

音声コミュニケーションシステムの開発と研究成果の社会還元

■概要

本開発室では、先進的音声翻訳研究開発推進センターの研究成果である音声認識、音声合成、言語翻訳などの技術を利用した各種統合システムを開発して広く世間に周知することにより、研究成果の成果展開と社会還元を進めている。具体的には、多言語音声翻訳システム、聴障者と健聴者とのコミュニケーション支援アプリ等を開発すると共に、それぞれの共通プラットフォーム化を図ることによりスムーズな成果展開に寄与している。さらに、今後の研究課題である同時通訳システムの研究プラットフォームの整備を行っている。

■平成29年度の成果

1. 多言語音声翻訳システムVoiceTraの機能拡張

VoiceTra (<http://voicetra.nict.go.jp/>) の機能拡張について以下に述べる。

(1) iOS*版定型文機能

よく使う文を事前に登録しておくことにより、相手に伝えたいことをリストから選ぶだけで翻訳できる機能をiOS版にも追加した。あらかじめ対訳を登録しておくことで正しい訳文を相手に示すことができるが、対訳が登録されていない場合は翻訳を実行して相手に示すことができる。また、こちらからの質問に対する相手の応答選択肢も登録することができ、相手はそれを指で選択するだけで応答することができる。図1から3に主な画面を示す。この機能は音声認識が困難な騒音下（例えば、搬送中の救急車の中）でも利用できること、相手が音声で応答することができない状況（例えば、呼吸困難状態）でも利用できること、あらかじめ正しい対訳を準備することにより誤訳が出ない（人命に関わる内容でも使える）ことから、全国の消防署の救急隊で活用されている。対訳の登録にWebサーバが必要であるため現時点で個人利用は想定していないが、今回iOS版をリリースしたことで外国人の生徒がいる学校などの教育現場でも活用される事が期待される。

* iOSは、米国およびその他の国におけるCisco社の商標または登録商標です。



図1 定型文一覧画面

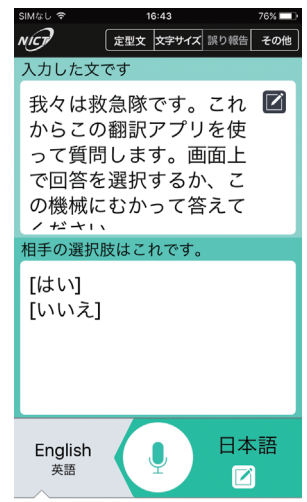


図2 定型文確認画面

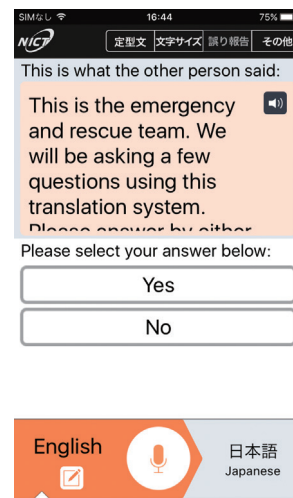


図3 相手に提示する画面

(2) OCRメニュー翻訳機能の改良

行き先案内板やレストランメニューなどの日本語が母国語に翻訳できると、海外からの旅行者にとっては大変便利である。前年度、NICT製のレストランメニュー用翻訳エンジンとパナソニック ソリューションテクノロジー株式会社製のOCR（光学文字認識）エンジンを組み合わせたレストランメニュー用のOCR翻訳機能を試作したが、平成29年度は主にユーザーインターフェースを改良した。スマートフォンでレストランメニューを撮影

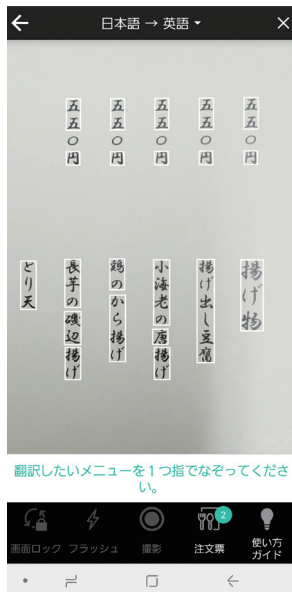


図4 OCRによる単語認識画面

するとOCRエンジンを用いて文字認識を行うが、レストランメニューの単語辞書をあらかじめ登録しておくことで、文字単位ではなく単語単位で認識結果の枠が表示されるようにした(図4)。これによって翻訳したいメニューの指定が楽に行えるようになった。また、平成28年度版では、日本語から英語への翻訳のみであったが、英語を含めて12言語への翻訳ができるようにした。対応言語は以下のとおりである。英語、中国語(簡体字、繁体字)、韓国語、タイ語、インドネシア語、ベトナム語、フランス語、スペイン語、ドイツ語、イタリア語、アラビア語。

(3) 男声・女声切替機能

翻訳結果の合成音声は主に女声のみであったが、男性が話した内容が翻訳されて女性の声で再生されるのは違和感があるため、男声と女声の両方の合成音声が可能となる言語については、オプション指定で自分側及び相手側について、独立に合成音声の性別を指定できるようにした。今後は、入力音声の性別を自動判定して、その結果に応じて合成音声の性別を自動的に切り替える機能に拡張していく予定である。

(4) 画面の回転機能

タブレット端末を傾けることで180度画面が回転するようにした。この機能により、対面で使用する場合に端末自体を回転して相手に見せる必要がなくなりコミュニケーションが円滑に行えるようになった。

(5) 多言語表記対応

VoiceTra(ボイストラ)の表示対応言語は、日本語、



図5 フランス語のVoiceTraチラシ

英語、中国語(簡体字)、韓国語の4カ国語のみに対応していたが、訪日外国人観光客の利用促進を目的として、タイ語、インドネシア語、ベトナム語、フランス語、スペイン語、中国語(繁体字)を追加し、合計10言語の表記に対応した。合わせて日本語を除く9言語のアプリ紹介チラシを作成して国際空港、観光案内所等に配置している(図5)。これらの対応によりVoiceTraを通じて収集される外国語音声データの増加が期待される。

2. 同時通訳プラットフォームの研究開発

2020年度の完成を目標に同時通訳プラットフォームの研究開発を進めている。そのための第一段階として一括音声翻訳方式のプロトタイプを作成した。現在のVoiceTraでは、音声認識、翻訳、逆翻訳、音声合成の4回のリクエストを順次サーバに送って処理結果を受け取っている。一括音声翻訳方式では、入力音声を音声翻訳リクエストとしてサーバに送ると、その1回のリクエストに対してサーバ内で必要な処理が全て実行されて、音声認識結果、翻訳結果、逆翻訳結果、音声合成結果が1個のレスポンスとして順次返却されるため、これまでの4回のリクエストとレスポンスの送受信に比べて反応速度が向上する。ヨーロッパで比較評価した結果、従来方式で約6秒の反応速度が、一括音声翻訳方式では、約2秒となった。また、日本国内でも約2秒から約1.5秒となった。今後、一般公開しているVoiceTraに一括音声翻訳方式を組み込む。さらにこの方式を拡張して連続音声入力に対応させることで、講演音声等の同時通訳が実行できるシステムを開発し、同時通訳の研究プラットフォームとして整備していく予定である。

