

19

世界中の言葉を聞き分けます

～多言語音声識別・認識技術～



どんな研究

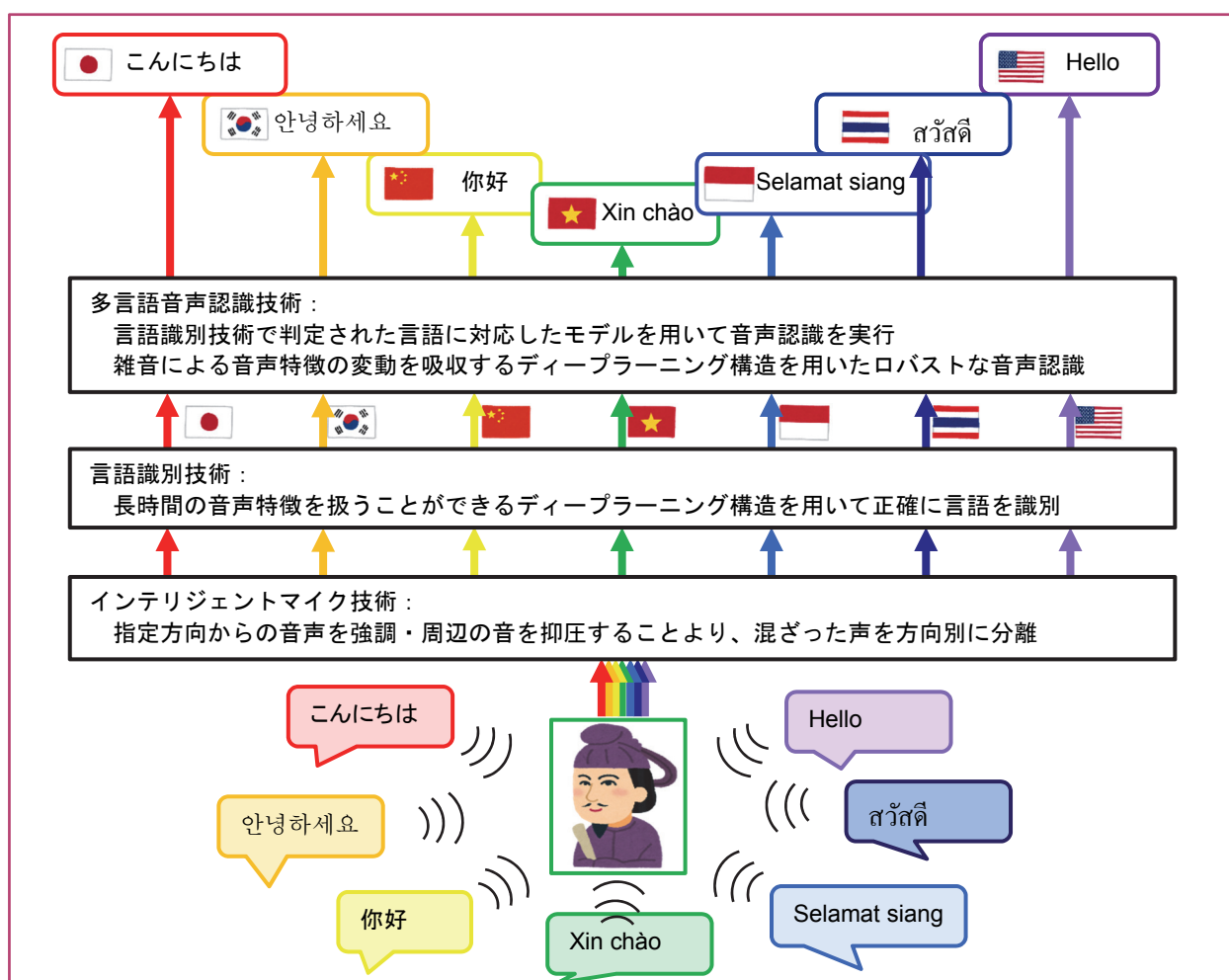
同時に話された**世界中の言葉を瞬時に聞き分けて文字に変換する研究**です。声を分離して聞き取るインテリジェントマイク技術、音声がどの言語で話されたかを判定する言語識別技術、音声を文字化する多言語音声認識技術から構成されています。

どこが凄い

インテリジェントマイクは、**最大7人が同時に話した声を分離**することができる技術です。また、雑音下での音声認識の国際技術評価で世界1位の精度を達成した音声認識技術を多言語対応し、言語識別技術と組み合わせ**てアジア・パシフィック10言語の認識**を可能としました。

めざす未来

今後は対応言語を増やし、**世界のあらゆる言葉の認識**を目指します。また、対話技術・感情認識技術などを組み合わせることにより、日本を訪れる外国からの観光客に対して、**自然で気の利いた応対**ができるような技術の実現を目指します。



関連文献

- [1] T. Yoshioka, N. Ito, M. Delcroix, A. Ogawa, K. Kinoshita, M. Fujimoto, C. Yu, W. J. Fabian, M. Espi, T. Higuchi, S. Araki, T. Nakatani, "The NTT CHiME-3 system: Advances in speech enhancement and recognition for mobile multi-microphone devices," in *Proc. ASRU2015*, pp. 436-443, 2015.
- [2] R. Masumura, T. Asami, H. Masataki, Y. Aono, "Parallel Phonetically Aware DNNs and LSTM-RNNs for Frame-by-Frame Discriminative Modeling of Spoken Identification," in *Proc. ICASSP2017*, 2017.

連絡先

政瀧 浩和 (Hirokazu Masataki) メディアインテリジェンス研究所
E-mail: masataki.hirokazu(at)lab.ntt.co.jp