

# アクセント型とモーラを考慮することによる感情を含んだ音声変化の解析

## Analysis of the voice including the feeling by taking an accent type and mora into consideration

越智 雅人\*<sup>1</sup>  
Masato Ochi

黒田 久泰\*<sup>1</sup>  
Hisayasu Kuroda

\*<sup>1</sup> 愛媛大学  
Ehime University

\*<sup>1</sup> 愛媛大学  
Ehime University

### 1. はじめに

感情を含んだ音声はさまざまな韻律的特徴の変化が見られ、それを調べた研究が数多く存在する。過去の研究[1][2][3]で、基本周波数やフォルマント周波数の構造、アクセントなどの振幅構造、文の長さなどの時間構造などが感情を分類するうえで有用であることが証明されている。しかし、アクセント型とモーラの両方を考慮した感情認識は少なく、またモーラ単位で特徴量を定義しているものは少ない。

本研究では、感情を含むことによって、アクセント型とモーラにどのような影響があるのかを、さまざまな感情を含んだ音声集合について、特徴量を抽出し主成分分析を行うことによって明らかにする。

### 2. 音声変化の解析

#### 2.1 特徴量

本研究では、感情音声の特徴量としてピッチ(基本周波数)、振幅、モーラ時間を使用する。重永[3]らの研究では以下の特徴量が用いられているので、本研究でも使用する。

pm:発話中の基本周波数の最大値(Hz)  
am:発話中の振幅の最大値(db)  
tm:発話中の開始から終了までの時間(sec)

また、上記の特徴量とかなりの相関があると考えられるが、必ずしも不必要とはいえないので補填的な意味で以下の特徴量を使用する。

pa:発話中の基本周波数の平均値(Hz)  
aa:発話中の振幅の平均値(db)

本研究では、モーラを考慮した分析を行うため、モーラごとの特徴量も考慮する。

pmmi:発話中の第*i*モーラの基本周波数の最大値  
ammi:発話中の第*i*モーラの振幅の最大値  
pami:発話中の第*i*モーラの基本周波数の平均値  
aami:発話中の第*i*モーラの振幅の平均値  
tmmi:発話中の第*i*モーラの時間

さらに、本研究では隣接モーラの間隔を明らかにするために、以下も特徴量として加えることにする。

pmami:発話中の2モーラ間の基本周波数の最大値の差  
amami:発話中の2モーラ間の振幅の最大値の差

連絡先: 〒790-8577 松山市文京町3番

paami:発話中の2モーラ間の基本周波数の平均値の差  
aaami:発話中の2モーラ間の振幅の平均値の差  
tmami:発話中の2モーラ間の時間の差

隣接モーラ間における特徴量で3モーラの場合、*i*が1のときは2-1モーラ間、2のときは3-1モーラ間、3のときは3-2モーラ間を表す。

#### 2.2 実験手法

本研究では、音声解析の対象として頭高型、中高型、平板型の3モーラ音声であるmidori, nagame, nagameを使用する。この各単語に対して、2.1で述べた35次元の特徴量を用いて主成分分析を行う。その結果を考察することで、感情を含むことで音声の韻律的特徴がどのように変化するかを明らかにする。

#### 2.3 音声資料

音声資料として、国立情報学研究所音声資源コンソーシアムを通して公開されている東京工芸大学森山剛助教らが作成した慶應義塾大学研究用感情音声データベースを利用した。このデータベースには舞台経験のある男性に対して、さまざまな感情を含んだ20単語を録音したものである。本研究においては、midori, naname, nagameに対して47の感情音声データを用いた。感情の種類を表1に示す。

表1 47感情の内訳

1 怒り	2 喜び	3 嫌悪
4 悔り	5 おかしい	6 心配
7 優しい	8 安堵	9 憤慨
10 羞恥	11 穏やか	12 憧れ
13 苛立ち	14 不平	15 切望
16 気の毒な	17 寛容	18 ほくそえむ
19 失望	20 叱責	21 悲しい
22 恐れ	23 憎い	24 軽蔑
25 嬉しい	26 皮肉	27 無関心
28 賞賛	29 誇り	30 愛
31 嘆き	32 媚び	33 満足
34 退屈	35 苦しい	36 期待
37 幸福	38 好き	39 嫌い
40 いや	41 落胆	42 非難
43 不安	44 驚き	45 慌て
46 あきれ	47 平静	

### 3. 実験結果

#### 3.1 寄与率

midori, nagame, nagame に対して主成分分析を行った結果、得られた寄与率を表 2 に示す。この結果から、第1主成分及び第2主成分が大きな寄与率を示した。

#### 3.2 因子負荷量

midori, nagame, nagame において、第1主成分に対する因子負荷量を表した図を図 1, 2, 3 にそれぞれ示す。図において、赤い部分が第1主成分と強い相関を示している。

#### 3.3 主成分得点

midori, nagame, nagame に対して主成分分析を行った結果、得られた主成分得点の一部を表 3 に示す。ただし、1列目の番号は表 1 に対応している

#### 3.4 考察

図 1 において、pm, pmm2, pmm3, am, amm2, amm3, pa, pam2, pam3, aa, aam2, aam3 の特徴量が第1主成分と強い負の相関を示していることがわかる。つまり、頭高型(midori)では第1主成分の主成分得点が低いほど、2,3拍目のピッチと振幅が高いことを示している。他、図 2, 3 からも同じように推測すると、中高型(naname)では第1主成分の主成分得点が低いほど、1拍目に比べて2,3拍目の方がピッチと振幅が高いことを示しており、平板型(nagame)では第1主成分の主成分得点が低いほど、1,2拍目に比べて3拍目の方がピッチと振幅が高いことを示している。

### 4. まとめ

ここまでのことから、感情を含んだ音声にはどのような韻律的特徴の変化が見られるのかを、アクセント型とモーラ単位で抽出した特徴量の視点から分析した。その結果、アクセント型の違う midori, naname, nagame では同じ感情を含んだ音声でも、韻律的特徴に違いがあることがわかった。

本研究では、3モーラ音声でしか試していないため、それ以外のモーラ数に対して同様の実験をしたいと考えている。

表 2 midori, nagame, nagame に対する寄与率

	midori	naname	nagame
第1主成分	37.15	37.05	33.54
第2主成分	20.42	24.59	28.43
第3主成分	14.72	11.95	11.10
第4主成分	8.8	10.03	8.14
...	...	...	...

表 3 midori, nagame, nagame に対する主成分得点

	midori		naname		nagame	
	PC1	PC2	PC1	PC2	PC1	PC2
1	-5.807	3.259	-7.223	-0.579	-5.263	2.827
2	-3.303	2.872	-5.49	-0.735	-1.15	2.223
3	1.628	2.473	0.156	-2.43	-4.124	-3.18
4	4.75	-3.788	4.85	1.424	4.036	0.166
...	...	...	...	...	...	...

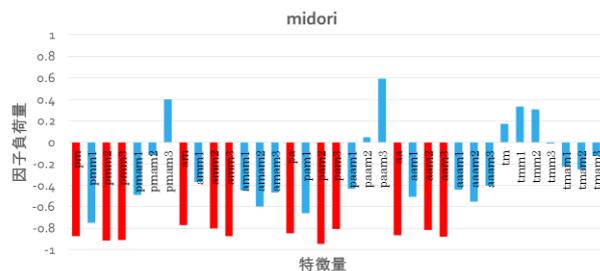


図 1 midori に対する因子負荷量

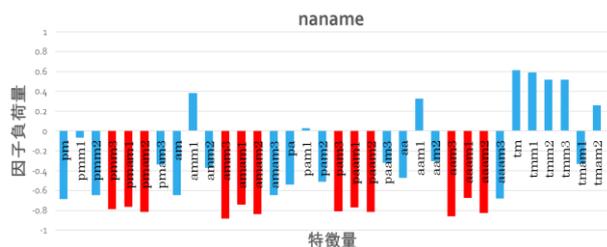


図 2 naname に対する因子負荷量

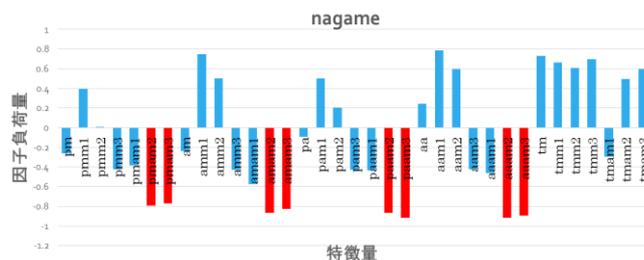


図 3 nagame に対する因子負荷量

### 謝辞

素晴らしい感情音声データベースを提供してくださった森山助教らに感謝いたします。

### 参考文献

- [1] 上床弘幸, 小林豊, 新美康永: 音声の感情表現の分析とモデル化, 電子情報通信学会技術研究報, SP92-131, pp.65-72, 1993
- [2] 川波弘道, 広瀬啓吉: 態度・感情音声における韻律的特徴の考察, 電子情報通信学会技術研究報, SP97-67, pp.73-80, 1997
- [3] 重永実: 感情の判別分析からみた感情音声の特性, 電子情報通信学会論文誌, Vol.J83-A, No.6, pp.726-735, 2000