

# クラウドソーシング支援による選択制自動議事録作成手法の提案

## Proposal for a Crowd-Sourcing Supported Choice-Based Automated Minute Taking Methodology

黒岩 舜<sup>†</sup>      中平 一六<sup>†</sup>      大島 千尋<sup>‡</sup>      小板 隆浩<sup>†</sup>  
 Shun Kuroiwa    Kazumu Nakahira    Chihiro Oshima    Takahiro Koita

\*

### 1. はじめに

近年、クラウド音声認識を利用した議事録作成は大きな注目を集めている。しかし、多人数会話において、クラウド音声認識による議事録作成では、1人1つのマイクを使用したピンマイク録音方式では高い精度である一方、会議において机の中央に設置したマイクにより録音する全周囲録音方式では精度が低い [1]。本研究では提案手法として、クラウド音声認識システムにより作成した議事録をクラウドソーシングのワーカーに提示する。ワーカーは人間の感性に近い議事録を選択し、議事録を修正することによって、議事録の作成方法を提案する。

提案手法により、精度の低い全周囲録音方式による議事録の精度向上が期待される。また、提案手法ではクラウド音声認識による議事録を作成し、クラウドソーシングを用いてワーカーが議事録を選択する。その後、選択した議事録を修正する。本研究では、提案手法の有用性を実証し、効果的な議事録作成手法を検討する。

### 2. 先行研究

クラウド音声認識クラウドソーシングによる議事録作成を比較した結果、コンテンツ認識においてクラウドソーシングの有用性を確認した [2]。ここで、コンテンツとは議事録におけるタイムスタンプと話者を除いた話者の言葉のみが書かれた議事録である。さらに、音源分割手法による議事録作成の有用性を確認した。ここで、音源分割方式とは、図 1 に示すように、会議音源を分割してクラウドソーシングにより議事録作成を行う手法である。以上より、クラウドソーシングにおいて、議事録を作成する際には長時間の議事録を 1 つのタスクで作成するより、短時間の議事録を複数のタスクとして作成し、まとめることにより議事録を作成することで精度が上がる事が示された。

しかし、河原らの研究によると、議事録に要求される音声認識精度は 75% であり [3]、黒岩らの研究では到達していない。本研究ではクラウド音声認識をクラウドソーシングと組み合わせて議事録を作成することで精度の改善を目指す。

音源分割方式

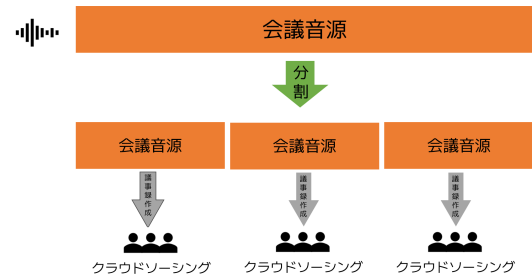


図 1: 音源分割方式

### 3. 提案手法

クラウドソーシングとは不特定多数の人間 (ワーカー) にプラットフォームを通して仕事 (タスク) を依頼する仕組みである。提案手法では、会議音源を複数のクラウド音声認識したものを、ワーカーに提示する。ワーカーは音源を聞き、提示された議事録から一つ正しいと思う議事録を選択し、議事録を修正する。

### 4. 実験

本研究では複数のクラウド音声認識により議事録を作成後、クラウドソーシングにおいて、クラウド音声認識により作成された議事録から、ワーカーが最も正しい選択肢を選択し、議事録を修正する。高精度の議事録作成を目指し、作成された議事録の精度について評価を行い、クラウドソーシングの有用性及び利用方法を検討する。

#### 4.1 方法

本実験では、会議音源として技術研究組合 新情報処理開発機構 (RWCP) により制作された RWCP-SP01 の会議 7 を用いる [4]。音源は会議音源からランダムに 20 秒切り取った音声を 6 つ用いる。クラウド音声認識は既存システムの Rimo Voice [5], Amazon Transcribe [6], whisper [7] を用いる。

方法を以下に示す。まず、3 つのクラウド音声認識によって音源を議事録作成し、ワーカーに提示する。その後、クラウドソーシングのワーカーが、音源を聞き、提示された 3 つの議事録から最も適切な議事録を選択する。選択した議事録をクラウド音声認識の議事録とする。さらに、ワーカーはクラウド音声認識の議事録を修正する。クラウドソーシングのプラットフォームとしてクラウド

<sup>†</sup> 同志社大学大学院, Doshisha University

<sup>‡</sup> 同志社大学, Doshisha University

ワークス<sup>†</sup>を用いる。クラウドソーシングのタスクは1つのタスクにつき3音源とし、2つのタスクを依頼する。報酬は1人辺り44円とし、10人のワーカーに依頼する。

## 4.2 評価指標

本実験ではコンテンツを評価する。

音声認識における一般的な評価指標として、WER(Word Error Rate)が挙げられる。WERは単語誤り率で式(1)で算出される。

$$WER = \frac{S_i + I_i + D_i}{T_i} \times 100 \quad (1)$$

ただし  $T_i$  は正解単語数、 $S_i$ 、 $I_i$ 、 $D_i$  はそれぞれ正解議事録と比較して発生した置換語数、挿入語数、脱落語数である。

本実験では WER 及び修正率、WER から算出される改善率を用いて評価する。

修正率とは修正後の議事録を正解としたとき、クラウド音声認識の議事録の単語誤り率である。修正率は式(2)で算出される。

$$\text{修正率} = \frac{A_i + B_i + C_i}{K_i} \times 100 \quad (2)$$

ただし  $K_i$  はクラウドソーシングによる修正後の議事録の単語数、 $A_i$ 、 $B_i$ 、 $C_i$  はそれぞれ修正後の議事録と比較して発生した置換語数、挿入語数、脱落語数である。

## 5. 結果

WER について述べる。本実験のタスクごとの WER を図2に示す。WER について、音源によって16%-40%の値をとっている。また、クラウド音声認識と比較して平均で4.7%改善している。次に、提案手法と先行研究を比較した結果を図3に示す。提案手法は黒岩らの先行研究と比較して1.94%と改善している。

修正率について述べる。本実験のタスクごとの修正率を図4に示す。修正率について、音源によって3%-25%の値をとっている。6つの音源の修正率の平均は14.38%である。

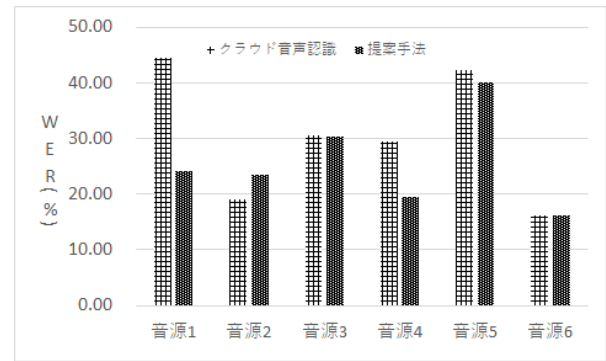


図 2: WER

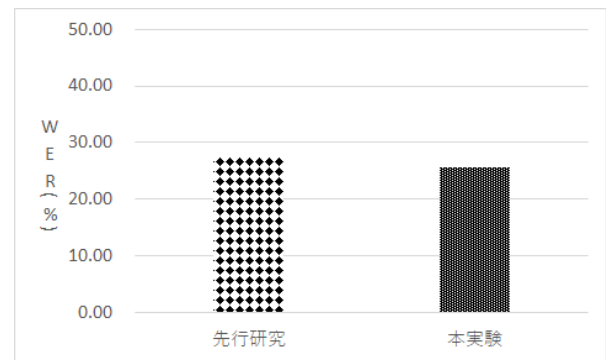


図 3: 先行研究との比較

## 6. 考察

音源1について、WERは20%改善している。修正率が25%であることから提案手法により改善された。クラウド音声認識では難しい音源を人間によって改善できている。

音源2について、WERが4.39%悪化している。クラウド音声認識の議事録が優秀であった一方、人間には議事録作成が難しい音源であった可能性がある。

音源5について、WERは他の音源と比較して9%以上高い。また、改善率も2.24%と低い。以上より、音源5は提案手法で改善できない困難な音源であると考えられる。全体を通して、修正率が高いほどWERがクラウド音声認識と比較して改善する傾向にある。これは選択肢をワーカーに示すことで、ワーカーが議事録をイメージしやすくなることで、議事録を修正しやすくなるが考えられる。ただし、音源の長さが短いことによる音源ごとの修正率、WERの差があるため、クラウドソーシングを行う音源を慎重に見極める必要がある。

## 7. まとめ

本研究ではクラウド音声認識による選択肢を使用した議事録を、クラウドソーシングによって修正する提案手法を示した。音源によってはWERが大きく改善した

<sup>†</sup><https://crowdworks.jp/>

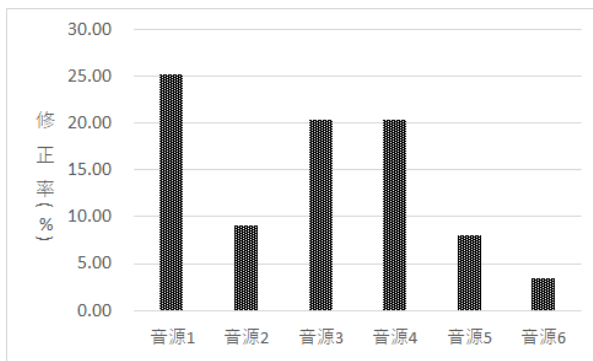


図 4: 音源ごとの修正率

一方、クラウド音声認識がクラウドソーシングより悪化する場合も見受けられた。ただし、修正率が高ければ、WER は改善する傾向にあるため、提案手法の有効性は示した。今後の課題は、いかにクラウドソーシングによる修正が必要な音源を切り分け、作業コストを安く、迅速に精度の高い議事録を作成するかが求められる。

## 参考文献

- [1] 中林一貴, 益井博史, 谷口忠大. 発話権取引を用いた会議の音声認識精度向上コミュニケーション場のメカニズムにおける制約活用による人工知能活用. 人工知能学会論文誌, Vol. 35, No. 5, pp. G-K31-1, 2020.
- [2] 黒岩舜. クラウドソーシングによる議事録作成支援の検討. 同志社大学, 2023.
- [3] 河原達也. 話し言葉の音声認識の進展—議会の会議録作成から講演・講義の字幕付与へ—. メディア教育研究, Vol. 9, No. 1, pp. 1-8, 2012.
- [4] RWCP (Real World Computing Partnership) 知的資源 WG. Rwcp 会議音声データベース (rwcp-sp01), 2006.
- [5] Rimo voice. <https://rimo.app/about/voice>.
- [6] Amazon transcribe. <https://aws.amazon.com/jp/transcribe/>.
- [7] Whisper 株式会社. <https://openai.com/blog/whisper/>.