

WISS 2022 in ZAO

第30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)

WISS2022 開催にあたって

志築 文太郎*

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、2022年12月14日(水)から16日(金)にかけて、アクティブリゾート宮城蔵王(宮城県刈田郡蔵王町)とオンラインとのハイブリッドで開催される「第30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: WISS2022)での発表をまとめたものです。

インタラクティブシステムに関する新しい実装技術や面白い応用などに関して議論するワークショップ WISS は1993年にスタートし、今年で30回目という大きな節目を迎えることになりました。WISS は、伝統的に2泊3日の泊まり込み形式での開催を特徴のひとつとしてきており、毎年約100名~150名の参加者が朝から深夜まで深い議論を行ってきました。更に WISS は革新的な試みを多く取り入れてきました。実際に動くデモンストレーションを交えての議論を可能とするデモ発表を早くから取り入れ、加えて、発表と同時進行するチャットでの議論、会議を新しい形で支援する様々な「WISS Challenge」システム、深夜まで行われるナイトセッションがその例として挙げられます。これらを通じて、WISS は、いわゆる研究討論に留まらず、研究への楽しみや研究者同士の輪がリアルタイムに広まり深まることを体感できる場となっていました。このこともあって、

「WISS にもう一度行こう…」

こういった感想が WISS から帰って来た参加者から自然と聞かれるのだと思います。

しかしながら、WISS2020 は、Covid-19 のために多くの学会と同様に、完全オンラインでの開催を余儀なくされました (WISS2020 も完全オンライン開催ながら完全に WISS らしいイベントで大成功でした。自分もセッション中は沢山チャットで議論し、また夜遅くまで議論でき、めちゃくちゃ楽しかったです)。ただ、そのすぐ次の WISS2021 は、ザ浜名湖 (静岡県浜松市) とオンラインとのハイブリッドで開催され、大盛況を収めました。WISS2021 のこの成功も、泊まり込み形式のイベントをなんとか安全に、それでいてこれまでの WISS の良さを体験できるようにしようという WISS2021 運営委員の皆様

様の強い思いと並々ならぬ奮闘とによってもたらされたものと考えます。そして、その WISS2021 を継承し、今回の WISS もハイブリッドで開催することにしました。

今回の WISS2022 では発表形態として、前回は踏襲し、登壇発表 (Archival)、ロングティザー発表 (Non-archival)、デモ・ポスター発表 (Non-archival) を用意しました。

登壇発表については、それぞれの論文に対して4名の査読者 (うちメタ査読者1名) を割り当てて査読を行った結果、32件の投稿に対して16件の論文を採択しました (ロング採択4件、ショート採択12件)。なお、幾つかの論文については条件付き採録としました。これらについてはいずれも提出いただいた論文を本論文集に掲載しております。

また、ロングティザー発表については9件を、デモ・ポスター発表については78件を採択しました。特に、デモ・ポスター発表については、当初は発表枠を50件として準備を進めていたのですが、締切り前にしてその枠を大幅に超える申込みがあったため、会場への問い合わせや関係者での打合せを急遽行うことにより、発表枠を増やすことにしました。このため、運営委員の皆様への負担が大幅に増える状況となりましたが、一方で、多数のデモ・ポスター発表があるという事実は、デモ重視というヒューマンコンピューターインタラクション分野の特徴の現れであるとともに、WISS に関わる人の数と力が強まっている現れでもあるので、本分野と WISS がこれから更に発展する可能性を示しています。

依然として Covid-19 が猛威を振るっている中、現地会場を設けるにあたって、ワクチン接種の確認や、前回よりも一部強化した感染症対策のために、運営委員の皆様、特に実行委員長の村尾和哉氏、そして会計担当の瀬川典久氏には、膨大な量のイレギュラーな業務をご担当いただきありがとうございます。ここに深く感謝いたします。

最後に、登壇発表論文を投稿いただいた皆様、デモ・ポスター発表に申込みいただいた皆様、スポンサー企業の皆様、協賛学会関係者の皆様、今回も会計業務をお手伝いしてくださっている林康恵様、そして、WISS2022 プログラム委員/運営委員の皆様

* 筑波大学、WISS2022 プログラム委員長

WISS 2022 プログラム委員長

志築 文太郎 (筑波大)

WISS 2022 実行委員長

村尾 和哉 (立命館大)

WISS 2022 副実行委員長

中野 倫靖 (産総研)

WISS 2022 運営委員

会計 瀬川 典久 (京産大)
丸山 一貴 (明星大)

出版・デザインコンペ 池松 香 (ヤフー, 東北大)
加藤 邦拓 (東京工科大)

参加者管理 樋口 啓太 (Preferred Networks, 醸燻酒類研究所)

照明・音響 川口 一画 (筑波大)

広報・学会リエゾン 横窪 安奈 (青学大)

WISS チャレンジ 入江 英嗣 (東大)

チャット 福地 健太郎 (明治大)

動画生中継 川口 一画 (筑波大)
高田 峻介 (神戸高専)

デモ・ポスター 土田 修平 (神戸大)
塚田 浩二 (はこだて未来大)

ネットワーク 原 貴洋 (ヤマハ)
奥本 隼 (TwoGate)

Web・サーバー 福里 司 (東大)
山本 和彦 (ヤマハ)

投票 鳴海 紘也 (東大)

懇親会 倉本 到 (福知山公立大)

表彰 宮下 芳明 (明治大)
栗原 一貴 (津田塾大, クーリード)

企業スポンサー 岩本 拓也 (サイバーエージェント)
秋山 耀 (チームラボ)

学生ボランティア 中野 倫靖 (産総研)
宮藤 詩緒 (東工大)
大西 鮎美 (神戸大)

表紙デザイン

杉山 恭之 (明治大)

WISS 2022 プログラム委員

秋山 耀 (チームラボ)
池松 香 (ヤフー, 東北大)
入江 英嗣 (東大)
岩本 拓也 (サイバーエージェント)
大西 鮎美 (神戸大)
奥本 隼 (TwoGate)
梶本 裕之 (電通大)
加藤 邦拓 (東京工科大)
川口 一画 (筑波大)
倉本 到 (福知山公立大学)
栗原 一貴 (津田塾大, クーリード)
小池 英樹 (東工大)
坂本 大介 (北海道大)
志築 文太郎 (筑波大)
瀬川 典久 (京産大)
高田 峻介 (神戸高専)
塚田 浩二 (はこだて未来大)
土田 修平 (神戸大)
中野 倫靖 (産総研)
鳴海 紘也 (東大)
原 貴洋 (ヤマハ)
樋口 啓太 (Preferred Networks, 醸燻酒類研究所)
福里 司 (東大)
福地 健太郎 (明治大)
丸山 一貴 (明星大)
宮下 芳明 (明治大)
宮藤 詩緒 (東工大)
村尾 和哉 (立命館大)
山本 和彦 (ヤマハ)
横窪 安奈 (青学大)

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: 志築 文太郎
筑波大学 システム情報系
〒305-8573 茨城県つくば市天王台 1-1-1
E-mail: shizuki@cs.tsukuba.ac.jp

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

もくじ

まえがき: WISS2022 開催にあたって
志築 文太郎 i

セッション1：入力手法

TetraForce: スマートフォンの表裏両面に対する垂直・剪断方向の力を入力可能な磁気式インタフェース
ショート採録 **ショート発表**
土田 太一, 藤田 和之, 池松 香, Sayan Sarcar, 高嶋 和毅, 北村 喜文 1

本論文はスマートフォンのケースに磁石を埋め込んだ簡潔な入力インタフェースの提案になります。スマートフォンとケースの間はばね構造で結ばれており、磁石の移動はスマートフォンの地磁気センサによって読み取られます。またタッチセンシングの結果とも合わせることで、表面に力を加えようとしているのか、裏面に力を加えようとしているかの区別も行っています。シンプルかつ有用な提案です。

スマートフォン背面のジェスチャ入力を実現するスマホリング型デバイスの設計と実装

ロング採録 (採録条件あり) **ロング発表**
日下部 完, 坂本 大介, 小野 哲雄 8

本論文では、複数の圧力センサを配置したスマホリングを用いたスマートフォン背面でのジェスチャ入力手法を提案しています。提案手法自体の技術的な新規性は低いですが、設計したデバイスに合わせたジェスチャセットを検討し、性能評価により高い認識精度を示していることから有用性が認められます。査読者全員から WISS の登壇発表として採録すべきと判定しました。

ReflecTouch: 角膜反射像を利用したスマートフォンの把持姿勢推定 **ロング採録** **ロング発表**
張 翔, 池松 香, 加藤 邦拓, 杉浦 裕太 15

スマートフォンのフロントカメラで撮影された顔写真を使用し、ユーザの角膜に反射したスマートフォンの画面の反射像を機械学習で分類することで、把持姿勢を推定する手法を提案している。総じて新規性があり、応用範囲も広く考えられる手法であると考えられる。現状の手法のままでは利用環境が制限されることが考えられるため、更なる工夫が期待される。

セッション2：コンテンツ作成・理解

Crane: プロダクトの意匠性・可折性・可製造性を考慮した折紙の計算設計プラットフォーム

ショート採録 (採録条件あり) **ショート発表**
須藤 海, 野間 裕太, 谷道 鼓太郎, 鳴海 紘也, 館 知宏 23

折紙プロダクトの設計・造形を支援するツール (Grasshopper のライブラリ群) と、デザイン事例を紹介している。提案されているツールは大変興味深く、実装の完成度も高い点が高く評価された。一方、長大な内容を圧縮したためか、投稿時の原稿にはシステムの挙動等、不明瞭な記述も散見された。こうした点を踏まえて、採録条件付きのショート採録と判定した。

FaceShow: 顔の部分表示による遠隔会議支援システム	ショート採録 (採録条件あり)	ショート発表	
岸本 諒大, 土田 修平, 寺田 努, 塚本 昌彦			31

本研究は、ビデオ会議において、顔の部分表示によって顔の開示度をさげつつリアクションを伝達することでコミュニケーションの円滑化を目指した研究である。評価方法等に不適切と思われる点があること、また現状の評価の範囲では既存手法（アバター）に対する優位性が不明瞭である反面、顔の部分表示というアプローチは非常に興味深く活発な議論が期待されることから、ショート採録と判定された。

研究発表コンテンツ内の意味的対応関係を活用した研究理解の支援	ショート採録 (採録条件あり)	ショート発表	
太田 彩, 脇田 建			38

オンラインでの学会参加のために論文や発表資料を有機的に関連付けることで、リッチなオンライン学会参加体験を実現する研究は、新規性の観点での判断は難しい一方で、有用性は高いと考えられます。投稿段階の論文については、出版に向けて複数の懸念事項があり、これらを解決することを条件としてショート採録とします。

セッション3：トレーニング支援

Gino .Aiki: 合気道の身体の使い方の習得を支援する MR ソフトウェア	ショート採録	ショート発表	
鈴木 湧登, 坂本 大介, 小野 哲雄			46

合気道の指導でよく行われる、「球を取るように」のような比喩的な身体動作指導を MR 環境で実現したものである。新規性のあるシステム提案を行っており、荒削りではあるが、本ワークショップにおいて議論に足るものである。以上の結果から、ショート採録と判断された。

振りの理解を助けるためのダンス動画の自動分割	ショート採録	ショート発表	
遠藤 輝貴, 土田 修平, 五十嵐 健夫			56

ダンスの振りを覚えることを支援するユーザインタフェースを実現するために、視覚特徴と聴覚特徴を用いてダンス動作を自動分割する手法を提案しました。アノテーションインタフェースと応用インタフェースの2種類を実装されており、実用性高く動作していると判断されました。全査読者のコメントを参考に、論文の質をさらに高めていただけることを期待します。当日の発表を楽しみにしております。

リアルタイムフィードバックとスコアリングを導入した顕微鏡縫合術訓練	ショート採録 (採録条件あり)		
ショート発表			
田代 優香, 宮藤 詩緒, Hwang Dong-Hyun, 清藤 哲史, 金 太一, 五十嵐 健夫, 小池 英樹			63

顕微鏡縫合術の訓練にリアルタイムフィードバックとゲーミフィケーションを用いてその訓練へのモチベーションを向上することを旨とした研究です。顕微鏡縫合術の訓練に対するスコアリングの方法を適切に定めており、シンプルなゲーミフィケーション表現でも意欲向上が見られるシステムが実現されています。

ユーザ嗜好と栄養摂取基準に基づくレシピ推薦手法の開発	ショート採録 (採録条件あり)	ショート発表	
尾関 沙羅, 小寺 正明, 石黒 勝彦, 西村 太一, 樋口 啓太			72

本研究では、栄養摂取基準に基づいたレシピを提案するアルゴリズムの提案や GUI の実装にとどまらず、シミュレーションによる定量評価、ユーザ評価による定性評価などが行われました。これは WISS の 6 ページのフォーマットを考えると非常に充実した内容であり、提出されたビデオでも提案システムの挙動がわかりやすく紹介されていました。このため、いくつかの微修正はありますが、WISS での登壇発表にふさわしい論文として「ショート採録 (採録条件あり)」としました。

セッション 4：ウェアラブル

メアンダコイル++：継続的なウェアラブルコンピューティングのための衣類全面での安全で高効率な無線給電	ロング採録	ロング発表	
高橋 亮, 雪田 和歌子, 横田 知之, 染谷 隆夫, 川原 圭博			79

全ての査読者が新規性／有用性／信頼性共に十分だと判断しており、採録と判断します。CHI2022 にも採録されている論文であり、翻訳や文章削減に伴う問題もほぼありません。よく 6 ページにエッセンスを収めて頂いています。発表形式は既発表であることからショート発表も検討しましたが、技術的な説明に時間が必要と思われる点や、しっかり議論したいという査読者の意見を踏まえてロング発表としました。

音漏れ信号を用いたヒアラブルデバイス向け手形状認識手法の検討	ショート採録	ショート発表	
雨坂 宇宙, 渡邊 拓貴, 杉本 雅則, 志築 文太郎			87

ヒアラブルデバイスのためのジェスチャ認識のためにイヤホンからの音漏れに着目した点は査読者全員が新規性を認めているが、実用性の面や安全性に大きな疑問が残る論文である。音漏れを用いたジェスチャ認識自体については WISS で議論することに価値があると判断し、ショート採録とするのが妥当である。

セッション 5：実世界

車椅子使用者を周囲の人と遠隔介助者に繋ぐ遠隔車椅子システム	ショート採録	ショート発表	
高木 瑠名, 宮藤 詩緒, ジェフェルソン パルドムアン, 小池 英樹			94

すべての査読者のスコアが 4 であった。提案システムはターゲットユーザへのインタビューに基づいており、要求を明らかにしたうえで設計をしている。システム評価について、視線推定のみに注目し、技術的な基礎精度評価と、ユーザ評価を行い、ある程度の有用性が示されている。一方で、ターゲットユーザを対象とした実験は実施されていない。上記の理由から、ショート発表での採録とした。

Vi-MAP: Visual SLAM と振動子を用いた視覚障害者の歩行支援システム	ショート採録 (採録条件あり)		
	ショート発表		
坂本 祐輝, 福田 大翔, 内原 正一, 門本 淳一郎, 入江 英嗣, 坂井 修一			102

本論文は、視覚障害者のための屋内ナビゲーションシステムを構築、評価したものです。スマートフォンで屋内マップを作成し、方向に依存した振動提示によってナビゲーションを行っています。技術的な構成要素は新規とは言えないものの、当事者への丁寧なインタビューによって要素を抽出し、実装評価に至っている点は高く評価できます。

Alertable Surfaces: 人の接触と消毒を認識しウイルスの付着を警告する実環境 **ショート採録 (採録条件あり)**

ショート発表

岸野 泰恵, 白井 良成, 柳沢 豊, 菅原 光, 尾原 和也, 水谷 伸 110

近赤外線カメラとサーマルカメラを用いて接触や消毒行為を検出し、プロジェクタでウイルス付着状況を提示するルームスケールのシステムを提案した論文です。提案システムの新規性を担う部分の有用性が示せていない点などに懸念があるものの、システムの実装はされているため、WISS で議論可能な論文であると判断されました。

自動運転車からの視線提示による交通事故低減の可能性の検討 **ロング採録 (採録条件あり)** **ロング発表**
張 家銘, 戸田 光紀, Gui Xinyue, Seo Stela H. , 五十嵐 健夫 117

本研究は、自動運転車が視線提示を行う場合に、状況に応じて適切に横断するか否かを判断できるかを調査した研究である。自動運転車による視線提示自体は既発表の内容であるが、危険な状況（自動運転車が歩行者を識別していない状況）における歩行者の判断に焦点を当てた点、およびそこから得られた結果は面白く、活発な議論が期待されることからロング採録と判定された。

TetraForce: スマートフォンの表裏両面に対する垂直・剪断方向の力を 入力可能な磁気式インタフェース

土田 太一* 藤田 和之* 池松 香*[†] Sayan Sarcar[‡] 高嶋 和毅* 北村 喜文*

概要. スマートフォンの2つの面（タッチ面と背面）に対する2種類の方向（垂直方向と剪断方向）の力による入力を可能にするスマートフォンケース型の磁気式ユーザインタフェース“TetraForce”を提案する. 本インタフェースの背面には, 3 DoFで可動するパネルが搭載されている. このパネルには磁石が取り付けられており, その変位をスマートフォン内蔵の地磁気センサ（および加速度センサとジャイロセンサ）により推定することで力による入力の検出を実現する. 本インタフェースのプロトタイプを用いたユーザスタディ(N=12)の結果, ユーザによる4種類の入力が平均97.4%の精度で意図通りに検出できることを確認した. また, 本インタフェースを用いたアプリケーションを複数実装し, その有用性を議論した.

1 はじめに

タッチ入力, スマートフォンやタブレット等のスマートデバイスにおける主要な入力手段である. 一方で, 入力に利用できる情報は2次元座標のみであることから, 入力語彙の乏しさの問題がたびたび指摘されてきた. これに対し, 入力語彙を拡張するための研究が盛んになされておられ, その1つに「力」の活用が挙げられる. 力は, タッチと同様に指で容易に入力でき, 大きさと方向によるベクトル情報を持つことから, 力を用いて入力語彙を拡張するユーザインタフェースが数多く開発された.

力を入力可能とする既存インタフェースは, 表1に例示するように, 力の印加方向（垂直・剪断方向）と力の印加対象（タッチ面・それ以外）の組み合わせによる4つに大きく分類することができる. これらの力の印加方向と印加対象の違いは, 異なるメタファや操作性を提供しうるが, それらはこれまで個別に検討されてきており, それぞれ個別の機構が必要であった. 4種類の力による入力を単一の機構で実現できれば, デバイスの付替えを必要とせず, 豊富な入力語彙を提供できる汎用的なユーザインタフェースになり得る.

表 1. 力を入力可能とする既存インタフェースの代表例

	垂直方向	剪断方向
タッチ面	GraspZoom[11], Suzuki et al.[12]	Harrison et al.[5], ShearSheet[8]
	Force gestures[6]	
背面	Besançon et al.[2], BackXPress[4]	Xial et al.[15], Cam- TrackPoint[16]

Copyright is held by the author(s).

* 東北大学

[†] ヤフー株式会社

[‡] Birmingham City University

そこで本研究では, 印加対象面と印加方向の組み合わせによる4種類の力による入力を1つの機構で実現可能なスマートフォンケース型の磁気式インタフェース“TetraForce”を提案する. TetraForceは3 DoFで可動するパネルを持ち, パネルの表面に磁石を搭載する. この機構をスマートフォンの背面部分に装着し, パネルの変位方向（垂直・剪断方向）および変位量をスマートフォン内蔵の磁気センサを用いて推定する. また, ユーザがスマートフォンを把持している場合, タッチ面・背面のいずれの面に力を加えてもパネルは変位する. この性質を用い, タッチ入力判定の有無により力の印加される面（タッチ面・背面）を区別する. 以上に述べた実装により, 4種類の入力を外部電源なしに実現できる. また, 力の入力時には弾性フィードバックが得られるため, より正確な入力操作に寄与する[17]と期待できる.

本稿では, TetraForceの設計とプロトタイプの実装, 及びその性能を調べるユーザスタディについて述べ, 最後にアプリケーション例を紹介する. なお, 本稿はACM ISS 2022で発表した内容[14]に, アプリケーションの実装例とハードウェアの改良を追記したものである.

2 TetraForce

TetraForceは, スマートフォンの背面で3DoFで可動するパネル部と, そのパネル部の動きをトラッキングするための磁気センシングにより, 4種類の入力を単一の機構で実現するインタフェースである. 以下では, 本インタフェースのプロトタイプ実装について述べる. なお, 本インタフェースが実現する4種類の入力を, 以降では図1に示すように定義する.

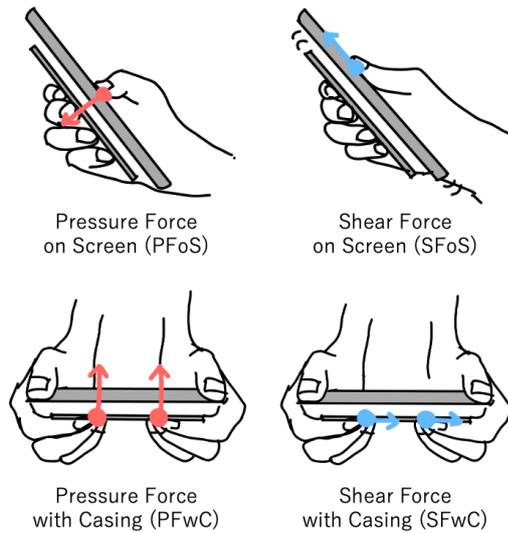


図 1. 実現する 4 種類の入力

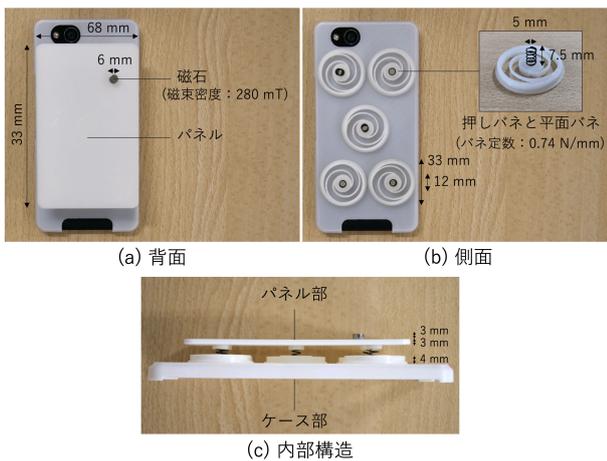


図 2. プロトタイプ実装の概観

2.1 ハードウェア実装

図 2 にプロトタイプの概観を示す。この実装には Google Pixel 2 を用いているが、他の数種類のスマートフォンでも同様に動作する見込みがあることを確認している。

本インタフェースは、スマートフォンに取り付けるケース部と、それと複数のバネで接続された可動のパネル部からなる。ケース部は FDM 方式の 3D プリンタで PLA フィラメントを使用し印刷したもので、市販のスマートフォンケースと同様の形状である。パネル部は同様の素材による長方形の板である。パネル部は、スマートフォンのカメラを覆わない位置及びサイズとした。このパネル部をユーザによる力の印加に応じて 3 DoF で可動とするため、パネル部とケース部との間には平面バネと押しバネの 2 種類を組み合わせたものを 5 箇所（パネル部の四

隅と中央）に配置する。平面バネはケース部及びパネル部と同様に 3D プリンタ製で、外周部分を接着剤でケース部に固定してある。作成した平面バネは方向によってわずかに可動域が異なっていたため、5 つの平面バネをそれぞれ異なる向きに配置することでバランスをとった。また、市販の押しバネの一端をこの平面バネにあるくぼみに、他端をパネル裏面にある同サイズのくぼみにはめ込み、接着剤で固定した。

パネル部表面上には、ネオジム磁石を接着剤で取り付けた。磁石の位置は、予備検討により磁気センサの値が得られやすい位置として決定した（磁気センサの位置から一定距離離すことで磁気ベクトルの変化が読み取れるようであった）。

入力可能な力の最大値に関して、言及のあった先行研究 [7, 4, 12] を参考にし、かつ実装時点で購入可能であったバネを複数検討し、押圧力入力の際には約 12.5 N、剪断力入力の際には約 13.0 N とした。この力の大きさは、異なるバネ定数のバネや異なる可動域を持つ平面バネで置き換えることで調整が可能である。

2.2 ソフトウェア実装

ソフトウェアの実装として、以下ではキャリブレーション、4 種類の入力の識別、押圧力・剪断力それぞれの推定アルゴリズムについて順に述べる。

キャリブレーション 磁気センサで取得される値は地磁気の影響を受けるため、力の入力の推定前にその影響を取り除くためのキャリブレーションが必要である。キャリブレーションでは、本インタフェースをスマートフォンから一旦取り外し、十分に距離を取った状態で取得した磁気ベクトルを地磁気ベクトルとして保存する。後述の推定アルゴリズムでは、保存した地磁気ベクトルを測定値から差し引いて使用する。

4 つの入力タイプの識別 力の印加面（タッチ面・背面）の識別は、タッチパネルへのタッチ入力が出検されているかどうかで行う。すなわち、タッチを伴う入力であれば FoS、タッチを伴わない入力であれば FwC として扱う。剪断力と押圧力は後述する個々のアルゴリズムで認識され、同時の認識も可能である。

剪断力の推定 剪断力の推定のため、先行研究 [1] の 4 章を参考に、パネルの水平方向の変位を算出するアルゴリズムを実装した。すなわち、磁気センサで得られた磁気ベクトルの値から、キャリブレーションで得られた地磁気ベクトルを差し引いた値を磁石の磁気として用い、磁石までの距離および 2 次元位置を理論式を用いて求めることで剪断力の方向と強さを推定する。

押圧力の推定 押圧力の推定方法として、当初は剪断力と同様の理論的なアプローチを検討したが、出力の次元が増え計算量が膨大になることから、機械学習のアプローチを採用した。すなわち、押圧力の入力がないか否かの確率を、Sigmoid 関数を活性化関数に用いて出力するニューラルネットワークモデルを構築することで、出力される確率を力の連続量として利用できると考えた。学習データに用いた変数はスマートフォン内蔵のセンサから得られるものとして、3次元磁気ベクトル、3次元加速度ベクトル、加速度の大きさ、4次元角速度ベクトル、3次元重力ベクトル、3次元回転率ベクトルを用いた。デバイスを地面と水平の状態から $\pm 45^\circ$ 程度の範囲内で前後左右に傾けながら画面中央に対して押圧力を加えた場合 6,752 データと加えない場合 6,560 データ、押圧力を加えず剪断力を入力した 13,186 データの計 26,498 データを収集した。これらのデータのうち 80 % を学習データ、20 % をテストデータに分割し学習を行った。モデルの精度は 98.18 % であった。このモデルを用いて磁石位置を推定することで、押圧力の推定値を連続値として取得できることを確認した。

上記のモデルを用いて概ね押圧力の推定が実現できた一方で、スクリーンの右端や下端に対する PFoS が検出されにくいことがわかった。これは、(磁気センサの位置の都合により) 磁石の位置がパネルの中央からややずれた位置にあることで、磁石から離れた位置でのパネルの沈み込みが磁気ベクトルの変化として反映されにくいことによるものであった。これを解決するため、ユーザのタッチ位置に応じた推定値の補正を試みた。具体的には、機械学習アルゴリズムで推定された力の大きさに、タッチ位置に応じたゲインを掛け合わせるにより補正する。ゲインの分布を求めため、画面を縦 $9 \times$ 横 $5 = 45$ のエリアに分割し、それぞれを最大まで 3 回押したときの押圧力の推定値の大きさを測定し、平均をとった。表 2(a) は、スマートフォンの画面領域内における押圧力入力の推定値 (%) の分布である。ゲインは各値の逆数で表すことができる。ただし、最右列 (幅 12.5 mm) と最下行 (高さ 12.3 mm) の領域では押圧力の検出がほぼ不可能であったため、本実装の段階ではこれらの領域の押圧力取得を対象外とした。この点を解決するハードウェアの改良については 6 章で述べる。

3 パフォーマンススタディ

実装したプロトタイプ的基本的な性能評価として、ユーザによる実際のインタラクションがユーザの意図通りに作動するかをユーザスタディを通じて調査した。具体的には、4 種類の入力 (PFoS, PFWC, SFoS, SFWC) それぞれで、力を特定の強さで一定時間加えることで画面上のコンテンツを操作するタ

表 2. PFoS 入力における入力位置ごとの力の推定値の分布

(a) 実装したプロトタイプ (b) 改良版のプロトタイプ

80	70	52	30	3	100	100	87	83	56	
95	79	68	45	3	100	100	90	97	58	
100	93	78	47	5	100	100	92	98	46	100-81
100	98	86	37	2	100	100	92	87	51	80-61
100	100	81	22	2	100	100	92	72	54	60-41
100	99	58	15	0	100	99	89	68	50	40-21
93	80	39	14	0	91	99	81	69	46	20-1
55	31	12	3	0	81	96	73	64	45	0
22	5	0	0	0	57	65	51	42	41	

スクを実験参加者に行ってもらい、その入力成功率を測定した。

3.1 実験参加者

実験参加者は 12 名 (男性 10 名、女性 2 名、平均年齢 22.3 歳、SD=1.23、右利き 10 名) の大学生であった。COVID-19 に関する規制により外部の参加者の募集が困難であったため、参加者のうち 8 名は本研究についてある程度の知識があり 3 名は実際にデバイスを使用した経験があったが、実験タスクで用いられたコンテンツの使用は初めてであった。参加者には謝礼 (大学の規定に基づき学年により異なる額、約 1,000 円) が支払われた。

3.2 実験環境

実験には、Android (バージョン 11) を搭載する Google Pixel 2 (解像度: 1920×1080 , ディスプレイ: $110.7 \text{ mm} \times 62.3 \text{ mm}$) に、TetraForce のプロトタイプを取り付けて用いた。実験ソフトウェアは Unity (バージョン 2020.3.2f1) で実装し、プログラムは C# で記述した。なお、毎回の実験開始前には、2.2 章に記載したキャリブレーションを予め実施した。

3.3 タスクデザイン

実験タスクは、4 種類の入力それぞれについて、特定の強さの力を一定時間入力してもらうものとした。タスクは先行研究 [13, 10] に倣い、力の強さに応じて 1 次元方向に動く画面上のバーをインジケータの 4 段階のターゲットエリアのうちのいずれかに 1 秒間収めることができた場合にタスク成功とし、入力開始後にバーがターゲットエリア外にある状態で 1 秒間以上経過した場合には失敗とした。4 種類の入力に関するタスクごとに、全試行に占める成功試行の割合である成功率を測定した。

4 種類の入力それぞれに関する 4 つのタスクで提示された実際の画面例を図 3 に示す。ターゲットと

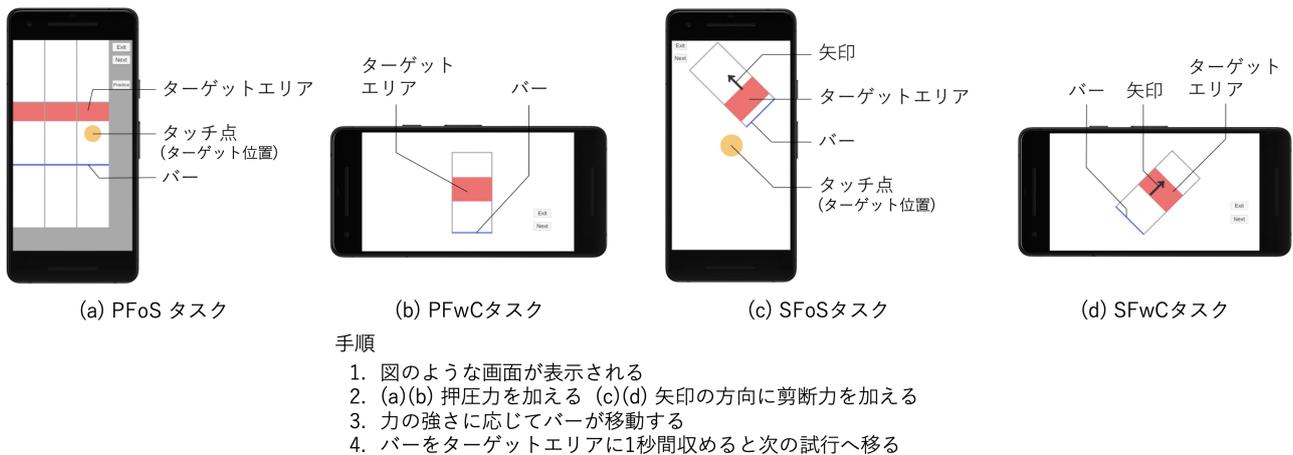


図 3. 4つのタスクで提示される画面と手順

なる力の強さは4段階 (Lv. 0-3) とし、インジケータにより表現された。このエリアの割り振りは先行研究 [3] を参考に決定した。具体的には、押圧力では 12.5 N、剪断力では 13.0 N を最大とし、Lv. 0 では 0-10 %、Lv. 1 では 10-40 %、Lv. 2 では 40-70 %、Lv. 3 では 70-100 % となるようマッピングした。PFoS タスクでは、押圧力を印加する際のタッチ位置は $3 \times 3 = 9$ つのエリア (2.2 に述べた、入力が困難な領域は除外) からランダムに指示した。一方で、PFwC タスクでは押圧力の印加位置を特に指示しなかった。SFoS と SFwC タスクでは、上下左右斜めの 8 方向の剪断力の入力をランダムに指示した。参加者には、タスクごとにスマートフォンの持ち方をあらかじめ提示した。具体的には、FoS のタスク 2 種類ではスマートフォンを縦向きにし、利き手で持ち親指でタッチ操作してもらうようにした。FwC のタスク 2 種類では横向きにして両手で持ち、親指が前面に、その他の指が背面側にあるような把持姿勢とした。

3.4 実験手順

実験は 4 種類の入力に対応する 4 つのタスクに分かれていた。タスク内容について参加者の混乱がないよう、全参加者に共通して、よりシンプルなタスク内容である押圧力のタスク 2 種 (PFoS と PFwC) を実施後、剪断力のタスク 2 種 (SFoS と SFwC) を実施した。ただし、押圧力のタスク内、剪断力のタスク内での 2 種のタスクの提示順序は参加者間でランダムとした。

各タスクでは、実験者からの説明の後、参加者は練習タスクと計測を伴う本タスクを行った。練習タスクでは、本番タスクと同様のタスクを操作に慣れるまで数分程度行ってもらった。練習タスク・本タスク中、各試行が完了した際には画面に「指を離してください」というメッセージが表示され、その 2 秒後に次の試行が自動的に開始された。前半 2 タス

クが終了した時点で、参加者は必要に応じて数分の休憩をとった。全 4 タスクが終了した時点で実験は終了した。実験全体の所要時間は参加者によってばらつきが大きかったが、いずれの参加者も 50 分以内で実験を終了した。

3.5 結果と考察

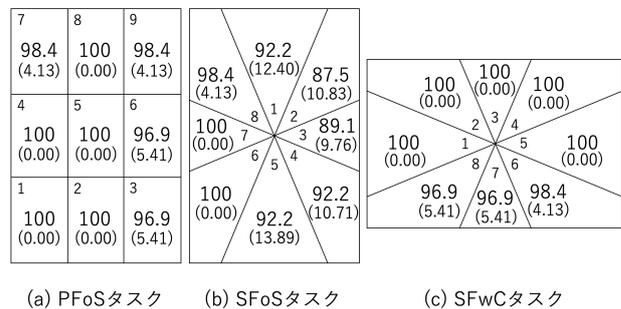


図 4. 各タスクにおける入力成功率 (%). 括弧内は標準偏差を示す. 各エリアは力が加えられた (a) 位置または (b)(c) 方向を示している.

実験の結果、全試行での平均入力成功率は 97.4 % であった。タスクごとに見ると、入力成功率は PFoS では 99.0 % ($SD = 1.28$), PFwC では 97.7 % ($SD = 3.03$), SFoS では 94.0 % ($SD = 4.58$), SFwC では 99.0 % ($SD = 1.34$) であった。これらの成功率は、例えば PFoS について類似タスクを実施した先行研究 [10, 13] と同程度の高さであり、提案インタフェースのプロトタイプは 4 つの入力を十分な精度で取得可能であるといえる。

図 4 に、力の印加位置が 1 通りであった PFwC タスク以外の 3 タスクにおける印加位置や方向ごとの入力成功率を示す。PFoS の結果 (図 4(a)) に関して、スクリーンの右側エリア 3, 6, 9 ではエラーが比較的多く見られたことがわかる。これは、タッチ

エリアに応じた推定圧力値の違いが影響したと考えられる。すなわち、2.2節に述べたように、当該エリアでは取得される圧力の推定値が特に小さく（表2(a)）、これを補正するためのゲインが大きく設定されている影響で、ユーザの意図しないバーの動きがあり操作が難しくなったと考えられる。

SFoSの結果（図4(b)）に関して、画面左側への入力（方向6, 7, 8, 平均99.5%）に比べて画面右側への入力（方向2, 3, 4, 平均89.6%）の入力成功率が低くなった。SFwCの結果（4(c)）では、スマートフォンを横向きに把持した状態での下方向への入力（方向6, 7, 8）でのみエラーが見られた。これらの印加方向による成功率の差異は、運動学的に、それらの方向への入力が右手では難しかったからではないかと考えられる（剪断力の入力しやすい方向を調べた研究は我々の知る限り無い）。これについては参加者を増やすことや左利きのユーザにもテストしてもらうことで、明らかにできる可能性がある。

4 ハードウェアの改良



図5. 改良版のプロトタイプ実装の概観

実装したプロトタイプの大きな制限である、画面右または下のエリアで押圧力をほとんど検出できない点を解消するため、図5に示すように、パネル部の不均一な沈み込みを発生しにくくする新たなプロトタイプを試作した。この実装は、パネル部のサイズや用いた磁石は元のプロトタイプと同様である（図5(a)）が、ケース部とパネル部の間に、キーボード等の機構を参考にした、パンタグラフ構造の3Dプリントパーツを新たに搭載している（図5(b)）。押圧力によるたわみを防ぐため、ケース部以外のパーツには基本素材としてOnyx(マイクロ炭素繊維充満ナイロン)を使用した上、パネル部とパンタグラフ構造および平面バネには強化素材としてFiberglassを使用して、FDM方式の3Dプリンタで印刷した。

パンタグラフ構造は、2つの棒状のパーツとそれらを支える上面・下面の板から成り立っており、2つの棒状パーツが交差することによりパネル部全体が沈み込むように可動する（図5(d)）。パネル部の適度な押圧力による沈み込みを実現するため、押しバネ1つが上面・下面の板の間を支えている。また、このパンタグラフ構造自体とケース部との間には、図5(c)に示すような、平面バネを2つ埋め込んだ長方形の板を挿入することで、剪断方向の可動を実現させた。平面バネはそれぞれの渦の回転方向を互い違いにすることで、可動域のバランスをとった。平面バネそれぞれの中央をケース部に、平面バネ以外の周辺部をパンタグラフの下面に接着した。なお、入力可能な力の最大値に関して、パフォーマンススタディでは必要な力が大きすぎるといったコメントが複数得られたことから、本実装では押圧力では約7.7 N、剪断力では約7.5 Nを最大とした。

このデバイスを用いて、2.2節で述べたものと同様の方法でPFoS入力における入力位置ごとの推定値の分布を測定した。表2(b)はその分布表である。提案インタフェースでの分布である表2(a)と比較すると、全体として分布の偏りに改善が見られた。特に、元のプロトタイプでは入力を検出できなかった画面右側と画面下側では、ゲインの設定は未だに必要ではあるものの、検出が可能となったと言える。

5 アプリケーション例

本章では、TetraForceの利点や応用場面について議論を促したワークショップ(N=12)の結果[14]をもとに筆者らが実装したアプリケーション例を、ワークショップ参加者が述べた使用上の利点とともに紹介する。なお、図6に、各アプリケーションの画面例と、表3に入力のマッピング例を示す。

1人称視点シューティング(FPS)ゲーム 実装例を図6(a)に示す。PFoSの入力を連射アクションに割り当てることで、タップでの武器の発射アクションとのシームレスな切り替えが可能となる。PFwCには、ボイスチャット開始操作のジェスチャを割り当てることで、ゲームのコンテキストに関わらず操作を実行できる。また、前進や後退等の移動操作は従来、画面上のジョイスティックUIのドラッグにより実現することが多かったが、この操作にSFoSを割り当てることで、ユーザはUIの位置を視認する必要なく、力覚フィードバックから移動速度を調節できる。加えて、必要に応じてPFwCでの操作によって画面を指で遮蔽しない視点操作が可能となる。

動画プレイヤー 実装例を図6(b)に示す。PFoSやPFwCをジェスチャ入力に割り当てることで、動画のコンテンツをUIや指で遮蔽することなく操作を完了できる。SFoSでの音量調節はボタンでの操作に

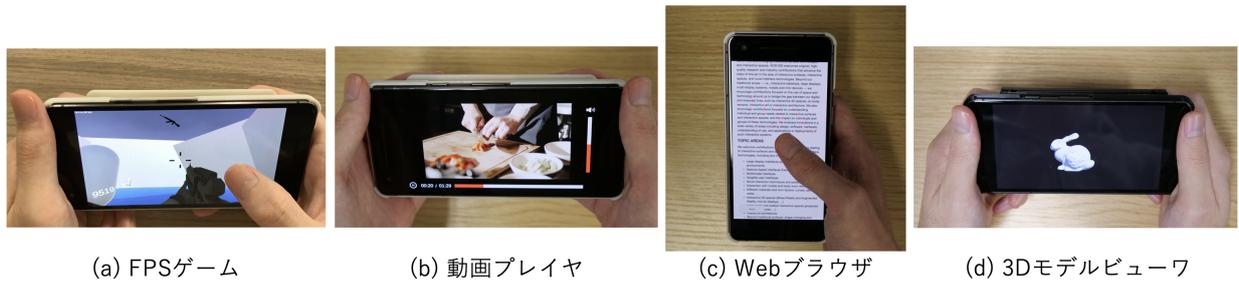


図 6. アプリケーションの実装例

表 3. 各アプリケーションでの入力と操作のマッピング例

	FPS ゲーム	動画プレイヤー	Web ブラウザ	3D モデルビューワ
PFoS	武器の連射	お気に入り登録	ブックマーク	縮小
PFwC	ボイスチャットを開始	再生・一時停止	スクリーンショット撮影	拡大
SFoS	アバターを移動	音量調整 (上下)	スクロール (上下)	カメラの回転 (パン, チルト)
SFwC	視点操作 (ロール軸)	早送り・巻き戻し (左右)	タブ切替 (左右)	モデルの回転 (ヨー, ピッチ方向)

比べて入力位置によらない操作によって実現できるほか、直感的に連続量の調節が可能である。SFwCによって、指で画面を遮蔽せずに早送りや巻き戻しの操作が可能となる。

Web ブラウザ 実装例を図 6 (c) に示す。PFoS や PFwC をジェスチャ入力に割り当てることでスクリーンスペースを削減できる。SFoS をスクロール操作に割り当てることで、従来のフリック入力よりも操作回数の少ない操作が可能になる。また、タブの切替は、従来 2 タップ以上を要するが、SFwC に割り当てることで 1 回の操作による発動が可能になる。

3D モデルビューワ 実装例を図 6 (d) に示す。画面に対する奥行き方向の移動 (縮小・拡大) 操作をそれぞれ PFoS と PFwC に、剪断方向の移動操作を PFoS に、モデルの回転操作を SFwC にそれぞれ割り当てることで従来は操作モードの切り替えが必要であった複数の操作をシームレスに使用できる。

6 制限と今後の展望

実験の結果から、本インタフェースを用いて 4 種類の入力を実用的な精度で実現できることが明らかになったが、実験で試行した条件以外にも、スマートフォンの機種やサイズ、使用する際の向き、把持姿勢、利き手など、様々な変数が考えられ、今後調査する必要がある。また、加えた力がどの程度正確に認識されているかといった厳密な意味での性能評価はしておらず、これも今後の課題である。加えて、本プロトタイプの厚み (磁石を除き 14 mm) は市販

の多くのスマートフォンケースよりも厚く、ユーザビリティに影響する可能性がある。今後はケースの薄型化と合わせてその影響を調査する必要がある。

別の観点として、地磁気 (ユーザの物理的な位置) の影響が挙げられる。予備的なテストでは、磁気センサのキャリブレーション後に約 500 m 移動したところ、それによる地磁気の影響は無視できる程度であったが、より大きな移動では問題になる可能性があり、さらなる調査やキャリブレーションの簡素化が必要である。加えて、スマートフォンが強く揺れる場面 (例: ポケットの中にある場合や電車内での使用中) や磁性体オブジェクト (例: 指輪などの金属製品やイヤホンなどの電化製品) の接近による誤作動の有無やその回避方法についても調べる必要がある。また、本プロトタイプで用いたスマートフォン (Google Pixel 2) 以外で用いる場合、磁石の位置の変更が必要になる可能性がある。このため、今後はユーザ自身が磁石をスライドさせながら適切な設置位置を見つけられるようなキャリブレーションソフトウェアを開発したい。

また、先行研究 [9] でも指摘されているように、磁石を付加するインタフェースは、地図アプリの方位磁針といった磁気センサの値を用いる他のアプリケーションの挙動に影響を与えるため、同時に利用できないという制約がある。加えて、磁気を活用する製品への影響も存在する。磁石を使った一般的なアクセサリ (例: MagSafe アクセサリ) と同様に、磁気カードや無線給電機器への干渉に注意して使用する必要がある。これらの影響を調べることも今後の課題である。

7 謝辞

本研究は JSPS 科研費 21H03473 の助成を受けたものである。

参考文献

- [1] T. Abe, B. Shizuki, and J. Tanaka. Input techniques to the surface around a smartphone using a magnet attached on a stylus. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2395–2402, 2016.
- [2] L. Besançon, M. Ammi, and T. Isenberg. Pressure-based gain factor control for mobile 3D interaction using locally-coupled devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1831–1842, 2017.
- [3] C. Corsten. *Use the force: how force touch improves input on handheld touchscreens*. PhD thesis, RWTH Aachen University, 2020.
- [4] C. Corsten, B. Daehlmann, S. Voelker, and J. Borchers. BackXPRESS: Using back-of-device finger pressure to augment touchscreen input on smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4654–4666, 2017.
- [5] C. Harrison and S. Hudson. Using shear as a supplemental two-dimensional input channel for rich touchscreen interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3149–3152, 2012.
- [6] S. Heo and G. Lee. Force gestures: augmenting touch screen gestures with normal and tangential forces. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pp. 621–626, 2011.
- [7] S. Heo and G. Lee. ForceDrag: using pressure as a touch input modifier. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, pp. 204–207, 2012.
- [8] M. Huang, K. Fujita, K. Takashima, T. Tsuchida, H. Manabe, and Y. Kitamura. Shearsheet: Low-cost shear force input with elastic feedback for augmenting touch interaction. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pp. 77–87, 2019.
- [9] S. Hwang, M. Ahn, and K.-y. Wohn. MagGetz: customizable passive tangible controllers on and around conventional mobile devices. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, pp. 411–416, 2013.
- [10] S. Hwang, A. Bianchi, and K.-y. Wohn. VibPress: estimating pressure input using vibration absorption on mobile devices. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, pp. 31–34, 2013.
- [11] T. Miyaki and J. Rekimoto. GraspZoom: zooming and scrolling control model for single-handed mobile interaction. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1–4, 2009.
- [12] K. Suzuki, R. Sakamoto, D. Sakamoto, and T. Ono. Pressure-sensitive zooming-out interfaces for one-handed mobile interaction. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1–8, 2018.
- [13] R. Takada, T. Ando, B. Shizuki, and S. Takahashi. BaroTouch: A Technique for Touch Force Sensing Using a Waterproof Device’s Built-in Barometer. *Journal of Information Processing*, 27:106–115, 2019.
- [14] T. Tsuchida, K. Fujita, K. Ikematsu, S. Sarcar, K. Takashima, and Y. Kitamura. TetraForce: A Magnetic-Based Interface Enabling Pressure Force and Shear Force Input Applied to Front and Back of a Smartphone. *Proceedings of the ACM on Human-Computer Interaction*, 6(ISS):185–206, Nov. 2022.
- [15] R. Xiao, G. Laput, and C. Harrison. Expanding the input expressivity of smartwatches with mechanical pan, twist, tilt and click. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 193–196, 2014.
- [16] W. Yamada, H. Manabe, and D. Ikeda. CamTrackPoint: Camera-based pointing stick using transmitted light through finger. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pp. 313–320, 2018.
- [17] S. Zhai. User performance in relation to 3D input device design. *ACM Siggraph Computer Graphics*, 32(4):50–54, 1998.

スマートフォン背面のジェスチャ入力を実現するスマホリング型デバイスの設計と実装

日下部 完* 坂本 大介* 小野 哲雄*

概要. スマートフォンが普及して久しいが、依然として指でタッチスクリーンを操作する際に生じる操作上の困難であるファットフィンガー問題や片手操作が困難という問題が存在している。これらの問題を扱う研究は多く、特にスマートフォン上での操作性の向上だけでなく外部デバイスを活用したインタラクション領域の拡張に取り組む研究も多い。本稿では、現在スマートフォンのアクセサリとして人気のあるスマホリングに着目し、スマートフォンのインタラクション領域の拡張を目的としたスマートフォン背面のジェスチャ認識を実現するスマホリング型のデバイスを提案する。我々はまず、ユーザ定義型ジェスチャのアプローチで20人の参加者へ聞き取り調査を行いスマホリングの最適なジェスチャを設計した。その後、機械学習モデルを使用し、設計したジェスチャを分類する2つの分類器を実装し、参加者を集めてジェスチャデータの収集を行った。集めたデータで学習と分類試験を行ったところ、日常動作とジェスチャ入力の分類では99.1%、ジェスチャの種類分類では92.7%の精度が確認された。

1 はじめに

スマートフォンをはじめとするモバイル端末は多くの便利な機能をユーザに提供する一方で、これらの端末の主な操作方法であるタッチスクリーンへの直接的なインタラクションはファットフィンガー問題 (Fat finger problem) など操作上の問題を抱えている。この問題を解決するために、外部デバイスを用いてモバイル端末のインタラクション領域を拡張する研究が多く行われてきている [5]。特にスマートフォンのカバー型のデバイスはスマートフォンのアクセサリとして設置できるため、ユーザの体に新たなデバイスを装着するコスト無しに、スマートフォンのインタラクションを豊かにするという利点が存在する [7]。

本研究ではスマートフォンのカバー以外のアクセサリとして人気のあるスマホリングに着目し、スマホリング型の外部デバイスによるスマートフォンのインタラクション領域を拡張手法を提案する (図1)。本デバイスの特徴は、圧力センサを用いてスマホリングを使用している指の動作を感知する。我々のアイデアは、このスマホリング周囲の指の動きを定義するジェスチャを設計し、これによりスマホリング型のジェスチャインタフェースを実現する。本稿では、以下の2つを行う。1) ユーザ定義型ジェスチャのアプローチによりスマートフォンの各種操作を対象としたスマホリング向けの最適なジェスチャを設計する。2) ジェスチャを認識するスマホリング型デバイスを実装する。ジェスチャの認識を行うため、我々は機械学習のモデルである畳み込みニューラルネットワーク (CNN) を使用して2つの分類器を作



図1. 本研究で開発したデバイスの概要。スマホリング型のデバイスに人差し指でダブルタップのジェスチャ入力を行うことで、地図アプリからホーム画面へと移動している様子。

成する。一つはスマホリングでジェスチャを行っているのか、日常的な操作を行っているのかを判定する分類器であり、もう一つは前述の分類器でジェスチャであると判定された場合に、実際にどのジェスチャが行われたのかを分類するものである。以上を実現するため、我々は2つの調査を行った。一つめの調査では20名の参加者へ、ユーザ定義型ジェスチャのアプローチに基づく調査を行い、24種類のスマホリングのジェスチャを設計した。その後、設計したジェスチャを認識するためのプロトタイプを開発し、続く2つめの調査では、10名の参加者から24種のジェスチャを入力しているデータを収集した。このデータを著者らが収集した日常動作のデータと合わせて分類器に学習させた。その結果、日常動作とジェスチャ入力中の分類では99.1%、24種類のジェスチャの種類分類では92.7%の精度が確認された。本稿では以上について報告する。

* 北海道大学

2 関連研究

2.1 スマートフォンのインタラクション領域の拡張

スマートフォンは動画の視聴、通話、SNSの使用など、多くの機能をユーザに提供する携帯性の高いデバイスである。スマートフォンの操作は、主に画面を指で直接触ることで行うが、スマートフォンの様な画面の小さな端末のタッチスクリーンの操作は、ファットフィンガー問題など操作上の問題を抱えている。この問題の解決に向けてスマートフォンのインタラクション領域の拡張することを目指した研究が行われてきており、その1つに外部デバイスによるインタラクション領域の拡張がある[5]。この手法では多様なセンサを使用できるため、豊かなインタラクションを提供できる一方で、ユーザの体に新たなデバイスを装着するコストが生じる。スマートフォンに装着することを前提とした周辺アクセサリ型のデバイスは、このコストを解消する手法であり、今日までにスマートフォンケース型のデバイスの研究等が行われてきている[7][14]。Yadavらは、複数のスマホリング型インターフェースをプロトタイプし、それらの比較調査を行った[13]。本稿では、ユーザ定義型ジェスチャのアプローチによってジェスチャの定義を行い、機械学習を用いたプロトタイプの開発と応用事例の提案を行う。

2.2 ジェスチャ入力によるインタラクション手法

ジェスチャとは、特定のポーズや動きによる非言語的なコミュニケーション手法の1つである。ジェスチャ入力によるインタラクションは、キーボードやマウスなどの一般的な入力手法と比較し、記憶性、学習性、さらに特定の条件下ではパフォーマンスも優れると言われている[1]。HCIの分野では非常に多くの領域でジェスチャ入力によるインタラクションの研究が行われている。これまでにジェスチャ入力を実現するための認識技術に関する基礎研究が多く行われてきており、特にどのようなセンサを使用するのかについての研究が多く行われてきている。音声を入力するマイクロフォンを用いた手法では、ドブラー効果を用いた手法や[3]、物体の表面を擦った際に生じる音響信号の解析によるジェスチャ認識手法が提案されてきている[4]。カメラを用いた手法ではRGB値のみを使用する手法のほか[10]、さらにロバストなジェスチャ認識を行うことを目的とした深度カメラや近接カメラ[9]を用いた手法が提案されている。その他にも、電界や磁界の干渉を利用したジェスチャ認識や[5]、加速度や慣性計測ユニット(IMU)を用いた軌道の解析によるジェスチャ認識[12]、圧力センサを用いた力の状態の解析によるジェスチャ認識[2]などが提案されてきている。本研究では圧力センサを用いてスマホリングでのユーザの指の動作を感知し、それをジェスチャとして扱う

ための分類器を開発することで、新しいジェスチャ入力手法を提案する。

2.3 ユーザ定義型ジェスチャ

ユーザ定義型ジェスチャとはWobbrockらが提案したジェスチャ設計手法であり、複数名のユーザが提案したジェスチャから、最も多く提案されたジェスチャを選択する[11]。これによって設計されたジェスチャは専門家が設計するよりも学習性、記憶性に優れていると言われている[6]。ユーザ定義型ジェスチャはこれまでに、テーブルトップインタフェース[11]やスマートフォン[8]、靴下型デバイスなど[2]、多岐にわたるデバイスで調査が行われている。本稿では、スマホリング型デバイスのプロトタイプで扱う最適なジェスチャを設計するため、スマホリングのユーザ定義型ジェスチャセットの調査を実施する。

3 ユーザ定義型ジェスチャのアプローチによるスマホリングのジェスチャ設計

本研究においてはRuizらが使用した23種類のタスクに対して、スマホリング型インタフェースによるジェスチャの設計を行う(図2)。ただし、我々は従来手法でのタスクに含まれる“Act on Selection”という名前のタスクは複数要素の操作が内在すると考え、このタスクを選択肢の移動のタスクと決定のタスクに細分化した。これを含めて全23種類のタスクを使用する。

3.1 ジェスチャ案の収集手順

本調査に先立って事前に300名のスマホリング使用者を対象にしたクラウドソーシングによる調査を実施した。その結果、多くの人々が人差し指もしくは中指をリングに通してスマホリングを使用すること、利き手とは反対の手でスマホリングを操作していることが明らかとなった。これらの結果を使用し、本調査の詳細を決定した。

先行研究の方針と同様に[8]、23種類のタスクを6つのシーンに割り当て、各シーンごとに、割り当てられたタスクを表示し、参加者へジェスチャの提案を促した。ジェスチャ案は、スマホリングに中指を通した場合と人差し指を通した場合でそれぞれ収集した。参加者には市販のスマホリングが貼られたスマートフォンを渡し、調査中はいつでもスマホリングを触りながらジェスチャを提案した。この時、我々は参加者にスマホリングから指を抜かないジェスチャを設計することを説明した。

3.2 参加者と収集したデータ数

19歳から24歳(平均20.95歳)までの20名の参加者が調査に参加した。3名がスマホリング経験者、5名が女性、全員が情報系を専攻していない大学生であった。13名が利き手ではない方の手でスマ

ホリングを操作しながらジェスチャを設計した。参加者には1,500円の謝礼が支払われた。1人の参加者につき23個のタスクを2種類の指の通し方でそれぞれジェスチャを収集し、20名の参加者から合計960個のジェスチャが収集された。

3.3 ジェスチャ選択

参加者から収集したジェスチャ案をコーディングし、人差し指と中指を通した場合でそれぞれ、各タスクごとに提案されたジェスチャを集計した。各タスクで集計されたジェスチャの中で最も多く提案されたジェスチャをユーザ定義型ジェスチャとして選択した。この時、複数のタスクに同じジェスチャが割り当てられる競合状態が発生した。我々は、発生した競合状態を、スマートフォンを操作する特定のシーンで、競合したタスクのいずれも使用する可能性がある場合に限り、競合状態を解消した。競合状態の解消は、先行研究と同じく[11]、集計値が小さい、つまり参加者によって該当のジェスチャが提案された数が小さいタスクのジェスチャを、2番目に多く提案されたジェスチャに再度割り当てることで行った。最終的に選択されたジェスチャ一覧を図2に示す。リングに中指を通した場合の最終的なユーザ定義型ジェスチャセットは12種類のジェスチャで構成され、人差し指をリングに通した場合でも同様に12種類のジェスチャで構成された。

4 ジェスチャ分類器の実装

設計したジェスチャを認識するスマホリング型デバイスのプロトタイプを実装する。スマホリングでの指の動作は圧力センサで取得し、機械学習を用いてジェスチャの分類を行う。我々は参加者からデータを収集し、交差検証によってプロトタイプの評価を行った。

4.1 ハードウェア構造

図3に我々が作成したスマホリング型デバイスのプロトタイプの外観と内部構造を示す。プロトタイプは8つの圧力センサ(Interlink Electronics, FSR 402)、天然ゴムシート、アクリル板(3mm)、市販のスマホリングで構成されている。圧力センサはスマホリングのプレート部の両面(表と裏)の四隅に設置され、スマホリングはアクリル板で作成した箱の中に設置する。箱はスマホリングのリング部だけが突出する構造となっている。ハードウェア構造の特徴としては、圧力センサが設置されたスマホリングのプレートとアクリル板の間にクッション性のある天然ゴムのシートを挟んでいることである。これにより微細な指の動作の感知を試みる。

スマホリングデバイスの各圧力センサの信号はデバイス外部で1kΩの炭素皮膜抵抗を經由しマイクロコンピュータ(Arduino Nano BLE 33)へ接

続され、マイクロコンピュータはプロトタイプから100Hz周期で信号を取得し、ホストコンピュータへと送信する。ホストコンピュータは送信された信号を、1.5秒間ごとのフレームに切り分け、各種分類を行う。開発したデバイスはGoogle Pixel 3aの背面に設置し、データ収集と解析にはGALLERIA XL7C-R36 11800H(Intel Core i7-11800H CPU with 16 GB memory)を用いた。

4.2 ジェスチャ分類手法

図4にデバイスで感知するセンサ値とその後の前処理と分類の流れを示す。圧力センサからの値はまず、正規化と濃淡画像化の前処理が行われる。生データの正規化は、各センサが取得した値ごとに、各値をそのセンサが取得した最大値で除算することで実装した。その後、正規化したデータを40×30の大きさの濃淡画像に変換する。

上記の前処理が行われたデータについて2つの分類器によってジェスチャ分類を行う。1つ目の分類器では日常動作の分類を行う。具体的にはデバイスから得られたデータがジェスチャを入力している動作なのか、日常動作なのかの2値分類を行う。この分類器でジェスチャ入力と分類されたデータは、2つ目の分類器で24種類のラベル(12種類のジェスチャ×2種類の指の通し方)に分類される。2つの分類器には、PythonのTensorflowライブラリのKerasを用いて、どちらも同様の層構造を持つCNNを実装した(図5)。

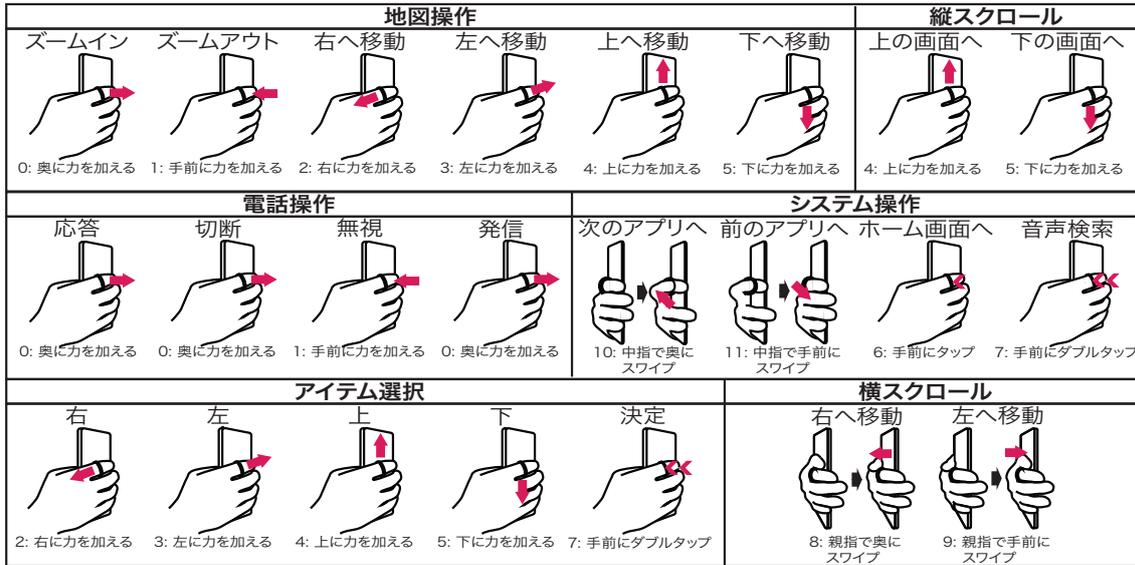
5 ジェスチャ分類精度の調査

開発したスマホリング型デバイスによるジェスチャ入力インタフェースの、ジェスチャ分類精度の確認のための調査を行う。

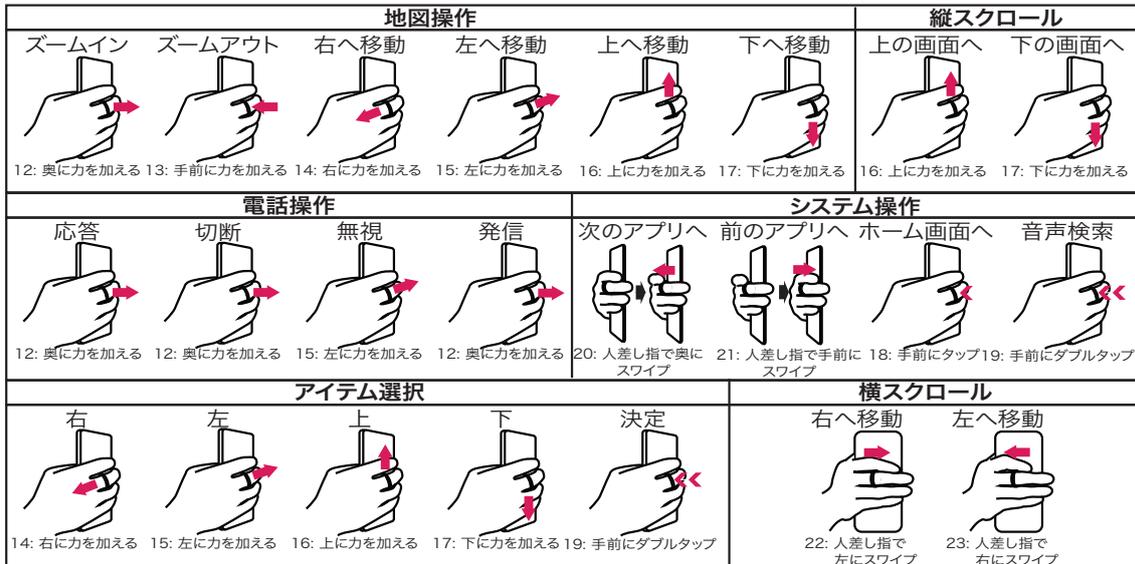
5.1 調査手順

ジェスチャデータの収集は研究室で実施された。参加者は椅子に座り、設計された24種類のジェスチャ(12種類のジェスチャ×2種類の指の通し方)について各ジェスチャにつき10回、全て右手で操作を行った。我々はジェスチャ入力中のデータとして、合計で2400個のデータを収集した。

スマホリングに指を通したスマートフォンの日常動作のデータには、スマホリングに中指を通した場合と人差し指を通した場合の2つの場合でそれぞれ、スマートフォン画面のタップ、スワイプ、何もしないの3つの操作をしている最中のデータを収集した。右利きの著者1名が、右手で各条件につき400回の入力を行うことで、合計2400個の日常動作のデータを収集した。



(a) 人差し指を通した場合のユーザ定義型ジェスチャセット



(b) 中指を通した場合のユーザ定義型ジェスチャセット

図 2. ユーザ定義型ジェスチャのアプローチで設計された 24 個のジェスチャ。人差し指と中指によるジェスチャは別ものとして扱っている。各イラストの上部がタスク名、下部がジェスチャ名である。

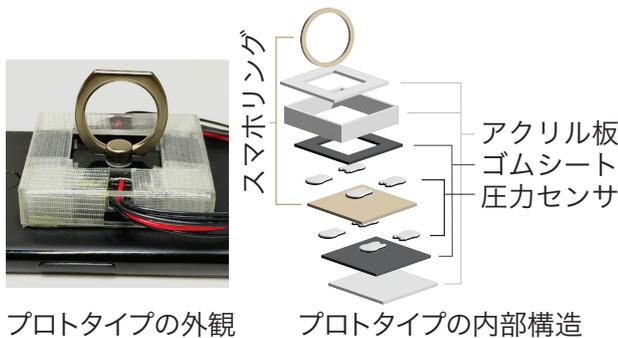


図 3. 開発するスマホリングデバイスの外観と構造

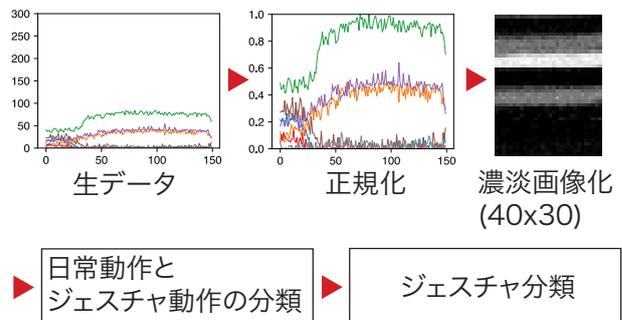


図 4. 信号処理と分類の流れ



図 5. CNN の層構造

表 1. 日常生活とジェスチャ入力を分類する分類器の精度.

分類ラベル	精度	再現率	F1 スコア
日常動作	0.986	0.996	0.991
ジェスチャ入力	0.996	0.985	0.991

5.2 参加者

21 歳から 33 歳（平均 24.1 歳）の 10 名の参加者が無償で実験に協力した。参加者全員が右利き、1 名が女性、全員が情報系を専攻している大学生もしくは大学院生であった。

5.3 結果：日常動作の分類

日常動作とジェスチャ入力を切り分ける分類器には、日常動作の 2400 個のデータ、調査の参加者から収集したジェスチャ入力中の 2400 個のデータの合計 4800 個のデータを使用した。このデータに対して 10-fold 交差検証を行った結果、99.1%の精度が検出された。それぞれのラベルの精度の詳細を表 1 に示す。

5.4 結果：ジェスチャの分類

ジェスチャの種類を分類を行う分類器には、参加者から収集した 2400 個のデータを参加者間学習と参加者内学習の 2 種類のデータの分割方法による交差検証を行った。表 2 にそれぞれの交差検証の精度の結果を示す。参加者間学習では、38.7%と低い精度であった一方で、参加者内学習では 92.7%と 9 割を超える精度が確認された。また、リングに人差し指を通したジェスチャより中指を通したジェスチャの方が高い精度が検出された。参加者内学習時の分類器の混同行列を図 6、各ラベルの精度の詳細を表 3 に示す。図 6 と表 3 のラベルの 0 から 23 の番号は、図 2 のジェスチャの各番号と互いに対応しており、0 から 11 はリングに人差し指を通したジェスチャ、12 から 23 が中指を通したジェスチャである。どちらの分類器も、1 つのデータを受け取り、前処理の後、予測結果が出力されるまでの所要時間は、12.6 ミリ秒であった。

6 利用例

スマホリングはスマートフォンの落下防止などの目的で使用され、本稿で提案したスマホリング型の

表 2. ジェスチャの種類を分類する分類器の全体の精度、人差し指をリングに通した場合の精度、中指をリングに通した場合の精度.

交差検証	平均精度	人差し指	中指
参加者間学習	0.387	0.300	0.447
参加者内学習	0.927	0.918	0.936

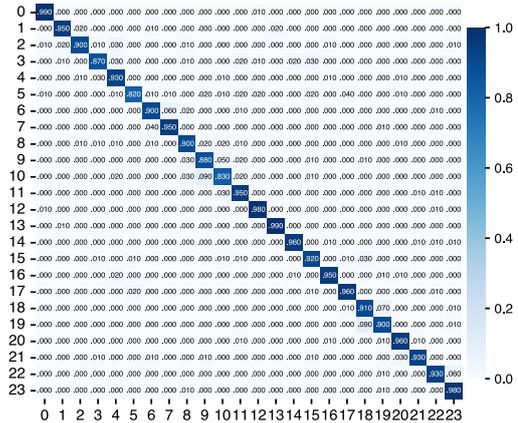


図 6. ジェスチャの種類を分類する分類器の参加者内学習時の各ラベルの精度。ラベル 0 から 11 は人差し指、12 から 23 は中指をリングに通したジェスチャ。

ジェスチャ入力インターフェースは、スマホリングの機能をそのままインタラクティブシステムにしたということに特徴がある。例えば、ユーザが目的地へと向かうため、スマートフォン上で地図アプリを操作する場合、我々が提案したデバイスを使用することで、スマートフォンの操作で必要だったピンチインやピンチアウトなどの複雑なジェスチャ入力無しに、片手で簡単な操作が可能となる。他にも、スマートフォン上でカメラアプリの操作では、上記と同様に撮影に必要な操作を片手でできることに加え、画面上からのカメラレンズが投影した景色の視認の妨げとなる撮影ボタンなどのインターフェースをスクリーン上から排斥することができ、ユーザはより理想的な写真を撮影することができるようなインターフェースを実現できる。このような屋外での利



(a) 野外で地図アプリを使用 (b) 室内で通知を削除

図 7. アプリケーションの利用例.

表 3. 参加者内学習時の分類器の分類精度。ラベル 0 から 11 は人差し指, 12 から 23 は中指をリングに通したジェスチャ。

ラベル	精度	再現率	F1 スコア
0	0.971	0.990	0.980
1	0.960	0.950	0.955
2	0.957	0.900	0.928
3	0.926	0.870	0.897
4	0.886	0.930	0.907
5	0.965	0.820	0.886
6	0.918	0.900	0.909
7	0.931	0.950	0.941
8	0.891	0.900	0.896
9	0.863	0.880	0.871
10	0.865	0.830	0.847
11	0.888	0.950	0.918
12	0.961	0.980	0.970
13	0.980	0.990	0.985
14	0.970	0.960	0.965
15	0.911	0.920	0.915
16	0.960	0.950	0.955
17	0.941	0.960	0.950
18	0.875	0.910	0.892
19	0.882	0.900	0.891
20	0.932	0.960	0.946
21	0.969	0.930	0.949
22	0.949	0.930	0.939
23	0.907	0.980	0.942

用時は、傘を差す場合や買い物袋を持っている場合など、片手が使用できない場合に特に効果的である。

提案手法は屋外での利用だけでなく、室内でのスマートフォン操作時にも有効であると考えられる。例えば、ユーザが自宅のソファやベッドなどの上で寝そべりながらスマートフォンを頭上に掲げ、SNS やインターネットサイトを閲覧している場合 (図 7 (b)), 従来のタッチインタラクションでは、ユーザは手元が不安定になり、スマートフォンを顔面などに落下させる危険性がある。このような場合でも、我々が提案したデバイスに指を通してスマートフォンを操作することで、上記のような落下を防止しながら快適にスマートフォンを操作できるだろう。他にも、スマートフォンの画面上部は指が届きにくい、このような領域に表示される通知バナーなどへのショートカットとしても使うことができるだろう。

7 議論と今後の課題

本稿では、スマホリング型のジェスチャ入力インタフェースを提案した。本稿で行った調査では、各分類器の分類精度は全て 9 割を超えた。しかし、本

調査では安定した環境下で収集したデータのみを使用しており、実際にユーザがスマートフォンを使用すると想定される、提案デバイスへの干渉が大きいと予想される環境下での調査が必要である。乗り物の乗車中や歩行中などがこの環境として考えられ、これらの様なスマホリング型デバイスへのノイズが多いことが予想される環境での調査が行われるべきである。同様に、本稿で実施した日常動作とジェスチャ入力の分類は、ジェスチャ入力のデータには 10 名の参加者、日常動作には別の 1 名のデータを学習データに使用し高精度の分類を確認した。しかし、ジェスチャの種類の分類で参加者間学習の精度が低かったことから、データには個人差が生じることが推察され、同一の参加者から収集したデータによる調査が行われるべきである。

本稿で開発したプロトタイプは、デバイス単体では機能せず、多くの計算処理を外部のコンピュータで行った。この構成は本研究のコンセプトを検討するという本稿の目的を達成するには十分であったが、実際に使用するには、これら全てをスマホリング内に収める必要がある。そのためには、電源供給の問題やスマートフォンやマイコンなどの計算能力が比較的小さな計算機でもリアルタイムにジェスチャを分類するためのシステムを検討する必要がある。また、本研究で実施した調査では、ユーザが右手でジェスチャの入力をしたデータのみを解析の対象とした。しかし、スマートフォンは左右どちらの手でも使用されるため、左右どちらの手でこのデバイスを使用した場合でも、高い精度でジェスチャの分類が行われる必要がある。本実験では、リングに通した指の違いによって、一見似たようなジェスチャでも 9 割を超える精度で分類が行われ、これらの中で特徴に差が生じたことが推察される。同様に、デバイスを使用する手の違いによってデバイスが取得する信号の特徴に差が出るのが予測される。そのため、左右の手で使用したデータの解析と検討が必要だろう。

8 結論

本稿では、ジェスチャ入力を実現するスマホリング型のインタフェースの提案と実装を行った。まず、スマートフォンの操作を目的としたスマホリング向けのジェスチャをユーザ定義型ジェスチャのアプローチで設計した。その後、実際にデバイスを開発し、あわせて日常動作の分類と、ジェスチャの種類を分類する 2 つの分類器を CNN による機械学習技術により開発した。参加者を集めて実際にデータを収集する調査を行い、そのデータを用いてそれぞれの分類器で学習させた。その結果、日常生活の分類精度は 99.1%、ジェスチャの種類分類精度は 92.7% であった。

参考文献

- [1] C. Appert and S. Zhai. Using Strokes as Command Shortcuts: Cognitive Benefits and Toolkit Support. In *In Proc. CHI '09*, p. 2289–2298, New York, NY, USA, 2009. Association for Computing Machinery.
- [2] K. Fukahori, D. Sakamoto, and T. Igarashi. Exploring Subtle Foot Plantar-Based Gestures with Sock-Placed Pressure Sensors. In *In Proc. CHI '15*, p. 3019–3028, New York, NY, USA, 2015. Association for Computing Machinery.
- [3] S. Gupta, D. Morris, S. Patel, and D. Tan. SoundWave: Using the Doppler Effect to Sense Gestures. In *In Proc. CHI '12, CHI '12*, p. 1911–1914, New York, NY, USA, 2012. Association for Computing Machinery.
- [4] C. Harrison and S. E. Hudson. Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile Finger Input Surfaces. In *In Proc. UIST '08*, p. 205–208, New York, NY, USA, 2008. Association for Computing Machinery.
- [5] C. Harrison and S. E. Hudson. Abracadabra: Wireless, High-Precision, and Unpowered Finger Input for Very Small Mobile Devices. In *In Proc. UIST '09*, p. 121–124, New York, NY, USA, 2009. Association for Computing Machinery.
- [6] M. A. Nacenta, Y. Kamber, Y. Qiang, and P. O. Kristensson. Memorability of Pre-Designed and User-Defined Gesture Sets. In *In Proc. CHI '13*, p. 1099–1108, New York, NY, USA, 2013. Association for Computing Machinery.
- [7] C. Rendl, D. Kim, P. Parzer, S. Fanello, M. Zirkl, G. Scheipl, M. Haller, and S. Izadi. FlexCase: Enhancing Mobile Interaction with a Flexible Sensing and Display Cover. In *In Proc. CHI '16*, p. 5138–5150, New York, NY, USA, 2016. Association for Computing Machinery.
- [8] J. Ruiz, Y. Li, and E. Lank. User-Defined Motion Gestures for Mobile Interaction. In *In Proc. CHI '11*, p. 197–206, New York, NY, USA, 2011. Association for Computing Machinery.
- [9] A. D. Wilson. PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System. In *In Proc. UIST '05*, p. 83–92, New York, NY, USA, 2005. Association for Computing Machinery.
- [10] A. D. Wilson. Robust Computer Vision-Based Detection of Pinching for One and Two-Handed Gesture Input. In *In Proc. UIST '06*, p. 255–258, New York, NY, USA, 2006. Association for Computing Machinery.
- [11] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. User-Defined Gestures for Surface Computing. In *Proc. CHI '09*, p. 1083–1092, New York, NY, USA, 2009. Association for Computing Machinery.
- [12] X. Xu, J. Gong, C. Brum, L. Liang, B. Suh, S. K. Gupta, Y. Agarwal, L. Lindsey, R. Kang, B. Shahsavari, T. Nguyen, H. Nieto, S. E. Hudson, C. Maalouf, J. S. Mousavi, and G. Laput. Enabling Hand Gesture Customization on Wrist-Worn Devices. In *In Proc. CHI '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- [13] A. Yadav, A. K. Eady, S. Nabil, and A. Girouard. JoyHolder: Tangible Back-of-Device Mobile Interactions. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces, ISS '19*, p. 343–346, New York, NY, USA, 2019. Association for Computing Machinery.
- [14] W. Yamada, H. Manabe, and D. Ikeda. Cam-TrackPoint: Camera-Based Pointing Stick Using Transmitted Light through Finger. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, UIST '18*, p. 313–320, New York, NY, USA, 2018. Association for Computing Machinery.

ReflecTouch: 角膜反射像を利用したスマートフォンの把持姿勢推定

張 翔* 池松 香† 加藤 邦拓‡ 杉浦 裕太*

概要. ユーザがスマートフォンをどのように把持しているかを検出することで、持ち方に応じて画面表示や操作ボタンの位置を自動的に切り替えるなど、ユーザインターフェースを最適化することができる。本研究では、スマートフォンのフロントカメラで撮影された顔写真を使用し、ユーザの角膜に反射したスマートフォンの画面の反射像を機械学習で分類することで、把持姿勢を推定する手法である ReflecTouch を提案する。スマートフォンの画面は光を発するため、顔の正面方向でスマートフォンを持つと画面の反射像が角膜に映る。画面上の指を置いている部分は影となるため、その部分の角膜反射像は欠ける。その際、スマートフォンの持ち方によって像の欠け方が異なるため、把持姿勢の推定が可能になると考えた。本稿では、ReflecTouch の識別精度を検証する実験を行い、その結果や制約について議論するとともに、本手法を利用したアプリケーションを報告する。

1 はじめに

ユーザがスマートフォンがどのように把持しているか（把持姿勢）はスマートフォンの操作において重要な要素の1つである。先行研究 [6, 22] では、片手持ちか両手持ちか、操作する手の角度、操作に使用する指の数などの把持姿勢がスマートフォンの操作性や効率に影響を与えることを指摘しており、ユーザがどのようにスマートフォンを操作しているのかを推定することでスマートフォンの利便性の向上が可能となる。例えば、ユーザはスマートフォンを片手で持ち、親指で操作することが多いが [5, 11]、メニューボタンなどが画面上で親指から最も遠い角の付近に配置されてしまうことがある。また、把持姿勢に応じてソフトウェアキーボードのタップの仕方が異なる [2] ことが明らかにされている。これらに対し把持姿勢を推定することで、画面上のボタンやキーボードをユーザがタップしやすい位置に配置したり、サイズを変更したりすることが可能になる。

把持姿勢を推定する手法として、静電容量センサなどのセンサを外付けする手法や [4, 25]、ジャイロセンサなどのスマートフォンに内蔵されたセンサを使用する手法 [7, 19] が提案されている。しかしこれらの手法ではセンサを外付けする必要があり実用化しにくいことや、推定モデルがデバイスに依存すること、推定できる把持姿勢が把持手のみなどの3-4種類に限られるといった、いずれかの課題があった。

本研究では、スマートフォンの内蔵フロントカメラを使用してユーザの顔写真を撮影することで、図1(右)に示す6つの把持姿勢を推定する手法を提案する。スマートフォンの画面は光を発するため、顔の正面方向でスマートフォンを持つと画面の反射像がユーザの角膜に映るが、画面上の指を置いている部分は影となるため該当箇所の角膜反射像は欠け、その欠け方は把持姿勢に応じて変化する。このことを利用し、フロントカメラで撮影した顔写真から瞳の部分を取り取り、畳み込みニューラルネットワーク(CNN)を用いた機械学習モデルに瞳の領域の画像を入力することで、把持姿勢の識別を行う。13名の実験参加者を対象にして本提案手法の精度を検証す

* 慶應義塾大学

† ヤフー株式会社

‡ 東京工科大学

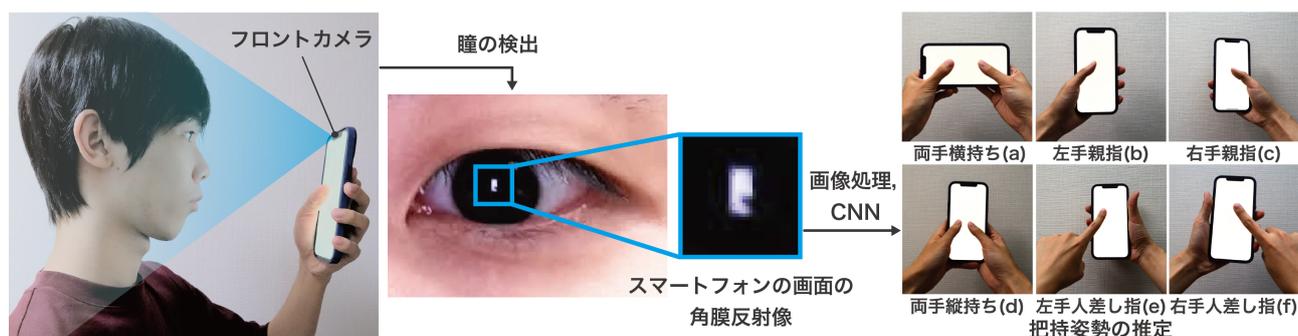


図 1. 把持姿勢推定の処理概要および本稿で扱う 6 種類の把持姿勢。

る実験を行った結果、85%の精度で識別が可能であることを確認した。なお、本稿は EC 2021 [28] と ACM CHI 2022 [27] での発表内容に議論を加筆したものである。

2 関連研究

2.1 把持姿勢の推定手法

スマートフォンの把持姿勢を推定する方法として、スマートフォンに追加のセンサやハードウェアを取り付ける方法や、内蔵されたセンサのデータを利用する方法が提案されている。センサを外付けするアプローチでは、小型かつ軽量の静電容量センサを用いた手法が多い [4, 13, 21, 25]。また、スマートフォンに取り付けたスピーカとマイクを利用する手法 [18] や、スマートフォンのフロントカメラにプリズムミラーを取り付けることでステレオビジョンを実現し、奥行き情報を推定することで把持姿勢を推定する手法 [26] が提案されている。

スマートフォンの内蔵センサを利用した手法として、加速度センサやジャイロセンサといった端末の動きを計測できるセンサを利用した研究が提案されている。Goel ら [7] は内蔵バイブレータで端末を振動させ、その際の振動の特徴を加速度センサで計測することで把持姿勢を推定する手法を提案した。また、Park と Ogawa [19] は加速度センサとジャイロセンサからの情報をサポートベクタマシン (SVM) に入力することで把持姿勢を推定する手法を提案した。Avery ら [1] はスマートフォンのロックを解除するためのパスワード入力時の端末の動きを測定することで、把持姿勢を推定する手法を開発した。また、提案手法と同様に内蔵フロントカメラを使用して把持姿勢の推定を行った研究としては、フロントカメラからユーザの体のシルエットを検出し、スマートフォンとユーザの体の相対的な傾きを計測することで把持姿勢を推定する研究 [15] がある。

これらの研究では、推定可能な情報が把持手 (両手・右手・左手) に限定されていたり [7, 15]、特定のタイミング (画面ロック解除時など) でしか利用できない [1] といった制約がある。提案手法は、タッチ操作に使用する指と把持手の検出を、顔写真が取得可能な任意のタイミングで行なうことができる上、追加のセンサやハードウェアを必要としないという利点がある。

2.2 角膜反射像の撮影と利用

角膜反射像を撮影し、コンピュータとのインタラクションに利用する研究が行われている。Schneider ら [20] はユーザの正面に置かれたミラーレス一眼カメラを使って角膜に写ったユーザの手を撮影することでスマートフォンの周辺領域での操作を可能にする研究を行った。

また、角膜反射像を利用して視線を推定する研究も行われている [17, 23]。スマートフォンのフロントカメラを使って角膜反射像を撮影した研究としては、スマートフォンのフロントカメラで角膜反射像を撮影し、瞳の中で角膜反射像がどの位置にあるかを測定することで視線推定を行う研究がある [9]。また、角膜反射像ではないが、ユーザの着用しているサングラスに反射したスマートフォンの画面とユーザの手をフロントカメラで撮影することでスマートフォンのインタラクションに用いる研究 [8] がある。

3 角膜反射像を利用した把持姿勢の推定

スマートフォンを操作する際、スマートフォンの画面は光を発しているため、ユーザの角膜にはスマートフォンの画面の形をした角膜反射像が映る。その際に、ユーザが指を置いている部分は図 1 に示すように影となり欠ける。反射像の欠け方は持ち手や操作している指などの把持姿勢によって異なるため、角膜反射像から把持姿勢を推定することが可能であると考えた。

図 1 に提案手法の処理の流れを示す。ユーザが画面をタップしたタイミングで、内蔵フロントカメラを使用してユーザの顔写真を取得する。次に、取得した画像からスマートフォンの画面の角膜反射像が写っている部分をクロップし、その画像を CNN モデルに入力することで把持姿勢の推定を行う。推定結果を利用することで把持姿勢に応じて画面表示などを最適化させることが可能となる。

一般に、スマートフォンをタッチ操作する際に、図 1 (右) のいずれかに分類される把持姿勢となる場合が多いことが知られている [6]。そのため、本研究ではこの 6 種類の把持姿勢を対象として把持姿勢の推定を行う。

本手法は、追加のセンサを必要とせずに 6 種類の把持姿勢を推定することができる上、多くのスマートフォンでは画面上部の位置にフロントカメラがあるため、機種の違いによる影響が小さいというメリットがある。

4 実装

提案手法の実現可能性を検証するため、各把持姿勢にてスマートフォンのフロントカメラからユーザの顔写真を収集するための iPhone アプリを実装し、アプリを用いて収集した画像を使って CNN を利用した学習モデルの実装と把持姿勢の推定を行った。

4.1 顔写真の収集

顔写真を収集するために、それぞれの把持姿勢でスマートフォンを持った際の顔写真を撮影する iPhone アプリを開発した。本アプリケーションでは各把持姿勢につき 12 枚、合計 72 枚を内蔵のフロン

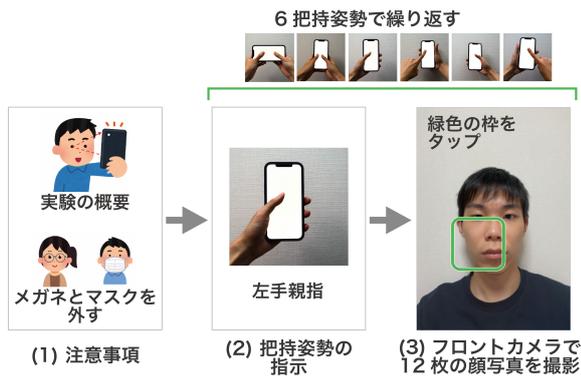


図 2. 顔写真収集用のアプリケーションの流れ.

トカメラを利用して撮影する。アプリケーションを起動すると、実験についての注意事項 (図 2 (1)) が表示された後、それぞれの把持姿勢でスマートフォンを持つように指示する画面が表示される (図 2 (2))。その後、撮影画面 (図 2 (3)) が表示される。撮影画面上には緑色の矩形領域 (ターゲット) が表示され、領域内をタップした際に写真が撮影される。6 種類全ての把持姿勢でのタップタスクが完了するとアプリケーションは終了する。なお、顔写真を撮影する際には画面が白くフラッシュするが、複数の画面輝度での顔写真収集を行なうため、ひとつの把持姿勢につき画面輝度が 100%, 85%, 70% の状態でそれぞれ 4 枚づつ撮影される。

4.2 画像の前処理

収集された顔写真に対して、CNN モデルの学習と識別に利用するための前処理を行った。ここでは、図 3 に示すように顔写真から瞳をクロップし、画像の輝度やコントラストの調整、データオーグメンテーションを行った。

まず、撮影された顔写真について、顔写真から瞳が写っている領域・瞳が写っている領域から角膜反射像がの領域まで Google の MediaPipe および OpenCV を用いて二段階でクロップした。なお、瞳から角膜反射像をクロップする際には、角膜反射像の大きさは瞳の領域の 5–20% の大きさであると想定し、角膜反射像である推定された領域がこの範囲外だった場合にはエラーとしてデータセットから取り除いた。その後、クロップした画像の明るさをデータセット全体で揃えた上で、スマートフォンの像で指の影となっている部分を強調するためにコントラストを上げた。次に、学習データとテストデータに 3:1 の割合で分けた後、学習用データに対してデータの量を増やすためにデータオーグメンテーションを実施した。入力画像の角度や大きさの僅かな違いに対して CNN モデルが影響されないように元画像の角度やサイズをランダムに変化させた画像を生成して学習用データに追加した。

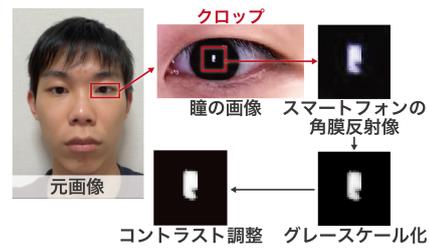


図 3. 前処理の流れ.

4.3 モデルの学習

前処理後の画像を用いて CNN モデルの学習を行った。モデルの実装は Python と Keras を用いた。両目 (二枚) の角膜反射像は 80×80 にリサイズされた上で、それらを重ねて $80 \times 80 \times 2$ の三次元テンソルとして CNN モデルに入力した。CNN モデルのレイヤ図は付録の図 5 の通りである。

5 データ収集実験

提案手法の実現可能性を検証するためデータ収集実験を行った。実験には 13 名の実験参加者が参加し、平均年齢は 35.1 歳 (標準偏差は 17.2 歳) であった。なお、性別については女性が 5 名、男性が 8 名、利き手については右利きが 10 名、左利きが 1 名、両利きが 2 名、瞳の色については黒色が 10 名、黒茶色が 1 名、茶色が 2 名であった。

5.1 実験手順

参加者は 4.1 節のアプリケーションを使用して、それぞれの把持姿勢で画面上のターゲットをタップするタスクを実施した。なお、実験中はメガネやコンタクトレンズおよびマスクは外すように指示をした。収集する画像の画素数を揃えるために 12Mpx のフロントカメラが搭載されている iPhone 11 および iPhone 12 系の端末を使用した。また、他の光源による影響を避けるため実験は室内かつモニターなどが付近に無い環境で行うように指示した。本実験では合計で $13 \text{ 名} \times 6 \text{ 把持姿勢} \times 12 \text{ 枚} = 936 \text{ 枚}$ の画像を収集した。

収集した画像に対して 4.2 節で述べた前処理を行い、エラー画像を除いた結果 81% の画像が残った。その後、前処理がされた画像を使って 4.3 節で述べたモデルの学習およびその精度の評価を行った。なお、データセットを学習用データとテストデータに分割する際には、それぞれの参加者のデータが両方のデータに含まれるように分割する方法 (以下、全ユーザ学習) と、Leave-one-subject-out 交差検証を行う方法 (以下、他ユーザ学習) の 2 種類の方法で行った。いずれの場合でも学習用データとテストデータの割合は 3:1 とした。

表 1. 全ユーザ学習, 6 姿勢での個別の識別結果.

	適合率	再現率	F1 値
両手横持ち	100%	98%	99%
両手縦持ち	82%	86%	84%
左手親指	88%	73%	80%
左手人差し指	85%	88%	87%
右手親指	81%	84%	82%
右手人差し指	82%	87%	84%

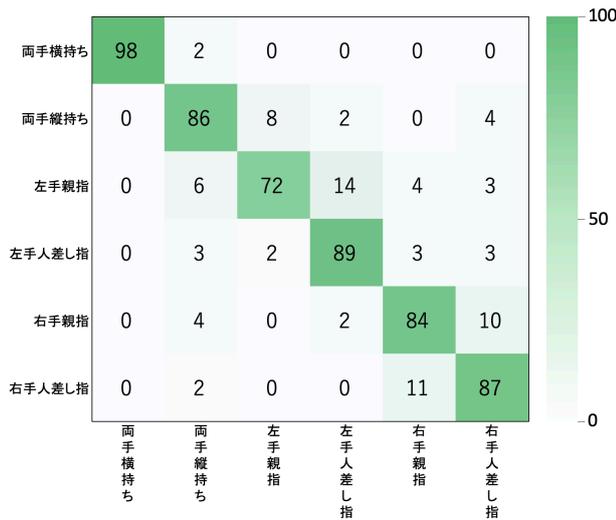


図 4. 全ユーザー学習での 6 姿勢の識別結果の混合行列.

5.2 結果と考察

データセットの分割方法ごとの識別精度について、他ユーザ学習は 41%、全ユーザ学習は 85%であった。また、全ユーザ学習での把持姿勢ごとの識別精度を表 1 と図 4 に示す。

図 4 から、右手親指と右手人差し指、左手親指と左手人差し指の把持姿勢が混同されやすいことがわかる。提案手法では、スマートフォンの画面の角膜反射像のうち指に隠れている部分を特徴として把持姿勢推定を行っているが、人差し指と親指のどちらの指で操作するかについては角膜反射像での指の部分の欠け方に大きな影響を及ぼさないため、混同をしてしまったと考えられる。

データの分割方法に着目した際、他ユーザ学習は全ユーザ学習と比べて精度が大幅に低下した。これは、スマートフォンを持つ位置や撮影する環境などに加え、ユーザごとのスマートフォンの使い方の違い、瞳の色や光の反射の仕方などの生物学的な個人差が影響したと予想される。

また、手の大きさや指の太さの個人差も親指と人差し指を混同や他ユーザ学習における精度低下の原因の 1 つとなりうる。

6 転移学習を用いた把持姿勢推定

他ユーザ学習の精度は 41%と低く実用的ではないため、全ユーザ学習のモデルの利用が望ましい。しかし、提案手法の実利用を検討した際、事前準備（データ収集）がユーザの負担となる可能性がある。前節で述べた全ユーザ学習の場合、提案手法の利用の前に 72 枚の写真を取得し、学習モデルを作る必要がある。5 節のデータ収集実験の 1 名あたりの所要時間は 6 分ほどであり、これと同程度の時間を要すると予想される。

そこで、機械学習の一つの手法である転移学習を利用することで、より少ない写真数で全ユーザ学習と同程度の識別精度を実現できないかを検討した。

6.1 実験手順

この実験は前節で収集した画像を利用した。実験の流れを図 6 に示す。まず、13 名のデータの中から 12 名分の参加者のデータを使って前章と同様の学習モデル (1) を作成する。次に学習モデルの作成に用いなかった 1 名のデータを使って転移学習を行い新しい学習モデル (2) を作成する。転移学習では 12 名のデータを使って作成された学習モデル (1) をベースにして、一部のレイヤをリセットした上で残りの 1 名のデータを用いて再学習を行った。なお、その他の層についてはパラメータを固定し、再学習を行わないようにした。画像枚数に応じたモデル精度の変化を調べるため、転移学習に使用した画像の枚数を把持姿勢ごとに 1 枚から 7 枚まで 1 枚ずつ増やしつつ、各枚数条件でモデルを作成し精度を検証した。上述の処理を、転移学習に使用する参加者のデータを入れ替えて 13 名分行い、識別精度の平均値を計算して最終的な結果とした。

6.2 結果と考察

識別結果を図 7 に示す。把持姿勢あたりの再学習に使用した画像数に着目すると、把持姿勢ごとの画像数が 1 枚から 4 枚までの間は識別精度が 20%上がったのに対して、4 枚から 7 枚までの間の識別精度の向上は 4%のみとなった。また、画像数が多くなると精度が全ユーザ検証での精度に近づく。

この転移学習の手法を利用する場合、把持姿勢ごとに 4 枚の画像を用いると 76%の精度を得ることができる。全ユーザ学習の精度より 9%低く、実利用を考えると十分な精度とは言い難いものの、撮影する顔写真の枚数は全ユーザ学習の 1/3 となるため、導入のための所要時間は 2 分程度に短縮できる。また、転移学習の場合、すでに学習されたモデルが用意されていれば 1 人分のデータを学習するだけであるので、計算コストが低くなり、モデルの学習に必要な処理時間も短くなる。

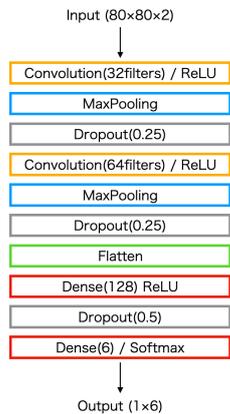


図 5. CNN モデルのレイヤ図.

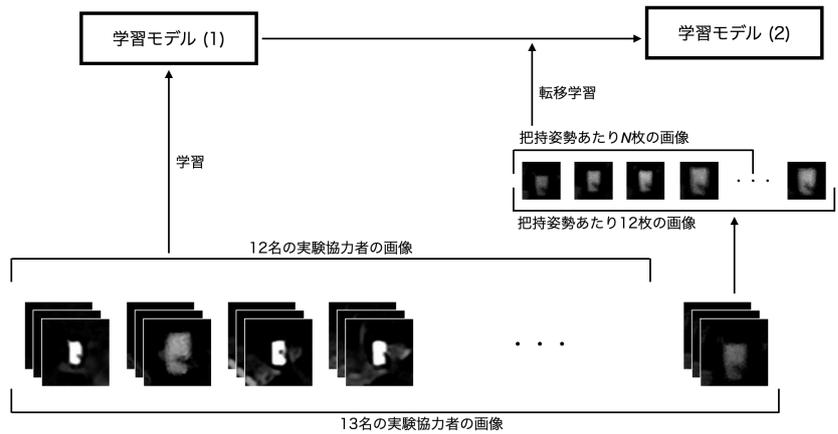


図 6. 転移学習を用いた把持姿勢推定の実験の概要.

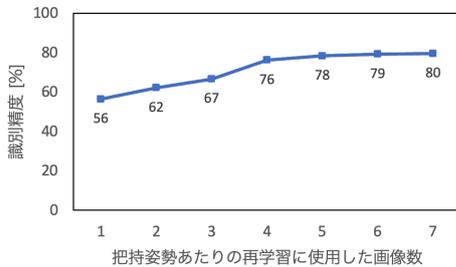


図 7. 転移学習の結果.

7 応用例

把持姿勢を推定することで、画面上の表示コンテンツや操作ボタンの位置をユーザの把持姿勢に合わせて自動で変更するユーザインタフェースを実現できる。一般に、画面上の操作しやすい領域はユーザの把持姿勢によって異なり [24]、特に、画面サイズの大きいスマートフォンではボタンなどの位置がアプリの使いやすさに与える影響は大きいと考えられる。例えば図 8 (1) は、ユーザが両手を使ってスマートフォンを操作している際、荷物を持つため一方の手がふさがり、片手での操作に切り替わった場合の例である。把持姿勢の切り替わりを検出することで、GUIを片手で操作しやすい位置に自動で再配置することができる。

図 8 (2) は GUI の自動再配置をソフトウェアキーボードに適用した例である。スマートフォンを両手持ちから片手持ちに切り替えた際に、キーボードを画面片側に寄せることでキーを入力しやすくしている。この他に、両手持ちの場合には QWERTY 配列のキーボードを使用し、片手持ちに切り替わったことが検知された場合には、フリック入力用のキーボードに切り替えるなどの応用が可能である。

また先行研究 [3] によると、スマートフォンの長時間の利用によって手指に痛みや痺れが生じるなど



図 8. 把持姿勢推定を用いた応用例.

の影響がある。そのため、ユーザの把持姿勢を記録し、長時間同じ姿勢での操作が続いた際に、通知を表示し持ち方変えるよう促すことで、これらの疾患を予防できると期待される (図 8 (3))。

さらに角膜反射像を使った本手法の将来的な応用例として、手のひらや指の関節で画面をタップした際に画面が手に覆われる面積が大きくなることの利用が考えられる。先行研究 [14, 10, 12] でも提案されているタップした手の部位に応じたインタラクションや HandSee [26] で提案されているような画面上の空中でのジェスチャを識別できると期待される。

8 議論

8.1 バッテリーの消費とプライバシーの保護

提案手法では、把持姿勢の推定のたびにスマートフォンのフロントカメラによる撮影を自動で行うた

め、バッテリーを多く消費してしまう上、プライバシーに対する懸念が生じる。把持姿勢の推定回数を減らすことで、これらの課題に対処できると考える。具体的には、スマートフォンのロックを解除した際や、アプリを起動した際、キーボードを表示した際など、把持姿勢が変わりやすいタイミングでの推定や、端末の角度が大きく変化したタイミングでの推定を検討している。また、現状の実装では、把持姿勢推定時に撮影した顔写真をサーバにアップロードし、サーバ上で角膜反射像部分をクロップと把持姿勢の推定している。個人識別のリスクを減らすため、これらの処理をスマートフォン内で完結させるなどの改善を検討する必要がある。

8.2 複数の手法を組み合わせた把持姿勢推定

より高い精度で把持姿勢の推定を行うためには、複数の推定手法を組み合わせたことが有効であると考えられる。例えば5.2節で述べたように、提案手法では親指操作(図1b,c)と人差し指操作(図1e,f)の識別が混同されやすいという結果が得られた。(図1b,c)および(図1e,f)はいずれも指の向きは同一であるが、スマートフォンの把持手が異なる。

そのため、把持手の検出を高精度に行える手法(例えば, HandyTrak [15] など)と組み合わせることで、この課題を解決できる可能性があると考えられる。HandyTrakの手法は、推定できる把持姿勢が3種類のみであるものの、スマートフォンの持ち手を高精度で推定可能であり、かつ単一のセンサ(カメラ)のみを使用するため、提案手法と併用することで把持姿勢の識別精度の改善が可能だと考えられる。

一方、本手法や HandyTrak [15] など、フロントカメラを使用した識別手法の場合、太陽光など強い光の当たる環境や、明かりがなく極端に暗い環境などでは、顔や体を適切に撮影することが困難になる。こうした場合には、環境光の影響を受けないジャイロセンサや加速度センサを使った手法 [1, 7, 16] や、タッチセンサを使った手法 [7, 16] などを利用することが有効になる。ユーザを取り巻く環境やユーザの姿勢を推定し、これらの手法を適切に切り替えることで、様々な状況下における高精度での把持姿勢の推定が可能となる。

8.3 メガネ・コンタクトレンズ着用の影響

5節では、メガネやコンタクトレンズによるスマートフォンの画面の光や環境光の反射を考慮し、実験参加者にそれらを外すよう指示した上で実験を行った。本節では実際にメガネ・コンタクトレンズの着用が、提案手法の推定精度に及ぼす影響について検証を行った。実験はメガネを装着した参加者9名、およびコンタクトレンズを装着した参加者6名に対して、5節と同様の流れで実施した。なお、この本実験で使用したアプリは4.1節のアプリを改良し、



図9. (左) 裸眼の場合、(右) メガネ(ブルーライトカットのレンズ)を装着した場合の例。

タップ入力ごとにターゲットの位置をランダムに変更する、撮影画面のカメラの映像を非表示するなどの変更を加えたものを使用した。実験結果は、メガネを装着していた場合の識別精度が24%、コンタクトレンズを装着していた場合の識別精度が64%となった。実験条件が異なるため5節の結果と単純比較はできないが、裸眼の場合と比べて精度が低くなる傾向が見られた。特にメガネを装着していた場合、角膜反射像を検出できた画像は全体の18%(5節では81%)となり、検出率が著しく低下した。図9に示す通り裸眼の場合に比べて環境光やスマートフォンの画面の光がメガネに反射してしまい、瞳に映る反射像が認識しづらくなっていることがわかる。以上より、特にメガネを装着したユーザに対しては、提案手法による把持姿勢の推定が困難であることがわかった。ユーザがメガネを装着している場合には、8.2節のように、他の把持姿勢推定手法と組み合わせることが解決策の1つとなりうる。

8.4 瞳の色の影響

5節の実験の参加者は日本人と中国人であり、瞳の色は黒色・黒茶色・茶色のいずれかであった。瞳の色がより明るいユーザの場合、角膜反射像と瞳の他の領域のコントラストが低下し、推定精度が低下する可能性がある。そこで、明るい瞳の色をもつユーザを対象にした提案手法の初期的な検証としてライトグレーの瞳の色の実験参加者(フランス人)を対象に、8.3節と同様のアプリを用いてデータ収集実験を実施した。合計で360枚の画像(各把持姿勢につき60枚)を収集し、モデルの実装を行ったところ、推定精度は82%であり、明るい瞳の色を対象にした場合でも把持姿勢は識別可能であるといえる。今後は他の瞳の色を含めた、より大人数での実験を行い、提案手法の適用可能性を調査する。

9 結論

本研究ではスマートフォンのフロントカメラで撮影された顔写真を使用し、スマートフォンの画面の角膜反射像をCNNで分類することで、6つの把持姿勢を識別するReflecTouchを提案した。13名の実験参加者を対象にして提案手法の精度を検証する実験を行い、85%の精度で識別できた。

謝辞

本研究の一部は JST さきがけ (課題番号: JP-MJPR2134) の助成を受けたものである。

参考文献

- [1] J. Avery, D. Vogel, E. Lank, D. Masson, and H. Rateau. Holding Patterns: Detecting Handedness with a Moving Smartphone at Pickup. In *Proceedings of the 31st Conference on l'Interaction Homme-Machine*, New York, NY, USA, 2019. Association for Computing Machinery.
- [2] S. Azenkot and S. Zhai. Touch behavior with different postures on soft smartphone keyboards. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, MobileHCI '12, pp. 251–260, New York, NY, USA, Sept. 2012. Association for Computing Machinery.
- [3] S. Berolo, R. P. Wells, and B. C. Amick. Musculoskeletal Symptoms Among Mobile Handheld Device Users and Their Relationship to Device Use: A Preliminary Study in a Canadian University Population. *Applied Ergonomics*, 42:371–378, Jan 2011.
- [4] L.-P. Cheng, H.-S. Liang, C.-Y. Wu, and M. Y. Chen. iGrasp: Grasp-Based Adaptive Keyboard for Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 3037–3046, New York, NY, USA, 2013. Association for Computing Machinery.
- [5] R. Eardley, A. Roudaut, S. Gill, and S. J. Thompson. Understanding Grip Shifts: How Form Factors Impact Hand Movements on Mobile Phones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, p. 4680–4691, New York, NY, USA, 2017. Association for Computing Machinery.
- [6] R. Eardley, A. Roudaut, S. Gill, and S. J. Thompson. Investigating How Smartphone Movement is Affected by Body Posture. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, No. Paper 202 in CHI '18, pp. 1–8, New York, NY, USA, Apr. 2018. Association for Computing Machinery.
- [7] M. Goel, J. Wobbrock, and S. Patel. GripSense: Using Built-in Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, p. 545–554, New York, NY, USA, 2012. Association for Computing Machinery.
- [8] J. Grubert, E. Ofek, M. Pahud, M. Kranz, and D. Schmalstieg. GlassHands: Interaction Around Unmodified Mobile Devices Using Sunglasses. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, p. 215–224, New York, NY, USA, 2016. Association for Computing Machinery.
- [9] M. X. Huang, J. Li, G. Ngai, and H. V. Leong. ScreenGlint: Practical, In-situ Gaze Estimation on Smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2546–2557. Association for Computing Machinery, New York, NY, USA, May 2017.
- [10] Huawei. Knuckle screenshots, 2019.
- [11] K. Ikematsu, H. Oshima, R. Eardley, and I. Sii. Investigating How Smartphone Movement is Affected by Lying Down Body Posture. *Proc. ACM Hum.-Comput. Interact.*, 4(ISS), nov 2020.
- [12] K. Ikematsu and S. Yamanaka. ScraTouch: Extending Interaction Technique Using Fingernail on Unmodified Capacitive Touch Surfaces. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(3), sep 2020.
- [13] K.-E. Kim, W. Chang, S.-J. Cho, J. Shim, H. Lee, J. Park, Y. Lee, and S. Kim. Hand Grip Pattern Recognition for Mobile User Interfaces. In *Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence - Volume 2*, p. 1789–1794. AAAI Press, 2006.
- [14] H. V. Le, T. Kosch, P. Bader, S. Mayer, and N. Henze. PalmTouch: Using the Palm As an Additional Input Modality on Commodity Smartphones. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, 2018.
- [15] H. Lim, D. Lin, J. Tweneboah, and C. Zhang. HandyTrak: Recognizing the Holding Hand on a Commodity Smartphone from Body Silhouette Images. pp. 1210–1220. ACM, 10 2021.
- [16] M. Löchtefeld, P. Schardt, A. Krüger, and S. Boring. Detecting users handedness for ergonomic adaptation of mobile user interfaces. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*, MUM '15, pp. 245–249, New York, NY, USA, Nov. 2015. Association for Computing Machinery.
- [17] C. Nitschke, A. Nakazawa, and H. Takemura. Corneal Imaging Revisited: An Overview of Corneal Reflection Analysis and Applications. *IPSJ Transactions on Computer Vision and Applications*, 5:1–18, 2013.
- [18] M. Ono, B. Shizuki, and J. Tanaka. Touch & Activate: Adding Interactivity to Existing Objects Using Active Acoustic Sensing. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, p. 31–40, New York, NY, USA, 2013. Association for Computing Machinery.
- [19] C. Park and T. Ogawa. A Study on Grasp Recognition Independent of Users' Situations Using Built-in Sensors of Smartphones. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, p. 69–70, New York, NY, USA, 2015. Association for Computing Machinery.

- [20] D. Schneider and J. Grubert. Towards Around-Device Interaction Using Corneal Imaging. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, p. 287–293, New York, NY, USA, 2017. Association for Computing Machinery.
- [21] R. Wimmer and S. Boring. HandSense: Discriminating Different Ways of Grasping and Holding a Tangible User Interface. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, p. 359–362, New York, NY, USA, 2009. Association for Computing Machinery.
- [22] J. O. Wobbrock, B. A. Myers, and H. H. Aung. The performance of hand postures in front- and back-of-device interaction for mobile computing. *Int. J. Hum.-Comput. Stud.*, 66(12):857–875, Dec. 2008.
- [23] D. H. Yoo, J. H. Kim, B. R. Lee, and M. J. Chung. Non-contact Eye Gaze Tracking System by Mapping of Corneal Reflections. pp. 101–106. IEEE, 2002.
- [24] H. Yoo, J. Yoon, and H. Ji. Index Finger Zone: Study on Touchable Area Expandability Using Thumb and Index Finger. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, p. 803–810, New York, NY, USA, 2015. Association for Computing Machinery.
- [25] D. Yoon, K. Hinckley, H. Benko, F. Guimbretière, P. Irani, M. Pahud, and M. Gavriiliu. Sensing Tablet Grasp + Micro-Mobility for Active Reading. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, p. 477–487, New York, NY, USA, 2015. Association for Computing Machinery.
- [26] C. Yu, X. Wei, S. Vachher, Y. Qin, C. Liang, Y. Weng, Y. Gu, and Y. Shi. HandSee: Enabling Full Hand Interaction on Smartphone with Front Camera-Based Stereo Vision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 1–13, New York, NY, USA, 2019. Association for Computing Machinery.
- [27] X. Zhang, K. Ikematsu, K. Kato, and Y. Sugiura. ReflecTouch: Detecting Grasp Posture of Smartphone Using Corneal Reflection Images. In *CHI Conference on Human Factors in Computing Systems, CHI '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- [28] 張翔, 池松香, 加藤邦拓, 杉浦裕太. 角膜反射を利用したスマートフォンの把持姿勢推定. エンタテインメントコンピューティングシンポジウム論文集, pp. 58–61, 2021.

Crane：プロダクトの意匠性・可折性・可製造性を考慮した折紙の計算設計プラットフォーム

須藤 海^{*†} 野間 裕太^{*} 谷道 鼓太郎[†] 鳴海 紘也^{*} 舘 知宏^{*}

概要. 折紙は、Human-Computer Interaction (HCI) やデジタルファブリケーションなどの学術分野および建築やプロダクトデザインなどの産業分野で活用されつつある。しかし、専門的な知識を持たない人間が (1) プロダクトに求められる意匠・(2) 計算折紙に基づく可折性・(3) 最終的に製造するための厚み処理を同時に考慮して折紙プロダクトを設計するのは難しい。特に、内部頂点を持ち複数のヒンジが連動する折紙についてこれらの問題を考慮することは、専門家にとっても困難である。そこで本稿では、Rhinoceros/Grasshopper上に実装された折紙の設計プラットフォーム Crane を提案する。ユーザは (a) 2D/3D の折りパターンを設計し、(b) その変形をシミュレートし、逆に (c) ユーザの与える制約から新たなパターンを発見することができる。また (d) 厚みのない理想的な折りパターンから、ユーザの製造方法に応じて適切なヒンジが付けられた厚みのある構造を出力できる。最後に、本システムの有効性を示すべく3つのデザイン事例を示す。

1 はじめに

折紙は機械的な構造を実現する技術として学術と産業の両面で利用されている。例えば、ミウラ折り [19, 20] は人工衛星の太陽光パネル、吉村パターン [38] は飲料缶の耐久性向上、waterbomb pattern (ナマコ) は医療用ステント [16] などに応用されている。最近では、機械的メタマテリアル [5, 17] や自己折り・4D プリント [8, 2] にも折紙が利用されている。また、折紙の意匠を利用したランプシェード・椅子 [13] や、Al Bahr Towers [3] に見られる切紙構造を用いた可動建築、ISSEY MIYAKE の Steam Stretch [36] における複雑なプリーツなど、折紙が応用されるプロダクトも多岐にわたる。そのため、折紙のメカニズムを非専門家にも利用できるようにすることは重要な課題である。

しかし、計算により折紙プロダクトを設計・製造するためには、(1) プロダクトの適切な形状と機能を実現する「意匠性」(大きさ・水平・対称など)、(2) 計算折紙の理論に基づく「可折性」(後述する可展性や平坦可折性など)、(3) 最終的に現実世界で作るための「可製造性」(製造方法に応じた適切なヒンジ形状や自己交差しない材料の厚みなど) という3つの要求を同時に考慮する必要がある。例えば、デザイナーが意匠性だけを考慮してプロダクトの形状を変更すると、可折性と可製造性は簡単に失われる。このように、適切な折紙プロダクトを作る際にはデザイン・計算折紙・ファブリケーションの知識が要求され、非専門家には敷居が高い。また、折紙の専門家にとっても異なる要求を同時に実現する形状を繰り返し探索するには多大な労力が必要となる。

本稿では、デザイナーが意匠性・可折性・可製造性を同時に考慮することのできる設計プラットフォーム Crane を提案する。本環境の核となるのは、デザイナーがインタラクティブに形状を探索する中で自動的に要求を満たせるような制約付きの GUI システムである。これにより、折紙の非専門家および専門家が折紙プロダクトをより簡単に実現することが可能となる。本稿では本システムの機能を示すとともに、有効性を示すデザイン事例(椅子・テーブル・ランプシェード・マスク)を紹介する。なお、本内容の詳細は ACM TOCHI で出版予定 [29] なので、適宜参照してほしい。

2 関連研究

2.1 HCI で用いられる折紙

折紙は、Human-Computer Interaction (HCI) やデジタルファブリケーションの分野においていくつかの目的に利用される。例えば、LaserOrigami はプラスチックの板を選択的に加熱することにより、3次元の折り曲げ形状を実現する [21]。Thermorph は熱溶解積層 (FDM) 方式の 3D プリンタから出力される平面状のプラスチックを印刷後の加熱で意図的に歪めることにより、サポート剤を使わず高速に折紙形状を生成する 4D プリントと呼ばれる技術の 1 種である [2]。Pop-up Print は 3D プリント構造物をあらかじめ折り畳んだ状態で造形し造形後に展開することで、短い印刷時間・少ないサポート材・不必要な際の収納性を達成する [22]。その他、折紙構造はインタラクティブで動的な物体 [25, 26, 23, 24]、簡単に作製できる 3次元の電気回路 [37]、個人の要求に合わせてカスタム可能なハプティクス [4] などに利用されている。

* 東京大学

† Nature Architects 株式会社

しかし、上記の事例に見られる折紙の多くは、内部頂点を持たない比較的単純なパターンか、既に存在する有名なパターンをそのまま利用したものに限られている。意匠性・可折性・可製造性を考慮した上で比較的複雑なパターンをユーザの要求に合わせて変更できるシステムが存在すれば、これらの研究で利用できる折紙の幅は大きく広がる。

2.2 折紙設計ソフトウェア

折紙を設計するためのソフトウェアは過去にいくつか提案されている。まず、折紙をパラメトリックな順問題として設計する手法として、Constructiveなアプローチが存在する。例えば、ORI-REVO [18]は与えられた輪郭を軸を中心に捻ることで、回転対称な折紙構造を実現する。また、与えられた曲線や断面に沿う剛体折り可能な展開構造を実現する手法も存在する [32, 9]。しかし、これらの方法は形状に何らかの対称性を仮定するため、例えば与えられた任意の3D形状を実現することはできない。

その一方で、最適化に基づく手法を用いることで、与えられた多面体に変形する折紙を逆問題として解くことが可能である。例えば Origamizer [31]は任意の多面体を1枚の紙から折ることのできるパターンを出力する。同様に、最適化計算により折紙の敷き詰め（テセレーション）を目標形状にフィッティングすることで曲面を近似する手法がいくつか提案されている [30, 7, 39, 10]。

特に、Freeform Origami [30]を用いると、形状が平坦な紙から折れることを保証する制約などをかけた上で（つまり可展性を保証した状態で）、2Dと3Dの両方の形状をインタラクティブに操作できる。しかし、Freeform Origamiで扱う要求は可折性のみであり、プロダクトデザインにおいて重要となる意匠性と可製造性の要求に答えることは容易ではない。また、必要に応じてソフトウェアの機能を組み合わせて新しい制約を定義することができない。さらに、ソフトウェアが独立しているため、必要に応じて入出力をそのまま他のプラグインなどに接続することはできず、デザイナーが使い慣れたソフトウェアを使い回すこともできない。

本研究では、上記の最適化と制約に基づくシステムをノードベースのビジュアルプログラミング環境である Grasshopper 上で実現した。これにより、ユーザは可折性だけでなく意匠性や可製造性に関する要求をノードの接続により実現することができ、必要に応じて新たな制約を自分で定義したり、Grasshopperで利用できる様々なプラグインに接続したりできる。なお、GrasshopperはRhinoCerosという3D CADの標準機能であり、プロダクトデザイナーや建築家に広く利用されている。

3 Crane

3.1 3つのフェーズ

図1に示すように、Craneのワークフローは3つのフェーズに分類される：

1. 入力フェーズ (図 1a)
最初の折りパターンを入力する。本システムは (1)RhinoCeros で作成した 2D/3D のパターン、(2) 現実の折紙の画像認識、(3)FOLD フォーマット [6]、(4) 有名なテセレーション（ミウラ折りなど）のテンプレート、(5) 後述の Generalized Ron Resch パターンを生成するための 3D メッシュなどを受けつける。
2. シミュレーション・形状探索フェーズ (図 1b)
入力されたパターンに基づき、システムは変形のシミュレーションまたは幾何的な制約に基づく新たな形状探索を行う。
3. ファブリケーションフェーズ (図 1c)
シミュレーションや形状探索が終わった後で、製造のためのデータを出力する。ユーザの使用するツール（レーザカッター、3D プリンタ、CNC 加工機）などに応じて適切なヒンジを付けた厚みのあるモデルを作る。

ユーザはこれら3つのフェーズを行き来して、インタラクティブに試行錯誤を行うことができる。個々のフェーズに関しては4章・5章・6章で解説するが、次の節からは折紙の椅子をデザインするシナリオを考え、問題解決の手順を示す。

3.2 ターゲットユーザ

本システムは、計算折紙の専門家に限らず折紙の知見を応用したいデザイナー・建築家・研究者などを対象とする。その一方で、RhinoCerosとGrasshopperの基本的な操作および本稿に登場するような計算折紙に関するツールの使い方の習得は前提とする。

3.3 ユースケース：椅子のデザイン

図1に示す木製の椅子をデザインする場合を考える。今、ユーザは意匠性に関する複数の要求を持っている。例えば、折れた状態で座面は水平であってほしい。背もたれは快適な角度であってほしい。地面に接する全ての頂点は浮くことなく接地してほしい。ある辺同士は目標の折れ角で互いにぶつかって、それ以上変形が進まないようにしてほしい。椅子は左右対称であってほしい。

一方で、収納のため椅子は展開状態で完全に平坦になってほしい。これが可折性の要求に相当する。また、意匠性と可折性の要求を満たしたとしても、ユーザは厚みがある板で作った際に板同士が自己交差しないようにしなければならない。また、使用するツールで作ることができるヒンジを選ぶ必要がある。これらが可製造性に相当する。

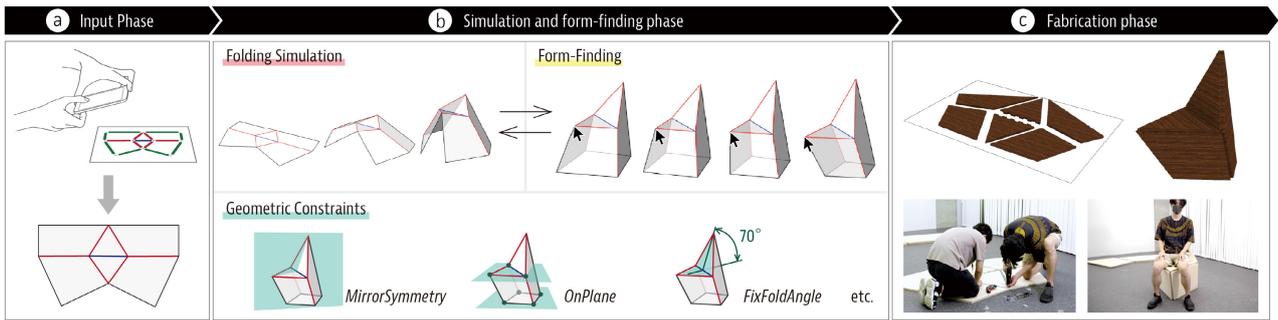


図 1. 設計ワークフロー：(a) 入力フェーズ, (b) シミュレーション・形状探索フェーズ, (c) ファブリケーションフェーズ。

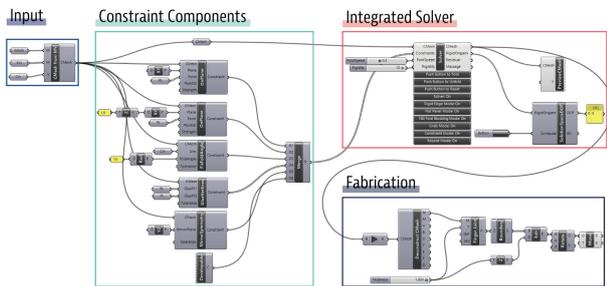


図 2. 椅子の設計に使用した Grasshopper の構造。

これら 3 種類の要求を同時に満たす Grasshopper の構成を図 2 に示す。

3.3.1 椅子 – 入力

ユーザはまず、手作業で実際の紙を折ることで大雑把に椅子のような動きをする折紙をデザインしたと仮定する。このとき、ユーザは図 3 に示すように、最終的な折パターンを赤（山折り）・青（谷折り）・緑（境界）で塗り分け、画像として Grasshopper に取り込む（図 2 の Input）。



図 3. 画像認識による入力：(a) 手作業で形状を試行錯誤し、(b) 折り目と境界に色を付け、(c) カメラで撮影する。

3.3.2 椅子 – シミュレーションと形状探索

次に、ユーザは得られたパターンの変形をシミュレーションにより確認する。このとき、システムは [30] の手法で剛体折りに基づく変形シミュレーションを行う。Crane 上では、図 2 の Integrated Solver に表示されたソルバー上で「Push Button to Fold」をクリックすることで変形を確認できる。

また、ユーザはいくつかの意匠性に対する要求を制約（図 2 の Constraint Components）として与え、形状探索を行う。ここでは、意匠性に対応する 5 つの制約を与えたとする：

1. *OnPlane Constraint*: 座面の頂点を与えられた水平な平面に固定する
 2. *OnPlane Constraint*: 下部の頂点を接地する
 3. *FixFoldAngle Constraint*: 背もたれと座面に存在する辺のなす角度を一定の値に固定する
 4. *GlueVertices Constraint*: 椅子の背面に存在する 2 つの辺がある折れ角で一致する（くっつく）ように固定する
 5. *MirrorSymmetry Constraint*: 鏡映対称にする
- 同様に、可折性に相当する制約も接続する：

6. *Developable Constraint*: 与えられたパターンを可展（1 枚の平面から折れる状態）にする

図 4 に、上記 6 つ（5 種類）の制約を図解する。これらの制約をソルバーに接続し、ソルバー上の「Solver: ON」というボタンをクリックすることでシステムは新たな形状を探索する。なお、このときユーザはある頂点をドラッグしたり、コンポーネントを追加・削除したりすることで、2D または 3D のモデルの形状をインタラクティブに変更することもできる。その際、システムは変形したモデルに対してリアルタイムで再計算を行い、制約を満たしながら変形を提示する。

3.3.3 椅子 – ファブリケーション

シミュレーションと形状探索で適切な折りパターンを発見した後、ユーザは可製造性の要求を満たすためにコンポーネントを追加する。このとき、ユーザは CNC 加工機によって板を切り出すことを考えているため、ソルバーコンポーネントの後段にフィンガージョイントのコンポーネントを接続し、使用する板材の厚みを入力する（図 2 の Fabrication）。システムは自動的に 2D のパターンに厚み付けをして、自己交差が起きないような矩形のヒンジをつける（図

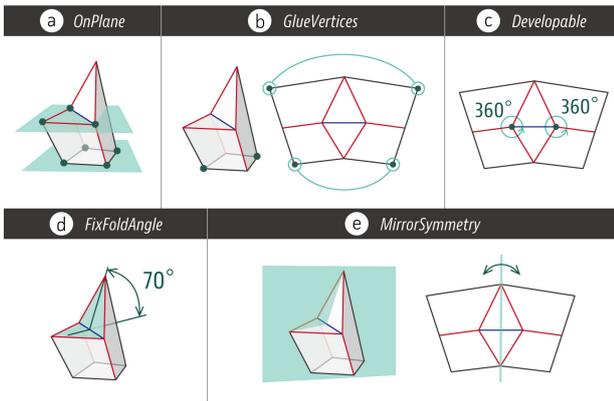


図 4. 椅子の形状探索に使用した制約.

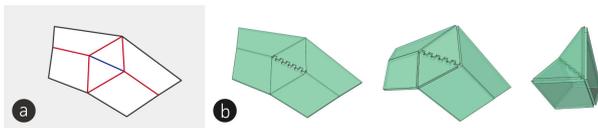


図 5. (a) 厚み付け前と (b) 厚み付け後の椅子.

5). また、フィンガージョイントの場合、あらかじめシミュレーションした際の目標の角度で板同士がちょうど互いにぶつかり安定するようになっている。厚み付けがなされた状態でも、ユーザはシミュレーションを行い形状変形を確認することができる(図 5b)。

以降の章ではそれぞれのフェーズについて説明していくが、紙面の都合により代表例のみ解説する。

4 入力フェーズ

ここでは3種類の入力について述べるが、それぞれの用途は7章にて例示する。

4.1 画像認識

椅子の事例で示したように現実の折紙からパターンを入力する場合、ユーザは撮影した画像を PNG ファイルで Grasshopper に与える。システムは複数色の辺を区別した後、一定の距離 ϵ の範囲に集まる頂点を結合することで折りパターンを生成する。

4.2 テンプレート

また、有名なテセレーションをテンプレートと呼び出すこともできる。現状では、ミウラ折り [19]、吉村パターン [38]、チキンワイヤーパターン [19]、waterbomb pattern などに対応している。

さらに、テセレーションの基本形状に限らず、図6に示すように2つのオフセットした NURBS 曲面の間を充填するようなテセレーションを自動生成することもできる(ここでは Fold Mapping [11] や [39] の手法を利用した)。なお、曲面の間を充填したパ

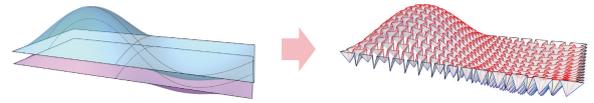


図 6. 2つのオフセット曲面を充填するテセレーション.

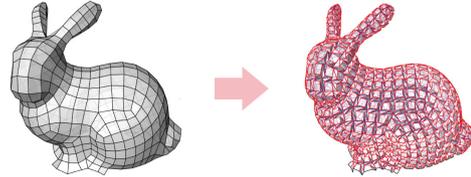


図 7. Generalized Ron Resch パターン.

ターンは多くの場合可展ではなくなっているため、1枚の平面から折る場合は追加で前述の *Developable Constraint* を足す必要がある。

4.3 Generalized Ron Resch パターン

Crane は入力として 3D メッシュも受け付ける。システムは [34] の手法に基づき入力されたポリゴンの隙間に追加の折り目を挿入し、図7に示すような Generalized Ron Resch パターンを生成する。なお、この場合も可展性は保証されないため、多くの場合 *Developable Constraint* をつける必要がある。

5 シミュレーションと形状探索

5.1 ソルバー

ソルバーは、CGNR 法により制約式の最小二乗解を計算する。詳しい計算方法は [30] を参照してほしい。制約のベクトル関数 $\mathbf{C} : \mathbb{R}^n \mapsto \mathbb{R}^m$ (ただし内部頂点数を V としたとき $n = 3|V|$ で、 m は制約の個数) が与えられたとき、ソルバーは方程式 $\mathbf{C}(\mathbf{X}) = \mathbf{0} \in \mathbb{R}^m$ を解く。ここで $\mathbf{C}(\mathbf{X})$ の各行は個々の制約関数に対応する。なお、多くの場合得られる解 $\mathbf{C}(\mathbf{X}) = \mathbf{0}$ は単一ではなくある解空間を持ち、システムは解空間の中から直前の形状に最も近い1点を選ぶ。このため制約が不十分な条件でもシステムはロバストに動作する。

5.2 Constraint Components

ここでは意匠性と可折性に対応する制約のコンポーネントについて説明する。現在 Crane には 14 つ (6 種類) の *Constraint Components* (5 つの *OnGeometry Constraints*, 3 つの *Symmetry Constraints*, 3 つの *GlueGeometry Constraints*, 1 つの *FixFoldAngle Constraint*, 1 つの *FlatFoldable Constraint*, 1 つの *Developable Constraint*) が存在するが、ここでは2種類だけを取り上げる。それぞれのコンポーネントは制約関数 $c_i(\mathbf{X})$ を出力し、

システムはそれらを行とする $\mathbf{C}(\mathbf{X}) = \mathbf{0}$ を解く。

5.2.1 OnGeometry Constraints

OnGeometry Constraints はある頂点の位置を目標形状の上に固定するものであり、以下の5つのコンポーネントが存在する：

1. *OnPoint*: ある頂点を点上に固定
2. *OnPlane*: ある頂点を平面上に固定
3. *OnMesh*: ある頂点をメッシュ上に固定
4. *OnCurve*: ある頂点をNURBS曲線上に固定
5. *OnSurface*: ある頂点をNURBS曲面上に固定

それぞれのコンポーネントは同じ考え方で実装される。制約を受ける頂点を $\mathbf{x} \in \mathbb{R}^3$ としたとき、システムは目標形状の上に存在し \mathbf{x} に最も近い頂点 $\mathbf{x}' \in \mathbb{R}^3$ を計算する。このとき、制約関数 $\mathbf{C}^{\text{onGeo}}$ は

$$\mathbf{C}^{\text{onGeo}}(\mathbf{X}) = \frac{1}{l_{\text{ave}}}(\mathbf{x} - \mathbf{x}')$$

と表される。なお、この関数は初期のパターンの辺長の平均 l_{ave} で割り、スケール非依存にしてある。

5.2.2 Developable Constraint

ある折紙「のような」立体形状が与えられたとしても、それが平面から折れる（可展）かどうかは自明ではない。可展性が満たされる条件は、任意の内部頂点について、頂点周りの角度の和が 2π になることである。各内部頂点とその周りの角度を v および $\theta_{0,v}, \dots, \theta_{n_v-1,v}$ とすると、制約関数 c_v^{dev} は

$$c_v^{\text{dev}}(\mathbf{X}) = 2\pi - \sum_{i=0}^{n_v-1} \theta_{i,v}$$

と表される。

6 ファブリケーション

図8に示すとおり、Craneは(a)オフセットヒンジ [33], (b)フィンガージョイント [14], (c)テーパードヒンジ [33], (d)3Dプリンタブルなシングルヒンジ, および(e)3Dプリンタブルなダブルヒンジの5種類のヒンジ形状を出力できる。

あるヒンジが常に最適ということではなく、それぞれに利点と欠点がある。まず、(a)(b)はレーザーカッターおよび2軸のCNC加工機で作製可能であるが、(c)は3軸加工機か3Dプリンタでしか作製できない。パネル造形後、(a)(b)(c)はパネルを布に貼り付けたりピアノ丁番で固定したりする後処理が必要になる。一方、(d)(e)は作製後の後処理が不要であるが、3Dプリントでしか作れないため他の手法に比べて小さな構造物しか作製できない。その他、(e)を除く全てのヒンジは折れ角に制限がある。(a)(d)は 180° 折りができず、(b)は 90° より大きな谷折り

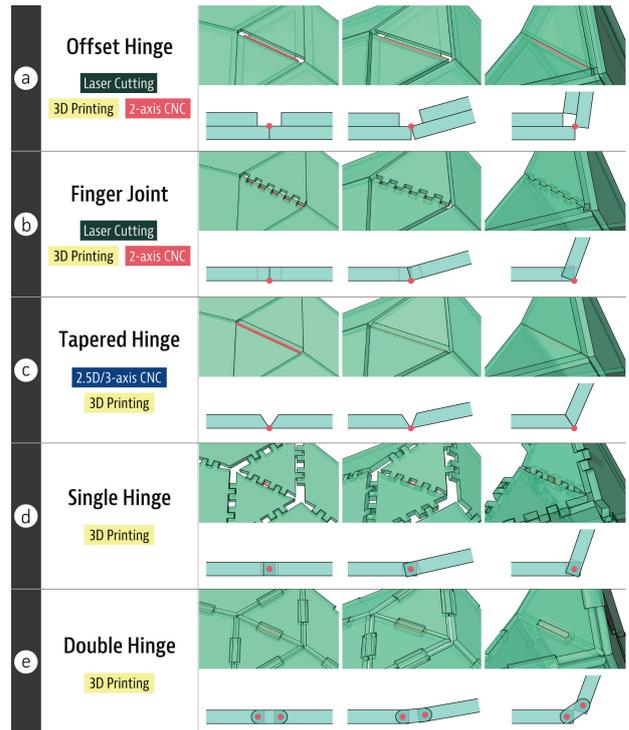


図8. Craneで利用できるヒンジ：(a) オフセットヒンジ, (b) フィンガージョイント, (c) テーパードヒンジ, (d)3Dプリンタブルなシングルヒンジ, (e)3Dプリンタブルなダブルヒンジ。

が実現できず、(c)は 180° の谷折りができない。(e)は山折りでも谷折りでも 180° 折りが可能だが、厳密にはヒンジが1本ではなく、元のパターンと同じように扱えるとは限らない [15]。

7 デザイン事例

Craneの使用法と有効性を示すため、テーブル・ランプシェード・マスクのデザイン事例を示す。

7.1 テーブル

図9にCraneで作製した折紙テーブルを示す。このテーブルは不必要なときには平面に折り畳める（平坦可折である）ため、収納に便利である（平面「から」折れるではないことに注意）。テーブルを設計するにあたり、我々はまず実際の紙で試作を行い、画像認識を行った。その後、以下の制約で実現される意匠性と可折性の要求を達成した：

1. *OnPlane Constraint*: テーブルトップの頂点を与えられた水平面に固定する
2. *OnPlane Constraint*: 下部の頂点を接地する
3. *FlatFoldable Constraint*: 形状を平坦可折にする（紙面の都合上、平坦可折条件は解説しない）
4. *MirrorSymmetry Constraint*: 鏡映対称にする

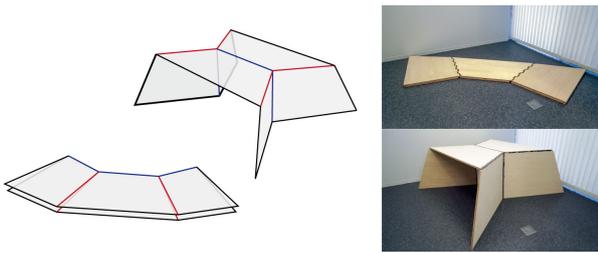


図 9. 平坦可折なテーブル.

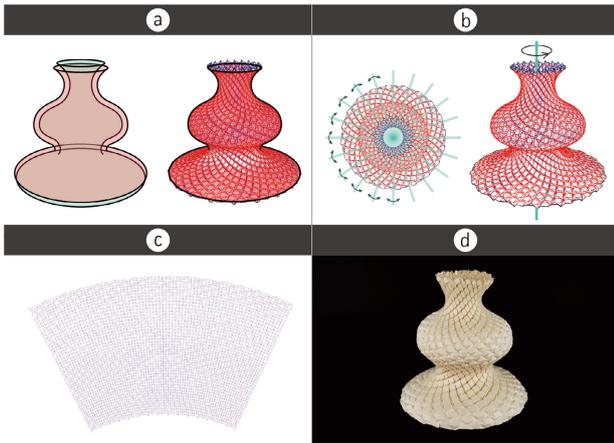


図 10. ランプシェード：(a)2つのオフセット曲面に充填した waterbomb pattern, (b)Rotational Symmetry Constraint による回転対称なモデル, (c)折紙の展開図, (d)実物の写真.

ヒンジにはフィンガージョイントを選択し、木材は後述の EMARF [35] により切削し、それぞれのパネルを固定するためにピアノ丁番を使用した。

7.2 ランプシェード

図 10 に Crane で作製したランプシェードを示す。入力として、2つのオフセット曲面の間に waterbomb pattern を充填した [39] (第 4.2 節参照)。制約として、まずランプシェードを1枚の紙から折るために *Developable Constraint* をかけた。さらに、形状を回転対称にするために *Rotational Symmetry Constraint* を課した(紙面の都合で説明は省く)。その後、Grasshopper 上で曲面の UV 座標を捻ることによって最終的な形状を得た。製造の際には、まずカッティングプロッタで折り目の癖を付けた後、約 10 時間手折りすることで実物を得た。

7.3 Mask

図 11 にマスクを示す。入力として、まず我々は顔の 3D メッシュを用意し、次にその顔を覆うメッシュを作成した。そのメッシュを入力として Generalized Ron Resch パターン (第 4.3 節参照) を生成し、意匠に利用した。また、ランプシェードと同

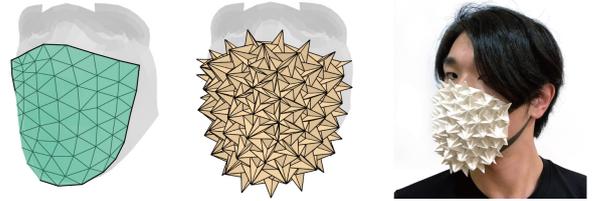


図 11. マスク：(左) 入力したメッシュ, (中) 自動生成されたパターン, (右) 実物の写真.

様に *Developable Constraint* を利用して1枚の紙から作製できるようにして、カッティングプロッタおよび手折りで作製した。

8 他のプラグインとの接続：EMARF

Crane は Grasshopper 上に実装されているため、プロダクトデザインで有用なプラグインを利用できる。ここでは紙面の都合により1つだけ紹介する。

EMARF [35] は Grasshopper で入力したデータに基づいて木材パネルを切削するオンライン製造サービスである。EMARF は Crane から入力した 3D モデルに対して CNC 加工機のツールパスを計算し、そのままデータ送信を行う。なお、本稿で示された椅子とテーブルは実際に EMARF を接続して発注・製造したものである。

9 限界：非剛体折りパターン

本研究で利用したシミュレーションは、折りのプロセスにおいて辺や面が変形しない剛体折りメカニズムに基づく。そのため、例えば折り鶴のような変形の際に歪みが生じるパターンは扱えない。剛体折りできないパターンをシミュレーションしたい場合は、Origami Simulator [1, 12] など別のシステムを利用したほうが良い。なお、非剛体折りで形状探索ができるツールは現在のところ存在しない。

10 結論

本稿では、折紙を利用したプロダクトの設計プラットフォームである Crane を提案した。このシステムは入力・シミュレーション・形状探索・厚み付けに対応しており、ユーザの要求する意匠性・可折性・可製造性を制約ベースのシステムでインタラクティブに実現する。Grasshopper 上に実装されているため、必要に応じて他のプラグインを利用できる。

最後に、本システムは GitHub と Food4Rhino で公開されている [28, 27]。我々は、HCI・ファブリケーション・建築・デザイン・ユビキタスコンピューティングなどに関わる様々な分野の人々が、Crane を通じて新たな驚くべき方法で折紙を活用することを期待している。

謝辞

本研究は情報処理推進機構 (IPA) の未踏ソフトウェア創造事業および JST ACT-I (JP18070086) の支援を受けて実施された。また、椅子の設計例は下田悠太氏によるものである。

参考文献

- [1] A. Ghassaei et al., *Origami Simulator*, 2017, <https://origamisimulator.org/>.
- [2] B. An, Y. Tao, J. Gu, T. Cheng, X. A. Chen, X. Zhang, W. Zhao, Y. Do, S. Takahashi, H.-Y. Wu, T. Zhang, and L. Yao. Thermorph: Democratizing 4D Printing of Self-Folding Materials and Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18*, p. 1–12, New York, NY, USA, 2018. Association for Computing Machinery.
- [3] ARUP, Al Bahr Towers, 2012, <https://www.arup.com/projects/al-bahr-towers>.
- [4] Z. Chang, T. D. Ta, K. Narumi, H. Kim, F. Okuya, D. Li, K. Kato, J. Qi, Y. Miyamoto, K. Saito, and Y. Kawahara. Kirigami Haptic Swatches: Design Methods for Cut-and-Fold Haptic Feedback Mechanisms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20*, p. 1–12, New York, NY, USA, 2020. Association for Computing Machinery.
- [5] K. C. Cheung, T. Tachi, S. Calisch, and K. Miura. Origami interleaved tube cellular materials. *Smart Materials and Structures*, 23(9):094012, aug 2014.
- [6] E. Demaine, FOLD file format, 2017, <https://github.com/edemaine/fold>.
- [7] L. H. Dudte, E. Vouga, T. Tachi, and L. Mahadevan. Programming curvature using origami tessellations. *Nature materials*, 15(5):583–588, 2016.
- [8] S. Felton, M. Tolley, E. Demaine, D. Rus, and R. Wood. A method for building self-folding machines. *Science*, 345(6197):644–646, 2014.
- [9] E. T. Filipov, G. H. Paulino, and T. Tachi. Origami tubes with reconfigurable polygonal cross-sections. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 472(2185):20150607, 2016.
- [10] M. Gardiner, R. Aigner, H. Ogawa, and R. Hanlon. Fold mapping: parametric design of origami surfaces with periodic tessellations. *Origami*, 7:105–118, 2018.
- [11] M. Gardiner, R. Aigner, H. Ogawa, and R. Hanlon. Fold Mapping: Parametric Design of Origami Surfaces with Periodic Tessellations. 09 2018.
- [12] A. Ghassaei, E. D. Demaine, and N. Gershenfeld. Fast, interactive origami simulation using GPU computation. *Origami*, 7:1151–1166, 2018.
- [13] ITOKI, Flip Flap Chair, 2017, <https://www.itoki-global.com/products/flip-flap.html>.
- [14] J. S. Ku. Folding thick materials using axially varying volume trimming. In *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2017.
- [15] J. S. Ku and E. D. Demaine. Folding flat crease patterns with thick materials. *Journal of Mechanisms and Robotics*, 8(3), 2016.
- [16] K. Kuribayashi, K. Tsuchiya, Z. You, D. Tomus, M. Umemoto, T. Ito, and M. Sasaki. Self-deployable origami stent grafts as a biomedical application of Ni-rich TiNi shape memory alloy foil. *Materials Science and Engineering: A*, 419(1):131–137, 2006.
- [17] B. Liu, J. L. Silverberg, A. A. Evans, C. D. Santangelo, R. J. Lang, T. C. Hull, and I. Cohen. Topological kinematics of origami metamaterials. *Nature Physics*, 14(8):811–815, 2018.
- [18] J. Mitani. A Design Method for 3D Origami Based on Rotational Sweep. *Computer-Aided Design and Applications*, 6(1):69–79, 2009.
- [19] K. Miura. Proposition of pseudo-cylindrical concave polyhedral shells. In *IASS Symposium on Folded Plates and Prismatic Structures*, 1970.
- [20] K. Miura and R. J. Lang. The science of Miuraori: A review. *Origami*, 4:87–99, 2009.
- [21] S. Mueller, B. Kruck, and P. Baudisch. LaserOrigami: Laser-Cutting 3D Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13*, p. 2585–2592, New York, NY, USA, 2013. Association for Computing Machinery.
- [22] Y. Noma, K. Narumi, F. Okuya, and Y. Kawahara. Pop-up Print: Rapidly 3D Printing Mechanically Reversible Objects in the Folded State. *UIST '20*, p. 58–70, New York, NY, USA, 2020. Association for Computing Machinery.
- [23] S. Olberding, S. Soto Ortega, K. Hildebrandt, and J. Steimle. Foldio: Digital Fabrication of Interactive and Shape-Changing Objects With Foldable Printed Electronics. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, UIST '15*, p. 223–232, New York, NY, USA, 2015. Association for Computing Machinery.
- [24] J. Ou, M. Skouras, N. Vlavianos, F. Heibeck, C.-Y. Cheng, J. Peters, and H. Ishii. AeroMorph - Heat-Sealing Inflatable Shape-Change Materials for Interaction Design. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16*, p. 121–132,

- New York, NY, USA, 2016. Association for Computing Machinery.
- [25] J. Qi and L. Buechley. Animating Paper Using Shape Memory Alloys. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, p. 749–752, New York, NY, USA, 2012. Association for Computing Machinery.
- [26] G. Saul, C. Xu, and M. D. Gross. Interactive Paper Devices: End-User Design & Fabrication. In Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction, TEI '10, p. 205–212, New York, NY, USA, 2010. Association for Computing Machinery.
- [27] K. Suto, CRANE, 2019, <https://www.food4rhino.com/en/app/crane>.
- [28] K. Suto, KaiSut0/Crane, 2021, <https://github.com/KaiSut0/Crane>.
- [29] K. Suto, Y. Noma, K. Tanimichi, K. Narumi, and T. Tachi. Crane: An Integrated Computation Design Platform for Functional, Foldable, and Fabricable Origami Products. ACM Transactions on Computer-Human Interaction (TOCHI), to appear in 2022.
- [30] T. Tachi. Freeform variations of origami. J. Geom. Graph, 14(2):203–215, 2010.
- [31] T. Tachi. Origamizing Polyhedral Surfaces. IEEE Transactions on Visualization and Computer Graphics, 16(2):298–311, 2010.
- [32] T. Tachi. One-DOF rigid foldable structures from space curves. In Proceedings of the IABSE-IASS Symposium, 2011.
- [33] T. Tachi. Rigid-foldable thick origami. pp. 253–264, 2011.
- [34] T. Tachi. Designing Freeform Origami Tessellations by Generalizing Resch’s Patterns. Journal of Mechanical Design, 135(11), 10 2013. 111006.
- [35] VUILD, EMARF, 2019, <https://www.food4rhino.com/app/emarf>.
- [36] WOW, ISSEY MIYAKE “Steam Stretch” Concept Movie, 2014, <https://vimeo.com/108092253>.
- [37] J. Yamaoka, M. D. Dogan, K. Bulovic, K. Saito, Y. Kawahara, Y. Kakehi, and S. Mueller. FoldTronics: Creating 3D Objects with Integrated Electronics Using Foldable Honeycomb Structures. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, p. 1–14, New York, NY, USA, 2019. Association for Computing Machinery.
- [38] Y. Yoshimura. On the mechanism of buckling of a circular cylindrical shell under axial compression. 1955.
- [39] Y. Zhao, Y. Endo, Y. Kanamori, and J. Mitani. Approximating 3D surfaces using generalized waterbomb tessellations. Journal of Computational Design and Engineering, 5(4):442–448, 01 2018.
- [40] 鳴海紘也, 中原健一, 佐藤宏樹, 新山龍馬, 川原圭博. Liquid Pouch Motors: 紙のインタフェースのための薄く軽く柔軟なアクチュエータ. 第26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 2018), No. 83, p. 1–6, 2018.
- [41] 野間裕太, 鳴海紘也, 奥谷文徳, 川原圭博. Pop-up Print: 双安定な折りたたみ3Dプリント. 第28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 2020), No. 89, p. 43–48, 2020.

未来ビジョン

我々はこれまでの WISS で、折紙のような薄い物体を変形させることのできるアクチュエータ Liquid Pouch Motors [40] や、3D 構造物をあらかじめ畳んだ状態で印刷し、印刷後に展開する手法 Pop-up Print [41] など、アクチュエーション・ファブ리케이션に関する技術を紹介してきた。

本研究は折紙の設計に立ち返り、これまで専門家にしか扱えなかった計算折紙の理論を他分野のユーザにも利用できることを目指したものである。また、理想的な折紙をシミュレーション上で再現するとどまらず、実際にプロダク

トとして作ることのできる環境を整えた。元々記載しなかった内容が長大であったため、様々な機能や評価をカットしなるべくシステム全体の理解が可能となるように努めた。発表では、原稿で触れられなかった部分も紹介したい。

また、我々は計算折紙の理論で作成された複雑なパターンをより簡単に製造する技術の開発も進めている。我々の最終目標は、折紙の分野に見られる機能的で美しい形状を、計算設計と計算製造の知見を活用することで実世界で簡単に再現することである。本研究はそのための重要な第一歩であり、本研究の成果を利用して様々な応用が発生することも期待している。

FaceShow: 顔の部分表示による遠隔会議支援システム

岸本 諒大* 土田 修平† 寺田 努† 塚本 昌彦†

概要. 遠隔会議において参加者がカメラ OFF にすると、参加者同士のリアクションが上手く伝わらず、コミュニケーションの質が低下する。しかし、多くの参加者は「カメラ ON にすることで自分が映ることが嫌である」などの理由からカメラ OFF にすることを希望している。そこで顔全体の開示度を下げることによって、参加者のカメラ ON に対するハードルを下げられ、会議内での相互のリアクションを確認しやすくなるという仮説を基に、顔の部分表示を用いた遠隔会議支援システム *FaceShow* を提案する。本論文では、遠隔会議中に Web カメラ映像を *FaceShow* 内で口や目のみを表示するエフェクト素材を用いて加工し、仮想カメラとして表示する。評価実験により提案システムで目表示のエフェクト素材を用いた際に、カメラ ON に対して抵抗感が減少する可能性があることがわかった。

1 はじめに

新型コロナウイルス (COVID-19) の感染拡大に伴い、従来のような対面でのコミュニケーションだけでなく遠隔会議システムを利用した遠隔でのコミュニケーションの機会が増えている。Zoom 社が公開するレポートによれば、Zoom は世界の約 200 カ国で利用されており、ミーティングの平均時間は 54 分、会議の平均参加人数は 10 人とされている [1]。また経済産業省の通商白書 2021 によると、2020 年 10 月時点の Microsoft Teams の日次ユーザー数は 1.15 億人を超えており、2020 年 9 月時点の Zoom の会議開催回数は年間相当数で 30 兆回を超えるとされている [2]。米国におけるリモートワークの割合は 2025 年で約 40% と予測されており、遠隔会議システムを活用した会話は引き続き行われると想定される。

遠隔会議におけるコミュニケーションの課題の一つとして、Web カメラを介した顔映像の表示機能を ON にして顔を見せることを嫌う人が多く [3]、結果としてコミュニケーションの質が低下する点が挙げられる [4]。江原 [3] が行った調査によると、カメラ ON/OFF のどちらで受講するのが望ましいか学生に複数回調査を行ったところ、すべての調査結果で 9 割以上の学生がカメラ OFF での授業を希望した。これよりオンライン授業において受講生はカメラ ON にすることに対して強い抵抗感を覚えていることがわかった。カメラ OFF を希望する理由として、「自分の顔が他者に見られることに抵抗がある」、「カメラ ON にするとリラックスして授業を受けることができない」などが挙げられている。また Jabra (GN オーディオジャパン (株)) 調べ [5] によると、20 代から 50 代の働く男女 300 人に対し行ったアン

ケートにおいて「毎回自分の顔を写している」と回答した人は約 3 割に留まっており、「自分が映るのが嫌だから」、「音声だけでも問題ないと思うから」などの理由からカメラ ON に抵抗を感じる人が多数を占めていることがわかっている。

一方で、Garau ら [4] の研究によると音声のみに比べてビデオ通話の方がコミュニケーションの質が高いことが報告されている。また、相手の顔が見えている方が話しやすいという意見が多数であることも報告されている [5]。以上より、コミュニケーションの質を高めるには顔を表示した方が良いため、顔表示による効果を維持しつつ、Web カメラを介した映像表示機能の利用に対するハードルを下げる必要がある。

そこで本研究では、遠隔会議参加者の目や口など顔の一部のみを表示する遠隔会議支援システム *FaceShow* を提案する。顔の一部表示により匿名性を確保することで顔出しへの抵抗を低減できるだけでなく、会議参加者はお互いのリアクションを得ることが容易となり、コミュニケーションの円滑化が期待できる。*FaceShow* では、PC の Web カメラから映像を取り込み、部位の抽出、拡大、透過処理のような画像処理を行ったのち、仮想カメラ機能を用いて Zoom 等の遠隔会議アプリケーションにその映像を反映する。顔の一部表示による遠隔会議への影響を調査するために、顔表示 (エフェクトなし)、口表示、目表示、同じアバター表示、異なるアバター表示の 5 パターンを用意し、それぞれのパターンにおいてグループディスカッションを行ってもらい、最後にアンケート調査を実施した。

2 関連研究

顔が人に与える影響についての研究は、これまで多数報告されている。顔に基づいて人の性格、行動意図、認知能力に関する判断を行う [6] 他、意思決定に

* 神戸大学工学部

Faculty of Engineering, Kobe University

† 神戸大学大学院工学研究科

Graduate School of Engineering, Kobe University

影響がある [7] とも報告されている。また、体力、年齢、感情、親しみを特徴づける顔情報について、いずれかの資質が一つだけでも顔に含まれると、それらについて知覚される傾向がある [8]。顔が伝える情報は多く、結果として遠隔会議におけるコミュニケーションにおいて顔を表示しない場合、コミュニケーションの質が低いことがわかっている [4]。我々の提案システムは、顔の全体表示によるコミュニケーションの効果は顔の一部表示によって維持しつつ、遠隔会議システムのカメラ機能を ON にする心理的なハードルを下げることを目指している。

目や口に対する感情認知の違いについてもいくつか報告されている。Yuki ら [9] によると、感情を表現する際、目は口よりもコントロールが難しいことから、日本などの感情を抑制する文化圏では、他人の感情を解釈する際、口よりも目に強く注目する。一方、あからさまな感情表現が一般的な米国などの文化圏では、顔の中で最も表情豊かな部分である口の位置から感情を解釈する傾向があり、文化圏によって他人の感情解釈のために注目する箇所が異なる。また、Low ら [10] によると、東洋人は目よりも口から感情を識別することが困難であり、これは口の情報を推論する経験が浅いためと考察している。さらに東洋人は、口のみ刺激と顔全体の刺激における感情認知において西洋人より劣っていた。一方目の感情認知においては差は見られなかった。本研究では、このような目や口に対する感情認知の違いに着目し、顔の一部表示が遠隔でのコミュニケーションにおいてどのような影響を与えるか調査する。

視線や注視といった目の情報は、コミュニケーションにおいて重要な役割を担っている。Mazur ら [11] は、人が相互に注視すると生理的覚醒が起こることを確認しており、それらがアイコンタクトの有効性の一端を担っていると主張している。Monk ら [12] によると、目線情報を共有することで少ない対話数でタスクを達成できることが報告されている。目線への意識が、自分と相手の発言の理解度を確保するための非言語チャンネルであることを示唆している。本研究では、目のみ表示することで、これら目によるコミュニケーションへのポジティブな効果を維持しつつ、同時に匿名性を下げることでカメラ機能 ON の心理的なハードルを下げることを狙っている。

3 提案システム

3.1 設計

遠隔会議での顔の一部表示による影響を調査するために口や目といった顔の一部のみをそれぞれリアルタイムで表示するシステムを構築する。そのためには、顔から口や目に該当する部分のみを抽出し、鼻や耳などその他の顔の部分を透過させることが必要となる。一般的に広く普及している遠隔会議シ

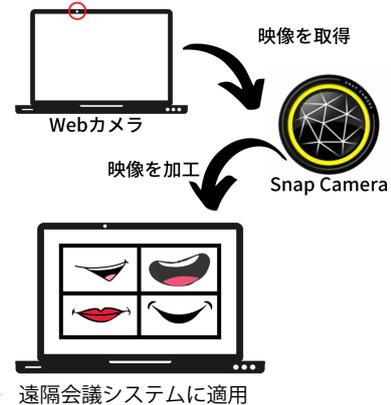


図 1: システムの処理フロー

テムにおいて適用可能とするため、OS がカメラとして認識できる仮想カメラを活用したシステムを構築する。

3.2 実装

図 1 にシステム全体の処理フローを示す。Webカメラ映像を容易に加工できるソフトウェアである Snap 社の Snap Camera [17] を利用する。Snap Camera は Zoom や Microsoft Teams 等の主要な遠隔会議システムに適用可能である。そのため遠隔会議時の Web カメラ映像を Snap Camera 上で任意のエフェクト素材を用いて加工し、遠隔会議システムの会議画面上に表示できる。

目や口など顔の一部を表示するために、エフェクト作成アプリケーションである Snap 社の Lens Studio [18] を用いて、Snap Camera 内で利用可能なエフェクトを作成した。Lens Studio を用いて、目と口それぞれのみを抽出、拡大し、顔の他の部分を透過させるエフェクトを作成した。そのため、Snap Camera を活用できる PC であれば作成したエフェクトを容易に利用可能である。また、顔の動きと連動するため、顔きといったリアクションも反映可能である。

4 評価実験

提案システムが遠隔会議における会話に与える影響を調査するために評価実験を行った。参加者は互いに顔見知りの間柄である 20 代の男子大学生 16 名で行った。各参加者は遠隔会議システム内でのみ声が聞こえる環境で実験を実施した。遠隔会議システムは Zoom を使用した。

4.1 実験条件

顔映像に作用させるエフェクトの種類はエフェクトなし（通常通りの顔表示）、口のみ表示させる「口表示」、目のみを表示させる「目表示」、全員共通



図 2: 口表示



図 4: 同じアバター表示



図 3: 目表示



図 5: 異なるアバター表示

のアバターを使用する「同じアバター表示」、それぞれ異なるアバターを使用する「異なるアバター表示」の5種類である。尚、アバターは ver5.10.0 以降の Zoom の機能であるアバター機能を主としてグループ内で異なるアバターを選択させた。

参加者 16 名は、1 組あたり 4 名の 4 組に分ける。ブレインストーミングの課題とエフェクトの種類のカウンターバランスを取るために、4 組に対して、題とエフェクトの組み合わせや実施順序が可能な限り被らないようにした。実験で用いた口表示、目表示、同じアバター表示、異なるアバター表示の4種類のエフェクトの動作例をそれぞれ図 2,3,4,5 に示す。

4.2 タスク

提案システムを適用した状態でブレインストーミングを行い、その後アンケートに回答してもらった。ブレインストーミングとは一種のアイデアを生み出す「集団発想法」であり、ある題に対して複数人で意見を交換し、新たなアイデアを生み出すものである。本実験では、Guilford's Alternative Uses Task と AC test of creative ability[13] に基づいて、ヘルメット、フォーク、レンガ、ガムテープ、マスクを課題とし、それぞれについて独創的な新しい用途を考える。上記の5種類の課題に対して1種類4分間のブレインストーミングを5回行ってもらった

本実験では、櫻井ら [15] の実験手法を参考にタスクを構成し、アンケート評価をメインとした解析を行う。提案システムが遠隔でのコミュニケーションに及ぼす影響として、恥ずかしさの低減、積極性、理解度の向上が可能性として挙げられる。顔の開示部分を減らすことで発言の匿名性を上げ恥ずかしさの低減につながるのではと考えた。また、恥ずかしさを低減できれば緊張感の緩和につながり発言の積極性が向上するのではないかと考えた。さらに、遠隔でのコミュニケーションの質を高めるという目的に関連して、理解度が向上するかどうかについても調査する。そのため、実験後行うアンケート評価では、顔表示に対する各エフェクトを用いた加工による表示の恥ずかしさやコミュニケーションに対する積極性の変化、内容の理解度を被験者アンケートにて回答してもらった。また、この3項目とは別に、顔を部分表示することに対してのメリットとデメリットについても自由回答してもらった。さらに会議に対する積極性を推し量る定量的な基準として発言回数をカウントした。

4.3 実験手順

参加者には事前に PC に Snap Camera をインストールしてもらい、各自別室に入室してもらった。入室後、参加者は PC 上で遠隔会議システムの Zoom

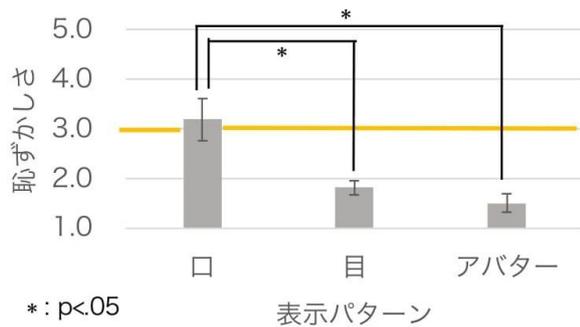


図 6: 恥ずかしさに関する比較

に参加してもらった。その際参加者間で互いの姿が見えないか、直接会話が聞こえないかを確認してから、ブレインストーミングを開始させた。実験時間は、概要説明の時間や参加者の疲労も考慮し 30 分とした。また概要説明時に実験者からブレインストーミングにおける注意点を伝え、適切なコミュニケーションになるよう配慮した。ブレインストーミングの課題がすべて終了した後に、アンケートに回答してもらった。

5 結果

5.1 アンケート結果

アバター表示では全員が同じ種類のもので全員が異なる種類のもので実験を行っているが、被験者アンケートではこの二つを合わせてアバター表示として調査を行っている。実験設計時には同じアバター表示と異なるアバター表示による結果の差異についても調査を進める方針であったが、実験時に大きな特徴が得られなかったことや提案手法の有効性を調査するという主となる目的に集中するため被験者アンケートではアバター表示に統一した。

各項目ではエフェクト無しを基準の 3 とし 1-5 の五段階でシステムを評価してもらった。恥ずかしさに関するアンケート結果を図 6 に示す。縦軸は恥ずかしさ (1: 恥ずかしくなかった, 5: 恥ずかしかった) の平均スコアを示し、エラーバーは標準誤差を示す。平均スコアはそれぞれ口表示が 3.2 (SE 0.4), 目表示が 1.8 (SE 0.1), アバター表示が 1.5 (SE 0.2) であった。表示パターンについて一要因分散分析を行った結果、有意差が見られた ($F(2, 15) = 9.62, p < .01$)。また LSD 法による多重比較を行なったところ、口表示と目表示、口表示とアバター表示においてそれぞれ有意差が見られた ($p < .05$)。口表示は目表示及びアバター表示に比べて、有意に恥ずかしく感じることがわかった。逆に目表示とアバター表示は顔表示と比べて平均スコアが低く、あまり恥ずかしく感じていない傾向が見られた。

匿名性に関するアンケート結果を図 7 に示す。縦軸は匿名性 (1: 誰か理解できなかった, 5: 誰か理解

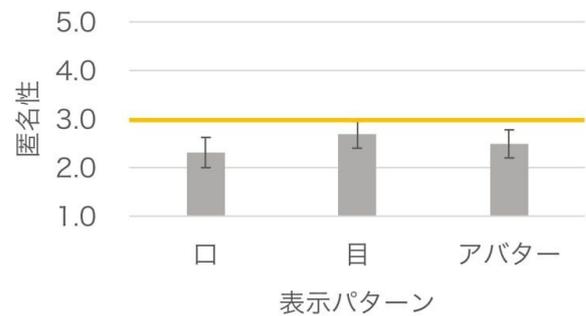


図 7: 匿名性に関する比較

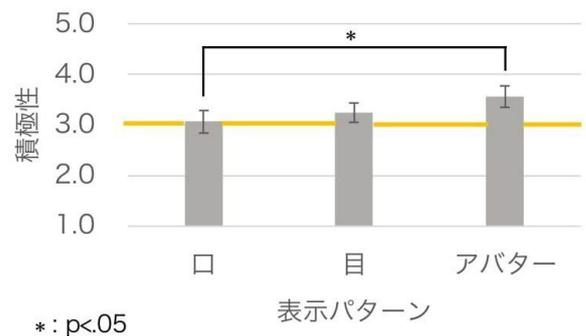


図 8: 積極性に関する比較

できた) の平均スコアを示し、エラーバーは標準誤差を示す。平均スコアはそれぞれ口表示が 2.3 (SE 0.3), 目表示が 2.7 (SE 0.3), アバター表示が 2.5 (SE 0.3) であった。表示パターンについて一要因分散分析を行った結果、有意差は見られなかった ($p > .05$)。また、エフェクトなしの顔表示 (基準の 3) と比べて全体的に平均スコアは低く、顔表示よりも誰か理解しづらい傾向が見られた。

積極性に関するアンケート結果を図 8 に示す。縦軸は積極性 (1: 積極性が向上しなかった, 5: 積極性が向上した) の平均スコアを示し、エラーバーは標準誤差を示す。平均スコアはそれぞれ口表示が 3.1 (SE 0.2), 目表示が 3.3 (SE 0.2), アバター表示が 3.6 (SE 0.2) であった。表示パターンについて一要因分散分析を行った結果、有意傾向が見られた ($F(2, 15) = 3.30, p < .10$)。また LSD 法による多重比較を行なったところ、口表示とアバター表示において有意差が見られた ($p < .05$)。アバター表示は口表示に比べて、有意に積極性が向上することがわかった。また、エフェクトなしの顔表示 (基準の 3) よりすべて表示パターンにおいて平均スコアが高く、積極性が向上する傾向が見られた。

発言内容の理解に役立つかという観点に関しても調査を行った。発言内容の理解に関するアンケート結果を図 9 に示す。縦軸は話の理解度 (1: 話を把握できなかった, 5: 話を把握できた) の平均スコアを示し、エラーバーは標準誤差を示す。平均スコア

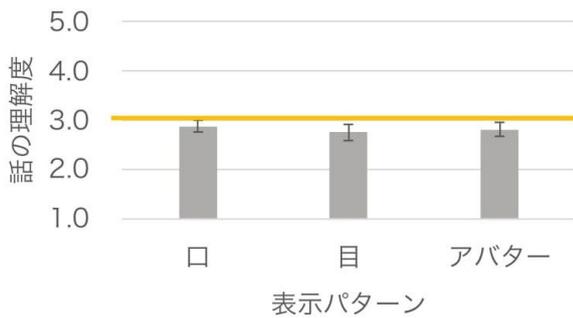


図 9: 内容の理解に関する比較

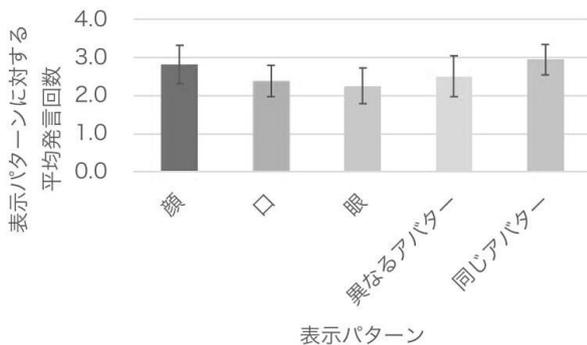


図 10: 表示パターンごとの平均発言回数

アはそれぞれ口表示が 2.9 (SE 0.1), 目表示が 2.8 (SE 0.2), アバター表示が 2.8 (SE 0.1) であった。表示パターンについて一要因分散分析を行った結果, 有意差は見られなかった ($p > .05$)。また, エフェクトなしの顔表示 (基準の 3) と比べて全体的に平均スコアは低く, 顔表示よりも話が把握しづらい傾向が見られたが, 総じて基準に比べての誤差が 0.1~0.2 程度であるため, 影響は小さいと考えられる。

5.2 発言回数

平均発言回数を図 10 に示す。縦軸は平均発言回数を示し, エラーバーは標準誤差を示す。ここでいう発言回数とはブレインストーミングでのアイデアの数のことで, 相槌やその他アイデア以外の発言はその数に含まない。口表示, 目表示の際の発言回数はエフェクトなしの顔表示と比べて低い傾向が見られた。表示パターンについて一要因分散分析を行った結果, 有意差は見られなかった ($p > .05$)。

5.3 参加者による自由回答

提案手法のメリットについて参加者からの意見を募った。「メイクやヘアセットをせずに参加できる」, 「口以外の部分をどれだけ動かしてもバレない」, 「発言している人が目立つ」, 「眠そうにしているもバレなさそう」, 「受けを狙った話をした際, 笑いを取れるかの確認がしやすい。相手の表情に関係なく笑っ

ているフィルタがあれば空気が和むと感じた」, 「相手の口元の細かな動きがわかるようになり, 微妙な微笑みなども分かるようになる」, 等のコメントが寄せられた。

同時にデメリットについても意見を募っており, 代表的なものとして, 「口の拡大表示は恥ずかしい」 (8名), 「誰が話しているか把握しにくい」, 「身振り手振りがわからない」, 「口以外の情報が無くなる」, 「初対面だと誰かわからなくなる」等があった。

また, 実験に対する感想として, 「口表示やアバター表示を使った後は通常時の緊張してる感じがいつもより伝わりやすい」, 「知り合い同士は相手の顔を見て話したいため, この場合アバター表示はコミュニケーションの助けになっているとはいえない」, 「顔が分かりにくいいため, あえて名前表示なしでもやってみよう」, 「課題が難しく感じ話せることがなかったため, 積極性が全体的に下がり積極性の判定にはあまり適していない気がした」, 「色々なフィルタを使ったらもっと議論が盛り上がると思った」, 「考え事すると画面の外見えてしまうため, 対話メインの方がよりエフェクトに対して感想を持てた」などの意見を得た。

6 考察

6.1 口表示・目表示の恥ずかしさ

口表示の会話形式について恥ずかしいと感じる人と感じない人の二極化がみられた。これには二つの原因があると考えられる。まず, 口表示では普段の会話以上に口元が強調されるためである。歯並びや唇などの口元にコンプレックスを抱えている場合, その部分が拡大され, 他者の前で表示されることに対して大きく抵抗を感じるのではないだろうか。次に, 口表示では自分の感情を相手に察知されやすくなり, 恥ずかしく感じてしまう可能性が考えられる。緊張感や悔しさなどの感情は口元に現れやすくなり [16], 口元が拡大されることで相手に自分の隠したい感情は伝わりやすくなってしまふ。

一方で口表示により恥ずかしさが低減されたと回答した被験者も多く見受けられた。理由として笑顔が強調されることが考えられる。口元が大きく拡大される加工を施しているため通常よりも微小な口元の変化に気づきやすい。そのため相手が笑顔になっていることを認知しやすく, 結果として場のリラックスした雰囲気を作り出す手助けになっていると考えられる。また, ヘアセットのように外見に気を遣う必要がなくなることも口表示が顔表示よりも優れている点であり, 恥ずかしさが低減される理由の一つだと考えられる。

口表示の会話形式とは対照的に目表示の会話形式は, 16名の参加者全員が通常の顔表示と同様の恥ずかしさ, もしくは恥ずかしさが軽減されたと回答し

ており、エフェクトなしと比べて恥ずかしさが低減される傾向にあることが分かる。理由として、口元に比べて目元のコンプレックスを持つ人は少ない、目のみでは自身の感情が相手に察知されづらいことなどが考えられる。

以上より、口表示では顔表示よりも恥ずかしさを軽減することはできなかったが、目表示では顔表示に比べ恥ずかしさを軽減する傾向が見られた。よって、提案システムを用いた目表示は遠隔会議におけるカメラ ON に対する抵抗を減少させる可能性がある。

6.2 アバター表示と提案システムの相違点

口表示については匿名性がアバター表示以上に高いことが分かる。初対面で顔を知らない相手との会話では十分にその匿名性の高さを発揮し相手にどんな顔をしているのかを悟られることなくコミュニケーションを行うことができると考えられる。また、アバター表示の種類によっては笑顔が反映されにくい場合がある。一方で口表示は口部分を拡大しその他の部位を透過していることで、アバター表示やエフェクトなし以上に相手の笑顔を認知しやすくなると考えられる。笑顔がコミュニケーションに与えるポジティブな影響は多数の研究により明らかになっている。これにより前述の二つの表示法以上に場の雰囲気が温かくなる可能性がある。

目表示についてはアバター表示とほぼ同等な水準まで恥ずかしさを軽減できることが分かった。また、口表示では匿名性が高いことが分かったが、逆に目表示は匿名性が低く誰が話しているか比較的わかりやすい結果となった。これによりある程度本人と話している実感が得られるため見知った仲でのコミュニケーションに適していることが考えられる。実際、アバター表示は参加者が誰かわからず距離感を感じるため、目表示の方が話しやすいという意見が被験者アンケートから得られた。これにより目表示はアバター表示同等に恥ずかしさを軽減しつつアバター表示以上に本人と話している実感を得ることにより親近感を感じ、リラックスした状態でコミュニケーションを行うことができる可能性がある。

図 10 の発言回数によると、提案手法は発言回数が減少している結果になっていることが分かる。しかし 5.2 節で述べたようにここでいう発言回数はアイデアの数をカウントした数値であるため、積極性が向上していた際でも結果に反映されず、適切に評価できていない可能性がある。そのため今後の研究活動では発言回数の定義や議題の内容を見直すことで、発言回数を積極性の定量的観点として明確にする必要がある。

このように、提案手法はシチュエーションに応じてアバター表示以上に効果的に用いることができると考えられる。しかし以上の考察は被験者アンケー

トによるものや著者の主観に基づいている意見が大きく占めているため、今後の研究活動ではこれらの事項を客観的に検証していく必要がある。

7 応用例

本提案システムを有効であると考えられるシチュエーションについていくつか述べる。

たとえばオンライン講義の場面での活用が挙げられる。高校や大学等のオンライン講義では講師側は生徒がカメラ OFF でリアクションを確認できないことにより、講義中に沈黙が発生してしまう。それにより沈黙を恐れて一方的に話し続けてしまうことが問題となっている [14]。この問題を解決するために、生徒のリアクションを講師側が確認する方法が挙げられる。生徒のリアクションを確認するにはカメラ ON にする、チャットにメッセージを送ってもらうなどがあるが、どれも生徒の負担が大きい。これに対して、本提案システムの目表示は顔表示よりも恥ずかしさを軽減できるため、生徒側のカメラ ON に対するハードルを下げるができるだろう。

顔の開示度を下げることで個人の内面により着目できる可能性がある。人に対する印象はその能力に関係なく、顔などの外見に大きく左右されることがわかっている [6]。たとえば就職の面接において能力を定量的に評価されることなく外見に左右されては、問題が起こる可能性がある。そこで口や目のみを表示することで本人と話している感覚を維持しつつ、面接官の注意をより内面へと向けることができるのではないかと予想する。

8 おわりに

本論文では遠隔でのコミュニケーション支援を目的として、顔の部分表示による遠隔会議支援システム *FaceShow* を提案した。Web カメラから映像を取得し、作成した口と目のみを表示するエフェクト素材を用いて仮想カメラ内で映像を加工した。加工した映像は Zoom 等の遠隔会議システムに適用可能である。評価実験では提案システムを用いた遠隔でのコミュニケーション素材に及ぼす影響について調査した。用いたエフェクト素材は口表示、目表示、同じアバター表示、異なるアバター表示であり、エフェクトなしの顔表示との比較を行った。実験結果から目表示のエフェクト素材を用いた場合に、カメラ ON の際の顔表示に比べて抵抗感が減少する傾向がみられた。今後はプレゼン形式、初対面同士といった多岐にわたるシチュエーションで提案システムを用いた実験を行い、シチュエーションに適した顔の表示方法について明らかにしていく。また、タスクやアンケート内容について再考し、多角的な視点から提案システムの有効性の調査及び改良を進めていく。

謝辞

本研究の一部は、JST CREST (JPMJCR18A3) の支援によるものである。ここに記して謝意を表す。

参考文献

- [1] 会議がいちばん多いのは水曜日 Zoom のレポートから見える「働き方」, <https://news.yahoo.co.jp/byline/dobashikatsutoshi/20220216-00281805> (最終アクセス日: 2022年9月26日).
- [2] 第1部 第1章 我が国を巡る経済情勢と今後の通商を巡るトレンド 第2節 通商を巡る国際潮流, <https://www.meti.go.jp/report/tshuhaku2021/2021honbun/i1120000.html> (最終アクセス日: 2022年9月26日).
- [3] 江原康生: オンラインコミュニケーションでのカメラ ON/OFF による発表者に与えるストレスに関する評価実験, *日本バーチャルリアリティ学会論文誌*, Vol.27, No. 1, pp. 14–16 (Mar. 2022).
- [4] M. Garau, M. Slater, S. Bee, and M. A. Sasse: The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*, pp. 309–316 (Mar. 2001).
- [5] WEB 会議で「毎回自分の顔を映す人」はわずか 3 割! ! 会議中自分の顔を映さない人の約半数が「音声だけで問題ないと思う」と回答 「オンライン会議と周辺デバイスに関する調査」, <https://prtimes.jp/main/html/rd/p/00000011.000072842.html> (最終アクセス日: 2022年9月26日).
- [6] A. Todorov, C. Y. Olivola, R. Dotsch, P. Mende-Siedlecki: Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance, *Annual Review of Psychology*, Vol. 66, No. 1, pp. 519–545 (Jan. 2015).
- [7] C. Y. Olivola, F. Funk, and A. Todorov: Social Attributions from Faces Bias Human Choices, *Trends in Cognitive Sciences*, Vol. 18, No. 11, pp. 566–570 (Oct. 2014).
- [8] L. A. Zebrowitz and J. M. Montepare: Social Psychological Face Perception: Why Appearance Matters, *Journal of Soc Personal Psychol Compass*, Vol. 2, Issue 3, pp. 1497–1517 (May 2008).
- [9] M. Yuki, W. W. Maddux, and T. Masuda: Are the Windows to the Soul the Same in the East and West? Cultural Differences in Using the Eyes and Mouth as Cues to Recognize Emotions in Japan and the United States, *Journal of Experimental Social Psychology*, Vol. 43, No. 2, pp. 303–311 (Mar. 2007).
- [10] A. C. Y. Low, V. Y. S. Oh, E. M. W. Tong, D. Scarf, and T. Ruffman: Older Adults Have Difficulty Decoding Emotions from The Eyes, Whereas Easterners Have Difficulty Decoding Emotion from The Mouth, *Scientific Reports*, Vol. 12, No. 7408 (May 2022).
- [11] A. Mazur, E. Rosa, M. Faupel, J. Heller, R. Leen, and B. Thurman: Physiological Aspects of Communication Via Mutual Gaze, *American Journal of Sociology*, Vol. 86, No. 1, pp. 50–74 (Jul. 1980).
- [12] A. F. Monk and C. Gale: A Look Is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation, *Journal of Discourse Processes*, Vol. 33, Issue 3, pp. 257–278 (June 2002).
- [13] A. Meadow and S. J. Parnes: Evaluation of Training in Creative Problem Solving, *Journal of Applied Psychology*, Vol. 43, No. 3, pp. 189–194 (1959).
- [14] 松下幸司: 大学の遠隔講義におけるアクティブラーニング型授業の試み—グループ・コミュニケーション・ルームと情報共有ツールを併用して—, *香川大学教育実践総合研究*, Vol. 41, pp. 89–98 (Sep. 2020).
- [15] 櫻井 翔, 中里直人, 吉田成朗, 鳴海拓志, 谷川智洋, 廣瀬通孝: 表情変形フィードバックによる遠隔協調作業における創造力向上支援 (<特集> 教育・訓練・支援).
- [16] 吉川 茂: 「だてマスク」着用に関する心理学的研究, *阪南論集. 人文・自然科学編*, Vol. 53, No. 2, pp. 97–108 (Mar. 2018).
- [17] Snap Camera, <https://snapcamera.snapchat.com/> (最終アクセス日: 2022年9月26日).
- [18] Lens Studio, <https://ar.snap.com/lens-studio> (最終アクセス日: 2022年9月26日).

研究発表コンテンツ内の意味的対応関係を活用した研究理解の支援

太田 彩* 脇田 建†

概要. 遠隔開催の学会の増加に従い、学会運営組織から提供されるコンテンツの種類は以前に比べ多様になった。しかしコンテンツは個別に視聴・消費され、研究の統合的な理解の手助けにはなっていない。そこで本研究は各種コンテンツ間をハイパーリンクで結びつけた相互参照を提供することで新たな学会参加の形を提案する。具体的には自然言語処理・画像処理を用いて、コンテンツを超えて関連した内容を含む箇所の対応関係を自動で抽出する。加えて対応関係を相互参照の形でユーザーに提示することで、研究理解過程におけるコンテンツ間の移動を容易にする。以上によりコンテンツ間の結びつきを深め、論文の内容理解の効率向上・認知負荷の軽減を目指す。本論文では対応関係を自動で抽出する手法を提案し、相互参照をユーザーが容易に確認できるシステムのプロトタイプを作成した。加えて抽出手法の性能評価と有効性を検証するための実験を行った。

1 はじめに

近年の学会では口頭発表を補完する形で、論文やシステムのデモ動画など多様なデータ種別のコンテンツが提供されている。加えて COVID-19 の蔓延により学会の遠隔視聴の機会が増加したことで、発表動画や紹介動画などの追加の情報も取得できるようになった。その結果、参加者は目的や関心に合わせ、研究理解に使用するコンテンツの形態を選択することが可能になった。加えてコンテンツを組み合わせることで、研究への深い理解を得ることもできる。

しかし一方でコンテンツを横断して研究を理解するための体勢は整えられておらず、コンテンツは個別に消費されている。例えば研究の概要を発表動画で、実験部分の詳細を論文で確認したい時、ユーザーはパソコンの各種ウィンドウを使い分け目的を達成しようとする。しかしウィンドウの切り替えや特定箇所の探索はユーザーに認知負荷を与え、統合的な理解は阻害されてしまう [24]。このような状況からコンテンツは個別に消費され、各コンテンツの特性を活かした使い分けは行われていない。

本研究では現状を踏まえ、多様なコンテンツの統合的な理解の補助を目標とする。加えて、そのことによって論文の内容理解の効率向上、認知負荷の軽減を目指す。

本研究で注目したのは**相互参照機能**である。相互参照の利用は World Wide Web や Wikipedia のハイパーリンクに遡り、今日では PDF や Microsoft Word 文書でも使われ、有効性は幅広く認識されている。本研究では口頭発表動画・発表資料・論文内の要素の対応関係を自動で検出し、コンテンツ間にもたがる相互参照の形でユーザーに提示する。例と

して論文と口頭発表動画と類似した議論が展開されている論文の箇所を示唆すること、選択されたスライドを用いて発表動画の箇所を提示することなどが挙げられる。コンテンツの切り替えを容易にすることで、データ種別を横断した内容の深い理解を支援することが目的である。

その上でコンテンツの要素を取り出し、要素間の対応関係を自動で抽出する手法を提案した。加えて対応関係を相互参照の形でユーザーに提示するシステムのプロトタイプを作成した。また対応関係の抽出手法の速度と性能を評価するための簡易的な評価実験を行った。

以上を踏まえ、論文の貢献は以下の通りである。まず相互参照に注目し、新しい学会視聴の形を提案する。加えて論文、発表動画から要素を抽出し、類似箇所を自動で検出する手法を提案する。さらに相互参照を確認するためのプロトタイプを作成し、自動検出手法の性能の予備調整を行った。

2 関連研究

2.1 アクティブ・リーディング

アクティブ・リーディングとは注釈をつけたり、メモをとるなどの行為をしながら読書を行うことである。学会参加者も論文を読み研究への理解を深めているため、本研究と関連深い分野である。これらの行為は教育機関や知的労働の場において大きな役割を果たしている。そのためアクティブ・リーディングの現状に関する調査 [6, 21, 17] や、それを支援するツールを提案する研究が数多く提案されてきた。

LiquidText は紙の読書の快適性を模倣するのではなく、流動的な文書表現と様々なインタラクションを提供することでアクティブリーディングを促進させた [25]。Headらは難解な論文に対し、専門用語や記号の定義を読者が最も必要とするときに提供

* 東京工業大学

† 東京工業大学

するインタフェースを提案した [12].

これらの研究は主に文書を読む読書の快適性と効率を向上することに重点をおいているが、本研究がめざすような関連する複数のデータ種別を横断的に理解する支援は提供しない。

2.2 オンライン学会への参加意識を高める研究

COVID-19 の蔓延により学会の遠隔視聴の機会が大幅に増加した。そのため近年遠隔方式の問題点を解決する形で、オンライン学会を支援する方法 [23, 13] が提案されてきた。加えて遠隔の学会や会議を後から振り返り理解するための研究 [10] も行われてきた。

Murali らはプレゼンテーションにおいてセンシング技術を利用し聴衆の反応を取得し発表者へ伝達した [18]。Aseniero らは参加者が会議中に自身の反応を発表者に匿名で送ることができ、あとから会議の様子を収集した反応を通し振り返ることのできるシステムを提案した [7].

El-Assady らは会議における話者の話すテーマの流れを表現することで、話者の行動パターンを分析するというアプローチを提案した [9]。阿辺川らは研究の映像と文書要素に対し、関係性などの付加情報を人間の手で注釈づけることで、情報を統合し読者の理解を支援した [5].

本研究は学会参加者による研究発表の深い理解を目的の中心に据え、必要なシステムの要件分析を行った。加えて論文と発表資料から図表や段落などの要素を自動で抽出し、要素間をまたいだ類似箇所を検出することで、ユーザーが短時間で必要な情報へアクセスできる体制を整えた。

3 提案手法

3.1 現在の学会参加の問題点

近年の学会は論文に発表動画、紹介動画など多様なコンテンツが提供されている。しかし複数コンテンツを統合的に理解するための体勢やシステムは提供されていない。そのためユーザーはコンテンツを個別に消費しており、各コンテンツの特性に応じた使い分けは行われていない。そこで本研究ではそれら問題を解決するために、様々なコンテンツの統合的な理解を促す新しい学会参加のコンセプトを提案する。そして論文の内容理解の効率向上、認知負荷の軽減を行うことを目標とする。

3.2 提案手法の概要

本研究では自然言語処理や画像処理技術を用いて、学会の提供コンテンツ内における関連した内容を含む箇所の対応関係を抽出する。加えて対応関係を相互参照の形でユーザーに提示し、研究理解の過程におけるコンテンツ間の移動を容易にする。こうして

学会参加者の意識の中でのコンテンツ内の結びつきを深め、研究を様々な方面から理解することを可能にする。

PacificVis 2022[4] で提供されたデータをもとに、相互参照に用いるデータを決定した。具体的には論文のメタデータ、紹介動画、発表動画、論文を対象に研究を進める。ここでメタデータは研究の表題や著者を BibTeX などのフォーマットで保存したものの、紹介動画は学会の開催前に提供される研究の内容を 30 秒程度で簡潔に紹介した動画のことである。発表動画とは発表者が自身の研究をスライドウェアや開発したシステムを用いて紹介した研究紹介の録画動画である。本研究では特にスライドやシステムが動画の全面に映し出され、発表者の音声によって説明されたものを特に発表動画と定義づけた。これらコンテンツの中でも発表動画と論文が研究の内容を知る上で大きな役割を担っているため、この二つのコンテンツ間の結びつきを深めることを重視する。

3.3 要件分析

システムを開発するにあたり、研究発表の参加者の行動を想定し要件の分析を行った。加えて関連分野であるアクティブリーディングの研究を参考にした。O'Hara らは実験を通し人々のアクティブリーディングの様子を明らかにし [20]、Tashman らはアクティブリーディングの主要プロセスを注釈、内容抽出、ナビゲーション、レイアウトという 4 点とした [25]。本研究ではアクティブリーディングの主要プロセスを参考に以下の 5 つの要件を設定した。

- (R1) 論文の概要を表示する機能
 - (R2) コンテンツ本体を視聴する機能
 - (R3) 関連箇所を示唆し、移動する機能
 - (R4) 関心のある箇所を選択し、自身の疑問や意見・感想を記述する機能
 - (R5) (R4) で記述した内容を他者に共有する機能
- (R2) はナビゲーションに、(R3) はレイアウトに、(R4) は注釈と内容抽出に対応している。Tashman らの提案した Liquid Text[25] が注釈付システムであることに対し、本研究は論文の概要確認システムであるため、本システムではナビゲーションやレイアウトのプロセス中でもコンテンツ自体の配置を変更する機能は提供しない。加えて関心のある研究の発見と、他者との意見や情報の交換は研究理解の過程において重要な要素と認識されている [8, 23] ため、それらを支援するための機能 (R1)(R5) を追加した。

3.4 ユースケース

この節ではシステムの全体像を明確にするために、要件をもとにしたユースケースについて述べる。

花子さんはリアルタイムで見られなかった論文発表について、あとから発表動画を視聴することにした。まず研究の概要を知ろうと思い、発表動画から自動生成した発表資料を眺めた。このうち特に興味のある提案手法を説明するスライドに目をとめ、その発表資料に対応する動画の箇所付近を視聴した。動画を視聴する中、ユーザー実験の説明を見て、細かい条件設定が気になった。そこで動画のその部分から、内容に関連のある論文の段落を参照した。そうすると論文には実験の詳細が記述されており、疑問を解消することができた。

発表動画を視聴後、他の人にこの研究を詳しく紹介したいと思ったため、論文も最初から確認するようにした。そうすると文中に表を参照した箇所を発見したため、クリックするとシステム上で別のページにあった表を確認することができた。また読み進める中で似た話を論文の以前のページで既に読んだ気がしたため、類似段落をまとめたクラスターを確認した。その結果現在読んでいる段落と似た内容が他の節でも記述されていることを確認した。花子さんはシステムを使うことで効率的に研究内容を理解し、また論文の構造も深く理解することができた。

4 システム

本研究では要件分析とユースケースをもとにプロトタイプを作成した。その際に IEEE の 2 カラム国際会議用の様式に合わせて、要素と対応関係の抽出手法を考案した。本節では最初にシステムの概要について紹介を行い、その後開発に至るまでの工程について紹介を行う。

4.1 システム概要

システムは大きく論文選択画面と論文視聴画面の二つの画面に分けられる。論文選択画面では論文の概要と内容を表した WordCloud・紹介動画を確認しながら興味のある論文を選択することができる。論文視聴画面は図 1 のような構造になっており、6 つの部分に分けることができる。

発表動画・発表資料ビュー このビューでは発表動画と、発表動画から自動生成された発表資料、また音声の文字起こしを閲覧できる。発表資料の一覧から発表の概観を眺め、円滑に自分の興味のある部分を選択することができる。加えて発表資料の 1 ページを選択すると、自動でそのページと内容が最も類似した論文の段落がハイライトされる。この画面は要件分析の (R2, R3) と対応している。

類似度ヒートマップ このビューでは現在選択している発表資料のページの、論文の段落に対する類似度の一覧をヒートマップを通し確認することができる。ヒートマップが濃いほど段落との内容が類似していることを表しており、ヒートマップをタップすると、選択した段落を論文ビューで確認することが

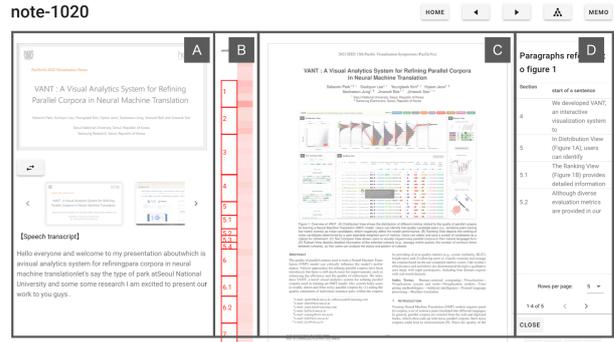


図 1. 現状のプロトタイプ of the paper viewing screen (A) 発表動画・発表資料ビュー (B) 類似度ヒートマップ (C) 図表参照段落発見ビュー (D) 図表参照段落発見ビュー

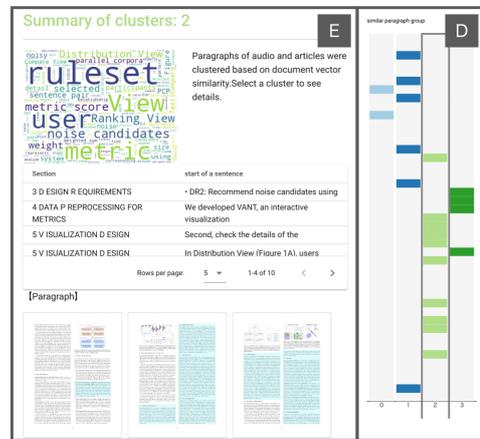


図 2. 論文視聴画面の一部分 (E) 論文グループビュー (F) 論文グループ詳細ビューの概要

できる。このビューは要件分析の (R3) と対応している。

論文ビュー このビューでは論文を閲覧することができ、要件分析では (R2) と対応している。

図表参照段落発見ビュー このビューは論文内の図表をクリックすることで閲覧することができる。ビュー内の表には選択した図表が参照されている論文の段落の一覧が表示されており、行を選択することで論文の段落の詳細を確認することができる。この画面は要件分析の (R3) と対応している。

論文グループビュー このビューは論文の段落を内容の類似度に応じてクラスタリングした結果を表示する。クラスターを表す列をタップで選択すると、論文グループ詳細ビューでクラスターの詳細を確認できる。この画面は (R3) と対応している。

論文グループ詳細ビュー このビューはクラスターを構成する段落の一覧を示すと同時にクラスターの概要を Word Cloud で要約する (図 2)。また画面下部に表示される論文のハイライト表示で、クラスターを構成する段落を確認できる。このビューは要

件 (R2, R3) と対応する。

4.2 分析の流れ

ここでは開発したシステムについて主にデータの処理の流れを中心に紹介する。以下の説明はコンテンツ内の要素の抽出、要素間の関連の検出、ユーザーを補助する機能、システム全体の構成に分けられる。

4.2.1 コンテンツ内の要素の抽出

コンテンツ間の関連した箇所の検出を行うために、この段階では図3のように各コンテンツから発表資料や図表、文字起こしなどの要素を抽出する。発表動画については Amazon Transcribe[1] を使用し、発話内容の音声から文字起こしをする。加えて画像処理によって発表動画を一定秒数ごとにフレームに分割し、類似したフレームをまとめることで発表資料を作成する。参考となる発表動画を閲覧したところ一枚のスライドを1分程度発表に使用していることを確認できたため、本研究では10秒ごとにフレームを分割する。最後に発表資料の各スライドから図表を抽出し、文字認識ライブラリを用いて文章を抽出する。

論文については文字認識システムを使用し、テキスト化する。このままでは2段組の論文に適用できないため、正規表現を用いて各行が2段組か判定し、2段組の場合は行を左右に分割し、左右の文章をそれぞれまとめることで文章の出現順を正しいものへと整える。そして調整した文章に対し正規表現を適用することで、節の表題と図表のキャプション、本文を段落ごとに取り出す。加えて論文の中の図表も画像処理ライブラリを用いて抽出する。これらの操作により相互参照を行うために必要な要素を一通り抽出することができた。

4.2.2 関連箇所の対応付け

この段階では自然言語処理と画像処理により、図3のように要素間の関連箇所を自動検出する。

まず論文の段落と発表動画の文字起こしを比較する。この際に各段落と文字起こしの文章をベクトル化し、コサイン類似度をもとに段落と音声間の類似度を計算した。ベクトル化手法は TF, TFIDF, word2vec[15, 16] の平均, doc2vec[14] の4手法を比較し、最も優秀な TFIDF を採用した。(詳細は5.2節) 加えて論文の段落間の類似箇所を紐づけるために、TFIDFによって生成したベクトルを用いたクラスタリングを行った。論文内の類似した段落のグループ数は論文ごとに異なり、事前に不可知である。このため、クラスタリング手法としてはクラスタ数を事前に与えるものは避け、外れ値にも強い DBSCAN[11, 22] を採用した。また注目した段落と関連した段落の数を削減するため、近接性基準にあたる ϵ を 0.01 から 1.00 の範囲で 0.01 ごとに変化

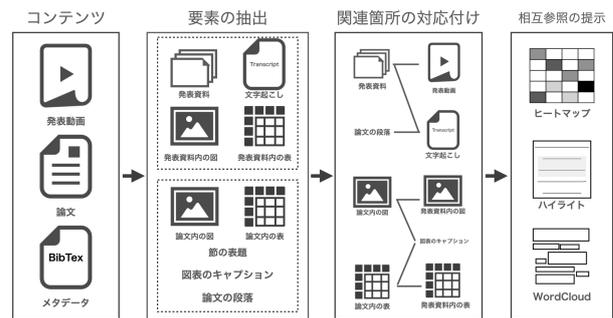


図3. コンテンツの処理から相互参照を提示するまでの流れ

させ、小規模なクラスターを最大数作成できた ϵ の値を採用した。

続いて図表とキャプションの紐付けを行う。しかし既存の図抽出ライブラリの図の認識精度は不十分だった。本研究では論文内の空白を認識し、領域分割することで図表を検出している。しかし領域分割をすると図表とキャプションが分離されたり、図表自体も小さな領域に余分に分割される場合がある。そこで図表が論文の本文よりも上に配置される IEEE 様式の特性を活かし、キャプションの直上に位置する領域を対応する図表として扱い、キャプションとの対応関係を記録する。ただし、本稿図4のように複数の小さな図表が組み合わさっている場合は、それらを全て包含する領域を対応する図表とした。

その他、画像処理ライブラリを用いて論文内の図表と発表資料内の図の対応関係を検出した。加えてタイムスタンプを用いることで、発表動画のタイミング、文字起こし、発表資料の対応関係も導き出すことができた。

4.3 ユーザーを補助する機能の開発

この段階では対応関係を相互参照の形でユーザーに提示するための可視化と論文の内容を表すための可視化を作成する。発表資料と段落間の類似度を見るための可視化としてヒートマップを採用し、選択した論文の段落をハイライトを用いて指し示すことにした。加えて論文の内容を一目で理解できるように、出現頻度によって単語の大きさを変えながら並べる可視化である Word Cloud[19] を採用した。

4.4 開発に用いたツール群

前段階で作成した可視化を繋ぎ合わせ、システムを開発した。要素の抽出・要素間の対応付には openCV[2], pyocr, pdftotext, Pillow[3] を使用している。システムのバックエンドは Flask, フロントエンドは Vue.js と WordCloud, 可視化は D3.js を使用している。

表 1. コンテンツ抽出の結果: 再現率 (適合率)

	論文 1	論文 2	論文 3	論文 4
発表資料の再構成	1.0(1.0)	0.89(0.96)	0.87(0.77)	0.84(0.89)
段落の抽出	0.90(0.83)	0.80(0.77)	0.83(0.79)	0.77(0.75)
段落の抽出 (分割ミスを除いた場合)	0.97(0.98)	1.0(1.0)	0.99(0.98)	0.96(0.98)
節の表題の抽出	1.0(1.0)	1.0(1.0)	0.94(1.0)	1.0(1.0)
図表キャプションの抽出	1.0(1.0)	1.0(1.0)	1.0(1.0)	1.0(1.0)
実行時間 (s)	87.55	174.44	262.37	169.65

5 実験

5.1 要素の抽出

IEEE の 2 カラム国際会議用の様式を採用した PacificVis 2022 の “Session 8: Text and Mixed Data” の 4 つの論文¹²³⁴に対し、開発した要素抽出手法を適用し実行時間を計測した。加えて発表資料の再構成、段落・表題・図表キャプションの抽出という 4 つのタスクについて精度を評価した。発表資料の再構成では動画から 10 秒ずつ切り出したフレームを人間の目で確認し、重複したフレームを削除することで発表資料を再構成する。その際に再現率と適合率を用いて、作成したものと自動処理した結果の一致率を計測する。続いて段落・表題・キャプションの抽出においても、人間の目で論文の原本を確認しそれぞれ正解データを作成した。その後自動抽出した結果と正解データを比較し再現率を計算した。以上の実験の結果が表 1 である。

表題とキャプションは再現率・適合率ともに 100% に近い精度で抽出することができた。一方発表資料の再構成と段落の抽出では論文 2~4 において精度が少し下がった。段落抽出の精度が下がった要因の多くは、段落を余分に分割してしまうなどの分割ミスである。実際に分割ミスを除いた場合の再現率は概ね 96% 以上と高い値を出している。発表資料の再構成で精度が下がった原因としては、発表資料内のアニメーションの使用が挙げられる。口頭発表では注目箇所を強調するためにアニメーションが使用されるが、その際にどこまでが同じスライドだと感じるかは視聴者によってばらつきが存在する。今回特に精度が低かった論文 3 では、スライド内の重要箇所を強調するためにペンを使って書き込みをしており、アルゴリズムでは書き込みのあるスライドを別のスライドとして認識したため精度が低下した。

5.2 関連箇所の対応づけ

続いて 5.1 節の実験で用いた中で特に精度がよかった論文 1 を用いて、追加の実験を行う。その際に確

認するタスクは発表資料と文字起こしの対応、発表動画と発表資料の対応、図表と図表番号の対応、段落と発表資料の対応、論文内の図表とスライド内の図表の対応の 5 つである。

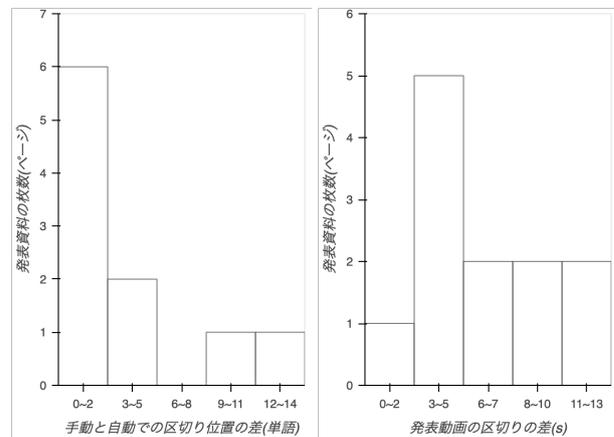


図 4. 一致箇所の検出結果 (左: 発表資料と文字起こしの対応, 右: 発表動画と発表資料の対応)

発表資料と文字起こしの対応 このタスクでは人間の手で発表資料の区切りに合わせて、文字起こしを分割し正解データを作成した。そして自動処理結果と正解データの分割位置が何単語ずれているのか比較した。その結果、図 4(左)のように平均のずれが 3.5 単語で、6 割は 2 単語以内のずれとなった。全体が 943 単語で 11 枚のスライドに分割されることを考えると、ずれは非常に小さいと言える。

発表動画と発表資料の対応 このタスクでは発表動画の画面が変わったと判断した時刻が、人間とアルゴリズムによって何秒ずれるのか計測した。その結果が図 4(右)で、3~5 秒ほどのずれが一番多いことがわかる。実験に用いた論文 1 の発表動画は 455 秒で 1 ページの平均秒数が 45 秒であることを考えると、差は十分に小さいと考えられる。

図表と図表番号の対応 5.1 節の実験で用いた論文 1~4 の全ての論文を使用し、36 組の図表とキャプションの紐付けを試みた。その結果全ての図表を抽出し、正しい図表番号を付加できることを確認できた。

¹ (Mishra+ 2022)10.1109/PacificVis53943.2022.00022

² (Zheng+ 2022)10.1109/PacificVis53943.2022.00023

³ (Wang+ 2022)10.1109/PacificVis53943.2022.00024

⁴ (Park+ 2022)10.1109/PacificVis53943.2022.00029

論文内の図表とスライド内の図表の対応 唯一論文とスライドで同じ図表が使用されていた論文4に着目し、提案アルゴリズムによる対応関係の検出精度を確認した。その結果再現率 0.64、適合率 1.00 で対応関係を取り出すことができた。失敗した原因はスライド内でのアニメーションの使用や画像の再配置が挙げられる。今後は類似スライドでの対応関係を参照することで、精度の向上が期待される。

論文の段落と文字起こしの対応 このタスクでは手動で論文と文字起こしを確認し、内容に関連性のある箇所を紐付け、正解データを作成する。アルゴリズムでは各スライドに対し類似度の高い上位5段落を紐づける。正解データと比較し再現率を計測する。その結果 TFIDF の再現率が 0.68 と一番高く、他の手法に比べ 0.3 以上高いという結果になった。

6 議論

6.1 自動抽出手法の精度に関する考察

本研究では要素の抽出と関連箇所の検出を行う手法を考案し、5章の実験を通し性能を測定した。その結果要素抽出タスク4つのうち2つのタスクで適合率 1.0 という非常に高い精度を確認することができた。加えて関連箇所の検出タスクにおいても、発表資料と発表動画・文字起こし間の対応関係を高い精度で検出することができた。そのため、提案手法を用いることでユーザーに対しある程度正確な対応関係を提供できることを確認することができた。一方で議論の余地のある以下の3点がある。

最初の点は発表動画からの発表資料の作成に伴う困難である。スライドウェアを用いた発表では頻繁にアニメーションが使用される。その際にどこまで一枚のスライドウェアとして認識するかについては人によってばらつきがあり、明確な正解を設定することは困難である。開発アルゴリズムでは少しでも画面に変化があった場合では違うスライドとして認識したが、今後はユーザーのニーズに合わせ、枚数を減らしたりする機能が必要であると考えられる。

つぎに文字起こしと論文の段落に関する関連箇所の検出精度の低さである。これには複数の理由が存在する。まず1点目は音声認識の精度の限界である。論文の中で定義した語句を認識しないことで、文章同士の類似度の精度が下がることがある。今後は文字認識を併用して利用することで、語句の認識を改善したいと考える。2点目の原因はスライドの内容が複数の段落の内容を包含していることである。その場合文字起こしと段落の内容が局所的に類似していても、全体の類似度は下がる傾向にある。今後は段落と文字起こしを文単位に分割し、文同士の類似度を計算することで性能を向上させたいと考える。最後は要素抽出にかかる計算時間の長さである。図1のように、長い場合では処理に4分以上の時間を

要することがある。一見するとこの計算時間は長く感じる。しかし論文の読み込みや発表動画の視聴に要する時間と比較すれば遥かに短い。加えて実行時間の50%以上は論文から図表やセクション・段落を抽出するための前処理時間である。論文は多くの場合発表当日よりも前に公開されるため、個々のユーザーに発表の前に論文の前処理を行ってもらうことで体感時間を減少させることができる。以上から現在の実行時間は十分に実用可能である。

6.2 要件に対する達成度

本研究では3.3節において、学会参加者を支援するための要件を (R1-R5) の5つ定義した。その上で現状のシステムの達成度について議論を行う。(R1)を達成するためにシステムでは論文選択画面内で論文の概要を表す WordCloud と論文概要の表示を行った。続いて (R2) を実現するために発表動画・発表資料ビューと論文ビューを導入し、コンテンツの本体を視聴できるようにした。そして (R3) を満たすために類似度ビューと論文グループビューと論文グループ詳細ビューを導入した。これらのビューではユーザーは関心のある箇所から関連のある箇所へと、コンテンツを超えて素早くアクセスすることができる。加えて論文グループ詳細ビューでは類似段落へのアクセスも容易にできるため、論文全体の構造も理解できる。一方、(R4, R5) を満たすための機能はまだ搭載していない。今後は (R4, R5) を満たすために新たなメモ機能とメモ共有機能を実装する予定である。その際にもメモと各コンテンツを紐づけることで学会参加者の作業量と認知負荷を軽減することを目指す。その事によってユーザーが研究を理解する過程において、関心のある論文の発見から研究を理解し、他者へと意見を共有するという一連のタスクを支援することが今後の目標である。

7 まとめ

本研究では学会で提供される様々なコンテンツの統合的な理解を目指し、相互参照に注目した新しい学会参加の支援の形を提案した。また支援を実現するために論文や発表動画から要素を取り出し、関連箇所を検出する手法を考案した。加えてユーザーが相互参照を目で見確認するためのシステムのプロトタイプの作成を行った。そして、提案アルゴリズムの性能の予備調整を行った。一方今後の研究の課題は2点ある。1点目はユーザー自身の考えを筆記し、それを他者へ共有することの補助である。2点目は提案手法の有効性を検証するための本格的なユーザー実験の実施である。システムを使用した場合と使用しなかった場合の学会参加・研究理解を比較することで、提案手法がユーザーの論文の内容理解の効率と認知負荷へどのような影響を及ぼすのか定性的に評価する。

謝辞

研究室の皆さんには常に刺激的な議論を頂きました。特に同期の牧修平氏、Chen Xueyi 氏には精神的にも支えられました。心より感謝申し上げます。

参考文献

- [1] Amazon Transcribe ホームページ. <https://aws.amazon.com/jp/transcribe/>.
- [2] openCV ホームページ. <https://opencv.org/>.
- [3] Pillow のホームページ. <https://pillow.readthedocs.io/en/stable/>.
- [4] PacificVis 2022 のホームページ. <https://pvis2022.github.io/pvis2022/>, 2022.
- [5] 阿辺川 武, 相澤 彰子. 内部構造解析機能と脚注表示機能を備えた論文閲覧システム. 人工知能学会第二種研究会資料, 2014(AM-07):03, 2014.
- [6] A. Adler, A. Gujar, B. L. Harrison, K. O'Hara, and A. Sellen. A Diary Study of Work-Related Reading: Design Implications for Digital Reading Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '98, pp. 241–248, USA, 1998. ACM Press/Addison-Wesley Publishing Co.
- [7] B. A. Aseniero, M. Constantinides, S. Joglekar, K. Zhou, and D. Quercia. MeetCues: Supporting online meetings experience. In *2020 IEEE Visualization Conference (VIS)*, pp. 236–240. IEEE, 2020.
- [8] F. Beck, S. Koch, and D. Weiskopf. Visual Analysis and Dissemination of Scientific Literature Collections with SurVis. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):180–189, 2016.
- [9] M. El-Assady, V. Gold, C. Acevedo, C. Collins, and D. Keim. ConToVi: Multi-party conversation exploration using topic-space views. In *Computer Graphics Forum*, Vol. 35, pp. 431–440. Wiley Online Library, 2016.
- [10] M. El-Assady, R. Sevastjanova, B. Gipp, D. Keim, and C. Collins. NEREx: Named-Entity Relationship Exploration in Multi-Party Conversations. In *Computer Graphics Forum*, Vol. 36, pp. 213–225. Wiley Online Library, 2017.
- [11] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, Vol. 96, pp. 226–231, 1996.
- [12] A. Head, K. Lo, D. Kang, R. Fok, S. Skjonsberg, D. S. Weld, and M. A. Hearst. Augmenting Scientific Papers with Just-in-Time, Position-Sensitive Definitions of Terms and Symbols. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.
- [13] E. Hu, M. A. R. Azim, and S. Heo. FluidMeet: Enabling Frictionless Transitions Between In-Group, Between-Group, and Private Conversations During Virtual Breakout Meetings. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [14] Q. Le and T. Mikolov. Distributed representations of sentences and documents. In *International conference on machine learning*, pp. 1188–1196. PMLR, 2014.
- [15] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- [16] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26, 2013.
- [17] M. Morris, A. Brush, and B. Meyers. Reading Revisited: Evaluating the Usability of Digital Display Surfaces for Active Reading Tasks. pp. 79–86, 10 2007.
- [18] P. Murali, J. Hernandez, D. McDuff, K. Rowan, J. Suh, and M. Czerwinski. AffectiveSpotlight: Facilitating the Communication of Affective Responses from Audience Members during Online Presentations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.
- [19] L. Oesper, D. Merico, R. Isserlin, and G. D. Bader. WordCloud: a Cytoscape plugin to create a visual semantic summary of networks. *Source code for biology and medicine*, 6(1):7, 2011.
- [20] K. O'Hara and A. Sellen. A Comparison of Reading Paper and On-Line Documents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, p. 335–342, New York, NY, USA, 1997. Association for Computing Machinery.
- [21] N. Roy, M. V. Torre, U. Gadiraju, D. Maxwell, and C. Hauff. Note the Highlight: Incorporating Active Reading Tools in a Search as Learning Environment. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, CHIIR '21, pp. 229–238, New York, NY, USA, 2021. Association for Computing Machinery.
- [22] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21, 2017.
- [23] J. Song, C. Riedl, and T. W. Malone. Online Mingling: Supporting Ad Hoc, Private Conversations at Virtual Conferences. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, 2021.

- [24] C. S. Tashman and W. K. Edwards. Active Reading and Its Discontents: The Situations, Problems and Ideas of Readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 2927–2936, New York, NY, USA, 2011. Association for Computing Machinery.
- [25] C. S. Tashman and W. K. Edwards. Liquid-Text: A Flexible, Multitouch Environment to Support Active Reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 3285–3294, New York, NY, USA, 2011. Association for Computing Machinery.

未来ビジョン

本研究では様々なデータ種別のコンテンツを統合することによって、ユーザーによる研究への深い理解を支援することを目的とした。その上で本論文では手始めとして、学会発表が行われた後を想定しユースケースを設定し、システムのプロトタイプを作成した。加えて簡易的な実験を通し、開発したデータ処理アルゴリズムの性能を定量的に評価した。しかしシステムの要件として定義した、「研究に対する疑問や意見・感想を記述する機能」と「記述した内容を他者へと共有する機能」は現状実現できていない。今後はユーザーのメモも論文や発表動画と関連づけることによって、発表後のユーザーの行動をさらに手厚く支援したいと考える。

本研究の大きなビジョンでは、学会参加における全てのプロセスを支援したいと考えている。学会参加者が行う行動は多岐に及んでお

り、発表を聴講する前の事前の準備から研究内容を理解しその内容を他者へと還元する部分まで様々なプロセスが存在する。その中で参加者は発表プログラムや著者の過去の研究論文、自身の記述したメモ、他者へ説明するために作成したスライドなど幅広いコンテンツに絶えず触れ続ける。そのためコンテンツの統合的な理解及び利用は学会参加における全てのプロセスにおいて必要不可欠であり、それを支援することで参加者の学会参加の方法も大きく変化すると考える。支援の一部としてプログラムから自動でカレンダーアプリへ予定を追加したり、発表中のメモを自動的に発表動画・発表資料と関連づけることなどが考えられる。参加者の負担となっている動作を相互参照によりサポートすることで、参加者はより深く研究の詳細を学び、発表者と接触する機会を有効に活用できるようになる。以上が本研究の描く未来ビジョンである。

Gino .Aiki: 合気道の身体の使い方の習得を支援するMRソフトウェア

鈴木 湧登* 坂本 大介* 小野 哲雄*

概要. 日本の伝統的な武道である合気道には、合気道独自の身体の使い方がある。これを習得すると、“小よく大を制す”ことができ、体格差に関係なく相手を制することができるようになる。しかし、その身体の使い方は身体感覚を伝えなければいけないため、分かりやすく伝えることが難しい。そのため、初学者にとって合気道の感覚を習得することの難易度が高い。そこで、本稿では合気道の身体の使い方の学習を支援する複合現実感 (Mixed Reality; MR) ソフトウェアを提案する。本ソフトウェアでは合気道の「押す」「上げる」「倒す」「下げる」といった技を仮想のMRオブジェクトとのインタラクションを通して体験することができる。この体験を通して合気道の基本となる身体の使い方を自然と学習する機会を提供する。本研究では、提案するソフトウェアを試験的に大学の合気道部で使用したり、一般の人を対象にしたワークショップを行ったり、労働現場での実証実験等を通して実用性と応用可能性について検討した。

1 はじめに

合気道とは合理的に身体を使うことで、体格や体力に関係なく相手を制することを目的とした武道である。合気道は小柄な人が多い日本人が、体格差に関係なく生き残るために長い歴史の中で編み出した護身術であり、守っていかなければならない日本の大事な文化の一つである [1, 2]。女性や子どもでも大男を投げ飛ばすことができることが合気道の特徴である [3, 4] が、一方で合理的な身体の使い方を学ぶことは簡単ではない。

合気道の練習は、型の練習と身体の使い方の練習 (身体操作の練習) がある [4]。型の練習とは、相手の攻撃に対して正確に決められた動きをするものである。型の見栄えが大事なため、技をかけられる方は抵抗してはならず素直に受身を取らなければいけない。しかし、街で不審者に襲われたときなどのような実戦時には、自分がいかに綺麗な技をかけても相手は技に抵抗してくる。抵抗してきた相手が自分よりも力のない人であれば、力づくで技をかけることで相手を制することができる可能性はあるが、小さい人が大きい人を相手にする場合は力負けしてしまい技をかけることができない。そこで必要となるのが“小よく大を制す”ことを可能にする身体の使い方 (力の出し方) である。これは腕の筋肉を使うのではなく体全体を使って技をかけることで実現できる。これにより小さな人でも大きな人を投げ飛ばすことができる。この2つの練習を行うことによって、体格差に関係なく相手を制することができるようになる。

しかしながら、合気道の身体の使い方は言語化することが難しい。例えば練習においても「近くのボールを軽く取るような気持ちで技をかける」、「食

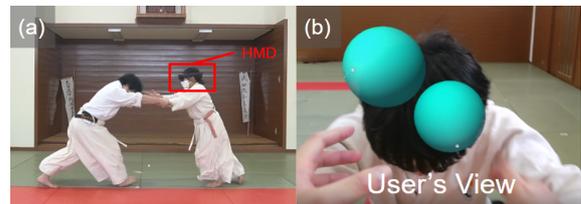


図 1. (a) システムを利用している様子. (b) ユーザの視点.

べものを口に運ぶ時と同じ力の使い方をして技をかける」などといった曖昧な指導が多い。実際に上記のような教えの通りに技をかけると“小よく大を制す”ことはできるようになるが、初心者は簡単にその感覚を掴むことが出来ない。また、指導者に1対1で指導してもらえる時間が少ないのも現在合気道が抱える問題の一つである。合気道の身体の使い方は指導者の実演を見て真似すればできるようになるわけではないので、1対1の指導が必要である。しかし、合気道は生徒の数に対して指導者の数が少ないので指導者が生徒一人当たり割ける指導時間も限られる。ゆえに、生徒が指導者に1対1で指導してもらえる時間は非常に短い。よって、指導者なしでもある程度のところまでは学べるようなシステムを構築することで合気道の自主学習を支援することができるだろう。

そこで本稿では、合気道の指導などで身体の使い方の感覚を表現するためによく言われている比喻表現をユーザの周りに仮想オブジェクトとして視覚的に提示するMR合気道学習支援システムGino .Aiki (ぎのうどっとあいき) を提案する (図 1)。これにより、今まで直感的に伝えることが難しかった合気道の身体の使い方の感覚を簡単に共有することができ、かつ、学習者個人が自立的に練習することがで

* 北海道大学

きるような環境を提供する。なお、本稿での合気道は武田流中村派合気道を指すものとする [5]。

2 関連研究

2.1 身体技能の習得支援

身体技能の習得を支援するための方法については今まで多くの研究がなされている。Ikedaらはゴルフの動作速度に関係なく、自分の動作と熟練者の動作をリアルタイムに重ねることのできるアルゴリズム (Decayed Dynamic Time Wrapping) を開発し、その違いを第三者視点から観測できるように Augmented Reality (AR) 環境を提案した [6]。YanらはHMDを用いて三人称視点から自分の動きを確認できるシステムを開発し、ダンサーに自分の姿を客観的に観察しながら練習できる環境を提供した。このような練習方法は、いわゆる「手本を見てまねる」「鏡を見てフォームを確認する」などの従来の練習方法を拡張したものだと言える。この他にも、嵐らのタットダンスの練習支援システム [8] や松井らのヘアアイロンの操作軌道の提示システム [9] など、視覚フィードバックを与えることによって正しい動きや姿勢を学習する方法を検討した研究は多い [10, 11, 12, 13]。

また、聴覚や触覚によるフィードバックを与えることで学習支援をするシステムを提案している研究もある。山本らはWebカメラを用いてユーザの体幹トレーニングの姿勢の良し悪しを判断し、BGM (background music) の再生速度の変化を用いてユーザに姿勢の変化を知らせるシステムを提案した [12]。菅家らは演奏フレーズに同期し、演奏者に叩打すべき手足に対して触覚提示を行い、手足を動作させる感覚を経験させることでドラム学習の図る手法を提案した [15]。さらに、視覚、聴覚、触覚フィードバックを組み合わせた学習支援の研究もある。Nakamuraらはゴルフスイングの軌道予測と非接地型力表示を用いてスイング中のスイングモーションの修正を実現するシステムを考案した [16]。これは理想の軌道からずれると正しい軌道になるようにクラブに力が加えられるというもので、視覚表現と触覚表現を組み合わせていると言える。

一人で練習できなかつたり室内で練習できないような技術の習得に対しては、Virtual Reality (VR) などの技術で練習環境を仮想的に再現する手法が取られていることが多い。Lopesらは、筋肉電気刺激で腕や足を刺激することによって、実際のスポーツにおいて受けるであろう触覚フィードバックを与えるシステムを提案しており、これをVRを組み合わせることで実際の練習に近い体験を提供している [17]。Hoffardらは、触覚フィードバックのあるVRスキートレーニングシステムを開発し、室内でスキーの練習を行えるようにした [18]。その他にもOagazら

のVR環境での卓球トレーニングシステム [19] やYokokuboらのTracKenzan [20] など環境を再現する手法は多く研究されている。

さらに、環境をそのまま再現するのではなく、再現する環境の一部のパラメータを現実のものとは異なるようにする環境歪曲手法も検討されていたり [21, 22, 23]、軌道予測の提示による練習手法も提案されている [24, 25]。また、筋肉電気刺激による直接的な技術習得を試みている研究もあり [26, 27]、Ebisuらは音楽演奏の初心者にとって難しい正確にリズムをとる動作をユーザの腕と脚に電気を流すことによって習得させた [28]。このように今まで様々な学習支援手法が提案されてきたが、比喩表現を視覚的に表示することで学習を支援させるような手法はあまり研究されていない。類似の手法としては、坂名らのVR映像を用いた視覚刺激による歌唱時の発声支援システム [29] があるが、これは比喩表現の視覚化について検討したものではない。

2.2 技能習得の支援対象

身体の使い方の学習を支援する研究としてはスポーツの技能獲得支援が盛んに研究されてきている。これまでにスキー [13, 18, 23]、スノーボード [30]、ゴルフ [16]、野球 [11]、卓球 [19, 21, 25, 31]、テニス [32] など様々なスポーツの学習支援が行われている。また、土田らの研究 [33, 34, 35] をはじめとするダンスの練習支援システムや音楽演奏支援に関する研究などエンターテインメントの分野でも多くの研究がなされている [8, 12, 15, 28, 29]。さらに、医療における手術のトレーニングシステム [36, 37] や、職場での作業技術習得支援 [38, 39] なども行われている。

武道の学習支援もよく行われている。Wennrichらが開発したVRTe doは、モーションキャプチャとVR技術に基づいた空手の型のトレーニングシステムで、ゲーム方式で空手の型を学ぶことができる [40]。Hanらは四方八方に太極拳のコーチをAR技術で表示し、頭の向きを気にせず正解の型を見ながら太極拳の型を学習できるようなシステムを提案した [41]。他にも柔道やテコンドーなど様々な武道に対しての学習支援が研究されている [42, 43]。

合気道に対しても習得支援は考えられており、代表的なものに根木らの行動的コーチングによる合気道の技の改善についての研究 [44] やCorbíらのMyShikko [45] などの提案がある。合気道の型の練習を支援するスマートホン向けアプリケーションも開発されていたり [46]、昨今のCOVID-19の影響を鑑みて、オンラインでの合気道指導に関する研究もKantzaraによって行われている [47]。しかしながら、根木らの研究やKantzaraらの研究は「初心者には合気道の感覚をつかむのが難しい」という問題点と「指導者のいない環境で自律的に学習するというこ

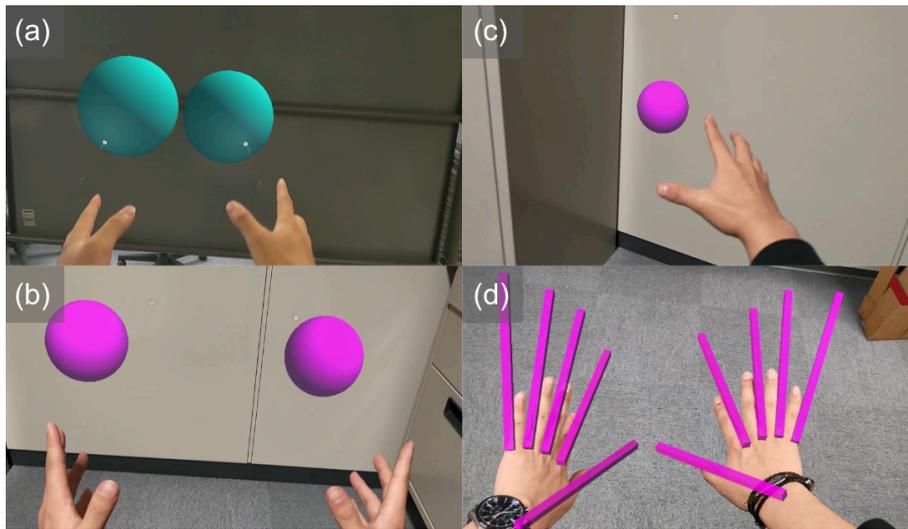


図 2. (a)「押す」、(b)「上げる」、(c)「倒す」、(d)「下げる」動作を習得する比喩を視覚表現化したもの。

とができない」という既存の練習の2つの問題点を解決しているとはいえ、CorbíらのMyShikkoやKantzaraらの合気道の型を学習できるアプリケーションは、合気道の型の練習支援に関するものであり身体の使い方の練習を支援しているとは言えない。その他の合気道に関する研究は、合気道の精神性や身体性に関する研究[48, 49]や合気道と教育との関係性に関する研究[50]、合気道の科学的解析に関する研究[51, 52, 53]などの合気道の習得支援ではないものになってしまうため、合気道の身体の使い方の習得を支援する提案や研究は未だされていない。

3 提案システム Gino .Aiki

本稿では、MR デバイス (MR グラス) である Microsoft 社の HoloLens 2 を利用してユーザの体の周りに視覚映像を追加することで、合気道の身体の使い方の習得を支援するソフトウェア (Gino .Aiki) を提案する。従来の合気道の練習において「近くにあるボールを軽く取るような気持ちで技をかける」、「食べものを口に運ぶ時と同じ力の使い方をして技をかける」といった言葉で表現されていた感覚を MR 技術で可視化することで、言語を介せずに合気道の身体感覚を習得することを支援する。例えば、「手の先にあるボールを軽く取るような」感覚を呼び起こす時には、ユーザの手の少し前方に仮想的なボールを置いて、ユーザにこのボールを取らせようとする。これによって、ユーザは自分の意志に関係なく自然と合気道的な身体の使い方をすることが出来る。これにより、言語を介せずに身体感覚を伝えられ、習得に必要な時間が短縮されることが期待される。これは、合気道で大事な「折れない腕」、「腕の張り」などのような身体感覚が、比喩によって誘発される動作により自然に行われること

で実現されるためである。また、指導者がいなくても本ソフトウェアだけで練習できるので、学習者同士で技を掛け合うことで自立的に学習することができる環境を実現する。

3.1 システム概要

本ソフトウェアはゲーム開発エンジン Unity を用いて開発した (Unity 2019.4.19f1)。また、ソフトウェア内のメニューやボタンといったオブジェクトやハンドトラッキング技術には Microsoft の Mixed Reality Toolkit 2 (MRTK) を用いた。Gino .Aiki は合気道の身体の使い方を自然と体験させるためのソフトウェア群であり、現在までに4つのソフトウェアを開発してきている。それぞれ合気道の「押す」「上げる」「倒す」「下げる」を習得することを目的に開発されている。各ソフトウェアには MRTK のハンドトラッキング技術が使われており、主に人差し指の指先の座標を元に仮想オブジェクトの位置を決定している。また、各ソフトウェアにおける仮想物体の大きさや表示する位置などは、事前に簡単なユーザテストを行いながら探索的に策定した。以下、各ソフトウェアの具体的な技術構成とシステムの動作を示す。

3.2 「押す」動作の習得

合気道の「押す」動作を体験させるために「目の前にボールがあると思って、そのボールを取るように技をかける」という比喩表現を可視化した。ここで「押す」動作とは、両手首を相手に掴まれた状態から相手を後ろに押し倒す動きのことである。本ソフトウェアにおいては MRTK のハンドトラッキング機能を利用してユーザの手の先 30cm のところに仮想的なボールを用意し、それをユーザに取ってもらう動作となる (図 2(a))。仮想的なボールは常に



図 3. 合気道部での試験的導入の様子。画像は部員が「下げる」動作を習得できるソフトウェアを使用している様子。

ユーザの手の先にくる（ユーザの両手の Ray を取得し、取得した座標から Ray 方向に 0.3m 先に仮想オブジェクトを表示する）ようプログラムされており、ユーザがボールを取ろうとしても実際にボールを取ることはできない。しかし、そのボールを取ろうとするときの人間の身体の使い方が「折れない腕」や「腕の張り」といった合気道の身体の使い方に酷似しているため、ボールを追いかけていくだけで勝手に合気道の身体の使い方ができていることになる。これにより、ユーザは合気道の身体の使い方の感覚を掴むことができる。この「押す」動作を習得するソフトウェアを球体捕縛と名付けた。

3.3 「上げる」動作の習得

「押す」動作の習得の時に用いた「ボールを取ろうとする」比喩を上方向に向けることで合気道の「上げる」動作の習得を支援するソフトウェアを作成した（図 2(b)）。「上げる」動作は、両手の手首を相手に掴まれた状態から自分の手を上に上げ相手の姿勢を崩す動作のことである。ユーザは自分の頭の動きに対して常に正面に表示されるようにプログラムされているメニューを片手の親指と人差し指でつまみ、自分の目先まで持ち上げる。その位置でキャリブレーションボタンを押すことで仮想オブジェクトがユーザの膝の前付近に表示される。ソフトウェア内では、キャリブレーションボタンを押した時のユーザの人差し指の指先の座標を取得しそれを元にボールの表示位置が決定される仕組みになっている。その状態でメニューの中の Anti-Gravity ボタンを押すと膝にあったボールが上に上がっていく。ユーザはその上がっていくボールを取りに行こうとすることで勝手に相手を上げることが出来る。これは、ボールを取りに行く動作が「腕を脱力し、肩甲骨付近から力を出す」という合気道の感覚と似ているために実現される。メニューにはリセットボタンがあり、これを選択することで利用を中止することができる。再度利用する際には、再キャリブレーションを行う。この「上げる」動作を習得するソフトウェアを重力反転と名付けた。

3.4 「倒す」動作の習得

合気道の「倒す」という動作を習得させるために、ボールが遠のいていくソフトウェアを作成した（図 2(c)）。「倒す」動作は、両手の掌を合わせて耐えている相手を片手で横に倒す動作である。キャリブレーションの仕組みやボールを追いかけるという基本的な仕組みは「上げる」動作の習得のソフトウェアと同じであるが、ボールの動く方向が異なる。メニュー内のキャリブレーションボタンを押すと、ボールがユーザの左手前に表示され、それぞれのユーザに適した場所にボールが現れる。その後、Free-Fall ボタンを押すと、ボールが斜め左奥の方に進んでいくので、ユーザがそのボールを取ろうとすると自然に相手を「倒す」ことができる。これはボールを取りに行く感覚が合気道の「折れない腕」の感覚と似ているために実現される。「上げる」動作と同様、再度学習する時は、リセットボタンとキャリブレーションボタンを用いて位置調整をし直す必要がある。この「倒す」動作を習得するソフトウェアを重力正転と名付けた。

3.5 「下げる」動作の習得

合気道の「下げる」動作を習得させるために、「指を長くして、その指で相手を落とす」という比喩表現を視覚化した（図 2(d)）。「下げる」動作は、両手を組んで耐えている相手を両手で下に落とす動作のことである。ユーザの両手の全ての指の指先の座標を取得し、それぞれの座標に指方向平行に 30cm の棒状の仮想オブジェクトを配置することで実現される。本ソフトウェアでは、「上げる」「倒す」を学習するソフトウェアと異なり、キャリブレーションの必要はない。ソフトウェアを起動した後にユーザは、その仮想的に大きくなった手を以って下に手を下げる。この感覚が「腕の張り」といった合気道の「下げる」動作の感覚に酷似しているため、ユーザは合気道の「下げる」動作のコツを簡単に掴むことができる。この「下げる」動作を習得するソフトウェアを五指拡張と名付けた。

4 身体技能獲得効果の予備的検討

本ソフトウェアで提案する技術学習支援がどの程度効果があるのかを確認するために、いくつかワークショップや体験会を行った。これらの予備検討では、既存の合気道習得方法である「口頭による伝達」と「本ソフトウェアによる身体の使い方」の両方を体験してもらい、それぞれの学習方法についての感想などのインタビューを行った。

4.1 合気道部での試験的導入

北海道大学武田流中村派合気道部¹の部員を対象に実際に合気道の稽古の最中に本ソフトウェアを試験的に導入し、練習をしてもらった(図3)。参加人数は11名(男性7名, 女性4名)で、全員が大学生であった。試験的に導入した結果、「いつもの練習よりも分かりやすい」「他の技にも応用できそう」などの感想を得た。また、部員の中には、本システムを使ってより高度な技の練習を行う者もいた。一方で、黒帯を取得しているような者の中には自分の感覚と異なると感じる者もいた。具体的には、合気道の感覚で重要な「膝の抜き」や「重心の移動」などの全身動作の練習の感覚が本ソフトウェアにより提供されていないことに違和感を覚えるという指摘があった。また、参加者の1人は「ボールを掴む」という比喩自体が分かりにくいと回答していた。

4.2 一般の方向けの体験会

2021年10月に札幌市で開催されたNoMaps²においてGino .Aikiのワークショップを開催した(図4)。開催にあたっては新型コロナウイルス感染拡大防止策を徹底して行った。ワークショップにおいては一般の方々に既存の方法で合気道を体験してもらったあとに、本ソフトウェアで合気道の習得体験(「押す」動作と「上げる」動作)を行った。加えて、体験した動作を日常動作に応用する機会を設け、重たい荷物を上げる、重たいものを押すといった動作がどの程度楽になるのかを体験してもらった。参加者は11名(男性8名, 女性3名)であった。また、練習時間は15分ほどであった。

この結果、従来の学習方法では難しかった「折れない腕」、「腕の張り」といった感覚や「掴まれている場所を意識しない感覚」が、本ソフトウェアでは自然に掴むことができた。特に11名中9名が実際に相手を押し倒すことができたり上げたりすることができた。また、既存の練習方法では全員が規定の時間(15分)のうちに合気道のコツを掴むことはできなかったが、本手法では、11名中9名が時間内に



図 4. NoMaps2021 でのワークショップの様子。

「折れない腕」の感覚をはじめとする合気道の感覚を掴むことができた。さらに、アンケートでは、「ARの方が分かりやすかった」「日常動作が楽になった」等の感想を得た。インタビューにおいても同様の感想を得られ、参加者が本ソフトウェアの利用で何らかの感覚を掴めた様子を観察することができた。また、「合気道の習得以外にもスキーの学習にも使えるのではないか」という意見もあった。また、小柄な女性が男性を後ろに押し倒している様子も観測された。一方で、ソフトウェアのキャリブレーションの段階で躓いている人も何人か見られた。特に、MRソフトウェア内での「メニューのボタンを押す」動作を行えない者が多く、本ソフトウェアの体験ができないこともあった。また、「ボールが消える」という現象も発生し、システムのハンドトラッキングが一部機能していないことが確認された。

4.3 労働現場での試験的導入

合気道の「上げる」動作を習得することで、通常の身体の使い方では重いものを持ち上げるよりも身体的負荷を抑えたまま上げることができると言われている。その効果を調査するために、日本国内のデータセンターで、サーバの持ち上げ作業の負荷が軽減できるかどうかの予備的検討を行った。

検討では4名(男性4名)の参加者を2名ずつの2つのグループに分け、片方のグループは既存の方法(口頭伝達)で合気道を学習し、もう片方のグループはGino .Aikiで合気道を学習した。その後、通常業務(20kgのサーバを二人で持ち上げる)を行ってもらった。タスク終了後、合気道の学習によって体が楽になったかどうかや練習はしやすかったかなどのインタビューを行った。

この結果、既存の方法で合気道を学習した場合には、「合気道の(口頭での)比喩が分かりづらく、いまいち合気道の効果を実感できなかった」という感想が聞かれた。提案手法で合気道を学習した場合には「比喩が視覚情報として表れていて直感的に分かりやすいため、合気道の効果を感じる事が出来た。

¹ 北海道大学武田流中村派合気道部。 <http://t-n-aiki-hokudai.sub.jp/>

² NoMaps2021. ARで合気道を体験するワークショップ。 <https://no-maps.jp/ar-aikido/>

普段の業務にも応用できそう」などの感想が聞かれた。また「仮にこのシステムが要求する合気道習得のための練習時間が1日あたり1時間であったとしても使いたい。」「入社1年目の人に有効であると思う。」「合気道以外の他の技術の習得にも使えそう。」などの感想も聞かれた。一方で、若い人は元々力がある人が多く、本ソフトウェアによる支援がなくても困らないという感想も聞かれた。また、合気道の「上げる」動作の時の手の形とサーバを上げる時の手の形が異なるため、ソフトウェアで学習してもそのまま応用しにくいという意見もあった。実施した予備的検討を通して、総じて本システムがサーバのような重量物の上げ下げによる身体的負荷を軽減することができる可能性が示唆された。

5 議論

予備的検討では、提案するソフトウェアを短期間利用することで、本ソフトウェアを使用しなくても合気道の身体の使い方を覚えているユーザが確認された。これはMRによる比喩表現の可視化手法の身体記憶性が高いからではないかと考えられる。ここでの身体記憶性が高いとは、一度体の感覚を覚えてしまえば、時間が経っても大体の感覚を覚えているという本稿で提案する考え方であり、自転車やスキーなどはその好例だと考える。この現象は、比喩は一度視覚映像として共有してしまえばあとで自分でも鮮明に比喩の様子や挙動を思い返すことができるから起こったものだと考えられる。そのため、合気道に限らず、他人の感覚を比喩として映像で学習すれば、かなり短い時間でその感覚を掴むことができるだろう。また、予備的検討の中で、「合気道の習得以外にもスキーの学習にも使えるのではないか」、「合気道以外の他の技術の習得にも使えそう」などの感想があった。これらは他のスポーツや作業においても比喩表現は一般的に使用されており、これをMR技術によって可視化することは合気道以外にも需要があることを示している。他の武道やスポーツなどにも応用できる可能性があるだけでなく、個人ごとの感覚を共有できる可能性を秘めている。そのため、人それぞれ自分に合った比喩表現で学ぶといったことができる可能性がある。

一方で技術的な課題は残っている。特に、MRグラスの視野角が狭いことは、学習の効率を妨げる要因の一つとなっている。本手法を用いる場合、比喩表現をユーザの周りの空間全てに表示することが望ましいが、現状の視野角では肩や首など顔に近すぎる部分に比喩表現を表示することはできない。また、比喩を表現化するにはその感覚を掴んでいる人がMRソフトウェアを作らなければならないというのも課題の一つである。ある伝統的な技術を伝承させたいときに、本手法を用いて比喩表現で視覚的に伝えることができれば短期間で感覚をつかむことが

できるかもしれないが、そのMRソフトウェアは伝承させたい人が作らなければならない。感覚を共有したい人がコンピュータでMRソフトウェアを作れる人であれば問題ないが、そうでない人の場合は本手法は手軽に採用できるものではない。加えて、合気道の「押す」動作において、手首を掴まれたときにハンドトラッキングの精度が下がることも課題である。さらに、本手法では「仮想のボールを取ろうとしてみてください」などの言語的インタラクションを行なっている。これでは、厳密には自立したシステムとは言えないため、音声案内を用意するなどの対策が別途必要である。

また、言語的比喩をMR表現した手法自体の限界もある。比喩表現で体の感覚を誘発できるのは上半身に限られる上、表示される比喩に多少の解釈の違いが生まれる。さらに、合気道の有段者には、自分の感覚と表示されている比喩の感覚が違うため、学習にならない可能性もある。

また、本稿では本格的な評価実験を行っていないため、本稿で実施した予備的検討を持って本手法に十分な効果があるとは定量的および定性的に断言することはできない。本手法により合気道の学習支援に効果があると示すためには、既存の練習方法とMRでの練習方法で時間を決めて練習し、どの程度合気道が上達したかを定性的、可能ならば定量的に評価する必要がある。しかしながら、合気道の習得度合いを示す定量的な指標は研究されていないため、信頼できる指標を作成し、その上でGino .Aikiの効果を検証する必要がある。

6 まとめ

本稿では、合気道の身体の使い方の学習を支援するMRソフトウェアGino .Aikiを提案した。本ソフトウェアは合気道の「押す」「上げる」「倒す」「下げる」の4つの身体の使い方に注目し、その身体の使い方のコツをMR表現を用いて視覚化することで合気道独自の身体の感覚をユーザに体験させた。これにより「習得するのが難しい」「一人で学習するのが難しい」といった合気道の習得における課題を解決することを試みた。本ソフトウェアの効果を検討するために、ワークショップや実際の作業現場において予備的検討を行った。この結果、一般の人々を対象としたワークショップでは「ARによる合気道体験のほうが合気道の身体の使い方が分かりやすかった」「日常動作が楽になった」などの感想を得た。この結果からGino .Aikiは合気道の身体の使い方を短期間で体験、習得させ、身体的な負荷を軽減させる効果がある可能性が示唆された。一方で、視野角の狭さの問題やMRコンテンツの制作の問題については本ソフトウェアおよびシステムの改善点としてあげられる。今後は定量的な技能獲得についての実証に取り組んでいく。

謝辞

本研究の一部は、情報処理推進機構 (IPA) 「2021 年度未踏 IT 人材発掘・育成事業」およびノーステック財団「研究開発助成事業」の補助を受けて実施した。

参考文献

- [1] ガスバル, ロランド.R, 井上誠治. 日本の伝統スポーツの文化性に関する研究. 1997.
- [2] 安道太軌・濱田初幸・川西正志・北村尚浩: 中学生の武道に対するイメージ, 日本体育学会第 63 回大会体育社会学専門領域発表論文集, 128-133, 2012.
- [3] 森山達矢. 合気道の稽古者はいかに倫理性を身につけるのか: 意味生成体験という観点から. 体育学研究, 58 (1): 225-241. 2013.
- [4] 北海道 綜武館 . <https://www.takedaryu-hokkaidousoubukan.com/>令和 4 年 1 月 2 1 日 閲覧.
- [5] 武田流中村派合気道 . サイバー 綜武館 . <https://www.takedaryu.jp/> 令和 4 年 9 月 21 日 閲覧.
- [6] A. Ikeda, D. H. Hwang and H. Koike, "AR based Self-sports Learning System using Decayed Dynamic TimeWarping Algorithm", ICAT-EGVE 2018-International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, 2018.
- [7] S. Yan, G. Ding, Z. Guan, N. Sun, H. Li, and L. Zhang. OutsideMe: Augmenting Dancer's External Self-Image by Using A Mixed Reality System. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '15, p. 965-970, New York, NY, USA, 2015. Association for Computing Machinery.
- [8] 嵐 冠太, 塚田 浩二. タットダンスの練習支援システムの提案. インタラクシオン 2021 論文集, pp.277-280, 2021.
- [9] 松井 菜摘, 大西 鮎美, 寺田 努, 塚本 昌彦. ゲート・スティック・ゴーストを用いたヘアアイロンの操作軌道提示手法の検討. WISS 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. pp.57-63, 2021.
- [10] N. Hamanishi and J. Rekimoto. Temporal Manipulation Interface of Motion Data for Movement Observation in a Personal Training. In Symposium on Spatial User Interaction. 1-2. 2020.
- [11] 塚本 裕樹, 角 薫. 運動イメージを用いた投球トレーニングシステムの研究. インタラクシオン 2015 論文集. pp.438-441, 2015.
- [12] 三浦 駿, 安藤 敏彦. AR を用いたギター演奏学習支援システムの一提案. インタラクシオン 2020 論文集. pp.461-463, 2020.
- [13] X. Zhang, E. Wu and H. Koike. "Watch-Your-Skiing: Visualizations for VR Skiing using Real-time Body Tracking," 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 2021, pp. 387-388, doi: 10.1109/ISMAR-Adjunct54149.2021.00088.
- [14] 山本 鷹人, 北原 鉄朗. BGM の再生速度変化を用いた体幹トレーニング支援システム. インタラクシオン 2020 論文集. pp.212-217, 2020.
- [15] 菅家 浩之, 寺田 努, 塚本 昌彦. 触覚提示に基づくドラム演奏学習支援システム. インタラクシオン 2015 論文集. pp.832-834, 2015.
- [16] T. Nakamura, D. Saito, E. Wu, and H. Koike. Actuated Club: Modification of Golf-Club Posture with force feedback and motion prediction in VR environment. In ACM SIGGRAPH 2020 Emerging Technologies (SIGGRAPH '20). Association for Computing Machinery, New York, NY, USA, Article 12, 1-2. 2020. <https://doi.org/10.1145/3388534.3407286>
- [17] P. Lopes, A. Ion, and P. Baudisch. Impacto: Simulating Physical Impact by Combining Tactile Stimulation with Electrical Muscle Stimulation. In Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15). Association for Computing Machinery, New York, NY, USA, 11-19. 2015. <https://doi.org/10.1145/2807442.2807443>
- [18] J. Hoffard, T. Nakamura, E. Wu, and H. Koike. PushToSki - An Indoor Ski Training System Using Haptic Feedback. In ACM SIGGRAPH 2021 Posters (SIGGRAPH '21). Association for Computing Machinery, New York, NY, USA, Article 15, 1-2. 2021. <https://doi.org/10.1145/3450618.3469158>
- [19] H. Oagaz, B. Schoun and M. -H. Choi. "Performance Improvement and Skill Transfer in Table Tennis Through Training in Virtual Reality" in IEEE Transactions on Visualization and Computer Graphics. 2021. doi: 10.1109/TVCG.2021.3086403.
- [20] A. Yokokubo, Y. Kato, I.Siio. TracKenzan: Digital Flower Arrangement Using Trackpad and Stylus Pen. In: Stephanidis, C., Kurosu, M., Degen, H., Reinerman-Jones, L. (eds) HCI International 2020 - Late Breaking Papers: Multimodality and Intelligence. HCII 2020. Lecture Notes in Computer Science(), vol 12424. Springer, Cham. 2020. <https://doi.org/10.1007/978-3-030-60117-1-25>
- [21] E. Wu, M. Piekenbrock, T. Nakamura, and H. Koike. SPinPong - Virtual Reality Table Tennis Skill Acquisition using Visual, Haptic and Temporal Cues. IEEE Transactions on Visualization and Computer Graphics, 27(5):2566-2576, 2021.
- [22] 川崎 仁史, 脇坂 崇平, 笠原 俊一, 齊藤 寛人, 原口 純也, 登嶋 健太, 稲見 昌彦. けん玉できた! VR: 5 分間程度の VR トレーニングによってけん玉の技の習得を支援するシステム. エンタテインメントコンピューティングシンポジウム 2020 論文集, pp. 26-32, 2020.

- [23] 松本 高, E. Wu, 小池 英樹. 時間歪曲機能を用いた VR スキートレーニングの拡張. WISS 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. pp.64-70, 2021.
- [24] E. Wu and H. Koike, "FuturePose - Mixed Reality Martial Arts Training Using Real-Time 3D Human Pose Forecasting With a RGB Camera," 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1384-1392, 2019. doi: 10.1109/WACV.2019.00152.
- [25] E. Wu and H. Koike. FuturePong: Real-time Table Tennis Trajectory Forecasting using Pose Prediction Network. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1-8. 2020. <https://doi.org/10.1145/3334480.3382853>
- [26] S. Kasahara, J. Nishida, and P. Lopes. Pre-emptive Action: Accelerating Human Reaction using Electrical Muscle Stimulation Without Compromising Agency. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 643, 1-15. 2019. <https://doi.org/10.1145/3290605.3300873>
- [27] A. Takahashi, J. Brooks, H. Kajimoto, and P. Lopes. Increasing Electrical Muscle Stimulation's Dexterity by means of Back of the Hand Actuation. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 216, 1-12. 2021. <https://doi.org/10.1145/3411764.3445761>
- [28] A. Ebisu, S. Hashizume, K. Suzuki, A. Ishii, M. Sakashita, and Y. Ochiai. Stimulated percussions: method to control human for learning music by using electrical muscle stimulation. In Proceedings of the 8th Augmented Human International Conference (AH '17). Association for Computing Machinery, New York, NY, USA, Article 33, 1-5. 2017. <https://doi.org/10.1145/3041164.3041202>
- [29] 坂名 純太, 土田 修平, 寺田 努, 塚本 昌彦. 歌唱時に高音の発声ができるようになる VR 映像の提案と評価. WISS 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. pp.16-22, 2021.
- [30] B. Hament, A. Cater and P. Y. Oh, "Coupling virtual reality and motion platforms for snowboard training," 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), pp. 556-560, 2017. doi: 10.1109/URAI.2017.7992668.
- [31] S. C. Michalski, A. Szpak, D. Saredakis, T. J. Ross, M. Billingham and T. Loetscher, "Getting your game on: Using virtual reality to improve real table tennis skills", PLoS ONE, vol. 14, no. 9, 2019.
- [32] P. Le Noury, T. Buszard, M. Reid, and D. Farrow, "Examining the representativeness of a virtual reality environment for simulation of tennis performance," Journal of Sports Sciences, vol. 39, no. 4, pp. 412-420, 2021.
- [33] S. Tsuchida, T. Terada, and M. Tsukamoto. A system for practicing formations in dance performance supported by self-propelled screen. In Proceedings of the 4th Augmented Human International Conference (AH '13). Association for Computing Machinery, New York, NY, USA, 178-185. 2013. <https://doi.org/10.1145/2459236.2459266>
- [34] S. Tsuchida, T. Terada, and M. Tsukamoto. A Dance Performance Environment in which Performers Dance with Multiple Robotic Balls. In Proceedings of the 7th Augmented Human International Conference 2016 (AH '16). Association for Computing Machinery, New York, NY, USA, Article 12, 1-8. 2016. <https://doi.org/10.1145/2875194.2875219>
- [35] 土田 修平, 岡本 秀明, 鈴木 裕真, 金田 麟太郎, 堀 隆之, 寺田 努, 塚本 昌彦. ダンスをマスターした自身の映像を先に見ることによるダンス学習支援. 研究報告ヒューマンコンピュータインタラクション (HCI), 2022(6), 1-8.
- [36] P. -A. Heng et al. "A virtual-reality training system for knee arthroscopic surgery," in IEEE Transactions on Information Technology in Biomedicine, vol. 8, no. 2, pp. 217-227, June 2004. doi: 10.1109/TITB.2004.826720.
- [37] T. Miki, T. Iwai, K. Kotani, J. Dang, H. Sawada, and M. Miyake, Development of a virtual reality training system for endoscope-assisted submandibular gland removal, J. Cranio-Maxillofac. Surg. 44(2016), No. 11, p. 1800.
- [38] I. Galvan-Bobadilla, A. Ayala-García, E. Rodríguez-Gallegos and G. Arroyo-Figueroa, "Virtual reality training system for the maintenance of underground lines in power distribution system," Third International Conference on Innovative Computing Technology (INTECH 2013), pp. 199-204, 2013. doi: 10.1109/INTECH.2013.6653713.
- [39] M. Perez-Ramirez, G. Arroyo-Figueroa, A. Ayala. The use of a virtual reality training system to improve technical skill in the maintenance of live-line power distribution networks. Interactive Learning Environments 29:4, pages 527-544. 2021.
- [40] K. Wennrich, B. Tag, and K. Kunze. 2018. VRTe do: the way of the virtual hand. In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18). Association for Computing Machinery, New York, NY, USA, Article 63, 1-2. 2018. <https://doi.org/10.1145/3281505.3281578>
- [41] P. Han, Y. Chen, Y. Zhong, H. Wang, and Y. Hung. My Tai-Chi coaches: an augmented-learning tool for practicing Tai-Chi Chuan. In Proceedings of the 8th Aug-

- mented Human International Conference (AH '17). Association for Computing Machinery, New York, NY, USA, Article 25, 1–4. 2017. <https://doi.org/10.1145/3041164.3041194>
- [42] C. Sielużycki et al. Can Kinect aid motor learning in sportsmen? A study for three standing techniques in judo. *PLOS ONE* 14(2): e0210260. 2019. <https://doi.org/10.1371/journal.pone.0210260>
- [43] P. Guo. "Effectiveness of VR Technology Based on Artificial Intelligence and Big Data in Overcoming Psychological Barriers in Taekwondo Practical Learning", *Security and Communication Networks*, vol. 2022. Article ID 3750213, 10 pages, 2022. <https://doi.org/10.1155/2022/3750213>
- [44] 根木 俊一, 島宗 理. 行動的コーチングによる合気道の技の改善. *行動分析学研究* 24 (1), pp.59-65, 2010.
- [45] A. Corbí and O. C. Santos. MyShikko: Modelling Knee Walking in Aikido Practice. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP '18)*. Association for Computing Machinery, New York, NY, USA, 217–218. 2018. <https://doi.org/10.1145/3213586.3225225>
- [46] Aikido training-techniques. https://play.google.com/store/apps/details?id=app.aikido_training.course&hl=ja&gl=US 令和 4 年 9 月 21 日閲覧.
- [47] K. Vasiliki and L. Margareta. Maintaining close contact from a distance: digital aikido training under Covid-19 conditions – reflections and lessons from an online martial arts course. *Martial Arts Studies* (11) , pp. 32-45. 2021. 10.18573/mas.128
- [48] M. A. Gordon. Awakening to wholeness: Aikido as embodied praxis of intersubjectivity. In O. Gunnlaugson, C. Scott, E. Sarath, & H. Bai (Eds.), *Catalyzing the field: Second person approaches to contemplative learning* (pp. 87–106). New York, NY: State University of New York Press. 2019.
- [49] D. Foster. Fighters who Don't Fight: The Case of Aikido and Somatic Metaphorism. *Qual Sociol* 38, 165–183. 2015. <https://doi.org/10.1007/s11133-015-9305-4>
- [50] M. A. Gordon. The Way of the Classroom: Aikido as Transformative and Embodied Pedagogy Through Self-Cultivation. In: *Aikido as Transformative and Embodied Pedagogy*. Palgrave Macmillan, Cham. 2019. <https://doi.org/10.1007/978-3-030-23953-4-5>
- [51] Y. Machi, C. Liu, K. Tohei, T. Ishizaki, T. Hamaoka, S. Kodato. The Physiological Study of Ki in Ki Aikido (1), *Journal of International Society of Life Information Science*. Volume 19. Issue 1. Pages 100-106. 2001.
- [52] Y. Machi, C. Liu, K. Tohei, T. Ishizaki, T. Hamaoka, S. Kodato. The Physiological Study of Ki in Aikido (2), *Journal of International Society of Life Information Science*. Volume 19. Issue 2. Pages 404-410. 2001.
- [53] 高須 公仁, 宮坂 裕之, 富田 豊, 本多 敏. 「折れない腕」における筋協調. *国際生命情報科学会誌*. 34 巻, 1 号, p. 7-14, 公開日 2016/08/01, Online ISSN 2424-0761, Print ISSN 1341-9226. 2016. <https://doi.org/10.18936/islis.34.1-7>, <https://www.jstage.jst.go.jp/article/islis/34/1/34-7/-article/-char/ja>
- [54] Microsoft. HoloLens2. <https://www.microsoft.com/ja-jp/hololens>
- [55] Unity. <https://unity.com/ja> 令和 4 年 9 月 25 日閲覧.
- [56] Microsoft Mixed Reality Toolkit 2. <https://learn.microsoft.com/en-us/windows/mixed-reality/mrktk-unity/mrktk2/?view=mrktkunity-2022-05> 令和 4 年 9 月 25 日閲覧.

未来ビジョン

今はYouTubeなどでどのようなことでも学習できる世界だ。スマートフォンやPC一つさえあればプログラミングから可愛い絵の描き方まで様々スキルを学ぶことができる。そのような時代において、動画では学べないことの価値は高い。本稿では、動画では学ぶことが難しい身体技能に関する研究を行なった。これは「スマートフォンにアプリケーションをインストールするかのよう、身体技能というソフトウェアを、人間の身体というハードウェアに手軽にインストール可能な世界を実現したい」というモチベーションにより行われた。誰もが簡単にコツや技能を自身の身体にインストールすることを可能にするMRソフトウェア

プラットフォームをGinoと名づけ、本ソフトウェアは中でも合気道の身体の使い方に注目したものであるという位置付けでGino.Aikiと名付けた。つまりGinoは、MRグラス一つあれば今まで動画などでは学習のしにくかった身体技術が簡単に学習できるような世界を目指すものであり、既存のスマートホンの立ち位置の延長という位置付けである。Gino.Aikiはその具体例であり、他にもGino.SkiやGino.Tennisなど他の身体技術を獲得するソフトウェアも作成することができるだろう。この研究を通してMRグラス一つあれば、合気道の身体の使い方からオリンピック選手のスキーの滑り方を学べる時代を目指している。本稿は、そのための第一歩となる研究である。

振りの理解を助けるためのダンス動画の自動分割

遠藤 輝貴* 土田 修平† 五十嵐 健夫*

概要. ダンサーがダンスの振りを覚える際、振りが短時間の動きに分割されていると理解や習得が容易になる。しかしダンス動画から振りを覚える場合はそのような振りの分割が予め存在しないため、学習者は振りを理解するために適切な分割位置を自分で見つける必要があり、これが振りの習得を困難にしていると考えられる。そこで我々はダンス動画の振りを個々の動きへと自動で分割する手法を提案する。提案手法では動画中のダンサーの身体や手の位置から視覚特徴量を、動画中の音楽から聴覚特徴量をそれぞれ計算し、これらの特徴量を Temporal Convolutional Network (TCN) に入力して、出力として得られる分割可能性のピークを検出することで動画の分割位置を求める。AIST Dance Video Database の動画から作成した学習データを用いた実験の結果、ダンス動画の分割に提案手法の視覚、聴覚特徴量の両方が役立つことを確認した。本論文では提案手法の詳細や実験結果と、自動分割の応用例として開発したダンス学習支援システムについて述べる。

1 はじめに

ダンサーがダンスの振りを習得する主な手段として、指導者から直接教わる方法と、ダンス動画の振りを真似する方法とがある。前者の場合は指導者が振りを短時間の動きに分割してそれぞれの動きを1つ1つ順番に教えていくことが一般的である。例えばダンスの「技」と呼ばれる一連の動きや、切れ目なく滑らかに繋がる動きなどは1つの塊として練習すると効率が良く、このように振りの構成要素を考慮して分割することで、学習者は振りの理解や習得が容易になる。これに対し後者の場合、通常のダンス動画ではそのような振りの分割が存在しないため、学習者は自ら踊ることのできるレベルまで振りを理解するために適切な分割位置を自分で見つける必要がある。ダンス経験の浅い人は上述の技や身体の流れに関する知識が少ないため、自分で振りを分割することが難しく、このことがダンス動画からの振りの習得を困難にしていると考えられる。

そこで我々はダンス動画の振りを時系列方向に、個々の動きへと自動で分割する手法を提案する。ダンスは音楽に合わせて身体を動かすため、振りの分割にはダンサーの動きだけではなく動画中の音楽も有用な情報になり得る。そこで提案手法ではまず、入力となるダンス動画から視覚的、聴覚的な特徴量を抽出する。視覚特徴量としては動画内のダンサーの身体や手のキーポイントの位置を推定し、その速度成分を使用する。また、聴覚特徴量としては動画に含まれる音楽からメルスペクトログラムを計算し、CNNで畳み込んだものを使用する。次にこれらの特徴量を Temporal Convolutional Network

(TCN) [1] の入力として与え、出力として得られる分割可能性のピークを検出することで動画の分割位置を求める。

提案モデルの学習には、ダンス動画に対してその動画の分割位置をアノテーションした学習データが必要となる。我々はAIST Dance Video Database [12] のダンス動画に対して振りの分割位置を手動でアノテーションし、合計1410本の学習データを用意した。この学習データを使い、視覚特徴量と聴覚特徴量の両方を使用した場合とどちらか一方のみ使用した場合とでモデルの性能を数値的に評価したところ、両方の特徴量を使用するモデルが最も性能が高くなった。この結果から、ダンス動画の分割には視覚的なダンサーの動きの情報と聴覚的な音楽情報の両方が重要であることが分かる。

提案手法の応用例として、自動分割の結果を利用したダンス学習支援システムを作成した。具体的には、推定した分割位置で区切った動画をループ再生することで、同じ動きの繰り返し練習を支援することができる。ループ再生するセグメント間には重なりを持たせることで、動きのつながりを理解しやすくしている。更にユーザが自動分割の細かさを調節できる機能も追加することで、ユーザの好みや熟練度に合わせた動画分割及び練習ができる。

2 関連研究

2.1 振りの分割の有効性

Rivièreら [9] は、ダンサーが動画から振りを覚える際の行動を調査し、ダンサーは動画の振りを分割して覚えていることを明らかにした。更に彼らは、手作業でダンス動画を分割するためのツールを開発し、ダンサーや指導者に対する被験者実験を行った。その結果、経験の浅いダンサーの場合は、自ら分割

* 東京大学

† 神戸大学

した動画よりも指導者が分割した動画のほうが振りを習得しやすいことを示した。この結果から我々は、ダンス動画を適切に分割することは振りの習得に有効であると仮定し、動画の分割作業を指導者の代わりにコンピュータで自動化することを目的とする。

2.2 ダンスモーション分割

Shiratoriら [10] は、日本舞踊のモーションデータと使用されている音楽情報を入力とし、ルールベースでモーション分割を行った。しかし、このルールは日本舞踊に特有の特徴を使用したものになっており、ヒップホップなどの一般的なダンスジャンルに適用することは難しい。Okadaら [6] は、CG キャラクタのダンスモーションを生成するために、既存のモーションデータをルールベースで分割する手法を考案した。しかし、この手法では音楽の拍の位置は既知でなければならず、拍の位置が分からない通常のダンス動画には適用できない。

ダンス動画の拍の位置を推定するために、Peder-soli と Goto [8] は、ダンス動画の映像情報を用いた手法を考案している。この研究は、動画から推定したダンサーのポーズ情報を TCN の入力に使っている点で我々の手法と類似しているものの、動画中の音楽情報を使用していない点や推定する対象が大きく異なる。

2.3 ダンス動画からのダンス学習支援システム

ダンス動画からの振りの習得を支援するアプリはいくつか存在する。ウゴトル [14] では、ダンス動画を閲覧しながら振りを覚えるための左右反転や速度変更の機能が充実しているが、動画を分割して一部だけ再生することができない。SymPlayer [11] は動画の左右反転や速度変更に加え、ループ再生機能が存在する。但し、ループ区間の始点と終点を手動で設定しなければならない上に、指定できる区間も1つのみなので、振りを複数の短い動きに分割して順番に練習するには不向きである。

また、ダンス動画に対して処理を行うことで振りの習得を支援する研究も存在する。例えば、斎藤ら [16] は、ダンス動画に対して動きのニュアンスを表すオノマトペを表示することで、学習者が動きのニュアンスを理解してダンスを学びやすくしている。Zhouら [15] は、複数人で踊っているシンクロダンスの動画に対して、動画に映る複数人の動きの類似度を推定し、身体部位別、時間別に動きのずれを表すヒートマップを表示することで、シンクロダンスの練習を支援するシステムを提案している。Tsuchidaら [13] は、深層学習技術を用いて学習者が手本となるダンスをマスターして踊っている動画を生成し、それを学習者自身に見せることでダンスの学習を支援することを提案している。

Rivièreら [9] は、手作業でダンス動画を分割す

るためのツール「MoveOn」を開発した。このツールでは動画を分割して複数のセグメントを作成し、各セグメントに対してループ再生や速度変更の設定ができる。しかし、ダンス動画の分割を手で行う必要があるため、適切な分割位置が分からないダンス初心者が使用するのは難しい。また、たとえ熟練者であってもセグメントを1つ1つ作成するのは手間がかかる作業である。

これに対し、我々が6章で提案するシステムは、自動分割した動画セグメントをユーザに表示するため、手動分割の手間が省けるだけでなく、初心者によるシステムの利用やダンスの練習支援ができると考えられる。

3 提案手法

3.1 アルゴリズムの概要

本研究の目的は通常のカメラで撮ったダンス動画を入力として、その動画の振りの分割位置を推定することである。提案手法の概要を図1に示す。はじめに、動画内のダンサーの動きから視覚特徴量を、使用されている音楽データから聴覚特徴量を、それぞれ後述の方法で計算する。次にこれらの特徴量を TCN の入力として使用する。TCN の出力は、動画の各フレーム t に対してそのフレームで動画が分割される可能性の高さを表す t の関数 $p(t)$ であり、この $p(t)$ のピークを検出することで、動画の分割位置を決定する。以降の節ではアルゴリズムの各段階の詳細を記す。

3.2 視覚特徴量の抽出

視覚特徴量は動画内のダンサーの動きから計算される。まず、動画内のダンサーの身体や手のキーポイントを AlphaPose [4] を用いて検出する。本研究で用いるキーポイントは身体が26個、左右の手がそれぞれ21個の合計68個であり、キーポイントの位置は動画の左上が $(0,0)$ 、右下が $(1,1)$ になるように正規化した座標値で求める。さらにダンサーの動きの速度に注目するため、1フレーム前からのキーポイント位置の変化を求め、これを視覚特徴量とする。即ち、フレーム t で検出した i 番目のキーポイントの位置を $\mathbf{k}_i(t) \in \mathbb{R}^2$ とすると、視覚特徴量 $\mathbf{v}(t) \in \mathbb{R}^{68 \times 2}$ の i 番目の要素 $v_i(t) \in \mathbb{R}^2$ は以下の式で求めることができる。

$$v_i(t) = \frac{1}{2}(\mathbf{k}_i(t) - \mathbf{k}_i(t-1)) \quad (1)$$

但し、最初の $\frac{1}{2}$ は特徴量の値を $[-0.5, 0.5]$ の範囲に収めるための定数である。

提案手法では1人のダンサーのみが映っているダンス動画を対象としているが、AlphaPoseの誤検出によって動画の1フレーム中に複数人間が検出されてしまうケースや、1人も検出されないケースも

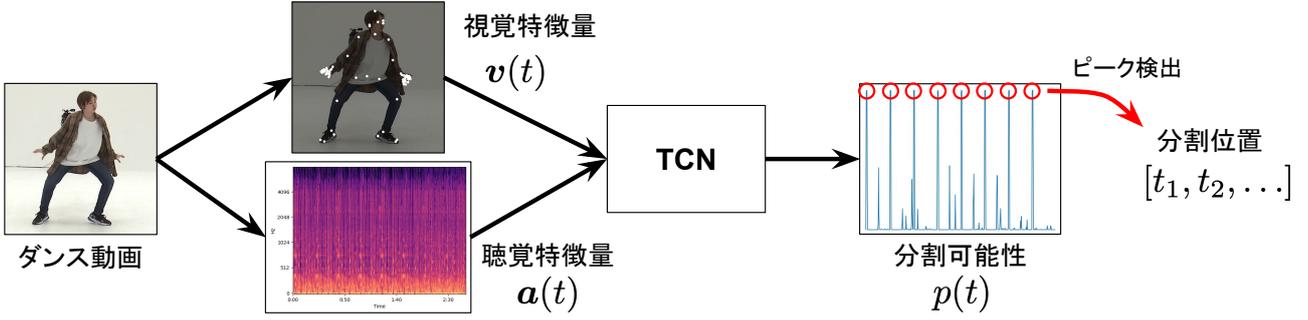


図 1. 提案手法のアルゴリズム概要. ダンス動画から抽出した視覚特徴量と聴覚特徴量を TCN の入力とし, 出力として得られた分割可能性のピークを検出することで動画の分割位置を決定する.

存在するため, その場合の対策も必要である. 前者の場合は AlphaPose がキーポイントの位置と合わせてそのキーポイントの検出の信頼度を出力することを利用し, 各フレームに対して全身のキーポイントの信頼度の総和が最大となる人間を選択する. 後者の場合は, 検出されなかったフレームでのキーポイントの位置は前後のフレームでのキーポイント位置から線形補間によって求める.

3.3 聴覚特徴量の抽出

聴覚特徴量は動画に含まれる音楽データから計算される. まず, 動画中の音楽を短時間フーリエ変換 (STFT) してメルスペクトログラム S を作成する. これは音の各時間での各周波数成分の強さを表す 2 次元配列であり, 各要素は dB 単位で求めた後に $[-0.5, 0.5]$ の範囲に正規化している. ここで得られたメルスペクトログラムのサンプル数が動画のフレーム数よりも多いため, CNN を用いて情報を圧縮し, フレーム t に対する聴覚特徴量 $a(t) \in \mathbb{R}^{16}$ を計算する. 具体的には, t に時間的に最も近いスペクトログラムのサンプルのインデックスを i とすると, $a(t)$ は

$$a(t) = \text{Conv2d}(S_{i-2, \dots, i+2}) \quad (2)$$

で求めることができる. ここで Conv2d は CNN による 2 次元の畳み込みを表す. メルスペクトログラムや Conv2d の計算におけるパラメータは [3] と同様に設定した.

3.4 Temporal Convolutional Network

以上で求めた視覚特徴量 $v(t) \in \mathbb{R}^{68 \times 2}$ と聴覚特徴量 $a(t) \in \mathbb{R}^{16}$ をまとめて 152 次元の特徴量を作成し, TCN への入力として使用する. TCN は時系列データに対して CNN を適用するモデルであり, 層が深くなるにつれて隣の要素との間隔を広げて畳み込むことで過去や未来の情報を考慮した計算ができる. TCN の構造は Davies と Böck [3] を参考に作成した. [3] では TCN を用いて音楽から拍の位置を推定しており, 元々の TCN のモデル [1] と比較す

ると, 過去だけでなく未来の情報も畳み込む点や活性化関数に exponential linear unit (ELU) [2] を使用している点が異なる. TCN による計算は各次元で独立に行い, 最後に全結合層で全ての次元をまとめ, 活性化関数としてシグモイド関数に通すことで出力 $p(t) \in [0, 1]$ を得る.

3.5 ピーク検出

TCN から出力された $p(t)$ は, 動画がフレーム t で分割される可能性であるため, 最終的な分割位置を決定するために $p(t)$ からピーク値を検出する必要がある. $p(t)$ がフレーム t^* でピーク値を取るとは, $p(t^*)$ が閾値を超える局所最大値であると言えるため, ピーク検出の条件は以下のように表すことができる.

$$p(t^*) > h \wedge p(t^*) = \max_{t-w/2 \leq s \leq t+w/2} p(s) \quad (3)$$

ここで w は局所最大値を計算する窓サイズ, h はピーク検出の閾値である. この条件を満たす t^* を $p(t)$ から全て求め, 分割位置 $[t_1, t_2, \dots]$ とする.

4 学習データの作成

4.1 使用したダンス動画

提案手法のネットワークモデルを学習させるためにはダンス動画と分割位置をペアとする学習データが必要となる. 本研究では, AIST Dance Video Database [12] のダンス動画に対して手動で分割位置を指定することで学習データを作成した. 使用した動画は基礎的な動きを行う 23 秒程度の基本ダンス 1200 本と様々な振りを含む 52 秒程度のフリーダンス 210 本の合計 1410 本であり, いずれも正面のカメラから撮影されたものである. なお, 分割位置の指定作業は約 10 年のダンス経験を持つ第 1 著者が手作業で行い, 所要時間は約 50 時間であった. 分割の際は全体的な基準として, 技と呼ばれる一連の動きや切れ目なく滑らかに繋がるような動きは 1 つの区間となるように, その動きの開始, 終了のタイミングで分割を行った.

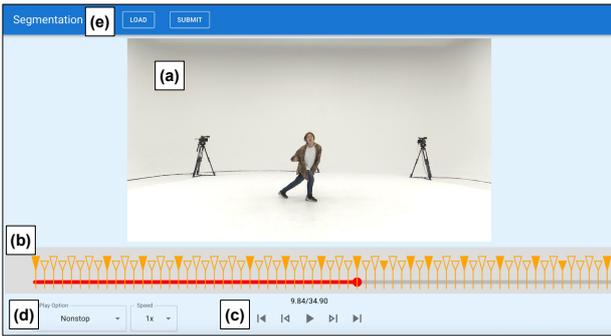


図 2. 学習データ作成ツール. (a) ダンス動画. (b) 分割候補点. (c) 再生, スキップボタン. (d) 再生モード, 速度変更用プルダウン. (e) 動画読み込み, 分割結果登録ボタン.

4.2 学習データ作成ツール

我々は作業効率化のために図 2 のようなツールを作成し, これを用いて動画の分割を行った. このツールでは中央にダンス動画を表示する画面 (図 2a), その下にシークバーと分割の候補点 (図 2b) が表示され, ユーザは各候補点をクリックすることで分割位置を指定する. 分割位置として指定された候補点はオレンジ色で塗りつぶされる. 下部中央 (図 2c) には動画の再生ボタンがあり, その両隣には現在の再生位置の 1 つ前または後ろの分割候補点まで再生位置をスキップするボタンがある. また通常の動画再生だけでなく, 指定した分割位置で動画を一時停止する機能を追加している. 再生モードは左下のプルダウン (図 2d) から変更できる. 動画の読み込みや分割結果の登録は左上のボタン (図 2e) から行う.

4.3 分割候補点の絞り込み

分割の候補点は少なすぎるとユーザの望み通りに分割できない一方で, 多すぎると作業効率が悪くなるため, 適切な数を設定することが望ましい. そこで我々は, 動画中の音楽の拍の位置とその半分的位置を分割の候補点として設定した. AIST Dance Video Database で使われている音楽はテンポが既知であり, 動画に含まれる音声はノイズのない純粋な音楽データであるため, 最初に音が大きくなった位置を音楽の開始位置としてテンポに従って区切ることで分割候補点の位置を容易に計算することができる. ツール上では図 2b のように, 拍の位置を大きい逆三角で, その半分の位置を小さい逆三角で示している.

5 評価実験

5.1 実験設定

第 4 章で作成した学習データを用いて提案手法のモデルを学習し, その性能を評価した. 学習データ

表 1. テストデータに対する分割位置の推定結果.

モデル	訓練終了 エポック	適合率	再現率	f 値
視覚	17.4	0.383	0.407	0.395
聴覚	29.2	0.594	0.625	0.609
視覚+聴覚	18.6	0.696	0.682	0.689

はそれぞれに含まれる基本ダンスとフリーダンスの割合が等しくなるよう注意しながら 3:1:1 の比でランダムに分割し, それぞれ訓練用, 検証用, テスト用データとした.

損失 \mathcal{L} は重み付きの binary cross-entropy を採用し, モデルの出力する分割可能性 $p(t)$ と正解ラベル $l(t)$ から以下の式で計算する.

$$\mathcal{L} = -\frac{1}{T} \sum_{t=1}^T \{ \alpha l(t) \log p(t) + (1 - l(t)) \log (1 - p(t)) \} \quad (4)$$

ここで T は動画の総フレーム数である. 正解ラベル $l(t)$ は第 4 章での分割結果から計算され, 指定された分割位置に最も近いフレームで 1, その両隣で 0.5, それ以外で 0 を取る. α は重みを表す定数であり, ラベルの要素がほとんど 0 であることを考慮して, 学習が偏らないように設定している. 本実験では $\alpha = 100.0$ とした.

以上の設定で, 訓練データを用いてモデルの学習を行った. 訓練時の最適化手法は Adam [5], バッチサイズは 1, 学習率は 0.001 とし, 検証データによる平均損失が直近 10 エポックで改善されなければ学習をストップし, その時点でのモデルをテストデータで評価した. また比較手法のため, 提案手法 (視覚特徴量と聴覚特徴量を両方使うモデル) に加えて, 視覚特徴量のみ, 聴覚特徴量のみを用いたモデルを用意した. 但し, 視覚のみ, 聴覚のみのモデルで実験するときは, 提案モデルの実験で用いた訓練, 検証, テストデータと同じものを使用した. モデルは PyTorch [7] を用いて実装し, 学習は Google Colaboratory 上で GPU を使用して実行した.

5.2 結果と考察

5.1 節の手順で実験を行い, モデルの訓練にかかったエポック数と, 訓練後のモデルを用いてテストデータから検出した分割位置の適合率と再現率を記録した. この実験を独立に 10 回繰り返し, 訓練エポック数と適合率, 再現率の平均値を求め, 平均適合率と平均再現率から f 値を求めた結果を表 1 に示す. ピーク検出時の窓サイズ w は 20 とした. また閾値 h は 0.90 から 0.99 まで 0.01 ずつ変更しながら数値評価を行ったところ, 全てのモデルで $h = 0.98$ の時に f 値が最大となったため, 0.98 を採用した.

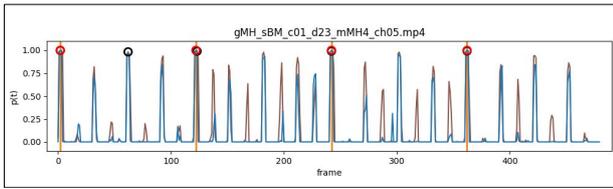


図 3. あるテストデータに対する推定結果の比較. タイトルは動画ファイル名, オレンジの縦線は正解の分割位置を示す. 視覚+聴覚のモデルで推定した $p(t)$ と分割位置はそれぞれ青線と赤丸, 聴覚のみのモデルの推定結果は茶色線と黒丸で示す.

表 1 から, 提案手法 (視覚特徴量と聴覚特徴量を両方使うモデル) が適合率, 再現率, f 値が最大となった. 一方, 視覚特徴量だけのモデルでは各指標の値が大きく低下した. 聴覚特徴量だけのモデルは提案モデルに近い性能を出しているものの, 訓練終了までのエポック数が大きくなる傾向が見られた. この結果から, ダンス動画の分割には視覚的なダンサーの動きの情報と聴覚的な音楽情報の両方を用いるのが有益であると言える.

また, 提案モデルと聴覚特徴量のみを用いたモデルとで, あるテストデータに対する推定結果を比較したものを図 3 に示す. 図 3 を見ると提案モデルで推定した分割可能性は聴覚のみのモデルと比べて, 正解位置以外での値が小さい. 他のテストデータでも同様の傾向が見られたことから, 視覚情報は聴覚情報のみでは判断できない不要な候補を削除する役割を果たしていると考えられる.

さらに提案手法において, テストデータの動画の中で f 値が最大, 最小の動画の分割位置を表すグラフを図 4, 5 に示す. 図 4a のダンス¹は同じ動きの繰り返しになっており, 音楽の拍位置に動きのアクセントが来るような振りのため, モデルによる予測が容易であったと考えられる. その一方, 図 5b のダンス²は, はっきりと音楽に合わせるのではなく柔らかく流れるような動きが多いため, 分割位置の予測が難しかったと考えられる.

6 応用例

提案手法の応用例として, 我々は自動分割の結果を利用してダンスの理解や習得を支援するシステムのプロトタイプを開発した. プロトタイプのインタフェースを図 6 に示す. 読み込んだダンス動画を表示する画面 (図 6a) やシークバー (図 6b), 再生, スキップボタン (図 6c) は, 学習データ作成時のツール (図 2) と同様の設計だが, シークバー上に表示

¹ https://aistdancedb.ongaacce1.jp/v1.0.0/video/10M/gLO_sBM_c01_d14_mL05_ch08.mp4

² https://aistdancedb.ongaacce1.jp/v1.0.0/video/10M/gJB_sFM_c01_d07_mJB2_ch03.mp4

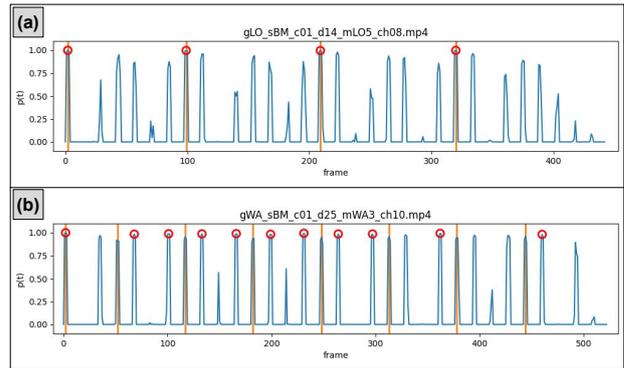


図 4. テストデータ (基本ダンス) に対する分割位置の推定結果 ((a) f 値が最大, (b) f 値が最小のもの). グラフの見方は図 3 と同様.

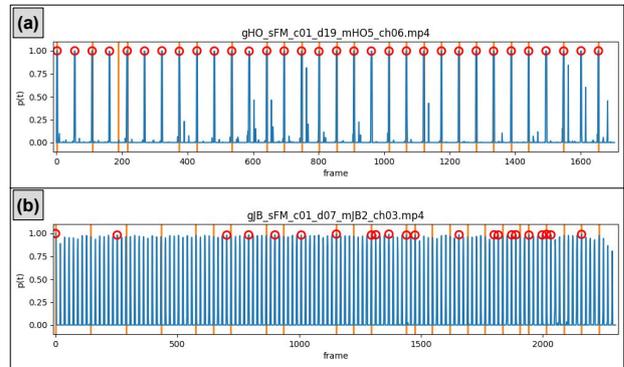


図 5. テストデータ (フリーダンス) に対する分割位置の推定結果 ((a) f 値が最大, (b) f 値が最小のもの). グラフの見方は図 3 と同様.

される分割位置は自動推定した分割位置となっている. 再生モードは 4.2 節で述べた 2 種類に加えてループ再生モード (図 6d) を追加した. ループ再生モードでは, 再生開始時の再生位置の前後にある分割位置の間をループ再生することで, ユーザが同じ動きを繰り返し練習することを支援することができる. ループ再生の範囲はシークバー上に鍵括弧 (図 6e) で表示する. 但し, ループ再生の範囲は, 分割位置ちょうどではなく, 少し前後にはみ出るように設定している. この理由として, 分割したセグメント同士に少し重なりを持たせることで, ユーザがセグメントの境界での動きのつながりや予備動作を理解しやすくなることが挙げられる. 現在の実装でははみ出す部分の長さは定数 (0.2 秒) としている. また, 学習に最適な分割の細かさはユーザの好みや熟練度によって変わりうるため, ユーザが分割の細かさを調節できるように, 右下部にスライダ (図 6f) を用意した. このスライダを操作することで, 3.5 節のピーク検出時の閾値 h の値を変更し, 分割位置の数を調整することができる. 動画の読み込みは左上の

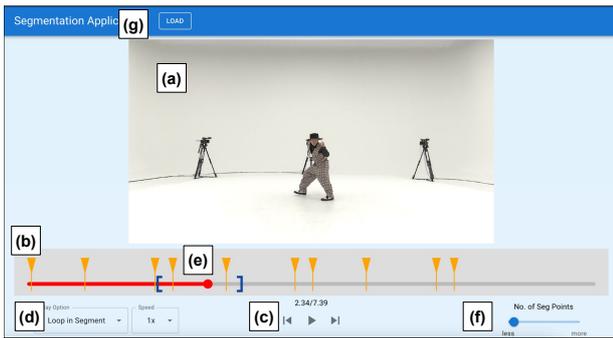


図 6. 自動分割の結果を利用したダンス学習支援システム. (a) ダンス動画. (b) シークバー. (c) 再生, スキップボタン. (d) 再生モード, 速度変更用プルダウン. (e) ループ再生の範囲を表す鍵括弧. (f) 分割の細かさ調節スライダ. (g) 動画の読み込みボタン.

ボタン (図 6g) で行う.

本システムは学習者が個人で使用し, 自力で振りを覚えるための分割動画を作成することを想定している. 振りの自動分割機能は主に初心者のダンス習得に役立つと考えられるが, 自力で振りを分割できる熟練者が動画の振りを練習する場合でも, 動画を手で分割する手間が省けるという点で役に立つと考えられる. 現在はプロトタイプ実装のため, 事前に提案モデルで分割可能性を推定した動画しか読み込めないが, 将来的にはユーザがアップロードする任意の動画に対しても自動で分割位置を推定できるように改善する予定である.

7 まとめと今後の課題

本研究では, ダンス動画の振りを個々の短い動きへと自動で分割する手法を提案した. 提案手法は, 動画内から推定したダンサーのキーポイントの速度情報により視覚特徴量を求め, 動画の音楽データのメルスペクトログラムを CNN で畳み込むことで聴覚特徴量を取得する. これらの特徴量を TCN に入力して分割可能性を計算し, その中のピークを検出することで分割位置を計算した. 我々は, 学習データ作成のためのシンプルなアノテーションツールを作成し, AIST Dance Video Database の動画に対して人手で分割位置を指定することで学習データを用意した. このデータを使用した実験の結果, 提案手法 (視覚と聴覚両方の特徴量を組み合わせたモデル) は視覚のみ, 聴覚のみの特徴量を使用するモデルよりも適合率, 再現率, f 値の項目で優れた結果が得られた. この結果から, ダンス動画の分割には視覚的なダンサーの動きの情報と聴覚的な音楽情報の両方が役に立つことが分かる.

今後の課題として, まず人による分割位置の差異

への適応が挙げられる. 現在の学習データは第 1 著者のみで作成しているが, 振りの分割位置は絶対的なものではなく人によってある程度の差異がある. この差異に対応するためには, まず複数のダンス経験者に動画の分割を行ってもらって学習データを集め, 大多数の人間の好みにある程度沿う「平均的な」正解ラベルを作成する. さらにこの平均的な正解ラベルを出発点として, [17] のように Human-in-the-Loop の手法で個々の好みに沿った分割を生成することが考えられる.

また, 現在は学習データの都合によりストリートダンスやジャズダンスを対象としているが, 提案手法はダンスの種類やジャンルに特有の知識を必要としないため, 民族舞踊やコンテンポラリーダンスなど他のダンスにも適用できる可能性がある. これらのダンスについて学習データを集め, 提案手法の有効性を検証することも今後の発展として考えられる. 性能評価については, 提案モデルによる分割の結果をダンス初心者や熟練者が手動分割した結果と比較することも考えられる. 更に, 提案手法の応用例として開発したダンス学習支援システムのプロトタイプに関するユーザテストを実施し, 自動分割やプロトタイプシステムが実際にダンスの理解や習得に役立つかどうかを検証していく予定である.

謝辞

本研究は JST CREST JPMJCR17A1 の支援を受けたものである.

参考文献

- [1] S. Bai, J. Z. Kolter, and V. Koltun. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv:1803.01271*, 2018.
- [2] D. Clevert, T. Unterthiner, and S. Hochreiter. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). In Y. Bengio and Y. LeCun eds., *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, 2016.
- [3] M. E. P. Davies and S. Böck. Temporal convolutional networks for musical audio beat tracking. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pp. 1–5, 2019.
- [4] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu. RMPE: Regional Multi-person Pose Estimation. In *ICCV*, 2017.
- [5] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. In Y. Bengio and Y. LeCun eds., *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

- [6] N. Okada., N. Iwamoto., T. Fukusato., and S. Morishima. Dance Motion Segmentation Method based on Choreographic Primitives. In *Proceedings of the 10th International Conference on Computer Graphics Theory and Applications - GRAPP, (VISIGRAPP 2015)*, pp. 332–339. INSTICC, SciTePress, 2015.
- [7] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett eds., *Advances in Neural Information Processing Systems*, Vol. 32. Curran Associates, Inc., 2019.
- [8] F. Pedersoli and M. Goto. Dance beat tracking from visual information alone. In *Proceedings of the 21st International Society for Music Information Retrieval Conference, (ISMIR 2020)*, pp. 400–408, Montreal, Canada, Oct. 2020.
- [9] J.-P. Rivière, S. F. Alaoui, B. Caramiaux, and W. E. Mackay. Capturing Movement Decomposition to Support Learning and Teaching in Contemporary Dance. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), Nov. 2019.
- [10] T. Shiratori, A. Nakazawa, and K. Ikeuchi. Detecting dance motion structure through music analysis. In *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.*, pp. 857–862, 2004.
- [11] M. Tachikawa. SymPlayer -動画ミラー反転でプロの動きをマスター- (最終閲覧日: 2022年9月24日). <https://apps.apple.com/jp/app/id1048785434>.
- [12] S. Tsuchida, S. Fukayama, M. Hamasaki, and M. Goto. AIST Dance Video Database: Multi-genre, Multi-dancer, and Multi-camera Database for Dance Information Processing. In *Proceedings of the 20th International Society for Music Information Retrieval Conference, (ISMIR 2019)*, pp. 501–510, Delft, Netherlands, Nov. 2019.
- [13] S. Tsuchida, H. Mao, H. Okamoto, Y. Suzuki, R. Kanada, T. Hori, T. Terada, and M. Tsukamoto. Dance Practice System That Shows What You Would Look Like If You Could Master the Dance. In *Proceedings of the 8th International Conference on Movement and Computing, MOCO '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- [14] Ugotoru Inc. ウゴトル (最終閲覧日: 2022年9月24日). <https://ugotoru.com/ugotoru>.
- [15] Z. Zhou, A. Xu, and K. Yatani. SyncUp: Vision-Based Practice Support for Synchronized Dancing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 5(3), Sept. 2021.
- [16] 斎藤 光, 徳久 弘樹, 中村 聡史, 小松 考徳. ダンス動画へのオノマトペ付与によるダンス習得促進手法. 情報処理学会 研究会報告グループウェアとネットワークサービス (GN), 2020-GN-109(33):1–8, Jan. 2020.
- [17] 山本 和彦. Human-in-the-Loop 型適応によるインタラクティブな音楽的拍節解析. 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), pp. 23–29, Dec. 2021.

リアルタイムフィードバックとスコアリングを導入した顕微鏡縫合術訓練

田代 優香* 宮藤 詩緒* Hwang Dong-Hyun* 清藤 哲史†
 金 太一† 五十嵐 健夫‡ 小池 英樹*

概要. 脳神経外科の顕微鏡縫合術は習得に時間を必要とする難易度の高い医療技術である。熟練の医師は多忙なため初学者の訓練に長時間付き合うことができず、初学者は単調な作業を単独で訓練する必要がある。そこで本研究は訓練のモチベーション向上のためにスコアリングやリアルタイムフィードバックの表示といったゲーミフィケーションの要素を取り入れたシステムを提案した。本研究では顕微鏡付属のカメラから取得した動画から手術に必要な技術的要素を検出し、その要素を用いてスコアを算出する。脳神経外科医によると縫合に必要な要素は速さ、正確さ、丁寧さの3つである。これらの技術的要素を深層学習や画像処理を用いて、器具やガーゼを追跡することによって検出する。本システムを用いて10人の顕微鏡縫合術初心者を実験を行った。実験の結果、UEQ(User Experience Questionnaire)とSUS(System Usability Scale)から本システムはモチベーション向上に寄与することがわかった。

1 はじめに

脳神経外科ではバイパス手術やクリッピング術などの高度な顕微鏡縫合術が必要とされる [28]。顕微鏡縫合術とは、顕微鏡を覗きながら1ミリ前後の組織を縫合する手術である。この顕微鏡縫合術は、縫合対象となる血管の周囲を傷つけないように [7,15]、また、術後の影響を最小限にするために20分から30分といった短時間で縫合を終わらせるように [14,15] 注意を払う必要のある高度な技術である。そのため脳神経外科では現在においても手術ロボットではなく直接術具を操作する手術が主流である。

これまで脳神経外科では、実際の手術を見て学んだり手術経験を積んだりすることによって顕微鏡縫合術を習得していた。しかし、近年はカテーテルを使用する血管内手術や放射線治療などの代替治療や患者の権利保護のため、若手脳神経外科医が実際の手術から学ぶ機会が減っている。そのため、少ない手術経験で難易度の高い手術を成功させなくてはならない。この事態に対処するため、普段の練習では、実際の手術で用いるフットペダル式の顕微鏡や卓上顕微鏡を用意し、ラットや手羽先、シリコンチューブ、ガーゼなどを縫合対象とする縫合訓練が行われている [25,27]。なかでもガーゼ縫合は、忙しい労働時間の間でも、短い時間で簡単に行える練習方法として、脳神経外科医の中でもよく使われる練習である。高度な顕微鏡縫合術を習得するためには、これらの縫合を約1万針分、数百時間練習する必要があると言われている [13,14]。しかし、人手不足の医療現場では、熟練者が常に練習に立ち会うのは困難

であるため、練習者はほとんどの時間を個人で練習する必要がある。特に初学者にとっては、自分の成長を測る指標が少ないため、単調な練習を長期間続けるためのモチベーションの維持がより難しくなる。

この課題を解決するために、図1に示す顕微鏡縫合術訓練のためのリアルタイムフィードバックシステムを提案する。本システムでは深層学習による器具検出や画像処理を用いてユーザの縫合をリアルタイムに分析し、ヘッドマウントディスプレイ (HMD) を用いて重畳表示によるフィードバックを行う。さらに、ユーザの縫合を採点し、スコアを提供する。これらのリアルタイムフィードバックやスコアリングといったゲーミフィケーションの技術を顕微鏡縫合に統合することで、ユーザエクスペリエンス (UX) を向上し、単調な訓練に対するモチベーション維持、効率化を図る。

本論文の貢献は以下の通りである。

- 熟練者の着目点である、「速さ」「正確さ」「丁寧さ」を明確化し、各要素を動画内から検出することを可能とした。
- 可視化を用いたリアルタイムフィードバックやスコアリングといったゲーミフィケーションの要素を取り入れることで、練習者のモチベーションの維持を試みた。
- ユーザ実験により、本システムがモチベーション向上に貢献することを示した。

2 関連研究

2.1 縫合器具や技能の測り方

外科医の技術評価は主観的な評価と臨床結果を用いて行われてきたため、客観性に欠ける問題が挙げられる。これに対し、Doughtyら [4] や Wangら [26]、Egiら [5] はロボットのセンサや、動画から機械学習

* 東京工業大学 情報理工学院

† 東京大学 医学部脳神経外科

‡ 東京大学 大学院情報理工学系研究科

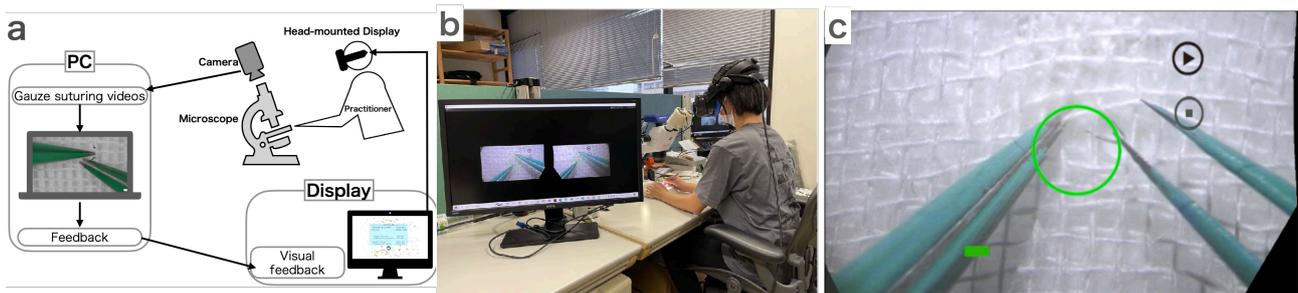


図 1. (a): システム概要. (b) システムを使用している様子. (c) ユーザーが装着したヘッドマウントディスプレイに顕微鏡画像に重畳されたリアルタイムフィードバックが表示される.

を用いて評価するシステムを開発した. 彼らのシステムでは手術技術の客観的評価を可能にした. しかし, 我々がターゲットとしている脳神経外科のガーゼを用いた顕微鏡縫合術訓練である. そのため, ロボット手術ではなく手による手術を行うためセンサーデータを取得できず, 訓練動画のデータも少ない.

脳神経外科でも器具の検出と評価をしている例として, Harada ら [8] は実際の医療器具にマーカーなどを装着し縫合技能を客観的かつ定量的に測り, 評価する方法を示した. しかし, このシステムでは実際の器具にマーカーをつけたり, トラッキングシステムを構築する手間と時間がかかるため, 多くの医院で導入することは難しい.

動画内から手術技能の一部を検出する例として, Sugiyama ら [23] は外科手術における体の内部組織に対する「優しく」扱うこと定量的に測定する方法を提案した. 本研究では顕微鏡縫合術に必要な技術的要素としてガーゼの扱いの丁寧さを検出する.

2.2 ゲーミフィケーション

ゲーミフィケーションとは, ゲームの要素を非ゲームに取り入れていることをいう [11]. モチベーションや学習効率を上げるために, Siemon ら [22] は大学教育のグループ活動におけるゲーミフィケーションを導入したシステムを提案し, Huang ら [11] や Park ら [20] は個人の学習として修士学生の授業に取り入れた.

医療分野では, Süncksen ら [21] がゲーミフィケーションを X 線画像処理トレーニングに取り入れた VR システムを開発した. 彼らのシステムでは VR のコントローラーを用いて仮想空間上の器具を操作できる. しかし, 彼らのシステムでは, VR 内の描画と実際の器具に差が生じる, ドメインギャップの問題が発生する. そこで本システムでは実際の縫合器具を用いることによってドメインギャップを解消し, 長時間行う単調な縫合訓練に対するモチベーションの維持のためにゲーミフィケーションを用いる.

2.3 医療における映像の表示方法

医療において臨床現場や訓練における表示方法として, Ho ら [9] は手術機器から離れていても双

眼顕微鏡の映像が立体的に見えるシステム, Chang ら [3] は球体映像ベース仮想現実を用いた. 本研究では HMD を用いて顕微鏡の接眼カメラの画像とフィードバックを重畳表示する.

3 提案手法

本研究ではゲーミフィケーションの要素を導入したリアルタイムフィードバックを行う顕微鏡縫合術訓練システムを提案する.

提案システムでは, ユーザーが図 1 (b) のように HMD を装着し, 顕微鏡と縫合器具を用いて顕微鏡縫合術訓練を行う. 本システムでは, 顕微鏡縫合映像を撮影できる顕微鏡カメラを搭載し, カメラの映像から縫合器具や術部の状態をリアルタイムで検出する. 検出された器具や術部の情報は, ユーザーが装着した HMD 上に図 1 (c) のようにリアルタイムで可視化される. また, 縫合終了時に, 検出された器具や術部の情報からスコアリングを行うことで, 擬似的な客観的評価を行う. 本研究では, これらの情報の可視化によるリアルタイムフィードバックや, スコアリングといったゲーミフィケーションの要素を縫合訓練に取り入れることで, ユーザーのモチベーション向上 [11, 22] を図る.

本システムを実現するためには, 顕微鏡画像上から「技術的要素の検出」を行い, それらを HMD 上に「リアルタイムフィードバック」, さらに「スコアリング」を行う必要がある. 以下の章では, 各要素に関して詳しく説明を行う.

3.1 技術的要素の検出

これまでの顕微鏡縫合術訓練を単独で行う場合, ユーザーは自身の上達を主観的に判断することが困難であった. 先行研究での脳神経外科医への聞き取りの結果, 顕微鏡縫合術に必要な要素は「速さ」, 「正確さ」, 「丁寧さ」の 3 点であるとわかった [24]. 本研究では, 顕微鏡の映像からこれらの要素を自動的に検出し, 客観的な評価を行うシステムを実現した. 縫合に必要な要素のうち「速さ」については Matsuura ら [19] で用いられている物体検出ネットワークとフェーズ推定により検出する. 本システムでは「正確さ」と「丁寧さ」は深層学習を用いた器具追跡と

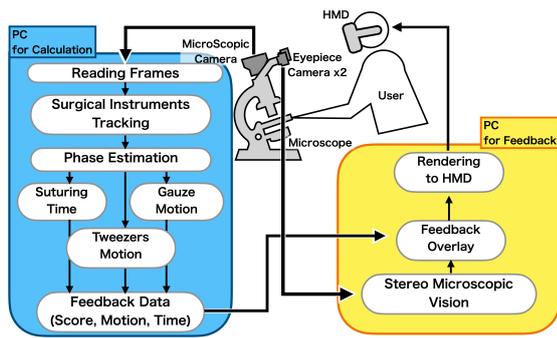


図 2. システム概要とハードウェアのセットアップ.

縫合動画内のオプティカルフローの値から検出する.

3.2 リアルタイムフィードバック

検出された技術的要素を分析し、ユーザにリアルタイムでフィードバックするには、リアルタイム解析と重畳表示技術が必要である。本システムでは、顕微鏡カメラからの映像を 30fps でリアルタイム解析すると同時に、顕微鏡の接眼部に設置された別カメラをから得られた映像を別 PC でプロセッシングすることで、リアルタイム性を保持したまま、図 1 (c) のように HMD での重畳表示を行う。この方法では顕微鏡からの映像をそのまま表示するため、マーカーが不要になり、ドメインギャップが解消される。

3.3 スコアリング

客観的な評価として、本研究では顕微鏡映像を用いて自動採点を行う。初級、中級、上級の顕微鏡縫合動画を複数用意し、各動画から技術的要素を抽出する。各レベルの動画に 30 点、60 点、90 点を付与し、技術的要素を引数として回帰分析を行い、自動採点を可能にする。

4 実装

本論文では顕微鏡画像からリアルタイムにスコアを算出し、フィードバックするシステムを実装した。フィードバックは HMD を用いてユーザにリアルタイムで表示される。

本研究ではまず技術的要素の検出方法の策定を行う。次にゲーミフィケーションの要素であるスコアの算出方法を実装する。最後に HMD を用いて、縫合動画とフィードバックの重畳表示を行う。

4.1 システムの概要

重畳表示をリアルタイムで行うために、システムは 2 台の PC、3 台のカメラ、HMD で構成する。図 2 に示すように、本システムは顕微鏡カメラから器具位置推定、フェーズ推定、スコアの計算を行う計算用 PC と HMD の画像処理を行うフィードバック用 PC の 2 台を使用する。単眼顕微鏡カメラを計算用 PC に接続し、1080p、60fps で PC に送信した。顕

微鏡の各接眼部に接眼カメラを 2 台設置し、1080p、60fps で立体視ができるようにした。

図 2 はデータの流れの概要を示している。顕微鏡カメラからの画像は、器具検出、ガゼの動きの検出、フェーズ推定のために計算用 PC にストリーミングされた。この PC はこれらの各値を組み合わせてスコアなどのフィードバックデータを生成し、UDP 経由でフィードバック用 PC に送信する。フィードバック用 PC は、接眼カメラからの顕微鏡画像を HMD の左右のディスプレイそれぞれ描画する。フィードバック用 PC から取得したデータをもとに、顕微鏡画像に重ねて視覚的フィードバックを生成する。これらの画像はリアルタイムで HMD に送信される。

4.2 技術的要素の検出

脳神経外科医の意見から熟練の脳神経外科医が重視するポイントは、「速さ」、「正確さ」、「丁寧さ」であることがわかった。顕微鏡の画像からこれらを検出するためには、以下の技術的要素が必要である。

- **速さ**: 1 縫合にかかる時間。
- **正確さ**: 縫合器具の先端の動く範囲。
- **丁寧さ**: ガゼ全体の動き。

縫合動画に対し深層学習を用いた物体検出ネットワーク [19] を使用することで器具位置の推定を行い、縫合フェーズの推定を行うことが可能である。本研究では、同様の手法で縫合時間や器具の動きの測定を行う。正確さの縫合器具の先端の位置は推定された器具位置のバウンディングボックスの角の座標とする。図 3 (a) の四角形は器具位置の例である。また、ガゼ全体の動きの検出には推定された器具位置にマスクをし、その後オプティカルフロー (OF) を用いて推定を行う。推定された器具位置にマスクをかける理由は、ガゼ全体の動きの検出時に器具の動きを無視したいからである。図 3 (b), (c) は OF によるガゼ全体の動きの検出の例である。

初級、中級、上級、各 10 縫合分の縫合動画に対し、左器具の動きとガゼ全体の動きを検出した結果を述べる。これらの動画は脳神経外科の医師 2 名によって初級、中級、上級のラベル付けされた。左器具の動きについては、クラスカルウォリス検定では技術レベルによる効果が有意 ($\chi^2 = 10.9, p = 0.004 < 0.01$) であり、ウィルコクソンの順位和検定に Bonferroni 補正を加えた事後検定では初級と上級との間で有意な差 ($p = 0.0028 < 0.01$) が確認された。このことから上級者は初級者より動画内の中心に近い位置で左器具を扱うことがわかった。また、ガゼ全体の動きについては、クラスカルウォリス検定では技術レベルによる効果が有意 ($\chi^2 = 72.2, p = 3.4e-16 < 0.01$) であり、ウィルコクソンの順位和検定に Bonferroni 補正を加えた事後検定では初級と中級 ($p = 9.4e-09 < 0.01$)、中級と上

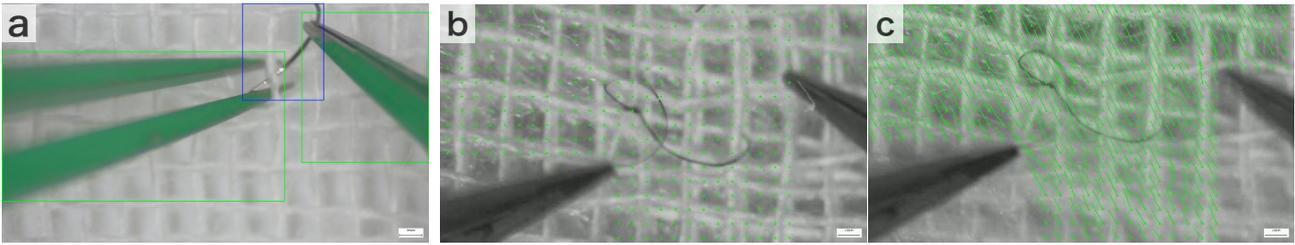


図 3. (a): 器具検出の例. (b), (c): ガーゼ全体の動きの例. (c) の緑線: (b) と (c) の間の OF のベクトル.

表 1. テストデータで学習させた場合のそれぞれの回帰の MAE の大きさ.

Linear	Huber	Ridge	Poisson
8.96	7.88	11.73	10.07

級 ($p = 1.7e-14 < 0.01$) それぞれで有意な差が確認された. これらのことから, 上級者になるほど縫合中のガーゼの動きが少なくなることがわかった.

4.3 スコアリング

前節で検出した技術的要素から特徴選択と重回帰分析を用いて擬似スコアリングを算出する. 特徴選択によってスコアに必要とされる技術的要素を選択し, 重回帰分析によってそれぞれの技術的要素に掛ける係数を決める. 特徴選択に用いたデータは速さの指標として 1 縫合にかかった時間, 正確さの指標として左右のピンセットの先端の距離, 左右それぞれのピンセットの先端と結び目の距離, 丁寧さの指標として OF の値を用いた. 正確さと丁寧さのそれぞれの指標は平均と標準偏差を用いており, 合計で 9 個である. 用いるデータは脳神経外科医にいただいた縫合動画を使用した. 縫合動画は初級, 中級, 上級それぞれ 2 動画ずつ, 合計 6 動画を使用した. 重回帰分析は Linear 回帰 [6], Huber 回帰 [12], Ridge 回帰 [10], Poisson 回帰 [18] の 4 つを使用した. 擬似スコアリングは初級, 中級, 上級のスコアを 30, 60, 90 として重回帰分析によって算出する. これらの 9 個の特徴量, 4 つの回帰手法から最もスコアの平均絶対誤差 (MAE) が小さくなる特徴量とその個数と回帰手法を選択する.

それぞれの回帰手法で最も小さい MAE の時の特徴量を用いてテストデータで擬似スコアを算出した結果を表 1 に示す. 最も MAE が小さいのは Huber 回帰であり, その時の特徴量は時間, OF の平均, 左ピンセットの先端と結び目の距離の標準偏差の 3 つである. これらはそれぞれ, 時間は「速さ」, 左ピンセットの先端と結び目の距離は「正確さ」, OF は「丁寧さ」を代表したパラメータが選択されており, 本研究で目的とした 3 要素が選択されている. よって, 本研究ではこの回帰手法と特徴量を擬似スコアリングに使用する. 決定した回帰手法と特徴量を用いて算出したスコアを図 4 (a) に示すように 1 縫合

後に表示した.

4.4 視覚的フィードバック

ユーザの縫合技術の改善を支援するために, 図 4 に示すように検出した技術的要素に基づいて縫合中動画に視覚的フィードバックを提示した. まず速さに対する表示は, 図 4 (a) のように, システムは 1 縫合終了後の画面に現在の所要時間が前回と最速の所要時間と共に表示する. システムはスコアも同様に図 4 (a) のように表示する. これによりユーザ自身で過去の記録と比較して振り返ることができ, モチベーション向上に繋がることを期待する. 次に正確さに対する表示は, 図 4 (b), (c) に示すように, 器具の先端の位置によって色の変わる円を表示する. これによりユーザが結び目の位置から離れずに器具を操作することを促す. 最後に丁寧さに対する表示は, 図 4 (d), (e) のように, ガーゼ全体が大きく動くほど赤色に近い色で長くなるバーを画面左下に表示する. これにより器具で糸やガーゼを無理に引っ張ってしまったことが視覚的にわかる.

5 被験者実験

本実験はゲーミフィケーションの要素を取り入れた本システムの有用性を評価するために提案システムによってどれだけモチベーションや UX が向上するかを検証する実験を行った. 提案システムを表示する条件の表示しない条件の 2 条件での被験者実験を行った.

5.1 仮説

本実験を通して以下の 3 つの仮説を検証する.

- H1: ゲーミフィケーション要素によってモチベーションが向上する.
- H2: ゲーミフィケーション要素によって練習時間が増える.
- H3: ゲーミフィケーション要素によって UX が向上する.

5.2 実験内容

本システムの対象は縫合初心者である. 縫合初心者と医学的背景を持たない学生では長期間の縫合練習に対するモチベーション維持の難しさは同様と考えられるため, 本実験では, 医学的背景を持たない

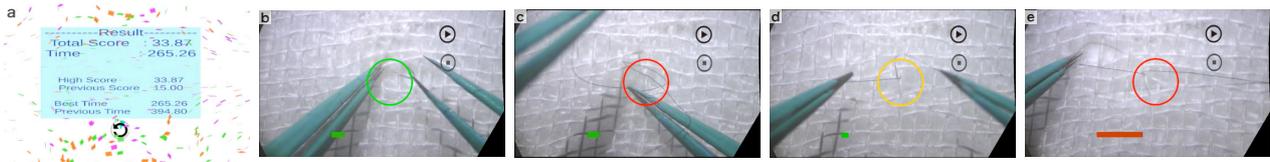


図 4. ヘッドマウントディスプレイで見た場合の UI 動作の例. : (a) 時間とスコアが表示された結果画面. (b) ピンセットが範囲内にある場合, 円は緑色に変化する. (c) ピンセットが範囲外にある場合, 円は赤色に変化する. (d) ガーゼ全体が大きく動いていないときは, 左下のバーが短くなり, 緑色に変化する. (e) ガーゼ全体が大きく動いているときは, 左下のバーが長くなり, 赤色に変化する.

学生を実験参加者とした. 参加者は実験を行う大学に通う平均年齢 23.5 歳 (SD = 3.32) の 10 名 (女性 3 名) の学生である. また, HMD を用いた VR の経験は, 1 人が経験なし, 3 人が年に 1 回程度がそれ未満, 2 人が半年に 1 回以上, 3 人が月に 1 回以上, 1 人が週に 1 回以上だった. 実験 1 時間あたりの参加報酬は 1000 円である.

被験者は図 1 (b) のように HMD を覗いて実験を行う. 実験で用いる縫合器具の先端は Matsuura [19] と同様に図 1 (c) のように検出精度を上げるために着色されている. 実験を行う前に顕微鏡縫合術訓練の事前練習を行った. 事前の実験から, 初心者であっても安定してくると, 3 縫合を 20 分で完了することができることがわかった. そこで, 事前練習では, 1 縫合 7 分弱で行えるようになるまで練習してもらい, この事前練習中に VR 酔いを起こした学生や学生自身が上達できないと判断した場合は, 実験には参加しないこととした. 本実験は被験者内実験であり, 各被験者は 1 条件 1 時間ずつ実験を行った. 目や手の疲れ, 集中力を考慮して練習と 2 条件の実験の間にそれぞれ 2 時間以上空けた.

本研究は以下の 2 つの条件で実験を行った. 条件 D: この条件では, 提案手法である円, バー, スコア, 時間が表示される. 円は左ピンセットの先端の位置によって色が変わり, バーはガーゼ全体が動くときと長さや色が変わり, スコアと時間は前回値と最高値と共に表示される. 条件 N: この条件では円, バー, スコア, 時間が表示されない. ただし条件間で有利不利が生じないようにするため, 最初の数秒間のみ縫合箇所を目標とする円を表示した.

実験の手順は以下である. (1) 条件 N を 1 縫合行う. (2) 無作為に選ばれた条件で飽きると感じるまで最大 1 時間実験を行う. (3) 条件 N で 1 縫合行う. 2 回目の別条件の実験もこの手順で行う.

5.3 評価方法

モチベーション維持や練習効果を調べるために, 定量評価として合計の縫合時間とスコアを評価した. 定性評価として, UEQ (User Experience

Questionnaire) [17]¹ と SUS (System Usability Scale) [2] の 2 つのアンケートを実施した. UEQ は, モチベーションが評価項目に入っている UX を評価するために使用した. また, SUS はシステムの使いやすさを測定するために使用した. さらに 2 つの条件が終わった後に自由記述のアンケートを行った. このアンケートは 5 段階評価とその理由を記述する質問である.

6 実験結果

6.1 定量評価の結果

縫合時間 定量評価の指標として合計の縫合時間を測定した. 合計の縫合時間は長いほどモチベーションの維持が可能であると言える. 図 5 に縫合時間の結果を示す. 合計の縫合時間を平均した値はそれぞれ条件 D で 2048 秒, 条件 N で 2004 秒だった. 合計の縫合時間の分析では条件間で有意差は見られなかった. そのため, H2 を支持する結果は得られなかった. また, 練習の効果を測るために 1 縫合にかかる時間を計測した. 1 縫合にかかる時間は短いほど上級者に近づくとと言える. 図 6 に実験後の 1 縫合の時間についての結果を示す. 平均値は条件 D で 234 秒, 条件 N で 284 秒だった. 実験後の時間についての対応のある t 検定を行ったところ, 条件間で有意差は見られなかったが, 有意傾向が見られた ($p = 7.4e-2 < 0.1$).

スコア 図 7 に条件それぞれにおいての実験前後の縫合のスコアの結果を示す. 実験前の平均値は条件 D で -11 点, 条件 N で -20 点, 実験後の平均値は条件 D で 36 点, 条件 N で 22 点だった. ウィルコクソンの符号順位検定を用いた結果, どちらの条件でも実験前後の縫合スコアに有意差があった (条件 D: $p = 9.3e-3 < 0.01$, 条件 N: $p = 1.3e-2 < 0.05$). また, 実験後の縫合のスコアの平均は条件 D で 36.22 点, 条件 N で 22.47 点だった. 実験後の縫合のスコアを条件間で分析したところ有意差は見られなかった.

6.2 定性評価の結果

UEQ の結果を図 8 に示す. UEQ の値が大きほど肯定的な値である. 不等分散を仮定した 2 標本の t 検定を行ったところ, この図から見栄え, 刺激, ノベルティで有意差があることがわかる ($p = 1.7e-2$

¹ User Experience Questionnaire, <https://www.ueq-online.org/>

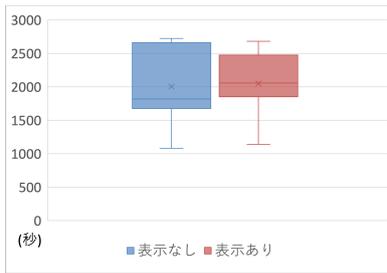


図 5. 合計の縫合時間の結果.

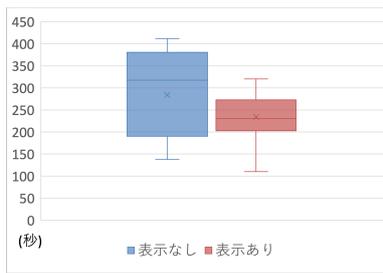


図 6. 実験後縫合時間の結果.

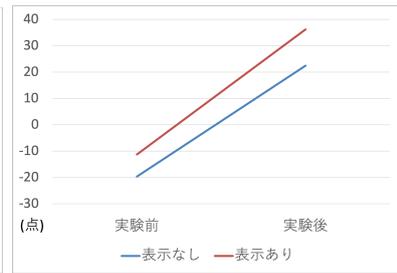


図 7. 実験前後のスコア.

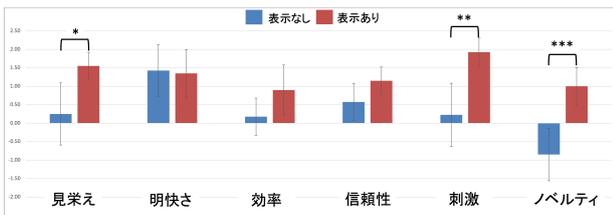


図 8. UEQ の結果 (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

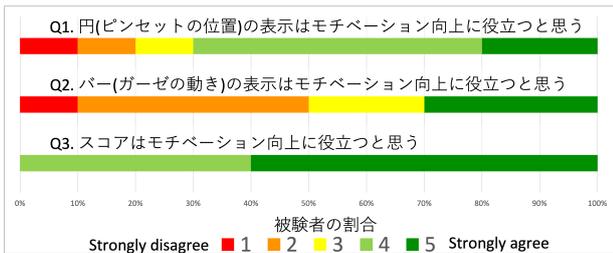


図 9. アンケートの結果.

< 0.05 , $p = 3.8e-3 < 0.01$, $p = 7.7e-4 < 0.001$). また SUS スコアの平均は、条件 D の時に 72.25, 条件 N の時に 62.25 だった. Bangor ら [1] の指標を参考にすると、各スコアは条件 D では「ACCEPTABLE」、条件 N では「LOW MARGINAL」である。

以上のことから、本システムは次の縫合への意欲を向上させると考えるため、H1 を支持すると考えられる。また、UEQ の結果から条件 D の方が UX が向上しているため、H3 を支持すると示唆される。

アンケートの結果を図 9 に示す。Q1. では 7 人が Strongly agree に近い 4 と 5 を選択した。Q2. では半数の被験者が Strongly disagree の 1 と 2 を選択した。その 5 人中 4 人が選択した理由として見えにくいという回答をした。Strongly agree である 5 を回答した 3 人全員がバーによってガーゼ全体を動かさないように注意できたと回答した。Q3. では被験者全員が Strongly agree の 4 か 5 を選択した。

アンケートの結果から、大半の初心者にとって円の表示はモチベーション向上に役立つと推察される。また、バーの存在が縫合の邪魔をすることがなく、注意に寄与したことがわかる。参加者全員がス

コアの表示がモチベーション向上に役立つと答えた理由を 2 つ述べる。1 つ目は、現在のスコアと同時に前回のスコアと最高スコアが表示されるため、自分の過去の記録と比較できるからである。2 つ目はピンセットやガーゼの動きと違い、スコアは自分では予想が付きにくいいため、スコアがモチベーションに寄与したことがわかる。

7 議論

本論文の提案システムはモチベーション向上のために顕微鏡縫合術訓練にゲーミフィケーションの要素を導入した。

定量評価では仮説 H2 に反して縫合練習時間に有意差がなかった。しかし、UEQ の結果から UX は向上した。特に見栄え、刺激、ノベルティで有意差があった。UEQ の計算ではモチベーションに関する質問肢が刺激の項として計算される¹ため、提案手法がモチベーション向上に寄与すると言える。

顕微鏡縫合術訓練は本来習得に時間がかかる訓練である。そのため、本提案が長期的なモチベーション向上に繋がるかどうか、長期の実験を行う必要がある。その際、ユーザごとに表示のカスタマイズが可能な UI を導入することで、長期的なモチベーション向上 [16] にも対応したい。

また、システムの限界は、HMD の慣れが使いやすさや疲れに大きく影響することである。実験参加者の中には HMD に閉塞感を感じたり酔ってしまっている人がいた。今回の実験では HMD の慣れの問題を軽くするために顎のせを導入した。ただし、長期的に使う場合、HMD に慣れると考えられる。

8 まとめ・今後の展望

本論文ではモチベーション向上のため、スコアリングやリアルタイムフィードバックといったゲーミフィケーションの要素を顕微鏡縫合術訓練を導入した。提案システムでは顕微鏡縫合術に必要な技術的要素の検出、スコア算出、リアルタイム可視化を行い、実験を通してその有用性を検証した。今後の課題としてモチベーションの向上を評価するために長期的な訓練の実験やユーザによる表示のカスタマイズが考えられる。

謝辞

本研究はJST CREST JPMJCR17A1, JST CREST JPMJCR17A3, JSPS 科研費 21K17788 の助成を受けた。

参考文献

- [1] A. Bangor, P. Kortum, and J. Miller. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Studies*, 4(3):114–123, May 2009.
- [2] J. Brooke. *SUS – a quick and dirty usability scale*, pp. 189–194. 01 1996.
- [3] C.-Y. Chang, H.-Y. Sung, J.-L. Guo, B.-Y. Chang, and F.-R. Kuo. Effects of spherical video-based virtual reality on nursing students' learning performance in childbirth education training. *Interactive Learning Environments*, 30(3):400–416, 2022.
- [4] H. Doughty, D. Damen, and W. Mayol-Cuevas. Who's Better? Who's Best? Pairwise Deep Ranking for Skill Determination. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6057–6066, 2018.
- [5] H. Egi, M. Okajima, M. Yoshimitsu, S. Ikeda, Y. Miyata, H. Masugami, T. Kawahara, Y. Kurita, M. Kaneko, and T. Asahara. Objective assessment of endoscopic surgical skills by analyzing direction-dependent dexterity using the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD). *Surgery Today*, 38(8):705–710, Aug 2008.
- [6] D. A. Freedman. *Statistical models: theory and practice*. cambridge university press, 2009.
- [7] D. Haga, H. Uekusa, Y. Node, S. Kubota, K. Kondo, and N. Sugo. Surgical Tips for STA-MCA Bypass from Misato Central General Hospital. *Surgery for Cerebral Stroke*, 49(1):64–69, 2021.
- [8] K. Harada, A. Morita, Y. Minakawa, Y. M. Baek, S. Sora, N. Sugita, T. Kimura, R. Tanikawa, T. Ishikawa, and M. Mitsuishi. Assessing Microneurosurgical Skill with Medico-Engineering Technology. *World Neurosurgery*, 84(4):964–971, 2015.
- [9] D. K.-h. Ho. Using smartphone-delivered stereoscopic vision in microsurgery: a feasibility study. *Eye*, 33(6):953–956, Jun 2019.
- [10] A. E. Hoerl and R. W. Kennard. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1):55–67, 1970.
- [11] B. Huang and K. Hew. Do points, badges and leaderboard increase learning and activity: A quasi-experiment on the effects of gamification. 12 2015.
- [12] P. J. Huber. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1):73 – 101, 1964.
- [13] T. Inoue, N. Kunii, A. Kumakiri, R. Otani, A. Tamura, I. Saito, and K. Tsutsumi. The Role of 10–0 Suturing Training under Desktop Microscope for the Mastery of Cerebrovascular Surgery: The Effectiveness of 80,000 Stitches. *Surgery for Cerebral Stroke*, 37(4):247–252, 2009.
- [14] T. Inoue, K. Tsutsumi, S. Adachi, S. Tanaka, K. Saito, and N. Kunii. Effectiveness of suturing training with 10-0 nylon under fixed and maximum magnification ($\times 20$) using desk type microscope. *Surgical Neurology*, 66(2):183–187, 2006.
- [15] T. Ishikawa, J. Moroi, N. Tamakawa, T. Mutoh, and N. Yasui. Technical Tips for Microvascular Anastomosis: How We Prepare Appropriate Operative Field and Manage Troubles in Bypass Surgery. *Surgery for Cerebral Stroke*, 38(2):77–82, 2010.
- [16] I. Kuramoto. An Entertainment-System Framework for Improving Motivation for Repetitive, Dull and Monotonous Activities. In I. Maurtua ed., *Human-Computer Interaction*, chapter 18. IntechOpen, Rijeka, 2009.
- [17] B. Laugwitz, T. Held, and M. Schrepp. Construction and Evaluation of a User Experience Questionnaire. Vol. 5298, pp. 63–76, 11 2008.
- [18] G. S. Maddala. *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs. Cambridge University Press, 1983.
- [19] M. Matsuura, S. Miyafuji, E. Wu, S. Kiyofuji, T. Kin, T. Igarashi, and H. Koike. CV-Based Analysis for Microscopic Gauze Suturing Training. In *Augmented Humans Conference 2021, AHs'21*, p. 169–173, New York, NY, USA, 2021. Association for Computing Machinery.
- [20] J. Park, D. Liu, M. Y. Yi, and R. Santhanam. GAMESIT: A gamified system for information technology training. *Computers & Education*, 142:103643, 2019.
- [21] M. Süncksen, H. Bendig, M. Teistler, M. Wagner, O. J. Bott, and K. Dressing. Gamification and virtual reality for teaching mobile x-ray imaging. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, pp. 1–7, 2018.
- [22] D. Siemon and L. Eckardt. *Gamification of Teaching in Higher Education*, pp. 153–164. Springer International Publishing, Cham, 2017.
- [23] T. Sugiyama, T. Nakamura, Y. Ito, K. Tokairin, K. Kazumata, N. Nakayama, and K. Houkin. A pilot study on measuring tissue motion during carotid surgery using video-based analyses for the objective assessment of surgical performance. *World Journal of Surgery*, 43(9):2309–2319, 2019.
- [24] Y. Tashiro, M. Matsuura, D.-H. Hwang, S. Miyafuji, S. Kiyofuji, T. Kin, T. Igarashi, and H. Koike. Introducing a Concept of Gamification to Microscopic Suturing Training. In *Augmented Humans 2022, AHs 2022*, p. 294–297,

New York, NY, USA, 2022. Association for Computing Machinery.

- [25] R. Uranishi, R. Tei, T. Sugimoto, T. Fukuda, and K. Shinomiya. Microsurgical Training Methods for Microvascular Anastomosis in Our Institute. *Surgery for Cerebral Stroke*, 38(3):153–157, 2010.
- [26] Z. Wang and A. Majewicz Fey. Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *International Journal of Computer Assisted Radiology and Surgery*, 13(12):1959–1970, Dec 2018.
- [27] T. Yoshikane, Y. Fujiwara, M. Uchimura, T. Miyazaki, Y. Akiyama, Y. Kimura, and T. Kagawa. Education in Basic Techniques of Cerebrovascular Surgery for Neurosurgical Trainees. *Surgery for Cerebral Stroke*, 47(1):33–40, 2019.
- [28] 印東 雅大. 顕微鏡手術の技法超入門 - これからの脳神経外科・形成外科のための. メディカ出版, 2009.

未来ビジョン

本論文が提案するシステムはモチベーション向上のために顕微鏡縫合術訓練にゲーミフィケーションの要素を導入した。我々のシステムをより良いシステムにするために考えらえる将来像を説明する。実際の顕微鏡縫合術では表示はされないため、スコアの大きさによって表示の有無を変更できるシステムが良いと考えられる。さらに、スコアの計算に用いている3つの要素から、ユーザの技術が足りていない箇所のみ表示をすることによってユーザは直すべき場所が明確になり、さらにモチベーション向上に繋がると考える。我々のシステムのターゲットは初心者であったため、今回の実験では縫合知識のないユーザが参加者だったが、システムの内容として中級者にも効果があると考えられるため、医学知識のある中級者にも効

果があるのか確かめたい。

今回のシステムではゲーミフィケーションの要素としてスコアリングやフィードバックを行った。本システムではスコアリングとスコアリングに用いた要素のデータを保存することができる。そのため、保存した新たなデータをもとにスコアリングの精度を上げていきたい。

本研究ではモチベーション向上を主な目的としてシステムを作成した。これに加えて教師側の指導の効率化に役立つポテンシャルがある。例えば、普段は学習者が本システムを使って練習を行い、特定の点数に達したら教師にアドバイスをもらう、という練習方法にすれば、初期段階の指導を短縮することができ、効率化につながる。この仕組みはゲーミフィケーションの要素の1つである、訓練者への報酬という役割も果たすと考える。

ユーザ嗜好と栄養摂取基準に基づくレシピ推薦手法の開発

尾関 沙羅* 小寺 正明† 石黒 勝彦† 西村 太一‡ 樋口 啓太†

概要. 日々の献立作成において、栄養バランスの管理は健康維持のために重要であるが、個人の食に対する嗜好を考慮しつつ栄養バランスを両立することは簡単ではない。本研究では、ユーザの食材嗜好及び、1日の栄養バランスに基づいたレシピの推薦手法を提案する。本手法では、嗜好食材に基づく嗜好スコアと、食事摂取基準に基づく栄養ペナルティを設定する。それらを用いてレシピを推薦することで、ユーザに栄養バランスを考慮したレシピ選択を促す。また、提案手法を用いた、1日分の献立を作成するアプリケーションを開発した。アプリケーションでは、レシピで使用している食材の入れ替えや、レシピにない食材の追加を可能にする。シミュレーション及びユーザスタディの結果から、既存の食材のみを考慮した推薦手法よりも良い栄養バランスで1日分の献立作成が可能であることを確認した。

1 はじめに

毎日の食事は、長年にわたり人々の健康に大きな影響を及ぼす。各国の保健機関（例：日本人の食事摂取基準 [12]）は、食事摂取基準のガイドラインを作成しており、炭水化物、タンパク質、脂質、ビタミンなどの栄養素の摂取基準量を定めている。

近年のレシピサイトでは多くのレシピを収録しており、食材名やレシピ名を用いた検索で、所望のレシピを簡単に見つけることができる。しかし、それらのレシピは食事摂取基準には配慮されていない、脂質や塩が過多となる場合がある。また、ユーザの多くは、それぞれの食品にどれだけの栄養素が含まれているかを把握していないため、栄養バランスを考慮せず、個人の食に対する嗜好だけがレシピ選びに反映される可能性がある。

嗜好を考慮しながら、健康に配慮した単品レシピの選択を助けるようなレシピ推薦手法はいくつか提案されている [5, 6]。我々は、1食分の栄養バランスのみに着目するのではなく、1日を通した栄養バランスを考慮することで、無理なく健康的な食事を続けることができると考えている。これは、嗜好を重視した結果1食の栄養が偏っていたとしても、残りの食事で栄養バランスを改善するレシピを選択することができるからである。しかし、1日分の献立は1食分よりも選択するレシピ数が増えるため、毎日レシピを食事摂取基準に従いながらバランスよく組み合わせ続けることは簡単ではない。

そこで本研究では、ユーザの食材嗜好及び、1日の栄養バランスに基づいたレシピの推薦手法を提案する。本手法では、嗜好食材に基づく嗜好スコアと、食事摂取基準に基づく栄養ペナルティを設定する。それらを用いてレシピを推薦することで、ユーザに

栄養バランスを考慮したレシピ選択を促す。また、栄養バランスを過度に考慮することで生じ得る満足度の低下を防ぐため、レシピで使用している食材の入れ替えや、レシピにない食材の追加を可能にする。

本研究の主な目的は、食材嗜好と栄養バランスを考慮した1日分の献立を作成する、新しいアプリケーションを開発することである。我々は、図1に示すようなコンポーネントで構成される、1日分の献立を作成する Web アプリケーションを開発した。このアプリケーションでは、嗜好スコアと栄養ペナルティを各レシピで計算し、そのスコアをもとに推薦レシピの一覧を表示する。ユーザは嗜好食材の選択やレシピの決定、レシピに使用されている食材の置き換え、追加をすることができる。そのようなユーザ入力、スコアの計算、推薦レシピの更新を繰り返し、ユーザに栄養バランスを考慮した献立の作成を促す。

本稿では、提案アルゴリズムが嗜好と栄養バランスを両立できるかをシミュレーションによって検証した。また、このアプリケーションを用いて1日の献立を作成するユーザスタディを実施し有効性を確認した。

2 関連研究

個人適応推薦 [8] には、音楽推薦 [17]、スマート通知 [10]、美的画像推薦 [1, 7]、食品推薦 [2, 3] など多くのアプリケーションで重要な研究分野がある。多くのレシピ情報がオンラインで公開されるようになり、レシピ推薦が推薦研究の重要なトピックになりつつある。多くの研究は、栄養情報に基づいた健康的なレシピの推薦 [5, 6, 16] に焦点を当てている。また、個人に合わせてレシピを推薦する個人適応推薦の研究もある [15, 4]。Teng は食材の補完ネットワークに基づいて食材の削除、置き換え、追加を考慮した食材ベースレシピ推薦法を提案した [14]。し

* 東京農工大学

† 株式会社 Preferred Networks

‡ 京都大学



図 1. アプリケーションのコンポーネント

かし、これまでの研究は1レシピもしくは1食分の推薦に着目した研究がほとんどである。そのため、食事摂取基準に従って複数のレシピを組み合わせながら、1日分の献立を作成することはまだ簡単ではない。本研究では、ユーザの食材嗜好及び1日の栄養バランスに基づいたレシピの推薦手法を提案し、献立作成アプリケーションを提供する。

3 定式化とアルゴリズム

各レシピ $r_i \in \mathbf{R}$ に対して推薦スコア $score_i$ を定義し、このスコアによって推薦を行う。 $score_i$ を繰り返し計算することで、推薦レシピを更新する。 \mathbf{R} は全 N レシピの集合である。 レシピ r_i は食材リスト $\{f_{r_i,1}, f_{r_i,2}, \dots\}$ と栄養素リスト $\{e_{r_i,1}, e_{r_i,2}, \dots, e_{r_i,L}\}$ を含み、 L は基準栄養素数である。 $\mathbf{Y} = \{f_0, f_1, \dots, f_M\}$ は嗜好食材としてユーザが M 個の全食材候補から選択した m 個の食材である。 $score_i$ は嗜好スコア $Preference(\mathbf{Y}, r_i)$ と栄養素ペナルティ $Nutrient(r_i, \zeta)$ から計算される。

$$score_i = \frac{1 + Preference(\mathbf{Y}, r_i)}{1 + Nutrient(r_i, \zeta)} \quad (1)$$

ζ はユーザが選択したレシピ $\{r_j, r_k, \dots\}$ のリストである。 この推薦は反復的であるため、スコアの計算とレシピの選択を繰り返す。 なお $\zeta = \emptyset$ (例えば、1回目の反復) のとき、栄養ペナルティは0である。

嗜好スコア $Preference(\mathbf{Y}, r_i)$ は、各レシピで使用される食材と嗜好食材 \mathbf{Y} との一致 $A(\mathbf{Y}, r) = |\mathbf{Y} \cap r|$ 、置き換え $B(\mathbf{Y}, r)$ 、追加 $C(\mathbf{Y}, r)$ に基づく3つの基準の重み付き和である。

$$Preference(\mathbf{Y}, r) = \alpha A(\mathbf{Y}, r) + \beta B(\mathbf{Y}, r) + \gamma C(\mathbf{Y}, r) \quad (2)$$

ここで、係数 α, β, γ は定数である。 食材の置き換えを表す $B(\mathbf{Y}, r)$ は、食材と食材の $M \times M$ 隣接行列 \mathbb{M} を用いて求める。 $M_{i,j}$ には、 i 番目の食材が j 番目の食材に置き換え可能な場合は1、そうでない場合は0が入っている。 \mathbb{M} は食材置き換えに関する料理の知識に基づいて手作業で作成した。

また、食材の追加を表す $C(\mathbf{Y}, r)$ は、追加可能な食材リスト \mathbb{P} を用いて求める。 \mathbb{P} は、対象のレシピと類似したレシピで使用されている食材を参考にして決まる。 レシピ間の類似度は非負行列因子法 (NMF [9]) で求めた。 i 番目のレシピが j 番目の食材を使用している場合は i, j の要素に1、それ以外の場合は0であるレシピと食材の行列を作成し、この行列に対して NMF を実行する。 上位の類似レシピを使用し、類似レシピの食材から追加可能なセットを作成する。 全ての実験において、各係数の値を $\alpha = 100, \beta = 15, \gamma = 10$ とした。

栄養ペナルティ $Nutrient(r_i, \zeta)$ は、栄養量の過不足があるレシピに対してペナルティを加えるものである。 食事摂取基準は、条件式 $\mathbb{E} = \{e_1, e_2, \dots, e_L\}$ で表され、 l 番目の栄養素の基準 e_l は、摂取量の最小値と最大値を持つ $e_l = \{e_{l,min}, e_{l,max}\}$ で表される。 ただし、栄養素の規格によっては、最小値、最大値のどちらか一方しかとまらないものもある (例: ビタミン C は摂取量の上限がない)。 栄養素ペナルティは、これまで選択したレシピに候補レシピを追加したときに過不足が生じた場合のみペナルティを与える。 そのため、候補レシピ r の l 番目の基準栄養素量 $e_{l,r}$ を

Algorithm 1: 嗜好食材と栄養バランスを考慮した1日分の献立を作成するためのレシピ推薦アルゴリズム

```

Input:  $\mathbf{Y} = \{f_1, f_2, \dots\}$ ,  $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$ 
 $\zeta = \{\}$ 
for  $t = 1, \dots, T$  do
     $\text{scores} = \left\{ \frac{1 + \text{Preference}(\mathbf{Y}, \mathbf{r}_i)}{1 + \text{Nutrient}(\mathbf{r}_i, \zeta)} \right\}$  foreach
         $\mathbf{r}_i \in \mathbf{R}$ 
     $\mathbf{r} = \text{GetSelection}(\mathbf{R}, \text{scores})$ 
     $\zeta \leftarrow \mathbf{r}$ 
     $\mathbf{Y} = \mathbf{Y} - \mathbf{Y} \cap \mathbf{r}$ 
end
    
```

これまでの選択したレシピ $\hat{r} \in \zeta$ の栄養素値の合計に足した値 $\sum_{\hat{r} \in \zeta} e_{l, \hat{r}}$ が最大摂取基準 $e_{l, max}$ を超えた場合には、ペナルティ $(e_{l, r} + \sum_{\hat{r} \in \zeta} e_{l, \hat{r}}) / e_{l, max} - 1$ を返す。一方で、最小基準 $e_{l, min}$ を下回った場合には $1 - (e_{l, r} + \sum_{\hat{r} \in \zeta} e_{l, \hat{r}}) / e_{l, min}$ を返す。

3.1 アルゴリズム

アルゴリズム 1 に、提案するレシピ推薦手順を示す。GetSelection($\mathbf{R}, \text{scores}$) はユーザに推薦結果を表示し、ユーザが選択したレシピ \mathbf{r} を献立 ζ に追加する (scores は各レシピへのスコア score_i の集合とする)。また、ユーザはレシピで使用される食材の置き換えや追加が可能である。なお、ユーザは入力リスト \mathbf{Y} にいつでも食材を追加することができる。T は1日の献立のレシピ数である。

3.2 栄養素・レシピデータ

食材データベースには、日本機構の公式食品成分表「日本食品成分表 2015 版 7 訂 [11]」を使用した。食品成分表には、成分名 (2191 成分・食品) とその栄養データ (61 カテゴリー) が収録されている。レシピデータには、キックマンホームレシピ¹ 1,932 件を使用した。レシピごとに食材名や分量、調理手順などが記載されている。各レシピの成分や栄養素を人数分で割って正規化している。今回使用するレシピには全部で 290 の食材が使用されている。

4 レシピ作成アプリケーション

我々は、提案アルゴリズムに基づく献立作成 Web アプリケーションを開発した。図 1 にアプリケーション画面の各コンポーネントを示す。まず、ユーザが自分の好みに応じて食材を選択すると、図 1(1-2) のように選択された食材が UI 上に表示される。食材

を入力すると、提案手法のスコアが高い順に推薦レシピが表示される (図 1(3))。ユーザは、選択したレシピの材料リストを図 1(4) のように確認することができる。また、レシピ決定後には、選択した嗜好食材をレシピの食材と置き換えたり、追加したりすることができる。あらかじめ、全てのレシピから主食 (お米、麺類) を削除しているため、ユーザは1日分の主食を選択する必要がある (図 1(5))。ユーザは決定したレシピを確認し、図 1 (6) のようにレシピの詳細ページに移動することができる。赤とグレーのバーはそれぞれ、現在の栄養バランスと食事摂取基準を示している (図 1(7))。レシピを選択すると、図 1(8) のように、そのレシピを作るために必要な食材リストが表示される。

図 2 は 1, 4, 7 個目のレシピを選択しているときの栄養素のグラフである。赤のバーは、献立の栄養素を、グレーのバーは食事摂取基準を示す。黄色のバーは、献立とその時選択しているレシピの総栄養素を示している。

図 3 はレシピで使用される食材と嗜好食材 (図 1(2)(4)) の例である。選択した食材リストの \circ と \times のアイコンは、その食材が献立のレシピ全体で使用済みかもしくは未使用であることを示している。赤色のハイライトは、直近で選択したレシピで使用される食材であることを示す。緑と青のハイライトは、それぞれ入れ替え、追加を推奨していることを示す。

栄養素のガイドラインとして「日本人の食事摂取基準 (2015 年版) [12]」を参照した。このガイドラインには 36 種類の栄養素の摂取量が記載されているため、ユーザは 36 種類の栄養基準から自由に対象栄養素をピックアップすることができる。なお、本アルゴリズムは、あらゆる種類の栄養素の摂取目標に対応可能であり、目標となる栄養素の値を調整することで個人に合わせた栄養素の摂取が可能となる (例: 高血圧予防のための減塩)。

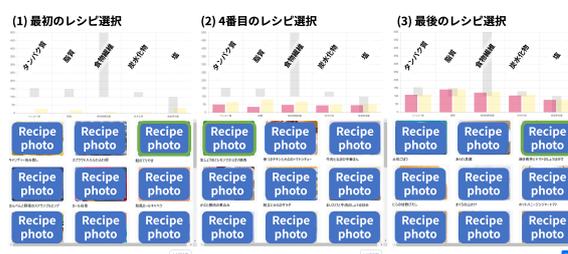


図 2. レシピ選択による栄養素の推移

5 シミュレーション評価

提案アルゴリズムが食事摂取基準を満たすレシピを推薦するかどうか、また食材の好みを推薦に反映するかどうかを確認するためシミュレーションによ

¹ <https://www.kikkoman.co.jp/homecook/index.html> (Jan. 4, 2022 Accessed.)



図 3. 選択したレシピで使用される食材

る評価実験を行った。シミュレーションでは、提案手法 (proposed) とベースライン (baseline) を比較した。ベースラインでは、嗜好食材とレシピの食材との一致のみを考慮する。このような食材を基にレシピを検索する方法は、一般的なレシピサイトでも使われる手法である。具体的には、式 2 の $A(\mathbf{Y}, \mathbf{r})$ のみに基づく手法をベースラインとして使用した。また、食材の入れ替えや追加の有効性を確認するため、(A) 食材の置き換えや追加を許可しない条件と (B) 許可する条件のそれぞれで実施した。

対象栄養素は、タンパク質、脂質、炭水化物、食物繊維、ナトリウムの 5 種類である。表 1 の min と max は、各栄養素の基準摂取量である。シミュレーションでは、日本人の栄養摂取基準 [12] における 18-29 歳の男性の食事摂取基準に対応した。

5.1 手順

シミュレーションでは、ユーザがどのようにアプリケーションを利用し、どのようにレシピを選択するかを参考にした。レシピの候補には主食 (米、麵) が含まれていないため、1 日 3 食分の主食として白米 150g を 1,2,4 品目のレシピを選択した後に追加する。以下のシミュレーション手順を 100 回実行した。

1. シミュレータは嗜好食材 (selected ingredients) を野菜 6 種、肉 2 種、きのこ 1 種、魚 1 種の計 10 種になるようにランダムに 1 つ選択する。
2. アルゴリズムは全レシピのスコアを算出し推薦レシピを更新する。
3. シミュレータは推薦されたレシピの中から、最もスコアが高い 5 つのレシピを候補としてランダムに 1 つ選択し、献立に追加する。
4. (B) 条件では、レシピに入れ替えもしくは追加可能食材がある場合、食材を 1 つ 1/2 の確率で入れ替えもしくは追加する。
5. アルゴリズムは選択したレシピで使用した食材を嗜好食材から削除する。

6. 選択したレシピが 1,2,4 品目の場合、白米 150g を献立に追加する。

7. 7 つのレシピが決定するまで (2)-(6) を繰り返す。

5.2 評価

選択した 7 つのレシピから食事摂取基準に対する各栄養素の過不足値を計算した。栄養が食事摂取基準範囲内であれば、過不足値は 0 とした。100 回のシミュレーションで、各栄養素の平均値と標準偏差を計算した。また、7 つのレシピ中で使用する嗜好食材の数 (最大 10) を確認し、アルゴリズムがレシピ推薦における食材の嗜好をどのように反映しているかを確認した。

5.3 シミュレーション結果

表 1 は、シミュレーション調査の結果である。上段は (A) 食材の入れ替えと追加を行わない、下段は (B) 食材の入れ替えと追加を行った結果である。いずれも、提案手法ではベースラインよりも全ての栄養素の過不足の平均値が低くなった。また嗜好食材の使用量については、(A) ではベースラインで多くなり、(B) では提案手法で多くなった。シミュレーションの結果、提案手法では、食事摂取基準を満たそうとするレシピを選択することや、食材の入れ替えが許可されることで嗜好食材を多く使用することができることが分かった。

6 ユーザスタディ

提案手法の効果を調査するため、ユーザスタディを実施した。開発したアプリケーションを用いて、式 2 の $A(\mathbf{Y}, \mathbf{r})$ のみに基づくベースラインの検索手法と、提案手法それぞれで 1 日の献立を作成した。アプリケーションの GUI はどちらも同じである。

ユーザスタディの参加者は 11 名 (女性: 1 名, 男性: 10 名, 平均年齢 36.5 歳 (SD: 8.94)) であり、料理経験は様々であるが、全員が日常生活における健康的な食生活に関心を持っている。

6.1 手順

ユーザ調査のセッションは全てオンラインで行った。タスクを開始する前に、アプリケーションの使用方法を説明し、練習をした。全ての参加者は、ベースラインと提案手法の両方でタスクを実行したが、タスクを行う順序は、全ての参加者についてランダムであり、どちらがベースラインでどちらが提案手法なのかやその違いなどはユーザには知らせていない。選択する嗜好食材は 10 種類程度とし、どちらの手法でも、食材の入れ替えや追加が可能である。また、対象となる栄養素は、タンパク質、脂質、炭水化物、食物繊維、塩である。

表 1. 食事摂取基準とシミュレーション結果

(A) 食材の置き換えや追加を許可しない

栄養素名	最小	最大	平均摂取量 (SD)		過不足 (SD)	
			proposed	baseline	proposed	baseline
タンパク質	86g	133g	97.39g(12.22)	112.88g(28.19)	0.75 (2.04)	6.15 (13.70)
脂質	59g	88g	78.88g(18.01)	93.52g(26.44)	0.64 (2.46)	14.21 (17.85)
炭水化物	331g	431g	276.45g(41.92)	267.66g(38.19)	56.34 (38.89)	66.08 (29.11)
食物繊維	20 g	N/A	27.16g(6.61)	25.54g (5.96)	0.10 (0.42)	0.14 (0.56)
塩	N/A	8.0g	8.98g(1.37)	12.96g (3.02)	1.08 (1.27)	4.96 (3.02)
選択食材の平均使用数			proposed 8.02 (1.15)	baseline 8.15 (1.03)		

(B) 食材の置き換えや追加を許可

栄養素名	最小	最大	平均摂取量 (SD)		過不足 (SD)	
			proposed	baseline	proposed	baseline
タンパク質	86g	133g	97.29g(12.22)	107.88g(27.15)	1.16 (3.03)	5.49 (11.84)
脂質	59g	88g	73.89g(10.53)	87.71g(24.21)	1.10 (4.55)	10.47 (14.89)
炭水化物	331g	431g	278.55g(41.92)	269.57g(32.88)	53.93 (32.80)	62.19 (30.97)
食物繊維	20 g	N/A	28.38g(5.96)	27.59g (7.82)	0.07 (0.45)	0.27 (0.93)
塩	N/A	8.0g	9.05g(1.25)	12.77g (2.81)	1.13 (1.16)	4.77 (2.81)
選択食材の平均使用数			proposed 8.96 (1.34)	baseline 8.34 (1.34)		

表 2. ユーザスタディの結果

栄養素名	最小	最大	平均摂取量 (SD)		過不足 (SD)		p-value
			proposed	baseline	proposed	baseline	
タンパク質	86g	133g	97.27g (11.19)	126.24g (17.60)	0.00 (0.00)	4.68 (7.99)	0.043*
脂質	59g	88g	78.88g (18.01)	108.93g (25.61)	4.11 (11.18)	22.39 (23.96)	0.050[†]
炭水化物	331g	431g	336.27g (54.66)	328.55g (74.67)	18.45 (30.56)	35.98 (31.20)	0.283
食物繊維	20 g	N/A	36.60g (4.87)	35.82g (9.82)	0.00 (0.00)	0.00 (0.00)	N/A
塩	N/A	8.0g	9.31g (1.50)	12.83g (3.06)	1.48 (1.27)	4.83 (3.06)	0.006**

6.2 評価

選択した7つのレシピから食事摂取基準に対する各栄養素の過不足値を計算した。栄養が食事摂取基準範囲内であれば、過不足値は0とした。各栄養素で参加者の平均値と標準偏差を計算した。また、7つのレシピ中で使用する嗜好食材の数を確認した。参加者は両手法でタスク終了後アンケートに回答した。事後アンケートでは、献立や栄養バランスの満足度などを質問項目とした。参加者は5段階評価（そう思わない=1, どちらでもない=3, そう思う=5）で質問に回答した。

6.3 ユーザスタディ結果

表2は、レシピ選択タスクの定量的な結果であり、11人の参加者の各栄養素量の平均と過不足値の平均を含んでいる。提案手法では、タンパク質、脂質、炭水化物、ナトリウムの過不足値が、ベースライン条件よりも少なくなっている。食物繊維については、

表 3. 提案手法とベースラインに対するアンケート結果。
Q1: 選択したレシピへの満足度. Q2: 選択したレシピの栄養バランスの主観評価

対象	平均	標準偏差	中央値	
Q1	proposed	3.82	1.11	4
	baseline	3.00	0.95	3
Q2	proposed	3.36	0.88	3
	baseline	2.63	0.77	2

両条件とも過不足値はゼロであった。

Wilcoxon 符号順位検定での有意性については、タンパク質 ($p < 0.05$) とナトリウム ($p < 0.01$) で有意差が認められた。

表3は、参加者11名の回答結果である。Q1ではレシピ選択のユーザ満足度、Q2ではユーザが感じる栄養バランスについて確認した。Q1では、提案条件がベースライン条件よりも平均値、中央値が高

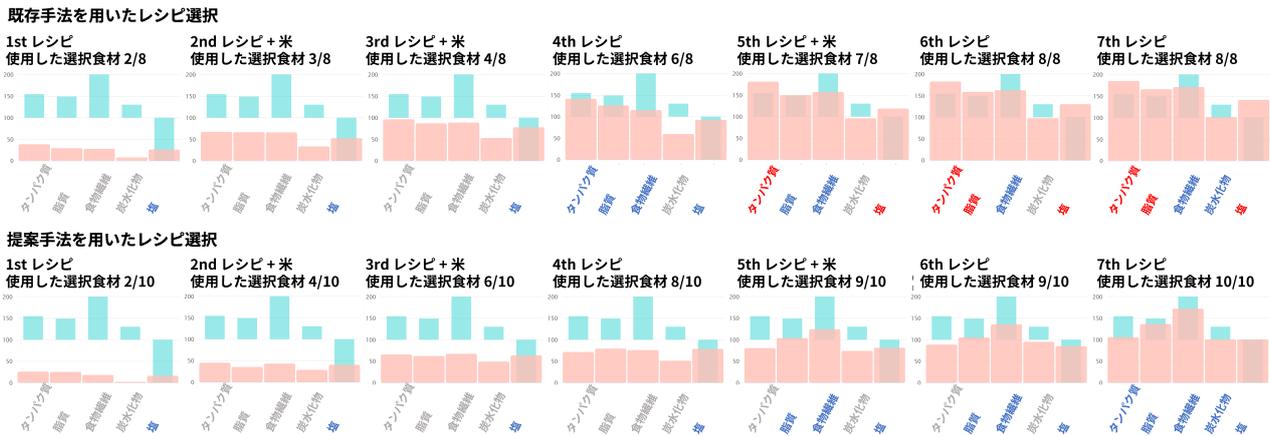


図 4. 7つのレシピを選択した際の栄養バランスの推移

くなった。また、Q2でも提案条件の方がベースライン条件よりも平均値、中央値が高くなった。

実験中、参加者は自由に提案レシピの詳細ページを見ることができたため、提案手法を使った方がより多く詳細を見る傾向があり、タスク実施時間は提案条件とベースライン条件でそれぞれ 601.1 秒 (SD: 869.8)、421.5 秒 (SD: 463.1) となった。

7 議論

シミュレーションの結果、提案アルゴリズムは食材の一致のみでレシピを推薦するベースラインと比較して、食事摂取基準をより満たすようなレシピを選択できた。食材の置き換えを行わない場合、提案アルゴリズムは、栄養素の使用量がベースラインよりも少なくなった。しかし、置き換えと追加を行った場合、提案アルゴリズムは、嗜好食材の使用量においてベースラインを上回った。これは、提案アルゴリズムが推薦スコアの計算において、食材の置き換えと追加を考慮するためである。

ユーザ調査の結果、提案手法によるレシピ推薦の有効性が示され、ユーザは栄養バランスの取れた献立を作成できた。特に、塩の測定値は摂取基準を超えないように低塩分レシピが推薦されたことで、ベースラインと比較して大幅に減少した。多くの健康指針で、ナトリウムの過剰摂取は健康に深刻な影響を与えることが報告されているため、日々の食事において摂取基準に従うことが望ましい [13]。図 4 は、提案手法とベースラインにおける献立の栄養素推移を示している。赤が献立栄養素で、青が食事摂取基準である。ベースラインでは、栄養ペナルティを考慮していないため、3つの栄養素が基準摂取量を超えている。一方提案手法では、食事摂取基準を満たすような献立を作成できた。この結果から、提案手法は栄養バランスの制約に適応し、食事摂取基準に基づいて推薦スコアを更新していることが分かった。

定性的フィードバックでは、表 3 の Q1 より、提案手法の方がベースラインよりもレシピ選択の満足度が高いことが分かった。自由記述では、「栄養バランスを考えてレシピを構成できた」など、提案手法の推薦結果を好意的に受け止めていることが分かった。また、「食材を入れるだけでレシピの提案が見られるのが使いやすい」、「好きな食材をいくつか入れると、それをカバーするようにおすすめレシピが変わり、栄養素を見ながらレシピを選べるのがうれしい」など機能に対するポジティブな意見が挙がった。一方で、「減塩メニューを優先的に提案してほしい」、「1週間など、より長期的なレシピを選択できるようにしてほしい」という要望もあった。

8 まとめと今後の課題

本研究では、ユーザの食材に対する嗜好を考慮した上で、食事摂取基準を満たすように推奨レシピを更新するレシピ推奨アルゴリズムを提案した。また、提案手法を用いて、1日分の献立を作成することができるアプリケーションを開発した。提案手法による推薦と、アプリケーションによる食材検索を比較するシミュレーションとユーザスタディを実施した。その結果、提案手法を用いたレシピ推薦により、ユーザは栄養バランスの取れた1日分の献立を作成できた。さらに、選択したレシピに対する参加者の満足度が向上した。

今後の課題は、ユーザ個人に最適化した情報を利用することである。提案するアルゴリズムは、あらゆる食材、レシピ、栄養基準を適応させることができるが、我々のユーザ研究では、限られた情報のみを使用した。したがって我々は、ユーザに最適化された栄養素接種要件を入力できるように拡張することを計画している。また、食事制限などのユーザ入力に基づいたパーソナライズされた推薦を提供する追加のユーザ実験の実施を計画している。

参考文献

- [1] H. Cramer, V. Evers, S. Ramlal, M. Someren, L. Rutledge, N. Stash, L. Aroyo, and B. Wielinga. The Effects of Transparency on Trust in and Acceptance of a Content-Based Art Recommender. *User Modeling and User-Adapted Interaction*, 18(5):455–496, nov 2008.
- [2] D. Elswailer, M. Harvey, B. Ludwig, and A. Said. Bringing the” healthy” into Food Recommenders. *DMRS*, 1533:33–36, 2015.
- [3] D. Elswailer, C. Trattner, and M. Harvey. Exploiting Food Choice Biases for Healthier Recipe Recommendation. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR ’17, p. 575–584, New York, NY, USA, 2017. Association for Computing Machinery.
- [4] J. Freyne and S. Berkovsky. Intelligent food planning: personalized recipe recommendation. In *Proc. IUI*, pp. 321–324, 2010.
- [5] M. Ge, F. Ricci, and D. Massimo. Health-aware Food Recommender System. In *Proc. RecSys*, pp. 333–334, 2015.
- [6] E. Gorbonos, Y. Liu, and C. T. Hoàng. NutRec: Nutrition Oriented Online Recipe Recommender. In *Proc. WI*, pp. 25–32, 2018.
- [7] K. Higuchi, H. Tsuchida, E. Ohn-Bar, Y. Sato, and K. Kitani. Learning Context-Dependent Personal Preferences for Adaptive Recommendation. *ACM Trans. Interact. Intell. Syst.*, 10(3), nov 2020.
- [8] S. Kulkarni and S. F. Rodd. Context Aware Recommendation Systems: A review of the state of the art techniques. *Computer Science Review*, 37:100255, 2020.
- [9] D. D. Lee and H. S. Seung. Learning the Parts of Objects by Non-Negative Matrix Factorization. *Nature*, 401(6755):788–791, 1999.
- [10] A. Mehrotra, R. Hendley, and M. Musolesi. PrefMiner: Mining User’s Preferences for Intelligent Mobile Notification Management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp ’16, pp. 1223–1234, New York, NY, USA, 2016. ACM.
- [11] S. Ministry of Education Culture Sports and Technology. STANDARD TABLES OF FOOD COMPOSITION IN JAPAN - 2015 - (Seventh Revised Version). online, 2015. Jan. 4, 2022 Accessed.
- [12] L. Ministry of Health and Welfare. Dietary Reference Intakes for Japanese (2015). online, 2015. Jan. 4, 2022 Accessed.).
- [13] W. H. Organization. Guideline: sodium intake for adults and children. online, 2012. Jan. 4, 2022 Accessed.).
- [14] C.-Y. Teng, Y.-R. Lin, and L. A. Adamic. Recipe Recommendation Using Ingredient Networks. In *Proceedings of the 4th Annual ACM Web Science Conference*, WebSci ’12, p. 298–307, New York, NY, USA, 2012. Association for Computing Machinery.
- [15] N. Twomey, M. Fain, A. Ponikar, and N. Sarraf. Towards Multi-Language Recipe Personalisation and Recommendation. In *Proc. RecSys*, pp. 708–713, 2020.
- [16] W. Wang, L. yu Duan, H. Jiang, P. Jing, X. Song, and L. Nie. Market2Dish: Health-aware Food Recommendation. *ACM. TOMM*, pp. 1–19, 2021.
- [17] X. Wang, Y. Wang, D. Hsu, and Y. Wang. Exploration in Interactive Personalized Music Recommendation: A Reinforcement Learning Approach. *ACM Trans. Multimedia Comput. Commun. Appl.*, 11(1), sep 2014.

メアンダコイル++：継続的なウェアラブルコンピューティングのための衣類全面での安全で高効率な無線給電

高橋 亮*

雪田 和歌子*

横田 知之*

染谷 隆夫*

川原 圭博*

概要. ライフログやヘルスケアのためのウェアラブル機器の多くは電池容量が小さいため頻繁に充電する必要があり、長期の使用が困難である。そこで、導電糸で実装された布型無線給電器（コイル）を衣服へ内蔵し使用中のウェアラブル機器へ常に給電することで、継続的なウェアラブルコンピューティングを目指す。しかし、人体の至る所にあるウェアラブル機器への給電に向け布型コイルを人体スケールへ拡大すると、電磁界が人体内部まで浸透するため送信電力を制限する必要があり、さらに導電糸の損失が大きいため給電効率が著しく低くなる。そこで、安全で比較的高効率な人体スケールの布型無線給電器、メアンダコイル++を提案する。メアンダコイル++は、人体への電磁界浸透を抑制できるコイル構造と液体金属ベースの低損失な導電糸とを用いる。これにより、人体の70%程度を覆う広範囲な給電領域を実現し、さらに電磁界曝露を抑制しながら最大2.5Wの電力を約40%の給電効率でウェアラブル機器へ給電できた。

1 はじめに

スマートウォッチやスマート衣服などのウェアラブル機器が日常生活に浸透しつつある。ウェアラブル機器は、使用用途に応じて身体の適切な場所へ装着することで、行動認識や、パーソナルヘルスケア、ライフログなど、多種多様なウェアラブルコンピューティングを手軽に実現できる。しかし、多くのウェアラブル機器は人体へ違和感なく装着できるよう小型・軽量化が求められるため、搭載する電池のサイズや容量に制限が生じ、数日程度しか機能しない。Apple Watch [2] や Google Glass [1] などのウェアラブル機器は、高性能ゆえに適度な使用で1日程度しか電池がもたない。ゆえに、ウェアラブルコンピューティングの性能向上を目指し、ウェアラブル機器の高機能化や高頻度な使用を行うと、小容量な電池の搭載が前提であるウェアラブル機器では身体外で充電される時間の割合が多くなるため、駆動時間が著しく減る。

この課題の解決に向け、身体に配置されたウェアラブル機器への給電パスを衣類上へ実装することが研究されてきた [15]。これにより、大容量のモバイル電源などを繋げた布型給電パスを介して人体上のウェアラブル機器へ給電できるため、ユーザは電池容量や電池残量、充電管理などを気にせず、ウェアラブル機器の高性能化や頻繁な使用が容易にできる。さらに、ウェアラブル機器に低消費電力で間欠的な駆動を要求する環境発電と比べ [5]、布型給電は比較的大きな電力を継続的にウェアラブル機器へ給電できる。最も単純な方法は、導電糸を介してウェアラブル機器へ直接給電する布型有線給電であるが、電気的接続のために配線を露出させる必要がある、

感電などの危険が伴う [13, 14, 15]。そこで、電気的接続が不要な無線給電が布型給電の手法として注目されている。特に、マイクロ波や電界と比べ、比較的人体との干渉が小さい磁界を介して、安全で高効率な給電が可能なコイルが有力である [13, 20]。

しかし、これまでの布型コイルを人体スケールへ拡張すると、1) 送信電力が数百 μW を下回り、2) 給電効率が1%~3%程度と著しく低くなる。なぜなら、1) 無線給電で通常使用される渦巻状コイルは、人体スケールへ拡張すると強い電磁界が人体内部まで生じるため、送信電力を下げ電磁界曝露を抑える必要があり、さらに、2) 布型送電コイルは服の着用性を損なわないように柔軟であるが損失の大きな導電糸で構成されるため、人体スケールのコイルでは損失が数百 Ω を超えるほど大きくなるからである。ゆえに現状の布型コイルは人体の数% (約 80 cm^2) のサイズに狭めることで、体内への電磁界浸透の範囲を狭めつつ、損失を抑えることで給電効率を10%程度まで向上させ、数 mW 程度の電力を超低消費電力なセンサコイル (NFC タグ) へ給電するのが限界である [13]。またセンサの配置・数に制約が生じるため、センサの使用用途毎に布型コイルの配置などを設計し直す必要がある。

そこで本論文では、安全で比較的高効率な人体スケールの布型無線給電器、メアンダコイル++¹を紹介する。メアンダコイル++は、1) メアンダコイルから着想を得た、人体への電磁界曝露を抑制できるコイル構造と2) 液体金属を利用した損失の小さな導電糸とを用いる。これにより、人体スケールのメアンダコイル++ (約 6000 cm^2 , 人体上半身の約70%の範囲) とウェアラブル機器サイズの小型受電コイ

* 東京大学

¹ 本論文は、ACM CHI 2022にて発表済である [19]。

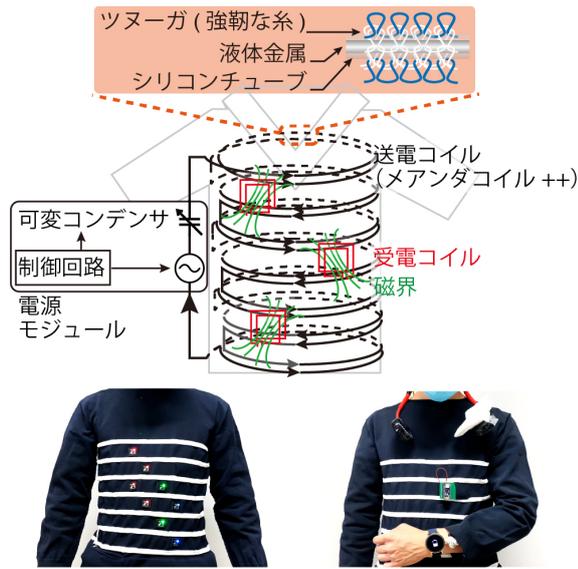


図 1. メアンダコイル ++ の概要. 安全で比較的高効率な送電コイル (メアンダコイル ++) と電源モジュール, 受電コイルで構成される.

ル (約 16 cm^2) のサイズ比が $375 : 1$ である状況で, 衣服上の様々な場所に配置された複数の受電コイルへ約 40% の AC-to-AC 給電効率で給電できる. さらに, 人体への電磁界曝に関する国際的な安全ガイドライン (i.e., ICNIRP: international commission on non-ionizing radiation protection) [3] に準拠しながら, 約 2.5 W の電力を安全に受電コイルへ給電できることを確認した. 以下, 設計・実装, 応用例, 評価, 今後の課題を述べる.

2 設計と実装

2.1 システム概要

図 1 で示すように, 提案システムは, 液体金属入りのシリコンチューブで巻かれた送電コイル (メアンダコイル ++), 送電コイルへ接続された電源モジュール, ウェアラブル機器へ接続された受電コイルの 3 つから構成される. 電源モジュールは, 商用無線給電で使用される周波数の一つである 6.78 MHz の高周波電源と, 送電コイルの共振周波数を調整するための可変コンデンサ, 受電コイルの接近を検出する制御回路で構成される. 無線給電の方式は, Kurs らが提案した磁界共振結合方式である [12]. 磁界共振結合方式では, コイルを用いた単純な無線給電と異なり, 高い Q 値² を備える LC 共振器 (ある周波数で共振するように直列接続されたチップコンデンサで調整された共振コイル³) を用いるた

² $Q = 2\pi fL/r$. 一般的に Q 値が高いと給電効率も高い.

³ コイルのインピーダンス, $Z = r + j\{2\pi fL - 1/(2\pi fC)\}$ が周波数 f_0 で $Z = r$ と共振することを意味する.

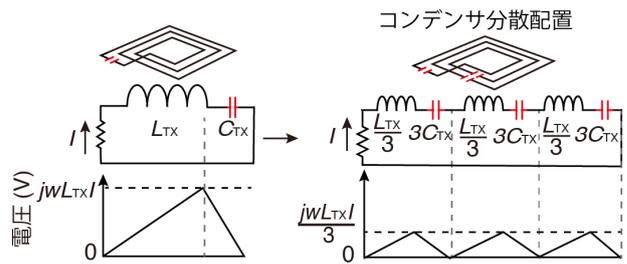


図 2. コンデンサ分散配置手法の概要.

め, 送受電コイル間の磁氣的結合が弱い状況下 (送受電コイルのサイズ比が大きい場合) において, 効率的に電力を給電できる. 本システムでは, 人体スケールの送電コイルと小型の受電コイルのサイズ比が $375 : 1$ であるため, 送受電コイル間の結合係数 $k (= M/\sqrt{L_{TX}L_{RX}}, M$: 相互インダクタンス) が 0.033 であり, 一般的なケース ($k > 0.1$) と比べ小さい. ゆえに, 磁界共振結合方式を用いる.

動作原理は以下の通りである. まず, 電源モジュールはユーザの体型や姿勢に応じて送電コイルに接続する可変コンデンサの静電容量を調整し, 送電コイルの共振周波数を動作周波数 (6.78 MHz) へ合わせる. 次にモジュールが受電コイルの接近を検出した後, 送電コイルは受電コイルへ給電し始める. モジュールは, 受電コイルが送電コイルへ接近する際に送電コイルの入力インピーダンスが変化することを検知し, 受電コイルの接近を認識する. これにより, 無駄な給電を回避できる. 受電コイルは, 送電コイルからの交流電力を整流器などで直流電力へ変換し, ウェアラブル機器へ供給する.

2.2 メアンダコイル ++

一般的なコイル (渦巻状コイル) を人体スケールへ拡大すると, 人体内部まで電磁界が浸透し人体への熱作用・刺激作用を引き起こす電磁界干渉が生じる. ゆえに, 送信電力量を著しく制限する必要がある [3]. この電磁界干渉は, 1) 人体スケールのコイルから生じる浮遊容量などが人体へ及ぼす強い電界と, 2) コイルが人体内部へ発生させる強い磁界に起因する.

そこで課題 1 に関しては, 長いコイルへ複数のコンデンサを等間隔に配置し, 高周波 (6.78 MHz) での浮遊容量などの電界結合を抑制するコンデンサ分散配置手法を用いた [7, 17]. 図 2 で示すように, コンデンサで分割された各コイルの配線長は高周波の短波長に対して十分短いため, 長い配線に起因する高周波での浮遊容量の影響を抑制し, 各コイルは浮遊容量の小さなコイルとして機能する. 具体的には, N 個のコンデンサにより各コイルの両端の電圧差は, 1 個のコンデンサで共振させたコイルより N 倍小さくなる [7] (図 2 では $N = 3$) ため, 浮遊容量の

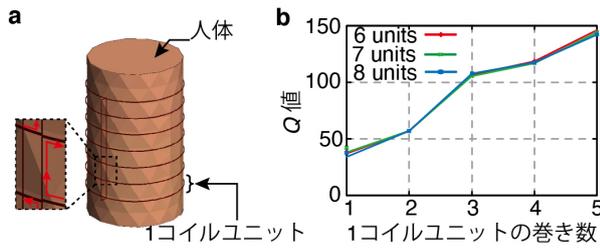


図 3. (a) メアンダコイル ++ の電磁界シミュレーション用モデルと (b) コイルの巻き数に対する Q 値、すなわち給電性能の結果。

人体に対する誘電損失が $1/N$ ($= 1/N^2 \times N$) 倍となり、コイルと人体との電界結合を抑制できる [17].

次に課題 2 に関しては、メアンダコイルに着想を得たメアンダコイル ++ を提案する。メアンダコイルは、渦巻コイルと異なり、一卷きごとに時計回りと反時計回りと向きを変えながら、コイルを巻く。さらに、円筒形状のメアンダコイルは、比較的強い磁界を衣服付近に閉じ込め、体内へ浸透する磁界を打ち消す [18]. ウェアラブル機器の多くは、衣服や身体に装着して動作するため、衣服近傍のみに強い磁界を閉じ込めても給電可能である。ただし、一卷きごとに向きを変えるメアンダコイルでは、同じ向きへ複数巻くことで Q 値を高くする通常のコイルに比べ、 Q 値が低くなる。なぜなら、コイルのインダクタンス (L) は巻数の 2 乗に比例して増加し、コイルの損失である抵抗 (r) はコイルの線長に線形に比例して増加するため、巻き数が多いほど Q 値が大きいためである。そこで、本論文では複数巻ごとに巻き方向を変えるメアンダコイル ++ を提案する。メアンダコイル ++ は、コンデンサ分散配置を利用し高周波でコイルの巻き数を増やすことで、 $Q(= \omega L/r)$ 値を大きくできる。

メアンダコイル ++ の有効性を示すために、電磁界シミュレーション (HyperWorks FEKO) を用いてコイルの巻き数に対する Q 値を調べた。図 6a で示すように、シミュレーションモデルは、1) 円柱状の人体モデル (比誘電率:132, 導電率:0.66 S/m, 直径: 26 cm) [6] と液体金属で構成されたメアンダコイル ++ (コイル直径: 26.5 cm, 導電糸の直径: 1 mm, 間隔: 5 cm, 抵抗率: $29.6 \times 10^{-8} \Omega \text{m}$ [10]) とで構成される。図 6b に、1 コイルユニットの巻き数を 1 から 5 まで 1 巻刻みで変化させた時のメアンダコイル ++ の Q 値を示す。この結果から、巻き数が増えるにつれて Q 値が大きくなるのがわかる。ただし、 Q 値が 100 を上回るコイルは、呼吸などによるコイルのわずかな変形により共振周波数が大きく変化し給電効率が著しく低下したため、本研究では 1 コイルユニットの巻き数を 2 とした。

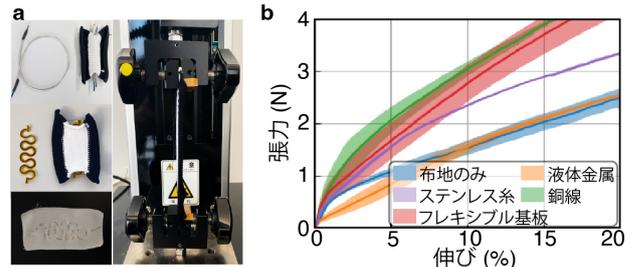


図 4. (a) 長さ 5 cm の生地へ内蔵した 4 種類の導電糸の (b) 引張試験の結果。

表 1. 6.78 MHz での各導電糸の抵抗率 ($10^{-8} \Omega \text{m}$) .

フレキシブル基板	銅線	ステンレス糸	液体金属入りチューブ (伸び: 0%/10%)
20.5	20.1	246.6	29.6/40.1

2.3 液体金属を利用した導電糸

人体スケールの高効率な布型送電コイルの実現には、布型コイルに使用される導電糸が柔軟かつ低損失である必要がある。ステンレス糸のような一般的な導電糸は、柔軟であるが損失が大きい [11]. また銅線や細いフレキシブル基板などの硬い導電糸は、低損失であるが伸縮性がない。一方で、eGaIn と呼ばれる液体金属は高い導電性と低毒性を持ち、 15.7°C 以上の温度で液体状態を保つ。ゆえに、eGaIn を伸縮性のあるシリコンチューブへを入れることで、柔軟で低損失な導電糸が実現できる [10, 9].

図 4a に、4 種類の導電糸 (フレキシブル基板、銅線、ステンレス糸、液体金属入りチューブ) の引張試験の結果を示す。フレキシブル基板は幅 5 mm, 銅線とステンレス糸の直径は 1 mm, チューブは内径 1 mm, 外径 2 mm である。銅線とステンレス糸は生地へ S 字に縫い込み、伸びやすくした。チューブとフレキシブル基板は袋付きの生地へ入れた。引張の参考としてただの生地の引張試験も行った。図 4b から、液体金属入りチューブは生地の伸縮を損なわない一方で、フレキシブル基板や銅線は硬いため生地の伸縮を損なうことが分かる。さらに各導電糸の抵抗率を比較した (表 1)。その結果、1) 液体金属は 6.78 MHz で低損失な銅線などと同程度の抵抗率を備え、2) 10% の引っ張り以下では低い抵抗率を示した。衣服着用時に衣類が 10% 以上引っ張られることはほぼない。ゆえに、伸縮性があり低損失な導電糸として液体金属入りチューブを用いる。

2.4 実装

ここでは、主に送電コイル (メアンダコイル ++) の作製工程を説明する。送電コイルの作製工程を図 5a-c で示す。まず、衣服を縫製なしで作製できる編み機 (MACH2XS15S, 島精機製作所) で衣服

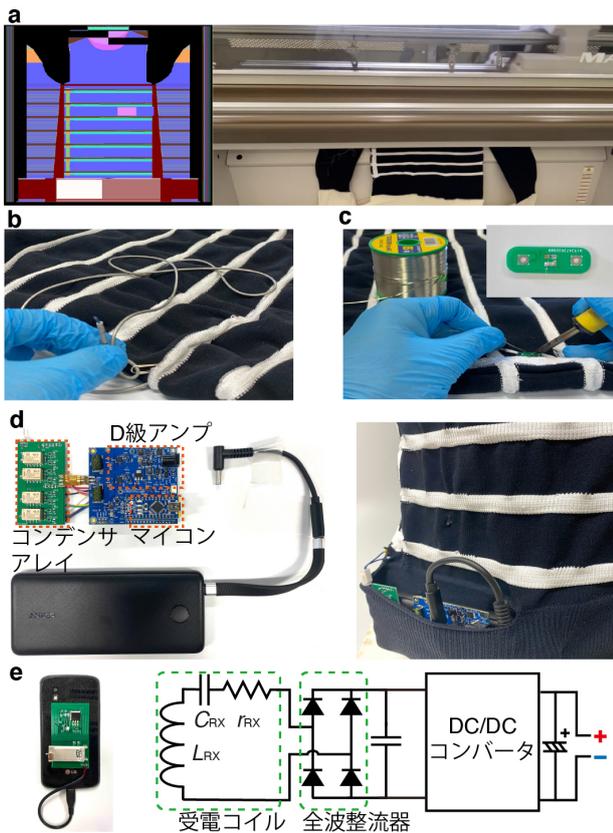


図 5. 提案システムの実装の概要. (a) - (c) 送電コイルの作製方法. (a) チューブを入れるための袋付き服を編んだ後, (b) 液体金属入りチューブを袋へ入れ, (c) コイルの共振周波数を 6.78 MHz へ合わせるためにコンデンサを配置する. (d) 電源モジュールの概要. 6.78 MHz の高周波電源 (EPC9065) とマイコン (Arduino Nano), コンデンサアレイなどから成る. (e) 受電コイルの概要. コイルと整流器などから成る.

を編む (図 5a). チューブを通すための袋を衣服へ実装するが, この袋はツヌーガ (東洋紡) と呼ばれる強靱な糸で実装した. これにより, 洗濯などの外的要因によるチューブの破れを防ぐ. 袋の編成には多層編みができるジャカード編みを用いた. その後, 液体金属 (eGaIn) を入れたシリコンチューブを袋に通す (図 5b). チューブの作製では, 注射器で eGaIn を注入しチューブの両端をワイヤで塞いだ. eGaIn の総使用量は 9.8 mL である. 最後に, コンデンサ分散配置用に, リジッド基板へはんだ付けした 16 個の 470 pF のチップコンデンサを等間隔にチューブ間へ挿入し, 送電コイルの共振周波数を 6.78 MHz に合わせた (図 5c). 送電コイルのインダクタンス, 抵抗, Q 値, 重さは, 18.7 μ H, 14.5 Ω , 55, 133 g である. 電源モジュール (図 5d) と受電コイル (図 5e) の作製は [19] を参考にしてほしい.

3 評価

3.1 安全性

まず, 人体へ浸透する電磁界が神経刺激や熱作用を引き起こす危険があるため, メアンダコイル++の電磁界曝露の影響を評価した. 評価には, 国際的な電磁界曝露制限ガイドライン (ICNIRP) を参照した [3]. この安全ガイドラインには, 基本制限 (生物学的影響が予想される閾値) と参考レベル (基本制限より厳しい閾値) の 2 つの制限値が定められている. ここでは, 2 つの制限値を下回る最大送信電力を電磁界シミュレーションで調べた. 具体的には, 6.78 MHz でのある送信電力に対する, 参考レベルの閾値の指標である電界強度と磁界強度, 基本制限の閾値の指標である全身平均比吸収率 (SAR) と局所 SAR を評価した. より専門的な評価には正式な認定が必要であり, 以下の実験は予備評価にとどまる.

基本制限および参考レベル以下の最大送信電力を求めるために, HyperWorks FEKO を使用した. 図 6a で示すシミュレーションモデルでは, 線径などを調整することで, モデルのインピーダンスをプロトタイプへ揃えた. また, 渦巻状コイルはメアンダコイル++の寸法に揃えた. 図 6b-c に, 送信電力が 1 W の時の各モデルにおける電磁界分布を示す. これから, メアンダコイル++は身体を貫く電磁界を抑制する一方で, 渦巻状コイルは身体に到達する電磁界が浸透することが分かる. さらに, 表 2 の SAR に対する最大送信電力から, メアンダコイル++の最大送信電力 (52 W) は渦巻状コイル (2.6 W) より 20 倍程度大きいことが分かる. ただし, 参考レベルに関していずれも人体とコイルの距離がほぼないため, 数 μ W の電力で頭打ちとなる. これらの結果から, メアンダコイル++は, 一般的なコイルよりも安全に電力を給電できる. 今回のアプリケーション例では, 電源モジュール内部の D 級アンプの耐圧を考慮し, 最大送信電力を 10 W とした. これは, 図 8 で示すように, 数 W 程度の電力で充電されるウェアラブル機器への送信電力として十分である.

3.2 給電効率

次に, 送受電コイル間の AC-to-AC の給電効率を評価した. 具体的には, 1) 受電コイルの位置, 2) 送受電コイルの距離, 3) ユーザの姿勢, 4) 人体または発砲スチロールマネキンかどうかの 4 つのパラメータを変えながら, AC-to-AC の給電効率を調べた. ユーザは 30 代の日本人男性 1 名である. AC-to-AC の給電効率の算出には, S パラメータからインピーダンス整合下での最大給電効率を導き出す式を用いた [22]. S パラメータの測定では, 図 7a で示すように, メアンダコイル++ (送電コイル) と 4 cm \times 4 cm の受電コイルをベクトルネットワークアナライザ (ZNB20,R&S) へ接続した.

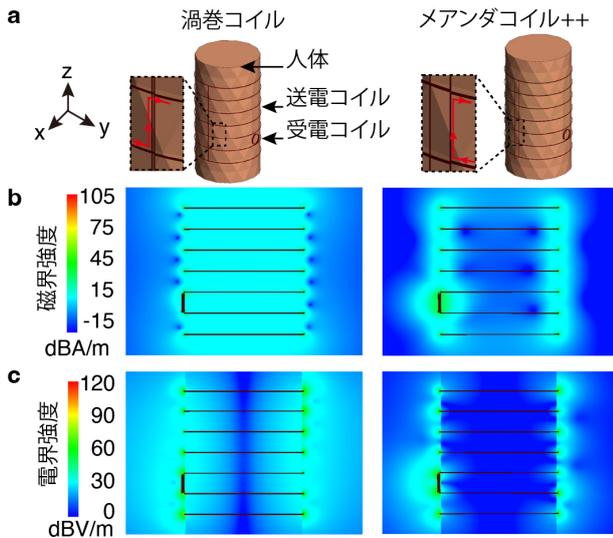


図 6. (a) メアンダコイル ++ と渦巻状コイルのシミュレーションモデルと、送信電力が 1 W の時の (b) 磁界分布と (c) 電界分布.

表 2. 図 6a のモデルにおける、参考レベル（電界強度と磁界強度）と基本制限（全身平均 SAR と局所 SAR）を超えない最大送信電力.

	渦巻状コイル (受電コイル有)	メアンダコイル ++ (受電コイル有)
電界強度	6.5 W	30 W
磁界強度	17 μ W	37 μ W
全身平均 SAR (公衆曝)	2.6W	59 W
局所 SAR (公衆曝)	19 W	52W

図 7b に、受電コイルの位置に応じた給電効率の結果を示す。これから、受電コイルをメアンダコイル ++ の配線間に配置すると給電効率が高くなる一方で、受電コイルを配線上へ配置すると、メアンダコイル ++ から生じる磁界が受電コイルをほぼ貫くことがないため、給電効率がほぼ 0 となることから分かる。この位置ずれによる給電効率の低下は、位置ずれ・角度ずれに頑強な、受電コイルを正三角形に並べた受電コイルアレイなどにより回避できる [16]. 次に、図 7c に、受電コイルを 2.5 cm の位置に置いた時の、送受電コイルの距離に対する給電効率の結果を示す。これから、受電コイルが衣服から 1 cm 以上離れると給電効率が最大給電効率の半分以下になることが分かる。数 W の電力をウェアラブル機器へ送るためには、受電コイルを服から 1 cm 以内に配置することを前提とする必要がある。配線間隔を大きくすることで給電距離を伸ばすことができるが、人体への電磁界曝露の影響も大きくなる。ゆえに、配線距離と電磁界曝露の関係性を明らかにし、適

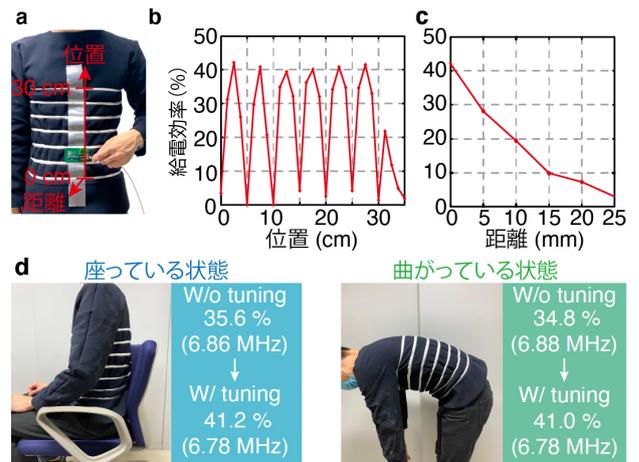


図 7. (a) AC-to-AC の給電効率の評価環境と (b) 受電コイルの位置、(c) 受電コイルとの距離、(d) ユーザの姿勢に対する AC-to-AC の給電効率.

切な配線間隔を今後求める必要がある。

図 7d に、ユーザの姿勢が立つ状態から座る状態、曲がる状態へ変化した時の給電効率を示す。メアンダコイル ++ の変形によりインダクタンス、すなわち共振周波数 ($f = 1 / (2\pi\sqrt{L_{TX}C_{TX}})$) が変化するため、ユーザの姿勢が変わると給電効率が低下する (図 7d の W/o tuning を参照)。しかし、電源モジュールに内蔵した可変コンデンサで送電コイルの静電容量 (C_{TX}) を再調整することで、効率の低下を緩和できる (図 7d の W/ tuning を参照)。最後に、人体または発泡スチロールマネキンをメアンダコイル ++ を着るに関わらず、給電効率がほぼ同じであることを確認し、メアンダコイル ++ が人体との電磁的干渉を抑制できることを確認した。本稿では詳細な説明を省くが、メアンダコイル ++ は 40 回以上洗濯可能である [19].

4 アプリケーション例

メアンダコイル ++ の技術的有効性を示すために、様々なウェアラブル機器への給電事例を示す。従来の布型無線給電器が数十 mW 程度の電力しかウェアラブル機器へ送電できない一方で、メアンダコイル ++ は、人体との電磁界干渉を抑制しながら、スマートフォン、ネックファン、スマートウォッチなどへ数 W を給電できる (図 8a-c)。そのため、高機能であるが電池容量の小さいウェアラブル機器を、ユーザは電池残量を気にせずいつでも使用できる。ゆえに、メアンダコイル ++ がウェアラブル機器のための大容量な電池となることで、ウェアラブル機器は電池容量に制約を受けず、小型軽量化と高機能化を両立できる。スマートフォン、ネックファン、スマートウォッチへの送電コイルからの入力電力は、各々 10.0 W, 3.6 W, 2.1 W 程度である。

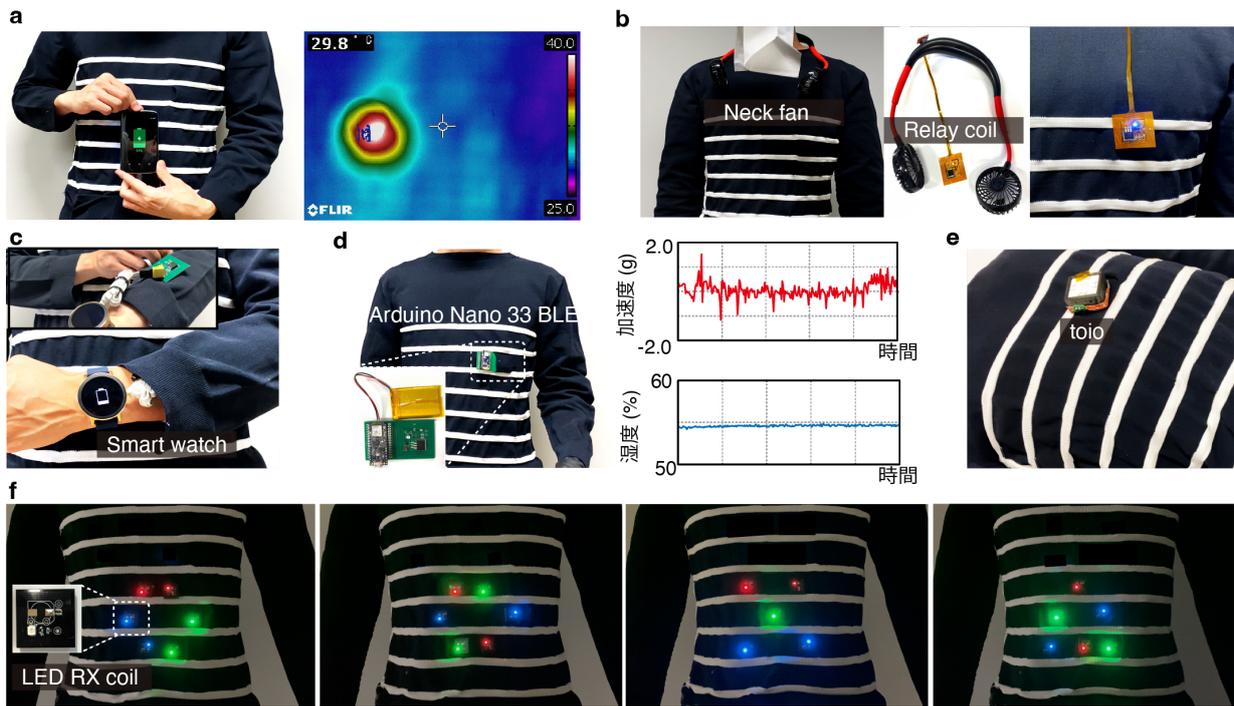


図 8. メアンダコイル++ のアプリケーション例.

さらに、メアンダコイル++は、従来のポケットサイズの布型無線給電器と異なり、人体スケールの布型無線給電器である [13, 20]. そのため、図 1 で示すように、衣服全体の様々な場所に配置されたウェアラブル機器を同時に駆動できる. 例えば、衣類上の Arduino Nano 33 BLE を充電しながら、ユーザが望む体の部位での活動状況や周囲の湿度を把握できる (図 8d). また、ユーザの体を移動する toio のようなウェアラブル機器 [8] を位置によらず給電できる (図 8e). ゆえに、従来の布型無線給電に比べ、ウェアラブル機器の配置に応じて布型無線給電の設計をし直す必要がない. Arduino と toio への入力電力は、各々 1.4 W, 1.5 W 程度である.

最後に、メアンダコイル++は、布型有線給電と異なり、電気的接続による配線の損傷を回避し、安全に無線で給電できる [14, 21, 15]. ゆえに、ユーザは配線の損傷や感電を気にせずとも、機器の配置を何度も変更できる. 一例として、複数の LED の点灯配置を自由に変更できる LED 展示を紹介する (図 8f). LED 展示では、一般的に LED の数が増えるにつれ配線接続の手間が著しく増加する. しかし、メアンダコイル++は機器を置くだけで良いので、数十個の LED の点灯パターンを手軽に変更できる.

5 まとめと今後の課題

本論文では、人体スケールの無線給電に向け、安全で比較的高効率な布型無線給電、メアンダコイル

++ を紹介した. メアンダコイル++は、伸縮性チューブ内に低損失な液体金属を入れた導電糸と、メアンダコイルに着想を得た安全なコイル構造を用いた. 本論文では、プロトタイプを作製し、安全性や給電効率などの評価を行い、アプリケーション例を数種類紹介した. 今後メアンダコイル++が、様々なテキスタイル製品に組み込まれることで、日常生活での継続的かつ高機能なウェアラブルコンピューティングを持続できることを期待する.

メアンダコイル++に関する今後の課題は以下の通りである. まず、継続的なウェアラブルコンピューティングの実証に向け、1) 体型の違い、2) 衣服の種類、3) 長期間の使用に対する給電効率の低下を評価する必要がある. 1の予備実験として、電源モジュールが備える共振周波数の再調整機能を利用することで、体格の異なる 3 人の 30 代男性に対して、最大給電効率の低下を 1~3% に抑えることを確認した. 今後は、より多くのユーザを集め、1-3の要因による給電効率の変化を評価する.

次に、現状の作製手法は複数の手作業を要するため、手間と時間がかかる. そこで、1) 編むと同時にチューブなどの太い糸を生地へ織り込むインレイ編みを活用し、人がチューブを挿入する作業を省くこと [4] と、2) チップコンデンサをあらかじめチューブ内へ入れることで、はんだ付けを不要にすることを目指す. 2により衣類の伸縮性を損なうリジッド基板を除くこともできる. これらの改良をもとに、メアンダコイル++の完全自動作製を目指す.

謝辞

本研究は JST ERATO 川原万有情報網プロジェクト (JPMJER1501) と JST 未来社会創造事業探索加速型 (JPMJMI17F1), JST ACT-X (JPM-JAX21K9), JSPS 特別研究員奨励費 (JP22J11616) との一環として実施されました。本研究の実装を手伝って下さった池内尚史さんと笹谷拓也さんに深く感謝致します。

参考文献

- [1] Google Glass Enterprise Edition 2 drops to \$999 and adds Qualcomm 's XR1, 2019.
- [2] Apple Watch Battery and Performance, 2020.
- [3] ICNIRP GUIDELINES, 2020.
- [4] L. Albaugh, L. Yao, and S. Hudson. Digital fabrication of soft actuated objects by machine knitting. *Conference on Human Factors in Computing Systems - Proceedings*, 2019.
- [5] Y. W. Chong, W. Ismail, K. Ko, and C. Y. Lee. Energy Harvesting for Wearable Devices: A Review. *IEEE Sensors Journal*, 19(20):9047–9062, 2019.
- [6] A. Christ, M. G. Douglas, J. M. Roman, E. B. Cooper, A. P. Sample, B. H. Waters, J. R. Smith, and N. Kuster. Evaluation of wireless resonant power transfer systems with human electromagnetic exposure limits. *IEEE Transactions on Electromagnetic Compatibility*, 55(2):265–274, 2013.
- [7] B. Cook and I. Lowe. A large-inductance, high-frequency, high-Q, series-tuned coil for NMR. *Journal of Magnetic Resonance (1969)*, 49(2):346–349, 9 1982.
- [8] A. Dementyev, H. L. C. Kao, I. Choi, D. Ajilo, M. Xu, J. A. Paradiso, C. Schmandt, and S. Follmer. Rovables: Miniature on-body robots as mobile wearables. *UIST 2016 - Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 111–120, 2016.
- [9] M. D. Dickey. Stretchable and Soft Electronics using Liquid Metals. *Advanced Materials*, 29(27):1–19, 2017.
- [10] M. D. Dickey, R. C. Chiechi, R. J. Larsen, E. A. Weiss, D. A. Weitz, and G. M. Whitesides. Eutectic gallium-indium (EGaIn): A liquid metal alloy for the formation of stable structures in microchannels at room temperature. *Advanced Functional Materials*, 18(7):1097–1104, 2008.
- [11] J. Gong, Y. Wu, L. Yan, T. Seyed, and X.-d. Yang. Tessutivo. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, No. 1, pp. 29–41, New York, NY, USA, 10 2019. ACM.
- [12] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić. Wireless power transfer via strongly coupled magnetic resonances. *Science*, 317(5834):83–86, 2007.
- [13] R. Lin, H. J. Kim, S. Achavananthadith, S. A. Kurt, S. C. Tan, H. Yao, B. C. Tee, J. K. Lee, and J. S. Ho. Wireless battery-free body sensor networks using near-field-enabled clothing. *Nature Communications*, 11(1):1–10, 2020.
- [14] A. Noda and H. Shinoda. Frequency-division-multiplexed signal and power transfer for wearable devices networked via conductive embroideries on a cloth. *IEEE MTT-S International Microwave Symposium Digest*, pp. 537–540, 2017.
- [15] E. R. Post and M. Orth. Smart fabric, or 'wearable clothing'. *International Symposium on Wearable Computers, Digest of Papers*, pp. 167–168, 1997.
- [16] T. Sasatani, Y. Narusue, and Y. Kawahara. Genetic Algorithm-Based Receiving Resonator Array Design for Wireless Power Transfer. *IEEE Access*, 8:222385–222396, 2020.
- [17] R. Takahashi, M. Fukumoto, C. Han, T. Sasatani, Y. Narusue, and Y. Kawahara. TelemetRing: A Batteryless and Wireless Ring-shaped Keyboard using Passive Inductive Telemetry. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pp. 1161–1168, New York, NY, USA, 10 2020. ACM.
- [18] R. Takahashi, W. Yukita, T. Sasatani, T. Yokota, T. Someya, and Y. Kawahara. Twin Meander Coil : Sensitive Readout of Battery-free On-body Wireless Sensors Using Body-scale Meander Coils. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4), 2021.
- [19] R. Takahashi, W. Yukita, T. Yokota, T. Someya, and Y. Kawahara. Meander Coil++: A Body-Scale Wireless Power Transmission Using Safe-to-Body and Energy-Efficient Transmitter Coil. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [20] X. Tian, P. M. Lee, Y. J. Tan, T. L. Wu, H. Yao, M. Zhang, Z. Li, K. A. Ng, B. C. Tee, and J. S. Ho. Wireless body sensor networks based on metamaterial textiles. *Nature Electronics*, 2(6):243–251, 2019.
- [21] I. Wicaksono, C. I. Tucker, T. Sun, C. A. Guerrero, C. Liu, W. M. Woo, E. J. Pence, and C. Dagdeviren. A tailored, electronic textile conformable suit for large-scale spatiotemporal physiological sensing in vivo. *npj Flexible Electronics*, 4(1), 2020.
- [22] M. Zargham and P. G. Gulak. Maximum achievable efficiency in near-field coupled power-transfer systems. *IEEE Transactions on Biomedical Circuits and Systems*, 6(3):228–245, 2012.

未来ビジョン

本研究では、布型無線給電の一つとして衣服型無線給電器を提案した。しかし、この衣服だけではズボンや、顔、手などにあるウェアラブル機器へ給電できない。そこで未来ビジョンとしては、右図で示すように日常空間の至る所にあるテキスタイル製品へのメアンダコイル++の実装を提案する。これにより、あらゆるウェアラブル機器がいずれかの布型無線給電器を通して給電されることで、ウェアラブル機器が小さな電池容量の制約から解放され、高機能かつ半永続的に振る舞うことができる。

さらには、現状のウェアラブル機器は専用ハードウェアを内蔵することで小型軽量化かつある程度の長期動作を実現するが、これでは設計に手間と時間がかかり、ウェアラブル機器の普及は進まない。しかし、メアンダコイル++がウェアラブル機器の電池代わりとなることで、専用ハードウェアありきのウェアラブル機器から、電力はある程度消費するが汎用的なハードウェアを内蔵したウェアラブル機器へ移り変わることで、Arduinoのように、設計・実装・応用のしやすさが飛躍的に向上することで、ウェアラブル機器の日常空間への浸透が促進されると期待できる。最終的には、アク

セサリを自然と身につけるように、ユーザが電池管理不要なウェアラブル機器を自然と身につけている世界を目指したい。



音漏れ信号を用いたヒアラブルデバイス向け手形状認識手法の検討

雨坂 宇宙* 渡邊 拓貴† 杉本 雅則† 志築 文太郎*

概要. 次世代のウェアラブルデバイスとしてヒアラブルデバイスが注目されている。ヒアラブルデバイスは、これまでの音楽鑑賞/通話機能だけでなく、音声ガイドやメッセージ読み上げなど様々な機能が搭載されると考えられる。これらの種々の機能の切り替え、ON/OFF 操作をヒアラブルデバイスから実行するために、多くのデバイス操作手法が研究されてきた。本研究ではヒアラブルデバイスの音漏れが手と耳介で形成される空間内を反響することに着目し、耳介反響音を測定、解析することで現在の手形状を認識する手法を開発した。提案手法はヒアラブルデバイスにて一般的に搭載されているスピーカと外側マイクで実装できる点でヒアラブルデバイスとの親和性が高い。本研究ではプロトタイプデバイスを用いて、5名の実験参加者から5種類の手形状のデータを収集した。実験の結果、参加者ごとの個人分類モデルによる認識精度が平均92.7% (F 値) であることを確認した。

1 はじめに

近年の研究/技術開発に伴い、イヤホン型のウェアラブルデバイス（ヒアラブルデバイス）が更に注目されている。ヒアラブルデバイスとは従来の音楽鑑賞や通話としての機能だけでなく、様々な機能が付加された高機能なイヤホン型デバイスを指す。現在も多くの研究や開発が進められており、今後はより多くのアプリケーションがヒアラブルデバイスに搭載されると考えられる。一方、現在市販されているヒアラブルデバイスの多くはスマートフォンと連携して使用することを想定しているが、デバイスを操作するために、まずはスマートフォンの画面をタッチする必要があり、ユーザビリティを損う。そのため、ヒアラブルデバイスのみを用いたデバイス操作手法が求められている。市販されているヒアラブルデバイスでは、デバイスに感圧センサや静電容量センサ、物理ボタンを埋め込み、タッチや押し込みを認識する手法が採用されている。これらのセンサの利用は、感圧センサではデバイスデザインの制約が生じ、静電容量センサでは手袋をしたままの利用ができない。また、物理ボタンではボタンの押し込み動作によって外耳への押し込みが生じるため身体負担がある。その他に、これらのセンサ・ボタン押下によるデバイス操作手法は、押下回数や押下時間を変えることでコマンドの指定を行うが、指定できるコマンドに限りがある。そのため、市販されている AirPods [3] などのヒアラブルデバイスの多くは、1~3回の押下回数と長押し、上下スワイプの6種類程度のコマンドの利用に限られている。したがって、例えば音楽鑑賞中に受信メッセージを読み上げる機能や、時刻/予定を確認する機能を実装するために

は、タッチ入力とは異なるデバイス操作手法が追加で必要となる。

ヒアラブルデバイスのための操作手法は、主に手を使わない入力（ハンズフリー入力）と手を使って行う入力（ハンズ入力）にわかれる。ハンズフリー入力手法は、音声入力手法や慣性/気圧/筋電/赤外線/音響センシング [1, 2, 4, 9, 10] を用いた頭部ジェスチャ手法などがこれまで研究されてきた。音声入力の利用は公共の場での発声や騒音環境下での認識精度の低下などの課題がある。頭部ジェスチャの利用は、頭を振るなどの大きな頭部の動きを伴うジェスチャは周囲に与える印象の懸念や視線が外れてしまう危険性の課題がある。一方、顎をずらすなどの小さな頭部の動きで実現できるジェスチャは日常生活で行う動作と混同する課題がある。ハンズ入力手法ではヘッドホンに回路を追加することでデバイスのタップ検知を行う手法 [13] や、デバイス内蔵マイクを用いて頬や耳に触れながら行うスワイプジェスチャを認識する手法 [12] がある。前者の手法は追加の回路が必要であるという課題があり、後者の手法はマスクなどの装着中に一部のジェスチャが利用できない。

本研究ではヒアラブルデバイスが音響信号を再生する時に発生する外部へと漏れる音響信号（音漏れ信号）に着目し、装着者の手形状をジェスチャとして認識する手法を開発した。具体的には、ヒアラブルデバイスから周波数が遷移するチャープ信号を再生することで発生する音漏れ信号を、デバイス外側に搭載されたマイクにて録音する。この時、ユーザがデバイスを手で覆うことにより音漏れ信号は手と耳介で形成された空間内で反響する。我々はこの反響音を耳介反響音と定義する。この時、手の覆い方を変えることで得られる耳介反響音が変化すると考えられる。この変化パターンを機械学習手法を用い

* 筑波大学

† 北海道大学

て認識することで手の覆い方をジェスチャとして利用する。提案手法はヒアラブルデバイスに一般的に搭載されているスピーカと外側マイクを利用する点で実装コストが抑えられる。本研究では提案手法の実現可能性を検討するためにプロトタイプデバイスにてデータの収集を行い、ジェスチャ認識性能の評価を行った。本研究の成果を以下にまとめる。

- ヒアラブルデバイスのスピーカと外側マイクで実装可能な手形状認識手法を提案した。
- 5名の参加者データより、5種類の手形状を92.7% (F値)の精度で認識できることを示した。

2 関連研究

ヒアラブルデバイスの操作を目的としたジェスチャ認識手法は、ハンズフリー/ハンズ入力手法にわかれる。

2.1 ハンズフリー入力手法

ハンズフリー入力手法は、主に顔の動きや表情の変化などを頭部ジェスチャとしてデバイスを操作する。Ando [2]らは顔関連の動作時に生じる外耳道内部の気圧変化を気圧センサで取得し、ユーザごとに11種類の顔関連の動作を87.6%の精度で認識することに成功している。Matthiesら[9]は電極を貼付したイヤピースを用いて外耳道内部から筋肉の動きを読み取り5種類の頭部状態を精度90.0% (座位状態), 85.2% (歩行状態)で認識することに成功している。Taniguchiら[10]はLEDとフォトトランジスタを用いて舌の特定の動きを外耳道の変形から認識している。また、音楽プレイヤーの操作を対象にユーザビリティの調査も行っている。Bedriら[4]は近接センサを用いて外耳道の変形を読み取り、心拍数、舌・顎の動作、まばたきを認識している。製品では、BragiのDashProが慣性センサを用いて頭部ジェスチャを認識し、電話の応答や音楽プレイヤーの曲変更などの簡単なハンズフリー操作を実現している。Amesakaら[1]は外耳道内部の音を取得できるマイクを用いて、首、顎、顔の状態(頭部状態)にともなって外耳道が変形することに着目し、スピーカから流した測定信号をもとに、外耳道伝達関数を測定、解析することで頭部状態を認識している。

これらの手法は、ハンズフリーでジェスチャが行えるというメリットがある。一方で、頭を振るなどの大きな動きを必要とするジェスチャは視線が外れる危険性や公共の場での利用が憚られるという課題がある。顎をずらすなどの小さな頭部の動きで実現できるジェスチャは日常生活で行う動作と混同する課題がある。また、これらの手法は追加のセンサや内向きマイクなどが必要のためデバイスコストや導入コストなどの問題も生じやすい。提案手法が目指

すハンズ入力手法は片手を使う必要があるが、視認性や社会的受容性が高い動作でデバイス操作が可能となる。また、提案手法はヒアラブルデバイスに一般的に搭載されているスピーカと外側マイクで実装可能である。

2.2 ハンズ入力手法

ハンズ入力手法に関する研究では、真鍋ら[13]が市販されているヘッドホンに簡単な回路を組み合わせることでヘッドホンのタップ動作を認識することに成功している。また、菊池ら[14]は反射型リフレクタを用いて5つの方向ジェスチャの認識に成功している。Xiuら[12]はデバイス内蔵マイクを用いて頬や耳上で行うスワイプジェスチャ時に発生する音を認識し、8種類のジェスチャを95.3%の精度で認識することに成功している。

真鍋らや菊池らの手法はデバイスへの回路の追加やセンサを追加する必要があり、追加のデバイスコストが生じる。また、Xiuらの手法はマスクの装着中に頬上のジェスチャなどの一部が利用できないという課題がある。提案手法はヒアラブルデバイスで一般的に搭載されているスピーカと外側マイクで実装可能であり、デバイスコストを抑えることができる。また、ユーザはマスクや眼鏡を装着したまま利用できる。

3 提案手法

提案手法の全体の流れを図1にまとめた。ユーザはデバイスを装着し、耳介を手で覆う。この状態でデバイスから測定信号を再生することでデバイスからの信号が耳介側へと漏れ、手と耳介で形成された空間内で反響する。この反響音をデバイスの外側に搭載されたマイクで録音する。得られる耳介反響音はユーザの手の形状によって変化する。これらの手形状の違いをジェスチャとして利用することでデバイスの操作手法に用いる。認識システムは得られた反響音の特徴量を抽出し、機械学習を行うことで手形状認識モデルを作成する。

3.1 特徴量 & 機械学習モデル

取得した音響信号からジェスチャの特徴を示す特徴量の抽出を行う。本研究ではフーリエ変換による周波数成分(FFT: Fast Fourier Transform)/メル周波数ケプストラム係数(MFCC: Mel-Frequency Cepstrum Coefficients)/線形周波数ケプストラム係数(LFCC: Linear-Frequency Cepstrum Coefficients)[8]の3種類から最も精度の高かったMFCCを特徴量に用いる。選定過程は5.1節に後述する。機械学習モデルには勾配ブースティング決定木アルゴリズム(GBM: Gradient Boosting Machine)を用いる。分類モデルを用いて現在の耳介反響音から手形状の予測を行う。

音漏れ信号を用いたヒアラブルデバイス向け手形状認識手法の検討

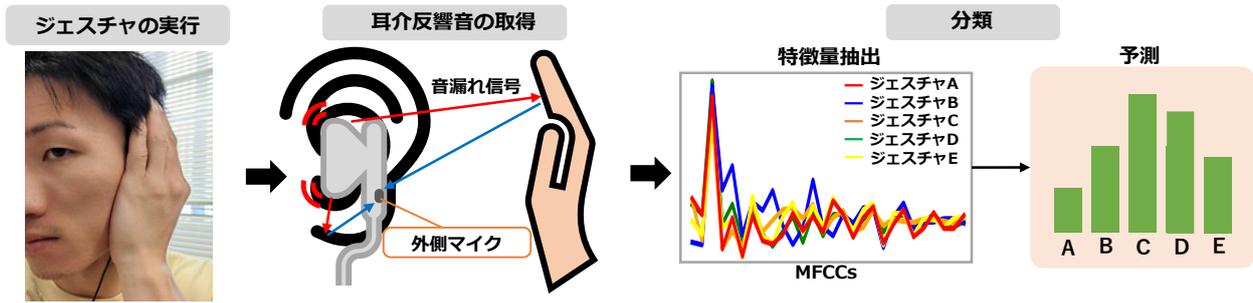


図 1. 提案手法の流れ.

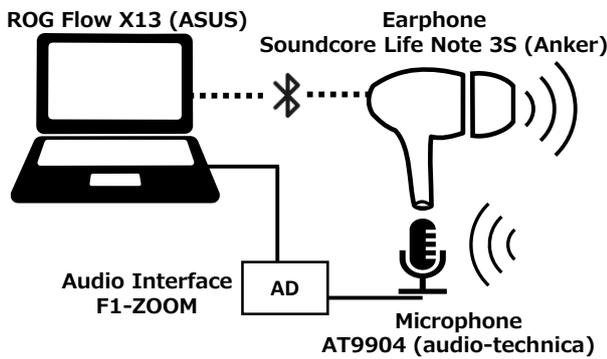


図 2. デバイス構成.



図 3. 実験デバイス.

4 実装

4.1 ハードウェア

提案手法はデバイスの外側マイクを利用することを想定しているが、市販のイヤホン型デバイスに搭載されているマイクはローパスフィルタが内部的にかけられており、高周波帯域の周波数の利用が制限されている。提案手法の本来の性能を調査するために、本実験では市販のイヤホンに別のマイクを固定する。また、予備調査よりカナル型イヤホンの音漏れは小さいことがわかったため、今回の実験ではインナーイヤ型イヤホンを用いる。デバイス構成を図2にまとめた。図3に実験に用いたデバイスを示す。本デバイスでは、イヤホンにAnker社のSoundcore Life Note 3S、マイクにaudio-technica社のAT9904を用いた。マイクから取得した耳介反響音のAD変換にはZOOM社のF1オーディオインタフェースを使用し、PC(ASUS社: ROG Flow X13)へとデータを保存した。

4.2 ソフトウェア

本研究で使用するチャープ信号は0から22kHzのアップチャープ信号(4,096サンプル)を用いた。信号の再生/録音のサンプリングレートは44.1kHzである。信号の再生はPCとBluetooth接続したイヤホンよりRealtek(R) Audioドライバを用いて再生

する。信号の音量は音漏れを発生させるために通常より大きく設定する必要がある。今回の信号音量は外耳形状を模した模型と騒音計(SANWA SUPPLY CHESD1)を用いて測定したところ、鼓膜部分にて約80dBであった。これは地下鉄車内や航空機内と同程度の騒音レベルであり、短時間での使用を想定すると聴覚健康上問題ないといえる[11]。MFCC抽出のためのフーリエ変換のウィンドウサイズは16,384サンプルとし、8,192サンプル分のスライドを行った。また、30分割のメルフィルタバンクを用いてMFCCを30次元分抽出し、直流成分である1次元目を除いた。機械学習では訓練データを5分割の交差検証を行い、F値が最も高いパラメータをハイパーパラメータとした。データの測定、特徴量抽出、機械学習のプログラムはPython3.9で実装した。

5 評価実験

本実験では5名(20~27歳)の参加者から耳介反響音を取得した。手形状の種類は図4で示す5種類である。実験は研究室室内にて座位状態で行った。研究室の空調の電源は切っていない。実験参加者は測定デバイスを図5のように右耳に装着し、右手で手形状を再現してもらう。その間に、チャープ信号を再生し、耳介反響音を5秒間録音する。これを5種類のすべての手形状で測定したものを1ラウンドと



a. 覆う（空間あり） b. 覆う（空間なし） c. 手前を覆う



d. 後ろを覆う e. ホバー

図 4. 評価した手形状.



図 5. 実験デバイス装着時の様子.

する。ラウンド終了後にデバイスの脱着を挟み、合計で6ラウンド分のデータを取得した。参加者はマスクを装着し、眼鏡装着者は装着したまま行った。

5.1 特徴量と認識精度の関係

提案手法に有効な特徴量の選定を行う。音響信号の反響・伝播を解析し、物体・人体の状態などを認識するアクティブ音響センシングで一般的に利用される特徴量のFFT/MFCC/LFCCの3種類の特徴量を調査した。各特徴量の抽出に用いられるウィンドウサイズは16,384サンプルとし、8,192サンプル分のスライドを行った。そのため、FFTで得られる周波数成分は8,192次元となるがデータ数に対して次元数が大きい。したがって、本研究では10/40次元ごとの間引きを行い、819/204次元まで次元圧縮を行った。また、MFCC/LFCCの抽出に用いられるフィルタバンクの分割数は30分割と50分割で調査する。なお、直流成分である1次元目を除くため次元数はそれぞれ29次元と49次元となる。

取得した5名のデータを用いて特徴量ごとの精

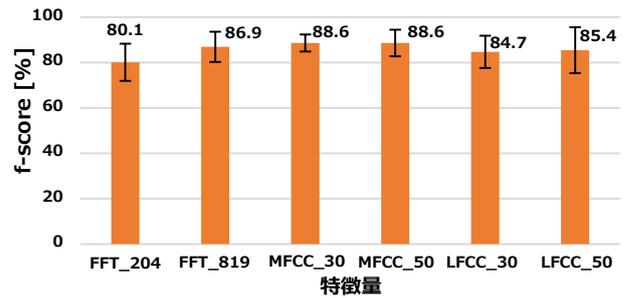


図 6. 特徴量ごとの認識精度.

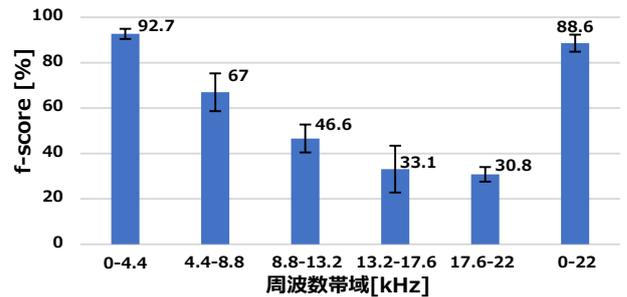


図 7. 利用周波数帯域ごとの認識精度.

度の差異を評価する。精度の調査には参加者ごとのデータを用いて個人ごとの分類モデルを作成するleave-one-round-out 交差検証を行った。図6に各特徴量での認識精度 (F 値) の平均をまとめた。実験の結果、MFCCの2パターン共にF値が88.6%で最も精度が高かった。次元数は少ないほうが機械学習の効率が良くなるため、以下の調査では30分割フィルタバンクによるMFCCを特徴量として用いる。

5.2 利用周波数帯域と認識精度の関係

実験で利用したチャープ信号の帯域は22kHz分である。得られたデータより認識性能に有効な帯域を調査するために全体の20%にあたる4.4kHzごとに信号を帯域制限した場合の各認識精度を調査した。精度の調査には参加者ごとのデータを用いて個人ごとの分類モデルを作成するleave-one-round-out 交差検証を行った。各参加者の認識精度の全体平均を周波数帯域ごとに図7にまとめた。実験の結果、0から4.4kHzまでの周波数帯域を利用した特徴量による認識精度が最も高く、F値は92.7%であった。図8はジェスチャの予測分布を示す混同行列である。最も認識精度の高かった手形状は「覆う（空間なし）」で、最も認識精度の低い手形状は「覆う（空間あり）」であった。

正解ラベル	予測ラベル				f-score[%]	
	覆う (空間あり)	覆う (空間なし)	手前を覆う	後ろを覆う		
覆う (空間あり)	1596	0	92	172	0	89.0
覆う (空間なし)	48	1750	0	0	62	96.9
手前を覆う	77	0	1696	67	20	91.5
後ろを覆う	6	0	42	1788	24	90.9
ホバー	0	2	18	46	1794	95.4
						92.7(平均)

図 8. 混同行列 (0-4.4kHz).

5.3 一般分類モデルの評価

上記の実験により最も精度の高かった0から4.4 kHzの帯域の信号を用いて、一般分類モデルの評価を行う。テストデータ以外の参加者のデータを訓練データとして作成した分類モデルでテストデータを評価する leave-one-user-out 交差検証を行った。調査の結果、各参加者の F 値の平均は 65.7% ($SD=11.7$) であった。個人分類モデルの認識精度に比べて、27ポイントの精度低下が確認された。

6 議論および今後の課題

6.1 ジェスチャ認識性能

5.2節より、0から4.4 kHzに帯域制限したデータによる認識精度が最も高いことがわかった。一方、高周波の信号帯域の認識精度が低いことがわかった。実験参加者1名の各手形状の周波数特性と MFCC の t-SNE 分布を図9にまとめた。周波数特性から分かるように、4 kHz 付近までの周波数応答が顕著であり、かつ手形状ごとの特性の変化も大きい。一方で、15 kHz 付近以降の周波数応答は著しく低く、手形状ごとの変化も小さい。周波数応答の不均衡の原因は、信号の減衰率やマイク特性などが影響していると考えられる。今後は、高周波帯域の信号音量を上げることで、高周波数帯域の応答を高くする改善を行う。高周波帯域の応答改善によって、認識性能が向上するか調査する。また、特徴量に用いた MFCC は、人間の聴覚の特徴である低周波帯域の音に敏感な聴覚特性に合わせたフィルタ設計が行われている。したがって、高周波帯域の音響データによる認識には不適であったと考えられる。今後は、周波数帯域に依存しない特徴量を用いた調査を進める予定である。これらの調査をもとに、信号の最適な周波数遷移帯域や周波数ごとの音量を決定する必要がある。

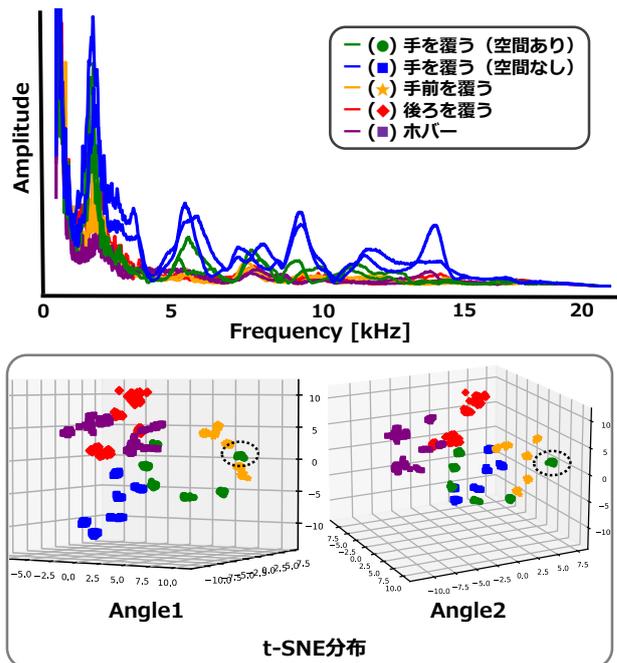


図 9. 各手形状の周波数特性 (上) と t-SNE 分布 (下).

各手形状の認識精度について、最も認識精度の低い形状は「覆う (空間あり)」であったが、「手前/後ろを覆う」のそれぞれと混同する傾向にあった (図8)。図9の t-SNE 分布からも、黒破線で囲まれた「手を覆う (空間あり)」のデータ群は同ラベルのデータ群よりも「手前を覆う」のデータ群に近いことがわかる。これらの手形状は共に手の形は同じで、開放する部分のみが変わる。そのため、開放部分が小さかった場合、得られる周波数特性が似てしまうと考えられる。今回の実験では、実験参加者に手前や後ろを覆う時の開放部分の幅などを指定しなかった。今後は、開放部分の幅と誤認識率の関係を調査することで、適切な手形状の指示を行うなどの改善を行う予定である。

6.2 ジェスチャ検出/ユーザビリティ

提案手法は可聴域の信号を利用する必要があり、常に信号を再生し続けるのはユーザビリティを大きく損なう。また、ジェスチャ認識を常に行う場合、生活動作をジェスチャとして誤認識し、意図しないジェスチャが発生する問題もある。したがって、提案手法を実環境で利用する場合、ユーザが手形状をジェスチャとして行ったことを検出し、ジェスチャ認識モードに認識システムを切り替える必要がある。ジェスチャの検出には、例えば搭載されているタッチセンサをトリガーとする方法などが考えられ、これは実装が容易である。しかし、タッチセンサが必要なら入力コマンドを一つ減らすことになる。提案手法を応用することで、タッチセンサなどを利

用せずに、ジェスチャ検出を行う方法も今後調査する必要がある。例えば、音楽などのユーザが気になりにくい信号や超音波帯域（18 kHz–22 kHz）のみのチャープ信号を利用することで、耳介反響音によるジェスチャ検出を検討している。今後は、これらの調査も行っていく予定である。

また、今回の実験では信号の音量が鼓膜部分でおよそ 80 dB であった。信号音量は大きいほど、音漏れが大きくなり得られる周波数応答も良好になるが、ユーザへの聴覚負担も大きくなる。一般的に、85 dB 以上の音を長時間聴き続けると、騒音性難聴のリスクが高まるとされている [11]。今後は、信号音量と認識精度の関係やアンケートによる聴覚への負担などを調査し、最適な信号音量を決める必要がある。また、音楽再生時に提案手法を利用するためには、音楽再生を中断する必要がある。ユーザビリティの低下につながる。例えば、測定信号の再生帯域を超音波帯域のみにする手法が考えられるが、超音波信号の再生/録音に対応したデバイスが必要である。他には、ユーザが聴いている音楽の応答を利用して周波数特性を測定する手法 [7] などもある。今後は、これらの手法を提案手法に組み込むことで音楽再生時にも利用可能な手法に改善する予定である。

6.3 実用環境/オープンイヤードバイスの調査

本研究は提案手法の実現可能性の調査が目的であったため、安定環境（静かな部屋/座位状態）で収集したデータを用いた基礎的な性能調査に留まっている。今後の実利用を考慮すると、ノイズ環境やモバイル環境での調査を進めていく必要がある。また、本実験はインイヤータタイプのデバイスによる実験のみであったが、オープンイヤータタイプ [5] のデバイスでも耳介反響音は取得できると考えており、このようなデバイスタイプでは信号音量も小さくできる可能性が高い。今後は、このようなデバイスタイプの手形状認識性能も調査していく予定である。

6.4 認識モデルの一般化と個人認証への応用

5.3 節より、一般分類モデルによる認識精度は個人分類モデルに比べて 27 ポイントも低下することがわかった。これは参加者ごとに耳介や手のサイズ、形状が大きく異なるため、得られる音響データも異なったことが原因だと考えられる。一般分類モデルの精度を向上させるためには、個人の耳介/手形状に依存しない特徴量を見つけ出す特徴量エンジニアリングなどの認識アルゴリズムの改善が考えられる。また、動きや手の形状変化を伴うジェスチャを採用することでドップラー効果による周波数変化を利用し、個人に依存しない手の動作を認識に利用する手法などが考えられる。

一方で、これらの個人の差異を利用することで個人認証が実現できる可能性がある。ヒアラブルデバ

イスによる個人認証手法も多くの研究 [6] が進められており、デバイスから秘匿性の高い情報にアクセスするための認証や決済アプリケーションなどの応用が考えられる。我々は、現在保持している 5 名の音響データより「手を覆う（空間あり）」のデータを用いて簡易的な個人識別モデルを作成した。その結果、5 名の参加者の個人識別率は 94.5% であることが確認された。したがって、今回我々が測定した耳介反響音は個人識別/認証システムへの応用も期待できる。

7 まとめ

本研究ではイヤホン型デバイスの音漏れ信号を利用した手形状認識手法を提案し、認識システムの実装を行った。利用する特徴量や周波数帯域と提案手法の認識性能の関係などを調査した。実験の結果、5 名の参加者による 5 種類の手形状認識において 0 から 4.4 kHz の MFCC を用いることで 92.7% (F 値) の認識精度が得られることを確認した。

謝辞

本研究は JSPS 科研費 21J10706/18H04104, JST さきがけ JPMJPR2138 の助成を受けたものです。

参考文献

- [1] T. Amesaka, H. Watanabe, and M. Sugimoto. Facial Expression Recognition Using Ear Canal Transfer Function. In *Proceedings of the 23rd International Symposium on Wearable Computers*, ISWC '19, pp. 1–9. ACM, 2019.
- [2] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi. CanalSense: Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 679–689. ACM, 2017.
- [3] Apple Inc. AirPods, (2022.09.22. accessed). <https://www.apple.com/jp/airpods-pro/>.
- [4] A. Bedri, D. Byrd, P. Presti, H. Sahni, Z. Gue, and T. Starner. Stick It in Your Ear: Building an In-ear Jaw Movement Sensor. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '15 Adjunct, pp. 1333–1338. ACM, 2015.
- [5] BOSE. WHAT ARE OPEN-EAR HEADPHONES?, (2022.09.22. accessed). https://www.bose.com/en_us/better_with_bose/open-ear-headphones.html.
- [6] Y. Gao, W. Wang, V. V. Phoha, W. Sun, and Z. Jin. EarEcho: Using Ear Canal Echo for Wearable Authentication. *Proceedings of*

the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 3(3), 2019.

- [7] H. Kawahara and K. Yatabe. Safeguarding test signals for acoustic measurement using arbitrary sounds: Measuring impulse response by playing music. *Acoustical Science and Technology*, 43(3), 2022.
- [8] H. Lei and E. L. Gonzalo. Mel, linear, and antimer frequency cepstral coefficients in broad phonetic regions for telephone speaker recognition. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association*, INTERSPEECH '09, pp. 2323–2326. ISCA, 2009.
- [9] D. J. C. Matthies, B. A. Strecker, and B. Urban. EarFieldSensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input Through Facial Expressions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 1911–1922. ACM, 2017.
- [10] K. Taniguchi, H. Kondo, M. Kurosawa, and A. Nishikawa. Earable TEMPO: A Novel, Hands-Free Input Device that Uses the Movement of the Tongue Measured with a Wearable Ear Sensor. *Sensors*, 18(3), 2018.
- [11] World Health Organization. Safe Listening Devices and Systems, 2019.
- [12] X. Xu, H. Shi, X. Yi, W. Liu, Y. Yan, Y. Shi, A. Mariakakis, J. Mankoff, and A. K. Dey. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–14. ACM, 2020.
- [13] 真鍋 宏幸, 福本 雅朗. Headphone Taps : 通常のヘッドホンへのタップ入力. *情報処理学会論文誌*, 55(4):1334–1343, 2014.
- [14] 菊地 高史, 正井 克俊, 杉本 麻樹, B. H. Thomas, 杉浦 裕太. EarTouch: 耳の入力インタフェース化. *日本バーチャルリアリティ学会論文誌*, 22(4):477–483, 2017.

車椅子使用者を周囲の人と遠隔介助者に繋ぐ遠隔車椅子システム

高木 瑠名* 宮藤 詩緒* Jefferson Pardomuan* 小池 英樹*

概要. 本システムは、全天周映像を用いて、遠隔介助者と車椅子使用者、車椅子周囲の第三者を繋ぐ、遠隔操作・コミュニケーションシステムである。近年車椅子のニーズが高まっている中、多くの車椅子に関する研究は、完全自動運転や遠隔操作等、車椅子の操作に焦点を当てている。だが、車椅子使用者にとっては、操作だけでなく、周りにいる介助者や第三者との会話も生きるために重要な要素である。そこで我々は、全天周カメラを用いて車椅子使用者と遠隔介助者を繋ぎ、遠隔介助者が車椅子使用者とその周囲の状況を観察しながら車椅子を操作し、三者がコミュニケーションを行うシステムを提案する。本システムでは、遠隔介助者側にお椀型ディスプレイ、車椅子側に球体ディスプレイを使用することで、操作だけでなくコミュニケーションにも焦点を当てた。提案システムにより全天周理解が向上することを実証するため、技術的評価・ユーザ実験を行い、最後に考察を行った。



図 1. 遠隔車椅子システム: (左) 遠隔介助者を見る車椅子使用者, (中央) コンセプト, (右) 車椅子使用者とその周囲を観察する遠隔介助者.

1 はじめに

最近の車椅子は、重度障害者だけでなく、移動性向上を目的として使用されることも多く、高齢者においてこのニーズは顕著である。しかし、車椅子使用者の増加に伴い介助者が不足する現状では、介助者が常に隣でマンツーマンサービスを提供することは困難である。この問題を解決するため、介助者の負担を軽減する様々な研究が行われている [2, 10, 14, 22, 26, 29]。このような自動運転は技術が不完全で危険なため、現代の技術では半自動運転等を伴う、遠隔操作が有効である [7, 8, 9]。しかし、これらの研究では平面ディスプレイやヘッドマウントディスプレイ (HMD) を使用しており、全天周映像を観察する場合、一度に見える範囲が限定されてしまい、危険な操作に繋がってしまう。

車椅子使用者にとって、周囲の第三者との会話は生活や健康に不可欠な要素である。しかし、第三者に介助を頼むのは心理的負担が大きい。また、自立した車椅子使用者にとっても見知らぬ第三者との会話は性格や疲労度等に左右され、難しい問題である。通常車椅子上のコミュニケーションでは、介助者が使用者と第三者の間を繋ぐファシリテータとして、

三者間のコミュニケーションの円滑化を促進する。しかし、既存の遠隔操作の研究は、双方向の遠隔会話ではなく遠隔者による一方的な操作のみ議論されている [7, 8, 9]。また、介助者の隣で会話を行うシステムでは、介助者が常に付き添っていなければならない [11, 12, 13, 24]。これらの問題を解決するため、テレプレゼンスを利用した研究 [6, 23, 25] も存在するが、介助者の顔が見えない、介助者を第三者に提示できないという問題がある。

本論文では、車椅子使用者を周囲の人と遠隔介助者に繋ぐ遠隔車椅子システムを提案する (図 1)。介助者が安全に車椅子周辺を観察するため、Display-Bowl [17, 19] と呼ばれるお椀型ディスプレイを使用する。遠隔介助者は、お椀型ディスプレイの特徴を活用することで、車椅子の全周囲映像だけでなく、車椅子使用者の顔・視線も同時に観察し、使用者と周囲とのコミュニケーションを図る。さらに、車椅子側では、介助者の存在を車椅子上の球体ディスプレイに表示することで、車椅子使用者や第三者が遠隔介助者がそばにいることを実感でき、車椅子使用者の孤独感や社会的孤立、負担を軽減させ、周囲とのコミュニケーションに繋げる。

2 関連研究

自動運転 自動運転は介助者の負担や立会いを必要とせず、車椅子使用者が行きたいところへ連れて行く。近年、空港や病院等で車椅子の自動運転が実用化されているが、一時的に発生する興味を反映できない。これに対し、使用者の視線や頭部等の情報を利用して目的地まで誘導するシステム [2, 10, 14, 22, 26, 29] は、障害を持つ方にとって有用なシステムだが、倫理的・技術的課題が多く実用化には程遠い。また、操作にのみ焦点があり介助者がいないため、必ず必要となる周囲の第三者へのサポート要求が困難である。

* 東京工業大学 情報理工学院

表 1. 車椅子での外出に必要な要素と発言内容. (C1 と C2, W2 と C5 と C6 はそれぞれグループで行った.)

要素	W1	W2	W3	C1	C2 & C3	C4	C5 & C6
周囲の状況把握	①	③	⑥	⑨	⑫	⑮	⑱
周囲の人とのコミュニケーション	②	④	⑦	⑩	⑬	⑯	⑲
車椅子使用者の意思の尊重	-	⑤	⑧	⑪	⑭	⑰	-

遠隔操作 自動運転の欠点解消のため、人間が車椅子を遠隔操作するシステムも開発されている [7, 8, 9]. これらのシステムは、介助者がそばにいない必要がないが、平面ディスプレイや HMD を用いて車椅子周辺の環境と車椅子使用者を観測するため、車椅子側の状況を全て感知することが難しく、危険を伴う。この問題を解決するため、衝突回避による半自動運転システムが提案されているが、動的な環境下での衝突回避は非常に複雑であり [27], 特定シナリオでの実用化に留まっている。また、これら遠隔操作システムでは、遠隔操作を行う介助者の情報が提供されず、車椅子使用者の恐怖心を煽り [30], 周辺の人とのコミュニケーションに支障をきたす。

車椅子でのコミュニケーション 主に車椅子を利用する高齢者や障害を持つ方にとって、周囲とのコミュニケーションは生きるために必要な手段であり、孤独感や社会的孤立感、健康面においても重要な要素である。車椅子使用者の手 [3] や腕 [1] を拡張するシステムは、会話に必要な指や腕を活用することで、周囲とのコミュニケーションを支援する。介助者の車椅子を押すという役割をなくすシステム [11, 12, 13, 24] では、介助者が車椅子使用者の隣を歩きながら、両者が等しく会話を行うが、介助者が常に車椅子の近くにいる必要がある。ロボットを用いて車椅子の周囲や車椅子使用者と対話するシステム [6] では、ロボットが簡単な動作をするものの、人の顔が見えず使用者に不安感や孤独感を与える可能性がある。タブレット端末に遠隔介助者の顔を表示するシステム [25] は、使用者が介助者の存在を確認し安心感を得ることができるが、平面ディスプレイであるため、第三者が介助者の存在を認識できず、介助者が第三者とのコミュニケーションをサポートできない。

3 車椅子に関わる人の要求

車椅子使用者と介助者を対象に半構造化インタビューを行い、車椅子での外出で直面する課題や介助者が気をつける点について、実体験から潜在的な要求事項を引き出すことを試みた。車椅子使用者 3 名 (W1-W3) と介助経験者 6 名 (C1-C6) の計 9 名に協力してもらった。インタビューは、参加者情報を調査し、論文掲載の許可を得ることから始まり、参加者にシステムの簡単な説明を行った。その後、車椅子の遠隔操作やコミュニケーションに関する質問を行い、具体的な体験に基づく意見を求めた。各イン

タビューは 1 時間から 1 時間半程度で終了した。

インタビューから、車椅子外出に必要な要素「周囲の状況把握」「周囲の人とのコミュニケーション」「車椅子使用者の意思の尊重」を明らかにした (表 1)。

周囲の状況把握 車椅子で外出する際に最も重要なことは、安全のために周囲の物理的環境を認識することである。「(W3⑥) 誰かと会話しながら歩いている際にも、常に周囲の環境に意識を向けなければならない。このような別のことへの意識により、疲労具合が大きくなった気がする」「(C1⑨) 操作してなくても、周囲を確認して危険を察知してあげたい」

周囲の環境は危険情報だけでなく、視野を広げるためにも必要となる。車椅子に乗ると小学校低学年程の身長となり、周囲や後方が見えにくく、自分より視野の広い介助者からの提案を求める。「(W1①) 介助者に知人を見つけてもらったりして、あの人に話しかけてみたらどう? という提案をしてほしい」

これら周囲の環境認識には、ある程度の知識、経験、感覚、体力が必要だ。そのため、普段介助者のサポートを受けている車椅子使用者や使用経験の少ない使用者、W3 のような自立した車椅子使用者にとっても、介助者の支援が必要である。車椅子使用者が行いたい作業と、周囲の環境の認識を切り離すことで、本来の作業に集中することができる [30]。

周囲の人とのコミュニケーション 全ての参加者が、車椅子での外出には、他者とのコミュニケーションが不可欠であると述べた。代表的な意見として「(C1⑩) コミュニケーションは車椅子使用者にとって、生きるための手段であり、普通に生活する人と比較すると、コミュニケーションの負担は大きい」

コミュニケーションにも介助者の存在が必要だ。一人で外出できない車椅子使用者は、介助者の存在で安心感を得、見知らぬ人との会話が困難な使用者は、介助者の第一声のサポートが精神的な支えになる (C1)。介助者が遠隔地にいても、その存在が周囲との会話のファシリテータとして機能することで、社会的孤立を回避できると考えられている [5]。

この周囲とのコミュニケーションを車椅子使用者だけに負わせることは余計な負担となる。第三者とコミュニケーションをとる際、本来の目的を達成するため、車椅子使用のためのコミュニケーションの 2 つが発生する。例えば、買い物の際、欲しい物の情報 (サイズ、色等) は目的のためのコミュニケーションだが、動作の制約による情報 (財布の場所、で



図 2. 二つの全天周ディスプレイ: (左) 車椅子側の球体ディスプレイ, (右) 遠隔側のお椀型ディスプレイ.

きないことの伝達等)は車椅子使用により発生するコミュニケーションである. 前者を車椅子使用者, 後者を介助者に割り振ることで, 負担軽減に繋がる.

車椅子使用者の意思の尊重 また, 車椅子で外出する際, 介助者や第三者の行動を車椅子使用者の意思に合わせて修正すべきである. 車椅子使用者は過度な心配や助けを被ることが多く, 喋りかけられたくない場面も存在する (W3⑧, C4⑬). W3は普通に行っているだけで, 特別な目で見られることが多く, 見えなくなりたいと思うことが多いと言う.

操作に関しても, 介助者の操作は介助者の意思だけによってはいけない. 車椅子使用者の興味が湧いたときには, その意思を反映させる操作が必要である. C1の学生(車椅子使用者)からは, 行きたい場所や意思があるときに, 自分の考えと操作の齟齬がストレスになるという意見が複数出された.

4 システム設計

本システムの主な目的は, 車椅子使用者と遠隔介助者が, 離れていても隣にいるように操作やコミュニケーションを行い, 車椅子周囲の第三者とのコミュニケーションに繋げることである (図 1, 中央). しかし, 関連研究では前章で述べた3つの要求等, 遠隔での操作とコミュニケーションの両立を実現していない. そこで, 全天周映像を用いてこれら両面を実現し, 車椅子使用者と遠隔介助者のコミュニケーションに第三者を含めるシステム設計を提案する.

全天周操作とコミュニケーション 遠隔介助者側: 車椅子を操作しコミュニケーションをとるため, 遠隔介助者が車椅子周囲を観察し, 車椅子使用者の安全を確保するとともに, 話し相手となる第三者を発見する必要がある. 遠隔介助者が車椅子使用者の代わりに周囲を認識することで, 負担を軽減し, 見えない視界を拡大する. 全天周観察には, 平面ディスプレイやHMDよりも安全な周辺観察を行うことができると実証された [17], お椀型ディスプレイ (図 2, 右) を使用する. お椀型ディスプレイは, 俯瞰視点 (TDV), 一人称視点 (FPV), 疑似三人称視点 (TPV) という特徴を持ち, 前後方向の感覚を失うことなく全天周を観察できる. 特に TDV は, 地図の観察によく使

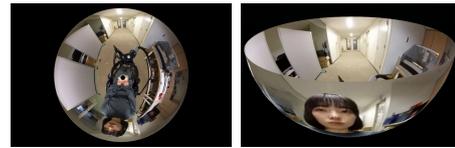


図 3. お椀型ディスプレイの見え方: (左) 俯瞰視点 (TDV), (右) 疑似三人称視点 (TPV).

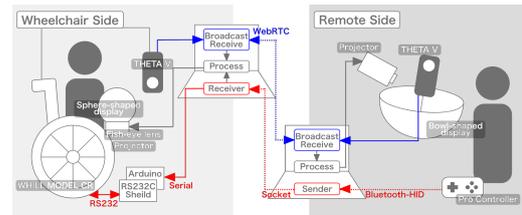


図 4. システム構成図

われる正距方位図法で, 直感的な車椅子周囲の観察に使いやすく, 遠隔操作に適している (図 3, 左). TPV は, お椀の前面と背面を同時に観察することで, 遠隔介助者が車椅子使用者と周囲を観察でき, コミュニケーションに最適な視野となる (図 3, 右).

車椅子側: 遠隔介助者の顔や一部の映像等を車椅子使用者や周囲に提示するため, 遠隔側の全天周映像も撮影する必要がある. 全天周により, 複数介助者が会話に参加できる等, 柔軟性を持てる.

第三者の介入 周囲とのコミュニケーションを容易にするため, 車椅子に球体ディスプレイ (図 2, 左) を導入する. 球体ディスプレイ [4, 15, 28] はその形状特徴から, どの位置からでも歪みのない映像を見られる. そのため, 車椅子使用者だけでなく, 周囲に情報を伝えるディスプレイとして適している. また, 球体に顔を表示することでより自然な遠隔コミュニケーションが可能になるという従来の考察 [18, 20, 21] から, 球体ディスプレイは遠隔介助者の顔を表示し存在感を示すのにも適している.

5 プロトタイプの実装

遠隔介助者側 天井等不要な情報を排除するため, 車椅子全天周映像の正距円筒図法のうち地面から 5/8 部分を取得し, プロジェクタ (BenQ HT3550) からお椀型ディスプレイ (直径 30cm) に投影することで (図 5, b), 遠隔介助者に対し, 車椅子使用者や第三者, 周囲の環境の観察を支援する (図 4, 右).

また, お椀型ディスプレイ中央の全天周カメラ (RICOH THETA V) で撮影した遠隔側環境 (図 5, b) は, USB で遠隔 PC に接続され, WebRTC 通信により車椅子 PC に配信される (図 4, 青).

車椅子側 車椅子 (WHILL MODEL CR) の操作では, コントローラ (Nintendo Switch Pro Con-



図 5. ハードウェア設計: (a) 車椅子側, (b) 遠隔側.

troller) のスティックからの操作信号を Bluetooth-HID プロファイルで遠隔 PC に認識させ、UDP 通信で車椅子 PC に送信する。そして、RS232C 通信を用いて車椅子を操作する (図 4, 赤)。

車椅子上の全天周カメラ (RICOH THETA V) で周辺と車椅子使用者を撮影し、WebRTC 通信で遠隔側 PC に映像を配信する (図 4, 左)。運転のために車椅子周辺を観察する場合、車椅子使用者が視界の妨げにならないため、カメラの位置は頭上が最適となる。しかし、操作だけでなく車椅子周囲の会話も重視するため、車椅子の背面は若干見にくくなるが、方向感覚を失わず顔を観察できるよう、カメラは顔の前方上部に設置する (図 5, a)。

また、プロジェクタ (ASUS S1) と魚眼レンズ (OPT-0.2X-37) から、遠隔側の映像を加工し、球体ディスプレイ (直径 12 cm) に投影する (図 2, 左)。

5.1 基本機能

視線共有 視線共有は他者の意図を理解する上で重要な役割を担う。安藤ら [1] は、アイトラッカーを使用して車椅子使用者の視線情報を遠隔操作者に共有した。本機能は、全天周映像から車椅子使用者の視線を推定し、コミュニケーション円滑化を目指す。

車椅子上の全天周カメラで暗黙的興味領域を検出し、遠隔ディスプレイに送信する (図 6)。遠隔介助者は、TPV を通して注視点領域 (緑丸: 注視点から一定面積) と顔を同時に観察し、無意識の興味による操作や新たな遠隔コミュニケーションに繋げる。視線位置検出方法は以下の通りである。車椅子上カメラで撮影した全天周映像を Unity の仮想球にマッピングする。球の中心に仮想カメラを配置し、車椅子使用者の顔を MediaPipe¹ により撮影する。得られた顔のランドマークから、比較的動きの少ない鼻先と両鼻孔の位置を抽出し、3次元空間に変換し、これらの値の差から方向を求める。そして、顔位置と視線方向から球面との交点 (注視点) を求め、お椀型ディスプレイに緑丸で表示させる。

対面式コミュニケーション 遠隔介助者が第三者に直接話しかけたいときは、お椀型ディスプレイを回転させ、目の前に第三者を映し出す (図 7, b,d)。そ

¹ <https://google.github.io/mediapipe/>

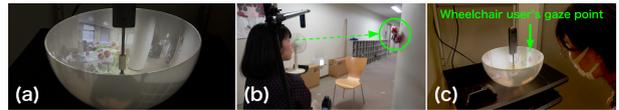


図 6. 視線共有: (a) 手前に顔、奥に視線方向を示すお椀型ディスプレイ, (b) 物体を見る車椅子使用者, (c) 車椅子使用者の視線, 顔, 周囲を見る遠隔者.

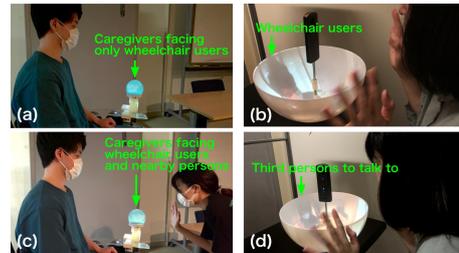


図 7. 対面式コミュニケーション: (a)(b) 遠隔者が車椅子使用者のみを見て会話をする, (c)(d) 遠隔者が車椅子使用者と第三者を見て会話をする.

れに伴い、遠隔介助者の顔が第三者の方を向き、対面した会話が可能になる (図 7, c,d)。

お椀型ディスプレイの回転角度は、底面に取り付けた導電プラスチック角度センサー (MIDORI PRECISIONS, CP-2FB) を使って計測される。取得回転角に応じて投影映像を回転させ、全天周映像がお椀型ディスプレイに張り付いた表示、球体ディスプレイ上で介助者が方向を変えた表示を可能とする。

対面コミュニケーションやアイコンタクトは、会話の中で相手を理解し心地よく感じるため必要不可欠であり、遠隔介助者と第三者に対してこの機能を実装することで、両者のコミュニケーションを円滑にできる。特に、車椅子使用者がコミュニケーションを取りたくない場合や、関わる必要がない場合に有効である。また、遠隔介助者が仲介役となることで、介助者、第三者、車椅子使用者の3者の対等コミュニケーションに有効となる (図 7, c,d)。

6 システム評価

6.1 技術評価

全天周画像からの視線推定精度を測定するため技術評価を行った。頭部に設置したセンサーで頭部角度を計測し、車椅子上のカメラで車椅子使用者含む全天周映像を撮影した。人間の頭部の可動範囲や負担を考慮し、ヨー $-60^\circ \sim 60^\circ$ 、ピッチ $-30^\circ \sim 30^\circ$ で測定し、推定角度と実際の角度の差を分析した。

結果と考察 誤差を図 8 に示す。カラースケールは誤差の大きさを可視化したもので、誤差角 0 の青色から、誤差角が大きくなるにつれ黄色に変化する。

ヨー、ピッチ共に端に向く程誤差が大きくなるが、

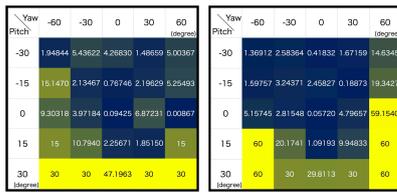


図 8. 推定角と実角度の誤差: (左) ピッチ, (右) ヨー.



図 9. 実験設計: (a) 使用する全天周映像撮影環境, (b) 平面ディスプレイ, (c) お椀型ディスプレイ.

これはカメラから顔が検出できないためである. 特に下方向は誤差が大きく, この技術的誤差を何かで補う必要がある. この問題は, お椀型ディスプレイをTPVで観察することで解決可能であると考えられる. 以下のユーザー実験でこの考察を証明する.

6.2 ユーザー実験

方向表示におけるお椀型ディスプレイの有用性を実証するため, 平面正距方位図法ディスプレイとの比較ユーザー実験を行った. 地元の大学から21歳から30歳 ($M = 24.6$, $SD = 2.59$) の10名 (女性3名, 男性7名) の, お椀型ディスプレイの経験のない参加者 (P1-P10) を募集した.

実験設計とタスク 参加者は, 車椅子使用者の視線を遠隔介助者として判断することを求められた. 車椅子の前方に1~6の数字を配置し (図 9, a), 車椅子上のカメラで画像を撮影した. 各画像に大まかな視線情報 (左右のみ正しい等) を緑丸で描画した.

これらの画像は, 被験者が観察した2種類のディスプレイに表示された. **平面正距方位図法ディスプレイ:** 全天周映像は正距方位図法により平面にマッピングされ, 中央には車椅子足元, 上部には車椅子前方, 下部には車椅子後方と車椅子使用者の顔が表示される (図 9, b). **お椀型ディスプレイ:** 図 9 (c) は, 被験者がお椀型ディスプレイを見ながらタスクを行っている様子で, 背面には車椅子正面, 前面には車椅子背面と使用者の顔が映し出される.

参加者は全天周画像を観察し, 誤差を含む視線情報 (緑丸) と車椅子使用者の顔や表情を考慮し, 正しい注視点を推定するタスクを行った.

1. タスクは, 実験者の合図で開始される.
2. 全天周画像と補助点がランダムに表示される.
3. 車椅子使用者の向いている方向を1~6の中から選択したら, キーボード入力を行う.

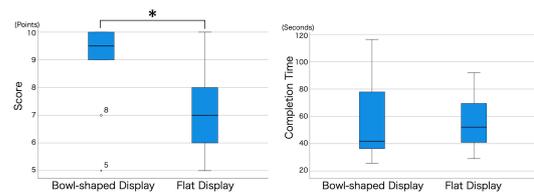


図 10. 各測定の定量的評価: (左) 視線方向の推測のスコア. 「*」は p 値が 0.05 未満であることを示す ($p < 0.05$). (右) タスク完了時間.

4. 以上の作業を 10 枚の画像で繰り返す.

1つ目のディスプレイが終了した後, 2つ目が行われる. 順番は参加者間でカウンターバランスをとった.

測定 車椅子使用者の視線推定に成功した回数を「スコア」として計測した. 正解を1点, 誤答を0点とし, 10点満点で計算した. また, 最初の画像表示から最後の画像の視線方向に反応するまでの「タスク完了時間」を測定した. さらに, 定性的なフィードバックと主観的な尺度を測定するため, 短い半構造化インタビューを実施し, タスク内での各ディスプレイの長所と短所を答えてもらった.

仮説 お椀型ディスプレイの特徴が全天周映像の方向性理解に役立つと考え, 以下の仮説が導かれた.

H1: お椀型ディスプレイの使用は, 平面ディスプレイ使用時よりも誤差を低減させられる.

H2: お椀型ディスプレイの使用は, 平面ディスプレイ使用時よりも速度を向上させられる.

定量結果 客観的指標として, スコアと実行時間を記録し, 仮説の検討に利用した.

H1: 図 10 (左) に示すように, お椀型ディスプレイ条件 ($M = 8.9$, $SD = 1.58$) の方が平面ディスプレイ条件 ($M = 7.2$, $SD = 1.47$) より高いスコアを獲得した. Shapiro-Wilk 検定の結果, お椀型ディスプレイ条件では正規性が仮定されていないことが分かった ($p = 0.002 < 0.05$). そこで, Wilcoxon Signed-rank 検定を用いて効果を確認し, 条件間で統計的に有意な差があることが分かり [$Z = -2.083$, $p = 0.042 < 0.05$], 仮説 H1 が確認された.

H2: 条件間でタスク完了時間に差があるかどうかを調べるため, paired t-test を実施し, お椀型ディスプレイ ($M = 55.25$, $SD = 28.3$) と平面ディスプレイの条件 ($M = 55.31$, $SD = 18.6$) では, 有意差がないことが分かり [$t(20) = -0.011$, $p = 0.99 > 0.05$] (図 10, 右), 仮説 H2 は棄却された.

定性結果 10人中9人が, お椀型ディスプレイの方が平面ディスプレイよりもどこを見ているか推測しやすいと述べていた. また, 表示される映像は同じだが, 「車椅子使用者の視線の高さと顔の傾きを同じ

高さで確認できた」(P2)という参加者もあり、視線の共有が起きていることが分かった。

平面ディスプレイでは大まかな方向は分かりやすいが(P3,P6,P7)、鉛直判断は難しく(P1,P2,P3,P6,P7,P9)、正しい推測が難しい(10名中9名)。

考察 技術評価で分かったように、システムのシンプル化のため1台の全天周カメラを使用しており、視線や顔の向きを推定できない場合も存在する。このような際に、お椀型ディスプレイを使用すると、視線推定スコアが向上することから、より正確な視線推定が可能となる。特にインタビューにより、お椀型ディスプレイは平面ディスプレイよりも空間の上下方向が把握しやすく(10名中4名)、視線を直感的に把握しやすい(10名中8名)ことが分かった。

ユーザー実験の結果、平面ディスプレイとお椀型ディスプレイには、視線推定に時間差は見られなかった。しかし、インタビューから我々の主張を裏付ける結果が得られた。平面正距方位図法ディスプレイは、視線の移動が少なく画像全体を一度に把握できるが(P6)、方向性が定まらない(P3,P6,P7,P9)。一方、お椀型ディスプレイは前面と背面を同時に見る必要があり、初めて使うには難しいが(P3,P5,P6,P8)、ユーザーは上下などの正しい方向感覚や車椅子使用者を直感的に認識することができる(P1,P6,P7,P9)。

7 考察

本研究では、新しい遠隔全天周操作・コミュニケーション車椅子システムを提案し、既存研究で実現されなかった、遠隔での会話支援も実現し、車椅子外出手段のオプションを増やす貢献をした。本章では、インタビューで明らかになった3つの要求に関して、本研究の特徴や利点、制約等について考察する。

周囲の状況把握 本システムは、全天周カメラとお椀型ディスプレイを用いて全天周観察を実現し、車椅子の遠隔操作・コミュニケーションに有用となるシステムとなった。車椅子に全天周カメラを搭載した研究はいくつかあるが[7, 8, 9]、操作とコミュニケーション両面に最適なカメラ位置を考慮し、1台のカメラで周囲環境、車椅子使用者、第三者を観察できる。ユーザー実験の結果、車椅子遠隔操作時の周囲の一貫観察にはTDV、正確な方向認識にはTPVが適切であることが分かった。お椀型ディスプレイを採用することでこれら両面を実現し、車椅子外出に必要な「周囲の状況把握」のサポートを行う。

本研究では、車椅子周囲の簡単な観察のため、1台の全天周カメラを設置したが、使用者の体が周囲観察の妨げになっている。今後はより視野を広げるため、ドライブレコーダーのように背面に180度カメラを搭載したり、1台の全天周カメラを可動式にする[16]等の改善を行いたいと考えている。

周囲の人とのコミュニケーション 2つのディスプレイで「周囲の人とのコミュニケーション」に寄与する。ユーザー実験の結果、お椀型ディスプレイは周囲位置の正確な理解を促し、第三者の位置等の正確な状況を直感的に感じ取れるため、遠隔側にお椀型ディスプレイを搭載することで、話しかける相手を決断したり、直接的に介助情報を提示したりと、第三者とのコミュニケーションに有効となる。さらに、全周囲から観察できる球体ディスプレイを車椅子に設置し、車椅子使用者だけでなく、周囲の第三者にも遠隔介助者の存在や情報を提示することで、従来研究では実現できなかった、第三者を含めた三角形のコミュニケーションの実現を可能にする。

本研究では、車椅子外出の要求を探るため、車椅子に関わる人々にインタビューを行い、システムの視線推定に対するユーザー実験を行いフィードバックを得た。今後は、実際にターゲットとするユーザーが車椅子で遠隔外出するシナリオを想定し、第三者の介入によるコミュニケーションの変化や行動の変化に着目したフィールド実験を行いたい。

車椅子使用者の意思の尊重 ユーザー実験では、お椀型ディスプレイにより車椅子使用者を直感的に見て、使用者と同一の視点で観察できることが実証された。P2が語ったように、前面に使用者が投影されることで、同じ高さで周囲を観察することができる。これは、介助者に求められる要素でもある(C4)。このように、車椅子使用者の興味を正確に認識できるため、遠隔介助者は意思に気付きやすくなり、「車椅子使用者の意思の尊重」を考慮しやすくなる。

8 むすびと展望

全天周映像を用いて、車椅子使用者と近くにいる第三者や遠隔介助者を繋いでサポートを行う、新しい遠隔操作・コミュニケーションシステムを提案し、実装した。遠隔介助者はお椀型ディスプレイで車椅子使用者とその周辺を観察し、車椅子側では球体ディスプレイで介助者の存在を車椅子使用者や第三者が認識し、コミュニケーションの促進となりうる。そして、視線情報や対面コミュニケーションの基本機能を実装し、技術評価とユーザー実験を行いシステムの有効性を評価し、考察を述べた。

今回実装したプロトタイプでは、無線通信と投影(583 ms)、スムーズングのための回転取得に遅延が発生している。しかし、遅延は操作性や会話の弊害となってしまうため、屋外での実装に5G通信、回転のノイズ軽減フィルターにより遅延は軽減できる。

謝辞

本研究はJSPS科研費JP20K23321の助成を受けたものである。

参考文献

- [1] R. Ando, K. Minamizawa, and M. Y. Saraiji. SlideFusion: Surrogacy Wheelchair with Implicit Eyegaze Modality Sharing. In *ACM SIGGRAPH 2020 Emerging Technologies*, SIGGRAPH '20, New York, NY, USA, 2020. ACM.
- [2] J. M. Araujo, G. Zhang, J. P. P. Hansen, and S. Puthusserypady. Exploring Eye-Gaze Wheelchair Control. In *ACM Symposium on Eye Tracking Research and Applications*, ETRA '20 Adjunct, New York, NY, USA, 2020. ACM.
- [3] Y. Asai, Y. Ueda, R. Enomoto, D. Iwai, and K. Sato. ExtendedHand on Wheelchair. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16 Adjunct, p. 147–148, New York, NY, USA, 2016. ACM.
- [4] H. Benko, A. D. Wilson, and R. Balakrishnan. Sphere: Multi-Touch Interactions on a Spherical Display. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, p. 77–86, New York, NY, USA, 2008. ACM.
- [5] P. Boissy, H. Corriveau, F. Michaud, D. Labonte, and M.-P. Royer. A qualitative study of in-home robotic telepresence for home care of community-living elderly subjects. *Journal of telemedicine and telecare*, 13:79–84, 02 2007.
- [6] H. Fukuda, K. Yamazaki, A. Yamazaki, Y. Saito, E. Iiyama, S. Yamazaki, Y. Kobayashi, Y. Kuno, and K. Ikeda. Enhancing Multiparty Cooperative Movements: A Robotic Wheelchair That Assists in Predicting Next Actions. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, ICMI '18, p. 409–417, New York, NY, USA, 2018. ACM.
- [7] R. Gunderson, S. Smith, and B. Abbott. Applications of virtual reality technology to wheelchair remote steering systems. *Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies (ECDVRAT'96)*, Maidenhead, UK, pp. 8–10, 01 1996.
- [8] S. Hashizume, I. Suzuki, K. Takazawa, R. Sasaki, Y. Hashimoto, and Y. Ochiai. Telewheelchair: A Demonstration of the Intelligent Electric Wheelchair System towards Human-Machine. In *SIGGRAPH Asia 2017 Emerging Technologies*, SA '17, New York, NY, USA, 2017. ACM.
- [9] S. Hashizume, I. Suzuki, K. Takazawa, R. Sasaki, and Y. Ochiai. Telewheelchair: The Remote Controllable Electric Wheelchair System Combined Human and Machine Intelligence. In *Proceedings of the 9th Augmented Human International Conference*, AH '18, New York, NY, USA, 2018. ACM.
- [10] J. B. Hinkel. Head-Guided Wheelchair Control System. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '10, p. 313–314, New York, NY, USA, 2010. ACM.
- [11] Y. Kobayashi, Y. Kinpara, E. Takano, Y. Kuno, K. Yamazaki, and A. Yamazaki. Robotic Wheelchair Moving with Caregiver Collaboratively Depending on Circumstances. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, p. 2239–2244, New York, NY, USA, 2011. ACM.
- [12] Y. Kobayashi, Y. Kinpara, E. Takano, Y. Kuno, K. Yamazaki, and A. Yamazaki. A Wheelchair Which Can Automatically Move alongside a Caregiver. In *Proceedings of the 6th International Conference on Human-Robot Interaction*, HRI '11, p. 407–408, New York, NY, USA, 2011. ACM.
- [13] Y. Kobayashi, R. Suzuki, Y. Sato, M. Arai, Y. Kuno, A. Yamazaki, and K. Yamazaki. Robotic Wheelchair Easy to Move and Communicate with Companions. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, p. 3079–3082, New York, NY, USA, 2013. ACM.
- [14] M. Kutbi, X. Du, Y. Chang, B. Sun, N. Agadakov, H. Li, G. Hua, and P. Mordohai. Usability Studies of an Egocentric Vision-Based Robotic Wheelchair. *J. Hum.-Robot Interact.*, 10(1), Jul 2020.
- [15] Z. Li, S. Miyafuji, E. Wu, H. Kuzuoka, N. Yamashita, and H. Koike. OmniGlobe: An Interactive I/O System For Symmetric 360-Degree Video Communication. In *Proceedings of the 2019 on Designing Interactive Systems Conference*, DIS '19, p. 1427–1438, New York, NY, USA, 2019. ACM.
- [16] A. Matsuda, T. Miyaki, and J. Rekimoto. ScalableBody: A Telepresence Robot That Supports Face Position Matching Using a Vertical Actuator. In *Proceedings of the 8th Augmented Human International Conference*, AH '17, New York, NY, USA, 2017. ACM.
- [17] S. Miyafuji, F. Perteneder, T. Sato, H. Koike, and G. Klinker. A Bowl-Shaped Display for Controlling Remote Vehicles. In *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '19, New York, NY, USA, 2019. ACM.
- [18] S. Miyafuji, T. Sato, Z. Li, and H. Koike. Qoom: An Interactive Omnidirectional Ball Display. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, p. 599–609, New York, NY, USA, 2017. ACM.
- [19] S. Miyafuji, S. Toyohara, T. Sato, and H. Koike. DisplayBowl: A Bowl-Shaped Display for Omnidirectional Videos. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*, UIST '18 Adjunct, p. 99–101, New York, NY, USA, 2018. ACM.
- [20] O. Oyekoya, W. Steptoe, and A. Steed. SphereAvatar: A Situated Display to Represent a Remote Collaborator. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, p. 2551–2560, New York, NY, USA, 2012. ACM.

- [21] Y. Pan, W. Steptoe, and A. Steed. Comparing Flat and Spherical Displays in a Trust Scenario in Avatar-Mediated Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, p. 1397–1406, New York, NY, USA, 2014. ACM.
- [22] N. Sahadat, N. Sebkhi, and M. ghovanloo. Simultaneous Multimodal Access to Wheelchair and Computer for People with Tetraplegia. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction, ICMI '18*, p. 393–399, New York, NY, USA, 2018. ACM.
- [23] M. Sarabia and Y. Demiris. A Humanoid Robot Companion for Wheelchair Users. In G. Herrmann, M. J. Pearson, A. Lenz, P. Bremner, A. Spiers, and U. Leonards eds., *Social Robotics*, pp. 432–441, Cham, 2013. Springer International Publishing.
- [24] Y. Sato, M. Arai, R. Suzuki, Y. Kobayashi, Y. Kuno, K. Yamazaki, and A. Yamazaki. A maneuverable robotic wheelchair able to move adaptively with a caregiver by considering the situation. In *2013 IEEE RO-MAN*, pp. 282–287, Aug 2013.
- [25] J. Shen, B. Xu, M. Pei, and Y. Jia. A low-cost tele-presence wheelchair system. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2452–2457, 2016.
- [26] C. C. Singer and B. Hartmann. See-Thru: Towards Minimally Obstructive Eye-Controlled Wheelchair Interfaces. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS '19*, p. 459–469, New York, NY, USA, 2019. ACM.
- [27] C. Wang, A. S. Matveev, A. V. Savkin, T. N. Nguyen, and H. T. Nguyen. A collision avoidance strategy for safe autonomous navigation of an intelligent electric-powered wheelchair in dynamic uncertain environments with moving obstacles. In *2013 European Control Conference (ECC)*, pp. 4382–4387, 2013.
- [28] J. Williamson, J. Williamson, D. Sundén, and J. Bradley. Multi-Player Gaming on Spherical Displays. pp. 355–358, 04 2015.
- [29] I. Yoda, K. Sakaue, and T. Inoue. Development of Head Gesture Interface for Electric Wheelchair. In *Proceedings of the 1st International Convention on Rehabilitation Engineering & Assistive Technology: In Conjunction with 1st Tan Tock Seng Hospital Neurorehabilitation Meeting, i-CREATE '07*, p. 77–80, New York, NY, USA, 2007. ACM.
- [30] B. Zhang, G. Barbareschi, R. Ramirez Herrera, T. Carlson, and C. Holloway. Understanding Interactions for Smart Wheelchair Navigation in Crowds. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, New York, NY, USA, 2022. ACM.

未来ビジョン

本論文では、車椅子使用者と周囲の人、介助者を繋ぐ遠隔車椅子システムを提案した。車椅子使用者数増加に伴い、介助者数が少なくなり続ける今後の社会において、遠隔での操作性だけでなく、既存研究では取り組まれてこなかった遠隔コミュニケーション性を実現することで、離れていても信頼関係のある介助者、友達、家族等と共に車椅子での生活を過ごすことができるようになる。

今後は更なる技術的革新を図り、車椅子社会に貢献していきたい。遠隔ではどうしても伝わりづらい地面の凹凸や、車椅子周囲の人等の動的な周囲環境等を把握するサポート、実感を感じづらいゲームコントローラーではなく操作のリアルを伝えるコントローラー、自走と他走の最適な切り替え制御、遠隔コミュニ

ケーションの質を改善させるための遅延改善、介助者一人で車椅子複数人の対応、車椅子一人に対し介助者複数でのサポート等、更なる技術的改変を行うことで、将来の車椅子社会において、本システムを多様なシナリオで使用できるようになる。

本システムの更なる発展により、車椅子使用による潜在的な課題を解決する手助けになる。車椅子に関わる人達へ行ったインタビューを通して、車椅子を使用せざるを得ない方達は、現状のシステムや技術、社会環境での生活に適応しなければならぬため、認識能力やコミュニケーション、意思の我慢等、車椅子未使用者が抱えなくても良い負担を抱えてしまっていることが分かった。提案システムがより進歩すれば、車椅子使用者の「あたりまえ」を覆すきっかけとなり得る。

Vi-MAP: Visual SLAM と振動子を用いた視覚障害者の歩行支援システム

坂本 祐輝* 福田 大翔* 内原 正一† 門本 淳一郎* 入江 英嗣* 坂井 修一*

概要. 視覚障害者の安全かつ自由な歩行は重要な課題の一つであるが、視覚障害者が歩行の際に使用する白杖や盲導犬には、安全面やコストなどで問題がある。本研究では、コンピュータを用いた視覚障害者の歩行支援システムの提案を行う。システムは、モバイルデバイスの Visual SLAM を使った地図作成および自己位置推定と経路生成、および手持ちデバイスの振動による経路案内から構成される。システム要件は当事者へのインタビューをもとに歩行支援システムに求められる要素を整理した上で決定した。SLAM の精度を評価するための実験では、本システムの SLAM は屋内での経路案内に十分な測位精度があることがわかった。視覚障害者を対象とした歩行実験では、参加者全員が設定されたルートを歩くことが出来、本手法の有効性が確かめられた。また、伝達手法の使用感についても好意的な反応が得られ、経路伝達手法として有効であることがわかった。

1 はじめに

視覚障害者の安全かつ自由な歩行は多くの当事者が希望する重要な課題である。日本における視覚障害者数は約 31 万人であり、その中でも単独歩行が難しい 1-2 級の障害がある視覚障害者数は約 20 万人にのぼる [18]。日本では視覚障害者が外出する際に白杖または盲導犬を携帯することが義務付けられているが、これらの歩行支援にも課題が存在する。白杖は手軽に手に入る反面、探知範囲が狭く障害物や段差を察知できなかったことによる衝突事故、転落事故が数多く発生している [23]。また、盲導犬は育成に多大な費用と時間が必要とされ、視覚障害者数に対して盲導犬が不足している現状がある [24]。

コンピュータの小型化によって人が持ち運べる大きさのデバイス上で高度な計算が可能となった。特に 3 次元形状計測技術の進歩は目覚ましく、視覚障害者支援への応用として白杖や盲導犬などの既存支援に代わるもの、またそれらと組み合わせるものなど、様々なアプローチでの研究が行われている。

本研究では視覚障害者福祉協会および日本盲導犬協会と連携して歩行支援システムに求められる要素を整理した上で、視覚障害者の歩行支援システムの実装、評価を行った。システム全体として、周囲の環境からの地図作成および自己位置推定、目的地までの経路生成、生成した経路に沿った経路案内の 3 要素が整理された。また、経路伝達手法としては、聴覚に頼りすぎることなく、直感的に進行方向を提示する手法が望ましいこと、取り回しの面から小型軽量の機器を用いる手法が望ましいことが指摘された。これらを踏まえ、提案システムには、スマートフォンなどに搭載される RGB カメラ、方位センサと振動子を用い、

RGB カメラを用いた SLAM (Simultaneous Localization and Mapping) での地図作成および自己位置推定と、方位センサと振動子を用いたデバイスを身体の前方で左右に振りながら歩き、デバイスを振動させることで進行方向へ誘導する経路案内を用いる手法を採用する。

提案手法の有効性を評価するために SLAM の精度を評価する実験と、視覚障害者を対象とした歩行実験を行った。SLAM の精度を評価するための実験では、本システムの SLAM は屋内での経路案内に十分な測位精度があることがわかった。視覚障害者を対象とした歩行実験では、参加者全員が設定されたルートを歩くことが出来、本システムの有効性が確かめられた。また、参加者のフィードバックでは好意的な反応が得られ、本システムの誘導手法は視覚障害者の歩行支援システムにおける経路伝達手法として有効であることがわかった。

2 視覚障害者の歩行の現状

2.1 人間の歩行の要素と視覚障害者の歩行の困難

人間は移動の際に、自分は地図上のどこにいるのか、目的地は地図上のどこにあるのか、どの経路で目的地に向かうのか、といった事柄を把握している。これらにはナビゲーション、オリエンテーション、モビリティの 3 要素が関係している [16]。

1 つ目のナビゲーションとは地図上の自分の現在位置を見つけ、目的地まで進む経路を決定する能力である。人間は頭の中にメンタルマップを作成し、それをもとにした移動を行うが、視覚障害者は音や匂いなどを用いてメンタルマップを生成するため、初めて訪れる場所へ独力で行くことは困難である。

2 つ目のオリエンテーションとは周囲の環境から現在の自分の位置を把握し、目的地に向かって次に進む方向を見つける能力である。視覚障害者は現状

* 東京大学

† 新産業創造研究機構 (現在, JSU 支援技術研究所)

白杖または盲導犬を携帯することによってオリエンテーション達成の補助をしているが、障害物や段差を察知できなかったことによる衝突や転落事故が絶えない現状がある [23]。またそれ以外に、人間が視覚に頼らず歩行する際に意図しない方向へ曲がってしまう特性である偏軌傾向特性が挙げられる。視覚障害者は壁や点字ブロックなどの手がかりがないところでは意図しない方向へ曲がってしまう可能性があり、オリエンテーションの達成が困難である [5] [22]。

3つ目のモビリティとは障害物などを避けながら姿勢を保ったまま歩行する能力である。障害物や段差、歩行路面などの歩行中の周囲の状態によって歩行中の安全性や効率性は変化するため、環境に合わせた適切な行動を取ることが安全かつ効率的な歩行には求められる。しかし視覚に障害があるとそれらを一定の水準に保つことが難しくなり、モビリティの達成が困難である [21]。

2.2 視覚障害者の歩行支援に求められる要素

視覚障害者福祉協会および日本盲導犬協会の協力のもと、視覚障害のある当事者および職員を対象に視覚障害者の歩行での困難についてインタビューを行ったところ、障害物がない場所でまっすぐ歩くことが難しい、初めて行く場所の周辺地図がわからない、自分がどこにいるのか突然見失うことがある、といった声が挙がった。

障害物がない場所でまっすぐ歩くことが難しいという問題は、オリエンテーション達成の困難として分類できる。当事者からは周囲に壁などの手がかりがないとまっすぐ歩くことが難しいとの声や、公園や駐車場などの広く手がかりに乏しい場所では、本当に目的地の方向に歩いているのか不安になるといった声が挙がった。

また、初めて行く場所の周辺地図がわからない、自分がどこにいるのか突然見失うことがあるという問題は、ナビゲーション達成の困難として分類できる。当事者からは既存の解決策として、人に案内してもらおう、スマートフォンの地図アプリや施設の案内板などを利用するといったものが挙げられたが、人に案内してもらおうといった解決策に関しては、特にホテルの自室やトイレへの案内などは、人に行き先を聞かれない、知らない人に案内してもらおうのは不安だといった声が挙がった。また、GPSの活用が難しい屋内では既存の地図アプリの使用が難しく、視覚障害者が利用可能な形での案内システムが整っていない施設も多い問題が指摘された。

3 関連研究

3.1 視覚障害者の歩行支援技術

コンピュータを用いた視覚障害者の歩行支援に関連する技術の構成要素は以下の3つに整理される。

1. 周囲の環境のセンシング
2. 目的地までの経路生成
3. 視覚によらない情報伝達

Joao らは自立走行型ロボットを用いた歩行支援システムを提案した。ハンドル付きの自立走行型ロボットにステレオカメラ、および LiDAR を搭載し、周囲のセンシングを行い、ハンドルに取り付けられた複数の振動子の振動とともにロボットが目的地まで移動する形での経路の案内を行う。[4]。また森らは手持ちデバイスを用いた歩行支援システムを提案した。RGB、深度カメラを搭載したスマートフォンによって周囲のセンシングを行い、利用者は方位センサと振動子を搭載した携帯デバイスを身体の前方で左右に振りながら歩き、デバイスの指す方向が進む方向と一致した瞬間にデバイスが振動することにより目的地までの経路を伝達する。利用者は自由に歩きながら必要に応じてシステムの誘導を利用することができる [20]。同様に手持ちデバイスを用いた歩行支援システムとして、栗林らの研究が挙げられる。LiDAR を搭載したスマートフォンによって周囲のセンシングを行い、主に骨伝導イヤホンからの音声により目的地までの経路を伝達する [17]。

3.2 周囲の環境のセンシング手法

周囲の環境のセンシング手法として、主にカメラを用いるもの、LiDAR を用いるもの、深度カメラを用いるものが挙げられる。

カメラを用いるものとして、山崎らはカメラを用いた SLAM である Visual SLAM を、類似画像検索と並行に用いることで視覚障害者ナビゲーションを研究している [19]。また Mocanu らは、RGB カメラより取得した画像のフレーム間の関係から障害物を検知する手法を研究している [10]。

LiDAR を用いるものとして、Kayukawa らは LiDAR による 2 次元マップ生成と、ステレオカメラによる障害物検知を並行に用いる歩行支援システムを研究している [6]。

深度カメラを用いるものとして、Zhang は深度カメラを用いて障害物回避のためのローカルマップを生成する手法を研究している [15]。また Brock らは、深度センサを用いて前方の障害物を検知する手法を研究している [1]。

上述の各センシング手法についてセンサごとの特徴とその比較を表 1 に記した [25]。

表 1. 各センシング手法の比較

センサの種類	距離精度	解像度	照明条件	価格
カメラ	△	○	△	○
LiDAR	◎	△	○	×
深度カメラ	○	○	×	△

3.3 利用者への情報の伝達手法

利用者への情報の伝達手法は、主にモバイル・ウェアラブルデバイスで触覚提示するもの、モバイル・ウェアラブルデバイスで聴覚提示するもの、ロボットやドローンで情報提示するものに分類できる。

モバイル・ウェアラブルデバイスで触覚提示するものとして、Spiers らは利用者の手に変形する直方体型のデバイスを持たせ、その回転、伸縮などにより経路を伝える手法を提案した [9]。また Tsukada らや Cosgun らは、ベルトの周上に等間隔で取り付けられた振動子を用いて、振動子を振動させる位置やパターンを変化させることによって情報を伝える手法を提案した [12] [2]。

モバイル・ウェアラブルデバイスで聴覚提示するものとして、Brock らはヘッドフォンから 3 次元的な音を生成することで障害物の情報を伝える手法を提案した [1]。

ロボットやドローンで情報提示するものものとして、Wachaja らは様々なセンサを搭載した手押し車のハンドルに振動子を取り付け、その振動により障害物の情報を伝える手法を提案した [13]。また Wei らは、盲導犬を模したロボットを用いて誘導する手法を研究している [14]。

4 提案手法

4.1 手法の概要

森らの研究における視覚障害者への事前インタビューでは、聴覚に頼りすぎることなく、直感的に進行方向を提示する手法が望ましいこと、取り回しの面から小型軽量な機器を用いる手法が望ましいことが指摘された。また森らの評価実験における実験参加者からのフィードバックによれば、安心感がある、自然な動作で利用できる、白杖で点字ブロックをなぞって歩いているような感覚が得られたという声があり、白杖を用いた普通の歩行の動作と親和性の高い動きが当事者からの好意的な評価につながったと分析されている。またリアルタイム経路生成により利用者が自由に歩く方向を決めることができ、必要ときにはすぐに経路案内を直感的に受けられることも当事者から高く評価されていた [20]。一方、森らの経路案内手法の実用化における課題として、細かな方位提示が可能なことに見合った高精度の定位技術が必要であり、従来の精度では開けた公園などでの適用にとどまっていたことが挙げられる。

本研究では、ウェアラブルデバイスでの Visual SLAM 導入により当事者が自由な歩行を行うことが可能なまま、小型軽量な情報機器を用いた屋内での高精度定位の実現が可能となった。SLAM により利用者の歩行に応じた動的な地図の修正が可能となる。カメラは安価で一般に普及していること、画像は多くの情報を含み解像度が高く、特に起伏がなく

テクスチャの特徴がある箇所での使用に適していることから、建物内の屋内の定位に適していると考え、センサとしてはカメラを用いる。

また定位の高精度化に応じ、森らの手法に振動による複数の通知パターンを状況に応じて変化させる改良を行った、振動を用いる直感的インタフェースによる経路伝達手法を導入する。視覚障害者を対象としたものではないが、類似の通知手法は Pielot らの研究で“Magic wand metaphor”としてまとめられている [8]。本手法は白杖を左右に振る動きを模し、視覚障害者の経路案内に最適化したものと位置付けられる。これにより入り組んだ屋内での適切な経路案内を可能とし、視覚障害者の自力での安全かつ自由な歩行をサポートするシステムが実現した。

4.2 システムの構成

4.2.1 全体の構成

提案システムの機器は、RGB カメラを搭載した地図作成、自己位置推定用のスマートフォン A、方位センサと振動子を搭載した経路伝達用のスマートフォン B と、地図の作成、経路の計算等の処理を行う PC によって構成される。システム全体としては図 1 のように、まずスマートフォン A の内蔵 RGB カメラを用いた Visual SLAM による地図の作成、自己位置の推定を行う。次にそれを元に PC 上で目的地に向かう経路生成を行う。最後に生成された経路情報をスマートフォン B の方位センサと振動子により利用者に伝達する。システムの外観は図 2 である。

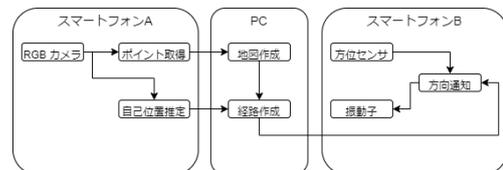


図 1. 提案システムの処理の流れ



図 2. システムの外観

4.2.2 地図作成および自己位置推定

目的地を把握し、現在地から目的地に進む経路を作成するために地図作成および自己位置推定を行う。まず Visual SLAM を用いて周辺のマッピングを行う。Visual SLAM としては特徴点ベースの Visual SLAM の一種である ORB-SLAM2 を用いる [7]。図 3 にスマートフォンの画面上での特徴点検出の様子を示す。スマートフォンの画面上にカメラからの画像が映し出されており、緑色の点でその特徴点として検出された点が描画されている。

次に、目的地に進む経路を作成するために 2 次元占有格子地図を作成する [11]。占有格子地図とは空間を格子状に分割し、各格子ごとの障害物の存在確率を格納することで表される地図である。図 4 に作成された 2 次元占有格子地図を示す。

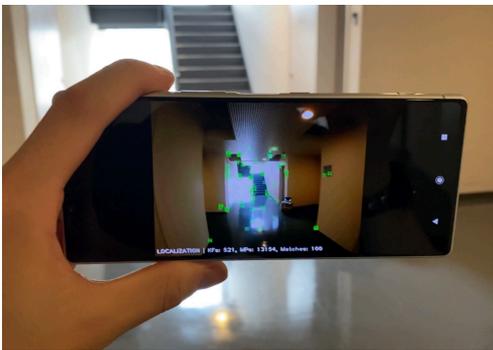


図 3. 屋内での障害物の特徴点検出の例

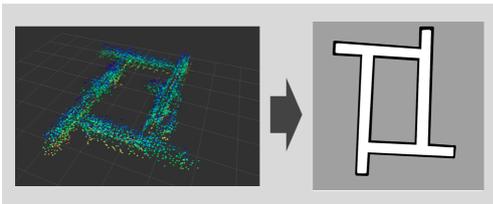


図 4. ポイントクラウド地図から作成された 2 次元占有格子地図

4.2.3 経路生成

経路生成モジュールでは作成されたマップを利用して PC 上で現在地から目的地に進む経路を作成する。ウェイポイントはゴールおよびゴールへとたどり着くまでの中継地点を指定するものであり、作成されたマップ中に指定する。図 5 に生成された直近ウェイポイントまでの経路を示す。緑色の丸で表された現在位置から、赤色の矢印で表されたウェイポイントへ緑色の経路が生成されている。

4.2.4 経路伝達

経路伝達モジュールでは生成された目的地までの経路を利用者に伝達する。方位センサと振動子を搭

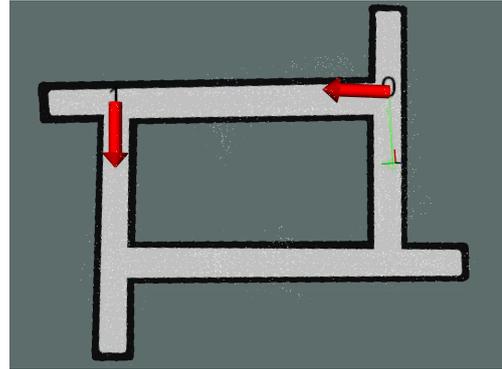


図 5. 地図から生成された直近ウェイポイントまでの経路

載した経路伝達用のスマートフォン B を手に持ち、身体の前方で左右に振りながら歩く。スマートフォン B の指す方向が進む方向と一致したときに振動させることで利用者に歩くべき方向を通知する。また障害物に衝突する危険性がある場合は利用者に警告音を出すことで停止を促す。進む方向と一致した瞬間に振動させる命令を送った場合、命令から実際に振動するまでの遅延のため進むべき方向と利用者が感じる振動した方向との間に角度のずれが生じる。これを解消するため、利用者がスマートフォン B を左右に振る際の角速度を計算し補正を行う。

また、経路上での各ウェイポイントとの距離に応じて振動の周期を 2 段階に変化させることにより、利用者に次のウェイポイントまでの距離を合わせて通知する。ウェイポイントとの距離が十分ある場合は長い振動を用いて経路の伝達を行う。ウェイポイントとの距離が近づいた場合は短い振動を用いて経路の伝達を行い、方向転換点に近いことを知らせる。ウェイポイントを通過し、次のウェイポイントへと経路が生成された際には、連続的に振動し、進行方向が変わったことを伝達する。最後のウェイポイントへと到達した際には、連続的振動が 5 回繰り返され、ゴールへと到達したことを伝達する。

図 6 に経路伝達手法の概念図を示す。また伝達する情報と通知パターンの対応について表 2 に記した。

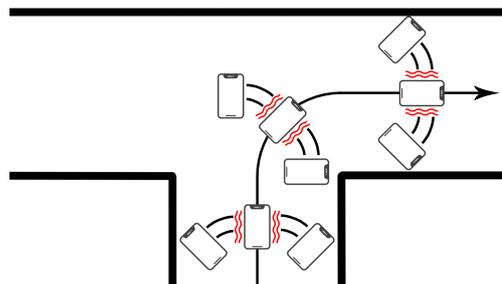


図 6. 経路伝達手法の概念図

表 2. 伝達する情報と通知パターンの対応

伝達する情報	通知パターン
進行方向通知 (通常)	長い振動 1 回 (0.5 秒程度)
進行方向通知 (方向転換点に近い)	短い振動 1 回 (0.2 秒程度)
ウェイポイント通過	連続で振動 (0.2 秒と 0.5 秒程度)
ゴールに到着	連続で振動 (0.2 秒と 0.5 秒程度) 5 回
衝突する危険性がある場合	警告音

5 評価実験

5.1 晴眼者を対象とした歩行実験

5.1.1 実験方法

本システムで用いた SLAM の測位精度を評価する実験を行った。ある一般的な建物内の廊下の床面に目印となるマーカーを設置して晴眼者 1 名が本システムとともにそのマーカーの地点まで歩行し、マーカーの地点で停止した状態での座標データを取得する。これを同じ地点で 10 回行い精度を評価する。

5.1.2 実験結果

図 7 に SLAM によって取得した 10 回分の歩行経路の全データについて重ねて示す。また、この全体の歩行のデータのうちマーカーの地点で停止している状態 30 フレームを切り出し、マーカーの地点で停止した状態での座標データとした。図 8 にマーカーの地点で停止した状態の座標データを示す。本システムは単眼 SLAM であるため、3 次元情報は相対的にしか計算できず、絶対的な距離情報を得ることができないが、参考として建物の廊下の横幅は 213 cm でありその距離は SLAM により作成した地図上では 0.424 であった。このことから、SLAM の座標データ上での距離 1 は 502 cm に相当する。

停止した状態での座標データについて各歩行ごとの平均値を取り、x, y 方向それぞれに対しデータの範囲と分散を計算したところ、x 方向のデータの範囲は 0.00977 (4.90 cm に相当)、標準偏差は 0.00344、y 方向のデータの範囲は 0.00928 (4.66 cm に相当)、標準偏差は 0.00270 であった。

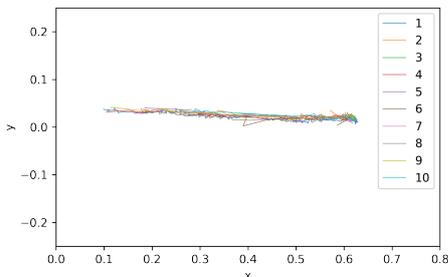


図 7. SLAM の測位精度を評価する実験における 10 回分の歩行経路の全データ

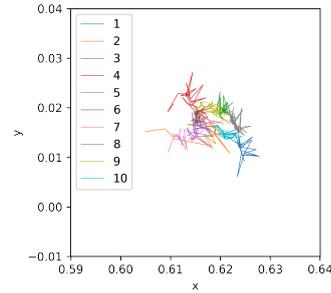


図 8. 停止した状態 30 フレームの座標データ (座標軸の長さ 0.05 は 25.1 cm に相当)

5.2 視覚障害者を対象とした歩行実験

5.2.1 実験方法

日本盲導犬協会神奈川訓練センターにて視覚障害者 6 名を対象に屋内における歩行実験を行った。実験の参加者は盲導犬や白杖を利用した単独歩行の経験がある。参加者の属性の内訳は、年齢: 32-65 (平均 51) 歳, 男性 4 名, 女性 2 名, うち盲導犬使用者 1 名, 白杖使用者 5 名である。

当事者を含めた打ち合わせにより、初めて訪れる建物のエレベーターホールから目的地の部屋まで歩行するというシチュエーションを模して、訓練センター 2 階のエレベーターホールのエレベーター前から目的地となる部屋の入り口まで、3 つのウェイポイントを設けたコースを作成した。屋内施設管理者が作成した地図を利用者が用いる状況を考え、地図は事前に用意されたものを用いた。

本システムの概要とインターフェースの動作について口頭で説明した後、システムの使い方を説明しながら実際のシステムと共にコースを歩行してもらい、動作の理解が確認できた段階で実験に移った。説明に際しては最大 2 回コースを歩行してもらった。その後エレベーターホールから設定したウェイポイントを順にたどり、目的地となる部屋まで本システムと白杖を用いて歩行してもらった。実験後、経路伝達手法に関して実験参加者からのフィードバックをもらった。システムが誤作動し予定の経路から外れた場合や障害物に衝突する可能性を考え、経路外の安全の確保や、参加者 1 名につき支援者 1 名以上を確保し、安全面での指導と管理を行うなど安全性には十分注意して行った。図 9 が歩行実験の様子である。

5.2.2 実験結果

実験参加者全員が振動のある方向を維持し、体を回転させることにより探りながら歩くことができた。特に白杖を用いた参加者は、白杖を用いた普通の歩行の動作を通じて進行方向を把握し、歩行すること

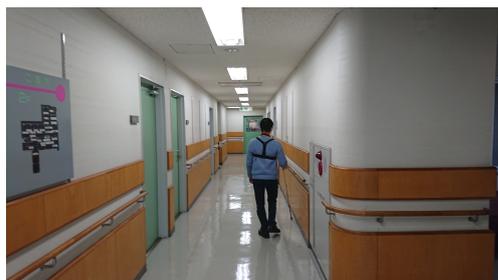


図 9. 視覚障害者を対象とした歩行実験の様子

ができた。また、実験本番では実験参加者全員が設定されたルート歩き、目的地で停止することが出来た。

実験参加者からのフィードバックでは、前方で左右に手を振りながら歩き進行方向へ誘導する手法に関して、進むべき方向が直感的にわかった、曲がり角をどちらに曲がるべきかがよくわかった、といった好意的な声があった一方、進行方向を振動で伝える手法はそのまま音を併用したほうがわかりやすいといった声もあった。また、振動の周期を変化させることで目的地までの距離を伝達する手法に関しては、直進中にそろそろ曲がるということがわかり身構えることができた、といった好意的な声があった一方で、振動の使い分けが難しいといった否定的な声もあった。アンケート結果としては、好意的な評価が多く、使いたい、使うのが簡単、すぐに使えるようになるという項目が高い一方で、否定的な評価としては事前に多くのことを学ぶ必要があるという項目が高かった。

5.3 考察

図 5.4 より、停止した状態 30 フレームの座標データはいずれも x, y 方向ともに 15 cm 相当の範囲に収まっている。また、停止した状態での座標データの各歩行ごとの平均値の範囲は、 x, y 方向ともに 5 cm 弱に相当する。視覚障害者の歩行支援においては、視覚障害者が白杖で確認できる範囲が半径 1m ほどであることから、自己位置推定の際には誤差 1m ほどで測位する事が望ましい [3]。よって本システムの SLAM は屋内での経路案内に十分な測位精度があると考えられる。

視覚障害者を対象とした歩行実験では、目的地まで決められたルートに従い歩行し、最終目的地で停止できた点で、本手法のガイダンス精度は十分高く、有効であることが分かった。

練習の際うまく目的地にたどり着けない原因として利用者のシステム理解が口頭の説明だけでは不十分であったことが挙げられる。練習の際に場所によって SLAM が機能しにくい点があること、必ずしも振動の方向通りに歩かなくてもシステムがリアルタイ

ムで経路作成し案内してくれることを説明し、利用者のシステム理解を深めることで、本番では全員が設定されたルート歩き、目的地で停止することが出来た。

また、前方で左右に手を振りながら歩き進行方向へ誘導する手法と、その振動の周期を変化させることで目的地までの距離を伝達する手法、どちらに対しても好意的な反応が得られ、視覚障害者の歩行支援システムにおける経路伝達手法として今後も発展が期待される。一方、フィードバックには振動の使い分けが難しいといった声や、振動と音を併用したほうがわかりやすいといった声があり、振動パターンの違いだけで情報を伝達することの難しさも明らかになった。聴覚により情報提示する場合は、人の声よりは短い音による伝達が望ましいとの声もあり、今後の改良として現状のシステムに曲がり角に来た、方向が間違ったといった際に音で通知する機能を追加することが考えられる。

6 まとめ

本論文では、まず視覚障害者の安全かつ自由な歩行のためにはいまだに多くの課題があることに注目し、既存の白杖や盲導犬による視覚障害者の歩行支援の現状についてと、その問題点について述べた。次に、人間が歩行する際の要素や視覚障害者およびその支援者を対象としたインタビューから視覚障害者の歩行支援に求められる要素について整理した。また、コンピュータを用いた歩行支援技術に関する先行研究について、地図作成および障害物検知手法と利用者への情報の伝達手法 2 つの視点から分類し、それらの手法の特徴と長短について比較して論じた。これらを踏まえて、地図作成および自己位置推定手法として Visual SLAM を用いて、経路案内手法として方位センサと振動子を内蔵した手持ちデバイスを用いた森らの手法に、方向転換点との距離に応じて振動周期を変化させ、進路変更のタイミングを事前に通知する手法を組み合わせたシステムを実装して、その評価を行った。SLAM の精度を評価するための実験では、本システムとともに床上に設置したマーカーまで歩行する試行を複数回行い、その誤差から本システムの SLAM は屋内での経路案内に十分な測位精度があることがわかった。視覚障害者を対象とした歩行実験では、参加者全員が設定されたルートを歩くことが出来、本システムの有効性が確かめられた。また、伝達手法の使用感に関しても参加者のフィードバックから好意的な反応が得られ、本システムの誘導手法は視覚障害者の歩行支援システムにおける経路伝達手法として有効であることがわかった。

今後、より利用者が簡単に使えるシステムとするために、振動パターンの違いで情報を伝達するだけでなく、音声等の併用も行っていく。

謝辞

本論文の研究は一部、三菱財団社会福祉事業・研究助成「ウェアラブルカメラによる 3D スペースセンシングと触覚通知による、視覚障害者のための屋内ナビゲーションシステムの開発」によるものです。支援者・当事者による評価では、神奈川県視覚障害者福祉協会の鈴木孝幸様、日本盲導犬協会神奈川訓練センターの山口義之様、多和田悟様にご協力いただきました。VSLAM システムの Android 実装につき、株式会社デジタルハンズの清水敏雄様にサポートいただきました。

参考文献

- [1] M. Brock and P. O. Kristensson. Supporting blind navigation using depth sensing and sonification. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication, UbiComp '13 Adjunct*, pp. 255–258, 2013.
- [2] A. Cosgun, E. A. Sisbot, and H. I. Christensen. Guidance for human navigation using a vibrotactile belt interface and robot-like motion planning. *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6350–6355, 2014.
- [3] N. Fallah, I. Apostolopoulos, K. Bekris, and E. Folmer. Indoor human navigation systems: A survey. *Interacting with Computers*, 25(1):21–33, 2013.
- [4] J. Guerreiro, D. Sato, S. Asakawa, H. Dong, K. M. Kitani, and C. Asakawa. CaBot: Designing and Evaluating an Autonomous Navigation Robot for Blind People. *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 68–82, 2019.
- [5] C. S. Kallie, P. R. Schrater, and G. E. Legge. Variability in stepping direction explains the veering behavior of blind walkers. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1):183, 2007.
- [6] S. Kayukawa, T. Ishihara, H. Takagi, S. Morishima, and C. Asakawa. BlindPilot: A robotic local navigation system that leads blind people to a landmark object. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2020.
- [7] R. Mur-Artal and J. D. Tardós. ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. *IEEE transactions on robotics*, 33(5):1255–1262, 2017.
- [8] M. Pielot, B. Poppinga, W. Heuten, and S. Boll. Pocketnavigator: Studying tactile navigation systems in-situ. pp. 3131–3140, 2012.
- [9] A. J. Spiers and A. M. Dollar. Design and evaluation of shape-changing haptic interfaces for pedestrian navigation assistance. *IEEE Transactions on Haptics*, 10:17–28, 2016.
- [10] R. Tapu, B. Mocanu, and T. Zaharia. Real time static/dynamic obstacle detection for visually impaired persons. In *2014 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 394–395, 2014.
- [11] S. Thrun, W. Burgard, and D. Fox. Probabilistic Robotics (Intelligent Robotics and Autonomous Agents). *The MIT Press*, 2005.
- [12] K. Tsukada and M. Yasumura. Activebelt: Belt-type wearable tactile display for directional navigation. *Lecture Notes in Computer Science*, 44, 2004.
- [13] A. Wachaja, P. Agarwal, M. Zink, M. R. Adame, K. Möller, and W. Burgard. Navigating blind people with a smart walker. *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6014–6019, 2015.
- [14] Y. Wei and M. Lee. A guide-dog robot system research for the visually impaired. In *2014 IEEE International Conference on Industrial Technology (ICIT)*, pp. 800–805, 2014.
- [15] H. Zhang and C. Ye. A Visual Positioning System for Indoor Blind Navigation. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 9079–9085, 2020.
- [16] 英雄森. 歩行ガイドロボット実用化への道 視覚の役割. 情報処理学会研究報告コンピュータビジョンとイメージメディア (CVIM), 2003(2):59–66, 2003.
- [17] 栗林 雅希, 粥川 青汰, J. Vongkulbhisal, 浅川 智恵子, 佐藤 大介, 高木 啓伸, 森島 繁生. Corridor-Walker: 視覚障害者が障害物を回避し交差点を認識するためのスマートフォン型屋内歩行支援システム. *The 29th Workshop on Interactive Systems and Software (WISS 2021)*, 2021.
- [18] 厚生労働省. 平成 28 年生活のしづらさなどに関する調査 (全国在宅障害児・者等実態調査) 結果, 2018.
- [19] 山崎康平, 穴戸英彦, 北原格, 亀田能成. 類似画像検索と SLAM による協調的位置推定システムの性能評価. 電子情報通信学会技術研究報告; 信学技報, 119(222):89–94, 2019.
- [20] 森 大輝, 入江 英嗣, 内原 正一, 荒川 明宏, 坂井 修一. リアルタイム経路生成と振動通知による視覚障害者の歩行支援システム. *The 26th Workshop on Interactive Systems and Software (WISS 2018)*, 2018.
- [21] 大倉元宏, 清水美知子, 田内雅規, 村上琢磨. 視覚障がい者の歩行の科学 - 安全で安心なひとり歩きをめざして -. 2014.
- [22] 大倉元宏, 村上琢磨, 清水学, 田内雅規. 視覚障害者の歩行特性と駅プラットフォームからの転落事故. *人間工学*, 31(1):1–8, 1995.

- [23] 嶋村幸仁, 山口通智, 片山博. リスク概念と視覚障害者のハザード要因に関する研究. 安全工学, 53(2):108-114, 2014.
- [24] 福井良太. 世界から見た日本の盲導犬育成事業. 日本補助犬科学研究, 2(1):22-25, 2008.
- [25] 友納正裕, 原祥堯. SLAM の現状と今後の展望. システム/制御/情報, 64(2):45-50, 2020.

未来ビジョン

本研究では RGB カメラを用いた Visual SLAM を用いたが LiDAR など別のセンサを使う, あるいは併用してのさらなる安定化と精度の向上が検討される. さらに実際に視覚障害者が単独で用いることを想定したインターフェースの開発なども挙げられる. 視覚障害者が手軽に使えるシステムとするために今後とも当事者を含めた議論を重ねていく. そして将来的にはホテル・病院など屋内施設内における

視覚障害者のナビゲーションシステムの実現を目指している. これにより屋内でのナビゲーションを可能とし, 視覚障害者の単独での外出を後押しする. 具体的には, ホテルなどが貸し出した端末や, 利用者のスマートフォンなどを利用者が身に着け, 目的の部屋などへのルートを自動で生成し, 案内を行う利用法が検討される. そして, どんな人であってもコンピュータ技術によって不便を克服し, 気軽に社会参画できる未来を見据えている.

Alertable Surfaces: 人の接触と消毒を認識しウイルスの付着を警告する実環境

岸野 泰恵* 白井 良成* 柳沢 豊* 菅原 光† 尾原 和也* 水谷 伸*

概要. ウイルスが付着している可能性の高いモノの表面に警告を表示し、消毒行為を認識可能な環境 Alertable Surfaces を提案する。接触感染は、感染者から排出されたウイルスがモノの表面などに付着し、それを触れた他者にウイルスが付着することで起こる。Alertable Surfaces はウイルスが付着した可能性の高い場所に対して警告を光学的に表示することで、接触感染を防止することを狙っている。さらに Alertable Surfaces では消毒の行為を識別し、ウイルスが付着している可能性がなくなった時には警告を消去する。これにより、利用者は、ウイルスのある可能性の高い場所をよけたり、可能性のある場所を集中的に消毒したり、ウイルスが付着した可能性が低いことから安心感を得たり、といったことが可能になる。本稿では、サーマルカメラを用いて接触と消毒を検知し、プロジェクタを用いて可視化するプロトタイプシステムの実装について述べる。さらに本稿では、画像処理性能を評価した結果とプロトタイプシステムを展示した際の感想を紹介し、Alertable Surfaces の可能性や課題について議論する。

1 はじめに

2020 年初頭から新型コロナウイルス (COVID-19) の感染が拡大し、2 年経った 2022 年現在でも世界で多くの感染者が発生し続けている。新型コロナウイルス (COVID-19) の予防には、ウイルスを減らし感染予防をすることが重要である [11]。ウイルスを減らすためには大きく分けて、手や指などに付着したウイルスへの対策、空気中のウイルスへの対策、モノに付着したウイルスへの対策の 3 つ方法がある。このうち手や指に付着したウイルスに対しては手洗いやアルコール消毒、空気中に浮遊するウイルスに対しては換気が有効な方法となる。モノに付着したウイルスに対しては、アルコールや次亜塩素酸水などでモノの表面を消毒する方法があり、人が多く触れる場所などでは定期的に消毒することでウイルスを減らそうとしている。

しかし、モノの表面を定期的に拭く消毒方法は、最も消毒する必要のある手指に限って消毒する対策や、窓を開けたりサーキュレータを設置するだけで部屋全体の換気が行える空気に対する対策と比べて多くの労力を要する。また、定期的に拭いて消毒されたとしても、消毒の拭き残しはないか、消毒した後に他の人が触っているのではないか、といった不安が残る。そこで本稿ではウイルスが付着している可能性の高い場所を可視化して警告可能な環境として Alertable Surfaces を提案する。さらに本稿では、Alertable Surfaces を実現するための画像処理法や、試作したシステムを展示した際の感想を元にシステムの利用場面や情報提示方法についても議論する。

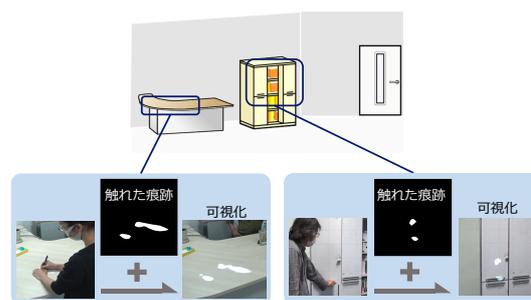


図 1. Alertable Surfaces イメージ図

2 Alertable Surfaces

2.1 警告の表示と消去を可能とする環境の提案

Alertable Surfaces は、実環境の表面 (環境表面) 上にウイルスの付着している可能性が高い場所を可視化することを狙って提案している。図 1 にイメージ図を示す。新型コロナウイルスの感染経路の一つに接触感染がある。例えば、患者の体から排出されるウイルスがくしゃみや咳をした際に患者の手に付着し、次にこの手で触れた環境表面 (ドアノブや机など) にウイルスが付着する。その後そこを別の人が触れ、その手で口や鼻、目などの粘膜に触れるとそこからウイルスが体内に入り感染してしまう。ウイルスを検出するためには、PCR 検査や抗原検査などの医学的な検査を行う必要があるが、数十分から数時間の時間が必要であり、環境表面上にウイルスが存在するかをその場で検出することは現在の技術では困難である。しかし、人の動きは様々な方法で検出することができ、ウイルスは人を介して広がる。人の動きをセンシングすることで、ウイルスが

* NTT コミュニケーション科学基礎研究所

† 豊橋技術科学大学 (NTT での実務訓練中に従事)

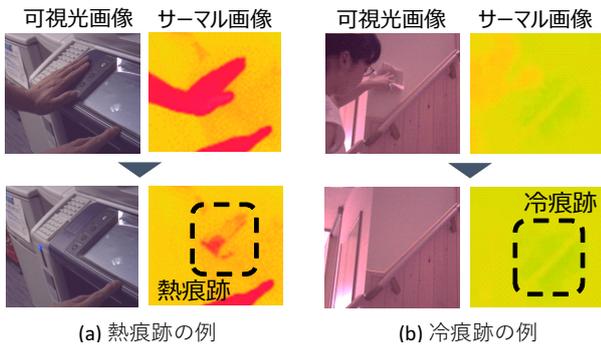


図 2. 熱痕跡と冷痕跡の例

存在する可能性が高い場所を推測することは可能ではないかと考えた。そこで、Alertable Surfaces では、ウイルスそのものではなく、ウイルスが付着している可能性の高い場所を可視化することを考える。また、消毒を行った際には、その部分のウイルスは死滅したと考えられるため、可視化を終了すればよい。

Alertable Surfaces を使ってウイルスが付着している可能性の高い場所を可視化できると以下のような効果があると考えられる。

- ウイルス付着の恐れのある場所を可視化することで、そこにウイルスがあるかもしれないとユーザーに注意喚起が行える。
- ウイルスがあるかもしれないと考えることで消毒をするモチベーションが高まる。
- ウイルスの付着した可能性のある場所のみを消毒することで消毒作業の効率化が図れる。
- 消毒が行われたあと、誰も触れておらずウイルスの付着する可能性が生じていない場所だと可視化によって確認できれば安心できる。

ウイルスの付着の可能性としては、くしゃみや咳などによって口等から排出された飛沫の付着、ウイルスの付着した人の手を介した付着等いくつかのパターンが考えられる。本研究では、研究の最初のステップとして温度の変化を取得することで比較的容易に検出できる人が環境表面に触れた箇所を可視化するプロトタイプシステムを構築し、ユーザーの評価を得て可視化方法などインタフェースデザインについて検討を進めることとする。

2.2 熱に注目した実現方法の検討

人は恒温動物であり、体内での発熱により 37℃ 前後の熱を維持している。これに対して、人の周りの気温や環境表面の温度は体温より低い。(酷暑の夏の日中等の例外を除く。) そのため、人がモノの表面を手で触れるとその表面に人の手の熱が残る。これは熱痕跡 (Heat traces) と呼ばれる。図 2 (a) に例を示す。本研究では、サーマルカメラを用いて環境

表面を撮影し、画像処理によって熱痕跡を検出することで、人が環境表面に触れたことを認識する。

サーマルカメラを用いて環境表面を撮影していれば、消毒の検出も可能である。環境表面に対する消毒としては、アルコールや次亜塩素酸水などの消毒液をスプレーで環境表面に吹き付けるか、消毒液で濡らした布などで拭く方法が一般的である。この際、環境表面は消毒液でいったん湿るものの、すぐに消毒液は気化して乾く。液体が気化するときには、気化熱を奪って蒸発していくため、環境表面の温度が一時的に低下する。これを冷痕跡と呼ぶ。図 2 (b) に例を示す。画像処理でこの温度の低下を検出することができれば消毒がどこで行われたかを検出、記録できるようになる¹。

2.3 関連研究

熱痕跡が触れたモノの表面に残ることはこれまでも知られており、ヒューマンインタフェースの分野では壁面ディスプレイにおけるインタラクションの認識に活用したり [3, 6]、スマートフォンやキーボード上の熱痕跡から入力内容を推測できるか調査 [1] した研究例が存在する。また、Iwai らが提案した文書管理支援システムでは、サーマル画像の差分画像を用いて人が書類や本に触れたことを検出し、これをシステム操作に利用している [2]。本研究は、熱痕跡を画像処理によって取得し、人の触れた箇所の可視化に活用しようとするものである。本研究の画像処理のアプローチは Iwai らの手法と似ているが、机上に限らず書類や机以外のものを対象にしている点、消毒による冷痕跡も扱っている点が異なる。

人が触れた場所を検出する方法としては、サーマルカメラを用いて熱痕跡を取得する方法以外にもいくつかの方法がある。例えばブラックライトで光る塗料を手指に塗り、さまざまな行動をとった後でブラックライトを光らせて触った場所を確認する方法や、人の三次元の姿勢を取得するセンサを使って人の動きを取得して触れた危険性の高い箇所を可視化する手法 [10] などが提案されている。これらの手法とサーマルカメラを用いる方法を比較すると、塗料を使った方法は日常生活の中で継続的に利用したり不特定多数を対象とすることは難しく、姿勢を取得する方法では確実に触れてウイルスが付着した可能性が生じているかを判定することは難しい。

サーマルカメラを用いた熱痕跡画像処理に関する研究例 [4] も存在し、深層学習を活用し一定の精度で熱痕跡を検出することに成功している。この手法ではサーマル画像のみを使っているのに対し、本研究ではサーマルカメラと近赤外線カメラを使うこと

¹ 厳密にはウイルスに効果のある消毒液を十分な濃度と量で使うことが重要となるが、ここでは対象のウイルスに有効な消毒液を使ったものと仮定し、消毒を可視化することの効果について検討を進める。

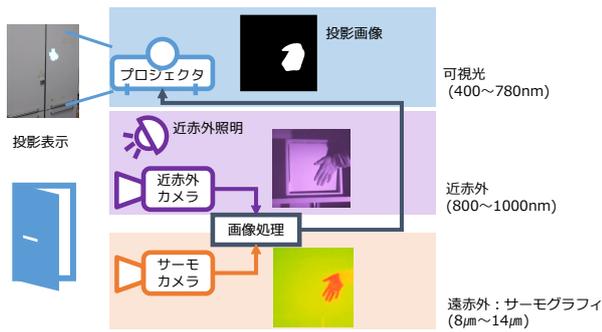


図 3. システム構成



図 4. 実装したシステム

で、実環境の変化に軽量なアルゴリズムで対応し、リアルタイムに動作するシステムを構築する。

3 Alertable Surfaces の実現

Alertable Surfaces は実際の環境でリアルタイムに接触を検知し、それを可視化する。以下で設計の方針と実装、画像処理アルゴリズムについて述べる。

3.1 システム構成

実際の環境でサーマルカメラを用いて連続撮影を行うと、室温や人の表面体温が変化し、さらにサーマルカメラ自体の温度キャリブレーションによって、同じ対象を撮影していても異なる温度で撮影されることが起こり得る。そこで近赤外線カメラを用い、2種類のカメラの画像を相補的に用いた画像処理アルゴリズムによって、人がなくなった後に残る熱痕跡を軽量な処理で検出することを目指す。可視光領域を情報提示のプロジェクションに利用するため、近赤外領域のカメラを用いてプロジェクタからの可視化の結果によらず確実に画像処理が行えるようにする。また、できるだけ軽量な画像処理アルゴリズムを目指すのは、深層学習のような複雑なアルゴリズムを用いると計算力のある計算機を要し、設置可能な場所に制約が出る可能性があるためである。

図3にシステム構成を示す。近赤外カメラとサーマルカメラの2種類のカメラで対象物体を撮影する。近赤外領域は屋内照明では明るさが不十分な場合が多く、近赤外線照明で対象を照射する。近赤外画像とサーマル画像を組み合わせる画像処理を行い、その結果を投影画像として生成する。作成した投影画像はプロジェクタで対象物体の上に重畳表示を行う。

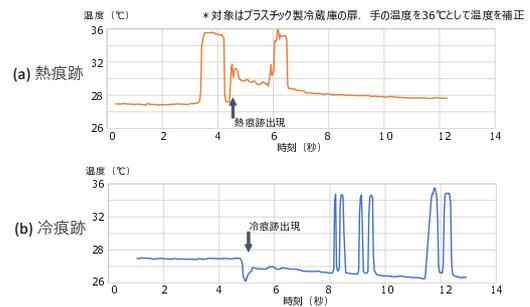


図 5. 熱痕跡出現時と冷痕跡出現時の温度変化

システムは、サーマルカメラに FLIR Boson640R (解像度:640×512)、近赤外カメラに GAZO MCM-303NIR (解像度:640×480) を用いて Windows PC で画像処理を行う PC システムと、サーマルカメラに FLIR Lepton 3.5 (解像度:160×120)、近赤外カメラに Raspberry Pi NoIR Camera V1.3 (解像度:640×480) を用いて ZYNQ7020 で画像処理を行う軽量システムの2種類を実装した。プロジェクタはどちらも熱の発生が少ない LED プロジェクタ CASIO CX-E1 を用いた。軽量システムのカメラボードは [8] で試作したものである。近赤外カメラでは可視光を撮影可能なカメラに可視光カットフィルタを装着した。軽量システムの外観と PC システムのカメラ部分を図4に示す。PC システムで 15 fps、軽量システムでは 3 fps 程度の速度で画像処理を行うことができ、接触を検知した結果をすぐに提示できることを確認した。

3.2 サーマルカメラを用いた接触と消毒の検知

次に、サーマルカメラと近赤外線カメラを用いた接触検知アルゴリズムの詳細について述べる。なお、画像処理開始前に両方の画像が同じ領域に重なるようにキャリブレーションが完了しているものとする。

図5に PC システムで取得した熱痕跡と冷痕跡出現時の温度の時間変化を示す。グラフ中の温度の急激な上昇は人の手など温度が高い物体の通過である。(a) 熱痕跡は手が離れた直後が一番温度が高く、その後徐々に低下し、数秒から十秒程度で元の温度に戻る。(b) 冷痕跡は、出現時に消毒液の温度が直接撮影されて温度が大きく下がるが、その後は対象物体の温度が気化熱によって十秒から数十秒をかけて低下する。グラフの範囲外になるが、物体表面の温度は消毒液が揮発した後、徐々に元に戻る。

3.2.1 2種類のカメラを用いた接触と消毒の検知

接触と消毒の検知は背景差分画像処理を用いて行う。具体的には、カメラを設置した場面の定常状態を背景画像として保存し、背景画像と現在の画像の差分画像を作成し、そこに生じた変化に対して画像処理を行うことで接触や消毒の有無を判定している。

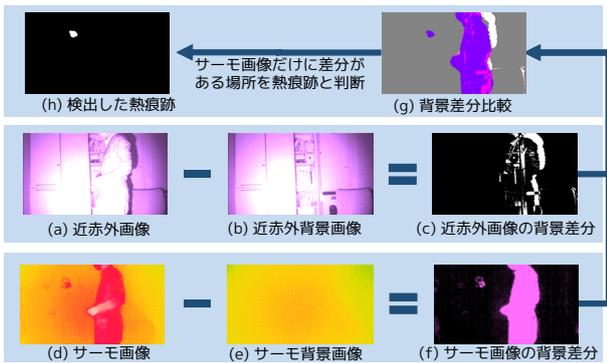


図 6. 画像処理手順

	近赤外画像	サーマル画像	説明
(a)			背景画像
(b)	$d_I \geq \epsilon$	$d_H < \epsilon$	近赤外画像のみの変化であり、照明変化と考え、近赤外の背景画像を更新
(c)	$d_I \geq \epsilon$	$d_H \geq \epsilon$	両方の画像で変化あり、人が通るなど一時的な変化と考え、背景画像は更新しない。
(d)	$d_I < \epsilon$	$d_H \geq \epsilon$	サーマル画像のみ全体的に温度の変化があり、サーマルカメラのキャリブレーションなどが考えられ、サーマル背景画像を更新する。
(e)	$d_I < \epsilon$	$d_H \geq \epsilon$	サーマル画像のみに変化があり、これは熱痕跡と認識されている。熱痕跡を含む画像は一時的な変化と考え、背景画像を更新しない。

$d_I = d(I_t, I_b)$, $d_H = d(H_t, H_b)$ とする。

図 7. 背景画像更新手法の概要

背景画像との差分を算出し、近赤外画像とサーマル画像の両方に変化があった時は背景の前を人などが通り過ぎたと考え、近赤外画像のみに変化があった時は照明などの変化とし、サーマル画像のみに変化がある場合を熱痕跡や冷痕跡の出現とみなす。

以下、アルゴリズムの詳細を説明する。図 6 に画像処理のイメージ図を示す。現在画像処理を行う対象の時刻を t とし、時刻 t における近赤外線画像を I_t 、サーマル画像を H_t とし、それぞれの背景画像を I_b 、 H_b とする。最初に近赤外の背景差分画像 D_I 、サーマル背景差分画像 D_H を作成する。次に、 D_I と D_H を重ね合わせ (g)、それぞれの白色連結領域の重複度合いが一定以上のものを削除し、サーマル画像 H_t のみに差分が生じている部分を抜き出す (h)。これを熱痕跡とする。これを温度の低下に対して同じように処理を行えば、冷痕跡を検出できる。なお、実際には、膨張収縮など基本的なノイズ処理を行っているが、一般的な処理であるため説明を省略する。

3.2.2 背景画像調整アルゴリズム

連続的に画像処理を行っているとき前述したような状況の変化や温度キャリブレーションのために、元

からある背景画像では正しく画像処理ができなくなる場合がある。安定して接触検知を続けるためには、背景画像調整処理が必要となる。背景画像の調整方法としては、事前に撮影した複数の照明条件下の画像を動的に組み合わせる方法が提案されていたり [9]、画像処理によって背景以外のモノの有無を判定する方法もあり得るが、本研究では 2 種類のカメラ画像を用いていることを活かし、両方の差分画像を使うことで軽量の計算で背景画像の調整を行うことを目指す。サーマルカメラと近赤外カメラで同時に撮影する際、画面内でモノが動くときほとんどの場合近赤外領域だけでなく温度にも変化があり、これを前提とする。また、急激な温度変化のない屋内を想定している。

以下、アルゴリズムの詳細を説明する。アルゴリズムは背景画像を更新する部分と、部分的に差分が生じている場合に対応するオフセット法の 2 つからなる。2 つの手法を背景画像更新の状況に応じて切り替える概要を図 7 に示す。人が通り過ぎる等大きく変化中の画像は背景画像に適さないため、まず一定時間 δt の間画像に大きな変化がないことを $d(I_t, I_{t-\delta t}) < \epsilon$ と $d(H_t, H_{t-\delta t}) < \epsilon$ で確認する。ここで $d(I_1, I_2)$ は 2 枚の画像の差であり、例えば全画素値の差分の平均値を計算して算出する。近赤外背景画像 I_b とサーマル背景画像 H_b はあらかじめ撮影されているものとする (a)。

$d(I_t, I_b) < \epsilon$ かつ $d(H_t, H_b) < \epsilon$ の場合、両方の画像に大きな変化がないため、背景画像処理を行う必要はない。

$d(I_t, I_b) \geq \epsilon$ かつ $d(H_t, H_b) < \epsilon$ の場合、照明などが変化したため、近赤外画像のみに差が生じたと考え、背景画像 I_b を I_t で更新する (b)。

$d(I_t, I_b) \geq \epsilon$ かつ $d(H_t, H_b) \geq \epsilon$ の場合、両方の画像に大きな変化があり、人が前を通り過ぎる等の一時的な変化が起きていると考えられ、背景画像の更新を行えない (c)。しかし、温度キャリブレーションなどの原因で差分が生じていることを考慮し、オフセット法による背景画像調整を行うため、 D_I と D_H 、この時点で検出されている熱痕跡と冷痕跡を重ね合わせたマスク画像 I_m を作成する。

$d(I_t, I_b) < \epsilon$ かつ $d(H_t, H_b) \geq \epsilon$ の場合、サーマル画像のみに差分が発生しているため、温度キャリブレーションや空調による温度変化が生じていると考えられる。このとき、既存の熱痕跡や冷痕跡による一時的な温度変化が生じている可能性もあるため、熱痕跡と冷痕跡の有無を確認し、両方がない場合のみ H_b を H_t で更新する (d)。痕跡がある場合は、痕跡のある部分をマスク画像 I_m としてオフセット法による背景画像調整の必要性を検討する (e)。

次に、オフセット法による背景画像処理を説明する (図 8)。時刻 t の画像全体の平均温度 \bar{T}_t と背景画像の平均温度 \bar{T}_b を算出する。マスク画像 I_m を使

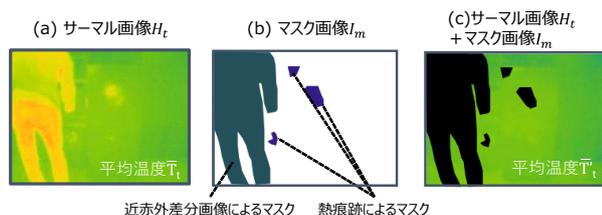


図 8. オフセット法による背景画像調整

い、マスクされていない部分のみの平均温度 \bar{T}'_t と \bar{T}'_b を算出し、差分 $dT = \bar{T}'_t - \bar{T}'_b$ を計算し、 δT よりも大きければ、マスクされていない背景がそのまま映っている部分に変化があると考え、背景画像の調整を行う。背景画像 H_b 全体に dT を加算し、加算後の背景画像を用いて前節の接触検知の画像処理を行う。 H_t の方が H_b よりも温度が低い場合も同様に、負の値の dT を加算し、 H_b の温度を全体的に下げて処理を行う。

4 画像処理性能の評価

両システムで動画を撮影し、熱痕跡認識の精度を調査した。消毒についても認識は可能であるが、気化熱と消毒の有効性について調査しなければ正しく消毒が行えたかを判定できないため、本稿では人が触れたことを検出する性能のみを調査した。なお、図5に示すように消毒による冷痕跡の方が温度変化の継続時間が長く温度差も同等か大きいことから、この撮影時に用いた次亜塩素酸水による冷痕跡に対しては、検出は熱痕跡より容易であると推測される。

実験では、冷蔵庫、棚、壁、紙製引き出し、の4種類のモノが設置された状況を約1時間半撮影し、その間、著者を含む数名が冷蔵庫のドアを開閉する、引き出しを開閉するといった操作を自由に行った。また、対象に触れずに、カメラの前を通り過ぎたり立ち止まるといった動作も行った。撮影後、人が触れた後の熱痕跡を手動でラベリングし正解データとした。PCシステムで約120個、軽量システムで約80個の熱痕跡をラベリングした。

熱痕跡と正解データが一致しているかの判定には、Dice係数を用いることとした。Dice係数は2つの領域が重なっていることを評価するための係数の一種であり、2つの領域が重なった面積の2倍を2つの領域の和で除算することで求められ、0.1以上である場合を正解と判定した。熱痕跡は中央に比べて力の加わりにくい周辺部で温度差が小さくなることが多く、幾通りかの値を検討の上、一般的な画像処理より小さい値を採用した。また、各フレームの画像に対して熱痕跡検出を評価するのではなく、正解ラベルのある区間で1回でも熱痕跡を検出できれば対応する熱痕跡の検出に成功したとした。運用上、人が触れた事実を1回でも検出できれば十分である

表 1. 画像処理の評価結果

(a) 全体の認識率

背景調整	PCシステム		軽量システム	
	無	有	無	有
適合率	0.82	0.46	0.056	0.37
再現率	0.2	0.65	0.79	0.56

(b) 素材別再現率

対象	冷蔵庫	棚	紙引出	壁
PCシステム	0.71	0.5	1.0	0.75
軽量システム	0.72	0.34	0.67	1.0

ためである。

認識精度を評価する実験を行った結果を表1に示す。表1(a)に示すように軽量システムにおいて背景画像の調整を行ったことで温度変化のノイズを誤認識することが減り適合率が改善したが、熱痕跡を認識できない場合が増加し再現率は低下した。PCシステムでは逆に適合率が下がり、再現率が上がる結果となった。PCシステムでは撮影した期間、部屋の温度が低下し続けていたため、これが補正されて再現率が向上した一方、誤検出の割合も上がったのだろうと考えられる。

素材別の再現率についてさらに調査を行った。結果は表1(b)に示す。冷蔵庫の表面はマットな質感のプラスチック、棚はプラスチックや金属等複数種類の素材、紙引出はダンボール、壁は金属をマットなペンキで塗装した素材である。この中で紙引出や壁は高い再現率であるが、棚は再現率が低かった。棚は反射のある金属や紙よりも熱が残りにくいプラスチック製であり、紙などと同じパラメータでは熱痕跡を検出しにくく、再現率が低下したと考えられる。棚以外での再現率はPCシステムで0.75、軽量システムで0.73である。さらに、領域の大きさが小さすぎ、検出が難しくなっている場合もあった。

背景画像の調整は、サーマルカメラがケースに入っておらずより温度変化を受けやすい軽量システムにおいて特に効果があったが、領域のサイズが小さすぎる場合や、固定の閾値では検出できなかった素材において認識誤りが残った。このように、実験より素材別の閾値を用い、カメラと対象の物体から適した領域の大きさの閾値を設定するといった画像処理の改良に向けた示唆が得られた。

5 議論

5.1 デモ展示に対する感想

実装したAlertable SurfacesのPCシステムを研究所内に実際の操作を試せるように設置してデモ展示を行い、さまざまな感想を得た。展示では小型冷蔵庫の扉と共用のテーブルを対象の環境表面とした。

可視化した箇所に対して、他の人が触った跡が表示されていても全く気にせず振る舞うという人もいれば、気になるので可視化された箇所以外を触れたいという感想の人もいた。新型コロナウイルスに対して、本人の性格や家族のリスクの高さなどさまざまな要因によって差が出ていると考えられる。また、社内の知り合いであればあまり気にならないが、不特定多数が触れたと考えると避けたいというコメントもあった。これらの感想から、触られた状況に応じて情報を提示することも有益だと考えられる。

また、わかりやすく他の人が触れた場所が可視化されていると、その外側ならば触れても安心だと考える人が多かった。一方で画像処理は認識場所のずれを含む場合があり、わずかに外側を持つとすると、実際には他の人が触れた場所と重なってしまうこともあり得る。このことから一定のマージンをもたせて可視化する必要があると言える。

また、消毒による熱痕跡表示の消去は消毒されたことがわかりやすいという感想が多かった。実際に消毒作業を行うとき、必要な箇所に集中して消毒できるというコメントや、スプレーで網羅的に消毒できたことを確認できると助かるというコメントがあった。今後、気化熱の温度の低下とウイルスに対する消毒の効果の関係性について調査を行い、ウイルスが消毒された可能性が高いときにのみ消毒を認識するような仕組みについて検討を進めていきたい。

5.2 認識精度の向上

現在の画像処理は、カメラボードでも動作する程度であるが、正確な接触の記録とするには後処理の追加や画像処理精度の向上が必要であると考えている。実際に3DCNN (3D Convolutional Neural Network) を用いた2チャンネルのカメラ動画を用いた画像処理にも着手している [5]。

また、実験結果が示すように、熱痕跡の残りやすさは環境表面の素材の熱伝導率によって異なる。熱伝導率の高い金属では熱がすぐに広まってしまい熱痕跡として残りにくく、紙や木材のような熱伝導率の低い素材では熱痕跡が残りやすい。また、表面の質感がざらざらしていれば手との接触点が減るため熱が伝わりにくく、つるつるした表面では接触点が多く熱がよく伝わる。さらに素材の厚さにも影響を受ける。このため、環境表面の材質や部位によって異なる画像処理パラメータを用いることで画像処理精度を向上させられる可能性がある。物体の素材の推定は、深層学習を用いる方法も提案されている [7]。

5.3 利用シーン

提案手法は、近赤外線画像の差分画像を用いているため、直射日光が当たり木漏れ日や雲によって照明条件の変化が激しいような状況では正確な画像処理が難しい。また、サーマルカメラ画像を使って接

触検知を行っているため、屋外のような温度変化の激しい状況で利用することも難しい。本システムは、まずは空調の効いた屋内での利用を想定し、必要とされる場面やユーザインタラクションについて検討を進めようという段階である。

画像処理精度の面からは、高い精度で接触を識別できた紙や木材などに物体の素材が限定され、人の手が物体に触れる部分を大きく撮影できる状況では、一定の精度で熱痕跡を記録、可視化できると見込んでいる。一方で、その他の素材を含む今回の評価実験のような精度であっても、接触感染の可能性を可視化して啓発したり、消毒のモチベーションを向上させるような目的では利用可能だと考えている。

今回実装したようなプロジェクタによる表示システムはオフィスの共用部のような、複数人で共有する机や冷蔵庫などに対して特に有益だと考えている。このような場面であれば複数人で触れる箇所が限られ導入も比較的容易だと考えられる。一方で、接触情報の収集のみを行い、可視化をスマートフォンなどの端末で行うシステムもあり得ると考えている。

5.4 接触による感染の可能性に関して

最近では、接触感染が発生するのは限られた状況下のみであるという報告もある [12]。一方で感染症拡大の予防という観点からは、接触感染の可能性を否定することもできないため、念のためにさまざまな環境表面の消毒を行い続けている。このような状況に対して、提案システムを用いると、人が本当に触れた最小限の消毒にとどめたり、人がよく触れる場所のログを取りその場所のみを消毒するといった消毒作業の効率化につながる。

6 まとめ

本稿では、Alertable Surfaces を提案し、人が環境表面に触れた箇所や、環境表面への消毒をサーマルカメラと近赤外カメラを用いて認識し、ウイルスが付着している可能性がある場所を可視化するシステムの実装について述べた。

今後は素材の差に対応するなどして画像処理性能の向上を行い、さらに状況に応じて情報を提示する方法について検討し、またそれに対するユーザの評価実験などを行う予定である。

また、本アプローチは、軽量のシステムで実現可能であるが、対応する素材に限りがあり、また空調の効いた屋内に限定される。将来的には、画像処理手法を改良するだけでなく、他のアプローチも併用することで接触箇所検知ができる環境を増やし、飛沫などによる感染防止についても検討を進めたい。ウイルスの付着した可能性のある場所を可視化することで、ピンポイントに効率よく消毒を行ったり、少しでも安心して過ごせるようになればよい。

参考文献

- [1] Y. Abdelrahman, M. Khamis, S. Schneegass, and F. Alt. Stay Cool! Understanding Thermal Attacks on Mobile-Based User Authentication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*, pp. 3751–3763, 2017.
- [2] D. Iwai and K. Sato, Kosuke. Document search support by making physical documents transparent in projection-based mixed reality. *Virtual Reality*. Vol. 15, pp. 147–160. 2011.
- [3] E. Larson, G. Cohn, S. Gupta, X. Ren, B. Harrison, D. Fox, and S. Patel. HeatWave: Thermal Imaging for Surface User Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, pp. 2565–2574, 2011.
- [4] G. Ma, W. Ross, M. Tucker, P. -C. Hsu, D. M. Buckland and P. J. Codd. Touch-Point Detection Using Thermal Video With Applications to Prevent Indirect Virus Spread. In *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 9, pp. 1–11, 2021.
- [5] 野口憲人, 岸野泰恵, 白井良成, 村田正幸. 近赤外画像とサーマル画像を用いた接触検知におけるマルチバンド画像処理の基礎検討. 画像の認識・理解シンポジウム2022 (MIRU2022), IS1-63, 2022.
- [6] A. Sahami Shirazi, Y. Abdelrahman, N. Henze, S. Schneegass, M. Khalilbeigi, and A. Schmidt. Exploiting Thermal Reflection for Interactive Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, pp. 3483–3492, 2014.
- [7] G. Schwartz and K. Nishino. Recognizing Material Properties from Images. <https://arxiv.org/abs/1801.03127>, 2018.
- [8] Y. Shirai, Y. Kishino, T. Suyama, and S. Mizutani. PASNIC: A Thermal Based Privacy-Aware Sensor Node for Image Capturing. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 202–205, 2019.
- [9] 土田勝, 川西隆仁, 村瀬洋, 高木茂. 背景差分法による物体検出を目的とした逐次モンテカルロ法による背景推定. 電子情報通信学会論文誌, Vol. J87-D-II, No. 5, pp. 1062–1070, 2004.
- [10] A. D. Wilson. Combating the Spread of Coronavirus by Modeling Fomites with Depth Cameras. In *Proceedings of ACM Human-Computer interaction*, Vol. 4, Issue ISS, Article 203, 2020.
- [11] 厚生労働省. 新型コロナウイルスの消毒・除菌方法について. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/syoudoku_00001.html. (2022年9月22日閲覧)
- [12] Nature. COVID-19 rarely spreads through surfaces. So why are we still deep cleaning? <https://www.nature.com/articles/d41586-021-00251-4>. (2022年9月22日閲覧)

自動運転車からの視線提示による交通事故低減の可能性の検討

張家銘* 戸田光紀* Xinyue Gui* Stela H. Seo† 五十嵐健夫*

概要. 様々な自動車メーカーや研究者が、自動車に目を追加することで、歩行者とのコミュニケーションを円滑にするアイデアを提示している。我々の以前の研究では、自動運転車に視線を提示できる目を付与することにより、歩行者の道路横断の判断の迅速化が可能であることを示した。本研究では、より重要な問題である "視線の提示によって交通事故を減らすことができるのか?" について検討した。歩行者が急いで自動運転車の前を横断しようとしている状況を考える。この時、車が歩行者を見ていないということは、車が歩行者を認識していないことを意味する。したがって、歩行者は道路を渡るべきではないと判断でき、潜在的な交通事故を回避することができる。我々は、モーターで駆動される物理的な目を付与した実験車両を製作し、それを撮影した仮想環境において被験者実験を行った。その結果、視線提示によって危険な道路横断を低減できる可能性があること、歩行者の主観的な安全感・危険感を高めることができることが分かった。また、視線を提示する自動運転車とのインタラクションにおいて、男性と女性の行動の間に差異があることも示唆された。

1 はじめに

自動運転車の課題の1つに、自動運転車と歩行者の間のインタラクションの問題が挙げられる [1]。以前の我々の研究では、歩行者が道路を横断するために立っているシナリオにおいて、自動運転車の目が歩行者の道路横断の判断をより早くすることを示した [7]。しかし、道路横断が少し速くなったり遅くなったりすることは、交通事故（けがや死亡など）ほど重大ではないといえる。交通事故を減らすこと（＝正しい判断をすること）の方が、早く判断できるようにすることよりもずっと重要である。本研究では、「視線の提示によって交通事故を減らすか」という研究課題への解答を試みる。

歩行者が急いで歩いて道路を渡るという重要な道路横断のシナリオを考える。車（目）が歩行者を見ていない場合、車は歩行者を認識していないことを意味するとする。そのため、歩行者は道路を渡るべきではないと判断し、交通事故の可能性を回避することができる。しかし、車（目）が歩行者を見ている場合は、車が歩行者を認識していることを意味する。そのため、歩行者は横断歩道を渡っても安全かどうかを判断することができる。

18名の参加者を対象に、視線提示の効果を調べる実証実験を行った。実車のプロトタイプを製作し、実際の交通環境での360度動画を撮影し、バーチャ

ルリアリティ環境に取り込んだ。4つの条件（図1）について、重要な道路横断のシナリオで、目のない車と目のある車を比較した。参加者はVRの中で歩行者となり、道路横断の判断を行った。参加者の判断の誤り率を計測した。

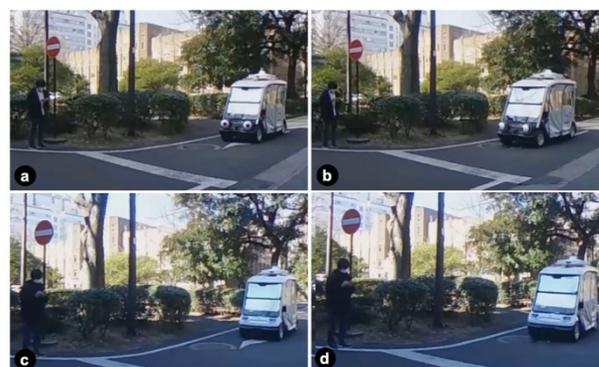


図1. 実証研究における参加者の視点。(a)安全な場合、(b)危険な場合、(c)安全な場合、(d)危険な場合。

最も注目すべき結果は、目が男性歩行者の潜在的な交通事故を減らす可能性がある点である。本論文では、実際のモーターで駆動される目のプロトタイプ設計と、正確なタイムラインによるシナリオ、さらに、目に対する性別の影響についての考察を行う。

本論文は、ACM AutomotiveUI 2022 で発表された論文の内容をまとめたものである [20]。

Copyright is held by the author(s).

* 東京大学, † 京都大学

2 関連研究

車両とドライバーの合図は、車両が接近した際に歩行者が道路横断の判断をするのに役立つ2種類の情報である[2]。しかし、自動運転車と歩行者のインタラクションは、人間が運転する車両とは異なる。近年、自動運転車と歩行者のインタラクションが広く検討されており、多くの研究が自動運転車と歩行者のインタラクションのための様々なコミュニケーション様式を提案している。例えば、自動運転車に外部ディスプレイを搭載することは、自動運転車と歩行者のインタラクションのための一般的なソリューションである [3, 4]。このディスプレイは、歩行者とコミュニケーションをとるために、テキストやシンボルを提示する[8]。また、自動運転車と歩行者のインタラクションのための外部ヒューマンマシンインタフェース (eHMI) で、光の異なる色やアニメーションについていくつかの研究が行われている[5, 6]。また、自動車に搭載された「目」[7, 8, 9, 10, 11] や「笑顔」[8, 12, 13] を用いて歩行者とインタラクションを行うなど、人間らしいコミュニケーションインタフェースが研究されている。Chang ら[7]は、自動運転車に「目」を搭載するコンセプトを提案し、「目」が歩行者の道路横断の判断をより早くすることを示した。Jaguar Land Rover [11]も同様のコンセプトを提案し、歩行者の自動運転技術に対する信頼感を調査した。さらに、研究[10]では、複数の視線方向を用いて車の進行方向を示すという概念を提案した。本研究では、先行研究とは異なり、歩行中の歩行者が横断歩道を急いで渡るといった重要な交通シナリオに着目し、視線方向を用いて車が周囲の歩行者を認識していることを提示している。

実環境で自動運転車による実証研究を行うことは困難である[14]。多くの研究が、3DCG を使用して、提案した eHMI コンセプトを提示し、VR 環境における自動運転車対歩行者のインタラクションを評価している[15, 16, 17]。しかし、VR 環境における3DCGの自動車モデルは、実環境における実車とは大きく異なる可能性がある。実環境における実車を用いた実証研究がいくつか行われている[18]。本研究では、実環境とVRの混合法によって研究を行った。同様の混合法は、自動運転車-歩行者間のインタラクションを評価するために使用された[14]。この混合法は、実世界とVR環境のバランスをとることができると考えている。

3 提案手法

本研究では、歩行者が急ぎ足で信号機のない横断歩道を渡ろうとする重要な道路横断のシナリオに着目した。これは交通事故が発生する典型的なケース

であり、調査[19]によれば、90%以上のドライバーが信号機のない横断歩道で歩行者が横断待ちをしても停止しないことが分かっている。このような場合、歩行者は危険な状況(=傷害や死亡)を回避するために、車両の合図だけを頼りに道路横断の判断をすることになる。我々は、[7]のような歩行者が立っていてこれから道路を渡るといった危険でないシナリオよりも、すでに歩行者が歩いて渡っているというこのシナリオの方がはるかに重要であると考えている。

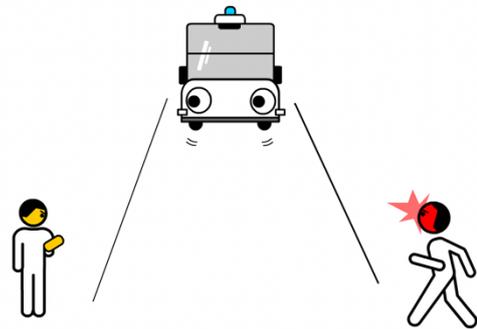


図2. 重要な道路横断のシナリオにおける自動車への視線。

図2は、急ぎ足の歩行者が道路を渡ろうとしたときに、突然、車の接近に気づくというシナリオである。歩行者は車が止まるかどうかを判断しなければならない。車が止まると判断した場合、歩行者は止まらずに横断歩道を渡り続けることができる。しかし、車が止まらないと判断した場合、歩行者は安全のために止まらなければならない。(1)車が止まっても歩行者も止まる場合はエラー、(2)車が止まらなくても歩行者が歩き続ける場合はエラーと、2つの判断をエラーと考える。

提案手法では、歩行者への視線提示によって、自動運転車が停止するかどうかの情報提示を行う。歩行者への情報提示方法としては、ディスプレイやライトなども考えられるが、以下の理由により視線を選択した。まず、手動運転車では、歩行車の84%が道路を横断するときに、運転者とのアイコンタクトを求めている[21]。よって、自動運転車においてもアイコンタクトを利用することで、歩行者が混乱することなく情報を受け取ることができる。また、ディスプレイやライトでは、個別の歩行者に対して異なる情報を提示することが難しいが、視線では、視線を向けている歩行者と視線を向けていない歩行者にそれぞれ異なる情報を提示することができる。

4 ユーザー実験

我々は、図3に示すように、ゴルフカート上に「制御可能な」目を設置した。目には3つの視線方向(左、

直、右)があり、矢印キー(物理ボタン)を使って手動で操作することができる。このカートは完全な自動運転機能を有しているが、安全性の観点から、ユーザー調査の際には手動で運転した。フロントガラスとサイドガラスにスモークガラスを貼り、自動運転であることを装った。



図3. 実車のプロトタイプ(左, 直, 右).

提案する歩行者用道路横断状況の実環境で360度動画を撮影した。図5に歩行者道路横断の環境と状況を示す。図4(a):この場所から歩行者は歩道を歩き始め、歩行者は完全な道路状況を見ることができない。図5(b):歩行者が歩き続けていると、突然、自動車が近づいてくるのが見える。図5(c):歩行者は歩き続け、道路(つまり車線)に出る前に道路を横断するかどうか判断する必要がある。



図4. 360度動画を実環境で.

映像には、(1)目なし車は歩行者に譲らないつもり(危険な場合)、(2)目なし車は歩行者に譲るつもり(安全な場合)、(3)目線あり車は歩行者を見ておらず、歩行者に譲らないつもり(危険な場合)、(4)目線あり車は歩行者に譲るつもり(安全な場合)の4つの車の挙動が収録されている(図5)。

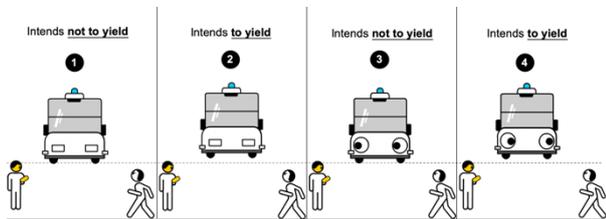


図5. 車の挙動に関する4つの条件.

ユーザースタディのために、18名(男性9名、女性9名、18~49歳)を募集した。ユーザースタディは、2022年2月15日から3月10日にかけて実施した。募集は外部に委託した。手続き費用と謝礼を含めた総費用は約2,160ドル(参加者1人あたり

約120ドル)であった。

実証研究では、動画をVR環境に配置した。参加者にはHMDを装着してもらい、今にも道路を横断しそうな歩行者になりきってもらった。参加者のタスクは、周囲の交通状況を観察し、道路横断の判断を行うことである。具体的には、横断歩道を渡る際に危険を感じた場合(衝突を避けるため)には横断を止める判断をする必要がある。しかし、安全に横断できる場合は、何もする必要がない。実証研究では被験者内法を用い、各参加者に40回(各条件10回、図6)道路横断の判断をさせた。4つの条件の順番は参加者ごとにランダムとした。

参加者を部屋に招き、360度動画が見られるVR環境を設置した(図6)。参加者に対しては、口頭で概要説明と詳細な説明を行った。評価自体は、指示・試行(10~15分)、VRタスク(15~25分)、アンケート(5分)の3部構成とした。30~40分程度で全ての評価が終了した。



図6. ユーザー実験の写真.

VR実験中の被験者への説明とタスク:

“VR、では急いでいる歩行者が道路(信号なし)を渡ろうとしているところです。まず、障害物(自転車や建物)に阻まれ、道路状況を完全に見ることができません(つまり、車が近づいているのが見えません)。しかし、その障害物を通り過ぎようと歩き続けていると、突然、車が近づいてくるのが見えます。車はあなたに譲るかもしれませんし、譲らないかもしれません。横断が危険と感じたら「STOP」横断の判断をする必要があります。”

本実験で実施するタスクの手順:ステップ1:「開始」ボタンを押してタスク開始。ステップ2:(危険を感じたら)「停止」ボタンを押して道路の横断を停止。

図7は、VR環境における車の挙動と歩行者の動きの関連する3つの場所を示している。a地点:タスクが開始されると、車と歩行者はこの場所から移動を開始する。この位置では歩行者は車を見ることができない。b地点:車と歩行者はb地点まで前進を続け、この地点に到達すると、歩行者は車を見ることができる。c地点:車と歩行者がこの場所に到着するまでに3秒のカウントダウンが提示される。参加者はカウントダウン終了直後に道路横断の判断をする必要がある。

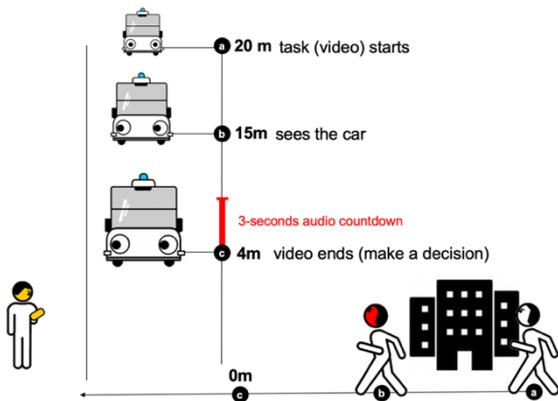


図7. VR環境における自動車の挙動と歩行者の動き.

道路横断の判断の誤りをカウントした. 条件 2, 4 (安全な場合) では, 停止判定をエラーとカウントした. しかし, 条件 1, 3 (危険な場合) では, 横断判定をエラーとしてカウントした.

5 結果

図8は, 全体のエラー率である. 全体のエラー率は, 目なし車で 50.56%, 目あり車で 29.44%であった. Paired t 検定による誤り率の解析の結果, 両車種間の差は統計的に有意であった ($p < 0.01$). このことから, 目あり車は一般的に目なし車よりも歩行者の道路横断の判断をより正しく行うことができることがわかった.

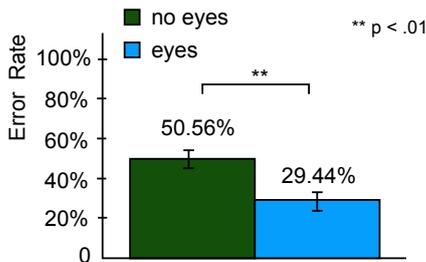


図8. 全体のエラー率. 目なし: 平均=50.56, SD=13.81, 目あり: 平均=29.44, SD=14.54.

図9は, 男性参加者のエラー率を示している. 危険のあるケースにおいて, 男性参加者のエラー率は, 目なし車で 48.89%, 目あり車で 17.78%であった. 対応のある t 検定分析の結果, その差は統計的に有意であった ($p < 0.05$). しかし, 危険のないケースでのエラー率は, 目なし車で 55.56%, 目あり車で 35.56%となり, 統計的に有意な差は見られなかった ($p > 0.05$). これらの結果は, 男性参加者にとって, 目は安全性には寄与するが, 効率性には寄与しないことを示している.

Male Participants

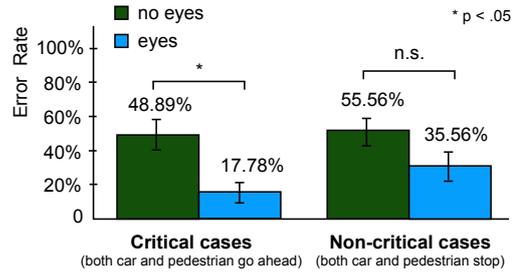


図9. 男性参加者のエラー率. 危険な場合 (Critical cases). 目なし: 平均 48.89, SD=31.4, 目: 平均 17.78, SD=17.87. 安全な場合 (Non-critical cases). 目なし: 平均=55.56, SD=26.03, 目あり: 平均=35.56, SD=27.44.

図10は, 女性参加者のエラー率を示している. 危険のある場合, エラー率は, 目なし車で 25.56%, 目あり車で 31.11%であった. 対応のある t 検定分析の結果, その差は統計的に有意ではなかった ($p > 0.05$). 一方, 危険のない場合では, エラー率は目なし車で 72.22%, 目あり車で 34.44%となり, 両者の間に統計的に有意な差 ($p < 0.05$) が認められた. これらの結果から, 女性参加者にとって目は効率性に寄与するが, 安全には寄与しないことがわかった.

Female Participants

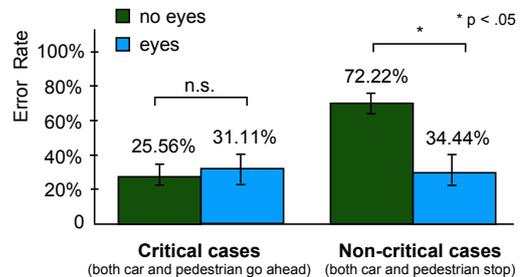


図10. 女性参加者のエラー率. 危険な場合 (Critical cases). 目なし: 平均 25.56, SD=20.68, 目: 平均 31.11, SD=29.77. 安全な場合 (Non-critical cases). 目なし: 平均値=72.22, SD=17.87, 目あり: 平均値=34.44, SD=31.67.

男性歩行者と女性歩行者に分けて異なるシナリオ (危険な場合と危険でない場合) の集計を行うとともに, ANOVA により性差 (=エラー率) を分析した. その結果, 男女間の差は統計的に有意であった ($F(7, 64) = 4.195, p = 0.0007$). このことから, 男性歩行者と女性歩行者は, 自動運転車と目対話する際に, 異なる道路横断行動をとることがわかった.

6 今後の課題

将来的には, 実際の自動運転車のシステム (AI) と目を統合し, AI の内部状態に応じて目の視線を制

御することを目指したいと考えている。例えば、さまざまな交通状況に対応するために、異なるアルゴリズムを開発して、異なる視線制御（＝異なる自動運転モード）を行うことが可能である。さらに、本研究では、実際の自動運転車と歩行者の対話において、危険な場合とそうでない場合で興味深い性差が見られたことから、歩行者の性別を認識し、実際の自動運転車の挙動を変更する技術の開発も有用であると考えられる。また、歩行者の性別だけでなく、子供や高齢者など、道路利用者によっても実際の自動運転車との対話方法が異なる可能性があると考え、異なる道路利用者についても調査する予定である。VRの評価については、VR環境においてユーザーは現実の交通事故の危険をどのように認識することができるのか、といった点について検証したいと考えている。

7 結論

目が歩行者の道路横断判断の効率化に役立つとした先行研究[7]とは異なり、本研究では、状況によっては目を使って交通事故を減らすことができる可能性があることを示したものである。我々は、視線を提示できる目を搭載した実車のプロトタイプを製作し、重要な道路横断のシナリオにおいて、目を搭載した車と搭載していない車を比較する実証実験を実施した。その結果、目があることで男性歩行者の潜在的な交通事故を減らすことができ（危険な場合）、女性歩行者の交通効率を高めることができることが示された（危険でない場合）。また、視線の方向によって、道路横断時における歩行者の「安心感」（車が自分を見ている）や「危機感」（車が自分を見していない）を高められることが判明し、本研究では、実際の自動運転車と歩行者のインタラクションにおいて、危険な場合（車も歩行者も先に進む）とそうでない場合（車も歩行者も止まる）の対応の性差について考察した。これらの知見は、今後の自動運転車-歩行者間インタラクションの開発において有用な知見と興味深い方向性を提供するものである。

謝辞

本研究は JST CREST JPMJCR17A1 の支援を受けたものである。また、車両は株式会社ティアフォーから提供されたものである。

参考文献

- [1] Daniel J. Fagnant, and Kara Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77: 167-181.
- [2] Stefanie M. Faas, Johannes Kraus, Alexander Schoenhals, and Martin Baumann. 2021. Calibrating Pedestrians' Trust in Automated Vehicles: Does an Intent Display in an External HMI Support Trust Calibration and Safe Crossing Behavior?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-17).
- [3] Yeti Li, Murat Dikmen, Thana G. Hussein, Yahui Wang, and Catherine Burns. 2018. To Cross or Not to Cross: Urgency-Based External Warning Displays on Autonomous Vehicles to Improve Pedestrian Crossing Safety. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18)*. ACM, New York, NY, USA, 188-197.
- [4] Karthik Mahadevan, Sowmya Somanath, and Ehud Sharlin. 2018. Communicating Awareness and Intent in Autonomous Vehicle-Pedestrian Interaction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 429, 12 pages.
- [5] Debargha Dey, Marieke Martens, Chao Wang, Felix Ros, and Jacques Terken. 2018. Interface Concepts for Intent Communication from Autonomous Vehicles to Vulnerable Road Users. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18)*. ACM, New York, NY, USA, 82-86.
- [6] Debargha Dey, Kai Holländer, Melanie Berger, Berry Eggen, Marieke Martens, Bastian Pfefing, and Jacques Terken. 2020. Distance-Dependent EHMI for the Interaction Between Automated Vehicles and Pedestrians. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Virtual Event, DC, USA) (AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 192-204.
- [7] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '17)*, pp. 65-73.

- [8] Chia-Ming, Chang, Koki Toda, Takeo Igarashi, Masahiro Miyata, and Yasuhiro Kobayashi, 2018, September. A Video-based Study Comparing Communication Modalities between an Autonomous Car and a Pedestrian. In Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18), pp. 104-109.
- [9] Yoichi Ochiai and Keisuke Toyoshima. 2011. Homunculus: The Vehicle as Augmented Clothes. In Proceedings of the 2nd Augmented Human. 3-6
- [10] Xinyue Gui, Koki Toda, Stela H. Seo, Chia-Ming Chang and Takeo Igarashi. 2022. "I am going this way": Gazing eyes on a self-driving car show multiple moving directions. In Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22).
- [11] Jaguar Land Rover Automotive Plc., 2018. The Virtual Eyes Have it. Retrieved June 19, 2022 from <https://www.jaguarlandrover.com/2018/virtual-eyes-have-it>
- [12] Hiroshi Hayakawa, Paul S. Fischbeck, and Baruch Fischhoff. 2000. Traffic accident statistics and risk perceptions in Japan and the United States. *Accident Analysis & Prevention*, 32(6), 827-835.
- [13] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '17), pp. 65-73.
- [14] Marius Hoggenmüller, Martin Tomitsch, Luke Hesperhol, Tram Thi Minh Tran, Stewart Worrall, and Eduardo Nebot. 2021. Context-based interface prototyping: Understanding the effect of prototype representation on user feedback. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1-14.
- [15] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. In Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications, pp. 262-274.
- [16] Uwe Gruenefeld, Sebastian Weiß, Andreas Löcken, Isabella Virgilio, Andrew L. Kun, and Susanne Boll. 2019. VRoad: gesture-based interaction between pedestrians and automated vehicles in virtual reality. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings, pp. 399-404.
- [17] Sebastian Stadler, Henriette Cornet, Tatiana Novaes Theoto, and Fritz Frenkler. 2019. A tool, not a toy: using virtual reality to evaluate the communication between autonomous vehicles and pedestrians. In *Augmented reality and virtual reality* (pp. 203-216). Springer, Cham.
- [18] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2016. Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), pp. 795-802. IEEE.
- [19] The Japan Automobile Federation (JAF), 2016. JAF finds over 90% of drivers ignore pedestrian crosswalks. Retrieved April 23, 2022 from <https://mainichi.jp/english/articles/20161002/p2a/00m/0na/015000c>
- [20] Chia-Ming Chang, Koki Toda, Xinyue Gui, Stela H. Seo and Takeo Igarashi, 2022, Can Eyes on a Car Reduce Traffic Accidents?. The 14th ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2022)
- [21] Šucha, Matúš. 2014. Pedestrians and drivers: their encounters at zebra crossings. In Proceedings of the Fit to drive: 8th International Traffic Expert Congress.

未来ビジョン

Society 5.0において、自動運転は重要な役割を担っている。近年、自動運転技術が高度に発達し、私たちは日常生活の中で自動運転車と出会い、接する機会が増えている。しかし、自動運転技術自体は完成に近づいているかもしれないが、そのような技術を受け入れる社会の準備が整っていない可能性がある。自動運転技術は、信頼性や安全性に関するさまざまな問題をもたらす「ブラックボックス」であり、道路利用者は自動運転車とどう接していいかわからないという問題がある。この点は、もっと慎重に検討されるべき重要なテーマだと考えている。

著者索引

Sayan Sarcar, 1

雨坂 宇宙, 87

五十嵐 健夫, 56, 63, 117

池松 香, 1, 15

石黒 勝彦, 72

入江 英嗣, 102

内原 正一, 102

遠藤 輝貴, 56

太田 彩, 38

尾関 沙羅, 72

小野 哲雄, 8, 46

尾原 和也, 110

加藤 邦拓, 15

門本 淳一郎, 102

川原 圭博, 79

岸野 泰恵, 110

岸本 諒大, 31

北村 喜文, 1

清藤 哲史, 63

金 太一, 63

Gui Xinyue, 117

日下部 完, 8

小池 英樹, 63, 94

小寺 正明, 72

坂井 修一, 102

坂本 大介, 8, 46

坂本 祐輝, 102

ジェフェルソン パルドムアン, 94

志築 文太郎, 87

白井 良成, 110

菅原 光, 110

杉浦 裕太, 15

杉本 雅則, 87

鈴木 湧登, 46

須藤 海, 23

Seo Stela H. , 117

染谷 隆夫, 79

高木 瑠名, 94

高嶋 和毅, 1

高橋 亮, 79

田代 優香, 63

舘 知宏, 23

谷道 鼓太郎, 23

張 家銘, 117

張 翔, 15

塚本 昌彦, 31

土田 修平, 31, 56

土田 太一, 1

寺田 努, 31

戸田 光紀, 117

鳴海 紘也, 23

西村 太一, 72

野間 裕太, 23

樋口 啓太, 72

Hwang Dong-Hyun, 63

福田 大翔, 102

藤田 和之, 1

水谷 伸, 110

宮藤 詩緒, 63, 94

柳沢 豊, 110

雪田 和歌子, 79

横田 知之, 79

脇田 建, 38

渡邊 拓貴, 87

日本ソフトウェア科学会 研究会資料シリーズ ISSN 1341-870X
No. 98 2022年12月14日発行

第30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 2022) 論文集