

## 研究テーマ：LPC ケプストラムを用いた母音音声の認識

73 番 亀田 康介

## 1、研究内容

本研究では、高度な技術である連続音声認識システムの作成を行う上で、基礎的な知識となる、簡単な音声認識処理として、/a/、/i/、/u/、/e/、/o/という日本語の母音をコンピュータによる計算によって認識するシステムの作成を行う。認識を行う際の特徴量としては、LPC ケプストラムを用いる。

## 2、システム概要

LPC ケプストラムとは、LPC 分析で求めたスペクトルより求めたケプストラム係数のことである。認識は、あらかじめ5つの母音音声について LPC ケプストラムを求めておき、各母音について認識対象の音声の LPC ケプストラムとのユークリッド距離を求め、その値が最小になる母音を認識結果とすることで行う。

## 3、実験方法と結果

実験は、まず LPC ケプストラムの次数  $m$  を何次まで認識に使用するかを決定するシミュレーションを行い、その後、本人の声、同性の声、異性の声、他人の声での認識精度について調べる。

音声は、男性 5 名(m0～m4)、女性 5 名(f0～f4)について、一人あたり aueo という音声を 5 回ずつ録音し、それを /a//i//u//e//o/ のそれぞれの部分に切り出したものを用いる。切り出しは各母音の時間波形が定常なところで行う。

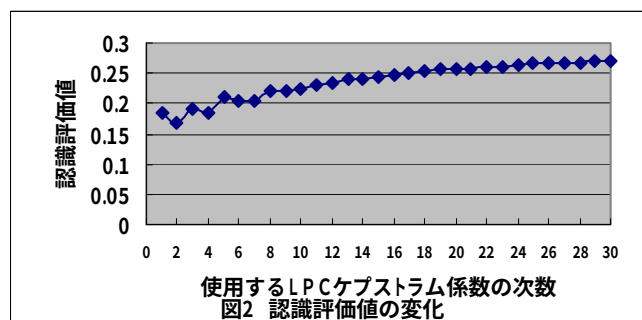
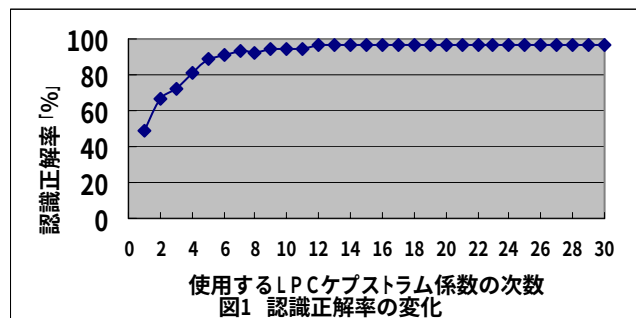
ではまず、LPC ケプストラム係数系列を何次まで認識に使用するかを決定するための簡単なシミュレーションを行う。音声テストデータとしては、男性 2 人、女性 2 人、計 4 人の各 5 通りの音声を用いる。その方法はまず、各人それぞれで 1 つの音声テストデータを標準パターンとし、それ以外の 4 つの音声に対する、認識正解率と認識評価値を使用する LPC ケプストラムの次数  $m$  を 1 から 30 まで変化させて求める。そして、使用する次数  $m$  ごとに求められた 4 人の認識正解率と認識評価値を平均し、その値を用いて LPC ケプストラムの次数の最適性を評価する。

ここで、認識評価値としては、次式を提案した。

$$F(m) = \frac{\sum_{i=1}^5 \frac{d_{ii}^m / 5}{\sum_{\substack{j=1 \\ j \neq i}}^5 d_{ij}^m} / 20$$

( $d_{ij}^m$  =  $i$  と  $j$  の間の次数  $m$  までのユークリッド距離、 $i, j = \{a/i/u/e/o\}$ )

この、シミュレーションの結果は図 1、2 のようになり、これは、LPC ケプストラム係数を 12 次に達すれば、認識正解率は安定し、認識評価値については LPC ケプストラムの次数と共に単調増加するような結果となった。これより、後の評価実験で LPC ケプストラム係数の次数として 12 次を選択することとした。



次に、このシステムの認識精度について調べると、表 1 のようになった。これから、本人の音声は認識正解率約 96%と高いものとなったが、他人の音声については約 58%と改善の余地のある結果となった。

|    | 認識正解率[%] |
|----|----------|
| 本人 | 96.5     |
| 同性 | 68       |
| 異性 | 46       |
| 他人 | 57.2     |

表 1 本システムの認識精度

## 4、まとめ

本研究では、母音音声の認識プログラムを作成し、その評価を行った。その結果、本人の音声では 96%以上、他人の音声では 58%程度の正解率を持つシステムの作成に成功した。この結果から、他人の音声認識についてはこの方法では限界があり、改善の必要があることがわかる。

しかしながら、音声認識についての基礎原理の学習としてはとても有益な実験であり、今後もこの知識を深めていきたいです。

## 参考文献

- [1]音声認識 共立出版 新美康永 1979 年  
[2]音声認識システム オーム社 鹿野清宏ら著 2001 年