

June 28, 2019

音声が感情を含むことによる声質変化の解析

沼田 賢一

富山県立大学 情報基盤工学講座

1. はじめに
2. 手順
3. まとめ

June 28, 2019

背景

音声は、複雑な情報が含まれていて様々な分野で研究が行われている。

先行研究

従来 of 感情音声の研究の多くは、音声の韻律的特徴を調査したものが多い。

しかし、全く異なる心理状態でも韻律的特徴が類似していて声質によってのみ区別されるものが存在することもあることが問題点であった。

目的

本研究では、感情を含む音声の声質がどのように変化するかわかるようにするために音節を分けて主成分分析を行うことで、その変化する主たる声質成分を明らかにする。

感情に特徴的な声質の変化とは、静的な周波数構造の変化と時間的に変動する動的な変化の組み合わせである。

そこで、静的な周波数構造を表現するパラメータとしてメル周波数ケプストラム係数 (MFCC) を用いる。

メル周波数ケプストラム係数

$$f_{met} = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (1)$$

f_{met} :メル尺度

次に、音声波形に対し高速フーリエ変換をし、メル尺度に基づいた L 個の帯域フィルタを周波数軸に配置したものと掛け合わせる。最後に (2) に示すフィルタバンク分析によって得られた L 個の帯域におけるパワーに対して離散コサイン変換をすると MFCC を得られる。

$$c_i = \sqrt{\frac{2}{3}} \sum_{l=1}^L A_l \cos \left[\left(l - \frac{1}{2} \right) \frac{i\pi}{L} \right] \quad (2)$$

N :フレーム長 A_j 対数フィルタバンクの振幅 i :MFCC の次数

次に、J 個の音声で MFCC を K フレームにわたり算出したものをもとめる。

そして、変数の類似性やポジショニングを明らかにするために、主成分分析を行って主成分得点を求める。

主成分分析とはある変数群に対してその変数相互の関係から新しい概念のファクタを導き出す統計的手法のこと。

主成分数の決定法

主成分得点の分散である固有値が大きいほど、その主成分の説明力は高くなる。第 k 主成分の固有値 λ_k が固有値の合計に占める割合を寄与率として次式で表す。

$$r_k = \frac{\lambda_k}{\sum_{l=1}^{p \times K} \lambda_l} \quad (3)$$

第 k 主成分までの固有値の合計が全固有値の合計に占める割合を累積寄与率とする。本研究では、70 % を超える主成分まで抽出する。

実際に様々な感情を含んだ音声を用意し、主成分分析まで行う。また、この結果に対してアクセント核に対するものと非アクセント核に対するもののとの比較を行った。話者は一名で、発言内容は「なま/nama/」という 2 モーラ頭高型の語とする。また感情表現の手がかりとして表 1 のような 47 種類の感情語を教示した。音声の長さは 20 フレームで音声で 13 次の MFCC にする。

表 1 日常の感情を表現する言葉

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

アクセント核/na/に対して分析を行った結果を表2に示す。
図1及び図2におけるマーカとそれらに対応する感情語を表3に示す。また、図1に横軸を第1主成分、縦軸を第2主成分とした感情ごとの声質ベクトル（主成分得点ベクトル）の布置図、図2に横軸を第1主成分、縦軸を第3主成分とした感情ごとの声質ベクトルの布置図を示す。

表2 アクセント核/na/に関する主成分分析

	寄与率[%]	累積寄与率[%]
第1主成分	31	31
第2主成分	28	59
第3主成分	12	71
第4主成分	5	76
第5主成分	5	81
...

表3 マーカと対応する感情語

平穏	△	安堵	○	気の毒な	×	軽蔑	*	媚び	□	いや	◇
怒り	△	憤慨	○	寛容	×	嬉しい	★	満足	□	落胆	◇
喜び	△	羞恥	○	ほくそえむ	+	皮肉	*	退屈	□	非難	▽
嫌悪	△	穏やか	○	失望	+	無関心	★	苦しい	□	不安	▽
悔り	△	憧れ	×	叱責	+	賞賛	*	期待	◇	驚き	▽
おかしい	△	苛立ち	×	悲しい	+	誇り	*	幸福	◇	慌て	▽
心配	○	不平	×	恐れ	+	愛	□	好き	◇	あきれ	▽
優しい	○	切望	×	憎い	+	嘆き	□	嫌い	◇		

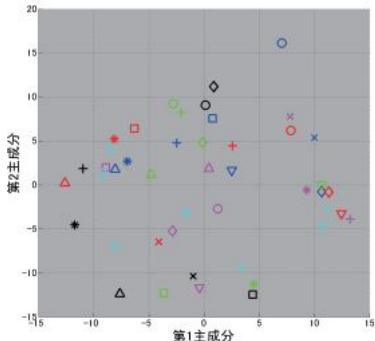


図1 第1—第2主成分平面への感情の布置

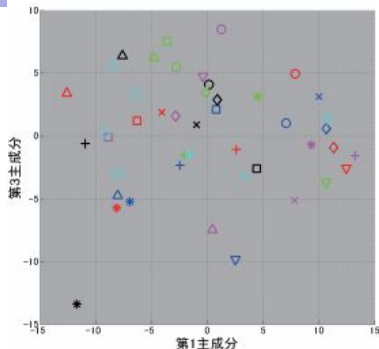


図2 第1—第3主成分平面への感情の布置

両布置図において、「恐れ」「慌て」「切望」は近傍にプロットされ、対極に「喜び」「憎い」「誇り」がプロットされた。前者は息混ぜでスピード感をもって発話されているのに対して、後者の音声は逆に息をあまり使わずに発話される特徴があることから、第一主成分は生き混ぜに関する軸と考えられる。

次に、非アクセント核/ma/に対して分析を行った結果を表4に示す。図3に横軸を第1主成分、縦軸を第2主成分とした感情ごとの声質ベクトル（主成分得点ベクトル）の布置図、図4に横軸を第1主成分、縦軸を第3主成分とした感情ごとの声質ベクトルの布置図を示す。

表4 非アクセント核/ma/に関する主成分分析

	寄与率[%]	累積寄与率[%]
第1主成分	30	30
第2主成分	24	54
第3主成分	18	72
第4主成分	6	78
第5主成分	3	81
...

実験結果

10/12

はじめに
手順
実験
おわりに

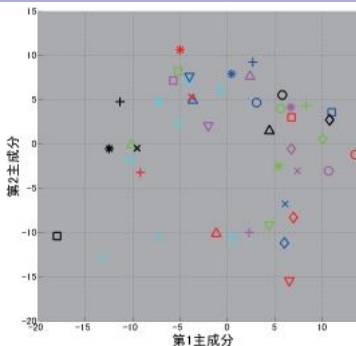


図3 第1—第2主成分平面への感情の布置

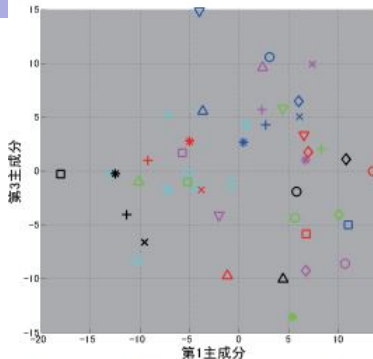


図4 第1—第3主成分平面への感情の布置

10/12

両布置図上において、「愛」「安堵」「落胆」といった感情が、第1主成分軸の正に大きな得点を有し、「苦しい」「悲しい」「誇り」といった感情が、その対極において負の大きな得点を有している。
よって、非アクセント核における最も主要な声質の変化は、「穏やかさ」軸であると解釈できる。

13 次元 20 フレームの時間軌跡 260 次元で表されていた声質を, 主成分分析を行うことで 5 次元程度の低次元で表すことができた.

また, 第一主成分に関して息混ぜによる声質変化が最も主要なもので, 第二主成分に関しては穏やかさによる声質変化が最も主要なものであるとわかった.

まとめ

- ① 音声の声質が、感情を含むことによってどのように変化するかを、様々な感情を含んだ音声のメル周波数ケプストラム係数 MFCC の時間変動に対して主成分分析を行うことによって明らかにした.
- ② 音節がアクセント核であるか否かで感情を含むことによる声質の変化に関して主要な成分が異なることが示された.