

徳田 恵一

名古屋工業大学 大学院工学研究科
教授

コンテンツ生成の循環系を軸とした次世代音声技術基盤の確立

§1. 研究実施体制

(1) 名古屋工業大学・徳田グループ

- ① 研究代表者: 徳田 恵一 (名古屋工業大学 大学院工学研究科、教授)
- ② 研究項目
 - ・研究グループ間の連携調整
 - ・基盤技術および関連ソフトウェアの高速化
 - ・ユーザによるコンテンツ生成環境の構築

(2) 名古屋工業大学・山本グループ

- ① 主たる共同研究者: 山本 大介 (名古屋工業大学 大学院工学研究科、准教授)
- ② 研究項目
 - ・音声対話サービスのための情報基盤ネットワークの設計と活用
 - ・統一データベースに基づく音声対話コンテンツの動的作成
 - ・Web インタフェースを用いた音声対話コンテンツの編集と共有

(3) エジンバラ大学・レナルズグループ

- ① 主たる共同研究者: Steve Renals (University of Edinburgh, School of Informatics, 教授)
- ② 研究項目
 - ・多言語化
 - ・対話コンテンツの自動インデキシング／リンキングおよびそれに基づくユーザ分析
 - ・ユーザが作成した対話コンテンツデータベースを利用した能動的対話システム

§2. 研究実施の概要

本プロジェクトでは、魅力的な音声対話システムを作成するための基盤技術、基盤ソフトウェアの高度化、音声対話コンテンツを生成するための仕組み作り、実証実験に基づくコンテンツの分析と評価に取り組んでいる。本年度は、継続的に基盤技術と関連ソフトウェアの高度化に取り組むと同時に、ユーザがより手軽に音声対話コンテンツを作成することを可能にする音声対話コンテンツ生成環境の開発に取り組み、さらに、公共空間における実証実験および音声対話システム構築に関する講習会を実施した。

1) 音声基盤技術の改良と関連ソフトウェアの高度化

本年度は、音声合成のためのより高精度なモデル学習手法や、音声認識のためのリカレントニューラルネットワーク型言語モデルの高精度化等の音声言語情報処理の基礎研究を継続的に実施した。また、音声対話システム基盤ソフトウェアとして、各種オープンソースソフトウェア(音声対話システム構築ツールキット MMDAgent、音声認識エンジン Julius、HMM 音声合成ツールキット HTS、HMM 音声合成エンジン hts_engine API、日本語音声合成システム Open JTalk、歌声合成システム Sinsy、音声信号処理ツールキット SPTK)を改良し、新バージョンを公開した。また、英語版の各種ソフトウェアを開発し、英語版音声対話システムの構築を進めた。

2) ユーザによる音声対話コンテンツ生成環境の研究開発

ユーザ生成型の音声対話コンテンツを実現するためには、ユーザがネットワークを介して、手軽に音声対話コンテンツを作成できることが重要である。そこで、タブレット端末を用いて、音声とタッチインタフェースで手軽に音声対話コンテンツを作成できるシステムの開発を実施した。さらに、近年流行しつつあるクラウドソーシングの概念を音声対話コンテンツの作成に取り入れた、クラウドソーシングモデルの検討と試作を実施した。これは、インターネットを介して複数のユーザで対話シナリオの作成を分担するための仕組みである。

また、既存のスマートフォン向け音声対話システムは、端末単体で動作するか、サーバとの通信しかできず、効率的に複数台の端末が連携した複雑な対話シナリオを構築することができなかった。そこで、複数台のスマートフォン版 MMDAgent を搭載した端末間の連携を実現するために、BluetoothとNFCを用いたP2P通信機構を用いた通信が可能なエージェントフレームワークを開発することにより、簡単な記述でエージェント間のメッセージ送受信を可能な仕組みについても研究した。

3) 実証実験

半田市市役所新庁舎と名古屋工業大学学生スペース(夢ルーム)に実証実験システムを設置した。学生スペースに設置した提案システムは、PCだけでなくNFCを用いてスマートフォンからも容易に音声対話コンテンツを投稿可能である。

さらなる音声対話システムの普及を促進するためには、MMDAgent が多くの人に使用されることが重要である。そこで、人工知能学会と連携することにより、大阪と東京の二か所で一般の人を対

象とした MMDAgent 講習会を実施した。講習会の資料として MMDAgent に関するマニュアルを整備すると同時に Web で公開した。また、JST 共催のモノアプリハッカソンスーパーに参加し、MMDAgent をシーズとして提供した。優秀な成績を収めた多くのチームにおいて提供した MMDAgent が利用され、提案システムの拡張性の高さ、可用性の高さが示された。

1. Daisuke Yamamoto, Keiichiro Oura, Ryota Nishimura, Takahiro Uchiya, Akinobu Lee, Ichi Takumi, Keiichi Tokuda, “Voice Interaction System with 3D-CG Virtual Agent for Stand-alone Smartphones,” Proceedings of the 2nd International Conference on Human Agent Interaction (HAI 2014), ACM digital library, pp.320-330, October 2014.
2. Shinji Takaki, Yoshihiko Nankaku and Keiichi Tokuda, “Spectral modeling with contextual additive structure for HMM-based speech synthesis,” IEEE Journal of Selected Topics in Signal Processing, vol. 8, issue. 2, pp. 229-238, April 2014.
3. Siva Reddy Gangireddy, Fergus McInnes, and Steve Renals, “Feed-forward pre-training for recurrent neural network language models,” proceedings of Interspeech 2014, pp. 2620-2625, Singapore, September, 2014.