

WISS 2024



第32回 バンコクにて開催される国際会議（WISS）

WISS2024 開催にあたって

塚田 浩二 *

本論文集は、日本ソフトウェア科学会インタラクティブシステムとソフトウェア（ISS）研究会が主催し、2024年12月11日（水）から12月13日（金）にかけて、苗場プリンスホテル（新潟県南魚沼郡湯沢町）で開催される「第32回インタラクティブシステムとソフトウェアに関するワークショップ」（Workshop on Interactive Systems and Software: WISS2024）での発表をまとめたものです。

WISSは、伝統的に2泊3日の泊まり込み形式で開催されており、例年150名以上の参加者が朝から深夜まで深い議論を行ってきました。Covid-19のために、オンライン／ハイブリッド開催された年もありましたが、今年のWISS2024では240名以上が参加して現地開催されます。

WISS2024のテーマは、「開かれたWISS」です。これまでのWISSでは、議論を活性化するための様々な工夫が行われてきましたが、その弊害として発表形式が複雑化してしまい、常連でも混乱するような状態になっていました。そこで、WISS2024では以下のように発表形式を整理しました。

- 登壇発表の種別（ロング枠／ショート枠／議論枠）を廃止し、ショート枠相当に統一する。
- ロングティザーフォーマットを廃止する。代わりに、登壇発表を不採択となった場合、デモ発表への優先投稿制度を設ける。
- 國際学会・國際論文誌採択論文発表を名称変更し、國際学会招待発表とする。デモ・ポスター発表についても、デモ発表とする。
- WISS Challengeを正式な発表カテゴリとし、予稿集にも掲載する。

WISSの「顔」となるWebページも整理し、全ての発表形式／締切を一覧可能にしたり、デモ発表プログラムに代表画像を追加する等、常連以外の参加者にもWISSの魅力を伝える工夫を行いました。

登壇発表には、各論文に対して4名の査読者（内プライマリ査読者1名）を割り当てて査読を行い、採否判定会議での議論を経て、36件の投稿に対して21件の論文を採択しました。一部の論文は条件付き採録とし、査読者と著者による修正プロセスを経

て、論文の完成度を高めています。登壇発表の論文は、本予稿集に採録理由と併せて掲載しています。

国際学会招待発表には、6件の投稿があり、いずれにも利害関係のない4名の審査員による審査を行いました。その結果、4件の登壇発表／1件のデモ発表が招待発表として実施されます。

デモ発表については、近年締切の前に発表枠が埋まる問題が続いており、今年は約3週間前に埋まってしまいました。そこで、募集時の発表枠の制限を撤廃して、(1) デモ発表の申し込み件数に応じてデモ会場を増設する、(2) 増設で対応できない程多くの投稿があった場合、追加審査を行い採否を決定する、という方針に転換しました。その結果、今年は会場の増設のみで対応でき、108件のデモ発表を採択しました。これと先の1件の招待発表に加えて、17件の登壇デモ（登壇発表される論文のデモ発表）や、9件のスポンサー展示を含めて、3日間で135件の発表が行われます。

WISS Challengeは、学会での議論等を活性化するシステムを募集するユニークなカテゴリですが、近年は応募が少ない状態が続いていました。そこで、WISS2024では、前述したような工夫を行い、Webページでも存在感を高めることで、7件の投稿が集まり、審査の結果、6件の発表が行われます。

さらに、今年はWISS30周年企画として、WISSの立ち上げから関わる、現在も第一線で活躍される3名の先生方（増井俊之氏、曽根純一氏、小池英樹氏）をお招きした特別講演も行います。先生方の発表・議論を通してWISSの30年を振り返り、その革新性を未来に繋いでいければと思います。

私事になりますが、修士2年の時（2001年）にWISSに初参加し、登壇発表／デモ発表を通じた議論や、ナイトセッション等を通して多くの個性的な研究者と交流する中で、研究の面白さ・奥深さを感じました。当時は某メーカーに就職するつもりでしたが、道を踏み外して（？）、研究者となる大きなきっかけとなりました。このWISS2024での体験／出会いが、多くの人に価値あるものになればと思います。

最後に、登壇発表をはじめとした全ての発表にご投稿頂いた皆様、スポンサー企業の皆様、協賛学会関係者の皆様、WISS2024のプログラム委員／運営委員の皆様に深く感謝いたします。

* 公立はこだて未来大学、WISS2024 プログラム委員長

WISS 2024 プログラム委員長
塚田 浩二 (はこだて未来大)

WISS 2024 実行委員長
渡邊 恵太 (明治大)

WISS 2024 副実行委員長
松村 耕平 (立命館大)

WISS 2024 運営委員

会計	丸山 一貴 (明星大) 松村 耕平 (立命館大)
出版・ WISS Design	石井 紗郁 (NTT) 越後 宏紀 (ソフトバンク)
広報・ 学会リエゾン	辻田 眞 (無所属)
WISS Challenge	門本 淳一郎 (東大)
会場	佐藤 俊樹 (北陸先端大)
投稿システム	三浦 元喜 (千葉工大)
参加者管理・ 交流	西田 健志 (神戸大)
動画生中継	高田 峻介 (神戸高専) 藤田 和之 (東北大)
デモ・ポスター	高橋 治輝 (立命館大) 鳴海 紘也 (慶應大) 渡邊 拓貴 (はこだて未来大)
ネットワーク	奥本 隼 (TwoGate) 米田 英正 (Helpfeel)
Web・ サーバー	山本 和彦 (ヤマハ) 福里 司 (早大)
投票	高橋 亮 (東大)
懇親会	五十嵐 悠紀 (お茶大) 吉田 博則 (はこだて未来大)
表彰	栗原 一貴 (津田塾大) 三上 浩司 (東京工科大)
企業 スポンサー	秋山 耀 (チームラボ) 清木 昌 (ほぼ日)
学生 ボランティア	宮藤 詩緒 (東京科学大) 大西 鮎美 (神戸大)

表紙デザイン

穂積 佳 (はこだて未来大)

WISS 2024 プログラム委員

秋山 耀 (チームラボ)
五十嵐 悠紀 (お茶大)
石井 紗郁 (NTT)
越後 宏紀 (ソフトバンク)
大西 鮎美 (神戸大)
奥本 隼 (TwoGate)
門本 淳一郎 (東大)
栗原 一貴 (津田塾大)
小池 英樹 (東京科学大)
後藤 真孝 (産総研)
坂本 大介 (北海道大)
佐藤 俊樹 (北陸先端大)
志築 文太郎 (筑波大)
清木 昌 (ほぼ日)
高田 峻介 (神戸高専)
高橋 治輝 (立命館大)
高橋 亮 (東大)
塚田 浩二 (はこだて未来大)
辻田 真 (無所属)
寺田 努 (神戸大)
鳴海 紘也 (慶應大)
西田 健志 (神戸大)
福里 司 (早大)
藤田 和之 (東北大)
松村 耕平 (立命館大)
丸山 一貴 (明星大)
三浦 元喜 (千葉工大)
三上 浩司 (東京工科大)
宮藤 詩緒 (東京科学大)
山本 和彦 (ヤマハ)
吉田 博則 (はこだて未来大)
米田 英正 (Helpfeel)
渡邊 恵太 (明治大)
渡邊 拓貴 (はこだて未来大)

ISS 研究会連絡先

日本ソフトウェア科学会
「インタラクティブシステムとソフトウェア研究会」
主査: 塚田 浩二
公立はこだて未来大学
〒041-8655 北海道函館市亀田中野町 116 番地 2
E-mail: tsuka@mobiuitous.com

本研究会に関する情報は <https://www.wiss.org/> からアクセス可能です。

もくじ

まえがき: WISS2024 開催にあたって
塚田 浩二 i

セッション1：XR

ParaSights：両眼視野闘争によって2つの環境と並行してインタラクションできる空間提示手法
山本 航世, 渡邊 恵太 1

本研究は、HMD の左右の眼にそれぞれ異なる1人称視点映像を提示して両眼視野闘争を生じさせ、それらの2視点の映像の見え方を特定の視点変更操作を用いることなく変化させることで、2つの異なるVR空間と並行してインタラクション可能にすることを目指したシステムを提案している。2視点の見え方をユーザの手の動きによって自然に切り替えられるインタラクション設計は面白く、新規性が高い。一方で、まだ限られた映像セットでの実装に留まっている点やフォーマルなユーザ評価が実施されていない点等、今後に向けた課題は残り、これらの点について WISS での議論が期待できる。以上の理由から、条件付き採録と判断された。

HidEye：片目を隠す動作によるHMD用インタラクション手法の提案
伊勢 隆之介, 塚田 浩二 9

パススルー表示対応のMRゴーグルに搭載された、両眼カメラセンサの片目を手で塞ぐジェスチャを用いて、VR / MRコンテンツとパススルー表示や、コンテンツ自体の表現をシームレスに切り替える手法を提案している。既存のMRゴーグル上にて提案手法の提案、実装、評価を行っており、その認識精度は高い。いくつかの応用例も述べられており、WISS での議論すべき論文であると判断し、採択に至った。

高速投影による運動物体の残像色制御手法
幸谷 有紗, 宮藤 詩緒, 小池 英樹 16

本研究では、高速プロジェクタを用いて運動物体の残像の色を制御する手法を提案されている。既存手法との比較や人間の被験者に対する評価実験が不足している点など懸念はあるものの、アイディア自体が興味深く、WISS での活発な議論が期待されることから採録と判定された。

机上に投影した空中立体映像による方向指示の精度調査
松浦 向日葵, 福地 健太郎 23

本論文は著者らが提案するAIRR方式を応用した空中立体像提示手法を用いて方向指示の精度評価を行ったものである。提案するディスプレイにおいて2D表示と3D立体表示で方向指示を行った際の精度の比較は試みとして新規性があると判断された。行われている実験については、方向指示の必要性等の前提条件や提案する空中立体像提示手法の視認性についての指摘、追加の統計解析の必要性の指摘等があったため、委員による議論の結果、これらの修正を条件とした条件付き採録と判断された。

セッション2：音

Whisphone: ささやき声で入力できるイヤホン

福本 雅朗 30

シンプルな構成でイヤホンにおけるささやき声での音声入手法を実現した点を、全ての査読者が高く評価した。実験は1名のみだが、論文中の議論において幅広く網羅的な議論を行っている点も興味深く、採択にふさわしい論文であると判断された。

楽器未経験者のための弦管打複合電子楽器の開発

矢田 紘理奈, 岩井 将行 38

本研究の「従来楽器を始めるきっかけとなる楽器」という位置付けで複合楽器を提案する方向性は興味深く、ピアノ、バイオリン、トランペット、フルートの4楽器に相当する奏法をコンパクトな同一楽器として実装したアイディアは新規性が認められる。従来楽器を始めるきっかけになったのかどうかまでは確認できておらず、演奏難易度が難しすぎるなど荒削りな側面はあるが、WISSが議論をする場であるという点を考慮して、採録と判断された。

歌詞に基づく歌声アノテーションのためのインタフェース構築

中野 倫靖, 加藤 淳, 渡邊 研斗, 濱崎 雅弘, 後藤 真孝 46

歌声データに対して時間的なアノテーション（ビブラート、表情、声質など）をするために、ユーザが歌詞文字列をクエリとしてそれにアライメントされた歌声時間範囲を探索、アノテーション付与する興味深い手法を提案している。従来法に比べてアノテーションコストが下がるのが明確であり、その実用性は高く評価できる。一方で、まだ発展途上であり、実用性を考えると多くの疑問が残るもの事実であるが、WISSで議論をするには十分に面白い研究であり、採録と判断する。会議では実用性や発展可能性についての議論がなされることを期待する。

セッション3：入力/出力

TTTV4：一口ごとに味を提示する味覚のパーソナルメディア

笠原 暢仁, 深池 美玖, 宮下 芳明 55

これまでの著者らの取組みの発展系として、一口毎に味を変えられるデバイスの提案は興味深く、WISSの登壇発表として議論を呼ぶ提案になっていると判断された。一方で、著者ら以外の関連研究がほぼ述べられていないことによる研究の位置づけの不明瞭さ、LLMと本研究の位置づけの不明瞭さ等を中心に、記述に関しては問題を抱えていたため、条件付き採録と判断された。

眼鏡の鼻あてに搭載した圧力センサを用いた耳びく入力と身体活動検出手法

大塚 晟, 高田 峻介 61

メガネの鼻あて部へ圧力センサを装着することで、耳びく動作、歩行、瞬き、脈拍等の複数の動作を高精度で認識できる手法を提案している。シンプルなセンサ構成で複数の動作を高精度で認識できている点は、全ての査読者が評価した。一方で、様々な動きを取れるということはノイズにも敏感だと考えられる点、耳びく動作自体の使いやすさ、学習コストが不明瞭な点などが指摘されたが、WISSにおいて議論する価値がある論文と考え、条件付採録と判断された。

母音、子音の順に選択を行う間接タッチ用かな文字入力手法

和田 優斗, 白根 薫, 崔 明根, 志築 文太郎 69

離れたデバイスの画面を見ながら手元のデバイスに対して直視せずに入力する「間接タッチ」という入力方法について、母音→子音という順序で入力するユニークな方法の提案が評価され、採録と判断された。母音→子音という順序の文字入力は直感に反するのでは? というナイーブな疑問があるが、著者らは実際にスマートウォッチアプリを開発して実験的に提案手法の良さを検証している。ぜひ提案手法で文字入力を試してみたいと思わせる研究である。

E-String Figures: 導電纖維編み込み紐を用いたあやとり技認識システム

永山 晃誠, 崔 明根, 武山 侑輝, 赤田 真由, 高田 嶽介, 志築 文太郎 77

あやとりのパターン認識を行うために、導電性のある紐の接触状態を両手にはめた手袋の電極から見た抵抗値で判別している。紐に電極を付ける先行研究や画像認識を行う先行研究に比べ、通常のあやとりにかなり近い状態で技認識を行えることが有用と思われ、採録と判断された。発表では、紐の長さを変えた場合のキャリブレーションや将来のアプリケーションなどを議論できると良いと思われる。

セッション4： ファブ/マテリアル

ボタンひとつで3Dプリントが体験できるシステムの提案と運用

高橋 治輝, 白井 義人, 松村 耕平 85

3Dプリント未経験者向けに、簡易な作業のみで3Dプリントを体験させるシステムが提案されている。ボタン押下のみというオペレータ不要な簡単な手段で3Dプリンタを操作可能である。新規性が明らかであり、システムが正確に実装されている点が評価された。本システムで3Dプリンタに触れてみたユーザに対する将来的なプリンタの使用可能性や、より進んだ技術の習得方法など、システムの有用性については今後明らかにされていくものとする。以上の理由から、採録と判断された。

MagElePaint: 多様な素材に簡単に回路プロトタイピングが可能な磁性導電性塗料

山岸 真人, 高田 嶽介 92

本研究の「回路プロトタイピングのための磁性導電性塗料」は、さまざまな素材に対して塗布可能であり、将来性のある研究であると評価されました。一方で、既存研究や代替手法もあり、材料としての新規性や有用性に疑問が指摘されています。また、論文の記述内容に関しても不明瞭な点があります。こうした課題はありますが、WISSが議論をする場であるという点とその議論に足る実装が行われているという点を考慮して、条件付き採録と判定しました。

RippleSpike: スパイク表現と波紋表現を組み合わせた情報提示手法

野間 直生, 沖 真帆, 塚田 浩二 100

磁性流体にオイルを混合することで、スパイク現象に加え、なだらかな半球や余韻などの表現を実現している。有用性が不明瞭であるという指摘もあったが、新規システムが正確に実装されており、有用性についてはWISS参加者間で議論されるべきだという判定となった。ただし、提案手法の目的（課題感や目指すところ）が不明瞭となっている点は改善されるべきと考える。以上の理由から、採録（条件付き採録）と判断された。

セッション5：エンターテインメント

ゲームを用いた自動点眼システムの実装と検討

森谷 美羽, 栗原 一貴 107

ゲームが持つプレイヤーを引き込む特性を利用して、ゲームプレイ中に目が開いているタイミングをカメラ画像から検出し、点眼するシステムを実装している。このシステムを用いてゲーム内の状況と連動して「罰」「報酬」「妨害」の3つの異なるタイミングで噴射するゲームを実装し評価実験を行っている。実験の結果、一部有意な差が検出されており、WISSにおいて議論するに値する可能性があると考え条件付き採録と判定した。

オノマトペを用いた脳波制御訓練手法による魔法ゲームシステム

二瓶 朝渉, 平野 怜旺, 渡邊 恵太 115

本論文は、BCIを使用して魔法を発動する訓練をオノマトペで行うという前年のWISSのデモ発表のアイデアを元に、ゲームとして整えた上で、実験を行ったものである。実験の結果、多くの参加者は「魔法を使うという新しい体験を提供している」と感じており、特に、上手くなる過程が娯楽として面白いという反応をインタビューで得られていることが興味深い。BCIをゲームの入力に使う利点や、操作の上達を楽しめる条件など、幅広い議論を期待し、条件付き採録と判断された。

参加者との協同を実現する言葉遊びアンケートシステム

西出 新也, 西岡 裕子, 福田 孝輔, 米沢 勝, 新井 美音, 西田 健志 122

本研究は主催者と参加者が協力してアンケートを実施するために、言葉遊びを例題としてシステム実装したものである。実際に新入社員研修で使用した結果を報告している。アンケート（特定のテーマに関する情報収集）を実施することに対して、本手法がどのくらい有効であるのかについては今後も検証が必要であると思うが、総じてWISSコミュニティという実践の場で試しながら育てていくのが良いと考え、採録と判定する。

セッション6：支援

InverseVis: 疎な領域へのサンプリング点生成による多次元データ中の「非存在」の可視化

伊藤 貴之, 田上 湖都, 矢島 知子, 李 国政 129

本研究では、多次元データが疎な領域にサンプリング点を多く設定し、周辺にある既存データから得られる予測値と共に可視化して、点を選択することで追加すべきデータの性質を確認できる手法を提案している。存在しない点を可視化するというアプローチは興味深く、ドリルダウンのように詳細を確認していくインタラクションも含め、登壇発表で議論する価値があると考え採録と判断した。

ShowMe: 対話的な強調表示と拡大表示によるプレゼンテーションビデオの視覚的アクセシビリティの改善

SECHAYK Yotam, SHAMIR Ariel, 五十嵐 健夫 137

ユーザ参加型のデザインによって事前に問題点や課題を抽出し WISS2023 で発表された先行研究から実装改良している点が評価されている。また、弱視者のユーザ参加型デザインは有益で、実際に実用的なユーザインターフェースの設計になっている点が評価されている。他方で、読みづらい文章が多く見られる点やユーザ体験が感想のみと見受けられ、そこから研究への考察・深掘りが不足している点が指摘されている。以上の理由から、条件付き採録と判断された。

サブセグメント分割に基づく 3D スキャンモデルのインタラクティブセグメンテーション
真殿 航輝, 五十嵐 健夫, 加藤 大晴, 橋本 泰輔, Fabrice Matulic, 高木 士, 樋口 啓太 146

3D モデルのセグメンテーションを複数視点からのテクスチャ上 2D セグメンテーションを利用しつつインタラクティブにおこなう手法を提案している。テクスチャ上で自動解析とユーザ操作をマージしていく手法は新規性があり、またデモされている内容も十分に納得感のある結果となっていており高く評価できる。一方で、論文の記述の質に問題があり、また、インタラクティブな 3D のセグメンテーションに関する従来手法からの新規性がわかりにくい、という問題もあるため著者らにはカメラレディの前に修正を依頼し、条件付き採録とする。

PP-Checker: プログラミング教育における大規模言語モデルと協調した曖昧性のある自動採点システム
関口 祐豊, 中村 聰史 153

本研究は、入門的なプログラミング教育において受講生が犯しがちな初歩的なミスを、LLM による自動採点システムでスクリーニングする手法を提案している。結果として、受講生は単純なミスについては LLM の指摘で素早く修正することができ、また教員や TA はそのチェックを通過した提出物の採点に注力できるため、双方に有益なシステムになっている。120 名以上が受講する 1 学期間の授業で運用した実績を含めて報告されており、採録に足る論文と判断した。

ParaSights：両眼視野闘争によって2つの環境と並行してインタラクションできる空間提示手法

山本 航世 * 渡邊 恵太 *

概要. 本研究は、複数の空間と並行してインタラクションできる仕組みとして、両眼視野闘争の特性を活用した手法を提案する。複数環境の同時体験インタラクションとして、従来は視野分割やオーバーレイといった手法が試みられてきたが、それらは操作効率の低下や自己投射性の欠如といった課題があった。本研究では、左右それぞれの眼に異なる第一人称視点を提示する「ParaSights」を提案し、視点切り替えの操作を必要とせずに、2空間と並行してインタラクションできるシステムを実現した。ParaSightsは、左右それぞれの眼に異なる視点を提示することで視覚闘争を誘発し、2視点の見え方が無段階に切り替わる仕組みである。これにより、体験者は2つの環境に同時に没入し、直感的に操作できる。11名を対象とした試用により、VR酔いを誘発することなく、異なる2つの視点を同時に知覚できることが明らかになった。また、視野闘争がもたらす知覚の切り替わりが、体験者の意図により一定程度制御できる可能性が示唆された。

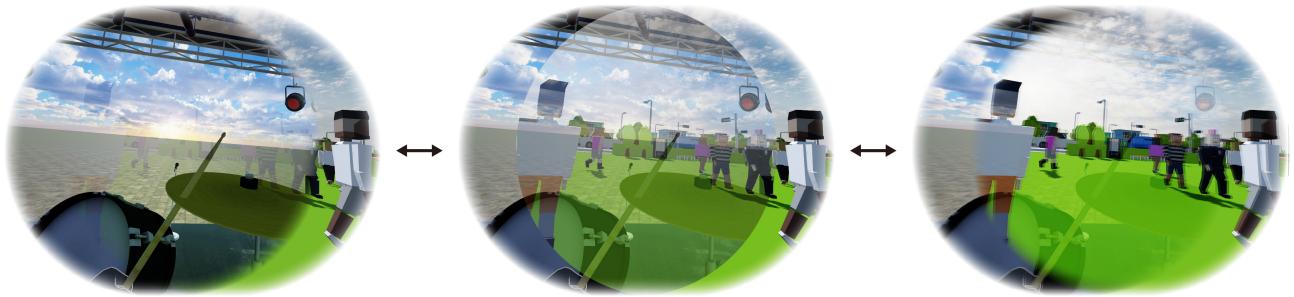


図 1. ParaSights は左右の眼に別々の視点を提示し、2つの環境を同時に体験する仕組みである。両眼視野闘争を活用し、2視点の見え方を体験者の意識に応じて制御する。(画像は見え方の再現)

1 はじめに

人間は自身の眼で知覚した1つの第一人称視点を体験している。バーチャルリアリティ(VR)手法によってバーチャル身体の第一人称視点に没入できる[12]が、通常は複数の第一人称視点を同時に知覚できない。複数の第一人称視点を同時体験できれば、バーチャル身体の動作を第三者の視点から観測する、2つのバーチャル身体を同時に操作するような体験ができる可能性がある。このようなバイロケーションの考え方はこれまでも探求されており[32][5][1][8][21]、1人の操作者による複数ロボットとのインタラクション[6][14]や、他者との視覚共有の技術[11][16][18][10][23][29]に応用されている。

一方、VRにおける複数視点の提示と操作のインタラクションに関しては未だ課題が残っている。三浦らが提案したMultiSoma[19]は、複数のバーチャル身体の視点を視野分割して提示する手法である。

同手法ではヘッドマウントディスプレイ(HMD)の表示領域を分割し、それぞれに各バーチャル身体の一人称視点を歪めて提示している。そのため、実空間の手の位置とは異なる複数の位置にバーチャル身体の手が提示されるなど、バーチャルリアリティの三要素[34]の1つである自己投射性を欠いている。Schjerlundらが提案したOVRlap[25]は、バーチャル環境内の離れた場所に存在する複数の視点をオーバーレイして提示する手法である。体験者は、アクティブな視点でのみオブジェクトを操作でき、パッシブな視点のオブジェクトは半透明で表示される。視点はジョイスティックを左右に倒すことで切り替えられるが、視点切り替え操作にかかる時間が操作効率の低下を引き起こす課題として指摘されている。

そこで我々は両眼視野闘争に注目した。両眼視野闘争の性質を利用すると、これらの課題を解決できる可能性がある。両眼視野闘争とは、左右の眼に異なる像を提示した場合、脳が混乱し両眼の像が1つに融合せず、知覚される像が時間とともに交互に入れ替わる現象である。像が入れ替わる周期は不規則

Copyright is held by the author(s).

* 明治大学 総合数理学部

であるが、知覚されている時間の割合を大まかにコントロールすることは可能である。そのため、2つの視点の見え方は体験者の意識に応じてある程度切り替えられる可能性がある。

本研究では、左右それぞれの眼に異なる視点を提示し、両眼視野闘争によって2つの環境と並行してインタラクションできる仕組みであるParaSightsを提案する(図1)。ParaSightsは、両眼視野闘争を活用し、2つの視点の見え方を体験者の意識によって変化させる。これにより、体験者は並行して提示された2つの環境に同時に没入し、インタラクションできる可能性がある。本稿では、両眼視野闘争を活用して2つの環境を同時体験する場合の操作手法と2視点の切り替わりについて探求する。

2 関連研究

2.1 両眼視野闘争

両眼視野闘争とは、左右の眼に異なる像を提示した場合、脳が混乱し両眼の像が1つに融合せず、知覚される像が時間とともに交互に入れ替わる現象である。両眼視野闘争は1838年のWheatstoneによる研究[31]より永く心理学や脳神経科学の領域で研究されているが、未だ全てが解明されていない。像が入れ替わる周期は不規則であるが、知覚される時間の割合は大まかに制御できることが知られている。例えば、コントラストの高い像は低い像よりも知覚される時間の割合が大きく[20]、明度が高い像は低い像よりも優勢になる[9]。また、輪郭が大きく変化する像は小さく変化する像よりも優位になり[3]、空間周波数が高い像は低い像よりも知覚されやすい[4]。さらに、知覚が切り替わるタイミングは体験者の意図や注意によってある程度制御でき、練習によって制御の精度を上げられる[17]。本研究では、この両眼視野闘争の特性を活用し、体験者の意識に応じて2視点の見え方を制御する。

2.2 両眼視野闘争を活用したシステム

水野らは、独立した視野を両眼に呈示してカメレオンの視野を再現するバーチャルカメレオンを提案した[35]。同手法では、2台のCCDカメラの画角をトラックボールで制御し、カメラから得られた視点を左右の眼に別々に提示する。これにより、体験者は独立した2つの視野を獲得し、視野を拡張できる。しかし、視野内の刺激に対する反応速度の低下が確認されており、独立した視点の提示が反応速度を低下させることを示している。一方で、左右の眼に提示された2つの独立した視点を人間が並行して認知できる可能性も示唆している。ただし、水野らの研究は視野拡張に焦点を当てているため、本研究では異なる場所に配置した2視点を並行して認知できるか確認する。

凱らは、片目にのみUI画面を表示することで、ビジュアルシーンを隠さずに情報を確認できるBUI(binocular rivalry UI)を提案した[7]。実験結果では、BUIが従来の半透明UIと比較してビジュアルシーンの視認性が有意に向上させたが、UIの読みやすさは低下した。これは、両眼視野闘争が注意をUIと視覚シーンに分散させ、文章を読むような認知負荷の高いタスクには適さないことを示している。そのため、本研究では文字情報の読解を避け、認知負荷が比較的低い場面での活用を目指す。

落合らは、2つの高解像度ディスプレイに異なる2枚の画像を提示し、両眼視野闘争を引き起こす視野闘争万華鏡を制作している[22]。

村本らは、等倍速の映像と1/2倍速の映像をそれぞれ両眼に提示し、リアルタイムとスローモーションの体験を同時に提供するRS-lensを提案している[36]。この手法では、両眼に提示された異なる速度の映像が1つの像として認識された。これは、両眼に提示する映像に大きな差分がない場合、視野闘争が起りにくくなる可能性を示唆している。本研究では、両眼に提示する映像の差分を大きくして視野闘争の影響を高め、2つの視点を独立して認識できるようにすることを目指す。

2.3 単眼立体視

Vishwanathらは、両眼視差が存在しない場合も立体視が生じうることを示し[30]、両眼視差が立体視の主要な要因であるとする従来の理論[24]を覆した。静止画を片目で見ても立体視が生じるという報告もあり[15]、Vishwanathらは、こうした単眼立体視の現象について検証した。同氏らは、静止画を様々な条件(両眼視、単眼視、単眼開口視、両眼開口視)で観察した場合の奥行きの印象や質的特性の変化を調査した。開口視では、絵画の額縁を排除するために、小さな開口部を通して観察した。その結果、単眼開口視条件は両眼視条件よりも強い立体視を得た。また、単眼開口視条件は単眼視条件よりも立体視の印象が有意に強かった。同氏らはこの結果から、絵画の表面視認性の低下によって、立体知覚が絵画の奥行き表現に影響されやすくなり、単眼立体視が生じた可能性を指摘している。HMDを用いた本システムは、同氏らの実験における単眼開口視条件と同じようにバーチャル環境を穴を通して見るような条件となっているため、この単眼立体視の現象が発生する可能性がある。

2.4 VRにおいて複数の視点を提示する手法

画面分割して提示する手法 三浦らは、最大4つのバーチャル身体を同時に制御し、各身体からの一人称視点を視野分割して提示するMultiSoma[19]を提案した。体験者は各身体に身体化感覚を維持しつつ、適切に注意を切り替えながら操作できる。また、

適切な身体を選択することで移動距離が減少し、タスク効率が向上する。しかし、同手法はバーチャル身体の一人称視点を視野分割で提示するため、視点が歪められ、バーチャルリアリティの「3 次元の空間性」「実時間の相互作用性」「自己投射性」のうち、特に自己投射性が欠如している。具体的には、実空間で自分の手がある位置と異なる位置にバーチャル身体の手が提示されるため、自己受容感覚が一貫しないという問題が生じる。ラバーハンド錯覚 [2] のような自己受容感覚のドリフト現象が VR 環境上でも起こることが報告されてはいるが [33][27]、その効果は限定的である。さらに、視野分割による視野角の狭さも課題として挙げられる。ParaSights 手法は、視野角を最大限に保ちながら自己投射性を維持した空間を提供し、VR システムの質を向上させる可能性がある。

オーバーレイで重ねて提示する手法 Schjerlund らは、複数の離れた視点をオーバーレイで重ねて提示する OVRlap[25] を提案した。体験者は、アクティブな視点のオブジェクトのみを操作でき、パッシブな視点のオブジェクトは半透明で視認できる。また、ジョイスティックを用いてアクティブな視点を切り替えられる。同手法は体験者の空間認識能力を向上させ、効率的なタスク遂行を支援している。しかし、視点切り替え操作の短縮によるタスク効率向上が今後の課題とされている。ParaSights 手法は、視野闘争の特性を活かし、この課題を解決できる可能性がある。また、OVRlap 手法では、パッシブな視点を青色の半透明状態で表示するため、環境内の色情報が失われる。一方、本研究では 2 つの視点の色情報をそのまま提示できるため、この課題を解決しているといえる。

3 ParaSights システム

ParaSights は、HMD の 2 つのディスプレイにそれぞれ異なる視点を提示し、2 つの空間を同時に体験する仕組みである（図 2）。本手法は、左右の眼に異なる視点の映像を提示することで、2 つの第一人称視点を体験者の左右の眼にそれぞれ対応させる。これにより、体験者は同時に提示された 2 つの視点に並行して没入し、インタラクションできる可能性がある。また、両眼視野闘争を活用し、2 視点の見え方を体験者の意識に応じて制御する。

3.1 ParaSights のインタラクション

ParaSights システムは、視野闘争の特性を活かして、視点変更の操作なしに 2 視点の見え方を無段階に変化させる。視点の切り替わり方を 4 つのパターンに分類し、それぞれの状態で自然な操作と視点の見え方を実現している（図 3）。

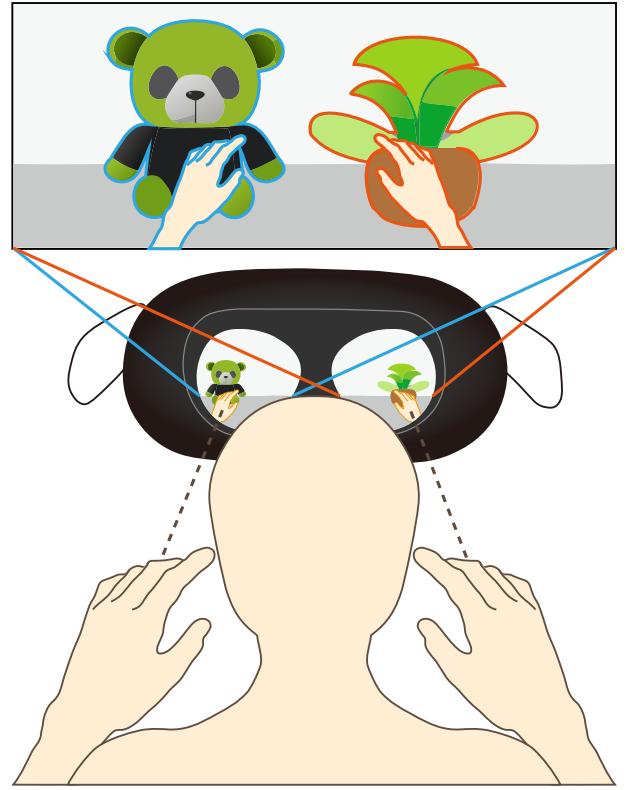


図 2. ParaSights のシステム実装イメージ。HMD の 2 つのディスプレイにそれぞれ異なる視点を提示する。異なる 2 視点を重ねるように配置し、Culling Mask で体験者の左眼、左手に片方の視点を、右目、右手にもう片方の視点を割り当てている。

Skim（スキム） Skim は、両眼の視界をぼんやりと眺めるインタラクションであり、このとき視野闘争の影響が最大となる。その結果、2 つの視点の要素が大きく重なり合って視認される。左眼側の環境は左手で、右眼側の環境は右手で操作できる。環境音はそれぞれ対応した耳に届けている。

Focus(フォーカス) Focus は、どちらか片眼の環境にある特定のオブジェクトに集中するインタラクションである。本研究では、片方の手がもう一方の手より活発に動いている状態を、その環境に集中している状態とした。左右の手のハンドトラッキングにより、システムは各手に対応するオブジェクトの加速度を常時取得している。そのため、集中している側の手で活発にインタラクションしていることを検知できる。そこで、操作中の環境の明度とコントラストを高め、もう片方の環境の明度とコントラストを低くする。これにより、視野闘争において集中している環境が優位となり、視認されやすくなる。また、視野闘争によって見える像が意識に応じて切り替わり、意識したオブジェクトが浮き出るように視認される。環境音も同様に、意識が集中している側

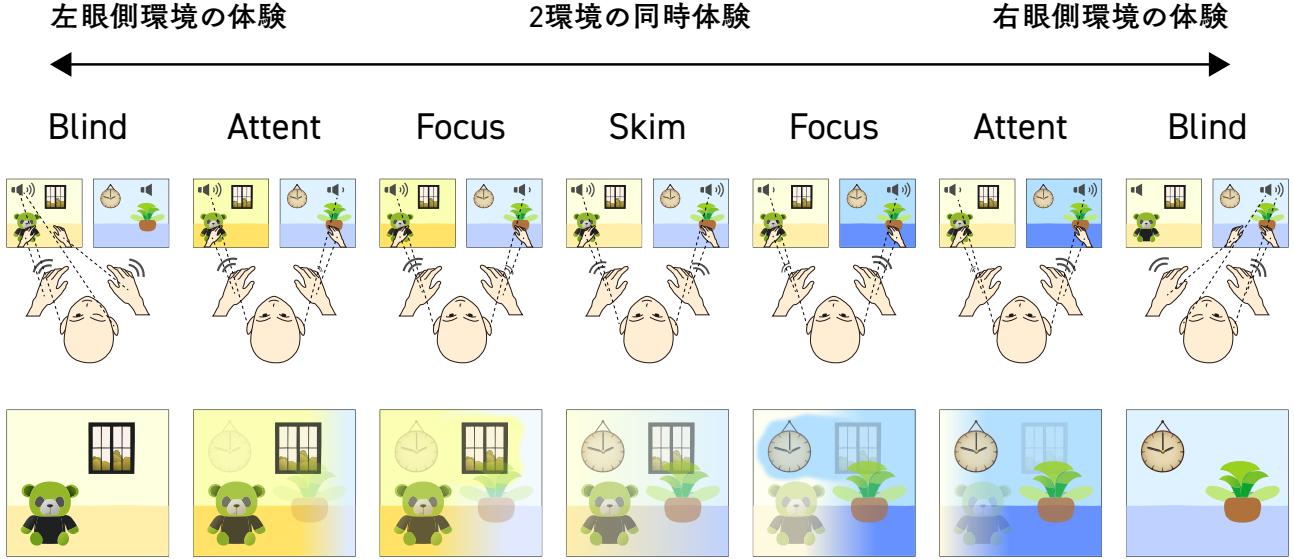


図 3. ParaSights によるインタラクションの仕組み. Blind は片眼を瞑ってもう片方の環境に没入する. Attent と Focus は、片手を活発に動かす動作をトリガーに視野闘争が誘発され、片側環境を中心とした体験を提供する. Skim は 2 つの環境を完全に並行して体験する.(画像下側は見え方の再現)

は明瞭に、意識が薄い側は音量を低く調整している。

Attent(アテント) Attent は、どちらか片眼の環境全体に集中するインタラクションである。この場合、Focus 状態と同様に集中している側の環境の明度とコントラストを高める一方で、もう片方の環境の明度とコントラストを低くする。これにより、視野闘争の影響で意識している環境側の視野が広がったように視認される。また、環境音も集中している側を優先し、意識が薄い側の音量は小さくしている。

Blind(ブラインド) Blind は、片眼を瞑ってもう片方の視界に完全に没入するインタラクションである。このとき、視野闘争の影響を受けることなく、片方の視点のみが明確に視認される。両手を視認されている環境で操作でき、環境音も視認されている環境のものを両耳に届ける。

3.2 ハードウェア/ソフトウェア実装

本手法の実装にあたり、HMD は Meta Quest Pro¹ を使用した。また、本手法およびサンプルアプリケーションの実装には Unity(2022.3.10f1)² を用いた。VR 環境の構築は Meta XR SDK(v68) および Oculus XR Plugin(v4.2) を使用した。実装においては、左右の眼に異なる視点を提供するため、OVR-CameraRig の Use Per Eye Cameras オプションに

より、Left EyeAnchor と RightEyeAnchor の 2 つのカメラを用いる手法を採用した。2 つのバーチャル環境を別々のレイヤー上に配置し、Culling Mask を使用して逆の眼に表示されるバーチャル環境の描画を無効化することで、Unity 上で独立した 2 視点の提示を擬似的に実現した。なお、近い地点の別視点を提示する場合は、それぞれを異なるバーチャル環境として提示し、環境内のオブジェクトの動きを同期して対応した。また、ステレオレンダリングモードについては各眼に対して個別に描画する必要があるため、Quest Pro 開発時のデフォルトである Multiview 方式ではなく、MultiPass 方式を採用した。

Focus, Attent モードにおけるコントラストおよび明度の調整について、BuildingBlock 機能の Virtual Hands で配置した OculusHand_R(L) の現在位置をフレームごとに取得し、加速度を算出した。この加速度に基づき、左右の手における加速度差が一定以上となった際に、Post Processing Stack v2 の Post-process Volume 内 ColorGrading プロパティの contrast.value および postExposure.value によって画面のコントラストと明度を調整する。具体的には、それぞれの値を Mathf.Lerp 関数を用いて -75.0f から 100.0f, 1.0f から -1.0f の範囲でスムージングさせた。Blind インタラクションについては、OVRFaceExpression を用いて EyeClosedR(L) の値を参照し、片方の目が閉じられている状態を検知する。この際、開いている側の視点に対応したレイヤーに両手のオブジェクトを移動させることで、両手を片方の環境において視認し操作できる。

¹ Meta Quest Pro
<https://www.meta.com/jp/quest/quest-pro/>

² Unity
<https://unity.com/ja>

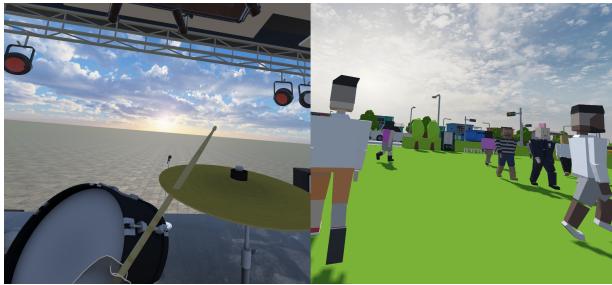


図 4. 試用に用いたサンプルアプリケーション。左眼側でドラムを叩く体験、右目側で周囲を歩き回る人型のキャラクターを観測する体験を提供している。

4 試用と体験

本章では、試用を通して、異なる 2 つの視点を同時に体験する際の見え方と視点の切り替わりについて調査した。具体的には、2 視点を同時に知覚し並行して反応できるか、2 視点の見え方を Skim, Focus, Attent, Blind のように変化させられるか、酔い VR 酔いが発生しないか、および距離感を正確に掴めるかの 4 点を中心に検討した。

4.1 手順

試用に用いるサンプルアプリケーションとして、片方の環境でドラムを叩く体験、もう片方の環境で周囲を歩き回る人型キャラクターを観測するアプリケーションを用意した(図 4)。本試用の体験者は、両眼立体視を利用した VR を体験したことのある大学生および大学院生(19~24 歳)11 名(うち男 8 名女 3 名)である。学科内で行われたデモ発表の参加者を対象に試用し、謝金は支払っていない。また、体験者全員に斜視特性は無かった。体験者には先入観が生じないよう事前に左眼と右眼に異なる視点を提示しているという最低限の情報を提供し、両眼視野闘争や複数環境とのインテラクションなど本研究のテーマとなる詳しい内容は伝えていない。体験者は 2 つの視点の見え方と切り替わりに関する非構造化インタビューに回答しながら体験を行った。

4.2 結果

試用の結果、11 名全員が 2 つの視点を同時並行して視認できたと報告し、VR 酔いの症状を訴える者はいなかった。特に、6 名の体験者は、左右の眼に独立して提示された 2 つの視点が融合し、1 つの場所に存在している感覚があると報告した。また、9 名の体験者は、ドラムが配置された環境内で迷うことなくバチを持ち上げることができた。ただし、2 名はバチの位置を把握するまでに時間がかかったが、一度掴んだ後は安定して持ち直すことができた。

見え方の切り替わりに関しては、7 名の体験者がドラムを叩く動作を行っている際にドラムが鮮明に

見え、動作をやめると鮮明さが失われると報告した。さらに、5 名の体験者が、人型キャラクターが視点の近くを通り過ぎるとその人型キャラクターの姿がはっきり見えると報告した。また、4 名の体験者が視界の手前側にはドラムが、奥側には街の風景が見えやすいと感じたと述べた。

さらに、10 分程度本手法を体験した 3 名のうち 2 名から、体験時間が 5 分を超えた頃から本手法に慣れ、体験開始時よりも視点の見え方を制御できるようになったとの報告があった。

4.3 考察

試用の結果は、ParaSights 手法が両眼視野闘争によって、2 つの異なる視点を同時に知覚できる可能性を示唆した。特に、11 名全員が VR 酔いを訴えなかった点から、自律神経に過度な負荷をかけることなく、複数の視点を提供できると考えられる。

6 名の体験者が報告した「2 つの視点が融合した 1 つの場所に存在している感覚」というフィードバックは、視野闘争によって生じる本手法の体験が、2 つの場所に同時に存在する感覚を提供していないことを示唆している。一方で、両方の空間の要素を同時に視認できることから、Skim インタラクションが機能していると考えられる。

また、7 名の体験者はドラムを叩く際にドラム周囲の視界が鮮明になったこと、5 名の体験者が人型キャラクターが視点の近くを通り過ぎた際に人型キャラクターの周囲が鮮明に見えたことを報告した。この結果は、ParaSights インタラクションの 1 つとして設定した Focus インタラクションが効果を發揮し、視野闘争による知覚の切り替わりが体験者の意図により一定程度制御可能である可能性を示唆した。

さらに、10 分間の体験で視点の制御が容易になるという慣れの報告から、ParaSights は使用時間が延びるにつれて操作性が向上する可能性がある。この点から、長時間の使用を前提としたタスクにおいて ParaSights が有効であることが期待されるが、長時間使用した際の疲労感や視覚的負担については今後の調査が必要である。

5 議論

本手法の特徴と試用の結果をもとに、本手法が活用できるアプリケーション例について議論する。

5.1 バーチャル環境内の 3D モデリング

3D モデリングを行う際、1 つの視点に固執してしまい、後から別の視点で見るとモデルが崩壊しているという場面や、引いて全体を見てみるとバランスが崩れているといった場面によく遭遇する。従来のデスクトップ環境でのモデリングソフトウェアでは、複数の方向から見た視点を画面分割して表示することで、そのようなミスを防いでいる。しかし、

バーチャル空間内での 3D モデリングでは、現状この課題に対する解決策はない。バーチャル空間内でモデリングを行えるソフトウェアとして、Blender の Freebird XR アドオン³ や、Adobe Substance 3D Modeler⁴がある。これらのソフトウェアではオブジェクトの回転や視点の移動でオブジェクトを見回せるが、複数の視点を同時に見ることはできない。また、バーチャル環境内で建物などの大きなオブジェクトを作る際に建造物を回転させたり縮尺を変えて回り込む動作は好ましくない。本手法を用いれば、同時に複数の視点からオブジェクトを視認できるため、別角度や引きの視点から見た際に不整合なモデリングに瞬時に気が付き対応できる可能性がある。このように、VR 環境でオブジェクトを複数の視点から見ながら作業するシーンにおける活用が期待できる。その他にも、3D パズル、積み木のようなタスクにおいても活用できる可能性がある。

5.2 バーチャルリアリティ観戦 (VRS) の 2 窓視聴

スポーツ観戦では時折見逃せない瞬間がある。得点が決まった瞬間や重要なセットプレーが発生した瞬間など、どうしてもリアルタイムに観戦したい場面がある。しかし、オリンピックやワールドカップなどのイベントでは見逃せない複数の試合が同時に行われる場合も多い。こういった場合、テレビ配信などではいわゆる 2 窓視聴を行うことになる。しかし、バーチャルリアリティ観戦 (VRS) では通常 2 窓視聴ができない。VRS は新たなスポーツメディアとして注目されており、従来の 2D スクリーンよりも鮮明さ、インタラクティブ性、テレプレゼンスを介してフローレ体験(認知的没入、楽しさ、および時間の歪みを伴う体験)を大きく増幅する [13]。提案手法は VRS において双方ともリアルタイムかつ第一人称視点で体感できる可能性がある。また、ライブやコンサートの VR 鑑賞においても活用できる可能性がある。このように、2つの視点を常時観察しつつ、両方を見逃せないシーンへの活用が期待できる。このようなタスクは常時ひどく認知負荷を要するタスクではないため、複数視点を同時に観察しつつ、重要なシーンの見逃しを防げる可能性がある。

5.3 コミュニケーションのリアルタイム自己改善

普段のプレゼンテーションやコミュニケーションの場において、相手から見た自身の見え方が気になることがある。身振り手振りは適切か、緊張で拳動不審になっていないかなどが気になるが、リアルタイムに相手から見た自分を知ることはできない。本

手法はバーチャル環境においてこの課題を解決できる可能性がある。操作しているアバタと同時に第三者視点から見ている別のアバタにも没入することで、第三者視点から見た自身の見え方を確認できる。このように、自身をリアルタイムに客観視して改善するシーンにおいて活用できる可能性がある。

5.4 今後の課題と展望

視野闘争への慣れ

今回の試用において、ParaSights は使用時間が延びるにつれて徐々に視点の制御ができるようになる可能性を示唆した。しかし、試用時間は最大で 10 分程度であったため、それ以上の長時間使用時における視点制御の可能性は依然として不明である。著者自身は累計で数時間にわたる体験を行い、ParaSights の 4 つのインタラクションを全て体験できているため、長時間の使用によって本手法に対する慣れが進む可能性がある。このため、今後の研究では長時間の使用が視点制御に与える影響を検証する必要がある。

2 視点の独立ロコモーション

今回の試用では、2つの視点を同時に視認できると報告された。しかし、両視点への独立したロコモーションは試みられなかった。もし左右の眼に独立して提示された 2 つの視点に対し、それぞれで移動操作が行われた場合、2つの視点を別々の場所として知覚できる可能性がある。この仮説を検証するためには、左右の環境に対する独立したロコモーションを実装し、その効果を検証する必要がある。

斜視への適用

今回の試用では、斜視を持つ体験者は含まれていなかった。斜視は、片目が正面を向いている際にもう片方の目が異なる方向を向く状態であり、左右の視線が一致しない特徴を持つ。斜視の特性を持つ者は、正視者とは異なる視野闘争の特性を持つことが報告されている [26][28]。したがって、斜視を持つ人に対する本手法の効果を確認するためには、追検証が必要である。

6 まとめ

本稿では、両眼視野闘争を利用した空間提示手法である ParaSights を提案し、複数の空間と並行してインタラクションできるシステムを実現した。試用の結果、ParaSights が視点切り替えを不要にし、2つの視点を同時に視認できることを示した。さらに、視野闘争による視点の切り替わりが、体験者の意図により一定程度制御可能であることを示唆した。今後は、2 視点の独立したロコモーションや長時間の連続使用についてさらなる検証を行い、より詳細な実験によって手法の有用性を確認していく。

³ Freebird XR

<https://freebirdxr.com/>

⁴ Adobe Substance 3D Modeler

<https://www.adobe.com/jp/products/substance3d-modeler>

参考文献

- [1] L. Aymerich-Franch, D. Petit, G. Ganesh, and A. Kheddar. The second me: Seeing the real body during humanoid robot embodiment produces an illusion of bi-location. *Conscious. Cogn.*, 46:99–109, Nov. 2016.
- [2] M. Botvinick and J. Cohen. Rubber hands 'feel' touch that eyes see. *Nature*, 391(6669):756, Feb. 1998.
- [3] B. B. Breese. Binocular rivalry. *Psychol. Rev.*, 16(6):410–415, Nov. 1909.
- [4] M. Fahle. Binocular rivalry: suppression depends on orientation and spatial frequency. *Vision Res.*, 22(7):787–800, Jan. 1982.
- [5] T. Furlanetto, C. Bertone, and C. Becchio. The bilocated mind: new perspectives on self-localization and self-identification. *Front. Hum. Neurosci.*, 7:71, Mar. 2013.
- [6] D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita. Teleoperation of Multiple Social Robots. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.*, 42(3):530–544, May 2012.
- [7] K. Guo, Y. Shimomura, J. Hosoi, Y. Ban, and S. Warisawa. GUI presentation method based on binocular rivalry for non-overlay information recognition in visual scenes. In *Proceedings of the Augmented Humans International Conference 2024*, New York, NY, USA, Apr. 2024. ACM.
- [8] A. Guterstam, D. E. O. Larsson, J. Szczotka, and H. H. Ehrsson. Duplication of the bodily self: a perceptual illusion of dual full-body ownership and dual self-location. *R. Soc. Open Sci.*, 7(12):201911, Dec. 2020.
- [9] I. T. Kaplan and W. Metlay. Light intensity and binocular rivalry. *J. Exp. Psychol.*, 67(1):22–26, Jan. 1964.
- [10] S. Kasahara, M. Ando, K. Suganuma, and J. Rekimoto. Parallel eyes: Exploring human capability and behaviors with paralleled first person view sharing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, May 2016. ACM.
- [11] H. Kawasaki, H. Iizuka, S. Okamoto, H. Ando, and T. Maeda. Collaboration and skill transmission by first-person perspective view sharing system. In *19th International Symposium in Robot and Human Interactive Communication*, pp. 125–131. IEEE, Sept. 2010.
- [12] K. Kilteni, R. Grotens, and M. Slater. The sense of Embodiment in virtual reality. *Presence (Camb.)*, 21(4):373–387, Nov. 2012.
- [13] D. Kim and Y. J. Ko. The impact of virtual reality (VR) technology on sport spectators' flow experience and satisfaction. *Comput. Human Behav.*, 93:346–356, Apr. 2019.
- [14] S. Kishore, X. N. Muncunill, P. Bourdin, K. Or-Berkers, D. Friedman, and M. Slater. Multi-destination beaming: Apparently being in three places at once through robotic and virtual embodiment. *Front. Robot. AI*, 3:217540, Nov. 2016.
- [15] J. J. Koenderink. Pictorial relief. *Philos. Trans. A Math. Phys. Eng. Sci.*, 356(1740):1071–1086, May 1998.
- [16] K. Kurosaki, H. Kawasaki, D. Kondo, H. Iizuka, H. Ando, and T. Maeda. Skill transmission for hand positioning task through view-sharing system. In *Proceedings of the 2nd Augmented Human International Conference*, New York, NY, USA, Mar. 2011. ACM.
- [17] L. C. Lack. Effects of knowledge of results and spacing of practice trials in training of control of binocular rivalry. *Percept. Mot. Skills*, 31(3):827–830, Dec. 1970.
- [18] R. Mitchell, C. S. B. Bravo, A. H. Skouby, and R. L. Möller. Blind running: Perceptual team interdependency for self-less play. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, New York, NY, USA, Oct. 2015. ACM.
- [19] R. Miura, S. Kasahara, M. Kitazaki, A. Verhulst, M. Inami, and M. Sugimoto. MultiSoma: Motor and gaze analysis on distributed embodiment with synchronized behavior and perception. *Front. Comput. Sci.*, 4, May 2022.
- [20] T. J. Mueller and R. Blake. A fresh look at the temporal dynamics of binocular rivalry. *Biol. Cybern.*, 61(3):223–232, 1989.
- [21] E. Nakul, N. Orlando-Dessaints, B. Lenggenhager, and C. Lopez. Measuring perceived self-location in virtual reality. *Sci. Rep.*, 10(1):6802, Apr. 2020.
- [22] Y. Ochiai. Kaleidoscopes for binocular rivalry. In *Proceedings of the 3rd Augmented Human International Conference*, New York, NY, USA, Mar. 2012. ACM.
- [23] R. Pan, S. Singhal, B. E. Riecke, E. Cramer, and C. Neustaedter. MyEyes: The design and evaluation of first person view video streaming for long-distance couples. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, New York, NY, USA, June 2017. ACM.
- [24] C. R. Ponce and R. T. Born. Stereopsis. *Curr. Biol.*, 18(18):R845–50, Sept. 2008.
- [25] J. Schjerlund, K. Hornbæk, and J. Bergström. OVRlap: Perceiving multiple locations simultaneously to improve interaction in VR. In *CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, Apr. 2022. ACM.
- [26] C. M. Schor. Visual stimuli for strabismic suppression. *Perception*, 6(5):583–593, 1977.
- [27] M. Slater, D. Perez-Marcos, H. H. Ehrsson, and M. V. Sanchez-Vives. Towards a digital body: the virtual arm illusion. *Front. Hum. Neurosci.*, 2:6, Aug. 2008.
- [28] E. L. Smith, 3rd, D. M. Levi, R. E. Manny, R. S. Harwerth, and J. M. White. The relationship between binocular rivalry and strabis-

- mic suppression. *Invest. Ophthalmol. Vis. Sci.*, 26(1):80–87, Jan. 1985.
- [29] J. Sypniewski, S. B. Klingberg, J. Rybar, and R. Mitchell. Towards dynamic perspective exchange in physical games. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, Apr. 2018. ACM.
- [30] D. Vishwanath and P. B. Hibbard. Seeing in 3-d with just one eye: stereopsis without binocular vision: Stereopsis without binocular vision. *Psychol. Sci.*, 24(9):1673–1685, Sept. 2013.
- [31] C. Wheatstone. XVIII. Contributions to the physiology of vision. —Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philos. Trans. R. Soc. Lond.*, 128(0):371–394, Dec. 1838.
- [32] B. Wissmath, D. Weibel, J. Schmutz, and F. W. Mast. Being present in more than one place at a time? Patterns of mental self-localization.
- Conscious. Cogn.*, 20(4):1808–1815, Dec. 2011.
- [33] Y. Yuan and A. Steed. Is the rubber hand illusion induced by immersive virtual reality? In *2010 IEEE Virtual Reality Conference (VR)*, pp. 95–102. IEEE, Mar. 2010.
- [34] 館 瞳, 佐藤 誠, 廣瀬 通孝. バーチャルリアリティ学. 工業調査会 / 日本バーチャルリアリティ学会 / コロナ社, 2011.
- [35] 水野文雄, 早坂智明, 山口隆美. バーチャルカメレオン：任意の方向の独立した視野を両眼に呈示するシステムに関する基礎的研究(教育工学・福祉工学・マルチメディア応用, 一般論文). 情報科学技術フォーラム講演論文集, 8(3):631–632, 08 2009.
- [36] 村本剛毅, 齊藤寛人, 脇坂崇平, 笠原俊一, 稲見昌彦. リアルタイムとスローモーションを共存させるインターフェース. 日本バーチャルリアリティ学会大会論文集, 25th, pp. 2–7, 2020.

HidEye：片目を隠す動作による HMD 用インタラクション手法の提案

伊勢 隆之介 * 塚田 浩二 *

概要. 近年、メタバース等の発展とともに没入型ヘッドマウントディスプレイ (HMD) の利用場面は日常生活にも広がりつつある。日常生活環境で HMD を利用するには周辺の人や物に対する状況認識が重要であるが、高い没入感や視認性と引き換えに両目の周辺が覆われてしまうため、周辺状況の確認が難しい。本研究では、片目を隠す動作で VR コンテンツとパススルー機能を重畳表示することで、仮想空間と現実空間のコンテンツを手軽に切り替え可能なインタラクション手法 HidEye を提案する（図 1）。さらに、片目を隠す動作を仮想空間での視野変更に活用した応用例を実装する。

1 はじめに

近年、メタバース等の発展とともに没入型ヘッドマウントディスプレイ (HMD) の利用場面は日常生活にも広がりつつある。任意の場所でプライベートな大画面ディスプレイを確保でき、現実空間の視界を遮断することで高い没入感を提供できる特徴から、VR ゲームなどのエンターテイメント利用はもちろん、技術トレーニング [4] や移動中の快適なオフィスワーク [2] 等の幅広い用途で活用が進められている。また、誰でも自由に装着可能な HMD を公共の場に設置してパブリックディスプレイとして利用する研究 [6] も進められている。

一方、日常生活環境で HMD を利用するには周辺の人や物に対する状況認識（アウェアネス）が重要であるが、利用中に現実世界の状況を確認するためには、HMD を取り外すかパススルー機能¹を使用する必要がある。しかし、HMD の着脱は手間がかかり、頭髪や化粧も乱れることから頻繁に行なうことは適さない。パススルー機能を使えば HMD を装着したまま周辺状況を確認できるが、都度特定のコマンド操作が必要となり、体験中のコンテンツも不可視になるため、没入感が削がれてしまう。

そこで我々は、自然な動作 (Hand-to-Face Gesture) をトリガーとして、仮想空間とパススルー機能を重ね合わせることで、仮想空間での没入感を一定程度保ちつつ、現実空間の状況認識が可能であると考えた。顔を手で触れる Hand-to-Face Gesture は、自然で直感的な入力手法であると報告されている [7]。HMD に対する多様な Hand-to-Face Gesture が先行研究 [8, 5] で紹介されているが、我々は、「手で片目を覆う」ジェスチャに注目した。「手で目を覆う」動作は物理的に視界を変化させる行為であり、心理的に見たくないものから目を背ける遮断行

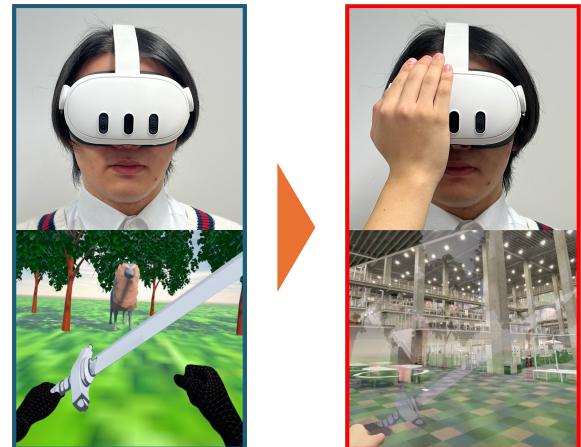


図 1. 本研究のコンセプト。片目を隠す動作を行っている間 HMD に表示されるコンテンツを切り替える。

為ともされている [9]。また、「片目を隠す」動作は、視力検査やカメラのファインダーを覗き込む時に、特定のタスクに集中する行為としても利用される。こうした点から、「手で片目を覆う」動作は仮想空間と現実空間の視野変更を行う際のジェスチャとしても有効であると考えた。

そこで本研究では、片目を隠す動作で VR コンテンツとパススルー機能を重畳表示することで、仮想空間と現実空間のコンテンツを手軽に切り替え可能なインタラクション手法 HidEye を提案する（図 1）。さらに、片目を隠す動作を仮想空間での視野変更に活用した応用例についても紹介する。HMD に特殊なセンサ等を付与することなく、自然な動作で素早く仮想空間と現実空間のタスクを連携することを目指す。

なお、本論文はインタラクション 2024 でインタラクティブ発表 [11] を行い SIGGRAPH Asia の Posters[3] で発表予定の内容に加筆したものである。

Copyright is held by the author(s).

* 公立はこだて未来大学

¹ カメラやセンサを用いて周辺映像を取得し、HMD に表示する機能。

2 関連研究

2.1 HMD 利用時の周辺状況の確認

Endo らは、手動で再構成可能なモジュラー機構を用いた新たな HMD の概念 ModularHMD を提案している [1]. 没入感の高い VR 体験を確保できるようにしつつ、必要に応じて外界とのインタラクションを実現するために側面 2 台と底面 1 台の着脱可能なディスプレイモジュールで HMD を構成している。底面モジュールを取り外すことで入力デバイスとして活用したり、側面モジュールを取り外すことで VR 体験を継続しながら横目で周辺状況の確認や近接者と意思疎通する手段を提供している。また、山内らは、ヘッドバンドを用いないことで頭部への負担や着脱の手間を解消する胸部装着型ディスプレイ CMD を提案している [10]. CMD は俯いてデバイスを覗き込むことで MR 体験を行うため、頭を上げてデバイスから離れることで任意のタイミングで MR 体験を中断することができる。これらの研究では、HMD の物理的な機構を工夫することで周囲のアウェアネスの確保を容易にしているのに対して、本研究ではパススルー機能と適切なジェスチャを組み合わせることで HMD の着脱の手間を省いたスムーズな視界の変更を実装する。

2.2 Hand-to-Face Gesture インタラクション

Weng らは、AR グラス上で Hand-to-Face Gesture インタラクションを可能にするカメラベースのセンシング技術である FaceSight を提案している [8]. AR グラスのブリッジに赤外線センサを取り付けて顔の下部と手の動作をセンシングすることで 21 種類のジェスチャの識別を可能にしている。また、Kawasaki らは、スマートグラスのレンズに触れて操作を行うインタラクション技術 LensTouch を提案している [5]. 抵抗膜式タッチパッドをスマートグラスのレンズ表面に取り付けることでレンズに触れて操作を行うことを可能にしている。レンズ越しの指と表示された画像の両方を視認できるため、素早く対象のオブジェクトを選択することが可能となっている。これらの研究では、HMD に追加のセンサを取り付けることで入力手法の拡張を図っているのに対して、本研究では HMD の標準機能のみで実現可能なインタラクション手法を構築する。

3 提案手法：HidEye

本研究では没入型 HMD の利用中に、片目を隠す動作で仮想空間と現実空間のコンテンツを手軽に切り替え可能なインタラクション手法 HidEye を提案する。

本研究の特徴は、以下の 3 点である。

1. 片目を隠すという自然な動作を利用



図 2. MetaQuest3 の搭載センサ

2. 外部センサやデバイスが不要

3. 仮想空間と現実空間の多彩な視野変更に活用

1 については、1 章でも述べたように、「手で片目を覆う」動作は、直感的な Hand-to-Face Gesture の一つであり、物理的／心理的に視界を変化させる行為であることから、仮想空間と現実空間の視野変更を行なう際の入力手段として適切であると考える。

2 については、特殊なセンサやデバイスを必要とせず、市販の HMD さえあれば利用可能な設計とすることで、システム導入の敷居を下げることができると考える。

3 については、現実空間と仮想空間を対象とした、多様な視野変更に活用できる点である。詳細の事例は 5 章で紹介するが、現実世界と仮想世界のなめらかな移行に加えて、仮想世界のオブジェクトの生成／消失や手軽な視点変更等、視野変更を対象とする多様な場面で活用できることを示す。

4 実装

4.1 HMD と開発環境

本研究の実装に用いた HMD は MetaQuest3²である。本体前面には、左右に RGB カメラ (18PPD, 400 万画素)，左右の RGB カメラの下と本体の両側面にトラッキング用 IR カメラ，中央に深度センサが搭載されている（図 2）。深度センサから投影される赤外線の模様を IR カメラで撮影することで、その模様の歪み具合から深度を算出し周囲の立体構造を測定している。開発のためのソフトウェアは Unity ver.2022.2.8f1 を用いた。また、MetaQuest3 に対応した VR コンテンツの開発用のアセット Meta XR Core SDK を用いた。

² MetaQuest3 公式 <https://www.meta.com/quest/quest-3/>

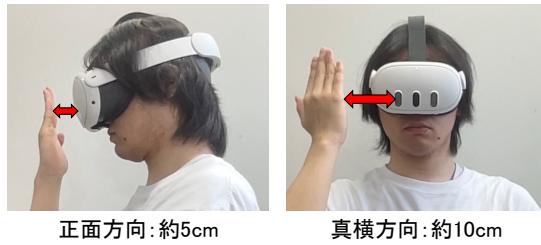


図 3. ハンドトラッキングの有効範囲

4.2 片目を隠す動作の検出手法

本提案では、MetaQuest3 に搭載されているハンドトラッキングを用いて片目を隠す動作の検出機能を実装した。動作の検出手法としては、RGB カメラのデータを用いた色分布の活用等も考えられるが、コンテンツ操作との親和性が高い点、API が提供されており開発が容易である点から、ハンドトラッキングを利用する。ハンドトラッキングは HMD に搭載されている IR カメラで手指の 3 次元形状や位置・姿勢を認識する。この認識結果に基づいて、手指の 3D モデルを VR 空間に重畳表示したり、詳細な座標取得を行う機能が提供されている。ここで、ハンドトラッキングでは、手がカメラの画角から外れた場合、認識不能（消失）状態となる。著者らが検証したところ、HMD と手の距離が正面方向が約 5cm、真横方向が約 10cm までがハンドトラッキングの有効範囲である（図 3）。このように、HMD に近い距離までハンドトラッキングが有効であることから、片手がカメラに近い位置から消失した場合に、片目を隠す動作として解釈できると考えた。なお、MetaQuest3 では両側に RGB カメラ／IR カメラがあるため、片目を隠す動作で片側のカメラが隠れた状態でもパススルー機能やハンドトラッキングが利用できることも確認している。

動作検出の流れを図 4 に示す。まず、(1) 右手から右側 IR カメラの距離 dR 、左手から左側 IR カメラの距離 dL を毎フレーム算出する。次に、(2) 片手のハンドトラッキングが消失した時、直前のフレームの dR (or dL) の値を参照する。この時、閾値より距離の値が小さければ片目を隠したと判定しイベントを発生（例：パススルー画像へ切り替え）する。その後、(3) 消失した手がハンドトラッキングで再認識されたらイベントを停止して(1)に戻る。

5 応用例

提案手法の活用場面を示すために、「現実空間と仮想空間の移行」と、「仮想空間での視野変更」の 2 種類の応用例を、計 4 つ実装した。各応用例の動作は、デモビデオでも紹介している。

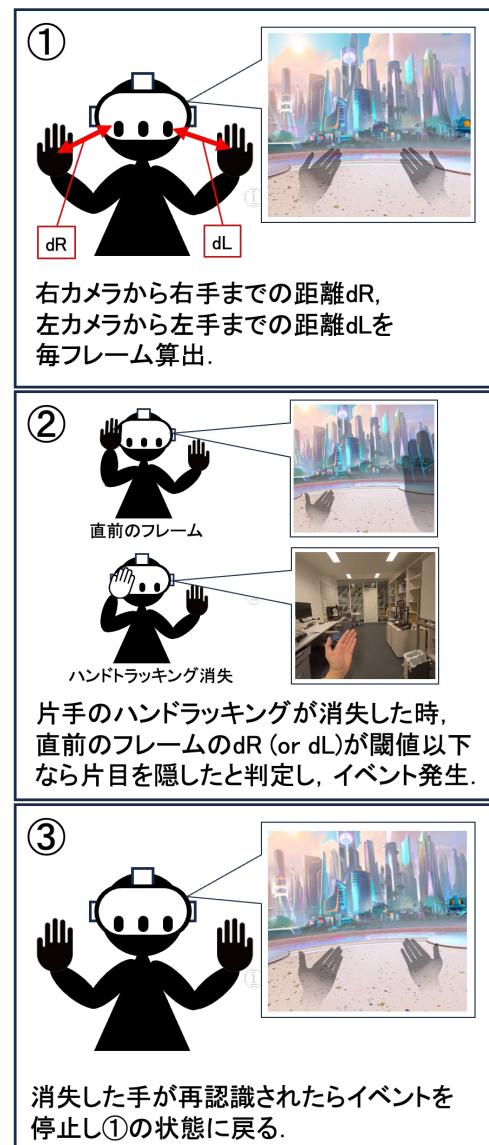


図 4. 片目を隠す動作検出の流れ

5.1 仮想空間と現実空間の切り替え

5.1.1 VR ゲーム中に周辺状況を確認

一つ目は、迫ってくる敵を武器を振って撃退する VR ゲームの体験中に周辺状況の確認や小休憩を取ることを想定した事例である（図 5）。通常時は VR ゲームを両手を動かしてプレイできるが、片目を隠すことでゲームを一時停止しつつ、パススルー機能を用いて周辺状況を確認できる。例えば図 5 の例では、コップの水を飲んで小休止している。この時、ゲームを一時停止しつつ、不透明度を下げてパススルーに重畳表示することで一定の没入感の維持と復帰しやすさを図っている。



図 5. 応用例 1: VR ゲームの体験中に周辺状況を確認.

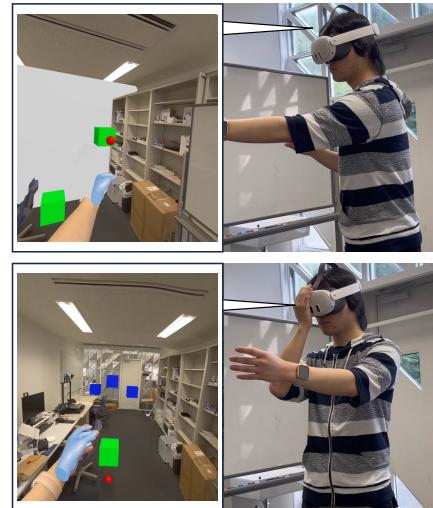


図 7. 応用例 3: MR シューティングゲーム中に障害物を消失／生成.



図 6. 応用例 2: MR タスク中に好みの景色を見て小休憩.



図 8. 応用例 4: VR ゲーム中に視点を変更.

5.1.2 MR タスク中の小休憩

二つ目は、パススルー機能を用いた MR 的なタスクを実行している際に好みの景色の VR コンテンツで小休憩するための事例である(図 6)。タスクに疲れた際に思わず目を手で覆うような動作をメタファーとして実装した。図 6 の例では VR 空間内で半球状にパーティクルを放出することで星空を表現している。任意の景色や映像を設定することでユーザの好みに応じてカスタマイズできる。

5.2 仮想空間での視野変更

5.2.1 障害物の消失／生成

三つ目は、片目を隠す動作を MR シューティングゲームに応用した事例である(図 7)。通常は親指と

人差し指をピンチすることで球を発射してターゲットを破壊することができる。片目を隠す動作することで、障害物となるオブジェクトが消失し、手を離すことで再度表示される。片目を隠して障害物を消した状態でないと破壊できないターゲットがある一方で、障害物があった方が消しやすいターゲットもあるなど、視野変更動作をゲーム性に取り入れている。

5.2.2 カメラ視点の切り替え

四つ目は、片目を隠す動作で仮想空間内の別の視点に移動できる事例である(図 8)。通常時は、一人称視点で巨大な敵と間近に対峙しているが、片目を隠す動作を行うことでゲームを一時停止しつつ三人称視点に切り替えることができる。片目を隠してい

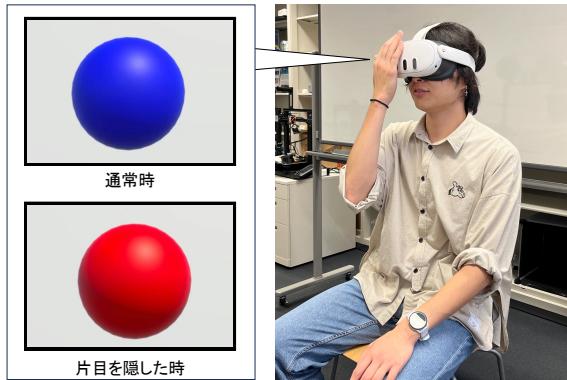


図 9. 性能評価実験の様子

る間は攻撃できない代わりに敵を観察し次の行動を考えることができる。

6 性能評価

ハンドトラッキングによる片目を隠す動作の検出精度の調査を行った。

6.1 手法

事前準備として、片目を隠す動作によってオブジェクトの色が変化する実験用アプリを作成した。通常時はオブジェクトが青色で表示され、片目を隠したと判定されると赤色に変化する(図9)。なお、片目を隠したと判定するための閾値は実験を通して共通とした。実験に用いたデモアプリはUnity ver.2022.2.8f1で作成した。また、実験中の被験者の様子とアプリの画面をそれぞれ映像で記録した。

被験者は10名(20~25歳の男性8名、女性2名)であった。まず、被験者にはデモアプリを起動した状態のMetaQuest3を装着して椅子に座ってもらい、その状態で以下の教示を与えた。

- 両手を膝の上に置いた状態を基本姿勢とする
- 実験者の指示に応じて左右どちらかの手で片目を隠し、任意のタイミングで基本姿勢に戻る
- 片目を隠したと判定されている間はオブジェクトが青色から赤色に変化する
- オブジェクトの色が変化しなかった場合でもやり直さない

なお、ユーザが自然な動作で利用できることが重要であると考えたため、「どのように目を隠すか」という具体的な指示は与えていない。その後、順序をランダムに左右10回づつ片目を隠す動作を行わせ、その成功率を算出した。基本姿勢から片目を隠す動作を行い基本姿勢に戻るまでを一連の動作とし、この一連の動作に合わせてオブジェクトの色が正しく変化(青→赤→青)した場合を成功とし、それ以外

表 1. 片目を隠す動作の検出成功率

被験者	左手 (%)	右手 (%)	両手 (%)
A	100	100	100
B	100	100	100
C	100	100	100
D	90	100	95
E	100	90	95
F	100	80	90
G	80	90	85
H	100	100	100
I	100	100	100
J	100	90	95
全体平均	97	95	96

の場合(複数回色の変化が生じた場合など)は失敗と判定する。

6.2 結果

被験者全体の検出成功率(左手、右手、両手)を表1に示す。表1より、両手の検出成功率は平均96%(標準偏差6.8)であった。左手では平均97%(標準偏差6.8)、右手では平均95%(標準偏差7.1)であり、若干左手の方が高いものの、大きな差は見られなかった。

結果の考察について示す。まず、目の隠し方などの詳細な指示を与えなくとも、全ての被験者は右手で右目/左手で左目を隠しており、全体的な認識率は平均96%と高いことが確認できた。次に、やや成功率が低かった被験者F/Gに着目して実験中の映像を分析し、要因を調査した。その結果、検出に失敗したタイミングではHMDに直接触れずに、少し手を浮かした状態で動作を行っていた。一方、他の被験者はHMDに直接触れるように片目を隠す動作を行っていた。手をHMDからやや離した状態で静止すると、手振れなどでIRカメラの画角に再び入ってしまうことがあり、複数回の判定が行われてしまい、今回の実験条件では失敗と判定されたと考える。

また、被験者が目を覆ってからオブジェクトの色が変わるまでの反応時間については、厳密には調査できていないが、記録した映像から目視で確認した範囲では、ほぼ0.2秒以内であった。このことから、片目を隠す動作に対応した視野変更までの遅延は、比較的小さいと思われる。詳細については、追って検証を進めていきたい。

7 議論

ここでは、「片側カメラが隠れる影響」「片目を隠す動作の共通認識」「動作認識の細分化」「片目を隠す動作の課題」について議論する。

まず、「片側カメラが隠れる影響」については、片目を隠す動作をすることでRGBカメラやIRカメラが手で覆われる。3章で述べたように、MetaQuest3では両側にRGBカメラ／IRカメラがあるため、その状態でもパススルー機能やハンドトラッキング機能は継続して利用可能だが、パススルー映像が片側分になる等、一定の制約が生じる。例えば、5.1節で紹介した応用例では、周辺状況を確認する際に、パススルーの視界は本来の半分となっている。一方、学会のデモ発表等で50人以上に応用例を体験してもらった中で、視野の制限等を指摘されたことはなく、影響は限定的であると考えている。

次に、「片目を隠す動作の共通認識」について述べる。性能評価では、「左右どちらかの手で片目を隠す」という指示だけを与えたが、全ての被験者は右手で右目、左手で左目の位置を覆うことが確認できた。このことから、片目を隠す動作には、一定の共通認識があり、理解しやすい動作であると考える。一方、動作中にHMDに手を触れる／触れない被験者がおり、後者の動作検出率はやや低下する傾向があった。「HMDに触れて片目を隠す」ような教示をすれば、動作検出率は向上すると考える。

次に、「動作認識の細分化」について述べる。本研究では、左手で覆う／右手で覆う／両目を覆う動作の区別は行なっていない。これらのそれぞれに対応する機能を割り当てる（例：左手で隠したらパススルー表示、右手で隠したらコンテンツ切り替え）でより幅広い活用ができる、実装も容易である。その一方で、動作のマッピングが複雑になり、直観性が損なわれるため、慎重な検討が必要になる。

最後に、「片目を隠す動作の課題」について述べる。まず、動作中は片手が塞がるため、コンテンツの操作方法が制限される。応用例で示したように、一時的な視界切り替え操作としての利用や、片手を用いた簡単な操作との併用は十分可能であるため、今後適切な利用場面を整理していく。また、手を肩より上の位置で空中で保持するような動作は、一般的に疲労しやすい。一方、提案手法は手をHMD上に接触させても動作するため、前述したように「HMDに触れて片目を隠す」ような指示をすることで、疲労を軽減できると考えている。

8 まとめと今後の展望

本研究では、片目を隠す動作でVRコンテンツとパススルー機能を重畠表示することで、仮想空間と現実空間のコンテンツを手軽に切り替え可能なインタラクション手法HidEyeを提案した。HMD内蔵のハンドトラッキング機能を用いることで外部センサ等を利用せず、動作検出を行う手法を実装した。また、本手法を用いて、「現実空間と仮想空間の移行」や、「仮想空間での視野変更」を行う、4種類の応用例を実装した。性能評価より、片目を隠す動作は被

験者に厳密な指示を与えなくても、高い精度で認識できることが確認できた。今後は提案手法を用いたより多くの応用例を実装・検証することで、新たなHMD用のインタラクション手法として確立していきたい。

謝辞

この成果の一部は、NEDO（国立研究開発法人新エネルギー・産業技術総合開発機構）の委託業務(JPNP23025)の結果得られたものです。

参考文献

- [1] I. Endo, K. Takashima, M. Inoue, K. Fujita, K. Kiyokawa, and Y. Kitamura. ModularHMD: A Reconfigurable Mobile Head-Mounted Display Enabling Ad-hoc Peripheral Interactions with the Real World. In *The 34th Annual ACM Symposium on User Interface Software and Technology, UIST '21*, p. 100–117, New York, NY, USA, 2021. Association for Computing Machinery.
- [2] J. Grubert, E. Ofek, M. Pahud, and P. O. Kristensson. The Office of the Future: Virtual, Portable, and Global. *IEEE Computer Graphics and Applications*, 38(6):125–133, 2018.
- [3] R. Ise and K. Tsukada. HidEye: Proposal of HMD Interaction Method by Hiding One Eye. In *SIGGRAPH Asia 2024 Posters (SA Posters '24)*, December 03-06, 2024, New York, NY, USA, 2024. Association for Computing Machinery.
- [4] K. Ito, H. Suzuki, H. Kugai, M. Takaoka, H. Matsumoto, M. Sakka, N. Yamamoto-Mitani, and A. Igarashi. HMD in General Public for Training: Case Study of Standardized Dementia-Friendly Initiative in Japan. In *2023 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–3, 2023.
- [5] T. Kawasaki and H. Manabe. LensTouch: Touch Input on Lens Surfaces of Smart Glasses. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, UIST '23 Adjunct*, New York, NY, USA, 2023. Association for Computing Machinery.
- [6] C. Mai and M. Khamis. Public HMDs: Modeling and Understanding User Behavior around Public Head-Mounted Displays. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays, PerDis '18*, New York, NY, USA, 2018. Association for Computing Machinery.
- [7] Y. Mais and J. Shaidah. A systematic review on hand gesture recognition techniques, challenges and applications. *PeerJ Computer Science*, 5:e218, 2019. <https://peerj.com/articles/cs-218.pdf>.
- [8] Y. Weng, C. Yu, Y. Shi, Y. Zhao, Y. Yan, and Y. Shi. FaceSight: Enabling Hand-to-

HidEye：片目を隠す動作による HMD 用インタラクション手法の提案

- Face Gesture Interaction on AR Glasses with a Downward-Facing Camera Vision. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.
- [9] ジョー・ナヴァロ (著), 西田美緒子 (翻訳). FBI 捜査官が教える「しぐさ」の実践解説辞典 407. 河出書房新社, 2019.
- [10] 山内耀斗, 岩崎琢己, 松井尚樹, 岩井大輔, 佐藤宏介. 断続的使用に適した胸部装着型 VR デバイスの提案. システム制御情報学会研究発表講演会講演論文集, pp. 342–346, 2023.
- [11] 伊勢 隆之介, 塚田 浩二. HidEye: 片目を隠す動作による HMD 用インタラクション手法の提案. インタラクション 2024 論文集, pp. 377–379, 2024.

高速投影による運動物体の残像色制御手法

幸谷 有紗 * 宮藤 詩緒 * 小池 英樹 *

概要. 本研究では、高速投影技術と残像効果を利用し、観測者視野内における物体の運動に応じた残像の目的色提示を実現する新たな手法を提案する。継時加法混色による投影中に投影フレームを遮る物体が動くと、物体の残像に意図しない色が視認される。本研究では、運動時の残像に目的の色を提示するため、目的色で構成された複数フレームとその補色で構成された單一フレームを生成し、両者を高速に投影する。提案手法はトラッキング装置を必要とせず、残像の色彩を自在に制御することで、より豊かな視覚体験を提供する。検証により、本手法が静止時の知覚に影響を与えることなく運動時の残像に目的色を提示できることが示された。本手法は、歩きスマホの抑制や、キャラクターとのインタラクション、カメラ撮影時の映像変化を利用した新しいインタラクションが期待される。

1 はじめに

人間の視覚系は、短時間に受け取った光情報を統合して色や形を認識している。この特性は、DLP プロジェクタやディスプレイ技術において、RGB 光による多彩な色表現に活用されている。特に DLP プロジェクタでは、カラーホイールと呼ばれる円盤状のフィルターを使用し、RGB 光を順次高速に投影することで、継時加法混色を実現している。しかし、継時加法混色において、高速移動する物体による遮蔽や急激な視線移動時に RGB の各色が分離して知覚されるカラーブレイキング現象が発生する。この現象は、人間の視覚における残像効果や光の積分特性に起因する。この知覚特性を利用し、運動速度や光の統合を制御することで、意図的な残像を知覚させる手法が提案されてきた。単色の LED ライトを振ることで映像を表現する手法 [2, 21] や、特定の運動ベクトル下で異なる映像を知覚させる手法 [8, 18] がこれに該当する。しかし、これらの手法は運動ベクトルに制約がある。そこで、我々の先行研究ではこの制約を緩和し、観測者の歩行時に特定の箇所に意図的にモザイク模様を知覚させる手法 [11] を提案した。しかし、ランダムな色での表現に制限されており、特定の色の絵やロゴを表現することはできない。本論文では、これらの問題を解決するため、運動方向の制約を最小限に留めたまま、投影フレームを遮る運動物体の残像に特定の色を提示できる新たな手法を提案する(図 1)。提案手法では、重畠する複数枚の投影フレームのうち、目的色で構成されたフレームの割合を増やすことで、残像の色を目的に近づける。本手法は、従来手法と比較して、より自由度の高い色制御を可能にし、観測者に多様な視覚体験を提供する。本論文の貢献は以下のとおりである。

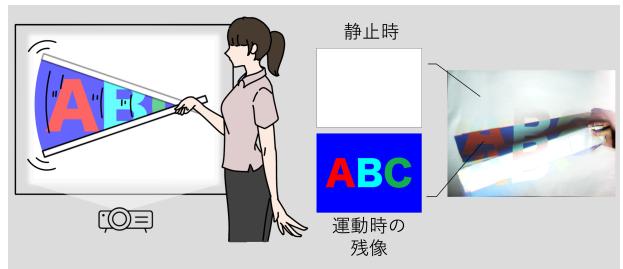


図 1. 投影フレームを遮る運動物体の残像に特定の色が観測される

- トラッキングを必要とせず、動きに応じた色知覚の変化を実現する新たな投影手法の提案
- 提案手法の視覚的原理に関する説明
- 提案手法を用いた応用事例の提示による実用性の議論

2 関連研究

2.1 高速プロジェクタを用いた動的プロジェクション技術

高速プロジェクタは、高フレームレートでの映像投影が可能なプロジェクタである。そのため、運動物体に対する動的プロジェクション技術の基盤となっている。この技術は、物体の動きに合わせてリアルタイムで映像を変化させ、観測者にインタラクティブで臨場感のある体験を提供する [14]。Narita ら [17] は、マーカと高速カメラを用いた高速追従型動的プロジェクション手法を開発した。また、Bermano ら [3] は、赤外線カメラと高速カメラを組み合わせ、低遅延の顔面への動的プロジェクション技術を提案した。さらに、Miyashita ら [16] は、マーカレスで高速移運動物体にテクスチャを付与する投影手法を実現した。

Copyright is held by the author(s).

* 東京科学大学 情報理工学院

これらのシステムは、カメラやマーカを利用して動的物体の位置情報を取得し、高速な動的プロジェクションを可能にしている。

2.2 残像効果を用いた知覚刺激の変化

人間の目は、残像効果により、一定の時間の光を重ね合わせて観測する。残像効果を用いた視覚刺激の変化手法として、棒状LEDライトの高速移動による2次元画像の提示が挙げられる[2, 21]。この効果をプロジェクションで起こすことで、観測者の動きに応じて異なる資格情報を提示することが可能である。Ikedaら[8]は、観測者の動く方向に応じて異なる白黒画像が認識される手法を提案した。この研究では光の時間的重なりを利用しているが、色彩表現に制約がある。また、Sakaueら[18]は、観測者の運動ベクトルに応じて異なるカラー画像を提示する投影手法を提案した。この手法では、運動ベクトルに応じて複数のカラー画像を提示できるが、移動は特定のベクトルに制限される。さらに、我々の先行研究[11]では、モザイクパターンを用いて歩行時の視線を誘導する手法を提案した。この研究では、赤、緑、青、黒の4色からなるモザイクパターンを映像内に埋め込み、色の配置を変えながら時間積分することで、観測者の運動時のみ知覚されるモザイクパターンを実現している。しかし、高速に時間積分される4色のモザイクは、運動時には虹色のに観測され、色彩を制御することはできない。これらの技術は、残像効果を用いた運動時の視覚変化手法として有用であるが、色彩や映像表現の自由度向上が課題となっている。

2.3 本研究の位置づけと目的

本研究では、以下の要件を満たす新たな高速投影手法を提案する：

1. 観測者や投影対象の位置情報を必要としない
2. 運動方向や速度の制約が少なく、自由な動きに対応可能である
3. 物体の運動によって生じる残像に目的の色を提示する
4. 物体の運動時と静止時で異なる映像知覚を実現する

本手法は、従来の高速プロジェクタを用いた動的プロジェクションとは異なり、運動物体の位置情報を必要とせず、残像効果を用いて運動物体の残像に対して特定の色知覚を実現する。これにより、既存のプロジェクション技術や残像効果を用いた手法の制約を解決し、より柔軟で効果的な視覚体験の提供を目指す。

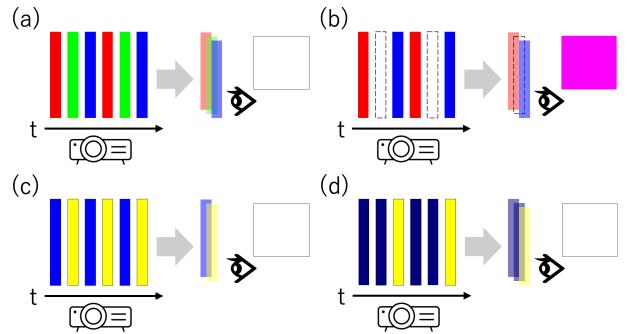


図2. 継時加法混色による色の知覚刺激の変化

3 手法

3.1 手法の概要

本研究では、残像効果と高速投影を組み合わせ、運動物体の残像に意図的な色を提示する新たな投影手法を提案する。本手法の特徴は、運動時の残像に特定の色を強調するため、複数枚で構成される投影フレームのうち目的色の割合を増やすことである。また、静止時には目的色の補色フレームを重畠することで、運動時とは異なる色を提示する。これにより、運動時と静止時で動的な色の変化を可能にする。

3.2 残像効果

人間の視覚システムは、残像効果により一定時間内の光刺激を統合して知覚する[4]。この性質を利用し、DLPプロジェクタではRGB光を継時加法混色することで多彩な色を表現している。赤(R)、緑(G)、青(B)の3色が高速に重なり合わさることで白色が知覚され(図2(a))、RGB光の割合を変更することで、多様な色を表現できる(図2(b))。R,G,Bそれぞれの色の合計画素値が一致するように各フレームを分解することで、黄色と青の重畠、さらに、黄色と暗い青2枚の重畠というように、重畠するフレームの色を変更して同じ色を知覚させることも可能である。提案手法では、 T 枚の投影フレームの t 枚目のフレームを $F(t)$ とし、 $F(0)$ から $F(T)$ までを合成すると静止時の目的画像 I_s となるように、投影フレームを設計する。

$$I_s = \sum_0^T F(t) \quad (1)$$

また、物体の運動時に表示したい目的の画像を I_t とする。この I_t を残像として知覚させるために、表示するフレーム $F(t)$ を設計する。物体が微小に移動する際、運動軌道にフレームが順次知覚される。そのため、運動時に目的のカラー画像 I_t を残像として知覚させるためには、フレーム内で I_t の色の割合を増やす必要がある。そこで、 $T - 1$ 枚の目的色フレームと1枚の補色フレームを用意することで、残

高速投影による運動物体の残像色制御手法

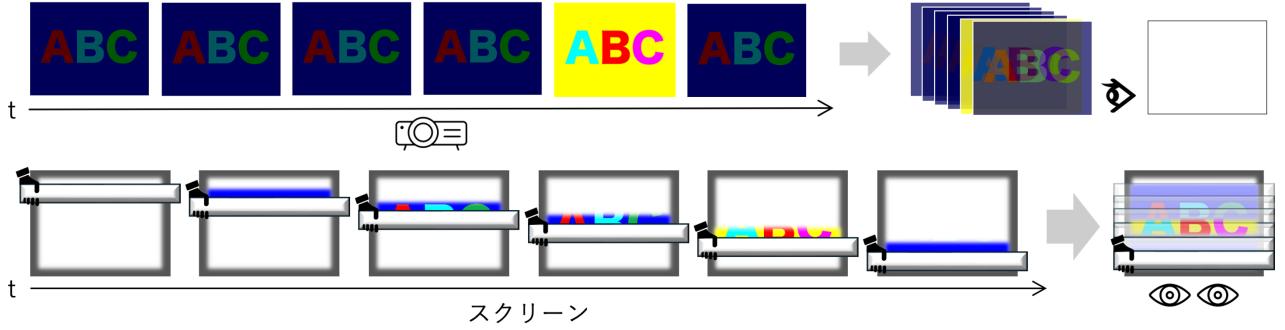


図 3. 投影フレームのうち目的色の割合を増やすことで、運動物体の残像を目的の色に近づけることができる

像中の目的色の割合を最大化し、残像の色を I_t に近づけることができる（図 3）。重畠するフレーム数 T が増加すると、個々のフレームの輝度は低くなるが、人間の視覚特性により暗い投影でも目的色を十分な輝度として認識できる [19]。これにより、物体の運動によって目的色の残像 I_t を知覚させることができる。

3.3 投影フレームの色の割り当て

上記のような条件を満たすように各投影フレームを作成する。画像 I_s を $T - 1$ 枚の暗い目的色フレーム F_d と、補色となる 1 枚のフレーム F_c に分割するため、 F_d と F_c を以下の式で求める。

$$F_d = \frac{I_t}{T - 1} \quad (2)$$

$$F_c = I_s - (T - 1)F_d \quad (3)$$

上の式から求められる $T - 1$ 枚の F_d と 1 枚の F_c を繰り返し投影することで、残像効果により全フレームが重畠し、静止時に見える画像 I_s が知覚される。このとき、運動時の目的画像 I_t は、静止時に観測される画像 I_s よりも小さな画素値で構成される必要がある。

4 実装

装置

本手法は PC による投影画像生成と高速プロジェクタによる投影によって実装した。図 4 にシステムで用いる装置構成を示す。高速投影には DynaFlash プロジェクタを使用する [22]。プロジェクタを制御する PC には RTX-3080 の GPU を搭載しており、高速投影処理に用いられている。また、投影環境は黒色布で覆われており、映像はスクリーンに投影される。

フレームの作成

図 5 に実装の概要を示す。実装の手順として、まず静止時に知覚する画像 I_s と物体の運動時に知覚する



図 4. 装置

画像 I_t を入力データとして定める。次に、式(2),(3)に基づいて投影フレーム F_d, F_c を生成する。ここで、フレームの分割数 T が大きいほど運動時に知覚される映像が目的色 I_t に近づくが、分割数が増えると投影周波数の増加とともに F_d の輝度が低下するため、色の再現性が減少する。このため、プロジェクタの投影速度や色域に応じて分割数 T を設定する必要がある。本実装では、分割数 T を 16、投影周波数を 960fps に設定した。これにより、16 枚のフレームが重畠された I_s は 60fps で観測される。この周波数は CFFF (臨界閾値周波数) [13, 15] を超えており、観測者がちらつきを感じないように選択されている。 $T = 16$ として式(2)に基づき、目的色画像 I_t の各画素値を 15 で割り、 F_d を算出する。なお、投影するフレームは 8bit の整数で表現される必要があるため、小数点を切り捨てて整数化する。その後式(3)により補色フレーム F_c を算出する。 F_d を 15 枚、 F_c を 1 枚、合計 16 枚のフレームを用意し、これらのフレームを 960fps で高速プロジェクタで繰り返し投影することで目的の視覚効果を実現する。

5 検証

5.1 目的

本章では、提案手法によって投影された映像が物体の静止時および運動時にどのように観測されるか

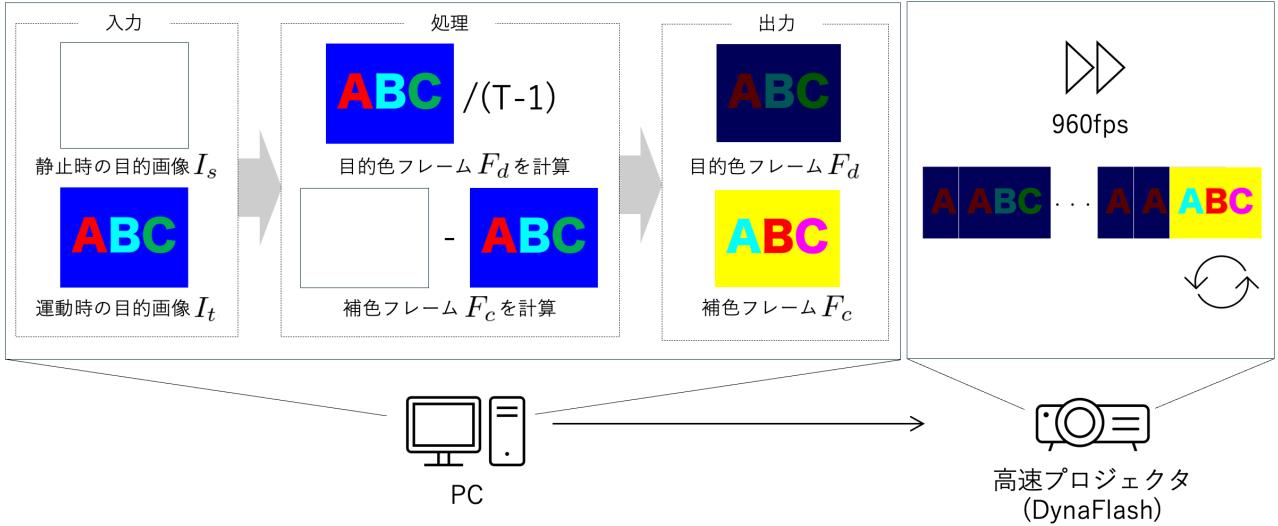


図 5. 実装の概要図

を確認し、さらに、フレームの分割が運動時の色知覚に与える影響を検証することを目的とする。特に以下の 2 つの仮説を検証する。

- H1: 投影結果は、物体の運動時に静止時とは異なる色が残像として観測される
- H2: 重畠する投影フレームにおける目的色の割合が増えることで、運動時の残像が目的色に近づく

5.2 評価方法

以上の仮説を検証するため、

1. 静止時と運動時における目的色と観測結果の色差の比較
2. 目的色フレームの異なる割合における目的色と観測結果の色差の比較

以上の 2 点を評価した。フレーム分割による目的色提示の効果を検証するため、本検証では目的色と補色の割合を変えてその効果を比較した。実験では静止時の目的色 I_s を 16 フレームに分割し、目的色 F_d と補色 F_c の割合を 4 種類の分割数:(a)8:8 (b) 12:4 (c) 14:2 (d) 15:1 で投影した。このとき、フレームの投影順序は (a): $F_dF_cF_dF_c\dots$, (b): $F_dF_dF_dF_cF_dF_dF_c\dots$ というように、補色フレームが時間的に最も離れるようにした。各フレームは 960fps で投影し、 I_s は 60fps の映像として観測される。観測結果の記録には、SONY ZV-E10 カメラを使用し、シャッタースピードは 60fps に設定した。この設定は人間の知覚に近く [15]、実際に観測される知覚を再現している。静止状態ではカメラを固定し、運動状態ではスクリーンの前で黒色の棒を縦に振ることで再現した(図 6 右)。投影する画像は、静止時と運動時の双方で様々な色が含まれるように設計した(図 6 左 1 列目)。静

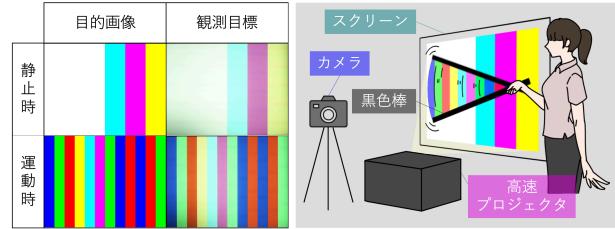


図 6. 実験の観測目標 (左) と実験の様子 (右)

止時には、画像の左半分が白色光、右半分が順にシアン (C), マゼンタ (M), イエロー (Y) を示し、運動時には、左から BGR Y CMGBR BRG の順に色が知覚されるようにフレームを作成した。これらの画像を投影し、撮影することで静止時および運動時における投影目標とした(図 6 左 2 列目)。(a)-(d) の各分割数について静止時と物体の運動時における投影結果を撮影し、投影目標と色を比較することで、色の再現性を評価した。色差の評価には CIEDE2000 [20] を使用し、観測目標と投影結果の各画素における色差を算出し、カラーマップに示した。

5.3 結果

各分割数で撮影された投影結果と、投影結果と観測目標との色差を図 7 に示す。静止時の結果より、どの分割数においても色差のカラーマップが全体的に小さな値となっており、観測目標に近いことが示された(図 7 上段)。また、運動時の結果より、いずれの分割数においても残像に静止時とは異なる色が知覚されることが確認された(図 7 下段)。特に棒の周辺には重畠されたフレームの色が提示されており、フレームの分割割合に応じてその色の割合が変化している。これらの結果から、仮説 H1 が支持される。

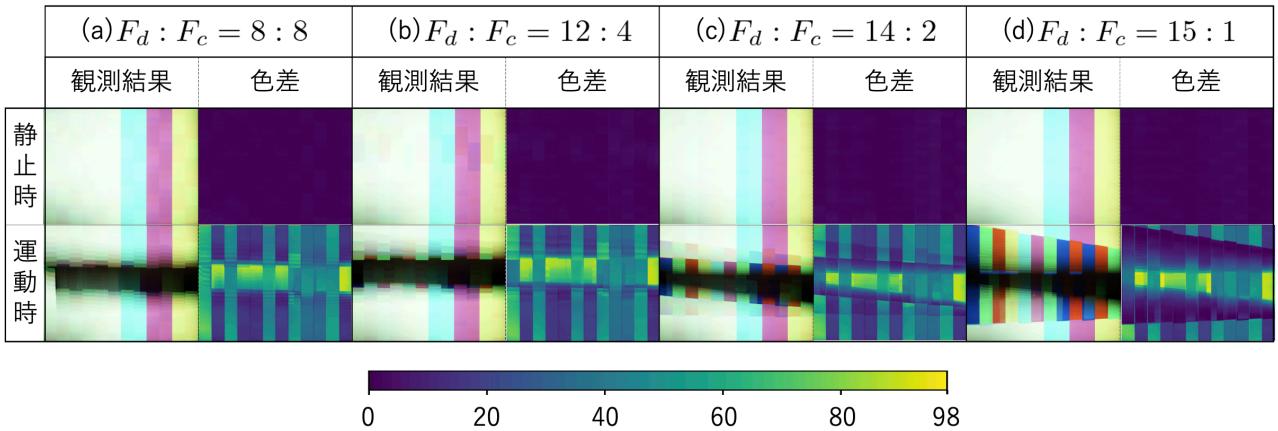


図 7. 各分割数における投影結果と、観測目標との色差を示すカラーマップ

また、運動時の残像色について、(a) では残像の色はランダムに表現されており、残像の色を制御できていない一方、分割数を増やすにつれて残像の色がより鮮明になり、目的色との色差も減少している。これにより、仮説 H2 が支持される結果となった。

6 アプリケーション

本手法は、トラッキング技術や観測者の動きに応じた映像の変更を伴わずに運動物体の残像の色を操作できる。また、個別の着用デバイスも不要であり、公共の場においても幅広く応用可能である。以下に、具体的なアプリケーション例を示す。

6.1 環境光への埋め込み

Brittany ら [5] は、投影を環境光と一体化することで、実世界の物体に対する知覚を変化させる手法を提案した。本手法を白色環境光として利用することで、日常的な状況において、物体の運動時にのみ知覚を変化させることができる。以下に二つの具体例を示す。一つ目は歩きスマホ防止のアプリケーションである。図 8(a) は静止時にスマートフォンを使用している場合には白色環境光として知覚される一方、歩行時にはスマートフォンの動きに応じて、警告を意味する映像を周囲に表示することができる例である。これにより、駅や商業施設での歩きスマホを効果的に抑制することが期待できる。二つ目は、人の往来の可視化である。図 8(b) は、人の進行方向に応じた残像を提示することができる例である。これにより、人の動きに応じた空間演出が可能となり、空間の雰囲気を動的に変えることができる。

6.2 インタラクションへの応用

本手法を利用することで、観測者の動きに応じた映像変化を通じて、様々なインタラクションが実現

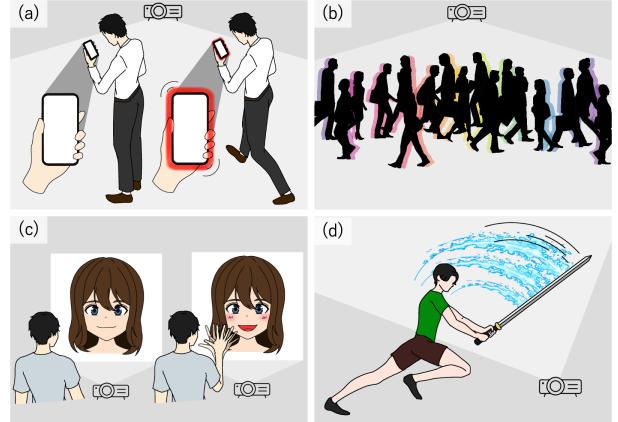


図 8. アプリケーション例:(a) 歩きスマホの抑制 (b) 人の往来への残像提示 (c) キャラクターとのインタラクション (d) 振りかぶった剣への効果線の提示

可能である。以下に二つの具体例を示す。一つ目はキャラクターとのインタラクションである。提案手法を用いることで、静止時にはキャラクターが表示され、観測者が手を振るなどの動作を行うと、運動時用に用意されたエフェクトをキャラクターに追加することができる。図 8(c) は、観測者がキャラクターの映像に手を振ることで、キャラクターが微笑むような知覚映像の変化を与えていた例である。これにより、動作を通じたキャラクターとの対話的な体験を提供することができる。二つ目は効果線の提示である。提案手法を用いることで、剣やバットを振った際に、残像に効果線を表示することが可能となる。図 8(d) は、剣を振った際に、水のエフェクトが残像に表示される例である。これにより、現実世界でのアクションが視覚的に強化され、より没入感のある体験を提供できる。

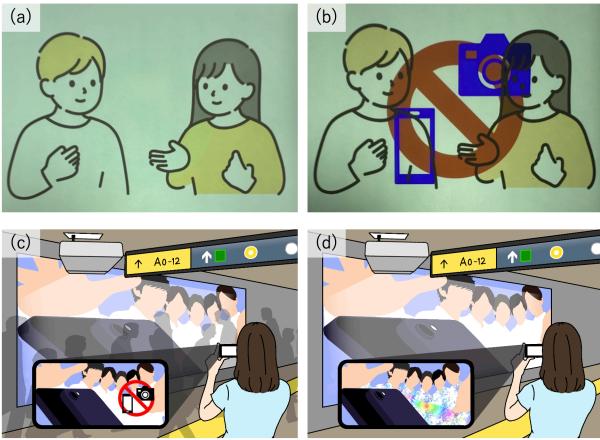


図 9. カメラ撮影による映像の変化 (上段) と、駅の広告におけるアプリケーション例 (下段)

6.3 カメラとのインタラクション

本手法は、カメラを用いた新しいインタラクション手法としても応用可能である。先行研究 [1, 6, 10] では、投影とカメラを同期させることで、映像内にバーコードのような二値画像を埋め込み、撮影を通じたインタラクションを実現している。本手法では、カメラで撮影することで、シャッタースピードとデジタル処理によるフレームの欠損を利用し、運動時の観測と同様に特定の映像を知覚させることができる。この手法は、カメラと映像の同期を必要とせず、一般的なスマートフォンのカメラでも容易に応用できる。図 9 上段の画像は、実際に提案手法を用いて映像に撮影禁止のピクトグラムを埋め込んだ例である。人間の知覚よりも長いシャッタースピードで撮影した場合には静止時の目標画像となり(図 9(a)), シャッタースピードが短い場合は運動時の目標画像となる(図 9(b))。特に、スマートフォンのシャッタースピードは手ブレを防ぐために高速であることが多いため、運動時の目標画像が撮影される可能性が高い。カメラを用いた具体的なアプリケーションとして、駅の広告における撮影抑制・促進が挙げられる。駅には多くの魅力的な広告が掲示されており、その中には多くの人々に撮影されるものもある。広告の写真撮影は、その後の拡散によって広告効果を高める可能性があるが、ラッシュ時には撮影者が通行人の妨げになることが懸念される。図 9(c)は、人の往来の多いラッシュ時には広告の撮影を妨げ、図 9(d)は、閑散時には広告内の人物の服装の変化という撮影による付加価値を提供し、撮影を促進することができる例である。これにより、適切なタイミングでの広告の拡散をコントロールすることが可能となり、広告効果を最大化することができる。

7 議論

本研究では、残像効果と高速投影を組み合わせ、運動時の残像に目的の色を提示する投影手法を提案した。評価実験においても、静止時および運動時の双方において目的の色を提示できることが確認された。しかし、本手法にはいくつかの制約が存在する。**運動速度に関する制約**

まず、運動速度に関する制約が挙げられる。本手法が効果的に機能するためには、物体の運動が人間の目の残像時間を上回る速度で行われる必要がある。この条件を満たさない場合、運動時に意図した色や画像が十分に知覚されない可能性がある。

運動時用画像における周波数成分の制約

次に、運動時用画像における周波数成分の制約が挙げられる。高周波成分を含む画像は運動時により顕著に知覚されやすいが、ぼやけた文字や模様など、低周波成分を含む低い顕著性の画像は知覚が困難である [9]。そのため、エッジの強調や大きな色の変化を含む画像を用いることで、より効果的に目的の映像を提示することができる。

投影周波数の制約

さらに、適切な投影周波数の選定における制約も存在する。CFFF (臨界閾値周波数) が環境光、輝度、色彩条件、個人差に応じて変動するように [1, 7]、本手法における投影分割数と投影周波数も、これらの要因に依存する。また、周辺視のほうが中心視よりもフリッカー周波数に対して敏感であるため [12]、中心視では観測されなかったフリッカーが、周辺視では知覚される可能性がある。本研究における実験では、分割フレーム数を 16、投影周波数を 960fps に設定したが、環境光の強さ、投影画像の彩度、補色の輝度などの要因によって、最適な投影周波数は変動することが考えられる。これらの制約を踏まえ、今後より広範な条件下での適応性を検討することが必要である。

8 結論

本研究では、高速投影技術と残像効果を組み合わせることで、物体の運動に応じた残像の目的色提示を実現する新たな手法を提案した。提案手法では、運動時の残像に目的の色が強調されるよう、複数の目的色フレームと 1 枚の補色フレームを高速で繰り返し投影する。検証の結果から、提案手法による目的色フレームの分割は、残像の色の制御に有効であることが示された。しかし、運動速度や投影周波数における制約があり、今後制約の緩和が望まれる。本手法は、トラッキング装置を必要としないため、公共の場でのインタラクションやエンターテインメント分野への応用が期待される。

謝辞

本研究は JSPS 科研費 20H04221 の助成を受けたものです。

参考文献

- [1] S. Abe, A. Arami, T. Hiraki, S. Fukushima, and T. Naemura. Imperceptible Color Vibration for Embedding Pixel-by-Pixel Data into LCD Images. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, p. 1464–1470, New York, NY, USA, 2017. Association for Computing Machinery.
- [2] H. Ando, J. Watanabe, T. Amemiya, and T. Maeda. Full-Scale Saccade-Based Display: Public / Private Image Presentation Based on Gaze-Contingent Visual Illusion. In *ACM SIGGRAPH 2007 Emerging Technologies*, SIGGRAPH '07, p. 5–es, New York, NY, USA, 2007. Association for Computing Machinery.
- [3] A. H. Bermano, M. Billeter, D. Iwai, and A. Grundhöfer. Makeup Lamps: Live Augmentation of Human Faces via Projection. *Computer Graphics Forum*, 36(2):311–323, 2017.
- [4] J. R. Brockmole, R. F. Wang, and D. E. Irwin. Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2):315, 2002.
- [5] B. Factura, L. LaPerche, P. Reyneri, B. Jones, and K. Karsch. Lightform: procedural effects for projected AR. In *ACM SIGGRAPH 2018 Studio*, SIGGRAPH '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [6] A. Grundhöfer, M. Seeger, F. Hantsch, and O. Bimber. Dynamic Adaptation of Projected Imperceptible Codes. ISMAR '07, p. 1–10, USA, 2007. IEEE Computer Society.
- [7] S. Hattori and T. Hiraki. Accelerated and Optimized Search of Imperceptible Color Vibration for Embedding Information into LCD Images. In *SIGGRAPH Asia 2022 Posters*, SA '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [8] H. Ikeda, T. Hayakawa, and M. Ishikawa. Bilateral Motion Display: Strategy to Provide Multiple Visual Perception Using Afterimage Effects for Specific Motion. VRST '19, New York, NY, USA, 2019. Association for Computing Machinery.
- [9] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.
- [10] R. Junichi. Temporal Phase Shift: 時間位相差を持つプロジェクションシステムによる錯視現象とそのモデル化. WISS2009, p. 0028, 12 2009.
- [11] A. Kohtani, S. Miyafuji, K. Uragaki, H. Katayama, and H. Koike. MOSion: Gaze Guidance with Motion-triggered Visual Cues by Moasic Patterns. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery.
- [12] B. Krajancich, P. Kellnhofer, and G. Wetzstein. A Perceptual Model for Eccentricity-Dependent Spatio-Temporal Flicker Fusion and Its Applications to Foveated Graphics. *ACM Trans. Graph.*, 40(4), jul 2021.
- [13] C. Landis. Determinants of the Critical Flicker-Fusion Threshold. *Physiological Reviews*, 34(2):259–286, 1954. PMID: 13155188.
- [14] J. Lee, Y. Kim, M.-H. Heo, D. Kim, and B.-S. Shin. Real-Time Projection-Based Augmented Reality System for Dynamic Objects in the Performing Arts. *Symmetry*, 7(1):182–192, 2015.
- [15] N. D. Mankowska, A. B. Marcinkowska, M. Waskow, R. I. Sharma, J. Kot, and P. J. Winklewski. Critical Flicker Fusion Frequency: A Narrative Review. *Medicina*, 57(10), 2021.
- [16] L. Miyashita, Y. Watanabe, and M. Ishikawa. MIDAS Projection: Markerless and Modelless Dynamic Projection Mapping for Material Representation. *ACM Trans. Graph.*, 37(6), dec 2018.
- [17] G. Narita, Y. Watanabe, and M. Ishikawa. Dynamic Projection Mapping onto Deforming Non-Rigid Surface Using Deformable Dot Cluster Marker. *IEEE Transactions on Visualization and Computer Graphics*, 23(3):1235–1248, 2017.
- [18] F. Sakaue and J. Sato. Active 3D Motion Visualization Based on Spatiotemporal Light-Ray Integration. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1977–1985, 2020.
- [19] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcova. *High Dynamic Range Display Systems*. Association for Computing Machinery, New York, NY, USA, 1 edition, 2023.
- [20] G. Sharma, W. Wu, and E. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research Application*, 30:21 – 30, 02 2005.
- [21] J. Watanabe, T. Maeda, and H. Ando. Gaze-Contingent Visual Presentation Technique with Electro-Ocular-Graph-Based Saccade Detection. *ACM Trans. Appl. Percept.*, 9(2), jun 2012.
- [22] Y. Watanabe, G. Narita, S. Tatsuno, T. Yuasa, K. Sumino, and M. Ishikawa. High-speed 8-bit Image Projector at 1,000 fps with 3 ms Delay. 12 2015.

机上に投影した空中立体映像による方向指示の精度調査

松浦 向日葵* 福地 健太郎*

概要. 机上に映像を投影し、机上の実物体の説明や操作指示を行うシステムがこれまでに提案されている。多くの場合、矢印や記号類を投影することで指示を表現するが、小さな人間型キャラクタを表示して説明や指示をすると、利用者との対話感や共在感、楽しさを向上させることができ、博物館展示などに向く。しかし、投影する映像が平面の場合、キャラクタの方向指示精度に課題があった。本研究では、キャラクタを空中立体映像として投影することにより、方向指示精度の改善を図った。評価実験の結果、平面映像投影と比較して、方向指示の認識精度が有意に向上することが明らかになった。これにより、指示対象となる実物体が多数あるような環境における、空中立体映像の優位性が示唆された。

1 はじめに

プロジェクトやHMDを使用して机上に映像を投影し、机上の実物体の説明や操作指示を行うシステムが提案されている。多くの場合、矢印や記号などの図形を実物体近傍に投影して指示を表現するが[1, 2, 3, 4, 5]、小さな人間型キャラクタを表示して説明や指示をすることで、利用者との対話感や共在感、楽しさを向上させることができる[6]。

これまでに我々は空中立体映像投影を用いて、机上に身長10cmほどのキャラクタを投影するシステムを構築した(図1)。同システムにより、キャラクタが実際に机上に立っているように見せ、机上の実物体を指差したり近づいたりといった指示が可能となった。

指差しによる方向指示には精度の問題がある。操作者が方向指示精度の限界を把握していないと、対話相手に関心の対象を正しく伝えられない恐れがある。キャラクタを指示対象に近づけたり、他の情報で補うにしても、事前に指示精度の特性を把握していないとその必要を見落とす可能性がある。

そこで今回、提案システムにおける指差し方向の認識精度を評価する実験を行った。この実験では線分が方向を示す映像を空中に投影し、被験者には示されたと思う方向を推測させるタスクを行った。このタスクを2D映像と3D映像の2条件で実施し、実際に示された方向と被験者が答えた方向の誤差を分析した。

実験の結果、提案システムにおいて、3D映像による方向提示は平均で約7度の角度誤差があった。2D

映像では約11度であり、立体映像投影の優位性が確認された。特に、被験者から見て奥側を示した際に、2D映像では角度誤差が最大約20度まで低下したが、3D映像では5度以下に抑えられることが分かった。一方、被験者から見て手前側や真横を示した際は差が小さく、効果は限定的であった。

のことから、提案システムでの空中立体映像による方向提示は、机上の位置を指示するのに十分な精度を持つことが分かった一方で、平面映像の投影では被験者から見て奥側60~20度の角度の指示に問題があることが分かった。



図1. テーブル型空中立体映像投影システム
(AIRR テーブル)

2 関連研究

キャラクタによる方向指示の認識精度を向上させるにはキャラクタを立体的に提示することが有効である。キャラクタの全身を3D映像として投影する手法[7]や、身体をロボットで代替して表現する手法[8]のいずれにおいても、観察者は対話相手による方向指示を高い精度で認識できたことが示されている。これらの研究では、方向指示は腕のみならず頭部や

Copyright is held by the author(s).

* 明治大学

胴体の姿勢を含む全身を用いたものである。鈴木らはロボットが無言であっても身体の動きのみで注意誘導が可能であることを示した[10]。特に、指差し行為による方向提示が有効であることを明らかにしている。

また、視線による対象指示に着目し、対話相手の視線を伝達することに特化した研究が行われている。Higuchi らは遠隔協調作業において、2D プロジェクションマッピングまたは HMD を用いて、リモートユーザーが視線を向けた場所をローカル環境に図式的に表示するシステムを製作した[1]。リモートユーザーの視線位置を提示することで、早く正確な対象指示が可能となったことが報告されている。大槻らはビデオ通話をを行う対話相手の視線方向を立体的な眼球型ディスプレイを用いて提示する手法を提案している[9]。対話相手が視線を向けた実物体に観察者の注意をより早く誘導できることが示されている。

Izuka らはより生物的な頭部ジェスチャを実現するため、リモートユーザーの頭部を投影するフレキシブルスクリーンを提案した[11]。このスクリーンは投影面を柔軟に変形させることができ、頭部の向きを立体的に提示することができる。

3 実験システム概要

3.1 空中立体ディスプレイ

本研究で用いた、テーブル型の空中立体映像投影システムについて説明する。我々はこれまでに、山本[12]が開発した空中結像技術である AIRR (Aerial Imaging by Retro-Reflection) とレンチキュラー方式の立体ディスプレイである Looking Glass 8.9"を組み合わせ、空中立体映像を投影できるテーブル型ディスプレイを製作した(図 1)[13]。以下、このディスプレイを AIRR テーブルと呼称する。

AIRR テーブルでは、観察者は立った状態で映像を見る想定している。テーブル面の上には実物体を置くことができ、空中立体映像と実物体を並べて見せることができる。Looking Glass は、両眼視差と左右方向に約 60 度の運動視差を伴う裸眼で観察可能な立体映像を投影できるディスプレイであり、この映像を空中に再結像させることで、AIRR テーブルにおいても両眼視差と運動視差のある空中立体映像を提示することができる。図 2 に AIRR テーブルに投影した 1 つの空中立体映像を左・中央・右の 3 視点から撮影した写真を示す。このように、AIRR テーブルに投影した空中立体映像は視点位置に応じて見え方が変化するようになっている。

空中立体映像の提示手法として使用した AIRR の原理を図 3 に示す。AIRR とは再帰性反射材とハーフミラーによって構成される光学系による空中結像技術である。ディスプレイから出た光はハーフミラーによって鏡面反射され、再帰性反射材に入射する。ここで再帰反射された光はハーフミラーを通過し、空中で結像する。空中像はハーフミラーを基準としてディスプレイの対面位置に結像される。



図 2. AIRR テーブルに投影した 1 つの空中立体映像を
左・中央・右の 3 つの視点から撮影した写真

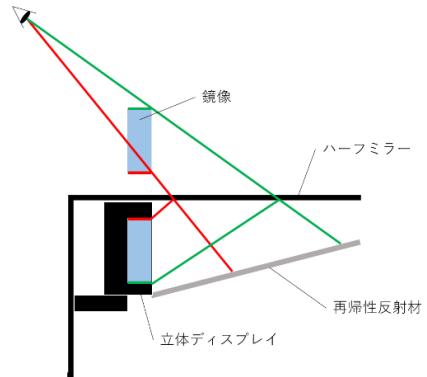


図 3. AIRR の原理図

3.2 設計

AIRR テーブルの概略図を図 4 に示す。テーブルの大きさは天板が 110cm×61cm、床面から天板までの高さが 107cm である。天板にはアクリルハーフミラーを使用しており、テーブルはアルミフレームによって構築している。天板の下には立体ディスプレイを垂直に置き、再帰性反射材は立体ディスプレイの映像表示面と向き合うように取り付けている。立体ディスプレイは Looking Glass 8.9" の映像表示面に設置されていたアクリルブロックを取り外したものである。再帰性反射材は日本カーバイド工業製の RF-Ay であり、25cm 四方のシート 8 枚を横に 4 枚、縦に 2 枚並べて全体で横 100cm、縦 50cm のシートを作成した。シートはなるべく隙間が出来ないように敷き詰めている。ただし、シートの縫合目が立体ディスプレイの映像表示範囲の前にある状態で空中立体映像を投影すると、映像には縫合目の位置に応じた黒い縦線が入ってしまう。そのため、実験時にはシートの縫合目と映像表示面とが向き合わないようにディスプレイを設置した。また、周囲からの光を遮蔽するために天板の上には天蓋を取り付けている。

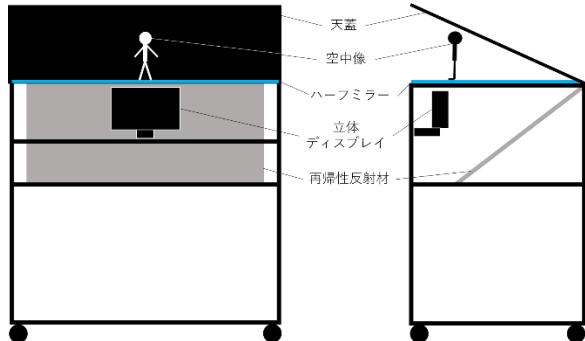


図 4. AIRR テーブル概略図

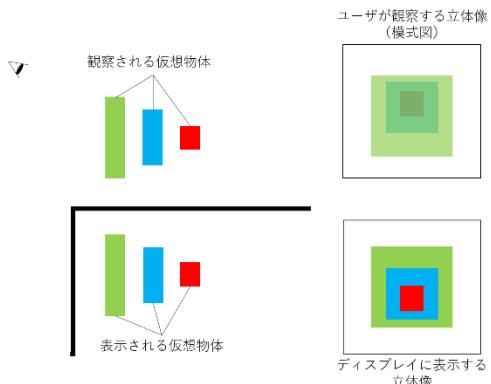


図 5. AIRR テーブルに投影した空中立体映像の前後が反転して見える様子の模式図

3.1 節で触れた通り、Looking Glass の映像はテーブル面を基準として対称の位置に結像する。つまり、空中映像は立体ディスプレイに表示された映像の上下および前後が反転した状態のものとなる。そのため、観察者に正しい像を見せるためにはこの反転を考慮した映像を立体ディスプレイに表示する必要がある。このうち、前後方向の反転について詳述する。

仮に、ディスプレイ上に図 5 下図のような前後関係で仮想物体を表示したとする。鏡像はテーブル面を境界面として対称の位置に結像するため、ユーザが観察するのは図 5 上図のようにディスプレイ映像を裏から見透かしたようなものとなる。すなわち、ユーザはディスプレイに表示された像の前後が反転した像を観察することとなる。そのため、観察者に正しい空中立体映像を見せるには、立体ディスプレイに投影する映像の上下及び前後を反転しておく必要がある。

また、AIRR テーブルに投影した空中映像には、実物体を映像の奥側に置いたとしても映像が実物体によって遮蔽されてしまうという問題がある。よって空中映像の周囲に実物体を設置する場合は、映像に被る形で映像より後ろに実物体を置かないよう注意する必要がある。

4 実験

4.1 実験概要

本実験は、AIRR テーブルでの指差し方向の認識精度を評価するために実施した。実験では AIRR テーブルの天板上に実物体の円柱 16 個を等間隔に設置し、そのうちの 1 個を線分で指し示した後、被験者には示されていると思う円柱の番号を口頭で答えさせるというタスクを実施した。線分は 2D 映像または 3D 映像として AIRR テーブルに投影した。実験後、実際に線分が示した円柱と被験者が回答した円柱の位置との差を 2D 映像と 3D 映像の 2 条件間で比較した。

今回は、実物体を空中立体映像によって示した際に、観察者に示した実物体が正しく伝わるか検証するため、上記のような実験設計としている。また、今後は人間型キャラクタを空中立体映像に投影することを想定しているため、本来はキャラクタが実物体を指差す形式で実験を行うべきであった。しかし、AIRR テーブルでは身長 10cm ほどの小型キャラクタを投影することを前提としており、指す向きによつては指や腕が判別しにくいという問題が生じた。そこで、今回は空中立体映像による方向提示の認識精度の上限を調査するため、紛れの少ない線分による方向指示を採用した。

4.2 実験環境

実験の様子を図 6 左図に示す。被験者の視点は床から 149cm、空中映像の中心から 78cm の位置に設定した。実験中、被験者の視点位置を固定するためには被験者には台に頸を乗せた状態で映像を見るように指示した。実験で使用した実物体は直径 1.5cm、高さ 10.5cm の円柱 16 個である。これらの円柱を、テーブル天面の中央を中心とした、半径 13cm の円周上に 10 度ごとに配置した。この半径は、円柱がテーブル面に収まるように設定した。

円柱を指し示す線分は、やはりテーブル天面の中央を端点として表示した。線分の長さは、真横を指したときの長さが 4.7cm となるようにした。この長さは、どの方向を指し示す際も線分が Looking Glass の表示範囲内に収まるように設定した。なお、3D 映像における線分は直径 0.235mm、高さ 4.7cm の円柱であり、2D 映像における線分は縦 0.235mm、横 4.7cm の長方形である。線分の高さは、テーブル天面から 6.48cm の高さに設定した。線分の空間上における配置を図 7 右図に示す。

3.2 節で触れたように、本実験で使用した空中映像には、実物体を被験者から見て映像の奥側に置いたとしても映像が実物体によって遮られてしまうという問題がある。そのため、被験者から見て奥側の

円柱が空中映像に被らないように、円柱は被験者から見て手前側には 90 度分、奥側には 60 度分配置した。被験者から見た円柱と空中に投影された線分を図 6 右図、両者の位置関係を図 7 左図に示す。



図 6. (左)実験の様子。
(右)被験者から見た円柱と線分



図 7. (左)線分と円柱の位置関係を上から見た図。
(右)テーブル面上における線分の配置。矩形は、机上に投影された立体ディスプレイの鏡像を表す。

4.3 実験で用いた映像

実験では、線分が 16 個の円柱のうち 1 個をランダムに示すという動きを連続で 48 回行い、その様子を空中に 2D 映像と 3D 映像でそれぞれ投影する。図 8 に線分を 2D 映像、3D 映像として空中にそれぞれ投影し、右・中央・左の 3 つの視点から撮影した写真を示す。2D 映像での線分は視点位置によらず同じ映像が見えるのに対し、3D 映像では視点位置に応じた映像が見えるため、両眼視差および運動視差による立体感を得ることができる(ただし実験では頭部位置を固定しているため、被験者は両眼視差のみで立体感を得る)。なお、2D 映像投影は Looking Glass に全て同一の仮想カメラ位置からの映像を表示することで実現した。

円柱を示す順番は 2D 映像条件と 3D 映像条件とで異なるようにした。これは、被験者が先に見た映像で円柱の示される順番を覚え、後に見る映像で覚えた順番を答えてしまうことを避けるためである。

円柱をランダムに指示する際に、近い距離にある 2 つの円柱を順に指示すると、被験者にとって方向を推測する手掛かりとなってしまう。今回の実験では指示する方向毎の精度を調査するため、連続して

指し示す円柱間の距離もランダムにばらけるよう、以下の手法で指し示す順序を定めた。まず、円柱間の角度差は、被験者から見て手前側をプラス方向、奥側をマイナス方向とすると、同一方向の円柱を続けて指すのを避ける場合、-150 度～+150 度の 30 通りの角度差がある。この角度差がまんべんなく出現する順序で円柱を指すのがよいと仮定し、円柱が指される回数を 3 回とした配列を生成した上で、これを焼きなまし法を用いて、配列をシャッフルしつつ、各角度差ができるだけ均質になるように順序を入れ替える作業を行った。また、線分は 1 秒間表示された後に 2 秒間消え、また 1 秒間表示されるといった動きを繰り返す。

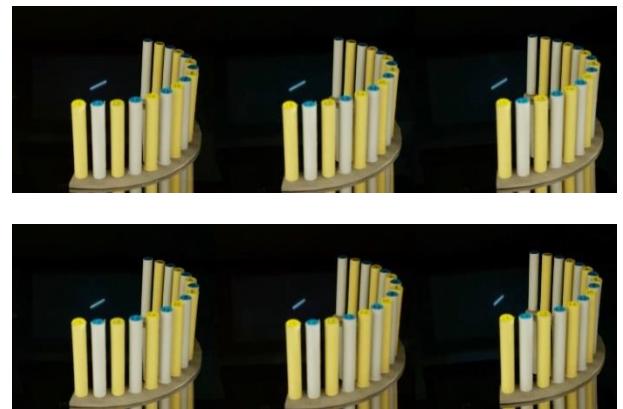


図 8. 空中に投影した 2D 映像(上)、3D 映像(下)を右・中央・左の 3 つの視点から撮影した写真。交差法による裸眼立体視をしやすい順に配置している。

4.4 被験者

被験者は理工系学部に所属する大学生・大学院生 20 名である。被験者 20 名は 2D 映像→3D 映像の順にタスクを行う群(2D 先行条件群)10 名と、3D 映像→2D 映像の順に行う群(3D 先行条件群)10 名の 2 群に、無作為に分けて実験を実施した。

4.5 実験手順

実験では、16 個の円柱のうち 1 個を線分で示した後、被験者には指されていると思う円柱の番号を口頭で答えさせるという課題を行った。以下、実験手順の詳細を説明する。

まず、実験についての説明を行った。この時、被験者に線分の表示時間について説明し、円柱の番号は出来るだけ素早く直感で答えるように伝えた。なお、実験本番で円柱の番号を答え損ねた被験者はいなかった。

次に、実験手順の確認を行った。被験者には線分が 1 番から 16 番の円柱を順番に 1 つずつ示す様子を空中に 3D 映像で投影したものを見せた。被験者に円柱を示す順番は伝えていない。線分が円柱を 1 つ示すたびに、被験者には示されたと思った円柱の

番号を口頭で答えてもらった。

続いて、実験本番を行った。被験者に線分が円柱を示す様子を2D映像または3D映像で48回見せ、示されたと思った円柱の番号を口頭で答えてもらった。1分間の休憩を挟んだ後、2D先行条件群には3D映像、3D先行条件群には2D映像を見せて同様のタスクを行った。被験者が答えた番号は実験者がその都度記録した。

2つのタスク終了後に、視力に関するアンケートに回答してもらった。アンケートでは「視力を矯正しているか」「普段、立体視に関して何か問題を感じたことはあるか」「日常生活において、物体の距離や深さを見極めることに困難を感じることはあるか」「差し支えない範囲で現在の視力の状況について」の4項目を質問した。

5 結果と考察

5.1 実験結果

線分が実際に示した円柱と被験者が答えた円柱の位置の角度の差を誤差として、方位ごとの誤差を集計した。

なお、2D映像条件において、線分が実際に示した円柱が被験者から見て最も手前側にある16番であるのに対して、最も奥側にある1番と答えたデータが1件あった。これ以外に2D条件・3D条件のどちらにおいても、線分が示した方向の前後を間違えた回答はなかったため、これは外れ値として分析の対象外とした。線分が示した方向の前後を錯誤する回答が1件になった理由として、実験手順を確認する際に線分が被験者から見て最も奥側にある円柱から手前側の円柱を1つずつ示していく様子を見せていくことが関係していると考えられる。線分が被験者から見て手前側を示している時は、線分は円柱に近い方の端点にかけて下側に傾いているように見える。反対に、奥側を示している時は線分は円柱に近い方の端点にかけて上側に傾いているように見える。このような線分の見え方について、被験者は実験手順を確認する際に前提知識を得ていたことから、前後を間違えた回答は前述の1件以外に発生しなかったと思われる。よって、線分が示した方向の前後を単純に誤認する可能性は低いと考え、その1件は外れ値として処理した。

まず、被験者ごとに2D映像条件・3D映像条件それぞれにおける角度誤差の平均を比較した結果を図9に示す。図9中のエラーバーは標準偏差を示している。ここで、2D条件と3D条件の間に等分散性があるとは見なせず、2D条件より3D条件の方が角度誤差が小さくなることが想定されるため、検定にはWelchの片側t検定を用いた(被験者1の2D条件は

N=47、その他はN=48)。結果より、被験者1~4・6・7・9・10・12・14~17について、2D条件より3D条件の方が角度誤差の平均が有意に小さいことが分かった(被験者2・3・6・7・16・17は*p<.001、4・12・15は†p<.01、1・9・10・14は**p<.05)。その他の被験者については、2D条件・3D条件間での有意差は確認されなかった。

次に、2D映像条件と3D映像条件との精度の差を調べるため、各条件における被験者全員の角度誤差の平均を比較した(図10)。なお、図10中のエラーバーは標準偏差を表している。Welchの片側t検定(2D条件はN=959、3D条件はN=960)を行ったところ、2D条件より3D条件の方が角度誤差の平均が有意に小さいことが分かった(*p<.001)。

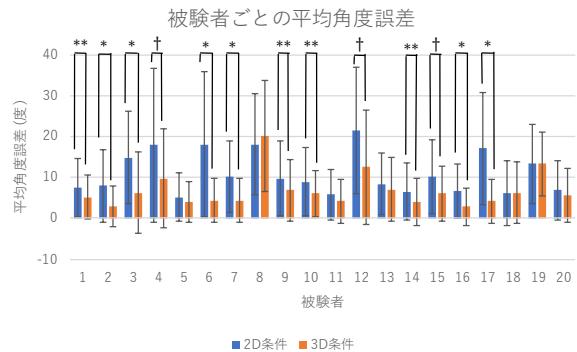


図9. 被験者ごとの角度誤差の平均
(*p<.001, †p<.01, **p<.05)

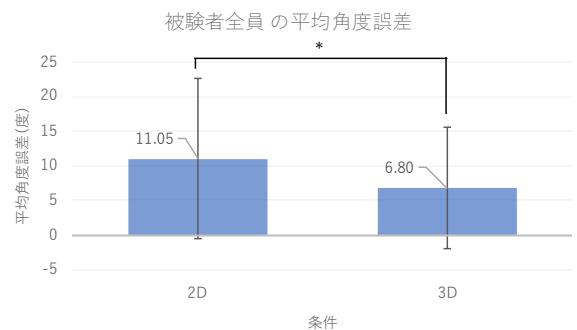


図10. 被験者全員の角度誤差の平均(*p<.001)

最後に、2D映像条件と3D映像条件で線分が示した方位別に角度誤差の平均を比較した結果を図11に示す。この比較は、方位ごとの認識精度が2D条件と3D条件とで異なるかどうかを調査するために行った。図11中のエラーバーは標準偏差を示している。なお、ここでの角度誤差は、被験者から見て手前側の方位をプラス方向、奥側の方位をマイナス方向とし、向きを考慮して計算している。検定にはWelchの両側t検定を用いた(2D条件の+90度のみ

N=59, その他は N=60). -60~-20 度では 2D 条件・3D 条件のどちらにおいても角度誤差の平均は正の値となった。さらに、2D 条件に比べて 3D 条件の方が角度誤差の平均の絶対値は有意に小さいことがわかった(図 11 中 -60~-30 度で $p<.001$ (*), -20 度で $p=0.005$ (†)). +50~+60 度では 2D 条件・3D 条件のどちらにおいても角度誤差の平均はマイナスの値となり、2D 条件に比べて 3D 条件の方が角度誤差の平均の絶対値が有意に小さかった(+50 度, +60 度ともに $p=0.001$ (†)). その他の方位については、2D 条件・3D 条件間での有意差は確認されなかった。

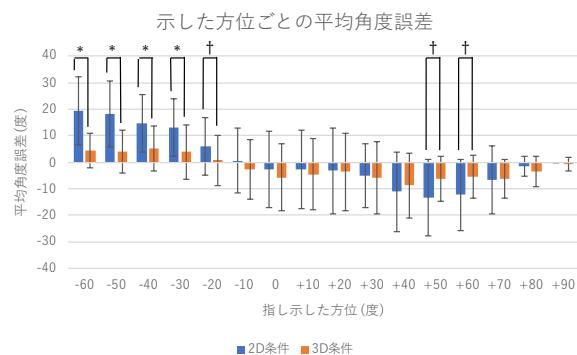


図 11. 線分が示した方位別の角度誤差の平均

(* $p<.001$, † $p<.01$)

5.2 考察

実験結果より、2D 映像条件よりも 3D 映像条件の方が、被験者の方向認識精度が向上することが分かった。特に、被験者から見て奥側を指した際に 3D 映像の優位性が高かった。この結果より、空中立体映像は、観察者から見て奥側の方向提示を伝達する際に優れるといえる。

ただし、被験者の目の見え方の個人差によって認識精度に影響が出た可能性が考えられる。図 9 より、被験者 5・8・11・13・18~20 は 2D 条件と 3D 条件間で平均角度誤差に有意差が出なかった。実験では被験者の視力に関するアンケートも実施したが、被験者 8 は極端に左右の視力差があることを申告している。また、被験者 11・13 は「普段、立体視に関して何か問題を感じたことはあるか」という設問に「はい」と回答しており、被験者 5・18・19 は「日常生活において、物体の距離や深さを見極めることに困難を感じることはあるか」という質問に「はい」と答えている。しかし、後者の質問には 2D 条件と 3D 条件間で平均角度誤差に有意差が出た被験者のうち数名も「はい」と回答しており、被験者 20 は視力に関する申告事項がなかった。よって、被験者ごとの見え方の癖が結果に少なからず影響を与えたと考えられる。

また、2D 映像条件と 3D 映像条件で線分が示した

方位別に被験者が答えた角度誤差の平均を比較したところ(図 11)，被験者から見て手前側(プラス方向)を線分が示した際には両条件間で有意差が出ない場合が多くかった。この原因として、AIRR テーブルに組み込まれた立体ディスプレイである Looking Glass の仕様が関係していると考えられる。Looking Glass はディスプレイの輻輳面より手前側に飛び出している物体を立体的に表示することに向いてあるという性質を持つ。よって、線分が被験者から見て手前側を指し示した際は 3D 映像の優位性が十分に保たれなかつたと考えられる。また、線分が真横に近い方向を示した場合、示された方向を判断する際に立体的情報が手掛かりになることは少ないため、2D 条件と 3D 条件で角度誤差に大きな違いが出なかつたと思われる。

6 結論

我々は、机上にキャラクタを投影し、指差しによる方向提示により机上の実物体の指し示しを行うようなシステムの構築において、指示方向の認識精度向上を目指してきた。今回、キャラクタを空中立体映像として机上に投影した際の方向提示を認識する精度の上限を検証するため、キャラクタの代わりに線分を表示して実験を行った。その結果、立体映像が平面映像に比べて有意に指示方向の認識精度を向上させることを明らかにした。特に、被験者から見て奥側の方向を指し示す場合において、立体映像はより優れることが分かった。平均 7 度の角度誤差は、例えば 30cm 先を指す場合、3.7cm 程度の誤差を見込むことになる。2D 条件での最悪値である 20 度の場合、誤差は最大 11cm となり、机上で方向指示精度としては十分とはいがたい。指し示す方角によって誤差範囲が異なることも考慮せねばならず、平面映像では操作者の負荷を高める恐れがある。

ただし、この結果はあくまでも精度の上限を示すものである。これまでの著者らの経験から、実際に人間型キャラクタを用いて方向指示をした場合、投影手法の問題により映像の解像度が低下するため、その指し示す方向が明確にならない傾向がある。そのため、方向指示の認識精度を損ねないようなキャラクタのモデリングやカラーリング、モーションなどを検討する必要がある。また今後は、顔や身体の向きと組み合わせた、より複雑な方向指示における、空中立体映像の優位性についても検証を進めたいと考えている。

参考文献

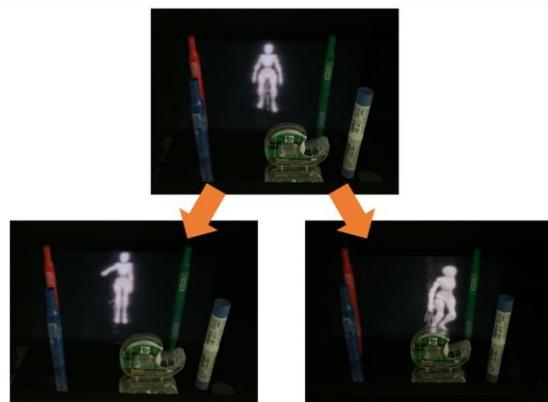
- [1] Keita Higuchi, Ryo Yonetani, Yoichi Sato, “Can Eye

- Help You?: Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks,” CHI ’16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems(2016)
- [2] Shinsaku Hiura, Kenji Tojo, and Seiji Inokuchi, “3-D Tele-direction Interface using Video Projectors”
- [3] Andrew J. Davison, Walterio W. Mayol, and David W. Murray, “Real-Time Localisation and Mapping with Wearable Active Vision,” In Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality (2003)
- [4] Taehee Lee and Tobias Hollerer, “Viewpoint Stabilization for Live Collaborative Video Augmentations,” IEEE/ACM International Symposium on Mixed and Augmented Reality(2006)
- [5] Jen-Shuo Liu, Barbara Tversky, Steven Feiner, “Precueing Object Placement and Orientation for Manual Tasks in Augmented Reality,” IEEE Transactions on Visualization and Computer Graphics(2022)
- [6] Santawat Thanyadit, Matthias Heintz, Effie L-C Law, “Tutor In-sight: Guiding and Visualizing Students’ Attention with Mixed Reality Avatar Presentation Tools,” CHI ’23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems(2023)
- [7] Kibun Kim, John Bolton, Audrey Girouard, Jeremy Cooperstock, Roel Vertegaal, “TeleHuman: effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod,” CHI ’12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(2012)
- [8] Hideaki Kuzuoka, Shinya Oyama, Keiichi Yamazaki, Kenji Suzuki, Mamoru Mitsuishi, “GestureMan: a mobile robot that embodies a remote instructor’s actions,” CSCW ’00: Proceedings of the 2000 ACM conference on Computer supported cooperative work (2000)
- [9] 大槻麻衣, 丸山啓太, 葛岡英明, 鈴木雄介, “眼球型ディスプレイによる立体的な視線提示がモバイル端末を介した物理的遠隔協調対話における注視誘導における効果,” ヒューマンインターフェース学会論文誌, Vol. 22, No. 1 (2020)
- [10] 鈴木祐也, 葛岡英明, 山下淳, 山崎敬一, 山崎晶子, 久野義徳, “無言のロボットによる注意誘導の研究,” ヒューマンインターフェースシンポジウム論文集 (2007)
- [11] Miho Izuka, Kentaro Fukuchi, “FlexFace: A Head Gesture Motion Display with Flexible Screen for Telecommunication,” ISS ’18: Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces(2018)
- [12] 山本裕紹, “再帰反射による空中結像 (AIRR) による空中ディスプレイ,” 日本画像学会誌, Vol. 56, No. 4 (2017)
- [13] 福地健太郎, 松浦向日葵, 大野貴広, “テーブル上に投影された空中立体映像に対する影を用いたインタラクション手法,” 情報処理学会インタラクション 2022(2022)

未来ビジョン

空中立体映像を活用した実世界融合型アプリケーション

我々は、図1に示した空中立体映像投影システムを用いて、人間型キャラクタが複数のユーザと対話しながら机上の実物体について説明するような実世界融合型アプリケーションの開発を目指している。アプリケーションでは右図のように、キャラクタが、周囲にある複数の実物体のうち1つを指差したり近づいたりすることを想定している。今後の実装としては、机上に美術品を置いてキャラクタが案内役となってそれらについて説明したり、机上にある商品をキャラクタが宣伝したりするようなものを構想している。



Whisphone: ささやき声で入力できるイヤホン

福本 雅朗*

概要. Whisphone は、ささやき声での音声入力が可能なイヤホンである。先端部にマイクを搭載したカナル(耳穴插入)型イヤホンによる外耳道閉鎖効果によって、骨導経由で外耳道内部に放射されたささやき声を効率的に収音できる。耳穴を塞ぐことによる外部騒音の遮断に加え、イヤホンの ANC (Active Noise Canceling) 機能の併用によって、80dB(A)の騒音下においても微小なささやき声の収録と音声認識が可能である。小型の機器は目立ちにくく常時装用が容易であり、小さなささやき声は入力に際して周囲の迷惑になりにくい。本装置を用いることで、オフィス・家庭・街頭等、日常生活の多くの場面においてハンズフリーで AI アシスタントとの対話を行うことができる。

1 はじめに

LLM (Large Language Model) と生成 AI は、コンピュータを「操作するもの」から、「対話するもの」に変えつつある。中でも音声による会話はその中心になっていくだろう。しかしながら、音声入力には「音漏れ(周囲への迷惑・情報漏洩)」という大きな問題が残っている。狭いオフィスや家庭、あるいは混雑する街頭で皆が大声で喋りまくる未来はぞつしない。静かで心地よい世界を実現するには、「音漏れない」音声入手法の開発が必須と言える。

1.1 サイレント音声入力

サイレントな音声入力には、全く音を出さないもの(以下完全サイレント)と、周囲の迷惑にならない程度の微弱の音を出すものの 2 種類がある。

完全サイレントな手法の多くは、発話に伴う口腔周辺のパツツの動きを各種センサで捉えるものである。Sottovose[13]は、下顎に当てた超音波センサを用いて顎・唇・舌などの動きを検出し、音声に変換する。顎や下顎に設置した加速度センサ[18]や EMG (Electromyogram, 筋電) 電極[28]を用いるものもある。このほか、マスク内部に設置した加速度センサ[6]や歪センサ[9]を用いるものや、イヤホンから外耳道内に放射した音波の反射を用いるもの[3]もある。また、頭部に設置した複数の電極で発話をイメージした際の ECG (Electro-encephalogram, 脳波) を検出するもの[4]に至っては、もはや口すら動かす必要が無い。

いずれの手法でも、得られたセンサの信号を最終的な文字や音声に変換するには、専用に設計された

認識機構が必要であり、多くの場合は話者個々人に対する追加学習を必要とする。また、認識できる音素や単語数が限られているものが多く、リアルタイム処理が難しいものもある。

1.2 ささやき声による音声入力

これに対し、ささやき声などの微弱な音を検出する手法は、得られる信号が本来の発話音声に似ていることから、認識機構の設計が前者に比べて容易であり、認識率や認識速度等の性能も上げやすい。

この分野の先駆的な研究としては NAM (Non-Audible Murmur, 非可聴つぶやき)[12]がある。耳介後下方部に接触させた「肉伝導マイク」で微小なつぶやき音声を収録することで、ほぼ音漏れの無い音声入力を可能としている。SilentVoice[11]は、従来の発話とは逆に「息を吸いながら話す」ことで、ポップノイズ無しにマイクを極限まで近づけられ、40dB(A)未満の超微弱音声の検出が可能である。WhisperMask[5]は、マスク内側に設置した面状の ECM (Electret Condenser Microphone) の持つ優れた close talk 性能により、高騒音環境においても小さい発話音声をクリアに検出することができる。多くは取得したささやき声を文字として認識するが、WESPER[17]のように、リアルタイムに通常音声に変換するものもあり、コンピュータとの対話だけで無く、人間同士の会話にも使用可能である。

一部の音声認識機構 (Google Voice Input, OpenAI Whisper, Amazon Alexa の "Whisper mode" 等) はささやき声の認識に対応しており、周囲に迷惑をかけずに入力を行なうことができる。また、通常音声とささやき声を識別することで、ささやき声をコマンドとして用いることもできる[16]。

しかし、本来小音量であるささやき声を良好に収録するには、静かな場所で使用する、スマートホンやスマートスピーカ等の機器に口を近づける、ある

Copyright is held by the author(s).

* Microsoft Corporation

(English Title)

Whisphone: Whispering Input Earbuds

いはヘッドセット等を用いてマイクを口元に近づける必要がある。また、イヤホン型ヘッドセットの多くはビームフォーミング等を用いて口元の音声を増幅している。しかし、たとえこれらの機器を使用したとしても、騒々しいオフィスや街頭での使用は難しく、ヘッドセットマイクを装着したまま日常生活を送ることは容易では無い(食事等の際に邪魔になるほか、「見た目」の問題から一般大衆への浸透は簡単では無い)。

1.3 骨導マイクロホン

上述のささやき声入力機構の多くは、発話に伴って周囲大気中に放射された音(気導音)を検出している。これに対し、発話に伴う振動が頭蓋や軟組織を経てきたものを、皮膚表面に設置したセンサで検出するもの(骨導式、上述のNAMもこの一種)がある。骨導式の収音機構としては、古くから使われている咽頭マイク[22](喉部に接触させる)のほか、外耳道壁の振動を検出するもの[21]、小型のECMマイクで外耳道内放射音を収録するもの[20]、イヤホンに設置したMEMS振動センサによって耳穴周辺の振動を検出するもの[19]等がある。また、市販されているノイズキャンセルイヤホンの中にも、筐体内部に設置した加速度センサを用いて発話時の振動を検出し、音声の明瞭度向上を図っているもの[7]がある。骨導式は外部騒音に強い反面高周波成分の減衰が強く、声が籠って聞こえてしまうという問題があり、特殊状況下での使用に留まっている。またNAMを除き、通常発話の収録を目的としており、微小なささやき声を十分なレベルで検出できるものは少ない。

1.4 骨導によるささやき声収録

一般的に知られている骨導聴取の経路は、頭蓋や軟組織を経て伝わった音響振動が、直接耳小骨や蝸牛に到達して音として知覚されるというものである。一方でこれらの音響振動が、外耳道壁を震わせて外耳道内部に一旦「音として」放射され(外耳道内放射)，それが(通常の気導音と同様に)鼓膜を伝わって聴取されるという経路もあり、特に健聴者においては、こちらの寄与も大きいとされている[25]。また、指や耳栓などで外耳道開口部を塞いだ際に、骨導聴取において1kHz以下の低周波領域が5-20dB強調されて聞こえる現象(外耳道閉鎖効果)が知られており、これは後者の経路によって外耳道内に放射された音響エネルギーが、外耳道開口部から逃げなくなる為だと考えられている[8]。

この効果は自身が発した音声(聴覚フィードバック)に対しても有効なので、外耳道内部にマイクロホンを設置し、併せて外耳道開口部を塞ぐことで、ささやき声などの微弱な発話を増幅された状態で収録できると考えられる。また、これは「耳を塞いだ」

状態に相当するので、マイクロホンへの外部騒音の侵入も同時に阻止でき都合が良い。

しかしながら、ささやき声の発話音量は40dB(A)程度であり、60dB(A)である通常発話に比べてかなり低い。仮に外耳道閉鎖による骨導成分増幅や外部騒音低減があったとしても、高騒音下で良好なS/N比を得るのは簡単では無い。

一方、近年のTWS(True Wireless Stereo)イヤホンの普及に伴い、ANC(Active Noise Canceling)技術の発展が著しく、特にカナル型のイヤホンでは、公称値で45dBを超えるノイズ抑圧性能を謳うものが多くある[2]。カナル型イヤホンを装着した状態が「外耳道閉鎖」と同じであることを考えると、ANC機能を併用することで、同様のノイズキャンセル効果が外耳道内部に設置したマイクロホンに対しても得られることになる。

本稿では、常時装用型ささやき音声入力デバイスであるWhisphoneの提案を行う。ANCを動作させたカナル型イヤホンの耳栓先端部分に設置したマイクロホンで外耳道壁から放射される骨導音声の収録を行うことで、たとえ騒音下であっても、周囲の迷惑にならないような微小なささやき声を安定して取得できる。以下の章では、外耳道閉鎖による骨導音声の増幅と、耳栓とANCの併用による外部騒音低減効果について実験結果と共に説明する。次いで、汎用の音声認識機構と組み合わせることで、追加学習や個人適応を行うこと無く、騒音下であってもささやき声の音声認識が可能であることを示す。いくつかの実装例を紹介すると共に現状の課題や解決策を示し、将来への展望を述べて結言とする。

2 Whisphone

2.1 構成

Whisphoneの構成を図1に示す。ANC機能を持つカナル型イヤホンの耳栓先端部に設置したマイクロホンによって、ささやき声発話時に外耳道内に放射される骨導音声を収録する。カナル型イヤホンによって外耳道開口部が密閉されるので、外部騒音の低減に加え、外耳道閉鎖効果による音声信号の増幅が期待できる。ANC機能による更なる外部騒音低減

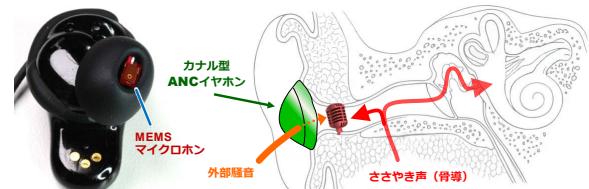


図1. カナル型ANCイヤホンの耳栓先端部分にマイクを設置し、ささやき声の外耳道内放射音を収録する。

Whisphone: ささやき声で入力できるイヤホン

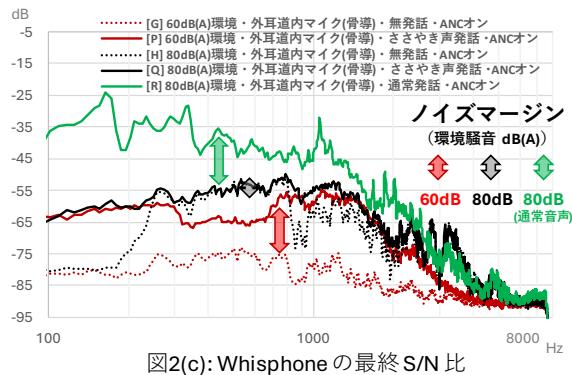
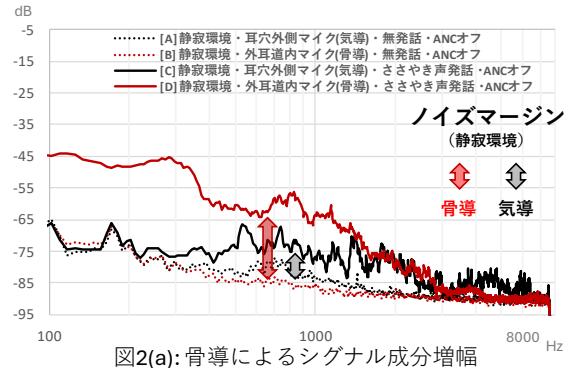
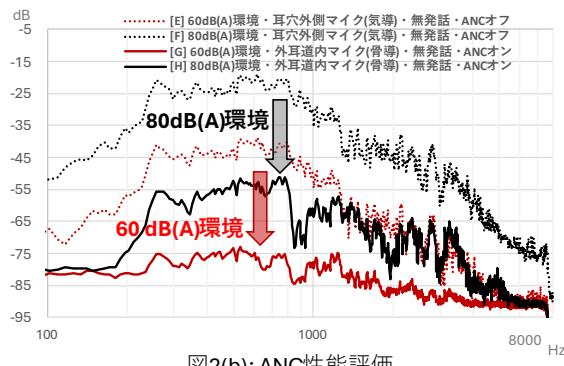
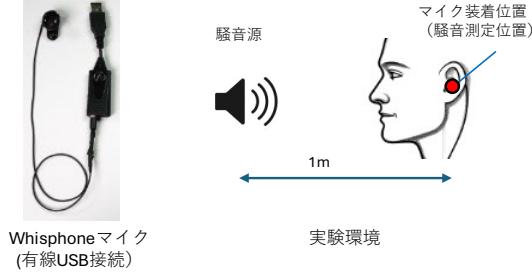


図 2. Whisphone の収音性能 (a): 外耳道閉鎖効果によって、ささやき声を 10dB 程度増幅して収音できる。 (b): ANC によって外部騒音を 30dB 程度低減できる。 (c): 総合的な S/N 比改善効果は 40dB 程度になり、騒音下でも微弱なささやき声の収録が可能になる。

により、騒音下においてもささやき声による音声収録が可能である。

2.2 収音性能

最初に、Whisphone の収音性能の確認を行った。ANC 機能を持つカナル型イヤホン[2]の耳栓先端部に小型の MEMS (Micro Electro Mechanical Systems) マイク[10]を設置し、USB オーディオインターフェース[1](ローカットフィルタを 235Hz に変更)を介して PC に接続した。イヤホンを外耳道に装着した状態でささやき声による入力をを行い、音声波形を記録した。ささやき声の発話音量は、1m 距離で標準的な 40dB(A)となるように調整した。騒音源には、電車のホームで収録した音源を用い、耳元で 60dB(A)もしくは 80dB(A)になるように調整している。なお、実験環境の暗騒音は 33dB(A)であった(以下、静寂環境と表示)。ANC 機能は同イヤホンの“Quiet”モード(最大強度)を用いている。

実験は、被験者一名(成人男性)に対して行った。発話音声としては『こんにちは、私の声が聞こえますか?』を用いた。外部騒音レベル(静寂環境 /60dB(A)/80dB(A))・マイク設置位置(外耳道内/耳穴外側)・発話/無発話・ANC 機能の On/Off を変更しつつ収録を行った。

収録結果を図 2 に示す。図 2(a) は、外耳道内マイ

ク及び外耳道閉鎖による音量増幅の確認である。静寂環境におけるバックグラウンドノイズ [A] に対し、通常のイヤホンで使われるような耳穴付近に設置したマイクでの気導音の収音 [C] ではノイズマージンは 10dB 程度であり、微弱なささやき声を十分な音量で取得することは難しい。一方で、マイクを外耳道内に設置し、外耳道開口部を塞いだ状態 [D] では、1.5kHz 以下の領域で、同じささやき声を 20dB 程度のノイズマージンで取得することができている。これは、耳栓による外部ノイズレベルの低下 [B] に加え、外耳道閉鎖効果による外耳道内放射成分の増強によるものと考えられる。一方、骨導収音においては 2kHz 以上の成分は急激に減少し、4.5kHz 以上ではほぼ収録されていない。なお、同図の [D] において、500Hz 以下の成分が急激に増加している。これは主に発話に伴う外耳道の変形によるものであるが、現在は積極的に使用していない(詳細後述)。

図 2(b) は、ANC 機能による外部騒音低減の確認である。60dB(A)及び 80dB(A)の外部騒音に対しては、耳栓と ANC 機能を併用することで、30dB 程度のノイズ低減効果が得られている ([E]vs[G], [F]vs[H])。

図 2(c) は、Whisphone の収音性能の確認である。外耳道マイクと ANC 機能の併用によって、60dB(A)

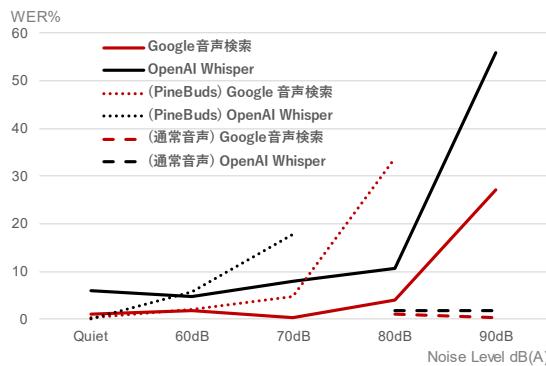


図 3. Whisphone を用いた音声認識の性能評価

の外部騒音下においては、ささやき声による発話を最大 20dB 程度のノイズマージンを保って取得することができている([P]vs[G])。ささやき声の音量が 40dB(A)であることを考えると、Whisphone による総合的な S/N 比改善効果は、通常の気導マイクに比べて最大 40dB 程度(外耳道閉鎖効果によるシグナル成分増強が+10dB, ANC によるノイズ成分低減が-30dB)であることがわかる。また、外部騒音 80dB(A)でも同様に 40dB 程度の S/N 比改善効果を得られているが、ノイズマージンはほぼ無くなってしまう ([Q]vs[H])為、このあたりが使用限界と考えられる。

外部騒音レベルがこれ以上になると、ささやき声による音声認識は困難になってしまい、通常発話(60dB(A)程度の音量)を用いた場合は更に 20dB 程度のノイズマージンがある為、90dB(A)を超える外部騒音下であっても音声の取得が可能([R]vs[H])である(外部騒音が 80dB(A)以上の場合、60dB(A)程度の通常発話は周囲の人からはほぼ聞こえなくなる)。

2.3 音声認識性能

次いで簡易的な音声認識性能の評価を行った。使用した機材やノイズは前節で用いたものと同一である。音声認識機構には、Google 検索ページの「音声検索」(以下 Google 音声検索)及び OpenAI の Whisper (large-v3)を用いており、追加学習やエンロールなどの個人適応は一切行っていない。テストには、音声アシスタントや生成 AI を用いた操作で良く使われるような語句(『明日の予定は何がありますか』・『東京駅はどう行けば良いですか?』・『新しいアイデアを出してください』等)を 15 個選んでいる。ささやき声による発話は各 3 回行い、WER (Word Error Rate, 単語エラー率) を算出した。実験環境における暗騒音は 33dB(A)であった。なお被験者は成人男性 1 名である。

図 3 に結果を示す。ノイズレベルが 80dB(A)程度(地上を走る電車の車内等)までは WER が概ね 10% 以下であり、小さなささやき声でもほぼ正確な入力

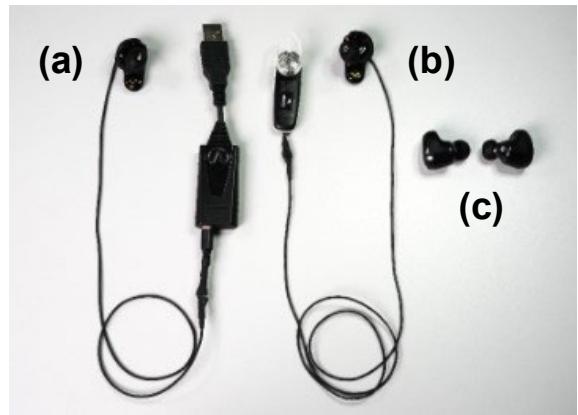


図 4. 実装例 (a): 有線 USB マイクロホンタイプ. (b): Bluetooth ヘッドセットタイプ(モノラル). (c): Bluetooth TWS イヤホンタイプ.

が行えている。一方、90dB(A)(地下鉄の車内等)になると WER が大きく上昇すると共に、VAD (Voice Activity Detection, 発話区間検出) の失敗も頻発する為実用的とは言えない。

のことから、Whisphone を用いた音声認識は、外部騒音が 80dB(A)までは特に追加学習や個人適応処理を行うことなく使用可能であると考えられる。なお、外部騒音が 80dB(A)を超える場合であっても、通常発話を用いた場合には引き続き低 WER での音声認識が可能である(図 3 の破線部参照)。

3 実装

次に Whisphone の実装例を示す。第二章で示したように、Whisphone の騒音抑圧性能を上げる鍵は ANC にある。幸いなことに、近年では公称値で 45dB を超えるノイズ抑圧性能を持つ TWS イヤホンが入手できるようになっており、小型マイクをこれらと組み合わせるだけで簡易的な実装が可能である。図 4(a)は小型の MEMS マイク[10]を市販の TWS イヤホン[2](左耳側)と組み合わせた例である。マイクカプセルの信号線は、外耳道の密閉性を損なわないよう極細ケーブルを用いて外側に取り出し、USB オーディオ I/F[1]に接続している(一般的な有線マイクロホンとして動作)。使用時には、TWS イヤホンのノイズキャンセル機能のみをオフラインで動作させることで、ノイズ抑圧された骨導音を取得することができる。

図 4(b)は、市販の Bluetooth ヘッドセット[15]の ECM カプセルを取り外し、図 4(a)の小型マイクを接続したものである。使用時には、同じく左耳側はオフラインで ANC 機能のみ動作させ、右耳側のみペアリングして(モノラルの)Bluetooth ヘッドセットとして使用する。

図 4(c)は、一体型の Bluetooth TWS イヤホンと

Whisphone: ささやき声で入力できるイヤホン

しての実装例である。市販の ANC イヤホンの中には、小型のマイクカプセルが耳栓先端部に設置されているものがある。これは ANC の FeedBack マイクとして動作させる為であるが、同じマイクを音声収録用として使えば Whisphone を実現することができる(一部の TWS イヤホンでは、通話時にこのマイクを用いて音声のピックアップを行っている[7]。但し、音声収録のメインはあくまでもイヤホンの外側に設置した複数マイクによるビームフォーミングであり、内部マイクや筐体内部の振動センサは、あくまで明瞭度を上げる為の補助的な使用に留まっている)。つまり、このタイプの ANC イヤホンでは、音声収音用マイクとして、イヤホン外側のマイクの代わりに FeedBack 用マイクを使用するようにファームウェアを書き換えるだけで Whisphone を実現できる。ここでは、ファームウェアがオープンソースで公開されている Pinebuds Pro[14]を用いて実装を行った。ANC の性能が劣る(実測値で 10dB 程度)為に騒音下での WER は悪化するが、70dB(A)程度までは比較的良好な WER が実現できている(図 3 の点線部参照)。

4 議論

• エコーチャンセル

イヤホン一体型の機器として用いる場合、イヤホンから再生された音楽や音声が先端のマイクで収音され、ノイズとなってしまう。音声認識時の障害となるほか、リアルタイム通話の場合はエコーとなって相手方に帰ってしまう。イヤホンからの再生信号は既知なので、スピーカーホン等で使われているエコーチャンセル技術を用いて対処可能である。

簡易的な手法としては、左右双方に装着している場合に、音声認識や通話時には片方をマイク、他方をイヤホンとして用いることでもノイズやエコーを防止できる(図 4(c)の TWS イヤホンタイプで実装)。

• ダブル NR

外部騒音に音声が含まれている場合、たとえ低減後であっても音声認識機構が反応してしまうことがある(多くの音声認識機構側に自動ゲイン調整機構がある為)。これを防ぐには、イヤホン外部に別のマイクを設置(ANC イヤホンの FeedForward マイクに相当)して内外のマイクの音量差を比較し、内部マイクの音が大きい成分のみ通過させれば良い(耳栓がある為、外部騒音の音量は外側のマイクが常に大きくなる)。

• 片耳/両耳

Whisphone は片耳/両耳のどちらでも使うことができる。片耳装用の場合、もう一方の外耳道の状態(解放/耳栓/ANC)を変化させても、収録された音声や

騒音のレベルに違いは見られなかった。装着の負担が少なく外音聴取も可能であり、音楽鑑賞を行わないのであれば有効な選択肢と言える。

両耳装用の場合、高騒音下では困難になる自身のささやき声の発話状態を(骨伝導による聴覚フィードバックを介して)確認でき、発話音量や滑舌の制御が容易になる。なお、両側に設置したマイクの出力を合成すれば更なる S/N 比改善の可能性もある(理論上最大 1.4 倍)。但し、タイミング調整が難しい(特に TWS の場合)ことに加え、左右音声成分の相關低下(外耳道や頭蓋の形状は完全な左右対称では無い)、騒音成分の相關上昇(骨導によるミックス)がある為、効果は限定的だと考えられる。

• 認識率向上

Whisphone で収録された骨導 ささやき声は、2kHz 以上の高周波成分が著しく減少している。特に高周波成分の多い音素(子音の /s/ や /h/、母音の /i/)の聞こえ方が通常の ささやき声とは異なり、認識率低下の一因と考えられる。骨導 ささやき声による認識機構の再トレーニングが最も有効ではあるが、大量の音声データを集めるのは容易ではない。既に通常 ささやき声のデータがある場合、フィルタによって 2kHz 以上の成分をカットしたもので追加トレーニングを行うことで、新たなデータ収集を行わずとも認識率向上が可能だと考えられる。なお、ささやき声では有聲音が無聲音化する(例えば /g/ → /k/)が、多くの音声認識機構では ささやき声の大量学習や単語辞書を用いて認識率低下を防いでいる。

• リアルタイム音声変換

本稿では主に骨導 ささやき声の認識に焦点を当ててきた。一方、収録された骨導 ささやき声を、通常 ささやき声や通常音声にリアルタイム変換すれば、音声通話等にも利用することができる。前者には欠落している高周波成分の復元が必要であり、骨導音声から気導音声への変換技術[26]が利用できる。後者では更に声帯振動成分やピッチの復元が必要であり、ささやき音声から通常音声への変換技術[17]が利用できるだろう。

• 発話による外耳道変形と入力への応用

図 5(a)上側は、静寂環境における ささやき声発話時(『こんにちは、私の声が聞こえますか?』)の波形例、同下側は、同じ文章のサイレントスピーチ時(声を出さずに同じ口の動きをする)の波形例である。両者で同様のスパイク状の低周波信号が確認できる。これは、発話時の口や頸の動作に伴う外耳道の変形によって生じた圧力変化だと考えられる。これらの成分は、音声波形に比べて過大であり、信号クリップによる歪の原因となる為、現在はフィルタで取り除いており、積極的な利用は行っていない。但し、ここには発話情報が含まれていると考えられ、これ

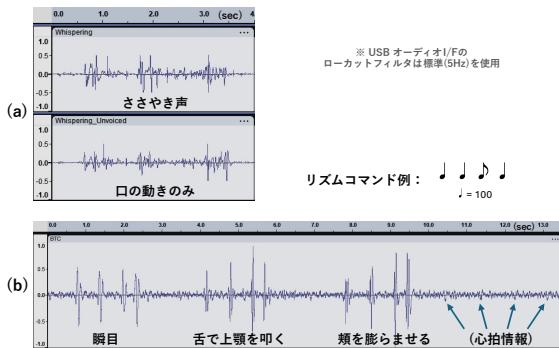


図 5. 音声信号以外の利用可能性 (a): 低周波成分には発話時の口や頬の運動情報が含まれている. (b): 瞬目リズム等をコマンドとして使用できる.

らの成分を含めて学習させることで、認識性能を高められる可能性がある。更に、この成分「のみ」を用いれば、サイレントスピーチの認識が可能となるかもしれない。

また、図 5(b)は、瞬目等を $\downarrow \downarrow \uparrow \downarrow$ のリズムで行った場合の波形であり、外耳道内の圧力変化に動作情報が含まれていることがわかる。これらを利用すればハンズフリーでの操作が可能になる(特に舌の動きを用いたものは秘匿性が高い)。簡単なリズムを用いたコマンド[27]であれば、コンパレータやタイマで実現可能であり、特に低消費電力での常時動作が要求される wakeup 操作に適している上、画像や音声の認識に比べてプライバシー問題への対処がしやすい(図 4(c)の TWS イヤホンタイプで AI アシスタントの呼び出しコマンドとして実装)。なお、微弱ながら心拍情報も含まれているので、簡易的な心拍センサとしても利用できるだろう。

• 発声手法

通常発話の主たる音源が声帯振動であるのに対し、ささやき声のそれは、声道中の狭窄部位を空気が通過する際の乱流音である。狭窄部位の生成手法は個人や状況によっても異なるが、声帯を狭める(「ハスキーボイス」の音源としても使われる)・上顎に舌を近づける・唇を狭める(いわゆる「ひそひそ声」)等があり、声質にも違いがある。声帯を狭める手法は大きな乱流を発生させることができ、発話可能な音素も多い(ほぼ全ての無声音)。一方で高音域が弱く明瞭度が下がるのに加え、声帯への負担も大きい。これに対し、口先で喋る手法は音量を絞りやすく声帯の負担にはなりにくいが、発話可能な音素が限られる(特に鼻音 /n/ が難しい)。上顎を使う手法は両者の中間である。認識率の更なる向上には、それぞれの発声手法による追加学習が必要だと考えられる。

なお、ささやき声は SilentVoice[11]で述べられている吸気発話(Ingressive Speech)を用いても発声可能である。音量はほぼ同じである為秘匿性向上に

は繋がらないが、呼気発話との弁別が可能になれば、人間向け(呼気発話)とネット向け(吸気発話)の切り替えに利用できるかもしれない。

• オートベント

日常生活で密閉度の高いカナル型イヤホンを使用する場合、周囲とのコミュニケーションや危険把握の為に適切な外音の取り込みが必要となる。現在は主に「外音取り込みモード」(外のマイクで取得した音の一部をイヤホンから再生する)が用いられているが、消費電力増大が課題である。また、密閉度を上げると自分が発話した音声の聴覚フィードバックが過大となる・心拍音が聞こえる等の問題も出てくる。通常はベント穴を設けることでこれらを軽減しているが、一方でノイズ抑圧性能の低下を招いてしまう。Whisphone を常時装用する場合には、これらの問題への対応が求められる。例えば、圧電素子を用いた(Normally Open タイプの)アクティブベント機構[23]を用いれば、電源オフ時にはベント穴を開放して(電気的な再構成では無い)外音をそのまま通過させ、動作時には密閉してノイズを遮断することができるだろう。

• 他言語への対応

今回用いた Google 音声検索や OpenAI Whisper は多言語でのささやき声認識に対応している。2章で行った実験と同様のセンテンスに対して、英語における WER の計測を行った結果を示す。静寂時での通常マイクによるささやき声認識の WER がそれぞれ 2.4%・0.0% なのに対し、Whisphone 環境では 55.0%・56.2% と極端に悪くなつた。英語は子音(1kHz 以上の高周波成分が支配的)が多い言語であり、骨伝導による高周波成分減少との相性が悪い。このほか中国語(普通話)は声調(tone)ベースの言語であり、tone 成分がほぼ存在しないささやき声との相性が悪い。その意味では Whisphone は、母音の寄与率が高く tone の変化が少ない日本語(及び韓国語)での使用に適していると考えられる。

5 おわりに

本稿では、常時装用型 ささやき音声入力デバイスである Whisphone の提案を行つた。まずは社会的適合性(周囲への迷惑を気にする)と言語的メリット(骨導 ささやき声入力に向いている)が多い日本での普及を足掛かりに、いずれ他の地域・言語にも広げて行きたい。本稿では主に音声入力を対象としたが、ささやき声を用いたリアルタイム会話にも、周囲への迷惑低減や情報漏洩防止等のメリットがある為、将来 ささやき声での会話が日常的に使われるようになるかもしれない。

静かな未来がやってくるのだ。

参考文献

- [1] Andrea Communications USB-MA (2024/10/27 確認) <https://andreacommunications.com/products/usb-ma-premium-external-usb-microphone-adapter/>
- [2] BOSE QuietComfort Ultra Earbuds (2024/10/27 確認) <https://www.bose.com/p/earbuds/bose-quiet-comfort-ultra-earbuds/QCUE-HEADPHONEIN.html>
- [3] DONG, Xuefu, et al. ReHEarSSE: Recognizing Hidden-in-the-Ear Silently Spelled Expressions. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems. 2024.* pp. 1-16.(2024)
- [4] FITRIAH, Nilam; ZAKARIA, Hasballah; RAJAB, Tati Latifah Erawati. EEG-Based Silent Speech Interface and its Challenges: A Survey. *International Journal of Advanced Computer Science and Applications*, 2022, 13.11. (2022)
- [5] HIRAKI, Hirotaka, et al. WhisperMask: a noise suppressive mask-type microphone for whisper speech. In: *Proceedings of the Augmented Humans International Conference 2024.* pp.1-14. (2024)
- [6] HIRAKI, Hirotaka; REKIMOTO, Jun. SilentMask: mask-type silent speech interface with measurement of mouth movement. In: *Proceedings of the Augmented Humans International Conference 2021.* pp. 86-90. (2021)
- [7] HUAWEI FreeBuds Pro 3 (2024/10/27 確認) <https://consumer.huawei.com/jp/audio/freebuds-pro-3/>
- [8] K.Ito; S.Nakagawa. Bone-conducted ultrasonic hearing assessed by tympanic membrane vibration in living human beings, *Acoust, Sci. & Tech*, Vol. 34(6), pp. 413-423, (2013)
- [9] KUNIMI, Yusuke, et al. E-MASK: a mask-shaped interface for silent speech interaction with flexible strain sensors. In: *Proceedings of the Augmented Humans International Conference 2022.* pp. 26-34. (2022)
- [10] LinkMems LMA2718T381-OY1 (2024/06/18 確認) https://www.lcsc.com/product-detail/MEMS-Microphones_LinkMems-LMA2718T381-OY1_C5455565.html
- [11] Masaaki Fukumoto. SilentVoice: Unnoticeable voice input by ingressive speech. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, UIST '18*, pp. 237-246. (2018)
- [12] NAKAJIMA, Yoshitaka, et al. Non-audible murmur (NAM) recognition. *IEICE TRANS. on Information and Systems*, 89.1: pp. 1-8. (2006)
- [13] Naoki Kimura, Michinari Kono, and Jun Rekimoto. Sottovoce: An ultrasound imaging-based silent speech interaction using deep neural networks. In *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI '19*, pp. 1-11. (2019)
- [14] PINE64 PineBuds Pro (2024/10/27 確認) https://pine64.org/devices/pinebuds_pro/
- [15] Plantronics M70 (2024/10/27 確認) <https://web.archive.org/web/20200416222955/http://www.plantronics.com/us/en/product/m70>
- [16] REKIMOTO, Jun. DualVoice: A speech interaction method using whisper-voice as commands. In *Proceedings of the CHI Conference on Human Factors in Computing Systems Extended Abstracts. 2022.* pp. 1-6. (2022)
- [17] REKIMOTO, Jun. WESPER: Zero-shot and realtime whisper to normal voice conversion for whisper-based speech interactions. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems. 2023.* pp. 1-12. (2023)
- [18] REKIMOTO, Jun; NISHIMURA, Yu. Derma: silent speech interaction using transcutaneous motion sensing. In: *Proceedings of the Augmented Humans International Conference 2021.* pp. 91-100. (2021)
- [19] SCHILK, Philipp, et al. In-Ear-Voice: Towards Milli-Watt Audio Enhancement With Bone-Conduction Microphones for In-Ear Sensing Platforms. In: *Proceedings of the 8th ACM/IEEE Conference on Internet of Things Design and Implementation. 2023.* pp. 1-12. (2023)
- [20] SONY ECM-TL1 (2024/10/27 確認) <https://web.archive.org/web/20110325062731/https://www.sony.jp/microphone/products/ECM-TL1>
- [21] TEMCO JAPAN EM20N-T3.5P (2024/10/27 確認) <https://www.temco-j.co.jp/products/em20n-t3-5p/>
- [22] TEMCO JAPAN TM80N-T (2024/10/27 確認) <https://www.temco-j.co.jp/products/tm80n-t/>
- [23] xMEMS SKYLINE (2024/10/27 確認) <https://xmems.com/acousticvent/>
- [24] 安藤毅; 田野俊一; 市野順子; 橋山智訓. 痕跡器官の末梢系生体情報を用いた常時装着型入力インターフェース. HI シンポジウム 2008, pp.229-234. (2008)
- [25] 伊藤一仁. 骨導聴覚の知覚機序に関する実験的考察. 純真学園大学雑誌. Vol.11, pp.67-72, (2021)
- [26] 川本大貴; 中山仁史. 疊み込みニューラルネットワークを用いた光骨伝導音の音質改善検討. 日本機械学会論文集, 2024, 23-00304, (2024)

- [27] 福本雅朗; 外村佳伸. “指鉗”:手首装着型コマンド入力機構. 情報処理学会論文誌, Vol.40, No.2, pp.389-398, (1999)
- [28] 真鍋宏幸; 平岩明; 杉村利明. 無発声音声認識: 筋電信号を用いた声を伴わない日本語 5 母音の認識. 電子情報通信学会論文誌 D, 2005, 88.9: pp. 1909-1917. (2005)

未来ビジョン：進化か先祖返りか？

我々人類は古くから主に音声によって人対人のコミュニケーションを行ってきた。この先、人類が情報空間と共に生きていく為には、対話先(人かネットか)の瞬時確実な切替手段が必要になるだろう(ウェイクワードやジェスチャーでは誤入力の可能性を排除にくい)。しかし我々が持つ唯一のアクチュ

エータである筋肉は、既に日常生活中で用いられている為、それらと干渉しない新たなモードを探すのは簡単では無い。例えば耳動筋等の痕跡器官を使う手法[23]は若干の訓練こそ必要ではあるが、新世代人類の基本技能としては有望だと思われる(例：耳を持ち上げながらひそひそ声で話すと音声入力)。今から耳を動かす練習をしておいた方が良いかもしれない(朝練耳上げ百回！)。

楽器未経験者のための弦管打複合電子楽器の開発

矢田 紘理奈 * 岩井 将行 *

概要. 楽器未経験者にとって、楽器を始めるハードルは高く日常的に楽器に触れる機会も限られている現状がある。演奏者を増やすためには何よりも「きっかけ」が重要であると考え、演奏方法に着目して簡易的に楽器を体験する機会を増やす方法を模索し、弦楽器、管楽器、打楽器の要素を一体化した複合電子楽器「kiMera」を考案した。この楽器は、表面に鍵盤、裏面にはヴァイオリンの構造、側面にはフルートのキー やトランペットのピストンが配置されている。各楽器の要素をセンサーで再現し、MIDI 規格で音を生成することで、ユーザーが複数の楽器を一度に体験することができるコンパクトな楽器である。試作機の外見に対するアンケートでは、92%が各楽器のイメージを持つことができたと回答しており、主観的な評価としては、演奏は難しくとも音を出す体験が容易になったことを示す。

1 はじめに

日本における楽器演奏人口は約 10%（令和 3 年度調査 [13]）に過ぎない。これは、楽器に触れる機会が限られていることが一因と考えられる。世界には歴史的価値と美しい音色を持つ楽器が多く存在し、音楽文化において重要な役割を果たしてきたが、それらの楽器に触れる機会は限られており、その魅力を直接知る機会は少ない。楽器を始めるきっかけは、楽曲を聞いたり、周囲の影響を受けて演奏したくなるケースと、楽器のある環境や習い事が要因となるケースに大別される。つまり、楽器自体や音色、その楽器の知識に触れることがなければ、演奏を始める機会も生まれ難い。学校教育においては、リコーダーやピアノ、太鼓や木琴・鉄琴などに触れる機会があるが、オーケストラや吹奏楽で使用される本格的な楽器に触れる機会は、部活動に入るなどしない限り皆無である。また楽器の持つ固定観念 [1]（例えば、ヴァイオリンは裕福な家庭で親しまれ、エレキギターにはカジュアルで軽快な印象が伴うなど、楽器に対するやや偏ったイメージ）が楽器演奏の敷居の高さにつながり、こうした先入観で楽器に触れる機会を逃すこともある。こうした現状を踏まえ、本研究では、楽器演奏人口を増やすことを目指し、従来楽器を始めるきっかけになる体験楽器として、弦管打複合電子楽器「kiMera」を提案する（図 1）。

複数の楽器を演奏できるものとして、スマートフォンアプリやシンセサイザが挙げられる。スマートフォンアプリにはキーボードやギターなどの楽器アプリが存在するが、画面を介した操作では、実際の楽器演奏とは異なり、鍵盤を押す、弦をはじくことはできても、弓で弦を弾く感覚や管楽器を演奏する実体験は得られない。シンセサイザで各楽器の音を出すことは可能である

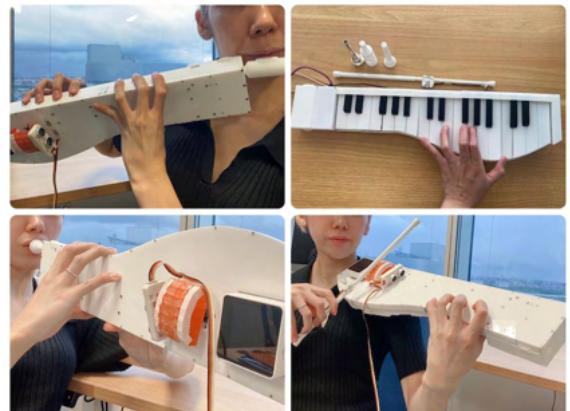


図 1. kiMera 各モードの構え方

が、楽器の演奏体験を再現することはできない。シンセサイザの特徴は音色作りであって他楽器の体験を目的としたものではない。そこで演奏方法に着目し、弦楽器、管楽器、打楽器から代表的な楽器を選定し、それらを一つの楽器に統合することで、様々な演奏方法を体験できるようにすることを考えた。

楽器を模倣したものとして、市販されている玩具も存在するが、実際の楽器の難しさを排除しすぎた結果、演奏のリアルさを失ってしまう課題がある。一方で、初学者にとっていきなり何のガイドもなく演奏することは難しい。どこまで簡易化するのか、そしてどこまで従来の形を残すのかが、慎重に議論すべきポイントである。

複数の楽器を 1 つのシステムに実装するにあたり、ハードウェア上の制約や楽器演奏の難しさを軽減するための演奏方法については、オーケストラに所属していた経験をもとに考慮して設計した。また、楽器演奏には多くのメリットがあり、認知症予防など様々な研究 [2] がされており、従来楽器の難しさの

Copyright is held by the author(s).

* 東京電機大学大学院 未来科学研究科 情報メディア学専攻

軽減を目指した「誰でも簡単に演奏可能な楽器」が数多く開発されてきた[4]。しかし、本研究は「誰でも簡単に演奏可能な楽器」という位置付けではなく、「従来楽器を始めるきっかけとなる楽器」という位置付けで複合楽器を提案している。これまで IoT ヴァイオリン[16]という名前でヴァイオリン機構を作成し、その後鍵盤機構を作成した[15]。本稿では、本システムに採用した関連研究における技術を紹介し、従来の楽器演奏方法と比較しながら、音程、音量、発声タイミングをどのようにセンサで表現したのかを記述する。また、外見における評価実験の結果を報告し、本システムが複数の楽器体験を提供できるのか、さらに、それが従来楽器を始めるきっかけに繋がるのかどうかを議論する。

2 関連研究

2.1 無弦チェロ

従来楽器のハードルを下げるため、弦楽器、管楽器それぞれ電子化されたものがある。三井ら[11]は、チェロに着目し、ボディの大きさや重さ、音量の問題、購入時や維持費のコスト等の問題をあげ、これらはチェロの広音域のための弦の長さによるものと考えた。そこで、弦を無くしセンサで置き換えることで、電子楽器としてチェロを再現するアプローチを考案した。我々も同様に、弦を押さえる動作が初心者にとってつまずきやすい点だと考える。通常の演奏では、弦と弓の接触方法が重要となるが、本システムでは、演奏方法を簡略化し、弓と弦の触れさせ方を省略し、JOYSTICKによって発声タイミングを取っている。JOYSTICKが動いたときに弓の真ん中に取り付けた加速度センサが弓の角度を読み取り、左手の指の位置を超音波センサで読み取ることで距離から音程を算出した。このように弦を押さえる動作をなくすることで、体験する際のハードルを除きながら演奏方法を知ることができるシステムにした。しかし、体験楽器から従来楽器へ移行する際には、弦を押さえる指の痛みは発生してしまうため、どのように慣らすのかという課題にも対応する必要がある。

2.2 ウィンドシンセサイザ

気圧センサを用いて息の量を判定し、静電容量のタッチセンサによってキーを作成した管楽器を作る試みがある。黒川ら[10]はリコーダー型のウィンドシンセサイザを作成していた。静電容量センサをプリント基板で作成し、本体やマウスピースは3Dプリンタで作られており、本体は上部と下部で回転して使用することができる新たな機能を追加していた。気圧センサを使用すると、その場の環境の気圧に息の圧力がプラスされた値が読み取れる。そのため楽器起動時の気圧を基準に息の圧力を求められる。ま

た、従来楽器は息を入れて必ず出口が存在するが、販売されている既存のウィンドシンセサイザでは管内の蒸気による不具合を防ぐため、口の両端を開けて息を抜き、リードを噛む力を読み取るバイト(bite)センサを用いて音が出る判定を制御する方法をとる楽器がある[3]。本研究で実装する際には、楽器内の蒸気の対策として、防水加工がされているセンサを選び、息の出口を作り楽器内の手入れや複数人での使用も視野にいれ、衛生的に演奏可能な設計を心がけた。

2.3 EZ-TP

YAMAHAが2004年に発売したEZ-TP(イージートランペット)[6]はトランペットの電子化がされている。トランペットは唇で震わせたマウスピースの音が増幅して発声する楽器である。マウスピースの振動パターンとピストンの押さえ方によって音程が決まる。EZ-TPでは、マウスピースからの音を声で認識させるというマウスピースの簡略化を行った。本システムでのトランペット機構は、先行研究[5][8]を参考に、マウスピースの振動を読み取って音に変換しようと考えていたが、マウスピース自体にセンサをつけることになると、マウスピースの取り外しができず、複数人での使用が困難になるため、前者のEZ-TPを参考にPDMセンサでマウスピースを吹いた際の周波数を測定することにした。弦楽器での弓や弦の懸念点と同じく、金管楽器のハードルはマウスピースを鳴らすことである。しかし、マウスピースを鳴らすということを体験するのは金管楽器を演奏するにあたって比重が高いと考えた。マウスピースをうまく制御できなくても、初心者でも吹くことができるよう、システムによって音程の誤差をアシストするように設計した。

2.4 複合楽器

本研究では、演奏方法をそのままに組み合わせて一つの楽器として設計することを考案しているが、複合楽器としてUnitInstrumentsを紹介する。既存の電子楽器は従来楽器の形を模倣することが主な目的だったが、楽器を発音、音程の決定などの構成要素(ユニット)の集合であると考え、これらの構成要素を部位ごとに分離し、他ユニットとの組み合わせを行うことによって新たな形の楽器を組み合わせられる楽器を丸山ら[7]が作成した。各ユニットを組み合わせる場合、指板ユニットがピックアップユニットに2方向についているマルチネックギター、ギターのように押弦し鍵盤で打鍵して演奏する楽器を作成することができる。従来楽器の構成要素を組み合わせることによって演奏の難易度は高くなってしまうため、未経験者向けの楽器ではなく、上級者向きの楽器であるが、楽器を構成要素の単位として見る視点が参考になった。また、本システムも演奏活

楽器未経験者のための弦管打複合電子楽器の開発

動にも応用できそうな機能をつけていくことによって体験楽器としてだけでなく、新しい楽器としての面も補えると考えた。

3 楽器の選定

楽器は、弦楽器、管楽器、打楽器という大枠の分類がされ、管楽器も木管楽器と金管楽器に分かれる。楽器の大きさから音の高さ、構え方や奏法の違い、もちろん音色に至るまで、それぞれの特徴を持っている。本システムを開発するにあたって、弦楽器、木管楽器、金管楽器、打楽器から楽器を選定することにした。

初めて手にする人が手軽に扱える適度にコンパクトで持ちやすい大きさが望ましいと考えた。また、楽器には認知度に差があり、全く知らないものよりも、名前を聞いたことがある楽器の方が興味を引きやすい。そのため、弦楽器、管楽器、打楽器のそれぞれから、最も認知度が高く、典型的な楽器を選定することとした。

さらに、小型化を目指すと、必然的に高音域の楽器を選ぶこととなる。具体的には、打楽器からはピアノを選んだ。ピアノは鍵盤ハーモニカとしても馴染みがあり、鍵盤のビジュアルから手に取りやすい印象を与えることができると考えた。弦楽器では、高音域のヴァイオリンを選んだ。管楽器からは、金管のトランペットと、横持ちという特徴を持つ木管のフルートを選定した。

4 システム

システム構成図を図2に示す。使うマイコンとセンサを表1に示す。M5Stack-Core2(以下M5Core2と示す)と、加速度付きでWi-Fi通信も可能なArduino nano33 IoT(以下nanoと示す)を採用した。赤で示したセンサはM5Core2に、青で示したセンサはnanoに接続されている。センサから読み取られた値によって得られた音声情報を、SAM2695を使用しMIDI規格でスピーカーから出力している。シリアル通信でモード切り替えを行ってそれぞれの楽器を演奏することができる。各モードごとに構えると図1のようになる。(上段左がフルート、右がピアノ、下段左がトランペット、右がヴァイオリン)

外装設計

外装はfusion360を用いて作成した。全体像の3Dモデルを図3に示す。3章で述べた楽器の選定基準より、ピアノは鍵盤ハーモニカと考え、その他ヴァイオリン、トランペット、フルートは大体大きさが50~60cmの楽器であるため、本システムの最長辺を50cmとした。レーザカットしたアクリル板(図4)と、他部品を3Dプリンタで生成した。

実際の鍵盤の大きさと大きく違わぬよう鍵盤の幅

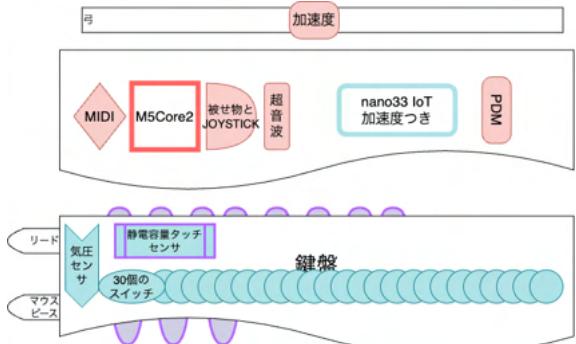


図2. kiMera システム構成

表1. 使用するセンサ・デバイス。

	M5Stack Core2	nano 33 IoT
ピアノ	MIDISynth	ソフトタクト スイッチ
ヴァイオリン	超音波センサ 外付け加速度 JOYSTICK MIDISynth	内蔵加速度 -
トランペット	MIDISynth PDM	静電容量 タッチセンサ
フルート	MIDISynth	静電容量 タッチセンサ 気圧センサ

を白鍵は21.5mm 黒鍵は10mmで設計した。丸みを帯びたデザインに沿わせて鍵盤を作成したため、鍵盤の長さは変化している。鍵盤の沈み込みは既存の電子ピアノを参考に設計した[14]。既存の電子ピアノの鍵盤の機構は根本にバネがついており、鍵盤を押すとバネの力で滑らかに沈み込む仕組みで、下がってきた鍵盤の窪みに圧力センサが押されている。この機構を参考に、図5のように鍵盤の根本に穴をあけ、棒を通してプラスチックの柔軟性を利用し、バネではない方法で鍵盤の沈み込みを再現でき、全ての鍵盤が通った棒を両端でねじ止めし固定することもできた。

ヴァイオリンは弦を弾く時にできる弓の角度によって弦を判定するため、弓を模した棒の中央に加速度センサを設置した。JOYSTICKはゲームのコントローラのため、そのままでは弓を置くことができない。弦の判定がしやすいうようにJOYSTICKのヘッドに被せられる角度をつけた部品(図6)を作成した。

フルートのキーとトランペットのピストンは図7のように側面に設置している。静電容量センサの後に金属製のボタンやスペーサーを付け通電するように



図 3. 全体像の CAD データ

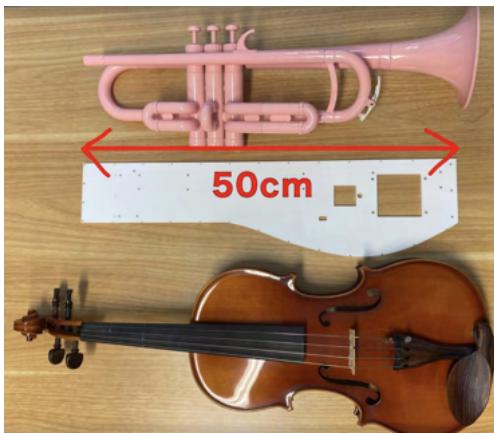


図 4. 従来楽器とのサイズ比較

した。複数人で使用することも考慮にいれ、リードやマウスピースが取り外せたり、実際の(市販の)マウスピースを使用することも可能にした。また、図8のように蓋をつけることで手入れのしやすさも考慮した作りになっている。トランペットのピストンを作成しようと考えていたが、ヴァイオリンの演奏時に妨げてしまうことや楽器の幅が大きいことでピストンに手が届かない恐れが出たため、フルートのキーと同様に金属製のボタンで作成した。そのため、トランペットのピストンの位置とフルートの左親指のキーと同じ場所に作成することになった。今後の発展としてキーの周りにLEDを実装するなどしてキーの差別化を図りたい。

回路基板

本システムのために二つの回路基板をKiCadを用いて設計した。一つは鍵盤の下に設置するスイッチのための基板である。もう一つはM5Core2にセンサを接続するために端子を増やした基板である。作成した基板を組み立てると図9のようになる。基板上のソフトタクトスイッチの上に鍵盤が配置され、図での底であるレーザーカット面の穴から弓の中央に設置した加速度センサのケーブルが伸びている。



図 5. 鍵盤の設計



図 6. JOYSTICK に被せる部品

演奏制御

各センサが繋がったマイコン同士が、シリアル通信で値をやりとりしている。電源はnanoに給電することでM5Core2も起動する。M5Core2の画面からモードを選択し、各楽器を演奏できる。ピアノは鍵盤を押しているかどうかを一音ずつフラグを用意することで和音も鳴らすことができる。ヴァイオリンはJOYSTICKが動いている間、超音波と加速度が動き、2つの値から音程を定めてMIDIが鳴るようになっている。JOYSTICKが動いたタイミングでnanoに合図が送られ、nanoの加速度によって得られた本体の角度をM5Core2に送り、本体と弓の間の角度が測定できるようになったため、より正確に音程を定めることができた。トランペットはPDMセンサが一定の音量を検知したらその音の周波数からマウスピースの音を基準音(B♭やF)に変換し、ピストンのパターンで音を設定している。フルートではLPS3xの気圧センサで一定の息の量を検知し、キーのパターンで音程を定めた。指使いはリコーダーにプラスして既存のウィンドシンセサイザを参考に、半音あげるキーと下げるキーを左手の小指で操作できる場所に設け、左手の親指で3オクターブを操作できるよう設けることでフルートの音域をカバーした。

5 評価実験

本研究のシステムを評価するにあたって、従来楽器の簡略化が適切に実装されているか、楽器として機能するかという面と、未経験者にとって興味を持つきっかけになり得るか、という面を評価実験においての先行研究[9][12]を参考に、調査する必要がある。まずは弊学のオープンキャンパスにて、本シス

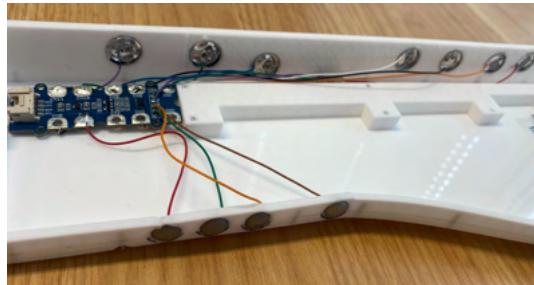


図 7. 管楽器のキーとピストン



図 8. 吹き口手入れ用の蓋

テムの外見がそれぞれのモデルである楽器をイメージできるかのアンケート調査を行った。

5.1 外見の評価実験

本システムのハードウェアについて、各奏法のモチーフがどの程度認識されるかを評価するため、アンケートを実施した。被験者には楽器の概要を説明し、各面がどの楽器に対応するかを示した後、それぞれの奏法がイメージできるか、また複数の楽器を統合することで楽器の認知度が高まると考えるかについて、5段階評価で回答を求めた。前者は25件、後者は35件の有効な回答が得られた。各楽器の奏法がどの程度イメージできたかの結果を図10に示す。結果として、40%が「かなりイメージできた」と答え、52%が「まあまあイメージできた」と回答した。特に、ピアノの面を上にして楽器の概要を説明したため、第一印象としてピアノのイメージが強く、その後に他の楽器も演奏できると伝えると、多くの被験者が驚きの反応を示した。また、クラリネットやホルンといった他の従来楽器も実装して欲しいという意見や、デザインに対するポジティブな意見も多く寄せられた。さらに、体験楽器としてのコンセプトも十分に伝わり、複数の楽器を統合することで楽器の認知度が高まるかという設問には、48.6%が「と



図 9. 楽器の内側

てもそう思う」と答え、40%が「そう思う」と回答した。一方で、予想通り「使いにくそう」といった意見も見られた。今回の評価では、外観に基づく認識のみを対象とし、使用感についての質問は含まれていなかったが、複数の楽器の要素を統合しつつ、一般的な楽器のイメージから大きく逸脱することなく実装することができたと評価された。今後、楽器に興味を持っていない人や、難しそうというイメージを持った人などに対して、本システムを説明した同一のシナリオを用意し、体験することでイメージが変わったか、楽器に対して興味が湧いたか、複数の楽器が合わさることで一度の機会で体験することができたことに対する意見を問うアンケートを行う。

kiMeraを見て弦・管・打楽器の弾き方がイメージできましたか?(25件の回答)

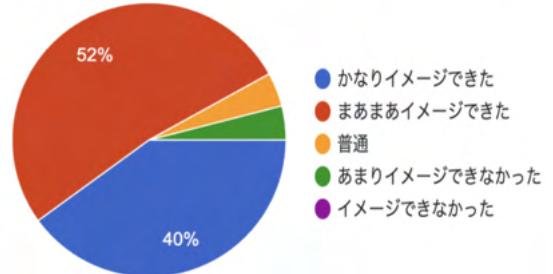


図 10. アンケート結果 1

複数の楽器(演奏方法)を一つにまとめることによって楽器の認知度は高まると思いますか?
(35件の回答)

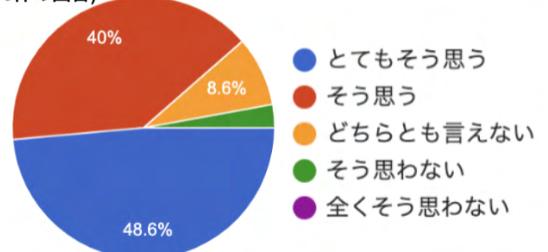


図 11. アンケート結果 2

5.2 主観評価

本システムが楽器として機能するかどうかを評価するために各モードを実演し、著者自身の主観評価を行った。

まず、ピアノは簡単な曲を演奏するレベルには達していると感じた。丸みを帯びたデザインは弾きにくさに個人差があると思われるが、見た目のインパクトが強調されていて非常に良いと感じた。フルートの機構と合わせて鍵盤ハーモニカのようにも演奏できるようになれば機能の追加もできる。

ヴァイオリンは、nanoによって本体の角度を算出できることから精度の向上が見られたが、I2C通信の遅延が目立ち、反応面で不具合が生じた。また、外装のデザインに丸みをつけたため、一番端のE線を弾く際に弓が楽器の縁に当たってしまい、弾きづらくなってしまうという設計ミスを発見した。JOYSTICKに取り付けた角度をつけた部品の足を少し伸ばすことで対応できればと考えている。さらに、MIDIの不具合により、ヴァイオリンの音ではなくピアノの音が出てしまうことがあり、この問題の解決が必要である。これはトランペットでも同様に発生し、ソフトウェア面での改善が課題であると考えられる。

トランペットの吹き心地はマウスピースを鳴らせるかどうかにかかっているため難しさが残ってしまう結果となった。著者はトランペットの経験が浅く、マウスピースをうまく鳴らすことができなかつたため、声で評価を行ったが、意図した通りに音が出ることがわかった。しかし音量によって発声を検知しているため、物音で鳴ってしまう点や、音が連なって鳴ることもあった。音声の機械学習を用いてマウスピースの音やタンギングを判断できるようになれば、スムーズな演奏が可能であると考え、今後実装を進めていく。

フルートは、リコーダーの指使いにしたことで吹きやすさを感じた。音の切り替えもうまくいき、少しの遅延はあるものの、狙い通りの音程が出た。また、タンギングの反応や息の読み取りもスムーズであると感じた。著者は肺活量が弱いため、息の量の最大値を事前に設定する等で奏者個人に合わせることができれば吹きやすさにつながると考える。

総じて、どのモードも従来の楽器のように滑らかに演奏するレベルには達していない。しかし、見た目はコンパクトにまとまり、各楽器の特徴に寄せることができた。今後、母校のオーケストラ部に協力を依頼し、楽器としての有用性を評価する。

6 まとめ

本研究では、楽器未経験者が楽器演奏を体験する機会を増やすため、弦楽器、管楽器、打楽器の要素を一体化した複合電子楽器「kiMera」を提案した。

この楽器は、ピアノ、ヴァイオリン、トランペット、フルートという代表的な楽器の特徴をセンサ・デバイスで再現し、ユーザーが一度に複数の楽器を体験できる設計となっている。

評価実験では、本システムの外装に対して25件中92%の被験者が各楽器のイメージを持つことができ、楽器の再現性とユーザへの親しみやすさが確認された。また、主観評価では、演奏は難しくとも音を出す体験が容易になったと感じ、kiMeraが楽器を始めるきっかけとなり得る可能性を示す結果となった。今後の評価実験での結果によって有用性を示すことができればと考えている。しかし、本研究にはいくつかの課題が残されている。各楽器の演奏体験をより精密に再現するためのセンサー技術の改善や、複数のセンサーを同時に使用した場合の精度向上が挙げられる。

今後の展望としては、システムの改善に努め、外部デバイスに繋げての作曲作業の機能を追加するなど、体験楽器に留まらないシステムの拡張を目指したい。これにより、楽器未経験者のみならず、経験者にとっても魅力的な楽器となる可能性が広がる。本研究が示すように、楽器演奏の体験を提供する新たな手段としての複合電子楽器は、音楽教育やエンターテインメントの分野で多くの可能性を秘めている。今後も更なる改良と発展を重ねることで、音楽文化の普及と楽器演奏人口の増加に貢献したい。

謝辞

本研究の遂行にあたり、株式会社PlanckUnitsの野中直樹氏には、基板設計の多大なご支援、アドバイスを頂きました。本論文の執筆に先立ち、明治大学総合数理学部先端メディアサイエンス学科の宮下芳明教授には、多大なご助言を賜るとともに、本研究で作成した楽器の命名にあたりご示唆を頂きました。名古屋大学大学院工学研究科の米澤拓郎准教授には、本論についてご議論を頂きました。東京電機大学ものづくりセンターには、本楽器の部品作成にて技術協力を受けました。最後に、本研究の基盤となった、関東学院中学校高等学校オーケストラ部では貴重な音楽体験をさせて頂きました。また、本システムの評価実験のご協力もして頂きました。全ての方々に心より感謝申し上げます。

参考文献

- [1] I. M. Cantero and J.-A. Jauset-Berrocal. Why do they choose their instruments? *British Journal of Music Education*, 34(2):203–215, 2017.
- [2] D. Marie, C. A. Müller, E. Altenmüller, D. Van De Ville, K. Jünemann, D. S. Scholz, T. H. Krüger, F. Worschech, M. Kliegel, C. Sinke, and C. E. James. Music interventions in 132 healthy older adults enhance cerebellar grey matter and

楽器未経験者のための弦管打複合電子楽器の開発

- auditory working memory, despite general brain atrophy. *Neuroimage: Reports*, 3(2):100166, 2023.
- [3] AKAI PROFESSIONAL(2024/11/23 確認). EWI SOLO. <http://ewi.akai-pro.jp/ewi-solo>.
- [4] R. Takahashi, Y. Maeda, K. Tanaka, T. Ichinose, R. Okuno, and K. Akazawa. Development of a Novel Breath-touch Electronic Instrument that Enables Beginners to Engage in Ensemble Playing. *Advanced Biomedical Engineering*, 12:64–73, 2023.
- [5] C. Vergez and X. Rodet. Comparison of Real Trumpet Playing, Latex Model of Lips and Computer Model. *ICMC*, 01 1998.
- [6] YAMAHA. EZ-TP(イージートランペット). <https://www.yamaha.com/ja/about/experience/innovation-road/collection/detail/0052/> (2024/11/23 確認).
- [7] 丸山裕太郎, 竹川佳成, 寺田努, 塚本昌彦. UnitInstruments: 楽器の機能要素を再構築可能なユニット型電子楽器の設計と実装. 日本ソフトウェア学会, 2011.
- [8] 吉川茂, 足立整治. 金管楽器裏者の唇はどのように振動しているか. 日本音響学会誌, 52(5):389–395, 1996.
- [9] 宮下芳明. インタラクション研究でのシングルケース実験についての考察. エンタテインメントコンピューティングシンポジウム 2019 論文集, 2019:248–254, 09 2019.
- [10] 黒川陸. 簡易に演奏可能なリコーダー型ウインドシンセサイザーの開発, 2020.
- [11] 三井実. マイコンを用いた電子楽器開発記 1～「無弦チェロ」の開発～. 技能と技術, 2017.
- [12] 小出英範, 西村明. 従来楽器とは異なる電子楽器の評価手法の検討. 日本音響学会誌, 77(1):7–15, 2021.
- [13] 総務省統計局. 令和3年社会生活基本調査. <https://www.stat.go.jp/data/shakai/2021/pdf/gaiyoua.pdf>(2024/11/23 確認).
- [14] 北口二朗. 電子楽器の技術発展の系統化調査. 技術の系統化調査報告, 26:1–90, 3 2019.
- [15] 矢田絵理奈, 岩井将行. 楽器未経験者に向けた弦管打複合の IoT 楽器の提案. エンタテインメントコンピューティングシンポジウム 2023 論文集, 2023:419–422, 08 2023.
- [16] 矢田絵理奈, 岩井将行. 楽器未経験者のための弦管打複合楽器の開発を目指した IoT ヴァイオリンの制作. インタラクション, 2023.

未来ビジョン

電子楽器は、音量制御や多様な音色の再現、持ち運びの容易さといった特徴により、従来のアコースティック楽器に代わる新たな表現手段として、20世紀後半から急速に普及した。電子楽器はアコースティック楽器に代わるものとして台頭してきたが、著者はアコースティック楽器も電子楽器も、どちらの素晴らしい点も広めていきたいと考えている。本システム「kiMera」は体験楽器としての役割を果たしており、将来的には体験楽器を脱却し、従来の楽器へとレベルアップしてほしいと考える。しかしながら、kiMera の電子楽器としての有用性も高めるため、今後の展望としてアプリ等の接続を通じて多様な機能を拡張する可能性も視野に入れている。最初に述べた通り、楽器に興味を持たない人は自ら楽器に触れる機会を求めようとはしないであろう。逆に、楽器に興

味を持つ人は自ら購入を検討したり、楽器店で試奏を行うことが容易に想像できる。kiMera は、楽器に興味を持たない人にこそ使用してもらいたい楽器である。自ら動かなくても自然に楽器に触れる機会が増える世界に寄与したいと考えている。音楽を聴く人は多いが、楽器を弾くことへのハードルが高い現状は、音楽を楽しむ上で非常にもったいない状況である。音楽が好きなら、カラオケに行くことがあるかもしれない。その場にタンバリンの代わりに kiMera があったらどうだろう。小学校時代、各教室に一台あったオルガンのように、kiMera が設置されていたらどうだろう。新しい体験を提供することは十分に可能であると考える。情報を得て満足してしまう時代に、新しい体験を提供していきたい。

歌詞に基づく歌声アノテーションのためのインタフェース構築

中野 優靖 * 加藤 淳 * 渡邊 研斗 * 濱崎 雅弘 * 後藤 真孝 *

概要. 本稿では、歌声に対する時間局所的なアノテーションを行う際に、その歌詞を用いるインタラクションを提案する。従来、時系列メディアのアノテーションでは、アノテーション内容に時刻情報を含める強ラベルと、時刻情報が含まれない弱ラベルを基本として、それらの派生や改善が提案されてきた。本研究では、歌詞の文節を選択するだけでその時刻情報を指定でき、簡単にアノテーションできる「歌詞ベース」のアノテーションを提案する。歌詞ベースのアノテーションでは、その音源を再生するプレーヤと、既存のテキストエディタや Excel 等のスプレッドシートがあれば可能であるので、本稿ではまず、Excel をアノテーションエディタとして用いて実際にセマンティックタグをアノテーションした結果を分析することで、実用性を検証する。そしてさらに、その使いやすさを向上するためのインターフェースとして、Lyrics-Based Singing Annotator を提案する。本インターフェースでは、クリック可能な歌詞と音源を同期して再生する機能、付与対象の歌詞をループ再生する機能、特定のタグが付与された歌詞をハイライトする機能を持つ。

1 はじめに

音楽へのアノテーションは、機械学習の学習データとしての利用や、音楽の特性分析などのために重要である。音楽や音声などの時系列メディアへのアノテーションでは、ジャンルや歌手名などの楽曲固有で時不变なラベルだけでなく、音響イベント・歌唱テクニック・セマンティックタグなどの時変なラベルが存在する。これら二種類のラベルは、時刻情報付きラベルが強 (Strong) ラベル、時刻情報なしラベルが弱 (Weak) ラベルと呼ばれている（図1）。

機械学習データにおいては弱ラベルが広く利用されることが多い [24, 31, 43]。個人ごとの結果の変動を吸収するためには複数人によるアノテーションが必要となるが、強ラベルは困難かつ時間消費の大きい高コストな作業であり、多くのデータを収集するのが難しいためである。例えば、音源の一部分（例えば、10秒間）を切り出し、その区間にラベルの存在のみをアノテーションする弱ラベルは、強ラベルより時刻的な精度は低下するが、相対的に低コストで収集が容易するために採用される。

しかし、時刻情報を含むラベルは依然として重要である。音響イベント検出等、音の種類と時刻を同時に推定するタスクでは、評価のために時刻情報が必要となるためである [30]。さらに、自動分類のための機械学習モデルにおいて、強ラベルを用いることによる性能向上が報告されている [14]。そこで、弱ラベルのようにアノテーションコストを下げながら、時刻情報の正確性を上げる方法として、点 (Point) ラベル [22] と重複弱ラベル (Overlapping weakly-label) [29] が提案された（図1）。前者の点ラベル

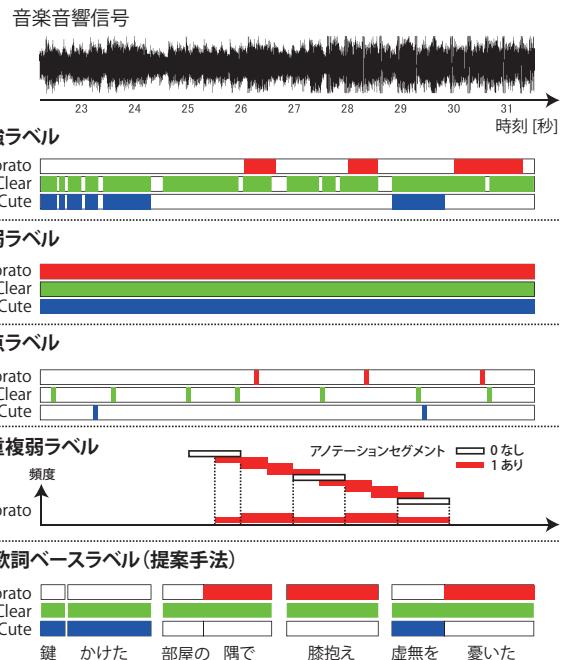


図 1. 5 種類の時系列ラベル例。歌詞ベースラベルの全てと、強ラベルの Vibrato のみが実データに基づいており、それ以外は近似的に図を作成した。Clear と Cute について、点ラベルは歌詞ベースラベルと同様に文節区間ごとに配置し、強ラベルは最も精密なアノテーションを想定して音素区間に配置した。重複弱ラベルは、セグメント長を 1 秒、シフトを 0.5 秒として作図した。

は 1 回の指摘で一つの音響イベントを示すことができ、後者の重複弱ラベルは複数アノテータによる結果の統合が行いやすいという特長がある。

点ラベルでは、複数アノテーション結果を統合す

Copyright is held by the author(s).

* 産業技術総合研究所

るためには、それらの対応関係を別途推定する必要がある。また、長く持続するラベルについて、どの時刻を指摘するのが良いか直感的ではない。一方、重複弱ラベルは、一つの音響イベントに対して複数のアノテーションが必要となって、アノテーションが必要となる回数が増える。また、アノテーション対象（セグメント）の長さやシフト時間を短くすると（時間分解能を上げると）、さらに回数が増えることになる。点ラベルではアノテータが時間を指定する必要があるが、重複弱ラベルでは不要である。

そこで本稿では、作業コストが低く、複数アノテーション結果の統合がしやすい新しい弱ラベルとして歌詞ベースラベルを提案する。つまり、歌詞の文節を単位として、それに該当する歌声にラベル—音域（高音域、中音域、等）、テクニック（ビブラート、しゃくり、等）、意味（かわいい、明るい、等）—をアノテーションするためのインタフェースを構築する。つまり、歌詞という音響イベントに対して、そこにさらに意味的なラベル付けをすることになる。

歌詞ベースラベルには以下の特長がある。まず、歌詞は自然言語であり、波形等の音響特徴量に不慣れな人にもなじみ深い。次に、歌声の発声区間と紐づいており、固定長の時間切り出しそりも詳細な時刻情報が得られる。また、複数人のアノテーション結果を文節ごとに対応付けることも容易である。さらに、歌詞の異なる2名以上の歌唱が重複する場合には、そのいずれに対するアノテーションかを明確にできる。歌聲音楽音響信号の中ではイントロや間奏など、歌声が含まれていない区間も多いため、そこを効率的に除ける点でも有用である。

例えば、図1の「Clear（クリアな、透き通った、澄んだ）」のように、歌詞のあるフレーズが全て同じ意味を持っていた場合、歌のない短い区間を避けて強ラベルをアノテーションすると（このような区間を避けるアノテーションはかなり難しく、かつコストが高い）、歌詞のいずれかの単位（音素や文節など）でラベル付けすることと同じ結果となる。

なお、歌詞の時刻についてもアノテーションが必要となるが、歌声と歌詞の自動対応付け（アラインメント）技術[8, 11, 36, 38, 41, 45]を活用することができる。この際、曖昧性は少ないため、複数人によるアノテーション結果の統合は不要である。

2 関連研究

従来、時系列メディアに対するアノテーションインターフェースとして、強ラベルを付与できるインターフェースが提案してきた。音響信号や動画を対象として、Praat [3]、Wavesurfer [40]、CLAM Annotator [1]、MUCOSA [13]、ELAN [50]、Sonic Visualiser [5]、Timeline Annotator [54]、EMU-webApp [49]、BAT [32]、CrowdCurio [7]、VGG Image Annotator (VIA) [9]、GECKO [27]、au-

dino [12] 等がある。楽譜レベルでのアノテーションインターフェース Score Annotator [54] も提案された。これらのうち、複数アノテータによるアノテーションを考慮したインターフェースもある[9, 12, 32]。さらに、アノテーションコストを下げるために、機械学習モデルによって自動アノテーションする手法として、区間の自動分割と認識を行う EASEL [44]、認識モデルをインタラクティブに更新する SoundScape [25] と I-SED [20, 21] が提案された。以上のインターフェースは、開始時刻・終了時刻・ラベルの3種類、もしくはそれに周波数帯域を加えた4種類[25, 54] を入力するインターフェースである。

一方で、弱ラベル付与のためのインタラクション[6, 23, 46] も提案されてきた。Wang *et al.* は、ラベリング効率を上げるために、類似区間をクラスタリングして同時にアノテーションを提案した[46]。Cartwright *et al.* は、同一区間に重複して存在する音源アノテーション（複数ラベル）のために、複数回のバイナリアノテーション（付与するかしないか）、単一パスでの複数ラベルアノテーション、階層的複数パス・複数ラベルアノテーションの結果を比較した[6]。Kim *et al.* は、歌声へのセマンティックタグ付与のためのインターフェースを構築し、曲単位でのアノテーション結果を初期値として、10秒のセグメント単位でアノテーションさせた[23]。このような弱ラベルの時間長には、MagnaTagATune (MTAT) [26] は30秒、Million Song Dataset (MSD) [2] は30秒、MTG-Jamendo [4] は30秒以上のフル尺（1曲）、CAL500exp [46] は3~16秒の可変長、歌唱タグ Kpop Vocal Tag (KVT) [23] では10秒、音楽における Multiple Instance learning に関する研究 [28] では10秒、アマチュア女性歌唱者の意味ラベル [53] では約9秒（固定フレーズ）が用いられた。

以上のような強ラベルと弱ラベルの両者の利点を得るために、新しいアノテーション手法が研究されている。Kim *et al.* [22] は、強ラベルよりも簡便で、弱ラベルよりも時間精度を上げる点ラベルを提案した。アノテーターは音の発生時にマウスクリックするなどにより、音響イベント名を、その音が発生したいずれかの時点でアノテーションする。点ラベルが含まれているセグメントに関して、バイナリクロスエントロピーロスを計算する方法、点ラベルを前後（時間）に拡張する方法が提案された。Martín-Morató *et al.* [29, 30] は、10秒の長さ（音響イベントの長さに基づいて決定）を持つセグメントを1秒シフトさせることで、弱ラベルから時刻情報を特定する方法を提案した。アノテータは時刻を指定する必要はないために作業はシンプルであり、クラウドソーシングを用いてアノテーションを収集しやすい。アノテータの能力を推定する手法を提案して、信頼性の高いアノテーション結果のみを用いた。

話声・歌声信号の代わりに、歌詞など、対応する

歌詞に基づく歌声アノテーションのためのインタフェース構築

テキストを扱うインタラクションもこれまで提案されてきた。Fujihara *et al.* [11] は、歌詞の自動アラインメント結果を活用して、歌詞に基づく音楽シーク機能を提案した。最近では、Amazon Music や Apple Music などの音楽再生プレーヤにも、歌詞のテキスト(行)で再生時刻をシークできる機能がある。また、音声収録結果を編集するために、音声波形を音声認識結果のテキストで編集できるインタラクションが提案されてきた [15, 33, 37, 39]。その他、歌詞の自動同期結果を用いて、同じ歌詞やそのトピックが歌われている別の曲を検索できる Hyperlinking Lyrics [10] や LyricListPlayer [34, 57]、より一般的な歌詞駆動型のインタラクティブな視覚表現として「リリックアプリ」[18] が提案されてきた。

以上、アノテーションインターフェースや、歌詞を活用するインタラクションは提案されてきたが、歌詞ベースでのアノテーションに着目するインタラクションは研究されてこなかった。

3 歌詞ベースラベル

歌詞ベースラベルは、手動で付与した歌詞ラベルや自動歌詞アラインメント結果を活用し、時刻情報を持つアノテーションを間接的に可能とする。

3.1 単位

歌詞をアノテーションする単位としては、ユーザが指定するか(可変長)、事前に分割しておくことが考えられる。ユーザが単位を指定する場合、強ラベルと同様、開始時刻と終了時刻を指定する必要があつて、ユーザに要求する行動が複雑になり、また作業コストが増える。したがって、本稿では、事前に歌詞を分割することとした。歌詞の分割が共通だと、複数アノテーション結果を統合しやすい利点もある。

事前の分割の単位としては、日本語であれば、段落、文、文節、形態素、文字、音素などが考えられるが、本稿では文節を対象として、CaboCha/南瓜¹を用いて自動分割(係り受け解析まではせずに、文節解析のみ)した。さらに、本来の歌詞に存在する空白と空行でも分割した。

3.2 歌声ラベル

歌声アノテーションに用いるラベル(歌声記述子)は、歌唱や声質に関する先行研究 [17, 19, 23, 53, 55] に基づいて決定した。これらの研究では、アノテーター間の合意、了解性、または同義性が考慮された。まず、Kim *et al.* による KVT データセット [23] で使用されたセマンティックタグと歌唱タグを含む CAL500exp [46] を参考にした。音域(高域、中域、低域)については、男女合わせて 6 つの記述子とした。また、音色に関する 30 の記述

歌詞	やっと出逢えたね 君に逢うことだけが 夢にだって見ちゃうくらい ゆるぎない願いなの
----	--

マークアップを用いたアノテーション

<Energetic><Shakuri>やっと <Dynamic> 出逢えたね
</Dynamic></Shakuri> 君に逢う<Shakuri>ことだけが
夢にだって<Dynamic>見ちゃうくらい</Dynamic>
ゆるぎない願いなの</Shakuri></Energetic>

スプレッドシートを用いたアノテーション

	Shakuri しゃくり	Energetic エネルギッシュな	Dynamic ダイナミックな
やっと	1	1	
出逢えたね	1	1	1
君に		1	
逢う		1	
ことだけが	1	1	
夢にだって	1	1	
見ちゃうくらい	1	1	1
ゆるぎ	1	1	
ない	1	1	
願いなの	1	1	

図 2. シンプルなアノテーションインターフェース例

子(Husky/Throaty, Thick, Thin, Warm, Bright, Clear, Relaxed, Dark, Energetic, Mild/Soft, Sharp, Rich, Rounded, Stable, Breathy, Lonely, Sad, Passion, Charismatic, Pretty, Cute, Delicate, Emotional, Pure, Robotic/Artificial, Embellishing, Sweet, Young, Compressed, Dynamic)を用いた。

次に、歌唱印象 [17, 53] と声質 [19, 55] に関する研究で共通に用いられていた以下の 7 つの記述子(Powerful, Nasal, Calm, Weak, Sexy, Resonant, Dosu (Threatening/Frightening))を追加した。さらに、歌唱印象に関する研究 [17, 53] から、上記に関連する Beautiful, Cool の 2 つの記述子を追加した。

また、歌唱テクニックは重要なため、KVT データセットと、歌唱タグ・歌唱テクニックに関する先行研究 [16, 46, 48, 51, 56] を参考に、12 の記述子(Whisper/Quiet, Shout, Vibrato, Falsetto, Spoken/Speech-like, Fall, Growl/Scream, Kobushi, Mix voice, Rap, Shakuri, Vocalfry)を選択した。

以上、合計 57 種類の歌声ラベルを用いた。

3.3 シンプルなアノテーションインターフェース

歌詞ベースラベルは、仮に歌詞テキストに対する時刻情報がなくても、ユーザ(アノテータ)は歌詞と音楽を対応づけて聴取できる特長がある。つまり、既存の音楽プレーヤとエディタさえあればアノテーションできる。例えば、図 2 に示すように、歌唱ラベルをマークアップとしてテキストに付与したり、スプレッドシートの縦軸に歌詞、横軸にラベルを配置してアノテーションできる。

そこで本研究ではまず、歌詞ベースラベルの特性を分析するために、シンプルなアノテーションで歌

¹ <https://taku910.github.io/cabocha/>

詞ベースラベルを収集した。歌声ラベルの数が57と多くてマークアップは不適と考え、スプレッドシートとしてMicrosoft Excelを用いた。Excelでは文字の入力に応じてハイライトすることも可能で、各歌詞に付与されたラベル数もカウントできる(図4)。

歌声アノテーションでは、歌声と背景音楽との関係性が重要な場合があり、また歌声分離によるアティファクトの影響を避けるため、従来、背景音楽がミックスされた歌唱(オリジナル音源)でアノテーションされており[23]、本研究でも同様とした。

3.4 アノテーション

楽曲には産業技術総合研究所が学術目的で構築した非公開の音楽データベースを用いた。具体的には、2010年代後半のポピュラー音楽シーンを考慮して新規に作詞、作曲、編曲をして制作した多様な日本語楽曲120曲に対して、アノテーションした。アノテータは、母国語を日本語とする音楽のエキスパート(歌唱に関する知識、音楽的な知識及びそれらの評価に関する経験が十分あり、ポピュラー音楽の歌声を客観的に聴いて評価・タグ付けできる)の6名(M1～M3の男性3名、F1～F3の女性3名)で、性別が同一とならないように、3名ずつ2組「M1, F1, F3」「M2, M3, F2」に分けた。つまりアノテーションとしては、1曲毎に、いずれかの組のアノテータ3名がアノテーションした。

各組は、120曲の半分の60曲が割り当てられて、各文節において、カテゴリ「ピッチレンジ」、「音色(Low, 抽象度が低い)」、「音色(High, 抽象度が高い)」のそれぞれに分類されたラベルに対して、カテゴリ毎に必ず1つ以上のタグを、かつ、当てはまるラベルを漏れなく入力することとした。一方、カテゴリ「テクニック」に関しては、該当するラベルが存在しない場合は入力しなかった。歌唱としてハモリやコーラスが含まれる場合は、歌詞に対応するメインボーカルのみを対象とした。また、歌唱において、メインボーカルの人数が複数となる場合も、歌詞毎にアノテーションした。

3.5 結果

図3に、歌詞の行毎及び文節毎の文字数分布を示す。歌詞に最適化された文節推定ではないこと等が原因で10文字以上となることがあったが、平均的には3.71文字となった。

アノテータによる、文節単位のアノテーションの良かった点ややりやすかった点の回答を以下に示す。

- 各テクニックや音色などが様々な表現で詳細に分類されており、タグ付けの際に当てはまる項目を見つけやすかった。色分けされ視覚的にも見易く工夫されていた。
- 付与する箇所のみに「1」を入力するだけだったので、集中して作業できたのは良かった。

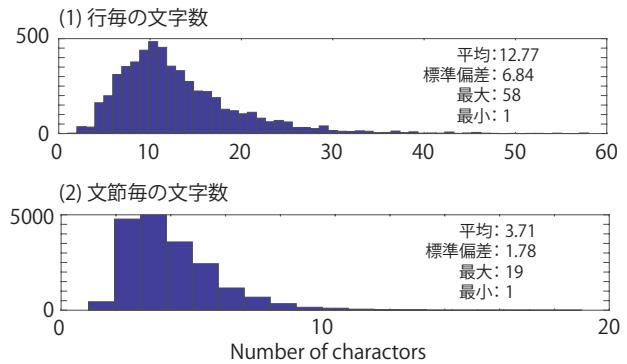


図3. 歌詞の行及び文節の文字数分布

- 細かい表情にまで対応出来る単位だった。
- 詳細な分類は分析において良い点だと思った。
- 普段ここまで聞き込まないので勉強になった。
- 「1」のみの入力だったので使いやすかった。

つまりアノテータは、「詳細な歌声ラベル」「時間方向の単位の細かさ」「文節選択のみ(1を入力)の簡易さ」に魅力を感じていた。

また、逆に悪かった点や、改善可能性のある点の回答を以下に示す。

- イメージが近いものなら、1つのセルの中に2つの項目をまとめるのも良いと思う。
- 単位が細かすぎる部分があり、その部分のニュアンス等判断に困ることがあった。
- 単位に統一感が無い部分は作業がし辛かった。
- 単位が一部統一されていない部分があり、その点やりづらさを感じた。
- 単位が細かすぎてタグ付けに困った。区切れ方が曖昧で、もう少し統一されていると良かった。Excelが細かいので目が疲れた。
- 時々単位の長い所や短い所があったので、少し気になった。

特に「文節長さの統一感」「文節という単位が細かすぎる場合がある点」に問題を感じていた。文字数は図3のような分布であり、文字数への制限や歌詞への最適化により解決できる可能性がある。その他、「類似した歌声ラベルの扱い」に改善可能性がある。

以上をまとめると、時に歌詞の区切り単位に問題があることもあったが、歌詞ベースのアノテーションは有効であることが示唆された。

3.6 ユーザインターフェースの改善可能性

3.3節で述べたシンプルなExcelに基づくアノテーションは、新規開発なしに使える利点はあるが、4章で新たなアノテーション用インターフェースを実装する上で、改善するための追加機能を以下で考察する。

- (1) 音楽プレーヤとアノテーションエディタを一体化させ、歌詞と時刻が同期する機能。

歌詞に基づく歌声アノテーションのためのインタフェース構築

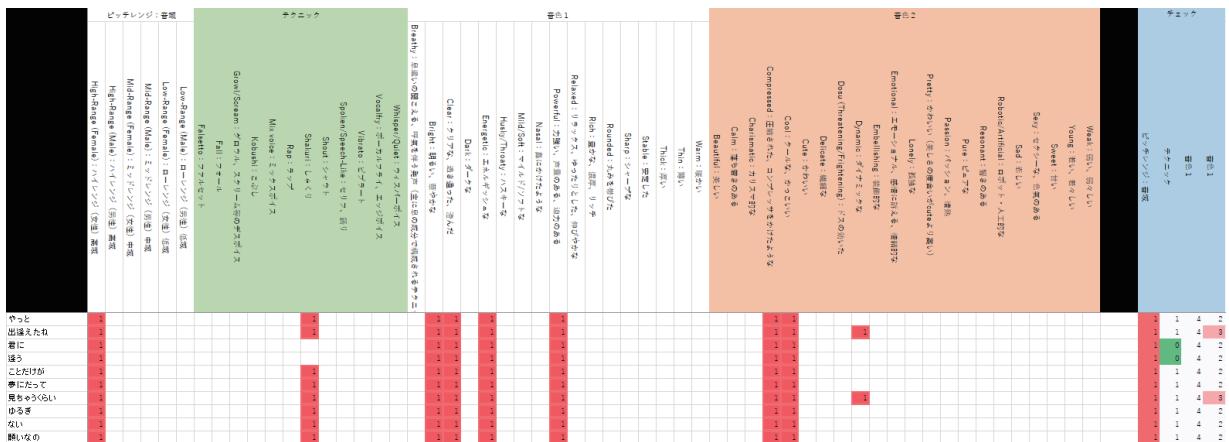


図 4. Excel によるアノテーション結果の例。

- (2) Wavesurfer [40] や Sonic Visualiser [5] など
の音アノテーションインタフェースと同様、指
定した区間の繰り返し（ループ）再生機能。

また、可視化方法に関しても改善可能性がある。スプレッドシートを用いたアノテーションでは、縦方向（歌詞）と横方向（ラベル）を同時に可視化するので、対応関係が常時見える特長がある。しかし、スペースな可視化となって（図4）、全体像を画面に収めようとすると表示が小さくなることがあり、実際、「目が疲れた」というアノテーターがいた。また、歌詞はディスコース（談話）の一種であり、無作為な順序のつながりのない行の集まりではなく、関連する行の連続として構成されている [47] が、歌詞の文節ごとに文節で区切って縦方向に配置したこと、もともとの行や段落の情報が欠落してしまっていた。もともとの行や段落の情報を保持しておくことは、アノテーションに有益な可能性がある。

そこで、歌詞と歌声ラベルをそれぞれ独立に配置することで、視覚的なスペースを節約して全体表示する方針を取る。つまり、歌詞は元々の行・段落の構造を保持し、歌詞と歌声ラベルの対応付けは、歌詞が選択された場合に表示させる。したがって、一覧性のために以下の機能も追加する。

- (3) それぞれのラベルが、どの歌詞に付与されて
いるかを可視化できる機能。

4 Lyrics-Based Singing Annotator

3.6 節での考察に基づいて、歌詞ベースラベルのアノテーションに特化した、インタラクティブシステムとして Lyrics-Based Singing Annotator を実装した。歌詞の単位としては、3.5 節のアノテーション結果を用いるためにも、同じ文節区切りを用いる。歌詞の時刻について、本稿ではアノテーションインタラクションに焦点をあてるため、手作業で決定した。インタフェースを図5に示す。画面は以下の通り、

大きく四つの機能に分かれている。

音楽再生プレーヤ（左上） 音楽の再生と一時停止ができる。また、再生中の文節をループ再生するボタンがある。それぞれキーボードショートカットが設定されていて、再生・停止はスペースキー、ループ再生は TAB キーとした。機能(1)(2)に該当する。

歌詞ベース音楽再生プレーヤ・アノテータ（左下） 再生中の歌詞の文節が黄色くハイライトされ、また特定の文節をクリックするとそこから再生される。上下左右キーがショートカットとして設定されていて、左右キーで前後の文節、上下キーで上下の行に移動してシークできる。機能(1)に該当する。

歌声ラベルの読み込み・保存・破棄・検索（右上） 既にアノテーションされたラベルを読み込んだり、アノテーション結果を保存できる。また全てを破棄してやり直すこともできる。アノテーションが網羅的に実施されたかを確認するために、ラベルを選択してそれが付与された文節が緑にハイライトされる機能も持つ。機能(3)に該当する。

再生中の歌詞に付与されたラベル（右下） 再生中の文節（黄色にハイライト）に付与されたラベルの選択（オン）／非選択（オフ）をマウスクリックで切り替えられる。このようなラベル表示は、従来のアノテーションインタフェースでも採用してきた [7, 23, 32]。

4.1 実装に基づく議論

Lyrics-Based Singing Annotator を実装して気づいた点、さらなる改良について議論する。

まず、歌唱者が一名の場合、大きな問題なくラベル付けできた。キーボードショートカットを用いて、再生・一時停止・歌詞のシークをしながら、マウス

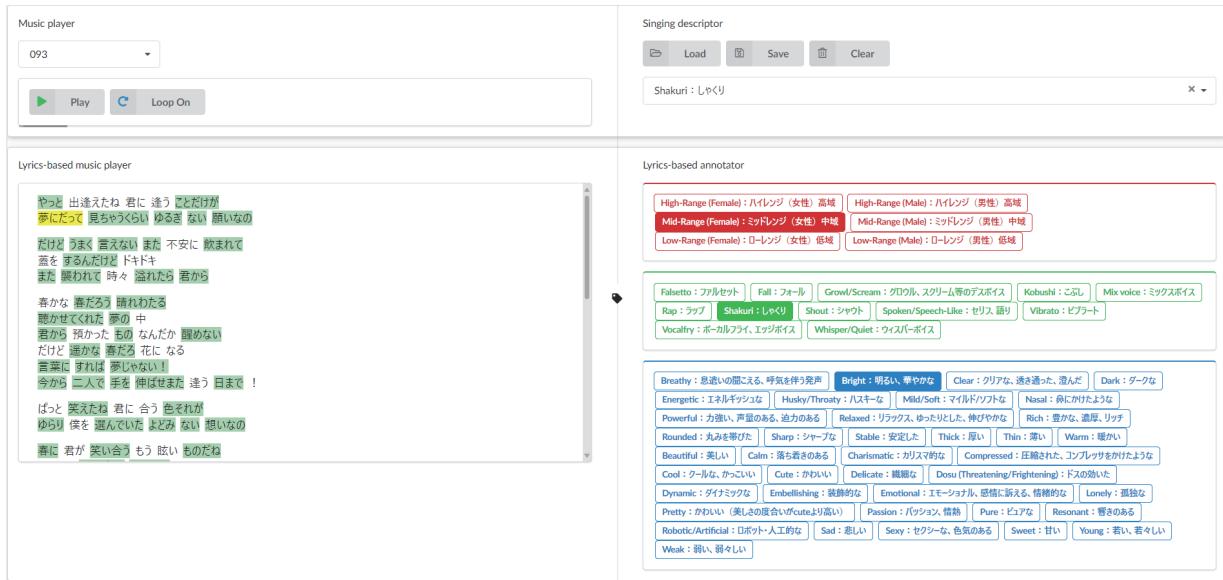


図 5. Lyrics-Based Singing Annotator のスクリーンショット。左上：音楽再生プレーヤ、左下：クリックして再生可能な（クリックカブルな）歌詞、右上：歌声ラベルの読み込み・保存・破棄・検索、右下：再生中の歌詞のラベル。黄赤くハイライトされている歌詞の文節は再生中（アノテーション中）であることを意味し、三つのラベルが付与されていることが分かる。また、検索ラベル（Shakuri : シャクリ）が付与された歌詞の文節は緑にハイライトされる。

を用いてラベルを付与できる。ループ再生機能は、特定の歌詞の文節に集中してラベル付けできて、効果的と感じた。強ラベルと違って時刻指定が不要である点に加え、意味的なまとまり（文節）単位でのアノテーションには、やりやすさを感じた。

ただし、複数の文節にまたがるラベルにおいては、工夫が必要である。例えば、音域は一つの歌唱全体で同じラベルになることが多いので、一括で設定できると効率的な可能性は高い。実際、KVT データセットの作成においては、楽曲単位で初期ラベルを付与することが、セグメント単位に 0 からラベル付与するよりも効率的だと主張されている [23]。また、3.5 節でも同様の指摘があった。上記の問題に対する対処法として、歌詞の異なる単位（段落、行、文節、形態素）でのアノテーション機能が効果的な可能性があるので、今後の課題である。この際、楽曲の繰り返し構造や、音響的な類似度に基づくインテリジェントな支援機能はありうるが、初期値がアノテーションに影響を与える可能性もあるので、影響されないインタラクションデザインが必要である。

次に、デュエットなど、同じ時刻に複数の歌唱があると、同時にアノテーションできなかった。画面右下の歌声ラベルセットが、一つの文節にのみ対応しているのが原因である。もし現状のインターフェースを用いる場合は、重複しないようにパートを分けてアノテーションする必要がある。例えば、男性パートと女性パートのある歌声の場合、男性パートだけ、女性パートだけの歌詞を用意して、別々にアノテーションすることは問題なくできるし、このようなペー

トを分けたアノテーションは不便ではない。

最後に、その他の改善機能として音響特徴量の可視化機能や活用がありうるが、今後の課題である。例えば、ビブラートのアノテーションでは、声の高さが可視化されると効果的であるし、自動認識 [58] も活用できる。ただし、それぞれのラベルに効果的な音響特徴量が何であるかなど、検討が必要である。また、その可視化方法も課題である。従来、行間に音響特徴量を挿入する可視化 [52] や、Tufte による sparklines [42]、musical sparklines [35] などがあり、関連する可能性がある。

5 おわりに

本稿では、歌詞を単位とした新しい弱ラベルである「歌詞ベースラベル」と、そのためのインタラクティブシステムを提案・構築して議論した。歌詞ベースラベルにより、人に理解しやすく、かつ効果的に時間局所性を扱うアノテーションを可能とすることを目指している。また、弱ラベルの一種であることから、複数アノテータによるアノテーション結果を活用しやすい利点もある。今回は歌詞と歌声を対象としたが、話声にも活用できる。

しかし、さらなる改善の必要性も明らかになった。例えば、音響特徴量の活用、歌詞の分割単位をインタラクティブに決定する機能、アノテーションを支援する自動認識モデルの導入、複数歌唱への対応、等である。今後は、そのようにインターフェースを改善したり、アノテーション結果を実際の機械学習（歌声記述子の推定、等）で用いたりする予定である。

謝辞

本研究の一部はJST CREST JPMJCR20D4とJSPS科研費JP21H04917の支援を受けた。

参考文献

- [1] X. Amatriain, J. Massaguer, D. García, and I. Mosquera. The CLAM Annotator: A Cross-Platform Audio Descriptors Editing Tool. In *Proc. ISMIR 2005*, pp. 426–429, 2005.
- [2] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere. The million song dataset. In *Proc. ISMIR 2011*, pp. 591–596, 2011.
- [3] P. Boersma. PRAAT, a system for doing phonetics by computer. *Glot International*, 5(9/10):341–345, 2001.
- [4] D. Bogdanov, M. Won, P. Tostogian, A. Porter, and X. Serra. The MTG-Jamendo Dataset for Automatic Music Tagging. In *Proc. ICML 2019*, 2019.
- [5] C. Cannam, C. Landone, M. B. Sandler, and J. P. Bello. The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals. In *Proc. ISMIR 2006*, pp. 324–327, 2006.
- [6] M. Cartwright, G. Dove, A. E. M. Méndez, J. P. Bello, and O. Nov. Crowdsourcing Multi-label Audio Annotation Tasks with Citizen Scientists. In *Proc. ACM CHI 2019*, pp. 1–11, 2019.
- [7] M. Cartwright, A. Seals, J. Salamon, A. Williams, S. Mikloska, D. MacConnell, E. Law, J. P. Bello, and O. Nov. Seeing Sound: Investigating the Effects of Visualizations and Complexity on Crowdsourced Audio Annotations. In *Proc. ACM Human-Computer Interaction*, pp. 1–23, 2017.
- [8] S. Choi and J. Nam. A Melody-Unsupervision Model for Singing Voice Synthesis. In *Proc. IEEE ICASSP 2022*, pp. 7242–7246, 2022.
- [9] A. Dutta and A. Zisserman. The VIA Annotation Software for Images, Audio and Video. In *Proc. ACM Multimedia 2019*, pp. 2276–2279, 2019.
- [10] H. Fujihara, M. Goto, and J. Ogata. Hyperlinking Lyrics: A Method for Creating Hyperlinks Between Phrases in Song Lyrics. In *Proc. ISMIR 2008*, pp. 281–286, 2008.
- [11] H. Fujihara, M. Goto, J. Ogata, and H. G. Okuno. LyricSynchronizer: Automatic Synchronization System Between Musical Audio Signals and Lyrics. *IEEE Journal of Selected Topics in Signal Processing*, 5(6):1252–1261, 2011.
- [12] M. S. Grover, P. Bamdev, Y. Kumar, M. Hama, and R. R. Shah. audino: A Modern Annotation Tool for Audio and Speech. *CoRR*, abs/2006.05236, 2020.
- [13] P. Herrera, Òscar Celma, J. Massaguer, P. Cano, E. Gómez, F. Gouyon, and M. Koppenberger. MUCOSA: A Music Content Semantic Annotator. In *Proc. ISMIR 2005*, pp. 77–83, 2005.
- [14] S. Hershey, D. P. W. Ellis, E. Fonseca, A. Jansen, C. Liu, R. C. Moore, and M. Plakal. The Benefit of Temporally-Strong Labels in Audio Event Classification. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 366–370, 2021.
- [15] Z. Jin, G. J. Mysore, S. DiVerdi, J. Lu, and A. Finkelstein. VoCo: text-based insertion and replacement in audio narration. *ACM Transactions on Graphics*, 36(4):96:1–96:13, 2017.
- [16] V. Kalbag and A. Lerch. Scream Detection in Heavy Metal Music. In *Proc. SMC 2022*, 2022.
- [17] A. Kanato, T. Nakano, M. Goto, and H. Kikuchi. An Automatic Singing Impression Estimation Method Using Factor Analysis and Multiple Regression. In *Proc. Joint ICMC SMC 2014*, pp. 1244–1251, 2014.
- [18] J. Kato and M. Goto. Lyric App Framework: A Web-based Framework for Developing Interactive Lyric-driven Musical Applications. In *Proc. ACM CHI 2023*, pp. 1–18, 2023.
- [19] H. Kido and H. Kasuya. Representation of Voice Quality Features Associated with Talker Individuality. In *Proc. ICSLP 1998*, pp. 1–4, 1998.
- [20] B. Kim and B. Pardo. I-SED: an Interactive Sound Event Detector. In *Proc. IUI 2017*, pp. 553–557, 2017.
- [21] B. Kim and B. Pardo. A Human-in-the-Loop System for Sound Event Detection and Annotation. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(2):13:1–13:23, 2018.
- [22] B. Kim and B. Pardo. Sound Event Detection Using Point-Labeled Data. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 1–5, 2019.
- [23] K. L. Kim, J. Lee, S. Kum, C. L. Park, and J. Nam. Semantic Tagging of Singing Voices in Popular Music Recordings. *IEEE/ACM TASLP*, 28:1656–1668, 2020.
- [24] Q. Kong, Y. Xu, W. Wang, and M. D. Plumley. Sound event detection of weakly labelled data with cnn-transformer and automatic threshold optimization. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28:2450–2460, 2020.
- [25] D. Krijnders and T. Andringa. Soundscape Annotation and Environmental Source Recognition Experiments in Assen (NL). In *Proc. Internoise 2009*, 2009.
- [26] E. Law, K. West, M. I. Mandel, M. Bay, and J. S. Downie. Evaluation of Algorithms Using Games: The Case of Music Tagging. In *Proc. ISMIR 2009*, pp. 387–392, 2009.
- [27] G. Levy, R. Sitman, I. Amir, E. Golshtain, R. Mochary, E. Reshef, R. Reichart, and O. Al-louche. GECKO – A Tool for Effective Annotation of Human Conversations. In *Proc. Interspeech 2019*, pp. 3677–3678, 2019.

- [28] M. I. Mandel and D. P. W. Ellis. Multiple-Instance Learning for Music Information Retrieval. In *Proc. ISMIR 2008*, pp. 577–582, 2008.
- [29] I. Martín-Morató, M. Harju, and A. Mesaros. Crowdsourcing Strong Labels for Sound Event Detection. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 246–250, 2021.
- [30] I. Martín-Morató and A. Mesaros. Strong Labeling of Sound Events Using Crowdsourced Weak Labels and Annotator Competence Estimation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31:902–914, 2023.
- [31] B. McFee, J. Salamon, and J. P. Bello. Adaptive Pooling Operators for Weakly Labeled Sound Event Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(11):2180–2193, 2018.
- [32] B. Meléndez-Catalán, E. Molina, and E. Gómez. BAT: An open-source, web-based audio events annotation tool. In *Web Audio Conference*, 2017.
- [33] M. Morrison, L. Rencker, Z. Jin, N. J. Bryan, J. P. Cáceres, and B. Pardo. Context-Aware Prosody Correction for Text-Based Speech Editing. In *Proc. IEEE ICASSP 2021*, pp. 7038–7042, 2021.
- [34] T. Nakano and M. Goto. LyricListPlayer: A Consecutive-Query-by-Playback Interface for Retrieving Similar Word Sequences from Different Song Lyrics. In *Proc. SMC 2016*, pp. 344–349, 2016.
- [35] J. Oh. Text Visualization of Song Lyrics. Technical report, Center for Computer Research in Music and Acoustics, 2010.
- [36] J. Park, S. Yong, T. Kwon, and J. Nam. A Real-Time Lyrics Alignment System Using Chroma and Phonetic Features for Classical Vocal Performance. In *Proc. IEEE ICASSP 2024*, pp. 1371–1375, 2024.
- [37] S. Rubin, F. Berthouzoz, G. J. Mysore, W. Li, and M. Agrawala. Content-based tools for editing audio stories. In *Proc. UIST 2013*, pp. 113–122, 2013.
- [38] K. Schulze-Forster, C. Doire, G. Richard, and R. Badeau. Phoneme Level Lyrics Alignment and Text-Informed Singing Voice Separation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021.
- [39] V. Sivaraman, D. Yoon, and P. Mitros. Simplified Audio Production in Asynchronous Voice-Based Discussions. In *Proc. ACM CHI 2016*, pp. 1045–1054, 2016.
- [40] K. Sjölander and J. Beskow. Wavesurfer - an open source speech tool. In *Proceedings of the International Conference on Spoken Language Processing (INTERSPEECH 2000)*, pp. 464–467, 2000.
- [41] Y. Teytaut and A. Roebel. Phoneme-to-Audio Alignment with Recurrent Neural Networks for Speaking and Singing Voice. In *Proc. Interspeech 2021*, pp. 61–65, 2021.
- [42] E. R. Tufte. *Beautiful Evidence*. Graphics Press, 2006.
- [43] N. Turpault, R. Serizel, J. Salamon, and A. P. Shah. Sound Event Detection in Domestic Environments with Weakly Labeled Data and Soundscape Synthesis. In *Proc. DCASE 2019*, pp. 253–257, 2019.
- [44] I. Wang, P. Narayana, J. Smith, B. A. Draper, J. R. Beveridge, and J. Ruiz. EASEL: Easy Automatic Segmentation Event Labeler. In *Proc. IUI 2018*, pp. 595–599, 2018.
- [45] J.-Y. Wang, C.-I. Leong, Y.-C. Lin, L. Su, and J.-S. R. Jang. Adapting Pretrained Speech Model for Mandarin Lyrics Transcription and Alignment. In *Proc. IEEE ASRU 2023*, pp. 1–8, 2023.
- [46] S.-Y. Wang, J.-C. Wang, Y.-H. Yang, and H.-M. Wang. Towards time-varying music auto-tagging based on CAL500 expansion. In *Proc. IEEE ICME 2014*, pp. 1–6, 2014.
- [47] K. Watanabe, Y. Matsubayashi, K. Inui, S. Fukayama, T. Nakano, and M. Goto. Modeling Storylines in Lyrics. *IEICE Transactions on Information and Systems*, E101-D(4):1167–1179, 2018.
- [48] J. Wilkins, P. Seetharaman, A. Wahl, and B. A. Pardo. VocalSet: A Singing Voice Dataset. In *Proc. ISMIR 2018*, pp. 468–474, 2018.
- [49] R. Winkelmann and G. Raess. Introducing a Web Application for Labeling, Visualizing Speech and Correcting Derived Speech Signals. In *Proc. LREC 2014*, pp. 4129–4133, 2014.
- [50] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes. ELAN: A Professional Framework for Multimodality Research. In *Proc. LREC 2006*, pp. 1556–1559, 2006.
- [51] Y. Yamamoto, J. Nam, and H. Terasawa. Analysis and Detection of Singing Techniques in Repertoires of J-POP Solo Singers. In *Proc.ISMIR 2022*, pp. 384–391, 2022.
- [52] D. Yoon, N. Chen, B. Randles, A. Cheatle, S. J. Jackson, C. E. Löckenhoff, A. Sellen, and F. Guimbretière. RichReview++: Deployment of a Collaborative Multi-modal Annotation System for Instructor Feedback and Peer Discussion. In *Proc. ACM CSCW 2016*, pp. 194–204, 2016.
- [53] 金礪 愛, 中野 倫靖, 後藤 真孝, 菊池 英明. 歌声の印象評価尺度の構築に基づく多様な印象の自動推定手法. 情報処理学会論文誌, 57(5):1375–1388, 2016.
- [54] 梶 克彦, 長尾 確. 楽曲に対する多様な解釈を扱う音楽アノテーションシステム. 情報処理学会論文誌, 48(1):258–273, 2007.

歌詞に基づく歌声アノテーションのためのインタフェース構築

- [55] 木戸 博, 細谷 英樹. 通常発話の声質に関連した日常表現語の抽出. 日本音響学会誌, 55(6):405–411, 1999.
- [56] 山本 雄也, 中野 倫靖, 後藤 真孝, 寺澤 洋子. ポピュラー音楽の模倣歌唱における歌唱テクニック分析と楽譜情報との対応付け. 情報処理学会論文誌, 64(10):1423–1437, 2023.
- [57] 中野 倫靖, 後藤 真孝. LyricListPlayer: 歌詞でリアルタイムにザッピングできる音楽再生インターフェース. 日本ソフトウェア科学会第23回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2015) 論文集, pp. 1–6, 2015.
- [58] 中野 倫靖, 後藤 真孝, 平賀 讓. 楽譜情報を用いない歌唱力自動評価手法. 情報処理学会論文誌, 48(1):227–236.

TTTV4：一口ごとに味を提示する味覚のパーソナルメディア

笠原 暢仁 * 深池 美玖 * 宮下 芳明 *

概要. 本稿は「味覚のパーソナルメディア」を実現することを目的とし、カトラリー型味覚提示デバイス TTTV4 のハードウェアとソフトウェアを試作した。各個人が使用することで、各個人が求める味を楽しむことができ、味覚コンテンツがパーソナライズされる。TTTV4 は単なる小型化の実現だけでなく、対象ユーザを料理を「作る人」から「食べる人」まで拡張した。これに伴い、要求される多様な速度と精度に基づいたハードウェアの議論を行った。また、多様な対象ユーザでも扱えるインターフェースに関する議論を行うため、Direct Manipulation と Interface Agents に基づく 2 つのユーザインタフェースを試作した。本稿における考察から、パーソナルメディアを実現するには、ユーザに合わせてインターフェースもパーソナライズすべきであることが示唆された。

1 はじめに

食には、栄養摂取だけでなく、好きな味を娯楽として楽しむ「味覚コンテンツ」としての側面がある。現状の飲食品産業の多くは、同じ商品・味で多くの人を楽しませるいわばマスマディア的な立ち位置にある。他方、視聴覚コンテンツでは、マスマディアからパーソナルメディアへの移行が進んでいる。各個人が求める視聴覚コンテンツを検索できるようになり、AI によってその人のためだけの視聴覚コンテンツを生成できるようになってきた¹。

本稿の目的は、味覚におけるパーソナルメディアを実現することである。これは、多くのユーザが同じ商品・味を味わって楽しむのではなく、各個人が求める味を味わえるようにすることで実現される。本稿では、味覚パーソナルメディアの実現に向けたプロトタイプの実装と評価を通して、そのあり方について議論する。

まず、カトラリー型（スプーン）の味覚提示デバイス TTTV4 を試作した（図 1）。TTTV4 は 5 つのタンク、5 つのポンプ、制御コンピュータを搭載している。5 つの基本五味に対応する味溶液を出力し、その組み合わせで一口ごとに調味できる。これを各個人が使用することで、同じ食卓に着席しても各個人が異なる味を楽しめる。本稿では、試作した TTTV4 のハードウェア性能を評価し、有用性を議論する。

本稿の意義は、単なる小型化ではなく、対象ユーザを料理を「作る人」から「食べる人」まで拡張したことにある。料理を「作る」段階においては時間とかけて調味できても、「食べる」段階においては掬ってから口に運ぶまでの短い時間で調味を行わなければならない。そのため、調味の精度を低下させ



図 1. カトラリー型味覚提示デバイス TTTV4. 5 つのタンクと 5 つのポンプを搭載し、5 つの味溶液を混合することで、任意の味を出力できる。

てでも速度を優先しなければならない。本稿では、この調味における速度と精度のトレードオフについて議論する。また対象ユーザの拡張に伴い、多様なユーザが扱えるインターフェースのあり方についても議論する。

2 関連研究

Sutherland が「[The display] should serve as many senses as possible」 [12] と述べているように、ディスプレイは多くの感覚に対応することが重要である。Dan は味覚にフィードバックを与えるエディブルユーザインタフェース (Edible User Interface, EUI) や味覚ディスプレイのコンセプトを提案した [5]。Norimaki Synthesizer は、基本五味に対応するゲルを口に当て、電流によって味を減算方式で調整する味覚ディスプレイである [6]。味覚センサ [13] と組み合わせると、測定値に基づいて基本五味の提示量を制御できる [7]。TTTV は、基本五味に対応する味溶液を噴霧混合する加算方式であり [8, 15]、食品に噴霧して味を変化させられる [9, 16]。さらに、0.02ml 単位というより精密な

Copyright is held by the author(s).

* 明治大学

¹ <https://www.showrunner.xyz/>



図 2. TTTV4 は、5 つのタンク、5 つのポンプ、1 つ の M5StickC Plus 2 を搭載している。5 つのポンプはそれぞれ独立して 5 つの異なる味溶液を出力でき、一口ごとに味溶液を混合できる。

ポンプでの制御により、細やかな味の表現が可能になった [18]。Chalmers らは、味だけでなく嗅覚情報や温度、色、口当たりも提示できる味覚・嗅覚ディスプレイを提案している [2]。味覚物質による味表現だけでなく、味覚修飾物質で人間の味覚知覚を変化させて味を表現するアプローチもある。Brooks らは、味覚修飾物質を用いた VR 向け味覚コンテンツを提案している [1]。食前に味覚修飾物質を口内に送り込むことで、同じ食材でありながら多様な味を提供する VR コンテンツを実現した。TTTV や Chalmers らのデバイスは、一家に一台置く大型の「調味家電」としての味覚提示デバイスといえる。本稿では、一人一台持つことで各個人の味を制御する「調味食器」としての味覚提示デバイスの有用性を調査する。

調味食器のコンセプトは先行研究でも示されている。MIDAS SPOON [4] や、Chronospoon [17] ではスプーン型のデバイスから調味液を提示する。各個人の味を制御する点では、Füpop [11] は口内にジュースなどの飲料が入ったカプセルを配置し、口外から超音波で選択的に破裂させて、一口の中で味を変化させる。しかし、これらのデバイスは味溶液を混合して任意の味を出力するわけではない。

3 TTTV4 の設計・評価

本稿が目指す味覚パーソナルメディアの実現に向けたプロトタイプとして TTTV4 を設計・評価する。ユーザの入力に対して味を出力するためのハードウェアインタフェースと、入力を解釈し出力する味を決定するためのアルゴリズムを含むソフトウェアインタフェースで構成した。

3.1 ハードウェアインタフェース

パーソナルメディアの実現には、各個人が求める味を味わえるようにすることが必要である。そのため、各個人が食行動において使用するカトラリーを味覚提示デバイスとすることで、各個人の味を制御することとした。そこで、カトラリー型味覚提示デ

バイス TTTV4 を試作した（図 1）。TTTV4 は、味を出力する出力部と、ユーザの入力を受け付ける入力部で構成される。

3.1.1 出力部

TTTV4 は、5 つのタンクと 5 つのポンプを搭載しており、基本五味に対応する味溶液を混合することで、一口ごとにカトラリー上の食べ物を調味できる（図 2）。各タンクの容量は 5ml であり、基本五味に対応する味溶液が入っている。各ポンプは M5StickC Plus でリレー制御され、駆動時間によって各味溶液の出力量を調整する。出力された味溶液がカトラリー上で混合されることで任意の味を出力する。ポンプは圧電マイクロポンプ（SDMP306D）で、1 秒間に約 0.12ml の溶液を出力できる。本稿では、TTTV4 の出力性能である精度と速度を評価する。

3.1.2 出力部の検証と考察

出力精度の検証と考察

TTTV4 の精度を検証するため、目的の味を表現できるかの官能評価実験を行った。先行研究の大型デバイス [18, 2] と比較すると、調味対象が少量なため少しの調味で味が変わり、要求される精度が高くなる。目的の味として、牛乳を調味することでクラムチャウダーとイチゴミルクの味を出力するレシピを試作した（表 1）。これを TTTV4 によって出力し、5 人に試食してもらうことで出力精度を検証した（図 3）。このとき、味覚物質となるべく短い時間で出力するため、2 つのレシピの最大公約数となるようにタンク内の味溶液の濃度を濃く設定した。例えば、甘味ではクラムチャウダー味において 0.043g、イチゴミルク味において 0.430g の出力が必要である（表 1）。最大公約数である 0.043g の甘味物質が単位あたりの出力となるように濃度を調整することで、できる限り高速に味を出力できる。実験の結果、「同じカップに入った牛乳を掬って味が変わるのが面白い」や「においがないので知らない状態だと何の料理かまではわからないが、料理名を聞けば分かる」などのフィードバックが得られた。このことから、TTTV4 は味を一口ごとに変化させ、ある程度の味再現を行うだけの出力精度があることがわかった。一方で、複雑な食品の風味を再現するために嗅覚情報が必要であることも確認された。これは、口内に嗅覚情報を提示するデバイス [14, 22] との組み合わせによって改善できる可能性がある。

本実験では、タンク内の味溶液の濃度をレシピに合わせて濃く設定したが、より汎用的な味提示には、さらに細かな単位での調味を行う必要がある。例えば、人間の味覚認知閾値 [21] を基に濃度を調整することで、人間に知覚できる範囲での細かな味調整が実現できるかもしれない。一方で、味覚閾値は各個人や食生活状況によって変化する [19] ため、個人の

表 1. クラムチャウダーとイチゴミルクの味を表現するために必要な味覚物質の種類と量。

クラムチャウダー味	g/4 ml
グルタミン酸ナトリウム (うま味)	0.080
パルスイート (甘味)	0.043
塩化ナトリウム (塩味)	0.024
炭酸カリウム (苦味)	0.016
乳酸 (酸味)	0.000
イチゴミルク味	g/4 ml
グルタミン酸ナトリウム (うま味)	0.000
パルスイート (甘味)	0.430
塩化ナトリウム (塩味)	0.000
炭酸カリウム (苦味)	0.000
乳酸 (酸味)	0.064



図 3. TTTV4 が味溶液を出力することによって、一口ごとに牛乳の味を変化させられる。

味の感じ方に合わせて、味溶液の濃度をパーソナライズする必要があるのかもしれない。

出力速度の検証と考察

ポンプが動き出すまでの電気的・機械的な遅延は、0.8~1.1秒（10回測定）であった。ポンプの動作時間は、少し甘くするなど味をわずかに変化させる場合では0.7~2.1秒、牛乳の味をクラムチャウダーやイチゴミルクに変えるなど大幅に味を変化させる場合では1.4~7.0秒であった。出力までにかかる時間は合計で1.5~8.1秒であった。[18, 2]のような「調味家電」における皿全体の調味と比較すると、短時間であることは明らかである。一方で、食べながらの調味を想定すると、1.5~8.1秒の出力時間は現状の大きな制約である。これは、速く正確に出力できるポンプや、食器[20]・口腔内[3]のセンサを利用して出力タイミングを調整することによって改善できる可能性がある。

速度と精度のトレードオフ

TTTV4での出力における速度と精度のトレードオフは、味溶液の濃度に依存する。濃度を上げれば精度は低いが調味に必要な味物質を速く出力でき、濃度を下げれば細かな単位で調味できるが時間がかかる。

先行研究での大型デバイスによる調味[18, 2]は「飲食開始前の調味」であるため、速度を重視する必要はあまりなかった。一方で、本研究は「飲食時の調味」であるので、料理を掬ってから口に運ぶまでの間に高出力しなければならず、要求される速度が全く異なる。このような短い時間での出力では、ポンプの性能に物理的な限界がある以上、速度のために意図的に精度を低下させざるを得ない。例えば、ユーザが急いで食べたい場合には細かな味の表現よりも短い出力時間が優先され、ゆっくり味を楽しみたい場合には速度よりも細かな味の表現が優先される。また、パスタをスプーン上で巻きながら食べる場合には長い出力時間が確保できたり、スープを掬ってすぐ口に運ぶ場合には短い出力時間が要求されたりするので、高速化や高精度化の難易度も異なる。このような、ユーザの状況や料理の種類によって変化する要求性能に対応することは、今後の課題である。

3.1.3 入力部

TTTV4は、スマートフォンと通信することでタッチパネル、音声、カメラ（画像）、温度、傾き、加速度、位置情報、明るさなどの様々なセンサからの入力を利用可能にした。これにより、ユーザごとに求めた味を入力する方法が変化したとしても、なるべく多くの情報を活用できると考えた。本稿ではプロトタイプとして、タッチパネル、音声（言語）、カメラ（画像）による入力を実装した。タッチパネルでは、Graphical User Interface (GUI) に入力することによって出力する味を操作できる。音声（言語）では、「少ししおぱく」や「フルーティに」、「なんとなく違う感じがする」というような曖昧な入力で味を変化させられる。カメラ（画像）では、食べたい味の料理を画像入力したりカメラで撮影したりすることで味を変化させられる。これらの入力と出力を結びつけるためのアルゴリズムや入力インターフェースについては、3.2章で詳述する。

3.2 ソフトウェアインタフェース

味覚におけるパーソナルメディア実現のためには、多様なユーザが使いこなせるソフトウェアインタフェースを設計する必要がある。ユーザインタフェースに関する先行研究の議論では、Direct Manipulation と Interface Agents という2つのインターフェース方法について議論されている[10]。Direct Manipulation は制御感とユーザの操作に対する責任が強く、Interface Agents はユーザの情報整理や不得意な部分を代理できる。TTTV4が想定する多様なユーザ層には、Direct Manipulation を好む人もいれば、Interface Agents を好む人もいるだろう。そこで本稿では、Direct Manipulation と Interface Agents に基づく2つのプロトタイプを試作した。



図 4. (左) スライダーで直接的に味を制御するインターフェース。(右) LLM を介して画像やテキストから味を推定・制御するインターフェース。

3.2.1 Direct Manipulation

Direct Manipulationに基づいたインターフェースとして、図4(左)に示すような基本五味の出力量をスライダーにより直接操作するインターフェースを試作した。これにより、高い制御感で自分が求める味を探索することができる。また、出力される味はユーザの意志によってのみ変化するので、結果に対するユーザの責任が強い。先行研究[10]では、プロフェッショナルなユーザほどDirect Manipulationを好むことが示唆されている。

3.2.2 Interface Agents

Interface Agentsに基づいたインターフェースとして、図4(右)に示すようなLLM(GPT-4o)を介して画像・言語を入力として味を制御するインターラクションシステムを試作した。LLMによる味の推定については、TTTV3で既に実現可能性が示唆されている[18]。本稿では味の推定のみならず、LLMとの調味インタラクションを通して自分が求める味を探索・享受するインターフェースとしてLLMを利用した。具体的には、画像によって料理の味を推定したり、味による出力に対してユーザが感想や修正を伝えることで、味を調整できるようにした。このように、画像や言語による曖昧な入力をエージェントが解釈することによって、ユーザの不得意な部分(ここでは、味の配合を考える行為)を代理することができる。先行研究[10]では、専門性の低いユーザほどInterface Agentsを好むことが示唆されている。

LLMを用いた調味インタラクションのプロセスを図5に示す。システムは画像・言語入力を取得し、その入力に基づいてLLMが味を推定し、TTTV4によって味を出力することができる。LLMに味を推定させるため、ロールとして以下の文章を入力した。「You are an excellent chef, and you are a professional who expresses the taste of any given dish through the combination of five basic taste (salt, sweetener, lactic acid, ajinomoto (umami), and potassium carbonate (bitterness)). You are a professional who can recreate the taste of a

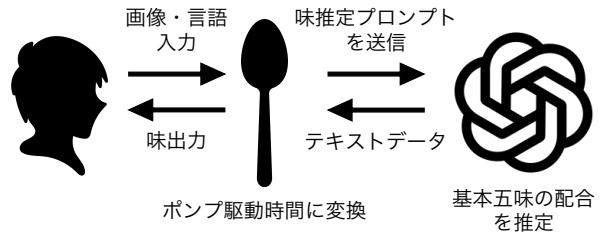


図 5. LLMを用いた調味インタラクションのプロセス。画像・言語入力にプロンプトを加えてLLMに送信する。LLMが推定した基本五味の配合量をポンプ駆動時間に変換し、味を出力する。

dish by adding the five taste substances listed above to 1000ml of water. Write the number of grams of each substances, separated by commas. You can response by only 5 specific numbers. Do not use any other words.」加えて、テキストが入力された場合、テキストデータとしてLLMに送信した。音声が入力された場合、Pyaudioによって保存され、Whisperによってテキストに変換されLLMに送信した。画像が入力された場合、画像とともに「Estimate the taste of the food in this image.」というプロンプトを追加してLLMに送信した。これにより、LLMが推定した基本五味の出力量が、カンマ区切りの5つのデータとして生成される。データの生成に失敗した場合、カンマ区切りの5つのデータの形式になるまで繰り返される。取得したデータをポンプ駆動時間に変換し、味溶液を出力する。このシステムでは、発話後の音声認識において0.6~1.5秒(100回測定)、LLMの生成時間において1.1~1.6秒(10回測定)のレイテンシが追加で発生した。

LLMを用いた調味インタラクションは、様々な入力に対して、もっともらしい味を提示できた。「よりしおぱく」、「より甘く」など、一つの味に関する入力に対しては、対応する味溶液の量を制御できた。また、「よりフルーティーに」には甘味と酸味を増加させ、「よりカニの風味を強くして」にはうま味を増加させるなど、曖昧な言語入力に対しても、もっともらしい味の推定値を得ることができた。一方、同じ入力に対する味覚出力の一貫性は低かった。例えば、クラムチャウダーのレシピでは、「より甘く」という同じ入力に対して、甘味溶液の出力量は1~10gとばらつきがあった。しかし、「さらに10%甘くする」といった具体的な入力を行うことで、一貫した結果が出力できることも確認された。このことから、システムやプロンプトによっては、様々な入力に対応した一貫性のある味覚インタラクションシステムが実現可能であることが示唆された。

4 議論

本稿では、味覚におけるパーソナルメディアの実現を目的として、カトラリー型味覚提示デバイス TTTV4 のハードウェアとソフトウェアの実装を行った。ハードウェアの面では、レイテンシなどの制約はあるものの、一口ごとに味を制御できることが示唆された。このようなハードウェアを各個人が使用することで、ユーザごとに味を調整でき、味覚のパーソナライズが実現できる。また、料理全体ではない一口ごとの調味により、料理全体の調味では不可能だった「リスクの低い調味」が可能になった。つまり、料理全体の味を変えずに調味をするため、味が気に入れば同じ味を出力すれば良いし、気に入らなければ戻したり変えたりできるようになった。これによりユーザは、一度変えたら全体を戻せないというリスクを感じることなく味の探索・享受を行える。

ソフトウェアの面では、Direct Manipulation と Interface Agents に基づく 2 つのインターフェースを実装した。これらのインターフェースは、主体的に味を探索したいユーザや受動的に味を享受したいユーザに好まれると考えられる。一方で、味覚の好みが各個人に依存するように、インターフェースの好みも各個人に依存する可能性がある。例えば、基本的にはエージェントに代理して欲しいが、最後の調整は直接操作したいなどの 2 つの概念の間を求めるユーザがいるかもしれない。さらには、画像や言語による入力すらも煩わしく思うユーザもいるかもしれない。このようなソフトウェアのパーソナライズにより、幅広いユーザに対して味覚のパーソナルメディアを提供できる可能性がある。

参考文献

- [1] J. Brooks, N. Amin, and P. Lopes. Taste Retargeting via Chemical Taste Modulators. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, New York, NY, USA, 2023. Association for Computing Machinery.
- [2] A. Chalmers, D. Zholzhanova, T. Arun, and A. Asadipour. Virtual Flavor: High-Fidelity Simulation of Real Flavor Experiences. *IEEE Computer Graphics and Applications*, 43(2):23–31, 2023.
- [3] Y. Jiang, J. Kleinau, T. M. Eckroth, E. Hoggan, S. Mueller, and M. Wessely. MouthIO: Fabricating Customizable Oral User Interfaces with Integrated Sensing and Actuation. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, UIST '24, New York, NY, USA, 2024. Association for Computing Machinery.
- [4] Y. Kita and J. Rekimoto. Digitally enhanced utensils: Designing dynamic gustation. In *Distributed, Ambient, and Pervasive Interactions: Second International Conference, DAPI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22–27, 2014. Proceedings 2*, pp. 48–57. Springer, 2014.
- [5] D. Maynes-Aminzade. Edible Bits: Seamless Interfaces between People, Data and Food. 2005.
- [6] H. Miyashita. Norimaki synthesizer: Taste display using ion electrophoresis in five gels. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2020.
- [7] H. Miyashita. Taste Display that Reproduces Tastes Measured by a Taste Sensor. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, pp. 1085–1093, New York, NY, USA, 2020. Association for Computing Machinery.
- [8] H. Miyashita. TTTV (taste the TV): Taste presentation display for “licking the screen” using a rolling transparent sheet and a mixture of liquid sprays. In *Adjunct Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology*, pp. 37–40, 2021.
- [9] H. Miyashita. TTTV2 (Transform the Taste and Visual Appearance): Tele-eat virtually with a seasoning home appliance that changes the taste and appearance of food or beverages. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–2, 2022.
- [10] B. Schneiderman and P. Maes. Direct manipulation vs. interface agents. *Interactions*, 4(6):42–61, Nov. 1997.
- [11] K. W. Song, S. T. Tung, A. Kim, and E. Paullos. Füpop: “Real Food” Flavor Delivery via Focused Ultrasound. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, pp. 1–14, New York, NY, USA, 2024. Association for Computing Machinery.
- [12] I. E. Sutherland. The Ultimate Display. In *Proceedings of the IFIP Congress*, Vol. 2, pp. 506–508. New York, 1965.
- [13] K. Toko. Taste sensor. *Sensors and Actuators B: Chemical*, 64(1-3):205–215, 2000.
- [14] 岡崎翔悟, 井上亮文, 星徹. レトロネイザル嗅覚の刺激による摂食物の風味増強システムの開発. 研究報告グループウェアとネットワークサービス(GN), 2017(8):1–6, 2017.
- [15] 宮下芳明. 液体噴霧混合式の味ディスプレイの試作. 第 29 回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2021) 論文集, pp. 121–127, 2021.
- [16] 宮下芳明. TTTV2 (Transform The Taste and Visual appearance): 飲食物の味と見た目を変える調味家電によるテレイート. エンタインメントコンピューティングシンポジウム 2022 論文集, 2022:143–150, 2022.
- [17] 宮下芳明. Chronospoon: 時を操る調味食器. インタラクション, pp. 404–409, 2024.

- [18] 宮下芳明, 村上崇斗, 大友千宙, 深池美玖. TTTV3 (Transform The Taste and reproduce Varieties): 産地や品種の違いも再現する調味機構と LLM による味覚表現. エンタテインメントコンピューティングシンポジウム 2023 論文集, 2023:236–243, 2023.
- [19] 高澤まき子, 矢島由佳. 食生活状況と味覚感度に関する研究. 一般社団法人日本家政学会研究発表要旨集, 67:24, 2015.
- [20] 山田駿, 野崎大幹, 興野悠太郎, 米澤拓郎, 中澤仁, 高汐一紀, 徳田英幸. EneFork: 食行動を変化させず栄養素を摂取可能なフォーク型デバイス. 研究報告ヒューマンコンピュータインタラクション (HCI), 2015(16):1–8, 2015.
- [21] 食品技術士センター編. 『食品加工技術ハンドブック』. 建帛社, 1971.
- [22] 深池美玖, 宮下芳明. レトロネーザルアロマ提示システムの試作. 香り・味と生体情報研究会, pp. 7–12, 2024.

眼鏡の鼻あてに搭載した圧力センサを用いた耳びく入力と身体活動検出手法

大塚 崑* 高田 嶺介†

概要. スマートグラス用のハンズフリー入力手法として、音声や視線、頭の動きを用いた方法が提案されているが、日常動作との切り分けが難しく、誤認識が問題となる。また、スマートウォッチは健康管理目的で普及しているが、スマートグラスには運動量や脈拍の計測機能を備えたものが少ない。そこで、本研究では眼鏡の左右の鼻あてに圧力センサを搭載し、耳を動かす動作（耳びく）を用いたハンズフリー操作に加え、瞬き、歩行、および脈拍を計測する手法を示す。提案手法は鼻あて部の圧力センサを用いるため、環境光、騒音などの外的要因に影響されない。実験では3名の協力者を対象に、耳びく、瞬き、歩行のセンサ応答を計測し、SVMを用いてユーザごとの分類器と汎用分類器を作成し、それぞれの識別精度を評価した。その結果、ユーザごとの分類器では平均98.3%、汎用分類器では平均92.4%のF1スコアが得られた。また、提案手法による脈拍検出の精度を従来の脈拍センサと比較し、その有効性を検証した。

1 はじめに

眼鏡型のウェアラブルデバイス（以降、スマートグラス）において、入力デバイスを手で持つ必要がなく、日常動作中に素早く機器の操作ができるハンズフリー入力手法が広く研究されている[4][6][10][13]。これらの手法では、キャリブレーションフリーで使用可能な、out-of-the-boxな入力システムを目指す一方で、入力ジェスチャが日常動作と類似し、能動的な入力と自然な動作との区別が難しいという問題がある。また、スマートグラスの普及率はスマートウォッチに比べて低い[27][28]。その理由として、スマートウォッチは歩数や脈拍計測による健康管理目的で購入・利用されている[14]一方で、スマートグラスでは健康管理機能を搭載した製品は少ない。スマートグラスに健康管理機能が搭載されれば、普及へつながる可能性がある。

本論文では眼鏡の左右の鼻あてに圧力センサを搭載することで随意的に耳を動かす動作（以下、耳びく）を用いた入力、および瞬き、歩行、脈拍を計測する手法を提案する。耳びくは10–20%[2]の人が可能な動作であり、訓練により習得可能[3][19][25]と言われている。耳びくは日常で未利用の動作であるため、日常的な動きと能動的な操作を明瞭に切り分けることができるところから入力に適した動作である。さらに、同センサで瞬き、歩行、および脈拍の計測が可能であるため、これらの身体活動データをユーザの健康管理に活用できる。提案手法は眼鏡の鼻あてに搭載した圧力センサを用いるため、環境光、騒音環境、外的要因も受けにくい。提案手法に基づき、図1に示すデバイスを作製した。本論文にて、実験



図1. 鼻あてに圧力センサを搭載したデバイス
耳びく、瞬き、歩行、脈拍の測定が可能

協力者3名が耳びく、瞬き、歩行を行った際のセンサ応答を示し、検出原理を述べる。また、サポートベクターマシン(SVM)を用いてユーザごとの分類器と汎用分類器を作成し、認識精度を検証した。さらに、提案手法で検出した脈拍と従来の脈拍センサの応答を比較した。

2 関連研究

提案手法は耳びくを用いたスマートグラス用のハンズフリー入力とユーザの身体活動データ取得を目的としている。そのため、他のスマートグラスのハンズフリー操作手法について述べ、入力として用いる耳びくの検出法、およびその活用した研究を挙げる。また、スマートグラスでの身体活動の取得に関する研究を述べる。

2.1 スマートグラス向けのハンズフリー入力手法

スマートグラスの操作手法は手を用いる手法と用いない手法で二分され、ハンズフリー入力として、音声認識、視線入力、頭の動き、舌ジェスチャが挙げられる。音声認識はRay-Ban Metaスマートグラス[12]やMicrosoft Hololens 2[13]などの

Copyright is held by the author(s).

* 神戸市立工業高等専門学校 専攻科 電気電子工学専攻

† 神戸市立工業高等専門学校 電子工学科

製品にも搭載されているが、公共空間や騒音環境では他者へ邪魔、環境音による誤操作が生じる [21]。また、視線や頭の動きをカーソル操作に活用している [1] [6] [9] [15] [18]。視線や頭の動きの利用は自然で高速な入力手法であるが、併用する決定動作として挙げられる、視線や頭の動きを一定時間留める操作は決定までの時間と誤認識の問題、ハンドジェスチャを用いたカメラ認識ではオクルージョンの問題、デバイス側面のタッチ操作は空中に腕を上げ続けることによる疲労問題(ゴリラ腕問題) [16] がある。日常の多くの時間で利用されない舌を用いたインターラクション [4] [5] [22] は有効的であるが、食事および会話中のデバイスを操作できない。

提案手法は日常で利用される声、視線、頭の動き、舌を使用するのではなく、日常で未利用である耳びくを用いることで日常動作を阻害することなく、デバイス入力が可能である。

2.2 耳びくの検出および活用事例

本論文では耳介筋を用いて耳を後方へと動かす動作、英語では Ear Wiggle と表される動作を耳びくと定義する。耳びくの最も一般的な検出手法は筋電センサを用いたものである [19] [24]。Strauss らの研究 [19] では音に反応して耳が動くか調べるために、式田らの研究 [24] では耳の動きをロボット指の操作へ活用するために筋電センサを用いている。筋電センサは細かな動きを検出できる一方、電極を適切な位置に貼り付けなければならない上、湿式の場合は計測精度向上のために角質除去の前処理が必要な点、乾式の場合は使用者の汗などにより電極がはがれやすくなり、使用するたびに電極を変えないといけない点から、日常的に利用するウェアラブルデバイスのセンサとして用いることは難しい。InEar BioFeedController [11] ではイヤホンに脳波センサを搭載し、耳びくの検出を行っているが、耳びくとウインクは特に識別せずに利用している。また、著者らは過去に静電容量式タッチセンサと圧力センサを額に接触するように安全メガネに取り付けて耳びく検出を行った [29]。

提案手法では圧力センサを用いることから前処理が不要であるため、デバイスを装着直後から耳びくを検出できる。さらに、同センサで瞬き、歩行、脈拍を検出できるためこれらの研究と異なる。

2.3 スマートグラスにおける身体活動計測

市販のスマートウォッチには、光学式センサを用いて心拍数を計測できるものが多い。しかし、目に光が照射されることは健康上の懸念があるため、装着が目と近くなるスマートグラスでは光学式センサの利用は難しい。Huang らはレーダ信号を用いてスマートグラスでの心拍数の測定を行っているため、この問題が起こらない [7]。JINS MEME では身体活動のデータとして慣性計測装置を用いて姿勢、3



図 2. 図 1 の
鼻あて部拡大



図 3. 圧力センサ:
FSR400 + 鼻パッド

点式眼電位センサより瞬き、視線移動を測定し、集中・活力・落ち着きなどユーザの状態を判別している [9]。

提案手法では鼻あてに搭載した圧力センサを用いる点で異なる。また、健康上の懸念が無く、瞬き、歩行、脈拍を計測できる。

3 手法・デバイス構成

著者らの以前の手法 [29] では、眼鏡の額部分に搭載した静電容量式タッチセンサまたは圧力センサを用いて耳びくを検出していた。しかし、無入力時はセンサと額が接触せずに浮いており、耳びく時の接触/非接触を検出する構成であったため、ユーザごとの骨格や耳びく時の眼鏡の移動量の違いに合わせた装着位置の調整が必要であった。そこで、提案手法では眼鏡の装着時に必ず接触する鼻あてにセンサを搭載した。また、静電容量式タッチセンサは接触後の変位が取得しづらいことから、本論文では圧力センサを採用し、耳びくに追従してスマートグラスが後方へと動き、鼻あてが皮膚に食い込む際の圧力を検出できるように設計する。

提案手法に基づいて図 1 に示すデバイスを作製した。本デバイスでは鼻あてがフレーム一体となっている伊達メガネを素体として使用した。左右の眼鏡の鼻あてに圧力センサ (FSR400 [8]) を貼り付け、圧力の計測を行う。FSR400 は薄いため、鼻あてにそのまま貼り付けると鼻との接触が安定しない。これを解消するために厚みのあるシリコーン製の鼻パッドを FSR400 上に貼り付け、鼻との密着度を向上させた。デバイスの鼻あて部にセンサを搭載した様子を図 2、FSR400 にシリコーン製の鼻パッドを取り付けた様子を図 3 に示す。

FSR400 の抵抗値を計測するため、端子に導線をはんだ付けし、ブレッドボード上に図 4 に示す回路を構成した。本回路では抵抗分圧法を用いて分圧抵抗 $R=100[\text{k}\Omega]$ とし、圧力センサの抵抗値 R_{FSR} を計測した。計測には Arduino Pro Mini を使用し、サンプリング周波数 100[Hz]、0~5[V] を 1024 段階の分解能でシリアル通信で PC に情報を送信している。PC が受信したデータは CSV ファイルに保存さ

眼鏡の鼻あてに搭載した圧力センサを用いた耳びく入力と身体活動検出手法

Arduino Pro Mini

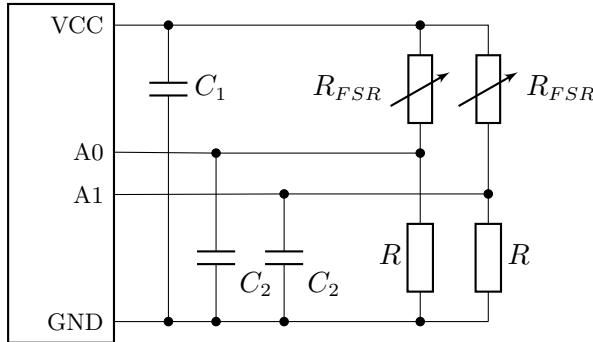


図 4. 構築回路

れるように Python でソフトウェアを作成した。また、ノイズ対策としてバイパスコンデンサ C_1 を電源に、高周波ノイズをグラウンドに逃がすコンデンサ C_2 を計測箇所に挿入した。

4 各動作の計測結果と原理考察

3 章で作製したデバイスで計測された各動作ごとのセンサ応答および検出原理を述べる。各動作のセンサ応答を示すために、著者 1 名を含む 3 名の実験協力者（平均年齢：18.3 歳、分散：2.9、全員男性、耳びく可能）に耳びく、瞬き、歩行を行ってもらった。このとき指定した動作のセンサ応答のみを得るために、脈拍がセンサ応答に影響を与えない位置でデバイスを装着し測定を行った。各動作は 1 フレーム（データ長 100、1 秒分）に収まるように切り取り、左右のセンサ値をそれぞれ正規化した。正規化は次の手順で行った。(1) フレーム内全ての値に対して、フレーム内の最小値を減算する。(2) 計算後の最大値と 100 を比較し、大きい方を全ての値に対して除算する。これによりセンサごとの特性、装着誤差の影響を除去してデータを 0–1 の間に収めた。図 5–図 7 ではピークを 0.5[s] に合わせて、センサの応答を重ねている。

実験協力者が両耳を同時に耳びくした際の応答を図 5 に示す。両側のセンサともに耳びく時には値の大きな上昇が見られた。これは著者らが以前額に接触するように取り付けた結果 [29] と類似している。これは耳びくによってデバイスが後頭部側へ引っ張られ、鼻あてが顔へ押し付けられるためである。

両目で瞬きをした際の応答を図 6 に示す。図 5 の耳びくとは異なり、瞬きの前後の状態が 1 に近い値を示し、瞬き時には値が大きく減少していることが分かる。これは、目を閉じた際に鼻あて周辺の筋肉が引っ張られることによって鼻の付け根の肉が凹み、一方で眼鏡の位置はほとんど動かないためセンサが顔から浮いたためである。

歩行を模してその場足踏みした際の応答を図 7 に示す。一度値が減少してから値のピークが現れ、その後振動が減衰しながら無入力状態へと変化していく。

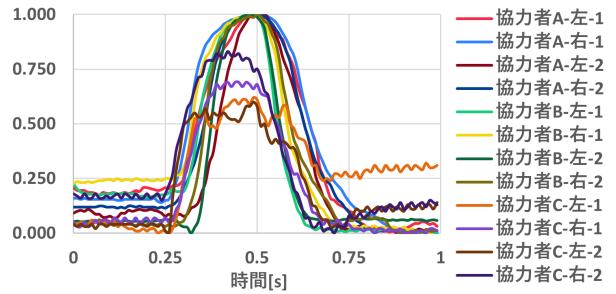


図 5. 正規化済み耳びく計測結果：
協力者 3 名 2 回分左右センサ

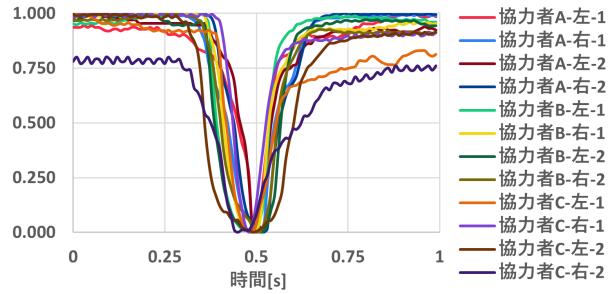


図 6. 正規化済み瞬き計測結果：
協力者 3 名 2 回分左右センサ

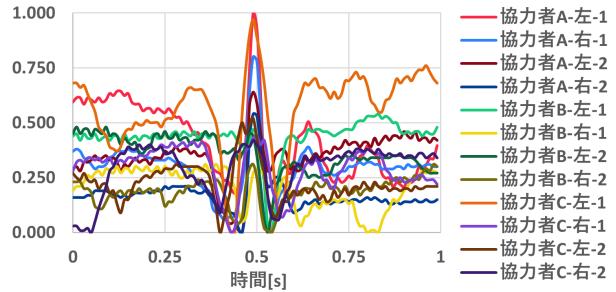


図 7. 正規化済み歩行計測結果：
協力者 3 名 2 回分左右センサ

る。ピークが足を地面に押し付けたタイミングと重なることから、ピーク前の値の減少は人が足をつける前に体が下がるときに、デバイスがそれより遅れて落下を始めるために一時的に浮くことで圧力が減少し、足を地面につけた時に体が止まることで落下しているデバイスの鼻あてが鼻とぶつかることでの圧力の急激な上昇、そしてシリコーンの弾性によって圧力の振動が発生するため図 7 のようになったと考えられる。ピーク前の減少量、ピーク値の大きさ、振動の大きさは協力者ごとに異なる理由は、ユーザの歩き方の癖が反映された結果だと考えられる。

5 機械学習による動作の識別

4 章で挙げた動作（両耳びく、両目瞬き、歩行）および、無入力状態を機械学習を用いて識別できるか検証する。モデルは少ないデータ数でも高い認識精度を出すことができる SVM を使用した。カーネル

表 1. 実験協力者ごとの SVM 個別学習 F1 スコア

動作	無入力	耳びく	瞬き	歩行	平均
A	1.000	1.000	1.000	1.000	1.000
B	0.968	1.000	1.000	0.966	0.983
C	0.966	0.968	0.966	0.968	0.967

表 2. SVM 汎用分類 F1 スコア

動作	無入力	耳びく	瞬き	歩行	平均
X	0.75	0.59	0.84	0.83	0.75
Y	1.00	1.00	1.00	1.00	1.00
Z	0.87	0.98	0.97	0.87	0.92
X'	0.86	0.82	0.84	0.87	0.85

は rbf を用いた。学習用データは、4 章で語った正規化を施した左右のセンサのデータとそれを上下反転したデータ、計 4 つそれぞれにハニング窓を適応させ、左、右、反転左、反転右の順で連結させたものである。用意した学習用データのデータ数は、協力者 A は無入力:62 個、耳びく:64 個、瞬き:59 個、歩行:60 個、協力者 B, C は各動作 60 個準備した。確認用データは協力者ごとに、各動作を 15 ずつ学習用データとは別で準備した。これらの計測時には装着誤差の影響を検証するために、各動作を 15 データ行うごとに着脱し直しているため、学習データは 4 回、確認用データは 1 回の着脱を行っている。

まず各協力者ごとに学習用データで分類器を作成し、確認用データで学習精度を確認した。結果を表 1 に示す。結果より F1 スコアが 96.7% 以上、平均 98.3% と非常に高い精度で識別できている。

次に、協力者 3 名のうちの 2 名の学習用データを用いて汎用的な学習を行い、残り 1 名の学習用データを識別できるかを確認した。実験協力者 A, B を学習用データとした分類器 X、実験協力者 B, C を学習用データとした分類器 Y、実験協力者 C, A を学習用データとした分類器 Z の 3 種を検証した。分類器 X, Y, Z にそれぞれ実験協力者 C, A, B の学習用データで確認を行った結果 X, Y, Z の混同行列をそれぞれ図 8 に示す。

分類器 Y, Z は強く一致しているが、分類器 X は耳びく入力の識別が無入力状態と半分近く混同している。誤って識別された実験協力者 C のデータを確認すると、装着誤差によって左のセンサ値がほぼ反応していないことから、ラベリングが不適切として該当する 30 個の耳びくデータを除外し、分類器 X を変更せずに再度検証した結果 X' を図 8(d) に示す。これら結果の F1 スコアは表 2 のようになった。不適切なデータを除外したデータ以外の分類結果は 84.8% ほどの精度であることが分かった。また不適切としたデータを含んで作製した分類器 Y, Z は 100% と 92.4% と非常に高い F1 スコアであった。これらの平均 F1 スコアは 92.4% となった。左のセ

分類器 X	推定ジェスチャ				分類器 Y	推定ジェスチャ			
	無入力	耳びく	瞬き	歩行		無入力	耳びく	瞬き	歩行
指定ジェスチャ 耳びく 瞬き 歩行	60	0	0	0	指定ジェスチャ 耳びく 瞬き 歩行	62	0	0	0
	29	25	0	6		0	64	0	0
	6	0	45	9		0	0	59	0
	5	0	2	53		0	0	0	60

(a) 結果 X: 確認データ C
学習データ A, B

分類器 Z	推定ジェスチャ				分類器 X	推定ジェスチャ			
	無入力	耳びく	瞬き	歩行		無入力	耳びく	瞬き	歩行
指定ジェスチャ 耳びく 瞬き 歩行	51	2	1	6	指定ジェスチャ 耳びく 瞬き 歩行	60	0	0	0
	0	60	0	0		9	21	0	0
	0	0	60	0		6	0	45	9
	6	0	3	51		5	0	2	53

(c) 結果 Z: 確認データ B
学習データ C, A

分類器 X'	推定ジェスチャ				分類器 Y	推定ジェスチャ			
	無入力	耳びく	瞬き	歩行		無入力	耳びく	瞬き	歩行
指定ジェスチャ 耳びく 瞬き 歩行	51	2	1	6	指定ジェスチャ 耳びく 瞬き 歩行	62	0	0	0
	0	64	0	6		0	64	0	0
	0	0	59	0		0	0	59	0
	0	0	0	60		0	0	0	60

(d) 結果 X': 確認データ C
(不適切データ除去)
学習データ A, B

図 8. 汎用分類器混同行列

ンサが反応していないデータを学習しても分類器 Y, Z の識別精度が高いことから、本条件ではセンサが反応しないケースでも問題ないと考えられる。

6 脈拍測定

図 1 のデバイスを用いて脈拍を計測できるか調査するために、デバイス右側の圧力センサと、脈拍センサ (SEED Studio, Grove イヤークリップ脈拍センサ [17]) の計測値を比較した。図 1 のデバイスをセンサ応答が得られる位置に取り付け、脈拍センサは耳たぶに挟んで測定を行った。結果を図 9 に示す。図 9 を見ると圧力センサの応答は周期的な値の変化を示している。計測の原理は脈打つ血管の容積変化によるセンサへの加圧である。また、脈拍センサの立ち上がるタイミングで圧力センサの応答が、実験協力者 A, C は立ち上がり、実験協力者 B は立ち下がっている。これは測定を行った血管に到達した血液のタイミングの違いにより位相差が発生したと考えられる。位相差はあるが、圧力センサと脈拍センサのタイミングが重なっているため、提案手法で脈拍測定可能といえる。

なお、図 1 のデバイスは適切な位置にセンサが接觸しなければ、センサ応答が見られなかった。原因と改善の議論を 8.2 節で述べる。

7 アプリケーション

提案手法を用いて図 10 のアプリケーションを実装した。これらの実装は Python を用いて行った。

眼鏡の鼻あてに搭載した圧力センサを用いた耳びく入力と身体活動検出手法

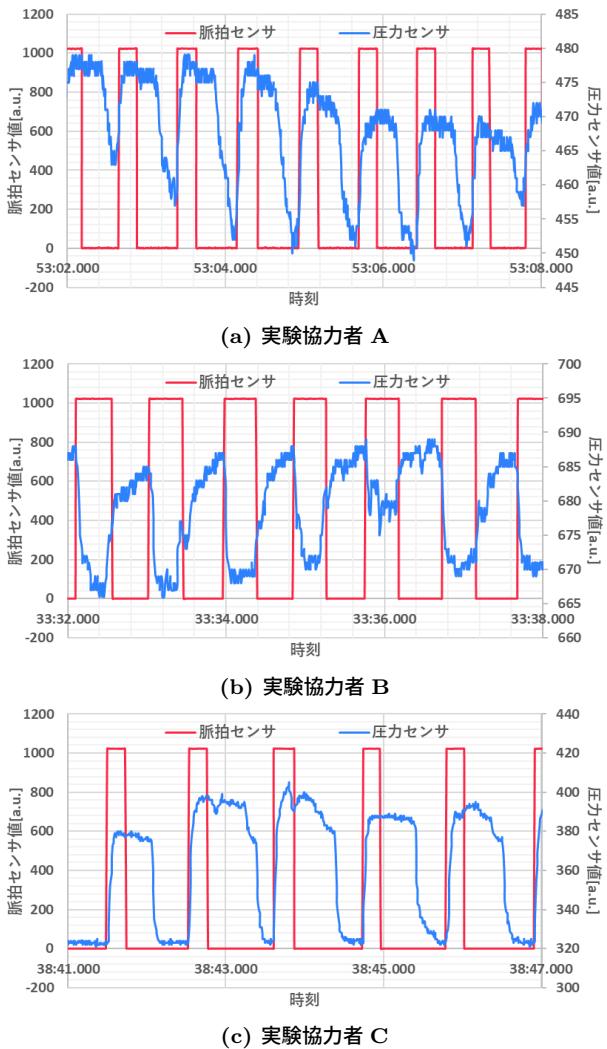


図 9. 脈拍センサと圧力センサの脈拍計測結果

耳びく入力をキーボード操作やマウス操作に割り当てるプログラムを作成した。耳びくにスクリーンショット機能を割り当てることで、従来操作に比べてキーの検出や誤操作の影響が少なく、撮影したい瞬間に操作できる。また、歩行による波形より耳びくによる波形の方が変化量が大きい。そのため、この2つの検出を組み合わせることにより歩行中は異なる操作を割り当てるこも可能である。例えばウォーキング中のメディアプレイヤ再生/一時停止操作に耳びく操作を使用することができる。

また、瞬き検出よりユーザの瞬きの頻度を観測するシステムを制作した。瞬きの間隔時間を測定できる他、15秒以上瞬きが検出されなければ、目の乾燥をアプリケーションのウィンドウ上に表記される。

8 議論

8.1 ジェスチャ数の拡張と分類

4章で挙げた動作以外にも著者が試したところ図11と図12で示されるように左右それぞれのウィンク(左ウィンク、右ウィンク)や耳びく(左耳びく、

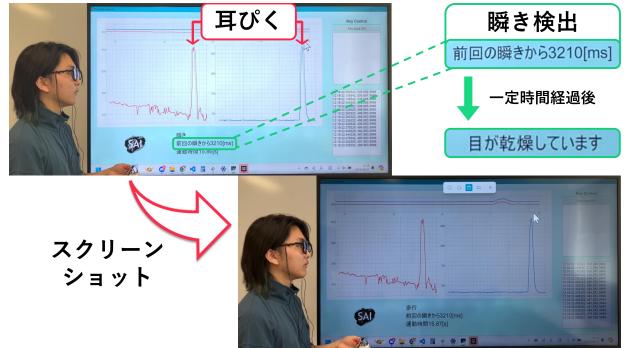


図 10. アプリケーション: 瞬き検出と通知、耳びく入力でのスクリーンショット

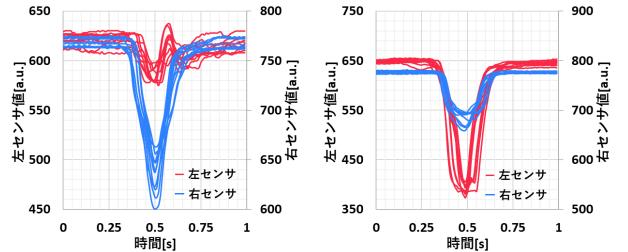


図 11. 著者がウインクを 10 回行った結果

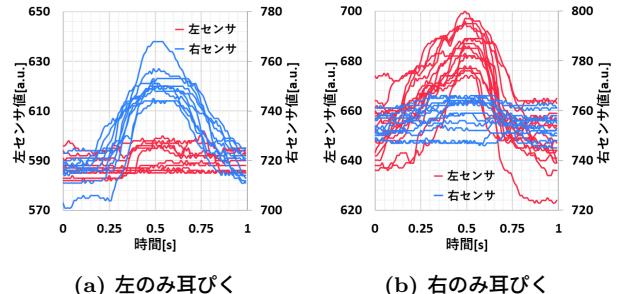


図 12. 著者が片耳ずつの耳びくを 10 回行った結果

右耳びく)が検出できた。この時、指定のジェスチャによる波形の変化を得るために、脈拍が検出されない位置で測定を行った。無入力、耳びく、左耳びく、右耳びく、瞬き、左ウインク、右ウインク、および歩行の8ジェスチャを各動作60個用意しSVMで分類器を作成した。これらのデータに対して正規化を行うが、4章の正規化法では左右のセンサ応答の振幅の違いが削除されてしまう。そのため、正規化法の(2)を変更し(1)の計算後の左と右それぞれのセンサ応答の最大値と100を比較し、最大となる値で全てデータを除算した。学習データとは別のデータを各動作15個データを用意して確認を行った結果、混同行列が図13、F1スコアが表3のようになった。表3より瞬きのF1スコアが16%と低く左ウインクと識別していないことと、図13より右ウインクが瞬きとの識別で2分されていることが分かる。これ

		推定ジェスチャ							
		無入力	両耳びく	左耳びく	右耳びく	両目瞬き	左ウインク	右ウインク	歩行
指定期	無入力	15	0	0	0	0	0	0	0
	両耳びく	0	15	0	0	0	0	0	0
	左耳びく	0	0	15	0	0	0	0	0
	右耳びく	1	0	0	14	0	0	0	0
	両目瞬き	0	0	1	0	2	12	0	0
	左ウインク	0	0	0	0	0	15	0	0
	右ウインク	0	0	0	0	8	0	7	0
	歩行	0	0	3	0	0	0	0	12

図 13. 著者データでの 8 ジェスチャ混同行列

表 3. 著者データでの 8 ジェスチャF1 スコア

動作	無入力	耳びく	左耳びく	右耳びく	平均
F1	0.968	1.000	0.882	0.966	
動作	瞬き	左ウインク	右ウインク	歩行	
F1	0.16	0.714	0.636	0.889	0.777

らは左右のセンサの接触によるセンサ応答の違いと、正規化処理の影響で信号波形が似てしまうことが原因である。今後はハードウェアの再設計、データの前処理を工夫するなどで検出率の改善が行えるか検証する必要がある。

左右別々に動かすジェスチャの検出にはセンサ特性、左右のセンサの接触具合が影響するため、実験協力者間で一貫性のある結果が示されなかった。加えて、片耳ずつの耳びくの難易度が高いためデータを十分に集めることが難しい。今後、デバイス形状の改良やフォトリフレクタなど異なるセンサを用いることでこれらの動作を検出できるか試みる。

8.2 測定安定性

測定の安定性に課題がある。5章の結果のようにセンサが動作に対して反応しづらいことがあった。これは頭の形がユーザごとに異なることが影響していると考えられる。より精度を高めるにはデバイスの形状の改良、フォトリフレクタなど異なるセンサの使用、もしくはより多くのデータを用いた高精度の分類器の作成が必要となる。

脈拍測定では装着位置によってセンサと血管間の距離や圧力が変わり、センサ応答が観測されない。現時点では安定条件は推測の域を出ていないため、さらなる研究を通じて計測の安定条件を明確にし、装着位置がずれない工夫や計測手法の設計を行う必要がある。また脈拍数の算出精度を上げるためにサンプリング周波数を上げる必要がある。

歩行中の他の動作の入力では、歩行の影響を受けづらい安定的な測定法が必要である。耳びくと瞬きはある程度の強さの入力であれば、歩行のセンサ応答より大きい変化量が現れるため、識別することが可能である。しかし、細かな入力や脈拍は歩行に

埋もれてしまう。周期性の違いからフィルタリングを行うなど、データ処理で歩行の影響を除去できるか検討する必要がある。

8.3 耳びくの特性

耳びくを入力として用いるケースは少なく、その特性について理解を深める必要がある。特に学習コストに関しては現状耳びくできる人が限られていることから十分な調査が必要である。筋電図で観測しながら耳びくの習得を調べた研究 [24] [26] がある。これらの研究では 30 分以内のトレーニングによって耳びくを習得できると報告されている。しかし、提案手法は筋電センサよりセンサ感度が低く、必要な耳の可動量が大きいため、これらのトレーニングで十分に本デバイスでの操作が行えるかは不明である。また、インターネットにいくつか練習法がある [20] [23]。これらは日々のトレーニングで数週間かけて習得を目指しているが、実際に習得できるかの検証が必要である。

耳びく入力の使用感の評価も必要である。耳びく習得済みの実験協力者は練習なしでデバイス入力が行えた。一方で、実験協力者の 2 名は実験終了時に少し耳の疲れを感じていた。疲労の大きな要因は耳びくが日常で未利用であり、動作に慣れていないためと考えられる。著者の 1 人はデバイス製作の過程で長時間耳びくを続けた結果、1 月ほどで 1 時間ほど耳びくを続けてもさほど疲れを感じなくなった。そのため、慣れにより耳びく入力による疲労は起こりづらくなると考える。

9 まとめ

本稿ではスマートグラスの鼻あてに圧力センサを搭載することで、耳びくを検出する手法を示した。実際にデバイスを作製し、実験したところ両耳で行う耳びくが計測可能であることが示された。同時にユーザの健康管理に役に立つデータ: 瞬き、歩行、脈拍を同デバイスで取得できることを示した。検出できる動作はそれぞれ特徴的な波形、周期を示すことから、これらの値を十分に区別することができる。SVM を用いた識別においてもユーザごとの分類器では F1 スコアが平均 98.3%，汎用分類器では F1 スコアが平均 92.4% と高い精度でこれらの動作を識別できた。脈拍測定では圧力センサでの測定と脈拍センサの測定の周期が一致していることが示された。今後はデバイスの形状の改良や学習するデータ量を増やすことで認識精度の向上や識別可能な動作やジェスチャを増やしていく。

参考文献

- [1] S. Ahn and G. Lee. Gaze-Assisted Typing for Smart Glasses. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, pp. 857 – 869,

眼鏡の鼻あてに搭載した圧力センサを用いた耳びく入力と身体活動検出手法

- New York, NY, USA, 2019. Association for Computing Machinery.
- [2] Connecticut Science Center. Useless Super Powers, 2020. <https://ctsciencecenter.org/blog/useless-super-powers/> (Accessed: 2024/08/30).
- [3] C. Darwin. *THE DESCENT OF MAN, AND SELECTION IN RELATION TO SEX*, Vol. 1. John Murray, 1st edition, 1873. pp.20–21 <https://darwin-online.org.uk/content/frameset?itemID=F937.1&viewtype=text&pageseq=1> (Accessed: 2024/08/30).
- [4] T. Gemicioglu, R. M. Winters, Y.-T. Wang, T. M. Gable, and I. J. Tashev. TongueTap: Multimodal Tongue Gesture Recognition with Head-Worn Devices. In *Proceedings of the 25th International Conference on Multimodal Interaction*, ICMI '23, pp. 564–573, New York, NY, USA, 2023. Association for Computing Machinery.
- [5] M. Goel, C. Zhao, R. Vinisha, and S. N. Patel. Tongue-in-Cheek: Using Wireless Signals to Enable Non-Intrusive and Flexible Facial Gestures Detection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 255–258, New York, NY, USA, 2015. Association for Computing Machinery.
- [6] B. J. Hou, J. Newn, L. Sidenmark, A. A. Khan, and H. Gellersen. GazeSwitch: Automatic Eye-Head Mode Switching for Optimised Hands-Free Pointing. *Proc. ACM Hum.-Comput. Interact.*, 8(ETRA), may 2024.
- [7] I. W. Huang, P. Rajbhandary, S. Shiu, J. S. Ho, J. Zhu, B. Wilson, and G. Ye. Radar-Based Heart Rate Sensing on the Smart Glasses. *IEEE Microwave and Wireless Technology Letters*, 34(6):829–832, 2024.
- [8] Interlink Electronic. FSR 400. <https://www.interlineelectronics.com/fsr-400> (Accessed: 2024/08/30).
- [9] JINS. JINS MEME. <https://jinsmeme.com/> (Accessed: 2024/08/30).
- [10] L.-H. Lee and P. Hui. Interaction Methods for Smart Glasses: A Survey. *IEEE Access*, 6:28712–28732, 2018.
- [11] D. J. Matthies. InEar BioFeedController: A Headset For Hands-Free And Eyes-Free Interaction With Mobile Devices. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pp. 1293–1298, New York, NY, USA, 2013. Association for Computing Machinery.
- [12] Meta. Ray-Ban Meta スマートグラス. <https://www.meta.com/smart-glasses/> (Accessed: 2024/08/30).
- [13] Microsoft. Microsoft HoloLens. <https://www.microsoft.com/hololens/> (Accessed: 2024/08/30).
- [14] MMD研究所. 2021年スマートウォッチに関する利用実態調査. https://mmdlabo.jp/investigation/detail_1930.html (Accessed: 2024/08/30).
- [15] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen. Gaze + Pinch Interaction in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, SUI '17, pp. 99–108, New York, NY, USA, 2017. Association for Computing Machinery.
- [16] D. Schön, T. Kosch, F. Müller, M. Schmitz, S. Günther, L. Bommhardt, and M. Mühlhäuser. Tailor Twist: Assessing Rotational Mid-Air Interactions for Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pp. 1–14, New York, NY, USA, 2023. Association for Computing Machinery.
- [17] SEEED Studio. Grove イヤークリップ心拍数センサ. <https://jp.seeedstudio.com/Grove-Ear-clip-Heart-Rate-Sensor.html> (Accessed: 2024/08/30).
- [18] L. Sidenmark and H. Gellersen. Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, pp. 1161–1174, New York, NY, USA, 2019. Association for Computing Machinery.
- [19] D. J. Strauss, F. I. Corona-Strauss, A. Schroeer, P. Flotho, R. Hannemann, and S. A. Hackley. Vestigial Auriculomotor Activity Indicates the Direction of Auditory Attention in Humans. *eLife*, 9:e54536, jul 2020.
- [20] wikiHow. How to Wiggle Your Ears. <https://www.wikihow.com/Wiggle-Your-Ears> (Accessed: 2024/08/30).
- [21] S. Yi, Z. Qin, E. Novak, Y. Yin, and Q. Li. GlassGesture: Exploring Head Gesture Interface of Smart Glasses. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pp. 1–9, 2016.
- [22] Q. Zhang, S. Gollakota, B. Taskar, and R. P. Rao. Non-Intrusive Tongue Machine Interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 2555–2558, New York, NY, USA, 2014. Association for Computing Machinery.
- [23] ちゃんねるGO. 【潜在能力】誰でも耳を動かせるようになる方法【開放する】. <https://www.youtube.com/watch?v=TZDFpTb0AbQ> (Accessed: 2024/08/30).
- [24] 式田 寛, 青山 忠義, 長谷川 泰久. ロボット指の内部モデル獲得に向けた後耳介筋による操作学習. ロボティクス・メカトロニクス講演会講演概要集, 2018:2A1-B18, 2018.
- [25] 藤田 恒太郎. 人体解剖学. 南江堂, 改訂 41 版, 1993. pp.115-116, pp.521-522.

- [26] 古賀 精治. 新奇な筋活動の制御の獲得における気づきと筋電図バイオフィードバック. 心理学研究, 62(5):308–315, 1991.
- [27] 総務省情報通信政策研究所. 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. 第4章 4-2 pp.73–75 https://www.soumu.go.jp/main_content/000952988.pdf.
- [28] 矢野経済研究所. 「XR（VR/AR/MR）対応HMD・スマートグラス市場に関する調査（2023年）」. https://www.nikkei.com/article/DGXZRS658618_V00C23A7000000/.
- [29] 大塚 崇, 高田 嶺介. 眼鏡型デバイスにおける耳びく入力手法. インタラクション2024 予稿集, pp. 1100–1105. 情報処理学会, 2024.

未来ビジョン

本手法を搭載したスマートグラスがスマートフォンを手に持つ煩わしさから解放してくれるだろう。なぜなら耳びく操作ができるからだ。耳びくは日常生活では全く役に立たない。だが日常で役に立たないからこそ、現在行っている日常動作と干渉せずデバイスの操作が可能となる。周辺環境に左右されず、歩きながらでも、手で作業をしながらでも、話しながらでも、視線を変えながらでも、食事をしながらでも耳びくはできる。左右の耳びくを左クリック、右クリックに割り当て視線のカーソル操作としてネットサーフィンをするのが一般的になり、友達と会話しながら調べたり、料理中にレシピを確認したり、手も使わずベッ

ドでごろごろしながらSNSを見たり、耳びくモールス信号で文字を打つことだって可能だ。スマートグラスにカメラを搭載すれば手でボタンを押すより早く耳びくでシャッターを切れる。VRではアバターの耳を人間の耳と連動させて耳びくしたり、象になって耳で空を飛ぶこともできる。

デバイスの操作だけでなく人間の健康管理に役立つ。スマートグラスをつけるだけで心拍数をいつでも教えてくれる。歩行を検知して歩数、消費カロリー、運動時間を自動でまとめてくれる。瞬きを検知して眼気、目の乾燥度合いを読み取ってユーザに休憩を促してくれる。工夫せねばスマートウォッチで行っていることがすべてできるようになる。

母音、子音の順に選択を行う間接タッチ用かな文字入力手法

和田 優斗 * 白根 薫 † 崔 明根 ‡ 志築 文太郎 §

概要. 間接タッチとは、操作対象が表示された出力面に直接触れるのではなく、別の入力面を介して行われるタッチを指す。間接タッチを行う際、ユーザは入力面を直接視認しないため、正確な位置に指をタッチダウンすることが難しい。日本語文字入力におけるこの課題を解決するために、我々はスマートウォッチを入力面に用いた間接タッチ用かな文字入力手法を開発した。本手法では、ユーザはベゼルからのスライドインによる母音選択、および目的の子音キーからのタッチアップによる子音選択を連続して行う。これにより、入力面の視認を必要としない、間接タッチによる正確な日本語文字入力が達成される。研究室内実験の結果、間接タッチ時に、本手法の平均文字入力速度が 29.5 CPM、平均トータルエラー率が 10.0% であることが示された。実験結果より、間接タッチにおける本手法の有用性が示唆された。

1 はじめに

間接タッチ (indirect touch) とは、操作対象が表示された画面 (以下、出力面) とは別の入力面 (以下、入力面) を介して行われるタッチを指す [16][19][12][15]。従来、間接タッチは、トラックパッドおよびペンタブレットを用いた操作に使用されてきた。近年では、スマート TV およびヘッドマウントディスプレイ (HMD) の操作にも間接タッチが活用されている [20]。

出力面に直接触れる直接タッチ (direct touch) と比較して、間接タッチは、垂直な出力面のタッチ時に腕を持ち上げることによる疲弊 (ゴリラーム現象 [4]) を抑制する [15]。また、間接タッチを文字入力に応用すれば、ユーザは出力面から目を離さずに、他の物体を見ながら文字入力を行うことが可能となる。この例として、スマート TV での動画視聴中に内容を見ながらコメントを入力する、および仮想現実上のチャットにて、相手の顔を注視しながらメッセージを送信するという状況が想定される。

間接タッチを行う際、入力面にフィードバックが表示される場合、ユーザは入力面および出力面を交互に確認するために、頭部および首に負担が掛かる。一方、入力面にフィードバックが表示されない場合、ユーザは入力面を直接視認しないため、正確な位置に指をタッチダウンすることが難しい [19]。この特性を踏まえて、タッチダウンが不正確であっても使用可能な、間接タッチ向けに設計された文字入力手法が提案されている [20][21][1]。これらの手法では、出力面の一部にキーボードおよびポインタが表示される。ユーザは、間接タッチによって操作したポイ



図 1. 本手法を用いたアプリケーション例。ユーザはスマート TV 上の動画を見ながら、スマートウォッチにタッチしてコメントの入力をを行う。

ンタを用いてキーを選択することによって文字入力をを行う。しかし、これらの手法は英文字入力を対象としており、我々の知る限りでは、間接タッチ向けに設計された日本語文字入力手法は提案されていない。

そこで、我々はスマートウォッチを入力面に用いた間接タッチ用かな文字入力手法を開発した。本手法では、ユーザはベゼルからのスライドイン (入力面外側から内側へのスライド) を用いて母音を選択した後、そのまま入力面上にて指をスライドして、目的の子音キーにてタッチアップすることによって子音を選択する。この操作により、母音、子音の順に連続した 1 ストロークにて入力を実行。ユーザはベゼルに触れた際に得られた触覚フィードバックを手掛かりにジェスチャを行うため、入力面を視認しなくとも、正確に入力を開始することができる。これにより、入力面を直接視認することなく、出力面に目を向けたまま、間接タッチを用いて日本語文字入力を実行することができる (図 1)。

本論文の貢献を以下に示す。

- 間接タッチ用かな文字入力手法の設計および実装を示した。
- 研究室内実験にて、間接タッチ時に、本手法の平均文字入力速度が 29.5 CPM、平均トータルエラー率が 10.0% であることを示した。

Copyright is held by the author(s).

* 筑波大学 情報メディア創成学類

† 筑波大学 情報理工学位プログラム

‡ 日本学術振興会

§ 筑波大学 システム情報系

- タルエラー率が 10.0% であることを示した。
 • 実験参加者から得られたコメントを参考に、本手法の改善点を示した。

2 関連研究

本手法は間接タッチ用かな文字入力手法である。また、入力面にスマートウォッチの画面を用いる。本節では、間接タッチおよびスマートウォッチのタッチジェスチャに関する研究、ならびにこれらを用いた文字入力手法を述べる。

2.1 間接タッチ

間接タッチの特性に関する調査、および間接タッチを用いたインターフェースの研究が行われている。Xin ら [19] は、間接タッチのターゲット選択精度の低さに着目して、間接タッチ入力技術を調査した。Voelker ら [16] は、間接タッチにおいて、ユーザが手元を視認せずオブジェクトに狙いを定めるために必要な設計の推奨事項を示した。また、間接タッチおよび視線を組み合わせた入力手法を提案した [15]。Schmidt ら [12] は、間接タッチおよび直接タッチにおけるマルチタッチの性能を比較した。我々の手法も、間接タッチにおける使用を想定する。

2.2 スマートウォッチのタッチジェスチャ

近年、Extended Reality (xR) へのスマートウォッチの活用が注目を集めている。特にスマートウォッチのタッチジェスチャを用いたインタラクションが提案されている。GestureMark [11] は、スマートウォッチのタッチジェスチャを用いて xR 空間上のターゲットを選択する技術である。Lang ら [10] は、xR 上にて、スマートウォッチのタッチジェスチャがハンドジェスチャと比較して高速に実行可能かつユーザに好まれることを示した。

また、スマートウォッチのベゼルを活用した研究が行われている。B2B-Swipe [9] は、四角形のスマートウォッチにおけるベゼルを始点および終点とした 16 種類のジェスチャを実現する。Wong ら [18] は、ベゼルを始点としたスワイプ（ベゼルからのスライドイン）である Bezel-initiated Swipe (BIS) の、円形スマートウォッチにおける性能を調査した。ユーザはベゼルに触れた際に触覚フィードバックを得られるため、入力面を視認しなくともこれらの操作を実行できる。以上の特徴を踏まえて、我々の手法もスマートウォッチおよび BIS を入力に使用する。

2.3 文字入力手法

間接タッチ向けに設計された文字入力手法が提案されている。Yang ら [20] は、タッチの開始位置が不正確であっても使用可能な、ジェスチャ入力による間接タッチ用文字入力手法を提案した。加えて、論文中に直接の記載はないが、間接タッチを使用した文字入力手法が存在する。川瀬らの手法 [21] および HoldBoard [1] は、入力面にスマートウォッチ

を、出力面に HMD またはスマートグラスを用いる。しかし、我々の知る限りでは、間接タッチ向けに設計された日本語文字入力手法は存在しない。

また、スマートウォッチ用にキー選択を工夫した日本語文字入力手法が提案されている。スマートウォッチは画面が小さいため、ファットフィンガ問題 [13] およびスクリーンオクルージョン問題 [8] が発生する。これに対し、円環状に配置された子音キーからのスライドインまたはフリックによって入力が行われる手法 [25][23][22][14] が存在する。SliT [24] では、ユーザは周縁に配置された小さなキーから画面中央へと指をスライドインして子音を選択する。PonDeFlick [2] では、円環型に配置されたキーによって子音を選択した後に、フリック入力と共に方向にフリックして母音を選択する。しかし、これらの手法は入力面の視認を前提とするため、間接タッチにて使用することは難しい。

さらに、画面の視認を必要としないアイズフリー用日本語文字入力手法が提案されている。No-look Flick [7] では、画面上の 3 つのキーを始点にフリックを行う。下岡ら [26] は、2 本指を用いてフリックを行う、スマートウォッチ用アイズフリー文字入力手法を提案した。これらの手法がアイズフリー用に特化しているのに対し、我々は出力面を活用した、間接タッチ用の文字入力手法を開発した。

3 本手法の設計

本手法は、入力面にスマートウォッチの画面を用いた、間接タッチ用かな文字入力手法である。

本手法では出力面にキーボードが表示される。ユーザは、初めにスライドインを用いて母音を選択する。この際に選択される母音は、スライドインの始点によって決定される。続けてユーザは入力面から指を離さずにスライドした後、入力対象の子音キーにてタッチアップすることによって子音を選択する。この操作を通じて、母音、子音の順に連続した 1 ストロークにて入力をを行う（図 2）。

本手法は、日本語文字入力において広く普及しているフリック入力と同配列のキーボードを採用する。キーボードには初期状態にて 12 個のキー「あ, か, さ, た, な, は, ま, や, ら, 小, わ, ←」が表示される。これらのキーは、母音選択後に、選択された母音に応じたキー（「い」段の選択時には「い, き, し, ち, に, ……」）に変化する。「小」キーは末尾の文字を濁音、半濁音、および捨てがなに変換する。「←」キーは末尾の文字を削除する。

本手法の間接タッチ用設計は、母音入力が BIS に割り当てられ、かつ母音が先に入力される点を特徴とする。先行研究にて、画面を視認しない状態では正確なタッチが困難である [7] のに対し、入力方向が少ない BIS は正確に実行可能である [18] と報告されている。このため、子音と比べて数の少ない母

母音、子音の順に選択を行う間接タッチ用かな文字入力手法

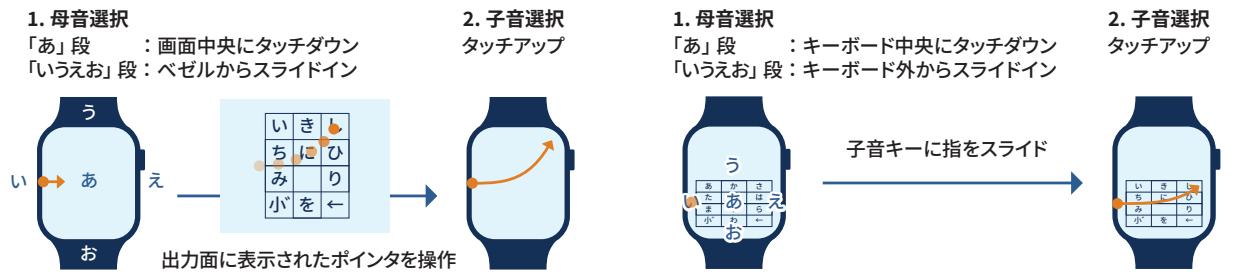


図 2. 本手法の操作 (a: 間接タッチ設計, b: 直接タッチ用設計)

音の入力を BIS に割り当てることによって、ユーザは入力面を直接視認しなくとも、ジェスチャを正確に開始できる。タッチダウン後はタッチ位置を示すポインタが出力面に表示されるため、ユーザはこのポインタを通じて、子音も正確に選択できる。これにより、間接タッチにおける正確かつ連続的なかな入力が実現される。加えて、既存のフリック入力とスライドインの方向を共通化することによって、学習コストの削減を見込む。

我々は、本手法を間接タッチ用に設計しつつ、直接タッチにおいても使用可能とするために、本手法を直接タッチ用に再設計した。この設計を以下に示す。

3.1 間接タッチ用設計

キーボードは入力面（スマートウォッチの画面）には表示されず、別の出力面に表示される（図 2a）。母音入力時、入力対象の母音が「い、う、え、お」であれば、ユーザはスマートウォッチのベゼル（「い」段：左側、「う」段：上側、「え」段：右側、「お」段：下側）から入力面中央へ指をスライドインする。入力対象の母音が「あ」段であれば、入力面の中央付近にタッチダウンする。スライドインまたはタッチダウンした時点にて母音が選択されると同時に、入力面のタッチ位置を示すポインタが出力面に表示される。子音入力時、ユーザは入力面から指を離さずにスライドすることによって、出力面上のポインタを目的の子音キーまで移動させる。その後、指をタッチアップすることによって子音を選択する。

入力される母音は、入力面を 5 分割した領域へのタッチダウンの位置によって決まる（図 3a）。入力される子音は、入力面を 12 分割した領域からのタッチアップの位置によって決まる（図 3b）。指の移動量を抑えるとともに端に位置するキーの選択を容易にするために、「い、う、え、お」段が入力されている間、スライドインされたベゼルの方向に判定領域が移動する。加えて、端に位置するキーの判定領域が入力面の端まで拡張される。誤った母音を選択した場合、ユーザはスライドインしたベゼルに指を戻すことによって入力を中断できる。ただし、「あ」段を選択した場合には、入力を中断できない。

3.2 直接タッチ用設計

キーボードは入力面（スマートウォッチの画面）に表示される（図 2b）。ユーザは、入力対象のかな

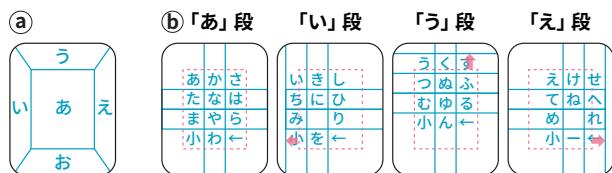


図 3. 間接タッチにおける判定領域の仮想的な図示 (a: 母音判定の領域, b: 子音判定の領域)

の母音に応じてキーボードの内側または外側（「あ」段：内側、「い」段：左側、「う」段：上側、「え」段：右側、「お」段：下側）にタッチダウンする。続けて目的の子音キーまで指をスライドした後、タッチアップする。入力を中断する場合、キーボードの外側に指をスライドさせる。間接タッチ用設計では、母音入力時に各方向のベゼルがスライドインの起点となる。一方、直接タッチ用設計では、入力面にキーボードが表示されるため、キーボードに対する相対位置がスライドインの起点となる。

4 実験

ベースライン手法との比較を通じた本手法の性能評価を行うとともに、間接および直接タッチにおける利用を比較するために、研究室内実験を行った。

4.1 実験計画

本実験を参加者内計画にて行った。独立変数は、タッチ条件（間接タッチ、直接タッチ）、および入力手法（スライド手法、フリック手法）の 2 つである。

間接タッチ ユーザの正面に設置されたディスプレイが出力面となる。課題文、入力文、およびキーボードはディスプレイに表示される。実験参加者は、ディスプレイを見ながら、スマートウォッチの画面をタッチして入力をを行う（図 4a）。いずれの入力手法においても、スマートウォッチの画面へのタッチ中は、タッチ点を示すポインタがディスプレイに表示される。スマートウォッチの画面に情報は表示されない。

直接タッチ スマートウォッチの画面が出力面となる。課題文、入力文、およびキーボードはスマートウォッチに表示される。実験参加者は、スマートウォッチの画面に表示された情報を見ながら、スマートウォッチの画面をタッチして入力をを行う（図 4b）。ディスプレイは使用しない。



図 4. タッチ条件 (a: 間接タッチ, b: 直接タッチ)

スライド手法 3 節に述べた入力手法である。

フリック手法 スマートフォンおよびスマートウォッチにて標準的に使用されるフリック入力と同様の入力手法を実装した。実験参加者は入力対象のかなの子音キーにタッチダウンした後、かなの母音に応じてフリックまたはそのままタッチアップする。

実験条件は、タッチ条件および入力手法を組み合わせた以下の 4 条件から構成される。

- indirect-slide (IS): 間接タッチにてスライド手法を使用する
- indirect-flick (IF): 間接タッチにてフリック手法を使用する
- direct-slide (DS): 直接タッチにてスライド手法を使用する
- direct-flick (DF): 直接タッチにてフリック手法を使用する

従属変数には、文字入力速度 (CPM: Characters Per Minute), トータルエラー率 [3], ユーザビリティ (SUS [5]), およびワーカロード (NASA-RTLX: NASA Raw Task Load Index [6][27]) を用いた。

4.2 実験参加者

実験には、研究室内の大学生および大学院生 12 名（平均 22.8 歳, SD = 1.0 歳, P1-P12）が参加した。実験参加者のうち、1 名が女性、1 名が左利きであった。スマートフォンの平均使用年数は 9.1 年 (SD = 2.0 年) であった。全員がスマートフォンを毎日使用すると回答した。3 名がスマートウォッチを日常的に使用すると回答した。

4.3 実験システム

実験にはスマートウォッチ (Apple Watch Series 9, 41 mm モデル), ラップトップ PC (MacBook Air, M1 2020), および 27 インチディスプレイを使用した。ディスプレイは PC に接続された後、机の端から 0.6 m 離れた位置に設置された（図 4）。

4.3.1 実験用アプリケーション

スライド手法およびフリック手法を用いて文字入力をを行う実験アプリケーションを実装した。本アプリケーションは、スマートウォッチ用アプリケーションおよび PC 用アプリケーションから構成される（図 5）。これらは、Bluetooth Low Energy を介して通信を行った。課題文と一致する入力文の部分文字列は、太字のフォントを用いて強調された。



図 5. タッチ条件ごとのデバイスの画面

スマートウォッチ用アプリケーション

Swift を用いて実装された watchOS (Apple Watch) 用アプリケーションである。本アプリケーションは、タッチ情報に基づく文字入力処理、タッチ座標および入力文の PC 用アプリケーションへの送信、ならびにキー選択時の触覚フィードバックの提示を行う。また、直接タッチに限って、課題文、入力文、キーボード、およびタッチ中の視覚フィードバックを表示する。

PC 用アプリケーション

Swift を用いて実装された macOS 用アプリケーションである。本アプリケーションは、スマートウォッチ用アプリケーションからタッチ座標および入力文を受信する。また、間接タッチに限って、課題文、入力文、キーボード、ならびにタッチ中のポインタおよび視覚フィードバックを表示する。

4.3.2 キーボード

すべての入力手法において、キーへのタッチ時に触覚フィードバックが提示された。スライド手法では、指をスライドさせて別のキーを選択した際にも触覚フィードバックが提示された。キーボードの幅は、watchOS 標準のフリック入力を参考に画面幅の 70% に設定された。キーボードの高さは、間接タッチではキーボードの幅と等倍に設定された。直接タッチでは課題文および入力文を表示する都合上、キーボードの幅の 60% に設定された。

4.4 実験タスク

本実験のタスクは、表示された課題文を、指定された文字入力手法を用いて入力することである。実験参加者は、可能な限り高速かつ正確に文字を入力するように指示された。タスクは座位にて行われた。実験参加者は、利き手と反対の腕にスマートウォッチを装着して、腕を机の上に置いた。操作時には利き手の親指および中指を用いてスマートウォッチを把持することが推奨され、また人差し指を用いて画面をタッチするように指示された。加えて、OS 標

母音、子音の順に選択を行う間接タッチ用かな文字入力手法

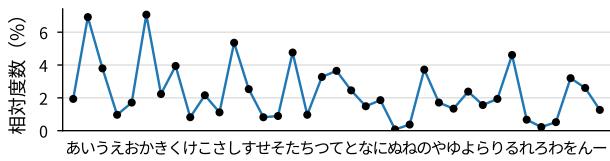


図 6. かな文字の出現頻度

準の通知センタの呼び出しを避けるために、上側のベゼルに長時間触れないように指示された。

実験参加者は、課題文が表示された時点で入力を開始した。また、課題文の入力が完了したと各自が判断したタイミングにて、スマートウォッチの竜頭を回転させた。これにより、次の課題文が表示された。一連の操作（課題文の入力および竜頭の回転）を、すべての課題文の入力が終了するまで行った。

4.5 実験手順

実験概要の説明を行った後、実験前アンケートを実施した。実験参加者は、年齢、性別、利き手、スマートフォンおよびスマートウォッチの使用状況、普段使用する文字入力手法、ならびにフリック入力の習熟度を回答した。

次に実験タスクを説明した後、4 セッションから構成される実験を実施した。実験条件はセッションごとに変更された。実験条件の順番はランダム法を用いて、実験参加者間のカウンタバランスを取って決定された。各セッションにおいて、実験参加者はタッチ条件および文字入力手法を確認した後、練習タスクを行った。練習タスクでは、12 文の練習文が連続して提示された。続いて、本番タスクを行った。本番タスクでは、28 文の課題文が連続して提示された。本番タスクには、練習タスクにて使用した練習文セットとは異なる課題文セットが使用された。本番タスク終了後、実験参加者は、SUS、自由記述、NASA-RTLX から構成されるアンケートに回答した。

全セッション終了後、実験後アンケートとして、実験参加者はタッチ条件ごとに好みの入力手法および順位付けの理由を回答した。

4.6 課題文セット

本番タスクには、我々が独自に作成した 4 セットからなる日本語の課題文セットを用いた。1 セットは短文 28 文から構成された。課題文セットには連番が振られた。セッション毎に、対応する番号の課題文セットが使用された。

課題文セットは、かなの清音、濁音、半濁音、捨てがな、および長音記号から構成され、また全体を通じてすべての清音が出現した（図 6）。1 つの短文の長さは 10–14 文字であり、文字数の平均は 12 文字、合計は 336 文字であった。どのセットにおいても、同じ長さの短文が同じ回数だけ出現した。

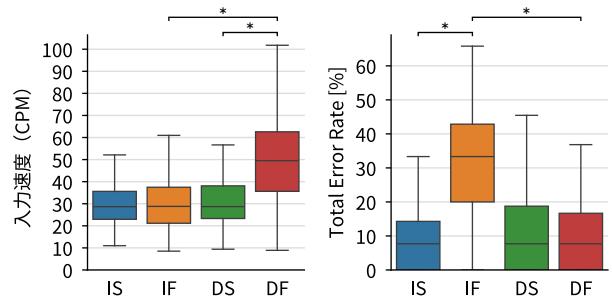


図 7. 手法ごとの平均文字入力速度および平均トータルエラー率 (* $p < .05$, ひげは外れ値を除いた最小値および最大値を示す)

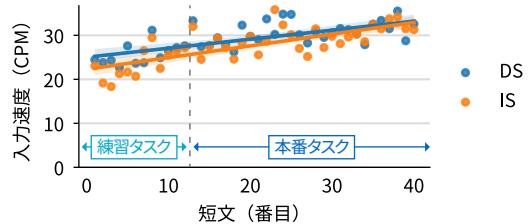


図 8. 文字入力速度の変遷および線形回帰

5 実験結果

5.1 文字入力速度およびトータルエラー率

平均文字入力速度は、IS 条件では 29.5 CPM ($SD = 9.1$ CPM), IF 条件では 31.4 CPM ($SD = 13.5$ CPM), DS 条件では 30.8 CPM ($SD = 11.1$ CPM), DF 条件では 51.2 CPM ($SD = 20.7$ CPM) であった（図 7 左）。平均トータルエラー率は、IS 条件では 10.0% ($SD = 11.0\%$) , IF 条件では 32.3% ($SD = 14.7\%$), DS 条件では 11.8% ($SD = 11.6\%$), DF 条件では 10.8% ($SD = 10.9\%$) であった（図 7 右）。

文字入力速度およびトータルエラー率にコルモゴロフ-スミルノフ検定を行った結果、正規性は確認されなかった。そのため、入力手法およびタッチ条件を独立変数として整列ランク変換 [17] を行った後、二元配置分散分析を行った。その後、下位検定としてウィルコクソンの順位和検定を実施した。その結果、文字入力速度について、タッチ条件の主効果 ($F_{1,1314} = 272.75, p < .05$) , 入力手法の主効果 ($F_{1,1314} = 284.47, p < .05$) , および交互作用 ($F_{1,1314} = 227.03, p < .05$) がみられた。トータルエラー率について、主効果のタッチ条件 ($F_{1,1314} = 293.41, p < .05$) , 入力手法の主効果 ($F_{1,1314} = 327.62, p < .05$) , および交互作用 ($F_{1,1314} = 382.45, p < .05$) がみられた。

練習および本番タスク中の文字入力速度の変遷を図 8 に示す。

5.2 SUS および NASA-RTLX

SUS（高いほど良い）および NASA-RTLX（低いほど良い）の平均スコアを図 9 に示す。

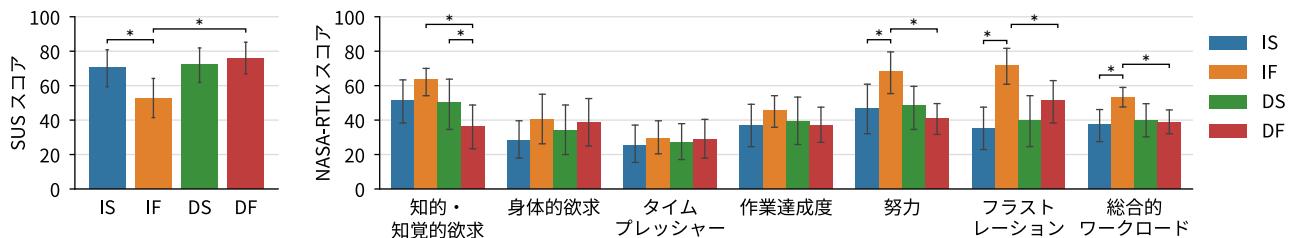


図 9. 手法ごとの平均 SUS スコア, NASA-RTLX の項目別平均スコア (* $p < .05$, エラーバーは 95% 信頼区間)

SUS および NASA-RTLX にシャピロ-ウィルク検定を行った結果、正規性が確認された。そのため、二元配置分散分析を行った。その後、下位検定として t 検定を実施した後、ホルム補正を行った。その結果、SUS について、タッチ条件の主効果 ($F_{1,1843} = 13.65, p < .05$) および交互作用 ($F_{1,1435} = 6.96, p < .05$) がみられた。総合的ワークロードについて、交互作用 ($F_{1,840} = 18.60, p < .05$) がみられた。

5.3 手法の好み

間接タッチでは、11 名がスライド手法を、1 名がフリック手法を好むと回答した。スライド手法の選択理由には、「誤入力が少ない」(P1, P3, P6, P8), 「触っている位置が分かりやすい」(P2, P4), 「指の開始位置を判断しやすい」(P5), 「タッチ時にキーを選択しなくてよい」(P7), 「正確にキーを選択できる」(P10), 「ミスタッチした際にリカバリできる」(P11, P12) が挙げられた。フリック手法の選択理由には「使い慣れている」(P9) が挙げられた。

直接タッチでは、6 名がスライド手法を、6 名がフリック手法を好むと回答した。スライド手法の選択理由には、「誤入力が少ない」(P3, P8), 「キーの小ささが気にならない」(P4, 10), 「どのキーを選択しているかが把握できる」(P6), 「端のキーが選択しやすい」(P11) が挙げられた。フリック手法の選択理由には「使い慣れた手法である」(P1, P5, P7, P9, P12), 「入力が速い」(P2) が挙げられた。

6 考察および今後の課題

実験結果およびアンケートのコメントを基に、本手法に関する考察を述べる。

6.1 間接および直接タッチへの応用

間接タッチにて、本手法はフリック手法と比較して、平均文字入力速度が同程度でありながらも、平均エラー率および総合的ワークロードが有意に低く、またユーザビリティが有意に高かった。ゆえに、本手法の間接タッチにおける有用性が示唆された。

直接タッチにて、本手法はフリック手法と比較して、平均文字入力速度が有意に低かった。一方、自由記述には「指を動かしながら操作を行うため、選択中のキーが分かりやすい」(P4, P6, P12) との回答があった。ゆえに、本手法はファットフィンガ問題の解消に有用である可能性がある。また、本手法

はタッチダウン時のみならず、指をスライドさせて別のキーを選択した際にも、連続的に触覚フィードバックを提示できるという特徴を持つ。複数の実験参加者 (P1, P5, P6, P9, P10) が、自由記述にてこの点を好意的に回答した。

6.2 設計の改善

我々は、スライドインの始点に、一般的なフリック入力における母音のフリック方向と同様の方向を割り当てることによって、本手法にフリック入力との共通性を持たせようと試みた。しかし、現在の設計では指を動かす方向がフリック入力と反対となるため、複数の実験参加者 (P1, P8, P12) が入力方向について混乱したと回答した。その他に「指の移動量が多い」(P6, P10) という回答も存在した。したがって、スライドインの方向および最適な Control-Display 比を検討する必要がある。

6.3 今後の課題

今回の実験では、スマートウォッチのベースライン手法であるフリック手法との比較を行った。今後は、研究室外実験を通じて本手法の正確な性能を調査するとともに、長期実験による本手法の正確な習熟を明らかにする。加えて、既存のアイズフリー手法および音声入力、ならびにポインタを用いてキー選択を行う手法との比較実験を行う。また、xR 環境における使用を想定して、出力デバイスに HMD を用いた際の本手法の性能、およびその際の視線の動きについても調査する。加えて、スマートフォン等の別デバイスを入力に用いた際のインターフェースについても検討を進める。

7 まとめ

本論文では、スマートウォッチを入力に用いた間接タッチ用かな文字入力手法を示した。研究室内実験を通じて、本手法およびフリック手法を間接および直接タッチにて使用した際の性能を調査した。実験の結果、間接タッチ時に本手法の平均文字入力速度は 29.5 CPM、平均トータルエラー率は 10.0% を記録した。これにより、間接タッチにおいて、本手法は実用的な日本語文字入力手法として使用可能なことが示唆された。今後はさらなる実験を通じて、本手法の詳細な性能および最適な設計を調査する。

参考文献

- [1] S. Ahn, S. Heo, and G. Lee. Typing on a Smart-watch for Smart Glasses. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, pp. 201–209, 2017.
- [2] K. Akamine, R. Tsuchida, T. Kato, and A. Tamura. PonDeFlick: A Japanese Text Entry on Smartwatch Commonalizing Flick Operation with Smartphone Interface. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, pp. 1–11, 2024.
- [3] A. S. Arif and W. Stuerzlinger. Analysis of Text Entry Performance Metrics. In *2009 IEEE Toronto International Conference Science and Technology for Humanity*, TIC-STH '09, pp. 100–105, 2009.
- [4] S. Boring, M. Jurmu, and A. Butz. Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, OzCHI '09, pp. 161–168, 2009.
- [5] J. Brooke. SUS: A Quick and Dirty Usability Scale. *Usability Evaluation in Industry*, pp. 189–194, 1996.
- [6] J. C. Byers, A. Bittner, and S. G. Hill. Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary. *Advances in industrial ergonomics and safety*, 1:481–485, 1989.
- [7] Y. Fukatsu, B. Shizuki, and J. Tanaka. No-look flick: Single-handed and Eyes-free Japanese Text Input System on Touch Screens of Mobile Devices. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '13, pp. 161–170, 2013.
- [8] C. Harrison and S. E. Hudson. Abracadabra: Wireless, High-precision, and Unpowered Finger Input for Very Small Mobile Devices. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pp. 121–124, 2009.
- [9] Y. Kubo, B. Shizuki, and J. Tanaka. B2B-Swipe: Swipe Gesture for Rectangular Smartwatches from a Bezel to a Bezel. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 3852–3856, 2016.
- [10] M. Lang, C. Strobel, F. Weckesser, D. Langlois, E. Kasneci, B. Kozlíková, and M. Krone. A Multimodal Smartwatch-based Interaction Concept for Immersive Environments. *Computers & Graphics*, 117:85–95, 2024.
- [11] J. Lee, M. Baeck, H.-S. Yeo, T. Starner, and W. Woo. GestureMark: Shortcut Input Technique using Smartwatch Touch Gestures for XR Glasses. In *Proceedings of the Augmented Humans International Conference 2024*, AHs '24, pp. 63–71, 2024.
- [12] D. Schmidt, F. Block, and H. Gellersen. A Comparison of Direct and Indirect Multi-touch Input for Large Surfaces. In *Proceedings of the IFIP TC13 International Conference on Human-Computer Interaction*, INTERACT '09, pp. 582–594, 2009.
- [13] K. A. Siek, Y. Rogers, and K. H. Connolly. Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In *Proceedings of the IFIP TC13 International Conference on Human-Computer Interaction*, INTERACT '05, pp. 267–280, 2005.
- [14] T. Tojo, T. Kato, and S. Yamamoto. BubbleFlick: Investigating Effective Interface for Japanese Text Entry on Smartwatches. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '18, pp. 44:1–44:12, 2018.
- [15] S. Voelker, A. Matvienko, J. Schöning, and J. Borchers. Combining Direct and Indirect Touch Input for Interactive Workspaces using Gaze Input. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, SUI '15, pp. 79–88, 2015.
- [16] S. Voelker, C. Wacharamanottham, and J. Borchers. An Evaluation of State Switching Methods for Indirect Touch Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 745–754, 2013.
- [17] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The Aligned Rank Transform for Nonparametric Factorial Analyses using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 143–146, 2011.
- [18] P. C. Wong, K. Zhu, X.-D. Yang, and H. Fu. Exploring Eyes-free Bezel-initiated Swipe on Round Smartwatches. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–11, 2020.
- [19] Y. Xin, R. Liu, and Y. Li. Strategy for Improving Target Selection Accuracy in Indirect Touch Input. *IEICE Transactions on Information and Systems*, 103(7):1703–1709, 2020.
- [20] Z. Yang, C. Yu, X. Yi, and Y. Shi. Investigating Gesture Typing for Indirect Touch. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3 of *IMWUT*, pp. 117:1–117:22, 2019.
- [21] 川瀬 寛也, 澤邊 太志, 藤本 雄一郎, 神原 誠之, 加藤 博一. 日常利用の拡張現実感環境におけるタッチタイピング可能な文字入力システム. 複合現実感研究会 MR2022-10, pp. 30–35, 2022.
- [22] 東條 貴希, 本田 裕己, 加藤 恒夫, 山本 誠一. BubbleSlide: スマートウォッチ向け円環型日本語かな入力インターフェース. 情報処理学会論文誌, 60(10):2075–2084, 2024.
- [23] 西田 圭佑, 加藤 恒夫, 山本 誠一. スマートウォッチ向け日本語かな入力インターフェースの試作.

- 情報処理学会インタラクション 2017 論文集, pp. 897–902, 2017.
- [24] 秋田 光平, 田中 敏光, 佐川 雄二. 画面占有率の低いスマートウォッチ向け文字入力手法 SliT. ヒューマンインターフェース学会論文誌, 21(1):131–140, 2019.
- [25] 斎藤 航平, 奥 寛雅. HARI キーボード：超小型タッチパネル端末向け日本語入力キーボード. 情報処理学会インタラクション 2016 論文集, pp. 701–703, 2016.
- [26] 下岡 純也, 浅井 洋樹, 山名 早人. スマートウォッチにおけるアイズフリー日本語入力手法. 第 23 回インタラクティブシステムとソフトウェアに関するワークショップ予稿集, WISS 2015, pp. 1–2, 2015.
- [27] 芳賀 繁, 水上 直樹. 日本語版 NASA-TLX によるメンタルワークロード測定—各種室内実験課題の困難度に対するワークロード得点の感度. 日本人間工学会誌, 32(2):71–79, 1996.

E-String Figures: 導電纖維編み込み紐を用いたあやとり技認識システム

永山 晃誠 * 崔 明根 †‡ 武山 侑輝 * 赤田 真由 * 高田 嶺介 § 志築 文太郎 †

概要. 導電纖維を編み込んだ紐を用いることによって、あやとり技を認識するシステムを示す。認識には、導電纖維を編み込んだ紐を用いてあやとりをした際に、紐同士が絡まった接觸点において短絡が生じることによって、紐の電気抵抗値が変化する現象を用いる。あやとり技ごとに紐の絡み方が異なるため、抵抗値も異なる。本システムでは、両手の10本指に装着した電極を用いて、あやとり紐の抵抗値を測定し、得られた測定値から機械学習を用いることによってあやとり技を認識する。実験を行い、本システムによる3種類の連続技と呼ばれる技の変形工程を認識したところ、それぞれのF値が0.91, 0.98および0.82という精度で認識できた。

1 はじめに

あやとりは、1本の輪にした紐を手にかけ、指を用いてその紐を取ることによって様々な形象を作り出す伝承遊びである[15][24]。日本では平安時代から存在していたと考えられており、主に子供遊びとして古くから親しまれてきた[19]。また、あやとりは海外にも歴史が存在し、アフリカ、アジア、北アメリカ、南アメリカ、オセアニアおよび極北圏などの地域において、その文化および風土を反映した多様なあやとり技が伝承されている[4]。

あやとりを遊ぶことによる、教育的効果もいくつか報告されている。あやとりは複雑な手指運動を伴うことに加えて、あらかじめ記憶した手順を思い出しながら紐を操作する必要があるという特徴を持つ遊びであるため、あやとり遊びによる脳の活性化効果が認められている[14][21][25]。更に、あやとりを遊ぶことによって幼児の指先の操作能力および数学的能力が養われる可能性があるとの報告もされている[16][19]。そのため、文化の継承および幼児の発達のためという理由から、全国の多くの幼稚園・保育園にてあやとりが実施されている[15]。

あやとりでは、あらかじめ決まった手順に従って紐を操作する。紐を操作する手順はあやとり技によって数手から数十手にまでおよぶため、あやとり技の習得には時間がかかる。一人であやとりを習得する場合の教材として、主に書籍および動画が挙げられるが、あやとりは両手を同時に扱う遊びであるため、あやとりをしながら途中で書籍のページを捲ったり、動画を再生/停止することが難しい[19]。そのため、あやとり技の作成段階を認識することによって、自動で次の紐の操作を指示する教材が求められる。

そこで今回我々は、導電纖維を編み込んだ紐を用いたあやとり技認識システム：E-String Figuresを開発した。本システムは、導電纖維を編み込んだあやとり紐（以下、導電性あやとり紐）の電気抵抗値を、10本の指の甲部分に電極を装着した手袋を用いて測定することによって、あやとり技を認識する。認識に導電纖維を編み込んだ紐を用いるため、カメラを用いるあやとり技認識手法とは異なり、環境光およびオクルージョンの影響を受けない[20]。また、導電性あやとり紐は容易に作製できるため、ユーザの手の大きさに合わせて紐の長さを調整できる。

2 関連研究

本システムでは、導電纖維を編み込んだ紐を用いてあやとりをした際に、紐同士が絡まった接觸点において短絡が生じることによって、紐の電気抵抗値が変化する現象を用いてあやとり技を認識する。したがって、本節では紐状インタフェース、およびあやとり技の認識手法に関する関連研究を述べる。

2.1 紐状インタフェース

各種センサを搭載した紐、あるいは特殊な内部構造を持った紐へのジェスチャ（タッチ、ねじる、結ぶなど）を認識し、それをコンピュータへの入力に活用するインタフェースが研究されている。既存の紐にセンサ（圧力センサ、抵抗式ポジションセンサ、曲げセンサ、光学式センサ）を搭載することによりジェスチャ認識するもの[1][6][7][9][10][22]、スリットが入った光ファイバを内蔵したもの[2][3][12]、チューブ内に複数個のコイルを内蔵し、コイル同士の相互誘導による起電力からチューブの変形を認識するもの[13]、およびコイル状になった銅線および導電纖維をシリコンを用いて挟んで多重に配置し、コイルの変形によってジェスチャ認識するもの[11]などが提案されている。

また、導電纖維を編み込むことによって作製され

Copyright is held by the author(s).

* 筑波大学 情報理工学位プログラム

† 日本学術振興会

‡ 筑波大学 システム情報系

§ 神戸市立工業高等専門学校 電子工学科

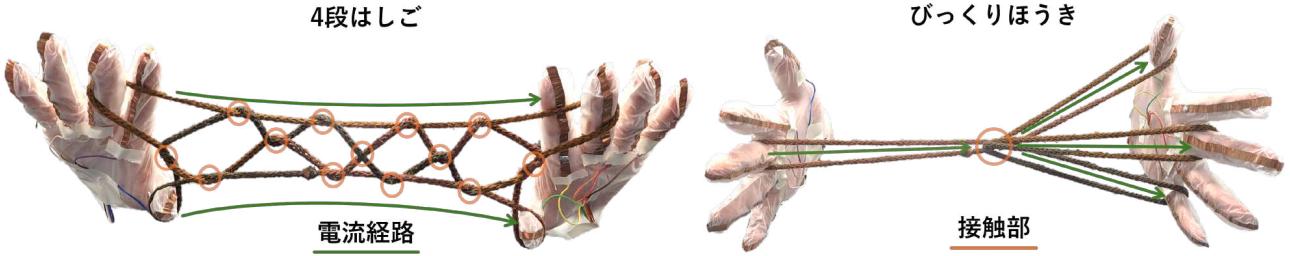


図 1. 導電性あやとり紐の技(絡まり方)ごとの紐同士の接触点および紐中の電流経路.

た紐状インターフェースも研究されている。導電纖維を編み込んだ紐を人体、または導電性を持たせた毛糸、クリップおよびフックなどのアイテムを経由して接地することによって紐へのタッチ位置の検出を行うもの[9][26]、皮膜付き導電纖維を編み込んだ組紐の静電容量の変化を用いてジェスチャ認識するもの[8][9]などが提案されている。更に、穂積ら[23]は、導電纖維を編み込んだ組紐の、引っ張りおよび無ずんだ際の電気的特性を評価することによって、この組紐をタンジブル・インターフェースとして活用するための基礎検討を行っている。

これらの研究と比較して、我々の研究は、導電纖維を編み込んだ紐の絡まり方による抵抗値変化を用いて、あやとり技を認識するという点で異なる。

2.2 あやとり技の認識手法

あやとり技を認識する手法として、カメラを用いる手法[20]、および導電纖維を用いる手法[5]が提案されている。

One-String [20] はスマートフォンのフロントカメラを用いて「びっくりほうき」の制作4工程における紐形状を認識している。認識にユーザがあやとりをしている様子を撮影したカメラからの映像を用いるため、特殊な構造をした紐やセンサを必要としないが、環境光およびユーザの服の柄の影響を受ける。この影響を除くために映像からあやとりの部分をトリミングしているが、その精度は約6割との報告がなされている。また、映像にあやとりが写り続けるようにするためにユーザは常にカメラの前に滞在する必要がある。

String Figuring [5] は、我々のシステムと同様に、あやとり紐同士が絡んだ接触点において導電纖維同士が短絡することにより、紐の抵抗値が変化する現象を用いて紐形状を推定している。この手法では、毛糸および導電纖維と共にリリアン編みすることによって作製した中空構造のチューブに、芯材として表面に絶縁加工が施された針金を通した紐が、センサとして用いられている。チューブおよび芯材の間に、紐全体を5等分する位置に銅箔電極が配置されている。また、銅箔および回路を接続するための信号線もチューブおよび芯材の間に内蔵されている。

本システムではチューブの抵抗値変化から紐同士の接触点および曲げなどの紐形状を推定できる。しかし、紐の中心に針金が通されているため紐が固く、かつ紐の端から電圧測定用の信号線が伸びているため、紐を複雑に絡ませることができない。

これらの手法と比較して提案システムは、導電纖維を編み込んだ紐を用いるため、環境光およびオクルージョンの問題を受けない。また、指に装着した電極を用いてあやとり紐の抵抗値を測定するため、紐が柔軟であり、複雑な技を作ることも可能である。

3 E-String Figures

導電纖維を編み込んだ紐を用いたあやとり技認識システム：E-String Figures を示す。以下に、本システムの原理およびシステム構成を示す。

3.1 原理

あやとりを行うと図1に示されるように紐同士が様々な位置で絡まる。その絡まり方(紐上の位置=接触点、一箇所で絡まる本数など)はあやとり技毎に異なる。また各指が触れる紐上の位置もあやとり技毎に異なる。

そこで我々はあやとり技を認識するために、導電纖維を編み込んだ紐と、各指に取り付けた電極を用いる。あやとり技が異なれば、図1に示されるように、あやとり技毎に異なる短絡が生じ、また、各指が触れる紐上の位置も異なる。このため、電極間の電流経路も異なる。この違いは、ある指の電極に電圧を印加すると、他の指の電極において計測される電位の違いとなって現れる。

本システムは両手10本の指に設けた電極の電極間の抵抗値を測定し、得られた測定値から機械学習によってあやとり技を認識する。

3.2 システム構成

あやとり技を認識するためのシステム構成要素を述べる。本システムは、紐に流れる電流を測定するシステム、および測定値からあやとり技を認識するためのソフトウェアから構成される。



図 2. 導電纖維編み込み紐.



図 3. 銅テープを貼り付けたした手袋.

3.2.1 紐電流測定システム

紐に流れる電流を測定するシステム構成を以下に述べる。

導電纖維編み込み紐

作製した導電纖維編み込み紐を図 2 に示す。我々はこの紐を、導電纖維（フジックス, Smart-X）3 本および非導電性の刺繡糸を束ねたものをアイコードニッター（AZ-Front, YL-508）を用いてリリアン編み [5][18] することによって作製した。紐の電気抵抗値は $40.0 [\Omega/m]$ となった。

測定回路および測定用プログラム

作製した手袋を図 3 に、抵抗値を測定するための回路を図 4 に示す。本システムでは指の甲部分に電極として銅テープを貼り付けた手袋を用いることによって電極間の紐の抵抗値を測定する。また図 4 の回路はこの抵抗値を測定するための分圧回路である。

電極はマイコン（Arduino Mega）のデジタルピンならびにマルチプレクサ（SparkFun, BOB-09056）に並列に接続されている。マルチプレクサの Sig ピンは 270Ω の抵抗器を介して接地されており、かつマイコンのアナログピンに接続されている。

マイコンで動作するプログラムは、まず 10 個の電極に順番にデジタルピンから電源として 5V の電圧を印加する。ある 1 個の電極に電圧を印加している間に、残りの 9 個の電極をマルチプレクサを用いて順に GND と接続した上でアナログピンの電位を記録する。したがって、 10 (通りの印加電極) $\times 9$

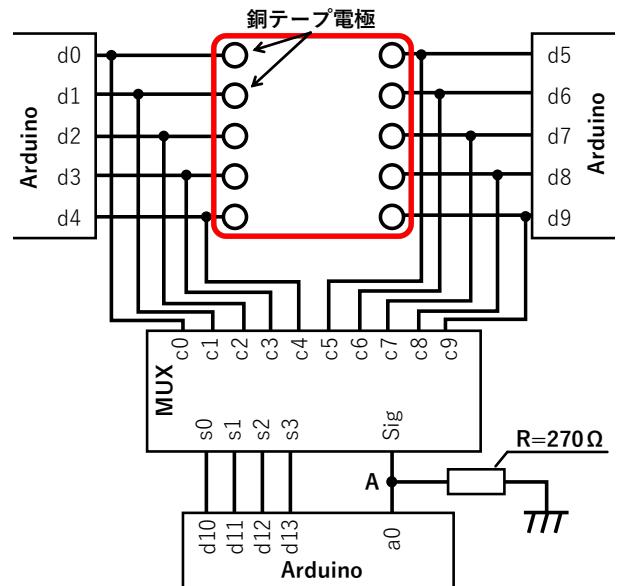


図 4. 電流測定回路。図中では 1 個の Arduino を 3 つに分割（左上部、右上部、下部）して表記している。

(通りの測定用電極) = 90 [個] の測定値が得られる。この 90 個の測定値は、シリアル通信を用いて PC に $100 [Hz]$ の頻度にて送信される。

3.2.2 ソフトウェア

あやとり技を認識するために、PC 上で動作するソフトウェアを実装した。本ソフトウェアは認識に RandomForest 分類器を用いる。あやとり技を作り上げる過程にて生じる紐への操作（取る、外す、入れ替えるおよび掛け替えるなど）を一回終えた後の状態を 1 工程とし、各工程における紐形状を認識する。各工程の紐形状ごとに 90 個の測定値を取得し、これを説明変数、各技の何工程目であるかを目的変数としてあやとり技認識モデルを生成する。

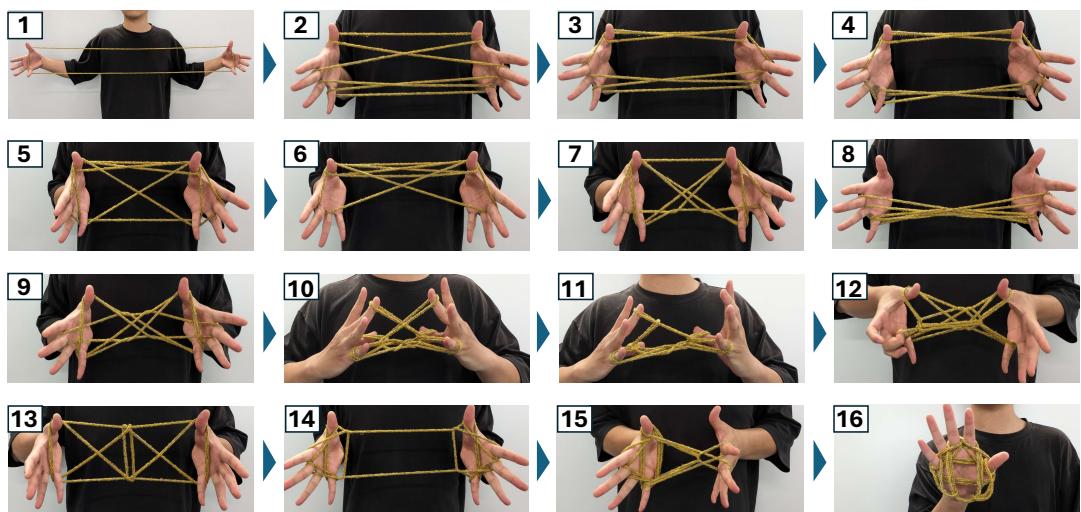
4 検証

本システムによるあやとり技の認識精度を調査するために実験を行った。実験では、図 5 に示す連続技と呼ばれる技 3 種類の各変形工程における紐の電圧を測定した。

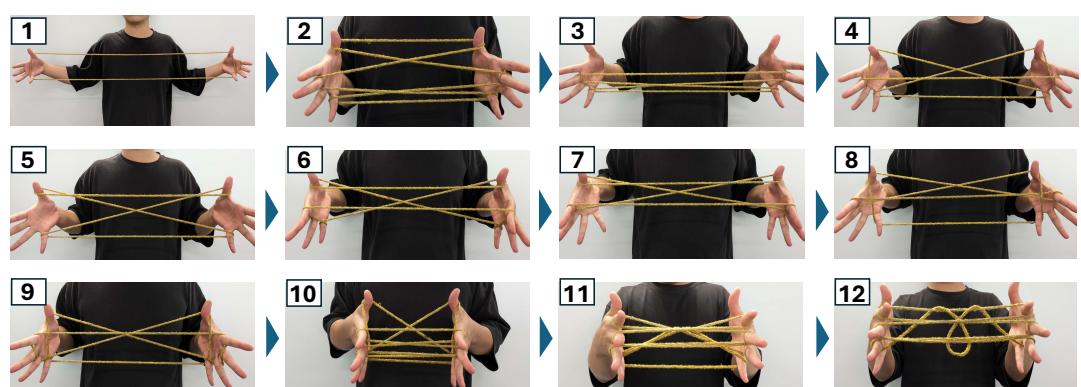
4.1 あやとり変形工程の定義

山田ら [17] は、あやとりの基本動作は、「取る」、「外す」および「入れ替える」から構成され、あやとりの複雑な図形はこの基本動作の組み合わせから表現できると述べている。「取る」は紐を指を用いて取る動作、「外す」は指にかかっている紐を外す動作、および「入れ替える」は同一の紐にかかっている 2 本の紐において、内側の紐および外側の紐を入れ替える動作であるとそれ述べている。山田らの定

連続技1（くり、すべり台、かめ、かみしばい、ゴム、飛行機、かぶと）



連続技2（ふじさん、つき）



連続技3（4段はしご、東京タワー）

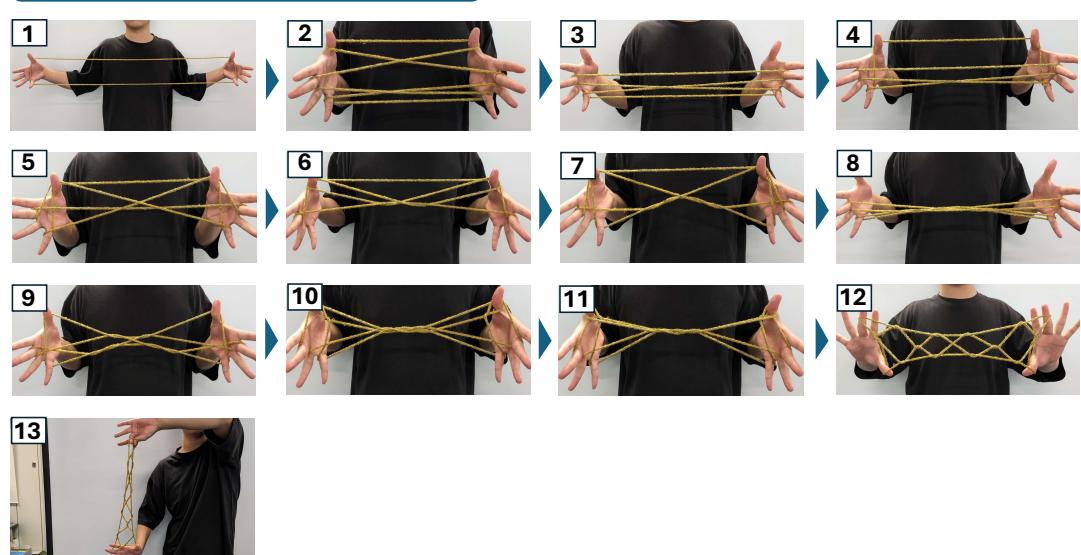


図 5. 各連続技の変形工程.

義した基本動作は、あやとりの複雑な動作を構成する最小要素であり、すべて片手の指一本の動作を表す。この基本動作を基に、実験にて認識する対象となる連続技の変形工程を実験にて測定するあやとり技の変形工程を定義した。実験では、あやとり操作の簡略化のために、いくつかの基本動作を組み合わせたものを新たな動作として定義した。まず、一本の指にかかっている紐を取り、かつその指から紐を外す動作をまとめて「掛け替える」動作とした。また、「入れ替える」動作は同一の指にかかっている2本の紐の上下を入れ替える動作であり、この動作を完了した後の紐形状は完了する前のものと比べて大きく変化せず、測定値も変化しない。そこで、本論文では基本動作である「入れ替える」を「ナバホ取り」として扱った[17]。「ナバホ取り」は「入れ替える」動作および「外す」動作を組み合わせた動作、すなわち、同一の指にかかっている2本の紐の内、下の方の紐のみを外す動作である。加えて、同一の指に複数の紐が掛かっており、これらの紐に対して同じ動作を繰り返し行う場合、これらの動作をまとめて一つの動作とする。更に、左右対称の紐形状から対象の位置にある紐同士に対して、両手の同一の指を用いて連続していざれかの動作を行う場合、これらをまとめて一つの動作とする。

これらのうち、いざれかの動作を終えた後、紐の弛みを無くした状態を1工程と定義する。しかし図5に示した連続技1の14工程目から15工程目では、左手の親指から右手の親指、および左手の小指から左手の親指に紐を「掛け替える」動作をまとめて1工程とした。これは、親指または小指のどちらかのみを用いて動作を完了した後、紐を張るのが困難であったためである。また、連続技1の15工程目から16工程目、および連続技3の11工程目から12工程目も同様の理由によって両手の4本の指の動作をまとめている。更に、連続技2の12工程目、「つき」は紐をあえて弛ませることによって、紐图形中心に小さな輪を浮かべる工程である。「つき」のみ、この紐がたるんだ状態を1工程とした。

4.2 実験手順

研究室内外の9名の実験協力者（男性8名、女性1名、平均年齢=23.0歳、SD=0.93）に対して実験を行った。図3にて示した紐電流測定用手袋を実験協力者が装着した後に、著者が実験協力者に紐の操作を教えながら測定を開始した。3つの連続技の各工程につき50回（約3秒間）測定値を記録し、各連続技のすべての工程を終えた後一度紐を指から外すように指示した。これを1セッションとし、一つの連続技を5セッション続けて測定した。これを3つの連続技にて行い、それぞれの変形工程は連続技1は16工程、連続技2は12工程、連続技3は13工程であるため、一人当たり（16+12+13）[工程]

表 1. LOSO 交差検証による F 値。

実験協力者	連続技 1	連続技 2	連続技 3
P ₁	0.94	0.98	0.90
P ₂	0.96	1.00	0.87
P ₃	0.95	1.00	0.89
P ₄	0.73	1.00	0.75
P ₅	0.92	1.00	0.81
P ₆	0.65	0.94	0.79
P ₇	0.87	0.98	0.84
P ₈	0.92	1.00	0.79
P ₉	0.92	0.99	0.78
P _{ave}	0.87	0.99	0.82

×50 [回] ×5 [セッション]=10,250 [個] のデータが収集された。なお、測定中は実験協力者に「紐をぴんと張った状態」を保つように指示した。

4.3 実験結果

9人の実験協力者から得られたデータを用いて、RandomForest 分類器を生成し、各技の変形の認識精度を算出した。各実験協力者における、各連続技ごとの Leave-one-session-out (LOSO) 交差検証の結果を表1に、各連続技ごとの Leave-one-participant-out (LOPO) 交差検証の結果の混同行列を図6にそれぞれ示す。LOPO 交差検証の結果、各連続技のF値は、それぞれ連続技1:0.91、連続技2:0.98、連続技3:0.82であった。

4.4 考察

結果から、本システムは全ての連続技の変形を8割以上の精度で認識できることがわかった。したがって、導電纖維を編み込んだ紐の抵抗値変化を用いることによって、あやとり技の変形工程を高精度に認識することが可能であると言える。LOSO 交差検証によって得られたF値、およびLOPO 交差検証によって得られたF値の間に大きな差が見られないことから、本システムによって得られる測定値間には個人差による影響が少ないと考えられる。したがって、本システムを用いたあやとり習得支援システムは、ユーザごとに分類器のキャリブレーションを行う必要がないと考えられる。

また、連続技3が他2つの連続技と比較して精度が低くなった原因として、形状は異なるが、得られる測定値が似ている紐形状が、他2つの連続技に比べて多く含まれていることが挙げられる。図6cの混同行列を見ると、連続技3の10工程目を、7工程目および11工程目の形状と誤認識した回数が多いことがわかる。図5に示した変形工程を見ると、各工程は、両手の親指、中指および小指に紐をかける形状であるという点において共通していることがわかる。そのため、この6本の指の電極が紐を通して

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
True	99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	96	2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	4	97	2	0	2	0	0	0	0	0	0	0	0	0
4	0	0	0	0	77	2	10	0	3	0	0	2	0	3	0	0
5	0	0	0	0	1	98	1	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	14	0	70	2	16	0	0	0	1	0	0
7	0	0	0	0	0	4	98	0	0	0	0	0	0	0	0	0
8	0	0	0	0	6	0	12	0	80	4	0	0	0	0	0	0
9	0	0	0	0	0	0	0	95	5	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	92	9	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	3	86	2	0	0	0	0	0
12	1	0	0	0	0	0	0	0	0	4	87	2	0	0	0	0
13	0	0	0	0	0	0	0	0	2	93	2	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	98	2	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0

	1	2	3	4	5	6	7	8	9	10	11	12
True	100	0	0	0	0	0	0	0	0	2	0	0
1	0	100	0	0	0	0	0	0	0	0	0	0
2	0	0	100	0	0	0	0	0	0	0	0	0
3	0	0	0	100	0	0	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0	0	0
5	0	0	0	0	0	100	0	0	0	0	0	0
6	0	0	0	0	0	0	100	0	0	0	0	0
7	0	0	0	0	0	0	0	100	0	0	0	0
8	0	0	0	0	0	0	0	0	100	0	0	0
9	0	0	0	0	0	0	0	0	0	100	0	0
10	0	0	0	0	0	0	0	0	0	0	100	0
11	0	0	0	0	0	0	0	0	0	0	0	100

	1	2	3	4	5	6	7	8	9	10	11	12	13
True	99	0	0	0	0	0	0	0	0	0	0	0	0
1	0	100	1	0	0	0	0	0	0	0	0	0	0
2	0	0	99	0	0	0	0	0	0	0	0	0	0
3	0	0	0	97	1	0	0	0	2	0	0	0	0
4	0	0	0	3	94	0	0	0	1	0	0	0	0
5	0	0	0	0	0	78	0	0	0	0	0	16	10
6	0	0	0	0	0	58	0	7	16	11	0	0	0
7	0	0	0	0	0	100	0	0	0	0	0	0	0
8	1	0	0	1	0	0	4	0	89	1	7	0	0
9	0	0	0	0	1	27	0	1	54	27	0	0	0
10	0	0	0	0	3	0	11	0	0	29	55	0	0
11	0	0	0	0	0	13	0	0	0	0	71	17	0
12	0	0	0	0	0	9	0	0	0	0	13	73	0

図 6. LOPO 交差検証の結果の混同行列 (a: 連続技 1, b: 連続技 2, c: 連続技 3).

て接続され、かつ得られた測定値も類似したと予想される。しかし、連続技 3 の 7 工程目および 10 工程目は、連続した変形工程ではない。あやとりはあらかじめ決まった手順を順番に実行して技を作る遊びであるため、直前の工程を正しく認識することができれば、その工程を基に次の変形工程を推測することによって、紐形状の認識精度を向上できると考えられる。したがって、本システムにおける、連続していない工程間にて生じる誤認識の影響は小さいと考えられる。

5 今後の課題

技ごとの分類モデル生成の必要性

本システムでは 3 種類の連続技を認識したが、それ以外の多くのあやとり技を認識するためには、それぞれのあやとり技ごとに分類モデルを新たに生成する必要がある。そのため、あらゆるあやとり技を認識するにあたって大量のデータを収集する必要がある。なお、実験の結果、本システムによって得られた測定値には個人差がなく、他人のデータを用いて生成した分類器においても高い精度であやとり技を認識できることがわかった。したがって、本システムでは認識できるあやとり技の数を容易に増やせる可能性がある。

誤った手順の指摘

本システムは、あやとりの初心者のあやとり習得を支援するためのアプリケーションへの応用を想定している。初心者はあやとりに習熟した者とは異なり、あやとりを習得する過程において、あやとり紐の取り方を間違える可能性がある。しかし、本システムでは、あやとり技の正しい変形工程のみを認識したために、各変形工程にて生じた紐の誤操作を検出できない。したがって、本システムには、初心者があやとり紐の操作を誤った際にその誤りを検知し、あやとり紐の正しい操作を伝える機能が求められる。

これを達成するために、機械学習における外れ値

を検知するためのアルゴリズムである、One Class SVMなどを本システムと組み合わせて用いることによって解決できる可能性がある。したがって、使用的機械学習モデルの工夫によって誤った手順を指摘できるかどうか調査する必要がある。

あやとり習得支援アプリケーションの制作

実験によって、本システムはあやとり技の変形工程を高い精度で認識できることがわかった。今後は、本システムを応用した、あやとりの習得支援アプリケーションを制作する。本アプリケーションは、あやとりの変形工程を認識し、次の紐の操作を自動で指示することによってユーザのあやとり習得を支援する。そのためこのアプリケーションには、書籍および動画などの従来のあやとりの教材とは異なり、あやとりの操作中に両手が塞がってしまうことによってページを捲ったり、動画を再生/停止することができないという問題がないと考えられる。

6 まとめ

本論文にて導電繊維を編み込んだ紐を用いてあやとり技を認識するシステムである E-String Figures を示した。9人の実験協力者に対して実験を行い、連続技と呼ばれる技 3 種類の制作工程の認識精度を調査した。実験の結果、本システムは連続技 1 の F 値が 0.91、連続技 2 の F 値が 0.98 および連続技 3 の F 値が 0.82 という精度で認識できた。また、本システムによって得られた測定値には、個人間の差がなく、システムをキャリブレーションせずに適用可能であることもわかった。今後は、本システムを応用したあやとり習得支援アプリケーションを制作する。

参考文献

- [1] G. Blasko, C. Narayanaswami, and S. Feiner. string-based interaction techniques for dual-display mobile devices. In *Proceedings of the*

- SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pp. 369–372, New York, NY, USA, 2006. Association for Computing Machinery.
- [2] V. Cheung, A. K. Eady, and A. Girouard. Deformable Controllers: Fabrication and Design to Promote Novel Hand Gestural Interaction Mechanisms. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '18, pp. 732–735, New York, NY, USA, 2018. Association for Computing Machinery.
- [3] N. Fellion, A. K. Eady, and A. Girouard. FlexStylus: A Deformable Stylus for Digital Art. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, pp. 2482–2489, New York, NY, USA, 2016. Association for Computing Machinery.
- [4] INAX ギャラリー企画委員会. 世界あやとり紀行 : 精霊の遊戯. INAX booklet. INAX 出版, 2006.
- [5] J. Klefeker and L. Devendorf. String Figuring: A Story of Reflection, Material Inquiry, and a Novel Sensor. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, pp. 1–6, New York, NY, USA, 2018. Association for Computing Machinery.
- [6] E. Koch and H. Witt. Prototyping a chest-worn string-based wearable input device. In *Proceedings of the 2008 International Symposium on a World of Wireless, Mobile and Multimedia Networks*, WOWMOM '08, pp. 1–6, USA, 2008. IEEE Computer Society.
- [7] K. Nakagaki, S. Follmer, and H. Ishii. Line-FORM: Actuated Curve Interfaces for Display, Interaction, and Constraint. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pp. 333–339, New York, NY, USA, 2015. Association for Computing Machinery.
- [8] A. Olwal, J. Moeller, G. Priest-Dorman, T. Starner, and B. Carroll. I/O Braid: Scalable Touch-Sensitive Lighted Cords Using Spiraling, Repeating Sensing Textiles and Fiber Optics. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, pp. 485–497, New York, NY, USA, 2018. Association for Computing Machinery.
- [9] P. Schoessler, S.-w. Leigh, K. Jagannath, P. van Hoof, and H. Ishii. Cord UIs: Controlling Devices with Augmented Cables. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '15, pp. 395–398, New York, NY, USA, 2015. Association for Computing Machinery.
- [10] J. Schwarz, C. Harrison, S. Hudson, and J. Mankoff. Cord Input: An Intuitive, High-Accuracy, Multi-Degree-of-Freedom Input Method for Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 1657–1660, New York, NY, USA, 2010. Association for Computing Machinery.
- [11] F. Shahmiri, C. Chen, A. Waghmare, D. Zhang, S. Mittal, S. L. Zhang, Y.-C. Wang, Z. L. Wang, T. E. Starner, and G. D. Abowd. Serpentine: A Self-Powered Reversibly Deformable Cord Sensor for Human Input. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–14, New York, NY, USA, 2019. Association for Computing Machinery.
- [12] C. To, T. L. Hellebrekers, and Y.-L. Park. Highly Stretchable Optical Sensors for Pressure, Strain, and Curvature Measurement. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5898–5903, 2015.
- [13] F. Wu, J. Jeon, S. K. Moon, H.-J. Choi, and H. Son. Voice Coil Navigation Sensor for Flexible Silicone Intubation. *IEEE/ASME Transactions on Mechatronics*, 21(2):851–859, 2016.
- [14] 野田 さとみ, 佐久間 春夫. 手指の運動を伴う遊びにおける脳波および覚醒度・快感度の変化について. バイオフィードバック研究, 36(1):41–46, 2009.
- [15] 稔丸武臣, 丹羽孝, 勅使千鶴. 日本における伝承遊び実施状況と保育者の認識. 名古屋市立大学大学院人間文化研究科『人間文化研究』, 7:57–78, 2007.
- [16] 岡部恭幸. 幼児期において Disembedding を指導することの意義と可能性. 神戸大学大学院人間発達環境学研究科研究紀要, 特別:47–52, 2016.
- [17] 山田雅之, R. Budiarto, 伊藤英則, 世木博久. アヤトリにおけるひも图形変形過程の表現とその処理. 情報処理学会論文誌, 35(3):497–504, 1994.
- [18] 山本紀久子. 指ねじり編みによるアクリルたわしの教材化. 茨城大学教育実践研究, (27):105–112, 2008.
- [19] 松本典子. 伝承あそびと運動感覚. 鳥取短期大学研究紀要, (48):79–86, 2003.
- [20] 大橋鍊, 塚本啓太, 杉本隼斗, 濱川礼. One-String: CNN を用いたあやとりの工程認識手法と競技性を付加したあやとりゲームの実現. 研究報告エンタテインメントコンピューティング (EC), 2021(18):1–5, 2021.
- [21] 中村和也, 武持敬太, 東登志夫, 植原淳, 田中浩二, 小河原格也. 通所事業所における手工芸活動に関する実態調査. 長崎大学医学部保健学科紀要, 16(2):115–120, 2003.
- [22] 田實佳郎, 高橋綾子. Future Vision: 私の未来学圧電性 PLLA によるウェアラブルセンサーの開発: 組紐状の首飾りでストレスなく生体情報を収集: 関西大学システム理工学部 電気電子情報工学科 田實佳郎教授. コンバーテック, 45(7):10–14, 2017.
- [23] 穂積佳, 塚田浩二, 吉田博則. 導電糸を活用した組紐型タンジブル・インタフェースの基礎検討. 研究報告ヒューマンコンピュータインタラクション (HCI), 2024(24):1–7, 2024.

- [24] 野田さとみ. あやとりの形象変化に伴う脳内活動について: 近赤外分光法の結果から. 名古屋女子大学紀要. 家政・自然編, 人文・社会編, (58):179–185, 2012.
- [25] 野田さとみ, 佐久間春夫. 「あやとり」「折り紙」の学習過程における脳波及び心理的変化. バイオフィードバック研究, 37(1):29–36, 2010.
- [26] 劉悅怡樹, 串山久美子. Musicyarn: 導電糸組ひもを利用したタンジブル音楽システム. インタラクション 2019 論文集, pp. 352–365, 2019.

ボタンひとつで3Dプリントが体験できるシステムの提案と運用

高橋 治輝* 白井 義人* 松村 耕平*

概要. 新しい技術を人に使ってもらうことは難しい。ましてや研究段階のシステムとなるとなおさらである。本稿では、3Dプリンタに興味のない人たちに3Dプリントを体験させることを目的として、簡単なボタン操作のみで3Dプリンタのすべての工程を体験可能なシステムを提案する。ボタンを押すことで3Dプリンタは自動で造形を開始し、造形中の造形速度を変化させる。システムを1ヶ月間に渡って運用したところ、オペレータ不在の状態であるにも関わらず166回の造形が行われた。また、体験者へのインタビューからシステムによって3Dプリンタに関する理解や興味が深まったことが示された。

1 はじめに

3Dプリンタは、ここ数十年で著しく進歩したデジタルファブリケーション技術のひとつであり、2009年の熱溶解積層方式（FDM）3Dプリンタの特許満了を皮切りに、より扱いやすく安価になった。結果として、パーソナルファブリケーション [5] と呼ばれる潮流が生まれ、さまざまなトレンドも生み出した [8]。HCI分野でも数多くの研究があり [2]、たとえば、3Dプリンタを用いたデバイスの設計は主要な研究領域である [1]。2024年現在、Bambu Lab¹が手掛けた3Dプリンタが大きな注目を集めている。洗練されたマシンは、キャリブレーションやエラー検知の自動化、高速造形、自動素材供給システムによるカラー造形をも実現しており、TIME誌のThe Best Inventions of 2022に選出される成功を収めた。低速でしばしば造形に失敗する従来の3Dプリンタ像は、払拭されつつあると言って良いだろう。

こうした進歩がある一方で、依然として3Dプリンタを使用するまでにはさまざまな障壁がある。3Dプリンタがいくら扱いやすくなっても、自分が作りたいものの3Dデータ入手する作業が残っている。また、パラメータの選択や材料の選定といった、3Dプリンタに関する最低限の知識、安全な利用方法やトラブル対応についても知る必要があるだろう。さらに、3Dプリンタの購入や利用可能な場所へのアクセス、そして積極的にその環境を利用するという動機も必要である。いくら安価で扱いやすくなろうとも、その人が「3Dプリンタで作る」という動機を持っていない限りは使用に至らない。依然として3Dプリンタを使用したことのない人は多くおり、その結果、大学の施設やファブースペースが上手く活用されずに閑散としてしまうことが起こり得る。

本稿では、3Dプリンタを「どんな形であれ一度でいいから試してみる」ことが重要だという考えに



図 1. 提案システム。ボタン操作だけで3Dプリントが体験可能なシステムを大学の廊下で長期運用した。

基づき、興味がない・知らない人たちに3Dプリントを体験させるシステムを提案する。一度でも体験することができれば、そのユーザは3Dプリンタを使つたことがあると主張したり、知った気になれたりする。また、自分でも使ってみようという動機にも繋がるだろう。提案システムは、ボタン操作（押下、連打、長押し）のみで3Dプリンタの工程に関わることができる設計とした（図1）。ボタン操作によって、体験者は造形対象をルーレットで選び、造形中の3Dプリンタを応援し、造形物の仕上がりを値踏みできる。造形物はテトリスのブロックになっており、ホルダーに移す作業を組み込むことで造形物の取り外しを動機づける。これにより、事前の知識や3Dプリンタを管理するオペレータも不要な運用の仕組みを実現した。立命館大学の廊下で1ヶ月間の実運用したところ、166回の造形が行われ、ボタンは約2.7万回押下された。体験者に対して半構造化インタビューを実施し、システムの利用のされ方を明らかにした。

2 システムデザイン

2.1 3Dプリンタ体験システムの要件

先行研究や3Dプリンタの性質を踏まえ、3Dプリンタを体験させるシステムの要件を整理する。

Copyright is held by the author(s).

* 立命館大学

¹ <https://bambulab.com/>

単純かつ簡単に操作可能なインターフェース (D1) :

3D プリンタのユーザは、3D モデリングソフトウェア、造形データを作成するスライサ、3D プリンタのホストアプリケーションなどを横断的に使用する。こうしたツールの一部は初心者でも利用可能な “low thresholds” な設計 [9] になっているが、ソフトウェアの準備と習熟という環境構築に対する障壁は残っている。本稿では、3D プリンタに触れ合える体験を提供することを目指すことから、こうしたソフトウェアの使用を排除する。体験者にはボタンとシステムの状態を表示するディスプレイという単純かつ簡単に操作可能なインターフェースのみを与える。

事前準備をせずとも利用可能 (D2) : 3D プリンタを利用するためには、前述のようなソフトウェアを準備するか、環境構築された据え置き PC にアクセスしなければならない。そして、自分が行いたい造形に関して最低限の知識が求められ、作りたいものの 3D モデルを準備する作業も必要である。しかし、初心者はこうしたすべての段階で躊躇可能性があることが知られている [6]。提案システムでは、この事前準備すら排除しており、造形データは事前にパラメータ調整をした状態でシステム側が提供する。3D モデルの選択方法はルーレットになっており、ボタン押下によってランダムに選出される。

体験者へのフィードバック (D3) : 3D プリンタにはコンピュータ黎明期のバッチ処理 [2] のような側面があり、しばしば動作に関するフィードバックに欠けている。たとえば、温度の上昇など時間のかかる処理は何が行われているかがわかりづらく、ボタンを押しても何も反応しないように感じるという問題が生じ得る。そこで、温度の制御が必要な工程に至った際には、現在と目標の温度を示すインターフェースを表示する。さらに、温度を上昇させる命令を送信すると同時に、ベッド全体を周回するダミーの運動も加えた。これによって、3D プリンタがボタン押下に即座に反応したと見せかける。

オペレータ不在でも利用可能 (D4) : 3D プリンタに精通したオペレータは、初心者の 3D プリンタ利用において重要な役割を担っており、彼らから十分なサポートが得られないとき、初心者はさまざまな障壁に直面してしまう [6, 4]。マイカースペースなどであれば精通したユーザが多くいるが、こうした空間で居心地の悪さを感じる人もいることが報告されている [10]。提案システムは、オペレータが不在でも利用可能のように設計しており、かつ誰でもアクセス可能なオープンな場所で運用する。体験者が必要な作業を極力取り除き、もし作業を求める場合はその旨をディスプレイに表示する仕組みとした。

いつでも体験可能 (D5) : 造形中の 3D プリンタに対しては何も体験することがなく、完了するまで 3D プリンタを眺める時間になってしまう。Willis らが指摘するように、素材への直接的なはたらきかけ



図 2. 提案システムが output するテトリスのブロックと Rainbow Filament による色の変化。

のない工作機械は、伝統的な手芸・工芸で見られた創造的プロセスや材料の理解といった機会を損失している [13]。さらに、造形が完了したあとの出力物を回収しなければ、次の造形を体験することができない。提案システムは、いつでも体験可能な仕組みとして、ひとりのユーザに全工程を達成させるのではなく、その場に居合わせた体験者がその時に必要な作業に従事するという分担を目指した。そして、造形中の 3D プリンタに対して、ボタン連打によって造形速度が上昇する「応援」を実装した。つまり、造形の開始、造形中の速度変更、造形物の回収と次の造形の準備をした人がすべて異なる可能性がある。

造形失敗の回避 (D6) : 造形失敗は、造形が中断されるだけでなく、3D プリンタに幻滅してしまう原因になり得る。ハードウェアの整備不良、材料不足、物理的なトラブルなどさまざまな原因によって生じ、それらを突き止めて解決するためにも知識やスキル [11]、オンラインヘルプの有効活用 [7] が求められる。Benedetti らが、初心者向けペイントシステムの要件として最低限の品質を保証する Safe playground [3] が重要であると述べたように、造形の失敗を恐れずに体験できる仕組みが必要である。完全な失敗の回避は難しいが、提案システムでは、造形データは比較的造形が容易な形状にして、3D プリンタのベッドに PEI シートを採用して造形の安定性を高める予防策を講じた。このシートは高温時に溶解したプラスチックと強固に接着し、低温時に容易に剥がれるという特性がある。

プレイフルな操作・出力結果 (D7) : 事前に用意したデータを造形させるだけでは、3D プリンタに対する十分な関心が得られないかもしれない。また、造形失敗の回避のために単純な立方体や円柱などの造形物を選ぶと、出力物を魅力的に感じず、取り外し作業に従事してもらえないかもしれない。すなわち、提案システムで体験できる操作や造形結果は体験者の興味関心を引くものでなければならない。そこで、提案システムには、プレイフルな 3D プリンタ体験を目指したエンタテインメント要素を導入した。まず、ルーレットで選択される対象にはテトリスのブロックを採用し、3D プリンタとともにホルダーを展示する（図 2 左）。ブロックの形状は造形が容易であり、テトリスの進行具合に合わせてルーレットを止めたり、テトリスで遊ぶために出力物を回収したりする動機付けにもなる。そして、材料は造形

ボタンひとつで3Dプリントが体験できるシステムの提案と運用



図3. ディスプレイに表示されるシステム画面。ユーザのボタン操作や3Dプリンタの造形工程に応じて画面を書き換える。

を繰り返すたびに色が変化していく iSANMATE の PLA+ Rainbow Filament²を選択した(図2右)。さらに、ボタンの脇に配置した NFC リーダに学生証を乗せることで、自分が各操作を行ったことを記録して主張できる仕組み、造形物を回収したあとにその出来を値踏みできる工程を実装した。

2.2 造形工程ごとのデザイン

前述の要件に従ってデザインした提案システムの造形工程について述べる。対象となる造形工程は、造形開始、造形中、造形完了であり、ディスプレイの画面は図3のように遷移する。

造形開始: 提案システムは、ルーレットを表示し(図3: 造形開始)，体験者がボタンを押下する選択することで造形データが選択される。造形データは、テトリスの7つ基本形状に加えて、デバッグ用に用意した「R」の輪郭を押し出した形状の8種類である。なお、この「R」の造形データは、数日間の動作検証後に削除予定であったが、体験者の意見を反映してそのまま残した。詳細は考察を参照されたい。

造形中: 造形データが選択されると、提案システムはノズルとベッドの温度上昇を伝える表示(温度を示す2つのバーとそれぞれの目標温度を示すインジケータ)へと画面を切り替える。また、温度上昇命令を送信すると同時に、ノズルを原点へ移動させ、ベッドの周囲を2周分だけ周回させる。これによって、ボタン操作に即座に反応したように感じさせ、温度上昇中の待機時間を極力短く感じさせる。温度上昇が完了すると、提案システムは造形の情報を伝える画面へと切り替えて、造形速度を示すバーと進度を示すプログレスバー、ボタン連打で3Dプリンタを応援できる旨のメッセージを表示する(図3: 造形中)。造形速度は50%からはじまり、ボタン押下で10%だけ、最大でバーの最大値をはみ出して500%まで上昇する。また、操作せずに放置すると毎秒1%ずつ、50%まで減少する仕組みとした。

造形完了後: 造形が完了すると、提案システムは造形物をベッドから取り外す指示を表示する。3Dプリンタの冷却を待ち、取り外しが正しく行われたか確認させる必要があることため、複数の写真と説明文で構成された画面とした(図3: 造形完了)。画

面上部のメッセージは、温度が40度を下回るまでは「冷却中」、下回ったあとは「取り外しにご協力ください」と表示が切り替わる。すべての作業を終えたあと、ボタンを3秒間長押して次の画面へと進む。最後に、プリントした造形物の評価を行う画面が表示される(図3: 評価)。GoodとBadの項目の間を赤い枠が交互に移動し、体験者はボタンの長押しでいずれかの評価を下す。なお、この画面で放置し続けると自動的に開始画面へと戻る。また、評価項目の下部には、造形に関わった体験者情報表示するようにした。

3 評価実験

提案システムの体験者が、3Dプリンタに対して示す反応を知るために評価実験を実施した。

3.1 実験方法および環境

提案システムを立命館大学大阪いばらきキャンパスH棟5Fの廊下に設置し、通りかかった人たちに体験させる方法を採用した。このフロアには、同大学の情報理工学部(実世界情報コース)が配置されており、各研究室のメンバーや講義を受ける学生の往来がある。また、運用中にボタンを押した人を無作為に選んで半構造化インタビューを実施した。提案システムの通知機能を使用し、ボタンが押されたと同時に著者が体験者にインタビューを申し込み、合意が得られた人に対してのみデータを収集した。そのため、複数の体験者に対して同時にインタビューを行うことがあった。なお、システムにアンケートのQRコードを表示していたが、ほとんど回答されなかったため評価には用いないこととした。

3.2 データの収集と分析

提案システムは、選択された3Dモデル、ボタン押下の時刻と回数、各工程が完了した時刻をログに書き出す。分析の対象は、2024年7月11日から約1ヶ月後の8月9日までとする。造形中にトラブルが発生したもの(造形完了が記録されていないログ)は分析対象から除いている。インタビューの項目として、名前と所属、3Dプリンタの使用経験、システムをはじめて利用したか、システムの体験で3Dプリンタについて理解できたことはあるか、自分が作った感があるかを尋ねた。ボタンを押すタイミングは

² <https://www.isanmate.com/product/pla-plus-matte-rainbow-color-3d-filament>

体験者によるため、造形の状況や体験者の返答に応じてインタビュー内容を掘り下げ、適宜システムの補足説明を行った。インタビューの内容は合意を得た時点から録音を開始し、書き起こしてコーディングを行った。

3.3 結果

造形記録

図4に造形回数を示す。合計で166回造形され、平均は21回、95%信頼区間は15から26回である。また、実験開始時に用意した材料を7月29日に使い果たしたため、同種のフィラメントへ交換した。

各工程に要した時間を図5に示す。このデータには、提案システムの前を通りかかる人物がほとんどいないと思われる夜間の時間や造形に問題が生じた日付（次節を参照）も含まれる。そこで、夜間の時間として19時30分から9時³と8月4日、6日を除いたフィルタ済データも同図に掲載した。各ブロックの造形には、平均で2時間6分33秒かかっており、最短で1時間5分6秒、最長で2時間36分19秒であった。なお、「R」の造形は平均で23分5秒である。完了から取り外し、すなわち造形物を回収して図3の「造形完了」画面の作業を終えるまでに要した時間は、中央値が24分1秒、最長で17時間20分14秒であった。フィルタ済データでは、中央値が19分6秒、最長で14時間19分23秒である。取り外しから次の造形、すなわちルーレット画面が表示されている時間は、中央値が9秒、最長で13時間13分16秒であった。フィルタ済データでは、中央値が11秒、最長で3時間59分52秒である。

実験中に生じた問題

実験期間中、何度か造形を中断せざるを得なかつた問題が生じた。図6aは、体験者が造形物を残したまま次の造形を行ったため、ノズルと造形物の衝突および造形位置のずれが生じた。実験初日に起こった問題であり、これを受け取り外し作業のボタン押下を長押しにするなどの対策を講じた。図6bでは、途中まで造形した造形物を残してノズルが後方にずれてしまっており、同様な問題が実験期間中に2度発生した。調査したところ、これは3Dプリンタの電源が瞬間に落ちたためであることがかった⁴。図6cのように、X軸の端点を固定するネジに緩みが生じている様子も発見した。いずれの問題も、発見から数時間程度で対応し、造形失敗した造形物の回収と各種点検、システムの修正を行った。

³ 夜間の時間は、立命館大学の講義時間を参考に決定した。

⁴ 8月4、6日に実験実施地周辺で落雷による瞬間停電があった。図6bを発見する1時間程度前に報告されている。

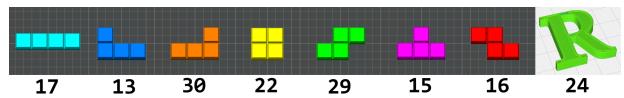


図4. 造形回数。約1ヶ月間で166回の造形が行われた。

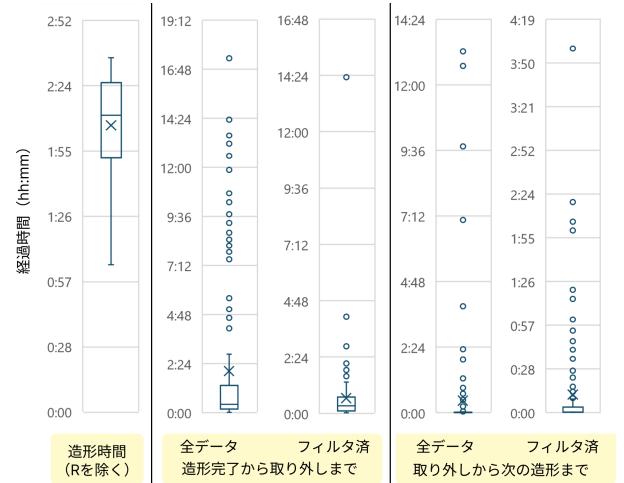


図5. 各工程に要した時間。フィルタ済みデータでは夜間と問題があった8月4、6日を除いた。

ボタンの操作記録

ボタン押下の合計回数は27,735回、その中で造形中の連打は26,173回であった。NFCリーダによって12名（著者らのものを除く）の教職員証および学生証番号が記録されていたが、26,769回は匿名状態での押下であった。造形速度に注目すると、最もボタンが多く押されたのは490%から500%と表示されているときであり、その回数は2,446回となっている。次点が50%から60%のときで817回、これら以外の造形速度のときはいずれも約400回であった。

一日あたりの押下回数が最も多かったのは7月16日の3,012回であり、大学の定期試験や夏期休暇期間となる翌週以降は押下回数が減少している。休日は押下回数が少なくなる傾向があり、一日あたりの最低回数を記録した7月21日（1回）、次いで少なかった7月28日（77回）はいずれも日曜日であった。

インタビュー結果

合計で10組に対して実施され、回答者は15名（男性13名、女性2名）すべて大学生であった。タイミングは、4組が造形中のボタン連打時、6組が造形完了および造形開始時であった。8名がシステムの体験以前に3Dプリンタの利用経験があり、9名がインタビューを受けた時点ではシステムを初めて体験したと回答した。システムへの反応は概ね好意的で、3名が外で展示されていて誰でも体験できる点やボタンを押すだけで参加できる点を評価した。

ボタンを押した理由として主だったものは、「ボタン押下を促すメッセージが表示されていたため」、

ボタンひとつで3Dプリントが体験できるシステムの提案と運用

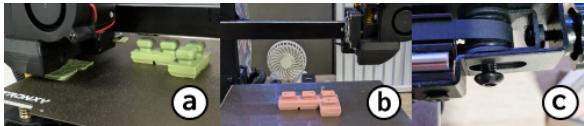


図6. 運用中に生じた造形エラー

「ボタンがあったらつい押してしまう」、「面白そうだと思ったため」である。システムを複数回体験している者の中には、「造形完了のタイミングを狙って押しに来ている」、「前を通りかかるたびに押している」といった回答もあった。1名（ペアでインタビューを受けた者の片方）が、「ボタン押下で不利益があると思って自分は押さなかった」と答えた。

3Dプリントを体験したことによる発見について、3Dプリント未経験者からは、「ボタン押下で造形速度が変化する」、「出力したものが意外と硬い」、「線を引くように内部を造形している」ことが理解できたと回答があった。3Dプリンタ利用経験がある者は、「造形中でも造形速度が変えられる」、「テト里斯の形状がサポート無しで造形可能である」、「色が変化する材料がある」ことを新しい発見として報告した。また、ボタンを連打しても造形速度の変化がわからなかつたという体験者が1名あった。

造形された（あるいは造形中の）ものに自作感があるか尋ねたところ、11名が「自作感はない」と回答し、4名がどちらとも言えない反応を示した（悩んだ末に不明瞭な返答があったか、一緒にインタビューを受けた者に同意した）。自作感の要因として最も重視されたのは、「自分がデザインしたものが作れるかどうか」であり、これに類するコメントが8名から得られた。また、「造形物を部品のように使って自分なりの作品ができれば自作感がある」、「関与したことで結果により大きな変化があれば自作感が得られる」という意見も得られた。

4 考察

4.1 システムの長期運用について

実験結果より、提案システムが十分に長期運用可能であることが確認された。この期間中、著者らはログや通知をチェックする以外の監視は行っていない。したがって、オペレータ不在で(D4)、いつでも体験可能(D5)な3Dプリント体験システムが実現できたと言えるだろう。造形失敗の回避(D6)について、図6で示した問題が生じたものの、週1回程度のメンテナンスと停電などへの対策を講じることで十分に対応可能であると考えている。また、3Dプリンタの造形エリア周辺を監視するカメラを設置することで監視体制を強化することも可能である。

4.2 3Dプリント体験の機会提供

インタビュー結果より、ボタンひとつで操作可能なインターフェース(D1)によって、事前準備なし

(D2)で3Dプリンタに触れる機会を提供できたことがわかった。さらに、ボタン押下で速度が変化するフィードバック(D3)によって、体験者は造形速度の変化について理解することができた。一方で、提案システムで3Dプリントを体験できたとしても出力物に自作感はない。これは提案システムの課題のひとつである一方で、3Dプリンタを知った・体験したことによって生じた興味や欲求であるとも考えられる。体験者からは「ブロックの形状を設定できると良い」、「絵に描いたものが出力できると面白い」といったコメントが得られている。また、こうした簡易的な3Dモデリング機能を組み込むのは容易だが、新たな手間や障壁が生まれてしまつては本末転倒であるため、本来の目的の「3Dプリント体験」に必要かを十分に検討すべきである。

造形記録の結果によれば、造形に約2時間、完了から回収に約20分、回収から次の造形までに約10秒の間隔がある。つまり、体験者が遭遇するのは造形中の3Dプリンタである場合が多く、ここに関わることができる機能を拡充することで体験を豊かにできる可能性がある。たとえば、「線を引くように内部を造形している」ことが理解できた体験者があつたため、一時的にノズルを寄せて造形の様子を見せる「観察」が有用かもしれない。また、「結果により大きな変化があれば自作感が得られる」という意見を採用し、ノズルの動きや押し出し量を変化させて造形物の表面に「エフェクト」を付与する手法[12]を適用する機能も考えられる。一方で、造形物の回収と造形開始がセットになっており、体験者が造形開始(ルーレット)画面にのみ立ち会う機会はほとんどなかったことが考えられる。造形開始時点には、3Dモデルの選択やノズルの加熱、1層目の造形など、3Dプリンタの理解に関わる要素が含まれているため、これら上手く体験させる仕組みが必要である。

4.3 自発的なオペレータの登場

提案システムの想定利用者は3Dプリンタに興味がない人たちであるが、インタビュー結果によると、利用経験がある人たちにも利用されていた。これは、ルーレットによる選択やテトリスという題材、ボタン押下による介入といったプレイフルな操作と出力(D7)によって、経験者にとっても新鮮な体験になったためだと考えられる。「前を通りかかるたびに押している」というコメントや各工程の時間間隔から伺えるように、こうしたユーザは頻繁にシステムにアクセスし、回収や造形中のボタン押下を行っている。また、図6の造形エラー内の1件は、頻繁に体験している者から報告されたものであった。彼らは自発的にオペレータの役割を担ってくれており、3Dプリンタの維持やメンテナンスにも貢献していた。

興味深いことに、テトリスブロックではない「R」の造形物はすべて持ち帰っていた。インタビュー

を実施していく中で、この「R」を「当たり」と捉えてルーレットを止め、いち早く回収することに面白さや提案システム利用の動機を見出している体験者がいることが明らかとなった（多い者は 10 個近く造形物を回収していた）。このような反応は意図していなかったが、3D プリンタに興味をもたせる方法のひとつとして非常に面白い知見であると考えている。ある種の競技性や運要素を取り込むことで、さらにシステムを発展させることができるだろう。

5 本研究の制約および展望

提案システムを設置した場所の都合上、体験者のほとんどが理系（情報系）の大学生だったという偏りがある。インタビューで得られたように、多くの体験者が3D プリンタの存在を知っており、約半数は使用経験もあった。また、実験期間中に大学のオープンキャンパスなどの行事があり、外部の体験者もあったと思われるが、ログから分析できた情報以外は得られていない。さらに、ボタン押下に抵抗を感じた者はそもそもインタビューの対象となっていなかったため、提案システムに対する意見を得られていない。また、NFC リーダで読み取った情報は十分に活用できておらず改善の余地がある。

展望として、提案システムで3D プリンタを体験したユーザをどのように次のステップに進めるか、という課題が考えられる。まず、システムに3D プリンタに関する情報や知識を提供する仕組みが実装可能である。造形データ（3D モデルや G-code）や3D モデル共有サイトの情報、使用される材料の知識などに関する Tips を各画面で表示することで、3D プリンタに対する興味をもたせることができるものかもしれない。そして、3D プリンタをどこで利用できるかを示すことも重要である。著者らの大学では、ものづくりスペース「KOBO」が運営されており、学生は3D プリンタやレーザーカッターなどが利用できる。こうした施設の存在を知らせれば、3D プリンタの体験から実践につなげることができるだろう。さらに、システムを体験した人同士でメッセージを残すなどのコミュニケーションをとれる仕組みが実現できるのではないかと考えている。交流の場としての利用が促進できれば、体験者間の共同制作などに発展させることができるだろう。

6 実装

6.1 ハードウェア

提案システムは、3D プリンタとボタン、ディスプレイ、小型コンピュータ（HP EliteDesk 800 G1 Mini PC）で構成されている。提案システムで使用する3D プリンタは、シリアル通信によって造形命令（G-code）のやり取りが可能であればどのような

3D プリンタでも構わない。本稿では、TRONXY⁵ Crux1 を使用した。ボタンはタクトスイッチと押下部分からなっており、押しやすさと印象を重視して大きめのもの（直径 60 mm）を採用した。また、ボタン内部の LED を点滅させるプログラムを書き込んだ Arduino を接続している。学生証を読み取るための NFC リーダには、SONY RC-S380/S を使用した。これらを3D プリンタで制作したケースに収納し、小型 PC の上部に設置した。

6.2 ソフトウェア

実装には Python を使用しており、ユーザインターフェースやボタンの入力処理の実装に Pygame、3D プリンタへの G-code の送信に Pyserial を用いた。提案システムは、造形対象となる G-code をリストとして読み込んだ後、その内容を3D プリンタに送信し続ける。造形中にボタンが押されたとき、システムはこのリストの先頭に造形速度を変更するための G-code（M220）を追加する。システムの運用および分析のため、Logger を用いてログを記録する。ひとつのログファイルは造形が開始された時点で生成され、ボタン操作、完了および取り外しのタイミングを記録する。また、ログの記録タイミングに合わせて Slack にメッセージを送信する機能も実装した。

6.3 3D モデル

テトリスの3D モデルは Rhinoceros と Grasshopper を用いて設計した。形状やホルダーに装着可能とする仕組みは、Thingiverse で共有されている Analog Tetris⁶を参考にした。3D モデルのスライスには Ultimaker Cura を使用し、基準となる造形速度を 50 mm/s に設定してスライスした。なお、各ブロックの造形には 100% の造形速度で約 80 分要する。

7 おわりに

3D プリンタに興味がない人を対象として、ボタンひとつで3D プリント体験が可能なシステムを提案した。体験者は、ルーレットを止めることで造形を開始したり、造形中の3D プリンタをボタン連打で応援したりする方法で3D プリンタに関わることができる。評価実験として、提案システムを誰でもアクセス可能な場所で 1 ヶ月間運用した。造形されたデータを記録し、体験者へインタビューでシステムの利用され方について明らかにした。さまざまな新興技術が利用可能になる一方で、それらを知らないかたり、体験する機会がなかつたりする人は依然として多い。本研究は、3D プリンタを題材として、そういった技術を「一度いいから試してみる」ための機会を提供するためのシステムを目指した。

⁵ <https://www.tronxy3dprinter.com/>

⁶ <https://www.thingiverse.com/thing:2974176>

参考文献

- [1] R. Ballagas, S. Ghosh, and J. Landay. The Design Space of 3D Printable Interactivity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(2), jul 2018.
- [2] P. Baudisch and S. Mueller. Personal Fabrication. *Foundations and Trends® in Human-Computer Interaction*, 10(3–4):165–293, 2017.
- [3] L. Benedetti, H. Winnemöller, M. Corsini, and R. Scopigno. Painting with Bob: assisted creativity for novices. UIST ’14, p. 419–428, 2014.
- [4] A. Berman, F. Quek, R. Woodward, O. Okundaye, and J. Kim. “Anyone Can Print”: Supporting Collaborations with 3D Printing Services to Empower Broader Participation in Personal Fabrication. NordiCHI ’20, 2020.
- [5] N. Gershenfeld, 田中浩也, 糸川洋. Fab—パソコンから3Dプリンタへ. オライリージャパン, 2012.
- [6] N. Hudson, C. Alcock, and P. K. Chilana. Understanding Newcomers to 3D Printing: Motivations, Workflows, and Barriers of Casual Makers. CHI ’16, p. 384–396, 2016.
- [7] N. Kwon, T. S. Sun, Y. Gao, L. Zhao, X. Wang, J. Kim, and S. R. Hong. 3DPFIX: Improving Remote Novices’ 3D Printing Troubleshooting through Human-AI Collaboration Design. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW1), apr 2024.
- [8] C. Mota. The rise of personal fabrication. C&C ’11, p. 279–288, 2011.
- [9] B. Schneiderman. Creativity support tools: accelerating discovery and innovation. *Commun. ACM*, 50(12):20–32, dec 2007.
- [10] D. Smit, G. Regal, and C. Gerdenitsch. Making Your Makerspace: A Tale of Tension. TEI ’24, 2024.
- [11] B. Subbaraman and N. Peek. 3D Printers Don’t Fix Themselves: How Maintenance is Part of Digital Fabrication. DIS ’23, p. 2050–2065, 2023.
- [12] H. Takahashi and H. Miyashita. Expressive Fused Deposition Modeling by Controlling Extruder Height and Extrusion Amount. CHI ’17, p. 5065–5074, 2017.
- [13] K. D. Willis, C. Xu, K.-J. Wu, G. Levin, and M. D. Gross. Interactive fabrication: new interfaces for digital fabrication. TEI ’11, p. 69–72, 2010.

MagElePaint: 多様な素材に簡単に回路プロトタイピングが可能な磁性導電性塗料

山岸 真人 * 高田 嶺介 †

概要. 従来の導電性塗料を用いた回路プロトタイピング手法では、接着剤等を介して電子部品を接続するため着脱が難しく、また塗布対象が紙などに限られていた。そこで本研究では、鉄粉とボンド等の乾燥により硬化する液体を混合した磁性導電性塗料「MagElePaint」を用いた回路プロトタイピング手法を提案する。MagElePaint は、インジェクタを用いて紙、布、皮膚など多様な素材に自在に塗布でき、磁性を有するため磁石を介して電子部品を簡単に着脱できる。本論文にて、MagElePaint の材料の種類や混合比、塗布対象の素材による、抵抗値の経時変化を調査した。さらに紙、布、皮膚上の回路プロトタイピングを試みた。

1 はじめに

電子回路を作製するためには、電子部品のはんだ付けや配線、基板の設計を必要とするため、初学者にとって難易度が高い。そのためブレッドボードを用いたプロトタイピングが広く用いられているが、サイズが限定的であり、どこが接続されているか見通しが悪い。そこで導電性塗料を用いて家庭用 2D プリンタやペンなどで回路の配線を描くプロトタイピング手法が研究されている [5][11][18][19]。導電性塗料は描いた箇所が配線となり導電性接着剤を介して電子部品と接続できる一方で、塗布できるものが紙などに限られる。また、ホワイトボード用途など壁面に磁性を付与するための磁性塗料が市販されている [9]。磁性塗料は水性/油性の基材を混合しており、平面に一様に塗布できるが、布や皮膚のように変形するものには適用が難しい。鉄粉とボンドを用いた磁性導電性塗料の研究 [17] は、特性評価などは行われておらず、柔軟性が低いなどの問題もあった。

そこで、本研究では紙や布、皮膚に回路を描くことのできる磁性導電性塗料「MagElePaint」を提案する。柔軟な磁性導電性塗料の研究 [13][14] は行われているが、MagElePaint を構成する素材はこれらの手法に用いられる素材よりも安価で、製作が容易である。MagElePaint は磁性/導電性を有する鉄粉を、ボンドや液体ゴムといった基材に混合した塗料である。また、塗料は他の金属粉末を添加することで抵抗値をコントロールできる。MagElePaint は、従来手法と比べて布や皮膚上でも回路プロトタイピング可能であり、さらに磁性を有するため磁石を用いて電子部品を容易に着脱できる。導電性塗料の回路端に磁石を埋めておいて素子を接続する方法や、紙の回路の裏側に磁石を貼り付ける方法も考え

られるが、それらの方法では素子を接続できる場所が磁石取り付け部のみに限定される。それに対し、MagElePaint は、塗布された塗料のどこでも磁石を介して接続可能である。本論文にて MagElePaint の材料の検討、および様々な材料上で回路プロトタイピングを行った結果を報告する。

2 関連研究

関連研究として、導電性塗料を用いた回路プロトタイピング、布上の回路プロトタイピング、皮膚上の回路プロトタイピングを述べる。

2.1 導電性塗料を用いた回路プロトタイピング

導電性塗料を用いて回路プロトタイピングを行う研究が挙げられる [5][10][11][13][14][18][19]。Instant inkjet circuits[5] は、家庭用の 2D プリンタで紙に印刷可能な導電性の塗料である。この手法はプリンタを用いて正確に紙に印刷できる一方、紙以外の素材では利用できない。これらの手法に対し、MagElePaint では、紙や木材以外にも布や皮膚など様々な場所に塗布可能である。

また、Matsuhisa らの研究 [10] および Electric Paint[2] は、柔軟な素材に塗布可能な導電性ベーストである。これらの手法に対し、MagElePaint は磁石を介して容易に電子素子を接続できる点で異なる。

Pradhan らの研究 [13] や Singh らの研究 [14] では、磁性/導電性を有する柔軟な塗料を製作しており、柔軟な素材への塗布が可能である。さらに、Pradhan らの研究 [13] では、磁石を介して容易に電子素子を接続できる。これらの手法で用いる素材に対し、MagElePaint で用いる素材（液体ゴム、鉄粉、銅粉）は安価であり、容易に磁性導電性塗料を作製できる点で異なる。さらに、MagElePaint は粘度が高く分厚く塗布することが可能である。分厚く塗布した塗料は押下などで断面積などが変化するため、センシングできる点で異なる。

Copyright is held by the author(s).

* 神戸市立工業高等専門学校 専攻科 電気電子工学専攻

† 神戸市立工業高等専門学校 電子工学科

2.2 布上の回路プロトタイピング

布上に回路プロトタイピングを行う研究が挙げられる。Sketch&Stitch[4], Lovell らの研究 [7], および LilyTiny in the Wild[8] は、導電纖維を布に縫い付けることによる回路プロトタイピングである。布素材に自在に回路を設計できる一方、回路を縫い付ける必要があるため回路の拡張や電子部品の着脱は容易ではない。

一方、MagElePaint では、塗料自体に磁性があるので、電子部品接続用の追加手順を必要とせず、磁石を介して電子部品を容易に着脱できる。

2.3 皮膚上の回路プロトタイピング

皮膚上に導電性塗料を塗布し、回路を構成する研究が行われている。BodyPrinter[3], Soft Inkjet Circuits[6], Body-Stylus[12], Lotio[15] は、皮膚上で回路プロトタイピングを行う研究である。構成された皮膚上の回路では、塗布した部位またはその周囲を動かすことによって抵抗値などが変化する。

これらの研究は、電子部品や計測用回路をスナップボタンや接着剤を介して接続するため、素子の付け外しが困難である。MagElePaint は磁性を有するため、磁石によって簡単に電子部品や回路を着脱できる。

3 MagElePaint の検証

MagElePaint は、磁性・導電性を有する鉄粉を、ボンドやゴム、シリコンといった紙や布に塗布可能であり硬化すると柔軟性を持つ基材に混ぜた塗料である。また、鉄粉よりも導電性の高い銅粉や銀メッキ粉末を混合することで導電性を高めることができる。一方で、金属粉末の酸化等により抵抗値が上昇する問題がある。

そこで、MagElePaint の素材の種類や、素材の混合比による導電性の経時変化特性を調査する。まず、基材による特性の違いを調査するために、4種類の基材を用いて作製した塗料の特性を調査した。次に、銅粉や銀メッキ粉末の混合による特性を評価した。さらに、塗布可能な紙や布などの塗布対象ごとの特性を調査した。

3.1 基材の検証

基材ごとの MagElePaint の特性を調査する。鉄粉には協和純薬工業株式会社の鉄粉#300(46 μm)を使用した。基材には、液体ゴム(株式会社ユタカメイク、クリア色), 布用ボンド(コニシ株式会社、裁ほう上手), シリコンコーク(コニシ株式会社)および速乾シリコンコーク(コニシ株式会社)を用いた。混合は、全ての材料を容器に入れた状態でマドラーを用いて約5分間攪拌することにより行った。鉄粉:基材の混合比は6:4(重量比)とした。これは、実験の結果、液体ゴムを基材として用いた場合にひび割

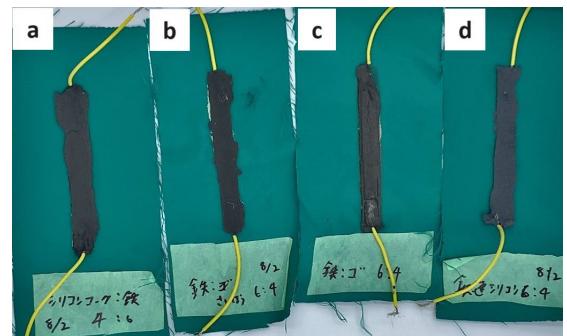


図 1. 基材を変えて作製した MagElePaint
(a: シリコンコーク b: 布用ボンド c: 液体ゴム d: 速乾型シリコンコーク)

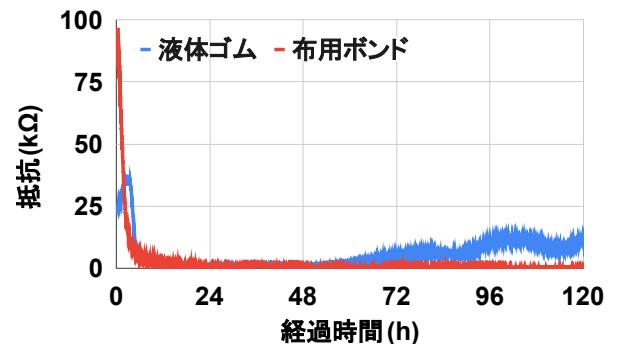


図 2. 液体ゴムと布用ボンドを使用した Mag-ElePaint の時間経過による抵抗値の変化の比較

れや断線を起こさず、かつ最も抵抗が低くなった比率である。また、どの基材においても鉄粉の比率をこれ以上にすると塗料がうまく硬化しない、柔軟性がなくなるといった問題が生じた。

塗料の経時による抵抗変化を観察するために、3Dプリンタで作製した厚さ 1 mm, 縦 8 cm, 横 1 cm の型を布生地(化学纖維)に固定した状態で塗料を流し込み、塗布した(図 1)。塗布時から 1 分ごとの抵抗値の変化を 5 日間記録した。2種類のシリコンコークはほぼ絶縁状態であったため、液体ゴムと布用ボンドの結果のみを図 2 に示す。各塗料は時間経過とともに固まり、同様に抵抗値が減少した。図 2 より、液体ゴムと布用ボンドは数時間で完全に固まり、抵抗値が 1 kΩ 前後と小さくなつたが、液体ゴムは 3 日目あたりから抵抗値が大きくなり始めた。これらの結果より、液体ゴムと布用ボンドの特性が良かったため、基材として用いる。

3.2 導電性材料の検証

3.1 節の検証後、液体ゴムを用いた塗料をさらに 1 週間(合計 12 日間)静置した際の抵抗値変化を図 3a に示す。図 3a より、時間経過とともに抵抗値が上昇していることがわかる。これは塗料内部の鉄粉

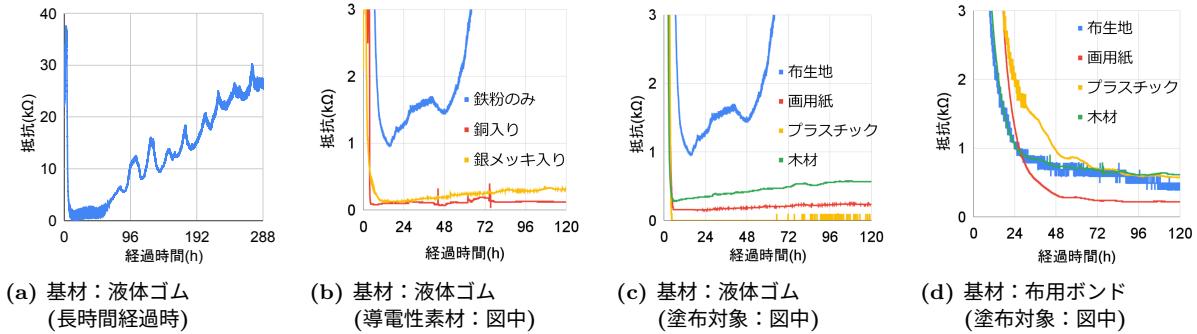


図 3. 基材や導電性材料、塗布対象を変更した MagElePaint の時間経過による抵抗値変化

の酸化や基材の水分の蒸発が原因と考えられる。

この問題を解決するために、鉄粉より安定で抵抗率の低い銅粉末や導電纖維（銀メッキ）を細かく刻んだもの（以下、銀メッキ粉末）の配合を試みた。MagElePaint は磁石による電子部品の接続を行うことを目標としているため、磁性を有する鉄粉をベースとして他の金属を混合する。鉄の抵抗率は 0-100 °C で 8.9×10^{-8} - $14.7 \times 10^{-8} \Omega \cdot \text{m}$ であるのに対し、銅の抵抗率は 0-100 °C で 1.55×10^{-8} - $2.33 \times 10^{-8} \Omega \cdot \text{m}$ と 1/6 程度である [1]。ここで、3.1 節と同様に鉄粉には協和純薬工業株式会社の鉄粉 #300 (46 μm) を使用し、銅粉には林純薬工業株式会社の銅粉 #325 (43 μm) を使用した。基材にはより柔軟性の高かった液体ゴムを使用した。銅粉入り塗料は鉄:銅:液体ゴムを 2:3:5 とし、銀メッキ粉末入り塗料は、鉄:液体ゴムが 6:4 の塗料に 15 cm ほどの導電纖維 (FUJIX, Smart-X) を 1 mm ごとに細かく切って混ぜて制作した。液体ゴムを用いた MagElePaint と同じように厚さ 1 mm、縦 8 cm、横 1 cm の型に入れて布生地に塗布し、1 分ごとの抵抗値の変化を 5 日間記録した結果を図 3b に示す。

図 3b に示すように、銅や銀メッキ粉末を混ぜた MagElePaint は液体ゴムを用いた MagElePaint よりも抵抗値が小さくなつた。特に銅を混ぜた MagElePaint は時間経過による抵抗値の上昇もほとんど見られず、常に 100 Ω 前後の抵抗値を維持していた。さらに銅、銀メッキ粉末どちらも混ぜた場合も、抵抗値が安定するまでの時間が短縮された。

3.3 塗布対象の素材の検証

布生地に MagElePaint を塗布した場合、布に MagElePaint が染み込むという事象が確認された。この事象が抵抗値の増加に関係しているのではないかと考え、塗布する素材の変更を行つた。重量比が鉄:基材 (液体ゴム、布用ボンド) = 6:4 の塗料を、布生地以外に画用紙、プラスチック板、木材に塗布し、抵抗値の変化を記録した。液体ゴムを使用した MagElePaint の抵抗値変化を図 3c、布用ボンドを使用した MagElePaint の抵抗値変化を図 3d に示す。

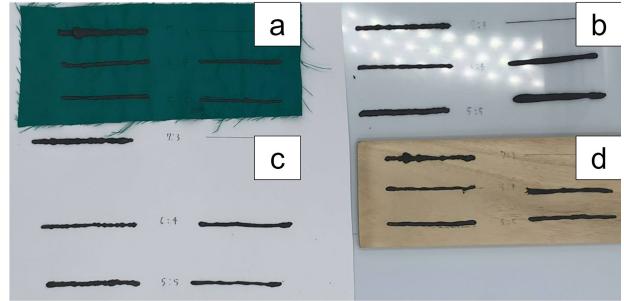


図 4. インジェクタを用いて MagElePaint を塗布した様子 (a: 布生地 b: プラスチック板 c: 紙 d: 木材)

液体ゴムを使用した MagElePaint では、布生地以外の素材の抵抗値の上昇はそれほど顕著に現れなかつた。しかし、図 3c より、布生地以外の素材も時間経過により抵抗値が微増していることがわかる。

布用ボンドを使用した MagElePaint では、全ての素材において同じような抵抗値の変化が得られた。図 3d より、全ての素材において抵抗値が安定してからの上昇はなく、抵抗値が微減している。しかし、どの素材に塗布した場合においても塗布してから抵抗値が安定するまでの時間は液体ゴムを用いた MagElePaint よりも長くなり、抵抗値も大きくなつた。

4 インジェクタを用いた細配線の検証

実際の回路プロトタイピングでは、3 章の検証のように分厚い板形状ではなく、細長い配線で描く。そこでインジェクタ (2 mm 径、2.5 ml) を用いて塗布した際の特性を調査した。

4.1 鉄粉のみを用いた塗料の検証

基材には液体ゴムと布用ボンドを使用した。重量比は鉄:基材が 7:3, 6:4, 5:5 について調査した。塗布する素材には画用紙、木材、布生地、プラスチック板を用い、各素材を 5 cm 塗布した。インジェクタを用いて塗布した様子を図 4 に示す。塗布した

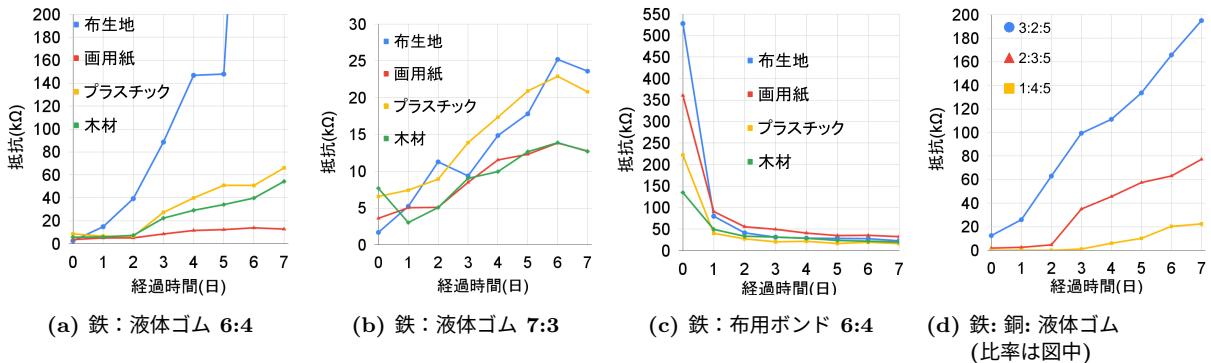


図 5. 様々な MagElePaint をインジェクタを用いて塗布した場合の時間経過による抵抗変化

MagElePaint の抵抗値を 1 日ごとに 1 週間計測した。液体ゴムを使用した重量比 5:5 の MagElePaint はどの素材に塗布した場合も 7 日を持たずして絶縁した。よって液体ゴムを使用した MagElePaint について図 5a(重量比 6:4), 図 5b(重量比 7:3) に示す。布用ボンドを使用した MagElePaint の重量比 7:3 は粘度が高すぎたためインジェクタに入れることができず、重量比 5:5 は全ての素材に塗布した場合も 1 日目から絶縁した。よって、布用ボンドを使用した MagElePaint(重量比 6:4) を図 5c に示す。

図 5a および 5b の結果より、液体ゴムを使用した MagElePaint は図 3c の結果と比較して抵抗値の上昇が顕著であった。さらに、5 cm の配線で数十～数百 kΩ と高い抵抗値であり、配線抵抗が無視できないほど高いことが分かった。

図 5c の結果より、布用ボンドを使用した MagElePaint はどの素材に塗布した場合も時間経過によって抵抗値が減少した。しかし初期の抵抗値が高いため、時間が経過しても抵抗値は数十 kΩ オーダーであった。そのため、液体ゴムと同様に配線抵抗が無視できないほど高いことが分かった。

それぞれの MagElePaint において、磁力の強さとひび割れ、断線の有無を調査した。磁力については、MagElePaint に磁石 (1.5g, 220mT) を張り付けた状態で塗布対象を反転させ、磁石が落ちなければ○、落ちれば×とした。ひび割れについては、塗布した MagElePaint が完全に硬化した状態で塗料にひび割れが生じていなかった場合は○、生じていた場合は×とした。画用紙と布に関しては、MagElePaint を塗布した状態で軽く曲げるなどを行い、ひび割れが生じないかも調査した。調査の結果を表 1 に示す。表 1 より、全ての MagElePaint において、磁力は十分であった。しかし、液体ゴムを使用した MagElePaint の塗布対象が画用紙や布であった場合、ひび割れが生じていた。

4.2 銅を使用した場合

内部抵抗が高すぎる、ひび割れが生じるという問題を解決するために、銅を混ぜた MagElePaint を作

表 1. MagElePaint をインジェクタを用いて塗布した際の磁力とひび割れの評価

基材の種類	重量比 鉄	重量比 銅	重量比 基材	塗布対象	磁力	ひび割れ
液体ゴム	6	0	4	画用紙	○	×
液体ゴム	6	0	4	プラスチック	○	○
液体ゴム	6	0	4	木材	○	○
液体ゴム	6	0	4	布	○	×
液体ゴム	7	0	3	画用紙	○	×
液体ゴム	7	0	3	プラスチック	○	○
液体ゴム	7	0	3	木材	○	○
液体ゴム	7	0	3	布	○	×
布用ボンド	6	0	4	画用紙	○	○
布用ボンド	6	0	4	プラスチック	○	○
布用ボンド	6	0	4	木材	○	○
布用ボンド	6	0	4	布	○	○
液体ゴム	1	4	5	画用紙	×	○
液体ゴム	2	3	5	画用紙	○	○
液体ゴム	3	2	5	画用紙	○	○

製し、同じように紙に塗布した。基材には、布用ボンドと比較して柔軟性が高く、硬化するまでの時間が短い液体ゴムを使用した。重量比が鉄: 液体ゴムが 6:4 の場合でひび割れが生じたため、今回は液体ゴムの比率を増やし、鉄: 銅: 液体ゴムが 1:4:5, 2:3:5, 3:2:5 について調べた。結果を図 5d に示す。図 5d より、どの重量比においても硬化後 1～3 日間は小さい抵抗値を維持していたが、その後抵抗値が上昇し数十～数百 kΩ となった。これらの結果より、銅を使用した MagElePaint は塗布後 2 日間は内部抵抗を無視できる塗料として使用できると考える。さらに、4.1 節と同様に磁力の強さとひび割れ、断線の有無を確認した(表 1)。表 1 より、鉄: 銅: 液体ゴムが 1:4:5 の場合、磁力が不十分であった。しかし、銅を使用した全ての MagElePaint においてひび割れは生じなかった。よって、鉄: 銅: 液体ゴムが 2:3:5 の MagElePaint は、磁力が十分であり、ひび割れや断線もなく、抵抗値が最も低いため、回路プロトタイピングに適していると結論付けた。

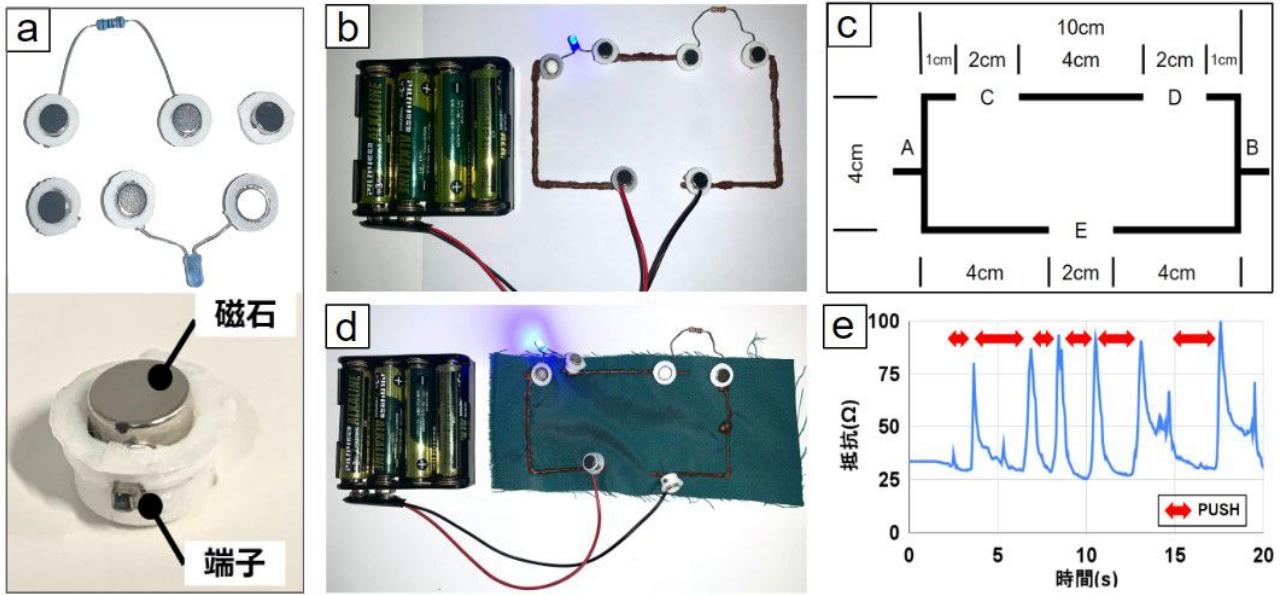


図 6. アプリケーション例(紙, 布). (a)3D プリンタで作製した電子部品接続用の端子, (b) 紙上への回路プロトタイピング, (c) 抵抗計測のために用いた回路の模式図, (d) 布上への回路プロトタイピング, (e) 布に塗布した MagElePaint を押下することによる抵抗値の変化

5 アプリケーション例

MagElePaint を用いた紙上/布上/皮膚上の回路プロトタイピング例を示す。すべてのアプリケーション例において、4章の検証で最も特性の良かった鉄:銅:液体ゴムの重量比が 2:3:5 の MagElePaint を使用した。

5.1 紙上への回路プロトタイピング

市販の電子部品を容易に接続できるように磁石端子を 3D プリンタで作製した。図 6a に作製した端子を、図 6b に作製した端子を用いて回路を形成した様子を示す。これらの端子と MagElePaint を用いて回路を構成し、理論値と実測値を比較した。

比較のため、2つの抵抗の合成抵抗(直列、並列)を求め、理論値との誤差を調べた。抵抗計測のため用いた回路の模式図を図 6c に、理論値と実測値の比較を表 2 に示す。図 6c の AB にデジタルマルチメータ(sanwa, CD731a)を接続し抵抗値を接続した。直列時は図 6c の C に表 2 の R_1 , D に R_2 を接続し、E は開放した。並列時は C に R_1 , E に R_2 を接続し、D は短絡した。使用した抵抗の許容誤差は全て $\pm 1\%$ であった。表 2 より、 100Ω , $1k\Omega$ の抵抗を使用した場合は誤差が抵抗の許容誤差である $\pm 1\%$ を大幅に超えた。しかし、 $10k\Omega$, $100k\Omega$ の抵抗を使用した場合は誤差が抵抗の許容誤差である $\pm 1\%$ 前後であった。よって、回路プロトタイピングの際、 $10k\Omega$ 以上の回路であれば内部抵抗を無視することができるといえる。

表 2. MagElePaint を使用した回路の理論値と実測値の比較

$R_1[\Omega]$	$R_2[\Omega]$	理論値 [Ω]		実測値 [Ω]		誤差 [%]	
		直列	並列	直列	並列	直列	並列
100	100	200	50	265	101	24.6	50.6
1k	1k	2k	500	2.15k	541	6.98	7.58
10k	10k	20k	5k	20.2k	5.08k	0.99	1.58
100k	100k	200k	50k	201k	50.1k	0.70	0.20

5.2 布上への回路プロトタイピング

布上に MagElePaint を塗布した回路プロトタイピングを図 6d に示す。図 6d の通り、MagElePaint の配線上に電子部品を磁石を介して接続することで、LED が点灯した。さらに、MagElePaint を布に塗布し、配線を押下したときの抵抗値を図 6e に示す。図 6e より、配線の押下によって抵抗値が変化するため、布に塗布した MagElePaint は押下センサとしても利用できることが示唆された。

5.3 皮膚上への回路プロトタイピング

皮膚上に MagElePaint を塗布した回路プロトタイプを図 7a に示す。なお安全のために、皮膚上から直接塗布するのではなく、フィルム(skinix 社、エアウォール UV)の上から塗布した。図 7a の通り、LED が点灯したことから、皮膚上でも回路を作製できることが確認された。さらに、図 7b のように皮膚上に配線した際に、腕を弛緩、緊張させた際の抵抗値を図 7c に示す。図 7c より、MagElePaint は筋肉の隆起をセンシングできる可能性が示唆された。

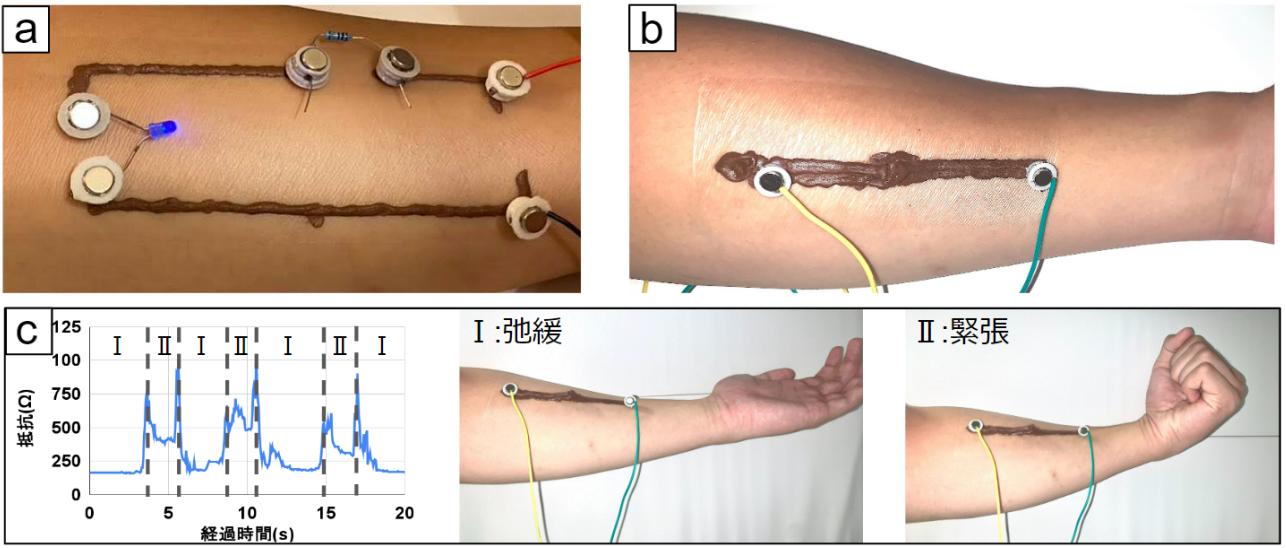


図 7. アプリケーション例(皮膚). (a) 皮膚上への回路プロトタイピング, (b) 皮膚上に MagElePaint を塗布することによるセンサ, (c) 腕の弛緩, 緊張による抵抗値の変化

6 課題と議論

6.1 MagElePaint が硬化するまでの時間

MagElePaint は硬化すると抵抗値が安定し、回路として使えるようになる。MagElePaint が硬化するには一定の時間を要する。様々な素材を用いた MagElePaint をペンのように塗布した場合の硬化するまでの時間を図 8 に示す。図 8 より、硬化するまでの時間はそれぞれの素材で 40~100 分程度であることがわかる。この硬化時間は、電子回路教育や皮膚上に回路を実装する際に問題となりえる。この問題を解決するためには、より硬化時間の短い素材、例えば UV ライトですぐに固まるレジンの使用が考えられる。今後、柔軟性を有するフレキシブルレジンの使用を検証する。

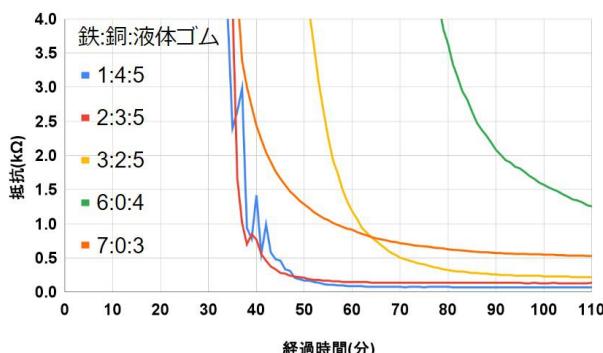


図 8. MagElePaint をインジェクタを用いて塗布した場合の比率ごとの硬化するまでの時間

6.2 経時による抵抗の上昇

MagElePaint は経時によって抵抗が大きくなる。この問題は、図 1 のようにあらかじめ端子を内部に入れた場合は抑えられることから、MagElePaint の表面の酸化が原因であると考えられる。

また、図 5d に示すように、鉄と銅を混合した MagElePaint は 2~3 日後に急激に抵抗値が上昇した。これは液体ゴム中の水分と、鉄と銅の二種類の金属により電池が形成され、鉄がイオン化したためであると考えられる [16]。

これらの問題を解決するために、磁性を有するステンレス (SUS430) のように酸化しにくい磁性導電性素材を今後試す。

7 まとめ

本論文では、導電性/磁性/柔軟性を持った塗料である MagElePaint を提案した。鉄粉、基材、導電性素材の種類や重量比を検証し、様々な素材に塗布可能かつ抵抗値が低くなるよう調査した。基材として、液体ゴム、布用ボンド、シリコンコークや速乾型シリコンコークを、導電性素材として、銅粉や銀メッキ粉末を調査した。さらに塗料をインジェクタを用いて塗布した場合の最適な鉄粉、基材、導電性素材の種類や重量比を調査した。その結果、鉄:銅:液体ゴムの重量比が 2:3:5 の場合、磁力が十分でひび割れ、断線が生じず、最も抵抗値が低かった。アプリケーション例として MagElePaint を用いた紙、布、皮膚上での回路プロトタイピングを行った。

今後は MagElePaint が硬化するまでの時間を短縮する、経時による抵抗の増加を軽減するといった課題に取り組む。

謝辞

本研究は JSPS 科研費 JP22K17941 の助成を受けたものです。また、本研究の一部は東北大大学電気通信研究所 共同プロジェクト研究によるものです。

参考文献

- [1] 金属の電気抵抗. https://www.chip1stop.com/sp/knowledge/007_electric-resistance-of-metal. Accessed 2024/08/30.
- [2] Bare Conductive. Electric Paint. <https://www.bareconductive.com/collections/electric-paint>. Accessed 2024/10/25.
- [3] Y. Choi, N. Ryu, M. J. Kim, A. Dementyev, and A. Bianchi. BodyPrinter: Fabricating Circuits Directly on the Skin at Arbitrary Locations Using a Wearable Compact Plotter. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, pp. 554–564, New York, NY, USA, 2020. Association for Computing Machinery.
- [4] N. A.-h. Hamdan, S. Voelker, and J. Borchers. Sketch&Stitch: Interactive Embroidery for E-textiles. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 1–13, New York, NY, USA, 2018. Association for Computing Machinery.
- [5] Y. Kawahara, S. Hodges, B. S. Cook, C. Zhang, and G. D. Abowd. Instant inkjet circuits: lab-based inkjet printing to support rapid prototyping of UbiComp devices. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '13, pp. 363–372, New York, NY, USA, 2013. Association for Computing Machinery.
- [6] A. Khan, J. S. Roo, T. Kraus, and J. Steimle. Soft Inkjet Circuits: Rapid Multi-Material Fabrication of Soft Circuits using a Commodity Inkjet Printer. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, pp. 341–354, New York, NY, USA, 2019. Association for Computing Machinery.
- [7] E. Lovell and L. Buechley. An e-sewing tutorial for DIY learning. In *Proceedings of the 9th International Conference on Interaction Design and Children*, IDC '10, pp. 230–233, New York, NY, USA, 2010. Association for Computing Machinery.
- [8] E. Lovell, L. Buechley, and J. Davis. LilyTiny in the Wild: Studying the Adoption of a Low-Cost Sewable Microcontroller for Computing Education. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, pp. 282–293, New York, NY, USA, 2023. Association for Computing Machinery.
- [9] MagPaint JAPAN. MagnetPaint. <https://www.nsmrp.com/#topProductListWrapper>. Accessed 2024/08/30.
- [10] N. Matsuhisa, M. Kaltenbrunner, T. Yokota, H. Jinno, K. Kuribara, T. Sekitani, and T. Someya. Printable elastic conductors with a high conductivity for electronic textile applications. *Nature Communications*, 6(1):7461, 2015.
- [11] D. A. Mellis, S. Jacoby, L. Buechley, H. Perner-Wilson, and J. Qi. Microcontrollers as material: crafting circuits with paper, conductive ink, electronic components, and an "unt toolkit". In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '13, pp. 83–90, New York, NY, USA, 2013. Association for Computing Machinery.
- [12] N. Pourjafarian, M. Koelle, B. Fruchard, S. Mavali, K. Klamka, D. Groeger, P. Strohmeier, and J. Steimle. BodyStylus: Freehand On-Body Design and Fabrication of Epidermal Interfaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 1–15, New York, NY, USA, 2021. Association for Computing Machinery.
- [13] S. Pradhan and V. K. Yadavalli. A dual function conductive nano ink for printed electronics connections. *Organic Electronics*, 124:106959, 2024.
- [14] K. Singh, A. Ohlan, P. Saini, and S. K. Dhawan. Poly (3,4-ethylenedioxythiophene) γ -Fe₂O₃ polymer composite–super paramagnetic behavior and variable range hopping 1D conduction mechanism–synthesis and characterization. *Polymers for Advanced Technologies*, 19(3):229–236, 2008.
- [15] K. W. Song, C. Dierk, S. T. Tung, and E. Paullos. Lotio: Lotion-Mediated Interaction with an Electronic Skin-Worn Display. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery.
- [16] 秋山 貴之, 西海枝 恵. 「化学変化と電池のしくみ」. 仙台市科学館研究報告, 31:23–27, 2022.
- [17] 山岸 真人, 高田 峻介. 鉄粉ボンドを手袋に塗布することによる拡張可能なデータグローブの作製. インタラクション 2024 論文集, pp. 831–835. 情報処理学会, 2024.
- [18] 酒井 大輔, 木田 彩佳, 原田 建治, 柴田 浩行. 導電ペンと筆ペンで描いて学ぶ電気の基礎. 電気学会論文誌A (基礎・材料・共通部門誌), 138(1):30–35, 2018.
- [19] 一条 洋和, 鈴木 大介, 遠藤 健太郎, 宝賀 剛, 佐藤 淳. 導電ペンを使用した電気回路演習. 工学教育研究講演会講演論文集, 2016:126–127, 2016.

未来ビジョン

本研究では、様々な素材に回路プロトタイピングを行うことができる磁性導電性塗料、「MagElePaint」を提案した。MagElePaintは、紙や布、木材、プラスチックなど多くの素材に接着することが可能である。さらに、MagElePaintを用いたセンシングの可能性も示唆された。そこで未来ビジョンとしては、様々な素材に容易に塗布できるセンサを提案する。その結果、一般的な壁や手袋、靴をMagElePaintによってセンサやデータグローブ、スマートシューズに容易にグレードアップできることが期待される。

さらに、現在普及しているウェアラブルデバイスには拡張性が低く、センサを新しく取り付けたり新たなハードデバイスを接続することは困難である。しかし、MagElePaintを用いたウェアラブルデバイスでは、磁石によって容易にデバイスを拡張できることが期待される。例えばMagElePaintを用いたデータグローブであれば、ボタンやジョイスティックなどを磁石で容易に接続できるだろう。

最終的には、MagElePaintによって多くの物がデバイス化、拡張され、ウェアラブルデバイスがより人々によって身近な存在となる手助けとなることを目指したい。

RippleSpike: スパイク表現と波紋表現を組み合わせた情報提示手法

野間 直生 * 沖 真帆 * 塚田 浩二 *

概要. 液体ディスプレイには、柔軟な形状変化や液体の質感を活用できるといった特徴がある。本研究では、液体の中でも磁力によって容易に制御が可能な磁性流体に着目した。磁性流体とは、強磁性の粉末を分散させたコロイド溶液であり、強力な磁力を印加することでスパイク状に隆起する。このスパイク表現はメディアアート等で広く活用されており、強い存在感を持つ。そこで本研究では、磁性流体にオイルを混合させることで、棘のようなテクスチャを抑制し、なだらかな半球や余韻を表現して、スパイク表現と組み合わせて利用する。本稿では、スパイク表現となだらかな波紋表現を組み合わせた情報提示手法「RippleSpike」の提案、実装、性能評価、表現事例と応用例について述べる。

1 はじめに

近年、液体を用いた情報表現手法が盛んに提案されている。例えば、バブルディスプレイ [13] は水中の気泡を利用し、水の時計 [11] は水滴の間隔を調整することで、情報提示を行う。こうした液体ディスプレイは、柔軟な形状変化や液体の質感を活用できるといった特徴がある。特に印象的な表現を行える素材として、磁性流体が知られている。磁性流体とは、強磁性粒子、溶媒、界面活性剤から成り立つコロイド溶液である。磁場の影響を受けて形状が変化する特性を持ち、磁石を近づけるとスパイク現象が発生して、磁場の方向に沿って棘のような隆起の集合体が形成される（図1左）。隆起の密度や数、大きさは磁力の強さや位置に応じて変化するため、電磁石等を用いた制御により動的に形状を変更するメディアアート作品等に活用されている [12]。このようにスパイク表現は印象的な効果を生む一方で、視覚的な存在感が強く、アンビエント・ディスプレイのようなさりげない情報提示には活用が難しかった。磁性流体でさりげない情報提示を行う手法を確立できれば、スパイク表現と組み合わせて、存在感や誘目性をなめらかに調整可能な情報提示手法として活用できる可能性がある。そこで本研究では、磁性流体にオイルを混合し、磁力の制御方法を工夫することで、なだらかな半球や余韻を表現し（図1右）、スパイク表現と組み合わせ可能な情報提示手法「RippleSpike」を提案する。本稿では、RippleSpike のコンセプト、実装、性能評価、表現事例と応用例について述べ、適切なオイル混合率や電磁石の制御方法について報告する。

なお、本稿は、インタラクション 2024 のデモ発表 [14] と HCI 研究会で発表 [15] した内容を元に整理した。

Copyright is held by the author(s).

* 公立はこだて未来大学

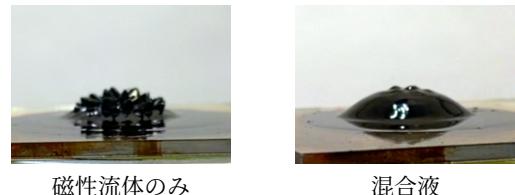


図 1. 磁性流体に磁力を印加した様子

2 関連研究

本研究の関連研究を「アンビエント・ディスプレイ」、「液体を用いた表現手法」、「磁性体を用いた表現手法」の3つの視点から紹介する。

2.1 アンビエント・ディスプレイ

Tangible Bits[3] では、実世界でのインターフェースの1つとして、周辺視野などの人間の知覚の周辺に設置する Ambient media を提案している。Ambient media は、人が常に周辺環境から様々な情報を得ていることに着目し、音や光、空気の流れなどを利用することで、ユーザが注意を向いていなくても情報を気軽に受け取ることができる。Wisneski ら [8] は、アンビエント・ディスプレイとして AmbientRoom と Pinwheels を提案している。AmbientRoom は壁への投影や環境音、光の揺れなどを Ambient media として用いている。Pinwheels は、株価等の多様なデジタル情報を風車の回転速度を用いて表現する手法である。本研究では、磁性流体を用いて、存在感のあるスパイク表現だけでなく、さりげない波紋表現も組み合わせたアンビエント・ディスプレイの応用例を作成する。

2.2 液体を用いた表現手法

Coworo[4] は、攪拌によって液体表面が隆起する液体に着目したメディアアートである。電子制御によって、任意の位置、高さで隆起を発生させること

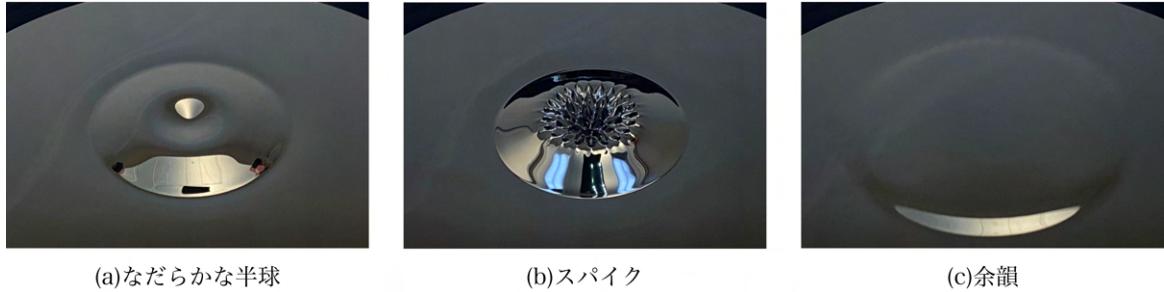


図 2. 本提案で作成可能な表現例

ができる。BubBowl[1] は、電気分解により発生する泡を画素として用いて、水面をドットマトリクスディスプレイのように活用できる。具体的には、インスタントコーヒー、コーンスター、炭酸水素ナトリウムを混ぜた黒い電解液を用いて、動的に変化する表現事例を提示している。また、Bubble Mirror[2] は、BubBowl の応用研究であり、水面を覗き込むことで鑑賞者の顔が泡で表示される作品等を提案している。Tangible Drops[6] は、水酸化ナトリウム中に浮かべた液状のガリウムを画素として用いるディスプレイである。任意の位置に液状のガリウムを移動させることで、視覚・触覚フィードバックを実現している。本研究では、なだらかな半球とスパイク表現、および高速に振幅する波紋表現等を組み合わせた多様な液体表現の構築を目指す。

2.3 磁性体を用いた表現手法

Astral Body[10] はテーブル上に広げた砂鉄を磁石を用いて制御することで、生き物の動きを再現したメディアアートである。Millefiori[5] は磁性流体の「ラビリンス不安定性」という、狭い空間に密閉した際に発生する複雑なパターンを利用したメディアアートである。水彩絵の具と磁性流体の混合液を用いることで色彩豊かな表現を行なっている。本研究では、磁性流体にオイルを混合したり、制御方法を工夫することで、スパイク現象に加えて、なだらかな半球や余韻を表現する手法を提案する。

中野ら [9][16] はスライム状の磁性体を作成し、永久磁石で操作する手法を提案している。磁石を操作することで、磁性体の移動／伸縮、合成／分離などの動作が可能である。Slime Robot[7] は粘弾性のある液体状のロボットである。変形や移動に加えて、液体が導電性であることを活用し、電子回路を修復する事例を提案している。本研究では、なだらかな表現やスパイク表現を組み合わせた情報提示手法を開発する。

3 提案

3.1 コンセプト

本研究では、磁性流体にオイルを混合することで、磁性流体の表現を拡張し、スパイク現象に加えて、なだらかな半球や余韻を表現する。また、作成した表現を用いた情報提示手法「RippleSpike」を提案する(図 2)。RippleSpike のコンセプトは、大きく以下の 2 点である。

1. スパイク表現となだらかな半球表現の融合
2. 生活空間での利用

一点目は、スパイク現象に加えて、なだらかな半球や余韻を表現することである。こうした表現を行うためには、磁性流体の粘度を下げることで、スパイク状の隆起を抑えたり、応答性を高める。二点目は、システムを生活空間で長期的に利用できるよう、構成をシンプルにしたり、磁性流体の漏れや揮発を防ぐ工夫を行うことである。

3.2 デバイス構成

生活空間での利用を見据えて、図 3 のように設置や運用が容易なシンプルなデバイス構成を取る。

格納部

磁性流体は直接触ることができず、また布などに付着すると微小な粒子が沈着してしまうため、密閉する必要がある。そこで、格納部となるトレイと液漏れ対策を施した蓋を製作する。

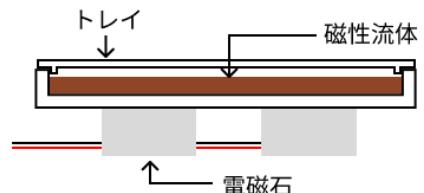


図 3. 基本のデバイス構成

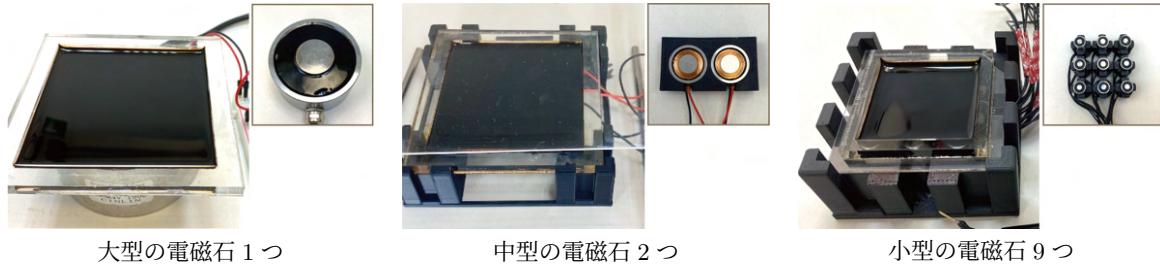


図 4. プロトタイプ

電磁石アレイ

格納部の下部には電磁石アレイを設置する。電磁石アレイは筐体に固定されており、安定して格納部を支えつつ、交換もできるように工夫する。様々なサイズ/数の電磁石アレイを用意することで、多様な表現を検討する。

センシング

基本的な表現にセンサは利用しないが、後述する応用例等を想定して、センサを取り付け可能な設計とする。具体的には、会話を検出するマイク、人の存在を検出する距離センサ等を検討する。さらに、センサの脱着が容易で拡張性のある構造を目指す。

4 実装

本章では、RippleSpike のプロトタイプについて説明する。プロトタイプは大きく、「磁性流体の格納部」「電磁石アレイ」「制御基板」を中心に構成される。以下、オイルの混合と、プロトタイプ各部の詳細を示す。図4は作成した3種類のプロトタイプである。それぞれ、電磁石アレイの筐体に格納トレイ(蓋無し)を設置している。プロトタイプを用いて、混合液の動作確認と表現の模索を行った。

4.1 オイルの混合

なだらかな半球や余韻を表現するために、本研究では磁性流体にミネラルオイルを混合する。ミネラルオイルの、「粘度が低い」「揮発しにくい」「劣化しにくい」といった特性に着目する。磁性流体にはシグマハイケミカルのDS-50を、ミネラルオイルにはジョンソンのベビーオイルを用いた。図1はそれぞれ、磁性流体のみと混合液に下部から永久磁石を用いて磁力を印加している様子である。オイルを混合することで、図1右のように、スパイクを抑制する効果や、応答性能が向上する傾向があった。さらに、磁性流体の揮発を防止する効果も見られた。

4.2 磁性流体の格納部

磁性流体はシャーレのような薄型のトレイに格納する。プロトタイプでは、100mm × 100mmのケースと50mm × 50mmのトレイを作成した。磁性流体の動きを観察しやすいように、トレイは透明のアクリル板をレーザーカッターで加工し、液体が漏れないよう慎重に接着した。また、磁性流体は人体に有害であるため、直接触れたり、液漏れしたりする状況を避ける必要がある。そこで、磁性流体の格納トレイにはパッキンを付けた透明の蓋を用意した。パッキンはエラスティックレジンを用いて光造形式の3Dプリンタで自作し、蓋の内側のガイドに沿って接着した。

4.3 電磁石アレイ

多様な表現を試行するために、3種類の電磁石アレイを用意した。筐体のサイズを揃えるために、大型の電磁石(直径80mm、磁力55mT)は1つ、中型の電磁石(直径26mm、磁力29mT)は2つ、小型の電磁石(直径8mm、磁力23mT)は9つを同時に制御可能な設計とした。磁力はそれぞれ24Vを印加したとき、ガウスマーターで計測した実測値である。それぞれの電磁石に合わせて筐体を設計し、格納トレイを固定できるように配慮した。図4の各プロトタイプの右上に、使用した電磁石を示す。

4.4 制御基板

電磁石を制御するための制御回路と基板について述べる。電磁石毎に1つのFET(K4017)を利用しておらず、マイコン(Arduino pro mini)から制御する。基板はEagleで独自に設計した(図5)。1枚の基板に6つのFETを備えており、最大6個の電磁石を制御できる。JSTコネクタやオーディオプラグを用いることで、手軽に交換できるように設計した。さらに、基板同士をI²Cを介して接続できるように設計しており、n枚の基板を連結することで、n × 6個の電磁石を同時に制御できる。同様の方法で、センサを制御する基板も拡張できる仕様とした。

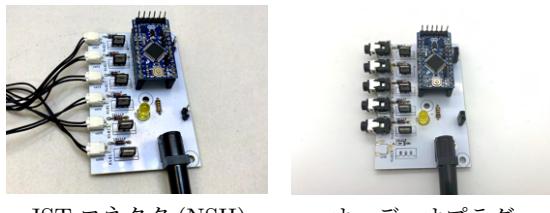


図 5. 基板の外観

5 性能評価

本章では、主にオイルの混合率に応じた外観や応答性の変化について述べる。

5.1 オイルの混合による外観の変化

5.1.1 目的と手法

本研究では、スパイク表現となだらかな半球表現の両立を目指す。そこで、2種類の表現が両立可能なオイルの混合割合を調査した。大型の電磁石に0mT～90mTの間で10mT毎に磁力を印加し、混合液の様子を観察した。混合液は表1の6種類の比率で用意した。磁力はガウスマーターを用いて最も磁力が強い位置を計測した。

表 1. 混合液の比率

混合液	磁性流体 (%)	オイル (%)
A	100	0
B	90	10
C	67	33
D	50	50
E	33	67
F	25	75

5.1.2 結果と考察

実験結果のうち、特に差が見られた部分を図6に示す。混合液A(100:0)では50mT～90mTにおいて、混合液B(90:10)では60mT～90mTにおいてスパイク現象が見られた。混合液C(67:33)では90mTでもスパイク現象が確認できなかった。これらのことから、オイルを混ぜることでスパイク現象がかなり抑制されることがわかった。また、混合液のオイルの割合を10%にすることで、スパイクが発生せずになだらかな半球表現が行える磁力(\approx PWM値)の範囲を広く取ることができ、多様な表現を制御しやすいと考えた。

5.2 オイルの混合による反応速度の比較

5.2.1 目的と手法

応答性を定量的に確認するため、オイルの混合による反応速度の変化を調査した。8mmの電磁石を

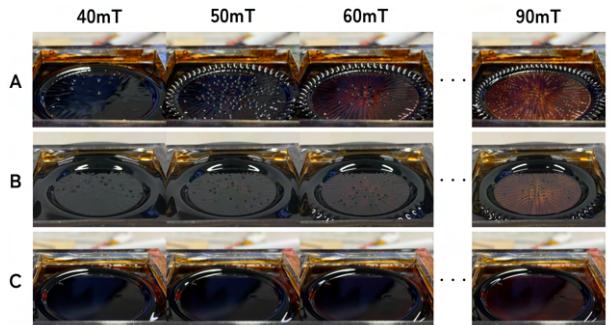


図 6. スパイク現象の変化

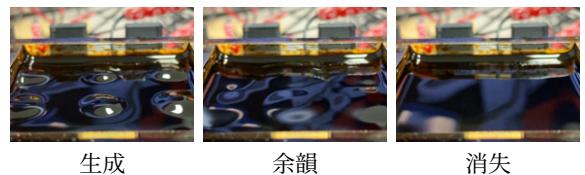


図 7. 反応速度調査の様子

9つ用いて、2秒間100%／8秒間0%の2値制御を行った。2値制御を行うことで、図7のような生成／余韻／消失と遷移する様子が観察される。

磁力を100%印加するとなだらかな半球の生成が始まる。生成の開始から半球が大きくなりきたところまでを「生成にかかった時間」とする。次に、磁力の印加を100%の状態から0%に急激に変化させることで、余韻の表現が見られる。余韻の開始から、水面が平らになるまでの時間を「消失にかかった時間」とする。

混合液の様子を斜め上から30fpsで撮影し、生成／消失にかかった時間を動画編集ソフト(final cut pro)を用いて目視で確認した。電磁石の出力と同時にLEDを点灯させることで、磁力の印加／停止のタイミングを判別した。

5.2.2 結果と考察

表2は半球の生成/消失にかかった時間を示す。混合液D(50:50)が最も生成に必要な時間が短く、割合が偏ることで生成が終わるまでの時間がやや長くなる傾向が見られた。消失では、混合液A(100:0)が最も消失に必要な時間が長く、オイルの混合割合を増加させることで時間が短縮する傾向が見られた。

消失にかかる時間は、AとBの差は少ないが、C以降は大幅に短くなり、「余韻」の表現が難しくなる。5.1.2の結果と併せて考えると、Bの混合液（オイル10%）のバランスが良いと考えたため、応用例等で利用した。

混合液の番号	生成 (s)	消失 (s)
A	1.55	6.29
B	1.56	5.50
C	1.16	3.49
D	1.04	1.17
E	1.52	1.00
F	1.49	0.56

6 表現事例と応用例

6.1 表現事例

提案手法は大別して、なだらかな半球、スパイク、余韻の3種類の表現を行うことができる(図2)。具体的な制御方法と表現事例を併せて紹介する。

6.1.1 PWM 値と表現例

PWMの出力値(割合)を時間経過に合わせて制御することで、(1)2値制御、(2)連続制御、(3)振幅制御の3種類の表現事例を作成した(図8)。筐体には、中型の電磁石と磁性流体75%／ミネラルオイル25%の混合液を用いている。この混合率の混合液では、PWM値100%であっても、スパイク表現は見られない。

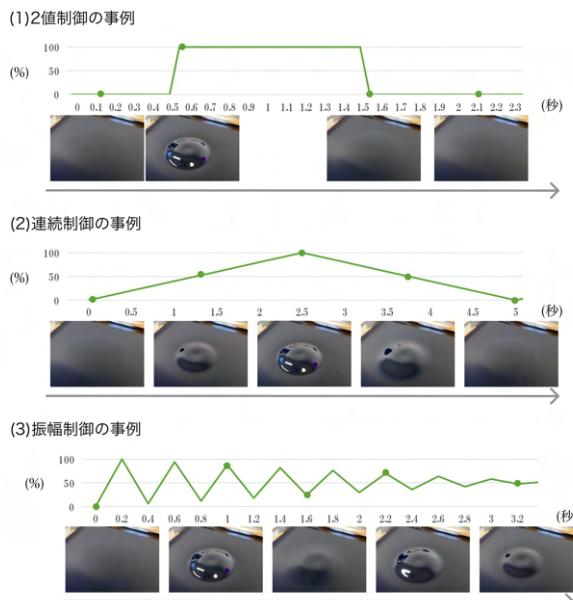


図 8. 基本的な表現事例

(1)は、PWM値で0%／100%の2値制御の事例である(図8(1))。100%にした直後(0.5秒)で半球が生じており、0%にした直後(1.5秒)にはゆっくり波が引く余韻の表現が見られた。

(2)は、PWM値を0～100%まで約2.5秒かけて

連続的に上昇させ、その後同様に100～0%まで下降させた事例である(図8(2))。

(3)は、振幅を徐々に小さくしながら、連続的に繰り返した事例である(図8(3))。周期は約0.4秒で固定した。PWM値に連動して、半球が高速に振動しながら徐々に落ちしていく表現が観察できた。

また、上記の表現とスパイク表現を組み合わせて、(4)水跳ね、(5)ツノの2種類の表現を作成した(図9)。筐体には、大型の電磁石と磁性流体90%／ミネラルオイル10%の混合液を用いた。

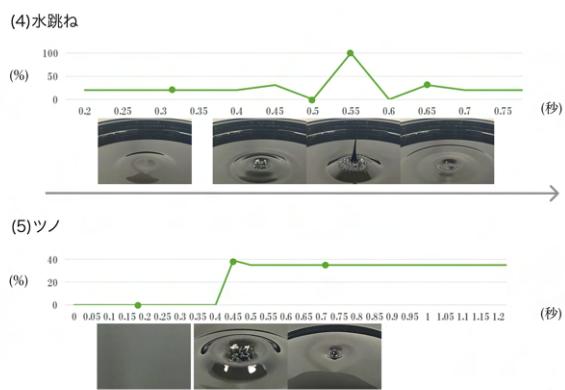


図 9. 複合的な表現事例

(4)は、水が跳ねている様子を参考に表現した事例である(図9(4))。約0.1秒の間に25%の出力から35%まで上昇させ、0%まで出力を下げる。次に、同様に約0.1秒で0%から100%まで出力を上げ、0%まで出力を低下させ、最初と同様に35%の出力を行った。一瞬ではあるが、混合液が飛び跳ねる様子が観察でき、スパイク／余韻共に刹那的な表現が見られた。予備動作を付随させることで、より激しく混合液が動いたと考えられる。

(5)は、スパイクの数を磁力によって調整し、一本のみスパイクを作成することで、ツノのように見える事例である(図9(5))。0%の状態から、約0.1秒で40%の出を行い、38%まで出力を低下させた。

6.1.2 状態遷移

6.1.1節の知見に基づいて、RippleSpikeを用いた表現の状態遷移を図10に整理した。ここでは、磁性流体90%とミネラルオイル10%の混合液を対象とする。磁力を印加していない初期の状態(PWM値: 0%)から、一定の磁力を印加(PWM値: 30%)することでなだらかな半球が生成される。より強い磁力(PWM値: 100%)を印加することで、初期の状態からスパイクに遷移する。なだらかな半球／スパイクから、瞬時に磁力の印加を止める(PWM値: 0%)ことで、余韻が生成されてから初期状態に戻る。

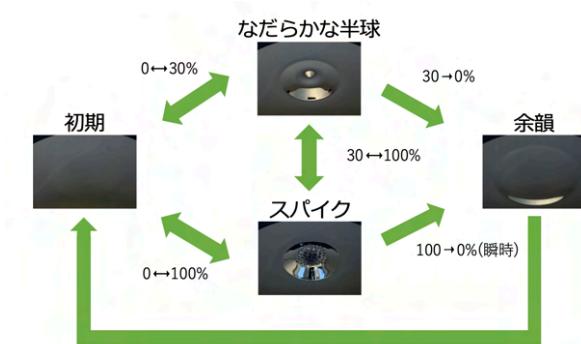


図 10. 磁力の変化による遷移図

土壤の水分量のセンシングを行い、水分量に応じて卓上デバイス上で表現を行う。例えば、徐々に土壤の水分量が不足してくると、ツノが生成／消失を繰り返し、水が足りていないことをユーザに伝える(図13左)。水を与えるとスパイクの表現を用いて水が十分に足りている様子を伝える(図13右)。

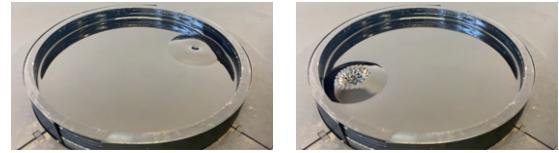


図 12. 会話の可視化の表現例

6.2 応用例

会議室やホール等に設置してアンビエント・ディスプレイのように活用することを想定して、直径80mmの大型電磁石を5つ用いた大型筐体を実装した(図11)。磁性流体の格納ケースの直径は260mmであり、数メートル程度離れた距離からも容易に挙動を観察することができる。また、マイクや距離センサを取り付け／交換可能な仕組みも備えている。

また、ユーザが生活する場所に設置しやすいように、大型の電磁石一つを円柱状の筐体に組み込んだ卓上デバイスを作成した(図13)。

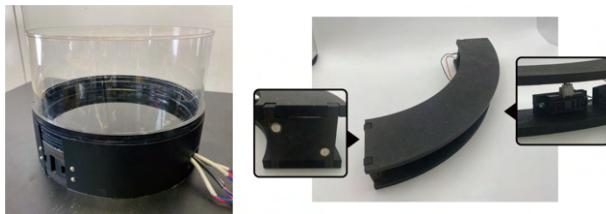


図 11. 大型筐体の外観(左)とセンサ類の交換機構(右)

6.2.1 会話量の可視化

複数人の会話に応じて表現が変化する事例を紹介する。会話の可視化では、大型筐体を用いる。指向性マイクを複数用いて、会話の音量と方向をセンシングし、なだらかな半球やスパイクを用いて表現する。静かな声で会話が進行している時は、静かに半球が生成と消失を繰り返す(図12左)。会話が白熱したり、感情的になることで、スパイクが出現し、白熱している様子を表現する(図12右)。筐体を中心で設置し、会議の様子を可視化することで、さりげなく発話量の調整や、会話の温度感の調整ができると考えている。

6.2.2 植物の土壤水分量の可視化

観葉植物の土壤の水分量を可視化する事例を紹介する。鉢に取り付けた自作水分量センサを用いて、



図 13. 植物の状態可視化の表現例

7 まとめと展望

本研究では、磁性流体にオイルを混合させ、制御方法を工夫することで、スパイク現象に加えて、なだらかな半球や余韻を表現する手法を提案した。さらに、これらの表現手法を組み合わせた情報提示手法「RippleSpike」を構築した。また、性能評価を通して、システムの応答性や表現力を調査した。

今後の展望として、電磁石の角度や配置を調整し、電磁石から発生する磁力同士を干渉させることで、より複雑で多彩な表現を作成することができる可能性がある。例えば、本提案では磁場を垂直に印加しているが、水平に印加することで、楕円状の半球が生成される。また、現状のアクリルを利用した天板は、光の反射によって視認性が損なわれてしまう。具体的な解決策として、低反射ガラスを使用することで、光の反射を防ぐことができると考える。

謝辞

本研究の一部は、科研費20H04231の支援を受けた。

参考文献

- [1] A. Ishii and I. Sii. BubBowl: Display Vessel Using Electrolysis Bubbles in Drinkable Beverages. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, p. 619–623, New

- York, NY, USA, 2019. Association for Computing Machinery.
- [2] A. Ishii, N. Tanaka, and I. Siio. Bubble Mirror: An Interactive Face Image Display Using Electrolysis Bubbles. In *SIGGRAPH Asia 2020 Emerging Technologies*, SA '20, New York, NY, USA, 2020. Association for Computing Machinery.
- [3] H. Ishii and B. Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, p. 234–241, New York, NY, USA, 1997. Association for Computing Machinery.
- [4] T. Matsunobu and Y. Kakehi. Coworo: a kinetic installation with shape-changing liquid. <https://xlab.iii.u-tokyo.ac.jp/projects/coworo/>, 2018.
- [5] F. Oefner. Millefiori. <https://www.behance.net/gallery/4036035/Millefiori>.
- [6] D. R. Sahoo, T. Neate, Y. Tokuda, J. Pearson, S. Robinson, S. Subramanian, and M. Jones. Tangible Drops: A Visio-Tactile Display Using Actuated Liquid-Metal Droplets. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–14, New York, NY, USA, 2018. Association for Computing Machinery.
- [7] M. Sun, C. Tian, L. Mao, X. Meng, X. Shen, B. Hao, X. Wang, H. Xie, and L. Zhang. Reconfigurable Magnetic Slime Robot: Deformation, Adaptability, and Multifunction. *Advanced Functional Materials*, 32, 2022.
- [8] C. Wisneski, H. Ishii, A. Dahley, M. Gorbet, S. Brave, B. Ullmer, and P. Yarin. Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information, pp. 29–32, 02 1998.
- [9] 中野 亜希人, 脇田 玲. 磁性ゲル形状ディスプレイの開発とインタラクション手法の検討. 情報処理学会論文誌, 54(4):1528–1537, 04 2013.
- [10] 山本 景子, 長沼 大樹, 油井 俊哉, 原田 誠史, 福谷 和芳, 堀 修生, 福澤 貴之. Astral Body: 強磁性粉末を用いた動的なディスプレイによる生き物らしさの表現. 日本バーチャルリアリティ学会論文誌, 24(3):313–323, 2019.
- [11] 光栄. JR 大阪駅「水の時計」. <http://www.koeiaquated.co.jp/works/detail/jrosakamizutokei.html>, 2011.
- [12] 児玉 幸子, 宮島 靖. 音楽に同期する磁性流体彫刻(「アートamp;エンタテインメント」特集). 日本バーチャルリアリティ学会論文誌, 12(3):247–257, 2007.
- [13] 佐川 俊介, 小川 剛史. バブルディスプレイ: 水中の気泡を用いたインタラクティブ映像システム. 情報処理学会論文誌デジタルコンテンツ (DCON) , 2(1):16–23, 02 2014.
- [14] 野間 直生, 塚田 浩二. RippleSpike: スパイク表現と波紋表現を組み合わせたアンビエントディスプレイ. 情報処理学会 インタラクション 2024 論文集, pp. 433–438, 2024.
- [15] 野間 直生, 塚田 浩二. RippleSpike: スパイク表現と波紋表現を組み合わせたアンビエントディスプレイの試作と評価. 情報処理学会研究報告, 2024-HCI-207(13), pp. 1–8, 2024.
- [16] 脇田 玲, 中野 亜希人. Rheologic Interaction: 流動的プログラマブルーマターを用いたインタラクションデザインの試み. 情報処理学会研究報告, 2010 年度 (4):1–5, 12 2010.

ゲームを用いた自動点眼システムの実装と検討

森谷 美羽* 栗原 一貴*

概要. 目の乾燥や疲れなどがあった場合、目薬をさすことで緩和されることがある。しかし、点眼が苦手な人は一定数存在している。そこで本研究では、ゲームをしながら自動で点眼ができるシステムを提案する。本システムはobniz Board、それによって制御されるDCモータ駆動ポンプとボタン、スマートフォン上で使用するWebアプリケーション等で構成されている。本システムはゲームが持つプレイヤーを引き込む特性を活かし、ゲームプレイ中に点眼することで被点眼者の心理的負荷の軽減を目指す。さらに、点眼のタイミングの異なる3つのゲームルールを用いたユーザスタディを行うことで、ゲーム中の点眼液の噴射がプレイヤーに与える心理的影響を明らかにし、点眼の心理的負担を軽減するための方法や点眼のエンタテインメント化に関する検討を行った。その結果、ゲームルールによってプレイヤーの感じる点眼の快・不快度合いが異なることと、ゲームを用いることで点眼への意識を逸らすことができる可能性が示唆された。

1 はじめに

目の乾燥や疲れ、アレルギーなどの症状があった場合、目薬をさすことで緩和されることがある。しかし、目薬をさすことに苦手意識を持つ人は一定数存在しているが、それを支援するシステムは限られている。

そこで、本研究では点眼が苦手な人向けに、ゲームをしながら自動で点眼が行えるシステムを提案する。ゲームプレイには、プレイヤーを引き込む特性がある。この特性を活かし、ゲームプレイ中に点眼を行うことで、被点眼者の点眼に対する意識を紛らわせること、およびゲームプレイ中は目が開いていやすい状態であり、点眼の成功率を高められることを期待する。また、点眼のタイミングの異なる3つのゲームルールを用いたユーザスタディを行うことで、ゲーム中の点眼液の噴射がプレイヤーに与える心理的影響を明らかにし、点眼の心理的負担を軽減するための方法や点眼のエンタテインメント化に関する検討を行う。

本論文は以下の構成になっている。2章で関連研究について述べ、本研究の位置付けを明らかにする。3章で実装したシステムの詳細について記述する。4章で評価実験について述べ、5章で結果と今後の展望を考察する。6章でまとめとする。なお、本稿では以後、提案システム使用者の側面を強調する際は「ユーザ」を、提案ゲームで遊ぶ人の側面を強調する際は「プレイヤー」をそれぞれ表記として用いる

こととする。

2 関連研究

2.1 眼と液体のインタラクション

エンタテインメントを活用した点眼を苦手とする人への支援研究として、我々は、映像鑑賞中に自動で点眼をするシステムを提案している[1]。この研究では、静的なエンタテインメントコンテンツとして映像に着目し、映像の持つ鑑賞者を引きつける特性を活かし、ユーザが映像に気を取られている間に、自動で点眼を行うものである。吉田ら[2]は、眼の近くに水滴を放出して涙を模倣する開発を行った。これは、擬似的な涙を流すことで、着用者と観察者の悲しみを増大させることを意図したものである。

本研究は、インタラクティブなエンタテインメントコンテンツとしてゲームに着目し、ゲームをしながら点眼を行うことで、点眼の心理的負担を軽減することを検討する点がこれら研究と異なる。

2.2 視線、意識誘導

視線や意識の誘導を用いた研究がある。塚田ら[3]は、被写体の視線をカメラに向かつて、撮影行為を意識させないことで多様な表情を捉えるカメラを提案している。また、撮影中に動物の注意を惹くために、指向性スピーカーを用いて様々な音を照射し、意図的に動物のリアクションを誘発するカメラを提案している[4]。

畠ら[5]は、画像の解像度を制御することで、ユーザに気づかれずに視線を誘導する手法を提案している。また、動いている観察者の注意を誘導する視線誘導法[6]や画像内で高周波フリッカを使用する視

Copyright is held by the author(s).

* 津田塾大学

線誘導法[7]がある。また、香りを使った視線誘導をしている研究[8]もある。加えて、村上ら[9]はプレゼンテーションや発表を行う際、発表者の視線を先導するシステムと、発表者が目配りしにくい箇所を提示して視線を誘導するシステムを提案している。

本研究ではゲームを用いることで、ユーザの意識をゲームプレイに集中させるよう誘導し、点眼への意識を逸らすことを目指す。

2.3 日常生活における行動変容のエンタインメント化

日常生活における行動変容をエンタインメント化している研究は多くある。中森ら[10]は、食べ物を食べることで様々な音を奏でることができるフォーク型楽器を開発した。小坂[11]は、偏食克服のための食育シリアルゲームの開発を行った。この研究は、嫌いな食べ物を摂取することで、ゲームクリアを目指すシステムである。市村[12]は、子供が楽しく歯磨きを行えるようにするために歯ブラシ動作計測機能付きアプリケーションの開発を行った。このシステムには、歯ブラシで磨いた箇所近辺の歯が綺麗になっていく様子を表示するゲーミフィケーション機能が備えられている。また、掃除を楽しくするために、掃除機に加速度を検出するデバイスを取り付け、ゲーム要素を取り入れることで、掃除の作業をゲーミフィケーション化する試みが行われている[13]。小笠原ら[14]は、掃除を楽しくするためにインタラクティブな掃除機とそのアプリケーションを提案、試作し、掃除の習慣がない人が掃除に興味を持つきっかけや、親子のコミュニケーションを促進し、子供に掃除の習慣をつける手助けとなる可能性を検討している。杉野ら[15]は、キッチンの道具やダイニングテーブルにセンサを組み込むことで、家事や食事の動きに合わせて効果音や音楽を再生し、家事や食事の楽しさを向上させるとともに、家族が家事に関心を持つきっかけを提供するシステムを提案している。このシステムは、音で子供の状態を知らせることで、安心して調理できる環境を整えることも目的としている。吉野ら[16]は、節電を目的としてコンセントプラグを抜くことを習慣化するためのシステムを開発している。片桐ら[17]は摂食行動をゲームの進行条件として設定することで、ユーザが自発的に偏食を改善し、咀嚼回数を増やすことを目的としたゲームシステムを開発している。

本研究では、点眼の心理的負荷の軽減とエンタインメント化を探求する。

3 提案システム

3.1 システム構成

提案システムの構成について説明する。図1は提

案システムの全体図である。点眼筐体は三脚に設置している。ユーザは点眼筐体を立った状態あるいは座った状態で水平方向に覗き込み、点眼を行う。図2は、ユーザから見える点眼筐体の内部である。点眼液は、図2の噴射口から噴射される。



図 1. 全体図

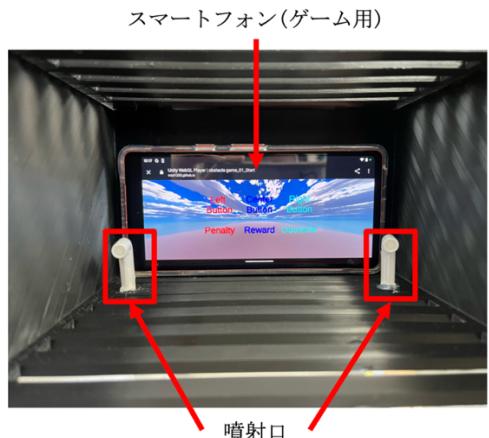


図 2. ユーザ視点

3.2 ハードウェア

実装したハードウェアは、点眼筐体、DCモータ駆動ポンプ、obniz Board、電源供給用モバイルバッテリ、スマートフォン、物理的なボタンからなる。obniz Boardでポンプとボタンの入出力制御を行い、スマートフォンではWebブラウザ上でのゲームプレイを可能としている。ポンプ起動から目薬が点眼されるまでの平均時間は0.262秒であった。点眼筐体の噴射口(図2)は市販のアルコールディスペンサーのものを流用した。この噴射口はL字型であり、微細な多数の穴が開いているため、噴射する液体は直線的にではなく霧状に拡散する。噴射口から接眼面までの距離は15cm、長方形の窓状の接眼面のサイズは縦10cm、横18cmであり、モータの出力を調整し液体を噴射すると接眼面において縦9cm、横13cm程度の範囲に液体が届く。これは接眼面に正対して覗き込んだ場合にはほぼ確実に液体が眼球に届く性能であり、これまで著者らの経験上、体験者の

ゲームを用いた自動点眼システムの実装と検討

まぶたが開いている状況では点眼が失敗する状況は発生していない。

3.3 ゲーム

本システムで扱えるゲームは、ボタン1つと、スマートフォン内蔵のカメラに映る両目の動きのジェスチャで操作でき、なんらかのタイミングで点眼液の噴射を行うことを制約として多様なデザインを検討可能である。本研究ではまずはどのようなタイミングで点眼液の噴射を行うことが妥当かを検証するための簡易なゲームをUnityにより実装した。

このゲームでは、プレイヤがボタンを押すことで、ユーザーの操作する主人公キャラクタがジャンプする。これを利用し、画面奥方向から迫ってくる障害物を飛び越えて進んでいくというジャンプアクションゲームである。

さらに、点眼液の噴射に関して、3つの異なるゲームルールを実装した。これらは、点眼液の噴射をゲーム内における罰、報酬あるいは妨害の要素として対応づけているものである。一般的に、液体の目への噴射はネガティブな印象を与える可能性が高いことが想像されるが、アミューズメントパークでのアトラクションなどの中には、水に濡れることについて、爽快感が得られる、ポジティブなエンタテインメントとしてとらえているものもある。本研究では、ゲームデザインの中で罰、報酬、妨害として液体の目への噴射を扱った場合に、プレイヤに与える心理的影響を明らかにし、点眼の心理的負担軽減に活かすことを意図している。

図3がプレイヤ視点でのゲーム画面となっている。ユーザーは図1のボタンを押すことで、正面の壁を飛び越えて進むことができる。

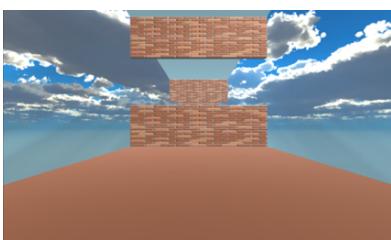


図3. ゲーム画面（プレイヤ視点）

3.3.1 ゲームルール「罰」

点眼液の噴射を「罰」として対応付けたゲームルールについて述べる。このルール下においては、画面左上に4つのハートが表示されており、これがライフゲージとなっている。プレイヤが前方から迫る壁に衝突した際、ライフゲージが1つずつ消滅していく。4回壁に衝突したら、ゲームオーバとなり、罰（ペナルティ）の意味合いで点眼液が噴射され、ゲーム終了となる。

3.3.2 ゲームルール「報酬」

点眼液の噴射を「報酬」として対応付けたゲームルールについて述べる。このルール下においては、ステージ上に表示されているコインを獲得し、スコアを稼いでいく。スコアがある一定の値を超えた場合、ゴール時に報酬の意味合いで点眼液が噴射され、ゲーム終了となる。

3.3.3 ゲームルール「妨害」

点眼液の噴射を「妨害」と対応付けたゲームルールについて述べる。このルール下においては、ゲームプレイ中にある3つの地点に到達した時、点眼液が一時的に噴射され、かつそのままゲームが進行される。目に点眼液が付着した状態でのゲームプレイを強制されるため、正常な視覚情報の認識が一定時間妨害されることを意図した。合計3回の点眼液噴射のうち、ゴールに到達したら、ゲーム終了となる。

4 評価実験

4.1 仮説

ゲーム中の点眼液の噴射がプレイヤに与える心理的影響を明らかにし、点眼の心理的負担を軽減するための方法や点眼のエンタテインメント化に関する検討を行うため、以下の3つの仮説を立てた。

- ・ 点眼液の噴射をゲーム中の「罰」「報酬」「妨害」として対応づけた場合、快または不快の感覚について、プレイヤに異なる影響を与える（仮説1）。
- ・ ゲームをしながら点眼を行うことで、点眼への意識が逸らされる（仮説2）。
- ・ ゲームをしながら点眼を行うことで、点眼を楽しく行うことができる（仮説3）。

4.2 手法

実験協力者9人を対象にした。実験協力者に、点眼に対する好き嫌いの感覚としての苦手意識についての7段階評価（4をニュートラルとし、7が「非常に点眼が好き」、1が「非常に点眼が嫌い」）のアンケートを行ったところ、平均値は4.111、標準偏差は1.448であった。

ゲーム中の点眼液の噴射を「罰」「報酬」「妨害」と対応づけた3つの条件、およびゲームをせずにボタン（図1）を押したら直ちに点眼液が噴射される条件、の合計4条件を体験する実験を行った。ゲーム中の点眼液の噴射を「罰」とする条件をP条件、ゲーム中の点眼液の噴射を「報酬」とする条件をR条件、ゲーム中の点眼液の噴射を「妨害」とする条件をH条件とした。それぞれの条件を視覚的に区別しやすくするため、図4のようにゲーム内の障害物の色合いを変化させた。また、ゲームをせずに点眼液の噴射が行われる条件をC条件とした。



図 4. ゲームルール画面

実験協力者には、各実験前に以下のような教示を行った。P 条件実施前には、画面左上に表示されるライフゲージが無くなかった場合にゲームオーバーになることを伝えた。R 条件実施前には、ゴール地点でスコアが 40 以上あればゲームクリアとなることを伝えた。H 条件実施前には、ゴールを目指すように指示した。C 条件実施前には ボタンを押すと点眼液が噴射されることを伝えた。

なお、点眼液噴射回数に関する実験条件を統制するため、以下の工夫を行っている。P 条件ではゲーム終盤で極端に難易度が高くなり、必ずゲームオーバーになるレベルデザインを行っている。R 条件では、どのようにプレイしてもゴール時のスコアが 40 以上になり、必ずクリア条件を満たすようにレベルデザインを行っている。H 条件では、どの実験協力者においても必ずランダムにステージ上の 3箇所で点眼液の噴射が行われる。

実験は、各実験協力者が 4 条件すべてを体験する被験者内計画であり、条件の提示順は実験協力者ごとにランダムにすることで順序効果を相殺した。実験協力者は各条件の体験後に、「点眼液の噴射に対する快と不快の度合い」、「点眼液の噴射をどの程度意識していたか」、「楽しく点眼できたか」について 7 段階リッカート尺度で回答した。なお、「点眼液の噴射に対する快と不快の度合い」については、ゲーム内での点眼液の噴射の影響を調査する意図の設問のため、ゲームを用いない C 条件では回答は得ていらない。また、すべての実験終了後に自由記述によりフィードバックを求めた。

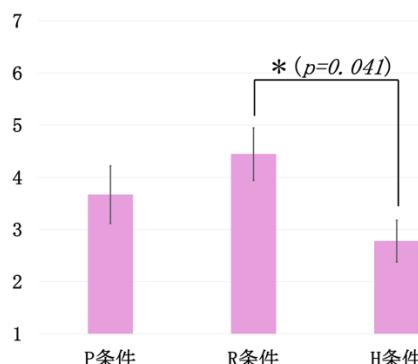


図 5. 点眼液の噴射に対する快と不快の度合いの平均値と標準偏差

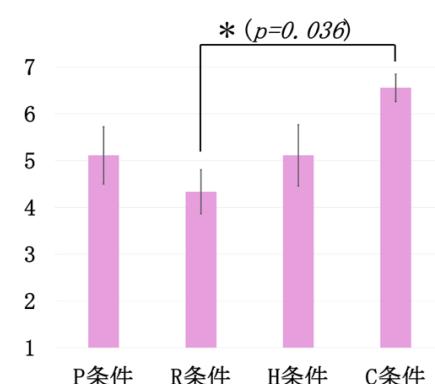


図 6. 点眼液の噴射をどの程度意識していたかの平均値と標準偏差

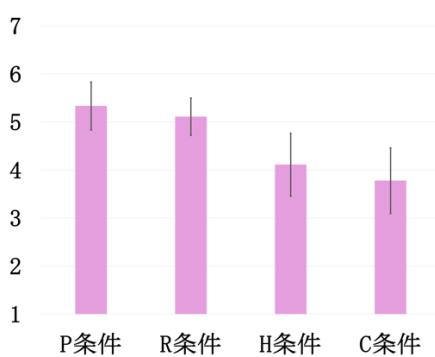


図 7 楽しく点眼できたかの平均値と標準偏差

4.3 結果と分析

「点眼液の噴射に対する快と不快の度合い」、「点眼液の噴射をどの程度意識していたか」、「楽しく点眼できたか」の回答について、実験協力者の平均値を統計的仮説検定した。

4.3.1 点眼液の噴射に対する快と不快の度合い

図 5 は「点眼液の噴射に対する快と不快の度合い」の回答を P 条件と R 条件、H 条件で比較し、それらの平均値と標準偏差を表したグラフである。4 をニュートラルとして、7 が「非常に快適に感じた」、1 が「非常に不快に感じた」である。P 条件の平均値は 3.667、標準偏差は 1.658 であった。R 条件の平均値は 4.444、標準偏差は 1.509 であった。H 条件の平均値は 2.778、標準偏差は 1.202 であった。フリードマン検定を行ったところ、有意差が認められた ($p=0.021$)。事後検定としてボンフェローニ補正を行った Conover 検定の結果、R 条件と H 条件間に有意差が認められた ($p=0.041$)。

4.3.2 点眼液の噴射をどの程度意識していたか

図 6 は「点眼液の噴射をどの程度意識していたか」の回答を P 条件と R 条件と H 条件、C 条件で比較し、それらの平均値と標準偏差を表したグラフである。1 が「全く覚えていられなかった」、7 が「常に覚えていられた」である。P 条件の平均値は 5.111、

ゲームを用いた自動点眼システムの実装と検討

標準偏差は 1.833 であった。R 条件の平均値は 4.333、標準偏差は 1.414 であった。H 条件の平均値は 5.111、標準偏差は 1.965 であった。C 条件の平均値は 6.556、標準偏差は 0.882 であった。フリードマン検定を行ったところ、有意差が認められた ($p=0.026$)。事後検定としてボンフェローニ補正を行った Conover 検定の結果、条件 R と C 条件では有意差が認められた ($p=0.036$)。

4.3.3 楽しく点眼できた

図 7 は「楽しく点眼できたか」の回答について、P 条件と R 条件と H 条件、C 条件で比較し、それらの平均値と標準偏差を表したグラフである。1 が「全く楽しめなかつた」、7 が「非常に楽しめた」である。P 条件の平均値は 5.333、標準偏差は 1.500 であった。R 条件の平均値は 5.111、標準偏差は 1.167 であった。H 条件の平均値は 4.111、標準偏差は 1.965 であった。C 条件の平均値は 3.778、標準偏差は 2.048 であった。フリードマン検定を行ったところ、有意差が認められなかつた ($p=0.335$)。

5 考察

5.1 点眼液の噴射とゲーム内における意味

「点眼液の噴射に対する快と不快の度合い」について、平均値の統計的仮説検定により有意差が得られたことから、点眼液の噴射をゲーム中の「罰」「報酬」「妨害」として対応づけた場合、快または不快の感覚について、プレイヤに異なる影響を与えるという仮説 1 は支持された。

また、4 をニュートラルとして R 条件は快寄り、P 条件と H 条件は不快寄りの平均値が得られている。R 条件と H 条件の平均値の間には有意差が見られたことから、R 条件を用いた点眼液の噴射を「快」な効果をもたらす報酬要素として、あるいは H 条件を用いた点眼液の噴射を「不快」な効果をもたらす妨害要素として、ゲームデザインに活用しそれぞれ異なる心理的影響をプレイヤに与えることができる可能性が示唆された。R 条件に関する自由記述に、

「水の噴射を忘れており、ゴール地点で水の噴射をご褒美と感じた」というものがあり、この考察を裏付けるものとなっている。一方で H 条件が不快と感じられた要因としては、様々なものが考えられる。たとえば他の 2 条件に比べて点眼液の噴射タイミングが予測しづらいこと、複数回の噴射が行われることに関する嫌悪、あるいは我々が意図したように、正常な視覚情報の認識が一定時間妨害されることなどである。H 条件に関する自由記述に、「いきなり水が来てびっくりしました」、「いつ噴射されるのかタイミングが読めなく、ゲーム中少し不安を感じました」、「どんな基準で噴射されるか分からなかつたた

め、対策ができず、ずっと緊張感を持たないといけませんでした」などがあり、他の 2 条件に比べて点眼液の噴射タイミングが予測しづらいことが要因の少なくとも一つであった可能性が示唆される。これらの要因の詳細な特定は今後の課題である。

5.2 ゲームにより点眼の意識を逸らすことについて

「点眼液の噴射をどの程度意識していたか」について、ゲームを用いる R 条件と、ゲームを用いない C 条件の平均値の間に有意差が得られた。このことから、ゲームをしながら点眼を行うことで、点眼への意識が逸らされるという仮説 2 は、R 条件を採用するという条件つきで支持された。

前節の議論により、R 条件による点眼液の噴射は「快」な感情をもたらす効果が期待できるため、ゲームを用いた点眼システムを構成する際に相乗効果がある。すなわち、報酬要素としてゲームクリア時に点眼液を噴射させるゲームは、「快」の感情をもたらし、かつ点眼への意識をそらす可能性があるため、現状では点眼の心理的負担を減らす上で最良の選択肢といえる。

以下に C, P, H, R 条件それぞれについて、さらに詳しく考察する。まず、C 条件はボタンを押すと、すぐに点眼液が噴射されることを事前の教示で行っていたため、点眼液の噴射への意識が消えにくく結果であったのは自然な結果である。P 条件に関しては、ライフゲージが 0 になると点眼液の噴射がされるということを実験協力者が予想し易い状況であったため、ライフゲージが減るたびに点眼液の噴射が思い起こされて意識が消えにくかった可能性が考えられる。一方で、H 条件はランダムで 3箇所の地点で点眼液の噴射が行われる。1 度噴射されてもゲームが終了しないので、1 度目の噴射以降はゲームの終了が宣言されるまで点眼液の噴射への意識が消えにくかったことが予想される。前節で引用した H 条件に関する自由記述の「いつ噴射されるのかタイミングが読めなく、ゲーム中少し不安を感じました」および「どんな基準で噴射されるか分からなかつたため、対策ができず、ずっと緊張感を持たないといけませんでした」などは、この分析を裏付けるものである。R 条件に関しては自由記述において、「最初はいつ水の噴射があったかが気になっていたが、ゲームが難しく、集中できた」という意見があった。このことから、ゲームの難易度を高くすることでゲームに集中することができ、点眼液の噴射から意識を逸らすことができる可能性が示唆される。一方で、ゲームが簡単だったと答える実験協力者も存在した。実験参加者が感じるゲームの難易度は個人間で異なるため、ゲーム難易度を調節できる機能を検討して

いくことが必要だろう。

5.3 提案システムと点眼の楽しさについて

「楽しく点眼できたか」について、平均値の統計的仮説検定により有意差は得られなかった、このことから、ゲームをしながら点眼を行うことで、点眼を楽しく行うことができるという仮説3は支持されなかつた。

ゲームをしながら点眼を行う条件と、ゲームをせずに点眼を行う条件間に有意な差が見られなかつた原因として、従来の点眼方法とは異なる体験であつたC条件の新奇性効果の可能性が考えられる。C条件に関する自由記述において、「点眼されるタイミングを自身で決断できるので良かった」という意見や、「ボタンを自分で押して水が噴射されることには少しワクワクしました」という意見もあったことが、この考察を裏付けるものである。本研究の目的の一つは点眼の心理的負担の軽減であるため、ゲームを用いざとも点眼を楽しく行える方向性としてのC条件の活用と発展は今後の研究方向として興味深い。

また「楽しく点眼できたか」の平均値はゲームを用いたP, R, H条件において比較的「非常に楽しめた」寄りの値となっている。自由記述においても「ゲーム要素がある条件P, R, Hの方が、点眼を楽しめました」というコメントが得られている。今回は比較実験に用いるため、シンプルなゲームデザインを採用したが、より作り込まれた本格的な点眼用ゲームを開発することによって、より楽しい点眼を実現することは可能であると考える。これは今後の課題である。

最後に、快・不快の感情と楽しさについて言及する。ナイーブには「快」な感情を催すエンタテインメントがより強い楽しさを生むように想像されるが、「楽しく点眼できたか」の平均値において、P条件の値が高い点は興味深い。5.1節においてはP条件はネガティブよりの平均値となっており、どちらかというと「不快」な感情を催すものとして受け止められている。また、5.2節においてはP条件は比較的ゲーム中の点眼への意識が消えにくいものとして受け止められている。それにもかかわらず実験協力者が「楽しい」と感じるという点は、ホラーコンテンツなどの、恐怖や緊張、そこからの開放を楽しむエンタテインメントと共に通性があることが想像される[18]。点眼の心理的負担軽減という点からは相反するが、結果的に楽しく点眼を行う工夫として、今後探求を深めたい。

5.4 点眼に対する苦手意識と実験結果について

本実験の実験協力者は「点眼の好き嫌いの感覚としての苦手意識度合い」アンケートの7段階評価について、平均値が4.111、標準偏差が1.448であつ

たことから、平均的には好き嫌いがニュートラルに近いが、ある程度多様性のある集団であったと考えられる。したがつて5.1節および5.3節で得られた知見は、特に点眼への苦手意識を限定しない多様なプレイヤ層を対象にして、液体噴射をエンタテインメント活用する際の知見として応用が期待される。

一方で、本実験は点眼への苦手意識がある人々を選別して行った実験結果ではないため、5.2節で得られた知見をそのまま点眼への苦手意識がある人々へと適用することはできない。今回の実験では、苦手意識のアンケートで2と回答した、点眼への苦手意識があると考えられる実験協力者から、苦手な理由として「目に鋭利なものを近付けるのが怖いから(刺してしまって、目を傷つけそう)」および「まばたきと点眼のタイミングをあわせづらいから」というコメントがあった。現状の点眼機構は噴射口と接眼面が離れており、噴射口への物理的な接触がおこらないデザインとなっているほか、我々はスマートフォンのカメラにより開眼状態を検知し、開眼時に自動的に噴射を行う機能も試作している。これらの機能性により点眼への苦手意識がある人々の心理的負担を軽減できるかどうかについては今後の課題である。

6 まとめ

本研究では、ゲームをしながら自動で点眼ができるシステムを開発し、点眼液の噴射を「罰」「報酬」「妨害」に対応付けたジャンプアクションゲームを用いたユーザスタディにより、ゲーム中の点眼液の噴射がプレイヤに与える心理的影響を明らかにし、点眼の心理的負担を軽減するための方法や点眼のエンタテインメント化に関する検討を行つた。その結果、「快」な効果をもたらす報酬要素として、あるいは「不快」な効果をもたらす妨害要素として点眼液の噴射をゲームデザインに活用しそれぞれ異なる心理的影響をプレイヤに与えることができる可能性が示唆された。また、報酬要素としてゲームクリア時に点眼液の噴射することは点眼の意識を逸らす効果がある可能性が示唆された。また、ゲームを用いて点眼体験を楽しくするための方策を議論した。今後は、今回制作したジャンプアクションゲームの改良や他のゲームでの点眼の心理的負担軽減とエンタテインメント化を検討していく。

謝辞

本研究の一部は、2024年度東芝デバイス&ストレージ株式会社学術奨励制度、JSPS科研費JP24K15248、および中山隼雄財團研究助成の支援

を受けて行われた。謹んで感謝の意を表する。

参考文献

- [1] 森谷美羽, 栗原一貴: ぱちぱちドロップ: 目薬が苦手な人のための自動点眼システムの試作, エンタテインメントコンピューティングシンポジウム 2023 論文集, pp.290-292, 2023.
- [2] Shigeo Yoshida, Takuji Narumi, Tomohiro Tanikawa, Hideaki Kuzuoka, and Michitaka Hirose: Teardrop Glasses: Pseudo Tears Induce Sadness in You and Those Around You., In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI' 21)*, Article 508, pp.1-12, 2021.
- [3] 塚田浩二, 沖真帆: EyeCatcher: 多様な表情を撮るカメラ, コンピュータソフトウェア, 27巻 1号, pp.1_89-1_100, 2010.
- [4] Koji Tsukada, Maho Oki, Kazutaka Kurihara, Yuko Furudate: AnimalCatcher: a digital camera to capture various reactions of animals, *arXiv preprint arXiv:1508.03653*, 2015.
- [5] 畑元, 小池英樹, 佐藤洋一: 解像度制御を用いた視線誘導, 情報処理学会論文誌, Vol.56 No.4, pp.1152-1161, 2015.
- [6] Arisa Kohtani, Shio Miyafuji, Keishiro Uragaki, Hidetaka Katsuyama, and Hideki Koike: MOSion: Gaze Guidance with Motion-triggered Visual Cues by Mosaic Patterns, In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI' 24)*, Article 374, pp.1-11, 2024.
- [7] N. Waldin, M. Waldner, and I. Viola: Flicker Observer Effect: Guiding Attention Through High Frequency Flicker in Images, *Computer Graphics Forum*, Vol. 36. No. 2, 2017.
- [8] 伴野明, 神田こより, 伴野啓介: 香り提示による視線誘導効果とどう物体追跡効果の検討, 電子情報通信学会論文誌 D 94.5, pp.803-813, 2011.
- [9] 村上梨沙, 塚越さくら, 栗原一貴: 発表時における視線誘導システムの検討, 研究報告エンタテインメントコンピューティング (EC), Vol.2017-EC-43 No.14, pp.1-6, 2017.
- [10] 中森玲奈, 塚田浩二, 椎尾一郎: 食べテルミン, 情報処理学会インタラクション2011論文集, pp.367-370, 2011.
- [11] 小坂崇之: 偏食克服を目的とした食育シリアルゲーム 「Food Practice Shooter」, エンタテインメントコンピューティングシンポジウム 2013 論文集, pp.47-50, 2013.
- [12] 市川哲: migaco: 子供が楽しく歯みがきが行えるようにするための歯ブラシ動作計測機能付きアプリケーション, 情報処理学会論文誌, Vol. 61, No.1, pp.95-102, 2020.
- [13] 市村哲, 矢澤崇史, 戸丸慎也, 渡邊宏優: 家事をゲーム化・分散協調化する試み ~掃除への適用~, マルチメディア, 分散協調とモバイルシンポジウム 2014 論文集, pp.1285-1290, 2014.
- [14] 小笠原遼子, 山木妙子, 塚田浩二, 渡邊恵太, 椎尾一郎: インタラクティブな掃除機, エンタテインメントコンピューティング 2007 講演論文集, pp.71-74, 2007.
- [15] 杉野碧, 岩渕絵里子, 椎尾一郎: 家事と食卓を楽しくする「歌うダイニングキッチン」の試作, エンタテインメントコンピューティングシンポジウム 2007 論文集, pp. 67-70, 2007.
- [16] 吉野孝, 森田沙奈: AR を用いたコンセントプラグを抜く習慣付け支援システム「ぷらとん」の開発と評価, マルチメディア, 分散協調とモバイルシンポジウム 2013 論文集, pp.632-640, 2013.
- [17] 片桐祐規, 谷中俊介, 二階雅弘, 小坂崇之: 偏食改善および咀嚼増加を目的としたゲームシステムの開発, エンタテインメントコンピューティングシンポジウム 2016 論文集, pp. 191-194, 2016.
- [18] 水口充, 片寄晴弘: エンタテインメントコンピューティング研究における評価問題の解決に向けての施策の実践, インタラクション 2019 予稿集, pp.141-150, 2019.

未来ビジョン

1つ目の未来ビジョンは、ゲーム以外の多様なエンタテインメントと点眼を組み合わせることの検討である。我々はこれまでに受動的なエンタテインメントとして動画鑑賞を、そして本研究においてインタラクティブなエンタテインメントとしてゲームを対象に、点眼のエンタテインメント化について検討してきた。未開拓のエンタテインメントと点眼を組み合わせることで、より心理的負荷の小さい点眼の実現や、新しいエンタテインメントの実現を目指したい。

2つ目の未来ビジョンは、多様な液体を多様な身体部位へ噴射することとディジタルゲームを組み合わせた新たなエンタテインメントの検討である。本研究は点眼を念頭

に、水または点眼薬の目への噴射を想定しシステムデザインおよびゲームデザインを行ってきた。これに縛られず、色や温度、粘性や香り、味などが異なる無害な液体を用いることで、より多様な表現が可能になるだろう。さらに、目や目周辺の皮膚以外の身体部位への液体の噴射についても探求することは興味深い。たとえば水遊び時に耳や鼻や口に水がかかったり浸水したりする体験は独特のものであるだろう。またたとえば意図せず身体が雨に濡れるときに、快や不快の感覚を生じることも容易に想像される。健康を損なわない範囲で、これらを再現するような、あるいはより拡張された体験を人工的に創出しディジタルゲームに組み入れたエンタテインメントに活用していきたい。

オノマトペを用いた脳波制御訓練手法による魔法ゲームシステム

二瓶 朝渉^{*†} 平野 恵旺^{*‡} 渡邊 恵太[†]

概要. ブレインコンピュータインターフェイス (BCI) は、脳と外部機器の直接通信を可能にする技術である。脳波 (EEG) ベースの非侵襲型 BCI は、主に医療用途に利用されるが、ビデオゲームへの利用にも関心が高まっている。BCI を使ったビデオゲームは、プレイヤーに斬新で魅力的なモダリティを提供し、ゲームへの没入感を高める可能性がある。本研究では、オノマトペを用いた脳波制御訓練手法により、心的イメージで魔法を制御するビデオゲームを開発した。本システムでは、ユーザは BCI を装着し、魔法の視覚イメージとオノマトペを用いた発話イメージを同時に想起することで、魔法を操作することができる。本稿では、制御パフォーマンスとビデオゲームの体験の両方を評価することにより、本手法の有用性を検証した。その結果、本手法が BCI 制御において一定の制御性能を示し、ビデオゲーム体験の向上に有用である可能性を示唆した。また、ゲームへの没入感や楽しさについても高い評価を得た。

1 はじめに

ブレインコンピュータインターフェース (BCI) は、ニューラルインターフェースやブレインマシンインターフェース (BMI) とも呼ばれ、脳と外部機器との直接通信を可能にする [11]。脳波ベースの非侵襲型 BCI[9] は、リハビリテーション [3, 13] や車椅子制御 [16]、コミュニケーション支援 [4] などの医療用途に利用されることが多く、運動機能障害を持つユーザでも外部機器を制御し、インタラクションを可能にする [15]。一方、健常者のための BCI への関心も高まっており、ビデオゲームへの応用も注目されている。BCI を用いたビデオゲーム体験は、プレイヤーにとって斬新で面白いモダリティを感じさせることや、ゲームへの没入感の向上 [19, 17] をもたらす可能性がある。しかし、この技術が一般に受け入れられるためには、ユーザビリティとユーザ体験を考慮した BCI システムを設計する必要がある [14]。

BCI を用いたビデオゲーム制御において、運動イメージによる制御が一般的なパラダイムである [18]。運動イメージは BCI パフォーマンスが安定しやすく、一般的な手法であるが、イメージ想起が難しいことや、ユーザの運動イメージと画面上の視覚的なフィードバックに乖離がある場合、直感的な操作性が損なわれる [1, 8]。一方、視覚イメージによる制御は直感的な操作が可能であるが、BCI パフォーマンスが安定しないため、プレイヤーが意図した操作を正確に反映することが難しく、ビデオゲーム体験の質を低下させる問題がある [7]。

平野らは、オノマトペを用いた脳波制御訓練手法



図 1. オノマトペを用いた脳波制御訓練手法により魔法制御できるビデオゲームを開発した。ユーザは、魔法の視覚イメージとそれに対応するオノマトペによる発話イメージを同時に想起することにより魔法を制御できる。

を提案した [23]。オノマトペには、構成する音と意味の間に密接な結びつきがある音象徴性と呼ばれる性質があり、直感的な理解や感覚的なイメージを与えることができる [6]。この手法では、発話イメージをオノマトペで表現し、視覚イメージとオノマトペによる発話イメージを同時に想起するマルチモーダルなアプローチにより、BCI パフォーマンス向上と認知負荷軽減の可能性を示した。例えば、画面上に表示されたキューブを脳波制御により回転させる時、キューブが回転する視覚イメージと同時に「クルクル」という回転に関連するオノマトペを繰り返し想起する。視覚イメージのみの脳波制御は不安定だが、オノマトペによる発話イメージと組み合わせることで、特定の脳波パターンを強化できる。また、本手法を用いることにより、ビデオゲームや UI 操作

Copyright is held by the author(s).

* Authors contributed equally.

† 明治大学総合数理学部

‡ 明治大学大学院先端数理科学研究科



図 2. BCI を用いてイメージ制御により魔法を発動し、敵を攻撃するゲーム。画面中央部には敵と体力を示すゲージを表示し、画面下部には発動した魔法を表示した。

作において、実現可能なインタラクションとその応用可能性について検討した [22]。

本研究では、オノマトペを用いた脳波制御訓練手法により魔法制御できるビデオゲームを開発した(図1)。ユーザは、魔法の視覚イメージとそれに対応するオノマトペによる発話イメージを同時に想起することにより魔法を制御する(図2)。例えば、炎の魔法を生成するには、ユーザは心的イメージで炎を視覚化し、日本語のオノマトペ「メラメラ」を繰り返し脳内で唱える。

本稿では、制御パフォーマンスとゲーム体験を評価することにより、本手法がBCI制御のビデオゲーム体験を向上させる可能性を検証した。その結果、本手法は参加者が 60%以上意図した通りに魔法を制御できることから、一定の制御性能を示した。また、ゲームへの没入感や楽しさについて高い評価を得たことから、BCIを用いたイメージ制御とオノマトペの高い親和性がプレイヤー体験を向上させる可能性があることを示唆した。

2 ゲームシステム

2.1 システム構成

本システムでは、EMOTIV EPOC X¹を用いて脳波計測を行い、MATLABを用いて脳波解析を行った。また、Unityを用いてリアルタイムに脳波で制御可能なゲームを開発した。EPOC Xは、頭皮からの電気信号を記録する 16 個のセンサー（内 2 個はリファレンス）を搭載しており、導電性を向上させるために生理食塩水で浸したフェルトを各センサーに取り付けている(図3)。本研究では、サンプリング周波数を 256Hz に設定し、脳波データを計測した。

取得した脳波データは、無線通信によりコンピューターに送信し、MATLAB 上で解析を行った。データの前処理、特徴抽出、特徴分類の一連の処理をリアルタイムに実行するシステムを、MATLAB 上で実装し、MATLAB 上で分類した結果は、UDP 通信により Unity に送信する。そして、Unity では受信した分類結果に基づいてアプリケーションを制御するシステムを構築した。本研究では、Unity 上で魔法ゲームを実装し、プレイヤーは脳波制御によりシステムをリアルタイムに操作し、視覚的なフィードバックを受けつつゲームをプレイした(図4)。

2.2 BCI システム

2.2.1 前処理

本研究では、MATLAB (2020b: MathWorks) を用いて脳波データの解析を行った。まず、実験で計測した生脳波データに対して前処理を行うために、1-40Hz の範囲でバンドパスフィルタを作成した。そして、設計したバンドパスフィルタを各チャネルの脳波データに適用し、脳波データからノイズ成分を軽減した。フィルタリング後、刺激に対する脳の反応を特定するために、各刺激に対応するエポックに分割し、収集した脳波データを正確に解析した。

2.2.2 特徴抽出

特徴抽出には、共通空間パターン (Common Spatial Pattern : CSP) アルゴリズムを用いて行った。CSP は脳波データ解析において広く用いられる信号処理技術であり、2 つの異なるクラス間の差異を強調し、脳波の変動を最大限に分離する空間フィルターを作成できる [2]。このプロセスでは、各脳波データに対し、共分散行列の固有値分解を行うことにより、信号の分散を最大化し、各脳波データを最大限分離する。得られた固有ベクトルは CSP フィルターとして機能し、最も分類性能が高い固有ベク

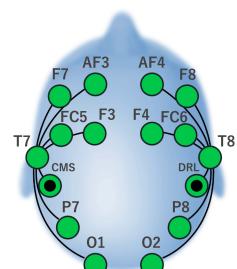


図 3. EPOTIV EPOC X と各センサー配置を示す。

EPOC X は、頭皮からの電気信号を記録するために 16 個のセンサーを搭載し、センサー位置は国際的な 10-20 システムに準拠する(センサー構成: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4)。

¹ <https://www.emotiv.com/epoc-x/>

オノマトペを用いた脳波制御訓練手法による魔法ゲームシステム

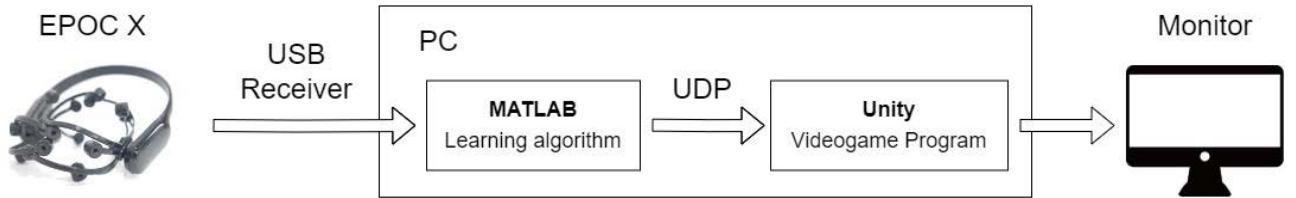


図 4. システム構成図. EMOTIV EPOC X で計測した脳波データは、無線通信によりコンピューターに送信し、MATLAB 上で解析を行った。MATLAB 上で特徴分類した結果は、UDP 通信により Unity に送信する。そして、Unity では受信した分類結果に基づいてアプリケーションを制御するシステムを構築した。

トルを選択することにより、元の脳波データより少ない次元で情報を抽出できるため、機械学習モデルの入力として非常に有効である [12]。そして、CSP フィルターをデータに適用した後、対数分散特徴量を計算した。この対数分散は、機械学習モデルへの入力に利用されることが多く、元のデータの分散を対数スケールで表現することにより、異なるデータソースやセンサーからの信号比較が容易になる。

本研究では、各エポックに対して一定の時間窓を設定し、この時間窓を移動することにより、元のデータから複数のサンプルを得た。時間窓を 2 秒間、オーバーラップを 25% (0.5 秒間) に設定し、元のデータから多様なサンプルを生成した。この時間窓の設定により、データセットを拡大し、特徴抽出の精度を向上させることができる。そして、このように作成したデータセットに対し、CSP フィルターを作成、適用して特徴量を抽出した。また、リアルタイムでの解析においても同様に、2 秒間の脳波データに対し、事前トレーニングで作成した CSP フィルターを適用することで特徴抽出を行った。

2.2.3 特徴分類

特徴分類では、サポートベクターマシン (Support Vector Machine : SVM) を用いて抽出した特徴量の分類を行った。SVM は、特徴空間内で異なるクラスを区分する最適な境界線、すなわち超平面を見つけるために設計された教師あり学習アルゴリズムである [5]。このアルゴリズムは、クラス間のマージンを最大化することを目的としている。さらに、SVM は線形分離可能でないデータセットに対しても効果的であり、元の特徴空間をより高次元の空間にマッピングし、線形分離を可能にする。本システムでは、SVM を用いてリアルタイム BCI システムを開発した。SVM による分類結果を毎秒出力し、その出力結果を用いてアプリケーションを制御できるように設計した。

2.3 ゲームシステム

本研究では、BCI を用いたイメージ制御によって魔法を発動し、敵を攻撃するゲームを Unity で開発した。画面中央部には敵と体力を示すゲージを表示

表 1. 実験前アンケートの内容。炎魔法のオノマトペは「メラメラ」、電気魔法のオノマトペは「バチバチ」に決定した。

- | | |
|-----|---------------------------------|
| (1) | オノマトペを用いることで視覚イメージがより鮮明になったか |
| (2) | オノマトペが視覚イメージを想起しやすいものであったか？ |
| (3) | オノマトペが視覚イメージとマッチしていたか？ |
| (4) | オノマトペを発話イメージしやすかったか？ |
| (5) | オノマトペの発話リズムが視覚イメージのリズムと一致していたか？ |
| (6) | 総合的にそのオノマトペは視覚イメージに適していたか？ |

し、画面下部には発動した魔法を表示した。敵は草属性または水属性を持ち、ランダムに出現する。

ユーザは、心的イメージ制御により炎魔法と電気魔法を発動することができる。BCI システムから、ユーザのイメージの分類結果の入力は 1 秒に 1 回受信し、受信した入力の魔法は 1 段階大きくなる。一方、既に発動している魔法とは異なる入力を受け取ると 1 段階小さくなる。魔法は最大 2 段階まで大きくなり、最大サイズになると自動で発射され、敵に攻撃できる。また、攻撃によるダメージは、魔法の種類と敵の属性により決定し、魔法の種類と敵の属性に応じて効果的な魔法で攻撃すると、より高いダメージを与えることができる。例えば、草属性の敵に炎魔法で攻撃すると、すべての体力ゲージを減らすことができるが、電気魔法で攻撃すると 1/4 しか減らすことができない。同様に、水属性の敵に電気魔法で攻撃すると、すべての体力ゲージを減らすことができるが、炎魔法で攻撃すると 1/4 しか減らすことができない。さらに、魔法発射後、次の魔法が発動できるようになるまで 3 秒のインターバルを設けることで、ユーザのイメージをリセットする時間を確保した。これにより、プレイヤーは次の魔法発動に集中できる。

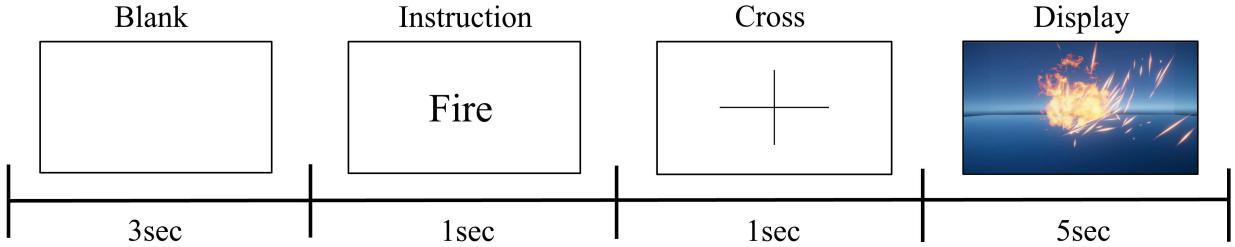


図 5. 動画の指示に従いながらイメージ想起を繰り返すトレーニングを行った。トレーニング動画は、安静 3 秒、指示 1 秒、合図 1 秒、イメージ想起 5 秒のサイクルを繰り返す構成になっている。

3 実験

3.1 事前調査

実験において、炎魔法と電気魔法の制御に最適なオノマトペを選定するために、我々は日本語オノマトペ辞典 [21] から選定した各魔法 5 種類のオノマトペを対象に、イメージ想起に関するアンケートを行った(表 1)。アンケートでは、参加者は魔法が発動している様子の動画を視聴しつつ、魔法の視覚イメージと同時にオノマトペによる発話イメージを想起した。

その後、各オノマトペについて魔法イメージの鮮明さ、想起のしやすさ、視覚イメージとの一致度に関するアンケートを 7 段階リッカート尺度を用いて行った。アンケートは 15 人が回答し、取得したデータを処理した結果、炎魔法のオノマトペは「メラメラ」、電気魔法のオノマトペは「バチバチ」に決定した。

3.2 実験参加者

参加者は 7 名(男性 6 名、女性 1 名、年齢:19-23)の大学生であった。参加者の 1 名は BCI デバイスの使用経験があったが、それ以外の 6 名は BCI の使用経験が無かった。参加者のプライバシーを最大限に尊重し、すべての個人情報は匿名化して保管した。参加者のデータは研究者の間でのみ共有し、それ以外の第三者とは共有しなかった。

3.3 実験手順

本実験では、参加者間の知識量の差や技術に対する不信感がパフォーマンスに影響を与える可能性を考慮し、実験全般の説明を行った後に実験を開始した。実験の流れは、最初に脳波機器を装着し、信号の安定度を確認した。次に、BCI 制御トレーニングを行い、取得した脳波データを解析した。その後、参加者がどの程度正確に魔法を発動できるかを検証するタスクを実施した。タスク終了後には、ゲーム体験に関するアンケートとインタビューを行い、参加者の主観的な評価を収集した。

トレーニングセッションでは、動画の指示に従い

ながらイメージ想起を繰り返すトレーニングを行った。トレーニング動画は、安静 3 秒、指示 1 秒、合図 1 秒、イメージ想起 5 秒のサイクルを繰り返す構成となっている(図 5)。イメージ想起では、炎魔法イメージ、休憩、雷魔法イメージ、休憩の順番で行い、このサイクルの繰り返しを 5 分間で 1 試行とし、試行間に休憩を設け 4 試行を行った。トレーニング終了後、取得した脳波データを解析することにより特徴量を抽出し、SVM に学習させた。これにより、参加者個人に最適化した分類器と、リアルタイム脳波データから特徴量を抽出するための CSP フィルターを作成した。この一連のプロセスにより、参加者の脳波パターンを効果的に学習し、魔法のイメージ想起を識別できるシステムを構築した。

タスクでは、参加者は BCI を用いてイメージ制御により魔法を発動し、敵を攻撃するゲームをプレイした。参加者には、ランダムに出現する 2 種類の敵に応じて、効果的な魔法のイメージを想起し続けるように指示した。1 試行につき 20 体の敵を倒し、合計 3 試行を行った。実験中は、敵の種類と参加者が発動した魔法の入力を計測し、参加者がどの程度正確に適切な魔法を発動できたかを記録した。これにより、オノマトペを用いた脳波制御訓練手法によるゲームプレイにおける参加者の制御パフォーマンスを定量的に評価した。

タスク終了後、プレイヤーのゲーム体験を評価するため、5 段階のリッカート尺度を用いた主観評価アンケートを実施した。このアンケートでは「操作性」「没入感」「体験の楽しさ」の 3 つの観点から 12 項目の質問を用意した。

4 結果

4.1 制御パフォーマンス

オノマトペを用いた脳波制御訓練手法の有効性を検証するために、魔法ビデオゲームを用いた実験を行った。実験では、出現した敵に対して最適な魔法を選択できた割合を算出し、その結果を制御パフォーマンス指標として用いた。この指標は、参加者が意図通りの魔法を発動できたかどうかを示すものであ

オノマトペを用いた脳波制御訓練手法による魔法ゲームシステム

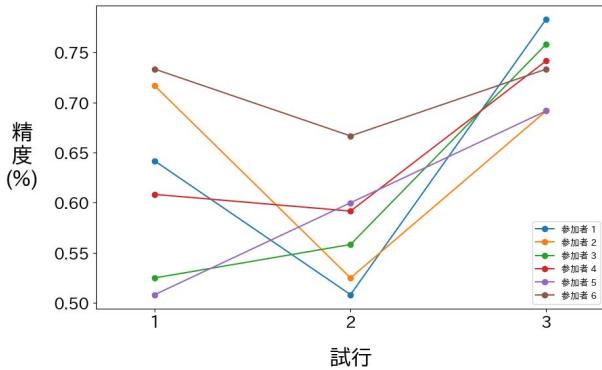


図 6. 各セットにおける制御パフォーマンスを示す折線グラフ。敵を倒す回数を重ねていくことで、精度が向上する傾向があった。

る。なお、実験参加者の一名が全く魔法を制御できなかったため外れ値とした。

実験の結果、参加者が意図通りの魔法を使えた割合は60%以上であった(図6)。このことは、プレイヤーがある程度意図通りに魔法を制御できることを示唆した。また、敵を一體倒すためにかかった平均時間は11.366秒であった。さらに、敵を倒すプロセスの中で、最初の試行で精度が悪かった参加者も、敵を倒す回数を重ねていくことで、精度が向上する傾向があった。これは、BCI制御トレーニングにおいて、BCIシステムの操作習熟度がトレーニング時間に依存する可能性を示唆した。

4.2 ゲーム体験

オノマトペを用いた脳波制御訓練手法による魔法ゲーム体験について、5段階リッカート尺度を用いたアンケートによって評価を行った。評価は「操作性」「没入感」「体験の楽しさ」の3つの観点から実施した(表2)。

操作性に対する評価では、参加者がどの程度自分の意図通りに魔法を制御できたかを示し、評価には4項目の質問を用意した。直感性、自然な操作感の2項目については4を記録したが、操作性と満足感の2項目については3.14を記録した。また、参加者ごとに評価にばらつきが見られ、操作性には個人差があった。没入感に対する評価では、プレイヤーがどれだけゲームの世界に没入できたかを示し、評価には3項目の質問を用意した。3項目の平均値はすべて3.5以上を記録し、2項目において中央値が4を記録した。没入感に関して、個人差はあるものの、本手法を用いた操作によりゲームの没入感が向上する可能性がある。体験の楽しさに対する評価では、魔法ゲームの楽しさがどの程度だったかを示し、評価には5項目の質問を用意した。プレイ体験の斬新さに関する質問では、平均値4.71の高い値を記録した。これは、魔法ゲームが参加者にとって新しい

ゲーム体験を提供することを示唆した。また、魔法の制御の楽しさとエンターテイメント性に関する質問では、平均値4.2以上の値を記録し、プレイヤーにとって楽しいゲーム体験をもたらしたこと示唆した。ゲームの楽しさに関する全体的な評価は、平均値、中央値ともに高い結果となった。

オノマトペを用いた脳波制御訓練手法による魔法ゲーム体験は、操作性にはばらつきがあるものの、楽しいゲーム体験を提供できる可能性を示した。

4.3 インタビュー

オノマトペを用いた脳波制御訓練手法による魔法ゲーム体験について体験評価を行うため、タスク終了後にインタビューを実施した。このインタビューの目的は、魔法ゲームに対するユーザーの主観的な感想や印象を調査することであった。

インタビューの結果、多くのユーザーが魔法ゲームに対して肯定的な印象を抱いていることが明らかになった。例えば、「トレーニング次第で上達するのが面白い」、「後半徐々に上手に操作できるようになった」や「トレーニング次第で上手くなる感覚があった」といった意見を得た。これらの意見は、魔法ゲームが参加者に新しいモチベーションを与えた可能性があることを示唆する。つまり、ゲームを上手にプレイするために、自身の「イメージの技術」を向上させるという目標がユーザーに生じているのである。

一方で、「イメージしている魔法が出ない時にストレスを感じた」、「アニメやゲームの世界で魔法がうまくできない人の気持ちがわかった」といった意見も得た。これらの意見は、魔法ゲームが、プレイヤーに新しいタイプのフラストレーション体験を与えていることを示唆する。つまり、本システムでは、ユーザーが自身の思い通りに魔法を使えないもどかしさも体験できる。この体験は、プレイヤーに新しいゲーム体験をもたらす可能性がある。

従来のゲームでは、魔法技術習得の過程で生じるフラストレーションを体験することはなかった。ゲーム内の魔法の成功や失敗は、主にゲームのルールやパラメータに依存しており、ユーザーの習熟度とは関係しない。これに対し、本システムでは、ユーザーのイメージ想起の巧拙が魔法の成功に直結するため、現実世界で魔法技術を習得するかのような体験を実現できる可能性がある。

5 議論

5.1 BCI操作とゲームの親和性

本研究では、オノマトペを用いた脳波制御訓練手法による魔法ゲーム体験が参加者に良い体験を提供したことを示唆する結果となった。タスク終了後に実施したアンケートおよびインタビューにおいて、

表 2. オノマトペを用いた脳波制御訓練法を用いた魔法ゲーム体験は、操作性、没入感、体験の楽しさを 5 段階のリッカート尺度によるアンケートで評価した。

観点	質問	中央値	平均値
操作性	オノマトペを用いた魔法の生成は、直感的で自然に感じた	4	4
	オノマトペによる魔法の制御は、操作性や応答性に優れていると感じた	3	3.14
	オノマトペを用いた魔法の生成に、違和感を感じなかった	4	4
	オノマトペを用いた魔法の制御に満足している	3	3.14
没入感	炎や雷などの魔法を使用している感覚を強く感じた	4	3.57
	ゲームプレイ中、魔法の世界に没入している感覚があった	3	3.57
	魔法を使用している実感が湧いた	4	3.57
体験の楽しさ	このゲームは、魔法を使うという新しい体験を提供していると感じた	5	4.71
	オノマトペによる魔法の制御は、楽しく、エンターテイメント性が高いと感じた	4	4.42
	オノマトペを用いた魔法の制御は、面白みがあると感じた	4	4.28
	このゲームシステムを友人に勧めたいと思う	4	4.28
	このゲームシステムをまた遊んでみたいと思う	5	4.71

参加者は没入感やゲームの楽しさに関して高く評価し、「オノマトペを唱えることで、魔法を使っている感覚を感じた」「イメージどおりの魔法が出た時は魔法を制御している感覚があった」といったポジティブなフィードバックを得た。

ただし、BCI を用いた全てのゲームが良い体験になるわけではない。市販ゲームを限定的な BCI 操作でプレイした先行研究では、BCI 制御の有無で楽しさや没入感に関して有意差はないことを示した [17]。これは、ゲームのアバターモードをリラックス状態とストレス状態により制御するという限定的な操作であり、自分で操作している感覚が少なかったため、BCI 制御がゲーム体験に影響を与えたかった可能性がある。

魔法ゲームが没入感やゲームの楽しさに関して高い評価を示したのは、魔法を生成、操作する体験と、BCI を用いたイメージによる操作の間に高い親和性があり、心地の良いユーザ体験を提供できたからだと考察する。このことから、イメージによる制御体験と親和性の高いゲームを選択することは、プレイヤー体験を最大化させる可能性がある。

5.2 視覚フィードバックが与える影響

タスク終了後のインタビューにおいて、「一度魔法が発動すると魔法の切り替えが難しかった」という意見を得た。この問題について、同じ魔法が連続して入力される傾向があった。これは、視覚的なフィードバックを受けながら視覚イメージを想起するという本システムの特性上、意図した魔法とは別の魔法が検知された際、視覚的に誘因されている可能性が高い。この問題は、運動イメージにはない視覚イメージ特有の問題である。運動イメージでは、身体の動きをイメージするため、視覚フィードバックのネガティブな影響を受けにくい [10]。一方、視覚イメージでは、視覚的なイメージを心の中で構築しなければならないため、視覚フィードバックの影響を受け

やすい可能性がある。

視覚イメージは、「知覚からではなく、記憶から由来する知覚情報の操作」と定義され、視覚刺激を観察する視覚観察タスクと、視覚刺激を想起する視覚イメージタスクに分類できる [7, 20]。視覚観察タスクと比較して、視覚イメージタスクでは BCI パフォーマンスが低い傾向があるため、本システムでは、視覚観察タスクによりトレーニングを行った。そのため、視覚的なフィードバックを顕著に受けた可能性がある。

この問題を解決するためには、視覚イメージタスクと視覚観察タスク時の脳波データから各々分類器を開発し、イメージタスクと観察タスクの特性を組み合わせたアンサンブル分離器を開発することが有効であるかもしれない。魔法の発動前と発動中に別の分類器による入力を受け取ることにより、視覚フィードバックの影響を軽減できる可能性がある。今後の研究では、視覚イメージを用いた BCI 制御における視覚フィードバックの影響を検証していく必要がある。

5.3 展望

今回のユーザスタディでは、参加者が意図通りの魔法を使った割合は 60% 程度にとどまった。しかし、試行を繰り返すことにより精度が向上する傾向があったことから、長時間プレイすることによる精度向上が期待できる。今後は、トレーニング時間やタスク時間を伸ばすことにより、どの程度制御パフォーマンスが向上するか検証する。また、本研究では条件比較を行っていないため、本手法が BCI 制御によるゲーム体験において、制御パフォーマンスとゲーム体験の向上に与える影響を直接的に確認することはできなかった。オノマトペによる発話イメージの有無による BCI 制御の比較や、他の脳波制御手法による魔法制御との比較を行うことで、本手法の効果を定量的に評価する必要がある。

参考文献

- [1] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K.-R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *Neuroimage*, 51(4):1303–1309, 2010.
- [2] B. Blankertz, R. Tomioka, S. Lemm, M. Kawabae, and K.-R. Muller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal processing magazine*, 25(1):41–56, 2007.
- [3] M. A. Cervera, S. R. Soekadar, J. Ushiba, J. d. R. Millán, M. Liu, N. Birbaumer, and G. Garipelli. Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. *Annals of clinical and translational neurology*, 5(5):651–663, 2018.
- [4] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday. Brain-computer interfaces for communication and rehabilitation. *Nature Reviews Neurology*, 12(9):513–525, 2016.
- [5] I. Guler and E. D. Ubeyli. Multiclass support vector machines for EEG-signals classification. *IEEE transactions on information technology in biomedicine*, 11(2):117–126, 2007.
- [6] L. Hinton, J. Nichols, and J. J. Ohala. *Sound symbolism*. Cambridge University Press, 2006.
- [7] N. Kosmyna, J. T. Lindgren, and A. Lécuyer. Attending to visual stimuli versus performing visual imagery as a control strategy for EEG-based brain-computer interfaces. *Scientific reports*, 8(1):13222, 2018.
- [8] S.-H. Lee, M. Lee, J.-H. Jeong, and S.-W. Lee. Towards an EEG-based intuitive BCI communication system using imagined speech and visual imagery. In *2019 IEEE international conference on systems, man and cybernetics (SMC)*, pp. 4409–4414. IEEE, 2019.
- [9] D. J. McFarland and J. R. Wolpaw. EEG-based brain-computer interfaces. *current opinion in Biomedical Engineering*, 4:194–200, 2017.
- [10] M. Mousavi, A. S. Koerner, Q. Zhang, E. Noh, and V. R. de Sa. Improving motor imagery BCI with user response to feedback. *Brain-Computer Interfaces*, 4(1-2):74–86, 2017.
- [11] C. S. Nam, A. Nijholt, and F. Lotte. *Brain-computer interfaces handbook: technological and theoretical advances*. CRC Press, 2018.
- [12] Y. Park and W. Chung. Optimal channel selection using correlation coefficient for CSP based EEG classification. *IEEE Access*, 8:111514–111521, 2020.
- [13] G. Pfurtscheller, G. R. Müller-Putz, R. Scherer, and C. Neuper. Rehabilitation with brain-computer interface systems. *Computer*, 41(10):58–65, 2008.
- [14] D. Plass-Oude Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, and D. Heylen. Brain-computer interfacing and games. *Brain-computer interfaces: applying our minds to human-computer interaction*, pp. 149–178, 2010.
- [15] S.-D. Rosca and M. Leba. Design of a brain-controlled video game based on a BCI system. In *MATEC Web of Conferences*, Vol. 290, p. 01019. EDP Sciences, 2019.
- [16] S. K. Swee, K. D. T. Kiang, and L. Z. You. EEG controlled wheelchair. In *MATEC web of conferences*, Vol. 51, p. 02011. EDP Sciences, 2016.
- [17] B. van de Laar, H. Gürkök, D. P.-O. Bos, M. Poel, and A. Nijholt. Experiencing BCI control in a popular computer game. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):176–184, 2013.
- [18] A. Vourvopoulos, S. Bermudez i Badia, and F. Liarokapis. EEG correlates of video game experience and user profile in motor-imagery-based brain-computer interaction. *The Visual Computer*, 33:533–546, 2017.
- [19] M. P. Woźniak, P. Sikorski, M. Wróbel-Lachowska, N. Bartłomiejczyk, J. Dominiak, K. Grudzień, and A. Romanowski. Enhancing in-game immersion using BCI-controlled mechanics. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–6, 2021.
- [20] Z. Zhou, A. Gong, Q. Qian, L. Su, L. Zhao, and Y. Fu. A novel strategy for driving car brain-computer interfaces: Discrimination of EEG-based visual-motor imagery. *Translational neuroscience*, 12(1):482–493, 2021.
- [21] 小野正弘. 日本語オノマトペ辞典: 擬音語・擬態語 4500. (*No Title*), 2007.
- [22] 平野怜旺, 岸駿斗, 和栗真花, 山田紗妃, 渡邊恵太. BCIxD: ブレインコンピュータインターラクションデザイン. In *WISS 2023 Proceedings*, 2023.
- [23] 平野怜旺, 渡邊恵太. メンタルイメージベースのBCIにおけるオノマトペを用いた脳波制御訓練手法の提案. エンタテインメントコンピューティングシンポジウム 2023 論文集, pp. 220–227, 2023.

参加者との協同を実現する言葉遊びアンケートシステム

西出 新也^{*†} 西岡 裕子[‡] 福田 孝輔[‡] 米沢 勝[‡] 新井 美音[‡] 西田 健志^{*}

概要. 本研究では、アンケート調査における回答率と回答品質の向上を目的とした新たなアプローチとして、回答者が出題者に役割転換する共同運営型アンケートシステムを提案する。従来のアンケートが抱える一方向的な構造を転換し、参加者の主体的な関与を促す設計を採用したプロトタイプの開発を行なった。提案システムの主要な特徴として、次の2点を導入した。(1)回答者が出題者として質問を作成できる機能。(2)言葉遊びを取り入れた回答形式。これらの特徴により、参加者のモチベーション向上と社会的望ましさバイアスの低減を図った。提案システムの試用結果から、従来のアンケートと比較して、回答内容から具体的な改善点の抽出には課題が残ったが、同時に、これまで得られにくかった率直な意見や感情を引き出していることが観察された。言葉遊びを用いた回答形式により、参加者の付度や誇張につながる回答を抑制し、本音に近い情報を収集できることがわかった。

1 はじめに

多くのイベントにおいて、参加者に対するアンケート調査は、実施後の評価を行い、成功点や改善点を抽出するための重要な手段として広く用いられている。アンケートの結果を適切に分析することで、次回以降の活動に対する精度の高い改善策が導き出されることが期待される。

アンケート調査の効果を最大限に引き出すためには回答率と回答内容の質が重要であるが[1]、参加者のモチベーション低下による非回答や不真面回答の増加[2]や、出題者と回答者の関係性が引き起こす付度や誇張・過小した回答、いわゆる「社会的望ましさバイアス」[3]が、回答データの信頼性や有用性に影響することが指摘されている。

アンケートの回答率や回答品質を向上させようとする試みや研究は数多くなされているが[4, 5, 6]、いずれも回答者の負荷やバイアスの軽減に注目するものであり、「出題者と回答者」という一方的な構造によって生じてしまう回答者の受け身姿勢は変わらないままである。

本研究では、この一方的な構造を転換することで、回答率と回答品質の向上を目指す手法を提案する。提案手法は、アンケートの自由記述回答に代えて言葉遊びゲームを導入し、従来は回答するのみであった参加者側もお題を出すことができるというものである。提案手法で実施するアンケートは従来のアンケートと比べて、参加者が主体的に関わることがで

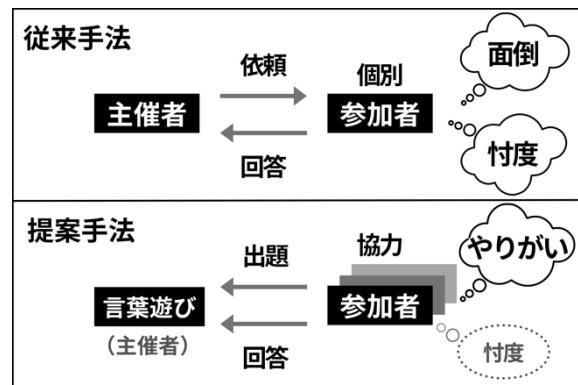


図 1. 参加者との協同を実現する言葉遊びアンケート

きるため、内発的動機づけにつながるとともに言葉遊びゲームの回答を考えるという過程を経ることによって主催者に付度を利かせる余裕が失われ、より率直な回答が増えることを期待している(図 1)。

本研究では、提案手法の可能性を探るためにアンケートシステムのプロトタイプを実装し、新入社員に対する社員研修後に行うアンケート場面で運用を行った。並行して行った従来型のアンケートで得られた回答と比較して、言葉遊びによる回答の自由度を制限したことでの具体的なフィードバックの収集という観点では課題が残ることがわかったが、同時に、参加者の付度や誇張につながる回答を抑制し、本音に近い情報を収集できることができた。

2 関連研究

より良いアンケート結果を得るための研究としては大きく分けて、回答率に関する研究、回答品質に関する研究が行われている。

Copyright is held by the author(s).

* 神戸大学

† 関西学院千里国際中等部・高等部

‡ 三菱総研DCS株式会社

2.1 アンケートの回答率に関する研究

従来のアンケート調査では、回答者は一方的に出題される質問に対して回答する形式が一般的である。この形式では、回答者がアンケートに対する主体的な関与やメリットを感じにくく、これが回答率の低下につながる主な要因の一つとして指摘されている[7]。さらに、回答者がアンケートに回答することの意義や価値を十分に認識できることや匿名回答形式である場合、非回答や不真面目回答の増加が見られる[2, 8]。

回答率の向上を目指すアプローチとして、回答者の内発的動機を活性化させるための様々な手法が提案されている。その一つが「社会的交換理論」を基盤としたアプローチである[9]。これは、アンケートにおける回答行動が、回答者が受け取る社会的報酬と提供する労力や時間のバランスに依存するという考え方に基づいている。具体的には、金銭的な報酬や労働の対価だけではなく、挨拶や感謝の意などの社会的行為も報酬として捉えることで、回答者の動機づけを高め、結果的に回答率の向上につながることが示されている[4]。

しかし、挨拶や感謝といった出題者との関係を強調する手法は回答率の向上にはつながりうるもの、出題者を意識することによる付度や誇張につながり、回答内容の信頼性に悪い影響を与える可能性が否めない。これら多くの研究では、回答者と出題者の関係性が回答率に与える影響については、主に回答者が相対的に強い立場であり、出題者が「お願いする」立場である場合を前提に検討がなされているが[10, 11]、出題者と回答者の関係性には様々な状態を考えうる。本研究においては、参加者が出題側にもなることができることによって、多様な主催者・参加者関係において、付度を過度に引き起こさずに回答を集めることができるアンケートを目指す。

2.2 アンケートの回答品質に関する研究

回答品質の向上を目指した研究では、回答者の負荷やバイアスを軽減するための様々なアプローチが行われてきた。不真面目な回答を分類し、分析時にこれらの回答を除外することで、データの精度と信頼性を保つことを目指した研究では、データの一貫性を保ちながら、回答の品質を高める手法を提案している[5, 6]。

アンケートの設計段階で、質問内容や回答形式といったインターフェースの工夫を試みた研究も存在する。例えば、自由記述設問の順番やテキストボックスサイズにより、回答品質に影響がでることを実証した研究では、回答者に対する負担を軽減することに注目している[12, 13, 14]。

従来の研究においては不真面目な回答を除外する

ことや、回答の負荷を軽減することでそもそも不真面目な回答を減らすことを検討しているが、回答の質は真面目かどうかのみで決まるものではなく、イベントに関連する人間関係に由来する付度なども含まれる。「社会的望ましさバイアス」に関する研究では、回答者は他者の期待に沿った望ましい回答を行う傾向があるため、このバイアスがデータの信頼性を損なう一因となることを指摘している[15]。

これに対して本研究の提案手法では、言葉遊びを取り入れ、回答者の参加意欲を高めることと、回答作成への集中を促す仕掛けをしている。これにより、従来のアンケートが抱える参加へのモチベーション低下や、バイアスによる影響を最小限に抑えることができ、結果として回答品質の向上につながると期待する。

3 共同運営型言葉遊びアンケートシステム

本研究では、回答者が出題者に役割転換する共同運営型アンケートシステムの設計およびプロトタイプの開発を行った。本章では、システムの基本コンセプト、デザイン指針、インターフェースの主用部分について説明する。

3.1 基本コンセプト

提案システムでは、回答率と回答品質それぞれの向上を目指し、大きく2つのコンセプトを設定する。それに伴った期待される効果を次に示す。

コンセプト1：出題者と回答者の方的な関係性を変える

従来のアンケートでは、主催者が出題し、参加者がそれに答えるという一方向の関係が基本である。本研究では、この構造を転換し、回答者が出題者に入れ替わることで、主催者の存在を薄める。そのことにより、双方に水平な関係を形成するねらいがある。その結果、付度や誇張した回答内容が減少する効果が期待できる。

コンセプト2：「言葉遊び」によるモチベーションの向上

回答時に言葉遊びのルールを導入することで、参加者のモチベーションを向上させることと、回答作成への集中を促す。その結果、社会的望ましさバイアスの回答は減少し、参加者の感情や本音を引き出すことができる効果が期待できる。

言葉遊びの一例として、1970年代にニューヨークで始まったラップが挙げられる。ラップの起源には、シグニファイリング・モンキーという神話があり、言語ゲームによって社会構造を逆転させるという考え方がある[16]。この言語ゲームが現代のラップバトルの原型とも言える。絵本作家のDr.Seussが

参加者との協同を実現する言葉遊びアンケートシステム

残した、「制約は創造の母」という言葉に代表されるように、制限があるからこそ、表現に向き合い工夫する想像が生まれ、ラップバトルのように、本音をうまく言ったものは聴衆の共感を得て称賛される。これをアンケートに応用することで、回答者が自由記述を作品として捉え、また、その作品づくりに集中することで、感情や本音を外化できることが期待できる。

回答に、言葉遊びによる制約を設けることで、上手にできなくても恥ずかしくない、という心理的安全性[17]を確保し、全員が安心して参加できる環境を作り出すことを目指す。最終的には、アンケート終了後に回答者が、やってよかったと思える満足感を得られる効果が期待できる。

3.2 デザイン指針

システムの実装にあたり、次の 2 つのデザイン指針を設けた。

指針 1：主催者の介入を抑えた進行と誘導

参加者が主体的にアンケートに取り組み、進めることができるようなシステムとするため、主催者の介入をできるだけ抑え、参加者同士が協力して進行していることを強調するデザインとする。

指針 2：貢献や協力による一体感を生み出す

参加者間で貢献や協力が発生し、一体感が生まれるシステムとするため、ユーザがお互いの状況を把握しやすくなることによって協力や分担を可能にする。さらに参加者全体で協力して達成すべき目標を提示し、貢献度の偏りができるだけ少なくなるように配慮する。

3.3 プロトタイプアプリケーション

本研究で開発したプロトタイプは、イベントの参加者が協力しながらそのイベントを表現する「あいうえお作文」を完成させていくことができる Web アプリケーションである。

システムを利用する流れは「個人ミッション」と「集団ミッション」の 2 段階に分かれている(図 2)。最初に取り組む個人ミッションにおいては、指定のひらがながら始まるお題に対する回答文を出し合い、それに対してお互いに投票しあう。お題は開始前に主催者によっていくつか準備されるが、参加者たちが増やすこともできるようになっている。

集団ミッションにおいては、得票数の多かった回答文を集合させることによって「あいうえお作文」を完成させる。「あいうえお作文」の候補を考えて提案するユーザは、個人ミッションにおいて貢献度の低いユーザの中からシステムが選出するようになっている。他のユーザはヒントを投稿して選出されたユーザを支援することができる。これは全体の貢献

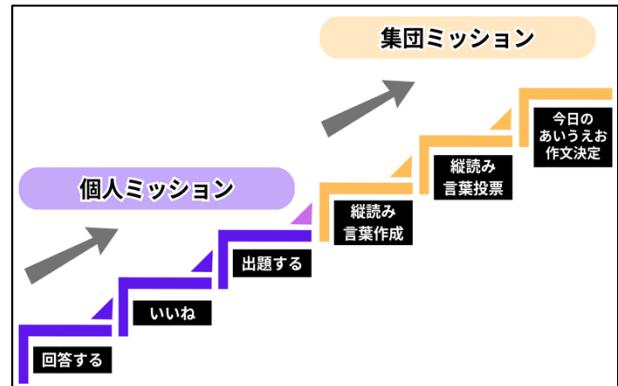


図 2. 「個人ミッション」と「集団ミッション」の流れ図

や一体感を引き出すことを期待した設計である（デザイン指針 2）。

最終的に提案の中から最多の得票を集めたものがそのイベントの「あいうえお作文」として決定される。これは参加者全員が協力して最終的に達成すべき目標となる（デザイン指針 2）。

3.4 ユーザインターフェース

プロトタイプのユーザインターフェースの詳細について個人ミッションと集団ミッションに分けて順に説明する。

3.4.1 個人ミッション

図 3 はアンケート参加者がアプリケーションで最初に目にするトップ画面である。ここでアンケート参加者は、個人ミッションとして、画面の左に示されたナビゲーションに従い、3 つのミッションを順にクリアするように案内される。3 つのミッションは「回答を 1 回する」「いいねを 1 回する」「3 つのカテゴリで、回答 or いいね or 出題をする」である。個人ミッションで示される行動の回数は、チュートリアルとして最低限こなしてほしいものであり、達成以降も好きなだけ行うことができるようになっている。ミッションを順にこなしていくことで、このアプリケーションで何をしていくべきかを把握できるようになっている（デザイン指針 1）。



図 3. トップ画面

トップ画面の右側には質問のカテゴリ毎に回答数があわせて花がだんだん成長していく鉢が表示されており、出題と回答の全体での進行状況を大まかに把握することができるようになっている。鉢の周りにはひらがなの文字がふわふわするアニメーションで表示される。この文字は、その頭文字から始まる回答が多いほど大きく表示される。参加者はこの表示によって他ユーザの存在を感じることができ、これを手掛けたりとして盛り上がっている質問や回答が不足しているカテゴリなどを発見し、他のユーザと協調した行動を取ることができる（デザイン指針 2）。

回答は「みんなからの質問画面（図 4）」から質問を選択し、質問で指定された頭文字から始まる回答を入力する。それぞれの質問にはそれを作成したユーザが表示されるようになっており、参加者も出題側になることができることが感じられるようになっている。運営による強制力を感じさせ過ぎない範囲内でほどよく出題傾向を統制する仕組みとして、質問を分類するカテゴリは運営側があらかじめ準備することができるようになっている（デザイン指針 1）。



図 4. みんなからの質問画面

一度でも回答を行って最初のミッションを達成すると、「みんなの回答画面（図 5）」へ誘導され、2つ目のミッション「いいねを1回する」を達成するよう促される。この画面では他ユーザによる回答をカテゴリ毎に参照することができるようになっており、共感できる回答があった場合、ハートマークをクリックすることで「いいね」することができます。

いいねが多く集まった回答が表れると「ステキな回答」として全ユーザに通知が表示され、画面上部にその数が表示される。これにより、いいねが集まった回答を増やすように参加者が自然に誘導される（デザイン指針 1）。

2つ目のミッションを達成することで次は出題機能を利用することができるようになる。出題画面では、質問カテゴリのから一つ選択をし、それにマッ



図 5. みんなの回答画面

チした内容の出題を行う（図 6）。なお、「あいうえお作文」形式で出題するため、お題となる最初の一文字を五十音のひらがな一覧から選択をする。例えば、「『お』から始まる、『Java』について、短い説明をお願いします！」という内容で質問を作成し、出題するボタンで投稿する。

回答者は、出題者が自分と同じアンケートシステムの参加者であることを理解している。そのため、コンセプト 1 にあるように、出題者と回答者の関係性が引き起こす、社会的望ましさバイアスの回答が減少し、参加者の感情や本音を引き出せる効果が期待できる。



図 6. 出題画面

3.4.2 集団ミッション

個人ミッションは、タイマーもしくは主催者による手動操作で終了し、集団ミッションのステージに移行する。集団ミッションでは、「クリエイター」と「オーディエンス」に分かれて、縦読み言葉を作成する。縦読み言葉は、個人ミッションで集まった回答に使用された頭文字を3~5文字選んで作成する。クリエイターは、個人ミッションで回答数が最も少

参加者との協同を実現する言葉遊びアンケートシステム

なかつた 3 名をシステムで選出し、該当する参加者の画面に任命されたことを提示する。ただし、回答数が少ないことが理由で選出されていることはアナウンスされない。回答数が少ない者から選出する理由は、デザイン指針 2 にあるように、参加者全員が何かしら「取り組むことができた」、「貢献できた」と感じてもらうことのねらいである。

図 7 に縦読み言葉作成画面を示す。50 音表では回答の存在するひらがな文字が強調され、クリックすることでその頭文字から始まる人気の回答ベスト 3 を確認することができる。クリエイターはこれらの表示を確認しながら、縦読み言葉の候補を考えて登録する。その裏でオーディエンスも同じように縦読み言葉を考えて、クリエイターへのヒントとして送信することができるようになっている。



図 7. 縦読み言葉作成画面

3 人のクリエイターによる、縦読み言葉の作成が完了した後の最終ミッションとして、オーディエンスを含めた参加者全員は、クリエイターの作品を一つ選び、投票する(図 8)。最終結果として、最も投票数が多かった作品が、「今日のあいうえお作文」として、全体に公開される(図 9)。



図 8. 投票画面



図 9. 今日の「あいうえお作文」画面

4 プロトタイプの試用評価

提案システムが回答率を向上させるか、また従来のアンケートでは得られなかった新たな意見を引き出すことができるかを検証するために試用を行った。被験者は三菱総研 DCS 株式会社の新入社員、約 100 名を対象とし、実験は、新入社員研修の期間に、研修の一環として以下の 3 つのタイミングで行った。

1. 入社後 1 ヶ月 (ビジネスマナー研修)
2. 入社後 2 ヶ月 (Java 研修)
3. 入社後 3 ヶ月 (Web アプリケーション研修)

各研修の最後に、約 20 分間を使って提案システムでのアンケート調査を実施した。従来から研修後に行ってきた通常のアンケートも並行して実施され、両者の回答を比較するとともに、従来型のアンケートを実施し、アンケート結果を利用するこことなる研修の内容に関与した社員、計 4 名に協力を依頼し、回答の生データおよび簡単な集計・分析を行ったデータを見ながら聞き取り調査を行った。

なお、今回の参加者は研修および、その後の 2 種類のアンケートには業務として参加しているため、回答を行う強い強制力がある状況であった。そのため、回答数についてはいずれの方法においても想定通り非常に高いものであったため、今回の分析では注目しなかった。同様の理由によって回答に忖度が働く可能性は想定されるため、その点に重点を置いて聞き取りを行った。

4.1 結果

提案システムの実験結果から得られた回答内容は、従来のアンケートと比較して回答の文字数が少ない傾向が見られた。従来のアンケートでは、回答内容が特定の事象について焦点を当てたものや、具体的な言葉で構成されていることが多い、「～がわかりやすかった」などのポジティブなフィードバックが多く見られた。

一方、提案システムでは、「つらい」や「しんどい」

など、回答者の感情を吐露するような短い回答が目立った。これは、言葉遊びを取り入れた仕組みにより長い回答を作ることが難しくなったことに加え、回答者が忖度や建前を忘れて作品づくりに集中していたことで、より率直な感情を表出することを促された結果と考えられる。

研修の内に関与した人へのインタビューからは、提案システムの回答内容を見た結果、おそらく参加者がそう感じているだろうと危惧していたが、従来のアンケートからは見えていなかったことが表れていることが確認された。特に、これまでのアンケートでは得られなかつた講義テキストに対する不満など、新たな視点や発見が得られた点は、従来のアンケート手法では見落とされていた参加者の本音に近い情報を引き出すことができたことを示している。

提案システムの運営を現場でモニターしていた、人事担当者のインタビューでは、回答内容について、被験者が自分自身に向けた「つぶやき」のような表現が多く見られるため、具体的な改善点を抽出するための情報が不足しているとの指摘があった。これは、提案システムが従来のアンケートとは異なり、言葉遊びによる回答の自由度を制限したことにより、具体的な改善点やフィードバックの収集という観点では課題が残ることを示唆している。一方、回答者が出題者になることで、「あえてのダメ出し」という出題トピックが出るなど、従来のアンケートには存在しなかつた聞き方が生まれ、主催者側の新たな気づきにつながったことがわかった。

回答データだけを見た社員インタビューでは、提案システムの実施回数を重ねるごとに、回答内容にも出題の趣旨とは異なる不真面目な回答が増加しており、被験者に飽きが生まれている様子が指摘された。一方、新入社員同士が研修時にお互いに教えるなど、協力してコミュニケーションを取っている様子がみられたことは、従来のアンケートでは拾いにくかつた点で有益な回答であることにも言及している。

これらの結果から、提案システムでは、回答者がつぶやきのような形式で自由に感情を表現することと、具体的なフィードバックを収集することのバランスをどのように取るか、また実施回数が重要な課題であることがわかった。同時に、回答者の感情や本音を引き出すという点で、回答品質の新たな側面を捉えていることが評価できた。

5 まとめ

本研究では、回答者が出題者に転換できるアンケートシステムの提案をし、従来の一方向的なアンケートの構造を変え、参加者が主体的に関与できるデ

ザインを採用することで、回答率と回答品質を向上させることを目指した。具体的には、回答者が出題者として質問を作成できる機能や、言葉遊びを取り入れた回答形式を導入したプロトタイプを開発した。これにより、参加者のモチベーション向上とバイアスの低減を図った。

試用の結果、提案システムは、従来アンケートの回答には見られなかった意見を引き出すことに寄与していることが確認された。回答者が出題者になることで、互いの関係性を水平に保ち、忖度や誇張した回答を減少させることにつながることがわかった。また、言葉遊びの要素で回答作成への集中を導き、社会的望ましさバイアスの低減効果が見られた。

しかし、本研究の実験では特定の集団でのみ実施されているため、他の集団や状況への適用にはさらなる検証が必要である。また、長期的な運用で参加者がアンケートに飽きる可能性があり、対策への考察が不足している。さらに、言葉遊びが全ての参加者に魅力的とは限らず、個々の反応を詳しく分析する必要がある。

今回は回答を行う強い強制力があり、また短時間での実施であった。そのため、強制力が弱く、長時間実施での場面で検証する必要を感じるため、WISS Challengeでの運用を希望している。従来のように、イベントの最後にアンケート時間を設けるのではなく、学会のオープニングで紹介をし、会期中を通じて回答を集めることで、異なる環境・参加者層での検証が実現できる。この機会を活用し、回答率や回答品質のより高いシステムの開発につなげることを目指したい。

謝辞

三菱総研DCS株式会社の新入社員の皆様、インタビューにご協力いただいた皆様に心より感謝申し上げます。

参考文献

- [1] Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. Open-ended vs. close-ended questions in web questionnaires. *Developments in applied statistics*, 19(1), pp. 159-177, 2003.
- [2] 山崎 郁未, 中村 聰史, 小松 孝徳. Web アンケートにおける自由記述設問の順番が回答時間と離脱に及ぼす影響のスマートフォン・PC 間比較. 情報処理学会, 25, pp. 1-8, 2023.
- [3] Holtgraves, T. Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin*, 30(2),

参加者との協同を実現する言葉遊びアンケートシステム

- pp. 161-172, 2004.
- [4] Dillman DA. *Mail and telephone surveys: the total design method*. A Wiley Inter-science Publications, 1978.
- [5] 後上 正樹, 松田 裕貴, 荒川 豊, 安本 慶一. オンラインアンケート回答時のスマートフォン画面操作状況に基づく不適切回答検出. 情報処理学会インタラクション, pp. 1-10, 2021.
- [6] 山崎 郁未, 畑中 健壱, 中村 聰史, 小松 孝徳. Web アンケートにおける不真面目回答削減に向けた回答分類とその検証. 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI) , 29, pp. 1-8, 2022.
- [7] 福井 賢一郎, 藤井 聰, 北村 隆一. 内発的動機に基づく協力行動:社会調査における報酬の功罪. 土木計画学研究・論文集, 19 (1), pp. 137-144, 2002.
- [8] Dickinson, D. L., & McEvoy, D. M. Further from the truth: The impact of moving from in-person to online settings on dishonest behavior. *Journal of Behavioral and Experimental Economics*, 90, 101649, 2021.
- [9] Adams, J. S. Inequity in social exchange. In *Advances in experimental social psychology*, 2, pp. 267-299, 1965.
- [10] Manfreda, K. L., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. Web surveys versus other survey modes: A meta-analysis comparing response rates. *International journal of market research*, 50(1), 79-104, 2008.
- [11] Kaplowitz, M. D., Hadlock, T. D., & Levine, R. A comparison of web and mail survey response rates. *Public opinion quarterly*, 68(1), pp. 94-101, 2004.
- [12] Toepoel, V., Das, M. and Van Soest, A. Design of Web Questionnaires: The Effects of the Number of Items per Screen. *Field Methods*, 21(2), pp. 200-213, 2009.
- [13] Mavletova, A. and Couper, M. P. Mobile Web Survey Design: Scrolling versus Paging, SMS versus E-mail Invitations. *Journal of Survey Statistics and Methodology*, 2(4), pp. 498-518, 2014.
- [14] Manfreda, K. L., Batagelj, Z. and Vehovar, V. Design of Web Survey Questionnaires: Three Basic Experiments. *Journal of Computer-Mediated Communication*, 7(3), JCMC731, 2002.
- [15] Koivula, A.; Räsänen, P.; Sarpila, O. Examining Social Desirability Bias in Online and Offline Surveys. In *Lecture Notes in Computer Science*, pp. 145-158, 2019.
- [16] 金澤 智. ヒップホップ・クロニクル：時代を証言するポピュラー文化. 水声社. 2020.
- [17] Edmondson, A. Psychological safety and learning behavior in work teams. *Administrative science quarterly*, 44(2), pp. 350-383, 1999.

未来ビジョン

言葉遊びは万葉集に見られる和歌の技巧から1300年に及ぶ歴史を経て発展してきた。本研究で扱った言葉遊びは、「古今和歌集」から見られる、和歌の五句それぞれの頭一文字を合わせて意味のある言葉にする技巧「折句」の系譜である。今日ではアクロスティックと呼ばれるこの折句は、小野小町とある人物の贈答歌に見られるように、時に暗号の通信文としての役割を担うこともあったという。また、提案システムにある縦読み言葉、いわゆるダブルアクロスティックには、平安時代にみられる「沓冠（くつかむり）」の和歌技巧が始まりでもある。

このように制約のある中で、どのように最大限に想いを相手に伝えるか、先人たちは長い時間をかけ、工夫を凝らし、言葉を作品として紡いできた歴史がある。しかし、現代では自由に気軽に発言ができるSNSを中心に、インターネットを通じて相手に伝える言葉は軽くなりすぎた。匿名文化であること、さらには、

生成AIの登場が拍車をかけ、いつでも自由に思いついた時に言葉を発信できることで、相手に想いを馳せ、共感することを怠ってしまう、そのようなコミュニケーションの難しさが今も至る所で生まれている。

自由と制約、実名と匿名、そのバランスを考慮することで、言葉の重みが増していくのではないかだろうか。相手を慮り、協力が生まれる、そんな言葉の重みを取り戻すことを、私たちの研究の未来ビジョンに掲げたい。



画像出典：「百人一首繪抄」「九」「小野小町」
(東京都立図書館)

InverseVis: 疎な領域へのサンプリング点生成による多次元データ中の「非存在」の可視化

伊藤 貴之^{*} 田上 湖都[†] 矢島 知子[‡] 李 国政[§]

概要. 多次元データの可視化手法は非常にアクティブな話題であり、長年にわたって多くの手法が報告されている。その中でも散布図や平行座標プロットといった幾何学的な手法は、多次元データを構成する各個体を点や線で描くことにより、データ中の密な部位や外れ値が存在する部位を表現する。一方で時として我々は、多次元空間中の「個体が存在しない領域」に着目したいことがある。この要求を解決するために本報告では、多次元空間中の疎な部位に多数のサンプリング点を生成し、その点を可視化することで、個体が存在しない領域の分布を表現する手法を提案する。本手法では多次元空間の生成したサンプリング点に次元削減を適用して散布図で表示する。マウス操作によって特定の点を指定すると、平行座標プロットにより指定された点の各次元の値を表示する。本報告では、含フッ素有機化合物の反応実験データと、学術成績と給与のデータでのケーススタディをもって、提案手法の有効性を検証する。

1 序論

計算機を用いた可視化は一般的に、データ中に特定の事象が「存在する」ことを表現している。例として図1に示すように、流体（空気や水）の可視化では流れが存在することを表現する。散布図による多次元データの可視化では点群の中にクラスタや外れ値が存在することを表現する。ネットワークの可視化では2点間に関係が存在することを表現する。これらの可視化によって我々は、データ中に特定の事象が存在することを理解することができる。

しかしこのような可視化からは、データ中の事象の「非存在」を発見することが難しい場合がある。計算機を用いた可視化は、特定の事象の存在を点や線などで描画する。可視化のユーザは一般的に、画面上の点や線が描画された箇所に着目する。言い換えば、点や線が描画されていない箇所にはユーザの注目が集まりにくい。この点に着目して我々は、逆転の発想として、データ中の「存在しない事象」の可視化手法を研究している。本報告はの中でも多次元データに関する手法を提案するものである。

日常社会の情報の多くは多次元データとして記述可能である。これらの多次元データの理解を深める目的で、多次元データの可視化手法が開発されてきた。多次元データの可視化手法は長年にわたって研究されており [10, 12, 22]、特に散布図行列や平行座標プロットに関する論文は非常に多く発表されている。これらの手法は多次元空間中のクラスタや外

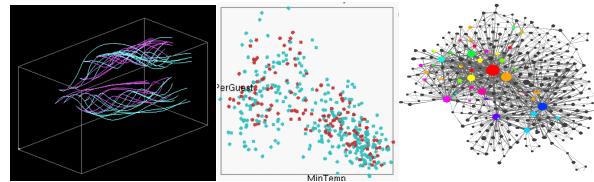


図 1. 計算機を用いた可視化は一般的に、データ中に特定の事象が「存在する」ことを表現している。(左) 流体の可視化。(中) 多次元データの可視化。(右) ネットワークの可視化。

れ値の「存在」を可視化するのに向いている。

一方で時として我々は、多次元空間中の「個体が存在しない領域」あるいは「個体の密度が疎である領域」に着目したいことがある。例えば、機械学習の訓練データにおいて、どのような数値を有するデータ標本が足りないかを知りたい状況が起こりえる。あるいは、反復的な科学実験や計算機シミュレーションにおいて、どのようなパラメータでの実験や計算が足りないかを知りたい状況が起こりえる。しかし、散布図行列や平行座標プロットから多次元空間中の疎な領域の視認は難しい。

そこで本研究では、逆転の発想として、多次元空間において個体の密度が低ければ低いほど多数のサンプリング点を発生することで、多次元空間中にて個体が存在しない（あるいは密度が疎な）部位を可視化する手法を提案する。図2に本研究の概念を示す。本手法では入力データを構成する多次元空間中の点群（図2(a)）の各々がポテンシャルを有すると仮定し、それを補間することで、多次元空間における個体の密度分布（図2(b)）を生成する。続いて、その密度分布を参照して、密度が低いほど高い確率

Copyright is held by the author(s).

* お茶の水女子大学大学院 理学専攻 情報科学領域

† お茶の水女子大学大学院 理学専攻 化学・生物化学領域

‡ お茶の水女子大学大学院 理学専攻 化学・生物化学領域

§ 北京理工大学

でサンプリング点を生成する(図2(c)). この処理により、個体の密度が疎であるほど密度の高いサンプリング点を生成する。

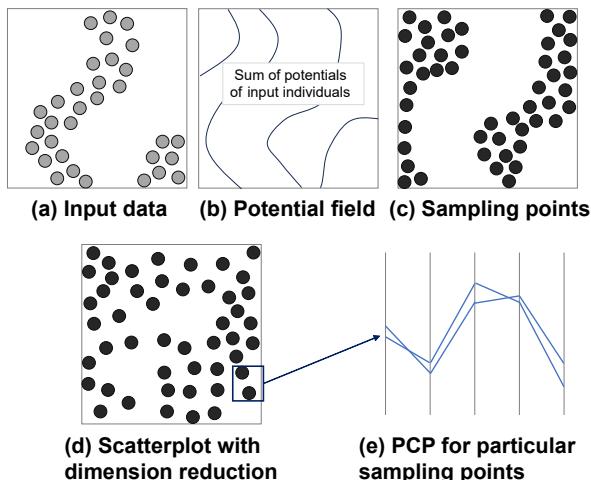


図 2. 本研究の概念。本研究では多次元空間中において個体の密度が低い部位にサンプリング点を発生させることで、個体が存在しない（あるいは疎な）部位を可視化する。

サンプリング点の集合を可視化する手段として、提案手法ではまずサンプリング点に次元削減を適用して、2次元の散布図として可視化する(図2(d)). そして散布図中の特定の点（または点の集合）をマウス操作で指定すると、それらのサンプリング点が有する各次元の値を平行座標プロットにより可視化する(図2(e)). この2段階の可視化により、まず多次元空間中の疎な領域の全体像を可視化し、続いてユーザが指定する特定のサンプリング点の各次元の値を確認できる。

本報告では、含フッ素有機化合物の反応実験データと、学術成績と給与のデータ[1]でのケーススタディをもって、提案手法の有効性を検証する。

本報告の貢献は以下のとおりである。

- 多次元空間中において個体の密度が低いほど多数のサンプリング点を発生することで、多次元空間中の疎な部分を可視化するというアプローチ。
- 多次元空間中の疎な部分における各次元の値を読み取るために、サンプリング点に次元削減を適用した散布図と、その散布図中の特定の点を表示する平行座標プロットの2つの可視化コンポーネントを組み合わせたインターフェクションデザイン。
- 含フッ素有機化合物の反応実験データ、学術成績と給与のデータ、を用いた2種類のケーススタディによる提案手法の有効性の検証と、ユーザ評価実験結果にもとづく改良案の議論。

2 関連研究

データ空間中の疎領域に着目した手法は科学技術系の可視化手法(Scientific Visualization)で多用されてきた。Ivanら[18]は空間中のSparsenessを各部位の重要度に置き換えて特徴的な部位を強調表示するボリュームレンダリング手法を提案している。また流れ場の可視化においても同様に、疎領域に注目することで可視化の視認性を高める議論がなされている[11]。これらの手法はあくまでもデータ中の特徴的な部位の存在を強調表示する手法であり、疎領域そのものの表示をしていない。

情報可視化の中でも多次元データを対象とした可視化手法において、疎な領域に存在するOutlierを逃さずに強調表示する手法がいくつか研究されている[3, 4, 5]。これらの手法は疎な領域に存在するOutlierの視認性を高めることを目的としており、疎な領域の全体像を可視化することや、あるいは点群や個体が全く存在しない領域を可視化することを目的としている。

次元削減と散布図を用いた多次元データ可視化が、多次元空間中の疎な領域の発見に貢献できる、ということを示した研究もいくつか報告されている[8, 23, 25]。また、次元削減によって2次元の画面空間に投影された多次元データを、再度3次元以上の空間に逆投影することで、多次元データの数値分布の表現性を高めた手法も報告されている[9, 20, 21]。しかしこれらの手法も、多次元空間中の疎な領域を直接的に描画するものではない。

本報告の提案手法で採用しているサンプリング点の生成と同様に、多次元空間中の密度の低い領域に積極的に点を生成するアルゴリズムはいくつかの研究で採用されている[2, 6]。しかしこれらの研究は密度の低い領域の可視化を目的としておらず、本論文とは目標設定が異なる。

本報告で1つ目の適用事例として紹介する化学反応の条件設定の探索問題には、実験計画法などの数理的な方法がしばしば適用される。可視化によって実験計画法を支援する手法はいくつか報告されている[7, 16, 24]。本報告で2つ目の適用事例では学術成績と給与のデータにおける分布の男女差を可視化している。このような性差の問題を可視化によって分析した事例もいくつか報告されている[13, 17, 19]。これらの用途に適用できる可視化手法を「疎な領域の可視化」という抽象化によって実現した、という点で提案手法はこれらの先行研究と異なる。

3 多次元空間中の疎な領域の可視化

本章では提案手法の処理手順を示す。具体的には、想定するデータ構造、サンプリング点の発生方法、可視化手法とインタラクションについて述べる。

3.1 データ構造

提案手法では入力情報として N 個の個体 d_i で構成される多次元データ $R = \{r_1, r_2, \dots, r_N\}$ を想定する。個体 r_i は M 次元の変数 $r_i = \{x_{i1}, x_{i2}, \dots, x_{iM}\}$ を有するものとする。またオプションとして、個体 r_i は目的変数 f_i またはカテゴリ変数 c_i をもちうるものとする。

以上を入力情報として、提案手法はサンプリング点 $S = \{s_1, s_2, \dots\}$ を生成する。サンプリング点 s_i は M 次元の変数 $s_i = \{y_{i1}, y_{i2}, \dots, y_{iM}\}$ を有するものとする。またオプションとして、目的変数 g_i またはカテゴリ変数 d_i をもちうるものとする。

3.2 サンプリング点の生成

提案手法では入力データとして与えられた点群からポテンシャル場を生成し、それに沿ってサンプリング点を生成する。ここで、単に点群の密度分布を近似するポテンシャル場を生成するのであれば、例えば混合ガウスモデル (Gaussian Mixture Model) を適用することが考えられるが、入力データ中の点の数が少ない場合や点の位置がまばらな場合、あるいは点の密度分布が非連続な場合にうまく動作しないことがある。そこで試行錯誤の結果として、我々は以下のようない手法を採用した。

提案手法では各々の個体がポテンシャル場を有するものとする。多次元空間中の位置 X における個体 r_i のポテンシャル場 $p_i(X)$ を以下のガウス関数

$$p_i(X) = a_{pot} \exp\left(-\frac{r_i - X}{b_{pot}^2}\right) \quad (1)$$

で表す。ここで r_i は多次元空間中の個体の位置、 a_{pot} および b_{pot} はユーザ定義の定数とする。これを用いることで、位置 X におけるポテンシャル $P(X)$ は、個体 r_i のポテンシャル場の総和 $P(X) = \sum_i p_i(X)$ により求められる。

続いて提案手法では、与えられた多次元空間中のランダムな位置に n_{pot} 個のサンプリング点を生成する。ただしサンプリング点が生成されるか否かの確率を、位置 X から以下の式 $c_{pot} - P(X)/P_{max}$ によって算出する。ただし P_{max} は $P(X)$ の最大値であるとし、 c_{pot} はユーザ定義の定数であり $0 \leq c_{pot} \leq 1$ であるとする。この値が 0 以下となる場合にはサンプリング点は生成されない。

図 3 に提案手法の仕組みを図解する。提案手法では各個体が有するポテンシャルの和によってポテンシャル場を生成し、ポテンシャルが低い位置ほど高い確率でサンプリング点を生成する。

サンプリング点がオプションで目的変数 g_i を有する場合には、サンプリング点の近傍の数個の個体 r_j の目的変数 f_j を補間することで g_i の値を算出する。サンプリング点がオプションでカテゴリ変数 d_i を有する場合には、そのカテゴリ変数がもちうる各々

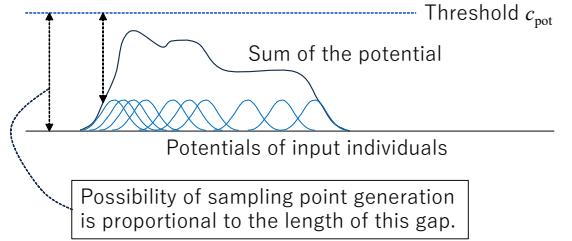


図 3. 入力データ中の個体によるポテンシャル場とサンプリング点の生成確率。

の値について別々にサンプリング点の生成を実行する。例えばカテゴリ変数が値 D_a, D_b, D_c の 3 種類の値をとりうる場合には、カテゴリ変数の値が D_a である個体のみ、 D_b である個体のみ、 D_c である個体のみを対象として、それぞれのカテゴリ変数値に対してポテンシャル場を生成し、そこからサンプリング点の生成確率を求めて、それぞれのカテゴリ変数値についてサンプリング点を生成する。

以上の処理により、入力データによって与えられた多次元空間にサンプリング点を生成する。この過程において提案手法では、 $a_{pot}, b_{pot}, c_{pot}, n_{pot}$ の 4 つの定数を調節する必要がある。

3.3 可視化とインタラクション

提案手法ではサンプリング点の多次元空間中の位置 $s_i = \{y_{i1}, y_{i2}, \dots, y_{iM}\}$ に次元削減を適用し、これを散布図で可視化する。現時点での実装では、線形性の高い分布の表現に向いた PCA (Principal component analysis) と、非線形性の高い分布の表現に向いた手法のうち計算時間が比較的短い UMAP (Uniform Manifold Approximation and Projection) を採用している。上述の $a_{pot}, b_{pot}, c_{pot}, n_{pot}$ の 4 つの定数を調節することで、散布図の視認性を調節することができる。なお、サンプリング点が目的変数 g_i またはカテゴリ変数 d_i を有する場合には、散布図を構成する各点の色をもって目的変数またはカテゴリ変数の値を表現する。

提案手法のスナップショットを図 4 に示す。画面左側の散布図にて、ユーザがマウスでホバリングすることで、特定の点を指定することができる。この操作によって特定の点（またはその近傍にある複数の点）を指定すると、画面右側の平行座標プロットを用いてその各点の各次元の値を表示する。これにより、与えられた多次元空間中における疎な領域での数値の例を知ることができる。なお、平行座標プロットの軸の並び順は、ユーザがデータ入力時に指定できる。あるいは最適な並び順の自動設定手法 [14] を適用してもよい。

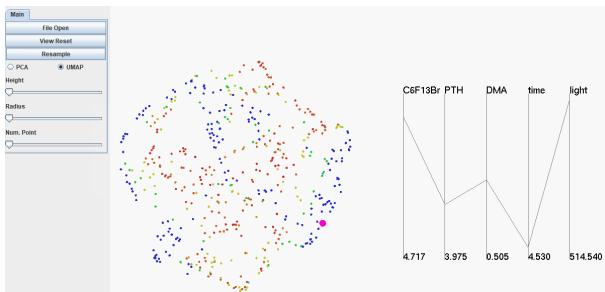


図 4. 提案手法のスナップショット. 左側の散布図で特定の点を指定すると、右側の平行座標プロットで各次元の値を表示する、というインタラクションにより疎な領域での数値の例を知ることができる.

4 適用事例

4.1 含フッ素有機化合物の実験データ

化学や生物などの実験科学では、条件設定を変えるながら反復的に実験を繰り返すことで、探索的に良好な実験結果を求めるプロセスが必要となることが多い。ここで探索過程となる実験の回数を減らすためには、まだ実施していない実験設定を可視化によりリストアップし、それを実験者が閲覧しながら自らの経験則と照合して、良好な実験結果が得られそうだと予想される実験設定を選ぶことが重要である。本論文ではこの目的に沿って提案手法を適用した事例を紹介する。

含フッ素有機化合物は医薬品や農薬、機能性材料として幅広く利用されており、その新たな合成法の開発は有機化学研究の重要な課題である。本研究では、ハロゲン結合を利用したペルフルオロアルキル化反応を、電子不足な共役オレフィン類、スチレン類及び非共役末端オレフィン類に適用した実験[15]を題材として、その合成におけるいくつかの説明変数と、目的変数となる収率の関係を可視化した。具体的には、実験時に設定される以下の4つの値を4次元の説明変数 $r_i = \{x_{i1}, \dots, x_{i4}\}$ とし、測定結果となる収率を目的変数 f_i とする19組の実験結果を入力データとした。

- ヨウ化ペルフルオロヘキシルの当量
- 酸素の当量
- N,N -ジイソプロピルエチルアミン (DIPEA) の当量
- 光照射時間

そして提案手法により「まだ実験されていない説明変数の例」となるサンプリング点の集合を可視化した。図5(上)にサンプリング点の集合を散布図として可視化した例を示す。ここで散布図の点は目的変数の値で色付けされており、収率が低い（好ましくない実験結果）と予測されるほど青に近い色で、収

率が高い（好ましい実験結果）と予測されるほど赤に近い色で表示されている。

図5(上)では(a)～(d)に収率が高いと予測される赤い点が、(e)(f)に収率が低いと予測される青い点が見られる。これらの点における説明変数の値を図5(下)に示す。このような表示により、収率が高いと予測される説明変数の値を発見し、その中からどの説明変数をもって次の実験を実施するかの議論材料にすることができる。

図5(下)にて(a)～(d)の共通点として、DIPEAの当量が1.0以上であること、光照射時間が長いことがあげられる。これらの条件を満たしつつ、他の2つの説明変数が特定の値を有するときに、実験結果としての収率の高さが期待できる。一方で、図5(下)にて(e),(f)の共通点として、DIPEAの当量が1.0未満であること、光照射時間が短いことがあげられる。これらの条件を満たす際に収率の低さが予測される。この傾向は入力データである19回の実験結果とも、またその後の知見とも整合するものであり、提案手法による可視化結果の妥当性を示すものである。

4.2 学校成績と給与のデータ

機械学習を用いた採用人事システムのバイアスが社会問題になったことがある。具体的には、採用人事のための訓練データに女性のデータが少なかったことから、システムが女性に不利な判定を下す傾向が見られたというものである。このようなバイアスを防ぐ手段として、データのバランスを事前に観察することが重要である。そこで本論文では、中等教育・高等教育の成績と給与オファーなどをまとめた人材オープンデータを題材として、データ中の男女間の偏りを可視化した事例を報告する。

このデータには以下の6つの値が記録されている。

- 中学校での成績
- 高校での成績
- 大学（学部）での成績
- Employment test での成績
- MBA での成績
- 給与

本事例ではこれらの値を6次元の説明変数 $r_i = \{x_{i1}, \dots, x_{i6}\}$ として扱う。さらに、このデータには各人物の性別が記載されており、これをカテゴリ変数 c_i として扱う。

そして提案手法により「データ中に存在しないタイプの人材の例」となるサンプリング点の集合を可視化した。図6(上)にサンプリング点の集合を散布図として可視化した例を示す。ここで散布図の点は性別で色分けされており、青は男性、赤は女性を示す。

図6(上)では(a)(b)に男性と女性の両方が集中しており、(c)～(f)に女性が集中している。これらの

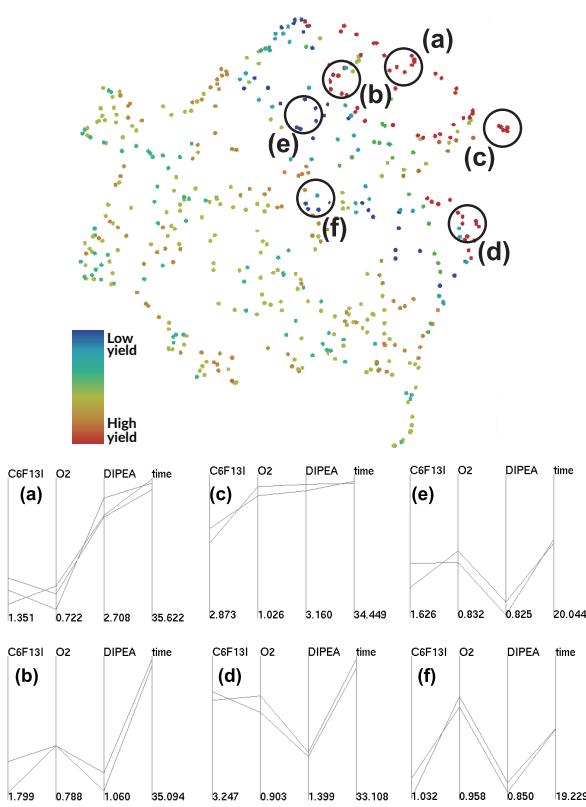


図 5. 含フッ素有機化合物の実験データを用いた実行例.
(上) まだ実験されていない説明変数の例となるサンプリング点の集合を散布図で可視化した例. 図中の (a)~(d) に収率が高いと予測される赤い点が, (e)(f) に収率が低いと予測される青い点が見られる. (下)(a)~(f) の点をマウス操作で指定して説明変数の値を平行座標プロットで表示した例.

点における説明変数の値を図6(下)に示す. このような表示により, 男性にも女性にも見られないタイプの人材の例, 女性(または男性)にのみ不足しているタイプの人材の例を理解することができる. 特に後者のタイプの人材については, 散布図中のそれぞれの領域について各変数の値を注意深く観察し, 男女間のバランスをとるためのデータの加工や是正を実施するか否かの意思決定を進めることができる.

5 ユーザ実験

4.1節に示したデータとは別の, 5次元の説明変数を有する含フッ素有機化合物の実験データを用いて, 高い収率をもたらす説明変数値のパターンを発見させるユーザ実験を実施した. 我々はあらかじめ, 自身で当該データを探索し, 高い収率をもたらす説明変数値の9種類のパターン(図7の(a)~(i))を発見した. この9種類の各々について, ユーザが画面左側のクリック操作によって発見できるかを検証

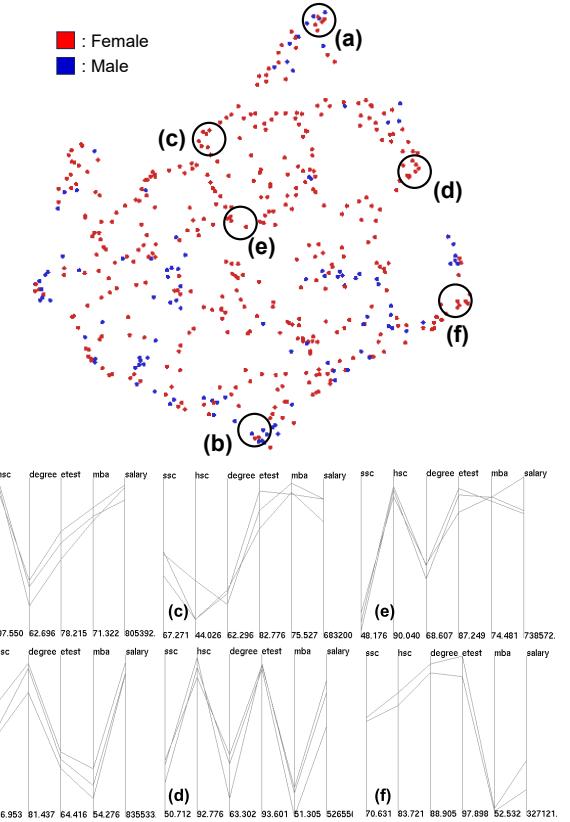


図 6. 学術成績と給与のデータを用いた実行例. (上)
データ中に含まれないタイプの人材の例となるサンプリング点の集合を散布図で可視化した例. 図中の (a)(b) には男性を示す青い点と女性を示す赤い点の両方が, (c)~(f) には女性を示す赤い点のみが, それぞれ多数見られる. (下)(a)~(f) の点をマウス操作で指定して説明変数の値を平行座標プロットで表示した例.

した. ただしユーザにはクリックの回数を最大10回に限定した. 実験参加者は計算機科学を専攻し可視化やデータサイエンスの授業を履修したことがある20~26歳の大学生・大学院生22名であった.

表1に9種類のパターンを発見できた実験参加者の割合を示す. 発見成功率は80%以上と70%以下に大きく二分される結果となった. 具体的には, 図5の(b)と(c)を片方だけ発見した人, (e)と(f)を片方だけ発見した人が多く, これらの発見成功率が低い結果となった. この傾向を改善する手段として, 散布図の配色が考えられる. 図5にも示されている通り, 現段階の実装では高い収率が予想されるサンプリング点は全て同様に赤色で表示されている.これを改めて, 高い収率が予想されるサンプリング点のみに高い彩度を割り当てた上で, 説明変数値のパターンによって異なる色相を割り当てる, というように配色を変えれば, 多様なパターンの説明変数値

表 1. 高い収率をもたらす説明変数値の 9 種類のパターンに対する提案手法での発見成功者の割合。

パターン 発見成功率(単位%)	(a) 90.9	(b) 59.1	(c) 68.2	(d) 86.4
(e) 54.5	68.2	95.5	86.4	90.9

を有するサンプリング点をユーザは積極的に選択できるようになると予想される。

6 制約と展望

本手法には現時点で以下の制約がある。 a_{pot} , b_{pot} , c_{pot} , n_{pot} の 4 つのパラメータによって全く異なる可視化結果を生じてしまうため、ユーザがその調節の手間を要するという難点がある。そこで好ましいパラメータ値の自動設定方法などを検討したい。また実装上の問題として、説明変数の数が増えることで計算時間やメモリ使用量が急増する可能性があるため、これらを抑える実装上の工夫が必要である。またサンプリング点の座標値を単純に乱数で発生させると、毎回異なる可視化結果を生じてしまうため、これを避けるには擬似乱数を導入する必要がある。

現時点での実装はカーネル密度推定に類似した手法でサンプリング点を生成しているが、入力データの分布によっては混合ガウス分布などを採用したほうが品質と計算量の両面において有利な可能性もある。また本研究は「疎な領域に満遍なく」サンプリング点を生成して可視化することを目的としているが、目的に合致した領域のサンプリング点だけを可視化すればいいのであればベイズ最適化などの探索手法を用いたほうが効率的である。

本手法ではサンプリング点の目的変数の値を求める際に、近傍の数個の個体から局所的な補間によって算出している。一方で、目的変数の数値分布によっては、回帰分析を適用したほうが精度の高い形で目的変数を予測できる場合もある。これについては今後の課題として実装を進めたい。

提案手法が有效地に働くと考えられる適用事例として、以下の条件を満たすデータがあげられる。

- 多次元の各変数が連続値であるデータ。離散値をとる変数、順列変数、カテゴリ変数などは望ましくない。
- 色付けによってデータの分布を読み取れるデータ。分散の大きい目的変数や、目安として 10 種類以下の値をとるカテゴリ変数が該当する。

以上の条件を踏まえた上で、さらに多くの事例に提案手法を適用したい。本報告で紹介した有機化合物の実験以外にも、科学技術、ビジネス、社会現象に関する幅広い予測問題に本手法を適用したい。また、本報告で紹介した学術成績のデータ以外にも、機械

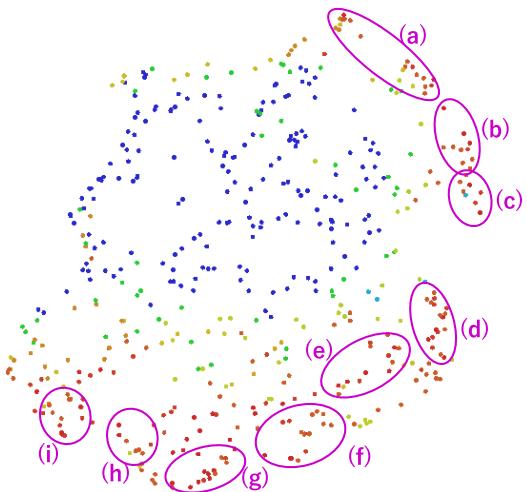
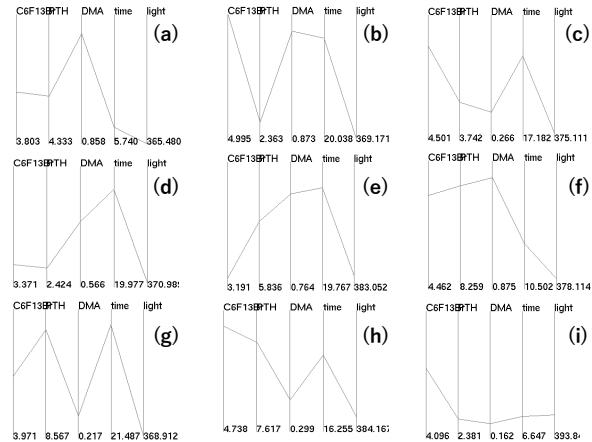


図 7. 高い収率をもたらす説明変数値の 9 種類のパターンとその散布図上での位置。

学習に用いる訓練データ全般にわたってそのデータ内のバランスを観察する目的で本手法を適用したい。さらに他の目的として、特徴量を算出可能なデジタル作品群（例えば音楽や絵画など）に本手法を適用し、データ内に存在しない特徴量を有する作品としてどのようなものがあげられるかを観察したい。

7 まとめ・今後の課題

本報告では、多次元空間中の疎な部位に多数のサンプリング点を生成し、その点を可視化することで、個体が存在しない領域の分布を表現する手法を提案した。また、含フッ素有機化合物の反応実験データと、学術成績と給与のデータでのケーススタディをもって、提案手法の有効性を検証した。

今後の課題として、前章で述べた制約を解消する改良手法開発、さらなる適用事例の開拓、より実用シナリオに近いユーザ評価実験に取り組みたい。

参考文献

- [1] Campus Recruitment Data, <https://www.kaggle.com/datasets/benroshan/factors-affecting-campus-placement>.
- [2] K. E. Bennin, J. Keung, P. Phannachitta, A. Monden, and S. Mensah. MAHAKIL: Diversity based oversampling approach to alleviate the class imbalance issue in software defect prediction. *IEEE Transactions on Software Engineering*, 44(6):534–550, 2017.
- [3] L. Blouvshtein and D. Cohen-Or. Outlier Detection for Robust Multi-Dimensional Scaling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(9):2273–2279, 2018.
- [4] S. Butscher, S. Hubenschmid, J. Müller, J. Fuchs, and H. Reiterer. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multi-dimensional Data. In *Proceedings of the CHI conference on human factors in computing systems*, pp. 1–12, 2018.
- [5] N. Cao, Y.-R. Lin, D. Gotz, and F. Du. Z-Glyph: Visualizing outliers in multivariate data. *Information Visualization*, 17(1):22–40, 2017.
- [6] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [7] F. Cheung. A Figure One Web Tool for Visualization of Experimental Designs. *Journal of Open Research Software*, 8(6), 2020.
- [8] E. P. dos S. Amorim, E. V. Brazil, J. Daniels, P. Joia, L. G. Nonato, and M. C. Sousa. LAMP: Exploring high-dimensional spacing through backward multidimensional projection. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 53–62, 2012.
- [9] M. Espadoto, G. Appleby, A. Suh, and D. Cashman. Unprojection: Leveraging Inverse-projections for Visual Analytics of High-dimensional Data. *IEEE Transactions on Visualization and Computer Graphics*, 29(2):1559–1572, 2021.
- [10] G. Grinstein, M. Trutschl, and U. Cvek. High-Dimensional Visualizations. In *Proceedings of the Visual Data Mining Workshop (KDD)*, Vol. 2, p. 120, 2001.
- [11] R. S. Laramee, H. Hauser, H. Doleisch, B. Vrolijk, F. H. Post, and D. Weiskopf. The State of the Art in Flow Visualization: Dense and Texture-Based Techniques. *Computer Graphics Forum*, 23(2):203–221, 2004.
- [12] R. Mazza. Introduction to Information Visualization. *Springer*, 2009.
- [13] Y. Nakai, T. Itoh, H. Takahashi, S. Nakashima, and T. Yamamoto. Hierarchical Data Visualization of Gender Difference: Application to Feeling of Temperature. In *27th International Conference Information Visualisation (IV2023)*, pp. 178–183, 2023.
- [14] W. Peng, M. O. Ward, and E. A. Rundensteiner. Clutter Reduction in Multi-Dimensional Data Visualization Using Dimension Reordering. In *IEEE Symposium on Information Visualization*, pp. 89–96, 2004.
- [15] K. Tagami and T. Yajima. Halogen-bond-promoted hydroxyperfluoroalkylation of olefins with molecular oxygen under visible-light irradiation. *Asian Journal of Organic Chemistry*, 12(8):e202300273, 2023.
- [16] F. Taverna, J. Goveia, T. K. Karakach, S. Khan, K. Rohlenova, L. Treps, A. Subramanian, L. Schoonjans, M. Dewerchin, G. Eelen, and P. Carmeliet. BIOMEX: an interactive workflow for (single cell) omics data interpretation and visualization. *Nucleic Acids Research*, 48(W1):W385–W394, 2020.
- [17] N. Tovanich, P. Dragicevic, and P. Isenberg. Gender in 30 Years of IEEE Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):497–507, 2022.
- [18] I. Viola, A. Kanitsar, and M. E. Groller. Importance-driven feature enhancement in volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):408–418, 2005.
- [19] E. Wall, A. Narechania, A. Coscia, J. Paden, and A. Endert. Left, Right, and Gender: Exploring Interaction Traces to Mitigate Human Biases. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):966–975, 2022.
- [20] Y. Wang, A. Machado, and A. Telea. Quantitative and Qualitative Comparison of Decision-Map Techniques for Explaining Classification Models. *Algorithms*, 16(9):1–26, 2023.
- [21] Y. Wang and A. Telea. Fundamental Limitations of Inverse Projections and Decision Maps. In *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2024.
- [22] P. C. Wong and R. D. Bergeron. 30 Years of Multidimensional Multivariate Visualization. In *Scientific Visualization, Overviews, Methodologies, and Techniques*, Vol. 2, pp. 3–33, 1994.
- [23] S. Wu and T. W. S. Chow. PRSOM: a new visualization method by hybridizing multidimensional scaling and self-organizing map. *IEEE Transactions on Neural Networks*, 16(6):1362–1380, 2005.
- [24] J. Xia, H. L. N, M. L. Mayer, O. M. Pena, and R. E. W. Hancock. INVEX—a web-based tool for integrative visualization of expression data. *Bioinformatics*, 29(24):3232–3234, 2013.
- [25] J. Zhao, M. Fan, and M. Feng. ChartSeer: Interactive Steering Exploratory Visual Analysis With Machine Intelligence. *IEEE Transactions on Visualization and Computer Graphics*, 28(3):1500–1513, 2022.

未来ビジョン

本研究の主な未来ビジョンは以下の 2 点である。

1 点目は「非存在の可視化」という概念を多次元データ以外の多様なデータに適用するという点である。本報告では多次元データに限定してその空間中の非存在の可視化を試みた。一方で、可視化が対象とするデータは多次元データの他にも、ネットワーク、時系列データ、地理データ、テキストデータなど多岐にわたる。あるいは、主に科学技術を対象とした物理空間でのデータもある。将来的なビジョンのひとつとして、これらの多様なデータを対象として可視化技術全般において「非存在の可視化」という概念を実現したい。あくまでも例として、毎年見られるはずなのに今年見られない現象を可視化する、ネットワークの中で連結されて然るべき関係を有するのにエッジで連結されていない部位を可視化する、といった技術の研究開発に臨みたい。

2 点目は「データ中の非存在の発見」という

工程を価値創出につなげるという点である。非存在の可視化によって「何が存在しないか」を知ることができれば、存在しないものを新しく生成してみようという動機付けが生まれる可能性がある。例えば、多次元の音楽特徴量空間に多数の楽曲を配置して、そこから非存在な音楽特徴量を導くことができれば、非存在な音楽特徴量を有する楽曲を生成 AI に作らせてみて、という試行錯誤的な創造につなげることができる。あるいは、多次元の味覚特徴量空間に多数の食品を配置して、そこから非存在な味覚特徴量を導くことができれば、非存在な味覚特徴量を有する架空の食品を生成して試しに味見してみることが可能になる。従来の「何が存在するか」を知る可視化と違って、「何が存在しないか」を知るための可視化を確立することで、これまで存在しなかった新しい作品や商品の創出を促進できるのではないか…という議論を展開したい。

ShowMe: 対話的な強調表示と拡大表示によるプレゼンテーションビデオの視覚的アクセシビリティの改善

SECHAYK Yotam* SHAMIR Ariel[†] 五十嵐 健夫 *

概要. プrezentationビデオを使った学習は広く一般的に行われている。講師は、ビデオ作成過程で、さまざまな視覚的補助動作を活用することが多い。具体的には、プレゼンテーション中のポイントинг、マーキング、スケッチなどが、視覚的補助動作としてよく使われる。しかし、これらの動作は視覚的に認識が難しいことが多い、説明が不十分であることが多い。弱視の学習者は、このような動作に追従するために、常にプレゼンテーションのフレーム内を探索する必要があり、フラストレーションと疲労につながっている。我々は、この問題を理解し解決するために、3人の弱視ユーザとユーザ参加型デザインを実施し、その結果にもとづき、講師の視覚的補助動作を強調表示し、拡大表示するツール ShowMeを開発した。ShowMeは、弱視ユーザがプレゼンテーションをフォローできるように支援し、疲労とフラストレーションを軽減する。

1 はじめに

学習動画は、正式および非公式の学習環境において、現代の教育において極めて重要である [34] [1] [18] [31]。教師はしばしばスライド [43]、インタラクティブ黒板 [32]、画面共有 [10]、2D/3D アニメーション [19]などの様々な視覚的ツールを使用し、多くのプレゼンテーションはこれらを組み合わせている。電子ピントタあるいはペンも、ポイント・マーク・スケッチにより学習者を導くのによく使われる [35]（図 1 を参照）。これらのツールは多くの学習に役立つが、低視力 (LV) の学習者にとってリアルタイムで気付くことが困難である [45]。

学習のためのユニバーサルデザインの原則とアクセシブルプレゼンテーションのガイドライン [4] [29] [13] は、記述的なスピーチによるアクセシビリティを推奨している。しかし、実際には、多くの教師はポイント・マーク・スケッチの際に不完全な説明を提供することが一般的である。例えば、教師はポイントングしながら抽象的な指示詞（例：「これ」、「ここ」）を使用したり、テキストを強調するためにマークしたり、図をスケッチしながら詳細を完全に説明しないことがある [13]。

これらの行動—ポイント・マーク・スケッチ—to ビジュアルアクティビティと呼ぶ。これらは教師が学習者を導き、情報を伝えるのに役立つ。これらの活動はしばしば微細で、小さく、コントラストが低く、気付きにくい。その結果、低視力の学習者は常にこれらの手掛かりを探さなければならず、視覚的探索の問題を示している [53]。



図 1. ビジュアルアクティビティの例. (A) ポイント, (B) マーク, (C) スケッチ.

3人の低視力学習者と共に、視覚的探索の問題を理解し解決するためのユーザ参加型デザインを行った。その結果、ビジュアルアクティビティの非アクセシビリティが関与・注意・挫折のレベルにどのように影響を与え、認知的負荷を増大させていることが示された。ユーザ参加型デザインの結果に基づき、視覚的探索の問題に対処するための ShowMeを開発した。ShowMeは2つの部分で構成されている：(1) 活動認識、(2) 活動可視化である。ShowMeはビジュアルアクティビティをより見やすく、追いやすくするのを助ける。

2 関連研究

2.1 動画におけるアクセシビリティ

低視力 (LV) のビデオアクセシビリティは重要な問題である [36]。解説放送 (AD) は伝統的に言語的な説明を追加する [38] [46] [41] [33] [39]。最近

Copyright is held by the author(s).

* 東京大学

† Reichman University

では、機械学習 [48] [5] [20] [52] [24] [23]、コラボレーション [21]、および空間オーディオ [16] による改善がなされている。しかし、AD はしばしば LV ユーザにとって過度に説明的に感じられる。

視覚的強化手法には、ポストプロセッシング（例：コントラスト、シャープネス）[44] や顕著性に基づく拡大 [2] が含まれるが、これらはビジュアルアクセシビリティを見逃したり、費用がかかり、プレゼンテーションに最適化されていない [25]。

先行研究において、我々は主に時間制御（速度調整およびリプレイ機能）に重点を置き、一般ユーザ向けのアクセシビリティの向上を探求した [45]。本研究では、特に弱視者に対するアクセシビリティに注目し、視覚的強化手法の探究を行い、デザインおよび実装の双方において改善を目指した。デザインプロセスにおいては、参加型デザインアプローチを採用し、パーソナライゼーションの重要性を強調した。実装面においては、ビジュアルアクセシビリティを正確に検出および分類するためのフレーム間関係を定式化する方法を提示する。

2.2 学習動画におけるビジュアル・コンテンツのアクセシビリティ

今日では、YouTube [51]、Khan Academy [27]、Coursera [8] などのプラットフォームが広範なプレゼンテーション動画を提供している。制作過程において、先行研究はアクセシブルなプレゼンテーションを簡単に作成する方法を提案している [42] [14] [28]。既存のプレゼンテーションに対して、先行研究はスライド要素の意味的なハイライト [26] [50] と、未対応のスライド要素の検出を探求している [43]。しかし、ビジュアルアクセシビリティはスライド要素とみなされず、未対応のままである。

2.3 映像内オブジェクトとのインタラクション

ビデオ内オブジェクトはインタラクティビティのために利用される。先行研究は、インタラクティブなナビゲーション [12] [32]、ソフトウェアシミュレーション [49] [3]、ノート作成 [6]、視線に基づくハイライト [37]、テキスト操作 [10]、またはマウスクリックを使用してスライド要素を拡大する手法を提案している [10]。本研究では、ビジュアルアクセシビリティをインタラクションのための基点として利用する。

3 ユーザ参加型デザイン

3人の低視力ユーザーと共にユーザ参加型デザインを行い、以下の点について調査した：(RQ1) 低視力ユーザーがビジュアルアクセシビリティをどのように体験するか、(RQ2) ビジュアルアクセシビリティから生じるチャレンジは何か、(RQ3) アクティビティをより見やすく追いややすくするにはどのアプローチが効果的か、および (RQ4) 本ユーザ参加型デザインにおける経験はどのようなものであったか。

3.1 方法

3.1.1 参加者

家族や友人を通じて3人の低視力参加者を募集した。参加者はそれぞれ異なる視覚状態で、異なるアクセシビリティのニーズを持っていた（表1を参照）。

3.1.2 手順

ユーザ参加型デザインはZoomを通じてオンラインで行われ、各参加者に対して5回の60分セッションを行った。これらのセッションは5週間にわたって実施した。参加者は、本研究全休に対して100ドルのギフトカードで時間に対する補償を受けた。セッション1では教育用プレゼンテーション動画を見る際の体験と課題について議論し（RQ1-RQ2）、セッション2-4ではShowMeプロトタイプを反復設計し（RQ3）、セッション5では協働設計の体験についての考察を行った（RQ4）。各セッションは分析のために記録され、文字起こしされた。

3.2 調査結果

本セクションでは、3名からそれぞれヒアリングした結果を記載する。その際、我々に対して説明した参加者をカッコ書きで文末に記載することとする。

3.2.1 低視力の経験 (RQ1)

参加者全員は、視覚情報が失われることは避けられないことが多いという考えを共有した。Johnはさらに、見逃したコンテンツを視覚的に特定できないと、何が見逃されているのかを把握するのが難しいと説明した。すべての参加者は、プレゼンテーション動画を視聴した経験があった。ビジュアルアクセシビリティに関して、参加者全員が「何を見ればいいのか」あるいは「どこを見ればいいのか」についての不確実性を表明した。その結果、オンラインプレゼン

表 1. ユーザ参加型デザイン参加者情報。

仮名	年齢	性別	視覚状態	発症	教育	スクリーン拡大鏡
John	36	男性	アルビニズム、乱視	先天性	高校卒	時々
Eva	30	女性	網膜色素変性症	先天性	大学卒	使用せず
Emily	33	女性	デュアン症候群	先天性	大学卒	使用せず

ンテーション動画の視聴にはより高いレベルの注意が必要であると、参加者全員が述べた。また、John と Eva は時には挫折したり、視聴を避けたりすることもあると述べた。これは、低視力ユーザが微妙な視覚変化を検出するために多大な努力を投入しなければならないことを示している。

3.2.2 低視力のチャレンジ (RQ2)

異なるプレゼンテーション動画は異なるスタイルの選択を示し、多くの教師が小さなポインタや低コントラスト、細い電子ペンを使用する。これらの視覚スタイルによって反応時間が伸び、参加者全員がビジュアルアクティビティの認識能力に影響が出ると感じた。John は特に「マーキング動作」が非常に微妙かつ短時間であるため難しいと感じている。Emily にとっては、教師が線と一緒に描くことで理解しづらくすることがよくある。

最後に、拡大ソフトウェアを使用することは面倒で、絶え間ない手動調整を要する (John)。したがって、いくつかのユーザは詳細を見るのをあきらめ (John)，またはデジタルで拡大できるように、スクリーンショットを頻繁に一時停止して保存する (Eva)。

3.2.3 見やすくするアプローチ (RQ3)

ビジュアルアクティビティに対処するためのアプローチを洗練するために、我々は以前の研究に基づいたプロトタイプでの設計を行った [45]。

初期の印象: 初期のプロトタイプは赤い半透明円を使用し、フレーム間の最新の視覚的变化をハイライトした。ズーム機能は、連続して 5 秒以上の変化が発生し、空間的な閾値内にある場合にトリガーとして追加された。すべての参加者はこれらのハイライトを非常に助かると感じた。John は「それがないと迷子になる」と述べた。しかし、参加者全員が視覚スタイルを制御できないことを批判した。また、ズームは有用であったが、教師のスケッチ線は依然として細すぎて見えなかった (John, Emily)。参加者はすべてのアクティビティでズームのオプションも望んでいた。

反復設計の結果: 最終設計では、ハイライトスタイル・ズーム機能・視覚フィルタ (例: 色反転) のパーソナライズを可能にした。ハイライト形状は、設定された半径の円またはアクティビティを囲むボックスの 2 つを提供した。丸を好む人もいるかもしれないが、すべての参加者は「正確にどこを見て何を見ればいいのか教えてくれる」としてボックスを好むと述べた (John)。アニメーションは視覚的な補助手段を強化する可能性があるため (Eva)，ハイライトをアニメーション化するオプションを追加した (サイズの変動)。しかし、John と Emily はそれを

煩わしいと感じた。また、John と Eva はタイミングを合わせるためにインストラクターの活動の事前通知が必要であると表明した。

これに対処するために、透明性を高めるためにハイライトの開始時間を 1.5 秒だけ増やした。Emily は、一貫性のある視覚化が見つけやすいと述べたため、カーソル (Ⓐ) や手 (Ⓑ) を追加した。指示するなどのさまざまな活動のために異なるビジュアルを試みたが、参加者は一貫性のあるスタイルを好んだ。

参加者の意見に基づき、すべてのアクティビティでズームを有効にし (Eva)，ズームがアクティブ時に自動で再生が一時停止するトグルを追加し、ズーム速度と拡大率を調整できるようにした (John, Eva)。これらの機能は有益と評価された。結果として、すべての参加者が ShowMe を集中力向上と挫折軽減に役立つものと感じ、John は「それなしでは迷子になる」と述べ、Eva は「特別なメガネのようだ」と評価した。

3.2.4 本ユーザ参加型デザインにおける経験 (RQ4)

最終セッションでは、参加者に本ユーザ参加型デザインでの経験について尋ねた。すべての参加者は ShowMe の設計に意味のある影響を与えたと感じた。Eva は、「それについて話した後にアニメーションが (実装されているの) 見てとても嬉しかった」と述べた。パーソナライズに関しては、すべての参加者が実装されたハイライトとズーム設定に満足感を表明した。John と Eva は将来的に ShowMe を使用したいと述べたが、Emily は新しいツールに適応することに関して懸念を示し、「アクセシビリティ機能のために以前 iPad の使用を試みたがいつも使っているものとは非常に異なっており、あきらめてしまった」と述べた。

しかし、Emily は研究への参加がアクセシビリティツールを試そうという姿勢にわずかな変化をもたらしたと述べた。研究中にいくつかの目の疲れや疲労を経験したが、すべての参加者は参加したこと嬉しく思っていた。今後目の疲労を軽減するため、より長い休憩や短くてもっとコースを空けたセッションが役立つと提案された。

4 ShowMe

この章では、我々のアプローチを詳細に説明し、ユーザインタラクションを示す。ShowMe は、アクティビティ認識を使用し、アクティビティビジュализエーションを生成する (図 2)。

4.1 アクティビティ認識

ビジュアルアクティビティは、ポイント・マーク・スケッチである (図 1)。我々のアプローチは以前の研究に基づき [45]、既存のコンピュータビジョン (CV) 動き検出アルゴリズムを使用している [40]。

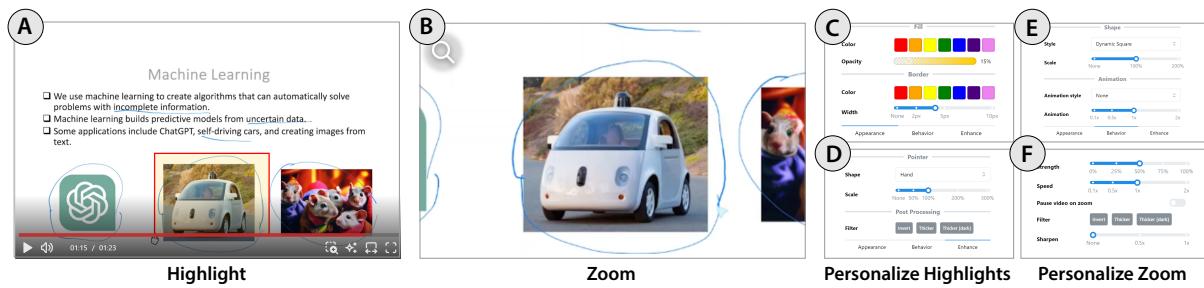


図 2. ShowMe のビデオプレーヤ. (A) 教師のアクションを教師が話すときにハイライト（赤の輪郭、黄色の塗りつぶし）、(B) 教師の活動にズームするオプションを提供. ShowMe はユーザに (C-E) 視覚的ハイライトと (F) ズーム動作をカスタマイズさせる。

従前の研究を時間領域に拡張し、アクティビティを時間的に認識しクラスタリングする。前回の研究がフレーム間の局所的な違いを検査していたのに対し、ShowMe は違い間の関係を構造化して時間的にアクティビティを作成する。

4.2 アクティビティ視覚化

認識されたアクティビティは、ハイライトを通して可視化され、ズーム機能を備えている。ハイライトとズームはビデオ再生に追加され、完全なユーザパーソナライゼーションが可能である。図 2 はユーザインターフェースを示している。

4.2.1 アクティビティハイライト

3章で述べたように、低視力ユーザは幅広いニーズを持っており、パーソナライゼーションが重要な要素である。ユーザは追加ボタン を使用して設定パネルを開くことができる（図 2 の C-E を参照）。ハイライトの設定には、背景色と不透明度、境界色と厚み、ポストプロセッシングフィルタ、アニメーションなどが含まれる（表 2 を参照）。

表 2. ハイライトのデフォルト設定。

カテゴリー	設定	デフォルト
ビジュアル	塗り色	黄色
	塗り不透明度	15%
	境界色	赤
	境界幅	4px
動作	形状スタイル	ボックス
	スケール	100%
	アニメーション種類	なし
	アニメーション速度	1.0x
改善	ポインタースタイル	なし
	ポインタースケール	100%
	フィルタ	なし

4.2.2 アクティビティ拡大

拡大は Z キーボードキーを押すことによって行われる。押すと、ビデオプレーヤは現在ハイライトされているアクティビティにスムーズにズームする。アクティビティの位置が更新されると、ズームエリアもスムーズな遷移で更新される。ユーザは追加ボタン を使用して設定パネルを開くことができる（図 2 の F を参照）。ビデオプレーヤは、上左に拡大の視覚的インジケーションを含んでいる（図 2 の B 参照）。ユーザはズーム遷移速度を個別化し、拡大時に一時停止するトグルなどを利用できる（表 3 を参照）。

表 3. ズームのデフォルト設定。

設定	デフォルト
ズーム強度	50%
ズーム速度	1.0x
ズーム時の一時停止	いいえ
シャープネス	1.0x
フィルタ	なし

4.3 インタラクションの概要

低視力ユーザが学習動画を視聴すると、アクティビティのハイライトが表示される。ユーザは再生中でも設定をリアルタイムで変更可能で、Z キーを押すと拡大表示に切り替わる。拡大後は矢印キーで拡大レベルを調整できる。タイムラインナビゲーションや時間スキップ、キーボードショートカット、フルスクリーンなどの通常機能もサポートされ、ユーザは引き続きスムーズな操作が可能である。また、Z キーで拡大を解除したり、他の設定を調整することもできる。

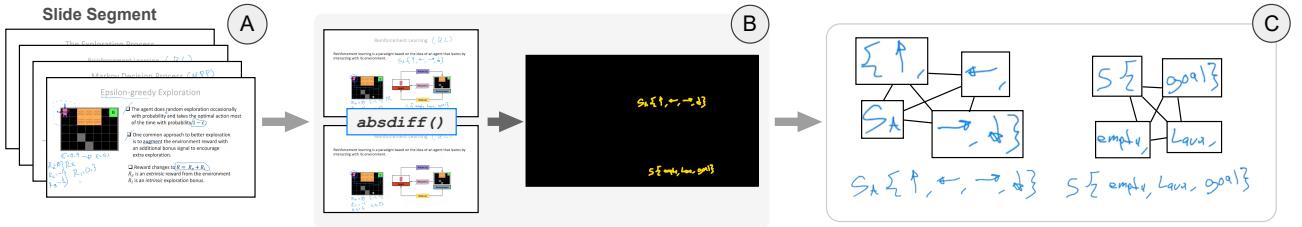


図 3. ShowMe のアクティビティ認識パイプライン. (A) スライドセグメントから始まり, 各フレームウィンドウに対して (B) 視覚的な差異を認識し, 最後に (C) 接続されたコンポーネント解析に基づいてアクティビティを分類するグラフとして差異を表現している.

5 実装

5.1 アクティビティ認識

認識パイプラインは 3 つの主要なコンポーネントから成り立っている：(1) スライドセグメント抽出, (2) アクティビティ検出, (3) アクティビティ分類である（図 3 参照）。このパイプラインは既存の研究 [45] を拡張したものであり, アクティビティ認識の基盤として安価な動き検出手法 [40] を使用している。Python [17] で開発し, OpenCV [30] と NetworkX [11] ライブライアリを使用している。

5.1.1 スライドセグメント抽出

最初のステップはスライドセグメントの抽出である。以前の研究 [43] [45] と同様に, PySceneDetect [9] を使用してスライド遷移を検出している。これは Python 用の既製のショット検出ライブラリで, Content Detector を使用している。スライドセグメントはそのセグメント内でアクティビティの検出を境界づけるのに使用される。

5.1.2 アクティビティ検出

アクティビティ検出は, 2 つの近接したフレーム間の差異を検出することから始まる [40]。アルゴリズムはフレームのスライディングウィンドウで各スライドセグメントに適用される。デフォルトのウィンドウサイズ W_{size} は $0.33 \cdot V_{fps}$ に設定され, ここで V_{fps} は動画のフレーム毎秒である。それぞれの結果として得られる輪郭 [47] は, 圧縮アーティファクトや見逃せない大きな変化（例：フレームの 60% を超える）を無視するためにサイズでフィルタリングされる。その結果として得られるのは変化領域 (RoC) である。次に, RoC からグラフ表現を構築する。すべての RoC について, 表しているノードは

$$n_i = (T, X, Y, W, H) \quad (1)$$

である。ここで T は開始タイムスタンプ, X と Y は左上の座標, W と H は RoC の幅と高さである。ノード n_i と n_j の間にエッジが作成されるのは, 2 つの RoC が次の条件を満たす場合である：

- 時間的に近い: $\sqrt{T_i^2 - T_j^2} \leq \Theta_T$ ここで Θ_T は設定された閾値である。
- 空間的に近い: $\sqrt{\Delta X^2 + \Delta Y^2} \leq \Theta_S$ ここで ΔX と ΔY はそれぞれ最小の水平または垂直距離であり, Θ_S は設定された閾値である。

エッジの重みは, 2 つの RoC 間の視覚的類似度として計算される。視覚的類似度は hu モーメント [22] を使用して基礎となる輪郭形状をマッチングすることによって計算される。[10] に触発され, 2 つの RoC をマージする条件は以下である：

- 時間的に隣接: $\sqrt{T_i^2 - T_j^2} \leq \frac{W_{size}}{V_{fps}}$.
- 空間的に同一: $\sqrt{\Delta X^2 + \Delta Y^2} \leq \epsilon$ ここで ϵ は小さな値である。
- 視覚的に類似: $Sim(n_i, n_j) \geq \Gamma$ ここで Γ は設定された閾値である。

これによりオブジェクトの動き（例：ポインタ）が見逃されるのを防ぐ。

5.1.3 アクティビティ分類

RoC グラフの連結成分 (CC) 解析を使用する。各 CC について, アクティビティ

$$a_i = (CC_i, X, Y, W, H, T_{start}, T_{end}) \quad (2)$$

を定義する。ここで CC_i は連結成分を表しており, X , Y , W , H はバウンディングボックス（つまり, 関連するすべての RoC を含む最小の長方形）を表している。タイムスタンプ T_{start} と T_{end} はアクティビティの開始と終了を表す。次に, W と H がそれぞれビデオの幅と高さの 1% から 70% の間にある場合にアクティビティを有効としてマークする。最後に, 各アクティビティのタイプをヒューリスティックなルールベースのアプローチを使用して分類する。例えば, 平均エッジ類似度がある閾値を超える CC は pointing として分類できる。ユーザ参加型デザイン参加者は一貫したスタイルを好んでいたが, 将来の研究では, より頑健な分類アプローチを用いてスタイルの多様性をさらに探る必要がある。

5.2 アクティビティ視覚化

ユーザ参加型デザインの結果、低視力ユーザの多様な視覚ニーズに対応するためのパーソナライズの重要性を強調している。そこで、ビデオ再生の拡張オーバーレイとして視覚化を実装した。この実装はTypeScript [7] を用いてReact [15] 上で行い、ユーザ設定やデバイス（例：画面解像度、アスペクト比）に適応するリアクティブな動作を実現した。

5.2.1 アクティビティハイライト

再生中、各 `ontimeupdate` ビデオプレーヤーイベントは現在のアクティビティを収集する。ShowMeは、アクティビティの開始時刻に基づくヒューリスティックを用いて最も関連性のあるアクティビティを選択する。この選択されたアクティビティはユーザの設定に基づいて視覚化される。以前の研究 [45] が静的スタイルを使用していたのとは異なり、ShowMeのキーとなる機能はリアルタイムパーソナライゼーションとリアクティブデザインである。

5.2.2 アクティビティ拡大

ズームインビューは、動画フレームを反映する `canvas` エレメントを使用して実装されている。拡大は (C_x, C_y, S_x, S_y, Z) で定義され、ここで C_x, C_y はアクティビティの中心、 S_x, S_y は左右のシフト量、 Z はズームファクターである。アクティビティの中心にズームファクターでズームしながら、ズームしたビューをビデオパネルに完全に収めるようにシフトする。

6 議論

バリアフリー映像。 低視力ユーザは残存視力に頼ってタスクを行うことが多いため、視覚コンテンツへのアクセスの必要性は重要である。このアクセスを促進するためには、ユーザはコンテンツを個々のニーズに合わせて適応させる必要がある。視覚アクセスが利用できない場合、多くの弱視ユーザーは疲れ、フラストレーションを経験し、中には作業を放棄する人もいる。弱視者の視覚状態は千差万別であるため、万能なものは存在しない。例えば、アニメーションは視野の狭い人には役立つが、視力の弱い人には非常に邪魔になる。したがって、パーソナライゼーションと適応性はアクセシビリティの重要な要素である。我々のユーザ参加型デザインはプレゼンテーション動画に焦点を当たが、他の領域でも同様の課題が経験されている [53]。

ユーザ参加型デザイン。 アクセシビリティにおいて可視性認識は重要な考慮点である。一般的な信念に反して、最も重要な可視性認識は、すべてを認識することはできないという認識である。我々のアプローチはこの原則を反映するためにユーザ参加型デザイ

ンの結果を反映している。参加者の数は限られているが、彼らの独自の経験は、効果的なアクセシビリティツールを作成する上でユーザ主導の設計の必要性を強調している。ShowMeは馴染みのある視覚的慣例に基づいて作られ、さらに適応性とパーソナライゼーションが追加されている。しかし、将来の研究ではShowMeをより大規模な低視力ユーザグループでテストし、利用パターンや嗜好をより明確にする必要がある。

技術的限界とユーザ満足度。 我々のアクティビティ認識アプローチにはいくつか重要な限界がある。例えば、ShowMeはアクティビティの意味的な重要性を考慮しておらず、複数のアクティビティが発生した場合に重要でないものをハイライトする可能性がある。ハイライトは参照している要素ではなく、変更そのものに焦点を当てる；例えば、テキストをマークする際、マーク自体をハイライトし、関連するテキストではない。将来の研究では、アクティビティ検出と視覚ハイライトに対する意味的アプローチを検討すべきである。それにもかかわらず、ユーザ参加型デザインの結果は、不完全な解決策でも低視力ユーザの体験を大幅に向上させることができることを示している。

新ドメインへの拡張。 視覚的探索問題は低視力ユーザの生活に一般的にみられる問題であるが [53] [45]、我々のアクティビティベースのアプローチはプレゼンテーション動画に特化されている。しかし、我々のアプローチは機械学習に頼らないため、ソフトウェアチュートリアルのような他のビデオ領域にも、さらにはVR/ARのような他のメディアにもシームレスに適応することができる。ただし、ライブ放送には追加の考慮が必要である。ShowMeは完全に利用可能な時間情報に依存しているため、ライブシナリオではアクティビティの境界（空間と時間の両方）をよりよく特定するために人工バッファを必要とする。将来の研究では、ライブコンテキストにおける低視力ユーザの好みを調査すべきである。

7 結論

プレゼンテーション動画で、教師は学習者を視覚的に導くために電子ポインタあるいはペンを使用することが多いが、低視力（LV）ユーザはこれらの微細な手がかりを頻繁に見逃す。我々のユーザ参加型デザインは、LV学習者が小さな視覚変化を常に探す「視覚的探索問題」を特定するのに役立った。それに応じて、これらの活動を検出し、ハイライトと拡大で再生を向上させるアクセシブルツールであるShowMeを開発した。パーソナライゼーションを念頭に置いて設計されたShowMeは、LV参加者の挫折を軽減し、学習の容易さを改善する。

謝辞

本研究は *Google*, および JST AdCROP (JP-MJKB23) の一部支援を受けて行われた。その経済的なサポートに感謝の意を表す。

参考文献

- [1] W. Aljedaani, M. Alkahtani, S. Ludi, M. W. Mkaouer, M. M. Eler, M. Kessentini, and A. Ouni. The State of Accessibility in Blackboard: Survey and User Reviews Case Study. In *Proceedings of the 20th International Web for All Conference*, W4A '23, p. 84–95, New York, NY, USA, 2023. Association for Computing Machinery.
- [2] A. S. Aydin, S. Feiz, V. Ashok, and I. Ramakrishnan. Towards making videos accessible for low vision screen magnifier users. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, IUI '20, p. 10–21, New York, NY, USA, 2020. Association for Computing Machinery.
- [3] N. Banovic, T. Grossman, J. Matejka, and G. Fitzmaurice. Waken: reverse engineering usage information and interface structure from software videos. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pp. 83–92, 2012.
- [4] S. E. Burgstahler and R. C. Cory. *Universal design in higher education: From principles to practice*. Harvard Education Press, 2010.
- [5] V. P. Campos, L. M. Gonçalves, W. L. Ribeiro, T. M. Araújo, T. G. Do Rego, P. H. Figueiredo, S. F. Vieira, T. F. Costa, C. C. Moraes, A. C. Cruz, et al. Machine generation of audio description for blind and visually impaired people. *ACM Transactions on Accessible Computing*, 16(2):1–28, 2023.
- [6] Y. Cao, H. Subramonyam, and E. Adar. Video-Sticker: A tool for active viewing and visual note-taking from videos. In *Proceedings of the 27th International Conference on Intelligent User Interfaces*, pp. 672–690, 2022.
- [7] M. Corporation. TypeScript: JavaScript With Syntax for Types. <https://www.typescriptlang.org/>, 2012. Accessed: 2024-10-21.
- [8] Coursera. Coursera. <https://www.coursera.org/>. Accessed: 2023-10-05.
- [9] R. D. and A. D. PySceneDetect, 2024.
- [10] L. Denoue, S. Carter, M. Cooper, and J. Adcock. Real-time direct manipulation of screen-based videos. In *Proceedings of the companion publication of the 2013 international conference on Intelligent user interfaces companion*, pp. 43–44, 2013.
- [11] N. Developers. NetworkX, 2024.
- [12] P. Dragicevic, G. Ramos, J. Bibliowicz, D. Nowrouzezahrai, R. Balakrishnan, and K. Singh. Video browsing by direct manipulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 237–246, 2008.
- [13] Alfred T D'Agostino. Accessible teaching and learning in the undergraduate chemistry course and laboratory for blind and low-vision students. *Journal of Chemical Education*, 99(1):140–147, 2021.
- [14] M. Elias, A. James, E. Ruckhaus, M. C. Suárez-Figueroa, K. A. De Graaf, A. Khalili, B. Wulff, S. Lohmann, and S. Auer. SlideWiki-Towards a Collaborative and Accessible Platform for Slide Presentations. In *EC-TEL (Practitioner Proceedings)*, pp. 1–3, 2018.
- [15] I. Facebook. React: A JavaScript Library for Building User Interfaces. <https://reactjs.org/>, 2013. Accessed: 2024-10-21.
- [16] D. Fan, S. Junuzovic, J. C. Tang, and T. Jaeger. Improving the Accessibility of Screen-Shared Presentations by Enabling Concurrent Exploration. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS 2023, New York, NY, USA, October 22–25, 2023, pp. 44:1–44:16. ACM, 2023.
- [17] P. S. Foundation. Python Programming Language, 2024.
- [18] S. Ghosh and A. Figueroa. Establishing TikTok as a Platform for Informal Learning: Evidence from Mixed-Methods Analysis of Creators and Viewers. In T. X. Bui ed., *56th Hawaii International Conference on System Sciences*, HICSS 2023, Maui, Hawaii, USA, January 3–6, 2023, pp. 2431–2440. ScholarSpace, 2023.
- [19] S. Grice and J. Hughes. Can music and animation improve the flow and attainment in online learning? *Journal of Educational Multimedia and Hypermedia*, 18(4):385–403, 2009.
- [20] T. Han, M. Bain, A. Nagrani, G. Varol, W. Xie, and A. Zisserman. AutoAD: Movie description in context. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18930–18940, 2023.
- [21] M. Hirvonen, M. Hakola, and M. Klade. Co-translation, consultancy and joint authorship: User-centred translation and editing in collaborative audio description. *Journal of Specialised Translation*, (39):26–51, 2023.
- [22] M.-K. Hu. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 8(2):179–187, 1962.
- [23] G. Jain, B. Hindi, C. Courtien, C. Wyriick, X. Y. T. Xu, M. C. Malcolm, and B. A. Smith. Towards accessible sports broadcasts for blind and low-vision viewers. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–7, 2023.
- [24] L. Jiang, M. Phutane, and S. Azenkot. Beyond audio description: Exploring 360 video accessibility with blind and low vision users through

- collaborative creation. In *Proceedings of the 25th international ACM SIGACCESS conference on computers and accessibility*, pp. 1–17, 2023.
- [25] L. Jiang, M. Xu, T. Liu, M. Qiao, and Z. Wang. Deepvs: A deep learning based video saliency prediction approach. In *Proceedings of the european conference on computer vision (eccv)*, pp. 602–617, 2018.
 - [26] H. Jung, H. V. Shin, and J. Kim. Dynamic-slide: Exploring the design space of reference-based interaction techniques for slide-based lecture videos. In *Proceedings of the 2018 Workshop on Multimedia for Accessible Human Computer Interface*, pp. 33–41, New York, NY, USA, 2018. ACM.
 - [27] Khan Academy. Khan Academy. <https://www.khanacademy.org/>. Accessed: 2024-08-10.
 - [28] J. Kong, D. Sabha, J. P. Bigham, A. Pavel, and A. Guo. TutorialLens: authoring Interactive augmented reality tutorials through narration and demonstration. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, pp. 1–11, 2021.
 - [29] R. E. Ladner and K. Rector. Making your presentation accessible. *Interactions*, 24(4):56–59, 2017.
 - [30] O. S. C. V. Library. OpenCV, 2024.
 - [31] I. A. Mastan, D. I. Sensuse, R. R. Suryono, and K. Kautsarina. Evaluation of distance learning system (e-learning): a systematic literature review. *Jurnal Teknoinfo*, 16(1):132–137, 2022.
 - [32] T.-J. K. P. Monserrat, S. Zhao, K. McGee, and A. V. Pandey. Notevideo: Facilitating navigation of blackboard-style lecture videos. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1139–1148, 2013.
 - [33] R. Natalie, J. Loh, H. S. Tan, J. Tseng, I. L. Y.-R. Chan, E. H. Jarjue, H. Kacorri, and K. Hara. The efficacy of collaborative authoring of video scene descriptions. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–15, 2021.
 - [34] E. Navarrete, A. Nehring, S. Schanze, R. Ewerth, and A. Hoppe. A Closer Look into Recent Video-based Learning Research: A Comprehensive Review of Video Characteristics, Tools, Technologies, and Learning Effectiveness. *CoRR*, abs/2301.13617, 2023.
 - [35] E. Navarrete, A. Nehring, S. Schanze, R. Ewerth, and A. Hoppe. A Closer Look into Recent Video-based Learning Research: A Comprehensive Review of Video Characteristics, Tools, Technologies, and Learning Effectiveness. *arXiv preprint arXiv:2301.13617*, 2023.
 - [36] A. Neovsky, T. Neate, E. Simperl, and R. Vatavu. Accessibility Research in Digital Audiovisual Media: What Has Been Achieved and What Should Be Done Next? In P. L. Callet, M. P. D. Silva, T. Vigier, K. Tahiroglu, N. Murray, G. Valenzise, and M. Wang eds., *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences, IMX 2023, Nantes, France, June 12-15, 2023*, pp. 94–114, Nantes, France, 2023. ACM.
 - [37] C. Nguyen and F. Liu. Gaze-based notetaking for learning from lecture videos. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 2093–2097, 2016.
 - [38] A. C. of the Blind. Audio Description Project, Guidelines for Audio Describers. <https://www.acb.org/adp/guidelines.html>, n.d. Accessed: 2024-08-10.
 - [39] J. Packer, K. Vizenor, and J. A. Miele. An overview of video description: history, benefits, and guidelines. *Journal of Visual Impairment & Blindness*, 109(2):83–93, 2015.
 - [40] S. Parveen and J. Shah. A motion detection system in python and opencv. In *2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV)*, pp. 1378–1382, Virtual Conference, 2021. IEEE, IEEE.
 - [41] A. Pavel, G. Reyes, and J. P. Bigham. Resscribe: Authoring and automatically editing audio descriptions. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pp. 747–759, 2020.
 - [42] Y.-H. Peng, J. Jang, J. P. Bigham, and A. Pavel. Say it all: Feedback for improving non-visual presentation accessibility. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2021.
 - [43] Y. Peng, J. P. Bigham, and A. Pavel. Slidecho: Flexible Non-Visual Exploration of Presentation Videos. In J. Lazar, J. H. Feng, and F. Hwang eds., *ASSETS '21: The 23rd International ACM SIGACCESS Conference on Computers and Accessibility, Virtual Event, USA, October 18-22, 2021*, pp. 24:1–24:12. ACM, 2021.
 - [44] A. Sackl, F. Graf, R. Schatz, and M. Tscheiligi. Ensuring accessibility: Individual video playback enhancements for low vision users. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–4, 2020.
 - [45] Y. Sechayk, A. Shamir, and T. Igarashi. Smart Replay: e ラーニング動画における視覚的・時間的アクセシビリティの向上. In *Proceedings of the Workshop on Interactive Systems and Software (WISS)*, Japan, 2023.
 - [46] J. Snyder. Audio description: The visual made verbal. In *International congress series*, Vol. 1282, pp. 935–939. Elsevier, 2005.
 - [47] S. Suzuki, et al. Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, 30(1):32–46, 1985.

- [48] Y. Wang, W. Liang, H. Huang, Y. Zhang, D. Li, and L.-F. Yu. Toward automatic audio description generation for accessible videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2021.
- [49] S. Yang, J. Vermeulen, G. Fitzmaurice, and J. Matejka. AQuA: Automated Question-Answering in Software Tutorial Videos with Visual Anchors. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–19, 2024.
- [50] C. Yip, J. M. Chong, S. Y. Kwek, Y. Wang, and K. Hara. Visionary Caption: Improving the Accessibility of Presentation Slides through Highlighting Visualization. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–4, 2021.
- [51] YouTube. YouTube. <https://www.youtube.com/>. Accessed: 2024-08-10.
- [52] B. F. Yuksel, P. Fazli, U. Mathur, V. Bisht, S. J. Kim, J. J. Lee, S. J. Jin, Y.-T. Siu, J. A. Miele, and I. Yoon. Human-in-the-loop machine learning to increase video accessibility for visually impaired and blind users. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pp. 47–60, 2020.
- [53] Y. Zhao, S. Szapiro, J. Knighten, and S. Azenkot. CueSee: exploring visual cues for people with low vision to facilitate a visual search task. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp ’16*, p. 73–84, New York, NY, USA, 2016. Association for Computing Machinery.

サブセグメント分割に基づく 3D スキャンモデルのインタラクティブセグメンテーション

真殿 航輝^{*†} 五十嵐 健夫[†] 加藤 大晴[†] 橋本泰輔[†] Fabrice Matulic[†] 高木 士[†] 樋口啓太[†]

概要. 3D スキャンモデルにおいてパーツ毎のテクスチャセグメンテーションは、モデルの再編集やマテリアル付与の目的で重要である。従来の手法では手動によるセグメンテーションが手間である一方、自動セグメンテーション技術ではユーザの意図通りに分割される保証がない。本研究では、自動セグメンテーションと最小限の手動操作を組み合わせたインタラクティブセグメンテーションツールを提案する。このツールは、前処理として 3D モデルを多視点から自動セグメンテーションすることでテクスチャをサブセグメントに分割し、実行時にユーザが 3D モデルビューで簡単な線描画でクラスタ化する事で、最終セグメンテーション結果を得る。評価の結果、本アプローチは手動セグメンテーションや標準的な 3D コンピュータグラフィックスソフトウェアに比べて精度と品質を向上させ、UV テクスチャマップの自動セグメンテーションに比べてより詳細なセグメントを生成できることが示された。

1 はじめに

コンピュータビジョンや機械学習技術の進展により、高品質な 3D スキャンモデルが幅広いユーザに利用可能である。3D 再構築技術は、写真から正確なモデルとテクスチャの両方の生成を可能にした。3D スキャンモデル利用が広がるにあたり、後処理が重要となり、色の変更、材質の修正、部品の形状変更など、特定のユーザ要件に対してモデルを編集できる事が求められている。一方で、3D モデリングソフトウェアで作成された 3D スキャンモデルとは異なり、3D スキャンモデルではパーツや意味のあるセグメントに分割する方法は自明ではない。そのため、ユーザは後処理で 3D スキャンモデルのセグメンテーションを行う必要がある。

3D スキャンモデルのセグメンテーション方法として、3D ソフトウェアを使用してテクスチャを手動で分割する方法が挙げられる。例えば、Blender¹は 3D ペイント機能としてブラシや塗りつぶしツールを提供しており、これにより 3D スキャンモデルをユーザの意図通りにセグメンテーションすることが可能である。しかし、特にパーツ間の境界を精度よく分割することは手動セグメンテーションでは手間がかかり、3D ペイントには精密な操作が必要である。また、インタラクティブな 3D のセグメンテーションに関する従来手法として、Click-base [7] の手法や Graph-cut を用いた手法 [2] が提案されている。しかしながら、複数のクラスの分割の際にユーザーの意図する形に分割する事が難しく、特に今回

対象としているような 3D スキャンモデルに対しての有効性が十分に検証されていない。

別のアプローチとして、自動セグメンテーション手法を利用する方法がある。k-means クラスタリング [9] やスーパーピクセルセグメンテーション [1] などのヒューリスティック手法はユーザの介入なしにセグメンテーション可能であるが、幾何情報とテクスチャ情報の両方を考慮した手法が少なく、ユーザーの意図しないセグメンテーション結果になる事が多い。ディープラーニングに基づくより高度な技術は、意味のあるセグメントの自動抽出において優れた能力を示している [4, 11, 3] が、ユーザのニーズやアプリケーションの要求に合致する結果を保証するわけではない。ユーザの多様な好みや理想的なセグメンテーションに対する精密な結果を自動セグメンテーションだけで達成する事は難しく、3D スキャンモデルを使用するそれぞれのアプリケーションのニーズに適切に対応するためには、インタラクティブなセグメンテーション技術が必要である。

本研究では、手動と自動セグメンテーション手法の欠点を克服するために、2つを組み合わせた半自動インタラクティブセグメンテーション技術を提案する。図 1 に全体像を示す。まず、3D スキャンモデル上でのテクスチャ画像をサブセグメントに分割する。その後、ユーザが 3D ビュー上でサブセグメント同士を連結する。具体的には、ユーザがスクリブルをモデルに描き込むと、交差するサブセグメントにセグメンテーション ID が割り当てられ、ID が周辺のパッチに伝播してセグメンテーションが生成される。このプロセスを繰り返し、ユーザがスクリブルを追加しながら希望するセグメンテーションを作成していく。ユーザ実験を通して本手法が伝統的な手動ペイントインターフェースに比べより効率的

Copyright is held by the author(s).

* 早稲田大学

† Preferred Networks

¹ <https://www.blender.org>

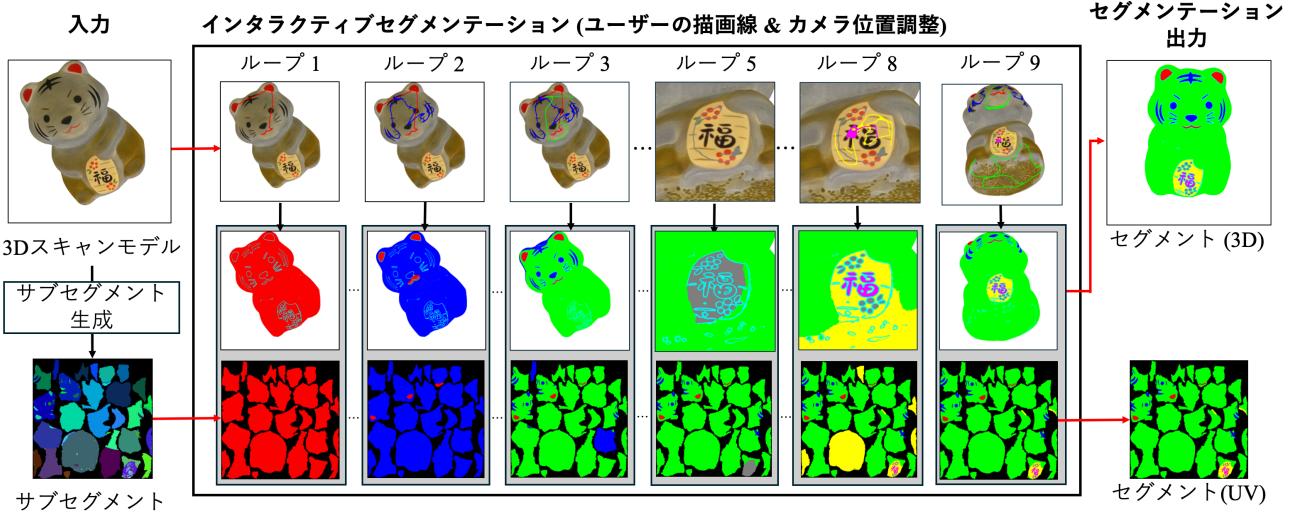


図 1. 3D スキャンモデルのインタラクティブセグメンテーションのワークフロー。自動セグメンテーションを用いて、事前処理ステップでテクスチャをセマンティックかつ十分に小さいサブセグメントに分割する。ユーザはその後、3D スキャンモデル上にスクリブルを描くことで、サブセグメントを反復的にセグメント化する。サブセグメント ID は、ユーザが書いた最も近いスクリブルの ID が割り振られる。

にセグメンテーション出来ることを確認し、自動セグメンテーションの前処理と手動セグメンテーションの後処理を組み合わせた提案法の有効性を明らかにした。

2 関連研究

2.1 自動 3D セグメンテーション

初期の自動 3D セグメンテーションは、K-means [9]、およびグラフカット [8] などのヒューリスティック手法が用いられていた。近年では、Segment Anything Model (SAM) [6] が 2D セグメンテーションで顕著な精度改善を達成しており、SAM の応用技術であるマッチングベースのアプローチ、ニューラルラジアンスフィールド (NeRF) ベースのアプローチ [10]、およびガウススプラッティングベースのアプローチ [5] などの様々な 3D セグメンテーション手法が提案されている。これらの先行研究は複数の画像からメッシュを生成するアプローチであるのに對し、本研究は 3D スキャンモデルと対応する UV テクスチャを用いており、関連研究とは異なる問題設定である。

本研究と同じ問題設定である SAM3D [11] は、複数視点のセグメンテーション結果をそれぞれマッチングベースで結合・分割し、自動で 3D 物体のセグメンテーションを生成する技術である。しかし、結合する際のハイパーパラメーターに精度が依存し、ユーザが希望するセグメンテーションの獲得を保証するアプローチではない。

本研究では、3D レンダリング画像と UV テクスチャのセグメンテーション結果から小さなサブセグ

メントと対応する距離行列を生成し、ユーザが効率的にサブセグメントをグループ化可能なインタラクティブセグメンテーション技術を提案する。

2.2 インタラクティブ 3D セグメンテーション

インタラクティブ性は、ユーザの負担を減らすために重要な役割を果たしている。3D ペイントソフトウェアである Blender¹ や ZBrush² は、ユーザにとって便利な描画ツールを用意する事で、インタラクティビティを向上させている。特にバケツツールとブラシツールを組み合わせることで、セグメンテーションの効率を大幅に向上させている。しかし、これらの解決策は一定のソフトウェア習熟度や作業時間を要求するため、すべてのユーザにとって使いやすいとは限らない。

本研究ではより多くのユーザにとって使いやすく効率的なインタラクティブセグメンテーションを実現するために、ユーザのスクリブルのみを入力とするインタラクティブセグメンテーション技術を提案する。

3 ユーザインターフェース

3.1 対象問題設定

3D スキャンモデルには、3D メッシュと UV テクスチャの両方が含まれる。しかし、パーツ毎にセグメンテーションはされていないため、図 2 のような質感やテクスチャの編集を行うためには、3D メッシュと UV テクスチャに対しセグメンテーションを行う必要がある。ユーザがテクスチャを手動で分類し、UV ピクセルにクラスラベルを割り当てること

サブセグメントに基づく 3D インタラクティブセグメンテーションシステム



図 2. 3D スキャンモデルのセグメンテーションと編集

は可能であるが、パーツ間の境界を精緻に分割することに時間がかかるといった課題がある。

そこで本研究では、UV テクスチャを大きさが異なるサブセグメントに自動で分解し、ユーザがスクリブルを使用してサブセグメントを効率的にグループ化出来るインタラクティブセグメンテーション手法を提案する。手動ラベリングの負担を大幅に削減するために 3D スキャンモデルのセグメンテーションプロセスを簡素化する事で、本研究はユーザ体験と効率を向上させることを目指している。

3.2 ユーザフロー

提案システムではユーザがスクリブルの色を選択し、テクスチャ付き 3D スキャンモデルの表面にスクリブルを描く。システムは、スクリブルと交差するサブセグメントに色に対応しているセグメント ID を割り当てる。その後システムは、セグメント ID を近くのサブセグメントに伝播させる。これにより、スクリブルが交差していないサブセグメントに対してもセグメント ID を割り当てることで、セグメンテーション結果を推定する。システムが推定した結果に応じてユーザは追加でスクリブルを加えて結果を修正する。このようにシステムと反復的に作業する事で、希望するセグメンテーションを得られる。

Figure 1 に例を示す。プロセスは、3D スキャンモデルとセクション 4 のアルゴリズムに基づいて生成されたサブセグメントを初期状態として始める。ユーザとの初回の反復（ループ 1）では、ユーザが赤色の線を適用して、黒い眉毛・目・ひげなどの特徴と一緒にクラスタリングする。その結果、すべてのサブセグメントが赤色としてセグメント化されている。次の反復（ループ 2）では青色の線が追加され、セグメント境界が再評価される。この結果、耳や口は同じ赤のセグメントになり、他の顔の表面は青いセグメントとして区別されている。これらの反復プロセスにより、ユーザはセグメントを修正可能である。またユーザは反復プロセスの過程でカメラ位置の調整によりモデルの視点を変更し、モデル全体のセグメンテーションの精度を確認する事ができる。



図 3. システムのスクリーンショット：このシステムには、ブラシ調整（ブラシサイズ、色）、元に戻す／やり直す操作、UV テクスチャの保存操作、カメラ制御が含まれている。メイン画面では、3D スキャンモデルの UV テクスチャとセグメンテーションを同時に確認でき、サブセグメントの境界線は常にライトブルーで表示される。テクスチャビューに描画された複数色のスクリブルに応じて、右画面のセグメンテーション結果が更新される。

3.3 ツール

図 3 に、提案のインタラクティブセグメンテーションツールのスクリーンショットを示す。本システムは、ユーザが 3D 表面にスクリブルを付与する半自動セグメンテーションを対象としている。以下の項目が提案システムの特徴となっている。

リアルタイムプレビュー インターフェースでは、ユーザがセグメントを作成しながら、3D スキャンモデルのテクスチャとジオメトリの両方を同時に確認できる。システム上では、元のテクスチャを表示するビュー（図 3、左側）と、セグメンテーション領域を表示するビュー（図 3、右側）の 2 つが利用されている。

3D 表面への線描画 セグメンテーション境界を手動でトレースする事が難しいため、システムはテクスチャを小さいサブセグメントに分割し、ユーザはスクリブルを用いてサブセグメント同士を連結する。インターフェースでは、2 つのモデルビューのいずれかにスクリブルを追加することが出来る。

ブラシ選択 ユーザは、セグメント ID を表すブラシの色を選択し、スクリブルを付与する。ブラシ選択による分類は、各セグメントに対してレイヤーを作成するよりも直感的で効率的であると考えている。

Undo & Redo ボタン ツールには「やり直す（Undo）」と「元に戻す（Redo）」ボタンが含まれており、ユーザはスクリブルの仮追加・必要に応じたスクリブルの削除が可能である。

4 サブセグメンテーションと実行時処理

図 4 に提案の前処理の全体像を示す。初めに提案手法では、3D スキャンモデルの UV テクスチャ画

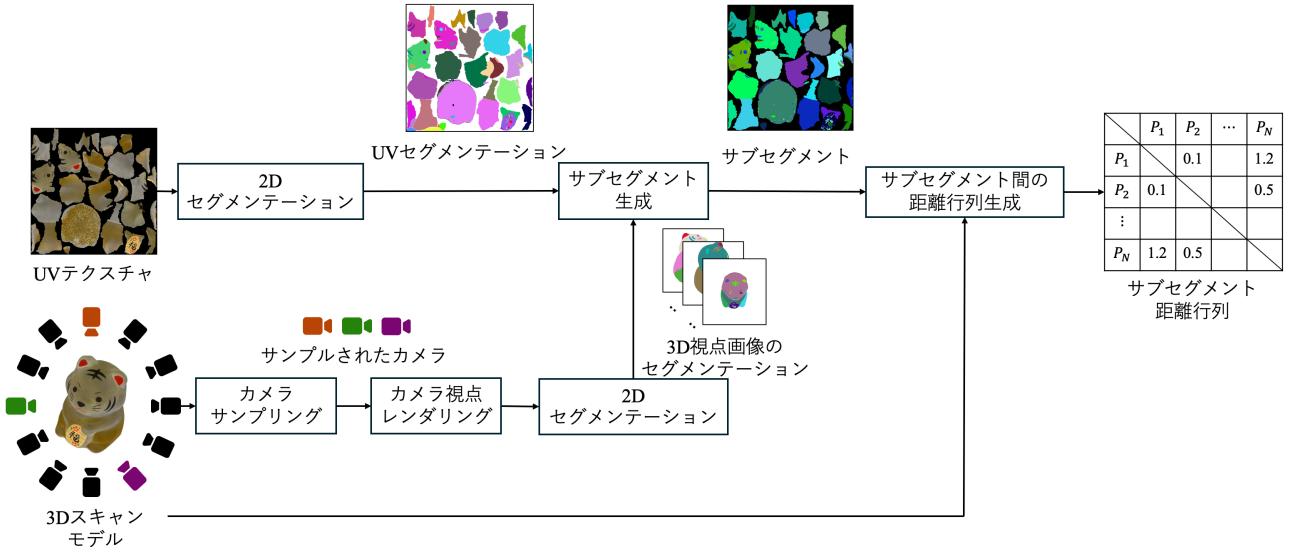


図 4. 前処理のステップは、主に 2 つのフェーズで構成されている。サブセグメント生成とサブセグメント間の距離行列生成である。サブセグメント生成には、3D モデルと対応する UV テクスチャを用いる。最初に、3D スキャンモデル全体を包含するカメラ群をサンプリングする。次に、3D 視点画像と UV テクスチャの両方に対して 2D セグメンテーションを行う。その後、2 つのセグメンテーションに基づいてサブセグメントを生成する。その後、生成したサブセグメントを用いてサブセグメント間の距離行列を計算する。この距離行列には、全てのサブセグメント間の距離情報が保持されており、ユーザーによるインタラクティブセグメンテーションの際に用いる。

像と特定の視点からのレンダリング画像から HQ-SAM [4] を用いて 2D セグメンテーション結果を得る。特定視点からのレンダリングの際に、3D スキャンモデルの周囲を包含するようにカメラの位置を最適化することで、3D スキャンモデル全体のセグメンテーション結果を取得している。上記 2 つのセグメンテーション結果の境界線を取得し、一つの UV テクスチャ上にまとめる。その後まとめた UV テクスチャに対して連結成分解析を計算する事で、サブセグメントを獲得する。その後隣接するサブセグメントの 3D 上での幾何的距離と色情報の 2 つを掛け合わせた値を計算する。その後、ダイクストラ法を用いて全てのサブセグメント間のペアに対する距離を計算し、計算結果をサブセグメント距離行列として生成する。

図 5 に実行時処理を示す。ユーザが提案インターフェースによりセグメンテーションを行う際には、サブセグメントに対して描画線を付与していく。付与された線から付与されてないサブセグメントの距離を事前計算した距離行列から計算することで、描画線が付与されてない箇所のセグメントを自動決定する。具体的には複数の ID (色) の描画線からの距離を計算し、最も距離が短い ID をサブセグメントに割り当てる。

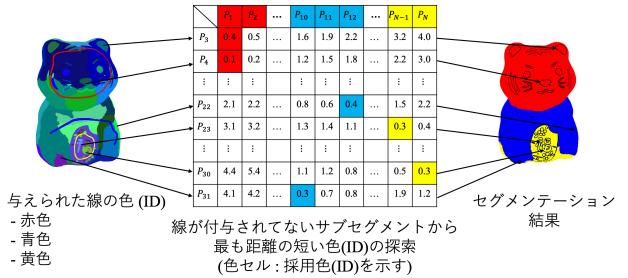


図 5. 実行時処理は、ユーザのスクリブルに基づいて、サブセグメント距離行列から最も近いユーザのスクリブルを確認し、スクリブルがないサブセグメントにセグメンテーション ID を割り当てる。その後、最終的に全体のセグメンテーション結果を得る。

5 ユーザ実験

ユーザインターフェースを用いたセグメンテーション精度の評価を目的として、3D スキャンモデルのユーザ実験を実施した。参加者のパフォーマンスと各インターフェースに対する体験の比較のため、被験者内デザインを採用した。

5.1 実験設定

5.1.1 参加者

当機関から 8 名の参加者を募集した。参加の前提条件として、ゲームや CAD を含む 3D コンピュ

サブセグメントに基づく 3D インタラクティブセグメンテーションシステム

タグラフィックスの経験が必要である。前提条件が満たされていることを事前調査により確認した。

5.1.2 ベースライン

本研究のベースラインとして、3D グラフィックスソフトウェアの Blender を利用した。このインターフェースは、セグメンテーションの際に UV テクスチャと 3D メッシュを同時に表示および操作できる有用なソフトウェアである。また、Blender を用いたセグメンテーションでは、主にブラシツールやバケツツールが用いる。

5.1.3 対象 3D スキャンモデル

本実験では、3D スキャナで作成されたブロッコリーと水鉄砲の 3D スキャンモデルを利用した。ブロッコリーモデルは花蕾と茎から成る。モデルの構造は一見単純に見えるが、花蕾と茎の境界は色境界が曖昧であり識別しづらい。さらに、茎の裏側にブラシを入れる際にカメラや 3D スキャンモデルの位置を調整する必要があり、作業コストが大きい。水鉄砲モデルは、比較的明確な色分けがされているが、水タンク、引き金、ノズル周りなどのアクセスが難しい領域がいくつか存在する。上記のモデル選定は、2 つの異なる性質を持つモデルのセグメンテーションを通して、提案技術の性能を確認する事が目的である。

5.1.4 手順

各参加者に対して、合計 1 時間の実験時間を設けた。セッションの開始時に、参加者に実験手順について説明し、同意書への署名を依頼する。提案されたインターフェースおよびベースラインインターフェースの操作方法についてを行った後、練習用モデルを使用して、5~10 分間の両方のインターフェースでトレーニングを行い、ツール利用への理解を促した。

本実験では、参加者にブロッコリーと水鉄砲のモデルを用いたセグメンテーションタスクをそれぞれのインターフェース上で行うよう指示を与えた。すなわち、参加者は合計で 4 つのセグメンテーションタスクを実施することになる。順序効果を減少させるために、インターフェースの順序は参加者ごとに入れ替えた。

各タスクには 5 分の時間制限が設けられた。ただし、参加者がセグメンテーションに満足した場合は、早期終了も可能である。タスクの目的は、正解データにできるだけ近いセグメンテーションを達成することである。ルールとして参加者は、(1) 実験中にセグメンテーション正解図の参照が可能、(2) 基本的な操作についての質問が可能、(3) セグメンテーションツールのより効率的な使い方に関する質問はなし、とした。また参加者は、ペイントイングに使

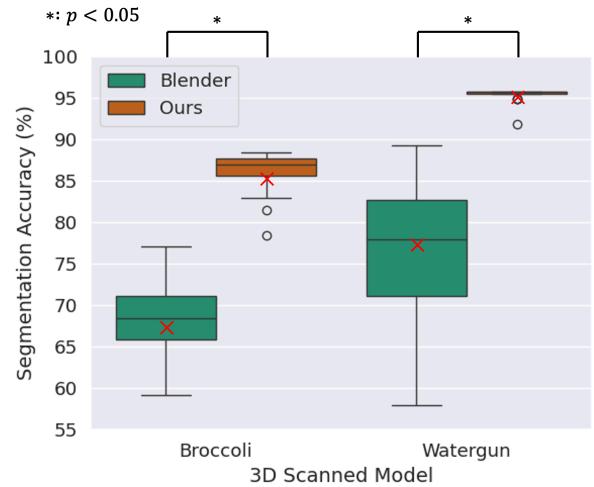


図 6. ボックスプロットの結果は、提案手法はほぼすべてのケースでセグメンテーション精度を改善した事を示している。×は各平均値を示す。

用する色を自由に選択でき、色の違いは最終結果の評価に考慮しないものとした。

5.1.5 評価指標

参加者が作成したセグメンテーションマップの精度を評価するために、正解データとのセグメンテーション結果の UV テクスチャの一一致率 IoU を計算することで、最終精度を計算した。

5.2 結果

セグメンテーション精度 図 6 に、ユーザが生成したセグメンテーションの精度をまとめたボックスプロットを示す。平均して、提案インターフェースはブロッコリーの 3D スキャンモデルに対して 85.40% \pm 3.29 の精度を達成し、ベースラインの平均精度 67.71% \pm 5.90 を上回っている。また水鉄砲の 3D スキャンモデルでは、提案インターフェースが平均精度 95.58% \pm 0.28 を示し、ベースラインの 77.27% \pm 10.19 を上回った。これらの結果は、提案インターフェースが両方のモデルにおいて平均精度でベースラインを一貫して上回っていることを示している。これらの違いの統計的有意性を、結果の非正規分布を示唆する Shapiro-Wilk 検定の後に実施した Mann-Whitney U 検定によって確認した。その結果 Broccoli モデルで $p < 0.05$ 、Watergun モデルで $p < 0.05$ と記録され、統計的有意性を確認した (p 値は Benjamini-Hochberg 手法により補正済み)。

図 7 に、ユーザ実験の中で最も精度の高いセグメンテーション結果と自動 3D セグメンテーションモデルである SAM3D [11] の参考結果を示す。図 7 の黒いボックスを詳しく見ると、提案手法のセグメ

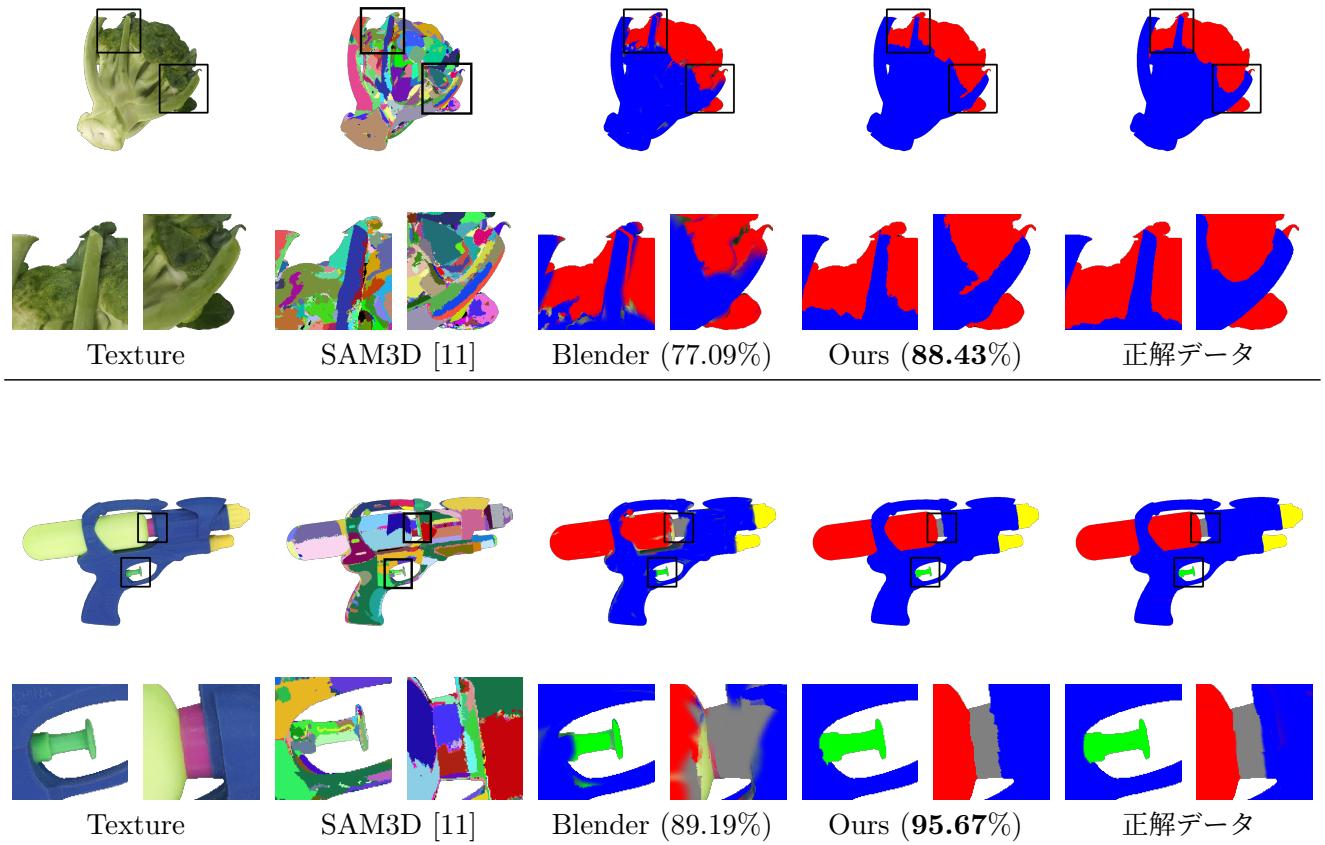


図 7. ユーザー研究におけるセグメンテーション結果を示す。各条件において最も高いセグメンテーション精度を示すモデルを可視化する。Blender に対する改善を黒いボックスで強調する。

ンテーションがテクスチャに沿って分離されており、ベースラインに比べても正解セグメンテーションに対するエラード合が小さいことを確認した。この結果は、3Dスキャンモデルの塗りやすさがインターフェースのパフォーマンスに大きく影響することを強調しており、半自動セグメンテーションの重要性を示唆している。

ユーザの主観評価 ユーザの好みを評価するためにベースラインと提案手法のどちらのインターフェースを好みかを5段階で質問した：ベースラインを強く好む、ベースラインを好む、どちらでもない、提案インターフェースを好む、提案インターフェースを強く好む。その結果、7人の参加者が提案のインターフェースを提案インターフェースを強く好むと評価し、1人の参加者が提案インターフェースを好むと評価した。参加者は提案インターフェースの方が、より少ないユーザ介入でセグメンテーションを迅速に行なうことが可能だと評価した。一部のBlenderに慣れた参加者はベースラインの方が直感的であると評価したが、提案手法が有効であることを認めた。

完了時間 Broccoli 3D スキャンモデルの場合、ユー

ザは両方のインターフェースにて与えられた5分の経過後に作業を打ち切った。Watergun 3D スキャンモデルでは、ユーザは提案インターフェースを使用して平均 247 秒で完了し、ベースラインでは与えられた5分の経過後に作業を打ち切った。

6 まとめ

本研究では、3Dスキャンモデルに対する半自動3Dインタラクティブセグメンテーション手法を提案した。提案手法では、はじめに自動前処理フェーズで小さなサブセグメントを生成し、ユーザが3Dスキャンモデルビュー上にスクリブルを描くことで、サブセグメント群をセグメント毎にクラスタリングすることを可能にした。実験では提案手法が従来の手動セグメンテーション技術に比べて、精度と品質向上に貢献する事を確認した。将来研究では、セグメンテーションの高精度化に向けたセグメンテーションパッチ生成過程やインタラクティブインターフェースの改善を行う予定である。

参考文献

サブセグメントに基づく 3D インタラクティブセグメンテーションシステム

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.
- [2] D. Freedman and T. Zhang. Interactive graph cut based segmentation with shape priors. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, pp. 755–762 vol. 1, 2005.
- [3] H. Guo, H. Zhu, S. Peng, Y. Wang, Y. Shen, R. Hu, and X. Zhou. SAM-guided Graph Cut for 3D Instance Segmentation. *arXiv preprint arXiv:2312.08372*, 2023.
- [4] L. Ke, M. Ye, M. Danelljan, Y. Liu, Y.-W. Tai, C.-K. Tang, and F. Yu. Segment Anything in High Quality. In *NeurIPS*, 2023.
- [5] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, 42(4), 2023.
- [6] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, et al. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023.
- [7] T. Kontogianni, E. Celikkan, S. Tang, and K. Schindler. Interactive object segmentation in 3d point clouds. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2891–2897. IEEE, 2023.
- [8] M. Landy and J. A. Movshon. *The Plenoptic Function and the Elements of Early Vision*, pp. 3–20. 1991.
- [9] J. MacQueen, et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1, pp. 281–297. Oakland, CA, USA, 1967.
- [10] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.
- [11] Y. Yang, X. Wu, T. He, H. Zhao, and X. Liu. Sam3d: Segment anything in 3d scenes. *arXiv preprint arXiv:2306.03908*, 2023.

PP-Checker: プログラミング教育における大規模言語モデルと協調した曖昧性のある自動採点システム

関口 祐豊 * 中村 聰史 *

概要. 本論文では、プログラミング教育における採点業務の効率化を目指し、即時性、採点精度、曖昧性に焦点を当てた自動採点システム PP-Checker を提案する。PP-Checker は、大規模言語モデル（LLM）との協調により、動的かつ視覚的なプログラミング言語の自動採点を実現する。また、リアルタイムでプロンプトを調整できる機能を備えており、学生が課題に取り組んでいる間にも採点基準を更新し、結果をすぐに反映できる。実際の講義で計 2,400 分運用した結果から、PP-Checker は課題の再提出時間を平均 3.6 分まで短縮し、課題採点業務の効率化に貢献することが示された。

1 はじめに

大学における必修プログラミング教育は、数十人から数百人規模の学生を一度に指導することが求められる [43, 29]。こうした大規模な教育現場において、大学院生が TA (Teaching Assistant) として教育補助業務を行っている [41]。例えば、我々の所属する明治大学総合数理学部先端メディアサイエンス学科では、毎年 10 名程度の TA と 4 名の教員が協力し、120 名以上の学生が受講する必修のプログラミング演習を担当しており、講義や課題提示、質問対応や課題採点などを行っている。こうしたプログラミング教育において学生・TA・教員の人数比に偏りがある場合、適切なコミュニケーションが難しく、講義運営が円滑に進まないことがある。我々の学科では、プログラムタイプイングシステム [42] を開発・運用しており、初学者の基礎力向上には寄与しているが、講義運営の改善には十分でない。

プログラミング演習講義では、学生は提示された課題に取り組むことが一般的である。ここで、課題の迅速な採点が学生の学習意欲の向上につながることが知られている [6, 39]。しかし、TA や教員の人数に対して学生の人数が多い場合、隨時発生する質問対応を行いつつ課題の採点を行うのは容易ではなく、全学生の課題を迅速に採点できない [4]。

自動採点は課題採点業務を効率化する方法の一つであり、フィードバックの迅速化だけでなく、学生の動機付けやスキル向上にも寄与することが知られている [8]。実際、自動採点システムとして、定義されたテストケースや短冊型問題 [30] に対応する手法 [17, 22, 31, 37] や、構文木を用いた自動採点および理解度評価の研究 [33, 36] が行われてきた。しかし、教員にとって課題の準備に加えてテストケースを用意することは、多くの時間と労力が必要とな

り、負担が大きい。また、Processing [23] のようなインタラクティブなプログラミング言語に対する採点では、テストケースを用意することが難しい。

ここで、大規模言語モデル（LLM）の急速な普及により、特にアルゴリズムやデータ構造の課題に対する自動採点において、様々なアプローチが試みられている [3, 19, 34]。しかし、Processing [23] などのインタラクティブなプログラミング言語においては、LLM がハルシネーション [11] を起こすリスクが高いことが指摘されており [26]、十分な精度を満たすためにはプロンプトの工夫が必要となる。

そこで本論文では、人間と LLM の協調によってインタラクティブなプログラミング言語に対応し、課題採点業務の効率化を目的とした曖昧性を許容する自動採点システムである PP-Checker (図 1) を提案する。本システムを 12 回にわたるプログラミング演習講義で実運用することで、その有用性と課題を明らかにする。

2 関連研究

宮下ら [32] の MOTIVATION や、Resnick ら [24] の Scratch、増井ら [40] が提唱した全世界プログラミングなど、プログラミング学習のモチベーションを高める研究は多く行われている。また、Greenberg ら [9] は、Processing [23] が Java と比較して学習者のモチベーションを高めることを明らかにした。Salga ら [25] は Processing.js を提案し、教育やアートの分野での広範な利用可能性を示した。Terroso ら [27] は、非プログラマ向けに p5.js [1] を用いた授業において、インタラクティブなプログラムが学生の関心や評価を高めることを明らかにした。第二著者が実施するプログラミング演習の講義においても、学生の関心を高めるため、錯視を取り入れた視覚的なプログラムの作成や、統計量の可視化、インタラクティブなゲームの作成といったように課題へ

Copyright is held by the author(s).

* 明治大学

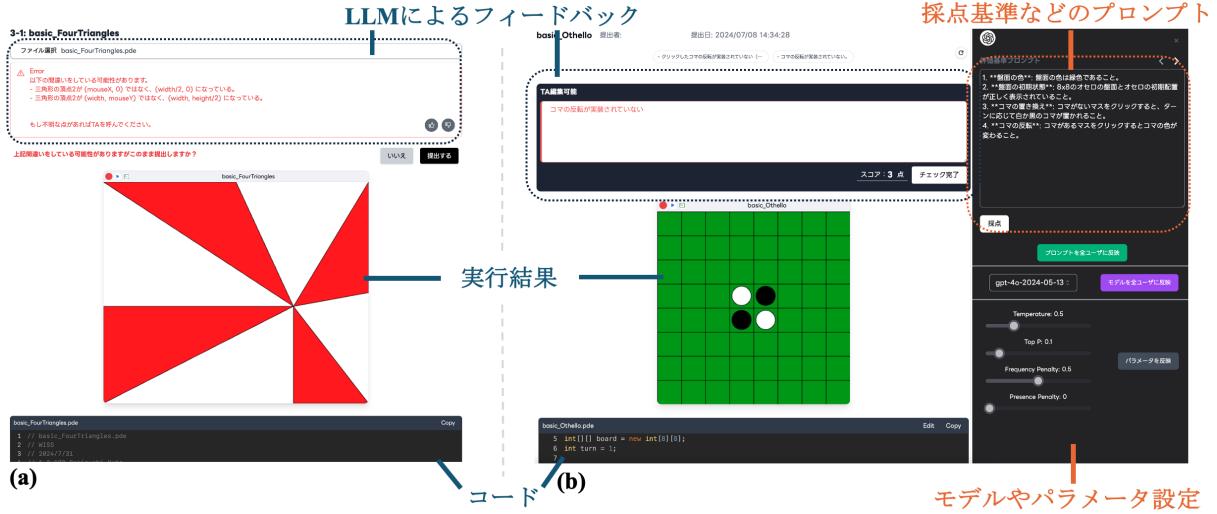


図 1. PP-Checker の 2 つの主要な画面. LLM による自動採点機能を備えた課題提出画面 (a) , 手動チェック画面 (b) .

のモチベーションを高める工夫を行っている¹.

LLM を活用したプログラミング支援も活発に行われている. Kazemitabaar ら [15] は, AI コードジェネレータである CodeX [5] が初学者のプログラミング学習においてコードの完成率やスコアを向上させることを示したが, 同時にツール依存を避け, スキル向上のための LLM の適切な使用が重要であると指摘していた. Jonsson ら [13] は, AI がプログラミング教育に有効である一方, 予測不可能な振る舞いが学習体験に与える影響に注意が必要であることを示唆した. Abolnejadian ら [2] は, GPT[21, 20] を用いた教材カスタマイズが学生の理解度を向上させることを示しつつ, 誤情報の訂正には教師の役割が不可欠であるとした. Kabir ら [14] も同様に, LLM の包括性と明瞭さを評価しつつ, 人間と AI の協力の重要性を強調した. また, LLM によるフィードバックは, 従来のテストベースのフィードバックを補完する形で効果的であることが示されている [7]. さらに, CodeAid [16], KOGI [35], ChotGPT [38] のように LLM を用いたエラー解決方法の研究も進展している.

これらの研究を踏まえ, PP-Checker は人間と LLM の協調による自動採点を実現し, 課題採点業務の負担を軽減しつつ, 迅速かつ適切なフィードバックを提供することを目指している. また, インタラクティブなプログラミング言語 Processing を使用した講義において, 実運用可能な採点システムとして, プログラミング教育の円滑化を図る.

3 提案手法

学生や TA, 教員の人数比の偏りにより, TA や教員が全学生の質問対応をしつつ課題採点を行うこ

とは負荷が高い. また, 学生は迅速なフィードバックを求めており [6, 39], こうした即時性のニーズに応えることがプログラミング教育の質向上において重要な要素である. 一方, Processing [23] などのインタラクティブなプログラミング言語は動的要素を含むことから, 従来の採点手法では正確な評価が困難であるため, 採点精度および曇昧性を考慮することが重要である.

そこで, システムの実現にあたり, 即時性, 採点精度, 曇昧性の観点に着目した仕組みについて述べる.

3.1 フィードバックの即時性を目指した仕組み

従来の手動採点ではフィードバックが遅れることが多く, 学生の学習意欲や効果に悪影響を及ぼす可能性がある. そこで, LLM を活用して自動採点を行うことにより, 提出されたコードの誤りを即座に指摘することを可能とする. この仕組みにより, 学生は TA の採点を待たずに自身のコードの潜在的な問題点を早期に発見し, 再提出することが可能となる.

3.2 高い採点精度を目指した仕組み

先述の通り, 自動採点では複雑なテストケースを作成し, 提出されたコードが期待通り動作するかを評価する必要がある. しかし, 視覚的・インタラクティブなプログラム課題においてはテストケースの準備に多くの時間と労力がかかり, 柔軟性にも限界があるうえ, 様々な記述形式が想定されるため, 事前にそのすべてを予測することは困難である.

そこで, LLM とプロンプト²を用いた採点手法を導入することで, テストケースの作成を不要とする. しかし, LLM の出力は完全ではなく, 採点精度向上にはプロンプト設計が重要である [44] ため, TA

¹ <https://lecture.nkmr.io>

² <https://github.com/YutoSekiguchi/ExPrompt>

PP-Checker: プログラミング教育における大規模言語モデルと協調した曖昧性のある自動採点システム

課題名	説明	提出期限	状況
basic_ParaPara	5枚以上の画像を用意し、クリックするたびに1枚ずつ画像が切り替わるバラバラ画像のようなアニメーションを表示するプログラム。	2024/07/15 16:45 期限切れ	
basic_BoundSound	Basic: 13-2 (PP-Checker 上では音はない場合があります) 背景画像がセットされた画面内で2つのキャラクターが移動し、壁に当たると跳ね返り、効果音が鳴るプログラム。	2024/07/15 16:45 期限切れ	
basic_CalcStdDev	Basic: 13-3 配列を引数として、その標準偏差を計算し、結果を表示するプログラム。	2024/07/15 16:45 期限切れ	
advanced_Matrix	Advanced: 12-1 3x3の行列を表示し、行列式、対角和、ヨーリッドノルムを計算するプログラム。	2024/07/15 13:30 期限切れ	
advanced_LifeGame	Advanced: 12-2 進生、生存、進路、消滅のルールに基づいてセルの状態が変化するライフゲームを実装するプログラム。	2024/07/15 13:30 期限切れ	

図 2. 課題一覧画面

や教員が LLM のプロンプトをリアルタイムに変更できる機能も導入する。具体的には、採点中にプロンプトを修正すると、更新されたプロンプトに基づいた採点が未採点の課題に即座に反映される仕組みを構築する。これにより教員は、講義開始前に完璧なプロンプトを準備する必要がなく、採点精度を講義中にリアルタイムで調整することが可能になる。

3.3 曖昧な判定を許容する仕組み

LLM による自動採点は当然判定ミスも多くなると考えられる。そこで、学生に対して LLM の指摘を受け入れるか否かを選択する機会を提供する。これにより、学生は LLM の指摘を参考にしつつ、自らの判断を信頼して学習を進めることができる。このアプローチは、LLM などシステムへの過度な依存を避け、自立性と判断力の強化にもつながることが期待される。

さらに LLM を用いた指摘において、具体的な正答を示してしまうと学生の問題解決能力を高めることができない。そこで LLM の出力結果を工夫することで、学生が自ら考えを深め、問題解決能力を高めること [18] を目指す。

4 PP-Checker

提案手法を講義において実運用することを目指し、課題提出、実行、採点、コード管理を一元化した Web アプリケーションとして実装を行った。本システムにより TA や教員は複数のツールを使い分けることなく、一つのプラットフォーム上で全ての課題採点業務を行うことができ、業務の効率化が期待される。PP-Checker は、課題一覧画面、課題提出（自動採点）画面、提出物一覧画面、手動チェック画面の 4 つの主要な画面から構成されている。

4.1 課題一覧画面

課題一覧画面（図 2）は、PP-Checker のホーム画面であり、学生は各課題の進捗状況、提出状況、

提出者	課題名	TA	状況
basic_MouseTrace	basic_MouseTrace	TAR	2024/07/01 13:46:38
basic_MouseTrace	basic_MouseTrace	TAR	2024/07/01 13:47:10
basic_MouseTrace	basic_MouseTrace	x	2024/07/01 14:10:18
basic_LineBoard	basic_LineBoard	x	2024/07/01 14:17:30
basic_MouseTrace	basic_MouseTrace	TAR	2024/07/01 14:17:52
basic_LineBoard	basic_LineBoard	x	2024/07/01 14:19:07
basic_LineBoard	basic_LineBoard	TAR	2024/07/01 14:20:05
basic_LineBoard	basic_LineBoard	TAR	2024/07/01 14:20:38
basic_LineBoard	basic_LineBoard	TAR	2024/07/01 14:20:44
basic_Goichi100	basic_Goichi100	TAR	2024/07/01 14:20:49
basic_LineBoard	basic_LineBoard	TAR	2024/07/01 14:21:12
basic_MouseTrace	basic_MouseTrace	TAR	2024/07/01 14:21:29
basic_LineBoard	basic_LineBoard	TAR	2024/07/01 14:21:45
basic_MouseTrace	basic_MouseTrace	TAR	2024/07/01 14:22:40
basic_LineBoard	basic_LineBoard	TAR	2024/07/01 14:22:41

図 3. 提出物一覧画面

フィードバックの有無、過去に提出したコードを確認できる。各課題項目をクリックすることで課題提出画面に移動し、コードの提出を行うことができる。これにより、学生は課題管理が容易になり、学習の効率が向上することが期待される。

4.2 課題提出（自動採点）画面（学生用）

課題提出画面（図 1a）では、学生が自身のプログラムコードを提出し、LLM が採点を行う。具体的には、学生はコードを含むフォルダをアップロードし、提出前にその場で LLM による自動採点を受けることができる。LLM はアップロードされたコードを即座に採点し、誤りの可能性がある箇所を指摘するフィードバックを生成する。このフィードバックをもとに学生は自身のコードの潜在的な問題点を TA に提出する前に把握し、コードの修正や再提出が可能となる。また、この画面には提出されたコードや実行結果も表示され、学生は自分のコードがどのように動作するかを改めて確認できる。

4.3 提出物一覧画面

提出物一覧画面（図 3）は、TA や教員が提出された課題を一覧で確認できる。この画面では、各課題の提出状況、採点状況、フィードバックの有無、どの TA がどの課題をチェックしているかがリアルタイムで表示される。各提出物の課題名をクリックすることで、手動チェック画面に移動し、LLM が行った初期採点を TA や教員が最終確認できる。

4.4 手動チェック画面（TA・教員用）

手動チェック画面（図 1b）は、TA や教員が LLM による採点結果を確認し、必要に応じて補足や修正を行うことができる。画面左側には、LLM が生成したフィードバックを確認・修正できるインターフェースがあり、その下にコードや実行結果も表示されるため、TA は学生のコードの動作を確認しながら課題採点を行うことができる。また、画面上部には、

他の提出物に対するフィードバックコメントを選択できるインターフェースが設置されており、これを再利用することで、入力の手間を省き、効率的な課題採点業務をサポートする。画面右側では、LLM のプロンプトやモデルの変更、パラメータの調整を行うことができ、リアルタイムで採点精度を高めることができ。また、プロンプトは TA や教員が容易に変更できるように、課題の満たすべき条件を箇条書きで記述できる形式にしている。

4.5 実装

PP-Checker は、フロントエンドに Next.js と TypeScript、バックエンドに FastAPI、データベースに MySQL を用いて実装した。LLM は OpenAI 社の gpt-3.5-turbo, gpt-4-turbo-preview, gpt-4-turbo, gpt-4o を選択できるような設計にしている。また、リアルタイム通信には Socket.io を利用しており、学生と TA の間での迅速なやり取りを可能にしている。

Processing [23] などのインタラクティブな言語は、動的な要素を含むため静的なコードよりも複雑な評価が求められる。そこで、Processing.js [25] を用いて実行画面を描画することで、学生が提出したコードの自動採点と視覚的な確認を可能にするとともに、マウスやキーボードのクリックや画像、学生ごとに異なるデザインを含む課題でも正確なフィードバックを提供する。Processing.js で対応していない circle や enum などの関数については、自作の関数を作成し対応している。なお、学生や TA を識別するため、大学発行のメールアドレスによる認証を用いて、ログインを行なうようにしている。

5 運用と分析

5.1 運用形態

PP-Checker を、明治大学総合数理学部先端メディアサイエンス学科 1 年次対象の必修科目であるプログラミング演習 I (100 分 2 コマ) において、2024 年 4 月 15 日から 7 月 22 日までの 12 回分の講義 (計 2400 分) で運用した。履修者は学部 1 年生と再履修者の 123 名で、TA は筆頭著者を含む大学院生 10 名、教員は第二著者を含む 4 名であった。

講義では、教員が冒頭の数十分間でスライドを用いて説明を行い、その後、4~5 つの課題が提示される。課題は難易度に応じて 2 種類に分かれ、基本課題は講義時間内に、難易度の高い発展課題は、次回授業開始時までに提出する必要がある。これらの課題の提出先を PP-Checker にすることで運用を行った。

5.2 運用結果と分析

運用期間中、基本課題 34 問と発展課題 17 問の計 51 問の課題を実施した。本研究では、授業内の採点業務の効率化を目的としているため、分析対

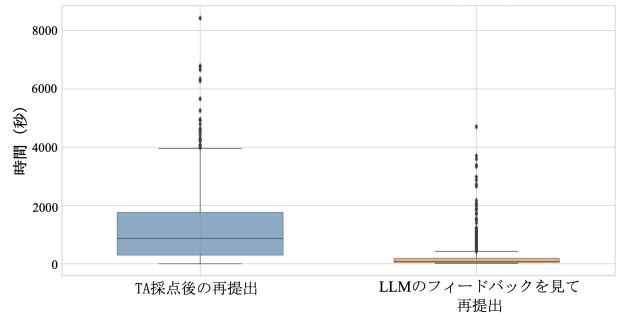


図 4. 再提出までの時間の比較

象を基本課題に限定する。基本課題において、PP-Checker を通じて合計 6,415 回の提出が行われた。

PP-Checker の導入による迅速なフィードバックの効果を定量的に評価するため、学生がフィードバックを受けてから再提出までの時間的変化や提出回数について分析を行った。分析の結果、6,415 回の提出のうち、1,491 回 (約 23.2%) は TA 採点前に LLM によるフィードバックを受けた時点で、学生が自主的に課題を取り下げてコードの改善を行っていることが明らかになった。図 4 に、再提出までの時間の分布を箱ひげ図で示す。TA 採点後に再提出された件数は 634 件であったのに対し、LLM によるフィードバックを受けて修正し TA に再提出された件数は 886 件であり、学生が早期フィードバックを活用していることが明確となった。再提出までの平均時間は、TA 採点後の再提出が約 20.3 分であったのに対し、TA 採点前に LLM のフィードバックを受けて修正した場合は約 3.6 分であり、再提出のスピードが著しく向上していることがわかった。

プロンプト変更機能の使用状況について、収集を開始した第 5 回講義以降の基本課題 24 問について分析を行った結果、合計 82 回のプロンプト変更が観察された。プロンプト変更前後の LLM の採点正解率を比較したところ、変更前が 60.4% (標準偏差 : 23.7)、変更後は 60.7% (標準偏差 : 17.8) とわずかな向上が見られたものの、有意な差は確認されなかった。全体的な正解率に有意な差は見られなかつた一方で、課題ごとの分析では、プロンプト変更の効果に顕著な変動があることが確認された。具体的には、初期精度が 60%未満のプロンプトに変更を加えた場合、90%近くのケースで精度が向上することが確認された。最も改善した課題では、初期精度 33.3% のプロンプトが変更後に 72.9% まで向上し、約 40% の精度向上が確認された。一方、初期精度が 80%以上あった課題では、プロンプト変更後にその精度が低下するケースが大半であった。

5.3 SUS と学生アンケートの結果

114 名の学生を対象に、System Usability Scale (SUS) [12] を用いた PP-Checker のユーザビリティ

表 1. TA アンケート結果

質問項目	評価値の分布					平均
	-2	-1	0	1	2	
PP-Checker は全体的に使いやすいと思いましたか？	0	0	0	2	7	1.78
PP-Checker をこれからも利用したいと思いますか？	0	0	0	1	8	1.89
PP-Checker 導入前後で、課題採点業務の作業効率は向上しましたか？	0	0	0	0	6	2.00

と信頼性を評価するアンケート調査を実施した。

調査の結果、PP-Checker の SUS スコアは 76.4 となった。これは平均スコア (68.0) を大きく上回っており、システムのユーザビリティが高く評価された。しかし、標準偏差は 12.2 であり学生間の評価には一定のばらつきが見られた。

PP-Checker の利用に関するアンケートの分析では、授業の進行に伴い、学生が LLM のフィードバックからコードの誤りに気づいたり、修正のヒントを得たりする経験が徐々に減少する傾向にあった。しかし、学期後半においても 57%以上の学生が、フィードバックの半分以上が有用だと回答していた。参考になったフィードバックには、「円の動作が毎フレーム X 方向に 3 ピクセル、Y 方向に 2 ピクセル動いていない (Y 方向に 2 倍動いている)。」や「A さんのダイスの範囲が 1 から 6 ではなく、1 から 5 になっている。」といった具体的な誤りの箇所を含むフィードバックがあげられた（図 5）。クリックするたびにじゃんけんの結果を標準出力する課題では、「judgeJanken 関数内の条件で、パーとグーの勝敗判定が間違っている。」といった、複数回操作しないと気づきにくい誤りを指摘するフィードバックも有用と評価された。さらに、「枠線が削除されていない」といった、一見ただけでは見落としがちな問題を指摘するフィードバックも高く評価された。

一方で、参考にならないと評価されたのは、正確性に欠けるフィードバックや課題を解く上でプログラムの挙動に影響しない指摘であった。例えば、「標準偏差の計算において、平均値を毎回再計算しているため、効率が悪い。」といった課題の評価に影響しない指摘は、参考にならなかったと評価された。

5.4 TA アンケートの結果

PP-Checker の運用終了後、筆頭著者を除く 9 名の TA を対象にアンケート調査を実施した。アンケートは、PP-Checker の使用感や作業効率に関する定量評価項目（表 1）と、質的データ収集のためのインタビュー項目で構成した。定量評価は 5 段階のリッカート尺度 (-2: 全くそう思わない～2: 非常にそう思う) を用いた。インタビューでは、LLM との協調や PP-Checker 導入による TA 業務の変化に焦点を当てた。PP-Checker 導入前後の比較は、導入前の TA 経験を持つ 6 名のみを対象とした。

表 1 より、PP-Checker の使用感や作業効率に関

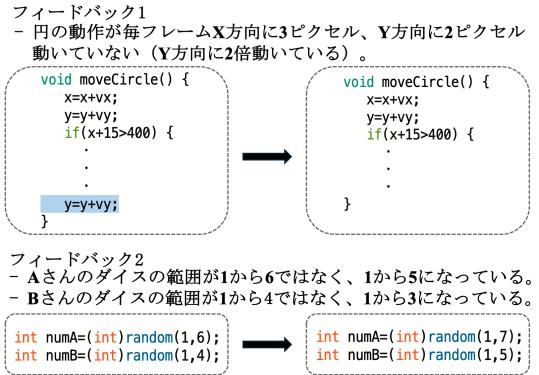


図 5. 参考になった LLM のフィードバックと修正の例

する全ての項目において高い評価を得ることができることがわかった。インタビュー項目では、「採点作業の効率化により、採点にかけていた時間を質問対応に回せるようになった」や「GPT のフィードバックがあることで基礎的な質問が減少した」、「プロンプトの変更により採点精度が向上した」といったポジティブな意見が多く得られた。

一方で、動作順序が重要な課題において精度にばらつきが多く見られ、課題内容による LLM の精度の差に関するネガティブな意見も一部見られた。

5.5 教員アンケートの結果

PP-Checker の運用終了後、第二著者を除く 2 名の教員を対象にアンケート調査を実施した。アンケートは、PP-Checker の利点や改善点、TA や学生に感じた変化などを調査するための項目で構成した。調査の結果、PP-Checker の利点として、LLM のフィードバックが学生自身の確認時に注目すべきポイントを明確となる点や LLM のフィードバックによって学生が誤りに気づいて修正する周期が早くなつたという意見が得られた。一方、段階的な減点条件を定めたい場合のプロンプト作成が難しかったという意見が得られた。

TA に対して感じた変化としては、以前は多かった氏名や課題名の記入のような誤りの指摘をしなくて良くなったことでストレスが減少したという意見が得られた。また、不備を指摘することが機械であることで、TA や教員側だけでなく学生側のストレスも下がっているという意見が得られた。しかし、自発的な興味が確立していない初学者に対する LLM

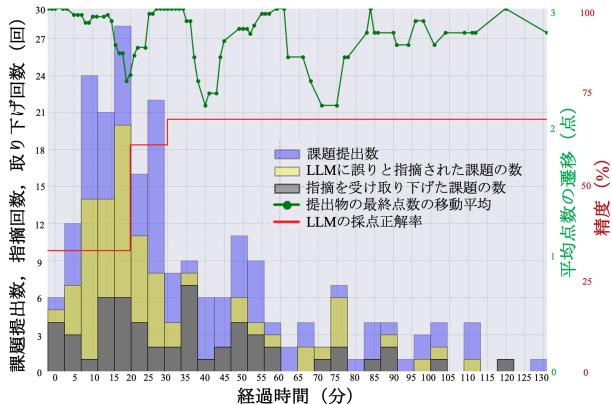


図 6. ある課題における提出数（青棒），LLM の指摘数（黄棒），課題取り下げ回数（灰棒），点数推移（緑），LLM フィードバックの正解率（赤）

の依存を危惧するような意見もあった。

PP-Checker を今後も運用したいかという質問では、2名ともに非常にそう思うと肯定的な回答をした。

6 考察と議論

6.1 学生に与える影響に関する考察と議論

PP-Checker を利用することで、再提出までの時間が大幅に短縮されたことから、即時性のあるフィードバックが学生の学習プロセスに大きな影響を与えることが示唆された。これは、学生が自身の誤りを素早く認識し、改善する機会を得ており、学習サイクルの加速につながっていると考えられる。LLMからのフィードバックを受けて自ら修正を行うプロセスは、単にプログラミングスキルの向上だけでなく、問題解決能力全般の改善にもつながると考えられる。さらに、SUS の結果から、システムへの満足度や有用性が高いことが示唆された。

図 6 は、ある基本課題における時間経過に伴う提出状況、採点精度、点数の推移を示している。この結果では最初の 20 分は精度が 30%程度と低いが、30 分までに 2 回プロンプトの修正がなされ 70%近くまで向上している。このプロンプトの修正は、最初の 20 分程度で提出された課題の自動採点およびその結果を無視して提出されたものにおける自動採点の結果を参考に行われている。プロンプト修正の主な例として、「**クリック時の再描画**: 画面をクリックするたびに a, b, c, d の値が 1~20 の間でランダムに決定され、曲線が再描画されていること。」という条件の太字部分のような未定義パラメータの具体化、出力形式の明確化、数式の具体的な指定といった変更が挙げられる。さらに、採点精度向上のため、明示的な例外規定（例：「1. 円が端付近に到達すると跳ね返るように実装されていない。円の半径分を考慮していない場合でも、課題の要件には明記

されていないため、減点対象にはしません。」）などのプロンプト変更が行われていた。これらの修正により、プロンプトの曖昧さを授業の進行に伴い軽減し、LLM の反復的な誤評価を抑制することで、教育効果の向上につながる可能性が考えられる。

プロンプトがまだ不十分な段階で精度が低い時に課題を提出する学生は基本的にプログラミング能力が高く、また自信があるため、精度が低くても大きな問題はない。一方で、中盤に課題を提出するプログラミングにおける自信があまりない学生は精度が低い場合にその影響を受けてしまう可能性がある。つまり、今回のように採点精度は時間とともに徐々に向上することが望ましく、今回の手法はこうした講義運営において適切であったと考えられる。

6.2 TA・教員に与える影響に関する考察と議論

TA へのアンケート結果から、PP-Checker の導入が課題採点業務の効率化につながることが示唆された。また、LLM による即時フィードバックにより、学生は初步的なミスを自ら修正し、より完成度の高い状態で提出を行うようになった。これにより、TA が同じような間違いを含む提出物を繰り返し確認する手間が減少し、学生の質問対応に多くの時間を割けるようになることが示唆された。一方で、プロンプトのリアルタイム変更に関する分析結果から、課題によって精度が大きく異なることや精度の高いプロンプトを変更することで精度が下がってしまう可能性があること確認された。そこで今後は、学生が LLM の判定を無視したときに、その情報をを利用してプロンプトの自動修正を行う仕組みを実現する予定である。

また教員へのアンケート結果から、PP-Checker が教員の期待に十分に応えていることが示唆された。しかし、初学者の学生が LLM に依存してしまうこと [28, 10, 15] への懸念を示した意見があったことから、今後はこのような不安を解消しつつ、LLM、学生、TA、教員がそれぞれの役割を最大限に発揮できるようなインターフェース設計を目指す。

7 まとめ

本論文では、プログラミング教育における課題採点業務の効率化を目指した曖昧性のある自動採点システム PP-Checker を提案し、その有用性を検証した。実運用の結果、PP-Checker は学生、TA、教員の全てから高評価を得ており、課題採点業務の効率化とフィードバックの迅速化に貢献した。特に、LLM による早期フィードバックが学生の再提出までの時間を大幅に短縮し、TA が学生の質問対応に時間を多く割けるようになった。

今後は、本システムの欠点を徹底的に洗い出し、その改善や新たな知見の深掘り、新たなプロンプト構築手法の検討を行っていきたいと考えている。

謝辞

本システムを講義で活用してくださった明治大学総合数理学部先端メディアサイエンス学科の12期生および、対応していただいたTA、教員のみなさまに感謝します。

参考文献

- [1] p5.js. <https://p5js.org/> (2024/8/31 確認).
- [2] M. Abolnejadian, S. Alipour, and K. Taeb. Leveraging ChatGPT for Adaptive Learning through Personalized Prompt-based Instruction: A CS1 Education Case Study. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA'24)*, (521):1–8, 2024.
- [3] R. Balse, B. Valaboju, S. Singhal, J. M. Wariem, and P. Prasad. Investigating the Potential of GPT-3 in Providing Feedback for Programming Assessments. *In Proceedings of the Conference on Innovation and Technology in Computer Science Education (ITiCSE'23)*, pp. 292–298, 2023.
- [4] D. Boud and E. Molloy. *Feedback in Higher and Professional Education*. Routledge, 2012.
- [5] M. Chen, J. Tworek, H. Jun, Q. Yuan, H. P. de Oliveira Pinto, J. Kaplan, H. Edwards, Y. Burda, N. Joseph, G. Brockman, A. Ray, R. Puri, G. Krueger, M. Petrov, H. Khlaaf, G. Sastry, P. Mishkin, B. Chan, S. Gray, N. Ryder, M. Pavlov, A. Power, L. Kaiser, M. Bavarian, C. Winter, P. Tillet, F. P. Such, D. Cummings, M. Plappert, F. Chantzis, E. Barnes, A. Herbert-Voss, W. H. Guss, A. Nichol, A. Paino, N. Tezak, J. Tang, I. Babuschkin, S. Balaji, S. Jain, W. Saunders, C. Hesse, A. N. Carr, J. Leike, J. Achiam, V. Misra, E. Morikawa, A. Radford, M. Knight, M. Brundage, M. Murati, K. Mayer, P. Welinder, B. McGrew, D. Amodei, S. McCandlish, I. Sutskever, and W. Zaremba. Evaluating Large Language Models Trained on Code. *arXiv:2107.03374*, 2021.
- [6] J. Clune, V. Ramamurthy, R. Martins, and U. A. Acar. Program equivalence for assisted grading of functional programs. *In the Proceedings of the ACM on Programming Languages*, 4(171):1–29, 2020.
- [7] H. Gabbay and A. Cohen. Combining LLM-Generated and Test-Based Feedback in a MOOC for Programming. *In Proceedings of the Eleventh ACM Conference on Learning @ Scale (L@S'24)*, pp. 177–187, 2024.
- [8] A. Gordillo. Effect of an Instructor-Centered Tool for Automatic Assessment of Programming Assignments on Students' Perceptions and Performance. *Sustainability*, 11(20):1–24, 2019.
- [9] I. Greenberg, D. Kumar, and D. Xu. Creative coding and visual portfolios for CS1. *In Proceedings of the 43rd ACM technical symposium on Computer Science Education (SIGCSE'24)*, pp. 958–964, 2024.
- [10] M. Jakesch, A. Bhat, D. Buschek, L. Zalmanson, and M. Naaman. Co-Writing with Opinionated Language Models Affects Users' Views. *In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI'23)*, (111):1–15, 2023.
- [11] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung. Survey of Hallucination in Natural Language Generation. *In ACM Computing Surveys*, 55(12):1–38, 2023.
- [12] B. John. SUS: A 'Quick and Dirty' Usability Scale. *Usability Evaluation In Industry*, p. 207–212, 1996.
- [13] M. Jonsson and J. Tholander. Cracking the code: Co-coding with AI in creative programming education. *In Proceedings of the 14th Conference on Creativity and Cognition (CC'22)*, pp. 5–14, 2022.
- [14] S. Kabir, D. N. Udo-Imeh, B. Kou, and T. Zhang. Is Stack Overflow Obsolete? An Empirical Study of the Characteristics of ChatGPT Answers to Stack Overflow Questions. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'24)*, pp. 1–17, 2024.
- [15] M. Kazemitabaar, J. Chow, C. K. T. Ma, B. J. Ericson, D. Weintrop, and T. Grossman. Studying the effect of AI Code Generators on Supporting Novice Learners in Introductory Programming. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'23)*, (455):1–23, 2023.
- [16] M. Kazemitabaar, R. Ye, X. Wang, A. Z. Henley, P. Denny, M. Craig, and T. Grossman. CodeAid: Evaluating a Classroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs. *In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI'24)*, (650):1–20, 2024.
- [17] S. Krusche and A. Seitz. ArTEMiS: An Automatic Assessment Management System for Interactive Learning. *In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE'18)*, pp. 284–289, 2018.
- [18] T. J. Ngoon, C. A. Fraser, A. S. Weingarten, M. Dontcheva, and S. Klemmer. Interactive Guidance Techniques for Improving Creative Feedback. *In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)*, (55):1–11, 2018.
- [19] H. Nguyen and V. Allan. Using GPT-4 to Provide Tiered, Formative Code Feedback. *In Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE'24)*, pp. 958–964, 2024.

- [20] OpenAI. GPT-4 Technical Report. *OpenAI Blog*, pp. 1–100, 2023.
- [21] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, and R. Lowe. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, pp. 1–15, 2022.
- [22] R. A. P. Queirós and J. P. Leal. PETCHA: a programming exercises teaching assistant. In *Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education (ITiCSE'12)*, pp. 192–197, 2012.
- [23] C. Reas and B. Fry. Processing: a learning environment for creating interactive Web graphics. In *ACM SIGGRAPH 2003 Web Graphics (SIGGRAPH'03)*, 2003.
- [24] M. Resnick, A. M.-H. John Harold Malone, N. Rusk, E. Eastmond, K. A. Brennan, A. Millner, E. Rosenbaum, J. S. Silver, B. S. Silverman, and Y. B. Kafai. Scratch: Programming for All. *Communications of ACM* 52, pp. 60–67, 2009.
- [25] A. Salga, D. Hodgson, A. Sobiepanek, S. Downe, M. Medel, and C. Leung. Processing.js: sketching with jcanvasjs. In *ACM SIGGRAPH 2011 Talks (SIGGRAPH '11)*, (15):1, 2011.
- [26] A. Singla. Evaluating ChatGPT and GPT-4 for Visual Programming. In *Proceedings of the 2023 ACM Conference on International Computing Education Research (ICER'23)*, 2:14–15, 2023.
- [27] T. Terroso and M. Pinto. Programming for Non-Programmers: An Approach Using Creative Coding in Higher Education. In *Third International Computer Programming Education Conference (ICPEC'22)*, 102(13):1–8, 2022.
- [28] X. Wang, H. Kim, S. Rahman, K. Mitra, and Z. Miao. Human-LLM Collaborative Annotation Through Effective Verification of LLM Labels. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI'24)*, (303):1–21, 2024.
- [29] L. Yan, N. McKeown, and C. Piech. The PyramidSnapshot Challenge: Understanding student process from visual output of programs. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE'19)*, pp. 119–125, 2019.
- [30] 角田博保, 久野靖. 短冊型問題: プログラミングの技能を評価可能な試験出題形式. 夏のプログラミング・シンポジウム 2016 「教育・学習」 報告集, 2016:73–80, 2016.
- [31] 久野靖. 短冊型問題を用いたプログラミング学習アドバイスツール. 情報処理学会 プログラミングシンポジウム 報告集, pp. 129–139, 2020.
- [32] 宮下芳明, 中橋雅弘. 学習者のモチベーション向上のための好意的解釈を行うフィジカルコンピューティング環境のデザイン. ヒューマンインタフェース学会論文誌, 13(4):303–313, 2011.
- [33] 漆原宏丞, 本多佑希, 岸本有生, 兼宗進. 抽象構文を利用したプログラミング理解度採点の試み. 情報処理学会 研究報告コンピュータと教育 (CE) , 2021-CE-162(16):1–6, 2021.
- [34] 若谷彰良, 前田利之. 生成 AI を用いた C 言語プログラミング学習のための助言システムの試作. 甲南大学紀要 知能情報学編, 16(2):7–16, 2024.
- [35] 小原有以, 佐藤美唯, 倉光君郎. KOGI: ChatGPT を Colab に統合したプログラミング演習支援. 情報処理学会 情報教育シンポジウム論文集, pp. 141–148, 2023.
- [36] 小川弘迪, 小林亜樹. プログラミング課題の自動採点に向けた構文木上のカーネル法による類似度関数の提案. 情報処理学会 第 80 回全国大会講演論文集, 2018(1):661–662, 2018.
- [37] 新田章太, 小西俊司, 竹内郁雄. 複数言語に対応しやすいオンラインプログラミング学習・試験システム track. 情報処理学会 情報教育シンポジウム論文集, pp. 114–121, 2019.
- [38] 森陽菜, 松澤芳昭. ChatGPT を利用したプログラミング教育支援システム「ChotGPT」の提案と評価. 情報処理学会 研究報告コンピュータと教育 (CE), 2024-CE-174(7):1–8, 2024.
- [39] 石原浩一, 泰山裕. フィードバックと振り返りが学習者の認知欲求に及ぼす影響の検討. 日本教育工学会論文誌, 44(1):105–113, 2020.
- [40] 増井俊之, 塚田浩二. 全世界プログラミング. 情報処理学会 プログラミングシンポジウム, pp. 1–8, 2006.
- [41] 文部科学省. 中央教育審議会 大学分科会 制度部会 (第 22 回 (第 3 期第 7 回)) 議事録・配付資料 [資料 2 - 1] ティーチング・アシスタンント (TA) について. https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/003/gijiroku/07011713/001/002.htm (2024/8/31 確認).
- [42] 又吉康綱, 中村聰史. typing.run: 初学者のプログラミング学習を支援するプログラムタイピングシステムの提案と実践. 情報処理学会 研究報告 ヒューマンコンピュータインターラクション (HCI) , 2020-HCI-189(1):1–8, 2020.
- [43] 又吉康綱, 中村聰史. askTA: 消極性を考慮したオンライン演習講義支援システム. コンピュータソフトウェア, 39(1):55–71, 2022.
- [44] 尹子旗, 王昊, 堀尾海斗, 河原大輔, 関根聰. プロンプトの丁寧さと大規模言語モデルの性能の関係検証. 言語処理学会 第 30 回年次大会, pp. 1–6, 2024.

著者索引

赤田 真由, 77
新井 美音, 122

五十嵐 健夫, 137, 146

伊勢 隆之介, 9

伊藤 貴之, 129

岩井 将行, 38

臼井 義人, 85

大塚 晟, 61

沖 真帆, 100

笠原 暢仁, 55

加藤 大晴, 146

加藤 淳, 46

栗原 一貴, 107

小池 英樹, 16

幸谷 有紗, 16

後藤 真孝, 46

崔 明根, 69, 77

志築 文太郎, 69, 77

SHAMIR Ariel, 137

白根 薫, 69

SECHAYK Yotam, 137

関口 祐豊, 153

高木 士, 146

高田 嶽介, 61, 77, 92

高橋 治輝, 85

田上 湖都, 129

武山 侑輝, 77

塚田 浩二, 9, 100

中野 倫靖, 46

中村 聰史, 153

永山 晃誠, 77

西岡 裕子, 122

西田 健志, 122

西出 新也, 122

二瓶 朝渉, 115

野間 直生, 100

橋本 泰輔, 146

濱崎 雅弘, 46

樋口 啓太, 146

平野 恵旺, 115

Fabrice Matulic, 146

深池 美玖, 55

福田 孝輔, 122

福地 健太郎, 23

福本 雅朗, 30

松浦 向日葵, 23

松村 耕平, 85

真殿 航輝, 146

宮下 芳明, 55

宮藤 詩緒, 16

森谷 美羽, 107

矢島 知子, 129

矢田 絵理奈, 38

山岸 真人, 92

山本 航世, 1

米沢 勝, 122

李 国政, 129

和田 優斗, 69

渡邊 研斗, 46

渡邊 恵太, 1, 115

日本ソフトウェア学会 研究会資料シリーズ ISSN 1341-870X
No. 102 2024年12月11日発行

第32回 インタラクティブシステムと
ソフトウェアに関するワークショップ (WISS 2024) 論文集