

# 人間親和性を持つ音響反応装置の研究 I

## Development of Audio-response Machine having Human-friendliness I

松原邦彦

Kunihiko Matsubara

### 1. 研究の目的

本研究は人間が聴覚において持つ感応や感性と音響パラメータとの関係を調べ、その結果に基づき人間親和性を持つ音響反応装置を実現しようとするものである。

ここで言う感応／感性とは音によって快い感じや、不快感を持ったり、また警戒感、注意心を生起したり、やすらぎ感を持ったりすることをいう。これら聴覚上の感応／感性を支配するパラメータを音響信号から分析、抽出し、人間の発する警戒的な叫び声や、機械の故障時に発生する摩擦音、破壊音などによって非常事態を判断することのできる音響反応装置を実現することを目標としている。

人間親和性とは「より人間に近い反応」を意味しており、音響による判断において存在するあいまいな状態に対してもなんらかの反応を示し、判断が確定しなければそれなりに情報をあたえる装置を目指している。従来の工学では感性について取り扱うことはほとんど無かった。高精度を目指す技術を追求してきた技術ではあいまいなものを持ち込むことを避けるのが賢明であったし、またあいまいさを持たないことが機械的な美しさだとされていた。ところが高精度の追求は逆に機械の多機能性、多様性への対応を要求し、最近では積極的にあいまいなものを取り込む方向へと転じてきた。心と感性の工学<sup>1)</sup>と呼ばれる分野が急速に開けつつある。

従来、音声による人間-機械間のインターフェースの研究はいわゆる「音声認識」として各国の研究機関において巨額の投資が行なわれ、長年にわたり数多く研究されてきた。これらは音声と語

彙を結びつける技術であり、音声によって伝達される感性を取り扱うものではなかった。自然言語の曖昧性のほかに、音声パターンの認識エラーや曖昧性が避けられないため、音声による計算機械との対話は円滑に進行するには至っていない<sup>2)</sup>。現在は特定話者認識のみ実用になっており、広く応用するには一定の困難が存在する。

この困難さは発音の多様性や、人間の言語中枢神経網の複雑な認識過程を従来の機械的処理で簡単に代替え出来ないことを示している。脳生理学が明らかにしたところによれば、脳には音刺激のみでなく、光刺激または皮膚刺激などの各種刺激を連合する「連合領」および「運動領」のあることが明らかにされている<sup>3)</sup>。従って音に対する動物の反応は音響情報だけでは決まらない。さらに、人間の脳における音声認識では、情報信号の概念形成、コード化、学習と呼ばれる高度な機能と結びついている。これら人間の言語中枢機能は十数億年という生物の進化発達の過程を経て形成されていると考えられ、複雑なものである。したがっていくつかのステージを経て、発達せられるべきものと考えられる。

筆者は、人間の聴覚と結びついた感応は本来的に生体防衛の機能として発達したものと考えた。事件を他の仲間に伝達するため、先ず擬態語が発達し、さらに情報伝達のため発声コードとして発達し、認識するための概念形成が行なわれ、言語として発達したと考えられた。生物が発達の過程で最初に獲得した聴覚上の感応は「危険察知」のためのものであったであろう。言語としての理解能力をコンピュータに背負わす前に、まず聴覚から受けるいくつかの感応や感性との関係を研究することが、人間親和性をもつ音響反応装置の第一

歩であると考えている。

音声や音楽、またある種の機械的音響は人間に何らかの生理的反応を引き起こす。著者は過去に主として音曲信号の持つ周波数スペクトル分析にもとづき、この反応を支配するパラメータを分析抽出する方法を研究してきた<sup>4)</sup>。そこでは、特殊フィルターを設計し、その生理的反応または感性与結びつける回路に、ファジイ集合による判定を研究してきた。しかしながら、産業上の応用、特に機械安全あるいは作業ロボット等での音響による危険状態察知など、人間の感応／感性が捉える状況を機械によってどの程度認識できるかの研究はまだ十分ではなかった。そこで、本研究では、様々な音響パラメータと危険察知、すなわち警戒感に結びついた感応との対応を分析することを試みている。

## 2. 研究の内容

本研究は次の4ステップより成る。

### ステップ1：

音響スペクトル分析システムの高精度化整備を行う。各種音響データを採取する。人間が異常状態と察知する様々な破壊音、摩擦音および異常状態を察知したときの叫び声についての採録とスペクトル分析を行なう。

### ステップ2：

スペクトル強度の時間的変化と生理的反応(感応)または感性との関係を研究する。

### ステップ3：

上記の研究の結果より得られた関係を、ファジイ理論にもとづいて表現するとともに、その感性評価方法を組み立てる。

### ステップ4：

これらの結果にもとづき機械安全、特に作業ロボットにおいて音響による危険状態、異常状態への察知への応用形態を研究する。

今回の報告はステップ2までを取り扱う。現在はステップ3の研究が進行中であり、シリーズIIとして次回に報告する。ステップ4は今後のテーマである。

## 2.1 音響スペクトル分析システムの開発

聴覚の感応／感性のための音響スペクトル分析システムは次の構成要素からなる。

(1) データレコーダ：音響サンプルの収録と編集に用いる。

Digital Audio Tape-corder (DAT) 周波数特性 20Hz-20KHz,  $\pm 1$ dB SN比 87dB

4-Track Recorder/Reproducer 周波数特性 40Hz-22KHz,  $\pm 3$ dB SN比 60dB

(2) 入力アンプ：データレコーダーからの出力をミックスしたり適切な振幅に調整、波形観察や視聴のための信号出力などのインターフェース。

(3) オシロスコープ：音響信号の観察用(20MHz, 2ch)。

(4) A/Dコンバーター：レコーダーからのアナログ信号を指定の周波数でサンプリングし時系列データとして、コンピュータに送る。

A/D変換ボード分解能 12bit 変換速度  $10\mu$  sec 逐次変換方式

(5) パーソナルコンピュータ：取り込んだ時系列データを記憶し、周波数スペクトル分析、ファジイ集合論による認識過程、その他の計算処理を行なう。

Memory 4MB, 数値演算コプロセッサ使用, PC-9801RA2

システム構成のブロック図を図1に示す。写真1はデジタルオーディオテープコーダ(DAT)およびアンプ、写真2はデータレコーダー、オシロスコープ等である。

上記分析システムのためのプログラム開発を行なった。それらは次の4種類である。

(1) A/D変換器制御およびデータ取り込みプログラムコード(ADCONVR)

サンプリング数、サンプリング周波数等のデータ採取条件の設定、採取データのファイル保存等を行なう。最高サンプリング周波数30KHzまで。このコードのフローチャートを図2 a、2 bに示す。図はP A D (Program Analysis Diagram) で示してある。

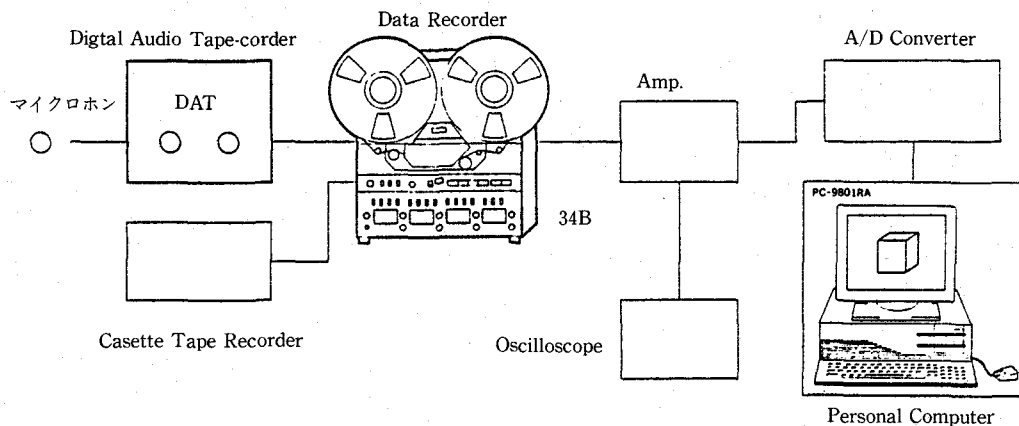


図1 音響収録分析システム構成



写真1 音響収録分析システムの一部 a



写真2 音響収録分析システムの一部 b

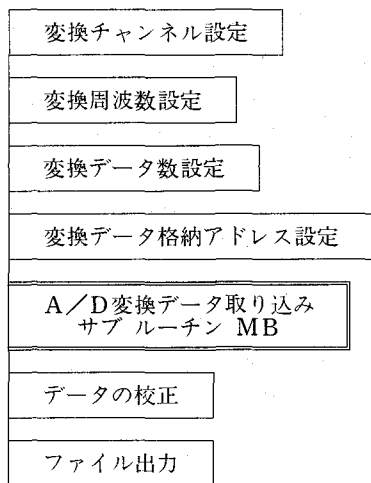


図2 a A/D変換プログラムの流れ (PAD)

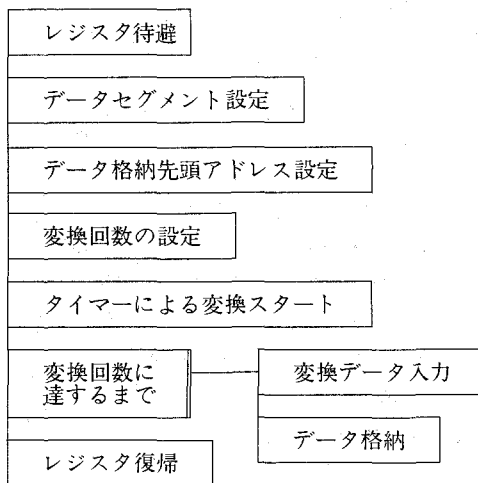


図2 b サブルーチンMBの流れ (PAD)

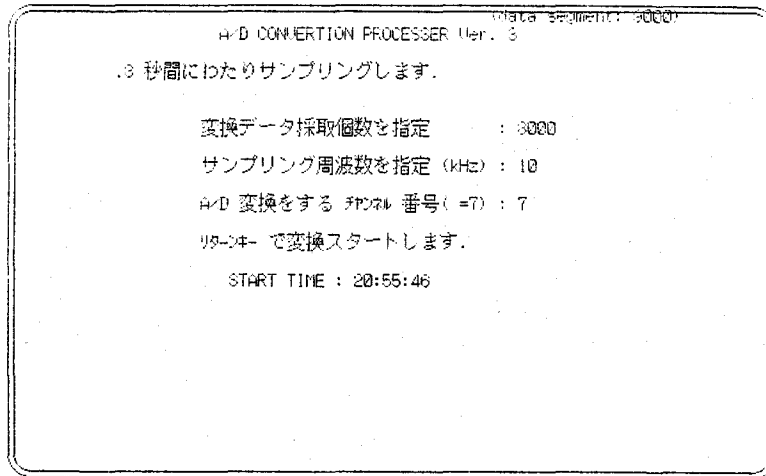
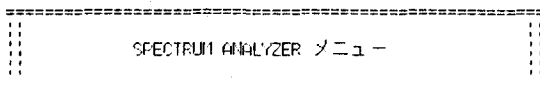


図3 A/D変換器制御プログラム ADCONVR の画面

変換データの取り込み部分はアセンブラで組み、機械語になおしてから BASIC98/Pro により組み込みを行ない、高速に処理する。(図3)

(2) 高速フーリエ変換、パワースペクトル計算プログラム (AUDION)

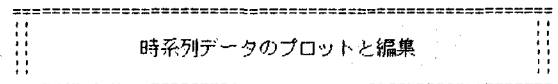
時系列データの任意の区間を切り出して編集したり、FFT演算、パワースペクトル演算、グラフ表示を行なう。プログラムは FORTRAN を用い、グラフィックライブラリは基本描画部分で C 言語で作成し、特別のインターフェースによって結合し FORTRAN から参照出来るように構成した。それぞれの出力データはファイルに落すことができ、後続の処理の入力として使用される。図4a、4bにメニューの一部を示す。



- (1) : 時系列データのプロットと編集
- (2) : フーリエスペクトルの計算
- (3) : パワースペクトルの計算
- (4) : スペクトル密度のプロット
- (5) : 終了

SELECT and ENTER job No. =

図4a AUDION のメインメニュー



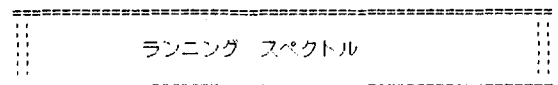
- (1) : 全時系列範囲をプロット
- (2) : 任意の時系列区間をプロット
- (3) : 任意の時系列区間を切り出す
- (4) : 終了

SELECT and ENTER job No. =

図4b AUDION のサブメニュー (1)

(3) ランニングスペクトル計算プログラム (AUDION 2)

時系列データを 10msec から数10msec 程度のフレームに分割し、各フレームのスペクトルを計算、各周波数成分の大きさを 8 階調のカラーによって表示する。この分析コードによってスペクトルの時間的変化の様子を観察することができる。図5にメインメニューを示す。



- (1) : 時系列データのフレーム分割
- (2) : フーリエスペクトルの計算
- (3) : パワースペクトルの計算
- (4) : スペクトル密度のプロット
- (5) : 終了

SELECT and ENTER job No. =

図5 AUDION 2 のメインメニュー

(4) 音響のパターン認識プログラム(AUDION3) スペクトルパターンの判定にファジイ推論を用いる。このためファジイパターンの組立ておよびファジイ推論を行なう。ランニングスペクトル強度の時間変化を求め、フレーム時間軸上のパターンを推論する。更にスペクトルパターンとフレーム時間パターンの両者の推定結果を合成し、最終的な推論に乗せるまでの処理を行なう。図6にこのプログラムのメニューを示す。この部分の処理の原理は本報告のそのIIで詳細に述べる。

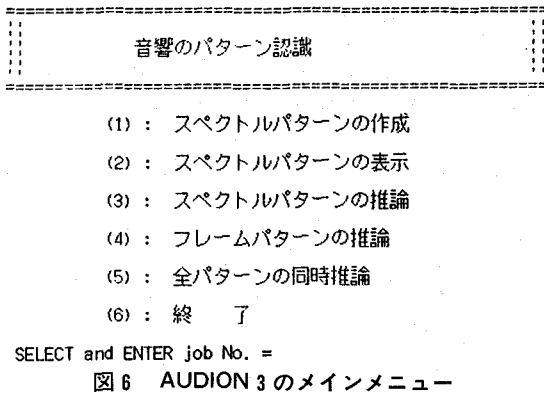


図6 AUDION 3のメインメニュー

## 2.2 聴覚における感応および感性

1に述べたように、聴覚から受けるいくつかの感応や感性との関係を研究するが、まず聴覚における感応及び感性の言葉を本研究の便宜上次のように使い分ける。

### 聴覚における感応：

聴覚による刺激から誘発される生体的防衛的反応で、高等動物に共通して見られる反応。

### 聴覚における感性：

聴覚による刺激から誘発される情緒的感情を意味し、記憶や理性と相まって発生する人間に固有の精神機能と考えられる反応。

これらについて分類したものを表1に掲げる。

表1 聴覚における感応および感性

(A)	闊達 警戒 沈静 恐れ	(B)	喜び 悲しみ 怒り いたわり
感 応	落ち着き いらだち	感 性	勇気 活気

なお、これらの分類にはいろいろな方法が考えられ、また感応と感性の間の境界も判然としたものではない。さらに動物にはここに言う感性を全く持たないとは現状では言いきれない。これらは本研究の便宜上のもので、確立した分類法ではない。(A)感応にかかげるものは人間をはじめ動物にとってその生体防衛上で大切な反応である。特に危険を察知する機能として警戒感、恐れ等の感応が備わっていると考えられる。警戒的な人の叫びや破壊音、破裂音に対して聴覚は非常に敏感である。(B)感性に関するものは、いくつかの記憶、過去の経験、時には理性さえ関係してくると考えられ、聴覚以外の感覚、たとえば目で見たり、触ったり、運動的な動作、その他あらゆる物理的な感覚動作の記憶やプロセスと関連している。したがって感応に関係するものに比べ、複雑な処理を必要とすると考えられる。脳の多くの部分と連合して認識していると考えられるから、人工的な認識装置への応用には更に多くの困難を伴う。

今回は、これらの中から警戒感を主体に取り上げる。これは、生体が身の安全を守るためにいち早く備えた音響による異常検知の手段であると考えられ、産業上の利用価値が高いと考えたからである。

## 2.3 各種音響データの採取

この研究では次の音響データを採取した。音響パターンの特徴を比較するために類似の感性を生起するもの、異なる感応、感性を生起するものも採取した。

### (1) 警戒感を生起するもの

- a) ガラス類の破壊音、金属板/プラスチック類の打撃音
- b) 自動車の警笛(人工的な警戒音)
- c) 人の叫び声：アーツ、キャー、その他(言語的意味を持たないもの)
- d) 人の叫び：「stop!」、「止まれ!」その他(言語的意味を持つもの)

### (2) 落ち着き感(リラックス)を与えるもの

- a) 波の音、小川のせせらぎ
- b) 音曲(子守歌、電子楽器音など)
- c) 柔らかい人の話声、呼び掛け

### (3) 活気を与えるもの

- a) 太鼓の音、行進曲
- b) 元気な人の話声
- (4) いらだちをあたえるもの
  - a) 各種雑音、連続する金属の摩擦音、鋸の音
  - b) 街頭の雑音

- タ分析結果を掲げる。
- a) 破壊／打撃音
  - b) 自動車の警笛
  - c) 人の叫び声1 (言語的意味をもたないもの)
  - d) 人の叫び声2 (言語的意味を持つもの)

2.4 スペクトル分析

警戒感をもつものとして次の4種についてデー

図7-図18に採取データの例を示す。

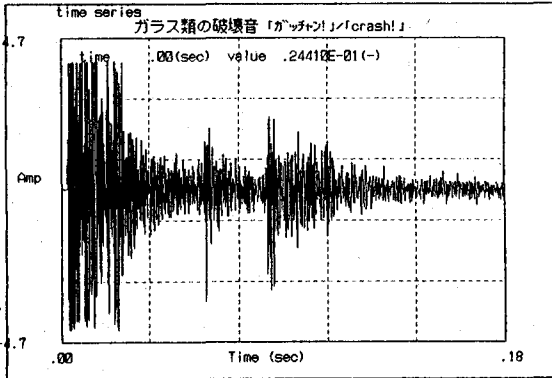


図7 破壊／打撃音時系列データの例

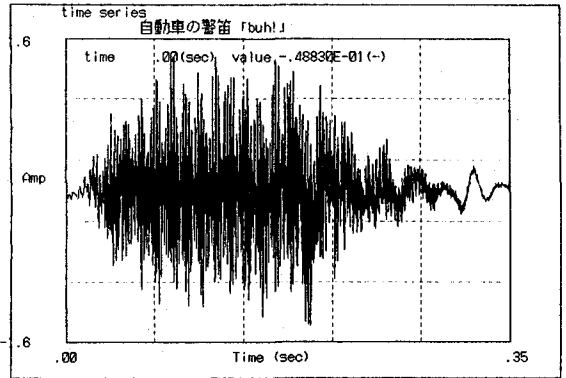


図8 警笛音時系列データの例

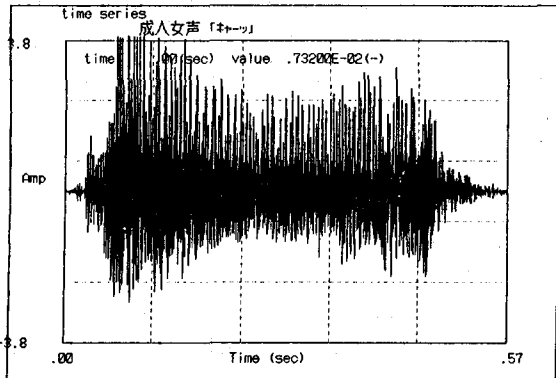


図9 人の叫び声時系列データの例 (言語的意味をもたないもの)

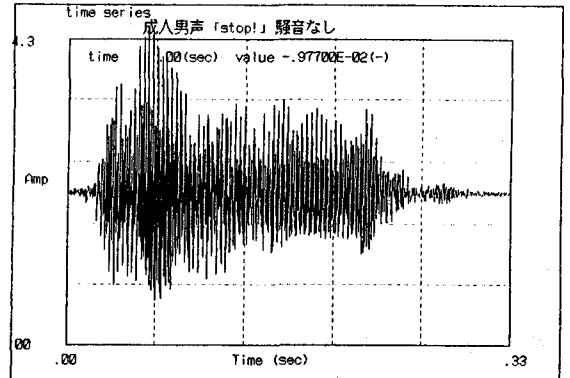


図10 人の叫び声時系列データの例 (言語的意味をもつもの)

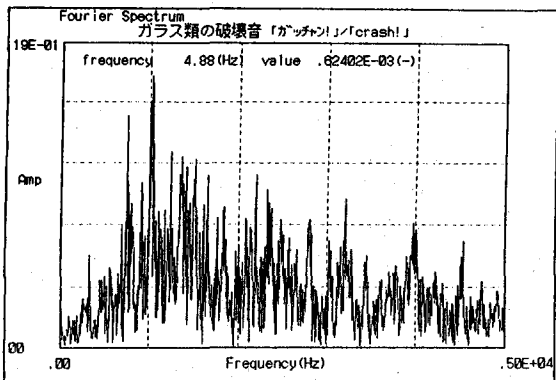


図11 破壊／打撃音スペクトルの例

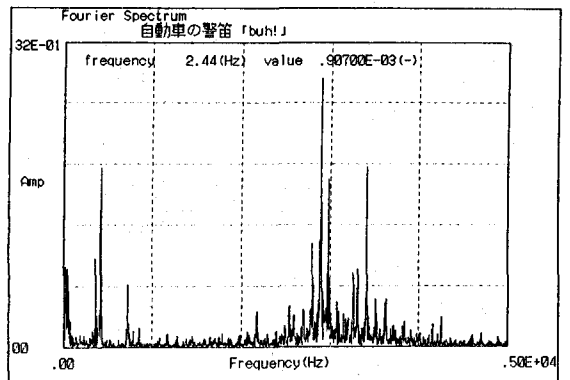


図12 警笛音スペクトルの例

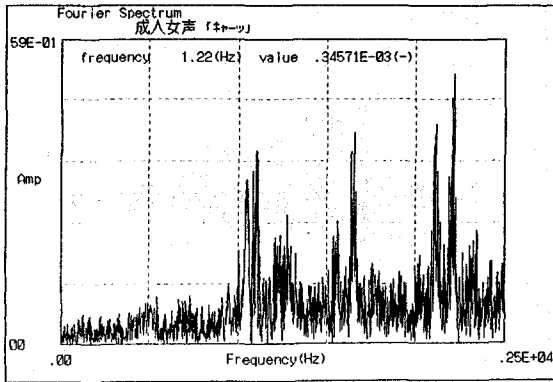


図13 人の叫び声スペクトルの例  
(言語的意味をもたないもの)

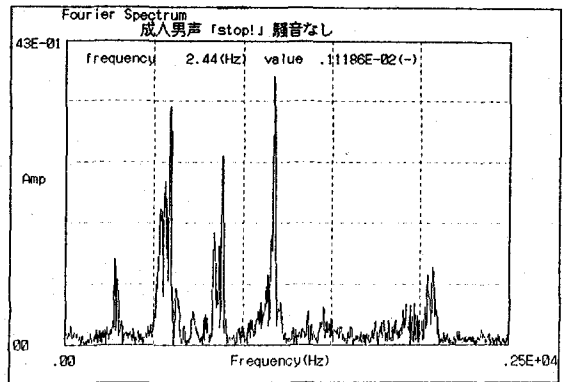


図14 人の叫び声スペクトルの例  
(言語的意味をもつもの)

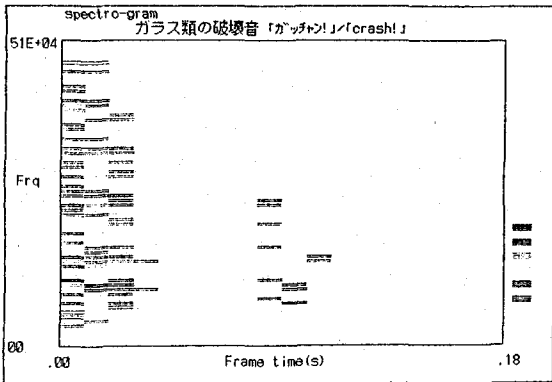


図15 破壊/打撃音ランニングスペクトルの例

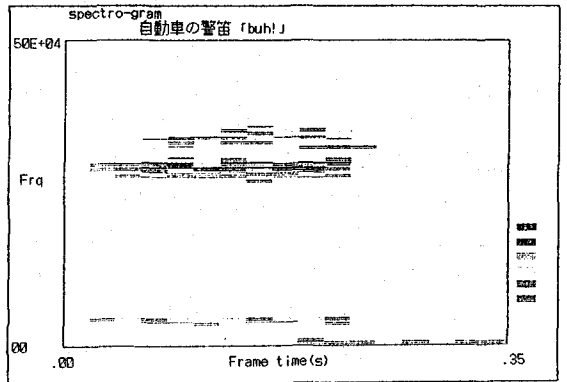


図16 警笛音ランニングスペクトルの例

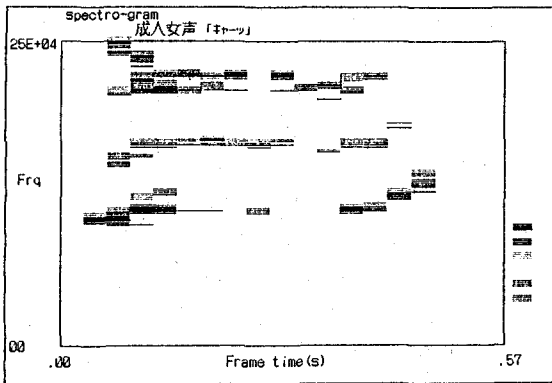


図17 人の叫び声ランニングスペクトルの例  
(言語的意味をもたないもの)

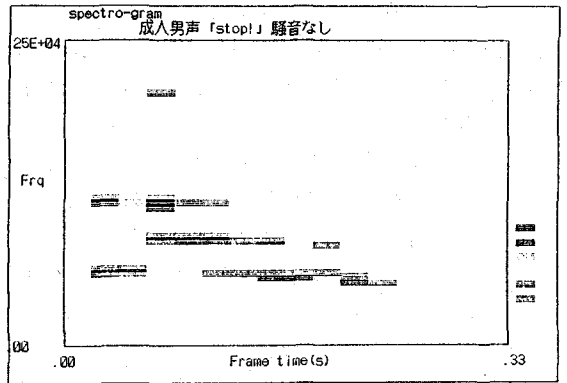


図18 人の叫び声ランニングスペクトルの例  
(言語的意味をもつもの)

これら警戒音と対照な感応を示す2例を比較のために示す。紙面の関係から1例は弦楽器の音響のスペクトル、もう1例は人の柔らかな呼掛けの声の時系列データにとどめる。図19はいわゆる楽音的と呼ばれる音響でスペクトル上に少数のピー

クを持つ特徴を示し、警戒感を生起しない。図20は柔らかな感じをもつ、人のはなしかけ言葉の例であり、やはり警戒感を生起しない。この特徴がその立ち上がり部と終局部におけるなめらかな変化にあることを示す例である。

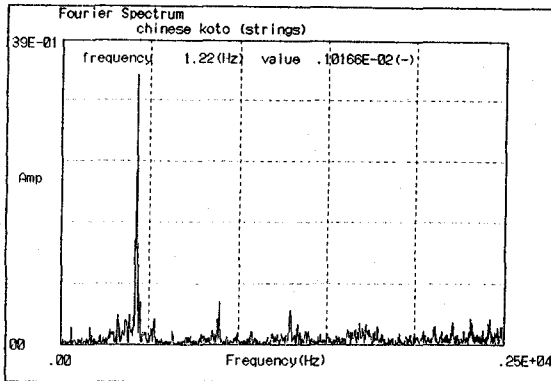


図19 弦楽器（中国古箏）スペクトルの例

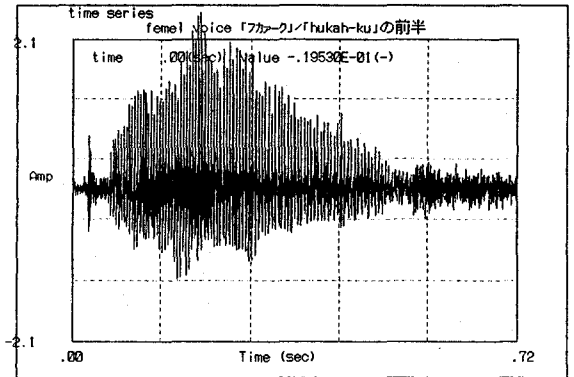


図20 人の柔らかな呼び掛け声の時系列データの例

## 2.5 ランニングスペクトルパターンの特徴抽出

多種類にわたる採取音響データから、ここに警戒感を生起する音響についての分析結果を述べる。ランニングスペクトルは図15-18に示すように、周波数軸と時間軸を持つ2次元空間上に表わされる。音声認識より原始的な感応の取扱いの範囲では、この2つの軸上のパターンは分離できるとの考えに立って分析を進める。データの分析の結果から、次のように規則性をまとめた。

### (1) 警戒音の周波数軸上でのパターンの特徴

- a) 破壊音や打撃音に特徴的なパターン
  - 連続的なスペクトル成分または非常に密に並んだ線スペクトル成分が多い。
  - いくつかの強い線スペクトルの混入がある。
  - 一般に周波数帯が広く、高周波成分まで伸びている。
- b) 自動車の警笛に特徴的なパターン
  - 周波数軸上の限られた部分に非常に密に並んだ線スペクトルの集団を持つ。
  - 一般にこの集団を複数持ち、集団数が多いほど警戒感を強く生起する。
- c) 人の叫び声に特徴的なパターン
  - いずれも比較的少数の離散的傾向を示す線スペクトルを持つ。
  - 言語的意味を持たない悲鳴になると a) のスペクトルパターンに近づく。
- d) 言語的意味のある叫びに特徴的なパターン
  - いずれも少数の離散的傾向の線スペクトルを持つ。
  - 大きなピークは中間領域 (100-1000Hz 程

度)の周波数帯に現われる。

- 線スペクトルの数が少ない程、音曲的になる。
- 声の破裂音、摩擦音の初期に相対的にレベルの小さい a) のパターンが現われる。

### (2) 警戒音のフレーム時間軸上でのパターンの特徴

- a) 破壊音や打撃音に特徴的なパターン
  - ステップ状に立ち上がった後、10-30msec程度の平坦部を経てなだらかに下降する。
  - レベルがバックグラウンド程度に戻る時間は0.1-0.2sec程度と短い。
  - 破壊の場合には、破片の飛散にもとづく2次の音響が発生する事が多い。
- b) 自動車の警笛などの人工的警戒音に特徴的なパターン
  - 立ち上がりに10-30msecを要し、のち一定のレベルを保つ。
  - バックグラウンドレベルに戻る時、立ち上がりと同程度の時間を要する。
- c) 人の叫び声に特徴的なパターン
  - 立ち上がりに10-30msecを要し、なだらかなピークを経て鞍部状に変化する。
  - 終局部のなだらかな盛り上がりの後急激にバックグラウンドレベルへ戻る。
- d) 言語的意味のある叫びに特徴的なパターン
  - c) のパターンにほぼ同じ。ただし、長さはまちまちである。

いずれの場合でも持続時間や繰り返しの回数により、警戒感に対して持つ意味が異なってくる。しかし警戒感を感じるためには少なくとも0.1sec



表2 感応/感性とフレーム時間軸パターン類型との関係

フレーム時間軸上のパターン	感応/感性	対応する叫び/擬態語	対応する音曲/音響
	警戒	ガチャン、ドカン、バン、ボン	破壊、破裂、打撃音
	警告	ブー、ビー、ピー、ワー、ジー、キー	警笛、叫び
	恐れ、怒り	ギャーッ、ウオーッ、カーッ、ウワーッ	悲鳴、叫び
	奮起	カーッ、ターッ、トーッ、ウオーッ	掛け声、吠え声
	沈静	ザー、サー、ネー、フー、シー	波の音
	闊達	(奮起音の繰り返し)	行進曲
	いらだち	(警告音の繰り返し)	街頭の雑音
	落ち着き	(沈静音の繰り返し)	子守歌

程度以上の持続を要する。

ここでわれわれは、これらをなんらかの形で定量化しなければならない。詳細な定量化はシリーズIIで述べることにし、ここではパターンの大まかな分類を試みる。表2は感応/感性とフレーム時間軸パターン類型との関係を整理したものである。

一定のピッチで3回以上続くと、たとえ警戒音であっても危険の意味が薄れ、符号的解釈を行なうようになる。これは符号化能力をもつ人間の特性によるのであろう。また打撃音では、脈拍に近いテンポで適切な強度が繰り返し与えられると、心が落ち着くことが知られている。

### 3. 研究の結果および成果

1で述べたように本研究の目的は、人間が聴覚において持つ感応や感性と音響パターンとの関係を調べ、その結果にもとづいて人間の発する警戒的な叫び声や、機械の故障時に発生する破壊音などを認識することを目標としている。2で述べたように本報告に関わるステップ1および2は、後続のステップ3においてファジイ理論等により感応/感性評価方法を組み立てるため、必要なデー

タ分析結果を用意することである。

まずステップ1では、音響スペクトル分析システムの整備を行なった。これらはランニングスペクトルを分析するためのもので、音響収録装置と直結している。録音部は十分に高測定精度を有するので、全体の測定精度はA/D変換器の分解能(12bit)で決まる。システム全体で高精度を保つ周波数範囲は20Hz-20KHzである。採録したデータは破壊音、打撃音、人の叫びなど警戒音を中心にしない、柔らかな人の話し声、様々な音曲等は警戒音の特徴との比較のために採録した。

ステップ2では、ランニングスペクトルのパターンと感応または感性との関係を研究し、警戒音を主体にその特徴を抽出した。我々はファジイ理論の適用を前提にしており、ファジイ推論<sup>5),6)</sup>に乗る形で結論を求めておく必要がある。ファジイ推論ではいわゆるあいまいさを含む知識をif.. then形式に表現するので、この形で主要な結果をまとめる。

#### 警戒音に共通のパターン上の特徴

##### a) 周波数軸上のパターン

- 「連続的なスペクトル成分を多く持つかまたは非常に密に並んだ線スペクトル成分が多いな

らば、その音響は警戒感が強い。」

- 「線スペクトルの数が少ないならば、その音響は音曲的になる。」
- 「スペクトルパターンが比較的少数の離散的傾向を示すならば、その音響は人の声である可能性が高い。」

ただし、ここでスペクトル上の大きなピークは中間領域(100-1000Hz程度)の周波数帯に現われるものを拾うものとする。

また、スペクトル上の相隣るピーク間の周波数差をdFiとすると、人間の叫び声に特徴的なdFiの帯域は男性成人では300-400Hz、女性、子供では350-500Hz。総じて人の声は300-500Hzの帯域にある。また人の声に特徴的な周波数帯は250-1.8kHzまでである。

b) フレーム時間軸上のパターン

- 「立ち上がり時間が短いならば、その音響は警戒感が強い」
- 「終局時の減衰カーブが急激ならば、それに応じて、警戒感が強い」
- 「音響の持続時間が約100msecより少ないと、警戒感が弱くなる。」
- 「終局時にバックグラウンドレベルに戻る時間が非常に短いならば、警戒的な人の叫び声である可能性が高い」

この研究では感応/感性とフレーム時間軸パターンとの関係を類型付け、表2の様にまとめた。この中から警戒音に特有なパターンとして図21に

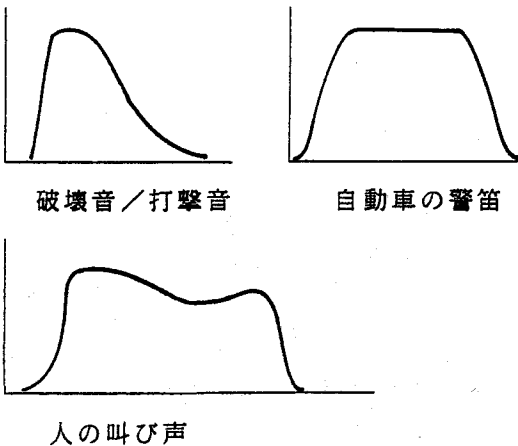


図21 警戒音のフレーム時間パターン

示す3種を挙げることができる。このパターンは時間軸上に固定したものではなく、時間軸上で様々な伸縮がある。

b) に記したように、フレーム時間軸上の変化率に対して人間の聴覚は非常に敏感である。従って、聴覚受容器あるいは神経回路のいずれかは時間変化の微分値を検出していると考えられる。そこで、図21のパターンの時間微分曲線を図22に示す。

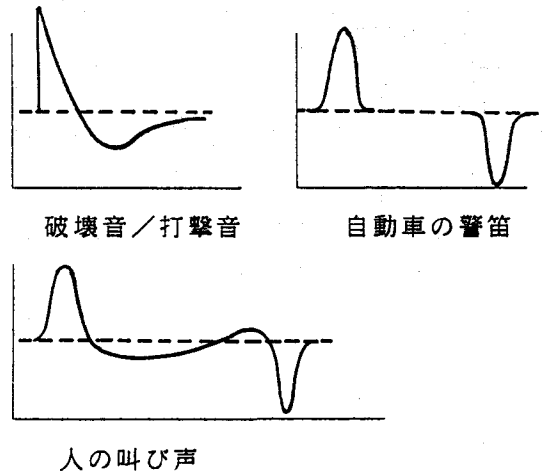


図22 警戒音のフレーム時間パターンの微分曲線

いずれにもプラス側のピークがあり、次にマイナス側の鋭いピークあるいはなだらかなピークがあり、これらが警戒感の生起に強い関係をもっているとみられる。

4. おわりに

本報告では主として、音響の感応/感性とランニングスペクトルパターンとの関係を分析した。以上により、ステップ3のファジイ理論による感応/感性評価方法の研究へと橋渡しされる。

本研究は、(財)長野県テクノハイランド開発機構および(財)浅間テクノポリス開発機構から研究助成を受けて行なったものです。関係各位に深く感謝いたします。

(まつばらくにひこ 教授)  
(1994.4.1 受理)

**参考文献**

- 1) 計測自動制御学会編、ミニ特集 心と感性の工学  
（『計測と制御』Vol.33, No.3, 計測自動制御学会、  
1994年）。
- 2) 竹林洋一「音声理解と対話における自然言語処理」  
（『情報処理』Vol.34, No.10, pp.1287-1296, 情報処理  
学会、1993）。
- 3) 電子通信学会編『聴覚と音声』p.28（電子通信学  
会発行、昭和45年）。
- 4) 松原邦彦(発明者)「音曲反応の方法」  
特許第1771641号（分類G10L 5/00）  
特許出願公告第050600号（平成4年）
- 5) 寺野寿郎、浅居喜代治、菅野道夫共編『ファジイ  
システム入門』p.45（オーム社、1988年）。
- 6) 水本雅晴『ファジイ理論とその応用』p.307(サイ  
エンス社、1989年）。