

HMD から取得可能な目に関する情報に基づく「楽しい」という感情の推定

赤田 真由* 西川 宜利* 磯本 俊弥*† 志築 文太郎‡

概要. 本研究では、VR 空間における目に関する情報の収集を通じて、ユーザの感情推定を行う手法を示す。本手法において、我々は視線の方向ベクトル、瞳孔径、および瞼の開き具合などの VIVE PRO EYE において取得可能な目に関する情報を用いることにより、楽しいかそうでないかの 2 種類を推定した。LightGBM アルゴリズムを用いた推定精度は、AUC が 0.84、Accuracy は 0.78 であった。また、推定した感情をリアルタイムに表示するシステムを開発した。

1 はじめに

人間の感情をインタフェースに適應すると、より豊富なインタラクションをユーザに提供できる。特に、人間の日常的な行動および表情といった無意識的に変化する情報を用いた感情推定は、ユーザビリティ向上の面において有用である。例として、テキスト解析を通してユーザの感情状態を把握し、その感情に基づいたインタラクションの提供が行われており [2]、ユーザのストレスおよび不安の軽減が試みられている。また、マルチプレイヤーオンラインゲームにおいてユーザの感情をカメラを用いて認識し、それに基づいてアバタの表情を制御するシステムも提案されている [6]。この研究のように、カメラを用いて表情から感情を推定する技術は多く研究されている (例: [5])。一方で、カメラを用いて顔全体を捉えることができない状況においては感情の推定が難しい。特に HMD を使用したインタラクションにおいては、表情から感情を推定することが難しいと考えられる。

そこで本稿では、HMD を装着した状態にて取得可能な目に関する情報を用いた感情推定技術を示す。近年、目に関する情報を検出するアイトラッキング技術を搭載した HMD が多く開発されており、外部のカメラなどの追加の機器を必要とせず HMD 単体での動作が可能である。本研究では目に関する情報として、アイトラッカから得られる目の方向、瞳孔の位置および径、ならびに瞼の開き具合を使用する。

我々はこれらの目に関する情報を実験から収集し、収集した情報を用いて機械学習による感情推定を行う。我々は HMD から取得可能な目に関する情報のみから行う感情推定の第一段階として、まず「楽しい」という感情の推定を、先行研究 [3, 7, 1] と同程

度の精度にて実現することを目的としている。本稿では、ユーザが感じた「楽しい」か「そうでない」かの 2 つの感情を、AUC=0.84 にて推定可能である機械学習モデルを作成した。そして、この機械学習モデルを用いた感情推定のアプリケーション例を示す。

2 感情推定システム

本稿では、目に関する情報からユーザの「楽しい」という感情を推定するシステムを示す。本システムを作成するために行った、データ収集の実験、特徴量の計算、および機械学習に関して述べる。

目に関する情報には、ユーザの目が向いている方向として視線の方向ベクトル、アイトラッカから見てユーザの瞳孔がどの位置にあるかを示す瞳孔の位置、ユーザの瞳孔の直径を示す瞳孔径、ユーザの目がどの程度開いているかを示す瞼の開き具合を採用した。これらの情報はいずれも、HTC VIVE PRO EYE [8] を通じて収集可能な情報である。以降、これら 4 つの目に関する情報を、単にデータと呼ぶ。

2.1 データ収集

本稿における感情推定は、実験を通じて得たデータをもとに機械学習を用いて行った。実験は、Unity 上で表示した動画を HMD に映すことによりユーザの感情を誘発し、その際の目に関する情報を収集した。なお、HMD には HTC VIVE PRO EYE を採用した。また、実験システムは Unity を用いて作成した。本実験には、著者らと同じ研究室に所属する大学生および大学院生 10 人が参加した。本実験のタスクとして、ユーザに 2 種類の動画を視聴してもらった。これらの動画は、VIVE PRO EYE を装着した状態で楽しい感情を誘発することを目的としている。動画視聴中、実験協力者が、感情が動いていると感じた時にスペースキーを押すことにより、感情誘発のタイミングを計測した。実験には約 30 分を要した。

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* 筑波大学 情報理工学位プログラム

† 現在、LINE ヤフー株式会社所属

‡ 筑波大学 システム情報系

表 1. 本システムに使用する計 84 個の特徴量.

データ	特徴量	計
視線の方向ベクトル	平均, 標準偏差, 振幅,	24
瞳孔の位置	最大値, 最小値,	40
瞳孔径	尖度, 歪度, 最後の値	8
瞼の開き具合	平均, 標準偏差, 振幅,	12
	最大値, 最小値, 最後の値	

2.2 特徴量計算

特徴量の計算は先行研究 [4] を参考に行った. 以下に示す特徴量の計算には, 取得した 2000 ms 分のデータを使用する. 本システムにて使用する HTC VIVE PRO EYE はデータを 90 Hz にて収集可能であるため, 計 180 個をデータを用いて特徴量の計算を行う. まず, 一般的にアイトラッカから得られる目に関する情報にはノイズが含まれているため, 180 個のデータを, 100 ms (9 個) 毎に平均を取った 20 個のデータに分割した. その後, 20 個のデータに対し, データの収集時の周囲の環境および個人差に対する依存性を少なくするため, 最初のデータを基準とした相対値を計算した. その後, 計算した相対値に対し, 表 1 に示す 84 個の特徴量を計算した.

2.3 機械学習

本稿では, 実験参加者が楽しいと回答した (つまり, ボタンが押下されている) 区間のデータを「楽しい」感情のラベルに紐付け, そうでない (つまり, ボタンが押下されていない) 区間のデータを「そうでない」感情のラベルに紐付け, これらのデータおよびラベルを用いて機械学習を実施した.

まず, 瞳孔径が本来正の値のみを取るにも関わらず, 負の値を取っているサンプルが全体の約 0.1% 存在したため, それらを外れ値としてデータから除外した. その後, 2.2 節に示すように, 2000 ms のデータを使用し特徴量の計算を行った. 2000 ms の区間は, 実験参加者が「楽しい」と感じボタンを押下した時刻を基準に -500–1500 ms の区間とした. また, 実験では, 「そうでない」感情の収集は明示的に行っていなかった. そこで, 本稿ではボタンが押されていない区間を「そうでない」感情として扱った. 結果として, 各クラス間のラベル数が偏ったため, 「そうでない」ラベルの中からランダムに 235 個抽出し, 各ラベル 235 個のデータセットを機械学習に用いた.

実験から収集したラベルおよびデータから計算した特徴量を用いて, 機械学習モデルを作成した. 機械学習のテストは, 全データセットを 4 対 1 の比率に分割した 5 分割交差検証を採用した. 機械学習アルゴリズムには, 決定木の勾配ブースティングアルゴリズムである LightGBM を採用した. 認識精度は, AUC が 0.84, Accuracy が 0.78, F1 Score が 0.77 であった.

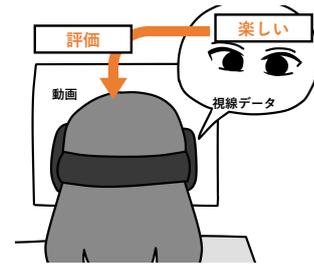


図 1. 目に関する情報から機械学習を行い, ユーザの感情を推定し動画の評価を行うアプリケーション.



図 2. 目に関する情報を用いて機械学習することにより, ユーザの感情を VR アバターへ反映させるアプリケーション.

2.4 アプリケーション

本システムを活用したアプリケーションの例として, 動画の自動評価システム (図 1) を示す. 現状, 動画コンテンツの評価指標としては, 視聴回数, 高評価数およびコメント数などの動画全体に対する指標が一般的である. 対して, 本研究で提案した感情推定技術を用いることにより, 目に関する情報の潜在的な変化から逐次感情を推定できるため, 動画の時間帯毎の評価を自動的に行うことができる. また, 図 2 のように, アバターにユーザの目に関する情報から推定した感情を投影させることにより, より現実に近い感覚のコミュニケーションを可能にすることも検討している.

3 まとめおよび今後の展望

本稿では, HMD から取得した目に関する情報を用いて, 「楽しい」および「そうでない」という感情を機械学習を用いて推定する手法を提案した. 実験の結果, 推定精度は AUC が 0.84 であり, 先行研究と同程度の性能であることを示した.

現在のシステムでは, ユーザの感情を推定し, その推定した感情を HMD 上に表示することができる. 今後は, 作成したモデルを用いたさらなるアプリケーションの開発およびその有用性の調査, 推定できる感情数の増加, ならびに推定精度の向上を試みる.

参考文献

- [1] H. Feng and X. Shen. A Random Forest Algorithm-Based Emotion Recognition Model for Eye Features. In *Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences, ISAIMS '22*, pp. 148–152, New York, NY, USA, 2022. Association for Computing Machinery.
- [2] W. Health. Woebot Health: Relational Agent for Mental Health, 最終閲覧日 2023 年 8 月 30 日. <https://woebothealth.com/>.
- [3] Z. Huang and Y. Li. Machine Learning-Based Study of Eye Features under the Emotion of Anger. In *Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering, EITCE '22*, pp. 1631–1635, New York, NY, USA, 2023. Association for Computing Machinery.
- [4] T. Isomoto, S. Yamanaka, and B. Shizuki. Dwell Selection with ML-Based Intent Prediction Using Only Gaze Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3):1–21, 2022.
- [5] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak. Emotion recognition using facial expressions. *Procedia Computer Science*, 108:1175–1184, 2017. International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland.
- [6] C. Zhan, W. Li, P. Ogunbona, and F. Safaei. Facial Expression Recognition for Multiplayer Online Games. In *Proceedings of the 3rd Australasian Conference on Interactive Entertainment, IE '06*, pp. 52–58, Murdoch, AUS, 2006. Murdoch University.
- [7] L. J. Zheng, J. Mountstephens, and J. Teo. Four-class emotion classification in virtual reality using pupillometry. *Journal of Big Data*, 7(43):1–9, 2020.
- [8] VIVE 日本. VIVE Pro Eye 概要. <https://www.vive.com/jp/product/vive-pro-eye/overview/>. 最終閲覧日 2023 年 8 月 30 日.