, $\sin \frac{\pi(f-m)}{N}$ はつぎのようになる.

$$\sin \frac{\pi(f-m)}{N} = \frac{\pi(f-m)}{N} \quad (\text{A.7})$$

以上より離散データから求められた z_m は連続データから求められた z_m に等しくなり, 連続データの場合と同様に m に関する 1 次関数と横軸との交点から周波数 f を求めることができる.

A.3 ピーク周波数同定の方法

フーリエ変換により求められた複素スペクトルのピークを形成する隣合う二つの成分を z_m と z_{m+1} とする. この二つのベクトルは, 理論的には前節で述べたように, 方向は ψ で, 向きがちょうど π だけ異なるが, 実際には他の周波数成分, ノイズ, 量子化誤差などの影響で, 図 A.3 に示すように位相の違いはちょうど π にはならない. そこで $u = \frac{z_{m+1} - z_m}{|z_{m+1} - z_m|}$ なる単位ベクトルを定義し, u と z_{m+1} , z_m の内積の逆数との関係は, 図 A.4 のようになり, 次式から周波数 \hat{f} を推定できる.

$$\hat{f} = m + \frac{(u, z_{m+1})}{(u, z_{m+1}) - (u, z_m)} \quad (\text{A.8})$$

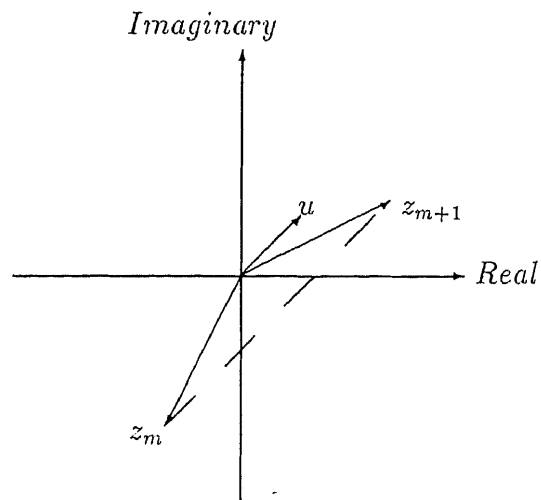


図 A.3: ピーク周波数成分でのスペクトル成分のベクトル表示

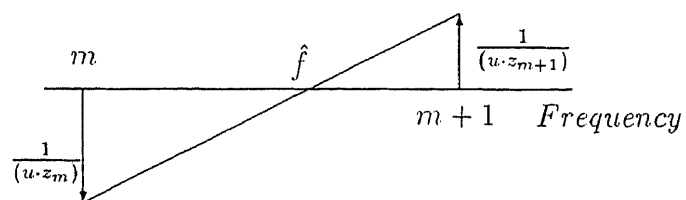


図 A.4: ピーク周波数成分付近のスペクトルの逆数表示によるピーク周波数成分の推定

参考文献

- [1] B.Widrow and S.D.Stearns: *Adaptive Signal Processing*, Prentice Hall Inc.(1985)
- [2] S.Haykin:*Introduction to Adaptive Filter*, Macmillan, New york(1984)
; 式部 幹訳: 適応フィルタ入門, 現代工学社 (1987)
- [3] B.Widrow, J.McCool and M.Bell:The Complex LMS Algorithm, Proc.IEEE, 719/720(1974)
- [4] 浜田晴夫: アダプティブフィルタの基礎(その1)(その2), 日本音響学会誌, 45-8, 624/630 (1989), 45-9, 731/738 (1989)
- [5] 羽鳥光俊: 適応フィルタの最近の動向, 計測と制御, 25-12, 1028/1088 (1989)
- [6] 梅本敏孝: 適応フィルタの応用に関する研究, 筑波大学工学研究科修士論文 (1991)
- [7] 梅本, 青島: 適応フィルタを用いたスペクトル分析法, 第29回 SICE 学術講演会予稿集, 341/342 (1990)
- [8] 梅本, 青島: 適応フィルタを用いた自動採譜, 第30回 SICE 学術講演会予稿集, 155/156(1991)
- [9] 梅本, 青島: 適応技術を用いた採譜に有効な周波数分析, 計測自動制御学会誌論文集, 28-5, 619/625(1992)
- [10] 梅本, 青島: 適応フィルタを用いた周波数分析法におけるステップサイズパラメータの選定, 計測自動制御学会誌論文集, 28-10, 1257/1262(1992)

- [11] 梅本, 青島: 採譜システムの構築と音高抽出精度の比較, 第32回SICE
学術講演会予稿集, 265/266(1993)
- [12] 梅本, 青島: 採譜システムの構築と音高抽出精度の比較, 計測自動制御
学会誌論文集, 29-10, 1227/1231(1993)
- [13] 梅本, 青島: 適応離散フーリエ変換法とその応用, 第36回自動制御連
合講演会予稿集, 387/388(1993)
- [14] 梅本, 青島: 適応離散フーリエ変換法とその応用, 計測自動制御学会論
文集,30-8 掲載予定
- [15] 井口征士: 音楽情報の処理 - 電算機を用いた自動採譜 -, 計測と制御,
音と計測制御小特集, 19-3, 314/319 (1980)
- [16] 原, 井口: 複素スペクトルを用いた周波数同定, 計測自動制御学会論文
集, 19-9, 718/723 (1983)
- [17] 新原, 今井, 井口: 歌唱の自動採譜, 計測自動制御学会論文集, 20-10,
68/73 (1984)
- [18] 片寄, 今井, 井口: 音楽における感情情報抽出の試み, 人工知能学会誌,
3-6, 748/754(1988)
- [19] 片寄, 井口: 知的採譜システム, 人工知能学会誌, 5-1, 59/66(1990)
- [20] 高澤嘉光: 計算機による採譜, 情報処理学会誌, 29-6, 593/598 (1988)
- [21] 岸恒行: 電子楽器の成り立ちとその変遷, トランジスタ技術2月号(1990)
- [22] 井口征士: 採譜と音楽情報処理, bit 別冊, 共立出版(1987)
- [23] 高澤嘉光: 採譜から自動演奏, bit 別冊, 共立出版(1987)
- [24] 小川他: 音楽理論
- [25] 原田, 河原田: 高分解能周波数分析による鼻子音の特徴抽出, 信学技報,
PRL84-10, 39/47(1984)

- [26] 原田, 河原田: 高分解能周波数分析による破裂音の特徴抽出, 信学技報, PRL84-21, 49/58(1984)
- [27] 原田, 河原田: 音声の高分解能周波数分析 $-/mi/$, $/ni/$ の特徴抽出-, 信学論 (A), J69-A, 2, 620/629(1986)
- [28] 梅本, 青島: 非線形適応フィルタを用いた画像信号に加わった周期性雑音の除去, 第 32 回計測自動制御学会講演会予稿集, 777/778(1993)
- [29] 梅本, 青島: 非線形適応フィルタを用いた画像信号に加わった周期性雑音の除去, 計測自動制御学会論文集,30-8 掲載予定
- [30] , 大松: 適応デジタルフィルタを用いた画像修復と信号抽出, 信学論, J73-A, 5, 932/938(1990)
- [31] 土井, 安藤: 画像処理論, 昭晃堂 (株) (1980)
- [32] 原島, 小田島, 宮川: ϵ -分離非線形デジタルフィルタとその応用, 信学論, J65-A, 4, 293/304 (1982)
- [33] 荒川, 原島, 宮川: ϵ -分離非線形デジタルフィルタの統計的解析, 信学論, J66-A, 1, 32/39(1983)

謝辞

本研究を行なう機会を与えて下さり，論文執筆に至るまで厚く御指導御助言を賜りました，筑波大学物理工学系 青島伸治教授に心から深く感謝の意を表します。論文執筆において御指導下さった同大学物理工学系 永井啓之介教授，同大学構造工学系 大田道男教授，佐々木公男助教授に深く感謝致します。また，日頃から御指導，御助力下さった同大学物理工学系 藪野浩司講師にも深く感謝致します。

さらに，同大学入学に際して御指導，御助力下さった防衛大学校 佐々木岱教授，柏木英一教授，高橋信明助教授に深く感謝致します。また，研修の機会を下された防衛庁各位の方々にも深く感謝致します。

最後に，本研究を行なうのにおいて影から支えてくれた妻に感謝する。