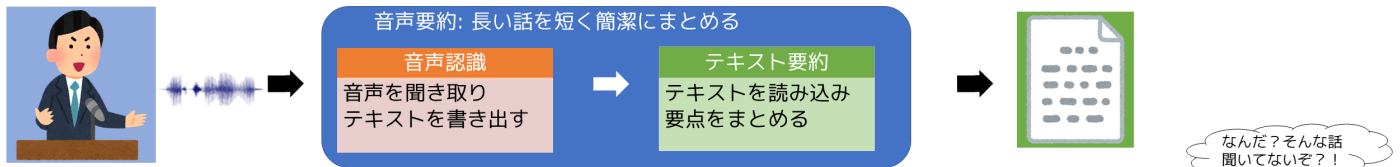


どんな研究	人の話の意図を正しく汲み取る機械の実現をめざして、 長い話の要点を的確にまとめる音声要約 の研究をしています。音声要約は、人の話を聞く音声認識と、聞いた内容をまとめるテキスト要約で構成されますが、前段の音声認識誤りを完全に無くすことは困難です。本展示では、 音声認識誤りに頑健な音声要約 技術を提案します。
どこが凄い	音声認識器がいくつかの単語を誤認識しても、話全体の文脈や単語間の関連性を考慮しながら正しく要約する音声要約を実現しました。複数の音声認識器を使うことで、それぞれの 認識結果に誤りが含まれていても、精度の高い音声要約を実現 します。また本技術は、音声要約のみならず音声翻訳などにも応用できることを確認しています。
めざす未来	人と人、人と機械がお互いに深く理解しあえる未来 をめざします。人の言っていることを一言一句聞き取るだけでなく、 人が何を意図しているのか、感じているのか 、を理解し、人と人、人と機械のより豊かなコミュニケーションを支える音声言語処理技術を研究していきます。

目的：聞き間違えても正しく理解する

*本研究の一部はカーネギーメロン大学との共同研究です

音声要約：講演などの長い音声から要点をまとめる技術
高度な音声理解を音声認識とテキスト要約を組み合わせることで実現



しかし、音声認識の聞き間違いがテキスト要約に大きな影響を与えることも！

発話内容: 当社のサーモン大好き弁当の、内容を再考します。検討結果はWEBで発表します。
認識結果: 当社の**さぁ問題** 好き弁当の、内容は**最高**です。検討結果はWEBで**発売**します。

要約結果: **大容量の「さぁ問題好き弁当」をオンラインで発売!**

提案：音声認識誤りに頑健な要約

聞き間違いが無い音声認識の実現は困難
・複数の音声認識器を使い、
正解単語がいずれかの結果に含まれるよう出力

聞き間違いがあるかもしれないという前提でテキスト要約
・最先端の言語処理器 (BERT*) を活用して、単語の意味や関係性を解析
・複数の認識結果を比較しながら、文脈を考慮して正確に要約

* Bidirectional Encoder Representations from Transformers

音声認識器A 当社の**さぁ問題** 好き弁当の、内容は**最高**です。検討結果はWEBで**発売**します。

音声認識器B 当社のサーモン**好き弁当**を再考します。検討結果はWEBで発表します。

音声認識器C 当社のサーモン大好き弁当は、内容を**最高**します。検討結果はWEBで**発佈**します。

文脈を考慮した仮説統合 (BERTを利用した音声要約器を構築)

1つ1つの認識は間違っても、認識器A,B,Cの単語の意味や内容を比較すると

- ・「さぁ問題好き弁当」より「**サーモン大好き弁当**」の方が文として自然
- ・話の後ろに、「検討結果」が続くということは「**再考**」であるべき
- ・「発表」と「発佈」も似た意味で、「検討結果」なら「**発表**」が自然

以前、コストカットをお願いしてた件が

要約結果: **サーモン大好き弁当の内容量を変更**

提案法により、
認識器の聞き間違いによる性能劣化を**最大30%低減**することができます!

関連文献

- [1] T. Kano, A. Ogawa, M. Delcroix, S. Watanabe, "Attention-based multi-hypothesis fusion for speech summarization," in *Proc. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 487–494, 2021.
- [2] 叶高朋, 小川厚徳, マーク・デルクロア, 渡部晋治, "音声要約のためのBERTを用いた認識仮説統合," 日本音響学会2022年春季研究発表会
- [3] T. Kano, A. Ogawa, M. Delcroix, S. Watanabe, "Integrating multiple ASR systems into NLP backend with attention fusion," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2022.

連絡先

叶 高朋 (Takatomo Kano) メディア情報研究部 信号処理研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp