

[招待講演] 音声感情認識の分野動向と 実用化に向けたNTTの取り組み

安藤 厚志^{1,a)}

概要：音声から話者の感情を推定する音声感情認識が盛んに研究されている。この技術は、コンタクトセンタにおける顧客満足度の自動分析、運転者の心理状態推定、対話システムにおける人間らしい振る舞いの実現など様々な応用が期待されている。一方で、感情は主観的かつ曖昧な概念であるため、正解の定義が困難である、感情発話の収集・アノテーションコストが大きいなど多くの実用上の課題が存在する。本招待講演では、これらの音声感情認識に関する研究動向から具体的な応用例までを概説する。まず初めに、音声感情認識の問題設定や近年の代表的な手法などの研究動向を述べる。その後、音声感情認識の応用例として、筆者らが研究・実用化したコンタクトセンタ通話における顧客満足度推定技術 [1,2] について紹介する。

[Invited Talk] Survey on Speech Emotion Recognition and its Application in NTT

参考文献

- [1] A. Ando, R. Masumura, H. Kamiyama, S. Kobashikawa and Y. Aono, “Hierarchical LSTMs with Joint Learning for Estimating Customer Satisfaction from Contact Center Calls,” *Proc. Interspeech*, pp. 1716–1720, 2017.
- [2] A. Ando, R. Masumura, H. Kamiyama, S. Kobashikawa, Y. Aono and T. Toda, “Customer Satisfaction Estimation in Contact Center Calls Based on a Hierarchical Multi-Task Model,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 715–728, 2020.

¹ NTT メディアインテリジェンス研究所
NTT Media Intelligence Laboratories, Yokosuka, Kanagawa
239-0847, Japan

^{a)} atsushi.ando.hd@hco.ntt.co.jp