

## 発話タイミングに注目した複数人自由対話要約手法の検討

山村 崇 徳永 陽 嶋田 和孝  
(九州工業大学 情報工学部 知能情報工学科)

## 1 はじめに

従来より、対話の内容を容易に把握するために、対話の要点のみをまとめた要約文の作成手法が検討されている。しかし、研究対象としてはコールセンタや会議のような話の内容や構成が形式化されたものが多く、雑談のような規則性の無い自由対話の要約はあまり行なわれていない。

徳永ら [1] は、このような自由対話の要約には自由対話特有の特徴を考慮する必要があるとし、複数人自由対話を構成する各発話の重要性や笑いなどの非言語情報を考慮し、発話の抜粋による自動要約手法を提案している。

本論文では、先行研究が触れていない自由対話特有の特徴を捉えることで、要約手法の精度の改善を目指す。例えば、話者の発話タイミングや発話速度、発話オーバーラップが話者に依存することも自由対話特有の特徴だといえる。発話タイミングが遅い発話は思考時間が長く、発話速度が遅い発話はゆっくり強調された発話が多く、発話者同士の発話が重なりあう発話オーバーラップが発生する状況では、活発に話されていることが多いといえる。これらが起きるような状況では、比較的重要な発話が多くなると考えられる。本論文では、このような発話の時間情報に着目し、その有効性を検証する。

## 2 先行研究の重要文抽出

まず、基本となる徳永らの手法について説明する。徳永らは、機械学習による分類器を用いた重要文抽出を行なっている。

機械学習による分類器では、自由対話の特徴に着目した発話の言語情報と非言語情報を合わせた計 30 の素性を用いて、重要文を抽出している。発話の言語情報の素性は、発話単体や発話間、照応性に関する素性である。また、非言語情報の素性は会議のような形式的な対話ではあまり見られないが、発話の自由度が高い自由対話特有の特徴として、笑いの有無や場の盛り上がりといった非言語情報を素性として用いている。

## 3 提案手法

前述のように先行研究の手法では、非言語情報である盛り上がりや笑いを自由対話特有の特徴の一つと捉え、これらを機械学習の分類器の素性として採用している。しかし、実験データに対し笑いを含まない発話が少なかったことなどから、素性として有効に機能しないという問題が実験結果より確認されている。

そこで本稿では、自由対話特有の特徴を捉え、かつ実験データに対しスパースでもない非言語情報として、発話時間情報を機械学習による分類器の素性に追加することで、機械学習による重要文抽出の精度改善を試みる。本研究では、発話タイミングが遅い発話と発話速度が遅い発話、発話オーバーラップが存在する発話には重要文が多いと仮定し、以下の 3 つを新たに発話時間に関する非言語情報の素性として追加する。

- $A_1$  : 発話速度
- $A_2$  : 発話タイミング
- $A_3$  : 発話オーバーラップ

発話速度は、発話文の句読点を除く文字数から、その発話の発話時間の除算とした。また、発話タイミングはある発話の発話開始時間から、その 1 つ前の発話の発話終了時間の差を用いて算出した。最後に、発話オーバーラップはある発話の発話開始時間と、それまでの発話の発話終了時間とを比較して算出した。

表 1: 重要文抽出実験結果

組み合わせパターン	F 値
(1) 先行研究 (ベースライン)	0.784
(2) (1) + $A_1$	<b>0.794</b>
(3) (1) + $A_2$	0.784
(4) (1) + $A_3$	0.781
(5) (1) + $A_1, A_2$	0.791

表 2: 表 1 の (1) と (2) における各対話毎の抽出結果

	対話 1	対話 2	対話 3	対話 4
(1) ベースライン	0.820	0.776	<b>0.810</b>	0.813
(2) (1) + $A_1$	<b>0.821</b>	<b>0.783</b>	<b>0.810</b>	<b>0.824</b>

  

	対話 5	対話 6	対話 7	対話 8
(1) ベースライン	0.713	0.827	0.744	0.771
(2) (1) + $A_1$	<b>0.730</b>	<b>0.839</b>	<b>0.756</b>	<b>0.785</b>

## 4 実験

4 人が特定の話題で自由に話している 8 対話 (全 1295 発話) を実験対象データとし、全 8 対話がそれぞれテストデータとなるような leave-one-out 法を用いて、得られた 8 回の結果を平均して F 値を算出した。正解データには、3 名のアナウンサーによる発話の重要度評価の平均に基づき生成したものを使用した。新たに素性を追加した機械学習による分類器を作成し、抽出結果を正解データと比較することで重要文の抽出の実験を行った。

結果を表 1 に示す。ここで、表 1 においてベースラインとは先行研究の重要文抽出結果を意味する。表 1 より、今回単一の素性を追加した場合には (2) の発話速度を追加したときだけが、(1) のベースラインの F 値より上回っていることが分かる。そこで、これらの違いを検証するために表 2 に、表 1 の (1) ベースラインと (2) 発話速度を追加したときにおける各対話ごとの抽出結果の F 値を示す。表 2 より、すべての対話において発話速度を素性に追加した場合の F 値がベースラインの F 値以上であることが確認できる。これより、発話速度は非言語情報の素性において、要約手法の精度改善に貢献しているといえる。

発話速度の有効性が確認された一方で、発話タイミングと発話オーバーラップが有効に機能しなかったその原因としては、発話タイミングは話の重要度だけではなく、発話者の外向性や発話者同士の社会的関係など様々な要因からも生じるためであると考えられる。また、発話オーバーラップについても同様であると考えられる。

## 5 おわりに

本論文では、複数人自由対話を対象とした要約手法において、話者の発話速度が要約手法の精度改善に貢献することが確認された。

## 参考文献

- [1] Yo Tokunaga, Kazutaka Shimada, "Multi-party conversation summarization based on sentence selection using verbal and nonverbal information", International Conference on Smart Computing and Artificial Intelligence (ICSCAI 2014), 2014.