

## MediaPipe およびタイピング動画像を用いた行動特徴抽出手法の検討

### Consideration of Behavioral Feature Extraction Method Using MediaPipe and Typing Video Images

○小野寺将太<sup>1</sup>, 廣田大輝<sup>2</sup>, 塚本新<sup>3</sup>

\*Shota Onodera<sup>1</sup>, Daiki Hirota<sup>2</sup>, Arata Tsukamoto<sup>3</sup>

Abstract: We extracted indirect hand points from typing video images using MediaPipe. The results showed that typing by the same person was similar under human recognition, suggesting that typing contains individual characteristics. In addition, it was suggested that individual differences could be extracted from the shape of the trajectory, the range of movement, and the measurement time.

#### 1. はじめに

近年, 個人情報や機密データのオンライン取り扱いが増加している. 従来の認証法としてパスワードやピンコードなどの情報を用いた認証が多岐にわたって利用されているが, これらの方法は本人自身の認証行為に基づかない不正アクセスのリスクがある. また指紋, 虹彩, 顔認証などの個人特徴を用いた生体認証技術が注目されているが, 物理的に偽造された生体情報も受容しうる可能性があり, 脆弱性が存在する<sup>[1]</sup>. そこで行動特徴を用いる認証法に着目した. 行動特徴は身体構造にも起因するが, 個人の経験や思想, 身体の内部構造などに依存し, 時系列の構造を持ち, 物質的情報として視認困難な, 複製難度の高い特徴である. 本研究では行動特徴の内, 特に動作中に無意識的に含まれる情報に着目し, タイピング動画像における行動特徴の抽出を検討する.

#### 2. 主要技術要素

##### 2.1 手指動作追跡

タイピング時の行動特徴を用いる認証技術のひとつとしてキーストローク認証があり, パソコンのキーボードでキー入力をしたときのキーストロークの速さ, 指の動きのパターンなどから個人的特徴を抽出し, 本人を識別する認証方法である. この認証方法の利点として認証行為中の時系列データに基づく特徴での認証が可能である点, スマートフォンなどによって撮影した動画でも認証可能な簡便さがあり, 特定のスマートフォンで認証するなどにより生体情報を用いた多要素認証の幅が広がると考える. 本検討ではキータイピング時に使用する手指の選択等を含むキーボード上における二次元的手指動作に注目し, 動画像での手指の間接点の追跡を行った.

##### 2.2 MediaPipe

代表的な姿勢・手指形推定ソフトとして MediaPipe<sup>[2]</sup> と OpenPose<sup>[3]</sup>が挙げられる. MediaPipe が OpenPose と比較して GPU を必要としない点, 認識速度が速いという点から本研究では Google 社が開発しているオーブ

ソンスの機械学習ライブラリである MediaPipe を用いた. MediaPipe は動画像のフレームごとに手指の特徴点を認識する. Fig.1 に示すように手指の認識では片手で計 21 個の特徴点を認識することができる.



Figure1. Hand landmarks of MediaPipe Hands

#### 3. 提案手法

本検討ではキータイピングの動画像から手指の間接点を抽出し, タイピングにおいて人差し指の先端が最も使用頻度が高いと考え, Fig.1 の landmark'8'に当たる特徴点に着目し, 間接点の分散を求めることで個人でのキータイピングのパターンに安定性, 一貫性があるかどうか, 特定のキーをより強調して打つかなどのが考えられるため landmark'8'のキータイピングの軌跡より特徴抽出を行った.

#### 4. 実験方法

動画撮影環境としては机上に設置したノートパソコンの直上からキーボードと両手の全指が写り, 両手の特徴点を検出できるようにスマートフォンにて 53 字のテキストをタイピングしている様子の撮影を行った. そしてフレームレートは 30fps として, 各フレームごとに特徴点の抽出を行った. また異なる測定者間で入力するテキストはすべて同文とし, 実験した. また本研究における仮定である無意識的に含まれる情報について測定者の間接点の軌跡から評価を行った.

#### 5. 結果および考察

##### 5.1 タイピングの軌跡

Fig.2, 3, 4 に同一測定者での landmark'8'の間接点のタイピングでの軌跡を示す. まず測定時間について 3 回の測定での平均測定時間は 24.62 [s]であり, Fig.2 のみ約 2.5 [s]測定時間が異なっていたが, いずれの結果も平均測定時間との差が 2 [s]以内であった. また軌

1 : 日大理工・学部・電子 2 : 日大理工・博士課程前期・情報 3 : 日大理工・教員・電子

跡の形状については観測者の識別にて LeftHand, RightHand 共に類似した形状を確認した. そして動作範囲については RightHand の y 座標のみ 30 [px]程度の違いがあったが LeftHand は x 座標, y 座標共に 10 [px]以内の差異であった. まず測定時間について Fig.2 が違う原因として本測定において使用したテキストは全て同じものを使用しており Fig.2 の測定は初回の測定であり, Fig.3,4 に比べ測定時間に大きな差が認められる. また軌跡の形状, 動作範囲についてはタイピング時の使う指と押すキーが複数回の測定で不変であり, 無意識に同じように手指を動かしていると考ええる. また Fig.3 の動作範囲が異なっている原因としてはタイプミスにより軌跡が右上に描画されたことにより動作範囲が拡大したと考え, タイプミスの回数によってタイピングの軌跡における分布, 動作範囲が変化するといえる. そして測定回数により軌跡の形状が不変であったことからタイピングの手指の動きに無意識的な行動特徴を含んでいると示唆される.

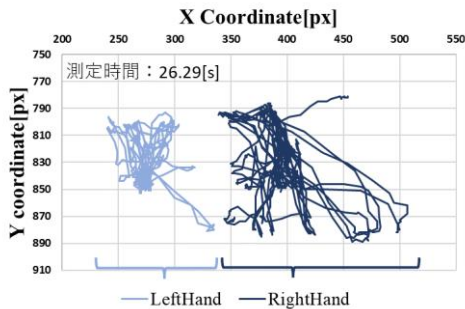


Figure2. First landmark '8' trajectory

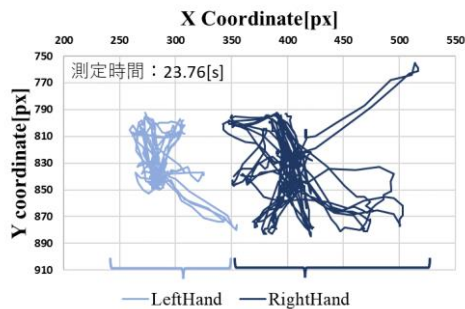


Figure3. Second landmark '8' trajectory

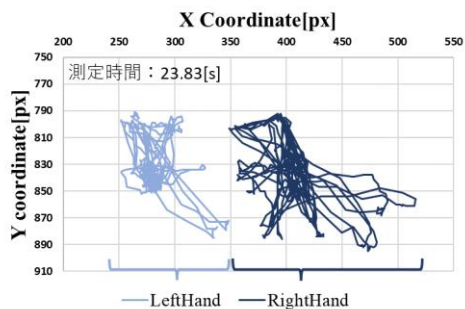


Figure4. Third landmark '8' trajectory

## 5.2 測定者による違い

測定者毎の「無意識的行動特徴」の検討を深めるため異なる測定者で検討を行った. Fig.5 に異なる測定者での landmark '8'の間接点のタイピングでの軌跡を示す. 5.1 と比較し, 測定時間の差の平均は 6.26 [s]であった. また軌跡の形状について 5.1 の測定者は RightHand が右下に延びているのに対して今回の測定者では右上に延びていることを確認した. そして動作範囲については LeftHand の x 座標が約 30 [px]の違いを確認した. 測定時間, 動作範囲についてはキーボードの使用頻度や熟練度によって変化したと考える. また異なる測定者で軌跡の分布, 動作範囲等が異なることから, タイピング動画像から本測定で landmark '8'において軌跡から確率密度の行動特徴を抽出できる可能性を示唆した.

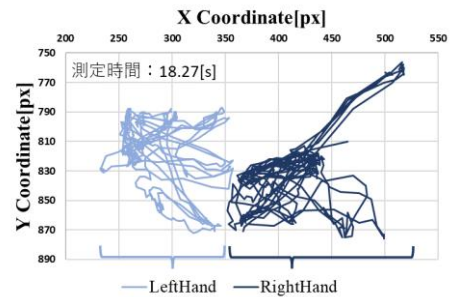


Figure5. Landmark '8' trajectory of different subjects

## 4. まとめ

今回タイピング時の手指動作に個人特徴を有するという仮説に基づき, MediaPipe を用いてタイピング動画像から手指の間接点の抽出を行った.定型文書タイピング時の二次元手指動作軌跡を描画し, 目視による比較を行った結果, 被験者毎で軌跡分布には大きな差異が認められると共に, 同一被験者では比較的類似性の高い結果が再現されることが示された. また, 被験者によって軌跡の分布や動作範囲から個人での差異が抽出できる可能性が示唆された. これらより, タイピング時の手指動作に, 個人で異なる無意識的な行動特徴が含まれていることが示唆された.

## 参考文献

- [1] 宇根正志, 松本勉「生体認証システムにおける脆弱性について:身体的特徴の偽造に関する脆弱性を中心に」『金融研究』(第24巻第2号)
- [2] MediaPipe: <https://github.com/google/mediapipe> (参照日 2023/7/27)
- [3] Z. Cao, T. Simon, S. Wei, and Y. Sheikh. Realtime Multi-Person 2D Pose Estimation using Part Affinity Field. In CVPR,