



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

Method for Increasing Familiarity of Avatar using Non-Verbal Information

船木烈¹⁾, 物部寛太郎²⁾

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 Google Meet 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. ここでは, 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. アバターの親近感を高めるためには, 非言語情報が重要となる. そこで本研究では, 非言語情報として, まばたき, うなずき, リップシンク, 呼吸の 4 つの動きを, Unity を用いてアバターに実装し, 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, 非言語情報, オンラインコミュニケーション

1. はじめに

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 などのオンラインツールを用いたオンライン授業が展開され, 現在の授業体制の基本となっている. 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. アバターとは, ゲームやアプリなどで使う, 自分の分身となる仮想キャラクターである. 本研究では, そのアバターに着目した. 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. これを利用し, アバターをカスタマイズして, 自分だけでなく, 他人にも受け入れられる, 親近感を感じられるアバターの作成を試みた.

アバターの親近感を高めるには, 非言語情報が重要である. 本研究では, 非言語情報のうち, まばたき, うなずき, リップシンク, 呼吸の 4 つの動きを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, 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を補助することを目的とする.

本研究は, Zoom でのアバター使用を想定して行ったが,

Zoom ではアバターの作成はできない. そこで, Unity と Live2D モデルを用いて, アバターを作成した. また, 音声取得のためのソフトとして, Voicemeeter を使い, Zoom への出力はライブ配信ソフトである, OBS Studio を介して行った.

2. 関連研究

2.1 相づちのリアルタイム性の分析と自動生成アルゴリズムの設計

この研究は, 人間の, 発話に対する相づちの特徴を分析し, そこから得た知見をもとに, 相づちの自動生成アルゴリズムを構築するというものである[1]. この論文では, 一定時間の発話のあと, 一定時間の沈黙を検知するとうなずき, 音響的手法を提案しているが, 本研究ではその音響的手法の再現を目指した. また, うなずき挿入頻度のデータも参考とし, うなずき頻度に補正をかけている.

2.2 キャラクターを使用した自動リップシンクシステムの開発

この研究は, 音声閾値を設定し, それを超える音声を検出すると, 画像切り替えによって, リップシンクするシステムの開発を目的としたものである[2]. 本研究においても, この研究同様に, 音声閾値を定め, それを上回る音声の検出時に, リップシンクを行う設定としている.

3. システムの開発

3.1 システムの概要

本研究では、Zoom でのアバターの使用を想定している。アバターには、まばたき機能、うなずき機能、リップシンク機能、呼吸機能の 4 つの機能を Unity にて実装し、Zoom に表示させることを目標としている。アバターのモデルは、Live2D モデルの無料配布サイトより、「響」を選び、使用した。

Live2D とは、アバターの 2D モデルの一つで、モデルの画像データは、目や口、髪の毛、手などのパーツを分けて保存されている。そのパーツ単体、あるいは数か所を、回転、変形などの動きをさせることで、平面の 2D でありながら、立体の 3D に近い動きをさせることができるのが特徴である。また、完全自作の 2D モデルが作りやすいのも特徴である。この Live2D モデルを Unity で動かすには、Unity の追加モジュールである、Cubism SDK for Unity が必要なので、ダウンロードとインポートしておく。

また、うなずき機能とリップシンク機能での音声取得のためのソフトとして、Voicemeeter、Zoom への高画質での出力のためのソフトとして、OBS Studio を使用した。しかし、追加モジュールである、Unity Capture が必要なので、これもダウンロードとインポートしておく。

システム全体の概要を、図 1 に示す。

3.2 まばたき機能

この機能では、時間経過に応じて、アバターがまばたきの動作を行う。東川らの研究[3]では、作業中の瞬目回数を記録し、その結果から疲労蓄積の度合いを導き、適切な休憩のタイミングを、提示するというものであるが、その結果、人間の平均的なまばたき間隔は 5 秒に 1 回というデータを得ていた。このデータを参考にしたが、ただ 5 秒に 1 回にするのではなく、乱数を用いて、まばたきの間隔に幅を持たせるほうが親近感は出ると判断した。そこで、平均 5 秒、幅は 3~7 秒とし、その間に 1 回まばたくように設計した。



図 1: システム全体の概要図

3.3 うなずき機能

この機能では、相手の音声を検知したときに、アバターがうなずきの動作を行う。香曾我部らの研究[1]では、一定時間の発話のあと、一定時間の沈黙を検知するとうなずき、音響的手法を提案している。本研究では、その音響的手法を再現するため、相手の 2 秒以上の発話のあと、2 秒以上の沈黙を検知するとうなずきシステムを開発した。しかし、香曾我部らの研究[1]によると、うなずき挿入頻度のデータから、うなずき頻度は 5~6 秒に 1 回が多く、それより短い間隔ではほとんどうなずかない。そこで、短い間隔でのうなずきを避けるため、うなずいてから 5 秒間は、再びうなずかないように設定した。なお、香曾我部らの研究[1]で用いている、音声認識アプリは、本研究では導入していない。

3.4 リップシンク機能

この機能では、マイクを介して入力した自分の音声を検知したときに、アバターの口が動き、リップシンクを行う。物部らの研究[2]に基づき、音声閾値を設定し、それを超える自分の音声を検知すると、アバターの口が開閉するシステムを開発した。

なお、3.3 と 3.4 では、音声閾値をプログラム内で設定しているが、それとは別に音声利得値の設定があり、これは Unity と Voicemeeter で設定できる。この利得値は、アバターの動作中も自由に設定を変えることができる。

3.5 呼吸機能

この機能では、アバターが呼吸の動作によって、一定のリズムで、体の揺れを表現する。アバターは静止させず、わずかに上下に動かすように設定した。これにより、人間の呼吸時の体の揺れを再現した。また、体だけでなく、髪の毛とリボンも連動して動かすようにすることで、現実での動きに近づけるようにしている。

4. 実験

本研究では、作成したアバターを Zoom で見てもらい、親近感を感じたかどうかについて、アンケート形式で評価してもらった。

4.1 実験方法

被験者は 8 人（男性 6 人、女性 2 人）である。まず、Zoom にて、都合の合わなかった男性 1 人を除く、7 人の被験者を招集し、そこに著者を加えた 9 名で、約 15 分のミーティングを行った。この際に、本研究で作成したアバターを表示し、被験者に見てもらった。ミーティングの内容は自由議論だったが、その中で、機能説明のための時間と、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合計 5 分程度設けた。後日、都合の合わなかった男性 1 人と著者で、1 対 1 の、約 10 分のミーティングを行い、同様にアバターを表示、機能説明と話者限定の時間も同様に設けた上で、アバターを見てもらった。実験終了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 3 に示す。



図 2: Zoom を用いた実験の様子 (ギャラリービュー時)



図 3: Zoom を用いた実験の様子 (スピーカービュー時)
(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)

4.2 アンケート設問内容

アンケートの設問はまず、年齢とアバターの使用経験の有無を問い、その後、アバターのモデル「響」が実験の対象として適切かどうか、本研究の4つの機能について、親近感を感じたかどうかを、動作の精度は良かったかを含めて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。

4.3 アンケート結果

年齢、アバター使用経験、4つの機能の5段階評価を、図4~9に示す。

アンケート結果より、被験者は全員20代、アバターの使用経験はその中の50%に至り、モンスターハンターなどのゲームでの使用が最も多く、他には、Live2DやVrchatでの使用経験があるという回答を得た。

アバターのモデル「響」の適切度については、全員が5の評価であった。

本研究の4つの機能について、5段階評価の平均はまばたき機能:4.75、うなずき機能:3.875、リップシンク機能:3.25、呼吸機能:4.375であった。それぞれに対する、記述式での改善点として、

まばたき機能：頻度がもう少し多いほうが良い
うなずき機能：動きが小さく、わかりづらい
リップシンク：精度が悪い、口の形の種類を増やすと良い
呼吸機能：不自然に止まる箇所があったので、常に動き続けると良い

という回答を得た。その他に、手の動きが欲しいという回答も得た。

全体の感想としては、上記の問題点に触れつつも、「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」といった回答を得た。

1. いままで、アバターを使ったことがありますか？
8件の回答

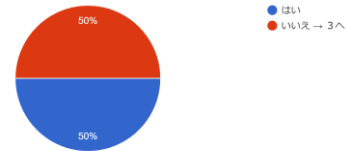


図 4: アバターの使用経験の有無

4-1. 今回のアバター (名前は「響」) は、システムのモデルとして適切でしたか？
8件の回答

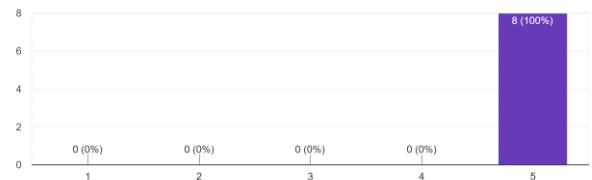


図 5: アバターのモデル「響」の適切度

5. アバターのまばたき機能で、親近感は高まった...ングが適切かどうかも含めて評価をお願いします
8件の回答

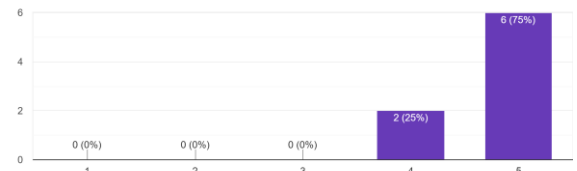


図 6: 実験における、まばたき機能の評価

6. アバターのうなずき機能で、親近感は高まった...ングが適切かどうかも含めて評価をお願いします
8件の回答

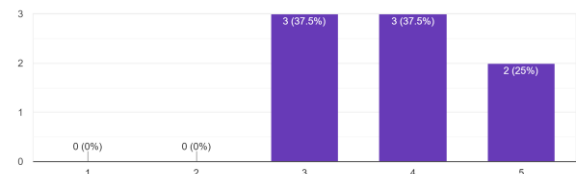


図 7: 実験における、うなずき機能の評価

7. アバターのリップシンク機能で、親近感が高ま...リップシンクの精度も含めて評価をお願いします
8件の回答

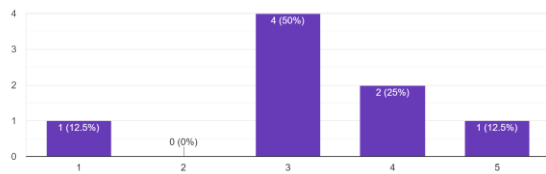


図 8: 実験における、リップシンク機能の評価

8. アバターの呼吸機能で、親近感が高まったと感じましたか？
8件の回答

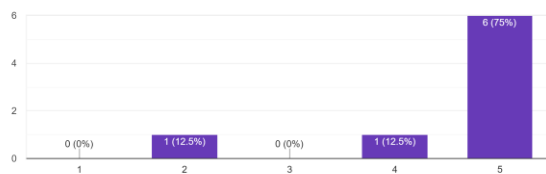


図 9: 実験における、呼吸機能の評価

5. 考察

アンケート結果より、まず、アバターのモデル「響」の適切度については、今回の実験では全員が5の評価であり、問題はなかった。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、評価の平均点は高かったが、改善点として、頻度が少ないという回答があった。それを踏まえて、頻度の調節は細かく行う必要がある。

うなずき機能は、動きがわかりづらいという回答があった。動きを大きくしても良いが、それよりはボイスやエフェクトの追加が良いという意見もあり、それらを使って、視認性や音でのわかりやすさを高めたほうが良いかもしれない。また、音声認識ソフトを導入して、香宗我部らの研究[1]で行っている、文末の品詞に応じて、うなずくか否かを定めるシステムの開発を目指すという方法もある。

リップシンク機能については、精度が悪いといった回答が多かった。この原因は、音声閾値の設定と、Unity、Voicemeeterでの音声利得値の設定にあると考えている。音声取得の際には、周りの雑音も拾ってしまうので、そうい

った環境の問題も考えつつ、設定を細かく修正していく必要がある。また、指摘を受けた、口の形の種類については、香宗我部らの研究[1]で使っている、音声認識ソフトの使用が不可欠であり、前述したうなずき機能への使用も踏まえて、導入を検討する必要がある。

呼吸機能については、常に動き続けたほうが良いという回答があった。今回の実験では、動きが時々止まるような設定になっていたが、その設定を変更する必要がある。

自由記述の回答で指摘を受けた、手の動きの導入については、2Dモデルでは困難であると考えている。理由は2Dモデルは平面構造で、親近感のある手の動きを表現しづらいからである。よって、手の動きを導入するのであれば、多方面から手を映せる3Dモデルに切り替えるという選択が必要となる。あるいは、2Dモデルで表現するなら、手の部分だけを画像切り替えて対応させる方法がある。また、手だけでなく、エフェクトやボイスの追加も指摘されており、導入を検討する必要がある。

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。2021年7月時点では、変異株の発見、緊急事態宣言やまん延防止等重点措置の再実施や期間延長などもあって、コロナ禍の収束が見えない。そうなる今後もオンライン授業やテレワークなど、オンラインでの交流は需要が伸びると考えている。そのオンライン交流において、本研究は貢献できるのではないかと考えている。

参考文献

- [1] 香宗我部多門, 山田節夫, 西野哲朗: 相づちのリアルタイム性の分析と自動生成アルゴリズムの設計, IPSJ SIG Technical Report, Vol.2017-MPS-112 No.24, 2017.
- [2] 物部寛太郎, 高橋淳也, 高橋まどか: キャラクターを使用した自動リップシンクシステムの開発, 情報処理学会東北支部研究報告, IPSJ Tohoku Branch SIG Technical Report, Vol.2012-4, 2013.
- [3] 東川知生, 山本景子, 倉本到, 辻野嘉宏: デスクワーク時における瞬目に基づく疲労蓄積の検出と適切な休憩タイミングの提示, IPSJ SIG Technical Report, Vol.2012-HCI-146 No.1, 2012.