

令和 6 年 6 月 7 日現在

機関番号：12103

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K11411

研究課題名（和文）手話の言語的特徴を考慮した深層学習による多次元時系列データを用いた手話認識

研究課題名（英文）Sign Language Recognition Using Multidimensional Time Series Data with Deep Learning Considering Linguistic Characteristics of Sign Language

研究代表者

白石 優旗（Shiraishi, Yuhki）

筑波技術大学・産業技術学部・准教授

研究者番号：00389214

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、ろう・難聴者（聴覚障害者）と聴者のコミュニケーション支援を目的とし、手話認識技術の開発に取り組みました。深層学習を用いて、指文字の識別や連続指文字の認識において高い精度を達成しました。また、車の運転シーンでの片手手話認識の基盤を構築し、ろう・難聴者の安全で自由な移動を支援するための技術的基盤を作成しました。今後は、この研究で得られた知見を活かし、実用的な手話認識システムの開発を目指します。

研究成果の学術的意義や社会的意義

本研究の成果は、手話認識技術の発展に寄与するものです。手話言語の特徴を工学的アプローチにより明らかにすることで、言語学の発展にも貢献し得ると考えます。また、手話認識の精度向上は、ろう・難聴者の社会参加を促進し、生活の質を高めることにつながります。これらの成果は、多様性を尊重し、誰もが平等に社会参加できる環境の実現に貢献するものと考えられます。本研究は、ダイバーシティ社会の実現に向けた重要な一歩となると信じています。

研究成果の概要（英文）：In this research, we aimed to support communication between deaf and hard-of-hearing (DHH) individuals and hearing individuals by developing sign language recognition technology. We achieved high accuracy in identifying finger-spelled letters and recognizing continuous finger-spelling. Moreover, we established a foundation for recognizing one-handed sign language in driving scenarios. These results are expected to promote the social participation of DHH individuals and contribute to realizing a society that respects diversity. In the future, we will utilize the insights gained from this research to develop a practical sign language recognition system.

研究分野：福祉工学、ヒューマンインタフェース、インタラクション、ソフトコンピューティング

キーワード：ろう・難聴者支援 手話認識 指文字認識 機械学習 深層学習 ダイバーシティ社会

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

昨今のダイバーシティ推進により、ろう・難聴者（聴覚障害者）と聴者が共に生活する機会の増加がより一層期待されるため、両者の円滑なコミュニケーションの実現は急務である。これまでの情報保障は、聴者の発生する音声を文字に変換してろう・難聴者に提示する方法が主流であった。しかし、コミュニケーションは双方向的であることが本質であるため、聴者からろう・難聴者への情報伝達を保障するだけでなく、ろう・難聴者の発話（手話）を聴者に文字または音声で提示することも必要となる。

現在、聴者が音声で講演し、ろう・難聴者が聴講する場合には、音声認識を活用した通訳が広く行われるようになってきている。一方で、ろう・難聴者が手話で講演する場合は、専門の手話通訳者による通訳が一般的であり、コストや手話通訳者の不足が問題となっている。手話認識は、音声認識と比較して認識精度が低く、少数の単語や簡単なフレーズの認識に留まっており、実用レベルには至っていないのが現状である。

このような背景を踏まえ、本研究では、ろう・難聴者と聴者の双方向コミュニケーションを実現するための手話認識システムの開発に取り組む。特に、講演やグループディスカッションなど、ろう・難聴者が多数の聴者に対して手話で発言する場面を想定し、実用的な手話認識の実現を目指す。

2. 研究の目的

本研究の目的は、ろう・難聴者と聴者の円滑なコミュニケーションを実現するために、高精度な手話認識システムを開発することである。具体的には、以下の3点を主な目的とする。

- (1) 深層学習の手法を用いた高精度な手話認識システムの開発：
研究代表者は、長年にわたって深層学習の研究に取り組んできた実績を有している。この知見を活かし、高い識別率を実現する手話認識システムの開発を目指す。深層学習は、複雑なパターンを認識するのに適しており、手話認識への応用が期待される。
- (2) 音声認識手法を拡張した手話認識の実現：
今回提案する手話認識では、データグローブを用いて取得した手の位置、向き、手指の形状、動きなどの情報を多次元時系列データとして扱う。一方、音声認識では、音声情報を一次元時系列データとして扱う。本研究では、音声認識の手法を拡張し、多次元時系列データへの適用を図ることで、手話認識を実現する。これにより、手話認識の精度向上が期待される。
- (3) 手話言語の特徴解明への貢献：
手話認識システムの開発過程において、音声言語とは異なる手話言語特有の特徴を工学的アプローチにより明らかにする。これにより、「言語とは何か」という根源的な問いの解明に寄与することを目指す。手話言語の特徴を理解することは、手話認識の精度向上につながるだけでなく、言語学の発展にも資するものと期待される。

これらの目的を達成することで、ろう・難聴者と聴者の円滑なコミュニケーションを実現し、ダイバーシティ社会の実現に貢献することを目指す。

3. 研究の方法

本研究では、手話認識を実現するために、データグローブを用いた手法を採用する。手話認識には、身体に対する手の相対位置、向き、手指の形状、動きなどの情報が不可欠である。データグローブを用いることで、これらの情報を安定して取得することができる。データグローブは、講演者が手に装着するだけで使用でき、環境光やオクルージョン（手と手の重なり）の影響を受けにくいという利点がある。また、使用場所を選ばずに自然な手話を行うことができるため、実用性が高い。

研究の進め方としては、段階的に手話認識の精度を高めていく計画とした。まず、動的指文字（手指を動かしながら提示する指文字）の識別実験を行う。動的指文字は、手話の基本要素であり、その識別精度を高めることは手話認識全体の精度向上につながる。次に、連続指文字（指文字の連続的提示）の認識手法について検討を行う。連続指文字の認識は、動的指文字の識別よりも難易度が高いため、様々な手法を比較検討し、最適な手法を見出すことを目指す。最終的には、こ

これらの知見を活かして手話認識システムを開発し、識別実験により精度の評価を行う。

認識手法には、研究代表者がこれまで長年にわたって取り組んできたニューラルネットワークと深層学習の手法を用いる。これらの手法は、複雑なパターン認識に適しており、高い識別率の達成が期待される。また、近年急速に発展している音声認識分野の研究成果を応用することで、手話認識の精度向上を図る。具体的には、音声認識で用いられている時系列データ処理の手法を拡張し、多次元時系列データである手話データに適用する。この独自のアプローチにより、手話認識の新たな可能性を切り開くことを目指す。

さらに、手話言語の特徴を深く理解することも重要な課題である。手話言語は、音声言語とは異なる独自の文法体系を持っており、その特徴を考慮した認識手法の確立が求められる。本研究では、工学的見地から手話言語の分析を行い、手話認識に活かすべき特徴を明らかにする。これにより、手話認識の精度向上だけでなく、手話言語の理解深化にも貢献することを目指す。

4. 研究成果

本研究では、動的指文字の識別から連続指文字の認識、そして手話認識へと段階的に研究を進展させてきた。当初は、講演やグループディスカッションなど、ろう・難聴者が多数の聴者に対して手話で発言する場面を想定していたが、研究の過程でより実用的な場面として車の運転シーンに着目し、片手手話認識に取り組むこととした。その過程で、以下のような成果を得ることができた。

動的指文字の識別においては、データグローブを用いて手の位置、向き、手指の形状、動きなどの情報を安定して取得する方法を確立した。20名のろう・難聴の実験協力者から収集した日本の指文字76文字（濁音、半濁音、長音、拗音を含む）のデータを用いて、深層学習による識別モデルを構築した結果、20分割交差検証法で評価した平均識別率は70.0%に達した。この成果は、動的指文字の識別精度としては高い水準にあり、手話認識の基盤技術として重要な意義を持つ。

連続指文字の認識においては、40名のろう・難聴の実験協力者から採取した64単語(2-5文字)のデータを用いて、CNNとLSTMを組み合わせたモデルを構築した。その結果、5-fold交差検証法で評価した認識率は92.1%という高い水準に達した。この成果は、9,305語分の連続指文字データに対して適用され、国際ジャーナルに採択されるなど、学術的にも評価されている。

手話認識に向けては、Bluetoothを用いたケーブルレスのセンサグローブを新たに開発し、自然な手話データの取得を可能にした。また、手話認識用のデータ採取実験における新たな実験プロトコルを作成し、効率的かつ体系的なデータ収集を実現した。

片手手話認識の基盤構築においては、以下のような成果が得られた。

(1) 片手手話の選定とデータ採取実験

車の運転シーンを想定して使用される可能性の高い片手手話単語58語と短文40文を選定した。選定作業は、運転免許を保有する3名のろう者と1名の聴者の議論に基づいて行われ、ニーズ調査に基づき、緊急度の高い情報、方向指示、雑談を含むものとした。

データ採取実験では、車内環境を模擬し、各短文について運転席では左手で、助手席では右手で片手手話を4回ずつ提示してもらった。カメラとグローブのそれぞれを用いて、10名と12名の実験協力者から片手手話データを収集した。本実験は、筑波技術大学倫理審査委員会の承認を得ている。

(2) 片手手話の評価とアノテーションルール

カメラによるデータ採取実験に参加した10名（手話経験平均11.6年、標準偏差5.76年）を対象に、選定した片手手話単語と短文についてアンケート調査を行い、その妥当性を検証した。

二項検定の結果、車内手話認識の必要性和片手手話の使用について、両質問とも有意水準5%で有意差が確認された($p=0.017$)。このことから、ろう・難聴者にとって車内手話認識の必要性和片手手話の使用が強く支持されていることが明らかになった。

また、57語中50語、40文中28文において、有意水準5%で片手手話表現の自然さが支持された。残りの単語と短文についても、受容可能性が確認された。

アノテーションルールでは、「単語」と「遷移」に大分類し、動的な手話表現と静的な手話表現における「単語」の開始点と終了点を定義した。このルールは、車の運転シーンにおける片手手話認識のためのデータセット構築において重要な役割を果たすものとする。

これらにより、ろう・難聴者のニーズに基づいた片手手話単語と短文の選定、データ収集実験、評価、アノテーションルールの策定を通して、車の運転シーンにおける片手手話認識の基盤を構築することができた。

本研究の一連の成果は、手話認識技術の発展に大きく寄与するものである。高精度な手話認識システムが実現されれば、ろう・難聴者と聴者のコミュニケーションが円滑になるだけでなく、ろう・難聴者の社会参加の促進や生活の質の向上にもつながることが期待される。特に、車の運転シーンにおける片手手話認識は、ろう・難聴者のモビリティ支援において重要な役割を果たすものとする。

また、本研究で得られた知見は、手話言語の理解を深める上でも重要な意義を持つ。手話言語の特徴を工学的アプローチにより明らかにすることで、言語学の発展にも資することができる。

今後は、データ採取実験で得られたデータとアノテーションラベルに基づき、手話認識システムを構築する。その際、手話言語の特徴を考慮した認識モデルを構築することで、高精度の認識性能の実現を目指す。本研究の成果は、ダイバーシティ社会の実現に寄与するものと考えられる。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件／うち国際共著 0件／うちオープンアクセス 2件）

1. 著者名 Yuhki Shiraishi, Akihisa Shitara, Fumio Yoneyama, Nobuko Kato	4. 巻 14
2. 論文標題 Sensor Glove Approach for Continuous Recognition of Japanese Fingerspelling in Daily Life	5. 発行年 2022年
3. 雑誌名 International Journal on Advances in Life Sciences	6. 最初と最後の頁 53-70
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Takashi Kato, Akihisa Shitara, Nobuko Kato, Yuhki Shiraishi	4. 巻 -
2. 論文標題 Sign Language Conversational User Interfaces Using Luminous Notification and Eye Gaze for the Deaf and Hard of Hearing	5. 発行年 2021年
3. 雑誌名 Proceedings of The Fourteenth International Conference on Advances in Computer-Human Interactions (ACHI 2021)	6. 最初と最後の頁 33-36
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計6件（うち招待講演 1件／うち国際学会 5件）

1. 発表者名 Akihisa Shitara, Taiga Kasama, Fumio Yoneyama, Yuhki Shiraishi
2. 発表標題 One-Handed Signs: Standardization for Vehicle Interfaces and Groundwork for Automated Sign Language Recognition
3. 学会等名 Proceedings of The Seventeenth International Conference on Advances in Computer-Human Interactions (ACHI 2024)（国際学会）
4. 発表年 2024年

1. 発表者名 Nobuko Kato, Yuito Nameta, Akihisa Shitara, Yuhki Shiraishi
2. 発表標題 Sign Language Writing System: Focus on the Representation of Sign Language-Specific Features
3. 学会等名 Proceedings of The Seventeenth International Conference on Advances in Computer-Human Interactions (ACHI 2024)（国際学会）
4. 発表年 2024年

1. 発表者名 佐々木 裕多, 張 建偉, 白石 優旗
2. 発表標題 Commonsense-aware AttentionとDiscrepancy Resolution Lossを用いたユーモア検出手法の提案
3. 学会等名 第15回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2023年

1. 発表者名 Yuhki Shiraishi
2. 発表標題 User-led Accessible Technology Research for the Deaf and Hard of Hearing in Human-computer Interaction
3. 学会等名 The Fourteenth International Conference on Advances in Computer-Human Interactions (ACHI 2021) (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Tomohiko Tsuchiya, Akihisa Shitara, Fumio Yoneyama, Nobuko Kato, Yuhki Shiraishi
2. 発表標題 Sensor Glove Approach for Japanese Fingerspelling Recognition System Using Convolutional Neural Networks
3. 学会等名 Proceedings of The Thirteenth International Conference on Advances in Computer-Human Interactions (ACHI 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Yuhki Shiraishi, Takuma Takeda, Akihisa Shitara
2. 発表標題 Alarm Sound Classification System in Smartphones for the Deaf and Hard-of-Hearing Using Deep Neural Networks
3. 学会等名 Proceedings of The Thirteenth International Conference on Advances in Computer-Human Interactions (ACHI 2020) (国際学会)
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Shiraishi Lab - Sign Language Recognition
<https://www.shiraishi-lab.org/home/project/sign-language-recognition>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------