

資料-3

(1/2)

研究題名	音声個人性の音響的特徴の研究	報告書作成者	桑原尚夫
研究従事者	桑原尚夫		
研究目的	<p>音声の「個人性」は、我々個人々の声を特定するだけでなく、様々な重要な情報を担っている。よく知っている人の声であれば、顔は見えなくても、我々は直ちに誰であるかを認めることができる。音声の個人性情報は明らかに音声波に含まれており、全ての音響的特徴量を総合すれば個人の機械による識別は可能である。しかし、個人性情報を担っている特定の音響的特徴量を探すのは必ずしも容易ではない。また、音声個人性の研究は、「個人性」だけに留まらず、声の「質」そのものについての洞察も含まれ、声質あるいは声の（特に合成音声の）自然性の研究への拡張が期待できる。</p> <p>人間の「声」に含まれる個人性の物理的および心理的特徴の抽出を目的に研究を行う。個人性は音声の付加情報として、我々の生活を豊かにすると同時に、技術的には厄介な代物ともなり得る。音声認識に於いては個人性は正規化ないし適応化の対象となり、音声合成に於いては個人性情報をどのように付加または制御するかが大きな問題である。音声個人性の基礎的研究の成果は、このような音声認識・合成のみならず、音声の自然性（人間の声らしさの特徴）や声による個人の同定、声質変換、音声翻訳技術等、様々な応用技術への貢献が期待される。</p> <p>個人性を決める要素には、発声器官の物理的・生理的性質に起因する音響的特徴と、発声の仕方や訛・方言等に起因する言語的特徴の2種類がある。本研究は前者のみを扱い、後者は研究対象から除外することとする。音声の生成過程から、個人性情報は声帯と声道特性の双方に含まれている。まず、対象とする音声信号から双方の音響的特徴を精度よく抽出しなければならない。</p>		

研究内容

音声個人性研究の全体の研究ステップは次のような順序で行う必要がある。

- (1) 種々の声質を持つ不特定多数の話者の音声を収集する。
- (2) 声の「個人性」あるいは「音色」を規定する幾つかの心理属性を設定し、それによって心理実験により声の分類を行う。
- (3) 各音声から声帯および声道特性の抽出を行う。これに対応する音響的特徴量としては、ピッチ周波数、ホルマント周波数、声帯波形、スペクトル包絡特性等がある。
- (4) 聴覚的な特徴（心理的属性）と音響的特徴との相関を調べ、ある属性に共通する音響的特徴を探る。

本助成金研究も上記ステップに沿って研究を行ったが、音響的特徴量をピッチ周波数に限定し、その変化と個人性情報との関係を聞き取り実験によって調べた。

まず、8名の成人男性話者の声を収集し、分析・合成方法によりピッチ周波数その他の音響的特徴量を抽出した。次に、同一方法により、他の音響パラメータは不変に保ったまま、ピッチ周波数のみを変えた音声を作成する。次ページ（様式-10）の図1に本実験で用いた分析・合成方式のブロック図を示す。また、同図2は、実際にピッチ周波数のみを変化させて音声を再合成するアルゴリズムである。これらの作業は全て計算機ソフトウェアで行っている。

各話者の音声に対し、ピッチ周波数を±40%の範囲で変化させ、各話者の個人性がどこまで保存されているかを聴取実験によって行った。ここで、+側はピッチ周波数を高く、-側は低くすることを意味する。聞き取り実験による結果を図3に示す。図3は8名の話者に対する結果を平均したものである。勿論、各話者によって多少の違いはあるが、傾向としては同じである。結果は、ピッチ周波数は個人性に鈍感であることを示している。即ち、ピッチ周波数を大幅に変化させても個人性はあまり変わらない。これは、以前に行ったホルマント周波数に対する結果と大きく異なる。

図1 分析・合成方式のブロック図

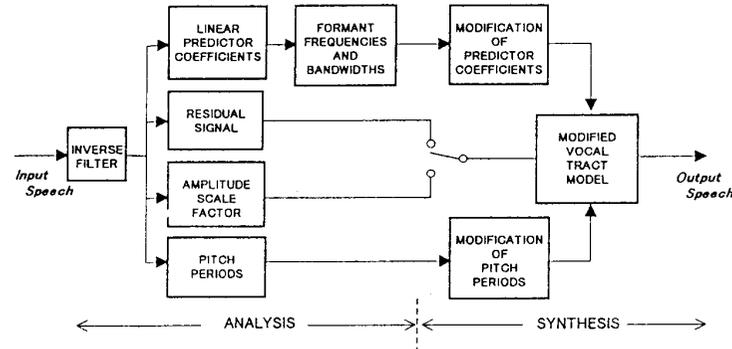


図2 ピッチ周波数のみを変化させて音声を再合成するアルゴリズム

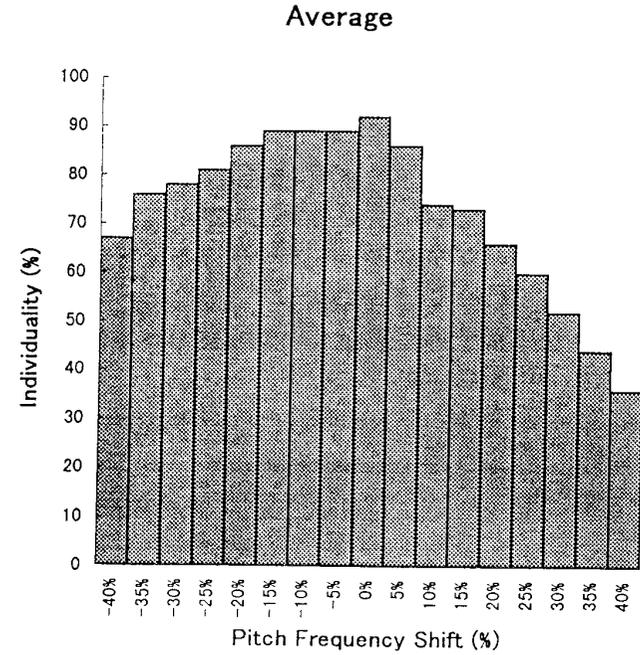
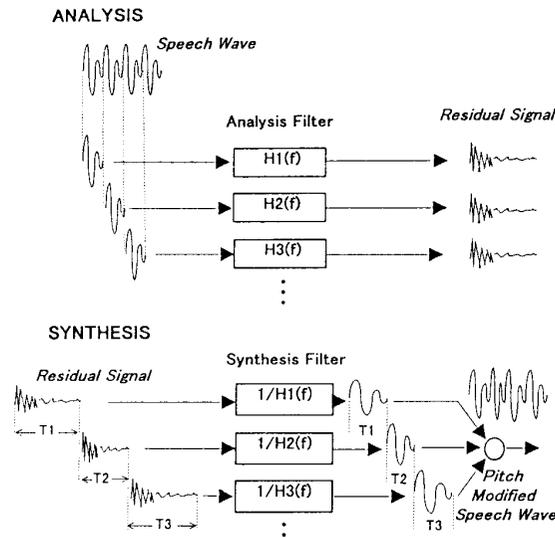


図3 ピッチ変化に対する個人性知覚実験結果。横軸はピッチ変化幅(0%の位置は原音声)、縦軸は個人性保存量

(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)