

音声可視化システム－観音

中室 健（中室 健二）

音声可視化 "観音"システムは難聴者のコミュニケーションをサポートするものであり、マイクから入力された音声信号に対して信号処理を施し、フォルマント、声の高さや強さに加えて、音声認識による文字情報をリアルタイムで画像表示する。本論文では特に最小相互エントロピー法によるスペクトル解析とその音声スペクトル解析への応用、そして、ニューラルネットワークを用いた音声認識について述べる。

始めに、日本における難聴者の実態を把握し、室内環境における難聴者の不便さをまとめた。そこから難聴者の要求を汲み取り、観音システムのコンセプトを提案し、観音システムの基本構造を示す。

次に音声信号の特徴抽出においては、音声のスペクトル推定法として最小相互エントロピー法を導入する。また、不確定な相関関数による拘束条件を用いた場合の最小相互エントロピー法を導出し、両手法の推定精度の比較を行った。さらに、最小相互エントロピー法とBurg法を組み合わせた手法を提案し、実音声を用いた実験により、スペクトルピークの推定における性能が高いことを示す。

一方、声の高さの情報であるピッチ周波数の推定法として、時間連続性を考慮したピッチ推定法についても述べる。

最後に以上により推定した音声特徴量を用いる音声認識システムを構築した。観音システムでは特に、非文法的な自由発生音に対する単語辞書に依存しない実時間音声認識を課題としている。そこで、階層型ニューラルネットワークにより構築した音声／無音、有声音／無声音認識システム、音素レベルの認識において高い認識率をもつ"時間遅れニューラルネットワーク"による日本語5母音の認識システムを構築し、実音声を用いた認識実験を行った。