

# 視線とジェスチャをモダリティに持つ オーケストラシステムの開発

久保田大和<sup>†1</sup> ピトヨ ハルトノ<sup>†2</sup>

**概要:** 先行研究で、人間のノンバーバルなコミュニケーションを人間と機械の間に拡張したオーケストラシステムを構築した。提案したシステムとボタン、キーボード、ジョイコンといった日常的に扱う機器を介させたインターフェースを用いて、複数の機械に対して操作実験を行った結果、提案システムが有意であることが分かった。しかし、実験回数が限られていた点や両システムのスコア差に対する統計的有意差の検証が不十分である点で課題が残った。そこで本研究では、先行研究で残された課題を基に2台のロボットアームを用いた逐次的な物体搬送タスクを両システムで実施した。そして、その結果に対して統計的仮説検定を用いた有意差の検証を行った。本稿では、両システムによるユーザビリティ評価実験の結果と本システムが有する直感性について、改めて評価した結果について報告する。

## 1. はじめに

近代社会では人間は様々な機械を用いて、仕事または生活をする。パーソナルコンピュータや携帯電話といった電子機器が一般家庭に普及したことで、人間が接する機械の種類と用途がさらに増えた。技術の進展に伴い計算能力や通信速度、作業効率は飛躍的に向上した一方、人間が機械を操作するためのインターフェースは依然として人間が機械の仕様や制約に合わせて設計されたものが多い。このように機械の種類と用途の増加に比べてインターフェースの発展は限定的であり、直感的な操作という点で課題が残る。

著者は先行研究において、視線およびジェスチャを入力手段とするインターフェースに着目し、オーケストラの指揮者と演奏者の関係を模倣したオーケストラシステム[1]を構築した。このシステムを用いて5種類の機械に対する操作実験を行い、日常的に輸入手段として扱う機器を混在させたシステムに対して、本システムが優位な結果を示した。しかし、この研究においては実験回数が限られていたこと、実験による両システムで得たスコア差の統計的優位性の検証に課題が残されていた。

以上の課題に対して、本研究では、新たにロボットを用いて対象物を運搬するタスクを設定し、本システムとリモコン型インターフェースとのユーザビリティ評価を行った。これにより、本システムが有する直感性を改めて評価し、その有効性を検証した結果について報告する。

## 2. 関連研究

身体動作に基づいたインターフェースの研究は多数報告されている。

例えば、Tokmurziyev ら[2]は、視線入力と深層学習による

物体認識を組み合わせ、ウェアラブル視線インターフェースによって協働ロボットアームをハンズフリーで操作する「GazeGrasp」を提案した。本システムは、ユーザの視線方向とカメラ画像上の物体位置、およびロボット作業空間との対応関係を補正するキャリブレーションと物体検出を統合し、視線カーソルを対象物中心へ吸着させる磁気スナップ効果を導入することで、視線選択の精度と効率を向上した。評価実験では、視線調整時間の短縮やタスク効率の向上を示し、視線のみで実機ロボットを直感的に操作できる可能性を確認した。

Taylor ら[3]は、視線入力とマイクロジェスチャを統合した新たなインタラクション手法を提案した。ここでは、視線による注視対象の選択に対して、小規模な指先ジェスチャを組み合わせる二つの戦略（静的割り当てと動的コンテキスト割り当て）を整理し、それぞれの利点と課題を検討した。提案手法は、視線の高速な対象選択能力とマイクロジェスチャの明確な操作意図を組み合わせることで、効率的かつ柔軟な操作が可能であることを示し、視線ベースUIの拡張に有用な知見を提供した。

また、ジェスチャ認識の評価について村尾ら[4]は、加速度センサによるジェスチャ認識において、ジェスチャの提示方法（テキスト・図・動画）の違いがユーザの動作に大きなばらつきを生み、特に抽象的な指示では動作の再現性が低下することを示した。これは、従来のジェスチャインターフェースが「開発者が定義した動作をユーザに模倣させる」構造に依存し、説明の理解差や身体動作の個人差に起因する不一致が避けられないことを意味する。

そこで、ユーザ自身がジェスチャを定義する手法として、Sriram ら[5]は、RGB カメラで取得した手の動きをその場で

<sup>†1</sup> 中京大学 工学研究科  
Graduate School of Engineering Chukyo University

<sup>†2</sup> 中京大学 工学部  
School of Engineering Chukyo University

学習し、ユーザが任意に定義したジェスチャを高精度に認識する手法を提案した。固定的な動作セットに依存しないため、ユーザ固有の自然な動きをそのまま操作に利用できる点が有効であることを示した。

Wang ら[6]は、ユーザが任意に定義した手書き文字列入力を手の軌跡として認識するジェスチャ入力手法を提案した。提案手法は、取得した軌跡データから特徴量を抽出し、機械学習モデルにより文字列パターンを分類することで、多様なユーザ固有のジェスチャに柔軟に対応できる点が特徴である。評価実験では、高い認識精度とユーザ適応性を確認し、特にユーザが個別にジェスチャを設定できる点が大きな強みとして示した。

以上に挙げた関連研究では、視線とジェスチャの身体動作を操作モダリティとして利用できる点では操作性に対して有用である一方、直感性が十分に議論されなかった。また、開発者が定義した動作方式では個人の直感性を損なう場合もあり、ユーザ定義型の手法でも動作の品質や身体的負荷が体系的に検討されなかった。これらの課題に対して本研究では、ユーザ自身の好みや感覚に基づいた操作基盤を確立し、個人専用のインターフェースにすることで直感的な機械操作を目指す。

### 3. オーケストラシステム

本システムは、視線インターフェースとジェスチャインターフェースで構成し、その概要図を図 1 に示す。

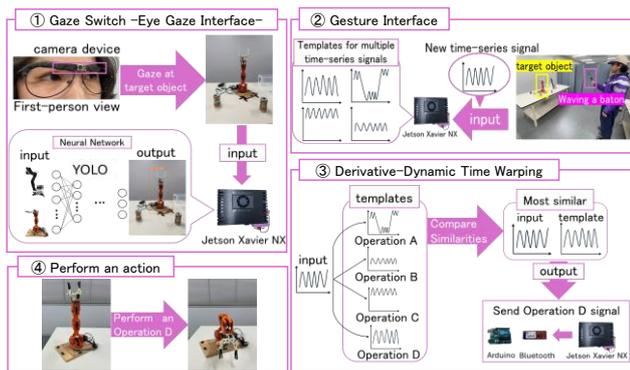


図 1 オーケストラシステムの概要図

図 1 (①) で、視線インターフェースを説明する。視線インターフェースの役割は制御対象を視線によって選択し、それを[7]で提案した視覚スイッチシステムで実現した。視覚スイッチシステムは、小型計算機 Jetson Xavier NX (以下 PC) と小型カメラで構成され、小型カメラをメガネに装着することでユーザの視野を捉え、それを PC に送信する。PC では、robot arm と black robot arm の 2 種類の機械を物体認識できるように学習されたニューラルネットワークがあり、視界中のロボットを検出する。ユーザが操作したい対象を注視し、その対象がカメラ画角の中央に収まった時点で、そ

の機械を制御対象として確定する。すなわち、本システムでは、ユーザが制御対象を視覚で選択できる。

次に図 1 (②, ③) より、ジェスチャインターフェースでは、Arduino と三軸加速度センサで構成された指揮棒を用いて操作を行う。ユーザは自身の感覚に基づき好みの振り方を予め決定し、テンプレートとして時系列信号を PC に保存する。このとき、保存したテンプレートは制御対象のそれぞれの動作と対応している。そして、制御対象に対して、特定の動作を実行させる場合、その動作に対応しているテンプレート通りの振り方で指揮棒を振り、PC に時系列信号を入力する。ここで、DDTW (Derivative-Dynamic Time Warping) を用いて、テンプレートとして保存した時系列信号のそれぞれと入力した時系列信号の類似度を算出する。DDTW は、時系列信号の形状に着目して類似度を評価する手法であり、入力した時系列信号と最も類似度の高いテンプレートに対応した動作を選択する。そして、その動作指示を制御対象へ送信することで特定の動作を実行させる。また、制御対象を切り替える際には、指揮棒を静止状態に保つことでジェスチャ操作が終了し、再度視覚スイッチシステムによる制御対象の選択工程へ戻るように設計した。

### 4. 実験

本実験では、本システムとキーボード操作とボタン操作が混在するリモコン型インターフェースを用いて動作実験とこれら 2 種類のインターフェースのユーザビリティ評価を行った。ユーザビリティ評価実験で得た結果に対して統計的仮説検証を行い、本システムの有する直感性を評価した。

#### 4.1 動作実験

動作実験では、本システムのジェスチャ操作における被験者が決定した振り方に対する操作性や妥当性を確認する目的で行った。

全 7 回で構成されたユーザビリティ評価実験 (i) の各回の実験開始前に動作実験を行った。ここでは被験者自身が決定した振り方に対する操作性を確認し、十分に操作が可能であると主観的に判断した時点で実験を終了した。

次に全 8 回で構成されたユーザビリティ評価実験 (ii) では、初回の実験前にものみ動作実験を行った。ここでは被験者自身が決定した振り方に対する妥当性について確認し、実験監督者からの動作指示を 4 回連続して遂行した時点で動作実験を終了した。以降、最終回まで同一の振り方でユーザビリティ評価実験 (ii) を行った。

#### 4.2 ユーザビリティ評価実験 (i)

ここでは、本システムとリモコン型インターフェースを用いて、4 つの対象物に対して robot arm と black robot arm による逐次的な運搬タスクを行った。ここで、本システムによるユーザビリティ評価実験 (i) の概念図を図 2 に示す。

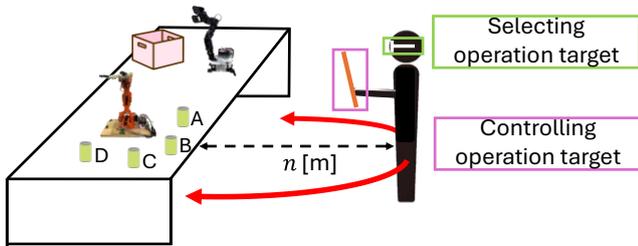


図 2 ユーザビリティ評価実験 (i) の概念図

図 2 より、ユーザビリティ評価実験 (i) では、まず robot arm の周囲に 4 つの対象物を A から D に配置し、ユーザは本システムを用いて robot arm を操作する。4 つの対象物のそれぞれを掴み取るための 4 種類の振り方を事前にテンプレートとして保存し、各動作の識別に用いた。続いて、ユーザは robot arm が掴み取った対象物を black robot arm の前まで移動させ、black robot arm を操作することで、所定の位置に収集させる。このとき、black robot arm は視覚スイッチシステムによって操作し、本システムにおける「操作対象の選択・確定」の用途ではなく、視線入力による直接操作のために用いた。また、ユーザビリティ評価実験 (i) においては、 $n$  を任意とし、ユーザが自由に決定した場所を固定して実験を行った。

本システムのジェスチャ操作における各振り方に対応した動作表を表 1 に示す。

表 1 robot arm の動作一覧 (i)

振り方	robot arm
1	A の対象物を掴む
2	B の対象物を掴む
3	C の対象物を掴む
4	D の対象物を掴む

以下に本システムによる実験手順を示す。

1. 被験者は、決定した振り方に関する 4 つの時系列信号をテンプレートとして PC に保存する。
2. テンプレートを保存した後に動作実験を開始する。
3. 動作実験終了後、システムを起動させ、操作時間の計測を開始する。
4. 被験者は、表 1 に示した動作一覧に倣って、オーケストラシステムにより robot arm を制御し、A に設置された対象物を掴みとり、引き渡し地点に設置する。
5. 引き渡し地点に設置された対象物を視覚スイッチシステムにより black robot arm を制御し、所定の位置に収集する。
6. 手順 4 から手順 5 を A から D に設置された対象物に対して順番に行う。
7. 全ての対象物が所定の位置に収集できた時点で、操作

時間の計測を終了し、システムを終了する。その後、アンケートに回答し、実験を終了する。

次にリモコン型インターフェースによるユーザビリティ評価実験 (i) の概要図を図 3、キーボード操作による動作表を表 2 に示す。

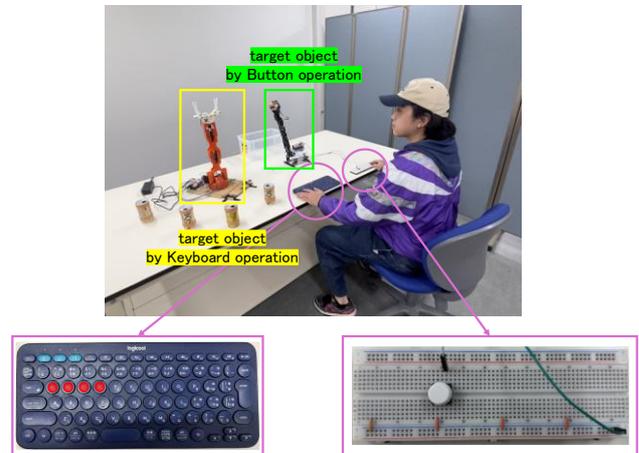


図 3 リモコン型インターフェースの概要図 (i)

表 2 キーボード操作の動作一覧 (i)

	robot arm
Q	A の対象物を掴む
W	B の対象物を掴む
E	C の対象物を掴む
R	D の対象物を掴む

ユーザビリティ評価実験 (i) のキーボード操作では、図 3 に示したキーボードを用いて、robot arm を操作する。キーボードの有効キーは、それぞれ「Q」、「W」、「E」、「R」であり、表 2 に示した動作を実行する。そして、図 3 に示されたボタンを押すことで、所定の位置に収集する。

次にリモコン型インターフェースを用いた実験の手順を示す。

1. 実験監督者の合図で実験を開始し、操作時間の計測を開始する。
2. 被験者は、表 2 に示した動作一覧に倣って、キーボード操作により robot arm を制御し、A に設置された対象物を掴みとり、引き渡し地点に設置する。
3. 引き渡し地点に設置された対象物をボタン操作により black robot arm を制御し、所定の位置に収集する。
4. 手順 2 から手順 3 を A から D に設置された対象物に対して順番に行う。
5. 全ての対象物が所定の位置に収集できた時点で、操作時間の計測を終了し、システムを終了する。その後、アンケートに回答し、実験を終了する。

ユーザビリティ評価実験 (i) では、21 歳から 23 歳の大学院生 2 名、学部生 2 名、研究生 1 名の男子学生 5 名を対象に計 7 回実験を行った。

#### 4.3 ユーザビリティ実験 (ii)

ユーザビリティ評価実験 (ii) では、ユーザビリティ評価実験 (i) と同様に robot arm と black robot arm を用いた逐次的な運搬タスクを実施した。ここでは、図 2 に示した 4 つの設置箇所を 2 つに減らし、A と C をそれぞれ X と Y に改め、ランダムに設置する仕様とした。このとき、ユーザが保存するテンプレートも 2 種類となる。また、ユーザは  $n = 1.5$  の場所からロボットを制御し、必要に応じて自由に動き回って操作することができる環境へと変更した。

本システムのジェスチャ操作における各振り方に対応した動作表を表 3 に示す。

表 3 robot arm の動作一覧 (ii)

振り方	robot arm
1	X の対象物を掴む
2	Y の対象物を掴む

以下に本システムによる実験手順を示す。

1. 被験者は、事前に 2 つの振り方を考え、動作実験を行う。
2. 動作実験終了後、決定した振り方に関する 2 つの時系列信号を PC に保存し、システムを起動する。
3. システム起動後、操作時間の計測を開始する。
4. 実験監督者が X と Y のいずれかにランダムに対象物を設置する。被験者は、設置された対象物に対して、表 3 に示した動作一覧に倣って、オーケストラシステムにより robot arm を制御し、対象物を掴みとり、引き渡し地点に設置する。
5. 引き渡し地点に設置された対象物を視覚スイッチシステムにより black robot arm を制御し、所定の位置に収集する。
6. 計 4 つの対象物に対して、手順 4 から手順 5 を繰り返し行う。
7. 全ての対象物が所定の位置に収集できた時点で、操作時間の計測を終了し、システムを終了する。その後、アンケートに回答し、実験を終了する。

次にユーザビリティ評価実験 (ii) のリモコン型インターフェースの概要を図 4、キーボード操作による動作表を表 4 に示す。

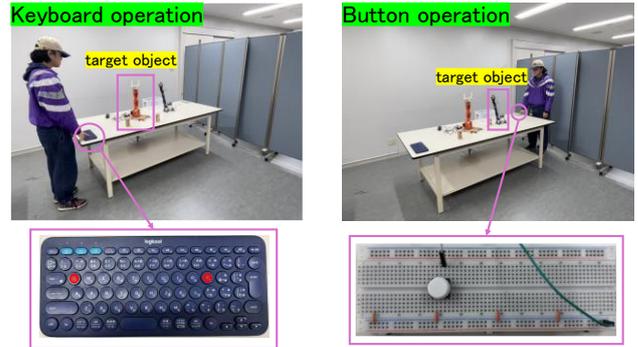


図 4 リモコン型インターフェースの概要図 (ii)

表 4 キーボード操作の動作一覧 (ii)

	robot arm
Q	X の対象物を掴む
P	Y の対象物を掴む

ユーザビリティ評価実験 (ii) におけるリモコン型インターフェースでは、ユーザビリティ評価実験 (i) と同様の仕様で robot arm と black robot arm を制御する。ここでは、図 4 より、キーボード操作の有効キーを「Q」、「P」の 2 つに変更し、表 4 に示した動作を実行する。このとき、キーボードとボタンは図 4 に示したように離れた場所に設置されており、その都度ユーザ自身が移動してそれぞれのロボットを制御する。

以下にリモコン型インターフェース用いた実験の手順を示す。

1. 実験監督者の合図で実験を開始し、操作時間の計測を開始する。
2. 実験監督者が X と Y のいずれかにランダムに対象物を設置する。被験者は、設置された対象物に対して、表 4 に示した動作一覧に倣って、キーボード操作により robot arm を制御し、対象物を引き渡し地点に設置する。
3. 引き渡し地点に設置された対象物をボタン操作により black robot arm を制御し、所定の位置に収集する。
4. 計 4 つの対象物に対して、手順 2 から手順 3 を繰り返し行う。
5. 全ての対象物が所定の位置に収集できた時点で、操作時間の計測を終了し、システムを終了する。その後、アンケートに回答し、実験を終了する。

ユーザビリティ評価実験 (ii) では、22 歳の学部生 1 名を追加し、男子学生 6 名を対象に計 8 回実験を行った。

#### 4.4 評価方法

各回の実験終了後にシステムの使いやすさを評価する方法として、システムユーザビリティスケール (SUS) を用い

たアンケートに回答してもらい、100点満点のスコア（SUSスコア）に変換し、各システムのユーザビリティをスコアリングした。また有意水準5%で、t検定、一元配置分散分析、二元配置分散分析を以下の内容で用いて、有意差と交互作用の有無について調査し、オーケストラシステムの直感性について評価した。

- ・被験者別の両システムのSUSスコアに対するt検定
- ・システム別のSUSスコアに関して、実験回数を要因とした一元配置分散分析
- ・システム別のSUSスコアに関して、被験者を要因とした一元配置分散分析
- ・両システムのSUSスコアに関して、システムと実験回数を要因とした二元配置分散分析
- ・両システムのSUSスコアに関して、システムと被験者を要因とした二元配置分散分析
- ・両システムの操作時間に関して、システムと実験回数を要因とした二元配置分散分析
- ・両システムの正確率に関して、システムと実験回数を要因とした二元配置分散分析

## 5. 実験結果と考察

ユーザビリティ評価実験 (i) における各回で得られた本システムとリモコン型インターフェースの平均 SUS スコア、平均操作時間、平均正確率を図 5 に示す。

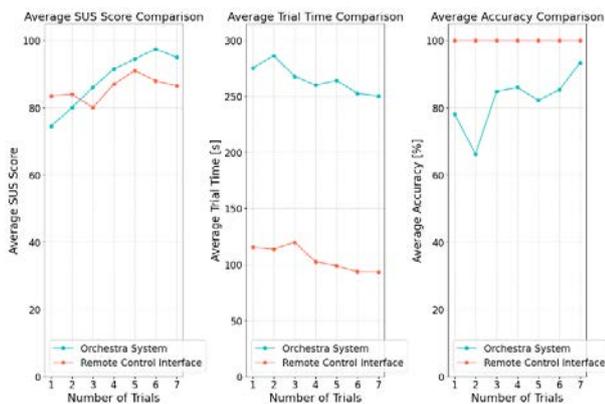


図 5 ユーザビリティ評価実験 (i) の結果

全7回の実験で得た被験者5名分のSUSスコアの平均は、本システムが88.4、リモコン型インターフェースが85.7であった。図5(左)より、本システムは、実験回数が増加するごとにSUSスコアの上昇を確認できる。また、図5(中)より、実験回数に応じて操作時間を短縮する傾向にあると読み取ることができる。これらの結果から、被験者は実験回数に伴って本システムにおける操作習熟度が上昇していると推測できる。図5(右)より、実験回数に伴い上昇傾向にあるが、実験中に被験者の意図に反した動作を実行してい

ることが多々あった。この原因として、本システムのジェスチャ操作において、三軸加速度センサからの三軸分の加速度を直列的に入力してしまっていることが挙げられる。

次に、SUSスコアに関して、システムと実験回数を要因とした二元配置分散分析を行った。ここでは、システム要因、実験回数要因ともに有意差を得ることはできず、交互作用も認められなかった。従って、両システムの使いやすさは実験回数に依存せず、概ね同程度であると分かる。ここで、実験回数を要因とした一元配置分散分析を両システムのSUSスコアに対して行った結果、本システムにおいて有意差を得た ( $F(6,34)=3.443, p=0.011$ )。そこで、Bonferroni法によるt検定の多重比較を行った結果、3回目と6回目におけるSUSスコアの比較において有意差を得た。従って、被験者は、実験回数に伴って本システムにおける操作習熟度が向上し、SUSスコアが上昇したといえる。

続いて、SUSスコアに関して、システムと被験者を要因とした二元配置分散分析を行った。その結果、被験者要因 ( $F(4,69)=11.45, p<0.001$ ) で有意差を得ることができ、交互作用 ( $F(4,69)=3.726, p=0.008$ ) も確認することができた。従って、被験者間におけるシステムの評価基準に大きな差が存在することが示され、交互作用が有意であることから、各被験者のシステムの好みが生じていることが分かった。また、被験者別に両システムに対してt検定を行った結果、本システムに対して有意な結果を得たのは1名、リモコン型インターフェースにおいて有意な結果を得たのは1名、他3名の被験者に関しては、各システムのSUSスコアの差に有意差はなかった。

最後に操作時間と正確率に対して、システムと実験回数を要因とする二元配置分散分析を行った。操作時間に対して、システム要因 ( $F(1,69)=399.3, p<0.001$ )、有意差を得た。図5(中)より、本システムとリモコン型インターフェースの操作時間には大きな差があり、本システムにおける視線操作とジェスチャ操作を直列処理で行っていることから操作の切り替えで生じる待機時間が原因として挙げられる。そのため、リモコン型インターフェースでは、操作の切り替えがスムーズに行えることから操作時間に大きな差が生じている。また、正確率に対しても、システム要因 ( $F(1,69)=25.85, p<0.001$ ) で有意差を得た。図5(右)より、リモコン型インターフェースに対して、有意な結果を得ていることが分かる。

次にユーザビリティ評価実験 (ii) における各回で得られた本システムとリモコン型インターフェースの平均 SUS スコア、平均操作時間、平均正確率を図 6 に示す。

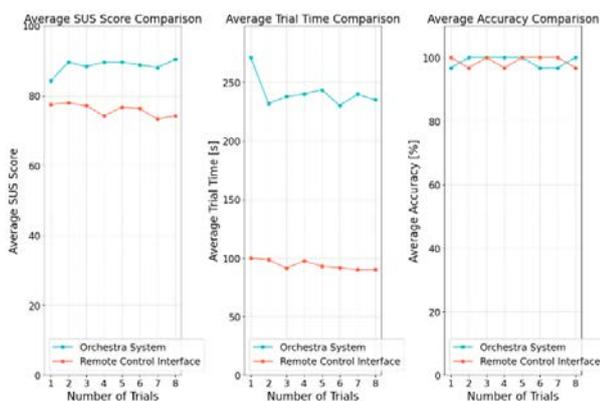


図 6 ユーザビリティ評価実験 (ii) の結果

図 6 (左) より、実験回数に伴う SUS スコアの上昇を確認することができなかった。また、図 6 (中) より、本システムの 1 回目から 2 回目までの大きな変動以降、リモコン型インターフェースも同様に実験回数に伴う変動を確認することはできなかった。しかし、全 8 回の実験で得た被験者 6 名分の SUS スコアの平均は本システムが 90.4、リモコン型インターフェースが 74.2 と本システムにおいて大差で良好な結果を得た。また、図 6 (右) において、本システムにおける平均正確率が図 5 (右) に示した結果に比べ、高水準で安定していることが確認できる。ここで、SUS スコアに関して、システムと実験回数を要因とした二元配置分散分析を行った。ここでは、システム要因 ( $F(1,95)=40.19, p<0.001$ ) で有意差を得ることができ、両システム間の SUS スコアに明確な差があることを確認した。

続いて、SUS スコアに関して、システムと被験者を要因とした二元配置分散分析を行った。その結果、システム要因 ( $F(1,95)=134.9, p<0.001$ )、被験者要因 ( $F(5,95)=9.521, p<0.001$ ) で有意差を得ることができ、交互作用 ( $F(5,95)=29.38, p<0.001$ ) も確認できた。従って、図 6 (左) より本システムの方が平均的に使いやすいと評価されたことが分かり、評価基準や操作能力の個人差が強く影響していることが分かる。さらに、交互作用が有意であることから、システムの良し悪しに対する感じ方が被験者ごとに大きく異なることが示された。また、被験者別に両システムに対して t 検定を行った結果、本システムに対して有意な結果を得たのは 5 名、リモコン型インターフェースにおいて有意な結果を得たのは 1 名という結果であった。リモコン型インターフェースに対して有意な結果を得た被験者はユーザビリティ評価実験 (ii) から参加した新たな被験者である。また、他 5 名に関しては、ユーザビリティ評価実験 (i) での経験を活かすことができたと考える。

最後に操作時間と正確率に対して、システムと実験回数を要因とする二元配置分散分析を行った。操作時間に対して、システム要因 ( $F(1,95)=1056, p<0.001$ ) で有意差を得た。図 6 (中) より、ユーザビリティ評価実験 (i) の結果と同様

のことが言える。また、正確率に対しては、システム要因並びに実験回数要因に伴って有意差を得ることができなかった。このことから、ユーザビリティ評価実験 (i) で得た結果に比べ、本システムが安定して操作できていることが分かった。

## 6. おわりに

ユーザビリティ評価実験の結果から、本システムは主観的な使いやすさにおいて高い評価を獲得し、操作内容そのものが直感的に理解しやすいことが示された。一方で、操作時間はリモコン型インターフェースに劣っていることが確認できた。この原因として、視線操作とジェスチャ操作を直列処理するアーキテクチャが直感的な操作フローを妨げていた可能性がある。従って、ユーザビリティ評価実験 (ii) のような単純な運搬作業を行う環境下で、本システムは操作概念としては十分に直感的であるが、リアルタイム性や応答性の面で改善の余地があると言える。

今後の展望として、本システムにおける視線操作とジェスチャ操作を並列処理で行うようなアーキテクチャを確立し、操作時間の短縮を狙うことで、ユーザに対するストレス軽減を目指す。また、視覚スイッチシステムに関して、ユーザの身長や視線を考慮して学習された個人専用のニューラルネットワークを用いて、個人に最適化された視線操作を導入する。より個人専用のインターフェースとして確立させることで、直感的な機械操作の実現を目指す。

## 参考文献

- [1] 久保田大和, ピトヨ ハルトノ, オーケストラシステムとしてのマン・マシンインターフェースの開発, 情報処理学会インタラクション 2025, pp.1391-1396, 2025.
- [2] I. Tokmurziyev, et al., GazeGrasp: DNN-Driven Robotic Grasping with Wearable Eye-Gaze Interface, IEEE/ACM International Conference on Human-Robot Interaction, pp.1680-1684, 2025.
- [3] K. Taylor, et al., Towards the Fusion of Gaze and Micro-Gestures, IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, pp.1474-1475, 2025.
- [4] K. Murao and T. Terada, Evaluating Instructions for Gesture Recognition with an Accelerometer, ヒューマンインターフェース学会論文誌 17 巻 1 号, pp.73-84, 2015.
- [5] S. K. Sriram and N. Sinha, GESTOP : Customizable Gesture Control of Computer Systems, COD COMAD 2021, pp.405-409, 2021.
- [6] J. Wang, et al., Hand gesture recognition for user-defined textual inputs and gestures, Universal Access in the Information Society, pp.1315-1329, 2025.
- [7] S. Kobayashi and P. Hartono, Eye Contact as a New Modality for Man-machine Interface, International Journal of Advanced Computer Science and Applications, 14(3), pp.42-49, 2023.