

認知症患者のコミュニケーション特性の評定における 会話エージェントの利用

酒井 洋一¹ 野中 裕子² 安田 清³ 中野 有紀子²

成蹊大学大学院 理工学研究科 理工学専攻¹

成蹊大学 理工学部 情報科学科²

千葉労災病院/京都工芸繊維大学³

1 はじめに

本稿では、認知症患者とその介護者を支援することを目的とし、アニメーションによる語りかけエージェントを提案・実装する。まず、基本的なコミュニケーション機能として、利用者の発話音声のピッチと音量による音声の特徴に基づき、頷きや相槌によるフィードバックを返す機構を開発する。次に、患者の発話音声や頭部の動きを認識し、エージェントとの会話中の患者の言語・非言語行動を記録するシステムを提案する。これにより、高齢者の認知状態を測定するために有益な情報を取得することができるかと期待される。



図1： エージェントのスナップショット

2 先行研究

認知症患者の数が増え続けており、患者の大半は高齢である。そのため、介護者の数が足りておらず、ケアが十分に行えないという問題がある。このような背景を受け、支援ロボットの研究でも、身体的な補助に加えて、ロボットのコミュニケーション機能も重要視されつつあり、会話を行うヒューマノイドを高齢者のケアに利用した研究も行われている[1]。

Pollack[2]は、認知障害を持つ高齢者を介助する機能として以下の3つを提案している。

- 高齢者が安全であり、日常の必要な活動を行っていることを確認し、そうでなければ介護者に知らせる
- 障害となっていることを補償し、日常生活を送ることができるよう支援する
- 高齢者の認知的状態を測定・評価する

我々は、3つ目の測定・評価において、人型のバーチャルエージェントを用いることが有用であると考えた。なぜなら、会話は高齢者に最も好まれる活動の一つであるため、会話を通して認知状態を測定する事ができれば、検査のストレスをかけない、最も自然な測定方法になる。そこで本稿では、患者の発話に対するフィードバックを返すことにより、発話を促すと同時に、認知状態を測定するために必要と思われる情報を取得するシステムのプロトタイプについて報告する。

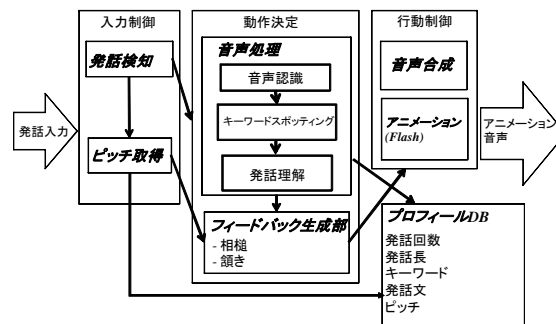


図2： システム構成図

3 語りかけエージェント

エージェントが患者によく受け入れられている場合、患者にとってよいコミュニケーションパートナーとなる。高齢者によるロボットエージェントの受容に関する過去の論文では、エージェントが笑顔や頷きといった社会的なシグナルを表出する事が、高齢者に受容される上で重要であることがわかっている[3]。また、患者の信頼の獲得や親密さを深めるためにも有効である[4]。そこで、初めに頷きと相槌のみを表出する単純な会話エージェントを実装した。

3.1 システム概要

語りかけエージェントに関して、図1にエージェントのスナップショットを、図2にシステム構成図を示す。本システムでは、ユーザの発話を引き出すために、エージェントが一連の質問を行う形式で会話を行う。質問群は医者や看護師の方のアンケート結果をもとに、体の状態や食事、子ども時代の思い出や故郷の事に関する典型的な質問を用意した。以下にシステムの各モジュールについて説明する

入力制御モジュール：発話認識部ではマイクから入力された音声の音量から、発話入力の有無を判定する。個人ごとに判定の閾値は適宜調整することとした。次に、

Employing Conversational Agents in Assessing Communication Status by People with Dementia
Yoichi Sakai¹, Yuuko Nonaka², Kiyoshi Yasuda³, Yukiko Nakano²

1: Graduate School of Science and Technology, Seikei University

2: Dept. of Computer and Information Science, Seikei University

3 Chiba Rosai Hospital/Kyoto Institute of Technology

表 1 フィードバック生成ルール

条件	エージェントの行動
現在のピッチ n が $n-1$ から $n-16$ までの平均ピッチより 5% 低下	頷き
発話終了後 2 秒以内の無音があり再度発話が検出される	相槌(ええ)
発話終了後 2 秒以上の無音 or 発話終了直前のピッチ下り幅が 2% 以内 or 発話終了直前のピッチが発話平均ピッチより 10% 以上上昇	相槌(はい)

ピッチ取得部では、発話入力であると判定された音声に対して、ピッチを算出する。これらの値は動作決定モジュールに出力される。

動作決定モジュール：行動決定部では、入力制御モジュールから得られた音声情報を用いて、エージェントの内部状態の更新と行動の決定を行う。エージェントは質問を実行すると待機状態に入り、ユーザの発話が検出されると、入力制御モジュールから得られたピッチの情報をフィードバック生成ルール（表 1）に適用し、相槌と頷きを決定する。フィードバック生成ルールは[5, 6]を参考にして作成した。エージェントの質問後、一定以上の時間が経過してもユーザからの応答がない場合には、次の質問に移行する。

行動制御モジュール：行動制御モジュールは、キャラクターのアニメーションの実行と音声ファイルの再生を行う。エージェントの 3D モデルは、Poser 7 を用いて作成された。エージェントの動きであるアニメーションは、Poser7 で作成されたアニメーションから画像ファイルを書き出し、これらをキーフレームとして Adobe Flash ver. 10.0 により、web ブラウザ上で連続描画することにより実装した。今回の実験では、エージェントは大きな動作は行わず、発話中の口の動き、頷き動作、またアイドリング動作として微細な体の揺れを表現するにとどめた。

4 音声認識・言語理解

フィードバックの表出に加えて、患者の認知的状態の評価に利用できるデータを取得するために、患者の発話の理解が必要である。そこで、音声認識・言語理解の機能と発話のログを取得する機能を加え、言語情報から取得できる情報をプロフィール DB に保存した。患者が単語を明瞭に発音しない場合があり、加えてシステムが頑健に動く事も重要であったため、簡単ではあるが安定した方式を採用した。

4.1 音声認識

認知症患者のリハビリ中の訴えの内容の書き起こし、これをもとに単語辞書・言語モデルを作成した。認知症患者の発話音量、発声の明瞭度には個人差はあるが、単語が明瞭に発音されないことが少なからずあったため、キーワードスポッティングを用いて、あらかじめ登録しておいたキーワードを検出し、キーワードのリストを音声認識結果とした。

4.2 言語理解

言語理解は、音声認識結果をあらかじめ用意した想定発話文と比較し、類似度の高いものを言語理解結果とする方式を用いた。類似度計算には、想定発話文の内容語の単語ベクトルと音声認識結果の単語リストのベクトルとのコサイン距離を使用した。その結果、最も類似度が高かったものが発話内容として理解される。さらに、各想定発話文に対して予め応答が設定されているので、言語理解が決定されると、それに対応した応答を音声とアニメーションにより出力する。

5 終わりに

本研究では、認知症患者の支援を目的とした会話エージェントの実現を目指し、質問を行うことにより患者に語りかけるシステムを実装した。本システムは、患者の発話に対して相槌や頷きを返すことに加え、患者の認知的状態を評価するための音声・言語情報を記録することができる。患者による会話エージェントの受容性について既に評価実験を行っており、その結果、患者は会話エージェントが行うフィードバックに対してより注意を向けている事が分かった。このことは、患者は社会的シグナルを表出するエージェントとコミュニケーションすることに、より興味を示すことを示唆している。今後は、言語情報に加えて患者の頭部動作情報に関する分析を行うとともに、ユーザにとって興味のある話題を選択できる機能についても検討していく予定である。

謝辞：本研究の一部は科研費新学術領域研究、および文部科学省私立大学戦略的研究基盤形成支援事業の支援によるものである。

<参考文献>

1. Bickmore, T., and Picard R. W. *Towards Caring Machines*. in *CHI 2004*. 2004.
2. Pollack, M., *Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairments*, in *AI Magazine*. 2005. p. 9-24.
3. Heerink, M., et al., *Studying the acceptance of a robotic agent by elderly users*. *International Journal of ARM*, **7**(3), 2006.
4. Cassell, J., *Nudge, Nudge, Wink, Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents*, in *Embodied Conversational Agents*, J. Cassell, et al., Editors. 2000, MIT Press: Cambridge. p. 1-27.
5. Gratch, J., et al., *Virtual Rapport*, in *6th International Conference on Intelligent Virtual Agents*. 2006, Springer: Marina del Rey, CA.
6. Tsukahara, W. and N. Ward. *Responding to Subtle, Fleeting Changes in the User's Internal State*. in *CHI 2001*. 2001: ACM.