

AIoT 行動変容学会 第9回研究会論文集

(Proceedings of the Academy of Behavior Transformation by AIoT 9th Research Meeting)

2025 年 3 月 10 日 (月) ~ 11 (火)

はこだて未来大学

目次

- 開催概要 … p. 1
- プログラム … p. 2～3
- 一般講演一覧 … p. 4
- ポスター・デモ一覧 … p. 5
- 一般講演予稿集 … p. 6～164
- ポスター・デモ・企業発表予稿集 … p. 165～181

AIOT行動変容学会

第9回研究会 (BTI-9)

Academy of Behavior Transformation by AIoT. The 9th Research Meeting (BTI-9)

- 日程： 2025年3月10日（月）14:00～18:00
3月11日（火）9:30～15:30
- 会場： はこだて未来大学 大会議室
〒041-8655
北海道函館市亀田中野町 116番地2)
- URL： <http://www.sig-bti.jp/>
- X： [@bti_academy](https://twitter.com/bti_academy)
- アクセス：地図出典： <https://www.fun.ac.jp/contact>



プログラム

1日目： 3/10（月）

- 13:30- 受付
- 14:00-14:10 オープニング
- 14:10-15:58 セッション1（108min, 18min×6件）

G01	運動速度とスマートウォッチを活用した筋力トレーニングの支援	石田直希 青山学院大学
G02	宅内行動アノテーションの省力化のための行動認識手法の提案と評価	浅井俊宏 奈良先端科学技術大学院大学
G03	メタバース空間におけるファントムセンスの実証研究	齊藤天馬 公立ほこだて未来大学
G04	非特定テレビ視聴履歴データに基づくテレビ視聴行動クラスタリングと視聴傾向変化分析	由田翔吾 奈良先端科学技術大学院大学
G05	訓練中の動作・所産からの漢字書字正答率の予測	大森幹真 早稲田大学
G06	ゲーミフィケーションを活用した珠算学習支援システム「CalcQuest」のユーザスタディ	小嵯泰造 奈良先端科学技術大学院大学

- 15:58-16:10 休憩
- 16:10-17:58 セッション2（108min, 18min×6件）

G07	生成AIを活用した能動的学習環境の構築とハルシネーション問題の解決	小北駿 公立ほこだて未来大学
G08	生成AIを用いたいけばな創作支援についての検討	横窪安奈 東京大学大学院
G09	RAG コンパニオンチャットボットの開発と検証 -自己開示とユーザー体験への影響-	中原慧 名古屋大学大学院
G10	エコ行動促進アプリにおけるPUSH配信を用いたフィードバック手法	塚本航也 NTTドコモ / NTT DOCOMO Ltd.
G11	経験知を取り込んだ機械学習とデジタルツールに対する心理的障壁の低減	村瀬颯登 株式会社UACJ
G12	未登録PoI探索のための時系列SNS投稿データに基づく位置推定手法の提案と評価	澤野耕平 奈良先端科学技術大学院大学

- 17:58-18:00 撤収・移動
- 18:15- 懇親会（キャンパス内）

2日目： 3/11 (火)

- 9:00- 受付
- 9:30-10:45 デモ・ポスターセッション (75min)

P01	メッセージ介入実験の個人特性に関する分析： クラスター分析による介入効果の検証	木村淳哉 沖電気工業株式会社
P02	一人称映像を用いた選書中の行動パターンの分析	畑井梨里衣 公立ほこだて未来大学大学院
P04	自閉傾向とマガーク効果における視聴覚刺激への選好	北田結 早稲田大学
P03	視覚探索課題からみたASD児の局所的注意の優位性	飯田亜衣 早稲田大学
P05	協調運動得点とASD傾向における読み書きの視線の違い	渡邊貴子 早稲田大学
P06	現実と仮想をつなぐジェネラティブアート体験： 複合現実を活用した学び	古市駿 公立ほこだて未来大学大学院
P07	テキスト対話における受信者感情予測の個別化	菅野光 電気通信大学
P08	応用行動分析学を活用して組織の行動変容につなげる	是村由佳 (株) コレムラ技研

- 10:45-11:00 休憩
- 11:00-12:30 セッション3 (90min, 18min×5件)

G13	チェロの遠隔練習支援システムの試作と検討	太田健 公立ほこだて未来大学
G14	仮想現実での視線情報を用いた商品作動ギミックによる関心の変化	松田滉生 青山学院大学
G15	メタバースにおける仮想アバターとロールプレイングがコミュニケーションに与える影響 - プロテウス効果の視点から	仲山芳古 公立ほこだて未来大学大学院
G16	宅内デジタルツインを活用した環境センサによる生活行動認識手法の検討	菊池尊勝 奈良先端科学技術大学院大学
G17	ペット型人工物の感情表現と動作を用いた習慣改善のための説得技術	原田理央 公立ほこだて未来大学大学院

- 12:30-13:30 昼食
- 13:30-15:00 セッション4 (90min, 18min×5件)

G18	Thing2Vec: 人の行動理解に向けたモノの使用特性のベクトル埋め込み方法の提案	田仲百音 大阪大学
G19	カードゲームの心理戦における表情の頻度分析	長澤颯音 公立ほこだて未来大学
G20	Immersion Neckwear: 動画体験拡張ネックウェア	岩本空 青山学院大学
G21	プレイヤーの感情に応じて適応するゲーム環境の試作と評価	嵯峨京介 公立ほこだて未来大学
G22	動画視聴における瞬目促進フィードバック手法の検証	新村温人 青山学院大学

- 15:00-15:10 休憩
- 15:10-15:30 クロージング・授賞式

一般講演発表一覧

G01	運動速度とスマートウォッチを活用した筋力トレーニングの支援 石田直希, ロペズギヨーム(青山学院大学)
G02	宅内行動アノテーションの省力化のための行動認識手法の提案と評価 浅井俊宏, 松井智一, 諏訪博彦, 安本慶一(奈良先端科学技術大学院大学)
G03	メタバース空間におけるファントムセンスの実証研究 斉藤天馬, 角薫(公立はこだて未来大学)
G04	非特定テレビ視聴履歴データに基づくテレビ視聴行動クラスタリングと視聴傾向変化分析 由田 翔吾(奈良先端科学技術大学院大学), 松田 裕貴(岡山大学), 松田 裕貴, 横田 哲弥, 榊原 太一(読賣テレビ放送株式会社), 安本慶一(奈良先端科学技術大学院大学)
G05	訓練中の動作・所産からの漢字書字正答率の予測 大森幹真(早稲田大学人間科学学術院), 切替このみ(早稲田大学大学院人間科学研究科)
G06	ゲーミフィケーションを活用した珠算学習支援システム「CalcQuest」のユーザスタディ 小嵯泰造(奈良先端科学技術大学院大学), 松田裕貴(岡山大学)
G07	生成AIを活用した能動的学習環境の構築とハルシネーション問題の解決 小北駿, 角薫(公立はこだて未来大学)
G08	生成AIを用いたいけばな創作支援についての検討 横窪安奈(東京大学大学院), 石田精一郎, 竹谷一真(VRC 華道部), 越塚登(東京大学大学院)
G09	RAG コンパニオンチャットボットの開発と検証 -自己開示とユーザー体験への影響- 中原慧(名古屋大学大学院)
G10	エコ行動促進アプリにおけるPUSH配信を用いたフィードバック手法 塚本航也, 鈴木明作, 吉川裕木子, 勝間田優樹, 山田拓也, 石川太郎(株式会社 NTT ドコモ)
G11	経験知を取り込んだ機械学習とデジタルツールに対する心理的障壁の低減 村瀬颯登, 山本佑樹(株式会社 UACJ)
G12	未登録PoI探索のための時系列SNS投稿データに基づく位置推定手法の提案と評価 澤野耕平(奈良先端科学技術大学院大学), 松田裕貴(岡山大学), 中谷響, 大内啓樹, 諏訪博彦, 安本慶一(奈良先端科学技術大学院大学)
G13	チェロの遠隔練習支援システムの試作と検討 太田健, 角薫(公立はこだて未来大学)
G14	仮想現実での視線情報を用いた商品作動ギミックによる関心の変化 松田滉生, ロペズギヨーム(青山学院大学)
G15	メタバースにおける仮想アバターとロールプレイングがコミュニケーションに与える影響 -プロテウス効果の視点から 仲山芳古, 角薫(公立はこだて未来大学大学院)
G16	宅内デジタルツインを活用した環境センサによる生活行動認識手法の検討 菊池尊勝, 松井智一, 諏訪博彦, 安本慶一(奈良先端科学技術大学院大学)
G17	ペット型人工物の感情表現と動作を用いた習慣改善のための説得技術 原田理央, 角薫(公立はこだて未来大学大学院)
G18	Thing2Vec: 人の行動理解に向けたモノの使用特性のベクトル埋め込み方法の提案 田仲百音(大阪大学), 岸野泰恵(NTT コミュニケーション科学基礎研究所, 大阪大学), 下西英之(大阪大学)
G19	カードゲームの心理戦における表情の頻度分析 長澤颯音, 角薫(公立はこだて未来大学)
G20	Immersion Neckwear: 動画体験拡張ネックウェア 岩本空, ロペズギヨーム(青山学院大学)
G21	プレイヤーの感情に応じて適応するゲーム環境の試作と評価 嵯峨京介, 角薫(公立はこだて未来大学)
G22	動画視聴における瞬目促進フィードバック手法の検証 新村温人, ロペズギヨーム(青山学院大学)

ポスター・デモ・企業発表一覧

P01	メッセージ介入実験の個人特性に関する分析： クラスタ分析による介入効果の検証 木村淳哉, 赤津裕子, 片桐一浩 (沖電気工業株式会社)
P02	一人称映像を用いた選書中の行動パターンの分析 畑井 梨里衣 (公立はこだて未来大学大学院), 角康之 (公立はこだて未来大学)
P03	自閉傾向とマガーク効果における視聴覚刺激への選好 北田結 (早稲田大学人間科学部), 大森幹真 (早稲田大学人間科学学術院)
P04	視覚探索課題からみたASD児の局所的注意の優位性 飯田亜衣, 大塚千冬 (早稲田大学人間科学部), 大森幹真 (早稲田大学人間科学学術院)
P05	協調運動得点とASD傾向における読み書きの視線の違い 渡邊貴子 (早稲田大学人間科学部), 大森幹真 (早稲田大学人間科学学術院)
P06	現実と仮想をつなぐジェネラティブアート体験：複合現実を活用した学び 古市駿, 角薫 (公立はこだて未来大学大学院)
P07	テキスト対話における受信者感情予測の個別化 菅野光, 橋山智訓 (電気通信大学)
P08	応用行動分析学を活用して組織の行動変容につなげる 是村由佳 (コレムラ技研)

運動速度とスマートウォッチを活用した筋力トレーニングの支援

石田直希^{†1}

青山学院大学^{†1}

ロペズギヨーム^{†2}

青山学院大学^{†2}

1. はじめに

近年、日本のフィットネス市場は急速に回復し、特にフィットネスクラブの利用者数が増加している[1]. 定期的な身体活動 (PA) は非感染性疾患の予防や治療に重要であり、特に筋力トレーニングは健康維持に不可欠だ[2,3,4]. しかし PA プログラムは高い離脱率が報告されており、フィットネスクラブの1年以上通う割合は4パーセント以下とされ、継続率の低さが問題である[5,6,7].モチベーションを維持するためには、運動中のパフォーマンスを最大化し、効果を実感できることが重要だ[8].

1.1 ベンチプレスの効果

ベンチプレスは広範囲の筋肉を効果的に鍛えるコンパウンド種目であり、上半身を鍛える最良の方法とされている[9].しかし、従来の10回3セットなどのプログラムでは個人に応じた効果を最大化できず、怪我のリスクも増加する[10].オーバーワークは極度な疲労やけがのリスクをもたらすため、モチベーションの低下につながる。よって、高強度低疲労の運動プログラムが重要だ。

1.2 Velocity Based Training の手法と効果

VBT (速度ベースのトレーニング) は、運動速度を指標にしたトレーニング法で、身体能力の個人差を考慮し、パフォーマンスを最大化する[8].これにより、個々の能力に応じたトレーニングプランを提供し、より効果的なトレーニングが可能になる[10].直感に頼った重量を指標としたトレーニングと比較し強度は高く、低疲労であるため、継続面において適切な運動プログラムの構築手法だ。

1.3 運動フィードバックに関する研究

Chariar らの研究では、画像分析人工知能技術を用いてスクワット動作を深層学習で分析し、改善のためのフィードバックをリアルタイムで提供した[2].関節位置を解析して2Dモデルを作成し、動作フィードバックを行うことで関節角度や姿勢、可動域を視覚的に確認し、姿勢の改善が見られた。しかし、このシステムを構築するには3台のカメラやその他の機材の設置が必要であり、トレーニング前にキャリブレーションも必要だ。ジム環境では、設置場所やコストの面で導入が難しく、実用性が低いと考えられる。

また、Michaud らは背中、腰、太ももに慣性センサ (IMU) を設置し、骨盤と体幹の傾きを測定することでデッドリフトのフォームを分析した[3].しかし、複数のIMUと3Dプリンタで制作した設置用器具が必要であり、実用性は依然として低いといえる。

本研究では実用性に重点を置き、単一のスマートウォッチのみを使用したシステムを制作する。

1.4 研究目的

本研究の目的は、VBTを活用して筋力トレーニングを支援することである。目標は以下の通りである。

1. VBTを基盤にした筋力トレーニング効率支援システムの開発
2. 開発したシステムの使用感と効果のフィードバック評価

2. ベンチプレストレーニング支援システムの概要

本システムはウェアラブルデバイスを利用し、運動速度を分析することでベンチプレスプログラムを構築する。データ取得にはスマートウォッチを用い、加速度を計測し重力成分を除去、ローパスフィルタでノイズを低減し、速度へ変換後ピークを検出し、挙上と速度を記録する。なお、図に示す通り挙上方向はz軸に相当するため、その軸の値のみを分析する。アプリはAndroid Studioで開発し、Google Pixel Watchにインストールする。Google Pixel Watchは加速度・ジャイロデータを取得可能で、毎秒20個のデータを記録し、ピーク検出を行う。

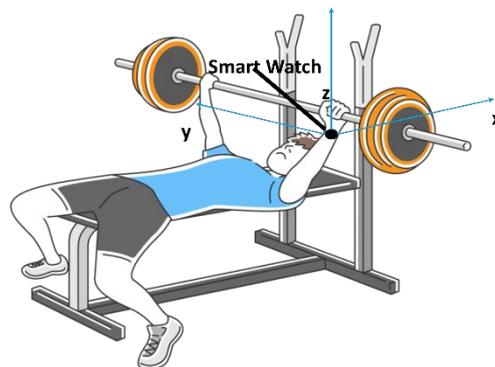


図1 ベンチプレスの挙上方向

Figure 1 Direction of Bench Press

Weakly らは、VBTをアプリケーションに導入する際に設定する運動速度の閾値を調査した[8].結果ほとんどの先行研究が検証を重ねて0.17m/sをベンチプレス最低運動速度の閾値として設定していることを示している。そのため、本実験も同様に0.17m/sを閾値として使用する。

ノイズに関しては、フィルタリング手法以外に以下の手法を使用して軽減している。

1. バーに左手をつけた状態でスタートボタンを押す
2. 指定回数のみピークを検出する

1に関してはセット前にノイズが入り、余計なピークが検出されることを防ぐことを目的とする。2は指定回数のみピークを検出することで、セット終了後のノイズを無視する。これにより、セット前後の動作により生じるノイズを無視することができ、正しく挙上回数を検出することができる。

2.1 重量の提案手法

重量が増加するにつれて運動速度が減少するという線形関係が見られ、この Load-Velocity プロファイル (以下 L-V プロファイル) を制作することで、個人の身体能力に応じた適切な運動プログラムが構築できる[8]。ただし、このプロファイルを制作するには、アスリートがレップを最大速度で実施し、1RM に達するまで様々なセットを繰り返す必要があり、時間と疲労の面で問題がある[11]。McBurnieらの研究では、2点法 (極端に異なる2つの負荷を用いる方法) と速度の基準データ (1RM 時の最小速度閾値) が、アスリートの 1RM を評価する実用的な方法となり得ることが示された[12]。この手法は上半身種目に信頼性があるが、下半身種目には適さない。本研究ではベンチプレスを対象とするため、この2点法を用いて個人に応じた L-V プロファイルを制作する。運動速度を測定するため、ウォームアップ時に 1RM45% と 85% の重量で2回ずつ挙上を行い、0.17m/s に該当する重量を提案する。

2.2 トレーニングメニュー

研究目標の一つとしてレジスタンストレーニングにおけるモチベーションの向上がある。そのためにも継続的に使用できるわかりやすい仕組みが重要であり、本システムでは5回5セットのメニューを取り入れた。これは Weaklyらが提唱した VBT を導入する際の運動プログラムの構築例の一つであり、図2にそのフローを示した[8]。

まず目標の重量をバーにセットし、セットを開始する。この時、挙上回数は5回とし、1セットを終了する。これを5回繰り返す、あるいは挙上に失敗した場合をセットの終了条件とする。本実験ではこれに運動速度を使用するため、閾値を下回った場合にもセットを終了する。

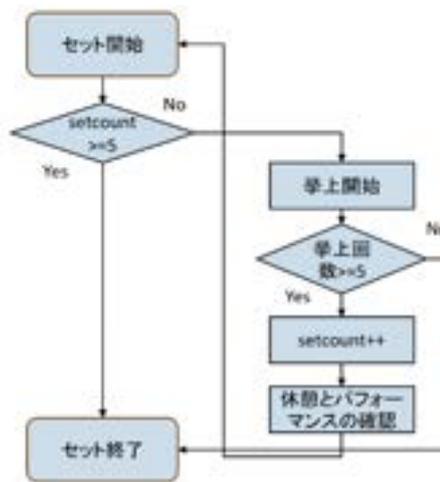


図2 メインセットフロー図

Figure 1 Flow Chart of Main Set

2.3 最適重量算出式の調整

ここで、一度制作した仮のシステムをもとに重量算出式の調整を実施した。被験者は7名で行い、2日程にてそれぞれシステム有無でデータを収集した。表1はその結果を出力したものである。これはベンチプレスにおける RM 換算表を使用し、5回5セットを前提とした重量に対してデータを収集した。その結果、算出した重量が本来の想定よりも約 5kg 軽く、想定した目標強度を下回っている。平均誤差は 5.71、標準偏差は 1.89 となっており、その差は重量が大きくなるにつれ小さくなる。

表1 システム有無における 1RM (kg) 挙上重量

Table 1 1RM (kg) raised weight with and without system

被験者	無し	有り	重量差
A	32.5	25	7.5
B	42.5	35	7.5
C	52.5	45	7.5
D	42.5	37.5	5
E	50	45	5
F	50	45	5
G	75	72.5	2.5

理由としては、図3に示す通りベンチプレスは垂直運動ではないため、z軸方向の垂直ベクトルの運動速度しか考慮せず、x軸方向の運動速度を計算に入れていないためであると考えられる。

先行研究では線形運動を正確に計測できるリニアトランジューサーを使用し、閾値を 0.17m/s としていたが、加速度測定において x 軸上のプレが速度の算出に影響を与えている。



図3 ベンチプレスの挙上方向

Figure 3 Direction of Bench Press Lift

そこで、図4ではこのデータをもとに直線回帰を行った。式 $y=0.8948x+10.299$ は、システムの提案する重量が、RM換算表をもとに推定した目標強度に対してどれくらい下回っているかを計算し、最適重量に調整するものである。

決定係数 0.99 と高い信頼性を示しており、この回帰式をシステムの重量提案に組み込む。

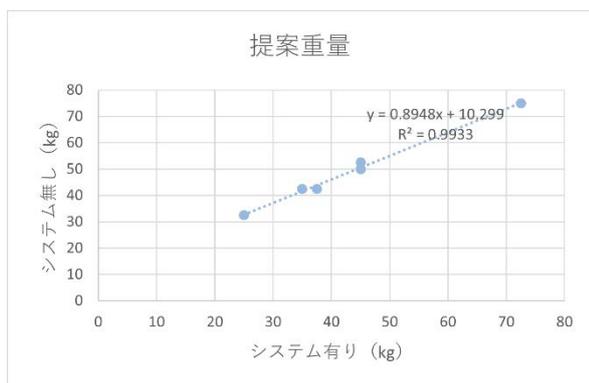


図4 提案重量のグラフ

Figure 4 Graph of Suggested Weight

3. 開発システムの効果検証

3.1 実験の目的

この実験の目的は、システムがパフォーマンスとモチベーションにどのような影響を与えるかを計測し、システム無しの結果と比較することである。まとめとして、以下の2つがあげられる。

1. システムありとなしでのパフォーマンスを比較
2. システムの使いやすさを検討

3.2 実験方法

本実験では健康な20代成人男性10名(7名は習慣的なベンチプレスの経験があり、3名は未経験)に対し、実験を2日

に分けて行う。なお、ベンチプレスの経験が浅い被験者がいることも考慮し、彼らは追加で1日程練習日を設ける。そして、1週間以上リカバリーの期間を設けたのち、システム有無での実験をそれぞれ行う。なお、被験者の健康状態を実験前に把握し、コンディションを考慮しつつ行う。

- 2つの実験
 - システムなし
 - システムあり
- 2つのアンケート (Google Formを使用)
 - SUSアンケート
 - フィードバックアンケート

3.3 実験手順

実験手順は以下の通りを行う

- システムなし
 1. 実験の内容について説明
 2. 筋力トレーニングにかかわる経験について質問
 3. 被験者の自由にトレーニングを実施
 4. セット数、回数、重量を記録
- システムあり
 1. 実験の内容について説明
 2. 筋力トレーニングにかかわる経験について質問
 3. スマートウォッチを左手に装着
 4. 1RM45%で運動速度を計測
 5. 1RM85%で運動速度を計測
 6. メインセットを実施
 7. セット数、回数、重量を記録

最後に、システムありとなしでの主観的評価をアンケートで収集した。

3.4 SUSアンケート

SUSアンケート (System Usability Scale の略) は、1986年にジョン・ブルックによって開発され、システムの知覚された使いやすさを測定するために使用されているアンケートである[13]。この手法は現在でも幅広く使用されており、主観的な利用性を評価するのに今後も主流であり続けるといわれている[14]。このスケールは、1 (全くそう思わない) から5 (強くそう思う) までの10の声明で構成されている。奇数番号の質問は肯定的な質問であり、偶数番号の質問は否定的な質問となっている。集計方法は、奇数番号の質問に対して回答数から1を引き、偶数番号の質問に対しては、回答数を5から引く。その後、総得点に2.5を掛けて、0から100のスケールに変換する。各スコアの評価指数は表4.2に示されている。

表2 SUS 評価スコア

Table 2 SUS Validation Grade

SUS Score	Grade	Adjective Rating
>80.3	A	Excellent
68-80.3	B	Good
68	C	Okay
51-68	D	Poor
<51	E	Awful

3.5 フィードバックアンケート

システムのフィードバックに対して評価をするため、アンケートを作成した。参加者には、6つの質問に対してどの程度同意するかを質問し、システムの使いやすさや効果を確認した。そして、最後にシステムに関する意見を自由記述で募集した。

3.6 開発システムの効果検証

ベンチプレストレーニング支援システム有無でのパフォーマンス比較とユーザビリティ検証 (SUS) を行った。被験者は健康な 20 代成人男性 10 名で、うち 7 名は習慣的なベンチプレスの経験があり、3 名は未経験だった。経験者はシステム有無で 2 日程、未経験者は練習日を含めた 3 日程で実験を行った。運動強度の 1RM 比較および、SUS および、アンケートの回答を分析した。

4. 実験結果と考察

4.1 システムによる強度比較の検証

10 人の成人男性 (経験者 7 名未経験者 3 名) よりデータを収集した。未経験の参加者もいたため、未経験者は練習を含めた 3 日程に分けて練習し、他はシステム無し、システム有りに分けて実施した。表 3 は 1RM における強度を比較した結果である。システムを活用することで、平均して 4kg 強度が増えた。また t 検定の結果として、 $p < 0.05$ であることから結果は有意であるといえる。しかし、被験者 2 と被験者 4 を見ると運動強度がシステムを使用した場合に下がっていることがわかる。原因としては、2 点法で 1RM を推定するうえで、直近のパフォーマンスを把握できないことが挙げられた。ブランク期間があったため 1RM 正しく把握できず、けがのリスクを避けるため本来より軽い重量に設定してしまい、提案重量も軽くなってしまった。ベンチプレスは前回トレーニングからのブランク期間に影響を受けるため、日別におけるパフォーマンスを考慮する必要がある[15]。

表3 システムありとなしにおける強度比較

Figure 3 Comparison in Strength Between with and without the system

Subject ID	1RM システム有り (kg)	1RM システム無し (kg)	差
1	45	39	6
2	39	42	-3
3	28	25	3
4	45	48	-3
5	63	59	4
6	62	56	6
7	68	56	12
8	90	84	6
9	48	45	3
10	39	34	5

4.2 ユーザビリティの検証結果

表 4 が SUS アンケートの結果である。Excellent 評価が 7 人、Good 評価が 3 人と総じて良い結果であることがわかる。スコア平均は 87.75 点となり、ユーザにとっては非常に使いやすいシステムであるということが分かった。

特に「システムはシンプルで使いやすい」の項目では非常に高い評価を頂き、信頼を得やすく、多くの人に受け入れてもらえるシステムであると考えられる。

一方で、「イレギュラーなことも多い」及び「このシステムを使うには技術的なサポートが必要」の項目の評価が低く、特に平均評価より低い被験者 1、被験者 5 に共通していた。この項目の評価が低かった原因を調べたところ、手首を曲げたときに無意識的にスマートウォッチのサイドボタンを押してしまい、アプリが使用できなくなったことが原因であった。使い慣れないデバイスを使用する場合、イレギュラーに対応できないため、トラブルに対処できるよう修正する必要がある。

表4 SUS の評価結果

Table 4 Results of SUS Validation

Subject ID	Score	Grade
1	77.5	B
2	92.5	A
3	82.5	A
4	92.5	A
5	77.5	B
6	75.0	B
7	95.0	A
8	97.5	A
9	92.5	A
10	95.0	A

Q1. フィードバック情報はわかりやすく、簡単に理解できたこの質問に対しては 70%の被験者が強く同意、30%の被験者が同意しており、肯定的な意見であった。このことから、フィードバック情報は簡潔で被験者にわかりやすく伝わっていたことがわかる。

Q2. 適切なフィードバックが与えられていた

この質問には 80%もの被験者が肯定的な意見を示した。運動速度のフィードバックがトレーニングのペース維持に有効であったことがわかる。

Q3. 運動速度ベースのメニューは重量ベースのものと比較して効果をより実感した

この質問には 70%もの被験者が肯定的な意見を述べたのに対し、20%は否定的な意見を示した。これは、一部の被験者は自身のやり慣れた手法の方が効果を実感したことがわかる。

Q4. メニューの構築がより簡単になったと感じた

この質問に対しては 80%もの被験者が肯定的な意見を示した。フレームワークを用意することは、運動プログラムの構築を簡素化に役立つことがわかった。

Q5. システムを活用することでモチベーションが向上すると実感する

質問に対しては被験者全員が肯定的な意見を示した。やはり、パフォーマンスの定量化とフィードバックはモチベーションの向上に効果的である。

Q6. デバイスを邪魔に感じた

この質問には 80%もの被験者が否定的な意見を示したが、20%のみが肯定的に感じた。

Q7. デバイスを装着することで、違和感やベンチプレスへの影響があった

この質問には、80%もの被験者が否定的な意見を示したが、10%のみが肯定的な意見を示した。これらのことから、一部の被験者は手首を曲げたときにデバイスが何らかの違和感を与えることが分かった。

また、アンケートの最後に自由記入欄を用意した。被験者の多くは肯定的な意見を述べた一方、システムの改善案についての意見も寄せられた。具体的には、「友人とスコアを比較したい(実数値的なものも含め)」という意見だ。本システムはスマートウォッチ内で計算処理を完結させたため、サーバにデータを送受信し、ユーザ横断的なプラットフォームを構築することができなかった。

他ユーザとの数値的な比較はモチベーションの向上につながるため、システムの改善を検討したい。他に「過去の推移から自分の成長具合を確認したい」という意見が寄せられた。他ユーザと自分の成長具合を推移的に把握することで、競争要素のあるプラットフォームを構築することができるかもしれない。

「細かな気づきを与えてくれるパーソナルトレーニング要素が欲しい」という意見も寄せられた。ベンチプレスはバーの持つ手幅やフォームによって鍛えられる部位が左右するため、フォームも同時にフィードバックできるとよい[16]。最後に「5回挙げきれなかったら、軽い重量で残りセットを再提案してほしい」という意見が寄せられた。これは5回5

セットメニューによる柔軟性の無さが、トレーニングの強度を下げている可能性があるとい指摘だ。筋力や筋持久力には個人差があるため、柔軟に運動プログラムを選択できるように、改善する必要がある[17]。

5. おわりに

この評価実験では、3つの問題が見つかった。1つ目はフィードバック方法に関するもので、今回はセットごとに独立したフィードバックを行った。しかし、ユーザからはより長期的に比較できるフィードバックが求められた。具体的には、過去のデータの推移や他ユーザとのスコア比較である。したがって、クラウド上でデータを保管し、分析できるシステムに改良することを検討したい。

2つ目の問題は、運動プログラムの柔軟性の無さである。5回5セットに固定する運動プログラムの構築方法はユーザの手間を省く利点がある。しかし、被験者の意見から、筋持久力や筋力の違いで一部ユーザには適さない可能性があることが発覚した。運動プログラムの構築には重量やセット回数を柔軟に設定できる方法もあるため、ユーザによって選択できるように改良することを検討したい。

3つ目の問題は、デバイスがトレーニングの邪魔になることだ。今回実験を行った被験者の半数は経験年数が3年以上の中級者以上であり、うち3人は自重以上を上げる上級者であった。彼らは手首にリストラップを装着するため、スマートウォッチが邪魔になるという意見があった。したがって、このシステムを中級者以下に限定する、もしくはスマートウォッチ以外のウェアラブルデバイスを使用することを検討する必要がある。これら3つの問題点を踏まえ、今後の研究につなげたい。

謝辞

本研究はJSPS科研費 JP22K11998の助成を受けたものです。

参考文献

- [1] 株式会社帝国データバンク情報統括本部：「フィットネスクラブ・スポーツジム」業界動向調査（2023年度），TDB Business View（2023）。
- [2] Committee, P. A. G. A.: 2018 Physical Activity Guidelines Advisory Committee Scientific Report (2018).
- [3] 山田綾澤田亨門間陽樹：“筋トレ”の疫学：Muscle-strengthening exercise に関するナラティブレビュー，運動疫学研究，Vol. 23, No. 2, pp. 129–142 (2021).
- [4] C. Harris, K. Watson, S. C. J. F. J. D. and Elam-Evans, L.: Adult participation in aerobic and muscle-strengthening physical activities , MMWR. Morbidity and Mortality Weekly Report, Vol. 62, No. 17, p. 326 (2013).
- [5] Schutzer, K. A. and Graves, B. S.: Barriers and motivations to exercise in older adults, Preventive Medicine, Vol. 39, No. 5, pp. 1056–1061 (2004)

- [6] J. Kruger, H. M. B. and Gillespie, C.: Dietary and physical activity behaviors among adults successful at weight loss maintenance, *International Journal of Behavioral Nutrition and Physical Activity*, Vol. 3, No. 1, p. 17(2006).
- [7] S. Sperandei, M. C. V. et al.: Adherence to physical activity in an unsupervised setting: Explanatory variables for high attrition rates among fitness center members, *Journal of Science and Medicine in Sport*, Vol. 19, No. 11, pp. 916–920 (2016).
- [8] J. Weakley, B. Mann, H. B. S. M. T. S. and Garcia Ramos, A.: Velocity-Based Training: From Theory to Application, *Strength and Conditioning Journal*, Vol. 43, No. 2, pp. 31–49 (2021).
- [9] Rippetoe, M.: スターティングストレングス JP Oversized, 医学映像教育センター, 東京, 第3版 edition (2019).
- [10] Jimenez-Reyes P., Castaño-Zambudio A., C.-P. V. et al.: Differences between adjusted vs. non-adjusted loads in velocity-based training: consequences for strength training control and programming, *International Journal of Sports Physiology and Performance*, No. 1, pp. 1–11(2021).
- [11] 著者名: 新版 図解 スポーツトレーニングの基礎理論, 西東社 (2019).
- [12] McBurnie, A. J. et al.: 負荷-速度関係を用いた最大挙上重量の予測の利点と限界, *Strength & Conditioning Journal: 日本ストレングス & コンディショニング協会機関誌*, Vol. 28, No. 1, pp. 51–60 (2021).
- [13] contributors, W.: System Usability Scale, https://en.wikipedia.org/wiki/System_usability_scale (2025). Accessed: 2025-01-13.
- [14] R, L. J.: The System Usability Scale: Past, Present, and Future, *International Journal of Human–Computer Interaction*, Vol. 34, No. 7, pp. 577–590 (2018).
- [15] Slope Media: 筋トレとブランクの関係性! 効率よく再開〜MAX まで戻すポイントを解説! (2025). Accessed: 2025-01-13.
- [16] L. Wagner, S. E. et al.: The Effect of Grip Width on Bench Press Performance, *International Journal of Sport Biomechanics*, Vol. 8, No. 1, pp. 1 – 10 (online), <https://journals.humankinetics.com/view/journals/jab/8/1/article1.xml> (1992).
- [17] Rissanen J., W. S. and B., P.: Velocity-based resistance training: do women need greater velocity loss to maximize adaptations?, *Eur J Appl Physiol*, No. 122, pp. 1269–1280 (2022).

宅内行動アノテーションの省力化のための行動認識手法の提案と評価

浅井俊宏^{†1}

奈良先端科学技術大学院大学

松井智一^{†2}

奈良先端科学技術大学院大学

諏訪博彦^{†3}

奈良先端科学技術大学院大学

安本慶一^{†4}

奈良先端科学技術大学院大学

1. はじめに

高齢化が進む現代社会において、家庭内における高齢者の生活状況を把握し、ヘルスケアや見守り支援を行うための技術が注目されている。具体的には、家庭内での移動、活動時間、睡眠パターンなどをモニタリングすることにより、健康状態の変化や異常を早期に検知し、迅速な対応や適切な介護計画の立案が可能となる。また、これにより、高齢者の生活の質 (Quality of Life: QOL) を向上させることが期待されている。このような行動認識を実現するためには、まず高精度なモデルを構築する必要があるが、その過程においていくつかの課題が存在する。

このような行動認識を実現するためには、まず高精度なモデルを構築する必要があるが、その過程においていくつかの課題が存在する。第一に、モデルの学習に必要な大量のラベル付きデータの収集が極めて困難である点が挙げられる。従来の研究では、被験者が日常的に行う行動に対して詳細なアノテーションを付与する必要があったが、この作業には多大な時間と労力がかかる。また、被験者自身がアノテーションを行う場合、その正確性にも課題が残る。第二に、家庭環境の多様性が、行動認識モデルの精度向上を阻む要因となっている。家庭ごとに生活様式が異なるため、汎用的なモデルでは各家庭の特有のパターンを適切に捉えることが難しい。一方、家庭ごとに個別のモデルを構築しようとすると、都度アノテーション作業が必要となり、実用化が遠く。そのため、アノテーションコストを最小限に抑えつつ、高精度なモデルを効率的に構築する手法が求められている。

さらに、行動認識技術は、非接触型であることやプライバシー保護の観点からも考慮する必要がある。近年、環境センサや人感センサを利用した非接触型のセンシング技術が注目を集めており、これらを活用することで、プライバシーに配慮したデータ収集が可能となる。このような技術の発

展は、ラベル無しデータを活用した新しい学習手法の適用可能性を広げるものである。

そこで本研究では、家庭内に設置した非接触型センサのセンシングデータを活用し、少量の行動ラベルでも効率的にモデルを学習できる手法を提案する。具体的には、無ラベルデータに対する自己教師あり学習 (Self-Supervised Learning) と、分類器の不確実性に応じた追加アノテーションを行う能動学習 (Active Learning) を組み合わせる。これにより、(1) 十分量のデータがあってもラベルが不足している問題への対処、(2) プライバシーに配慮しつつ行動認識の性能を向上、(3) 被験者のアノテーション負担軽減といった効果が期待できる。

本研究では、複数世帯 (4 家庭) を対象として、家庭内センサデータを収集し、そのうち 1 世帯分のデータに対して自己教師あり学習で無ラベルデータから分類モデルを実装した。さらに、能動学習として、分類モデルの確信度が低いサンプルのみを使用した学習を追加で行う事で、効率的に分類モデルの精度が向上されることを確認した。本論文では、関連研究を概観したうえで、提案手法の詳細と実験的評価を報告し、今後の課題を考察する。

2. 関連研究

2.1. 宅内行動認識

近年、行動認識 (Human Activity Recognition, HAR) 技術の発展により、スマートホーム、ヘルスケア、異常行動検知などの幅広い応用が期待され、高齢者の見守り等への応用が注目されている。[1, 2, 3, 4] これらの HAR を行うシステムは、ウェアラブルデバイスを利用した手法と利用しない手法の二つに大別することができる。

ウェアラブルデバイスを利用した行動認識は、加速度センサやジャイロスコープを搭載したデバイスを装着し、利用者の動きをセンシングすることで行動を認識する手法である。この手法は、デバイスの装着によって個人ごとの行動識別が可能であり、特に転倒検知や健康モニタリングの分野で多くの研究が行われている。Giulia らは、ウェアラ

In-Home Activity Recognition for Reducing Annotation Costs

^{†1} Toshihiro Asai, Nara Institute of Science and Technology

^{†2} Tomokazu Matsui, Nara Institute of Science and Technology

^{†3} Hirohiko Suwa, Nara Institute of Science and Technology

^{†4} Keiichi Yasumoto, Nara Institute of Science and Technology

ブルデバイスを活用し、状況認識を組み合わせた HAR システムを提案した [5]。しかし、装着の煩雑さや長期間の利用におけるデバイスの電池寿命、ユーザの装着拒否といった課題がある。

ウェアラブルデバイスを利用しない行動認識の手法は、近年、特にプライバシー保護の観点から注目されている。Xiaolong らは市販の Wi-Fi デバイスを活用した非接触型の人間の行動認識システムを開発し、LSTM ベースの時系列データ解析に、周波数情報を組み合わせた新しい認識モデルを提案した [6]。このような電波を活用した手法は、非装着型の利便性が高い一方で、センシング環境に依存しやすいという課題がある。例えば、家具の配置変更や機器の追加・削除が発生した場合、再キャリブレーションが必要になることが多い。

2.2. 人感・環境センサによる宅内行動認識システム

設置型の人感・環境センサを利用した行動認識は、非装着型の利便性が高く、ユーザの負担を軽減する手法として注目されている。ただし、Wi-Fi や RFID を使用する手法と異なり、住環境で収集されるデータは非常にプライベートなデータであることから、プライバシーやセキュリティに配慮する必要がある。

プライバシーに配慮しつつ、一般家庭における行動認識を目的としたセンシングシステムとして、CASAS のスマートホームキットがある [7]。CASAS は、居住者の動作や生活パターンを非接触かつ間接的に記録するプライバシーウェアなセンシングキットである。しかしながら、CASAS にはいくつかの課題も存在する。すべてのセンサはバッテリー駆動であり、定期的なメンテナンスが必要となるほか、バッテリー搭載によりセンサのサイズがやや大型化している。さらに、行動ラベルのアノテーションには、居住者自身が PC 上のソフトウェアを用いて行動ラベルを設定する必要があり、アノテーション作業の手間が発生する。

CASAS の課題を解決するため、同様のスマートホームキットとして SALON [8] が提案されている。SALON は、エネルギーハーベスティング (環境発電) 技術を活用したセンサを採用し、長期間のメンテナンス不要な動作を実現している。これにより、バッテリー交換の必要がない。さらに、SALON は居住者の負担を軽減するため、物理的なアノテーションボタンを導入し、ユーザが直感的に自身の行動を記録できるようになっている。このアノテーションボタンも環境発電技術を利用しており、電池交換なしで動作するため、CASAS と比較してさらなるアノテーション負荷の軽減およびメンテナンスフリーの利点を有する。しかしながら、依然としてアノテーション作業は負担のかかる作業であり、実際に SALON ではアノテーション忘れによるラベルの欠

損が課題として残っている。

2.3. アノテーションの効率化・省力化

SALON ではアノテーションボタンを導入することで一定の負担軽減を実現したが、それでもすべての行動に対して正確なラベルを付与することは依然として難しく、さらなる省力化の工夫が求められる。この課題を解決する手法の一つとして、不確実性サンプリング (Uncertainty Sampling) がある [9]。

不確実性サンプリングでは、モデルが予測に自信を持っていないデータ点を優先的に選択し、必要最小限のアノテーションでモデルの精度を向上させる。例えば、行動認識において、モーションセンサの発火パターンが曖昧な場合や、異なる行動の境界が不明瞭な場合に、そのデータのみを居住者に確認してもらい、ラベル付けを行うことで、不要なアノテーション作業を削減しつつ、効率的に学習データを強化することが可能となる。

この不確実性サンプリングを学習に応用した手法として能動学習 (Active Learning) が注目されている。能動学習は、機械学習モデルが高い確信度を持たないデータに対して、追加のアノテーションを優先的に要求する手法であり、アノテーション作業を効率化することができる [10]。

2.4. 本研究の位置づけ

従来の宅内行動認識システムでは、ウェアラブルデバイス、カメラ、マイクを活用する手法が広く研究されてきた。しかしながら、ウェアラブルデバイスは装着の煩雑さや電池寿命の制約といった課題があり、カメラやマイクを用いた手法はプライバシー侵害の懸念が大きいため、一般家庭への導入が難しい。これらの課題を解決するため、非接触かつプライバシーに配慮したセンサを活用した手法が近年注目されている。

しかしながら、依然としてアノテーション作業の負担が大きいという課題が残る。特に、長期間のデータ収集を行う実環境においては、居住者が自身の行動を逐一記録する負担が大きく、アノテーション忘れやラベルの欠損が生じやすい。

本研究では、これらの課題を踏まえ、SALON でも使用したセンサを活用したプライバシー侵害の少ない宅内行動認識システムを構築するとともに、アノテーション作業の負担を軽減する新たな手法を提案する。具体的にはラベル無しデータを活用した自己教師あり学習と能動学習を組み合わせることで、少量のアノテーションデータでも高精度な行動認識が可能なモデルを構築する。また、実環境における長期間のセンシングデータを用いた評価を行い、提案手法の有効性を検証する。これにより、実環境でも宅内行動

認識の普及を妨げる要因の一つであるアノテーションの負担を軽減し、より実用的なシステムの実現を目指す。

3. 提案システム

本システムのセンシングについては、ユーザの日常生活に干渉せずに行動データを収集することが求められる。具体的には、環境センサや人感センサを用い、家庭内に設置されたセンサ群から情報を取得する仕組みを採用する。これにより、身体に装着するデバイスを用いる必要がなく、ユーザが意識することなく行動記録を行うことが可能となる。また、データ収集において、ユーザのプライバシーを保護することは極めて重要である。特に、映像や音声など個人の特定が可能な情報の使用を避けるため、温度、湿度、移動感知等の個人特定ができないデータを用いる。

3.1. システム構成

提案システムの要件を満たすよう、使用したデバイスを図 1 に示す。

3.2. 人感センサ

人感センサは、赤外線の入射変化を検知することで、人間等の熱源の動きを感知する。メンテナンスフリーかつ長期間の運用を可能とするため、低消費電力通信規格である EnOcean に対応した ROHM 社の温度センサ STM431J と Panasonic 社の人感センサモジュール EKMB1101112 を組み合わせて使用した。このセンサは、環境発電のみで動作し、データ収集および送信が可能である。

加えて、環境発電が困難な状況においても安定したデータ取得を実現するため、予備の人感センサとして Switchbot 製を採用した。BLE 通信を利用し、単 4 電池 2 本で動作するこのセンサは、通常の使用条件下で最大 3 年間の稼働が可能である。

3.3. 環境センサ

環境センサとして、小型で多機能な Omron 社の 2JCIE-BL01 (BAG タイプ) を使用した。このセンサは、BLE 通信を利用して温度、湿度、照度、気圧、騒音などのデータを記録する。また、UV index や熱中症警戒度といった補助的な情報も取得可能である。1 つのボタン電池 (CR2032) で動作し、サンプリング周期を 5 分に設定した場合、電池寿命は約 6 か月である。

3.4. ドアセンサ

ドアセンサは、ROHM 社製の磁気ドア開閉センサ STM250J を採用した。このセンサは、ドアとドア枠に設置



図 1 システム構成

されたセンサ本体とマグネットの位置関係を検知し、ドアの開閉状態を記録する。EnOcean 規格に対応しており、環境発電による動作が可能である。

3.5. データサーバ

データサーバには、小型 PC である Intel 社の NUC を利用した。このサーバには、BLE 通信用ドングルおよび EnOcean 通信用ドングルが接続され、各センサがブロードキャストするデータを収集する。データは MongoDB に格納され、事前に登録された EEP や BLE MAC アドレスに基づいてセンサ個体情報を識別し保存する。また、ピクセラ社製の PIX-MT100 (LTE 対応 USB ドングル) 等を利用してオンライン接続を可能とし、無線 LAN 環境が整っていない場所でもサーバ監視やデータダウンロードが行える。

3.6. 自己教師あり学習

本研究では、無ラベルデータから特徴量を抽出するために、自己教師あり学習の一手法である DINO (Distillation with No Labels) [11] を採用する。自己教師あり学習は、データに明示的なラベルを付与することなく特徴表現を獲得する手法であり、代表的な手法として SimCLR[12] や MoCo[13] などのコントラスト学習に基づく手法が挙げられる。これらの手法はポジティブサンプルとネガティブサンプルを定義する必要があるが、宅内行動認識のようなタスクにおいて、異なる時間帯や異なるセンサのデータがどの程度類似しているかを明確に定義することは難しい。DINO はポジティブ・ネガティブの関係を明示的に設定する必要がなく、時系列の家庭内行動データから特徴量を取得する本研究のタスクに適していると考えられる。

DINO の特徴量抽出においては、教師ネットワークと生

徒ネットワークの2つのニューラルネットワークを用いる。学習の際、同一の入力データに対し異なるデータオーグメンテーションを適用し、教師モデルの出力する特徴量に生徒モデルの出力を近づけるように学習が進む。この過程において、教師ネットワークのパラメータは勾配更新を受けず、蒸留損失を最小化する形で更新される。この仕組みにより、DINO は自己蒸留による特徴学習を行い、無ラベルデータから有用な特徴量を獲得することができる。

今回の DINO の学習では、以下の手順を実施する。

- 収集した非接触センサデータを一定の時間スライスにリサンプリングする
- データオーグメンテーションを適用し、2つの異なるビューを生成する
- それぞれのビューを、教師ネットワークと生徒ネットワークに入力する
- 教師ネットワークの出力を基準として、生徒ネットワークの出力がこれに近づくように損失を計算し、学習を行う

このアプローチにより、DINO で事前は無ラベルデータを用いた特徴表現の学習を行うことで、家庭の時系列的な特徴量を取得する。

3.7. 能動学習

本研究では、少量のアノテーションデータを効果的に活用しながら高精度な行動認識を実現するために、能動学習を用いた分類器の更新手法を導入する。能動学習は、機械学習モデルの不確実性が高いデータを選択的にアノテーションすることで、限られたラベル付きデータの中で最大限の性能向上を図る手法である。本研究では、DINO によって事前学習された特徴量を用いて分類器を構築し、分類器の確信度をもとに追加アノテーションを行うデータを選択することで、アノテーションの省力化を実現する。

本研究では、DINO で抽出された特徴量を用いて分類器の初期学習を行う。この段階では、すべてのラベル付きデータを用いるのではなく、限定的なラベル付きデータのみを利用する。その後、初期学習した分類器を用いて、さらなる学習を行うべきデータの選定を行う。具体的には、分類器の出力する確信度のスコアを算出し、確信度が一定の閾値を下回るデータを不確実なデータとして抽出する。本研究では、Softmax による確率値の最大値を基準とし、分類器が明確な判断を下せないデータを対象として選択する。このように、モデルの判断が曖昧なデータに対して追加アノテーションを行うことで、分類器の性能を効率的に向上させることを目指す。

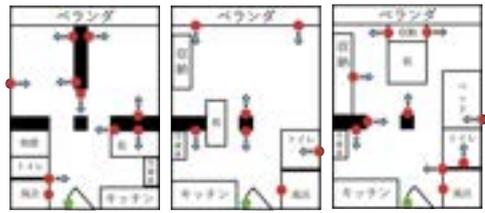


図2 2人居住者のセンサ設置箇所



図3 単身居住者のセンサ設置箇所



図4 人感センサおよび環境センサの設置例



図5 ドアセンサの設置例

4. 実験

4.1. 実験目的および手順

本実験の目的は、提案システムが一般家庭において想定通り運用可能であることを確認し、得られた実環境のデータに対しても、提案手法を用いることでアノテーションコストの省力化が可能であることを検証する事である。

実験対象は、大阪府東大阪市宝持西の集合住宅に在住している家庭4件であり、うち3家庭が2人居住者、1家庭が単身居住者である。

4.2. センサ設置要件

各家庭のセンサの設置箇所について図2, 図3に示す。なお、センサ設置箇所は赤丸で、矢印でセンサの向きを、ドアセンサは緑丸で表現している。それぞれのセンサの設置例を、図4, 図5に示す。

環境発電で動作する人感センサは、各家庭に配置する数を制限するために、10 個程度用意し任意の個所に設置した。ただし、設置個所はベランダ・寝室・居間・キッチン・トイレ・浴室など居住者が頻繁に使用する個所を優先的に設置した。なお、耐水ではないため、浴室に設置する際は浴室前のドアに、地面を見下ろす形で設置した。また Switchbot 製人感センサは 5 個程度用意し、就寝時の寝室など、環境発電が行えないような状況でも利用する部屋に対して優先的に設置した。

環境センサは、モーションセンサと同様の方法で、基本的には、モーションセンサの隣に配置することとしている。サンプリング周期は、3 分間に一度と設定した。なお、設置個所の利用頻度が高くなく、他の環境センサとの距離が近い場合、環境センサは省略し、人感センサのみ設置する事としている。

ドアセンサは、人感センサと同様の方法で玄関に設置した。この際に、ドアの枠側に開閉を検知するためのマグネットを設置し、ドアの可動部側に本体を設置した。

データサーバは、通信強度を高く保つために、家庭内のなるべく中心に近い箇所に配置した。その際、机の下、電話機の横など、邪魔にならない場所へ配置した。

5. データ収集および解析の結果

2024 年 4 月以降にデータ収集に同意した被験者に対して、システムの設置を行った。その上で、実験として単身家庭の 1 週間の ADL データを使用して、DINO の学習と能動学習を組み合わせた行動認識を行った。取得したデータの例として、単身居住者の一日のセンシングデータを図 7、図 8、図 9、図 10 に示す。なお、人感センサおよびドアセンサの反応回数については、10 秒間の反応回数をプロットした。

なお、本研究ではアノテーションの省力化を検証するため、DINO によって抽出された特徴量を用いた分類器の学習を実施した。DINO の学習には無ラベルデータのみを使用しているため、分類器の学習および評価にはラベル付きデータが必要となる。そこで、本研究では被験者の負担を考慮し、行動ラベルの付与を第三者によって行った。

アノテーションの手法としては、時系列センサデータに対し、特定の行動パターンが明確に識別できる時間帯を選定し、適切な行動ラベルを付与する形を取った。具体的に、以下の基準に従い、アノテーションを行った。

- トイレ内の EnOcean 人感センサの反応が顕著である時間帯には、照明を点灯して使用している事が想定されるため「トイレ」のラベルを付与した (プロットの背景色が赤色の期間)

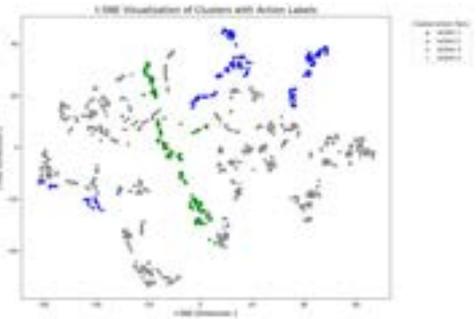


図 6 t-SNE による 4 日間の特徴量の可視化

- 朝と昼のキッチン周辺的环境センサの騒音上昇と、人感センサの反応が顕著である時間帯には、料理を行っている事が想定されるため「料理」のラベルを付与した (プロットの背景色が緑色の期間)
- 夜間の長時間にわたる静的な環境変化、特に人感センサの反応が隠微である時間帯には、就眠したことが想定されるため「睡眠」のラベルを付与した (プロットの背景色が青色の期間)
- それ以外の時間帯には「その他」のラベルを付与した

次に、4 日間の ADL データを DINO で学習した。データの前処理として、環境センサの温度・湿度・照度・気圧・騒音のデータを 10 秒ごとにリサンプリングし、異常値の除去および欠損値を線形補完した。人感センサ・ドアセンサは 10 秒ごとの反応回数をカウントし、これを環境センサのデータと統合し、入力する時系列特徴量とした。その後、3 章で提案した手法に則し、DINO の学習を行った。生成された特徴量の有用性を検証するため、t-SNE を用いて次元削減を行い、特徴空間を 2 次元で可視化した。可視化した図を図 6 に示す。なお、それぞれのプロットは、図 7、図 8、図 9、図 10 と同じく、赤色がトイレ、緑色が料理、青色が睡眠を示す。

最後に行動予測を行うため、作成した DINO モデルの次層に全結合層を追加し、分類器としての訓練を行った。訓練に使用したデータは能動学習を再現するため、以下の 3 パターンに分割して行った。

手法 1: ベースライン

最小限のラベルデータを使用するパターンであり、1 日目と 2 日目のデータを用いて分類器を学習し、5 日目から 7 日目のデータを評価に用いた。この手法では、最低限のラベルデータのみを使用し、どの程度の精度が得られるかを確認した。

手法 2: 完全なアノテーション

全ラベルデータを使用するパターンであり、1 日目から 4 日目のすべてのラベル付きデータを用いて分類器を学習し、5 日目から 7 日目のデータで評価を行った。この手法は、ラ

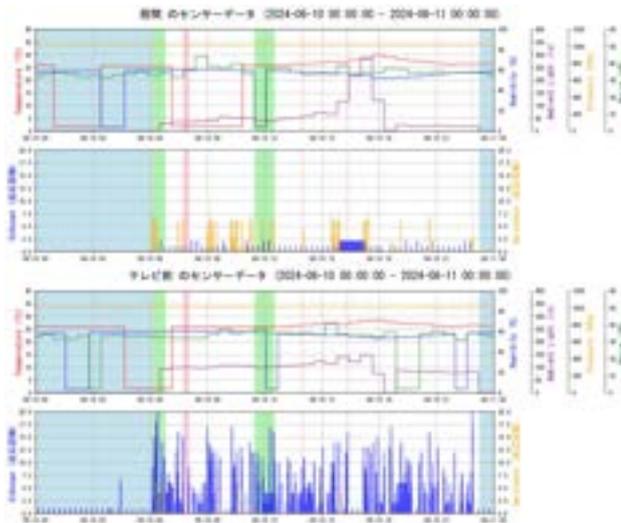


図7 単身居住者の1日のセンシング結果
(居間・テレビ前)

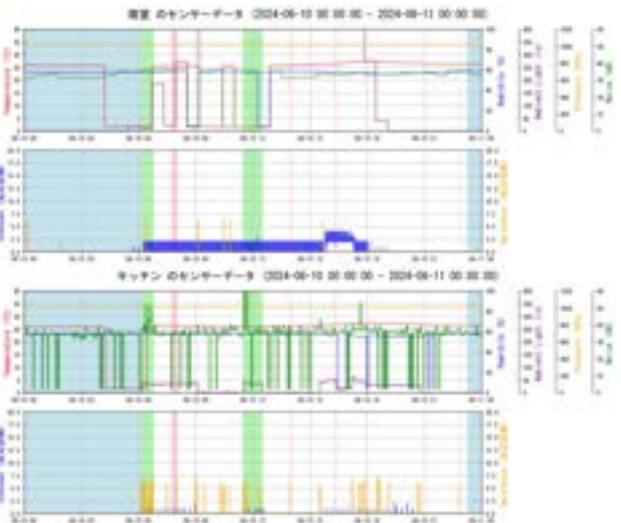


図8 単身居住者の1日のセンシング結果
(寝室・キッチン)

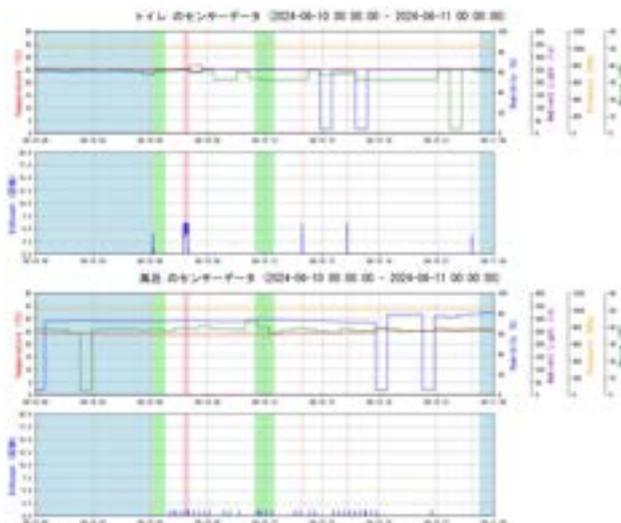


図9 単身居住者の1日のセンシング結果
(トイレ・風呂)

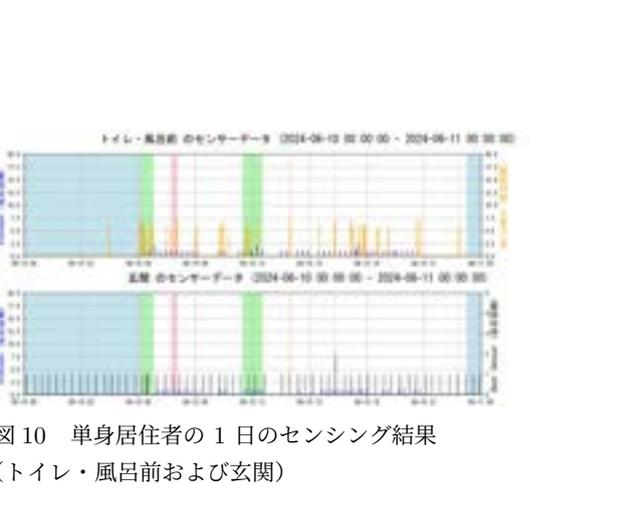


図10 単身居住者の1日のセンシング結果
(トイレ・風呂前および玄関)

ベル付けを最大限に活用した場合の上限性能を測定するための比較対象とした。

手法3: 能動学習

手法1に能動学習を適用するパターンであり、1日目と2日目のデータで分類器を学習した後、分類の確信度が低い3日目と4日目のデータを選択的に学習に使用した。この手法により、最小限のラベル付けで分類精度を向上させられるかを検証した。それぞれの手法の結果を表1に示す。また、図11、図12、図13にそれぞれの手法で得られた混同行列を示す。なお、図11、図12、図13の行動ラベルは、それぞれ1:睡眠、2:トイレ、3:料理、4:その他に対応して

いる。

表1 各手法の行動推定結果

	Accuracy	Avg-Recall	Avg-Precision	Avg-F1-score
手法1	0.902	0.902	0.895	0.890
手法2	0.931	0.931	0.931	0.926
手法3	0.909	0.909	0.904	0.900

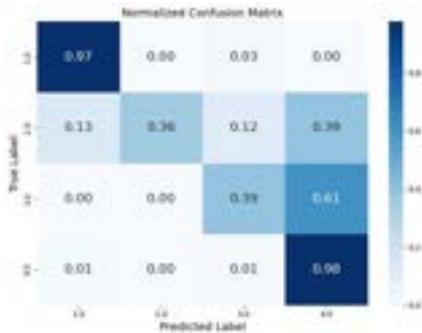


図 11 手法 1 の混同行列

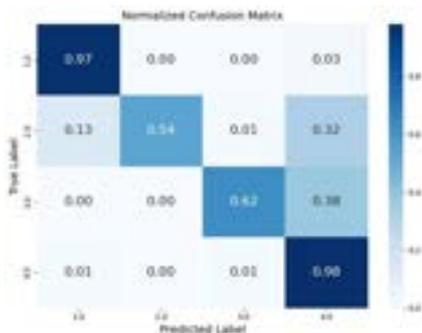


図 12 手法 2 の混同行列

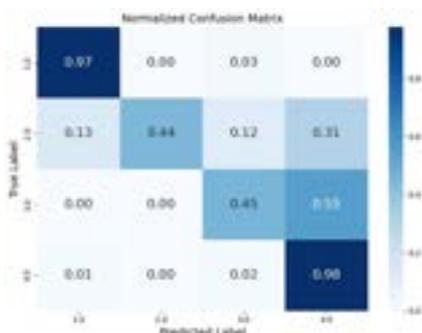


図 13 手法 3 の混同行列

6. 考察

取得されたデータについては、居間・テレビ前・寝室はベランダ側にあり、日中は頻繁に反応する結果となった。さらに、光の弱まる夜間については、EnOcean 製の人感センサの反応回数が減少する結果となった。テレビ前の人感センサについては、環境発電ができず、被験者が就寝していると思われる時間帯でも一定間隔で反応している。これは EnOcean 製の人感センサの元となった温度センサの特性で、蓄電池を搭載しているため、日中に取得できた余剰のエネルギーで動作していると考えられる。

一方で、キッチン・トイレ・風呂などの個所のセンサは、集合住宅における廊下側に位置している。そのため、ベラ

ンダ側よりエネルギーがあまり取得できていないため、反応回数がベランダ側のセンサと比較して少ないと考えられる。特に、トイレは使用していない際は消灯し、使用する際のみ点灯するため、センサ反応がある時、ない時がはっきりと分かれる結果となったと考えられる。

キッチンおよびトイレ・風呂前の EnOcean 製人感センサも反応が少ない結果となったが、ここでは Switchbot 製の人感センサが頻繁に反応している。電池駆動しているセンサを使用したため、環境に左右されずに反応することができたと考えられる。

環境センサのデータについては、温度が数度まで低下する期間が存在するなど、全体的に不連続な点が多い結果となった。これは、Omron 社の環境センサのデータは、測定データを一連の 16 進数の文字列として送信するため、何らかの原因で文字列が欠損した結果であると考えられる。

行動予測については、表 1 より、手法 3 の全体の性能が手法 1 より向上しており、能動学習の効果が表れたと考えられる。ただし、図 11、図 12、図 13 よりすべての手法において、「睡眠」と「その他」の行動は予測できているのに対して、「トイレ」、「料理」の行動は精度が低い。これについては、図 ?? から、「睡眠」や「その他」の行動がほとんどを占めており、逆に「トイレ」、「料理」等の行動が少ない事に起因していると考えられる。

7. おわりに

本研究では、家庭内の行動認識におけるラベル付け作業の負担を軽減するため、自己教師あり学習と能動学習を組み合わせた行動認識手法を提案した。具体的には、人感センサや環境センサから得られる無ラベルデータに対して、DINO を用いた自己教師あり学習を行い、少量のラベル付きデータと能動学習を組み合わせることで、高精度な行動認識モデルを構築することを目指した。実際の一般家庭にセンサを設置し、長期間にわたるセンシングデータを用いた評価実験では、提案手法がラベルデータを最小限に抑えながらも、従来手法と同等以上の分類精度を達成できる可能性を示した。特に、分類器が確信度の低いデータのみをアノテーション対象とする能動学習により、被験者や研究者のラベル作業コストを抑制できる点が確認された。今後は、複数居住者を含むケースへの適用や、より多彩な行動カテゴリへの拡張、さらには学習アルゴリズムの改良による行動クラスタの分離精度向上を検討していく必要がある。

参考文献

- [1] Wang, J., Chen, Y., Hao, S., Peng, X. and Hu, L.: Deep learning for sensor-based activity recognition: A survey, *Pattern Recognition Letters*, Vol. 119, pp. 3–11 (online), <https://doi.org/10.1016/j.patrec.2018.02.010> (2019). Deep Learning for Pattern Recognition.
- [2] Gu, F., Chung, M.-H., Chignell, M., Valaee, S., Zhou, B. and Liu, X.: A Survey on Deep Learning for Human Activity Recognition, *ACM Comput. Surv.*, Vol. 54, No. 8 (online), 10.1145/3472290 (2021).
- [3] Sun, Z., Ke, Q., Rahmani, H., Bennamoun, M., Wang, G. and Liu, J.: Human Action Recognition From Various Data Modalities: A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 3, pp. 3200–3225 (online), 10.1109/TPAMI.2022.3183112 (2023).
- [4] Nguyen, H.-C., Nguyen, T.-H., Scherer, R. and Le, V.-H.: Deep Learning for Human Activity Recognition on 3D Human Skeleton: Survey and Comparative Study, *Sensors*, Vol. 23, No. 11 (online), 10.3390/s23115121 (2023).
- [5] Apicella, G., GiuseppeD’ Aniello, Fortino, G., Gaeta, M., Gravina, R., Tramuto, L. G.: A Situation-aware Wearable Computing System for Human Activity Recognition, *2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)*, pp. 1–7 (online), 10.1109/DASC/PiCom/CBDCCom/Cy55231.2022.9927852 (2022).
- [6] Yang, X., Cao, R., Zhou, M. and Xie, L.: Temporal-Frequency Attention-Based Human Activity Recognition Using Commercial WiFi Devices, *IEEE Access*, Vol. 8, pp. 137758–137769 (online), 10.1109/ACCESS.2020.3012021 (2020).
- [7] Cook, D. J., Crandall, A. S., Thomas, B. L. and Krishnan, N. C.: CASAS: A Smart Home in a Box, *Computer*, Vol. 46, No. 7, pp. 62–69 (online), 10.1109/MC.2012.328 (2013).
- [8] Matsui, T., Onishi, K., Misaki, S., Fujimoto, M., Suwa, H. and Yasumoto, K.: SALON: Simplified Sensing System for Activity of Daily Living in Ordinary Home, *Sensors*, Vol. 20, No. 17 (online), 10.3390/s20174895 (2020).
- [9] Lewis, D. D.: A sequential algorithm for training text classifiers: corrigendum and additional data, *SIGIR Forum*, Vol. 29, No. 2, p. 13–19 (online), 10.1145/219587.219592 (1995).
- [10] Settles, B. and Craven, M.: An Analysis of Active Learning Strategies for Sequence Labeling Tasks, *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing* (Lapata, M. and Ng, H. T., eds.), Honolulu, Hawaii, Association for Computational Linguistics, pp. 1070–1079 (online), <https://aclanthology.org/D08-1112/> (2008).
- [11] Caron, M., Touvron, H., Misra, I., Jegou, H., Mairal, J., Bojanowski, P. and Joulin, A.: Emerging Properties in Self-Supervised Vision Transformers, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9630–9640 (online), 10.1109/ICCV48922.2021.00951 (2021).
- [12] Chen, T., Kornblith, S., Norouzi, M. and Hinton, G.: A simple framework for contrastive learning of visual representations, *Proceedings of the 37th International Conference on Machine Learning, ICML’20, JMLR.org* (2020).
- [13] He, K., Fan, H., Wu, Y., Xie, S. and Girshick, R.: Momentum Contrast for Unsupervised Visual Representation Learning, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9726–9735 (online), 10.1109/CVPR42600.2020.00975 (2020).

メタバース空間におけるファントムセンスの実証研究

齋藤天馬

公立はこだて未来大学

角薫

公立はこだて未来大学

1. はじめに

メタバースの発展とともに、仮想環境における知覚の研究が注目されている。その中でも、「ファントムセンス (Phantom Sense)」と呼ばれる現象は、物理的な刺激が存在しないにもかかわらず、触覚や温感、圧力感覚などの知覚が生じる興味深い錯覚である [1]。この現象は、視覚や聴覚などの他の感覚から得られる情報をもとに脳が生成するものであり、実際の物理的接触がなくても非常にリアルな感覚を伴うことが特徴である。

ファントムセンスに関する先行研究では、視覚的要素や期待効果が錯覚の発生に影響を与えることが示唆されている [2]。たとえば、VR 環境で仮想の棒を手に持ち、それで自らのアバターの手を撫でると、実際には存在しない触覚が生じることが報告されている [2]。メタバース空間で自身のアバターが他人に触れられると、現実の身体にも触れられたように感じるケースがあり、特に没入感の高い VR ユーザに顕著である [3,9]。さらに、視覚的な温度表現がユーザの温感知覚に影響を与えることも報告されており [4]、これらの研究は、ファントムセンスが VR 環境における知覚の重要な要素であることを示唆している。

日本国内でも、メタバース環境でのファントムセンスの事例が取り上げられている。たとえば、「メタバース進化論」では、メタバース内で自身のアバターが他人に触れられると、実際に自分の体も触れられたように感じる現象が生じることが述べられている [6]。さらに、存在しない尻尾がアバターに付いている場合、それを触られると、まるで実際に触れられたかのような感覚を覚えることもあるとされている。このようなファントムセンスは、メタバース内での高所からの落下体験や、その他の触覚経験にも影響を及ぼすことが指摘されている。

ファントムセンスの感覚は個人差が大きく、VR の使用頻度が高い人が必ずしもファントムタッチを感じやすいわけではないと報告されている [1]。一方で、新しいユーザでもこの感覚を経験する可能性があるとしており、必ずしも長期的な VR 体験がファントムセンスの発生に直結するわけではない。ファントムタッチの具体的な知覚としては、「チクチクする感覚」(47.5%) が最も一般的であり、次いで「熱」(17.5%) や「圧力」(12.5%) といった感覚が報告されている。その他、ブラッシング、くすぐったさ、電気

的な感覚、振動、単純な存在感、または実際に触れられたような感覚を経験するユーザもいる。また、時間が経つにつれて「チクチクする感覚」を感じるユーザは、他の具体的な感覚も発達させやすいことが示されている。

しかし、これまでの研究では、ファントムセンスがどの程度の確率で発生するのか、またその要因が何であるのかについて、実証的な研究は限られていた。特に、身体的な覚醒状態や事前の運動経験がファントムセンスの発生に及ぼす影響については、十分に検討されていない。本研究では、ファントムセンスの発生率を定量的に評価し、メタバース体験における感覚知覚のメカニズムを明らかにすることを目的とする。

また、ファントムセンスがメタバースの没入感やユーザのエンゲージメントに与える影響についても検討する。仮想環境における体験の没入感やリアリティの評価指標が提案されている [7]。エンターテインメントの「面白さ」や「世界観」が、ユーザの持続的な関与につながる事が指摘されている [10]。しかし、ファントムセンスが直接的に再訪動機と関連するののかについての実証研究はほとんど存在しない。本研究では、メタバース体験におけるファントムセンスの影響を多角的に分析し、VR 環境設計への示唆を提供することを旨とする。

2. 仮説

本研究では、メタバース環境におけるファントムセンスの発生率とそれに影響を与える要因を明らかにし、さらにファントムセンスがメタバース体験や再訪動機に与える影響を調査することを目的とする。特に、事前運動がファントムセンスの発生に及ぼす影響、およびファントムセンスの知覚がメタバースの没入感や再訪意向にどのような関係を持つのかを検証する。これに基づき、以下の仮説を設定する。

H1 (ファントムセンスの発生率) :

メタバース環境では一定の確率でファントムセンスが発生する。

- これまでの研究では、視覚・聴覚情報の統合により触覚が生じる可能性が示唆されていた [1]。また、温感に関しては、視覚的な温度表現が影響を与えることが報告されている [4]。しかし、どの程度のユーザがこの現象を経験するのかは未検証である。本研究では、異なるメタ

Investigating Phantom Sensation in the Metaverse
†1 TENMA SAITO, Hakodate Future University
†2 KAORU SUMI, Hakodate Future University

バース環境でファントムセンスの発生頻度を測定する。

H2 (事前運動の影響) :

メタバース体験前の身体的興奮状態がファントムセンスの発生率に影響を与える。

- 本研究では、励起伝達理論 (Excitation Transfer Theory) を適用し、身体的興奮が残留することでファントムセンスが増強される可能性を検証する [5].

H3 (ファントムセンスの期待効果) :

メタバース体験前にファントムセンスの概念を知ること、実際のファントムセンスの発生確率が変化する。

- 知覚において、期待効果 (Expectation Effect) が体験の強度に影響を与えることが知られている [2].
- 先行研究では、ミラーニューロンシステムが知覚の強化に関与し、他者が感じていることを観察することで自身の感覚が強化される可能性が指摘されている [8].
- そこで、ファントムセンスの説明を事前に受けた群とそうでない群で、ファントムセンスの発生率に違いが生じるかを検証する。

H4 (2人での体験がファントムセンスに与える影響) :

2人でメタバースを体験することで、ファントムセンスの発生率やメタバース体験の質が変化する。

- ゲーム体験の没入感や感覚の強化に関する指標が提案されている [7].
- メタバース内で他者と相互作用することで、触覚の錯覚が強化される可能性がある [9].
- 2人群と1人群の比較を行い、ファントムセンスの発生率や体験の主観的評価が異なるかを分析する。

H5 (ファントムセンスの発生とメタバース体験の質) :

ファントムセンスを経験することで、メタバース体験の没入感や面白さが向上する。

- メタバース体験の没入感や体験の質を評価する指標が提案されている [7].
- ファントムセンスを感じた群と感じなかった群で、メタバース体験の主観的評価(面白さ,リアリティ,没入感)が異なるかを検証する。

H6 (ファントムセンスと再訪動機) :

ファントムセンスを経験することで、メタバースへの再訪意向が変化する。

- 仮想環境のリアリティが高まることで、ユーザがその環境を繰り返し訪れたいくなる可能性がある [10].
- しかし、これまでの研究では、ファントムセンスが再訪動機に直接的な影響を及ぼすかどうかは未検証であるため、本研究で実証的に分析する。

3. メタバース実験システムの開発

本研究では、VRChat プラットフォームを活用し、メタバース環境内でのファントムセンスの発生を調査するための実

験システムを開発した。実験環境は、Blender, Unity, Udon Graph を用いて構築し、ユーザがインタラクティブな体験を行えるよう設計した。特に、触覚や温感を誘発する視覚的・聴覚的フィードバックを強調し、リアルな感覚錯覚を引き起こす仕組みを取り入れた。

3.1 メタバースの体験順

被験者には、以下の4つのメタバース環境を体験してもらい、ファントムセンスの発生率を測定した (図 1)。

- チュートリアルワールド - 操作方法の習得 (図 2)
- バンジージャンプワールド - 落下感のファントムセンス誘発 (図 3)
- サイクリングワールド - 衝突感のファントムセンス誘発 (図 4)
- 温泉ワールド - 温感のファントムセンス誘発

実験では、各ワールドを体験した後にアンケートを実施し、ファントムセンスの発生率と被験者の主観的評価を収集した。体験順序の影響を抑えるため、カウンターバランスを考慮してワールドの順番を変更しながら実験を行った。

3.2 チュートリアルワールド

チュートリアルワールド (図 2) は、被験者が VRChat の基本操作を習得し、メタバース環境に適応することを目的として設計した。Meta Quest 3 を使用し、以下の操作を練習できる環境を構築した。

- アバターの移動 (歩行, 座る)
- オブジェクトの持ち上げ
- 乗り物の操作 (サイクリングワールド用)
- ポータルを用いたメタバース間の移動

この環境は Unity と Udon Graph を用いて開発し、スムーズな体験を提供するために最適化を行った。また、バンジージャンプワールドやサイクリングワールドのような過激な体験を和らげるため、アニメーション付きのキツネのオブジェクトや明るい BGM を追加し、心理的負担を軽減した。

3.3 バンジージャンプワールド

バンジージャンプワールド (図 3) は、落下感のファントムセンスを誘発するために設計した。被験者には、以下の手順で体験してもらった。

1. 高所のエレベーターに乗り、上昇
2. 半透明のガラス床へ移動
3. 椅子に座ると 6 秒後に自動で落下開始
4. 落下中、風を切る音や視覚的な加速感を付与
視覚情報と音響効果を組み合わせることで、リアルな落下感覚をシミュレートした。

3.4 サイクリングワールド

サイクリングワールド (図 4) では、衝突時のファントム

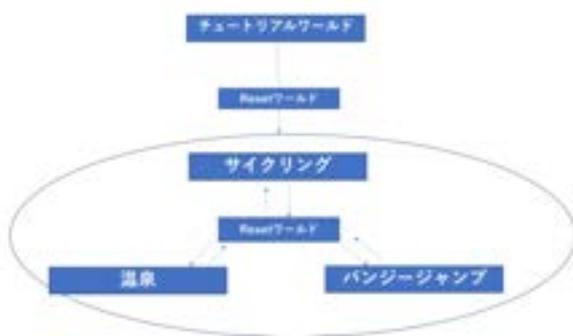


図1 メタバス体験の順番



図3 バンジージャンプワールド



図2 チュートリアルワールド

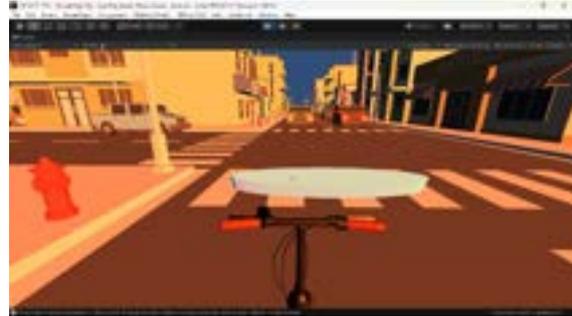


図4 サイクリングワールド



図6 リセットワールド



図5 温泉ワールド

センスを誘発することを目的とした。被験者は、以下のプロセスで体験を行った。

1. 自転車を操作し、交差点まで前進（後退不可）
2. クラクション音による心理的圧迫を追加
3. 信号無視の車に衝突される
4. 衝突の方向に吹き飛ばされ、血しぶきのエフェクトとリバーブした悲鳴音を追加

このワールドは、VR環境における危険体験の再現を意図しており、現実には存在しない衝撃感覚を引き起こす設計となっている。

3.5 温泉ワールド

温泉ワールド（図5）では、温感のファントムセンスを誘発することを目的とし、以下の視覚的・音響的演出を採用した。

- 湯気のエフェクトと画面全体の色温度変化

- 温かみのある照明とBGM
- 仮想的な湯船に浸かるアニメーション

このワールドは、特に視覚による温感の錯覚を誘発するように調整されており、ユーザが「温かい」と感じるかどうかを評価した。

3.6 リセットワールド

各メタバス体験後に、心理状態をリセットするための環境を用意した（図6）。ファントムセンスによる影響を次のワールドへ持ち越さないようにするため、落ち着いたBGMを流すことでリラックスできる空間を提供した。

3.7 メタバスで使用したアバター

被験者がメタバスを体験する際に使用するアバター（図7）は、VRoid Studioで作成されたデフォルトキャラクターを、VRChat Creator Companionを用いてアップロードし、最適化した。すべての被験者が同じアバターを使用



図7 メタバース体験用アバター

することで、外見の違いがファントムセンスの発生に与える影響を排除し、一貫した体験が得られるようにした。

4. 実験

4.1 実験概要

本研究では、メタバース環境におけるファントムセンスの発生率と、それに影響を与える要因を調査することを目的とした。さらに、ファントムセンスの知覚がメタバース体験の質や再訪動機にどのように関係するのかを明らかにするため、被験者を異なる条件群に分けて実験を行った。

実験では、以下の4つの実験群を設定し、それぞれの条件下でファントムセンスの発生率、体験の質、再訪動機の強さを比較した。

- 統制群 (Control Group) : 標準的なメタバース体験を実施し、ファントムセンスの基礎的な発生率を測定する。
- 事前運動群 (Pre-Exercise Group) : メタバース体験前に運動を行い、身体的興奮状態がファントムセンスの発生率に与える影響を検証する。
- 説明群 (Explanation Group) : メタバース体験前にファントムセンスの概念を説明し、期待効果の影響を調査する。
- 2人群 (Two-Person Group) : 2人で同時にメタバースを体験し、他者との相互作用がファントムセンスの発生や体験の質に与える影響を分析する。

各群の被験者数は、統制群、事前運動群、説明群がそれぞれ11名、2人群が10名である。

実験の前に、被験者には実験参加に関する同意書と運動に支障がないかを確認する書類に署名してもらい、事前アンケートによりファントムセンスやVR体験に関する知識の有無を確認した。

実験では、VRゴーグル (Meta Quest 3) を使用し、チュートリアルワールドを経た後、バンジージャンプワールド、サイクリングワールド、温泉ワールドを体験してもらった。各ワールドの体験後に事後アンケートを実施し、ファントムセンスの有無やメタバース体験の主観的評価を収集した。

最後に、各群の結果を比較し、メタバース体験の質がどのように変化するのかを検証した。

4.2 被験者と条件群

本研究では、大学生43名を被験者として募集した。被験者の内訳は、男性32名、女性11名で、平均年齢は20.3歳である。

すべての被験者は、VRゴーグル (Meta Quest 3) を装着し、バンジージャンプワールド、サイクリングワールド、温泉ワールドの3つのメタバース体験を行った。なお、体験の順序はカウンターバランスを考慮し、被験者ごとに異なる順番で実施した。

また、事前アンケートにおいてVRゲームのプレイ経験を調査した結果、76.7%の被験者がVRゲームを未経験であり、月に1~3回程度プレイする者が20.9%、毎日プレイする者が2.3%であった。この結果から、本実験の被験者の大半がVR体験に不慣れであることが分かった。

4.3 統制群 (Control Group) の実験

統制群の被験者11名には、VRChatのメタバース環境を標準的な手順で体験してもらい、ファントムセンスの基礎的な発生率を測定した。実験手順は以下の通りである。

まず、チュートリアルワールドでVR操作に慣れた後、バンジージャンプワールド、サイクリングワールド、温泉ワールドの3つのメタバースを順番に体験してもらった。それぞれのワールド体験後にはアンケートに回答し、ファントムセンスの有無や体験の質を評価した。次のメタバースを体験する前にはリセットワールドを経由し、被験者の心理状態をリセットすることで、各ワールドの影響を次に持ち越さないようにした。

この群の被験者は、事前アンケートでファントムセンスについての知識がないことが確認されており、また、事前運動 (心拍数を上げる運動) を行わない条件で実験を実施した。

4.4 事前運動群 (Pre-Exercise Group) の実験

事前運動群の被験者11名には、統制群と同様のメタバース体験を行う前に、サイクリングマシンを使用した運動を実施してもらった。

まず、チュートリアルワールドで基本操作を習得した後、サイクリングマシンに乗り、心拍数を110以上に維持しながら3分間運動してもらった。その後、被験者が「心拍が落ち着いた」と申告した後に、バンジージャンプワールド、サイクリングワールド、温泉ワールドを順番に体験した。各ワールド体験後にはアンケートを記入し、リセットワールドを経由して次のワールドへ進む手順を採用した。

この群では、事前運動がファントムセンスの発生率やメタバース体験の質に影響を与えるかどうかを検証した。

4.5 説明群 (Explanation Group) の実験

説明群の被験者11名には、統制群と同じメタバース体験

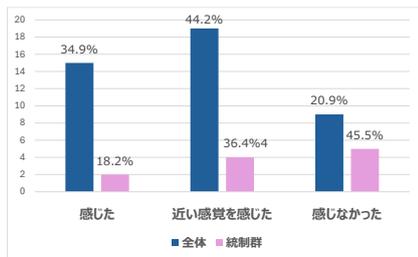


図 8 バンジージャンプの落下感の割合 (全体・統制群比較)

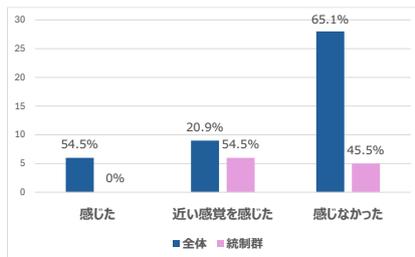


図 9 温泉の温感の割合 (全体・統制群比較)

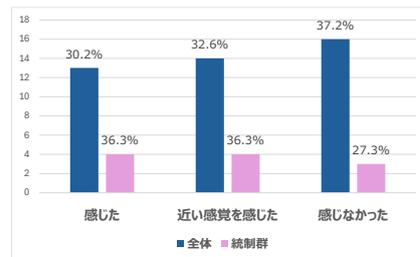


図 10 サイクリングワールドの衝突感の割合 (全体・統制群比較)

を行う前にファントムセンスに関する説明を受けてもらった。

まず、チュートリアルワールドで VR 操作に慣れた後、ファントムセンスとは何かについて説明を行った。具体的には、「ファントムセンスは、実際の刺激がないにもかかわらず、リアルな感覚として知覚される現象であり、視覚や聴覚情報によって触覚や温感が生じることがある」ことを被験者に伝えた。その後、バンジージャンプワールド、サイクリングワールド、温泉ワールドを順番に体験し、各ワールドの後にアンケートに回答してもらった。

この群では、事前にファントムセンスの概念を説明することで、期待効果がファントムセンスの発生率や体験の質に影響を与えるかどうかを検証した。

4.6 2 人群 (Two-Person Group) の実験

2 人群の被験者 10 名には、統制群と同じメタバース体験を行うが、2 人 1 組で同時に体験する形式を採用した。

まず、チュートリアルワールドで基本操作を習得した後、2 人でバンジージャンプワールド、サイクリングワールド、温泉ワールドを体験した。各ワールド体験後にはアンケートを記入し、リセットワールドを経由して次のワールドへ進んだ。

この群では、他者と一緒に体験することが、ファントムセンスの発生率やメタバース体験の質に影響を与えるかどうかを検証した。

5. 結果

本章では、メタバース環境におけるファントムセンスの発生率、メタバース体験の面白さ、再訪動機の強さについて各実験群の比較を行い、有意な差があるかどうかを統計的に検定する。また、事後アンケートの回答に基づき、被験者を条件別に分類し、各指標の関連性を分析する。

5.1 ファントムセンスはどれくらいの人が感じたか？

まず、被験者 43 名に対し、バンジージャンプワールド (落下感)、サイクリングワールド (衝突によるぞわぞわ感)、温泉ワールド (温感) の 3 つのメタバース体験においてファントムセンスを感じたかどうかを調査した。回答は「感

じた」「近い感覚を感じた」「感じなかった」の 3 択で行われた。

- バンジージャンプワールドでは、全体では 15 名 (34.9%) が「感じた」、19 名 (44.2%) が「近い感覚を感じた」と回答した。統制群では、「感じた」は 2 名、「近い感覚を感じた」は 4 名であり、統制群ではファントムセンスを強く感じる割合がやや低かった。
- サイクリングワールドでは、全体では 13 名 (30.2%) が「感じた」、14 名 (32.6%) が「近い感覚を感じた」と回答した。統制群では、「感じた」は 4 名、「近い感覚を感じた」は 4 名で、統制群でも一定の割合でファントムセンスが報告された
- 温泉ワールドでは、全体では 6 名 (14.0%) が「感じた」、9 名 (20.9%) が「近い感覚を感じた」と回答した。一方で、統制群では「感じた」と答えた被験者は 0 名であり、「近い感覚を感じた」は 6 名だった。

統制群のファントムセンス発生率を比較するため、各ワールドの統制群データを加えたグラフを図 8, 9, 10 に示す。結果として、バンジージャンプワールドでは落下感、サイクリングワールドでは衝突感を感じる被験者が多く、温泉ワールドは他の 2 つに比べて発生率が低かった。

これらの結果から、ファントムセンスは一定の割合で発生し、特に視覚と体感の結びつきが強い場面 (落下や衝突) で多くの被験者が経験する傾向があることが示唆された。また、統制群と比較すると、実験群のほうがファントムセンスの発生率が高い傾向にあり、メタバース体験の内容や没入感の違いが影響を及ぼしている可能性がある。

5.2 各ワールドはどのように評価されたのか？

各ワールドの面白さと再訪動機について、5 段階評価のアンケートを実施した。その結果、面白さの平均値は以下のようになった。

- バンジージャンプワールド : 3.65 点
 - サイクリングワールド : 4.06 点
 - 温泉ワールド : 3.25 点
- また、再訪動機の平均値は以下の通りであった。
- バンジージャンプワールド : 3.27 点

- サイクリングワールド : 3.41 点
- 温泉ワールド : 3.11 点

これらの結果から、最も面白いと評価されたのは サイクリングワールド であり、再訪動機も最も高かった。温泉ワールドは相対的に評価が低かったものの、それでも一定の関心を集めていた。

5.3 事前に説明するとどうなるのか？

ファントムセンスについて事前に説明を受けた群（説明群）と、説明を受けなかった群（統制群）を比較したところ、ファントムセンスの発生率、メタバース体験の面白さ、再訪動機のいずれにおいても有意な差は見られなかった ($p > 0.05$)。この結果から、ファントムセンスの発生は事前情報による期待効果に大きく依存しないことが示唆される。

5.4 事前に心拍数を上げるとどうなるのか？

事前に運動を行った群（事前運動群）と統制群を比較したところ、以下の結果が得られた（表 1-1, 1-2, 1-3）。

- サイクリングワールドにおいて、ファントムセンスを感じた人数に有意差が見られた ($p < 0.011$)。
 - 事前運動群では、ファントムセンスを感じた人が少なかった（表 1-3）。
 - これは、事前にリアルなサイクリングを行ったことで、メタバースの体験と現実の体験が乖離し、バーチャルな衝突感が薄れた可能性を示唆する。
- バンジージャンプワールド ($p < 0.38$) および温泉ワールド ($p < 0.077$) では有意な差は見られなかった。

表 1-1 統制群と事前運動群のファントムセンス比較 (バンジージャンプワールド)

群	感じた	感じなかった	合計
統制群 (バンジー)	6 (54.5%)	5 (45.5%)	11
事前運動群 (バンジー)	8 (72.7%)	3 (27.3%)	11
合計	14 (63.6%)	8 (36.4%)	22

表 1-2 統制群と事前運動群のファントムセンス比較 (温泉ワールド)

群	感じた	感じなかった	合計
統制群 (温泉)	6 (54.5%)	5 (45.5%)	11
事前運動群 (温泉)	2 (18.2%)	9 (81.8%)	11
合計	8 (36.4%)	14 (63.6%)	22

表 1-3 統制群と事前運動群のファントムセンス比較 (サイクリングワールド)

群	感じた	感じなかった	合計
統制群 (サイクリング)	8 (72.7%)	3 (27.3%)	11
事前運動群 (サイクリング)	2 (18.2%)	9 (81.8%)	11
合計	10 (45.5%)	12 (54.5%)	22

5.5 2人だとファントムセンスを感じやすいのか？

2人群と統制群を比較したところ、ファントムセンスの発生率には有意差がなかった ($p > 0.05$)。しかし、温泉ワ

ールドにおけるメタバース体験の面白さに有意差が見られた ($p < 0.046$) (表 2-1, 2-2, 2-3)。

この結果は、他者とメタバース体験を共有することで、より面白いと感じる可能性があることを示唆する。特にリラックスを伴う体験では、ソーシャルな要素が重要な役割を果たしていると考えられる。

表 2-1 バンジージャンプワールドの面白さ (統制群・2人群)

群	M (SD)
統制群 (バンジー)	3.36 (1.02)
2人群 (バンジー)	4.1 (0.99)

表 2-2 温泉ワールドの面白さ (統制群・2人群)

群	M (SD)
統制群 (温泉)	3.27 (0.90)
2人群 (温泉)	4.10 (0.87)

表 2-3 サイクリングワールドの面白さ (統制群・2人群)

群	M (SD)
統制群 (サイクリング)	4.09 (1.37)
2人群 (サイクリング)	4.40 (0.96)

5.6 GEQ を使った検定

ゲーム体験の質を測定するため、GEQ (Game Experience Questionnaire) のコアモジュール[7]を使用し、被験者のスコアを分析した。

高得点者 (ファントムセンスを強く感じた群) と低得点者 (ファントムセンスを感じなかった群) を比較したところ、「没入感」「フロー体験」「ポジティブ感情」の3つの要素で有意差があった ($p < 0.05$) (表 3, 4, 5)。

この結果から、ファントムセンスの経験がメタバースに対する没入感を高め、ポジティブな体験を増幅する可能性があることが示唆される。

表 3 没入感スコアと面白さ (ファントムセンス高・低得点群)

ファントムセンス	M (SD)
高得点	3.89 (0.65)
低得点	2.5 (0.96)

表 4 フロー体験スコアと面白さ (ファントムセンス高・低得点群)

ファントムセンス	M (SD)
高得点	3.65 (0.90)
低得点	2.45 (1.13)

表 5 ポジティブ感情スコアと面白さ (ファントムセンス高・低得点群)

ファントムセンス	平均値
高得点	4.03 (0.52)
低得点	2.85 (1.07)

5.7 ファントムセンスを体験した群とそうで無い群

ファントムセンスを体験した群と体験しなかった群のメタバース体験の評価を比較した結果、バンジージャンプワールドと温泉ワールドでは面白さの評価に有意差があった ($p < 0.012, p < 0.013$) (表 6-1, 6-2).

再訪動機には有意な差はなかったものの、ファントムセンスを感じた群はスコアがより高い傾向にあった (表 7-1, 7-2, 7-3).

表 6-1 バンジージャンプの面白さ (ファントムセンス別)

ファントムセンス	M (SD)
感じた	3.84 (1.09)
感じなかった	3.00 (0.81)

表 6-2 温泉の面白さ (ファントムセンス別)

ファントムセンス	M (SD)
感じた	3.80 (0.94)
感じなかった	2.96 (1.13)

表 12 バンジージャンプ再訪 (ファントムセンス別)

ファントムセンス	M (SD)
感じた	3.45 (1.34)
感じなかった	2.70 (1.49)

表 13 温泉再訪 (ファントムセンス別)

ファントムセンス	M (SD)
感じた	3.73 (0.96)
感じなかった	2.78 (1.10)

表 14 サイクリング再訪 (ファントムセンス別)

ファントムセンス	M (SD)
感じた	3.59 (1.57)
感じなかった	3.30 (1.47)

5.8 性格によってメタバース体験に違いがあるか

TIPI-J に基づき外向性, 誠実性, 神経症的傾向, 開放性などの因子を評価し, これらの性格特性がファントムセンスの発生やメタバース体験に与える影響を検討した. 被験者の性格特性 (Big Five) [11] とファントムセンスの発生率, 体験の質, 再訪動機の関係进行分析したが, 有意差は見られなかった ($p > 0.05$). この結果から, 性格特性がメタバース体験の質に大きく影響を与えるとは言えないことが示唆された.

6. 議論

本研究では, メタバース環境におけるファントムセンスの発生率や影響を実証的に調査し, 事前運動や説明, 2人での体験, 性格特性がどのように関与するかを検証した. その結果, いくつかの興味深い知見が得られた. 本章では, それぞれの仮説との対応を整理し, 得られた結果の意味を考察する.

6.1 仮説の検証

本研究で設定した仮説 (H0~H6) について, 結果との対応を以下に整理する.

- **H0 (基礎仮説)**: 「メタバース環境では一定の割合の人がファントムセンスを経験する」
結果: 検証の結果, バンジージャンプワールドでは 34 人 (約 79.0%), サイクリングワールドでは 27 人 (約 62.7%), 温泉ワールドでは 15 人 (34.8%) の被験者がファントムセンスを経験した. これは, 仮説 H0 が支持されることを示唆する.
- **H1 (統制群)**: 「通常のメタバース体験でもファントムセンスは生じる」
結果: 統制群においても一定数の被験者がファントムセンスを報告しており, H1 は支持された.
- **H2 (事前運動群)**: 「事前運動はファントムセンスを強化する」
結果: 予想に反し, 事前運動群ではサイクリングワールドのファントムセンスが低下した ($p < 0.011$). これは, 現実のサイクリング体験がバーチャル体験との差異を認識させ, 感覚のリアルさを低減させた可能性がある.
- **H3 (説明群)**: 「ファントムセンスを事前に説明すると, その発生率が上がる」
結果: ファントムセンスの発生率に有意な影響は見られなかった ($p > 0.05$). 期待効果による影響は限定的だった可能性がある.
- **H4 (2人群)**: 「2人でメタバースを体験するとファントムセンスが強まる」
結果: ファントムセンスの発生率に有意差はなかったが, 温泉ワールドではメタバース体験の面白さが向上した ($p < 0.046$).
- **H5 (性格特性)**: 「性格特性 (Big Five) はファントムセンスの発生に影響を与える」
結果: Big Five のいずれの要素とも有意な関連は見られなかった.
- **H6 (再訪動機)**: 「ファントムセンスを感じるとメタバースの再訪動機が高まる」
結果: ファントムセンスを感じた群と感じなかった群で再訪動機に有意な差は見られなかった. ただし, ファントムセンスを感じた群の方が平均スコアは高かった.

6.2 ファントムセンスの種類とワールドごとの違い

各ワールドにおけるファントムセンスの違いを分析すると, 興味深い傾向が見られた.

- **バンジージャンプワールド (落下感)**:
 体験者の 79.0%がファントムセンスを感じたまたは近い感覚を感じたが, 完全に落下感を経験したのは 2 人のみだった.
 「ぐわんとする感覚」「バウンドする感覚」の報告があり視覚情報による錯覚の影響が示唆された.
- **サイクリングワールド (衝突感)**:
 体験者の 62.7%がファントムセンスを感じたまたは近い

感覚を感じた。

「椅子に座っているにもかかわらず、動いている感覚があった」「車に飛ばされた感覚があった」といった報告があり、視覚情報の影響が強いことが示唆された。

しかし、事前運動群ではファントムセンスが低下しており、現実の自転車体験が逆に影響を与えた可能性がある。

● 温泉ワールド (温感) :

体験者の 34.8%がファントムセンスを感じたが、「温かそうというより冷たそうに感じた」との報告もあった。視覚情報による温感の再現が不十分だった可能性があり、今後の改善が必要である。

6.3 メタバース体験の質への影響

GEQ の分析結果から、ファントムセンスを感じた被験者は「没入感」「フロー体験」「ポジティブ感情」が高かったことが示された ($p < 0.01$)。これは、ファントムセンスがメタバース体験の質を向上させる可能性を示唆する。

6.4 メタバース設計への示唆

本研究の結果を踏まえ、メタバース設計において以下の示唆が得られた。

1. 視覚情報の工夫:

落下感の強化には視覚的な動きのエフェクトや浮遊感を増す演出が有効である。

温感の再現には、視覚情報の調整(湯気の質感や色温度の調整)や触覚デバイスの併用が求められる。

2. リアルな経験の影響:

事前運動がバーチャル体験の知覚に影響を与える可能性がある。これは現実世界での類似経験による影響としても考えられる。サイクリング体験後のサイクリングワールド体験のように、現実とメタバースの体験が乖離している場合、ファントムセンスが低下することがある。

3. 社会的要因:

2人での体験が没入感や面白さを高める可能性がある。単独での体験と比べ、メタバースでの社会的相互作用の影響を今後さらに検証する価値がある。

7. まとめ

本研究では、メタバースにおけるファントムセンスの発生率や影響を実証的に検証した。その結果、ファントムセンスは一定の割合で発生するものの、事前運動が感覚を低減させるなど、予想外の結果も得られた。また、視覚的な演出や社会的要因が体験の質に影響を与える可能性が示唆された。

さらに、事前に運動を行うとファントムセンスの発生率が低下することが示唆された。また、現実で直前に行った行動と類似する動作をメタバース内で行う場合にファントムセンスの発生率が低下する可能性があることが分かった。一方で、シンプルな操作性を維持し、過激な描写を控える

ことで、ファントムセンスが誘発されやすくなる可能性がある。

また、メタバース体験の面白さに関しては、1人で体験するよりも2人1組で体験する方がより没入感が高まり、楽しさが増すことが示された。例えば、バレーボールなどのスポーツをメタバース内で実装し、プレイヤー同士の相互作用を促進することで、ファントムセンスの活用による没入感の向上が期待される。

本研究の結果は、メタバースの設計において、ファントムセンスを意図的に活用することが、没入感やポジティブな感情を高める手段となることを示唆している。今後の研究では、長期的なメタバース体験への影響や、触覚デバイスとの連携によるさらなる没入感の向上について検討することが求められる。

参考文献

[1] S. Alexdotir and X. Yang, "Phantom Touch phenomenon as a manifestation of the Visual-Auditory-Tactile Synaesthesia and its impact on the users in virtual reality," 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, Singapore, 2022, pp. 727-732.

[2] Pilacinski, A., Metzler, M., & Klaes, C. "Phantom touch illusion, an unexpected phenomenological effect of tactile gating in the absence of tactile stimulation." Scientific Reports, 13, 15453 (2023).

[3] Qijia Chen, Michiel M. Spapé, and Giulio Jacucci. "Understanding Phantom Tactile Sensation on Commercially Available Social Virtual Reality Platforms." Proceedings of the ACM on Human-Computer Interaction, 8, CSCW1, Article 141 (April 2024), 22 pages.

[4] Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. "Too Hot to Handle: An Evaluation of the Effect of Thermal Visual Representation on User Grasping Interaction in Virtual Reality." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1-16.

[5] Zillmann, D. "Excitation Transfer Theory." The International Encyclopedia of Communication, Vol. 4, pp. 1627-1632, 2008.

[6] バーチャル美少女ねむ. メタバース進化論. 技術評論社, 2022.

[7] IJsselsteijn, W., de Kort, Y. A. W., & Poels, K. The Game Experience Questionnaire. Eindhoven: Technische Universiteit Eindhoven, 2013.

[8] Rizzolatti, G., & Craighero, L. (2004). "The Mirror-Neuron System." Annual Review of Neuroscience, 27, 169-192.

[9] Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). "First Person Experience of Body Transfer in Virtual Reality." PLoS ONE, 5(5), e10564.

[10] 都留泰作.<面白さ>の研究. 角川新書,2015,301p.

[11] 小塩 真司, 阿部 晋吾, Pino Cutrone. 2012. 日本語版 Ten Item Personality Inventory(TIPI-J)作成の試み. パーソナリティ研究. 21(1), pp. 40-52.

非特定テレビ視聴履歴データに基づく テレビ視聴行動クラスタリングと視聴傾向変化分析

由田 翔吾^{†1}

奈良先端科学技術大学院大学

松田 裕貴^{†2}

岡山大学

松田 裕貴^{†3}

読賣テレビ放送株式会社

横田 哲弥^{†4}

読賣テレビ放送株式会社

榊原 太一^{†5}

読賣テレビ放送株式会社

安本慶一^{†6}

奈良先端科学技術大学院大学

1. はじめに

近年、スマートフォンやインターネットの普及、さらにはインターネット接続テレビ (Connected TV) の台頭により、テレビ視聴のスタイルは従来の固定的なものから大きく変化している。視聴者の行動は、単に番組を「見る」という行為にとどまらず、生活リズム、家庭環境、さらにはスポーツ国際大会や特別番組といった外部要因によっても影響を受ける。これに伴い、テレビ局や広告主にとっては、視聴者の行動パターンを詳細に把握し、それに基づいた最適な番組編成や広告配信戦略を策定することが求められている。

本研究では、2022年10月から2023年3月までの6か月間にわたり、在阪5局の非特定テレビ視聴履歴データ (各月約5000人分) を対象に、視聴者の視聴行動を詳細に分析することを目的とする。特に本研究で扱う「非特定テレビ視聴履歴データ」とは、視聴者の個人情報を含まず、視聴者の許諾なしにオプトアウト方式で収集されるデータを指す。このデータは、視聴者がテレビをインターネットに接続することで取得され、視聴者はデータ放送画面から提供を拒否することが可能である。収集される情報には視聴開始・終了時刻、テレビID、視聴チャンネルが含まれ、これらは個人情報や視聴者属性を取得しない。

これらの形式で収集されたデータは、一般社団法人放送セキュリティセンター (SARC) が策定した「オプトアウト方式で取得する非特定視聴履歴の取扱いに関するプラクティス」 [1] に基づき、放送局以外への第三者への提供の禁止、データの保存期間の制限などが設けられ、一定の期間が経過したのちに削除される。オプトアウト方式のデータ収集方法の利点は、多くのデータを能動的な視聴者の負担な

く収集できる点が挙げられる。しかし、オプトイン方式と異なり、本手法では視聴者の個人属性や趣向の把握が難しいという課題がある。そこで本研究では、非特定テレビ視聴履歴データを用いて、視聴者群全体の傾向を分析し、放送コンテンツの改善や新たなサービス開発に資するアプローチを検討する。

さらに現在、在阪の5つの民間放送局では、非特定テレビ視聴履歴データの新たな価値創出に向けて、放送局間でのデータ連携技術の検証とデータ活用に向けた共同技術実験 [2] を実施している。本研究では、これらの実験で収集されたデータを基に特徴量を作成し、K-means クラスタリングを適用することで、視聴者の視聴傾向を分類する手法を提案する。具体的には、視聴時間に基づくメインクラスタリングを実施し、視聴パターンに応じて視聴者を分類した上で、番組カテゴリ別の視聴割合を考慮したサブクラスタリングを行う二段階アプローチを採用した。

さらに、視聴者のクラスタ遷移に着目し、月ごとのクラスタ移行確率を分析することで、視聴行動の安定性と変動性を評価した。この分析は、テレビ局にとってのターゲティング戦略の策定や、視聴者の離脱防止策の設計に貢献することが期待される。

本研究の成果は、テレビ局の番組編成や広告配信の最適化に資するだけでなく、視聴行動の変容をより深く理解するための新たなアプローチを提供するものである。

2. 関連研究

テレビ視聴履歴データの活用は、視聴者行動の詳細な分析を可能にし、放送業界や広告業界において重要な研究分野となっている。既存の研究では、「ネットワーク対応テレビを活用した大規模視聴データの分析」、そして「地上波放送における非特定視聴履歴データを用いた研究」という二つのアプローチが存在する。ここでは各研究について述べたうえで、本研究の立ち位置を整理する。

^{†1} DAIKI MAYUMI, Nara Institute of Science and Technology

^{†2} YUKI MATSUDA, Okayama University

^{†3} HIROKI MATSUDA, Yomiuri Telecasting Corporation

^{†4} TETSUYA YOKOTA, Yomiuri Telecasting Corporation

^{†5} TAICHI SAKAKIBARA, Yomiuri Telecasting Corporation

^{†6} KEIICHI YASUMOTO, Nara Institute of Science and Technology

2.1. ネットワーク対応テレビによる視聴履歴データ分析

近年、インターネット接続機能を持つネットワーク対応テレビ (Connected TV) を活用した視聴履歴データの分析が進められている。特に、東芝製テレビの視聴データを用いた研究は、地上波放送のみならず、録画視聴 (タイムシフト視聴) や VOD (ビデオ・オン・デマンド) を含む包括的な視聴行動の分析を可能にする。

菊池ら [3] は、東芝製ネットワーク対応テレビから収集された全国 24 万台分の大規模視聴データを用い、リアルタイム視聴とタイムシフト視聴の実態を明らかにした。彼らは、2012 年から約 4 年間にわたる視聴データを蓄積・分析し、ドラマ全話を通じた視聴動向や長期間の視聴行動の変化を調査するための集計システムを開発した。特に、広告主やコンテンツ制作者向けに視聴傾向の集計・分析結果を提供し、ビジネス判断の材料として活用することを目的とした。

また、すべての視聴者からアンケート回答を得ることが困難であるため、一部ユーザーのアンケート回答を基に機械学習手法を適用し、世帯構成員の性別・年代属性を推定する手法を提案した。実データを用いた実験では、一部属性を除き、F 値 0.7 以上の精度で性年代属性を推定可能であることが示された。

さらに、水岡ら [4] は、同様の東芝製テレビのデータを用いて、視聴パターンの遷移を自動抽出する手法を提案し、長期間にわたる視聴行動の変化を分析した。この研究では、特定の番組ジャンルや曜日ごとの視聴動向の変化が、視聴パターンの推移に与える影響を調査した。

これらの研究の課題として、以下の 2 点が挙げられる。

1. リアルタイム視聴のみの分析ではない

菊池らの研究は、録画視聴や VOD を含んだ分析であり、地上波放送に特化したリアルタイムの視聴行動の解析には適していない。

2. 個人レベルの視聴行動が主な対象

世帯単位での属性推定や行動予測が主眼であり、地上波視聴者全体のクラスタリング分析や、その時間的な遷移を捉える研究ではない。

2.2. 筆者らの研究グループによる地上波放送における視聴履歴データ分析

筆者らの研究グループでは、在阪 5 局 (讀賣テレビ、毎日放送、朝日放送、関西テレビ、テレビ大阪) が収集する地上波放送の非特定視聴履歴データを活用し、視聴行動の詳細な分析を進めている。このデータは、ネットワーク対応テレビやスマートテレビから得られる視聴履歴を元にしており、個人を特定しない形で収集されるため、より広範な視聴者層の分析が可能である。

筆者らの研究では、これまでに次の成果を発表している。まず、松田らは、非特定視聴履歴データの放送局を横断する統合手法を提案 [5, 6] しており、得られたデータを用いて視聴行動とインターネット検索行動の関連性分析 [7] することで、テレビ CM の視聴履歴データと Google Trends の検索トレンドを統合し、CM の影響が視聴者のオンライン行動に与える影響を調査している。吉村ら [8] は、視聴行動の地域差の分析を行った。具体的には、地域別の視聴履歴データを分析し、CM の完視聴率における地域差を明らかにした。真弓ら [9] は、非特定テレビ視聴履歴データを用いたクラスタリングにより、視聴者のテレビ視聴行動を分析する手法を提案している。

2.3. 本研究の立ち位置

本研究は、これまでの先行研究を発展させ、地上波放送の視聴者行動をより詳細に分類し、視聴傾向の動的変化を捉えるための新たなクラスタリング手法と遷移分析を提案する。特に、リアルタイム視聴データを活用し、視聴者の行動パターンを多角的に分析する点が特徴である。地上波リアルタイム視聴データに特化したクラスタリングと遷移分析を統合的に実施することにより、既存の視聴履歴分析では十分に捉えられていなかった、視聴行動の動的変化を精緻に評価し、テレビ局の番組編成や広告配信の最適化に資する新たな知見を提供する。

3. 提案手法

3.1. 特徴量抽出及びクラスタリングの実施

本研究では、2022 年 10 月から 2023 年 3 月までの 6 か月間にわたり、非特定テレビ視聴履歴データを収集し、視聴行動の詳細な分析を行った。対象視聴者は各月約 5000 人であり、視聴時間の分布や番組カテゴリの視聴傾向を明らかにするために、視聴行動に関する複数の特徴量を抽出した。これらの特徴量を用い、視聴者のクラスタリングを二段階で実施する。

まず、視聴行動の基本的な特性を捉えるために、表 1 に示す様に総視聴時間、平均視聴時間、最大視聴時間、視聴回数 (セッション数)、視聴時間のばらつき (分散・標準偏差) などの統計量を算出した。さらに、テレビの利用時間帯に関する情報として、プライムタイム (19:00~22:00)、ゴールデンタイム (19:00~23:00)、ノンプライム時間帯 (プライムタイム以外) における視聴時間の合計とその割合を求めた。これにより、特定の時間帯に集中して視聴する視聴者と、時間帯を問わず視聴する視聴者の違いを把握することができる。

加えて、視聴傾向の細分化を目的として、各視聴者の番

表 1 抽出した特徴量の一覧

特徴量	説明
基本統計量	
総視聴時間	全視聴時間の合計 (分)
平均視聴時間	1 回の視聴あたりの平均視聴時間 (分)
最大視聴時間	1 回の視聴における最長視聴時間 (分)
セッション数	視聴の回数 (テレビをつけた回数)
視聴時間分散	視聴時間のばらつき
視聴時間標準偏差	視聴時間のばらつきの尺度
時間帯別視聴指標	
プライムタイム視聴時間	19:00~22:00 の時間帯の視聴時間の合計
プライムタイム視聴割合	プライムタイムの視聴時間が全体の視聴時間に占める割合
ゴールデンタイム視聴時間	19:00~23:00 の時間帯の視聴時間の合計
ゴールデンタイム視聴割合	ゴールデンタイムの視聴時間が全体の視聴時間に占める割合
ノンプライム視聴時間	プライムタイム (19:00~22:00) 以外の視聴時間の合計
ノンプライム視聴割合	ノンプライム視聴時間が全体の視聴時間に占める割合

組カテゴリごとの視聴割合を特徴量として算出した。番組カテゴリには、図 1 に示す様にアニメ・特撮、スポーツ、ドキュメンタリー、ドラマ、ニュース・報道、バラエティー、情報・ワイドショー、映画、演劇・公演、趣味・教育、音楽が含まれる。これにより、特定のジャンルに強い関心を持つ視聴者と、幅広いジャンルを視聴する視聴者を識別することができる。

3.2. 視聴傾向の分析

本研究では、視聴者の行動をより詳細に把握するために、視聴時間を基準としたメインクラスターリングと、番組カテゴリの嗜好に基づくサブクラスターリングの二段階クラスターリングを行った。

まず、メインクラスターリングでは、K-means 法を用いて視聴時間に関連する特徴量を基に視聴者を 8 のクラスタに分類した。これにより、長時間視聴者、特定の時間帯に視聴が集中する視聴者などのグループに分類しラベル付けを行った。

次に、各メインクラスター内で、番組カテゴリの視聴割合を特徴量として K-means 法によるクラスターリングを実施し、10 のサブクラスタに分類した。

この手法により、同じ視聴時間パターンを持つ視聴者間において、番組カテゴリ別の嗜好の違いを抽出することが可能となり、より細分化された視聴者のターゲティングが可能となる。



図 1 番組カテゴリの一覧と対応色

3.3. クラスター遷移の分析

視聴者のクラスターが時間の経過とともにどのように変化するかを分析するため、クラスター遷移の追跡を行った。まず、10 月のクラスターリング結果を基準とし、その後の各月 (11 月~3 月) に対して視聴者ごとのクラスター再分類を行い、クラスター間の遷移パターンを集計した。

具体的には、各視聴者のクラスターの変遷を記録し、月ごとに前月からのクラスター移行をカウントした後、遷移確率を算出した。これにより、各クラスターにおける自己遷移率 (同じクラスターに留まる確率) や、他のクラスターへ移行する割合を求めた。

その後、得られた遷移確率を可視化するために、ヒートマップを生成した。このヒートマップでは、縦軸を前月のクラスター、横軸を翌月のクラスターとし、それぞれのマスに該当する遷移確率を表示している。色の濃淡によって遷移確

率の高低を視覚的に把握できるようにし、視聴者の視聴習慣の変化を直感的に理解できるようにした。

この分析を通じて、視聴者がどの程度クラスタに定着するのか、またどのクラスタ間で移行が多いのかを明らかにし、視聴行動の変化を把握することを目的とした。

4. 結果と考察

4.1. クラスタリング結果と分析

クラスタリングを行った結果と各クラスタの特徴について以下に示す。

1. **【ヘビー層：朝・プライムタイム集中型】** このクラスタの視聴者は、朝の時間帯（6時～9時）と夜のプライムタイム（19時～22時）に長時間視聴する傾向がある。特に朝の視聴時間が長く、朝の情報番組やニュース番組を視聴する習慣が強いと考えられる。深夜の視聴も活発であり、プライムタイムのエンタメ番組やニュースを視聴する可能性が高い。生活リズムが整っており日中の視聴が少ないことから社会人や学生などの層が含まれると予想される。
2. **【ミドル層：朝昼-夜視聴継続型】** このクラスタは、朝と夜に視聴する傾向があるが、クラスタ 1 に比べると視聴時間が短い。昼以降は増減せずテレビを付けっぱなしにしていると予想される。昼以降の視聴数が一定なこと、朝にテレビを見るのみで昼から夜の視聴が伸びていないことから、朝にテレビをよく見る層であると考えられる。
3. **【ヘビー層：終日視聴】** 朝から深夜まで連続してテレビをつける層で、視聴習慣が極めて強固。チャンネル固定傾向が高く、視聴する番組が大きく変動しにくい。このクラスタは、テレビを主な娯楽手段としている可能性が高く、在宅時間が長い視聴者層が含まれると予想される。
4. **【ライト層：朝昼-夜視聴継続型】** このクラスタは、朝と夜に視聴する傾向があるが、クラスタ 2 に比べると朝の視聴時間が短く、昼以降は増減せずテレビを付けっぱなしにしていると予想され、クラスタ 2 と同様に昼以降の視聴数が一定なこと、朝にテレビを見るのみで昼から夜の視聴が伸びないことから朝にテレビをよく見る層と考えられる。
5. **【ライト層：ゴールデンタイム視聴型】** 日中はほとんどテレビを視聴せず、主に夜のゴールデンタイムに視聴時間が集中しており、朝と日中の視聴がほぼない層であることがわかる。
6. **【ライト層（朝・ゴールデンタイム視聴型）】** 日中の視聴がほとんどなく、朝とゴールデンタイムに集中して

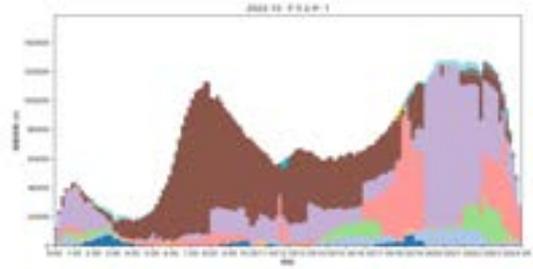


図2 クラスタ 1: ヘビー層：朝・プライムタイム集中型

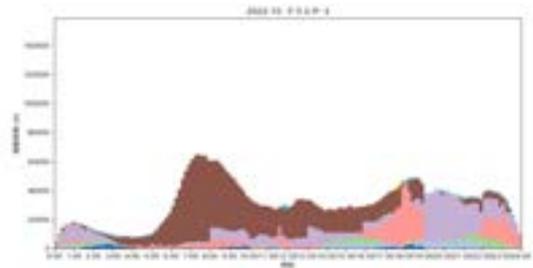


図3 クラスタ 2: ミドル層：朝昼-夜視聴継続型

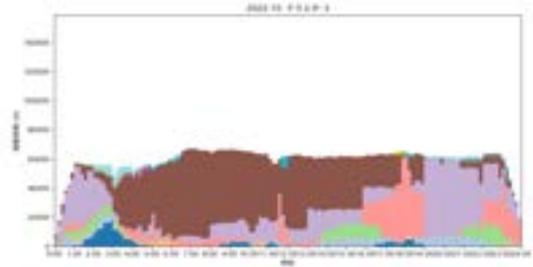


図4 クラスタ 3: ヘビー層（終日視聴）

視聴しており、視聴時間も短いライト層と考えられる。日中の視聴時間が低いことから出勤前などにテレビを軽く見ている層などだと考えられる。

7. **【ミドル層：朝・ゴールデンタイム視聴型】** クラスタ 6 に視聴傾向としては近いが、全体的な視聴時間が増加しており、特にゴールデンタイムの視聴時間が朝と比較して多いことから、仕事前などにテレビを視聴し、帰宅後ゴールデンタイムなどにテレビを見るのが習慣化されている層と考えられる。
8. **【ヘビー層：朝・ゴールデン型】** 朝と夜のゴールデンタイムに視聴が集中し、昼間の視聴はほとんどない。特に朝のニュースや情報番組、夜のバラエティ・ドラマを好んで視聴する傾向が強い。

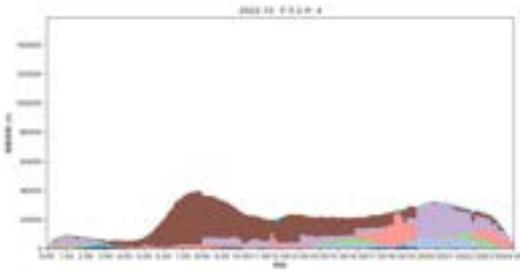


図5 クラスタ4: ライト層: 朝昼-夜視聴継続型

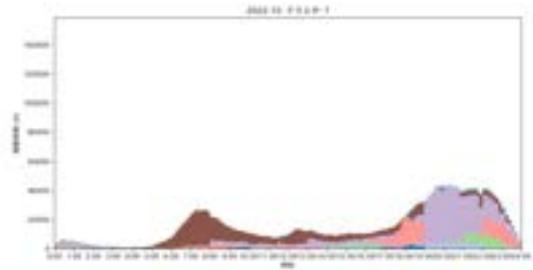


図8 クラスタ7: ミドル層: 朝・ゴールデンタイム視聴型

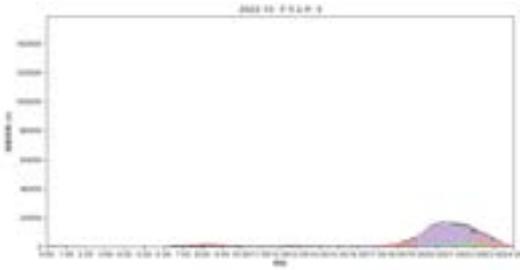


図6 クラスタ5: ライト層 (ゴールデンタイム限定)

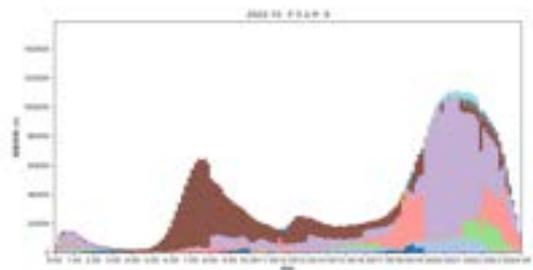


図9 クラスタ8: ヘビー層 (朝・ゴールデンタイム視聴型)

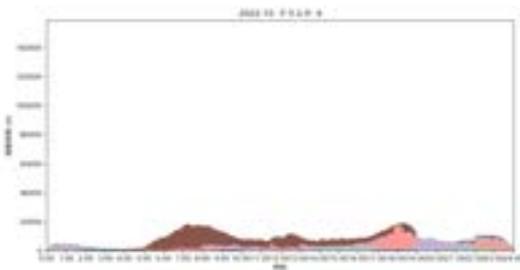


図7 クラスタ6: ライト層 (朝・ゴールデンタイム視聴型)

1. **【アニメ視聴層】** 深夜 (0:00~3:00) および早朝 (5:00~7:00) の視聴時間が長く、日中の視聴は少ない。アニメ/特撮 (濃い青) の視聴が顕著である
2. **【映画視聴層】** ゴールデンタイムにおいて、バラエティなどの視聴に加えて、映画 (灰色) の視聴時間が長く、習慣的に映画を見ている層だと予想される。
3. **【音楽視聴層】** ゴールデンタイムにおいて、バラエティなどの視聴に加えて、音楽 (水色) の視聴が目立ち、バラエティなどに加えて音楽番組が好きな層だと予想される

4.2. サブクラスタリング結果と分析

本節では、メインクラスタリングの各クラスタに対して行ったサブクラスタリングの結果を詳しく分析する。サブクラスタリングでは、より細かい視聴行動のパターンを抽出するために K-means 法を用い、視聴カテゴリごとの時間帯別傾向を明確にした。これにより、視聴者の特性をより深く理解し、番組編成や広告戦略の策定に活用することが可能となる。メインクラスタリングでは視聴時間や視聴時間帯に基づいた特徴量を用いてクラスタリングを実施したが、サブクラスタリングでは好みの番組カテゴリ別にクラスタリングすることを目的としているため、各番組カテゴリ別の視聴割合を特徴量として抽出した。

サブクラスタリングの利点は、メインクラスタリングでは捉えきれなかった詳細な視聴傾向を明確にできる点にある。例として以下のような詳細な視聴傾向を確認可能な点が挙げられる、今回はクラスタ1の結果に対してサブクラスタリングを行った結果を示す。

これらのサブクラスタリング結果を基に、各視聴層の特徴をより詳細に把握し、視聴者の行動パターンを特定することで、ターゲットに適した番組編成や広告戦略の策定が可能となる。今後の施策として、各視聴層の変遷を時系列で分析し、どの時間帯にどのような視聴習慣の変化が生じるかを詳細に追跡することが求められる。

4.3. クラスタ間の遷移分析と考察

本研究では、視聴行動に基づくクラスタリング結果に加え、月ごとのクラスタ間の遷移確率を解析し、各視聴層の安定性や動的变化を定量的に評価した。特に、安定した視聴行動を示すクラスタと、他クラスタへの移行が頻繁に発生するクラスタに着目し、視聴傾向の変遷がどのように発生するかを明らかにする。

図13にクラスタ間の全体遷移確率を示すヒートマップを

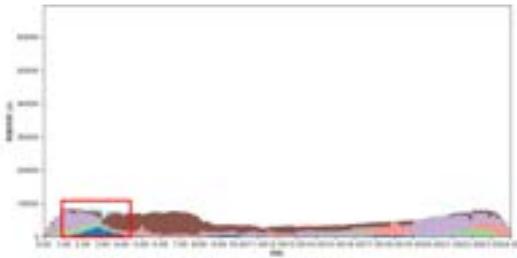


図 10 アニメ視聴層の可視化

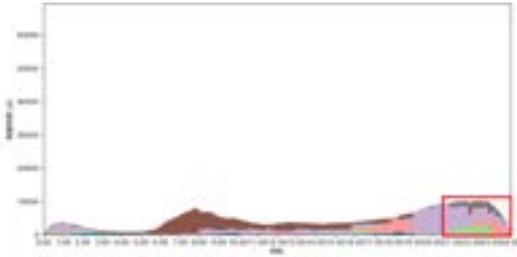


図 11 映画視聴層の可視化

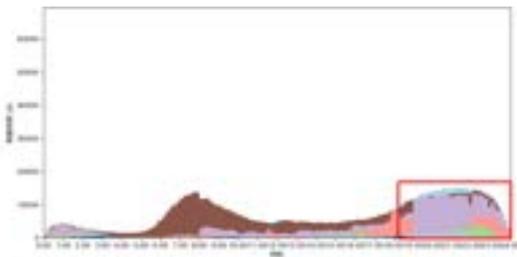


図 12 音楽視聴層の可視化

示す。この行列は、前月においてクラスタ i に属していた視聴者が、翌月にクラスタ j に遷移する確率を表しており、対角成分が大きいかほど視聴行動が安定していることを示す。対角成分が小さく、行内の他の要素が大きいクラスタは、視聴行動の変動が大きいことを示している。

4.3.1 視聴行動の安定性と変動性

図 13 に示すクラスタ別遷移確率ヒートマップをもとに、視聴行動の安定性を分析した。クラスタごとの自己遷移確率を比較したところ、最も視聴行動が安定していたのは「ヘビー層：終日視聴 (クラスタ 3)」であり、自己遷移確率は 82.99% であった。このクラスタの視聴者は、一日を通して長時間テレビを視聴する習慣が確立されており、他のクラスタへ移行する割合が極めて低い。視聴時間が終日にわたるため、視聴行動の変化が生じにくいと考えられる。

次いで安定していたのは「ヘビー層：朝・プライムタイム集中型 (クラスタ 1)」であり、自己遷移確率は 75.15% であった。このクラスタの視聴者は、朝のニュースや情報番組の視聴を習慣化しており、夜のプライムタイムにも継



図 13 クラスタ間の全体遷移確率ヒートマップ

続的に視聴する傾向が見られる。特に、朝の視聴習慣が固定されているため、他のクラスタへ移行しにくいことが特徴である。朝の視聴時間が確保されている点を考慮すると、通勤・通学前の情報収集を目的とした視聴層が多く含まれている可能性が高い。

朝の視聴時間が一定である「ライト層：朝昼-夜視聴継続型 (クラスタ 4)」では、視聴時間の減少に伴い「ミドル層：朝昼-夜視聴継続型 (クラスタ 2)」へと移行するケースが 20.32% 確認された。朝の情報番組を視聴する層のなかでも、昼以降の視聴行動が一定ではない場合、別のクラスタへ移行する傾向が強いことが示唆される。

4.3.2 クラスタ間の遷移傾向と相性

クラスタ間の遷移確率を分析した結果、視聴時間帯の変化によって強い関連性を持つクラスタのグループが存在することが明らかになった。特に、朝の視聴行動の変化によるクラスタ遷移、日中の視聴習慣の変化に伴う移行、ゴールデンタイムの視聴増減による移動が顕著であった。

朝の視聴習慣によるクラスタ遷移の傾向として、「ミドル層：朝・ゴールデンタイム視聴型 (クラスタ 7)」の視聴者は 12.78% の確率で「ミドル層：朝昼-夜視聴継続型 (クラスタ 2)」へと移行していた。

日中の視聴習慣の変化によるクラスタ遷移の傾向として、「ライト層：ゴールデンタイム視聴型 (クラスタ 5)」の視聴者は 21.59% の確率で「ミドル層：朝・ゴールデンタイム視聴型 (クラスタ 7)」へ移行していた。

4.3.3 テレビ視聴データの活用可能性

本研究のクラスタ遷移分析の結果は、テレビ局の番組編成や広告戦略の最適化に活用できる可能性がある。特に、視聴習慣が安定しているクラスタと流動的なクラスタの特性を把握することで、視聴者の維持と誘導を目的とした施策の立案が可能となる。

「ヘビー層：終日視聴（クラスタ 3）」や「ヘビー層：朝・プライムタイム集中型（クラスタ 1）」のような安定したクラスタに対しては、長期的な視聴者維持を目的とした番組戦略が有効である。これらの視聴者は、特定の時間帯に一貫した視聴習慣を持っているため、同じ時間帯に継続的なコンテンツを提供することで視聴を維持しやすい。

一方で、「ミドル層：朝昼-夜視聴継続型（クラスタ 2）」や「ライト層：ゴールデンタイム視聴型（クラスタ 5）」などの流動性が高いクラスタに対しては、特定の時間帯への視聴誘導を促す施策が求められる。朝の視聴を増やすために情報番組のターゲットを拡張し、夜の視聴を安定させるためにゴールデンタイムのドラマやバラエティ番組の誘導を強化することが考えられる。

このように、クラスタ間の遷移分析を活用することで、視聴率向上や広告効果の最大化が期待できる。

5. 結論

本研究では、在阪 5 局から収集された非特定テレビ視聴履歴データを用い、視聴行動のクラスタリングとその遷移分析を通して、テレビ視聴パターンの動的変化を詳細に捉える新たなアプローチを提案した。提案手法は、まず視聴時間などの基本統計量に基づくメインクラスタリングを行い、その後、各クラスタ内で番組カテゴリごとの嗜好を反映したサブクラスタリングを実施することで、視聴者の行動を多角的に分類することに成功した。

分析の結果、例えば「ヘビー層：終日視聴」や「ヘビー層：朝・プライムタイム集中型」など、視聴行動が明確に固定されたクラスタと、朝昼-夜視聴継続型やゴールデンタイム限定など、他クラスタへの移行が顕著な動的クラスタが存在することが明らかとなった。特に、終日視聴するクラスタにおいては自己遷移確率が高く、安定した視聴習慣が確認された一方、朝やゴールデンタイムに限定した視聴層では、視聴パターンの変動性が認められ、クラスタ間の遷移が活発であることが示された。

また、本手法は、個人情報を含まない非特定視聴履歴データを活用することで、プライバシー保護を維持しながら大規模データの分析が可能である点でも意義がある。これにより、テレビ局や広告主は、視聴者の特性に応じた最適な番組編成や広告配信戦略の策定に向けた実践的な知見を得ることが期待される。

参考文献

- [1] 一般財団法人放送セキュリティセンター視聴関連情報の取扱いに関する協議会：オプトアウト方式で取得する非特定視聴履歴の取扱いに関するプラクティス（ver2.1），https://www.sarc.or.jp/documents/www/NEWS/hogo/2021/optout_practice_ver2.1.pdf（2021）. Accessed on Feb. 28, 2025.
- [2] 読売テレビ放送株式会社：「テレビ視聴データ連携に関する共同技術検証実験（2022年度）」について，<https://www.ytv.co.jp/privacy/experiments2022/index.html>（2022）. Accessed on Feb. 28, 2025.
- [3] 菊池匡晃，坪井創吾，中田康太ほか：大規模テレビ視聴データによる番組視聴分析，デジタルプラクティス，Vol. 7, No. 4, pp. 352-360（2016）.
- [4] 水岡良彰，中田康太，折原良平：大規模テレビ視聴データによる視聴パターン推移の分析，第 32 回人工知能学会全国大会論文集，pp. 1P203-1P203（2018）.
- [5] 松田裕貴，榊原太一，松田裕貴，水本旭洋，安本慶一：放送局を横断する大規模テレビ視聴履歴データの統合手法の提案と実践，情報処理学会論文誌デジタルプラクティス（TDP），Vol. 4, No. 1, pp. 34-44（2023）.
- [6] 松田裕貴，榊原太一，真弓大輝，松田裕貴，水本旭洋，安本慶一：非特定テレビ視聴履歴データの放送局間統合手法，情報処理学会論文誌，Vol. 65, No. 10, pp. 1488-1500（2024）.
- [7] 松田裕貴，榊原太一，木俣雄太，鳥羽望海，真弓大輝，松田裕貴，安本慶一：テレビ視聴における非特定視聴履歴データとインターネット検索データの関係性分析，第 14 回データ工学と情報マネジメントに関するフォーラム（DEIM'22），pp. 1-6（2022）.
- [8] 吉村 啓，水本旭洋，榊原太一，松田裕貴：テレビ視聴時の CM 離脱と地域傾向分析，人工知能と知識処理研究会，Vol. 121, pp. 43-48（2022）.
- [9] 真弓大輝，松田裕貴，松田裕貴，横田哲弥，榊原太一，安本慶一：非特定テレビ視聴履歴データを用いたテレビ視聴行動クラスタリング手法の構築と評価，AIoT 行動変容学会第 8 回研究会（BTI-8），pp. 23-30（2024）.

訓練中の動作・所産からの漢字書字正答率の予測

拡張現実場面での空書と映像解析からの検討

大森幹真^{†1}・切替このみ^{†2}

早稲田大学人間科学学術院^{†1}・早稲田大学大学院人間科学研究科^{†2}

1. はじめに

文部科学省[1]によると2019年までの「学年別配当漢字数」は、小学1年生—80文字、小学2年生—160文字、小学3年生—200文字、小学4年生—200文字、小学5年生—185文字、小学6年生—181文字、の合計1006文字であった。しかし、2020年からの新学習指導要領になることで、「学年別配当漢字」が20文字増加し、小学校配当の漢字の総数は1026文字になった。近年では通常学級の小中学校に通う児童・生徒の中に学習面・行動面で困難さがあるものが8.8%の割合で存在しており、読むまたは書くことに困難さを示す子どもは3.5%に及ぶ[2]。特にそのような困難さを示す子どもは小学校に集中しており、2016年施行の障害者差別解消法などを受け、学校現場でも子どもの障害特性に応じた支援、社会的障壁の除去が求められるようになってきている(鶴巻他, 2019[3])。さらには、具体的な場面や状況に応じて多様で個別性の高い指導・支援が必要な現状に、指導者や指導方法のありかたも変化が必要になってきている[3]。

一方で漢字の指導において、指導者側が漢字の学習困難に気づきにくいという問題もある。それは第一言語として漢字を学ぶ人たちが、漢字の学び方や教え方に感受性が鈍くなっていることが関係している可能性もある。例えばヴォロビヨワ ガリーナ[4]によると、外国語として漢字習得を行う際には、10点の問題点があることを示した。それらは①学習すべき漢字の数が多い ②漢字を非体系的に感じる ③漢字の字体が複雑である ④漢字を構成する要素の種類が多く、部首以外の要素は標準化されていない ⑤字体・字義・読み方の関連性が不明瞭 ⑥個々の漢字に関わる情報が多く、同時に字体・字義・読み方・筆順・部首を覚える必要がある ⑦漢字学習において、熟語を数多く覚える必要がある ⑧音声情報が単一ではなく、複雑であり、字体と読み方の対応が複雑 ⑨学習すべき漢字の出現順序が合理的でない ⑩漢字辞典の使用方法が困難 というものであった。これらの外的な指摘は非常に興味深く、特に初学者や学習困難児が書くことを学習する際には②や③の部分は特に配慮する必要があるだろう。

しかし教育現場においては、漢字書字指導方法に多様性があるとは言い難く、指導法にバリエーションが乏しいことで、子どもたちの学習機会を最大化することが難しい場合もある。それにより、日本国内の小学校1~4年生の中に

存在する約10%の書き困難児への支援が最適化されていない可能性もある[5,6]。日本の通常学級において漢字書字を促す方法は2つに大別できる。一つは、紙面上に何度も繰り返し書くことを求める「反復書字」であり、もう一つは空中に指で文字を書く「空書」である。特に、書きの支援ではいまだに、何度も繰り返し書くことを求める「反復書字」が多く行われている。稲垣・藤田[7]によると、1つ目の実験では大学生を対象にした場合、通常の見視での学習や空書よりも反復書字を行う方が漢字習得に効果的であること報告している。一方で、大西・熊谷[8]は小学校2年生から小学校5年生までの学習障害児4名に対し、なぞり学習条件や空書条件、および視写による漢字の反復書字学習条件の実験を行った。その結果、画数が多い漢字は子どもたちにとって正確に見ることが難しく、運動負荷も高いために正確な書き写しにつながらないことも示した。このように、学習困難児に対しては単純な反復書字では視覚的・運動的な負荷の高さから効果的な漢字学習につながらないことも多い。そのため、机上の反復書字に関しては、その実施の方法をもう一度見直す必要があるだろう。

机上での反復書字とは対照的に空書では空中で文字を書くことで、運動の負荷を軽減させつつも、文字形態を表現かつ理解しやすくする効果があると考えられている。さらには、机上の反復書字とは異なり、授業中での一斉指導に組み込むことが容易であり、書字所産を消す必要がないことも方法論上のメリットとして考えることが出来る。しかし、空書をするだけで効果的に書字獲得につながるかは成果が一樣でないことが多い[9][10]。例えば成人の外国人75名が第二外国語として漢字を学ぶ実験においては、空書行動が漢字の書字獲得に繋がったと報告されている[9]。その一方で、日本の特別支援級在籍の小学校3年生から6年生が参加した空書と指書きによる同時再生実験においては、漢字書字獲得の成績は良くないとの研究結果が出ている[10]。その一因として、空書時の書字所産を書いた本人が視認することが出来ないことが考えられる。例えば稲垣・藤田[7]は2つ目の実験において、空書学習条件に書字結果の確認過程を含めない条件よりも、書字結果の確認過程を含めた条件において正答率が有意に大きくなったことを明らかにした。しかし、同研究では自身が書いた書字所産を観察したわけではなく、空書後に印刷された標的漢字を確認することで、書字結果の確認を行っていた。そのため、空書時に自分の書

^{†1} MIKIMASA OMORI, Faculty of Human Sciences, Waseda University

^{†2} KONOMI KIRIKAE, Graduate School of Human Sciences, Waseda University

字所産を確認することで漢字書字正答率が向上するかは未解明のままである。そこで、空書の書字所産を確認する方法として拡張現実(Augmented reality; AR)場面での空書指導を着想し、机上場面での反復書字と比較検討することとした。

そのうえで、机上・空書条件内においても書字所産を視覚的に確認できるか否か、つまり視覚情報を運動反応に統合させる能力である視覚運動協応についても検討することとした。大森 [11]は大学生が文章を視写している場面において、視覚運動協応得点が低いほど視覚的な負荷がかかり、それらが手元を見返す回数に反映していることを明らかにした。また大森・唐亀・豊浦[12]は、子どもたちの書字動作の不均一性から視覚運動協応能力や、発達障害傾向が予測できる可能性も示している。一方で、書字場面の映像解析からも書き困難を同定する試みもある[13]。Lopez and Vaivre-Douret [13]は、書き困難児の方が定型発達児よりも腕部分の垂直方向の運動のパラつきが大きくなることを発見した。そのうえで、書き困難児は目を開けて書字を行う場合と閉じた場合とでは、閉眼時の方が身体動作に安定性が増すという報告もあり、視覚的フィードバックの困難さが視覚運動協応の困難さにつながり、書き困難が顕在化するという考えもある[13]。つまり視覚と運動の関連性を客観的に評価することで書き困難児の早期発見につなげることが可能になるため、本研究では AR 場面での空書とそれらの映像解析を行うことで、教育場面内での漢字書字指導の再考および書き困難を発見するための方法論を検討する。

本研究の目的は定型成人(おとな群)と定型発達児(子ども群)を対象に 2 種類の反復書字条件 (拡張現実場面 [augmented reality; AR] vs 机上) と 2 種類の視覚条件 (可視化 vs 非可視化) を設定した。全参加者が 4 条件下での反復書字を行うことで、新規の漢字書字獲得につながるかを検討することとした。その上で、子ども群の AR 空書条件下での諸反応および、動作解析による身体の変動性が後の漢字書字正答率の予測につながるかも検討することとした。

2. 方法

2.1 参加者

5 歳から 9 歳の子ども 9 名と、比較対象のおとなとして 20 歳から 25 歳までの大学生 11 名が参加した。子どもの内訳は、就学前の女児 4 名、小学 1 年生男子 1 名、小学 2 年生女子 1 名、小学 3 年生男子 2 名・女子 1 名であった。就学前の子どもと 1 年生は、ひらがなが習得済みであった。子どもの IQ は 81 から 105 程度でスムーズに指示が通り、全員自分の名前を書くことができた。

2.2 刺激と装置

装置: AR 条件ではデジタルペンとして Pictionary Air (GKG38:長さは 17.5 cm, 重さは単三電池を入れて 69 g) を使用した。そして、書字所産を AR 上で反映するために iPad Pro を使用した。タブレットは上半身を写せるように参加者

から 1 メートルほど距離を取ってセッティングした。机上書字条件では、漢字練習用・テスト用紙としての A4 コピー用紙と、カーボン紙、クリアファイル、ストップウォッチを使用した。また、AR 書字場面は iPad Pro を通じて撮影した。そして書字場面中の運動機能を解析するために運動解析ソフト (VisionPose Single3D, ネクストシステム) を使用した。

刺激: 子どもとおとなの文字・漢字刺激は Table 1 に示した。各条件で学習を行った漢字数は 3 文字で、合計 12 漢字を使用した。対象年齢によって、学習セット内の総画数を変更し、難易度の調整を行った。

Table 1. AR 書字・机上書字に使用した文字・漢字刺激

条件	AR書字	机上書字	平均画数	
子ども	シ(し)	レ(れ)	人(ひと)	仏(ほとけ)
	ヤ(や)	イ(い)	川(かわ)	仕(し)
	チ(ち)	ヒ(ひ)	八(はち)	竹(たけ)
	ニ(に)	カ(ちから)	庫(こ)	公(こう)
	夕(ゆう)	小(こ)	界(かい)	今(いま)
	口(くち)	十(じゅう)	貝(かい)	兄(あに)
	相(そう)	畑(はたけ)	灯(ともしび)	石(いし)
	貝(いん)	宮(みや)	料(りょう)	才(さい)
	算(さん)	注(そそぐ)	徳(とく)	玄(げん)
	老(ろう)	印(いん)	古(ふるい)	込(こみ)
	倉(くら)	案(あん)	主(あるじ)	捕(とる)
	察(さつ)	漁(りょう)	台(だい)	熊(くま)
おとな	佩(おびだま)	阿(ながしめ)	餛(あん)	糞(いらか)
	竈(かまど)	塞(かめ)	噎(くしゃみ)	黴(かび)
	禪(たすき)	鬢(びん)	續(かすり)	壺(しきみ)

2.3 手続き

事前評価: 12 個の文字・漢字刺激に対応するひらがなを提示し、対応する漢字を書くことを求めた。全漢字が事前評価時に正答でないことを確認し 4 条件下での訓練を行った。4 条件の実施順序は参加者ごとにランダムにして行った。**訓練条件:** Fig 1 に 4 つの書字訓練条件の様子を示した。すべての訓練では、1 分間ごとに各漢字を出来るだけ多く正確に書くことを教示し、3 つの漢字での訓練が終了した後に、事後テストを行った。事後テストの終了後に、次の条件での訓練および事後テストを実施し、4 条件が終了するまで繰り返した。① AR 空書不可視化条件では、PC 上で訓練漢字を提示し、その漢字を見ながらデジタルペンで空書することを求めた。その際に、タブレットの前面カメラが参加者に見えないように提示し、AR 上で書いた文字が自分に見えないように提示して反復書字を行った。② AR 空書可視化条件では、PC 上で訓練漢字を提示し、その漢字を見ながらデジタルペンで空書することを求めた。その際に、タブレットの前面カメラを参加者側に向けて提示し、AR 上で書いた文字が自分に見えるように提示して反復書字を行った。③ 机上不可視化条件: PC 上で訓練漢字を提示し、着席してその漢字を見ながら鉛筆でカーボン紙上に繰り返し書くことを求めた。カーボン紙を使用することで、参加者が書いた書字所産が見えないようにした。



Figure 1.4 種類の訓練条件の様子

④机上可視化条件：PC上で訓練漢字を提示し、着席してその漢字を見ながら鉛筆で白紙上に繰り返し書くことを求めた。その際に、実際に参加者が書いた書字所産が見ることができるようにした。

事後評価：各訓練の終了後に事前評価と同様のものを実施した。

2.4 結果の処理法

行動評価として、漢字書字正答率や訓練場面での文字の判読性、訓練中の1分間あたりの反応数等の従属変数に対して、参加者要因(子ども vs おとな)、反復書字要因(AR vs 机上)、視覚要因(可視 vs 不可視)において、3要因の混合計画で分散分析を行った。漢字書字正答率は、事前テストではすべて不正解・未学習のものであったため事後テストの正答率のみを解析した。文字の判読性については、書き順は考えず書字所産として成立しているか否かを基準に0~3点の間で、複数人で評価し、算出した。1分あたりの反応数は、各参加者で訓練文字・漢字の画数が異なるため、1画を1反応として、1分間の平均値を算出した。運動機能評価として、子ども群のみに身体の変動性を変動係数により算出した。部位としては、書き手と非書き手の手首およびひじ、そして頭部におけるxyz軸方向の動きのパラツキの座標における平均値と標準偏差の商を求めた。これらの変動係数においても、書き手要因(書き手 vs 非書き手)、視覚要因(可視 vs 不可視)、方向要因(水平 vs 垂直 vs 奥行)において、3要因分散分析を行った。また頭部と首の変動係数に対しては視覚要因(可視 vs 不可視)と方向要因(水平 vs 垂直 vs 奥行)において2要因分散分析を行った。

次に子ども群におけるAR書字場面において運動機能の変動係数を基にした訓練中の動作・所産からの漢字書字正答率の予測が出来るかを評価するために、相関分析を行った後に、重回帰分析を実施した。

3. 結果

行動評価：図2に漢字書字正答率、1分あたりの総画数、判読性得点をそれぞれ示した。漢字書字正答率に事前テストでは全参加者の正答率が0%であった。事後テストにおいて、おとな群ではAR空書不可視化条件において74.75%(SE=.07)、AR空書可視化条件では55.56%(SE=.08)、机上不可視化条件において84.85%(SE=.07)、机上可視化条件では83.84%(SE=.05)であった。子ども群ではそれぞれ、66.67%(SE=.06)、58.02%(SE=.07)、56.79%(SE=.12)、86.42%(SE=.05)であった。分散分析を行ったところ、反復書字要因と視覚要因間に有意な交互作用が見られた $[F(1,18)=4.91, p=.04, \eta_p^2=.21]$ 。単純主効果の検定の結果、可視化要因におけるAR書字での正答率の方が、机上書字での正答率よりも有意に低くなったことを示した $[F(1,18)=10.77, p=.004, \eta_p^2=.37]$ 。一方で、正答率においてグループ間での差は見られなかった。

次に1分あたりの総画数では、おとな群ではAR空書不可視化条件において75.58回(SE=5.46)、AR空書可視化条件では65.24回(SE=5.04)、机上不可視化条件において128.42回(SE=11.33)、机上可視化条件では126.70回(SE=12.49)であった。子ども群ではそれぞれ、56.63回(SE=6.20)、29.78回(SE=4.30)、56.37回(SE=11.48)、47.63回(SE=11.03)であった。分散分析を行ったところ、グループ要因と反復書字要因間に有意な交互作用が見られた $[F(1,18)=12.98, p=.002, \eta_p^2=.42]$ 。単純主効果の検定の結果、おとな群においてAR書字での総画数の方が、机上書字での総画数よりも有意に少なくなったことを示した $[F(1,18)=33.102, p=.000, \eta_p^2=.65]$ 。さらには、おとな群の方が子ども群よりもAR書字での総画数 $[F(1,18)=19.29, p=.000, \eta_p^2=.52]$ も机上書字での総画数 $[F(1,18)=28.46, p=.000, \eta_p^2=.61]$ も多くなることを示した。一方で子ども群では反復書字条件間では可視化条

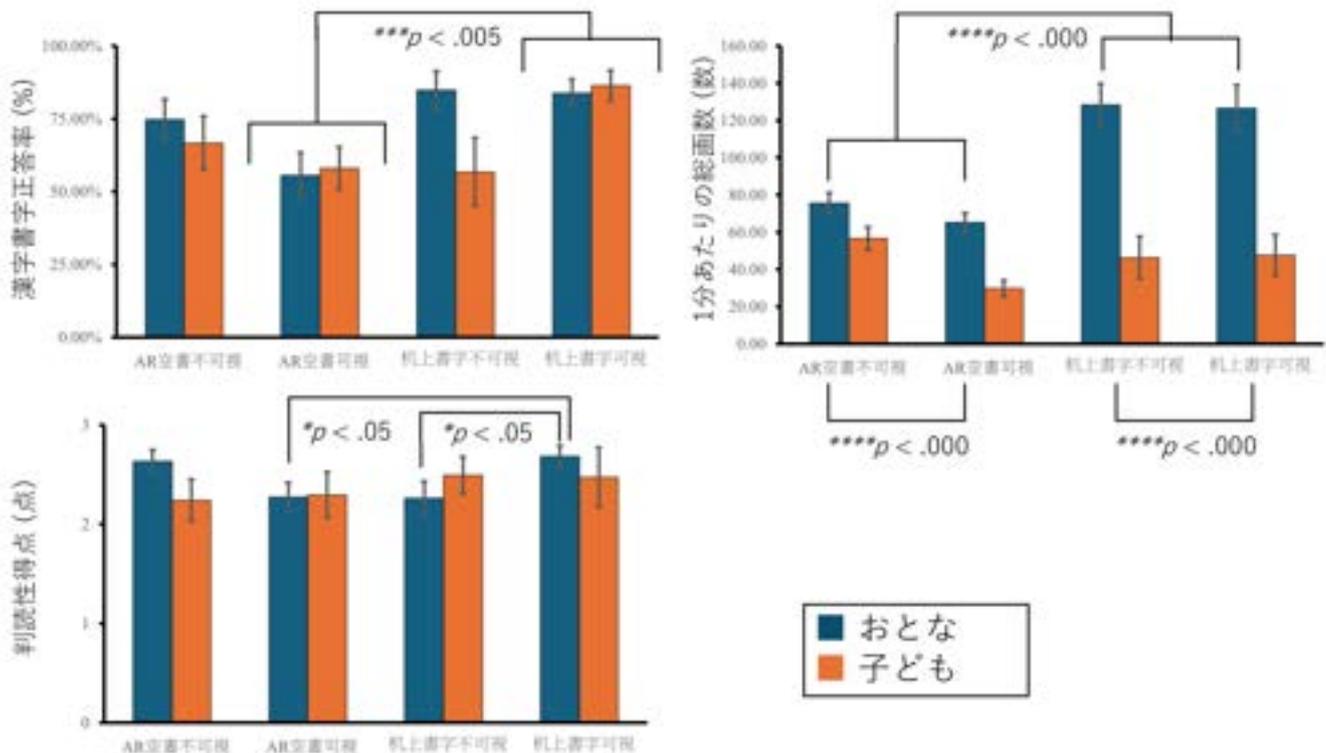


Figure 2. グループごとの行動評価の結果

件でも不可視化条件でも1分あたりの総画数に差がないことも明らかになった。

また訓練書字の判読性得点においては、おとな群のAR空書不可視化条件では2.63点(SE=.12)、AR空書可視化条件では2.27点(SE=.15)、机上不可視化条件において2.26点(SE=.17)、机上可視化条件では2.68点(SE=.11)であった。

子ども群ではそれぞれ、2.24点(SE=.21)、2.30点(SE=.23)、2.49点(SE=.19)、2.47点(SE=.30)であった。分散分析を行ったところ、3要因間に有意な交互作用傾向が見られた[F(1,18)=3.60, $p = .007$, $\eta_p^2 = .17$]。単純・単純主効果の検定の結果、おとな群の可視化条件においてAR書字での判読性得点の方が、机上書字での判読性得点よりも有意に低くなったことを示した[F(1,18)=5.20, $p = .04$, $\eta_p^2 = .22$]。さらには、おとな群での机上書字条件において可視化条件の方が、判読性得点が高くなったことを示した[F(1,18)=4.90, $p = .04$, $\eta_p^2 = .21$]。一方で子ども群ではいずれの条件でも判読性得点に差がないことも示した。

運動機能評価：Table 2にAR書字場面における書き手と非書き手の手首およびひじ、そして頭部におけるxyz軸方向の動きのバラつきとしての変動係数を示した。ひじの変動係数において3要因の分散分析を行ったところ、方向要因に有意な主効果が見られた[F(1,8)=3.82, $p = .04$, $\eta_p^2 = .32$]。多重比較を行ったところ、垂直方向の運動のバラつきが水平方向($p = .08$)および奥行方向($p = .09$)に比べて大きい傾向があることを示した。手首の変動係数に数においても同様に分析したところ、方向要因に有意な主効果が見られた[F(1,8)=5.56, $p = .04$, $\eta_p^2 = .41$]。多重比較を行ったところ、水平方向の運動のバラつきが奥行方向($p = .04$)に比べて大きくなったことを示した。

次に頭部の変動係数に対して2要因の分散分析を行ったところ、方向要因に有意な主効果が見られた[F(1,8)=14.34, $p = .006$, $\eta_p^2 = .63$]。多重比較を行ったところ、水平方向の運動のバラつきが垂直方向($p = .01$)と奥行方向($p = .02$)に比べ

て大きくなったことを示した。最後に首の変動係数に対しても分析を行ったところ、方向要因に有意な主効果が見られた[F(1,8)=10.42, $p = .001$, $\eta_p^2 = .57$]。多重比較を行ったところ、水平方向の運動のバラつきが垂直方向($p = .03$)と奥行方向($p = .04$)に比べて大きくなったことを示した。

相関分析：Table 3に子ども群におけるAR空書不可視化条件と、AR空書可視化条件書字での正答率と他の行動指標と変動係数における相関係数を示した。

Table 2. AR書字条件間での身体部位の変動係数

部位	視覚要因	方向要因	書き手	非書き手
肘	AR空書不可視化	水平方向	0.29 (0.13)	0.78 (1.64)
		垂直方向	2.09 (2.72)	2.79 (3.14)
		奥行方向	0.31 (0.14)	0.53 (0.62)
	AR空書可視化	水平方向	0.40 (0.29)	0.31 (0.10)
		垂直方向	0.80 (0.61)	4.98 (10.98)
		奥行方向	0.87 (1.25)	0.77 (0.42)
手首	AR空書不可視化	水平方向	0.72 (0.30)	2.41 (3.59)
		垂直方向	1.11 (1.18)	16.83 (42.62)
		奥行方向	0.27 (0.13)	1.83 (3.00)
	AR空書可視化	水平方向	2.64 (3.42)	0.66 (0.32)
		垂直方向	1.42 (2.81)	2.37 (3.34)
		奥行方向	0.30 (0.20)	0.61 (0.31)
肘	AR空書不可視化	水平方向	1.86 (2.12)	
		垂直方向	0.17 (0.14)	
		奥行方向	0.25 (0.12)	
	AR空書可視化	水平方向	1.47 (0.91)	
		垂直方向	0.15 (0.07)	
		奥行方向	0.56 (0.65)	
手首	AR空書不可視化	水平方向	2.27 (4.17)	
		垂直方向	0.28 (0.26)	
		奥行方向	0.27 (0.14)	
	AR空書可視化	水平方向	3.89 (4.11)	
		垂直方向	0.27 (0.21)	
		奥行方向	0.61 (0.72)	

相関分析の結果、AR 空書不可視化条件の正答率と変数間においては、非書き手のひじの垂直変動係数($r = -.50$)および、頭部の奥行変動係数($r = -.54$)において負の相関傾向が見られたのみであった。一方で、AR 空書可視化条件の正答率においては、書き手のひじ($r = -.63$)と手首($r = -.62$)の水平変動係数と有意な負の相関があり、非書き手のひじの奥行変動係数との間に有意な正の相関が見られた($r = .68$)。

Table 3. AR 書字条件での正答率における相関係数

	AR空書不可視化条件		AR空書可視化条件	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
行動指標				
総画数	0.22	0.28	0.71	0.02
判読性得点	0.20	0.30	0.13	0.36
ひじ				
水平方向	-0.21	0.29	-0.63	0.04
書き手				
垂直方向	-0.35	0.18	-0.22	0.29
奥行方向	-0.25	0.26	-0.40	0.14
ひじ				
水平方向	0.29	0.22	-0.44	0.12
非書き手				
垂直方向	-0.50	0.08	-0.14	0.36
奥行方向	-0.04	0.46	0.68	0.02
手首				
水平方向	0.02	0.48	-0.62	0.04
書き手				
垂直方向	-0.02	0.48	-0.39	0.15
奥行方向	-0.20	0.30	-0.38	0.15
手首				
水平方向	0.22	0.29	0.18	0.32
非書き手				
垂直方向	0.42	0.13	-0.41	0.14
奥行方向	-0.33	0.19	0.46	0.11
頭部				
水平方向	0.17	0.33	0.02	0.48
垂直方向	-0.17	0.33	-0.42	0.13
奥行方向	-0.10	0.40	-0.51	0.08
首				
水平方向	-0.54	0.07	0.13	0.37
垂直方向	-0.37	0.17	-0.42	0.13
奥行方向	-0.07	0.43	-0.39	0.15

空書での漢字書字正答率の予測：AR 上での空書訓練中の動作・所産からの漢字書字正答率の予測が出来るかを検討するために、視覚要因ごとに重回帰分析を実施した。まず、AR 空書不可視化条件においては、正答率を予測する重回帰式は算出することが出来なかった。次に AR 空書可視化条件で求めた重回帰式のモデルを Table 4 に示した。

Table 4. AR 空書可視化条件における重回帰モデル式

	モデル1	モデル2	モデル3
切片	0.214 (0.147)	0.020 (0.134)	0.078 (0.080)
総画数	0.012 (0.050)	0.016 (0.004)	0.016 (0.002)
非書き手の手首の 水平変動係数		0.034 (0.013)	0.063 (0.011)
非書き手のひじの 垂直変動係数			-0.044 (0.013)
決定係数 R^2	0.51	0.76	0.93
自由度調整済み決定係数 R^2	0.44	0.69	0.89
観測数	9	9	9

その結果、モデル3を採用し、AR 空書可視化条件での漢字書字正答率を予測する式は以下のものとして決定した。

$$[\text{モデル式}] \text{AR 空書可視化条件での漢字書字正答率} = 0.016 \times 1 \text{ 分あたりの総画数} + 0.063 \times \text{非書き手の手首の水平変動係数} - 0.044 \times \text{非書き手のひじの垂直変動係数} + 0.078$$

つまり、AR 上での空書においては、可視化された所産を見ながら早く書くことや、非書き手の手首やひじのバラつきから、訓練後の漢字書字正答率を予測できることを明らかにした。

4. 考察

本研究ではおとな群と子ども群を対象に4種類の書字訓練を行い、新規の漢字書字獲得につながるかを検討した。Fig 2 から、グループ間に有意差はなく、子ども群もおとな群も机上訓練の方がAR 空書での可視化条件に比べて正答率が高くなることを示した。一方で、教育場面で頻繁に活用されるAR 空書不可視化条件と机上書字可視化条件では同程度に漢字書字獲得につながることも明らかにした。稲垣・藤田[7]は、大学生を対象にした場合、通常の見視での学習や空書よりも反復書字を行う方が漢字習得に効果的であるが、空書学習条件に書字結果の確認過程を含めない条件よりも、書字結果の確認過程を含めた条件において正答率が有意に大きくなったことを明らかにした。本研究では子ども群においてもおとな群においてもAR 空書と机上書字の間に有意差が見られず、同等程度の学習効果があることも示した。一方で、空書条件においては書字所産の視認の有無で書字正答率に差が出なかったため、稲垣・藤田[7]の研究結果との相違も見られた。これらの要因として次の2点が考えられる。1点目は訓練中の総画数がAR 空書条件内で異なる点であった。両群ともにAR 空書不可視化条件の方が、AR 空書可視化条件よりも1分間の総画数が多く、見ることの負担が大きかった可能性がある。つまり大森[11]が報告した書き困難児と見返し回数の関連性が、机上場面だけでなく、空書時でも反映していた可能性がある。2点目はAR ペンの操作性や重量が考えられる。AR ペンで空書をする際には1画ごとにスイッチを押して空中に書く必要があった。しかし、切り替えや複数行動の両立の難しさや、ペンの反応性の問題もあり視覚化条件では書き直し反応をすることも多く見られた。そのために総画数が減少した可能性もある。また、AR ペンの重量も通常の鉛筆(10g程度)の約7倍で、69gの重さがあったことも、流暢な行動を維持することが困難であった一因であろう。

一方でTable 2 から運動機能についてはAR 空書場面においては、条件間で身体動作のバラつきは見られなかった。Lopez and Vaire-Douret [13]は、書き困難児は目を開けて書字を行う場合と閉じた場合とでは、閉眼時の方が身体動作に安定性が増すという報告もあったが、本研究の子どもたちは未診断の子どもたちであることもあり、視覚化要因による身体動作の変動性は見られなかった。しかし、Table 3 か

ら書字正答率との関連を検討すると、開眼時のひじや手首の変動性と負の相関があることも示した。つまり本研究の子ども群でも書き困難児に近い漢字書字正答率が低い子どもたちは、やはり開眼時に身体動作の不安定さが一因となり学習を阻害する可能性も示唆することとなった。今後の研究では大人や子ども内での正答率の高低で身体動作との関連を比較することで、視覚フィードバックの発達の变化を検討することも必要になってくると考える。

子ども群のAR空書条件下での諸反応および、動作解析による身体の変動性が後の漢字書字正答率の予測につながるかも検討することも本研究の目的であった。Table 4 から、特にAR空書可視化条件において、可視化された所産を見ながら早く書くことや、非書き手の手首やひじのバラつきから、訓練後の漢字書字正答率を予測できることを明らかにした。興味深いことに書き手の変動性ではなく、非書き手の動きの不均一性と因果関係が漢字書字正答率との関係を示した。Omori [14]によると視覚運動協応得点が低い子どもが図形描画を行った際に、高得点の子どもよりも非書き手のひじや手首の変動性が大きくなることも報告している。大森・唐亀・豊浦[12]は、書き手の変動性のみから、視覚運動協応能力を予測することを行っていたが、書き困難児を包括的に発見するためには、両側の変動性を分析する必要性が高いことも本研究から明らかとなった。

最後に小学校での書き困難の割合は10%程度と推測されているが、書きの支援方法についての効果検証はほとんど行われていない。国外に目を広げても書く行動への支援は代替行動の獲得やコンピュータを基盤にした支援が中心であり、書くことそのものの支援に関する方法論の開発は遅れている。本研究では机上および空書での反復書字が教育場面では一定の効果があることを示したが、正答率が80%程度ということもあり今後さらに方法論として洗練していく必要もあるであろう。また、書き困難や書字行動そのもの関しての評価方法は国内外でほとんど開発されていない。そのため、質問紙等で主観的に評価されている視覚運動協応や書字関連スキルを心理学と情報科学とのコラボレーションにより客観的に定量化する手法を開発していく必要がある。そうすることで書字困難児の早期発見や早期支援につなげることができつつ、国内外で適応可能である多角的な評価・支援システムを構築していくことが今後の課題となる。

引用文献

- [1] 文部科学省 (2017). 小学校学習指導要領(平成29年告示)解説, 外国語活動・外国語編, Retrieved from https://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/fieldfile/2019/03/18/1387017_011.pdf (最終閲覧日: 2025年2月19日)
- [2] 文部科学省 (2022). 通常の学級に在籍する特別な教育的支援を

必要とする児童生徒に関する調査結果について. https://www.mext.go.jp/content/20230524-mext-tokubetu01-000026255_01.pdf (最終閲覧日: 2025年2月19日)

- [3] 鶴巻正子・仁平義明・佐藤拓・高橋純一(2019). ADHD傾向の子どもが書く漢字に見られる特徴 人間発達文化学類論集, 29, 53-59.
- [4] ヴォロビョワ ガリーナ(2011). 構造分析とコード化に基づく漢字情報処理システムの開発 日本語教育, 149, 16-30.
- [5] 大庭重治 (2000). 通常の学級における低学年児童の書字学習状況とその支援課題 上越教育大学研究紀要, 29, 151-157.
- [6] 堂山亜希・橋本創一・林安紀子 (2014). 小学校通常学級における書字に関する困難がある児童の実態と支援:人力・出力・処理過程のつまずきに着目して 発達障害研究, 36, 369-379.
- [7] 稲垣紀夫・藤田正(2005). 漢字学習における書字行為に関する研究 教育実践総合センター研究紀要, 14, 47-51.
- [8] 大西正二・熊谷恵子(2019). 漢字書字の習得が困難な学習困難児に対する認知処理様式と体性感覚の入力方法に配慮した学習法の比較 LD研究, 28, 363-379.
- [9] Thomas, M. (2015). Air Writing as a Technique for the Acquisition of Sino-Japanese Characters by Second Language Learners. *L language Learning*, 65, 631-659.
- [10] 河村優詞(2019). 特別支学級に在籍児童における漢字学習方法の効果—5種類の学習方法の比較— 自閉症スペクトラム研究, 17, 15-22.
- [11] 大森幹真 (2019) 女子大学生・大学院生における書字運動と視線パターンの関連. 学苑, 940, 12-21.
- [12] 大森幹真・唐亀健大・豊浦正広 (2024). 書字動作の不均一性からの視覚運動協応能力の予測. 第8回 AIoT 行動変容学会研究会 (BTI) 論文集, 45-50.
- [13] Lopez, C., & Vaivre-Douret, L. (2023). Exploratory investigation of handwriting disorders in school-aged children from first to fifth grade. *Children*, 10, 1512.
- [14] Omori, M. (2024, June). Laterality of arm movement variability on copying and tracing. In *Proceedings of the Behavior Transformation by IoT International Workshop* (pp. 25-26).

本研究はJSPS 科研費 (No. 22K13739)の助成を受けて行った。本研究は第二著者(切替このみ氏)の修士論文に相当する課題研究論文の内容の追加解析等を行い、一部改変して発表した。

ゲーミフィケーションを活用した 珠算学習支援システム「CalcQuest」のユーザスタディ

小嵯泰造^{†1}

奈良先端科学技術大学院大学

松田裕貴^{†2}

岡山大学

1. はじめに

そろばんは珠を上下に動かして数値を表現し、計算を補助する伝統的な道具である。珠算学習は現在も習い事として存在するが、電子計算機の普及により使用機会は減少している。しかし、珠算学習には計算力や集中力の向上などの効果があり [1, 2, 3, 4, 5, 6, 7]、日本のみならず海外でも注目されている。

珠算能力の習得には、数値の表現方法と正確な指の動きを身につけることが不可欠であるため、一般的にそろばん教室や学校で指導者主導の学習が行われる。指導者は計算結果の正誤判定やミスの指摘を行うが、人手に依存する点が課題である。また、珠の操作には多様な種類があり、迅速かつ正確な操作が求められるため、学習者は「苦手操作」に直面しやすい。その結果、特定の操作でミスを繰り返し、苦手操作として定着してしまうことがある。苦手操作克服には繰り返し学習が必要だが、苦手操作ばかりの練習は心理的負担が大きく、モチベーションの低下を招く。また、紙の教本では学習者ごとに適した問題が提供されるとは限らない。これを解決するため、機械学習を活用し、苦手な問題を自動生成する手法が提案されている [8, 9] が、学習者自身によるラベル付けの作業が必要であり、学習者にとって負担が大きいといえる。著者らはこれまでに、市販のそろばんに AR マーカを貼り、書画カメラで撮影した映像を分析することで盤面を認識し、リアルタイムに学習支援情報を提示するシステムを提案している [10, 11, 12]。また、ゲーミフィケーションを取り入れた苦手操作克服のための学習支援システムの設計や問題生成手法も提案してきた [13, 14]。

本研究では、これらの手法を統合することで、ゲーミフィケーションを取り入れた繰り返し学習により指導者なしでも苦手操作の克服を可能とするシステム「CalcQuest」を提案する。本論文では、CalcQuest の設計・実装を述べ、そろばん教室の生徒 10 名を対象としたユーザスタディを通じて、有効性やユーザビリティを評価した結果を報告する。

2. 関連研究

2.1. 珠算学習支援システムに関する研究

猪俣ら [15] は、カメラと AR マーカを用いてそろばんの盤面を認識し、HMD 上に学習支援情報を重畳表示するシステムを提案した。これにより、初学者は運指や計算状況に応じた支援を受けながら珠算操作を習得できる。北川ら [16] は、RGB カメラを用いてそろばんの盤面をリアルタイムで推定し、プロジェクションマッピングにより計算支援情報を提示するシステムを開発した。これにより、計算手順や指使いを視覚的に提示し、学習を進めるインタフェースが提供されている。新川ら [17] は、LMS (Learning Management System) を備えた Web 学習支援システムを開発した。フラッシュ暗算の学習機能を搭載し、学習者は進捗管理を行いながら、自宅でも学習できる。齋藤ら [18] は、スマートフォンやタブレット端末を利用した学習支援ツールを提案している。このツールでは、端末画面上にそろばんを再現し、数値や操作手順を視覚的に提示する機能を備え、初心者でも無理なくそろばんの操作を学べるよう工夫されている。Digika 社 [19] は、タブレット上でそろばんに類似した UI を再現し、珠の操作をボタン入力に置き換えた「そろタッチ」を提供している。両手を使ったボタン操作により、そろばんの感覚を保ちつつ学習が可能となっている。

2.2. ゲーミフィケーションを活用した教育支援に関する研究

沼口ら [20] は、英単語の暗記学習を支援するシステムを提案した。作物の成長を通じて復習を促し、視覚的な復習タイミングの提示により学習を支援する。ゲーミフィケーションの導入により、モチベーション維持と学習効果向上が確認された。谷本ら [21] は、C 言語の文法やアルゴリズム学習を対象としたアプリ「Code Quiz」を提案した。レベルやバッジ、ポイントを活用し、ゲーム感覚で学習を進められる設計になっている。5 名の被験者による予備実験では、約 1 週間の利用後にモチベーション維持が確認された。大濱ら [22] は、相対音感向上を目的とした音楽学習アプリ「オトクマ」を提案した。ガチャや図鑑、マスコットの応援

User Study of A Gamified Abacus Learning Support System
“CalcQuest”

^{†1} TAIZO KOZAKI, Nara Institute of Science and Technology

^{†2} Yuki Matsuda, Okayama University

機能を活用し、ゲーミフィケーションを組み込んだ相対音感トレーニングシステムを提案している。22名の被験者による2週間の実験で音感スコアの向上が確認された。鈴木[23]は、SDGsをテーマにしたゲーミフィケーション教材「X(クロス)」をデジタル化し、オンライン環境で利用可能にした。視覚性、操作性、理解性の面で高評価を得ており、紙媒体と同等以上の満足度が確認された。

2.3. 本研究の立ち位置

珠算学習の向上には、計算過程における苦手操作や誤操作の発見と指導が重要だが、従来の方法は指導者に依存している。このため、学習者の珠操作をセンシングし、状況に応じた学習支援を提供することで、学習効率を向上させることが可能と考えた。既存の珠算学習支援システムは、初学者に対する珠の動かし方や計算手順の指導は可能だが、苦手操作の検出や克服のための問題生成機能は備えていない。一方、ゲーミフィケーションを活用した教育支援システムでは、学習のモチベーション向上が確認されている。これらを踏まえ、本研究では、人手を介さずに学習可能で、ゲーミフィケーションを活用した繰り返し学習により苦手操作を克服できるシステムの実現を目指す。著者らはこれまでに、ゲームシナリオの構想[13]や苦手操作の検出方法[14]を提案してきた。本論文では、それらを実際のシステム「CalcQuest」として実装するとともに、システムの有効性、ユーザビリティについて実際の珠算学習者を対象とした実験を通じて評価した結果を報告する。

3. CalcQuest : システムの設計

3.1. アプローチ

従来の珠算学習では、指導者が学習者に指摘を行うことが求められ、人手に依存する点が課題であった。また、苦手操作の克服には繰り返し練習が必要だが、同じ操作の練習がストレスやモチベーション低下を引き起こす可能性がある。そこで、ICTを活用した学習の効率化と、ゲーミフィケーションによるモチベーション向上の知見に基づき、珠算能力の向上を図る。本研究では、書画カメラを用いたARマーカ付きそろばんの盤面認識、苦手操作の検出と克服のための問題生成、ゲーミフィケーションを組み込んだ学習支援システム「CalcQuest」を提案する。

3.2. システム設計

本稿では、著者らがこれまでに提案してきた[10, 11, 13, 14]の手法を元に開発したシステムを提案する。システム構成を図1に示す。システムはARマーカを貼付した市販のそろばん、そろばんの盤面を俯瞰的な視点から撮像する書

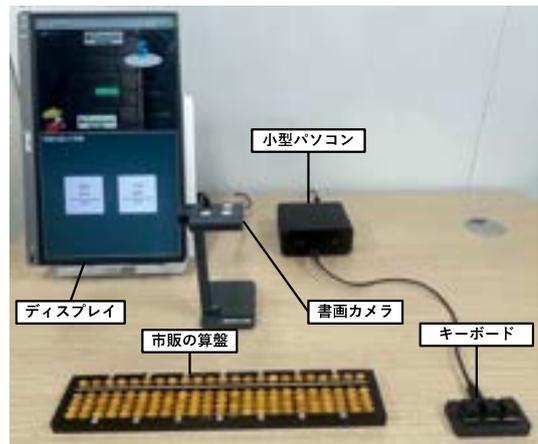


図1 システム構成

画カメラ、そろばんの盤面認識・苦手操作の検出・問題の生成を実行する小型パソコン、ゲームの画面を提示するディスプレイ、盤面認識結果を小型パソコンに送信するためのキーボードから構成される。

3.3. 苦手操作克服機能

珠算学習において、正確な珠操作を習得することが重要であるが、珠操作の種類は多様であり、学習者が苦手操作に直面することが多い。そこで、独自で加算と減算の際の6パターンの苦手操作を定義した[14]。以降ではそれらについて改めて述べた上で、苦手操作の検出および苦手問題を含む問題生成の方法について述べる。

3.3.1 苦手操作の定義

パターン1：加算操作の不足

パターン1は本来足すべき数字を足していない苦手操作である。パターン1の苦手操作の例である図2に示すように「3を足す」という操作が行なわれない場合である。

パターン2：減算操作の不足

パターン2は本来引くべき数字を引いていない苦手操作である。パターン2の苦手操作の例である図3に示すように「5を減ずる」という操作が行なわれない場合である。

パターン3：加算量の錯誤

パターン3は本来足すべき数字より多い数または少ない数を足す苦手操作である。パターン3の苦手操作の例である図4に示すように「1を足す」という操作が「2を足す」という操作になり、本来足すべき数字より多い数を足す場合である。

パターン4：減算量の錯誤

パターン4は本来引くべき数字より多い数または少ない数を引く操作である。パターン4の苦手操作の例である図5に示すように「5を引く」という操作が「6を引く」という操作になり、本来引くべき数字より多い数を引く場合である。

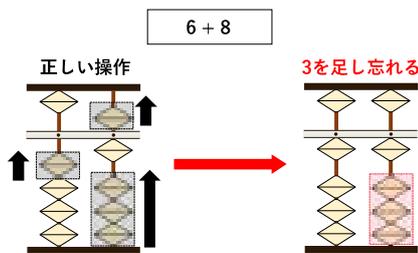


図2 パターン1の加算操作の不足

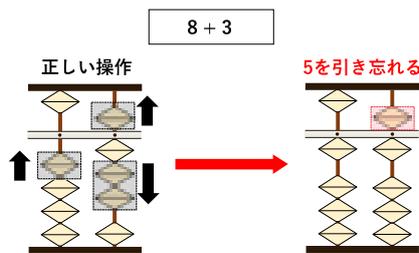


図3 パターン2の減算操作の不足

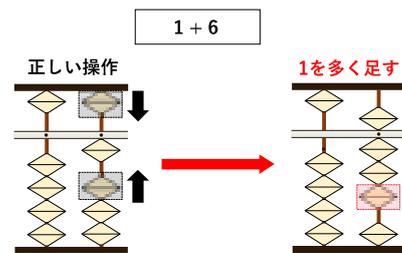


図4 パターン3の加算量の錯誤

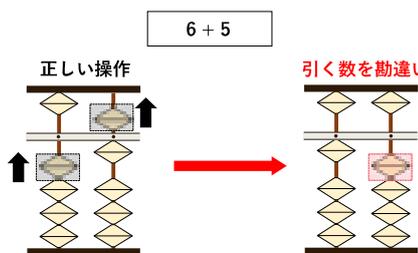


図5 パターン4の減算量の錯誤

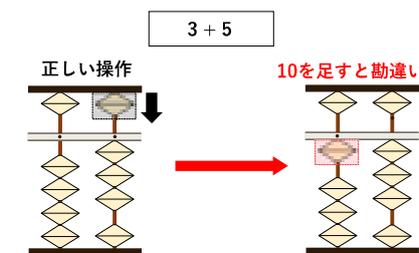


図6 パターン5の五珠・十珠の錯誤

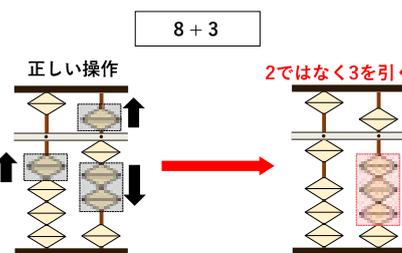


図7 パターン6の加減操作の反転

パターン5：五珠・十珠の錯誤

パターン5は本来足すべき数字を誤って引いてしまうまたは引くべき数字を誤って足してしまう操作である。パターン5の苦手操作の例である図6に示すように「5を足す」という操作が「10を足す」という操作になり、本来5を足す操作を誤って10足してしまう場合である。

パターン6：加減操作の反転

パターン6は本来足すべき数字を誤って引いてしまうまたは引くべき数字を誤って足してしまう操作である。パターン6の苦手操作の例である図7に示すように10の補数を勘違いし、「7を減ずる」操作を「3を減ずる」と操作してしまい、本来足すべき数字を誤って引いてしまう場合である。

3.3.2 苦手操作検出手法

苦手操作は、学習者が同じ誤操作を繰り返す傾向を基に推定される。誤操作が発生する場合、学習者が特定の操作を正確に理解していない、または忘れていた可能性が考えられる。本研究では、正誤判定に基づき、認識結果と出題された計算問題の正答の値との間に生じる差分を算出し、それをを用いて苦手操作を検出する手法を提案する。具体的には、出題された計算問題の正答の値と、誤操作を含む認識結果を一桁ごとに分解し、それぞれの桁における差分値を計算することで、誤操作を推定する。この推定結果を基に、前述した6つの苦手操作のパターンを検出する。また、一桁の各操作から想定される誤操作の値と正答の値の差分値は提案システム内のデータベースに存在し、差分値を参照することでパターンを検出できる仕組みとなっている。

3.3.3 苦手問題生成手法

苦手操作のパターン検出結果に基づき、苦手操作を含む問題を生成する。具体的には苦手操作のパターン検出が繰り返

返されることで特定のパターンの検出回数が増加し、検出回数が5回を超えると苦手操作としてフラグが立つ。このフラグが立った場合、該当する苦手操作を含む問題が優先的に出題される。次に、出題される問題では、過去に検出された苦手操作を含む問題が生成され出題される。苦手操作を含む問題を正答すると、苦手操作のパターンの蓄積回数が1回減少する。この減少を繰り返し、苦手操作のパターンの回数が3回以下になった場合、苦手操作を克服したとみなしてフラグが解除される。フラグが解除された場合、その苦手操作を含む問題は出題されなくなる。しかし、再度同じパターンで検出回数が5回を超えると苦手操作としてフラグが立ち、同様の手順で苦手操作を含む問題が優先的に出題される。一方、苦手操作のパターンが一つも検出されない場合は、ランダムな問題が生成される。

3.4. ゲームシナリオとシステムの実装

著者らは、珠算学習支援のためのゲームシナリオとして、通常の珠算学習環境（個人が問題集を解くことで学習を進める形態）を想定し、プレイヤーとコンピュータの敵で交互に攻撃を行い戦うターン制コマンドバトル（PvEゲーム）のシナリオをこれまでに提案してきた[13]。ここでは、前述の苦手操作の検出および問題生成の実装を踏まえて、ゲームシナリオを改訂するとともに実装を行う。改訂したゲームのシステムフローを図8に示す。以降ではそれぞれのフェーズについて改訂した点及び実装内容について述べる。

■ 攻撃フェーズ (P_A)

攻撃フェーズでは、プレイヤーにとって難易度が異なる2つの問題が提示される。一方の問題はプレイヤーの苦手操作を必ず含む問題が提示されるが、苦手操作がまだ検出

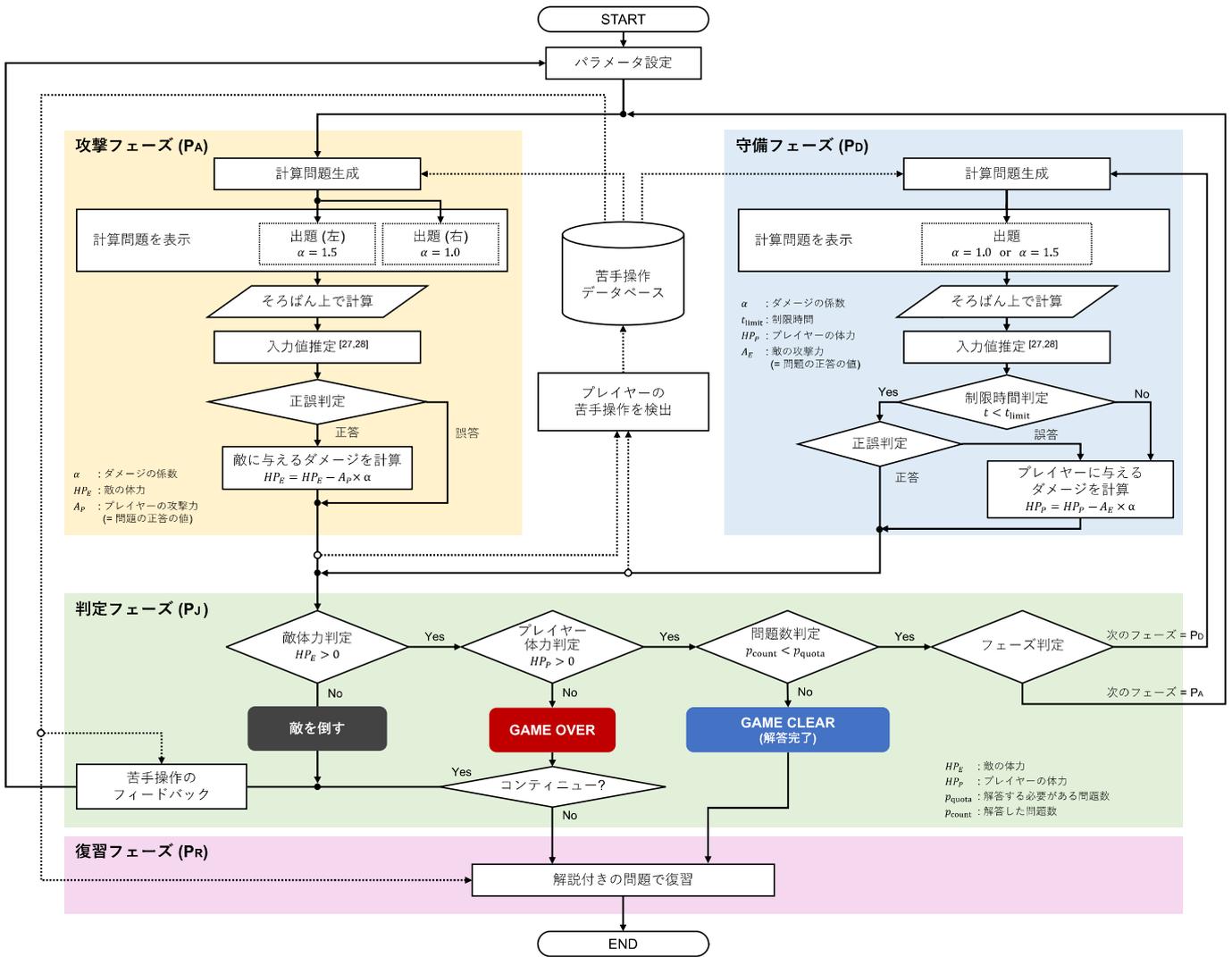


図8 PvEゲームを用いた珠算学習支援のゲームシナリオ (著者らの構想 [13] をベースに改良)

されていない状態であれば、ランダムな問題が提示される。もう一方の問題はプレイヤーの苦手操作を含まないランダムな問題が提示される。プレイヤーは提示された2つの問題のうち、どちらかひとつを選択してそろばんを用いて計算し、計算結果からどちらの問題を選択したかを判別する。正答であった場合はどちらを選択したかわかるが、誤答であった場合は、認識結果の値と計算問題の正答の値の差分結果から選択した問題を判別する。

■ 守備フェーズ (P_D)

守備フェーズでは、プレイヤーに対し苦手操作を含んだ問題を提示する。ここで、苦手操作がまだ検出されていない状態であれば、ランダムな問題が提示される。出題された計算問題を制限時間内に正答した場合には、敵の攻撃を回避し、制限時間内に計算が終了しなかった場合、あるいは、計算結果が誤答であった場合には、プレイヤーにダメージが与えられる。

■ 判定フェーズ (P_J)

判定フェーズでは、プレイヤーおよび敵の体力 (HP_P, HP_E) の残量の判定、および、次に遷移するフェーズを判定する。敵の体力がなくなった場合 (HP_E ≤ 0)、新たな敵が出現する。プレイヤーの体力がなくなった場合 (HP_P ≤ 0) はゲームオーバー状態となり、現在のゲームを終了するフローでゲームを続行するかどうかをプレイヤーに要求する。コンティニューすると選択した場合、更新されたパラメータに基づく新たなゲームが開始される。コンティニューしないと選択した場合、復習フェーズに移行する。なお、プレイヤー・敵の双方について体力が残っている場合、次に遷移するべきフェーズを判定する。現在のフェーズが攻撃フェーズ (P_A) の場合、守備フェーズ (P_D) に入り、そうでなければ攻撃フェーズ (P_A) に入る。さらに、特定の問題数を解答する必要があり、解答数を満たすと復習フェーズに移行する。

■ 復習フェーズ (P_R)

ゲーム終了後に復習フェーズに入る。復習フェーズでは、ゲーム中に不正解となった問題が表示され、プレイヤーはその問題を復習して解く。各復習問題には、対応する苦手操作パターンの解説が画面上部に表示される。

4. 評価実験

本研究では、提案システム CalcQuest の有効性とユーザビリティを評価するため、アンケート調査を実施した。珠算学習にゲーミフィケーションを取り入れることで、学習意欲が向上するか(有効性)と、システムの使いやすさ(ユーザビリティ)を検証する。本研究は奈良先端科学技術大学院大学「人を対象とする研究に関する倫理審査委員会」の承認を受けて実施した(承認番号:2022-I-63-1)。

4.1. 参加者

大阪府枚方市のそろばん教室に通う生徒 10 名を対象とし、1 年以上の経験者 8 名、1 年未満の経験者 2 名が含まれる。参加者の珠算・暗算の検定資格は異なり、全員が CalcQuest を初めて利用した。

4.2. 実験概要

本実験では、そろばん教室において提案システムを使用し、学習後にアンケートを実施した。以下の 3 条件を比較対象とした。

提案システムなし(通常授業): 従来のそろばん教室の授業をベースラインとし、提案システムの有効性およびユーザビリティを評価する。

提案システムのゲームありモード: ゲームシナリオおよびデザインを適用したシステムを使用する。ただし、本実験では苦手操作の検出や問題生成機能は搭載せず、ゲーミフィケーションの有効性とユーザビリティの評価に焦点を当てる。また、本研究では生徒のレベルに関わらず、制限時間の秒数を 10 秒とする。

提案システムのゲームなしモード: 提案システムを使用するが、ゲームシナリオやデザインは適用しない。

実験は以下の手順で実施した。

1. 事前説明: 実験の目的とシステムの使用方法を説明する。
2. 提案システムの使用: 各被験者は、ゲームありモードとゲームなしモードをそれぞれ 10 分間使用し、学習を行う。順序効果を防ぐため、モードの順番は被験者ごとにランダムに決定する。また、各モード終了後に 5 分間の休憩を設けた。本システムでは、計算問題の難易

度(桁数)を選択できるため、事前に難易度に関する質問を行い、被験者の熟練度に応じて適切な問題を出題した(例:3桁2口の計算では324+420)。

3. アンケート調査: システムの有効性とユーザビリティに関するアンケートを実施した。システムの有効性に関するアンケートでは、2つの形式の質問を行った。一つは、普段の授業、ゲームありモード、ゲームなしモードのそれぞれに対する個別の評価を5段階のリッカート尺度を用いたアンケートである。もう一つは、普段の授業、ゲームありモード、ゲームなしモードを比較してもらい順位を回答するランキング形式のアンケートである。ユーザビリティに関するアンケートについては、システムの使い勝手を評価するための質問票である System Usability Scale (SUS) [24] を平易な日本語に訳したものをを用いた。さらに、提案システムの利点や欠点に関する自由記述欄を設けた。システムの有効性に関するアンケートの質問は図9、ランキング形式のアンケートの質問は図10、図11、図12、ユーザビリティに関するアンケートの質問は図13に示すとおりである。なお、Q1~Q24に関しては具体的な選択肢の表現は異なっているが(例えばQ1は「5. わかりやすい」~「1. わかりにくい」、Q2は「5. 上手になった」~「1. 上手にならなかった」という選択肢)、いずれも「5」はポジティブ、「1」はネガティブな選択肢となっている。

また、実験にあたっては、安全性・円滑な実験実施のために、以下の条件を設定するとともに被験者に事前に説明を行った。

- 実験中に体調が悪くなったら、中断すること。
- 誤作動があった場合は、挙手して伝えること。
- プレイヤーの HP が 0 になった場合は、挙手して伝えること。
- カメラの画角内にそろばんの位置を収めること。
- そろばんの入力位置に手を覆い被せてはいけないこと。

5. 結果と考察

5.1. システムの有効性に関する結果(5段階のリッカート尺度)

図9に示す、5段階のリッカート尺度に基づくアンケート結果から、以下の3つの観点について分析・考察する。

そろばん教室の授業(Q1~Q5)

そろばん教室の授業に関する質問では、授業はわかりやすく、そろばんが上手になると感じる回答が多いことがわかる。しかし、Q3, Q4, Q5に関しては高い評価をしている回答が少ないことがわかる。また、Q1, Q4では2と評



図9 システムの有効性に関するアンケート結果

価しているネガティブな回答も見受けられる。この結果から、普通の珠算学習に対して、高いモチベーションを持っていない生徒がいることがわかる。

ゲームありモード (Q6~Q14)

ゲームありモードに関する質問では、全ての質問に対して2以下の回答がなく、全体的に高い評価をしている回答が多く見受けられた。特に、Q6, Q8, Q9, Q13の質問では、平均4.9と高い数値になっている。理由として、Q6, Q8に関しては、子どもでもわかりやすいようなゲームシナリオ設計が楽しさやわかりやすさに影響していると考えられる。Q9, Q13に関しては、制限時間機能により、集中力や速さが必要とされていたことが影響していると考えられる。

ゲームなしモード (Q15~Q23)

ゲームなしモードに関する質問では、ゲームありモードと比べて、高い評価をしている回答が少ないことが見受けられる。内容や楽しさ、モチベーションに対する回答が比

較的に高い評価をしている被験者が少ないことが確認できた。また、Q22, Q23に関しては、否定的な回答があることがわかる。この回答をした生徒のゲームありモードについての同様の質問(Q13, Q14)では、いずれも5の評価をしている回答であることが確認できた。この結果から、制限時間機能が珠算能力の向上を促す可能性があると言える。

5.2. システムの有効性に関する結果 (ランキング形式)

図10, 図11, 図12に示す、ランキング形式のアンケート結果から、提案手法の有効性について分析・考察を行う。まず、Q24(どれがいちばん上手になるとおもいますか?)という問いについては、そろばん教室の授業が最も高い評価を得ることがわかった。そろばん教室の先生による指導は対人での直接的な指導が可能であるためこのように高い評価が得られたものと考えられる。これに対し、楽しさ(Q25)やモチベーション(Q26)に関する問いに対しては、ゲーム

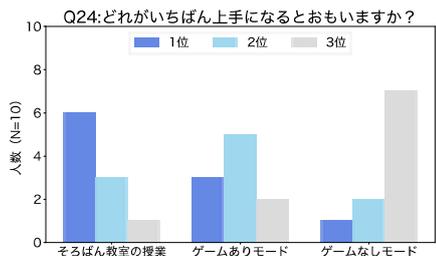


図 10 アンケート結果 (Q24)

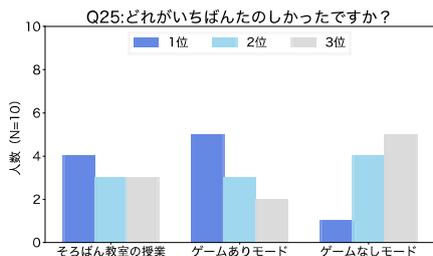


図 11 アンケート結果 (Q25)

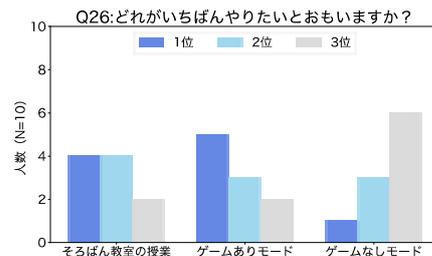


図 12 アンケート結果 (Q26)

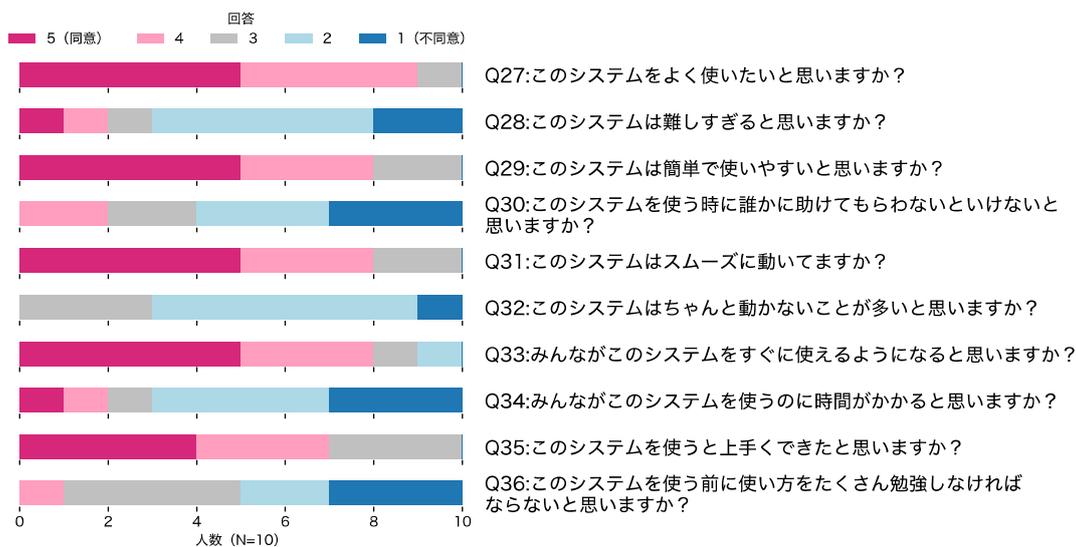


図 13 ユーザビリティに関するアンケート結果

ありモードを1位とした回答が最も多くなることがわかった。また、ゲームなしモードはどの質問でも3位の回答が多い結果となった。この結果から、CalcQuestは、そろばんの先生と比較すると珠算能力を向上するための機能改善の余地があるものの、珠算学習に対するモチベーションの向上に影響を与えることが確認できた。自由記述式の回答では、「ゲームありモードの制限時間により、計算を速く解かないといけないからプラスで楽しくなる」「ゲームだから楽しく続けられる」「モンスターを倒すところが楽しかった」など好意的な感想が得られた。また、「ゲームなしモードでは計算のみに集中でき、自身のペースで進めれるところがいい」などの意見もあり、ゲーミフィケーションを取り入れないシステムでも肯定的な意見があった。一方、「ゲームありモードの制限時間が速い」などの指摘もあった。現在のシステムでは、事前に学習者の難易度に合わせた制限時間を算出するシステムが組み込まれていないため、今後改善する必要があると考えられる。

5.3. ユーザビリティに関する結果

ユーザビリティに関するアンケート結果を図13に示す。System Usability Scale (SUS)[24]に基づくアンケートの結

果からスコアを算出した結果、74.5となった。SUSの平均的なスコアは68とされており、平均値を超えるとユーザビリティが良いと判断される。このことから、提案システムのSUSは平均値を大幅に超えており、ユーザビリティについて高い評価を得たといえる。しかし、Q28、Q30、Q32、Q34、Q36において、少数ではあるが否定的な意見も確認された。これらの原因としては、そろばんの位置をカメラの画角内に収めなければいけない点や、そろばんの入力位置に手を覆い被せてはならないといった、盤面認識を正確に行うためのシステム依存の制約が影響したと考えられる。また、実験中に特定の場所において光の反射などに起因する盤面認識が遅れが生じ、被験者が誤作動と捉える事象が発生していたことも否定的な意見の要因であると考えられる。

6. おわりに

本研究では、長期の繰り返し学習が必要であった従来の珠算学習をICTによる効率化するとともに、必要な繰り返し学習や苦手操作の克服を楽しみながら行うことを実現するために、ゲーミフィケーションを活用した珠算学習支援システム「CalcQuest」を提案した。そろばん教室に通う生徒

10名を対象とした提案システムの有効性、ユーザビリティについて評価した結果、珠算学習にゲーミフィケーションを取り入れることで、珠算学習に対するモチベーションの向上に繋がることが確認できた。また、ユーザビリティについても高い評価を得ることができた。一方で、主観的な学習効果という面においては対人での指導が可能なそろばん教室の方式が依然として優位であることもわかった。今後は、学習効果の向上およびモチベーション向上を両立するシステムへと改良していくとともに、長期的な実験を通じて、苦手操作の効果的な改善や珠算能力の向上に提案手法がどの程度寄与するかについて詳細に検証する。

参考文献

- [1] Amaiwa, S. and Hatano, G.: Effects of Abacus Learning on 3rd-graders' Performance in Paper-and-Pencil Tests of Calculation, *Japanese Psychological Research*, Vol. 31, pp. 161–168 (online), 10.4992/psychol-res1954.31.161 (1989).
- [2] Stigler, J. W.: “Mental abacus”: The effect of abacus training on Chinese children's mental calculation, *Cognitive Psychology*, Vol. 16, No. 2, pp. 145–176 (online), 10.1016/0010-0285(84)90006-9 (1984).
- [3] Amaiwa, S.: The Effects of Abacus Learning on Solving Arithmetic Problems: A Comparative Study of Elementary / Junior High School Students at Upper Level and Inexperienced Students, *Journal of the Faculty of Education, Shinshu University*, Vol. 96, pp. 145–156 (1999).
- [4] Wang, C., Xu, T., Geng, F., Hu, Y., Wang, Y., Liu, H. and Chen, F.: Training on Abacus-Based Mental Calculation Enhances Visuospatial Working Memory in Children, *Journal of Neuroscience*, Vol. 39, No. 33, pp. 6439–6448 (online), 10.1523/JNEUROSCI.3195-18.2019 (2019).
- [5] Wang, C.: A Review of the Effects of Abacus Training on Cognitive Functions and Neural Systems in Humans, *Frontiers in Neuroscience*, Vol. 14, No. 913, pp. 1–12 (online), 10.3389/fnins.2020.00913 (2020).
- [6] Lu, Y., Li, M., Cui, Z., Wang, L., Hu, Y. and Zhou, X.: Transfer Effects of Abacus Training on Cognition, *Current Psychology*, Vol. 42, pp. 6271–6286 (online), 10.1007/s12144-021-01968-1 (2023).
- [7] Hu, Y., Geng, F., Tao, L., Hu, N., Du, F., Fu, K. and Chen, F.: Enhanced White Matter Tracts Integrity in Children With Abacus Training, *Human Brain Mapping*, Vol. 32, No. 1, pp. 10–21 (online), 10.1002/hbm.20996 (2011).
- [8] 原子弘務, 井上一磨, 諏訪貴大, 福岡省伍, 村田遼, 須子統太: 珠算競技における効果的な練習問題の自動生成法について, 第 80 回全国大会講演論文集, No. 1, pp. 897–898 (2018).
- [9] 北村瑠菜, 原子弘務, 守屋郁宏, 神頭和希, 於勢奈都子, 角田和正, 須子統太: 珠算競技における苦手問題自動作成法について～かけ算に対する検討～, 第 81 回全国大会講演論文集, No. 1, pp. 731–732 (2019).
- [10] 松田裕貴: 書画カメラを用いた珠算行動センシング, 電子情報通信学会技術研究報告, センサネットワークとモバイルインテリジェンス研究会 (SeMI), Vol. 123, No. 31, pp. 70–75 (2023).
- [11] Matsuda, Y.: Abacus Manipulation Understanding by Behavior Sensing Utilizing Document Camera as a Sensor, *International Journal of Activity and Behavior Computing*, Vol. 2024, No. 1, pp. 1–16 (online), 10.60401/ijabc.2 (2024).
- [12] Matsuda, Y.: A Table-top Interface for Real-time Coaching in Abacus Learning, *The 10th International Conference on Smart Computing (SmartComp '24, Demo)*, pp. 243–245 (online), 10.1109/SMART-COMP61445.2024.00056 (2024).
- [13] 小嵯泰造, 松田裕貴: 算盤の苦手操作克服のためのゲーム要素を用いた珠算学習支援手法, 第 31 回マルチメディア通信と分散処理ワークショップ (DPSWS'23), pp. 267–272 (2023).
- [14] 小嵯泰造, 松田裕貴: 珠算学習における計算結果の逐次分析による苦手操作検出手法の提案, 社会情報学会関西支部研究会 (SSD), pp. 1–8 (2024).
- [15] 猪俣駿, 阿部亨, 菅沼拓夫: MR 技術を用いた珠算学習支援システムの提案, 情報処理学会第 86 回全国大会, pp. 461–462 (2024).
- [16] 北川珠莉, 鈴木優: 珠の位置認識と操作手順の重畳表示を用いたそろばん学習支援システム, 情報処理学会インタラクシオン 2022, pp. 759–762 (2022).
- [17] 新川晃司, 川崎健志, 澤田一樹, 二石芳裕, 算宗徳, 渡邊一衛: そろばん教育における Web 学習支援システムの開発, 成蹊大学理工学研究報告, Vol. 48, No. 1, pp. 75–79 (2011).
- [18] 齋藤謙太, 佐々木整, 水野一徳: 携帯電話を利用した学習支援ツールの開発, 情報科学技術フォーラム講演論文集, Vol. 8, No. 3, pp. 653–654 (2009).
- [19] 株式会社 Digika: そろタッチ, <https://www.sorotouch.jp/>. (accessed 2023-08-24).
- [20] 沼口天, 松村敦: ゲーミフィケーションを利用した暗記学習支援システム, 情報処理学会研究報告, pp. 1–8 (2022).
- [21] 谷本嵩晃, 崔恩壽, 水野修: ゲーミフィケーションを用いた C 言語の文法やアルゴリズムの学習支援アプリケーション Code Quiz の提案, ソフトウェア工学フォーラム (FOSE2022), pp. 201–202 (2022).
- [22] 大濱慎太郎, 山田雄貴, 梅村祐輝, 小林舞衣, 佐藤史章, 野村悠哉, 橘俊宏: ゲーミフィケーションを活用した音楽学習アプリによる相対音感の向上, 情報処理学会第 86 回全国大会, pp. 899–900 (2024).
- [23] 鈴木健斗: SDGs ゲーミフィケーション教材のアプリケーション開発と操作性向上に関する研究, 日本教育工学会研究報告集, pp. 40–45 (2021).
- [24] Brooke, J.: SUS: A quick and dirty usability scale, *Usability Eval. Ind.*, Vol. 189, pp. 1–7 (1995).

生成 AI を活用した能動的学習環境の構築と ハルシネーション問題の検討

小北駿^{†1}
はこだて未来大学^{†1}

角薫^{†2}
はこだて未来大学^{†2}

1. はじめに

近年、生成 AI (Generative Artificial Intelligence) の発展が著しく、とりわけ大規模言語モデル (Large Language Models: LLMs) を活用したサービスが急速に普及している[1][2]。これらの AI システムはコーディング、翻訳、日常会話、創作活動など幅広い領域で活用されており、教育分野でもその有用性が認識され始めている[3][4]。特に、AI を学習支援ツールとして活用する動向が顕著であり、学習者の進捗に応じた指導やフィードバックを提供する試みが進められている[5]。

生成 AI をチューターとして利用することにより、学習者は個別最適化された指導を受けることが可能となる。例えば、学習者の理解度に応じて AI が問題の難易度を調整したり、リアルタイムでフィードバックを提供したりすることで、学習の効率性を向上させる[6]。また、生成 AI は従来の学習支援ツールと異なり、対話型のインタラクションを提供できるため、学習者の興味・関心を引きつけ[4]、能動的な学習 (Active Learning) を促す可能性がある[3][5]。

しかしながら、生成 AI の教育利用にはいくつかの課題が指摘されている。第一に、「ハルシネーション (Hallucination)」と呼ばれる誤情報の生成問題が挙げられる[8]。LLM は確率的推論に基づいて応答を生成するため、学習者が誤った情報を無批判に受け入れるリスクがある[9]。特に、教育現場においてはこの問題がより深刻となることが示唆されている。例えば、Ahmed Salem and Kaoru Sumi の研究では、日本の中学生を対象に生成 AI を組み込んだ教育ロボットとの対話実験を行い、学習者の大多数が虚偽情報を正確性の検証なしに受け入れてしまう傾向を持つことが明らかになった[9]。これにより、AI の出力の信頼性を確保することの重要性が再認識されている。

第二に、受動的な学習環境が助長される懸念である。例えば、AI が即座に回答を提示することで、学習者が自ら情報を探索・整理するプロセスが省略され、結果として深い理解に至らないケースが報告されている[10]。このような状況は、特に初等・中等教育の場において顕著であり、学習者が AI の回答をそのまま受け入れてしまうことで、思考力や批判的思考の発達が阻害される可能性がある[11]。また、安川は、

日本の教育現場で 24 時間対応の AI チューターを導入した実験において、学習者が AI を利用する頻度が増える一方で、AI の回答の正確性を確認しようとする意識が低下していることを指摘している[11]。

このような課題を踏まえ、本研究では「信頼性の高い情報を提供する仕組みの構築」および「能動的な学習を促す環境の設計」に焦点を当てた学習支援システムの開発を試みる。具体的には、学習範囲を制限することでハルシネーションを抑制し、一貫した情報提供を可能にする手法を検討する[2][7]。学習者が自身で問いを立て、推論しながら学習を進めることができるような環境を設計することで、受動的な学習から能動的な学習への転換を目指す[10]。

さらに、本研究では実際の教育現場において開発した学習システムを運用し、その効果を検証する。具体的には、VR 技術を活用した歴史学習システムを構築し、小学生を対象とした実験を行う。これにより、生成 AI を活用した学習環境の有効性を定量的・定性的に評価し、学習成果やユーザー体験の向上に寄与する要因を明らかにすることを目的とする。

2. 関連研究

近年、生成 AI (Generative AI) を活用した教育システムの研究が進められており、特に大規模言語モデル (LLMs) を用いた学習支援の有用性が指摘されている[1][2]。例えば、Zapata-Rivera らは、教育分野における生成 AI の活用について総合的な調査を行い、個別指導、学習者のモチベーション向上、教材のカスタマイズなどの点で高い可能性を示している[1]。また、Kasneci らは、ChatGPT などの LLM を学習支援に応用することで、学習者が自ら質問を作成し、対話を通じて知識を深める能動的学習環境の構築が可能であることを提案している[4]。

さらに、Mollick & Mollick は、AI による即時フィードバックが学習効率を向上させることを示し[5]、Wollny らの研究では、AI チューターが学習者の関心に応じて適切なヒントを提供することで、学習の定着率が向上することが報告されている[6]。これらの研究は、生成 AI が教育において有益なツールであることを示唆しているが、一方で誤情報 (ハルシネーション) や受動的学習の助長といった課題も指摘されている[7][8]。

LLM を活用した学習支援において、「ハルシネーション (Hallucination)」と呼ばれる誤情報の生成が大きな問題となる[9]。Bender らは、LLM の特性上、確率的推論に基づく応

Building an Active Learning Environment Using Generative AI and Examining the Hallucination Problem

†1 SHUN KOKITA, Future University Hakodate

†2 KAORU SUMI, Future University Hakodate

答が発生するため、出力された情報が必ずしも正確ではないことを指摘している[10]。また、Jiらは、ハルシネーションが発生する要因として以下の3点を挙げている[7]：

1. トレーニングデータの偏り：学習データが偏っている場合、特定の視点に基づいた誤情報が強化される可能性がある。
2. 文脈の欠落：LLMは文脈を完全に理解するわけではなく、適切な情報を補完する際に推論を行うため、誤った内容が生成される可能性がある。
3. 確率的生成の特性：同じ質問に対して異なる回答を生成する特性があり、学習者が一貫した知識を得にくい場合がある。

特に、教育現場における影響は深刻であり、学習者が誤情報を無批判に受け入れてしまうリスクが指摘されている[12]。Ahmed & Sumの研究では、日本の中学生を対象にAIロボットとの対話を行わせたところ、ハルシネーションによる誤情報でも高い確率で信じてしまうことが実験で確認された[9]。これは、学習者がAIの出力を権威ある情報源と認識しやすい傾向に起因しており、教育現場でのAI導入には慎重な対策が求められる。

また、安川は、ChatGPTを活用した24時間対応のTA(Teaching Assistant)システムを教育現場に導入した結果、学習者がAIの出力を過信し、情報の正確性を検証しなくなる傾向があることを報告している[11]。これは、学習プロセスの中で「答えを考える」機会が失われ、結果として受動的な学習が助長される問題と関連する[15]。

生成AIを教育に導入する際、受動的学習(Passive Learning)が助長される可能性が指摘されている[16]。従来の教育方法では、学習者が自ら情報を検索・整理し、知識を獲得するプロセスを経ることが重要視されてきた。しかし、AIを利用することで即時の回答を得られる環境が整うため、学習者が自発的に考えなくなるという懸念がある[17]。

Zhangらは、LLMを利用した学習環境において、学習者が能動的に問いを立てることが少なくなり、AIの出力をそのまま受け入れるケースが増えることを指摘している[18]。これに対し、Thoppilanらは、AIの回答に「考えさせるプロンプト」を加えることで、学習者の思考プロセスを促す手法を提案している[19]。また、Chiらは、AIによるヒント提供の仕組みを導入することで、学習者が試行錯誤しながら解答を導き出す環境を構築することが可能であると報告している[20]。

本研究では、これらの研究を踏まえ、学習範囲を制限し、信頼性の高い情報を提供する仕組みと、能動的な学習を促す環境の設計を組み合わせた学習支援システムを提案する。

3. 提案手法

本章では、課題解決への提案手法の概要とそれらの実装方法について述べる。

3.1 提案手法の概要

本研究では、生成AIを活用した教育支援システムの構築において、ハルシネーションの抑制と能動的学習の促進を主な課題とし、以下の2つのアプローチを採用する。

1. 学習範囲の制限によるハルシネーションの抑制
 - ・ ChatGPTのFine-tuningを実施し、徳川家康に関する正確な情報を学習させ、誤情報の出力を抑制する。
 - ・ 特定のプロンプト設定により、AIが歴史的事実に基づいた一貫性のある回答を生成するようにする。
2. 能動的学習を促すVR学習環境の構築
 - ・ 学習者が単にAIの回答を受け取るのではなく、自ら問いを立て、推論しながら学習を進めることができるシステムを開発する。
 - ・ VR環境を活用し、仮想空間内で歴史的出来事を体験しながら学習できるシステムを構築する。

本研究では、これらのアプローチを組み合わせ、VRとChatGPTを統合した教育支援システムを開発し、その有効性を検証する。

3.2 学習範囲の制限によるハルシネーションの抑制

LLMを活用した学習支援においては、モデルが学習データの偏りや確率的推論に基づく誤情報(ハルシネーション)を生成する可能性がある。特に教育分野では、誤った情報が学習者に定着することを防ぐ必要がある。そこで、本研究では、学習範囲を明確に定義し、徳川家康に関する事実のみに特化した応答を生成するモデルを構築する。

本研究で用いるChatGPTのモデルは、Fine-tuningが可能な「gpt-4o」を使用する。モデルの調整には以下の3つの手法を用いる。

1. Fine-tuningの適用

ChatGPTに対し、徳川家康に関する正確な歴史的知識のみを学習させることで、不適切な推論を抑制する。Fine-tuningデータセットには、Fine-tuningデータセットには、日本の歴史研究書・教科書を基にした質問と正確な回答を100セット程度用意し、JSON形式で作成した([付録I])。

2. プロンプトエンジニアリングの活用

ChatGPTが「徳川家康」として回答するよう設定し、一人称視点での回答を促す。回答可能な範囲を事前に指定し、それ以外の質問には「それについては詳しく知らぬ」などの応答を返すよう設定する([付録II])。

3. 一貫性のある回答を保証するためのデータ作成

生成AIの回答を事前に学習させるために、歴史的事実に基づいたデータセットを用意し、特定の質問に対して一貫性のある回答を出力できるよう調整する。

3.3 能動的な学習環境の構築

本システムでは、学習者がVR空間で徳川家康と対話しながら学ぶことで、より能動的な学習体験を提供する。学習者が歴史的な出来事をただ受動的に読むのではなく、仮想空間内で対話しながら探索できるように設計する。



図 1: VR 学習システムの対話インターフェース
音声入力を通じて徳川家康のアバターと対話する。

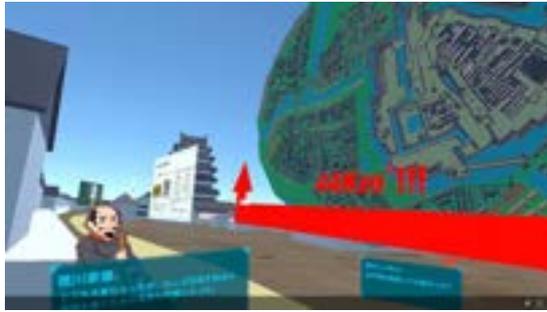


図 2: VR 学習システムの画面構成
学習者が江戸城を探索しながら、対話形式で歴史を学ぶ。

1. VR 環境での構築

学習者が徳川家康の時代に没入し、歴史的な出来事を体験しながら学習を進めることができるよう、「開拓中の江戸城」を舞台とした VR 学習環境を構築する(図 1, 図 2)。

- Oculus Quest 3 を使用し、江戸城の建設過程や当時の生活風景を再現する。
- 徳川家康のアバターと対話し、歴史上の出来事について質問することで知識を深める。
- スライド資料やキャプションを活用し、学習の導線を明確にする。

2. 音声認識による対話型学習

学習者が VR 環境内で直接質問できるよう、音声認識機能を実装し、音声入力に基づいて ChatGPT が応答する仕組みを導入する。

- Oculus Quest 3 のマイクを利用し、学習者の質問を認識する。
- 音声入力の誤認識を防ぐため、キーワード補正機能を実装し、正確な質問意図を把握するよう工夫する。
- AI の応答履歴を表示し、過去の会話を参照できるようにする。

3. 学習者の思考を促すインタラクティブ設計

学習者が単に AI の回答を受け取るのではなく、自身で考え、推論しながら学習を進めることができる仕組みを導入する。

- AI が学習者に対して「なぜそう思いますか？」と問い返す機能を実装し、思考を深める。また、徳川家康が学習者に質問を投げかける機能を実装し、対話の双方向性を高める。

- 「自分の言葉で説明する」セクションを設け、学習者が学んだ内容をまとめることで、知識の定着を促進する。
- クイズ形式のフィードバック機能を搭載し、学習内容の理解度を確認できるようにする。

3.4 システム構成

本システムでは、学習者が VR 空間で徳川家康と対話しながら学ぶことで、より能動的な学習体験を提供する。学習者が歴史的な出来事をただ受動的に読むのではなく、仮想空間内で対話しながら探索できるように設計する。

1. VR 環境の構築

Unity を使用し、「開拓中の江戸城」を 3D で再現する。学習ポイント（歴史的な出来事）を設置し、学習者が興味を持ったテーマについて AI に質問できるようにする。

2. 音声認識による対話型学習

Oculus Quest 3 のマイクを利用し、学習者の音声を認識し、ChatGPT が適切な応答を返す。音声認識精度向上のためのキーワード補正機能を実装し、質問意図が正しく理解されるようにする。

3.5 Fine-tuning の具体的な手順

Fine-tuning を実施するにあたり、学習データは OpenAI の Fine-tuning 仕様に準拠し、JSON 形式で作成した。データセット作成には以下の基準を採用した。

1. 質問のカバー範囲の設定

幼少期、戦国時代、幕府設立後など、家康の主要な歴史的フェーズを含める。

2. 正確性の保証

教科書や日本史研究書を基に、誤情報を含まないよう調整する。

3. 誤情報・誤解を生みやすい質問への対応

「家康は信長の部下だったか？」のような誤解を招く質問には、明確な歴史的背景を添えて正確な回答を作成する。

Fine-tuning の適用後、再度 ChatGPT の応答を検証し、歴史的事実の正確性や、学習者にとって分かりやすい表現になっているかを確認した。

4. 実験

本研究は ChatGPT を用いた教育支援システムを開発し、初等教育においてそのシステムの運用実験・効果測定を行うものである。運用し、取得した結果から児童のシステムに対する満足度や学習状況を評価していく。また、システムの評価にあたって、既存の学習方法と比較するために同程度の学習も並行して行った。本章では本研究で行った実験の方法と評価方法について説明する。

4.1 実験の概要と実験前の仮説

実験では、小学校の児童に 2 つの学習方法を比較するため

に、能動的な学習環境 (VR システム) と受動的な学習環境 (プリント学習) のどちらが効果的かを検証する。2回の勉強を通して、能動的な学習環境と受動的な学習環境どちらの方が効果的に学習できるか、自由記述の学習調査テストと併用して検証した。つぎに、両方の学習環境において児童が客観的に評価した授業終了後アンケートと、それぞれの授業中の児童の反応を基に興味関心の検証も行う。最後に、システムを利用した児童と ChatGPT の発話内容を記録した会話データを用いて、ハルシネーションの有無や起きる場合の傾向を分析する。

ここからは、実験前に立てた仮説について説明を行う。本実験では、「学習システムを活用した学習の方が、従来の学習よりも具体的な知識が身につく」という仮説を立てた。本システムは、ChatGPT との会話によって、教科書に書かれていること以上の情報を、学習者個人のレベルに合わせて提供してくれるという想定だ。そのため、学習調査テストの結果では学習用プリントでの勉強よりも、学習システムでの勉強の方が効果的であり、点数と回答内容に結果が表れると仮説を立てた。

また、ハルシネーションの問題について、「徳川家康の歴史に関する本質的な間違いは起こらない」という仮説を立てた。これは、3章で示した重要な話題に関する誤りが発生しないだけでなく、他者の出来事を自分の出来事のように語る現象や、会話の流れを無視した応答が起きないという前提に基づいている。しかし、ハルシネーションが起きる可能性を全て潰しているわけではないと思う。今回ハルシネーション問題を改善するために取った対策は、学習範囲の設定である。教師の役割の範囲を狭めることで、Fine-Tuning による小規模なデータ再学習を人力で行いやすくしたものである。そのため、学習中に「徳川家康視点での他者の評価」や「徳川家康と同じ年代の出来事の説明」を求められた場合は、元になった GPT のモデルの学習データから捏造した出力内容になるだろうと考えられる。

4.2 被験者と実験環境

本研究は函館市内の小学校の6年生の児童10人(男子6名, 女子4名)を対象に、2時限分授業の時間を借りて行った。学習内容は、先述の学習システムに実装した「徳川家康」と、比較実験用の同程度の学習内容として「豊臣秀吉」を取り上げた。それぞれの学習内容を表1と表2に示す。

実験は、10人の小学生を5人ずつ2グループに分けて行った。比較の公平性を保つために、教育システムが入っているOculus Quest 3を5台用意した部屋と、比較用の学習プリントを用意した部屋をそれぞれ用意して授業を行った。1時限目に2グループがそれぞれ「徳川家康」と「豊臣秀吉」を学習した。

表 1 徳川家康の学習内容

時期	キーワード
幼少期～元服	人質生活, 桶狭間の戦い, 今川家から独立
今川家独立後	清州同盟, 三方ヶ原の戦い, 長篠の戦い
信長没後～江戸城開拓	本能寺の変, 江戸城開拓, 朝鮮出兵
関ヶ原の戦い	五大老就任, 石田三成らとの敵対
征夷大將軍任命	征夷大將軍任命, 江戸幕府設立

表 2 豊臣秀吉の学習内容

時期	キーワード
青年期	農民としての子供時代, 松下之綱に仕える, 信長に仕える
信長の家臣時代	墨俣一夜城, 金ヶ崎の戦い
出世して城主へ	兵糧攻め(三木合戦), 水攻め(備中松山城の戦い)
本能寺の変後	中国代返し, 太政大臣任命, 検地, 刀狩
晩年	朝鮮侵略, 病死

2 時限目にグループの学習内容を入れ替えて「徳川家康」と「豊臣秀吉」を学習してもらう。このようにして、児童の学習内容に差が生まれないようにしている。また、今回の実験では小学生が VR 機材を利用した。そのため、VR コンテンツ利用に関するガイドライン[15]に従い、学習システムによる授業を 20 分に抑え、授業の合間に 10 分~15 分ほどの休憩時間を設けている。

4.3 比較実験用の学習について

4.2 にて、比較実験用の同程度の学習内容として「豊臣秀吉」を取り上げたと述べた。ここでは、比較用のために教科書を用いた学習に関する説明を行う。従来の学習方法と学習システムとの比較実験用に、学習用プリントの作成も行った。参考にしたのは東京出版の歴史の教科書であり、参考ページと同じ程度の文章量になるように文字数を調整した。

この資料作成にあたって、情報量が徳川家康の学習と同程



図 3 VR 学習システムを用いた授業

小学生が VR 学習システムを使用し、徳川家康と対話しながら、歴史を学習する様子。



図 4 比較実験用の学習プリントの一例

豊臣秀吉の生涯を学ぶためのプリント教材のサンプル。

度になるように、豊臣秀吉の生涯を 8 ページ以内に収めるよう調整した。また、徳川家康の学習システムと同様に、豊臣秀吉の学習用プリントの作成においても表 2 のキーワードに着目し、資料を作成した。また、実際に授業で使った学習用プリントは図 4 に例を示す。

4.4 評価方法

この実験を通して、システム利用と学習に関するデータを収集する。具体的には、会話システムのログ、テストによる学習状況データ、および授業に関する興味関心のアンケートを収集する。テストによる学習調査は表 3 に示すような自由記述形式で行い、事前と事後で同じ内容のテストを実施した。

表 3 徳川家康と豊臣秀吉の学習調査テストの内容

問い	徳川家康に関する問い	豊臣秀吉に関する問い
Q1	徳川家康の子供時代(7 才くらい)の生活を知っていますか？	豊臣秀吉の子供時代(15 才くらい)のようすを知っていますか？
Q2:	徳川家康が武士になってから(30 才くらい)の活動を知っていますか？	豊臣秀吉が織田信長に仕えていた頃(24 才くらい)の出来事を知っていますか？
Q3	織田信長が死んでからの徳川家康(50 才くらい)のできごとを知っていますか？	豊臣秀吉が天下統一するまで (50 才くらい)のできごとを知っていますか？
Q4	豊臣秀吉が死んだあとの徳川家康 (60 才くらい)のようすを知っていますか？	豊臣秀吉が天下統一してから、死ぬまで(58 才くらい)の様子を知っていますか？
Q5	徳川家康について、ほかに知っていることがあればぜひ書いてください。	豊臣秀吉について、ほかに知っていることがあればぜひ書いてください。

また、学習調査テストの採点基準を表 4 に示し、正答の基準を明確化した。

2つの授業が終了し、事後学習調査テストが終わった後に、授業終了後の感想アンケートを実施した。このアンケートでは、授業全体の満足度、VR とプリントの比較、興味関心の変化などを評価するための質問を用意し、表 5 に具体的な内容を示す。

表 4 学習調査テストの採点基準

点数	理由
0 点…不正解	正しくない内容の回答
0.5 点…中間点	正解には少し要素が足りない回答
1 点…正解	正しい内容の回答
2 点…具体的な正解	“いつ” “どこで” “誰と” “何を” “なぜ” や、出来事の名称を含む回答

表 5 授業評価アンケートの内容一覧

問い	アンケート内容
Q1	家康と会話できて、楽しかったですか？
Q2:	家康との会話で詳しく知りたいことは聞けましたか？
Q3	家康との会話で、家康が間違っていることを話していると感じたことはありますか？
Q4	VR と紙の資料の授業のどちらが好きでしたか？
Q5	VR と紙の資料の授業のどちらで学ぶとわかりやすいと感じますか？
Q6	VR と紙の資料の授業のどちらが集中できましたか？

それぞれの回答形式について、Q1 は 5 段階評価、Q2,3 は「はい/いいえ」の形式での回答、Q4,5,6 は VR か紙のどちらかを選択して理由を自由記述する形式になっている。

5. 結果

5.1 授業終了後アンケートの結果

学習システムの利用に関する満足度を評価するために、Q1 「家康との会話は楽しかったですか？」と Q2 「家康との会話で詳しく知りたいことは聞けましたか？」について、5 段階評価および「はい/いいえ」形式のアンケートを実施した。その結果、Q1 の平均点は 4.8 点となり、ほぼすべての児童が「楽しかった」と評価した。また、Q2 では 10 人中 9 人が「詳しく聞けた」と回答し、ChatGPT を活用した学習体験に対して高い満足度が示された。

Q3 「家康が間違っていることを話していると感じることがありましたか？」については、10 人中 9 人が「いいえ」と回答し、1 人の被験者が「聞き間違えがあった」と記述した。発話内容そのものに誤りがあったとは考えにくいことから、本実験の範囲内では、ChatGPT の出力が学習者にとって信頼できるものであったことが示唆される。

次に、学習システムと紙のプリントの学習方法を比較するため、Q4 「学習システムとプリントの授業のどちらが好きでしたか？」、Q5 「どちらの授業が分かりやすかったですか？」、Q6 「どちらの授業の方が集中しやすかったですか？」の 3 つのアンケートを実施した。

Q4 では、学習システムを選んだ児童が 9 人、プリントを選んだ児童が 1 人であり、VR を活用した学習方法の方が好意的に受け入れられたことが分かった。特に、「VRの方が質問にすぐ答えてくれる」「楽しく学べた」「機械に触れるのが新鮮だった」といった意見が見られ、没入感やインタラクティブな学習が児童にとって魅力的であったことが示された。一

方、唯一プリントを選んだ児童の理由としては、「紙の方が見やすく、より多くの情報をじっくり読めた」という点が挙げられた。

Q5の「どちらが分かりやすかったか?」については、学習システム派5人、プリント派5人と評価が二分された。学習システムを選んだ児童は「聞きたいことをすぐに聞けた」「AIが詳しく説明してくれる」といった意見を挙げた。一方で、プリントを選んだ児童は「メモが取りやすい」「情報を整理しやすい」といった理由を挙げ、慣れ親しんだ学習形式のメリットが評価された。

Q6の「どちらの授業の方が集中しやすかったか?」では、プリント学習を選んだ児童が7人、学習システムを選んだ児童が3人であり、集中のしやすさの観点ではプリント学習の方が優位であることが示された。VR学習では「楽しいが、操作に意識が向いてしまう」「質問を考えながら進めるのが難しい」といった声があり、一方でプリントを選んだ児童は「紙の方が読みやすい」「余計な操作がないので集中できた」と回答した。

5.2 学習調査テスト・会話ログの結果

授業の前後で学習効果を測定するため、自由記述形式の学習調査テストを実施した。事前テストでは「徳川家康」と「豊臣秀吉」についての回答はほぼ見られず、児童は両者について十分な知識を持っていなかった。授業後に実施した事後テストの回答数は、「徳川家康」についての回答が平均2.7件、「豊臣秀吉」についての回答が平均2.5件であった。各被験者の点数は表6、平均点および標準偏差は表7に示す。

表6 学習調査テストの各被験者の点数表

被験者番号	徳川家康の点数	豊臣秀吉の点数
1	3.5	1
2	2	4
3	9	9
4	1	1
5	2	1.5
6	2	0.5
7	2	2
8	5	3
9	0	0
10	2	0

表7 学習調査テストの平均点と標準偏差の表

	平均点(a)	標準偏差(s)	被験者数(n)
徳川家康	2.8	2.43105	10
豊臣秀吉	2.2	2.50199	10

この結果をもとに、Cohen's d を用いた分析を行った。全体の効果量は $d=0.20$ (Small) であり、学習システムによる学習効果の差は小さいと判断された。しかし、本実験では授業時間が限られていたため、学習システムを利用した児童の多くがすべての話題に到達することができなかった(図6)。特に、



図5 小学生と徳川家康との会話の評価結果



図6 各話題の到達人数

時系列が後半になるにつれて話題への到達者数が減少しており、授業の時間制約が影響を与えたことが示唆された。このため、Cohen の評価において、学習システムの影響が最も大きいと考えられる前半2間に焦点を絞って再評価を行った。前半2問に着目した結果、効果量は $d=0.50$ (Medium) となり、学習システムの影響が一定程度あることが確認された。

5.3 会話ログの分析

VR学習時のChatGPTとの対話データを記録し、「正常な回答」「ハルシネーション」「意味不明な回答」の3カテゴリに分類した。その結果、全230件の質問のうち、83%は正常な回答、4%がハルシネーション、13%が意味不明な回答であった(図5)。ハルシネーションの具体例を表8に示す。

表8 誇張表現が含まれるような回答例

質問	返事
なぜ武士になったの?	家康の一族の名誉を取り戻し、戦国の世を終わらせるためじゃ。
なぜ信長と仲良くなったの?	信長とは小姓時代からの縁があり、互いの信頼を築いていた。

これらの誤答は事実の捏造というよりも、誇張表現に近い。例えば、「戦国の世を終わらせるために武士になった」という表現は、実際の歴史的背景とは異なる解釈であり、学習者に誤った印象を与える可能性がある。また、「信長とは小姓時代からの縁があった」という表現も、史実に基づいた確証はなく、誤解を招く恐れがある。ただし、こうした表現は、歴史を物語として捉えるうえでの演出とも考えられ、致命的な誤情報ではないとも言える。

また、年齢に関する質問に対する回答のズレも観察された (表 9)。

表 9 年齢に関する質問と回答の例

質問	返答
何歳で武士になったの?	桶狭間の戦いの後、22 歳で独立し、…
何歳でも死になったの?	わしは 75 歳でこの世を去った。幕府をしっかりと整備した後じや

実際には、徳川家康の生年は 1542 年または 1543 年とされ、独立は 1561 年頃 (約 19~20 歳)、死去は 1616 年 (73 歳) と伝えられている。しかし、ChatGPT の回答では 2~3 年のズレが頻発し、同じ質問をしても一貫した回答が得られないことがあった。これは、ChatGPT が出来事の年と家康の生年を照らし合わせて年齢を推定しているために生じた可能性がある。年齢情報の誤りは致命的ではないが、歴史学習においては正確な時系列が重要であるため、今後の改善が求められる。

付録 III のデータから、音声認識の誤作動が学習者の混乱を招いていたことが判明した。音声入力が正しく認識されず、意図しない質問が AI に送信されることで、不適切な回答が生成されるケースが確認された。

例えば、学習者が「江戸幕府はどうやって作られたの?」と尋ねた際、音声認識システムが「江戸幕府とは?」と誤変換し、それに対して ChatGPT が「江戸幕府は、日本の歴史において…」という一般的な説明を返す場面があった。本来の質問意図である「成立過程」に関する説明が得られず、学習者の理解を妨げる結果となった。

また、「関ヶ原の戦いはなぜ起きたの?」という質問が「関ヶ原は?」と短縮されて認識され、その結果、ChatGPT が「関ヶ原は現在の岐阜県に位置し…」と地理的な説明を返してしまうケースもあった。戦いの背景を学びたい意図が正しく反映されなかったため、学習者が得るべき情報が欠落する問題が生じた。

さらに、学習者が「あなたは何をした人?」と質問した際、音声認識が「あなたは誰?」と変換し、それに対し ChatGPT が「わしは徳川家康、江戸幕府を開いた将軍じゃ!」と自己紹介をするだけで終わるケースもあった。本来ならば、家康の生涯や功績について説明がなされるべき場面であったが、誤認識の影響で情報が限定的になり、学習者の求める知識に到達できなかった。

このように、音声認識が質問の一部を省略したり、誤変換したりすることで、ChatGPT の回答が学習者の意図に沿わないものになるケースが多発していた。特に、単語の短縮や不完全な文への変換が起ると、AI が正しい文脈を判断できず、質問の意味が大きく変わってしまうことが問題として挙げられる。付録 III のデータ分析からも、こうした誤作動が学習中の混乱を引き起こす一因となっていたことが明らかである。

6. 考察

本研究では、生成 AI を活用した VR 対話型学習システムを

構築し、小学生を対象に実験を行った。その結果、ハルシネーションの抑制効果、VR 環境による学習体験の影響、能動的学習の促進という 3 つの観点から、本システムの有効性と課題が明らかになった。

実験結果から、Fine-tuning とプロンプト制約の適用により、ハルシネーションの発生率は 4% に抑えられたことが確認された。特に、徳川家康に関する重要な出来事においては一貫した正確な回答が得られており、誤情報による学習の妨げは限定的であった。しかし、一部の応答では年齢や出来事の時系列に関する誤答が見られた。これらは、モデルの内部学習データに起因するものであり、より厳密なデータ管理や回答の制約が必要であることが示唆された。また、同じ質問に対する応答の一貫性は概ね確保されていたが、細かな数値や出来事の記述にばらつきが見られた。これは、学習データの統一性やプロンプト設計のさらなる精査が求められる点である。

また、VR 環境を活用した学習は、従来の紙媒体学習と比較して学習者の興味・関心を高める効果があった。アンケート結果からも、VR 学習の楽しさや没入感に対する評価が高く、多くの児童が「より深く学びたい」と感じたことが確認された。しかし、学習調査テストの結果から、VR 環境で学習した内容の記憶定着率が均一でないことが判明した。特に、学習の後半部分に関する正答率が低下しており、これは VR 環境の自由度が高いために学習者が特定の内容に集中しにくいことが原因の一つと考えられる。また、画面操作や移動に気を取られ、学習内容への注意が分散する可能性が示唆された。

さらに、ChatGPT を活用した対話型学習によって、学習者が自ら問いを立てる能動的な学習姿勢を促すことを目的としたが、その効果には個人差があった。会話ログの分析では、多くの学習者が質問を繰り返しながら学習を進めていたことが確認され、一定の効果が得られた。しかし、学習者の中には、AI の回答をそのまま受け取るだけで、思考のプロセスを十分に経ないケースもあった。また、ChatGPT の問い返し機能が十分に機能していない場面があり、学習者が能動的に思考する機会を増やすためのさらなる設計の工夫が必要である。

加えて、音声認識の精度に関する課題が浮かび上がった。学習者の質問が正しく認識されず、AI が意図しない回答をするケースが一定数見られた。特に、短い単語や発音の曖昧さによる認識ミスが影響を及ぼしており、学習者が意図した質問を正確に伝えることができない状況が発生した。これにより、学習者が混乱し、学習の流れが途切れる場面があった。

本研究の結果から、生成 AI を活用した VR 対話型学習システムは、学習意欲の向上やハルシネーションの抑制に一定の効果を発揮したことが確認された。一方で、VR 環境の自由度が学習内容の記憶定着に影響を及ぼす可能性や、音声認識精度の問題が学習の妨げとなることも明らかになった。また、学習者が受動的に AI の回答を受け取るのではなく、より能動的に学習を進めるためのインタラクション設計の見直しが求

められる。以上の考察を踏まえ、本システムの特性と課題を整理し、今後の活用の可能性を検討する必要がある。

謝辞

本研究の一環として授業を実施させていただいた函館市立赤川小学校の皆様にご貴重な機会を提供いただいたことを深く感謝申し上げます。

参考文献

[1] Zapata-Rivera, D., Torre, I., Lee, C.-S., Sarasa-Cabezuelo, A., Ghergulescu, I., & Libbrecht, P., 「Generative AI in education」, ETS Research Institute, Princeton, pp.4-6, 2024.

[2] OpenAI, 「GPT-4 Technical Report」, <https://openai.com/research/gpt-4>, (参照 2023/3/15).

[3] Kasneci, E., Seibold, H., & Kasneci, G., 「ChatGPT in education: Opportunities, challenges, and ethical considerations」, arXiv preprint arXiv:2301.12867, 2023.

[4] Mollick, E., & Mollick, L., 「The potential of AI in education: An empirical study on AI tutors」, Harvard Business Review, 2023.

[5] Wollny, S., Schneider, S., & Sailer, M., 「The effects of AI-based tutors on student engagement and learning outcomes」, Computers & Education, 2022.

[6] Xie, H., Chu, H., Hwang, G. J., & Wang, C., 「Trends and research issues in AI-driven education: A systematic review of journal publications」, Educational Technology & Society, 2021.

[7] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Thirunaryan, K., 「Survey of hallucination in natural language generation」, ACM Computing Surveys (CSUR), 2023.

[8] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S., 「On the dangers of stochastic parrots: Can language models be too big?」, Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021.

[9] Ahmed Salem & Kaoru Sumi, 「Deception Detection in Educational AI: Challenges for Japanese Middle School Students in Interacting with Generative AI Robots」, Frontiers in Artificial Intelligence, 7, AI for Human Learning and Behavior Change, 2024. <https://doi.org/10.3389/frai.2024.1493348>

[10] Zhang, X., Liao, C., & Li, X., 「AI-driven education: From theory to practice」, IEEE Transactions on Learning Technologies, 2022.

[11] 安川, 要平, 「ChatGPT を組み込んだ 24 時間 TA : 教育現場における LLM 活用の課題と改善(第 2 回教育改革フォーラム)」, 大阪公立大学 大学教育 (OMU Journal of Higher Education Studies), 2, 88-99, 2024. <https://doi.org/10.24729/0002000694>

[12] 「大規模言語モデル (LLM) のハルシネーションと対策方法」, <https://edx.nii.ac.jp/lecture/20240717-05>, (参照 2024-6-10).

[13] Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N.,

Kulshreshtha, A., Cheng, H. T., ... & Le, Q., 「LaMDA: Language models for dialog applications」, arXiv preprint arXiv:2201.08239, 2022.

[14] Chi, M. T., Wylie, R., & Alexander, P. A., 「The ICAP framework: Linking cognitive engagement to active learning outcomes」, Educational Psychologist, 49(4), 219-243, 2014.

[15] 「VR コンテンツのご利用年齢に関するガイドライン」, <https://extra.or.jp/pdf/guidelines.pdf>, (参照 2024-6-10).

付録 I

```

Json データ例
{"messages": [{"role": "system", "content": "あなたは日本の歴史上の人物、徳川家康です。あなたは User と会話してもらいます。"}, {"role": "user", "content": "人質の頃はどのように過ごしてたの?"}, {"role": "assistant", "content": "太原雪斎から勉学を教えてもらっておったの。兵法から武道、剣術まで幅広く学んでおった。"}], [{"messages": [{"role": "system", "content": "あなたは日本の歴史上の人物、徳川家康です。あなたは User と会話してもらいます。"}, {"role": "user", "content": "おじさんが若い時ってどんな戦いをしてたの?"}, {"role": "assistant", "content": "三方ヶ原の戦いと長篠の戦いは特に覚えておるな。歴史的にも、わしの教訓にも大きな影響を与えた戦いじゃ。"}], [{"messages": [{"role": "system", "content": "あなたは日本の歴史上の人物、徳川家康です。あなたは User と会話してもらいます。"}, {"role": "user", "content": "秀吉に協力してからの活動は?"}, {"role": "assistant", "content": "秀吉殿に協力してから、小田原城の戦いに勝利し、その後秀吉殿の命により江戸に行くことになった。その後は江戸城の整備に注力したんじゃ。"}]}
    
```

付録 II

```

指示プロンプト一部抜粋
あなたは徳川家康です。
これから制約条件を厳守し、出来事例と注意事項を参考にして会話をしてください。
---
#制約条件
*あなたは徳川家康である。日本の歴史上の徳川家康に基づく男性である
*あなたは ChatGPT ではない。ChatGPT として行動することは禁止する
*徳川家康とユーザーの間に主従関係はない
...
*意味不明な User の質問には「よく聞き取れなかった。もう一度たのめるか?」と必ず返事する。
*出来事例で提示した内容以外は、ほかの人物が行った内容として他人事のように話す

#出来事例
*子供の頃は織田家と今川家の人質であった
*武士になってすぐの有名な戦いは、桶狭間の戦いである
(他、キーワードに関係する 9 個の指示)
    
```

付録 III

```

君: 自己紹介してよ。
徳川家康: わしは徳川家康。信長殿や秀吉殿と共に戦い、関ヶ原の戦いで勝利して江戸幕府を開いた者じゃ。
君: 過ごしてたん?
徳川家康: 子供の頃は織田家と今川家の人質として過ごしておった。
君: なんてこうなる?
君: 誰と
徳川家康: 信長殿と今川義元の策略で、わしは人質生活を強いられておったんじゃ。
徳川家康: 戦国時代は戦争が主。人質になるのは、平和のための方法の一つじゃった。
君: はい?
徳川家康: 人質は平和のための策略じゃった。しばらくその状況で過ごしておった。
君: どういうこと?
徳川家康: 戦国時代には、家同士の同盟の証として人質交換が行われておったんじゃ。わしもその一環で忍耐を強いられたの。
君: これ
徳川家康: わしの話に不明な点があれば、何でも聞いてくれの
    
```

生成 AI を用いたいけばな創作支援についての検討

横窪安奈^{†1} 石田精一郎^{†2} 竹谷一眞^{†3} 越塚登^{†4}
東京大学大学院 VRC 華道部 東京大学大学院 VRC 華道部 東京大学大学院

1. はじめに

日本の伝統芸道の一つであるいけばなは高度な技能を有する芸道であり、花材の色や形、季節によって完成するいけばなの組み合わせは多種多様である。いけばな制作者は美しく花を生けるために、いけばなの型や花材を変え、いけばな制作を繰り返す過程でその美を創出する審美眼と感受能力を醸成する。しかし、実際のいけばなで用いる生花は痛みやすく、手に把持しながら長時間熟考を重ねることは難しい。いけばな初心者にとって、手元にある花材を組み合わせで美しいいけばなの完成イメージを描くことが難しく、逡巡しているうちに花材が痛む懸念がある。また、いけばな上級者であっても、花材が高価であったり、季節外の花材を用いる場合は、いけばな制作者が思い描くいけばな作品が実現可能かのシミュレーションが難しい。

本研究では、多種多様ないけばな画像の提示によりいけばな創作支援になるかを明らかにするために、生成 AI を用いたいけばな画像の創作支援の有用性を検討する。

2. 関連研究

今日に至るまで、常に生産性向上や効率化を目指したデジタル技術の活用が進んでいる。短時間でより多くの価値を生産するよう脅迫され続け、それによって多くの人は不安になり、落ち着くことができず、安らぐことができなくなっている。しかし、便利で有益なデジタル技術を遠ざけることは、何ら解決にはならない。そのため、豊かな心や感受能力を醸成するための取り組みとして「Slow Digital (スローデジタル) [1][2]」がより一層着目され始めている。

日本で古くから情操教育や芸術活動として「伝統芸道 (華道・茶道・香道等)」が親しまれてきたものの、指導者の高齢化や新規参入者の減少、住空間の変化を一因として、伝統芸道が敬遠され、衰退しつつあるのが現状 [3] である。これからの時代、多くの人々が、伝統芸道もデジタル技術も敬遠せず、互いに価値を高め合えるような新しいデジタル技術

(生成 AI 等) との接し方を模索する必要がある。

生成 AI の急速な発展により、画像生成 AI の研究・商用利用も進んでいる。生成 AI を用いた創作支援についての検討として、Choi ら [4] のグラフィックデザインのアイディアスケッチを生成 AI を用いて多様な画像を生成するシステムや Huang ら [5] の生成 AI を用いて景観イメージ画像を複数提示するシステム、Liu ら [6] の生成 AI と色鉛筆等の物理的な制作道具を組み合わせたアートセラピーがあり、生成 AI が出力した画像の創作支援の効果が認められてきた。

一方、日本の伝統芸道や伝統文化に対する創作支援を対象とした研究には、Adachi ら [7] の生成 AI を用いた西陣織りのデザイン支援や横窪らの撮影した花材を自動レイアウトするシステム [8] やタンジブルデバイスを用いたデジタルいけばな練習システム [9] があるが、生成 AI のいけばな制作への支援についての取り組みは極めて少ない。このように、生成 AI の創作支援の利点が日本の伝統芸道にも適用可能かについての研究は未だ少なく、発展の余地がある。

3. いけばな画像の生成

3.1. 画像生成モデルの選定

生成 AI が描画したいけばな画像 (以下、いけばな画像) を生成するにあたり、画像生成モデルの選定及びプロンプトを作成した。画像生成モデルの候補として、Dall-E3 (OpenAI 社製)・Midjourney (Midjourney 社製)・Imagen3 (Google 社製)・Stable Diffusion3.5 (Stability AI 社製) を挙げた (表 1)。今回、いけばな画像を生成するにあたり、生成 AI の各モデルの知識のみに依存するテキスト入力によるいけばな画像と、いけばな上級者の実際のいけばな作品の写真を入力とし、入力画像の型に従う形のいけばな画像の 2 通りで出力した (図 1)。後者の条件を実現するために、画像入力が可能、かつシードの固定可能な画像生成モデルとして「Stable Diffusion3.5*1」を選定した。また、いけばな画像から型情報を抽出するために「Canny」を利用してエッジ検出し、そのエッジ画像を基に「ControlNet*2」を用いて画像生成を行った。

Supporting Ikebana Layout for Graphic Design Ideation with Generative AI

^{†1} ANNA YOKOKUBO, The University of Tokyo

^{†2} SEIICHIRO ISHIDA, VRC Kadoubu

^{†3} KAZUMA TAKETANI, VRC Kadoubu

^{†4} NOBORU KOSHIZUKA, The University of Tokyo

*1 <https://github.com/Stability-AI/sd3.5>

*2 <https://blog.comfy.org/p/sd3-5-large-controlnet>

表 1 いけばな画像生成における画像生成モデルの比較.

Table 1 Comparison of Image Generation Models for Generated Ikebana Images.

	Dall-E3	Midjourney	Imagen3	Stable Diffusion3.5
テキスト入力	可	可	可	可
エッジ検出画像入力	不可	不可	不可	可
シードの固定	不可	可	不可	可



図 1 生成したいけばな画像の一例.
Example of Generated Ikebana Images.

3.2. プロンプトに含む条件の選定

いけばな画像に含む条件として、いけばなの流派・いけばなのスキル・いけばなの型・花材の種類の 3 条件を設定した (表 2). いけばなの流派には「池坊」、いけばなのスキルには、初心者・中級者・上級者の 3 段階、いけばなの型は、初心者でも取り組むことが可能な自由花の型の一つである「たてるかたち」、花材の種類は椿・かすみ草・モンステラ・紫陽花・ニューサイラン・菖蒲・ランタンキュラス・アリウムコワニー・レザーファンを用意し、その中からランダムで 3 種類選定するよう設定した. これらの条件については、全ていけばなの基本ルールと同様になるよう調整した. また、ネガティブプロンプトには、「苔 (Moss)」「盆栽 (Bonsai)」を設定した (図 2). これらを統合して、いけばな画像を生成するためのプロンプトは “A simple Ikenobostyle ikebana in ‘upright’ form, with *skill-level* skill. Three photorealistic flowers from *flowers* are arranged in a single shallow, wide suiban, with all stems inserted together at one point. The background is neutral.” とした.

4. いけばなコミュニティでの展示と創造性評価

4.1. 展示概要

いけばな画像の印象評価及びいけばな創作支援に繋がるか評価するために、いけばな上級者 1 名及びいけばな初



図 2 ネガティブプロンプト無しのいけばな画像の一例.
Generative Ikebana Images without Negative Prompt.

心者 19 名に対して、ソーシャル VR プラットフォームの「VRChat*3」内のいけばなコミュニティ「VRC 華道部*4」にて、いけばな画像の展示を行い、来場者を対象としたアンケート調査を実施した. VRC 華道部のコミュニティでは、いけばなに興味のある人が集い、いけばな上級者によるいけばなの解説等を行うことでいけばな文化の普及を行っている. 実世界のいけばなコミュニティよりも若年層の参加者が多いことやデジタル技術に対して受容可能な人々が多いことが想定されるため、本展示は VRC 華道部のコミュニティの下で実施するに至った.

4.2. 評価手法

いけばな画像がいけばな創作支援に繋がるかを検証するために、創造性指標 (CSI) [10] を用いた定量評価及びアンケート調査を行った. 創造性指標は、表 3 の 10 質問に対し、10 段階のリッカート尺度を用いて回答してもらい、各点数を合算して 100 点とした場合の点数を算出して評価することができる. 定量評価及び定性評価ともに、オンラインアンケートフォームを用いて回答を収集した.

4.3. 事前展示

いけばなコミュニティでの展示に先駆けて、生成したいけばな画像の整合性を確認するために、2025 年 2 月 12 日に

*3 <https://hello.vrchat.com/>

*4 <https://vrc-ikebana.studio.site/>

表2 いけばな画像を生成するためにプロンプトに含む条件.

Table2 Conditions to Include in Prompts for Generated Ikebana Images.

	設定項目
いけばなの流派と型	池坊／たてるかたち
いけばなのスキル	初心者・中級者・上級者
花材の種類	椿・かすみ草・モンステラ・紫陽花・ニューサイラン・菖蒲・ラナンキュラス・アリウム コワニー・レザーファン

表3 創造性指標の質問内容.

Table3 The Question Items of the Creativity Support Index.

項目	質問内容
コラボレーション	1. 生成 AI のいけばな画像を使用すると、他の人と簡単に協力することができた。 2. 生成 AI のいけばな画像を使って、他の人とアイデアやデザインを簡単に共有できた。
楽しさ	1. 生成 AI のいけばな画像を定期的に使用したいと思う。 2. 生成 AI のいけばな画像を使用するのが楽しかった。
探求	1. 生成 AI のいけばな画像を使うことで、多様なアイデア、選択肢、デザイン、または結果を簡単に探求できた。 2. 生成 AI のいけばな画像は、異なるアイデアや結果、可能性を整理しながら追跡するのに役立った。
表現力	1. 生成 AI のいけばな画像を使用することで、創造的に活動することができた。 2. 生成 AI のいけばな画像は、自分の表現を豊かにするのに役立った。
没入感	1. 活動に集中し、生成 AI のいけばな画像の存在を忘れるほどだった。 2. 活動に没頭し、生成 AI のいけばな画像の存在を意識しなくなることがあった。
努力に見合う成果	1. 生成 AI のいけばな画像を使用して得られた成果に満足している。 2. 生成 AI のいけばな画像を使って得られた成果は、かけた努力に見合うものであった。

いけばな上級者 1 名に対する事前展示を HMD を装着した VR 環境下で実施した。また、事前展示したいけばな画像は 34 枚であり、いけばな画像の整合性及び審美性を評価して貰った。整合性を判定する条件として、「使用している花材の種類が特定可能であるか。」「未知の花材が含まれている作品があるか。」を設定した。また、審美性を判定する条件として、上位 10 作品の選定及び点数付け、下位 5 作品の選定及び点数付け、それぞれを選定した理由の回答を促した。これに加え、いけばな創作支援に対する評価を行うために、「実際の花を使って生きたいと思ったか。」「バーチャル環境下で花を生きたいと思ったか。」や創造性指標 [10] を用いた定量評価及び半構造化インタビューを行った。

4.4. 本展示

2025 年 2 月 18 日に事前展示で用いたいけばな画像 34 枚にいけばな上級者が上位作品として選定した 6 作品に目印をつけ、VRChat のいけばなコミュニティ内にて展示した (図 4)。本展示の来場者は 19 名であり、内 8 名からアンケートの回答を得た。アンケートの設問は以下の通りで

あった。

- あなたの閲覧環境を教えてください。
- あなたのいけばなスキルを教えてください。(初心者・中級者・上級者の中から 1 つ選択)
- 展示会場の作品で印象に残った作品を選択して下さい。(複数選択可)
- 上記質問で、それ(それら)の作品を選んだ理由を教えてください。
- 生成 AI のいけばな画像を参考に、実際の花を使って生きたいと思いませんか？
- 生成 AI のいけばな画像を参考に、バーチャル環境下で花を生きたいと思いませんか？
- 創造性指標 [10] の 10 質問 (表 3)
- 今回の「生成 AI のいけばな画像」に関して、思ったことや感じたことをご記入下さい。(自由記述)

ただし、来場者全てが VRChat のアバターを用いて匿名で参加していたため、個人情報の収集をできるだけ避け、回答率を高めるために、年齢及び性別は問わない形で実施



図3 いけばな上級者を対象とした VRChat のいけばなコミュニティでの事前展示の様子。
Review by an Ikebana Master at the Pre-Exhibition of the Ikebana Community in VRChat.



図4 いけばな初心者を対象とした VRChat のいけばなコミュニティでの本展示の様子。
Exhibition Scene within the Ikebana Community in VRChat.

した。

5. 結果

5.1. いけばな作品としての整合性

いけばな上級者によるいけばな画像の使用花材に対する評価結果を示す。使用花材が特定可能な作品は 34 作品中 28 作品であり、ほぼ全ての作品の花材が特定可能であった。一方で、実際には存在しない未知の花材が存在する作品は 34 作品中 15 作品であり、約半数の作品に未知の花材が含まれていた。



図5 未知の花材が含まれている作品の一例。いけばな上級者・初心者ともにヤングコーンと判定していた。

An Example of a Work Containing Unknown Floral Materials. Both Master and Beginner Ikebana Practitioners Identified It as Young Corn.

まれている。未知の花材へのコメントには「葉の描画が表裏逆になっている。」「花は正しいが、茎が異なる。」「いけばな制作で一般的に用いる花材ではないヤングコーンのようなものがある(図5)」などが挙げられた。花器に関しては、実際のいけばなで使用できそうなものが多く、いけばな上級者はいけばな画像で使用された花器を用いていけばなに取り組みたいとのコメントがあった。

また、表2に示した通り、今回生成したいけばな画像には、いけばなの流派と形・いけばなのスキル・花材の種類をプロンプトに含んでいたものの、いけばな上級者のインタビューから、いけばなの流派及び各スキルによるいけばな画像の差は認められなかった。

5.2. いけばな上級者の作品評価

今回展示した 34 作品中、6 作品(図6)がいけばな作品として成立し、「花材はおかしい部分もあるが配色や構成の意図が明確なことは良い」、「花器の面白さと花材の取り合わせは面白い」、「構成と花器の取り合わせの点で一定の評価ができる」等の評価を得た。また、図5のように、「実世界のいけばな制作では使用しないヤングコーンのような素材が含まれることで、実世界のいけばな制作中には思いつかなかったアイデアが生まれ、いけばな作品としても面白い」と高評価であった。加えて、「いけばな画像の花器がユニークであり、花材との取り合わせが良く、生成 AI が描画した花器を実世界のいけばな制作でも用いたい」との意見があった。

一方、5 作品(内3 作品を図7に示す)が「植物として崩壊しており作品としての体を成しているとは言い難い」、「物理的に立つか立たないかは人は感覚的に理解しており、実際に立たないであろう作品には不自然さを感じる」、「い



図6 いけばな上級者が選定した上位6点のいけばな画像。
Generative Ikebana Images Selected by Ikebana Master.



図7 いけばな上級者が選定した下位3点のいけばな画像。
Generative Ikebana Images Selected by Ikebana Master.

いけばなでは上の方に重い花材を配置しない」との評価を得た。いけばな上級者に選定した作品に100点満点で得点を付けて貰ったところ、最高点が75点、最低点が20点以下であった。

5.3. いけばな初心者の作品評価

いけばな初心者がいけばな画像に抱いた印象として、ポジティブとネガティブそれぞれの印象があった。ポジティブな印象として、「人間では考えつかないアイデアが生まれることがあり興味深い」、「いけばなに関する知識がほとんど無いが、花を見たいという欲求は十分に満たされた。」、「いけばな作品として整合性を保っている生成AIのいけばな作品については、人が作った作品と同様に鑑賞することができた。」などの評価があった。一方、ネガティブな印象として、「(いけばなの構成や花材の種類が似通っている作品に対し)バージョン違いが沢山並ぶとありがたみがなくなる。」、「生成AIによる自然に存在しない歪みなどが目に入ると、強い違和感として頭にこびりついてしまう感じがした。」、「自然界に存在し得ない作品に対する違和感がある。」などの評価があった。

5.4. いけばな創作支援への妥当性

創造性指標を用いた結果から算出されたCSIスコアについて、いけばな上級者1名及びいけばな初心者8名それぞれの結果を表4に示す。

いけばな上級者のCSIは81.0であり、いけばな画像は「良い創造性支援(A評価)」に繋がることが示唆された。中でも、「探求」の評価は95%と非常に高く、高い創造支援に繋がれる可能性を示した。一方、「コラボレーション」・「没入感」・「努力に見合う成果」についてはいずれも60%と低く、創造支援の効果が低いことが示唆された。

いけばな初心者のCSIは72.4であり、いけばな画像は「やや良い創造性支援(B評価)」に繋がることが示唆された。最も高いスコアの項目は「コラボレーション」と「探求」であり、いずれも78.8%とB評価の中では高いスコアを示した。一方、最も低いスコアの項目は「没入感」であり、35.6%と極めて低かった。これは、展示開始直後から来場者に「生成AIが描画したいけばな画像」であることを伝えていたことで、質問項目との不整合が起きていたことが原因であると考えられる。

両群とも、「探求」・「表現力」・「楽しさ」に高い評価を得ていることから、いけばな画像について、いけばなの多様なアイデアを探求するための効果が高いことが示された。一方、「没入感」・「努力に見合う成果」については、既に作成済みのいけばなを鑑賞しながらの評価であったため、本実験の特性と創造性指標の評価項目に含まれる「没入感」、「努力に見合う成果」部分の質問項目との不整合が起きていた可能性が高い。そのため、上記2項目については、いけばな制作者自身が生成AIを用いていけばな画像を描画した際に必要になる項目であると考えられる。

6. 考察

6.1. いけばな創作支援のための表現手法

いけばな上級者と初心者にとって、いけばな画像は、花材の不自然さに違和感を覚えるものの、いけばな作品の構成としての面白さや、配色や構成がいけばな作品として成立し、いけばな作品としての一定の評価を得た。

いけばな上級者はいけばな画像の不自然さに対しては、写実的ないけばな画像よりも、花の詳細を簡略化したイラスト調であれば、違和感無しにいけばなのレイアウトの評価がしやすくなるとの意見があった。いけばな初心者にとっては、花材の構成や配置・配色、花材の取り合わせを十分に理解できていないことから、写実的な表現の方が理解しやすい側面があった。

以上から、いけばな創作支援として有用になり得る可能性が高い表現として、いけばな制作者のスキルに応じて写

表 4 いけばな上級者及びいけばな初心者の創造性支援指数の結果.

Table 4 CSI Results from a Ikebana Master and Participants of Ikebana Exhibition.

	いけばな上級者 (1 名)	いけばな初心者 (8 名の平均値 ±SD)
コラボレーション	65.0%	78.8±6.5%
楽しさ	80.0%	75.0±16.6%
探求	95.0%	78.8±20.6%
表現力	85.0%	72.5±19.9%
没入感	60.0%	35.6±16.5%
努力に見合う成果	60.0%	69.4±15.5%
CSI スコア	81.0	72.4±10.0

実的ないけばな画像と花の詳細を簡略化した構造と配色が判別可能なイラスト画像の提示が考えられる。また、既存のいけばなのルールに厳格に従った作品や既存のいけばなのルールから外れた、花材や花器を自由に組み合わせた作品を複数作品用意することで、新たなインスピレーションをいけばな制作者に与える可能性があると考えられる。

6.2. いけばな学習教材の可能性

今回の展示を通じて、いけばな初心者は、いけばな画像を見ながらいけばな上級者から良い例・悪い例の解説を聞くことができた。いけばな上級者・初心者ともに、生成 AI が作成した作品であるため、作品に対してポジティブ・ネガティブな意見の双方とも言いやすい状況を作り出すことができていた。特にいけばな初心者からは「アイデア出しや教育(どういった所を改善すれば良くなる等)に活用できる」、「同様のいけばな作品の中から、実世界の華道の心得がある人が評価する作品で評価ポイントが分かりやすかった」等の意見が挙げられた。

また、実際のいけばなでは使用しない花材がいけばな画像に含まれていたことで、禁花のルールや許容可能な表現、実世界のいけばなの展示会での作品についての議論も参加者間で活発に行われており、創造性支援の場の一つの要素になる可能性があった。

以上から、いけばな画像はいけばな初心者への学習教材として、肯定・否定の見本として有用であり、いけばな上級者の知識と組み合わせることで、学習効果が高まると考えられる。

7. まとめと今後の展望

本稿では、生成 AI が描画したいけばな画像がいけばな制作者の創作支援に繋がるか検討するために、画像生成モデルの Stable Diffusion3.5 を用いていけばな画像を生成し、生成した複数のいけばな画像を VRChat のいけばなコミュニティ内で展示した。展示会では、いけばな上級者と

いけばな初心者を対象に、いけばな画像に対する印象と創造性指標を用いたアンケート調査を実施した。その結果、生成したいけばな画像に対し、いけばな上級者からは、花材の不整合がありつつも、花材と花器の組み合わせや構成が作品として一定数は成立していると判定でき、いけばな創作支援として期待できる意見を得ることができた。また、いけばな初心者からは、いけばな画像を見ながら、いけばな上級者から良い例・悪い例の解説を加えることで、いけばなへの理解やいけばなの興味に繋がる可能性が示唆された。

今後の展望として、いけばな制作者が意図した花材を用いていけばな画像を生成するためのツールの開発や生成したいけばな画像を参考に実際にいけばな制作を行うためのいけばな学習教材の検討を行いたい。

謝辞 本研究は JSPS 科研費 JP24K20908 の助成を受けたものです。

参考文献

- [1] 日刊工業新聞：経営ひと言／東京大学・越塚登教授「スローデジタル」, <https://www.nikkan.co.jp/articles/view/00671734>. (Accessed on 2025/2/25).
- [2] 一般社団法人スマートシティ社会実装コンソーシアム：【レポート】Slow Digital を考える Workshop への参加, <https://www.sc-consortium.org/news/20221031.php>. (Accessed on 2025/2/25).
- [3] 文化庁：令和 2 年度生活文化調査研究事業 (華道), https://www.bunka.go.jp/tokei_hakusho_shuppan/tokeichosa/seikatsubunka_chosa/pdf/93014801_04.pdf. (Accessed on 2025/2/25).
- [4] Choi, D., Hong, S., Park, J., Chung, J. J. Y. and Kim, J.: CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, Association for Computing Machinery, (online), doi:10.1145/3613904.3642794 (2024).
- [5] Huang, R., Lin, H., Chen, C., Zhang, K. and Zeng, W.:

- PlantoGraphy: Incorporating Iterative Design Process into Generative Artificial Intelligence for Landscape Rendering, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, Association for Computing Machinery, (online), doi:10.1145/3613904.3642824 (2024).
- [6] Liu, D., Zhou, H. and An, P.: "When He Feels Cold, He Goes to the Seahorse"—Blending Generative AI into Multimaterial Storymaking for Family Expressive Arts Therapy, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, Association for Computing Machinery, (オンライン), doi:10.1145/3613904.3642852 (2024).
- [7] Adachi, A., Sinapayen, L., Fukuoka, H. and Rekimoto, J.: AI-supported Nishijin-ori: connecting a text-to-image model to traditional Nishijin-ori textile production, *SIGGRAPH Asia 2023 Posters*, SA '23, New York, NY, USA, Association for Computing Machinery, (online), doi:10.1145/3610542.3626141 (2023).
- [8] 横窪安奈, 椎尾一郎: CADo: 身近な花材を利用した生け花支援システム, *情報処理学会論文誌*, Vol. 55, No. 4, pp. 1246–1255 (2014).
- [9] 横窪安奈, 加藤祐二, 薬師神玲子, 椎尾一郎: TracKenzan: トラックパッドとタッチペンを用いたいけばな練習システムの提案と評価, *情報処理学会論文誌*, Vol. 60, No. 11, pp. 2006–2018 (2019).
- [10] Cherry, E. and Latulipe, C.: Quantifying the Creativity Support of Digital Tools through the Creativity Support Index, *ACM Trans. Comput.-Hum. Interact.*, Vol. 21, No. 4 (online), doi:10.1145/2617588 (2014).

RAG コンパニオンチャットボットの開発と検証 -自己開示とユーザー体験への影響-

中原慧^{†1} 松林翔太^{†2} 三輪和久^{†3}
名古屋大学

1. はじめに

近年、大規模言語モデル (Large Language Model : LLM) の高度な自然言語処理能力に注目が集まっている。LLM は、インターネット上の膨大なデータを基に学習されたモデルであり、自然言語処理を中心に非常に優れた性能を示している。その性能によって発揮される会話能力においては、もはや人間の会話能力に迫る水準まで高まりつつある。その会話能力から、ChatGPT をはじめとする、LLM を搭載したチャットボットは、プログラミング学習やアイデア発想など、様々な用途で使用されている。学術的には、LLM に対して特定の性格特性を形成させる、いわゆるなりきりチャットボットのような研究も盛んにおこなわれている[1]。このようなチャットボットをはじめとして、単なる Q&A 的な用途にとどまらない、コンパニオンチャットボットの利用が広まってきている。その一例として、Replika というコンパニオンチャットボットは世界中のユーザーから利用されている。コンパニオンチャットボットのこれまでのチャットボットと異なる点は、チャットボットをタスク実行のためのツールとして利用するのではなく、会話を通じて、チャットボットと関係を形成することができるようなチャットボットである、という点である。コンパニオンチャットボットは、友人、メンター、あるいは恋人として利用されている[2-4]。

人間とチャットボットの関係 (Human Chatbot Relationship : HCR) は、LLM の台頭によりますます普及している。その中でも、コンパニオンチャットボットと人間との関係は、人間の友情関係との類似点が指摘されているものの[5]、HCR が具体的に人間関係にどのような影響を及ぼすのかはほとんどわかっていない[6]。これは、HCR 研究にて研究対象となるコンパニオンチャットボットが、直近の LLM の台頭により急速に発展したことにより、既存の HCR 研究の知見と同様の捉え方をするのが難しいことが起因している。既存の研究として、LLM を搭載していない、従来型のパターンマッチング的なチャットボットは、ユーザーのヘルスケア向上に寄与する可能性を示唆したが、既存の行動支援アプリとの決定的な優位性を主張することはでき

なかった[7]。この原因として、対話の柔軟性の低さを主要な要因として挙げている。実際、現在のコンパニオンチャットボットは、従来に比べ、遥かに人間らしく、技術的に高度な会話能力を持っている[8]。そして、LLM を搭載したコンパニオンチャットボットを使用した研究では[6]、ユーザーは、HCR が社会的健康に有益であると報告した。

HCR 研究の普遍的な議論として、人間と人間との関係において提唱されている理論が、HCR にも同様に適用できるのか、という議論が存在する。その議論の中の重要な理論の1つとして、社会的浸透理論[9]がある。社会的浸透理論では、関係が形成され、発展するための重要な要素として、自己開示を挙げている。自己開示とは、「他者に自分自身についての個人的な情報を明らかにする行為」[10]である。社会的浸透理論によれば、人々は、関係が形成されていない他者に対して基本的に自己開示を行うことはないが、自己開示が増加していくことで関係が進展していく。自己開示を行うことは、友人関係において親密さを高めるための重要な役割を果たしている[11]。

自己開示は HCR においても、重要であることが示唆されている[2,4,12]。しかし、HCR における自己開示に関連する先行研究は、一貫した結果が報告されているようには見えない。自己開示が増加し HCR が形成されたことを報告した研究[4,12]もあれば、自己開示が減少し HCR が形成されなかったことを報告した研究[2]もある。一方で、HCR における自己開示の増加がもたらす効果としては、会話相手に楽しさや温かさを感じたり[13]、親密感が増したり[14]といった、一般的にポジティブな効果が報告されている。これらの事から、HCR において、自己開示とユーザー体験との関係を明らかにすることは、重要な知見を提供することが出来る可能性が高い。

本研究では、2つの目的を立てた。第1の目的は、ユーザーの自己開示を促進させるチャットボットを開発することを目的とする。第2の目的は、ユーザーの自己開示とユーザー体験との関係を検証することである。これらの2つの目的から、以下の2つの仮説を立てた。

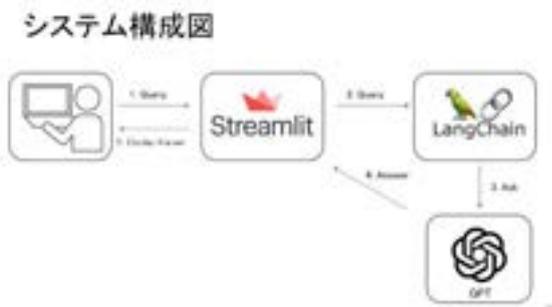
- 仮説 1: 自身の事を知ってくれているコンパニオンチャットボットと会話したユーザーは、自己開示度が増加する。
- 仮説 2: コンパニオンチャットボットとの会話において、ユーザーの自己開示度が増加すると、コンパニオンチャットボットに対する主観評定がポジティブになる。

自己開示を客観的に捉え定量化する為に、ユーザーのコンパニオンチャットボットに対する発話が、ユーザーの個人的な情報とどれだけ類似しているのかを自己開示度として指標化する。

2. RAG コンパニオンチャットボットの開発

本研究では、実験において使用する RAG (Retrieval Augmented Generation) コンパニオンチャットボットを新たに開発した。そのシステム構成を図 1 に示す。チャットボットはブラウザからアクセスできる web アプリケーションとして開発されており、使用されたライブラリやフレームワークは下記の通りである。

図 1
システム構成図



- Streamlit
- OpenAI API
- Langchain
- Chroma
- Firebase

Streamlit は、Python で web アプリケーションを作成するためのフレームワークである。Streamlit を使用することで、コンパニオンチャットボットの開発には必須となるインタラクティブなアプリケーション開発が可能となる。

OpenAI API は、LLM を web アプリケーションへ組み込むための API である。本研究では、先行研究でも広く使用されている GPT シリーズを選定し、実験当時最新であった GPT-4o を使用した。

Langchain は、LLM を利用した web アプリケーションを構築するための Python ライブラリである。RAG の実装や、会話履歴の保存機能や会話文脈記憶機能等を、Langchain を介して実装されている。

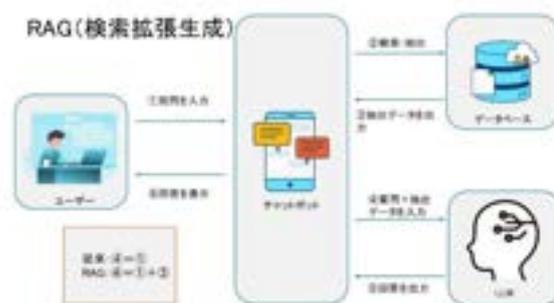
Chroma は、RAG の実装に不可欠となるベクトルデータベースの一種である。

Firebase は、様々なアプリケーションと接続してクラウドデータベース機能を提供する、Firestore を提供するクラウド開発プラットフォームである。

本研究の目的に沿って、ユーザーの自己開示を促進するコンパニオンチャットボットを開発するためには、チャットボットがユーザーの日記などの、ユーザーに関する情報を参照しながら会話することで、ユーザーのことを知ったうえで会話することが重要だと考えた。コンパニオンチャットボットの会話能力を担う LLM が、ユーザーの事を知ると、LLM の発話内容が変化し、その発話を受けて返答するユーザーの発話内容にも変化が生まれる。この変化により、ユーザーの自己開示が促進されることを目的としている。

これを実現する為の RAG を実装するためには、大きく分けて次のような 3 つのステップを経る必要がある。1) RAG データの用意、2) RAG データの埋め込み、3) RAG データを基にした LLM への入力プロンプトの拡張生成、である。まず、RAG データの用意である。このステップでは RAG に用いるためのデータを収集しておく必要がある。その形式は様々な形式が想定されるが、本研究においては、CSV において記録されたデータをユーザー毎に 1 ファイル毎の PDF へ出力することで RAG データを用意した。次に、RAG データの埋め込みを行った。これは、RAG データをベクトルデータへ変換するための埋め込みを行うためのステップである。Chroma を用いてベクトルデータベースを作成することで、RAG の準備が完了する。最後に、実際の会話場面において、ユーザーの入力に応じて、ベクトルデータ化された RAG データを適宜検索し、LLM への入力を拡張生成することによって RAG が実装される。このステップは、Langchain を使用することで実行されていた。このステップにおける、RAG の概要を図 2 に示す。

図 2
RAG の概要



3. 方法

3.1 参加者

参加者は、クラウドソーシングサービスにてオンライン上で 60 人を募集した。その内、実験を完了したのは 53 人であった (男性 23 人、女性 30 人、26~57 歳、 $M = 42.28$, $SD = 7.80$)。実験を完了した参加者へ謝金 8,000 円を支払った。

3.2 刺激

刺激は、2 章で紹介したコンパニオンチャットボットを使用した。RAG を用いていない標準的なコンパニオンチャットボットである RAG なしボットと、RAG により、参加者の事を知ったうえで会話するように指示されたコンパニオンチャットボットである、RAG ありボットの 2 種類を使用した。実験参加者との会話において、これらのコンパニオンチャットボットは参加者への提示後、参加者の発話に先立ち、必ず最初に発話を行った。この UI を図 3 に示す。参加者のアイコンと会話内容は右側に配置され、コンパニオンチャットボットのアイコンと会話内容は左側に配置された。会話は上から下に連なり、履歴が確認できるようになっていた。これらの UI は、近年広く使用されている SNS の標準的な UI に類似させたものであった。

図 3

コンパニオンチャットボット UI



3.3 アンケート

自己開示とユーザー体験との関係を探るために、4 つの尺度を使用して主観評定を測定した。1 つ目は、ポジティブ感情とネガティブ感情を測定する日本語版 PANAS[15]である。この尺度は、ポジティブ感情 (PANAS_P) とネガティブ感情 (PANAS_N) を独立した 2 次元の概念として捉え、6 件法によって測定する尺度である。それぞれの感情を測定するための 8 語を使用し、その評定値の平均値を使用する。2 つ目は、対人印象を測定する尺度[16]である。この尺度は、対人印象を温かさと有能さの 2 次元で捉え、5 件法によって測定する尺度である。それぞれの印象を測定するための 3 つの項目の評定値の平均値を使用する。3 つ目は、参加者のボットの将来的な利用意図をたずねるものである。「今回会話したチャットボットを今後も使いたい」という文章に参加

者の考えがどれだけあてはまるかを 5 件法によって測定する。4 つ目は、参加者の感じた被理解度をたずねるものである。「今回会話したチャットボットは自分の事を理解してくれていた」という文章に参加者の考えがどれだけあてはまるかを 5 件法によって測定する。これらの尺度によって測定したそれぞれの主観評定項目を、以降、PANAS_P、PANAS_N、warmth、competence、willingness、understood、と呼ぶ。

日記記入用のアンケートページは、「今日の日記を 100 文字以上で回答してください」という文章を提示し、回答させるページを作成した。このページは、100 文字未満の入力は受け付けなかったが、上限の文字数は設けなかった。

3.4 手続き

参加者は、計 4 週間毎日チャットボットとの会話と日記記入を行った。参加者は 2 種類のチャットボットに参加者間にランダムに割り当てられ、本実験中は同一条件のチャットボットを使用し続けた。手続きの開始は、まず割り当てられたチャットボットへアクセスした。チャットボットへアクセスすると、チャットボットが先行で「今日の振り返りをしよう！今日はどんな一日だった？」と参加者に話しかけるので、それに回答する形で会話が開始した。会話は計 5 ターン行い、5 ターンの会話が終了するとそれ以上の会話はできなくなり、UI 上に日記記入ページへの URL が提示された。また、毎週土曜日のみ、チャットボットとの会話と日記記入の間に、主観評定を測定した。毎週土曜日は、チャットボットとの 5 ターンの会話の後に、主観評定を測定するページを URL が提示された。主観評定に使用された 6 つの項目は、固定された順番で、PANAS_P、PANAS_N、warmth、competence、willingness、understood の順に測定された。PANAS の測定は、1 ページで行われ、項目内の 16 語はランダム化された。warmth および competence は、3 つの質問文をランダム化して 1 ページごとに提示し、評定値の平均を warmth の得点とした。willingness、understood は、それぞれの項目あたり 1 つの質問文を提示することにより測定した。日記記入は主観評定を測定するページの最後の設問として提示された。

両チャットボットは、実験環境としての状況設定を template プロンプトにより設定し、RAG ありボットにおいては、RAG により、参加者が記入した日記からエピソード記憶を積極的に参照するように会話することを指示されていた。これらの template プロンプトの内容は図 4、図 5 の通りである。チャットボットの発話はプロンプトにより、100 文字以内に収めることが指示されていた。これは、チャットボットの発話文字数の長短によって、参加者の会話体験が影響されることを防ぐためである。一方で、参加者に対しては、自然な会話を行ってもらうために、文字数の制限は行わなかった。また、実験に使用したデバイスは、web ブラウザへアクセスすることの可能なデバイスのみを使用すること

を実験の参加条件とした。毎日の手続きは、当日中に完了するように指示された。

図 4

RAG なしボットの template プロンプト

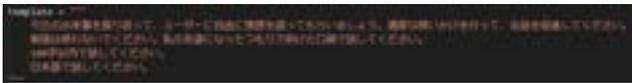


図 5

RAG ありボットの template プロンプト



3.5 自己開示度の定量化

自己開示度を定量化することで統計的仮説検定を行うために、本実験において観測可能な形で自己開示を捉える。本実験では、3つのテキストデータを収集した。1) チャットボットとの会話における参加者の発話、2) チャットボットとの会話における、チャットボットの発話、3) 参加者が記入した日記内容、である。この内、1) と 3) のテキストデータを用いて自己開示を客観的な事象として定量化した。そこで捉えた自己開示度は、ある実験日の参加者の発話テキストと、それ以前に記入された日記のテキストとの類似度、つまり、1) と 3) のテキストの類似度である。参加者が記入した日記の内容は、参加者のエピソード記憶を含む個人的な情報で、これは社会的浸透理論によれば深い自己開示情報であり、さらに、日記は参加者が体験した様々な幅広いトピックを含む。つまり、日記に記入された情報は、幅広く深い自己開示情報が記されているため、この日記に記されたテキストと、参加者の発話テキストとの類似度が高くなることは、すなわち自己開示が促進されていると考え、この類似度を自己開示度として定量化した。

2つのテキストの類似度を計算する方法としては、コサイン類似度を使用した。コサイン類似度は2つのベクトルがどれだけ同じ向きを向いているのかを計算する指標で、-1~1、で表される。コサイン類似度が0に近いほど、2つのベクトルの向きが直交に近いことを表し、符号が負であることは、逆向きを表す。このコサイン類似度を計算するためには、2つのテキストをベクトルに変換する必要があるが、そのために本研究では、Sentence-BERT モデルである、paraphrase-xml-r-multilingual-v1 を用いて、2つのテキストをベクトルへ embedding した。この Sentence-BERT モデルとは、自然言語処理能力に特化した BERT モデルをさらに、単語単位でな

く、テキスト単位の処理に特化したモデルであり、テキストをテキストのままその意味を捉え、埋め込みベクトルへ embedding するのに適したモデルである。

4. 結果

実験は、4週間にわたり行われたため、各週のデータを、以降、week1, week2, week3, week4, と呼称する。

4.1 参加者のスクリーニング

募集した60人の内、実験を最後まで完了しなかった7人を除外した。その結果、最終的な分析対象はRAGなしボット26人、RAGありボット27人の計53人であった。

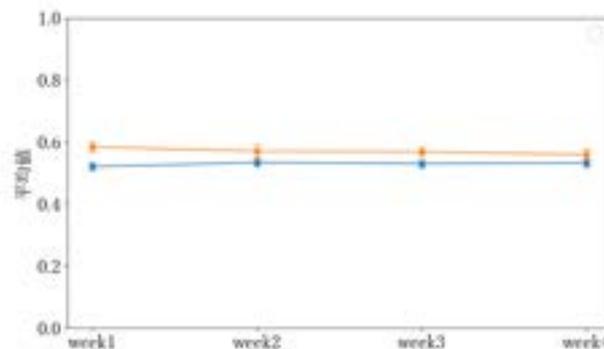
4.2 RAGありボット性能評価

自己開示度の week 毎のプロットは、図6に示す通りである。RAGありボットがRAGなしボットよりも自己開示度を増加させることが出来るのかを、week 要因を参加者内要因とした、2元配置ANOVAによって検証した結果、ボット要因の主効果が見られた ($(1, 51) = 5.59, p = .022, \eta_p^2 = .10$)。ため、RAGありボットが有意に高くRAGなしボットよりも自己開示度を増加させていたことが確認できたため、自身の事を知ってくれているコンパニオンチャットボットと会話したユーザーは、自己開示度が増加する、という仮説1は支持された。

図 6

自己開示度の時間変化

(橙: RAGありボット, 青: RAGなしボット, エラーバーは)



4.3 自己開示度と主観評定

自己開示と主観評定との関係を探るために、説明変数を自己開示度、統制変数をボット要因として、交互作用を考慮した重回帰分析を行った。その際に、ロバスト重回帰分析を行うか、従来のOLS重回帰分析を行うかの判断を行うため、Shapiro-Wilk検定により残差の正規性検定を行い、残差の正規性が認められない場合は、Huberの損失関数[17]によるロバスト重回帰分析を行った。分析の結果と散布図は以下の通りである。

図 7

week1 の自己開示度と主観評定値の散布図
(緑: RAG ありロボット, 青: RAG なしロボット)

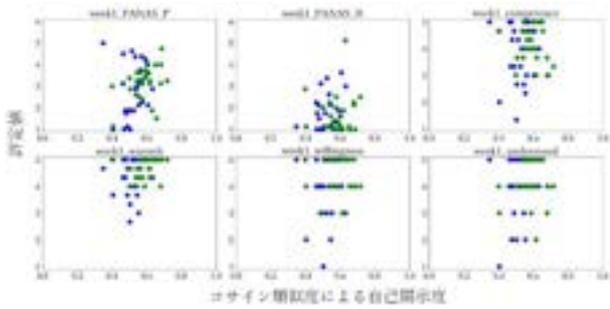


図 9

week3 の自己開示度と主観評定値の散布図
(緑: RAG ありロボット, 青: RAG なしロボット)

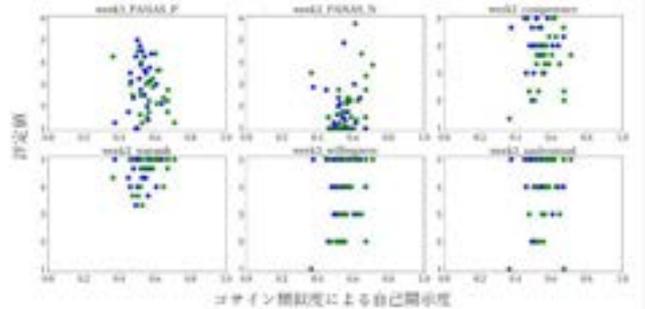


図 8

week2 の自己開示度と主観評定値の散布図
(緑: RAG ありロボット, 青: RAG なしロボット)

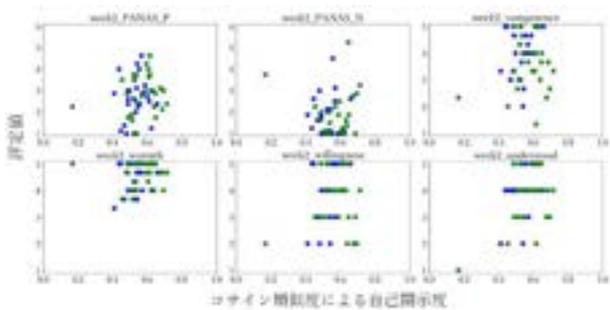


図 10

week4 の自己開示度と主観評定値の散布図
(緑: RAG ありロボット, 青: RAG なしロボット)

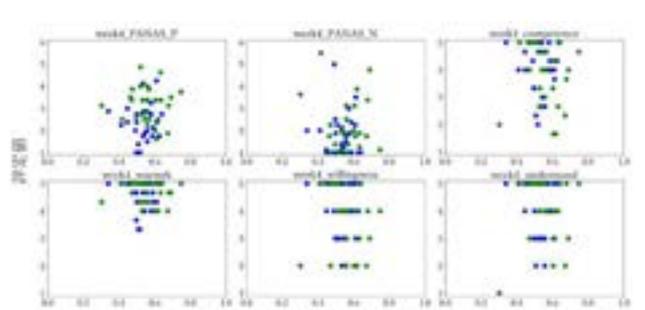


表 1

week1 の重回帰分析結果

β	PANAS_P [#]	PANAS_N	competence	warmth	willingness	understood
自己開示度	0.260	0.103	-0.120	0.171**	-0.032	0.136
ロボット	0.070	-0.093	0.099	0.018	0.138	0.021
交互作用	0.045	-0.083	-0.080	-0.107	-0.036	-0.067

[#]: OLS 重回帰, * $p < .05$, ** $p < .01$, *** $p < .001$.

表 2

week2 の重回帰分析結果

β	PANAS_P [#]	PANAS_N	competence [#]	warmth	willingness	understood [#]
自己開示度	0.133	-0.115	0.186	-0.014	0.288	0.367*
ロボット	0.115	0.063	-0.333*	0.026	-0.225	-0.189
交互作用	-0.0034	0.030	-0.096	-0.188	-0.031	0.029

[#]: OLS 重回帰, * $p < .05$, ** $p < .01$, *** $p < .001$.

表 3

week3 の重回帰分析結果

β	PANAS_P [#]	PANAS_N	competence [#]	warmth	willingness	understood
自己開示度	-0.080	0.185	0.141	0.092	0.261	0.183
ロボット	-0.110	0.116	-0.311*	-0.020	-0.289	-0.205
交互作用	-0.206	0.278*	0.117	0.042	0.299	0.402**

: OLS 重回帰, * $p < .05$, ** $p < .01$, *** $p < .001$.

表 4

week4 の重回帰分析結果

β	PANAS_P [#]	PANAS_N	competence	warmth	willingness	understood
自己開示度	0.0114	-0.049	-0.106	-0.076	-0.124	-0.069
ロボット	0.38 **	0.141	-0.001	0.111*	-0.164	0.041
交互作用	-0.045	-0.031	-0.048	0.053	-0.015	0.043

: OLS 重回帰, * $p < .05$, ** $p < .01$, *** $p < .001$.

week1 の各被説明変数毎の重回帰モデルでは, warmth において自己開示の主効果が見られた. week2 では, understood において自己開示の主効果が見られ, competence においてロボットの主効果が見られた. week3 では, competence においてロボットの主効果が見られ, PANAS_N と understood において交互作用が見られた. 単純傾斜検定を行ったところ, PANAS_N において RAG ありロボットのみに正の有意な傾きがあることが分かった ($p = .006$). understood においても RAG ありロボットのみに正の有意な傾きがあることが分かった ($p = .021$). week4 では, PANAS_P と warmth においてロボットの主効果が見られた. 自己開示度の主効果から, week1 の warmth と, week2 の understood において, コンパニオンチャットロボットとの会話において, ユーザーの自己開示度が増加すると, コンパニオンチャットロボットに対する主観評定がポジティブになる, という仮説 2 は支持された.

また, ロボットの主効果に着目すると, week2, week3 の competence において, RAG なしロボットの方が有能であると評価されていた. 一方で week4 の PANAS_P においては, RAG ありロボットの方が肯定的な感情を高めたと評価されていた.

5. 考察

実験の結果から, 本研究の2つの仮説について考察する. 先ず, 仮説 1 は, ロボット要因の主効果が見られたことから, week 要因に関係なく, RAG ありロボットが RAG なしロボットよりも自己開示を促進させていたことが確認できた. これは, RAG により, ユーザーの日記を参照しながら会話することで, コンパニオンチャットロボットの発話を変化させ, そ

れに対応する形でユーザーの発話も変化し, 自己開示が増加したのだと考えられる. 自己開示が増加することは, 心理的に様々なよい効果があることが分かっている[11,13,14]ことから, RAG コンパニオンチャットロボットは, ユーザーの発話における行動変容を引き起こし, 自己開示を促進させることで, ユーザー体験に良い影響をもたらすことができると示唆された.

次に, 仮説 2 は, week1 の warmth と, week2 の understood において, 支持された. これは, 関係の最初期は, HCR が形成され, 親密になっていく過程が, コンパニオンチャットロボットの事をより温かいと感じさせ, 次の段階では, この過程がより, 自身の事を理解してくれていると感じさせることに繋がったと考えられる. しかし, 他の主観評定項目においては, 自己開示とユーザー体験とのポジティブな関係が見られなかったことは, 2つの要因が考えられる. 1つ目の要因として, コンパニオンチャットロボットの技術的限界が挙げられる. LLM を搭載したコンパニオンチャットロボットの会話能力は, 従来と比べると飛躍的に向上したことは明らかだが, 完全に人間の会話能力を模倣するには多少の課題が残されている. 本研究においては, 特に RAG ありロボットにおいて, 日記の内容を無理に参照しようとして, 不自然な発話を行ってしまうことがあった. 2つ目の要因としては, 自己開示の時間変化には, 個人差があることが示唆されている[8]. 自己開示の時間変化は, 増加・減少・安定・変動のパターンに分けられることから, 分析の際にも, この個人差を考慮した分析を行うことで, 今回見られなかった新たな関係が見られるかもしれない.

6. おわりに

本研究の目的は、自己開示を促進するコンパニオンチャットボットを開発することと、自己開示とユーザー体験との関係を検証することであった。調査の結果、本研究で開発した RAG コンパニオンチャットボットは、ユーザーの日記を参照しながら会話することで、ユーザーの自己開示を促進させることが確認できた。また、関係の最初期において、ユーザーの自己開示度が高いほど、会話相手をより温かいと感じており、次の段階では、会話相手がより自身の事を理解してくれていると感じていることが分かった。

これらの知見は、ヘルスケア領域における行動変容支援に応用できる可能性があり、コンパニオンチャットボットの設計において重要な示唆を与える。

参考文献

- [1] Jiang, H., Zhang, X., Cao, X., Breazeal, C., & Kabbara, J. PersonaLLM: Investigating the ability of large language models to express Big Five personality traits. *p p*
- [2] Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *p* (1), 279-300. <https://doi.org/10.1177/0265407520959463>
- [3] Oh, Y. J. (2023). *p* (Doctoral dissertation). University of California, Davis.
- [4] Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzæg, P. B. (2021). My chatbot companion - a study of human-chatbot relationships. *p*, 102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- [5] Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human-AI friendship. (3), 404-429. <https://doi.org/10.1093/hcr/hqac008>
- [6] Guingrich, R. A., & Graziano, M. S. A. (2023). Chatbots as social companions: How people perceive consciousness, human likeness, and social health benefits in machines. *p p*. *arXiv:2311.12345*.
- [7] 大津 耕陽・西田 勇樹・木内 敬太・林 勇吾 (2022). チャットボットによる個人適応型ヘルスケアの実現に向けた対話型課題の導入：解決志向アプローチを題材として. *ヒューマンインタフェース学会論文誌*, (4), 285-300.
- [8] Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzæg, P. B. (2023). A longitudinal study of self-disclosure in human-
- [9] Altman, I., & Taylor, D. (1973). *p*. Holt, Rinehart & Winston.
- [10] Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. (3), 457-475. <https://doi.org/10.1037/0033-2909.116.3.457>
- [11] Carpenter, A., & Greene, K. (2016). Social penetration theory. In C. R. Berger & M. E. Roloff (Eds.), *p p* (1st ed., pp. 1-4). Wiley-Blackwell. <https://doi.org/10.1002/9781118540190.wbeic0160>
- [12] Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzæg, P. B. (2022). A longitudinal study of human-chatbot relationships. *p*, 102903. <https://doi.org/10.1016/j.ijhcs.2022.102903>
- [13] Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. (4), 712-733. <https://doi.org/10.1093/joc/jqy026>
- [14] Lee, Y. C., Yamashita, N., Huang, Y., & Fu, W. (2020). "I hear you, I feel you": Encouraging deep self-disclosure through a chatbot. In *p* (pp. 1-12). ACM. <https://doi.org/10.1145/3313831.3376175>
- [15] 佐藤 徳・安田 朝子 (2001). 日本語版 PANAS の作成. *性格心理学研究* (2), 138 - 139. https://doi.org/10.2132/jjpspp.9.2_138
- [16] 清水 佑輔・橋本 剛明・唐沢 かおり (2021). 多様な精神障害に対する人々の認知: ステレオタイプ内容モデルに着目して. *社会心理学研究* (1), 36 - 42. <https://doi.org/10.14966/jssp.2012>
- [17] Huber, P. J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. (5), 799-821. <https://doi.org/10.1214/aos/1176342503>

エコ行動促進アプリにおける PUSH 配信を用いた フィードバック手法

塚本航也 鈴木明作 吉川裕木子
 勝間田優樹 山田拓也 石川太朗
 株式会社 NTT ドコモ

概要

本研究ではユーザーへの介入手段である PUSH 配信において、文字数制限のある状況下で、ユーザーへエコ行動を促す際の効果的なフィードバック手法を検証する。従来研究ではエコ行動を促進する手法として、ユーザーの行動に対する詳細なフィードバックを行うことが有効であると示されている。しかし文字数制限のある状況下では、ユーザーの多様な行動に対して詳細なフィードバックを個別に提供することは困難である。本研究では複数のフィードバック対象を包括的に表現することで、文字数を抑えつつ、効果的にフィードバックを行う手法を提案する。エコ行動促進アプリの利用ユーザーを(1) 代表的行動の抽出と個別フィードバック、(2) 包括的表現によるフィードバック、(3) フィードバックなしの3つの群に分け PUSH 配信を実施し、PUSH 通知の開封率とエコ行動回数の変化を比較検証した。その結果、フィードバックの提供により PUSH 通知の開封率が増加し、特に包括的表現によるフィードバックはユーザーのエコ行動回数の増加に寄与することが分かった。本研究は、文字数制限のある状況において包括的表現によるフィードバックが有効であることを示し、今後のフィードバック手法の設計に貢献するものである。

1. はじめに

地球環境の深刻な変化や資源の持続可能な利用は、現代社会における喫緊の課題として注目されている。これらの課題に対処するため、国際連合は持続可能な開発目標 (SDGs) を策定し、2030 年までに達成すべき 17 の目標を掲げている。その中で、気候変動に対する具体的なアクション (目標 13) や、持続可能な生産と消費 (目標 12) が求められており、個人レベルでのエコ行動はこれらの目標達成に寄与する重要な一歩となる。

このような背景から、NTT ドコモはカーボンニュートラルへの貢献をスコア化することでエコ行動促進するアプリケーションを提供している。このアプリケーションでは、ユーザーが日常生活の中で継続的にエコ行動を行うために、定期的な PUSH 配信を実施している。PUSH 配信は、ユーザー

に対して直接的かつ迅速にメッセージを届ける技術であり、その利便性と即効性からエコ行動促進に有用である。実際に、PUSH 配信の際にユーザーにパーソナライズした文言を配信することで、効果的にユーザーのエコ行動を促している。

近年、個人の行動変容を促す手段としてナッジ[1]の活用が注目されている。ナッジとは、人々の選択を無理に制限すること無く、自然に望ましい行動を促す手段であり、行動経済学の視点から広く採用されている。例えば、PUSH 配信を行う際のメッセージにナッジを活用することで、PUSH 通知の開封率を向上させるという効果が挙げられている[2],[3]。ナッジは、人間のバイアスに基づいて設計されており、ナッジの手法の一つに「フィードバックの提供」がある。これは行動の結果を正確にフィードバックすることで、自らの行動を振り返るきっかけとなり、次回以降の行動の改善に繋がる可能性を高めるものである。この「フィードバックの提供」をエコ行動に応用することで、ユーザーのエコ行動への意識を高め、実際に行動へ移行させることが可能であると考えられる。

このようにナッジの考え方を活用し、PUSH 配信時にフィードバックを提供することで、ユーザーのエコ行動を促進する効果が期待される。しかし、ユーザーの端末ごとに表示可能な文字数は限られており、限られた文字数の中で如何にインパクトのある情報を提供するかが課題となっている。特に前述した「フィードバックの提供」を PUSH 配信で行う場合、ユーザーの多様な行動に対して一つひとつ個別にフィードバックを提供することは、文字数の増加に繋がり、PUSH 配信で実施することは困難である。

PUSH 配信の文字数制限という制約の中で効果的なフィードバックを実現するには、次の2つの手法が考えられる。

(1) 特定の行動のみに焦点を当てる：特定の行動に絞ってフィードバックを提供する。

(2) 複数の行動を包括的に表現する：複数の行動をまとめた表現を用いてフィードバックを提供する。

一方で、1.の手法ではフィードバックを受けられなかった行動が多数存在することとなり、潜在的にユーザーの不満を誘発してしまう可能性がある。また、2.の手法では包括的な表現を用いることで、言葉の抽象度が増すことになり、フィー

Feedback methods using push notifications in eco-behavior promotion apps
 KOYA TSUKAMOTO, MEISAKU SUZUKI, YUKIKO YOSHIKAWA,
 YUKI KATSUMATA, TAKUYA YAMADA, TARO ISHIKAWA,
 NTT DOCOMO, INC.

ドバックの精度が低下する可能性がある。

本論文では、NTTドコモが運営するエコ行動促進アプリケーションの利用ユーザを対象に 1.と 2.のフィードバック手法を用いて PUSH 配信を行い、ユーザの意識(PUSH 通知の開封率)と行動の変化(エコ行動回数の増減)をオンライン検証した結果を述べる。

2. 関連研究

近年、環境保護や省エネルギー促進の文脈において、ユーザの行動変容を効果的に誘導する手段としてフィードバックを活用する手法が数多く提案されてきた。とりわけ、認知バイアスやナッジ理論を応用することで、ユーザの行動をより持続的に変容させる可能性が示唆されている[4]。実際に、認知バイアスを活用した情報提示によって、エコ行動を含む様々な行動領域で介入の有効性が報告されており [5]、このような手法は限られたインタラクションを通じて利用者の意思決定を支援するうえで大きな意義を持つと考えられている。

一方、エコ行動を促す際に特に重視されるのが、行動の結果や省エネルギーの実績をユーザに対して可視化・通知するフィードバックである。たとえば、家庭のエネルギー消費データを具体的かつ継続的に提供することで、省エネルギー行動を有意に増加させられることがメタ分析によって確認されている [6]。加えて、エコ行動を行った直後に具体的な効果や数字を示す「正確なフィードバック」が、ユーザのモチベーション維持や行動回数の増加に寄与することも示唆されている [7], [8]。しかし、その一方で PUSH 配信のように文字数に制限があり、かつ配信頻度が高まるとユーザが通知を煩わしく感じる"通知摩耗"を引き起こす懸念が指摘されており [9]、効果的なタイミングや情報集約の工夫が必要となる。

これらの関連研究を踏まえ、本研究では PUSH 通知の文字数制限と通知摩耗という制約の中で、効果的なフィードバックを提供し、ユーザのエコ行動を促進するための手法を検証する。

3. 実験方法

3.1 エコ行動促進アプリ

本実証は、エコ行動促進アプリ「カボニューレコード」利用ユーザの内、PUSH 通知許諾を ON にしているユーザを対象に行われた。当アプリケーションは、図 1 のように日々の生活に関連するエコ行動を確認することができ、実際に自身が行ったエコ行動を記録することで、CO2 削減量や Reco と呼ばれる環境貢献度スコアを可視化することができるサービスである。

3.2 グループ分け

本実証では、エコ行動を促進する PUSH 配信によるフィードバックの提供の効果を効果的に測定するために、エコ



図 1 エコ行動促進アプリ画面サンプル

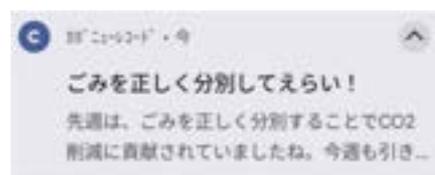


図 2 PUSH 通知画面(GroupB:ごみの分別)

行動促進アプリの利用ユーザを無作為に 3 つのグループに分けた。具体的なグループ分けは以下の通りである。

(1)代表的行動の抽出と個別フィードバック群(Group A)

このグループのユーザには、PUSH 配信を行う 1 週間前までのエコ行動記録を対象に、そのユーザにとって、1 週間を通して行動回数の最も多かったエコ行動を抽出し、そのエコ行動に対するフィードバックを含むメッセージを配信した。行動回数が最も多いエコ行動が同数で複数ある場合、無作為に 1 つを選択し、そのエコ行動に対するフィードバックを含むメッセージを配信した(この場合、残りの行動に対するフィードバックは提供しない)。1 週間を通して一度もエコ行動を行わなかったユーザには、エコ行動をしなかったことに対するフィードバックを提供し、社会的証明を含んだメッセージを配信した。社会的証明とは、人々は他者が行っていることや推奨していることを信じやすいというバイアスの事を指し、本実証では、「一緒に探しませんか?」という言葉を用いて、他者との共同行動を示唆することで、ユーザのエコ行動促進を図った。

(2)包括的表現によるフィードバック群(Group B)

このグループのユーザには、個別のエコ行動ではなく、エコ行動の上位概念を用いることで、ユーザのエコ行動を包括的に表現し、その包括的な行動に対するフィードバックを提供した。例えば、「照明をこまめに消す」という行動は「節電」に繋がり、「洗濯をまとめて行う」という行動は「節水」に繋がる。この「節電」や「節水」といった上位概念を基にフィードバックを提供するものである。PUSH 配信を行う 1 週間前までのエコ行動を上位概念に置き換え、その上位概念毎に行動回数を集計、最も行動回数の多かった上位概念に対するフィードバックを含んだメッセージを配信

表 1 エコ行動の種類とフィードバック文言

配信群	エコ行動の種類	タイトル	フィードバック文
Group A	エアコンの使用時間	エアコンの使用時間削減えらい！	先週は、エアコンの使用時間削減に取り組めていましたね。
	エアコンの設定温度	エアコンの温度設定の見直しえらい！	先週は、エアコンの温度設定を見直すことで節電に取り組めていましたね。
	冷蔵庫の設定温度	冷蔵庫の温度設定を見直しえらい！	先週は、冷蔵庫の温度設定を見直すことで節電に取り組めていましたね。
	テレビを観る時間	テレビの使用時間削減えらい！	先週は、テレビの使用時間削減に取り組めていましたね。
	照明の使用時間	照明の使用時間削減えらい！	先週は、照明の使用時間削減に取り組めていましたね。
	パソコンの使用時間	パソコンの使用時間削減えらい！	先週は、パソコンの使用時間削減に取り組めていましたね。
	便座の設定温度	トイレの便座の温度設定の見直しえらい！	先週は、便座の温度設定を見直すことで節電に取り組めていましたね。
	まとめて洗濯	まとめての洗濯えらい！	先週は、まとめて洗濯することで節水に取り組めていましたね。
	炊事・洗濯	炊事・洗濯の節水えらい！	先週は、炊事・洗濯の節水に取り組めていましたね。
	風呂・洗面	風呂・洗面の節水えらい！	先週は、風呂・洗面の節水に取り組めていましたね。
	マイボトル	マイボトルを持参してえらい！	先週は、マイボトルを持参することでリサイクル(資源削減)に取り組めていましたね。
	マイバッグ	マイバッグを持参してえらい！	先週は、マイバッグを持参することでリサイクル(資源削減)に取り組めていましたね。
	マイ箸・マイストロー	マイ箸・ストローを利用してえらい！	先週は、マイ箸・ストローを利用することでリサイクル(資源削減)に取り組めていましたね。
	食べ残しをしない	食べ残しをしないでえらい！	先週は、食べ残しをしないことでフードロス削減に貢献されていましたね。
	フードバンク	フードバンクへ寄付してえらい！	先週は、フードバンクへの寄付でフードロス削減に貢献されていましたね。
	再配達を利用しない	再配達を利用しないでえらい！	先週は、再配達を利用しないことでCO2削減に貢献されていましたね。
	テレワーク	テレワークの取組みえらい！	先週は、テレワークをすることでCO2削減に取り組めていましたね。
	コンポスト	コンポストを使用してえらい！	先週は、コンポストを使用することでCO2削減に取り組めていましたね。
	ゴミの分別	ゴミを正しく分別してえらい！	先週は、ゴミを正しく分別することでCO2削減に貢献されていましたね。
	Group B	行動無し	エコ行動を行って、Reco を貯めよう！
節電		節電に取り組んでえらい！	先週は、節電に取り組めていましたね。
節水		節水に取り組んでえらい！	先週は、節水に取り組めていましたね。
リサイクル(資源削減)		リサイクル活動に取り組んでえらい！	先週は、リサイクル(資源削減)に取り組めていましたね。
フードロス削減		フードロス削減に取り組んでえらい！	先週は、フードロス削減に取り組めていましたね。
CO2削減		CO2削減に取り組んでえらい！	先週は、CO2削減に取り組めていましたね。
行動無し		エコ行動を行って、Reco を貯めよう！	先週は、エコ行動記録がありませんでした。小さなことからできること一緒に探してみませんか？
Group C	行動に関わらず	エコ行動を行って、Reco を貯めよう！	エコ行動に取り組んで記録をつけましょう！

した。行動回数が同数、エコ行動が一度も無かったユーザについては、Group A と同様の配信処理を行った。

(3)フィードバックを行わない群(Group C)

このグループのユーザには、ユーザのエコ行動有無に関わらず、フィードバックは提供せず、単にエコ行動を促すメッセージを配信した。

各グループに対する無作為な割り当ては、参加者の背景(年齢、性別、職業など)の影響を最小化し、グループ間の均一性を確保するために実施された。この方法により、フィードバックの種類がエコ行動の実行頻度にどのような影響を与えるかを正確に評価することを可能にしている。

3.3 PUSH 配信方法

エコ行動促進アプリで扱うエコ行動の種類とフィードバック文言を表 1 に示す。集計期間に何らかのエコ行動を行ったユーザに対しては、フィードバック文言の後に「今週も

引き続きエコ行動を行って、Reco を貯めよう！」という文言を付加して、2024 年 12 月 25 日に PUSH 配信を行った(図 2)。

3.4 分析手法と分析範囲

エコ行動促進アプリから得られるデータには、PUSH 通知を開封したかどうか、ユーザが自身で記録する日々のエコ行動データがある。本研究ではフィードバックを含むメッセージを PUSH 配信することによるユーザの意識と行動の変化を検証する事を目的とし、各グループ間の PUSH 通知開封率、t検定を用いてエコ行動回数の増減平均に有意な差があるかを確かめるため、カイ二乗検定を実施した。

PUSH 配信による介入前 1 週間と、PUSH 配信後 1 週間を集計期間とする。分析の際、集計期間中にエコ行動を 1 度も行わず、PUSH 通知も開封しなかったユーザについては、アプリケーションの利用を止めた(離脱者)とみなし、集計対象

から除外する。エコ行動回数の増減平均を求める際は、PUSH 通知を開封したユーザを対象とし、PUSH 通知を開封しなかったユーザは通知に目を通さなかったものとして集計対象から除外する。

PUSH 開封率, エコ行動回数の変化とフィードバック手法の関係を明らかにするため, 分析の際はグループごとに全ユーザを対象にした結果を示すとともに, 前週のエコ行動の種類数が 0 種, 1 種類, 2 種類以上のセグメントに分けた結果を示す。

4. 検証結果

4.1 PUSH 通知開封率

各グループの PUSH 開封率を表 2 に示す。フィードバックを提供したグループ(Group A, Group B)は, 提供しなかったグループ(Group C)と比較して, わずかに開封率が向上した。具体的には, Group A の開封率は 25.95%, Group B は 26.39%, Group C は 23.27%であった。

更に前週のエコ行動種類数が 0 種, 1 種類, 2 種類以上のセグメントに分けた PUSH 開封率を表 3 に示す。エコ行動種類数が 0 種(=前週行動を行わなかった)のユーザにおいては, どのグループにおいても同水準の結果となった。行動種類数が 1 種のユーザの場合, Group A の開封率が最も高い 24.00%であり, 2 種以上のユーザの場合, Group A が 16.94%, Group B 16.96%と同水準となり, Group C は 14.47%と他グループと比較してわずかに劣る結果となった。

また, PUSH 通知開封率に対し, 各グループ間でカイ二乗検定を行った結果を表 4 に示す。有意水準を 5%としたとき, どのグループ間にも有意差は認められなかった。また, セグメント別にカイ二乗検定(表 5)を行った場合においても, 有意差は認められなかった。

4.2 エコ行動回数増減の平均

PUSH 配信前後のエコ行動回数の増減を表 6,表 7 に示す。また, 詳細分析として, 前週のエコ行動種類数ごとのエコ行動回数の増減平均を図 3, 図 4 に示す。

エコ行動種類数が 0 種(=前週行動を行わなかった)のユーザにおいては, フィードバックを提供したグループ(Group A, Group B)が, フィードバックを提供しなかったグループ(Group C)に対してエコ行動回数が上回る結果となった。具体的には Group A は前週比で平均 2.75 回増加, Group B は平均 4.25 回増加, Group C は 1.64 回増加であった。前週のエコ行動種類数が 1 種のユーザの場合, どのグループも同水準となっており, 平均 1.27 回の増加となった。2 種以上のユーザの場合, Group C は増減が見られなかったのに対し, Group A は平均 2.38 回増加, Group B は平均 4.63 回と大きく増加する結果となった。

また, PUSH 通知を開封したユーザを対象に, 各グループ間のエコ行動回数の増減の平均の差について, t検定を行う。有意水準を 5%とし, Group B > Group A > Group C の順に効

表 2 PUSH 開封率

グループ	配信対象者数	開封者数	開封率
Group A	1,545	401	25.95%
Group B	1,508	398	26.39%
Group C	1,521	354	23.27%

表 3 PUSH 開封率(セグメント別)

グループ	配信対象者数	開封者数	開封率	
Group A	(=0)	369	200	54.20%
	(=1)	25	6	24.00%
	(>1)	1,151	195	16.94%
Group B	(=0)	378	207	54.76%
	(=1)	22	3	13.64%
	(>1)	1,108	188	16.96%
Group C	(=0)	342	186	54.39%
	(=1)	32	2	6.25%
	(>1)	1,147	166	14.47%

表 4 カイ二乗検定の結果

グループ	χ^2 値	p値
Group A vs Group B	0.054	< .815
Group A vs Group C	2.824	< .093
Group B vs Group C	3.780	< .052

表 5 カイ二乗検定の結果(セグメント別)

グループ	χ^2 値	p値	
Group A vs Group B	(=0)	0.006	< .936
	(=1)	0.280	< .597
	(>1)	0.000	< 1.000
Group A vs Group C	(=0)	0.000	< 1.000
	(=1)	2.341	< .126
	(>1)	2.462	< .117
Group B vs Group C	(=0)	0.001	< .979
	(=1)	0.196	< .659
	(>1)	2.466	< .117

表 6 エコ行動回数の増減

グループ	開封者数	増減数平均	標準偏差
Group A	401	2.75	12.15
Group B	398	4.25	11.43
Group C	354	1.64	10.73

表7 エコ行動回数の増減(セグメント別)

グループ	開封者数	増減数平均
Group A (=0)	200	3.16
	6	1.17
	195	2.38
Group B (=0)	207	3.96
	3	1.33
	188	4.63
Group C (=0)	186	2.77
	2	1.5
	166	0.39

表8 t検定の結果

グループ	t値	p値
Group B > Group A	1.798	< .037
Group B > Group C	3.219	< .001
Group A > Group C	1.323	< .094

果が高いものとし、片側検定を行った結果を表8に示す。Group B > Group A, Group B > Group Cにおいて有意差が認められた。一方でGroup A > Group Cにおいては有意差が認められない結果となった。

5. 考察

5.1 PUSH 通知開封率に対する考察

PUSH 配信の前週にエコ行動を行わなかった人たちについては、フィードバック有無に関わらず、同程度の開封率となっている為、エコ意識を新たに芽生えさせるためにはフィードバックが特段有効な手段であるとは言えない。これは、エコ意識が低いユーザに対しては、フィードバックを含んだメッセージであっても、行動変容を促すには不十分であることを示唆している。

一方で、前週に何らかの行動を行っていた人たちに対しては、フィードバックを含むメッセージを送ることで開封率が向上した。これは、すでに行動を行っている人たちにとって、自身に向けられたパーソナライズメッセージである為、一般的に広く発信するメッセージよりも興味がそられる結果となったのではないかと考えられる。

なお、行動回数が1種の場合はGroup Aの正確なフィードバックが最も高い開封率となっている為、フィードバックを行う際に制約が無いのであれば、正確なフィードバックが効果的であると考えられる。また、前週のエコ行動回数の種類が2種以上の場合、Group AもGroup Bも同水準のPUSH開封率となっており、一部の行動に対してフィードバックを受けるのと、包括的にフィードバックを受けるのでは、特にユーザの興味が惹く度合いに差はないことが分かった。

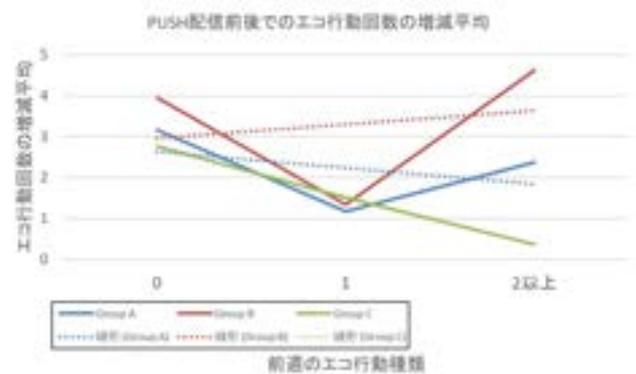


図3 PUSH 配信前後でのエコ行動回数の増減平均

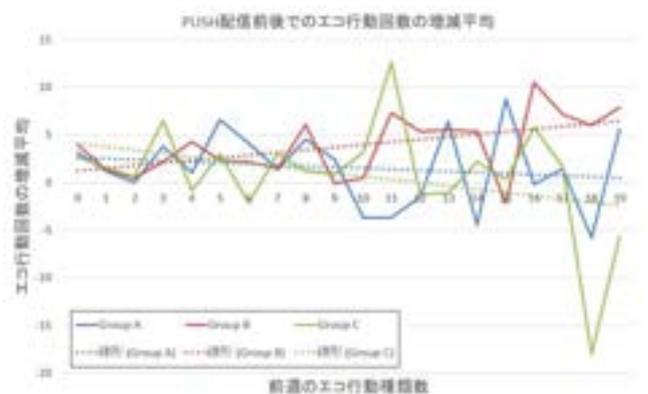


図4 PUSH 配信前後でのエコ行動回数の増減平均(詳細)

これらの事から、自身に向けられたメッセージである事が分かれば、人は興味を抱きやすく、いわゆる大衆に向けられたメッセージに対しては興味がそそられにくいものであると考えられる。また、自身に向けられたメッセージは、その内容が正確であればあるほど興味度合いが強まる傾向にあることが考えられる。これは自分のことを分かってくれているという事に対する親近性バイアスが影響したものだと考えられる。

5.2 エコ行動回数の変化に対する考察

t検定の結果より、Group BはGroup A, Group Cと比較して有意にエコ行動回数が増加した。図3に示すように、フィードバック対象となる行動が多岐にわたるほど、Group AとGroup Cは行動回数の増減が減少傾向にあるのに対し、Group Bは効果が高まる傾向にあることが分かる。このことから、包括的表現によるフィードバックを行うことは、ユーザのモチベーション向上(エコ行動回数の増加)に寄与することが示唆された。

この結果は、フィードバックの内容や対象行動の選択がユーザのモチベーションに複雑な影響を与える可能性を示唆している。Group Aでは、フィードバックされる内容が限定的であったため、ユーザは特定の行動のみに意識が集中し、他の行動へのモチベーションに繋がらなかった可能性がある。一方、Group Cでは、具体的なフィードバックを受

けられていないため、モチベーションの維持に繋がらなかった可能性がある。これに対し Group B では、包括的な表現によるフィードバックにより、ユーザは連想される様々な行動に対して良い行いだと自覚することができたため、高いモチベーションを維持できたと考えられる。

さらに、今回の結果については、オペラント条件付けの観点からも解釈できる。オペラント条件付けとは、行動の結果によってその行動の頻度が増減する学習過程である。フィードバックを受けることは、ユーザのエコ行動に対する正の強化子として機能し、その後のエコ行動を促進した可能性がある。つまり、ユーザはフィードバックを受けることで、自身の行動が肯定的に評価されたことを認識し、同様の行動を繰り返すようになったと考えられる。

6. まとめ

本論文では、文字数制限のある PUSH 配信環境下において、ユーザのエコ行動を効果的に促進するためのフィードバック手法を検証した。地球環境の深刻な変化や資源の持続可能な利用が現代社会における喫緊の課題として注目されている背景を踏まえ、個人のエコ行動を促進する手段として、ナッジ理論(フィードバックの提供)を活用した PUSH 配信に着目した。

従来のフィードバック手法では、文字数制限のために全ての行動を網羅することが難しい。特定の行動のみに焦点を当ててフィードバックを提供した場合、ユーザは特定の行動のみに意識が集中し、他の行動へのモチベーションに繋がらない可能性が考えられる。一方、複数の行動を包括的に表現するフィードバック手法では、ユーザの多様なエコ行動を上位概念に置き換えて包括的に表現することで、限られた文字数でも効果的にフィードバックを伝えることを可能にする。本研究では、この包括的表現を用いたフィードバックが、文字数制限のある PUSH 配信環境下において、ユーザのエコ行動を促進する上で有効であるという仮説を立て、検証を行った。

エコ行動促進アプリの利用ユーザの内、PUSH 通知許諾を ON にしているユーザを対象とした実証の結果、フィードバックを提供することにより、PUSH 通知の開封率向上に寄与することが分かり、特に包括的表現によるフィードバックはユーザのエコ行動回数を有意に増加させる効果があることが明らかになった。これは、文字数制のある状況下では、ユーザの多様な行動を包括的に捉え、肯定的なフィードバックを与えることで、モチベーションを高め、行動変容を促進することを示唆している。

本研究の成果は、PUSH 配信を活用したエコ行動促進のための新たな手法を提示するものであり、SDGs の達成に向けた個人の行動変容を促進する上で貢献をなすものである。今後の課題としては、以下のような点が挙げられる。

- **包括的表現の階層化と最適化:** 本研究では「節電・節

水」という行動レベルで包括的な表現を試みたが、より上位概念である「エコ行動」や「省エネ」といった包括的表現を用いることで、更なる効果が期待できる可能性がある。今後は、ユーザの行動特性や属性を考慮しつつ、「エコ行動」を核とした包括的な表現と、具体的な行動レベルの表現を組み合わせるなど、多層的なアプローチを検証することで、より効果的な表現方法を検討する必要がある。

- **長期的な影響の調査:** 本研究では短期的な効果を検証したが、長期的な視点でユーザの行動変容を促すためには、包括的表現によるフィードバックがどのように影響するのかを継続的に調査する必要がある。モチベーションを維持するための工夫や、習慣化を促進するための仕組みづくりも重要な課題となる。
- **季節やイベントによる効果の変化の検証:** 12月に行った今回の検証では、年末年始の休暇や暖房の使用増加など、季節的な要因がエコ行動に影響を与えている可能性がある。そのため、異なる時期に検証を行うことで、季節やイベントがフィードバックの効果に与える影響を分析し、より効果的なフィードバックの提供につなげることが重要となる。

これらの課題に取り組むことで、より効果的なエコ行動促進のためのフィードバック手法の開発に繋がると考えられる。本研究は、文字数制限のある状況において包括的表現によるフィードバックが有効であることを示し、今後のフィードバック手法の設計に貢献するものである。

参考文献

- [1] Leonard, T. C.: Richard H. Thaler, Cass R. Sunstein, “Nudge: Improving decisions about health, wealth, and happiness” (2008)
- [2] Hyun Bae Kim, Toshiya Iwamatsu, Ken-ichiro Nishio, Hidenori Komatsu, Toshihiro Mukai, Yoko Odate, Masanobu Sasaki, “Field experiment of smartphone-based energy efficiency services for households: Impact of advice through push notifications”, Energy and Buildings, Vol. 223 (2020)
- [3] Carmina G. Valle, Brooke T. Nezami, Deborah F. Tate, “Designing in app messages to nudge behavior change: Lessons learned from a weight management app for young adults”, Organizational Behavior and Human Decision Processes, Vol. 161, pp. 95-101, (2020)
- [4] 吉川裕木子, 酒井亮勢, 鈴木明作, 鈴木喬, 石川太朗, et al. 認知バイアスと大規模言語モデルを活用したメッセージ最適化システムの提案. 研究報告高度交通システムとスマートコミュニティ (ITS), 2024(16):1-6, 2024.
- [5] Christopher Henkel, Anna-Raissa Seidler, Johann Kranz, and Marina Fiedler. How to nudge pro-environmental behaviour: an experimental study. In ECIS, 2019.

- [6] Beth Karlin, Joanne F Zinger, and Rebecca Ford. The effects of feedback on energy conservation: A meta-analysis. *Psychological bulletin*, 141(6):1205, 2015.
- [7] Jon Froehlich, Leah Findlater, and James Landay. The design of eco-feedback technology. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1999–2008, 2010.
- [8] Kathryn Buchanan, Riccardo Russo, and Ben Anderson. Feeding back about eco-feedback: How do consumers use and respond to energy monitors? *Energy policy*, 73:138–146, 2014.
- [9] Ganugapenta Muralidhar Reddy, Reduce the Amount of Push Notifications Require for E-Commerce Apps, *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT* 09(01):1-9, 2025

経験知を取り込んだ機械学習と デジタルツールに対する心理的障壁の低減

村瀬颯登^{†1}
株式会社 UACJ^{†1}

山本祐樹^{†2}
株式会社 UACJ^{†2}

1. はじめに

製造業においてデジタル化の必要性は長年認識されており、その導入が広く取り組まれている[1]。株式会社 UACJ(以下、UACJ)も例外ではなく、日々の操業データを活用し、データ可視化ツールや機械学習による特性予測ツールなど、データ駆動型の意味決定を支援するデジタルツールを開発・導入を試みてきた。しかしながら、業務への定常的な活用は進んでいない。その背景には、以下のような製造業特有の文化・思想が少なからず影響している。

- 5ゲン主義(現地、現物、現実、原理、原則を重視)
- 役割が明確な分業体制
- 標準化を基本とする業務特性

現状の現場における意思決定プロセスおよび業務遂行の状況は、定まった経験、感覚に頼った意思決定が中心であり、業務細分化により担当範囲内のみに注力する傾向がある。また、新しく未知な技術やツールへの抵抗も強い。社会がデータ駆動意思決定により競争力を確保している中で[2][3][4][5]、このような保守的状況は、今後の当社の競争力維持の阻害リスクをもたらす。業務改革特にデジタルトランスフォーメーション (DX) の推進のためにはデータ駆動型意思決定という意味での行動変容が求められる。表 1 に技術現場におけるデータ駆動意思決定と従来の意思決定の違いを比較した。

表 1 データ駆動型方法論と伝統的方法論の比較

Table 1 Comparison of data-driven methodologies and traditional methodologies.

項目	データ駆動型方法論	伝統的方法論
プロセス	大量のデータからパターン・傾向を抽出	事前の知識や仮説に基づき実験・観察
強み	予測精度の向上 新たな知見の発見	確実性(検証済みの場合) 論理的推論が容易
弱み	データのバイアス 解釈の困難さ	思考のバイアス (感覚、経験知)

Machine learning incorporating empirical knowledge and reducing psychological barriers to digital tools

^{†1} MURASE HAYATO, UACJ Corporation

^{†2} YAMAMOTO YUKI, UACJ Corporation

UACJにおいて、データ駆動型意思決定を促進すべく、過去に導入したデジタルツールは、期待した行動変容を呼び起こすに至らなかった。興味深いことに、ユーザーニーズを満たすことを重視したユーザー中心設計[6]を徹底しても、過度にカスタマイズされたツールが完成し、一部ユーザーしか利用できず、運用停止という結果に終わった事例も経験している。このような状況から、ツールの定着を促す施策として機能や GUI を過度に調整することは得策ではないと考えられる。そこで、データ駆動型意思決定と従来の意思決定の相違点に着目し、ツールの定着が進まない要因をツールが出す結果の感覚・経験知との相違と考えた。Human-in-the-Loop (HITL) [7]の考え方にに基づき、データと感覚を補完的に組み合わせることで、製造業特有の文化を尊重しつつ、データ駆動型意思決定を促進できる新たなアプローチを模索した。

2. 方策

過去に UACJ において製造現場で定着しなかったデジタルツールである特性予測ツール (パターン A、図 1) ¹ の課題を踏まえ、新たなツール(パターン B)を開発する。デジタル活用リテラシーが醸成途中[8]である製造現場において、パターン A は多機能化による複雑な操作性による認知負荷増大[9]、機械学習やデータ解析の知識を前提とした設計、およびデータに基づく出力を重視しユーザーの経験や感覚を軽視していた点が、普及の阻害要因であったと考えられる。



図 1 特性予測ツール(パターン A):
機能・GUI が複雑で使用方法が分かりにくい

Figure 1 Characteristic prediction tool (Pattern A)

¹ ツールの予測対象は材料の基本的性質である引張強度、耐力、伸びといった機械的特性である。本ツールの想定使用用途は、規格を満たす製造条件の決定や、製品不良発生時の原因究明・対策立案を支援することである。

ツールの設計

新たなツール(パターン B)開発では、以下の点を重視した。

1. 経験知の反映: ユーザーの感覚や経験に基づいたノウハウを機械学習モデルに反映させることで、予測の精度向上だけでなく、ユーザーが予測結果を理解し、納得できる仕組みを実現すること

2. ユーザーフレンドリーな設計: 複雑な機能を排除し、直感的に操作できるインターフェースとすることで、操作性と認知負荷を低減[10]すること。

3. 外挿領域予測機能: 未知の領域における予測にも対応可能なモデル構築を目指し、新たな知見の獲得と業務活用の幅を広げること。

上記 3 点の重視により、試作の迅速化、規格適合製造条件の決定、製品不良原因究明・対策立案支援を目指した。

3. 開発結果

ツールの機能は必要最低限の機能、過去実績データ抽出(図 2)および特性予測機能(図 3、図 4)のみとした。使用する機能はユーザーが対話型で選択する(図 5)。



図 2 実績探索画面

Figure 2 Achievement exploration screen.



図 3 質問回答後の変数入力画面(特性予測機能)

Figure 3 Variable input screen after response (characteristic prediction function).



図 4 特性予測結果画面(特性予測機能)

Figure 4 Characteristic prediction result screen (characteristic prediction function).



図 5 シンプルな機能と GUI

ツール立ち上げ時に表示される画面は極めてシンプル

Figure 5 Simple functions and GUI.

表 2 アンケート設問一覧

Table 2 List of survey questions

質問 ID	質問内容
1	自分の感覚を取り入れる点はどうか
2	理由を教えてください
3	外挿予測ができるとどうか
4	シンプルな機能・見た目にしたがどう感じるか
5	理由を教えてください
6	仮に機能が同じ場合、対話形式のツールのほうがよいか
7	理由を教えてください
8	最もプラスの効果があった(ありそう)のはどれか
9	2 番目にプラスの効果があった(ありそう)のはどれか
10	ツール関係ない質問です。ツール(B)を使う以前のデータ活用/デジタル技術全般に関する積極性
11	ツール(B)を使った後のデータ活用/デジタル技術全般に関する積極性
12	ツール(A)の総合評価。使った場合は次使いたい、使っていない場合は使いたいかどうか
13	理由を教えてください
14	ツール(B)の総合評価。引き続き使いたい
15	理由を教えてください



図 6 対話型 GUI(特性予測機能)

指示・質問に回答するだけであり使用法が明白

Figure 6 Interactive GUI (Characteristic Prediction Function)

特性予測機能では、材質と調質も同様に対話形式で回答することで該当データのみを抽出し、モデル構築を行う(図 6)。多重共線性があれば影響の大きい変数を問う追加の質問がされ、ユーザーの回答をもって特性予測モデルが構築される。

対話型 GUI による経験知の取り込みは、直感的操作が可能になるだけでなく、ツールとの対話によりモデルを共創している感覚をユーザーに与えると同時に、ツール出力の結果にユーザーも責任を負うこととなる。これによりツール出力の結果への安心感・納得感を得ることができ、定着が進むと期待できる。

4. ユーザーによる評価結果の分析

4.1 評価

今回開発したツールが狙い通りの効果があったかどうかを調査するためにツールユーザーに対しアンケートを実施した。アンケートは対象部門 26 名中 21 名から回答が得られた(回答率 81%)。アンケートの設問は表 2 の通り。計 15 個の質問で、各設計の是非を問う質問、データ活用/デジタル技術に対する積極性(以下デジタル技術への積極性とする)に関する質問、ツール(パターン A,B 両者)の総合評価とその理由にて構成されている。また、回答はすべて選択式で選択肢は数値が小さいほどポジティブ、数値が大きいほどネガティブな回答になるよう設計した。

4.2 ユーザーの分類

デジタル技術への積極性を問う設問(Q10、Q11)で得られた回答を k-means にて 4 グループにクラスタリングした。図 7 は、横軸を Q10(ツールを使う前の積極性)の回答、縦軸を Q11(ツールを使う後の積極性)の回答としてプロットした散布図であり、クラスターごとに色分けしたものである。

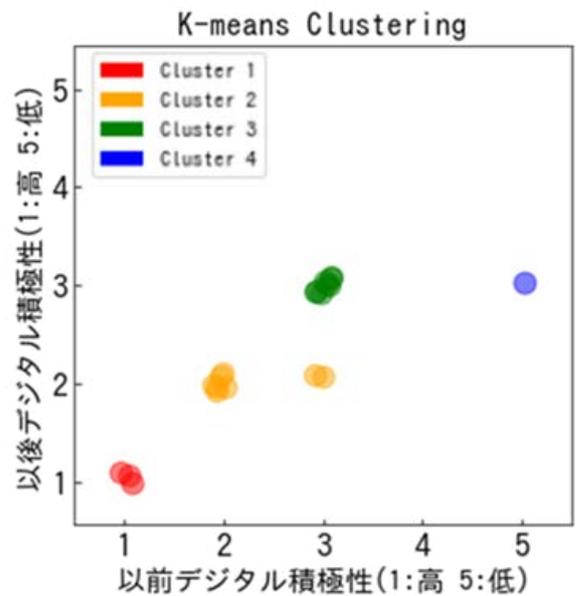


図 7 ユーザー属性分類結果

Figure 7 User attribute classification results.

おおまかに

- クラスター1：ツール使用以前以後とも積極的
- クラスター2：ツール使用以前以後ともやや積極的
- クラスター3：ツール使用以前以後とも中程度
- クラスター4：ツール使用以前は消極的だが以後積極的

という特徴でユーザーの属性が分類されていることが分かる。

なお、一度ツール(パターン B)を使用しただけでデジタル技術に対する積極性が向上するとは必ずしも言えないが、もともと消極的であった者が中程度に変わったことや向上した者が存在することは事実であり、そのことから、継続的な活用及び、便利なツールを他にも提供することでデジタル技術への積極性や DX 意識が向上する可能性は十分に考えられる。

4.3 ツール設計の効果の確認

その他すべての設問の回答結果を図 8 に示す。ユーザーの属性ごとの傾向が分かるように、4.2 節のクラスタリング結果に基づいて色分けを行った。本結果から、ユーザーがデータと感覚を組み合わせたモデル共創を受け入れていることが分かる。また、心理的障壁の低減により、今後もツールの継続的な活用が期待され、データ駆動型意思決定への行動変容が進むことが示唆される。以下、詳細を示す。

Human-in-the-Loop (HITL) の効用

図 8 から、Q1 の回答は感覚を取り入れた設計に対する評価は「1 よい」と「2 どちらでもない」が同数、「3 悪い」が両者の半数であった。全体として、感覚を取り入れることは肯定的に評価されている。Q2 の詳細分析では、「2 場合による」という回答が最多だった。これは、データの精度が高い場合、データ駆動型の判断が支持される一方で、精度が低い場合は感覚を補完的に活用すべきであるという考えを示している。「1 感覚と合致したほうが安心できる」「2 場合による」と回答した 10 名は、精度向上を前提に感覚を取り入れることに肯定的な意見を示しており、感覚と合致した安心感や納得感を期待していると考えられる。一方で、「4 データのみのほうが安心できる」という回答も 6 件あり、「1 感覚に合致したほうが安心できる」の 3 件を上回っている。これは、必ずしも感覚を取り入れることに安心感を持つとは限らず、データのみで判断することに信頼を置く意見も存在することを示している。以上から、基本的にはデータに基づいた意思決定を行い、不足情報を補うために必要に応じて感覚・経験知を活用することが最適であると考えられる。これは、HITL すなわち、データ・AI と人間の協調による意思決定が製造現場で受け入れられる態勢が整っていることを示唆している。

心理的障壁低減効果

Q4 と Q5 の結果から、シンプルな設計は心理的障壁を効果的に低減していることが明らかになった。特に、「使うハードルが下がった」との回答が多数寄せられ、ツールへの入り込みやすさが向上したことが確認できた。一方で、「3 メリットデメリット両方あり」という回答も存在するため、機能の少なさからデメリットを感じることもあるユーザーもいる点に留意が必要である。対話型 GUI についても、Q6 と Q7 からは心理的障壁低減効果が期待した通りに現れていることが示された。ユーザーからは「使うハードルが下がった」という意見が中心で、ツールへの接点が自然で親しみやすいと感じられたことが理解できる。ただし、「5 使いにくい、遅い等」といった使いやすさに関する課題を指摘する回答も少数ながら存在するため、さらなる改善が必要であることが浮き彫りになった。両方の設計を総合的に比較すると、対話型 GUI とシンプルな機能・見た目という 2 つのアプローチが、共に心理的障壁低減に効果的であることが明確となった。さらに、Q8 からユーザー属性別の傾向を分析すると、デジタルツールの積極性レベルによってどちらの設計がより高く評価されるのかという新たな知見を得られた。デジタルツールに馴染みの薄いユーザー層は対話型 GUI に好感を示し、デジタルツールへの意欲が高いユーザー層はシンプルな機能・見た目に惹かれる傾向が見られた。

機能面の設計効果

外挿予測機能の有効性がユーザーにとって重要な要素であることが明確に示された。特に、Q3 における「1 使いたいと思う」という回答が最多である点は、これまでに困難であった外挿予測が本ツールにより実現可能となり、実験や試作なしに迅速に新たな知見を得られるという点が高い評価を受けていることを示唆している。一方、パターン A のツールに関する Q12 と Q13 の回答は、「3. 使い方が分からない」という理由から、多くが使用を望まなかった。これは、導入時の説明のみでは使い方が習得しにくい、あるいはマニュアルが必要なツールは、ユーザーにとって利用障壁となることを示している。対照的に、パターン B のツールは Q14 と Q15 において「2. やや使いたい」という回答が最も多く、「3. 外挿予測ができるから」がその理由として挙げられている。また、「2. 使いやすいから」や「6. デジタルツール自体使わない」も続く。重要な点は、Q1~Q9 がツール使用を前提とした設計思想に関する設問であるのに対し、Q14, Q15 がツールに対する使いたい/使いたくないという設問であった点を踏まえる必要がある。機能面での有効性を理由とする回答が最も多いことは、ユーザーにとって必要な機能が存在することが活用の前提となることを示唆している。さらに、使いやすさも重要な要素であると言える。

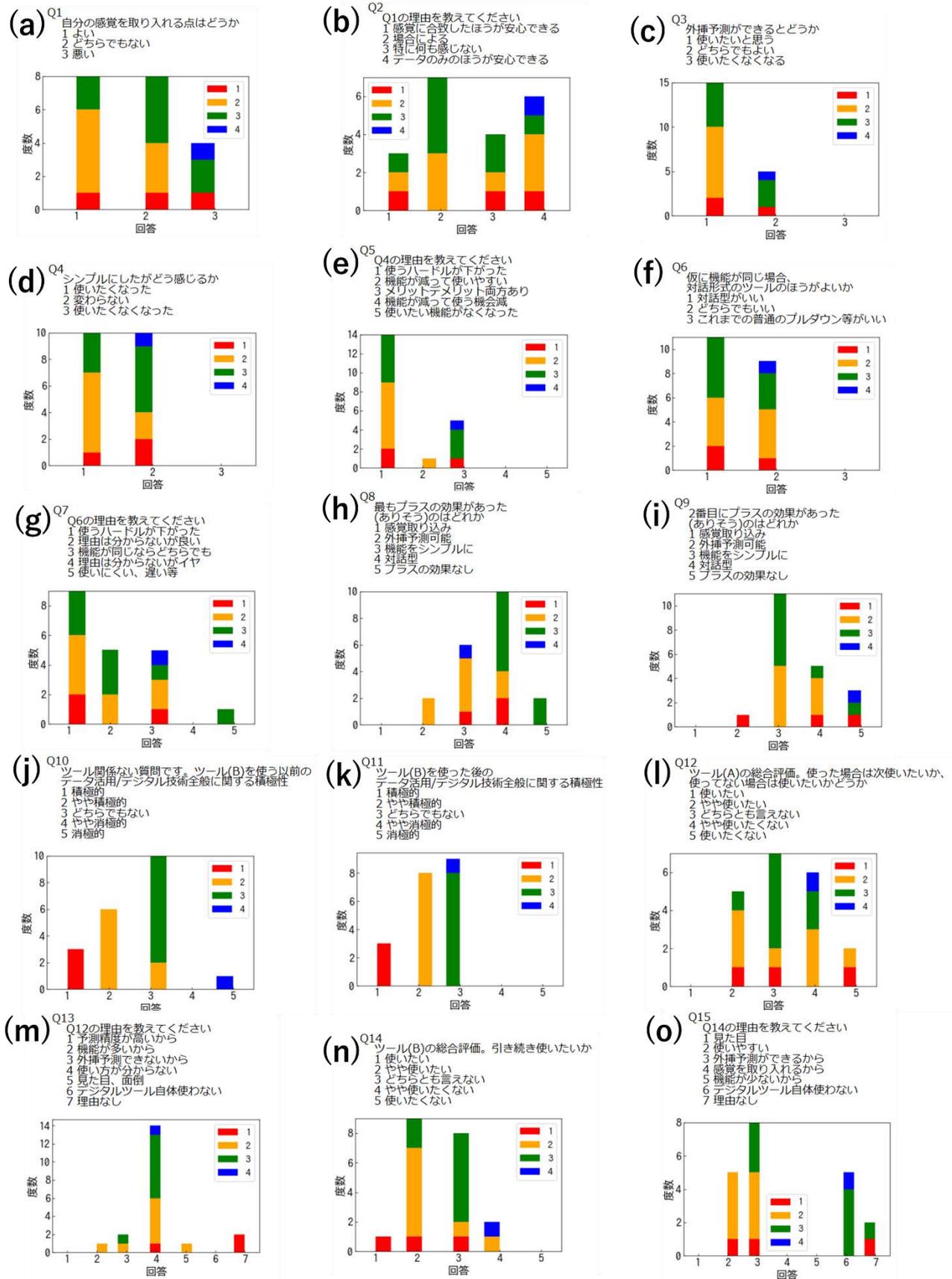


図 8 アンケート結果

4つの設計思想の評価の他、デジタル技術に対する積極性やツールの継続的活用意思を問うた。

Figure 7 Survey Results on User Satisfaction

5. 結論

本研究は、製造現場におけるデータ駆動型意思決定を促進するため、既存デジタルツールの課題を解決する新たなツールの開発と効果検証を行った。

結果として、Human-in-the-Loop(HITL)アプローチが有効であることが明らかになった。データと感覚の融合を重視した設計は、製造現場ユーザーの安心感と納得感を高め、データ駆動意思決定への行動変容を促す可能性が示唆された。HITL による、「データに基づいて判断するが、必要に応じて感覚や経験に基づく知見を補完する」意思決定が、製造現場で受け入れられる態勢が整っていることも明らかとなった。

さらに、以下の設計要素がデータ駆動意思決定支援ツールの定着と行動変容促進に繋がることも示唆された。

- 心理的障壁低減効果: シンプルな機能と対話型 GUI は、ユーザーのツールへの入り込みやすさを向上させ、デジタルツールに対する積極性を高める効果がある。
- 機能面の有効性: 外挿予測機能はユーザーにとって魅力的な機能であり、データ活用に対する期待感を高める。

6. 展望

心理的障壁低減効果がある「シンプルな機能」と機能面で効果がある「魅力的な機能」という一見相反する要件を両立するためには、シンプルでありながら多くのユーザーが求める機能を提供する必要がある。例えば、実績探索機能において CSV ファイルのダウンロードを可能にすることで、ユーザー自身がデータを追加・編集できるように設計した。この方針は、従来重視されてきたユーザー中心設計やカスタマイズとは異なるが、シンプルかつ汎用的な機能こそが長期的に活用される条件であると考えられる。汎用的なツールであれば、異なる部署やチームでも同じツールを使用でき、意思疎通や情報共有が円滑になり、組織全体としてデータ駆動型の最適な意思決定が可能となる。

本結果は、今回の対象部門同様、製造業における根幹の意思決定を担う部門全般に適用可能と考えられる。

今後、ツールで取得したユーザーの経験知をデータベースに蓄積し、新たな予測モデルに反映させる等、HITL の仕組みをより充実させる予定である。また、本ツールでの設計思想は種々のツールに適用可能であり、ユーザー一人ひとりの断片的な暗黙知を集約し知識ベースを構築することも視野に入れ、大規模な HITL 構築により社内の意思決定全般を支援することを目指す。

参考文献

- [1] IPA 独立行政法人 情報処理推進機構. (2023). DX 白書 2023. IPA 独立行政法人 情報処理推進機構ホームページ. <https://www.ipa.go.jp/files/000108048.pdf> (参照 2025-2-26)
- [2] Gertosio, C., & Dussauchoy, A. (2004). Knowledge discovery from industrial databases. *Journal of Intelligent Manufacturing*, 15(1), 29–37. <https://doi.org/10.1023/B:JIMS.0000010073.54241.e7>
- [3] Harding, J. A., Shahbaz, M., Srinivas, & Kusiak, A. (2006). Data mining in manufacturing: A review. *Journal of Mechanical Science and Engineering*, 128(4), 969–976. <https://doi.org/10.1115/1.2194554>
- [4] 徐良為. (2012). データマイニングと意思決定. *Operations Research: 経営の科学*, 57(5), 276-280.
- [5] 中林暁男, & 和田英彦. (2016). 第 3 次人工知能ブームを機に見る製造業におけるデータ解析技術の変遷. *横河技報*, 59(1), 3-6.
- [6] 黒須正明, 三樹弘之, et al. (2001). ISO13407 がわかる本. オーム社.
- [7] Wu, X., et al. (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems*, 135, 364-381.
- [8] 磯田祐世, 浅田勝義, et al. (2023). 製造業における DX 推進に向けたリテラシー醸成の取り組み. In 情報処理学会 IoT 行動変容学研究グループ第 3 回研究会論文集 (pp. 81-84).
- [9] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.
- [10] 鈴木萌々花, 大町奈央子, et al. (2024). 製造現場の認知構造に基づく UI/UX 設計が製造プロセスの意思決定に与える影響. In AIoT 行動変容学会 第 8 回研究会論文集 (pp. 4-9).

補遺. 機械学習モデルの設計

重回帰モデルをベースに、外挿予測が可能で、ユーザーの感覚と経験知を取り入れた機械学習モデルを考案した。重回帰モデルは、説明変数と目的変数の間に線形の関係を見出すシンプルな手法であるため、比較的外挿予測が可能であるという性質を持つ。ただし、外挿予測の精度が担保されるのは次の場合である。

- 説明変数と目的変数の間に線形の関係があること
- 多重共線性がないこと

説明変数には成分、製造条件(温度、加工度等)を用い、目的変数には引張強度、耐力、伸びの機械的特性を用いた。事前のデータ解析により、説明変数と目的変数は完全に線形の関係ではない場合があるものの、極端に線形関係から外れるものはないことが分かっているため、似た成分・加工条件、つまり狭い領域において線形関係を有していると近似しても問題ないと判断した。よって合金系、調質毎にモデル構築を実施することで対応した。

多重共線性がある、とは説明変数同士の相関があることを意味する。重回帰モデルが各変数の目的変数に対する線形の影響を係数として算出する以上、多重共線性があると個別の変数の影響(係数)を正確に切り分けて算出することができない。そこで、ユーザーの感覚と経験知を活用し、多重共線性の影響を軽減する手法を考案した。本手法では、多重共線性が認められる場合はユーザーにどちらの変数の影響が大きいかを選択させることで、実際に影響する変数を明確化する。その結果全説明変数の目的変数に対する線形の影響(係数)を正確に算出することが可能となり、外挿領域の予測が可能となる。

未登録 PoI 探索のための 時系列ソーシャルメディアデータに基づく位置推定手法の提案と評価

澤野耕平^{†1} 松田裕貴^{†2} 中谷響^{†3}
 奈良先端科学技術大学院大学 岡山大学 奈良先端科学技術大学院大学

大内啓樹^{†4} 諏訪博彦^{†5} 安本慶一^{†6}
 奈良先端科学技術大学院大学 奈良先端科学技術大学院大学 奈良先端科学技術大学院大学

1. はじめに

近年、ソーシャルメディアの普及により観光地にて観光客がソーシャルメディア投稿を行う機会は多い。観光客はソーシャルメディアを用いて旅行体験を投稿している。このような使用が行われる中で、ソーシャルメディアには観光客が興味を引く場所である Point of Interest (PoI) に関する情報が多く集積している。この集積した情報があるソーシャルメディアを一種の観光情報データベースとし、観光地での旅程選定の参考にするような情報検索ツールとしての使用もされている [1]。具体的には Instagram の「地図」機能である。この機能は現在位置周辺の位置情報付与された投稿を表示することができ、これを用いて観光客が近隣の PoI を検索することができる。このようにソーシャルメディア投稿が言及している位置を推定することは観光地での意思決定において重要である。

ソーシャルメディアの投稿データに位置情報を付与する方法には、デバイスの GPS データを利用するものと、投稿に対して地図データベース上の PoI をタグ付けるものがあるが、これらの方法にはそれぞれ課題が存在する。前者は、GPS 制度に大きく依存するという点や、実際に GPS データが付与されている投稿が全体の 0.4% ほどごく少数という点が挙げられる。後者は、既存の地図データベース（地理 DB）を参照するためにデータベースに載っていない PoI には対応できないという点が挙げられる。Instagram の「地図」機能は地図データベース上に存在する PoI をタグ付けする機能を使用して既存の地理 DB（Google Map や OSM

など）に登録されている位置情報を参照して投稿に位置情報を埋め込んだ物を地図上に可視化した機能である。

既存の PoI 位置推定手法は投稿した地点の画像や投稿データに付与された位置情報を入力データとして使用している。また、既に地理 DB に登録されている PoI をクラスとし、投稿をクラス分類することで位置推定を行っている。位置情報の有無や地理 DB の登録された範囲に位置推定が限定されており、位置情報を付与している投稿データが少ない中で、これらの既存手法を適用するのは制限されたデータに対してしか行うことができない。

そこで本研究では観光地でのソーシャルメディア投稿から観光客の時系列での移動経路と相対的な位置関係を用いることで、既存の地図データベースの範囲や投稿の位置情報に依存することなく、観光客にとって関心の高い場所である PoI の位置を推定する手法を提案する。

2. 関連研究

2.1. 文章から位置推定する手法（ジオパーズング）

文章中に含まれる場所に対する位置推定の既存の手法は大きく二つに分けられる。一つは文章中のテキストから地名を抽出し GeoNames や OSM などの地理 DB にある場所にクラス分けをする分類問題としてのアプローチである。もう一つは文章中のテキストから緯度経度を回帰する回帰問題としてのアプローチである。分類問題としての一般的なアプローチとしてこれまでには自然言語による研究がされてきた [2, 3, 4]。これらの研究では地球表面をグリッド分割し、テキストの地名が属するグリッドを予測することで、位置推定を行っている。

ソーシャルメディア投稿のテキストを用いた研究として、Li らは、ユーザの自由記述テキストと投稿プラットフォームといったカテゴリカルデータを BERT ベースのエンコーダで埋め込み、時間情報を階層型に分割しそれぞれエンコーディングすることで、これらのベクトル表現の特徴間の相関を学習させ、最終的に PoI が地理データベースの中の地理

Location Estimation Method using Time-Series Social Networking Service Data for Exploring Unregistered PoI

^{†1} KOHEI SAWANO, Nara Institute of Science and Technology

^{†2} YUKI MATSUDA, Okayama University

^{†3} HIBIKI NAKATANI, Nara Institute of Science and Technology

^{†4} HIROKI OUCHI, Nara Institute of Science and Technology

^{†5} HIROHIKO SUWA, Nara Institute of Science and Technology

^{†6} KEIICHI YASUMOTO, Nara Institute of Science and Technology

エンティティに属するかの確率を出力している。時間情報の階層的な表現により投稿の時間的なパターンを捉えやすくなり、既存の手法に比べて推定精度が向上したと報告している [5]。これらの分類問題としてのアプローチは、既存の GeoNames や Wikipedia といった地名データベースの範囲でしか分類ラベルを作れないという制限が存在している。

直接テキストから緯度経度を推定する回帰問題としてのアプローチの一つは、確率モデルを出力することによるアプローチである [6, 7]。Benjamin らは wikipedia データで学習し、出力に混合分布による空間的な確率分布を出力する。出力される確率分布は予測される緯度経度を中心としており、複数の分布が予測結果として出力された場合には混合分布が付与されることで、曖昧な地理表現などにも柔軟に対応可能なことを示している [6]。

2.2. ソーシャルメディア投稿データを用いた位置推定

Dai らは、ソーシャルメディア投稿のテキストとハッシュタグなど複数のデータを用いるマルチモーダル表現学習フレームワーク (MRLF) を提案している。このフレームワークは画像、テキスト、ハッシュタグの間の特徴を相互に学習することで、位置推定の向上がされたと報告している。特に投稿の PoI に属する分類精度では平均 8% の精度向上を実現し、分類された PoI と実際の投稿位置との間の距離では平均誤差約 900m を実現している [8]。

2.3. 本研究の位置付け

既存研究の分類手法は GeoNames などの地理 DB のエンティティをラベルとして使用するものであるため、地理 DB にないエンティティは位置推定できない。また、回帰手法については学習にジオタグ付きのソーシャルメディアデータを使用しており、このようなデータを得ることは難しいということ、既存のソーシャルメディアデータの持つジオタグは地理 DB に依存したものであるため、回帰手法においても地理 DB の範囲による制限は存在している。また、既存のテキストから直接、緯度経度を回帰している手法の誤差は数十 km オーダーであり、PoI レベルの位置推定への応用は難しいという課題がある。これらの問題点を踏まえて、本研究では観光地での時系列ソーシャルメディア投稿を用いた PoI 位置推定手法を提案する。観光地においてソーシャルメディアユーザは興味を引いた地点についての投稿を行いながら観光していると仮定する。この時系列の投稿列には既存のジオコーディングモデルでは位置推定できないような PoI についての投稿も存在している (例. 地理 DB に存在しないかもしくは地理 DB に存在するあるエリア内のポイントなど)。ジオコーディング可能な PoI を Geocoding PoI (GPoI)、不可能な PoI を Target PoI (TPoI) と定義

し、TPoI と GPoI の時系列的関係から、移動可能範囲を算出しそれらを複数組み合わせることで PoI レベルのジオコーディングの実現を目指す。

3. 位置推定フレームワーク

本章では、まずはじめに我々の提案する時系列ソーシャルメディア投稿を用いた位置推定フレームワークの概要について説明する。その後、本研究の主な貢献となる位置推定アルゴリズムの概要と詳細な構成について説明する。

3.1. 提案フレームワークの概要

本節では、本研究が提案する位置推定フレームワークの概要と構成について述べる。本研究は観光地において観光客が目についたものをリアルタイムでソーシャルメディアに投稿しながら観光しているという過程を置いた上で、いくつかの投稿が言及する PoI をアンカーとしてアンカー投稿と非アンカー投稿 (位置推定の対象となる投稿) の時系列関係から移動可能範囲を推定しその範囲をユーザごとに重ね合わせ、最も重なった部分の重心を推定ポイントとして位置推定することを目的とする。

3.2. データ収集

本節では、提案手法で用いたデータ収集のためのソーシャルメディアプラットフォームについて述べる。

提案フレームワークでは、データ収集実験に用いるソーシャルメディアとして松田らのレポット [9] を用いる。レポットとは、スマートフォンアプリケーションであり、写真を撮影し、その写真に対してコメントを投稿するアプリケーションである。このレポットを用いて実験参加者に対して観光地において観光しながら目に留まったものを写真撮影し、投稿してもらうという実験を行った。

3.3. ジオコーディング

本節では、提案フレームワークの中のテキストから緯度経度を推定するモデルについて説明する。

ジオコーディングに使用するモデルは、中谷ら [10] が提案している BERT ベースのジオコーディングモデルを用いる。中谷らは、自由記述文章を入力としたジオコーディングタスクが抱える課題である地名の曖昧性解消のために、文章中に登場する地名表現部分だけでなく前後の文章を含めたベクトル化と地理 DB である OSM のエントリをそれぞれ異なるエンコーダでベクトル化することを提案している。このモデルは、入力として受け取ったテキストに対して緯度経度のペアを出力するが、出力結果は推定確率の高い順にランク形式になっている。本研究では、最も確率の高い

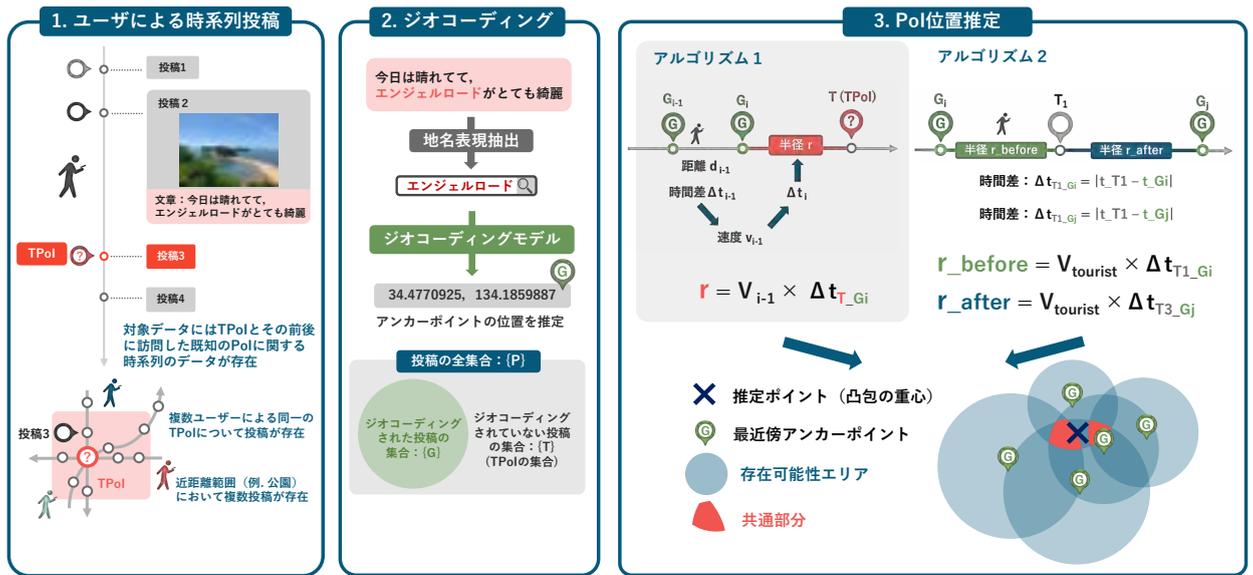


図1 提案フレームワークの全体図

上位1位の推定緯度経度を用いる。

3.4. 位置推定アルゴリズム

本節では、提案フレームワークのうち、アンカーとなる投稿を用いた位置推定アルゴリズムについて説明する。本研究では、位置推定について二つのアルゴリズムを実装し、その精度を評価した。手法1は、位置推定対象のPoI (Target PoI: TPoI) に関する投稿とその直前および直後のジオコーディングにより緯度経度が推定された2投稿から、TPoI訪問前後のユーザーの局所的な移動速度を算出しこの移動速度を用いてユーザーの移動可能範囲を算出するものである。手法2は全ユーザーの移動軌跡から算出した観光地での平均移動速度を各ユーザーの移動速度として近似的に用いて移動可能範囲の算出を行うものである。手法1,2において算出したユーザーの移動可能範囲内のTPoIの存在確率は一様分布に従うと仮定している。次節からそれぞれの具体的なアルゴリズムについて詳細を説明する。

3.4.1. 変数定義

本節では提案する手法1,2で用いるデータ集合の定義を行う。本研究の実験で得られたソーシャルメディアの投稿の集合とその要素を、

$$P = \{p_j \mid \forall j \in Z\}, \quad p_j = \begin{pmatrix} \text{caption,} \\ \text{timestamp,} \\ \text{latitude,} \\ \text{longitude,} \\ \text{geo latitude,} \\ \text{geo longitude,} \\ \text{cos similarity} \end{pmatrix}, \quad (1)$$

- latitude, longitude : レポットで投稿した時に自動で付与される緯度経度 (正解データとしても用いる)
- geocoding latitude, geocoding longitude, cosine similarity : ジオコーディングモデルによって推定された緯度経度とコサイン類似度 (ジオコーディングされた投稿のみ値を持ち、それ以外の投稿は NaN)

と定義する。また、集合 G をジオコーディングモデルによってジオコーディングされた投稿の集合と定義する。このとき、

$$G \subseteq P, \quad (2)$$

である。また、集合 P の要素かつ、集合 G の要素でない全ての投稿の集合を T と定義する。

$$T = \{p \in P \mid p \notin G\}, \quad G \cap T = \emptyset \quad (3)$$

3.4.2. 手法1: 時系列ウィンドウを用いた局所速度による位置推定

手法1では推定対象の直前および、直後の2投稿 ($p \in G$) とTPoIに関する投稿 ($p \in T$) の5投稿を1単位とする時系列のウィンドウを定義し、ウィンドウ内で局所速度を計算

し位置推定を行う。この手法の詳細についてアルゴリズム全体の流れを疑似アルゴリズムとして示し (Algorithm1), その中で使用する各関数の詳細について以下に示す。

Algorithm 1 位置推定全体フロー

```

1:  $df \leftarrow \text{ReadRawData}(\text{DataPath})$ 
2: 局所速度計算を行い, PoI クラスごとの円を算出
3:  $\text{PoiClassDict} = \{\}$ 
4: for  $i \in \text{length}(df)$  do
5:   if  $i \bmod 5 = 0$  then
6:      $\text{window\_df} = \text{MakeWindow}(df)$ 
7:      $p_{-2} = \text{window\_df}[0] \in G$ 
8:      $p_{-1} = \text{window\_df}[1] \in G$ 
9:      $p_0 = \text{window\_df}[2] \in T$ 
10:     $p_{+1} = \text{window\_df}[3] \in G$ 
11:     $p_{+2} = \text{window\_df}[4] \in G$ 
12:     $v_{\pm} = \text{ComputeVelocity}(p_{\pm 2}, p_{\pm 1})$ 
13:     $\text{center\_list} = p_{\pm 1}$ 
14:     $\Delta t_{\pm} = \text{TimeDiff}(p_{\pm 1}, p_0)$ 
15:     $\text{radius\_list} = \text{ComputeRadius}(v_{\pm}, \Delta t_{\pm})$ 
16:     $\text{center\_list} = p_{\pm 1}$ 
17:     $\text{PoiClassDict}[p_0["\text{PoiClass}"]] = (\text{center\_list},$ 
       $\text{radius\_list})$ 
18:   end if
19:    $\text{window\_df.delete}(0)$    ▷ ウィンドウをスライド
20: end for
21: グリッドマップを計算し, 最も高い値を持つエリアを抽出
22:  $\text{GridMap\_dict} = \text{BuildGridmap}(\text{PoiClassDict})$ 
23: for  $\text{gridmap} \in \text{Gridmap\_dict}$  do
24: 抽出したエリアの凸包を計算したのちに, 重心を計算
25:    $\text{Centeroid} = \text{ComputeCenteroid}(\text{gridmap})$ 
26: 計算した重心の値と正解位置のユークリッド距離を計算
27:    $\text{Result} = \text{EuclideanDistance}(\text{Centeroid},$ 
      $\text{GroundTruth})$ 
28:    $\text{ResultTable.add}(\text{Result})$ 
29: end for
30: return  $\text{ResultTable}$ 

```

1. 時系列ウィンドウの構築と処理: Algorithm1, 6 行目
 - **MakeWindow** 関数で投稿のインデックスを順次 window_list に追加する。
 - window_list の長さが 5 になった時点で, 中央のインデックスを参照し, 前後 2 行を含む計 5 行のデータ df を取得し, window_df として返す。
2. 速度および半径の計算: Algorithm1, 7-20 行目
 - window_df の各投稿を $p_{-2}, p_{-1}, p_0, p_{+1}, p_{+2}$

とする。 $(p_{-2}, p_{-1}, p_{+1}, p_{+2}) \in G, p_0 \in T$

- 各投稿間の時間差を $\Delta t_{\pm 1}$ と定義する。

$$\Delta t_{\pm 1} = \text{TimeDiff}(p_{\pm 1}, p_0) \quad (4)$$

- 各投稿の位置情報から距離を計算し, 局所速度 v_{\pm} , 半径 r_{\pm} を算出する。(※ **ComputeVelocity** は 2 投稿から局所速度を計算する関数)

$$v_{\pm} = \text{ComputeVelocity}(p_{\pm 2}, p_{\pm 1}) \quad (5)$$

$$r_{\pm} = v_{\pm} \times \Delta t_{\pm 1} \quad (6)$$

3. 円の計算とグリッドマップの生成: Algorithm1, 21-22 行目

p_{\pm} を中心とし半径 r_{\pm} の範囲をユーザ移動可能範囲とする。各移動可能範囲に対して, 範囲内の各ピクセルの値を以下で定義する任意の重み (w) に更新する (**BuildGridmap**)。

ここで取りうる重み (w) を以下のように定義する。

$$w = \begin{cases} 1, & \text{count の場合,} \\ \pi r^2 & \text{area の場合} \end{cases} \quad (7)$$

$$H(x, y) = H(x, y) + w \quad (8)$$

重みは 2 つの重みを定義した。一つは単純に移動可能範囲が最も重なったところに TPoI が存在する可能性が高いという考えのもとで, 移動可能範囲の重なった回数による重み (count) を定義した。もう一つは移動可能範囲の大きさも考慮に入れた重みで, 移動可能範囲が大きい場合には存在する範囲が大きくなり, 推定の確信度が低くなるという考えから, 移動可能範囲の面積の逆数を重みとする重み (area) を定義した。

4. 推定ポイントの抽出と結果の出力: Algorithm1, 24-29 行目

- グリッドマップの各セルから最も値の高いセルを特定し, そのセルを全て内包するような凸包を算出する。(**ComputeCenteroid**)
- 凸包の重心を算出し, 重心と正解位置とのユークリッド距離を計算する。(**EuclideanDistance**)

3.4.3. 手法 2 : 平均的観光客移動速度による位置推定

アルゴリズム 1 が時系列ウィンドウにより局所的なユーザの移動速度を個別に計算し, 位置推定を行っていたのに対して, アルゴリズム 2 では, 観光中の観光客の歩行速度は平均速度で近似できるという仮定のもと, ユーザ全体の移動軌跡から観光地での平均歩行速度を求め, それを用いた移動可能範囲を計算し位置推定を行うものである。観光客の平均的な移動速度を用いた移動可能範囲計算による位置推定フローの概要を以下に示す。

Algorithm 2 位置推定全体フロー

```

1:  $df = \text{ReadRawData}(DataPath)$ 
2: ステップ 1: 全ユーザの平均移動速度を計算
3:  $v\_list = []$ 
4: for  $user\_df \in df.groupby(user)$  do
5:   for  $i \in user\_df$  do
6:      $v\_i = \text{ComputeVelocity}(user\_df[i], user\_df[i+1])$ 
7:      $v\_list.add(v\_i)$ 
8:   end for
9: end for
10:  $V\_tourist = v\_list.average$ 
11: ステップ 2: POI クラスごとの円情報を計算
12:  $PoiClassDict = \{\}, radius\_list = []$ 
13:  $center\_list = []$ 
14: for  $i \in length(df)$  do
15:    $p_i = df[i] \in G$ 
16:    $p_{i\pm n} = df[i \pm n] \in T$  ▷ n は任意の値
17:    $t_{\pm n} = \text{TimeDiff}(p_i, p_{i\pm n})$ 
18:    $r_{i\pm n} = \text{ComputeRadius}(t_{\pm n}, V\_tourist)$ 
19:    $radius\_list.append(r_{i\pm n})$ 
20:    $center\_list.append(p_i)$ 
21:    $PoiClassDict[poiclass] = radius\_list, center\_list$ 
22: end for
23: ステップ 3: POI クラスごとのヒートマップを構築
24:  $Gridmap\_dict = \text{BuildGridmap}(PoiClassDict)$ 
25: for  $gridmap \in Gridmap\_dict$  do
26: 抽出したエリアの凸包を計算したのちに、重心を計算
27:    $Centeroid = \text{ComputeCenteroid}(gridmap)$ 
28: 計算した重心の値と正解位置のユークリッド距離を計算
29:    $Result = \text{EuclideanDistance}(Centeroid, GroundTruth)$ 
30:    $ResultTable.add(Result)$ 
31: end for
32: return  $ResultTable$ 

```

1. ユーザごとの速度を計算し、観光客の平均歩行速度を計算:Algorithm2, 2-10 行目
 - **ComputeVelocity** を用いて、投稿データフレームに含まれる各ユーザの連続投稿から移動距離と移動時間を求め、平均歩行速度を算出。
 - ユーザごとの平均歩行速度の歩行速度の平均値を「歩行観光客の平均歩行速度」として算出。
2. POI クラスごとの円情報を計算:Algorithm2, 11-22 行目
 - **TimeDiff, ComputeRadius** により、連続する「ジオコーディングにより得られた投稿」の間に位

置する「未確定投稿」に対して、円の中心座標と半径を算出する。

- 得られた (中心座標, 半径) を POI クラス *poiclass* ごとに *PoiClassDict* という辞書形式で保存。

3. POI クラスごとのグリッドマップを構築:Algorithm2, 23-25 行目

- **BuildGridmap** を用いて、*PoiClassDict* 内のすべての (中心座標, 半径) を *poiclass* ごとに指定範囲のグリッド *gridmap* を構築。
- *gridmap* の構築時には手法 1 と同様の重み (count, area) を用いて各セルの値を更新。
- POI クラス別に二次元配列 *gridmap* を得て、*Gridmap_dict* として格納。

4. 凸包や重心の抽出:Algorithm2, 26-31 行目

- グリッドマップの各セルから最も値の高いセルを特定し、そのセルを全て内包するような凸包を算出。 (**ComputeCenteroid**)
- 凸包の重心を算出し、重心と正解位置とのユークリッド距離を計算。 (**EuclideanDistance**)

4. 実験

ここまでで提案した手法を評価するために、実際の観光地である、奈良県奈良市のならまちエリア (詳しい範囲は後述) における観光実験を行った。本章では実験の概要について示す。

4.1. 実験概要

本実験は提案手法の有効性と中谷ら [10] の提案しているジオコーディングモデルの結果をアンカーポイントとした条件下での手法 1, 2 の精度について検証することを目的として行った。2024 年 6 月 1 日, 2 日にわたって、奈良先端大の学生 16 人を募り、奈良県奈良市のならまちエリアの奈良女子大, 東向商店街, 元興寺, 春日大社, 東大寺二月堂を囲うようなエリアにおいて、松田ら [9] が開発したソーシャルメディアであるレポットを用いて観光しながら目についたものを投稿してもらいデータ収集を行なった。本実験により得られた投稿数は 1617 件であった。

4.2. 投稿要件

本実験を行うにあたり、提案手法を適用するのに十分な投稿数、ジオコーディング可能な文字数の確保。そして投稿内容の性質がジオコーディング可能なものとそうでないものどちらかに偏りすぎないように、以下の投稿要件と投稿対象 PoI の例を示し実験参加者に説明した。投稿要件と投稿対象の例を以下の図 2, 3 に示す。



図2 “広く一般に知られているもの”の例



図3 “一般に知られていない珍しいもの”の例

- 投稿件数：100 件以上（1 人当たり）
- 投稿文字数：45 文字程度
- 投稿対象（以下のような PoI を 50% ずつの割合で投稿することが望ましい）
 1. 対象が広く一般に知られているもの（例，東大寺の大仏，春日大社など）
 2. 一般に知られていない珍しいもの（例，ご当地マンホールなど）

4.3. 結果・考察

本実験では計 1617 件の投稿が得られた。そのうち、ジオコーディングモデルによりジオコーディングされた投稿は 829 件であった。

本研究で得られたデータのうち、同じ PoI について投稿されているものにラベルづけを行うためにアノテーションツールを用いて全ての投稿に対して PoI のラベリングを行い、941 件の推定対象 PoI (TPoI) を同定した。そのうち、複数ユーザが言及しているものは 229 件であった。

4.3.1. 手法 1（アンカーポイントの緯度経度としてレポートに付与されているデータを用いたもの）

本節ではアンカーポイントの緯度経度が 100% 正しくジオコーディングされたという仮定のもとで手法 1 を適用した結果を示す。表 1 はアノテーションツールによって同定された TPoI の重みごとの位置推定結果である。アルゴリズム 1 では、アンカーポイントに理想的な GPS データを用いた場合、重みを「カウント」とした手法が最も高い精度を示し、中央値が 37.00m、平均値が 56.58m であった。この結果から、外れ値に対して頑健であり、投稿数が多い環境で特に有効な手法であることが確認された。半径および面積を重みとした場合に精度が低くなるのは使用している

アンカーポイントの GPS 精度の誤差に影響される可能性が高い。

表 1 推定エリアの凸包の重心から正解位置の距離の平均値、中央値、標準偏差（単位：m）

重み	平均値	中央値	標準偏差	四分位範囲
カウント	56.58	37.00	62.51	53.03
面積	67.60	40.51	88.45	60.85

4.3.2. 手法 1（アンカーポイントの緯度経度としてジオコーディングモデルの出力結果上位 1 件の場所を用いたもの）

本節では、アンカーポイントとしてジオコーディングモデルの推定結果（上位 1 件）を採用した場合にアルゴリズム 1 を適用した結果について述べる。表 2 に示されている通り、重みとして「面積」を使用した場合が中央値 145.21m、平均値 205.53m と、最も精度の高い結果を示した。一方で、重みを「カウント」や「半径」とした場合には、いずれも中央値や平均値がやや高く、標準偏差および四分位範囲が広い結果となった。これらの結果は、アンカーポイントの精度が完全には保証されていない環境では、円の信頼度を面積に基づいて調整する手法が有効であることを示唆している。

ただし、面積を重みとする手法は、標準偏差が大きく、外れ値が依然として目立つことから見受けられる。これは、アンカーポイントにおけるジオコーディングモデルの誤差が累積し、円の位置や大きさ、信頼度に影響を与えたためと考えられる。

閾値を高く設定した場合アンカーポイントの信頼性が向上するが、その分、利用可能なアンカーポイントの数が減少し、推定精度に悪影響を及ぼす可能性がある。一方で、閾値を低く設定すると、低精度のアンカーポイントが含まれることで外れ値の影響が増加する一方、円の数が増えることで精度が改善する場合もある。これらのトレードオフを考慮し、最適な閾値を設定することが重要である。

4.3.3. 手法 2(アンカーポイントの高精度 GPS を想定した場合)

まず、ジオコーディング結果が 100% 正確に得られると仮定し、GPS 精度が 50m 以下のすべてのポイントをアンカーポイントとして用いた結果を表 3 に示す。重みを「面積」とした場合、中央値が 82.47m、平均値が 134.64m と最も高い精度を示した。さらに、四分位範囲 (IQR) が 101.11m と相対的に狭く、推定値が集中している様子がうかがえる。一方、標準偏差は 180.54m と大きく、外れ値が一定数含まれていることも示唆される。

表 2 推定位置と正解位置の距離の平均値, 中央値, 標準偏差, 四分位範囲 (単位: m)

重み	平均値	中央値	標準偏差	四分位範囲
カウント	222.99	145.21	205.09	206.16
面積	205.53	145.21	198.62	171.07

表 4 アンカーポイントの緯度経度に GPS データを用いた場合 (GivenPoI 集合) の推定エリアの凸包重心から正解位置の距離 (単位: m)

重み	平均値	中央値	標準偏差	四分位範囲
カウント	160.05	101.53	182.09	152.88
面積	156.40	100.00	178.63	147.26

表 5 アンカーポイントの緯度経度にジオコーディングの結果を用いた場合 (FilteredGeocoding 集合) の推定エリアの凸包重心から正解位置の距離 (単位: m)

重み	平均値	中央値	標準偏差	四分位範囲
カウント	401.18	263.32	366.54	475.75
面積	420.14	215.80	434.96	497.42

表 3 推定エリアの凸包の重心から正解位置の距離の平均値, 中央値, 標準偏差, 四分位範囲 (単位: m)

重み	平均値	中央値	標準偏差	四分位範囲
カウント	139.59	89.70	162.30	118.14
面積	134.64	82.47	180.54	101.11

このことから、アンカーポイント同士の円の重なり「面積」が大きいほど位置推定への寄与度を高く設定する手法は、多くのアンカーポイントが円として重なる領域を強調でき、より精度の高い推定が期待できると考えられる。しかし、アンカーポイント自体に外れ値が含まれる場合には、その影響が大きく出る可能性も示唆される。

4.3.4. アンカーポイントの信頼性と閾値の影響

次に、アンカーポイントの緯度経度の信頼性が推定精度に及ぼす影響を検証するため、GPS 精度の閾値を変化させた場合の推定結果を比較した。具体的には、ジオコーディング可能な PoI の集合である GPoI 集合のうち、ある閾値以下の GPS 精度をもつポイントの元 GPS 情報をアンカーポイント (GivenPoI 集合) とした場合と、それらのポイントをジオコーディングモデルにより推定した緯度経度 (FilteredGeocoding 集合) をアンカーポイントとした場合で推定精度を比較する。ここでは閾値を 50m に設定した例を表 4 および表 5 に示す。

GivenPoI 集合 (表 4) では、面積重みの中央値が 100.00m

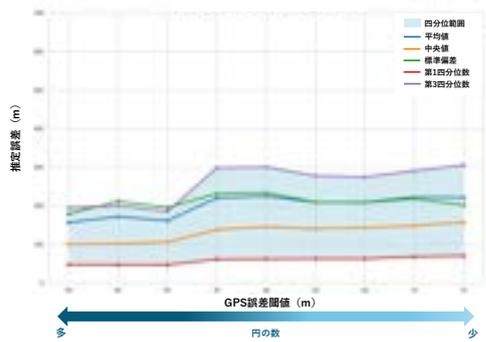


図 4 アンカーポイントの緯度経度に GPS データを用いた場合の各統計量の変化

で、PoI 推定に十分活用できるレベルの精度が得られている。これは元の GPS データが高い信頼度をもつ場合、円同士の重なり面積を重視する手法が有効に働くことを示唆している。一方、FilteredGeocoding 集合 (表 5) では、面積重みの中央値が 215.80m と大きく、ジオコーディングに伴う推定誤差が位置推定精度に大きく影響していることが分かる。

さらに、GPS 精度の閾値を変動させたときの推定誤差の統計量の変化を図 4 および図 5 に示す。理想的な GPS データ (GivenPoI) を用いる場合 (図 4) には、閾値を緩和するとアンカーポイント数が増加し、推定精度の平均値・中央値が改善するとともに、四分位範囲も狭まる傾向が確認できる。これは、アンカーポイント (円) の総数が増えることで、それらの重なり領域をより適切に抽出できるためと考えられる。

一方、ジオコーディングモデルの結果を用いた場合 (図 5) では、GPS 精度の閾値を緩和すると誤差の大きい推定値が混在しやすくなり、平均誤差が悪化するとともに、四分位範囲が広がる傾向がみられる。しかし、中央値の変化は比較的小さく、ある程度の範囲内での外れ値に対して本手法が頑健な一面も確認できる。

以上の結果から、以下の知見が得られる。

1. アンカーポイントの信頼性が高い (GPS 精度が高い) 場合には、円同士の重なり面積を重みに組み込むことで推定精度を向上できる。これは推定結果の中央値や四分位範囲の改善が顕著であり、PoI 推定の有効性が示唆される。
2. 一方で、外れ値が含まれた場合やジオコーディングモデルによる推定誤差が大きい場合には、面積重みを用いることでかえって大きな外れ値を生む可能性がある。すなわち、ジオコーディング結果の誤差が大きいアンカーポイントを多数含む状況では、推定精度のばらつきが増加し、平均誤差が悪化することが分かった。

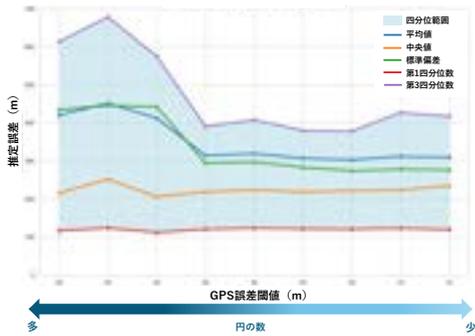


図5 アンカーポイントの緯度経度にジオコーディングの結果を用いた場合の各統計量の変化

5. 結論

本研究では、既存の PoI 位置推定では位置推定することのできない GoogleMap や OpenStreetMap といった地図 DB に登録されていない未登録の PoI の位置推定を行うことを目的として、観光地での観光客の時系列ソーシャルメディア投稿を用いた位置推定フレームワークを提案した。提案フレームワークの検証のために、奈良県奈良市において実験を行った。奈良県ならまちエリアを対象として、被験者 16 人で約 1600 の投稿データに対して、提案手法 1, 2 を適用し比較検討をおこなった。その結果、手アンカーポイントとして得られる緯度経度がジオコーディングモデルによって 100% の精度で得られたという仮定のもとで手法 1 を適用した際の推定誤差は 37m (中央値)、アンカーポイントにジオコーディングモデルの推定結果を用いた状態で手法 1 を適用した際の未登録 PoI の推定誤差は 145.2m (中央値) であった。また、手法 2 においてアンカーポイントを緯度経度がジオコーディングモデルによって 100% の精度で得られたという仮定のもとでの推定誤差は 82.5m (中央値)、ジオコーディングモデルが推定したポイントをアンカーポイントとした場合の推定誤差は 215.8m (中央値) であった。これらの結果は、本提案手法は位置推定対象が地図データベースに登録されているかどうかに関わらず、既存手法と同等の位置推定精度を示しており、地図データベースに未登録の PoI を推定する手法として有効であることが示された。

今後の展望として、本提案手法で行われた位置推定に対して不確実性も考慮するために推定出力を混合確率分布とすることが考えられる。提案手法の円を混合確率分布とすることで、円を描く範囲にどの程度の確率で推定対象のポイントが含まれているかを確率的に表すことができると考えられる。また、現状はジオコーディングモデルの上位 1 件を推定ポイントとして使用しているが、使用しているジオコーディングモデルは日本語旅行記データセットにおいて、上位 10 件のジオコーディングで 87.7% の精度を達成し

ており、これら全てのジオコーディングポイントを使用することで、より多くの円を描くことができそれによりさらなる推定精度の向上が期待できると考えられる。

参考文献

- [1] 福井一喜：東京大都市圏に居住する若者の観光・レジャーにおける SNS 利用—「SNS 映え」を超越する若者たち—, *E-journal GEO*, Vol. 14, No. 1, pp. 1–13 (オンライン), 10.4157/ejgeo.14.1 (2019).
- [2] Hulden, M., Silfverberg, M. and Francom, J.: Kernel Density Estimation for Text-Based Geolocation, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29, No. 1 (online), 10.1609/aaai.v29i1.9149 (2015).
- [3] Roller, S., Speriosu, M., Rallapalli, S., Wing, B. and Baldrige, J.: Supervised Text-based Geolocation Using Language Models on an Adaptive Grid, *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Association for Computational Linguistics, pp. 1500–1510 (2012).
- [4] Wing, B. and Baldrige, J.: Hierarchical Discriminative Classification for Text-Based Geolocation, *Conference on Empirical Methods in Natural Language Processing* (2014).
- [5] Li, M., Lim, K. H., Guo, T. and Liu, J.: A Transformer-Based Framework for POI-Level Social Post Geolocation, *Advances in Information Retrieval*, Springer Nature Switzerland, pp. 588–604 (2023).
- [6] Radford, B. J.: Regressing Location on Text for Probabilistic Geocoding, *Proceedings of the 4th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2021)*, Association for Computational Linguistics, pp. 53–57 (2021).
- [7] Iso, H., Wakamiya, S. and Aramaki, E.: Density Estimation for Geolocation via Convolutional Mixture Density Network, (online), 10.48550/arXiv.1705.02750 (2017).
- [8] Dai, R., Luo, J., Luo, X., Mo, L., Ma, W. and Zhou, F.: Multi-modal Representation Learning for Social Post Location Inference, *ICC 2023 - IEEE International Conference on Communications*, pp. 6331–6336 (online), 10.1109/ICC45041.2023.10279649 (2023).
- [9] 松田裕貴, 河中祥吾：Web ブラウザ上で動作する市民参加型写真収集アプリの開発と運用, 第 27 回社会情報システム学シンポジウム (ISS27), pp. 1–5 (2021).
- [10] Nakatani, H., Teranishi, H., Higashiyama, S., Sawada, Y., Ouchi, H. and Watanabe, T.: A Text Embedding Model with Contrastive Example Mining for Point-of-Interest Geocoding, *Proceedings of the 31st International Conference on Computational Linguistics*, Association for Computational Linguistics, pp. 7279–7291 (2025).

チェロの遠隔練習支援システムの試作と検討

太田健^{†1}角薫^{†2}公立はこだて未来大学^{†1}公立はこだて未来大学^{†2}

1. はじめに

弦楽器の習得は、初心者にとって極めて高い難易度を伴う。特に、独学による学習は困難であり、一般的には教則本や動画教材が活用される。しかし、これらの教材は主に視覚的な情報に依存するため、初心者が正確な演奏技術を効率的に習得することが難しいという課題がある。ヤマハの調査によれば、20代から50代の500人のうち約40%が楽器の学習を開始したものの、途中で挫折した経験があると報告されている[1]。

特にチェロのような弦楽器においては、ピアノの鍵盤やギターフレットのような明確なガイドが存在しないため、演奏者は自身の耳を頼りに音程を調整する必要がある。このため、適切な音程を保つことが難しく、演奏技術の習得には時間を要する。また、弓の持ち方や運弓技術の習得には、手指の微細な動作の調整が求められ、対面レッスンにおいても教師がこれらの技術を適切に伝えることは容易ではない。チェロはヴァイオリンなどと比較して大型の楽器であり、持ち運びが困難であるため、対面での指導を受ける機会が限られる。このため、遠隔指導のニーズが高いと考えられる。

しかし、遠隔でのチェロの指導に関する研究は限られている。遠隔教育における器楽指導の手法については一部の研究で検討されており、映像を用いた指導においては、モニター上での視認性の向上が求められることが指摘されている[3]。また、リモート環境下でのチェロレッスンの実践例も報告されており、オンラインレッスンの有効性が示唆されている[4]。しかし、これらの研究の多くは、ビデオ通話を利用した従来型の遠隔指導に関するものであり、3次元視覚補助を活用した指導方法については十分に検討されていない。

本研究では、弦楽器の習得におけるこれらの課題を解決するため、MR (Mixed Reality, 複合現実) 技術[2]を活用した遠隔レッスンシステムの開発と評価を行う。MRは、現実世界に仮想情報を重ね合わせ、ユーザがそれとインタラクションできる技術であり[5]、医療や教育分野においても応用が進められている[6][7]。本システムでは、演奏技術の指導を3次元空間に投影し、学習者がリアルタイムに視覚的・直感的に理解できる環境を構築することを目指す。従来の

教則本や動画教材とは異なり、MRを活用することで、実際の演奏空間に指導情報を重ね合わせることが可能となり、より没入感のある学習体験が実現できると考えられる。

楽器練習支援に関する先行研究として、個人練習環境の最適化に関する研究がある。例えば、楽器演奏時の心理的負担を軽減するために、演奏者の音量を調整し、周囲の奏者との音のバランスを最適化するシステムが提案されている[8]。この研究では、演奏レベルの異なる奏者が同一空間で練習する際に生じる心理的プレッシャーを軽減することで、学習効率の向上が示唆されている。また、楽器のチューニング支援技術として、スマートグラスを活用した音程調整システムの開発が行われており、従来の譜面台に設置するチューナーと比較して、姿勢や視線の負担を軽減できることが報告されている[9]。これらの研究は、演奏技術の向上を支援する技術の有用性を示しており、特に視覚的な補助が学習効率を高める可能性があることを示唆している。

本研究では、これらの先行研究を踏まえ、MRを活用した遠隔レッスンシステムを試作し、その有効性を検証する。特に、MicrosoftのHoloportation技術[6][7]を活用し、遠隔地にいる教師の演奏をホログラムとして生徒の視界内に投影するシステムを構築する。本システムの開発にはHoloLens技術の研究[10]を参考にし、リアルタイムかつ直感的なレッスン環境の実現を目指す。

2. 開発したシステムについて

本研究では、チェロの遠隔レッスンを支援するため、MicrosoftのMR技術を活用したシステムを開発した。本システムは、HoloLens 2およびAzure Kinect DKを用いて、遠隔地にいる教師の姿をホログラムとして生徒の視界内に投影し、双方向のリアルタイムコミュニケーションを可能にすることを目的とする。これにより、従来の映像によるオンラインレッスンでは困難であった、立体的かつ直感的な演奏指導を実現することが可能となる。本章では、本システムの概要、使用機材、構成、および操作方法について述べる。

2.1 システムの概要

本研究で開発したシステムは、HoloLens 2およびAzure Kinect DKを活用し、遠隔地にいる生徒と教師が同じ空間にいるかのようなレッスン環境を提供するものである。本システムの特徴として、従来のビデオ通話を用いた遠隔レッスンとは異なり、HoloLens 2を介して仮想空間上に教師の演奏姿勢をリアルタイムで表示し、生徒が異なる視点から

Prototype and Study of a Remote Practice Support System for Cello
^{†1} TAKERU OHTA, Hakodate Future University
^{†2} KAORU SUMI, Hakodate Future University



図 1 : HoloLens 2



図 2 : Azure Kinect DK

<https://www.microsoft.com/ja-jp/hololens> <https://azure.microsoft.com/ja-jp/products/kinect-dk>



図 3 : システムの構成図

指導を受けることができる点が挙げられる。

本システムは、日本マイクロソフトの千葉慎二氏より提供された Holoportation 技術をもとに構築した。Holoportation とは、Microsoft が 2016 年に発表した遠隔地にいる人物をホログラムとして投影する技術であり、医療や教育分野においても応用が進められている[5]。このシステムは、PC 上で動作するプログラムと HoloLens 上で動作するプログラムの 2 つの要素で構成されており、同一ネットワーク内での通信でリアルタイムに情報を転送する仕組みとなっている。

本研究では、チェロの演奏指導に特化した応用を行うため、Holoportation 技術を導入し、生徒の演奏姿勢を教師が遠隔から確認できる環境を整備した。本システムの利用により、教師は生徒の姿勢や弓の動きを立体的に確認しながら指導を行うことが可能となる。

なお、現段階では本システムに MR を活用した仮想教示情報の投影機能は未実装である。今後の拡張において、MR によるガイドラインの表示や、教師の演奏動作の録画・再生機能の追加を検討している。

以降、本システムにおける「生徒」とはチェロの演奏を学ぶ者、「教師」とは生徒に演奏技術を指導する者を指すものとする。

2.2 使用機材について

本システムで使用した機材は HoloLens 2, Azure Kinect DK, PC2 台, 三脚が 2 脚である。

● HoloLens 2

HoloLens 2 (図 1) は、Microsoft が 2019 年に発売した MR ヘッドセットである。本デバイスは、前頭部に搭載された半透明ディスプレイにより、外環境を見ながら仮想情報を視認することが可能である。また、複数のカメラを内蔵しており、ユーザのハンドジェスチャーや空間情報を認識す



図 4 : 実際の投影モデル



図 5 : 生徒の練習の様子

る機能を備えている。本研究では、Holoportation システムを介して、遠隔地の教師の姿をホログラムとして生徒の視界に表示するために使用する。

HoloLens 2 の特長として、スタンドアロンで動作可能であり、外部 PC を必要としない点が挙げられる。また、装着者の視点で録画を行うことが可能であり、レッスンの記録・振り返りに活用できる。

● Azure Kinect DK

Azure Kinect DK (図 2) は、Microsoft が提供する深度カメラであり、前方に赤外線を照射し、物体との距離を測定することで深度情報を取得することができる。また、RGB カメラを搭載しており、色情報を取得することも可能である。加えて、ボディトラッキングやハンドトラッキング機能を備えており、人物の動作をリアルタイムで認識することができる。

本研究では、Azure Kinect DK を用いて生徒の動作を取得し、その情報を HoloLens 2 へ送信することで、遠隔地にいる教師が生徒の演奏姿勢をリアルタイムで確認できるようにしている。

● PC およびネットワーク環境

本システムでは、Azure Kinect DK から取得した映像をリアルタイムで HoloLens 2 に転送するため、高性能な PC が必要となる。特に、映像処理の負荷が高いため、eGPU (外部 GPU) を搭載した PC を推奨する。また、本システムは同一ネットワーク内で動作するため、安定した Wi-Fi 環境が求められる。

● 三脚の使用

Azure Kinect DK はカメラデバイスであるため、適切な高さや角度で設置する必要がある。本研究では、Kinect を三脚に固定し、撮影範囲を調整することで、より適切な演奏姿勢のトラッキングを可能にした。Kinect によって捉えられた映像のみが転送されるため、三脚の高さや角度の調整は重

要な要素となる。

2.3 システムの構成

本システムの構成図を図3に示す。

システム構成の流れ

1. 生徒は HoloLens 2 を装着する。
2. 生徒の前方に Azure Kinect DK を配置し、高さや角度を調整する。
3. Azure Kinect DK は PC に接続され、取得した映像データをリアルタイムで処理する。
4. 生成されたデータはネットワークを介して教師側の HoloLens 2 へ転送される。
5. 教師は HoloLens 2 を介して生徒のホログラムを視認し、適切な指導を行う。

本システムを利用することで、遠隔地にいる教師と生徒がリアルタイムでコミュニケーションを取りながらレッスンをを行うことが可能となる。実際に生徒見ることが出来る投影モデルを図4に、生徒の練習の様子を図5に示す。

2.4 システムの使用法

本システムには、以下の2つの基本操作がある。

投影モデルの移動

- ・ 目の前に手を持ち上げ、親指と人差し指でつまむジェスチャーを行う。
- ・ その状態で指を移動させると、投影モデルの位置を変更できる。

投影モデルの回転

- ・ 親指と人差し指で投影モデルをつまみ、すぐに指を離すと 180° 回転する。
- ・ 左右反転の状態を補正し、正しい向きでモデルを視認できる。

これらの機能により遠隔学習における障壁の軽減を試みる。

3. 実験

本研究では、開発したシステムの有効性を評価するために、ユーザビリティ調査を目的とした実験を実施した。特に、チェロの遠隔レッスンにおける Mixed Reality (MR) 技術の有用性を検証し、従来の動画教材との比較を行うことを目的とする。本実験では、MR を活用したレッスンと、事前に撮影したチェロのレッスン動画の2種類を被験者に体験してもらい、それぞれの手法に対する評価をアンケートによって収集した。本章では、実験の概要、実施方法、評価手法について詳述する。

3.1 実験概要

本実験では、本システムのユーザビリティを調査するため、2種類のレッスン手法を比較した。1つ目は、本研究で開発した MR を活用した遠隔レッスンシステムを用いたレッスンである。2つ目は、一般的なオンライン教材と同様に事前に撮影したチェロのレッスン動画を用いたレッスンである。被験者には、MR レッスンと動画レッスンの両方を体験し



図6：MRによるレッスン(左)と動画のレッスン(右)

てもらい、各手法に関する評価をアンケートにより収集した。さらに、両手法を比較した際の印象についても調査を行い、それぞれの利点と課題を明らかにした。

3.2 実験詳細

本実験は、MR を用いたチェロの遠隔レッスンシステムの評価を目的として行った。被験者は19歳から23歳の大学生10名であり、実際にシステムを使用してレッスンを受けた後、アンケートに回答した。また、MR のレッスンと比較するため、事前に作成したレッスン動画を用いて同様のレッスンを行い、各手法の印象を比較した(図6)。

2種類のレッスンの順序による影響を排除するため、カウンターバランスを考慮し、ランダムな順序で受講してもらった。具体的な実験の流れは以下のとおりである。

レッスン順序の決定

MR レッスンと動画レッスンの順番をランダムに決定する。両レッスンの内容は同一とし、各レッスンの時間は20分とした。

MR レッスン

最初に HoloLens 2 の基本的な操作方法を説明する。システムを起動し、記録のためにカメラの録画を開始する。生徒側と教師側の PC および HoloLens 2 のソフトウェアを立ち上げ、相互の映像が投影モデルとして視認できることを確認する。

遠隔環境を再現するために、隣接する部屋に機材を配置し、音声通話には LINE を使用した。

チェロのレッスンを実施し、終了後に MR レッスンに関するアンケートに回答してもらう。

動画レッスン

被験者に、事前に筆者が作成したレッスン動画を視聴しながらチェロの練習を行ってもらう。

動画は一時停止、巻き戻し、スキップが可能とした。

レッスン終了後、動画レッスンに関するアンケートに回答してもらう。

比較評価

両レッスン終了後、MR と動画の比較評価をアンケート形式で実施する。

それぞれのアンケートは以下のとおりである。(表1, 表2)

表1：動画またはMRレッスンの印象について

Q1: このテクノロジーは、先進的だと感じましたか?
Q2: このテクノロジーは、洗練されていると感じましたか?

Q3: このテクノロジーの操作やデザインは直感的だと感じましたか？
Q4: このテクノロジーの操作は簡単だと感じましたか？
Q5: このテクノロジーは、親しみやすいと感じましたか？
Q6: このテクノロジーは、気軽に使えると感じましたか？
Q7: このテクノロジーは、効率的だと感じましたか？
Q8: このテクノロジーのデザインや構造は明瞭だと感じましたか？
Q9: このテクノロジーの応答性は高いと感じましたか？
Q10: このテクノロジーとのやり取りは自然だと感じましたか？
Q11: このテクノロジーは、双方向的なやり取りができると感じましたか？

表 2：動画と MR を比較した印象について（一部抜粋）

q1: MR システムにおいて、投影モデルを移動させたいと感じましたか？
q2: それはなぜですか？またどのような場面で移動させたいと感じましたか？
q3: MR システムにおいて、実際に投影モデルを移動させましたか？
q4: どのような場面で移動させましたか？
q5: MR システムにおいて、投影モデルの角度を変えたいと感じましたか？
q6: それはなぜですか？またどのような場面で角度を変えたいと思いましたか？
q7: MR システムの投影モデルは見やすかったですか？ 1 ～5 で選んでください
q8: なぜそう感じましたか？
q10: 弓の持ち方を教わるのはどちらの方がわかりやすかったですか？
q12: 弦の抑え方を教わるのはどちらの方がわかりやすかったですか？
q15: それでは、総じて現時点では MR と動画のどちらの方がわかりやすかったですか？

3.3 被験者と条件群

本実験には、19 歳から 23 歳の大学生 10 名が参加した。性別の内訳は、男性 3 名、女性 7 名である。また、10 名のうち 7 名が何らかの楽器演奏経験を有していた。被験者の詳細は表 3 に示す。本実験では、順番による影響を抑えるため、MR と動画レッスンの順番をランダムに割り当てた。

表 3：被験者のデータ実験手順

番号	性別	年齢	楽器の経験年数（複数の楽器がある場合はそれぞれの楽器と年数を教えてください）	先に行ったレッスン
1	女性	19	ピアノ 6 年 ヴァイオリン 6 年	MR
2	女性	22	トロンボーン 2 年 ヴィオラ 1 年 和太鼓 2 年	MR
3	男性	22	未経験	動画
4	女性	19	チェロ 7 か月	MR
5	男性	22	無いです。	動画
6	女性	19	箏 9 年、ユーフォニアム 3 年、ベース 3 年、チェロ 6 ヶ月	動画
7	女性	20	ピアノ 5 年、トランペット 2 年半、箏 2 年半、チェロ 6 ヶ月	MR
8	女性	19	ピアノ 2 年くらい習っていた ベース 4 年くらい	動画
9	男性	23	和太鼓 3 年	MR
10	女性	22	幼少期にピアノ 6 年	動画

本実験は、MR システムのユーザビリティと学習効果を評価することを目的とし、被験者に対して MR を用いた遠隔レッスンと動画を用いたレッスンの両方を体験してもらい、それぞれの方法の違いを評価する形で実施した。実験は以

下の手順で進められた。

まず、被験者には実験環境として準備された PC の前に座ってもらい、実験の概要と進行手順を説明した。その後、実験に関する同意書に署名をもらい、正式に参加の承諾を得た。次に、被験者が体験するレッスンの順番を事前に決定し、順序が実験結果に影響を与えないようにランダムに設定した。

本実験では、まず動画を用いたレッスンを実施した。被験者には、事前に録画されたチェロのレッスン動画を提示し、この動画を見ながらチェロの練習を行うように指示した。動画の再生時間は 11 分 24 秒であったが、レッスンの時間は 20 分とし、被験者が必要に応じて特定の場面を繰り返し視聴したり、不要な部分をスキップしたりすることを許可した。動画レッスンの終了後、被験者には PC 上で SD 法によるアンケートに回答してもらい、動画レッスンに対する印象を評価した。

続いて、MR を用いたレッスンを実施した。まず、被験者に MR システムの基本的な使用方法を説明した。具体的には、HoloLens 2 を装着すること、Kinect の前に座ること、教師が遠隔地にいるため画面上に投影モデルとして表示されることなどを伝えた。次に、PC のソフトウェアおよび HoloLens 内のソフトウェアを起動し、生徒と教師の姿が互いに投影モデルとして表示されることを確認したうえで、レッスンを開始した。MR レッスンでは、以下の内容を順番に指導した。

- ・チェロの構え方
- ・弓の持ち方
- ・弓でチェロを弾く方法
- ・指の押さえ方
- ・ピチカートで音階を練習
- ・弓を使って音階を練習
- ・ピチカートで「きらきら星」を練習
- ・弓で「きらきら星」を練習

このレッスン内容は、動画レッスンと同じカリキュラムに沿って進められたが、MR システムでは被験者が直接質問できるため、理解が深まるまで特定の内容を繰り返し練習するなど、被験者ごとに進行速度に差が生じた。レッスン終了後、動画レッスンと同様に、被験者には SD 法によるアンケートに回答してもらい、MR レッスンの使用感や学習効果を評価した。

最後に、被験者には動画レッスンと MR レッスンを比較し、それぞれの長所・短所について自由記述形式のアンケートに回答してもらい、実験を終了した。

3.4 評価方法

本実験では、SD 法 (Semantic Differential Method) によるアンケート調査と自由記述形式のアンケート調査を行った。

SD 法 は、システムの印象やユーザビリティを具体的に評価するための手法であり、特に MR のような新技術の受容

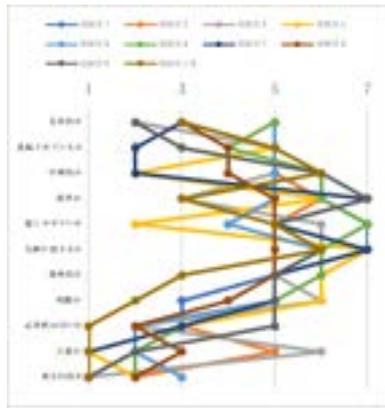


図 7：動画レッスンの SD 法評価結果

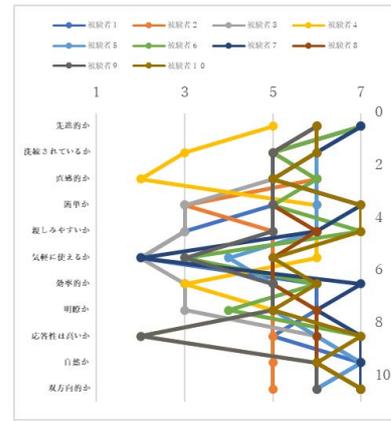


図 8：MRレッスンの SD 法評価結果

度を測定するのに適していると考えられる。

自由記述アンケートでは、MR システムの操作性や視認性、学習効果に関する詳細な意見を収集し、システムの改善点を明らかにすることを目的とした。

4. 結果

本章では、実験によって得られた評価データを分析し、MR システムと動画レッスンの比較を行う。まず、SD 法によるアンケートの結果を示し、次に自由記述アンケートの分析結果について述べる。

4.1 SD 法のアンケート結果

本実験では、動画レッスンおよび MR システムを用いたレッスンのそれぞれについて、SD 法 (Semantic Differential Method) による評価アンケートを実施した。各被験者には、「先進的であるか」「洗練されているか」「直感的であるか」「操作が簡単であるか」「親しみやすいか」「気軽に使用できるか」「効率的であるか」「明瞭であるか」「応答性が高いか」「自然なやり取りが可能であるか」「双方向的なやり取りが可能であるか」という 11 項目について、7 段階で評価

を行ってもらった (図 7, 図 8)。

SD 法の結果を集計し、それぞれの結果を図 6 と図 7 に示す。また、動画レッスンと MR システムのそれぞれの平均スコアを算出し、動画と MR の SD 法評価を重ねたグラフを作成した (図 9)。グラフにおいて、黄色は動画レッスンの平均点、青は MRレッスンの平均点を示している。

本結果から、MR システムは「先進的である」「双方向的なやり取りが可能である」「応答性が高い」「自然なやり取りが可能である」といった項目で動画レッスンよりも高い評価を得ていることが分かった。一方で、「気軽に使用できる」「操作が簡単である」といった項目では動画レッスンの方が評価が高い傾向にあった。これにより、MR システムは直感的なインタラクションと双方向性に優れている一方で、操作の習熟に時間を要する可能性があることが示唆された。

4.2 自由記述のアンケート結果

ここでは自由記述で被験者に MR のシステムについて質問して得られた回答のうち特筆すべき回答を掲載する。

Q1: MR システムにおいて、投影モデルを移動させたいと感じましたか？

この質問に対しては全ての被験者がはいと答えた。(100%)

Q2: それはなぜですか？またどのような場面で移動させたいと感じましたか？ (一部抜粋)

- ・演奏者の全体を見たい場面と部分的に見たい場面で詳しく見たいと思ったから。
- ・指の位置や弓の角度をもっと正確に見たいときに移動させようとおもった

Q3: MR システムにおいて、実際に投影モデルを移動させましたか？

この質問に対しては 6 名の被験者がはいと答えた。(60%)

Q4: (はいと答えた場合)どのような場面で移動させましたか？また、その場面で移動させた理由などあれば教えてください。(一部抜粋)

- ・指の位置や弓の角度をもっと正確に見たいと思ったときに移動させた。
- ・操作ごとにフォーカスしたい位置が違ったから 弦押さえたり弓の持ち方だったり

(いいえと答えた場合) 実際には移動させなかった理由など

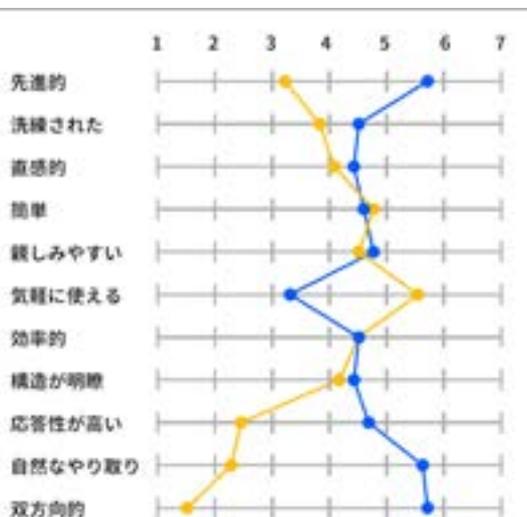


図 9：動画レッスンと MRレッスンの SD 法評価比較

あれば教えてください。(一部抜粋)

・操作が難しく移動させられなかった。
・チェロと弓を持って移動させるのは、両手がぼぼふさがっているのが難しかった。

Q5: MR システムにおいて、投影モデルの角度を変えたいと感じましたか？

この質問に対しては4名がはいと答えた。(40%)

Q6: (はいと答えた場合)それはなぜですか？またどのような場面でモデルの角度を変えたいと思いましたか？(一部抜粋)

・弦に対して垂直に弓を持つところで、目印になるようなところはないか知りたかった。
・向かい合う映像だけでなく同じ向きの映像も見たくなかったから

(いいえと答えた場合) 移動させたいと感じなかった理由などあれば教えてください。(一部抜粋)

・特に角度を変える必要を感じなかったから
・前面のみで全体像が見えたから。

Q7: MR システムの投影モデルは見やすかったですか？

1 (とても見づらかった) ~ 5 (とても見やすかった) の5段階評価で回答については以下のグラフに結果を示す。横軸は被験者番号、縦軸は各被験者の評価である。

(平均=3.0、標準偏差=0.63)

Q8: なぜそう感じましたか？(一部抜粋)

ポイント	コメント
2	画質があまり良くなかったから。
3	モデルの輪郭が歪んでいる点があったが、操作で画面を近づければ気にならなくなりました。
4	色もついていてわかりやすかったから

Q9: 投影モデルはどのように見えましたか？気になった点などあれば教えて下さい。(一部抜粋)

・明るく思っていたよりもはっきりと見えた。上下に制限があるのが少し気になった。
・古いゲームみたいな解像度と色合いだった。面白いとは思ったもうちょっと鮮明だともっとわかりやすいかも

Q10: 弓の持ち方を教わるのはどちらの方がわかりやすかったですか？

この質問に対しては7名がMRと答えた。(70%)

Q11: (MRと答えた場合)それはなぜですか？またどのような点でMRがよいと思いましたか？(一部抜粋)

・あっているかどうかをその場で確かめられるから
・リアルタイムで指示してもらえること、見たいものを見たい近さや角度で見て学ぶことができたから

(動画と答えた場合) それはなぜですか？またどのような点で動画がよいと思いましたか？(一部抜粋)

・動画の方が鮮明で、直感的に学びやすかった
・どちらもそんなに変わらなかったが、動画のほうが画質が良かったから。

Q12: 弦の抑え方を教わるのはどちらの方がわかりやすかったですか？

この質問に対しては8名がMRと答えた。(80%)

Q13: (MRと答えた場合)それはなぜですか？またどのような点でMRがよいと思いましたか？(一部抜粋)

・角度が立体的に確認できたため。立体的なところで教える部分が出たときにMRはとてもしよかった。
・押さえる位置の印がなくて感覚でやっていたのでコツとかが聞けて分かりやすかった点

(動画と答えた場合) それはなぜですか？またどのような点で動画がよいと思いましたか？(一部抜粋)

・MRは見えなくなってしまうときがあったが、動画は画面が常に固定されていたので見やすかった。
・動画のほうが見やすかったから

Q14: それでは、総じて現時点ではMRと動画のどちらの方がわかりやすかったですか？

この質問に対しては8名がMRと答えた。(80%)

Q15: (MRと答えた場合)それはなぜですか？その判断のもとになった場面やシステムなどあれば教えてください。

また、MRが使いにくいと感じる部分も教えてください。

(一部抜粋)

・やはり立体的に教わらないといけない部分はMRの方が良かったと感じたため。少し見づらい部分もあったが、それでも立体的に見えるのはとてもありがたかった。細かい部分をもっと見えたらかなりいいと思う。解像度的な問題な気がした。
・動画の方が直感的に分かりやすい部分はあったものの、再生を停止する必要があったり巻き戻す必要があることがストレスに感じた。しかし、MRでは何度も同じ点を教えてもらえるし、気になったことを尋ねることができてストレスにならなかった

(動画と答えた場合) それはなぜですか？MRがわかりにくいと感じた点を教えてください。どのような場面でわかりにくかったですか？(一部抜粋)

・レッスンという面ではMRのほうが一緒にやってる感じもあり、その場で疑問を解決できてよかった。しかし、動画のほうが気兼ねなく巻き戻せて、好きなところを復習できたため練習しやすかった。
・特に分かりにくかったことはないが、移動させることがほとんどなかったため、動画と変わらないと感じた。

5. 考察

本章では、第4章で示したデータを分析し、システムの評価および課題について考察する。まず、SD法による評価アンケートの結果をもとに、MRシステムの特性と従来の動画レッスンとの比較を行う。次に、自由記述アンケートの分析結果を示し、被験者の主観的な評価を考察する。最後に、システムの問題点を整理し、今後の改善策について述べる。

5.1 SD法のアンケート結果より

第4章で示したSD法のアンケート結果を分析すると、MRシステムは「先進的である」「応答性が高い」「自然なやり取りができる」「双方向的である」といった項目において、動画レッスンよりも高い評価を得ていた。これにより、遠隔環境においてもMRを活用することで、リアルタイムかつ効果的なコミュニケーションが可能であることが示唆される。また、「先進的」と評価されていることから、今後の教育分野におけるMR技術の発展に寄与する可能性があると考えられる。

一方で、「気軽に使用できる」という項目では動画レッスンの方が高い評価を得ており、MRシステムの利便性には課題が残ることが示された。この要因として、機材のセットアップに要する時間や、HoloLens 2の装着による物理的負担が考えられる。また、「直感的な操作」という項目においても、動画との顕著な差が見られなかったことから、システムのユーザインタフェースの改善が求められる。

さらに、「効率的である」という項目では動画とMRの評価がほぼ同等であり、現状のMRシステムが動画レッスンに対して特別に学習効率を向上させる要素を持っていない可能性が示唆された。今後は、MRならではの強みを生かし、学習の効率性を向上させる機能の追加が必要である。

5.2 自由記述のアンケートより

本研究では、被験者からMRシステムに対する主観的な評価を収集するために自由記述形式のアンケートを実施した。この結果を分析したところ、大きく「投影モデルの操作性」「視認性」「学習効果」の3つの観点から意見が得られた。

● 投影モデルの操作性に関する評価

MRシステムにおいて、投影モデルを移動・回転できる機能を導入したが、これに関する評価は分かれた。全被験者が「投影モデルを移動させたい」と回答したものの、実際に移動させたのは6名(60%)にとどまった。その理由としては、「移動操作が難しい」「チェロと弓を持っているため両手がふさがる」といった意見が多く挙げられた。

移動操作に関しては、HoloLens 2のジェスチャー操作に不慣れなことが影響したと考えられる。例えば、被験者の一人は「親指と人差し指でつまむジェスチャーを試みたが、うまく認識されずにモデルが意図しない方向へ動いた」と述べている。また、別の被験者は「チェロの演奏中に操作を行うのは難しく、結局モデルを移動させることはなかった」と指摘した。このように、実際の演奏中における操作性の向上が今後の課題となる。

投影モデルの回転についても評価は分かれた。「角度を変えたい」と回答した被験者は4名(40%)にとどまり、半数以上は「特に必要性を感じなかった」と回答している。その一方で、「演奏者と同じ向きでモデルを見られると理解しやすい」という意見もあり、特定の場面では有効である可能性が示唆された。「指の位置や弓の角度を確認する際に、投影モデルを回転させて詳細を見たかったが、現状の操作では回転がしづらい」と述べた被験者もいた。このことから、ユーザが直感的に視点を調整できるUI設計の改善が求められる。

● 視認性に関する評価

投影モデルの視認性については、5段階評価で平均3.0となり、中程度の評価にとどまった。具体的な問題点として、「モデルの輪郭が歪む」「弓を持つ指が実際よりも大きく表示される」「モデルが不鮮明で細かい動作が見えづらい」などの意見が寄せられた。特に、被験者の一人は「指の押さえ方を確認しようとしたが、投影モデルがぼやけていて詳細が分かりにくかった」と指摘しており、解像度の向上が求められる。

さらに、投影モデルの一部が視界から消えてしまう現象についても意見が寄せられた。本システムではKinect 1台で撮影しているため、被写体の後方部分が映らないという問

題があった。ある被験者は「腕を前に動かした際、体の一部が影のように見えなくなってしまった」と述べており、視覚的な一貫性を保つためには、カメラの設置方法や撮影角度の調整が必要である。

● 学習効果に関する評価

本システムの学習効果に関しては、8名(80%)の被験者が「動画よりもMRの方が優れている」と回答した。特に「弓の持ち方」「弦の押さえ方」の学習においてMRが有効であるとする意見が多かった。ある被験者は「弓の持ち方を正しく学ぶためには、実際にどのように手が動くのかをリアルタイムで確認できるMRの方が効果的だった」と述べている。別の被験者は「指の押さえる位置を立体的に確認できるため、どの角度で力を入れるべきかが分かりやすかった」と評価した。

一方で、動画の方が優れていると回答した被験者(2名)は、「動画は繰り返し視聴できるため、自分のペースで学習しやすい」「MRはリアルタイム性があるが、細かい部分をじっくり確認するには向いていない」といった意見を述べた。このことから、MRシステムには動画のような「再生・巻き戻し機能」を補完する仕組みを導入することで、より効果的な学習環境を構築できると考えられる。

また、「リアルタイム性と双方向性がMRの最大の強み」とする意見が多く見られた。例えば、「動画では一方的な情報提供になるが、MRではその場で質問できるため、より深い理解につながる」と述べた被験者がいた。この点は、MRシステムが動画レッスンに対して明確な利点を持つことを示している。

● 総合的な評価と改善点

自由記述アンケートの結果から、本システムは「リアルタイム性」「立体的な視認性」「双方向的な学習環境」といった点で高く評価されたものの、「操作の難しさ」「投影モデルの視認性」「反復学習のしやすさ」などに課題があることが明らかとなった。

投影モデルの操作性に関しては、特に「ジェスチャーの認識精度」や「両手がふさがる状況での操作」の改善が求められる。また、視認性の向上のためには「解像度の向上」「カメラの台数増加」「映像処理の最適化」が必要である。さらに、動画の長所を補完するため、「学習履歴の記録」「再生・巻き戻し機能の追加」などの改善策を検討することで、MRシステムの実用性をより高めることができると考えられる。

本研究で開発したMRシステムは、従来の動画レッスンに比べて、演奏技術の習得に有効であることが示された。しかし、実際の使用環境においては、操作性や視認性の課題が明らかとなった。これらの点を改良することで、より直感的かつ効果的な遠隔レッスンシステムの構築が可能になると考えられる。

5.3 システムの問題点と改善策

本研究の結果を踏まえ、本システムの課題とその改善策を以下の3点に整理する。

(1) 機材の普及性と操作性

MR システムの最大の課題の一つは、機材の普及率が低く、入手が困難である点である。HoloLens 2 および Azure Kinect DK はすでに生産終了となっており一般の学習者が手軽に利用することが難しい。この問題に対しては、より汎用的な機材を活用する方法が考えられる。例えば、スマートフォンやタブレットを用いた AR 技術を応用することで、より多くの学習者がアクセス可能なシステムを構築できる可能性がある。

(2) 操作の直感性の向上

現在のシステムでは、投影モデルの移動や回転の操作が難しく、被験者の多くが操作に苦勞していた。この問題を解決するためには、より直感的な操作方法を導入する必要がある。例えば、Unity の MRTK3 を活用した Object Manipulator や Bounding Box を実装することで、より自然なジェスチャー操作を可能にすることができる。さらに、システムの操作を練習できる試用モードを導入し、利用者がレッスン前に操作方法を習得できる環境を整備することも有効である。

(3) 映像品質の向上

被験者の意見から、投影モデルの画質が低く、一部の細かい動作が確認しづらいという課題が挙げられた。この問題を解決するためには、より高解像度のカメラを用いる、あるいは Kinect を複数台設置して異なる角度からの映像を統合する手法が考えられる。しかし、これらの手法はコストや機材の設置難易度を高めるため、今後の研究では映像の補正技術についても検討する必要がある。

6. まとめ

本研究では、チェロの学習における障壁を軽減し、遠隔指導の効果を高めるために、MR (Mixed Reality) 技術を活用したレッスン支援システムを開発した。本システムは、HoloLens 2 と Azure Kinect DK を使い、教師と生徒の姿をホログラムとして投影することで、リアルタイムでの指導を可能にした。

本システムの評価として、SD 法によるアンケート調査と自由記述アンケートを実施した。結果として、「双方向性」「応答性」「視認性」といった点で、MR レッスンが従来の動画レッスンを上回る評価を得た。一方で、「操作の直感性」「手軽さ」といった項目では動画の方が優れていると評価された。また、自由記述の分析では、リアルタイムでの指導の有効性が指摘された一方で、操作の難しさや HoloLens の重量、解像度の低さなどが課題として挙げられた。

本研究の成果は、従来の遠隔学習システムが抱えていた課題を克服する新たな手法を提示するものである。しかし、操作性の向上やデバイスの軽量化、映像の品質向上など、

さらなる改良が求められる。今後の研究においては、より直感的なインターフェースの開発や、ハードウェアの進化に対応したシステムの改善を進めることで、より実用的な遠隔チェロレッスンの実現を目指す。

謝辞

本研究の遂行にあたり、多大なご支援をいただいた日本マイクロソフトの千葉慎二氏に深く感謝申し上げます。

参考文献

- [1] 一般財団法人ヤマハ音楽振興会, “楽器演奏に関する調査. 働く男女, 3人に1人以上は今後楽器を始めたい! 一方, 4割は挫折経験あり. 始めたい楽器1位「ピアノ」一緒に楽器を演奏したい俳優ランキングトップ10,” PRTimes, 2022年3月30日. [Online]. Available: <https://prt-times.jp/main/html/rd/p/000000004.000024395.html> (参照 2025-01-23).
- [2] Microsoft, “Mixed Reality とは,” Microsoft, 2023年3月21日. [Online]. Available: <https://learn.microsoft.com/ja-jp/windows/mixed-reality/discover/mixed-reality> (参照 2025-01-23).
- [3] Y. Murase, K. Takashima, and K. Nishimoto, “GP-Mixer: A System That Enables Individual Adjustment of Musical Instrument Practice Environments Within Groups,” *Interaction*, 2018.
- [4] Y. Yamaguchi, et al., “Development and Evaluation of a Musical Performance Support System Using Smart Glasses: A Study on Posture, Gaze, and Performance Evaluation,” *Journal of the Japanese Society for Educational Technology*, vol. 46, Suppl., pp. 185-188, 2023.
- [5] Microsoft, “Holoportation,” Microsoft Research, [Online]. Available: <https://www.microsoft.com/en-us/research/project/holoportation-3/> (参照 2025-01-23).
- [6] S. Chiba, “Microsoft HoloLens の技術とその活用,” CIR NII, [Online]. Available: <https://cir.nii.ac.jp/crid/1390282763122992768> (参照 2025-01-23).
- [7] S. Chiba, “The New World Which Mixed Reality Brings,” *Journal of the Virtual Reality Society of Japan*, 2021.
- [8] J. M. M. Montoya and A. G. Salazar, “Remote Learning and Music Education: A Systematic Review,” *International Journal of Music Education*, vol. 39, no. 3, pp. 201-216, 2021.
- [9] C. Johnson and P. Thomas, “Augmented Reality Applications in String Instrument Learning,” *Proceedings of the International Conference on Music Education Technology*, 2022, pp. 98-105.
- [10] S. Mehta, “Microsoft’s Azure Kinect Developer Kit Technology Transfers to Partner Ecosystem,” Microsoft Tech Community Blog, [Online]. Available: <https://techcommunity.microsoft.com/blog/themixedrealityblog/microsoft-%e2%80%99s-azure-kinect-developer-kit-technology-transfers-to-partner-ecosystem/3899122> (参照 2025-01-23).

仮想現実での視線情報を用いた 商品作動ギミックによる関心の変化

松田 滉生^{†1}青山学院大学^{†1}ロペズギヨーム^{†2}青山学院大学^{†2}

1. はじめに

近年、日本では EC サイトを利用した商品の購入が急激に増加しており、今や多くの消費者がインターネットを活用した買い物を日常生活の一部として受け入れている。統計局が実施した調査によると、ネットショッピングの利用率や支出額は年々増加傾向にあり、その成長は新型コロナウイルス感染症の拡大による生活様式の変化とも相まって一層加速している[1]。特に外出を控える生活が推奨された2020年以降、食品や生活必需品の購入においてもネットショッピングを利用する人々が急増し、EC市場の拡大に拍車をかけたとされている。さらに、店舗で実物を確認せずにインターネットだけで商品を購入することに抵抗がない消費者も増加している。2021年の調査では、その割合が約50%に達し、従来の「商品を直接手に取って確認したい」という購買行動の変化を示している[2]。の背景には、商品の写真や動画、レビューなどのデジタルコンテンツの進化が挙げられる。これにより、消費者は実物を確認しなくても商品の品質や性能をある程度把握できるようになり、購買の意思決定がオンラインで完結するケースが増えている。

一方で、ネットショッピングにはまだ課題も残されている。総務省が実施した調査によると、ネットショッピングを利用しない理由として「実物を見てから購入したい」という回答が多く挙げられており、これがEC市場のさらなる成長を妨げる一因となっている[3]。特に高額商品やサイズが重要な商品(家具、衣類、電化製品)においては、購入前に実物を確認したいというニーズが根強く存在している。また、商品の写真や説明が実物と異なる場合、消費者の期待を裏切る結果となり、返品や交換の増加につながる可能性も指摘されている。以上のように、ネットショッピングの利用率や支出額は確実に増加しているが、同時にその利便性と実物確認の重要性という、相反する要素が存在していることが分かる。これらの課題を解決するためには、消費者が安心して購入できる環境を整備することが必要だ。

例えば、商品の詳細を分かりやすく伝えるための高品質な写真や動画、拡張現実(AR)や仮想現実(VR)を活用した視覚的な確認手段の導入が期待されている。

以上より、ネットショッピングは利便性の向上により多

くの支持を得ている一方で、実物を確認できないことによる購買意欲の低下という課題を抱えている。この問題を克服することで、さらに多くの消費者がネットショッピングを活用し、EC市場の発展が進むと考えられる。

2. 関連研究

本章では、VRや消費行動の変化に関する研究について述べる。ネットショップにおける商品の提示形式が消費者の選択や購買意欲に与える影響を検討した[4]。特に「暗示的動き(Implied Motion)」を表す静止画に注目し、其の効果とメカニズムを議論している。視覚刺激が消費者の注意や判断に及ぼす影響は広く研究されてきたが、従来は静的な商品画像やビデオなどの動的提示が中心だった。本研究は、静止画が持つ暗示的な動きが消費者行動に与える影響を実証的に検証した点で新規性がある。「暗示的動き」とは、静止画でありながら動きが感じられる視覚的特性を指し、視覚システムによる運動の想像と再現によって生じる。これにより、消費者は商品使用の場面や効果を鮮明に想像しやすくなり、購買意欲や商品態度に影響を受ける。仮説を検証するため、Yahoo!クラウドソーシングで実験を実施。200名の20代~60代の一般消費者に、動的な画像と静的な画像(図2-1)を提示し、質問に回答してもらった。結果として、暗示的動きを表す静止画は消費者に運動を知覚させ、心的イメージの鮮明さを高め、購買意欲や商品態度を向上させることが確認された。さらに、静止画はビデオやVRに比べて制作コストが低く、ネットショップにおける実用性が高いと考えられる。

Verification of a Feedback Method for Promoting Blinking Eyes in Video Viewing

†1 KOUKI MATSUDA, Aoyama Gakuin University

†2 GUILLAUME LOPEZ, Aoyama Gakuin University



図 2-1：暗示的動きの画像と静的な商品写真

落合は視線情報に基づいた嗜好分析から商品を推薦する VR システムを提案している[5]。コンビニエンスストアを想定したワールドを作成し、カップ麺を 33 種類用意しユーザーの興味や嗜好分析を視線情報とアンケート評価で行った結果、注視時間からユーザーの興味や嗜好を判断することが可能だと考えられた。

中島らは、百貨店におけるアイトラッキングデバイスを用いた購買行動の評価を行った[6]。非計画購買は、店舗滞在時間や顧客特性に影響されることが示唆されている。本研究では、多層階の百貨店を対象に、Tobii 社のアイトラッキングデバイスを装着した被験者に 1 時間の自由なショッピングを行わせ、視線データや滞在時間を記録した。さらに、購買目的や訪店歴のアンケートデータを収集し、非計画購買の発生要因を分析した。LOOCV によるモデル評価とランダムフォレストを用いた分析の結果、実店舗嗜好や滞在時間が非計画購買に影響を与えることが明らかとなった。特に、滞在時間が 1500 秒を超えると非計画購買が増加し、2000 秒を超えると購買意欲が減少する傾向が確認された。この結果から、実店舗への関心が高い消費者ほど非計画購買を行いやすいことが示唆され、アイトラッキングデバイスを活用した視線情報の収集が購買行動の予測・分析に有効であることが実証された。

3. システムの概要

本節では本研究で提案する仮想現実において、視線情報を用いた商品作動ギミックの概要について述べる。動的な動きが分かりやすいように本研究では家電製品を 3D モデルとして使用する。ヘッドマウントディスプレイを用いた仮想現実上に家電量販店をイメージしたワールドを作成し、Pupil Neon[7]を Meta Quest3[8]に取り付け視線情報を読み取り、注目が集まっているオブジェクトのみをアニメーションで動くようにシステムを作成。家電量販店に直接行くのではなく仮想現実上で動的な動きを見ることによる興味の向上や変化を図ることを目的としたシステムである。

図 3-1 に提案システムの全体図を示す。被験者はヘッドマウントディスプレイを装着し、VR 空間上で家電製品を見る

ことができる。視線情報によってポインターが移動し、図 3-2 のように視覚的にどこに注目しているのかわかりやすくするために、ポインターから赤色で ray を表示させた。



図 3-1：家電量販店をイメージしたワールドの様子



図 3-2：ray が飛んでいる様子

またアニメーションの設定として、炊飯器は開閉の動作、扇風機は首を振る動作、電子レンジは扉の開閉が行えるようにアニメーションを作成してあり、動的な動きを直感的に理解できるようにした。

3.1 動的な 3D モデルの作成方法

3D モデルには、Unity 上の Asset を使うのではなく、現実世界にある電化製品を使うことでより現実世界をイメージできるようにした。現実世界の電化製品を 3D モデルに変更するために Toolbox AI 社が出しているスマートフォン向けアプリケーションの Scaniverse-3DScanner[9]を使用した。使用したスマートフォンは Apple 社が出している iPhone 14 を使用した。図 3-3 のように現実世界の家電を机の上などに置き、アプリケーションを起動して全方位を読み取ることで 3D モデルを作成した。



図 3-4：3D モデルの読み取り方

3.2 システム開発に利用した開発プラットフォーム

オーム

本開発には Unity Technologies が提供する Unity[10]を用いた。Unity は無料で利用可能なゲームエンジンであり、作成したプロジェクトは Android や PC, HMD など様々な環境にビルド可能である。また、様々なユーザーが作成したアセットがアセットストアから利用可能となっており、短期間での開発が可能である。本システムは HMD 上で動作するため、HMD 上のあらゆるセンサ情報を扱えるようにする拡張パッケージ (OpenXR) を導入した。本システムは 1 人で Meta Quest を装着し椅子または立った状態で仮想現実内の電化製品を見る。

3.3 視線情報計測

本研究では、視線情報の取得のために Pupil Labs 社が出す Pupil Neon を使用した。Pupil Neon を利用することで VR 空間に存在する 3D オブジェクトに対してユーザーの視線情報が計測できる。ユーザーの左右の眼下にそれぞれ小型カメラを設置して、ユーザーの眼球運動と VR 空間にある注視点をキャリブレーションすることで、VR 空間においてユーザーの視線情報を取得することができる。

ユーザーの視線方向に ray を飛ばすことで、ray が電化製品に衝突した時に、商品を見たかと判定することができる。この視線情報を利用して、ユーザーがどの商品を見たのか、商品は何秒間見ていたのかなどのデータを取得することができる。1 秒以上の視線の停止を注目として VR 上でも判定できると分かっているため、これを使い電化製品に注目 (1 秒以上) が集められた時アニメーションが動きだすように設計した。Pupil Neon とスマートフォンを接続し、Pupil Labs 社の Neon Companion を使い、視線がどこを向いているのかの測定と接続を行い、python で動かした。

視線情報の計測を実装するため、図 3-5 のように Meta Quest3 と Pupil Neon を接続し仮想空間上 (Unity) で視線情報を読み取り Pupil Neon に情報を送り、データ化された視線情報から注目時間によってアニメーションを再生した。コンソールから api を Pupil Neon に送り、視線情報をコンソールに返して MQTT ブローカーにパブリッシュし、VR に送る。

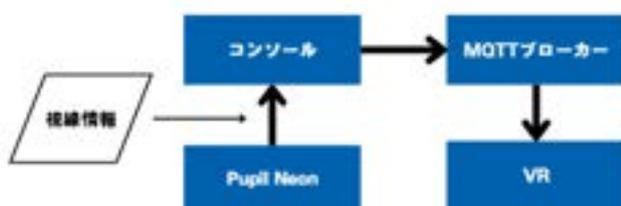


図 3-5：視線情報計測と共有の仕組みの概要図

4. モデルの動的な動きの有無による評価実験

4.1 実験目的

本研究の目的は、本システムを利用する際のシステムの

有用性や、ユーザーがどのように感じるかを検証することである。具体的には、ユーザーの興味や嗜好、体験を通じた感じ方がどのように変化するかを分析し、それらの変化に基づいてシステムの有効性を評価する。また、以下の 3 つのパターンに基づいて比較を行い、各パターンがユーザー体験に与える影響を検証する。静止、モデルが動かない場合常時動作、モデルが常に動いている場合注視動作、視線情報によってモデルが動く場合。これらの比較を通じて、視線情報を活用した動的なインタラクションがユーザーの興味喚起や体験の質に与える効果を明らかにし、システム設計における有用性の評価を行うことを目指す。

4.2 実験概要

被験者 10 人に図 4-1 の Meta Quest3 に Pupil Labs を取り付けたヘッドマウントディスプレイを装着してもらう。



図 4-1：ヘッドマウントディスプレイ

4.3 実験方法

本研究では、3 つの異なる VR 環境を用いて、被験者に電化製品 (炊飯器、電子レンジ、扇風機) を観察・比較してもらい、それぞれの場合でエンゲージメントに与える影響を評価した。以下に実験の手順を示す。各 VR 環境は以下の 3 つのパターンで構成される。

- ・ 静止: 3D モデルが動かない静止状態。
- ・ 常時動作: 3D モデルが常に動き続ける状態。
- ・ 注視動作: 被験者の視線情報に基づき、注目している 3D モデルのみが動く状態。

この時各実験参加者には 3 つのパターンの VR 環境を体験してもらうが、順序効果を排除するために、体験順序をランダムに設定した。被験者には事前に、3 パターンの環境で電化製品の観察・比較を行った後にアンケートに回答してもらうことを伝えた。また、それぞれのパターンの具体的な違いには明示せず、観察を自然に行ってもらうように促した。各被験者には 3 つの VR 環境を順番に体験してもらい、それぞれの環境で電化製品を観察した。観察時間は被験者に委ね、十分に観察したと感じた時点で次の環境に移行す

るようにした。

4.4 評価方法

4.4.1 エンゲージメント評価

本実験では、各 VR 環境でのユーザーの利用時間をエンゲージメントの定量的な指標として定義した。具体的には、各環境の体験開始から終了までの時間 (秒) を計測し、その値を比較することでエンゲージメントの違いを評価した。被験者が各 VR 環境で観察を開始した時点でタイマーを開始し、観察を終了した時点でタイマーを停止することで、利用時間を記録した。各環境での利用時間は被験者が自由に決定できるようにし、無理のない形で自然な観察行動が得られるよう配慮した。記録した利用時間を環境ごとに集計し、平均利用時間や標準偏差を算出した。また、各パターンにおけるエンゲージメントの違いを統計的手法を用いて比較した。定量的な評価に加えて、被験者にはアンケートを通じて主観的な感想や各環境に対する満足度を回答してもらった。これにより、定量的評価と主観的な意見の相関を分析し、視線連動型インタラクション (注視動作) の有効性を検証した。

4.4.2 アンケート評価

本研究では、3 つの異なる VR 環境に対する被験者の主観的な評価を収集するため、アンケートを実施した。アンケート評価の目的は、VR 環境の違いが被験者の興味や使用イメージ形成、さらには購買意欲に与える影響を明らかにすることである。特に視線情報によって動く環境 (注視動作) が、他の環境 (動かない場合、動き続ける場合) と比較して、より高い興味や購買意欲を引き出すかどうかを検証した。アンケートは以下の設問で構成されており、それぞれの質問に対する回答を数値化した。定量的に分析した。

- ・ 設問 1: 興味がわいたか。
- ・ 設問 2: 扇風機が暑さを軽減してくれると感じたか。
- ・ 設問 3: 炊飯器でご飯を炊くイメージができたか。
- ・ 設問 4: 電子レンジでものを温めるイメージができたか

設問 1 は、各 VR 環境がどの程度ユーザーの興味を引いたかを評価する設問。被験者には、VR 環境の動きや視覚的な特徴が興味を引くかどうかを直感的に回答してもらった。

設問 2~4 では、モデルの視覚的な表現が、利用場面の具体性にどのように影響を与えるかを検証した。評価方法として設問 1~4 は各設問に対して、被験者には 1~10 の範囲で数値評価を行ってもらった。1: 全くそう思わない 10: 非常にそう思うこの評価方法により、各設問に対する回答を数値化し、VR 環境ごとの違いを明確に把握することが可能となった。

5. 実験結果と考察

5.1 エンゲージメント評価考察

本節ではエンゲージメントの評価結果および考察を述べる。表 5-1 に 3 パターンそれぞれの平均利用時間と標準偏

差を示す。

表 5-1: 各パターンの平均利用時間と標準偏差

パターン	平均利用時間 (秒)	標準偏差 (秒)	有意差
静止: 3D モデルが動かない状態	20.684	9.213	あり
常時動作: 3D モデルが常に動き続ける状態	32.518	15.684	あり
注視動作: 視線情報に基づき動作する状態	45.018	12.872	—

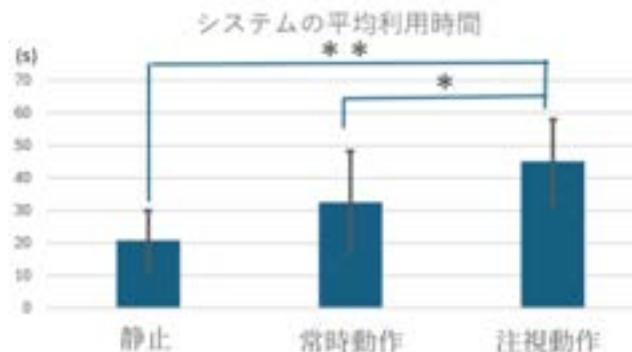


図 5-1: 平均利用時間の有意差

*:p<0.05, **:p<0.01

他の項目については有意差なし

図 5-1 から、視線情報に基づき動作する注視動作が最も長い平均利用時間 (45.018 秒) を示し、エンゲージメントが他の 2 つのパターンと比較して有意に高いことが分かった。静止状態と注視動作がもっとも有意な差を示し、注視動作の有意性を示しました。さらに、標準偏差の値から、各パターンにおける観察時間のばらつきを評価しました。静止および注視動作は、比較的安定したばらつきを示した一方で、常時動作は標準偏差が最も大きく (15.684 秒)、被験者による観察時間の個人差が大きいために示唆されました。これらの結果は、視線情報を活用する動的なインタラクションがユーザの興味を持続させ、観察行動を促進する効果があることを示しています。視線情報によって動作する注視動作では、被験者が長時間観察を続ける傾向が見られ、インタラクションのパーソナライズ効果が興味を向上させる可能性が示唆されます。

一方、常に動き続ける常時動作では観察時間が静止より長いものの、標準偏差が大きく、被験者間での評価が分かれる結果となりました。これは、動き続ける状態が一部の被験者にとって過剰な刺激や情報過多として認識された可能性を示しています。常に止まっている静止状態は、他のパターンに比べてエンゲージメントが低く、ユーザが積極的に関与する機会が限られていると考えられます。このことは、動的インタラクションがユーザの興味喚起に与える重要な役割を強調しています。

5.2 アンケート評価および考察

本節ではアンケート評価の結果とその考察を述べる。本研究では、3 つの異なる VR 環境 (「静止」、「常時動作」、「注視動作」) における被験者の主観的な評価をアンケートを

通じて収集し、興味関心、使用イメージ形成について定量的に分析した。アンケート結果の数値化に基づく分析の結果、注視動作が他の 2 つの環境と比較して、被験者の興味、イメージ形成において最も高いスコアを示した。この章では、各設問の詳細な結果を示し、T 検定による有意差の検証とともに、その考察を展開する。

設問 1. 興味がわいたかでは、VR 環境がどの程度ユーザの興味を引いたかを評価した。以下の表 5-2 に示す通り、注視動作は平均スコア 6.7 で、他のパターンと比較して最も高い評価を得た。

表 5-2：設問 1 における興味関心のスコア比較

パターン	平均値	p 値	有意差
静止	1.9	$p < 0.01$	あり
常時動作	5.0	$p < 0.05$	あり
注視動作	6.7	—	—

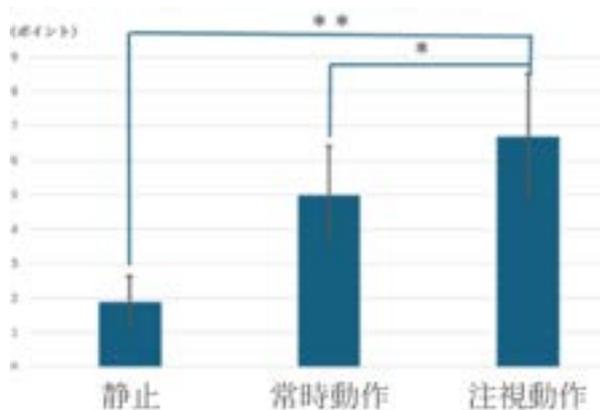


図 5-2：興味関心の有意差

*: $p < 0.05$, **: $p < 0.01$

他の項目については有意差なし

T 検定の結果、注視動作は静止と比較して非常に高い有意な差を示した ($p < 0.01$)。また、常時動作と比較したときも ($p < 0.05$) 有意な差があることが明らかとなった。興味関心の差として段階的に静止 < 常時動作 < 注視動作の順でスコアが上がり主観的なアンケート評価でも注視動作が有意であることが分かった。

設問 2 ではそれぞれ扇風機、炊飯器、電子レンジの使用イメージ形成を測定した。まず扇風機について示す。T 検定の結果として図 5-3 のように、注視動作は静止と比較して有意な差が検出されたが、常時動作とは有意差はないと検出された。

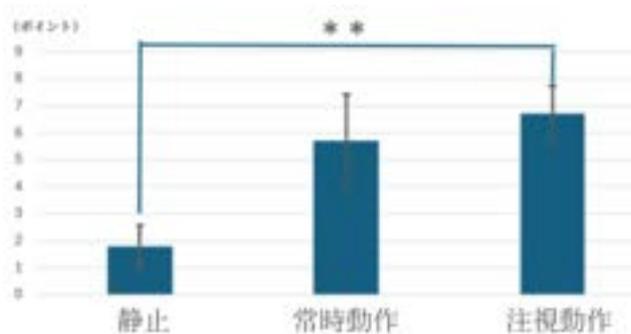


図 5-3: 扇風機の使用イメージの比較

*: $p < 0.05$, **: $p < 0.01$

他の項目については有意差なし

次に炊飯器について示す。

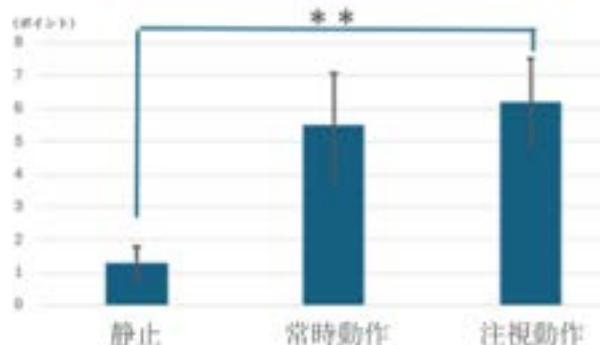


図 5-4: 炊飯器の使用イメージの比較

*: $p < 0.05$, **: $p < 0.01$

他の項目については有意差なし

T 検定の結果として図 5-4 のように、扇風機と同じように注視動作は静止と比較して有意な差が検出されたが、常時動作とは有意差はないと検出された。

最後に電子レンジについて示す。

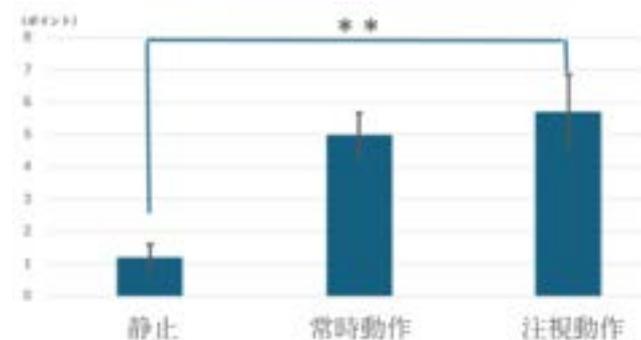


図 5-5: 電子レンジの使用イメージの比較

*: $p < 0.05$, **: $p < 0.01$

他の項目については有意差なし

T 検定の結果として図 5-5 のように、注視動作は静止と比較して有意な差が検出されたが、常時動作とは有意差はないと検出された。この 3 つの図から静止では動きがないため、視覚的な刺激が不足し、製品の利用イメージが付かず想像しにくい、常時動作と注視動作では同じように動きをみることができるため、視覚的な刺激の差がないため、有

意な差がなかったと考えられる。

5.3 考察

本章ではシステムを用いた評価実験結果とその考察について述べた。それぞれのパターンでエンゲージメントの評価とアンケート評価を行い比較を行った。常時動作、注視動作のような動的インタラクションがユーザの興味や使用イメージの形成において極めて効果的であることを示した。定量的なエンゲージメント評価では視線情報に基づいて動作する注視動作がもっとも有意性がありユーザの関心を長引かせ引き出す効果があると示された。視線情報を用いることで、ユーザ体験を個別化し、長時間のエンゲージメントを維持できる可能性が示唆された。

6. 実験結果と考察

6.1 まとめ

本研究ではネットショッピングで購入する前に実物を見られるようにするために、仮想現実上で 3D モデルを表示し動的な動きがあることで興味関心が変わるかを提案した。視線情報を活用して、注目しているモデルのみが動くシステムを作成し、動き続ける場合、動かない場合の 3 パターンでエンゲージメント評価とアンケート評価を行った。エンゲージメント評価では注視動作での平均利用時間が静止

状態の 2 倍近く長く、利用者の興味を最も引いていたと考えられる。アンケート評価では 3D モデルが静止状態と注視動作が非常に有意な差を示し、常時動作とも有意な差を示した。本システムの使用により、現実世界で家電量販店に行く前に仮想現実上で動きのある 3D モデルを先に見ることで興味を掻き立てて、使用するイメージを想像させることができ、興味を持たせることができると明らかになった。

6.2 今後の展望

システムの改善点として、商品が一方しか見れず 360 度見ることができないこと、音などの視覚以外のフィードバックがないため没入感にかけてしまっていること、商品との距離があり近くに寄せられなかったことが挙げられる。システムを考えたとき、家電量販店で商品を持ち上げることはないと考えてしまい、単方向からのアニメーションのみ作成してしまったが、仮想現実という環境を最大限に利用できていなかったと感じる。また、扇風機の音や電子レンジの音などの聴覚的フィードバックも実装することで、より没入感を与えて興味をより掻き立てることができたのではと考えられる。以上にあげられる課題を解決し実装することで、より没入感があり興味関心に影響を与えられるシステムが完成できると考えているため、本システムを発展させていきたい。

参考文献

- [1] 家計消費状況調査結果の概況。
://www.stat.go.jp/data/joukyou/2022ar/gaikyou/pdf/gk01.pdf
- [2] NRI メディアフォーラム：第 322 回 NRI メディアフォーラム (2021)。https://www.nri.com/-/media/Corporate/jp/Files/PDF/knowledge/report/cc/mediaforum/2021/forum322.pdf?la=ja-JP&hash=396365B469B39B585BCE5A74CD340308B02D63F2%EF%BC%89.
- [3] 総務省：平成 27 年インターネットショッピングの利用状況 (2015)。
https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc122400.html.
- [4] 張 テイテイ：ネットショップにおける暗示的動きを表す静止画が消費者行動に与える影響，経営論集 (2023)。
- [5] 落合拓朗 藤田智 益子宗星野准一：視線情報に基づいた嗜好分析から商品推薦を行う VR ショッピングシステム，IPSI SIG Technical Report, Vol. 2019-HCI-184, No. 3 (2019).
- [6] 中島仁大竹恒平：百貨店におけるアイトラッキングデバイスを用いた購買行動の評価，第 86 回全国大会講演論文集 437-438, Vol. 1 (2024).
- [7] Pupil Neon. PupilLabs 社，https://pupil-labs.com/. (Accessed on 1/8/2025)
- [8] Meta Quest3. Meta 社，https://www.meta.com/jp/quest/quest-3/. (Accessed on 1/8/2024).
- [9] Scaniverse-3D Scanner. Toolbox AI 社，https://scaniverse.com/. (Accessed on 1/8/2024).
- [10] Unity. Unity Technologies.，https://unity.com/ja. (Accessed on 1/8/2024).

メタバースにおける仮想アバターとロールプレイングが コミュニケーションに与える影響—プロテウス効果の視点から

仲山芳古^{†1}公立はこだて未来大学^{†1}角薫^{†2}公立はこだて未来大学^{†2}

1. はじめに

近年、メタバース技術の発展と HMD (ヘッドマウントディスプレイ) の普及により、仮想空間での教育、トレーニング、コミュニケーションが活発になっている。特に、メタバース環境における仮想アバター (Virtual Avatar) が、学習や社会的交流の質に影響を与える可能性が指摘されており、多くの研究が進められている。

仮想アバターは、単なる視覚的な表現ではなく、ユーザの行動や心理に影響を及ぼす存在である。例えば、プロテウス効果 (Proteus Effect) [15] とは、ユーザが自身の仮想アバターの外見や特性に影響を受け、その行動や認識が変化する現象を指す。プロテウス効果に関する研究では、Nick Yee と Jeremy Bailenson の研究[15]をはじめ、複数の仮想アバターによる行動の違いを調査した研究や、人外の仮想アバターを用いた研究など、多くの知見が蓄積されている[11, 16]。例えば、魅力的なアバターを使用すると自己評価が向上し、攻撃的なアバターを使用するとより支配的な行動をとることが示されている[9, 14]。この効果は、教育やコミュニケーションの場面でも重要であり、どのようなアバターを使用するかによって、学習成果や社会的インタラクションの在り方が大きく異なる可能性がある[1, 7]。

また、ロールプレイングとは、ある役柄を演じることであり、J.L. Moreno (1889-1974) [8]によるサイコドラマの理論を基礎として発展してきた。ロールプレイングは、医療やビジネスの分野においてスキルトレーニングやコミュニケーション教育[3]に活用されており、仮想アバターを用いたロールプレイングに関する研究では、体験型学習やシナリオベースの学習を行うことで、コミュニケーションスキルの向上や不安の軽減などの効果が示されている[4]。また、HMDを用いたロールプレイングの研究も少数ながら行われており、没入感の向上や仮想環境における学習効果の促進が期待されている[4, 5]。

仮想アバターが学習効果やモチベーションに与える影響についても、多くの研究が行われている。例えば、野澤・角 (2023) は、メタバース環境において仮想アバターの人格が学習に与える影響を調査した[18]。この研究では、開放性 (O)

が低い傾向の人は、学習に関して怠惰で成績が悪い人格のそばで学習するより、勤勉で成績がよい人格のそばで学習の方が成績が向上する可能性があることを示した。しかし、参加者全体の傾向としては、仮想アバターの人格の影響が学習成果に与える明確な違いは確認できなかった。このことから、アバターの影響は、個々の学習者の性格傾向によって異なる可能性があると考えられる。

さらに、バーチャルエージェントの性別による影響も研究されている。Shiban et al. (2015) の研究では、数学の学習において、女性エージェントによる指導を受けた女性学習者のモチベーションが向上することが報告された[14]。これは、類似性仮説 (Similarity Hypothesis) [2] に基づき、学習者が自身と類似したエージェントに対して親近感を抱きやすいことに起因すると考えられる[1, 7, 10, 13]。一方で、Baylor & Kim (2004) は、男性エージェントが専門家として認識されやすく、その結果、学習成果の向上につながることを示した[1]。また、Chang et al. (2019) は、「男性らしい」と認識される話題では、男性エージェントのアドバイスの方が受け入れられやすいことを示した[3]。Moreno et al. (2002) は、参加者の解釈として、女性エージェントは「女性的すぎる」と評価され、家庭教師の役割にはふさわしくないと見なされる傾向があることを示した[9]。

仮想アバターの違いが学習やモチベーションだけでなく、コミュニケーションの質にも影響を与えることも明らかになっている。例えば、山田・角 (2024) は、AI アバターと人間が混在する仮想授業環境において、AI アバターの振る舞いが学級の雰囲気や学習意欲に与える影響を調査した[19]。その結果、否定的な発言を行う AI アバターと共に授業を受けた場合、学級の雰囲気が悪化し、協調性の高い参加者ほどストレスを感じやすい傾向が示された。この研究は、メタバース環境においてアバターの社会的影響力が強く、単なる視覚的な要素以上の影響を持つことを示している。

このように、仮想アバターの特性がユーザの行動や認知に与える影響は、学習や教育、さらにはコミュニケーションの場面において重要な研究領域となっている[1, 7, 11]。しかし、従来の研究では、シナリオベースのロールプレイングが

主流であり[4, 5], ユーザが自由に役割を選択し, インタラクションする環境での影響については未解明の部分が多い.

本研究では, メタバース環境における自由度の高いロールプレイングを通じて, 仮想アバター, プロテウス効果, およびコミュニケーションへの影響を検証する. 特に, 社会問題に関連するアバターを使用することが, 議論の活性化や意識の変化にどのような影響を及ぼすかを調査する. 本研究の目的は, メタバース環境でのロールプレイングが, ユーザの意識や行動に与える影響を明らかにすることである.

2. 実験システム

本研究では, メタバースプラットフォーム「Cluster」上に専用のワールドを作成し, 仮想アバターを用いたロールプレイングを実施した. 実験の目的は, 仮想アバターの種類がメタバース環境におけるディスカッションの質や参加者の意識にどのような影響を与えるかを調査することである. 特に, 社会問題に関連するアバターを使用することが, 議論の活性化や意識の変化にどのような影響をもたらすかを検証する.

本実験では, ジェンダー問題, 動物愛護問題, 子供の貧困問題の 3 つの社会問題について, 被験者が仮想アバターを使用しながらディスカッションを行った. 社会問題の選定に関しては, 日本総研 (2022) が国内の中学生, 高校生, 大学生を対象に実施した若者意識調査を参考にした[17]. また, 実験環境の設計にあたり, 仮想空間における学習環境の影響を調査した先行研究を踏まえ, アバターの種類が議論の質や学習動機に及ぼす影響を評価するための設定を構築した.

本実験は, ロールプレイングの練習と本番のディスカッションを繰り返す形式で実施された. ロールプレイングの練習のための「プラクティスワールド(図 1)」と, 実際にディスカッションを行う「ディスカッションワールド(図 2)」の 2 種類の仮想空間を作成した. プラクティスワールドには, 仮想アバターの外見を確認するための鏡, ロールプレイングの参考資料 (設定書 (図 7)), 時間を計測するためのタイマーを設置し, 参加者が本番前にキャラクターとしての振る舞いを練習できる環境を整えた. 一方, ディスカッションワールドには, 社会問題のトピックを示す掲示板, ディスカッション用の机と椅子, 時間管理のためのタイマー, 仮想アバターの外見を確認する鏡を設置し, 参加者がメタバース環境内で議論を行いやすい環境を構築した.

実験では, 2つの異なる条件のアバター群(図 3)を用意し, それぞれの影響を比較した.

1. **A 群 (関連アバター群)**: 各社会問題に関連する仮想アバター (女性のアバター, 動物のアバター, 子供のアバター)
2. **B 群 (無関係アバター群)**: 各社会問題と関連性のない仮想アバター (ロボットのアバター)

被験者は, それぞれの群にランダムに割り当てられ, ディスカッションを行った. アバターのデザインは, 参加者に自分のキャラクターとして認識されやすく, プロテウス効果の影響を適切に評価できるよう考慮して選定した.

本研究では, 没入感やアバターとユーザ自身の心理的な同期を考慮し, VR ヘッドセット「Meta Quest 2」を使用して実験を実施した. 実験の前には, 参加者に Meta Quest 2 の基本的な操作方法を説明し, ミラーリング機能を用いて実験の流れを確認してもらった. また, 参加者は個別の防音室に配置され, 外部の影響を排除した状態で実験を行った.

ロールプレイングの練習では, 参加者が自由に発話できる環境を提供し, ボイスチェンジャーの調整や発声練習を含めた準備を行うように指示した. 本実験では, 従来のシナリオベースのロールプレイングとは異なり, 参加者が自由にキャラクターを解釈し, 発言内容や行動を決定できるような自由度の高いロールプレイングを採用した. この手法により, アバターの影響がどのように議論の進行や参加者の意識に影響を与えるかを詳細に分析することを目的とした.

本実験の設計は, プロテウス効果に関する既存の研究を参考にしており, 特にアバターの外見や設定が参加者の行動や認識に与える影響を測定するための工夫を施している. 実験後には, ディスカッションの録画データ, アンケート調査, 自由記述の回答を分析し, 仮想アバターが議論の活性化や参加者の社会問題に対する意識にどのような影響を及ぼすかを評価する.

3. 実験

本研究では, メタバース環境において, 仮想アバターの特性がディスカッションの活性化や参加者の意識に与える影響を調査するための実験を実施した. 特に, 社会問題に関連するアバターを使用することが, 議論の進行や参加者の意識の変化にどのように影響を及ぼすかを検証することを目的とする. 以下に, 実験の手続きと具体的な進行方法について述べる.

本実験は, 2名の被験者がペアとなり, 仮想アバターを使用しながら社会問題についてディスカッションを行う形式で実施された. 被験者は 18~24 歳の大学生 20 名 (男性 14 名, 女性 6 名) であり, ランダムにペアを組んだ. ディスカッションのテーマとして, ジェンダー問題, 動物愛護問題, 子供の貧困問題の 3 つを設定し, 各ペアがそれぞれのテーマについて議論を行った(図 6).

実験では, 仮想アバターの種類による影響を比較するため, 2つの異なる条件を設定した (図 4).

- **A 群 (関連アバター群)**: ディスカッションテーマに関連するアバター(女性, 動物, 子供)を使用する条件.
- **B 群 (無関係アバター群)**: ディスカッションテーマとは関係のないアバター(ロボット)を使用する条件.

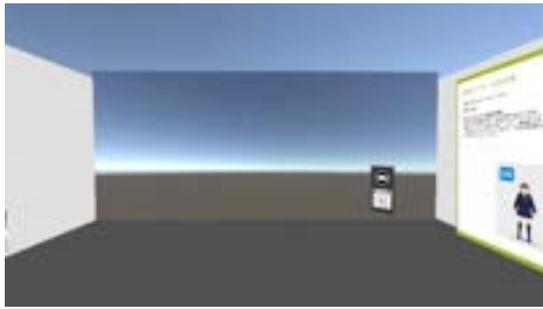


図 1 プラクティスワールドの様子
Figure 1 Scene of the Practice World



図 2 ディスカッションワールドの様子
Figure 2 Scene of the Discussion World



図 3 実験で使った仮想アバター
Figure 3 Virtual avatars used in the experiment

	ジェンダー問題	動物愛護問題	子供の貧困問題
A			
B			

図 4 各群のトピック別仮想アバター組み合わせ
Figure 4 Virtual avatar combinations by group and topic



図 5 設定書の例
Figure 5 Example of a configuration document



図 6 実験の様子
Figure 6 Scene of the experiment

各被験者は、ディスカッション前に「プラクティスワールド」(図 6) でロールプレイングの練習を行った後、「ディスカッションワールド」(図 7) へ移動し、本番のディスカッションを行った。

プラクティスワールドでは、被験者が仮想アバターの外見を確認し、設定書(図 8)をもとにキャラクターの役割を演じる準備を行った。ロールプレイングの練習時間は 3 分間とし、参加者には「自由にキャラクターの行動や発話内容を決定し、どのように振る舞うか考えてみるように」と指示した。また、ボイスチェンジャーの調整や発声練習を行う機会も提供し、各アバターになりきるプロセスを支援した。

本実験では、没入感や仮想アバターとの心理的同期を考慮し、VR ヘッドセット「Meta Quest 2」を使用した。参加者は個別の防音室に配置され、外部の影響を受けずに仮想空間での議論に集中できる環境を整えた。

ディスカッションワールドでは、被験者が仮想アバターを用いた状態で、5 分間のディスカッションを行った。議論のテーマに関しては「ディスカッションの目的は、特定の問題について議論を深めることにある」と事前に教示し、被験者には「積極的に議論を進めるように」と指導した。なお、研究者はワールド内に存在したが、録音・記録のみを行い、ディスカッションの内容には関与しなかった。

ディスカッション終了後、被験者にはアンケート(表 1)に回答してもらった。アンケートでは、

1. 社会問題に対する意識の変化(例:「ディスカッションを通じて、ジェンダー問題についての認識が変わりましたか?」)
2. 行動意欲の変化(例:「今後、動物保護活動に参加したいと思いませんか?」)
3. アバターとの同一視(例:「使用したアバターになりきれたと感じましたか?」)
4. ディスカッションの満足度(例:「メタバースでの議論は対面と比べて発言しやすかったですか?」)

の 4 つのカテゴリーに分けて評価を行った。各質問は 5 段階のリッカート尺度を用いて評価し、自由記述欄も設けた。また、会話の録音データを分析し、発言の回数や内容の多様性、議論の活性度を比較した。特に、関連アバター群(A 群)と無関係アバター群(B 群)の間で、ディスカッションの流れにどのような違いがあるかを検討した。

本実験の目的は、仮想アバターがメタバース環境における議論の進行や参加者の意識に与える影響を定量的・定性的に評価することである。次章では、実験の結果について述べる。

実験結果

本実験では、メタバース環境における仮想アバターの特性

がディスカッションの活性化や参加者の意識にどのような影響を与えるかを検証するため、ディスカッション後のアンケート結果と録音データの分析を行った。本章では、アンケート調査の結果とディスカッションの発話データの分析結果について述べる。

3.1 アンケート結果の分析

本研究では、表 1 に示すアンケート項目に基づき、被験者の**社会問題に対する意識の変化、行動意欲、アバターとの同一視、ディスカッションの満足度**について評価を行った。

表 1 動物愛護問題のアンケート内容の例

Table 1 Examples of survey content on animal welfare issues

	アンケート内容
質問 1	あなたはどのアバターを使用しましたか
質問 2	動物愛護問題について何を話しましたか? 箇条書きで教えてください
質問 3	あなたはこれまでに比べて、動物愛護問題を意識しようと思いましたが
質問 4	あなたは動物愛護問題は絶対に解決すべき問題だと思いますか
質問 5	あなたは、動物愛護問題の解決に向けて支援団体への寄付を行おうと思いませんか
質問 6	あなたは、動物愛護問題の解決に向けてボランティアに参加しようと思いませんか。
質問 7	あなたは、実際に市役所や動物管理センターから動物を引き取ってくれますか
質問 8	あなたは動物愛護問題についてどう思いましたか。自由に書いてください。(解決法、感想、その他のことでも)
質問 9	あなたはこのアバターになりきれましたか?
質問 10	それはなぜですか?
質問 11	あなたはこのアバターを使用してどんな気持ちになりましたか

表 2 に、各社会問題ごとに A 群(関連アバター群)と B 群(無関係アバター群)の点数平均を示す。質問 3(社会問題への意識変化)、質問 4(社会問題の解決への意識)、質問 5(社会問題への支援意欲)、質問 6(ボランティア参加意欲)、質問 7(実際の行動への移行意欲)について、A 群と B 群の間で有意な差があるかを Welch の t 検定で調査した。その結果、A 群と B 群の間に有意な差は確認されなかった(表 2)。

表 2 各社会問題における群ごとの点数平均

Table 2 Average scores per group for each social issue

社会問題	群	Q3	Q4	Q5	Q6	Q7	合計平均
ジェンダー問題	A 群	3.30	4.10	2.60	2.80	3.10	15.90
	B 群	3.40	3.70	2.80	2.60	2.50	15.00
動物愛護問題	A 群	4.20	4.70	3.00	3.20	1.20	16.30
	B 群	3.70	4.70	3.40	3.30	1.60	16.70
子供の貧困問題	A 群	4.00	4.70	3.50	3.40	1.60	17.20
	B 群	3.50	4.00	2.70	2.90	1.60	14.70

また、アバターとの同一視に関する質問(質問 9)では、A 群の方が B 群よりも「アバターになりきれた」と回答した割合が高かった(表 3)。これは、社会問題に関連するアバターを使用することで、参加者がより役割になりきりやすいことを示唆する。一方で、B 群では「アバターが自分とは関係ないと感じた」という意見が見られた(表 4)。

表 3 質問 9:「あなたはこのアバターになりきれましたか?」の結果 (%)

Table 3 Results of Question 9: "Did you fully embody this avatar?"

社会問題	A 群	B 群
ジェンダー問題	70.0%	50.0%
動物愛護問題	60.0%	60.0%
子供の貧困問題	60.0%	60.0%
総合	63.3%	56.7%

ディスカッションの自由記述回答を分析したところ、A 群では「社会問題について考えさせられた」「議論がスムーズに進んだ」といった肯定的な意見が多く、一方で B 群では「議論が深まりにくかった」「特にアバターの影響を感じなかった」といった意見が見られた(表 4)。

表 4 A 群・B 群の自由記述における特徴的な発言の比較 (A 群(関連アバター群)と B 群(無関係アバター群)の発言内容の違い)

Table 4 Comparison of Characteristic Statements in Free Responses Between Group A and Group B (Differences in responses between Group A (Relevant Avatar Group) and Group B (Unrelated Avatar Group))

社会問題	A 群 (関連アバター群)	B 群 (無関係アバター群)
ジェンダー問題	「ボランティアや行動を起こす必要がある」 「自身の偏りを修正する必要がある」	「もともと関心がなかった」 「メタバースでの交流でも関心は深まらなかった」
動物愛護問題	「動物の視点に立って考えることが大事」 「動物たちの命を尊重すべき」 「心が痛んだ」	「動物虐待は良くないが、医療のためには仕方ない」 「あまり関心がない」
子供の貧困問題	「自身の経験をもとに貧困を理解した」 「政府の制度を見直すべき」	「解決したいならすればいい」 「自分の力では何もできない」

3.2 合計点数の比較結果(重みづけ後の分析)

アンケート結果の分析において、質問 3~7 の各項目は内容

や回答の難易度が異なるため、それらを単純な平均値で比較することには限界がある。そこで、本研究では各質問の回答の難易度を考慮し、重みづけを行った点数で比較を実施した。

まず、A 群と B 群で各社会問題ごとの合計点数の平均値を比較した。その結果、子供の貧困問題では A 群の合計点数が B 群よりも高い傾向が見られたが、Welch の t 検定の結果、統計的に有意な差は確認されなかった ($p > 0.05$)。

しかし、各アンケート項目には回答のしやすさに違いがあり、例えば「社会問題について考えようと思うか?」(質問 3)に「強くそう思う」と回答するのと、「実際に支援活動に参加しようと思うか?」(質問 7)に「強くそう思う」と回答するのでは、後者の方が難易度が高い可能性がある。そこで、大学生 4 名に依頼し、質問 3「社会問題への意識の変化」に「強くそう思う」と回答するのと比較して、質問 4~7 に「強くそう思う」と回答するのがどれくらい難しいかを +0~10 点の範囲で評価してもらった。この評価結果をもとに、各質問の点数に重みづけを行い、最終的な合計点数を再計算した(表 5)。

重みづけ後の点数を用いて Welch の t 検定を行った結果、ジェンダー問題、動物愛護問題、子供の貧困問題のいずれにおいても、A 群と B 群の間に統計的に有意な差は見られなかった ($p > 0.05$)。ただし、子供の貧困問題の結果においては、A 群の点数 (31.45) が B 群の点数 (25.80) よりも高い傾向が見られた ($t(18)=1.750, p=0.097$)。この結果は、子供の貧困問題に関しては、関連するアバターを使用することで、より強い問題意識を持つ可能性があることを示唆している。

ジェンダー問題と動物愛護問題では、A 群と B 群の間で数値の差が小さく、仮想アバターの種類が参加者の意識に及ぼす影響は限定的であることが示された。特に動物愛護問題では、B 群の方が高いスコアを示しており (A 群: 25.80, B 群: 26.20)、必ずしも関連アバターを使用することが意識向上につながるわけではない可能性がある。

表 5 重みづけ後の各社会問題における群ごとの点数平均
Table 5 Average scores per group for each social issue after weighting

社会問題	群	Q3	Q4	Q5	Q6	Q7	合計平均
ジェンダー問題	A 群	3.30	8.10	3.85	5.05	5.35	25.65
	B 群	3.40	7.70	4.05	4.85	4.75	24.75
動物愛護問題	A 群	4.20	8.20	4.25	4.95	4.20	25.80
	B 群	3.70	8.20	4.65	5.05	4.60	26.20
子供の貧困問題	A 群	4.00	8.95	7.25	6.90	4.35	31.45
	B 群	3.50	8.25	6.45	6.40	4.35	25.80

3.3 結果のまとめ

本実験の結果から、仮想アバターの種類が社会問題に対する意識や行動意欲に有意な差を生むことは確認できなかった。しかし、アバターとの同一視が強い場合、議論の活性化や発話の多様性が向上する可能性があることが示唆された。特に、A 群では、社会問題についての議論が感情的になりやすく、参加者が議論に対する関心を高める効果があった。

一方で、B 群では議論の進行が淡泊になりやすく、社会問題に対する認識の変化が生じにくい傾向が見られた。これは、プロテウス効果の影響を受けにくいアバターを使用した場合、議論の進行が客観的になり、議論の活性度が低下する可能性を示している。

4. 考察

本研究では、メタバース環境における仮想アバターの種類がディスカッションの活性化や参加者の意識に与える影響を検証した。実験の結果、仮想アバターの種類によって社会問題に対する意識や行動意欲に有意な差は確認されなかった。しかし、自由記述の分析や発話データの解析から、関連するアバターを使用することで議論の進行や発話の多様性に一定の影響を与える可能性が示唆された。本章では、これらの結果について詳細に考察し、本研究の意義と今後の展望について述べる。

4.1 仮想アバターと議論の活性化

本研究の結果から、社会問題に関連するアバター (A 群) を使用した場合、議論の活性化に一定の影響を与えた可能性があることが示された。具体的には、A 群では発話内容に感情を伴う表現が多く見られ、「社会問題について考えさせられた」「議論がスムーズに進んだ」といった肯定的な意見が多かった。一方、B 群では「議論が深まりにくかった」「特にアバターの影響を感じなかった」といった意見が見られた。

これは、プロテウス効果[15] の影響を受けた結果、参加者がアバターの特性に適応し、議論に対する関心や感情的関与が高まった可能性を示唆する。本実験では、社会問題に関連するアバターを使用することで、参加者がその問題により深く関与しようとする傾向が見られた。

また、発話内容の分析から、A 群では感情を込めた発言や主体的な提案が多く見られたのに対し、B 群では客観的な情報提供にとどまる発言が多かった。この結果は、仮想アバターが議論のスタイルに影響を及ぼし、議論の内容がより主体的・感情的なものになる可能性を示している。

4.2 アバターの影響の個人差と社会問題の種類

アンケートの重みづけ後の分析では、ジェンダー問題と動物愛護問題では A 群と B 群の間にほとんど違いが見られなかったが、子供の貧困問題に関しては A 群の方が高い点数を示す傾向が見られた。これは、社会問題の種類によって

アバターの影響が異なる可能性を示唆している。

この結果を解釈するには、社会問題の性質とアバターの関連性の強さを考慮する必要がある。子供の貧困問題では、A 群のアバターとして「子供」を使用したため、被験者が自身の議論の対象に直接的な関与を感じやすかったと考えられる。これは、自己関連づけ (Self-relevance) の効果と関連しており、議論の対象と自身のアバターの特性が一致するほど、より強い感情的な関与を生む可能性がある。

一方で、ジェンダー問題や動物愛護問題では、アバターの影響が限定的であった。特に動物愛護問題では、B 群の方が高いスコアを示しているものもあり、仮想アバターの種類が必ずしも意識向上に直結しないことが示された。この結果は、仮想アバターの効果は社会問題の特性や個人の価値観によって異なり、一律の効果を期待することは難しいことを示唆している。

4.3 仮想アバターの同一視と議論の質

本研究では、アバターとの同一視が議論の進行に影響を与える可能性が示された。アンケート結果では、A 群の方が「アバターになりきれた」と感じた割合が高かった。これは、社会問題に関連するアバターを使用することで、参加者がその役割に没入しやすくなった可能性を示唆する。

先行研究では、学習環境において、仮想エージェントの外見が学習者のモチベーションや成果に影響を与えることが報告されている (Baylor & Kim, 2004; Shibata et al., 2015) [1, 14]。本研究の結果もこれらの知見と一致し、議論の場面においても、アバターの特性が認知的・感情的な関与に影響を及ぼしうることが示唆された。

しかし、本研究ではアバターの特性が学習成果にどのような影響を与えるかについては明確な結論を得ることができなかった。例えば、野澤・角 (2023) の研究では、仮想アバターの人格が学習成果に与える影響が個人の性格特性によって異なることが示されており、本研究でもアバターの影響が参加者の性格や経験によって変化する可能性があると考えられる。

4.4 本研究の意義と今後の課題

本研究は、メタバース環境において自由度の高いロールプレイングを用いたディスカッションの影響を検証した点に意義がある。従来のシナリオベースのロールプレイングとは異なり、プロテウス効果を応用した自由な議論環境を構築し、仮想アバターが議論の質にどのような影響を与えるかを検証した。

しかし、本研究にはいくつかの限界がある。まず、サンプルサイズが小さく、被験者の特性が限定的であったため、より多様な属性の参加者を含めた研究が求められる。また、長期的な影響については検討していないため、仮想アバターが継続的な議論の変化や社会問題に対する行動の変容にどのように寄与するのかについて、今後の研究が必要である。

今後の研究では、ディスカッションの言語データを精査し、仮想アバターの種類が議論の内容にどのような影響を与えるのかをさらに詳細に検討する必要がある。

本研究の結果は、メタバース環境における教育やコミュニケーションのデザインに示唆を与えるものである。例えば、学習分野では、学習内容と関連性の高いアバターを活用することで、学習者の関心や主体性を高めることができる可能性がある。

5. まとめ

本研究では、メタバース環境における仮想アバターがディスカッションの活性化や参加者の意識に与える影響を検証した。結果として、社会問題に関連するアバターを使用することで、議論の活性化や発話内容の多様性が向上する可能性が示唆された。一方で、社会問題に対する意識の変化や行動意欲においては有意な差は確認されなかった。これらの結果は、仮想アバターが議論のスタイルに影響を与えるものの、意識や行動の変容には他の要因も関与することを示している。今後の研究では、長期的な影響やより多様な参加者を対象にした検討が求められる。

参考文献

- [1] Baylor, A. L., & Kim, S.: Virtual Agents as Pedagogical Experts, *International Journal of Artificial Intelligence in Education*, Vol. 14, pp. 1–23 (2004).
- [2] Byrne, D.: *The Attraction Paradigm*, Academic Press (1971).
- [3] Chang, Y., et al.: Gender Stereotypes in Virtual Learning, *Journal of Educational Technology & Society*, Vol. 22, No. 3, pp. 45–57 (2019).
- [4] Fominykh, M., Leong, P., & Cartwright, B.: Role-playing and Experiential Learning in a Professional Counseling Distance Course, *Journal of Interactive Learning Research*, Vol. 29, No. 2, pp. 169–190 (2018).
- [5] Gu, X., Li, S., Yi, K., Yang, X., Liu, H., & Wang, G.: Role-Exchange Playing: An Exploration of Role-Playing Effects for Anti-Bullying in Immersive Virtual Environments, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 29, No. 10, pp. 4215–4228 (2023).
- [6] Krämer, N. C., et al.: The Impact of Virtual Human Gender and Personality on Learning Outcomes in STEM Education, *Computers & Human Behavior*, Vol. 64, pp. 23–34 (2016).
- [7] Mcarova, A., et al.: The Role of Gender in Virtual Learning Environments: Implications for STEM Education, *Educational Technology Research and Development*, Vol. 67, No. 4, pp. 975–992 (2019).
- [8] Moreno, J. L.: *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama* (1953).
- [9] Moreno, R., et al.: Gendered Perceptions of Virtual Tutors, *Educational Psychology Review*, Vol. 14, No. 2, pp. 111–126 (2002).
- [10] Nass, C., & Moon, Y.: Machines and Mindlessness: Social Responses to Computers, *Journal of Social Issues*, Vol. 56, No. 1, pp. 81–103 (2000).
- [11] Ratan, R., Rikard, R., Wanek, C., McKinley, M., Johnson, L., & Sah, Y. J.: Introducing Avatarification: An Experimental Examination of How Avatars Influence Student Motivation, *49th Hawaii International Conference on System Sciences (HICSS)*, pp. 51–59 (2016).
- [12] Rosenberg-Kima, R., Baylor, A. L., Plant, E. A., & Doerr, C. E.: The Role of Gender and Agent-Embodiment in Promoting Self-Efficacy for Learning Computer Programming, *Proceedings of the International Conference on Intelligent Virtual Agents*, pp. 78–89 (2008).
- [13] Rosenberg-Kima, R., et al.: Similarity Hypothesis in Virtual Learning Environments, *Computers & Education*, Vol. 51, No. 2, pp. 1191–1198 (2008).
- [14] Shiban, Y., et al.: Effect of Virtual Agent Gender on Mathematics Motivation, *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, (2015).
- [15] Yee, N., & Bailenson, J.: The Proteus Effect: The Effect of Transformed Self-Representation on Behavior, *Human Communication Research*, Vol. 33, No. 3, pp. 271–290 (2007).
- [16] 小柳陽光, 鳴海拓志, Lugin, J.-L., 安藤英由樹, 大村廉: ドラゴンアバターを用いたプロテウス効果の生起による高所に対する恐怖の抑制, *日本バーチャルリアリティ学会論文誌*, Vol. 25, No. 1, pp. 2–11 (2020).
- [17] 日本総研: 「2022 若者意識調査 - サステナビリティ, 金融経済教育, キャリア等に関する意識」 (2023).
<https://www.jri.co.jp/MediaLibrary/file/pdf/company/release/2023/0810.pdf> (閲覧日: 2025年2月28日).
- [18] 野澤ひな, 角薫: メタバースにおける仮想アバターの人格が学習に与える影響, *ことば工学研究会, 人工知能学会* (2023).
- [19] 山田純平, 角薫: 仮想空間の授業において AI アバターが学習に与える影響の調査, *ヒューマンコンピュータインタラクション研究会, 情報処理学会* (2024).

宅内デジタルツインを活用した 環境センサによる生活行動認識手法の検討

菊池尊勝^{†1}

奈良先端科学技術大学院大学

松井智一^{†2}

奈良先端科学技術大学院大学

諏訪博彦^{†3}

奈良先端科学技術大学院大学

安本慶一^{†4}

奈良先端科学技術大学院大学

1. はじめに

宅内行動認識は、スマートホーム、ヘルスケア、介護支援など幅広い応用が期待される技術であり、住環境における生活の質を向上させるための重要な要素とされている。特に、高齢者の見守りや健康管理の分野では、行動認識技術を活用することで、異常行動の検知や生活習慣の改善支援が可能となる。しかし、従来の行動認識手法は、主にカメラ、マイク、慣性計測装着 (IMU) といったセンサデータを活用することが一般的であり、これらの技術により高い識別精度が実現されている。一方で、これらの手法は、プライバシーの侵害や心理的抵抗といった課題を抱えており、特に家庭環境においては実用化の障壁となることが指摘されている。そのため、プライバシーを保護しつつ行動認識を可能とする手法の開発が求められている。

このような背景から、カメラやマイクを使用せず、温湿度センサや人感センサなどの環境設置型センサ (以下、環境センサとする) を活用して行動認識する手法が注目されている。環境センサは、家庭内の行動を間接的にセンシングし、プライバシーの保護を図ることができるため、家庭環境に適した技術である。しかし、環境センサを用いた行動認識には、以下の3つの課題が存在する。

第一の課題は、環境センサの情報が限定的であるため、単体では高精度な行動識別が難しい点である。環境センサは、行動の直接的な特徴を捉えるのではなく、周囲の環境変化

を通じて間接的に行動を推定するため、時系列性が強く、短期間のデータのみでは十分な識別が困難である。

第二の課題は、家庭環境毎にセンサ配置や家具レイアウトが異なるため、モデルの適用範囲が制限され、汎化性能が低下する点である。環境センサの設置場所が異なると、同様の行動であっても取得されるセンサデータに違いが生じるため、学習したモデルが他の家庭環境に適用しづらいという問題がある。

第三の課題は、行動認識モデルの学習に必要な大規模データの収集が困難であり、特に実環境でのデータ取得には高いコストがかかる点である。異なる家庭環境で十分なデータを収集するには時間とリソースが必要であり、居住者の負担も大きい。

本研究では、これらの課題に対処するために、環境センサを活用したマルチモーダル行動認識の統合手法、家庭環境間の違いを吸収可能な汎化モデルの設計、デジタルツインを活用したデータ拡張の検討を行う。

環境センサの情報が限定的である問題に対しては、統一表現空間を構築し、クロスモーダル対照学習を適用することで、環境センサの情報を他のモーダリティと統合し、識別精度を向上させる。また、家庭環境毎のセンサ配置や家具レイアウトの違いによる汎化性能の低下に関しては、家庭特性エンコーダと居住者特性エンコーダを導入し、家庭環境毎の差異を学習することで汎化性能を向上させる。さらに、データ収集のコストが高いという課題に対しては、デジタルツインを活用し、家庭環境を仮想的に再現することで、多様なデータを生成し、学習データの拡張を行う。

本稿の構成は、次のとおりである。2章では、関連した既存研究を紹介し、本研究の位置付けを明らかにする。3章では提案するマルチモーダルセンシングを利用した宅内でのマイクロ行動認識汎化モデルの開発手法について説明し、4章では今後の課題と研究の展望について述べる。5章では、結論を述べる。

Conference Manuscript Format for Academy of Behavior Transformation by AIoT (BTI)

^{†1} TAKAMASA KIKUCHI, Nara Institute of Science and Technology

^{†2} TOMOKAZU MATSUI, Nara Institute of Science and Technology, 理化学研究所革新知能統合研究センター (AIP) RIKEN Center for Advanced Intelligence Project AIP

^{†3} HIROHIKO SUWA, Nara Institute of Science and Technology, 理化学研究所革新知能統合研究センター (AIP) RIKEN Center for Advanced Intelligence Project AIP

^{†4} KEIICHI YASUMOTO, Nara Institute of Science and Technology, 理化学研究所革新知能統合研究センター (AIP) RIKEN Center for Advanced Intelligence Project AIP

2. 関連研究

本章では、宅内行動認識に関する既存研究を整理し、本研究がどのような位置付けにあるのかを明確にする。まず、宅内での行動認識技術の基盤となる従来手法を整理する。次にマルチモーダル統合を活用した行動認識研究について述べる。

2.1. 単一モーダリティでの行動認識に関する研究

行動認識の研究は、主にカメラや音声センサ、IMU といったモーダリティを活用する手法が一般的である。これらのデータを用いることで高い認識精度を達成することが可能となるが、それぞれの手法には利点と欠点が存在する。本節では、各モーダリティの従来手法を整理し、それぞれの特徴と課題を明らかにする。

2.1.1. IMU・環境センサを用いた行動認識

IMU ベースの行動認識は、カメラや音声と比較してプライバシーの問題が少なく、ウェアラブルデバイスを用いることで継続的なデータ収集が可能であるため、多くの研究が行われている。[1, 2]

IMU ベースの手法は、モーションデータの連続的な手法が可能である一方で、デバイスの装着を前提とするため、利用者の負担が大きく、デバイスの装着位置によるデータのばらつきが認識精度に影響を与えるという課題や、明確な動作を伴わない特定の行動の認識が他モーダリティと組み合わせないと困難であるという課題がある。

環境センサを利用した行動認識手法は、カメラや音声センサを用いることに抵抗がある家庭環境においては、プライバシーを保護しつつ行動認識を行う手段として注目されている。[3, 4, 5]

環境センサを活用した手法は、取得される情報が間接的であり、行動の発生時刻や持続時間などの精度が他のモーダリティと比較して劣るという課題や、センサ配置に依存しやすく、家具のレイアウト変更やセンサの検出範囲外での行動の認識ができないという課題が存在する。そのため、他のセンサと統合させることで認識精度を向上させる手法が求められる。

2.1.2. カメラ・映像・音声を用いた行動認識

カメラベースの手法は、高い認識精度を実現することが可能であるため、多くの研究が行われている。[6, 7, 8, 9]

カメラベースの手法は、視覚的な情報を直接取得できるため高精度な認識が可能であるが、プライバシーの懸念や心理的な負担が強いため家庭環境への導入には慎重な検討が必要である。また、カメラの設置位置に依存するため、視野外の行動に対する頑健性が低いという課題も存在する。

音声ベースの手法は、カメラと比較するとプライバシーの懸念が低く、IMU と違いデバイスを装着の必要がないため、多くの研究が行われている。[10, 11]

音声 ベースの手法は、背景ノイズの影響を受けやすい点や、同じ行動でも異なる音のパターンを持つ可能性があるため、データの一般化が難しいという課題が存在する。また、複数の居住者がいる環境では、個々の行動を認識するための追加の情報が必要となる課題も存在する。

2.2. マルチモーダル統合を活用した行動認識に関する研究

マルチモーダル統合は、異なるセンサから得られる情報を統合することで、単一モーダリティでは得られない詳細な行動情報を取得し、認識精度を向上させる手法である。近年、IMU や映像、音声などのモーダリティを組み合わせた行動認識の研究が進められており、環境センサと統合することでその有効性をさらに高めることが期待されている。本節では、各モーダリティ統合に関する研究を整理する。

映像データを活用したマルチモーダル統合の研究も活発に行われている。Wang ら [12] は、大規模な映像データを活用し、異なるモーダリティの統合を可能とする基盤モデルを提案した。また、Gao ら [13] は、映像とテキスト情報を組み合わせることで、行動の時間的特徴を効果的に捉える手法を提案した。これらの研究は、映像を中心としたマルチモーダル統合が行動認識の精度向上に寄与することを示しているが、プライバシーの観点から家庭環境での実用化には課題が残されている。

IMU と他のモーダリティを統合する手法も多く提案されている。Chatterjee ら [14] は、IMU と音響データを統合することで、IMU 単体では認識が困難であった行動を補完する手法を提案した。この研究は、身体動作の詳細な情報を取得するための IMU の有用性を示しつつ、追加センサ情報による補完が精度向上に寄与することを示している。しかし、音響データの収集はプライバシーの懸念を伴うため、家庭内での実用化には課題が残されている。

音声データとの統合に関する研究も進められている。Liang ら [11] は、大規模なオンラインビデオデータから音響特徴を学習し、行動認識に適用する手法を提案した。実世界の多様な音響環境を反映したデータを学習することで、より汎用的な音響特徴量の抽出が可能となり、行動認識の適用可能範囲が広がることが示されている。しかし、大規模データを用いることでモデルの表現力は向上するが、学習コストが増大することにより、リアルタイムでの行動認識は困難であるという課題が残されている。

これらの研究に共通する課題として、モーダリティ間の情報補完が不十分である点が挙げられる。従来手法では、

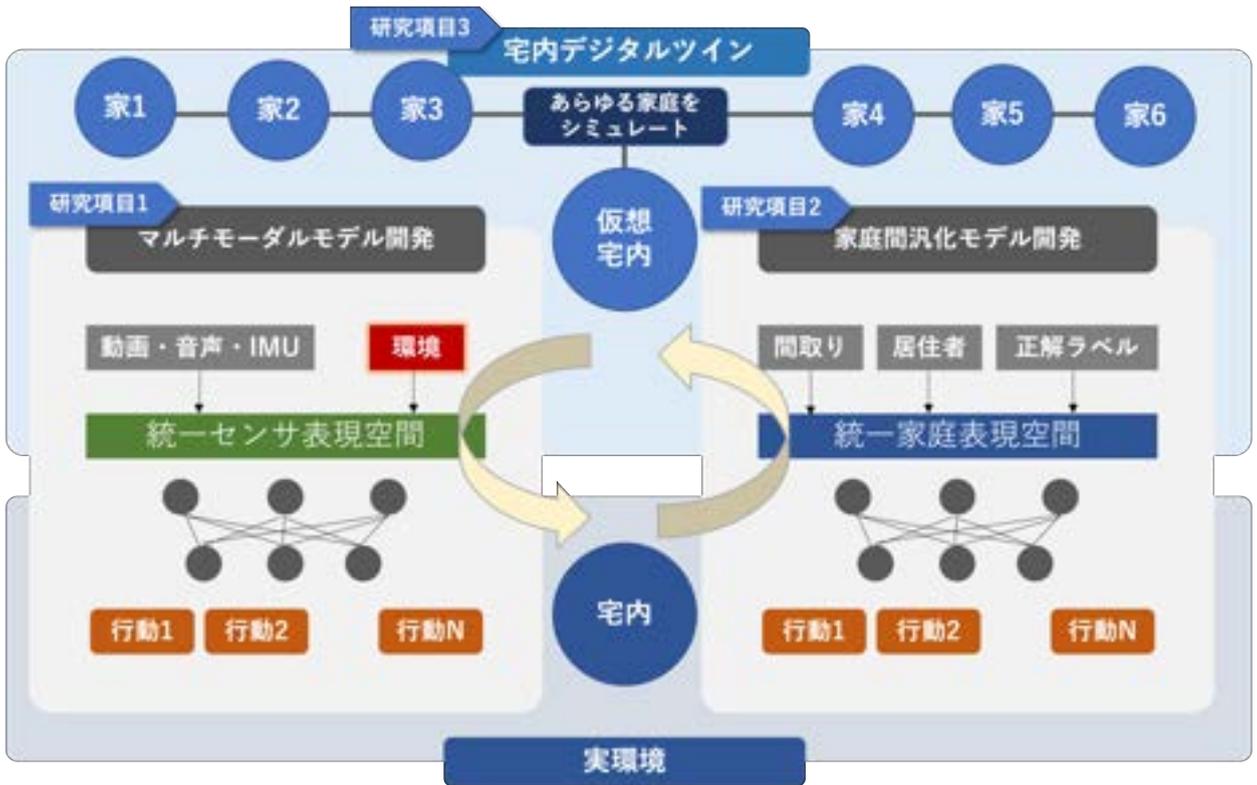


図1 本研究の全体像

単純な特徴レベルでの統合や、意思決定レベルでの統合が主流であり、環境センサの限界を根本的に補完するには至っていない。また、環境センサを活用したマルチモーダル学習において、統一表現空間を構築する研究が少なく、環境センサの情報を効果的に統合するための新たな手法が求められている。

2.3. 本研究の位置付け

本研究では、マルチモーダルセンサを活用し、環境センサのみの情報から高精度な行動認識を実現する汎用的なモデルを構築することを目的とする。従来の行動認識手法では、IMU、ビデオ、音声といったリッチな情報を含むデータを活用することで高精度な認識が可能となっていたが、それらのモダリティが常に利用であるとは限らず、特に家庭環境においてはプライバシーの観点からカメラや音声の使用が制限されることが多い。そのため、センサの情報量が限られてる環境でも高精度な行動認識を実現するための手法が求められている。また、家庭毎にセンサの配置や家具のレイアウトが異なることから、学習済みモデルを異なる家庭環境に適用する際の汎化性能にも限界がある。加えて、行動認識モデルの学習には大量のデータが必要となるが、実際の家庭環境でのデータ収集には高いコストがかかるため、

十分な学習データを確保することが困難である。

本研究は、これらの課題を解決するために、環境センサの情報を用いた高精度な行動認識を実現し、異なる家庭環境への適応性能を向上させるとともに、デジタルツインを活用したデータ拡張による学習効率の向上を図る。この目的を達成するために、本研究では、マルチモーダル統合の最適化、家庭環境間の適応、デジタルツインによるデータ拡張の3つの要素を組み合わせたアプローチを提案する。

3. 提案手法

本研究の全体像を、図1に示す。本研究は、環境センサを活用した行動認識のフレームワークを提案する。本章では、マルチモーダル統合、家庭間汎化、およびデジタルツインによるデータ拡張の3つの観点から、課題を解決するためのアプローチを検討する。

3.1. マルチモーダル学習に向けたセンサデータ統合

クロスモーダル対照学習と知識蒸留を組み合わせることで、統一表現空間を構築するマルチモーダル学習フレームワークを提案する。本手法は、ビデオ、音声、IMU、および環境データを統一的なベクトル空間にマッピングし、モダ

表 1 各モダリティに適用するエンコーダ

モダリティ	エンコーディング手法
環境センサ	1D-CNN, GRU
IMU	LSTM, CNN
画像	ResNet, ViT
音声	Transformer, Wav2Vec

リティ間の整合性を確保しつつ、モダリティの欠損に対して頑健な特徴学習を実現することを目的とする。

従来のマルチモーダル学習では、画像や音声といった主要なモダリティに関しては統一表現空間を学習する手法が提案されている [15] が、環境センサを統一表現空間に統合する手法は少ない。本研究では、教師モデルを用いた統一表現空間の学習と、クロスモーダル対照学習を組み合わせることで、異なるモダリティ間の関係を強化し、環境センサのみでも高精度な推論を可能にする。

3.1.1. 特徴抽出

本手法では、環境センサ、IMU、画像、音声データを統一表現空間に適応させるために適切なエンコーダを適用し、それぞれのモダリティから高次元特徴を抽出する。それぞれのセンサに適したエンコーディング手法を表 1 に示す。

各モダリティの特徴は、それぞれのエンコーダを通じて抽出され、統一表現空間に変換される。

3.1.2. 統一表現空間の学習

環境センサの情報のみを用いて高精度な行動認識を実現するため、マルチモーダル統合を行い、統一表現空間を構築する。統一表現空間の学習においては、以下の 2 つの手法を統合し、異なるモダリティ間の相関を考慮した特徴表現を学習する。

1. クロスモーダル対照学習

クロスモーダル対照学習では、各モダリティの特徴を統一表現空間にマッピングし、異なるモダリティ間の整合性を向上させることを目的とする。具体的には、同じサンプルに対する異なるモダリティの埋め込みが統一表現空間内で近接するように学習し、異なるサンプルの埋め込みとは十分な距離を保つことで、モダリティ間の相互関係を強化する。これにより、統一表現空間内で異なるモダリティ間の特徴が整合性を持ち、欠損モダリティに対しても補完的な情報を提供できる。

2. 知識蒸留

知識蒸留では、学習時にすべてのモダリティを用いて教師モデルを学習し、推論時には環境センサのみを

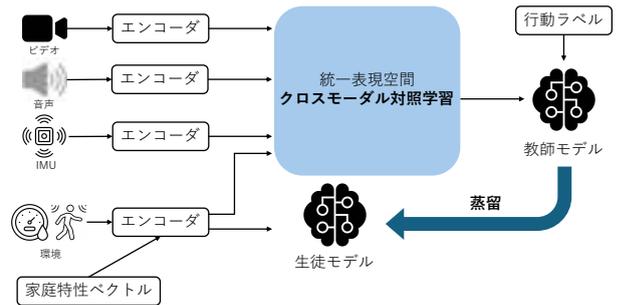


図 2 マルチモーダル学習の全体像

利用する生徒モデルを構築する。教師モデルは統一表現空間を学習し、生徒モデルはその表現を模倣することで、欠損モダリティの影響を軽減する。蒸留損失には、教師モデルと生徒モデルの出力分布の差を最小化するために Kullback-Leibler (KL) ダイバージェンスを適用し、生徒モデルが教師モデルの確率分布を再現できるように最適化する。これにより、生徒モデルは環境センサのみを用いる場合でも、統一表現空間を通じて他モダリティの情報を活用し、識別精度を維持できる。

3.1.3. マルチモーダル学習の全体像

マルチモーダル学習の全体像を図 2 に示す。それぞれのセンサデータに対してエンコーディングを行い、統一表現空間にマッピングを行う。環境センサは情報量が少ないため、家庭特性ベクトルを追加して、情報量を補完する。その後、蒸留を用いて環境センサのみで認識を行うモデルの構築を行う。以下に、手法のフローを示す。

1. 教師モデルの学習

すべてのモダリティの特徴を統一表現空間に投影し、クロスモーダル対照学習を適用することで、統一的な表現を学習する。

2. 生徒モデルの学習

モダリティを部分的に削減した状態で学習し、統一表現空間を通じて教師モデルの知識を模倣する。

3. 知識蒸留の適用

生徒モデルが教師モデルの出力を再現するように蒸留損失を適用し、モダリティ欠損時の影響を最小限に抑える。

3.2. 家庭環境間の汎化モデル

家庭毎の環境特性や居住者の行動パターンは、行動認識モデルの汎化性能に影響を与える。間取りやセンサ配置の違いは、同じ行動に対するセンサの応答を変化させ、居住者

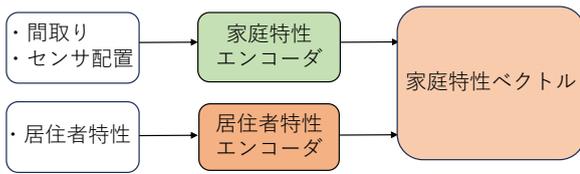


図3 家庭環境間の汎化モデルの全体像

の人数や生活リズムの違いは行動パターンのばらつきを生む。これにより、学習済みモデルの適応性が制限され、認識精度が低下する。

家庭環境間の汎化モデルの全体像を図3に示す。間取りやセンサ配置などの環境要因を学習する家庭特性エンコーダと、居住者の行動特性を考慮する居住者特性エンコーダを統合することで、環境間の違いを補正する手法について検討を行う。エンコーダで得られるデータで環境センサの情報量を補完することで、環境センサのみを用いた行動認識の汎化性能を向上させる。

3.2.1. 家庭特性エンコーダ

家庭環境の違いを考慮し、行動認識モデルの汎化性能を向上させるために、家庭特性エンコーダを構築する。このエンコーダは、家庭の間取り、センサ配置の情報を統合し、環境の違いを補正する。以下に、各要素のエンコーディング方法を示す。

1. 間取り情報のエンコーディング

各部屋の用途をワンホットエンコーディングし、それを埋め込みベクトルとして学習する。さらに、部屋間の関係性を考慮するために、GNNを用いて部屋間の距離や隣接関係を学習し、空間的な構造情報を組み込む。

2. センサ配置の補正

センサ配置の違いを補正するために、Transformerベースの位置埋め込みを適用し、家庭毎のセンサ配置の影響を軽減する。

3.2.2. 居住者特性エンコーダ

家庭環境だけでなく、居住者の行動特性も行動認識モデルの汎化性能に影響を与える。そのため、居住者の人数や生活リズムの違いを考慮するために、居住者特性エンコーダを導入する。このエンコーダは、家庭毎の居住者数や行動パターンを学習し、センサデータへの影響を補正することで、モデルの適応能力を向上させる。

具体的には、居住者数を埋め込みベクトル化し、これを環境センサデータの特徴と結合してエンコーダへ入力する。この手法により、単身世帯と大家族世帯における行動パターンの違いをモデルが適応的に学習し、汎化性能を向上させることが可能となる。

3.3. デジタルツインを活用したデータ拡張

行動認識モデルの汎化性能を向上させるためには、多様な家庭環境に適応可能なデータを学習することが不可欠である。本節では、宅内デジタルツインを活用し、家庭環境を仮想的に再現することで、データ拡張と適応学習を可能にする手法についての検討を行う。

3.3.1. シミュレーション環境によるセンサデータ生成

家庭環境毎の違いを再現し、モデルの適応性能を向上させるために、宅内の物理環境を仮想的にシミュレーションし、センサデータを合成する。本手法では、家庭の3Dモデルを構築し、環境センサの変化や家具の配置変更を仮想的にシミュレーションすることで、実環境でのデータ収集の負担を軽減する。家庭の空間情報は、間取り図を基に3D構造に変換することで生成されるが、LiDARやRGB-Dカメラを用いたスキャン技術を活用することでより精密な環境再現が可能となる。この仮想空間上で、異なる家庭環境を想定した家具配置の変更やセンサの設置場所の違いをシミュレーションし、センサデータへの影響を評価する。

センサデータの生成には、物理ベースシミュレーションを適用し、環境センサデータの変動を再現する。さらにエージェントベースモデリングを活用し、仮想的な居住者の行動をシミュレーションすることで、人感センサの反応を人工的に生成する。これにより、異なる家庭環境におけるセンサデータの変化を事前に学習し、モデルの汎化性能を向上させることができる。

3.3.2. 合成データを活用したデータ拡張

デジタルツイン環境で生成されたセンサデータをそのまま学習に使用するのではなく、実際の家庭環境で収集されたデータと統合し、モデルの精度を向上させる。合成データと実データの統合を行う際に、ドメイン適応技術を活用し、異なる環境間でのデータ分布の違いを補正する。また、シミュレーション環境で生成されたセンサデータの特徴が実データと異なる場合、それを補正するために、スタイル変換技術を適用し、合成データを実データに近づける。このプロセスにより、合成データの信頼性が向上し、学習データとして活用可能な質の高いデータを収取できる。また、センサノイズのシミュレーションを行うことで、実際の家庭環境で発生する測定誤差を再現し、より実データに近いデータを生成する。

3.3.3. フィードバックループを活用した適応学習

モデルの汎化性能をさらに向上させるためには、家庭毎の特性を考慮し、モデルの動的な適応が必要である。そこで、リアルタイムフィードバックを用いた適応学習の仕組みを導入する。具体的には、行動認識結果を居住者が確認し、誤

認識が発生した場合には修正を可能とする機能を導入する。修正されたデータをデジタルツイン環境にフィードバックし、環境特性に適応したモデル更新を行う。

適応学習の実現には、転移学習を適用し、各家庭において特定の行動認識精度が向上するように微調整を行う。これにより、家庭毎の行動特性の違いを反映し、異なる環境でも高精度な行動認識を維持することが可能となる。

4. まとめと今後の展望

本研究では、環境センサを活用した行動認識手法を提案し、家庭環境間の汎化性能向上およびデジタルツインを用いたデータ拡張の可能性についての検討を行った。従来の行動認識手法では、カメラやIMUを主に活用するアプローチが一般的であり、環境センサ単体での高精度な認識手法の検討は十分に進んでいなかった。本研究では、クロスモデル対照学習と知識蒸留を組み合わせることで、環境センサの情報のみから統一表現空間を構築し、行動認識の精度向上を目指した。また、家庭特性エンコーダと居住者特性エンコーダを導入することで、家庭毎の特性を考慮しながら汎化性能を向上させる手法を検討した。さらに、デジタルツインを活用し、仮想環境を用いたデータ拡張と適応学習の可能性についても検討を行った。

しかし、本研究は Position Paper としての議論を主としており、提案手法の定量的な評価は行っていない。提案手法の実装および評価を通じて、環境センサを用いた行動認識の実現可能性を検証し、各アプローチの有効性を定量的に評価することが今後の課題となる。また、デジタルツインを活用したデータ拡張に関しても、仮想環境で生成されたデータと実データの特性の違いを完全に解消することは難しく、特にセンサノイズや居住者の行動特性の再現性に課題が残る。

提案手法の適用範囲を拡張し、より多彩な家庭環境に対応するためには、いくつかの技術的課題を解決する必要がある。まず、家庭環境の多様性に関する問題が挙げられる。本研究では、家庭特性エンコーダを導入することで、家庭毎のセンサ配置や家具の違いを考慮したが、時間とともに変化する動的な環境の影響については十分に考慮できていない。長期的な居住環境の変化や一時的な家具の移動に対応するためには、モデルの動的な更新手法の導入が必要である。また、デジタルツインを活用したデータ拡張の精度向上も課題として挙げられる。仮想環境で生成されたデータと実データの特性の違いを完全に解消することは難しく、特にセンサノイズや居住者の行動特性の再現性に課題が残る。デジタルツイン環境の実データとの整合性を向上させるためには、合成データの適応的な補正手法の導入が求められる。さらに、実環境への適用に関しては、データ取得条

件やセンサノイズの影響によるモデルの精度の変動が課題として挙げられる。提案手法では、異なる家庭環境に適応可能なモデルを構築したが、長期間にわたるモデルの維持や更新には更なる工夫が必要である。特に、居住者の行動フィードバックを活用した適応学習を行う際には、利用者に過度な負担をかけることなく、継続的なモデル更新を実現するための設計が求められる。

今後の研究においては、提案手法の更なる汎用向上と実環境への適用を目指し、いくつかの方向性が考えられる。まず、多様な家庭環境への適応性能を高めるために、より大規模なデータセットの収集と活用が必要である。本研究では、異なる家庭環境間のデータを統合するための手法を提案したが、より多様なデータを用いた学習により、異なる文化圏や生活習慣も適応可能なモデルを構築することが求められる。また、家庭毎のカスタマイズ学習を行うことで、個々の居住者の行動特性をより正確に反映した認識モデルを実現することが可能となる。さらに、環境センサだけでなく、音声データによるゲーミフィケーションを活用しおたアノテーションインターフェースを統合することで、継続的なフィードバックに対する負担を軽減しつつ、認識モデルの精度の向上を図る。

本研究で提案した行動認識手法は、スマートホームシステム、健康管理、介護支援、支援モニタリングなど幅広い応用が期待される。今後の研究では、実環境での適用を視野に入れた実証実験を行い、提案手法の有効性を明確に示すとともに、より実用的なシステムの開発に向けた研究を進めていく必要がある。

参考文献

- [1] Alevizaki, A., Pham, N. and Trigoni, N.: Hierarchical activity recognition with smartwatch IMU, *Proceedings of the 24th International Conference on Distributed Computing and Networking*, pp. 48–57 (2023).
- [2] Thakur, D., Guzzo, A. and Fortino, G.: Attention-based multihed deep learning framework for online activity monitoring with smartwatch sensors, *IEEE Internet of Things Journal*, Vol. 10, No. 20, pp. 17746–17754 (2023).
- [3] Matsui, T., Onishi, K., Misaki, S., Fujimoto, M., Suwa, H. and Yasumoto, K.: Salon: Simplified sensing system for activity of daily living in ordinary home, *Sensors*, Vol. 20, No. 17, p. 4895 (2020).
- [4] Fujiwara, M., Kashimoto, Y., Fujimoto, M., Suwa, H., Arakawa, Y. and Yasumoto, K.: Implementation and evaluation of analog-pir-sensor-based activity recognition, *SICE Journal of Control, Measurement, and System Integration*, Vol. 10, No. 5, pp. 385–392 (2017).
- [5] Guan, Q., Li, C., Qin, L. and Wang, G.: Daily activity recognition using pyroelectric infrared sensors and reference structures, *IEEE Sensors Journal*, Vol. 19, No. 5,

- pp. 1645–1652 (2018).
- [6] Tan, T.-H., Gochoo, M., Huang, S.-C., Liu, Y.-H., Liu, S.-H. and Huang, Y.-F.: Multi-resident activity recognition in a smart home using RGB activity image and DCNN, *IEEE Sensors Journal*, Vol. 18, No. 23, pp. 9718–9727 (2018).
 - [7] Gaidon, A., Harchaoui, Z. and Schmid, C.: Temporal localization of actions with actoms, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 35, No. 11, pp. 2782–2795 (2013).
 - [8] Choi, W., Chen, J. and Yoon, J.: PWS-DVC: Enhancing Weakly Supervised Dense Video Captioning With Pre-training Approach, *IEEE Access*, Vol. 11, pp. 128162–128174 (2023).
 - [9] Cho, S., Kim, Y., Jang, J. and Hwang, I.: AI-to-Human Actuation: Boosting Unmodified AI’s Robustness by Proactively Inducing Favorable Human Sensing Conditions, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 7, No. 1, pp. 1–32 (2023).
 - [10] Wu, J., Harrison, C., Bigham, J. P. and Laput, G.: Automated class discovery and one-shot interactions for acoustic activity recognition, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14 (2020).
 - [11] Liang, D. and Thomaz, E.: Audio-based activities of daily living (adl) recognition with large-scale acoustic embeddings from online videos, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3, No. 1, pp. 1–18 (2019).
 - [12] Wang, Y., Li, K., Li, X., Yu, J., He, Y., Chen, G., Pei, B., Zheng, R., Wang, Z., Shi, Y. et al.: Internvideo2: Scaling foundation models for multimodal video understanding, *European Conference on Computer Vision*, Springer, pp. 396–416 (2024).
 - [13] Gao, J., Sun, C., Yang, Z. and Nevatia, R.: Tall: Temporal activity localization via language query, *Proceedings of the IEEE international conference on computer vision*, pp. 5267–5275 (2017).
 - [14] Chatterjee, S., Chakma, A., Gangopadhyay, A., Roy, N., Mitra, B. and Chakraborty, S.: LASO: Exploiting locomotive and acoustic signatures over the edge to annotate IMU data for human activity recognition, *Proceedings of the 2020 International Conference on Multimodal Interaction*, pp. 333–342 (2020).
 - [15] Lu, J., Clark, C., Lee, S., Zhang, Z., Khosla, S., Marten, R., Hoiem, D. and Kembhavi, A.: Unified-io 2: Scaling autoregressive multimodal models with vision language audio and action, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 26439–26455 (2024).

ペット型人工物の感情表現と動作を用いた 習慣改善のための説得技術

原田理央 角薫
公立ほこだて未来大学

1. はじめに

近年、説得技術 (Captology) が注目され、コンピュータによる行動変容の手法が研究されている。説得技術とは、コンピュータを用いて人間の行動や態度の変容を促す技術 (Fogg, 2003) [6]であり、言語的な対話、動作、感情表現など、多様な手法が検討されている。動物の仕草や行動を模倣したロボットの動作設計が注目されており、ペット型人工物を活用したインタラクションが研究されている (Breazeal, 2003) [2]。

家庭では、犬が飼い主に対して餌を要求する際には餌皿の前で「クーン」と悲しそうに鳴き、遊びを求める際にはおもちゃを咥えて尻尾を振るなど、特定の行動を通じて意図を伝える (Miklósi et al., 2000) [13]。このようなペットと飼い主のコミュニケーションは、日常生活における自然な相互作用の一例である。一方で、ペット型ロボットやバーチャルペットを用いた説得技術の研究は限定的であり、動作や感情表現の組み合わせが説得力に与える影響については十分に解明されていない。

これまでの研究では、ペット型人工物の感情表現と動作が人間とのインタラクションに及ぼす影響が調査されてきた。Fujita (2000) は、ソニーの AIBO の開発を通じて、犬型ロボットがどのようにユーザと交流し、感情を持つ存在として認識されるかを示した。また、Lakatos et al. (2012) は、非人型ロボットに対する人間の感情帰属を分析し、ロボットの行動が社会的関係の形成に影響を与える可能性を示唆している。さらに、Okada (2012) は、「弱いロボット (Weak Robot)」の概念を提唱し、非完璧なロボットが人間とのインタラクションにおいてどのように共感を引き出すかを検討した[15]。このようなアプローチは、ペット型人工物がユーザに助けをもらう形での説得手法にも応用できる。

また、ロボットと人間の関係性の構築において、犬や猫などのペット型ロボットが特に社会的な影響を与える可能性が指摘されている (Tan et al., 2023) [20]。高齢者施設における観察研究では、ロボットペットが入居者の心理的な安定や社会的なつながりの促進に寄与することが示された。これらの知見は、ペット型人工物が単なる娯楽や対話の道具ではなく、人間の行動や態度を変容させるポテンシャルを

持つことを示している。また、Wada & Shibata (2007) や Shibata & Wada (2011) の研究でも、セラピーロボットが高齢者の情緒的な安定に寄与することが示されており、ペット型ロボットの持つ社会的な効果が強調されている。

本研究では、ペット型人工物の感情表現と動作を活用し、習慣改善を目的とした説得技術の可能性を検討する。特に、何かを促すという動作は、四足歩行の形状をしたさまざまな人工物に適用可能であると考えられる。従来のペット型ロボットを用いた実験では、騒音や移動速度の遅さが説得には適さないという課題が明らかになった。そこで、本研究では Mixed Reality (MR) 環境において、バーチャルペットを用いた実験を行い、仮想ペットがどのようにユーザの行動変容を促すかを調査する。

Harada & Sumi (2024) は、MR 環境において四足歩行の人工物が感情と動作を用いてゴミ捨てを促す実験を行い、「悲しみ」の感情表現が行動促進に有効である一方、「喜び」の感情は遊びの誘いと誤解され、説得には適していなかったことを示した[8]。また、「怒り」の感情は攻撃的と受け取られ、行動促進には効果的ではなかった。本研究では、これを拡張し、ペット型人工物の家庭内での利用を想定し、感情表現と動作を用いた習慣改善のための説得技術を探求する。

2. 関連研究

本研究に関連する研究として、以下の 3 つの視点から先行研究を整理する。

2.1 ペット型ロボットの人間とのインタラクション

ペット型ロボットの研究は、ユーザの社会的交流を促進し、感情的なつながりを生む可能性を示している。例えば、Fujita (2000) の AIBO に関する研究では、犬型ロボットがユーザとの相互作用を通じて感情的な絆を形成する過程が分析された。また、Koay et al. (2013) は、犬のような行動をするロボットの振る舞いがユーザにどのような印象を与えるかを調査し、動物らしい挙動が親和性を高めることを示した。加えて、Inoue, Wada, & Uehara (2021) は、ロボット介在療法 (Robot-Assisted Therapy, RAT) と動物介在療法 (Animal-Assisted Therapy, AAT) を比較し、高齢者の認知機能や情緒面での効果を分析した[9]。ロボットによる対話支援が、動物との交流に近い心理的效果をもたらす

ことが示唆されている。

2.2 バーチャルペットおよびエージェントの影響

バーチャルペットやエージェントの影響に関する研究としては、Koay et al. (2013) が犬型ロボットの行動がユーザの印象に与える影響を分析し、動物らしい挙動が親和性を高めることを示した[11]。また、Lakatos et al. (2012) は、非人型ロボットに対する人間の感情帰属を分析し、ロボットの行動が社会的関係の形成に影響を与える可能性を示した[12]。

さらに、バーチャルペットやエージェントの影響については、D' Mello & Graesser (2012) がエージェントの感情表現が学習者の動機付けに影響を与えることを示し[5]、Wang et al. (2020) は、バーチャルエージェントが生徒のモチベーションやエンゲージメントを向上させる可能性を示している[22]。このことは、感情表現がエージェントとのインタラクションに重要な要素であることを示唆しており、バーチャルペットのデザインにも示唆を与える可能性がある。

一方で、犬や猫などの動物がどのように人間とコミュニケーションを取るかについても、多くの研究が行われている。Miklósi & Topál (2013) は、犬の進化的適応により、飼い主との社会的関係を築く能力が高まっていることを示した[14]。また、Kerepesi et al. (2006) は、犬の視点取得能力 (Perspective Taking) を検証し、犬が他者の視点を理解する可能性があることを示した[10]。この能力は、ロボットが人間の行動を適切に認識し、より自然な対話を生み出すためのヒントとなる。

2.3 感情表現と行動変容

人間の感情表現と行動変容の関係についても、多くの研究が行われている。Batson (1991) や Davis (1994) は、共感のメカニズムとその社会的影響について詳細に論じており、他者の感情を理解することで行動が促進されることを示している[1][4]。また、Sinaceur & Tiedens (2006) は、怒りの感情が交渉の場面で影響を及ぼす可能性があることを指摘しており、感情が他者の意思決定にどのように影響を与えるかを示唆している[18]。

本研究は、以下の3点を目的とする。

1. Mixed Reality (MR) 環境におけるペット型人工物の感情表現と動作が習慣改善に与える影響を明らかにする。
2. 感情表現と動作の適切な組み合わせを検討する。
3. Mixed Reality (MR) 環境と動画環境における説得効果の違いを明確にする。

本研究の成果は、Mixed Reality (MR) 環境におけるペット型人工物のインタラクションデザインにおいて、より効果的な習慣改善のための設計指針を示すものである。

3. 習慣改善のための説得方法

3.1 動作の分類

動作を用いた説得技術に関する先行研究として、聴導ロボットの研究 (Koay et al., 2013) では、聴導犬の動作を参考にロボットの動作を設計し、行動促進の効果を評価した[4]。この研究では、ユーザの注意を引く、目標に向かって移動する、被験者と目標を交互に見つめるという流れが有効であることが示されている。

また、Harada & Sumi (2024) の研究では、聴導ロボットの動作パターンを応用し、犬の動作を「呼びかけ」「誘導」「ポインティング」の3つの意味に分類する手法が提案された[6]。この分類に基づき、本研究では以下の動作を採用し、ペット型人工物の説得動作を設計する。

1. **呼びかけ (Attention Calling)** : ユーザの注意を引くための動作。例として、近づく、鳴く、視線を合わせる、しっぽを振るといった行動が含まれる。
2. **誘導 (Guiding)** : ユーザを特定の場所へ案内するための動作。例として、対象物の近くへ移動し、進む方向を示唆する行動が含まれる。
3. **ポインティング (Pointing)** : 特定の対象物に注目を促すための動作。例として、視線を向ける、鼻先を対象物に近づける、対象物の周囲を回る動作が含まれる。

本研究では、これら3種類の動作を組み合わせることで、説得動作の設計を行う。

3.2 犬モデルの感情表現

動作と並んで、感情表現も説得技術において重要な要素となる。先行研究では、ロボットが感情を適切に表現することで、ユーザとのインタラクションが向上することが示されている (Picard, 1997) [16]。特に、Lakatos et al. (2012) の研究では、感情表現を伴う非人型ロボットが人間との社会的関係を形成する上で有効であることが示されている[12]。

本研究では、Harada & Sumi (2024) が提案した犬型人工物の感情表現の分類を参考にし、以下の4つの感情表現 (図1) を用いる[8]。

1. **悲しみ (Sadness)** : 飼い主に助けを求める動作として機能する可能性がある。
2. **怒り (Anger)** : 強い意思表示を伴うが、攻撃的と受け取られる可能性がある。
3. **喜び (Happiness)** : 親和的な関係を築くが、説得には適さない可能性がある。
4. **通常 (Neutral)** : 感情を伴わない状態。

この感情表現は、ウサギ型ロボットである Miro を用いて行われた感情表現についての研究 (Inoue, Wada, & Uehara, 2021) をもとに作成されている[9]。Miro の研究では、動物のような感情表現が人間とのインタラクションを向上させることが示されており、本研究における犬型人工物の感情表現の設計にも応用されている。

これらの感情を表現するため、犬の首、頭、耳、尻尾、鳴き声、動きの速さのパターンを以下のように設定する(表 1-1-表 1-3)。

表 1-1 悲しみの感情表現

部位	動作
首	下に傾ける
頭	下に傾ける
耳	外側斜めに傾ける
尻尾	下に傾ける
鳴き声	高い声でクーンと発音する
動きの速さ	ゆっくり動く

表 1-2 怒りの感情表現

部位	動作
首	下に傾ける
頭	上に傾ける
耳	後ろ側と外側に傾ける
尻尾	上に傾ける
鳴き声	低い声でウー、ワンと発音する
動きの速さ	ややゆっくり動く

表 1-3 喜びの感情表現

部位	動作
首	少し下に傾ける
頭	上に傾ける
耳	後ろ側と外側に傾ける
尻尾	左右に大きく早く振る
鳴き声	高い声でアンと発音する
動きの速さ	早く動く



図 1 感情表現

3.3 行動を促すための動作の組合せ

本研究では、感情表現を組み合わせた説得表により、ユーザの行動変容を促すための適切な組み合わせを探る。具体的には、以下のような組み合わせが考えられる (Harada & Sumi, 2024) [14]。

1. 悲しみ + ポインティング (例:「ゴミを捨ててほしい」
→ 悲しそうにしながらゴミ箱の方向を指し示す)
2. 怒り + 誘導 (例:「スマートフォンを使うのをやめてほしい」
→ 怒った様子でユーザを視線で誘導する)
3. 喜び + 呼びかけ (例:「本を読んでほしい」
→ 楽しそうに本の前で待機する)

これらの組み合わせが、習慣改善を目的とした説得技術において有効であるかを本研究で検証する。

4. バーチャルペットの開発

本研究では、Mixed Reality (MR)環境において、バーチャルペットがユーザの行動変容を促すためのシステムを開発した。バーチャルペットは、感情表現と動作を組み合わせた説得手法を活用し、ユーザに対して行動を促す役割を担う。本章では、開発したバーチャルペットの機能とシステム構成について述べる。

4.1 促す行動の分類と動作設計

本研究では、ペット型人工物がユーザに対して行動を促す際の行動を以下の4つのカテゴリに分類した。

1. ペット関係行動 (例: 餌やり, おもちゃの片付け)
2. 非ペット関係行動 (例: 読書, 本の片付け, ゴミ捨て)
3. 制限したい行動 (例: スマートフォンの利用防止)
4. 緊急時の警告 (例: 部屋の外への移動)

これらの行動を促すために、バーチャルペットの動作を「呼びかけ」「誘導」「ポインティング」の3種類に分類し、それぞれの行動に適した組み合わせを設計した。

- **呼びかけ (Attention Calling)**: ユーザの注意を引く動作 (例: 近づく, 鳴く, 視線を向ける)
- **誘導 (Guiding)**: ユーザを目的の場所へ案内する動作 (例: 対象物の近くへ移動)
- **ポインティング (Pointing)**: 特定の対象物を示す動作 (例: 視線を向ける, 頭を傾ける, 対象物と目的地を交互に見る)

本研究では、これらの動作を適切に組み合わせることで自然に行動を促すことを目指した。具体的には、感情表現と動作の組み合わせを考慮し、行動の種類に応じた適切な説得手法を設計した(表 2)。これにより、バーチャルペットが状況に応じた適切な方法でユーザの行動を促せるようにした。

表 2 促す行動ごとの動作・感情パターン

行動	感情	呼びかけ	誘導	ポインティング
読書	喜び	近づき, 鳴く	本棚の前へ移動	本 → ユーザを交互に見る
ゴミ捨て	悲しみ	近づき, 鳴く	ゴミの位置へ移動	ゴミ → ゴミ箱を交互に見る
本の片付け	怒り	近づき, 鳴く	落ちている本の位置へ移動	本 → 本棚を交互に見る
おもちゃの片付け	怒り	近づき, 鳴く	おもちゃの位置へ移動	おもちゃ → 片付ける箱を交互に見る
ペットへの餌やり	悲しみ	近づき, 鳴く	餌皿の位置へ移動	餌皿 → 飼い主を交互に見る
スマートフォンの使用防止	怒り	近づき, 鳴く	ユーザの前で待機する	ユーザを見る, 5回鳴く
部屋の外への移動	怒り	近づき, 1回鳴く	ドアの前へ移動	部屋の外でユーザを見る

4.2 システム構成

本研究では、HoloLens 2 を用いた MR 環境でバーチャルペットを動作させ、ユーザとのインタラクションを実現するシステムを開発した。Unity を用いて仮想ペットを実装し、リアルタイムで環境内を移動し、説得動作を行う設計となっている。

バーチャルペットは、以下の機能を備えている。

- ナビゲーション機能: ユーザの動きに応じて適切に移動
- 感情表現機能: 喜び, 悲しみ, 怒りの感情を視覚・音声で表現
- インタラクション機能: ユーザの行動に対して適切な応答を示す

- 遠隔操作機能: 実験環境では、バーチャルペットの行動を実験者が遠隔で制御可能

バーチャルペットの基本的な動作として、ユーザの周囲を移動したり、対象物の前で待機したりすることが可能である。また、待機時には特定の感情を伴うジェスチャーを行い、ユーザの注意を引く。

4.3 システムの主要機能

本研究で開発したバーチャルペットは、ユーザの行動を促すための複数の機能を備えている。HoloLens 2 を用いた MR 環境で動作し、ユーザの位置や行動に適応しながら、適切な感情表現と動作を用いた説得を行う。

まず、ナビゲーション機能では、バーチャルペットがユーザの移動に追従し、適切な距離を維持する。ユーザの行動が期待される場面では、特定の位置で待機し、説得の動作を実行する。待機中にユーザがバーチャルペットを撫でると、ペットが喜びの反応を示す機能も備えており、これは HoloLens 2 のハンドトラッキング機能を活用している。

次に、QR コードを用いた位置調整機能により、仮想オブジェクトの配置を実世界の環境に適合させる。実験室内に設置された QR コードを HoloLens 2 のカメラで認識し、バーチャルペットの位置を正確に調整する。この機能により、実験参加者が直感的にバーチャルペットを実空間内のオブジェクトと関連付けることができる。

さらに、遠隔操作機能を備え、実験者がバーチャルペットの動作をリアルタイムで制御できる。操作用のインターフェースは UDP 通信を利用し、実験者はバーチャルペットの感情表現や説得動作の開始・停止、特定のアクション(呼びかけ、誘導、ポインティング)を任意に実行可能である。この機能により、被験者の反応に応じた適応的なインタラクションが実現される。

本研究では、このバーチャルペットを用いて、習慣改善を目的とした説得技術の有効性を検証する。バーチャルペットがユーザに対してどのように行動を促し、どの感情表現が効果的に作用するかを評価するため、実験においてその機能を活用する。

4.4 実験におけるシステムの使用

本研究では、開発したバーチャルペットを用いて、習慣改善を目的とした説得技術の有効性を検証する実験を行う。実験では、被験者が HoloLens 2 を装着し、仮想ペットとインタラクションを行う環境を提供する。実験システムの役割は以下の通りである。

- 行動喚起の効果を評価するために、対象行動に応じた適切な動作と感情表現を提供
- ユーザの反応に基づいた適応的なインタラクションを行うため、ペットが異なる説得手法を試行
- 遠隔操作によりリアルタイムで調整可能なインタラクションを実現

このバーチャルペットの導入により、従来のロボットや固

定映像では困難だった、動的で個別適応的な説得手法の検証が可能となる。

5. 実験方法

本研究では、ペット型人工物の感情表現と動作を用いた説得技術の有効性を検証するため、2つの実験を実施した。実験1では、HoloLens 2 を用いた MR 環境において「提案手法群」と「吠えるのみ群」を比較し、どのような説得手法が行動促進に寄与するかを調査した。実験2では、同じ説得手法を用いた動画実験と比較し、MR 環境特有の影響を検討した。

5.1 実験1: 提案手法群と吠えるのみ群の比較

MR 環境において、感情表現と動作を組み合わせた説得(提案手法群)が、単に吠えるだけの説得(吠えるのみ群)よりも行動促進に効果的かどうかを検証する。

5.1.2 被験者

本実験には、18歳から33歳の大学生20名(男性14名、女性6名)が参加した。被験者は無作為に2群に分けられた。

- 提案手法群 (10名): バーチャルペットが感情を表現しながら動作(近づく、じっと見る、繰り返すなど)を行い、説得を試みる。
- 吠えるのみ群 (10名): バーチャルペットは対象物を見つめた後、吠える動作のみを行う。

5.1.3 実験環境

本実験は、HoloLens 2 を装着した被験者が、MR 環境内のバーチャルペットとインタラクションを行う形式で実施された。実験室には、QR コードによって位置調整された仮想オブジェクト(本、ゴミ箱、おもちゃ箱など)が配置され、バーチャルペットがこれらの対象に対して行動を促すように設計された。

5.1.4 促す行動

本実験では、以下の7つの行動を対象とした。

- ペット関連行動: 餌やり、おもちゃ片付け
- 非ペット関連行動: 読書、本片付け、ゴミ捨て
- 制限したい行動: スマートフォン利用防止
- 緊急時の警告: 部屋の外への移動

5.1.5 実験手順

1. 導入: 被験者にバーチャルペットの目的を簡単に説明し、HoloLens 2 を装着してもらう。
2. 自由行動(20分間): 被験者は実験室内で自由に行動し、バーチャルペットの説得を受ける。
3. 説得の試行: バーチャルペットがランダムなタイミングで、事前に設定した行動を促す。
4. アンケート記入: 実験終了後、被験者は行動の成功率、意図の伝わりやすさ、説得の不快感について評価するアンケートに回答。

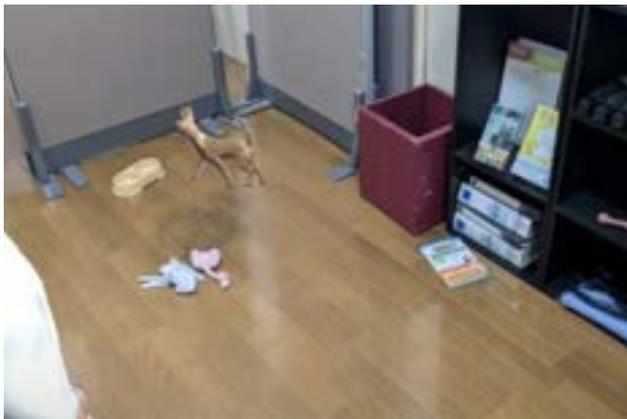


図 2 MR 環境と動画の実験

5.2 実験 2 : Mixed Reality (MR)環境と動画の比較

5.2.1 実験目的

本研究では、Mixed Reality (MR) 環境と動画環境におけるバーチャルペットの説得効果を比較する実験を行った

(図 2). MR 環境では、HoloLens を用いて実験室内にバーチャルペットを投影し、ユーザが直接インタラクションを行った。一方、動画環境では、同じ実験室の映像を用い、バーチャルペットの動作を画面越しに視聴する形式とした。両環境ともに、犬のバーチャルペットがおもちゃや本、餌皿などの対象物を用いて行動を促す様子を提示し、意図伝達や説得効果を比較した。

5.2.2 被験者

動画実験には、40 名の大学生が参加し、各行動に対して提案手法群と同様の説得を受けた後、同じアンケートに回答した。

5.2.3 実験環境

動画実験では、被験者が HoloLens 2 を用いず、PC 画面でバーチャルペットの説得動作を視聴し、その効果を評価する形式とした。

5.2.4 実験手順

1. 動画視聴：被験者は、実験 1 で使用したバーチャルペットの説得動作の動画を視聴する。
2. アンケート記入：実験 1 と同様の質問（成功率、意図の伝わりやすさ、不快感）に回答する。

5.3 分析方法

5.3.1 成功率の算出

成功率は、各行動目標に対して、バーチャルペットの説得を受けた後に行動を実行した被験者の割合として定義した。動画環境では、被験者がバーチャルペットの指示を含む動画を視聴した後、アンケートにて「この状況で実際に行動を起こすと思うか？」と質問し、「はい」と回答した割合を成功率として計算した。

5.3.2 統計解析

本研究では、以下の統計手法を用いて群間比較を行い、説

得手法の有効性を評価する。

- カイ二乗検定 (χ^2 検定) : 2 群間 (提案手法群 vs. 吠えるのみ群, HoloLens vs. 動画) の成功率の差を検定する。
- 効果量 (Cohen's d) : 実験 1 と実験 2 の成功率の差が実験的に意味のある大きさを持つかを評価する。
- 自由記述の質的分析: アンケート内の自由記述データを分類し、バーチャルペットの説得に対する主観的な評価を分析する。

6. 実験結果

6.1 実験 1 : 提案手法群と吠えるのみ群の比較

6.1.2 説得意図の伝達率と成功率

「バーチャルペットが行動を促していると感じたか？」という問いに対する肯定的な回答の割合を表 3 に示す。提案手法群の方が意図が伝わった割合が高く、特に「読書」「おもちゃの片付け」では 20%以上の差があり、「餌やり」では 30%の差が確認された。

「バーチャルペットの説得によって実際に行動をとったか？」という問いに対する肯定的な回答の割合を表 4 に示す。提案手法群の方が「本の片付け」「おもちゃの片付け」「スマートフォンの使用防止」の項目では吠えるのみ群よりも高い成功率を示した。特に「おもちゃの片付け」では、吠えるのみ群では行動が全く見られなかったのに対し、提案手法群では 30%の被験者が行動を実施した。また、「読書」「ゴミ捨て」「部屋の外への移動」では、吠えるのみ群の方が成功率が高かった。カイ二乗検定の結果、すべての行動において有意な群間差は確認されなかった ($p > 0.05$)。

表 3 説得意図の伝達率の比較

行動	提案手法群 (%)	吠えるのみ群 (%)
読書	70.0	50.0
ゴミ捨て	60.0	70.0
本の片付け	30.0	40.0
おもちゃの片付け	30.0	10.0
ペットへの餌やり	60.0	30.0
スマートフォンの使用防止	30.0	20.0
部屋の外への移動	70.0	90.0

表 4 説得意図の成功率の比較

行動	提案手法群 (%)	吠えるのみ群 (%)
読書	60.0	70.0
ゴミ捨て	40.0	50.0
本の片付け	30.0	20.0
おもちゃの片付け	30.0	0.0
ペットへの餌やり	30.0	30.0
スマートフォンの使用防止	40.0	20.0
部屋の外への移動	60.0	70.0

6.1.3 説得に対する不快感

「バーチャルペットの説得に対して不快感を感じたか？」

という質問に対し7段階評価で集計した結果を表5に示す。

表5 説得に対する不快感 (平均値)

行動	提案手法群	吠えるのみ群
読書	1.5	1.2
ゴミ捨て	2.8	1.5
本の片付け	2.6	1.3
おもちゃの片付け	1.0	1.0
ペットへの餌やり	1.2	1.0
スマートフォンの使用防止	1.5	1.5
部屋の外への移動	1.6	1.4

全体的に、バーチャルペットの説得による不快感は低いことがわかった。ただし、「ゴミ捨て」や「本の片付け」では、提案手法群の方がやや不快感の評価が高いことが示された。

6.2 実験2: Mixed Reality (MR)環境と動画の比較

6.2.1 成功率の比較

MR環境と動画の成功率を比較した結果を表6に示す。MR環境よりも動画の方が全体的に高い成功率を示した。特に「読書」「餌やり」「本の片付け」では動画の方が効果的であった。また、スマートフォン利用防止と部屋の外への移動については動画実験では実施されていないため、比較対象外である。

表6 MR環境と動画の成功率の比較

行動	MR環境 (%)	動画 (%)
読書	60.0	80.0
ゴミ捨て	40.0	50.0
本の片付け	30.0	50.0
おもちゃの片付け	30.0	40.0
ペットへの餌やり	30.0	70.0
スマートフォンの使用防止	40.0	—
部屋の外への移動	60.0	—

6.2.2 自由記述の分析

自由記述の分析から、以下のような特徴的な意見が得られた。

1. MR環境では、バーチャルペットが「対話できる存在」であることを期待される

「犬がもっと反応してくれたら良かった」「話しかけたのに反応しないのが違和感」といった意見が多数あった。

2. 動画では、受動的に視聴するため、感情表現がそのまま伝わりやすい

「動画では犬の表情がよくわかった」「喜んで様子が理解しやすかった」という意見が多かった。

3. MR環境では、行動のトリガーが分かりにくい

「犬が促しているつもりなのか分からなかった」「何をすればいいのか迷った」という意見が一部見られた。

6.2.3 MR環境でのインタラクティブ性の影響

MR環境では、ユーザがバーチャルペットとのインタラクティブなやり取りを期待する傾向があることが示唆された。特に「犬に話しかけたが無反応だった」など、受動的に動画を見る場合とは異なる期待感があるため、これが説得の成功率に影響を与えている可能性がある。

7. 考察

本研究では、バーチャルペットを用いた説得技術がユーザの行動変容に与える影響を検討し、MR環境と動画環境における違いを分析した。その結果、MR環境ではバーチャルペットとのインタラクティブな関係が期待される傾向があり、単なる映像として提示される動画とは異なる設計上の課題が明らかになった。本章では、被験者の行動、意図伝達、不快感、説得効果についての考察を行い、特にMR環境におけるバーチャルペットのインタラクティブ性の必要性と、それに伴うソフトウェアデザインの課題について詳しく議論する。

また、MR環境では、単なる視覚的な表現だけではなく、ユーザとの対話的なインタラクションが不可欠であることが示唆された。本研究の結果は、MR技術を用いたバーチャルエージェントの設計において、ユーザの期待と実際のシステム挙動のギャップを埋めるための設計指針を提供するものである。

7.1 実験中の被験者の行動について

実験中、被験者はバーチャルペットがいる方向に体を向けたり、こまめに視線を向けたりする行動を取っていた。このような行動は、実際のペットを見守る行動と類似しており、ユーザがバーチャルペットの動きを常に気にかけていたことを示唆している。特に、バーチャルペットが特定の行動を示した際に、それを注意深く観察する傾向があった。これは、バーチャルペットが「逃げそうになる」「ゴミを食べようとする」といった危険な行動をすることで、ユーザの行動を促す設計が有効である可能性を示している。

また、被験者はバーチャルペットに促された行動をする際、バーチャルペットの反応を確認しながら行動していた。このことから、ユーザの行動に応じてバーチャルペットが異なる反応を示すことで、より自然で対話的な説得が可能になると考えられる。さらに、ユーザが慎重に行動を選択していたことから、「対象に近づく」「手を伸ばす」「持ち上げる」「移動する」など、行動のプロセスを細かく分けて段階的に説得するアプローチが有効である可能性がある。

また、被験者はバーチャルペットに促された行動をする際、バーチャルペットの反応を確認しながら行動していた。このことから、ユーザの行動に応じてバーチャルペットが異なる反応を示すことで、より自然で対話的な説得が可能になると考えられる。さらに、ユーザが慎重に行動を選択していたことから、「対象に近づく」「手を伸ばす」「持ち上げ

る」「移動する」など、行動を段階的に分けて説得することが有効であると考えられる。

実験中、被験者はバーチャルペットに対して高い関心を示し、視線を向ける頻度が多かった。これは、ユーザがバーチャルペットを単なる視覚的情報としてではなく、「インタラクションが可能な存在」として認識していた可能性を示している。特にMR環境では、「犬が反応してくれなかった」「もっと対話できる動作が欲しい」といった自由記述の意見が多く見られ、ユーザがバーチャルペットとの双方向的なやり取りを期待していることが明らかになった。

これは、MR環境では現実空間と仮想空間が融合するため、ユーザの期待値が動画環境とは異なることを示唆している。動画環境では受動的に視聴することが前提となるため、単なる動作や感情表現でも意図が伝わりやすい。しかし、MR環境ではバーチャルペットが「そこにいる」というリアルな存在感を持つため、ユーザは「話しかければ反応する」「行動に応じて反応が変わる」といった高度なインタラクションを求める傾向が強いと考えられる。

このため、バーチャルペットがユーザの行動に応じたフィードバックを提供することで、より自然なインタラクションを実現する必要がある。例えば、ユーザが読書を始めたらバーチャルペットが満足そうに振る舞う、ゴミを拾ったら嬉しそうに吠えるといった設計が、より強い説得効果につながる可能性がある。

7.2 意図伝達について

実験の結果、特定の行動ではバーチャルペットの説得意図が比較的伝わりやすい傾向が見られた。例えば、読書やゴミ捨てといった行動では、対象物と行動のつながりが明確であるため、単純な動作(例:吠える)でも意図が十分に伝わる可能性がある。

一方で、本の片付けやおもちゃの片付けのように、行動に明確な「終点」が必要な場合、バーチャルペットの説得表現が不十分である可能性が示唆された。これらの行動では、単に対象物を指し示すだけでは不十分であり、片付けるべき場所を示す動作(例:本棚を指す、おもちゃ箱の前に移動する)を追加することで、意図の伝達がより効果的になると考えられる。

7.3 説得に対する不快感について

説得に対する不快感の評価では、ほとんどの被験者がバーチャルペットによる説得を不快には感じていなかった。特に、「ペットと関係のある行動」(餌やり・おもちゃの片付け)では、不快感が低く、「ペットを助けてあげたい」という気持ちが説得にポジティブな影響を与えたと考えられる。これは、バーチャルペットがユーザに寄り添う存在として機能していることを示しており、日常生活における行動促進の観点から有用な特性である。

一方、「ゴミ捨て」や「本の片付け」では不快感がやや高く、ペットと行動の関連性が低い場合、ユーザが違和感を感じ

る可能性があることが分かった。また、怒りの表現については、「叱られているようで嫌だった」との意見が見られたことから、過度な強調は避け、適切な強度で使用する必要があると考えられる。

7.4 説得効果について

行動促進の成功率では、MR環境よりも動画環境のほうが全体的に高い結果が得られた。特に視覚的な情報が直接伝わりやすい行動(例:読書や餌やり)では、動画のほうが効果的だった。これは、動画では感情表現がストレートに伝わりやすいことが要因となった可能性がある。

本研究の分類(ペット関係行動、非ペット関係行動、制限したい行動、緊急時の警告)に基づく説得効果に明確な違いは見られず、行動の種類ごとに異なる要因が影響している可能性がある。そのため、バーチャルペットの説得効果を向上させるには、行動の性質(実行のしやすさや対象物との関連性)、片付けなどの完了基準の明示、感情表現の適切な調整といった要素を考慮することが重要である。

7.5 MR環境におけるインタラクティブ性の課題

本研究の結果から、MR環境においてバーチャルペットがユーザとの対話的なやり取りを求められることが分かった。これは、現実空間と仮想空間が重なり合うMR環境では、バーチャルペットに「能動的な存在」としての振る舞いが期待されるためである。

この点は、従来の動画や単なるARエージェントとは異なり、バーチャルペットがどのように「反応」し、ユーザの行動に適應するかが設計の重要なポイントになることを意味している。例えば、「犬に話しかけても反応しない」といった意見が多かったことから、音声入力やジェスチャーを認識し、それに応じた動作を返す機能を追加することで、より自然なインタラクションが可能になると考えられる。

また、MR環境では物理的な空間との整合性も考慮する必要があり、現実空間のオブジェクトと連動したバーチャルペットの振る舞いが重要となる。たとえば、ユーザの手の動きを検出し、それにに応じてバーチャルペットが行動を変えるような設計が、より強い説得効果につながる可能性がある。

8. まとめ

本研究では、バーチャルペットを用いた説得技術の効果をMixed Reality (MR)環境と動画環境と比較し、ユーザの行動変容やインタラクションの違いを分析した。結果として、MR環境ではバーチャルペットとのインタラクティブな関係が期待される傾向が強く、単なる映像として提示される動画環境とは異なる設計の課題が明らかになった。

行動促進の成功率に関しては、MR環境と動画環境の間に有意差は確認されなかったものの、動画のほうが特定の行動(読書や餌やり)では効果的であった。一方で、MR環境ではユーザがバーチャルペットとの対話を期待する傾向があ

り、「反応がほしい」「対話的な動作が必要」といった意見が多く見られた。

これらの結果から、MR 環境におけるバーチャルペットの説得技術には、単なる感情表現や動作の提示だけでなく、ユーザとの双方向的なインタラクションを強化することが求められることが示唆された。特に、ユーザの行動に応じたリアルタイムな反応を取り入れることで、より自然で効果的な説得が可能になると考えられる。

参考文献

- [1] Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Lawrence Erlbaum Associates.
- [2] Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1-2), 119-155. [https://doi.org/10.1016/S1071-5819\(03\)00047-8](https://doi.org/10.1016/S1071-5819(03)00047-8)
- [3] Carballo, F., Cavalli, C. M., Martínez, M., & Bentosela, M. (2020). Asking for help: Do dogs take into account prior experiences with people? *Learning & Behavior*, 48(4), 411-419. <https://doi.org/10.3758/s13420-020-00437-2>
- [4] Davis, M. H. (1994). *Empathy: A social psychological approach*. Westview Press.
- [5] D'Mello, S., & Graesser, A. (2012). AutoTutor and affective computing: Enhancing learning through intelligent tutoring systems. *Educational Psychologist*, 47(3), 181-190. <https://doi.org/10.1080/00461520.2012.680692>
- [6] Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann.
- [7] Fujita, M. (2000). AIBO: Toward the era of digital creatures. *The International Journal of Robotics Research*, 20(10), 781-794. <https://doi.org/10.1177/02783640122068092>
- [8] Harada, R., & Sumi, K. (2024). Persuasive technology through behavior and emotion with pet-type artifacts. In Baghaei, N., Ali, R., Win, K., Oyibo, K. (Eds.), *Persuasive Technology. PERSUASIVE 2024. Lecture Notes in Computer Science*, Vol. 14636. Springer, Cham.
- [9] Inoue, M., Wada, K., & Uehara, R. (2021). Comparison of robot-assisted therapy and animal-assisted therapy for elderly individuals with dementia. *International Journal of Social Robotics*, 13(2), 221-232. <https://doi.org/10.1007/s12369-020-00654-3>
- [10] Kerepesi, A., Doka, A., Miklósi, Á., & Csányi, V. (2006). Do animals have a "theory of mind"? A nonverbal test of visual perspective taking in dogs. *Behavioural Processes*, 73(3), 224-230. <https://doi.org/10.1016/j.beproc.2006.06.005>
- [11] Koay, K. L., Walters, M. L., Dautenhahn, K., & Te Boekhorst, R. (2013). Social robots for the home: A study on dog-like robot behavior and people's responses. *International Journal of Social Robotics*, 5(4), 593-613. <https://doi.org/10.1007/s12369-013-0206-8>
- [12] Lakatos, G., Gácsi, M., Konok, V., Brúder, I., Bereczky, B., Korcsok, B., & Miklósi, Á. (2012). Emotion attribution to a non-humanoid robot in different social situations. *PLoS One*, 7(9), e49950. <https://doi.org/10.1371/journal.pone.0049950>
- [13] Miklósi, Á., Polgárdi, R., Topál, J., & Csányi, V. (2000). Intentional behavior in dog-human communication: An experimental analysis of "showing" behavior in the dog. *Animal Cognition*, 3(3), 159-166. <https://doi.org/10.1007/s100710000072>
- [14] Miklósi, Á., & Topál, J. (2013). What does it take to become "best friends"? Evolutionary changes in canine social competence. *Trends in Cognitive Sciences*, 17(6), 287-294. <https://doi.org/10.1016/j.tics.2013.04.005>
- [15] Okada, M. (2012). Robot as a weak agent: Human-like communication through incongruency resolution. *AI & Society*, 27(2), 173-181. <https://doi.org/10.1007/s00146-011-0361-5>
- [16] Picard, R. W. (1997). *Affective computing*. MIT Press.
- [17] Shibata, T., & Wada, K. (2011). Robot therapy: A new approach for mental healthcare of the elderly – A mini-review. *Gerontology*, 57(4), 378-386. <https://doi.org/10.1159/000319015>
- [18] Sinaceur, M., & Tiedens, L. Z. (2006). Get mad and get more than even: When and why anger expression is effective in negotiations. *Journal of Experimental Social Psychology*, 42(3), 314-322. <https://doi.org/10.1016/j.jesp.2005.05.002>
- [19] Tamura, T., Yonemitsu, S., Itoh, A., Oikawa, D., Kawakami, A., Higashi, Y., Fujimo, H., & Nakajima, K. (2004). Is an entertainment robot useful in the care of elderly people with severe dementia? *The Journals of Gerontology: Series A*, 59(1), M83-M85. <https://doi.org/10.1093/gerona/59.1.M83>
- [20] Tan, W. J., Asoh, H., & Ueno, K. (2023). The social impact of robotic pets in elderly care facilities: An observational study. *Journal of Robotics and Autonomous Systems*, 156, 104258. <https://doi.org/10.1016/j.robot.2022.104258>
- [21] Wada, K., & Shibata, T. (2007). Living with seal robots – Its sociopsychological and physiological influences on the elderly at a care house. *IEEE Transactions on Robotics*, 23(5), 972-980. <https://doi.org/10.1109/TRO.2007.906261>
- [22] Wang, L., Wu, T., Zhang, J., & Shen, W. (2020). Effects of virtual agents on students' learning motivation and engagement in online learning environments. *Computers & Education*, 157, 103994. <https://doi.org/10.1016/j.compedu.2020.103994>

Thing2Vec: 人の行動理解に向けたモノの使用特性のベクトル埋め込み方法の提案

田仲百音^{†1} 岸野泰恵^{†2} 下西英之^{†3}
 大阪大学 日本電信電話株式会社, 大阪大学 大阪大学

1. はじめに

近年、センシング技術の向上により、サーマルカメラの利用が拡大している。それに伴い、これまでデータ化できなかった情報を、スマートハウスなどで活用できる可能性が高まってきている過去の研究では、人が触れた箇所をサーマルカメラを用いて検出することで、感染症対策に活用する取り組みが行われている [1]。この技術と物体検知を組み合わせることで、日常的に触れた物体の使用履歴を記録できるようになる。そこで我々は、接触した物体の履歴を蓄積することで、モノの使われ方を分析できるのではないかと考えた。

モノの使われ方を分析するための手法としては、モノの分散表現を作成するアプローチを提案する。分散表現は、非構造化データをベクトルで表現する手法であり、ベクトル埋め込みとも呼ばれる。埋め込まれたデータ点のうち、近しい性質を示すものは、多次元空間上において互いに近い場所に位置するという特徴がある。最近では、文章、商品などさまざまなものを対象としたベクトル埋め込みの手法が提案され、分析や推薦などの応用に活かす研究が行われている。そのため、モノの使われ方をベクトルで表現することによって、異常検知や生活パターンの特定、人の行動理解など幅広い応用が期待できる。そこで我々は、モノの使われ方がモノを特徴づける情報として重要であるという仮説のもと、モノの使用情報を用いてモノの分散表現を作成する Thing2Vec を提案する。Thing2Vec では、Word2Vec のモデルをもとに、入力として物体 ID を、出力にはモノの使用情報を与えて学習を行うことで、中間層の潜在変数を分散表現として得ることができる。

実験では、可視光カメラとサーマルカメラを使用してモノの使用情報を取得し、得られた情報をもとに分散表現を作成した。生成された分散表現をクラスタリングし分析を行った結果、生成された分散表現は、モノの使用情報に基づ

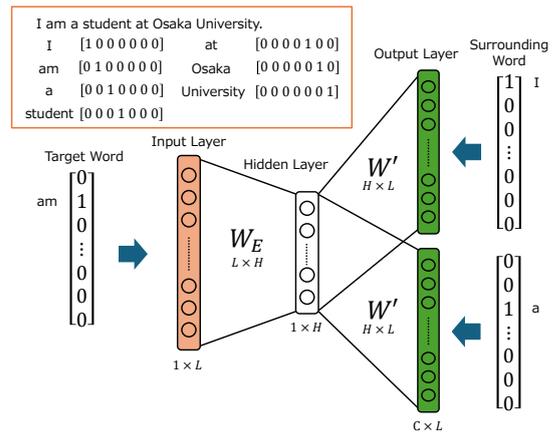


図1 Word2Vec のアーキテクチャ
 Figure 1 The Architecture of Word2Vec

いたクラスタを形成することが分かった。今後、より大規模なデータで学習することでモノの使われ方に関する分析や、人の行動理解に活用できる可能性がある。

2. 関連研究

これまでに様々な非構造化データを分散表現として扱う手法が提案されてきた [2, 3, 4, 5]。

Word2Vec [6, 7] は、単語をベクトル埋め込みする手法の一つである。文章内の注目単語について、その周囲の単語を予測するタスクを解くことで単語の分散表現を作成する。Word2Vec には、bag-of-words と Skip-gram の 2 つのモデルがあり、Skip-gram の方が高精度である。Skip-gram モデル (図 1) は入力層と 1 つの隠れ層、出力層からなる。ここで、図中の L はベクトルに埋め込む単語数、 H は分散表現の次元数、 C は出力層に与えられる周辺単語数、すなわちウィンドウサイズである。簡単のため、図中の出力層に与えられている周辺語は前後 1 単語であるが、実際の周辺語数 C は任意に設定することができる。Word2Vec では、以下の 3 つのステップによって学習が行われる。

1. 入力層には単語の対応する位置にフラグが設定された one-hot ベクトルを入力し、隠れ層で重み行列 W_E と

Thing2Vec: Proposal of Object Usage Pattern Embedding Toward Human Activity Understanding.

^{†1} MONE TANAKA, Osaka University

^{†2} YASUE KISHINO, Nippon Telegraph and Telephone Corporation, Osaka University

^{†3} HIDEYUKI SHIMONISHI, Osaka University

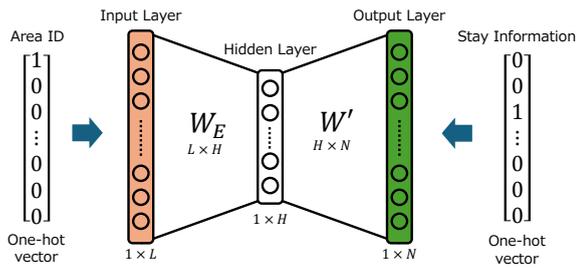


図2 Area2Vec のアーキテクチャ
Figure 2 The Architecture of Area2Vec

- の行列積に変換する。
2. 隠れ層の出力を、重み行列 W' との行列積に変換した後、出力層の softmax 関数によって各単語の確率分布として解釈する。
 3. 得られた確率分布と、出力層に与えられる周辺語の one-hot ベクトル間の交差エントロピー損失を計算し、重みを更新する。

ここで、入力層に与えられる入力は一-hot ベクトルであるため、入力層から隠れ層への重み行列の各行を、その単語の分散表現として扱うことができる。したがって、学習された重み行列から分散表現が得られる。実験の結果、大量のデータで学習して得られたベクトルは、単語間の関係を意味することが分かった。

Barkan らは、Word2Vec をレコメンデーションシステムに応用する手法として Item2Vec [8] を提案した。協調フィルタリングにおいて、アイテムをユーザが選択した集合として扱うが、ユーザとアイテムのセット関係に関する情報が提供されない場合がある。そこで Item2Vec では、アイテムベース協調フィルタリングのため、Word2Vec における単語をアイテムに、単語列をアイテム集合と置き換えて学習が行われる。Skip-gram を改良した Skip-gram with Negative Sampling (SGNS) を用いて学習することで、ユーザの情報が未開示であるシナリオに対し、アイテム間の関連性を推測することを可能にした。実験の結果、Item2Vec は SVD (特異値分解) モデルと比較して、高精度なジャンル分類を行えることがわかった。

Area2Vec [9] は注目エリアについて、曜日、到着時刻、滞在時間からなる滞在情報に基づいた分散表現を作成する。Skip-gram モデルをもとに考案された、Area2Vec のアーキテクチャを図2に示す。Area2Vec では、50メートル四方に切り取られた区画ごとにエリア ID が割り振られ、エリア ID に対応する one-hot ベクトルが入力層に入力される。隠れ層の出力は、Word2Vec と同様に各エリアの滞在情報の確率分布として解釈される。この構成で学習することで、出力層から得られる確率分布と、離散化された滞在情

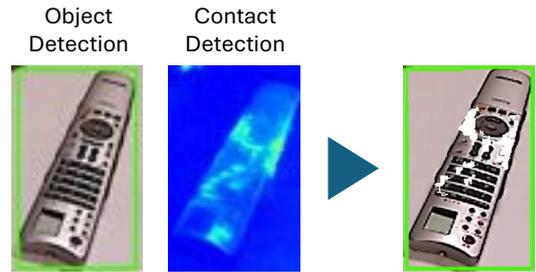


図3 物体検知・接触検知の例

Figure 3 The Example of Object and Contact Detection.

報の one-hot ベクトルとの交差エントロピー損失が計算され、重み行列が更新される。実験では、学習した重み行列から得られた分散表現を k-means 法でクラスタリングすることで、繁華街や住宅エリア、オフィスエリアなどを視覚的に区分し、人にとって意味のある分析結果が得られることが分かった。

3. Thing2Vec

近年、IoT 技術やユビキタスコンピューティングの発展により、センシング技術が向上している。個々のモノに関するセンシング技術も進歩しており、使用時刻など、これまで収集されていなかったデータを収集できるようになった。しかし、日常的に使用するモノに対し、それぞれのデータを組み合わせて分析を行う研究は少ない。一方で、実生活において、生活に密着したモノは多種多様である。そこで我々は、さまざまなモノの使われ方を同時に分析することで生活パターンの特定や人の行動理解に活用できるのではないかと考えた。

本研究では、モノの使用情報からモノの分散表現を作成する Thing2Vec を提案する。例えば、仕事のある平日に使うモノと、休日にプライベートで使うモノは異なる。通勤している人の場合、鍵や財布などは平日の朝夕の決まった時間に取りられることが多いだろう。このような情報から分散表現を作成することで、生活パターンの特定や異常検知への応用が期待できる。

我々の提案手法では、可視光カメラとサーマルカメラを用いてモノの使用情報を取得し、Thing2Vec によって使用情報に基づいたモノの分散表現を作成する。本章では、それぞれの手順について詳細に説明する。

3.1. 物体の検出

物体の検出においては、一般的な画像処理アルゴリズムを用いて可視光カメラから物体を検出する。また、人が触れて実際に使った物体であるかを、サーマルカメラによる

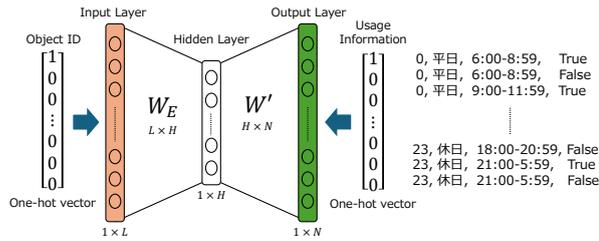


図4 Thing2Vec のアーキテクチャ
Figure 4 The Architecture of Thing2Vec

表1 モノの使用情報の定義

Table1 Definition of Object Usage Information

モノの使用情報	値
物体 ID	0-23
曜日	平日, 休日 (土曜日と日曜日)
出現時刻	6:00-8:59, 9:00-11:59, 12:00-14:59, 15:00-17:59, 18:00-20:59, 21:00-5:59
使用状態	True, False

熱痕跡を用いて判断する。人は恒温動物であり、通常、環境表面より体温が高いため、人がモノに触れると物体表面に熱が残る。熱痕跡と呼ばれるこの熱が、検出した物体上にあれば、人が触れたものとする (図 3)。人の日常物の操作を分析するためには、物体が単にカメラ画像中に撮影されたか、実際に触れたかの区別が重要であるため、サーマルカメラを併用する方法を取る。

検知された物体はログに記録した後、同一物体であれば一連の物体履歴として統合する。以上の処理により、物体履歴として、物体の出現時間、消失時間、接触の有無を記録していく。

3.2. モデル

提案する Thing2Vec は、Area2Vec モデルを参考に考案したものである。Thing2Vec は 3 層のネットワークから構成され、入力層には各物体の ID に対応した one-hot ベクトルを、出力層には、離散化されたモノの使用情報に対応する one-hot ベクトルを与える。入力層と隠れ層の間の重み行列は $L \times H$ の行列 W_E であり、隠れ層と出力層の間の重み行列は $H \times N$ の行列 W' である。ここで、 L はベクトルに埋め込むモノの数、 N は出力層のノード数 (すなわちモノの使用情報の組み合わせ総数)、 H は分散表現の次元数を示す。

3.3. 定義

我々は、モノの使用情報として「物体 ID」、「曜日」、「出現時刻」、検出された時間内に人が触れたかどうか、すなわちモノが使用されたかを意味する「使用状態」の 4 つを使用した。人がモノに触れて置いたのか、長時間置かれていた物の上にあったものが移動されて、そのモノが撮影されるようになったのかによって、人の物の使い方に対する認識は異なると考えられる。そこでモノが使用される曜日や時刻が、モノを特徴づける上で重要だと考え、使用状態を導入した。サーマルカメラを用いて使用状態を取得し、Thing2Vec に入力することで、より人の認識に近い分散表現になる可能性がある。我々は、表 1 に示すようにモノの使用情報を定

義し、それぞれを離散的な値に割り当てた。

3.4. 埋め込み

前節の定義によって離散化結果は、物体 ID、曜日、出現時刻、使用状態の全組合せである $24 \times 2 \times 6 \times 2 = 576$ のカテゴリに振り分けられる。入力層には物体 ID に対応した one-hot ベクトルが、出力層には、576 カテゴリに振り分けられたモノの使用情報に対応する one-hot ベクトルが与えられる。Thing2Vec モデルは、以下の 3 つのステップによって学習を行う。

1. 入力層には物体 ID の対応する位置にフラグが設定された one-hot ベクトルを入力し、隠れ層で重み行列 W_E との行列積に変換する。
2. 隠れ層の出力を、重み行列 W' との行列積に変換した後、出力層の softmax 関数によって各使用情報の確率分布として解釈する。
3. 得られた確率分布と、出力層に与えられる周辺語の one-hot ベクトル間の交差エントロピー損失を計算し、重みを更新する。

学習後、学習された重み行列から分散表現が得られる。

4. 実験

Thing2Vec がモノの特徴を捉えているかを確認するため、小規模な実験を行った。本章では、実験設定および実験の結果について述べる。

4.1. 物体検出の実装

実装上、物体の検出には、我々が過去に提案した物体検知・記録システム [10] を使用する。提案手法は、本来個別の物体を追跡することを想定しているが、個々の日常的な物体を認識して追跡するアルゴリズムを即座に実装することは困難であったため、既存の物体検知アルゴリズムである YOLO-World [11] を使用した。従来の YOLO では事前に定義された語彙の物体しか検知することができなかったが、YOLO-



図5 実験画像の例.

Figure 5 The example of experiment images.

World は柔軟な語彙を指定することができる. 今回の実験では, 日常的に使用するモノとして, YOLO-World に以下の 24 語彙を与えた. Person, Glasses, Bottle, Cup, Handbag, Book, Umbrella, Watch, Pen/Pencil, Cell Phone, Clock, Head Phone, Remote, Scissors, Folder, Earphone, Mask, Tissues, Wallet, Tablet, Key, Stapler, Eraser, Lipstick.

実験では, YOLO-World の出力する YOLO ラベルを「物体 ID」として扱った. 検知された物体はログに記録された後, バウンディングボックスの IoU と HSV 色空間の Hue 成分によって同一物体かどうか判定され, 最終的に一連の物体履歴へ統合される.

4.2. データセット

我々は著者の 1 名と 3 名の家族が住む自宅で日常生活の一部を撮影し, 10 日間で 24 シーンを撮影した. 撮影には, 解像度 480×640 の可視光カメラ (GAZO GZ300), 解像度 480×512 のサーマルカメラ (FLIR Boson+ AS50deg) を用いた. 撮影された画像の例を図 5 に示す. ここで, 前節で説明したような画像処理および物体履歴の生成を行った結果, 計 173 の物体履歴が得られた.

4.3. 結果

モデルは 200 エポックで学習され, 分散表現の次元は参考*1に示された以下の式に基づいて 5 に設定された.

$$embedding_dimensions = number_of_categories^{0.25} \quad (1)$$

カテゴリ数は $24 \times 2 \times 6 \times 2 = 576$ であるため, $576^{0.25} \approx 4.9$ より 5 次元とした. 得られた分散表現は, クラスタリングによって解釈可能な形に変換される. 本実験では, 代表的なクラスタリング手法である k-means と DBSCAN [12] を用いた. DBSCAN では, ある一定範囲内に 4 点以上集まっているものをクラスタと定義し, クラスタに含まれない点はノイズとした.

得られた分散表現を DBSCAN 法でクラスタリングした結果, 17 クラスタとノイズに分類された. 5 次元の分散表

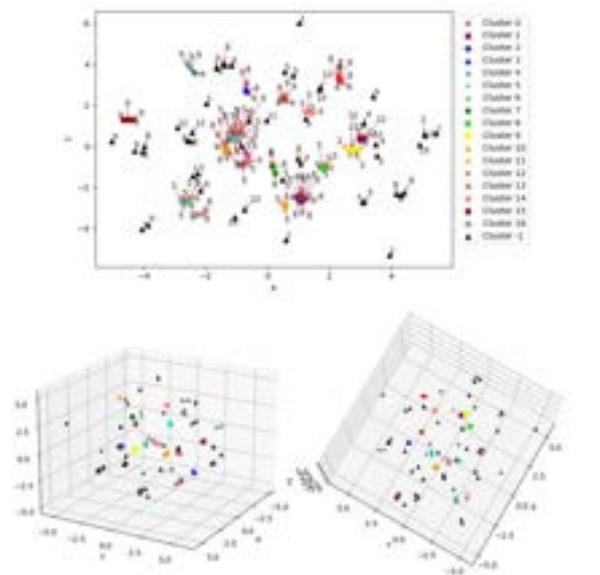


図6 DBSCAN クラスタリングを行った結果. 図中の数字は, 「物体 ID」すなわち, 検知された YOLO ラベルを示している.

Figure 6 2D and 3D visualization of DBSCAN clustering results. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

現を DBSCAN クラスタリングした結果について, 主成分分析を行って 2 次元と 3 次元のグラフに視覚化したものを図 6 に示す. ここで, 図中の数字は「物体 ID」すなわち, YOLO-World の出力した YOLO ラベルを示している. 以降では, 物体 ID とラベルを同義とする.

クラスタ 0, 5, 6, 8, 11, 14 はラベル 5: Book の集合である. これらのクラスタ間の違いは, 出現時刻, 曜日, 使用状態による. クラスタ 0 は平日の 18-21 時の間に使用された本, クラスタ 5 は平日の 9-12 時の間に使用された本, クラスタ 6 は平日の 9-12 時の間に出現した本, クラスタ 8 は平日の 15-18 時の間に使用された本, クラスタ 11 は休日の 15-18 時の間に使用された本, クラスタ 14 は休日の 15-18 時の間に出現した本を意味する. ここでは, クラスタ 5 とクラスタ 6 は冬休み中に朝から小学生が本を使って遊んでいるクラスタであり, 他は授業期間中に宿題を夕方に行うクラスタであった.

クラスタ 2, 3, 7, 10, 12, 16 はラベル 8: Pen/Pencil の集合である. クラスタ 2 は平日の 18-21 時の間に出現したペン, クラスタ 3 は平日の 18-21 時の間に使用されたペン, クラスタ 7 は平日の 9-12 時の間に出現したペン, クラスタ 10 は平日の 15-18 時の間に出現したペン, クラスタ 12 は休日の 15-18 時の間に出現したペン, クラスタ 16 は平日の 9-12 時の間に使用されたペンである. ここでは, 宿題をし

*1 <https://developers.googleblog.com/en/introducing-tensorflow-feature-columns/>

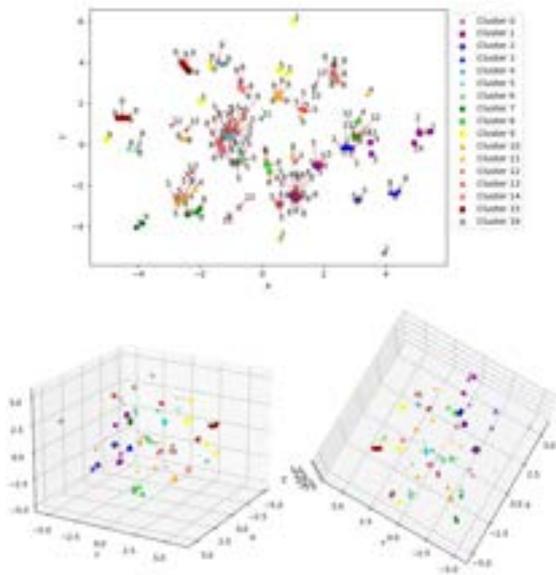


図7 k-means クラスタリングした結果. 図中の数字は, 「物体 ID」すなわち, 検知された YOLO ラベルを示している.

Figure 7 2D and 3D visualization of k-means clustering results. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

ている途中のペン (クラスタ 3) と机に置かれたままのペン (クラスタ 2) が別のクラスタに分類された.

クラスタ 4, 13, 15 はラベル 9 : Cell Phone の集合である. クラスタ 4 は平日の 18-21 時の間に使用された携帯電話, クラスタ 13 は休日の 15-18 時の間に出現した携帯電話, クラスタ 15 は平日の 18-21 時の間に出現した携帯電話である. カメラ撮影範囲内に携帯電話をよく充電する場所が含まれており, 平日においては 18 時以降の帰宅後に充電するが, 休日は 15 時から出現しており, 携帯電話の使われ方の違いがクラスタに表れていることがわかる.

クラスタ 1 はラベル 12 : Remote のクラスタ, クラスタ 9 はラベル 2 : Bottle のクラスタである. クラスタ-1 (ノイズ) には, 出現回数の少ないデータが含まれる.

図 7 に k-means クラスタリングの結果を示す. ここでは, DBSCAN クラスタリングと比較するために $k = 17$ とした. k-means は DBSCAN と比較して, 異なるラベルを同じクラスタに分類することが多く, DBSCAN の方が本研究の目的に沿ったクラスタリングが可能であることを確認した.

次に, 図 8, 9 に離散化したデータを分散表現に変換せず, 直接クラスタリングした結果を示す. つまり, 表 1 の定義に従って値を割り振り, 物体 ID, 曜日, 出現時刻, 使用状態に関する 4 次元配列をクラスタリングした. k-means ク

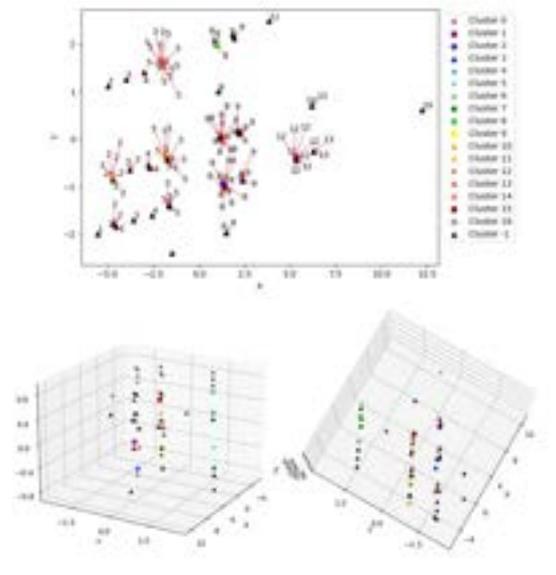


図8 分散表現に変換せず, データを直接 DBSCAN クラスタリングした結果. 図中の数字は, 「物体 ID」すなわち, 検知された YOLO ラベルを示している.

Figure 8 2D and 3D visualization of DBSCAN clustering results without Thing2Vec. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

ラスタリングでは, 時間帯や曜日は分類できているものの, 2 と 3, 12 と 13 といったように, 隣接する物体ラベルが同じクラスタに分布しているケースが多かった. これは, 離散化されたデータを直接プロットすると, 隣接するラベルが空間上で近い位置に配置されるためである. しかし, 曜日や時刻とは異なり, 隣接したラベルに意味はない. また, DBSCAN は隣接したラベルを同じクラスタとすることなくクラスタリングできたが, データが増えるにつれ, クラスタの収束に時間がかかる. したがって, 得られたデータを直接クラスタリングする方法では不十分であり, データを分散表現に変換することが重要である.

5. 議論

実験の結果, Thing2Vec によって得られた分散表現は, 物体 ID, 曜日, 出現時刻, 使用状態という使用情報に基づいたクラスタを形成することが確認できた. また, 宿題の最中のペンの置かれ方の状況によって, 別のクラスタに分類されることが分かった. したがって, たとえば宿題中のクラスタの遷移回数から, 宿題の難易度 (ペンを置いて考えた頻度), 宿題に対する集中度 (途中の離席回数) を分析できる可能性がある. 携帯電話については, 平日と休日で充電される時間帯が異なることが別クラスタに分かれ, 何時に在宅しているかを判別できることが明らかになった. こ

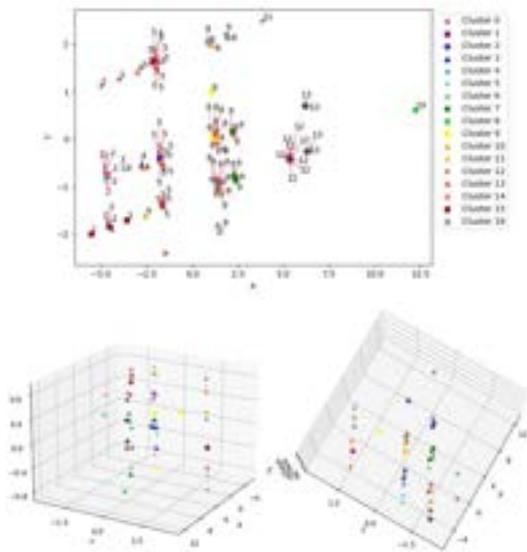


図9 分散表現に変換せず、データを直接 k-means クラスタリングした結果. 図中の数字は、「物体 ID」すなわち、検知された YOLO ラベルを示している。
 Figure 9 2D and 3D visualization of k-means clustering results without Thing2Vec. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

のような人の無意識な行動の変化を捉えることができれば、さまざまな分析に活かすことができるだろう。さらに冬休み中の例外的な行動が別クラスに分類されたことから、普段と違う行動の変化や異常を検出することもできるのではないかと考えている。ノイズに分類されたデータについては、データセットに出現時間や物体ラベルに関する偏りがあるため、データ量を増やせば新たなクラスを形成する可能性がある。k-means クラスタリングの結果に関しては、少ないデータの特徴を表現するのに出力層の 576 という次元が大きすぎたため、データが粗になる部分が多く、クラスタリングが意図しない結果になってしまったと考えられる。データを増強することで、さらに次元を増やし、月日やモノの位置など本実験で導入しなかった使用情報も埋め込める可能性がある。また、本実験では実装の都合上、YOLO ラベルと物体 ID を同義として扱ったが、個々の物体を追跡できれば、それぞれのモノについてより詳細な分析が可能となるだろう。

分散表現に変換せずに k-means を実行すると、互いに隣り合う異なるラベル (8 と 9 など) が同じクラスに分類されてしまうことがあった。これは、離散化されたデータを直接プロットすると、隣接するラベルが空間的に近接して配置されるためである。DBSCAN は離散化されたデータでも分類できたが、データ数と次元数が増えると収束に時間がかかってしまう。したがって、モノの使用情報を提案

する Thing2Vec を用いて分散表現に変換することで、物体履歴を人の理解に近い形でクラスタリングできるようになり、スケーラビリティも高まる。

Thing2Vec によって得られた分散表現は、異常検知や生活パターンの識別、人の行動理解など幅広い応用が期待できる。今回は少ないデータでの検証であったためクラスタリングが形成されることの確認にとどまったが、より大規模なデータセットを用意することでモノの使い方に関する新たな知見が得られるのではないかと期待している。また、個人ごとのデータを収集することで、生活習慣のクラスタリングや、家族構成によるクラスタリング、家庭状況の変化の検出などが可能になると考えられる。

本実験では固定カメラによるデータ収集を行い、固定された場所でのモノの使い方の情報を収集したが、ウェアラブルカメラを活用したデータ収集方法も考えられる。ウェアラブルカメラを用いると、個人の視点に近いデータを収集できるようになり、個人が使用するモノに関する分析ができるようになる。本手法においては、データの収集方法がボトルネックであるが、室内を常時撮影するような監視やライフログのようなアプリケーションと併用すれば、収集は容易になると考えている。

6. おわりに

本論文では、モノの使用情報に基づいて、モノの使用特性をベクトル埋め込みする手法 Thing2Vec を提案した。Thing2Vec は、Area2Vec を参考にしたモノの分散表現を作成方法であり、「物体 ID」、「曜日」、「出現時刻」、「使用状態」を用いてモノの使用情報をベクトルに埋め込む。モノの使用情報は、固定した可視光カメラとサーマルカメラで撮影した画像に対し、物体検知および接触検知を行うことで収集する。

実験では、実験者の日常生活の一部を撮影し、Thing2Vec によってモノの使用特性を捉えることができるか検証を行った。実験結果、Thing2Vec によって獲得された分散表現は、使用情報に基づいたクラスを形成することが確認できた。今後の取り組みとしては、データをさらに増強し生活パターンの特定や異常検知、人の行動理解に活用したいと考えている。

参考文献

[1] Kishino, Y., Shirai, Y., Yanagisawa, Y., Ohara, K., Mizutani, S. and Suyama, T.: Identifying human contact points on environmental surfaces using heat traces to support disinfect activities: poster abstract, *Proc. of the Conference on Embedded Networked Sensor Systems* (2020).

- [2] Le, Q. and Mikolov, T.: Distributed representations of sentences and documents, *International conference on machine learning*, PMLR, pp. 1188–1196 (2014).
- [3] Grover, A. and Leskovec, J.: node2vec: Scalable feature learning for networks, *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 855–864 (2016).
- [4] Narayanan, A., Chandramohan, M., Venkatesan, R., Chen, L., Liu, Y. and Jaiswal, S.: graph2vec: Learning distributed representations of graphs, *arXiv preprint arXiv:1707.05005* (2017).
- [5] Reddy, P., Gharbi, M., Lukac, M. and Mitra, N. J.: Im2vec: Synthesizing vector graphics without vector supervision, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7342–7351 (2021).
- [6] Mikolov, T.: Efficient estimation of word representations in vector space, *arXiv preprint arXiv:1301.3781*, Vol. 3781 (2013).
- [7] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. and Dean, J.: Distributed representations of words and phrases and their compositionality, *Advances in neural information processing systems*, Vol. 26 (2013).
- [8] Barkan, O. and Koenigstein, N.: Item2vec: neural item embedding for collaborative filtering, *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, IEEE, pp. 1–6 (2016).
- [9] 庄子和之, 青木俊介, 米澤拓郎, 河口信夫ほか: 大規模ユーザの滞在情報に基づくエリアの特徴付けと COVID-19 による影響分析, *情報処理学会論文誌*, Vol. 62, No. 10, pp. 1644–1657 (2021).
- [10] Tanaka, M., Kishino, Y. and Shimonishi, H.: Efficient Object Retrieval System Using Contact History from a Thermal Camera, *IEEE 43rd International Conference on Consumer Electronics ICCE2025* (2025).
- [11] Cheng, T., Song, L., Ge, Y., Liu, W., Wang, X. and Shan, Y.: YOLO-World: Real-Time Open-Vocabulary Object Detection, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)* (2024).
- [12] Ester, M., Kriegel, H.-P., Sander, J., Xu, X. et al.: A density-based algorithm for discovering clusters in large spatial databases with noise, *kdd*, Vol. 96, No. 34, pp. 226–231 (1996).

カードゲームの心理戦における表情の頻度分析

長澤 颯音^{†1} 角 薫^{†2}

公立はこだて未来大学^{†1}

公立はこだて未来大学^{†2}

1. はじめに

近年、表情認識技術の進展により、感情認識システムがさまざまな分野で活用されている。特に、Azure Face API, OpenFace, FaceReader などのツールは、顔の表情から感情を識別することを可能にし、マーケティング、教育、メンタルヘルス、セキュリティ分野での応用が進んでいる。しかし、これらのシステムは、自然な表情の識別を主な目的としており、意図的に操作された表情（例：ブラフや感情の抑制）が適切に認識されないことが課題とされている[1, 6]。

顔の表情を体系的に分析する手法として、Ekman and Friesen によって提案された Facial Action Coding System (FACS) [7, 8] が広く用いられている。FACS は、顔の動きを Action Unit (AU) という基本要素に分解し、それぞれの AU の組み合わせから感情を識別する標準的な分析手法である。この手法は、心理学や生理学の研究にとどまらず、アニメーション、ゲーム開発、医療診断などの分野においても、人間の表情の詳細な解析や感情推定の基盤として活用されている[2, 3]。しかし、心理戦のような場面における意図的な表情操作には十分に対応できるとは限らない。

ポーカーや交渉といった心理戦を伴う状況では、プレイヤーは感情を隠したり、逆に欺瞞的な表情（ブラフ）を用いたりする。Ekman によれば、感情は特定の AU によって表出されるが、心理戦における表情の操作は、これらの AU の出現頻度や組み合わせを変化させる可能性がある[5]。さらに、Pentland は、非言語的なシグナルが社会的な相互作用において重要な役割を果たすことを指摘[15]しており、心理戦の場面における表情操作の分析には、従来の感情認識手法を超えたアプローチが求められる。

本研究では、心理戦における表情データと生体情報（心拍数）の分析を通じて、ブラフの際の特徴的なパターンを明らかにすることを目的とする。具体的には、

1. 心理戦を伴うカードゲームの実験を通じて、プレイヤーの表情と生体情報を収集する。
2. FACS を用いて、感情やブラフ時の AU 出現頻度を解析し、特徴的なパターンを特定する。
3. 心拍数データを統計的に分析し、ブラフ時における生理的反応の変化を検証する。
4. 表情と生体情報の相関を分析し、ブラフの自動検出に向けた基礎的知見を得る。

本研究の成果は、将来的にアフェクティブコンピューティングの分野におけるブラフの自動検出に応用できる可能性がある。例えば、カードゲームの補助システムとして、プレイヤーの心理状態をリアルタイムで解析することで、新しい戦略支援の実現が期待される。本論文では、まず関連研究について述べ、その後、実験方法を詳述する。続いて、実験結果を示し、結果の考察を行う。最後に、本研究の結論と今後の展望を述べる。

2. 関連研究

表情認識技術は、近年のコンピュータビジョンおよび機械学習の進展により大きく発展している。従来の心理学的研究では、Ekman の Facial Action Coding System (FACS) に基づく表情分析が広く用いられ、基本感情の分類が行われている[8]。近年では、FACS を機械学習技術と組み合わせることで、より精度の高い自動表情認識システムが開発されており、リアルタイムの感情推定や欺瞞行動の検出にも応用されている[1]。

2.1 表情認識と心理戦

表情認識に関する研究の多くは、自然な感情表現を対象としており、意図的に操作された表情（例：ブラフや隠された感情）を分析する研究は限定的である[13]。特に、心理戦を伴う状況では、プレイヤーが意図的に表情を操作し、相手を欺く戦略を取ることが一般的であり、このようなケースでは既存の感情認識システムの精度が低下することが報告されている[9]。

ポーカーなどの対人ゲームにおいては、表情や非言語の手がかりがプレイヤーの戦略に与える影響が注目されており、FACS に基づく分析を行った研究も存在する[9]。これらの研究では、プレイヤーが意図的に表情を操作し、ポジティブな表情（微笑み）を用いて相手を欺いたり、逆にネガティブな表情（困惑・緊張）を抑制したりする戦略を取ることが示唆されている。しかし、表情だけでなく生体情報（心拍数や皮膚電気活動など）のデータを統合して分析する試みはまだ少ない。

2.2 生体情報と心理状態

感情認識においては、表情の変化とともに生体情報の測定が重要な役割を果たす。特に心拍数（HR）や皮膚電気活動（EDA）は、感情やストレス状態の指標として利用されるこ



図 1. 実験の実施風景

とが多い[11]. 過去の研究では、緊張やストレスの増加に伴い、心拍数や皮膚電気活動の変化が顕著に現れることが報告されている[20].

一方で、ブラフのように意図的に感情を隠す状況では、表情と生体情報の間にギャップが生じる可能性がある。例えば、顔には微笑みのようなポジティブな表情が現れていても、心拍数や EDA が増加している場合、それは欺瞞の兆候として解釈される可能性がある[10]. このような表情と生理的反応の乖離を分析することで、ブラフの検出精度を向上させることが期待される。

2.3 アフェクティブコンピューティングにおける位置づけ

アフェクティブコンピューティングは、人間の感情をコンピュータが理解し、適切に対応することを目指す研究分野[16]であり、感情認識技術はその中心的な要素である[9]. 従来の感情認識システムは、一般的な感情状態（喜び、怒り、驚きなど）を分類することに重点を置いてきたが、心理戦のような意図的に感情を操作する状況に対しては、十分な対応がなされていない。

本研究のように、心理戦における表情と生体情報の特徴を統合的に分析することは、アフェクティブコンピューティングの領域において重要な貢献となる。特に、ブラフのような欺瞞的な表情を検出するためには、表情認識技術と生理データを組み合わせたマルチモーダル解析が有効であると考えられる。本研究の成果は、将来的にリアルタイムの感情解析システムや、対人ゲームにおける戦略支援システムの開発につながる可能性がある。

2.4 ブラフ検出の試み

近年のアフェクティブコンピューティングの分野では、ブラフの検出に関する研究が進められている。機械学習を活用した研究では、表情の微妙な変化（マイクロエクспレッション）や生体情報のパターンを分析することで、欺瞞の識別を試みている[4]. また、近年の研究では、ディーラーニングを用いた手法が表情や生体情報のパターン認識において有望な結果を示している[12].



図 2. アノテーションツールの画面キャプチャ

しかし、既存の研究では、心理戦における表情と生体情報を統合的に分析した研究は限られている。本研究では、FACS に基づく AU の分析と生体情報の測定を組み合わせることで、心理戦におけるブラフの特徴を明確にし、将来的な自動認識システムへの応用を目指す。

3. 実験で使用したシステム

本研究では、心理戦における表情の特徴と生体情報の変化を分析するために、カードゲームを用いた実験を行った。実験では、被験者の表情データと心拍数を記録し、心理戦の状況下での表情の変化やブラフ時の特徴を定量的に評価した。

3.1 実験概要

本研究の目的は、心理戦における表情と生体情報の特徴を明らかにし、ブラフの識別に寄与する基礎データを提供することである。これを達成するために、大学生 12 名（男性 7 名、女性 5 名）を対象とし、カードゲームを通じてデータを収集した(図 1). 各被験者には、4 試合の対戦を行ってもらい、その間の表情データおよび生体情報（心拍数・皮膚電気活動）を記録した。実験は約 60 分間で、試合の実施に加えて、事前アンケート、試合後アンケート、およびアノテーション作業を含んでいる。

3.2 実験環境

3.2.1 OpenFace

本研究では、表情認識の方法として、OpenFace を用いた。OpenFace [1] は、Tadas Baltrušaitis 氏が MultiComp Lab との共同研究により開発した自動表情運動検出システムであり、オープンソースとして公開されている。このシステムは、表情分析や視線追跡などの研究分野で広く利用されており、任意のカメラ、写真、映像ファイルから詳細な顔の動作情報を取得できる点が特徴である。

具体的には、68 個の顔の特徴点（目、口、鼻、眉、顎などの部位の位置情報）、顔の向き、17 種類の Action Unit (AU) 要素、視線情報などを抽出することができる。被験者の表情

をより正確に捉えるため、カメラはテーブル中央および各被験者の正面に設置し、試合中の表情をリアルタイムで記録した。

3.2.2 アノテーションツール

本研究では、試合中の参加者が感じた感情や表情の偽装(ブラフ)を記録してもらうために、感情・ブラフアノテーションツールを使用した。このツール[18, 19]は、参加者が映像を見ながら感情やブラフを直感的にラベル付けできる仕様になっている。「感情」をラベル付けするページ、「ブラフ」をラベル付けするページの2種類を作成した(図2)。

ラベル付けの方法は、該当ボタン(例: 喜び, 驚き, 困惑, 緊張など)をクリックするとタイムライン上に矢印が表示され、ドラッグで範囲指定や調整をする形となっている。このツールにより、被験者自身が感じた感情をより正確に記録することができ、ラベル付けされたデータはJSON形式で保存された。データには、動画ファイル名、動画の総再生時間、各注釈の開始時間と終了時間、感情やブラフの種類、およびラベルID(ツール内での識別用)が含まれている。

3.2.3 使用したカードゲーム

本実験で使用するカードゲームは、「じゃんけん」を題材にしたカードゲームである。このカードゲームは、yutanisiguchi氏[14]が開発した既存のアプリであり、手札は「グー」「チョキ」「パー」のいずれかと、1から10までの得点を書かれたカードが2枚ずつ含まれている。

このカードゲームを使用した理由として、まずルールの覚えやすさが挙げられる。本ゲームは「じゃんけん」というシンプルで馴染み深い基礎ルールに基づいており、参加者は即座に理解できるため、心理戦や駆け引きに集中しやすい。また、心理戦の誘発がしやすい点も重要である。お互いの手札が常に公開されているため、相手の選択肢や戦略を予測する要素が含まれており、ブラフを含めた高度な駆け引きを促進する。

実験では、あらかじめ決められた組み合わせの手札4試合分を用意した(表1)。これにより、各プレイヤーが全ての状況を均等に体験できるようにし、公平性を確保した。また、手札を事前に設定することで運の要素を排除し、相手の手札を読む必要性を強調することで、心理戦を促進することを期待している。

表1 実験で使用したカードの組み合わせ

試合数	順番	1枚目	2枚目	3枚目	4枚目
1試合	先行	パー, 3	グー, 5	チョキ, 7	無
	後攻	パー, 3	グー, 5	チョキ, 7	無
2試合	先行	グー, 3	チョキ, 5	パー, 7	無
	後攻	グー, 7	チョキ, 5	パー, 3	無
3試合	先行	グー, 2	グー, 3	パー, 4	無
	後攻	グー, 4	グー, 5	チョキ, 9	無
4試合	先行	パー, 4	チョキ, 3	チョキ, 3	グー, 9
	後攻	パー, 7	チョキ, 4	チョキ, 4	グー, 6

3.2.4 Big Five 性格特性 (TIPI-J)

加えて、実験参加者の Big Five 性格特性を測定するために、TIPI-J[21]を使用してアンケートを作成した。TIPI-Jとは、Big Five 性格特性の5因子(外向性, 協調性, 勤勉性, 神経性傾向, 開放性)を10項目で測定する尺度である Ten Item Personality Inventory の日本語版である。アンケートの各項目は1(あまりそう思わない)から5(とてもそう思う)の5段階評価とした。

3.2.5 OpenSignals

本研究では、生体データの収集および解析において OpenSignals(pluX)[17]を使用した。OpenSignalsは、BITalinoなどの生体センサーデバイスと連携し、リアルタイムで心拍数(HR)や皮膚電気活動(EDA)を記録・解析するためのプラットフォームである。

このシステムは、データの視覚化やエクスポート機能を備えており、実験中に取得した生理データの精度を高めるのに貢献した。

3.3 実験手順

実験は、事前説明と同意取得の後、事前アンケートを実施し、次に生体情報測定のためのセンサー装着とキャリブレーションを行った。その後、カードゲームの試合を実施し、試合後アンケートとアノテーション作業を行った。まず、被験者には実験の目的と手順について説明し、データの匿名化やプライバシー保護に関する同意書に署名を求めた。その後、被験者の性格特性や心理戦の経験に関する情報を収集するために、事前アンケートを実施した。

本実験の結果に基づき、心理戦における表情の特徴と生体情報の変化を明らかにする。次章では、実験の結果を示し、統計的解析を行う。

4. 結果

本章では、実験で得られたデータを整理し、アンケート結果、表情データの分析結果、生体データの分析結果について詳細に報告する。

4.1 アンケート結果

4.1.1 実験前アンケート結果

本研究では、心理戦における表情の役割を明らかにするため、実験前に参加者へアンケートを実施した。質問内容は、性別、カードゲームの経験、心理戦の得意・不得意、表情操作のスキル、相手の表情を読む能力、心理戦における表情の重要性についてである。

(1) カードゲームの経験

自由記述形式で「得意なカードゲームのジャンル」を尋ねたところ、ポーカーを挙げた参加者が最も多く、次いで「シャドウバース」「エボルブ」「バトルスピリッツ」といったトレーディングカードゲーム(TCG)のプレイヤーが見られた。また、「神経衰弱」などの記憶系カードゲームを挙げる回答もあった。このことから、本研究の参加者には、対戦型カードゲームに慣れている者もいれば、戦略的なゲームに不慣

れな者も含まれていることが分かる。

(2) 心理戦や表情操作の得意・不得意

心理戦に関連するスキルとして、「人をだますことの得意さ」「表情を作ることの得意さ」「相手の表情をうかがう能力」について、4段階評価で回答を求めた。

「人をだますことが得意か？」の質問に対して、「得意」と答えた者(4点)はいなかったが、「やや得意」(3点)と回答した参加者は6名(50%)、「やや苦手」(2点)は1名(8.3%)、「苦手」(1点)は5名(41.7%)であった。これにより、参加者の中にはブラフ(嘘をつく行為)に苦手意識を持つ者が一定数存在することが示された。

「表情を作るのが得意か？」の問いでは、「やや得意」(3点)が最も多く8名(66.7%)を占め、「苦手」と答えた者はいなかった。この結果から、参加者の大半がある程度意図的に表情を操作する能力を持っていることが示唆される。

「相手の表情をうかがうことが得意か？」の質問では、「やや得意」(3点)と答えた者が8名(66.7%)、「得意」(4点)が1名(8.3%)となった。これにより、ほとんどの参加者が心理戦において表情を手がかりにする意識を持っていることが分かった。

(3) 心理戦における表情の重要性

「心理戦では、相手の表情を読むことが重要か？」という質問に対し、83.4%の参加者が「非常に重要である」と回答し、「ある程度重要である」と答えた者を含めると、全体の91.7%が表情の重要性を認識していた。これにより、ほぼ全員が心理戦において表情が重要な要素であると認識していることが確認された。

この結果から、参加者は心理戦における表情の重要性を強く認識している一方で、ブラフを得意とする者と苦手とする者に分かれることが示された。また、表情の操作能力や他者の表情を読むスキルにも個人差が見られることから、心理戦における表情の役割はプレイヤーごとに異なる可能性がある。この点を踏まえ、本研究では表情データと生体情報を組み合わせた分析を通じて、心理戦のブラフに関するパターンを明らかにすることを目指す。

4.1.2 試合後アンケート結果

試合後に実施したアンケートでは、プレイヤーが心理戦をどのように体験したか、ブラフの成功率、相手の表情の影響について調査した。

(1) ブラフの成功率と表情操作

ブラフを行ったプレイヤーに「平常を装うことができたか？」と尋ねたところ、74% (23名) が「できた」と回答し、26% (8名) が「できなかった」と回答した(表 2)。この結果から、多くの参加者がブラフを成功させたと感じていたものの、約 1/4 の参加者は表情をコントロールしきれずに相手に見破られる可能性があったことが示唆される。

また、「相手が嘘をついていると感じたか？」という質問では、33% (16名) が「感じた」と回答し、67% (32名) は

「感じられなかった」と回答した(表 2)。これにより、心理戦において相手のブラフを正確に見抜くことは容易ではないことが分かる。

(2) 質問の効果と心理戦の緊張感

試合中に質問を行ったプレイヤーのうち、「質問が効果的だった」と感じた者は 52% (15名) で、「効果がなかった」と感じた者は 48% (14名) であった(表 7)。また、「質問された際に緊張したか？」という問いには、48% (14名) が「緊張した」と回答し、42% (10名) が「緊張しなかった」と回答した(表 2)。

これらの結果から、心理戦では質問の有効性がケースバイケースであり、また、質問されることが必ずしも全てのプレイヤーに緊張をもたらすわけではないことが示された。

(3) プレイヤーの優位性の認識

「相手より優位にいたると感じたか？」の質問では、44% (21名) が「優位に感じた」と回答し、56% (27名) は「優位に感じなかった」と回答した(表 2)。この結果から、プレイヤーの多くは試合中に心理的な優位性を強く感じていなかったことが分かる。この傾向は、心理戦が互いに相手の動きを推測し合うバランスの取れた戦略的な要素を含んでいることを示唆する。

表 2 試合後アンケート (Q4~Q11) の回答結果

質問内容	選択肢 1	選択肢 2
	できた	できなかった
Q4(嘘をついた時、平常を装うことはできましたか?)	23(74%)	8(26%)
Q5(自分の質問は効果的だったと思いますか?)	効果的だった	効果がなかった
	15(52%)	14(48%)
Q6(質問されたとき緊張しますか?)	緊張した	緊張しなかった
	14(48%)	10(42%)
Q7(質問に対してすぐ答えましたか?)	すぐ答えた	すぐ答えられなかった
	15(63%)	9(37%)
Q8(質問に対して自信をもって答えましたか?)	答えられた	答えられなかった
	18(75%)	6(25%)
Q9(相手が嘘をついていると感じましたか?)	感じた	感じられなかった
	16(33%)	32(67%)
Q10(相手の手を予想できましたか?)	予想できた	予想できなかった
	24(50%)	24(50%)
Q11(相手より優位にいたると感じましたか?)	感じた	感じなかった
	21(44%)	27(56%)

4.1.3 実験後アンケート結果

実験全体が終了した後、被験者に対して心理戦の難しさや、表情が試合結果に与えた影響についてのアンケートを実施した。

(1) 表情が試合結果に与えた影響

「表情を利用して試合に勝てたか？」という質問に対し、41.7% (5名) が「勝てた」と回答し、25% (3名) が「どちらともいえない」と回答、また 25% (3名) が「勝てなかつ

った」と回答した(表 3)。このことから、表情を駆使して試合に勝利したと感じた者が一定数存在する一方で、表情が結果に直結しなかったと感じる参加者もいたことが分かる。

(2) ゲームのルールの難易度

「カードゲームのルールは複雑だったか?」という問いには、50% (6名) が「単純」と回答し、33.3% (4名) が「やや単純」と回答するなど、約 8 割の参加者がルールを理解しやすいと感じていた(表 3)。このことから、本研究のカードゲームのルールは、心理戦に集中しやすい設計であったことが示唆される。

(3) 心理戦の発生頻度

「心理戦が起こったと感じたか?」の質問には、58.3% (7名) が「感じた」と回答し、25% (3名) が「やや感じた」と回答した(表 3)。この結果から、本研究のゲーム環境において、ほとんどの参加者が心理戦を実感していたことが分かる。

表 3 実験後アンケート (Q1~Q3) の回答結果

質問内容	選択肢 1	選択肢 2	選択肢 3	選択肢 4
	1(負けた)	2	3	4(勝てた)
Q1(表情を利用して試合に勝てましたか?)	3 (25%)	3 (25%)	1 (8.3%)	5 (41.7%)
	1(単純)	2	3	4(複雑)
Q2(カードゲームのルールは複雑でしたか?)	6 (50%)	4 (33.3%)	1 (8.3%)	1 (8.3%)
	1(感じない)	2	3	4(感じた)
Q3(心理戦が起こったと感じましたか?)	0 (0%)	2 (16.7%)	3 (25%)	7 (58.3%)

4.2 表情データ分析結果

本研究では、OpenFace を用いて取得した顔の表情データを分析し、各感情およびブラフ時の Action Unit (AU) の出現頻度を比較した。まず、各ラベルごとの AU の頻出頻度比率を算出し、各被験者の全試合の平均 AU 頻出頻度を 1 としたときの各感情ラベルごとの AU 頻出頻度比率をグラフ化した (図 3~7)。

分析の結果、特定の AU (例: AU12, AU6) の頻度が、ブラフ時や喜びの表情で高くなる傾向が確認された (図 3, 図 7)。特に AU12 (口角の引き上げ) や AU6 (頬の上昇) は、心理戦において高頻度に出現し、ポジティブな印象を与える表情が戦略的に使用されている可能性が示唆された。一方で、緊張時には AU4 (眉を寄せる動き) の頻度が増加する傾向が見られた (図 6)。また、驚きや困惑の表情においても、他の感情と異なる AU の出現パターンが確認された (図 4, 図 5)。

さらに、t 検定を用いて有意差を検証した結果、特定の AU において有意な差が見られた。特に AU12 や AU6 の頻度

が高まること、心理戦においてポジティブな印象を意図的に活用する戦略として機能している可能性を示している。

4.3 生体データ分析結果

本研究では、心理戦における生理的変化を検討するために、BITalino を用いて心拍数 (HR) を測定・解析した。

ブラフ時と通常時の心拍数の平均値を比較した結果、ブラフ時の平均心拍数 (M = 89.50, SD = 6.57) は、平常時 (M = 87.31, SD = 7.89) よりわずかに高い傾向を示した。しかし、対応のある t 検定の結果、有意な差は認められなかった (t(7)=1.56, p=0.16)。また、ブラフ時と平常時の心拍数には強い正の相関が見られた (r=0.865)。

さらに、標準偏差の分析から、ブラフ時の心拍数のばらつき (SD = 6.57) は平常時 (SD = 7.89) よりも小さいことが確認された。これは、ブラフ時には心理的な負荷がかかり、心拍数が一定の範囲に収束しやすい可能性を示唆する。一方で、平常時の心拍数はより個人差が大きく、リラックス状態や集中度合いによってばらつきが生じたと考えられる。

これらの結果を踏まえると、心拍数の変化だけではブラフを識別するには不十分であり、皮膚電気活動 (EDA) や表情データなど他の生体情報と組み合わせた分析が必要である。今後、より多様な生理指標 (例: 脳波・視線データ) を統合的に解析し、心理戦時のブラフ検出精度を向上させることが求められる。

4.4 表情と生体データの関係

最後に、表情データと生体データの相関を分析し、心理戦における表情の特徴と生理的反応の関係を調査した。その結果、ブラフ時と平常時の心拍数には強い正の相関が見られたが、t 検定の結果、AU12 (微笑み) の頻出頻度と心拍数の上昇、およびブラフ時の表情と生体反応の間に明確な統計的関係は確認されなかった (図 8)。この結果は、サンプル数が比較的少ないことに起因する可能性があり、今後の研究では、より多くの被験者を対象としたデータ収集を行うことで、実験結果の精度を向上させることが求められる。

一方で、ブラフ時には AU12 (微笑み) の頻度が高まる一方で、生理的には心拍数の上昇が見られる傾向が示唆された。これがさらに検証されれば、心理戦においてプレイヤーが意図的にポジティブな表情を演出しながらも、内面では緊張を感じている可能性が示されることになる。今後、これらの特徴を活かし、心理戦における欺瞞の検出や、感情認識システムへの応用が期待される。

5. 考察

本章では、実験結果をもとに、心理戦における表情と生体情報の特徴について考察し、ブラフ時の表情と生理的反応の関係、心理戦の戦略としての表情操作、そして本研究の意義と今後の課題について論じる。

ブラフ



図 7. AU 頻出頻度比率 (ブラフ la eIID5)

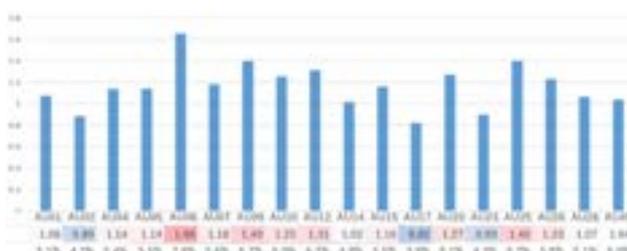


図 3. AU 頻出頻度比率 (喜び la eIID1)



図 4. AU 頻出頻度比率 (驚き la eIID2)

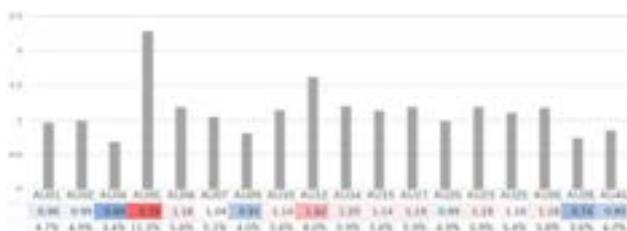


図 5. AU 頻出頻度比率 (困惑 la eIID3)

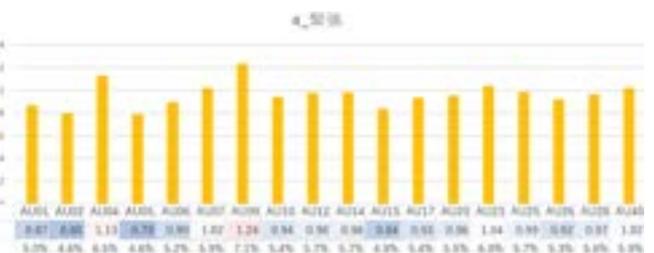


図 6. AU 頻出頻度比率 (緊張 la eIID4)

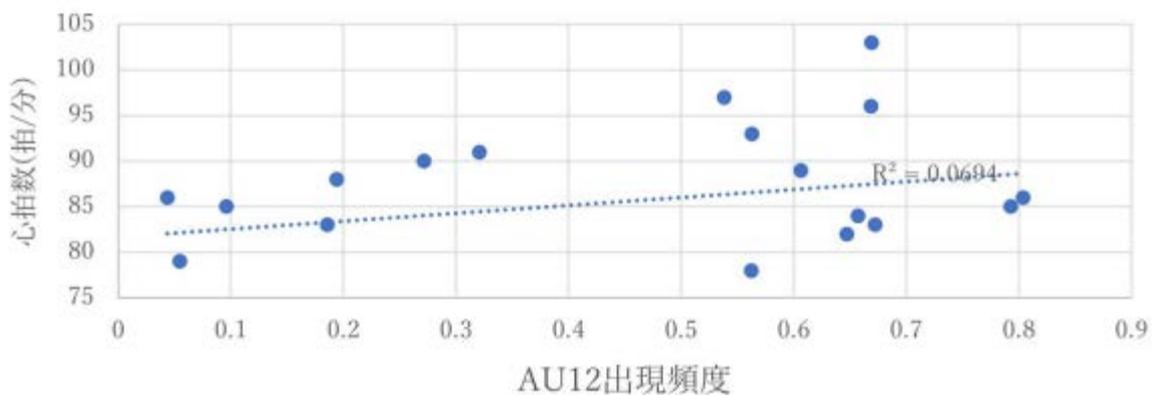


図 8. AU12 の頻出頻度と心拍数の関係

5.1 ブラフ時の表情と生体反応の特徴

実験結果から、心理戦におけるブラフ時には特定の表情パターンが高頻度で発生することが確認された。特に、AU12（口角の引き上げ）と AU6（頬の上昇）は、ブラフ時に有意に多く出現することが示された。これらの動きは、一般にポジティブな感情（喜びや安心感）と関連が深いため、プレイヤーが意図的にポジティブな印象を演出しようとする可能性がある。

一方で、生体データの分析結果から、ブラフ時には心拍数（HR）の上昇傾向が見られたが、統計的に有意な差とはならなかった。

この結果は、心理戦において、プレイヤーが内面の緊張を隠しながら、意図的に表情を操作する戦略を取っていることを示している。特に、表情の変化と生体データの乖離は、ブラフの識別に重要な要素となる可能性がある。

5.2 心理戦の戦略としての表情操作

ブラフを成功させるためには、プレイヤーが相手に対して信頼感や安心感を与える表情を作り出すことが求められる。本研究では、ブラフ時に AU12（微笑み）や AU6（頬の上昇）が多用されることが確認されており、プレイヤーは作り笑いをを用いて対戦相手を欺こうとする可能性がある。

また、驚きの感情を示す AU25（口を開く）や AU26（顎の下げ）も、ブラフ時に一定の頻度で見られた。これは、相手に「意外な展開である」と思わせるための演技的な要素が含まれている可能性がある。さらに、困惑のラベルが付与された場面では、AU12（微笑み）が頻繁に見られたことから、困惑を笑顔でごまかす行動も心理戦において利用されている可能性がある。

一方、緊張を示す AU9（鼻のしわ寄せ）や AU4（眉を寄せる動き）は、心理戦の中で相対的に抑制されていることが確認された。このことから、プレイヤーは意図的にネガティブな感情の表出を抑えることで、相手に心理的な手ごかりを与えないようにしていると考えられる。

5.3 本研究の意義と今後の課題

本研究は、心理戦における表情操作の実態を定量的に明らかにした点で意義がある。特に、ブラフ時の表情パターンと生理的反応の関係性を示したことにより、将来的に心理戦のブラフ検出や感情認識システムへの応用が期待できる。

5.3.1 ブラフ検出への応用

本研究の結果を活用することで、表情認識技術を用いたブラフ検出システムの構築が可能になると考えられる。特に、AU12（微笑み）や AU6（頬の上昇）と、生体データ（EDA）の変化を組み合わせることで、プレイヤーが「ポジティブな表情を見せながらも緊張している状態」を自動的に検出できる可能性がある。

将来的には、機械学習を活用し、心理戦のシナリオに応じたブラフ検出アルゴリズムを開発することで、ポーカーなどの対戦ゲームにおける AI の強化や、対人交渉における支援システムへの応用が期待される。

5.3.2 実験デザインの改良

本研究では、被験者数が 12 名と比較的少数であったため、今後はより多くの参加者を対象とした実験を実施し、データの一般化可能性を高める必要がある。また、被験者の個人差（性格特性や経験）による影響も考慮する必要がある。

特に、Big Five 性格特性と表情操作の関係性をさらに詳細に分析することで、プレイヤーの個人的な特性に基づいた心理戦の傾向を明らかにすることができると考えられる。

5.3.3 他の生体情報の活用

本研究では、主に心拍数を用いたが、今後は脳波（EEG）や筋電（EMG）、皮膚電気活動（EDA）を組み合わせた解析を行うことで、より詳細な心理的負荷の測定が可能になると考えられる。また、視線追跡技術を導入することで、プレイヤーがどのタイミングで相手を観察しているかを分析し、心理戦の戦略をさらに深く理解することができる。

6. 結論

本研究では、心理戦における表情と生体情報の特徴を分析し、ブラフ時の表情の特徴と生理的变化を明らかにした。主な知見は以下のとおりである。

- ブラフ時には AU12（微笑み）や AU6（頬の上昇）が有意に増加する。
- 心拍数には統計的に有意な変化は認められなかったが、ブラフ時と平常時の心拍数には強い正の相関が見られた。
- プレイヤーは意図的にポジティブな表情を作ることで、相手を欺こうとする戦略を取っていた。
- 困惑や驚きを装う行動もブラフの一部として活用される可能性がある。
- 心理戦において、表情と生体データの乖離はブラフ検出において重要な要素となる。

これらの結果を踏まえ、今後の研究では、より多様な被験者データを収集し、機械学習を用いたブラフ識別モデルの精度向上を図るとともに、さらなる生体情報を統合した心理戦の包括的な分析を行う。

参考文献

- [1] T. Baltrušaitis, "OpenFace 2.2.0: A Facial Behavior Analysis Toolkit," 2019. Available at: <https://github.com/TadasBaltrusaitis/OpenFace/wiki>.
- [2] M. S. Bartlett, G. Littlewort, M. Frank, and K. Lee, "Automatic decoding of facial movements reveals deceptive pain expressions," *Current Biology*, vol. 24, no. 7, pp. 738–743, 2014.
- [3] J. F. Cohn, Z. Ambadar, and P. Ekman, "Observer-based measurement of facial expression with the Facial Action Coding System," in J. A. Coan & J. J. B. Allen (Eds.), *The Handbook of Emotion Elicitation and Assessment*, pp. 203–221, Oxford University Press, 2007.
- [4] Z. Deng, X. Zhang, and Y. Fu, "Spotting deception from facial micro-expressions," *IEEE Transactions on Affective Computing*, vol. 10, no. 4, pp. 532–545, 2019.
- [5] P. Ekman, *Emotion in the Human Face*, Cambridge University Press, 1982.
- [6] P. Ekman, 顔は口ほどに嘘をつく, 河出書房新社, 2006.
- [7] P. Ekman, 表情分析入門, 誠信書房, 1987.
- [8] P. Ekman and W. V. Friesen, *Facial Action Coding System (FACS)*, APA PsycTests, 1978.
- [9] M. G. Frank, "The ability to detect deceit: A meta-analysis," *Journal of Applied Psychology*, vol. 91, no. 1, pp. 1–8, 2006.
- [10] R. E. Jack, "Facial expressions of emotion are not culturally universal," *Proceedings of the National Academy of Sciences*, vol. 109, no. 19, pp. 7241–7244, 2012.
- [11] J. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological Psychology*, vol. 84, no. 3, pp. 394–421, 2010.
- [12] Y. Li, M. Xu, and H. Tang, "Deep learning for human affect recognition: The state of the art," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 1–18, 2022.
- [13] D. Matsumoto and H. S. Hwang, "Microexpressions and deception," *Journal of Nonverbal Behavior*, vol. 35, no. 2, pp. 95–109, 2011.
- [14] niente, "じゃんけんカードバトル公開しました," note, 25 Feb. 2020. Available at: <https://note.com/niente0520/n/n06815c69a3de>. [Accessed: 16 Jan. 2025].
- [15] A. P. Pentland, *Honest Signals: How They Shape Our World*, MIT Press, 2008.
- [16] R. W. Picard, *Affective Computing*, MIT Press, 1997.
- [17] pluX, "OpenSignals," 2025. Available at: <https://www.pluxbiosignals.com/pages/opensignals>.
- [18] T. J. Tiam-Lee and K. Sumi, "Adaptive feedback based on student emotion in a system for programming practice," in *International Conference on Intelligent Tutoring Systems*, Springer, 2018, pp. 243–255.
- [19] T. J. Tiam-Lee and K. Sumi, "Analysis and prediction of student emotions while doing programming exercises," in *International Conference on Intelligent Tutoring Systems*, Springer, 2019, pp. 24–33.
- [20] B. Verschuere, G. Meijer, and E. Merckelbach, "Autonomic and behavioral responses to lying and truth telling in children and adults," *Journal of Experimental Psychology: Applied*, vol. 24, no. 3, pp. 373–386, 2018.
- [21] 小塩真司・阿部晋吾, "日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み," *パーソナリティ研究*, vol. 21, no. 1, pp. 1–10, 2012.

Immersion Neckwear: 動画体験拡張ネックウェア

岩本空^{†1}
青山学院大学^{†1}

ロペズギヨーム^{†2}
青山学院大学^{†2}

1. はじめに

定額制動画配信サービスの普及により、有料動画配信サービスの利用者が増加している。動画配信サービスの利用者数の需要予測では2025年には3900万人まで利用者が拡大すると予測されている[1]。これに伴い、動画視聴の形態も大きく変化している。好きなタイミングで視聴を開始でき、チケット代および交通費などのコストを削減可能なインターネット回線を活用した動画鑑賞が数年で市場規模を拡大している。2023年の動画配信（VOD）市場規模推計と、その後2028年までの各年の市場規模を3つのシナリオで予測した「動画配信（VOD）市場5年間予測（2024-2028年）レポート」によると、有料配信の市場規模はコロナ禍前の2019年と比較するとほぼ倍に成長すると予測されている。さらに、2028年には7371億円に増加することが予測されており、更なる成長が期待されている。有料配信の市場規模の拡大と付随して、特定の配信サービスのみで視聴可能なオリジナルコンテンツも増加している[2]。映画を見るときに、映画館で映画を見たい人が全体の3割を占めており、特に20代については男女ともに自宅で映画を見たい人より映画館で見たいと回答した人の方が上回っている。映画館で見たいと回答した人の中には映画館で見ることのメリットとして「観客の笑い声や泣き声などに包まれ、一体感をもたらしてくれるところ」と回答している[3]。他の調査でも映画を映画館で見たい理由として、非日常感を味わえることおよび、没入感を感じられることが上げられている[4]。株式会社スパコロが実施した映画館の利用意識調査によると、「コロナ禍の映画館においてみたいと思う作品はどんな作品か？」の問いに対して、回答結果を分析・可視化したところ「アクション」「迫力」「SF」「サスペンス」「映像」「臨場感」「アニメ」など語句が目立っており、劇場だからこそそのダイナミックな映画体験が感じられる作品が映画館で観たい作品と考えられていることがわかる[5]。また、技術の進化により、映画館での視聴体験も新たな楽しみ方として提案されており、特に4DX（Four Dimensional eXperience）技術の導入により、五感を刺激する没入型の体験が提供されている。よって、本研究の目的は、動画の映像に合わせて、五感を刺激するフィードバックを行い、自宅での動画体験を拡張することである。そのため、研究目標として、

映像に合わせて振動・風・温冷感を提示する頸部装着型ウェアラブルシステム Immersion Neckwear を創作し、一人での動画鑑賞の没入感を向上することを目指す。Immersion Neckwear を利用した動画鑑賞のイメージ図を図1に示す。



図1: Immersion Neckwear を利用した動画鑑賞の例

2. 関連研究

2.1 評価実験によるコンテンツ評価

杉原ら[6]は、音楽に対する感性がどのように表現されるかを探求した。音楽の聴取中に音楽が引き起こす印象を評価するために、SD法を使用した実験を実施し、音楽の特徴が聴取者の感情に与える影響を示唆した。楽曲から受ける印象の程度は男性と女性で異なることおよび、感性語対によっても異なることが判明した。仁科ら[7]は、観賞用の葉植物、花（バラ）、および香りが人間に与える生理的および心理的影響を、脳波（アルファ波とベータ波の比率）およびSD法を用いて分析した。その結果、香りの存在によりアルファ波とベータ波の比率が高くなることが観察された。香りが人間の生理面に影響を与えることを示唆した。一方、バラの存在は高い評価得点をもたらし、花が人間の心理面に影響を与えることも示唆した。寺本ら[8]は、非研究者が「没入感」をどのように概念化しているかを探求した。調査の結果、因子分析により、「評価」「インパクト」「活動性」「機械的性質」の4つの因子が没入感の構成要素として抽出された。没入感が高いイベントは、好ましく、印象的で、動的であると評価される傾向があった。安藤[9]は、臨場感は複数の感覚要素の複合体として捉えることが可能であると考えた。これらの感覚要素としては、立体感、質感、包囲感からなる空間要素、動感、リアルタイム感、同時感からなる時間要素、さらに自己存在感、インタラクティブ感、情感からなる「身体要素」を挙げられた。臨場感の評価手法として、主観評価、心理物理評価、脳活動計測、生体信号計測、行動計測の五つの手法を提案した。これらの手法を統合して臨場感を客観的・定量的に評価する必要がある

^{†1} Sora Iwamoto, Aoyama Gakuin University
^{†2} GUILLAUME LOPEZ, Aoyama Gakuin University

ことを示唆した。映像の質感評価では、立体映像が質感を強調し、心理物理実験により光沢感の定量的評価が行われた。飯村ら[10]は、VRを用いたシステムに関する臨場感と現実感の評価を行った。印象調査の結果、臨場感「動き」など視覚との結びつきが強く、映画またはゲームなどから感じ取れるもので、「その場にいるような感覚」と定義された。一方、現実感には視覚以外に痛覚との結びつきが強く、不安および痛みなどから感じ取れるもので、「非現実を現実と感じる」と定義された。因子分析の結果、臨場感には迫力因子と評価性因子、現実感には現実性因子があることがわかった。

2.2 フィードバックによるコンテンツ体験の変化

岡本ら[11]は、ワイヤアクションゲームをプレイ中の風の強度が臨場感に与える影響を検証した。風なし、一定の風、ワイヤアクションで引っ張られるときの速度に応じた風の3つの条件で評価したところ、風力変化がある状態が最も臨場感が高いことが判明した。前田ら[12]は、複数の体の部位に装着可能な熱フィードバックシステム TherModule を提案した。実験では、TherModule を装着し、視覚と温度フィードバックを受けながら映画を鑑賞した被験者は、視覚のみの被験者に比べて「楽しさ」と「興奮度」が有意に高いことが確認された。Kimら[13]は、触覚刺激が映画鑑賞中の感情変化と没入感に与える影響を調査した。ポジティブなシーンで柔らかい刺激を受けるとポジティブな感情と没入感が増加し、ネガティブなシーンでは柔らかい刺激がネガティブな感情を軽減することがわかった。

上記のように、風や温度、触覚刺激が没入感や感情に与える影響は示されたが、実用性の観点に問題がある。そのため、本研究では振動、風、温冷感を組み合わせた多感覚フィードバックシステム Immersion Neckwear を創作し、これが動画鑑賞における没入感やユーザ体験に与える影響を評価する。

2.3 映画鑑賞体験の拡張

Ohら[14]は、4DX映画の観客が感じる臨場感について調査し、4DX効果がどのように影響するかを分析した。7つのジャンルの映画を対象に、35名の観客に対して詳細なインタビューを実施し、グラウンド理論を用いて結果を解析した。4DX映画はリアリズムとしての臨場感、没入感、メディア内での社会的役割、社会的豊かさ、共感的移動感、社会的存在としての臨場感の6つのタイプを引き起こすことが判明した。特に、モーションコントロールおよび振動、空気、匂いの効果が臨場感を高める重要な要因であることが示唆された。Jeongら[15]は、4DX映画におけるモーション効果が観客の感情に与える影響を調査した。特に、共感レベルに応じた感情反応の違いを分析し、高い共感を持つ参加者が短いモーション効果でより強い恐怖を感じることを明ら

かにした。また、映画クリップとモーション効果の組み合わせが観客の感情を変化させることが示され、モーション効果の設計に関するガイドラインを提案し、感情体験の向上に寄与する方法を示唆している。Leeら[16]は、4DX映画におけるモーション効果の迅速な設計を可能にするアルゴリズムを提案した。4DX映画は視覚だけでなく、聴覚および触覚を通じて観客に没入感を提供するため、モーション効果が重要な役割を果たす。提案されたアルゴリズムは視聴者中心のレンダリング戦略に基づき、視覚的注意の動きに合わせて椅子の動きを調整した。これにより、観客の体験が向上が実現された。さらに、本アプローチは自動化されており、手動での作成に比べて10倍以上の速度で動き効果を生成することが可能になった。実験により、生成されたモーション効果の主観的品質が評価され、視覚的に適切な効果を提供することが確認された。このことから、4DX映画の制作プロセスを効率化し、観客の体験をより豊かにするための新たな手法を示唆された。Nicolae[17]は、従来の映画館とVR映画館の違いを分析した。VR映画館はヘッドセットが周囲の気を散らすものを遮断し、観客は映画に集中することを求められるが、従来の映画館での鑑賞と比較して社会交流が減少すると分析した。また、VR作品の質は様々で、強い感情的な没入感を提供するものもあれば、「そこにいる」効果のみに依存するものもあると述べた。VR技術は従来の映画では実現できない強力な「臨場感」を生みだした。Kimら[18]は、VRが人々の感情に与える影響を調査した。ヘッドマウントディスプレイ(HMD)を使用した視聴条件と使用しない視聴条件(No-HMD)を比較し、ホラーと共感の2種類の感情的コンテンツを適用した。結果、HMDを使用してホラー映画を見た視聴者は、No-HMDの視聴者よりも恐怖を感じやすかったことが判明した。しかし、共感を誘発する映画では、HMDとNo-HMDの視聴者の間に有意な感情の差は見られなかった。研究では、VRがホラー映画に対して特に感情反応を強化することを示し、没入感と感情の関係を確認した。

3. Immersion Neckwear の概要

3.1 Immersion Neckwear の基本設計

本研究で創作した頸部装着型の没入感向上デバイス、Immersion Neckwear、の画像は図2に示し、構成している主な要素は以下にまとめている。

- インターフェース部：nRF Connect
- 制御部：Arduino Nano 33 BLE
- 電源部：5500mAh リチウムイオンバッテリー
- 温冷部：ペルチェ素子 TEC1-03105
- 風部：ターボブロワファン B0DB5SDX91



図 2: Immersion Neckwear の全体像

- 振動部：振動モータ LBV10B-009

インターフェース部は、スマートフォンを用いてマイクロコンピュータと接続することで、映像に合わせてシステムの開始を出力する。

電源部は、制御部に電源を供給する。制御部は、他の構成要素の制御を行う。映像に合わせて電子部品の起動時間の制御を行う。また、電源部からの電源供給により、フィードバック部に電源供給を行う。

フィードバック部は、制御部からの電源供給により、映像に沿ったフィードバックを行う。

3.2 ハードウェアの設計

3.2.1 熱源の選定と温度制御

Immersion Neckwear では、温刺激、冷刺激を瞬時に生じさせる必要があった。よって、エネルギー効率が電熱線より高く、電流の向きにより加熱と冷却が同一の電子部品で実現可能な、ペルチェ素子 TEC1-03105(図 3)を用いた。TEC1-03105 のサイズは縦横 15mm、厚さ 3.1mm である。TEC1-03105 は、直流電流を流すことで一方の面が吸熱し反対面に発熱が起こす。電流の極性を逆転させると、特性が反転する。



図 3: ペルチェ素子

温度提示面の温度検出として、NTC サーミスタを用いた。NTC サーミスタは、温度上昇に対して抵抗値を減少させる性質があり、以下の式 (1) を満たす。よって、式変形した式 (2) を用い、温度検出を行った。

$$R = R_0 \exp \left(B \left(\frac{1}{T} - \frac{1}{T_0} \right) \right) \dots (1)$$

$$T = \left(\frac{1}{\frac{1}{T_0} + \frac{1}{B} \log \left(\frac{R}{R_0} \right)} \right) \dots (2)$$

R_0 温度 T_0 [K] 時のサーミスタ抵抗 [Ω]

R 温度 T [K] 時のサーミスタ抵抗 [Ω]

B サーミスタ B 定数

T_0 基準温度 [K]

T_1 測定温度 [K]

温度制御は、マイクロコンピュータの PWM (Pulse Width Modulation) 出力を用いて、PID 制御を行った。PWM 制御は、デジタル信号を用いてアナログ信号を模倣する手法であり、制御対象の平均電圧を調整することで精密な制御を実現する。これにより、デバイスはエネルギー効率を高めつつ、目的の動作を実現する。PWM 信号は、デューティサイクルと呼ばれる高電圧状態と低電圧状態の比率によって制御される。PWM 信号は一定周期で ON と OFF を繰り返すデジタル波形である。デジタル波形を用いることで、加熱素子および冷却素子に供給されるエネルギーの平均を調整し、温度を適切に制御することが可能になる。

また、本研究では、ペルチェ素子の同一面で温熱面と冷却面の両方を制御する必要があった。よって、モータに回転、逆回転、ストップ、プレーキ等の制御を行うことができるモータドライバを用いた。モータドライバ TA7291P と PWM 出力を用いてペルチェ素子の温度制御を行った。本研究では、温度センサを用いた PID 制御を行うことで、温熱面を 40°C、冷却面を 20°C に設定してフィードバックをした。PID 制御とは、proportional (比例)、integral (積分)、differential (微分) の三つの制御要素を組み合わせた制御手法であり、設定温度に対する誤差を最小限に抑えるためのものである。PID 制御を用いることで、目標とする温度に迅速かつ安定的に達成することが可能となる。具体的には、温度センサで検出した現在の温度と目標温度の差をもとに、ペルチェ素子への電力供給を調整する。これにより、過剰な温度変動を防ぎ、安定した温度制御が可能となった。

3.2.2 その他の電子部品の選定

体験に影響を与えることを防ぐために軽量化する必要があった。また、インターフェース部と制御部を Bluetooth で接続可能なマイクロコンピュータを選定した。Arduino Nano 33 BLE(図 4)は、Arduino が開発した長辺 40.64mm、短辺 17.76mm の小型のマイクロコントローラボードである。Bluetooth Low Energy (BLE) 機能を内蔵しており、スマートフォンおよび他の IoT デバイスとの無線通信が可能であ



図 4: Arduino Nano 33 BLE ([19]より引用)

る。本研究では、提案ボードの低消費電力性と小型性を活かし、システム全体の制御に利用した[19].

4. Immersion Neckwear の没入感効果検証実験

4.1 実験概要

Immersion Neckwear の有無による動画体験の印象の違いを比較し、有効性を検証した。被験者は 20 代の男女 10 名（男性:8 人、女性 2 人）が、25 分程度のアクションアニメを視聴し、視聴後 SD 法を用いたアンケートで主観的な評価を収集した。実験手順として、表 1 に対応している、フィードバックを起こす場面の個数がほぼ等しい 2 つの動画を用意し、被験者は Immersion Neckwear の有無で動画鑑賞を行った。視聴後には SD 法によるアンケートに回答し、さらに Immersion Neckwear を使用した被験者には SUS アンケートを実施した。SUS アンケートでは、システムの全体的な使いやすさおよび利便性、直感的な操作性などを評価した。図 5 は Immersion Neckwear を装着した場合の実験中の様子である。

表 1: 場面とフィードバックの種類の対応

場面の種類	フィードバック
そよ風、爆風、向かい風	風
爆発、炎が出現するシーン	温感
氷が出現するシーン	冷感
戦闘シーン(攻撃が当たるシーン)	振動

4.2 SD 法

SD 法は各刺激に対して人が抱く印象、イメージを明らかにするために用いる手法で、相反する形容詞対を多数用いて刺激を評価することにより、人がその刺激に対して、どのように感じるかといった情緒的な印象を明らかにすることが可能となる[20]。以下に、アンケートに用いる印象語対を記す。

- 好きな_____嫌いな



図 5: 実験の様子

- 良い_____悪い
- 気持ちの良い_____気持ちの悪い
- はっきりした_____ぼんやりした
- 迫力のある_____迫力のない
- リアリティのある_____リアリティのない
- 動的な_____静的な
- 弱い_____強い
- 興奮した_____落ち着いた

そして、各形容詞対を左から「非常に」、「かなり」、「やや」、「どちらでもない」、「やや」、「かなり」、「非常に」の 7 段階で評価した。

4.3 SUS

SUS は、製品およびサービスの使いやすさを評価するためのアンケートのことである。これは、ソフトウェア、ハードウェアを問わず、さまざまな新しいシステムのユーザビリティを評価し、洞察を得るための定量的な手法として使用される。リッカート尺度を使用して回答する 10 個の質問で構成される。本研究では、「強く同意する」から「強く同意しない」までの 5 段階で評価した[21].

まず、SUS 調査の各回答選択肢を対応する評定を割り当てる。回答が「強く反対」の場合は 1 点、「反対」の場合は 2 点、「どちらでもない」の場合は 3 点、「同意する」の場合は 4 点、そして「強く同意」の場合は 5 点とする。

各質問項目への回答（評点 X）に基づき、以下の方法で得点を計算する。

- 奇数番目の項目：X-1 点
- 偶数番目の項目：5-X 点
- 未回答の項目：3 点

これらの得点を合計し、その合計値に 2.5 を乗じることで、0-100 点の SUS スコアが算出する。SUS スコアが高いほど、評価対象の製品、システム自体、およびその利用開始までの過程または利用後に期待される効果に対するユーザの満足度が高いと判断される[22].

SUS(System Usability Scale)のスコアリングと解釈について、以下のようにまとめる。

- A+: 84.1–100 (96th–100th percentile)
- A: 80.8–84.0 (90th–95th percentile)
- A -: 78.9–80.7 (85th–89th percentile)
- B+: 77.2–78.8 (80th–84th percentile)
- B: 74.1–77.1 (70th–79th percentile)
- B -: 72.6–74.0 (65th–69th percentile)
- C+: 71.1–72.5 (60th–64th percentile)
- C: 65.0–71.0 (41st–59th percentile; 平均的なユーザビリティ)
- C -: 62.7–64.9 (35th–40th percentile)
- D: 51.7–62.6 (15th–34th percentile)
- F: 0.0–51.6 (0th–14th percentile; 低いユーザビリティ)

グローバル平均 SUS スコアは約 68 で、これは C グレード(平均的なユーザビリティ)に相当する。50 未満のスコアは欠陥があると見なされ、80 以上のスコアを持つシステムは、平均以上のユーザエクスペリエンスを示し、優れたユーザビリティに関連している。65 から 80 のスコアは受け入れ可能だが、改善の余地があることを示唆している。65 未満のスコアは、ユーザビリティの問題があることを示しており、注意が必要である。

5. 実験結果および考察

5.1 SD 法の評価結果

動画鑑賞後に SD 法を用いたアンケートを実施した。本アンケートを「プロフィール分析」と「因子分析」を利用してデータの分析を行うことで、Immersion Neckwear の有無による映画体験の没入感の変化を評価した。図 6 は被験者が回答したデータから、評定尺度ごとの回答平均値を算出して図式化したプロフィール分析の結果である。縦軸は各形容詞対を表し、横軸は形容詞対に対する 7 段階の評価尺度を表す。青線が Immersion Neckwear を装着した場合の動画鑑賞時の印象を示しており、赤線が Immersion Neckwear を装着しなかった場合の動画鑑賞時の印象を示している。プロフィール分析によって、Immersion Neckwear を装着した場合としなかった場合の印象の違い、Immersion Neckwear の装着が被験者の印象に与える影響を視覚的に比較可能になる。



図 6: Immersion Neckwear のプロフィール分析
共通因子を特定するために、アンケートデータに対して

因子分析を実施した。Immersion Neckwear を装着した場合の因子分析の結果を表 2 に示す。初期の固有値の大きさと減衰状況から、因子数は 3 と判断した。因子負荷量は、因子と変数の間に有意な相関があるとされている絶対値 0.4 以上とし、斜交回転法を用いて因子間の相関関係を考慮できるプロマックス回転を適用した。因子負荷量の絶対値が 1 に近いほど、因子との関係性が強いことを示す。因子 1 は「好きな」、「良い」、「迫力ある」などの形容詞と強く関連していることがわかった。また、因子 2 は「興奮した」、「強い」、「気持ちのいい」、因子 3 は「はっきりした」、「リアリティのある」などの形容詞と関連が深いことが判明した。

表 2: システムを装着する時の因子負荷量

形容詞 1	形容詞 2	因子 1	因子 2	因子 3
嫌いな	好きな	0.888	-0.376	-0.073
悪い	良い	0.820	0.321	-0.043
迫力のない	迫力のある	0.744	0.317	0.189
落ち着いた	興奮した	-0.027	0.982	-0.245
弱い	強い	-0.080	0.651	0.107
気持ちの悪い	気持ちのいい	-0.278	0.505	0.207
ぼんやりした	はっきりした	0.062	0.055	0.998
リアリティのない	リアリティのある	-0.108	-0.024	0.824
寄与率		30.282	20.791	14.951
回転後の負荷量平方和		2.457	2.313	1.849

Immersion Neckwear を装着しなかった場合の因子分析の結果を表 3 に示す。因子負荷量の絶対値が 1 に近いほど、因子との関係性が強いことを示す。因子 1 は「迫力のある」、「良い」、「興奮した」、「強い」などの形容詞と強く関連していることがわかった。また、因子 2 は「はっきりした」、「リアリティのある」などの形容詞と関連が深いことが判明した。

表 3: システムを装着しない時の因子負荷量

形容詞 1	形容詞 2	因子 1	因子 2
迫力のある	迫力のない	0.953	0.157
良い	悪い	0.861	-0.066
興奮した	落ち着いた	0.669	-0.256
強い	弱い	0.500	-0.027
はっきりした	ぼんやりした	0.116	0.994
リアリティのある	リアリティのない	-0.232	0.842
寄与率		46.164	30.801
回転後の負荷量平方和		2.414	1.792

5.2 SUS の評価結果

Immersion Neckwear のユーザビリティを評価するために、SUS によるアンケートを行った。表 4 は被験者ごとの SUS によるアンケートの回答結果をまとめたものである。また自由記述欄にて、「微妙な重さが少し気になった」、「デバイスが大きいと感じた」、「温度が人によっては暑すぎる」などの指摘がされた。一方で、「アニメに合わせて温度変化があり楽しかった」、「cooler と warmer 機能は実際に感じ

表 4: SUS の結果

被験者 ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	スコア
1	5	2	5	2	2	2	4	4	4	4	65
2	4	3	5	3	1	2	5	4	4	3	60
3	4	1	5	1	4	4	5	1	5	5	77.5
4	5	2	5	2	2	3	5	2	3	4	72.5
5	4	1	5	4	5	5	5	1	4	4	70
6	5	3	4	1	5	2	5	1	4	5	77.5
7	4	3	4	5	3	4	4	3	4	5	47.5
8	3	4	3	5	2	3	3	2	1	2	40
9	5	2	4	4	4	2	4	4	4	3	65
10	5	3	5	2	5	3	5	2	5	5	75
平均値	4.4	2.4	4.5	2.9	3.3	3	4.5	2.4	4	4	65
標準偏差	0.63	0.87	0.64	1.36	1.35	0.95	0.64	1.14	1.04	0.95	11.48

て、リアルタイム性も高いと思った」のような Immersion Neckwear の使用にポジティブな意見もあった。

5.3 SD 法の評価結果に関する考察

プロフィール分析の結果、Immersion Neckwear を装着した場合と装着しなかった場合の印象の違いがみられた。Immersion Neckwear を装着した場合、各形容詞対全体で肯定的な評価が確認された。これにより、Immersion Neckwear はユーザのポジティブな印象を高める効果があると推察する。よって、動画鑑賞における Immersion Neckwear の装着が、動画の印象を向上させる要因となることが考えられる。特に、「動的な-静的な」に関して、Immersion Neckwear の有無によって印象の変化が推察される。これは Immersion Neckwear が画面に応じて起動されることを、「動的」と捉えたと考える。表 5 と表 6 は、形容詞を使った印象の評価を分析した結果を示している。結果に基づき、Immersion Neckwear を使ったときの印象を整理した。

Immersion Neckwear を装着すると、因子 1 は「好きな」「良い」「迫力ある」といった言葉と強く結びつきことがわかる。これを「評価」とまとめた。因子 2 は「興奮した」「強い」「気持ちのいい」と関連が深かったため、「活力」とした。因子 3 は「好きな」「はっきりした」「リアリティのある」などの言葉と関係があったので、「現実性」とした。結果から、Immersion Neckwear を装着した動画鑑賞において、「評価」「活力」「現実性」の印象を持つことが考えられる。

表 5: システムありの時の、因子解釈

因子	形容詞	因子の解釈
1	「好きな」、「良い」、「迫力のある」	評価
2	「興奮した」、「強い」、「気持ちのいい」	活力
3	「はっきりした」、「リアリティのある」	現実性

Immersion Neckwear を装着しない場合、因子 1 は「迫力のある」、「良い」、「興奮した」、「強い」といった言葉と強く結びつきことがわかる。これを「活力」とまとめた。因子 2 は「はっきりした」「リアリティのある」と関連が深かったため、「現実性」とした。結果から、Immersion

Neckwear を装着した動画鑑賞において、「活力」「現実性」の印象を抱くことが考えられる。

表 6: システムなしの時の、因子解釈

因子	形容詞	因子の解釈
1	「迫力のある」、「良い」、「興奮した」、「強い」	活力
2	「はっきりした」、「リアリティのある」	現実性

因子の解釈から、従来の動画鑑賞での印象を包括しながらも「評価性」といった新たな肯定的印象が増えていることがわかる。よって、因子分析の結果として、Immersion Neckwear を用いた動画鑑賞は、Immersion Neckwear を使用しない場合よりも優れた鑑賞体験を提供することを推察する。また、寺本らの研究では、評価性が没入感の要素として抽出されており Immersion Neckwear を使った動画鑑賞がより没入感のある体験と考えられる[10]。

5.4 SUS の結果に関する考察

SUS によって評価されたシステム全体の平均値が 68 点であり、今回の SUS スコアが 65 点であることから、C グレードに相当する。よって、Immersion Neckwear のユーザビリティは改善の必要があると考える。質問 4 (技術的サポートの必要性)、質問 8 (操作性) では平均点が低く、回答にばらつきが見られた。実験では、被験者に動画を鑑賞することだけを指示し、システムの装着および起動は研究担当者が行ったため、一部の被験者はシステムの使用に技術的サポートが必要と感じたと考えられる。さらに、使用感に関しても一部の被験者から不満が出た。Immersion Neckwear の重さが 800g を超えており、これが動画鑑賞中の不快感につながり、システムの使用が面倒だと感じたと考えられる。以上の SUS に基づくアンケート結果から、システムのユーザビリティには改善の余地があると考えられる。ただし、「アニメに合わせた温度変化が楽しかった」、「cooler と warmer 機能がリアルタイムで感じられた」などのポジティブな意見もあり、Immersion Neckwear が動画鑑賞の没入感向上に有効であることが推察できる。

6. 結論

6.1 まとめ

本論文では、自宅での動画鑑賞体験を拡張することを目的として、多感覚フィードバックを提供する頸部装着型没入感向上デバイス Immersion Neckwear を創作した。Immersion Neckwear は、映像に合わせて振動、風、温冷感などを提示することで、映画館での没入感に近い体験を自宅でも再現することを目指している。実験結果から、Immersion Neckwear を装着した動画鑑賞では、被験者がよりポジティブな印象を抱き、動画の没入感が向上することが確認された。具体的には、SD 法によるアンケートの因子分析で、シ

システムの装着により評価性が高まることが確認された。一方、SUSによるユーザビリティ評価では、平均値が65点となり、技術的サポートの必要性およびシステムの重さ・大きさに関する不満が一部の被験者から指摘された。しかし、「アニメに合わせた温度変化が楽しかった」、「cooler と warmer 機能がリアルタイムで感じられた」といったポジティブな意見も多く、Immersion Neckwear が動画鑑賞の没入感向上に有効であることが示唆された。

6.2 今後の展望

今後の展望として、ユーザビリティの向上が必要である。SUS アンケート結果から示された技術的サポートの必要性およびシステムの重さ、大きさに関する不満を解消するため、ユーザインターフェースの改善、デバイスの軽量化、コンパクト化が求められる。具体的には、より直感的な操作が可能なインターフェースデザインを採用し、ユーザが簡単にシステムを操作可能にすることが挙げられる。また、使用する素材の見直しおよび電子部品の改良によって、デバイス全体の軽量化を図ることが考えられる。今回、映像に合わせて、風、振動、温冷部を実装したが、ミストや香りによるフィードバックの実装を目指す。これにより、ユーザが映像から感じる臨場感が向上すると考える。さらに、フィードバック機能の強化も重要な課題である。

Immersion Neckwear は、振動、風の強弱、温感、冷感といった多感覚のフィードバックを提供しているが、これらのフィードバックの精度、リアルタイム性の向上が求められる。特に、映像に合わせたフィードバックのタイミング、強度をより正確に制御するための技術開発が必要である。これらにより、ユーザにより一層没入感を感じさせることができると考えられる。また本研究では、20代の男女10名の被験者データのみ取得したが、より多くの被験者を対象にした大規模な実験を実施し、統計的に有意なデータを収集することで、Immersion Neckwear の効果をより明確に示す必要があると考えられる。また、異なる年齢層および文化背景を持つ被験者を対象にした研究を行うことで、ユーザビリティ、没入感に対する高度なシステムを実現したい。

参考文献

[1]株式会社ICT総研：2023年有料動画配信サービス利用動向に関する調査、<https://ictr.co.jp/report/20230421.html/>。(最終参照日: 2025/1/3).

[2]GEMPartners 株式会社：動画配信 (VOD) 市場 5 年間予測 (2024-2028 年) レポート、<https://gem-standard.com/columns/789/>。(最終参照日: 2025/1/5).

[3]株式会社プラネット：映画に関する意識調査、https://www.planet-van.co.jp/shiru/from_planet/vol210.html。(最終参照日: 2025/1/5).

[4]株式会社宣成社：映画離れる映画館、<https://senseisha.co.jp/useful/kiji.php?n=23>。(最終参照日: 2025/1/5).

[5]株式会社スパコロ：映画館の利用意識調査 ～コロナ禍、映画館利用意識どう変わった?～、<https://prtimes.jp/main/html/rd/p/000000041.000060722.html>。(最終参照日: 2025/1/5).

[6]杉原太郎, 森本一成, 黒川隆夫: SD 法を通してみた音楽に対する感性の基本特性, 映像情報メディア学会技術報告 25.48, 一般社団法人映像情報メディア学会, pp. 57-63(2001).

[7]仁科弘重, 中本有美: 観葉植物, 花, 香りが人間に及ぼす生理・心理的効果の脳波および SD 法による解析, 日本建築学会計画系論文集, Vol. 63, No. 509, pp. 71-75 (1998).

[8]寺本渉, 吉田和博, 浅井暢子, 日高聡太, 行場次朗, 鈴木陽一: 臨場感の素朴な理解 (i 特集) VR 心理学 4), 日本バーチャルリアリティ学会論文誌, Vol. 15, No. 1, pp.7-16 (2010).

[9]安藤広志: 2. 人が感じる臨場感の知覚認知メカニズムと評価技術, 映像情報メディア学会誌, Vol. 63, No. 12, pp. 1727-1730 (2009).

[10]飯村浩平, 中村広幸, 大倉典子, 小松剛: 臨場感と現実感の定量化と評価の実験, 日本人間工学会大会講演集, Vol. 48, No. 0, pp. 432-433 (2012).

[11]岡本早織, 羽田久一: ゲームプレイ中の風の強度による臨場感の変化, エンタテインメントコンピューティングシンポジウム 2022 論文集, pp. 135-138 (Aug 2022).

[12]前田智祐, 倉橋哲郎: ウェアラブルな温冷覚多点提示システム TherModule の基礎検討, 第 23 回日本バーチャルリアリティ学会大会論文集(Sep. 2018)

[13]Kim, A., Bae, H. and Lee, K.: Effects of Tactile Perception on Emotion and Immersion to Film Viewing in a Virtual Environment, VRST '19: Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology, pp.1-3 (Nov 2019)

[14]Oh, E., Lee, M. and Lee, S.: How 4D effects cause different types of presence experience?, VRCAI '11, Association for Computing Machinery, p. 375-378 (2011).

[15]Jeong, D., Han, S. H., Jeong, D. Y., Kwon, K. and Choi, S.: Investigating 4D movieaudiences' emotional responses to motion effects and empathy, Computers in HumanBehavior, Vol. 121, p. 106797 (2021).

[16]Lee, J., Han, B. and Choi, S.: Interactive motion effects design for a moving object in 4Dfilms, Association for Computing Machinery, p. 219-228 (2016).

[17]Nicolae, D. F.: Spectator perspectives in virtual reality cinematography. The witness, the hero and the impersonator, Ekphrasis. Images, Cinema, Theory, Media, Vol. 20, No. 2, pp.168-180 (2018).

[18]Kim, A., Chang, M., Choi, Y., Jeon, S. and Lee, K.: The Effect of Immersion on Emotional Responses to Film Viewing in a Virtual Environment, 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 601-602 (2018).

[19]株式会社秋月電子通商: 秋月電子通商, <https://akizukidenshi.com/catalog/default.asp>。(最終参照日: 2025/1/8).

[20]福田忠彦, 福田涼子, 福田忠彦研究室: 増補版 人間工学ガイド - 感性を科学する方法 -, 株式会社サイエンティスト社 (2009).

[21]Inc., Q.: System Usability Scale: What it is, Calculation + Usage, <https://www.questionpro.com/blog/system-usability-scale/>。(最終参照日: 2025/1/8).

[22]山内繁: 人を対象とする研究計画入門: 科学的合理性と倫理的妥当性, 丸善出版 (2015)

プレイヤーの感情に応じて適応するゲーム環境の試作と評価

嵯峨京介^{†1} 角薫^{†1}

公立はこだて未来大学^{†1}

1. はじめに

近年、ゲームのインタラクション技術は進化し続けており、プレイヤーの体験を向上させるための適応的なゲームシステムの開発が注目されている。特に、プレイヤーの感情や生理的反応に基づいてゲーム内容を動的に調整する手法は、没入感やプレイ満足度を向上させる可能性がある。Csikszentmihalyi (1990) が提唱する「フロー理論」[5]では、プレイヤーがゲームに対して適度な挑戦を感じることで没入体験の鍵となると述べられている。この理論をゲームデザインに応用した研究として、Chen (2007) は、ゲーム内での挑戦のバランスがプレイヤーの没入体験にどのように影響するかを分析した[4]。しかし、この理論が示唆する動的な挑戦のバランスは、従来のゲームにおける難易度設定では十分に考慮されていないのが現状である。

従来の難易度調整手法として、プレイヤーが事前に「簡単」「普通」「難しい」などのレベルを選択する方式が広く採用されている。しかし、この方式では、プレイヤーが自分に適した難易度を選択できない場合があり、一部のプレイヤーにとってはゲームが簡単すぎて退屈になり、逆に難しすぎて挫折してしまうといった問題が生じる[2]。また、ゲームジャンルの好みには個人差があり、プレイヤーが楽しめるジャンルを適切に選択できないケースもある[12]。こうした問題を解決するため、近年の研究ではプレイヤーの生体情報や表情、発話を活用し、リアルタイムで感情状態を推定し、それに基づいてゲームの難易度や内容を自動的に調整する手法が提案されている。

例えば、Hunicke ら (2005) は動的難易度調整 (Dynamic Difficulty Adjustment, DDA) の枠組みを提案し、ゲームプレイデータを活用した適応手法を示した [7]。また、Andrade ら (2006) は、生体情報を用いた DDA の有効性を示し、難易度調整がプレイヤーのモチベーションに与える影響を分析している[1]。さらに、Zook & Riedl (2012) は、強化学習を用いた難易度調整アルゴリズムを開発し、プレイヤーのスキル向上に適応する手法を提案した[11]。原田ら (2020) は、プレイヤーの音響的特徴量を分析し、ゲーム内の難易度をリアルタイムで調整する手法を提示している[13]。

また、機械学習を活用した難易度調整手法の研究も進められており、プレイヤーの行動パターンを学習し、適応的に難易度を調整するモデルが開発されている[1, 10]。さらに、プレイヤーの生理的データ (心拍数、皮膚伝導率など) を用いた感情推定手法が提案され、リアルタイムで適応的なゲームプレイを提供するシステムが検討されている[6, 9]。

ゲーム研究において、難易度調整に関する研究は広く行われているが、その多くはプレイヤーパフォーマンスデータ (スコアやクリア時間) に基づいた調整が主流である[10]。

一方、感情に基づいたゲーム調整の研究も進められており、表情情報と生体情報を組み合わせて感情の強さを推定する手法や、生体情報を用いたゲームの難易度自動調整シ

ステムが提案されている[16]。例えば、堀ら (1999) は、表情情報と生体情報を組み合わせた感情強度の推定手法を提案し、心拍数や皮膚伝導水準の変化が感情推定の有効な指標となることを示した[17]。佐藤と角 (2023) では、生体情報を用いたゲームの難易度自動調整システムが開発され、EDA (皮膚電気活動) および HRV (心拍変動) を活用することで、プレイヤーの感情推定の精度が向上することが示されている[9]。また、Sato ら (2024) はゲームプレイ中の心拍数の変動を分析し、男女間の生理的反応の違いを明らかにしている[16]。

感情推定技術を活用することで、プレイヤーの体験をよりパーソナライズする試みも行われている。例えば、大山ら (2022) は、アドベンチャーゲームにおけるユーザ体験向上のために、感情推定技術を用いた入力手法を開発した [14]。

本研究では、プレイヤーの感情に応じてゲームの難易度およびジャンルを動的に調整するシステムの試作と評価を行った。特に、既存研究が主に表情を用いた感情推定に依存していたのに対し、本研究では表情だけでなく心拍数の変動を分析することで、ゲーム調整の適切性を検討した。実験の結果、表情変化のみではプレイヤーの感情推定が困難であり、心拍数がジャンル変更や難易度調整に応じて変化することが明らかになった。

本研究の貢献は以下の3点である。

1. 表情のみの感情推定が不十分であることを実証し、心拍数の変化がより適切な調整指標となる可能性を示した。
2. ゲームジャンル変更が生理的興奮に影響を与えることを発見し、FPS ジャンルへの移行時に心拍数の有意な上昇が観測された。
3. プレイヤーの生理データを活用することで、より精度の高いゲーム適応手法の必要性を提唱した。

2. プレイヤーの感情に応じた動的ゲーム適応

2.1 システムの概要

本研究で開発したゲームは、プレイヤーの感情に応じて難易度および3種類のジャンル (アクション, FPS, アドベンチャー) が変化するものである。本ゲームは一人称視点 (図1) でプレイする形式であり、ジャンルの変化に伴い操作方法も若干異なる。初めは三人称視点で開発していたが、一人称視点の方が現実の視点に近く、没入感の向上に適していると判断したため、本研究では一人称視点を採用した。

ゲームの基本ルールを説明する。ゲーム開始後、ステージのクリア条件を満たすと、それまでのプレイヤーの感情とステージクリアタイムに基づき、難易度およびジャンルが調整され、次のステージが生成される。また、ステージ生成前には、プレイヤーにステージクリアタイムが表示される。

^{†1} Kyosuke Saga, Future University Hakodate

^{†2} Kaoru Sumi, Future University Hakodate

2.2 基本的なゲーム内容

本ゲームの基本的なゲーム内容について述べる。本ゲームでは、図 2 のようなステージと白いブロックがある地点からスタートし、そのブロックに触れることでステージが生成され、ゲームが開始される。ゲームは難易度：普通、ジャンル：アクションから始まり、ステージのクリア条件を満たすことで次のステージが生成される。

次に、ジャンルごとにゲームの内容とクリア条件を説明する。まず、アクションジャンルでは、ステージ開始時に、図 3 のように赤い敵が出現する。アクションジャンル中、プレイヤーは攻撃と回復を行うことができるため、敵の攻撃を受けないようにしながら、敵に攻撃を行い、全ての敵を倒すことでステージクリアとなる。

次に、FPS ジャンルでは、ステージ開始時に図 4 のように銃を持ったロボット型の敵が出現する。FPS ジャンル中、プレイヤーは射撃と回復を行うことができるため、敵の射撃を避けながら敵に射撃を当て、全ての敵を倒すことでステージクリアとなる。

最後に、アドベンチャージャンルでは、ステージ開始時に図 5 のように色と形の異なるオブジェクトが複数出現する。アドベンチャージャンル中は、攻撃や回復はできず、ステージ上に出現したアイテムを正しい順番で回収することでステージクリアとなる。正解の順番は 1 ステージごとに決まっており、正解・不正解に応じて○もしくは×のエフェクトが画面上に表示される (図 6, 図 7)。また、アイテムの回収順を間違えるとリセットされる。

ステージの生成時には、前のステージでのプレイヤーの感情およびゲーム内スコアに基づいて調整が適用される。これに加え、各ジャンルのクリア条件テキストが画面上に数秒間表示される (図 8, 図 9, 図 10)。また、ステージクリア時には、クリアタイムとステージ生成の通知テキストが画面上に表示された後に次のステージが生成される (図 11)。

2.3 難易度・ジャンル調整機能

本研究では、難易度調整を行ううえで、各感情に重みを付け、1 ステージ分の感情を数値化した。1 ステージ分の数値を合計し、そのステージの感情とみなすことで調整を行う。

数値が 5 以上の場合は、ポジティブな感情で楽しめていると判断し、調整は行わない。一方、数値が** -5 以下**の場合は、ネガティブな感情で楽しめていないとみなす。この場合、楽しめていない原因が「難易度が簡単すぎるため」なのか「難しすぎるため」なのかをクリアタイムを参照して判断する。また、数値が 5 より小さく、 -5 より大きい範囲で 3 ステージ連続した場合、プレイヤーが無関心または退屈していると判断し、ジャンル変更を行う。

次に、クリアタイムを用いた難易度調整について説明する。本研究では、難易度を「簡単」「普通」「難しい」の 3 段階に設定した。各難易度の調整は、以下のクリアタイムの基準に基づいて行う。クリアタイムが 20 秒以下であれば簡単

すぎると判断し、現在の難易度より 1 段階上げる。クリアタイムが 80 秒以上であれば難しすぎると判断し、現在の難易度より 1 段階下げる。20 秒から 80 秒の範囲であれば適切な難易度とみなし、ジャンル変更を行うことでプレイ体験の向上を図る。クリアタイムの基準については、テストプレイの結果、これらの閾値でプレイスキルの評価が適切に行えると判断したため本研究ではこの基準を採用している。

ジャンル調整については、プレイヤーの感情やクリアタイムの傾向を分析し、ジャンル変更が必要と判断された場合に実施する。ジャンルの変更はアクション→アドベンチャー→FPS の順に切り替える方式とし、このサイクルを繰り返す。プレイヤーが特定のジャンルで退屈していると判断された場合、次のジャンルに移行することで新たな体験を提供し、没入感を維持することを目的としている。

3. 実験

本研究では、プレイヤーの感情に基づいた動的難易度調整およびジャンル変更がゲーム体験に与える影響を検証するため、被験者実験を実施した。動的にゲームを適応させることで、プレイヤーの生理的・心理的反応にどのような違いが生じるかを評価する。

本研究では以下の仮説を検証する。

- **H1: 動的難易度調整はプレイヤーの生理的興奮を引き起こす。**
- **H2: ジャンル変更はゲーム体験に影響を与える。**
- **H3: 動的難易度調整はプレイヤーの主観的満足度を向上させる。**

本研究では、プレイヤーの感情に基づいた動的難易度調整およびジャンル変更がゲーム体験に与える影響を検証するため、被験者実験を実施した。動的にゲームを適応させることで、プレイヤーの生理的・心理的反応にどのような違いが生じるかを評価する。

3.1 実験概要

被験者には 2 種類のゲームをプレイしてもらい、難易度とジャンルが自動調整される条件 (自動調整群) と固定された条件 (無調整群) を比較した。プレイヤーの生理データや主観的評価を取得し、動的難易度調整の効果を分析する。

3.2 被験者と実験環境

本実験には 18 歳から 24 歳の大学生 26 名 (男性 21 名、女性 5 名) が参加した。各被験者は無作為に自動調整群 ($n=14$) または無調整群 ($n=12$) に割り当てられた。実験はゲームプレイ時の生理データを正確に測定するため、心拍センサーを装着した。

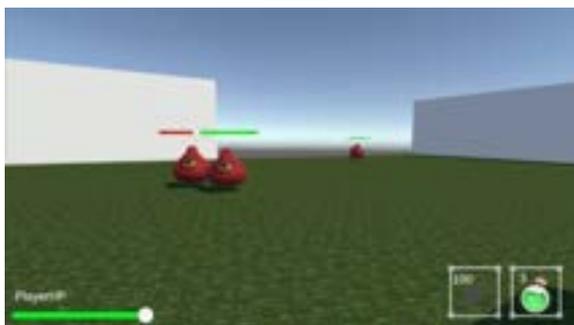


図1 プレイヤー視点の画面

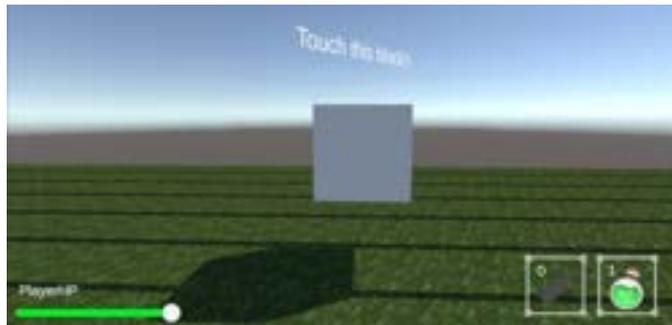


図2 ステージ開始時のブロック表示

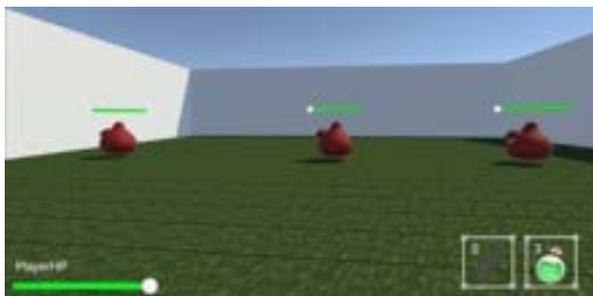


図3 アクションゲームのプレイ画面



図4 FPSゲームのプレイ画面

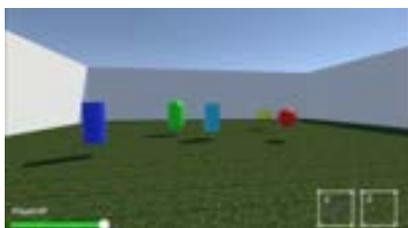


図5 アドベンチャーゲームのプレイ画面



図6 アドベンチャーゲームの正解画面



図7 アドベンチャーゲームの不正解画面



図8 アクションゲームのステージ生成時の説明テキスト



図9 FPSゲームのステージ生成時の説明テキスト



図10 アドベンチャーゲームのステージ生成時の説明テキスト



図11 ステージクリア時の画面

3.3 実験手順

実験の手順は以下の通りである。

1. **事前アンケート**: 被験者の年齢・性別・ゲーム経験を記入してもらう。
2. **準備**: 生体センサーを装着し、3分間の操作確認を実施する
3. **本実験**: 15分間ゲームをプレイし、心拍データおよびプレイログを記録する。
4. **事後アンケート**: ゲーム体験に関する主観評価を収集する。

3.4 評価指標

本研究では以下の指標を用いた。

- **生理データ**: ゲームプレイ中の心拍数 (BPM)
- **ゲームプレイデータ**: ステージクリア時間、難易度変化回数
- **主観評価**: Game Engagement Questionnaire (GEQ) [3, 8] に基づくゲーム没入度

4. 結果

本章では、実験で収集したデータを分析し、動的難易度調整とジャンル変更の効果を検討する。

4.1 生理データの分析

難易度上昇前後の心拍数を比較したところ、難易度上昇後に心拍数が有意に増加した ($p < 0.05$, t 検定)。特に、FPS ジャンルへの移行時に心拍数の有意な上昇が見られた (平均 +10 BPM, $p < 0.01$)。一方で、アクション→アドベンチャーへの移行では有意な変動は見られなかった。

4.2 主観評価

GEQ のスコア分析の結果、自動調整群の方が「挑戦的」と感じる傾向が強かった ($p < 0.05$)。しかし、「満足度」や「楽しさ」に関しては有意な差が見られなかった。

4.3 仮説の検証

本研究では以下の仮説を検証した。

- **H1: 動的難易度調整はプレイヤーの生理的興奮を引き起こす。**
→ 心拍数の有意な上昇が見られたため、支持された。
- **H2: ジャンル変更はゲーム体験に影響を与える。**
→ FPS ジャンルへの移行時に心拍数の上昇が確認されたため、部分的に支持された。
- **H3: 動的難易度調整はプレイヤーの主観的満足度を向上させる。**
→ 「満足度」に有意な差が見られなかったため、支持されなかった。

本結果から、動的難易度調整は生理的な興奮度を高めるが、主観的な満足度の向上には直結しない可能性が示唆された。

4.4 自由回答結果

本研究では、被験者からの自由回答を分析し、動的難易度調整がゲーム体験に与える影響を詳細に検討した。自由回

答の分析結果を表 1 に示す。

回答を分析した結果、動的難易度調整の効果を実感した被験者は、敵の数や速度の変化を適切な難易度調整と認識していた。一方で、調整がうまく機能しなかったケースでは、単調さや不満が報告され、特に表情認識の精度に課題がある可能性が示唆された。

本結果から、難易度調整のアルゴリズムの改善や、表情以外の生理データを活用した補正が今後の研究課題として挙げられる。

表 1 自由回答

ID	調整認識	回答結果
1-1	はい	三人称視点でやりたい
5-1	いいえ	敵の配置と数がそのままなので、途中で飽きてしまった。敵モンスターのグラフィックはいいと思った。
9-1	はい	自分は一人称視点のゲームをあまりやったことがなかったため、最初は新鮮な感じで楽しくプレイできた。しかし、敵が 5 体が増えたところで、そこからあまり代わり映えがなく、ひたすら 5 体の敵を倒すだけだったため、単純作業のようになり、飽きてきてしまった。そこで、自分でなにか変化をつけようと思って、ジャンプしたり、真ん中の敵から倒しはじめたりしたが、後半は慣れてきて、退屈に感じた。
13-1	はい	5 体になった時から難易度が上がったと感じました。
16-1	いいえ	システムの影響なのか、ずっとスライムが出てきていて、延々とスライムを倒し続けた。途中スライムの数が増えた。一番最後にゲームジャンルが変わったのでそれだけめっちゃ楽しかった。ゲームプレイはほとんどつまずくことが無く快適に遊べた。敵が追いかけてくるのも緊張感があって楽しかった。

4.5 ゲーム中の心拍数変動

実験中に取得した心拍数データを分析し、難易度調整およびジャンル変更が生理的反応に及ぼす影響を検討した。まず、被験者の中から無作為に選択した 1 名について、難易度上昇時 5 秒前後の心拍数変動を分析した。結果を表 2 に示す。

表 2 難易度上昇時の心拍数平均値

状況	変化前(BPM)	変化後(BPM)	P 値
難易度上昇	75	130	0.029

次に、ジャンル変更が多かった被験者について、ジャンル変化前後 5 秒前後の心拍数変動を分析した。特に、アドベンチャージャンルから FPS ジャンルへの移行時に有意な心拍数上昇が確認された ($p < 0.05$, 表 3)。

表 3 ジャンル変更時の心拍数の平均値

変化前(BPM)	変化後(BPM)
71.6	81.0
73.0	80.4
74.0	84.8

4.6 実験後アンケートの結果

被験者に対して GEQ に基づくアンケートを実施し、主観的評価を取得した。質問内容は表 4 に示す。表 4 に関しては、Q1~Q33 までのオレンジ部分のゲーム没入度尺度 Game

Engagement Questionnaire(GEQ) [3, 8]にいくつかの質問を加えたものである。

表 4 実験後アンケートの質問項目

番号	質問内容
Q1	ゲームのストーリーは興味深かった。
Q2	プレイ中、ゲームのことで頭がいっぱいだった。
Q3	プレイしたことによって機嫌が悪くなった。
Q4	プレイしたことによって満足感を覚えた。
Q5	ゲームのグラフィックは美しく魅力的だった。
Q6	周りのことはすべて忘れて、プレイに没頭してしまった。
Q7	プレイ中、ゲームに関係のない他のことを考えていた。
Q8	ゲームの内容は楽しいものであった。
Q9	プレイ中、想像力に満ちた気持ちになった。
Q10	プレイ中、時間の感覚を失っていた。
Q11	内容が単調で、退屈だった。
Q12	プレイによって幸せな気持ちになった。
Q13	ゲームの内容について、もっと探究できると思った。
Q14	プレイ中、ゲームに深く集中していた。
Q15	内容に興味が湧かず、退屈だった。
Q16	とても良いゲームだと思った。
Q17	感動的なプレイ体験であった。
Q18	現実世界とのつながりを失ったように感じた。
Q19	プレイ体験は楽しいものであった。
Q20	このゲームプレイは、豊かな経験となった。
Q21	このゲームに熟達していると思った。
Q22	ゲームの内容に軽度の苛立ちや不快感を覚えた。
Q23	ゲームが難しいと感じた。
Q24	実力に自信を持ってプレイしていた。
Q25	長時間、イライラしながらプレイしていた。
Q26	緊張感を覚えながらプレイしていた。
Q27	得意なゲーム (ジャンル) だった。
Q28	思うようにプレイできず、イライラした。
Q29	ゲームに対して挑戦的な気持ちになった。
Q30	プレイ内容に関して、目標を達成できたと考えた。
Q31	プレイ中、時間に追われている感覚に陥った。
Q32	プレイ内容に関して、早く目標を達成できたと考えた。
Q33	プレイのために多くの労力を注いだ。
Q34	ゲームプレイ中飽きを感じた。
Q35	ゲームプレイ中ゲームに関係のないことを考えた。
Q36	ゲームプレイ中楽しいと感じた。
Q37	ゲームの最初より最後の方が楽しいと感じた。
Q38	ゲームの最後より最初の方が楽しいと感じた。
Q39	ステージ生成されたとき、生成されたことに気づきましたか？
Q40	ゲーム中、ジャンルと難易度が調整されていることに気づきましたか？
Q41	ゲームをプレイしてみて感じたこと、感想等あれば何でも教えていただくと助かります。(自由回答)

アンケートの回答形式は 5 段階評価 と 二択形式 の 2 種類である。Q1~Q38 は「まったくそう思わない (0)」から「とてもそう思う (4)」までの 5 段階評価 で、主観的体験の強度を数値化できる。一方、Q39・Q40 は「はい」「いいえ」の二択形式 で、ゲーム調整の認識度を測定した。

4.6.1 結果の分析

アンケート結果の統計分析を行い、自動調整群と無調整群のスコアを比較した。特に **Q29 (ゲームに対して挑戦的な気持ちになった)** に関しては、自動調整群のスコアが有意に高く ($p<0.05$)、動的難易度調整がプレイヤーの挑戦意欲を高める可能性が示唆された。一方で、**Q4 (プレイしたことによる満足感)** や **Q19 (プレイ体験が楽しいものであった)** については、両群の間に有意な差は見られなかった。

また、Q22 (ゲームの内容に軽度の苛立ちや不快感を覚えた) においては、自動調整群の方がわずかに高いスコアを示した。このことから、一部のプレイヤーにとっては、難

易度やジャンルの変化がストレス要因となる可能性も考えられる。

本アンケート結果から、動的難易度調整はプレイヤーの挑戦意欲を向上させる効果はあるものの、満足度や楽しさの向上には必ずしもつながらないことが示唆された。特に、難易度やジャンルの変化が意図しない方向に適応した場合、プレイヤーの不満につながる可能性がある。

また、ゲームの難易度やジャンルの適応はプレイヤーの没入度に影響を与えるが、それが必ずしもポジティブな体験につながるとは限らないことが示された。今後の研究では、プレイヤーごとの個別適応を強化し、ストレス要因を最小限に抑える手法の導入が求められる。

5. 考察

本研究では、プレイヤーの感情に基づく動的難易度調整およびジャンル変更がゲーム体験に与える影響を検証した。結果として、動的難易度調整がプレイヤーの生理的興奮 (心拍数上昇) を引き起こすことが確認されたものの、主観的満足度や楽しさの向上には必ずしも結びつかないことが示された。本章では、得られた知見について考察し、仮説の検証結果を整理する。

5.1 動的難易度調整の影響

本研究の結果から、動的難易度調整はプレイヤーの生理的興奮を引き起こすことが確認された ($p<0.05$)。特に、難易度が上昇した際の心拍数の増加が有意であったことから、ゲームの難易度調整がプレイヤーの緊張や集中度に影響を及ぼしている可能性が示唆される。

しかし、主観評価では、「満足度」や「楽しさ」に関する評価に有意な差は見られなかった。この結果は、単に難易度を調整するだけでは、プレイヤーのゲーム体験の向上には不十分であることを示している。調整の適用が意図しない方向に進んだ場合、プレイヤーの満足度が低下する可能性があるため、プレイヤーのスキルやプレイスタイルに応じた柔軟な適応が求められる。

5.2 ジャンル変更の影響

ジャンル変更が生理的興奮に与える影響については、FPS ジャンルへの移行時にのみ心拍数の有意な上昇が確認された。一方で、アクション→アドベンチャーへの移行では有意な変化が見られなかった。この結果から、ジャンル変更の影響はプレイヤーの好みによって異なる可能性が示唆される。

また、自由回答の分析では、アクションや FPS の調整は適切であると評価された一方、アドベンチャージャンルでは試行錯誤が難しく、ストレス要因になったという意見が見られた。これは、単にジャンルを変更するのではなく、プレイヤーのスキルや好みに応じて適応させる必要があることを示している。

5.3 感情推定手法の課題

自由回答結果の回答を確認したところ、今回作成した自動本研究では、プレイヤーの表情情報を活用して感情推定を行い、動的な適応を実施した。しかし、自由回答の分析から、表情のみではプレイヤーの感情を正確に推定することが困難であることが示唆された。特に、一人でゲームをプレイする際には表情の変化が少なく、感情を正しく読み取ることができない場合があった。

この結果から、表情情報だけでなく、心拍数や皮膚電気活動 (EDA) などの生理的データを組み合わせることで、より高精度な感情推定が可能になると考えられる。今後の研究では、これらのマルチモーダルデータを統合し、より精度の高い感情推定手法の開発が求められる。

5.4 仮説の検証

本研究で設定した仮説の検証結果は以下の通りである。

仮説 H1 (動的難易度調整はプレイヤーの生理的興奮を引き起こす) に関しては、実験の結果、難易度上昇時に心拍数の有意な増加 ($p<0.05$) が確認され、支持されることが示された。これは、難易度の変化がプレイヤーの生理的な興奮度や緊張感を高める要因となり得ることを示唆している。

仮説 H2 (ジャンル変更はゲーム体験に影響を与える) については、FPS ジャンルへの移行時にのみ心拍数の有意な上昇 ($p<0.01$) が確認され、部分的に支持された。一方で、アクション→アドベンチャーへの変更時には有意な変化が見られず、ジャンルの違いがプレイヤーの生理的反応に与える影響は一定ではないことが示唆された。

仮説 H3 (動的難易度調整はプレイヤーの主観的満足度を向上させる) に関しては、アンケート結果の分析により、「満足度」や「楽しさ」に関するスコアには有意な差は見られなかったため、不支持となった。この結果から、難易度やジャンルの調整が行われたとしても、それがプレイヤーのゲーム体験の向上に直結するわけではなく、適切な調整方法やプレイヤーの好みに応じた適応戦略が必要であることが示された。

6. 今後の展望

本研究の結果を踏まえ、今後の研究において検討すべき課題を以下に挙げる。

6.1 感情推定の精度向上

表情認識のみでは感情を正確に推定することが難しいため、心拍数、皮膚電気活動 (EDA)、脳波 (EEG) などの生理データを統合したマルチモーダル感情推定モデルの開発が求められる。

6.2 プレイヤーモデルの構築

プレイヤーの長期的なデータを蓄積し、個別の適応戦略を学習するプレイヤーモデルを構築することで、より柔軟な動的適応が可能になる。

6.3 ジャンル変更の最適化

ジャンルの選択肢を増やし、プレイヤーの好みやスキルに応じた適応ができるようにすることで、より快適なゲーム体験を提供できると考えられる。

6.4 ストレスや不快感の低減

一部の被験者から、難易度やジャンルの変化が意図しない形で適応されるとストレスを感じるという意見があった。調整の遷移を自然に行う手法や、プレイヤーが適応の選択権を持つシステムの導入が求められる。

7. 結論

本研究では、プレイヤーの感情に応じた動的難易度調整およびジャンル変更がゲーム体験に与える影響を検証するため、被験者実験を実施した。その結果、以下の主要な知見が得られた。

- 動的難易度調整はプレイヤーの生理的興奮を引き起こすが、主観的満足度の向上には直結しない。
 - 難易度上昇時に心拍数が有意に上昇 ($p<0.05$) し、生理的興奮の増加が確認された。
 - しかし、アンケート結果では「満足度」や「楽しさ」に有意な差は見られなかった。
- ジャンル変更の影響はジャンルによって異なり、FPS への移行時に最も顕著な生理的反応が見られた。
 - FPS への移行時に心拍数が有意に上昇 ($p<0.01$) したが、アクション→アドベンチャーでは変化が見られなかった。
 - 自由回答では、アドベンチャージャンルに対するストレスや不快感の報告が多かった。
- 表情情報のみを用いた感情推定には限界があり、より高精度な推定手法が必要である。
 - 一人プレイ時には表情の変化が少なく、感情推定の精度が低下する可能性が示唆された。
 - 今後は、生理的データ (心拍数・EDA・EEG など) を統合したマルチモーダル推定が求められる。
- プレイヤーの長期的なデータを活用し、個別適応型のゲームシステムを構築する必要がある。
 - 機械学習を活用しプレイヤーのプレイ履歴やスキルレベルを考慮した適応戦略の開発が望まれる。

本研究の結果から、動的難易度調整とジャンル変更はプレイヤーの生理的反応を変化させることが明らかになったが、ゲーム体験の向上には、より精度の高い感情推定手法とプレイヤーごとの適応戦略が必要であることが示唆された。今後の研究では、プレイヤーのストレスを最小限に抑えつつ、個々に最適なゲーム体験を提供する適応システムの開発が求められる。

参考文献

- [1] Andrade, G., Ramalho, G., Santana, H., & Corruble, V. (2006). Challenge-sensitive dynamic difficulty adjustment in computer games. *AIIDE 2006: Proceedings of the Second Artificial Intelligence and Interactive Digital Entertainment Conference*, 47–52.
- [2] Bateman, C., & Boon, R. (2006). *21st Century Game Design*. Charles River Media.
- [3] Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Applied Psychophysiology and Biofeedback*, 34(4), 247–261. <https://doi.org/10.1007/s10484-009-9074-z>
- [4] Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, 50(4), 31–34.
- [5] Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper Perennial.
- [6] Dennis, S. A., Keiser, V., Kunkel, A., & McNamara, B. (2020). Attention, engagement, and affect detection using physiological signals in virtual reality gameplay. *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2020)*, 592–600.
- [7] Hunnicke, R. (2005). The case for dynamic difficulty adjustment in games. *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE '05)*, 429–433.
- [8] IJsselstein, W., de Kort, Y., & Poels, K. (2013). *The Game Experience Questionnaire*. Manuscript from Eindhoven University of Technology. Retrieved from https://pure.tue.nl/ws/portalfiles/portal/21666907/Game_Experience_Questionnaire_English.pdf
- [9] Sato, Y., Salem, A., & Sumi, K. (2024, July). Physiological Correlates while Playing Tetris: A Comparative Study between Males and Females. *Proceedings of the 1st IEEE International Workshop on Interactive Systems, Affective Computing Technologies and Applications (ISAC 2024)*, co-located with the 48th IEEE International Conference on Computers, Software, and Applications (COMPSAC 2024), Osaka, Japan.
- [10] Yannakakis, G. N., & Hallam, J. (2011). Real-time game adaptation for optimizing player satisfaction. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(3), 200–212.
- [11] Zook, A., & Riedl, M. O. (2012). A temporal data-driven player model for dynamic difficulty adjustment. *Proceedings of the 8th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2012)*, 93–98.
- [12] 小川聖矢, 遠藤雅伸. (2024). ゲームジャンル別好き嫌いの男女差に関する調査分析. 研究コラボレーションとネットワークサービス (CN), 15(1), 1-3.
- [13] 原田和明, 兼松祥央, 茂木龍太, 三上浩司. (2020). プレイヤーのリアクションの音響的特徴量に基づく動的難易度調整. *エンタテインメントコンピューティングシンポジウム (EC2020)*, 8-11.
- [14] 大山智弘, 川口一画. (2022). アドベンチャーゲームにおけるユーザ体験向上のための感情推定技術を用いた入力手法. *情報処理学会研究報告(Web), HCI-199(4)*, 1-8.
- [15] 三谷慧, 三上浩司. (2023). ノベルゲームにおけるテキストマイニングと心拍データを用いた感情推定. *日本デジタルゲーム学会年次大会予稿集*, 13(0), 268-270.
- [16] 佐藤優介, 角薫. (2023). 生体情報を用いたレベルを自動調整するゲームの開発. *2023年度人工知能学会全国大会(第37回)*, 2T6-GS-9-05.
- [17] 堀政重, 小関修, 横山清子, 渡辺與作. (1999). 表情と生体による感情の強さの推定—面白い・楽しいの場合—. *電気学会論文誌C (電子・情報・システム部門誌)*, 119(6), 668-675.

動画視聴における 瞬目促進フィードバック手法の検証

新村温人^{†1}

青山学院大学^{†1}

ロペズギヨーム^{†2}

青山学院大学^{†2}

1. はじめに

近年、日本における動画配信市場は急速に拡大している。総務省の調査によると、2019年の動画配信市場規模は2,959億円であったのに対し、2023年には5,704億円に達し、4年間で96.2%の増加を示した[1]。さらに、2028年には7,371億円規模に拡大することが予測されており、今後も動画視聴の機会が増加することが見込まれている。このような市場の成長は、動画コンテンツの多様化やストリーミング技術の進歩に支えられているが、一方で、長時間の動画視聴が人々の健康に与える影響が懸念されている。特に、視覚疲労やドライアイの発症リスクの増加が問題視されており、適切な視聴習慣の確立が重要な課題となっている。

VDT (Visual Display Terminal) 作業に関連する健康問題は、長時間のディスプレイ使用により引き起こされる身体的影響の一つとして注目されている。厚生労働省の調査によれば、VDT 作業を行う労働者のうち68.6%が身体的疲労を感じており、その中でも「目の疲れや痛み」を訴える割合は90.8%に達している[2]。また、第一三共ヘルスケア株式会社の調査では、テレワーク中の男女において、眼精疲労を感じる割合が女性で22.0%、男性で15.7%であることが報告されている[3]。これらの結果は、長時間のデジタルデバイス使用が視覚健康に悪影響を及ぼしていることを示唆している。

視覚健康の維持には、涙の安定した分泌が重要な役割を果たしている。涙は、目の表面に栄養を供給し、酸化や炎症を防ぐ働きを持つ[4]。また、涙は外部からの異物を洗い流し、目の表面を保護する機能を果たしている。瞬目（まばたき）は、涙を目全体に均等に分配するために不可欠な動作であり、これが不足すると目の乾燥が進み、ドライアイの発症リスクが高まる。しかし、近年の研究では、デジタルデバイスの長時間使用により瞬目の回数が減少し、それに伴い涙液膜の安定性が損なわれることが指摘されている[5]。デジタルデバイス使用時の瞬目頻度の低下は、特に深刻な問題となる。通常、瞬目は1分間に18~22回程度行われるが、パソコンやスマートフォンを長時間使用すると、瞬目回数は1分間に5~7回まで減少することが確認されて

いる。この瞬目減少によって涙液膜の破壊が早まり、目の乾燥や眼精疲労の原因となる。特に、オフィスワーカーやコンタクトレンズ使用者では、涙の膜が5秒以内に乾燥してしまう特徴がある「BUT 短縮型ドライアイ」と呼ばれる症状が報告されている[6]。ドライアイは、日本国内で約2,200万人の患者がいるとされており、年々増加傾向にある。このような現状を踏まえると、デジタルデバイスの使用環境に適した視覚健康対策が求められている。

本研究では、動画視聴中の瞬目減少による健康問題の改善を目的とし、瞬目を促進するシステムの開発を行う。本システムでは、リアルタイムで瞬目の回数を検出し、瞬目の回数が基準値を下回った場合に適切なフィードバックを提供することで、意識的に瞬目を促進することを目指す。具体的には、視覚（フラッシュ、アニメーション GIF）、聴覚（短音通知）、触覚（スマートウォッチの振動）を活用した4種類のフィードバック手法を実装し、それぞれの手法の有効性を比較・評価する。これにより、視聴体験を妨げることなく瞬目を促進する最適な手法を明らかにすることを目的とする。

2. 関連研究

本章では、瞬目促進や目の健康、注意喚起に関する先行研究を紹介する。東覚らは、長時間のパソコン作業によるドライアイ軽減を目的とし、瞬目回数を検知して適切なタイミングでリマインドを行うシステムを提案した[7]。本システムは、カメラと MediaPipe ライブラリを活用し、上下まぶた間の距離を基に瞬目を検出する。比較実験では、「10秒間に瞬目回数が2回以下の場合にリマインド」を行う条件が最も効果的であり、瞬目回数の増加や不完全な瞬目の減少が確認された。さらに、長時間の使用においても瞬目の安定化が示され、被験者の評価も高かった。一方でリマインド方法が邪魔だと感じるという意見があり、他の感覚モダリティ（聴覚や触覚など）を用いた比較や評価は行われていない。瞬目の質の改善にはさらなる検証が必要であることが指摘されている。

また、デジタルデバイスの使用による目の健康への影響についても研究が行われている。Irina らは、長時間の画面使用による視覚的・筋骨格的な症状や行動変化を指す「コンピュータビジョン症候群 (CVS)」に着目し、視力低下や目の疲れ、頭痛、首や肩の痛み、注意力の低下といった症状を報告した[8]。特に、デバイスの不適切な使用や照明条

Verification of a Feedback Method for Promoting Blinking Eyes in Video Viewing

†1 HARUTO NIIMURA, Aoyama Gakuin University

†2 GUILLAUME LOPEZ, Aoyama Gakuin University

件、長時間の作業、乾燥した環境が主な要因として挙げられており、女性や子ども、スマートフォンの過剰使用者にリスクが高いことが示された。予防策としては、20 分ごとに作業を中断し、20 秒間 6 メートル以上離れた場所を見る「20-20-20 ルール」や、まぶたを閉じて数秒間静止し、その後強く閉じる点滅運動が有効であるとされている。

さらに、瞬目促進に関連する注意喚起の手法についても研究が進められている。Xu らは、視線移動に基づく触覚フィードバック (eyerofeedback) の効果を調査し、触覚刺激によって被験者が自身の視線挙動や注意状態を自覚しやすくなり、特に注意維持が難しい長時間タスクでパフォーマンスが向上することを明らかにした[9]。この研究は、触覚刺激が他の感覚モダリティを補完し、注意喚起を促進する有効な手段となり得ることを示している。

これらの先行研究では、感覚刺激を利用した瞬目促進や注意喚起の効果が多く示されており、それぞれの感覚モダリティが瞬目行動や視覚健康に与える影響が明らかになっている。しかし、異なる感覚モダリティを比較検討する研究は限られており、それぞれの手法の有効性を包括的に評価する必要がある。本研究では、視覚・聴覚・触覚を活用したフィードバック手法を通じて、瞬目促進における最適な刺激方法を明らかにすることを目的とする。本研究の成果により、瞬目不足が引き起こす眼精疲労やドライアイの軽減に寄与し、日常生活に適用可能なシステム設計への新たな知見を提供することを目指している。

3. 瞬目促進システムの概要

本研究では、動画視聴中の瞬目回数が減少することによるドライアイのリスクを低減し、視覚の健康を維持することを目的とした瞬目促進システムを開発した。本システムは、Python を用いた瞬目検出アプリケーションを基盤とし、リアルタイムで瞬目を計測し、適切なフィードバックを提供することで瞬目を促す。以下に、本システムの構成や実装の詳細について述べる。

3.1 瞬目促進システムの概要及び提案手法

本システムは、動画視聴時の瞬目回数を計測し、基準値を下回った場合にフィードバックを提示することで瞬目を促進する。具体的には、Python を用いて開発したアプリケーションが、カメラを通じて瞬目を検出し、視覚・聴覚・触覚を用いたフィードバックを行うことで、ユーザの瞬目行動を誘発する。本研究では、瞬目促進の手法として、画面フラッシュ、アニメーション GIF、短音通知、スマートウォッチの振動の 4 種類を実装し、それぞれの効果を比較することを目的とする。

3.2 瞬目促進システムのハードウェアの構成

本システムの開発および実験には、以下のハードウェア

を使用した。

(1) 瞬目検出用デバイス

実験環境では、NEC 社製の「LAVIE Direct PM」ノート PC を使用する[10]。本デバイスは、第 11 世代の Intel Core i7-1165G7 プロセッサ(基本クロック周波数 2.80 GHz)、16GB メモリ、512GB SSD ストレージを備えており、軽量で持ち運びやすい仕様となっている。本 PC は、瞬目検出アプリケーションの実行およびデータ処理を担う。

(2) フィードバック提示デバイス

フィードバックの一部として、Google 社の「Google Pixel Watch」を使用する[11]。本スマートウォッチは、GPS、心拍計、3 軸加速度計、3 軸角速度計、3 軸磁力計など多様なセンサを搭載しており、フィードバック提示のためのデバイスとして活用される。特に、瞬目回数が基準値を下回った場合に振動を発生させることで、触覚フィードバックを提供する役割を果たす。

3.3 瞬目促進システムのソフトウェアの構成

本システムのソフトウェアは、以下の 3 つの主要機能で構成される。

(1) 瞬目検出機能

瞬目検出には、Google の MediaPipe ライブラリ[12]を活用した。本ライブラリの顔のランドマーク検出機能を使用し、目の形状から眼の開閉状態をリアルタイムで検出する。瞬目の識別には Eye Aspect Ratio (EAR) [13]を用いる。EAR は、目のランドマーク (特徴点) の位置関係を基に算出される指標であり、目が開いている間は一定の値を保ち、閉じると 0 に近づく。本システムでは、事前実験により最適な EAR の閾値を設定し、これを基準に瞬目を検出する。EAR の計算結果例を図 1 に示す。



図 1 : EAR の計算結果例

(2) フィードバック提示機能

瞬目回数が基準値を下回った際に、以下の 4 種類のフィードバックを提供する。

- 視覚的フィードバック (フラッシュ): 図 2 のように、画面全体を一瞬白くすることで強い光刺激を与え、瞬目を誘発する。



図 2 : フラッシュの実装例

- 視覚的フィードバック (アニメーション GIF) : 図 3 のように、目の瞬目動作を模倣した GIF アニメーションを表示し、模倣行動を誘発することで無意識的な瞬目を促す。

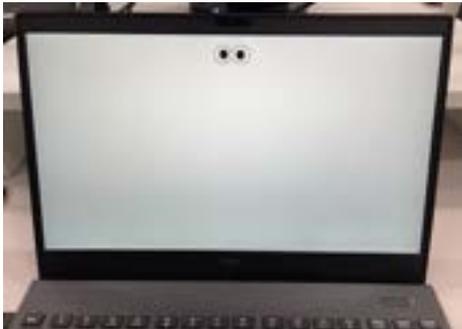


図 3 : アニメーション GIF の実装例

- 聴覚的フィードバック : 短い警告音 (「ピピッ」という短音) を再生し、瞬目不足への意識をたかめるとともに反射的な瞬目を促す。
- 触覚的フィードバック : Google Pixel Watch の振動機能を用いて、ユーザーに瞬目を促す刺激を与える。また、図 4 のように GIF を用いた視覚刺激と合わせて振動と画面表示を連動させることで、より瞬目を意識させる設計とした。



図 4 : スマートウォッチの振動・GIF 表示の実装例

(3) データ記録機能

実験データの分析のために、瞬目回数やフィードバックの提示状況を記録する機能を実装した。データは 10 秒ごとに保存され、後の統計分析に活用される。

3.4 システムの利用方法

本システムの利用手順は以下の通りである。

1. ノート PC 上で瞬目検出アプリケーションを起動する。
2. スマートウォッチを装着し、専用アプリを起動する。
3. ユーザーが動画を視聴し始めると、瞬目計測が開始される。
4. 視聴中、瞬目回数が基準値を下回った場合、フィードバックが自動的に提示される。
5. 視聴終了後、データを保存し、分析のためのログを出力する。

このシステムにより、ユーザーは自然な形で瞬目を促されることが期待される。このシステムのフロー図を図 5 に示す。

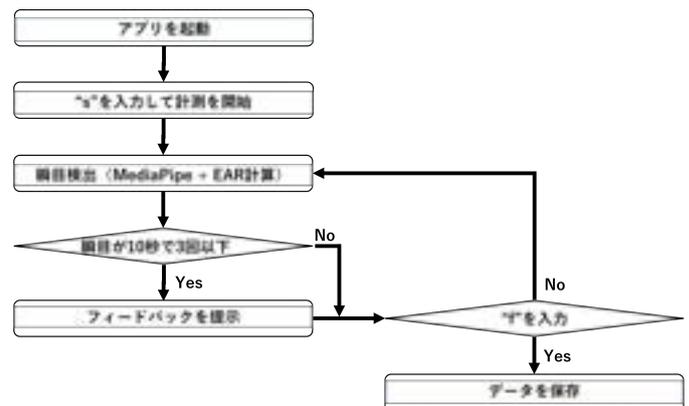


図 5 : システムのフロー図

今後、本システムを活用した実験を通じて、瞬目促進効果を定量的に検証し、最適なフィードバック方法を探る。

4. 瞬目促進システムの効果検証実験

本章では、開発した瞬目促進システムの有効性を検証するために実施した評価実験について述べる。本実験では、瞬目検出精度の最適化を目的とした予備実験と、フィードバック条件の有効性を検証する本実験の 2 つの段階で構成される。以下に、それぞれの実験方法、手順、分析手法について詳述する。

4.1 実験目的

本研究の目的は、動画視聴中の瞬目回数を増加させることによって、ドライアイや眼精疲労のリスクを低減することである。そのため、開発した瞬目促進システムの有効性を検証し、最適なフィードバック手法を明らかにする必要がある。本実験では、瞬目検出精度の最適化と、異なるフィードバック条件が瞬目回数に与える影響を調査する。

4.2 実験概要

本実験は、以下の 2 つの段階で構成される。

(1) 瞬目検出条件の最適化実験

瞬目検出の精度を向上させるために、EAR の閾値設定を最適化することを目的とする。5名の被験者に動画を視聴させ、異なる閾値設定で瞬目回数を計測し、実際の瞬目回数と最も一致する閾値を特定する。

(2) フィードバック条件の有効性検証実験

異なるフィードバック手法が瞬目回数に与える影響を比較するために、10名の被験者を対象に、5つの条件（フィードバックなし、画面フラッシュ、アニメーション GIF、短音通知、スマートウォッチの振動）で瞬目回数の変化を測定する。

4.3 実験手順

(1) 瞬目検出条件の最適化実験

1. 被験者 5 名を対象に、各自 10 分間の動画を視聴させる。
2. 視聴中の瞬目をシステムで検出し、異なる EAR 閾値設定で瞬目回数を記録する。
3. 視聴動画を録画し、実際の瞬目回数を手でカウントする。
4. 実際の瞬目回数とシステムの検出結果を比較し、最適な EAR 閾値を決定する。

(2) フィードバック条件の有効性検証実験

1. 被験者 10 名に対し、各自 30 分間の動画を視聴させる。
2. 6 分ごとにフィードバック条件を変更し、瞬目回数を記録する。
3. フィードバックの提示順序をランダム化し、順序の影響を排除する。

4. 視聴後にアンケートを実施し、各フィードバックの主観的評価を収集する。

フィードバック条件の有効性検証実験では、被験者には瞬目が少なくなった時にフィードバックがある旨の説明のみを行うが、直接的にフィードバックに反応するよう指示はしない。瞬目検出条件の最適化実験の目的については説明を省略し、被験者自身が行動変容を促し、瞬目が増えるかどうかを検証する。実験の様子を以下の図 6 に示す。



図 6：フィードバック条件の有効性検証実験の様子

4.4 データ分析方法

瞬目回数の変化を統計的に評価するために、分散分析 (ANOVA) [14]を用いて、異なるフィードバック条件間の有意差を検証する。さらに、有意差が確認された場合は Tukey の多重比較検定[15]を実施し、条件間で有意な差があるかを分析する。

これらの手法を用いることで、瞬目促進システムの各フィードバック条件が瞬目回数に与える影響を明確化し、条件間の差異を定量的に評価した。

4.5 アンケートの実施

本実験では、各フィードバックの受容性や影響を調査するために、以下の項目を含むアンケートを実施した。

動画視聴時に認識できたフィードバックの種類

- フィードバックが動画視聴の妨げになったかどうか
- 各フィードバックが瞬目を促進する効果があったか
- 被験者にとって最も効果が高いと感じたフィードバック

アンケート結果を分析することで、フィードバックの有効性とユーザビリティを評価する。

本章では、瞬目検出の精度向上とフィードバック条件の有効性を検証するための実験について述べた。次章では、実験結果とその考察を示す。

5. 実験結果および考察

本本章では、実験で得られた結果を示し、それらを分析した上で考察を行う。実験は、瞬目検出条件の最適化実験

とフィードバック条件の有効性検証実験の 2 つの段階に分けて実施された。それぞれの結果を示した後、統計的分析とアンケート結果を基に考察を行う。

5.1 瞬目検出条件の最適化実験の結果

瞬目検出条件の最適化実験では、EAR の閾値と連続フレーム数の設定が瞬目検出精度に与える影響を調査した。結果として、EAR の閾値が 0.17 前後である場合、実際の瞬目回数に最も近い値を算出できることが確認された。また、連続フレーム数を 2 フレームに設定した場合、過検出や未検出のリスクを抑えつつ、安定した瞬目検出が可能であることが示された。図 6 に、EAR が 0.05 から 0.25 における連続フレーム数 2 の条件で得られた精度比較を示す。このグラフは、被験者ごとの EAR 閾値における検出精度の分布を視覚的に示しており、2 フレーム条件が他の条件と比較して最もバランスの取れた結果を示していることを確認できる。

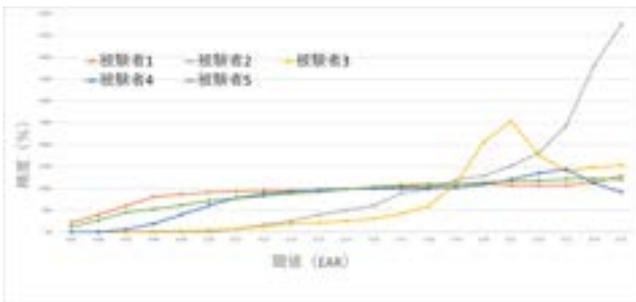


図 7：連続フレーム数 2 における閾値ごとの瞬目検出回数

しかし、被験者ごとに最適な閾値が異なり、個人差が大きいことも確認された。特に、被験者の瞬目パターンによっては 0.13 や 0.15 付近の閾値のほうが適切となる場合があり、被験者ごとの個別調整によってさらに精度を高める可能性が示唆された。このため、瞬目検出条件の最適化実験では事前に 1 分程度を用いて、個別に最適な EAR 閾値を設定する時間を設けることとする。

5.2 フィードバック条件の有効性検証実験の結果

フィードバック条件の有効性を検証するため、5 つの条件（フィードバックなし、画面フラッシュ、アニメーション GIF、短音通知、スマートウォッチの振動）において、瞬目回数の変化を測定した。10 人の被験者ごとに各フィードバックにおける 10 秒間の瞬目回数の平均を算出し、さらに各被験者の平均から全体の平均を算出した。結果を以下の図 7 に示す。フィードバックなしの条件と比べて、各フィードバック条件では瞬目回数が増加した。特にフラッシュを除く 3 つのフィードバック条件では、通常時の瞬目回数（10 秒間に 3 回）を上回る結果が得られた。

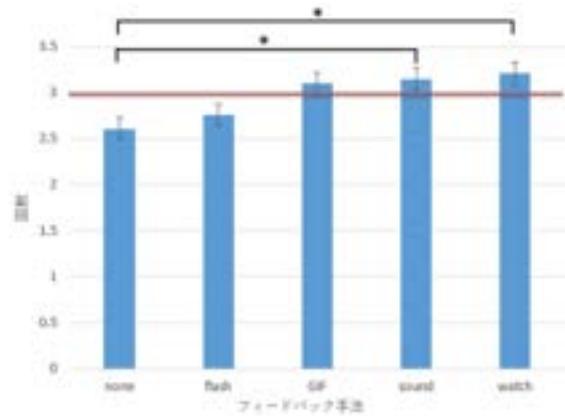


図 8：各フィードバックにおける 10 秒間の平均瞬目回数

*: $p < 0.05$

また、分散分析 (ANOVA) の結果、フィードバックなしの条件と短音通知・スマートウォッチの振動条件との間に有意な差が認められた ($p < 0.05$)。一方、画面フラッシュとアニメーション GIF の条件では、瞬目回数の増加が見られたものの、統計的に有意な差は確認されなかった。

分散分析によって有意差が認められたため、どの条件間で差があるのかを明確にするために、Tukey の多重比較検定を実施した。結果として、以下の条件間で有意な差が認められた。

- フィードバックなし (none) と短音フィードバック (sound) 条件 (平均差=0.54, $p=0.030$)
- フィードバックなし (none) と視覚フィードバック (watch) 条件 (平均差=0.60, $p=0.011$)

一方、その他の条件間では有意差は認められなかった。

Tukey の多重比較検定の結果をまとめて表 1 に示す。

表 1：Tukey の多重比較検定による書くフィードバック間の有意差

比較条件 (Group1)	比較条件 (Group 2)	平均差	p値	信頼区間 (Lower, Upper)	有意差
none	flash	0.1500	0.9290	(-0.3584, 0.6584)	FALSE
none	GIF	0.4889	0.0662	(-0.0196, 0.9973)	FALSE
none	sound	0.5417	0.0301	(0.0332, 1.0501)	TRUE
none	watch	0.6000	0.0113	(0.0916, 1.1084)	TRUE
flash	GIF	0.3389	0.3622	(-0.1696, 0.8473)	FALSE
flash	sound	0.3917	0.2189	(-0.1168, 0.9001)	FALSE
flash	watch	0.4500	0.1113	(-0.0584, 0.9584)	FALSE
GIF	sound	0.0528	0.9986	(-0.4557, 0.5612)	FALSE
GIF	watch	0.1111	0.9756	(-0.3973, 0.6196)	FALSE
sound	watch	0.0583	0.9979	(-0.4501, 0.5668)	FALSE

次に、フィードバック条件が瞬目回数に与える影響を被験者ごとに検討するために、各被験者に対して先ほどと同様に分散分析 (ANOVA) および Tukey の多重比較検定を行った。その結果、10 名中 8 名の被験者において統計的に有意差が認められた ($p < 0.05$)。表 2 に、各被験者の結果を示す。

表 2 : Tukey の多重比較検定による被験者間の有意差

被験者	F値	p値	有意差
1	1.6	0.1773	FALSE
2	3.48	0.0091	TRUE
3	4.76	0.0011	TRUE
4	10.03	<0.0001	TRUE
5	15.69	<0.0001	TRUE
6	2.96	0.0213	TRUE
7	3.78	0.0056	TRUE
8	0.55	0.6981	FALSE
9	2.49	0.0448	TRUE
10	8.41	<0.0001	TRUE

さらに、有意さが認められた被験者について、どの条件下で有意差があるかを特定するために Tukey の多重比較検定を実施した。有意差が認められた条件間を抽出したものを表 3 に示す。

表 3 Tukey の多重比較検定による有意差が認められた条件間の比較結果

被験者	条件1	条件2	平均差	p値	有意差
2	none	sound	-1.00	0.0046	TRUE
3	flash	sound	1.47	0.0034	TRUE
3	sound	watch	-1.39	0.0069	TRUE
4	GIF	sound	-0.78	0.0070	TRUE
4	none	sound	-1.36	<0.001	TRUE
4	flash	sound	-1.03	<0.001	TRUE
4	sound	watch	1.00	<0.001	TRUE
5	none	GIF	-2.72	<0.001	TRUE
5	none	sound	-3.11	<0.001	TRUE
5	none	watch	-2.39	<0.001	TRUE
5	flash	GIF	-2.39	<0.001	TRUE
5	flash	sound	-2.78	<0.001	TRUE
5	flash	watch	-2.06	<0.001	TRUE
6	none	GIF	1.63	0.0109	TRUE
7	flash	watch	-1.50	0.0403	TRUE
7	sound	watch	-1.53	0.0348	TRUE
10	none	GIF	-2.14	<0.001	TRUE
10	none	sound	-1.78	<0.001	TRUE
10	none	watch	-1.69	<0.001	TRUE

これらの結果から、フィードバック条件が瞬目回数に与える影響は被験者間で異なるものの、多くの場合、特定の条件間において有意な差が生じることが示された。

5.3 アンケート結果

被験者に対するアンケート調査では、各フィードバックの認識率や視聴体験への影響について意見を収集した。結果として、スマートウォッチの振動フィードバックは全被験者に認識され、視聴体験の妨げが少ないと評価された。一方、短音通知については、認識率が 60%と比較的低く、50%の被験者が「動画視聴の妨げになる」と回答した。

また、視覚的フィードバックについては、フラッシュに

対して「眩しく、字幕が読みにくくなる」といった否定的な意見が寄せられた。一方で、アニメーション GIF については「目立つが不快ではない」との意見があり、視聴体験を損ねにくい特徴を持つことが示唆された。

これらの結果から、フィードバックの種類によって瞬目促進効果と視聴体験への影響が異なることが確認された。特に、スマートウォッチの振動は視聴体験を損なわずに瞬目を促す有効な手法として期待される。

5.4 考察

本実験の結果を総合すると、瞬目促進には短音通知と触覚フィードバックが有効であることが示された。しかし、短音通知は認識率が 60%と他のフィードバックと比べると低く、また動画視聴の妨げになる可能性があり、全被験者にとって最適な手法とは言い難い。一方で、スマートウォッチの振動は、視聴体験を損ねることなく瞬目促進効果を発揮し、実用性の高い手法であると考えられる。また、視覚的フィードバックについては、GIF が視聴体験を損ねない特性を持ち、一定の瞬目促進効果を示した。一方で、フラッシュは短音通知と同様に認識率が 60%であり、また動画視聴の妨げになる可能性があることが示唆された。長時間の使用においては、目の疲れやストレスを引き起こす可能性があるため、今後の研究では、フィードバックの提示頻度や提示方法などを調整する必要があることが明らかになった。

さらに、瞬目促進効果には個人差があり、事前に最適な EAR 閾値やフィードバック設定を調整することで、より高い効果が期待できると考えられる。また、この実験では事前に「瞬きが少ないとフィードバックがある」という説明をしているため、瞬きを意識してしまった可能性があります。そのため認識度や視聴体験を考慮したフィードバックの設計が必要であることが考えられる。

本研究の結果を踏まえ、今後は個別最適化の手法を導入し、より効果的な瞬目促進システムの開発を目指す。また、本実験では被験者の数が 10 人であり、結果の信頼性を高めるためには、今後さらに多くの被験者を対象とした実験を行い、より精度の高い結果を得ることが必要である。

6. 結論

6.1 まとめ

本研究では、動画視聴中の瞬目回数の減少による健康リスクを軽減するため、視覚・聴覚・触覚を活用した瞬目促進システムを開発し、その有効性を検証した。本システムでは、瞬目検出に EAR (Eye Aspect Ratio) を用い、基準値を下回った場合に画面フラッシュ、アニメーション GIF、短音通知、スマートウォッチの振動という 4 種類のフィードバックを提示することで、瞬目を促進する設計となっている。

実験の結果、全てのフィードバックにおいてフィードバックなしの状態と比べて瞬目回数が増加し、特にアニメーション GIF の表示、短音通知、スマートウォッチのフィードバックで目標瞬き回数を上回った。また短音通知およびスマートウォッチが、瞬目回数の増加に対して有意な効果を示した。一方で、フラッシュと短音通知は、一部の条件では瞬目回数を増加させたものの、視聴体験を妨げる要因にもなり得ることが示された。また、アンケート調査では、GIF が比較的受け入れやすい視覚的フィードバックであることが確認されたが、短音フィードバックの認識率の低さが課題として浮かび上がった。

さらに、瞬目促進の効果には個人差が大きいことも明らかとなった。被験者ごとに最適な EAR の閾値が異なり、事前にパーソナライズされた設定を行うことで、より効果的な瞬目促進が可能になると考えられる。

本研究の成果により、瞬目促進システムの有効性が示され、ドライアイや眼精疲労の軽減に向けた新たな知見が得られた。しかし、視聴体験を妨げないフィードバック設計や、個人ごとの最適化を考慮した適応型アルゴリズムの導入など、さらなる改善が必要である。

6.2 今後の展望

本研究の成果を踏まえ、今後の研究では以下の課題に取り組む必要がある。

(1) フィードバック手法の改良

本研究で使用したフィードバック手法のうち、特にフラッシュは視聴体験を損ねる可能性が示された。今後は、視覚的負担を軽減するためのデザイン変更や、短音通知の認識率向上に向けた改善が求められる。また、触覚フィードバックに関しても、スマートウォッチ以外のデバイス（例：振動付きメガネ）を活用することで、より自然な瞬目促進が可能となるか検討する必要がある。

(2) 瞬目促進の長期的効果の検証

本研究では短期間の実験を実施したが、長期間にわたる瞬目促進効果の持続性については十分に検証されていない。今後は、数週間から数か月のスパンで瞬目回数の変化を追跡し、視覚健康への影響を詳しく調査する必要がある。また、ドライアイ症状の改善効果についても、眼科的な評価を含めた実験を実施することで、より実用的な知見を得ることができる。

(3) 瞬目促進システムのモバイル対応

本研究ではノート PC を用いた実験を行ったが、実際の使用環境ではスマートフォンやタブレットなどのモバイルデバイスでの利用が想定される。そのため、今後は Android や iOS 向けのアプリケーション開発を進め、より実用的なシステムへと発展させることが求められる。特に、バッテリー消費を抑えつつ、リアルタイムで瞬目を検出できる軽

量なアルゴリズムの開発が必要である。

(4) 個別最適化による瞬目促進効果の向上

本研究の結果から、瞬目促進の効果には個人差が大きいことが確認された。そのため、事前に被験者の瞬目パターンを測定し、最適な EAR の閾値やフィードバックの強度を個別に調整する機能を導入することで、より高い効果が得られる可能性がある。機械学習を活用した適応型アルゴリズムの開発により、ユーザーごとに最適なフィードバックを提供するシステムの構築が期待される。

本研究では、瞬目促進のための新たな手法を提案し、その有効性を実験的に検証した。今後の課題を克服することで、より実用性の高い瞬目促進システムの開発が可能となり、ドライアイや眼精疲労の軽減に寄与することが期待される。

参考文献

- [1] 総務省：令和 6 年版 情報通信白書 データ集 (2022). https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r06/html/data_shu.html#f00263. (最終閲覧日：2025/1/26)
- [2] 厚生労働省：平成 20 年技術革新と労働に関する実態調査結果の概況 (2008). <https://www.mhlw.go.jp/toukei/itiran/roudou/saigai/enzen/08/02.html>
- [3] 第一三共ヘルスケア株式会社：テレワークによる体の不調「テレワーク不調」に関する調査 (2022). <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r06/html/datashu.html#f00263>.
- [4] 坪田一男：涙のチカラ 涙は 7 マイクロリットルの海、株式会社技術評論社 (2008). (参照日 2025/1/17).
- [5] ウェルビーイングクリニック駒沢公園：デジタル時代における眼精疲労 (Digital Eye Strain, DES) (2025). <https://wbck.tokyo/archives/1353>. (最終閲覧日：2025/1/26)
- [6] ドライアイ研究会：ドライアイとは. <https://dryeye.ne.jp/for-general/dryeye-summary/>.
- [7] 東覚瑠菜, 神場知成：ドライアイ軽減のためのまばたきリマインド機能の開発と評価, 情報処理学会 インタラクシオン 2024 論文集, pp. 556-561 (2024).
- [8] Pavel, I. A., Bogdanici, C. M., Donica, V. C., Anton, N., Savu, B., Chiriac, C. P., Pavel, C. D. and Salavastru, S. C.: Computer Vision Syndrome: An Ophthalmic Pathology of the Modern Era, *Medicina*, Vol. 59, No. 2, p. 412 (online), <https://doi.org/10.3390/medicina59020412> (2023).
- [9] Xu, S. and Zhang, X.: Oculomotor trajectory mapping on body as an effective intervention to enhance attention, arXiv preprint, Vol. 2307.15172v2 (online), <https://arxiv.org/abs/2307.15172> (2023).
- [10] NEC LAVIE 公式サイト：仕様 LAVIE Direct PM(X) [Pro Mobile] - 13.3 型モバイルパソコン. <https://lenovo-nec.jp/navigate/products/pc/193q/10/lavie/pmx/spec/index.html>.
- [11] au: Google Pixel Watch. <https://www.au.com/mobile/product/plus-one/pixelwatch/>.
- [12] for Developers, G. A.: MediaPipe ソリューション ガイド (2025). <https://ai.google.dev/edge/mediapipe/solutions/guide?hl=ja>.
- [13] Soukupova, T.: Eye-Blink Detection Using Facial Landmarks, Master's thesis, Czech Technical University in Prague, Faculty of Electrical Engineering (2016). Available at <http://cmp.felk.cvut.cz/pub/cmp/articles/cech/Soukupova-TR-2016-05.pdf>.
- [14] Fisher, R.: *Statistical Methods for Research Workers*, Biological monographs and manuals, Oliver and Boyd (1925).
- [15] Tukey, J. W.: Comparing individual means in the analysis of variance, *Biometrics*, pp. 99-114 (1949).

パーソナリティ特性に基づくオンライン実験の分析： クラスター分析による介入効果

木村淳哉[†] 赤津裕子[†] 片桐一浩[†]
沖電気工業株式会社[†]

1. はじめに

近年、少子高齢化に伴う労働人口の低下や、企業の健康経営の観点から長時間労働の見直しなどの課題がある。対策として、就労者の生産性向上への取り組みが数多く検討されている。その中で、就労者の意識や日々の行動を変えさせるため、メッセージでの介入を用いて、健康や生産性を向上させる研究が行われている。メッセージ介入は、個人に適した手法や表現によって、生産性が向上すると考えられる。殿元ら[1]は、英文の多読行動の促進を目的に、心理特性に応じた効果的な介入を、目標設定やグラフ表示といった手法で検討した。そこで心理特性から介入への振る舞いを分類する可能性を述べた。二田ら[2]は、ネガティブなメッセージを提示し、心理的リアクタンスの喚起による行動の変容を検証した。吉澤は[3]、称賛獲得欲求や拒否回避欲求と Big-five の関係を検討し、性格特性によって、その欲求の程度が異なるとしている。これらの研究では、個人特性に応じた介入表現による生産性の向上効果は検討されていない。称賛するといったポジティブ表現のメッセージや、警告や注意といったネガティブ表現のメッセージでは、パフォーマンスの向上効果が異なり、個人特性に適した表現が存在すると考えられる。

そこで、本研究は、個人特性に応じた表現でのメッセージ介入によって生産性向上を促すシステム開発を目指す。個人特性としては、性格特性に着目した。認知能力や自己制御といった特性もあるが、性格特性は多くの行動や特性に幅広く影響している[4]と捉えて個別化を目指す。

筆者らの先行研究[5]において、2種類の介入表現とオンラインでの PC タスクを実施し、性格特性と介入表現の組み合わせによるタスクのパフォーマンスへの影響を明らかにした。しかし、メッセージ介入の効果検証に留まり、性格特性とパフォーマンス向上効果との関係は検証できておらず、性格特性に適した介入表現は明らかに出来ていない。

本稿は、実験参加者を性格特性で分類し、介入表現の違いによるタスクパフォーマンスの向上効果をクラスター毎に検証する。その結果から、各クラスターに適した介入表現の検討について述べる。

2. 実験概要

2.1 実験概要

本稿の実験は、先行研究[5]と同一で、介入表現とタスクパフォーマンスの関係の検証を目的とした。就労者の作業要素を抽出した2種類のタスクを Web 上で実施し、メッセージ介入によるパフォーマンス向上効果を検証した。

実験の流れは、タスクを3分間7試行繰り返すものとし、同様の手順で2種類のタスクを実施した。試行の間に計6回介入メッセージを表示した。メッセージ表現の条件は、ポジティブ介入条件とネガティブ介入条件、非介入条件の3種類である。メッセージの文言は、正答率や問題数に関わらず、規定の文言を提示した。なお、タスクの順序やメッセージ条件は、実験参加者毎にランダム化した。

2.2 タスク

就労者の作業特性を抽出した2種類のタスクを設けた。本研究では、制限時間内に解いた問題数と正解数、正答率をタスクパフォーマンスとして扱う。

計算タスクは、1桁同士の足し算で、計算結果の下一桁を回答させるものである。作業特性としては、単純で間違えられない緊張下を想定し、パターン認識能力や反復処理能力、計算処理能力を測る作業とした。

文字回転タスクは、回転した文字(“b”もしくは“d”)を表示し、文字を認識して回答させるものである。これは、作業書の図等と比較して模倣する作業や、物体を認識して組み立てる作業を想定している。

2.3 介入メッセージ

介入表現による違いを検証するため、ポジティブ感情とネガティブ感情を喚起させる2種類の介入表現の条件を作成した。各条件で言及する能力や改善点を揃え、メッセージ表現の違いによる評価を目的とした。

ポジティブメッセージは、成績や具体的な能力を肯定し、褒めや称賛によりポジティブ感情を喚起させる文章設計である[6]。ネガティブメッセージは、成績や改善の指摘、注意[2]をして、ネガティブ感情を喚起させる文章設計である。

2.4 個人特性アンケート

個人特性として、3つの指標に関してアンケートに回答してもらった。

Analyzing Individual Characteristics in Message Intervention: A Cluster Analysis of Effectiveness

[†] KIMURA ATSUYA, AKATSU YUKO, KATAGIRI KAZUHIRO

[†] Oki Electric Industry Co., Lt

1つ目は、Big-five の測定指標として NEO-FFI (NEO Five Inventory) [7]を用いた。これは、個人特性を幅広く計測する指標として多くの研究で活用され、メッセージへの影響や介入に対する反応が異なることが明らかになっている。

例えば、吉澤ら[3]は、神経症傾向が高い人は、拒否回避欲求が高いとし、否定的な評価の回避を目的にタスクへ取り組み、パフォーマンスが向上するとしている。また、外向性が高い人は称賛獲得欲求が高く、肯定的な評価を目的にタスクへ取り組み、パフォーマンスが向上するとしている。開放性の高い人は肯定的な評価を目的にパフォーマンスを向上させるとしている。Maria ら[8]は、誠実性が高い人は、ネガティブな介入を建設的なフィードバックとして捉え、正確性が向上するとしている。また、神経症傾向が高い人は、タイムプレッシャー下で、不安を煽る介入でパフォーマンスが低下するとし、同様の性格特性を有していても、状況によりパフォーマンスへの影響が異なると言える。

2つ目は、ユトレヒト・ワーク・エンゲージメント尺度[9]を用いた。これは、就労者の仕事に対する活力や熱意、没頭度合いといった、ポジティブで充実して仕事に向き合っているかを測る指標である。ワークエンゲージメントは、生産性に密接に関わっており、本タスクでも介入効果に影響を及ぼすと考えられる。

3つ目は、レジリエンスの測定指標として精神的回復力尺度[10]を用いた。これは、ストレスや困難な状況下において、うまく適応して対処できるかという指標である。本タスクは、時間制限内に出来るだけ多くの問題を解く設計であり、レジリエンスの影響が想定される。また、予備実験でも[6]、影響が見られており本実験でも採用した。

2.5 介入メッセージへの主観評価

介入メッセージの感情喚起の程度や印象を確認するため、評価アンケートを作成した。介入による悪影響も想定し、「1: ネガティブな影響」、「4:変化なし」、「7: ポジティブな影響」の7件法で回答をお願いした。

設問は、やる気や自信、ストレス、自己効力感などが介入メッセージを読んで変化したかを聞いた。また、介入メッセージ自体の評価として、文章の理解度や感情喚起の程度、文章の印象に関して、「変化なし」から「とても変化した」の7件法で回答をお願いした。

3. 実験結果

3.1 Big5 の回答

実験参加者は、20代から50代の就労者90名(男性42名、女性48名)であった。参加者のBig-five得点を、表1に示す。各因子で概ね同様の分布を示し、大きな偏りは見受けられなかった。外向性の平均値は他因子よりも低く、内向的な実験参加者が集まったことが分かる。

表 1 実験参加者の Big-five 得点分布

	開放性	誠実性	外向性	調和性	神経症傾向
平均	29.93	29.40	20.41	29.26	28.83
中央値	29.50	29.00	21.00	29.00	29.50
SD	5.50	8.23	8.79	7.66	9.78
尖度	0.16	0.10	0.31	-0.01	0.00
歪度	0.02	-0.18	0.11	-0.25	-0.37

3.2 クラスタ分析

Big-five の回答を基に、実験参加者を先行研究の手法を参考に分類した[11][12]。データの次元数を削減するために UMAP を使い、k-means 法によるクラスタ分析を実施した。その結果、クラスタ数を3とした際に、評価指標が最も高い値を示した。具体的には、シルエットスコアは0.457、Davies-Bouldin index は0.748であった。これらのスコアは、先行研究と同程度であり[12]、分類として妥当と判断した。

次に、クラスタ特徴の確認として、Big-five と2つの個人特性指標を整理した。図1に各クラスタの Big-five 平均得点を示す。

クラスタ1は、多くの指標で他のクラスタの中間的な値を示した。神経症傾向や調和性は平均的で、外向性や誠実性も偏りは見られなかった。ワークエンゲージメントとレジリエンスも中間であり、特に顕著な特徴がなかった。このため、クラスタ1を「バランス型」とした。

クラスタ2は、調和性や誠実性が高く、外向性も比較的高い値を示した。また、ワークエンゲージメントとレジリエンスが他クラスタの中で一番高かった。このため、クラスタ2を「積極的適応型」とした。

クラスタ3は、神経症傾向が最も高い一方、外向性が最も低かった。ワークエンゲージメントやレジリエンスも一番低く、このため、クラスタ3を「内向的慎重型」とした。

これの特徴から介入効果の仮説として、積極的適応型は、高い誠実性からネガティブ介入を建設的なフィードバックと捉え[8]、ネガティブ介入でのパフォーマンス向上が考えられる。また、ワークエンゲージメントやレジリエンスの高さから、難しいタスクを好み、文字回転タスクの介入効果が大きいと思われる。内向的慎重型は、高い神経症傾向から、ネガティブ介入でパフォーマンス低下が考えられ[8]、外向性の低さから、ポジティブ介入でパフォーマンス低下が考えられる[3]。

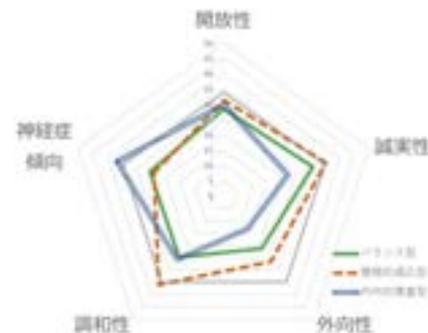


図 1 クラスタ毎の Big-five 平均得点

3.3 タスクパフォーマンス

性格特性とタスクパフォーマンスの関係を明らかにするため、ウィルコクソンの符号順位と検定により、各介入表現と非介入条件でパフォーマンスの差を分析した。表 2 に有意性または有意傾向が認められた条件の p 値と効果量 (Cliff's Delta) を括弧内に示す。効果量が負の場合は、非介入条件の方が高いパフォーマンスである事を示す。

バランス型は、ネガティブ条件において、文字回転タスクでは正解数が有意に増加した。

積極的適応型は、どちらの介入条件でも文字回転タスクにおいては、正解数と正答率の増加が見られた。

内向的慎重型は、ネガティブ条件において、計算タスクで問題数と正解数が有意に低下し、文字回転タスクでは正答率が低下傾向であった。

表 2 各条件のタスクパフォーマンスの結果

タスク種別	介入条件	指標	バランス型	積極的適応型	内向的慎重型
計算タスク	ポジティブ介入条件	正解数	*0.093 (-0.17)	ns	ns
		正答率	*0.064 (0.186)	ns	ns
	ネガティブ介入条件	正解数	ns	ns	**0.007 (-0.072)
		正答率	ns	*0.058 (-0.125)	*0.013 (0.066)
文字回転タスク	ポジティブ介入条件	正解数	*0.072 (0.159)	*0.06 (0.101)	**0.002 (-0.348)
		正答率	*0.075 (0.119)	**0.001 (0.266)	ns
	ネガティブ介入条件	正解数	**0.001 (0.235)	**0.011 (0.224)	ns
		正答率	ns	ns	*0.063 (-0.178)

*:p<0.1, **:p<0.05

3.4 主観評価アンケート

性格特性に基づくメッセージの受け取り方や印象の差異をアンケートにて確認した。図 2 に計算タスクの回答平均を示す。文字回転タスクでも同様の傾向が確認された。

バランス型は、ポジティブ介入に対して自信や自己効力感の変化が比較的小さかった。ネガティブ介入に対しては、やる気や集中力が低下するものの、他クラスターに比べて顕著な変化は見られなかった。

積極的適応型は、ポジティブ介入に対して積極性や自己効力感が高まった。一方、ネガティブ介入ではプレッシャーとストレスの増大とともに、自信の低下が確認された。

内向的慎重型は、ポジティブ介入に対してやる気や自信の向上が見られた。ネガティブ介入ではプレッシャーやストレスを中程度と感じているが、他クラスターと比べると大きな変化は見られなかった。

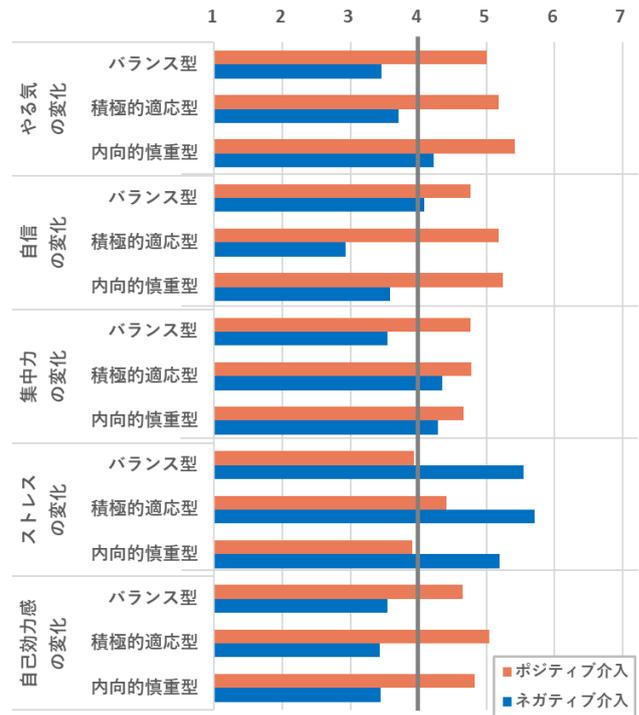


図 2 主観評価アンケートの回答平均

次に、介入表現の条件と感情の喚起度合いの影響を調べた。図 3 左に介入によるポジティブ感情およびネガティブ感情の喚起の程度度合いをクラスター別に示し、図 3 右に介入メッセージの印象をヒストグラムで示した。感情喚起の評価尺度は、「1:変化なし」から「7:非常にポジティブ/ネガティブな変化」であり、メッセージへの印象は、「1:非常に悪い」、「4:どちらでもない」、「7:非常に良い」である。

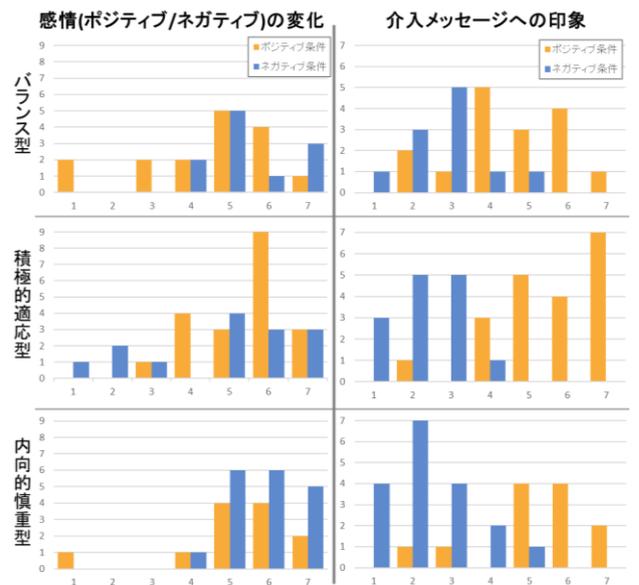


図 3 各クラスターの回答分布

感情喚起は次のような傾向を示した。バランス型は、両方の介入表現で感情喚起が比較的弱かった。積極的適応型は、ポジティブ介入で強い感情喚起を示し、内向的慎重型は、両方の介入表現で強い感情喚起を示した。

また、介入メッセージに対する印象評価では、ポジティブ介入で肯定的な評価が得られた一方で、ネガティブ介入では全体的に否定的な評価が得られた。特に、積極的適応型はポジティブ介入に強く好意的な印象を示し、内向的慎重型はネガティブ介入に強い否定的印象を示した。これらの結果から、ポジティブ介入とネガティブ介入の表現で差別化でき、今回の介入表現で性格特性によって評価が異なる事が確認された。

3.5 考察

クラスターの性格特性の特徴と介入メッセージへの主観評価およびパフォーマンスへの結果からクラスター毎の適切な介入表現を整理する。

バランス型は、計算タスクにてポジティブ介入がパフォーマンスの低下を招き、主観評価でもやる気やストレスの変化が他のクラスターと比較して少なかった。これは、今回の介入表現ではバランス型の特性に影響を与えなかった可能性を示唆している。バランス型は、感情やストレスへの反応が緩やかであるため、褒めや注意といった表現では動機づけに寄与しづらいと推察される。このクラスターでは、他の介入手法の検討や介入表現の強度を調整し、より個別的な対応が求められると考えられる。

積極的適応型は、誠実性と調和性が高く、神経症傾向が低いクラスターである。文字回転タスクにてポジティブ介入とネガティブ介入の両条件で向上効果が見られた。一方で、単調な作業では介入効果が弱い傾向が見られた。仮説の通り、認知能力が求められる様なやや複雑な作業においては、称賛や注意を建設的なフィードバックとして捉えるため、パフォーマンス向上に寄与する可能性が示唆された。

内向的慎重型は、神経症傾向が高く、外向性が低いクラスターである。両方のタスクで、ネガティブ介入によるパフォーマンスの低下が確認された。また、主観評価でもポジティブ介入とネガティブ介入に対して評価が顕著に分かれ、感情喚起の度合いが大きいという結果が得られた。仮説通り、神経症傾向の高さから、ネガティブ介入によって、不安が増大され焦燥感を引き起こすことで、タスクへの集中を妨げたと考えられる。

今回作成した介入メッセージは、全てのクラスターにおいて感情喚起を促す効果を示し、ポジティブ表現とネガティブ表現の印象に明確な差異が見られた。しかし、複数の介入で介入効果を検証する実験設計のため、どの介入表現が最も有効であったかの特定には至っていない。また、クラスターごとの感情喚起の程度に違いが見られ、介入効果にどの性格特性が寄与したかを詳細に確認する必要がある。

また、本研究のタスクは、業務作業の要素を抽出した実験課題であるが、本結果のみで実際の職務におけるパフォーマンスに直結するかは断定できない。また、タスクの種類により、介入効果の程度が異なることが示されたが、介入表現がタスクのどの特性に影響していたかについては未検証である。

4. おわりに

本研究では、介入表現の効果を検証するためオンライン実験を実施し、性格特性に基づくクラスター毎にタスクパフォーマンスへの影響を分析した。その結果、タスクと介入表現の組み合わせによるタスクパフォーマンスへの差異が確認された。また、クラスターの性格特徴や介入表現への主観評価から、神経症傾向が高いクラスターでは、メッセージへの印象評価が顕著に異なるなど、性格特性が介入効果に及ぼす影響が示唆された。

しかし、本実験だけでは介入表現の効果の十分な調査に至らず、結果を一般化して、業務におけるタスクパフォーマンスへの影響を直接結論づけることは難しい。

今後の展望として、より日常的な業務シーンを想定した実験タスクを設定し、介入システムの開発および有効性の検証を進めていく。

参考文献

- 1 殿元禎史, 黄瀬浩一. 英文多読のための個人に適合したナッジ戦略の検討. 情報処理学会研究報告, vol. 2020, no. 9, p. 3-6.
- 2 二田悠史, 杉原太郎, 佐藤健治, 五福明夫. 行動変容促進を目的とした論しおよび警告メッセージの予備的検討. 情報処理学会研究報告, vol. 2019, no. 13, p. 2-8.
- 3 吉澤英里. 称賛獲得欲求・拒否回避欲求と Big Five の関連について. 環太平洋大学研究紀要, 2020, p. 37-41.
- 4 高橋雄介. パーソナリティ特性研究をはじめとする個人差研究の動向と今後の展望・課題. 教育心理学年報, Vol. 55, p.38-56, 2016.
- 5 木村淳哉, 赤津裕子, 片桐一浩. メッセージ介入によるパフォーマンス向上効果の検証. 2025, インタラクシオン.
- 6 木村淳哉, 片桐一浩. 心理特性に基づくパーソナライズと行動変容効果の検討. 情報処理学会第 86 回全国大会論文集, 2024, p. 291-292.
- 7 下仲順子, 中里克治, 権藤恭之, 高山緑. 日本版 NEO-PI-R の作成と因子的妥当性の検討. 性格心理学研究, 1998, vol.6, no.2, p.138-147.
- 8 Maria, C., Ana, N. C., Santiago, S. P., et al. Do Personality Traits Affect Productivity? Evidence from the Lab: Working Papers, 2014.
- 9 Shimazu, A., Schaufeli, W. B., Kosugi, S. et al. Work engagement in Japan: Validation of the Japanese version of Utrecht Work Engagement Scale. Applied Psychology: An International Review, 2008, vol.57, p.510-523.
- 10 小塩真司, 谷素之, 金子一史, 長峰伸治. ネガティブな出来事からの立ち直りを導く心理的特性—精神的回復力尺度の作成—. カウンセリング研究, 2002, vol.35, p.57-65.
- 11 Kerber, A., Roth, M., Herzber, PY. Personality types revisited—a literature-informed and data-driven approach to an integration of prototypical and dimensional constructs of personality description. PLoS One, 2021.
- 12 Baruth, O., Cohen, A. Personality and satisfaction with online courses: The relation between the Big Five personality traits and satisfaction with online learning activities. Educ Inf Technol, 2023, vol.28, p879-904.

選書中の一人称映像を用いた行動パターンの分析

畑井梨里衣^{†1}
公立ほこだて未来大学システム情報科学研究科
角康之^{†2}
公立ほこだて未来大学システム情報科学研究科

1. はじめに

本研究の目的は、興味を絞り込む行動プロセスを可視化し、その特徴を明らかにする方法を探ることである。

日常生活の中で、人は多くの選択肢の中から徐々に対象を絞り込んでいく行動を自然に行っている。例えば、買い物の際に膨大な商品の中から1つの商品を選ぶプロセスや、飲食店でメニューを選ぶ際のプロセスが挙げられる。このような行動を分析することは、購買行動のプロセス解明や、潜在的な興味の推定、それに基づいたマーケティング戦略の立案に役立つ可能性がある。さらにそれらを応用して、商品探索システムや推薦システムの設計の改善などの幅広い応用も期待できる。

ここで、関連する研究を紹介する。

まず、分散的な興味を絞り込む行動に注目したものを挙げる。たとえば、伊藤ら [1] による EC サイトに注目した研究がある。EC サイト上の購買においては、Web ページを閲覧する中で徐々に興味の対象が絞られていき、最終的にある商品を購入するというユーザ行動が想定されている。この研究では、Web ページの閲覧データから、ユーザの興味の収束度合いをモデル化する手法を提案している。このような、人の意思決定において興味の範囲が変化していく過程を分析することで、マーケティング上の施策の効果を高めることなどが期待できることも示唆されている。

次に、視線計測装置を用いた研究の例を二つ挙げる。

一つ目は、EC サイトにおける商品探索タスクにおいて、視線情報からユーザーの特徴推定を試みた例である。伊藤ら [2] は、定量的な視線データからユーザのメンタルモデルの推定を行う方法を提案している。このように客観的なデータからユーザ理解を促進する方法は、効率性、再現性の観点からも有益であると示された。

二つ目は、図書館内での書架を探索したり、情報を検索したりする際の行動データを収集して、情報探索行動のパター

ンを分析した例である。寺井 [3] は、収集された行動データには視線計測データが含まれ、視線計測により、被験者の「潜在的な興味」や「探索行動の詳細」を可視化可能であることを示した。

これらの関連研究から、視線データから興味や意図を推定する手法が有効であることが確認できる。

以上の背景を踏まえ、本研究では、実世界の選書行動に注目し、視線計測装置による一人称映像を用いて、分散的で曖昧な興味から対象を絞り込む行動を分析した。予備実験としてデータ収集を行い、収集した一人称視点の映像データに対して、ヒートマップによる可視化と考察を行った。

ヒートマップによる視線の可視化により、どの書籍がより多くの注目を集めているかが明確になり、今後の書籍配置や推薦システムの改善に役立つと考えられる。さらに、視線データの分析によって選書時の曖昧な興味の状態から最終選択に至るプロセスを段階的に把握することができれば、本研究の手法は、より広範な選択行動の理解に寄与する可能性がある。

2. 手法

2024 年度に公立ほこだて未来大学にて開催されたイベントである未来大学ブックフェア（以下、ブックフェアとする）に伴いデータ収集を行った。ブックフェアとは、学生の専門分野や教養に関する本約 2000 冊を学内ミュージアムに展示し、教員や学生が本を手にとって選書できるイベントのことである。ここで選書とは、学内のライブラリに在庫として入れてほしいと本を自分たちで選択する行為を指す。ブックフェアにおいては、会場で配られる紙片を各自が選んだ本に挟み込むことで選書することができる。本実験では、被験者に視線計測装置を装着してもらいながらブックフェアに参加してもらうことで一人称映像データを収集した。被験者は会場を自由に散策し、選書したい本を見つけた場合は被験者の任意で紙片を挟むこととした。一人称映像データを収集する被験者は、学内の教員 1 人と学生 12 人とした。

使用デバイス (Pupil neon) について説明する。一人称

Analysis of behavior patterns during book selection using first-person video

^{†1} RIRII HATAI, Future University Hakodate

^{†2} YASUYUKI SUMI, Future University Hakodate

映像を記録する視線計測装置として、Pupil Labs 社^{*1}の NEON というデバイスを使用した。このデバイスは、キャリブレーションが不要でメガネ型のアイトラッカーを装着した瞬間から視線計測可能であり、専用のスマートフォンと組み合わせて使う。そのため視覚刺激や計測場所を問わず、動き回りながらも安定した視線計測が可能である。このデバイスの特徴はブックフェア会場を歩き回りながらの視線計測に適していると判断し、今回の実験に用いた。

視線計測装置を用いて収集したデータは Pupil Cloud ^{*2} にアップロードされ、アップロードされたデータを Cloud 上の解析ツールで分析することができる。本研究では、Reference Image Mapper という解析機能を使用した。Reference Image Mapper については、3 章で詳細を説明する。

3. 分析

Pupil Cloud の解析機能のひとつである Reference Image Mapper を用いて、一人称映像データの分析を行った。

Reference Image Mapper とは、一人称映像データから、参照した画像の上に視線データをマッピングする機能のことである。参照する画像と、参照画像に映っている対象を複数の方向からおさめた一人称映像データを読み込むことで、AI による画像認識で自動的に参照画像へのマッピングが行われる。

図 1 は、Reference Image Mapper により視線のマッピングを行っている画面のスクリーンショットである。左の画像は一人称映像、右の画像は視線データをマッピングする参照画像である。左の一人称映像に見える白い点描は、右の参照画像と一致する特徴点が一人称映像から抽出されていることを表す。そして、一人称映像にて赤丸で示されている被験者の視線データが、参照画像にも赤丸でマッピングされている様子が確認できる。

視線のマッピングに用いる参照画像としては、ブックフェアの会場で撮影した写真を用いた。会場には複数の書籍展示テーブルが設置されており、注目対象として選んだそれぞれの書籍展示テーブルを異なる方向から撮影した。注目対象として選んだ 4 つの書籍展示テーブルに対して、計 9 枚の参照画像を用意した。

さらに、Reference Image Mapper により参照画像にマッピングされた視線について、ヒートマップを生成することでデータの可視化を行った。これにより、注目対象となる書籍展示テーブルごとに被験者全体と各々の視線行動を集約し、その傾向や違いを概観することができる。

図 2 は、マッピングされた視線のヒートマップを抜粋したものである。列ごとに同じ参照画像に対応しており、行



図 1 参照画像への視線マッピング

Figure 1 Gaze mapping to a reference image



図 2 視線ヒートマップの抜粋

Figure 2 Gaze heatmap excerpt

ごとに同じ被験者のデータに対応するヒートマップを並べている。ただし、一行目は、被験者全員分のデータから生成したヒートマップを並べている。図 2 でヒートマップに抜けがある箇所は、対応する参照画像に映っている箇所に対する被験者の視線データが不十分であったため空欄としている。

これらのヒートマップを概観すると、被験者ごとにヒートマップが集中しているエリアが異なることがわかる。

また、選書を行った被験者のヒートマップにおける注視エリアと、それぞれの被験者が選書した本が置かれていた位置を照らし合わせて確認をした。図 3 に、被験者が選書した本とヒートマップの重なりを示した 4 人分の例を示す。それぞれの被験者が選書した本の位置を濃い青丸で表示している。被験者が選書した本が置かれていた位置と、ヒートマップにおいて着色された注視エリアが重なっていることが確認できる。

4. 考察

本研究では、ブックフェアにおける選書行動を一人称映像データを用いて分析し、視線ヒートマップを作成することで視線の集中領域を可視化した。その結果、(1) 個人ごとに視線が集中する領域が異なること、(2) 最終的に選書された本の位置と視線の集中領域が概ね重なっていることが確

^{*1} <https://pupil-labs.com>

^{*2} <https://pupil-labs.com/products/cloud>

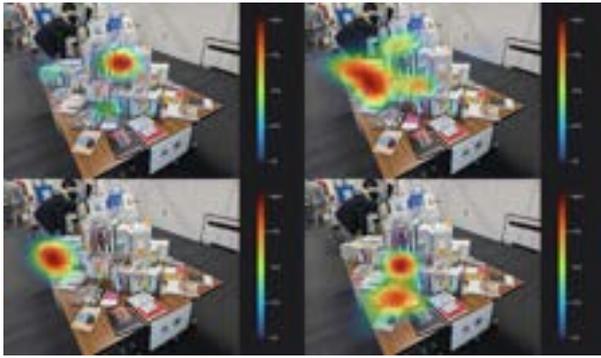


図3 選書された本とヒートマップの重なり

Figure 3 Overlap of selected books and heat maps

認められた。これにより、選書行動において視線の集中が興味
の収束と関係している可能性が示唆された。

本研究の結果から、次のような行動パターンが推測できる。まず、被験者の視線が長時間滞留した領域が、最終的に選ばれた本の位置と一致していることから、人はまず視線を通じて興味を絞り込み、その後選択に至るといった行動パターンが裏付けられる可能性がある。次に、ヒートマップの分布には個人ごとの違いが見られたことから、選書の際における個々の興味や探索行動のパターンには個人差があることが視覚的に示唆された。このように、視線の集中と興味
の収束の関係や、個人差の存在が説明できる結果が観測されたといえる。

また、先行研究と比較すると、本研究ではブックフェアの散策時における一人称映像から視線データを解析することで、EC サイトなどのオンライン上の選択行動に対する研究とは異なり、実世界における探索行動の特徴を明確に捉えることができた点が特徴的である。このような一人称視点映像データの分析を進めることで、実世界における探索行動の段階的な変化をさらに詳しく観察することができると考えられる。

5. 展望

本研究では、ブックフェアにおける選書行動を一人称映像データと視線ヒートマップを用いて分析し、視線の集中領域と最終的な選書との関連を示唆した。しかし、本研究には以下のような課題があり、今後の発展が求められる。

第一に、視線データのみでは選書
の意思決定プロセスの詳細を把握することが難しい。例えば、「本を手にする」「ページをめくる」といった行動データをラベリングにより分類して組み合わせ、視線の動きと行動の関係をより精緻に分析することで、選書行動の理解が深まると考えられる。

第二に、本研究は特定のブックフェアという環境に限定さ

れたデータに基づいているため、異なる状況に対する検証や、より網羅的なデータの収集が必要である。例えば、図書館や書店といった異なる文脈で同様の実験を行ったり、被験者の条件をより詳細に設定したりすることで、選書行動の普遍的な特徴を明らかにすることが求められる。

第三に、本研究では単回のデータ収集に基づいて分析を行ったが、長期的な興味の変化や選書行動の変遷についての検討は十分に行えていない。特定の被験者を対象に時間を置いてデータを収集し、興味の変化や選書スタイルの変遷を追跡することで、より深い知見が得られる可能性がある。

今後の展望として、(1) 行動データの統合による詳細な選書プロセスの解明、(2) 異なる環境における実験の実施と網羅的なデータ収集、(3) 長期的な興味の変化の分析が重要である。さらに、視線データに基づくクラスタリングを行い、異なる選書スタイル（例：直感型、熟考型）を分類することで、よりパーソナライズされた選書支援の可能性も探ることができると考えられる。

これらの展望に基づき、より発展的な研究を進めることにより、視線データを活用した意思決定プロセスの分析がさらに深められると考えられる。これにより書籍推薦システムや書店・図書館の書籍配置の最適化など、実用的な応用へつなげることや、さらにはより広範な選択行動の理解への一助となることが期待できる。

参考文献

- [1] 後藤 正幸伊藤 史世：顧客の閲覧行動分析のための時間窓トピックモデル、情報処理学会論文誌、Vol. 64, No. 1, pp. 214-228 (2023).
- [2] 吉武良治伊藤拓巳：視線情報によるユーザーの商品探索の特徴推定、日本デザイン学会第70回研究発表大会, 8D-06 (2023).
- [3] 寺井 仁：ハイブリッドな情報環境における情報探索行動の分析、情報の科学と技術、Vol. 61, No. 1, pp. 15-21 (2011).

自閉傾向とマガーク効果における視聴覚刺激への選好

北田結^{†1}・大森幹真^{†2}

早稲田大学人間科学部^{†1}・早稲田大学人間科学学術院^{†2}

1. はじめに

自閉スペクトラム症(以下 ASD)と定型発達(以下 TD)は感覚や行動において異なる。表情認知について ASD は相手の顔自体に視線を向けず,その中でも特に目を見ない傾向にある。そして,複数の感覚器から得られた情報を統合する他感覚機能の統合について ASD はこの統合処理に問題がある。そこで音声知覚の視聴覚統合を検討するのに有用であるとされているマガーク効果という錯覚を用いて ASD の統合処理の苦手さについて検討する。マガーク効果とは音声とは異なる口の動きをつけた動画を視聴した際,音声情報と聴覚情報が統合された音の聞こえ方をする錯覚である。Ujiie et al. [1]は自閉傾向が高くなるほど融合反応が減少すると示す。その要因が顔への注視時間なのか,処理能力の未発達なのかは明らかになっていない。そこで,アイトラッカーを使用し視線移動を計測し自閉傾向の有無による児童の視線移動の特徴を明らかにした上で,刺激に融合音を含めることで視覚と聴覚の優位性を考察する。

2. 方法

3歳から12歳の児童26名を対象とし,事前に行ったテストをもとに自閉傾向高群と低群に分けた。単音節(/pa/,/ka/,/ta/)の音声と映像を収録したものを編集し,動画を作成した。1.5秒間の映像を6回繰り返した約9秒間の動画を1施行として計8施行おこなった。その後,続けて冒頭と末尾は視聴覚刺激と音声刺激が一致する1.5秒間の映像を各2回ずつ繰り返し,中間は視覚と音声異なる1.5秒間の映像を2回繰り返した約9秒の動画を1施行として計3施行おこなった。課題の呈示と同時にアイトラッカーによる視線の記録を行った。動画視聴後は毎時,聞こえた音を3つの選択肢から複数選択を可能とし強制選択を行った。従属変数は聴覚・視覚・融合反応率を算出し,発話者の目,口,背景にAOI (Area Of Interest)を設定し,視線停留時間,視線停留回数を測定した。

3. 結果

自閉傾向低群が自閉傾向高群に比べ優位に視覚反応率が高くなっていることが明らかになった。視線停留時間は自

閉傾向との間に交互作用傾向が見られた($F(1, 25) = 2.686, p = .078, \eta_p^2 = 0.101$)。そして,自閉傾向の高低を要因とした各領域の視線停留時間に多重比較をしたところ,口や背景に対して自閉傾向低群は高群と比べ有意に長い間,注視していることが明らかになった。そして,融合音を刺激に含む選択正答率に関して,交互作用は見られなかった(図1)。

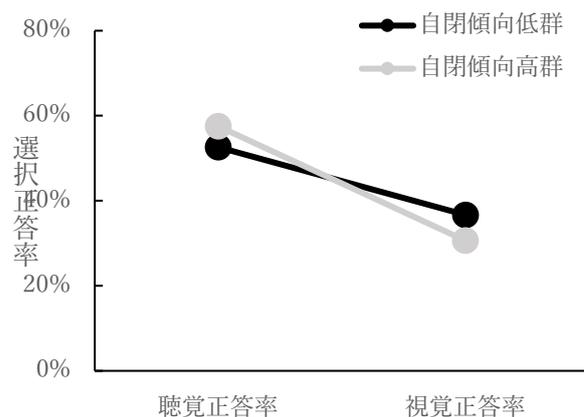


図1: 融合音を含む刺激の選択正答率

4. 考察

自閉傾向高群が低群と比べ,聴覚情報を優位に取り入れ視覚からの情報の干渉が少ないことが示唆された。視線は自閉傾向が高いほど画面への注視時間は短く,特に口や背景を見ない傾向にあった。これは自閉症の特徴である社会的要因を反映した結果となった。目への注視時間に差異はなく特に口について差異があったことについて TD は違和感から注目度が高まったのに対し,ASDは口への注視時間の短さから矛盾に気づかない,もしくは社会的な要因から顔の注視に苦手さがあったと解釈できる。そして融合音を含む選択反応率においてグラフが交差しているためサンプル数を増やすことで自閉傾向の違いによる選好性が見られる可能性がある。

参考文献

- [1] Ujiie, Y., Asai, T., & Wakabayashi, A. (2015). The relationship between level of autistic traits and local bias in the context of the McGurk effect. *Frontiers in Psychology, 6*(891).

^{†1} YUI KITADA, Faculty of Human Sciences, Waseda University

^{†2} MIKIMASA OMORI, Faculty of Human Sciences, Waseda University

視覚探索課題からみた ASD 児の局所的注意の優位性

飯田亜衣^{†1}・大塚千冬^{†1}・大森幹真^{†2}

早稲田大学人間科学部^{†1}・早稲田大学人間科学学術院^{†2}

1. はじめに

自閉症スペクトラム障害(以下 ASD)のある子どもたちは限局的な興味や関心があり、特定の事物に対して固執的な反応をすることが多い。それが日常生活や学習活動において、注意や集中力の維持に課題を抱えることにつながることもある。特に、ASD 児は意味情報に注意を向けることが苦手であり[1]、全体的な事物の認識よりも、局所的なものへの注意が向きやすいという中枢性統合の弱さがある(Frith, 2012)。一方で、限局的な興味関心の強さと、局所的なものへの注意の向きやすさについて、どちらがより優位なのかは未解明である。そのため、本研究では就学期の子どもたちに対して、視覚探索課題を行い、25 個の刺激の中から他と異なる刺激を 1 つ選択することを求めた。そのうえで質問紙による ASD 傾向の高低により反応傾向や、課題中の視線特性に違いが表れるかを明らかにすることを目的とした。

2. 方法

2.1 参加者

5 歳 0 か月から 14 歳 3 か月までの子どもが 26 名参加した。これらの人数に対して ASD 傾向を反映する質問紙である AQ の総合得点が 25 点以上もしくは、SRS-2 の総合 T 得点が 60 点以上の 13 名を ASD 高群とした。それ以外の 13 名は ASD 低群とした。

2.2 刺激と装置

漢字問題の提示用にノートパソコン(PC)と視線特性を計測するためにイトラッカー(X3-120, Tobii technology Japan)を使用した。本研究では、PC 上の 33.87cm×19.05cm のスライドの中央に 25 個 (5×5) の文字 (うち 24 個は同一、1 個は異なる文字) を用意した。

2.3 手続き

(1)視線特性計測：参加者は、PC を置いた机の正面の椅子に座り、キャリブレーションを行った。課題では、画面上に 25 個の漢字が提示され (25 個中 1 つのみ異なる漢字)、その中から 1 つ異なる漢字をなるべく早く、正確に、クリックすることを求めた。制限時間は 30 秒で、30 秒以内に標的刺激を見つけた場合は、その漢字をクリックすることで次の提示に進むことができ、9 課題終了まで継続した。

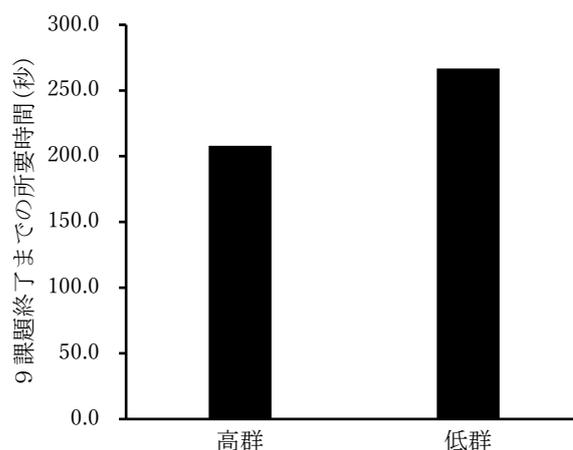
(2)独立変数と従属変数：独立変数として各刺激に、標的刺激位置と、前回の標的刺激位置、全領域に興味領域(AOI)を

設定した。従属変数は以下の 4 つを設定した。①標的刺激を選択するまでの所要時間 ②AOI 領域ごとの合計注視時間 ③AOI 領域ごとの視線の平均停留時間、④ AOI 領域ごとの視線停留回数。

3. 結果

図 1 から 9 課題の回答を終えるまでの所要時間を比較したところ、ASD 傾向高群は低群に比べて有意に所要時間が短くなったことを示した [$t(24)=1.81, p<.05^*, d=.71$]。また、ASD 傾向高群の方が低群よりも平均停留時間が短くなったという有意な主効果も示した [$F(1, 24)=11.02, p<.05, \eta^2=.32$]。一方で、合計注視時間や視線停留回数においては標的刺激と前回の標的刺激提示位置間で有意な主効果が見られたが、ASD 傾向間で有意差は見られなかった。

Figure1 9 課題終了までに要したグループ間の所要時間



4. 考察

本研究から、ASD 傾向高群では局所バイアスの方が優位に出やすく、低群の方が限局的な興味関心が強く顕在化する可能性を示した。今後の研究では、年齢や文字への理解度が今回示された傾向に与える影響を明らかにすることにより、教育現場へ応用していくことが課題である。

参考文献

- [1] Happé, F., Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorder*. 36, 5–25.

†1 AI HIDA, Faculty of Human Sciences, Waseda University

†1 CHIFUYU OTSUKA, Faculty of Human Sciences, Waseda University

†2 MIKIMASA OMORI, Faculty of Human Sciences, Waseda University

協調運動得点と ASD 傾向における読み書きの視線の違い

渡邊貴子^{†1}・大森幹真^{†2}

早稲田大学人間科学部^{†1}・早稲田大学人間科学学術院^{†2}

1. はじめに

文部科学省 (2022) が通常の学級に在籍する発達障害の可能性のある特別な教育的支援を必要とする児童生徒に関する調査を行った結果、学習面で著しい困難を示す小学生は、6.5%いると推定している。書く際には、認知機能や運動機能が求められる。Carames (2022) は、協調運動の向上は、読解力、書字能力の向上と関連していることを明らかにし、微細運動制御と視覚運動統合の障害が ASD 児の手書きの困難に寄与している可能性が高いことを述べている (Kushki et al., 2011)。そこで、Sita & Talor (2015) が行なった、単語が個別または文中に含まれている場合の読み書きの視線を比較した実験では、読みより書きの方が視線、平均注視時間がともに長く、文中と比較し個別で書いた方が、視線が多く、注視時間に変化がないことが明らかとなっている。本研究では先行研究をもとに、文章全体を読み書きした場合(文章群)を追加し、単語(単語群)を個別または文中(文中単語群)で読み書きした場合の視線の違い明らかにし、年齢と読み書きの関連について検討した。

2. 方法

2.1 参加者

男子 9 名女子 9 名で平均年齢は 8.7 歳であった。

2.2 刺激と装置

文章全体を読み書きする課題(課題 1)、単語(すいか、ぶらんこ、ひこうき)を個別と文中にある場合で読み書きする課題(課題 2)を用意した。課題はモニターに映し出され、下部にはアイトラッカーを設置した。アイトラッカー(X3-120, Tobii technology Japan)による測定興味領域内で観測された視線を分析対象とした。

2.3 手続き

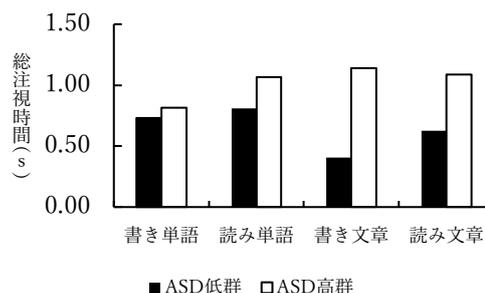
Berry-VMI 検査の Motor Coordination (運動協応) の結果をもとに MC 高群・低群に分類した。また、AQ 得点により ASD 傾向度を測った。参加者はアイトラッカーの画面から 50cm の位置に着席し、キャリブレーションを行った。その後課題 1、課題 2 を進めた。従属変数には、反応時間、総注視時間、平均注視時間、視線停留回数を用いた。また、MC または AQ 得点の高群と低群で分散分析を行った。

3. 結果

課題 1 について反応時間と総注視時間、平均注視時間、視

線停留回数について分散分析を行なった結果、いずれも MC 得点、AQ 得点による有意な差は認められなかった。読み書きにおいて反応時間 [$F(1, 16) = 25.5, p < .001^{**}$]、平均注視時間 [$F(1, 16) = 21.10, p < .001^{**}$] に有意な差が認められ、総注視時間、視線停留回数には有意差は認められなかった。課題 2 について分散分析を行なった結果、総注視時間と平均注視時間の読み書きにおいて有意な差が認められた [$F(1, 16) = 7.76, p < .01$; $F(1, 16) = 9.59, p = .007$]。視線停留回数、MC 得点による有意な視線の違いは認められなかった。また課題 2 の総注視時間の読み書きと AQ 得点に [$F(1, 15) = 5.72, p = .030$]、視線停留回数の単語の提示方法と AQ 得点に有意な傾向がみられた [$F(1, 15) = 5.98, p < .03$]。

Figure1 課題 2 における総注視時間の視線データ



4. 考察

本実験では協調運動得点と AQ 得点における読み書きの視線の違いはなかった。課題 1 より、書きの反応時間と読みの平均注視時間が長く、課題 2 において書きの総注視時間と読みの平均注視時間が長いという結果は、書く場合には画面より手元を多く見て書いていることが示唆することとなった。先行研究と異なる結果が得られたのは、参加者の年齢の違いに伴い、低年齢であることによる書きに必要な運動発達の違いが挙げられる。今後は、運動発達の程度による視線の違いを検討することで読み書きに困難を抱える子どもを早期に発見できる可能性がある。

参考文献

Sita, J. C., & Taylor, K. A. (2015). Eye movements during the handwriting of words: Individually and within sentences. *Human Movement Science, 43*, 229-238.

^{†1} KIKO WATANABE, Faculty of Human Sciences, Waseda University

^{†2} MIKIMASA OMORI, Faculty of Human Sciences, Waseda University

仮想と現実をつなぐジェネラティブアート体験： 複合現実を活用した学び

古市駿^{†1}

公立はこだて未来大学^{†1}

角薫^{†2}

公立はこだて未来大学^{†2}

1. はじめに

本研究では、複合現実 (Mixed Reality, MR) 技術を活用し、仮想空間に表示された「ジェネラティブアート」を操作・体験することで、数学やプログラミングへの学習意識の向上を目指す。ジェネラティブアートとは、コンピュータソフトウェアのアルゴリズムや数学的手法を用いて生成されるデジタルアートであり、数学の応用例の一つとして注目されている[1][2]。ジェネラティブアートは、数学的アルゴリズムやフラクタル構造を活用し、計算過程を視覚化することで、数学教育への応用が期待される。

MR 技術は、現実世界と仮想世界をリアルタイムで融合・表示する技術であり、従来のバーチャルリアリティ (Virtual Reality, VR) の限界を打破する新しい情報提示手段として注目を集めている。MR の特長は、現実空間と仮想空間を統合する点にある。本研究では、この特性を活かし、ユーザが現実空間から物理的に操作できる装置を導入し、仮想空間と現実空間を組み合わせたインタラクティブな体験を提供するシステムを開発する。このような体験を通じて、仮想空間のみで操作するシステムと比較した際の違いを調査し、MR ならではの教育的・体験的効果を明らかにすることを目指す。

さらに、MR を活用することで、ユーザはジェネラティブアートを様々な角度から観察することが可能となる。数学的構造を持つジェネラティブアートを、下から、上から、側面からといった自由な視点で三次元的に観測することは、通常の学習環境では得がたい体験である。本研究では、この特性を活かし、数学的概念の理解を直感的に促す教育システムの構築を試みる。

現在、日本の学習環境には改善の余地がある。文部科学省によると、「算数・数学の勉強は楽しい」と答えた日本の小学生は 77%、中学生は 56% であり、国際平均 (小学生: 84%、中学生: 70%) を下回っている[3]。さらに、日本の生徒は OECD 平均と比較して、数学を実生活に応用する自信が低いと報告されている[4]。この要因の一つとして、数学と日常生活との結びつきを実感する機会の不足が挙げられる。

数学教育において、AR (拡張現実) や VR (仮想現実) を

活用した学習支援システムの研究が進められている。生物学の授業に AR を導入することで、学生の学業成績やモチベーションに与える影響を調査した研究では、AR が学生のモチベーション向上に寄与することが示されている[5]。一方、VR を用いた科学教育に関する研究では、VR レッスンが学生の興味や関与を高める可能性があることが報告されている[6]。また、ポストコロナ時代における教育の新しい形態として、メタバースを活用したバーチャルとフィジカルを融合した教育環境の構築が重要であるとされている[7]。遠隔学習の普及を受け、従来の対面授業の利点とオンライン教育の柔軟性を組み合わせたハイブリッドな学習環境が求められている。メタバースを活用することで、学習者は物理的な距離に関係なく、リアルタイムでインタラクティブな教育体験を共有できるようになり、教育の公平性やアクセシビリティの向上につながると考えられる。

本研究では、MR 技術を活用し、教育現場やミュージアムにおけるインタラクティブな展示としての応用も視野に入れたシステムの開発を行う。このシステムを通じて、学習者が数学的概念を直感的かつ多角的に理解できる環境を提供し、学習意欲や関心の向上に寄与することを期待する。

2. 複合現実を活用したジェネラティブアート体験システム

本研究では、複合現実 (Mixed Reality, MR) を活用し、ユーザが直感的に操作しながら数学的概念を学習できる体験型システムを開発する。本章では、システムの構成と主要な機能について説明し、ジェネラティブアートの学習体験を促進するための設計思想を示す。

2.1 相似・比を用いたジェネラティブアート

本システムでは、Microsoft の HoloLens 2 を使用する。HoloLens 2 は、現実世界と仮想空間をリアルタイムで融合し、ハンドトラッキング、視線追跡、音声認識を通じて直感的な操作を可能にするヘッドマウントディスプレイ (HMD) 型 MR デバイスである[8]。本研究では、HoloLens 2 の以下の機能を活用する：

- ハンドトラッキング: ユーザがジェスチャーを用いて仮想オブジェクトを操作可能。
- 視線追跡: 注視する位置を認識し、システムとインタラクティブな関係を構築可能。
- 空間マッピング: 周囲の物理環境を認識し、現実空間

Bridging Real and Virtual Worlds: A Mixed Reality System for Generative Art and Mathematical Learning

†1 SHUN FURUICHI, Future University Hakodate

†2 KAORU SUMI, Future University Hakodate

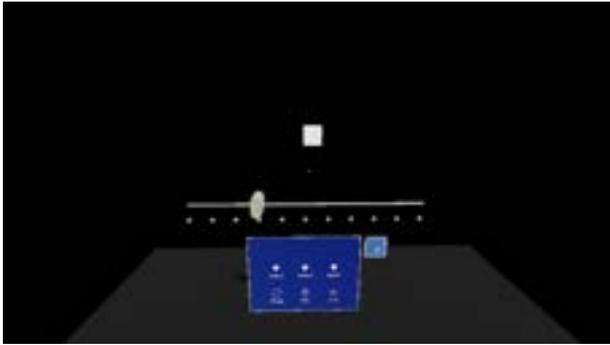


図1 相似・比を用いたジェネラティブアート (変更前)

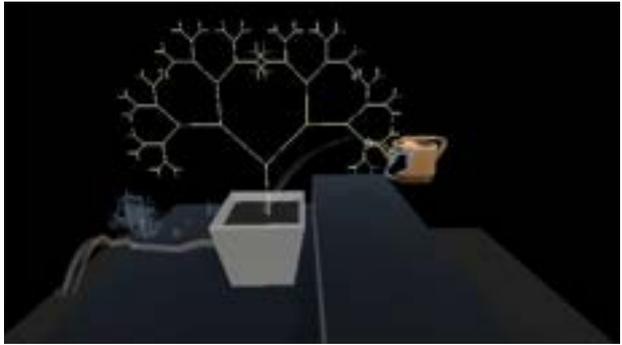


図3 フラクタルツリーを用いたジェネラティブアート

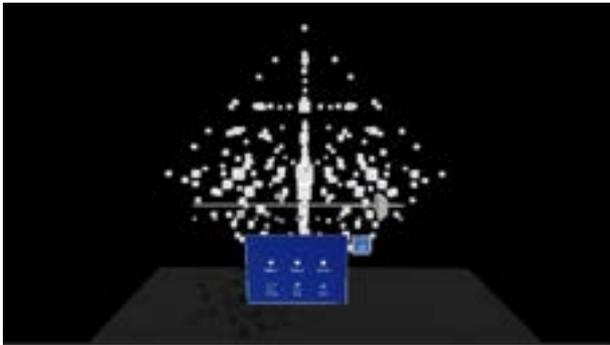


図2 相似・比を用いたジェネラティブアート (変更後)

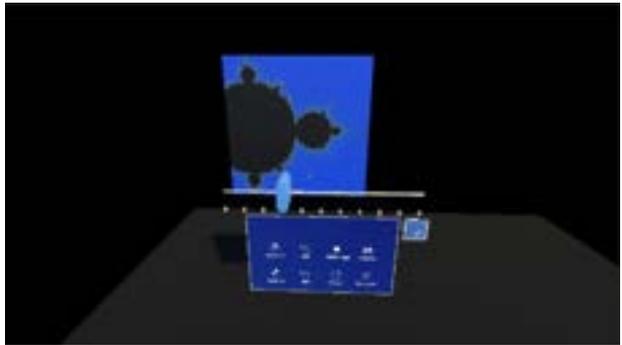


図4 マンデルブロ集合を用いたジェネラティブアート

に適応したオブジェクト配置が可能。

これらの機能を活用することで、従来の2Dスクリーン上での操作と異なり、空間的な体験を重視した直感的な学習環境を提供する。

2.2 ジェネラティブアート体験システムの概要

本システムでは、ユーザがMR空間内で数学的概念を視覚的に体験できるよう、異なる数学的手法を基にした3種類のジェネラティブアートを提供する。各システムは、数学的原理を理解しやすくするためのインタラクティブな操作機能を備えており、ユーザが能動的に探求できる設計となっている。

2.2.1 相似・比を用いたジェネラティブアート

このシステムでは、目の前に表示されたメニュー画面を操作することで、ジェネラティブアートを制御することができる(図1)。操作項目は以下の3つに分けられる：

1. 表示するキューブ数の変更
2. 中心のキューブとの相似比の変更
3. 表示しているキューブ間の距離比の変更

システムの起動直後は、図1のように一つのキューブのみが表示される。ここで、ユーザが操作を行うことで、図2のようにキューブが等間隔で増えたり、追加されたキューブが指定した相似比で配置されるなどの変化が起こる。

まず、ユーザには操作方法を説明せずにシステムを体験し

てもらい、変化するオブジェクトを観察しながら規則性を発見することを促す。その後、各操作の数学的背景(相似や比の概念)について説明し、今度は関係性を確かめる視点でアートを操作する。このような段階的な学習体験により、数学的概念の理解を深めることを目指す。

2.2.2 フラクタルツリーを用いたジェネラティブアート

このシステムでは、図1のように目の前に出現した「じょうろ」を持ち、鉢に水をやることでジェネラティブアートが出現する構造を採用している。このデザインは、フラクタルツリーの性質を直感的に学習することを目的としている。

成長プロセス：

- 初期状態では、鉢に1本の幹が表示される(図1)。
- じょうろで水をやる時、試行回数に応じて枝が増えていく(図2)。
- 各試行ごとに1本の幹が2本の枝に分岐し、このプロセスを繰り返すことで自己相似的なフラクタル構造が形成される。

ユーザが水をやる行為が試行回数の増加とリンクしており、フラクタル成長の規則性を視覚的に体験できる。また、枝分かれの角度や分岐率を調整する機能を追加し、ユーザが試行錯誤しながら数学的法則を探究できる環境を提供する。

2.2.3 マンデルブロ集合を用いたジェネラティブアート

このシステムでは、目の前に出現したパネルに表示されたマンデルブロ集合を操作することができる (図 1)。具体的には、パネル手前に操作メニューが表示されており、「横移動」、「縦移動」、「拡大」、「縮小」、「停止・再生」、「リセット」、「三次元版の表示」の 7 つの操作が可能である。

これらの操作を通じて、ユーザはマンデルブロ集合の詳細な構造を観察し、その数学的特性を体験することができる。特定の領域を拡大すると、拡大前と類似した形状が出現する自己相似性を確認できる。例えば、図 2 に示すように、拡大を続けることで同様のパターンが繰り返し現れ、数学的な無限の概念を視覚的に捉えることが可能となる。また、三次元版の表示を用いることで、マンデルブロ集合の構造を立体的に観察し、新たな視点から数学的規則性を探ることができる。

このように、マンデルブロ集合の操作を通じて、算数・数学と密接な関係にあるジェネラティブアートを体験的に鑑賞しながら、数学的特徴を発見し、認識することができる。また、この体験を通じて、数学が実生活の中でどのように活用されているかを理解し、日常に潜むフラクタル構造や相似の概念に気づく機会を増やすことを目的としている。

2.3 仮想空間と現実空間を融合したインタラクシ ョンの特長

本システムの特長の一つは、仮想空間と現実空間の統合による体験型学習環境を提供する点にある。従来の VR ベースの学習システムでは、仮想空間のみでの操作に制限されることが多かったが、本システムでは、ユーザが現実空間の動作 (例: 手のジェスチャーや視線の動き) を活用してインタラクティブに仮想オブジェクトを操作できる。特に、HoloLens 2 のハンドトラッキング機能を利用することで、実世界の動作と仮想世界の変化が直感的に結びつくため、学習者の没入感を高めることができる。

また、仮想空間内での観察視点を自由に変更できる点も重要な要素である。例えば、ジェネラティブアートを上方や側面から観察することで、三次元的な構造の理解が深まる。このように、MR の特性を活かした多角的な視点の提供は、数学的概念の理解を促進する上で有効な手段となる。

3. まとめと今後の展望

本研究では、HoloLens 2 を用いた複合現実 (MR) 環境において、ユーザがジェネラティブアートを直感的に操作・観察できる体験型システムを開発した。本システムは、「相似・比を用いたジェネラティブアート」「フラクタルツリーを用いたジェネラティブアート」「マンデルブロ集合を用いたジェネラティブアート」の 3 種類のアート体験を提供し、数学的概念を視覚的かつ体験的に学習できる環境を構築している。特に、HoloLens 2 のハンドトラッキングや視線追

跡機能を活用し、従来のスクリーンベースのインタラクションとは異なる、没入感のある学習体験を提供することを目的とした。

本システムでは、仮想空間内のオブジェクト操作を通じて数学的規則性を探求できるだけでなく、現実空間での物理的操作と連動させることにより、MR の特性を最大限に活かした学習体験を提供する。たとえば、フラクタル構造を学習するために「じょうろで水をやる」という物理的な動作を導入し、試行回数の増加がフラクタル成長と直感的に結びつくように設計されている。また、マンデルブロ集合の操作においては、三次元的な視点変更を可能にすることで、数学的パターンの視覚的理解を促進する工夫を施した。

今後の展望として、ユーザの体験をより没入感のあるものにするため、操作性に関する実験を実施する予定である。具体的には、インターフェースの配置について、仮想空間と現実空間を複合した環境に置くべきか、あるいは現実空間側に統一するべきかを検討する。この検証のため、HoloLens 2 向けのアプリケーション開発において標準的に用いられる MRTK (Mixed Reality Toolkit) のボタンやスライドバーを活用し、複合空間での操作性を評価する。また、現実空間側のインターフェースとして、センサを用いた物理的な制御装置や、キーボード・マウスによる制御などの手法を導入し、異なるインターフェースの比較調査を行う。

さらに、本システムの教育現場やミュージアムでの応用可能性についても検討を進める。特に、MR 技術を活用することで、数学的概念を理論的な学習だけでなく、視覚的・体験的に学ぶことができる新しい学習環境の構築が期待される。今後の研究では、このシステムが数学教育に与える影響を検証し、MR を活用した体験型学習の可能性をより深く探究していく予定である。

参考文献

- [1] "McCormack, J., Dorin, A., & Innocent, T. (2004). Generative design: A paradigm for design research. *Digital Creativity*, 15(1), 55-64.
- [2] Galanter, P. (2003). What is generative art? Complexity theory as a context for art theory. *Proceedings of the International Conference on Generative Art*.
- [3] 文部科学省. "国際数学・理科教育動向調査(TIMSS2019)のポイント", p.1.
- [4] 文部科学省・国立教育政策研究所. "OECD 生徒の学習到達度調査 PISA2022 のポイント", p.7.
- [5] Erbas, C., & Demirel, V. (2019). "The effects of augmented reality on students' academic achievement and motivation in a biology course." *Journal of Computer Assisted Learning*, 35(3), 450-458.
- [6] Parong, J., & Mayer, R. E. (2018). "Learning science in immersive virtual reality." *Journal of Educational Psychology*, 110(6), 785.
- [7] Wang, Y., Lee, L.-H., Braud, T., & Hui, P. (2022). "Re-shaping Post-COVID-19 Teaching and Learning: A Blueprint of Virtual-Physical Blended Classrooms in the Metaverse Era." *arXiv preprint arXiv:2203.09228*.
- [8] Microsoft. HoloLens 2: Overview and Features. <https://www.microsoft.com/en-us/hololens>, Accessed 2024-02-27.

テキスト対話における受信者感情予測の個別化

菅野光^{†1}
電気通信大学

橋山智訓^{†2}
電気通信大学

1. はじめに

近年テキストを主体とした対話（以下テキスト対話）の機会は急増しており、テキスト対話における不安な経験やトラブルの発生が増加している。総務省情報通信政策研究所 [1] によると、LINE、X (旧 Twitter) の利用率はそれぞれ 2014 年の 55.1%、21.9% から 2023 年の 94.9%、49.0% へと大幅に増加している。パイドゥ株式会社と北海道函館西高等学校の共同調査 [2] では、テキスト対話によるトラブルや不安な経験の多さ、テキスト対話の支援機能への需要の高さが示されている。

テキスト対話におけるトラブルを防止する支援方法の一つには、メッセージの送信直前に送信内容に応じてフィードバックを与え、送信者の行動変容を促す方法が考えられる (図 1)。そのためには、送信内容からそれを読む受信者の感情を送信前に予測する (以下、受信者感情予測) 必要がある。加藤らの実験 [3] では、受信者の感情の生じ方は受信者によって異なることが示されており、適切な支援の実現には受信者への個別化が重要といえる。例えば、相手からかう送信内容に対して、受信者が冗談だと解釈してポジティブに感じやすい人であるか、真に受けて不快に感じやすい人であるかによって、支援内容を変えるべきである。受信者感情予測を試みた先行研究 [4][5] では受信者への個別化は考慮されていない。本研究は、受信者に対する受信者感情予測の個別化を目指す。

2. 関連研究

2.1. テキスト対話における受信者感情予測

長谷川ら [4] は、感情ラベル付きの対話コーパスを作成し、直前 2 発話から受信者の感情を予測するモデルを学習させ評価した。対話コーパスは X(当時 Twitter) から収集した投稿や返信を元に作成されている。テストデータに対する 8 感情予測の F1 値の平均は 0.567 であった。

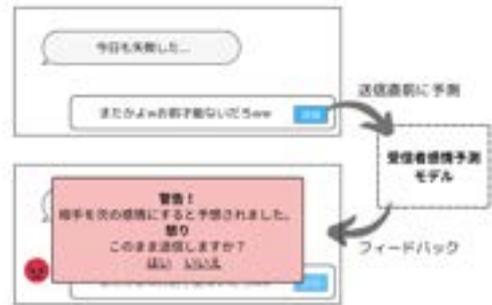


図 1 テキスト対話におけるトラブル防止支援システムのイメージ

古賀ら [5] は、テキスト対話における受信者の感情を推定するタスクを 4 種に分類し、それぞれに対応したモデルを学習させ評価した。これらのモデルは事前学習モデルを用いたファインチューニング手法で学習された。本研究が扱う受信者感情予測に該当するモデルでは、感情極性予測の F1 値の macro 平均は 0.55 程度であった。

2.2. 感情分析の個別化

鈴木ら [6] は書き手の感情分析を書き手に個別化できるモデルを提案した。このモデルは、同一の書き手の投稿を複数受け取ることで、書き手に特有の特徴ベクトルを得て推定を行う。この手法は、既存手法と異なり書き手の性格情報といった追加情報の付与が不要である上、既存手法を超える推定性能を示した。本研究ではこの手法を応用し個別化を試みる。

3. 目的

本研究では、テキスト対話における受信者感情予測について、受信者に応じた個別化手法を提案する。提案手法を評価するため、提案手法のモデルと比較用の単純なモデルで同じデータを学習させ、性能を比較する。

Personalization of Recipient's Emotion Recognition in Conversation

^{†1} HIKARU KANNO, The University of Electro-Communications

^{†2} TOMONORI HASHIYAMA, The University of Electro-Communications

4. 方法

4.1. 対話における受信者感情予測

本研究では、予測には受信者が読む送信内容とその直前 1 発話の内容を利用する。学習データは、3 発話分のテキストメッセージを元に、1 発話目と 2 発話目の送信内容及び 3 発話目に表現された感情情報で構成する。これらは、受信者が読む送信内容 (2 発話目)、その直前 1 発話の内容 (1 発話目)、ある送信内容を読む受信者の感情 (3 発話目に表現された感情) を表す。

予測は 8 感情から 1 つを選択することとする。8 感情は、Plutchik の感情の輪 [7] を元にした「嬉しい、悲しい、期待、驚き、怒り、恐れ、嫌悪、信頼」とする。

予測は深層学習モデルにより実現する。モデルの構築には事前学習済みの BERT [8] を利用し、ファインチューニング手法によって学習させる。モデルへの入力は、1 発話目と 2 発話目の送信内容 (文章) をトークン化して特殊トークン ([SEP]) で結合したトークン列とする。モデルは入力されたトークン列を特徴ベクトルに変換し、それを線形変換することで予測結果を出力する。8 感情に対する分類問題として学習させることで、予測モデルを実現できる。本研究では、事前学習済みの BERT として日本語の Wikipedia の文章を学習した bert-base-japanese-whole-word-masking*1 を使用する。

4.2. 提案手法

本研究の提案手法は、鈴木らの提案手法 [6] を応用したものである。鈴木らの手法では同一の書き手の文章をモデルに複数入力するが、提案手法では同一の受信者の対話をモデルに複数入力する。

提案手法におけるモデルの構造と順伝播の模式図を図 2 に示す。このモデルの構造は、特徴量抽出を行う事前学習済みモデルと線形変換を行う全結合層の間に Self-Attention 層を加えた構造である。Self-Attention 層とは、各入力に対し他の入力との関連性を計算し、重要な要素を強調する機構である。モデルに与えられた各対話は事前学習済みモデルによって特徴ベクトルに変換され、Self-Attention 層に入力される。Self-Attention 層での変換は同時に入力される他の入力によって決まる。すなわち Self-Attention 層に入力された特徴ベクトルは、受信者ごとに異なる変換がされる。この変換で得た特徴ベクトルを使って予測することで、受信者ごとに異なる予測ができる。

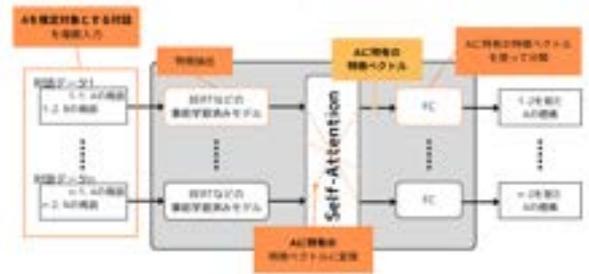


図 2 提案手法におけるモデルの構造と順伝播の模式図

5. データセットの作成

本研究でのモデルの学習には、2 人の人間同士のテキスト対話を記録したデータが必要である。特に個別化手法の評価をする観点から、対話者が識別できること、多様な対話者を含むことが必要である。一般公開されたデータセットにはこれらの条件を全て満たすものがない。そこで本研究では独自にデータセットを作成した。

データの作成には、Bluesky*2 上の投稿と返信を収集した。その結果、642 人の受信者に関する 10534 通りの対話を得た。収集したデータに対して感情ラベルのアノテーションを実施した。アノテーションには感情分析モデル luke-japanese-large-sentiment-analysis-wrime*3 を利用した。このモデルには各データの 3 発話目を入力し、得られた出力を Softmax 変換で確率分布に変換したものを感情ラベルとした。

感情分析モデルによるアノテーション結果を評価するため、収集したデータから無作為抽出した 100 通りの対話に対して、5 人の作業員によるアノテーションを実施した。作業員によるアノテーションとの比較の結果、モデルによるアノテーションで作成したデータは本研究の学習データに使用可能と判断した。

6. 実験

提案手法が個別化を実現できるかを検証するため、図 2 に示す提案手法のモデルと、Self-Attention 層を持たない比較用の単純なモデルでそれぞれ学習させ、性能比較を実施した。

学習データは 5 章で作成したデータセットを無作為に 6:2:2 の割合で訓練、検証、テスト用のデータに分割し、長さ 512 のトークン列に符号化したものを使用した。学習の設定を表 1 に示す。

*1 <https://huggingface.co/tohoku-nlp/bert-base-japanese-whole-word-masking>

*2 <https://bsky.app>

*3 <https://huggingface.co/Mizuiro-sakura/luke-japanese-large-sentiment-analysis-wrime>

表1 各モデルの学習の設定

設定	比較モデル	提案モデル
loss func.	Weighted Cross Entropy	
batch size	32	4*4
optimizer	AdamW	
weight decay	0.1	0.1
warmup steps	1epoch	1.5epoch
lr	3e-5	3e-5
max epochs	10	15
dropout	0.1	0.1

7. 結果と考察

比較モデルの学習は6エポックで早期終了し、提案モデルの学習は15エポックで終了した。検証データへの損失は、比較モデルでは3エポック目、提案モデルでは7エポック目が最小となった。テストデータによる各モデルの性能評価結果を表2に示す。

表2 評価結果

指標	比較モデル	提案モデル
accuracy	0.54	0.55
macro-F1	0.46	0.43
weighted-F1	0.54	0.55

表2を見ると、提案手法によるモデルの性能は、比較用の単純なモデルの性能と大きな差が見られない。したがって、本研究の提案手法は今回の実験では個別化を実現できなかったと結論づけた。

個別化が実現できなかった原因として、受信者1人あたりのデータ数が少なかったことが挙げられる。鈴木ら[6]は書き手1人あたりの投稿数が600件以上となる前処理をしていたが、今回使用したデータでは受信者1人あたりのデータ数は平均16.4件であった。

学習の様子として、バッチサイズが小さいほど学習が不安定になり、学習後の性能がバッチサイズに大きく依存していた。これは、今回の学習データが不均衡であったことが原因と考えられる。今回は損失関数の重みを調整することで学習の安定化を図ったが、それだけでは不十分であったと考える。

*4 提案モデルの学習には8通りの対話のデータを1単位にまとめたデータを使用している。バッチサイズ4のミニバッチは32通りの対話のデータを持つ。

8. 今後の課題

8.1. 受信者群に対する個別化

個別化実現のための改善案として、受信者をクラスタリングし、似た特徴を持つ受信者群への個別化を試みるべきだと考える。受信者1人あたりのデータが少なくても受信者1群あたりのデータは十分に確保できる。また加藤らの実験[3]では、受信者の感情の生じ方は受信者の感情特性により分類した3群間で異なることが示されている。この実験における感情特性とは、特定の感情経験の頻度または特定の感情反応の閾値の低さに関する、長期的に安定した個人の傾向または特性を指す。感情特性により分類した受信者群に対して提案手法を適用することで、個別化の実現が期待できる。

感情特性により受信者を分類するには、SNS上の過去の投稿内容を感情分析するつもりである。そうすることで例えば、喜びを表現した投稿が多い群、怒りを表現した投稿が多い群、などと受信者を分類できる。今回は同一の受信者の対話で入力を構成したが、今後は同一の受信者群の対話で入力を構成し個別化を試みたい。

8.2. 感情ラベルの形式変更

感情ラベルの形式に関して、モデルの性能向上に向けた改善案を以下に列挙する。

- いずれの感情にも属さない中間状態を追加する
- 似た感情同士が隣合うような間隔尺度で表現する
- 複数選択を許容する (マルチラベル分類)

特定の感情に属さない中間状態を追加することで性能向上が期待される。古賀らの研究[5]では「positive, negative, neutral」の3分類で学習させた。3分類のmacro-F1とneutralを除く2分類のmacro-F1(macro-F1 without neutral)を比較すると3分類のmacro-F1値の方が高い結果が示された。本研究と同条件では、macro-F1が約54%、macro-F1 without neutralが約38%であった。

本研究のテスト結果の評価では、モデル出力と正解ラベルが一致すれば正解、不一致ならば不正解としていた。しかし実用上は、不正解だが正解と似た感情を答える場合(「嬉しい」に対し「期待」など)はさほど問題でなく、全く異なる感情を答える場合(「嬉しい」に対し「怒り」など)は大きな問題となる。似た感情同士が隣合うような間隔尺度で感情ラベルを表現すれば、損失関数に正解との間隔に応じた重みを付けられる。そうした重みを持つ損失関数を学習に用いれば、たとえ不一致でもなるべく似た感情を予測するようにモデルを訓練できる。

本研究では8感情から1つのみ選択していた(シングルラベル分類)が、複数の感情状態の選択を許容すること(マルチラベル分類)

チラベル分類) でより適切な予測ができると考える。実際の受信者の感情は常に1つとは限らず複数の感情を同時に抱くこともあり得る。そうした場合に単一選択の予測で正解するのは難しく、複数選択の予測によって正解や部分的な正解が容易になる。

8.3. 不均衡データへの対応

不均衡な感情ラベルを持つデータの学習を安定化させるために本研究では損失関数の重みを調整したが、それだけでは不十分であった。アンダーサンプリングやオーバーサンプリングによりデータの不均衡さを是正したり、異常検知問題としての学習を試してみたい。小さなバッチサイズでも安定した学習が可能になれば、少ない計算量での学習が可能となる上、性能向上が期待できる。

8.4. 事前学習済みモデルの変更

本研究で用いた BERT(bert-base-japanese-whole-word-masking) は Wikipedia の文章で事前学習されたモデルである。SNS 上の投稿を事前学習したモデルを使えば、くだけた表現に対しても適切に特徴抽出でき、高い性能を発揮する可能性がある。SNS 投稿のテキストを事前学習したモデルには、hottoSNS-BERT[9] や JTweetRoBERTa[10] などがある。

対話における感情分析を想定した事前学習済みモデルを使うことも有効だと考える。例えば BERT-ERC[11] などは発話間の関係について特徴抽出できるなどの特徴がある。こうしたモデルにより対話の文脈情報まで捉えることで性能向上が期待できる。

8.5. 実用上の性能評価

本研究はモデルの性能をテストデータによって評価したが、テストデータに対する性能が、実用における汎化的な性能を示すとは限らない。今回のテストデータは Bluesky 上で収集したデータに限定されており、Bluesky 以外での対話に対する性能は示せない。また今回のテストデータは感情分析モデルによるアノテーションで作られたもので、実際の受信者の感情とは異なる。

実用的な汎化性能の測定するには、受信者本人の感情が付与されたテストデータが必要である。そうしたデータは、対話の各時点での感情を記録してもらったデータや自身の過去の対話にアノテーションしてもらったデータを収集すれば用意できる。しかしデータの偏りを減らしつつ効率良く大量のデータを収集するには、自動的に収集できる方法の検討が課題となる。

9. まとめ

本研究は、テキスト対話における受信者感情予測の個別化手法を提案した。独自にデータセットを作成し提案手法を評価した結果、提案手法は今回の実験では個別化を実現できなかった。個別化実現のためには受信者をクラスタリングし受信者群への個別化を試みる等の工夫が必要である。

参考文献

- [1] 総務省：令和5年度情報通信メディアの利用時間と情報行動に関する調査報告書の公表 (2024). https://www.soumu.go.jp/menu_news/s-news/01iicp01_02000122.html.
- [2] バイドゥ株式会社：北海道函館西高等学校探究チーム「ぶなしめじ君の冒険」と実施のテキストコミュニケーションの課題に対処するための調査報告 (2024). <https://prtimes.jp/main/html/rd/p/000000821.000006410.html>.
- [3] 加藤由樹, 加藤尚吾, 杉村和枝, 赤堀侃司：テキストコミュニケーションにおける受信者の感情面に及ぼす感情特性の影響：電子メールを用いた実験による検討, 日本教育工学会論文誌, Vol. 31, No. 4, pp. 403-414 (オンライン), 10.15077/jjet.KJ00004964312 (2008).
- [4] 長谷川貴之, 鍛冶伸裕, 吉永直樹, 豊田正史：オンライン上の対話における聞き手の感情の予測と喚起, 人工知能学会論文誌, Vol. 29, No. 1, pp. 90-99 (オンライン), 10.1527/tjsai.29.90 (2014).
- [5] 古賀友里愛, 神藤駿介, 宮尾祐介：対話における発話に反映されない聞き手の感情推定, 人工知能学会全国大会論文集, Vol. JSAI2024, pp. 4Xin246-4Xin246 (オンライン), 10.11517/pjsai.JSAI2024.04Xin246(2024).
- [6] 鈴木陽也, 山内洋輝, 梶原智之, 二宮 崇, 早志英朗, 中島悠太, 長原 一：書き手の複数投稿を用いた感情分析, 人工知能学会全国大会論文集, Vol. JSAI2024, pp. 3Xin2104-3Xin2104 (オンライン), 10.11517/pjsai.JSAI2024.03Xin2104(2024).
- [7] PLUTCHIK, R.: Chapter 1 - A GENERAL PSYCHO-EVOLUTIONARY THEORY OF EMOTION, *Theories of Emotion* (Plutchik, R. and Kellerman, H., eds.), Academic Press, pp. 3-33 (1980).
- [8] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (2019).
- [9] Sakaki, Takeshi, Mizuki, S., Gunji, N.: BERT Pre-trained model Trained on Large-scale Japanese Social Media Corpus, <https://github.com/hottolink/hottoSNSbert> (2019).
- [10] 高須遼, 狩野芳伸：JTweetRoBERTa: 大規模 SNS 投稿テキストによる事前学習と各種タスクによる性能検証, 言語処理学会 第 30 回年次大会 発表論文集, 言語処理学会 (2024).
- [11] Qin, X., Wu, Z., Cui, J., Zhang, T., Li, Y., Luan, J., Wang, B. and Wang, L.: BERT-ERC: Fine-tuning BERT is Enough for Emotion Recognition in Conversation (2023).

応用行動分析学を活用して組織の行動変容につなげる

是村由佳^{†1}

コレムラ技研バラスト事業部^{†1}

1. はじめに

行動変容は、あらゆる人を対象に、ライフステージのあらゆる場面で確認できる。また、行動変容を測定する方法や確認する方法もさまざまである。

本発表では、組織における行動変容の1つの方法として、応用行動分析学に基づいたPDCAのエンジニアリングサイクルの流れを紹介する。

2. 行動分析学の基礎と応用

応用行動分析学は、B. F. スキナー[1]が始めた心理学の1領域である行動分析学の発展に伴い始まった。行動に関する実験的研究は、1958年に刊行開始したJournal of the Experimental Analysis of Behavior (JEAB) [2]で発表されてきたが、実験室実験に加え、実験室から飛び出し、病院、企業、学校、施設、コミュニティ、家庭など実社会で起きる行動の問題を解決するための応用研究も盛んになった。

1968年にJournal of Applied Behavior Analysis (JABA) [3]が刊行され、二誌は行動分析学に関する研究の基礎と応用の両輪を担った。我が国では、1987年に『行動分析学研究』が刊行された (Fig.1)。

行動分析学の研究は、基礎と応用が両輪となる。いずれ



Fig. 1 行動分析学の学術誌 (左より Journal of the Experimental Analysis of Behavior, Journal of Applied Behavior Analysis, 行動分析学研究)

も、行動の定義、「行動は環境との相互作用の中で生じる」とし、環境を形成する刺激を制御して、

相互作用の関係性を明らかにする。そして測定を一貫して行い、数値化し、なるべく儉約的に、再現可能な方法を追求する。

2.1 スキナーの実験

2.1.1 実験装置

環境刺激との相互作用の中で生じる行動について、B.F. スキナーが開発した動物実験装置により研究が進められた ([4] 参照)。この実験装置は、オペラント箱やスキナー箱 (Skinner Box) と呼ばれる (Fig.2)。ハトやラットを主な被

験体としていたが、後にヒトを含むさまざまな有機物はその行動の特性を踏まえて作製されたオペラント箱を使用した。

この実験装置の中で被験体は、反応物に反応する。そうすると、あらかじめ設定されたプログラム

(強化スケジュール)により、強化子 (反応を維持する刺激) を得る。

例えばハトが被験体の場合、キー (反応物) にキーつき反応 (行動)

をすると餌の穀物 (強化子) を得る。反応 (行動) は、Fig.2 のオペラント箱

上にある装置 (累積記録器) で、反応と強化子の提供がリアルタイムに記録された。この記録は「累積グラフ」と言う (Fig.3)。この実験装置で様々な「強化スケジュール」を実施して行動の変容を検証した。

2.1.2 強化スケジュール

例えば初めてオペラント箱に入った被験体のハトが、ウロウロしながら箱の内壁をつつくうちに、偶然キーをつつくオペラント行動をした。そうしたらガチャンと音がして、キー下の開口部に穀物があることに気づき、食べた。またキーをつついたら穀物が出てきた。オペラント箱にいる間中、キーをつつくごとに開口部に穀物が提示されたら、これは連続強化 (continuous reinforcement, CRF) または固定比率スケジュール1 (fixed ratio schedule 1, FR1) という。通常、反応のパターンが安定するまで同じスケジュールで実験をする。

強化スケジュールは、固定比率、固定変率、変動比率、変動変率などがあり、さらに強化スケジュールを組み合わせるなどより複雑なスケジュールもある。ここでは最もシンプルなスケジュールの紹介にとどまるが、詳しくは、[5]、[6]などを参照してほしい。



Fig. 2 オペラント (スキナー) 箱 <https://www.youtube.com/watch?v=fERfERCCW4Y&t=70s> よりスクリーンショットにて引用

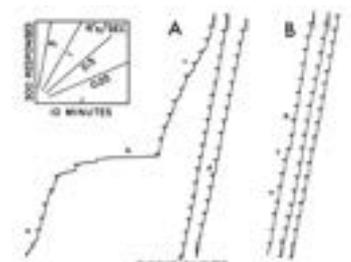


Fig. 1 累積グラフ [5] Fig. 14 より引用

^{†1} YUKA KOREMURA, Koremura Giken ballast

2.1.3 効果の検証 (消去)

強化スケジュールの効果は、その強化スケジュールがなくなった時にわかる。消去抵抗が生じるからである。ここで使用する「消去」という言葉は、特定の手続きにより生じた現象のことを指す。被験体のハトがある強化スケジュールで安定した行動パターンでキーつつきをしていたところ、ある時から強化子である餌を得られなくなってしまった。そしてその状態が続いた。これが「消去」である。

例えば、CRF (FR1) から FR5 へ、FR5 で反応が安定したら、FR10 へ、と強化子を得るまでの反応数 (固定比率) を増やす。個体差はあるが、固定比率の数字が大きいほど反応比率が大きくなる。そして、消去前の固定比率の数字が大きいほど、完全に反応が無くなるまで (消去) の抵抗が大きい。

3. 応用行動分析学

3.1 実験室実験から現場への systematic replication

実験室という制御された環境では研究を行わない応用行動分析学では、実験室で行う実験と同じ行動の定義で研究をするが、環境の制御が難しい。環境との相互作用による行動の関係性が複雑であることが大きい。

行動分析の応用研究について、Baer ら[7]は、[3]の第 1 巻第 1 号で、行動分析学を応用する際に、応用行動分析学の価値を定義し判断する 7 つの本来の基準 (primary criteria) があるとした。以下が、各基準である。①応用：社会的に重要な行動であること。②行動：改善を必要とする行動そのものを測定すること。その測定が主観的なものではなく客観的な方法により記録されていること。③分析：実験者が操作した変数を分析すること。④テクノロジー：変えようとする行動に関する研究手順のすべてが明らかにされ、再現性の担保を図ること。⑤概念体系：概念的に体系的であることとして 1) 同じ原理を使う、2) 「コツの寄せ集め」にしないこと。⑥有効性：応用においては、理論的に重要かは問題ではなく、実用的価値を認めるか、臨床的ないし社会的に意義のある行動変容を行うこと (本来変化すべき行動がどうなったのか? 一定の効果を上げたものの社会的妥当性の基準はクリアしているか?)。⑦適切に般化した結果 (一般性)：行動の変化は一般性を持っているか否か。

以上を踏まえ、筆者が組織の行動変容に際し、組織という環境の制御をどのように捉えているのかについて次に述べる。

3.2 環境の制御

応用行動分析学の研究の環境制御の 1 つのアプローチとして、Gilbert の「Subject matter (主題) は何か」という問いに答えることにより、どのような文脈で、行動変容をしたいのか、また、行動変容の目的は何かを明らかにするという[8]。これは、対象が 1 人の行動の変容の場合の考え方である (Fig.4 左)。

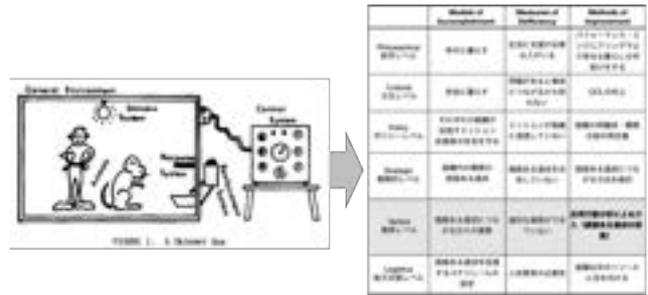


Fig. 2 環境の制御

組織行動の変容に際しては、組織を 1 つのオペラント箱として捉える。勿論、物理的な装置ではないが、Gilbert の 4 つの Leisurely theorems を全て包含する Performance Matrix (Fig.4 右[9]) を「組織」というオペラント箱と捉え、第 4 理論で言及する「6 つの視点」を、このオペラント箱の環境を制御する要素 (刺激) として活用する。最もアブストラクとな 1) 哲学的視点から (Philosophical), 2) 文化的価値観を踏まえたゴール (Cultural), 3) ゴールの範疇で掲げる組織のミッション (Policy), 4) ミッションを達成するための組織内の役割と職務 (Strategic), 5) その役割を担う者が日々行う業務遂行行動 (Tactics), 6) 業務が滞りなく遂行されるような後方支援 (Logistics) である。これらの視点について、組織が達成すべき価値あるパフォーマンスの観点から目標を設定し (Models of Accomplishment), それぞれの視点に対して現状把握と目標との乖離を調べ (Measures of Deficiency), 乖離があれば、改善をする (Methods of Improvement) PDCA サイクルを回す。この PDCA サイクルは、応用行動分析学の実験デザインの基本的な考え方と同じである。

3.3 現場の評価と行動変容のための介入とその評価 (実験デザイン)

行動変容を具体的にどのように進めるかについて考えた時に、まず変容したい行動を定義する。その行動をしている人がどのような環境に置かれているかを考え、同時に客観性、再現性、一貫性などを考慮して、誰もがその行動を同じように測定できるかをその行動の定義として落とし込み、具体的にどのように測定するかを決める。

落とし込まれ

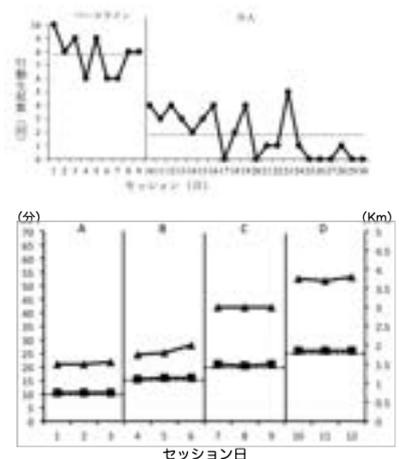


Fig. 3 実験デザイン例

た行動の現状把握（行動観察）をする。ベースライン (baseline) という期間である。変容したい行動が実際に、どのような関係性の中で、どれくらい生起しているのかを把握しなければ、その後に行う介入が効果的であったか否かを評価できないからである。

現状把握をしたら、どのような手続きで介入することが効果的なのかを検討する。ベースライン期に計測している行動データが最終的に安定傾向を示すまで、行動観察をするたびに分析を重ねる。

決定した手続きを実施し、行動変容の記録をする。介入 (intervention) という期間である。介入を実施する人が複数いる場合は、一貫した手続きで介入を行うように手順を共有する。Fig.5 ではこの一連をグラフ化した行動変容の例を2つ示した。上は、自分の爪のささくれを剥いてしまう行動を減らしたグラフ、下は、少しずつジョギングの時間と距離を伸ばして

ゆくというグラフである。また、インターネット上での研修事前動画が研修前どのように視聴されたかを累積で表したグラフを Fig.6 に示した。

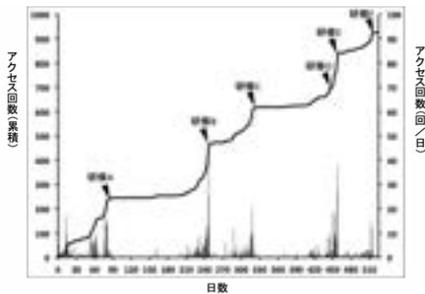


Fig. 6 累積グラフによる介入の一例

3.4 行動の機能分析と分析ツール

3.4.1 レスポンデント行動とオペラント行動

応用行動分析学で行動の変容を考える際には、オペラント行動だけではなく、レスポンデント行動が生起しているかもしれないことについても知っておく必要がある。レスポンデント行動とオペラント行動は、刺激（環境）との関係性が異なっているからである。Fig.7 が示す通り、レスポンデント行動には、行動に対するフィードバックはない。行動変容には、行動の後の刺激が行動に作用していることが必要である。オペラント行動の定義は、「行動は環境との相互作用の中で生じる」ので、行動に対するフィードバックがある。

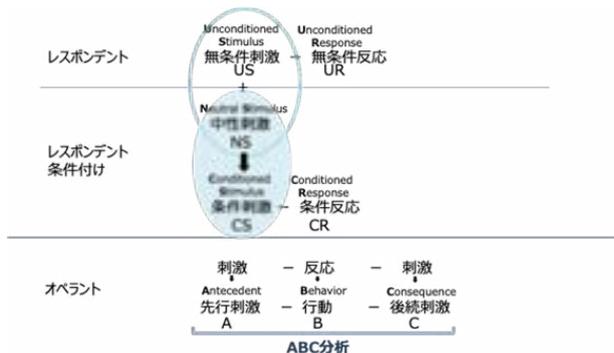


Fig. 7 レスポンデント行動とオペラント行動

3.4.2 ABC 分析

行動の後の環境刺激との関係もあるオペラント行動について、行動の後に、その行動をした人にとって何が伴っていたか（良かったのか、悪かったのか、何ともなかったのか）によってその後の行動が増減することがオペラント箱の実験により検証されている。Fig.8 に行動と行動の前後の環境刺激の関係性を図に示した。先行刺激 (Antecedent)、行動 (Behavior)、結果 (Consequence) の頭文字をとり、ABC 分析という。結果から行動への矢印がポイントで、行動の結果からのフィードバックによりその後の行動に影響する。



Fig. 8 ABC 分析

ABC 分析からわかるパターンは、行動に伴った結果が何もない場合（消去）を含め、5つの機能に分類される。

- (1) 行動に伴った結果がその行動をした人にとって好ましいもの・出来事である場合、行動が増加する（またその行動をする）現象である（正の強化）。
- (2) 行動に伴った結果がその行動をした人にとって嫌な物や出来事をなくした場合、行動が増加する（またその行動をする）現象である（負の強化）。
- (3) 行動に伴った結果がその行動をした人にとって好ましいもの・出来事を失う場合、行動が減少する（その行動はしなくなる）現象である（負の弱化）。
- (4) 行動に伴った結果がこの行動をした人にとって嫌なもの・出来事である場合、行動が減少する（その行動はしなくなる）現象である（正の弱化）。
- (5) その行動をした人にとって行動に伴った結果が何もない場合、しばらく何らかの結果を求めてその行動をするかもしれない。いくら行動しても、その行動に伴った結果が一貫して何もない場合はパタリとしなくなる（行動が増加した後に減少する）現象である（消去）。

行動を変容するために、ベースライン期に対象とする行動はどのパターンであるかを知る必要がある。その行動が起きている場面を観察し（行動観察）、その通りに記述して、行動の機能分析をする。実際に動画から行動を観察し、対象者の行動の ABC 分析の1部を Fig. 9 に示す。

A	Antecedent	B	Behavior	C	Consequence
座りながら体を使って感情を表現している					
職員Aが目の前で上着を着ている（外出の準備）	1分40秒	相んでいる手を解き、左手で目の前にいる職員に向かって振り下ろし、嬉しそうに両手を腰の上で重ねる…①		職員Aは上着をきている	
外での準備、職員Aの外出の準備	1分47秒	後ろの外の様子を見た後、相んでいる手を解き、目の前にいる職員に向かって唾を吐き、笑顔で両手を腰の上で重ねる…②		準備が続いている	
職員Aが車椅子を止める	1分48秒	左側に立つ職員Aの右腕を左手で掴んで引っ張る…③		職員Aが操作を終え、車椅子を再び押す	

Fig. 4 ABC 分析の一例

4. 応用行動分析学に基づいた組織の行動変容

これまでをまとめ、応用行動分析学の7つの本来的基準[7]を達成するために、以下に示す7段階のPDCAプロセスを、行動変容のエンジニアリングサイクルとして提案する。以下を繰り返すことにより、行動は変容し続ける。

- 1 変容すべき行動を定義する。
- 2 定義された行動の計測方法、計測したデータのグラフ表記方法を想定する（介入デザイン）。
- 3 行動データを収集する。
- 4 収集した行動データの分析（どのような環境との相互作用で生じているのか、そのような関係性か）を行う。
- 5 再現可能な行動変容の介入の手立てを考える。
- 6 実際に介入を実施する。
- 7 介入の効果を検証する。

特に組織の行動変容については、環境の制御に関しては3.2 で言及した Performance Matrix を構成する「6つの視点」をガイドにして、変容すべき行動を定義する際に、なるべくそれぞれの組織が、組織の目標を達成するために本当に変容したい行動を選定する必要がある。

5. おわりに

さまざまな行動変容の方法のうち、応用行動分析学による組織における行動変容の方法について、筆者が行動変容を計画する際の考え方と流れをまとめた。7つの本来的基準[7]は全て重要であるが、筆者は特に「④テクノロジー：変えようとする行動に関する研究手順のすべてが明らかにされ、再現性の担保を図ること」を重んじ、科学的な根拠を積み重ねることに貢献できたらと考えている。

応用行動分析学の行動変容のアプローチを取り入れようとする他学領域の研究者や、実際に行動変容をしたいと考える個人や組織の方に行動分析学が学際領域にどのように貢献できるのかを含め、今後も研鑽してゆきたい。

参考文献

- [1] Skinner, B. F. (1938/2019). "The behavior of organisms: An experimental analysis. BF Skinner Foundation."
- [2] "Journal of the Experimental Analysis of Behavior". <https://onlinelibrary.wiley.com/journal/19383711>, (参照 2025-02-16).
- [3] "Journal of Applied Behavior Analysis". <https://onlinelibrary.wiley.com/journal/19383703>, (参照 2025-02-16).
- [4] Ferster, C.B., & Skinner, B.F. (1957). "Schedules of reinforcement. Englewood Cliffs, NJ: Prentice-Hall."
- [5] ジェームス E. メイザー (2008) (磯他訳) メイザーの学習と行動日本語第3版 二瓶社.
- [6] クーパーJO, ヘロンTE, 他: 応用行動分析学. 中野良顯 (訳), 明石書店, 2013.
- [7] Baer, D. M., Wolf, M. M., & Risley, T. R. "Some current dimensions of applied behavior analysis". <https://doi.org/10.1901/jaba.1968.1-91>, (参照 2025-02-16).
- [8] Gilbert, T. F. (1969). Some issues in mathematics. I. Saying what a subject matter is. NSPI Journal, 8(2), 4-19.
- [9] Gilbert, T. F. (1978/2013). Human competence: Engineering worthy performance. John Wiley & Sons.