

遠隔操作型アンドロイドの笑い動作の付加効果

An Effect of Automatically Generated Laughing Motion of Teleoperated Android

船山智*

港隆史†

石井カルロス寿憲‡

石黒浩§

Tomo Funayama

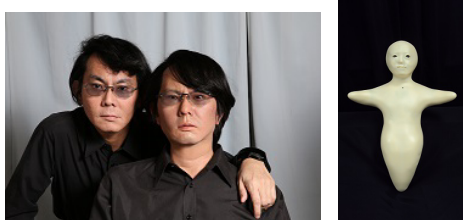
Takashi Minato

Carlos Toshinori Ishi

Hiroshi Ishiguro

1. 序論

我々の人間社会において人と人とのコミュニケーションは日々多様化していき、電話やEメールなどの通信メディアを用いることで、時間、場所を問わず誰かとコミュニケーションをとることができる。そのような通信メディアの一つとして遠隔操作型アンドロイドが開発されている [1](図1)。これらは電話と同様の音声による対話が可能であると共に、アンドロイドが操作者と同期した動作を行うため、対話者にとってはまるで操作者と対面しているかのように感じられる。また、実体があり触れることもできるため、他のメディアに比べ対話相手の存在感を強く感じることができる。



(a) ジェミノイド HI-4 とそのモデル (b) テレノイド

図1 遠隔操作型アンドロイド

遠隔操作型ロボットに操作者と同期した動作をさせるための動作生成方法は、操作者の動作をセンシングし同様の動作となるようにロボットの関節角を動かす動作複製手法が一般的である。しかし動作複製手法では、アンドロイドにとって不自然な動作となることがある。その理由として、操作者の音声から期待されるアンドロイドの動作と実際のアンドロイドの動作の不一致が考えられる。従来手法である動作複製手法で、この不一致が起こる原因としてアンドロイドの自由度が人間よりも少ないことが挙げられる。

アンドロイドの動かせる関節軸は、人間と比較して少ないため人間と同様に動くことができない(自由度の詳細については2章で記述する)。つまり従来手法で、人間である操作者の動作を複製したとしてもその動作をアンドロイドで完全に再現できるわけではなく、アンドロイドは常に操作者の動作よりも動きが制限された状態にある。例えば、遠隔操作型アンドロイドであるテレノイドには顔内部に目と口を動かすアクチュエータしか存在

しないため、操作者が笑顔を作りながら笑い声を出した場合、テレノイドの声は笑い声だが、表情は笑顔でないという状態が生じる(テレノイドの可動部分は図3を参照)。このようなハードウェアの制約により、音声と動作の不一致が生じていると考えられる。

この音声と動作の不一致を解決するために、音声からの動作自動生成手法を提案する。アンドロイドの自由度の少なさを解決する方法として、単純にアンドロイドのアクチュエータを増やすことが考えられる。しかし、アンドロイド内部のスペースや配置の問題でこれ以上アクチュエータを増やすことは困難である。そこで、操作者の音声に合わせてアンドロイドにとっての自然な動作を自動生成することで、この問題を解決する。

本研究では、従来手法で不自然となっている動作の一つとして、アンドロイドの笑いに取り組む。笑いは対話において頻出するものであり、本来コミュニケーションをポジティブに導くものである。しかし、動作複製のみでは音声から期待される動作との不一致が生じ、操作者が笑うことで対話者にネガティブな印象を与えてしまう可能性がある。これは対話メディアとして致命的でありその改善は必須である。また、ロボットを自律的に動かす場合にも笑う際の動作として利用できる知見となり得る。よって本論文では、笑い声と同期したアンドロイドの笑い動作を自動生成するシステムの開発を目指す。ここで問題となるのは、自由度が不足しているアンドロイドでどのように笑い動作を表現するかという点である。本論文では、人の笑い動作においてアンドロイドでも実現できる動き成分を抽出し、それを誇張して表現する手法を提案する。

音声から期待される動作は、アンドロイドの見かけの影響も受けると考えられる。NowakらはCGアバターエージェントを用いた対話実験を行い、エージェントの擬人化の度合いが低い方がより存在感が強かったと報告している [2]。これは、擬人化の度合いが高いエージェントの方がより人間らしく動くこと期待され、実際の動作がその期待に一致しなかったためだと考察されている。つまり、エージェントの見かけが人間に近づくほど、人間らしい動作をするはずだという期待は高まっていく。実体を持つロボットでも同様の関係が考えられ、ロボットの見かけが人間に近いほど、人間らしい動作を期待される。よって、ロボットの見かけによって適切な動作は変化すると考えられる。そこで、見かけが実在の人間に酷似しているジェミノイドと、抽象的な人間のデザインであるテレノイドの2体のアンドロイドを用いて、ロボットの見かけの違いによって、今回生成した笑い動作の効果がどのように変化するか検証を行った。

*大阪大学, Osaka University

†国際電気通信基礎技術研究所, ATR

‡国際電気通信基礎技術研究所, ATR

§大阪大学, Osaka University

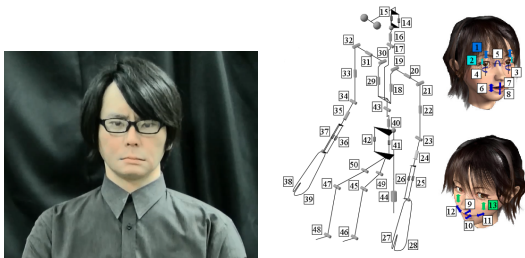
2. 遠隔操作型アンドロイド

本章では、今回提案手法の検証実験で用いた遠隔操作型アンドロイドであるジェミノイドとテレノイドについて、従来の動作複製手法を用いた遠隔操作システムについて紹介する。

2.1 遠隔操作型アンドロイドの概要

2.1.1 ジェミノイド

ジェミノイドは実在の人間に非常に近い見かけをしたアンドロイドであり、今回の検証実験では40歳代の男性をモデルとしたジェミノイド HI-2 を使用した(図2(a)). ジェミノイド HI-2 は全身に空気圧駆動のアクチュエータが50個配置されており、人間のような柔軟な動作が可能となっている(図2(b)). しかし、空気圧駆動では電気駆動程の素早く細かい動作ができないという欠点がある。また、マイクとスピーカは内蔵されていないため、外部機器を用いている。



(a) ジェミノイド HI-2 (b) ジェミノイド HI-2 の自由度

図2 ジェミノイド HI-2

2.1.2 テレノイド

ジェミノイドが実在の人間そっくりなのに対して、テレノイドは人間の個性をそぎ落とした誰にでも見えるようなデザインのアンドロイドである(図3)。ジェミノイドは見かけが特定の個人にそっくりなため、対話者から見て操作者の音声とその見かけにギャップが生じて不自然さを感じさせることがある(例: 男性の見かけであるジェミノイド HI-2 から女性の声がした場合)。そこで、誰が操作していても音声と見かけとの不一致を解消できるように、人間の最小限のデザインをコンセプトとしてテレノイドは開発された。この性別も年齢もわからないような中立的なデザインにより、誰が操作していてもロボットに操作者の姿を想像でき音声との不一致が生じない。自由度についても、コミュニケーション上最低限必要だと考えられた9つのアクチュエータが配置されている(図3)。マイクは顔の横の耳の位置に、スピーカは胸の位置に内蔵されている。また、アクチュエータは電気駆動のものを用いている。

2.2 遠隔操作システム

動作複製手法を用いたアンドロイドの遠隔操作システムを図4に示す。図ではアンドロイドとしてテレノイドを描いているが、ジェミノイドにおいても同様のシステムである。このシステムにより、アンドロイドに生成さ

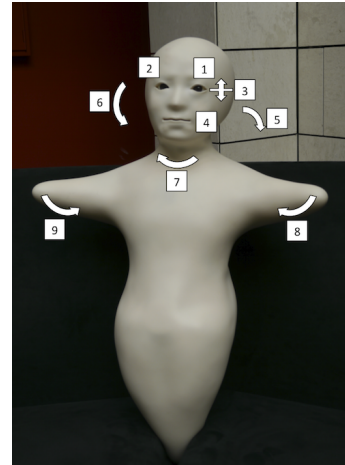


図3 テレノイドとその自由度

れている動作は操作者の首と口の動きである。操作者が地磁気センサと加速度センサを搭載したヘッドセットを装着することで、その首の動きを認識し、アンドロイドが同様の動作となるように各アクチュエータに指令値を送信する。口の動きは石井らが開発した音声から口の形状を推定、生成するシステムを用いている[6]。

また、ヘッドセットのマイクとスピーカをアンドロイド側のスピーカとマイクに接続することで音声通話が可能である。カメラについては、ジェミノイドとテレノイド共に内蔵されていないため、外部機器を用いることで操作者はアンドロイド側の映像を見ることが可能である。このシステムでは、操作者側とアンドロイド側のデータ通信により遠隔操作を実現しているため、インターネットを利用することで、どれだけ離れた場所であってもアンドロイドを遠隔操作可能である。

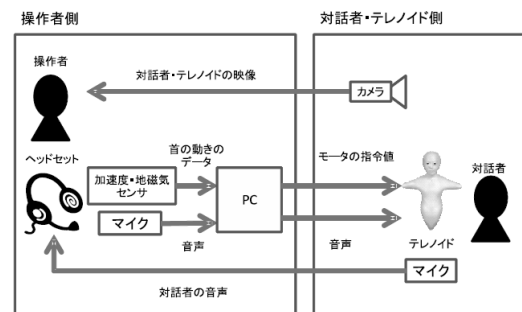


図4 アンドロイドの遠隔操作システム

3. 関連研究

3.1 アニメーションにおける誇張表現

アニメーションでは、キャラクターが現実の物理法則を無視した動きをしていることが多い。キャラクターの動きを誇張して描写することで、動作のわかりやすさや自然さ、躍動感を生み出している。

アニメーターの描くそのような誇張動作をコンピュータで自動生成する研究は様々に行われており、北爪らは

アニメーションにおける誇張動作生成手法として以下の様に分類している [3].

1. 角度制御による動作誇張手法 [7]
2. 軌跡制御による動作誇張手法 [8]
3. 加速度制御による動作誇張手法 [9]
4. タイミング制御による動作誇張手法 [10]
5. カトゥーンブラー [11]

この内 2, 3, 4 はモーションキャプチャデータを, 1 はキーフレームによる補間アニメーションを誇張元としている. 5 のカトゥーンブラーは非写実的な 3D アニメーションに動きの効果を生成する手法であり, 上記のどちらも誇張元とすることができる. これら 2, 3, 4 からわかるように, アニメーションではモーションキャプチャのような現実の人間と全く同じ動作は表現として不十分であり, 人間にはないような誇張された動作が必要であることがわかる. 人間ではないキャラクタを表現するものとして, アンドロイドでも同様に誇張された動作が有効なのではないかと考えられる.

3.2 文楽人形における誇張表現

文楽人形は日本の古典芸能である人形浄瑠璃で使用されている人形であり, 人間である人形遣いが動かすことで演目を演じる (図 5). 文楽人形は表情が動かない. しかし, 全身動作によって多様な感情を表現している. 中川は, アニメーションの感情表現の動きを文楽人形で実演する実験を行い, あるアニメーションの 1 シーンにおいて文楽人形の動作量がアニメーションの登場人物よりも, 首の動きが 4 倍, 動作角度が 1.8 倍, 上下の動きが 2.5 倍, 手の動きが 36 倍であったことを報告している [4]. このことから, 表情のない文楽人形はアニメ以上に誇張された動作で感情を表現していることがわかる. 文楽人形はまさに関節軸数の不足を誇張動作で補えていると言える. また, 中川は文楽人形の感情と動作の規則を抽出し, その感情表現動作をロボットに応用している. その中の幸・喜びの感情動作において, 笑う時には”上下に震え, 時折上方向を向く”としている.



(a) 文楽人形 (b) 人形遣い

図 5 文楽人形と人形遣い

アニメーションと文楽人形の例から, エージェントの動作について, 自由度の不足を誇張した動作で補うことができると考えられる. アニメーションでは様々な誇張手法があるが, 文楽人形でも見られるように動作角度と動作量を誇張する手法を, 今回アンドロイドの笑い動作に利用することとする. これにより, 文楽人形と同様に, 表情などの自由度の不足を全身動作で補うことを考える.

3.3 音声入力による遠隔操作型ロボットの動作生成

石井らは, 操作者音声からそれに同期した遠隔操作型アンドロイドの口唇動作を生成する手法を提案し, 提案手法が画像認識やモーションキャプチャによる操作者の動作複製手法よりもより自然な動作生成が可能であることを示した [6]. 例えば, 操作者があまり口を動かさずに発話した場合, 従来手法で複製するだけではアンドロイドの口唇動作は乏しくなり音声と合致しない. しかし提案手法は, 操作者の口唇動作とは無関係に音声に同期した口唇動作をアンドロイドで生成するため, より自然であるという評価を得たと考えられる. これはアンドロイドにおける音声と口の同期の重要性を示している.

また, 境らはロボットの遠隔操作において, 操作者音声の談話機能に適合した頷きなどの頭部動作を自動生成するシステムを開発し, 遠隔操作型アンドロイドであるテレノイドを用いてその有効性を示した [12]. この遠隔操作システムでは, 操作者はディスプレイを見ながら対話相手と対話することになり, その頭部動作は対面对話時よりも少なくなると考えられる. そこで, 音声の言語的意味と一致する頭部動作を自動生成することで, より適切なアンドロイドの動作とすることができたと考えられる. これは, 遠隔操作型ロボットにおいて, その動作と操作者の音声の言語的意味との一致が重要であることを示している.

これらの先行研究から, 音声とのタイミング, 言語的意味の一致が実現されていれば, アンドロイドの動きが操作者の動きと同一でなくても, 自然な動きとなることがわかる. 本研究ではそれらの一致を実現するため, 音声から笑いの生成を目指す.

4. アンドロイドにおける笑い動作のデザイン

本章では, アンドロイドの自然な笑い動作をどのように設計したかを説明する.

まず人間の笑い動作を調べるため, 対面对話を行っている人間の映像を解析したところ, 人間が笑う際に伴う動作として, 身体全体の上下方向の振動が多く観察された. これは, 笑い声の発声時に肺や腹部が伸縮することにより起こると考えられ, 様々な笑い動作の中でも一般的なものであると考えられる. また, 文楽人形の笑い表現についても上下方向の震えが報告されている [4] ため, この上下方向の振動動作をアンドロイドの笑い動作として生成する. ジェミノイドでは肩などの身体のアクチュエータを多数動かして振動することができるが, 今回はテレノイドと同じ動作とするため, 頭部動作での上下振動とした.

人間のこの振動動作は発声により胴体部分で起こると考えられ, その振動に付随する頭部動作の動く角度は非常に小さい. 関連研究の知見から, 誇張動作がアンドロイドにおいても動作表現として有効だと考え, その頭部の動作量を明らかに大きくした誇張動作を生成する. この誇張動作により, アンドロイドの表情の乏しさを, 文楽人形のように補い, 自然な笑いの表現ができるのではないかと考えられる. 生成動作では振動角度は正面から上下プラスマイナス 10 度程度, 振動回数は笑い声の長さに合うよう 2 回とした.

5. 検証実験

提案手法により笑い声に同期して笑い動作を自動生成することで、音声から期待される動作との不一致を解消し、自然な笑いを表現できているかを検証するため、被験者が遠隔操作による対話を行っているアンドロイドの映像を見て、対話中の遠隔操作型アンドロイドの動作の印象評価を行った。生成する笑い動作は、人間によく生じる動作を誇張したものであり、ロボットの笑い動作として適切であると考えられる。しかし、序論でも論じたように、ロボットの見かけによって期待される動作は変わると考えられ、この笑い動作の有効性も変化すると考えられる。今回の検証実験では、この動作による笑い表現の効果が、ロボットの見かけにより変化するかどうかを、ジェミノイドとテレノイドの2体を用いて検証した。

5.1 実験目的

笑い声に同期した笑い動作を自動生成する提案手法の有効性を、次の2つの仮説を検証することで明らかにする。

仮説1：動作要因の影響

操作者の複製動作とは別に笑い声に同期した笑い動作を生成することで、操作者の動作とは異なる動作となるが、アンドロイドのより自然な笑いを表現できる。

仮説2：見かけ要因の影響

音声から期待される動作は、見かけによっても異なり、生成すべき笑い動作も見かけにより決まる。

5.2 実験手法

実験における条件、生成動作、手順、対話音声についてそれぞれ説明する。

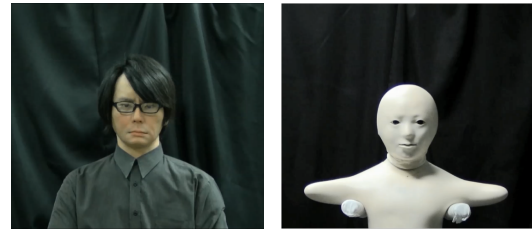
5.2.1 条件

今回はあらかじめ用意した対話音声の笑い声に手動でラベルを付け、音声再生時に笑い声部分で動作を生成するシステムを用いて実験を行った。対話データとして、2章で述べた遠隔操作システムを用いて行った人とテレノイドの対話を記録したものを用意した。記録したデータは操作者と対話者それぞれの対話音声、操作者の頭部動作である。操作者の頭部動作に、操作者の音声を入力とした[6]の口唇動作生成システムの出力を加えることで、対話時の操作者音声と頭部動作をアンドロイドで再現することができる。さらにその再現された動作に、笑い声のタイミングで特定の動作を加えた再生も可能である。これにより、笑い声に同期して笑い動作を付加した対話中の動作を作り出すことができる。ここでは、操作者の動作を再現したアンドロイドの対話映像と、それに笑い動作を付加したアンドロイドの対話映像の2条件を用意した。それぞれの映像は、アンドロイドの正面から撮影した(図6)。この2条件により、複製動作のみものと複製動作に生成動作を付加したものととの比較ができる。笑い動作は4章で述べたものであり、首のピッチ方向の軸のみによって実装している。首のロール方向の軸とヨー方向の軸は、笑い声の間も操作者の動作をそのまま複製することで、笑い声の部分の動作とそれ以外の動

作が違和感なく滑らかに繋がるようにした。さらにジェミノイドとテレノイドのそれぞれについてその2条件の映像を用意し、表1に示すようなG1, G2, T1, T2の合計4種類の映像を用意した。2つの仮説で提起したアンドロイドの見かけ要因と動作要因による印象への影響を検証するため、このように2×2の4条件を用意した。今回自由度の少ないテレノイドと動作の統制をとるため、動作複製は頭部動作のみとしている。また、笑い

表1 各映像の条件

| 見かけ要因/動作要因 | 複製動作のみ | 生成動作付加 |
|------------|--------|--------|
| ジェミノイド | G1 | G2 |
| テレノイド | T1 | T2 |



(a) ジェミノイド HI-2

(b) テレノイド

図6 実験に用いた映像画像

に必要な不可欠である表情生成については、4条件の全てで行っている。表情生成は、ジェミノイドは頬と口角を上げた表情の生成と口を開ける動作、テレノイドは口を開ける動作のみとしている。口を開ける動作を加えているのは、口唇動作生成のみでは口の開きが小さく、笑った時のように大きく開くことができないからである。これにより、ジェミノイドの方が動かす関節軸が多くなっている。

5.2.2 対話音声について

映像で使用している対話音声は、文学部と理学部の大学生二人による対話である。長さは1分半ほどでその中で操作者(文学部学生)の笑い回数は7回である。

5.2.3 実験手順

具体的な実験手順を図7に示す。

検証映像の前にアンドロイドの映像を見るのは、アンドロイドの見た目やどのように動くかに慣れてもらうためである。この映像も検証映像と同様に、対話中の操作者の動作を再現したアンドロイドの映像である。正面から撮影したもので、操作者は検証映像での操作者と同一である。各被験者は、4条件の映像を評価するが、各映像の対話内容はすべて同一である。そのため、最初の1回だけは、対話内容を初めて聞くことになり、2回目以降は、すでに知っている内容を聞くことになる(1回目ほどで笑いが起こるかかわからないが、2回目以降ほどで笑いが起こるかかわっていることになる)。この点

について統制するため、検証映像を見る前に、映像に使われている音声のみを被験者に聞かせることにした。

音声のみを聞く前に、対話が文学部と理学部の大学生2人によるものであることを説明し、さらに検証映像を見る前に文学部の人間がアンドロイドを遠隔操作していることを説明している。これにより、音声のどちらの人間の声が操作者であるかをはっきりとわかるようにした。映像を見る順番は順序効果を打ち消すようカウンターバランスをとった。最終アンケートでは、対話内容(音声)が十分理解できるものであったかどうかを7段階で評価してもらった。

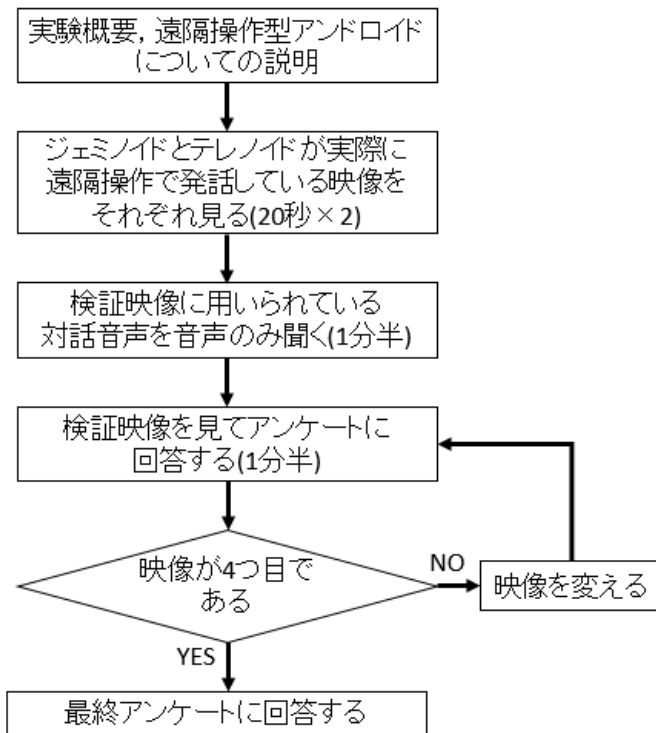


図7 実験手順

5.3 主観評価

アンケート項目は以下のものである。

1. ロボットの動作は自然でしたか
2. ロボットは相手の話を聞いていましたか
3. ロボットは笑っているように見えましたか
4. ロボットの頷き頻度はどうでしたか
5. ロボットの印象は良かったですか
6. ロボットがどのような感情を持っていると感じましたか

項目1から5は7段階評価とし、1, 2, 3の項目については1が”全くそう思わない”, 7が”非常にそう思う”としている。項目4は1が”少ない”, 7が”多い”とし、項目5は1が”悪い”, 7が”良い”としている。いずれも4点がどちらでもない中立の点数となっている。2と4の項目は被験者に笑いの評価をしているという意識を与えないためのダミーの項目である。

また、項目6を複数回答ありの選択肢で用意した。項

目8の選択肢は、嬉しさ・悲しさ・楽しさ・怒り・愉快さ・つまらなさ・嫌悪・驚き・その他(自由記述)とした。

5.4 実験結果

被験者は20歳から23歳までの16名(男性12名, 女性4名), 平均年齢は21.75歳(標準偏差1.0歳)の大学生である。アンケートの項目の内、ダミー設問を除いた1, 3, 5の項目について回答データを解析した。各映像の項目毎の平均値を図8で示す。

図8より3つの項目全てにおいて、ジェミノイドとテレノイド双方で動作生成手法の方が評価平均値が高い結果となった。また、総じてジェミノイドよりもテレノイドの方が評価点が高い。この差が有意であるかどうかを確かめるため、4群のデータに対して見かけ要因と動作要因の2要因分散分析を行った。その結果を表2に示す。

表2 4つの映像評価値の評価項目毎の2要因分散分析結果のp値(*: p<0.05)

| 要因 | 動作の自然さ | 笑っている度合い | 印象 |
|-------|---------|----------|-------|
| 見かけ要因 | 0.0386* | 0.514 | 0.225 |
| 動作要因 | 0.013* | 0.0479* | 0.105 |

表2より項目1の動作の自然さについては、ロボットの種類要因と動作要因の双方に有意水準5%で主効果がみられた。つまり、ジェミノイドよりもテレノイドの方が動作が自然であると言える。項目3の笑っているように見えたかについては、動作要因のみ有意水準5%で有意差がみられた。これにより、ロボットによらず笑い動作を自動生成した方が笑いをより表現できていることが言える。この有意差が見られた項目・要因のグラフを図9に示す。

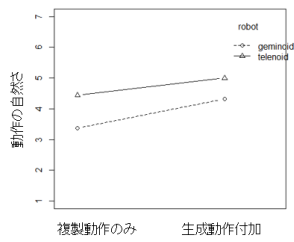
項目6の結果を図10に示す。

生成するピッチ軸の動きは全ての笑いで同じものとしたが、ロール軸とヨー軸の複製動作により、頭部全体としては全く同じ動作となるわけではない。また動作量について、首のピッチ軸の動作量のみ注目すると、今回検証で用いた対話では、笑い声の部分の複製動作での動作量は、7回の笑いの平均値が約2.86度であった。それに対し、提案手法で生成した笑い動作の動作量は約7.14度であり、複製動作よりも動作量が大きいことがわかる。

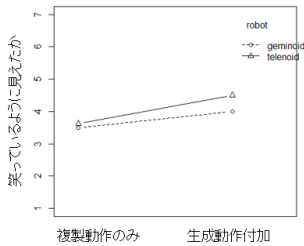
また、最終アンケートで対話音声十分理解できたかを7段階で評価してもらった結果、被験者16人(後述)の平均値が6.25であったことから音声の内容自体には評価する上で問題はなかったと考えられる。

5.5 考察

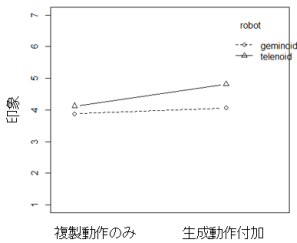
今回の評価実験により、遠隔操作型アンドロイドの笑いについて、操作者の動作複製に加えて、笑い声に同期した笑い動作を自動生成することが動作複製のみよりも笑いの表現について有効であることが確認でき、仮説1を検証できた。笑う時に表情や動きが乏しかった不自然さを改善できたと言える。特に、テレノイドについては笑っているように見えたかの評価項目の点数が従来手法ではマイナスであったところから、提案手法ではプラス



(a) 動作の自然さ

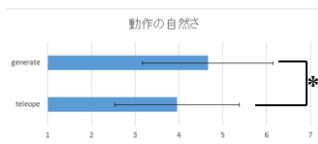


(b) 笑っている度合い

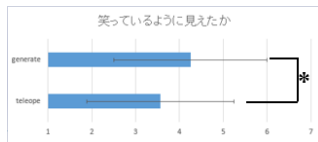


(c) 印象

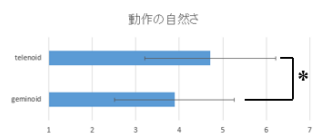
図 8 各項目の 2 要因のプロット図



(a) 動作の自然さ (動作要因)



(b) 笑っている度合い (動作要因)



(c) 動作の自然さ (見かけ要因)

図 9 2 要因分散分析で有意差が見られた項目・要因の棒グラフ

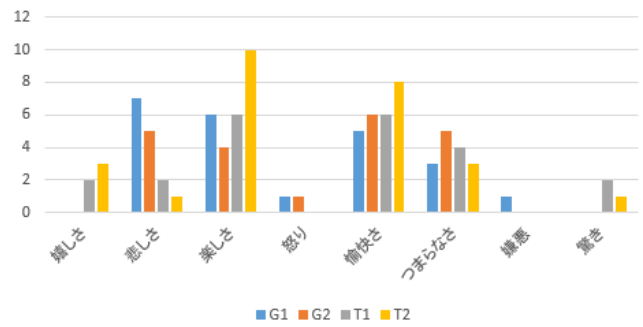


図 10 ロボットが持っていると感じた感情の回答人数

に転じており、表情が全くない不自然さを解消できたと考えられる。また図 10 の感情について、笑いに関係があると考えられる嬉しさ・楽しさ・愉快さの回答数が T1 よりも T2 で多くなっている。笑いにも様々な感情が考えられる(楽しい笑いだけでなく、卑下するような笑いもある)が、この結果により、今回生成した動作が、話の内容に一致した楽しさ、愉快さが伝わる表現であったことがわかる。

ジェミノイドは頬と口角による表情を生成したため、テレノイドよりも笑っているように見ると考えられるが、映像 G1 の評価は 1, 3, 5 の 3 項目共中立点 (4 点) を下回る点数となっている。感情の回答を見てみると、悲しさの感情を答えた人数が T1 で 2 人、T2 で 1 人であるのに対し、G1 では 7 人、G2 では 5 人であった。これは、ジェミノイド HI-2 のデフォルトの表情が険しい表情であり、また変化も乏しかったことから笑っているように見えなかったと考えられる。また、今回の検証映像では大学生の対話音声を用いたことで、ジェミノイドの見かけとの不一致が生じている。印象の自由記述回答では、“学生のように見えない”という回答があり、この見かけと音声の不一致が、見かけ要因の主効果に影響しているとも考えられる。

ジェミノイドについて自然さの得点を見ると、提案手法により自然さが向上しているとは言え、なお得点は中立点 (4 点) 以下である。これは、今回生成した笑い動作ではジェミノイドの見かけから期待される動作として不十分であったことが考えられる。つまり、生成した笑い動作は、テレノイドにおいては被験者の期待を満たしていたが、ジェミノイドにおいては期待を満たしていなかったということである。これにより、期待される動作が見かけによっても異なり、生成すべき笑い動作も異なる可能性があることが言える。しかし、ジェミノイドの低評価については、上述した声と見かけの不一致という要因も考えられ、笑い動作の不十分さだけが要因とは限らない。よって、今回の結果からは仮説 2 の一部として、必要な笑い動作が異なる可能性があることが言え、他の要因については今後検証していかなければならない課題となる。また、ジェミノイドの笑い動作の改善策としては、頭部のみではなく身体全体を使った振動動作とすることが考えられる。

テレノイドは、その見かけから期待される動作量がジェ

参 考 文 献

ミノイドほど高くないため、誇張した頭部動作で期待を満たし自然な笑い動作にできたと考えられる。特に、その表情のなさを動作で補い表現できたことは、テレノイドのような関節軸数を制限されたロボットの表現として有用であると考えられる。また、ロボットの設計において、見かけを調節することでハードウェアの制約を解決することも考えられる。

今回の実験では生成した笑い動作は1種類のみであったが、システム開発にあたっては音声情報を利用することで笑い声に合わせて生成動作を調整することも考えられる。全ての笑いを同じ表現にした場合、例えば、操作者が愛想笑い程度のつもりで笑ったものが、アンドロイド側で非常に笑っているように見えてしまうと、操作者の意図と異なり不適切となる。よって、笑い声を検出するだけではなく分類し、そのレベルに応じた表現のマッピングが必要となる。

今回提案したような遠隔操作型ロボットの動作自動生成手法による遠隔操作の半自律化は、アンドロイドの動作を改善できる利点があるが問題点も存在する。アンドロイドの遠隔操作では、操作者がアンドロイドを自分の身体のように感じられる身体感覚転移と呼ばれる現象がある [5]。この現象は、操作者が自分と同様に動くアンドロイドを見ることで、アンドロイドを自分の身体の一部であるかのように感じられる現象であり、これにより操作者も対話者と対面しているかのように感じることができる。しかし、提案手法によりアンドロイドが操作者と少しでも異なる動作をすることで、この操作者の身体感覚転移が弱まる可能性がある。よって開発したシステムを評価する際は、対話者からの評価のみではなく操作者の評価も必要となる。今後の研究では、どこまで動作を自動生成しても操作者の身体感覚転移を損なわないのか、そのバランスも含めて遠隔操作の半自律化についても研究を進めていく。

6. 結論

本論文では遠隔操作型アンドロイドにおいて、制限された動作自由度の下、音声と一致するアンドロイドの自然な動作を自動生成することの有効性を述べ、笑いを一例として検証を行った。音声と一致する動作の生成手法として、アンドロイドで使用できる自由度だけで誇張した動作を生成する手法を提案した。生成した笑い動作は操作者と異なる誇張された動作であったにもかかわらず、実験では、自然さや笑っているように見える度合いについて従来手法より高い評価を得られたことから、本提案手法の有効性が示された。

また、見かけが異なるアンドロイドでは、生成する笑い動作を変えなければならない可能性があることがわかり、今後の課題として見かけの違いによる影響を検証していかなければならない。

さらにこれらの結果を踏まえ、今後の研究としてリアルタイムに笑い声を検出し笑い動作を生成するシステムの開発を目指す。

謝辞

本研究の一部は、JST,CREST および JST,ERATO の一環として行われたものである。

- [1] 西尾修一, 石黒浩, “人として人とつながるロボット研究” 電子情報通信学会誌, vol.91, no.5, pp.411-416, 2008
- [2] Kristine L.Nowak, Frank Biocca, “The Effect of the Agency and Anthropomorphism on Users’ Sense of Telepresence,Copresence,and Social Presence in Virtual Environments”, 2003
- [3] 北爪剛志, 脇田龍平, 館野圭, 今野俊博, 近藤邦雄, “アニメーションの動作分類と誇張制御手法”, 日本図学会, vol.41, no.1, pp.187-190, 2007
- [4] 中川志信, “人間共存型ロボットにおける最適なモーションデザイン基本法則の抽出”, 2013
- [5] 渡辺哲矢, 西尾修一, 小川浩平, 石黒浩, “遠隔操作によるアンドロイドへの身体感覚の転移”, 電子情報通信学術論文誌, vol.94, no.1, pp.86-93, 2011
- [6] C.T. Ishi, C. Liu, H. Ishiguro, and N. Hagita, “Evaluation of formant-based lip motion generation in tele-operated humanoid robots,” IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.2377-2382, 2012
- [7] Mitsuhiro Kobayashi, Kunio Kondo, Hisashi Sato, “Emphasized Expressions Using Motion Filter in Creating Animation”, Proceedings of the 8th ICECGDG Conference vol.2, pp.451-454, 1998
- [8] 初山和秀, 近藤邦雄, “3D キャラクタアニメーションのための動作誇張モデル”, 情報処理学会第 64 回全国大会, pp.847-848, 2002
- [9] Koie Yoshiyuki, Kondo K. and Toshihiro K, “Motion Emphasis Filter for Making Mental Motion of 3D Characters”, Siggraph 2004 Sketch, 2004
- [10] 館野圭, 近藤邦雄, 今野俊博, “タイミング制御による誇張動作生成手法”, 日本図学会 2006 年度大会論文集, 2006
- [11] Obayashi Syoichi, Kondo K, Konma T, “Non-Photorealistic Motion Blur for 3D Animation”, Siggraph 2005 Sketch, 2005
- [12] 境くりま, 石井カルロス寿憲, 港隆史, 石黒浩, “発話者の音声に対応する動作生成と遠隔操作ロボットへの動作の付加効果”, 人工知能学会, 2014