

音声翻訳や音声言語対話による 情報アクセスについての研究

住田 一男 研究室

● 専門分野：自然言語処理、情報検索・フィルタリング



研究目的

人が話す音声や人が文章として書き記す自然言語文は、人同士のコミュニケーションや人が他の人に自分の持っている情報や知識を伝えるための最も自然な手段です。インターネットの普及に伴い、メールやブログ、SNS（Social Network Service）による情報のやり取りが加速度的に増え、パソコンやスマートフォンの利用者の増大に伴い、これら端末を介した音声や自然言語によるコミュニケーションや情報アクセスの機会も同様に増えてきています。本研究室では、音声や自然言語を対象として、人と人との間のコミュニケーションや、人とコンピュータシステムとの間の情報のやり取りを支援する技術を研究しています。音声や自然言語という一般的な手段によって、人のコミュニケーション力や知的生産性などの能力を飛躍的に高めることが目的です。

研究テーマ

1. 音声翻訳についての研究

社会のグローバル化につれて、個人が海外を旅行する機会が増えるとともに、海外とのビジネス機会も増加しています。利用者が話した音声を他の言語に翻訳する音声翻訳技術を研究開発しています。音声翻訳は、利用者が話した音声を文字化する音声認識技術、音声認識結果を他の言語に翻訳する機械翻訳技術、翻訳結果を音声で読み上げる音声合成技術といった3つの主要な要素技術から構成されます。

図1は、旅行での利用を想定し試作した、電波の届かない環境でも動作可能なスマートフォン単体で動作する音声翻訳アプリケーションです。耐雑音性を強化した音声認識、大量の対訳コーパスに基づく統計翻訳、高音質な音声合成が、計算リソースの小さい環境で動作することが技術的な特徴です。日英中の3言語に関してすべての言語方向での動作を実現しており、利用者が話す1文ないしは2文程度の短い朗読調の音声を対象に音声翻訳を行います。実証実験を通じて、旅行で必要となるコミュニケーションに役立つことを実証しました。

現在実現している技術レベルは、旅行会話のような比較的簡単なコミュニケーションを支援するものですが、今後は、利用者が自由に話す自由発話を対象とした話し言葉同時通訳を、業務や会議など専門性の高い内容で実現することが課題です。

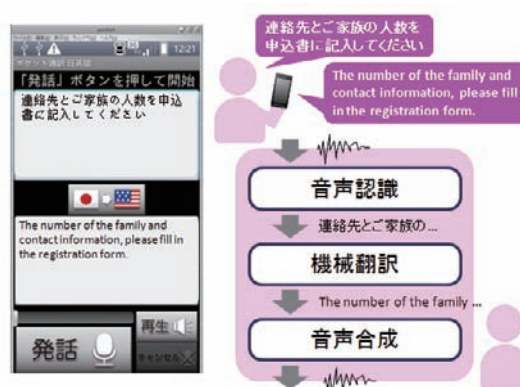


図1 スマートフォン単体で動作する音声翻訳ソフト

2. 音声対話に基づく情報アクセス

テレビ放送の多チャンネル化に伴い、録画した大量のテレビ番組や、VOD (Video On Demand) で提供される大量のビデオコンテンツ等、従来に比べて利用者がテレビで視聴可能な映像コンテンツが増加しています。このような大量の映像コンテンツから利用者が所望のものを検索する場合、現状では、コンテンツプロバイダが事前に作成したディレクトリをたどって探す方法や、テレビ画面に表示されたソフトキーボードからリモコンを使って一文字ずつ文字を選択し入力したキーワードや、外付けしたキーボードからキー入力したキーワードに基づいてキーワード検索する方法が一般的です。音声による検索や指示を可能にすることにより、より簡単に利用者が所望のコンテンツを検索できる可能性があります。

図2は、テレビ番組を対象に試作した音声対話システムです。録画したコンテンツの検索や録画の予約を利用者が音声入力で指示することができます。放送局から電子的に得られる電子番組表には、各番組の放送時間や、番組名、出演者名等の情報が含まれています。しかし、利用者が入力する音声に含まれるキーワードとは一致しない場合が多く、このような場合に対応するため、出演者等の愛称、番組名等の略称言い換えのための言語知識を、電子番組表やWWW (World Wide Web) から自動的に抽出し、音声認識や発話意図の解釈の処理に反映させることが技術的な特徴です。ここで示したシステムは一例ですが、音声対話による情報アクセスは様々な分野への適用が考えられます。

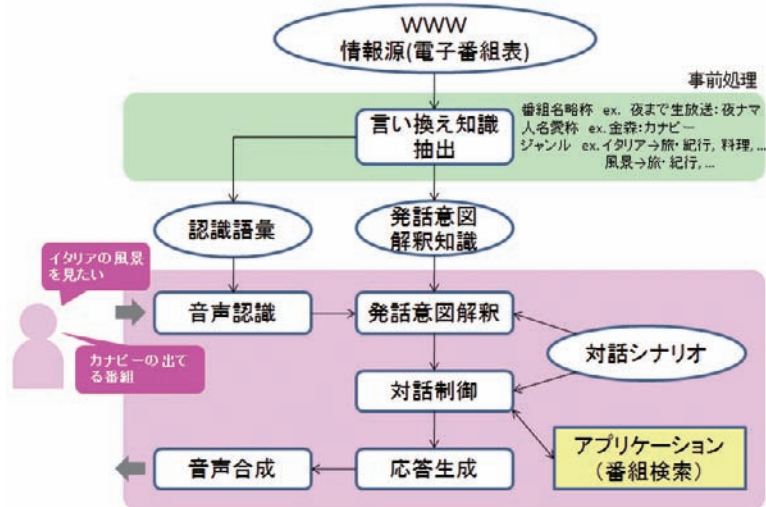


図2 音声対話に基づくテレビ番組検索

● 教員からのメッセージ

教員が所属する(株)東芝 研究開発センター 知識メディアラボラトリーでは、「ことば」を解析・理解し、知識として活用する技術の研究開発を行っています。具体的には、上の例で示した研究以外に、雑音に強い音声認識技術、多様な声を作り出すことができる音声合成技術、膨大な情報の中からユーザが求める情報を簡単に発見することができる情報検索技術、膨大な文章データから言語資源や知識データを抽出するマイニング技術などの研究を行っています。

音声や自然言語に基づく人と人とのコミュニケーション、人とコンピュータシステムとのインタラクション、WWW等の大量の文書データからの言語知識資源や知識データのマイニングに関心のある学生を歓迎します。

● 参考文献

1. 大内一成, 若木裕美, 屋野武秀, 住田一男, 土井美和子, “人名と番組名の言い換えに対応する音声認識インタフェース”, 情報処理学会論文誌, 51, 3, pp.846-855 (2010)
2. H. Wakaki, H. Fujii, M. Suzuki, M. Fukui, K. Sumita, “Abbreviation Generation for Japanese Multi-Word Expressions,” Proc. the 2009 Workshop on Multiword Expressions, ACL-IJCNLP 2009, pp.63-70 (2009)
3. 知野哲朗, 釜谷聡史, 降幡建太郎, 住田一男, “日中英3言語6方向音声翻訳システム,” 情報処理学会研究会資料, NL, 2008(46), pp.15-22, (2008)