

P105 機械音の線形予測係数を基にした  
ニューラルネットワークによる機械の故障診断

三井造船(株) ○橋 弘幸, 林 弘能, 竹田 博

### 1. はじめに

製品検査や設備診断の分野においては、機器の発生する音を基に、検査員の聴覚によってその状態を判定する方法が広く行なわれている。たとえば、製品検査では自動車用エンジンやモーターなどで、組み立て後に試運転を行ない、検査員が発生音を聞いて良不良を判断している。また、設備診断ではプレス機等の生産設備の点検時に、実際に運転して異常音の有無を確認する項目が含まれていることが多い<sup>1)</sup>。

このように、機械が発生する音はその機械の状態を知る重要な手がかりの一つとなるものであり、聴覚による検査、診断のニーズは高い。しかしながら、聴覚に限ったことではないが、人間の官能に頼った判定方法では、以下のようないくつかの問題点がある。

- ・主観的な検査：検査員の主觀に頼っているため、検査員の間のばらつきはもちろんのこと、同じ検査員であってもその時の体調などによって検査ミスが起きる可能性がある。また、検査基準を客観的なデータとして残すことができない。

・検査員の養成、確保：官能による検査では、検査員は経験で技能を高めていくしかなく、熟練した検査員の養成に時間がかかる。

このため、現場では発生する音を統一的、客観的に評価できる自動診断技術が求められている。また、ユーザに対しても、検査結果を数値で表示できることで製品に対する信頼性の向上が期待できる。すでに、FFTによる周波数解析やウェーブレット解析を取り入れた製品はあるが、実用上十分ではない。

そこで、本研究では人間の機能に倣った聴覚式パターンマッチングによる音響診断システムの開発を進めてきた。本稿では、その技術と運用方法、適用事例について説明する。

### 2. 人間の機能に倣った検査・診断技術の開発

人間が聴覚によって物事を判断するまでの過程は、情報を収集する耳の機能と、その情報を判断する脳の機能に分けられる。まず、耳では入ってくる音の情報を時間軸と周波数軸で解析し、音の特徴を顕在化する。そして、脳では顕在化された特徴を過去の情報と照合することで、

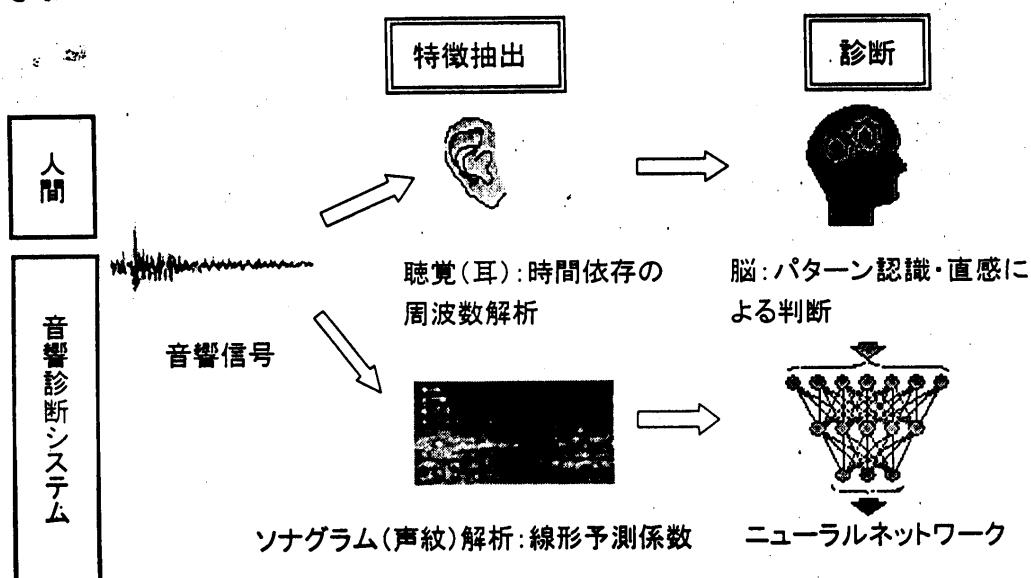


図1：人間の処理機能と音響診断システムの対比

音の特徴をしていると考えられている。

そこで、本研究では人間の耳と脳に相当する処理を、それぞれ線形予測解析、ニューラルネットワークに置き換えることで、人間の判断能力に匹敵する音響診断システムを目指して開発を行なってきた(図1)。以下、本診断システムの核となっているそれぞれの解析方法について説明する。

2.1 線形予測係数 音響(または振動)信号の特徴を解析する最も一般的な方法は、FFTを用いたスペクトル解析であろう。さらに、周波数情報の時間依存性まで調べるために、短時間FFTを利用したスペクトログラム(声紋)解析やウェーブレット解析が用いられる。しかしながら、これらの解析結果をそのまま用いて機器の状態を診断するにはデータ量が多く、後の処理に時間がかかるてしまう。

一方、パラメトリックなスペクトル解析方法として、線形予測法がある。これは、計測された信号の出力源として線形予測フィルタというモデルを適用するものである。信号の特徴は線形予測フィルタの係数(線形予測係数)で表現される。線形予測係数を時系列に並べることで、信号の時間変化も表現できる<sup>2)</sup>。

線形予測フィルタでは、時系列信号  $x(n)$  が、過去の信号列  $x(n-1), \dots, x(n-m)$  によって、次式のように表現できるものとする。

$$x(n) = \sum_{k=1}^m a_k x(n-k) + \varepsilon(n) \quad (1)$$

ここで、 $a_k$  は  $k$  次の線形予測係数、 $\varepsilon(n)$  は予測残差である。そして、予測残差が最小になるように係数  $a_k$  を求める。線形予測係数を決定する方程式は、ユール・ウォーカー方程式と呼ばれており、次式のようである。

$$\begin{pmatrix} r_s(0) & r_s(1) & r_s(2) & \cdots & r_s(p-1) \\ r_s(1) & r_s(0) & r_s(1) & \cdots & r_s(p-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_s(p-1) & r_s(p-2) & r_s(p-3) & \cdots & r_s(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} r_s(1) \\ r_s(2) \\ \vdots \\ r_s(p) \end{pmatrix} \quad (2)$$

ここで、 $r_s(l)$  は時間差  $l$  の自己相関係数である。実際にはユール・ウォーカー方程式を直接解くことはせず、Burg 法と呼ばれる方法で、1 次の

係数から順次漸化的に求めている。

線形予測係数が決定すると、パワースペクトル  $P(j\omega)$  は次式で計算することができる。

$$P(j\omega) = \frac{E_p}{2\pi} \left| \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}} \right|^2 \Big|_{z=e^{j\omega}} \quad (3)$$

ここで、 $E_p$  は予測残差の 2乗和である。つまり、線形予測係数はスペクトル情報を圧縮した量と考えることができる。

なお、スペクトルを精度良く予測するためには式(1)において過去にさかのぼる点数(次数)  $m$  をどのように選ぶかは重要な問題である。一般的には、最適な次数を決定する基準として AIC 基準が広く用いられている。しかし、機器の状態を診断する場合には、正常と異常で線形予測係数に違いが観測されれば十分であるため、本法では正常と異常で違いが出る次数まで計算する。

2.2 ニューラルネットワーク ニューラルネットワーク<sup>3)</sup>とは、脳の神経回路網の働きをコンピュータ上で再現することで、脳の情報処理機構をシミュレートするものである。

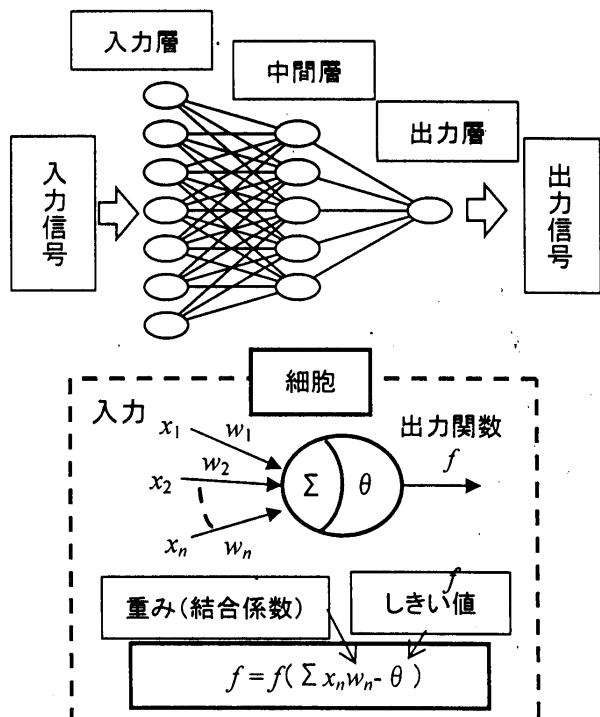


図2：階層型ニューラルネットワークの構造

まず、神経回路網を構成するそれぞれの細胞(ニューロン)は、その他のニューロンからの出力を重み(結合係数)が掛かった形で受け取る。そして、入力和と細胞内部のしきい値の差を独立変数とする関数の出力値を他のニューロンに伝える。

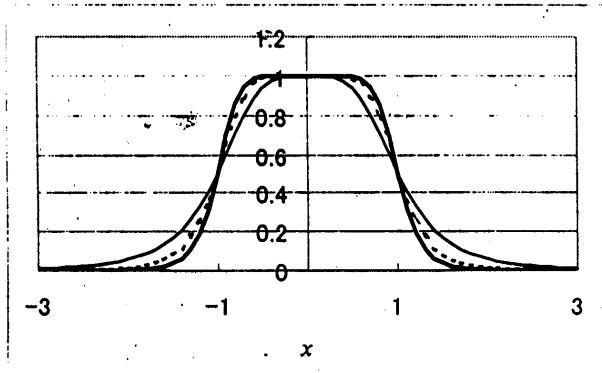
この細胞を階層型に並べたものを階層型ニューラルネットワークと呼ぶ。一般的には外からの信号を受け付ける入力層と、結果を出す出力層、そして中間層が1層の3層階層型が用いられる。

ニューラルネットワークの出力関数には、シグモイド関数を適用する場合が多い。シグモイド関数はしきい値を境に0から1に出力が変化する単調増加関数である。これに対し、本法では出力関数として次式で表されるような限局型関数を用いている<sup>4, 5, 6)</sup>。

$$f(x, \theta) = \frac{1}{1 + c(x + \theta)^n} \quad (3)$$

ここで、 $c$ は定数、 $n$ は正の偶数である。これはしきい値の近傍のみ有限値を持つ関数で(図3)、パターンマッチング用途に対してシグモイド型よりも良好な結果が得られることが分かっている。

図3：限局型出力関数の例。式(3)において、



$c=1, \theta=0$ 。また、実線、点線、太線はそれぞれ $n=4, 6, 8$ を示す。

ニューラルネットワークでは、学習という過程を経ることで結合係数としきい値を調整して望ましい入出力関係を得る必要がある。学習方法には、一般的な誤差逆伝播法(バックプロパゲーション)を用いている。

入力層にあるデータを与えたとき、出力層での教師データとの誤差は、次式で計算される。

$$E = \frac{1}{2} \sum_{k=1}^m (T_k - O_k)^2 \\ = \frac{1}{2} \sum_{k=1}^m \left\{ T_k - \frac{1}{1 + c \left( \sum_{j=1}^l V_{kj} H_j + \theta_k \right)} \right\}^2 \quad (4)$$

ここで、

$T_k$ :教師データ

$O_k$ :出力層の細胞 $k$ の出力値

$V_{kj}$ :中間層の細胞 $j$ と出力層の細胞 $k$ の結合係数

$H_j$ :中間層の細胞 $j$ の出力値

$\theta_k$ :出力層の細胞 $k$ のしきい値

$l$ :中間層の細胞数

$m$ :出力層の細胞数

である。

学習によって誤差を最小にする結合係数は、次式で得られる。

$$\frac{\partial E}{\partial V_{kj}} = \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial V_{kj}} \\ = (T_k - O_k) \frac{n c \left( \sum_{j=1}^l V_{kj} H_j + \theta_k \right)^{n-1} H_j}{\left\{ 1 + c \left( \sum_{j=1}^l V_{kj} H_j + \theta_k \right)^n \right\}^2} \\ = \gamma_k (T_k - O_k) O_k (1 - O_k) H_j \quad (5)$$

そこで、最急降下法を適用して、結合係数 $V_{kj}$ の更新値 $\Delta V_{kj}$ を求める。

$$\Delta V_{kj} = -\alpha \left( \frac{\partial E}{\partial V_{kj}} \right) \\ = -\alpha \gamma_k (T_k - O_k) O_k (1 - O_k) H_j \quad (6)$$

ここで、

$$\gamma_k = \frac{n}{\sum_{j=1}^l V_{kj} H_j + \theta_k}$$

$\alpha$ :定数

である。同様に中間層と入力層の結合係数、さらにしきい値についても更新値を計算し、この操作を繰り返して教師データとの誤差がある一定値以内に収まれば、学習を終了したとみなす。

ところで、従来のニューラルネットワークによるパターンマッチングでは、学習に用いるデータを多数準備しなければならない。たとえば、ある機械の正常音と異常音を判定する場合、その機械の正常音と異常音のサンプルを多数用意し、そのサンプルデータのうちいくつかを学習データとしてニューラルネットワークに学習させなければならない。構築したニューラルネットワークが妥当であるかを確認するには、学習データに使用した以外のサンプルデータを検定データとしてニューラルネットワークに入力して、その出力結果を望ましい出力値と比較する必要がある。

しかし、この方法ではあくまで学習したデータ近傍のデータについては正しく判断するが、まったく異なるデータに対してはどのような出力になるのか分からぬ。ユーザとしては、少なくとも学習したカテゴリには入っていないと判断してほしいところである。また、機械の異常音のようなサンプルでは、サンプル自体の収集が難しい場合もある。

そこで、本法ではあるカテゴリに属するサンプルは統計的に似通ったデータになることに着目して、ニューラルネットワークを構築する方法を考案した。カテゴリ境界設定法<sup>4, 5, 6)</sup>と呼んでいるこの方法は、あるカテゴリに属する学習データの統計的性質を利用してカテゴリの境界データを擬似的に作成し、学習データと境界データを同時に学習させる方法である。概念図を図4に示す。具体的には、あるカテゴリに属する学習データの分布が正規分布に従うと仮定して、平均値と標準偏差を利用して擬似的にそのカテゴリの境界データを作成する。そして、学習データであれば1を出力し、境界データであれば0を出力するように同時に学習させる方法である。

この方法であれば、境界データの作成の仕方によってある程度カテゴリの範囲を制御することができ、結果としてニューラルネットワーク構築の効率化と判定精度の向上を図ることができる。

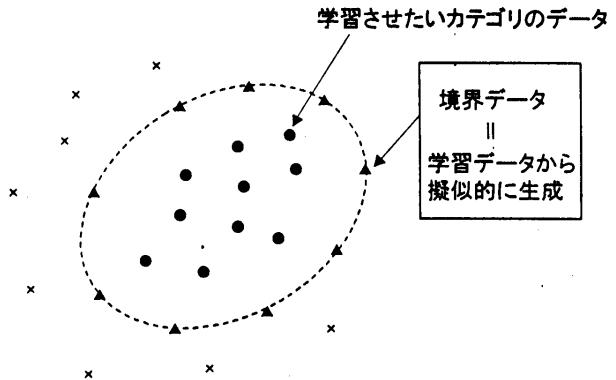


図4：カテゴリ境界設定法の概念図

### 3. 運用

音響診断システムを実際に運用するには、以下のようなステップを踏む。

- 1) データ収集：対象となる機械の運転音を収集する。製品検査に利用する場合は、多数の製品の発生音を収集できる。一方、設備診断の場合は多数の製品の発生音を収集することはできないため、時間を変えて同一機械の発生音を多数収集することでそれに代えることができる。
- 2) 特徴解析：解析ツールを用いて、その機械音の特徴がよく現れるような線形予測係数のパラメータを探る。特徴が現れる周波数帯域が事前に判明している場合は、線形予測解析の前に帯域フィルタを通することで、より特徴を出すことができる。正常音と異常音の判定のようにカテゴリ分けをしたい場合は、同一カテゴリ内では同じような特徴、異なるカテゴリでは異なる特徴が見られるようにすることができる。
- 3) ニューラルネットワーク構築：収集したデータの中からニューラルネットワークに学習させるデータを選び、学習を行なう。学習が終了したら、残りのデータを検定データとしてニューラルネットワークに入力し、所望の出力値が得られているか確認する。そうでない場合は、ニューラルネットワークのパラメータや学習データを代えて学習と検定を繰り返す。
- 4) システムへの組み込み：こうして得られた

線形予測係数、ニューラルネットワークの解析パラメータをシステムに組み込んで、運用を開始することになる。

ニューラルネットワークに学習させるデータの数は多いほど望ましいが、経験上、数十例から百例程度あれば精度よく構築できるようである。また、使用する学習データはカテゴリ内をまんべんなく学習できるよう、できるだけばついているものを使用する方がよい。

#### 4. 適用事例

ここでは、本法を自動車エンジンの擬音に対して適用した例を示す。

(社)自動車技術会発行の擬音調査報告<sup>7)</sup>には、エンジンの擬音として表1に示す17種類が挙げられており、付属CD-ROMには実際の音源も収録されている。この17種類の音を本法で判定することを試みた。

まず、各擬音の音源データから線形予測係数を計算した。その例を図5に示す。擬音の種類が異なると、線形予測係数のパターンも異なっていることが分かる。次に、計算された線形予測係数をニューラルネットワークに学習させて、データの収集、線形予測係数の計算、そしてニューラルネットワークによる判定をリアルタイムで行なうシステムを構築した。

従来のニューラルネットワークと判定精度を比較するため、従来法（シグモイド型関数、カテゴリ境界なし）と本法（限局型関数、カテゴリ境界あり）でニューラルネットワークの判定結果を比較した。その結果、擬音そのものについては両者とも十分な精度で判定することができた。ただし、それぞれの擬音の前に収録されている擬音の説明ナレーションの部分について、従来法のニューラルネットワークでは多数の誤判定（17種類の擬音のどれかと判定する）が起きた。たとえば、ゴロゴロ音、ヒューン音2、ウーン音の説明ナレーションでそれぞれ10回、3回、4回誤判定が起きた。これに対し本法のニューラルネットワークでは、先に挙げた3種類の説明ナレーションに対し誤判定はいずれも0回と、本法によって適切なカテゴリが構築できていることを確認した。

表1：17種類の自動車エンジン擬音

番号	和名	現象
1	カタカタ音	ピストンスラップ音
2	ガ一音	燃焼音
3	ゴロゴロ音／ガ一音	往復慣性力、吸気音
4	ゴロ音	加速時パワープラント騒音
5	ゴロゴロ音	往復慣性力、吸気音
6	モ一音	車体内こもり音
7	割れ音	往復慣性力
8	ボ一音	燃焼、吸気音管内共鳴
9	ウォンウォン音	タイミングベルト振動
10	ヒューン音1	チーンノイズ
11	ヒューン音2	チーンノイズ
12	ガラ音／ジャラ音	タイミングギア歯打ち
13	ピチピチ音	インジェクター部
14	ウ一音	オイルポンプ噛合
15	ブーン音／ボーン音	電動ファン
16	ウーン音	フューエルポンプ
17	ヒューン音	ターボチャージャー

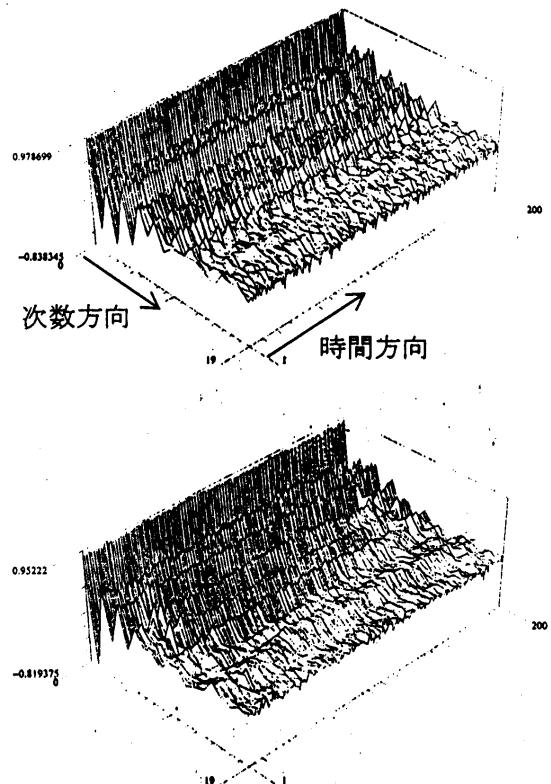


図5：自動車エンジン擬音の線形予測係数の例

(上) ヒューン音1, (下) ヒューン音2

また、実際に自動車エンジン、モータ等の製品検査に関して、導入前の試験結果では高精度に正常品と異常品の判定が行なえることを確認している。

## 5. おわりに

機械の自動診断技術として、機械音の線形予測係数とニューラルネットワークを組み合わせた音響診断方法を開発した。音響診断の分野では特徴解析にFFTやウェーブレット解析を適用した例が多いが、本法は線形予測係数を適用することで周波数成分の時間変化を効率よく解析できる特長がある。また、ニューラルネットワークについても独自の改良を行なうことで判定精度を高めている。

これまで主に製品検査分野を対象に開発を行なってきたが、今後は設備診断分野を対象として技術開発を行なう予定である。設備診断では同じ機械を対象に長期間モニタリングを行なうという用途になるが、ニューラルネットワークの出力値をモニタリングすることにより、機械の発生音で劣化度、さらには余寿命が診断できるシステムの開発を目指す。

## 参考文献

- 1) 労働省安全衛生部安全課編、「動力プレス機械特定自主検査マニュアル」中央労働災害防止協会、1999.
- 2) 森下、小畠、「信号処理」、第6章、計測自動制御学会、1994.
- 3) 中野、「ニューロコンピュータ」、p.38、技術評論社、1989.
- 4) 辻、佐藤、竹田、SICE'94、p467.
- 5) 竹田、佐藤、辻、第38回自動制御連合講演会（平成7年）、講演番号3010.
- 6) 橘、竹田、坪井、日本機械学会2001年度年次大会、講演番号F-1132.
- 7) 社団法人自動車技術会、「くるまの擬音報告—乗用車騒音を主体として—」、1992.