

### 1. 担当 PM

稲見 昌彦（東京大学 先端科学技術研究センター 教授）

### 2. クリエータ氏名

高橋 亮太（慶應義塾大学 大学院）

### 3. 委託金支払額

2,304,700 円

### 4. テーマ名

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

### 5. 関連 Web サイト

なし

### 6. テーマ概要

高橋氏は手話サークルでの活動を通じて、手話の初学者がろう者とのコミュニケーションに苦勞している現状を認識した。それは初学者が直面する大きな障壁の一つは、使われている手話の意味を自分で調べることの難しさである。既存の手話辞典やクラウドサービスには、手話から日本語の意味を調べるための逆引き機能が十分に備わっていない。この逆引き機能を実現する上での課題は主に 2 つある。まず、手話動画から単語の意味を予測するためには、大規模なデータセットが必要であり、日本手話の公開データセットが不足していることが挙げられる。次に、手話には地域差や個人差が大きく、標準形以外の動作が多用されるが、これを全て網羅することは現実的ではない。

そこで本プロジェクトでは、姿勢推定技術と手話言語学の知見を取り入れた逆引き検索が可能なクラウド型手話辞典の開発を目指す。この辞典は、様々な手話動作から意味を逆引きできるよう、集合知を活用したアルゴリズムと統合される。手話辞典の信頼性を保ち、ユーザが正しい内容を学習できるように、出力の精度には特に注意を払う。これにより、手話学習者が直面する課題に対処し、ろう者との円滑なコミュニケーションの促進を目指す。

## 7. 採択理由

英和辞典と和英辞典が存在するように、手話翻訳にも順引きと逆引きがある。提案者は大学の手話サークルに所属し、手話理解の困難さが初学者にとっての大きなボトルネックとなっていることに気付いたとのことである。手話は残念ながら日本手話、アメリカ手話、イギリス手話など地域ごとに分かれている。さらには同じ日本手話でも差があると言われており、プロジェクトの実現のためのハードルは高い。しかしながら、提案者は手話認識・翻訳するためのシステムをライフワークとしてでも取り組みたいという強い熱意を有しており、その熱意を応援する気持ちも込めて採択した。

## 8. 開発目標

手話は聴覚障害者のコミュニティで主に使用される自然言語であり、独自の文法を有する。手話が単なるジェスチャーと見なされがちなのは誤解であり、実際には日本手話と日本語は異なる言語である。手話は地域によっても異なり、日本手話とアメリカ手話の間には言語的な違いが存在し、さらに地域差、コミュニティ間、個人間での手話の形の違いも多様である。

この手話の多様性が原因で、「手話辞典で学んだ手話とろう者が使用する手話が異なるため、会話中に相手の手話を読み取れない」という状況が頻繁に発生する。手話動作から日本語の意味を検索できる逆引き手話辞典の不在は、「手話学習者が手話動作の意味を自ら調べられない」という深刻な問題に直面している状態を生み出している。本クリエイター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引き起こしていると認識し、逆引き検索が可能な手話辞典の開発が必要であると感じた。

そこで、手話認識を活用した逆引き検索が可能なクラウド型手話辞典の開発を提案した。このシステムの開発と普及により、手話学習者が手話動作から日本語の意味を知るための自立した手段を確立し、手話コミュニケーションの障壁を低減することを目指した。

## 9. 進捗概要

本プロジェクトで開発した手話の逆引き検索システムは、約 3,700 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いる、ユーザが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な Web アプリケーションである。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）。



図 1：開発した手話の逆引き検索システム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、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いえる。

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。ユーザは左右の手型の情報を入力して検索ボタンを押すと、予測結果の候補が更新され、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。例えば、「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場合、手話認識モデルの出力では「学生」の手話は 2,312 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できる。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いえる。

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。ユーザは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できる。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2)。

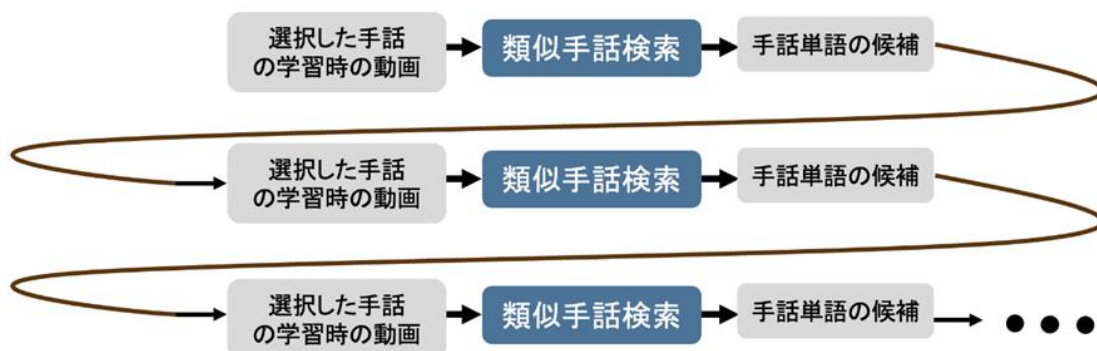


図2: 「類似手話検索機能」を用いた連鎖的な検索のイメージ

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聞いた。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、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「声なし交流会」という交流会に参加したようで、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、という悩みを打ち明けた。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番目の候補としてそれが出力され、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。システムの使用後、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」「本システムを早く公開して欲しい」という声を得られた。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を得られた。

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 3,700 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。

## 10. プロジェクト評価

高橋氏は、語彙の手話単語を含む大規模な日本手話データセットを用いて学習した手話認識モデルを開発し、動画から手話の意味を逆引き検索できる Web アプリケーションを実装した。これは問題解決に直結する適切なアプローチであった。さらに、手型による絞り込み機能や類似手話検索機能など、ユーザの利便性を考慮した機能が追加されており、システムの使いやすさと学習効果の向上に寄与するものと評価できる。

協力者からのヒアリングにおいて、システムへの高い評価が得られたことは、開発されたシステムが手話学習者のニーズに合致していることを示している。また、協力者からのフィードバックを踏まえ、スマートフォンへの対応を進めることで、システムのアクセシビリティと利便性をさらに高めることが期待される。

本システムを通じて、これまで無意識下で失われていた手話の学習機会を提供し、手話の普及と定着に貢献できる可能性が示された。さらに、手話の地域性・国際性・専門性など、様々な課題解決にも応用可能であることから、今後の発展的な研究・開発が期待される。

以上のように、本プロジェクトは、手話学習者が直面する問題を的確に捉え、適切なアプローチでシステムを開発し、ユーザのニーズに合致するソリューションを提供することに成功した。開発されたシステムは、手話学習者の学習機会を拡大し、手話の普及と定着に寄与する可能性を持っている。

## 11. 今後の課題

本プロジェクトの指導を通して PM 自身が得た知見として、国による手話の違いがあるだけでなく、日本国内にも多くの「方言」があるという点がある。そのような意味では多様な手話のデータベースの構築が大きな課題として今後も残るだろう。

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ため、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すと聞いている。PM としてその成果を見守っていきたい。