

クラウドソーシングによる手話文字通訳に関する研究

若月大輔²⁾, 田中康平¹⁾, 皆川洋喜²⁾

筑波技術大学 大学院技術科学研究科¹⁾, 産業技術学部²⁾

キーワード: 手話文字通訳, 聴覚障害, 情報保障, クラウドソーシング

1. 背景と目的

我々は、講演者が手話で講演する際に、クラウドソーシングを活用して手話を文字に通訳する方法について検討している。日本では、手話通訳者が手話を読んで発声し、その音声をもとに字幕生成者が字幕を生成するのが一般的である。しかし、この方法は人件費と字幕表出の遅延が単純に2倍になる。そこで、クラウドソーシングで専門家ではない人々をワーカーとして手話を字幕に変換するシステムを試作した。

ワーカーは手話を読むことができる聴覚障害者などが適任である。しかし、彼らはタイピングが得意とは限らない。我々のシステムは、手話のライブビデオを細かく分けて配信し、部分的に文字化してもらい、それらの文を統合して字幕を生成する。これは、手話を直接的に字幕化できるだけでなく、同時に聴覚障害者が自ら手話がわからない人々を助けることができる機会を提供する環境を実現する。

本報告では、本研究で試作した手話文字通訳システムについて説明し、評価実験を行った結果について述べる。

2. 関連研究

クラウドソーシングを活用した手話文字通訳について、ワーカーが現場で手話文字通訳を行い、通訳を担当する区間（以下、通訳区間）の区切り方は区切担当ワーカーが意味の切れ目を意識して手動で区切る方法が提案されている [1]。この方法では、まず、1つの区間に3名以上の複数の入力担当ワーカーを割り当てて手話を文字化する。次に、入力された複数の結果を評価担当ワーカーが投票し最終的に字幕として表出する。しかし、区切担当ワーカーや内容によって、区間の長さが一定ではなく、入力ワーカーの能力に合わない長さの通訳区間が与えられることもある。また、1区間に必要な入力ワーカーは3名以上で、5グループ以上必要とされており、評価担当ワーカーも含めると、人的コストが高くなってしまふ。

そこで、我々は、ワーカーの負担を減らすために、通訳区間を一定時間で自動的に区切り、手話を動画として配信し、

1区間を1名のワーカーで担当して通訳を行う方法を提案した [2]。区間が区切られる位置によっては手話が途切れる可能性があるが、通訳漏れを防ぐために通訳区間を重複させた。また、各通訳区間の間で文が自然につながるように、前後の区間を担当する入力状況を確認できるようにした連係入力機能を導入した [3]。実験によって、試作したシステムの機能の評価と、通訳結果と実験参加者の操作ログの分析結果を示し、議論する。

3. ライブ動画配信に対応した手話文字通訳システム

本研究では、ライブ動画配信に対応し PC で手話文字通訳タスクを実行できる手話文字通訳システムを試作した。まず、手話話者が YouTube Live [4] でライブ動画を配信する。この際、通訳区間の管理を行う管理画面を用いて、動画のタイムコードをベースに通訳区間を自動で設定し、動画と同時に配信する。次に、通訳を担当するワーカーは、入力画面で配信された動画を見ながら、通訳区間が与えられたら通訳を始める。そして、通訳が終わったら、遅れた時間を取り戻すために早送りリアルタイムに追いつく。追いついたワーカーは通訳区間が与えられるまで通常の再生速度で動画を見ながら待機する。

図1に本システムの入力ワーカー用インタフェースを示す。画面左上に配信された手話のライブ動画、左下に文字を入力する領域（中段）と、前後の通訳区間を担当するワーカーの入力状況を確認できる領域（上段/下段）がある。



図1 入力ワーカー用インタフェース

右に動画をコントロールするボタン（再生速度の変更，1秒戻す，拡大／縮小など）を配置している。動画のコントロールについては，キーボードでも操作できるようにショートカットを割り当てた。

4. 実験と結果 [2][3]

4.1 実験の概要

試作したシステムの評価を行うために，手話が読める聴覚障害者をワーカとした手話文字通訳実験を行った。実験の流れは，まず，ワーカのタイピング速度 [5] を計測する。次に，本システムに慣れるための手話文字通訳，そして，本実験のための手話文字通訳，最後にアンケートに回答してもらう。

ワーカは聴覚障害がある本学の大学生 4 名とし，ノート PC で入力してもらった。本システムはライブ動画配信が可能だが，実験にはあらかじめ撮影した手話動画を使用した。通訳区間は 9 秒，通訳区間の重複時間を 1 秒とした。分析のために各ワーカの操作や入力ログを記録した。

4.2 結果

実験参加者の日本語の入力速度は，平均 1.9CPS (characters per sec., SD 0.3CPS) で，最大 2.4CPS，最小 1.6CPS であった。通訳区間の文字化に要した時間は，平均 26.0 秒 (S.D. 13.1 秒) であった。通訳区間の間の文字化の漏れは軽微なものも含めて約 33% で発生し，重複についても約 33% で発生していた。実験参加者の通訳区間の担当数は平均 12.3 回 (SD 5.4 回) であり，最大 21 回，最小で 7 回であった。

主観評価において，タスクの楽しさの 5 段階評価では，「やや楽しい」が 1 名，「どちらでもない」が 2 名，「やや楽しくない」が 1 名であった。前後の通訳区間の入力をリアルタイムで確認できる機能については 1 名のみが活用したと回答した。

4.3 議論

通訳時間の標準偏差から，ワーカによって通訳時間が異なることが分かった。通訳時間が早いワーカは入力速度も速く，通訳時間が遅いワーカは入力速度が遅い傾向にあったため，デバイスの慣れが 1 つの理由として考えられる。今後は，普及しているスマートフォンを用いたシステムを検討したい。

通訳区間の間の文の漏れと重複については，前者が 33%，後者が 33% であった。「前後の通訳区間の文を確認していた」というコメントから，自然に文をつなげることを意識していた者もいたが，「入力の負担が大きく前後の文を確認する余裕がなかった」という者もいた。今後は，入力とは別に修正を行う専用のタスクを設けてワーカを割り当て

ることによって，通訳漏れや重複を解決できるか検討したい。

ログを分析すると，すべてのワーカが 1 回目の通訳を行った後に，待機中になることなく連続して手話文字通訳タスクを行っていた状況が見られた。通訳が「やや楽しくなかった」と回答したワーカは，2 連続で通訳する状況が多く，負担が大きかったと考えられる。一方，「やや楽しかった」と回答したワーカが担当した通訳区間の数は先のワーカの 2 分の 1 程度であり，負担が小さかった。平均通訳時間が一番遅かったワーカを基準にシミュレーションした結果から，負担を下げるために必要なワーカ数は 11 名であった。今後は，ワーカの人数を増やすことでワーカの負担が実際にどの程度軽くできるかについて検討したい。

5. まとめ

ライブ動画配信を活用したクラウドソーシングによる手話文字通訳システムを試作し，実験により，本システムの機能の評価と，実験参加者の操作ログの分析結果を示し，議論した。

ワーカの入力速度が通訳時間に与える影響，通訳区間の間の文の漏れや重複の発生率および主観評価と手話文字通訳タスクの負担の関係について結果を示し議論した。

今後の課題はスマートフォンを入力デバイスとした効率的な手話文字通訳タスクの検討と，修正タスクの導入による手話文字通訳結果の質の向上である。

謝辞：本研究は，筑波技術大学 2019 年度学長のリーダーシップによる教育研究等高度化推進事業の助成を受けた。また，一部は JSPS 科研費 JP15K01056 ならびに JP19K02996 の助成を受けた。

参考文献

- [1] 白石 優旗，張 健偉，熊井 克仁，森嶋 厚行，“クラウドソーシングに基づく聴覚障害者によるリアルタイム文字情報保障手法の検討”，DEIM Forum 2016 E8-6
- [2] 田中 康平，若月 大輔，皆川 洋喜，“クラウドソーシングによる手話文字通訳のためのタスクの基礎的検討”，信学技法，vol.118，no.491，WIT2018-83，pp.165-170，2019 年 3 月。
- [3] 田中 康平，若月 大輔，皆川 洋喜，“クラウドソーシングを用いた手話文字通訳 ～ ライブ動画配信を活用したシステムの試作 ～”，信学技報，vol.119，no.322，WIT2019-39，pp.95-100，2019 年 12 月。
- [4] YouTube Live，<https://www.youtube.com/live?hl=ja&gl=JP>，2019/10/26
- [5] 日本語入力スピード測定，<https://flickromaji.okaryunoske.work/>，2019/10/23