



THE VIRTUAL REALITY SOCIETY OF JAPAN

頭部装着型ディスプレイにおける視線を考慮した表情認識の検討

村上雅明¹⁾, 正井克俊¹⁾, 杉浦裕太¹⁾, 杉本麻樹¹⁾

1) 慶應義塾大学大学院理工学研究科 (〒 223-8522 神奈川県横浜市港北区日吉 3-14-1, {mmurakami, masai, sugiura, sugimoto}@imlab.ics.keio.ac.jp)

概要: バーチャル環境において、ユーザーの視線及び表情のデータをアバターに反映させることで、バーチャル環境内のユーザー同士のコミュニケーションを円滑にすることができる。本研究では、反射型光センサを用いて頭部装着型ディスプレイ内部のユーザーの皮膚表面の形状変化を取得して表情を認識する研究において視線が認識精度に与える影響を検討した。視線に応じた訓練データを用いることで表情認識の精度を向上させることができるか実験を行い、左右両端に視線を向けた状態の表情の識別率が視線情報によって向上することを確認した。

キーワード: 表情認識, 視線, Virtual Reality

1. 序論

バーチャル環境において、ユーザーはアバターを表示させることで自身を投影することが出来る。投影したアバターを介して人と人とが互いにコミュニケーションを取る際、ユーザーの表情や視線の情報をアバターに反映させることで、インタラクションの質を向上させることが出来ると考えられる。また、ユーザーの表情と視線を同時に取得することで、ユーザーの興味の対象と対象に対する感情等を推定することが出来ると考えられる。表情認識の手法としては顔画像を用いた手法が一般的であるが、バーチャル環境においては頭部装着型ディスプレイ (以下 HMD) が目とその周辺を隠してしまうため、顔画像による表情認識は困難である。この問題の解決策として、HMD 内部にセンサーを取り付けることによって顔の形状変化を取得する手法がある。ユーザーの顔変形を取得することで表情を認識することが出来るが、同じ表情でも視線方向に応じて顔の形状には差異が出てくるのが予想される。そこで本研究では、HMD 装着下における表情認識の手法として反射型光センサーによる手法を用い、これに対して視線情報を考慮して識別を行うことで、表情認識精度の向上が見られるかどうか検討を行った。本研究においては装着者の表情を 5 つの状態 (Neutral, Smile, Angry, Sad, Surprise) に分け、サポートベクターマシンを用いることによって表情の認識を行った。

2. 関連研究

表情認識の手法としては、画像認識による手法が広く用いられている [1][2]。しかし、これらの手法は HMD 装着下で利用することは非常に難しい。そのため、HMD 内部にセンサーを取り付ける形で表情を認識する手法などを用いる必要がある。そのなかでも反射型光センサーを用いたものは、

直接皮膚に触れる部分がないため HMD の装着感に変化を与えない点、センサーが小型かつ省電力であるため HMD 内部に取り付けやすいという利点がある。反射型光センサーを用いて表情を認識する研究に、正井らの AffectiveWear[3] がある。AffectiveWear では、作成した眼鏡型デバイスの縁部分に計 17 個の反射型光センサーを取り付け、その値を取ることによって 8 種類の表情を識別している。また、AffectiveWear と同様の仕組みを HMD 内部で使用できるようにした鈴木らの AffectiveHMD[4] という研究もあり、こちらでは Oculus Rift Development Kit 2[5] (以下 Oculus DK2) の内部に反射型光センサーを 16 個取り付けることのできる基板を配置することによって HMD 装着者の表情認識を行うことを可能にしている。また、HMD 内部で視線を取得する研究として、pupil-labs[6] の研究が挙げられる。この研究では、赤外光 LED を用いて眼球を照らし、赤外線カメラを用いて眼球の画像を取得、瞳孔の楕円を検出することによって視線の推定を行っている。

3. 理論

3.1 反射型光センサー

反射型光センサーは、赤外光を発する発光部と赤外光を受け取る受光部に分かれており、反射強度を測定することが出来るセンサーである。受け取った光の強さに応じて出力が変化するため、センサーの正面にある物体が反射する光の量を測定することが可能である。正面の物体までの距離や物体の反射係数によって受光部の照度が変化する。皮膚などの反射係数の変化が少ない部位に、このセンサーを用いると距離に応じた反射強度情報を測定することが可能となる。

3.2 顔形状変化の取得

HMD 内部に反射型光センサーを複数取り付けた基板を配置することで、皮膚表面とセンサーの距離を測定するこ

Masaaki MURAKAMI, Katsutoshi MASAI, Yuta SUGIURA, and Maki SUGIMOTO

とが可能となる。

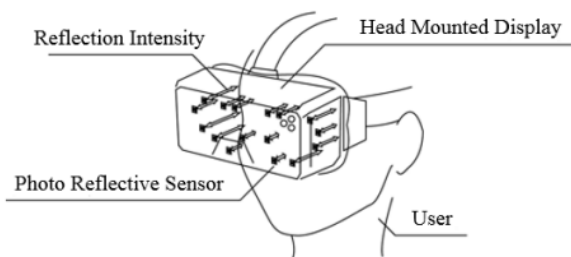


図 1: HMD 内部におけるセンサーを用いた表情識別 [4]

3.3 表情の推定

得られたセンサー値に対してサポートベクターマシンを用いることで表情の推定を行う。サポートベクターマシンは教師あり学習の一種であり、与えられたデータに対し、全てのサンプルからのマージンが最も大きくなる境界を求めることによって分類を行う手法である。

4. 実装

4.1 ハードウェア

本研究では顔形状変化の取得のため、皮膚表面からの反射強度を計測する鈴木らのセンサー配置 [4] を用いた。このデバイスは、16 個の反射型光センサーを用いることでセンサーと皮膚表面との距離の測定を行うことができるデバイスである。使用したセンサーの配置を図 2 に示す。7 と 15 のセンサーに対応する頬の位置には、HMD の底部に左右 1 つずつのセンサーを装着した。このセンサーアレイを視線計測装置を組み込んだ Oculus DK2 の接眼部周辺に配置した。ハードウェアの外観を図 3 に示す。

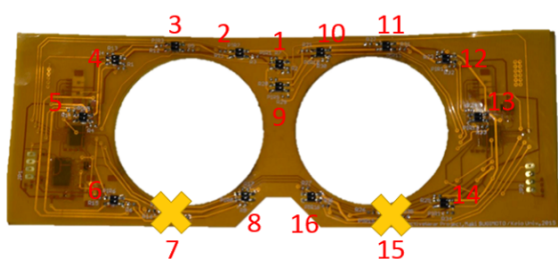


図 2: AffectiveHMD のセンサー配置 [4]

4.2 ソフトウェア

HMD に搭載したマイコンを用いて 16 個のセンサーの値を出力し、これを Serial 通信を用いて PC 側で受けとることによってセンサー値の取得を行った。また、Unity を用いて Oculus DK2 装着者への画面の提示を行った。サポートベクターマシンの実装としては、python のライブラリである scikit-learn を用い、表情の認識を行った。



図 3: ハードウェアの外観

5. 実験

表情認識の精度を評価するため、実験を行った。実験参加者は、1 名 (20 代男性) で、普段は特に表情を意図的に作る訓練を受けていない学生である。実験では、HMD を装着し、HMD 内に提示された注視点 15 点の内、指定された 1 点を視線位置として、音声の指示に従って 5 つの表情をするように指示した。注視点に対する各表情のセンサー値の取得を 1 試行として、これを繰り返し 5 試行のデータを取得した。試行ごとのサンプル数は 240 として、1 人当たり 90,000 のデータ (視線位置 15 点 × 5 表情 × 240 サンプル × 5 試行) を取得した。

表 1: 実験条件

表示点数	15 点 (水平 5 × 垂直 3)
表情セット	5 表情 Neutral, Smile, Angry, Sad, Surprise
サンプル数	240
試行回数	5 試行
総サンプル数	90,000

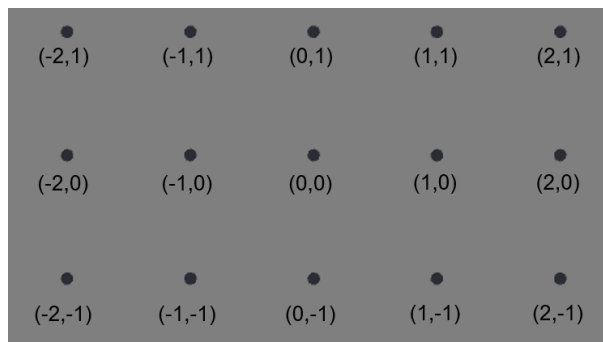


図 4: HMD 内に提示した注視点位置

6. 結果・考察

識別精度を評価するため、取得したデータを試行間で5つに分割して4試行分のデータを訓練データ（視線位置1点×5表情×240サンプル×4試行）として表情毎の識別器を構築し、残り1試行分のデータ（視線位置1点×5表情×240サンプル×1試行）をテストデータとすることで、各視線位置に対する識別精度の検証を行った。試行間の訓練とテストに使用するデータの組み合わせを入れ替えることで、全試行のデータに対する評価を行った。表2に中央を向いた状態のデータを訓練データとした際の各方向における識別精度を示す。識別率の平均は83.5%であった。また、各方向を向いた状態のデータをそれぞれ訓練データとした際の各方向における識別精度を表3に示す。識別率の平均は85.7%であった。尚、各表の識別精度の記載位置は、図4の注視点位置に対応している。

表2: 中央を向いた状態のデータを教師データとした場合の各方向における識別精度

83.3%	87.4%	89.1%	90.7%	83.8%
82.1%	88.2%	89.8%	83.1%	82.8%
70.2%	83.2%	90.6%	72.5%	75.4%

表3: 各方向を向いた状態のデータを教師データとした場合の各方向における識別精度

92.0%	83.4%	87.3%	84.5%	91.0%
90.2%	85.9%	89.8%	88.7%	72.3%
92.7%	80.8%	80.4%	78.4%	90.8%

中央(0,0)のデータのみで学習した表2を見ると、中央付近で精度が良好であるが、左右両端の識別精度は低下している。単一方向のデータのみを訓練データとして用いた場合、視線位置が大きく変化した場合には、表情識別の精度が低下することが確認できる。また、表2と各視点位置のデータを訓練データとして用いた表3を比較すると、中央付近の結果には大きな差異が見られないが、左右両端を見た際の識別率は向上していることが分かる。同じ視線位置の訓練データを持つことで、識別精度の向上を図ることができる。

全方向のデータから等しい数の訓練データを抜き出し、総サンプル数を先の評価と揃えて訓練データ（視線位置15点×5表情×16サンプル×4試行）とした場合の各方向における識別率は平均して90.2%となった。表4に視線位置ごとの識別率を示す。

表2で示したように視線位置の変化が小さい場合、完全に同一の視線位置のデータを訓練データとしなくても一定の識別精度が得られる。また、表情を作ることに熟練していない装着者が意図的に同一の表情をした際にも、幾何形

状にはばらつきが大きいいため、本実験で取得した反射型光センサーからのデータは、試行内で同一の視線位置・表情において取得している240サンプルにおいてはばらつきが小さく、試行間においてはばらつきが大きい。

このため、同一の視点位置・表情におけるサンプル数が少なくても、周囲の視線位置の訓練データが貢献すること、多くの試行の幾何形状のばらつきに対応した学習が行えることによって、各視線位置の表情の識別精度が向上し、全体としても良好な平均識別率が得られたと考えられる。

表4: 全方向のデータを教師データとした場合の各方向における識別精度

90.6%	91.3%	92.0%	90.6%	92.1%
86.6%	92.2%	87.5%	90.2%	92.7%
90.3%	92.9%	94.6%	88.1%	82.1%

全方向のデータを教師データとして用いることで、視線位置が変化した場合でも精度を大きく落とすことのない学習データを作成することが出来た。しかし、中央と左右両端等の一部視線位置においては、各視線ごとに学習データを作成した方が識別精度が高くなるという現象が確認された。これは著しく視線位置が変化した場合に、同じ表情でもセンサー値に変化が現れることを示していると考えられ、これにより著しく視線位置が変化する場合には新たなデータセットを用意して学習を行えばよいと考えられる。

7. 今後の課題

本研究では視線情報を既知の値として用いたが、pupil-labsのデバイス[7]等を用いることによって視線情報もHMD内部で実際に取得する必要があると考えられる。

また、本研究では15視点に対してそれぞれ学習とテストが行えるようにデータセットの用意を行ったが、学習に用いる点の数を減らしても性能が保証される範囲があると考えられ、視点数と精度のトレードオフを検討していく必要がある。また、本実験で用いた表示点を注視した状態のみでなく、表示点と表示点の間の位置に視線を向けた状態における認識精度が向上するかどうか検討することにより、本システム全体の有効性を再確認する必要があると考えられる。

8. 結論

反射型光センサーを用いてHMD内部での表情を識別する手法において、視線の影響を検討するため、視線位置を変化させて学習データを取得した。本研究においては、実験を通じて単一の視線位置の訓練データを用いた場合、視線位置の変化が大きいテストデータにおいて識別精度が低下することが確認できた。また、視線位置に応じた訓練データを用いる場合、左右両端に視線を向けている状態においても、精度が向上することが確認できた。

謝辞

本研究の一部は、科学研究費補助金 16H05870 および JST CREST 課題番号 JPMJCR14E1 の助成を受けたものである。

参考文献

- [1] Gwen Littlewort, Ian Fasel, Marian Stewart Bartlett, Javier R. Movellan: Fully automatic coding of basic expressions from video, Tech report 2002.03, MPLab, University of California, San Diego, CA
- [2] Ying-li Tian, Takeo Kanade, Jeffrey F. Cohn: Recognizing Action Units for Facial Expression Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 23 Issue 2, February 2001, Pages 97-115
- [3] 正井 克俊, 杉浦 裕太, 尾形 正泰, クンツェ カイ, 稲見 昌彦, 杉本 麻樹: AffectiveWear: 装着者の日常的な表情を認識する眼鏡型装置, 日本バーチャルリアリティ学会論文誌, 2016, 21 巻, 2 号, p. 385-394.
- [4] 鈴木 克洋, 中村 文彦, 大塚 慈雨, 正井 克俊, 伊藤 勇太, 杉浦 裕太, 杉本 麻樹: AffectiveHMD: 組み込み型光センサを用いた表情認識とバーチャルアバターへの表情マッピング, 日本バーチャルリアリティ学会論文誌, 2017, 22 巻, 3 号, p. 379-389.
- [5] LLC Oculus VR: Oculus rift — oculus. <https://www.oculus.com/rift/>.
- [6] Moritz Kassner, William Patera, Andreas Bulling: Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction, In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pages 11511160. ACM, 2014.
- [7] Pupil Labs : Pupil labs. <https://pupil-labs.com/>.