

音響特徴量を用いた自閉症児と 定型発達児の識別*

☆石井良, 高島遼一, 滝口哲也, 有木康雄 (神戸大),
中井靖 (川崎医療短期大), 高田哲 (神戸大)

1 はじめに

音声認識技術は、これまでの多くは成人を対象としたものであったが、現在では子供や高齢者などの発話スタイルが成人と異なる人を対象とした場合や[1], また車内や会議室といった様々な環境下や場面を対象とした場合など多様化し、使用される機会が増加している。本論文で対象としている自閉症患者のように、日本には現在多くの発達障害を患った人がいること、また近年の高齢化社会の問題から、福祉分野における情報技術の発展が重要である[2]。障害を患った人など、健常者とは異なる人を対象とした研究への期待は大きいといえる。

自閉症とは、先天性の脳機能障害であり、社会性や他者とのコミュニケーション能力に困難が生じる発達障害の一種である。早期幼児自閉症、小児自閉症、カナー自閉症などと呼ばれ、日本には現在約36万人の自閉症を患った人がいる。

そもそも自閉症は、現代医学では根本的な原因を治療することは困難とされている。しかし、病気において早期発見と早期治療が重要とされているように、自閉症に対してもその早期発見は非常に重要なテーマとなっている。

本論文では、幼稚園児から小学校4年生までの自閉症児を対象に、その早期発見を目指した音響的な側面による識別実験の結果に関して報告する。

具体的には、openEAR[3]を用いて自閉症児、定型発達児の音声から989次元の特徴量を抽出し、その特徴量に関して Forward selection method[4]により適切な特徴量を選択した後、SVM (サポートベクターマシン) によって識別を行う。

2 openEAR

本論文では、従来感情音声認識[5]などに用いられていたツールキットである openEAR によって抽出された音響特徴量を用いて、音声識別実験を行う。openEAR は音声ファイルから989次元の音響特徴量を抽出することができる。しかし、抽出された次元が大きすぎることによる計算時間の増加や、不必要な情報が多く含まれていることなどの問題点があげられる。

3 Forward selection method

本論文では、2章で述べた次元過多の問題点を解決するための手法として、Forward selection method を用いる。

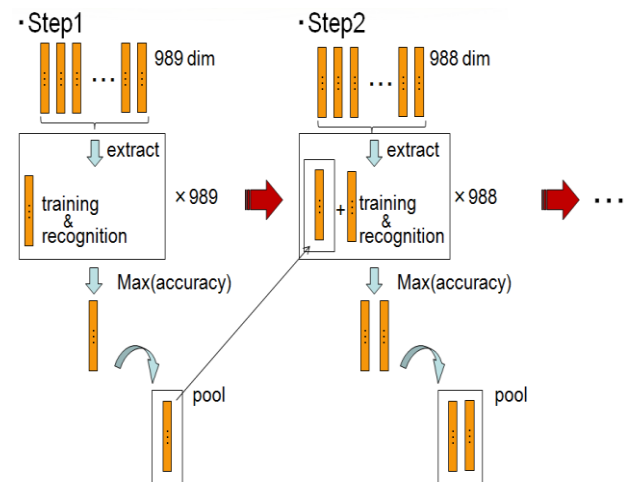


Fig. 1 Forward selection method

Fig. 1 において、まず Step1 として989次元ある特徴量から順に1次元ずつ取り出し、その取り出された次元ごとに単一で識別精度を算出する。本研究では識別器として SVM を用いている。989次元それぞれの識別精度を算出した後、最も精度の高かった次元を取り出す。この取り出した次元を仮に A 次元とすると、Step2 では A 次元を取り除いた988次

* Classification of autistic children and ordinary children using acoustic features , by Ryo Ishii, Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Arika (Kobe University), Yasushi Nakai (Kawasaki College of Allied Health Professions), Satoshi Takada (Kobe University)

元の特徴量の中から、それぞれ1次元ずつ取り出し、A次元と組み合わせた2次元で識別精度を算出する。以上の工程を、StepN-1における識別精度よりStepNの識別精度が劣るまで継続する。これにより、openEARによって抽出された特徴量から、自閉症児と定型発達児の識別に最適な特徴量を選択する。本研究では、Forward selection methodにおける特徴量選択は10次元で収束した。

4 実験

4.1 ベースライン

本論文では、比較対象として、特徴量をMFCC、識別器をSVMとした識別実験の結果を述べる。

特徴量としてMFCC(12次元)を用い、SVMの学習には自閉症児14人(1121単語)、定型発達児20人(1114単語)を用い、学習には用いていない自閉症児5人(471単語)、定型発達児5人(462単語)で識別実験を行った。

4.2 提案手法

自閉症児、定型発達児それぞれから、openEARにより音響特徴量を抽出し、4.1節で示した音声データと同条件の基、forward selection methodにより最適な特徴量を選択する。その後、forward selection methodに用いていない自閉症児6人(498単語)、定型発達児8人(541単語)で識別実験を行った。

4.3 実験結果

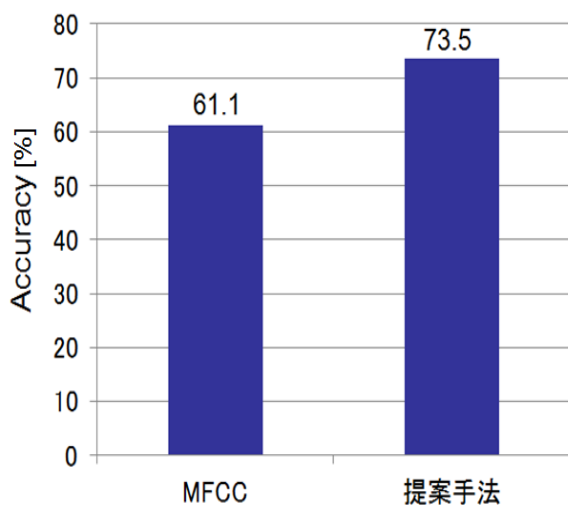


Fig. 2 Classification results

Fig. 2 はベースラインでの識別精度、提案手法での識別精度を示している。図より、ベ

ースラインに比べて、提案手法による自閉症児と定型発達児の識別精度の向上を確認することができた。

5 おわりに

実験結果として、ベースラインとして設定したMFCCによる識別精度が61.1%、本論文における提案手法での識別精度が73.5%という優れた結果を示すことができた。

今後の課題として、更なる認識率の向上を図るために、ランダムプロジェクトン[6]とForward selection methodの組み合わせによる識別実験を行いたい。またmultiple kernel learning[7]に基づく特徴量選択についても検討する。さらには、openEARで抽出された特徴量に関して、どのような特徴が自閉症児と定型発達児の識別に効果的なのかを考察し、今後のこの分野の研究発展に努めていきたい。

参考文献

- [1] 馬場朗 他, “高齢者音響モデルによる大語彙連続音声認識” 電子情報通信学会論文誌, Vol. J85-D-II, No. 3, pp. 390-397, 2002.
- [2] 市川薫 他, “福祉と情報技術,” オーム社, 2006.
- [3] B. Schuller et al., “The INTERSPEECH 2009 Emotion Challenge,” Proc. Interspeech 2009, pp. 312-315, 2009.
- [4] 森 裕一 他, “変数選択に基づく主成分分析,” 計算機統計学, 第11巻, 第1号, pp. 1-12, 1998.
- [5] A. Hassan et al., “Multi-Class and Hierarchical SVMs for Emotion Recognition,” Proc. Interspeech 2010, pp. 2354-2357, 2010.
- [6] T. Takiguchi et al., “Evaluation of Random-Projection-Based Feature Combination on Speech Recognition,” Proc. ICASSP 2010, pp. 2150-2153, 2010.
- [7] R. Takashima et al., “Feature Selection Based on Multiple Kernel Learning for Single-Channel Sound Source Localization Using The Acoustic Transfer Function,” Proc. ICASSP 2011, pp. 2696-2699, 2011.