

## 誰がどのように話しても正確に聞き取ります

～話者や発話スタイルの多様性に頑健な音声認識技術～

## どんな研究

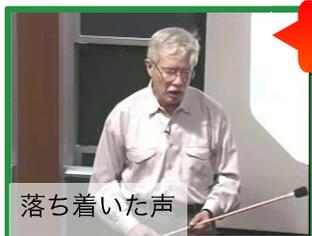
音声認識技術は高度な進化を遂げ、丁寧な話し方で正しい文を話せば高い精度で認識できるようになってきています。しかし、未だに話す人や発話の丁寧さ等によって性能が上下してしまいます。本展示ではこのように**多様な音声を適切に処理**する音声認識技術について紹介致します。

## どこが凄い

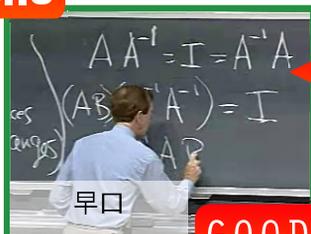
音の多様性を吸収する**多段の情報抽出方法**を大量の音声データから自動で学習し、その上で、音声と言語の対応を、音声言語一体型統計モデルにより表現する手法を考案しました。この手法で音声と言語の両面から多様性を吸収することで、世界トップクラスの性能を実現しました。

## めざす未来

誰が何処で何をどのように話しても正確に聞き取ることのできる、頑健な音声認識技術を目指します。このような音声認識技術を用いることで、対話エージェント技術、音声データマイニング技術のさらなる高度化を可能にし、**様々な情報機器を自然に活用できる社会**を目指します。



Hello



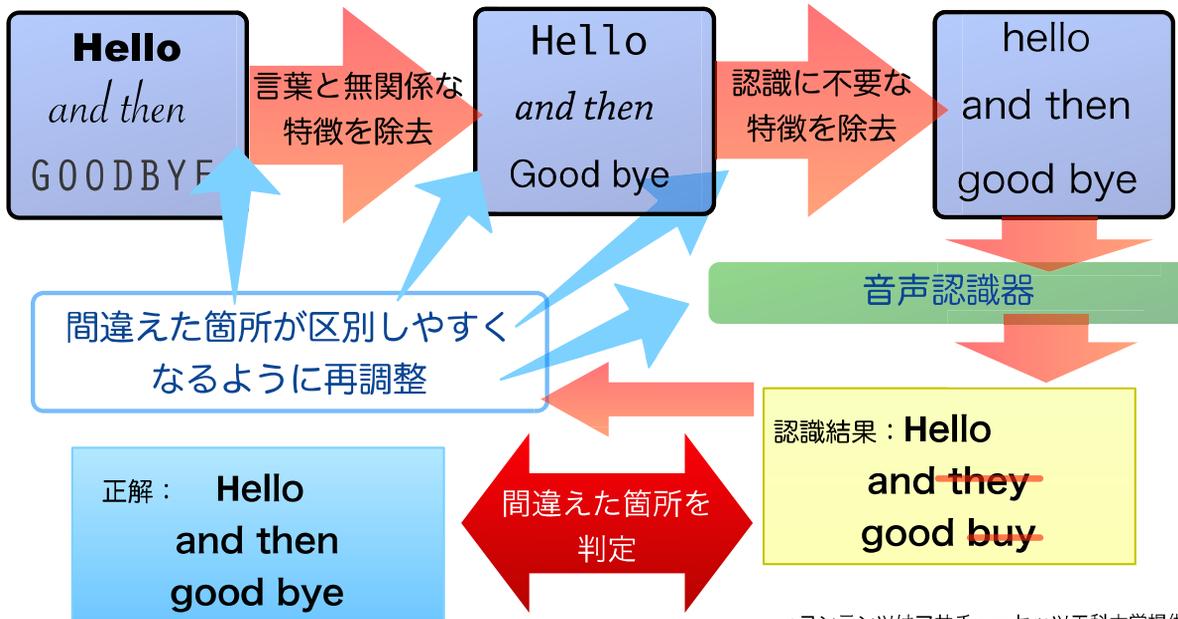
and then



GOODBYE

音声は話者や発話スタイルによって多様に変化する

シンプルな処理を多段に積み重ねることで、音声の多様性を抑圧



※コンテンツはマサチューセッツ工科大学提供

## 関連文献

- [1] Y. Kubo, T. Hori, A. Nakamura, "Integrating deep neural networks into structured classification approach based on weighted finite-state transducers," in *Proc. 13th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2012.
- [2] Y. Kubo, T. Hori, A. Nakamura, "Large vocabulary continuous speech recognition based on WFST structured classifiers and deep bottleneck features," in *Proc. The 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.

## 連絡先

久保 陽太郎 (Yotaro Kubo)   メディア情報研究部 信号処理研究グループ  
E-mail: kubo.yotaro{at}lab.ntt.co.jp ({at}の部分>@に置き換えてください)