

真似て選択するデジタルサイネージにおけるジェスチャ入力方式の比較

令和7年度 卒業論文

日本大学 文理学部 情報科学科 宮田研究室

坂内 武希

概要

デジタルサイネージに複数のコンテンツが表示されている状況において、デジタルサイネージとスマートフォンを連携させる場合、ユーザはデジタルサイネージから対象のコンテンツを1つ選択する必要がある。しかし、NFCタグやカメラ等を用いた従来方式では、ユーザへの制約・操作負担が大きいという問題があった。この問題を解決するために我々は、デジタルサイネージに表示されている複数のコンテンツの中から、スマートフォンと連携させるものを選択する方法として、デジタルサイネージ上に表示されたアイコンの動きを真似ることによって選択できる方式を提案した。しかし、アイコンの動きを真似る際に利用シーンによって適切なものが異なると考えられる。よって本研究ではアイコンの動きを真似る際に、腕を振ってジェスチャを行う方式と、タッチディスプレイ上で指を滑らせてジェスチャを行う方式の比較を2つの仮説を基に行い、特徴を明らかにすることを研究課題とした。本稿では、2つの方式を用いた比較実験を行った。その結果、腕を振ってジェスチャを行う方式はスマートフォンのタッチディスプレイ上で指を滑らせる方式に比べて、精度が低く操作負担が生じることが分かった。また、精度に関する問題を解消するために、腕を振ってジェスチャを行う方式のディープラーニングを用いた性能向上の検討を述べる。

目 次

第1章 序論	1
1.1 研究の背景	2
1.2 研究の目的	2
1.3 本論文の構成	2
第2章 デジタルサイネージにおける選択方式とジェスチャインターフェースによる選択方式に関する研究事例	4
2.1 デジタルサイネージにおける選択方式に関する研究事例	5
2.2 ジェスチャインターフェースによる選択方式に関する研究事例	7
第3章 研究課題	9
3.1 問題の定義	10
3.2 研究課題の設定	10
第4章 研究デザイン	11
4.1 研究の流れ	12
4.2 仮説の定義	12
第5章 デジタルサイネージ向けのアイコンの動作を真似る選択システムの実装	14
5.1 システムの全体像	15
5.2 スマートフォンクライアント	15
5.2.1 腕を振ってジェスチャを行う方式	15
5.2.2 スマートフォンの画面上で指を滑らせてジェスチャを行う方式	15
5.3 デジタルサイネージクライアント	16
5.4 サーバ	17
第6章 比較実験	18
6.1 実験の目的	19
6.2 実験の概要	19
6.3 実験の手順	19
6.4 評価指標・測定方法	20
6.5 実験の結果	20
6.6 考察	21

6.6.1 ジェスチャマッチング精度の考察	21
6.6.2 操作負担の考察	21
第7章 性能向上の検討	23
7.1 性能向上のためアプローチ	24
7.2 実装	24
7.3 データセットの作成	24
7.4 モデル構築	25
7.5 今後の予定	25
第8章 結論	26
謝辞	28
参考文献	30
付録	33
研究業績	34

図 目 次

5.1	システム構成図	15
5.2	腕を振ってジェスチャを行う方式	16
5.3	スマートフォンの画面上で指を滑らせてジェスチャを行う方式	16
5.4	サイネージ画面	17
5.5	アイコンの動作パターン	17
6.1	実験協力者別ジェスチャマッチング成功率	20
6.2	方式別 NASA-TLX スコア	21

表 目 次

第1章 序論

1.1 研究の背景

昨今、デジタルサイネージは商業施設や公共施設など多くの場所に設置されている。その中でも、デジタルサイネージと自身のスマートフォンを連携させるシステムが主流となっている。デジタルサイネージに複数のコンテンツが表示される場合、ユーザは最初に対象のコンテンツを近づいて1つ選択する必要がある。しかし、操作者以外の人間は目の前のデジタルサイネージから情報を得ることは不可能である。

デジタルサイネージに表示されている複数のコンテンツから1つ選択する時の占有率を解消した、NFCタグやカメラ等を用いた従来方式では、デジタルサイネージとユーザ間の距離や人・物の存在に関する制約・操作負担が大きいという問題がある。

1.2 研究の目的

この問題を解決するために先行研究 [1, 2, 3] では、デジタルサイネージ上に複数のコンテンツが存在するシーンにおいて、各コンテンツに一意に関連付けられたアイコンの動きを真似るジェスチャを行うことで任意のコンテンツを選択できる方式を提案してきた。これは、ユーザがスマートフォンを把持して各デジタルサイネージ上のアイコン動作と同じタイミングで同じ動きのジェスチャを腕を振るようにして行うことで、対応するデジタルサイネージ上のコンテンツを選択できる方式である。

この方式を各学会等の場でデモンストレーションしていたところ、腕を振ってコンテンツを選択する操作に楽しさを覚える人がいた。一方で、公共空間で腕を振るジェスチャを行うのは操作負担が生じる等の意見もあり、どのようなジェスチャがユーザに受け入れられるのか不明瞭であった。

よって、先行研究 [4] では、タッチディスプレイ上で指を滑らせてジェスチャを行う方式の実装を行った。これは、ユーザがスマートフォンを把持して各デジタルサイネージ上のアイコン動作と同じタイミングで同じ動きのジェスチャをスマートフォンの画面上で指を滑らせるようにして行うことで、対応するデジタルサイネージ上のコンテンツを選択できる方式である。

本稿では、腕を振ってジェスチャを行う方式 [3] と、タッチディスプレイ上で指を滑らせてジェスチャを行う方式 [4] の比較実験を行い、得られた結果から、両方式の精度、操作負担を明らかにする。

1.3 本論文の構成

本論文の構成は次のとおりである。

2章では、デジタルサイネージとスマートフォンを連携して利用する事例に関する研究事例について述べる。

3章では、本論文における問題の定義と研究課題について述べる。

4章では、本論文における研究デザインを述べる。

5章では、デジタルサイネージ向けのアイコン動作を真似る選択システムに関する実装について述べる。

6章では、デジタルサイネージ向けのアイコン動作を真似る選択システムに関する比較実験・考察について述べる。

7章では、性能向上の検討について述べる。

最後に8章にて、本論文の結論を述べる。

第2章 デジタルサイネージにおける選択 方式とジェスチャインターフェー スによる選択方式に関する研究 事例

本章では、デジタルサイネージにおける選択方式とジェスチャインターフェースによる選択方式に関する研究事例について述べる。

2.1節では、デジタルサイネージにおける選択方式に関する研究事例について紹介する。

2.2節では、ジェスチャインターフェースによる選択方式に関する研究事例について紹介する。

2.1 デジタルサイネージにおける選択方式に関する研究事例

デジタルサイネージにおける選択方式に関する研究事例として [5], [6], [7], [8], [9], [10], [11], [12], [13], [14] が挙げられる。

Hardy ら [5] は、タイル状に敷き詰めた NFC タグ上に映像を投影するデジタルサイネージシステムを提案した。このシステムは、ユーザが NFC リーダ搭載のスマートフォンをデジタルサイネージ上の任意の位置に接触させることで該当位置をポインティングできるものである。

Boring らは、Shoot & Copy [6] として、スマートフォンのカメラで撮影したデジタルサイネージ上の領域を判定し、その領域にあるコンテンツをスマートフォン上にコピーするシステムである。撮影領域の判定は、撮影画像とデジタルサイネージ上の各領域の画像類似度に基づいて行われる。

また、Touch Projector [7] として、デジタルサイネージ上のオブジェクトをスマートフォンのビデオカメラで撮影することで、当該オブジェクトをポインティングできる仕組みである。ユーザが複数あるデジタルサイネージの中から任意の1つを撮影すると、デジタルサイネージ上のオブジェクトの配置が画像処理によって認識され、どのデジタルサイネージが撮影されているか判定される。

Shirazi ら [8] は、デジタルサイネージの前に設置したカメラでスマートフォンのフラッシュライトを検出することで、ユーザがデジタルサイネージ上のどの位置にスマートフォンをかざしているか判定する方式を提案した。これにより、ユーザはデジタルサイネージ上の任意の位置をポインティングすることができる。

Schmidt らの PhoneTouch [9] は、スマートフォンでディスプレイをタッチすることで、ディスプレイ上の任意位置をポインティングすることができるシステムである。ユーザがスマートフォンでタッチした位置は、ディスプレイの役割を果たすスクリーンの裏側に設置したカメラを用いて検出する。このシステムはテーブルトップインターフェース向けであるが、原理的にはデジタルサイネージにも応用可能である。

Hannah ら [10] は、デジタルサイネージのジェスチャベースのインタラクションにおいて、空中触覚フィードバックがユーザのエンゲージメントに与える影響の調査を行っている。被験者 17 名を対象とした実験では、インタラクティビティのレベル（高/低）とハaptic キューの有無（オン/オフ）を変数として、ユーザエンゲージメントを測定した。結果として、空中触覚フィードバックがユーザの集中力、使いやすさ、美的魅力、報酬感を有意に向上させることが示され、特に高いインタラクティビティがエンゲージメントを高めることが確認された。

Georgios ら [11] は、超音波を利用して空中に触覚を与える、空中ハapticインターフェースを用いて、デジタルサイネージやキオスクでの非接触型の操作体験の向上について述べている。実験では、この技術がユーザの感情状態やエンゲージメントに与える影響が調査された。結果、空中でのフィードバックがコンテンツの認知やユーザの記憶を向上させることができることが確認された。さらに、マルチユーザの体験やアクセシビリティの向上についての調査も行われ、特に障害を持つユーザに対して有効であることが示唆された。

Abe ら [12] は、ユーザが距離を置いても両手の指を使って操作できる、空中マルチインターフェースを提案している。実験では、広視野カメラと高速パンチルトカメラを使って、両手の骨格検出を行い、指の動きを取得した。また、12種類の異なる手のポーズを実験協力者が行い、その検出精度を評価した。結果として、広視野カメラを使用した場合と比較して、提案手法の方が高い検出率を出した。また、中指のみや、中指・薬指・小指を伸ばしたポーズでは、一部失敗があったものの、それ以外のポーズでは高い検出率を記録した。加えて、骨格検出の動機によって、指の動きの正確性が向上し、スムーズな軌跡をかけることが確認された。また、このインターフェースを活用したドローイングアプリケーションと写真配置アプリケーションを実装した。残者では複数の指を使った画像の削除が可能で、後者では写真の移動、ズーム、回転が簡単に行える。今後の課題として、複数ユーザでの操作対応、既存のインターフェースとの比較によるユーザビリティ評価があげられている。この技術は直感的な操作性と高精度な動作かかのうなことから、デジタルサイネージにも応用可能である。

Liao ら [13] は、携帯端末における「SMART」というシステムを提案している。これは、画面と周囲光センサを活用して、追加のハードウェアを必要とせずにジェスチャ認識を可能にしている。実験では、ジェスチャ認識の精度と実用性を評価した。14名の実験協力者に9種類のジェスチャを各20回行い、異なる照明環境下で実験を行った。結果、94.3%から96.9%の精度の達成し、動的な照明環境下でも93%以上の精度を記録した。さらに、ユーザ間の多様性や未経験の環境下でも、高い認知制度を示した。これらの結果から、SMARTは様々な環境や条件でも高い信頼性と汎用性を持つことが確認された。この技術は、原理的にはデジタルサイネージにも応用可能である。

Otaki ら [14] は、指先に装着するデバイスを用いて、ジェスチャ入力を可能にするインターフェースを提案した。このデバイスは、3つのタッチセンサとひずみゲージを備え、指先に取り付けられ、5種類のジェスチャと4方向の力の入力を行える。提案された指先装着デバイスのジェスチャ入力制度と、日常生活での誤入力の発生状況を確認するために、2つの実験が行われた。1つ目の実験では、6人の実験協力者が5種類のジェスチャ左右のスワイプ、上下のスワイプ、タップ)を10回ずつ行い、そのデータを認識精度の評価に使用した。結果として、全参加者のデータを使用した場合の平均認識精度は94.6%，個々のデータを使用した場合は93.6%となり、高い認識精度が確認された。2つ目の実験では、日常生活の中で誤入力が発生する状況を調査した。デバイスを装着したまま日常的な活動を行ったところ、特定の動作が誤入力を引き起こすことが分かったが、壁やデーブルに力を加えるジェスチャでは誤入力が発生しなかった。これらの結果から、微小なジェスチャと力の入力をサポートするデバイスは、日常生活における効率的なインターラクションを

提供する可能性が示された。この技術は、原理的にデジタルサイネージにも応用可能である。

2.2 ジェスチャインターフェースによる選択方式に関する研究事例

ジェスチャインターフェースによる選択方式に関する研究事例として [15], [16], [17], [18], [19], [20] が挙げられる。

Tao ら [15] は、ユーザが空中で手を動かすことによって入力を行う、フリーハンドジェスチャで任意の項目を選択する時のマッチング精度の向上を行った。手首の傾きと、複数のピンチジェスチャによって制御され、選択する項目に傾きとピンチジェスチャをマッピングすることにより実現している。

Jason ら [16] は、足を使ったジェスチャ入力を用いている。タッチやスワイプといった、一般的なモバイルデバイスのコマンドをジェスチャにマッピングすることによって実現している。ユーザの手が汚れていてふさがっている、上半身のスペースが確保できず使えない状況での代替的なジェスチャとして有効である。

Li ら [17] は、スマートスピーカーを使用した部屋規模の手のジェスチャー認識システム「SpeakerGesture」を提案している。従来の音響センサーは近距離での認識に制限されていたが、新しい信号処理技術を開発し、認識範囲を 1 メートルから 4-5 メートルに拡大した。実験では、16 人の参加者から 1440 のジェスチャーを収集し、強い干渉がある状況でも 90% 以上の認識精度を達成した。このシステムは、音楽再生や音声コントロールと同時に機能し、ユーザーの体験を向上させることが期待される。

Viollet ら [18] は、LinLED の性能を評価するための実験を行った。LinLED は、手や指の動きを高精度で検出でき、かつ低遅延の非接触型ジェスチャインターフェースであるリニアステージと 3D プリントされた手を用いて、手を 0.8mm ずつ動かしながら 32 個のフォトダイオードの信号を記録した。結果、位置検出の分析能力が 1mm と非常に高精度であることが分かった。また、回転ブレードを用いた実験により、信号処理の遅延が 0.4ms、全体の遅延が 1ms 以下と測定された。さらに、異なる肌のトーンやサイズの手を使用した検証では、応答が線形で滑らかであることが確認され、LinLED の優れた性能が示された。

Waugh ら [19] は、非接触型ユーザインタフェースにおける、スライダーコントロールジェスチャを「プッシュ」と「ピンチ」の 2 つの方法で比較をしている。実験では、非接触型ユーザインタフェースでのスライダーコントロールにおける 4 つのジェスチャ（エアプッシュ、カーソルウェル、ピンチ、ピンチウェア）の効果の評価を行った。結果として、ピンチウェアジェスチャが最も早くスライダ操作を出来ることが分かり、特にスライダのハンドルを取得する時間が短縮された。また、ピンチジェスチャは、継続的なコントロールを必要とする操作において、信頼性が高く効率的な手段であることが示された。

Yan ら [20] は、動的な手のジェスチャ認識を最適化された特徴選択を通じて向上させるために、ダブルマシンラーニング (DML) を用いた方法を提案している。実験では、DML

を用いた特徴選択方法の有効性を評価するために、12人の実験協力者が10種類の手のジェスチャを各10回行った。収集されたデータは4つの異なるデータセットに整理され、DMLを含む4つの特徴選択方法（PCA, SFM, VAR）と、6つの機械学習モデル（ロジスティクス回帰、k近傍法、ランダムフォレスト、エクストラツリー、ヒストグラムベースの勾配ビースティング、ライトGBM）を用いて分類した。結果として、DMLを用いた場合、特に骨の特徴を使用したデータセットで最も高い精度を示し、他の方法と比較して分類のパフォーマンスが向上した。

第3章 研究課題

本章では、本研究における問題の定義と研究課題について述べる。

3.1 問題の定義

先行研究 [1, 2, 3] は、デジタルサイネージ上に一定の動きをするアイコンを表示し、ユーザがスマートフォンを持って、デジタルサイネージ上に表示されたアイコンの動きを、腕を振って真似るジェスチャを行うことによって、対応したコンテンツを選択できる方式である。

この方式を各学会等の場でデモンストレーションしていたところ、腕を振ってコンテンツを選択する操作に楽しさを覚える人がいた。一方で、腕を振るジェスチャを行うのは操作負担が生じる等の意見もあり、どのようなジェスチャがユーザに受け入れられるのか不明瞭であった。

ユーザが身体を用いてジェスチャを行う方法は、足を用いる方式 [16] や手を用いる方式 [15] 等多数存在する [21]。この中でも、先行研究 [4] では、ジェスチャを行う方式としてよく用いられており、入力時の負担が少ないと思われる、手を用いる方式に着目した。腕を振ることによって生じる操作負担を解消する目的として、スマートフォンの画面上で指を滑らせてジェスチャを行うことによって対応したコンテンツを選択できる方式の実装をした。

しかし、実際に指を画面上で滑らせる方式が、腕を振る方式より操作負担が生じないのがが検証されておらず。また、三次元空間上で行う腕を振るジェスチャを行う方式と、二次元空間上で行う画面上で指を滑らせる方式の精度が対等になっているかが不明瞭であった。

3.2 研究課題の設定

3.1 節で述べたように、両方式の精度、操作負担の差が不明瞭であるという問題があり、各ジェスチャ方式の精度、操作負担を明確にする必要があると考えられる。

本研究では、腕を振ってジェスチャを行う方式 [1] と、スマートフォンのタッチディスプレイ上で指を滑らせてジェスチャを行う方式 [22] の比較を行うことによって精度・操作負担を明らかにすることを研究課題とする。

第4章 研究デザイン

本章では、本論文における研究デザインを述べる。

4.1 研究の流れ

本節では、本研究の流れについて述べる。本稿では、先行研究のシステム [3, 4] の比較実験を仮説を基に行い、得られた結果の比較、考察を行い各方式の特徴を明らかにすることを目標としている。

実験では3.2節で定義した研究課題を達成するために、2つの方式を用いた比較実験を行った。1つ目は、スマートフォンを利き手で持ち、腕を振ることによってデジタルサイネージから任意のコンテンツを取得できる方式である。2つ目は、スマートフォンの画面上で指を滑らせることによってデジタルサイネージから任意のコンテンツを取得できる方式である。以降は、これらを方式1、方式2とする。実験は両方式とも、大学の研究室内で行った。研究室内にある60インチモニターをデジタルサイネージとして、そこに向かってジェスチャを行ってもらった。実験協力者は、大学生18名である。データの測定には、実験協力者に行ってもらったジェスチャのデータと、各方式終了後のアンケートを用いた。ジェスチャのデータの中で、特にコンテンツ選択の成功率に着目し、精度の検証に用いた。アンケートでは、操作負担を計測するために、日本語版NASA-TLX[23, 24]を用いた。本稿では、これらの結果を比較することによって、各方式の精度、操作負担を明らかにする。

4.2 仮説の定義

本節では、本研究での着眼点及び、それに基づく仮説について述べる。本稿では、比較実験を行うにあたって、2つの点に着目した。1つ目は、三次元空間上で行う腕を振るジェスチャを行う方式と、二次元空間上で指を滑らせる方式の精度の差という点である。方式1では計測時、スマートフォンの加速度センサを用いて値を取得している。その際、スマートフォンを傾けることによって、軸のブレが生じてしまい、正確な値を取得できない可能性がある。それに比べて、方式2では、画面上で指を滑らせて取得した座標の値を取得しているため、軸のブレがない。

2つ目は、腕を振る動作が必要な方式1と、画面上で指を滑らせるだけで済む方式2の操作負担の差という点である。方式1では、コンテンツを正確に取得するには、スマートフォンの軸をずらさずに腕全体を振る必要がある。しかし、方式2では画面上で指を滑らせるだけでコンテンツを正確に取得することができる。

本稿では、これらの特徴から、比較実験するにあたって2つの仮説を定義する。

H1 三次元空間上の動作のマッチングを行う方式1より、自由度が相対的に小さい二次元空間上の動作のマッチングを行う方式2の方がマッチング精度が高い。

H2 腕を振る動作が必要な方式1より、画面上で指を滑らせるだけの動作で済む方式2の方がユーザは操作負担が生じにくい。

この2つの仮説の検証を行うことを本稿の比較実験の目的とする。

第5章 デジタルサイネージ向けのアイコンの動作を真似る選択システムの実装

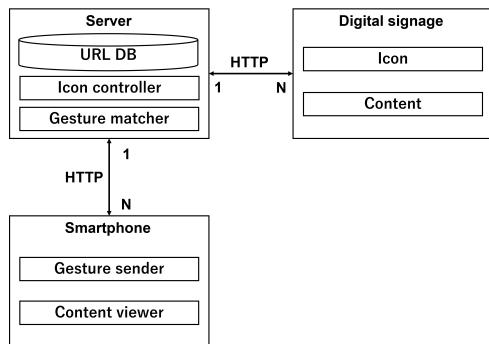


図 5.1: システム構成図

本章では、デジタルサイネージ向けのアイコンの動作を真似る選択システムに関する実装について述べる。

5.1 システムの全体像

全体のシステム構成図を 5.1 に示す。システムは、サーバ、デジタルサイネージクライアント、スマートフォンクライアントからなるクライアントサーバモデルで実装する。

5.2 スマートフォンクライアント

スマートフォンクライアントでは、画面上部にトリガーが配置されている。ユーザはトリガーを押しながら腕を振る、またはトリガーの中で指を滑らせるジェスチャによって計測された時系列データをサーバへ送信し、アイコンの動きとのマッチング結果として返ってきた URL を開き、コンテンツの表示を行う。

5.2.1 腕を振ってジェスチャを行う方式

図 5.2 のように、スマートフォンの画面上部にボタンが表示されている。ボタンを押した時に、加速度センサ値の計測を開始し、ボタンが押されている間に加速度センサ値の計測を行い、ボタンを離した時に時系列データとしてサーバへ送信する。ユーザはボタンを押しながらアイコンの動きを真似て腕を振るジェスチャを行うことによって、マッチングを行うことができる。

5.2.2 スマートフォンの画面上で指を滑らせてジェスチャを行う方式

図 5.3 のように、スマートフォンの画面上部に枠が表示されている。枠内で指を滑らせている間の画面上の軌跡の座標を計測し、計測した軌跡の座標を時系列データとしてサー

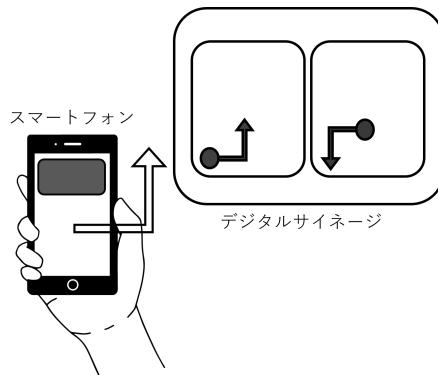


図 5.2: 腕を振ってジェスチャを行う方式

バへ送信する。ユーザは枠内でアイコンの動きを真似るように指を滑らせるジェスチャを行うことによって、マッチングを行うことができる。

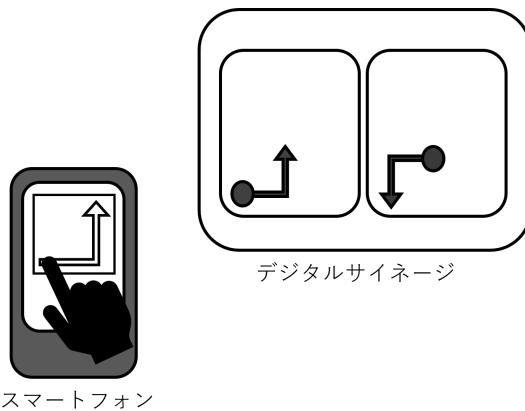


図 5.3: スマートフォンの画面上で指を滑らせてジェスチャを行う方式

5.3 デジタルサイネージクライアント

デジタルサイネージクライアントでは、デジタルサイネージ上にコンテンツとそれに対応したアイコンを表示する。図 5.4 に実際のデジタルサイネージの画面を示す。表示するアイコンは図 5.5 のようなシンプルなものである。アイコンは指定された動作時間帯にのみ表示されており、アイコンの動作時間以外はアイコンの軌跡のみが表示される。これは、ユーザがアイコンの動きを記憶しなくて済むようにするためである。表示するコンテンツは URL パラメータで、任意のサイネージコンテンツを表示することができる。

デジタルサイネージクライアントは JavaScript で Web アプリケーションとして実装しているため、Web ブラウザがあれば、OS を問わず様々なデバイスをデジタルサイネージとして使用することができる。

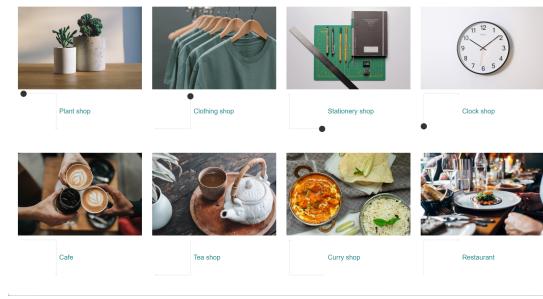


図 5.4: サイネージ画面

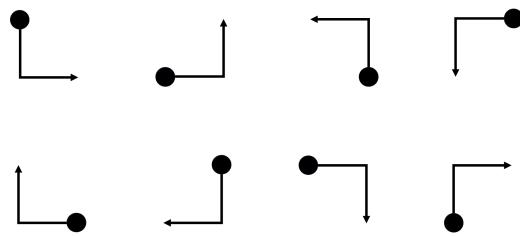


図 5.5: アイコンの動作パターン

5.4 サーバ

サーバでは、スマートフォンクライアントから送られてきた時系列データを受け取る。取得したデータと図 5.4 に示されたアイコンの動きの類似度を算出し、アイコンの動きと結び付けられたコンテンツの URL をスマートフォンクライアントへ送信する。1つのアイコンに対して、1つの URL と類似度を算出するために用いる正解データが結び付けられている。類似度の算出には、高速に計算でき、データポイントが多い時系列データに対しても効率的に使用できるユークリッド距離を採用した。

算出した類似度が最大となるアイコンの動きと対応するコンテンツをマッチング結果として、アイコンの動きに結び付けられた URL をスマートフォンクライアントへ送信する。なお、両方式ともに、スマートフォンクライアントから受け取った時系列データには、スケーリングを施している。加えて、方式 1においては、計算量削減のために、次元削減などを施している。

第6章 比較実験

本章では、デジタルサイネージ上のコンテンツ選択のためのジェスチャ入力方式に関する比較実験、結果・考察について述べる。

6.1 実験の目的

本実験では、2つの方式を比較し、精度、操作負担を明らかにするにあたり、4.2で定義した仮説の検証を行うことを実験の目的とする。

6.2 実験の概要

実験システムは5章で述べた通りに実装した。方式は腕を振ってジェスチャを行う方式とスマートフォンのタッチディスプレイ上で指を滑らせるジェスチャを行う方式の2つである。実験協力者はスマートフォンを使い慣れている大学生18名である。実験は大学の研究室で行い、60インチディスプレイに5.4のように8つのコンテンツを同じサイズで2行4列に並べた。実験協力者はディスプレイから1.5メートル程離れて立ち、腕を振ってジェスチャを行う方式では利き手、スマートフォンのタッチディスプレイ上で指を滑らせるジェスチャを行う方式では利き手の逆でスマートフォンを持ち、利き手の人差し指でジェスチャを行った。

6.3 実験の手順

比較方式は腕を振ってジェスチャを行う方式とスマートフォンのタッチディスプレイ上で指を滑らせるジェスチャを行う方式の2つとする。

Step 1 実験協力者はデジタルサイネージに関する説明を受ける

Step 2 実験者がランダムに選択した方式を実験協力者へ提示する

Step 3 実験協力者は使用する方式の説明を受ける

Step 4 実験協力者は1分間の練習を行う

Step 5 実験協力者はディスプレイに表示された8種類のコンテンツにそれぞれ3回ずつ（計24回）ジェスチャによる選択を行う

Step 6 方式を入れ替えてStep 2からStep 5を行う

Step 5の時、選択時にマッチングが失敗した場合でもやり直しは行わず、各コンテンツの選択回数は3回ずつにした。

順序効果の相殺のために、8種類のうち選択するコンテンツの順番は重複がないようにランダムに設定した。

6.4 評価指標・測定方法

本実験では、H1の仮説検証のために、両方式の精度の検証として、ジェスチャマッチング結果を用いる。

H2を検証するための評価指標における操作負担については、三宅ら[24]によって日本語に翻訳されたNASA-TLX[23]を用いた。

- 1 どの程度の知的・知覚的活動(考える、決める、計算する、記憶する、見るなど)を必要としましたか。
- 2 仕事のペースや課題が発生する頻度のために感じる時間的切迫感はどの程度でしたか。
- 3 作業指示者(またはあなた自身)によって設定された課題の目標をどの程度達成できたと思いますか。目標の達成に関して自分の作業成績にどの程度満足していますか
- 4 作業成績のレベルを達成・維持するために、精神的・身体的にどの程度いっしょくじめいに作業しなければなりませんでしたか
- 5 作業中に、不安感、落胆、いろいろ、ストレス、悩みをどの程度感じましたか。

これは、タスクの負荷を評価するために使用される主観的な測定ツールである。タスクを遂行する際の認知的・身体的・精神的負荷を測定する。

6.5 実験の結果

ジェスチャマッチング成功率を、図6.1に示す。縦軸はマッチング成功率、横軸は実験協力者IDを表している。各方式におけるジェスチャマッチング成功率は、24回行ったジェスチャで何%マッチングが成功したかを表している。ジェスチャマッチング成功率は、方式1で62.5%、方式2で99.5%となった。

方式1より方式2の方がジェスチャマッチング成功率が高いことから、H1が支持された。

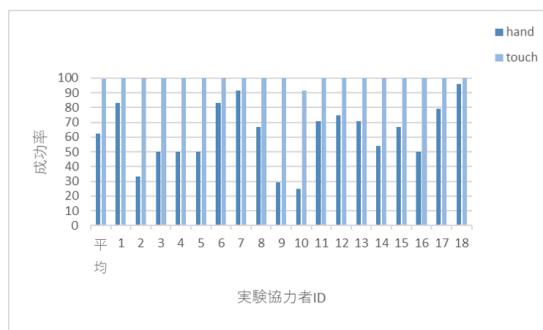


図 6.1: 実験協力者別ジェスチャマッチング成功率

NASA-TLX スコアの結果を、図「TLX スコア結果」に示す。縦軸は NASA-TLX における評価項目、横軸は平均スコアを表している。方式1のスコアが 50.97、方式2のスコアが 32.50 となった。これらのスコアを有意水準 5% で t 検定を行ったところ、有意差があった。このことから、H2 が支持された。

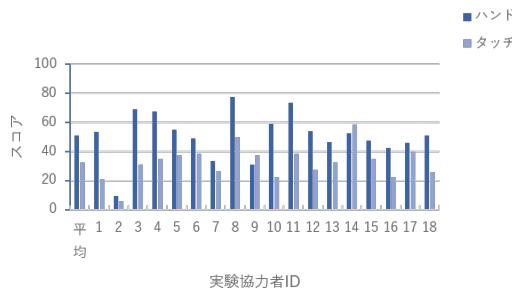


図 6.2: 方式別 NASA-TLX スコア

6.6 考察

本節では、実験結果に基づいた考察を述べる。

6.6.1 項では、ジェスチャーマッチング精度についての考察を述べる。

6.6.2 項では、操作負担についての考察を述べる

6.6.1 ジェスチャーマッチング精度の考察

方式1と方式2のジェスチャーマッチング成功率を比較したとき、方式1の方が方式2に比べて 37% 低いという結果が出た。また、方式1の成功率に満たしていない実験協力者について、成功率を前半と後半で分けて比較したとき、前半の成功率が 36.9% であるのに対して、後半の成功率が 50% と、13.1% の差が生じている。このことから、方式1は方式2に比べて習得に時間がかかると考えられる。

6.6.2 操作負担の考察

方式1と方式2の NASA-TLX のスコアを比較したとき、18.47 の差が出た。この結果は、腕を振る動作で身体的負担が生じてしまった点が大きいと考えられる、また、ジェスチャーマッチング成功率も 37% も差が出てしまっている。NASA-TLX の項目にある、達成感の項目において、方式1では 52.5/100、方式2では 87.78/100 と、35.28 もの差が出た。また、ジェスチャーマッチング成功率と達成感の項目を比較した時、ジェスチャーマッチング成功率が平均より低い実験協力者の達成感のスコアが平均よりも低く回答されていること

が分かった。さらに、NASA-TLX にあるストレスの項目において、方式1では49.72/100、方式2では8.05/100と、41.67 もの差が出た。このことから、ジェスチャマッチング成功率を向上させることによって、操作時のストレス、達成感の問題を解決でき、操作負担が軽減されると考えられる。

第7章 性能向上の検討

本章では、ジェスチャマッチング精度に問題があった方式1の性能向上の検討について述べる。

7.1 性能向上のためアプローチ

本節では、現在性能向上のために行っている実装について述べる。システムの全体像は5.1で述べたものと変更はしない。本稿では、性能向上のために、現在ジェスチャのマッチングに用いているユークリッド距離から、ディープラーニングを用いた方式を検討している。現在用いているマッチング方式の場合、スマートフォンで取得した加速度データを時系列データとして送信し、サーバで送信された時系列データと正解データの類似度をユークリッド距離で算出して、マッチングしたコンテンツのURLをスマートフォンへ返している。

現在のマッチング方法では、スマートフォンを傾けることによって生じる軸のブレに対応できていないという問題が生じている。また、腕の動かし方には人によって異なるため、腕を動かすという選択方法に得意不得意が生じていしまう。

これらの問題を解決するために、現在ディープラーニングを用いたジェスチャのマッチング方法を検討している。

7.2 実装

本節では、現在行っている実装及び、その予定について述べる。

7.3 データセットの作成

本項では、作成するデータセットについて述べる。

本研究で用いているデジタルサイネージに表示されているコンテンツは5.4節で述べた通り、8つである。1つのコンテンツに1つの動作パターンが結び付けられているため、今回はこの8パターンを対象として実装を進めていく。

訓練データとして、1パターン当たり20件×3人分の合計60件を想定している。検証用データとして、検証用データとして、1パターン当たり3件×3人の合計9件を想定している。テストデータとして、本稿で行った比較実験で取得したデータを用いる。よって、1パターン当たり3件×18名の合計54件を想定している。また、それぞれのデータは正規化せず、生データを用いることを想定している。今回の実装では、教師ありデータのみを用いる。

7.4 モデル構築

本項では、実際に行う予定のモデル構築について述べる。今回の実装では、一般データ分類ニューラルネットワークを用いる。主に行う処理としては、前処理、訓練、推定とする。まず、前処理として、訓練に不要なデータを取り除き、ディープラーニングに必要なデータだけを含むフォーマットに変換する。しかし、ここで含まれているデータは生データであるため、高精度なモデル構築を行うために、ノイズ除去、正規化を行う。前処理の最後に、これらの処理が正常に行えているかの確認を行うために、可視化を行う。

次に訓練として、推定器の構築を行う。今回は、エンコーダを用いず、前処理で正規化された教師ありデータで推定器の構築を行う。

最後に訓練した推定器を用いて、前処理で正規化されたテストデータに対して推定を行う。また、推定処理後、推定結果と推定精度集計結果を出力し、推定器の性能を確認できるようにする。

7.5 今後の予定

本項では、今後の実装の予定について述べる。

現在は、データセットの作成に取り掛かっている。今後は、作成したデータセットを元に、ハイパーパラメータの調整を行い、方式1のジェスチャマッチング精度の向上を目指す。

第8章 結論

本稿ではスマートフォンとデジタルサイネージを連携させる方として、アイコンの動きを真似るデジタルサイネージ選択方式における、腕を振ってジェスチャを行う方式と、スマートフォンのタッチディスプレイ上で指を滑らせてジェスチャを行う方式の比較を行うための実験を行った。結果として、腕を振ってジェスチャを行う方式は、スマートフォンのタッチディスプレイ上で指を滑らせてジェスチャを行う方式に比べて、精度が低く、操作負担が大きいということが分かった。

今後は、7章で述べた性能向上のための実装を進め、精度を対等にし、より正確に比較実験を行えるようにした上で、操作負担以外の項目に着目していく。また、今回は腕を振ってジェスチャを行う方式はスマートフォンのタッチディスプレイ上で指を滑らせる方に劣っている結果が出てしまったが、それぞれの長所・短所に加えて、適した利用シーンも調査していく。

謝辞

本研究は、ソフトバンク株式会社との共同研究の成果である。

参考文献

- [1] 宮田章裕. 真似て選択するデジタルサイネージ. マルチメディア, 分散協調とモバイルシンポジウム 2016 論文集, Vol. 2016, pp. 1350–1356, 2016.
- [2] 須賀美月, 松井優季, 新山はるな, 呉健朗, 森岡優一, 古野雅人, 宮田章裕. デジタルサイネージ向けのアイコン動作を真似る選択方式の検証. マルチメディア, 分散, 協調とモバイルシンポジウム 2023 論文集, Vol. 2023, pp. 1110–1114, 2023.
- [3] Mizuki Suga, Yuuki Matsui, Haruna Niyyama, Kenro Go, Masato Furuno, and Akio Miyata. Study on a content selection method by mimicking icon movements for digital signage. *Poster Proc. The 29th International Conference on Collaboration Technologies and Social Computing (CollabTech '23)*, pp. 25–30, 8 2023.
- [4] 坂内武希, 須賀美月, 土岐田力輝, 呉健朗, 古野雅人, 市川裕介, 宮田章裕ほか. デジタルサイネージ上のコンテンツ選択のための真似るジェスチャ入力方式の比較. マルチメディア, 分散, 協調とモバイルシンポジウム 2024 論文集, Vol. 2024, pp. 1573–1578, 2024.
- [5] Robert Hardy, Enrico Rukzio, Matthias Wagner, and Massimo Paolucci. Exploring expressive nfc-based mobile phone interaction with large dynamic displays. In *2009 First International Workshop on Near Field Communication*, pp. 36–41, 2009.
- [6] Sebastian Boring, Manuela Altendorfer, Gregor Broll, Otmar Hilliges, and Andreas Butz. Shoot & copy: Phonecam-based information transfer from public displays onto mobile phones. In *Proc. the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology (Mobility '07)*, pp. 24–31, 2007.
- [7] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: Mobile interaction through video. In *Proc. the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, p. 2287–2296, 2010.
- [8] Alireza Sahami Shirazi, Christian Winkler, and Albrecht Schmidt. Flashlight interaction: A study on mobile phone interaction techniques with large displays. In *Proc. the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09)*, 2009.

- [9] Dominik Schmidt, Fadi Chehimi, Enrico Rukzio, and Hans Gellersen. Phonetouch: A technique for direct phone interaction on surfaces. In *Proc. the 23rd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*, pp. 13–16, 2010.
- [10] Hannah Limerick, Richard Hayden, David Beattie, Orestis Georgiou, and Jörg Müller. User engagement for mid-air haptic interactions with digital signage. In *Proceedings of the 8th ACM international symposium on pervasive displays*, pp. 1–7, 2019.
- [11] Orestis Georgiou, Hannah Limerick, Loïc Corenthy, Mark Perry, Mykola Maksymenko, Sam Frish, Jörg Müller, Myroslav Bachynskyi, and Jin Ryong Kim. Mid-air haptic interfaces for interactive digital signage and kiosks. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2019.
- [12] Yuta Abe, Taishi Iriyama, Takashi Komuro, Kohei Shimasaki, and Idaku Ishii. A mid-air multi-touch interface using an ultrafast pan-tilt camera. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2024.
- [13] Zimo Liao, Zhicheng Luo, Qianyi Huang, Linfeng Zhang, Fan Wu, Qian Zhang, and Guihai Chen. Gesture recognition using visible light on mobile devices. *IEEE/ACM Transactions on Networking*, 2024.
- [14] Tsubasa Otaki and Hiroyuki Manabe. A nail-tip device for gesture and force input. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction*, pp. 1–2, 2024.
- [15] Tao Ni, Doug A Bowman, Chris North, and Ryan P McMahan. Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. *International Journal of Human-Computer Studies*, Vol. 69, No. 9, pp. 551–562, 2011.
- [16] Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 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)*, pp. 1229–1238, 2012.
- [17] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. Room-scale hand gesture recognition using smart speakers. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, pp. 462–475, 2022.
- [18] Stephane Viollet, Chauvet Martin, and Ingargiola Jean-Marc. Linled: Low latency and accurate contactless gesture interaction. In *Companion Publication of the 25th International Conference on Multimodal Interaction*, pp. 61–65, 2023.

- [19] Kieran Waugh, Mark McGill, and Euan Freeman. Push or pinch? exploring slider control gestures for touchless user interfaces. In *Nordic Human-Computer Interaction Conference*, pp. 1–10, 2022.
- [20] Keyue Yan, Chi Fai Lam, Simon Fong, João Alexandre Lobo Marques, Qun Song, and Huafeng Qin. Enhancing dynamic hand gesture recognition through optimized feature selection using double machine learning. In *Proceedings of the 2024 8th International Conference on Big Data and Internet of Things*, pp. 19–25, 2024.
- [21] Alessandro Carfi and Fulvio Mastrogiovanni. Gesture-based human–machine interaction: Taxonomy, problem definition, and analysis. *IEEE Transactions on Cybernetics*, Vol. 53, No. 1, pp. 497–513, 2021.
- [22] 坂内武希, 須賀美月, 土岐田力輝, 呉健朗, 古野雅人, 市川裕介, 宮田章裕. 真似て選択するデジタルサイネージにおける適切なジェスチャ方式の調査. 情報処理学会インタラクション2024論文集, pp. 662–666, 2024.
- [23] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, Vol. 52, pp. 139–183. Elsevier, 1988.
- [24] 三宅晋司, 神代雅晴. メンタルワークロードの主観的評価法 nasa-tlx と swat の紹介および簡便法の提案. 人間工学, Vol. 29, No. 6, pp. 399–408, 1993.

付録

研究業績

研究会・シンポジウム

- (1) 坂内武希, 須賀美月, 土岐田力輝, 呉健朗, 古野雅人, 市川裕介, 宮田章裕: 真似て選択するデジタルサイネージにおける適切なジェスチャ方式の調査, 情報処理学会インタラクション 2024 論文集 pp.662-666 (2024 年 3 月).
- (2) 坂内武希, 須賀美月, 土岐田力輝, 呉健朗, 古野雅人, 市川裕介, 宮田章裕: デジタルサイネージ上のコンテンツ選択のための真似るジェスチャ入力方式の比較, 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル 2024 論文集, Vol.2024, pp.1573-1578 (2024 年 6 月).

受賞

- (1) 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル (DICOMO2024) ヤングリサーチャー賞, デジタルサイネージ上のコンテンツ選択のための真似るジェスチャ入力方式の比較, 受賞者: 坂内武希 (2024 年 6 月)