

指骨格検出モデルを用いたリアルタイム日本語指文字認識システムに関する検討

A Study on a Real-Time Japanese Fingerspelling Recognition System Using a Hand Skelton

○落合幹太¹, 関弘翔², 本郷羽琉³, 細野裕行²

*Kanta Ochiai¹, Hiroto Seki², Haru Hongou³, Hiroyuki Hosono²
 節毎に3次元での速度、加速度をそれぞれ求めて9次元の関節特徴量を作成した。これを時間軸・関節軸・関節特徴量のテンソルとして2層の2D-CNNに入力し、時間と関節空間上の局所の特徴を抽出した後、その出力を時刻方向系列としてLSTMに入力し、分類結果を出力する構成にした。

This study investigated a system using MediaPipe Hands to extract 3D hand landmarks and recognize Japanese finger spelling in real time. The extracted features were preprocessed using a CNN and analyzed over time using an LSTM to attempt recognition of the 50 phonetic signs.

1. まえがき

令和4年(2022年)の統計によれば、日本での聴覚障害者は約37.9万人である^[1]。聴覚障害者は会話に手話や指文字を用いるため、健常者との円滑なコミュニケーションには手話通訳が不可欠であるが、令和6年(2024年)時点で通訳者は4,198人に留まっている^[2]。電話での対応は特に困難であり、既存の電話リレーサービスは仲介者を介して会話を補助するが、費用や仲介者の負担、利用のしにくさが課題となっている。

この問題解決に向けて、指骨格検出モデルを応用した手話の自動認識研究が進められている。指骨格検出モデルとして代表的なMediaPipe Handsは、画像から手を検出し、手のひらと21か所の3次元ランドマークをリアルタイムに推定できる仕組みを持つ^[3]。日本語手話認識システムに関する2024年のShinらの研究は、静止画像を入力とし、MediaPice HandsとCNNを組み合わせることでSVMにより判別を実現した。しかし、時系列処理を含むリアルタイム対応は未解決の課題である^[4]。

本研究では、MediaPipe HandsとLSTM(Long short-term memory)を組み合わせたリアルタイム指文字・手話認識システムを検討した。

2. 提案手法

本研究では、自身で撮影した50音それぞれの指文字動画(各音10本、合計460本)を用い、MediaPipe Handsにより各動画フレームから手の3次元ランドマークを抽出した。具体的には、片手21点のランドマークを取得し、手首位置を原点とした相対座標(63次元)および手の表裏(1次元)の計64次元を特徴量とした。これらの時系列データを固定長T=50の線形補間によってリサンプリングし、動画毎にCSV形式で保存した。ここで、訓練データを強化するため、各動画から複数のデータ拡張サンプルを生成した。

さらに相対座標を窓サイズ7で平滑化したうえで関

3. 研究結果

本研究では、まず「あ・い・う・え・お」の5音に限定した縮小実験を行い、次に50音すべてを対象とする分類モデルへ拡張して比較検証を行った。データは、各音8対2の割合で学習とテストにランダム分割した。

5音モデルではテストデータに対する分類精度0.98と高い精度を示した。リアルタイム検証において「あ」「い」「う」では高い認識率が得られたが、「え」「お」では誤認識が多く、汎化性能の低さが確認された。50音モデルへ移行したところ、テストデータに対する分類精度は同様に0.98と高かったものの、リアルタイム検証での認識率は低下し、5音実験で比較的安定していた「あ」「い」「う」においても誤認識が増加した。

4. まとめ

本研究では、MediaPipe Handsによる手の3次元ランドマークとLSTMによる時系列モデルを組み合わせ、日本語指文字(50音)のリアルタイム認識を試みた。

構築したデータセットに対しては高い分類精度が得られたものの、実際のリアルタイム環境下では認識率が十分に確保できず、汎化性能に課題が残った。

今後は過学習抑制やモデル構造改善により、実用的なリアルタイム手話・指文字翻訳システムへの応用を目指す。

参考文献

- [1] 厚生労働省「令和4年生活のしづらさなどに関する調査(全国在宅障害児・者等実態調査)結果」,2024(参照2025-09-29)
https://www.mhlw.go.jp/toukei/list/dl/seikatsu_chousa_c_r04.pdf
- [2] 社会福祉法人聴力障害者情報文化センター,「手話通訳士登録者数」,2025(参照2025-09-29)
https://www.jyoubun-center.or.jp/jbmain/wp-content/uploads/2025/01/si_tourokushasu250115.pdf
- [3] Zhang, et al., arXiv preprint arXiv:2006.10214, 2020.
- [4] J. Shin, et al., CMES, 139(3):2605-2625, doi:10.32604/cmcs.2023.046334, 2024.