

音声感情と顔表情のギャップ時における影響

情19-0114 折坂 幸司
指導教員 米澤 朋子

1. はじめに

現在、コロナ禍で不要不急の外出が減る中でメタバース、VR等の技術が急速に浸透した。その中で、アバターを用いることで他者とのコミュニケーションをマルチモーダルに行うことも出来る一方で、表現の幅に限界があるゆえに、感情が正確に伝わらない恐れもある。そこで本研究では、他者の感情を判断には音声感情と顔表情のどちらが強く影響するかを検証し、エージェントの感情判断モデルの作成を目指す。

今回は特に笑うという表現に注目し、音声感情と顔表情の間にギャップが生じたときのユーザの反応を調べた。1) 声と顔で笑いの度合いが違うとき、その差分により、愛想笑いなどの複雑な「偽物」的な感情を感じ取る、2) 関係性に応じて本音の出現率が異なる、という仮説を立てた。

2. 検証用システム

本システムの開発はwindowsPCでProcessingを使用した。笑顔度0%の無表情と、笑顔度100%の3Dキャラクタの顔を先に用意し、これらを基準に目元と眉の角度、口角、口の開き度合いを変え、笑顔度30%、70%の表情を作成した。音声は人間の笑い声を11パターン録音し、事前調査を実施した。笑いの度合いが10%と100%の音声と決めたものを基準に残りの9パターンの音声を評価してもらい、笑いの度合いが100%、70%、30%、10%と解釈される音声を選択した。それぞれの笑いの度合いの表情と音声の組み合わせで笑うアバターをシステム内で再現可能とした。



図1 笑顔度100%

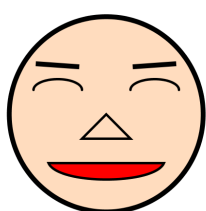


図2 笑顔度70%

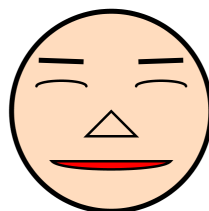


図3 笑顔度30%

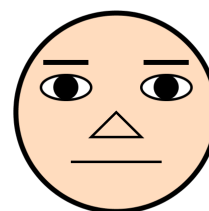


図4 笑顔度0%

3. 検証

音声感情と顔表情のギャップとその頻出度合いによる影響を調べるため、16～23歳の男女21人(男性19人、女性3人)を対象に実験を行った。実験参加者は、音声の笑い度と笑顔度が一致したものとして用意した動画(各10%、30%、70%、100%)を視聴した後に実験に参加した。評価方法として0～99の100段階で回答するVAS法を用いた。これは間隔尺度として評価値を用いるためである。

まず、実験1では1)の仮説を設定し、A. 表情要因: 笑顔の度合いの違い(100%、70%、30%、0%)、B. 音声要因: 笑い度の違い(100%、70%、30%、10%)の計2要因16条件の被験者内実験計画により評価した。その結果、両要因に全項目で有意差が見られた。表情要因では笑顔度が0%時に「違和感」と「不信感」が高まっていた。これは他の条件に対して口が開いていないことで人としての不自然さがあったためと考えられる。音声要因では、笑い度が10%時に不信感を抱

かせ、人間味を感じさせない傾向があった。さらに、交互作用において、声と顔の笑いの度合いの差分が大きいほどネガティブな印象を与える傾向が見られた。つまり、声と顔で笑いの度合いが違うとき、その差分により、ユーザが「偽物」的な感情を感じ取れることを示唆している。また、実験参加者のコメントとして、表情要因の0%時と音声要因の100%時の組み合わせが最も不気味に感じたとの意見があった。理由としては実際の人間ではありえない挙動であるからとあったが、Vtuberで同様な状況でも不気味に感じないという意見が多くあり、これは通信の影響であると推測できるために不気味に感じないと推測される。このことから、不連続時の表情と変化のある連続性の表情では感情の受け取り方が異なる可能性があると考えられる。

実験2では、2)の仮説を設定し、100%と30%の度合いの笑い声と笑顔を組み合わせた笑いを5回遷移させる中で、顔と声が合致する笑いの出現率(100%, 60%, 20%)の1要因3条件の被験者内実験計画により評価した。その結果、全項目で有意差が見られた。合致する笑いの出現率が100%時では「相手に気を許した」、「このキャラクタと親しくなりたいと感じた」の項目の評価が高く、「本心から笑っていた」の項目の評価も高くなっていた。一方で、合致する笑いの出現率が60%, 20%時では両条件の間には有意差は無く、「相手に気を許した」、「このキャラクタと親しくなりたいと感じた」の項目の評価が低く、「本心から笑っていた」の項目の評価も低くなっていた。つまり、関係性に応じて本音の出現率が異なっているが、条件が不十分であるため2)の仮説を示唆しているとは言い難い。

4. おわりに

本研究では、他者の感情を判断には音声感情と顔表情のどちらが強く影響するかを検証し、エージェントの感情判断モデルの作成を目的とし、特に笑うという表現に注目し、音声感情と顔表情の間にギャップが生じたときのユーザの反応を調べた。検証の結果、音声感情と顔表情が一致しているほどポジティブな印象を与え、不一致時の差分が大きいほどネガティブな印象を与えることが示唆された。今後は、リアルな表情を表現するアバターではなく、音声感情よりも表情を誇張することで親しみやすいアバターのモデルを目指し、検証していきたい。

参考文献

- [1]角所考, 伊藤淳子, 美濃導彦. 会話エージェントのためのノンバーバル表現間の相互依存性のモデル化. 社団法人情報処理学会研究報告. 2003巻100号, pp. 31-38, 2003.
- [2]白井暁彦. VR4. 0-リアルタイムバーチャルキャラクターが牽引するVR産業における基盤研究. DHUJOURNAL2019, vol. 06, pp. 82-92, 2019

