

[研究助成 (C)]

高齢者の自然発話における特徴量を用いた認知機能推定手法の開発

2247005



研究代表者

東京大学 大学院
新領域創成科学研究科

後期博士課程

五十嵐 俊 治

[研究の目的]

近年、家具や室内など周辺環境に配置されたアンビエントセンサ等を活用したスマートハウス技術への意識が高まっている。人が居住する環境を電子化することによって、居住者の活動を記録して遠方に住む家族に共有し、電力消費量のマネジメントを行うことなどが可能になっている。また、場合によっては居住者に行動変容を促すなどの情報推薦を行うケースもある。本プロジェクトは、アンビエントセンサーを活用し、高齢者の生活支援および認知機能モニタリングを行うスマートハウスエージェントの開発を目的としている。

提案者らの研究では、世界的にも例を見ない9ヶ月間の長期介入研究を実施したが、会話内容を日毎に変えたとしても、次第に使用頻度が落ちていくことが確認された。

介入終了後に、研究参加者を集めたグループディスカッションを実施し、意見を集めたところ、「会話の内容がこちらの状況を把握したものではないため会話のモチベーションがあがらない」「適時適切な受け答えができない」といった意見が挙げられた。これは ChatGPT の API を活用した会話型エージェントでも同様であり、システム応答性の高い最新の対話エージェントである Cotomo を用いても超えられていない課題である。しかし、リザーブコンピューティングは、主に出力層でのみ学習が行われるため、大規模なデータセットに対しても

高速に学習を行うことができる。この特性を活かし、リアルタイムまたはリソース制限のある環境での応用が可能となると考えている。

[研究の内容、成果]

本システムは、高齢者の聴覚・認知特性に最適化された対話を実現するため、複数の技術要素を統合した設計となっている。

本システムは、クラウド環境とローカルアプリケーションを組み合わせたハイブリッド構成を採用している。ユーザーの音声はローカルアプリケーションで処理され、Amazon API Gateway を経由して AWS クラウドに送信される。クラウド側では、AWS Lambda および Amazon ECS を活用し、音声データの解析と適応処理を実施する。解析結果は Amazon Simple Storage Service (S3) に格納され、管理画面 (Next.js 14 を用いた AWS Amplify ベース) を通じてモニタリングが可能となっている。更に、Echo State Network (ESN) は音声の高低・音量調整をリアルタイムに行い、適切な話速 (mora 単位) への調整を可能にする。また、音声から推定された感情ラベルをプロンプトに挿入し、LLM (Amazon Bedrock や Azure GPT) を通じて対話の自然さを向上させる。これにより、従来の固定的な音声合成と比較して、話者に応じたダイナミックな発話が可能となる。

1. 開発構想・計画

まず、要件定義のために、暗黙知化されている医療・介護現場での形式化を行った。現場で10年以上の経験を持つ言語聴覚士および看護師にインタビューを行い、現場の知見を活かして寄り添った応対に最適化できるように、会話の[速度][大きさ][周波数]、使用する言葉づかい、対話構成（傾聴/質問/情報提供）についての調整を行うようにした。

本システムは、高齢者との対話を支援し、より円滑なコミュニケーションを実現するために設計を行った。本システムはAWSを基盤とし、音声対話、センサデータの統合、およびクラウドベースの処理を組み合わせることで、リアルタイムな対話の最適化を目指している。機能としては下記の機能を持つシステムとして要件を作成した。

本プロジェクトの目的は、高齢者を含む様々なユーザーに適応可能な対話型エージェントシステムを開発することである。対話型エージェントは音声特性のリアルタイム調整により、話速・音量・周波数を適応的に調整し、高齢者にとって聞き取りやすい対話を実現する。また、認知特性に配慮した対話構成を取り入れることで、傾聴・質問・情報提供を組み合わせた自然な会話の流れを作り、対話の継続性を向上させる。さらに、センサーを活用し、室温や時間帯に応じた発話を生成することで、生活環境に適応した対話システムを実現することを目指す。

プロジェクトの目標を下記のように設定した。

[1]. 家庭内での居住者の活動状況のモニタリング

温湿度、照度、電流、赤外線といった複数のセンサを包含したマルチセンサを開発し、自宅内のコンセントに設置したセンサーから居住者の活動情報を収集し、「ユーザーが今どういう状態であるか」を推定するシステムを開発する。

[2]. エージェントのインタラクションシステムの開発

スマートスピーカーなどのような受動的なシステムではなく、リアルタイムで室内環境を反映した会話を行うリザーバー・エージェント・インタラクション・システムを開発する。

[3]. 実証実験の実施

開発したシステムの応用・実用化にあたっては、協力の許可を得られている有料老人ホームやサービス付き高齢者向け住宅などに導入を行い、ユーザーや施設管理者からのフィードバックを得ることを最終目標とする。

2. センシングデータ、発話履歴のデータベース設計

センサデータおよび会話されたデータをサーバー内に取り込むためのバックエンドの開発とAWSでの設定を行った。本システムは、クラウド環境とローカルアプリケーションを組み合わせたハイブリッド構成を採用している。ユーザーの音声はローカルアプリケーションで処理され、Amazon API Gatewayを経由してAWSクラウドに送信される。クラウド側では、AWS Lambda および Amazon ECS を活用し、音声データの解析と適応処理を実施する。解析結果はAmazon Simple Storage Service (S3) に格納され、管理画面（Next.js 14 を用いた AWS Amplify ベース）を通じてモニタリングが可能となっている。

さらに、データベースは管理画面で一覧として確認できるように設計を行った。

3. ユーザーの健康状態推定システムの開発

本研究では、広島市立大学感情音声コーパス（HCUIDB）を用い、音声データから抽出したMFCC 特徴量を入力として、ESN（Echo State Network）によるリザーバーコンピューティング手法で処理し、話者の発話感情とパラ言語（話速、音量）を検出し、応答で自動調律するシステムを構築した。

ユーザーの発話中に特定のフレーズが検出された場合、その結果に基づいて koeiromap のパラメーターを動的に調整する仕組みを実装している。LSM (Liquid State Machine) もリザバーコンピューティングの一手法として知られており、生物学的神経回路の動的特性を模倣する点に特徴がある。しかし、LSM はスパイクニューロンなどを用いた複雑な内部ダイナミクスを持つため、実装やチューニングが困難である。一方、ESN は連続値を扱う単純なニューロンモデルを採用するため、実装の容易さと高速な学習、さらにリアルタイム処理への適性が高い。これらの理由から、本研究では LSM ではなく、ESN を採用した。

4. 電流センシングデバイスの改良

センサ要件を下記のように設定し、基盤設計を行った。

- ・ 温湿度、照度の計測ができること
- ・ ESN が走るマイコンが実装できること
- ・ wifi connect されたら、確認できること

本システムでは、電流センサ、照度センサ、音声センサ、赤外線送受信機、リレースイッチを搭載したコンセント設置型のマルチセンサを用い、カメラを使用せずに居住者の活動を推定する。また、エージェントは単なるスマートスピーカーではなく、リアルタイムで居住環境を反映した対話を行うリザバー・エージェント・インタラクション・システムとして設計されている。

プロジェクトの成果として、

(1) 居住者の活動状況のモニタリング可能なセンサの開発、(2) エージェントによるインタラクション・システムの開発、(3) 高齢者施設などでの実証実験の計画が挙げられる。特に、システムがユーザーの設定や入力を待つのではなく、エージェントから能動的な関与を可能にすることで、ユーザーに対する対話の質と継続性を向上させる点が特徴的である。高齢者の社会

的孤立を軽減し、認知機能のモニタリングを可能にすることで、医療機関との連携による認知症の早期発見・予防にも貢献できると考えられる。

本システムの Echo State Network (ESN) は音声の高低・音量調整をリアルタイムに行い、適切な話速 (mora 単位) への調整を可能にする。また、音声から推定された感情ラベルをプロンプトに挿入し、LLM (Amazon Bedrock や Azure GPT) を通じて対話の自然さを向上させる。これにより、従来の固定的な音声合成と比較して、話者に応じたダイナミックな発話が可能となった。

[成果の発表、論文など]

【著書】

1. ○ Toshiharu Igarashi, Estimating Cognitive Function Using Spontaneous Speech in Older People Living in the Community, Pages: 105, Published: 2024-09-25, ISBN: 978-99993-2-042-9, ELIVAbooks.

【査読あり論文誌】

2. ○ Toshiharu Igarashi, Katsuya Iijima, Kunio Nitta, Yu Chen. Qualitative Analysis of Responses in Estimating Older Adults Cognitive Functioning in Spontaneous Speech: Comparison of questions asked by AI agents and humans. Healthcare. Vol. 10. No. 10. MDPI, 2024.
3. ○ Toshiharu Igarashi, Katsuya Iijima, Kunio Nitta, Yu Chen. Detailed Analysis of Responses by Interviewers in Estimating Cognitive Function of the Elderly through Natural Speech: Comparison of Questions by AI Agents and Humans. International Journal of Environmental Research and Public Health 2024.
4. ○ Toshiharu Igarashi, Yoichi Ochiai. Comparative Analysis of Wine Taste Augmentation: Artificial Climate Room Versus General Meeting Room in Evaluating Temperature and Humidity Effects. International Journal of Augmented Human 2024.
5. ○ Toshiharu Igarashi, Katsuya Iijima, Kunio Nitta, Yu Chen. Estimation of Cognitive Functioning of the Elderly by AI Agents: A Comparative Analysis of the Effects of Psychological Burden of Intervention. Healthcare. 2024.
6. ○ Toshiharu Igarashi, Yu Chen. Analysis of Intervention Effects of Prototyping Utilizing UI Stencils in Elementary School Students. Journal of Science Education and Technology 2024.