

テレプレゼンスロボットにおけるモーフィングを用いた

存在感の余韻強化の試み

Enhancing the Lingering Presence of Remote Speakers
by Visual Morphing in Telepresence Robot.吉田 寛和^{*1}
Hirokazu Yoshida田中 文英^{*2}
Fumihide Tanaka^{*1} 筑波大学 システム情報工学研究科
Graduate School of Systems and Information Engineering, University of Tsukuba^{*2} 筑波大学 システム情報系
Faculty of Engineering, Information and Systems, University of Tsukuba

Telepresence robot can give a sense such like remote person is really existing there. However, this sense is suddenly disappeared after the call. Therefore, the talker will feel loneliness. This research explores "tele-vestige" in which talkers can feel weak presence even after a call finish. To give this "tele-vestige", remote person's face is morphed to a telepresence robot's face after the call. We tested a pilot experiment to investigate the effect of this morphing on loneliness.

1. はじめに

テレビ電話のような遠隔通信メディアの普及により、遠隔地間でもコミュニケーションが容易に行えるようになってきた。対面と同等のコミュニケーションを実現するには未だに至っていないが、それを補う技術として近年注目されているのが物理的な身体を持ったテレプレゼンスロボットである[大塚 2013]。テレプレゼンスロボットはディスプレイ表示やジェスチャによって、遠隔者だけでなく、相手である通話者にも遠隔者がその場にいるかのような存在感を感じさせることができる[妻木 2012]。

しかし、テレプレゼンスロボットは通話が終了すると、遠隔者を特定できる要素がなくなり、遠隔者の存在感が急激に消失してしまう。この急激な存在感の変化が、通話者に通話終了後の寂しさをより顕著に感じてしまうのではないかと考えられる。

本研究は、通話が終了した直後に遠隔者の存在感を希薄しながら通話者に提示することで、存在感に「余韻」を残し、テレプレゼンス終了後の寂しさを軽減することを目的としている。その手法として、通話を終了する際に、遠隔者を特定できる特徴を、ロボットの対応する特徴と混ぜ合わせながら通話者に提示する。本稿では、遠隔者を特定できる要素として顔に注目し、通話終了後に人の顔とロボットの顔を混ぜ合わせながら切り替えを行う。

2. 関連研究

一時的に存在感や臨場感が消失する問題を解決する研究は多くある。テレプレゼンスシステムの多くは遠隔者の映像を平面ディスプレイに表示するため、通話者がディスプレイの正面にいない場合に、存在感が消失してしまう。この問題を解決する手段として、ディスプレイの形状を特殊なものに変えるということがあげられる。SphereAvatar[Oyekoya 2012]は、球面ディスプレイを使用し、顔を球面全体に表示することで、どの位置からでも遠隔者を見ることができる。しかし、バーチャルアバターの映像を使用しており、人間の顔を表示した際に存在感を感じさせるこ

とができるかの調査は行っていない。遠隔者の存在感をより高めるために、三澤らは人の顔形状のスクリーンに遠隔者の顔を表示するテレプレゼンスシステムである LiveMask[三澤 2012]を開発した。顔形状のスクリーンは 3D スキャナーを用いて得られたデータをもとに 3D プリンタによって作成しているため、通話が終了した後も遠隔者を特定できる要素は残っている。しかし、通話終了後の存在感についての調査は行っていない。

身体性を工夫した存在感の研究も行われている。ChameleonMask[Misawa 2015]では iPad air と HMD から構成されるヘルメットを代理人となる人間がかぶり、遠隔者の顔を iPad air の画面に表示するテレプレゼンスシステムを提案している。このシステムでは、代理人に遠隔者と同じような体形の人間を使用し、人間を特定するための主要要素である顔をマスクのように装着することで遠隔者の存在感を提示し、通話者に装着している代理人があたかも遠隔者であるように感じさせることができたことを報告している。しかし、代理人がマスクを外すと存在感が急激に消失してしまい、遠隔者がこの場にはいないことを感じさせてしまう。

本研究のように遠隔者の存在感を継続することを目的とした研究も行われている。田中らはテレプレゼンスモードから自律モードに切り替わる際のギャップが存在感の低減になることを明らかにし、自律システムが遠隔者を装って通話することによって、このギャップが軽減されることを報告している[田中 2015]。本研究は遠隔者とはっきりわかる状態からの切り替えを行い、存在感を残すことを目的にしているという点においてこの研究と異なっている。

3. 切り替え方法

存在感に余韻を残すためには、テレプレゼンスモードから自律モードへの切り替えが急激に行われるのではなく、遠隔者を特定できる要素とロボットのそれに対応する要素を混ぜ合わせながら徐々に切り替わっていくことが必要であると考えられる。本研究では、遠隔者を特定できる要素として、顔を選択し、2 つの顔を混ぜ合わせながら切り替えるためにモーフィングを使用する。モーフィングとは、映像技術で使用される変形手法[藤代

2015]で、画像を切り替える際に特徴点をもとに画像の形状を変化させながら画素値も変化させる手法である。以下にその手順を示す。

3.1 顔画像の取得

ロボットの頭部に表示する遠隔者の顔は遠隔地に設置したカメラを用いて取得する。本研究では Microsoft LifeCam Studio を使用する。カメラ画像の取得には、OpenCV 2.4.13.5 を使用する。

3.2 特徴点の取得

遠隔者の顔画像とロボットの顔画像間の対応点を取得するために特徴点の取得を行う。人間の顔特徴点の取得には、C++向けのツールキット Dlib に用意されている HOG と SVM を用いた検出器を使用する。取得できる特徴点の数は 68 個である。

ロボットの顔画像ではこの顔検出器が使用できないので、人手であらかじめ特徴点を指定した。人間の顔の特徴点との対応付けを行うために、こちらも特徴点の数は 68 個である。

3.3 正規化

2 次元画像では、カメラから顔までの距離が変化することで画像中における顔の大きさも変化してしまう。そこで、3.2 で取得した特徴点をもとに、遠隔者の顔の高さと幅をロボットの顔の高さと幅に正規化する。

3.4 モーフィング

モーフィングによる顔の変形は、顔を単純な形状に分割し、それらを変形することによって行う。本研究ではドローネー分割によって顔を三角形に分割する。画像 1 における特徴点の x 座標 x_1 から画像 2 における特徴点の x 座標 x_2 への変形において、混合率 α ($0 \leq \alpha \leq 1$) のとき、モーフィング画像における特徴点の x 座標 x_M は以下の式によって求められる。

$$x_M = (1 - \alpha)x_1 + \alpha x_2 \quad (1)$$

y 座標も同様に、

$$y_M = (1 - \alpha)y_1 + \alpha y_2 \quad (2)$$

これによって移動した特徴点を基準に、分割した三角形をアフィン変換することで顔全体を変形する。これと同時に画素値の混合率も変化させる。混合率 α ($0 \leq \alpha \leq 1$) における点 (x, y) の画素値 $I_M(x, y)$ は以下の式によって求められる。

$$I_M(x, y) = (1 - \alpha)I_1(x, y) + \alpha I_2(x, y) \quad (3)$$

これらの処理によって得られた画像をロボットの顔として表示する。

4. 実験内容

モーフィングを用いた手法が寂しさの軽減に有効であるか検証をするために、実験を行った。実験では、実験者が遠隔者に、実験参加者が通話者にあたる。実験は 20 代の学生 7 名に協力していただいた。

4.1 使用ロボット

ロボットには、SoftBank Robotics 社の Pepper を使用する。Pepper のディスプレイは頭部ではなく胸部に搭載されている

で、本実験に使用すると顔が 2 つになってしまう、実験参加者に違和感を与えてしまうことが考えられたので、Pepper の頭部にディスプレイを搭載するための外装を作成した。外装を取り付けたロボットの画像を Fig.1 に示す。ディスプレイには Surface を使用する。外装を取り付けることで Pepper のカメラが使用できなくなり、実験者が実験室の様子を見ることができなくなってしまうため、Pepper の頭部の上に USB カメラを取り付ける。使用カメラは Microsoft LifeCam Studio である。顔画像の送受信は UDP 通信によって行う。音声の送受信には Skype を使用する。ロボットの発話は Choregraphe2.5.5 の say モジュールを使用して発話したものを録音し、再生したものを Skype で送信する。

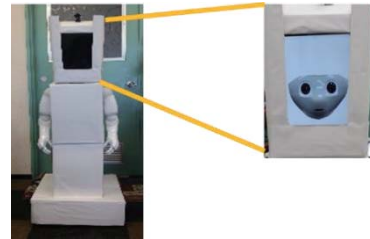


Fig.1: 外装を取り付けた Pepper

4.2 実験条件

実験条件は以下の 3 つで行う。

- A) 切断条件: テレプレゼンス終了後、人間の顔がすぐにロボットの顔に切り替わる。
- B) フェード条件: テレプレゼンス終了後、人間の顔が徐々に薄くなり、ロボットの顔が徐々に表れてくる(変形は行われない。)
- C) モーフィング条件: テレプレゼンス終了後、人間の顔をロボットの顔にモーフィングしながら切り替わる。

A はベースライン条件として行う。B は C による効果が式(3)のみではないことを確認するために行う。

B, C の切り替わりに要する時間は約 20 秒である。実験は参加者内配置で行う。順序効果を考慮して実験参加者ごとに条件の順番を入れかえる。

4.3 実験手順

実験は以下の手順で行う。

1. 実験参加者は実験者とともに実験室に入り、椅子に座ってもらう。
2. 実験参加者には、アンケートに回答してもらう。アンケート回答終了後、実験参加者にしばらく待機するように伝え、実験者は実験室を出て別室に移動する。
3. 実験者が実験室を出た約 1 分後にロボットが「電話が来たいです」と発話し、テレプレゼンスモードに切り替わる。この際、画面はすぐに遠隔者の顔に切り替わる。
4. 実験参加者はロボットを介して実験者とともにタスクを行う。タスクとして、実験参加者には実験者と Yes・No ゲームをしてもらう。Yes・No ゲームは以下のルールで行われる。
・プレイヤー人数: 2 人
・1 人が出題者、もう 1 人が回答者となる。出題者は問題を出し、回答者はその答えを考える。

- ・問題はあらかじめ選択肢が用意されたものを使用する。選択肢にはイラストと名称が記載されており、出題者はその中から答えを1つ決める。
- ・出題者が答えを決めたら、回答者はそれが何かを当てるために質問をする。出題者は質問に対して「はい」または「いいえ」で返答する。「いいえ」の場合は、回答者の質問内容に関連したヒントを出す。
- ・答えは必ずしもイラストの通りでなくてもよい。ただし、嘘の返答をしてはいけない。
- ・質問によるやり取りを3回繰り返した後、回答者が回答する。正解した場合は回答者の勝ち、不正解の場合は出題者の勝ちとなる。

このゲームを3回行う。本実験では実験参加者を出題者、実験者を回答者とする。回答後の正誤の判定に対して実験者は単調な反応をする。

5. タスク終了後、実験者は実験参加者に自分も実験室へ戻ることを伝え、通話を終了する。終了の際に、ロボットは切り替え手法を実行する。ただし、実際に終了するのは実験者の顔画像の表示と声の出力のみで、実験者は実験参加者側のカメラ映像と音声を確認できる。
6. 切り替え終了から約30秒後にロボットが「アンケートに記入をお願いします」と発話する。実験参加者にはアンケートに記入してもらう。アンケート記入後、実験参加者には記入が終わったことをロボットに伝えてもらう。記入が終わったことを確認した後、実験者は実験室に戻る。
7. 1～6の手順を他の2つの条件でも行う。

各条件における実験参加者の心的状態の変化を調査するため、アンケートには一時的気分尺度 TMS (Temporary Mood Scale) [徳田 2011]を使用する。TMS は「活気」、「疲労」、「怒り」、「抑うつ」、「緊張」、「混乱」の6つの下位尺度において今現在の気分変化を評定するアンケートである。アンケートの回答形式は「1. 全くあてはまらない」から「5. 非常にあてはまる」の5段階とする。通話前のアンケートでは TMS のみ行い、通話後のアンケートでは TMS と自由記述を行う。各尺度3項目の合計値が合計得点となり、通話前後の合計得点の変化量で評価する。

通話終了直後にアンケートを行うと、切り替え手法の違いではなく、タスクのやり取りや画像に対する印象が影響する可能性がある。そのため、画面が切り替わった後に30秒間の待ち時間を設定する。

実験を行うにあたり、以下の仮説を立てた。

仮説 1: 切断条件における「抑うつ」の増加量はモーフィング条件における増加量よりも多い。
(切断条件 > モーフィング条件)

仮説 2: フェード条件における「抑うつ」の増加量はモーフィング条件における増加量よりも多いか、同等である。
(フェード条件 \geq モーフィング条件)

5. 実験結果

TMS によって得られたアンケート結果から通話前と通話後での気分の平均変化量を Fig.2、「抑うつ」の各質問項目における平均変化量を Fig.3 に示す。「活気」における変化量は増加すればよい結果となり、「疲労」、「怒り」、「抑うつ」、「緊張」、「混乱」における変化量は減少していればよい結果となる。サンプル数

が十分ではないので、本研究では分散分析を行わず、数値比較のみを行った。

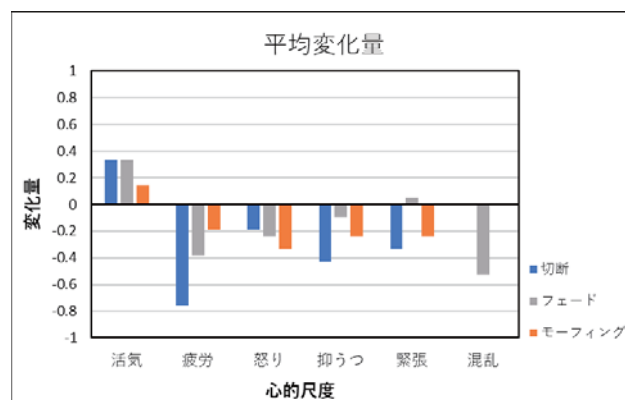


Fig.2 心的尺度の平均変化量

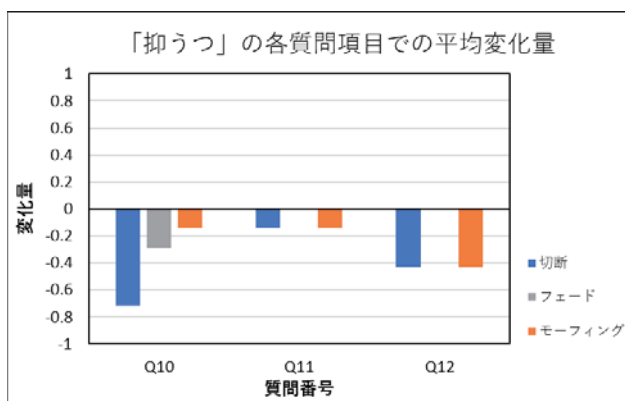


Fig.3 「抑うつ」の各質問項目での平均変化量

6. 考察

抑うつの項目において、変化量の平均が「切断条件 < フェード条件 < モーフィング条件」となった。これは仮説1、仮説2を棄却する結果である。しかし、すべての条件において、通話前よりも通話後のほうが「抑うつ」が軽減しており、本研究の目的に最も関連のある質問である「孤独でさびしい」においても平均変化量が減少していた。これは条件の違いで寂しさの軽減量が変化したわけではなく、もともと寂しさを感じていなかったためであると考えられる。実際、アンケートの回答でも実験前、実験後ともに1や2を選択する実験参加者が多く見られた。このような結果になった原因として、本実験が通話後の寂しさを感じにくい設定だった可能性が考えられる。通常の遠隔コミュニケーションでは、通話が終了した後、しばらくはその人物と会うことがない。しかし、本実験では実験進行のために通話が終了した数分後に通話者である実験者と会うということがわかってしまっている。そのため、寂しさを感じにくかったのではないかと考えられる。

「活気」、「疲労」、「混乱」においては、フェード条件がモーフィング条件よりもわずかながらよい結果となった。この要因として、モーフィングによる不気味さが考えられる。実験後のインタビューでは「具体的な要因はわからなかったが、フェード条件よりモーフィング条件のほうが気持ち悪かった」や「フェード条件では画像がずれていたため、切り替えに時間がかかっているのだろうと感じたが、モーフィング条件では気味悪さから早くこの場から立ち去りたいと思った」との意見が得られた。MacDorman らの実

験[MacDorman 2006]においても、アンドロイドとヒューマノイドロボットの中間的な顔の画像は不気味さの評価が高くなることが報告されており、今回の実験のような動的なモーフィングでも同様の印象を与えてしまったと言える。このモーフィングでの顔の変形による不気味な印象が、「活気」に影響を与えた結果、「疲労」、「混乱」にも影響が出たのではないかと考えられる。しかし、条件間での差はそれほど大きいものではなく、サンプル数を増やしても有意な差は見られないと考えられる。

その原因として、モーフィングによる切り替えを通話終了後のみに実施したことが考えられる。通話終了後に存在感の余韻を残すことを目的としていたため、人からロボットへの切り替えにのみ変化を加えた。そのため、実験参加者は通話終了時に初めてモーフィングによる切り替えを目にすることになる。これまで何も変化がなかったにも関わらず、突然容姿が変化したことにより、切り替えの際にギャップを感じてしまったのではないかと考えられる。この改善方法として、通話終了時だけでなく、通話開始時にもモーフィングによる切り替えを行うことが挙げられる。実験参加者からは、「通話中にもモーフィングを行った方がよいのではないか」という提案があった。しかし、先述の通り、モーフィングによる顔の変化は通話者に不気味さを感じさせてしまう可能性がある。そこで、通話者の視線に注目しようと考えている。通話者が視線を向けている状態、つまり意識的にロボットを見ている際は変化せず、視線が外れ、意識がロボットにあまり向いていない状態のときにわずかにモーフィングをすることで、通話中に不気味さを感じさせず通話中のモーフィングを行えると考えられる。

今回行ったアンケート手法も適切ではなかったと考えられる。アンケートは、回答を求める際に実験参加者の意識に介入する必要があり、それによって実験参加者の主観的な心的状態が変化してしまう可能性がある。通話時におけるハグの効果を調査する実験[Sumioka 2013]では、通話前後のストレスホルモンの変化量を計測することによって評価を行っている。このように生理的な変化を用いて評価する方が適切であると考えられる。IJzerman によると、孤独感によって体温が低下することがわかっている[IJzerman 2012]。そこで、今後は、通話前後における体温の変化を計測することによって評価を行うことを検討している。

7. まとめ

本研究では、テレプレゼンスロボットにおいて通話終了後に遠隔者の存在感が急激に消失することによる寂しさを軽減するために、存在感に余韻を残すことを提案した。その具体的な方法として、通話終了の際にロボットの顔に表示されている遠隔者の顔をロボットの顔とモーフィングしながら切り替えることを考え、実験を行った。比較条件として、切断条件、フェード条件を設けた。通話前と通話終了後にアンケートを実施し、各条件での心的状態の変化量を調査した。しかし、条件間で大きな差は見られなかった。

その原因として、実験が寂しさの変化を評価するのに適していなかったことや単純にモーフィングを用いるだけでは存在感の余韻を提示することはできず、通話者がロボットに対して不気味さを感じてしまっただけとなったことが考えられる。

8. 今後の展望

これらの結果を踏まえ、今後実験手順を改善した再実験を行うことを検討している。まず、実際の使用で想定される通話直後は会うことができないという状況を作り出すことが重要であると考

えられる。そこで、実験者と遠隔者の役割を分割することを検討している。また、通話の回数を増やしすぎると次第に通話後の寂しさを感じにくくなる可能性もあり、条件は通話前後にのみモーフィングを行う条件と通話中にもモーフィングを行う条件の2条件比較に変更しようと考えている。

評価にはアンケートを使用せず、サーモセンサを用いた実験参加者の体温の変化量を使用することを検討している。待機時間も現在の30秒では通話後の孤独を感じにくい可能性がある。そこで、IJzerman らの実験を参考に、待ち時間を300秒程度に設定することを検討している。

謝辞

本研究はJSPS 科研費 17K19993 の支援を受けて行われた。

参考文献

- [大塚 2013] 大塚 和弘, 熊野 史朗, 松田 昌史, 大和 淳司: MM-Space: 頭部運動の物理的補強表現に基づく会話再構成, 情報処理学会論文誌, Vol.54, No.4, 1450-1461, 2013.
- [妻木 2012] 妻木 勇一: 遠隔コミュニケーションとテレロボティクス, 日本ロボット学会誌, Vol.30, No.6, 606-608, 2012.
- [Oyekoya 2012] Oyewole Oyekoya, William Steptoe, Anthony Steed: SphereAvatar: A Situated Display to Represent a Remote Collaborator, Proceedings of the SIGCHI conference on human factors in computing systems, 2551- 2560, 2012
- [三澤 2012] 三澤 加奈, 石黒 祥生, 暦本 純一: LiveMask: 立体顔形状ディスプレイを用いたテレプレゼンスシステムにおけるコミュニケーションの評価, インタラクション 2012, 41-48, 2012.
- [Misawa 2015] Kana Misawa, Jun Rekimoto: ChameleonMask: Embodied physical and social telepresence using human surrogates, Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, 401-411, 2015.
- [田中 2015] 田中 一晶, 宇野 弘晃, 山下 直美, 中西 英之, 石黒 浩: ロボット操作者の偽存在感によるソーシャルテレプレゼンスの生成, インタラクション 2015, 28-37, 2015.
- [藤代 2015] 藤代 一成ほか: コンピュータグラフィックス[改訂新版], 公益財団法人 画像情報教育振興協会, 2015.
- [徳田 2011] 徳田 完二: 一時的気分尺度 (TMS) の妥当性, 立命館人間科学研究, 第 22 号, 1-6, 2011.
- [MacDorman 2006] Karl F. MacDorman, Hiroshi Ishiguro: The uncanny advantage of using androids in cognitive and social science research, Interaction Studies, Vol.7, No.3, 297-337, 2006.
- [Sumioka 2013] Hidenobu Sumioka, Aya Nakae, Ryota Kanai, Hiroshi Ishiguro: Huggable communication medium decreases cortisol levels, Scientific Reports, Vol.3, 3034, 2013.
- [IJzerman 2012] Hans IJzerman, Marcello Gallucci, Wim T.J.L. Pouw, Sophia C. Weißgerber, Niels J. Van Doesum, Kipling D. Williams: Cold-blooded loneliness: Social exclusion leads to lower skin temperatures, Acta psychologica, Vol.140, No.3, 283-288, 2012.