

環境音を用いた人の動作推定に関する検討

T19M057 北迫 祐樹

指導教員 三好 力 教授

1. はじめに

近年スマートフォンやスマートスピーカーといった音声操作を目的とした入力デバイスが普及している。しかし、音声認識による操作には同じサービスを受けるために同じ音声を繰り返す必要があるという不便な点もある。

本研究では、マイクロフォンを内蔵し音声認識による操作を目的とした音声 AI 機器を利用し環境音を用いて人の動作を推定するシステムの提案、システムの実現に必要な環境音の識別、識別し記録された特定の環境音のパターンが分類可能であるかについて検証を行う。

2. 提案システム

特徴的な環境音をシンボル音として分類し、ユーザーの家に設置された複数台の音声 AI が搭載された機器をつかい、命令やシンボル音の種類と発生した時間とを記録し収集する。収集したデータを用いることで、特定のシンボル音の並びが類似した時間間隔で発生する度に特定の命令が行われているといったパターンを分類し、頻度の高い特定の命令の前に発生したシンボル音のパターンとそれに類似したシンボル音の発生パターンを認識したとき、特定の命令を行うことをユーザーに提案する。

システムの実現のため①特定の環境音の識別が可能であるかの検証、②識別し記録された特定の環境音のパターンが分類可能であるかについて検証を行う。

3. 実験①

3つの特徴音をそれぞれ120個ずつ録音したデータを用意した。録音したシンボル音はスペクトログラム画像になるように短時間フーリエ変換を行い、380 * 380 pix の画像データに変換した。1つの音につき学習用データ80個、検証用データ20個のスペクトログラム画像で学習を行った結果、CNNモデルの識別精度が未確認データに対して平均97.8%と高い精度で

識別が可能であることがわかった。

4. 実験②

録音した音を基に検証用のダミーデータを作成し、発生順の項目それぞれが発生した時間間隔の項目のデータに変換しスケーリングしたデータ列を用意した。主成分分析による次元削減を行い、図1の様なグラフを作成した。

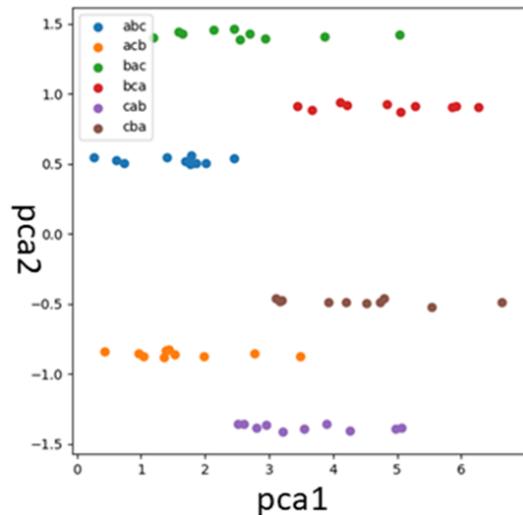


図1: 主成分分析の結果グラフ

第1主成分には発生時間の間隔における項目、第2主成分には発生順の項目が関係していると考えられる。これにより、識別により収集したデータから命令が行われる前に発生したシンボル音の時間的なパターンと順序的なパターンの分析が可能であり同一のパターンを含んだクラスターを見つけ出すことで類似したシンボル音のパターンから特定の命令を見つけ出すことが可能になるといえる。

5. おわりに

環境音を用いて人の動作を推定するシステムの提案、システムの実現に必要な検証として①特定の環境音の識別、②識別により収集したデータから特定のシンボル音のパターンが分類可能であることを検証した。