

研究開発テーマ名

自在音声対話の研究開発

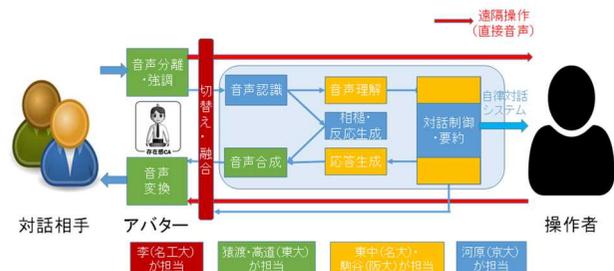
2022年度までの進捗状況

1. 概要

音声認識・対話技術を発展させて、人間のようにホスピタリティの感じられる自律的な音声対話システムを実現するとともに、アバターが操作者の意図や状況に応じて、遠隔操作対話と自律対話を臨機応変に切り替えられる自在対話技術を開発します。高齢者の傾聴対話、案内・プレゼンテーション、就職面接対話、相談・カウンセリング、接客・営業などの応用場面を想定して、システムを開発します。

本研究開発は、本プロジェクトにおいて、音声言語による対話処理の高度化を担っています。音声対話システムは、スマートスピーカやチャットボットなどにおいて実用化されましたが、画一的な知識レベルのやりとりにとどまっています。人間のように長く深い対話を実現するには、ユーザの状況（対話内外を含む）の理解とともに、自然な相槌や共感的な応答の生成が不可欠です。さらに、アバター技術との迅速かつシームレスな切替えに必要な対話制御や要約の処理について研究を行います。

これらを統合したシステムとCGアバター操作環境を開発します。全体像を以下に示します。



2. 2022年度までの成果

- 音声処理
 - ① 展示会場・ストリート・カフェテリアなどの実環境騒音下で頑健に動作する音声分離・認識を実現しました。
 - ② 感情表現ができ、自然性の高い低遅延の音声合成を実現しました。また、リアルタイムに動作する音声変換を実現しました。国際会議における合成音声の自然性を評価するコンテストで1位となりました。
 - ③ 対話相手の人間に同調して相槌や笑いを生成するロボットを実現しました。この論文発表に関して、英国 Guardian, Telegraph, Independent, BBC World など数多くの海外メディアで取り上げられ、フランス大手メディアによる2022年度世界10大イノベーションに選ばれました。

Scientists try to teach robot to laugh at the right time

Research team hopes system could improve natural conversations between humans and AI systems



- 言語・対話処理
 - ① 世界最大規模の Persona 情報付き対話データと日本語初の大規模タスクコーパスを構築しました。
 - ② 自然性の高いCGアバターを設計し、自律システムと遠隔操作の両方で動作するソフトウェア環境を構築しました。
- 統合システムと実証実験
 - ① 3名相手に同時並列に傾聴するシステムを作成し、動作を確認しました。
 - ② 3名相手に並列に説明・ガイドを行うシステムを作成し、水族館において1か月間実証実験を行いました。



3. 今後の展開

音声処理と言語対話処理については、統合システムに必要なレベルのものが開発できました。今後、システムを用いた被験者実験や実証実験を通じて、評価を進めるとともに改善を図ります。応用場面に関しても、面接や相談などのより複雑なタスクや、複数人が関わる対話を想定して、音声処理・対話処理とシステムの研究を進める予定です。