

機械学習を用いた e スポーツ実施中の高齢者における感情発生区間推定に関する検討

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

近年、超高齢社会の到来に伴い、高齢者の健康寿命延伸が必要とされている。健康寿命の延伸を妨げている要因の1つは認知症であり、認知症の予防や認知機能を維持・向上させることが重要である。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]。また、ゲームと認知機能の関係性について調査した研究も多く存在し、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大幅に減少する結果が認められている^[2]。そこで現在、高齢者の認知機能維持・向上を目的として、エレクトロニック・スポーツ(以下、e スポーツ)を用いる事例が増加している^[3]。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や健康寿命延伸に向けた e スポーツの評価や改善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の生成を目的とする。

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^[4]。その結果、顔の左頬領域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、感情発生区間推定に使用した一部のテストデータにおいて、推定結果が大きく低下している様子が認められた。この原因として、学習に使用した左頬領域画像内にほうれい線による影が含まれるデータが存在することが考えられる。そこで本稿では、影を含まない左頬領域画像を生成し学習に使用することで、影領域が感情発生区間推定に及ぼす影響について検討を加え、モデルの精度向上を図る。

2. データ取得

2020年10月および11月に1回ずつ(1名の被験者は11月のみ)データ取得実験を実施した。被験者6名(60歳以上、女性6名)がレースゲーム(GRAN TURISMO SPORT)^[5]をプレイする様子を、可視カ

A Study on Estimating Emotional Arousal Intervals in the Elderly when Playing eSports Using Machine Learning
 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

メラ(4K ビデオカメラ HC-VX2M)を用いて撮影した(10月:1280×720画素, 11月:1920×1080画素; 各30fps)。また実験時、被験者には心理・体調に関するアンケート調査およびゲームプレイ中における感情発生の有無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: 1,450~1,550 lx
- ・被験者正面の光: 650~750 lx

なお、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」に基づいてデータを取得している。

3. データ解析手法

本稿では、L*a*b*表色系^[6]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^[7])および Long short-term memory(以下、LSTM^[7])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試みる。

3.1 左頬領域画像の作成および彩度算出

左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^[8,9]に搭載されている顔検出機能を使用し、顔の輪郭、鼻および口の位置情報(以下、顔の位置情報)を取得する。さらに、取得した顔の位置情報を基に、顔の左頬領域を切り取る。この時、左頬領域内に影領域が含まれないよう調整を施している。そして、切り取った画像にリサイズ処理を施したものを左頬領域画像とする。本稿で作成した影を含まない左頬領域画像、ならびに従来使用していた影を含む左頬領域画像の例を図1内に示す。最後に、L*a*b*表色系を用いて、作成した左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、撮影動画内にて「感情が発生した」と評価が得られた区間における左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フレーム、後14フレームの合計30フレームの連続

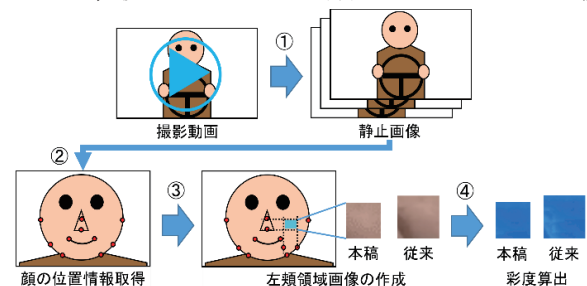


図1 左頬領域画像の作成および彩度算出の流れ

した左頬領域画像を取得し、これらを1つのデータセットとする。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うことで、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設定した。具体的には、被験者から感情発生の評価が得られたフレームには「感情発生」ラベルを、ゲーム開始前および前の感情発生から5秒以上経過し、被験者から何も評価が得られていないフレームには「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感情発生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 128 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、学習済みの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か、感情未発生区間であるかの確率を算出する。そして、感情発生確率が感情未発生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、感情発生推定フレーム)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のグループ分けを行っており、学習データに 4 名分のデータ、テストデータに 2 名分のデータを使用した合計 5 つのグループにおいて、CNN-LSTM モデルの学習と感情発生推定フレームの算出を行っている。

3.4 感情発生区間の評価方法

本稿では、算出した感情発生推定フレームが、指定したフレーム区間内においてどれだけの割合を占めるか(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 フレーム、後 14 フレームの合計 30 フレーム(以下、密集率算出フィルタ)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、密集率算出フィルタは 30, 60, 90 フレームの 3 種類、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っており、最も良好であると判断した値をそれぞれ使用している。

4. 感情発生区間推定結果

推定された感情発生区間が、被験者から「感情が発生した」と評価が得られた区間(以下、評価区間)とどれだけ一致しているかという一致率を算出し、一致率が 40%以上であった区間を感情発生の推定に成功した区間として、その件数と推定成功率を算出した。なお、一致率の閾値は 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。

影を含まない左頬領域画像を学習に使用したものを手法①、影を含む左頬領域画像を学習に使用したものを手法②として、それぞれの推定成功件数および推定成功率を表 1 に示す。グループ 1 においては推定成功件数および成功率が低下したものの、残り 4 グループにおいてはどちらも増加する結果が得られた。さらに、5 グループ全体としては推定成功率が 80%以上まで増加する結果が得られた。この結果は、本手法において影領域を取り除くことは、感情発生区間推定の精度を向上させることに寄与することを示唆している。

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	推定成功件数		推定成功率	
		手法①	手法②	手法①	手法②
1	28	20	26	71%	93%
2	26	26	18	100%	69%
3	26	21	18	81%	69%
4	23	18	17	78%	74%
5	23	20	7	87%	30%
合計	126	105	86	83%	68%

■ : 手法②よりも値が増加した事例

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。

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] "シルバーe スポーツ協会が世界で初めてできました", 特定非営利活動法人(NPO) さいたま市民ネットワーク, [https://nposcn.com/info/世界初!シルバーe スポーツ協会ができました/](https://nposcn.com/info/世界初!シルバーeスポーツ協会ができました/) (Accessed 2023/01/09)
- [4] 菊地 亮太, 景山 陽一, 白井 光, 石沢 千佳子, 末廣 健二, 高橋 伸明, 齊藤 洋樹, 小林 拓也, 渡邊 文人, 佐竹 久美, 佐藤 直子: "高齢者を対象とする e スポーツ実施時の顔色特徴解析ならびに感情発生区間推定手法の検討", 2022 年度 情報処理学会東北支部研究会, 9, 2022
- [5] "GRAN TURISMO SPORT", POLYPHONY DIGITAL, <https://www.gran-turismo.com/jp/gtsport/top/> (Accessed 2023/01/09)
- [6] 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
- [7] Antonio Gulli: "直観 Deep Learning", オライリージャパン, 2018
- [8] "dlib C++ Library", <http://dlib.net/> (Accessed 2023/01/09)
- [9] 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