

筑波大学大学院博士課程  
理工情報生命学院  
システム情報工学研究群修士論文

スマートフォンの  
背面カメラを用いて認識可能な  
空中人差し指ジェスチャによる片手操作手法

清 佑輔  
修士（工学）  
（情報理工学位プログラム）

指導教員 志築 文太郎

2022年3月

## 概要

タッチスクリーンを備えたスマートフォンへの操作は、主にタッチジェスチャにて行われる。ただし、片手にてスマートフォンを持ち、そのスマートフォンを持った手にて操作する片手使用の際には、マルチタッチジェスチャが難しいため、スマートフォンを持った手と異なる手にて操作する際に比べて容易に実行可能な操作の種類が少ない。スマートフォンの片手使用時における操作性の向上のためには、片手にて実行可能な操作の種類を拡張する必要がある。本研究では、片手使用時の操作の種類を拡張するために、スマートフォン背面後方の空中にて実行される人差し指によるジェスチャ（空中人差し指ジェスチャ）を操作に用いる。スマートフォン背面後方の空中を操作に用いるため、既存のタッチジェスチャによる操作、タッチジェスチャを拡張した操作、および背面上における操作と共存できる操作である。この操作の実現のため、7種類の空中人差し指ジェスチャを設計した。さらに、背面カメラ画像から取得される指の動作に関する情報を用いて、空中人差し指ジェスチャを認識するシステムを2種類試作した。まず、1つ目のシステムとして、ジェスチャ中に取得された画像内にて人差し指の領域が移動することに基づいて、閾値にて空中人差し指ジェスチャを認識するシステムを試作した。この試作システムの認識精度の評価を行ったところ、2種類の空中人差し指ジェスチャの認識精度は66.2%であった。さらに、2つ目のシステムとして、深層学習モデルを使用して空中人差し指ジェスチャを認識するシステムを試作した。この試作システムの認識精度の評価を行ったところ、7種類の空中人差し指ジェスチャの認識精度は、5分割交差検証にて平均79.9%、leave-one-participant-out 交差検証にて平均66.7%であった。これに対して、ジェスチャを4種類に限定したところ、認識精度は、5分割交差検証にて平均87.9%、leave-one-participant-out 交差検証にて平均78.3%であった。

# 目次

|            |                                   |           |
|------------|-----------------------------------|-----------|
| <b>第1章</b> | <b>はじめに</b>                       | <b>1</b>  |
| 1.1        | スマートフォンへの操作および片手使用における課題          | 1         |
| 1.2        | 先行研究におけるアプローチ                     | 2         |
| 1.3        | 目的およびアプローチ                        | 2         |
| 1.4        | 貢献                                | 3         |
| 1.5        | 本論文の構成                            | 4         |
| <b>第2章</b> | <b>関連研究</b>                       | <b>5</b>  |
| 2.1        | タッチジェスチャを拡張した操作                   | 5         |
| 2.2        | スマートフォン背面上における操作                  | 6         |
| 2.3        | 空中におけるジェスチャ操作                     | 7         |
| <b>第3章</b> | <b>空中人差し指ジェスチャによる片手操作手法</b>       | <b>8</b>  |
| 3.1        | 空中人差し指ジェスチャの概要                    | 8         |
| 3.2        | 空中人差し指ジェスチャの設計                    | 8         |
| 3.3        | 利用例                               | 11        |
| 3.3.1      | 通知の表示および端末設定の変更のショートカット           | 11        |
| 3.3.2      | アプリケーションごとのショートカットおよびグローバルショートカット | 12        |
| <b>第4章</b> | <b>試作システム1：閾値を用いるシステム</b>         | <b>14</b> |
| 4.1        | 実装環境                              | 14        |
| 4.2        | 実装                                | 14        |
| 4.2.1      | 人差し指領域の抽出                         | 15        |
| 4.2.2      | ジェスチャの認識                          | 15        |
| <b>第5章</b> | <b>試作システム1の認識精度評価</b>             | <b>18</b> |
| 5.1        | 参加者                               | 18        |
| 5.2        | 実験環境                              | 18        |
| 5.3        | タスクおよび実験手順                        | 20        |
| 5.4        | 結果                                | 20        |
| 5.5        | 考察                                | 21        |
| <b>第6章</b> | <b>試作システム2：深層学習モデルを用いるシステム</b>    | <b>22</b> |

|              |  |           |
|--------------|--|-----------|
| 6.1          | 実装環境 . . . . .                                 | 22        |
| 6.2          | 実装 . . . . .                                   | 22        |
| 6.2.1        | 事前動作の認識 . . . . .                              | 23        |
| 6.2.2        | 空中人差し指ジェスチャの認識 . . . . .                       | 23        |
|              | MHI の生成 . . . . .                              | 23        |
|              | 深層学習モデルを用いたジェスチャ認識 . . . . .                   | 26        |
| <b>第 7 章</b> | <b>試作システム 2 の認識精度評価</b>                        | <b>28</b> |
| 7.1          | 参加者 . . . . .                                  | 28        |
| 7.2          | 実験環境 . . . . .                                 | 28        |
| 7.3          | タスクおよび実験手順 . . . . .                           | 28        |
| 7.4          | 画像データ . . . . .                                | 32        |
| 7.5          | 結果 1 : 7 種類の空中人差し指ジェスチャの認識精度 . . . . .         | 33        |
|              | 7.5.1 5 分割交差検証 . . . . .                       | 33        |
|              | 7.5.2 leave-one-participant-out 交差検証 . . . . . | 34        |
| 7.6          | 結果 1 に関する考察 . . . . .                          | 34        |
| 7.7          | 試作システム 2 の改良の展望 . . . . .                      | 36        |
| 7.8          | 結果 2 : 4 種類の空中人差し指ジェスチャの認識精度 . . . . .         | 36        |
|              | 7.8.1 5 分割交差検証 . . . . .                       | 37        |
|              | 7.8.2 leave-one-participant-out 交差検証 . . . . . | 37        |
| 7.9          | 結果 2 に関する考察 . . . . .                          | 38        |
| 7.10         | 実験における制限 . . . . .                             | 39        |
| <b>第 8 章</b> | <b>議論</b>                                      | <b>41</b> |
| 8.1          | 空中人差し指ジェスチャを認識するシステムの改良 . . . . .              | 41        |
| 8.2          | 空中人差し指ジェスチャの難易度 . . . . .                      | 42        |
| 8.3          | 追加で実施すべき調査 . . . . .                           | 42        |
| <b>第 9 章</b> | <b>おわりに</b>                                    | <b>44</b> |
|              | 謝辞 . . . . .                                   | 45        |
|              | 参考文献 . . . . .                                 | 46        |
| <b>付録 A</b>  | <b>混同行列</b>                                    | <b>55</b> |
| A.1          | 結果 1 : 7 種類の空中人差し指ジェスチャの認識精度 . . . . .         | 55        |
|              | A.1.1 5 分割交差検証 . . . . .                       | 55        |
|              | A.1.2 leave-one-participant-out 交差検証 . . . . . | 58        |
| A.2          | 結果 2 : 4 種類の空中人差し指ジェスチャの認識精度 . . . . .         | 62        |
|              | A.2.1 5 分割交差検証 . . . . .                       | 62        |

|       |  |    |
|-------|--|----|
| A.2.2 | leave-one-participant-out 交差検証 . . . . . | 64 |
|-------|--|----|

# 目次

|     |   |    |
|-----|---|----|
| 1.1 | 空中人差し指ジェスチャによる片手操作の概観. . . . .  | 2  |
| 1.2 | 操作に用いる領域. a: 既存のタッチジェスチャによる操作およびタッチジェスチャを拡張した操作に用いる領域, b: 背面上における操作に用いる領域, c: 空中人差し指ジェスチャによる操作に用いる領域. . . . .   | 3  |
| 1.3 | 空中人差し指ジェスチャの認識の概観. . . . .  | 4  |
| 3.1 | 空中人差し指ジェスチャによる操作の流れ. 本研究においては, 背面カメラの前にて実行される空中人差し指ジェスチャを操作に用いる対象とする. 背面カメラに映るように空中人差し指ジェスチャを実行するために, ユーザは背面カメラに向けて人差し指を伸ばす事前動作を行う必要がある. . . . .  | 9  |
| 3.2 | 7種類の空中人差し指ジェスチャ. . . . .  | 10 |
| 3.3 | 空中人差し指ジェスチャによる通知の表示および端末設定の変更のショートカット. . . . .  | 11 |
| 3.4 | 地図アプリケーションのショートカットおよびグローバルショートカット. . . . .  | 13 |
| 4.1 | Samsung Galaxy S7 edge の背面カメラにて撮影されたジェスチャ実行時の人差し指の軌跡. a: Bend, b: Swipe. . . . .  | 15 |
| 4.2 | ジェスチャを実行した際に得られる2値画像. a: Bend, b: Swipe. 試作システム1では, 空中人差し指ジェスチャを認識するために, ジェスチャ実行前に, 人差し指領域にフィッティングした直線 $L$ および重心 $C$ , ジェスチャ実行中に, 重心 $C'$ を取得する. さらに, 有向線分 $CC'$ にて表されるベクトル $V$ および $L$ のなす角の内, 小さい方の角度 $\theta$ をジェスチャの認識に用いる. . . . . | 16 |
| 5.1 | 4種類の照明条件下における試作システム1を用いた実験の様子. a: 屋内点灯, b: 屋内消灯, c: 屋外日向, d: 屋外日陰. . . . .  | 19 |
| 6.1 | 事前動作の認識の流れ. . . . .   | 24 |
| 6.2 | 空中人差し指ジェスチャの認識の流れ. a: 取得した35枚の画像に対して, 人差し指領域の抽出およびグレイスケール化の処理を行う. b: Motion History Image を作製する. c: CNN モデルを用いて, Motion History Image から空中人差し指ジェスチャを認識する. . . . .   | 25 |
| 6.3 | 試作システム2に用いるCNNモデルの構成. . . . .   | 27 |

|     |   |    |
|-----|---|----|
| 7.1 | 試作システム2を用いた実験の様子. . . . .   | 29 |
| 7.2 | 試作システム2を用いた実験にて, 参加者が行ったタスク1. タスク1は, 空中人差し指ジェスチャを実行するタスクである. 参加者は, はじめの1秒にて人差し指を背面カメラの前に伸ばす空中人差し指ジェスチャの事前動作を行い, その後の1秒にて空中人差し指ジェスチャを実行するように指示された. このタスクにて, 空中人差し指ジェスチャ実行中の画像を取得した. . . . .                      | 30 |
| 7.3 | 試作システム2を用いた実験にて, 参加者が行ったタスク2. タスク1は, 空中人差し指ジェスチャを実行するタスクである. 参加者は, 2秒間継続してスマートフォンを自由に持ち替えるように指示された. なお, 人差し指を背面カメラの前に伸ばす空中人差し指ジェスチャの事前動作および空中人差し指ジェスチャを行わないように注意を受けた. このタスクにて, 空中人差し指ジェスチャ未実行時の画像を取得した. . . . . | 31 |
| 7.4 | 空中人差し指ジェスチャを実行中の44枚の画像からMHIを生成する流れ. 時系列順の44枚の画像について, 35枚をウィンドウサイズとするスライディングウィンドウにて分割した後, ウィンドウ毎にMHIを生成した. . . . .   | 32 |
| 7.5 | 試作システム2および改善案におけるBend2の認識の流れ. . . . .   | 36 |
| 7.6 | 実験にてSwipeが実行された際に生成されたMHI. a: カメラの画角内にて実行された場合. b: カメラの画角の端にて実行された場合. . . . .   | 39 |

# 表目次

|     |   |    |
|-----|---|----|
| 3.1 | 空中人差し指ジェスチャを通知の表示および端末設定の変更のショートカットに割り当てる例. . . . .                                     | 12 |
| 3.2 | 空中人差し指ジェスチャをアプリケーションごとのショートカットおよびグローバルショートカットに割り当てる例. . . . .                           | 13 |
| 5.1 | 試作システム1を用いた実験における参加者の手の大きさおよび人差し指の長さ. . . . .   | 18 |
| 5.2 | 4種類の照明条件下における照度. . . . .  | 19 |
| 5.3 | 取得した全試行と認識結果に基づく混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .   | 20 |
| 5.4 | 4種類の照明条件下にて取得した試行と認識結果に基づく混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                               | 20 |
| 5.5 | 参加者毎の認識精度. . . . .  | 21 |
| 7.1 | 試作システム2を用いた実験における参加者の手の大きさおよび人差し指の長さ. . . . .   | 30 |
| 7.2 | 5分割交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 33 |
| 7.3 | leave-one-participant-out 交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                     | 33 |
| 7.4 | leave-one-participant-out 交差検証にて, 各参加者のデータを検証データとした際の認識精度. . . . .                      | 35 |
| 7.5 | 4種類の空中人差し指ジェスチャに限定した5分割交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                        | 37 |
| 7.6 | 4種類の空中人差し指ジェスチャに限定したleave-one-participant-out 交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す. . . . . | 37 |
| 7.7 | 4種類の空中人差し指ジェスチャに限定したleave-one-participant-out 交差検証にて, 各参加者のデータを検証データとした際の認識精度. . . . .  | 38 |
| A.1 | 5分割交差検証にて得られた, 分割1を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                            | 55 |
| A.2 | 5分割交差検証にて得られた, 分割2を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                            | 56 |



|      |  |    |
|------|--|----|
| A.3  | 5 分割交差検証にて得られた, 分割 3 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                          | 56 |
| A.4  | 5 分割交差検証にて得られた, 分割 4 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                          | 56 |
| A.5  | 5 分割交差検証にて得られた, 分割 5 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .                          | 57 |
| A.6  | leave-one-participant-out 交差検証にて得られた, P1 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 58 |
| A.7  | leave-one-participant-out 交差検証にて得られた, P2 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 58 |
| A.8  | leave-one-participant-out 交差検証にて得られた, P3 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 59 |
| A.9  | leave-one-participant-out 交差検証にて得られた, P4 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 59 |
| A.10 | leave-one-participant-out 交差検証にて得られた, P5 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 59 |
| A.11 | leave-one-participant-out 交差検証にて得られた, P6 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 60 |
| A.12 | leave-one-participant-out 交差検証にて得られた, P7 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 60 |
| A.13 | leave-one-participant-out 交差検証にて得られた, P8 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 60 |
| A.14 | leave-one-participant-out 交差検証にて得られた, P9 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 61 |
| A.15 | leave-one-participant-out 交差検証にて得られた, P10 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . . | 61 |
| A.16 | 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 1 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .    | 62 |
| A.17 | 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 2 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .    | 62 |
| A.18 | 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 3 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .    | 63 |
| A.19 | 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 4 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .    | 63 |
| A.20 | 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 5 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .    | 63 |

|  |    |
|--|----|
| A.21 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P1 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 64 |
| A.22 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P2 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 64 |
| A.23 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P3 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 64 |
| A.24 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P4 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 65 |
| A.25 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P5 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 65 |
| A.26 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P6 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 65 |
| A.27 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P7 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 65 |
| A.28 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P8 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 66 |
| A.29 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P9 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . .  | 66 |
| A.30 4種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P10 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す. . . . . | 66 |

# 第1章 はじめに

スマートフォンは、近年最も広く普及している情報通信端末である [総務 21]。ユーザは、通話およびメールの送受信など、従来までの携帯端末の用途に加えて、ソーシャルネットワーキングサービス (SNS)、Web ブラウジング、およびゲームなどの用途にスマートフォンを利用できる。このように様々な用途に利用されること、さらに、パーソナルコンピュータおよびタブレット端末に比べて携帯性に優れることから、ユーザがスマートフォンを使用する場面は多岐に渡る。例えば、椅子に座っている時またはベッドに横たわっている時などの日常の場面、および歩行時または電車に乗っている時などの移動の場面にて、ユーザはスマートフォンを利用する。

このようなスマートフォンについて、ユーザは片手にてスマートフォンを持ち、そのスマートフォンを持った手にて操作する片手使用を好むことが報告されている [KBCV06, YYJ15, LMBH18]。これは、荷物を持っている時、電車にて吊革を握っている時、飲食の時、およびメモを取る時など、スマートフォンへの操作以外にもう片方の手を使いたい場面があるためである。本論文においては、このスマートフォンの片手使用における課題に焦点を当てる。また、その課題を解決するための手法を示す。

本章では、スマートフォンへの操作および片手使用における課題、先行研究におけるアプローチ、本研究における目的およびアプローチ、貢献、および本論文の構成を示す。

## 1.1 スマートフォンへの操作および片手使用における課題

タッチスクリーンを備えたスマートフォンへの操作は、主にタップおよびスワイプなどのタッチジェスチャにて行われる。また、複数本の指によるタッチを検知できるマルチタッチスクリーンを搭載したスマートフォンに対して、ユーザはマルチタッチジェスチャを用いて操作できる。例えば、ピンチアウトおよびピンチインは、文書、画像および地図などのコンテンツを閲覧する際に、拡大および縮小の操作として実行される。

スマートフォンの片手使用においては、主に親指のタッチジェスチャが用いられる。これは、親指以外の指が背面からスマートフォンを支えるような持ち方 (以降、把持姿勢) になるためである。ただし、片手使用時にマルチタッチジェスチャを行う場合、背面にてスマートフォンを支える指を前面に移動させる必要があるため、不安定な把持姿勢となる [BLC<sup>+</sup>12]。よって、スマートフォンを持った手と異なる手にて操作する際、つまり両手使用時に比べて、容易に実行可能なタッチジェスチャの種類は少ない。このため、スマートフォンの片手使用時における操作性の向上のためには、片手にて実行可能な操作の種類を拡張する必要がある。

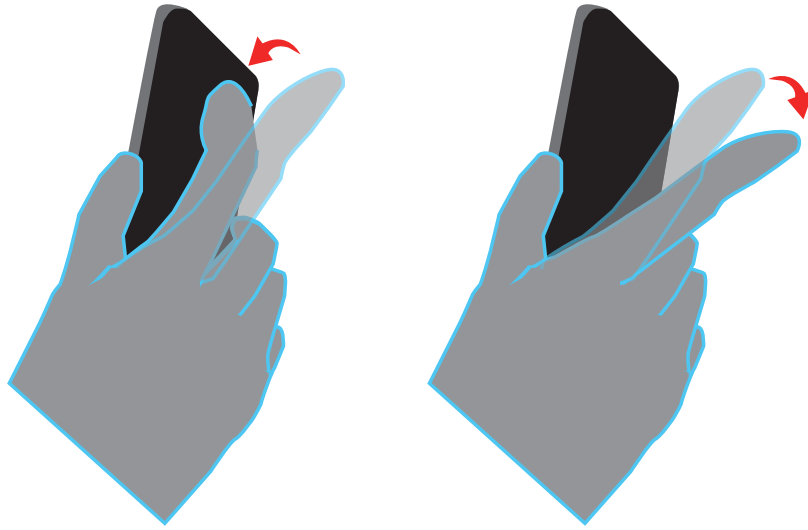


図 1.1: 空中人差し指ジェスチャによる片手操作の概観.

## 1.2 先行研究におけるアプローチ

これまでに、スマートフォンにおける操作の種類を拡張する研究が行われている。そのアプローチの1つとして、タッチジェスチャを拡張する手法が検討されてきた。例えば、既存のタッチジェスチャのタッチ位置を用いた手法 [RT09,HGL14,HIST19]、タッチする指および手の特徴を用いた手法 [RLG09,BLC<sup>+</sup>12,LKB<sup>+</sup>18]、スマートフォン本体の傾きおよび動作を用いた手法 [HS11] があげられる。タッチジェスチャを拡張した操作は、タッチスクリーンおよびスマートフォン内蔵のセンサのみを用いて認識できるため、実装のコストが低いという特徴がある。

さらに、スマートフォン背面上における操作が研究されてきた。これまでに、スマートフォン内蔵のセンサ [XHW13,SD14,ZGZ<sup>+</sup>15,WFZ16,GL16,GL17,YMI18,SZWX18,HLN<sup>+</sup>21] または追加のセンサ [HFST15,HBH16,LMH18,XBZN21] を用いて操作を認識する手法が研究されている。これらの研究では、背面上をタッチジェスチャによる操作に用いることにより、片手かつスクリーンを遮らずに実行可能な操作を実現している。また、操作に用いる領域が背面であることから、操作によってスクリーン上のコンテンツを遮らない [WFB<sup>+</sup>07,BC09] ため、片手使用において有用な操作手法である。

## 1.3 目的およびアプローチ

本研究の目的は、片手使用時の操作の種類を拡張することである。そのために、スマートフォン背面後方の空中を操作の領域として用いることを検討している。本研究では、スマートフォン背面後方の空中にて実行される人差し指によるジェスチャ（以降、空中人差し指ジェスチャ）をスマートフォンへの操作に用いる（図 1.1）。空中人差し指ジェスチャによる操作

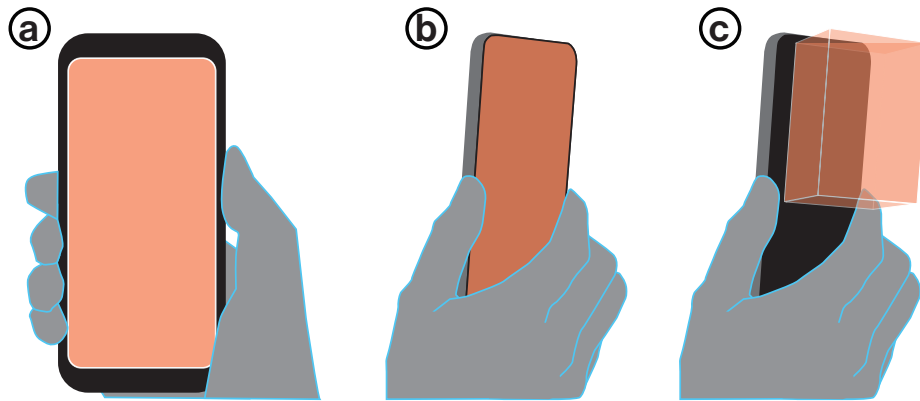


図 1.2: 操作に用いる領域. a: 既存のタッチジェスチャによる操作およびタッチジェスチャを拡張した操作に用いる領域, b: 背面上における操作に用いる領域, c: 空中人差し指ジェスチャによる操作に用いる領域.

に用いる領域は, 図 1.2 に示すように, スマートフォンのタッチジェスチャおよび背面上における操作と異なる. そのため, 既存のタッチジェスチャによる操作, タッチジェスチャを拡張した操作, および背面上における操作と共存できる. つまり, 既存の操作と同時にスマートフォンへ導入可能である.

この操作を実現するために, 本研究では, 背面カメラ画像から取得される指の動作に関する情報を用いて, 空中人差し指ジェスチャを認識するシステムを開発した (図 1.3). 背面カメラ画像を用いることにより, 背面カメラを備えた既製のスマートフォンに容易に導入可能なシステムを目指している. これまでに, ジェスチャ中に取得された画像内にて人差し指の領域が移動することに基づいて, 閾値にて空中人差し指ジェスチャを認識するシステムを試作した [SS21]. さらに, 畳み込みニューラルネットワーク (Convolutional Neural Network, 以降, CNN) を用いた深層学習モデルを使用して, 空中人差し指ジェスチャを認識するシステムを試作した [清 21, 清 22]. このモデルでは, 背面カメラ画像から作製された Motion History Image [BD01] (以降, MHI) を特徴量とする.

## 1.4 貢献

本研究の貢献は以下の通りである.

- 空中人差し指ジェスチャによる片手操作手法を考案した.
- 7種類の空中人差し指ジェスチャの設計およびそれらによる操作をショートカットに用いる利用例を示した.
- 空中人差し指ジェスチャを認識する試作システムとして閾値を用いるシステムおよび深層学習モデルを用いるシステムの実装を示した.

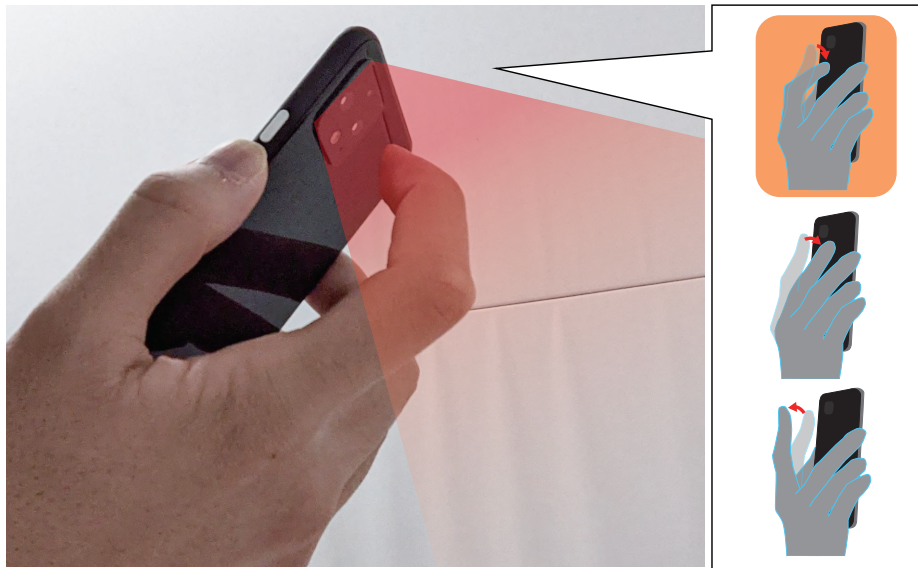


図 1.3: 空中人差し指ジェスチャの認識の概観.

- 空中人差し指ジェスチャの認識精度の評価を通じて，試作システムの課題および改良案を示した。

## 1.5 本論文の構成

本論文の構成を述べる．第 1 章にて，本論文の背景，本研究の目的およびアプローチを述べる．第 2 章にて，本研究に関連する研究および本研究の位置付けを述べる．第 3 章にて，空中人差し指ジェスチャによる片手操作手法およびその利用例を述べる．第 4 章および第 5 章にて，本研究にて試作した空中人差し指ジェスチャを認識するシステムの内，閾値を用いるシステムの実装および認識精度の評価を述べる．さらに，第 6 章および第 7 章にて，深層学習モデルを用いるシステムの実装および認識精度の評価を述べる．第 8 章にて，試作システムの改良点および空中人差し指ジェスチャの設計を議論する．最後の第 9 章は，本研究のまとめである．

## 第2章 関連研究

本研究では，スマートフォンにおいて片手にて実行可能な操作を拡張するために，空中人差し指ジェスチャを操作に用いる．また，操作の実現のために，背面カメラ画像を用いて空中人差し指ジェスチャを認識するシステムを開発している．本章では，まず，片手にて実行可能な操作として，タッチジェスチャを拡張した操作に関する研究を述べる．次に，スマートフォン背面上における操作に関する研究を述べる．さらに，スマートフォンのカメラを用いて認識される操作として，空中におけるジェスチャ操作に関する研究を述べる．

### 2.1 タッチジェスチャを拡張した操作

タッチスクリーン上にて実行されるタッチジェスチャを拡張することにより，新たな操作が実現されている．RothおよびTurner [RT09]は，ベゼルからのスワイプジェスチャである *Bezel Swipe* を提案した．Roudautら [RLG09]は，指をタッチスクリーン上にて転がすようなタッチジェスチャを示した．HinckleyおよびSong [HS11]は，スマートフォンの傾きおよび動作をタッチジェスチャに組み合わせる操作手法を提案した．HeoおよびLee [HL11]は，通常のタップおよび強いタップを区別することにより，操作の種類を拡張を試みた．Boringら [BLC<sup>+</sup>12]は，タッチスクリーン上の親指の接触面積を利用する操作手法を示した．この手法では，親指によるタッチ時に，その接触面積に応じて操作モードを切り替える．Heoら [HGL14]は，ダブルタップを拡張した操作手法である *Ta-tap* を提案した．ユーザは，1度目のタップ地点から離れた地点にて2度目のタップを行うことにより，マルチタッチジェスチャに代わる操作が可能である．Leら [LKB<sup>+</sup>18]は，手の平によるタッチを新しい操作として用いる手法を示した．この操作の認識は，指によるタッチと手の平によるタッチの差異に基づく．Hakkaら [HIST19]は，ターゲット内にて終点を持つスワイプジェスチャである，*Bounded Swipe* を提案した．

タッチジェスチャを拡張した操作は，既存のタッチジェスチャと同じくタッチスクリーン上にて実行される．一方で，本研究では，空中人差し指ジェスチャを操作に用いる．よって，操作領域が異なるため，タッチジェスチャを拡張した操作および本研究における操作は共存可能である．つまり，これらの操作は同時にスマートフォンへ導入可能である．

## 2.2 スマートフォン背面上における操作

これまでに、スマートフォン背面上における操作として、スマートフォン内蔵のセンサを用いて認識したジェスチャによる操作が研究されてきた。Xiao ら [XHW13] は、スマートフォンの背面カメラを覆うジェスチャおよび背面カメラ上における人差し指によるスワイプジェスチャによる操作手法を示した。Seipp および Devlin [SD14] は、人差し指、中指、および親指を用いてスマートフォン背面および側面を叩くジェスチャによる操作手法を示した。叩くジェスチャの認識にはスマートフォンに内蔵されたマイクおよびジャイロセンサが用いられる。Zhang ら [ZGZ<sup>+</sup>15] は、背面における人差し指のタップ、人差し指のスワイプ、親指のタップによる操作を示した。これらのジェスチャは、スマートフォン内蔵の加速度センサ、ジャイロセンサおよびマイクを用いて認識される。さらに、Zhang ら [ZGZ<sup>+</sup>16] はこれらのジェスチャの設計および認識精度を示している。Wong ら [WFZ16] は、スマートフォンのカメラと反射鏡を用いて認識される、背面における人差し指のタップおよびスワイプによる操作を提案した。この認識は、スマートフォン背面に取り付けた視覚的パターンに基づく。Granell と Leiva は、背面へのタップの認識に使用する特徴量について、特徴量エンジニアリングを行った [GL16]。その結果に基づいてタップを認識するシステムを実装し、さらに、タップによる操作の利用方法を示した [GL17]。Yamada ら [YMI18] は、スマートフォンの内蔵カメラ上における指の移動方向に基づく操作を示した。この操作の認識には 3D プリントされたリングが用いられるため、ユーザは物理的なフィードバックを得ながら方向入力ができる。Sun ら [SZWX18] は、スマートフォン内蔵のスピーカおよびマイクを用いて、背面上におけるジェスチャを音響信号にて認識する手法を示した。この手法により、背面上におけるスワイプ、タップ、およびスクロールが可能となる。Huang ら [HLN<sup>+</sup>21] は、深層学習モデルを用いて背面上におけるタップを認識する手法を示した。この手法により、タップの方向、指の接触部位、タップの位置に基づく操作を可能にした。提案された CNN モデルにおける入力は、スマートフォン内蔵の慣性計測センサから得られたデータおよびスマートフォンのフォームファクタである。

さらに、スマートフォン背面に取り付けた追加のセンサを用いて認識したジェスチャによる操作が研究されてきた。Hakoda ら [HFST15] は、スマートフォンの背面にフォトリフレクタを埋め込んだ穴を配置し、その穴を塞ぐジェスチャを認識した。Hidaka ら [HBH16] は、スマートフォンの背面に取り付けたフォトリフレクタを用いて、背面における上下のスワイプによる操作を認識した。この操作にてスクリーン上のコンテンツを下部に縮小させることにより、片手使用にて届きにくい画面上部への操作が容易になる。Le ら [LMH18] は、スマートフォンの背面全体に静電容量式のタッチセンサを取り付けることにより、タッチスクリーンを拡張したプロトタイプを示した。このプロトタイプにて、背面に対するタッチジェスチャによる操作を可能にした。Xiao ら [XBZN21] は、スマートフォンの背面上に追加したトラックパッドに対するタッチジェスチャによる操作を示した。このトラックパッドはタッチスクリーンから伸ばした電極を用いて実装される。

本研究では、空中人差し指ジェスチャを操作に用いる。この操作は、片手使用時の把持姿勢にて、背面に位置する人差し指により実行される。そのため、スマートフォン背面上における操作と同じく、操作によってスクリーン上のコンテンツを遮らないという特徴を持つ。ま



た、操作領域が異なるため、背面上における操作および本研究における操作は共存可能である。つまり、これらの操作は同時にスマートフォンへ導入可能である。

## 2.3 空中におけるジェスチャ操作

Baldauf ら [BZFR11] は、空中における指先によるポインティング操作を示した。操作の認識は、カメラ画像を用いた指先の位置の検出に基づく。さらに、Lasinger らにより、指先の検出手法の堅牢性および操作のパフォーマンスが評価されている [LLZ12]。Yang ら [YHBI13] は、スマートフォン周囲における手の空中ジェスチャによる操作を実現している。ジェスチャの認識には、360度撮影可能なレンズを取り付けたスマートフォンのカメラが用いられた。Song ら [SSP<sup>+</sup>14] は、スマートフォン内蔵のカメラを用いて、手のジェスチャを認識するアルゴリズムを提案した。さらに、このアルゴリズムを用いて既存のタッチジェスチャに空中ジェスチャを組み合わせるアプリケーションを示した。石井ら [石井18] は、スマートフォンを用いたバーチャルリアリティ環境向けのターゲット選択手法を提案している。この手法において、ポインタは背面カメラにて撮影した手の位置に応じて表示される。また、ユーザは手のジェスチャによりターゲットの選択が可能である。Yu ら [YWV<sup>+</sup>19] は、直角プリズムミラーを取り付けたフロントカメラを用いて、スマートフォンを持つ手およびタッチする手の状態の認識を可能にした。

これらはスマートフォンを把持している手と異なる手のジェスチャを操作に用いる。一方で、本研究では、スマートフォンを把持している手の人差し指によるジェスチャを操作に用いる。

本研究と同様に、空中にて実行される人差し指ジェスチャを操作に用いた手法として、Oh および Hong [OH13] の操作手法があげられる。この操作手法では、空中人差し指ジェスチャを3次元のUI操作に用いた。また、ジェスチャの認識は、背面カメラ画像を特徴量とする機械学習モデルを用いた指先の検出に基づく。一方で、本研究におけるジェスチャ認識システムでは、背面カメラ画像から取得される人差し指の動作を用いてジェスチャを認識する。

## 第3章 空中人差し指ジェスチャによる片手操作手法

本章にて、空中人差し指ジェスチャの概要、空中人差し指ジェスチャの設計およびその利用例を示す。

### 3.1 空中人差し指ジェスチャの概要

本研究では、空中人差し指ジェスチャ（図 1.1）をスマートフォンの片手使用時の操作に用いる。このジェスチャは、スマートフォンを持っている手の人差し指を、スマートフォンの背面後方にて動かすジェスチャである。

本研究においては、背面カメラの前にて実行される空中人差し指ジェスチャを操作に用いる対象とする。これは、背面カメラを用いて空中人差し指ジェスチャを認識するためである。よって、背面カメラに映るように空中人差し指ジェスチャを実行するために、ユーザは背面カメラに向けて人差し指を伸ばす事前動作を行う必要がある（図 3.1）。

### 3.2 空中人差し指ジェスチャの設計

本研究では、7種類の空中人差し指ジェスチャを設計した（図 3.2）。これらは、スマートフォンの背面カメラの前にて実行することを前提に、スマートフォンを把持しながら実行可能なジェスチャおよびそれを2回繰り返すジェスチャから構成されている。なお、今回は、著者がスマートフォンの背面後方にて実行可能なジェスチャからこれらが採用された。

7種類の空中人差し指ジェスチャについて、ユーザは人差し指を事前動作の後に次のように動かす。

- Stretch：伸ばしたまま保持する（図 3.2a）
- Bend：端末方向に曲げる（図 3.2b）
- Bend2：Bend を2回行う（図 3.2c）
- Swipe：中指方向に払う（図 3.2d）
- Swipe2：Swipe を2回行う（図 3.2e）

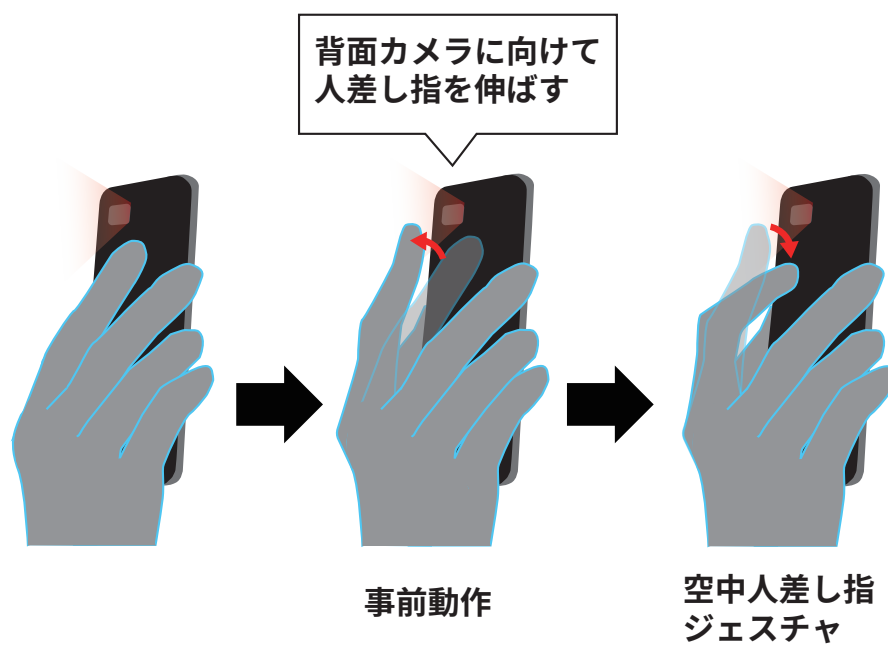


図 3.1: 空中人差し指ジェスチャによる操作の流れ. 本研究においては, 背面カメラの前にて実行される空中人差し指ジェスチャを操作に用いる対象とする. 背面カメラに映るように空中人差し指ジェスチャを実行するために, ユーザは背面カメラに向けて人差し指を伸ばす事前動作を行う必要がある.



図 3.2: 7 種類の空中人差し指ジェスチャ.

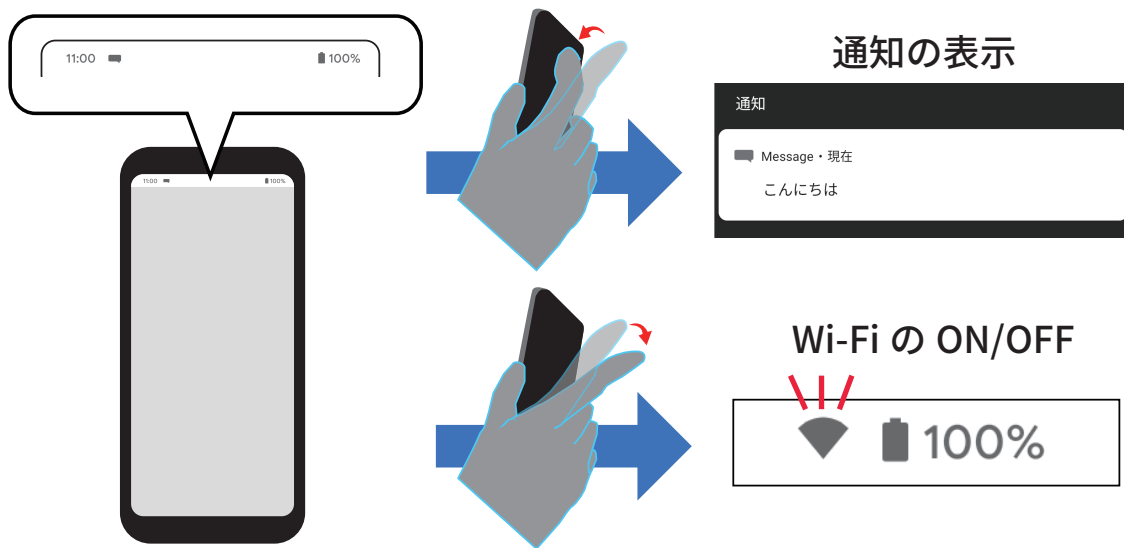


図 3.3: 空中人差し指ジェスチャによる通知の表示および端末設定の変更のショートカット.

- Release : 端末反対方向に伸ばす (図 3.2f)
- Release2 : Release を 2 回行う (図 3.2g)

### 3.3 利用例

本研究では、空中人差し指ジェスチャによる片手操作の利用方法として、ショートカットを想定している。本節では、空中人差し指ジェスチャによる操作の利用例として、通知の表示および端末設定の変更のショートカット、および、アプリケーションごとのショートカットおよびグローバルショートカットを述べる。

#### 3.3.1 通知の表示および端末設定の変更のショートカット

空中人差し指ジェスチャを用いた操作の利用例として、通知の表示および端末設定の変更のショートカットが挙げられる (図 3.3)。例えば、ユーザは、Bend を実行することにより通知を表示できる。さらに、Release を実行することにより Wi-Fi 接続機能の ON/OFF を切り替えられる。設計した 7 種類のジェスチャは、表 3.1 のようにショートカットを割り当てられる。ここで、Stretch は、ユーザが誤って事前動作を行った際にキャンセルする操作として用いられることを想定している。この割り当てにより、誤ってショートカットが実行されることを防ぐ狙いがある。よって、残りの 6 種類の空中人差し指ジェスチャを、通知の表示および端末設定の変更のショートカットに割り当てられる。

表 3.1: 空中人差し指ジェスチャを通知の表示および端末設定の変更のショートカットに割り当てる例.

|          |                   |
|----------|-------------------|
| Stretch  | ショートカットのキャンセル     |
| Bend     | 通知の表示             |
| Bend2    | 端末設定変更パネルの表示      |
| Swipe    | Bluetooth 接続の切り替え |
| Swipe2   | 位置情報サービスの切り替え     |
| Release  | Wi-Fi 接続の切り替え     |
| Release2 | 機内モードの切り替え        |

通知の表示および端末設定の変更は、多くのスマートフォンにて、画面上部のステータスバーに収められている機能である [Goo21c, App21b, Goo21a, App21a]. ただし、片手使用において、ステータスバーに親指を伸ばす操作は、不安定な把持姿勢を伴う。これは、スマートフォンの片手使用において、把持姿勢を変えることなくタッチジェスチャにて操作可能な領域が限られているため [BLO14, LMBH18] である。一方で、本ショートカットを用いることにより、ユーザは届きにくい画面上部に親指を伸ばすことなく、これらの機能を使用できる。

### 3.3.2 アプリケーションごとのショートカットおよびグローバルショートカット

空中人差し指ジェスチャを用いた操作の利用例として、アプリケーションごとのショートカットおよびグローバルショートカットが挙げられる (図 3.4). 地図アプリケーションにおいて、例えば、ユーザは、**Bend** を実行することにより地図を拡大できる。また、**Release** を実行することにより地図を縮小できる。さらに、グローバルショートカットとして、**Bend2** を実行することによりスクリーンショットを撮影できる。また、**Release2** を実行することにより、地図アプリケーションを閉じてホーム画面へ移動できる。このように、繰り返しが1回のジェスチャをアプリケーションごとのショートカット、繰り返しが2回のジェスチャをグローバルショートカットに使用できる。設計した7種類のジェスチャは、地図、Web ブラウザ、文書閲覧などのアプリケーション上にて、表 3.2 のようにショートカットを割り当てられる。なお、通知の表示および端末設定の変更のショートカットと同じく、**Stretch** は、ユーザが誤って事前動作を行った際にキャンセルする操作として用いられることを想定している。

空中人差し指ジェスチャは、スマートフォンの背面後方の空中にて実行される。よって、これらのショートカットは片手使用時に画面を遮ることなく利用可能である。また、画面の拡大および縮小の操作は、通常ピンチジェスチャに割り当てられている操作である。このピンチジェスチャは、人差し指をスマートフォンの背面から前面に移動させる把持姿勢の変更が必要となるため、片手使用時には実行が難しい [BLC<sup>+</sup>12]. 一方で、空中人差し指ジェスチャによる操作にショートカットを割り当てることにより、ユーザは片手使用時においても画面の拡大および縮小を容易に行える。



図 3.4: 地図アプリケーションのショートカットおよびグローバルショートカット。

表 3.2: 空中人差し指ジェスチャをアプリケーションごとのショートカットおよびグローバルショートカットに割り当てる例。

|          |               |
|----------|---------------|
| Stretch  | ショートカットのキャンセル |
| Bend     | 画面の拡大         |
| Bend2    | スクリーンショットの撮影  |
| Swipe    | 検索窓の表示        |
| Swipe2   | 画面の輝度の上昇      |
| Release  | 画面の縮小         |
| Release2 | ホーム画面へ移動      |

## 第4章 試作システム1：閾値を用いるシステム

空中人差し指ジェスチャによる操作を実現するために、ジェスチャを認識するシステムを試作した。後述する深層学習モデルを用いたジェスチャ認識システムとの区別のため、以降、本システムを試作システム1と呼称する。

本章では、試作システム1の実装環境および実装を述べる。

### 4.1 実装環境

本研究では、システムの試作の第1歩として、設計した7種類のジェスチャのうち、BendおよびSwipeの認識を検討した。よって、試作システム1は、スマートフォンの背面カメラにて撮影した人差し指の画像から、BendおよびSwipeを認識するシステムである。試作システム1の実装には、Samsung Galaxy S7 edge (150.8 mm × 72.6 mm × 8.3 mm, 5.5 インチ, Android 8) を用いた。このスマートフォンの背面カメラは、背面上部の中央（背面左端から35.4 mm, 上端から24.2 mm）に配置されている。また、画像処理ライブラリとしてOpenCV（バージョン4.3.0）を使用した。

### 4.2 実装

BendおよびSwipeを背面カメラにて撮影した際の人差し指の軌跡を図4.1に示す。BendおよびSwipeを認識するために、ジェスチャ実行中にスマートフォンの背面カメラにて撮影した画像における、人差し指の傾きおよび動きに着目した。Bendの場合、ジェスチャを実行する前の人差し指の傾きに沿って人差し指が移動する。一方で、Swipeの場合、ジェスチャを実行する前の人差し指の傾きに対して、垂直に近い方向へ人差し指が移動する。このことから、画像中の人差し指の傾きおよび人差し指領域の重心の移動を用いて、単純な閾値にてBendおよびSwipeを認識できると考えた。よって、試作システム1では、人差し指領域の抽出を経て、人差し指の傾きおよび人差し指領域の重心を算出し、これらを用いてBendおよびSwipeを認識する。

本節では、試作システム1における人差し指領域の抽出、およびジェスチャの認識を述べる。



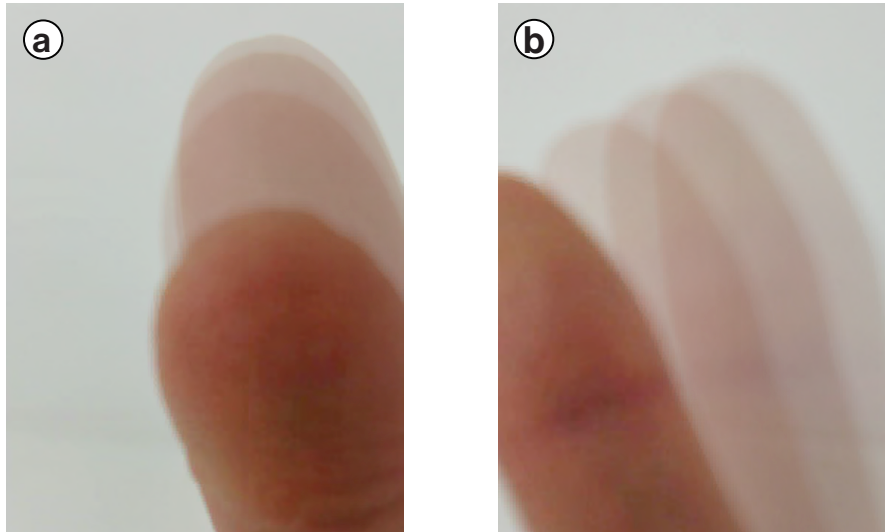


図 4.1: Samsung Galaxy S7 edge の背面カメラにて撮影されたジェスチャ実行時の人差し指の軌跡. a : Bend, b : Swipe.

#### 4.2.1 人差し指領域の抽出

人差し指領域を抽出するために、試作システム 1 は背面カメラを用いて 20 Hz にて撮影した RGB 画像 (720 × 1280 pixels) から肌の領域を抽出する。抽出には、次の Song ら [SSP<sup>+</sup>14] の式 (4.1) を用いた。

$$S(u) = \begin{cases} 1 & (\min(R(u) - G(u), R(u) - B(u)) > \tau) \\ 0 & (\text{otherwise}) \end{cases} \quad (4.1)$$

この時、 $u$  は画像中の座標、 $R(u)$ 、 $G(u)$ 、 $B(u)$  はそれぞれ  $u$  における画素の RGB 値である。また、 $\tau$  は閾値である。画像に対して式 (4.1) を適用することにより、座標  $u$  が肌の領域である場合に  $S(u) = 1$ 、それ以外の場合に  $S(u) = 0$  となる。この  $S(u) = 1$  となる画素群のうち、画像中にて最大の面積である領域を人差し指の領域とする。この処理により、人差し指領域より小さいノイズを除去する狙いがある。

式 (4.1) を用いる際には、環境光およびユーザの肌の色に応じて  $\tau$  を調整する必要がある。ただし、閾値が 1 種類であるため、容易に調整可能である。したがって、ユーザが試作システム 1 を使用する前に、 $\tau$  を調節することを想定している。

#### 4.2.2 ジェスチャの認識

ジェスチャを認識するために、試作システム 1 は、人差し指領域の抽出により得られた 2 値画像 (図 4.2) について、人差し指の傾きおよび人差し指領域の重心を取得する。人差し指の

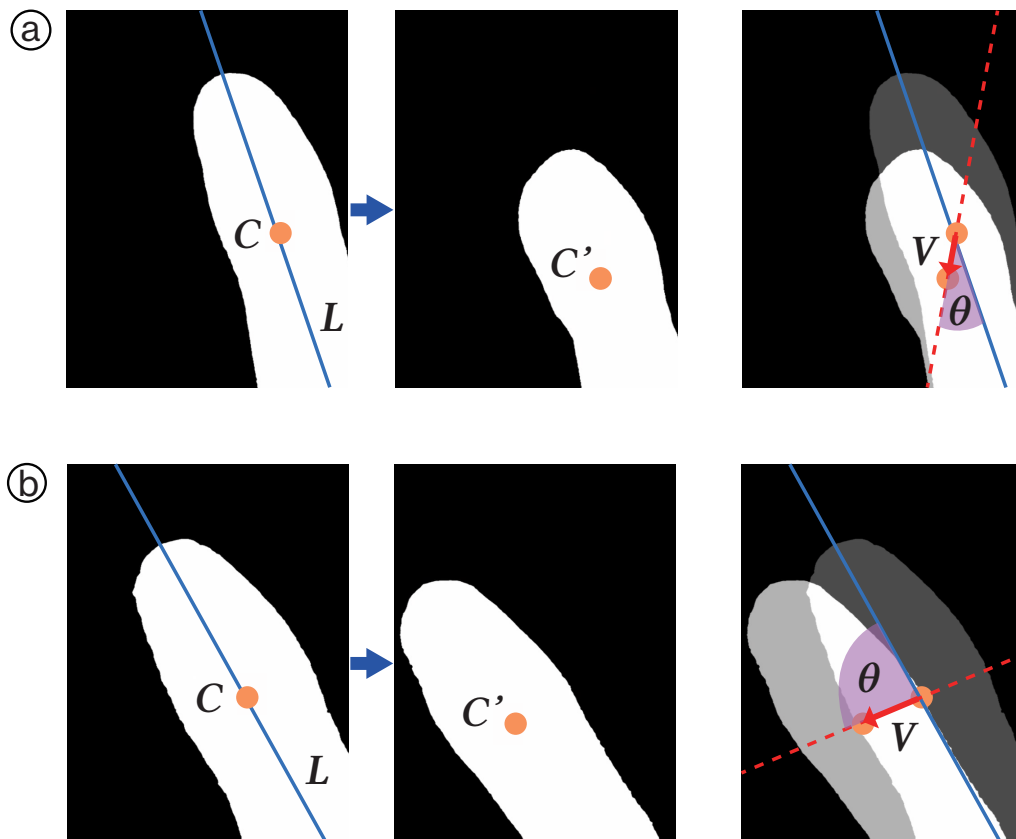


図 4.2: ジェスチャを実行した際に得られる 2 値画像. a: Bend, b: Swipe. 試作システム 1 では, 空中人差し指ジェスチャを認識するために, ジェスチャ実行前に, 人差し指領域にフィッティングした直線  $L$  および重心  $C$ , ジェスチャ実行中に, 重心  $C'$  を取得する. さらに, 有向線分  $CC'$  にて表されるベクトル  $V$  および  $L$  のなす角の内, 小さい方の角度  $\theta$  をジェスチャの認識に用いる.

傾きは、画素値が1の画素について、最小二乗法にてフィッティングした直線  $L$  の傾きとして算出する。また、人差し指領域の重心は、2値画像中における画素値が1の画素の重心として算出する。試作システム1では、ジェスチャ実行前に、人差し指の領域にフィッティングした直線  $L$  および重心  $C$ 、ジェスチャ実行中に、重心  $C'$  を取得する。なお、ジェスチャが実行されたかの判定には、人差し指領域の重心の移動速度を用いる。本実装では、重心の移動速度が  $1.2\text{pixels/ms}$  となった時にジェスチャが実行されたと判定する。これは、著者が実際にジェスチャを実行した際の経験に基づく。

ジェスチャの認識には、画像中における有向線分  $CC'$  にて表されるベクトル  $V$  およびジェスチャ実行前に、取得された  $L$  を用いる。この  $V$  および  $L$  のなす角の内、小さい方の角度  $\theta$  について、Bend の場合、Swipe に比べて角度が小さくなる。したがって、試作システム1は、 $\theta$  に関する閾値にてジェスチャを認識する。つまり、 $\theta$  が閾値未満であれば Bend、閾値以上であれば Swipe となる。なお、この閾値はユーザにて調節可能である。

## 第5章 試作システム1の認識精度評価

試作システム1の認識精度を評価するために、実験を行った。本章では、本実験における参加者、実験環境、タスクおよび実験手順、結果、および考察を述べる。

### 5.1 参加者

参加者はボランティアとして参加した研究室の大学生および大学院生4人（P1–P4、平均年齢22.8歳、SD=1.0歳）である。参加者の内、男性が3人、女性が1人であり、左利きが1人であった。デジタルノギス（シンワ19976）を用いて計測された参加者の手の大きさは平均186.6mm（SD=12.0mm）であり、人差し指の長さは平均72.6mm（SD=3.8mm）であった。各参加者の手の大きさおよび人差し指の長さを表5.1に示す。なお、手の大きさおよび人差し指の長さは、それぞれ日本人の手の寸法データ [河内12] における計測項目L02およびL10の長さに対応している。

### 5.2 実験環境

実験の様子を図5.1に示す。参加者は、照明条件として、ブラインドを下ろした室内にて照明を点けた状態（屋内点灯）または消した状態（屋内消灯）、晴れた日の屋外にて日向の状態（屋外日向）または日陰の状態（屋外日陰）にて実験タスクを行った。それぞれの照明条件において、マルチメータ（MASTECH MS8229）の照度計を用いて計測した照度は、表5.2の通りである。参加者全員が、すべての照明条件において、立位かつ右手にて実験タスクを行った。実験では、スマートフォンとして Samsung Galaxy S7 edge を用いた。

表 5.1: 試作システム1を用いた実験における参加者の手の大きさおよび人差し指の長さ。

| 参加者 | 手の大きさ     | 人差し指の長さ  |
|-----|-----------|----------|
| P1  | 191.12 mm | 76.69 mm |
| P2  | 192.19 mm | 73.62 mm |
| P3  | 168.69 mm | 67.47 mm |
| P4  | 194.52 mm | 72.81 mm |

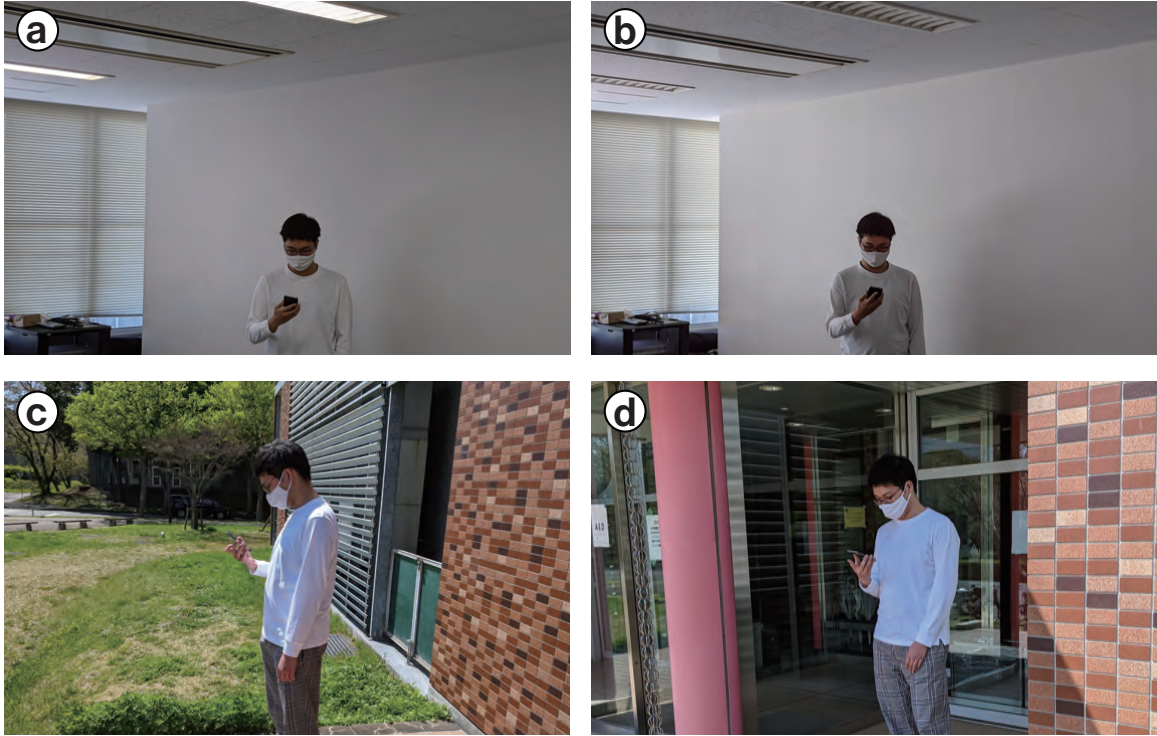


図 5.1: 4 種類の照明条件下における試作システム 1 を用いた実験の様子. a: 屋内点灯, b: 屋内消灯, c: 屋外日向, d: 屋外日陰.

表 5.2: 4 種類の照明条件下における照度.

|        | 屋内点灯        | 屋内消灯      | 屋外日向            | 屋外日陰          |
|--------|-------------|-----------|-----------------|---------------|
| 照度計上向き | 260–280 lux | 30–60 lux | 20000–32000 lux | 1500–4000 lux |
| 照度計下向き | 35– 40 lux  | 5–20 lux  | 2000– 7600 lux  | 300–1000 lux  |

表 5.3: 取得した全試行と認識結果に基づく混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | Bend  | Swipe | 未認識   |
|----------|-------|-------|-------|
| Bend     | 60.6% | 19.4% | 20.0% |
| Swipe    | 17.8% | 71.7% | 10.5% |

表 5.4: 4 種類の照明条件下にて取得した試行と認識結果に基づく混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 屋内点灯  |       |       | 屋内消灯  |       |       |
|----------|-------|-------|-------|-------|-------|-------|
|          | Bend  | Swipe | 未認識   | Bend  | Swipe | 未認識   |
| Bend     | 66.9% | 21.9% | 11.3% | 56.3% | 18.1% | 25.6% |
| Swipe    | 11.3% | 86.3% | 2.5%  | 20.0% | 68.1% | 11.9% |

| 真値 \ 予測値 | 屋外日向  |       |       | 屋外日陰  |       |       |
|----------|-------|-------|-------|-------|-------|-------|
|          | Bend  | Swipe | 未認識   | Bend  | Swipe | 未認識   |
| Bend     | 58.8% | 21.9% | 19.4% | 60.6% | 15.6% | 23.8% |
| Swipe    | 14.4% | 71.3% | 14.4% | 25.6% | 61.3% | 13.1% |

### 5.3 タスクおよび実験手順

本実験にて, 参加者が行ったタスクは次の通りである. まず, 空中人差し指ジェスチャの事前動作として, 参加者は背面カメラの前に人差し指を伸ばす. 次に, 参加者は空中人差し指ジェスチャを行う. この空中人差し指ジェスチャを開始するタイミングは, スマートフォンの振動にて知らされた.

本実験において, 1セッションは40試行(2種類のジェスチャ×繰り返し20回)であった. 実行するタスクはスマートフォンの画面表示にてランダムな順に提示された. 参加者全員がそれぞれ1回の練習セッションおよび2回の本番セッションを行った. 参加者は練習セッションにてジェスチャ認識に用いられる $\theta$ に関する閾値を調整した. 実験を通して, 1,280試行(40試行×本番セッション2回×照明条件4種類×参加者4人)分の認識結果を得た.

### 5.4 結果

本実験では, ジェスチャ実行中に Bend または Swipe が認識されないまま, 背面カメラの画角から人差し指領域が消失した場合, 認識結果を未認識とした. 表 5.3 に取得した全試行と認識結果に基づく混同行列を示す. 認識精度は 66.2%であった. また, 表 5.4 に4種類の照明条件下にて取得した試行と認識結果に基づく混同行列を示す. 認識精度は, 屋内点灯条件にて 76.6%, 屋内消灯条件にて 62.2%, 屋外日向条件にて 65.0%, 屋外日陰条件にて 60.9%であった.

表 5.5: 参加者毎の認識精度.

| 参加者 | 認識精度  |
|-----|-------|
| P1  | 68.8% |
| P2  | 70.6% |
| P3  | 61.3% |
| P4  | 64.1% |

## 5.5 考察

参加者が空中人差し指ジェスチャを実行する様子を観察したところ、ユーザの把持姿勢によっては、人差し指の移動方向が Bend および Swipe にて同様であることが分かった。人差し指の付け根がスマートフォンの側面に位置するように把持した場合、人差し指の移動方向は、いずれも取得される画像中にて右上から左下となる。さらに、ジェスチャ実行前の人差し指の傾きは Bend および Swipe にて同様である。よって、背面カメラ画像中の人差し指の傾きおよび人差し指領域の重心の移動を用いて、これらを区別することは難しい可能性がある。

本実験にて、未認識の割合は 15.2%であった。表 5.4 に示すように、特に、屋内消灯、屋外日向、屋外日陰の条件にて、Bend の未認識の割合が高い（それぞれ、25.6%、19.4%、23.8%）。これらの結果は、試作システム 1 によるジェスチャ認識が照明条件の影響を受けることを示している。しかし、未認識の場合、操作は実行されないため、ユーザが再度空中人差し指ジェスチャを行うことにより、ユーザの意図通りにスマートフォンを操作できる可能性がある。

参加者毎の認識精度を表 5.5 に示す。参加者の手の大きさおよび人差し指の長さに対する認識精度の相関係数を算出したところ、手の大きさおよび認識精度の相関係数は 0.69（正の相関）、人差し指の長さおよび認識精度の相関係数は 0.81（強い正の相関）であった。このことから、試作システム 1 の認識精度がユーザの手の大きさおよび人差し指の長さに依存することが示唆された。手が大きいまたは人差し指が長いほどジェスチャ実行中の人差し指の動作が大きく、また、背面カメラに映りやすいことが要因と考えられる。

さらに、ジェスチャを実行する人差し指の位置によって、空中人差し指ジェスチャを誤認識することが分かった。人差し指が背面カメラの画角の端に近づくと、ジェスチャ実行中の人差し指の動きを正確に捉えることが難しい。この問題を解決するために、実利用時においては、人差し指がジェスチャを開始できる範囲に入ったときに、フィードバックする必要がある。フィードバックする方法として、例えば、バイブレーションおよび画面表示が考えられる。

## 第6章 試作システム2：深層学習モデルを用いるシステム

試作システム1の実装では、背面カメラ画像中の人差し指の傾きおよび人差し指領域の重心の移動に着目して、閾値にて Bend および Swipe の認識を試みた。しかし、閾値を用いる実装では、認識精度が66.2%と低いことが分かった。これに対して、既存研究にて、深層学習モデルを用いることにより、カメラにて撮影された画像から手および指のジェスチャを高い精度にて認識可能であることが示されている [Zha19, IHuIA19, PRS20]。よって、本研究におけるジェスチャ認識システムにおいても、既存研究と同様に深層学習モデルを使用することにより、空中人差し指ジェスチャを認識する精度が向上すると考えた。

本研究では、深層学習モデルとして CNN モデルを用いて、空中人差し指ジェスチャを認識するシステムを試作した。試作システム1との区別のため、以降、本システムを試作システム2と呼称する。

本章では、試作システム2の実装環境および実装を述べる。

### 6.1 実装環境

試作システム2は、スマートフォンの背面カメラにて撮影した人差し指の画像から、7種類のジェスチャを認識するシステムである。試作システム2の実装には、Google Pixel 4 (147.1 mm × 68.8 mm × 8.2 mm, 5.7 インチ, Android 11) を用いた。このスマートフォンの背面カメラは、背面上部の左側（背面左端から 17.4 mm, 上端から 11.8 mm）に配置されている。また、画像処理ライブラリとして OpenCV (バージョン 4.5.0) を用いた。さらに、深層学習モデルの作製にはクラウド上で実行可能な環境である Google Colaboratory [Goo21b] を使用した。また、機械学習ライブラリとして TensorFlow (バージョン 2.4.1) を用いた。

### 6.2 実装

試作システム1では、ジェスチャを実行する人差し指の位置によって、空中人差し指ジェスチャを誤認識する課題があった。この課題を解決するためには、人差し指がジェスチャを開始できる範囲、つまり背面カメラの画角内に十分に入った時に、フィードバックする必要がある。よって、試作システム2では、空中人差し指ジェスチャの認識の前に、その事前動



作の認識を行う。また、この認識により、空中人差し指ジェスチャおよび日常の手の動作を切り分ける狙いもある。

本節では、試作システム2における事前動作の認識、および空中人差し指ジェスチャの認識を述べる。

### 6.2.1 事前動作の認識

事前動作の認識の流れを図6.1に示す。事前動作の認識は、スマートフォンの背面カメラを用いて、50 Hzにて取得された画像(240 × 320 pixels)に基づく。まず、取得した画像内の人差し指領域を抽出する。この処理は、4.2.1節に示した、試作システム1の人差し指領域の抽出と同じ処理を用いる。次に、画像内における人差し指領域の面積を算出する。

事前動作の認識には、ある一定時間に取得した画像を使用する。取得したすべての画像について、人差し指領域の面積の最大値および最小値を算出する。この最大値および最小値が閾値以内である場合に、事前動作として認識する。なお、画像を取得する時間および認識に用いる閾値は、利用時にユーザにより調節されることを想定している。

### 6.2.2 空中人差し指ジェスチャの認識

空中人差し指ジェスチャの認識の流れを図6.2に示す。空中人差し指ジェスチャの認識は、事前動作を認識した直後の一定時間にて取得した画像に基づく。この一定時間は、7種類の空中人差し指ジェスチャを実行可能な時間として、著者の経験に基づき700 msとした。試作システム2では、700 ms間に取得した35枚の画像(240 × 320 pixels)を用いて、MHIを生成する。さらに、MHIを特徴量とする深層学習モデルにてジェスチャを認識する。

### MHIの生成

取得した35枚の画像に対して、まず、人差し指領域の抽出およびグレースケール化の処理を行う(図6.2a)。人差し指領域の抽出には、4.2.1節に示した、試作システム1の人差し指領域の抽出と同じ処理を用いる。抽出された人差し指領域をマスク画像として、取得したRGB画像を次の式(6.1)を用いてグレースケール化する。

$$Y(u) = \begin{cases} 0.299 \times R(u) + 0.587 \times G(u) + 0.114 \times B(u) & (M(u) = 1) \\ 0 & (M(u) = 0) \end{cases} \quad (6.1)$$

この時、 $u$ は画像中の座標、 $R(u)$ 、 $G(u)$ 、 $B(u)$ はそれぞれ $u$ における画素のRGB値である。また、 $M(u)$ はマスク画像の画素値である。つまり、 $M(u) = 1$ となる画素群が人差し指領域を表す。式(6.1)におけるRGB値からグレースケールへの変換は、ITU-R BT.601 [ITU11]に規定された算出方法に倣っている。

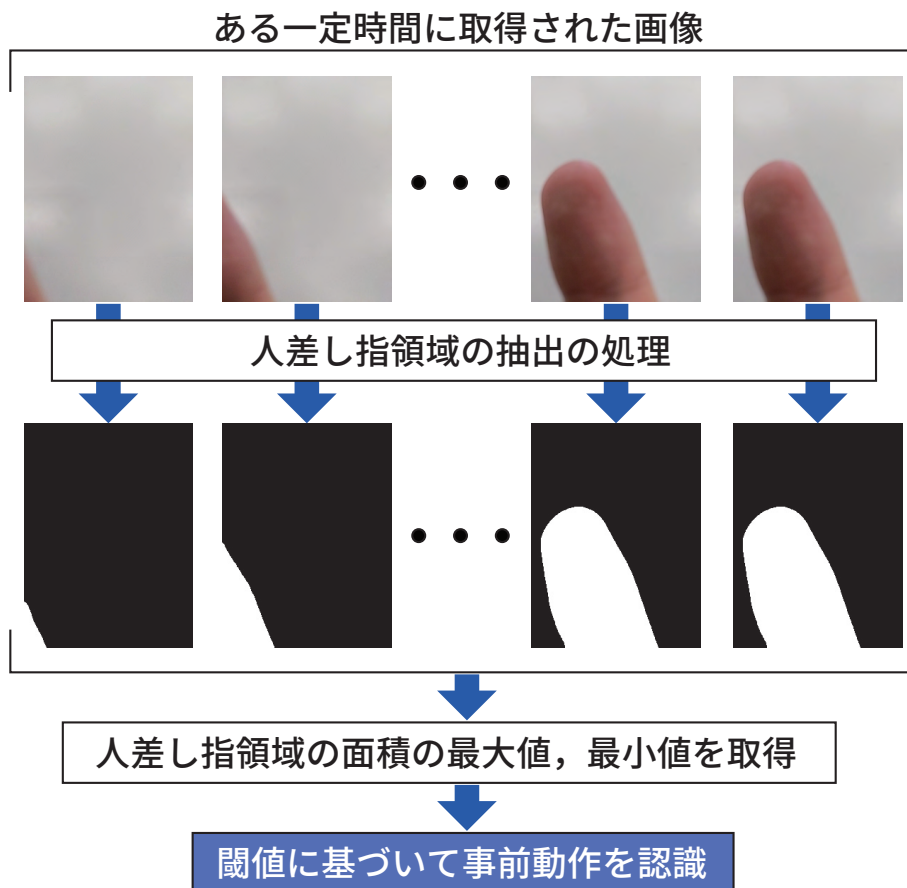


図 6.1: 事前動作の認識の流れ.

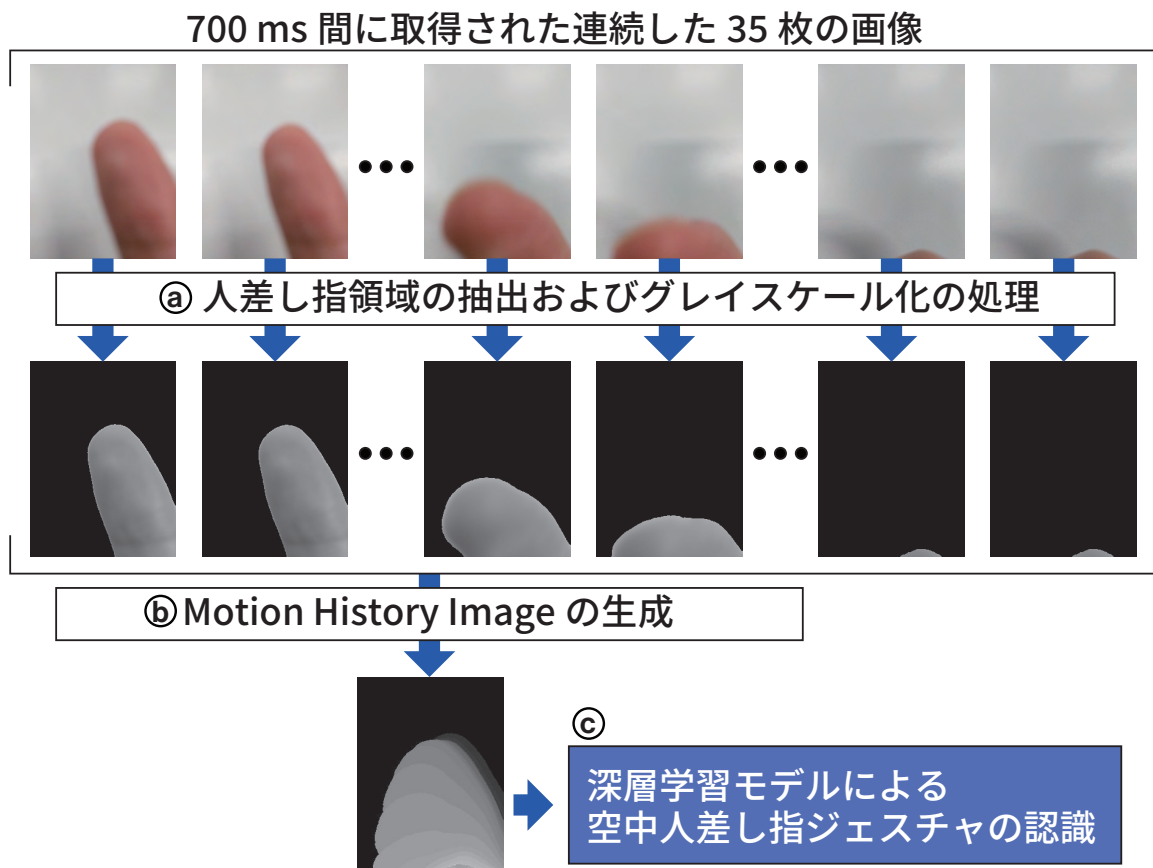


図 6.2: 空中人差し指ジェスチャの認識の流れ. a: 取得した 35 枚の画像に対して, 人差し指領域の抽出およびグレースケール化の処理を行う. b: Motion History Image を作製する. c: CNN モデルを用いて, Motion History Image から空中人差し指ジェスチャを認識する.

次に、これらの処理が行われた 35 枚の画像を用いて、MHI が生成される (図 6.2b)。MHI は動作履歴を輝度値にて表現する画像である。つまり、MHI にて新しい動作領域は輝度値が高く、古い動作領域は輝度値が低く示される。一時的な動作が画像中に描画されることから、MHI はこれまでも、深層学習モデルによるジェスチャ認識にて利用されている [YWL<sup>+</sup>19, WYY<sup>+</sup>20]。MHI の生成には、次の式 (6.2) を用いる。

$$H_{\delta}(u, t) = \begin{cases} \delta & (D(u, t) = 1) \\ \max(0, H_{\delta}(u, t - 1) - 1) & (\textit{otherwise}) \end{cases} \quad (6.2)$$

この時、 $\delta$  は動作履歴を保存する期間を決定するパラメータであり、 $H_{\delta}$  は MHI を表す。また、 $u, t$  はそれぞれ画像中の座標および時間を表す。 $D(u, t)$  は時系列にて連続した画像の差分から算出される、動作領域を示す。つまり、座標  $u$  について、時系列にて 1 枚前の画像と差分がある場合に  $H_{\delta}(u, t) = \delta$ 、差分がない場合に  $H_{\delta}(u, t) = \max(0, H_{\delta}(u, t - 1) - 1)$  となる。

これらの処理を通して、空中人差し指ジェスチャの動作履歴を含む MHI を得る。なお、Google Pixel 4 を用いて、35 枚の  $240 \times 320$  pixels の RGB 画像から  $240 \times 320$  pixels の MHI を生成するために必要な時間は、5 ms 程度である。

#### 深層学習モデルを用いたジェスチャ認識

試作システム 2 では、MHI から空中人差し指ジェスチャを認識するために、深層学習モデルを用いる (図 6.2c)。試作システム 2 の実装では、深層学習モデルとして、CNN モデルを使用する。空中人差し指ジェスチャの認識に使用される、CNN モデルの構成を図 6.3 に示す。このモデルは、 $240 \times 320$  pixels の MHI を入力として、認識したジェスチャの種類またはジェスチャが未実行であるかを出力する。

計算資源が制限されるスマートフォンに導入されることから、CNN モデルは軽量なモデルである必要がある。このことから、本実装における CNN モデルを、主に 3 層の畳み込み層および 2 層の全結合層から構成した。このモデルのすべての畳み込み層において、フィルタ数は 32、活性化関数は ReLU、ストライドは  $2 \times 2$  である。また、すべてのプーリング層におけるプーリングには、Max Pooling 法を採用した。プーリングにおけるカーネルサイズは  $2 \times 2$  である。ドロップアウト層では、データを 50% の確率にて無効にする。この層を用いることにより、過学習を抑える狙いがある。全結合層における活性化関数は ReLU である。このモデルにおける最適化手法には、Adam を採用した。また、学習率は 0.001 である。なお、Google Pixel 4 を用いて、 $240 \times 320$  pixels の MHI から CNN モデルを介して認識結果を得るまでに必要な時間は、30 ms 程度である。

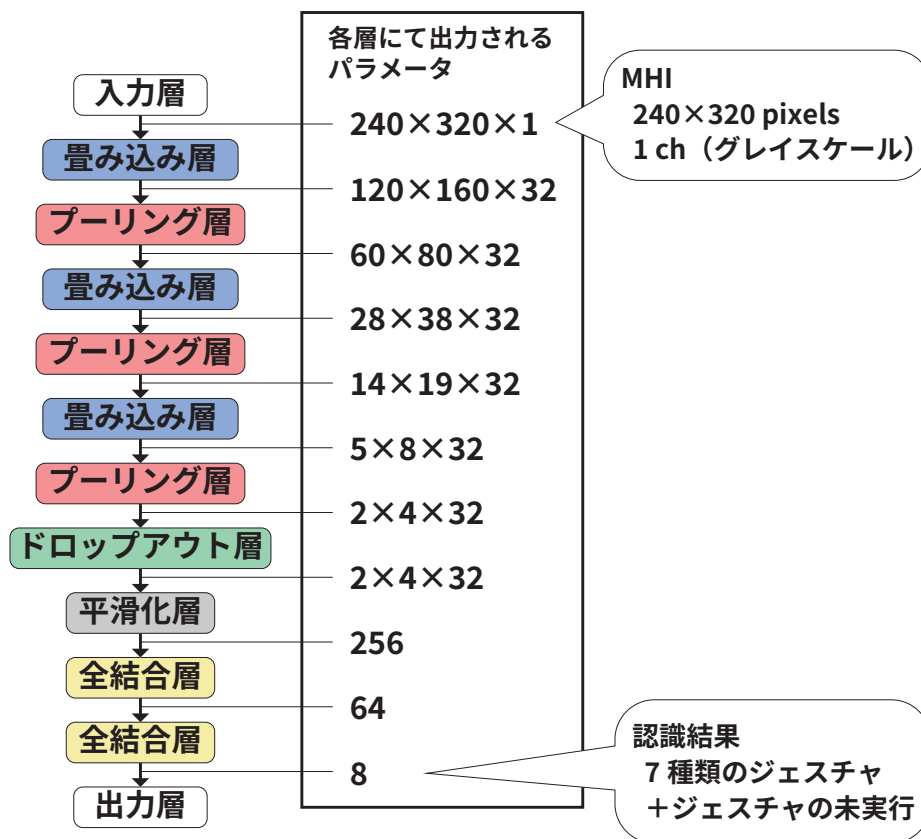


図 6.3: 試作システム 2 に用いる CNN モデルの構成.

## 第7章 試作システム2の認識精度評価

試作システム2の認識精度を評価するために、実験を行った。本実験では、空中人差し指ジェスチャの画像データの収集およびそのデータを用いてCNNモデルにて学習および検証を行った。本章では、本実験における参加者、実験環境、タスクおよび実験手順、および使用した画像データを述べる。また、7種類の空中人差し指ジェスチャの認識精度およびその考察を述べる。この認識精度に基づいて、試作システム2の改良の展望を示した後、4種類の空中人差し指ジェスチャに制限した際の認識精度およびその考察を述べる。最後に、本実験における制限を述べる。

### 7.1 参加者

参加者はボランティアとして参加した研究室内の大学生および大学院生10人（P1–P10、平均年齢22.5歳、SD = 0.8歳）である。参加者の内、男性が9人、女性が1人であり、全員が右利きであった。デジタルノギス（シンワ19976）を用いて計測された参加者の手の大きさは平均183.9mm（SD = 8.9mm）であり、人差し指の長さは平均71.6mm（SD = 5.6mm）であった。各参加者の手の大きさおよび人差し指の長さを表7.1に示す。なお、手の大きさおよび人差し指の長さは、それぞれ日本人の手の寸法データ [河内12] における計測項目L02およびL10の長さに対応している。

### 7.2 実験環境

実験の様子を図7.1に示す。参加者は、同じ室内にて、椅子に腰掛けて、右手にて実験タスクを行った。実験中はブラインドが閉じられた上で、照明が点けられた。同室内にて、照度計（サトテック デジタル照度計 CENTER530）を用いて計測した照度は、上向きにて530–650 lux、下向きにて180–260 luxであった。

実験では、スマートフォンとしてGoogle Pixel 4を用いた。また、深層学習モデルの作製および交差検証にGoogle Colaboratory [Goo21b]を使用した。

### 7.3 タスクおよび実験手順

本実験では、参加者が1回タスクを行うことを1試行とした。本実験にて、参加者が行ったタスクは次の2通りである。



図 7.1: 試作システム 2 を用いた実験の様子.

表 7.1: 試作システム 2 を用いた実験における参加者の手の大きさおよび人差し指の長さ.

| 参加者 | 手の大きさ     | 人差し指の長さ  |
|-----|-----------|----------|
| P1  | 191.24 mm | 75.33 mm |
| P2  | 167.50 mm | 62.80 mm |
| P3  | 193.44 mm | 71.24 mm |
| P4  | 185.95 mm | 71.60 mm |
| P5  | 187.00 mm | 72.29 mm |
| P6  | 186.29 mm | 73.24 mm |
| P7  | 180.57 mm | 66.14 mm |
| P8  | 174.10 mm | 67.78 mm |
| P9  | 177.24 mm | 71.95 mm |
| P10 | 195.24 mm | 83.43 mm |

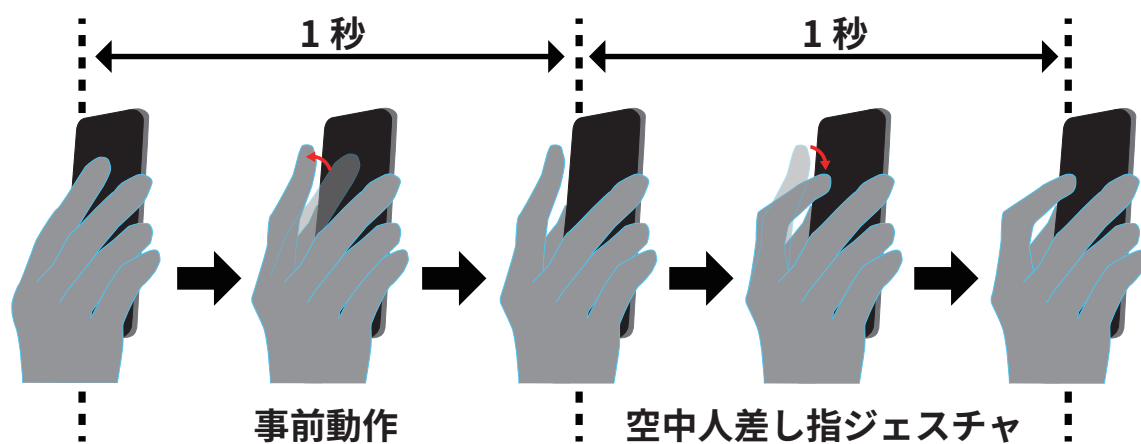


図 7.2: 試作システム 2 を用いた実験にて, 参加者が行ったタスク 1. タスク 1 は, 空中人差し指ジェスチャを実行するタスクである. 参加者は, はじめの 1 秒にて人差し指を背面カメラの前に伸ばす空中人差し指ジェスチャの事前動作を行い, その後の 1 秒にて空中人差し指ジェスチャを実行するように指示された. このタスクにて, 空中人差し指ジェスチャ実行中の画像を取得した.



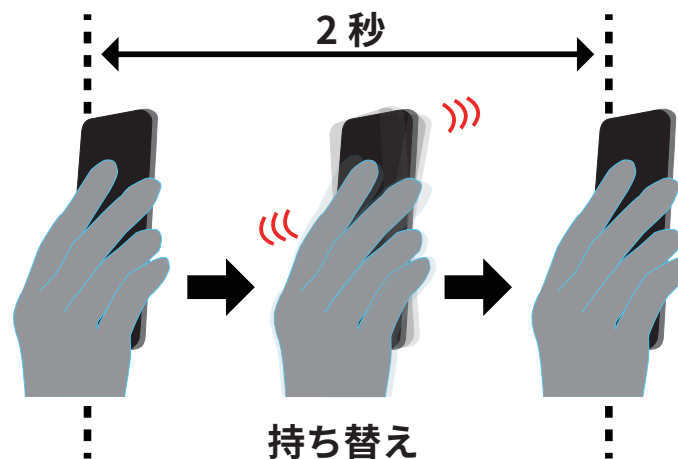


図 7.3: 試作システム 2 を用いた実験にて、参加者が行ったタスク 2. タスク 1 は、空中人差し指ジェスチャを実行するタスクである。参加者は、2 秒間継続してスマートフォンを自由に持ち替えるように指示された。なお、人差し指を背面カメラの前に伸ばす空中人差し指ジェスチャの事前動作および空中人差し指ジェスチャを行わないように注意を受けた。このタスクにて、空中人差し指ジェスチャ未実行時の画像を取得した。

**タスク 1** 空中人差し指ジェスチャを実行するタスク (図 7.2)

**タスク 2** スマートフォンを持ち替えるタスク (図 7.3)

タスク 1 にて、参加者は、まず、空中人差し指ジェスチャの事前動作として背面カメラの前に人差し指を伸ばす。参加者は、この動作をタスク開始から 1 秒経過するまでに実行するように指示された。事前動作を行う 1 秒が経過した後、参加者は空中人差し指ジェスチャを行う。参加者は、提示された空中人差し指ジェスチャを 1 秒経過するまでに実行するように指示された。なお、参加者は、事前動作または空中人差し指ジェスチャを正しく行えなかったと判断した場合にそのタスクをやり直すように指示された。このタスクにて、空中人差し指ジェスチャ実行中の画像が取得された。

また、タスク 2 として、参加者は 2 秒間継続して自由にスマートフォンを持ち替えるタスクを行った。この際、参加者は、人差し指を背面カメラの前に伸ばす空中人差し指ジェスチャの事前動作および空中人差し指ジェスチャを行わないように指示された。このタスクにて、空中人差し指ジェスチャ未実行の画像が取得された。

本実験においては、タスク 1 またはタスク 2 を実行することを 1 試行とした。1 セッションは 80 試行 ((7 種類のジェスチャ × タスク 1 + タスク 2) × 繰り返し 10 回) であった。参加者が実行するタスクはスマートフォンの画面表示にてランダムな順に提示された。参加者全員がそれぞれ 3 セッション分のタスクを行った。各セッション間に 5 分程度の休憩が設けられた。実験に要した時間は、参加者 1 人当たり 45 分程度であった。

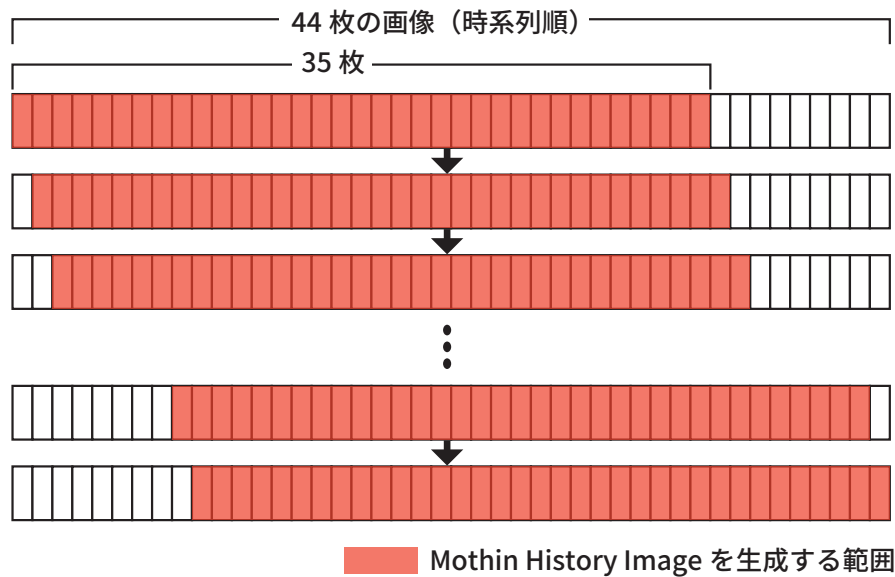


図 7.4: 空中人差し指ジェスチャを実行中の 44 枚の画像から MHI を生成する流れ. 時系列順の 44 枚の画像について, 35 枚をウィンドウサイズとするスライディングウィンドウにて分割した後, ウィンドウ毎に MHI を生成した.

## 7.4 画像データ

本実験において, 1 試行にて取得した画像の枚数は 100 枚 (2 秒  $\times$  50 Hz) である. その内, 参加者がタスク 1 にて, 事前動作を行った際に収集された画像は時系列順に 1-50 枚目の 50 枚, 空中人差し指ジェスチャを実行した際に収集された画像は時系列順に 51-100 枚目の 50 枚である. 本実験では, 空中人差し指ジェスチャを実行した際に収集された 50 枚の画像について, 時系列順に 1-44 枚目を空中人差し指ジェスチャを実行中の画像として扱った. これは, 45 枚目以降の画像は, 参加者がジェスチャを完了した後に取得されたと予想したためである. つまり, 人差し指が静止しているもしくは画角外にて移動している際に取得された画像と考えられる.

空中人差し指ジェスチャを実行中の 44 枚の画像から MHI を生成する流れを図 7.4 に示す. まず, 時系列順の 44 枚の画像について, 35 枚をウィンドウサイズとするスライディングウィンドウにて分割した. 次に, ウィンドウ毎に人差し指領域の抽出およびグレイスケール化の処理, ならびに MHI の生成を行った. なお, 人差し指領域の抽出において, 式 (4.1) に用いる閾値  $\tau$  を 0.1 とした. これらの処理により, 10 枚の MHI を得た. また, タスク 2 にて取得した画像についても, 同じ処理を行うことにより, 10 枚の MHI を得た. 実験を通して, 画像データとして 24,000 枚 (10 枚  $\times$  80 試行  $\times$  3 セッション  $\times$  10 人) の MHI を得た.

表 7.2: 5 分割交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 99.1% | 0.3%    | 0.4%  | 0.1%  | 0.1%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.7%  | 94.4%   | 0.0%  | 0.3%  | 0.3%  | 0.1%   | 3.2%    | 1.0%     |
| Bend     | 5.2%  | 0.0%    | 75.6% | 8.8%  | 6.3%  | 1.1%   | 0.9%    | 2.1%     |
| Bend2    | 0.0%  | 0.0%    | 3.6%  | 73.0% | 0.4%  | 10.4%  | 0.4%    | 12.2%    |
| Swipe    | 9.3%  | 0.2%    | 24.8% | 4.4%  | 53.9% | 6.4%   | 0.2%    | 0.8%     |
| Swipe2   | 0.0%  | 0.1%    | 0.9%  | 23.3% | 1.5%  | 70.6%  | 0.1%    | 3.5%     |
| Release  | 0.2%  | 1.9%    | 2.7%  | 0.6%  | 0.1%  | 0.1%   | 84.4%   | 10.0%    |
| Release2 | 0.0%  | 0.4%    | 0.1%  | 7.4%  | 0.0%  | 0.6%   | 3.3%    | 88.2%    |

表 7.3: leave-one-participant-out 交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 98.1% | 1.1%    | 0.1%  | 0.2%  | 0.2%  | 0.0%   | 0.2%    | 0.1%     |
| Stretch  | 1.2%  | 90.0%   | 0.1%  | 0.1%  | 0.8%  | 0.5%   | 6.6%    | 0.7%     |
| Bend     | 5.7%  | 0.4%    | 46.7% | 6.2%  | 32.1% | 4.4%   | 2.9%    | 1.5%     |
| Bend2    | 0.0%  | 0.0%    | 3.0%  | 34.2% | 2.4%  | 33.9%  | 2.6%    | 23.9%    |
| Swipe    | 9.7%  | 0.1%    | 13.6% | 2.3%  | 63.0% | 8.0%   | 1.6%    | 1.8%     |
| Swipe2   | 0.0%  | 0.5%    | 1.1%  | 14.9% | 6.8%  | 66.2%  | 0.5%    | 10.0%    |
| Release  | 1.1%  | 10.6%   | 4.4%  | 2.2%  | 1.2%  | 1.2%   | 72.0%   | 7.3%     |
| Release2 | 0.2%  | 0.8%    | 0.5%  | 18.2% | 0.6%  | 5.1%   | 11.3%   | 63.3%    |

## 7.5 結果 1 : 7 種類の空中人差し指ジェスチャの認識精度

取得した画像データを用いて, 7 種類の空中人差し指ジェスチャの認識精度を調査した. この調査では, 5 分割交差検証および leave-one-participant-out 交差検証を行った.

### 7.5.1 5 分割交差検証

画像データ 24,000 枚を用いて 5 分割交差検証を行った. 分割においては, 24,000 枚の画像データを 4,800 枚ずつの 5 つのデータセットとした. このとき, データセット間にて, 空中人差し指ジェスチャの種類および画像の枚数のバランスを取った. また, 同じ試行から得られた画像が同じデータセットに含まれるように分割を行った. この交差検証においては, 5 つのデータセットのうち, 1 つを検証データ, 残りの 4 つを学習データとした. ここで, 学習データにのみ拡張を行った. この拡張にて学習データの多様性を向上させることにより, 過学習を防ぐ狙いがある. データセット内の画像に対して, 中心を基準として 1.2 倍に拡大した画像

および0.8倍に縮小した画像を学習データに追加した。これにより、学習データに含まれる画像の枚数を3倍に拡張した。よって、検証に用いられる画像は4,800枚、学習に用いられる画像は57,600枚であった。なお、CNNモデルの学習におけるエポック数は80であった。

交差検証の結果、認識精度は平均79.9% (SD = 1.2%)であった。表7.2に混同行列を示す。なお、分割毎の認識精度は、79.7%, 81.8%, 80.1%, 78.6%, 79.4%であった。表A.1–A.5に分割毎の混同行列を示す。

## 7.5.2 leave-one-participant-out 交差検証

画像データ24,000枚を用いてleave-one-participant-out交差検証を行った。この交差検証においては、1人分の画像データを検証データ、残りの3人分の画像データを学習データとする。ここで、5分割交差検証と同じく、学習データにのみ拡張を行った。データセット内の画像に対して、中心を基準として1.2倍に拡大した画像および0.8倍に縮小した画像を学習データに追加した。これにより、学習データに含まれる画像の枚数を3倍に拡張した。よって、検証に用いられる画像は2400枚、学習に用いられる画像は64,800枚であった。なお、CNNモデルの学習におけるエポック数は80であった。

交差検証の結果、認識精度は平均66.7% (SD = 8.7%)であった。表7.3に混同行列を示す。なお、分割されたデータセット毎の認識精度は、67.0%, 56.7%, 77.6%, 76.1%, 69.3%, 59.5%, 53.9%, 64.7%, 64.0%, 78.5%であった。表A.6–A.15に分割毎の混同行列を示す。

## 7.6 結果1に関する考察

5分割交差検証の結果、認識精度は平均79.9%であった。このことから、空中人差し指ジェスチャをMHIからある程度の精度にて認識可能であることが分かった。ただし、Stretchを除く空中人差し指ジェスチャにて精度が低いことから、試作システム2は実利用には耐え難いと考えられる。

5分割交差検証における混同行列(表7.2)より、Swipeの認識精度が低いことが分かった。特に、Bendと誤認識する割合が高い。このことから、MHIからSwipeの特徴が捉えきれなかった可能性がある。これは、5.5節の考察にて述べた、把持姿勢による影響と考えられる。つまり、人差し指の付け根がスマートフォンの側面に位置するように把持した場合、人差し指の移動方向は、BendおよびSwipeのいずれも取得される画像中にて右上から左下となる。そのため、同様な動作履歴が表現されたMHIが生成された可能性がある。

また、同じ動作を2回繰り返すジェスチャ(Bend2, Swipe2, およびRelease2)同士を誤認識する傾向が見られた。特にBend2およびSwipe2, Bend2およびRelease2にて、この傾向が顕著である。このことから、繰り返すジェスチャを行った際に生成されたMHIに同様な特徴があると考えられる。さらに、この傾向の原因として、MHIにおいて、同じ動作を正確に繰り返した場合、過去の動作が新しい動作に上書きされることが挙げられる。そのため、Bend2,

表 7.4: leave-one-participant-out 交差検証にて、各参加者のデータを検証データとした際の認識精度.

| 参加者 | 認識精度   |
|-----|--------|
| P1  | 67.00% |
| P2  | 56.70% |
| P3  | 77.60% |
| P4  | 76.10% |
| P5  | 69.30% |
| P6  | 59.50% |
| P7  | 53.90% |
| P8  | 64.70% |
| P9  | 64.00% |
| P10 | 78.50% |

Swipe2, および Release2 を実行した際に生成される MHI において、指の動作に関する情報が損なわれた可能性がある。

leave-one-participant-out 交差検証の結果、認識精度は平均 66.7%であった。このことから、未学習のユーザの空中人差し指ジェスチャに対する認識精度は低いことが分かった。よって、実利用の際には、ユーザ向けのシステムとするために、ユーザ毎にジェスチャを学習する必要がある。

leave-one-participant-out 交差検証における混同行列（表 7.3）より、Bend および Bend2 の認識精度は、いずれも平均 50%未満と低いことが分かった。5 分割交差検証においては、いずれも平均 70%以上であるため、大きく精度が落ち込んでいる。この傾向から、Bend および Bend2 について、ユーザ間にてその実行方法に個人差があることが考えられる。

また、Release が Stretch に誤認識される傾向がある。この傾向について、Release が実行された際の人差し指の動作が小さかったため、Stretch と同様な特徴が学習された可能性がある。なお、参加者より、Release の実行が難しかったと報告されていることから、Release の難易度が高いために人差し指をうまく動かせなかったことが原因として考えられる。

leave-one-participant-out 交差検証にて、各参加者のデータを検証データとした際の認識精度を表 7.4 に示す。参加者の手の大きさおよび人差し指の長さに対する、その参加者のデータを検証データとした際の認識精度の相関係数を算出したところ、手の大きさおよび認識精度の相関係数は 0.71（強い正の相関）、人差し指の長さおよび認識精度の相関係数は 0.67（正の相関）であった。つまり、試作システム 2 の認識精度は、ユーザの手の大きさおよび人差し指の長さに依存することが分かった。これは、ユーザの手の大きさが小さいまたは人差し指の長さが短いほど、背面カメラの画角内に指が伸ばしにくいことが原因と考えられる。ただし、より多くのユーザによるジェスチャを学習することにより、改善する可能性がある。

表 7.2 および表 7.3 の混同行列が示すように、5 分割交差検証および leave-one-participant-out

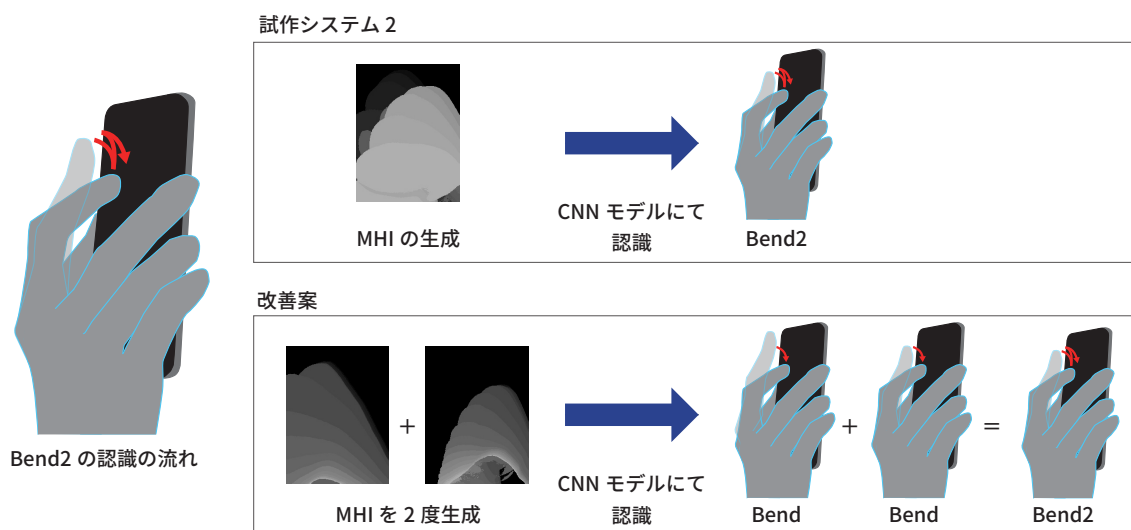


図 7.5: 試作システム 2 および改善案における Bend2 の認識の流れ.

交差検証のいずれにおいても、ジェスチャの未実行は高い精度にて認識された。このため、事前動作が誤って認識された場合においても、空中人差し指ジェスチャが認識される可能性は低いと考えられる。

## 7.7 試作システム 2 の改良の展望

交差検証にて、現状の試作システム 2 は実用には耐え難いことが分かった。今後、試作システム 2 を改良するにあたって、同じ動作を 2 回繰り返すジェスチャ (Bend2, Swipe2, および Release2) に注目した。試作システム 2 において、それぞれ同士を誤認識するため、これらの認識精度は低いことが分かっている。これらは、Bend, Swipe, および Release を 2 回繰り返すジェスチャであるため、システムの工夫により認識できる可能性がある。例えば、指を曲げる動作の度に MHI を生成して Bend が 2 回実行されたことを認識することにより、Bend2 を認識できる (図 7.5)。ただし、この改良を行う前提として、同じ動作を 2 回繰り返すジェスチャを除いた場合において、ジェスチャの認識精度が向上することを確認する必要がある。

## 7.8 結果 2 : 4 種類の空中人差し指ジェスチャの認識精度

試作システム 2 の改良に先立ち、同じ動作を 2 回繰り返すジェスチャを除いた、4 種類の空中人差し指ジェスチャ (Stretch, Bend, Swipe, および Release) に限定して認識精度を調査した。この調査では、同じ動作を繰り返すジェスチャの MHI を除いた 15,000 枚の画像データを用いて、5 分割交差検証および leave-one-participant-out 交差検証を行った。

表 7.5: 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Swipe | Release |
|----------|-------|---------|-------|-------|---------|
| 未実行      | 99.1% | 0.5%    | 0.3%  | 0.1%  | 0.1%    |
| Stretch  | 0.5%  | 94.9%   | 0.1%  | 0.3%  | 4.1%    |
| Bend     | 5.2%  | 0.0%    | 78.3% | 13.3% | 3.1%    |
| Swipe    | 9.1%  | 0.5%    | 15.6% | 73.6% | 1.2%    |
| Release  | 0.1%  | 3.2%    | 2.6%  | 0.4%  | 93.7%   |

表 7.6: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Swipe | Release |
|----------|-------|---------|-------|-------|---------|
| 未実行      | 98.6% | 0.8%    | 0.3%  | 0.1%  | 0.3%    |
| Stretch  | 0.8%  | 87.5%   | 0.2%  | 1.3%  | 10.3%   |
| Bend     | 5.4%  | 0.6%    | 73.3% | 13.5% | 7.2%    |
| Swipe    | 9.6%  | 0.3%    | 38.3% | 48.9% | 2.9%    |
| Release  | 0.2%  | 10.2%   | 4.8%  | 1.4%  | 83.4%   |

### 7.8.1 5 分割交差検証

画像データ 15,000 枚を用いて 5 分割交差検証を行った. 画像データの分割および学習データ拡張の方法は, 7.5.1 節に示した方法と同じである. これにより, 検証に用いられる画像は 3,000 枚, 学習に用いられる画像は 36,000 枚であった. なお, CNN モデルの学習におけるエポック数は 80 であった.

交差検証の結果, 認識精度は平均 87.9% (SD = 2.2%) であった. 表 7.5 に混同行列を示す. なお, 分割されたデータセット毎の認識精度は, 86.6%, 90.9%, 85.1%, 89.2%, 87.9% であった. 表 A.16–A.20 に分割毎の混同行列を示す.

### 7.8.2 leave-one-participant-out 交差検証

画像データ 15,000 枚を用いて leave-one-participant-out 交差検証を行った. 画像データの分割および学習データ拡張の方法は, 7.5.2 節に示した方法と同じである. これにより, 検証に用いられる画像は 1,500 枚, 学習に用いられる画像は 40,500 枚であった. なお, CNN モデルの学習におけるエポック数は 80 であった.

交差検証の結果, 認識精度は平均 78.3% (SD = 8.6%) であった. 表 7.6 に混同行列を示す. なお, 分割されたデータセット毎の認識精度は, 86.2%, 61.2%, 79.1%, 87.4%, 75.6%, 74.5%, 74.6%, 72.8%, 81.7%, 90.3% であった. 表 A.21–A.30 に分割毎の混同行列を示す.

表 7.7: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて、各参加者のデータを検証データとした際の認識精度。

| 参加者 | 認識精度   |
|-----|--------|
| P1  | 86.20% |
| P2  | 61.20% |
| P3  | 79.10% |
| P4  | 87.40% |
| P5  | 75.60% |
| P6  | 74.50% |
| P7  | 74.60% |
| P8  | 72.80% |
| P9  | 81.70% |
| P10 | 90.30% |

## 7.9 結果 2 に関する考察

空中人差し指ジェスチャを 4 種類に限定した場合、5 分割交差検証の結果、認識精度は平均 87.9%であった。ここで、7 種類の空中人差し指ジェスチャについて、認識精度は平均 79.9%であったため、ジェスチャを限定することにより精度が向上することが分かった。主に Bend, Swipe, および Release の認識精度が向上していることから、これらのジェスチャが Bend2, Swipe2, および Release2 として誤認識されることがなくなったことが要因と考えられる。

leave-one-participant-out 交差検証においては、空中人差し指ジェスチャを 4 種類に限定した場合、7 種類のジェスチャの場合に比べて、Swipe の認識精度に低下が見られた。7 種類の場合に 63.0%、4 種類の場合に 48.9%である。表 7.6 の混同行列から、主に Bend に誤認識されていることが原因である。ユーザ毎に空中人差し指ジェスチャ、特に Bend および Swipe を学習し直して、ユーザ向けにカスタマイズすることにより、改善する可能性がある。

leave-one-participant-out 交差検証にて、各参加者のデータを検証データとした際の認識精度を表 7.7 に示す。参加者の手の大きさおよび人差し指の長さに対する、その参加者のデータを検証データとした際の認識精度の相関係数を算出したところ、手の大きさおよび認識精度の相関係数は 0.77 (強い正の相関)、人差し指の長さおよび認識精度の相関係数は 0.82 (強い正の相関)であった。つまり、認識精度がユーザの手の大きさおよび人差し指の長さに依存する傾向が、7 種類のジェスチャの場合に比べて強いことが分かる。これを改善するためには、より多くの参加者のジェスチャのデータを学習することにより、学習時に手の大きさおよび人差し指の長さに関して多様性を確保する必要がある。また、学習データの拡張の際に、ユーザの手の大きさおよび人差し指の長さに合わせて、拡大縮小率を変えるといった工夫により改善する可能性がある。



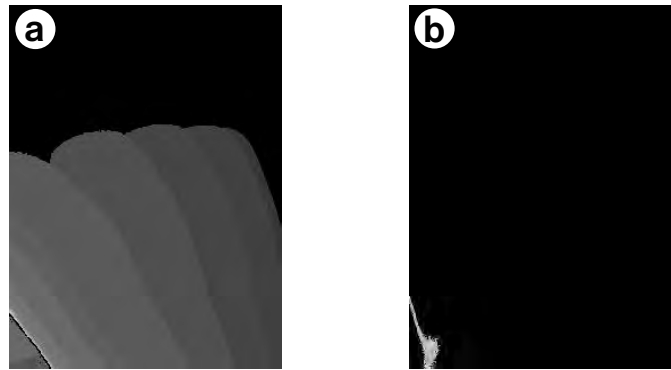


図 7.6: 実験にて Swipe が実行された際に生成された MHI. a: カメラの画角内にて実行された場合. b: カメラの画角の端にて実行された場合.

## 7.10 実験における制限

今回の実験においては、空中人差し指ジェスチャの事前動作の認識および事前動作が正しく行われているかのフィードバックを行わなかった。そのため、参加者が空中人差し指ジェスチャを実行する前に人差し指が画角の端または画角外にあることを確認できなかった。背面カメラの画角の端にてジェスチャが実行された場合には、画角内にて実行された際に生成される MHI (図 7.6a) に比べて、指の動作をうまく捉えられていない MHI (図 7.6b) が作製される。これが、7 種類および 4 種類の空中人差し指ジェスチャの認識において、Bend および Swipe の精度が低い原因となった可能性がある。特に、leave-one-participant-out 交差検証にて、Bend および Swipe がジェスチャの未実行として認識されやすい参加者がいる (表 A.7, 表 A.10, 表 A.13, 表 A.22, 表 A.25, 表 A.28) ことから、指の動作をうまく捉えられていない MHI が画像データに複数存在すると考えられる。よって、実験にて得られたデータについて、正しくラベル付けし直した上で、精度を確認する必要がある。

今回は、スマートフォンとして Google Pixel 4 のみを用いて実験を行った。そのため、他の機種における認識精度の評価が不十分である。なお、実験結果より認識精度がユーザの手の大きさおよび人差し指の大きさに依存することから、背面カメラにおける人差し指の映り具合による認識精度への影響が考えられる。よって、人差し指の映り具合が異なる、本体のサイズおよび背面カメラの位置が異なるスマートフォンにて評価する必要がある。また、機種によって精度が大きく異なる場合には、機種毎にて CNN モデルの作製またはファインチューニングといった追加の実装が求められる。

試作システム 1 における認識精度の評価と異なり、試作システム 2 の照明条件に対する評価を行っていない。人差し指領域の抽出においては、試作システム 1 および試作システム 2 において同様の手法を用いている。よって、試作システム 2 においても、照明条件による認識精度への影響が考えられる。これに対して、先行研究 [BZFR11, OH13, RCGS16, JC19] にて示された肌領域の抽出方法を用いることより、照明条件による影響が小さくなる可能性があ

るため、調査が必要である。また、今後のスマートフォンの発展により、背面カメラとして深度カメラが導入されると、照明条件に依存することなく人差し指領域を抽出できる。

## 第8章 議論

本章では、空中人差し指ジェスチャを認識するシステムおよび空中人差し指ジェスチャによる操作手法に関する今後の課題を述べる。

### 8.1 空中人差し指ジェスチャを認識するシステムの改良

試作システム2における空中人差し指ジェスチャの認識精度を向上させるために、CNNモデルの改善が必要である。特に、認識精度はユーザの手の大きさおよび人差し指の大きさに依存することが分かっているため、ユーザの個人差に依らず認識可能にする対応が求められる。これに対して、学習に用いるデータの多様性を高めることにより、精度が向上する可能性がある。例えば、追加のデータを参加者から取得すること、また、7.6節および7.9節にて述べたように、より多くの参加者からデータを取得すること、および学習データの拡張方法を増やすなどの方法が考えられる。さらに、CNNモデルの構成に関して、畳み込み層およびプーリング層といった層の構造、ならびに最適化手法などについて、パラメータチューニングによる改良が考えられる。また、スマートフォン内蔵センサにて計測された加速度および角速度を特徴量に含めることにより、画像中の背景におけるノイズに対して堅牢なモデルを作製できる可能性がある。

試作システム2にて生成される  $240 \times 320$  pixels のMHIには、人差し指の動作に関する情報の他に、人差し指の概形および陰影などの情報も含まれる。そのため、CNNモデルの学習時に、このMHIからそれぞれのジェスチャの特徴に関連性の低い情報が学習される可能性がある。これに対して、画素数を少なくすることにより、MHIから得られる情報を少なくできるため、人差し指の動作に関する情報に絞れると考えられる。また、MHIの生成およびCNNモデルを用いた認識がより高速に処理できると予想する。

7.7節にて、試作システム2の改良の展望を述べた。これに加えて、4種類の空中人差し指ジェスチャに限定することにより、試作システム2の認識精度が向上することが分かっている。よって、空中人差し指ジェスチャについて、人差し指の動作毎にMHIの生成およびCNNモデルにてジェスチャを認識する処理を導入することにより、システムの改善が図れる。さらに、この処理により、例えば、Bendの後にSwipeを実行するジェスチャおよびBendを3回繰り返すジェスチャなど、新たなジェスチャの設計も可能であると考えられる。ただし、人差し指の動作毎に処理を行うためには、人差し指の動作を検知する必要がある。この人差し指の動作を検知する方法として、試作システム1の実装にて用いた人差し指領域の重心の移動または先行研究における指先の検出 [OH13] が利用できると考えられる。

## 8.2 空中人差し指ジェスチャの難易度

スマートフォンの片手使用時において、ユーザは人差し指にて背面から支えるようにスマートフォンを持つ。つまり、人差し指がスマートフォンを持つために用いられる。このことから、ユーザが意図して人差し指を動かすことは少ないと考えられる。そのため、スマートフォンを持ちながら実行する空中人差し指ジェスチャは難易度が高い可能性がある。実験においても、参加者の多くは7種類の空中人差し指ジェスチャの中に実行が難しいジェスチャがあったという感想を述べた。

容易に実行可能な操作を実現するためには、空中人差し指ジェスチャの設計に改善の余地がある。アプローチの1つとして、より小さい動作のジェスチャにて設計することが挙げられる。現状の設計と比べて小さい動作にて実行可能であれば、ユーザの人差し指への負荷が小さくなる。これにより、実行難易度が低くなると考えられる。また、ユーザからジェスチャを誘発する実験 [WMW09, MWW10, MDD<sup>+</sup>14, SMSJ<sup>+</sup>15] にて、ユーザ定義のジェスチャを得るアプローチも考えられる。ただし、これらのアプローチにて設計したジェスチャについて、背面カメラ画像にて人差し指の動作を捉えられるかを確認する必要がある。

## 8.3 追加で実施すべき調査

本論文では、スマートフォンの背面カメラを用いて空中人差し指ジェスチャを認識するシステムを示した。この実装により、既製のスマートフォンに容易に導入されうるという利点がある。ただし、現状の実装では、電力消費が大きいという課題がある。これは、空中人差し指ジェスチャの事前動作も背面カメラ画像を用いて認識することから、背面カメラを常時起動する必要があるためである。この課題を解決するアプローチとして、事前動作をスマートフォン内蔵の加速度センサおよびジャイロセンサを用いて認識することが挙げられる。事前動作を背面カメラと異なるセンサにて認識することにより、空中人差し指ジェスチャを認識する時にのみ背面カメラを起動する実装が可能になるため、電力消費を抑えることができる。この実現のために、加速度センサおよびジャイロセンサを用いて事前動作を認識できるか確認する必要がある。

空中人差し指ジェスチャを認識するシステムの実装として、CNNモデルの特徴量にMHIと異なる画像を使用する実装が考えられる。例えば、ジェスチャ実行中に取得した複数枚の画像について、単に並べた画像またはそれぞれの画像の画素を1チャンネルとした複数チャンネルの画像などが考えられる。また、実装に背面カメラを用いる方針の他に、スマートフォンの背面にて追加のセンサおよび装置を取り付けるという方針が考えられる。例えば、受動型の光学センサ、測距センサ、および先行研究 [YWV<sup>+</sup>19] のような3Dプリンタにて印刷した装置などが考えられる。これらの利用可能性を調査するために、実際に実装して認識精度の評価を行う必要がある。

スマートフォンへの操作は、把持姿勢を安定させながら実行可能であることが理想である。一方で、実験にて参加者が空中人差し指ジェスチャを実行する際に、把持姿勢を変更する場合があった。この変更により、不安定な把持姿勢にて空中人差し指ジェスチャによる操作を

行った可能性がある。よって、把持姿勢の変更による、空中人差し指ジェスチャによる操作の使いやすさへの影響を評価する必要がある。例えば、3.3節に示したような、空中人差し指ジェスチャによる操作が可能なアプリケーションを実装し、既存のタッチジェスチャと把持姿勢の安定度および実行速度を比較するという評価方法が考えられる。また、スマートフォンを使用する際にユーザは様々な姿勢をとることが予想されるため、立位時、座位時、および歩行時などの様々な姿勢にて、評価する必要がある。

## 第9章 おわりに

本論文では、片手使用時の操作の種類を拡張するために、空中人差し指ジェスチャによる操作手法を示した。空中人差し指ジェスチャによる操作は、スマートフォン背面後方の空中を用いるため、既存のタッチジェスチャによる操作、タッチジェスチャを拡張した操作、および背面上における操作と共存できる操作である。つまり、既存の操作と同時にスマートフォンへ導入することにより、片手使用時の操作の種類をさらに拡張しうる。この操作の実現のため、7種類の空中人差し指ジェスチャを設計した。また、本論文にて、ショートカットに用いる利用例を示した。さらに、背面カメラ画像から取得される指の動作に関する情報を用いて、空中人差し指ジェスチャを認識するシステムを2種類試作した。まず、ジェスチャ中に取得された画像内にて人差し指の領域が移動することに基づいて、閾値にて空中人差し指ジェスチャを認識するシステムを試作した。この試作システムの認識精度の評価を行ったところ、2種類の空中人差し指ジェスチャの認識精度は66.2%であり、また、その認識精度は照明条件に影響を受けることが分かった。さらに、MHIを特徴量とするCNNモデルを使用して空中人差し指ジェスチャを認識するシステムを試作した。認識精度の評価を行ったところ、7種類の空中人差し指ジェスチャの認識精度は、5分割交差検証にて平均79.9%、leave-one-participant-out交差検証にて平均66.7%であり、実用には耐え難い精度であることが示された。そのため、同じ動作を繰り返すジェスチャを認識するためのシステムの改良を考案した。改良に先立ち、ジェスチャを4種類に限定して認識精度の評価を行ったところ、認識精度は、5分割交差検証にて平均87.9%、leave-one-participant-out交差検証にて平均78.3%であり、7種類の場合に比べて精度が向上することが示された。

## 謝辞

本研究を進めるにあたり、志築文太郎先生、高橋伸先生、川口一画先生には多大なご意見とご指導を頂きました。心から感謝いたします。特に、主指導教員である志築文太郎先生には、研究の進め方、論文執筆、発表の方法といった研究に関することのみならず、日常生活における研究者の心構えなど多岐に渡るご指導を頂きました。先生の熱心なご指導のおかげで、これまでの研究成果の発表および本論文の執筆ができました。ここに深く感謝の意を表します。

インタラクティブプログラミング研究室の後輩、同輩、先輩方には研究生生活においてお世話になりました。特に WAVE チームの皆様には、チームゼミにおけるご意見や論文の添削といった研究に関する多くのご支援に加えて、研究室においても研究の進め方や研究生生活について多くのご助言を頂きました。研究室の皆様のおかげで、3年間充実した研究生生活を送ることができました。深く感謝いたします。

最後に、学生生活においてお世話になった皆様、そして、私の学生生活を支えて頂いた家族に深く感謝を申し上げます。

## 参考文献

- [App21a] Apple Inc. Access Control Center on your iPhone and iPod touch - Apple Support, 2021. <https://support.apple.com/HT202769> (最終参照日:2022年01月07日).
- [App21b] Apple Inc. View and respond to notifications on iPhone - Apple Support, 2021. <https://support.apple.com/guide/iphone/view-and-respond-to-notifications-iph6534c01bc/15.0/ios/15.0> (最終参照日:2022年01月07日).
- [BC09] Patrick Baudisch and Gerry Chu. Back-of-Device Interaction Allows Creating Very Small Touch Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pp. 1923–1932, New York, NY, USA, 2009. ACM.
- [BD01] Aaron F. Bobick and James W. Davis. The Recognition of Human Movement Using Temporal Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 3, pp. 257–267, 2001.
- [BLC<sup>+</sup>12] Sebastian Boring, David Ledo, Xiang ‘Anthony’ Chen, Nicolai Marquardt, Anthony Tang, and Saul Greenberg. The Fat Thumb: Using the Thumb’s Contact Size for Single-handed Mobile Interaction. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pp. 39–48, New York, NY, USA, 2012. ACM.
- [BLO14] Joanna Bergstrom-Lehtovirta and Antti Oulasvirta. Modeling the Functional Area of the Thumb on Mobile Touchscreen Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 1991–2000, New York, NY, USA, 2014. ACM.
- [BZFR11] Matthias Baldauf, Sebastian Zambanini, Peter Fröhlich, and Peter Reichl. Markerless Visual Fingertip Detection for Natural Mobile Device Interaction. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pp. 539–544, New York, NY, USA, 2011. ACM.
- [GL16] Emilio Granell and Luis A. Leiva. Less Is More: Efficient Back-of-Device Tap Input Detection Using Built-in Smartphone Sensors. In *Proceedings of the 2016 ACM Inter-*



- national Conference on Interactive Surfaces and Spaces*, ISS '16, pp. 5–11, New York, NY, USA, 2016. ACM.
- [GL17] Emilio Granell and Luis A. Leiva.  $\beta$ Tap: Back-of-Device Tap Input with Built-in Sensors. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, pp. 52:1–52:6, New York, NY, USA, 2017. ACM.
- [Goo21a] Google LLC. Change settings quickly on your Android phone - Android Help, 2021. <https://support.google.com/android/answer/9083864> (最終参照日: 2022年01月07日) .
- [Goo21b] Google LLC. Google Colaboratory, 2021. <https://colab.research.google.com/> (最終参照日: 2022年01月07日) .
- [Goo21c] Google LLC. Notifications Overview - Android Developers, 2021. <https://developer.android.com/guide/topics/ui/notifiers/notifications> (最終参照日: 2022年01月07日) .
- [HBH16] Shiori Hidaka, Tetsuaki Baba, and Paul Haimes. IndexAccess:A GUI Movement System by Back-of-Device Interaction for One-Handed Operation on a Large Screen Smartphone. *International Journal of Asia Digital Art and Design Association*, Vol. 20, No. 2, pp. 41–47, 2016.
- [HFST15] Hiroyuki Hakoda, Yoshitomo Fukatsu, Buntarou Shizuki, and Jiro Tanaka. Back-of-Device Interaction Based on the Range of Motion of the Index Finger. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, OzCHI '15, pp. 202–206, New York, NY, USA, 2015. ACM.
- [HGL14] Seongkook Heo, Jiseong Gu, and Geehyuk Lee. Expanding Touch Input Vocabulary by Using Consecutive Distant Taps. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pp. 2597–2606, New York, NY, USA, 2014. ACM.
- [HIST19] Kyohei Hakka, Toshiya Isomoto, Buntarou Shizuki, and Shin Takahashi. Bounded Swipe: Swipe Gesture Inside a Target. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, OZCHI '19, pp. 312–316, New York, NY, USA, 2019. ACM.
- [HL11] Seongkook Heo and Geehyuk Lee. ForceTap: Extending the Input Vocabulary of Mobile Touch Screens by Adding Tap Gestures. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pp. 113–122, New York, NY, USA, 2011. ACM.

- [HLN<sup>+</sup>21] Michael Xuelin Huang, Yang Li, Nazneen Nazneen, Alexander Chao, and Shumin Zhai. TapNet: The Design, Training, Implementation, and Applications of a Multi-Task Learning CNN for Off-Screen Mobile Input. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 282:1–282:11, New York, NY, USA, 2021. ACM.
- [HS11] Ken Hinckley and Hyunyoung Song. Sensor Synaesthesia: Touch in Motion, and Motion in Touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 801–810, New York, NY, USA, 2011. ACM.
- [IHuIA19] Md. Zahirul Islam, Mohammad Shahadat Hossain, Raihan ul Islam, and Karl Anderson. Static Hand Gesture Recognition using Convolutional Neural Network with Data Augmentation. In *2019 Joint 8th International Conference on Informatics, Electronics Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, pp. 324–329. IEEE, 2019.
- [ITU11] *Recommendation ITU-R BT.601-7 Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios*, Geneva, Switzerland, March 2011. International Telecommunications Union.
- [JC19] Chuanrui Jiang and Jianxin Chen. A MLP-Based Skin Segmentation Method for Fingertip Recognition. In *2019 IEEE 5th International Conference on Computer and Communications (ICCC)*, pp. 1690–1695. IEEE, 2019.
- [KBCV06] Amy Karlson, Ben Bederson, and José Contreras-Vidal. Studies in One-Handed Mobile Design: Habit, Desire and Agility. In *Proceedings of 4th ERCIM Workshop on User Interfaces for All, UI4ALL*, pp. 1–10. Citeseer, 2006.
- [LKB<sup>+</sup>18] Huy Viet Le, Thomas Kosch, Patrick Bader, Sven Mayer, and Niels Henze. PalmTouch: Using the Palm as an Additional Input Modality on Commodity Smartphones. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 360:1–360:13, New York, NY, USA, 2018. ACM.
- [LLZ12] Katrin Lasinger, Matthias Baldauf Lasinger, and Sebastian Zambanini. Evaluating Camera-based Methods for Mobile Markerless Fingertip Detection. In *Proceedings of Workshop on Smart Mobile Applications, 10th International Conference on Pervasive Computing*, SmartApps '12, pp. 1–6, 2012.
- [LMBH18] Huy Viet Le, Sven Mayer, Patrick Bader, and Niels Henze. Fingers' Range and Comfortable Area for One-Handed Smartphone Interaction Beyond the Touchscreen. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 31:1–31:12, New York, NY, USA, 2018. ACM.

- [LMH18] Huy Viet Le, Sven Mayer, and Niels Henze. InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, pp. 779–792, New York, NY, USA, 2018. ACM.
- [MDD<sup>+</sup>14] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, m. c. schraefel, and Jacob O. Wobbrock. Reducing Legacy Bias in Gesture Elicitation Studies. *Interactions*, Vol. 21, No. 3, pp. 40–45, May 2014.
- [MWW10] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. Understanding Users' Preferences for Surface Gestures. In *Proceedings of Graphics Interface 2010*, GI '10, pp. 261–268, CAN, 2010. Canadian Information Processing Society.
- [OH13] Byung-Hun Oh and Kwang-Seok Hong. Finger Gesture-based Three-Dimension Mobile User Interaction Using a Rear-facing Camera. *International Journal of Multimedia and Ubiquitous Engineering*, Vol. 8, pp. 119–130, September 2013.
- [PRS20] Neethu P.S., Suguna Ramadass, and Divya Sathish. An Efficient Method for Human Hand Gesture Detection and Recognition Using Deep Learning Convolutional Neural Networks. *Soft Computing*, Vol. 24, pp. 15239–15248, October 2020.
- [RCGS16] Romi Fadillah Rahmat, Tengku Chairunnisa, Dani Gunawan, and Opim Salim Sitompul. Skin Color Segmentation Using Multi-Color Space Threshold. In *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*, pp. 391–396. IEEE, 2016.
- [RLG09] Anne Roudaut, Eric Lecolinet, and Yves Guiard. MicroRolls: Expanding Touch-Screen Input Vocabulary by Distinguishing Rolls vs. Slides of the Thumb. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pp. 927–936, New York, NY, USA, 2009. ACM.
- [RT09] Volker Roth and Thea Turner. Bezel Swipe: Conflict-Free Scrolling and Multiple Selection on Mobile Touch Screen Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pp. 1523–1526, New York, NY, USA, 2009. ACM.
- [SD14] Karsten Seipp and Kate Devlin. BackPat: One-Handed Off-Screen Patting Gestures. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services*, MobileHCI '14, pp. 77–80, New York, NY, USA, 2014. ACM.
- [SMSJ<sup>+</sup>15] Shaikh Shawon Arefin Shimon, Sarah Morrison-Smith, Noah John, Ghazal Fahimi, and Jaime Ruiz. Exploring User-Defined Back-Of-Device Gestures for Mobile Devices.

- In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, pp. 227–232, New York, NY, USA, 2015. ACM.
- [SS21] Yusuke Sei and Buntarou Shizuki. Expanding One-Handed Input Vocabulary for Smartphone Using In-Air Gesture of Index Finger Captured by Rear Camera. In *Proceedings of the 2021 Symposium on Emerging Research from Asia and on Asian Contexts and Cultures*, AsianCHI '21, pp. 61–63, New York, NY, USA, 2021. ACM.
- [SSP<sup>+</sup>14] Jie Song, Gábor Sörös, Fabrizio Pece, Sean Ryan Fanello, Shahram Izadi, Cem Keskin, and Otmar Hilliges. In-air Gestures Around Unmodified Mobile Devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pp. 319–329, New York, NY, USA, 2014. ACM.
- [SZWX18] Ke Sun, Ting Zhao, Wei Wang, and Lei Xie. VSkin: Sensing Touch Gestures on Surfaces of Mobile Devices Using Acoustic Signals. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, MobiCom '18, pp. 591–605, New York, NY, USA, 2018. ACM.
- [WFB<sup>+</sup>07] Daniel Wigdor, Clifton Forlines, Patrick Baudisch, John Barnwell, and Chia Shen. LucidTouch: A See-through Mobile Device. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, pp. 269–278, New York, NY, USA, 2007. ACM.
- [WFZ16] Pui Chung Wong, Hongbo Fu, and Kening Zhu. Back-Mirror: Back-of-Device One-Handed Interaction on Smartphones. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*, SA '16, pp. 10:1–10:5, New York, NY, USA, 2016. ACM.
- [WMW09] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-Defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pp. 1083–1092, New York, NY, USA, 2009. ACM.
- [WYY<sup>+</sup>20] Erwin Wu, Ye Yuan, Hui-Shyong Yeo, Aaron Quigley, Hideki Koike, and Kris M. Kitani. Back-Hand-Pose: 3D Hand Pose Estimation for a Wrist-Worn Camera via Dorsum Deformation Network. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, pp. 1147–1160, New York, NY, USA, 2020. ACM.
- [XBZN21] Chang Xiao, Karl Bayer, Changxi Zheng, and Shree K. Nayar. BackTrack: 2D Back-of-Device Interaction Through Front Touchscreen. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 7:1–7:8, New York, NY, USA, 2021. ACM.

- [XHW13] Xiang Xiao, Teng Han, and Jingtao Wang. LensGesture: Augmenting Mobile Interactions with Back-of-Device Finger Gestures. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction, ICMI '13*, pp. 287–294, New York, NY, USA, 2013. ACM.
- [YHBI13] Xing-Dong Yang, Khalad Hasan, Neil Bruce, and Pourang Irani. Surround-See: Enabling Peripheral Vision on Smartphones during Active Use. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13*, pp. 291–300, New York, NY, USA, 2013. ACM.
- [YMI18] Wataru Yamada, Hiroyuki Manabe, and Daizo Ikeda. CamTrackPoint: Camera-Based Pointing Stick Using Transmitted Light Through Finger. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, UIST '18*, pp. 313–320, New York, NY, USA, 2018. ACM.
- [YWL<sup>+</sup>19] Hui-Shyong Yeo, Erwin Wu, Juyoung Lee, Aaron Quigley, and Hideki Koike. Opisthenar: Hand Poses and Finger Tapping Recognition by Observing Back of Hand Using Embedded Wrist Camera. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST '19*, pp. 963–971, New York, NY, USA, 2019. ACM.
- [YWV<sup>+</sup>19] Chun Yu, Xiaoying Wei, Shubh Vachher, Yue Qin, Chen Liang, Yueting Weng, Yizheng Gu, and Yuanchun Shi. HandSee: Enabling Full Hand Interaction on Smartphone with Front Camera-based Stereo Vision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19*, pp. 705:1–705:13, New York, NY, USA, 2019. ACM.
- [YYJ15] Hyunjin Yoo, Jungwon Yoon, and Hyunsoo Ji. Index Finger Zone: Study on Touchable Area Expandability Using Thumb and Index Finger. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, MobileHCI '15*, pp. 803–810, New York, NY, USA, 2015. ACM.
- [ZGZ<sup>+</sup>15] Cheng Zhang, Anhong Guo, Dingtian Zhang, Caleb Southern, Rosa Arriaga, and Gregory Abowd. BeyondTouch: Extending the Input Language with Built-in Sensors on Commodity Smartphones. In *Proceedings of the 20th International Conference on Intelligent User Interfaces, IUI '15*, pp. 67–77, New York, NY, USA, 2015. ACM.
- [ZGZ<sup>+</sup>16] Cheng Zhang, Anhong Guo, Dingtian Zhang, Yang Li, Caleb Southern, Rosa I. Arriaga, and Gregory D. Abowd. Beyond the Touchscreen: An Exploration of Extending Interactions on Commodity Smartphones. *ACM Transactions on Interactive Intelligent Systems*, Vol. 6, No. 2, pp. 16:1–23, August 2016.

- [Zha19] Felix Zhan. Hand Gesture Recognition with Convolution Neural Networks. In *2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI)*, pp. 295–298. IEEE, 2019.
- [河内 12] 河内まき子. AIST 日本人の手の寸法データ, 2012. <https://www.airc.aist.go.jp/dhrt/hand/index.html> (最終参照日: 2022 年 01 月 07 日) .
- [清 21] 清佑輔, 志築文太郎. スマートフォンの背面カメラを用いて認識可能な空中人差し指ジェスチャによるショートカットの検討. 第 29 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2021) . 日本ソフトウェア科学会, 2021. 2 pages.
- [清 22] 清佑輔, 志築文太郎. スマートフォンの背面カメラ画像を用いた空中人差し指ジェスチャ認識の精度評価. 情報処理学会研究報告, 第 2022-HCI-196 巻. 情報処理学会, 2022. 8 pages.
- [石井 18] 石井晃, 阿達拓也, 島佳吾, 中前秀太, 志築文太郎, 高橋伸. カードボード VR のための空中操作を用いたターゲット選択手法. *コンピュータ ソフトウェア*, Vol. 35, No. 2, pp. 64–78, 2018.
- [総務 21] 総務省. 令和 3 年版情報通信白書, 2021. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/html/nd100000.html> (最終参照日: 2022 年 01 月 07 日) .

# 著者論文リスト

## 本論文に関する論文および発表

- 査読あり国際会議論文
  1. Yusuke Sei, Buntarou Shizuki. Expanding One-Handed Input Vocabulary for Smartphone Using In-Air Gesture of Index Finger Captured by Rear Camera. CHI 2021 symposia on Asian CHI Symposium: HCI Research from Asia and on Asian Contexts and Cultures, ACM, pp. 61–63, May 2021.
- 査読あり国内会議論文
  1. 清佑輔, 志築文太郎. スマートフォンの背面カメラを用いて認識可能な空中人差し指ジェスチャによるショートカットの検討. 第29回インタラクティブシステムとソフトウェアに関するワークショップ論文集 (WISS2021), 日本ソフトウェア科学会, 2021年12月, 3 pages.
- 査読なし国内会議論文
  1. 清佑輔, 志築文太郎. スマートフォンの背面カメラ画像を用いた空中人差し指ジェスチャ認識の精度評価. 情報処理学会研究報告, 第2022-HCI-196巻, 情報処理学会, 2022年1月, 8 pages.
  2. 清佑輔, 船越南斗, 志築文太郎. スマートフォン背面の人差し指を用いた片手操作向け入力拡張手法の検討. 第28回インタラクティブシステムとソフトウェアに関するワークショップ論文集 (WISS2020), 日本ソフトウェア科学会, 2020年12月, 6 pages.

## その他論文

- 査読あり国際会議論文
  1. Takuma Hidaka, Yusuke Sei, Buntarou Shizuki. Pilot Study on Notification Using Phantom Sensation on Hand. CHI 2021 symposia on Asian CHI Symposium: HCI Research from Asia and on Asian Contexts and Cultures, ACM, pp. 157–159, May 2021.

2. Yusuke Sei, Minto Funakoshi, Buntarou Shizuki. Expanding Input Vocabulary Using Index Finger on and Above Back of Smartphone. CHI 2020 symposia on Asian CHI Symposium: HCI Research from Asia and on Asian Contexts and Cultures, ACM, April 2020, 4 pages.
  3. Yusuke Sei, Minto Funakoshi, Buntarou Shizuki. Expanding One-Handed Touch Input Vocabulary Using Index Finger on and Above Back-of-Device. 31st Australian Conference on Human-Computer-Interaction (OzCHI '19), ACM, pp. 585–589, December 2019.
- 査読なし国内会議論文
1. 齋藤圭汰, 清佑輔, 志築文太郎. Dial-B-Locks: ダイヤル付きブロックに基づく長さの換えられる入力インタフェース. インタラクション 2021 論文集, 情報処理学会, pp. 557–558, 2021 年 3 月.
  2. 日高拓真, 清佑輔, 志築文太郎. 手・指におけるファントムセンセーションを用いた通知のための事前実験設計. インタラクション 2021 論文集, 情報処理学会, pp. 607–610, 2021 年 3 月.
  3. 柳原直貴, 清佑輔, 志築文太郎. 磁力の異なるブロックを用いた種類の認識可能なブロック構造物認識システム. インタラクション 2020 論文集, 情報処理学会, pp. 236–241, 2020 年 3 月.
  4. 清佑輔, 舩越南斗, 志築文太郎. 片手操作時の端末背面における人差し指を用いたタッチ入力拡張手法の提案. 第 27 回インタラクティブシステムとソフトウェアに関するワークショップ論文集 (WISS2019), 日本ソフトウェア科学会, 2019 年 9 月, 2 pages.



## 付録A 混同行列

付録として、第7章にて示した実験において得られた混同行列を掲載する。

### A.1 結果1：7種類の空中人差し指ジェスチャの認識精度

7種類の空中人差し指ジェスチャの認識精度に関して、交差検証した際に分割毎にて混同行列が得られた。

#### A.1.1 5分割交差検証

5分割交差検証においては、データを5個に分割した。それぞれを分割1-5として、検証データに用いた際の混同行列を表A.1-A.5に示す。

表A.1: 5分割交差検証にて得られた、分割1を検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 99.7% | 0.0%    | 0.3%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 1.7%  | 93.8%   | 0.0%  | 1.3%  | 0.0%  | 0.3%   | 1.3%    | 1.5%     |
| Bend     | 7.3%  | 0.0%    | 72.7% | 8.7%  | 5.0%  | 1.0%   | 2.5%    | 2.8%     |
| Bend2    | 0.0%  | 0.0%    | 5.0%  | 65.2% | 1.7%  | 14.2%  | 0.0%    | 14.0%    |
| Swipe    | 10.0% | 0.0%    | 21.8% | 3.7%  | 55.0% | 8.0%   | 0.0%    | 1.5%     |
| Swipe2   | 0.0%  | 0.0%    | 1.3%  | 16.7% | 0.3%  | 77.7%  | 0.0%    | 4.0%     |
| Release  | 0.0%  | 3.7%    | 2.7%  | 1.3%  | 0.0%  | 0.0%   | 82.0%   | 10.3%    |
| Release2 | 0.0%  | 0.5%    | 0.0%  | 6.5%  | 0.0%  | 0.2%   | 1.5%    | 91.3%    |

表 A.2: 5 分割交差検証にて得られた, 分割 2 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 98.7% | 0.0%    | 0.8%  | 0.3%  | 0.0%  | 0.0%   | 0.0%    | 0.2%     |
| Stretch  | 1.7%  | 93.0%   | 0.0%  | 0.0%  | 1.7%  | 0.0%   | 3.7%    | 0.0%     |
| Bend     | 2.8%  | 0.0%    | 74.2% | 12.0% | 7.8%  | 2.0%   | 0.7%    | 0.5%     |
| Bend2    | 0.0%  | 0.0%    | 3.2%  | 78.0% | 0.2%  | 8.2%   | 1.0%    | 9.5%     |
| Swipe    | 7.8%  | 0.0%    | 19.0% | 5.8%  | 61.5% | 4.2%   | 0.0%    | 1.7%     |
| Swipe2   | 0.0%  | 0.3%    | 1.2%  | 21.8% | 2.5%  | 71.5%  | 0.0%    | 2.7%     |
| Release  | 0.0%  | 0.0%    | 1.7%  | 0.0%  | 0.0%  | 0.0%   | 94.3%   | 4.0%     |
| Release2 | 0.0%  | 0.8%    | 0.0%  | 11.3% | 0.0%  | 1.3%   | 3.5%    | 83.0%    |

表 A.3: 5 分割交差検証にて得られた, 分割 3 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 99.0% | 0.0%    | 0.7%  | 0.0%  | 0.3%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%  | 95.8%   | 0.0%  | 0.2%  | 0.0%  | 0.0%   | 3.3%    | 0.7%     |
| Bend     | 4.2%  | 0.0%    | 82.0% | 7.7%  | 3.8%  | 0.0%   | 0.2%    | 2.2%     |
| Bend2    | 0.0%  | 0.0%    | 4.3%  | 66.3% | 0.0%  | 10.8%  | 0.5%    | 18.0%    |
| Swipe    | 8.2%  | 0.0%    | 28.0% | 6.3%  | 54.5% | 2.7%   | 0.0%    | 0.3%     |
| Swipe2   | 0.0%  | 0.0%    | 0.0%  | 27.8% | 1.7%  | 65.8%  | 0.0%    | 4.7%     |
| Release  | 0.5%  | 2.5%    | 4.0%  | 0.0%  | 0.5%  | 0.0%   | 84.5%   | 8.0%     |
| Release2 | 0.0%  | 0.3%    | 0.3%  | 0.7%  | 0.0%  | 0.0%   | 5.8%    | 92.8%    |

表 A.4: 5 分割交差検証にて得られた, 分割 4 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%   | 92.8%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 6.0%    | 1.2%     |
| Bend     | 5.5%   | 0.0%    | 78.5% | 9.2%  | 3.0%  | 1.8%   | 0.0%    | 2.0%     |
| Bend2    | 0.0%   | 0.0%    | 1.5%  | 71.8% | 0.0%  | 12.5%  | 0.3%    | 13.8%    |
| Swipe    | 8.8%   | 0.3%    | 34.3% | 2.8%  | 42.2% | 11.2%  | 0.0%    | 0.3%     |
| Swipe2   | 0.0%   | 0.0%    | 0.0%  | 20.8% | 1.5%  | 72.2%  | 0.0%    | 5.5%     |
| Release  | 0.3%   | 0.2%    | 3.7%  | 0.3%  | 0.2%  | 0.0%   | 78.8%   | 16.5%    |
| Release2 | 0.0%   | 0.3%    | 0.0%  | 6.7%  | 0.0%  | 0.0%   | 0.8%    | 92.2%    |

表 A.5: 5 分割交差検証にて得られた, 分割 5 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 98.3% | 1.7%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%  | 96.5%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 1.8%    | 1.7%     |
| Bend     | 6.2%  | 0.0%    | 70.7% | 6.3%  | 11.7% | 0.7%   | 1.3%    | 3.2%     |
| Bend2    | 0.0%  | 0.0%    | 4.2%  | 83.7% | 0.0%  | 6.5%   | 0.0%    | 5.7%     |
| Swipe    | 11.7% | 0.7%    | 20.7% | 3.5%  | 56.2% | 6.0%   | 1.0%    | 0.3%     |
| Swipe2   | 0.0%  | 0.0%    | 2.0%  | 29.3% | 1.5%  | 66.0%  | 0.5%    | 0.7%     |
| Release  | 0.0%  | 3.2%    | 1.7%  | 1.3%  | 0.0%  | 0.5%   | 82.3%   | 11.0%    |
| Release2 | 0.0%  | 0.0%    | 0.0%  | 12.0% | 0.0%  | 1.5%   | 4.7%    | 81.8%    |

### A.1.2 leave-one-participant-out 交差検証

leave-one-participant-out 交差検証においては、データを参加者毎の 10 個に分割した。参加者 P1-P10 から得られたデータをそれぞれ検証データに用いた際の混同行列を表 A.6-A.15 に示す。

表 A.6: leave-one-participant-out 交差検証にて得られた、P1 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 98.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 2.0%    | 0.0%     |
| Stretch  | 0.0%  | 96.7%   | 0.0%  | 1.0%  | 0.0%  | 2.3%   | 0.0%    | 0.0%     |
| Bend     | 0.0%  | 0.0%    | 24.3% | 14.3% | 53.0% | 4.3%   | 4.0%    | 0.0%     |
| Bend2    | 0.0%  | 0.0%    | 1.3%  | 36.7% | 0.3%  | 24.7%  | 0.0%    | 37.0%    |
| Swipe    | 0.0%  | 0.0%    | 0.0%  | 2.0%  | 88.3% | 9.7%   | 0.0%    | 0.0%     |
| Swipe2   | 0.0%  | 0.0%    | 0.0%  | 23.0% | 0.7%  | 65.7%  | 0.0%    | 10.7%    |
| Release  | 0.0%  | 0.7%    | 0.0%  | 2.3%  | 0.3%  | 0.0%   | 67.3%   | 29.3%    |
| Release2 | 0.0%  | 0.3%    | 0.0%  | 25.0% | 0.0%  | 14.0%  | 2.0%    | 58.7%    |

表 A.7: leave-one-participant-out 交差検証にて得られた、P2 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%   | 53.3%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 44.3%   | 2.3%     |
| Bend     | 3.7%   | 0.0%    | 56.0% | 16.3% | 3.7%  | 1.7%   | 7.7%    | 11.0%    |
| Bend2    | 0.0%   | 0.0%    | 2.7%  | 30.3% | 0.0%  | 10.3%  | 18.0%   | 38.7%    |
| Swipe    | 10.3%  | 0.0%    | 14.3% | 1.7%  | 53.3% | 10.7%  | 0.3%    | 9.3%     |
| Swipe2   | 0.0%   | 0.0%    | 1.0%  | 13.0% | 5.7%  | 61.0%  | 0.0%    | 19.3%    |
| Release  | 9.7%   | 3.3%    | 18.7% | 5.0%  | 10.7% | 1.3%   | 43.7%   | 7.7%     |
| Release2 | 2.3%   | 0.0%    | 0.0%  | 19.7% | 0.0%  | 10.0%  | 12.3%   | 55.7%    |

表 A.8: leave-one-participant-out 交差検証にて得られた, P3 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%   | 99.7%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.3%    | 0.0%     |
| Bend     | 0.0%   | 0.0%    | 83.3% | 0.0%  | 13.3% | 0.0%   | 3.3%    | 0.0%     |
| Bend2    | 0.0%   | 0.0%    | 12.7% | 55.0% | 1.3%  | 7.3%   | 5.0%    | 18.7%    |
| Swipe    | 0.0%   | 0.0%    | 7.7%  | 4.0%  | 67.0% | 13.7%  | 5.7%    | 2.0%     |
| Swipe2   | 0.0%   | 0.0%    | 0.0%  | 18.0% | 0.3%  | 81.7%  | 0.0%    | 0.0%     |
| Release  | 0.0%   | 14.0%   | 17.3% | 0.3%  | 0.0%  | 0.0%   | 64.3%   | 4.0%     |
| Release2 | 0.0%   | 0.0%    | 0.0%  | 17.3% | 0.0%  | 3.3%   | 9.3%    | 70.0%    |

表 A.9: leave-one-participant-out 交差検証にて得られた, P4 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 4.0%   | 96.0%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Bend     | 3.0%   | 0.0%    | 57.7% | 14.7% | 24.0% | 0.7%   | 0.0%    | 0.0%     |
| Bend2    | 0.0%   | 0.0%    | 6.0%  | 33.7% | 3.7%  | 48.7%  | 0.0%    | 8.0%     |
| Swipe    | 1.3%   | 0.3%    | 9.3%  | 4.7%  | 71.3% | 11.0%  | 0.0%    | 2.0%     |
| Swipe2   | 0.0%   | 0.0%    | 0.0%  | 4.0%  | 14.0% | 78.0%  | 0.0%    | 4.0%     |
| Release  | 0.7%   | 1.7%    | 0.0%  | 0.7%  | 0.0%  | 0.0%   | 92.0%   | 5.0%     |
| Release2 | 0.0%   | 0.0%    | 0.0%  | 8.0%  | 3.3%  | 0.0%   | 8.7%    | 80.0%    |

表 A.10: leave-one-participant-out 交差検証にて得られた, P5 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 96.7% | 3.3%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 8.3%  | 68.3%   | 0.7%  | 0.0%  | 7.7%  | 2.7%   | 7.7%    | 4.7%     |
| Bend     | 10.3% | 0.0%    | 39.0% | 1.3%  | 45.7% | 3.7%   | 0.0%    | 0.0%     |
| Bend2    | 0.0%  | 0.0%    | 0.0%  | 48.3% | 0.0%  | 13.7%  | 0.0%    | 38.0%    |
| Swipe    | 30.3% | 0.0%    | 3.7%  | 0.0%  | 65.0% | 1.0%   | 0.0%    | 0.0%     |
| Swipe2   | 0.0%  | 0.0%    | 0.7%  | 29.0% | 0.3%  | 67.3%  | 0.0%    | 2.7%     |
| Release  | 0.0%  | 11.0%   | 0.0%  | 2.0%  | 0.0%  | 0.0%   | 75.7%   | 11.3%    |
| Release2 | 0.0%  | 0.0%    | 0.0%  | 2.3%  | 0.0%  | 2.3%   | 1.7%    | 93.7%    |

表 A.11: leave-one-participant-out 交差検証にて得られた, P6 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 93.3% | 2.0%    | 0.0%  | 1.7%  | 2.0%  | 0.0%   | 0.0%    | 1.0%     |
| Stretch  | 0.0%  | 99.7%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.3%    | 0.0%     |
| Bend     | 2.3%  | 0.0%    | 36.3% | 10.7% | 42.0% | 3.0%   | 4.3%    | 1.3%     |
| Bend2    | 0.0%  | 0.0%    | 0.0%  | 12.3% | 7.0%  | 34.7%  | 0.0%    | 46.0%    |
| Swipe    | 0.7%  | 0.7%    | 6.3%  | 0.0%  | 59.7% | 18.3%  | 9.7%    | 4.7%     |
| Swipe2   | 0.0%  | 0.0%    | 0.0%  | 9.3%  | 15.0% | 38.0%  | 1.0%    | 36.7%    |
| Release  | 0.0%  | 25.0%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 73.3%   | 1.7%     |
| Release2 | 0.0%  | 0.7%    | 0.0%  | 10.7% | 0.0%  | 5.3%   | 20.3%   | 63.0%    |

表 A.12: leave-one-participant-out 交差検証にて得られた, P7 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%   | 86.3%   | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 13.7%   | 0.0%     |
| Bend     | 2.3%   | 0.0%    | 34.0% | 2.7%  | 38.0% | 20.0%  | 3.0%    | 0.0%     |
| Bend2    | 0.0%   | 0.0%    | 2.7%  | 23.0% | 0.3%  | 67.3%  | 3.0%    | 3.7%     |
| Swipe    | 7.0%   | 0.0%    | 30.3% | 2.7%  | 55.7% | 4.3%   | 0.0%    | 0.0%     |
| Swipe2   | 0.0%   | 5.0%    | 2.7%  | 20.7% | 9.3%  | 40.7%  | 2.7%    | 19.0%    |
| Release  | 0.0%   | 23.7%   | 1.7%  | 7.7%  | 0.0%  | 10.3%  | 56.0%   | 0.7%     |
| Release2 | 0.0%   | 7.0%    | 0.0%  | 28.3% | 1.7%  | 6.3%   | 21.3%   | 35.3%    |

表 A.13: leave-one-participant-out 交差検証にて得られた, P8 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%   | 100.0%  | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Bend     | 33.7%  | 4.0%    | 22.7% | 0.0%  | 34.3% | 2.7%   | 2.7%    | 0.0%     |
| Bend2    | 0.0%   | 0.0%    | 4.0%  | 32.7% | 8.3%  | 51.3%  | 0.0%    | 3.7%     |
| Swipe    | 47.0%  | 0.0%    | 2.7%  | 0.0%  | 50.3% | 0.0%   | 0.0%    | 0.0%     |
| Swipe2   | 0.0%   | 0.0%    | 2.7%  | 11.3% | 13.3% | 68.3%  | 1.0%    | 3.3%     |
| Release  | 0.7%   | 18.0%   | 2.3%  | 1.3%  | 1.0%  | 0.0%   | 76.7%   | 0.0%     |
| Release2 | 0.0%   | 0.0%    | 2.7%  | 7.0%  | 0.7%  | 2.0%   | 21.0%   | 66.7%    |

表 A.14: leave-one-participant-out 交差検証にて得られた, P9 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|--------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%   | 100.0%  | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Bend     | 1.3%   | 0.0%    | 45.7% | 1.0%  | 44.3% | 7.7%   | 0.0%    | 0.0%     |
| Bend2    | 0.0%   | 0.0%    | 0.0%  | 32.7% | 2.7%  | 62.0%  | 0.0%    | 2.7%     |
| Swipe    | 0.0%   | 0.0%    | 37.3% | 8.0%  | 50.3% | 4.3%   | 0.0%    | 0.0%     |
| Swipe2   | 0.0%   | 0.0%    | 3.3%  | 10.7% | 7.3%  | 78.3%  | 0.0%    | 0.3%     |
| Release  | 0.0%   | 3.7%    | 4.3%  | 2.7%  | 0.0%  | 0.0%   | 83.3%   | 6.0%     |
| Release2 | 0.0%   | 0.0%    | 2.7%  | 58.7% | 0.0%  | 3.0%   | 14.3%   | 21.3%    |

表 A.15: leave-one-participant-out 交差検証にて得られた, P10 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe | Swipe2 | Release | Release2 |
|----------|-------|---------|-------|-------|-------|--------|---------|----------|
| 未実行      | 93.3% | 5.7%    | 1.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Stretch  | 0.0%  | 100.0%  | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 0.0%    | 0.0%     |
| Bend     | 0.0%  | 0.0%    | 68.3% | 1.3%  | 23.0% | 0.7%   | 4.0%    | 2.7%     |
| Bend2    | 0.0%  | 0.0%    | 1.0%  | 37.7% | 0.0%  | 18.7%  | 0.0%    | 42.7%    |
| Swipe    | 0.0%  | 0.3%    | 24.3% | 0.0%  | 68.7% | 6.7%   | 0.0%    | 0.0%     |
| Swipe2   | 0.0%  | 0.0%    | 0.3%  | 10.3% | 2.3%  | 83.3%  | 0.0%    | 3.7%     |
| Release  | 0.0%  | 4.7%    | 0.0%  | 0.0%  | 0.0%  | 0.0%   | 87.7%   | 7.7%     |
| Release2 | 0.0%  | 0.0%    | 0.0%  | 4.7%  | 0.0%  | 4.7%   | 1.7%    | 89.0%    |

## A.2 結果2：4種類の空中人差し指ジェスチャの認識精度

4種類の空中人差し指ジェスチャの認識精度に関して、交差検証した際に分割毎にて混同行列が得られた。

### A.2.1 5分割交差検証

5分割交差検証においては、データを5個に分割した。それぞれを分割1-5として、検証データに用いた際の混同行列を表A.16-A.20に示す。

表A.16: 4種類の空中人差し指ジェスチャに限定した5分割交差検証にて得られた、分割1を検証データとした際の混同行列。横軸が真値，縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Swipe | Release |
|----------|-------|---------|-------|-------|---------|
| 未実行      | 99.3% | 0.7%    | 0.0%  | 0.0%  | 0.0%    |
| Stretch  | 1.7%  | 93.7%   | 0.0%  | 0.0%  | 4.7%    |
| Bend     | 6.5%  | 0.0%    | 79.5% | 11.7% | 2.3%    |
| Swipe    | 10.0% | 0.8%    | 17.5% | 69.8% | 1.8%    |
| Release  | 0.0%  | 5.5%    | 4.0%  | 0.0%  | 90.5%   |

表A.17: 4種類の空中人差し指ジェスチャに限定した5分割交差検証にて得られた、分割2を検証データとした際の混同行列。横軸が真値，縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%   | 93.2%   | 0.5%  | 1.2%  | 5.2%  |
| Bend     | 4.3%   | 0.0%    | 88.8% | 5.5%  | 1.3%  |
| Swipe    | 10.8%  | 0.0%    | 12.2% | 76.2% | 0.8%  |
| Release  | 0.0%   | 2.5%    | 1.3%  | 0.0%  | 96.2% |



表 A.18: 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 3 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe |
|----------|-------|---------|-------|-------|-------|
| 未実行      | 98.0% | 0.0%    | 1.3%  | 0.0%  | 0.7%  |
| Stretch  | 0.8%  | 93.2%   | 0.0%  | 0.5%  | 5.5%  |
| Bend     | 5.2%  | 0.0%    | 77.2% | 11.0% | 6.7%  |
| Swipe    | 7.2%  | 0.0%    | 27.8% | 64.2% | 0.8%  |
| Release  | 0.0%  | 5.0%    | 2.0%  | 0.0%  | 93.0% |

表 A.19: 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 4 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe |
|----------|-------|---------|-------|-------|-------|
| 未実行      | 99.7% | 0.0%    | 0.0%  | 0.3%  | 0.0%  |
| Stretch  | 0.0%  | 97.7%   | 0.0%  | 0.0%  | 2.3%  |
| Bend     | 5.3%  | 0.0%    | 77.3% | 15.0% | 2.3%  |
| Swipe    | 8.2%  | 0.2%    | 10.8% | 79.3% | 1.5%  |
| Release  | 0.5%  | 2.0%    | 5.2%  | 0.5%  | 91.8% |

表 A.20: 4 種類の空中人差し指ジェスチャに限定した 5 分割交差検証にて得られた, 分割 5 を検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe |
|----------|-------|---------|-------|-------|-------|
| 未実行      | 99.7% | 0.0%    | 0.0%  | 0.3%  | 0.0%  |
| Stretch  | 0.0%  | 97.7%   | 0.0%  | 0.0%  | 2.3%  |
| Bend     | 5.3%  | 0.0%    | 77.3% | 15.0% | 2.3%  |
| Swipe    | 8.2%  | 0.2%    | 10.8% | 79.3% | 1.5%  |
| Release  | 0.5%  | 2.0%    | 5.2%  | 0.5%  | 91.8% |

## A.2.2 leave-one-participant-out 交差検証

leave-one-participant-out 交差検証においては、データを参加者毎の 10 個に分割した。参加者 P1-P10 から得られたデータをそれぞれ検証データに用いた際の混同行列を表 A.21-A.30 に示す。

表 A.21: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P1 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Swipe | Release |
|----------|-------|---------|-------|-------|---------|
| 未実行      | 98.0% | 0.0%    | 0.0%  | 0.0%  | 2.0%    |
| Stretch  | 0.0%  | 96.7%   | 0.7%  | 2.7%  | 0.0%    |
| Bend     | 0.0%  | 0.0%    | 95.7% | 3.7%  | 0.7%    |
| Swipe    | 0.0%  | 0.0%    | 59.0% | 41.0% | 0.0%    |
| Release  | 0.0%  | 0.0%    | 0.3%  | 0.0%  | 99.7%   |

表 A.22: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P2 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%   | 37.0%   | 1.0%  | 0.0%  | 62.0% |
| Bend     | 1.7%   | 0.7%    | 68.3% | 4.3%  | 25.0% |
| Swipe    | 8.3%   | 1.0%    | 44.3% | 39.3% | 7.0%  |
| Release  | 1.3%   | 4.7%    | 23.7% | 9.0%  | 61.3% |

表 A.23: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P3 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%   | 90.0%   | 0.0%  | 0.0%  | 10.0% |
| Bend     | 0.0%   | 0.0%    | 96.3% | 1.7%  | 2.0%  |
| Swipe    | 0.0%   | 0.0%    | 62.3% | 23.0% | 14.7% |
| Release  | 0.0%   | 0.3%    | 13.7% | 0.0%  | 86.0% |

表 A.24: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P4 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 3.3%   | 96.3%   | 0.0%  | 0.0%  | 0.3%  |
| Bend     | 3.0%   | 0.7%    | 78.0% | 15.0% | 3.3%  |
| Swipe    | 1.3%   | 0.3%    | 30.7% | 67.7% | 0.0%  |
| Release  | 0.7%   | 3.7%    | 0.7%  | 0.0%  | 95.0% |

表 A.25: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P5 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 4.7%   | 68.0%   | 0.0%  | 10.0% | 17.3% |
| Bend     | 9.3%   | 0.0%    | 73.0% | 15.0% | 2.7%  |
| Swipe    | 32.3%  | 0.3%    | 17.7% | 49.0% | 0.7%  |
| Release  | 0.0%   | 12.0%   | 0.0%  | 0.0%  | 88.0% |

表 A.26: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P6 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe |
|----------|-------|---------|-------|-------|-------|
| 未実行      | 93.3% | 3.3%    | 2.0%  | 0.7%  | 0.7%  |
| Stretch  | 0.0%  | 98.7%   | 0.0%  | 0.0%  | 1.3%  |
| Bend     | 4.0%  | 3.3%    | 66.7% | 23.0% | 3.0%  |
| Swipe    | 0.7%  | 1.0%    | 36.7% | 58.3% | 3.3%  |
| Release  | 0.0%  | 44.3%   | 0.0%  | 0.0%  | 55.7% |

表 A.27: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた、P7 のデータを検証データとした際の混同行列。横軸が真値、縦軸が予測値を表す。

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%   | 88.0%   | 0.0%  | 0.0%  | 12.0% |
| Bend     | 2.3%   | 0.0%    | 60.0% | 24.0% | 13.7% |
| Swipe    | 7.0%   | 0.0%    | 42.0% | 50.3% | 0.7%  |
| Release  | 0.3%   | 16.0%   | 4.3%  | 4.7%  | 74.7% |

表 A.28: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P8 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%   | 100.0%  | 0.0%  | 0.0%  | 0.0%  |
| Bend     | 32.7%  | 1.0%    | 49.3% | 11.3% | 5.7%  |
| Swipe    | 46.3%  | 0.0%    | 20.0% | 33.7% | 0.0%  |
| Release  | 0.0%   | 14.0%   | 5.0%  | 0.0%  | 81.0% |

表 A.29: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P9 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行    | Stretch | Bend  | Bend2 | Swipe |
|----------|--------|---------|-------|-------|-------|
| 未実行      | 100.0% | 0.0%    | 0.0%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%   | 100.0%  | 0.0%  | 0.0%  | 0.0%  |
| Bend     | 1.3%   | 0.0%    | 61.3% | 37.3% | 0.0%  |
| Swipe    | 0.0%   | 0.0%    | 49.7% | 48.0% | 2.3%  |
| Release  | 0.0%   | 0.7%    | 0.0%  | 0.0%  | 99.3% |

表 A.30: 4 種類の空中人差し指ジェスチャに限定した leave-one-participant-out 交差検証にて得られた, P10 のデータを検証データとした際の混同行列. 横軸が真値, 縦軸が予測値を表す.

| 真値 \ 予測値 | 未実行   | Stretch | Bend  | Bend2 | Swipe |
|----------|-------|---------|-------|-------|-------|
| 未実行      | 95.0% | 4.3%    | 0.7%  | 0.0%  | 0.0%  |
| Stretch  | 0.0%  | 100.0%  | 0.0%  | 0.0%  | 0.0%  |
| Bend     | 0.0%  | 0.0%    | 84.3% | 0.0%  | 15.7% |
| Swipe    | 0.0%  | 0.3%    | 20.7% | 79.0% | 0.0%  |
| Release  | 0.0%  | 6.7%    | 0.0%  | 0.0%  | 93.3% |