

# 手指の形と動きの分析による静的・動的指文字認識手法の検討

山口 陽功<sup>†</sup>澤野 弘明<sup>†</sup><sup>†</sup> 愛知工業大学

## 1 はじめに

耳に障がいを持つ人は日常会話の場面で手話や指文字を用いることで会話をする。その中でも指文字は、特定の手の形や動作で仮名文字を一文字ずつ表現する視覚言語の一つである。指文字に馴染みのない人にとっては、常に表現され続ける日常会話を間違えることなくすべて理解することは困難である。池田らはモーションセンサでの関節 25 か所の三次元座標を特徴量として指文字を認識する手法 [2] を提案している。池田らの手法は認識精度が 70.1% であり、有用性は高いがモーションセンサを扱うことによる普及しづらいことが課題である。本研究では日常的に普及しやすい観点から、Web カメラを扱う指文字の自動認識手法に着目する。

渡辺らは Web カメラから得られた画像から骨格座標を抽出して、機械学習を用いて指文字を一文字単位で認識する手法 [1] を提案している。この手法では連続フレームで同一文字が認識されるので単語認識には対応できない。また、動作状態の指文字を認識できないという課題があった。そこで本研究では動きをもつ指文字の認識手法、連続するフレームで認識された指文字が持続する特徴(フレーム持続性)を考慮して指文字単語を認識する手法を提案する。本稿では、提案手法に対して評価実験を行い、今後の課題について述べる。

## 2 提案手法

### 2.1 概要

指文字には、動きを必要としない、形だけで表現する静止状態の指文字と、動きと形を組み合わせる動作状態の指文字が存在する。静止状態の指文字は手の形で一文字を表す。静止状態の指文字の場合、手の形を固定して胸の前で表現される。動作状態の指文字である濁点の表現は、静止状態の形を固定して横に移動させることで表現する。半濁点の場合は上に移動させる。動作状態の指文字には、仮名文字一つを表す場合も存在する。その場合も形を作り特定の動きをすることで一文字を表現する。本研究における動作状態の指文字は、濁点(゛)と半濁点(゜)の表現に限定する。

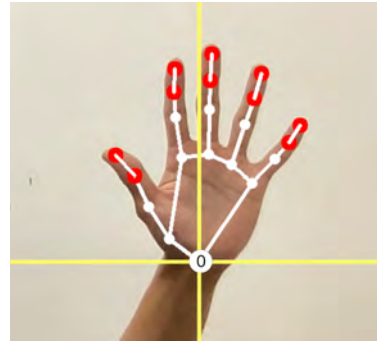


図 1: 検出された指関節の骨格座標

### 2.2 静止状態の指文字の判別

本節では、静止状態の指文字を認識する手法について述べる。Web カメラから得た画像に対して、MediaPipe\*により指関節の骨格座標を三次元座標で取得する。取得された指関節の骨格座標を図 1 に示す。空間の位置に依存させないために、手首の中心を原点に設定して、認識される 21 個の指関節の骨格座標の相対座標を扱う。相対的な指関節の骨格座標から、各指ごとの座標の距離と座標間の角度を算出して、算出された距離と角度をデータとして教師あり機械学習アルゴリズム Support Vector Machine により学習モデルを作成して認識する。

### 2.3 濁点・半濁点の指文字の判別

本節では動作状態の指文字である濁点、半濁点を認識する手法について述べる。前節で述べた静止状態の指文字から濁点、半濁点の付くことが可能な文字に対してのみ処理を行う。静止状態の指文字の認識結果では、毎フレームで指文字の認識処理をするため、複数フレームで変化がなければ同一文字が連続して認識される。そのため連続したフレーム間で手首の中心が閾値以上移動した場合を判別する。濁点の模式図を図 2 に示す。図 2 の模式図のように形を維持して横方向なら濁点、上方向なら半濁点として指文字を認識する。

### 2.4 指文字単語の判別

本節では、フレームで持続する特徴を利用して指文字単語として認識する手法について述べる。2.3 節で述べたように複数フレームで同一文字が連続して認識さ

A Study of Static and Dynamic Fingerspelling Recognition Method Based on Analysis of Hand Shape and Movement

<sup>†</sup>Aichi Institute of Technology



図 2: 濁点の指文字表現方法「び」

表 1: 一文字違いの誤認識 表 2: 過検出文字の誤検出

正解単語	認識結果	正解単語	認識結果
うそ	ゆそ	えき	てえき
くるま	くしま	なつ	になつ
こめ	よめ	はる	はしる
といれ	とゆそ	ふゆ	ふゆみ
へた	へむ		

れるので濁点や半濁点と類似した処理をする。図3に示すように一連の同じ文字の認識結果が閾値以上連続した場合に、それらを一文字の指文字単語を認識する。

### 3 実験と考察

本節では、提案手法を用いた指文字単語の認識結果、濁点や半濁点の認識結果について述べる。指文字単語認識では、1フレームごとに認識処理を行うため、静止状態で表現する指文字41文字に限定して考察する。濁点や半濁点の認識については、濁点や半濁点が存在する文字に限定して考察する。ここで評価実験の入力には図3に示すように、指文字表現者と正解単語が提示

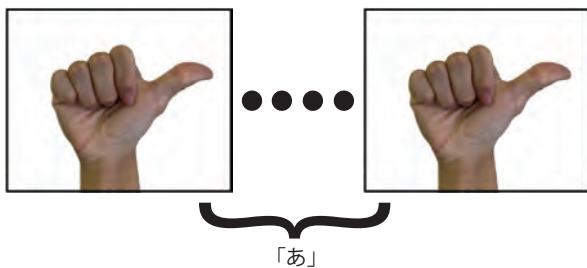


図 3: 一文字として認識する文字の例



図 4: 入力動画 (チャンネル名: こだ YouTube 手話サークル)\*

されており、2-4文字の単語が指文字で表現されている動画を用いる。認識結果数が15フレーム以上続いた場合を一文字として認識する。正解単語と認識された指文字単語が一致している場合を正認識、誤った場合を誤認識とする。実験の結果単語数18に対して正認識は8、誤認識は10であった。表1のように、一文字の認識違いが5単語確認された。また、表2のように正解単語に対して一文字追加された誤認識が4単語あった。そのほかの誤認識は「けいさつ」の単語が「てあつ」と誤って判定された。濁点と半濁点の認識結果については、静止状態の指文字が正認識されている場合に認識された。動作状態時の移動による手のブレで座標が正しく認識されず、他の文字として認識してしまい連続性が保てない場合に誤認識となった。本研究の今後は単語認識結果に対して自然言語処理の予測変換により、認識精度を向上させる方法、濁点、半濁点以外の動作状態の指文字を認識する手法について検討する。

### 4 おわりに

本稿では、手指の形と動きの分析による静的・動的指文字認識手法について述べた。指文字単語の評価実験では、単語数18に対して8単語正認識、10単語誤認識となった。濁点や半濁点の認識では、静止状態の文字がフレーム間で連続して正しく認識される場合に正認識となった。誤認識と過検出された文字の削減や濁点や半濁点以外の動作状態の指文字の認識手法の検討が今後の課題である。指文字単語の誤検出の削減については自然言語処理による精度向上が挙げられる。

### 参考文献

- [1] 渡邊聡, 五十嵐悠紀: “Web カメラを用いた指文字自動認識システム”, WISS 2021, No. 2-B04 (2021)
- [2] 池田聡哉, 高橋正信: “リアルタイム指文字認識システムの研究”, 信学会東京支部学生会研究発表会, No. 154, pp. 1-5 (2018)

\*<https://youtu.be/Yk2O6wsXWYA>