

# 脳計測情報の代用特性としてのマルチモーダル感情推定技術の提案

Emotion estimation method using multi-modal information as substitute for electroencephalogram

大道 博文<sup>†</sup> 見尾 和哉<sup>‡</sup> 段原 優和<sup>‡</sup>

Hirofumi Omichi Kazuya Mio Yuto Danbara

† 広島市立大学 言語音声メディア工学研究室

Language and Speech Research Laboratory, Hiroshima City University

‡ 広島経済大学 石野研究室

Ishino Laboratory, Hiroshima University of Economics

概要 高精度の感性可視化技術として脳計測情報からの感情推定が注目されている。しかし日常生活において常時脳計測装置を装着することは現実的ではない。そこで我々は脳計測情報の代用特性として表情や発話音声を用いた感情推定システムの構築を目指している。本稿では脳情報とマルチモーダル情報の同時計測実験および感情評定手法について説明する。さらに、時系列変化を考慮した感情推定手法として、LSTM(Long short-term memory)を用いた手法についても説明する。

## 1 はじめに

現在広島大学医学部と共同で脳計測情報を用いた感情推定技術[1]の開発を行っているが、日常生活において接触型脳計測装置を装着し続けられない状況も少なからず存在する。そこで我々は、脳計測情報の代用特性として、表情・口調・発話テキストといったマルチモーダル情報を用いた感情推定技術[2]を提案している。そのために、ある感情喚起状況において脳情報の計測とマルチモーダル情報の計測を同時に行ったデータを収集し、それぞれの手法による感情推定結果を比較する必要がある(図1)。各感情喚起状況において被験者に生じた感情は、客観評定ではなく被験者自身による事後評定によりラベリングされる。本発表では、脳情報とマルチモーダル情報の同時計測実験および感情評定手法について説明する。さらに、時系列変化を考慮した感情推定手法として、LSTM(Long short-term memory)を用いた手法についても説明する。

## 2 脳情報とマルチモーダル情報の同時計測データの収集実験

脳計測情報の代用特性としてマルチモーダル情報を用いるためには、同一の感情喚起状況において脳情報とマルチモーダル情報を同時に収集したデータが必要となる。そこで本研究では、情動喚起刺激としてテレビゲームをプレイしてもらい、ゲーム中に感情が喚起した際の脳情報、表情、発話音声を同時に収録する。実験の様子を図2に示す。

以下にデータ収集実験の手順を示す。

1. 4人の被験者(21歳~46歳)に2人1組になってもらい、その中の1人に脳計測装置を装着する
2. 脳計測装置を装着した被験者の前後にビデオカメラを配置し、高指向性マイクをその被験者に向ける

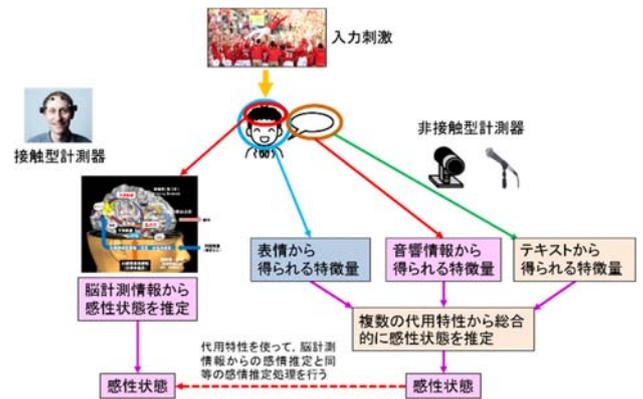


図1: 脳計測情報の代用特性としてのマルチモーダル情報



図2: 情動喚起時のデータ取得風景

3. (1), (2)の状態でゲームソフト「マリオカート」をプレイしてもらい、プレイ中の脳波・表情・発話音声を収録する
4. 1人目の収録が終了したらもう1人の被験者に(1)~(3)の手順を同様に行う
5. 両者の収録終了直後に、収録風景の動画を見ながら感情が喚起した時点の発話とその時の生起感情を自己評定アンケートにより回答してもらう

本実験では情動喚起刺激にテレビゲームを用いているため、IAPSのように刺激と喚起される情動の対応付けが行われているわけではない。そこで被験者自身によって事後評定を行ってもらう。本実験で用いた事後アンケート用紙を図3に示す。事後アンケートについては感情生起中に回答してもらうことが望ましいが、回答作業による感情生起傾向や脳計測情報に乱れる可能性があり、また、データ収集後

発話があった時間	発話文字列	備考
生じた感情の属性		
全く感じない	非常に感じた	
0	100	
快の度合		
不快の度合		
活気がある		
期待感		
わくわく度		
生じた感情の種類と強さ		
全く感じない	非常に感じた	
0	100	
怒り		
嫌悪		
恐れ		
幸福		
悲しみ		
驚き		
その他		
( )		
その他		
( )		
その他		
( )		

図 3: 事後評定用アンケート用紙(VAS 方式)

に自己評定する場合も時間経過による自己評定精度の低下が考えられるため、本実験ではデータ収集直後に収録した動画を提示した。また回答方式としては VAS (Visual Analog Scale) を用いた。

なお今回の収録では、一人当たり 20 件程度、合計 94 件のデータを収集できた。

### 3 時系列変化を考慮した感情推定手法

現在構築しているマルチモーダル感情推定システムでは、表情、音声ともに静的な特徴量を用いて話者感情を推定している (図 4)。それらの特徴量の中に音量、基本周波数、表情推定の尤度の変化の一次導関数や近似直線の傾きを使うことで数値の変化傾向を捉えようとしていたが、より細かい時系列での変化を把握するため、現在 LSTM(Long Short-Term Memory)を用いた感情推定手法 (図 5) を検討している。

LSTM に入力する音響特徴量の変化を計算するため、openSMILE[3]を使って 25msec の窓に対して“音圧”、“基本周波数(F0)”、“自己相関関数から算出される声らしさ”の値を算出する。この窓を 10msec ごとにシフトさせることで、各数値が時間経過に応じて変化していく。

また表情については、従来は各フレームにおける表情尤度 (無、喜、驚、怒、悲) の変化をグラフ化したものに対して、平均、標準偏差、近似直線の傾きなどの静的特徴量を算出して機械学習の入力に用いていた。しかし、Ekman[4]の squelched expression (抑圧された表情) のように顔の部位ごとの感情表出やその程度まで考慮するため、LSTM への入力として OKAOVision[5]によって検出可能な顔部位の特徴点の座標 (左目頭、左目尻、右目頭、右目尻、鼻左、鼻右、口上、口元左、口元右) を用いる予定である。

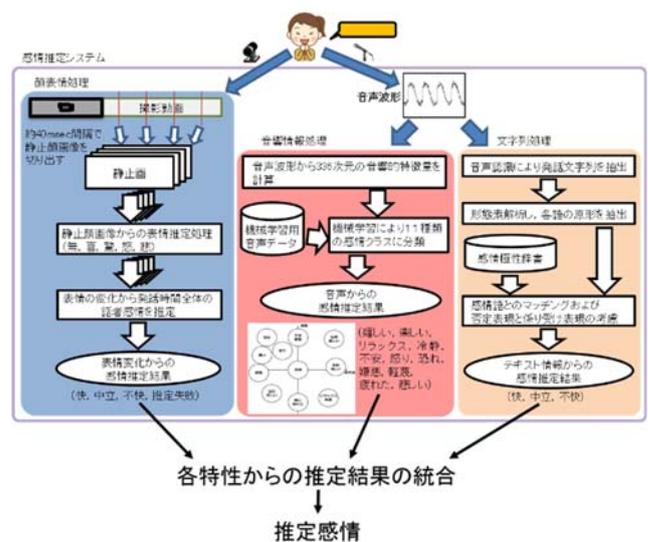


図 4: マルチモーダル感情推定システム

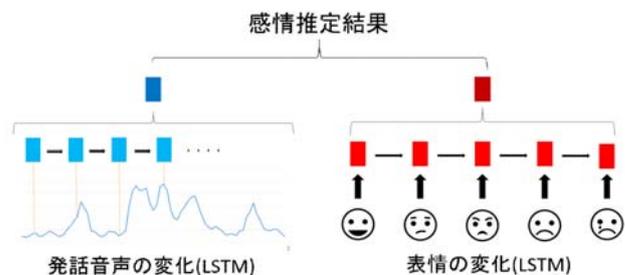


図 5: LSTM を用いた表情特徴量の時系列

### 4 今後の課題

今後は収集したデータを用いて提案するシステムで実験を行う予定である。また、脳計測情報を正解とした感情推定実験も検討している。

### 参考文献

- [1] 町澤昌宏, 脳波振動シグナルを用いた脳情報のリアルタイム可視化に纏わる課題と解決方法:BEI を用いた R&D ツール開発へ向けて, 第 4 回感性脳工学研究会 (2017).
- [2] Uemura, et al., Suppressed Negative-Emotion-Detecting Method by using Transitions in Facial Expressions and Acoustic Features, in Proc. of EDO 2017, pp.122-127 (2017).
- [3] F. Eyben, et al.: openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor, ACM Multimedia Conference – MM, pp.1459-1462 (2010).
- [4] P. Ekman, Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage (Revised Edition). W.W. Norton (2009).
- [5] OKAO Vision | オムロン画像センシングサイト, <http://plus-sensing.omron.co.jp/technology/>, (2018/7/18 アクセス).