

TV 雑談型対話ロボットのためのユーザとの 継続対話を目的とした雰囲気共有

西村 祥吾¹ 神原 誠之¹ 萩田 紀博^{1,2}

Shogo Nishimura¹, Masayuki Kanbara¹, and Norihiro Hagita^{1,2}

¹ 奈良先端科学技術大学院大学 情報科学領域 環境知能学研究室

¹ Nara Institute of Science and Technology, Information Science, Ambient Intelligence Laboratory

² 国際電気通信基礎技術研究所 知能ロボティクス研究所

² Advanced Telecommunications Research Institute International, Intelligent Robotics and Communication Laboratories

Abstract: 対話ロボットとユーザとの日常的な対話を促進するための取り組みとして、TV を見ながらともに雑談することが可能なロボットについての研究が行われている。

このロボットに対し、ユーザと場の雰囲気に応じたロボットの振舞いを決定する雰囲気共有を適用することにより、ユーザの継続対話意欲の促進を目指す。今回は主に笑いの雰囲気、盛り上がる雰囲気を中心に、雰囲気共有の評価を行う。

1 はじめに

未婚化やライフスタイルの変化に伴い、年齢層に関わらず独居世帯の数が増加している[1]。独居世帯者は日常生活における会話機会が減少する傾向にある。日常会話が不足すると、脳機能の低下による認知症やうつ病の原因、また生活環境の悪化による身体機能の低下の原因になる恐れがある。そのため、日常会話が不足している人の会話機会を増やすための対策が重要な課題となっている。

このような課題の対策として、日常的に対話ロボットを利用してコミュニケーションを活性化させる取り組みが従来から数多く研究されてきた。神田ら[2]は、ロボットが社会的な存在として人々に受け入れられるためには人とロボットが友好的な関係を築く必要があると述べている。このように日常的に対話ロボットを利用してもらうためには人とロボットの友好的な関係構築が重要であると考えられる。

これに対し我々はユーザに継続的に対話ロボットを利用してもらうために、ユーザの継続対話意欲の向上を目的とした研究を行っている。TV 視聴時に共に雑談し合うことが可能な TV 雑談型対話ロボット（以下、TV 雑談ロボット）を用いて、ユーザのロボットに対する対話意欲向上を図った[3]。また、スポーツ中継のようなテレビ番組において、得点シーンなどの場面で複数体のロボットによりパブリックビューイングのような盛り上がり表現することで、ユーザに共感の印象を与え対話意欲を向上させた

[4]。このように、TV 視聴時における特定シーンにおいて、ロボットがユーザと感情を共有しているように振舞うことで、ユーザのロボットに対する対話意欲が向上することが示唆されている。そこで本誌では、場の雰囲気に応じロボットにユーザと感情を共有するための振舞いを決定する機能を“雰囲気共有”と定義し、TV 雑談ロボットに特化した雰囲気共有について説明する。また、雰囲気共有の中で特に、盛り上がる雰囲気および笑いの雰囲気に注目して対話ロボットに適用した際の印象評価を行う。

2 関連研究と本研究の位置付け

2.1 対話ロボットの印象評価

ロボットの印象評価についての研究は数多く行われている。神田ら[5]は、ソーシャルロボットに関する研究動向から、ロボットが実世界の情報を伝える際は CG で描画されたエージェントよりも、存在感、社会的な印象、信頼される、有能であると思われる、といった点で有用であり、人間に働きかける上でも効果が高いと報告している。一方で、角ら[6]は、モチベーションを維持するための説得を行う場面においてはキャラクターエージェントの方がロボットよりも影響力が強いと結論づけている。これは、目的によりロボットとエージェントの優位性が変動することを示唆している。

2.2 TV 雑談型対話ロボット

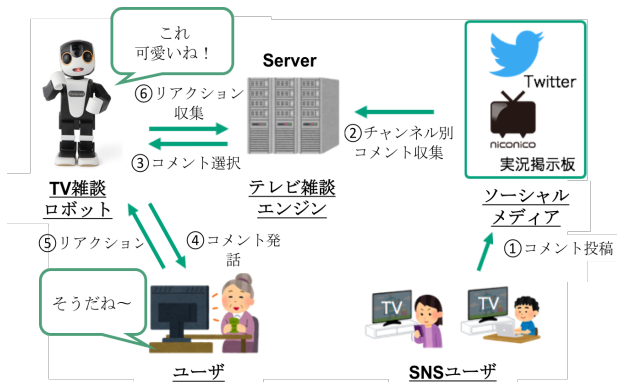


図1 TV 雑談システム

TV 雑談ロボットの開発の経緯として、ロボットの魅力的な発話のため、ユーザが SNS コメントを用いることでユーモアを含んだ発話を行う TV 雑談ロボットを提案した。このロボットはユーザと共に TV 視聴を行いながら TV の内容についてロボットが発言することで、ユーザの発話を促すことを目的に開発され、ユーザの継続的な対話を促す[3]。図 1 に TV 雑談システムの概要図を示す。また、ユーザとの共感や一体感を演出するための機能として、複数体の TV 雑談ロボットにおける盛り上がり共有を提案した。複数体のロボットの服装および発話音声に個性を持たせた上で、ユーザと TV 番組の盛り上がり場面を共有することで、ユーザに共感の印象を与えることができ、ユーザの対話意欲が向上することを示した。しかし、複数体のロボットにおける盛り上がり共有の印象比較は行われていない。

2.3 本研究の位置付け

本研究では不特定多数の SNS ユーザが視聴中のテレビ番組に対して感じる感情が、TV 雑談ロボットユーザと一致すると仮定した際の雰囲気共有の実装を行う。雰囲気共有で扱う感情としては、盛り上がりおよび笑いを扱う。雰囲気共有を盛り上がり・爆笑環境において適用した TV 雑談ロボットに対し、アピランスの異なるロボットを用いた場合の被験者の印象の差を比較することである。表 2 に示すよう

表 1 雰囲気共有によるロボットの振舞いと要素

振舞 要素	ロボットの発話・モーション		
	When	What	How
適用区間	○		
レベル		○	○
発話文		○	
韻律情報			○
感情			○

に、2 種類の雰囲気（盛り上がり・爆笑）に対し、雰囲気共有適用の有無について 2 条件、ロボットのアピランスにおいて 2 条件（バーチャル・フィジカル）の計 4 条件についてそれぞれ比較を行う。

3 雰囲気共有の実装

3.1 雰囲気共有の定義

本誌では対話ロボットとユーザが共有する空間においてユーザが喚起するであろう感情を、対話ロボットがリアルタイムに追跡し、共感している印象をユーザに与えることと定義する。雰囲気には怒り、悲しみ、恐怖、興奮など様々なものが考えられるが、今回扱う雰囲気には、スポーツ中継で発生することが多い興奮（盛り上がり）とバラエティ番組などで見られる誘い笑い（爆笑）を扱う。各雰囲気を構成すると考えられる要素とそれに関連して決定される振舞いの要素を表 1 に示す。これらの要素を基に、ロボットの振舞いを大きく発話およびモーションに分類した。各要素を推定するためのソースとして、ユーザ情報を取得する方法が挙げられるが、今回はユーザ以外の情報として、TV 雑談システムにて収集される SNS コメントを用いる。ユーザ以外の情報から雰囲気共有によるロボットの振舞いを決定することで、ユーザの負担を軽減できるが、大衆の意見を用いて場の雰囲気を推定するため、ユーザ個人に特化するのではなく、汎化された雰囲気共有を試験的に実装する。これを TV 雑談ロボットに対し適用することで、雰囲気共有がユーザに与える印象の評価を行う。

3.2 ロボットの振舞いを決める要素の推定

本誌では雰囲気共有を TV 雑談ロボットにあたり、盛り上がりおよび爆笑の 2 条件を扱う。ロボットの振舞いとして、主に発話・モーションのタイミングおよびレベルを決定する方法について述べる。それらの要素を基にロボットの振舞いを設計する方法について述べる。

3.2.1 振舞いを行うタイミング

今回は TV 雑談ロボットユーザが放送中の TV 番

表 2 雰囲気共有評価時の条件

雰囲気	アピランス	雰囲気共有なし	雰囲気共有あり
盛り上がり	バーチャル	1 対 1	1 対多
	フィジカル	1 対 1	1 対多
爆笑	バーチャル	1 対 1	1 対多
	フィジカル	1 対 1	1 対多

組に対する感情の変化が、SNS 上にリアルタイムにコメントを投稿する不特定多数の SNS ユーザの感情変化と一致すると仮定して、雰囲気共有適用時のロボットの表現を決定する。使用する SNS リソースには、SNS ユーザのコメントをリアルタイムに収集するという点で安定性のあるニコニコ実況を用いる。

ロボットが盛り上がり表現を行う場合、SNS 上に投稿されるコメント数が一斉に増加するタイミングで盛り上がり区間の判定を行う。

中澤ら[7]は、Twitter のコメント機能である Tweet 数から TV 番組中の重要シーンを検出する手法を提案している。この手法を参考に、TV 番組放送時にニコニコ実況に書き込まれたコメント数 C_i を 5 秒ごとに取得し、以下の式(1)(2)により盛り上がり区間の判定を行う。ここで μ , ρ は、判定区間から 15 分前までの盛り上がりでないデータ群およびコメント数 C_i の増加量が負であったデータ群の平均と標準偏差である。

$$Threshold = \mu + 2\rho \quad (1)$$

$$C_i > Threshold : Enthusiasm \quad (2)$$

取得したコメント数が設定した閾値以上の場合、その区間を盛り上がり区間とし、区間内にロボットに盛り上がり表現を行わせるものとする。

一方ロボットが爆笑表現を行う場合は、SNS 上に投稿される特定のコメントに注目する。特定のコメントとして、笑いを表現するキーワード（今回は 'w','W','w','W','笑','草','芝','わろた','ワロタ','わら','吹いた','フイタ','藁' が含まれるコメントとした）数 L_i を、盛り上がり表現と同様に 5 秒ごとに取得し、以下の式(3)(4)により爆笑区間の判定を行う。ここで μ , ρ は、判定区間から 15 分前までの笑いを表すキーワードを含まないデータ群およびコメント数 L_i の増加量が負であったデータ群の平均値および標準偏差である。

$$Threshold = \mu + \rho \quad (3)$$

$$L_i > Threshold : \text{爆笑区間} \quad (4)$$

3.2.2 振舞いを行う際のレベル

3.2.1.で述べた手法により盛り上がり判定された区間においても、盛り上がりの程度に差があると考えられる。そのため、盛り上がりの程度を盛り上がりレベルと定義し、レベルを判定する手法について述べ。我々はこれまでに Twitter において 5 秒間に取得された Tweet 数と基準となる閾値との差から、盛り上がり区間における盛り上がりレベルの判定を行っている[8]。この手法を用いて、以下の式(5)~(8)より盛り上がり区間におけるレベル判定を行う。ここで C_i はニコニコ実況において 5 秒間に書き込まれたコメント数、 $Threshold$ は 3.2.1.の手法により求め



図2 盛り上がり時のロボットの振舞い

た閾値である。

$$C_i < Threshold : Level 0 \quad (5)$$

$$Threshold < C_i < 1.3 * Threshold : Level 1 \quad (6)$$

$$1.3 * Threshold < C_i < 1.6 * Threshold : Level 2 \quad (7)$$

$$1.6 * Threshold < C_i : Level 3 \quad (8)$$

同様に爆笑にもレベルが存在すると仮定し、レベル 0 から 3 の 4 段階に設定した。以下の式(9)~(12)より爆笑区間におけるレベル判定を行う。ここで L_i はニコニコ実況において 5 秒間に書き込まれたコメント数、 $Threshold$ は 3.2.1.の手法により求めた閾値である。

$$L_i < Threshold : Level 0 \quad (9)$$

$$Threshold < L_i < 1.33 * Threshold : Level 1 \quad (10)$$

$$1.33 * Threshold < L_i < 1.66 * Threshold : Level 2 \quad (11)$$

$$1.66 * Threshold < L_i : Level 3 \quad (12)$$

3.3 ロボットの振舞いの決定

本章では、盛り上がり判定手法をもとに盛り上がりを表現する手法について述べる。雰囲気共有に基づく振舞いを行うロボットのアピランスは、ともに人型であり、2 次元のバーチャルエージェントと身体性を伴ったロボットとを用いる。図に示すようなバーチャルエージェントには TV 雑談ロボットをスマートフォンアプリケーションに実装したアプリケーション “Dabelive” (株式会社 amirobo tech)をもとに構成した。身体性を持つロボットには pip 製のうなずきかぼちゃんを用い、肩関節部にモータを挿入することで腕を大きく振る動きができるよう改造した。本誌ではアピランスに関わらず、人とコミ



図3 爆笑時のロボットの振舞い（バーチャル）
コミュニケーションを行う媒体をロボットと表現する。

3.2.2 発話内容の決定

発話内容の決定にはリアルタイムに放送されている番組に対する SNS 上のコメントを収集可能な TV 雑談サーバを用いる。TV 雑談サーバが収取する SNS コメントのソースとしては、高いリアルタイム性が安定して確保できるニコニコ実況を主に用いる。収集された各局の番組に対するコメントから、過激なコメントを除外したコメント群からランダムに選択した 1 つのコメントを 7 秒おきにクライアント（TV 雑談ロボット）に送信し、音声合成によってコメントを発話させる。本研究では、ニコニコ実況からコメントを取得する際に 3.2 で述べた手法を用いて、各種区間とレベルを判定し、その結果をコメントサーバに記録する。コメントサーバは区間とレベルに応じて、クライアントに送信するコメントを雰囲気に合わせて変更する。

盛り上がりの例では、盛り上がり区間中複数のロボットに対して、たたみかけるように連続して発話を行わせることで興奮状態を表現する。クライアントに対し発話コメントをレベルによって以下のように変更する。

- **Level 0** : 7 秒につき、1 体のロボットが発話 1 コメントを発話
- **Level 1** : 7 秒につき、2 体のロボットがそれぞれ 1 コメントずつ発話
- **Level 2** : 7 秒につき、3 体のロボットがそれぞれ 1 コメントずつ発話
- **Level 3** : 7 秒につき、5 体のロボットがそれぞれ 1 コメントずつ発話

バーチャル環境におけるロボットの場合は、図 2

に示すようにレベルに応じて発話するロボットの数を増やす。身体性を伴うロボットにおいても同様にロボットの台数を増やすことで対応する。

一方、爆笑区間中のロボットの振舞いでは、通常の 7 秒間に 1 回発話を行うロボットに加え、笑い声を発話するロボットの数と笑い声をレベルに応じて変更していく。笑い声は事前に録音された人の笑い声を、主観的に大笑い・中笑い・小笑いに分類した。レベルごとに以下の変更を行う。

- **Level 0** : 7 秒につき、1 体のロボットが 1 コメントを発話
- **Level 1** : 7 秒につき、1 体のロボットが 1 コメントを発話し、別の 1 体が小笑いを発声
- **Level 2** : 7 秒につき、1 体のロボットが 1 コメントを発話し、別の 2 体がそれぞれ中笑、小笑いを発声
- **Level 3** : 7 秒につき、1 体のロボットが 1 コメントを発話し、別の 4 体がそれぞれ別の大笑いを発声

バーチャル環境でのロボットの振舞いは図 3 に示す通りである。笑い声の発声タイミングは、3.2.1 で述べた通りであるが、複数のロボットが笑い声を発生する場合は各ロボットの発生タイミングを 2 ミリ秒ずつずらすことで、笑い声が重なりエコー掛かることがないようにした。

4 盛り上がり表現の評価実験

本節では、TV 番組の盛り上がり場面においてバーチャル・フィジカル両環境においてロボットを用いて盛り上がりの表現を行い、ユーザのロボットに対する印象の差を比較について詳述する。被験者にはアピランスの異なるロボットと共に TV 番組を視聴してもらい、各番組の視聴後にアンケートに回答してもらう。

4.1 実験構成

4.1.1 映像コンテンツ

映像コンテンツは盛り上がるタイミングが明確であり、かつ盛り上がりの種類に個人差が少ないサッカー中継とした。TV 番組には 2018 年 10 月 16 日に放送されたサッカーキリン杯 2018 の日本対ウルグアイ戦を用いた。番組放送時に TV 雑談エンジンによって収集された SNS コメントを用いて、3 章にて述べた手法により盛り上がり区間と盛り上がりレベルを検出し、盛り上がりと判定された場面の中の 1 つを実験映像コンテンツとした。

4.1.2 実験に使用した TV 雑談ロボット

本実験で用いたフィジカルな TV 雑談ロボットには、これまでの実験[10]を参考に、うなづきかぼちゃ



図4 盛り上がり評価実験時の様子
(左：フィジカル1対多，右：バーチャル1対多)

んを使用する。雰囲気共有時は5台のロボットが識別可能なように、音声合成には VoiceText Web API を用いて異なる音声モデルを設定し、見た目での識別も考慮し異なる色の服装にした。

また、本実験では実験時の再現性を保つため、時系列に並べたコメントを事前に作成し、7秒間隔でTV雑談ロボットがコメントを読み上げるようデザインした。コメントは、本実験で使用する映像コンテンツが放送されていた間にTV雑談エンジンにて収集されたコメントを用いる。

4.1.3 実験条件と評価項目

11名の被験者(20～30代男女)にフィジカル・バーチャルの2条件および盛り上がりの雰囲気共有あり・なしの計4条件のロボットと共にTV視聴を行ってもらった際の印象をアンケートによって評価する。4.1.1にて述べた映像を事前に1度視聴してもらい、映像内容を把握してもらう。その後、

- ・ 雰囲気共有なし条件・バーチャルロボット
- ・ 雰囲気共有あり条件・バーチャルロボット
- ・ 雰囲気共有なし条件・フィジカルロボット
- ・ 雰囲気共有あり条件・フィジカルロボット

の各条件で計4回視聴してもらう。図4に実験時の様子を示す。視聴順序による影響を考慮し、被験者ごとに各条件にて視聴する順序を変更し、カウンターバランスをとった。視聴後は逐次以下5項目に対して7段階のリッカード尺度による評価値を付けてもらう。

(1:全くそう思わない, 7:非常にそう思う)

- ・ ロボットは盛り上がっているように感じましたか？
- ・ ロボットは親しみやすかったですか？
- ・ ロボットの盛り上がり違和感は無かったですか？
- ・ ロボットを継続的に利用したいと感じましたか？
- ・ ロボットの発話内容に共感できましたか？

4.2 実験結果

7段階で行なった主観評価について、各項目の平均スコア、標準偏差を図5に示す。各項目に対し、

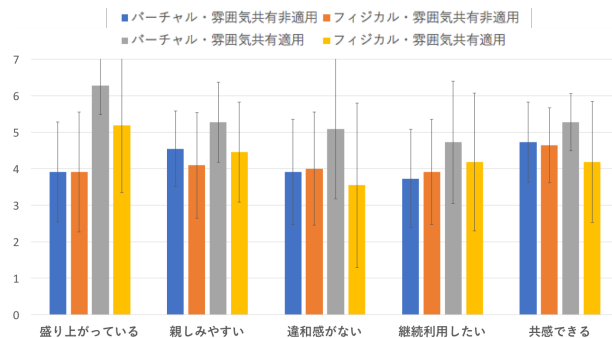


図5 盛り上がり評価実験に使用した
アンケートスコアと標準偏差

- ① 雰囲気共有なし条件・バーチャルロボットと雰囲気共有なし条件・フィジカルロボット
- ② 雰囲気共有なし条件・バーチャルロボットと雰囲気共有あり条件・バーチャルロボット
- ③ 雰囲気共有なし条件・フィジカルロボットと雰囲気共有あり条件・フィジカルロボット
- ④ 雰囲気共有あり条件・バーチャルロボットと雰囲気共有あり条件・フィジカルロボット

の4つの組に分けてt検定を行なった。その結果を図6に示す。ここで、*： $p < 0.05$ (有意差), †： $0.05 < p < 0.1$ (有意傾向)とする。

①の組み合わせに関して、全ての項目で有意差は認められなかった。②に関して、「ロボットは盛り上がっていたか」の項目について有意差が認められた。他の項目に関しても有意差は認められないものの雰囲気共有あり条件の方が好印象であった。③に関して、「ロボットは盛り上がっていたか」について有意傾向は認められたものの、②の比較ほど差は見られなかった。④の比較に関して、「ロボットは盛り上がっていたか」、「ロボットの盛り上がり違和感は無かったか」、「ロボットの発話内容に共感できたか」の3項目で有意差が認められた。残りの項目に関しても、複数体のロボットの方が好印象であった。

4.3 考察

結果から、本実験における盛り上がり共有の場面では、雰囲気共有が適用されていないロボットを用いるよりも、適用されたロボットを用いた方がユーザにより盛り上がっている印象を与えることがわかった。主観評価における自由記述をまとめると、バーチャル・フィジカルロボットの両方の場合において、雰囲気共有適用時には「興奮するようなシーンで畳み掛けるように発話していたから」、「ワイワイしている感じあった」とあり、1体の場合では「淡々と話していたから」、「盛り上がりそうなところでも一定の周期で発言しているから」とある。つまり、

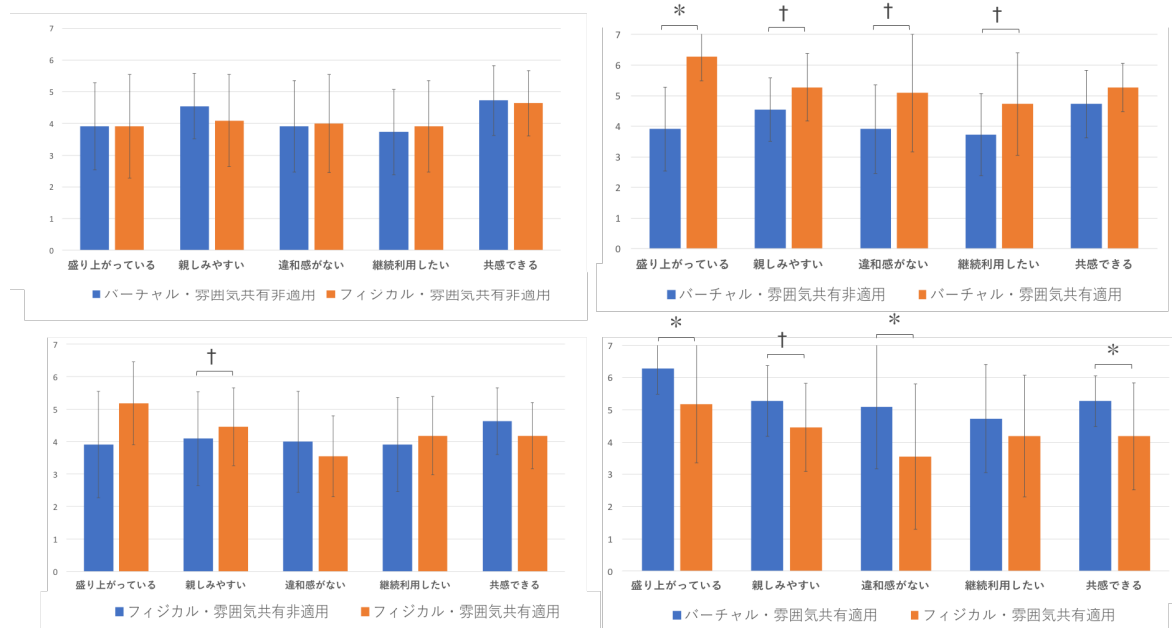


図6 盛り上がり評価実験における各条件の組み合わせごとの比較

(左上：雰囲気共有なし条件・バーチャルロボットと雰囲気共有なし条件・フィジカルロボット，右上：雰囲気共有なし条件・バーチャルロボットと雰囲気共有あり条件・バーチャルロボット，左下：雰囲気共有なし条件・フィジカルロボットと雰囲気共有あり条件・フィジカルロボット，右下：雰囲気共有あり条件・バーチャルロボットと雰囲気共有あり条件・フィジカルロボット)

ロボットのアピランスに関わらず，盛り上がりの雰囲気共有で表現した発話者数と発話頻度の増加がユーザへの盛り上がり印象の向上の要因になっていると考えられる。

ロボットのアピランスの比較に関して，雰囲気共有なしの場合は，バーチャルなロボットに対し「字幕があると?」とか「!」「w」を確認できるので，ただ音声聞くだけの時と比べて印象が変わった，「文字で表示される点はよかった」といったポジティブな意見と，「あまり表情が変わらなかったり，物理的に人形がないのがやや不満だった」といったネガティブな意見が見られた。ロボットに対しては「一緒にテレビを見てくれる感じがしたから」，「膝に乗せて一緒にテレビを見ていたい」といったポジティブな意見と，「字幕がないため，何を話しているのかあまりわからないところがあった」といったネガティブ意見があった。これらの意見から，ロボットは発話に合わせた字幕表示が高評価を受け，実体がないことに低評価を受けている。また，ロボットは実体があることに高評価を受け，字幕がないことに低評価を受けている。これらの点は，相反するものであり，評価が分かれたために結果として，条件間のスコアに有意差が認められなかったと考えられる。

次に，雰囲気共有適用時のロボットのアピラン

スの比較に関する考察を述べる。「盛り上がっているように感じたか」，「親しみやすかったか」の両項目において，アピランスの違いに関わらず「盛り上がっている感が出ていた」というコメントがある一方で，バーチャルロボットでは「それぞれのキャラがピコピコ動いていたため」のようなスマートフォン上でロボットの数が変化することに関するコメントがあり，ロボットでは「擬音が違和感」，「返事や相槌がなかったところがすこし違和感を感じる」のような発話や動作に関するコメントがあった。これらのことから，バーチャルなロボットが盛り上がり時に追加で出現することが，盛り上がり時とそうでない時の視覚差を生み，ユーザが好印象を持つ要因になっていると考えられる。またフィジカルなロボットは，バーチャル条件では感じられなかった発話の違和感や，返事や相槌といった動作に対する違和感が生じていたことから，実体を持つことでユーザがロボットに対する期待感を増幅させ，これにより違和感が生じ，スコアが低くなる要因となったと考えられる。

5 爆笑表現の評価実験

次に，TV 番組の誘い笑いを喚起する場面においてバーチャル・フィジカルの両環境においてロボットを用いて爆笑表現を行い，ユーザのロボットに対す

る印象の差を比較について詳述する。4 章での評価方法と同様、被験者にはアピランスの異なるロボットと共に TV 番組を視聴してもらい、各条件での番組視聴後にアンケートに回答してもらう。

5.1 実験構成

5.1.1 映像コンテンツ

映像コンテンツは比較的万人に対して明確に笑いを誘う場面が多く見られた、2018 年 12 月 30 日に放送されたバラエティ番組“アメトーーク!”中の“アメトーーク大賞 2018 ビジュアル映像部門”の一部分を用いる。番組放送時に TV 雑談エンジンによって収集された SNS コメントを用いて、3 章にて述べた手法により爆笑区間とレベルを検出し、爆笑区間と判定された場面が含まれたシーンを実験映像コンテンツとした。なお今回は映像コンテンツの時間とロボットが発話を行うタイミングが被験者によって同期したものとなるよう、図 7 に示すように実験映像コンテンツと各条件のロボットが並んだ状態の映像を視聴することでロボットの評価を行う。

5.1.2 実験に使用した TV 雑談ロボット

4.1.2 と同様、実験で用いたフィジカルな TV 雑談ロボットには、これまでの実験[10]を参考に、うなづきかぼちゃん使用する。雰囲気共有時は 5 台のロボットが識別可能なように、音声合成には VoiceText Web API を用いて異なる音声モデルを設定し、見た目での識別も考慮し異なる色の服装にした。また、バーチャルなロボットにおいて、ロボットが笑い声を発生する際、ロボットの表情のみ図 7 に示すような笑顔に変化した状態となる。

ロボットの発話文には、時系列に並べたコメントを事前に作成し、7 秒間隔で TV 雑談ロボットがコメントを読み上げるようデザインした。コメントは、本実験で使用する映像コンテンツが放送されていた間に TV 雑談エンジンにて収集されたコメントを用いる。

5.1.3 実験条件と評価項目

7 名の被験者(20~40 代男女)にフィジカル・バーチャルの 2 条件および爆笑の雰囲気共有あり・なしの計 4 条件のロボットと共に TV 視聴を行ってもらった際の印象をアンケートによって評価する。5.1.1.にて述べた映像を事前に 1 度視聴してもらい、映像内容を把握してもらう。その後、

- 雰囲気共有なし条件・バーチャルロボット
- 雰囲気共有あり条件・バーチャルロボット
- 雰囲気共有なし条件・フィジカルロボット
- 雰囲気共有あり条件・フィジカルロボット

の各条件で計 4 回視聴してもらう。視聴順序による影響を考慮し、被験者ごとに各条件にて視聴する順



図 7 爆笑の評価実験に用いる映像
(上：雰囲気共有あり・バーチャルエージェント、
下：雰囲気共有あり・フィジカルエージェント)

序を変更し、カウンターバランスをとった。視聴後以下 6 項目に対して 7 段階のリッカー尺度による評価値を付けてもらう。

(1:全くそう思わない, 7:非常に思う)

- ロボットはは番組を楽しんでいるように感じましたか？
- ロボットと番組を視聴して楽しいと感じましたか？
- ロボットは親しみやすかったですか？
- ロボットと同じ感情を共有しているように感じましたか？
- ロボットの言動に違和感はなかったですか？
- ロボットを継続的に利用したいと感じましたか？

5.2 実験結果

7 段階で行なった主観評価について、各項目の平均スコア、標準偏差を図 7 に示す。各項目に対し、

- ① 雰囲気共有なし条件・バーチャルロボットと雰囲気共有なし条件・フィジカルロボット
 - ② 雰囲気共有なし条件・バーチャルロボットと雰囲気共有あり条件・バーチャルロボット
 - ③ 雰囲気共有なし条件・フィジカルロボットと雰囲気共有あり条件・フィジカルロボット
 - ④ 雰囲気共有あり条件・バーチャルロボットと雰囲気共有あり条件・フィジカルロボット
- の 4 つの組に分けて t 検定を行なった。その結果を

図 8 に示す。ここで、*： $p<0.05$ (有意差)，†： $0.05< p<0.1$ (有意傾向)とする。

①の組み合わせに関して、ロボットとの共感について有意傾向が見られたが、有意差は得られなかった。一方で継続利用意欲の促進に関して有意差が認められ、バーチャルな見た目のロボットがフィジカルなそれに比べ高いスコアを得ることが示された。

②に関して、「ロボットが楽しんでいる」の項目について有意差が認められた。他の項目に関して有意差は認められないものの雰囲気共有あり条件の方が一貫して好印象であった。

③に関して、「ロボットが楽しんでいる」、「共感している」、「継続利用意欲」の項目について有意差が認められた。しかしフィジカルなロボットを用いる場合において、雰囲気共有の有の場合に一貫して有意となる傾向は見られなかった。

④の比較に関して、「ユーザが楽しいと感じる」、「ロボットは親しみやすい」、「ロボットの言動に違和感がない」の 3 項目で有意差が認められた。雰囲気共有を適用した条件においてロボットはバーチャルな見た目の方がよりユーザにとって好印象であることがわかった。また、4 条件のロボットのうち、同じような TV 番組を視聴する際にどの条件のロボットを最も使いたいかという質問を被験者に行ったと

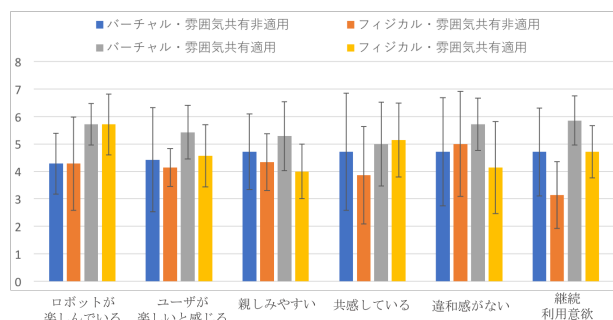


図 8 爆笑表現の評価実験に使用した項目に対するアンケートスコアおよび標準偏差

ころ、7 名中 5 名が雰囲気共有あり条件かつ、バーチャルロボットを用いた場合の条件で利用したいという回答を得た。

5.3 考察

今回の評価実験結果から、雰囲気共有において爆笑を扱う場合、ロボットの評価はバーチャルエージェントであるほうが好ましいことが示された。この理由として、今回の実験では特にロボットの表情が変化すること、発話文章が表示されること、爆笑区間外ではロボットが画面から消えること、などバーチャルロボットにおいてのみ適用される表現が含ま

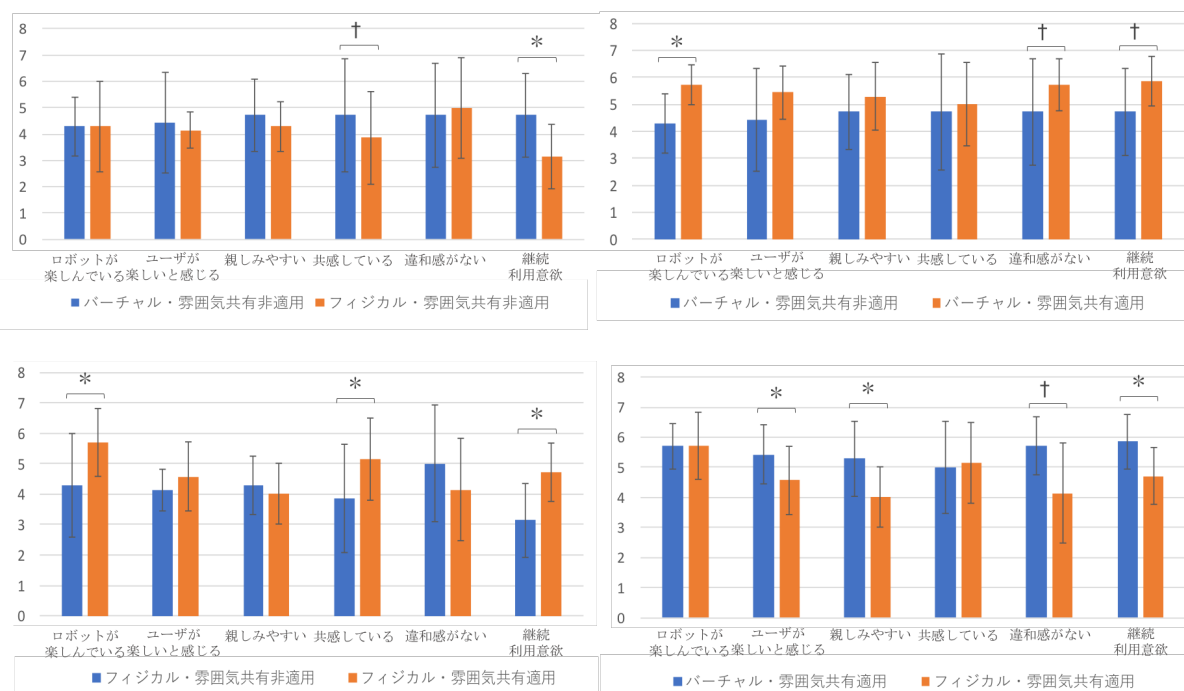


図 9 盛り上がり評価実験における各条件の組み合わせごとの比較

(左上：雰囲気共有なし条件・バーチャルロボットと雰囲気共有なし条件・フィジカルロボット，右上：雰囲気共有なし条件・バーチャルロボットと雰囲気共有あり条件・バーチャルロボット，左下：雰囲気共有なし条件・フィジカルロボットと雰囲気共有あり条件・フィジカルロボット，右下：雰囲気共有あり条件・バーチャルロボットと雰囲気共有あり条件・フィジカルロボット)

れていたことから、このような結果を誘導することになったのではないかと考えられる。

今回実装した爆笑表現について、雰囲気共有を適用した際のロボットについての被験者の自由記述からポジティブな意見として、「自分が面白いと感じる場面でロボットが笑っており、同じ感情を共有しているように感じた」、「ロボットの笑い声に誘われて笑ってしまうことがあった」、「笑い声が聞こえると楽しく感じる」などが得られた。

一方、ネガティブな意見として、「少し笑い声が多すぎる」、「(大笑)などの字幕によってややしらけた」、「ロボット内で笑っていて取り残されてる感があつた」、「笑い声がロボットの見た目とあっておらず、違和感を感じた」、といった意見が得られ、被験者によっては爆笑表現によって継続利用意欲が抑えられてしまう可能性が示された。

以上から、爆笑表現によってロボットに対し共感を感じるユーザが見られ、バラエティ番組のような笑いを喚起する内容のコンテンツと合わせて用いることで、ユーザとの親和性を高めることに貢献する可能性を示した。また、爆笑表現自体に違和感を感じるユーザも見られ、笑い声のバリエーションやロボットの見た目をユーザの好みに合わせるなどといったカスタマイズ性が根本的に必要であることが示された。

6 おわりに

本研究では、TV 番組の盛り上がり場面および爆笑場面において、SNS 上の TV 番組に関連するコメント情報を用いた雰囲気共有機能の開発および評価を行った。雰囲気共有を TV 雑談ロボットに適用した際、ユーザがロボットから受ける印象の差を主観評価により比較した。結果、盛り上がり・爆笑の両条件において、雰囲気共有をロボットに適用した場合、非適用時に比べユーザが雰囲気に対して受け取る感情の変化に、ロボットが追従し、ユーザから共感を得られる可能性が示唆された。盛り上がり表現の評価においては、雰囲気共有非適用時よりも、適用した際のロボットの方が、ユーザが受ける盛り上がりの印象を強めることがわかった。雰囲気共有におけるロボットの視差がユーザに対する盛り上がり・爆笑の印象の向上に影響する可能性が示唆された。一方でロボットが実体を持つことにより、ユーザがロボットに対して持つ期待値が上昇する可能性が示唆された。雰囲気共有を評価する際は、ロボットのアピランスや声質を事前にユーザの好みにあったものに變更可能な状態に設定する枠組みの必要性も示された。

謝辞

本研究の一部は JSPS 科研費 18H03274 および、国立研究開発法人科学技術振興機構研究成果展開事業世界に誇る地域発・実証拠点(リサーチコンプレックス) 推進プログラムの助成を受けたものである。

参考文献

- [1] 内閣府, “高齢者の生活環境,” in 平成 28 年版高齢社会白書, 2016.
- [2] 神田崇行, 佐藤留美, 才脇直樹, 石黒浩, “対話型ロボットによる小学校での長期相互作用の試み,” ヒューマンインターフェース学会論文誌, vol.7, no.1, pp.27-37, 2005.
- [3] H. Minami, et al, “Chat Robot Coupling Machine Responses and Social Media Comments for Communication Conversation,” International Conference on Multimedia and Expo Workshop on Multimedia Services and Technologies for E-health (MUST-EH), 2016.
- [4] S. Nishimura, et. al., “A TV chat robot with time-shifting function for daily-use communication,” in International Conference on Social Robotics, Springer, pp. 516–525, 2017.
- [5] 神田崇行, “HRI におけるソーシャルロボット研究の動向,” 日本ロボット学会誌, vol.29, No.1, pp.2–5, 2011.
- [6] 角薫, 長田瑞恵., “モチベーション維持におけるキャラクターエージェントとロボットの比較,” HAI シンポジウム 2011, 2011.
- [7] 中澤昌美, 帆足啓一郎, 小野智弘, “Twitter によるテレビ番組重要シーン検出およびラベル付与手法,” DEIM Forum 2011, pp. 517-519, 2011.
- [8] 栗田優輝, 木俣大樹, 西村祥吾, 神原誠之, 萩田紀博, “TV 視聴型雑談ロボットのための盛り上がり解析および検証,” 信学技法, vol. 118, no. 184, 2018.
- [9] HOYA 株式会社, “VoiceText Web API”, <https://cloud.voicetext.jp/webapi>.
- [10] 西村 祥吾, 川波弘道, 神原誠之, 萩田紀博, “日常的な対話継続を目的としたテレビを共同視聴するロボットとの盛り上がり共有,” 信学技報, vol. 116, no. 488, LOIS2016-88, pp. 147-152, 2017