



バーチャル笑い場における エンゲージメント向上のための観客アバタの反応設計

Designing the Reactions of Audience Avatar to Enhance Engagement in Virtual Laughter Place

大原嶺, 楊期蘭, 鳴海拓志, 葛岡英明

Ryo OHARA, Chi-Lan YANG, Takuji NARUMI, and Hideaki KUZUOKA

東京大学 (〒 113-0033 東京都文京区本郷 7-3-1, rohara, chilan.yang, narumi, kuzuoka@cyber.t.u-tokyo.ac.jp)

概要: ライブイベントの視聴時における他の参加者の存在は、明示的な相互作用が無いにも関わらず視聴者のエンゲージメントに影響を与えることが知られている。本研究では、劇場を模した VR 空間「バーチャル笑い場」において、観客アバタの視聴覚の手がかりが共在感覚、エンジョイメント、エンゲージメントに与える影響について調査した。実験の結果、観客アバタの反応の視聴覚の手がかりの付与とその一貫性が共在感覚の向上に寄与することが示唆された一方、エンジョイメントとエンゲージメントは条件間で有意な差が見られなかった。実験参加者のコメントより、参加者の笑いとお客アバタの笑いのタイミングが一致することで観客アバタが効果的に作用する可能性が示唆された。

キーワード: 笑い, 他者存在感, エンゲージメント, ノンバーバルコミュニケーション

1. 序論

バーチャルリアリティ技術のエンタテインメント応用が進み、バーチャル空間におけるライブイベントの開催事例が増えている。例えば、メタバース内でお笑いライブを視聴できる吉本興業社の月面劇場¹や、音楽ライブイベント SANRIO Virtual Festival²は人気を博している。これらの取り組みは、従来のライブイベントの物理的な制約を取り払い、自由度の高い演出や遠隔参加を可能とする点で意義深い。

しかし、遠隔ライブイベントでは他の参加者との共在感覚(ソーシャルプレゼンス)が希薄となり、エンゲージメントが低下するという問題が指摘されている [1]。対面のライブ会場では参加者間で微細な反応が非明示的に共有されており、そうした無自覚的な相互作用が共在感覚を強化して情動伝染や場の一体感を生み出すために、観客全体のエンゲージメントが向上すると考えられる。他方、現状の遠隔ライブイベントでは、多くの場合、参加者のアバタには参加者自身の表情や仕草、笑い声などのライブコンテンツに対する自然な反応が反映されない。そのため、バーチャル空間におけるライブイベントでエンゲージメントの低下を防ぐには、共在感覚や情動伝染を効果的に起こせるように、自分以外の参加者のアバタ(本研究では観客アバタと呼ぶ)に適切な反応を付与することが有効であると考えられる。

そこで本研究では、劇場を模したバーチャル空間である「バーチャル笑い場」において、共在感覚とエンゲージメントを高めるために、ライブコンテンツに対する観客アバタ



図 1: 本研究で実装したバーチャル笑い場

の反応をどのように設計すれば良いかを明らかにすることを狙う。本稿では、観客アバタの発するノンバーバルな反応として、視覚の手がかりである表情や動作と、聴覚の手がかりである笑い声の有無を制御できるシステム(図 1)を利用した実験の結果に基づいて、他者存在感、笑いによるエンジョイメント、エンゲージメントを高める上で、どのように視聴覚の手がかりを用いることが有効かを議論する。

2. 関連研究

2.1 遠隔ライブイベントにおける共在感覚の設計

遠隔ライブイベントにおいて共在感覚を強化することは、エンゲージメントを向上させる上で重要である。Webb ら [1] は、物理的な共存空間が無い場合には参加者間の相互作用が希薄となり、ライブイベントへのエンゲージメントが低下すると指摘した。エンゲージメントの低下はエンジョイ

¹月面劇場 <https://x.fany.lol/getsumen/>

²SANRIO Virtual Festival 2024 in Sanrio Puroland <https://v-fes.sanrio.co.jp/2024>

メントの低下に繋がる [2]. 対策として、ソーシャルメディアを併用することで共在感覚を高める事例 [3] がある。

バーチャル空間におけるライブイベントでは、観客アバタを活用して参加者体験を高める取り組みが存在する。特に音楽ライブイベントにおける観客アバタの動作に関する検討が進められており、参加者との動作の同期 [4] や、機械学習による楽曲のビートからの動作生成 [5] により、臨場感や共在感覚を高める手法が提案されている。ただし、こうした手法が音楽以外のコンテンツにも有効かは定かではない。

2.2 笑いの伝染効果

他者の特定の感情表出を知覚することで、自分自身も同じ感情を経験する現象を情動伝染という。情動伝染に効果的な感覚は情動の種類によって異なるが、De Weck ら [6] は笑い刺激を3つの感覚モダリティ (視覚, 聴覚, 視聴覚) で提示する実験により、視聴覚条件で最も伝染性が高いこと、視覚条件よりも聴覚条件で伝染性が高いことを示した。

笑いの伝染性を工学的に応用する研究も進められている。福嶋ら [7] は、笑いの初期動作を検出してラフトラックを同調的に再生することで、笑いを増幅させるためのシステムを提案した。Suzuki ら [8] は、遠隔会議システムにおいて画像変形技術により笑顔の伝染を擬似的に再現し、ポジティブな心理効果を誘発する手法を提案した。本研究も、観客アバタからの笑いの伝染に着目してバーチャル笑い場におけるエンゲージメントを向上させるという点でこうした流れの上に位置する。笑いの伝染に効果的と考えられる観客アバタの笑い反応の視聴覚の手がかりを操作し、これらが視聴者のエンゲージメントに与える影響を評価する。

3. 観客アバタの設計

3.1 概要

ヒトの笑い反応は主に (a) 表情, (b) 全身動作, (c) 笑い声によって特徴付けられる [9]. 本研究では、観客アバタの笑い反応を視覚の手がかり (a,b) と聴覚の手がかり (c) に分類し、それらを独立して提示できるシステムを構築した。以下、各手がかりの実装手法について述べる。

3.2 視覚的手がかり

観客アバタの表情は、46種類の顔面筋ユニット (AU) により表情を定量化する Facial Action Coding System (FACS) [10] に基づいて実装された。Radoslaw ら [11] は、笑い表情の FACS 特徴量として頬を上げる (AU6), 口角を上げる (AU12), 唇を開く (AU25) 等を挙げており、この知見に従って笑い表情が作成された。観客アバタと AU の制御による笑い表情の実装は Character Creator 4³により行われた。

観客アバタの全身動作には、Mixamo⁴において配布されている座位の笑い動作が採用された。先行研究では、座位の笑い動作は主に胴体の動き, 肩の回転, 手の動き, 頭部の動きによって特徴付けられるとされており [12], 採用した動作がこれらの条件を満たしていることを確認した。実装

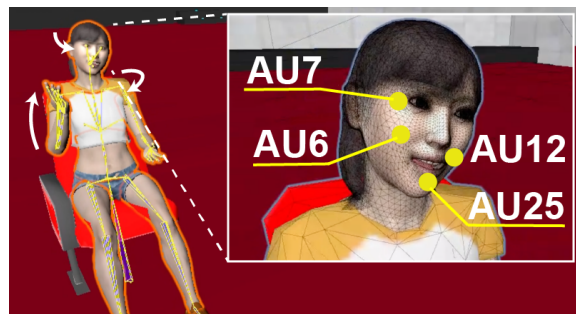


図 2: 観客アバタの笑い反応の視覚的手がかり

した笑い表情および笑い動作を図 2 に示す。

3.3 聴覚的手がかり

観客アバタの笑い声には、Xin ら [13] の作成した日本人の笑い声コーパスの一部が採用された。このコーパスは、インターネット上の動画から抽出された日本人の笑い声が 7,489 種類含まれている。本研究ではそのうち、単一話者の笑いであること、再生時間が採用した笑い動作の時間と同じ 4 秒程度であることを条件に、20 種類の笑い声を選択された。

観客アバタの笑い声の音量は、ユーザと観客アバタの距離に応じて以下のモデルにより減衰させた V が用いられた。

$$V = V_0 \cdot \left(1 - \frac{d - d_{\min}}{d_{\max} - d_{\min}}\right) \quad (1)$$

ここで、 V_0 は初期音量, d は音源からの距離を表す。 d_{\min} は最大音量を維持する距離, d_{\max} は音の届く最大距離であり、今回はそれぞれ 1 m と 20 m に設定された。

3.4 観客アバタの座席配置

バーチャル環境は Unity により作成された。観客アバタは、ヨシモト∞ドーム⁵のステージ I を模した空間に計 30 体配置された。観客アバタの視覚的手がかりを十分に視認できるよう、ユーザのアバタは 3 列目中央部の座席に配置された。ユーザのアバタの腰と両足は固定され、姿勢は頭部と両手部のトラッカより逆運動学に基づき制御された。

4. 実験

4.1 実験デザイン

実験は、視覚的手がかりと聴覚的手がかりを 2 要因とする参加者内計画で行われた。視覚的手がかりは 3 水準で、観客なし (No visual cues; NV) 条件, アイドリングモーションのみを取る観客あり (Low visual cues; LV) 条件, 笑い表情と動作を取る観客あり (High visual cues; HV) 条件が設定された。聴覚的手がかりは 2 水準で、笑い声なし (No auditory cues) 条件, 笑い声あり (Auditory cues) 条件が設定された。順序効果を相殺するため、実験参加者に対する 6 条件の提示順序はラテン方格法により決定された。

4.2 動画刺激とバーチャル観客の笑いタイミング

動画刺激には YouTube チャンネル「ジャルジャルタワー⁶」でライブストリーミング配信された 6 つのコントが採用さ

³Character Creator 4 <https://www.reallusion.com/jp/character-creator/>

⁴Mixamo <https://www.mixamo.com/>

⁵ヨシモト∞ドーム <https://mugendai-dome.yoshimoto.co.jp/>

れた。これは、同一コント師による作品に統一することで動画間の平等性が担保される点、配信に投稿されたコメントより観客アバタの笑いタイミングを機械的に決定できる点が理由にある。配信において笑いを意味する「w」「笑」「草」を含むコメントが投稿されたタイミングで、30体のうちランダムな観客アバタを笑わせるよう実装された。動画の長さは2分30秒に統一され、観客アバタの総笑い回数は、最も笑いコメントの少なかった動画に合わせ57回に統制された。

4.3 手続き

実験は18名(男性13名, 女性5名, 平均32.2 ± 9.1歳)に対して行われた。実験意図を知ることによる結果の偏りを防ぐため、実験参加者はコンテンツ視聴時の振る舞いの解析が実験目的であると伝えられた。実験参加者は、動画視聴とアンケートへの回答からなる試行を計6回行った。キャリアオーバー効果を防ぐため、試行間に1分間の四則演算課題に取り組んだ。全ての試行が終了した後、実験全体に関するインタビューに回答し、謝礼としてAmazonギフト券1,500円を受け取った。実験の所要時間は60分程度だった。

4.4 評価手法

観客アバタの反応の種類が共在感覚、笑いの伝染による楽しさ、そしてエンゲージメントに与える影響を評価するために、ソーシャルプレゼンス、エンジョイメント、エンゲージメントをそれぞれ質問紙によって計測した。ソーシャルプレゼンスは、先行研究[3, 4]で用いられた15項目の質問を用い7段階リッカート尺度で評価された。エンゲージメントおよびエンジョイメントは、先行研究[2]で用いられた8項目の質問を用い7段階リッカート尺度で評価された。

5. 結果

5.1 ソーシャルプレゼンス

正規性と等分散性の仮定が棄却されたため、整列ランク変換を用いた二元配置分散分析を行った結果、視覚的手がかりの主効果($p < .001$, $\eta_p^2 = 0.514$)と交互作用($p < .001$, $\eta_p^2 = 0.485$)が有意であった一方、聴覚的手がかりの主効果は有意ではなかった($p = .430$)。交互作用が有意であったため、視覚的手がかりについてフリードマン検定による多重比較を行った結果、全ての水準間でスコアに有意な差があった(NV vs. HV, LV vs. HV: $p < .001$; NV vs. LV: $p < .05$)。聴覚的手がかりについてウィルコクソンの符号順位検定を行った結果、全ての視覚的手がかり条件で単純主効果が有意であった(NV, LV: $p < .001$; HV: $p < .05$)。ただしLV条件では聴覚的手がかりの効果が逆転した。結果を図3の(A)に示す。

5.2 エンジョイメントとエンゲージメント

エンジョイメントのスコアについて、正規性と等分散性の仮定が棄却されたため、整列ランク変換を用いた二元配置分散分析を行った結果、視覚的手がかりの主効果($p = .757$)、聴覚的手がかりの主効果($p = .798$)、および交互作用($p = .580$)は有意でなかった。結果を図3の(B)に示す。

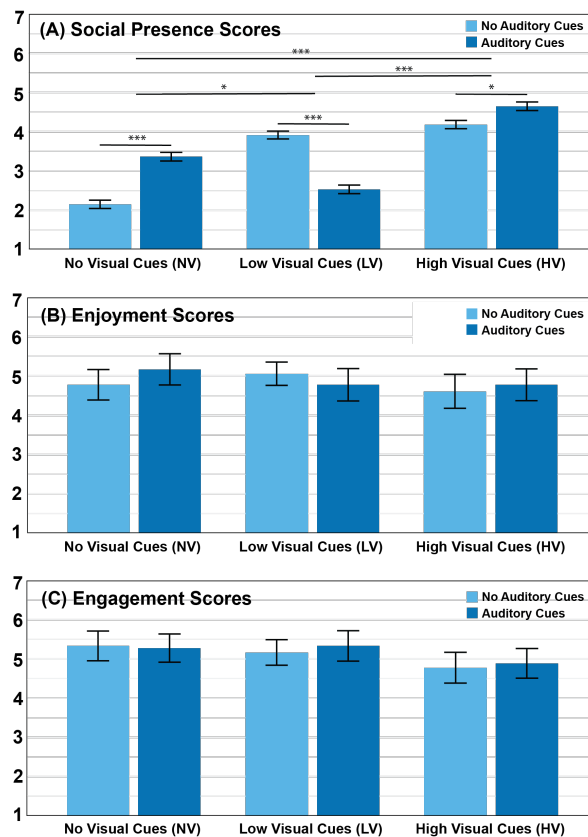


図3: (A) ソーシャルプレゼンス, (B) エンジョイメント, (C) エンゲージメントの結果 (* $p < .05$, *** $p < .001$)

エンゲージメントのスコアについて、正規性と等分散性の仮定が棄却されたため、整列ランク変換を用いた二元配置分散分析を行った結果、視覚的手がかりの主効果($p = .112$)、聴覚的手がかりの主効果($p = .733$)、および交互作用($p = .671$)は有意でなかった。結果を図3の(C)に示す。

6. 議論

6.1 ソーシャルプレゼンス

観客アバタの視覚的手がかりの増加は、概してソーシャルプレゼンスの向上に寄与することが示された。実験参加者から得られた「アバタがリアクションを取っている時は、皆で楽しさを分かち合っている感覚に包まれた(ID=7)」「笑い声と動作の両方あったほうが劇場で観ている感じがあった(ID=9)」といったコメントはこの結果を支持する。

一方、LV条件では聴覚的手がかりの付与によりソーシャルプレゼンスが低下したが、これはモダル間の情報の不一致に起因すると考えられる。実際に、実験参加者からは「心霊現象のように怖いところがあった(ID=1)」「動きのない人々が実際にそこに居るとは思えなかった。誰も居ないより不気味だった(ID=18)」といったコメントが得られた。

以上より、観客アバタのソーシャルプレゼンス向上には、視覚的手がかりの付与とその一貫性が重要であることが示唆された。実際にバーチャル笑い場を構成する際、参加者の表情や全身動作、笑い声の共有は技術面やプライバシー面で困難な場合もある。一方で本結果より、参加者の笑い

6 ジャルジャルタワー <https://www.youtube.com/@jarujarutower365/>

を検知した上で、事前定義された反応を再生する場合でも一定のソーシャルプレゼンス向上の効果が期待できる。

6.2 エンジョイメントとエンゲージメント

エンジョイメントとエンゲージメントに関しては、条件間で有意な差が見られなかった。実験参加者のコメントより、観客アバタの笑いタイミングが影響していた可能性が示唆されたため、本稿ではこの観点で議論する。

実験参加者からは「自分が面白いと思ったタイミングで笑い声が上がると気持ち良いが、逆だとなぜ笑っているかと混乱する (ID=16)」「自分の好みと違うところで笑っている不自然さが印象に残っている (ID=10)」といったコメントが得られた。今回、観客の笑いタイミングは YouTube Live に投稿された不特定多数のコメントより生成されたが、テキストベースのコメントを笑い変換することに一定の限界があること、不特定多数のコメントを反映したことにより、個々の実験参加者の笑いタイミングとの不一致が生じた可能性が示唆された。適切なタイミングでの笑い声はエンジョイメントとエンゲージメントを高める一方で、不適切なタイミングでの笑い声はそれらを低下させる可能性があり、これらのポジティブな効果とネガティブな効果が相殺し、有意な差として表れなかったものと考えられる。

実験環境とは異なり、実際のライブイベントでは共通の動機を持った参加者が集うため、ある程度この問題は回避されると考えられるが、バーチャル笑い場特有の対応も想定される。例えば、参加者の笑い嗜好に基づいた座席配置の最適化や、参加者の笑い反応に基づいた周囲の参加者の笑い反応の強度やタイミング調整が有益である可能性がある。

7. 結論

本研究は、バーチャル笑い場において、観客アバタの反応の視聴覚の手がかりが共在感覚、エンジョイメント、エンゲージメントに与える影響について調査を行った。実験の結果、観客アバタの反応の視聴覚の手がかりの付与とその一貫性が共在感覚の向上に寄与することが示唆された一方、エンジョイメントとエンゲージメントは条件間で有意な差が見られなかった。実験参加者のコメントより、参加者の笑いとお客アバタの笑いのタイミングが一致することで観客アバタが効果的に作用する可能性が示唆された。

なお、本研究では実験室環境における基礎調査を行ったが、実際のバーチャルライブイベントを通じた実践的な知見の蓄積も重要だと考えている。例えば、本システムをソーシャル VR プラットフォーム上に実装することで、プロパフォーマーとのコラボレーションや日常利用時の効果検証も可能となるだろう。引き続き、効果的な観客アバタの設計論についての議論を深めるとともに、バーチャル環境ならではの笑い場の構成法についても検討を進めていきたい。

謝辞 本研究は、JST 共創の場形成支援プログラム (COINNEXT) の支援を受けて行われた。

参考文献

- [1] A. M. Webb et al., Distributed liveness: Understanding how new technologies transform performance experiences, Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing, 2016, pp.432-437.
- [2] E. W. K. See-To, User Experience on Mobile Video Appreciation: How to Engross Users and to Enhance Their Enjoyment in Watching Mobile Video Clips, Technological Forecasting and Social Change, 2012,79(8), pp.1484-1494.
- [3] J. Cameron, The On-Screen Water Cooler: Effects of Televised User-Generated Comments on Cognitive Processing, Social Presence, and Viewing Experience, Doctoral dissertation, University of Tennessee, 2016.
- [4] Y. Guang 他, バーチャルライブにおけるアバタの動作を通じた観客同士のインタラクションが臨場感に与える影響, 情報処理学会インタラクション 2023, 2023, 3P-58.
- [5] H. Yakura, M. Goto, Enhancing Participation Experience in VR Live Concerts by Improving Motions of Virtual Audience Avatars, 2020 IEEE International Symposium on Mixed and Augmented Reality, 2020, pp.555-565.
- [6] De W. Micaela et al., Modality-Specific Contagion of Laughter and Yawning in the Absence of Others, Frontiers in Psychology, 2022, 13, 780665.
- [7] 福嶋ら, 笑い増幅器: 笑い増幅効果の検証, ヒューマンインタフェース学会論文誌, 2010, 12 巻, 3 号, p.199-207.
- [8] K. Suzuki et al., FaceShare: Mirroring with Pseudo-Smile Enriches Video Chat Communications, Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017, pp.5313-5317.
- [9] S. Cosentino et al., Quantitative laughter detection, measurement, and classification—A critical survey, IEEE Reviews in Biomedical engineering, 2016, 9, pp.148-62.
- [10] P. Ekman, W. V. Friesen, Facial action coding system, Environmental Psychology & Nonverbal Behavior, 1978.
- [11] R. Niewiadomski, C. Pelachaud, Towards Multimodal Expression of Laughter, International Conference on Intelligent Virtual Agents, 2012, pp.231-244.
- [12] H. J. Griffin et al., Perception and Automatic Recognition of Laughter from Whole-Body Motion: Continuous and Categorical Perspectives, IEEE Transactions on Affective Computing, 2015, 6(2), pp.165-178.
- [13] D. Xin et al., Laughter Synthesis Using Pseudo Phonetic Tokens with a Large-Scale in-the-Wild Laughter Corpus, Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2023, pp.17-21.