

携帯端末を活用した 自動通訳実証実験

花沢 健・長田 誠也・後藤 由希子

要 旨

自動通訳技術は、言語の壁を越えた相互理解を促進する重要な技術です。現在、旅行会話の通訳に関しては実用レベルに達しつつあります。今回、総務省により、自動通訳技術を中心とした産業の発展と、それを活用した地域振興の両方を目的として「地域の観光振興に貢献する自動音声翻訳技術の実証実験」が行われることとなり、NECも同実証実験に参加しています。本稿では、実証実験の概要と、そこで使用されるNECの自動通訳システムについて紹介します。

キーワード

●自動通訳 ●音声認識 ●機械翻訳 ●音声合成 ●ウェアラブル

1. まえがき

NECではこれまで、自動通訳技術の研究開発を行ってきました^{1~3)}。自動通訳技術は、音声認識・機械翻訳・音声合成の各技術を融合させることで、異なる言語間でのコミュニケーションを実現するものです。近年、例えば「旅行会話」のように限定された領域では自動通訳技術は実用レベルに到達しつつあります。

このような状況の下、2009年度総務省により、自動通訳技術を中心とした産業の発展と、それを活用した地域振興との両方を目的として「地域の観光振興に貢献する自動音声翻訳技術の実証実験」が行われることとなり、NECも同実証実験に参加しています。

本実証実験は、自動通訳技術の観光産業での有用性を実証するためのものです。すなわち、実際の外国人観光客との接客場面での自動通訳精度や速度といった技術水準の確認と課題の抽出、ハードウェアやユーザインタフェースなど適切なシステム構成の検証、自動通訳システムを利用した結果のアンケートから、ユーザの満足度の評価やニーズの抽出を行います。

本稿では、実証実験の概要と、観光産業での利用におけるユーザの利便性を考慮してNECが提案する、携帯端末を活用した自動通訳システムについて説明します。

2. 山梨県における実証実験の概要

2.1 背景

2009年度総務省から、自動通訳技術を中心とした産業の発展と、それを活用した地域振興の両方を目的として「地域の観光振興に貢献する自動音声翻訳技術の実証実験」プロジェクトの公募が行われました。実証実験は、全国5カ所（北海道、関東、中部、関西、九州）で同時に実施されます。NECは、株式会社JTBグローバルマーケティング&トラベル（以下、JTB）、独立行政法人情報通信研究機構（以下、NICT）とともに、3者共同で関東地区の実証実験を提案し採択されました。全体統括と現地での実証実験のオペレーションは主にJTBが行い、NECは実証システム全体の開発を担当します。また、NICTとNECはそれぞれ独自の自動通訳技術を端末に搭載します。

2.2 実験計画

関東地区の実証実験は、関東の中でも、欧米、アジアなど各国で人気の高い富士山・富士五湖観光圏を有する、山梨県で行います。山梨県及びびやまなし観光推進機構の協力の下に、山梨県下の複数のエリア、例えば富士山の鑑賞スポットである河口湖や山中湖、石和温泉など外国人観光客に人気の高いエリアにおいて自動通訳システムを配備し、訪日外国人客に対する接客場面において、通訳技術を検証します。以下の場

所に設置します。

- ・ 宿泊施設
- ・ 観光施設
- ・ 飲食業施設

それぞれの設置場所における利用ケースに応じて適切なシステム構成があると考え、以下の3種類のシステムを実証実験で使用します。

(1) 据置型

ホテルのフロントやコンシェルジュデスク、お土産屋のレジ近辺、観光案内デスクなどは場所が固定されておりシステムを移動させる必要がないため、比較的高性能・大画面のPCを据置型として用います（図1参照）。外国人観光客が現地の観光案内担当者や接客スタッフと会話する際、設置されているスタンドマイクで音声入力し、PCの画面を観光客と観光事業者で共有しつつ自動通訳でコミュニケーションを行います。通訳機能は、NICTが提供する日英韓中四カ国語の通訳サーバにネットワーク接続して実現します。

(2) モバイル型

宿泊施設やレストラン、あるいはお土産屋などの店員が通訳システムを利用することを想定した場合、動き回りながら利用可能であることが求められるため、モバイル型が適しています（図2参照）。店員が携帯端末を持ち、外国人観光客との会話を進めます。モバイル型では、NECの保有する日英通訳機能（通信不要）とNICTが提供する通訳サーバ（LANなどの通信環境を利用）による四カ国語通訳機能を1台の端末で状況（通信環境や対応言語）に応じて切り替えて動作させます。



図1 据置型の利用イメージ



図2 モバイル型の利用イメージ



図3 ウェアラブル型の利用イメージ

(3) ウェアラブル型

接客中に通訳システムを利用することを想定した場合、両手がふさがった状態でもシステムを用いてコミュニケーションができること、お互いの顔を見ながら会話できること、によりコミュニケーションの満足度がより高まる可能性があります。このような仮説を検証するため、お互いにウェアラブル端末を身につけた観光客とスタッフが対面で会話するウェアラブル型を用意します（図3参照）。ウェアラブル型については、NECで事業化を進めているTele Scouter^{4,5)}上に、NECの日英通訳機能を搭載します。

今回は、据置型、モバイル型、及びウェアラブル型の3種類の端末で合計300台を用意し、実証実験中できるだけ多くの方に使っていただけるよう環境を整えています。

3. NEC版自動通訳システムの詳細

今回の実証実験においては、モバイル型とウェアラブル型の2種類において、我々が開発した低演算量・省メモリな自動通訳システムを利用します。本通訳システムは、旅行会話を対象とする日英双方向の音声認識・機械翻訳・音声合成機能を統合したものです。

以下、本システムの音声認識・翻訳性能と、モバイル型・ウェアラブル型それぞれのシステムの特徴を説明します。

3.1 音声認識・翻訳性能

NEC版自動通訳システムの大きな特長は、数万単語規模の大語彙連続音声認識と機械翻訳を、市販の携帯電話レベルのリソースで高速に実行可能な点です。特に、通信を介さず端末単体で音声認識・翻訳処理を行うことにより、高速なレスポンスを実現しています。また、音声認識・翻訳機能のコア部分は言語依存性を排しており、言語の切り替えは辞書・モデルの差し替えのみで対応することが可能です。

処理速度については、これら音声認識・翻訳機能を市販の携帯電話相当の端末に搭載して評価し、音声入力終了してから翻訳結果が表示されるまで、1秒程度とほぼリアルタイムで動作することを実際に確認しています。メモリ使用量は音声認識と翻訳あわせて1言語当たり20MBと非常に小さく、コンパクトかつ高速な動作を実現しています。

3.2 モバイル型への搭載

本実証実験において、モバイル型の端末としては、市販のSmartPhone（Windows Mobile端末）を使用します。SmartPhoneの内蔵マイクから音声入力を行い、音声認識結果とその翻訳結果を画面表示するとともに、内蔵スピーカから翻訳結果を合成音声出力します。1台の端末を日本語話者と英語話者が共有するため、端末上で日英方向の通訳と英日方向の通訳とを切り替えて用います。

3.3 ウェアラブル型への搭載

ウェアラブル型では、日英それぞれの話者が1台ずつ端末を装備します。入力話者側のヘッドセットから音声を入力し、

入力話者側の端末で音声認識と機械翻訳を行った後、翻訳結果を相手側の端末へ送信し、相手側端末で音声合成を行って相手側のヘッドセットから合成音声出力します。このとき、音声認識結果とその翻訳結果はお互いのヘッドマウントディスプレイ上で共有します。これを双方向で行うことで、お互いの顔を見た状態での自動通訳を実現します。

4. むすび

自動通訳システムの観光産業における有用性を実証するために、山梨県で行われる総務省の実証実験と、そこで使用されるNEC製自動通訳システムの詳細について説明しました。ここで得られる知見をもとに、今後引き続き自動通訳技術の開発・事業化を推進していく予定です。

*Windows Mobile は、米国 Microsoft Corporation の米国及びその他の国における登録商標または商標です。

参考文献

- 1) 長田、ほか：「携帯電話試作機上で動作する旅行会話向け自動通訳システムの開発」、情報処理学会全国大会、2009年
- 2) 花沢、ほか：「携帯電話試作機上で動作する旅行会話向け音声認識システム」、情報処理学会全国大会、2009年
- 3) 花沢、ほか：「携帯電話試作機上で動作する旅行会話向け音声認識」、日本音響学会全国大会、2009年
- 4) NECプレスリリース：「世界初、メガネ型網膜走査ディスプレイとウェアラブルコンピュータ端末を活用したユビキタス業務支援システム「Tele Scouter」を発売」、2009年
- 5) メガネ型網膜走査ディスプレイとウェアラブルコンピュータ端末を活用したユビキタス業務支援システム「Tele Scouter」
<http://www.nec.co.jp/solution/telescouter/index.html>

執筆者プロフィール

花沢 健
共通基盤ソフトウェア研究所
主任
情報処理学会
人工知能学会
日本音響学会各会員

長田 誠也
共通基盤ソフトウェア研究所
主任
情報処理学会
言語処理学会各会員

後藤 由希子
市場開発推進本部
主任