

生体信号と顔特徴点分析による複数の映像視聴者情動分類に関する考察  
 A Consideration on Classification of Video Viewers' Emotion  
 by Analyzing Bio-signals and Facial Feature Points

田上 結衣<sup>†</sup>  
Yui Tagami

菅沼 睦<sup>‡</sup>  
Mutsumi Suganuma

亀山 渉<sup>‡</sup>  
Wataru Kameyama

サイモン クリピングデル<sup>\*</sup>  
Simon Clippingdale

## 1. はじめに

近年、動画配信サービスが普及したことで、さまざまな映像コンテンツの量が増え、視聴者が自分の嗜好に合った映像にたどり着くのが困難になっている。これを改善するために、映像視聴履歴などから映像を推薦するコンテンツ推薦システムが存在している[1]。しかし、同じ映像に抱く感情は人それぞれで違うため、本当に個人の嗜好に合った映像を推薦できていないとは限らない。そこで、生体情報や顔特徴点を利用し、視聴者の情動を基にしたより精度の高いコンテンツ推薦システムの開発を本研究では目指している。筆者らの従来研究[2]では、生体信号及び顔特徴点のデータを被験者ごとにスパースコーディングを用いて解析することで、各被験者での情動分類の可能性が示唆されたが、被験者間での特徴の共通性の検討が課題として残っている。そこで本稿では、全被験者の瞳孔径、基礎律動、RRI 変化率、及び顔特徴点データにスパースコーディング施しクラスタリングを行うことで、被験者間での特徴の共通性を検討する。

## 2. 実験

### 2.1 実験概要

被験者 19 名 (男性 15 名, 女性 4 名, 平均年齢 21.5 歳, 標準偏差 1.14) に、驚く、癒される、楽しい、感動する、悲しい、怖い、面白い (2 種類)、退屈の感情が生起されると予測した計 9 個の映像コンテンツを視聴してもらい、その際の瞳孔径、脳波、RRI 変化率、顔特徴点を計測した。各映像の長さは約 5 分であり、各映像の視聴後には約 5 分間の休憩をはさんだ。

ここで、瞳孔径は文献[3]の対光反射補正方式を適用し、輝度による瞳孔径変動を補正している。

### 2.2 実験装置

四つの機器を用いて各データを測定した。瞳孔径測定には Tobii Technology 社製アイトラッカ Tobii X60 (60Hz 計測)、脳波測定には B-Bridge International 社製簡易脳波計測定器 Brain Athlete (512Hz 計測)、RRI 測定には Polar 社製心拍センサ H7 及び同社製スポーツウォッチ V800、顔特徴点測定には Microsoft 社製 Kinect v2 を用いた。

## 3. 解析手法

### 3.1 解析データ

瞳孔径は、まず対光反射補正を施し、1 秒単位で、左右を平均したものを使用した。瞳孔径の大きさは刺激を受け

てから変化するまで遅延があるため[4]、基礎律動と同時刻、1 秒後、2 秒後の三つを瞳孔径のデータとして使用した。

基礎律動は、1 秒ごとの周波数成分を使用し、 $\alpha$  波 (7-12Hz)、 $\beta$  波 (12-30Hz)、 $\gamma$  波 (30-45Hz) の各帯域の周波数成分の全体に対する割合を算出し、「 $\beta$  の割合 /  $\alpha$  の割合」と「 $\gamma$  の割合 /  $\alpha$  の割合」を解析に用いた。

RRI 変化率は、RRI データを線形補間し 1 秒毎の値を求め、「毎秒の RRI / 全データの平均 RRI」とした。

顔特徴点のデータは目、口、鼻に関するデータを用いた。目、口、鼻からそれぞれ 8, 6, 4 点の特徴点を取り出し、各点の x, y 両座標値から被験者ごとに頭部モデルの推定を行った。そして、フレームごとに推定されたモデルと計測されたデータの差分の値を 1 秒毎に平均した値をデータとして用いた[2]。

したがって、生体信号指標が 6 個、顔特徴点指標が 36 個、計 42 指標をデータセットとした。各指標でレンジが異なるので、全ての指標を z-score で正規化した。被験者ごとに正規化したデータセットと全被験者で正規化したデータセットを解析したが、どちらも同様な結果が得られた。各指標の絶対値よりも指標同士の割合が重要であることが理由として考えられる。本稿では被験者ごとに正規化した場合の結果を示す。

### 3.2 解析手法

作成したデータセットを入力信号としてスパースコーディングを行い、得られたスパース係数を k-means 法 ( $k=50$ ) を用いてクラスタリングを行った。スパースコーディングの基底数は 150 とした。なお、スパースコーディングの実装には統計計算環境 R と R のライブラリ SPAMS[5, 6]を用いた。辞書作成には L1/L2 正則化と Lasso を用いた。その際、スパース係数のスパース性が 90%以上となるようパラメータを調節した。実際のスパース性は 90.2%であった。

## 4. 解析結果

### 4.1 クラスタリング結果

被験者 1 のクラスタリング結果を図 1 に示す。縦軸はクラスタ番号、横軸は累計時間である。図中にある縦線はコンテンツの区切りを表しており、左から順に、驚く映像、癒される映像、楽しい映像、感動する映像、悲しい映像、怖い映像、面白い映像、面白い映像、退屈な映像となっている。各被験者、一つの映像からのデータのみで構成されるクラスタを赤、二つの映像からのデータのみで構成されるクラスタを緑、それ以外を黒で表示した。

図 1 のように、赤色のクラスタはすべての被験者に存在し、複数被験者で赤色であるクラスタも多数存在した。この結果から、複数の被験者に類似した特徴があることが考えられる。

<sup>†</sup> 早稲田大学大学院基幹理工学研究所, Waseda Univ.

<sup>‡</sup> 早稲田大学理工学術院, Waseda Univ.

\* NHK 放送技術研究所, NHK

次に、複数被験者で赤色であるクラスタを抽出し、被験者ごとに映像の種類で色付けをした。その結果を表 1、色と映像の対応関係を表 2 に示す。表 1 において、色のついていない部分は、複数の映像からのデータで構成されていた、或いは一つもデータが無かったことを示す。また、赤色のクラスタであってもデータ数が 1~4 個の場合はノイズと判断し、色付けしなかった。

表 1 から、4, 7, 21, 36, 42 のクラスタは、複数の被験者の面白い映像のデータから構成されていることが分かる。また、1, 20, 32, 34 のクラスタにおいても、一部の被験者で同一の映像データから構成されていることが分かる。このことから、特定の感情において、複数の被験者に共通性があることが示唆された。

4.2 スパース係数の分析

表 1 にある 27 個のクラスタに関して、スパース係数の分析を行った。分析方法は、各クラスタ番号に存在するスパース係数を抽出し、対応する基底番号ごとにスパース係数の絶対値の平均値をとり、強く反応している基底を調べた。スパース係数の絶対値の平均値をグラフ化したものの一例を図 2 に示す。

27 個のクラスタのうち、25 個のクラスタにおいて図 2 のように一つの基底のみ値が大きく、他はほとんど反応していなかった。この結果から、複数の被験者で共通して用いられる基底が存在するため、複数の被験者に対し、共通の辞書を用いた情動分類が行える可能性が示唆された。

5. まとめ

全ての被験者のデータにスパースコーディングを施し、得られたスパース係数をクラスタリングすることにより、特定の感情において複数の被験者の特徴に共通性があることが示唆された。また、スパース係数の分析を行った結果、ほとんどのクラスタが一つの基底のみに強く反応していることから、複数の被験者で共通して用いられる基底が存在することが分かった。これらの結果から、一部の被験者間で共通性を見出すことができ、共通の辞書を用いた情動分類が行える可能性が示唆された。

今後の課題としては、映像ラベルでの分類ではなく、被験者が映像視聴時に感じた感情に基づく分類が必要である。

謝辞

本研究の一部は JSPS 科研費 15H01684 の助成を受けたものです。

参考文献

- [1] 北川 和裕, 新井 範子, “情報量規準を用いた協調フィルタリング”, 日本計算機統計学会大会論文集, Vol.19, pp.17-20, May 2005.
- [2] 田上 結衣, 菅沼 睦, 亀山 渉, サイモン クリピングデル, “生体信号と顔特徴点のスパースコーディングによる映像視聴時情動分類に関する一考察”, HCG シンポジウム 2017, C-1-4, Dec. 2017
- [3] 加藤敦士, 菅沼睦, 亀山渉, “ディープラーニングによる瞳孔径の対光反射補正方式に関する検討”, 電子情報通信学会 2017 年総合大会, Mar. 2017
- [4] J.L. Andreassi, “Psychophysiology-human behavior and physiological response”, 5th ed., Mahwah, New Jersey, 2007.
- [5] J. Mairal, F. Bach, J. Ponce, G. Sapiro, “Online dictionary learning for sparse coding”, Proc. of 26th International Conference on Machine Learning, Montreal, Canada, pp.689-696, Jun. 2009

- [6] J. Mairal, F. Bach, J. Ponce, G. Sapiro, “Online learning for matrix factorization and sparse coding”, Journal of Machine Learning Research, Vol.11, pp.19-60, 2010

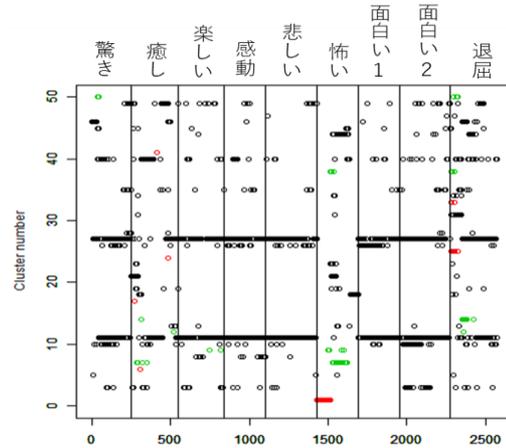


図 1 クラスタリング結果 (被験者 1)

表 1 特徴的な分類が行われたクラスタ

クラスタ番号	被験者																			
	102	103	104	106	107	110	112	113	114	116	118	120	124	125	130	131	132	134	135	
1																				
4																				
6																				
7																				
12																				
16																				
17																				
18																				
19																				
20																				
21																				
22																				
23																				
24																				
25																				
28																				
29																				
30																				
32																				
33																				
34																				
36																				
37																				
39																				
41																				
42																				
43																				

表 2 映像コンテンツと色の対応関係

驚き	癒し	楽しい	感動	悲しい	怖い	面白い1	面白い2	退屈

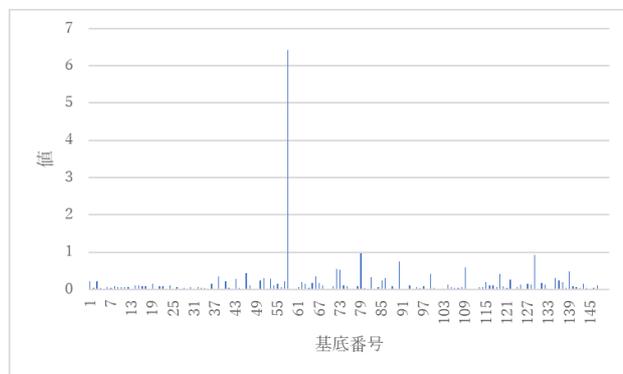


図 2 クラスタ番号 7 に分類されたデータの各基底に対するスパース係数の絶対値の平均値