

e スポーツ実施中の高齢者における顔色情報を用いた感情発生区間推定法の検討

菊地 亮太^{†1} 景山 陽一^{†1} 白井 光^{†1} 石沢 千佳子^{†1}
 末廣 健二^{†2} 高橋 伸明^{†2} 斉藤 洋樹^{†2} 小林 拓也^{†2}
 渡邊 文人^{†3} 佐竹 久美^{†3} 佐藤 直子^{†3}
 秋田大学^{†1} 秋田ケーブルテレビ^{†2} ALL-A^{†3}

1. 背景・目的

現在、日本は世界で最も高齢者の人口比率が高い国である。このため、高齢者の健康寿命延伸が必要とされるが、これを妨げる要因の1つとして、認知症が挙げられる。日本の65歳以上の高齢者における認知症発症の割合は年々増加傾向にあるため、認知症の予防と、認知機能を維持・向上させることが課題となっている。

認知症予防を目的とした研究はこれまでも行われており、例えば、社会との多様な繋がりを持つことは、認知症発症リスクを低減させることが明らかにされている^[1]。さらに、近年ではエレクトロニックスポーツ(以下、eスポーツ)への注目が高まっている。ゲームが認知機能に与える影響について調査した研究は多く存在し、例えば、アクションゲームをプレイした被験者において、視覚的な応答時間が速くなる結果が認められている^[2]。しかしながら、eスポーツを行う高齢者を対象とした研究はまだ少なく、eスポーツが高齢者にどのような影響を与えるのかを理解することは重要であると考えられる。

eスポーツから得られる刺激や感情を定量的に検出することは、対象ゲームをどれだけ楽しめているかを定量的に判断する指標として活用できる。さらに、この指標は高齢者の認知症予防に向けたeスポーツの評価に役立てられると考える。そこで本研究では、eスポーツから得られる刺激や感情を定量的に判断する指標を生成し、eスポーツが高齢者の認知機能に良い影響を及ぼすのかどうか検討することを目的とする。

本研究グループではこれまでに、eスポーツ実施中に感情が発生した区間を推定する手法について検討を行った^[3]。その結果、顔の左頬および右頬における彩度の時系列変化に着目し、その特徴を機械学習を用いて学習することで、感情発生区間を推定できる可能性があることを明らかにした。しかしながら、学習時には左頬と右頬の情報を個々に使用(以下、従来手法)しており、両頬の情報を併用して検討を行うまでには至っていない。そこで本稿では、新たに両頬の情報を有するデータセットを作成し、モデルの精度向上を図ったので報告する。

2. データ取得

実験は2020年に2回、2021年に1回実施した。被験者は60歳以上の方々であり、2020年6名(女性、1名は1回のみ参加)、2021年4名(女性、2名は2020年に参加いただいた被験者と同一)である。実験では、レースゲーム(GRAN TURISMO SPORT^[4])をプレイする様子を、可視カメラ(4Kビデオカメラ HC-VX2M)を用いて撮影した(1920×1080画素;30fps、一部1280×720画素)。ゲームプレイ後には、自身のゲーム中の様子を視聴していただき、感情が発生した区間について評価していただいた。データ取得時の照度を以下に示す。

- 被験者上方の光

2020年: 1,260~1,750 lx, 2021年: 860~1,190 lx

- 被験者正面の光

2020年: 630~960 lx, 2021年: 790~1,150 lx

本研究は「秋田大学手形地区における人を対象とした研究に関する倫理規程第6条第2項」に基き、新型コロナウイルス感染症対策を十分に施したうえで、データを取得している。

3. 提案手法

本稿では、L*a*b*表色系^[5]を用いて算出された頬領域における彩度情報を、Convolutional neural network(以下、CNN^[6])およびLong short-term memory(以下、LSTM^[6])を組み合わせた機械学習モデル(以下、CNN-LSTM)に入力し、彩度変化の特徴を学習させることで、感情発生区間の推定を行った。

3.1 頬領域画像の作成および彩度算出

頬領域画像の作成および彩度算出の流れを図1に示す。はじめに、撮影動画を静止画像に分割する。次に、全ての静止画像を対象とし、オープンソースライブラリのdlib^[7,8]に搭載される顔検出機能を使用し、輪郭や鼻など、顔の位置情報を取得する。さらに、取得した位置情報を基に頬領域を抽出する。この時、頬領域内にほうれい線の影響による陰影が含まれないよう調整を施す。そして、切り取った画像にリサイズ処理を施し、左頬と右頬を結合したものを頬領域画像とする。なお、左頬と右頬の結合

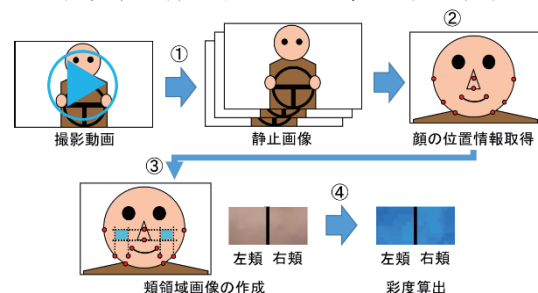


図1 頬領域画像の作成および彩度算出の流れ

A Study on Estimation of Emotional Arousal Intervals Using Facial Color Information of Older Adults during eSports
 Ryota Kikuchi^{†1}, Yoichi Kageyama^{†1}, Hikaru Shirai^{†1},
 Chikako Ishizawa^{†1},
 Kenji Suehiro^{†2}, Nobuaki Takahashi^{†2}, Hiroki Saito^{†2},
 Takuya Kobayashi^{†2},
 Fumito Watanabe^{†3}, Hisami Satake^{†3}, Naoko Sato^{†3}
^{†1}Akita University, ^{†2}Cable Networks Akita, ^{†3}ALL-A

部には、黒画素(2画素)を挿入している。最後に、作成した頬領域画像の各画素における彩度を算出する。

3.2 データセット作成

はじめに、各感情発生区間における頬領域画像群から着目フレームを決定する。次に、着目フレームの前15、後14の合計30フレームの頬領域画像を1つのデータセットとして取得する。この処理を、着目フレームを1フレームずつシフトしながら行い、各感情発生区間におけるデータセットを作成する。また、各データセットには着目フレームにおけるラベルを付与し、「感情発生」ラベルと「感情未発生」ラベルをそれぞれ設定した。

3.3 CNN-LSTMモデルの学習および推定フレームの算出

はじめに、作成したデータセットをCNN-LSTMモデルに入力し、着目フレームにおける正解ラベルを出力するように学習を行う。モデルの構造は32, 64次元の2層からなるCNN層と、中間層のユニット数n個からなるLSTM層に設定する。また、勾配法にはAdamを使用し、学習回数は100回まで検討を行い、最も良好であると判断したものを使用する。次に、学習済みモデルに対してテストデータセットを入力し、着目フレームが感情発生区間である確率が、感情未発生区間である確率よりも大きかった場合、着目フレームを感情が発生した可能性が高いフレーム(以下、推定フレーム)として算出する。なお、本稿では学習データとテストデータが重複しないように被験者をグループ分けしており、学習データに6名、テストデータに2名を使用した合計6つのグループにおいて学習・テストを実施している。また、学習時におけるモデルのバッチサイズ、ならびにLSTM層のユニット数は、 $64 \cdot 128 \cdot 256$ の3種類すべての組み合わせで検討を行い、グループごとに最も良好であると判断した値を使用している。

3.4 感情発生区間の評価方法

本稿では、指定した区間内に占める推定フレームの割合(以下、推定フレーム密集率)を算出し、これを基に感情発生区間の推定を行う。はじめに、撮影動画内から着目フレームを決定する。次に、着目フレームの前15、後14の合計30フレームの区間を指定し、区間内における推定フレーム密集率を算出する。そして、推定フレーム密集率が30%以上であった場合、着目フレームを感情が発生したフレームとして検出する。この処理を1フレームずつシフトしながら行い、検出されたフレームが連続する区間を感情発生区間として推定する。なお、推定フレーム密集率の閾値は10%~90%(10%刻みで変更)までの範囲で検討を行い、最も良好であると判断した値を使用している。

最後に、推定された感情発生区間と、被験者が「感情が発生した」と評価した区間(以下、評価区間)との一致率を算出した。本稿では、一致率が40%以上であった区間を推定に成功したと判断した。

4. 感情発生区間推定結果

グループごとにモデルの学習を行い、学習後のモデルを使用してテストデータにおける感情発生区間推定を行った。感情発生区間の推定成功件数および

推定成功率算出結果を表1に示す。なお、表1における「左頬」および「右頬」は、従来手法^[1]を用いて算出した結果、「両頬」は提案手法にて算出された結果をそれぞれ示している。

提案手法は、6グループ中4グループにおいて推定成功率が増加する結果が得られた。さらに、6グループ全体においても、推定成功率が従来手法より増加する結果が得られた。この結果は、提案手法が感情発生区間推定の精度向上に寄与することを示唆している。従来手法では、左頬と右頬、どちらか一方の情報を使用した際には推定に失敗するが、もう一方の情報を使用した際には推定に成功する区間が存在する場合を認めた。両頬の情報を一度に学習することで、これらの区間が推定可能となり、成功率が向上したと考える。

5. 謝辞

本研究はJSPS科研費JP19K12909, JP22K12215の助成を受けて行われた。また、データ取得実験に参加下さった被験者の皆様に深く感謝申し上げます。

表1 推定成功件数および推定成功率算出結果

グループ	評価区間数	推定成功件数			推定成功率		
		左頬	右頬	両頬	左頬	右頬	両頬
1	28	25	25	27	89%	89%	96%
2	39	34	31	33	87%	79%	85%
3	39	33	28	34	85%	72%	87%
4	36	29	31	32	81%	86%	89%
5	36	29	31	32	81%	86%	89%
6	32	31	32	32	97%	100%	100%
合計	210	181	178	190	86%	85%	90%

■ : 従来手法より増加

参考文献

- [1] T. Saito, C. Murata, M. Saito, T. Takeda and K. Kondo: "Influence of social relationship domains and their combinations on incident dementia: a prospective cohort study", *J Epidemiol Community Health*, Vol. 72, Issue 1, pp. 7-12, DOI:10.1136/jech-2017-209811, 2018
- [2] C. V. Hutchinson, D. J. K. Barrett, A. Nitka and K. Raynes: "Action video game training reduces the Simon Effect", *Psychonomic Bulletin and Review*, Vol. 23, pp. 587-592, 2016
- [3] R. Kikuchi, Y. Kageyama, H. Shirai, C. Ishizawa, K. Suehiro, N. Takahashi, H. Saito, T. Kobayashi, F. Watanabe, H. Satake, N. Sato: "Emotional Arousal Interval Estimation Method for Older Adults Playing eSports by Focusing on Cheek Saturation", *The 10th IIAE International Conference on Intelligent Systems and Image Processing 2023 (ICISIP2023)*, GS2-2, 2023
- [4] "GRAN TURISMO SPORT", POLYPHONY DIGITAL, <https://www.gran-turismo.com/jp/gtsport/top/> (Accessed 2024/01/06)
- [5] 日本色彩学会編: "新編 色彩ハンドブック(第3版)", 東京大学出版会, 2011
- [6] Antonio Gulli: "直観 Deep Learning", オライリージャパン, 2018
- [7] "dlib C++ Library", <http://dlib.net/> (Accessed 2024/01/06)
- [8] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou and M. Pantic: "300 Faces In-the-Wild Challenge: Database and Results", *Image and Vision Computing*, Vol. 47, pp. 3-18, 2016