

研究概要報告書

資料 - 3

(/)

研究題名	雑音に乱された音声信号の再生とその認識への応用	報告書作成者	小畑 秀文
研究従事者	小畑 秀文 ・ 李 勝		
研究目的	<p>音声言語によるマンマシンインターフェースはヒューマンフレンドリーなものとして、極めて広い応用分野を持つ。しかし、音声を用いられる環境では、一般に背景雑音が極めて高いにもかかわらず、雑音処理技術は極めて未熟な状態にある。本研究では、実際の環境レベルの雑音が加わった音声から、元の音声波形を高精度に再生するための基本技術の確立と、その応用として低品質音声の認識法の開発を目的としている。本研究の進展により、高騒音下で収録した音声や高雑音の回線を経由した音声の再生、高雑音下でも高精度に動作する音声認識装置の実現などが期待できる。</p>		
研究内容	<p>雑音が音声スペクトルに及ぼす影響のうち、処理が困難なものに雑音自身の短時間内変動と雑音と音声との相互干渉とがある。これらは確率的なもので、さらに悪いことにその影響が非常に大きいにもかかわらず、これまでその処理の重要性が見過ごされてきた。本研究では、その課題を正面から取り上げ、前述の確率変動を吸収し、自然性の高い音声信号の再生を実現する手法の確立を目指した。</p> <p>その目的実現のために、音声信号のモデルとして正弦波の重ね合わせたものを提案し、低品質音声信号からの音声信号の復元問題をそれら正弦波の振幅項の推定問題として定式化した（説明書（1））。さらにそのモデルパラメータの最尤推定法を理論的に導いた。</p> <p>次に、現実の処理においては、短時間の分析フレームを用いた処理が基本になるため、フレーム間の推定音声スペクトルに大きな揺らぎが生じる。そこで、ホルマント構造を保ち、かつ音声スペクトルが滑らかに変化するように、新しい時間および周波数の2次元空間で二次元平滑化手法を考案し、再生音声の聴感上の品質を大幅に向上させることに成功した（説明書（2））。また、再生音声の品質評価の一つの尺度として、単音明瞭度テストを行い、SN比0 dBにおいて8.2%の明瞭度向上を確認した（説明書（3））</p> <p>さらに、低品質音声に見られるロンバート効果を含む音声の新しい認識法についても検討を行った。これは周波数領域においてDPマッチングを行い、ホルマントの遷移を認識段階で吸収しようとするもので、実験により本手法の有効性を確認し、問題の一つを解決した。</p>		

様式 - 9

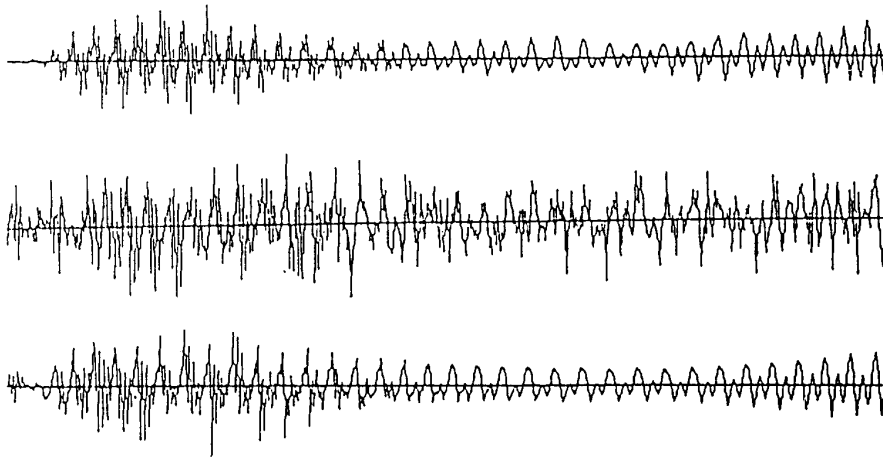
(1) 本研究で用いる音声信号モデル $s(t)$ は次のような正弦波の重ね合わせである。

$$s(t) = \sum_{n=1}^M (a_n \cos n\omega_0 t + b_n \sin \omega_0 t)$$

音声信号の再生問題はこのモデルの係数 a と b の推定問題となる。

(2) 二次元平滑化手法

ケプストラム領域における平滑化を行うことにより、音声スペクトルの滑らかな変化を可能にすることができる。推定音声波形の一例を下に示す。上が元音声、中央がSN比0 dB、下が再生音声である。



(3) 明瞭度試験の結果

16人の被験者にSN比が5 dBと0 dBの低品質音声の単音およびその処理音声を不規則な順序で提示し、聞き取りテストを行った。

表1 明瞭度増減の分布

SN比	明瞭度増加	明瞭度不変	明瞭度増加
5 dB	8人	1人	7人
0 dB	11人	2人	3人

表2 明瞭度試験結果 (16単音×16名)
256回答の中の正答数と率を示す。

SN比	未処理音声	処理音声	明瞭度増加率
5 dB	136	141	1.9%
0 dB	85	107	8.2%

表1は明瞭度の増減の人数内訳、表2は明瞭度の増減内訳である。

(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)