

令和4年度
博士論文

人間の身体の内的変化に着目した社会的つながりを
活性化させる環境知能システム

京都産業大学大学院先端情報学研究科
先端情報学専攻
博士後期課程3年生

086123
御手洗 彰

2023年3月

Doctoral Thesis

An environmental intelligence system that activates social connections
by focusing on internal changes in the human body

by

Sho Mitarai

Faculty of Frontier Informatics
Graduated School of Kyoto Sangyo University
March, 2023

Abstract— The purpose of this research is to realize an environmental intelligence system that activates social connections by focusing on internal changes in the human body. Internal changes in the body refer to changes in a person’s internal state that do not appear externally, such as biological signals that change with emotional states. Activating social connections refers to improving a person’s motivation for social activities (e.g., conversation, interaction, recreation, etc.). In other words, the goal is to improve people’s motivation to engage in social activities by means of an environmental intelligence system that focuses on internal changes in the human body.

Specifically, we conducted the following research to achieve this goal.

1. Internal changes in people during social activities and the extraction of their elements
2. Investigation of whether support by environmental intelligence systems can contribute to the activation of people’s social connections
3. Consideration of an interface that allows seamless input without interfering with a person’s social activities

The purpose of this study is to realize an environmentally intelligent system that activates people’s social connections, and to do so, it is necessary to understand how the system behaves to bring about changes in people’s behavior. In Study 1, we investigated the relationship between the activity and internal change in the social activities of multiple people and the extraction of elements necessary for activation. Specifically, we used a wargame as a verbal communication task between people in different positions with different goals, and investigated the relationship between the internal changes of players using skin electrical activity and the playing situation. The results of the analysis showed that the players of the wargame synchronized their internal changes during important events common to all players, and that players with a high winning rate said significantly more than other players, indicating that they were able to convey more information necessary for other players to make decisions. These results indicate that the presentation of appropriate information in appropriate situations is necessary for human behavior change (persuasion of other players in the experiment). Based on these findings, Study 2 investigated whether user support by the system can contribute to the activation of users’ social connections, assuming a situation in which an environmentally intelligent system is realized. Specifically, recreation was targeted as a multi-person communication task, and the goal of the study was to evaluate the system’s ability to improve the user’s motivation for recreation. In this study, a humanoid robot was used as a recreation assistant, and the motivation of both participants was investigated during a two-person recreation (a VR game based on ”Darumasan Koronda”). For the ”appropriate scene” and ”presentation of appropriate information” as settings of the environmental intelligence system obtained in Study 1, we implemented ”a scene in which one participant was in a disadvantageous situation during the game” and ”information about a trick to make the disadvantaged participant in an advantageous situation”, respectively. Specifically, the robot as an intermediary of the environmental intelligence system behaved according to the above settings, and the effect on the motive of the robot’s behavior was evaluated. The results of the evaluation experiment showed that the robot’s behavior according to the above settings significantly increased the motive for recreation (e.g., obsession with winning, entertainment value, etc.). The experimental results suggest that we were able to verify the effectiveness of the environmental intelligence system in activating human social connections by focusing on internal changes in the human body. In Study 3, we examined the interface as an annotation input for users, which is necessary to demonstrate the feasibility of the environmental intelligence system. In the systems that have been proposed so far, the methods used to collect user

information have mainly utilized human communication media. However, conventional methods have problems such as inhibiting real-world human social activities (e.g., users cannot use voice input interfaces while conversing with others, and users cannot use their fingers when their hands are occupied). Therefore, in this study, we focused on isometric contraction of muscles without joint motion to create an interface that enables seamless input even in various environments. The proposed interface was shown to be capable of inputting data without interfering with the main activities of humans (accuracy of 93.13%) under various real-world situations.

In this thesis, we have demonstrated the feasibility of an environmentally intelligent system that can activate social connections through studies 1 through 3. Although this research focused on entertainment, the activation of social connections has potential applications in various fields other than entertainment, such as education and medicine, and the findings of this research are an important contribution with the potential for their development.

Keywords: human robot interaction, werewolf game, hand gesture, electromyography, electrodermal activity

概要: 本研究では、人間の身体の内的変化に着目した社会的つながりを活性化させる環境知能システムの実現を目的とする。身体の内的変化とは、外観に現れない人間の内的状態の変化を指し、例えば情動状態に伴って変化する生体信号などが挙げられる。社会的つながりを活性化するというのは、人の社会的な活動（人との会話、交流、レクリエーションなど）への動機を向上させることを指す。つまり、人間の身体の内的変化に着目した環境知能システムによって、人の社会的な活動への動機の向上を実現することを目的とする。具体的には、この実現を目指し以下の研究を行った。

1. 社会的活動時の人の内的変化とその要素抽出
2. 環境知能システムによる支援が人の社会的つながりの活性化に貢献し得るかの調査
3. 人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討

本研究の目的は、人の社会的つながりを活性化させる環境知能システムの実現であり、そのためには、システムがどのように振る舞うことで、人の行動に変化をもたらすことができるのかを理解する必要がある。研究1では、複数人の社会的活動において、その活性と内的変化の関係性および活性化に必要な要素抽出を行った。具体的には、様々な目的を持った立場の異なる人同士の言語的コミュニケーション課題として人狼ゲームを用い、皮膚電気活動を用いたプレイヤーの内的変化と、プレイ状況との関係を調べた。分析の結果、人狼ゲームではプレイヤー間で共通した重要なイベントにおける内的変化の同期がみられ、勝率の高いプレイヤーはその他のプレイヤーより有意に発言数が多いことがわかり、他プレイヤーの判断に必要な情報を多く伝えることができていたことが理解できた。この結果から、人間の行動変容（実験においては他プレイヤーの説得）には、適切な場面における適切な情報の提示が必要であることがわかった。この知見を踏まえて、研究2の研究では環境知能システムが実現された状況を想定して、システムによるユーザ支援がユーザの社会的つながりの活性化に貢献し得るのかを調査した。具体的には、複数人コミュニケーション課題としてレクリエーションを対象とし、レクリエーションに対する動機の向上を目標として設定して評価した。本研究ではレクリエーション援助者としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者の動機について調査した。研究1で得られた環境知能システムの設定として「適切な場面」と「適切な情報の提示」について、それぞれ「一方の参加者がゲーム中に不利な状況下となった場面」と「不利な状況下にある参加者が有利な状況となるための仕掛けに関する情報」の提供を実装した。具体的には、環境知能システムの仲介役としてのロボットが、上記の設定に応じた振る舞いをとり、その際の動機に及ぼす効果を評価した。評価実験の結果、ロボットの上記設定に応じた振る舞いがレクリエーションの動機（勝利への執着やエンタテインメント性など）を有意に高めたことが示された。この実験結果から、人間の身体の内的変化に着目した人の社会的なつながりを活性化させる環境知能システムの効果を検証できたと考えられる。研究3では、環境知能システムの実現可能性を示すために必要なユーザのアノテーション入力としてのインタフェースの検討を行った。これまで提案されてきたシステムでは、ユーザの情報収集手法として、主に人間のコミュニケーションメディアを利用した手法が用いられてきた。しかし従来の手法では、実世界の人間の社会的活動を阻害するといった課題がある（他者との会話中に音声入力インタフェースを使用できない、両手が塞がった状態では手指を使った操作ができない等）。そこで本研究では、様々な環境下でもシームレスに入力可能なインタフェースを目指し、関節動作を伴わない筋肉の等尺性収縮に着目しインタフェースの構築を行なった。そして提案するインタフェースが実世界で起こり得る様々な状況下において、人間の主活動を阻害せずに入力可能（精度93.13%）であることを示した。以上本学位論文では、研究1〜3の研究によって、社会的なつながりを活性化しうる環境知能システムの実現可能性を示した。本研究では、エンタテインメントに着目して研究を行ったが、社会的なつながりを活性化することは娯楽以外にも教育や医療など様々な分野において応用可能性があり、本研究の知見はそれらの発展可能性を有する重要な貢献である。

キーワード: ヒューマンロボットインタラクション, 人狼ゲーム, ハンドジェスチャ, 筋電位, 皮膚電気活動

目次

第1章	はじめに	1
1.1	本研究の背景	1
1.2	本稿の構成	2
第2章	関連研究	3
2.1	人への介入を行う人工物に関する研究	3
2.1.1	ヒューマンロボットインタラクションに関する研究	3
2.1.2	人同士のコミュニケーションを仲介するロボットに関する研究	4
2.1.3	これまでのヒューマンロボットインタラクション研究のまとめ	4
2.2	複数人コミュニケーションを用いたゲームの観察を行う研究	4
2.2.1	人狼ゲームをプレイするエージェント	5
2.2.2	人狼ゲームプレイヤの思考過程分析	6
2.2.3	人狼ゲームにおける皮膚電気活動の分析	6
2.2.4	これまでの人狼ゲーム研究のまとめ	6
2.3	様々な利用文脈で利用可能な入力インタフェース	7
2.3.1	手を用いた入力モダリティ	7
2.3.2	足を用いた入力モダリティ	7
2.3.3	目を用いた入力モダリティ	8
2.3.4	口を用いた入力モダリティ	8
2.3.5	身体を用いた入力モダリティ	8
2.3.6	これまでの入力インタフェース研究のまとめ	9
2.4	本研究の位置付け	11
第3章	人-人間インタラクションの観察	12
3.1	人狼ゲーム	12
3.2	実験1: 人狼ゲームにおけるプレイ体験の分析	13
3.2.1	実験参加者	13
3.2.2	実験手順	13
3.2.3	SCL計測	14
3.2.4	結果	15
3.2.5	考察	17
3.3	実験2: 覚醒水準による分析	21
3.3.1	プレイヤのSCL分析	21
3.3.2	考察	22
3.4	実験3: 人狼ゲームにおけるハイライト分析	25

3.4.1	実験参加者	25
3.4.2	ハイライトの抽出方法	25
3.4.3	結果	26
3.5	本章のまとめ	28
第4章	ロボットを介した人-人間インタラクションの調査	31
4.1	研究手法	31
4.1.1	概要	31
4.1.2	実装環境	31
4.1.3	ゲームのルール	32
4.1.4	ゲームの操作方法	32
4.1.5	ヒューマノイドロボット	33
4.2	評価実験	33
4.2.1	実験環境	34
4.2.2	実験条件	34
4.2.3	アンケート	35
4.2.4	被験者	35
4.2.5	実験手順	36
4.2.6	結果	36
4.3	考察	37
4.3.1	仮説の検証	37
4.3.2	他のアンケート項目について	39
4.4	本章のまとめ	40
第5章	人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討	42
5.1	インタラクション阻害要因の調査実験	42
5.1.1	データ収集実験	42
5.1.2	分析	46
5.2	提案手法: ハンドジェスチャ列認識手法	49
5.2.1	偽陽性に対する堅牢性	50
5.2.2	学習手法	51
5.2.3	ハンドジェスチャ測定センサ	52
5.2.4	入力の分割	52
5.2.5	機械学習アルゴリズム	53
5.2.6	プロトタイプ	53
5.3	予備実験: 提案手法に適したハンドジェスチャの選定	54
5.3.1	データ収集実験	54
5.3.2	ハンドジェスチャ列の作成	56
5.3.3	機械学習アルゴリズム	57
5.3.4	特徴量の計算	57
5.3.5	結果	58
5.4	評価実験	58

5.4.1	被験者	59
5.4.2	把持条件とハンドジェスチャ列	59
5.4.3	データ収集と分類	60
5.4.4	実験手順	60
5.4.5	結果	60
5.4.6	考察	61
5.4.7	ハンドジェスチャと識別精度について	61
5.4.8	入力時間の妥当性	62
5.4.9	疲労の影響	63
5.4.10	ハンドジェスチャ列の設計	63
5.5	フィールド実験	63
5.5.1	概要	64
5.5.2	結果	65
5.5.3	考察	66
5.6	本章のまとめ	66
第6章	総合的な議論	67
6.1	本研究の成果とその意義	67
第7章	おわりに	69

第1章 はじめに

1.1 本研究の背景

あらゆる人工物がネットワークに接続され、人々の行動や環境情報を様々なセンサによって収集することで、人工物の自律的なユーザ支援を行う環境知能システムの実現が目指されている [115]. これにより、環境情報（環境の配備されたカメラ）を利用することで、ロボットの死角になる場所の情報を推論したり、人の動作（加速度、カメラ情報）を取得することで、人の行動認識が容易となったりするなど、これまでのロボット単体では実現できなかったタスクが可能となっている。

一方で、これらの情報を取得することで、ロボットのタスク支援だけでなく、人間の社会的な活動を支援することも可能であると考えられる。人間の社会的な活動を支援することはその人の健康に寄与する。例えば、高齢者の社会的なつながりの希薄さが、生活の質の低下や孤立死につながるものが問題として指摘されている [107]. また、地方における若者において人とのつながりが幸福感に影響すると考えられている [118]. 他にも、教育や仕事においても人とのつながりが活性化されることで学習効率や生産性の向上につながるものがわかっている [73, 11]. 環境知能システムの技術的基盤があればこのようなユーザの社会的つながりを支援するような応用が技術的に可能であるにも関わらず実現されていない。

そこで本研究では、社会的つながりを活性化させる環境知能システムの実現を目的とする。社会的つながりを活性化するというのは、人の社会的な活動（人との会話、交流、レクリエーションなど）への動機を向上させることを指す。人間の動機を向上させるためには、これまでの環境知能システムで扱われていた人間の身体の外的变化（加速度やカメラ情報）だけでは不十分であると考えられる。そのため、本研究では人間の身体の内的变化に着目する。身体の内の変化とは、外観に現れない人間の内的状態の変化を指し、例えば情動状態に伴って変化する生体信号などが挙げられる。つまり、人間の身体の内の変化に着目した環境知能システムによって、人の社会的な活動への動機の向上を実現することを目的とする。

具体的には、この実現を目指し以下の研究を行った。

1. 社会的活動時の人の内的変化とその要素抽出
2. 環境知能システムによる支援が人の社会的つながりの活性化に貢献し得るかの調査
3. 人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討

本研究の目的は、人の社会的つながりを活性化させる環境知能システムの実現であり、そのためには、システムがどのように振る舞うことで、人の行動に変化をもたらすこ

とができるのかを理解する必要がある。研究1では、複数人の社会的活動において、その活性と内的変化の関係性および活性化に必要な要素抽出を行う。具体的には、様々な目的を持った立場の異なる人同士の言語的コミュニケーション課題として人狼ゲームを用い、皮膚電気活動を用いたプレイヤーの内的変化と、プレイ状況との関係を調べる。この結果を踏まえて、2の研究では環境知能システムが実現された状況を想定して、システムによるユーザ支援がユーザの社会的つながりの活性化に貢献し得るのかを調査する。具体的には、複数人コミュニケーション課題としてレクリエーションを対象とし、レクリエーションに対する動機の向上を目標として設定して評価した。本研究ではレクリエーション援助者としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者の動機について調査する。研究3では、環境知能システムの実現可能性を示すために必要なユーザのアノテーション入力としてのインタフェースの検討を行う。これまで提案されてきたシステムでは、ユーザの情報収集手法として、主に人間のコミュニケーションメディアを利用した手法が用いられてきた。しかし従来の手法では、実世界の人間の社会的活動を阻害するといった課題がある（他者との会話中に音声入力インタフェースを使用できない、両手が塞がった状態では手指を使った操作ができない等）。そこで本研究では、様々な環境下でもシームレスに入力可能なインタフェースを目指し、関節動作を伴わない筋肉の等尺性収縮に着目しインタフェースの構築を行う。

1.2 本稿の構成

本論文は以下の構成から成る。第2章では、本論文と関連する研究を紹介し、これらに対する本論文の位置づけを示す。第3章では、複数人の社会的活動において、その活性と内的変化の関係性および活性化に必要な要素抽出を行う。具体的には、様々な目的を持った立場の異なる人同士の言語的コミュニケーション課題として人狼ゲームを用い、皮膚電気活動を用いたプレイヤーの内的変化と、プレイ状況との関係を調べる。第4章では、第3章の知見をふまえて環境知能システムが実現された状況を想定して、システムによるユーザ支援がユーザの社会的つながりの活性化に貢献し得るのかを調査した。具体的には、複数人コミュニケーション課題としてレクリエーションを対象とし、レクリエーションに対する動機の向上を目標として設定して評価した。本研究ではレクリエーション援助者としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者の動機について調査する。第5章では、環境知能システムの実現可能性を示すために必要なユーザのアノテーション入力としてのインタフェースの検討を行う。第6章では、本研究で得られた知見・成果をまとめる。第7章では、各章の外観を振り返り、本論文をまとめる。

第2章 関連研究

本研究は3つの研究に分かれており、まずそれぞれにおける関連研究を示す。まず人への介入を行う人工物に関する研究について述べ、人間の社会的なつながりを活性化させるような人工物の振る舞いがどこまで研究されているかを示す。次に、複数人コミュニケーションを用いたゲームに関する研究について述べ、人のコミュニケーションの観察からどのようなことがわかるか、本研究で使用する皮膚電気活動を用いることでどのような情報を得ることができるかを示す。最後に、様々な文脈で利用可能なインタフェースについて網羅的に紹介し、環境知能システムの入力として、様々な環境下で入力可能なインタフェースの実現のために、関連研究で解決できていない点を明らかにする。

2.1 人への介入を行う人工物に関する研究

人の支援を目的として介入を行う人工物に関する研究は様々な形式で行われている。本研究では、人間への働きかけとしてヒューマノイドロボット（以降、ロボットとする）と人間がどのようにインタラクションを行い、人の行動を変容させているかについてまとめる。まず、人間とロボットとのインタラクションを研究するヒューマンロボットインタラクション分野に関する研究について述べ、ロボットの振る舞いが人間に対してどのように影響するかをまとめる。次に、人同士のコミュニケーションを仲介するロボットに関する研究について述べ、ロボットの振る舞いが人同士のコミュニケーションにも影響しうることを示し、本研究の位置づけを明らかにする。

2.1.1 ヒューマンロボットインタラクションに関する研究

人間とロボットが共存する社会を目指してこれまでに多くの研究が行われてきた。まず、人とロボットのインタラクションの出発点としてロボットが人にどのように接近するべきかを調査する研究が行われている。Satakeらは単に人に近づいて声をかけるだけでは人とロボットのインタラクションは成立しないとしている [78]。ロボットは対話のモチベーションを持つ人を特定し、対話の意思を身体言語のような非言語で示す必要がある。Mizumaruらはマナー違反者（タバコのポイ捨て、歩きスマホ）に注意するためのヒューマノイドロボットの行動戦略について検討した [57]。その結果、対象者に刺激を与えないような友好的な態度よりも、ロボットが目の前に現れる接近戦略の方がマナー違反者が効果的に注意を聞き入れることがわかった。Kuzuokaらはヒューマノイドロボットが人とインタラクションする際の適切な立ち位置について検討した [42]。Vacentらはロボットが人に助けを求める際、より助力を得られるための礼儀戦略（politeness strategies）の調査した [87]。Sakamotoらは人-人-ロボットの3者におけるインタラクションを調査し、ロボッ

トの意見がロボットの印象だけでなく、人の印象にも影響することを確認しており [75], ロボットの行動によって人の行動や人が抱く印象を変えることが可能である. ロボットを人間のコミュニケーションの仲介役とすることで, より良い人間関係の構築を試みる研究もある.

2.1.2 人同士のコミュニケーションを仲介するロボットに関する研究

ロボットを介した人同士のコミュニケーションは, いくつかのアプローチで研究されている. ロボットを小さなグループで活用することで, 能力の異なるチーム間のタスク分配の公平性を維持 [22] やチームの生産性を向上 [11], 教育環境における学習効果を向上 [73] を目指す研究が行われている. これらの研究では, ロボットがグループメンバーの能力を考慮して適切なタスク分配や発話を促すことで, 多くの人が効率的に行動できるように支援した.

ロボットが集団に影響を与えるインタラクションには, 主に言語的インタラクションと非言語的インタラクションの2種類がある. 言語的インタラクションは, 人間同士の衝突を調停する行動を設計したり [38, 84], 会議において発言の少ない受動的な参加者に意見を促したり [83] するために用いられる. 非言語的インタラクションでは, 視線やアイコンタクトを用いて人間の発話を促すなど暗黙的なインタラクションに利用される [22].

上記のような複数人における作業や教育に加えて, レクリエーションを支援するロボットについても研究されている. 特に高齢化社会である日本の高齢者施設にロボットが導入されている [64, 63]. Hamada らは高齢者介護施設において, ロボットセラピーを用いたレクリエーションによるリラクゼーションを実現するための研究を行った [26]. Itai らはシナリオベースのロボットレクリエーションにより, 認知症高齢者の参加率やコミュニケーション頻度, ポジティブな感情を表現する時間が増加することを確認した [35].

2.1.3 これまでのヒューマンロボットインタラクション研究のまとめ

関連研究で示した人同士のコミュニケーションを仲介するロボットは, 本研究で目的とする社会的なつながりを活性化ことで同一である. 一方で, それぞれの研究では人同士のコミュニケーションの観察から症的にロボットの振る舞いを提案しており, まだ社会的なつながりを活性化しうる要素については明らかにされていない. そこで, 本研究では, 人の身体の内的変化に着目し, 複数人コミュニケーションの観察とともにその要素を明らかにすることを試みる. それを踏まえて, 実際にその要素を取り入れて, 人同士のコミュニケーションに対してロボットを導入させることで, 社会的なつながりを活性化しうるかを調査する.

2.2 複数人コミュニケーションを用いたゲームの観察を行う研究

複数人コミュニケーションは我々の社会の至るところで行われている. 一方で, 研究としてそれらを観察する場合, 何らかの課題無しに自然なコミュニケーションを成立させることは難しい. そのため, 自然なコミュニケーションを行うための課題としてゲームが挙

げられる。ゲームは結果（勝敗など）が明確に決まり、ルールベースで記述可能なため、明確に場面を分けてコミュニケーションの要素を分析可能である。ゲームを題材とする研究は、比較的新しいが、重要な研究である。特徴として、被験者がゲームに真剣に取り組むことで、その人間の本質的な部分をかいまみることができる。

本研究では、複数人コミュニケーションの課題として、人狼ゲームを用いた。人狼ゲームとは複数人で会話しながら行う自然言語ベースのパーティゲームであり、“だまし”、“嘘を見破る”、“協力”、“説得”など複雑なコミュニケーションを要し、その戦略は多岐にわたるゲーム性となっている。人狼ゲームに関する多くの研究が人狼知能プロジェクトによって行われている。人狼知能プロジェクトとは、不完全情報コミュニケーションゲームである人狼ゲームにおいて「人間と自然なコミュニケーションを取りながら人狼をプレイできるエージェントの構築」を目指しているプロジェクトである [122]。このような目標の実現のためには、エージェント設計や人工知能設計、認知心理学、自然言語処理、ヒューマン・エージェントインタラクション（以降、HAI）などといった複合領域の技術を要する。本章では、これまで人狼ゲーム研究でどのような試みがなされてきたかを紹介し、人狼ゲーム研究における本研究の位置づけを明らかにする。

2.2.1 人狼ゲームをプレイするエージェント

人狼ゲームをプレイする様々なエージェントが提案されている。人狼ゲームは、他プレイヤーの役職がわからない状態で村人陣営は人狼を処刑する必要があり、人狼陣営は処刑を回避する必要があるため、役職の推定が重要である。そのため、様々なアプローチでプレイヤーの役職を推定する研究が行われてきた。推定のパラメータとして、人狼ゲーム中の発言情報 [109, 121, 114] や身振り [112] などの身体的な情報や、占い結果などのゲーム中に得られる定量的な情報 [109] を用い、サポートベクターマシン [109] や強化学習 [108]、深層学習 [116, 106] を適用したエージェントが開発されてきた。

一方で、人狼知能プロジェクトの目標は人間らしい自然なプレイを行うエージェントの開発であるため、論理的・確率的な判断だけでなく、人間らしい判断も必要とされている。Nakamura ら [62] は、人狼をプレイするエージェントとして、心理学的モデルを加えたエージェントを開発した。彼らのアプローチでは、他プレイヤーの役職を確率的に推論するモデルをベースとして人間らしい振る舞いを実現するために、ゲーム中の特定の状況における他プレイヤーの役職推定のアンケートを取り、それを推論のパラメータとして用いることを試みた。しかし、彼らは自然なコミュニケーションを取るエージェントの実現にはさらなる人間性の追求を要すると述べており、人間の思考過程分析の必要性を示唆している。つまり、人とプレイするエージェントを考えた時、人がゲームを楽しむことも重要な要素となるため、エージェントが下す判断に人間目線の妥当性や論理性が求められる。従って、プレイ中に得られる様々な情報を用いたエージェントの判断・意思決定には、人狼プレイヤーの意思決定を探ることが必要である。

2.2.2 人狼ゲームプレイヤーの思考過程分析

人狼ゲームは情報が不完全な状態で人狼を探す必要があるため、複雑な思考を要する。そのような思考をエージェント化するために、人狼プレイヤーの思考を分析する試みが行われている。稲葉ら [105] は、プレイヤー同士の議論中における同調と反駁に着目して分析した結果、人狼役職のプレイヤーは「人狼を対象とする意見」に対して自然に反駁できるが、自然に同調することは難しいことがわかった。丹野ら [117] は、会議中にプレイヤーがつく嘘をどのように見抜くかを調査した。人狼ゲーム未経験者と経験者を対象として、嘘の手がかり信念にどのような差が見られるか調査した。その結果、人狼ゲーム未経験者は嘘の手がかり信念として体の動きや目線などの身体的特徴を重視する傾向があり、経験者は、発言量や発言内容を重視する傾向があることがわかった。伊藤ら [104] は人狼ゲームを発話プロトコル分析によって人狼ゲームにおけるプレイヤーの意思決定過程をモデル化した。彼らは人狼ゲームにおいて、始めから得ている情報が少ない「村人」と「狂人」役職のプレイヤーに着目して分析し、意思決定モデルを提案した。このモデルの特徴的な点は論理的合理性や経済的合理性によって決められない場合に、他プレイヤーに対する共感・信頼という非論理的な意思決定によって判断する点である。

これまでの研究では、人間の発話情報を中心に人狼ゲーム中における人間の意思決定過程を分析することで人狼プレイヤーの傾向・特徴や意思決定モデルを示してきた。一方で、人狼ゲームの特性上、プレイヤーは、他プレイヤーに余計な情報を与えないように発話内容に自身の思考を入れないようにしたり、時には嘘をついたりする。そのため、発話情報のみでプレイヤーの意思決定過程を明らかにすることは難しい。

2.2.3 人狼ゲームにおける皮膚電気活動の分析

先行研究では、人狼ゲームの特性を明らかにするために皮膚電気活動を用いた様々な試みが行われている。皮膚電気活動はプレイヤーの意思や思考を表す一つの指標である。人狼ゲームのような、騙し合いが必要なゲームでは、他プレイヤーに思考を読まれないよう表情を作ってしまう傾向があるため、外界に表出する情報から思考を読み取ることが困難である。皮膚電気活動は、不随意の測定系のため偽ることができず、客観的なデータを測定することができるため、嘘偽りのないプレイヤーのゲーム体験を得ることができる。これまでの研究では、人狼ゲームをプレイ中の SCL を解析することによって所属する陣営によってプレイヤーの SCL が有意に異なる（人狼陣営が村人陣営よりも心的な負荷が高く、SCL が高い）ことやプレイヤー同士が話し合いをする場面において、人狼陣営が勝利を確信している場合には、SCL が低くなること（定常状態により近づく、心的負荷が低い）など、ゲーム全体の情勢（どちらの陣営が有利か）における SCL の傾向が確認されている [111]。また、風間ら [120] は、皮膚電気活動からプレイヤーの意思決定過程を調査した。その結果、疑いが強まる程、皮膚電気活動の反応が検出されることがわかった。

2.2.4 これまでの人狼ゲーム研究のまとめ

先行研究では、ゲームを網羅的に見渡し、分析や解析を行った。また、分析では、発話やジェスチャなど外界に表出する情報しか扱っていない。人狼ゲームの特性上、嘘をつく

ことや相手に思考を読まれない等の要素があり、プレイヤーがどのような点に着目しているか、客観的に得ることはできていない。その結果、人狼ゲームのゲーム特性や分析において着目すべき点が明らかになっておらず、ゲームの分析に膨大な時間を要する。そこで本研究では、皮膚電気活動によって内的変化を得ることでプレイヤー目線で重要な場面や人狼ゲームのゲーム特性から、人間の行動を変容しうる要素を明らかにすることを試みる。

2.3 様々な利用文脈で利用可能な入力インタフェース

ユーザがコンピュータとインタラクションを行うための手段、入力インタフェースは人が持つ様々な入力モダリティによって実現されている。本章では、多様な入力モダリティについて体系的に整理し、どのような特性があり、どのような点で本研究で用いる人の筋肉による入力モダリティが優れているかを示す。

2.3.1 手を用いた入力モダリティ

手をベースとした入力インタラクションは多くの状況・場面で使われてきており、コンピュータへの入力インタフェースとして最も使用されてきた。現在においても、デスクトップコンピューティング環境では、キーボード・マウスを用いた手によるボタン押下入力、モバイルコンピューティング環境においては、静電容量センサによるタッチ入力主流である。

手は柔軟であり多様な動作が可能である。そのため、自然な手の動作によるコンピュータの入力はHCI分野における重要な課題である [40]。そのような課題の解決アプローチとしてハンドジェスチャ認識に関する研究が多くの研究者によって行われてきた。ハンドジェスチャ認識は、手を特徴的な形状へと変化させることで、コンピュータへ特定の入力を行うための技術である。これまでの研究ではハンドジェスチャ認識の入力手法として、手首の屈伸 [46, 53, 54] や手話 (アメリカン・サイン・ランゲージ) などのような手の形状変化 [70, 56, 36] が用いられてきた。測定手法は、ユーザの手の甲に装着し直接的に手の動作を測定する手法と、腕の前腕に装着して間接的に測定する手法がある。前者では、圧力センサ [54] や曲げセンサ [14]、加速度センサ [68] などが用いられる。また、公共の場におけるハンドジェスチャ入力の利用を考慮すると小さな動作のジェスチャが望ましい。そこで、指の動作のような微細な動きをハンドジェスチャとして利用するアプローチも存在する [3, 14, 77, 32, 52]。これらのアプローチでは、手の形状変化やユーザの手首の動作を高精度に識別可能であり、デバイスを手に把持せずともコンピュータとのインタラクションを可能にしている。

2.3.2 足を用いた入力モダリティ

足をベースとしたインタラクション技術は、ユーザの手が実世界の作業に占有されている時に、コンピュータへの入力として用いることができる [82, 12]。入力手法としてキック動作 [65, 27] や地面など平面へのタッピング、足の底面の特定部位に圧力をかける Foot planter based-gesture [67, 19] などが挙げられる。足をベースとしたインタラクション技術

は、それ単体で使用されるだけでなく、他の入力モダリティと組み合わせても使用される。例えば、精度の必要としないタスクにおいて、手を用いた入力の補助として有用であることが報告されている [66, 81, 98, 85]。他にも身体（主に胴体部分）を使ったジェスチャ[44]や目を使ったジェスチャ[23]などと組み合わせて、主となる入力モダリティの補助として足を使ったジェスチャが用いられている。

2.3.3 目を用いた入力モダリティ

目ベースとしたインタラクションの多くはメガネ型のウェアラブルデバイスを介してユーザからの入力を認識している。認識技術としてはカメラ [39, 91] や眼電図 [9, 18] などのセンサでユーザの目の動作を追跡することで実現しており、視線入力や注視による入力などを可能としている [69, 89, 43, 31]。また、眼球運動 [48, 92, 49] やまぶた [37] を使ったジェスチャを認識するアプローチも存在する。これらの技術の応用例としては、筋萎縮性側索硬化症（ALS）などの運動機能障がい者のための入力手段として用いられている [100, 71, 90]。

2.3.4 口を用いた入力モダリティ

認知アシスタントのような音声インタフェースは、近年の新しいインタラクション手法として注目されている。音声インタフェースはすでにスマートフォンやPC、スマートスピーカなど様々な製品 [5, 2, 55, 24] に搭載されており、運転中などの全身の入力モダリティが使用できない場合においても、入力可能である。既存の音声入力は音声で行うため、他人に聞かれるなどプライバシーの問題がある [59] が、これに対して、サイレントスピーチインタフェース [16] が提案されている。この技術の着想としては、ユーザがほとんど音声を発さずに入力を行い、コンピュータが音声内容を推定することで入力を行うというものである。提案されている手法としては、聞き取れない音声（inaudible voice）を使う手法 [20] やカメラ [88, 34]、筋電図 [94]、超音波 [17, 15] などのセンサを用いて身体の構造変化から音声内容を推定する手法がある。音声インタフェースは音声内容によって多様な入力が可能ではあるが、入力完了までの時間を要ことから、スイッチのオン/オフ操作には不向きであると指摘されている [21]。

他のアプローチとして、口をベースとしたインタラクション手法についても多くの研究がなされている。口を使ったジェスチャでは、息を吸う（sip）と吐く（puff）を組み合わせたジェスチャ[60]や顎・口腔における音響を利用した入力 [96, 95]、口腔内の歯の接触音を利用した入力 [58, 41, 6]、唇・舌の動作 [47, 13, 51, 99, 45]、顎の動作 [8, 4] など多様な入力手法が提案されている。これらの技術は、いずれもハンズフリー、アイズフリーのインタラクションを実現するもので、モバイル環境において有用な技術である。

2.3.5 身体を用いた入力モダリティ

オン・ボディインタラクション (on-body interaction)[29] は身体の様々な部位と連係してユーザの利用文脈に応じて、使用部位を変えて入力できるようにする技術であり、光学

トラッキング [28, 25] や静電容量センシング [79] など様々なアプローチが提案されている。オン・ボディインタラクションは、インタラクション空間が制限されるモバイル環境において、広いインタラクション空間を提供可能であることが利点であるとされている [50]。入力手法としては、身体をタッチパネルのように使用するもの [30, 97, 101] や、身体の内蔵構造の変化を認識するもの [50, 79] など様々である。最近では、ユーザの移動経路をずれを入力とする（車線を変更するように右や左にずれた結果を入力とする）技術も提案されている [61]。

2.3.6 これまでの入力インタフェース研究のまとめ

表 2.1 に上記に記した入力手法について、利用可能な場面と利用不可能な場面をまとめた。各入力モダリティにはそれぞれの特性に応じた様々な利点がある。例えば、足を用いた入力モダリティは、机上におけるコンピューティング環境において利用可能であり、マウス/キーボード操作の延長として使用することができる。口や目を使った入力モダリティは、運動障害のある人にとって有益である。本節で紹介したすべての入力モダリティはモバイル環境において重要であるとされており、特に手、足、身体のインタラクションを用いたモバイルインタフェースが多く提案されている。

一方で、すべての入力モダリティは実世界の活動によって占有されており、インタラクションが不可能な状況がある。これまでの研究では、他の入力モダリティとの組み合わせ（マルチモーダル）によってこの問題の回避を試みてきた。この場合、主要な入力モダリティを補助するために二次的な入力モダリティを使用するケースが多くある。例えば、キーボード操作を補助するために足のジェスチャを用いる [66]、主要な入力モダリティが利用できないときに簡単な入力を行うための手段として用いる [1]。しかし、これまでのアプローチでは、二次的な入力モダリティについて適切に議論されてこなかった。例えば、主要な入力モダリティが利用できない場合、二次的な入力モダリティも同様に利用できない可能性がある。このような状況下では、コンピュータとの入力が不可能となり、インタラクションが阻害されてしまう現状がある。そこで、本研究では入力モダリティが占有されている状況下であっても入力可能なインタフェースの開発を行い、様々な状況下で入力インタフェースが利用できない課題の解決を試みる。

表 2.1: Summary of input modalities

Modality	Input types	Usable situations	Unavailable situations	With other modalities
hand	wrist flexion [46, 54, 53]	mobile [46, 54, 53, 40, 70, 3, 14, 77] indoor space [14]	holding object	foot [66, 81, 98, 85]
	hand shape [40, 70, 56, 36]			
	finger gesture [3, 14, 77, 32]			
foot	foot movement [65, 27, 85]	mobile [65, 27, 98, 82, 12] indoor space [7, 44, 81] desktop computing [23, 67, 19, 66, 85]	walking [19] driving [19]	body [44] eye [23] hand [66, 81, 98, 85]
	tap [7, 44, 12]			
	plantar gesture [67, 19, 81] heel/toe rotation [98, 82]			
eye	eye tracking [91, 39, 9]	mobile [48, 92, 37, 39, 91, 9, 18, 43] desktop computing [37, 49] with impairments [18, 100] virtual reality (VR) [69, 31]	concentrating [86] driving	foot [23] head [48, 18, 43, 93] face [89]
	eye-lid [37]			
	eye movement [48, 92, 49, 100] gaze [69, 89, 43, 31]			
mouth	sip and puff [60]	mobile [6, 51] with impairments [58, 99, 45, 60, 13] music [47] eating [8]	talking eating running	-
	tooth contact sound [58, 41, 6]			
	tongue movement [99, 45, 86] jaw movement [8, 4] lip movement [47, 13, 51, 47]			
body	on-body gesture [28, 79]	mobile [28, 25, 50, 61]	holding object	foot [44]
	body touch [50]			
	spatial interaction [25] path of travel [61]			

2.4 本研究の位置付け

人間の社会的なつながりの活性化を実現している関連研究は多く存在するが、その要素については未だにわかっていない。本研究では、人間の身体の内的变化に着目して、人のコミュニケーションを調査することでそれが明らかにできると考えた。これより、人間の内的変化に着目した社会的なつながりを活性化させる環境知能システムを提案していることは本研究の新規性であるといえる。次に、複数人コミュニケーションの分析のための課題として人狼ゲームに着目した。特に、社会的なつながりを活性化させるためには人の思考や行動に対して何らかの変容を促す必要がある。本研究では、人狼ゲームの説得行動に着目して分析する。関連研究では、ゲーム中の会話や行動ログを分析することでプレイヤーの思考過程を分析する試みがなされていたが、人狼ゲームでは、相手を騙したり、状況に応じて説得方法を変えるといった特性上、ゲーム中のプレイヤーの振る舞い（発話内容、表情、リアクションなど）に必ずしも真意が含まれていないため、分析が難しい。そこで、本研究では皮膚電気活動を用いることで、心理的な負荷や疑い、緊張、闘争反応など外界に表出しない情報を計測できるため、人狼ゲームにおいてプレイヤーによって隠されていた内的変化を分析する。人狼ゲームの皮膚電気活動を分析することでプレイヤーの内的変化から、説得行動がどのような条件下で成功するのか調査することが本研究の新規性である。最後に、環境知能システムに対する情報収集には、人の身体の内定変化だけでなく、その人のアノテーションも重要となる。日常的に環境知能システムを用いることを考えると、様々な状況下で入力できることが望ましい。現状の入力インタフェース研究をまとめてみると、個々の研究ではそれぞれの状況下での入力を実現しているが、一つの入力インタフェースで様々な状況下で入力可能なものはない。本研究では、実世界の活動によって入力インタフェースのインタラクション阻害される原因を調査し、その原因を解消するための提案手法を考案することで、様々な状況下で入力可能なインタフェースを開発する。入力インタフェースのインタラクションが阻害される原因を調査すること、その結果をふまえて入力モダリティが占有されていても入力可能なインタフェースの開発を試みる点に、本研究の新規性がある。

以上をまとめると、本研究では人間の社会的なつながりの活性化を実現するために、人間の身体の内的变化に着目して、人間の行動を変容しうる要素を調査し、その要素を取り入れたロボットが人同士のコミュニケーションに介入することで、社会的なつながりを活性化しうるのかを調査する。さらに、そのようなシステムを実現するにあたり必要な環境知能システムの実現にあたり必要な様々な状況下で入力可能なインタフェースの開発を試みる。

第3章 人-人間インタラクションの観察

人間の社会的なつながりを活性化させるにあたり、システムがその人に対して変化をもたらす必要がある。つまり、人工物がどのように振る舞うことで、人の行動に変化をもたらすことができるかを理解する必要がある。そのために、本研究では人同士のコミュニケーションの観察から人の行動を変容しうる要素を調査する。具体的には人同士のコミュニケーションで人の行動を変容させる行動として説得に着目した。様々な目的を持った立場の異なる複数人の言語的コミュニケーションに着目し、それぞれがどのように他者の考えを変えようとしているかを調査する。コミュニケーションの課題として人狼ゲームを用い、ゲーム中における他者の考えを変える説得行動に着目して分析を行う。本章では、まず人狼ゲームにおけるプレイヤーのゲーム体験を調査することで、人狼ゲームにおける重要な場面やその状況下におけるプレイヤーの振る舞いを分析する。次に覚醒水準の指標として皮膚電気活動（SCL）を用いて、SCL 値の傾向や重要な場面におけるプレイヤーの言動を分析する。最後に人狼ゲームにおけるハイライトを SCL 値より抽出し、人狼ゲーム経験やプレイヤーごとの個別の傾向を分析する。

3.1 人狼ゲーム

人狼ゲームとは複数人で会話しながら行う自然言語ベースのパーティゲームである。正式なゲーム名は「汝は人狼なりや？」であるが、本稿では人狼ゲームと記す。簡単なルールを説明すると、ゲーム開始時にプレイヤーにはそれぞれ役職が与えられる。例として村人、人狼、占い師などがあり、他プレイヤーの役職を知ることができない。役職に応じて村人陣営（村人、占い師）と人狼陣営（人狼）に分かれそれぞれが勝利を目指すゲームである。ゲームの勝利条件は陣営によって異なり、村人陣営は誰が人狼役であるかを当てること、人狼陣営はゲーム終了まで人狼を当てられないことである。人狼陣営は村人のふりをしながら最後まで正体を見破られないようにふるまい、ゲーム中の会話をヒントに村人陣営は誰が嘘をついているかを探し、というゲームである。以上から、人狼ゲームは“だまし”、“嘘を見破る”、“協力”、“説得”など複雑なコミュニケーションを要し、その戦略は多岐にわたるゲーム性となっている。

近年、人工知能や認知科学、ヒューマンインタフェース、エンタテインメントなど多岐に渡る学術分野でゲームの研究が行われてきた [123, 110, 102, 103]。ゲームは結果（勝敗など）が明確に決まり、ルールベースで記述可能なため、プレイヤーの熟達、認知・思考過程の分析など様々な研究課題を有するためである。チェスや囲碁、将棋などの完全情報ゲームでは、ゲームをプレイするエージェントがプロプレイヤーに勝利するほど研究が進んでいる。一方で、人狼ゲームなどの不完全情報ゲームについては未開の部分が多い現状がある。特に人狼ゲームは、自然言語の会話で行われるため、ゲームのログなどの分析が難しい課

題である。また、人狼ゲームの相手を騙したり、状況に応じて説得方法を変えるといった特性上、ゲーム中のプレイヤーの振る舞い（発話内容、表情、リアクションなど）に必ずしも真意が含まれていないため、分析が難しい。

本研究では、人狼ゲームのゲーム特性の調査に皮膚電気活動（Skin Conductance Level: SCL）を採用した。SCLは一過性の正負の興奮（arousal）を検知する生体信号であり、精神性発汗量を指標として測定できる。人間の情動である興奮を定量的に計測可能であり、心理的な負荷や疑い、緊張、闘争反応など外界に表出しない情報を計測できるため、人狼ゲームにおいてプレイヤーによって隠されていた内的情報を分析可能である。

3.2 実験1: 人狼ゲームにおけるプレイ体験の分析

人狼ゲームのプレイ体験を分析するために、人狼ゲーム中におけるプレイヤーに対して、皮膚コンダクタンス水準 (SCL) を測定する実験を実施した。実験では、5人のプレイヤーで行う人狼ゲーム（以降より、5人人狼）を行い、計15回人狼ゲームを実施し実験中に映像・音声・SCLを記録した。

3.2.1 実験参加者

実験参加者は6人であり、全員が情報系の学科に所属する大学生もしくは大学院生であった。実験は、6人がそれぞれ最低10回はゲームに参加するように調整し、交代しながらゲームを行なった。参加者は普段から5人人狼をプレイしており、ルールの理解不足や不慣れによる参加者の心理的な影響はなかったと考えられる。

3.2.2 実験手順

図3.1のように円卓上に座った状態で人狼ゲームを試行した。実験では、映像・音声・皮膚電気活動を記録した。本実験で実施した人狼ゲームの進行手順と各手順の所用時間を以下に示す。尚、各手順や所用時間に関しては鳥海らの設定を参考にした [119]。

1. GM（ゲームマスター）がプレイヤーに役職カードを配る（1分）
2. 夜フェーズ（1分）
3. 会議フェーズ（4分）
4. 投票フェーズ（1分）
5. ゲームの勝敗が決まるまで、2～4を繰り返す

簡単なゲームの流れを説明すると、GMがプレイヤーに役職カードを配り、プレイヤー全員が自分の役職を確認する。この時、役職は実験参加者ごとに偏りがでないように選定した。夜フェーズでは、人狼による他プレイヤーの殺害と占い師による他プレイヤーの（人狼か否かの）役職確認が行われる。会議フェーズでは、プレイヤー全員が話し合い、投票フェーズで



図 3.1: 実験環境

Figure 3.1: Experimental Environment.

誰を処刑するか決める。投票フェーズでは、人狼と思しきプレイヤーに投票し、投票数が一番多いプレイヤーが処刑される。同票で最多票となった場合は、決選投票により決める。決選投票では、投票の対象となるプレイヤーは順に 20 秒間の弁明を行い、他プレイヤーはそれを聴聞し対象のプレイヤーに投票を行う。処刑されたプレイヤーが人狼であった場合は村人陣営の勝利、人狼でなかった場合は手順 2 (夜フェーズ) へと戻る。5 人人狼では、2 回の投票フェーズを経て人狼が生き残っていた場合、人狼陣営の勝利となる。どちらかの勝敗が決まるまでの手順を 1 試行として、本実験では計 15 試行した。SCL の計測時間は GM がプレイヤーにカードを配った時点から計測開始とし、GM の勝敗宣言を計測終了とした。本実験における役割と割り当てられる人数は以下とした。

- 人狼 (1 人) : 夜フェーズに他プレイヤーを殺害する
- 狂人 (1 人) : 村人として数えられ、人狼に加担する村人
- 占い師 (1 人) : 夜フェーズに他プレイヤーが人狼であるか否かを知ることができる
- 村人 (2 人) : 特殊な能力は持たない

1 試行で 10 分程度の時間を要するため、15 試行すると 150 分程度の時間を要する。そのため本実験では疲労による実験への影響を考慮し、5 試行ごとに 20 分の休憩を設けた (以降、1 試行を 1 ゲームとする)。従って、実験時間は合計で 210 分程度であった。

3.2.3 SCL 計測

実験では、参加者の手掌に SCL 測定センサの電極を装着することで、SCL の計測を行った。SCL の計測装置は Affectiva 社の Q Sensor を参考に実装したものを使用した。SCL

センサは、電極から微弱な電流を流すことで、手掌のコンダクタンスの変化をセンサ値として出力する。人体の特性上、皮膚のコンダクタンスの変化には個人差があり、実験参加者ごとの個体間変動を分析するため、実験中に得られたセンサ値を0.0~1.0の範囲で正規化した。ここで、SCL値0.0は安静時を含む実験中の最低値を表し、SCL値1.0は最高値を表す。

電極の装着位置について、参加者の利き手とは逆の手掌部に装着した。電極や計測装置は参加者のジェスチャなどの振る舞いを妨げることをないように図3.2のように装着・配置した。

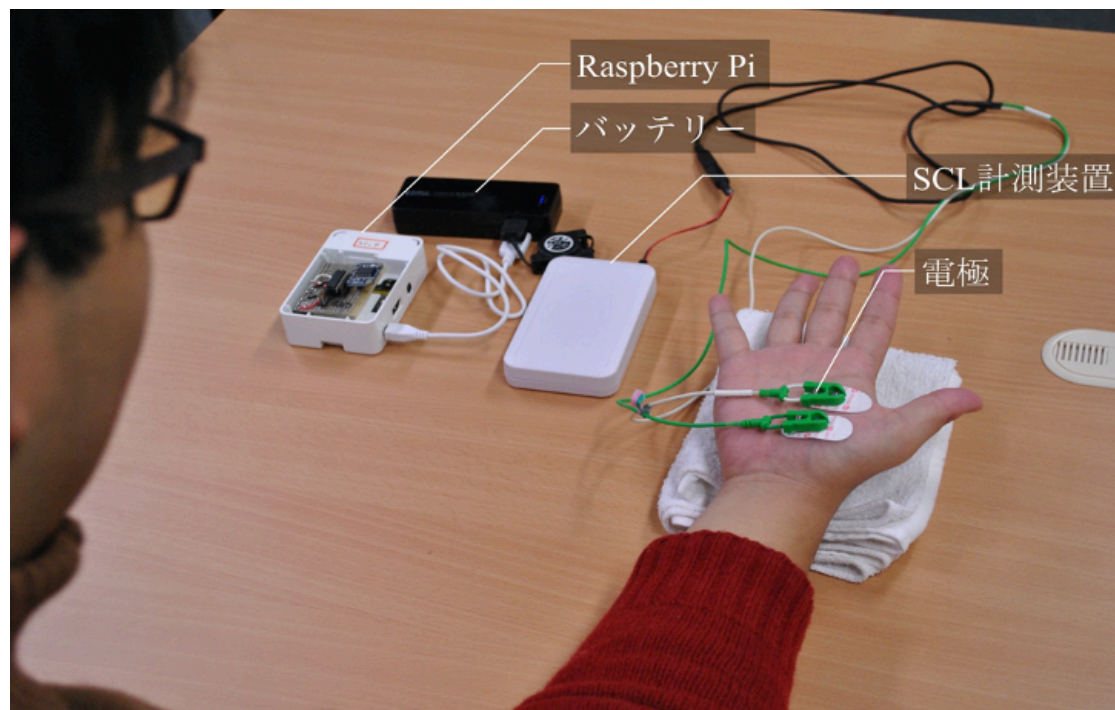


図 3.2: 計測システムと電極装着位置

Figure 3.2: Measurement System and Electrodes Placement.

3.2.4 結果

本実験では、プレイヤーにより自然な姿勢でゲームをプレイしてもらうため、行動に制限を設けるなど厳密な実験統制を行わなかった。そのため、被験者は自由に手や腕を動かせる状態であり、電極部分を強く押ししたりするなどして、5ゲーム分の測定データが欠損してしまった。そのため、本研究では15ゲーム中10ゲームを分析対象とする。

実験で行なった人狼ゲームの勝敗結果を表3.1に示す。表ではA~Fの被験者（以降、プレイヤー）について、それぞれの試合数と勝率を表している。表より、プレイヤーごとの勝率に差が生じたことがわかった。つまりプレイヤーA, B, Cは勝率の高かったプレイヤー、プレイヤーD, E, Fは勝率の低かったプレイヤーであることがわかった。また、本実験におけ

表 3.1: 各プレイヤーの試合数と勝率

Figure 3.2: Number of Games and Win Rate for Each Player.

プレイヤー	A	B	C	D	E	F
試合数	15	10	10	10	15	15
勝率 [%]	66.7	70.0	70.0	40.0	40.0	40.0

表 3.2: 会議フェーズ中の議論内容

Figure 3.2: Discussions During the Meeting Phase

	概要	平均所用時間
カミングアウト (CO)	占い師が名乗り出て、占い結果を伝える.	14 秒
推理	占い結果から、誰がどの役職かを仮定・推理する. この際、新たな占い師が CO する場合もあった.	1 分 34 秒
疑い	誰が人狼陣営かを疑う. 疑う理由や疑われたプレイヤーの弁明が議論の主旨となる.	1 分 36 秒
総括	今まで議論した情報を整理して、誰に投票するかを議論する.	36 秒

る会議フェーズでは、実験を通して議論の内容に傾向がみられたため、会議フェーズ中のイベントとして4つに分割した。それぞれの概要を表3.2に示す。

次に、実験で得られた人狼ゲームのSCLを図3.3に示す。これは各群（勝率の高かった/低かったプレイヤー）の会議フェーズ（初日）から投票フェーズ（初日）までの10ゲーム分のSCLの平均値を1秒間隔でプロットした図である。縦軸は、0.0~1.0の値に正規化したSCL値を表し、横軸は時間（秒）を表す。グラフの実線と点線は、それぞれ勝率の高かったプレイヤーと低かったプレイヤーのSCLの平均値を表し、ヒートマップは、全プレイヤーのSCLの平均値を表している。また、図3.3の上部は、会議フェーズと投票フェーズの時間幅を表している。会議フェーズは表3.2の平均所用時間で4分割した。ゲームのシステム上、最大2日目まで会議と投票が行われる場合があるが、多くのゲームが1日目で終了したため1日目の会議フェーズと投票フェーズに着目した。

図3.3のヒートマップより、SCLの高い領域（赤色領域）は人狼ゲームにおいて全プレイヤーに興奮が生じていたことを表す。4分割された会議フェーズと投票フェーズの領域におけるSCLを比較するため、それぞれのSCLの平均値を図3.4に示す。1要因参加者内分散分析により領域間の関係を分析したところ、主効果が認められた ($F(4, 20) = 8.04, **p < 0.01$)。次にLSD法による事後分析を行い、領域間の関係を明らかにしたところ、COイベントは、その他の会議イベント（推理、疑い、総括）よりも有意に高いSCL値であることが示された ($*p < 0.05$)。また、投票フェーズは、一部の会議イベント（疑い、総括）よりも高いSCL値であることが示された ($*p < 0.05$)。この結果より、赤色領域のうち、会議フェーズにおけるCOイベントと投票フェーズにおけるSCLが高い値であったことが示された。皮膚電気活動の原理[113]から、COイベントと投票フェーズにおいて、プレイヤーに共通して強い興奮が生じていたことが理解できた。さらに、COイベントと投票フェーズについ

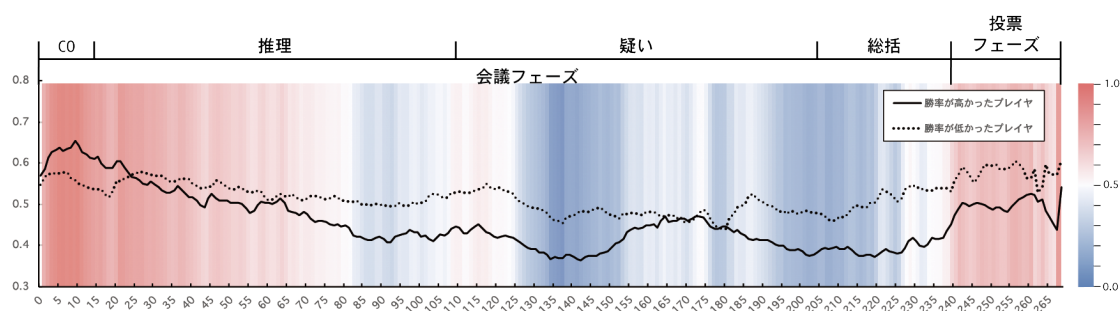


図 3.3: 会議フェーズ（初日）から投票フェーズ（初日）の時系列データの平均値（縦軸：正規化された SCL 値，横軸：時間（秒））

Figure 3.3: Mean SCL Values from the Discussion Phase (first day) to the Voting Phase (first day) (Vertical Axis: Normalized SCL Values, Horizontal Axis: Time (second)).

て、勝率の高かったプレイヤーと低かったプレイヤーを一要因参加者間分散分析によって比較してみると、CO イベントについては勝率の高かったプレイヤーが有意に高く ($**p < 0.01$)、投票フェーズについては勝率の低かったプレイヤーが有意に高い ($**p < 0.01$) ことが示された (図 3.5)。従って、勝率で分けたプレイヤー群においてより強い興奮が生じるイベントは異なることが示された。

次に、プレイヤーの行動としてゲーム中における発言数を分析するため、勝率で分けた群間の発言数について図 3.6 に表す。発言数は、実験映像から、発言内容を文字起こしを行い、一つの文章を 1 つの発言として数え上げた。勝率で分けた群における発言数の関係について一要因分散分析を行った結果、勝率の高い群の方が発言数が有意に高いことが示された ($*p < 0.05$)。

以上をまとめると、人狼ゲームにおいてプレイヤーに強い興奮を生じる場面には傾向があり、それが CO イベントと投票フェーズであることが示された。さらに勝率によって分けたプレイヤー群によって CO イベントと投票フェーズの SCL 値に差があることが示された。さらに、ゲーム中の発言数においても、勝率で分けたプレイヤー群で差があることがわかり、プレイヤーの内的状態とゲーム中の振る舞いのどちらにおいても差があることが理解できた。

3.2.5 考察

本研究では、人狼ゲームのゲーム体験を理解するために、皮膚電気活動 (SCL) を用いて、プレイヤーの情動反応を分析した。まず、CO イベントでみられた特徴として、必ずしも占い師のみが名乗り出るわけではなく、人狼陣営のプレイヤーが情報を攪乱するために嘘の情報を共有する場合があった。そのため、占い師が他プレイヤーを説得することが村人陣営の勝利に重要となる (人狼陣営にとっては嘘の情報を信じさせることが重要となる) イベントであった。図 3.4 より、プレイヤーの SCL 値がその他の会議フェーズのイベントより有意に高い結果から、CO イベント中はより全プレイヤーに共通して大きな興奮が生じていたことを示している。皮膚電気活動はプレイヤーの緊張 [111] や疑い [120] が生じる際に反応

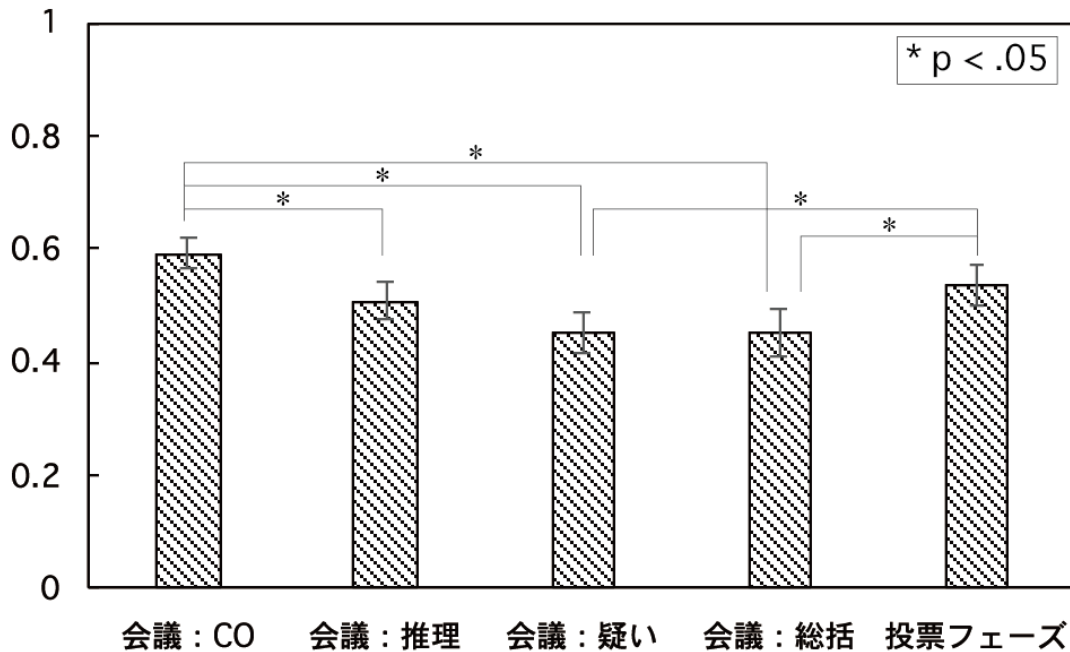


図 3.4: ゲーム全体の各場面における SCL の平均値 (縦軸: 正規化された SCL 値, 横軸: ゲーム場面, エラーバーは標準誤差を表す, CO はカミングアウトを表す)。

Figure 3.4: Average SCL for Each Event in the All Games (Vertical Axis: Normalized SCL Values, Horizontal Axis: Game Event; Error Bars Show Standard Errors).

を示す傾向があることがわかっている。CO イベントでは、占い師は他プレイヤーを説得するために働きかけ、人狼や狂人は占い師を騙り村人陣営を攪乱するような動きがほとんどのゲームでみられたため、緊張や疑いが入り混じるようなゲーム体験になっていたと考えられる。一方で、CO イベント内における発話内容や SCL の挙動はプレイヤー間で異なっていた。CO イベントでは CO した (占い師と名乗り出た) プレイヤーに対して「なぜその人を占ったか?」という内容が議論の主題であった。それに対する CO したプレイヤーの返答は「<プレイヤー名>が人狼だと後々面倒になりそうだから」といった理由を明示するプレイヤーと「適当に決めた」などの理由を明示しなかったプレイヤーの 2 種類に分かれた。実験時の勝率をもとに分析すると、理由を明示する返答をしたプレイヤーの勝率が高い結果となった。さらに CO イベント時のプレイヤーの SCL 平均値を分析すると、勝率が高かった (占い先の根拠について理由を明示した) プレイヤーの SCL 平均値の方が有意に高い結果となった (図 3.5)。実験記録映像を分析すると、理由を明示する場合は両陣営ともに、その理由から話を展開して他プレイヤーを説得したり (村人陣営)、騙そうとしたり (人狼陣営) する行動がみられたことから、興奮反応がみられたと考えられる。その興奮には、緊張や疑いのほかに、誰かを処刑する (あるプレイヤーに向けて投票させる) ために戦略を練る、といった闘争反応 [10] に類似した反応もあったのではないかと考えられる。一方、理由を明示しなかった場合 (勝率が低かったプレイヤー群) は次の話題へと移っていったため、CO イ

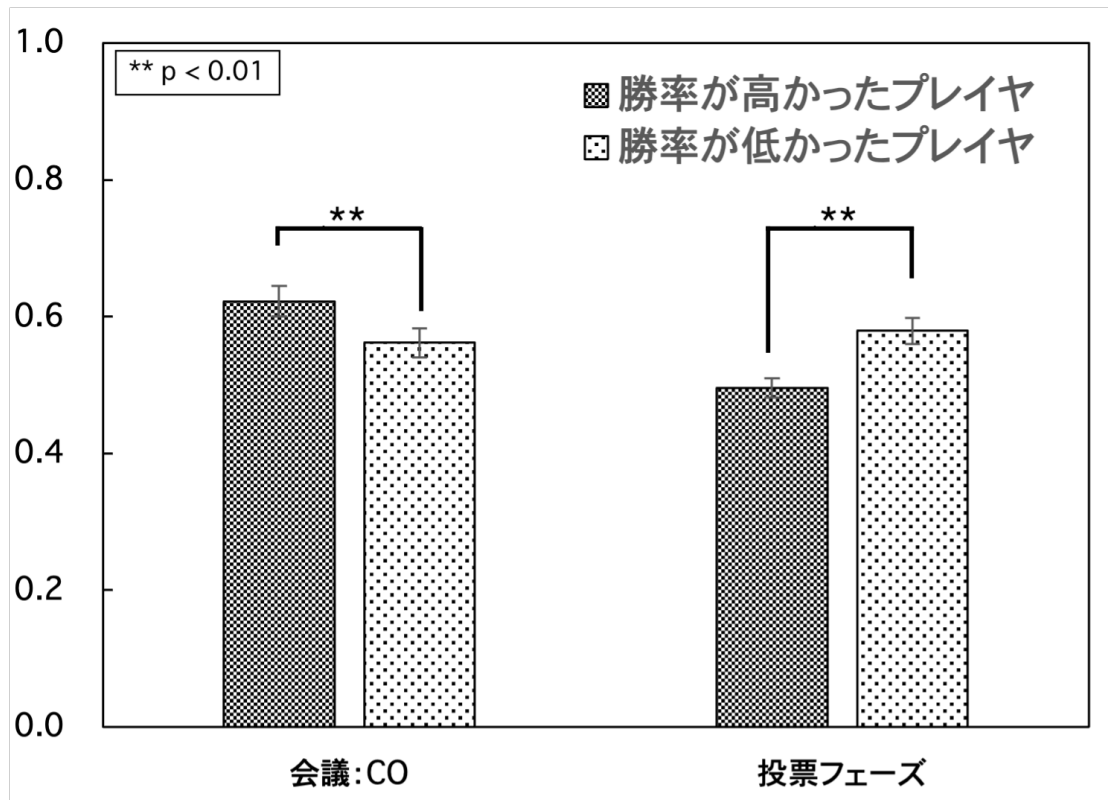


図 3.5: 勝率で分けた群における CO イベントと投票フェーズの SCL 値の比較 (縦軸: 正規化された SCL 値, 横軸: ゲーム場面, エラーバーは標準誤差を表す, CO はカミングアウトを表す).

Figure 3.5: Comparison of SCL in Coming Out Event and Voting Phase (Vertical Axis: Normalized SCL Values, Horizontal Axis: Game Event; Error Bars Show Standard Errors).

イベントにおける興奮反応が勝率が高かったプレイヤー群と比較して小さかったと考えられる。

次に投票フェーズでは CO イベントと同様に全プレイヤーに共通して大きな興奮が生じていた。SCL は発話などの能動的な行動も反応の一要素となるため、発話がほとんど行われないにも関わらず、会議フェーズの他イベントより高い値を示したことから、プレイヤーの内省によって大きな興奮が生じたことがわかる。その要因としてまず考えられることは、投票先に関する思考である。投票先について、村人陣営は誰が嘘をついているか、人狼だと疑わしいか、などを考えて人狼である確率が最も高いプレイヤーを決定する必要がある。人狼陣営は、誰が味方 (狂人、人狼) であるか、誰に投票が集まりそうか、などを考えて人狼以外に投票が集まるように投票先を決定する必要がある。そのため、プレイヤーの内省であっても CO イベントと同様に緊張と疑いが混じるようなゲーム体験となっていたと考えられる。また、勝率の高かったプレイヤーと低かったプレイヤーの SCL の平均値を比較してみると、勝率が低かったプレイヤーの SCL 値が有意に高いことが示された (図 3.5)。このことから、勝率が低かったプレイヤーは投票フェーズ時の内省によって大きな興奮が生じ

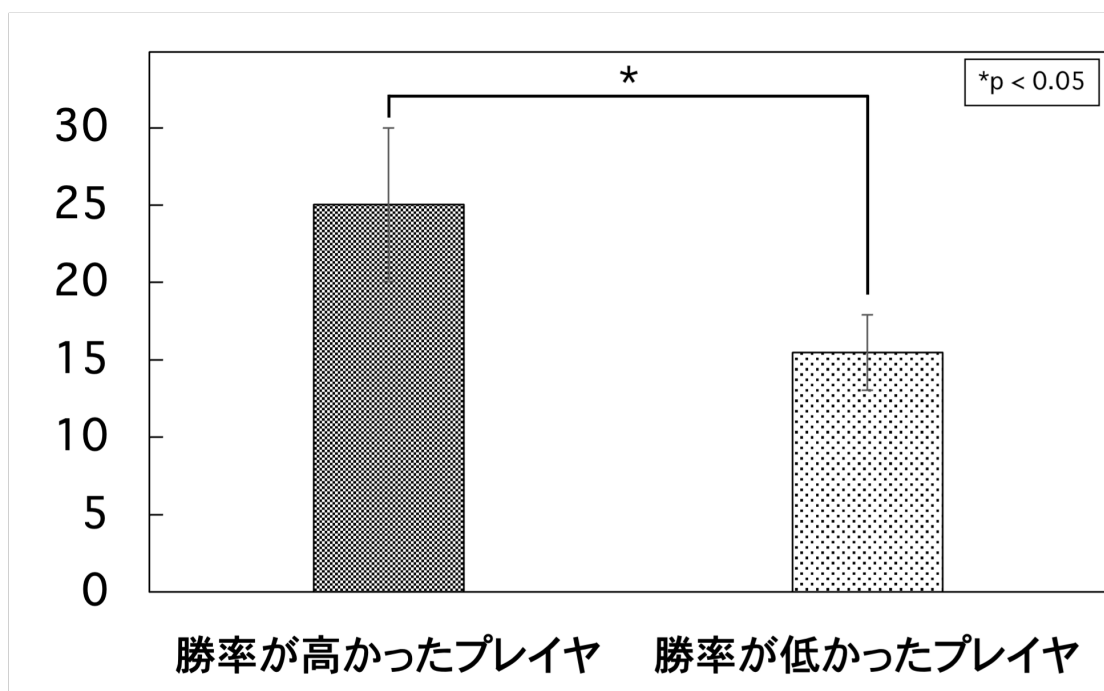


図 3.6: 勝率で分けた群における発言数の比較（縦軸：発言数，横軸：勝率で分けた群，エラーバーは標準誤差を表す）。

Figure 3.6: Comparison of Utterance Count (Vertical axis: Utterance Count, Horizontal Axis: Divided Group of Winning Rate; Error Bars Show Standard Errors).

ていたことがわかった。この要因について、実験記録映像からプレイヤーの発言を分析してみると、会議フェーズ終了時点や投票直前において投票先を決めかねている発言が確認できた。一方で、勝率の高かったプレイヤーは会議中に人狼を仮定するような発言がみられ、人狼（投票先）についてある程度目星をつけながらゲームを進行していたことがわかった。つまり、勝率の高かったプレイヤーは会議フェーズから徐々に投票先を絞り込むような推測をしており、勝率の低かったプレイヤーは会議フェーズ終了後に集まった情報から推測をする、といった投票先の考え方について群間で異なっていたと考えられる。

また、ゲーム中のプレイヤーの行動として発言数を分析すると、勝率が高かったプレイヤーの方がゲームを通して発言数が多いことが示された。これは、COにおけるプレイヤーのSCLの挙動と一致する。つまり、勝率が高かったプレイヤーは、CO時に理由を明示したり、会議フェーズのそのほかにおいても積極的に発言をしていたということである。逆に、勝率が低かったプレイヤーはCO時に理由を明示しないなど、発言数を抑えていた。本研究の範囲内では、CO時の理由の明示や積極的な発言が勝率を高めるために重要であったかについて断ずることはできないが、少なくとも両群においては、プレイスタイルや戦略の違いがあり、それがSCLや発言数に影響をしていたことは確かである。具体的には、勝率が高かったプレイヤー群は他者に情報を開示し、議論を自身の有利な方向へ誘導する戦略を取り、勝率が低かったプレイヤー群は情報をあまり開示せず、敵へ情報を提供しないような戦略を取っていたのではないかと考えられる。以上のように、SCLと他の情報を組み合わせ

せることで、より詳細なプレイヤーの行動分析が可能であることが理解できた。発話以外にも視線やジェスチャなどと SCL を組み合わせて、プレイヤーの行動や思考を分析することは今後の課題である。

人狼ゲームのコミュニケーションが行われる会議フェーズは分析において重要視されていることは先行研究 [104, 120, 111] から明らかであるが、時間が限られている会議フェーズの中で具体的に何に着目すべきかはこれまでわかっていなかった。本研究の分析から会議フェーズを 4 つのイベントに分割したところプレイヤーが着目しているイベントが明らかになった。また、会議フェーズ中の SCL を時系列データ (図 3.3) をみると、全プレイヤーの SCL 平均値は、CO イベントと投票フェーズにおいて高い値を示すことがわかった。この SCL の高さは興奮反応、つまり着目する場面や勝負を仕掛けていることであると考えられる。これらの分析によって、人狼のゲーム特性として、プレイヤーが着目する場面を SCL によってを定量的に示すことができた。一方で、プレイヤーの発話内容や人狼の推測の仕方についてはプレイヤー間で異なっており、SCL を比較するとプレイヤー間で違いがみられた。具体的には、CO イベントでは勝率の高かったプレイヤーが高い SCL 値を示し、投票フェーズにおいては勝率の低かったプレイヤーが高い値を示した。人狼ゲームでは、プレイヤーがそれぞれ持っている情報 (役職により得られる情報など) に違いはあるが、会議のフェーズにおいてプレイヤーから発せられる情報は同一である。しかし、投票先や弁明の言動と SCL を分析すると同じ情報量でも立場に違いがあることで異なる行動や SCL の反応を示すためであると考えられる。一方で大局的な傾向でみると、SCL の反応は一致していることがわかった。人狼ゲームはその性質上、陣営によって分かれており、それぞれの陣営が勝利を目指すため、プレイヤーによって情報の捉え方が違うのは当然の結果であると言える。しかし、プレイヤーによって考え方、捉え方が違うのにも関わらず、プレイヤー全体の大局的なゲーム体験は一定の傾向を持っていることは本研究で実験的に示せた。これは本研究の人狼ゲームのゲーム特性に対する貢献である。

3.3 実験 2: 覚醒水準による分析

3.3.1 プレイヤーの SCL 分析

実験結果から初日の会議における CO に着目して強い人狼役職プレイヤーと弱い人狼役職プレイヤーの SCL 値の類似性について分析する。

SCL 値の類似性の指標として、動的タイムワーピング (DTW: Dynamic Time Warping) 距離 [76] を用いた。DTW 距離はサンプリング数の違いや時間軸上のずれを許容して、2 データ間の距離を算出できるため、SCL 値のような時系列データの類似性の指標として適している可能性がある。DTW 距離は距離のため、DTW 距離が小さいほど類似性が高く、大きいほど類似性が低いと判断できる。

事前に行った実験で SCL 値の類似性が大きい例と小さい例を 2 例用意し、DTW 距離の有効性を確認した。結果の SCL 値のグラフと DTW 距離を以下に示す。図 3.7, 3.8 は、類似性の高い SCL 値のグラフと類似性の低い SCL 値のグラフを示している。SCL 値間の DTW 距離 d を計算すると、図 3.7 は $d = 0.94$ 、図 3.8 は $d = 2.04$ であった。このように、類似性の高い DTW 距離と低い DTW 距離には差がみられ、類似性が高いほど DTW 距離

は小さく、低いほど DTW 距離は大きくなるのが理解できる。従って、SCL 値の類似性の指標として DTW 距離を用いることは適切であると考えられる。

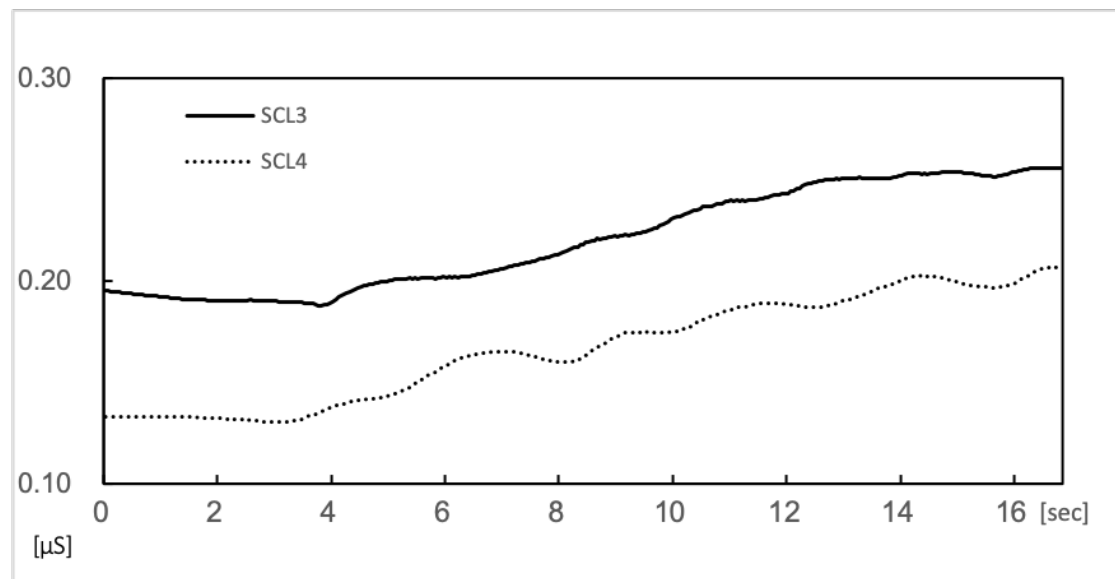


図 3.7: 類似性の高い SCL 値 ($d = 0.94$) の挙動

DTW 距離を指標値として、有効なデータ 10 ゲーム分で DTW 距離を求めた。求めた DTW 距離を 3 群に分けた。3 群の内訳は、「強いプレイヤー同士の DTW 距離」(以下, 1 群), 「弱いプレイヤー同士の DTW 距離」(以下, 2 群), 「強いプレイヤーと弱いプレイヤーを組み合わせた DTW 距離」(以下, 3 群) である。1-3 群に対して、一元配置分散分析を行い群間の関係を明らかにした (図 3.9)。その結果、1 群は 2 群よりも DTW 距離が小さいことが有意水準 1% で示された ($F(1, 12) = 36.56, p < .01$)。1 群は 3 群よりも DTW 距離が小さいことが有意水準 10% で示された ($F(1, 31) = 3.80, p < .10$)。2 群は 3 群よりも DTW 距離が小さいことが有意水準 1% で示された ($F(1, 29) = 20.65, p < .01$)。

結果を整理すると、3 群間の DTW 距離の関係は、図 3.9 に示すように 2 群 < 3 群 < 1 群となった。よって、各群の類似性の関係は、強いプレイヤー同士の SCL 値の類似性 (1 群) が一番高く、弱いプレイヤー同士の SCL 値の類似性 (2 群) が一番低い結果となった。また、強いプレイヤーと弱いプレイヤー間の SCL 値の類似性 (3 群) は強いプレイヤー同士の類似性 (1 群) よりも低い結果となった。つまり、強いプレイヤーの SCL 値は一定の傾向を持ち、強いプレイヤーと弱いプレイヤーの SCL 値は異なることが分かった。これらの結果から強いプレイヤーは弱いプレイヤーと比較して、CO 時の SCL 値の挙動に関して、傾向を持つことが示された。

3.3.2 考察

前章の結果より、強いプレイヤーは弱いプレイヤーと比較して SCL の挙動に傾向を持つことが示唆された。本章では、強いプレイヤーのどのような行動が傾向としてみられたかを SCL の特性と実験時の言動から考察・議論する。最後に、みられた傾向と勝因の結びつきにつ

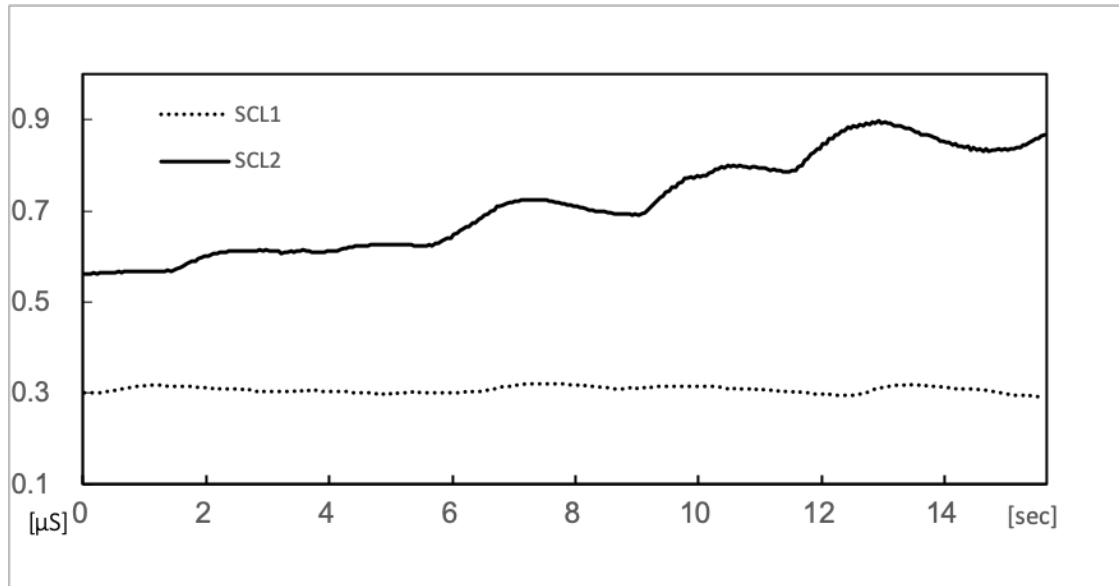


図 3.8: 類似性の低い SCL 値 ($d = 2.04$) の挙動

いて議論して、人狼ゲームにおける強さの要因について考察する。

SCL 値の傾向

前章の結果より、強いプレイヤーは SCL の挙動に傾向を持つことが示された。実験で測定した強いプレイヤーと弱いプレイヤーの代表的な SCL の波形を図 3.10, 3.11 に表す。図の横軸は CO 開始を 0 秒とした経過時間であり、縦軸は SCL 値である。縦軸の基準値は被測定者によって個人差が生じるため、波形に着目して考察する。図 3.10 をみると、両波形ともに SCL 値の上昇と下降を繰り返すような波形となっている。一方、図 3.11 は片方の波形は上昇した後、下降していないことから継続的に反応を示しており、もう片方はほとんど上昇や下降の反応を示していないことがわかる。

ここで覚醒水準から、強いプレイヤーと弱いプレイヤーの違いを考察する。覚醒水準では、ベストなパフォーマンスを発揮するためには、適切な水準を保つ必要があり、水準が下がり過ぎても上がり過ぎてもパフォーマンスは落ちるとされている。図 3.12 に強いプレイヤーと弱いプレイヤーの覚醒水準の比較を示す。横軸は強いプレイヤーと弱いプレイヤーの各群、縦軸は CO 時の SCL 平均値を示している。図より弱いプレイヤーは $0.2\mu\text{S}$ 付近と $0.7\mu\text{S}$ 付近のグラフの両端に多くのプロットがあることに対し、強いプレイヤーは $0.4\mu\text{S}$ ~ $0.5\mu\text{S}$ 付近に多くのプロットがあることがわかる。つまり、弱いプレイヤーは覚醒水準が下がり過ぎ、もしくは上がり過ぎていること、強いプレイヤーは適切な水準を保っていることが理解できる。

CO 時のプレイヤーの言動と SCL 値の関係

CO 時のプレイヤー同士の会話をプロトコル分析を行なった結果、強いプレイヤーと弱いプレイヤーの CO の捉え方の違いがみられた。強いプレイヤーは、他のプレイヤーに「○○君は誰

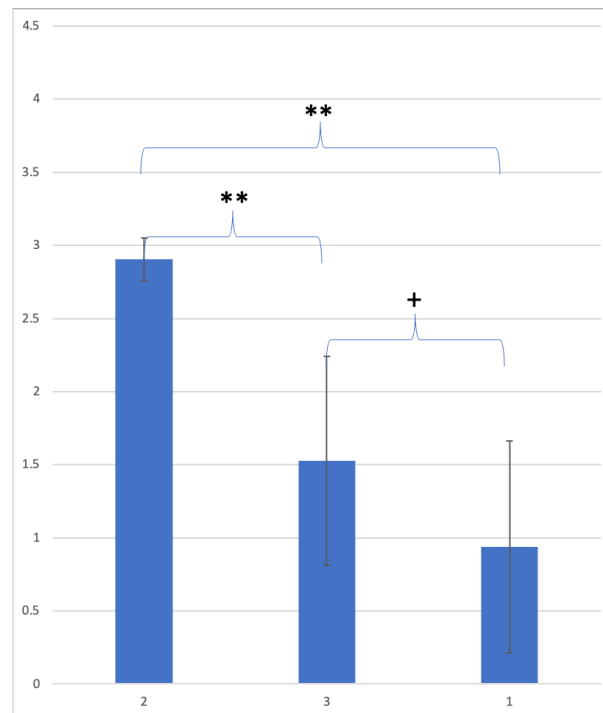


図 3.9: 各群における DTW 距離の平均値の比較

さしてた？」や「なんでその人占った？」などの質問（原文ママ）があった場合、「〇〇君に黒出した」や「〇〇君が人狼だと後々面倒になりそうだから」（原文ママ）など理由を明確にして答えている。しかし、プレイヤー F は同じような質問をされた時、「誰かに黒出した」、「んー適当」（原文ママ）など自身が行った行動にも関わらず、理由が明確でない回答が多い。

強いプレイヤーと弱いプレイヤー間の発言意図の明確性が差としてあったと考えられる。丹野らの研究 [117] も経験豊富なプレイヤーは嘘を見抜く手がかりは、発言に重きをおいていることから、強いプレイヤーは発言の意図に明確性、正当性を持たせることで周りへの説得を行っていると考えられる。

CO 時における人狼プレイヤーの強さの要因

前節までは、強いプレイヤーの行動を皮膚電気活動と言動の分析から考察した。本節では、それらの結びつきを議論することで、プレイヤーの強さの要因について論じる。

CO は初日の会議に必ず発生しており、占い師による占い結果が開示される等、最初の戦略が構築されるイベントである。図 3.12 に示された覚醒水準の高まりから、強いプレイヤーは CO 時に高いパフォーマンスを発揮できる状態であった。そのため情報を素早く処理し、戦略を練る、発言内容やタイミングを考える等、勝つために必要な行動を行っておりそれらが有利に働いていたことが示唆される。

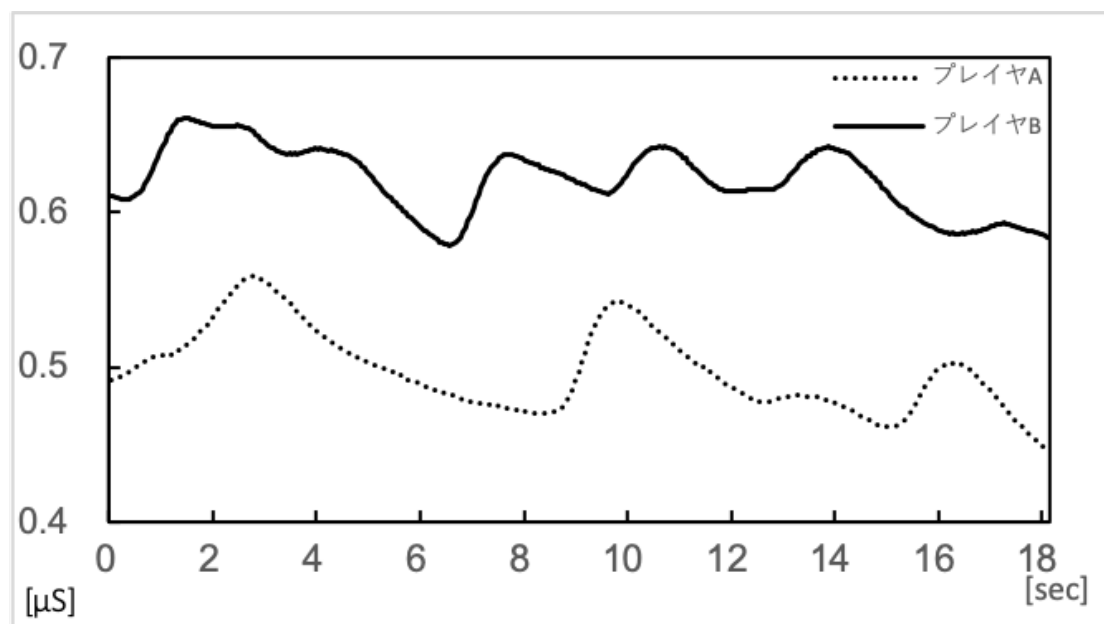


図 3.10: 強い人狼役職プレイヤーの SCL 値の挙動

3.4 実験3: 人狼ゲームにおけるハイライト分析

人狼ゲームのプレイヤー視点での重要な場面を分析することでゲームの特性を明らかにするために、人狼ゲーム中のプレイヤーの皮膚コンダクタンス水準 (SCL) を測定する実験を実施した。実験では、計 8 回の 5 人人狼を実施し、複数のカメラやマイクを用いて映像・音声を記録し、各プレイヤーに SCL 測定装置を装着させそれぞれ記録した。その他の実験手順に関しては実験 1 と同様の手順にて行った。

3.4.1 実験参加者

実験参加者は実験 1 とは異なる男性 7 人であり、全員が情報系の学科に所属する学生もしくは教員であった。7 人のうち一部のプレイヤーが交代をしながら 5 人人狼を行なった。本実験では、それぞれの参加者について、A~G の記号で表す。人狼ゲームの経験が豊富である参加者が 3 名 (A, C, G)、人狼ゲームの経験は数回程度の参加者が 4 名 (B, D, E, F) いることを本人の自己申告によって確認した。いずれの参加者も人狼ゲームのルールについては問題なく理解しており、ルールの理解不足による参加者の心理的な影響はなかったと考えられる。

3.4.2 ハイライトの抽出方法

ハイライトの抽出方法として、本研究ではそれぞれのプレイヤーにおける各試行ごとの SCL 最大値を抽出した。SCL は一過性の興奮を検出し、SCL 値が高い程より覚醒度の高い情動状態であることを示す。そのため、1 つの試行内において覚醒度の高まりの最大値

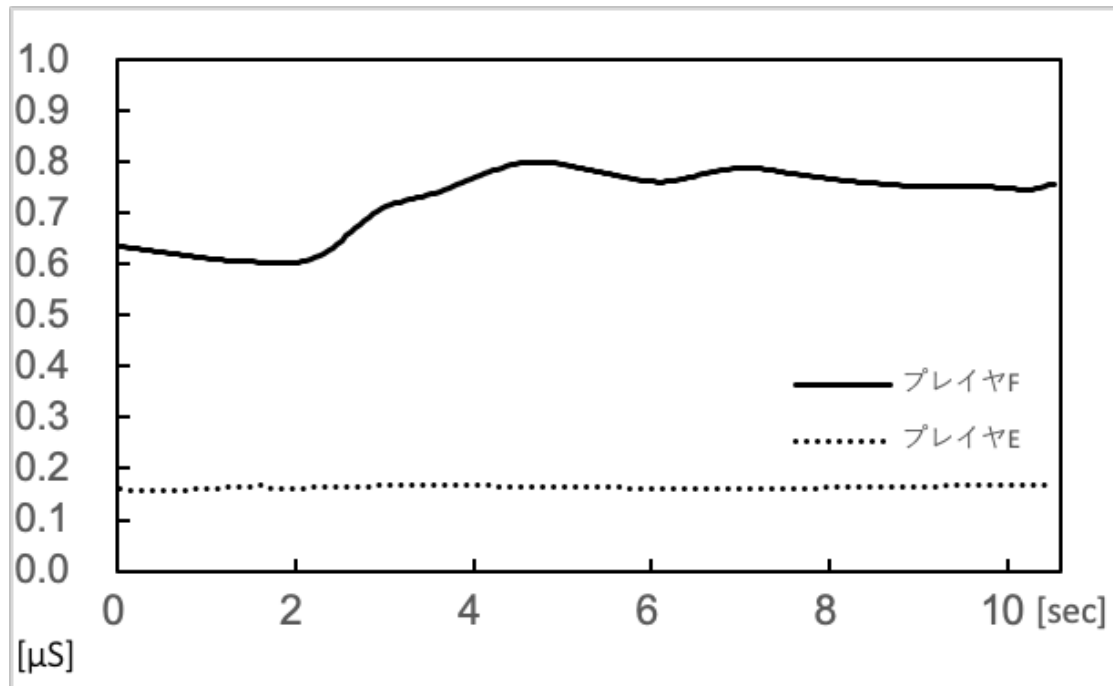


図 3.11: 弱い人狼役職プレイヤーの SCL 値の挙動

が検出された場合、そのプレイヤーにとっての重要な場面であるととらえ、ハイライトとして示した。つまり、本研究では、プレイヤーにとって最も覚醒度の高い情動状態がみられた箇所をハイライトと定義した。これをそれぞれのプレイヤーについて 8 試行分集計することで、各プレイヤーが着目している場面やその傾向について分析する。

3.4.3 結果

本実験では、合計で 8 試行の人狼ゲームを実施した。そのうち、著しいノイズの混入がみられた 1 名の実験参加者の 1 試行分の測定データについては欠損として扱い、分析から除外した。また、本実験においては、全ての試行において初日議論の一番最初に役職カミングアウト (CO) が行われていた。

まず、各試行におけるプレイヤーの SCL 平均値から抽出されたハイライト場面を表 3.3 に示す。この表より、各試行において全プレイヤーが共通して高い SCL 値を示していた場面が投票と CO であることがわかる。CO は、誰が嘘をついているか、などがわかる重要な場面であり、投票は、勝敗が決着する重要な場面であるため、その共通認識があったため、全プレイヤーの SCL 平均値が最大となる場面となったと考える。実験 1 において SCL の平均値で分析した際に、CO や投票の場面が他の場面と比較して高い値となっていることが示されており、これを支持する結果が得られた。

次に、実験で計測した SCL からそれぞれのプレイヤーの各試行において抽出されたハイライトの場面の数を図 3.13 に示す。図は、実験全体でハイライトとして抽出できた場面を集計しており、横軸は役職確認から夜、議論序盤・中盤・終盤・投票・ゲーム終了まで時系列順のイベントを示しており、縦軸は各イベントごとのハイライトの集計値を示して

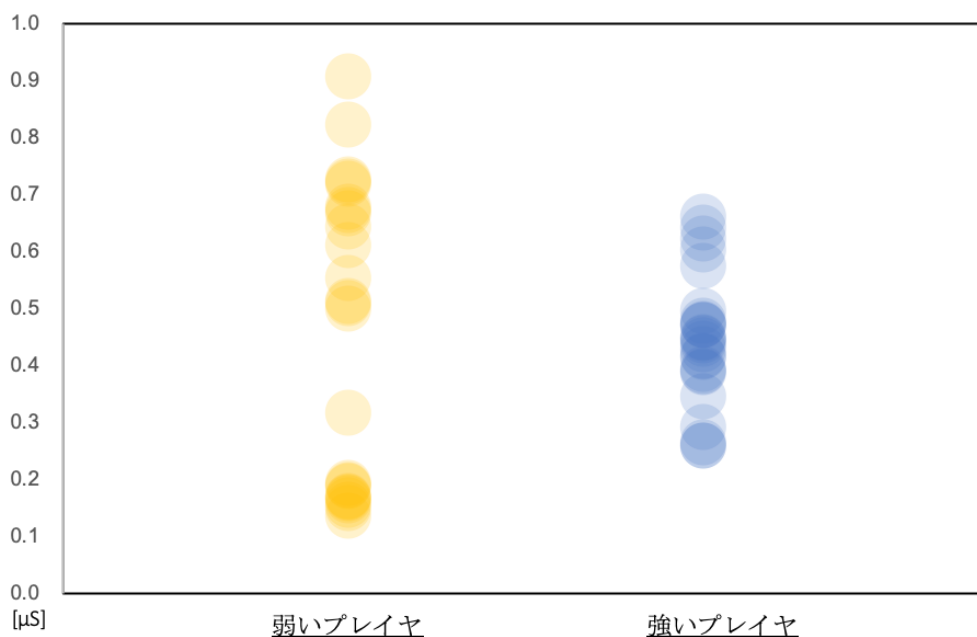


図 3.12: 強いプレイヤーと弱いプレイヤーの覚醒水準の比較

いる。この図より、ハイライトとして抽出された場面は、議論の開始 (CO) から徐々に増加することがわかる。つまり、プレイヤーの議論によって、どのプレイヤーが疑わしいか、など情報が増加するにつれて覚醒度の高い情動状態となった。特に、投票の場面においては、顔を覆って考える動作をする、など投票先に迷うような行動が多くみられており、決断を下す前の悩みや迷い等が SCL 値の上昇に影響を与えたと考えられる。また、他プレイヤーとの会話において、人狼役のプレイヤーが他プレイヤーを疑う発言をする、人狼役のプレイヤーが CO を撤回する、占い師役のプレイヤーの発言を聞いて村人プレイヤーが CO を撤回する、など戦略的な行動をとった場面がハイライトとしていくつか抽出された。以上の結果から、SCL 値を元にハイライト抽出を行うことで、人狼ゲームの進行における重要場面の時系列変化の分布を得ることができた。また、プレイヤー視点における重要場面について、特に戦略的な行動を要する場面が多く抽出される結果となった。

次に、各プレイヤーごとに抽出されたハイライト場面を表 3.4 に示す。この表は、ハイライト場面を行として、実験に参加した 7 人のプレイヤー A~G を列として示したものである。また、実験中にハイライトが抽出されなかった場面は、“-”として表している。また、ハイライト場面については、人狼ゲームの進行に沿って上から順に並べている。この表から、各プレイヤーのハイライト場面を比較すると、それぞれ異なった視点で人狼ゲームを捉えていることがわかる。例えば、プレイヤー C と D を比較すると、プレイヤー C は議論、投票などでの他プレイヤーへの声かけや他プレイヤーからのほたらきかけなどがハイライトとして抽出された。一方、プレイヤー D は役職確認や夜の行動、勝利の瞬間など、自身の内省による情動変化がハイライトとして抽出されていることがわかる。次に、表 3.4 のハイライト場面について詳細を説明する。「役職確認」について、本人の申告ではあるが人狼ゲームに慣れているプレイヤー A, C, G においてはハイライトは確認されなかった。一方で、プレイ

表 3.3: プレイヤの SCL 平均値から抽出したハイライト場面

試行	ハイライト場面
試行 1	投票
試行 2	CO
試行 3	投票
試行 4	投票
試行 5	投票
試行 6	投票
試行 7	CO
試行 8	投票

ヤ B, D, E, F において、「役職確認」にハイライトが確認された。確認されたハイライト 6 件のうち 5 件は役職が与えられた（人狼 2 件，狂人 2 件，占い師 1 件）場面であった。また、「占い結果確認」でみられたハイライト 2 件はどちらも占い結果として人狼を当てた場面であった。また，CO のハイライト 2 件（役職は狂人 1 件と村人 1 件）は，どちらも CO の際にプレイヤー 5 人全員が手を挙げて占い師であると名乗り出ており，その際に笑いが起こっていたことが確認された。議論については議論序盤から終盤まで時間で区切ってハイライトの抽出を行ったところ，結果として序盤が 4 件（人狼 2 件，村人 2 件），中盤が 5 件（人狼 4 件，村人 1 件），終盤が 6 件（狂人 4 件，占い師 2 件）となった。序盤は村人と人狼で，中盤は人狼で，終盤は狂人と占い師で，ハイライトが確認されており，時系列的に変化していることが示された。特に終盤で確認されたハイライト 6 件について，全件でそのプレイヤーが CO を行っていたことが確認されており，それぞれ占い師として投票に向けた主張を行うなど他プレイヤーから注目された場面であった。「投票」では計 10 件のハイライトが確認されており，役職別では人狼 1 件，占い師 3 件，村人 6 件となっており，村人が多い結果となった。村人は誰が占い師で，誰が人狼であるのかの情報を一切もたないため，判断に悩むなど投票への負荷が高まったことが一つの要因として考えられる。「勝利の瞬間」については，プレイヤー B, D, G において計 3 件のハイライトが確認されており，B, G のプレイヤーは喜びを表現するポーズがみられ，D は初勝利であった。勝敗の確定の際の情動表出には理由があると考えられるので，今後より詳細に調査する必要がある。このように抽出されたハイライト場面の傾向については，プレイヤーの人狼経験，熟達度，プレイスタイルなどに応じて変わると考えられる。特に人狼ゲームにおいてプレイヤーがどのような場面を重視するプレイスタイルなのかについては定量データとして測定することが難しいため，本研究におけるハイライト分析が有効である可能性があると考えられる。

3.5 本章のまとめ

本章では，人間の社会的なつながりを活性化させるにあたり，人工物がどのように振る舞うことで，人の行動に変化をもたらすことができるかを理解するために，人同士のコミュニケーションの観察から人の行動を変容しうる要素を調査した。具体的には人同士のコミュニケーションで人の行動を変容させる行動として説得に着目し様々な目的を持った

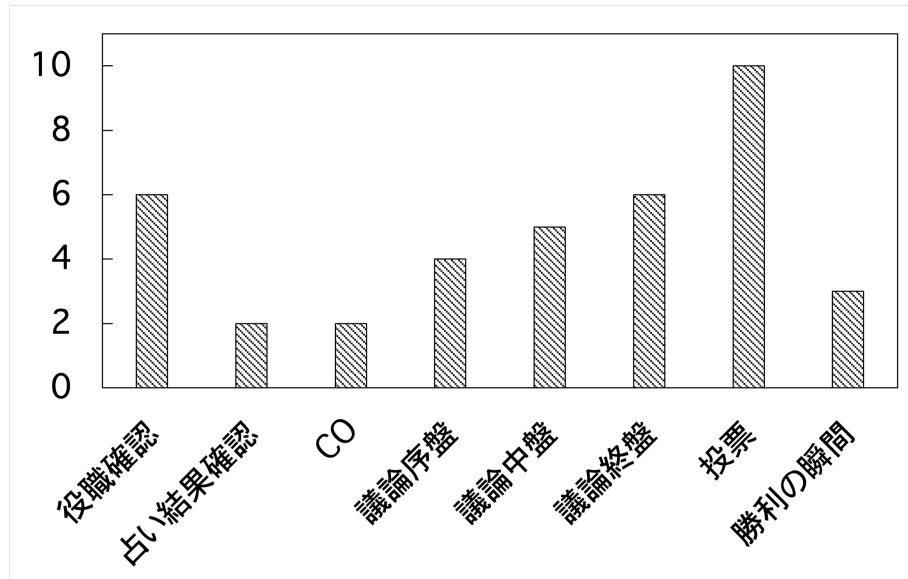


図 3.13: 抽出されたハイライト場面の集計 (縦軸: 全ゲームで全プレイヤーから抽出されたハイライトの合計値, 横軸: ゲーム中の各場面)

立場の異なる複数人の言語的コミュニケーションの課題として人狼ゲームを用い、ゲーム中における他者の考えを変える説得行動に着目して分析した。その結果、人狼ゲームにおける重要な場面やその状況下におけるプレイヤーの振る舞いを理解することができた。人狼ゲームにおいて、他者への説得は重要な要素となっており、重要な場面の多くでその行動がみられた。説得方法にはプレイヤーのゲーム経験や個性によって、様々であるが特に多くのプレイヤーにみられた要素として「適切な状況把握」による「適切なタイミングによる適切な介入」が他者への説得につながるということが理解できた。

表 3.4: 各プレイヤーにおけるハイライト場面の集計表

	A	B	C	D	E	F	G
役職確認	-	1	-	2	1	2	-
占い結果確認	-	-	-	1	-	-	1
CO	-	-	-	-	-	-	2
議論序盤	-	1	1	-	-	2	-
議論中盤	1	-	-	-	2	-	2
議論終盤	-	2	1	-	2	1	-
投票	2	3	4	-	-	1	-
勝利の瞬間	-	1	-	1	-	-	1

第4章 ロボットを介した人-人間インタラクションの調査

本章では人工物の振る舞いによってユーザの社会的つながりの活性化を実現できるかを調査する。環境知能システムが実現された状況を想定して、前章の結果を踏まえて人工物の振る舞いを実現することで、人間の行動の変容を促すことができるかを調査する。具体的には、複数人コミュニケーションの課題としてレクリエーション活動を用い、レクリエーション活動に対するモチベーションの向上を目標として設定して評価する。レクリエーション援助を行う人工物としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者のモチベーションを向上できるかを調査する。提案手法として、適切なタイミングと適切な介入方法として、それぞれ片方がゲーム中に不利な状況下と不利なプレイヤーが有利になる情報の提供を設定し、ロボットに対して振る舞わせたときの効果を評価する。

4.1 研究手法

4.1.1 概要

図4.1にロボットがレクリエーション援助者として介入を行うシステムの概念図を示す。本研究では、レクリエーションとして「だるまさんがころんだ」をベースとしたVRゲームを開発した。「だるまさんがころんだ」を選択した理由は、伝承遊びとして広く認知されており、実験参加者も理解しやすい点が挙げられる。また、VRゲームとした理由として、狭い室内でも体を動かして遊べる仕掛けを含めることが容易である点、外での遊びに近い体験が可能である点が挙げられる。本ゲームはVR環境内で実行されるため、クリアタイムやプレイヤーの動作などのデータの取得が容易であり、これらの情報をネットワークを介してレクリエーション援助ロボットへ伝達することで、ロボットはゲームの状況に応じてリアルタイムに判断・介入することが可能となる。

4.1.2 実装環境

本研究で作成したゲームは、Unity 2021.1.25で開発された。VRゲームの開発を行うにあたり、XR Interaction Toolkitを用いた。実行環境は、OS: Windows, CPU: Intel(R) Core(TM) i7-10700, 実装メモリ: 64GB, GPU: RTX2070 SUPERを搭載したPCで行った。ヘッドマウントディスプレイ(HMD)はVIVE Pro2 (HTC社, 片眼2448x2448px, リフレッシュレート90/120Hz, 視野角120度)を使用してプレイヤーに視覚情報の提示した。また、ゲーム中のプレイヤーの行動を認識するために、HMDから頭部座標、両手に把持し



図 4.1: Robot mediation in recreation

たコントローラから両手の座標を取得した。

4.1.3 ゲームのルール

次にゲームのルールについて説明する。ゲームの外観を図 4.2 に示す。図右側の白い球体がプレイヤーであり、左側が鬼役のキャラクターである。プレイヤーは、鬼役の方を向き進行する。図 4.2 は俯瞰図であるが、実際のプレイヤーの主観視点であり、図 4.3 のようにみえる。また、プレイヤー視点では制限時間や失敗回数が確認できる。鬼役のキャラクターはプレイヤーとは反対方向を向いているが、「だるまさんがころんだ」の掛け声の後にプレイヤーの方向へ振り返る。鬼役が振り返っている際にプレイヤーが動くと、ゲーム失敗となる。また、プレイヤー以外の操作によってプレイヤーの進行方向に障害物を設置する妨害行動が存在し、障害物へ触れた場合においてもゲーム失敗となる。ゲームに失敗した場合、再度、初期位置（図 4.2 右側）からのやり直しとなるが、ゲームクリア時にペナルティが課される。1 回のゲームにおけるスコアは、「ゲームクリア時間 [秒] + 5[秒]*ゲーム失敗回数」で計算される。

4.1.4 ゲームの操作方法

プレイヤーのゲームの操作方法は 2 種類ある。それぞれ、歩行操作とジャンプ操作である。歩行操作はコントローラを持ちながら走るように腕を前後に振ることで実行される。歩行操作では、プレイヤーは鬼役のキャラクターに近づくことができる。歩行操作の条件としては、図 4.4 において、右手もしくは左手座標が、頭部座標より X 軸方向で前に移動すると、歩行操作が実行される。次に、ジャンプ操作は、障害物が出現しているときのみ可能である。障害物が出現すると、プレイヤーの前にジャンプゲージが出現する。ゲージが満タンの状態で、両手のコントローラを後ろから前に振るとジャンプ操作が実行される。ジャンプ操作では、ゲーム中に出現した障害物を回避することができる。ジャンプ操作の条件は、図 4.4 において、頭部座標が前方を向き、かつ 0.1 秒以内に両手の座標が X 軸方向で前に移動すると操作が実行される。



図 4.2: Appearance of the game "Daruma-san ga Koronda"

4.1.5 ヒューマノイドロボット

レクリエーション援助役のロボットとしてソフトバンクロボティクス社が製造しているヒューマノイドロボットである Nao[72] を用いることとした。Nao のジェスチャや発話は Choregraphe (Nao を制御するための専用開発環境) を用いて実装した。ロボットの介入として、適切なタイミングと適切な介入方法として、それぞれ片方がゲーム中に不利な状況下と不利なプレイヤーが有利になる情報の提供を設定し、ロボットに対して振る舞わせた。

4.2 評価実験

ロボットの介入がプレイヤー間のインタラクションに与える影響を調査するために実験を行った。著者が開発した「だるまさんがころんだ」をベースとした VR ゲームを 2 人組に体験させた。被験者はゲームのプレイ方法を教示されたあと、それぞれ 3 回ずつゲームを体験した。実験者は、被験者に対してより良いゲームスコア (ゲームクリアまでの時間) を目指すように指示を与え、被験者同士で競い合わせた。実験終了後に実験時の体験に関するアンケート調査を被験者に回答させた。

レクリエーション援助者として、ロボットは被験者の心身の状態を把握し、適切に介入することで、レクリエーション参加者の目標達成を支援することが求められる。そこで、レクリエーション支援に対するロボットの介入効果について以下の仮説を立て、評価を行う。

H1: ロボットの介入によりレクリエーション参加者はより積極的にゲームに参加するようになる

H2: ロボットの介入により参加者間のコミュニケーションが活性化される



図 4.3: Player's view while playing game

4.2.1 実験環境

実験の概要を図 4.5 に示す。図左はゲームをプレイする実験参加者（以降、ゲームプレイ役）である。ゲームプレイ役はヘッドマウントディスプレイ（以降、HMD）を被り、表示されているゲームをプレイする。図右はゲームプレイ役を妨害する実験参加者（以降、妨害役）である。本実験で用意したゲームでは、通常の「だるまさんがころんだ」のルールに加えて、妨害要素として「他者による障害物の設置」を設定した。妨害役は、ゲームプレイ役と別室におり、ゲーム画面をモニタ上で確認しながら、任意のタイミングで妨害することができる。妨害行動は 2 種類あり、一つが「障害物の設置」、もう一つが「フェイント」である。「障害物の設置」ではバナナ型の 3D オブジェクトを青色の煙と共にプレイヤーの前に出現させる。「フェイント」では、青色の煙のみをプレイヤーの前に出現させる。「障害物の設置」の際、バナナ型の 3D オブジェクトは煙が出現している間は見えなため、煙が消えるまでゲームプレイ役は妨害行動が「障害物の設置」か「フェイント」かがわからない。これによって「フェイント」の妨害行動が成立する。妨害行動には回数制限があり「障害物の設置」はゲームの各試行において 1 回まで、「フェイント」は 3 回までである。

4.2.2 実験条件

実験条件としてロボットによる介入なしと介入あり条件を設定した。実験条件は各組ごとに異なり、介入なし条件は 6 組、介入あり条件は 6 組であった。各実験条件の違いは、妨害行動の教示方法にある。介入なし条件では、実験者が実験を始める際のゲームの操作方法とともに妨害の存在と妨害のための操作方法を教示する。一方で、介入あり条件では、

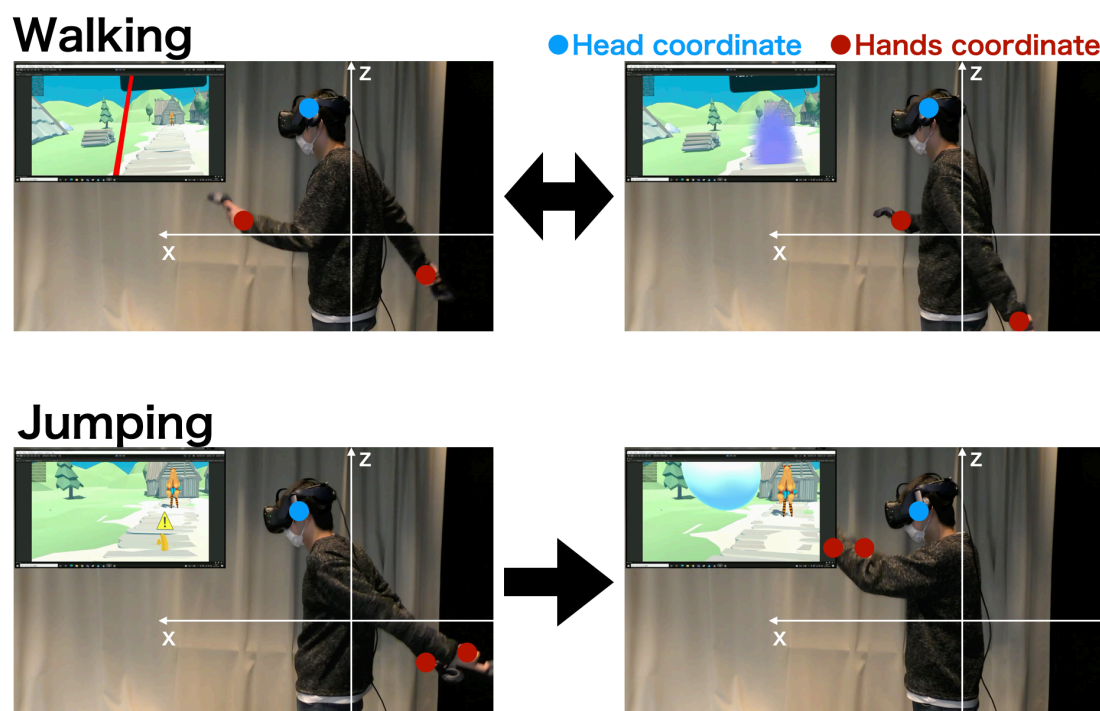


図 4.4: Operations of the game

実験を始める際には妨害についてのみ、つまり「ゲームプレイ中に障害物が出現する可能性がある」旨を教示し、誰がその操作を行うかについては明示しない。実験参加者両名が1回目のゲームを終えて勝敗を決めた際に、敗者にのみ、妨害行動の操作方法をロボットが教示する。2回目のゲームを終えた際に、1回目のゲームの勝者が敗北した場合には、その実験参加者にも妨害行動の操作方法をロボットが教示する。つまり、介入あり条件では、妨害によって敗者が勝者に対して有利な条件でゲームを行うことができる。

4.2.3 アンケート

ゲーム体験終了後に被験者にアンケートに回答させた。アンケートの目的は、レクリエーションへの積極性や相手とのコミュニケーション意欲における条件間（介入なし条件/介入あり条件）の差異を調査することである。アンケート項目は表 4.1 に示すとおりである。アンケートの回答様式は6段階（0：全くそう思わないー5：とてもそう思う）のリッカート尺度を用いた。ただし、問 9, 10 は選択式であり、問 11 は記述式であった。

4.2.4 被験者

実験参加者は 24 人 (12 組) であった (男性 17 人, 女性 7 人, 平均 21.1 歳)。各組はお互いに親しく、共にゲームを楽しめる関係性であることを条件として募集した。また、実験参加者は「だるまさんがころんだ」のルールについてあらかじめ理解しており、実験中

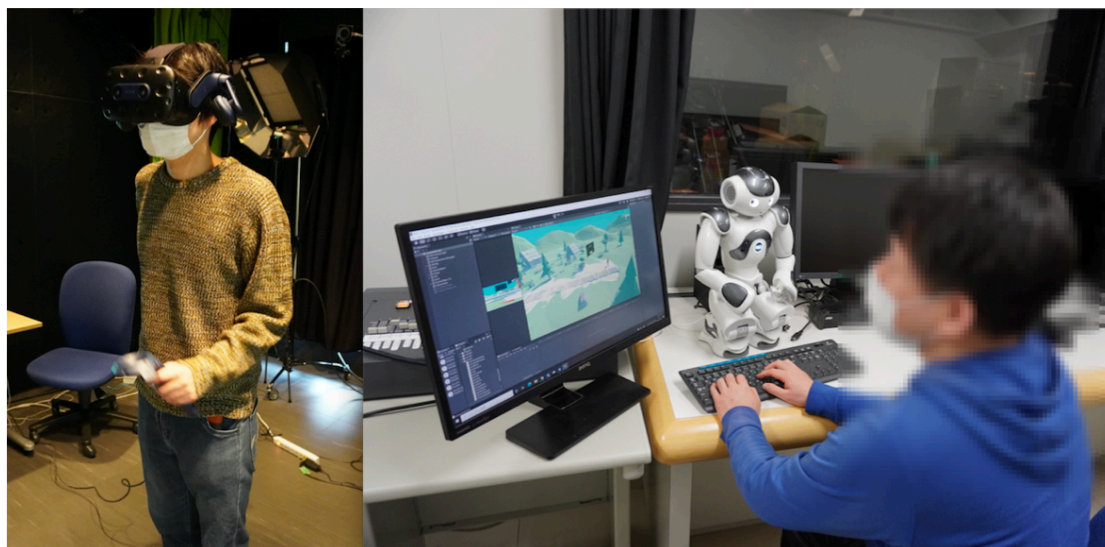


図 4.5: 実験概要図 (図左：ゲームプレイ役の実験参加者, 図右：妨害役の実験参加者)

にルールの不理解による不備は生じなかった。実験参加者の男女の組み合わせは、男性 2 人のペアが 7 組、女性 2 人のペアが 2 組、男女 1 人ずつのペアが 3 組であった。

4.2.5 実験手順

まず、実験参加者を実験室に入室させた後、実験者が実験説明を行った。実験の目的、内容、取得データなどの説明を受け、実験参加者 2 人とも実験参加に同意可能な場合にのみ実験を行った。次に、WOZ 法を用いてロボットを操作し、ロボットが本実験の司会を務めること、最初に操作説明を行うことなどを教示した。操作説明では、ゲームのルール説明や操作方法について説明を行った。この時、介入なし条件では妨害行動の存在と操作方法を両実験参加者に対して説明した。次に、実験参加者に HMD を交代で装着させ、ゲーム操作の練習を実験参加者が操作に十分に慣れたと感じるまで行った。その後、どちらが先にゲームプレイ役となるかを決め、交互のゲームを体験させた。両実験参加者が 1 回ずつゲームを体験させた時点で、WOZ 法にてロボットを操作して、結果発表を行った。2 回目のゲームは勝者から先に体験させた。この時、介入あり条件では敗者に対して妨害行動の操作方法を WOZ 法にてロボットを操作して教示した。これを両実験参加者が 3 回ずつゲームを体験するまで繰り返し、最後にアンケートに回答させた。

4.2.6 結果

アンケート調査の結果を図 4.6 に示す。横軸はアンケートの各項目、縦軸はリッカート尺度の点数である。エラーバーは標準偏差を示す。全体的な結果として、条件に関わらず、各項目の点数が高く参加者にとって、エンタテインメント性の高い体験であることがわかった。アンケートの各項目について、条件間の関係を明らかにするために一要因分散分

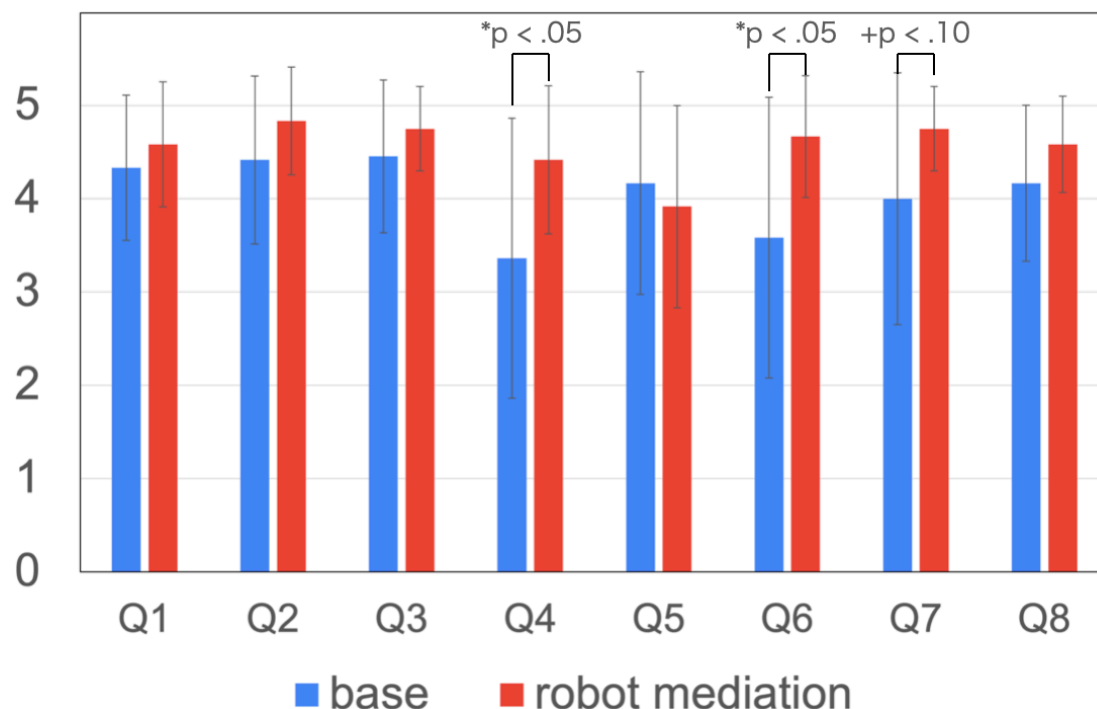


図 4.6: The results of questionnaire (Q1 - Q8)

析を行った。結果、Q4 と Q6 において介入あり条件のほうが有意に高い点数であることが示された ($*p < .05$)。また、Q7 において有意傾向が示された ($*p < .10$)。

図 4.7 に、各条件におけるゲーム間のプレイ時間の差 (秒) を示す (エラーバーは標準偏差を示す)。プレイ時間について、条件間の関係を一要因分散分析で分析すると、有意差が示された [介入なし条件: $F(2, 22) = 4.06$, $*p < .05$, 介入あり条件: $F(2, 22) = 2.84$, $+p < .10$]。事後分析として LSD 法による多重比較を行ったところ、介入なし条件では第 3 ゲームのプレイ時間が第 1 ゲームより長く ($*p < .05$)、介入あり条件では第 1 ゲームのプレイ時間が第 3 ゲームより長く ($*p < .05$) なることが示された。

4.3 考察

4.3.1 仮説の検証

まず、実験から得られたアンケート結果について述べる。アンケート調査において 2 条件間の比較からいくつかの項目で有意差・有意傾向がみられた。具体的には「相手に負けたときくやしかった (Q4)」と「次は勝とう」という気持ちになった (Q6/Q7)」という項目であった。これらの項目は感情の動きや競争の促しの要素を表しており、ロボットの介入あり条件においては、被験者の感情の動きや勝ち負けへの執着が高い状態であったことがわかった。つまり、ロボットの介入が被験者のレクリエーションへの参加意欲を高めたと考えられ、仮説 H1 は支持された。結果から、ロボットの介入によって、被験者はよ

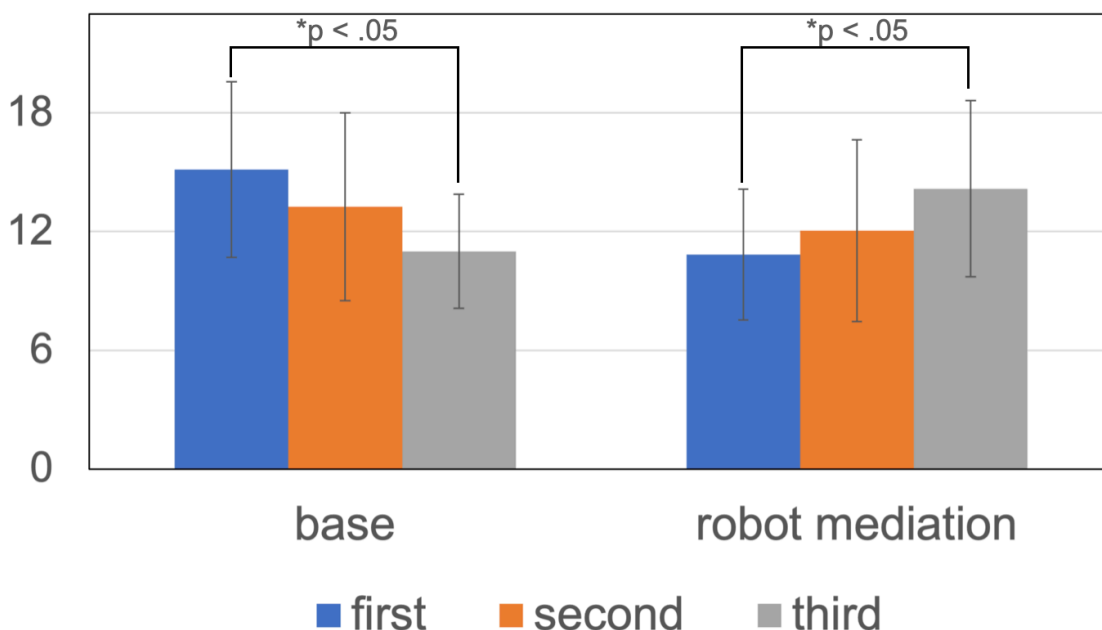


図 4.7: Gameplay time for each play

り積極的にレクリエーションへ参加するようになることがわかった。実験では、介入なし条件と介入あり条件の違いは妨害行動を教示するタイミングのであった。介入なし条件では、実験開始時に両被験者に対して妨害行動の操作方法を教示する。一方、介入あり条件では、ゲームに負けた被験者に妨害行動の操作方法を教示する。そのため、2つの条件におけるロボットの行動の差はわずかであった。しかし、2条件間におけるロボットの行動の差は、わずかな差でもレクリエーションの参加者に影響を与えることを示している。一方で、この実験では、ロボットはレクリエーションの参加者に対して良い影響（モチベーションを高める）を与えたが、介入の方法を間違えると悪い影響を与える可能性がある。実験では妨害行動はゲームの敗者のみに教示され、勝者は敗者が妨害行動を行っていることを知らなかった（口頭のインタビューでは、勝者はコンピュータが妨害行動をランダムに生成していると考えていた）。もし、実験中に敗者だけが妨害行動を行えることを知っていたら、勝者はレクリエーションや援助者であるロボットに対して不公平さを感じる可能性があり、レクリエーションに対するモチベーションは向上しなかったと推測される。

次に、コミュニケーションの活性効果に関連するQ8は、本実験で有意な差がみられなかった。したがって、本実験の結果はH2を支持しない。Q8の結果を考察すると、両条件で高い点数であった。各条件の中央値は4（そう思う、介入なし条件）と5（とてもそう思う、介入あり条件）であり、実験の条件に関わらず、被験者は相手とコミュニケーションを取りたいと考えていたことが理解できる。被験者の選定条件としてゲームで競い合うことができる程度に親しい2人組としたため、もともとコミュニケーションを取る意欲の高い組み合わせであったことは差がみられなかった一因であると考えられる。

全体を通して、実験におけるレクリエーションの体型は、被験者にとってエンタテインメント性の高い体験であった一方で、各被験者にとって初めての体験による認知バイアス

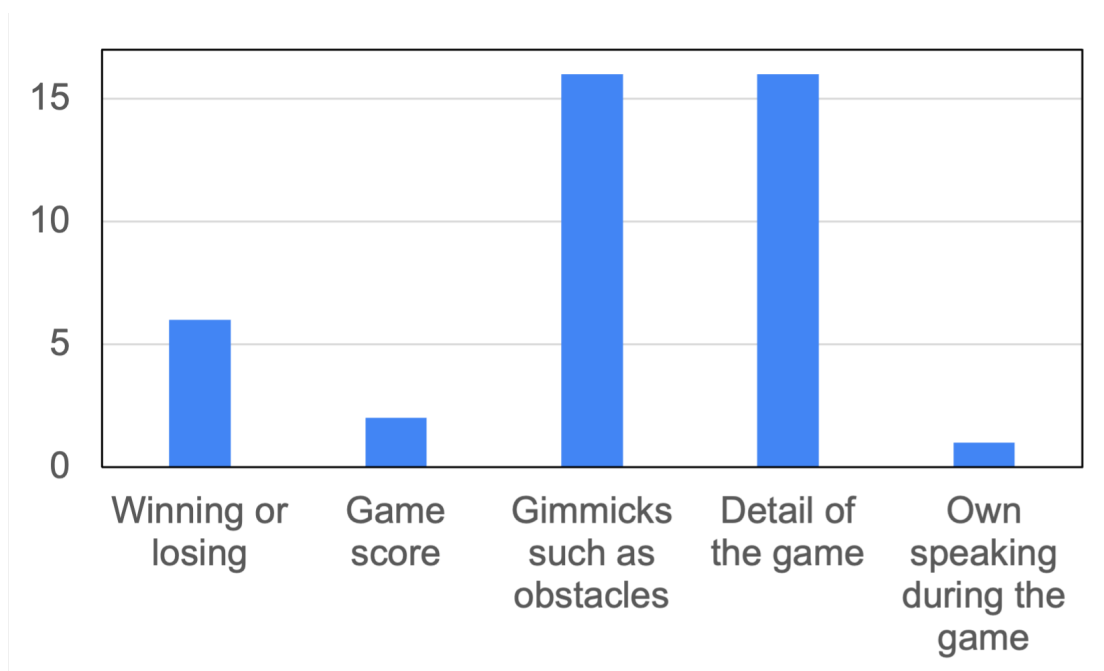


図 4.8: The result of questionnaire (Q9)

が存在し、体験のエンタテインメント性を正しく評価することが難しかった可能性がある。被験者全員にとって、実験のレクリエーションは初めての体験であり、かつほとんどの被験者にとって初めてのVR体験であった。そのため被験者はレクリエーションから得られるエンタテインメント性を過大に評価した可能性がある。今後、長期的な実験で被験者にレクリエーションを体験してもらうことで、このようなバイアスは排除する必要がある。

4.3.2 他のアンケート項目について

本節では、残りのアンケート項目 (Q9~Q11) について考察する。まず、Q9の結果を図4.8に示す。Q9はゲーム終了後に被験者が相手と話したい内容について質問した項目である。この結果から、被験者は障害物などのゲーム内容について話したいと考えていることがわかった。コミュニケーション活性化の促進という観点からみると、ゲームの勝ち負けなどの結果より、内容や仕掛けについて相手と話したくなるようなレクリエーション設計が重要であることがわかった。

Q10とQ11は実験中に被験者の心身の健康に悪影響がなかったかどうかを確認するための項目である。Q10では24人中4人がゲーム中に「不快感を感じた」と回答した。Q11はその不快感の内容に関する項目であり、ある参加者は「止まっているけど進んでいることが違和感だったのと、バランス感覚が狂うこと。」や「ジャンプする動作で少しコードに腕が引っかかりかけた」と回答している。これより、実験環境としてVR酔い対策や配線の取り回しなどに工夫が必要であることがわかった。また、被験者1人から「負けた時、ギミックを配置したであろう相手について苛立ちなどの不快さを感じた」と回答を得てお

り、相手の妨害行動に不快感を覚えた被験者もいることがわかった。

4.4 本章のまとめ

本章では人工物の振る舞いによってユーザの社会的つながりの活性化を実現できるかを調査するために、システムによるユーザ支援がユーザの社会的つながりの活性化に貢献し得るのかを調査した。複数人コミュニケーションの課題としてレクリエーション活動を用い、レクリエーション活動に対するモチベーションの向上を目標として設定して評価した。レクリエーション援助を行う人工物としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者のモチベーションを向上できるかを調査した。提案手法として、適切なタイミングと適切な介入方法として、それぞれ片方がゲーム中に不利な状況下と不利なプレイヤーが有利になる情報の提供を設定し、ロボットに対して振る舞わせたときの効果を評価した。評価実験の結果、ロボットが行った介入によって、レクリエーションの参加や勝利へのモチベーションが有意に高まったことが示された。これらの研究から、人間の身体の内的変化を活用してユーザの社会的なつながりを活性化させる環境知能システムの効果を示すことができた。

表 4.1: アンケートの設問内容

設問内容	
問 1	ゲームに集中できた
問 2	ゲームにのめり込む感覚があった
問 3	相手に勝った時うれしかった (勝ちなしの場合は飛ばしてください)
問 4	相手に負けたときくやしかった (負けなしの場合は飛ばしてください)
問 5	ゲームプレイ中, 緊張感やスリルを感じた
問 6	1回目のゲーム後, 「次は勝とう」という気持ちになった
問 7	2回目のゲーム後, 「次は勝とう」という気持ちになった
問 8	実験後にペアの実験協力者と話したいと思った
問 9	(問 8 でとてもそう思う, そう思う, ややそう思うと回答した方へ) どのような場面・要素について話したいですか? (5つの選択肢から選択 (複数回答可): ゲームの勝敗, 障害物 (バナナ) などのゲームギミック, プレイ内容 (うまくプレイできたシーンなど), ゲームのスコア, その他)
問 10	実験を通して, 不快感や嫌な気持ちを感じることがありましたか? (2つの選択肢から選択: 不快感を感じるがあった, 不快感はなかった)
問 11	(不快感を感じるがあったと回答した方へ) 差し支えなければ不快に感じた内容を記入してください

第5章 人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討

本章では、システムの実現可能性を示すために人の社会的活動を阻害せずシームレスに入力できるインタフェースを検討する。これまで提案されてきた人間の情報収集手法として、主に人間のコミュニケーションメディアを利用した手法が用いられてきた。しかし従来の手法では、実世界の人間の活動を阻害するといった課題がある（他者との会話中に音声入力インタフェースを使用できない、両手が塞がった状態では手指を使った操作ができない等）。そこで本研究では様々な環境下で入力可能なインタフェースを目指し、関節動作を伴わない筋肉の等尺性収縮に着目しインタフェースの構築を行う。具体的には、まずインタラクションがなぜ阻害されるのか、その阻害の原因は何であるかについて調査を行う。その後、阻害要因について考察を行い、解消するための提案手法を検討する。提案手法を用いたハンドジェスチャ入力の識別について、シミュレーション評価とユーザ評価を行い、インタラクションの阻害を解消できたかを評価する。

5.1 インタラクション阻害要因の調査実験

本実験では、ユーザの手が荷物の把持により占有されている状況で、従来のジェスチャ認識手法がどのように機能するかを確認する。Saponas ら [77] はユーザの手が占有されている状況においてハンドジェスチャ認識技術の評価を行ったが、その状況のジェスチャデータは従来手法により学習されていた。そこで、その状況におけるジェスチャデータの有無による識別精度に着目して従来手法を評価した。

5.1.1 データ収集実験

前述の分析を行うためにはハンドジェスチャのデータを収集する必要がある。Saponas らの手法は我々の手法とは異なるが、彼らの手法は特定の手のジェスチャを認識し、我々の手法は特定のジェスチャ列を認識する、という点で同じである。我々はこれらの実験用にジェスチャデータを集めるためのデータ収集実験を実施した。その実験風景を図 5.1 に示す。

設備

実験の主な設備として、Bluetooth が利用可能なコンピュータ（Apple Macbook Pro, Retina, 15-inch, Mid 2015）と商用で入手可能な筋電位測定センサである Myo を使用し



図 5.1: データ収集実験の概要

た。また、ジェスチャデータを計測するためのシステムを Processing 3 を用いて実装した。この実装の中で、”Myo for Processing” ライブラリの ”myoOnEmgData” メソッドから得られた 8 次元の離散値を EMG データとした。EMG データはそれぞれ 200Hz で取得され、11 bit で -1024 ~ 1023 の値で量子化される。

実験は明るい室内で実施された。実験参加者（以後、参加者とする）はディスプレイと対面した状態で立っており、ディスプレイから情報を得ながら実験を進めた。

ハンドジェスチャ

実験で行うハンドジェスチャとして、複数の先行研究で共通して使われているジェスチャを用いた [3, 54]。図 5.2 に実験で用いたハンドジェスチャの様式を示し、以下にそれぞれのジェスチャについて簡単に説明する。

1. Dorsal(DR)：手の甲を自身の身体へと仰け反らせるような動作
2. Palmar(PL)：手の平を自身の身体へと向けるような動作
3. Open Hand(OH)：手の平を思い切り開くような動作（物を把持している場合は親指を立てる動作で代替した）
4. Fist(FT)：握り拳を作るような動作

これらのジェスチャは従来手法において高精度に認識可能であることが確認されている。また、これらのジェスチャは見た目の違いがわかりやすく、参加者が容易に覚えることができた。



図 5.2: 実験で用いるハンドジェスチャ

実験設計と手順

実験の条件として、参加者の手の状態が異なる4つの条件を用意した。条件1では、フリーハンドの状況のジェスチャを記録した。他の3つの条件では、ユーザが物を把持している状況で行うジェスチャを記録した。それぞれの条件は、傘（条件2）、バッグ（条件3）、箱（条件4）を把持する条件であった。それぞれの条件で参加者は80回のジェスチャを4セットに分割（各セット20回ずつ）して入力した。1セットでは、4種のハンドジェスチャを5回ずつとし、合計で20回とした。各セット間では、参加者の希望があれば休憩を取った。我々はそれぞれの条件で合計80回のジェスチャを記録し、実験を通して320回のジェスチャを記録した。

実験のはじめに参加者は実験室に入った後、上記で説明した実験手順の説明を受けた。それには、ジェスチャを行う回数や各条件をどの順番で実施するかが含まれていた。次に、実験で参加者の肌に触れる物質のアレルギーの有無を確認した。その後、参加者の右前腕を酒精綿で軽く吹きとり、Myo デバイスをおおよそ右前腕上部に装着した。それから、我々は参加者にディスプレイに映し出したビデオクリップを見せながら、ジェスチャの行う方法について指導した。このビデオクリップでは、それぞれの条件におけるそれぞれのジェスチャが再生され、参加者にはそれを見てもらいながら真似させた。最後に、我々は参加者にジェスチャ収集方法について説明した。

ジェスチャの記録は、図5.3のようなジェスチャ記録システムを製作し、これを使って記録した。また、このシステムを使って参加者にジェスチャを行うタイミングや行うジェスチャの種類を指示した。

1. ディスプレイに行うべきジェスチャが表示される
2. ジェスチャの開始のタイミングを知らせる音が鳴る

3. ディスプレイに表示されているジェスチャを行う
4. 図 5.3 の緑色のバーが右端に到達するまで (2.5 秒間) ジェスチャを維持する
5. 4.5 秒のインターバル後, 1. に戻る

このような様式で参加者はジェスチャを 20 回繰り返し入力し, 我々はその時の EMG を記録した. タスクのジェスチャは各 5 回ずつランダムな順番で表示された. 疲労によるジェスチャの認識精度への影響を排除するため, 参加者は必要であれば 5~10 分程度休憩を取った後に, 次のセットを行った. 全ての参加者は実験が終わるまで Myo デバイスを取り外さなかった.

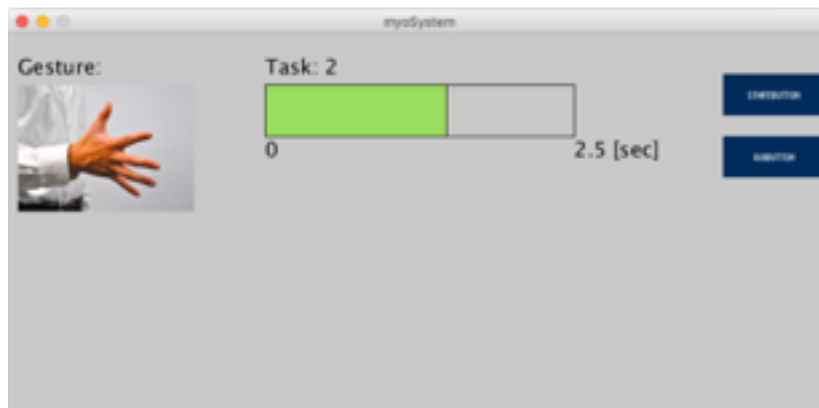


図 5.3: ジェスチャ記録システム

実験参加者

12 名の参加者 (男性 7 名, 女性 5 名, 平均年齢: 23.5 歳) にデータ収集実験に協力してもらった. 全ての参加者は日常的にコンピュータを使用しており, 実験で参加者の肌に触れる物質についての皮膚アレルギーを持っておらず, 利き腕は右腕であった. 実験時間はおおよそ 1 時間 30 分程度を要した (休憩時間を含む). データ収集実験では, 参加者にはフリーハンドの状況と物を把持した状況すべてでジェスチャを行ってもらった.

結果

我々は, 参加者一人あたり 320 個のジェスチャをした時の EMG を収集した (合計で 3,840 個の EMG). 320 個の内訳は, 20 回 \times 4 条件 \times 4 種類のジェスチャとなっている. 1 個のデータには 1 回のジェスチャをした時の 8ch の EMG が記録されている. データの詳細を述べると, 8ch \times 500 サンプルの二次元配列となっている. また, データの値は -3.3V ~ 3.3V の値を -1024 ~ 1024 の範囲で離散化されている. 収集したデータの例を表 5.1 に載せる.

まとめると, 4 種類のハンドジェスチャの EMG を 4 つの条件で収集した. 条件はフリーハンドの状況と 3 つの物を把持した状況である. 収集したデータに欠損値はなかった.

表 5.1: 収集した EMG の例

	1ch	2ch	3ch	4ch	5ch	6ch	7ch	8ch
1	4	-28	-20	-12	-12	-20	-60	-28
2	-12	-12	-4	4	-4	12	-68	4
.
.
.
499	365	-445	-60	-12	381	-44	-421	-493
500	-614	-269	-52	-317	-261	-236	-542	-261

5.1.2 分析

従来手法において、物を把持した状況におけるジェスチャ認識の実現可能性を検討するために実験を行った。我々は、SVM を用いて従来のジェスチャ認識システムを実装し、データ収集実験で集めた EMG を使って、その性能を検証した。SVM はジェスチャ認識における手法として良く用いられており、従来のアプローチであれば多くの先行研究で高精度の認識精度を示しているため、本実験で採用した。

本実験では 2 つの条件 A, B で従来手法の識別器の性能を評価する。条件 A ではデータ収集実験で集めた 4 つの状況ごとに独立して学習とテストを行う。この条件は先行研究 [77] で行われた識別手法と同様であり、高い認識精度でジェスチャを識別できると予想される。したがって、この条件の結果から我々の実装した従来手法が正しく動作しているか、またデータ収集実験におけるデータの正当性を確認する。条件 B ではフリーハンドの状況で識別器を作成し、物を把持した 3 つの状況でテストを行う。この条件は、物を把持した状況が未学習である識別器を用いたジェスチャの識別であり、先行研究ではどのような結果となるか確認されていない。我々は、2 つ目の条件の結果から従来手法における未学習の状況に対する適用可能性を検討する。

特微量抽出と学習

この節では、学習とテストのための特微量について説明する。我々は収集したデータから、それぞれの ch の二乗平均平方根（以後、RMS）値を計算した。RMS 値は以下の式にて、計算された。

$$RMS_i = \sqrt{\frac{1}{N} \sum_{j=1}^N EMG_{ij}^2} \quad (5.1)$$

$$RMS = [RMS_1, RMS_2, \dots, RMS_8] \quad (5.2)$$

EMG はジェスチャデータ全体、 i は ch 番号、 N はサンプルサイズを表す。1 つのデータのサンプルサイズは 500 なので、 $N=500$ である。したがって、(式 4.1) より RMS 値を各 ch ごとに 500 サンプルを 1 つのデータに圧縮し、(式 4.2) のようにそれらの特微量として 1 つの EMG データを 8 次元のベクトルであらわした。

ジェスチャ認識器は、プログラミング言語 Python を用いて実装した。その際、機械学習ライブラリである”scikit learn”¹を用いて認識器の学習とテストを実装した。条件 A では、各状況（1種：フリーハンド、3種：物を把持）ごとに独立して、学習とテストを行った。各状況において、80 個のデータ（特徴量）のうち、60 個（15 個 × 4 種のハンドジェスチャ）を訓練データとし、20 個（5 個 × 4 種のハンドジェスチャ）をテストデータとした。条件 B では、フリーハンドの状況の 60 個（15 個 × 4 種のハンドジェスチャ）のデータを訓練データとして認識器を作成し、3 つ物を把持した状況のデータから各 20 個ずつ（3 種の状況 × 5 個 × 4 種のハンドジェスチャ）をテストデータとした。条件による学習データとテストデータを図 5.4 に示す。また、それぞれの条件や状況における識別器は 10-分割交差検証を用いて作成した。具体的には、”scikit learn”の sklearn.grid_search.GridSearchCV 内の cv パラメータを 10 と設定することで実装した。また、識別器の作成や精度の評価は各参加者ごとに独立して行われ、訓練データとテストデータは参加者特有のデータが使用された。

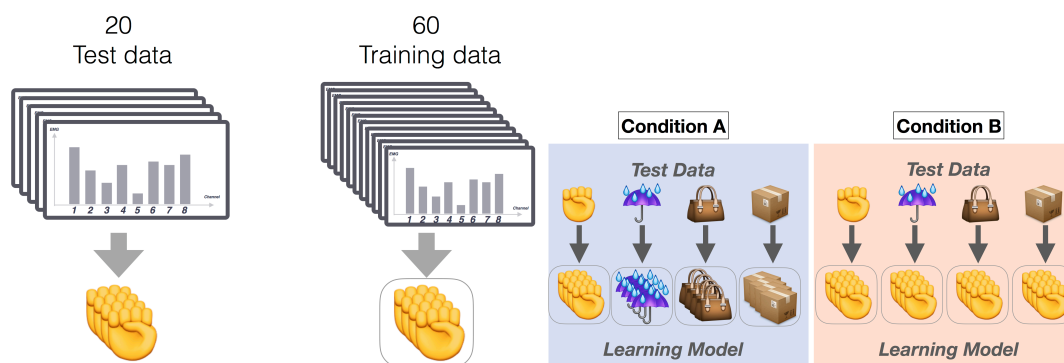


図 5.4: 条件 A, B における学習・テストデータ

結果

条件 A, B における、認識精度の結果を図 5.5 に示す。エラーバーは標準偏差を表す。まず、条件 A における認識精度を確認すると各状況において 90%以上の高い認識精度を示しており、このことから従来手法のジェスチャ認識器は正しく動作していることが確認できた。ここで、両条件を比較すると物を把持した状況において認識精度の急激な変化が起こっていることがわかった。具体的には、条件 A から B において傘は 100%から 55%、バッグは 99%から 77%、箱は 92%から 45%へと認識精度が減少した。条件 A では、物を把持した状況においてジェスチャを認識精度は高かったことから、条件 B では学習していたフリーハンドジェスチャの EMG とテストに用いた物を把持した状況のジェスチャの EMG が異なっていたことがわかる。これにより、物を把持することでジェスチャを行った際の EMG が変化し、その結果、条件 B では未学習のジェスチャを正しく認識することができなかった。

識別精度について 2 要因参加者内（条件：2 種 × 物を把持した状況：3 種）における

¹<http://scikit-learn.org/stable/>

表 5.2: 条件 A, B における各状況の認識精度と状況における平均認識精度

	フリーハンド	傘	バッグ	箱	平均 (フリーハンドを除く)
条件 A	99.1%	100.0%	99.1%	92.5%	97.2%
条件 B		55.0%	77.5%	45.0%	59.2%

分散分析の結果, 有意な交互作用 [$F(2, 22) = 7.08, p < .01$], 条件の主効果 [$F(1, 11) = 123.08, p < .01$], 状況の主効果 [$F(2, 22) = 14.95, p < .01$] が認められた. これより, 物を把持している状況について条件 A が条件 B より有意に精度が高いことが示された. また, Bonferroni 法により p 値調整を行った t 検定における多重比較の結果, バッグを把持している状況では他の 2 状況より有意に精度が高いことが示された.

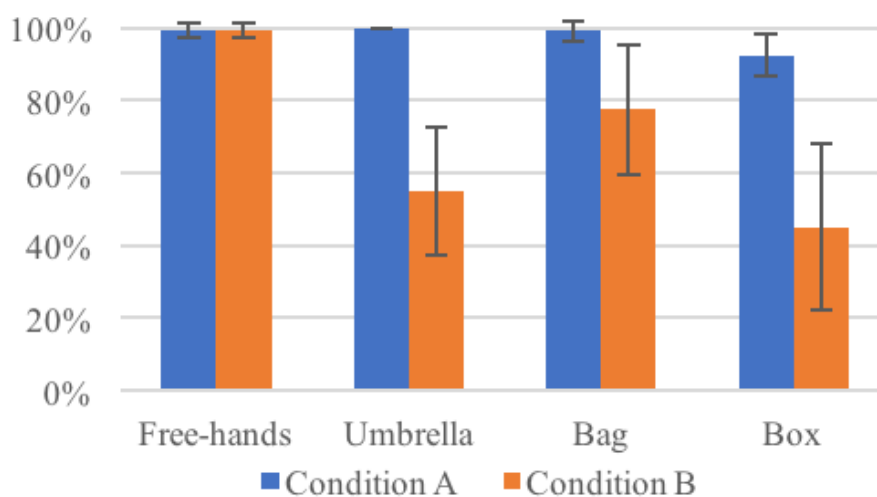


図 5.5: 条件 A, B における各状況の認識精度

考察

結果より, 条件 B, つまり物を把持した状況を未学習の識別器では, 物を把持した状況におけるハンドジェスチャを正しく分類することができなかったことがわかった. スマートホーム環境においてもユーザの行動によって様々な生活文脈 (料理, 食事, 掃除, 洗濯など) によってユーザが把持する物は異なる. このような状況下で未学習の状況下では, 正しくハンドジェスチャを分類することはできない. 本節では, ハンドジェスチャが誤分類された原因について考察する.

実験時の観察の結果から, 把持条件では, 把持する対象物が被験者の入力様式に影響を与えていた. 被験者の行動を具体的に述べると, 重心の不安定な傘を把持する条件では地面に対してそれを垂直に保とうとしたり, 重みのあるバッグを把持する条件ではハンドジェスチャの実行によって発生するバッグの揺れを防ごうとしたり, 箱を把持する条件では, 手首や腕, 上半身の姿勢を維持したりするなどの特徴がみられた. つまり, 把持する対象物の大きさや重さ, 形状だけでなく, それのもつ機能が被験者の入力様式に影響を及ぼす

結果となった。特に識別精度の低い傘の把持条件に着目すると、それを把持する被験者の行動から、傘としての機能を維持するためにその位置を固定しようとする振る舞いが観察された。その無意識的に働く力が意図しない筋収縮として検知され誤認識の原因となったと考えられる。その一方で、傘の把持条件において高い認識率が確認された被験者もあり、ハンドジェスチャの識別精度は対象物の形状のみならず、個人の経験や習慣による影響も受けることが理解できた。まとめると、把持する対象物の大きさや重さ、形状などの物理的な特性やそれのもつ機能等の特徴により、被験者の振る舞いや入力様式に影響を及ぼしていた。さらには各被験者のハンドジェスチャの形状（例として、FTジェスチャのような手を握り込む動作において親指を手のひらの中へ収めるか、外へ出すか、という違い）、物の把持の仕方（例として、バッグを把持する時に指に引っ掛けるように持つか、しっかりと握りしめるか、という違い）、といった経験や習慣の違いから、識別精度の傾向が異なるといった結果もみられた。

以上より、調査実験の結果から誤認識の原因となる問題が理解できた。しかし、日常に起こりうる全ての要因に対して事前に学習した識別器を用意することは難しい。この問題を解決するために次節の提案手法より解決を試みる。

5.2 提案手法: ハンドジェスチャ列認識手法

本節では、物を把持した状態における入力を可能とするため手法を提案する。具体的には、日常生活における筋肉の非活動部位を入力モダリティとして用いるハンドジェスチャ認識手法を提案する。

入力モダリティとは人がコンピュータへ入力を行う際に用いるための手段である。例えば、マウスやキーボード、タッチ入力、ハンドジェスチャ入力などは手の入力モダリティを用いる。他に足の入力モダリティを用いることでフットスイッチの入力や、口の入力モダリティを用いることで音声入力が可能となる。他にも様々な入力モダリティが存在し、また新たな入力モダリティが提案されている。入力モダリティは我々がコンピュータとインタラクションを行う唯一の手段であり、これらを使い分けたり同時に複数種類用いることで様々な状況で入力を行うことができる。一方で、これらの入力モダリティは日常生活の動作に占有されている場合に使用できない問題がある。この問題は既存の入力モダリティが対象の身体部位を占有しなければ使用できないため、日常生活の動作と共存できない点に原因がある。そのため、身体部位の一部を使用するような入力モダリティがあればこの問題は解決できる。そこで、本節では、日常生活における筋肉の非活動部位を入力モダリティとして用いることとした。

本提案手法の概念図を 5.6 に示す。この図ではユーザがバッグを把持している時の筋肉の使用量とバッグを把持しながらハンドジェスチャを行った時の筋肉の使用量を模式的に表している。この図からわかるように、ユーザは物の把持に常に 100%の筋肉を使用しているわけではなく未使用、つまり非活動な筋肉が存在する。この非活動な部位の筋肉（図 5.6 赤領域）を入力モダリティとして用いることができれば物を把持しながら入力を行うことができると考えられる。理論上は、上記のアプローチで実現可能ではあるが、実用にあたり懸念事項が多く存在する。例えば、入力モダリティを共存している状態では、偽陽性の問題、つまり意図しない入力は顕著に現れると考えられる。次節ではこれらの事項を

解決し、最終的に本提案手法の実装を記す。

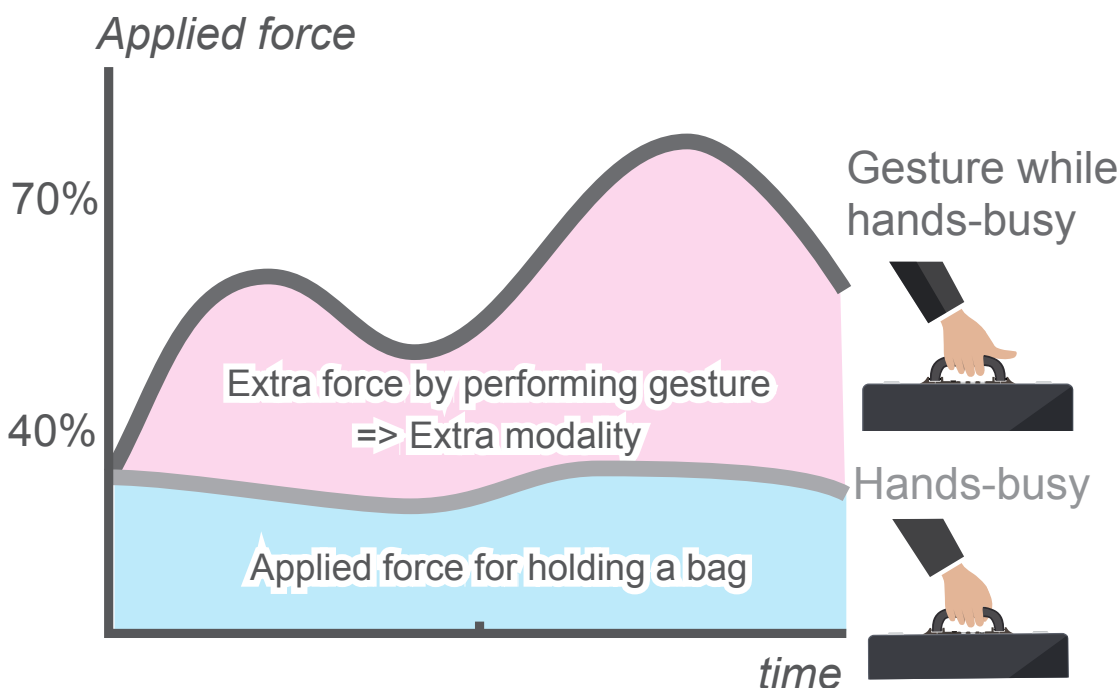


図 5.6: Comparison of the EMG for a gesture performed while busy and the typical hands-busy situation, demonstrating the extra capacity for input between the two. We thus use this capacity as an “extra modality.”

5.2.1 偽陽性に対する堅牢性

ジェスチャベースの入力手法においてユーザの日常動作は意図しない入力（偽陽性）を引き起こす問題となりうる。偽陽性を抑えるための様々な手法が提案されている [74]。しかし、Fukabori ら [19] はユーザの日常動作は活動によっては誤検知につながる事が多く適切な取り扱いが難しいと述べており、完全に偽陽性を無くすことは難しい。一方で、Whack Gestures[33] と呼ばれるインタラクション手法では偽陽性に対して堅牢性を確保している。Whack Gestures はハンドジェスチャの入力手法でありハンドジェスチャ入力の前後に「whack gesture (自身の身体を軽く叩くようなジェスチャ)」を入力する、つまりハンドジェスチャを3回行うことによって日常動作で意図しないハンドジェスチャが行われてもそれが3回繰り返されない限り入力として認識されないため、偽陽性に対して堅牢である。本提案手法では、この手法を参考にして3回のハンドジェスチャをハンドジェスチャ列として扱い、それを認識する入力手法を採用する。ハンドジェスチャ入力インタフェースにおいて、一つの入力に対して複数のハンドジェスチャの実行を要することはユーザビリティの低下を招くと考えられる。つまり、複数のハンドジェスチャの実行による偽陽性に対する堅牢性の確保はユーザビリティとトレードオフの関係にあるといえる。その点において、本提案手法は、意図しない入力が発生しやすい特性があるゆえに、偽陽性に対す

表 5.3: Three-gesture sequences used in this study. A and B in the table represent different hand gestures.

First gesture	Second gesture	Third gesture	Command
A	A	A	AAA
		B	AAB
	B	A	ABA
		B	ABB

る堅牢性の確保に重きを置く必要があり、ユーザビリティは限定的であるといえる。しかし、ユーザが物を把持しているなど実世界で主要なタスクを行っている最中に、複雑な操作は認知負荷の観点から難しいと考えられるため、シンプルなコマンドを起動するようなユースケースに限られると考えられる。その範囲内では本提案手法のユーザビリティは実用可能な範囲を確保できていると考えられる。

5.2.2 学習手法

センサベースのハンドジェスチャ認識は主に手の形状変化を電気信号や内部構造、外部構造を測定することによって実現されている。そのような認識手法は手の微細な変化を敏感に検出できる。一方で、そのような特性から物を把持しているような状況においては対象物の形状や重さ、持ち方、ユーザの姿勢によって、ユーザが同じジェスチャを行っても測定される信号は変化する。

Saponas ら [77] は物を把持しながら実行可能なハンドジェスチャ認識手法とジェスチャの学習方法を開発した。しかし、彼らの手法では把持の対象物は学習済みの物に限定されており、未学習の物では正確にハンドジェスチャを識別できない問題がある。人が把持する対象物は多種多様であり、どのような学習手法を用いたとしても全ての候補を事前に学習することはできない。さらに、このような学習手法は、把持の対象物 × 把持方法 × 対象物の重量 × 腕の姿勢のように多くのパラメータの組み合わせを考慮した非常に大きな学習データセットを必要とするため、把持の対象物を事前に学習するというアプローチは難しいと考えられる。

上記より、ハンドジェスチャ認識の精度を安定させるためには上記のパラメータから独立した学習手法が必要である。そこで本提案手法では複数回入力されるハンドジェスチャ間の差分を学習させる手法を考案した。具体的には二つのハンドジェスチャを比較してそれらが同じジェスチャであるか、異なるかを学習させる。同じハンドジェスチャ同士は状況が変わっても同じ入力信号となり、異なるハンドジェスチャ同士は異なる入力信号となるため、どのような状況においてもハンドジェスチャ同士の比較を正確に行えると考えられる。また、この学習手法が成り立つ前提条件は「上記のパラメータが全て同一である」であるが、本提案手法でハンドジェスチャ入力は短時間に連続で行われるため入力時間は5,6秒程度であると予測できる。この間に、把持の対象物や把持方法、対象物の重量や腕の姿勢が変わることはほとんど無いと考えられるため、実用においても前提条件を満たすことは容易であると考えられる。そのためハンドジェスチャの差分を学習する手法であれば物を把持した場合に考慮が必要なパラメータから独立した学習が可能である。提案手法では3つのハンドジェスチャ入力をを用いる。連続したハンドジェスチャの一番目と二番目

のハンドジェスチャ、二番目と三番目のハンドジェスチャのペアごとの差分から入力を識別する。簡単のために、AABやABAのようなモルルス信号に似たハンドジェスチャ列に対応するコマンドを表5.3に示す。

5.2.3 ハンドジェスチャ測定センサ

センサにはそれぞれ特性を持っており、用途にあったセンサを使う必要があるため、センサの選定は重要である。どのようなセンサが適しているかはどのような入力モダリティやジェスチャを用いるかに依存する。また、センサ感度も重要な要素である。力覚センサは感度が高く、微細なハンドジェスチャを検出することができるが、意図しない入力が識別される可能性が高くなる [14, 19]。

Saponas ら [77] はユーザが物体を把持しているときの筋電信号（以後、EMG）を用いたハンドジェスチャ認識の実現可能性を示した。EMGを用いることで、微細なハンドジェスチャを検出することができ、実世界での作業を妨げることなくコンピュータとのインタラクションを実現することができる。そのため、本提案手法ではEMGをハンドジェスチャ列の認識に用いることとした。また、EMGセンサとして8チャンネルのEMGを計測可能な市販のデバイスであるMyoを用いる。

5.2.4 入力の分割

ハンドジェスチャ列の認識には連続したハンドジェスチャ入力を分割するための区切り文字が必要である。区切り文字の具体例として、しきい値 [33] や無入力状態の学習 [19] が挙げられる。提案手法ではEMG信号に基づいて二つの状態を定義してそれを区切り文字として用いた。一つ目は、ハンドジェスチャが入力されている状態である「ジェスチャ状態」、二つ目は、ハンドジェスチャが入力されていない状態である「リラックス状態」である。EMG信号からハンドジェスチャとリラックス状態を区別するために、二人の被験者に素手の状況と物を把持（傘、バッグ、箱を把持）した状況の両方で、被験者がちょうどよいと感じる強さでハンドジェスチャを繰り返し行わせ、その時の前腕のEMGを測定した。次に、被験者に対して最大限の力を発揮してハンドジェスチャを行わせ、EMGを測定した。各状況における適切な強さと最大限の力を発揮した際のEMGから、発揮した力の強さを百分率で表す最大随意筋収縮（%MVC）を計算し、ハンドジェスチャを行った際にどの程度の力が発揮されるかを分析した。その結果、ジェスチャを行った際の平均値は94.8%MVCであり、力を入れていない時の平均値は32.2%となった。結果より、70%MVCをジェスチャ状態のしきい値として、40%MVCをリラックス状態のしきい値として定義した。つまり、EMGが40%以下から70%MVC以上に変化するとき、ハンドジェスチャが入力されたと認識した。また、意図しない入力を無視するために、3秒以内に別のハンドジェスチャが入力されなければ、ジェスチャの入力回数をリセットすることとした。

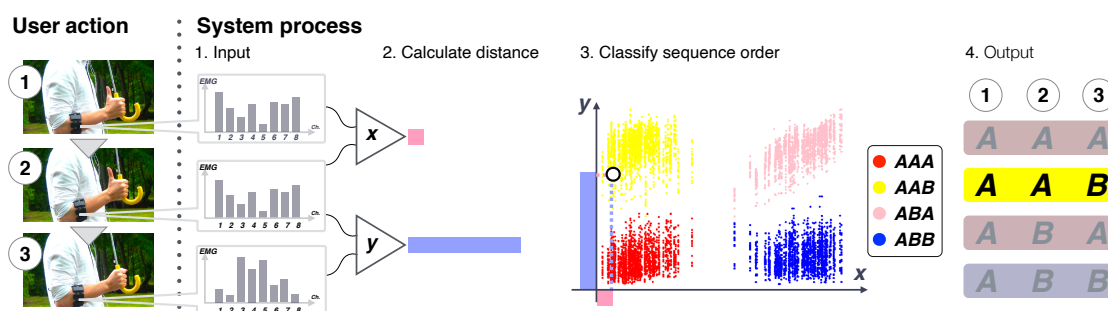


図 5.7: Interaction sequence for hand gesture sequence recognition. (1) The user performs three hand gestures as input. (2) The differences in each of the two pairs (the differences between the first and second gestures and between the second and third gestures) are calculated as Euclidean distances. The two distances are used as the classification features for machine learning. (3) The k-NN algorithm classifies the three-gesture sequence. The data points were acquired by the user to perform hand gesture sequences in the hands-free situation. (4) Finally, according to the classification result, an application-specific operation such as playing music is performed.

5.2.5 機械学習アルゴリズム

連続したハンドジェスチャ列を認識するための機械学習アルゴリズムとして、本提案手法ではk近傍法を用いた。一番目と二番目に入力されたハンドジェスチャ間、二番目と三番目に入力されたハンドジェスチャ間のユークリッド距離を用い、それらを合わせた二次元ベクトルを特徴量として用いた。

5.2.6 プロトタイプ

提案手法であるハンドジェスチャ列認識の技術的な実現可能性を評価するためにこれまでに記した要素をあわせてプロトタイプを実装した。図 5.7 に提案手法のインタラクションシーケンスを示す。1) ユーザは3つのハンドジェスチャから成るコマンドを実行する。それぞれのハンドジェスチャはEMG信号が40%MVC以下から70%MVC以上となったときに入力と認識される。次の入力はEMG信号が40%MVC未満となるまで認識されない。2) 連続するハンドジェスチャのペアごとの距離をユークリッド距離として計算し、これをハンドジェスチャ列の特徴量とする。3) k近傍法の学習モデルを用いて特徴量からハンドジェスチャ列がどのコマンドに対応するか推論する。4) 推論によって決定したコマンドに応じた特定の操作命令(例:音楽の再生,電話への応答,ショートメッセージの返信などのアプリケーション操作)がコンピュータへ入力される。

図 5.7 では、ユーザが「親指を立てる」、「親指を立てる」、「拳を握りしめる」というハンドジェスチャを行う例を示している。したがって、ハンドジェスチャ列として入力されるコマンドはAABとなり、ハンドジェスチャAは「親指を立てる」、ハンドジェスチャBは「拳を握りしめる」に対応する。なお、本提案手法では、ハンドジェスチャA,Bとして必ずしも前述の組み合わせを要求するわけではない。例えば、「手の甲側に曲げる(図

5.2 : Dorsal)」、「手の甲側に曲げる (図 5.2 : Dorsal)」、「手のひら側に曲げる (図 5.2 : Palmar)」という入力も同様に AAB として識別される。提案手法が特定のハンドジェスチャを学習するのではなく、連続したハンドジェスチャ間の差異から、ハンドジェスチャの並びのみを学習するためである。結果として提案手法は特定のハンドジェスチャに依存しないため、未学習の把持方法や物を把持した状況においても、ハンドジェスチャ列を正確に識別することができる。

5.3 予備実験: 提案手法に適したハンドジェスチャの選定

提案手法では2種のハンドジェスチャA, Bを選定する必要がある。そのため、提案手法に適したハンドジェスチャを選定することを目的とした予備実験を実施した。本提案手法の目的は、未学習の状況下においても物を把持しながら行うハンドジェスチャを高精度に識別することである。予備実験では把持していない状態のハンドジェスチャを学習させ、把持している状態のハンドジェスチャを正しく識別できるかを評価した。また、選定するハンドジェスチャを4種用意し、組み合わせ、ジェスチャ組として12種用意した。4条件(把持なし、傘・バッグ・箱を把持)でハンドジェスチャを実行した際のEMG信号を実験データとして収集し、把持なし条件を学習データとして、把持条件をテストデータとして用いた。次に、12種のジェスチャ組から作成したハンドジェスチャ列の識別精度を比較して最も精度の高いジェスチャ組を選定した。最後に、予備実験を通してハンドジェスチャの組み合わせや学習、識別精度の評価は被験者ごとに独立した行った。

5.3.1 データ収集実験

ハンドジェスチャ列データを収集するために、シミュレーションによるアプローチを採用した。具体的には、単一のハンドジェスチャを収集し、後でそれらを組み合わせるシミュレーション的にハンドジェスチャ列を作成した。データ収集としてハンドジェスチャ列(3回のハンドジェスチャ)を被験者に実行させる手法では、被験者に対してジェスチャ組の種類 × コマンドの種類 × 条件数 × 試行回数、の数だけハンドジェスチャを行わせる必要がある、非常に多くの時間と被験者への負担を要する。例えば、試行回数を100回とすると、 $12(\text{ジェスチャ組}) \times 4(\text{コマンド}) \times 4(\text{条件数}) \times 100(\text{試行回数}) = 19,200$ 回の試行が必要となる。一回のハンドジェスチャ列の実行を5秒と仮定すると約26時間の時間が必要である。このような実験を被験者に課すことは現実的ではないため、シミュレーション的にハンドジェスチャ列を作成する手法を採用した。具体的には各被験者から一つの状況あたり80回のハンドジェスチャ(4種のハンドジェスチャ × 20試行)を収集した。データ収集後に80種のハンドジェスチャを組み合わせることで、ハンドジェスチャ列データを $82,160$ 個分作成した(${}_{80}C_3 = 82,160$)。

ハンドジェスチャのデータはEMG信号を計測して表現した。計測にはEMG計測デバイスであるMyoを用い、下記の筋肉部位の活動を計測した。

- Anconeus
- Brachioradialis

- Extensor carpi radialis longus
- Extensor carpi ulnaris
- Extensor digitorum
- Extensor digiti minimi
- Extensor pollicis longus
- Extensor radialis brevis
- Flexor carpi radialis
- Flexor carpi ulnaris
- Flexor digitorum superficialis

EMG 信号の計測は Processing 3 を用いた自作の計測システムを用いた。システムでは Processing ライブラリである「myoOnEmgData²」を用いて Myo からの EMG 信号の受信を行い、8 チャンネルの EMG 信号データを収集した（サンプリング周波数：200 [Hz], サンプリング解像度：8 bit）。

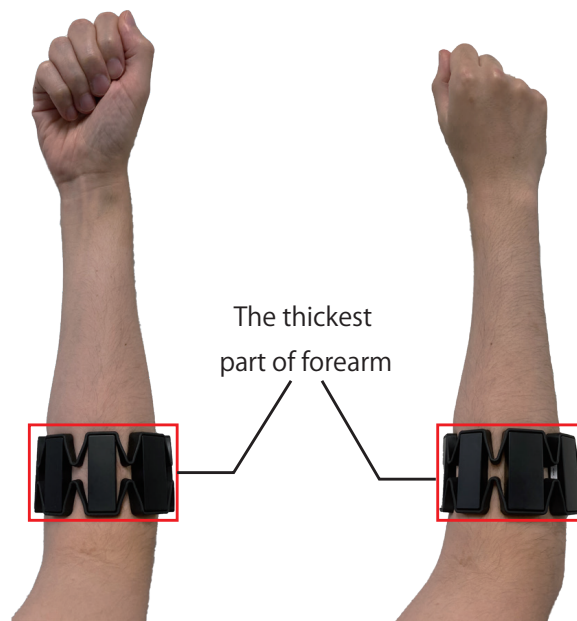


図 5.8: Placement of a the Myo device.

12 人の被験者（男性 7 名，女性 5 名，平均年齢 23.5 歳，全員右利き）がデータ収集に参加した。被験者は日常的にコンピュータを扱っており，計測システムの操作などに関する不慣れなどによる計測の不備は生じなかった。また，被験者は実験前に実験内容や計測

²<https://github.com/nok/myo-processing>



図 5.9: Hand gesture set for data collection.

データ，データの取り扱いに関する説明を受けており，実験の実施に同意した被験者のみ参加した．実験者は被験者に対して皮膚のアレルギー等がないことを確認した後に計測デバイス（Myo）を利き腕の前腕部（図 5.8 参照）に装着させた．

実験者は被験者に対して最大随意筋収縮（100%MVC）を計測するために計測デバイスの装着箇所に最大限の力を入れるように指示した．次に被験者は計測システムの指示に従って把持なし条件と把持条件の両方で利き手でハンドジェスチャを行った．被験者が行ったハンドジェスチャを図 5.9 に示す．被験者は 4 種のハンドジェスチャを 4 つの条件下で行った．被験者が実施したハンドジェスチャの詳細を以下に示す．

- Dorsal extension (DE): 手首を手の甲側に曲げる動作を行う．
- Palmar flexion (PF): 手首を手のひら側に曲げる動作を行う．
- Open hand (OH): 手を広げる動作を行う．把持条件では対象物を落とさないようになるべく把持なし条件と同じ動作を再現するように被験者に指示した．結果として親指を挙げるようなハンドジェスチャが行われた．
- Make a fist (MF): 拳を握りしめる動作を行う．

被験者一人あたりデータ収集は休憩時間も含めて 1.5 時間を要した．

5.3.2 ハンドジェスチャ列の作成

データ収集では 4 種のハンドジェスチャを収集したため 12 種のジェスチャ組が存在する (${}_4P_2$)．一例として，把持なし条件におけるハンドジェスチャ A=DE，ハンドジェスチャ B=PF の場合にどのように組み合わせるかを説明する．ハンドジェスチャ A，B ともに 20 回分のデータがあり，総当たりで順列を作成する．各ハンドジェスチャ列の計算式を表 5.4 に示す．ハンドジェスチャ列のコマンド AAA はハンドジェスチャ A=DE の 20 データか

表 5.4: The equation and resulting number of hand gesture sequence combinations.

	Equation	Number of gesture sequence
AAA	${}_{20}P_3$	6840
AAB	${}_{20}P_2 * {}_{20}P_1$	7600
ABA	${}_{20}P_2 * {}_{20}P_1$	7600
ABB	${}_{20}P_1 * {}_{20}P_2$	7600

ら3つを選んだ順列により作成した. 同様にコマンド (AAA, ABA, ABB) は A=DE の 20 データから 2 つ選んだ順列と B=PF の 20 データから 1 つ選んだ順列の掛け算により作成した. 最後にコマンド ABB は A=DE の 20 データから 2 つ選んだ順列と B=PF の 20 データから 2 つ選んだ順列の掛け算により作成した. 結果として, 1 条件につき 29,640 個のハンドジェスチャ列のデータを得た.

5.3.3 機械学習アルゴリズム

ハンドジェスチャ列の分類には k 近傍法による分類器を実装して行った. 10-fold 交差検証によってパラメータと分類モデルのチューニングを行った. 具体的には, k の値を 5~30 まで 1 刻みでグリッドサーチを行い, 最も精度の高くなる k の値と分類モデルを評価に使用した. パラメータチューニングは被験者ごとに独立して行っているため, k の値はそれぞれ異なるが, 平均して 12.9 であった (標準偏差は 10.7). 分類モデルの学習には把持なし条件のハンドジェスチャ列データを用い, 精度の評価には把持条件 (傘, バッグ, 箱を把持の 3 条件) のハンドジェスチャ列データを用いた.

5.3.4 特徴量の計算

収集したハンドジェスチャ列データから k 近傍法による分類器の学習・推論のための特徴量を設計した. まず, 収集した EMG 信号の源波形に含まれるノイズを低減するために, 各チャンネルの信号の RMS 値を以下のように算出した.

$$RMS_i = \sqrt{\frac{1}{N} \sum_{j=1}^N EMG_{ij}^2}, \quad (5.3)$$

$$\mathbf{RMS} = [RMS_1, RMS_2, \dots, RMS_8], \quad (5.4)$$

ここで, EMG は EMG 信号の源波形の値, i はチャンネル数, N はサンプル数である. 上記の式を計算することで 8 チャンネルの EMG 信号から RMS 値の 8 次元のベクトルを得た.

次に, ハンドジェスチャ列の連続する 2 つのハンドジェスチャ間ユークリッド距離を計算した. 距離は RMS 値ベクトルから求め, 以下のように計算した.

$$d(\mathbf{RMS}_p, \mathbf{RMS}_q) = \sqrt{\sum_{i=1}^8 (RMS_{pi} - RMS_{qi})^2}, \quad (5.5)$$

表 5.5: Mean recognition accuracy of hand gesture pairs.

A \ B	B				Mean
	DE	PF	OH	MF	
DE	–	90.0%	73.4%	73.9%	79.1%
PF	89.4%	–	78.2%	74.9%	80.8%
OH	72.5%	82.0%	–	75.0%	76.5%
MF	72.2%	78.4%	77.4%	–	76.0%
Mean	78.0%	83.4%	76.3%	74.6%	78.1%

ここで、 RMS は RMS 値の 8 次元ベクトル、 i はチャンネル数、 p と q は入力されたハンドジェスチャの順番を表す。前節で述べたように、特徴量は一番目のハンドジェスチャと二番目との距離、二番目と三番目との距離からなる。したがって、特徴量は次式より抽出された。

$$feature = [d(RMS_1, RMS_2), d(RMS_2, RMS_3)]. \quad (5.6)$$

結果として、2次元の特徴量を得た。

5.3.5 結果

表 5.5 に各状況における平均識別精度をジェスチャ組ごとに示す。最も精度の高いジェスチャ組は、ハンドジェスチャA=DE, ハンドジェスチャB=PF であり、把持条件である 3 条件の平均精度は 90.0%であった。平均精度は 4 種のハンドジェスチャ列 (表 5.3 参照) の平均値である。結果より、ハンドジェスチャDE と PF を最も精度の高いジェスチャ組と決定した。

5.4 評価実験

前節の予備実験では、ハンドジェスチャ列認識に用いるジェスチャ組 (ハンドジェスチャA=DE, ハンドジェスチャB=PF) を決定した。さらに、90.0%の平均識別精度でハンドジェスチャ列を識別可能であり、未学習の把持条件において高い精度で識別可能であることが確認できた。しかし、ハンドジェスチャ列をシミュレーション的に作成したため、得られたデータセットには均質性が存在し、評価された学習モデルでは過学習を引き起こしていた可能性がある。さらに、把持条件の数は 3 種類に限られていた。そこで、提案手法であるハンドジェスチャ列認識が実世界上で利用可能であるか、高い精度を実現可能かを評価する実験を行った。具体的には、把持の古典分類学を参考に体系的に網羅された把持条件を設定し、実世界上で入力されたハンドジェスチャ列を学習し、識別精度を評価した。

予備実験時と同様に把持なし条件を学習データ、把持条件をテストデータとしてハンドジェスチャ列の収集を行うユーザスタディを実施した。実験では、実際の使用を想定し、ハンドジェスチャ列の識別精度と入力時間を計測し、実現可能性を評価した。把持条件として 6 つの対象物を用意し、被験者に各状況下でハンドジェスチャ列を実行させた。

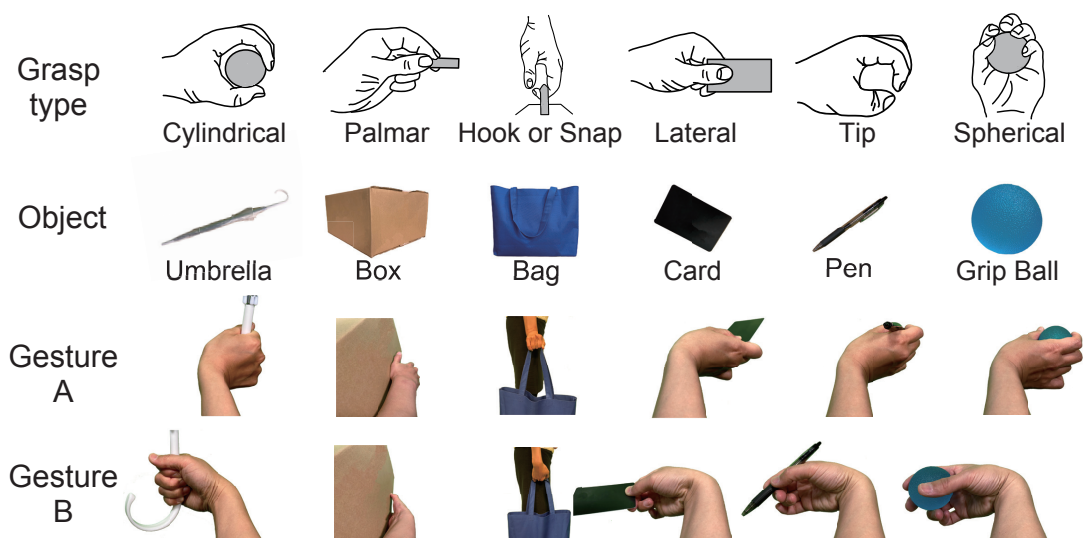


図 5.10: Grasp types, grasped objects, and gesture set considered in the user study. The grasp type graphs are classic taxonomies [80]. The object in the row denotes the object grasped by the participants. Gestures A and B in the row were the dorsal extension and palmar flexion gestures, respectively, which were determined in the gesture selection phase of this study.

5.4.1 被験者

12人の被験者（男性：8人，女性：4人，平均年齢23.2歳）が実験に参加した。被験者は事前に実験内容や測定データ，データの取扱について説明を受け，実験の実施に同意した被験者のみ参加した。被験者は全員日常的にコンピュータを使用していた。また，被験者に皮膚アレルギーなどが無いかをEMG計測装置を装着する前に確認した。ハンドジェスチャ列の実行は被験者の利き腕（10人の被験者が右利きであり，2人の被験者が左利きであった）で行った。評価実験の所要時間は休憩含めて一人あたり1.5時間であった。

5.4.2 把持条件とハンドジェスチャ列

図 5.10 に把持条件と被験者が行うハンドジェスチャを示す。これらの物はSchlesingerの把持の古典分類[80]をもとに設定した。図ではジェスチャ組としてハンドジェスチャA（DE）とハンドジェスチャB（PF）について，それぞれの把持条件におけるハンドジェスチャの実行例を示している。被験者は表 5.3 に示す4種のハンドジェスチャ列を各条件下で行った。各把持条件の詳細を以下の通りである。

- Cylindrical: 円筒状の物体を把持する
- Palmar: 手のひらを対象物へ接地させながら把持する
- Hook or Snap: 重さのある物体をぶら下げるように把持する

- Lateral: 薄く平面の物体を把持する
- Tip: 細長い物体をつまむように把持する
- Spherical: 球状の物体を把持する

評価実験では学習用データとして把持なし条件にて被験者一人あたり 80 個のハンドジェスチャ列のデータ (4 種のハンドジェスチャ列 × 20 試行) とテストデータとして各把持条件にて被験者一人あたり 80 個のハンドジェスチャ列のデータを収集した。

5.4.3 データ収集と分類

評価実験でははじめに把持なし条件におけるハンドジェスチャ列を収集した。次に 6 つの端条件からハンドジェスチャ列を収集した。被験者の実験中の疲労が識別精度やハンドジェスチャ列の入力時間に与える影響を考慮して、把持条件の順番は被験者ごとにランダムとした。ハンドジェスチャ列を入力した際の入力分割は提案手法の節で説明した手法で行った。収集したハンドジェスチャ列は、予備実験時に用いた機械学習手法と特徴量生成の方法で行った。

5.4.4 実験手順

被験者は実験室に入室後、評価実験に関する説明を受けた。具体的には、実行タスク (ハンドジェスチャ列の試行回数) や把持条件、データの取り扱いなどの同意事項であり、説明後実験参加に同意可能かを確認した、次に口頭で皮膚アレルギーの有無を確認し、EMG 計測デバイス (Myo デバイス) を被験者の利き腕の前腕部に装着した。次にハンドジェスチャ列の実行方法を被験者に教示した。その後、利き腕に最大限の力を加えてもらい、MVC を測定した。その後、ハンドジェスチャ列の入力を被験者に練習させ、入力に問題がないことを確認してから評価実験を開始した。評価実験では、把持なし条件を最初に行い、その後、把持条件を被験者ごとにランダムな順番で行った。実験中、被験者の希望に応じて適宜休憩を取り、疲労によるハンドジェスチャ列の実行への影響を軽減した。計測終了後、実験内容に関する口頭インタビューを行い、実験を終了した。

5.4.5 結果

表 5.6 に被験者ごとの 6 つの把持条件における識別精度を示す。各識別精度は 4 種のハンドジェスチャ列の 80 入力分の平均値である。全体の平均精度は 93.13% であり、未学習の把持条件下であっても高い識別精度で入力を認識可能であることがわかった。一要因分散分析より各把持条件の関係性を明らかにしたところ、把持条件 (箱) の識別精度が他条件より有意に低いことがわかった [$F(5, 55) = 9.67, p < .05$]。把持条件 (箱) は他の把持条件に比べて精度が低い原因についてみると、4 人の被験者が著しく低いことがわかった (表 5.6 下線部)。次に表 5.7 に把持条件ごとのハンドジェスチャ列の平均入力時間を示す。全条件の平均入力時間は 3.96 秒であり、一要因分散分析の結果、入力時間に関して把持条件間の有意差は認められなかった [$F(5, 55) = 1.31, n.s.$]。

表 5.6: Mean gesture recognition accuracy per participant and grasp condition (P in the column denotes the participant).

#	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	Mean
Bag	96.30%	71.30%	93.80%	82.50%	97.50%	82.50%	97.50%	83.80%	98.80%	73.80%	91.30%	97.50%	88.88%
Ball	98.80%	98.80%	100.00%	100.00%	100.00%	98.80%	98.80%	98.80%	98.80%	97.50%	100.00%	100.00%	99.19%
Box	6.30%	8.80%	7.50%	97.50%	97.50%	87.50%	87.50%	91.30%	100.00%	78.80%	80.00%	1.30%	77.00%
Card	97.50%	98.80%	100.00%	100.00%	98.80%	95.00%	98.80%	97.50%	100.00%	97.50%	98.80%	100.00%	98.56%
Pen	100.00%	98.80%	98.80%	98.80%	97.50%	98.80%	100.00%	96.30%	100.00%	98.80%	100.00%	98.80%	98.88%
Umbrella	100.00%	93.80%	97.50%	100.00%	95.00%	95.00%	100.00%	97.50%	100.00%	80.00%	98.80%	97.50%	96.26%
Mean	93.15%	86.72%	87.93%	96.47%	97.72%	92.93%	97.10%	94.20%	99.60%	87.73%	94.82%	89.18%	93.13%

表 5.7: Mean input time and standard deviation (S.D.) for hand gesture sequences for each grasp condition

Object	Umbrella		Box		Bag		Card		Pen		Ball	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Input time (s)	4.00	0.82	4.13	0.98	4.00	0.79	3.81	0.68	3.99	0.84	3.81	0.90

評価実験の結果、実装したプロトタイプは把持なし条件のハンドジェスチャ列で学習した分類モデルを用いて、ほぼ全ての把持条件においてハンドジェスチャ列を正しく分類でき、かつ現実的な時間で入力可能であることが示された。また、ハンドジェスチャ列認識手法を用いることによる疲労に関して、被験者は実験中に平均 0.83 回休憩を取り（6 名は休憩無し）、疲労によってハンドジェスチャの継続が困難となった被験者はいなかった。

5.4.6 考察

評価実験では識別精度・入力時間については技術的な実現可能性を示すことができたと考えられる。一方で、ハンドジェスチャ列認識手法は一部の条件における精度低下や実生活の環境における検証など改善の余地がある。本節では、提案手法で得た結果や解決すべき課題について議論する。

5.4.7 ハンドジェスチャと識別精度について

提案手法は、未学習の把持条件下でハンドジェスチャ列を正確に識別することができ、平均精度 93.13%を実現したことが評価実験で示された。一方で、評価実験から、把持条件（箱）の識別精度が他条件より有意に低いことが示された。把持条件（箱）において精度が低かった 4 人の被験者について分析し、精度低下の原因を考察する。図 5.11 に把持条件（箱）における 4 人の被験者のハンドジェスチャ列の識別精度を混合行列で示す。混合行列ではハンドジェスチャ列 AAB, ABA, ABB が AAA と誤分類され、これが誤分類の大部分を占めている。ハンドジェスチャ列 AAA は同じハンドジェスチャが 3 回繰り返されることを意味する。したがって、ハンドジェスチャ A と B を行った際に、筋活動において明確な差異が出なかったため、ハンドジェスチャと A と B が同じハンドジェスチャであると識別されてしまったことが誤分類の原因であることがわかった。これが箱を把持する条件のみに起きた理由は、箱の把持方法にあると考えられる。被験者は箱の面を手のひらに接地させて包み込むように把持を行ったため、手首が可動域が制限されていた。評価

実験で用いたハンドジェスチャはどちらも手首を曲げる必要があったため、十分にハンドジェスチャを実行することができず、これが筋活動の差異が低減した理由であると考えられる。この問題は、手首を使わないハンドジェスチャの使用など、筋活動、EMG 信号の違いがより明確なジェスチャ組を使用することで解決できると考えられる。また、圧力センサのような僅かな動きでも敏感なセンサを用いることでも解決できる可能性がある。そのため、EMPress[54]のような圧力センサと EMG センサを組み合わせるようなアプローチが有効であると考えられる。

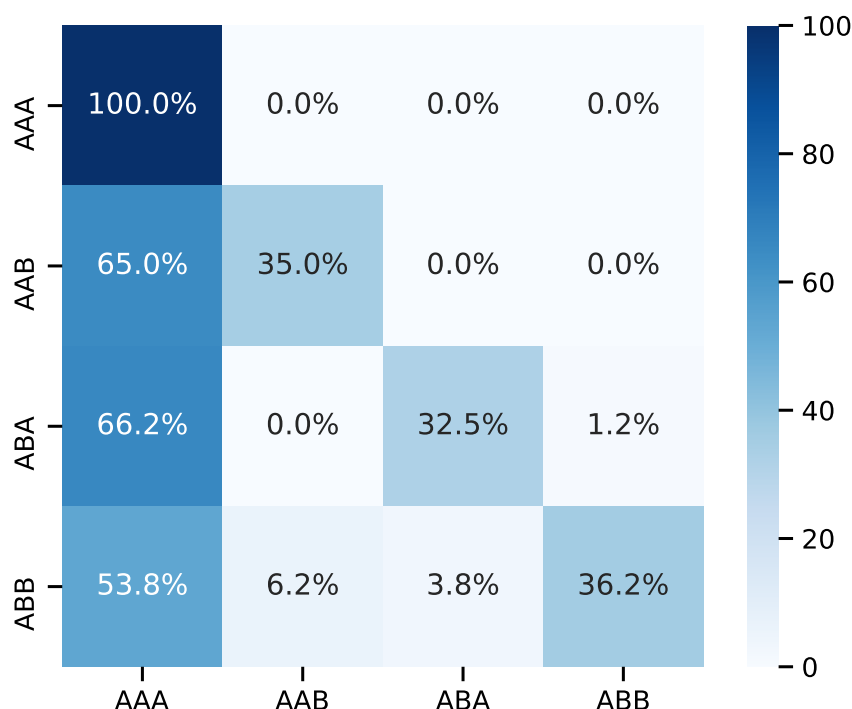


図 5.11: Confusion matrix of gesture recognition for the four participants who showed low accuracy while grasping a box.

5.4.8 入力時間の妥当性

ハンドジェスチャ列の入力にかかる時間は、ユーザビリティに大きな影響を与える。通常のハンドジェスチャ入力と比較して、ハンドジェスチャ列入力ではハンドジェスチャを3回実行するため単純計算で3倍の時間を要する。そのため、評価実験で得られた入力時間が現実的に妥当であるかを考察する。実験で得られたハンドジェスチャ列入力の平均時間は3.96秒であり、単一のハンドジェスチャ入力（1.5秒程度）と比較すると入力時間は長い。しかし、ハンドジェスチャ列認識手法を使用する場面を考えると、物を把持している状況などユーザが主要なタスクを行っている最中である。その場面下でコンピュータへ

命令するタスクは簡潔なものであると考えられる。例えば、スマートスピーカへの音声入力など1つのコマンドで命令が完結するようなタスクである。スマートスピーカへの入力時間は3~5秒程度（ウェイクワード³を含む）であり、ハンドジェスチャ列認識手法と同程度であることから、入力時間は現実的に妥当であると考えられる。

5.4.9 疲労の影響

ハンドジェスチャ列入力では単一のハンドジェスチャ入力と比較して、入力に多くの労力がかかる。そのため、ハンドジェスチャ列入力が被験者にどの程度の疲労を与えたかについては検証する必要がある。被験者は評価実験の90分間で560回のハンドジェスチャ列を行った（1把持なし条件 × 80試行 + 6把持条件 × 80試行）。ハンドジェスチャ列1回の入力で3回のハンドジェスチャを入力する必要があるため、合計で1680回のハンドジェスチャを行った。実験中、被験者は平均0.83回の休憩を取り、疲労によって入力が困難となった者はいなかった。しかし、実験後の口頭インタビューにて、全ての被験者が実験中にいくらかの疲労を感じていたと述べた。また、多くの被験者が把持条件によって疲労を感じる度合いは変わると述べた。全ての被験者が重い物（把持条件（バッグ））と不安定な物（把持条件（傘））を把持する条件で他条件より疲労を感じたと述べた。これより、ハンドジェスチャ列認識手法は、文字タイピングのような入力頻度の高いタスクより、音楽プレイヤーの操作のような一回の入力で命令を実行可能なシンプルなタスクの方が適していると考えられる。

5.4.10 ハンドジェスチャ列の設計

一般に、ジェスチャ入力はコントローラとなる端末を把持することなく使用できることが利点として挙げられるが、欠点としてユーザが操作方法のジェスチャを覚える必要がある。提案手法は単一のハンドジェスチャより入力手法が複雑であるため、それと比較してユーザが入力を覚えることは難しくなる。しかし、関連研究として人の眼球運動を入力として用いるアイジェスチャを複数回入力する手法では、4種のコマンドで音楽プレイヤーの操作を実現しており、提案手法であるハンドジェスチャ列入力においても同様に実現可能であると考えられる。

5.5 フィールド実験

本実験では、差分ジェスチャ認識法の評価をあらかじめ取得したジェスチャデータから精度を算出しているため、実用とは異なる可能性がある。そのため、本研究では予備実験として差分ジェスチャ認識法のプロトタイプシステムを実装し、精度を評価した。

³音声コマンドを開始するためのキーワード

5.5.1 概要

この評価実験では、提案手法のジェスチャ認識手法について、実環境に近い設定で認識率を評価することで、提案手法の実現可能性を示す。具体的には、図 5.12 に示した実験環境 (廊下) を歩行しながら、実験と同様のジェスチャを用いて、素手と 3 種の物を把持した状況において評価を行った。入力するジェスチャ組は 2 種の異なるジェスチャからなる。実験でも用いた 4 種 (図 5.9) の内、異なる 2 種の全ての組み合わせである以下の 6 種とした。

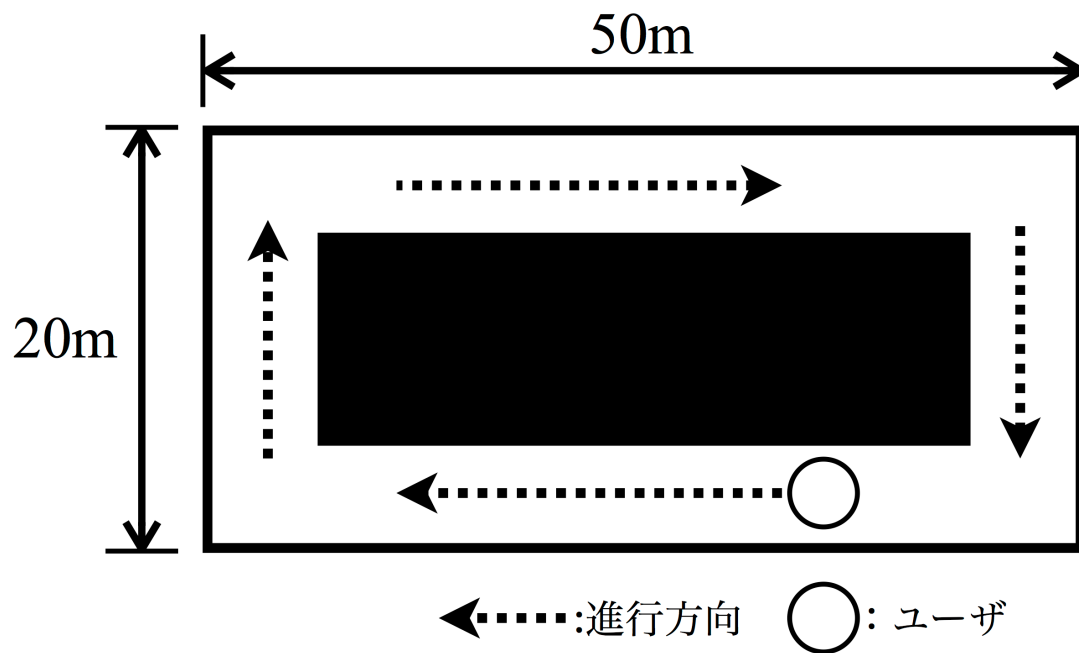


図 5.12: 実験環境の模式図

- Dorsal-Palmar
- Dorsal-OpenHand
- Dorsal-Fist
- Palmar-OpenHand
- Palmar-Fist
- OpenHand-Fist

参加者は 5 名 (男性 5 名, 平均年齢: 24 歳) とし, 筋電センサを用いたジェスチャ入力について十分練習を行った後に実験を実施した。全ての参加者は, 情報系の大学に所属しており, 日常的にコンピュータを使用している。

表 5.8: ジェスチャ組×実験条件の認識率の結果 (%)

	素手	傘	バッグ	箱
Dorsal-Palmar	99.2	94.4	99.2	77.6
Dorsal-OpenHand	86.4	69.2	86.0	52.8
Dorsal-Fist	83.2	76.3	81.1	53.9
Palmar-OpenHand	87.4	76.2	79.6	51.4
Palmar-Fist	88.4	75.4	79.4	48.5
OpenHand-Fist	83.9	72.9	81.2	49.5

次に、実験手順について説明する。図 5.12 の廊下で、各ジェスチャ組 (6 種) について各状況 (1 種: フリーハンド, 3 種: 物を把持) の下それぞれ 25 回ずつ入力させた。実験の所要時間は 1 人あたり 1 時間 30 分程度であった。

5.5.2 結果

実験の結果得られた認識率の内訳を表 5.8 に示す。結果の分析は図 5.13 に示す (*: $p < 0.05$)。左側は認識率の平均値を条件毎に比較したグラフ、右側はジェスチャ組毎に比較したグラフとなる。エラーバーは標準偏差を表している。2 要因参加者内 (ジェスチャ組: 6 種 × 実験条件: 4 種) における分散分析の結果、有意なジェスチャ組 ($F(5,20) = 12.45$, $p < 0.01$) の主効果と実験条件の主効果 ($F(3,12) = 14.40$, $p < 0.01$) が認められた。さらに、ホルム法による多重比較によって群間の関係を示したところ、ジェスチャ組において、[DR-PL: 92.6%, DR-OP: 73.6%, DR-FT: 73.6%, PL-OP: 73.7%, PL-FT: 72.9%, OP-FT: 71.9% ($F(5,20) = 12.45$, $p < 0.01$)] となり、多重比較から DR-PL の認識率が他のジェスチャ組の認識率より 5% 水準で有意に高いことが認められた。また、条件において、[素手: 88.1%, 傘: 77.4%, バッグ: 84.4%, 箱: 55.6% ($F(3,12) = 14.40$, $p < 0.01$)] となり、多重比較から箱を把持した条件の認識率が他の条件の認識率より 5% 水準で有意に低いことが認められた。

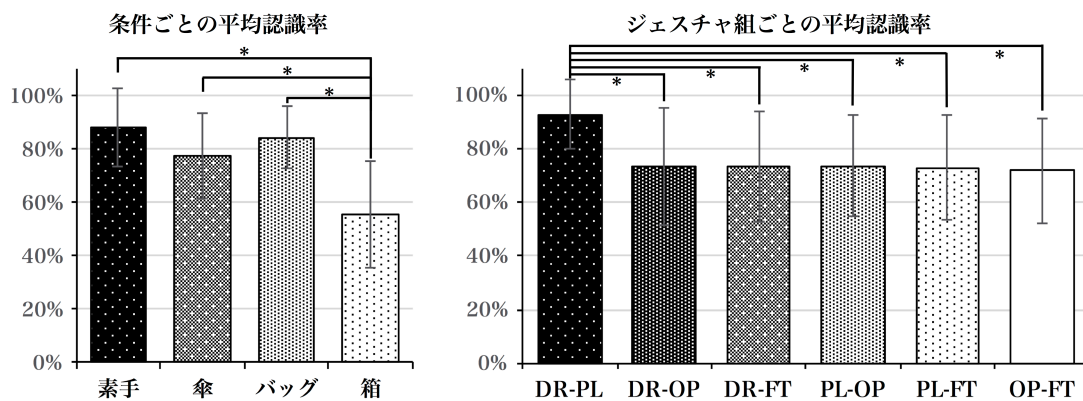


図 5.13: 評価実験の結果

5.5.3 考察

素手条件において平均の認識率は88.1%となった。この結果は、先行研究 [77] と比較しても低い認識精度ではないことが理解できる。先行研究とは実験の条件や使っているハードウェア、認識アルゴリズムなど本研究とは異なる点が多数あるため、単純に比較することはできないが、「ユーザが行うジェスチャをシステムが正しく認識できる」という点では同じである。そのことから本研究の提案手法の実現可能性が示された。次に把持条件において、多重比較の結果から傘とバッグを把持した条件については素手条件との間に有意な認識率の差は認められなかった。これより、この2つの条件においては物を把持することによる認識率の減少がみられなかったと理解できる。この結果から、特に実験1で認識率の低かった傘を把持した条件においては、提案手法を用いた実用に近い場面で精度の改善を確認することができた。また、本実験で行ったジェスチャ組については、多重比較の結果から Dorsal-Palmar が他のジェスチャ組と比較して有意に認識率 (92.6%) が高かった (図 5.13: 右図)。このことから、実験設定の範囲内においては、様々な要因に関わらず、高い精度で識別できるジェスチャ組の存在が確認された。

その一方、実験結果から、さらに調査や検証の必要な点や問題点が示された。把持させた対象物の中でも、特に箱を把持した条件におけるジェスチャ組が有意に低い認識率 (55.6%) となったことが示された。評価実験では、実環境に則した設定で評価を行うため、歩行時におけるジェスチャの認識精度を調査した。その際、箱を把持した条件は他の物を把持した条件に比べて、手や腕だけではなく姿勢などの自由度も奪われていたことが影響し、ジェスチャ組の認識率の著しい低下を招いたと考えられる。実環境において、姿勢が固定されるような状況は度々起こりうるため、新たな課題として姿勢変化による前腕の筋電に与える影響なども調査する必要がある。加えて、ユーザの様々なニーズに応えるため、提案手法の入力パターンについて、より多くの種類を確保することも重要な課題となる。

5.6 本章のまとめ

本章では、システムの実現可能性を示すために人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討を行った。本研究では様々な環境下で入力可能なインタフェースを目指し、関節動作を伴わない筋肉の等尺性収縮に着目しインタフェースの構築を行った。まずインタラクションがなぜ阻害されるのか、その阻害の原因は何であるかについて調査を行った。その結果、把持する対象物の大きさや重さ、形状などの物理的な特性やそれのもつ機能等の特徴により、被験者の振る舞いや入力様式に影響を及ぼしており、それが誤認識の原因となっていたことが理解できた。この問題を解決するための手法として、筋肉の非活動部位を入力モダリティとして用いる新たな入力手法とそれを認識するための筋電入力インタフェースを構築した。実験室内で行われた評価実験の結果、93.13%の入力識別精度を示した。従来手法で同様の分類タスクを行った際の精度は 59.17%であったため、33.96%の精度向上を実現できた。また、フィールド実験として屋内の廊下を歩きながら、提案手法の入力精度、識別精度を評価する実験を行い、評価実験と比較して精度の低下はあるものの、実世界での実現可能性を確認することができた。

第6章 総合的な議論

6.1 本研究の成果とその意義

本研究では、人間の身体の内的変化を着目した社会的つながりを活性化させる環境知能システムの実現を目的とした。この実現のために、まず複数人の社会的活動において、その活性と内的変化の関係性および活性化に必要な要素抽出を行った。具体的には、様々な目的を持った立場の異なる人同士の言語的コミュニケーション課題として人狼ゲームを用い、皮膚電気活動を用いたプレイヤーの内的変化と、プレイ状況との関係を調べた。分析の結果、人狼ゲームではプレイヤー間で共通した重要なイベントにおける内的変化の同期がみられ、勝率の高いプレイヤーはその他のプレイヤーより有意に発言数が多いことがわかり、他プレイヤーの判断に必要な情報を多く伝えることができていたことが理解できた。この結果から、人間の行動変容（実験においては他プレイヤーの説得）には、適切な場面における適切な情報の提示が必要であることがわかった。

この知見をふまえて、次の研究で環境知能システムが実現された状況を想定して、システムによるユーザ支援がユーザの社会的つながりの活性化に貢献し得るのかを調査した。具体的には、複数人コミュニケーション課題としてレクリエーションを対象とし、レクリエーションに対する動機の向上を目標として設定して評価した。本研究ではレクリエーション援助者としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者の動機について調査した。前述の研究で得られた環境システムの設定として「適切な場面」と「適切な情報の提示」について、それぞれ「一方の参加者がゲーム中に不利な状況下となった場面」と「不利な状況下にプレイヤーが有利になる情報の提供を設定し、ロボットに対して振る舞わせたときの効果を評価した。評価実験の結果、ロボットが行った介入によって、レクリエーションの参加や勝利へのモチベーションが有意に高まったことが示された。これらの研究から、人間の身体の内的変化を活用してユーザの社会的なつながりを活性化させる環境知能システムの効果を検証できた。

最後に、環境知能システムの実現可能性を示すために人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討を行った。これまで提案されてきたシステムでは、ユーザの情報収集手法として、主に人間のコミュニケーションメディアを利用した手法が用いられてきた。しかし従来の手法では、実世界の人間の社会的活動を阻害するといった課題がある（他者との会話中に音声入力インタフェースを使用できない、両手が塞がった状態では手指を使った操作ができない等）。そこで本研究では、様々な環境下でもシームレスに入力可能なインタフェースを目指し、関節動作を伴わない筋肉の等尺性収縮に着目しインタフェースの構築を行なった。そして提案するインタフェースが実世界で起こり得る様々な状況下において、人間の主活動を阻害せずに入力可能（精度 93.13%）であるこ

とを示した。

以上をまとめると、

- 様々な目的を持った立場の異なる人同士の言語的コミュニケーション課題として人狼ゲームを用い、皮膚電気活動を用いたプレイヤーの内的変化と、プレイ状況との関係を調べることで、人間の行動変容（実験においては他プレイヤーの説得）には、適切な場面における適切な情報のが必要であることがわかった。
- 複数人コミュニケーション課題としてレクリエーションを対象とし、前述の研究で得られた環境システムの設定として「適切な場面」と「適切な情報の提示」について、それぞれ「一方の参加者がゲーム中に不利な状況下となった場面」と「不利な状況下にプレイヤーが有利になる情報の提供を設定し、ロボットに対して振る舞わせたとときの両参加者の動機に与える影響を評価し、ロボットが行った介入によって、レクリエーションの参加や勝利へのモチベーションが有意に高まったことが示した。
- 環境知能システムの実現可能性を示すために必要なユーザのアノテーション入力としてのインタフェースの検討を行い、提案するインタフェースが実世界で起こり得る様々な状況下において、人間の主活動を阻害せずに入力可能（精度 93.13%）であることを示した。

以上本学位論文では、これらの研究によって、社会的なつながりを活性化しうる環境知能システムの実現可能性を示した。本研究では、エンタテインメントに着目して研究を行ったが、社会的なつながりを活性化することは娯楽以外にも教育や医療など様々な分野において応用可能性があり、本研究の知見はそれらの発展可能性を有す重要な貢献である。

第7章 おわりに

本論文では、人間の身体の内的変化に着目した社会的つながりを活性化させる環境知能システムの実現を目的として、社会的活動時の人の内的変化とその要素の抽出、それを踏まえて環境知能システムによる支援が人の社会的つながりの活性化に貢献し得るかの調査を行った。さらに環境知能システムの実現可能性を示すために、環境知能システムの実現可能性を示すために必要なユーザのアノテーション入力として、人の社会的活動を阻害せずシームレスに入力できるインタフェースの検討を行った。以降、各章を外観し、本論文をまとめる。

第3章では、複数人の社会的活動において、その活性と内的変化の関係性および活性化に必要な要素抽出を行った。具体的には、様々な目的を持った立場の異なる人同士の言語的コミュニケーション課題として人狼ゲームを用い、皮膚電気活動を用いたプレイヤーの内的変化と、プレイ状況との関係を調べた。分析の結果、人狼ゲームではプレイヤー間で共通した重要なイベントにおける内的変化の同期がみられ、勝率の高いプレイヤーはその他のプレイヤーより有意に発言数が多いことがわかり、他プレイヤーの判断に必要な情報を多く伝えることができていたことが理解できた。この結果から、人間の行動変容（実験においては他プレイヤーの説得）には、適切な場面における適切な情報の提示が必要であることがわかった。

この知見を踏まえて、第4章では環境知能システムが実現された状況を想定して、システムによるユーザ支援がユーザの社会的つながりの活性化に貢献し得るのかを調査した。具体的には、複数人コミュニケーション課題としてレクリエーションを対象とし、レクリエーションに対する動機の向上を目標として設定して評価した。本研究ではレクリエーション援助者としてヒューマノイドロボットを用いて、2人で行うレクリエーション（「だるまさんがころんだ」をベースとしたVRゲーム）中に両参加者の動機について調査した。第3章で得られた環境知能システムの設定として「適切な場面」と「適切な情報の提示」について、それぞれ「一方の参加者がゲーム中に不利な状況下となった場面」と「不利な状況下にある参加者が有利な状況となるための仕掛けに関する情報」の提供を実装した。具体的には、環境知能システムの仲介役としてのロボットが、上記の設定に応じた振る舞いを取り、その際の動機に及ぼす効果を評価した。評価実験の結果、ロボットの上記設定に応じた振る舞いがレクリエーションの動機（勝利への執着やエンタテインメント性など）を有意に高めたことが示された。この実験結果から、人間の身体の内的変化に着目した人の社会的なつながりを活性化させる環境知能システムの効果を検証できたと考えられる。

第5章では、環境知能システムの実現可能性を示すために必要なユーザのアノテーション入力としてのインタフェースの検討を行った。これまで提案されてきたシステムでは、ユーザの情報収集手法として、主に人間のコミュニケーションメディアを利用した手法が用いられてきた。しかし従来の手法では、実世界の人間の社会的活動を阻害するといった

課題がある（他者との会話中に音声入力インタフェースを使用できない、両手が塞がった状態では手指を使った操作ができない等）。そこで本研究では、様々な環境下でもシームレスに入力可能なインタフェースを目指し、関節動作を伴わない筋肉の等尺性収縮に着目しインタフェースの構築を行なった。そして提案するインタフェースが実世界で起こり得る様々な状況下において、人間の主活動を阻害せずに入力可能（精度 93.13%）であることを示した。

これらの研究から、本研究で提案する人間の身体の内的変化に着目した社会的なつながりを活性化しうる環境知能システムの実現可能性を示した。本研究では、エンタテインメントに着目して研究を行ったが、社会的なつながりを活性化することは、人の健康に寄与する要素であり、若者から高齢者まで広く適用できる。さらに応用分野としては娯楽以外にも教育や仕事、医療など様々な分野で可能性があり、本研究で得られた知見はそれらの発展可能性を有す重要な貢献である。

参考文献

- [1] Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 2012. Putting Your Best Foot Forward: Investigating Real-World Mappings for Foot-Based Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1229–1238. DOI:<http://dx.doi.org/10.1145/2207676.2208575>
- [2] Amazon. 2020. Amazon Alexa Official Site: What Is Alexa? Retrived August 20, 2020 from <https://developer.amazon.com/en-US/alexa>. (2020).
- [3] Christoph Amma, Thomas Krings, Jonas Böer, and Tanja Schultz. 2015. Advancing Muscle-Computer Interfaces with High-Density Electromyography. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 929–938. DOI:<http://dx.doi.org/10.1145/2702123.2702501>
- [4] Toshiyuki Ando, Yuki Kubo, Buntarou Shizuki, and Shin Takahashi. 2018. CanalSense+: Face-Related Movement Recognition and Identification System Based on Air Pressure in Ear Canals. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1. DOI:<http://dx.doi.org/10.1145/3170427.3186600>
- [5] Apple. 2020. Siri - Apple. Retrived August 20, 2020 from <https://www.apple.com/siri/>. (2020).
- [6] Daniel Ashbrook, Carlos Tejada, Dhwanit Mehta, Anthony Jiminez, Goudam Muralitharam, Sangeeta Gajendra, and Ross Tallents. 2016. Bitey: An Exploration of Tooth Click Gestures for Hands-Free User Interface Control. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 158–169. DOI:<http://dx.doi.org/10.1145/2935334.2935389>
- [7] Thomas Augsten, Konstantin Kaefer, René Meusel, Caroline Fetzer, Dorian Kanitz, Thomas Stoff, Torsten Becker, Christian Holz, and Patrick Baudisch. 2010. Multitoe: High-Precision Interaction with Back-Projected Floors Based on High-Resolution Multi-Touch Input. In *Proceedings of the 23rd Annual ACM Symposium*

- on User Interface Software and Technology (UIST '10)*. Association for Computing Machinery, New York, NY, USA, 209–218. DOI:<http://dx.doi.org/10.1145/1866029.1866064>
- [8] Abdelkareem Bedri, David Byrd, Peter Presti, Himanshu Sahni, Zehua Gue, and Thad Starner. 2015. 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)*. Association for Computing Machinery, New York, NY, USA, 1333–1338. DOI:<http://dx.doi.org/10.1145/2800835.2807933>
- [9] Andreas Bulling, Daniel Roggen, and Gerhard Tröster. 2008. It's in Your Eyes: Towards Context-Awareness and Mobile HCI Using Wearable EOG Goggles. In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*. Association for Computing Machinery, New York, NY, USA, 84–93. DOI:<http://dx.doi.org/10.1145/1409635.1409647>
- [10] Walter Bradford Cannon. 1914. The emergency function of the adrenal medulla in pain and the major emotions. *American Journal of Physiology-Legacy Content* 33, 2 (1914), 356–372.
- [11] Houston Claire, Yifang Chen, Jignesh Modi, Malte Jung, and Stefanos Nikolaidis. 2020. *Multi-Armed Bandits with Fairness Constraints for Distributing Resources to Human Teammates*. Association for Computing Machinery, New York, NY, USA, 299–308. <https://doi.org/10.1145/3319502.3374806>
- [12] Andrew Crossan, Stephen Brewster, and Alexander Ng. 2010. Foot tapping for mobile interaction. *Proceedings of HCI 2010 24* (2010), 418–422.
- [13] P. Dalka and A. Czyzewski. 2010. Human-Computer Interface Based on Visual Lip Movement and Gesture Recognition. *Int. J. Comput. Sci. Appl.* 7 (2010), 124–139.
- [14] Artem Dementyev and Joseph A. Paradiso. 2014. WristFlex: Low-Power Gesture Input with Wrist-Worn Pressure Sensors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. Association for Computing Machinery, New York, NY, USA, 161–166. DOI:<http://dx.doi.org/10.1145/2642918.2647396>
- [15] B. Denby, Y. Oussar, G. Dreyfus, and M. Stone. 2006. Prospects for a Silent Speech Interface using Ultrasound Imaging. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, Vol. 1. I–I.
- [16] B. Denby, T. Schultz, K. Honda, T. Hueber, J.M. Gilbert, and J.S. Brumberg. 2010. Silent speech interfaces. *Speech Communication* 52, 4 (2010), 270 – 287. DOI:

- <http://dx.doi.org/https://doi.org/10.1016/j.specom.2009.08.002> Silent Speech Interfaces.
- [17] B. Denby and M. Stone. 2004. Speech synthesis from real time ultrasound images of the tongue. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 1. I–685.
- [18] Murtaza Dhuliawala, Juyoung Lee, Junichi Shimizu, Andreas Bulling, Kai Kunze, Thad Starner, and Woontack Woo. 2016. Smooth Eye Movement Interaction Using EOG Glasses. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI '16)*. Association for Computing Machinery, New York, NY, USA, 307–311. DOI:<http://dx.doi.org/10.1145/2993148.2993181>
- [19] Koumei Fukahori, Daisuke Sakamoto, and Takeo Igarashi. 2015. Exploring Subtle Foot Plantar-Based Gestures with Sock-Placed Pressure Sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3019–3028. DOI:<http://dx.doi.org/10.1145/2702123.2702308>
- [20] Masaaki Fukumoto. 2018. SilentVoice: Unnoticeable Voice Input by Ingressive Speech. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 237–246. DOI:<http://dx.doi.org/10.1145/3242587.3242603>
- [21] Pablo Gallego Cascón, Denys J.C. Matthies, Sachith Muthukumarana, and Suranga Nanayakkara. 2019. ChewIt. An Intraoral Interface for Discreet Interactions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. DOI:<http://dx.doi.org/10.1145/3290605.3300556>
- [22] Sarah Gillet, Ronald Cumbal, André Pereira, José Lopes, Olov Engwall, and Iolanda Leite. 2021. Robot Gaze Can Mediate Participation Imbalance in Groups with Different Skill Levels. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21)*. Association for Computing Machinery, New York, NY, USA, 303–311. DOI:<http://dx.doi.org/10.1145/3434073.3444670>
- [23] Fabian Göbel, Konstantin Klamka, Andreas Siegel, Stefan Vogt, Sophie Stellmach, and Raimund Dachselt. 2013. Gaze-Supported Foot Interaction in Zoomable Information Spaces. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 3059–3062. DOI:<http://dx.doi.org/10.1145/2468356.2479610>
- [24] Google. 2020. Google Assistant, your own personal Google. Retrived August 20, 2020 from <https://assistant.google.com/intl/en-US/>. (2020).

- [25] Sean Gustafson, Daniel Bierwirth, and Patrick Baudisch. 2010. Imaginary Interfaces: Spatial Interaction with Empty Hands and without Visual Feedback. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. Association for Computing Machinery, New York, NY, USA, 3–12. DOI:<http://dx.doi.org/10.1145/1866029.1866033>
- [26] Toshimitsu Hamada, Hiroki Okubo, Kazuya Inoue, Joji Maruyama, Hisashi Onari, Yoshihito Kagawa, and Tomomi Hashimoto. 2008. Robot therapy as for recreation for elderly people with dementia - Game recreation using a pet-type robot -. In *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*. 174–179. DOI:<http://dx.doi.org/10.1109/ROMAN.2008.4600662>
- [27] Teng Han, Jason Alexander, Abhijit Karnik, Pourang Irani, and Sriram Subramanian. 2011. Kick: Investigating the Use of Kick Gestures for Mobile Interactions. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. Association for Computing Machinery, New York, NY, USA, 29–32. DOI:<http://dx.doi.org/10.1145/2037373.2037379>
- [28] Chris Harrison, Hrvoje Benko, and Andrew D. Wilson. 2011. OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 441–450. DOI:<http://dx.doi.org/10.1145/2047196.2047255>
- [29] Chris Harrison, Shilpa Ramamurthy, and Scott E. Hudson. 2012. On-Body Interaction: Armed and Dangerous. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (TEI '12)*. Association for Computing Machinery, New York, NY, USA, 69–76. DOI:<http://dx.doi.org/10.1145/2148131.2148148>
- [30] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body as an Input Surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 453–462. DOI:<http://dx.doi.org/10.1145/1753326.1753394>
- [31] Teresa Hirzle, Jan Gugenheimer, Florian Geiselhart, Andreas Bulling, and Enrico Rukzio. 2019. A Design Space for Gaze Interaction on Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. DOI:<http://dx.doi.org/10.1145/3290605.3300855>
- [32] Donny Huang, Xiaoyi Zhang, T. Scott Saponas, James Fogarty, and Shyamnath Gollakota. 2015. Leveraging Dual-Observable Input for Fine-Grained Thumb Inter-

- action Using Forearm EMG. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15)*. Association for Computing Machinery, New York, NY, USA, 523–528. DOI:<http://dx.doi.org/10.1145/2807442.2807506>
- [33] Scott E. Hudson, Chris Harrison, Beverly L. Harrison, and Anthony LaMarca. 2010. Whack Gestures: Inexact and Inattentive Interaction with Mobile Devices. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '10)*. Association for Computing Machinery, New York, NY, USA, 109–112. DOI:<http://dx.doi.org/10.1145/1709886.1709906>
- [34] Thomas Hueber, Elie-Laurent Benaroya, Gérard Chollet, Bruce Denby, Gérard Dreyfus, and Maureen Stone. 2010. Development of a silent speech interface driven by ultrasound and optical images of the tongue and lips. *Speech Communication* 52, 4 (2010), 288 – 300. DOI:<http://dx.doi.org/https://doi.org/10.1016/j.specom.2009.11.004> Silent Speech Interfaces.
- [35] Shiroh Itai, Atsushi Shimoda, Toshihiko Yoneoka, and Toshimitsu Hamada. 2020. Development of Scenario-Type Robot Recreation Program for the Elderly with Dementia and Its Evaluation. In *2020 5th International Conference on Universal Village (UV)*. 1–7. DOI:<http://dx.doi.org/10.1109/UV50937.2020.9426218>
- [36] Hassaluk Jirathampradub, Chakarida Nukoolkit, Kaneungjit Suriyathumrongkul, and Bunthit Watanapa. 2020. A 3D-CNN Siamese Network for Motion Gesture Sign Language Alphabets Recognition. In *Proceedings of the 11th International Conference on Advances in Information Technology (IAIT2020)*. Association for Computing Machinery, New York, NY, USA, Article 29, 6 pages. DOI:<http://dx.doi.org/10.1145/3406601.3406634>
- [37] Ricardo Jota and Daniel Wigdor. 2015. Palpebrae superioris: exploring the design space of eyelid gestures. In *Proceedings of Graphics Interface 2015 (GI 2015)*. Canadian Human-Computer Communications Society, Toronto, Ontario, Canada, 273–280. DOI:<http://dx.doi.org/10.20380/GI2015.35>
- [38] Malte F. Jung, Nikolas Martelaro, and Pamela J. Hinds. 2015. Using Robots to Moderate Team Conflict: The Case of Repairing Violations. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. Association for Computing Machinery, New York, NY, USA, 229–236. DOI:<http://dx.doi.org/10.1145/2696454.2696460>
- [39] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-Based Interaction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. Association

- for Computing Machinery, New York, NY, USA, 1151–1160. DOI:<http://dx.doi.org/10.1145/2638728.2641695>
- [40] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. Digits: Freehand 3D Interactions Anywhere Using a Wrist-Worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. Association for Computing Machinery, New York, NY, USA, 167–176. DOI:<http://dx.doi.org/10.1145/2380116.2380139>
- [41] Koichi Kuzume. 2010. Input Device for Disabled Persons Using Expiration and Tooth-Touch Sound Signals. In *Proceedings of the 2010 ACM Symposium on Applied Computing (SAC '10)*. Association for Computing Machinery, New York, NY, USA, 1159–1164. DOI:<http://dx.doi.org/10.1145/1774088.1774335>
- [42] Hideaki Kuzuoka, Yuya Suzuki, Jun Yamashita, and Keiichi Yamazaki. 2010. Re-configuring spatial formation arrangement by robot body orientation. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 285–292. DOI:<http://dx.doi.org/10.1109/HRI.2010.5453182>
- [43] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billingham. 2018. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. DOI:<http://dx.doi.org/10.1145/3173574.3173655>
- [44] Joseph J. LaViola, Daniel Acevedo Feliz, Daniel F. Keefe, and Robert C. Zeleznik. 2001. Hands-Free Multi-Scale Navigation in Virtual Environments. In *Proceedings of the 2001 Symposium on Interactive 3D Graphics (I3D '01)*. Association for Computing Machinery, New York, NY, USA, 9–15. DOI:<http://dx.doi.org/10.1145/364338.364339>
- [45] Zheng Li, Ryan Robucci, Nilanjan Banerjee, and Chintan Patel. 2015. Tongue-n-Cheek: Non-Contact Tongue Gesture Recognition. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks (IPSN '15)*. Association for Computing Machinery, New York, NY, USA, 95–105. DOI:<http://dx.doi.org/10.1145/2737095.2737109>
- [46] Zhiyuan Lu, Xiang Chen, Zhangyan Zhao, and Kongqiao Wang. 2011. A Prototype of Gesture-Based Interface. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. Association for Computing Machinery, New York, NY, USA, 33–36. DOI:<http://dx.doi.org/10.1145/2037373.2037380>
- [47] Michael J. Lyons, Michael Haehnel, and Nobuji Tetsutani. 2003. Designing, Playing, and Performing with a Vision-Based Mouth Interface. In *Proceedings of the*

2003 Conference on New Interfaces for Musical Expression (NIME '03). National University of Singapore, SGP, 116–121.

- [48] Hiroyuki Manabe, Masaaki Fukumoto, and Tohru Yagi. 2015. Conductive Rubber Electrodes for Earphone-Based Eye Gesture Input Interface. *Personal Ubiquitous Comput.* 19, 1 (Jan. 2015), 143–154. DOI:<http://dx.doi.org/10.1007/s00779-014-0818-8>
- [49] Katsutoshi Masai, Kai Kunze, and Maki Sugimoto. 2020. Eye-Based Interaction Using Embedded Optical Sensors on an Eyewear Device for Facial Expression Recognition. In *Proceedings of the Augmented Humans International Conference (AHs '20)*. Association for Computing Machinery, New York, NY, USA, Article 1, 10 pages. DOI:<http://dx.doi.org/10.1145/3384657.3384787>
- [50] Denys J. C. Matthies, Simon T. Perrault, Bodo Urban, and Shengdong Zhao. 2015. Botential: Localizing On-Body Gestures by Measuring Electrical Signatures on the Human Skin. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 207–216. DOI:<http://dx.doi.org/10.1145/2785830.2785859>
- [51] Denys J. C. Matthies, Bernhard A. Strecker, and Bodo Urban. 2017. EarField-Sensing: 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)*. Association for Computing Machinery, New York, NY, USA, 1911–1922. DOI:<http://dx.doi.org/10.1145/3025453.3025692>
- [52] Fabrice Matulic, Brian Vogel, Naoki Kimura, and Daniel Vogel. 2019. Eliciting Pen-Holding Postures for General Input with Suitability for EMG Armband Detection. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces (ISS '19)*. Association for Computing Machinery, New York, NY, USA, 89–100. DOI:<http://dx.doi.org/10.1145/3343055.3359720>
- [53] Jess McIntosh, Asier Marzo, Mike Fraser, and Carol Phillips. 2017. EchoFlex: Hand Gesture Recognition Using Ultrasound Imaging. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1923–1934. DOI:<http://dx.doi.org/10.1145/3025453.3025807>
- [54] Jess McIntosh, Charlie McNeill, Mike Fraser, Frederic Kerber, Markus Löchtefeld, and Antonio Krüger. 2016. EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Com-

- puting Machinery, New York, NY, USA, 2332–2342. DOI:<http://dx.doi.org/10.1145/2858036>.2858093
- [55] Microsoft. 2020. Cortana - Your personal productivity assistant. Retrived August 20, 2020 from <https://www.microsoft.com/en-us/cortana>. (2020).
- [56] Pranav Mistry and Pattie Maes. 2009. SixthSense: A Wearable Gestural Interface. In *ACM SIGGRAPH ASIA 2009 Sketches (SIGGRAPH ASIA '09)*. Association for Computing Machinery, New York, NY, USA, Article 11, 1 pages. DOI:<http://dx.doi.org/10.1145/1667146>.1667160
- [57] Kazuki Mizumaru, Satoru Satake, Takayuki Kanda, and Tetsuo Ono. 2019. Stop Doing It! Approaching Strategy for a Robot to Admonish Pedestrians. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI '19)*. IEEE Press, 449–457.
- [58] Tamer Mohamed and Lin Zhong. 2006. *Teethclick: Input with teeth clacks*. Technical Report. Technical Report. Rice University.
- [59] Aarthi Easwara Moorthy and Kim-Phuong L. Vu. 2015. Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. *International Journal of Human-Computer Interaction* 31, 4 (2015), 307–335. DOI:<http://dx.doi.org/10.1080/10447318.2014.986642>
- [60] I. Mougharbel, R. El-Hajj, H. Ghamlouch, and E. Monacelli. 2013. Comparative study on different adaptation approaches concerning a sip and puff controller for a powered wheelchair. In *2013 Science and Information Conference*. 597–603.
- [61] Florian Müller, Martin Schmitz, Daniel Schmitt, Sebastian Günther, Markus Funk, and Max Mühlhäuser. 2020. Walk The Line: Leveraging Lateral Shifts of the Walking Path as an Input Modality for Head-Mounted Displays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. DOI:<http://dx.doi.org/10.1145/3313831>.3376852
- [62] N. Nakamura, M. Inaba, K. Takahashi, F. Toriumi, H. Osawa, D. Katagami, and K. Shinoda. 2016. Constructing a Human-like agent for the Werewolf Game using a psychological model based multiple perspectives. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. 1–8.
- [63] Tomoko Nariai, Shiroh Itai, and Hiroaki Kojima. 2021a. Evaluating Effectiveness of Robot-Assisted Recreation for Older Adults by Speech Analysis. In *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*. 240–243. DOI:<http://dx.doi.org/10.1109/LifeTech52111.2021.9391826>

- [64] Tomoko Nariai, Shiroh Itai, and Hiroaki Kojima. 2021b. Speech Analysis to Evaluate Robot-Assisted Recreation of Older Adults with Dementia. *Journal of Robotics and Mechatronics* 33, 4 (2021), 777–783. DOI:<http://dx.doi.org/10.20965/jrm.2021.p0777>
- [65] Volker Paelke, Christian Reimann, and Dirk Stichling. 2004. Kick-up Menus. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1552. DOI:<http://dx.doi.org/10.1145/985921.986128>
- [66] Toni Pakkanen and Roope Raisamo. 2004. Appropriateness of Foot Interaction for Non-Accurate Spatial Tasks. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1123–1126. DOI:<http://dx.doi.org/10.1145/985921.986004>
- [67] G. Pearson and M. Weiser. 1988. Exploratory Evaluation of a Planar Foot-Operated Cursor-Positioning Device. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '88)*. Association for Computing Machinery, New York, NY, USA, 13–18. DOI:<http://dx.doi.org/10.1145/57167.57169>
- [68] J. K. Perng, B. Fisher, S. Hollar, and K. S. J. Pister. 1999. Acceleration sensing glove (ASG). In *Digest of Papers. Third International Symposium on Wearable Computers*. 178–180.
- [69] T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst. 2017. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 36–39.
- [70] J. Rekimoto. 2001. GestureWrist and GesturePad: unobtrusive wearable interaction devices. In *Proceedings Fifth International Symposium on Wearable Computers*. 21–27.
- [71] Microsoft Research. 2015. Hands-Free Keyboard - Microsoft Research. Retrieved August 21, 2020 from <https://www.microsoft.com/en-us/research/project/hands-free-keyboard/>. (2015).
- [72] Softbank Robotics. 2022. NAO the humanoid and programmable robot — Soft-Bank Robotics. Retrived June 30, 2022 from <https://www.softbankrobotics.com/emea/en/nao>. (2022).
- [73] Rinat B. Rosenberg-Kima, Yaacov Koren, and Goren Gordon. 2020. Robot-Supported Collaborative Learning (RSCL): Social Robots as Teaching Assistants for Higher Education Small Group Facilitation. *Frontiers in Robotics and AI* 6 (2020). DOI:<http://dx.doi.org/10.3389/frobt.2019.00148>

- [74] Jaime Ruiz and Yang Li. 2011. DoubleFlip: A Motion Gesture Delimiter for Mobile Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2717–2720. DOI:<http://dx.doi.org/10.1145/1978942.1979341>
- [75] Daisuke Sakamoto and Tetsuo Ono. 2006. Sociality of Robots: Do Robots Construct or Collapse Human Relations?. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI '06)*. Association for Computing Machinery, New York, NY, USA, 355–356. DOI:<http://dx.doi.org/10.1145/1121241.1121313>
- [76] H. Sakoe and S. Chiba. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26, 1 (1978), 43–49. DOI:<http://dx.doi.org/10.1109/TASSP.1978.1163055>
- [77] T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. 2009. Enabling Always-Available Input with Muscle-Computer Interfaces. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)*. Association for Computing Machinery, New York, NY, USA, 167–176. DOI:<http://dx.doi.org/10.1145/1622176.1622208>
- [78] Satoru Satake, Takayuki Kanda, Dylan F. Glas, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. How to Approach Humans? Strategies for Social Robots to Initiate Interaction. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI '09)*. Association for Computing Machinery, New York, NY, USA, 109–116. DOI:<http://dx.doi.org/10.1145/1514095.1514117>
- [79] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 483–492. DOI:<http://dx.doi.org/10.1145/2207676.2207743>
- [80] G. Schlesinger. 1919. *Der mechanische Aufbau der künstlichen Glieder*. Springer Berlin Heidelberg, Berlin, Heidelberg, 321–661. DOI:http://dx.doi.org/10.1007/978-3-662-33009-8_13
- [81] Johannes Schöning, Florian Daiber, Antonio Krüger, and Michael Rohs. 2009. Using Hands and Feet to Navigate and Manipulate Spatial Data. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. Association for Computing Machinery, New York, NY, USA, 4663–4668. DOI:<http://dx.doi.org/10.1145/1520340.1520717>
- [82] Jeremy Scott, David Dearman, Koji Yatani, and Khai N. Truong. 2010. Sensing Foot Gestures from the Pocket. In *Proceedings of the 23rd Annual ACM Symposium*

- on User Interface Software and Technology (UIST '10)*. Association for Computing Machinery, New York, NY, USA, 199–208. DOI:<http://dx.doi.org/10.1145/1866029.1866063>
- [83] Ameneh Shamekhi and Timothy Bickmore. 2019. A Multimodal Robot-Driven Meeting Facilitation System for Group Decision-Making Sessions. In *2019 International Conference on Multimodal Interaction (ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 279–290. DOI:<http://dx.doi.org/10.1145/3340555.3353756>
- [84] Solace Shen, Petr Slovak, and Malte F. Jung. 2018. "Stop. I See a Conflict Happening.": A Robot Mediator for Young Children's Interpersonal Conflict Resolution. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. Association for Computing Machinery, New York, NY, USA, 69–77. DOI:<http://dx.doi.org/10.1145/3171221.3171248>
- [85] A. L. Simeone, E. Velloso, J. Alexander, and H. Gellersen. 2014. Feet movement in desktop 3D interaction. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 71–74.
- [86] Ronit Slyper, Jill Lehman, Jodi Forlizzi, and Jessica Hodgins. 2011. A Tongue Input Device for Creating Conversations. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 117–126. DOI:<http://dx.doi.org/10.1145/2047196.2047210>
- [87] Vasant Srinivasan and Leila Takayama. 2016. Help Me Please: Robot Politeness Strategies for Soliciting Help From Humans. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4945–4955. DOI:<http://dx.doi.org/10.1145/2858036.2858217>
- [88] Ke Sun, Chun Yu, Weinan Shi, Lan Liu, and Yuanchun Shi. 2018. Lip-Interact: Improving Mobile Device Interaction with Silent Speech Commands. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 581–593. DOI:<http://dx.doi.org/10.1145/3242587.3242599>
- [89] Veikko Surakka, Marko Illi, and Poika Isokoski. 2004. Gazing and Frowning as a New Human–Computer Interaction Technique. *ACM Trans. Appl. Percept.* 1, 1 (July 2004), 40–56. DOI:<http://dx.doi.org/10.1145/1008722.1008726>
- [90] Tobii. 2020. Assistive technology for communication/AAC - Tobii Dynavox. Retrieved August 21, 2020 from <https://www.tobiidynavox.com/>. (2020).

- [91] Marc Tonsen, Julian Steil, Yusuke Sugano, and Andreas Bulling. 2017. InvisibleEye: Mobile Eye Tracking Using Multiple Low-Resolution Cameras and Learning-Based Gaze Estimation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1 (09 2017), 1–21. DOI:<http://dx.doi.org/10.1145/3130971>
- [92] Vytautas Vaitukaitis and Andreas Bulling. 2012. Eye Gesture Recognition on Portable Devices. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 711–714. DOI:<http://dx.doi.org/10.1145/2370216.2370370>
- [93] Oleg Špakov and Päivi Majaranta. 2012. Enhanced Gaze Interaction Using Simple Head Gestures. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 705–710. DOI:<http://dx.doi.org/10.1145/2370216.2370369>
- [94] Michael Wand, Christopher Schulte, Matthias Janke, and Tanja Schultz. 2013. Array-based Electromyographic Silent Speech Interface. In *BIOSIGNALS*.
- [95] David M. Watt. 1966a. Clinical applications of gnathosonics. *The Journal of Prosthetic Dentistry* 16, 1 (1966), 83 – 95. DOI:[http://dx.doi.org/https://doi.org/10.1016/0022-3913\(66\)90115-6](http://dx.doi.org/https://doi.org/10.1016/0022-3913(66)90115-6)
- [96] David M. Watt. 1966b. Gnathosonics—A study of sounds produced by the masticatory mechanism. *The Journal of Prosthetic Dentistry* 16, 1 (1966), 73 – 82. DOI:[http://dx.doi.org/https://doi.org/10.1016/0022-3913\(66\)90114-4](http://dx.doi.org/https://doi.org/10.1016/0022-3913(66)90114-4)
- [97] Robert Xiao, Teng Cao, Ning Guo, Jun Zhuo, Yang Zhang, and Chris Harrison. 2018. *LumiWatch: On-Arm Projected Graphics and Touch Input*. Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3173574.3173669>
- [98] Tetsuya Yamamoto, Tsutomu Terada, and Masahiko Tsukamoto. 2011. Designing Gestures for Hands and Feet in Daily Life. In *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia (MoMM '11)*. Association for Computing Machinery, New York, NY, USA, 285–288. DOI:<http://dx.doi.org/10.1145/2095697.2095757>
- [99] Qiao Zhang, Shyamnath Gollakota, Ben Taskar, and Raj P.N. Rao. 2014. Non-Intrusive Tongue Machine Interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2555–2558. DOI:<http://dx.doi.org/10.1145/2556288.2556981>

- [100] Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017. Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2878–2889. DOI: <http://dx.doi.org/10.1145/3025453.3025790>
- [101] Yang Zhang, Junhan Zhou, Gierad Laput, and Chris Harrison. 2016. SkinTrack: Using the Body as an Electrical Waveguide for Continuous Finger Tracking on the Skin. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1491–1503. DOI: <http://dx.doi.org/10.1145/2858036.2858082>
- [102] 伊藤毅志 and 松原仁. 2004. 将棋熟達者の発話にみる思考と認知. Technical Report 64(2004-GI-012). 電気通信大学情報工学科, はこだて未来大学システム情報科学部.
- [103] 伊藤毅志, 松原仁, and ライエルグリーンベルゲン. 2004. 将棋の認知科学的研究 (2) – 次の一手実験からの考察. *情報処理学会論文誌* 45, 5 (may 2004), 1481–1492.
- [104] 伊藤毅志 and 杉本磨美. 2020. (OS 招待講演) 人狼プレイヤーの意思決定過程. *人工知能学会全国大会論文集 2020* (2020), 2F4OS20a01–2F4OS20a01. DOI: http://dx.doi.org/10.11517/pjsai.JSAI2020.0_2F4OS20a01
- [105] 稲葉通将, 鳥海不二夫, 大澤博隆, 片上大輔, 篠田孝祐, and 西野順二. 2014. 同調と反駁に着目した人狼ゲームの分析. *人工知能学会全国大会論文集 2014* (2014), 1E4OS23a1–1E4OS23a1. DOI: http://dx.doi.org/10.11517/pjsai.JSAI2014.0_1E4OS23a1
- [106] 王天鶴 and 金子知適. 2018. 人狼エージェントにおける深層Qネットワークの応用. *ゲームプログラミングワークショップ 2018 論文集 2018* (nov 2018), 16–22. <https://ci.nii.ac.jp/naid/170000178462/>
- [107] 岡本秀明. 2012. 都市部在住高齢者の社会活動に関連する要因の検討: 地域におけるつながりづくりと社会的孤立の予防に向けて. *社会福祉学* 53, 3 (2012), 3–17.
- [108] 梶原健吾, 鳥海不二夫, and 稲葉通将. 2015. 人狼における強化学習を用いたエージェントの設計. *人工知能学会全国大会論文集 2015* (2015), 1F22–1F22. DOI: http://dx.doi.org/10.11517/pjsai.JSAI2015.0_1F22
- [109] 梶原健吾, 鳥海不二夫, 稲葉通将, 大澤博隆, 片上大輔, 篠田孝祐, 松原仁, and 狩野芳伸. 2016. 人狼知能大会における統計分析とSVMを用いた人狼推定を行うエージェントの設計. *人工知能学会全国大会論文集 2016* (2016), 2F41–2F41. DOI: http://dx.doi.org/10.11517/pjsai.JSAI2016.0_2F41
- [110] 梶浦久江 and 中山伸一. 2010. ブロック崩しゲームにおけるプレイヤーとゲームを見る人のフロー体験に与える音楽の影響. *デジタルゲーム学研究* 4, 2 (2010), 13–22. DOI: http://dx.doi.org/10.9762/digraj.4.2_13

- [111] 御手洗彰, 水丸和樹, 本田健吾, 棟方渚, 坂本大介, and 小野哲雄. 2018. 人狼プレイヤーの皮膚電気活動の解析: 情動変化を利用したソシオメータの実現へ向けて. *インタラクシオン 2018* (feb 2018), 885–888.
- [112] 高山周太郎 and 大澤博隆. 2019. 実世界人狼ゲームのジェスチャーの分析. *人工知能学会全国大会論文集 2019* (2019), 3F3OS14a01–3F3OS14a01. DOI:http://dx.doi.org/10.11517/pjsai.JSAI2019.0_3F3OS14a01
- [113] 沼田恵太郎, 宮田洋, and others. 2011. 皮膚電気条件づけ: その意義と研究動向. *人文論究* 61, 2 (2011), 55–88.
- [114] 清水大輔 and 長谷部浩二. 2020. プレイヤーの発言内容に関するルールに基づいた人狼ゲームの役職推定. *人工知能学会全国大会論文集 2020* (2020), 2F5OS20b05–2F5OS20b05. DOI:http://dx.doi.org/10.11517/pjsai.JSAI2020.0_2F5OS20b05
- [115] 総務省. 2017. 総務省 ICT スキル総合習得プログラム - 総務省 ICT スキル総合習得教材. (2017). https://www.soumu.go.jp/ict_skill/pdf/ict_skill_1_1.pdf.
- [116] 大川貴聖, 吉仲亮, and 篠原歩. 2017. 深層学習を用いて役職推定を行う人狼知能エージェントの開発. *ゲームプログラミングワークショップ 2017論文集 2017* (nov 2017), 50–55. <https://ci.nii.ac.jp/naid/170000176048/>
- [117] 丹野宏昭 and 児玉健. 2015. 人狼ゲーム経験による嘘の手がかり信念の差異: 大学生, 人狼ゲーム愛好者, 人狼ゲーム舞台役者の比較. *パーソナリティ研究* 24, 1 (2015), 88–90. DOI:<http://dx.doi.org/10.2132/personality.24.88>
- [118] 長尾由希子. 2020. 地方の若者の幸せと都市の若者の幸せ——出身県による幸福度の構造と高等教育の解釈. *日本教育学会大会研究発表要項* 79 (2020), 214–215.
- [119] 鳥海不二夫, 片上大輔, 大澤博隆, 稲葉通将, 篠田孝祐, and 狩野芳伸. 2016. *人狼知能: だます・見破る・説得する人工知能*. 森北出版. <https://ci.nii.ac.jp/ncid/BB21844125>
- [120] 風間祥光, 畑雅之, and 松原仁. 2018. *人狼ゲーム内の議論での行動がプレイヤーに与える影響*. Technical Report 16. 公立はこだて未来大学, 株式会社 HiSC, 公立はこだて未来大学.
- [121] 福田宗理 and 穴田一. 2020. 15 人狼ゲームにおける会話情報による役職推定. *人工知能学会全国大会論文集 2020* (2020), 2F5OS20b01–2F5OS20b01. DOI:http://dx.doi.org/10.11517/pjsai.JSAI2020.0_2F5OS20b01
- [122] 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, and 松原仁. 2015. 人狼知能プロジェクト (<特集> エンターテイメントにおける AI). *人工知能 = journal of the Japanese Society for Artificial Intelligence* 30, 1 (jan 2015), 65–73. <https://ci.nii.ac.jp/naid/110009884678/>

- [123] 三宅 陽一郎. 2018. デジタルゲームにおける可視化技術. *可視化情報学会誌* 38, 151 (2018), 28-33. DOI:http://dx.doi.org/10.3154/jvs.38.151_28

発表論文

- [1] Designing Hand Gesture Sequence Recognition Technique for Input While Grasping an Object, Sho Mitarai, Nagisa Munekata, Daisuke Sakamoto, Tetsuo Ono, Transactions of the Virtual Reality Society of Japan, 26(4), pages 333-344, December 2021.
- [2] 物を把持した状態における筋電センサを用いたハンドジェスチャ入力の問題抽出と新手法の提案, 御手洗 彰, 棟方 渚, 小野 哲雄, *日本バーチャルリアリティ学会論文誌*, 22(1), pages 41-50, March 2017.
- [3] 皮膚電気活動を指標とした人狼プレイヤーのゲーム体験に基づくハイライト抽出の試み, 御手洗 彰, 棟方 渚, *人工知能学会全国大会論文集 第 36 回全国大会 (2022)*, June 2021.
- [4] 音楽聴取中における振動提示手法の差異が音楽体験の向上に及ぼす影響の実験的調査, 増田 智也, 御手洗 彰, 棟方 渚, *研究報告エンタテインメントコンピューティング (EC)*, June 2022.
- [5] ロボットを用いたレクリエーション援助がプレイヤーのモチベーションに与える影響の調査, 御手洗 彰, 棟方 渚, *研究報告エンタテインメントコンピューティング (EC)*, June 2022.
- [6] スタンプラリー型イベントの参加者の回遊に関するモチベーションの実験的調査, 坂本 唯斗, 御手洗 彰, 棟方 渚, *研究報告エンタテインメントコンピューティング (EC)*, March 2022.
- [7] ユーザによる外観制作に基づくロボット外観特徴分類及びユーザ嗜好抽出の試み, 加藤 雄大, 御手洗 彰, 棟方 渚, *Human-Agent Interaction シンポジウム 2022*, March 2022.
- [8] 筋肉の非活動部位を活用した外科手術用の筋電入力インタフェースの検討, 御手洗 彰, 棟方 渚, *研究報告ヒューマンコンピュータインタラクション (HCI)*, January 2022.
- [9] 筋肉の非活動部位を入力モダリティとして用いたハンドジェスチャ認識手法の提案, 御手洗 彰, 棟方 渚, *研究報告ヒューマンコンピュータインタラクション (HCI)*, November 2021.
- [10] プレイヤーのゲーム体験分析による 5 人狼の特性抽出, 御手洗 彰, 山本 浩隆, 棟方 渚, *研究報告ヒューマンコンピュータインタラクション (HCI)*, August 2021.
- [11] ロボットの外観制作による親近性要因の抽出, 加藤 雄大, 御手洗 彰, 棟方 渚, *研究報告ヒューマンコンピュータインタラクション (HCI)*, August 2021.
- [12] 特殊詐欺の実音声を用いた被疑者の発話情報の調査と分析, 辻 健一郎, 御手洗 彰, 棟方 渚, *人工知能学会全国大会論文集 第 35 回全国大会 (2021)*, June 2021.

- [13] ロボットの外観制作によるユーザ嗜好の抽出と分類, 加藤 雄大, 御手洗 彰, 棟方 渚, Human-Agent Interaction シンポジウム 2021, March 2021.
- [14] 覚醒水準による人狼プレイヤーの特徴分析の試み, 山本 浩隆, 御手洗 彰, 棟方 渚, 人工知能学会全国大会論文集 第 34 回全国大会 (2020), pages 2F4OS20a02-2F4OS20a02, June 2020.
- [15] 人狼プレイヤーの皮膚電気活動の解析: 情動変化を利用したソシオメータの実現へ向けて, 御手洗 彰, 水丸 和樹, 本田 健悟, 棟方 渚, 坂本 大介, 小野 哲雄, 情報処理学会 インタラクション 2018, pages 885-888, March 2018.
- [16] Cami-log: 咀嚼の改善を促す筋電情報を用いたアプリケーションの提案, 黒澤 紘生, 御手洗 彰, 棟方 渚, 小野 哲雄, 情報処理学会 インタラクション 2017, pages 907-912, March 2017.
- [17] 前腕の表面筋電を用いたジェスチャ認識の実験的検討, 御手洗 彰, 棟方 渚, 吉田 彩乃, 櫻沢 繁, 小野 哲雄, 研究報告ヒューマンコンピュータインタラククション (HCI), 2016.13, pages 1-6, June 2016.

目次

3.1	実験環境	14
3.2	計測システムと電極装着位置	15
3.3	会議フェーズ（初日）から投票フェーズ（初日）の時系列データの平均値（縦軸：正規化された SCL 値，横軸：時間（秒））	17
3.4	ゲーム全体の各場面における SCL の平均値（縦軸：正規化された SCL 値，横軸：ゲーム場面，エラーバーは標準誤差を表す，CO はカミングアウトを表す）	18
3.5	勝率で分けた群における CO イベントと投票フェーズの SCL 値の比較（縦軸：正規化された SCL 値，横軸：ゲーム場面，エラーバーは標準誤差を表す，CO はカミングアウトを表す）	19
3.6	勝率で分けた群における発言数の比較（縦軸：発言数，横軸：勝率で分けた群，エラーバーは標準誤差を表す）	20
3.7	類似性の高い SCL 値 ($d = 0.94$) の挙動	22
3.8	類似性の低い SCL 値 ($d = 2.04$) の挙動	23
3.9	各群における DTW 距離の平均値の比較	24
3.10	強い人狼役職プレイヤーの SCL 値の挙動	25
3.11	弱い人狼役職プレイヤーの SCL 値の挙動	26
3.12	強いプレイヤーと弱いプレイヤーの覚醒水準の比較	27
3.13	抽出されたハイライト場面の集計（縦軸：全ゲームで全プレイヤーから抽出されたハイライトの合計値，横軸：ゲーム中の各場面）	29
4.1	Robot mediation in recreation	32
4.2	Appearance of the game "Daruma-san ga Koronda"	33
4.3	Player's view while playing game	34
4.4	Operations of the game	35
4.5	実験概要図（図左：ゲームプレイ役の実験参加者，図右：妨害役の実験参加者）	36
4.6	The results of questionnaire (Q1 - Q8)	37
4.7	Gameplay time for each play	38
4.8	The result of questionnaire (Q9)	39
5.1	データ収集実験の概要	43
5.2	実験で用いるハンドジェスチャ	44
5.3	ジェスチャ記録システム	45
5.4	条件 A, B における学習・テストデータ	47

5.5	条件 A, B における各状況の認識精度	48
5.6	Comparison of the EMG for a gesture performed while busy and the typical hands-busy situation, demonstrating the extra capacity for input between the two. We thus use this capacity as an “extra modality.” . . .	50
5.7	Interaction sequence for hand gesture sequence recognition. (1) The user performs three hand gestures as input. (2) The differences in each of the two pairs (the differences between the first and second gestures and between the second and third gestures) are calculated as Euclidean distances. The two distances are used as the classification features for machine learning. (3) The k-NN algorithm classifies the three-gesture sequence. The data points were acquired by the user to perform hand gesture sequences in the hands-free situation. (4) Finally, according to the classification result, an application-specific operation such as playing music is performed.	53
5.8	Placement of a the Myo device.	55
5.9	Hand gesture set for data collection.	56
5.10	Grasp types, grasped objects, and gesture set considered in the user study. The grasp type graphs are classic taxonomies [80]. The object in the row denotes the object grasped by the participants. Gestures A and B in the row were the dorsal extension and palmar flexion gestures, respectively, which were determined in the gesture selection phase of this study.	59
5.11	Confusion matrix of gesture recognition for the four participants who showed low accuracy while grasping a box.	62
5.12	実験環境の模式図	64
5.13	評価実験の結果	65

表 目 次

2.1	Summary of input modalities	10
3.1	各プレイヤーの試合数と勝率	16
3.2	会議フェーズ中の議論内容	16
3.3	プレイヤーの SCL 平均値から抽出したハイライト場面	28
3.4	各プレイヤーにおけるハイライト場面の集計表	30
4.1	アンケートの設問内容	41
5.1	収集した EMG の例	46
5.2	条件 A, B における各状況の認識精度と状況における平均認識精度	48
5.3	Three-gesture sequences used in this study. A and B in the table represent different hand gestures.	51
5.4	The equation and resulting number of hand gesture sequence combinations.	57
5.5	Mean recognition accuracy of hand gesture pairs.	58
5.6	Mean gesture recognition accuracy per participant and grasp condition (P in the column denotes the participant).	61
5.7	Mean input time and standard deviation (S.D.) for hand gesture sequences for each grasp condition	61
5.8	ジェスチャ組×実験条件の認識率の結果 (%)	65