

氏名	西田 麻希子
授与した学位	博士
専攻分野の名称	工学
学位授与番号	博甲第135号
学位授与の日付	令和2年3月24日
学位論文の題目	発話感情表現に基づく反応動作と音声相槌に関する研究
学位審査委員会	主査 渡辺 富夫 副査 伊藤 照明 副査 石井 裕 副査 山崎 大河

## 学位論文内容の要旨

現在、音声対話システムの研究開発が盛んに行われ、音声入力システムは日常的に使用できる身近な技術である。音声操作可能なシステムとして受付ロボットなどの案内エージェントや、Apple 社の「Siri」、Google 社の「Google Home」などが実サービスとして展開されている。音声認識システムに身体エージェントをもたせた「スマートメイちゃん」のような汎用性のあるプラットフォームも運用されつつあり、今後音声対話システムは様々な分野で発展すると期待される。しかし、現在利用されている音声対話システムは、身体的リズムの同調など人同士の対話に存在するノンバーバル情報が十分に考慮されておらず、使用者が違和感を感じることもある。

人の対面コミュニケーションでは、言葉によるバーバル情報だけではなく、音声に対するうなずきや身振り・手振りなどのノンバーバル情報が相互に同調することで、対話者同士が互いに引き込み合い、円滑なコミュニケーションが行われている。また情動変動と密接に関連した心拍変動の引き込みや呼吸の引き込み等、生理的側面での生体リズムの引き込みも、コミュニケーションに重要な役割を果たしている。人とのインタラクションを行う対話エージェントの研究開発においてもこれらの身体的なかかわりはインタラクションに重要な役割を果たすと期待される。

また、今後さらに音声対話システムを発展させていくには、身体的リズムの同調に加えて、使用者の状態に即した応答を行うことも重要である。相手の感情を汲み取った応答は人同士の対話における重要な要素のひとつであり、近年、身体動作や情動表出による社会的相互作用に着目して、人やロボットとのインタラクションに応用する試みも研究が進められており、使用者の感情推定に基づく応答動作を行うことで、より使用者に寄り添った対話エージェントの開発が期待されている。

渡辺らはこれまでに、発話音声と身体動作が同期する身体的引き込みに着目し、インタラロボット技術 iRT を開発してきた。iRT は音声と身体動作の関係をモデル化することで、発話音声の有無（音声の呼気段落区分での ON-OFF パターン）から身体的引き込み

を伴ううなずきなどのコミュニケーション動作を自動生成し、コミュニケーションを支援するものである。この iRT を CG キャラクタに組み込んだ音声駆動型身体的引き込みキャラクタ **InterActor** を開発し、コミュニケーション支援の有効性を示している。従来の **InterActor** は、発声の有無に基づいて、うなずきなどのキャラクタの身体動作が生成されることによる効果を確認してきたが、使用者の状態によって反応を変化させる機能は実装されていなかった。そのため、使用者が否定的な発話をした場合においてもキャラクタがうなずき動作を行ってしまい、使用者がうなずきを肯定動作と解釈した場合、その否定的な発話感情を肯定し、助長してしまう可能性がある。

そこで本研究では、音声入力からうなずきなどのコミュニケーション動作を自動生成する音声駆動型身体的引き込みキャラクタ **InterActor** に対して、身体的リズム同調を損なうことなく、使用者の発話感情に応じた反応動作または音声相槌を提示するシステムを開発し、その評価を行った。以下、本論文における各章の成果をまとめる。

1 章では、研究の背景と目的、論文の構成を述べた。

2 章では、身体的リズム同調を損なうことなく使用者の状態に寄り添った応答を行うことを目的に、従来の音声駆動型身体的引き込みキャラクタ **InterActor** に音声認識を導入し、話者の発話内単語の感情極性から使用者の状態を推定し、結果に基づいて反応動作を変化させる音声駆動型身体的引き込みキャラクタシステムを開発した。

3 章では、2 章で開発したシステムの評価実験を行った。Negative または Positive と判定されたシナリオに基づくキャラクタへの語りかけによる自動応答エージェントシステムとしての評価実験および二者対話によるコミュニケーションインタフェースシステムとしての評価実験を行った結果、システムの有効性が示された。

4 章では、使用者の多様な感情に対応することを目的に、2 章で開発したシステムに発話時間率に基づく活性度を加えて話者の状態を推定する状態推定モデルを定義し、結果に基づき反応動作を行う身体的引き込みキャラクタシステムを開発した。発話活性度を考慮した状態推定の有効性を検討するために、2 章で開発したシステムとの比較実験を行ったが、評価実験における活性度の状態変化が小さいことでモードの判別がつきにくかった可能性もあり、十分な効果は認められなかった。

5 章では、iRT によるうなずきの出力タイミングに基づく音声相槌出力について検討した。従来の **InterActor** には音声による応答機能は実装されていなかったが、うなずきに相当する音声相槌を意味的な解釈として出力させることで使用者の発話感情に対応したシステムとして活用できる可能性がある。そこでまず従来の **InterActor** に音声相槌を付加したキャラクタシステムを開発し、自動生成する音声相槌についてその出力タイミングとキャラクタ表示による効果に着目した評価実験を行った。その結果、うなずき反応モデルによる音声相槌の推定は応答として有効に機能していることが確認され、表示されたキャラクタのうなずき動作に音声相槌を伴うモードがうなずき動作のみのモードよりも高く評価された。次に、日本語対話での応答特性を考慮した頻度で

音声相槌を自動生成するキャラクタシステムを開発して評価実験を行った。その結果、提案手法を用いたモードが高く評価され、システムの有効性が示された。

6章では、5章で開発したうなずき動作に音声相槌を伴う音声駆動型身体的引き込みキャラクタシステムにおいても、否定的な発話感情を助長する可能性を回避する手法を提案した。まず、うなずき動作に対する音声相槌の好ましい提示タイミングを検討するための評価実験及び **Negative** な発話が行われた際の音声相槌の提示タイミングを検討するための評価実験を行った。その結果、**InterActor** が推定したうなずき動作の開始から **600 ms** 程度までの音声遅延が許されることが示された。また、**Negative** な発話内容に対する音声相槌のタイミングは **900 ms** 程度の遅延まで許容されることが示唆された。次に、うなずき動作に音声相槌を伴う音声駆動型身体的引き込みキャラクタシステムに、発話内単語の感情極性から話者の状態を推定する状態推定モデルを導入し、音声認識により得られた発話内単語の感情極性に基づいて、**Negative** の場合に **900ms** の応答遅延を伴う音声相槌を行う身体的引き込みキャラクタシステムを構築した。

7章では、本研究を通じて得られた成果を総括するとともに、今後取り組むべき課題について述べた。

## 主業績

No.1	
論文題目	発話内単語の感情極性に基づき反応動作を行う身体的引き込みキャラクターシステム
著者名	西田麻希子, 太田靖宏, 渡辺富夫, 石井裕
発表誌名	日本機械学会論文集, Vol.83, No.846 (2017), pp.1-14, DOI: 10.1299/transjsme.16-00148.
No.2	
論文題目	音声相槌を行う音声駆動型身体的引き込みキャラクターシステム
著者名	西田麻希子, 渡辺富夫, 石井裕
発表誌名	日本機械学会論文集, Vol.85, No.880 (2019), pp.1-12, DOI: 10.1299/transjsme.19-00159.

## 副業績

No.1	
論文題目	発話活性度および感情極性に基づき反応動作を行う身体的引き込みキャラクターシステムの開発
著者名	西田麻希子, 渡辺富夫, 石井裕
発表誌名	日本機械学会 2018 年度年次大会講演論文集(2018), G1200101, pp.1-5.
No.2	
論文題目	身体的引き込みキャラクターシステムにおける音声相槌タイミングの評価
著者名	西田麻希子, 渡辺富夫, 石井裕
発表誌名	日本機械学会 2019 年度年次大会講演論文集(2019), S12102, pp.1-5.
No.3	
論文題目	発話内単語の感情極性に基づく音声相槌を伴う身体引き込みキャラクターシステムの開発
著者名	西田麻希子, 渡辺富夫, 石井裕
発表誌名	第 169 回ヒューマンインタフェース学会研究会研究報告集(2019), Vol.21, No.7 (2019), pp.21-24.

## 論文審査結果の要旨

本論文は、発話音声からうなずきなどの聞き手の引き込み反応を自動生成し、発話感情を推定して反応を変化させることで、身体的リズム同調を損なうことなく使用者の状態に寄り添ったインタラクション・コミュニケーション支援を行う手法とシステムの研究開発について論じたものである。

現在、音声操作可能な様々なシステムが実用化され、音声対話システムの研究開発も盛んに行われている。しかし、今後さらに音声対話システムを発展させていくには、身体リズムの同調などの人間同士の対話に存在するノンバーバル情報を十分に考慮しながら、使用者の状態を汲み取った応答を行うことでより使用者に寄り添ったシステムを設計することが重要である。本論文で提案するシステムは、身体リズムが同調する身体的引き込みに加え、主に使用者の否定的な発話感情を助長しないという観点から、使用者の状態に即した応答を行うことで人とエージェントとのインタラクション・コミュニケーションを促進するものである。

具体的には、まず反応動作による応答として、従来の発話音声に基づいてコミュニケーション動作を自動生成するキャラクタInterActorに音声認識を導入することで、発話内単語の感情極性から使用者の状態を推定し、結果に基づいて反応動作を変化させる音声駆動型身体的引き込みキャラクタシステムを開発している。官能評価によりシステムの有効性を示し、使用者の状態に即した反応動作による応答の重要性を示している。次に音声による応答として、従来のInterActorに音声相槌を付加したキャラクタシステムを開発し、自動生成する音声相槌についての効果を明らかにするとともにシステムの有効性を示している。さらに、うなずき動作に対する音声相槌の好ましい提示タイミングおよび否定的な発話が行われた際の音声相槌の提示タイミングを官能評価により示している。その上で、発話内単語の感情極性に基づいて、負極性の場合に応答遅延を伴う音声相槌を行う身体的引き込みキャラクタシステムを構築している。

本研究で開発した使用者の感情表現に基づく音声対話システムは、新たなインタラクション・コミュニケーションインタフェースの基盤になるものである。よって、本論文は博士（工学）の学位論文として価値あるものと認める。