

「賢い」音声合成による 次世代音声情報処理技術の開拓

コンテンツ科学研究系 准教授

山岸 順一



研究背景・目的

音声合成・音声認識を核とする音声インタフェースが身近に利用できるようになってきました。スマートフォン等に音声で話しかけると言葉を認識し、必要な情報を音声合成でユーザに提示する音声対話システムや、音声認識を行い、自動翻訳し、多言語に翻訳した結果を音声合成にて出力する音声翻訳システムが良い例でしょう。人間の自然な音声を生成する技術である「音声合成」において、現在重要視されていることは、その人の声で、もしくは、その言葉を話すのにふさわしい人物の声で、あるいは、周囲の状況にふさわしい音声で出力すること。ユーザの希望に合わせて、状況に合わせることで、以下では、このように賢く臨機応変な音声合成技術の研究成果および社会還元への取り組みについて紹介します。

研究内容

現在の音声合成は、ベイズ理論や隠れマルコフモデルと言った統計理論を用いた枠組みが主流になりつつあります。私たちの研究は、統計的音声合成に必要なデータ量をこれまでの100分の1以下に減らすことを可能にした「適応」技術を基礎としています。私たちはこの適応技術について10年以上にわたり研究を行っており、この技術により、音声合成は、「楽しい」といった感情のこもった喋り方を模倣する事ができるようになったことは勿論、車内などの騒音下ではより聞き取りやすい様、喋り方を自動的に変えることが可能になりました。また、この適応技術を発展させることにより、喉頭がんやALSなどの病気により自分の声を失う人に、もともとの自分の自然な声で会話を行うことが可能な個人用音声合成システムを提供する事も可能になりました。更に、自分の声で外国語が喋れるよう、音声翻訳システムの出力をパーソナライズする研究も行われています。単に「読み上げる」機能を超えた賢い音声合成が実現しつつあります。

産業応用の可能性

- 自分の声を利用した音声障害者用会話補助器、ボイスバンク（声の“保存”）
- 騒音下でも聞き取りやすいカーナビゲーション、高齢者にも聞き取りやすい音声ガイド
- 自分の声を利用した音声翻訳システム、自分の声を利用した吹き替えシステム
- ユーザの動きに応答的に音声出力を変化させ、声を楽器の様に操る声の演奏システム



図1 自分の声を搭載した音声合成アプリ



図2 体の動きを利用した声の演奏システム

連絡先：山岸 順一 [コンテンツ科学研究系 准教授] URL <http://researchmap.jp/read0205283/> Email : [jyamagis\[at\]nii.ac.jp](mailto:jyamagis[at]nii.ac.jp)