

音環境分析と音声認識に基づく 議論振り返り支援インターフェースの開発

幸浦 弘昂¹ 白松 俊¹ 水本 武志² 林 加代子³

¹名古屋工業大学 大学院工学研究科情報工学専攻 ²ハイラブル株式会社 ³株式会社ソーシャルアクティ

1. はじめに

対面の議論による合意形成やアクティブラーニングを円滑にする上で、ファシリテーションの重要性が増している。ファシリテーションとは、会議やミーティングが生産的かつ効果的に進むように発言を促進し、話の流れを整理し、また参加者の合意形成をサポートする行為を表す。ファシリテーションの普及を目指し、日々ファシリテーションの研鑽を行う日本ファシリテーション協会 (FAJ) では、研鑽のための議論の振り返りが行われている。その際、議論のターニングポイントとなった重要発言を探りたいというニーズが存在する。

本研究では、そのような対面の議論の振り返りを支援するためのユーザインターフェースを開発する。音環境分析 (音源分離, 音源定位) に基づくハイラブル社の議論評価サービス (DAS; Discussion Assessment Service) [1]では、議論参加者の発言量遷移を図1のように可視化する機構を有するが、これを振り返りの手掛かりとして用いる。これに音声認識結果を併用し、議論のプロセスの変化や韻律情報から感情の変化や重要発言を自動抽出する。これらの、情報を図1の発現量遷移図に重ねることで、音環境分析と音声認識を併用した議論振り返り支援インターフェースを開発する。これにより、議論の内容面、プロセス面の両方から議論の振り返りを支援することを目指す。

2. システム概要

本研究で実装しているインターフェースの概要を図2に示す。図2右のグラフは各ユーザーの総発言量グラフであり、以下の2要素から構成される。

- グラフの形式として棒グラフを採用した、議

Developing User Interface for Supporting Review of Face-to-Face Discussion based on Audio Scene Analysis and Speech Recognition

Hiroaki KOURA¹, Shun SHIRSMATSU¹, Takeshi MIZUMOTO², Kayoko HAYASHI³
Nagoya Institute of Technology¹, Hylable Inc.², Social Acty Inc.³



図1: DASの発言量変化グラフ

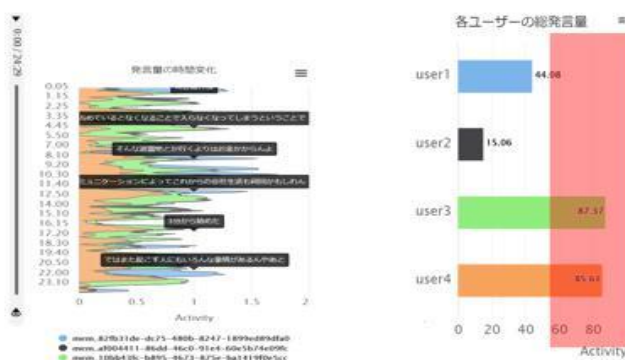


図2: 振り返りインターフェース概要

論中の各ユーザー総発言量を表したグラフ

- グラフ右にある、ユーザーのしゃべりすぎを警告する赤いゾーン

次に、図2右のグラフは発言量の時間変化を表しており、以下の3要素から構成される。

- 頭出しのための音声再生バー
- 縦軸を時間、横軸を発言量、グラフの形式として積み立てグラフを採用した、各ユーザーの発言量を時間変化で表した積み立てグラフ
- Google Cloud Speech-to-Text による音声認識結果をグラフ上に表示させた議事録

また、このグラフは図3のようにズームイン・ズームアウトすることが可能である。ズームイン前のグラフ上議事録には、その議論においてターニングポイントとなった重要発言だけを表示させておき、ズームインすることによって表示区間の発言すべてを標示させ

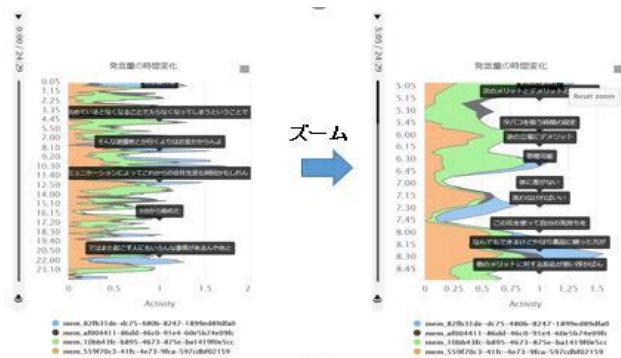


図 3: ズーム機構

る仕様になっている。重要発言の選定基準及び、選定基準に基づいた議論の重要発言抽出手法については後述する。

また、未だ実装中であるが、グラフ上に各ユーザーの発言感情に基づいたアイコンを付与する予定である。感情認識には、事例[2]を参考にして CNN を用いている。声優による感情発話のデータセット¹を用い、特徴量は MFCC、メルスペクトログラムを用いたところ、喜びについては適合度 0.75、再現率 0.88 で認識できた。

3. ヒアリング調査及び提案手法

ファシリテーション素人の我々が重要発言の抽出及び議論のスコア化の選定基準を設けるのは困難である。そこで、プロのファシリテータ集団である FAJ の協力を仰ぎヒアリングを実施し、選定基準を設けることとする。

3.1. 議論の重要発言選定基準

ヒアリングにより得られた基準を以下に示す。

1. 議論のプロセスを表す発言

「何が問題だっけ」「そもそも議題は…」
「そろそろまとめようか」のような、その時点での議論がどの段階にいるのかを示す語を含む発言

2. 賛否を表す発言

賛意を表す「うんうん」「なるほど」、否定を表す「いやいや」「違うよ」など。意見に納得しているのか否か、納得していない場合は議論が広がり、納得している場合は収束に向かうことが多い。

3. 頻出語など共通認識を持ちやすい言葉を含む発言

議論内容を表す重要語になりうる。

4. 発言量の増加などの議論のプロセスに変化をもたらす発言

発言量の量や参加者間分布に変化が起きる

直前にされた発言。プロセスが大きく変化しているという点で議論のターニングポイントとなりやすい

3.2. 議論の重要発言抽出提案手法

上記の 1~4 の選定基準における議論の重要語抽出手法については以下の手法を提案する。

1. キーワードマッチングによる抽出

1, 2 の議論の重要語発言選定基準に関しては、賛否を表す単語、プロセスそのものを表す単語をリスト化しておき、リストに含まれる単語を含む発言を重要語発言として抽出する。

2. TF-IDF による議論の重要語抽出

3 の議論の重要語発言選定基準に関しては、TF-IDF 議論中を用いて単語の重要度を測り、重要度の高い単語を含む発言を重要発言として抽出。

3. 発言量の変化点による抽出

4 の議論の重要語発言選定基準に関しては、窓を二つ用意し、二つの窓区間の類似度を測り、類似度が低くなった箇所の発言を重要発言として抽出。類似度調査にはユークリッド距離を用いる。

上記の 1~4 の手法で抜き出した発言をすべて重要語発言とするには発言数が多すぎるため、それぞれの発言に重みづけを行う。提案手法の 1, 4 によって抜き出した、主にプロセスに関する発言に関しては 2 の重みを与える。2, 3 の提案手法によって抜き出した発言には 1 の重みを与える。最終的に重みが大きいものを優先的に重要語発言とする。

4. おわりに

本稿では、音環境分析、音声認識に基づいた議論の振り返り支援インターフェースの実装を行った。まだファシリテーション支援への有用性の証明はできていないため、今後は研究目的でも挙げたように「議論の振り返りの手助け」「若手ファシリテータが重要発言を探す際の教育・研鑽のためのツール」としての有用性を実証していくことが課題に挙げられる。

謝辞 本研究の一部は JSPS 科研費 (17K00461) と JST CREST (JPMJCR15E1) の支援を受けた。

参考文献

- [1] 水本武志. 音環境分析に基づく議論評価サービスとその応用可能性. 人工知能学会第二回市民共創知研究会「みらいらぼつしま」予稿集, pp. SIG-CCI-002-04, 2017.
[2] 音声分類を色々なモデルや特徴量でやってみた - Qiita <https://qiita.com/kshina76/items/5686923dee2889beba7c>, (参照 2019-11-06)

¹ <https://voice-statistics.github.io/>