

### 3.4.4 知識創成コミュニケーション研究センター 音声コミュニケーショングループ

グループリーダー 河井 恒 ほか 33 名

#### ナチュラル言語コミュニケーション技術に関する研究開発

##### 【概要】

誰が、いつ、どこで、どのような表現で、何語で話そうとも、音声や身振り・手振りなどの人間にとって自然な言語・非言語表現によって情報を補いながら、息の合ったコミュニケーションを実現するナチュラル言語コミュニケーションの構成技術の研究開発を推進している。本年度は、研究の基盤となる対話コーパスの整備、非言語情報を利用するための画像処理技術の開発、音声認識・合成技術の改良および多言語化など、音声対話システムを構成する要素技術の開発を進めた。さらに、これらの技術を統合して対話処理のプロトタイプシステムを開発した。また、言語翻訳グループと共同で多言語音声翻訳システムの開発ならびに国際標準化活動を推進した。

##### 【平成 22 年度の成果】

#### スマートフォン用多言語音声翻訳アプリケーション “VoiceTra” を全世界で無料公開

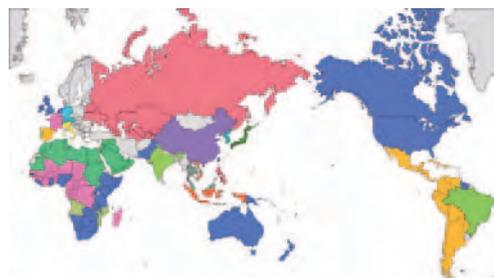
これまで研究開発を行ってきた旅行会話の音声翻訳に関する成果を社会に還元し、同時に利用時に蓄積されるログデータを用いて技術の改良を行うことを目的として、スマートフォン用多言語音声翻訳アプリケーション “VoiceTra” (図 1) を開発し、平成 22 年 7 月 31 日付けで無料公開した。公開後、平成 22 年度末時点でダウンロード件数 46 万、アクセス件数 470 万を記録し、非常に多くの人々に使われた。

翻訳可能な 21 言語 (下線の 5 言語は音声入出力が可能)

日本語、ヒンディ語、スペイン語、英語、インドネシア語、タガログ語、中国語(普通話)、イタリア語、タイ語、中国語(台湾)、マレー語、ベトナム語、韓国語、ポルトガル語、アラビア語、フランス語、ポルトガル語(ブラジル)、オランダ語、ドイツ語、ロシア語、デンマーク語



翻訳可能言語が話される地域



46 万ダウンロード、470 万アクセスを達成 (2011 年 3 月 31 日現在)  
(サポートページ <http://mastar.jp/translation/index.html>)

図 1 スマートフォン用多言語音声翻訳アプリケーション “VoiceTra”

#### 世界初のネットワーク型音声翻訳技術の標準化を達成

平成 21 年 10 月より ITU-T SG16<sup>1</sup>、WP2<sup>2</sup> においてネットワーク型音声翻訳技術の標準化に向けて活動を行ってきたが、平成 22 年 10 月 14 日に勧告 F.745 および勧告 H.625 (表 1) として承認された。

表 1 ITU-T で承認されたネットワーク型音声翻訳技術に関する勧告

草案名	タイトル	スコープ	最終勧告名
F.S2STreqs	Functional Requirements for Network-based S2ST	<ul style="list-style-type: none"> <li>■ ネットワーク型音声翻訳の定義</li> <li>■ サービス要求条件と機能</li> </ul>	F.745
H.S2Starch	Architectural Requirements for Network-based S2ST	<ul style="list-style-type: none"> <li>■ アーキテクチャにおける要求条件</li> <li>■ ネットワーク型音声翻訳におけるインタフェースの定義</li> </ul>	H.625

<sup>1</sup> Study Group 16: Multimedia coding, systems and applications について検討する研究部会。

<sup>2</sup> Working Party 2: Applications and systems について検討する作業部会。

### 音声処理技術の研究

音声認識技術に関して、音声認識モデル学習コストの低減、多言語化、話者年齢層の拡大を目的として、以下の研究を行った。

- ・ 信頼度を用いた音声認識用音響モデルの自動学習手法を開発し、実利用データで有効性を確認した。
- ・ 韓国語音声認識に関して統計的手法による発音変形修復手法を考案し、実用性能を達成した。
- ・ 子供対応の音声認識用音響モデルの構築を行い、実利用データで効果を確認した。

音声合成技術に関しては、自然な対話音声の実現、多様な話者性の実現を目的として、以下の研究を行った。

- ・ 対話音声コーパスから音響モデルを学習することで、対話に適した韻律で音声を合成する方法を確立した。
- ・ 混合励振源モデルの枠組みを確立し、合成音声のブザー感、雑音感の軽減に成功した。
- ・ ネットワーク上の音声コンテンツから音響モデルを自動構築するシステムを試作し、有望な結果を得た。

### 音声対話システムの研究開発

音声対話処理技術に関して、次の研究を行った。

- ・ 音声対話コーパスより統計的対話制御モデル（重み付き有限状態トランスデューサ）を学習し、これを用いて対話制御を行うことで、人間に近い自然な対話を実現する技術を確立した。
- ・ ユーザの頭部検出に基づく対話相手の属性推定、顔向き推定とそれに基づく興味推定、システムの誤応答に対する反応の検出を行う技術を開発した。
- ・ 観光スポットに関する2,000名の被験者の選好評価構造を定量化し、観光スポット推薦システムを開発した。

これらの研究結果を統合して音声対話システム（図2）を開発し、学会・展示会で動態展示するとともに、被験者実験により非言語情報処理を考慮した音声対話システムの有効性を確認した。

また、音声対話・機械学習技術の機能実証としてロボカップ世界大会 家庭用ロボット部門に参加した結果、家事動作の模倣学習技術が評価され、24 チーム中で優勝した。



図2 2つの音声対話システムを開発