



音声情報を用いたリアルタイム感情識別システムの検討

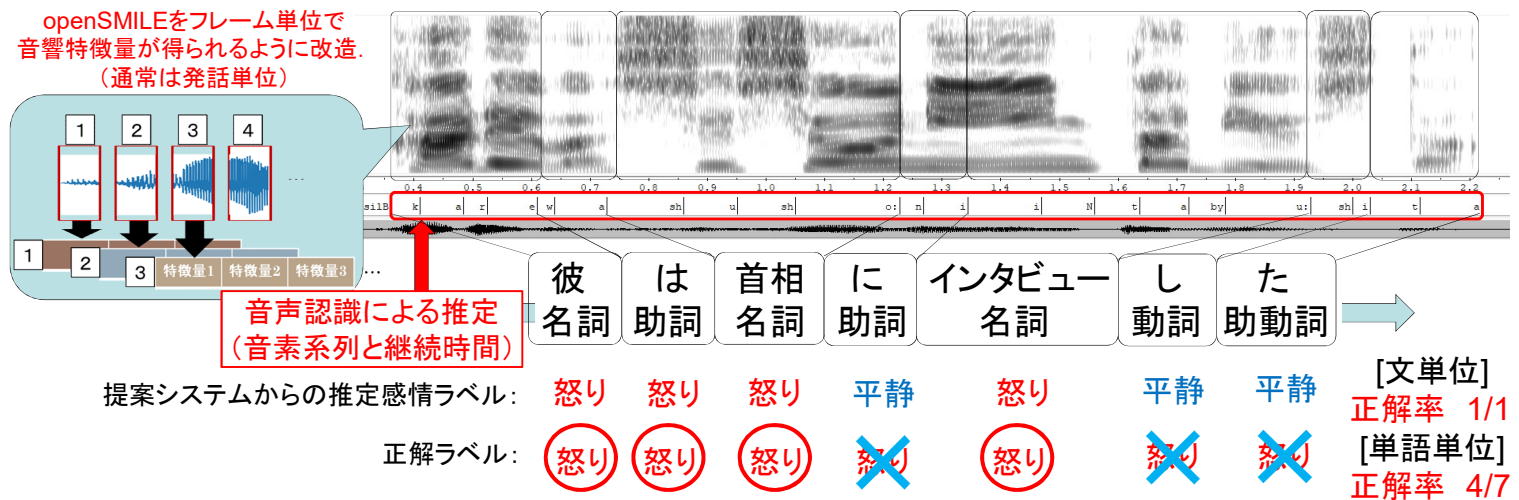
福吉 力, 中山 仁史

広島市立大学 大学院情報科学研究科

1. はじめに

近年、コミュニケーションやミーティングをオンラインで行うことが推奨され、これまでと異なる急激な環境変化にともなう心身の不調や精神的ストレスなどが問題視されている。そこで我々は、遠隔地より参加するユーザの精神的・心的状態を講演者やユーザにフィードバックし、ユーザが精神的・不安を感じる際には休憩や退出を促すためのユーザ支援を目標とする。従来の感情識別では単語、文節などの単位による分析が用いられているが、ヒトの感情は対話中に瞬時的に変化する場合もある。そこで、音声を瞬時的なフレームに分割・分析し、音声の強弱・韻律など様々な音響的变化を時々刻々と推定するリアルタイム感情識別システムを提案する。瞬時フレーム単位で音声処理するため、発話単位として音声情報を残さずに棄却することでプライバシーも担保することができる。

2. リアルタイム感情認識システムの動作例

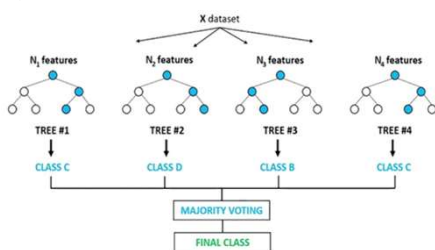


3. 単語単位による文単位の感情識別実験

機械学習による感情識別モデル

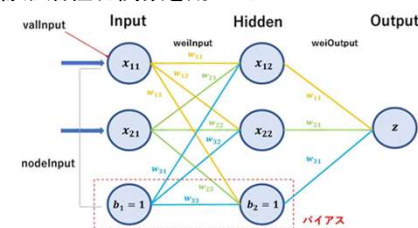
1. Random Forest

弱学習器に決定木を用いて、並列的に学習するバギングタイプのアンサンブル手法。



2. Multi Layer Perceptron Classifier(MLP)

順伝播型ニューラルネットワークの一つ。非線形活性化関数を用いたニューロン。



実験条件

対象コーパス:

日本語共感的音声対話コーパス (STUDIES)

100種類のセリフ

感情: Angry, Happy, Normal, Sadの計4種類

音響特徴量:

openSMILE 3.0emobase(136種類)

開発環境:

Visual Studio Code 1.56.2

Python 3.7.6

単語単位による文単位の感情識別結果

	RandomForest				MLPClassifier			
	Angry	Happy	Normal	Sad	Angry	Happy	Normal	Sad
Angry	1071	43	34	97	1069	41	67	68
Happy	47	1021	76	101	54	1057	79	55
Normal	66	43	1049	87	49	68	1078	50
Sad	65	28	29	1123	53	25	51	1116
	Predicted label				Predicted label			

発話文単位の感情識別結果

Model	Accuracy
Random Forest	86.7 %
MLP Classifier	85.6 %

いずれの方法においても、
85%以上の識別精度を達成

4. おわりに

連続した瞬時フレームを用いて音声の強弱・韻律変化などの音響的变化に注目したリアルタイム感情識別システムを構築した。本研究では、瞬時フレームで分析した単語単位の感情識別結果を用いて文単位での感情識別を評価した。結果として、いずれの方法においても約85%以上の精度を達成した。今後はより細かな音素や音節単位での感情識別の構築と、発話中に感情変化がともなう音声を対象とした検討を行う予定である。