

指文字の音声出力システムの開発背景

高原佑哉† 長野有斗武† 足立ゆきの† 落田綾†
高槻高等学校†

1. はじめに

当初、モーショキャプチャや Python での OpenCV を用いた画像処理などについての研究を行いたいと考え、画像処理を用いた骨格検出について研究していた。今日の骨格検出研究の多くはスポーツや姿勢に関するものであり、社会問題の解決につながるものは少ない。そこで、何か実用的なものを開発し、社会問題を解決できないだろうかと考え、先行研究等を調査したところ、「Kinect for Windows」による手話認識という研究に興味を持った。その研究によると、手話認識システムを作るためには学習データの作成に大きな課題があるとされており、また特殊なカメラが必要とされていた。そこで、特別な機器を用いず、一般的なスペックの PC で動かせる手軽なシステムを作れば幅広い場合で使用できると考えた。画像認識による骨格検出の研究を進めていたため、それを応用して手話を認識したいと考えた。

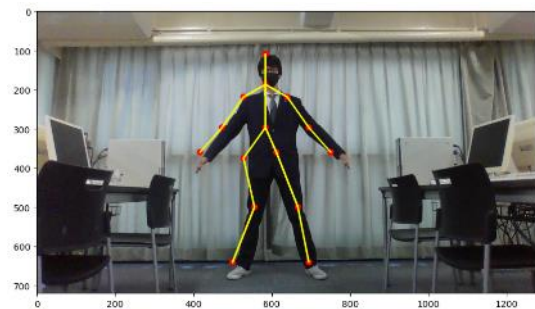


図1 骨格認識の結果

2 システムの構想

ウルトラリープモーションは手指の 3D データを取得可能にするセンサーと赤外線カメラで撮影・解析することで手の動きを検知するカメラである。手の動きを検出するウルトラリープモーションを動作確認した際、比較的高性能なデスクトップ PC においては問題なく動作したが、一般的なスペックのノート PC では動作が遅く、しばしばタイムラグが起こった。より実用性の高いシステムにするには動作を軽くする必要があると考えた。そこで、五十音手話(指文字)に着目した。一般的な手話は動きが多いのに対し、五十音手話は動きがほぼなく、骨格を認識することで対応可能で、学習データが少なくよいため、システムを作成しやすいのではないかと考えた。

2.1 五十音手話とは

五十音手話とは、五十音を手や指の形で表現する方法であり、一文字ごとに手の形が割り振られている。また、一

般に広く使用される手話とは違い、動作があまりないため画像処理がしやすい。さらに、一般の手話と異なり、地域差がないこと、アルファベットや数字を表せること、また固有名詞を表現する際に使われているので手話利用者なら使えることを踏まえて、本システムでは五十音手話を採用した。

2.2 システム概要

聴覚障がい者から健常者へ言葉を伝えるために、五十音手話をパソコンのカメラで読み取り、文字起こしをして、音声として出力する。

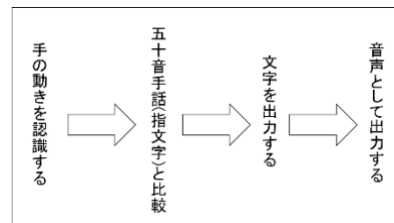


図2 システムの流れ

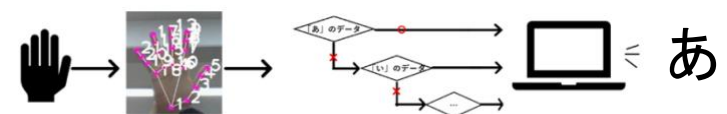


図3 システムの概要図

3. 今後の展望

今回考案したシステムは開発中であり、現在一般的なスペックのノート PC でタイムラグなしで動作している。もし、このシステムが完成すれば、聴覚障がい者と健常者がより円滑に意思疎通を行えるようになると考えられる。また、文字の出力機能により議事録の作成が容易になり、通訳の必要がなくなるため、DX 化やペーパーレスに貢献できると考える。

参考文献

- Python OpenCV で骨格検出 AI OpenPose <https://miyashinblog.com/openpose/> (2023.2.7)
- SureTalk <https://www.suretalk.mb.softbank.jp/> (2023.2.7)
- Voicevox <https://voicevox.hiroshima.jp/> (2023.2.7)
- Kinect for windows による手話認識 <https://www.mizuhort.co.jp/seminar/lecture/pdf/lecture140322.pdf> (2023.2.14)
- ハンドトラッキングでポーズ認識 <https://is-ai.jp/?p=314> (2023.9.20)
- 身体障害者・児実態調査結果の概要 <https://www.dinf.ne.jp/doc/japanese/resource/physical/%E5%B9%B4%E8%BA%AB%E4%BD%93%E9%9A%9C%E5%AE%B3%E8%80%85%E3%83%BB%E5%85%90%E5%AE%9F%E6%85%8B%E8%AA%BF%E6%9F%BB%E7%B5%90%E6%9E%9C.pdf> (2024.1.14)
- WIRED Audi INNOVATION AWARD2017 https://wired.jp/waia/2017/23_junt (2023.2.7)