

## 再構成法を用いた音声分析と認識\*

3 R-6

○齊藤安行 志田晃一郎 藤川英司 山田新一 †  
武蔵工業大学 ‡

## 1. はじめに

音声情報処理の基礎は音声の分析であり、音声の認識と合成がその直接の応用例である。現在の音声処理の方法としては、フィルタ群を用いた周波数別のパワースペクトルからの特徴抽出や、A/D変換器のサンプリングデータから線形予測係数を求めスペクトル解析を行うものが主である。しかし線形分析に基づく音声情報技術が飽和しつつあると指摘する報告も多くなされている今、新しい視点での音声生成の物理現象を再検討することも必要であると考えられる。そこで本研究では音声の生成現象と決定論的カオスの関係という観点から新たな音声分析を考えてみる。

## 2. 母音波形の再構成

生体システムの中で脳波・脈拍など多くのカオス性がわかっているなかで音声もまたカオス性を持つと考えてもおかしくない。事実、母音の揺らぎの時間的推移は全く不規則に揺らいでいるのではなく、何らかの秩序にしたがっていることがわかっている。そこで音声分析に時系列データのカオス性を調べる時に良く用いられる再構成法を用いた。これは元のデータからだけではわからない内包された規則性を見つけるのに有効である。以下にその方法を示す。まず音声信号を1定の時間遅れ $\tau$ ごとの差分を取りそれを遅延座標系、

$$X_t = (x_t, x_{t-\tau}, \dots, x_{t-(d-1)\tau}) \quad (1)$$

( $d$ :埋め込み次元,  $\tau$ :時間遅れ)の各座標に振り分け再構成軌道を描く。つまり1本の時系列データをもとに、仮想的にアトラクタを描くことになる。さらに主成分分析法でアトラクタの構造を変換することで、構造解析は遅延座標系の選び方に依存しなくなる。再構成軌道を主成分空間に変換するために以下の処理を行う。遅延座標系に再構成された母音の軌道の分布は共分散行列:

$$\frac{1}{N-(d-1)\tau} \sum_{t=1+(d-1)\tau}^N X_t^T X_t \quad (2)$$

で表される。共分散行列の固有値を

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_d \geq 0 \quad (3)$$

と大きさの順に並べ、各々に対応する固有ベクトルを

$$u_1, u_2, \dots, u_d \quad (4)$$

と書くとき、以下の変換をする。

$$Y = (y_1, y_2, \dots, y_d) = X[u_1|u_2|\dots|u_d] \quad (5)$$

回転変換 $[u_1|u_2|\dots|u_d]$ により、主成分空間 $Y$ 上で、母音の軌道分布の共分散行列は、分散(軌道変動)の大きい順に $\{u_1, u_2, \dots, u_d\}$ と対角化される。以上の手順で新たな座標空間に再構成した音声信号の軌道が描かれる。今回の研究では単一の母音信号を対象として用いることにし、埋め込み次元 $d$ は、視覚的にも軌道状態がわかるように $d=3$ の3次元空間とした。

\* Analysis and Recognition of Voice by Using Reorganization

† Yasuyuki Saito, Koichiro Shida, Hideji Fujikawa, Shin'ichi Yamada

‡ Musashi Institute of Technology

### 3. 再構成軌道からの特徴抽出

再構成した音声信号の軌道を観測し、その特徴抽出を試みる。その音声信号が何の音か判別できるような特徴抽出である必要がある。よって次のような特徴抽出を考えた。

1) 各主成分軸方向の分散の大きさの比率を求める。主成分空間に変換する事により第1主成分軸方向の分散の大きさが一番大きく、第2、第3主成分軸方向となるごとに分散の大きさは小さいなっていく。そこで分散の大きさの比率を各音声信号の固有の特徴として用いる。

2) 再構成空間をある平面で区切り、その平面を軌道が次々と横断する様子を調べる。(ポワンカレ写像)

3) カオスの予測不可能性の指標であるリアプノフ指数を求める。リアプノフ指数は遅延座標系に再構成された音声の軌道の各固有接ベクトル方向に対する誤差の平均拡大率である。

以上のような特徴抽出を行い、それを各音声信号特有の判別材料とする。

### 4. 母音/a/の再構成

図1は母音/a/の時系列波形である。これを埋め込み次元 $d=3$ 、時間遅れ $\tau=0.4\text{msec}$ で位相空間に再構成したものが図2となる。

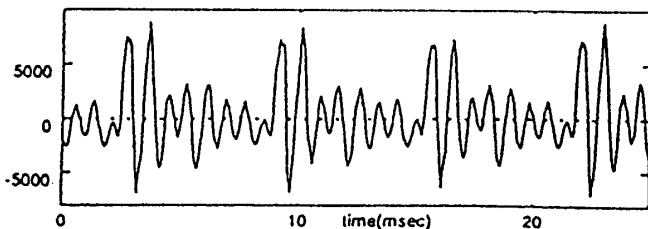


図1 母音信号/a/の時系列波形

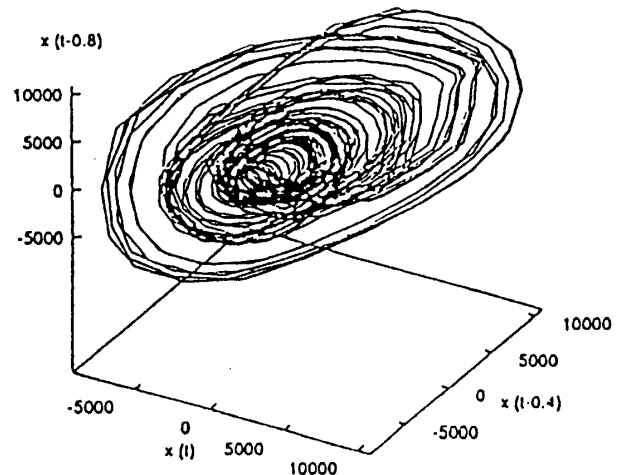


図2 母音信号/a/の再構成軌道

再構成軌道はある決まった領域の中でアトラクタを描いている。各音素ごとにその形状、軌道が占める領域は異なり、それぞれ特有の情報として用いることができると推測できる。

### 5. まとめ

今回の発表では、音声分析の新たな方法として音声信号の位相空間への再構成法を提案した。

これからこの方法により、各音素ごとの特徴抽出を行い、音声認識の判別材料として用いる事にする。そして最終的には子音を含めたレベルまで拡張し、音声認識の上で問題となるセグメンテーションや調音結合について解決策が得られないか考えている。

### 参考文献

- [1] 三木 信弘 : 音声生成過程の音響理論の最近の進歩, 日本音響学会誌, Vol. 48-1, 1992, pp. 15-19
- [2] 合原 一幸 : カオス-応用を目指して-, 数理科学, No. 348, JUNE, 1992, pp. 5-9
- [3] 伊福, 橋場, 松島 : 母音の自然発生における「波形ゆらぎ」の役割, 日本音響学会誌, Vol. 47-12, 1991, pp. 903-910
- [4] 徳田 功 : 音声とカオス, 数理科学, No. 381, MARCH, 1995, pp. 53-59