

手話認識による逆引き検索が可能なクラウド型手話辞典の開発

ー 手話の逆引き検索の自立を実現 ー

1. 背景

手話は独自の文法をもつ自然言語であり主に聴覚障害者のコミュニティで用いられている。つまり、よく誤解されがちであるが手話はジェスチャーではなく、日本手話と日本語は別言語である。また、地域によっても手話は異なる。これは日本手話とアメリカ手話の間の言語的な違いだけでなく、方言による語彙の違い、さらにはコミュニティ間や個人間での手話形の違いなど様々である。

このような手話の多様性が理由で「手話辞典で勉強した手話とろう者が使用する手話が異なるため、ろう者との会話中に相手の手話を読み取ることが出来ない」という状況がよく発生する。さらに、手話の世界では手話動作から日本語の意味を検索できるような逆引き手話辞典が存在しないため「手話学習者が手話動作から日本語の意味を自身で調べられない」という深刻な問題に直面している。本プロジェクトのクリエイターの高橋は、学部時代に所属していた手話サークルでの経験から、この問題が多くの手話学習者の挫折を生み出していると考え、逆引き検索可能な手話辞典の必要性を痛感した。

2. 目的

手話認識による逆引き検索が可能なクラウド型手話辞典を開発することで「手話学習者が手話動作から日本語の意味を自身で調べられない」という問題を解決する。本システムの開発と普及を通じて、手話動作から日本語の意味を知るためには有識者に直接尋ねるしか方法がないという現状を改善し、手話の逆引き検索の自立を目指す。

3. 開発の内容

図 1 は開発した逆引き検索システムの Web アプリケーションのスクリーンショットである。本システムには約 3700 語彙の手話単語を含む大規模な日本手話データセットで学習した手話認識モデルが搭載されており、ユーザは撮影した動画から手話の意味を逆引き検索することが可能である。また、本システムには直感的な「動画による逆引き検索機能」、網羅的な「手型による絞り込み機能」、連鎖的な「類似手話検索機能」の 3 種類の手話検索機能が実装されている。本プロジェクトで開発した手話認識による大規模な逆引き検索システムは「豊富な語彙が登録された手話データベース」と「直感的・網羅的・連鎖的な逆引き検索機能」の両立により手話学習者に新たな学習体験を提供する。



図 1：開発した逆引き検索システムの Web アプリケーションのスクリーンショット

1つ目の「動画による逆引き検索機能」は、撮影した動画から手話を検索する機能である。まず、図 1 の「①手話動画撮影」でユーザは意味を調べたい手話の動画を撮影して検索ボタンを押すと手話認識による逆引き検索が開始される。その後、手話認識により予測されたスコアが高い順にデータベースに登録されている手話動画とその手話の意味が「②予測結果表示」に複数表示され、それらの手話動画が全てループ再生される。これにより、ユーザは表示された複数の手話動画の候補と自分の手話動作を比較するだけで手話の意味を調べることができる。つまり、「動画による逆引き検索機能」は直感的な検索方法であると言える。

2つ目の「手型による絞り込み機能」は、左右の手型を入力することで逆引き検索の結果を絞り込むことができる機能である。ユーザが図 1 の「③手型で絞り込み」で左右の手型の情報を入力して検索ボタンを押すと「②予測結果表示」の候補が更新され、既に表示されている候補の順序を維持したまま入力した左右の手型の情報と同じ手話動画のみが表示される。

図 2 は、「学生」の手話を「動画による逆引き検索機能」により検索した後に「手型による絞り込み機能」による絞り込み検索を行った例である。手話認識モデルの出力では「学生」の手話は 2312 番目の候補として予測されており、動画による逆引き検索では 257 ページ目に表示されることになるため、ユーザは「学生」の手話を発見することが困難である。しかし、手型による絞り込みを行うことで 1 ページ目に「学生」の手話を表示することが出来ている。このように、ユーザは「動画による逆引き

検索機能」による候補の中に調べたい手話動作が見つからなかった場合でも、「手型による絞り込み機能」を用いることで最終的に調べたい手話動作の意味を発見することができる。つまり、「手型による絞り込み機能」は網羅的な手話検索機能であると言える。



図 2: 「学生」の手話の検索結果の比較

3つ目の「類似手話検索機能」は、予測結果に表示されている手話と動作が似ている手話を検索する機能であり、手話学習者の知的好奇心を刺激して学習効果を高める役割がある。ユーザが図 1 の「④類似手話検索」のボタンを押すことで、撮影した動画の代わりに選択した手話動画を入力として「動画による逆引き検索機能」の結果が表示される。そして、「類似手話検索機能」は図 3 のように繰り返し実行可能であり、ユーザは興味が尽きるまで似ている手話を調べ続けることが出来る。動作が似ている手話は語彙の意味が関連していることが多いため、動作が似ている手話を連鎖的に検索することが可能な本機能は、非常に高い学習効果が期待される。

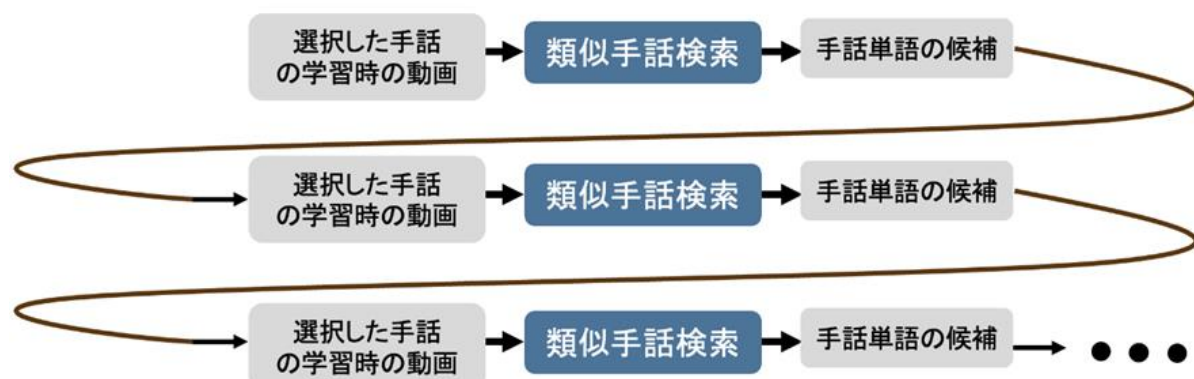


図 3: 「類似手話検索機能」を用いた連鎖的な検索のイメージ

本システムの有効性を検証するために、2月上旬に男女2名の協力者にヒアリングを行い、手話の逆引きに関する悩みや本システムの感想を伺った。本システムの使用前のヒアリングでは、2名とも手話の逆引きを行えない現状について問題意識を持っていた。また、男性の協力者はヒアリングの実施日の1週間前に「声なし交流会」と言う、聴者と聴覚障害者が集まり、声を用いずに手話や筆談などの手段によって互いに交流を行う交流会に参加していた。その際に話し相手の手話が分からないことがあったが、複数人で交流を行っていたため相手に直接手話の意味を聞くことができず、結局手話の意味が分からないまま話が進んでいってしまった、とその男性協力者は悩みを打ち明けた。この体験談からも有識者に直接手話の意味を尋ねることは心理的ハードルが高いことがわかる。そして、実際にその分からなかった手話である「趣味」という手話動作を本システムに入力したところ、1番目の候補として出力され、2名の協力者は驚愕していた。また、2名の協力者が連鎖的な類似手話検索機能を繰り返し用いて積極的に手話を学習していく様子も観察することができた。システムの使用后、2名の協力者から「(「趣味」の手話の意味を)声なし交流会の時に知りたかった」という声や「本システムを早く公開して欲しい」という声を頂いた。また、「PCだと外出時に使用することや調べたい時にすぐに使用することが難しいので、本システムをスマートフォンでも使用できるようにして欲しい」というシステムの改善案についてのフィードバックも頂いた。

4. 従来の技術（または機能）との相違

本プロジェクトで開発した手話認識による大規模な逆引き検索システムの最大の価値は「豊富な語彙が登録された手話データベース」と「直感的・網羅的・連鎖的な逆引き検索機能」の両立である。既存サービスはどちらか一方しか存在しなかった。例えば、クラウド型手話辞典は豊富な語彙が登録されているものの逆引き検索機能が不十分であり、手話認識を活用したサービスは登録されている語彙が少ないためユーザが手話辞典として使用することができなかった。本プロジェクトでは、約3700種類の手話単語を含む日本手話データセットを使用し、わずかなデータセットから頑健な手話認識モデルを学習する手法を確立することでこれらの両立を実現した。さらに、本システムは単に手話の逆引きが可能だけでなく、直感的・網羅的・連鎖的な逆引き検索機能の実現により新たな手話学習体験をユーザに提供することに成功している。

5. 期待される効果

本プロジェクトで開発した手話認識による大規模な逆引き検索システムは手話学習者の逆引き検索の自立を実現するとともに新たな学習体験を提供する。手話学習者にとっては自分が分からない手話の意味をろう者や有識者に何度も尋ねるのは心理的なハードルが高く、ろう者にとっても何度も話を遮ってまで手話の意味を尋ねられるのは心理的な負担が大きい。有識者の協力に依存しない本システムの逆引き検索機能は、心理的な要因によりこれまで無意識下で失われていたであろう手話の学習機会を手話学習者に提供することが可能であり、これは手話の普及や語彙の定着に貢献すると考えられる。また、本システムは任意の手話単語データセットを追加することで様々な国や地域の手話に対応することが可能であり、手話の地域性・国際性・専門性に関する様々な課題の解決に応用可能である。

学術的な観点から見ても、開発した「わずかなデータセットから学習可能な手話認識モデル」は手話認識分野の発展に大きく貢献した。本プロジェクトでは手話単語の逆引き検索に限定した上で語彙の大規模化に焦点を当てたが、その応用性として大規模な手話フレーズ認識や大規模な自動手話翻訳が挙げられる。

6. 普及（または活用）の見通し

本プロジェクトで使用したデータセットである「工学院大学 多用途型日本手話言語データベース (KoSign)」は学術研究目的での使用に限定されていた為、プロジェクト期間中は本システムの公開を行わなかった。本プロジェクトを通じて、各語彙についてわずかな手話動画のサンプルがあれば大規模な逆引き検索システムを構築可能であることが分かったので、今後は大規模な手話データセットを持つ機関と連携することで手話認識による逆引き検索システムの社会実装を目指していく予定である。

7. クリエータ名（所属）

高橋 亮太（慶應義塾大学大学院）

（謝辞）

本プロジェクトでは、国立情報学研究所の IDR データセット提供サービスにより工学院大学から提供を受けた「工学院大学 多用途型日本手話言語データベース (KoSign)」を利用した。