

高齢者を対象とした音声解析による健常者と認知症患者の識別

Discrimination of the elder healthy and dementia patients by speech analysis

宮川 大[†] 花沢 明俊[†]
Masaru Miyagawa Akitoshi Hanazawa

1. はじめに

近年、認知症患者の増加が懸念され、様々な研究がなされている。その中でも認知症の早期発見に関する研究は多く、様々なアプローチから早期発見を目指している。本研究では、診断手法について着目し、従来の診察方法であるMMSE[1]や HDS-R[2]や血液、MRI 検査とは異なる、コストや手軽さを追求した、音声による認知症診断を考案した。さらに、Talerらは、非常に軽度のアルツハイマー病患者を対象に発話において、文法及び、感情韻律上に障害が認められることを報告している[3]。これらの報告は、韻律特徴の解析を、認知症のスクリーニング検査として用いることができる可能性を示唆している。

2. 関連研究

黒川らの研究では、認知症患者に関する音響的特徴を分析した研究が行われ、k、s、h、q、nの音素の発話を苦手とする傾向が認められ、u、oの母音でのフォルマントに有意差が見られることが示された[4]。また、Benbaらの研究では、音声解析によって3つのタイプの神経疾患患者同士を90%以上の精度で識別した研究がある[5]。この研究では神経疾患の認知機能が由来の発声障害に着目し、声の震えや、発音等の韻律特徴を扱い、高い識別精度を達成した。

これらの研究から、本論文では認知障害を患っている認知症患者にも発声障害が見られるため、その韻律特徴をBenbaらの研究で扱われたメル周波数スペクトラム係数(MFCC)と、フォルマントに関する特徴量を抽出することができる線形予測(LPC)分析を採用した。また、この2種類の手法を用いて認知症を識別した前例はない。本論文では、先述した2つの手法についての有効性について議論するため、健常者と認知症患者の音声からそれぞれ特徴量を算出し、ラベル付けを行い、分類器にかけ識別精度を検証し、有効な特徴量やパラメーターを決定する。今回は分類器として、2クラス間の識別精度が優れているSVMを採用した。

3. 特徴抽出

3.1 MFCC

人間の音声信号は、準定常と呼ばれており、10~30ms程度の短い区間に観察すると、かなり安定しているように見える[6]。したがって、音声信号は短時間スペクトル分析を行い、特徴付けることが適切である[6]。そこでMFCCでは解析する音声信号を分割して処理し、音声信号から高域周波数成分である音源特性を除き低域周波数成分である声道特性についての特徴を抽出する。また、MFCCではケプストラム分析を行い、音声波形の振幅スペクトルについて計算を行い、この計算によって導出された低次元成分がMFCCとなる[6]。また、MFCCの計算はMelscaleに基づい

ており、周波数の計算領域を人間の周波数知覚に則している[6]。

3.2 LPC 分析

LPC分析では自己相関関数を算出し、LPC係数を導出する。音素を特徴づける周波数成分は、声道の共振周波数に対応し、フォルマントと呼ばれている[7]。また、音声のスペクトルで時間的に移動している複数のピークのことをいい、周波数の低い順に第一フォルマント、第二フォルマントと呼ばれる。ここでは、そのフォルマントについての特徴量を導出する。このLPC分析では声道を音響管連結として考え、自己回帰モデルと捉え相関関数を算出している。

3 実験

本研究の実験として、前述した2種類の特徴量抽出手法であるMFCCとLPC分析を用いて、収集した高齢の健常者と認知症患者の音声解析し、採取したそれぞれの特徴量からSVMを用いてクラスタリングし、2種の識別精度を求める。また、より精度が高く出された各パラメーターを求める。

3.2 音声データ

本研究で扱うデータは60歳以上の認知症患者20名と50歳以上の軽度認知症患者を含む、健常な高齢者20名による音声を集めた。録音内容は、知能テストの音声を扱う。また、HDS-Rを実施することで音声を得ると同時にその高齢者の認知症度を測ることができる。それに基づいて音声のラベル付けや、評価を行う。扱うラベルは健常者と認知症患者だが、今回は軽度の認知症傾向がある被験者が居た。しかし、点数が1、2点低い程度であるので健常者とラベル付けをした。表1に音声データの概要を示す。

認知症患者	20 個
健常者	16 個
軽度の認知症傾向	4 個
データ総数	40 個

表1 音声データの概要

3.3 識別方法

今回使用した識別器として、機械学習の一手法であるサポートベクターマシン(SVM)を採用した。SVMは2クラス間の識別精度が非常に優れている学習モデルである。このSVMによって健常者と認知症患者の韻律的特徴について、クラスタリングを行うことにした。また、MFCCとLPC分析を用いてSVMを実行する上で、より良い精度にするため、各種パラメーターの調整を行った。

SVMでは、ハイパーパラメーターであるC、カーネル、 γ の値、種類をグリッドサーチによって決定した。MFCC、

[†]九州工業大学大学院 工学府 先端機能システム工学専攻

LPC 分析におけるパラメーターの調整は、データ全体から求められた特徴量を元に、次元数と音声信号を解析する際のフレーム長、FFT のサンプル数を調整しながら精度を求め、平均してより高い精度を得ることができるパラメーターを決定した。

4. 結果

表 3 に求められた最適なパラメーターを示している。このパラメーターに設定し、高齢の健常者と認知症患者の識別精度を求めた。また、データ数が少ないことにより実行ごとに精度が変化するので 100 回実行した平均を出した。以下の表 3 に実行結果を示す。

次元数	フレーム長	サンプル数	学習:テスト	カーネル	C	γ
20	0.5	4096	36:4	線形	1	-
48	0.02	1024	34:6	RBF	1	0.1

表 2 MFCC(上)、LPC(下)での最適なパラメーター

	精度
MFCC	約 97.3%
LPC	約 99.8%

表 3 MFCC と LPC それぞれの識別結果

また、多次元尺度構成法によって多次元に渡る MFCC や LPC の特徴量を 2 次元に圧縮し、音声データ群の大まかな分布図を作成した。多次元尺度構成法とは、データの類似性を調べるための可視化によく用いられる統計的な手法で、対象間の類似度をもとに、その関係を可能な限り保つように対象を低次元空間の点で表すことができる[21]。

この手法を用いる事で、健常者と認知症患者のそれぞれの特徴量に関する類似度を可視化した。

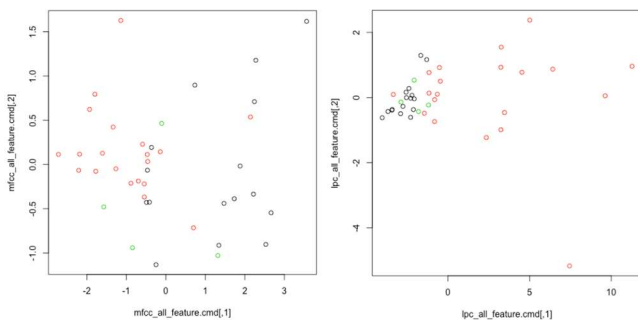


図 1 多次元尺度構成法で圧縮した各被験者の MFCC (20 次元)、LPC 係数 (48 次元) の散布図 (認知症 (赤)、健常者 (黒)、軽度認知症 (緑))

5. 結論・考察

本論文での研究結果として、LPC 分析を採用した SVM での最高精度を 99.8% で達成することができ、分類器での高齢の健常者と認知症患者において、健常者と認知症患者との間に存在するフォルマントに関する発話特徴が非常に有効であることが実証された。また、MFCC での分類でも 97.3% の識別精度を達成することができ、その有効性を確

認することができた。また、LPC 分析で得た被験者それぞれの特徴量を圧縮した多次元尺度構成法における分布図において、健常者と認知症患者がはっきりと分かれている。さらに、健常者の中の軽度な認知症傾向が見られた被験者の位置として健常者と認知症患者の間に分布しているのが見て取れる。これにより、LPC 分析が多次元尺度構成法を用いた分布図から認知症傾向が見られる可能性を示唆している。しかし、音声データの録音環境が悪いことからノイズ等が影響してこの分布図にバラツキとして現れている。よって、ノイズの少ない音声データを使用することでより正確な分布図が得られると考えられ、有効性をさらに裏付けられると考える。

6. おわりに

この研究によって、認知症の早期発見に貢献できるようにするには、軽度認知症と健常者を識別できるかを検証する必要がある。また、この音声解析では特徴量から両者の境界線を求める SVM によって識別を行なっているため、より多くのデータがあれば精度が増し、この研究の実証性も増す。しかし、機械学習の留意すべき点として過学習を避けなければならない。データが多いからといってデータの質が悪ければ、認知症との関連性のない特徴量を拾ってしまい、精度が悪化する恐れがある。よって、より質の高いデータを入手する必要がある。無音区間の時間や音節毎の基本周波数等の特徴も有意差が認められているので、精度の向上に貢献すると考えられる。

参考文献

- [1] M. F. Folstein and S. E. Folstein, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician", *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, (1975).
- [2] H. Buschke, G. Kuslansky, M. Katz et al., "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician
- [3] Taler, V., Baum, S. R., Chertkow, H., and Saumier, D., "Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease, *Neuropsychology*", Vol. 22, No. 2, pp. 188–195 (2008)
- [4] Yuki Kurokawa, "音響的特徴を利用した高齢者の認知傾向の分析", 愛知県立大学情報科学部卒業論文(2016)
- [5] Benba Achraf, Jilbab Abdelilah, Hammouch Ahmed, "Detecting multiple system atrophy, Parkinson and other neurological disorders using voice analysis," *IEEE Conference Proceedings Vol.2016 No.ICEIT Page.406-409* (2016)
- [6] Benba, A., Jilbab, A., & Hammouch, A. "Discriminating between patients with Parkinson's and Neurological diseases using cepstral analysis." *IEEE transactions on neural systems and rehabilitation engineering* (Vol. 24, No. 10, pp. 1100–1108(2016)).
- [7] 中山弘二, "音声分析合成の理論と仕組み", 岡山理科大学総合情報学部情報科学科卒業論文 p11~12, 2007