

ウェブベース遠隔手話通訳のための手話映像生成に関する基礎的検討

若月大輔

筑波技術大学 産業技術学部 産業情報学科

キーワード：遠隔手話通訳, 聴覚障害者, 遠隔情報保障, ウェブアプリケーション

1. はじめに

平成 28 年 4 月の障害者差別解消法の施行を控え、聴覚障害者の社会進出支援の新たな取組や効率化が求められている。本研究では、手話通訳者の確保が困難な地域や、専門的な内容に対して手話通訳を実施するために、ICT 技術を用いた遠隔手話通訳について研究を重ねてきた。我々の遠隔手話通訳システムでは、提示資料やキーワードを手話通訳映像に合成するシステム、現地の話者の映像や利用者の映像を合成して手話通訳者に提示するシステムを導入して、利用者と通訳者とがともに利用しやすい環境を実現している。しかし、高価な機材の導入や、高速で安定なネットワークの利用ならびに通訳者以外に技術支援を行うスタッフが必要になるなど、運用面のコストについて課題がある。

そこで、特別機材の導入やネットワークの制限を受けにくいウェブベースで遠隔手話通訳が可能なシステム構築を目指す。本研究では、特にウェブブラウザで遠隔手話通訳を実現するにあたって重要な技術的要件の 1 つである手話映像の送受信方法について試作システムを構築し、基礎的な検討を行った。通常の遠隔手話通訳では通訳者の実写映像を利用者側に送信するが、本研究では実写と CG を組み合わせて手話表現を行う方法（以後、ハイブリッド CG 手話）を提案し、実装した。

2. ハイブリッド CG 手話

手話はその名前から手指だけが注目されがちだが、目や口、眉の動きを含めた顔も非常に重要な要素である。一方で、手話者の体型は、手指や顔にくらべて重要度は低い。先の研究で提案した CG で匿名手話を生成する方法 [1] を応用し、本研究では、手話者の身体を CG で表し、手指と顔を実写で表現するハイブリッド CG 手話を提案する。

ハイブリッド CG 手話に必要なデータを得るために Kinect v2 センサを利用した。図 1 のように人体を認識し、骨格データを取得し、右手、左手および顔の実写画像を抽出す

る。これらのデータを用いて、モーションデータより身体と腕を CG で生成し、左右の手と顔の画像を合成することによってハイブリッド CG 手話映像を生成する。

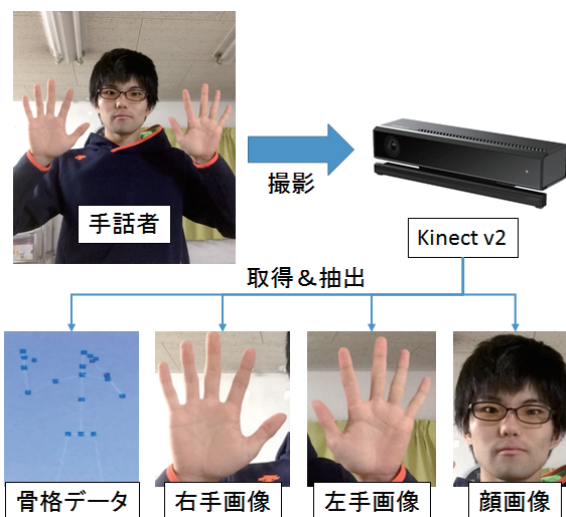


図1 Kinect v2によるデータの取得・抽出

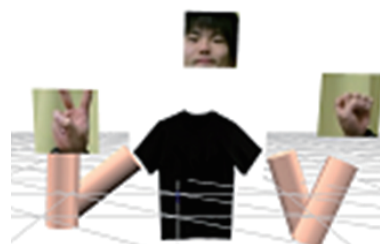


図2 ハイブリッドCG手話の例

3. ハイブリッド CG 手話の送受信

ウェブベースでハイブリッド CG 手話を送受信するために、HTML5 と WebSocket を用いて実装を行った。試作したシステムの概要を図 3 に示す。手話者側で Kinect v2 で得られたデータを WebSocket で送受信サーバに送り、同サーバに接続している利用者のブラウザに対して送信を行

う。ブラウザで受信したデータを HTML5 の canvas 機能を用いてハイブリッド CG 手話を表示する。

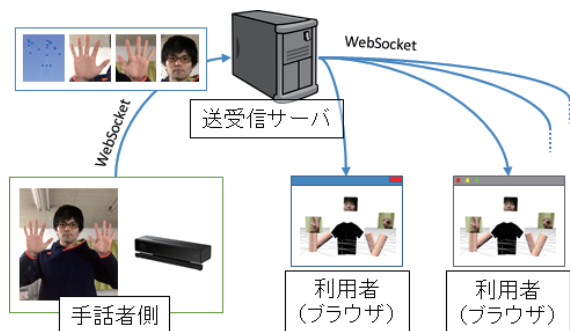


図3 試作システム概要

今回の実装では、骨格データを浮動小数点型配列、画像を Jpeg で圧縮し BASE64 でエンコードしたものを送信した。

4. 性能評価

試作システムの性能を評価するために、1 台の計算機 (Let's Note CF-AX2) に送受信サーバを動作させ、手話者側の送信用プログラム、利用者のブラウザを接続してハイブリッド CG 手話の送受信実験を行った。送信用プログラムは 50fps でデータを送信し、それを受信するブラウザには Chrome を利用し、通信速度シミュレーションで表示のフレームレートを計測した。表 1 にその結果を示す。

表1 ハイブリッドCG手話のフレームレート

通信速度シミュ設定	フレームレート
未設定	37fps
4G	33fps
3G	24fps

未設定では 37fps であり、送信時の約 75% のフレームレートになってしまった。これは、ブラウザ上での処理速度の課題であると考えられる。設定を 4G, 3G に変更するとフレームレートが下がる。ネットワーク通信速度によって、データの送受信に影響が出るためである。しかし、3G でも 24fps であり、動画像を見るうえで最低限のフレームレートが確保されることがわかった。

5. 本システムの利点と今後の課題

ハイブリッド CG 手話表現では手話者の両手、顔のみの実写画像を用いるため、手話者の身体や腕は直接表現されない。一般的な手話通訳者は手話の見やすさに配慮し

てダークトーンの色を着用することが多いが、本システムを介して手話通訳を行う場合はこの配慮が不要になり、利用者も見やすい CG の身体を自由に選択することが可能になる。

試作システムでは通信に WebSocket, 表示に HTML5 を利用したので、様々なネットワーク環境において、ウェブブラウザだけで手話映像を閲覧することが可能になった。音声と映像を手話通訳者側に届ける方法については課題が残されているが、利用者のスマートフォンなどのウェブブラウザに対して、通訳者が自宅などから手話映像を配信することが可能になった。

ハイブリッド CG 手話表現は、すべて実写で表現する方法と比較して、身体部分をモーションデータ（時系列の骨格データ）に置き換えることができるため、通信量の削減が期待できる。試作システムでは、実写画像をモーション Jpeg のような形式で送信したため、従来法との通信量の直接比較をしなかった。今後は従来法と同様なエンコードを行い、通信量がどの程度削減されるか評価を行う必要がある。

6. まとめ

本研究では、ウェブベースの遠隔手話通訳システムの構築を目指し、ウェブブラウザで遠隔手話通訳を実現するための重要な技術的要件の 1 つである手話映像の送受信方法の基礎的な検討を行った。ハイブリッド CG 手話で手話映像を配信することができる試作システムを構築し、表示フレームレートについて評価を行った。今後は、現場の音声と映像を手話通訳者に届ける方法および遠隔手話通訳に必要な機能の検討と実装を行う予定である。

参考文献

- [1] 松岡, 若月, 河野, 読みやすさを考慮した匿名手話映像の生成法, HCG シンポジウム 2014 講演論文集, 2014;p.507-512.