



# フェイストラッキングを用いた アバターの親近感を高める手法に関する研究

Research on Method for Increasing Familiarity of Avatar using Face-Tracking

船木烈<sup>1)</sup>, 物部寛太郎<sup>2)</sup>

Retsu FUNAKI and Kantaro MONOBE

- 1) 東北学院大学大学院 電子工学専攻 (〒985-8537 宮城県多賀城市中央 1-13-1, s2194305@g.tohoku-gakuin.ac.jp)  
2) 東北学院大学 工学部 情報基盤工学科 (〒985-8537 宮城県多賀城市中央 1-13-1, monobe@mail.tohoku-gakuin.ac.jp)

概要: 現在, 新型コロナウイルスの感染拡大を防ぐために, Zoom や Teams などのオンラインツールを用いた遠隔授業やテレワークが行われている. そこでは, 自分の顔の代わりに, アバターを映して参加することがある. 本研究では, Unity を用いて, フェイストラッキング機能を実装し, 顔の動きをリアルタイムでアバターに投影した. 著者らが開発した, 音声や時間経過に応じてアバターの表情を動かす従来手法と, 本研究のフェイストラッキング手法を比較することで, オンラインチャットにおける円滑なコミュニケーションに必要な親近感の向上を目指した.

キーワード: アバター, 親近感, フェイストラッキング, オンラインコミュニケーション

## 1. はじめに

現在, 3次元の仮想空間内のサービスであるメタバースをはじめとして, アプリやゲーム, 動画配信などで, 自らの分身となるキャラクター「アバター」の重要性が急速に高まっている. 一方, 2020年から現在に至るまで, 新型コロナウイルス COVID-19による感染症が拡大している. そのため, 人間同士の接触を避けるために, Zoom や Google Meet などのツールを用いたオンライン授業が展開されている. また, 授業に限らず, 企業でも, テレワークが推奨されている. さらに, オンライン飲み会, オンライン卒業式など, イベントもオンラインで行う形が主流となっている. このオンライン授業やテレワークなどにおいて, 自分の顔を映す代わりに, アバターで参加することがある. このように, アバターの用途は様々であるが, 本研究では, Zoom などのオンラインチャットでのアバターの使用を想定し, 表情の操作に比重をおくことで, 親近感を高めるシステムを開発した. アバターを用いた交流においては, そのアバターの親近感が, 交流の場の雰囲気を作るため, 親近感重要なパラメータとなる. アバターは, カスタマイズ性の高さが特徴で, 外見を自分好みに設定することができる他, 様々な動きを追加できるため, 相手に与える印象や親近感に変化をつけやすい. そこで著者らは, 非言語情報を用いて, まばたき, うなずき, リップシンク, 呼吸の4つの動きを2Dアバターに追加することで, アバターの親近感を向上させることを目的としたシステムの開発

を行った[1].

本研究では, 著者ら[1]の先行研究で開発した4つの機能を3Dアバターに実装する. 3Dを選んだ理由は, 3Dよりも多彩かつ繊細な動きが可能であり, 2Dよりも豊かな動きや表現を可能とするからである. また, それにより今後の機能の拡張性も高まる. さらに, フェイストラッキングを用いて, 3Dアバターを動かすシステムも開発した. 2つのシステムを比較し, 親近感を高めるための手法を模索することが本研究の目的である.

本研究は, Zoom でのアバター使用を想定して行ったが, Zoom では, アバターに機能の追加はできない. そこで, Unity と 3D アバターを用いて, システムを開発した.

## 2. 関連研究

### 2.1 非言語情報を用いたアバターの親近感を高める手法に関する研究

著者ら[1]の先行研究であり, 2Dアバターに, まばたき機能, うなずき機能, リップシンク機能, 呼吸機能を実装し, 親近感の調査を行った. 本研究においても, 上記の4つの機能を実装し, フェイストラッキング手法との比較を行う.

### 2.2 自己顔表情同期アバタによる心理的影響の調査

倉橋ら[2]は, 日本人の平均顔のアバタと, 被験者の顔から生成したアバタの心理的影響, および被験者とアバタの表情の同期の有無による心理的影響の比較を行っている.

この研究における心理的影響の定義と、本研究の親近感の定義は似ているため、参考にした。

### 2.3 遠隔対話者の身体動作の提示による音声コミュニケーションの円滑化

尾上ら[3]は、音声コミュニケーションにおいて、視覚情報を身体動作と外見に分け、それらの有無が及ぼす影響を、発話速度や途切れの回数から計測している。この研究では、アバターを使用することで、外見情報は取り入れず、身体動作として、話者の頭部と口の動きを取り入れたシステムを開発している。その結果、身体情報のみを取り入れたアバターを使用することで、発話速度の上昇や途切れ回数の減少が見られ、コミュニケーションを円滑にできることが確認された。本研究においても円滑なコミュニケーションを可能とするために、親近感の向上を目指しているため、この結果を参考にした。

## 3. システムの開発

### 3.1 システムの概要

本研究は、Zoom でのアバターの使用を想定している。開発には、Unity Technologies 社が開発、提供を行っているゲームエンジン Unity[4]を用いた。3D アバターは、ニコニコサービスが無料提供している、ニコニコ立体公式マスコットキャラクター「アリス・ソリッドちゃん」[5]を導入している。導入したアバターの制御は、Unity のスクリプトで行っている。

本研究では、「従来手法システム」と「フェイストラッキングシステム」の2つのシステムを開発した。従来手法システムでは、音声の取得、調節のためのソフト「Voicemeeter」[6]を、フェイストラッキングシステムでは、使用者の表情撮影のため、PC 内蔵カメラを使用する。両方のシステムに対して、モジュール「Unity Capture」[7]を導入し、Unity の画面を録画したものをライブ配信ソフト「OBS Studio」[8]に流し、その画面を Zoom で出力する手法をとっている。これにより、高画質での出力が可能となり、アバターの動きを細部まで相手に見せることが可能となる。

### 3.2 従来手法システム

著者ら[1]の先行研究で実装した、まばたき機能、うなずき機能、リップシンク機能、呼吸機能を従来手法システムとして実装した。基本的な動きは先行研究と同じであるが、2D アバターで実装したものを、本研究では 3D アバターに実装しているため、パラメータの調節や、スクリプトの内容に修正を加えている。実験で得られた意見から、うなずき機能とリップシンク機能は、手動で ON・OFF の切り替えができるように設定した。まばたき機能は、平均 5 秒はそのままに、幅を 2~9 秒に設定した。うなずき機能は、瞳孔を首と真逆に動かすことで、視線が下を向かないように調節した。呼吸機能は、髪の毛のパーツやアクセサリが細分化されているため、その全てに連動させたいので、動きの量やベクトルをパーツごとに調整している。

### 3.3 フェイストラッキングシステム

本システムでは、パソコン内蔵のカメラで表情を撮影し、その表情の動きに合わせて、アバターの表情も連動して動く機能を実装した。撮影された表情からは、図 1 のように、全体の輪郭と、目の開閉、瞳孔の向き、口の開閉および開いたときの形をデータとして取得できる。このデータを、アバターの表情に投影し、同期させることで、使用者の表情に合わせて、アバターの表情も動く。撮影には、PC 内蔵カメラを使用し、表情データの取得のために、2つのモジュール「DlibFaceLandmarkDetector」[9]と「OpenCV for Unity」[10]を、さらにアバターの表情に投影、同期させるためのモジュールとして、「CVV TuberExample」[11]を取り入れた。従来手法システムでは、時間設定や音声取得を必要としたまばたき、うなずき、リップシンクを、全てフェイストラッキングによる表情データを入力情報とすることで実現できる。本研究は、表情のみのトラッキングであり、体のトラッキングは行っていない。そのため、上半身を動かす必要のある呼吸機能は、従来手法[1]で実装した。

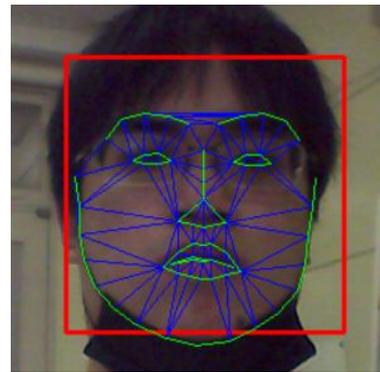


図 1: フェイストラッキングによる表情の取得

## 4. 実験

### 4.1 実験方法

本研究では、東北学院大学の、20 代の学生 7 名を対象に、実験とアンケート調査を行った。

実験環境として、Zoom にて、7 人の被験者を招集し、そこに著者を加えた 8 名でオンラインチャットに臨んだ。著者および、被験者のうちの 4 人は大学の研究室から参加し、被験者の残り 3 人は自宅から参加した。

### 4.2 実験手順

従来手法システム、フェイストラッキングシステムともに、約 10 分のミーティングを行った。そこで、本研究で開発したアバターを、著者の顔の代わりに表示した。ミーティングの内容は自由議論であるが、その中で、機能説明のための時間と、従来手法システムに関しては、うなずき機能とリップシンク機能の確認のため、話者を一時的に限定する時間を設けた。実験中および実験後に、アンケートに回答してもらった。実験における、Zoom の画面をキャプチャしたものを図 2 に示す。



図 2: Zoom を用いた実験の様子 (ギャラリービュー時)  
(プライバシー保護の観点から、被験者の氏名や顔の表示部分にモザイク処理を施している)

#### 4.3 アンケート設問内容

アンケートの設問として、アバターの使用経験などの被験者の情報に加えて、システムのアバターが適切かどうか、2つのシステムを比較して親近感を感じたかどうかなどについて、5段階で評価してもらった。また、フェイストラッキングシステムの全体的な精度も5段階評価してもらった。最後に、本研究の良かった点、改善点、全体を通しての感想を記述式で回答してもらった。

#### 4.4 アンケート結果

アンケート結果より、アバターの使用経験は被験者の28.6%であり、InstagramやLINEなど、SNSでの使用経験があるといった回答が多かった。

アバターの適切度については、図3に示すように平均4.42の評価であった。

従来手法システムの4つの機能について、5段階評価の平均は図4に示すようにまばたき機能:4.42、うなずき機能:3.57、リップシンク機能:3.57、呼吸機能:4.00であった。

フェイストラッキングシステムについては、5段階評価の平均は図4に示すようにまばたき機能:4.85、首の動きに対する評価:4.71、リップシンク機能:4.71、呼吸機能:3.71であった。

フェイストラッキングシステムの全体的な精度の評価は、図5に示すように平均3.85であった。

全体の感想としては、「親近感は得られた」、「顔出ししたくないときは便利なシステム」「フェイストラッキングの精度が高い」といった回答を得た。一方で、「表情の動きを追加してほしい」という意見も多数あった。

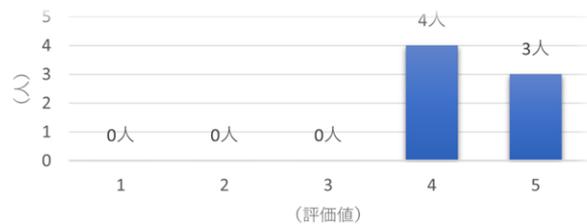


図 3: アバターの適切度

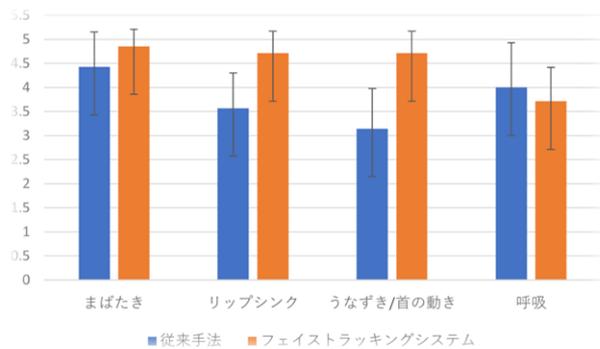


図 4: 4つの機能の比較評価

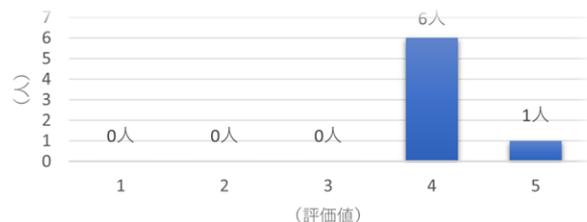


図 5: フェイストラッキングシステムの精度の評価

## 5. 考察

アンケート結果より、まず、アバターの適切度については、今回の実験では平均4.42の評価であった。「髪型がショートカットのほうが好み」という意見があり、個人の嗜好が結果に表れたと見ることもできる。アバターには多くの種類があり、自分で好きなモデルを選ぶことができる、あるいは自作することができるというメリットがある。しかし、使用する場面に応じて、ある程度は規律に則ったアバターのモデルを選択しなければならない。

まばたき機能は、両方のシステムにおいて、5段階評価の平均点は高かったが、より高かったのはフェイストラッキング側であった。ただし、使用者が眼鏡着用の場合、視線をうまく捉えることができないという結果が出た。そのため、上記の例も含め、視線導入が難しい場合は、従来手法システム側の手法に切り替える必要があると感じた。従来手法システム側の改善点としては、頻度やまばたき速度の調整が挙げられる。

リップシンク機能については、従来手法システム側は平

均 3.57 であった。精度については特に指摘がなかったが、音声取得の際は、閾値や利得値の設定を、細かく行う必要がある。一方、フェイストラッキングシステム側は平均 4.71 の高評価であった。フェイストラッキングを利用すれば、音声取得の設定の問題、口の形の種類に関する問題をすべて解決できるので、今回の結果も踏まえると、リップシンクは、フェイストラッキング手法を用いた方が、親近感が高まると考えられる。

従来手法システムのうちき機能の評価は低かったが、フェイストラッキングシステムによる首の動きの評価は高く、首はフェイストラッキングで動かした方が親近感が高まると考える。従来手法システム側の改善点として、動きを大きくする、ボイスやエフェクトの追加などの手法がある。精度については低いといった指摘はなかった。

呼吸機能については、従来手法システムの方が 5 段階評価では上回った。フェイストラッキングシステムでは、「オーバーに見える」という意見があった。

フェイストラッキングシステムの全体的な精度評価は平均 3.85 であり、自由記述では「高い精度だった」との意見があった。しかし、フェイストラッキングシステムの作動中は、Unity のフレームレートが大幅に低下し、処理落ちによる速度の低下が見受けられた。このことから、さらに精度を高めるうえでは、機能を入れすぎない、すでにある手法で簡略化できるところは簡略化する、アバターの表示方法を変えるなどの工夫が必要と考える。

自由記述の回答では、本研究では実験の対象とはしなかったが、キー操作によるアバターの腕の動きを示したところ、自由記述で高評価が得られ、追加で手の動きの導入を求める意見が多く得られた。腕の動きについては、キー操作で動かす手法や、ハンドトラッキングを導入する手法などがあり、導入の上では様々な手法から選択できる。また、腕だけでなく、エフェクトやボイス、視線の動きに対応した機能の追加導入を検討する必要もあると考える。

全体的な評価としては、リップシンクと首の動きはフェイストラッキングシステムの方が評価は高く、フェイストラッキングを用いて、使用者の表情データを入力情報として取り入れた方が、親近感が高まると考える。一方、まばたき機能はフェイストラッキング、従来手法の両方を取り入れると、幅広い状況に対応できると考える。また、首の動きではなく、ボイスやエフェクトによる表現を用いたうなずき機能であれば、従来手法のうちき機能とフェイストラッキングによる首の動きの両立ができると考える。呼吸機能はフェイストラッキングと従来手法との相性が悪く、フェイストラッキングシステムで導入するならば、動きの内容や変化量の調整が必要と考える。

## 6. まとめ

本研究では、Unity で親近感を高める動きや機能を搭載したアバターを開発し、Zoom に出力するシステムを開発した。本研究では、2つのシステムを開発し、従来手法とフェイストラッキング手法のどちらが良いかを比較した。現在は、メタバースやアバターに関する研究や開発が急速に進んでいる。そのため、フェイストラッキングシステムをもとに、新機能を開発、導入して、独自性を出す必要があると感じた。また、本研究の実験を踏まえて、従来手法の方が親近感を感じやすい機能もあったため、フェイストラッキングだけでなく、従来手法も柔軟に取り入れた方が、より親近感を高めることができると考える。

## 参考文献

- [1] 船木烈, 物部寛太郎:非言語情報を用いたアバターの親近感を高める手法に関する研究, 第 26 回日本バーチャルリアリティ学会大会論文集, 2021.
- [2] 倉橋鉄平, 小川奈美, 鳴海拓志, 葛岡英明:顔表情同期アバタによる心理的影響の調査, 第 25 回日本バーチャルリアリティ学会大会論文集, 2020.
- [3] 尾上聡, 山本健太, 田中一晶, 中西英之: 遠隔対話者の身体動作の提示による音声コミュニケーションの円滑化, 情報処理学会論文誌, Vol.54, No.4, pp.1462-1469, 2013.
- [4] Unity, <https://unity.com/ja> (入手 2022.7.26)
- [5] ニコニ立体マスコットキャラクター アリシア・ソリッド 公式サイト <https://3d.nicovideo.jp/alicia/> ©DWANGO Co., Ltd. (入手 2022.7.26)
- [6] VB-AUDIO Software/Audio Apps/Voicemeeter <https://vb-audio.com/Voicemeeter/> (入手 2022.7.26)
- [7] schellingb/UnityCapture <https://github.com/schellingb/UnityCapture> (入手 2022.7.26)
- [8] OBS Open Broadcaster Software® <https://obsproject.com/ja/download> (入手 2022.7.26)
- [9] Unity Asset Store/ DlibFaceLandmarkDetector <https://assetstore.unity.com/packages/tools/integration/dlib-face-landmark-detector-64314> (入手 2022.7.26)
- [10] Unity Asset Store/ OpenCV for Unity <https://assetstore.unity.com/packages/tools/integration/opencv-for-unity-21088> (入手 2022.7.26)
- [11] Unity Asset Store/ CVVTuberExample <https://assetstore.unity.com/packages/templates/tutorials/cvvtuber-example-118186> (入手 2022.7.26)