

戦略的創造研究推進事業 CREST

研究領域

「共生社会に向けた人間調和型情報技術の構築」

研究課題

「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」

研究終了報告書

研究期間 平成23年10月～平成29年3月

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)

§ 1. 研究実施の概要

(1) 実施概要

近年、音声情報処理技術の発展に伴い、スマートフォンなどを用いて音声対話システムを誰もが簡単に利用可能になりつつある。一方で、音声対話の内容(コンテンツ)は主に技術者あるいはそれに近い人達によって作られ、画一的になっているものが多い。本プロジェクトの目的は、最新の音声技術を基盤として、ユーザーの創意工夫で魅力的な音声対話システムを作成・利用できる環境を用意することで、「ユーザー生成型コンテンツ」として音声対話システムの多様な応用・展開を広く促し、もって次世代の音声情報処理技術基盤を確立することにある。本プロジェクトの期間中、基盤技術・基盤ソフトウェアの高度化に取り組むと同時に、音声対話コンテンツ共有サービスの開発、音声対話コンテンツ生成環境の拡充に取り組み、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などを実施した：

1) 基盤技術の高度化(主に徳田グループ)

より魅力的な音声対話システムを構築するためには音声基盤技術や基盤ソフトウェアの高度化が必要である。本プロジェクトの期間中に、コンテキストの加算構造を考慮した HMM 音声合成や、ディープニューラルネットワークに基づく音声合成の品質改善、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による音声認識性能の改善をはじめとした各種音声言語情報処理の基礎研究を実施した。また、音声インタラクションシステム構築ツールキット MMDAgent をはじめとした関連ソフトウェア(HMM 音声合成ツールキット HTS, 音声認識エンジン Julius, 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 等)を継続的に改良し、新バージョンを公開し続けた。さらに、MMDAgent のマルチプラットフォーム化などを実施した。

2) コンテンツ生成のための仕組み作り(主に山本グループ)

ユーザー生成型の音声対話コンテンツを実現するためには、ユーザー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。タブレット端末による音声とタッチインター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、クラウドソーシングによる複数ユーザーの対話シナリオ作成分担の仕組みの試作、エージェント間のメッセージ送受信を可能とする Bluetooth と NFC を用いた P2P 通信機構を持つエージェントフレームワークの開発をはじめとする各種研究開発を実施した。また、音声対話コンテンツをユーザー間で共有可能にするサービス MMDAgent SHARE を立ち上げた。この他、利用頻度の高い対話や実用性を考慮した対話などの基本的な音声対話コンテンツをまとめた基本音声対話ライブラリの構築、MMDAgent で動作可能なライセンスフリー 3D モデルの作成、モジュール型音声対話システムの構築などに取り組んだ。

3) 実証実験に基づくコンテンツの分析と評価(主にレナルズ・山岸グループ)

言語に依存しない知見を得るためには、日本語のみならず他の言語による実証実験を並行して進める必要がある。英語の対話に適した音声コーパスの構築、英語による音声対話システムの開発、ヨーロッパ圏で受け入れられやすいリアリスティックな 3D エージェントの作成、ユーザーの発話内容に応じて応答内容を自動作成するクエリシステムの開発などを実施した。また、コンテンツの魅力を解明するために、名古屋工業大学正門前、半田市観光協会などに提案システムを設置・運用する等の社会実装を推進し、実際に一般ユーザー自身が音声対話コンテンツを作成し運用する環境を実現した。

これらの活動を通じて、実証実験の場だけでなく、インターネット上でも公開されたソフトウェアツール群を用いた実に多様にコンテンツが生成されていることが確認された。更に、これらの成果を

踏まえて最終年度には、音声対話コンテンツの集積・分析を行うための音声対話コンテンツサーバー「MMDAgent SHARE」を10月1日に公開し、ユーザーの意見収集やユーザー間の情報共有のためにMMDAgent ユーザーミーティング「MMDAgent DAY!」を同じく10月1日に開催した。

プロジェクトの最終成果すべてをまとめ、プロジェクト後につなげるため、ソフトウェア群やリファレンス・マニュアルだけでなく、コンテンツ制作支援ツール、講習会用のスライド、MOOC/OCW 用の動画、魅力ある音声対話システムがユーザー生成により生み出されていくための「設計指針」などを併せ、統合的にまとめたオールイン・パッケージとして「エンサイクロペディア MMDAgent」を公開した。

(2) 顕著な成果

<優れた基礎研究としての成果>

1. 【音声関連分野にインパクトを与えた基盤技術】

本プロジェクトでは音声対話システムの基盤技術として音声情報処理技術の研究を進めており、特に徳田らによって提案された「統計的パラメトリック音声合成方式」に関する研究が大きな成果を上げている。本方式は、合成音声における感情や話者性などの多様性を容易に実現できる等の優位性から、次世代の音声合成技術として国際的にも認知され、国内外の重要な賞を複数受賞するなど高い評価を受けている。この研究成果は音声研究に新たな研究分野を切り開くとともに、音声対話システムを始めとする多くの関連研究においても大きなインパクトを与えるものである。

2. 【オープンソース戦略に基づく研究用プラットフォーム】

音声信号処理に関する一連の基礎研究を土台として、複数の音声基盤ツール群の開発、ならびにオープンソースによる公開を実施してきた。その一つである、音声対話構築ツールキット MMDAgent では、有限状態遷移(FST)記法により柔軟な音声対話を可能にすると同時に、デジタルサイネージやスマートフォンによる動作も実現してきた。これらのツール群はCREST研究期間中に世界中から延べ 80 万件以上ダウンロードされ、国内外で広く認知されるだけでなく、人工知能学会主催の講習会が開催されるなど、学術的な研究用プラットフォームとして広く利用されている。

3. 【音声対話システムにおけるコンテンツ分離の実証】

本プロジェクトでは、専門知識の無い一般ユーザーが音声対話コンテンツをモジュール(辞書・ボイス・対話・モデル)ごとに自由に組み替えて選択・構築・共有できるフレームワークを提案した。これにより、これまで考えられてこなかったクラウド上での音声対話コンテンツの共同編集などが実現可能になった。さらに、コンテンツ群のパッケージ化を推し進め、ユーザーはワンクリックで様々な音声対話コンテンツを体験できるようになった。最終年度には、音声対話コンテンツを共有できるサービスである「MMDAgent SHARE」を一般公開し、音声対話システムのコンテンツ分離を実証した。

<科学技術イノベーションに大きく寄与する成果>

1. 【一般ユーザーによる音声対話システムの実運用】

本プロジェクトでは、開発した音声対話システムを、キャンパス内を始めとした複数の公共空間で運用するとともに、サイエンスアゴラや名工大テクノフェアへの展示を行う等、様々なシーンにおいて容易にシステムを適応し、運用が可能であることを示した。さらには、2014年7月にイギリス・ロンドンで開催された HYPER JAPAN を始めとした様々なイベントにおいて、本プロジェクトとは独立に、ユーザーの手による様々な対話システムの構築が進んでおり、ユーザーによるコンテンツ生成のための基盤が確立しつつある。

2. 【新産業の創出】

本プロジェクトにおける研究成果を学内ベンチャーが橋渡しをすることによって、CeVIO プロジ

ェクトと連携した産業応用を強力に推進した。また、CeVIOプロジェクトが販売している音声対話デジタルサイネージや音声創作ソフトウェア以外にも、フューレックのスマートフォンアプリやシャープのモバイル型ロボット電話、Google の Android OS を始めとして多くの外部企業の製品にも本プロジェクトの研究成果が搭載されており、音声対話システムに関連した新しい産業の創出につながっている。

3. 【新分野へのスピノフ】

本プロジェクトにおける基盤技術の研究開発は非常に高い成果を得ることができた。論文投稿等による関連学術分野への貢献はもちろんのこと、これらの成果をベースとして、多様なスピノフプロジェクトが始まっている。声のドナーを集めることで医療分野に貢献する NII VoiceBank プロジェクトのように社会実装に繋がっているものや、声の成りすましに焦点を当てた国際ワークショップである ASVspoof Challenge のように社会に新しく問題提起しているものなど、様々な分野における新展開の起点となっている。

§ 2. 研究実施体制

(1) 研究チームの体制について

① 徳田グループ

研究参加者

氏名	所属	役職	参加時期
徳田 恵一	名古屋工業大学	教授	H23.10～H29.3
李 晃伸	同上	教授	H23.10～H29.3
南角 吉彦	同上	准教授	H23.10～H29.3
大浦 圭一郎	同上	研究員	H24.4～H29.3
橋本 佳	同上	研究員	H24.4～H29.3
玉森 聡	同上	D2～4	H23.10～H26.3
中村 和寛	同上	D1～3	H23.10～H26.3
高木 信二	同上	D1～3	H23.10～H26.3
伊藤 直晃	同上	D1～5	H23.10～H28.3
沢田 慶	同上	D1～2	H25.4～H27.3
吉村 建慶	同上	D1～2	H27.4～H29.3
夏目 龍司	同上	M2	H27.12～H28.3
小中 彩貴	同上	M1～2	H27.12～H29.3
村上 拓也	同上	M1～2	H27.12～H29.3

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
- ・ ユーザーによるコンテンツ生成環境の構築

② 山本グループ

研究参加者

氏名	所属	役職	参加時期
山本 大介	名古屋工業大学	准教授	H23.10～H29.3
内匠 逸	同上	教授	H23.10～H29.3
打矢 隆弘	同上	准教授	H23.10～H29.3
西村 良太	同上	研究員	H24.4～H27.3
高橋 直久	同上	教授	H24.4～H29.3
Pablo Martinez Lerin	同上	D2～D3	H23.10～H25.3
伊藤 翔太	同上	M2	H23.10～H24.3
中島 悠太	同上	M2	H23.10～H24.3
金子 智哉	同上	M2	H23.10～H24.3
奥村 文雄	同上	M2	H23.10～H24.3
板津呂 翔	同上	M2	H24.4～H25.3
立松 直也	同上	M2	H24.4～H25.3
吉田 真基	同上	M1～M2	H24.4～H26.3
柴川 元宏	同上	M1～M2	H24.4～H26.3
石原 佑哉	同上	M2	H24.4～H25.3
田川 達司	同上	M2	H24.4～H25.3
若林 敬太郎	同上	M1～M2	H26.4～H28.3

石川 博規	同上	M1～M2	H26.4～H28.3
清水 政嗣	同上	M2	H26.4～H27.3
水野 恵輔	同上	M2	H26.4～H27.3
森 智哉	同上	M1～M2	H26.4～H28.3
岸上 友樹	同上	M2	H26.4～H27.3
久米 雄介	同上	M2	H26.4～H27.3
仲野 良佑	同上	M2	H26.4～H27.3
松下 祐一	同上	M1～M2	H26.4～H28.3
大竹 聡史	同上	M1～M2	H26.4～H28.3
加藤 正隆	同上	M1～M2	H27.4～
小林 隆弘	同上	M1～M2	H27.4～
日比野 雅人	同上	M2	H27.4～
可児 龍平	同上	M1～M2	H27.4～
田中 佑太郎	同上	M1～M2	H27.4～
大野 健	同上	M2	H28.4～
加藤 義隆	同上	M2	H28.4～
宮原 悠輔	同上	M2	H28.4～
石田 大明	同上	M2	H28.4～
佐藤 清隆	同上	M1	H28.4～
杉寄 諒吾	同上	M1	H28.4～
石田 雄登	同上	M1	H28.4～
堤 修平	同上	研究員	H27.4～
浮田 弥	同上	M1	H28.4～
加藤 りか	同上	M1	H28.4～
河原 誠斗	同上	M1	H28.4～
田中 亮佑	同上	M1	H28.4～
山口 大介	同上	M1	H28.4～
佐野 敦志	同上	M2	H28.4～
大橋 洋介	同上	M2	H28.4～

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 インターフェースを用いた音声対話コンテンツの編集と共有

③ レナルズグループ

研究参加者

氏名	所属	役職	参加時期
Steve Renals	エジンバラ大学	教授	H23.10～H29.3
Junichi Yamagishi	同上	研究員	H23.10～H29.3
Simon King	同上	教授	H23.10～H29.3
Hiroshi Shimodaira	同上	講師	H23.10～H29.3
Qiang Huang	同上	研究員	H25.3～H28.3
Rasmus Dall	同上	博士課程学生	H24.9～H29.3
Siva Reddy Gangireddy	同上	博士課程学生	H24.9～H29.3

David Braude	同上	博士課程学生	H26.4～H28.4
Kathrin Haag	同上	博士課程学生	H26.4～H29.3
Cristophe Veaux	同上	研究員	H24.3～H28.9
Jennifer Oxley	同上	技術補助員	H24.4～H28.4
Magda Konkiewicz	同上	技術補助員	H24.7～H24.9
Gergely Bakos	同上	プログラマ	H28.1～H28.8
Kirsten MacDonald	同上	技術補助員	H24.1～H24.6

研究項目

- ・ 多言語化
- ・ 音声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ー分析
- ・ ユーザーが作成した音声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

④ 山岸グループ

研究参加者

氏名	所属	役職	参加時期
山岸 順一	国立情報学研究所	准教授	H27.4～H29.3
高木 信二	同上	研究員	H27.4～H29.3
Fernando Villavicencio	同上	研究員	H27.4～H28.3
高木 かおり	同上	研究員	H27.4～H28.3
Xin Wang	同上	博士後期課程学生	H27.10～H29.3
Gustav Eve Henter	同上	研究員	H28.4～H29.3

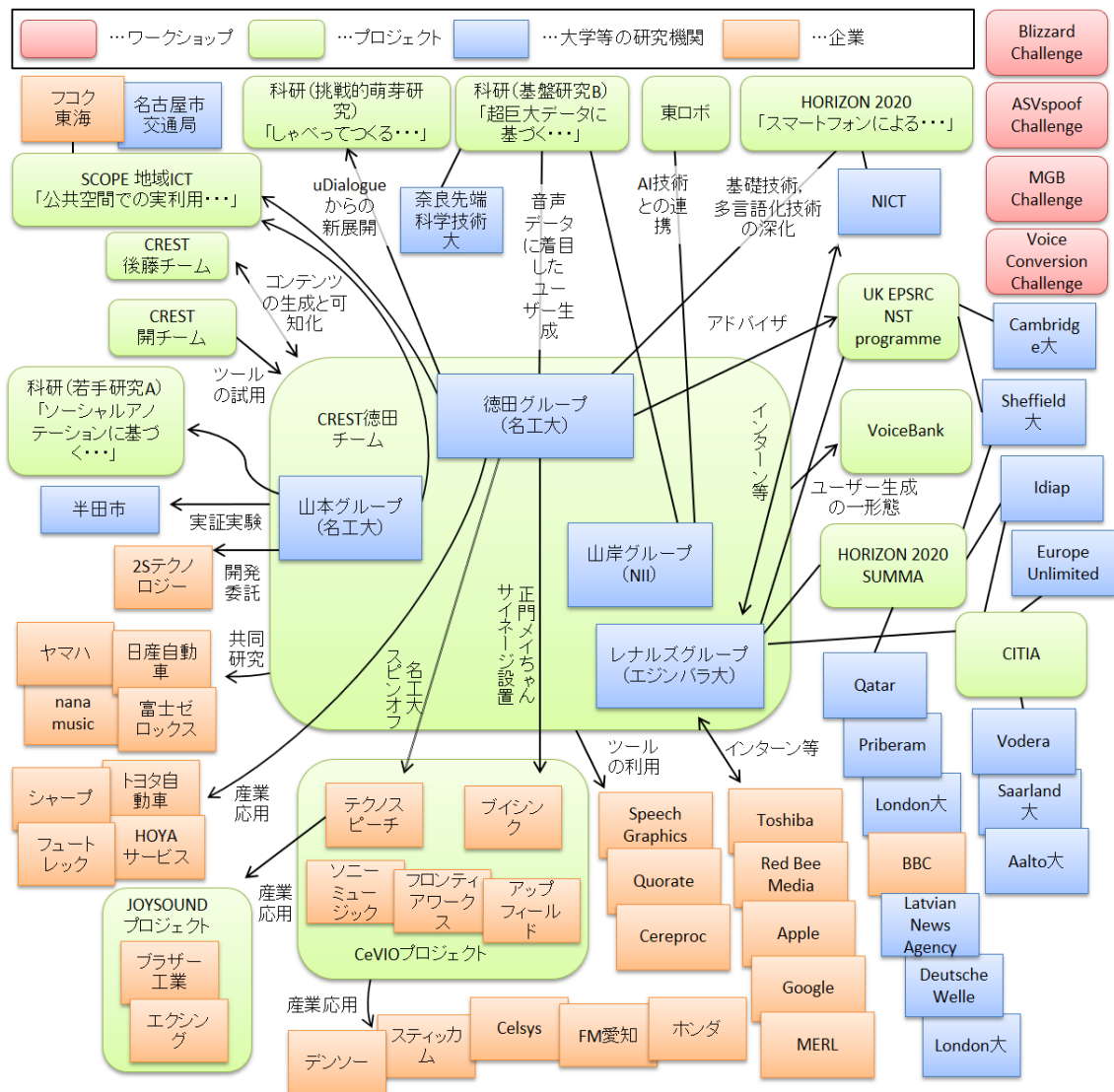
研究項目

- ・ 音声合成とユーザのフィードバックに関する基礎研究

(2) 国内外の研究者や産業界等との連携によるネットワーク形成の状況について

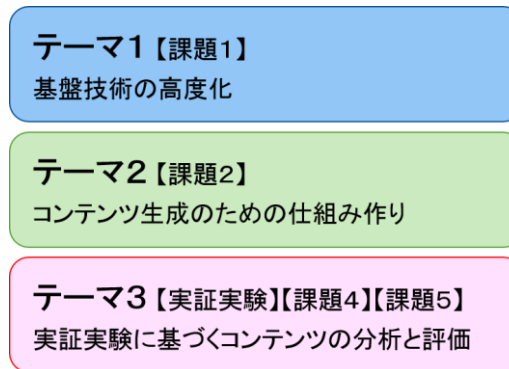
本プロジェクトにおいて、研究者の相互訪問などにより、名古屋工業大学、エジンバラ大学、国立情報学研究所は非常に強力な連携を実現しているが、プロジェクト内のコミュニケーションだけではなく、国内外の様々な関連プロジェクト、企業とも連携している。図 2 に連携状況を示す。図の上側の通り、本プロジェクトから派生した科研基盤研究 B、科研挑戦的萌芽研究、科研若手研究 A 等の国内のプロジェクトはもちろんのこと、欧州連合やイギリス等の海外プロジェクトにも参加し、奈良先端技術大学院大学、NICT、Sheffield 大学、Cambridge 大学、Saarland 大学、Aalto 大学、Idiap 等の、当該分野を代表する大学、研究機関と連携しながらプロジェクトを進めている。また、本プロジェクトのスピノフとして、Blizzard Challenge、ASVspoof Challenge、MGB Challenge などの国際ワークショップを主催している。図の下側の通り、Google、Apple、株式会社東芝等の大企業ともインターンシップ等によって連携している。産業応用については、学内ベンチャーである株式会社テクノスピーチを橋渡しとして、CeVIO プロジェクトと連携して強力に進めており、本田技研工業株式会社の車内システム、デンソー株式会社の車載エージェント、ブラザー工業株式会社のカラオケ機器などの実用化が始まっている。ヤマハ株式会社、日産自動車株式会社、富士ゼロックス株式会社、株式会社 nana music、名古屋市交通局等とは共同研究などの形で連携しており、本プロジェクトの成果の一部が実用化に繋がった例としては、トヨタ自動車株式会

社や HOYA サービス株式会社, シャープ株式会社等の製品を上げることができる。



§ 3. 研究実施内容及び成果

コンテンツ生成の循環系を実現するために、我々は様々な課題に取り組んできた。プロジェクト終盤に向けてターゲットを絞り込み、課題および実証実験を図 3 に示す 3 つの研究テーマに分類し、テーマごとの重点課題に取り組んだ。また、音声インタラクションシステム構築ツールキット MMDAgent を用いたあらゆる創作・応用に関する情報共有の場として「MMDAgent DAY!」と題する MMDAgent ユーザーミーティングを開催し、ユーザーとの意見交換やユーザー同士の交流の場を設けることで成果の普及と関連分野の発展に貢献した。さらに、プロジェクトの最終成果として、ソフトウェア、マニュアル、音声対話システムの設計指針等を総合的・統合的にまとめたオールインワン・パッケージ「エンサイクロペディア MMDAgent」を作成した。以下で、3 つの研究テーマと MMDAgent DAY!, エンサイクロペディア MMDAgent の詳細について述べる。なお、本プロジェクトでは、全グループが緊密な連携の下、人的な交流も行いながら進めたため、各テーマを全グループが融合的に取り組む形となっている。このため、以下では、グループ毎ではなく、テーマ毎に記述する形をとった。



※【課題3:多言語化】は全テーマに関係

図 2 3 つの研究テーマ

3. 1 【テーマ1 基盤技術の高度化】(主として名古屋工業大学大学 徳田グループ)

魅力的な音声対話システムを実現するためには、音声対話システムがユーザーにとって魅力的であり得る技術基盤を持たなければならない。そこで、本研究項目では、音声認識や音声合成などの音声対話システムにおける基盤技術の高度化、音声対話システムを実現するための基盤ソフトウェアの高度化、音声対話システムを構築するためのコーパスとエージェントの設計に取り組んだ。

3. 1. 1 音声インタラクションシステム構築ツールキット MMDAgent

音声インターフェースにおける魅力等の諸要素を統合的に研究するためには、音声での情報のやりとりだけでなく、表情やしぐさ、声質、間合いなどの人間的かつ高度な処理が必要とされる領域までシステムが踏み込む必要がある。そのためには、音声処理系のみならず画像処理やエージェント表現部と密統合したプラットフォームが必要である。

これまでに我々は、HMM 音声合成ツールキット HMM-based Speech Synthesis System (HTS), HTS をベースに構築された日本語テキスト音声合成システム Open JTalk, 音声認識エンジン Julius 等、音声技術に関連したオープンソースの研究基盤ソフトウェアを継続的に開発・公開してきた。これらのソフトウェア群をベースに、物理演算を含む高度な 3D・CG によるエージェント表現をリアルタイムに描画可能なグラフィック表示部、および、有限状態トランスデューサ (Finite State Transducer; FST) に基づくインタラクション記述部を新たに実装し、密結合することで、「音声インタラクションシステム構築ツールキット MMDAgent」を構築し、オープンソースのフ

リーソフトウェアとして公開した。

MMDAgent は、音声技術の専門家だけではなく、一般の人々により広く音声技術を利用したもののづくりを楽しんでもらうこと、また更には、魅力あふれる音声対話コンテンツを次々と作成してもらうことを強く意識した設計方針が採られており、これが MMDAgent の大きな特徴となっている。音声認識や音声合成に用いる辞書やモデルだけでなく、エージェントの 3D モデルやそれを動かすモーションデータ、FST の定義にもオープンフォーマットを採用しており、あらゆる構成要素をユーザーあるいは開発者が自由に作成、編集し、入れ換えることが可能である(図 4)。ユーザーが音声対話コンテンツを作成するという新しい考え方は、ソフトバンク株式会社の家庭用会話ロボット「Pepper」などでも採用されており、MMDAgent は新しい音声対話システムの枠組みの先駆けとなったと考えられる。



図 3 音声インタラクションシステム構築ツールキット MMDAgent

MMDAgent は、高速・高精度、かつ CG 表示を伴った表情豊かな音声対話を実現可能であり、本プロジェクトの技術基盤として継続的に開発してきた。開発当初は Windows のみの対応となっていたが、現在では Windows に加え Mac OS, Linux, Android OS, iOS といったモバイルデバイスを含むマルチプラットフォームに対応した(図 5)。音声から画像まですべてのモジュールがスマートフォンやタブレット上で動作しており、応答遅延の少ない音声対話システムをスマートフォンやタブレットにおいて利用可能である。なお、スマートフォンやタブレット単体で動作可能な 3D エージェント付き音声対話システムを実現するオープンソースソフトウェアは世界初の成果と考えられる。また、対話記述言語である有限状態トランスデューサ (FST) を拡張し、正規表現によるマッチングや変数の取扱いを可能とすることで、複雑な対話をより単純な FST スクリプトで記述することが可能になった。また、文字コードの UTF-8 への統一やコンテンツパッケージの導入により、多様な音声対話コンテンツをあらゆる場面で利用するユニバーサルな環境を実現した。

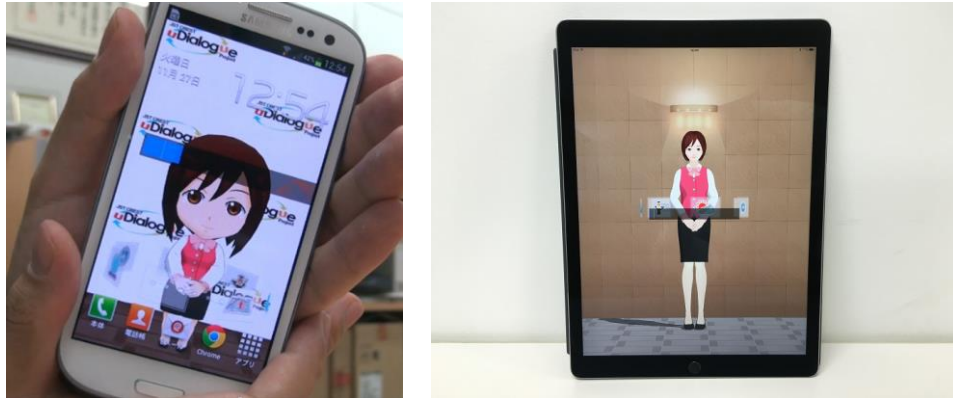


図 4 Android 版 MMDAgent(左), iOS 版 MMDAgent(右)

3. 1. 2 基盤技術

音声合成, 音声認識等の音声情報処理技術に関する多くの基礎研究をレナルズグループ, 山岸グループと連携しながら実施し, 基盤技術の高度化に取り組んだ。ここで得られた研究成果については学会等で積極的に発表した。以下で代表的なものを述べる。

- HMM 音声合成のための特徴抽出とモデリングの統合手法
近年の音声合成では, スペクトルなどの音声の特徴をモデル化する音響モデルとして, 隠れマルコフモデル (hidden Markov model; HMM) が広く用いられる。本研究では, これまで独立であった音声の特徴抽出と HMM を統合し, 音声波形を直接モデル化する手法を提案した。システム全体が一貫した基準により最適化されるため, 合成音声の品質が改善された。
- 加算構造によるスペクトラムモデルの改良
音声特徴と言語特徴の複雑な関係性をモデル化するために, 音声特徴(スペクトラム)に加算的な構造があると仮定した新しいモデルを提案した。また, 提案モデルを実現するための効率的な学習アルゴリズムを提案した。提案モデルによって合成音声の品質を大幅に改善することを示した。
- 対話調の音声合成手法
人と人との音声対話には, 言い淀みやフィラー(会話の隙間を埋める「あー」等の発話)が存在する。そこで, 音声対話における言い淀み等の影響について検証し, 合成音声にこれらを自動挿入する手法を検討した。提案法により, 言い淀み等の挿入箇所を高精度に予測することが可能となり, より自然な対話調の音声合成が可能になった。
- リカレントニューラルネットワーク言語モデルを用いた音声認識の高精度化
音声認識において, 言語的な特徴をモデル化する言語モデルに, リカレントニューラルネットワークを用いる研究が進められている。本研究では, 音声認識に適したニューラルネットワークの学習法を提案し, 音声認識の精度を改善した。さらに, リカレントニューラルネットワークの入力に韻律などの音響的な特徴を付加することにより, 言語モデルの性能を改善する手法を提案した。
- ディープニューラルネットワークに基づく音声合成の改善
近年, 音声認識だけでなく, 音声合成においてもニューラルネットワークの導入による合成音声の品質改善が報告されている。本研究では, 音声合成に適したニューラルネットワークの学習手法の検討や, 複数のニューラルネットワークを統合する手法, 音響特徴量と音素継続長を同時に学習する手法を提案し, 合成音声の品質を改善した。
- 発音情報が未知の言語におけるテキスト音声合成システムの構築法
様々な言語における音声合成システムの構築を考えた場合, 言語体系や発音情報(音素などの読みの情報)が明確ではない言語が数多く存在する。本研究では発話情報が未知の言語において, 大量の音声波形とテキスト情報のセットから自動的に発話情報を推定し, 音声

合成システムを構築する手法を提案した。これにより、より多くの言語において容易に音声合成システムを構築することが可能となった。

- 多様な声質を表現可能な因子分析に基づく音声合成の改善
これまで、話者性や感情などの声質を自由にコントロールできる音声合成を実現するために、因子分析を導入したモデル化手法を提案してきた。本研究では、この提案手法においてモデル学習やモデル構造選択の改良に取り組み、声質の再現性や合成音声の自然性を改善することができた。
- トピックモデルを用いた音声対話コンテンツの分析
ユーザー生成型の音声対話コンテンツにおける効率的なコンテンツ生成・管理のためには、コンテンツの分類、類似コンテンツの検出、人気コンテンツの推薦などの実現が必要となる。しかし、音声対話コンテンツは多種多様であり、規則に基づくタグ付け等は困難である。本研究では、言語モデルの一種であるトピックモデルを応用することで、音声対話コンテンツの統計的に自動分類する手法を提案した。

この他にも、多様な感情を表す音声合成や音響モデルの適応手法等、音声対話システム構築に必須となる様々な基盤技術の提案および改善に取り組んだ。

3. 1. 3 基盤ソフトウェア

基盤技術の高度化において得られた成果を研究基盤ソフトウェアとしてまとめ、オープンソースソフトウェアとして公開した。以下に、我々が開発・公開しているオープンソースソフトウェアを示す(図 6)。

- 音声対話システム構築ツールキット MMDAgent (<http://www.mmdagent.jp/>)
(CREST 期間中累計 77,000 ダウンロード)
- 音声認識エンジン Julius (<http://julius.sourceforge.jp/>)
(CREST 期間中累計 246,000 ダウンロード)
- HMM 音声合成ツールキット HTS (<http://hts.sp.nitech.ac.jp/>)
(CREST 期間中累計 426,000 ダウンロード)
- HMM 音声合成エンジン hts_engine API (<http://hts-engine.sourceforge.net/>)
(CREST 期間中累計 54,000 ダウンロード)
- 日本語音声合成システム Open JTalk (<http://open-jtalk.sourceforge.net/>)
(CREST 期間中累計 60,000 ダウンロード)
- 音声信号処理ツールキット SPTK (<http://sp-tk.sourceforge.net/>)
(CREST 期間中累計 56,000 ダウンロード)
- 日本語歌声合成システム Sinsy (<http://sinsy.sourceforge.net/>)
(CREST 期間中累計 3,400 ダウンロード)

ダウンロード数は 2011 年 10 月 (CREST 開始) から 2017 年 3 月までの累計である。



図 5 開発・公開中のオープンソースソフトウェアおよび
CREST 期間中の累計ダウンロード数

これらのオープンソースソフトウェアの開発は継続的かつ力強く行っており、新バージョンを順次公開してきた。また、これらのソフトウェアは最先端の技術を含む研究基盤ソフトウェアであり、ダウンロード数が示すように、既にデファクトスタンダードの一つとしての地位を確立している。実際に、図 7 に例示するように、学术论文やソフトウェア開発、イベント等、様々な場面で広く利用されている。

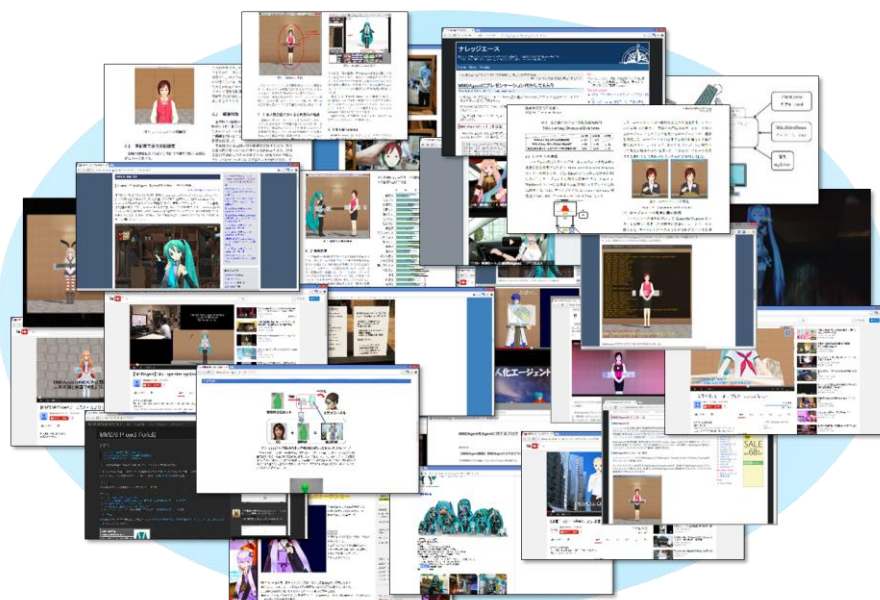


図 6 公開したオープンソースソフトウェアの利用事例

申請時は、産業応用はプロジェクト終了後に展開という想定であったが、音声対話システム構築ツールキット MMDAgent を始めとした音声情報処理基盤ソフトウェアを核として既に産業応用が始まりつつある。我々も大学発ベンチャーの株式会社テクノスピーチを通して CeVIO プロジェクトに参加するなどすることによりこのような展開に貢献してきた。

3. 1. 4 コーパス／エージェント

言語に依存しない普遍的な知見を得るためには、様々な言語による実証実験を平行して進める必要がある。この際には、言語的な違いだけでなく、文化的な違いに関しても十分考慮する必要がある。そこで、日本語とイギリス英語を対象として音声対話システムを開発し、日本語と英語の音声対話システムによる実験を行うことにより、言語や文化に依存する点、依存しない点について検証を行った。

まず、音声対話システムの開発に必要な日本語音声コーパスとイギリス英語音声コーパスを構築した。

- 日本語による男性声優コーパス

本コーパスは男性声優による日本語3000文を収録した音声コーパスである。収録文は新聞を含む複数のドメインから選択した。5種類の発話スタイルを含んでおり、平静を1800文、怒り、悲しみ、喜び、囁きの4つの発話スタイルをそれぞれ300文収録した。

- CSTR VCTK コーパス

本コーパスは、様々なイギリス英語アクセントをもつ話者109名による合計約60時間の大規模音声コーパスである。収録内容は各話者400文程度であり、新聞記事を含む複数のドメインから選択した。本コーパスは、イギリス英語研究用音声データベースとして現在標準的であるWSJCAM0(LDCより有償で配布)の1.5倍の規模であり、今後、標準音声データベースとして様々な分野において多様な用途に利用されると期待される。

- イギリス・エジンバラ英語による女性声優コーパス

エジンバラ訛りのイギリス英語を話すプロの女性声優による4600文を収録した。2種類の話速の音声は収録されており、収録音声のうち、800文は話速の早い音声、800文は話速の遅い音声である。また、100文に及ぶ自発的な会話を収録している。さらに、通常発話、早口発話などの4つの異なるスタイルの同一文を400文収録した。

次に、音声対話に適した対話エージェントを作成した。ユーザーに受け入れられる音声対話システムにするためには、文化的な違いを考慮したシステムを作る必要がある。日本ではアニメティックな3Dエージェントに抵抗が少ないと考えられたため、2.5頭身のエージェントや、男性のエージェントを作成した(図8)。一方、様々な議論の結果から、ヨーロッパ圏においてはリアリスティックな3Dエージェントが、より受け入れられやすいと考え、イギリス英語版の音声対話システム用の対話エージェントを作成した。

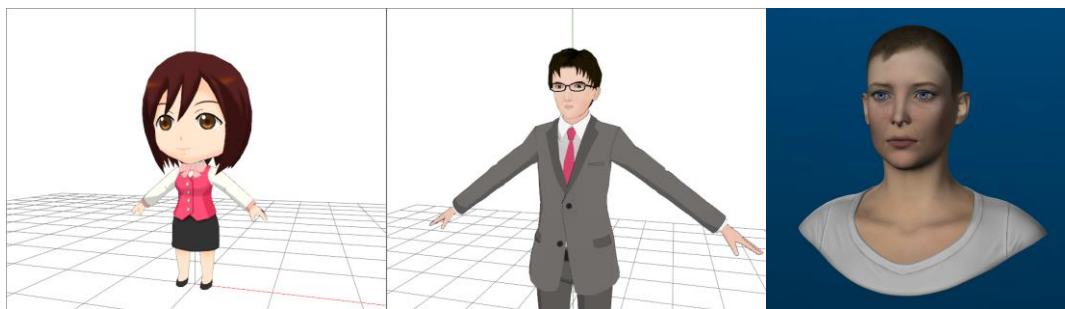


図 7 アニメティックな音声対話エージェント(左, 中),
リアリスティックな音声対話エージェント(右)

3. 2 【テーマ2 コンテンツ生成のための仕組み作り】(主として名古屋工業大学 山本グループ)

音声対話コンテンツは、図 9 に示すように、3 つの階層から構成される。マテリアル層では、音声モデル・言語モデルなどといった専門的なモデルデータや、画像・音楽・3Dモデル・モーシオンなどのバイナリファイルが含まれる。アクション層では、一問一答形式の簡単な挨拶や、天気予報のパネルの表示など、一続きの短い動作が含まれる。シナリオ層では、アクション層の動作を組み合わせることによって、より複雑な音声対話シナリオを実現する。シナリオ層やアクション層は FST 形式で記述され、マテリアル層はそれぞれの素材の独自の形式で格納されている。

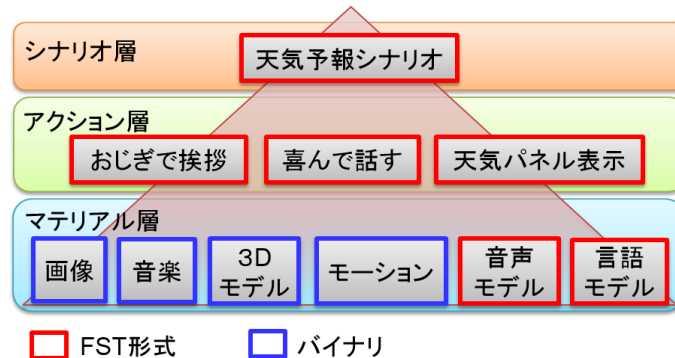


図 8 音声対話コンテンツの階層

本研究項目では、音声対話コンテンツにユーザー生成の概念を取り入れ、音声対話コンテンツを作成および共有するための仕組みの構築に取り組んだ。また、図 10 が示すように、音声対話システムは単体で閉じているのではなく、複数のシステムが互いに連携しあうことも想定し、音声対話システムの連携に関する研究にも取り組んだ。以下、3. 2. 1～3. 2. 3でこれらについて述べる。

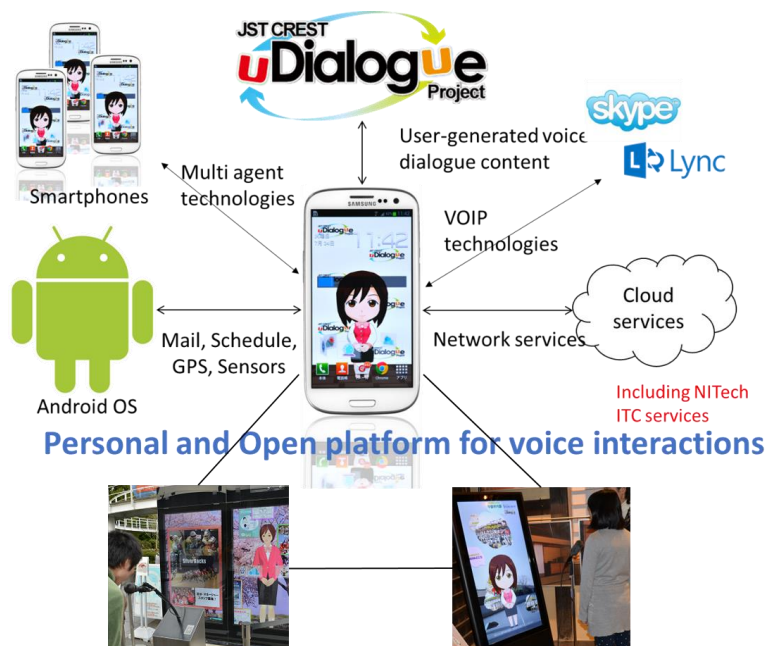


図 9 音声対話システムとのネットワーク連携

3.2.1 コンテンツ制作支援ツールの構築

音声インタラクションシステム構築ツールキット **MMDAgent** では、**FST** 形式のスクリプトファイルを記述することによって、音声対話シナリオの制御を行っている。しかしながら、**FST** 形式のスクリプトファイルを一般のユーザーが直接記述することは簡単ではない。そこで、音声対話コンテンツを手軽に編集できる仕組みとして対話スクリプト作成ツールおよび音声インタラクションビルダの構築に取り組んだ。

対話スクリプト作成ツール

ユーザー生成型の音声対話コンテンツを実現するためには、ユーザー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、我々は、初心者向けのタッチインターフェースを用いた音声対話コンテンツ作成ツールEFDE (Extended FST based Dialogue Editor) (図11(左))と、Web ブラウザを用いた上級者向けの音声対話コンテンツ作成ツールMMDAE (MMDAgent Editor) (図11(右))を開発した。

● 音声対話コンテンツ作成ツールEFDE

初心者を対象としたAndroid 端末向けのインターフェースであり、状態遷移図で表されたFST スクリプトをタッチインターフェースで編集可能である。編集結果をオンデマンドに実行可能にしたことにより、一般ユーザーが手軽に音声対話コンテンツの動作を確認しつつ編集が可能である。また、よく使われるFSTスクリプトをテンプレート化することにより、少ない状態数で複雑な音声対話を実現可能にした。さらに、音声による対話文やキーワードを入力可能な機能を追加することにより、音声とタッチインターフェースのみで音声対話コンテンツを容易に作成可能な仕組みである。

● 音声対話コンテンツ作成ツールMMDAE

上級者を対象としたインターフェースであり、Webブラウザを用いたFST レベルの詳細な編集が可能である。FSTスクリプトの入力補助と、FSTスクリプトの構造化を実現することにより、容易な編集が可能になった。また、Webブラウザを基盤として用いることで様々なプラットフォームで動作可能となり、ネットワークを介して、手軽にコンテンツを作成できる環境が整った。

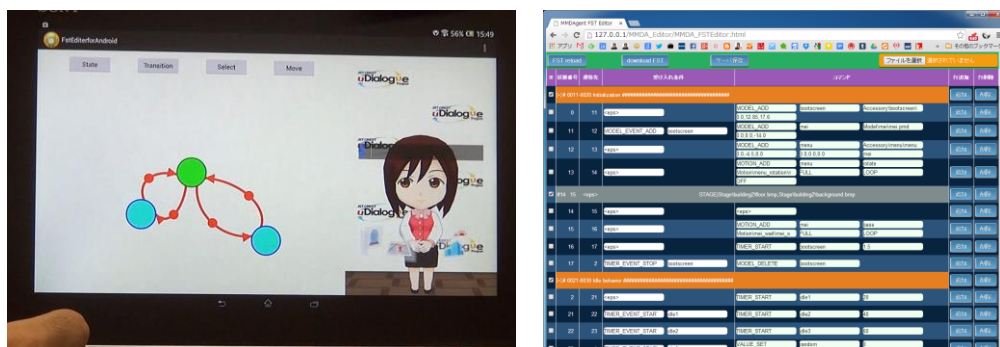


図 10 タブレットを用いた音声対話コンテンツ編集インターフェース EFDE(左), Web ブラウザを用いた音声対話コンテンツ編集インターフェース MMDAE(右)

音声インタラクションビルダ

ユーザーによる魅力的で優秀なコンテンツの創造を推進するためには、音声対話コンテンツを詳細な部分まで自分の思い通りに作り込めるようなクリエイター向けの環境も必要である。そこで、インタラクティブな音声対話コンテンツを発話のタイミングや詳細な部分まで作りこみ、動作検証を行うことが可能な開発環境として音声インタラクションビルダを試作した(図12)。この音声インタラクションビルダは以下の3つのコンポーネントから構成されている。

1. 状態遷移図の3次元空間における可視化・ブラウジングによるFSTスクリプトの構造把握機能
2. イベント入力シミュレーションによる動作確認機能
3. 入出力イベント系列の保全と再現による時系列インタラクションの検証機能

主観実験の結果、既存の環境に比べて、複数のモーションをイベントに従って非同期に組み合わせるような複雑なインタラクションを含む高度なコンテンツが作成しやすいことが示された。

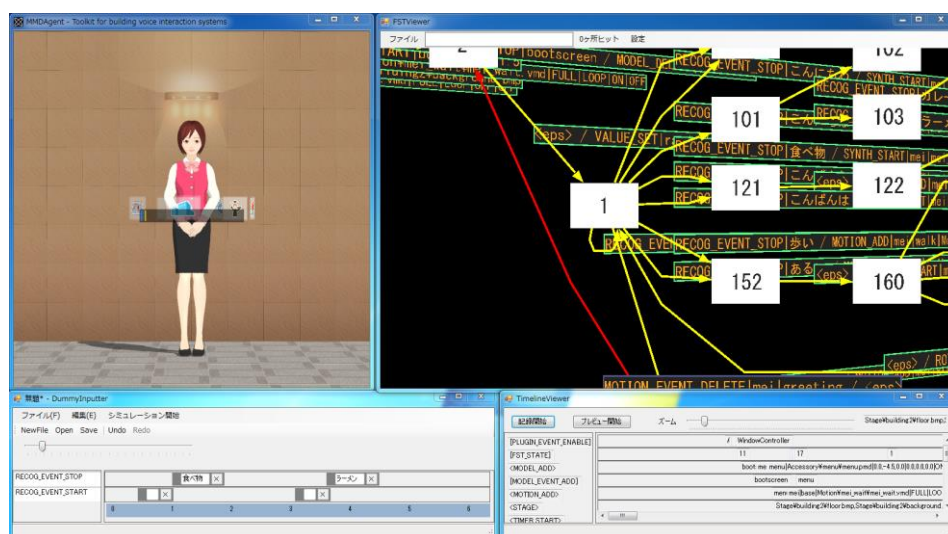


図 11 音声インタラクションビルダ

3. 2. 2 クラウドでのコンテンツ共同編集環境の構築

ユーザー生成型メディアの特徴の一つは、他者とコラボレーションして作品を構築することが頻繁に行われることである。音声対話コンテンツにおいても、ネットワークを介して音声対話コンテンツを共同作成・編集する環境を構築することで、他者と協力して大きく複雑なコンテンツを作り上げる、既存のコンテンツを流用・拡張するといったことが容易となる。ここでは、クラウド環境を利用した音声対話コンテンツの共同編集環境の構築に取り組んだ。

履歴付き音声対話コンテンツの共同編集システムの構築

音声対話コンテンツをユーザー間で共同構築する環境・システムを考えると、例えば人工無脳に代表されるように、質問とそれに対する応答の組をユーザーが登録していく一問一答型のシステムが最も簡便な形態の一つである。この形式は誰でも気軽に構築に参加でき、様々なトピックの対話を広く集めることに適しているが、対話の流れや履歴を考慮しない一度切りのやりとりしか扱えないため、連続した会話やシナリオを持ったやりとりのような、対話本来の楽しさや魅力を提供するシステムとしては明らかに不十分である。一方、長い会話やシナリオを破たん無く設計するにはシステム発話に対するユーザーの反応を予測する必要がある、一般ユーザーが気軽に構築することは難しい。

そこで、履歴を持つ会話をユーザー間で容易に共同構築可能なユーザー生成型音声対話システムを開発した(図13)。対話の登録単位はキーワードと応答の組であるが、親と子を設定できるようにすることで、複数回のやり取りを連続した一問一答としてユーザーが協同構築できるようにした。登録された音声対話コンテンツ全体は、キーワードと応答のペアを単位とする木構造となる。登録画面では、誰でも対話の流れを直感的に把握できて編集できるよう、SNSのチャット風のWebインターフェースを採用した。このシステムにより、ユーザーが共同して連続した長い対話を作り込むことが可能となり、さらに他者が登録した対話に別の対話分岐や後続の対話を追加することができるようになった。主観評価実験より、従来法よりも大幅に高い面白さと期待感をユーザーに与えることが示され、対話本来の魅力をもつシステムをユーザー生成的に構築することの可能性が示された。

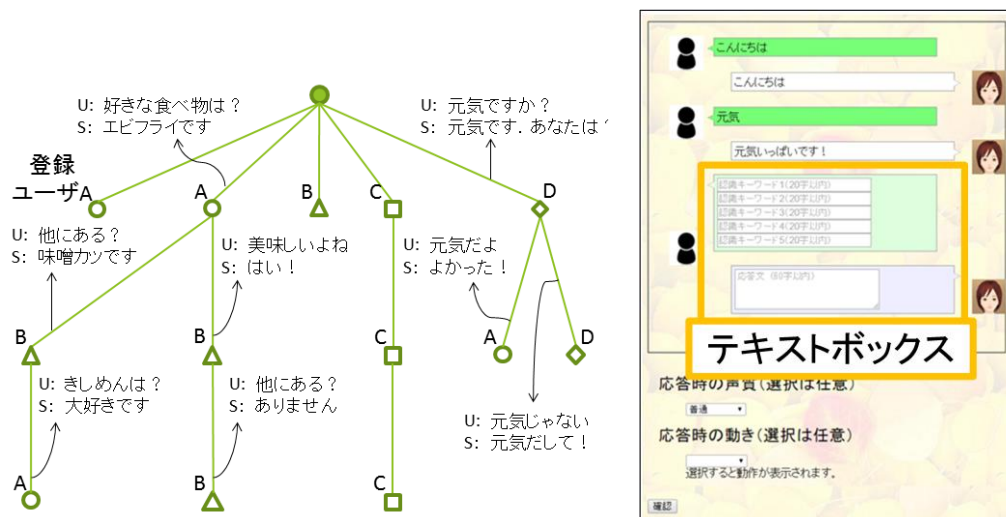


図 12 履歴付き音声対話コンテンツの概念図(左)，
履歴付き音声対話の登録インターフェース(右)

クラウド方式の音声対話コンテンツの編集環境の構築

一般にMMDAgentにおける対話のボキャブラリーを増やすためには多数の対話シナリオを記述することが重要である。そこで、クラウドソーシングの概念に基づき、Webを介して対話シナリオ作成タスクを受発注する、音声対話シナリオ作成に特化したクラウドソーシングシステムを開発した(図14)。このシステムにより複数人の分担作業によるシナリオ作成が可能となった。また、対話シナリオ作成タスクを受注したユーザーが手軽にシナリオ作成を行えるよう、前述のMMDAE とSkype版エージェント動作確認ツールの連携機能を強化した。開発したシステムの評価実験を通して、クラウドソーシングに基づく対話シナリオ作成の有効性を確認した。また、実際に本システムを用いて対話シナリオを多数作成し、それらを基本音声対話ライブラリとしてまとめた。

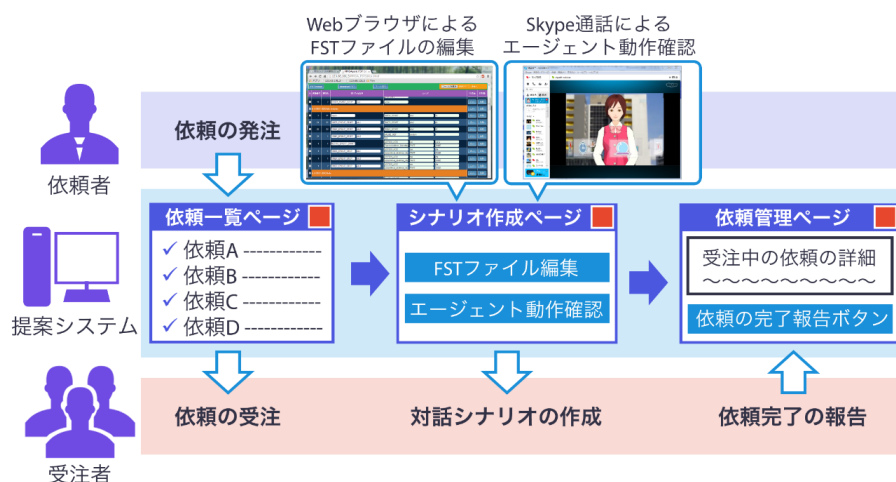


図 13 対話シナリオ作成に特化したクラウドソーシングシステム

3. 2. 3 モジュール化およびエージェント間連携基盤の構築

ユーザーが作成した音声対話コンテンツを、サーバーを介してネットワーク上で配信・共有するための仕組みとして、音声対話コンテンツのパッケージ化に取り組んだ。また、複数の音声対話システムが協調動作するための環境の構築に取り組んだ。

音声対話コンテンツのパッケージ化

一般に音声対話コンテンツはユーザーや技術者が個々に作成しているのみで、それらを効果的に流通させるための手段はあまり考慮されていない。そこで、ユーザー生成の概念に基づき、ユーザーが作成した音声対話コンテンツを、サーバーを介してネットワーク上で配信・共有するための仕組みについて提案した。まず、FSTスクリプトを拡張し、複数並列化に基づくパッケージ化手法について検討した。これにより、FSTスクリプトの保守性が向上すると同時に、従来困難であったFSTスクリプトの部分更新や機能単位の配信が可能になった。さらに、図15に示す、音声対話コンテンツのパッケージ化に基づく配信の仕組みを試作することによって、パッケージ単位でのコンテンツ循環の仕組みについて検討した。

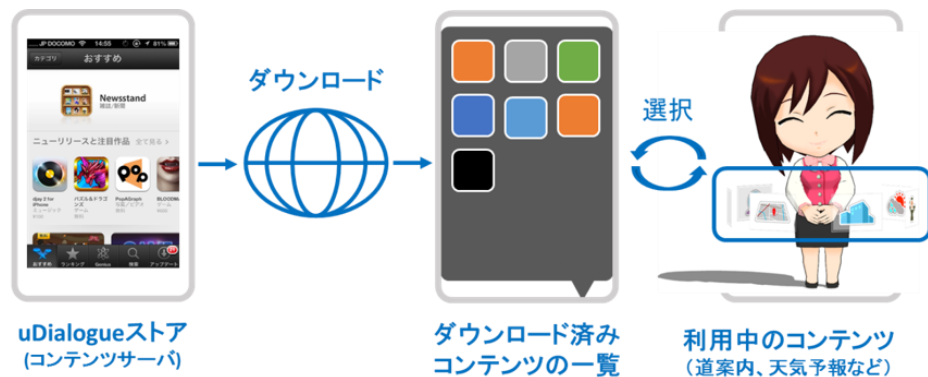


図 14 音声対話コンテンツのパッケージ化

また、対話スクリプトだけでなく単語辞書やボイス、3Dモデルやモーションといった音声対話コンテンツ全体の各構成要素をユーザーが自由に組み替えて選択・構築できる環境の実現を目指し、ユーザー生成型音声対話システムのモジュールアーキテクチャの提案および構築を行った(図16)。音声認識や音声合成・対話管理・エージェントモデル等の全てのコンテンツを汎用モジュールとして統一的に扱い、パッケージとして動作の整合性やモジュール間の依存関係を取り扱うことが可能である。実際にMMDAgentに対してモジュールをパッケージで管理する仕組みを提案し、対話スクリプト内のモデル名指定のようなモジュール名指定に対するスコープの設定や、パッケージ間の依存性や競合をヘッダ記述により自動的に検出する仕組みを実装した。

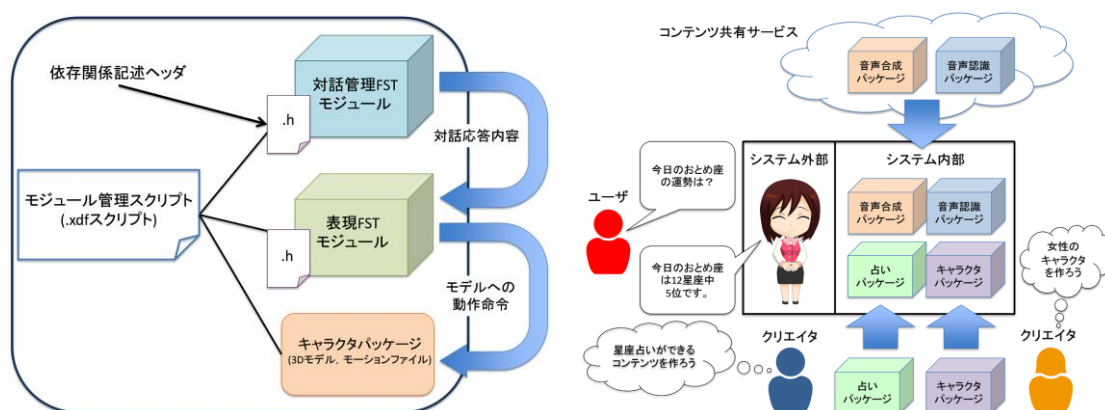


図 15 音声対話モジュールのパッケージ化(左)、
モジュールベースの音声対話コンテンツ共有(右)

ネットワークエージェント技術を用いた音声対話システム間の連携基盤

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、サーバー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、下記のシステム連携基盤を導入した。

- 複数のスマートフォン端末上のMMDAgent が協調動作するための環境を構築した。具体的には、エージェント/NFC/Bluetooth技術に基づく対話システムのネットワーク接続機構を開発した(図17)。スケジュール調整システムを試作し、複数対話システム連携型サービスの有効性を確認した。
- 複数のサイネージ上のMMDAgentが協調動作するための環境を構築した。この環境は、多数のサイネージの接続を考慮し規模拡張性にも優れた設計となっている。対話回数共有システムを試作し、サイネージ間システム連携の有効性を確認した。
- サイネージ上のMMDAgentとスマートフォン上のMMDAgentが協調動作するための環境を構築した。スマートフォンを用いた音声対話により、サイネージ上に存在する「案内地図」等のコンテンツをスマートフォンに取り込むことで、サイネージを離れても案内コンテンツの参照が可能となった。これにより、道案内などのコンテキスト依存型の音声対話サービスの実現が容易となった。

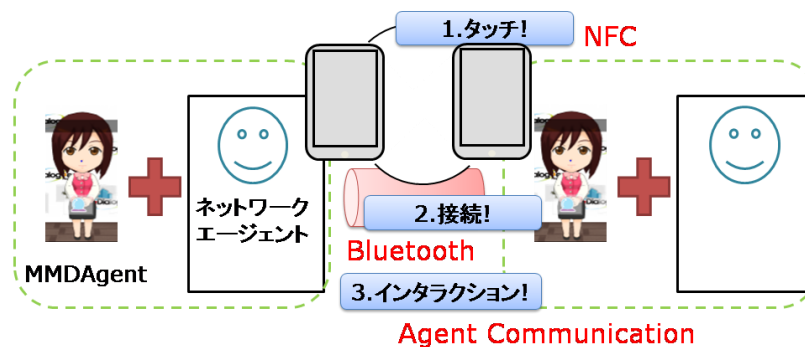


図 16 P2P 型ネットワーク接続機構

3.3 【テーマ3 実証実験に基づくコンテンツの分析と評価】(主としてエジンバラ大学 レナルズグループおよび国立情報学研究所 山岸グループ)

ここでは、コンテンツ利用を促進するための仕組みづくりに取り組んだ。さらに、提案システムを用いた実証実験を行い、提案システムの社会実装の可能性を検証した。

3.3.1 コンテンツ利用を促進する仕組みづくり

ユーザー生成型の音声対話コンテンツでは、ユーザーによる多種多様で多量の音声対話コンテンツが作成されることが期待できる一方で、それらのコンテンツ群を適切に分類・関連付けすることができなければ、それらの利用価値を最大限に高めることは困難である。ここでは、コンテンツの推薦や検索等の技術を実現し、音声対話システムの利用支援や作成支援等の実現するための音声対話コンテンツ群の分析・分類・関連付け手法の研究成果について述べる。

音声対話システムの利用履歴に基づく関連ワード推薦

音声対話システムの利用に不慣れな初心者に対する対話システムの利用促進を目指して、ユーザーに次に話すべきキーワードの候補群を提示する仕組みを検討した。そして、サーバーに蓄積されたユーザーの対話履歴から、情報推薦技術等を用いて分析し、関連ワードを獲得する手法を実現した(図18(左))。被験者実験より、関連ワードを提示することにより、ユーザーの音声対話システムの円滑な利用を支援することを示唆する結果を得た。また、関連ワードの推薦手法として、履歴ベース推薦、内容ベース推薦、ランキング推薦、ランダム推薦の4つの方式、および、それらを統合した複合推薦の方式を比較検討した。被験者実験により、複合推薦方式が内容ベース推薦やランダム推薦単体よりもよい総合結果が得られることを確認した。

利用履歴のユーザー間共有による相互インセンティブの強化

一般に、ユーザー生成型メディアにおいては、ユーザー間で活動の記録がシェアされることでユーザーが互いに刺激を受け、システムの利用やコンテンツの作成が誘引されるという特徴がある。音声対話コンテンツにおいても、あるユーザーの対話システムとのやりとりやコンテンツの利用履歴を別のユーザーが共有することで、ユーザー同士で相互に利用意欲を高めることが期待される。このため、一問一答型のユーザー参加型音声対話システムにおいて、以下の2つの機能を実装した。

1. ユーザーが何を発話しているかのオンライン逐次共有
2. 対話の再生回数やキーワード発話回数を記録しランキング表示

上記の2つの機能を実装することによって、ユーザーが相互にインセンティブを強化する仕組みを構築した(図18(右))。音声対話コンテンツの利用者と作成者にこれらを提示することにより、利用者が必要としているキーワードを作成者が知ることが可能になった。また、コンテンツ作成者が、ランキング結果に基づきユーザーが必要としている対話を手軽に登録する仕組みも実現した。被験者実験により、これらの仕組みが利用者とコンテンツ作成者の両方のモチベーション(インセンティブ)の向上に繋がる可能性を確認した。



図 17 関連ワード提示手法(左),
発話履歴の共有と音声対話コンテンツのランキング提示(右)

3. 3. 2 公共空間での実証実験

提案システムを名古屋工業大学正門前や半田市観光協会など、様々な場所で設置・運用することにより、提案システムの社会実装の可能性を検証した。

学内での実証実験

我々が開発した音声インタラクションシステム構築ツールキットMMDAgentを用いて構築した「全天候型双方向音声案内デジタルサイネージ(メイちゃん)」を、大学正門入ってすぐのオープンスペースに設置した(図19)。複数のカメラを用いた顔画像認識技術によるキャラクターの視線制御や、焦電センサーを用いた人体検知によるキャラクター側からの能動的な呼びかけ等の機能が実現されている。表示テキストや画像、発話テキストに加えて、案内時のキャラクターのモーションや発話スタイルまでも統合したコンテンツを、サーバーから動的に更新することが可能であり、情報基盤センターのデータベースに登録された学内イベント情報の中から、季節や時間、内容に応じて適切に取捨選択された情報がタイムリーに案内される(図20)。想定されるユーザーは、来学者、学生・教職員である。



図 18 双方向音声案内デジタルサイネージ「メイちゃん」(左),
名古屋工業大学正門前(右)



図 19 情報基盤システムとの連携

2011年4月6日にシステムの稼働を開始し、2011年6月15日にコンテンツ登録システムを学内に公開した。2011年11月15日までに100件以上のコンテンツが登録されており、1日当たりのユーザー発話も、休日を含めた平均も約35 発話を超えている。学生や教職員から243件(内学生43件)の音声対話コンテンツの投稿があった。図21に投稿されたパネル画像の例を示す。

また、2014年9月からは、新たに名古屋工業大学のオープンスペース(夢ルーム)にも室内型双方向音声案内デジタルサイネージを設置した(図22)。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。設置から半年の間に学生などから計437件の音声対話コンテンツの投稿があった。



図 20 パネル画像

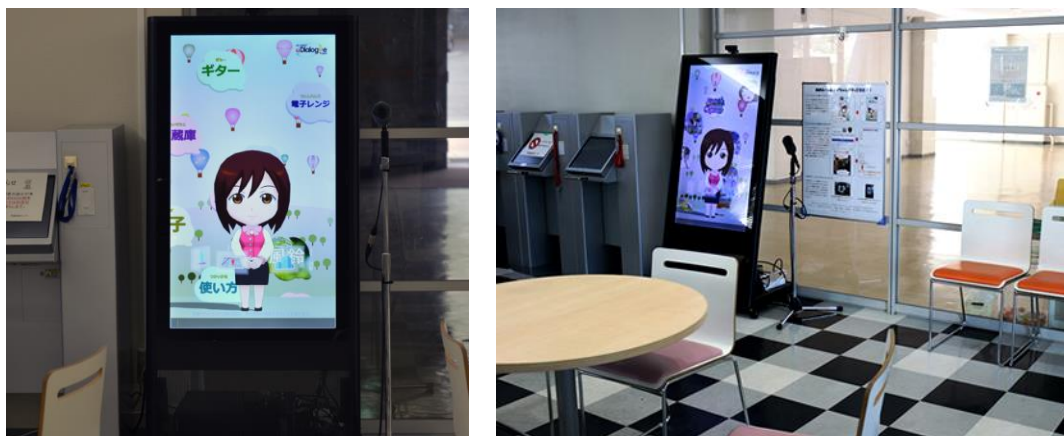


図 21 室内型双方向音声案内デジタルサイネージ(左),
名古屋工業大学オープンスペース(夢ルーム) (右)

また、「メイちゃん」について、学内での知名度や使用頻度などについて、Moodleのアンケートシステムを用いて調査を行った。この結果、262件の有効回答があり、その99%が「メイちゃん」を認知していた。また、対話経験も77%と高く、MMDAgentを利用した音声対話システムの注目度および実用性の高さを示している。自由記述アンケートへの回答率は33%であり、3分の1のユーザーが積極性をもって、音声対話システムの改良点などについての意見を寄せた。

学外での実証実験

2013年度には、新たに、室内型双方向音声案内デジタルサイネージを試作し、半田市観光協会蔵のまち案内所(国登録有形文化財小栗家住宅内)に設置した(図23)。小さいキャラクターを画面下部に置き、画面上部にパネル情報等を提示することによって、狭い室内でも快適に利用可能になるように工夫した。利用者は、観光客を中心とした一般の来訪者である。



図 22 半田市観光協会「蔵のまち案内所」に設置した音声対話システム(左),
半田市観光協会「蔵のまち案内所」の外観(右)

2014年度には、Webインタフェースを用いたより魅力的な音声対話を構築できる枠組みについて検討した。具体的には、従来からある認識キーワードとそれに対応する音声案内文だけでなく、声質やキャラクターの表情・モーションなども追加可能なWebインタフェースを試作した。サイエンスアゴラ2012では図24に示すデモシステムを実際に使用し、モーションや声質の組み合わせの変化が音声対話の魅力に大きく影響することを実感した。



図 23 サイエンスアゴラ

また、スマートフォンやWebブラウザを用いてユーザー(ここでは、観光案内所の職員)が簡単に音声対話コンテンツを追加・更新できる仕組みを実現することで、ユーザーの創意工夫で、現場に即した、より魅力的な音声観光案内が実現可能になった。実際に、観光協会の職員によって、多数の観光案内に関するコンテンツが登録された。なお、本実証実験システムは、テレビや新聞などで紹介される、半田市の公式広報誌の表紙に掲載されるなど、観光PRに繋がるものとして地域からも注目を集めた。さらに、同様のシステムを、半田市市役所新庁舎やNHK名古屋放送局(図25)、国立情報学研究所(図26)などに音声対話システムを設置した。NHK名古屋放送局ではTV番組やイベント等において実証実験を実施した。その際、複数のキャラクターを連携させて動作できるようにするなどの拡張を行った。国立情報学研究所では、国立情報学研究所のキャラクター「ビットくん」の3Dモデルを利用した音声対話システムを構築した。このキャラクターに合わせた音声データベースも構築し、キャラクター声による音声合成システムを構成した。



図 24 半田市市役所新庁舎(左), NHK 名古屋(右)



図 25 国立情報学研究所に設置した音声対話システム(情報犬ビット)

その他の実証実験

モバイル環境における音声対話システムとユーザー生成型音声対話コンテンツ作成環境に関する実証的な研究を実施した。

- Skypeのビデオ電話機能とMMDAgentを連携させることによって、iPhoneなどのスマートフォンを用いてCGキャラクターと音声による対話を可能にするシステムを開発した(図27(左))。第74 回情報処理学会全国大会において公開実験とアンケート調査を実施し、120名分のアンケート結果から、モバイル向け対話システムはデジタルサイネージ版より対話の応答速度は劣るものの、親しみやすさが向上していることが判明した。
- 中学生を対象として、音声対話に基づくモバイル学内見学支援システムの実証実験を2014年1月に行った(図27(右))。スマートフォン上の音声対話エージェントが、中学生に対して名古屋工業大学のキャンパスや設備に関して音声案内を行うシステムである。なお、実験で用いた音声対話コンテンツは、名古屋工業大学の案内担当の職員2名によって作成された。主観評価実験の結果、提案システムに基づく音声案内の有効性が示された。
- より一般の人に音声対話システムを利用してもらうためには、公衆電話網等の連携が有用である。そこで、MMDAgentとVoIPクライアントを連携させる仕組みを開発した。学内に整備・展開されたSkype for business (Lyncの新バージョン)に接続するためのミドルウェアを作成し接続した。同時に、名古屋工業大学ユニファイドコミュニケーションシステムの一部であるVoice over IPによる新しい内線電話システムとも接続するために、Software Phone連携用ミドルウェアを開発し、相互接続試験を実施した。これらの仕組みを発展・評価すると同時に、学内電話網との連携の仕組みを検討した。



図 26 Skype 版音声対話システム(左), モバイル学内見学支援システム(右)

3. 3. 3 ネットワーク環境での実証実験

より幅広い社会環境におけるユーザー生成型音声対話コンテンツの実現のための手がかりを得るために、対象を一般のネットワーク環境へ広げた実証実験を行った。また、その前段階として、音声対話コンテンツがユーザー生成型メディアとして実際に社会で形成され成長していくために必要な、ユーザーの動機および誘引要素(インセンティブ)に関して分析を行った。

まず、ユーザーの利用動機および誘引要素の分析については、ユーザーをコンテンツへの関わり度合いから以下の4種類の層に分類し、それぞれの相互関係と移動の関係を分析した(図28(左))。

- 潜在者: システムを全く利用したことがない人。
- 消費者: システムを利用する・したことがある人。受け身なユーザー。
- 参加者: システムを積極的に利用しコメントやお気に入り、評価を下す人。アクティブなユーザー。
- 生産者: コンテンツを作成・投稿する人。

そして、ネットワーク環境において音声対話コンテンツがユーザー生成型メディアとして成立するために音声対話システムが満たすべき要件を以下のように定めた。

- 潜在者から消費者へ: スマートデバイス等マルチデバイス対応による音声対話コンテンツ再生環境の拡張。
- 消費者から参加者へ: ユーザー動線設計やSNSでの定期的な情報提供によるシステム全体への興味の喚起。
- 参加者から生産者へ: Twitterのリプライ機能を用いた「お題」提供型対話コンテンツ生成(図28(右))。



図 27 ユーザー層の分類とインセンティブ(左), お題提供型対話コンテンツ生成(右)

上記の改善を施したマルチプラットフォームの一問一答型対話コンテンツ配信・登録システムを構築し、社会実験を実行した(図29)。音声対話コンテンツを保持するコンテンツサーバーとコンテンツ登録システムを学外に向けて公開し、マルチプラットフォーム(Windows/Linux/Android)でのシステムを提供し、ユーザーが自由にダウンロードして音声対話を試せる環境を構築した。実際に一般ユーザーを誘引するための動線設計やサポート体制を整え、1ヶ月間の公開実験を行った。Twitterのリツイートによって約30,000人への周知を行い、6,400回の発話と232個の音声対話コンテンツ新規登録を得た。利用者の55%はAndroidからの参加者であり、15%がTwitterのリプライ機能を用いた「お題」提供に参加しライトなコンテンツ登録を行うなど、提案法の有効性が確認でき、またユーザー動向に関する貴重な資料を収集することができた。



図 28 社会実験における一般ユーザー向けの説明ページおよび参加動線の設計

3. 3. 4 音声対話コンテンツ共有サービスの公開と運用

音声対話コンテンツの製作者と利用者の間に音声対話コンテンツを共有する枠組みが用意されおらず、作成された音声対話コンテンツは様々な場所に分散した状態になっているのが現状であった。そこで、音声対話コンテンツを発表・共有する場を提供することで、より多くの音声対話コンテンツを生成することに繋がると考えた。具体的には、音声対話コンテンツをユーザー間で共有可能にするための仕組みとして、音声対話コンテンツ共有サービスMMDAnget SHARを立ち上げた(図30)。

公開日：2016年10月1日(土)

URL: <http://share.udialogue.org/>



図 29 音声対話コンテンツ共有サービス MMDAgent SHARE (<http://share.udialogue.org/>)

一般に、音声対話コンテンツを容易に共有できるようにするためには様々な課題をしなければならない。提案システムでは以下のような対応を行った。

1. 音声対話コンテンツは多数のファイル群から構成されるなど扱いが煩雑であったため、独自の MMDA形式を定義し、1ファイルで利用可能にしている。
2. MMDA形式のファイルをサーバー側で容易に生成できる機能を作成している。
3. Windows用のMMDAgentインストーラを新たに開発し、MMDA形式のコンテンツを1クリックで実行可能にしている。
4. サービスの利用を容易にするためにGoogle などのオープンIDとの連携を可能にしている。
5. コンテンツの投稿時に自動的にスクリプトのエラーやワーニングを検出する仕組みを開発している。
6. 本サービスの利用規約やプライバシーポリシーを慎重に作成している。
7. コンテンツに含まれるすべての付属文章 (READMEなどを) を自動的に本サービスから確認できるようにする仕組みを開発するなど、著作権に最大限配慮している。

図31(左)に共有された音声対話コンテンツの例を示す。音声対話コンテンツを分離し、また、音声対話コンテンツを手軽にWebで作成するためのサービスMMDAgent EDITも併せて試作した。これは、ユーザーが一般的なWebブラウザを用いて、画像やキーワード、文章などを入力していくことで、手軽に音声対話コンテンツを作成するための仕組みである(図31(右))。



図 30 共有された音声対話コンテンツの例(左), 音声対話コンテンツ作成サービス MMDAgent EDIT (<http://share.udialogue.org/agora/>) (右)

3. 4 MMDAgent ユーザーミーティング「MMDAgent DAY!」の開催

音声インタラクションシステム構築ツールキットMMDAgentを用いたあらゆる創作・応用に関する情報共有の場として「MMDAgent DAY!」と題するMMDAgentユーザーミーティングを開催した(図32)。

日時:2016 年 10 月 1 日(土)

場所:名古屋工業大学 4 号館ホール・ホワイエ

URL: <http://udialogue.org/events/MMDAgentDAY/>

ユーザーとの意見交換をすると同時にユーザー同士の交流の場を設けることで成果の普及と関連分野の発展に貢献した。



図 31 MMDAgent DAY!

3. 5 エンサイクロペディア

これまでに開発してきた音声インタラクションシステム構築ツールキットMMDAgentを中心とした基盤ソフトウェア群や音声対話コンテンツ制作支援ツール群だけではなく、キャンパス内外における複数の実証実験、解説論文執筆、MMDAgent講習会、インターネット上のブログ形式によるTIPSの発信等を通して、MMDAgentおよび音声対話コンテンツ制作に関する指導書・手引書、スライド、リファレンス・マニュアル、サンプルスクリプト等が充実してきた。これらの成果物を統合し、さらに拡張・整備することにより、ソフトウェア群、マニュアル、基本音声対話ライブラリ、設計指針等を総合的・統合的にまとめたオールインワン・パッケージ「エンサイクロペディアMMDAgent」を構成した。パッケージは以下の項目を含むものとなる。

1. 音声インタラクションシステム構築ツールキットMMDAgent (マルチプラットフォーム対応: Windows, Linux, Mac OS, Android OS, iOS)
2. MMDAgent概説 (日英)
3. MMDAgentクリエイター (コンテンツ制作者) 向けリファレンス・マニュアル (日英)
4. MMDAgentディベロッパ (開発者) 向けリファレンス・マニュアル (日英)
5. MMDAgent講習用スライド (日英)
6. MMDAgent講習用動画 (MOOC, OCW)
7. 音声対話コンテンツ制作支援ツール群 (Web ブラウザ上で利用可能な音声対話コンテンツ編集ツール, タブレット上で利用可能な音声対話コンテンツ編集ツール, 音声インタラクションビルダ, 他)
8. 音声対話コンテンツライブラリ (基本対話ライブラリ, サンプル用3Dモデル, 音声合成用モデル等を含む)
9. 音声対話コンテンツ設計指針

本パッケージは、マルチプラットフォームに対応したソフトウェア、ユーザーによるコンテンツ制作を支援するツール群、長期運用に基づいた実証実験に裏付けされた音声対話コンテンツの設計指針に、コンテンツ制作者・開発者向けリファレンス・マニュアルや講習用スライド、MOOC・OCW用チュートリアル動画、音声対話コンテンツライブラリを加え、有機的に統合したものであり、本パッケージによりユーザーコンテンツ生成環境を容易に構築可能とするものである。

§ 4. 成果発表等

(1) 原著論文発表 (国内(和文)誌 2 件、国際(欧文)誌 90 件)

【国際】

- [1] Zhizheng Wu, Junichi Yamagishi, Tomi Kinnunen, Cemal Hanilci, Md Sahidullah, Aleksandr Sizov, Nicholas Evans, Massimiliano Todisco, Hector Delgado, “ASVspoof: the Automatic Speaker Verification Spoofing and Countermeasures Challenge,” Special Issue on Spoofing and Countermeasures for Automatic Speaker Verification, IEEE Journal of Selected Topics in Signal Processing 11(4) 1-17, June 2017. (DOI: [10.1109/JSTSP.2017.2671435](https://doi.org/10.1109/JSTSP.2017.2671435))
- [2] Hieu-Thi Luong, Shinji Takaki, Gustav Eje Henter, Junichi Yamagishi, “ADAPTING AND CONTROLLING DNN-BASED SPEECH SYNTHESIS USING INPUT CODES,” Proc ICASSP 2017, March 2017.
<https://www2.securecms.com/ICASSP2017/Papers/ViewPapers.asp?PaperNum=1172>
- [3] Xin Wang, Shinji Takaki, Junichi Yamagishi, “AN AUTO REGRESSIVE RECURRENT MIXTURE DENSITY NETWORK FOR PARAMETRIC SPEECH SYNTHESIS,” Proc ICASSP 2017, March 2017.
<https://www2.securecms.com/ICASSP2017/Papers/ViewPapers.asp?PaperNum=1503>
- [4] Xin Wang, Shinji Takaki, Junichi Yamagishi, “Investigation of Using Continuous Representation of Various Linguistic Units in Neural Network based Text-to-Speech Synthesis,” The Institute of Electronics, Information and Communication Engineers (IEICE) Transactions on Information and Systems, E99-D(10) 2471-2480, October 2016. (DOI: [10.1587/transinf.2016SLP0011](https://doi.org/10.1587/transinf.2016SLP0011))
- [5] Kei Sawada, Chiaki Asai, Kei Hashimoto, Keiichi Oura, and Keiichi Tokuda, “The NITech text-to-speech system for the Blizzard Challenge 2016,” Blizzard Challenge workshop 2016, 2016.
http://www.festvox.org/blizzard/bc2016/NITech_Blizzard2016.pdf
- [6] Lauri Juvela, Xin Wang, Shinji Takaki, SangJin Kim, Manu Airaksinen, Junichi Yamagishi, “The NII speech synthesis entry for Blizzard Challenge 2016,” Blizzard Challenge workshop 2016, 2016.
http://www.festvox.org/blizzard/bc2016/NII_Blizzard2016.pdf
- [7] Keiichi Tokuda, Kei Hashimoto, Keiichi Oura, and Yoshihiko Nankaku, “Temporal modeling in neural network based statistical parametric speech synthesis,” 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016. http://ssw9.net/download/ssw9_proceedings.pdf
- [8] Xin Wang, Shinji Takaki, Junichi Yamagishi, “Investigating Very Deep Highway Networks for Parametric Speech Synthesis,” 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016.
http://ssw9.net/download/ssw9_proceedings.pdf
- [9] Xin Wang, Shinji Takaki, Junichi Yamagishi, “A Comparative Study of the Performance of HMM, DNN, and RNN based Speech Synthesis Systems Trained on Very Large Speaker-Dependent Corpora,” 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016.
http://ssw9.net/download/ssw9_proceedings.pdf
- [10] Shinji Takaki, SangJin Kim, Junichi Yamagishi, “Speaker Adaptation of Various Components in Deep Neural Network based Speech Synthesis,” 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016.
http://ssw9.net/download/ssw9_proceedings.pdf
- [11] Michael Pucher, Fernando Villavicencio, Junichi Yamagishi, “Development and

- evaluation of a statistical parametric synthesis system for operatic singing in German,” 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016. http://ssw9.net/download/ssw9_proceedings.pdf
- [12] Cassia Valentini-Botinhao, Xin Wang, Shinji Takaki, Junichi Yamagishi, “Investigating RNN-based speech enhancement methods for noise-robust Text-to-Speech,” 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016. http://ssw9.net/download/ssw9_proceedings.pdf
 - [13] Dall, R., Tomalin, M. and Wester, M. (2016). Synthesising Filled Pauses: Representation and Datamixing. 9th ISCA Workshop on Speech Synthesis (Satellite workshop after INTERSPEECH 2016), 2016. http://ssw9.net/download/ssw9_proceedings.pdf
 - [14] Masanari Nishimura, Kei Hashimoto, Keiichi Oura, Yoshihiko Nankaku, and Keiichi Tokuda, “Singing voice synthesis based on deep neural networks,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-1027](https://doi.org/10.21437/Interspeech.2016-1027))
 - [15] Naoki Hosaka, Kei Hashimoto, Keiichi Oura, Yoshihiko Nankaku, and Keiichi Tokuda, “Voice conversion based on trajectory model training of neural networks considering global variance,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-1035](https://doi.org/10.21437/Interspeech.2016-1035))
 - [16] Rasmus Dall, Kei Hashimoto, Keiichi Oura, Yoshihiko Nankaku, and Keiichi Tokuda, “Redefining the linguistic context feature set for HMM and DNN TTS through position and parsing,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-399](https://doi.org/10.21437/Interspeech.2016-399))
 - [17] Siva Reddy Gangireddy, Pawel Swietojanski, Peter Bell and Steve Renals. “Unsupervised adaptation of Recurrent Neural Network Language Models”. Interspeech 2016, pp. 2333-2337, San Francisco, Sep. 2016. (DOI: [10.21437/Interspeech.2016-1342](https://doi.org/10.21437/Interspeech.2016-1342))
 - [18] Takenori Yoshimura, Gustav Eje Henter, Oliver Watts, Mirjam Wester, Junichi Yamagishi, and Keiichi Tokuda, “A Hierarchical Predictor of Synthetic Speech Naturalness Using Neural Networks,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-847](https://doi.org/10.21437/Interspeech.2016-847))
 - [19] Xin Wang, Shinji Takaki, Junichi Yamagishi, “Enhance the word vector with prosodic information for the recurrent neural network based TTS system,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-390](https://doi.org/10.21437/Interspeech.2016-390))
 - [20] Fernando Villavicencio, Junichi Yamagishi, Jordi Bonada, Felipe Espic, “Applying Spectral Normalisation and Efficient Envelope Estimation and Statistical Transformation for the Voice Conversion Challenge 2016,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-305](https://doi.org/10.21437/Interspeech.2016-305))
 - [21] Cassia Valentini-Botinhao, Xin Wang, Shinji Takaki, Junichi Yamagishi, “Speech Enhancement for a Noise-Robust Text-to-Speech Synthesis System using Deep Recurrent Neural Networks,” Interspeech 2016, 2016. (DOI: [10.21437/Interspeech.2016-159](https://doi.org/10.21437/Interspeech.2016-159))
 - [22] Yuto Ishida, Takahiro Uchiya, Kouhei Yamamoto, Daisuke Yamamoto, Ryota Nishimura, Ichi Takumi, “Related Word Recommendation Mechanism for Speech Dialogue System,” Proc. of The 19-th International Conference on Network-Based Information Systems, pp.464-468, 2016. (DOI: [10.1109/NBiS.2016.53](https://doi.org/10.1109/NBiS.2016.53))
 - [23] Kei Hashimoto, Keiichi Oura, Yoshihiko Nankaku, and Keiichi Tokuda, “Trajectory training considering global variance for speech synthesis based on neural networks,” in Proc. ICASSP, pp.5600-5604, 2016. (DOI: [10.1109/ICASSP.2016.7472749](https://doi.org/10.1109/ICASSP.2016.7472749))
 - [24] Keiichi Tokuda and Heiga Zen, “Directly modeling voiced and unvoiced components in speech waveforms by neural networks,” in Proc. ICASSP, pp.5640-5644, 2016. (DOI: [10.1109/ICASSP.2016.7472757](https://doi.org/10.1109/ICASSP.2016.7472757))
 - [25] Kei Hashimoto, Junichi Yamagishi, and Isao Echizen, “Privacy-preserving sound to

- degrade automatic speaker verification performance,” in Proc. ICASSP, pp.5500-5504, 2016. (DOI: [10.1109/ICASSP.2016.7472757](https://doi.org/10.1109/ICASSP.2016.7472757))
- [26] R. Dall, S. Brognaux, K. Richmond, C. Valentini-botinhao, G. E. Henter, J. Hirschberg, J. Yamagishi, and S. King, “Testing the Consistency Assumption: Pronunciation Variant Forced Alignment in Read and Spontaneous Speech Synthesis,” In Proc. ICASSP, pp.5155-5159, Shanghai, China, March 2016. (DOI: [10.1109/ICASSP.2016.7472660](https://doi.org/10.1109/ICASSP.2016.7472660))
- [27] Takahiro Uchiya, Masaki Yoshida, Daisuke Yamamoto, Ryota Nishimura, Ichi Takumi, “Design and Implementation of Open-Campus Event System with Voice Interaction Agent,” International Journal of Mobile Multimedia, Vol.11, No.3&4, pp237-250, 2015. (DOI: [10.1109/NBiS.2015.87](https://doi.org/10.1109/NBiS.2015.87))
- [28] Keitaro Wakabayashi, Daisuke Yamamoto, Naohisa Takahashi, “A Voice Dialog Editor Based on Finite State Transducer Using Composite State for Tablet Devices,” Computer and Information Science 2015, Studies in Computational Intelligence, vol. 614, pp.125-139, 2015. (DOI: [10.1007/978-3-319-23467-0_9](https://doi.org/10.1007/978-3-319-23467-0_9))
- [29] Tomoya Mori, Shinsuke Kajioka, Takahiro Uchiya, Ichi Takumi, Hiroshi Matsuo, “Experiments of Position Estimation by BLE Beacons on Actual Situations,” Proc. of the IEEE GCCE 2015, pp.50-51, 2015. (DOI: [10.1109/GCCE.2015.7398727](https://doi.org/10.1109/GCCE.2015.7398727))
- [30] Takahiro Uchiya, Ryosuke Nakano, Daisuke Yamamoto, Ryota Nishimura, Ichi Takumi, “Extension with Intelligent Agents for the Spoken Dialogue System for Smartphones,” Proc. of the IEEE GCCE 2015, pp.298-299, 2015. (DOI: [10.1109/GCCE.2015.7398666](https://doi.org/10.1109/GCCE.2015.7398666))
- [31] Yuichi Matsushita, Takahiro Uchiya, Ryota Nishimura, Daisuke Yamamoto, Ichi Takumi, “Experiment and Evaluation of Crowd Sourcing Model for Creation of Voice Interaction Scenario,” Proc. of the IEEE GCCE 2015, pp.321-322, 2015. (DOI: [10.1109/GCCE.2015.7398647](https://doi.org/10.1109/GCCE.2015.7398647))
- [32] Kathrin Haag and Hiroshi Shimodaira, “The University of Edinburgh Speaker Personality and MoCap Dataset,” In proceedings of Facial Analysis and Animation, Vienna, No.8, September, 2015. (Doi: [10.1145/2813852.2813860](https://doi.org/10.1145/2813852.2813860))
- [33] Kei Sawada, Kei Hashimoto, Keiichi Oura, and Keiichi Tokuda, “The NITECH HMM-based text-to-speech system for the Blizzard Challenge 2015,” in Proc. Blizzard Challenge, 2015.
http://www.festvox.org/blizzard/bc2015/NITECH_bc2015.pdf
- [34] Takenori Yoshimura, Kei Hashimoto, Keiichi Oura, Yoshihiko Nankaku, and Keiichi Tokuda, “Simultaneous optimization of multiple tree structures for factor analyzed HMM-based speech synthesis,” in Proc. Interspeech, pp.1196-1200, 2015.
http://isca-speech.org/archive/interspeech_2015/i15_1196.html
- [35] Sva Reddy Gangireddy, Steve Renals, Yoshihiko Nankaku, and Akinobu Lee, “Prosodically-enhanced Recurrent Neural Network Language Models,” In Proc. of Interspeech, 2015. http://isca-speech.org/archive/interspeech_2015/i15_2390.html
- [36] M. Wester, M. Aylett, M. Tomalin, and R. Dall, “Artificial Personality and Disfluency,” In Proc. Interspeech, 2015.
http://isca-speech.org/archive/interspeech_2015/i15_3365.html
- [37] Shinji Takaki, SangJin Kim, Junichi Yamagishi, JongJin Kim, “Multiple Feed-forward Deep Neural Networks for Statistical Parametric Speech Synthesis,” In Proc. Interspeech, pp.2242-2246, 2015.
http://isca-speech.org/archive/interspeech_2015/i15_2242.html
- [38] Sayaka Shiota, Fernando Villavicencio, Junichi Yamagishi, Nobutaka Ono, Isao Echizen, Tomoko Matsui, “Voice liveness detection algorithms based on pop noise caused by human breath for automatic speaker verification,” In Proc. Interspeech, pp.239-243, 2015. http://isca-speech.org/archive/interspeech_2015/i15_0239.html
- [39] Takahiro Uchiya, Masaki Yoshida, Daisuke Yamamoto, Ichi Takumi, “Development of Open-Campus System Using Voice Interaction Agent,” Proc. of the 18-th

- International Conference on Network-Based Information Systems (NBIS-2015), pp.591-595, 2015. (DOI: [10.1109/NBiS.2015.87](https://doi.org/10.1109/NBiS.2015.87))
- [40] Yuichi Matsushita, Takahiro Uchiya, Ryota Nishimura, Daisuke Yamamoto, Ichi Takumi, "Crowdsourcing Environment to Create Voice Interaction Scenario of Spoken Dialogue System," Proc. of the 18-th International Conference on Network-Based Information Systems (NBIS-2015), pp.500-504, 2015. (DOI: [10.1109/NBiS.2015.74](https://doi.org/10.1109/NBiS.2015.74))
- [41] Fernando Villavicencio, Jorbi Bonada, Yuji Hisaminato, "Observation-Model Error Compensation for Enhanced Spectral Envelope Transformation in Voice Conversion," 2015 IEEE 25th International Workshop on MLSP, September 2015. (DOI: [10.1109/MLSP.2015.7324328](https://doi.org/10.1109/MLSP.2015.7324328))
- [42] M. Tomalin, M. Wester, R. Dall, B. Byrne, and S. King, "A Lattice-Based Approach to Automatic Filled Pause Insertion," In Proc. Disfluencies in Spontaneous Speech, 2015. http://disfluency.org/DiSS_2015/Programme_files/Tomalin-DISS2015.pdf
- [43] R. Dall, M. Wester, and M. Corley, "Disfluencies in Change Detection in Natural, Vcoded and Synthetic Speech," In Proc. Disfluencies in Spontaneous Speech, 2015. http://disfluency.org/DiSS_2015/Programme_files/Dall-DISS2015.pdf
- [44] M. Wester, M. Corley, and R. Dall, "The Temporal Delay Hypothesis: Natural, Vcoded and Synthetic Speech" In Proc. Disfluencies in Spontaneous Speech, 2015. http://disfluency.org/DiSS_2015/Programme_files/Wester-DISS2015.pdf
- [45] Tomohiro Yanagi, Daisuke Yamamoto, Naohisa Takahashi, "Development of Mobile Voice Navigation System Using User-Based Mobile Maps Annotations," Proc. of IEEE/ACS ICIS 2015, pp.373-378, 2015. (DOI: [10.1109/ICIS.2015.7166622](https://doi.org/10.1109/ICIS.2015.7166622))
- [46] Masaki Murase, Daisuke Yamamoto, Naohisa Takahashi, "On-demand Generalization of Guide Maps with Road Networks and Category-based Web Search Results," Proc. of W2GIS 2015, Springer LNCS, vol. 9080, pp.53-70, 2015. (DOI: [10.1007/978-3-319-18251-3_4](https://doi.org/10.1007/978-3-319-18251-3_4))
- [47] Kei Hashimoto, Keiichi Oura, Yoshihiko Nankaku, and Keiichi Tokuda, "The effect of neural networks in statistical parametric speech synthesis," in Proc ICASSP, pp.4455-4459, 2015. (DOI: [10.1109/ICASSP.2015.7178813](https://doi.org/10.1109/ICASSP.2015.7178813))
- [48] Keiichi Tokuda and Heiga Zen, "Directly modeling speech waveforms by neural networks for statistical parametric speech synthesis," in Proc. ICASSP, pp.4215-4219, 2015. (DOI: [10.1109/ICASSP.2015.7178765](https://doi.org/10.1109/ICASSP.2015.7178765))
- [49] Ryota Nishimura, Daisuke Yamamoto, Takahiro Uchiya, Ichi Takumi, "Development of a Dialogue Scenario Editor on a Web Browser for a Spoken Dialogue System," Proceedings of the 2nd International Conference on Human Agent Interaction (HAI 2014), ACM digital library, pp.129-132, 2014. (DOI: [10.1145/2658861.2658904](https://doi.org/10.1145/2658861.2658904))
- [50] Daisuke Yamamoto, Keiichi 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, 2014. (DOI: [10.1145/2658861.2658874](https://doi.org/10.1145/2658861.2658874))
- [51] Yuichi Matsushita, Takahiro Uchiya, Ryota Nishimura, Daisuke Yamamoto, Takumi Ichi, "Construction of Crowdsourcing Environment for Creation of Voice Interaction Scenario," Proceedings of the IEEE GCCE 2014, pp.689-690, 2014. (DOI: [10.1109/GCCE.2014.7031156](https://doi.org/10.1109/GCCE.2014.7031156))
- [52] Ryota Nishimura, Koji Sugioka, Daisuke Yamamoto, Takahiro Uchiya, Ichi Takumi, "A VoIP-based Voice Interaction System for a Virtual Telephone Operator Using Video Calls," Proceedings of the IEEE GCCE 2014, pp.529-530, 2014. (DOI: [10.1109/GCCE.2014.7031317](https://doi.org/10.1109/GCCE.2014.7031317))
- [53] Kei Sawada, Shinji Takaki, Kei Hashimoto, Keiichi Oura, Keiichi Tokuda, "Overview of NITECH HMM-based text-to-speech system for Blizzard Challenge

- 2014,” Blizzard Challenge 2014 Workshop, 2014.
http://www.festvox.org/blizzard/bc2014/NITECH_Blizzard2014.pdf
- [54] Kazuhiro Nakamura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda, “A mel-cepstral analysis technique restoring high frequency components from low-sampling-rate speech,” proceedings of Interspeech 2014, pp.2494-2498, 2014. http://isca-speech.org/archive/interspeech_2014/i14_2494.html
 - [55] Rasmus Dall, Marcus Tomalin, Mirjam Wester, William Byrne, Simon King, “Investigating Automatic & Human Filled Pause Insertion for Speech Synthesis,” proceedings of Interspeech 2014, pp. 51-55, 2014.
http://isca-speech.org/archive/interspeech_2014/i14_0051.html
 - [56] Rasmus Dall, Mirjam Wester, Martin Corley, “The Effect of Filled Pauses and Speaking Rate on Natural, Vcoded and Synthetic Speech,” proceedings of Interspeech 2014, pp.56-60, 2014.
http://isca-speech.org/archive/interspeech_2014/i14_0056.html
 - [57] 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, 2014.
http://isca-speech.org/archive/interspeech_2014/i14_2620.html
 - [58] Takahiro Uchiya, Daisuke Yamamoto, Ryota Nishimura, Ichi Takumi, Tetsuo Kinoshita “Proposal of Voice Control-based Interactive Design Method of Agent System,” Proceedings of the 13th IEEE International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC2014), pp.496-500, 2014. (DOI: [10.1109/ICCI-CC.2014.6921505](https://doi.org/10.1109/ICCI-CC.2014.6921505))
 - [59] Kazuhiro Nakamura, Kei Hashimoto, Yoshihiko Nankaku, and Keiichi Tokuda, “Integration of spectral feature extraction and modeling for HMM-based speech synthesis,” IEICE Transactions on Information and Systems, vol. E97-D, no. 6, 2014. (DOI: [10.1587/transinf.E97.D.1438](https://doi.org/10.1587/transinf.E97.D.1438))
 - [60] Rasmus Dall, Junichi Yamagishi, Simon King, “Rating naturalness in speech synthesis: The effect of style and expectation,” Proceedings of Speech Prosody 7, 2014. <http://fastnet.net.soc.ie/sp7/sp7book.pdf>
 - [61] Oliver Watts, Siva Reddy Gangireddy, Junichi Yamagishi, Simon King, Steve Renals, Adriana Stan, and Mircea Giurgiu, “NEURAL NET WORD REPRESENTATIONS FOR PHRASE-BREAK PREDICTION WITHOUT A PART OF SPEECH TAGGER,” proceedings of ICASSP 2014, pp. 2618-2622, 2014. (DOI: [10.1109/ICASSP.2014.6854070](https://doi.org/10.1109/ICASSP.2014.6854070))
 - [62] Kazuhiro Nakamura, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda, “HMM-based synging voice synthesis and its application to Japanese and English,” proceedings of 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014), pp. 265-269, 2014.
(DOI: [10.1109/ICASSP.2014.6853599](https://doi.org/10.1109/ICASSP.2014.6853599))
 - [63] Kanako Shirota, Kazuhiro Nakamura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda, “Integration of speaker and pitch adaptive training for HMM-based singing voice synthesis,” proceedings of 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014), pp. 2578-2582, 2014. (DOI: [10.1109/ICASSP.2014.6854062](https://doi.org/10.1109/ICASSP.2014.6854062))
 - [64] Shinji Takaki, Yoshihiko Nankaku and Keiichi Tokuda, “Contextual additive structure for HMM-based speech synthesis,” IEEE Journal of Selected Topics in Signal Processing, vol. 8, issue. 2, pp. 229-238, 2014.
(DOI: [10.1109/JSTSP.2014.2305919](https://doi.org/10.1109/JSTSP.2014.2305919))
 - [65] Shinji Takaki, Kei Sawada, Kei Hashimoto, Keiichiro Oura, and Keiichi Tokuda, “Overview of NIT HMM-based speech synthesis system for Blizzard Challenge 2013,” Blizzard Challenge 2013 Workshop, 2013 (web proceedings).
http://www.festvox.org/blizzard/bc2013/NITECH_Blizzard2013.pdf

- [66] Takenori Yoshimura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda, "Cross-lingual speaker adaptation based on factor analysis using bilingual speech data for HMM-based speech synthesis," 8th ISCA Speech Synthesis Workshop, pp. 317-322, 2013.
http://isca-speech.org/archive/ssw8/ssw8_297.html
- [67] N. d'Alessandro, J. Tilmanne, M. Astrinaki, T. Hueber, R. Dall, T. Ravet, A. Moinet, H. Cakmak, O. Babacan, A. Barbulescu, V. Parfait, V. Huguenin, E. S. Kalayci, and Q. Hu, "Reactive statistical mapping: Towards the sketching of performative control with data," IFIP Advances in Information and Communication Technology, Special Issue: Proceedings of the 9th eNTERFACE Workshop, 2013.
(DOI: [10.1007/978-3-642-55143-7_2](https://doi.org/10.1007/978-3-642-55143-7_2))
- [68] Motohiro Shibakawa, Takahiro Uchiya, Ichi Takumi, Tetsuo Kinoshita, "Design and implementation of multiagent-based distributed backup system," Proceedings of the 12th IEEE/ACIS International Conference on Computer and Information Science, pp. 234-240, 2013. (DOI: [10.1109/ICIS.2013.6607847](https://doi.org/10.1109/ICIS.2013.6607847))
- [69] Takahiro Uchiya, Syo Itazuro, Ichi Takumi, Tetsuo Kinoshita, "Design environment of intelligent multiagent systems," Proceedings of the 12th IEEE/ACIS International Conference on Computer and Information Science, pp.241-246, 2013. (DOI: [10.1109/ICIS.2013.6607848](https://doi.org/10.1109/ICIS.2013.6607848))
- [70] Shinji Takaki, Yoshihiko Nankaku, and Keiichi Tokuda, "Contextual partial additive structure for HMM-based speech synthesis," Proceedings of 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), pp. 7878-7882, 2013. (DOI: [10.1109/ICASSP.2013.6639198](https://doi.org/10.1109/ICASSP.2013.6639198))
- [71] Kazuhiro Nakamura, Kei Hashimoto, Yoshihiko Nankaku, and Keiichi Tokuda, "Integration of acoustic modeling and mel-cepstral analysis for HMM-based speech synthesis," Proceedings of 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), pp. 7883-7887, 2013. (DOI: [10.1109/ICASSP.2013.6639199](https://doi.org/10.1109/ICASSP.2013.6639199))
- [72] Akinobu Lee, Keiichiro Oura, and Keiichi Tokuda, "MMDAgent -a fully open-source toolkit for voice interaction systems-," Proceedings of 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), pp. 8382-8385, 2013. (DOI: [10.1109/ICASSP.2013.6639300](https://doi.org/10.1109/ICASSP.2013.6639300))
- [73] Keiichi Tokuda, Yoshihiko Nankaku, Tomoki Toda, Heiga Zen, Junichi Yamagishi, and Keiichiro Oura, "Speech synthesis based on hidden Markov models," Proceedings of the IEEE, vol. 101, no. 5, pp. 1234-1252, 2013.
(DOI: [10.1109/JPROC.2013.2251852](https://doi.org/10.1109/JPROC.2013.2251852)) (招待解説)
- [74] Wenpeng Wei, Hideyuki Takahashi, Takahiro Uchiya, Tetsuo Kinoshita, "Cooperation protocol design method for repository-based multi-agent applications," International Journal of Cognitive Informatics & Natural Intelligence, vol. 5, no. 2, pp. 1-14, 2013. (DOI: [10.4018/ijssci.2013040101](https://doi.org/10.4018/ijssci.2013040101))
- [75] Pablo Martinez Lerin, Daisuke Yamamoto, Naohisa Takahashi, "Encoding network-constrained travel trajectories using routing algorithms," International Journal of Knowledge and Web Intelligence, vol. 4, no. 1, pp. 34-49, 2013. (DOI: [10.1504/IJKWI.2013.052724](https://doi.org/10.1504/IJKWI.2013.052724))
- [76] Hiroya Mizutani, Daisuke Yamamoto, Naohisa Takahashi, "A preventing method for overlapping focuses in a Focus+Glue+Context map," International Journal of Knowledge and Web Intelligence, vol. 4, no. 1, pp. 50-69, 2013. (DOI: [10.1504/IJKWI.2013.052725](https://doi.org/10.1504/IJKWI.2013.052725))
- [77] Takahiro Uchiya, Ichi Takumi, Tetsuo Kinoshita, "Circulation mechanism of multiagent system based on agent repository," Proceedings of IEEE BWCCA-2012, pp. 668-672, 2012. (DOI: [10.1109/BWCCA.2012.116](https://doi.org/10.1109/BWCCA.2012.116))
- [78] Naoya Tatematsu, Takahiro Uchiya, Ichi Takumi, Tetsuo Kinoshita, "Enhancement of repository based agent framework for ubiquitous environment," Proceedings of

- IEEE BWCCA-2012, pp.673-678, 2012. (DOI: [10.1109/BWCCA.2012.117](https://doi.org/10.1109/BWCCA.2012.117))
- [79] Syo Itazuro, Takahiro Uchiya, Ichi Takumi, Tetsuo Kinoshita, “Design Environment of reinforcement learning agents for intelligent multiagent system,” Proceedings of IEEE BWCCA-2012, pp.679-683, 2012. (DOI: [10.1109/BWCCA.2012.118](https://doi.org/10.1109/BWCCA.2012.118))
- [80] Takahiro Uchiya, Tetsuo Kinoshita, “Enhancement of agent design environment for agent interoperability,” Proceedings of IW-STEIC2012, pp. 89-94, 2012.
- [81] Pablo Martinez Lerin, Daisuke Yamamoto, Naohisa Takahashi, “Mouse-sensitive following path suggestion for drawing travel routes in web map systems,” Journal of Geographic Information System, vol. 4, no. 5, pp. 393-402, 2012. (DOI: [10.4236/jgis.2012.45045](https://doi.org/10.4236/jgis.2012.45045))
- [82] Shinji Takaki, Kei Sawada, Kei Hashimoto, Keiichi Oura, Keiichi Tokuda, “Overview of NIT HMM-based speech synthesis system for Blizzard Challenge 2012,” Blizzard Challenge 2012 Workshop, 2012 (web proceedings). http://www.festvox.org/blizzard/bc2012/NITECH_Blizzard2012.pdf
- [83] Takafumi Hattori, Kei Hashimoto, Yoshihiko Nankaku, Keiichi Tokuda, “A Bayesian approach to speaker recognition based on GMMs using multiple model structures,” Proceedings of Interspeech 2012, 2012. http://isca-speech.org/archive/interspeech_2012/i12_1107.html
- [84] Viviane de Franca Oliveira, Sayaka Shiota, Yoshihiko Nankaku, Keiichi Tokuda, “Crosslingual speaker adaptation for HMM-based speech synthesis based on perceptual characteristics and speaker interpolation,” Proceedings of Interspeech 2012, 2012. http://isca-speech.org/archive/interspeech_2012/i12_0983.html
- [85] Syo Itazuro, Takahiro Uchiya, Ichi Takumi, Tetsuo Kinoshita, “Design support mechanism of learning agent on repository-based agent framework,” Proceedings of ICCI*CC2012, pp.300-304, 2012. (DOI: [10.1109/ICCI-CC.2012.6311165](https://doi.org/10.1109/ICCI-CC.2012.6311165))
- [86] Wenpeng Wei, Hideyuki Takahashi, Takahiro Uchiya and Tetsuo Kinoshita, “Repository-based methodology of cooperation protocol design for multi-agent system,” Proceedings of ICCI*CC2012, pp.283-288, 2012. (DOI: [10.1109/ICCI-CC.2012.6311162](https://doi.org/10.1109/ICCI-CC.2012.6311162))
- [87] Syo Itazuro, Takahiro Uchiya, Tetsuo Kinoshita, Ichi Takumi, “Development support of learning agent on repository-based agent framework,” International Journal of Software Science and Computational Intelligence, vol. 4, issue 3, pp. 62-79, 2012. (DOI: [10.4018/jssci.2012070104](https://doi.org/10.4018/jssci.2012070104)) **(Best Paper Award)**
- [88] Pablo Martinez Lerin, Daisuke Yamamoto and Naohisa Takahashi, “Encoding travel traces by using road networks and routing algorithms,” Proceedings of KES IIMSS 2012, Gifu, pp. 233-243, 2012. (DOI: [10.1007/978-3-642-29934-6_23](https://doi.org/10.1007/978-3-642-29934-6_23))
- [89] Hiroya Mizutani, Daisuke Yamamoto and Naohisa Takahashi, “A fusion of multiple focuses on a Focus + Glue + Context map fusion,” Proceedings of KES IIMSS 2012, pp. 11-21, 2012. (DOI: [10.1007/978-3-642-29934-6_2](https://doi.org/10.1007/978-3-642-29934-6_2))
- [90] Pablo Martinez Lerin, Daisuke Yamamoto and Naohisa Takahashi, “Making a pictorial and verbal travel trace from a GPS trace,” Proceedings of W2GIS 2012, pp. 98-115, 2012. (DOI: [10.1007/978-3-642-29247-7_9](https://doi.org/10.1007/978-3-642-29247-7_9))

【国内】

- [91] 大浦圭一郎, “統計モデルに基づいた歌声合成技術の最先端,” 電子情報通信学会誌, 2015. <http://ci.nii.ac.jp/naid/40020498942>
- [92] 大浦圭一郎, 山本大介, 内匠逸, 李晃伸, 徳田恵一, “キャンパスの公共空間におけるユーザ参加型双方向音声案内デジタルサイネージシステム,” 人工知能学会誌, vol.28, no.1, pp.60-67, 2013. (招待解説) <http://ci.nii.ac.jp/naid/110009563047>

(2) その他の著作物(総説, 書籍など)

- [1] 山岸 順一, “音声の障害者のための最先端音声合成技術,” JST 情報管理 57(12) 882-889, 2015. https://www.jstage.jst.go.jp/article/johokanri/57/12/57_882/pdf
- [2] 山岸順一, 徳田恵一, 戸田智基, みわよしこ, “情報研シリーズ 19 おしゃべりなコンピュータ—音声合成技術の現在と未来—”, 丸善ライブラリ, 2015. (ISBN: [978-4-621-05385-0](https://www.jstage.jst.go.jp/article/johokanri/57/12/57_882/pdf))
- [3] Korin Richmond, 山岸順一, Zhenhua Ling, “調音運動の機械学習に基づく応用,” 日本音響学会誌 解説記事, pp.539-545, 2015. <http://ci.nii.ac.jp/naid/110009987702>
- [4] 李晃伸, 大浦圭一郎ほか, “音響キーワードブック,” コロナ社, 2016. (ISBN : [978-4-339-00880-7](https://www.jstage.jst.go.jp/article/johokanri/57/12/57_882/pdf))
- [5] Nicholas Evans, Tomi Kinnunen, Junichi Yamagishi, Zhizheng Wu, Federico Alegre, Phillip de Leon, “Voice anti-spoofing,” (担当:分担執筆) Handbook of Biometric Anti-Spoofing, Elsevier, 2014. (ISBN:978-1447165231)

(3) 国際学会発表及び主要な国内学会発表

① 招待講演 (国内会議 22 件, 国際会議 17 件)

【国際】

- [1] Steve Renals, “Natural speech technology,” IEEE ASRU, Scottsdale AZ, USA, December 17, 2015.
- [2] Zhizheng Wu, Tomi Kinnunen, Nicholas Evans, and Junichi Yamagishi, “Spoofing and anti-spoofing: a shared view of speaker verification, speech synthesis and voice conversion,” the Asia-Pacific Signal and Information Processing Association (APSIPA) ASC 2015, The Hong Kong Polytechnic University, December 16, 2015.
- [3] Fernando Villavicencio, “Speech synthesis: from diphones to modern speech synthesis engines,” 2015 Speech Processing Courses in Crete, University of Crete, Greece, July 31, 2015.
- [4] Steve Renals, “Multilingual Interaction Technologies,” Riga Summit, Riga, Latvijas Republika, April 27, 2015.
- [5] Steve Renals, “Neural networks for distant speech recognition,” HSCMA 2014 workshop, May 12-14, 2014. (Invited keynote)
<https://hscma2014.inria.fr/steve-renals/>
- [6] Steve Renals, Speech Technology and Human Computer Interaction Workshop, Edinburgh, UK, March 27, 2014. (Invited keynote)
<http://www.sicsa.ac.uk/events/sicsa-hci-speech-technology-human-computer-interaction-workshop/>
- [7] Keiichi Tokuda, “Flexible speech synthesis based on hidden Markov models,” Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2013 (APSIPA ASC 2013), Kaohsiung, Taiwan, October 29-November 1, 2013. (Invited keynote)
http://www.apsipa2013.org/?page_id=659#talk3
- [8] Steve Renals, “(Deep) Neural Networks for Speech Recognition,” UKSpeech 2013, Cambridge, UK, September 17-18, 2013. (Invited tutorial)
<http://www.ukspeech.org.uk/workshop/cambridge2013.php>
- [9] Steve Renals, “Conversational Interaction Technologies - linking research & innovation -,” LT-Innovate Summit 2013, Brussels, Belgium, June 2013. (Invited panelist)
<http://www.lt-innovate.eu/event/LTi-events/lt-innovate-summit-2013-brussels/Presentations>
- [10] Steve Renals, “Neural Networks for Speech Recognition,” NOLISP 2013, Mons, Belgium, June 2013. (Invited keynote)

- <http://www.tcts.fpms.ac.be/nolisp2013/plenary.php>
- [11] Steve Renals, “Machine Learning and Natural Language,” Microsoft Machine Learning Summit 2013, Paris, France, April 2013. (Invited speaker)
<http://research.microsoft.com/en-us/um/cambridge/events/mls2013/agenda.htm>
 - [12] Steve Renals, “Current speech research in Edinburgh,” UKSpeech conference, Birmingham, UK, December 2012.
 - [13] Junichi Yamagishi, “HMM-based speech synthesis adapted to listener’s and talker’s conditions,” The Listening Talker Workshop, Edinburgh, UK, May 2012. (Invited speaker)
<http://listening-talker.org/workshop/abstracts/abstracts.html#yamagishi>
 - [14] Steve Renals, “Natural Speech Technology”, CREST Symposium on Human-Harmonized Information Technology -- Behavior, Interaction, Music, and UGC --, Kyoto, Japan, April 2012. (Invited talk)
<http://www.ar.media.kyoto-u.ac.jp/crest/sympo12/>
 - [15] Steve Renals, “Towards Natural Speech Technology,” Workshop on Innovation and Applications in Speech Technology, Dublin, Ireland, March 9-19, 2012. (Invited keynote)
<http://muster.ucd.ie/workshops/iast/#keynote>
 - [16] Steve Renals, “Natural Speech Technology,” Workshop on Image and Speech Processing, Hyderabad, India, December 17, 2011. (Invited keynote)
http://www.iitg.ernet.in/ece/emstlab/pdf/WISP_2011.pdf
 - [17] Keiichi Tokuda, “Speech synthesis as a statistical machine learning problem,” 2011 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2011), Hawaii, USA, December 14, 2011. (Invited talk)
<http://www.superlectures.com/asru2011/speech-synthesis-as-a-statistical-machine-learning-problem>

【国内】

- [18] 高木信二, “とても Deep なテキスト音声合成” 信学技報 116(414) 41-46 2017 年 1 月.
- [19] 山岸順一, “EMA データを利用した統計的音声合成における調音制御,” 音声研究会, 2016 年 1 月 14 日.
- [20] 徳田恵一, “統計的音声合成技術の現在・過去・未来,” 音声言語情報処理研究会, 2015 年 12 月 2 日.
- [21] Keiichi Tokuda, “Text-to-speech synthesis a machine learning problem,” Recent Advances in Machine Learning and Applications, 2015 年 11 月 5 日.
- [22] 山岸順一, “短時間で高品質なクローンを可能にする新しい音声合成技術,” 次世代デジタルコンテンツが拓く新ビジネス創出セミナー2015, 2015 年 9 月 18 日.
- [23] 山本大介, 高木信二, “AI ツール入門講座 2: 音声インタラクションシステム構築ツールキット MMDAgent,” 人工知能学会第 8 回, 2015 年 6 月 22 日.
- [24] 山岸順一, “音声合成の現在と音声翻訳のための新技術,” 第 69 回人工知能セミナー「音声翻訳技術の現在、未来」, 2015 年 5 月 1 日.
- [25] 山岸順一, “音声の障害者のための最先端音声合成技術,” 医療とニューメディアを考える会, 第 312 回勉強会 2014 年 5 月 27 日. (招待講演)
http://iryoutonewmedia.web.fc2.com/pdf/zisseki_b.pdf
- [26] 山岸順一, “音声合成の話者適応技術と福祉応用,” 音学シンポジウム 2014, 日本大学, 2014 年 5 月 24 日. (招待講演)
http://www.sigmus.jp/?page_id=2713
- [27] 徳田恵一, “テーマセッション:「音声合成」は今後こうなる!,” 情報処理学会音声言語情報処理研究会 (SIG-SLP 第 100 回記念シンポジウム), 2014 年 2 月 1 日. (パネリスト)
<http://sig-slp.jp/2014-SLP-100.html>
- [28] 李晃伸, “テーマセッション:「音声応用(対話・検索)」は今後こうなる!,” 情報処理学会音声

言語情報処理研究会(SIG-SLP 第100回記念シンポジウム), 2014年2月1日。(パネリスト)

<http://sig-slp.jp/2014-SLP-100.html>

- [29] 山本大介, “スマートフォン向け音声対話 3D エージェントの開発,” 情報処理学会東海支部講演会, 2013年10月23日。(招待講演)
<http://www.ipsj-tokai.jp/jigyoku/files/H25-poster20131023.pdf>
- [30] 山岸順一, “ここまで来た HMM 音声合成話者適応技術,” 日本音響学会 2013 年秋期研究発表会, 豊橋科学技術大学, 2013 年 9 月 27 日。(招待講演)
http://www.asj.gr.jp/annualmeeting/pdf/2013autumn_program.pdf
- [31] 徳田恵一, “ユーザ参加型双方向音声案内デジタルサイネージシステム:「メイちゃん」,” 電気関係学会東海支部連合大会, 2013 年 9 月 24 日。(招待講演)
- [32] 李晃伸, 大浦圭一郎, 徳田恵一, “音声インタラクションシステム構築のためのオープンソースツールキット MMDAgent,” 電気関係学会東海支部連合大会, 2013 年 9 月 24 日。(招待講演)
- [33] 山本大介, 大浦圭一郎, 西村良太, 打矢隆弘, 内匠逸, 李晃伸, 徳田恵一, “スマートフォンのためのスタンドアローン型モバイル音声対話エージェントの取り組み,” 電気関係学会東海支部連合大会, 2013 年 9 月 24 日。(招待講演)
- [34] 打矢隆弘, 山本大介, 西村良太, 内匠逸, 大浦圭一郎, 李晃伸, 徳田恵一, “映像音声ストリーミング技術を用いた音声対話エージェントの開発と実証実験,” 電気関係学会東海支部連合大会, 2013 年 9 月 24 日。(招待講演)
- [35] 山本大介, “音声対話 3D エージェント「メイちゃん」の開発と運用 - 大型サイネージからスマートフォンまで -,” 東海工学教育協会講演会, 2013 年 3 月 24 日。
- [36] 徳田恵一, “統計的機械学習問題としての音声合成,” 情報処理学会音楽情報科学研究会(音学シンポジウム), 2013 年 3 月 11 日。(招待講演)
http://www.sigmus.jp/?page_id=1683
- [37] 山岸順一, “EU および英国の研究費の仕組み: 国際的な研究活動にむけて,” 第4回若手研究者フォーラム, 東京大学, 2012 年 12 月 22 日。(招待講演)
<http://w3voice.jp/onsei-wakate/>
- [38] 徳田恵一, 大浦圭一郎, 李晃伸, 山本大介, 打矢隆弘, 内匠逸, “ユーザ参加型双方向音声案内デジタルサイネージシステムの開発・設置・運用事例”, 日本音響学会春季研究発表会, 東京, pp.119—122, 2013 年 3 月 14 日。(招待講演)
http://www.asj.gr.jp/annualmeeting/pdf/2013spring_program.pdf
- [39] 徳田恵一, “コンテンツ生成の循環系を軸とした音声技術基盤の構築を目指して”, 音声言語シンポジウム, 東京, 2011 年 12 月 20 日。(招待講演)
<http://www.ipsj.or.jp/kenkyukai/event/slp89.html>

② 口頭発表 (国内会議 125 件, 国際会議 5 件)

【国際】

- [1] Takahiro Uchiya, “Attractive Voice Interaction Agent on Smartphone and Campus Navigation Application”, International Conference for Leading and Young Computer Scientists (IC-LYCS2017), Okinawa, February 27, 2017.
- [2] Shuhei Koike, Akinobu Lee, “Spoken keyword detection using recurrent neural network language model,” 5th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, November 2016.
- [3] Xin Wang, Shinji Takaki, Junichi Yamagishi, “Investigation of using the highway network to predict the F0 trajectory for text-to-speech synthesis,” 5th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, November 2016.
- [4] Hieu Thi Luong, Shinji Takaki, SangJin Kim, Junichi Yamagishi, “A DNN-based

text-to-speech synthesis system using speaker, gender and age codes,” 5th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, November 2016.

- [5] Christophe Veaux “Voice reconstruction for individuals with vocal disabilities,” IAST: innovation and applications in speech technology, Dublin, March 9, 2012.

【国内】

- [6] 平野雄大, 打矢隆弘, 内匠逸, 西村良太, 桜井優, “音声対話により認知症を抑制する回想法システムの提案”, 情報処理学会第 79 回全国大会, 名古屋, 2017 年 3 月 17 日.
- [7] 村瀬栞, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “ニューラルネットワークに基づく音声合成における音響特徴量抽出条件の検討,” 日本音響学会 2017 年春季研究発表会, 2017 年 3 月 16 日.
- [8] 市川裕詞, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “DNN-GMM ハイブリッドモデルに基づく声質変換の検討,” 日本音響学会 2017 年春季研究発表会, 2017 年 3 月 15 日.
- [9] 佐野敦志, 堤修平, 山本大介, 高橋直久, “周辺環境とインタラクション可能な音声対話用 BLE ビーコンの開発,” 第 9 回データ工学と情報マネジメントに関するフォーラム, 岐阜, 2017 年 3 月 7 日.
- [10] 浮田弥, 山本大介, 高橋直久, “逆進検知機能を有する案内粒度変更可能な音声経路案内システム,” 第 9 回データ工学と情報マネジメントに関するフォーラム, 岐阜, 2017 年 3 月 7 日.
- [11] 山口大介, 堤修平, 山本大介, 高橋直久, “簡単化された実行履歴に基づく音声対話コンテンツ編集システム,” 第 9 回データ工学と情報マネジメントに関するフォーラム, 岐阜, 2017 年 3 月 7 日.
- [12] 田中佑太郎, 堤修平, 山本大介, 高橋直久, “音声対話コンテンツの利用履歴における頻出パターン解析手法の検討,” 第 9 回データ工学と情報マネジメントに関するフォーラム, 岐阜, 2017 年 3 月 7 日.
- [13] Xin Wang, Shinji Takaki, Junichi Yamagishi, “Investigation on an autoregressive recurrent mixture density network for parametric speech synthesis,” 第 115 回音声言語情報処理研究会 (SIG-SLP), 2017 年 2 月 17 日.
- [14] 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “DNN 音声合成における音響特徴量系列とその時間構造の同時モデル化,” 音声研究会, 2017 年 1 月 21 日.
- [15] Hieu Thi Luong, 高木信二, SangJin Kim, 山岸順一, “DNN に基づくテキスト音声合成における話者・ジェンダー・年齢コード利用の検討,” 音声研究会, 2016 年 10 月 27 日.
- [16] Xin Wang, Shinji Takaki, Junichi Yamagishi, “Investigating the impact of a neural network's depth on spectral and F0 modelling for parametric speech synthesis,” 音声研究会, 2016 年 10 月 27 日.
- [17] 興梠斗吾, 李晃伸, “音声対話システムにおける環境および知識の共有表出と話しかけやすさの関連調査,” 人工知能学会 音声・言語理解と対話処理研究会 SIG-SLUD, Vol.78, pp.125-128, 2016 年 10 月.
- [18] 佐藤翔平, 李晃伸, “話しやすい音声対話システム実現のための対人対話における心理特性の関連性調査,” 人工知能学会 音声・言語理解と対話処理研究会 SIG-SLUD, Vol.78, pp.129-134, 2016 年 10 月.
- [19] 山西元樹, 船谷内泰斗, 李晃伸, “音声対話システムのオープンコンテンツ化実現のためのモジュール仕様および管理手法,” 音声言語情報処理研究会, 2016 年 7 月 30 日.
- [20] 三原寛哉, 李晃伸, “ユーザフレンドリーな音声対話システム実現のためのユーザ話速および発話内容に基づくシステム話速制御手法の検討,” 音声言語情報処理研究会, 2016 年 7 月 30 日.
- [21] Wang Xin, 高木信二, 山岸順一, “巨大特定話者データを用いた HMM・DNN・RNN に基づく音声合成システムの性能評価,” 第 112 回音声言語情報処理研究会, 2016 年 7 月

28 日.

- [22] 浮田弥, 山本大介, 高橋直久, “中間点を用いた案内粒度変更可能な音声経路案内システム,” DICOMO2016, 三重, 2016 年 7 月 6 日.
- [23] 山口大介, 堤修平, 山本大介, 高橋直久, “Android 端末のための実効履歴を用いた音声対話コンテンツ編集システム,” DICOMO2016, 三重, 2016 年 7 月 7 日.
- [24] 田中亮佑, 堤修平, 山本大介, 高橋直久, “領域グラフと利用者の位置に基づく音声対話シナリオ更新手法,” DICOMO2016, 三重, 2016 年 7 月 7 日.
- [25] 河原誠斗, 堤修平, 山本大介, 高橋直久, “子供の負担を考慮した子供音声収集システム,” DICOMO2016, 三重, 2016 年 7 月 7 日.
- [26] 佐藤清隆, 打矢隆弘, 山本大介, 内匠逸, “BLE ビーコンを用いた学内見学支援システムの改善,” DICOMO2016, 三重, 2016 年 7 月 8 日.
- [27] 沢田慶, 伊神和輝, 浅井千明, 佐藤雄介, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “統計的パラメトリック音声合成のためのオーディオブックを用いた学習コーパス自動構築,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日. (学生優秀発表賞受賞)
- [28] 鈴木達也, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “条件付確率場に基づく音声の基本周波数推定の検討,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [29] 保坂尚紀, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “ニューラルネットワークに基づく声質変換における系列内変動を考慮したトラジェクトリ学習,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [30] 西村方成, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “Deep Neural Network に基づく歌声合成の検討,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [31] 大浦圭一郎, 橋本佳, 南角吉彦, 徳田恵一, “楽譜情報に基づく事前分布を用いた HMM 歌声合成の検討,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [32] 橋本佳, 山岸順一, 越前功, “話者照合性能を低下させる UBM に基づくプライバシー保護音の検討,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [33] 柴田翔平, 打矢隆弘, 内匠逸, “スマートグラスと音声対話による料理支援,” 情報処理学会第 78 回全国大会, 2016 年 3 月 10 日.
- [34] Wang Xin, 高木信二, 山岸順一, “Using continuous representation of various linguistic units for recurrent neural network based TTS synthesis,” 第 110 回音声言語情報処理研究会, 2016 年 2 月 6 日.
- [35] 村上拓也, 李晃伸, 西川由理, 小島良宏, 遠藤充, “音声対話システムにおけるシステムからの話しかけと他者性認知の関連性の調査,” HAI シンポジウム 2015, 2015 年 12 月 6 日.
- [36] 沢田慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “発音情報が未知の言語におけるテキスト音声合成システム構築法の評価,” 音声言語情報処理研究会, 2015 年 12 月 2 日.
- [37] 山本大介, 堤修平, 打矢隆弘, 内匠逸, “Web ユーザによる音声対話コンテンツ生成環境の構築とそれに基づく実証実験の評価,” 音声言語情報処理研究会, 2015 年 12 月 2 日.
- [38] 高木信二, 山岸順一, “統計的パラメトリック音声合成のための FFT スペクトルからの Deep Auto-encoder に基づく低次元音響特徴量抽出,” 音声言語情報処理研究会, 2015 年 12 月 2 日.
- [39] 徳田恵一, “これからの歌声合成(パネル討論),” 第 109 回音楽情報科学研究会, 2015 年 11 月 7 日.
- [40] 可児龍平, 打矢隆弘, 山本大介, 内匠逸, 西村良太, “音声対話デジタルサイネージ間の通信機構の開発,” 平成 27 年度電気・電子・情報関係学会東海支部連合大会, 2015 年 9 月 28 日.
- [41] 小林隆宏, 内匠逸, 船瀬新王, 打矢隆弘, 西村良太, “MMDAgent による対話型電話交換システムの音声認識率改善,” 平成 27 年度電気・電子・情報関係学会東海支部連合大会, 2015 年 9 月 28 日.
- [42] 加藤正隆, 内匠逸, 船瀬新王, 打矢隆弘, “臨場感に対する輻輳開散運動の効果について,” 平成 27 年度電気・電子・情報関係学会東海支部連合大会, 2015 年 9 月 28 日.

- [43] 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “ニューラルネットワークに基づく音声合成における系列内変動を考慮したトラジェクトリモデル学習,” 日本音響学会 2015 年秋季研究発表会, 2015 年 9 月 16-18 日.
- [44] 橋本佳, 山岸順一, 越前功, “自動話者照合システムの性能を低下させるプライバシー保護音の検討,” 日本音響学会 2015 年秋季研究発表会, 2015 年 9 月 16-18 日.
- [45] 橋本聖矢, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “複数の音響特徴量を用いた対数線形モデルに基づく話者認識の検討,” 日本音響学会 2015 年秋季研究発表会, 2015 年 9 月 16-18 日.
- [46] 沢田慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “発音情報が未知の言語におけるテキスト音声合成システム構築法の検討,” 日本音響学会 2015 年秋季研究発表会, 2015 年 9 月 16-18 日.
- [47] 中村和寛, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “音声の高帯域成分復元を考慮したメルケプストラム分析に基づく HMM 音声合成の検討,” 日本音響学会 2015 年秋季研究発表会, 2015 年 9 月 16-18 日.
- [48] Fernando Villavicencio, Jordi Bonada, Junichi Yamagishi, Michael Pucher, “Automatic Wide-Range Pitch Extraction for Singing-Voice Based on Spectral Amplitude Autocorrelation,” 第 107 回音声言語情報処理研究会, 2015 年 7 月 16 日.
- [49] 田中佑太朗, 山本大介, 高橋直久, “音声対話コンテンツにおける状態遷移ログの可視化と単純化手法,” DICOMO2015, 2015 年 7 月 9 日.
- [50] 可児龍平, 打矢隆弘, 西村良太, 山本大介, 内匠逸, “音声対話デジタルサイネージ間の通信機構の試作,” DICOMO2015, 2015 年 7 月 9 日.
- [51] 橋本佳, 山岸順一, 越前功, “自動話者照合システムの性能を低下させるプライバシープリーザービングサウンドの検討,” 音声研究会, 2015 年 6 月 16-17 日.
- [52] 松下祐一, 打矢隆弘, 西村良太, 山本大介, 内匠逸, “クラウド型音声対話シナリオ作成環境の構築,” 平成 26 年度電気・電子・情報関係学会東海支部連合大会, 2014 年 9 月 8 日.
- [53] 大竹聡史, 打矢隆弘, 山本大介, 西村良太, 内匠逸, “音声対話に基づくハイブリッド型ユーザー嗜好抽出手法,” 平成 26 年度電気・電子・情報関係学会東海支部連合大会, 2014 年 9 月 8 日.
- [54] 神谷翔大, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “H/L 型アクセント推定と音響モデリングを統合した HMM 音声合成の検討,” 日本音響学会秋季研究発表会, 2-7-3, pp. 237-238, September, 2014 年 9 月 3-5 日. (学生優秀発表賞受賞)
- [55] 吉村建慶, 橋本佳, 南角吉彦, 徳田恵一, “因子分析に基づく HMM 音声合成における基底クラスタリングの検討,” 日本音響学会秋季研究発表会, 2-7-4, pp. 239-240, 2014 年 9 月 3-5 日.
- [56] 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “ニューラルネットワークに基づく音声合成における生成モデルの利用の検討,” 日本音響学会秋季研究発表会, 2-7-7, pp. 245-246, 2014 年 9 月 3-5 日. (粟屋潔学術奨励賞)
- [57] 柳倫浩, 山本大介, 高橋直久, “ユーザによる案内情報の付与とそれに基づくモバイル音声道案内システム,” DICOMO2014, 新潟県, 2014 年 7 月 9 日. (優秀プレゼンテーション賞受賞)
- [58] 若林敬太郎, 山本大介, 高橋直久, “タブレット端末のための複合状態を用いた音声対話コンテンツ編集手法,” DICOMO2014, 新潟県, 2014 年 7 月 10 日. (ヤングリサーチ賞受賞)
- [59] 石川博規, 山本大介, 高橋直久, “音声対話コンテンツのパッケージ化とその配信システム,” DICOMO2014, 新潟県, 2014 年 7 月 10 日.
- [60] 仲野良佑, 打矢隆弘, 山本大介, 内匠逸, “エージェントを用いたスマートメイちゃんの拡張,” DICOMO2014, 新潟県, 2014 年 7 月 10 日.
- [61] 手塚祐太, 山本大介, 高橋直久, “モバイル音声案内のための 3 次元表示地図作成システム,” 2014 年電子情報通信学会総合大会, D-11-38, 2014 年 3 月 19 日.

- [62] 石川博規, 山本大介, 高橋直久, “音声対話コンテンツのパッケージ化に関する研究,” 2014 年電子情報通信学会総合大会, D-21-9, 2014 年 3 月 18 日.
- [63] 水野恵輔, 船瀬新王, 内匠逸, “GMM による加速度センサデータからの歩行状態推定,” 電子情報通信学会全国大会予稿集, B-18-3, 2014 年 3 月 18 日.
- [64] 大竹聡史, 打矢隆弘, 内匠逸, “音声インタラクションに基づくユーザの嗜好抽出,” 情報処理学会 第 76 回全国大会, 2014 年 3 月 13 日.
- [65] 松下祐一, 打矢隆弘, 内匠逸, “音声対話シナリオ作成のためのクラウドソーシング環境の構築,” 情報処理学会 第 76 回全国大会, 2014 年 3 月 13 日.
- [66] 伊神陽介, 李晃伸, 徳田恵一, 南角吉彦, “条件付き確率場に基づく仮説の逐次早期確定を用いた低遅延音声インタフェース,” 日本音響学会春季研究発表会, 2-4-7, pp. 63-64, 2014 年 3 月 10-12 日.
- [67] 小升章裕, 李晃伸, 南角吉彦, 徳田恵一, “統計的音声対話システムにおける登録キーワードの近傍単語を優先した仮説生成に基づく応答選択,” 日本音響学会春季研究発表会, 3-Q-13, pp. 221-223, 2014 年 3 月 10-12 日.
- [68] 佐藤雄介, 中村和寛, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “表現語空間を用いた連結固有声法に基づくクロスリンガル話者適応の検討,” 日本音響学会春季研究発表会, 3-6-2, pp. 325-326, 2014 年 3 月 10-12 日.
- [69] 鶴野高輝, 橋本佳, 南角吉彦, 徳田恵一, “GMM 事後確率に基づいた重み付き変換関数による声質変換の検討” 日本音響学会春季研究発表会, 3-6-3, pp. 327-328, 2014 年 3 月 10-12 日.
- [70] 虫鹿弘二, 中村和寛, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “HMM 歌声合成における音声データの誤りに頑健なモデル学習手法の検討,” 日本音響学会春季研究発表会, 3-6-7, pp. 335-336, 2014 年 3 月 10-12 日.
- [71] 有竹貴士, 中村和寛, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “HMM 音声合成における LSP に関連した特徴量表現の検討,” 日本音響学会春季研究発表会, 3-6-8, pp. 337-338, 2014 年 3 月 10-12 日.
- [72] 中村和寛, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “低周波数標準化音声データの広帯域成分復元を考慮したメルケプストラム分析の検討,” 日本音響学会春季研究発表会, 3-6-9, pp. 339-340, 2014 年 3 月 10-12 日. (学生優秀発表賞受賞)
- [73] 大浦圭一郎, 橋本佳, 南角吉彦, 徳田恵一, “状態レベルのコンテキストを用いた HMM 音声合成の検討,” 日本音響学会春季研究発表会, 3-6-10, pp. 341-342, 2014 年 3 月 10-12 日.
- [74] 中島陽祐, 李晃伸, “クラウドソーシングを用いたインタラクティブな音声対話システムのための大規模主観評価プラットフォームの構築,” 情報処理学会音声言語情報処理研究会 (SIG-SLP 第 100 回記念シンポジウム), vol. 2014-SLP-100, no. 4, 2014 年 1 月 31 日.
- [75] 柘植彬史, 橋本佳, 南角吉彦, 徳田恵一, “変分ベイズ法による特徴量生成を用いた対数線形モデルに基づく話者認識,” 電子情報通信学会技術研究報告, vol. 113, no. 404, pp. 13-18, 2014 年 1 月 23 日.
- [76] 吉田真基, 打矢隆弘, 山本大介, 内匠逸, “音声対話を用いた学内見学支援システムの試作,” 第 21 回 マルチメディア通信と分散処理ワークショップ, 2013 年 12 月 6 日.
- [77] 柘植彬史, 橋本佳, 南角吉彦, 徳田恵一, “ベイズ統計量を用いた対数線形モデルに基づく話者認識,” 日本音響学会秋季研究発表会, 2-8-7, pp. 73-74, 2013 年 9 月 25-27 日.
- [78] 中村和寛, 大浦圭一郎, 南角吉彦, 徳田恵一, “HMM に基づく英語歌声合成の検討,” 日本音響学会秋季研究発表会, 2-7-7, pp. 299-300, 2013 年 9 月 25-27 日.
- [79] 吉田真基, 打矢隆弘, 内匠逸, “音声対話を用いた学内見学支援アプリケーションの試作,” 平成 25 年度 電気関係学会 東海支部連合大会, 2013 年 9 月 25 日.
- [80] 山本康平, 打矢隆弘, 内匠逸, “音声対話システムにおける関連ワード推薦システムの提案,” 平成 25 年度 電気関係学会 東海支部連合大会, 2013 年 9 月 25 日.
- [81] 山本大介, 西村良太, 打矢隆弘, 内匠逸, “スマートメイちゃん: スマートフォンの機能を有

- するモバイル音声対話エージェント,” DICOMO 2013, pp. 300-304, 2013 年 7 月 10 日.
- [82] 西村良太, 山本大介, 打矢隆弘, 内匠逸, “音声対話エージェントのための Web ブラウザを用いたシナリオエディタの開発,” 情報処理学会 DICOMO2013, pp. 1796-1799, 2013 年 7 月 12 日.
 - [83] 大浦圭一郎, 南角吉彦, 徳田恵一, “Sinsy -隠れマルコフモデルに基づく歌声合成の現状と今後-,” 2013 Information Processing Society of Japan, vol. 2013-MUS-99, no. 52, pp. 1-3, 2013 年 3 月 11-12 日.
 - [84] 吉村建慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “HMM 音声合成のためのバイリンガルデータを用いた因子分析に基づくクロスリンガル話者適応,” 日本音響学会春季研究発表会, 1-7-3, pp. 267-268, 東京, 2013 年 3 月 13-15 日.
 - [85] Viviane de Franca Oliveira, Sayaka Shiota, Kei Hashimoto, Yoshihiko Nankaku, Keiichi Tokuda, “Cross-lingual speaker adaptation for HMM-based speech synthesis using joint-eigenvoices with a space of perceptual characteristics,” 日本音響学会春季研究発表会, 1-7-4, pp. 269-270, 東京, 2013 年 3 月 13-15 日.
 - [86] 中村和寛, 橋本佳, 南角吉彦, 徳田恵一, “HMM 音声合成におけるメルケプストラム分析を統合した音響モデリングの検討,” 日本音響学会春季研究発表会, 1-7-11, pp. 289-290, 東京, 2013 年 3 月 13-15 日.
 - [87] 桑子修一, 高木信二, 橋本佳, 南角吉彦, 徳田恵一, “HMM 音声合成における因子分析を用いた発話適応学習の検討,” 日本音響学会春季研究発表会, 1-7-12, pp. 291-292, 東京, 2013 年 3 月 13-15 日.
 - [88] 岡田知大, 高木信二, 南角吉彦, 徳田恵一, “HMM 音声合成における F0/有声・無声重みの個別クラスタリングによる合成音声の評価,” 日本音響学会春季研究発表会, 1-7-13, pp. 293-294, 東京, 2013 年 3 月 13-15 日.
 - [89] 喜多村翔斗, 中村和寛, 橋本佳, 南角吉彦, 徳田恵一, “音高正規化学習を用いた HMM 歌声合成における調子外れの補正,” 日本音響学会春季研究発表会, 2-7-10, pp. 337-338, 東京, 2013 年 3 月 13-15 日.
 - [90] 城田佳菜子, 中村和寛, 大浦圭一郎, 南角吉彦, 徳田恵一, “HMM 歌声合成における歌唱者適応学習の検討,” 日本音響学会春季研究発表会, 2-7-11, pp. 339-340, 東京, 2013 年 3 月 13-15 日.
 - [91] 打矢隆弘, 伊藤翔太, 内匠逸, 木下哲男, “発展型エージェントシステムにおける動作特性指標,” 第 75 回情報処理学会全国大会, 宮城, 2013 年 3 月 8 日.
 - [92] 岸上友樹, 打矢隆弘, 内匠逸, 木下哲男, “エージェントフレームワーク DASH におけるスナップショット管理機構の構築,” 第 75 回情報処理学会全国大, 宮城, 2013 年 3 月 7 日.
 - [93] 仲野良佑, 打矢隆弘, 内匠逸, 木下哲男, “マルチエージェントシステムを用いた節電機構に関する研究,” 第 75 回情報処理学会全国大会講演論文集, 宮城, 2013 年 3 月 8 日.
 - [94] 田川達司, 内匠逸, 打矢隆弘, “カルマンフィルタとマップマッチングを用いた歩行者経路の推定精度向上に関する研究,” 第 75 回情報処理学会全国大会, 宮城, 2013 年 3 月 6 日.
 - [95] 石原佑哉, 内匠逸, 打矢隆弘, “端末特性を考慮したベイズの定理による 5GHz 帯無線 LAN 位置推定に関する研究,” 第 75 回情報処理学会全国大会, 宮城, 2013 年 3 月 6 日.
 - [96] 尾前貴則, 山本大介, 高橋直久, “音声対話を用いた構内道案内システムの実現,” DEIM 2013, D1-4, 福島, 2013 年 3 月 3 日.
 - [97] 山本大介, 大浦圭一郎, 西村良太, 打矢隆弘, 内匠逸, 李晃伸, 徳田恵一, “スマートフォン単体で動作する音声対話 3D エージェント「スマートメイちゃん」の開発,” インタラクシオン 2013, IPSJ Symposium Series Vol. 2013, No. 1, pp. 675-680, 東京, 2013 年 3 月 2 日.
 - [98] 打矢隆弘, 山本大介, 柴川元宏, 吉田真基, 西村良太, 内匠逸, 松尾啓志, “オープンキャンパスにおけるモバイル端末向け音声対話システムの活用,” 大学 ICT 推進協議会, 神戸, 2012 年 12 月 17 日.
 - [99] 打矢隆弘, 山本大介, 柴川元宏, 吉田真基, 西村良太, 内匠逸, “ビデオ通話型音声対話

- サービス「モバイルメイちゃん」の開発,” 合同エージェントワークショップ & シンポジウム 2012, 静岡, 2012 年 10 月 24-26 日.
- [100] 立松直也, 打矢隆弘, 内匠逸, 木下哲男, “ユビキタス環境に向けたリポジトリ型マルチエージェントフレームワークの拡張,” 合同エージェントワークショップ & シンポジウム 2012, 静岡, 2012 年 10 月 24-26 日.
 - [101] 板津呂翔, 打矢隆弘, 内匠逸, 木下哲男, “知的マルチエージェントシステムにおける強化学習エージェント設計支援機構,” 合同エージェントワークショップ & シンポジウム 2012, 静岡, 2012 年 10 月 24-26 日.
 - [102] 服部貴文, 橋本佳, 南角吉彦, 徳田恵一, “複数のモデル構造を用いた変分ベイズ法に基づく話者認識,” 日本音響学会秋季研究発表会, 1-1-13, pp. 39-40, 長野, 2012 年 9 月.
 - [103] 真野翔平, 南角吉彦, 徳田恵一, “隠れ条件付確率場を用いた音声認識のためのアンローリングに基づく学習アルゴリズムの検討,” 日本音響学会秋季研究発表会, 2-1-10, pp. 79-80, 長野, 2012 年 9 月 19-21 日. (学生優秀発表賞受賞)
 - [104] 大浦圭一郎, 南角吉彦, 徳田恵一, “HMM 歌声合成における状態数可変のモデル構造の検討,” 日本音響学会春季研究発表会, 2-2-2, pp. 275-276, 長野, 2012 年 9 月 19-21 日.
 - [105] 高木信二, 南角吉彦, 徳田恵一, “HMM 音声合成のための加算モデル構造に基づく複数話者同時モデリングの検討,” 日本音響学会春季研究発表会, 2-2-5, pp. 281-282, 長野, 2012 年 9 月 19-21 日.
 - [106] 吉田真基, 打矢隆弘, 内匠逸, “スマートフォンを用いた学内見学支援アプリケーションの試作,” マルチメディア, 分散, 協調とモバイル(DICOMO2012)シンポジウム, 石川, 2012 年 7 月 6 日.
 - [107] 柴川元宏, 打矢隆弘, 内匠逸, 木下哲男, “エージェント技術を用いた分散バックアップシステムの開発,” マルチメディア, 分散, 協調とモバイル(DICOMO2012)シンポジウム, 石川, 2012 年 7 月 5 日.
 - [108] 打矢隆弘, 奥村文雄, 内匠逸, 木下哲男, “リポジトリ型エージェントフレームワークにおける協調エージェント開発支援機構,” マルチメディア, 分散, 協調とモバイル(DICOMO2012)シンポジウム, 石川, 2012 年 7 月 4 日.
 - [109] 高木信二, 南角吉彦, 徳田恵一, “コンテキストの部分的な加算構造に基づく HMM 音声合成,” 日本音響学会 2012 年春季研究発表会, 1-11-5, pp. 303-304, 2012 年 3 月 13-15 日.
 - [110] 橋本佳, 山岸順一, Peter Bell, Simon King, Steve Renals, 徳田恵一, “HMM 音声合成における変分ベイズ法に基づく線形回帰,” 日本音響学会 2012 年春季研究発表会, 3-11-12, pp. 403-404, 2012 年 3 月 13-15 日.
 - [111] Viviane de F. Oliveira, Sayaka Shiota, Yoshihiko Nankaku, Keiichi Tokuda, “Cross-lingual speaker adaptation for HMM-based speech synthesis using speaker interpolation based on perceptual characteristics,” 日本音響学会 2012 年春季研究発表会, 3-11-12, pp. 405-406, 2012 年 3 月 13-15 日.
 - [112] 石原佑哉, 内匠逸, 打矢隆弘, “端末特性を考慮したベイズの定理による無線 LAN 位置推定,” 情報処理学会第 74 回全国大会, 名古屋, 4E-1, 2012 年 3 月 7 日.
 - [113] 細井純太, 山本大介, 高橋直久, “携帯端末を用いた Focus+Glue+Context マップ操作手法,” 情報処理学会第 74 回全国大会, 名古屋, 1P-4, 2012 年 3 月 6 日.
 - [114] 加藤史也, 山本大介, 高橋直久, “Focus+Glue+Context マップのための 凹型多角形の Focus 生成手法,” 情報処理学会第 74 回全国大会, 名古屋, 1P-5, 2012 年 3 月 6 日.
 - [115] 伊藤広記, 山本大介, 高橋直久, “複数メッシュにまたがる周回道路データベース構築手法,” 情報処理学会第 74 回全国大会, 名古屋, 1P-6, 2012 年 3 月 6 日.
 - [116] 樋川貴一, 山本大介, 高橋直久, “XML を用いた WEB サービス用施設領域 DB サーバの構築,” 情報処理学会第 74 回全国大会, 名古屋, 1P-7, 2012 年 3 月 6 日.
 - [117] 原田麻紀, 山本大介, 高橋直久, “道や角の分かりやすさを考慮した構内道案内システム,”

- 情報処理学会第 74 回全国大会, 名古屋, 1P-8, 2012 年 3 月 6 日.
- [118] 小池克拓, 打矢隆弘, 内匠逸, “社会性を組み込んだ情報推薦システムの設計,” 情報処理学会第 74 回全国大会, 名古屋, 1C-2, 2012 年 3 月 6 日.
 - [119] 山本康平, 打矢隆弘, 内匠逸, “アルゴリズム切替による情報推薦システム,” 情報処理学会第 74 回全国大会, 名古屋, 1C-3, 2012 年 3 月 6 日.
 - [120] 吉田真基, 打矢隆弘, 内匠逸, “スマートフォンを用いた学内見学支援アプリケーション,” 情報処理学会第 74 回全国大会, 名古屋, 2C-6, 2012 年 3 月 6 日.
 - [121] 城山翔平, 打矢隆弘, 内匠逸, “AndroidOS 搭載端末向けエージェントプログラミング支援機構の開発,” 情報処理学会第 74 回全国大会, 名古屋, 1D-2, 2012 年 3 月 6 日.
 - [122] 秋元悟史, 打矢隆弘, 内匠逸, “大規模処理機構との連携を考慮した DASH エージェントの設計,” 情報処理学会第 74 回全国大会, 名古屋, 1D-3, 2012 年 3 月 6 日.
 - [123] 徳田恵一, 大浦圭一郎, “自動学習により人間のよう歌う音声合成システム -Sinsy-,” 第 49 回音楽情報科学研究会 (SIGMUS), Technical Report of IPSJ, vol. 2012-MUS-94, no. 1, pp. 1-6, 静岡, 2012 年 2 月.
 - [124] 才野慶二郎, 大浦圭一郎, 橘誠, 剣持秀紀, 徳田恵一, “ラップスタイル歌声合成の検討,” 第 49 回音楽情報科学研究会 (SIGMUS), Technical Report of IPSJ, vol. 2012-MUS-94, no. 1, pp. 1-6, 静岡, 2012 年 2 月 3-4 日.
 - [125] 大浦圭一郎, 間瀬絢美, 南角吉彦, 徳田恵一, “HMM 歌声合成における音高正規化学習の検討,” 第 49 回音楽情報科学研究会 (SIGMUS), Technical Report of IPSJ, vol. 2012-MUS-94, no. 1, pp. 1-6, 静岡, 2012 年 2 月 3-4 日.
 - [126] 中島悠太, 打矢隆弘, 内匠逸, 木下哲男, “エージェントリポジトリを用いた異種エージェント組織化手法,” ネットワークシステム研究会, 信学技報 NS2011-164, pp. 109-114, 沖縄, 2012 年 1 月 27 日.
 - [127] 金子智哉, 打矢隆弘, 内匠逸, 木下哲男, “発展型エージェントシステムのための監視アーキテクチャの実装,” ネットワークシステム研究会, 信学技報 NS2011-165, pp. 115-120, 沖縄, 2012 年 1 月 27 日.
 - [128] 奥村文雄, 打矢隆弘, 内匠逸, 木下哲男, “エージェント相互運用のための開発支援機構の設計,” ネットワークシステム研究会, 信学技報, NS2011-166, pp. 121-126, 沖縄, 2012 年 1 月 27 日.
 - [129] 李晃伸, 大浦圭一郎, 徳田恵一, “魅力ある音声インタラクションシステムを構築するためのオープンソースツールキット MMDAgent,” 音声言語シンポジウム, Technical Report of IEICE, vol. 2011-SLP-89, no. 27, pp. 1-6, 2011 年 12 月 20 日.
 - [130] 山本大介, 大浦圭一郎, 李晃伸, 打矢隆弘, 内匠逸, 徳田恵一, 松尾啓志, “双方向音声デジタルサイネージのための学内イベント登録システム,” 大学 ITC 推進協議会 2011 年度年次大会, CD-ROM, 2011 年 12 月 7-9 日.

③ ポスター発表 (国内会議 11 件, 国際会議 4 件)

【国際】

- [1] Siva Reddy Gangireddy, Pawel Swietojanski, Peter Bell and Steve Renals, “Unsupervised Adaptation of Recurrent Neural Network Language Models,” UK Speech Conference, Sheffield, 20-21 June 2016.
- [2] Ryuhei Kani, Takahiro Uchiya, Daisuke Yamamoto, Ryota Nishimura Ichi Takumi, “Proposal of the Communication Mechanism among Voice-Interactive Digital Signage Systems,” Proc. of 2016 IEEE 5th Global Conference on Consumer Electronics, 2016. (accepted as poster)
- [3] Oliver Watts, Siva Reddy Gangireddy, Juichi Yamagishi, Simon King, Steve Renals, Adriana Stan, Mircea Giurgiu, “Neural net word representations for phrase-break prediction without a part of speech tagger,” UKSpeech, Edinburgh, UK, June, 2014.
- [4] Rasmus Dall, Maria Astrinaki, Alexis Moinet, Junichi Yamagishi, Simon King,

Nicolas d'Alesandro, "Real-time Incremental Speech Synthesis Using Mage," UKSpeech, Cambridge, UK, September 2013.

【国内】

- [5] 後藤良介, 李晃伸, “ニューラルネットワーク言語モデルを用いた2パス型音声認識デコーダの実装,” 日本音響学会 2017 年春季研究発表会講演論文集, 1-Q-10, 2017 年 3 月.
- [6] 村上拓也, 李晃伸, “音声対話システムからの話しかけによる対話性認知の獲得 一話しかけ内容および心理特性との関連—,” 日本音響学会 2017 年春季研究発表会講演論文集, 2-P-7, 2017 年 3 月.
- [7] 浅井千明, 沢田慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “オーディオブックを用いた表現豊かな音声合成のための言語特徴の検討,” 音声研究会, 2017 年 1 月 21 日.
- [8] Juvela Lauri, Hirokazu Kameoka, Junichi Yamagishi, “Majorisation-minimization based composite autoregressive system optimization with a glottal source model prior,” 音声研究会, 2016 年 3 月 28 日.
- [9] 吉村建慶, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “クラウドソーシングによる音声収集のための収録ソフトウェアの設計,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [10] 橋本佳, 山岸順一, 越前功, “話者照合性能を低下させる UBM に基づくプライバシー保護音の検討,” 日本音響学会 2016 年春季研究発表会, 2016 年 3 月 9-11 日.
- [11] 夏目龍司, 李晃伸, “音声対話システムにおける音環境への反応表出による アフォーダンスの評価,” HAI シンポジウム 2015, 2015 年 12 月 5 日.
- [12] 小中彩貴, 李晃伸, “音声対話インタフェースにおける マルチタスク性の適切な表出方法の検討,” HAI シンポジウム 2015, 2015 年 12 月 5 日.
- [13] 小林大晃, 伊藤直晃, 李晃伸, “大語彙連続音声認識における単語信頼度に基づく単語固有ノードの枝刈り手法の検討,” 日本音響学会春季研究発表会, 2-Q4-15, pp.149-150, 2014 年 3 月 10-12 日.
- [14] 船谷内泰斗, 大浦圭一郎, 南角吉彦, 李晃伸, 徳田恵一, “ユーザ生成型音声対話コンテンツに向けた有限状態トランスデューサに基づく簡潔な対話記述法の検討,” 日本音響学会秋季研究発表会, 2-P-28, pp.223-224, 2013 年 9 月 25-27 日.
- [15] Robert A. J. Clark, Magdalena Anna Konkiewicz, Maria Astrinak, Junichi Yamagishi, “Reactive control of expressive speech synthesis using Kinect skeleton tracking,” IPSJ SIG Technical Report, 2012-SLP-94, vol. 30, pp.1-4, 2012 年 12 月 21 日.

(4) 知財出願

① 国内出願 (0 件)

② 海外出願 (0 件)

③ その他の知的財産権

- [1] 音声対話システム構築ツールキット MMDAgent
<http://www.mmdagent.jp/>
- [2] 汎用大語彙連続音声認識エンジン Julius
<http://julius.sourceforge.jp/>
- [3] HMM 音声合成ツールキット HTS
<http://hts.sp.nitech.ac.jp/>
- [4] 音声信号処理ツールキット SPTK
<http://sp-tk.sourceforge.net/>
- [5] HMM 音声合成エンジン hts_engine API

- <http://hts-engine.sourceforge.net/>
- [6] 日本語テキスト音声合成システム Open JTalk
<http://open-jtalk.sourceforge.net/>
- [7] 英語テキスト音声合成システム Flite+hts_engine
<http://hts-engine.sourceforge.net/>
- [8] 歌声合成システム Sinsy
<http://sinsy.sourceforge.net/>
- [9] 英語音声合成システム IDLAK
<http://sourceforge.net/p/kaldi/code/HEAD/tree/sandbox/idlak/>
- (5) 受賞・報道等
- ① 受賞
- [1] 日本音響学会東海支部, 優秀発表賞, 市川祐詞, 2017 年 3 月 13 日.
<http://www.asj-tokai.jp/activities/awardees>
- [2] 日本音響学会東海支部, 優秀発表賞, 梅村真由, 2017 年 3 月 13 日.
<http://www.asj-tokai.jp/activities/awardees>
- [3] * 日本学術振興会 日本学術振興会賞 山岸順一, 統計的音声合成における話者適応とその応用, 2016 年 12 月.
https://www.jsps.go.jp/jsps-prize/ichiran_13th.html
- [4] 一般社団法人情報処理学会 2016 年度 山下記念研究賞 高木信二, 統計的パラメトリック音声合成のための FFT スペクトルからの Deep Auto-encoder に基づく低次元音響特徴量抽出, 2017 年 3 月.
<https://www.ipsj.or.jp/award/yamashita2016.html>
- [5] 日本音響学会 第 13 回学生優秀発表賞, 沢田慶, 2016 年 9 月 15 日.
http://asj.gr.jp/recommending/07_gakusei.html
- [6] 日本音響学会東海支部, 優秀発表賞, 助川正人, 2016 年 3 月 25 日.
<http://www.asj-tokai.jp/activities/awardees>
- [7] * 電子情報通信学会, 業績賞, 徳田恵一, 2015 年 6 月 4 日.
<https://www.ieice.org/jpn/about/rekidai/gyosekishou.html>
- [8] 日本音響学会 第 37 回栗屋潔学術奨励賞, 橋本佳, 2015 年 3 月 17 日.
http://asj.gr.jp/recommending/03_awayu.html
- [9] 日本音響学会 第 10 回学生優秀発表賞, 神谷翔大, 2015 年 3 月 17 日.
http://asj.gr.jp/recommending/07_gakusei.html
- [10] 日本音響学会東海支部, 優秀発表賞, 有竹貴士, 2014 年 12 月 15 日.
- [11] 日本音響学会 第 9 回学生優秀発表賞, 中村和寛, 2014 年 9 月 4 日.
http://asj.gr.jp/recommending/07_gakusei.html
- [12] 優秀プレゼンテーション賞, DICOMO 2014, 柳倫浩, 2014 年 7 月 11 日.
<http://dicomo.org/commendation/>
- [13] ヤングリサーチ賞, DICOMO 2014, 若林敬太郎, 2014 年 7 月 11 日.
<http://dicomo.org/commendation/>
- [14] * 文部科学省 文部科学大臣表彰(科学技術分野) 若手科学者賞, 山岸順一, 2014 年 4 月 10 日.
http://www.mext.go.jp/b_menu/houdou/26/04/_icsFiles/afieldfile/2014/04/07/1346090_02.pdf
- [15] * IEEE Fellow, Institute of Electrical and Electronics Engineers (IEEE), 徳田恵一, 2014 年 1 月 1 日.
<http://www.signalprocessingsociety.org/awards-fellows/fellows-programs/>
- [16] * IEEE Fellow, Institute of Electrical and Electronics Engineers (IEEE), Steve Renals, 2014 年 1 月 1 日.

- <http://www.signalprocessingsociety.org/awards-fellows/fellows-programs/>
- [17] Best Paper Award, IEEE ASRU 2013, Liang Lu, Arnab Ghoshal, Steve Renals, December 2013.
 - [18] * ISCA Fellow, International Speech Communication Association (ISCA), 徳田恵一, 2013 年 8 月 26 日.
<http://www.isca-speech.org/iscaweb/index.php/honors/fellows?id=111>
 - [19] * 喜安記念業績賞, 情報処理学会, 徳田恵一, 戸田智基, 全炳河, 山岸順一, 大浦圭一郎, 2013 年 6 月 5 日.
<https://www.ipsj.or.jp/award/kiyasu.html>
 - [20] 日本音響学会 第 8 回独創研究奨励賞 板倉記念, 大浦圭一郎, 2013 年 3 月 14 日.
http://www.asj.gr.jp/recommending/04_itakura.html
 - [21] 日本音響学会 第 6 回学生優秀発表賞, 真野翔平, 2013 年 3 月 14 日.
http://asj.gr.jp/recommending/07_gakusei.html
 - [22] Best Paper, IEEE SLT2012, Pawel Swietojanski, Arnab Ghoshal, Steve Renals, “Unsupervised cross-lingual knowledge transfer in DNN-based LVCSR,” December 2012.
<http://www.slt2012.org/Papers/PublicSessionIndex3.asp?Sessionid=1007>
 - [23] 日本バーチャルリアリティ学会論文賞, 森井敦士, 山本大介, 舟橋健司, 2012 年 9 月 13 日.
http://www.vrsj.org/awards/best_papers/
 - [24] Best Paper Award in the 11th IEEE International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC'12), Syo Itazuro, Takahiro Uchiya, Ichi Takumi, Tetsuo Kinoshita, 2012 年 8 月 22 日.
 - [25] 電子情報通信学会・日本音響学会 2011 年度音声研究会研究奨励賞, 塩田さやか, 2012 年 6 月 22 日.
<http://www.ieice.org/iss/sp/jpn/>
 - [26] * 文部科学省 文部科学大臣表彰科学技術賞(研究部門), 徳田恵一, 2012 年 4 月 9 日.
http://www.mext.go.jp/b_menu/houdou/24/04/_icsFiles/afieldfile/2012/04/16/1319413_01.pdf
 - [27] 日本音響学会 第 4 回学生優秀発表賞, 間瀬絢美, 2012 年 3 月 14 日.
http://asj.gr.jp/recommending/07_gakusei.html
 - [28] 情報処理学会 2011 年度山下記念研究賞, 大浦圭一郎, 2012 年 3 月 7 日.
<https://www.ipsj.or.jp/award/yamashita2011.html>
 - [29] 日本音響学会東海支部 優秀発表賞, 大野博之, 2011 年 12 月 21 日.
 - [30] 日本音響学会東海支部 優秀発表賞, 彭湘琳, 2011 年 12 月 21 日.

② マスコミ(新聞・TV等)報道

- [1] NHK-FM, トーキング ウィズ 松尾堂「声」を作り上げる, 2017 年 3 月 12 日.
- [2] NHK 総合テレビ, くらし☆解説, 2017 年 3 月 3 日.
- [3] 日経 Linux 3 月号, 秘書の生成「3D の対話エージェントにする」, 2017 年 2 月 25 日.
- [4] 読売新聞 朝刊 31 面, 音声合成 語る亡父～「ドラえもんのように感情豊かに」～, 2017 年 1 月 4 日.
- [5] テレビ愛知, 夕方サテライト 2016 年 12 月 9 日.
- [6] 東海テレビ, みんなのニュース ONE, 2016 年 10 月 25 日.
- [7] CBC テレビ, イッポウ, 2016 年 10 月 24 日.
- [8] 中京テレビ, キャッチ, 2016 年 10 月 21 日.
- [9] 朝日新聞, 朝刊 22 面, 名工大の案内人、3D アニメのメイちゃん, 2016 年 10 月 19 日.
- [10] 日経 Linux 第 11 号, 今! オススメのフリーソフト Sinsy ～Linux が歌姫になる歌声合成～, 2016 年 10 月 15 日.
- [11] NATURE No.7618, Innovation from the frontier of research, 2016 年 9 月 1 日.
- [12] Forbes Japan, 声を失った人が「自分で会話する」方法はあるのか?, 2016 年 8 月 19 日.

- [13] テレビ朝日 報道ステーション SUNDAY, 「あの人を感じたい…」故人を再現 最新技術, 2016 年 7 月 17 日.
- [14] 日刊工業新聞, 投稿も受ける音声対話装置, 2016 年 6 月 15 日.
- [15] 中部経済連合会 Central Japan, Voice recognition and synthesis system in Mei-chan, 2016 年 5 月 26 日.
- [16] 朝日新聞, エジンバラ大学(イギリス)で国際共同研究に参加, 2016 年 4 月 29 日.
- [17] 熊本日日新聞(共同通信取材)(朝刊 13 面), 失った自分の声 再現へ 事前に録音 コンピュータ復元, 2016 年 4 月 13 日.
- [18] ITmedia, 故 hideさんをよみがえらせたのはどの音声合成技術, 2016 年 2 月 15 日.
- [19] 日刊工業新聞, 声を売買「ボイスバンク」, 2016 年 1 月 28 日.
- [20] AERA, 夢をかなえる技術 100, 2016 年 1 月 4 日.
- [21] 山形新聞(共同通信取材), 失った自分の声を再現, 2015 年 11 月 25 日.
- [22] 中部経済新聞, 人間のように歌い話す機械の実現, 2015 年 11 月 24 日.
- [23] 中日新聞, チーム愛忘れない, 2015 年 11 月 24 日.
- [24] 河北新報(共同通信取材), 録音 10 分 失った声を再現, 2015 年 11 月 21 日.
- [25] 山陰中央新報(共同通信取材), 失った声再現, 2015 年 11 月 20 日.
- [26] 福井新聞(共同通信取材), 失った自分の声 再現, 2015 年 11 月 17 日.
- [27] 中部経済新聞(共同通信取材), 失った自分の声を再現, 2015 年 11 月 16 日.
- [28] 四國新聞(共同通信取材), 失った自分の声を復元, 2015 年 11 月 13 日.
- [29] 山梨日日新聞(共同通信取材), 失った「自分の声」再現, 2015 年 11 月 12 日.
- [30] 岐阜新聞(共同通信取材), 失った自分の声 復元, 2015 年 11 月 12 日.
- [31] 大分合同新聞(共同通信取材), 失った自分の声再現, 2015 年 11 月 9 日.
- [32] 中国新聞 朝刊 111 面 (共同通信取材), 失われた声再現 特徴分析『平均』に反映, 2015 年 11 月 8 日.
- [33] 朝日新聞, 広がる理・工学の世界～中部のリーディング大学 理・工学リーダーが語る ものづくりの神髄～, 2015 年 10 月 30 日.
- [34] INTERNET Watch, KDDI 研究所、日本語音声合成ソフト「N2」の Windows/Linux 版 SDK 販売開始, 2015 年 10 月 29 日.
- [35] Yahoo ニュース, KDDI 研究所、日本語音声合成ソフト「N2」の Windows 版/Linux 版 SDK 販売開始, 2015 年 10 月 29 日.
- [36] Yahoo ニュース, KDDI 研究所、日本語音声合成ソフト「N2」の Windows 版/Linux 版を販売, 2015 年 10 月 29 日.
- [37] TBS 夢の扉+, もう一度話したい “そっくり声”作る新技術, 2015 年 10 月 18 日.
- [38] TBS 夢の扉+ YouTube チャンネル, TBS「夢の扉+」10 月 18 日 OA 番組の未放送映像, 2015 年 10 月 18 日.
- [39] another life, 「音」への好奇心を、社会の価値に繋げていく。研究者としての不安の先に見出した可能性。 , 2015 年 10 月 14 日.
- [40] NHK World Radio Japan Online, Radio Japan Focus - Recovering Lost Voices: Pitch-Perfect Voices Created by Computers, 2015 年 9 月 9 日.
- [41] Care Appointments, MND PATIENT GORDON AIKMAN WELCOMES VOICE SYNTHESISERS ANNOUNCEMENT, 2015 年 9 月 2 日.
- [42] Eveningtimes.co.uk, MND campaigner Gordon Aikman welcomes voice synthesisers announcement, 2015 年 9 月 2 日.
- [43] キーマンズネット, 失った声を音声合成で取り戻す「ボイスバンク」, 2015 年 9 月 2 日.
- [44] 三洋化成ニュース 2015 秋 No.492, 音声合成と数学のやさしい関係, 2015 年 9 月 1 日.
- [45] NHK, ほっとイブニング, 2015 年 8 月 28 日.
- [46] NHK NEWS WEB, 最新の音声合成技術でJ2社長の声を再生, 2015 年 8 月 27 日.
- [47] NHK ニュースウォッチ 9, 最新の音声合成技術でJ2社長の声を再生, 2015 年 8 月 26 日.
- [48] AV Watch, 小林幸子の声がなぜボカロに? “こぶしプラグイン”の秘密をヤマハに聞く,

- 2015 年 8 月 18 日.
- [49] ITmedia, 小林幸子さんの歌唱力が移植された「VOCALOID Sachiko」の秘密とは?, 2015 年 7 月 28 日.
 - [50] The Japan News by Yomiuri Shimbun, Synthetic vocal technology: A voice for the speechless, 2015 年 7 月 27 日.
 - [51] 教育情報誌「ドリームナビ」2015 年 9 月号, 夢のかなえかたを大人に聞いてみました, 2015 年 7 月 17 日.
 - [52] NHK NEWS WEB, もう一度自分の声を 音声合成技術, 2015 年 7 月 15 日.
 - [53] 月刊「潮」2015 年 8 月号, 世界を変えるニッポンの革新者たち 連載第 17 回「声の障がい乗り越える音声合成システム。」田原総一朗氏との対談, 2015 年 7 月 5 日.
 - [54] 読売新聞 夕刊, 科学「失った声を取り戻す」, 2015 年 6 月 25 日.
 - [55] 週刊文春 7 月 2 日号, ガン医療はここまで進化した!, 2015 年 6 月 25 日.
 - [56] 日経産業新聞, 先端技術 日本のイノベーター-数式用いて自然な音に, 2015 年 6 月 23 日.
 - [57] FC 岐阜プレスリリース, FC岐阜×日本語ボイスバンクプロジェクト, 2015 年 6 月 23 日.
 - [58] 全国大学生生活協同組合連合会 季刊「読書のいずみ」2015 年夏号, The 著者紹介「おしゃべりなコンピュータ」山岸順一・徳田恵一・戸田智基・みわよしこ, 2015 年 6 月 1 日.
 - [59] 日経サイエンス 7 月号, 丸善ライブラリー「おしゃべりなコンピュータ」, 2015 年 5 月 25 日.
 - [60] 日本経済新聞朝刊, (新書)『おしゃべりなコンピュータ』山岸順一ほか著, 2015 年 5 月 24 日.
 - [61] Jersey Evening Post, £10,000 donation to help motor neurone sufferers 'bank' their voices, 2015 年 5 月 20 日.
 - [62] 朝日小学生新聞, 失った声を取り戻せる「音声合成」の技術 開発進む,
 - [63] 中日新聞 知多版, だし丸くん 声で案内, 2015 年 5 月 11 日.
 - [64] 日本経済新聞電子版および朝刊, 音声合成技術に脚光 東芝は感情表現、香川大は発声ロボ, 2015 年 4 月 28 日.
 - [65] 週刊アスキー 5/12・19 合併号, 今週の気になる書籍「おしゃべりなコンピュータ」, 2015 年 4 月 28 日.
 - [66] 文教速報 第 8137 号, 情報研「おしゃべりなコンピュータ」を刊行, 2015 年 4 月 24 日.
 - [67] ITmedia, JOYSOUND、CeVIO に男性バラード「白咲 優大」、女性ポップ「黄咲 愛里」を追加, 2015 年 4 月 23 日.
 - [68] 文教ニュース, 国立情報学研究所「おしゃべりなコンピュータ」情報研シリーズ19(丸善ライブラリー)を刊行, 2015 年 4 月 20 日.
 - [69] Edinburgh News, Gordon Aikman records voice 'as insurance policy', 2015 年 4 月 19 日.
 - [70] 文教速報 第 8133 号, 最新刊『おしゃべりなコンピュータ』, 2015 年 4 月 15 日.
 - [71] シノドス, 自分の声を取り戻せ—「ボイスバンク」プロジェクト, 2015 年 4 月 9 日.
 - [72] 日刊工業新聞, 奈良先端大など、コミュニケーションを円滑にする訓練システム開発—認知療法を自動化, 2015 年 4 月 8 日.
 - [73] 日刊工業新聞、平成28年6月15日、拓く研究人 124、投稿も受ける音声対話装置、山本大介
 - [74] NHK, リハビリ介護を生きる(2)失った声を再現する, 2014 年 9 月 18 日.
<http://www.nhk.or.jp/heart-net/tv/calendar/2014-09/18.html>
 - [75] 雑誌: NII Today 第 65 号, 2014 年 8 月 27 日.
http://www.nii.ac.jp/userdata/results/pr_data/NII_Today/65/index.html
 - [76] CBC, イッポウ, 2014 年 7 月 29 日.
<http://hicbc.com/tv/ippou/index.htm>
 - [77] 読売新聞(朝刊, 28 面), “病で失った声、合成・再現「ドナー」を募集,” 2014 年 7 月 28 日.
 - [78] 中日新聞(朝刊, 31 面), “音声装置、録音協力者を募集 声のドナー,” 2014 年 7 月 27 日.

- [79] Web ニュース: INSIDE, “「MMDAgent」とプロジェクターで「ダメ男製造機」を三次元に！
自室で『艦これ』の雷と会話できる墮落システム完成,” 2014 年 7 月 8 日
<http://www.inside-games.jp/article/2014/07/08/78425.html>
- [80] The japan Times, “Phrase archive restores lost voices,” 2014 年 5 月 22 日.
<http://www.japantimes.co.jp/news/2014/05/22/national/phrase-archive-restores-lost-voices/#.VCe5BPnCTZE>
- [81] NHK FM ラジオ第一, ラジオあさいちばん(番組内コーナー「聞きときウォッチ」), “自分の声で思い伝える「ボイスバンク・プロジェクト」,” 2014 年 5 月 2 日.
<http://www.nhk.or.jp/r-asa/>
- [82] 会誌「ごきそ」, 2014 年 5-6 月号 24 ページ, “半田で「メイちゃん」が観光案内,” 2014 年 5 月 1 日.
- [83] 朝日新聞(朝刊, 26 面), “機械の声、より人間らしく,” 2014 年 4 月 24 日.
<http://www.asahi.com/articles/DA3S11101174.html> (WEB 版)
- [84] 朝日新聞(朝刊, 12 面), “ボイスバンク、幅広い参加を,” 2014 年 4 月 18 日.
- [85] 民間放送(新聞, 3 面), “音声合成ソフトのキャラクターが DJ,” 2014 年 4 月 13 日.
- [86] TBS ラジオ 土曜朝イチエンタ, 堀尾正明+PLUS! 人権 Today, 2014 年 4 月 12 日.
<http://www.tbs.co.jp/radio/horio/human/20140412.html>
- [87] 朝日新聞(朝刊, 25 面), “失った声 みんなで再現,” 2014 年 4 月 5 日.
- [88] 中日新聞(夕刊, 9 面), “若者呼び込み 秘策あり,” 2014 年 3 月 28 日.
- [89] 関西テレビ, “スーパーニュースアンカー 失った声を取り戻すボイスバンク,” 2014 年 3 月 12 日.
<http://www.ktv.jp/anchor/index.html>
- [90] 市報: はんだ市報, 2014 年 3 月 15 日. (半田市観光協会に設置した音声案内デジタルサイネージが紹介されました)
<http://www.city.handa.lg.jp/contents/66667009.html>
- [91] テレビ: NHK, NHK ニュース, 2014 年 3 月 3 日. (半田市観光協会に設置した音声案内デジタルサイネージが紹介された)
<http://www3.nhk.or.jp/news/>
- [92] 中日新聞(朝刊, 知多版), メイちゃん半田を案内～名工大の音声システム学外に初設置～, 2014 年 2 月 27 日.
- [93] Web ニュース: 毎日新聞, 2014 年 2 月 5 日. (音声合成技術が紹介された)
<http://mainichi.jp/universalon/news/20140205mog00m040009000c.html>
- [94] Web ニュース: ねとらぼ, 2013 年 12 月 26 日. (歌声合成システム Sinsy と音声対話システム MMDAgent が紹介された)
<http://nlab.itmedia.co.jp/nl/articles/1312/26/news095.html>
- [95] 中日新聞(朝刊, 21 面), 新聞わくわくシート, 2013 年 12 月 8 日.
- [96] ラジオ: ZIP-FM, STARTS!, 2013 年 12 月 5 日. (電話インタビュー)
<http://zip-fm.co.jp/program/starts/>
- [97] 東京新聞(夕刊, 6 面), 2013 年 11 月 25 日. (国際音声技術研究所の研究紹介記事)
- [98] 中日新聞(朝刊, 1 面), 2013 年 11 月 18 日. (国際音声技術研究所の研究紹介記事)
http://edu.chunichi.co.jp/?action_kanren_detail=true&action=education&no=3966
- [99] テレビ: フジテレビ, ニュース JAPAN, N の衝撃, “医療の現場へ広がる音声合成技術,” 2013 年 11 月 13 日. (音声合成技術が紹介されました)
http://www.fujitv.co.jp/nj/n_shock_08.html
- [100] テレビ: CBC, イッポウ, 2013 年 11 月 1 日. (音声合成技術が紹介された)
<http://hicbc.com/tv/ippou/index.htm>
- [101] 日本経済新聞(37 面), 人工音声をサポート, 2013 年 10 月 29 日.
- [102] Web ニュース: 日本経済新聞(Web 刊), “カラオケ、人工音声をサポート エクシングが機器に搭載,” 2013 年 10 月 28 日. (音声合成技術が紹介された)
<http://www.nikkei.com/article/DGXNZO61763870Y3A021C1L91000/>

- [103] Web ニュース: ITmedia ニュース, 2013 年 10 月 28 日. (音声合成技術が紹介された)
<http://www.itmedia.co.jp/news/articles/1310/28/news057.html>
- [104] Web ニュース: アキバ総研, 2013 年 10 月 28 日. (音声合成技術が紹介された)
http://akiba-souken.com/article/hobby/18039/?utm_source=twitterfeed&utm_medium=twitter
- [105] テレビ: NHK BS1, Cool Japan, 2013 年 8 月 11 日. (秋葉原に設置した音声案内デジタルサイネージが紹介された)
<http://www.nhk.or.jp/cooljapan/past/130811.html>
- [106] テレビ: フジテレビ, ニュース JAPAN, 2013 年 5 月 2 日. (音声合成技術が紹介された)
http://www.fujitv.co.jp/nj/n_shock_02.html
- [107] Web ニュース: ITmedia ニュース, 2013 年 4 月 26 日. (音声合成技術が紹介された)
<http://www.itmedia.co.jp/news/articles/1304/26/news133.html>
- [108] 朝日新聞(朝刊, 15 面), 名古屋工業大学の羅針盤(広告特集), 2013 年 4 月 22 日.
- [109] 日経新聞夕刊(中部版), “ネットのわくわく形に「天才プログラマー」お墨付き山本大介さん(32 歳),” 2013 年 3 月 16 日.
- [110] 読売新聞, “双方向音声案内デジタルサイネージ”メイちゃん”, 2012 年 10 月 13 日.
- [111] 旺文社 螢雪時代 11 月号, “大学学案内のためのスマートフォン向け音声対話サービス”モバイルメイちゃん”, 2012 年 10 月 13 日.
- [112] 中日新聞, “メイちゃんスマホで会話”, 2012 年 8 月 4 日.
- [113] 中部経済新聞, “3D キャラ大学案内”, 2012 年 8 月 2 日.
- [114] プレスリリース, “世界初! 大学案内のためのスマートフォン向け音声対話サービス～3D キャラクターとどこでも音声インタラクション～”, 2012 年 8 月 1 日.
https://www.sp.nitech.ac.jp/index.php?plugin=attach&pcmd=open&file=20120726_press_release.pdf&refer=%A5%DB%A1%BC%A5%E0%2F%CA%F3%C6%BB
- [115] 名古屋テレビ, “UP!”, 2012 年 7 月 31 日. (双方向音声案内デジタルサイネージが紹介されました)
<http://www.nagoyatv.com/up/special/backnumber.html?key=464f26744c54bd9948093555550b9f62>
- [116] 文教ニュース, “山中文部科学審議官によるメイちゃんの視察”, 2012 年 7 月 16 日.
- [117] テレビ朝日, “モーニングバード”, 2012 年 7 月 2 日. (双方向音声案内デジタルサイネージが紹介された)
<http://www.tv-asahi.co.jp/m-bird/>
- [118] 朝日新聞, “< be report > 広がるデジタルサイネージ”, 2012 年 6 月 30 日.
- [119] 中日新聞, “故人の声でメッセージ”, 2012 年 5 月 2 日.
- [120] NHK WORLD TV, “Science View,” 2012 年 3 月 11 日. (名古屋工業大学の歌声合成技術が紹介された)
<http://www3.nhk.or.jp/nhkworld/english/tv/scienceview/archives20120305.html>
- [121] NHK 総合, “クローズアップ現代「思いが伝わる声を作れ ～初音ミク 歌声の秘密～」,” 2012 年 2 月 28 日. (名古屋工業大学を中心として開発された最新技術として HMM 音声合成技術が紹介された。)
http://cgi4.nhk.or.jp/gendai/kiroku/detail.cgi?content_id=3166
- [122] NHK 教育, “大! 天才テレビくん,” 2012 年 2 月 22 日. (名古屋工業大学の歌声合成技術が紹介された。また、VTR 後半において、運動ニューロン疾患(MND/ALS)などの農神経変性病患者に個人用音声合成システムを提供するエジンバラ大の取り組みが紹介された)
<http://www9.nhk.or.jp/tvkun/>
- [123] 産総研プレスリリース “英語版 PodCastle 用の音声認識器をエジンバラ大が提供,” 2011 年 10 月 12 日.
http://www.aist.go.jp/aist_j/press_release/pr2011/pr20111012/pr20111012.html
<http://podcastle.jp/info/>

(6) 成果展開事例

① 実用化に向けての展開

- ・ 本プロジェクトで得られた技術の実用化例
半田市の観光案内デジタルサイネージ
<http://www.city.handa.lg.jp/contents/66667009.html>
フロンティアファクトリー株式会社のソフトウェア「Band-in-a-Box」
<http://www.pgmusic.com/>
HOYA サービス株式会社のサービス「VoiceText Web API」
<https://cloud.voicetext.jp/>
トヨタ自動車株式会社のサービス「ナニここ」
http://gazoo.com/drive/enjoy/Pages/nanikoko_guide.aspx
ソニー・ミュージックエンタテインメントのスマートフォンアプリ「おとあてきつず」
<http://otoate.hubbub-party.com/kids.html>
株式会社セルシスの「クリップスタジオアクション」
http://www.clip-studio.com/clip_site/download/clipstudioaction/clipstudioaction_top
FM 愛知の「ktr!アニチューン」
<http://fma.co.jp/f/prg/anitune/>
CeVIO プロジェクトの音声案内デジタルサイネージ「CeVIO Vision」
http://cevio.jp/about_cevio/
CeVIO プロジェクトの音声創作ソフト「CeVIO Creative Studio」
<http://cevio.jp/product/ccs/index.html>
株式会社エクシングのカラオケ機器「JOYSOUND」
<http://joysound.com/ex/index.htm>
- ・ 本プロジェクトで得られた技術の実用化に向けたデモンストレーション
ホンダ株式会社の車載システム「メイリン」
<http://www.honda-c86.jp/>
デンソー株式会社の車載エージェント「ハナちゃん」
<http://www.denso.co.jp/ja/news/event/tokiomotorshow/2013/booth/ar/index.html>
- ・ 欧州連合 ROCKIT 事業に採択され、実施中。
- ・ 本プロジェクトで得られた音声関連技術に関する共同研究
ヤマハ株式会社
日産自動車株式会社
富士ゼロックス株式会社
nana music 株式会社
- ・ 開発したプログラムやデータベースの公開
HMM 音声合成ツールキット HTS
<http://hts.sp.nitech.ac.jp/>
汎用第語彙連続音声認識エンジン Julius
<http://julius.sourceforge.jp/>
音声信号処理ツールキット SPTK
<http://sp-tk.sourceforge.net/>
音声合成エンジン hts_engine API
<http://hts-engine.sourceforge.net/>
日本語テキスト音声合成システム Open JTalk
<http://open-jtalk.sourceforge.net/>
音声インタラクションシステム構築ツールキット MMDAgent
<http://mmdagent.sourceforge.net/>
Text-to-Speech System Festival
<http://www.cstr.ed.ac.uk/projects/festival/>

Speech Recognition System Kaldi

<http://kaldi.sourceforge.net/>

CSTR VCTK Courpus

<http://www.udialogue.org/ja/download-ja.html>

② 社会還元的な展開活動

- ・ 本プロジェクトの研究成果を英語および日本語でインターネット (URL: <http://www.udialogue.org/>) で公開し、広く国内外に情報提供している。
- ・ 名古屋工業大学正門前に双方向音声案内サイネージシステムを設置し、イベント等の案内を行っている。
- ・ 本プロジェクトで開発した音声対話システムの使い方に関する大学生向けの講習会を 3 回行い、それぞれ 50 人, 100 人, 40 人の大学生が参加した。
- ・ 人工知能学会主催・第 6 回～第 8 回 AI ツール入門講座において本プロジェクトで開発した音声対話システムの使い方に関する講座を行い、計 70 名以上が参加した。
- ・ 半田市蔵のまち観光案内所に双方向音声案内サイネージシステムを設置し、半田市の観光案内を行っている。
- ・ 半田市蔵のまち観光案内所に設置した双方向音声案内サイネージシステムに関するシンポジウム「半田市地域活性化シンポジウム『双方向音声案内サイネージシステムを利用したまちづくり』～名工大メイちゃん街に出る～」において講演及びパネルディスカッションを行った。
- ・ 音声インタラクション技術と実応用システムに関するシンポジウム「平成 25 年度電気関係学会東海支部連合大会『先進的音声インタラクション技術と実応用システム』」において講演を行った。
- ・ サイエンスアゴラ 2012 に本プロジェクトで開発した「音声対話システム『みんなでメイちゃんをかしこくしよう!!』」を出展した。
- ・ 名工大・名市大合同テクノフェア 2012 に本プロジェクトで開発した音声対話システムを出展した。
- ・ CREST／さががけ合同シンポジウムにおいて、ポスターによる本プロジェクトの紹介を行った。
- ・ 日本バーチャルリアリティ学会「力触覚の提示と計算」第 7 回研究会ラボツアーにおいて本プロジェクトの紹介を行った。
- ・ 第 74 回情報処理学会全国大会において本プロジェクトで開発した音声対話システムの公開実験を行い、121 人が実験に参加した。
- ・ 名古屋工業大学のオープンキャンパスにおいて、音声対話システムと地図を連携した学内案内システムを用いた大学案内を行った。
- ・ シンガポールとの学生交流において、開発した音声対話システムを利用した。
- ・ 高校生向け公開講座で「メイちゃん：声で会話できるコンピュータを作ってみよう」と題する MMDAgent を用いた音声対話システムをつくる実習を行った。49 名が参加した。
- ・ 2016 年 NII オープンハウスにてビットくんの音声対話システムを展示した。また、大学共同利用機関シンポジウム 2016 においても展示した。

(7) その他の重要な成果

特になし

§ 5. 研究期間中の活動

5. 1 主なワークショップ、シンポジウム、アウトリーチ等の活動

年月日	名称	場所	参加人数	概要
2017 年 3 月 23 日	人工知能学会主催・第 11 回 AI ツール入門講 座「音声インタラクション システム構築ツールキット MMDAgent」	電気通信大 学	11 人	3D エージェントとの音声対 話を実現するツールキット MMDAgent のツール入門 講座
2016 年 10 月 1 日	MMDAgent DAY!	名古屋工業 大学		MMDAgent のユーザーミ ーティング
2016 年 9 月 3 日	「音声対話システム作成 のためのクラウドソーシ ング環境」	名古屋工業 大学	---	
2016 年 8 月 10 日	講演会	名古屋工業 大学	40 人	研究者を招いた講演会
2016 年 8 月 5 日	「メイちゃん:声で会話で きるコンピュータを作っ てみよう」	名古屋工業 大学	49 人	高校生向けに, MMDAgent を用いた音声 対話システムを作る演習
2016 年 6 月 2 日	講演会	名古屋工業 大学	62 人	研究者を招いた講演会
2015 年 7 月 6 日	講演会	名古屋工業 大学	約 40 人	研究者を招いた講演会
2015 年 6 月 22 日	人工知能学会主催・第 8 回 AI ツール入門講座 「音声インタラクションシ ステム構築ツールキット MMDAgent」	国立情報学 研究所	25 人	3D エージェントとの音声対 話を実現するツールキット MMDAgent のツール入門 講座
2015 年 3 月 28 日	ものアプリハッカソンス ーパー	大阪イノベ ーションハ ブ	本プロジェ クト内から 3 人参加	JST CREST の技術シーズ をつかったハッカソン
2015 年 2 月 27 日	人工知能学会主催・第 7 回 AI ツール入門講座 「音声インタラクションシ ステム構築ツールキット MMDAgent」	大阪大学	17 人	3D エージェントとの音声対 話を実現するツールキット MMDAgent のツール入門 講座
2014 年 7 月 10 日	人工知能学会主催・第 6 回 AI ツール入門講座 「音声インタラクションシ ステム構築ツールキット MMDAgent」	東京工業大 学	20 人	3D エージェントとの音声対 話を実現するツールキット MMDAgent のツール入門 講座
2014 年 4 月 12 日	半田市地域活性化シンポ ジウム「双方向音声案内 サイネージシステムを利 用したまちづくり」～名工 大メイちゃん街に出る～	アイプラザ 半田	約 40 人	半田市蔵のまち観光案内所 に設置した双方向音声案内 サイネージシステムに関する シンポジウム

2014 年 3 月 24 日	講演会	名古屋工業 大学	30 人	海外の研究者 2 名を招いた 講演会
2013 年 9 月 24 日	平成 25 年度電気関係 学会東海支部連合大会 「先進的音声インタラク ション技術と実応用シス テム」	静岡大学	90 人	音声インタラクション技術と 実応用システムに関するシ ンポジウム
2012 年 12 月 12 日	名工大・名市大合同テ クノフェア 2012	名古屋工業 大学	本プロジェ クト内から 4 人参加 (来場者多 数)	双方向音声案内デジタルサ イネージシステムの展示
2012 年 11 月 9～11 日	サイエンスアゴラ 2012	日本科学未 来館	本プロジェ クト内から 16 人参加 (来場者多 数)	音声対話システム「みんな でメイちゃんをかしこくし よう！！」の展示
2012 年 11 月 8 日	CREST／さがけ合同 シンポジウム	一橋大学	本プロジェ クト内から 12 人参加 (来場者多 数)	ポスターによる本プロジェクト の紹介
2012 年 11 月 7 日	名工大・名市大合同テ クノフェア 2012	JR 名古屋 駅	(来場者多 数)	双方向音声案内デジタルサ イネージシステムの展示
2012 年 4 月 1～2 日	CREST symposium on Human-Harmonaized Information Technology -Behaviour, Interaction, Music, and UGC-	京都大学	約 200 人	CREST4 プロジェクト(徳 田, 後藤, 武田, 河原)によ る共催シンポジウム
2012 年 3 月 6～8 日	モバイルメイちゃん公開 実験(第 74 回情報処理 学会全国大会)	名古屋工業 大学	121 人	会場音声案内「モバイルメイ ちゃん」公開実験
2011 年 11 月 25 日	日本バーチャルリアリテ ィ学会「力触覚の提示と 計算」第 7 回研究会ラボ ツアー	名古屋工業 大学	30 人	ラボツアーでの本プロジェク トに関する紹介

この他にも月 1 回程度ビデオ会議システムによる全グループでのチーム内ミーティングと月 3 回程度
の徳田グループと山本グループによるチーム内ミーティングを行い、定期的に進捗報告を行っ
た。

§ 6. 最後に

【研究の目標等から見た達成度、得られた成果の意義等の自己評価】

本プロジェクト開始以前、研究代表者は、新しいモデルを提案して有効性を評価するといった「インベンション型」の研究スタイルで研究を進めてきたが、開発した技術を一般の人々に広く喜んでも使ってもらうには何かが足りないのではという思いが募り、「ユーザー生成型音声対話コンテンツ」という着想を基本に本プロジェクトを提案した。当初は、少々突拍子もない考えだったのではないのか、雲をつかむような話ではないのか、と不安を感じていたが、公開したソフトウェアツール群により、実に多くの多様なコンテンツが生成されることが観測され、間違った考えではなかったと安堵するとともに、期待以上の成果が得られたと達成感を感じている。Apple 社の Siri などでは、音声対話コンテンツ作成チームが編成されているとことで、更には Pepper 等では、ユーザー自身がコンテンツをつくるという概念が導入されており、本プロジェクトはこのような流れの先駆けとなったと自負している。また、等身大のアニメ調キャラクターとデジタルサイネージあるいは巨大スクリーンなどで音声により対話するというシステムあるいは企画は、本プロジェクトが切っ掛けとなり、ひとつのブームの如くなっている。

その一方で、生成されたコンテンツの分析に関しては、想定と若干異なる展開となった。当初は、アドバイザーからの「こういったユーザー生成型のものは、違法コンテンツ・不適切コンテンツを含め、むしろ初めカオスになることによって大きく動き出す。但し、それに直接、関わるとまずいので、良い距離感を保ちながら進めるのがよいと思う。」という趣旨のアドバイスもあり、こちらでコンテンツをアップロードするためのサーバー等を用意することはしなかった。twitter, blog 等を利用することにより、ユーザーコミュニティと緩やかな連携を保ちつつ進めることができたが、様々な形でインターネット上に散逸したコンテンツの全体像を把握したり、分析したりすることが困難になってきた。このため、中間評価でのアドバイスもあり、若干の方向修正を行った。つまり、大量の音声対話コンテンツに基づいた統計的な分析によるアプローチから、ユーザーのコンテンツ利用履歴に基づく手法へ軸足を移した(4.3.1 に記載)。また、限定された状況での社会実験を行うとともに(4.3.3)、コンテンツ共有サーバー「MMDAgent SHARE」を立ち上げた。音声対話コンテンツをシステムから分離して共有可能であることの実証となっている。

【今後の研究の展開】

MMDAgent は、最新の音声合成、音声認識、3D-CG 表示、対話制御を統合したオープンソースで商用も含めて自由に利用可能なソフトウェアツールキットである。各種モデルファイル等のフォーマットもオープン仕様となっている。このようなソフトウェアは、他に見当たらず、今後、関連研究の研究基盤ソフトウェアとして大変重要な役割を担っていくと考えられる。既にデファクトソフトウェアツールとして研究目的だけでなく、産業応用としても広く利用されているため、本プロジェクト後も、これまで同様、ソフトウェアを保守するだけでなく、最新の関連技術を導入し続け、ソフトウェアをフレッシュな状態にキープしていきたいと考えている。

【研究代表者としてのプロジェクト運営】

本プロジェクト運営の特徴は、グループ毎に独立した課題を担当するのではなく、全グループが緊密に連携しながら各課題(あるいはテーマ)に取り組んだことである。このため、大学内での週次ミーティング、ビデオ会議システムを使った月例全体ミーティング、国際会議参加等の機会を活かした小ミーティング、メーリングリストによる情報共有、skype による個別確認、等を頻繁に行った。コミュニケーションコストが高くなるデメリットがあったが、その分、プロジェクトの目標や考え方を良く共有することができ、心をひとつとした運営ができたと考えている。また、何もなければ、出会わなかったであろう突出して優秀な若手研究者を複数擁し、彼らをコラボレーションさせることにより、これまでにないソフトウェアツールおよび研究成果を創出することができたと考えている。



図 32 大学正門でキャンパス案内をする「正門メイちゃん」

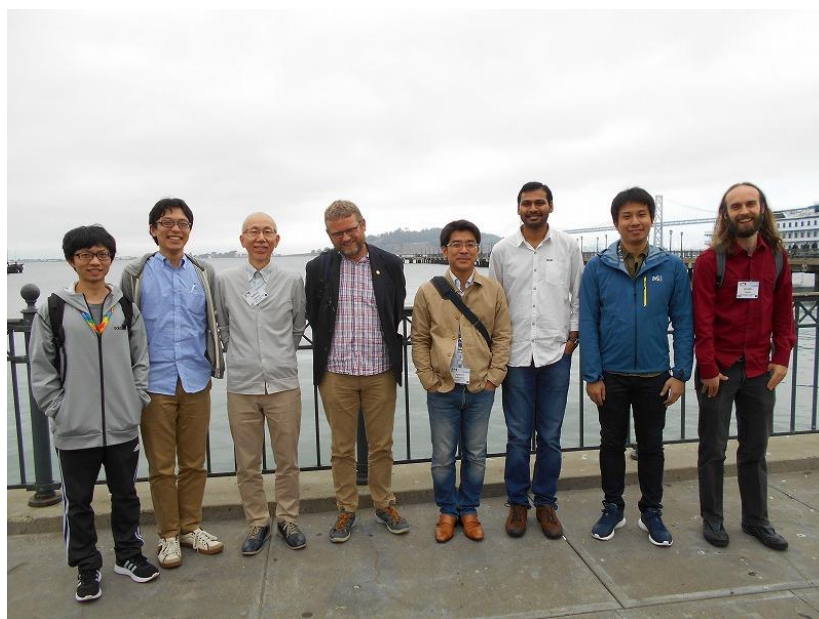


図 33 国際会議 INTERSPEECH 2016 での参加メンバー集合写真



図 34 サイエンスアゴラ出展時の参加メンバー集合写真