

Title	スピーチにおける感情的な知覚の多層ファジィ論理的なモデルの構築
Author(s)	黄, 純芳
Citation	
Issue Date	2004-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1884
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

ファジィ論理を用いた感情音声のための多層知覚モデルの構築

黄 純芳 (210204)

北陸先端科学技術大学院大学 情報科学研究科

August 13, 2004

キーワード: 感情的な音声知覚、ファジィ論理、多次元的尺度構成法、音響特徴.

1 序論

本論文では、感情音声知覚の能力をシミュレートするための多層知覚モデルを提案する。先行研究では、感情音声について音響特徴と感情知覚のみの関係に着目していた。しかしながら、これらの研究は1つの重要なポイントを見落としていると考えられる。それは、音声から感情を知覚するときは、直接音響特徴量に基づかないため、基本的な心理的特徴を考慮しなければならないということである。本論分で提案するモデルは、感情レイヤ、基本的な心理的特徴レイヤ、音響特徴レイヤの3つのレイヤを含んでおり、感情音声は基本的な心理的特徴を表す言葉により定義される。

提案モデルを完成するためには二つのアプローチを考える必要がある。トップダウンアプローチを採用してモデルの枠組みを構成すること、そして、既存の枠組みについて確かめるためにボトムアップ・アプローチを採用することである。本論文では、主にトップダウンアプローチに着目して、トップダウンアプローチでモデルを構築することを目的とする。このため、感情と基本的な心理的特徴、基本的な心理的特徴と音響特徴との二つの関係を調べた。

2 感情レイヤと基本的な心理的特徴レイヤとの関係

本研究では、感情レイヤに位置する五つの感情、Neutral、Joy、Cold Anger、Sadness と Hot Anger を調査した。感情レイヤと基本的な心理的特徴レイヤの関係を調べるために、三つの聴覚実験を行った。これらの実験の結果にファジィ理論を応用することによって、その関係を構成する。

2.1 3つの実験

最初の実験では、発話された音声の感情評定を行った。被験者は、171発声と5つの感情に評定した。実験の結果から171の発声は五つの感情との関係の強さが分かった。

2番目の実験の目的は、心理的距離モデルを構成することであった。一対比較の方法で実験を行った。被験者は一組二つの発声に関して知覚した類似度を評定した。実験の結果と多次元的尺度構成法(MDS)により、心理的距離モデルを作成した。3番目の実験を行う前に、予備実験で34個の形容詞が60個の形容詞から選ばれた。3番目の実験では、被験者は発話された音声ごとにその34の形容詞を評定した。評価結果と形容詞を心理的距離モデルに重ね、34個の形容詞がどのように感情に関連するかを相関値をもと計算し、15の形容詞を提案されたモデルの基本的な心理的特徴のメンバとして選択した。

2.2 ファジィ理論による感情と基本的な心理的特徴の関係の構成

3つの実験の結果をもとに、ファジィ論理を応用することによって、感情と基本的な心理的特徴との関係を築く。各感情において、ファジィ推論システム(FIS)が、MATLAB Fuzzy Logic Toolboxを使用することによって構成された。主に3つのステップでファジィ推論システムは構成される。最初に、初期モデルを構成するために、初期モデルのメンバーシップ関数とルールが選択される。2番目に、初期モデルが適応型のニューロファジィ・テクニクによって訓練された。3番目に、検査エラーとトレーニング・エラーに応じて、メンバーシップ関数は再調整される。入力データ(すなわち、基本的な心理的特徴)に応じた出力(すなわち、感情度合い)の曲線に対して、カーブに合うように回帰直線を計算することによって、基本的な心理的特徴と感情の関連性を調べる。回帰直線の傾きは感情と基本的な心

理的特徴との関係を示している。5つの感情に対して、5つのFISを作成した。

3 基本的な心理的特徴レイヤと音響特徴レイヤとの関係

知覚モデル内の基本的な心理的特徴層レイヤと音響特徴レイヤとの関係を見つけるために、音響特徴の性質ピッチ、ラウドネス、および音色に関する物理量であるF0、パワー、およびパワースペクトルがSTRAIGHTによって計算された。ピッチとラウドネスに関して、アクセント句と一般的な発声の両方に関係するいくつかのF0とパワーの音響特徴が測定された。測定された音響特徴と基本的な心理的特徴の間の相関係数を用いて、基本的な心理的特徴に最も関連したそれらの音響特徴が選択された。F0の分析により、選択された音響特徴は一番高いF0(HP)、F0の平均値(AP)、上昇時の傾きの平均値(RS)、第1句での上昇の傾き(RS1st)の音響特徴がある。パワーの分析により、選択された音響特徴はアクセント句のパワー範囲の平均値(PRAP)、パワー範囲(PWR)、最初上昇の傾き(RS1st)、高周波数の平均パワーと全周波数の平均パワーの比率(RHT)の音響特徴がある。

音色に関しては、完全にまだ分析できたというわけではないが、基本的な心理的特徴に影響する音色に関する音響特徴を見つけるために、分析の結果の一部が、非常に役に立つ手がかりを提供した。それらは平均パワースペクトルと母音パワースペクトルの変化である。

4 結論

本論文では、感情音声知覚をシミュレートするための三層レイヤの知覚モデルを報告した。このモデルを構成するために、感情レイヤと基本的な心理的特徴レイヤとの関係については3つの実験を行い、実験結果にファジィ論理を適用することによってモデルを構築した。それぞれのFISの結果が示したように、関係はよくモデル化できた。基本的な心理的特徴レイヤと音響特徴レイヤとの関係は、基本周波数とパワーについて部分的に築き上げられた。