発話状況認識に基づく遠隔対話支援エージェント

及川　開斗†　　福田　悠人†　　久野　義徳†　小林　貴訓†
†埼玉大学

1. はじめに

少子高齢化を背景に，介護が必要な高齢者が増加し，介護士への負担が大きくなっている．特に，日常的なコミュニケーションが不足する独居高齢者への介護負担増加から，対話支援を行うロボットの開発に期待が寄せられている．我々のグループでは，遠隔地の高齢者が介護士や家族との会話を円滑に進めることができるビデオ通話システムの開発を進めている[1]．本稿では，これまでのビデオ通話システムに加えて，利用者の対話を支援する会話エージェントの導入を行う．このエージェントは，発話に対してロボットの反応があると会話意欲が高まる効果[2]を利用する．

2. ビデオ通話システムとエージェントの作成

エージェントの作成には MMDAgent を用い，音声認識機能と発話機能，シナリオに沿ったモデルの行動指定を行う．これらの機能によって本エージェントは，ビデオ通話時の発話状況を把握し，3D モデルが発話者の方向を見てうなずく．これにより，利用者は発話に対する反応を感じ，発話意欲が増すと考えられる(図 2)．

図 2. システム使用例

MMDAgent の音声認識には話者の特定を行う機能が存在しないため，本システムにおいては，音声認識ソフト Julius を複数起動し，2つのマイクの音声認識結果を利用することで，ビデオ通話利用者のうちどちらの利用者が発話中であるか把握し，うなずく方向を決定している．

また，会話をやめてしまう原因の一つである両者の沈黙への対策として，両者が一定時間発話しなかった場合，MMDAgent の発話機能を用いて 3D モデルが適当な話題を提供する機能を設けた．

更に，取得した音声認識結果を保存しておき，利用者が相手の言葉を聞き返す発言をした際に，自動で繰り返す機能を実装した．これにより，利用者が会話をより円滑に行うことが出来ると考えられる．

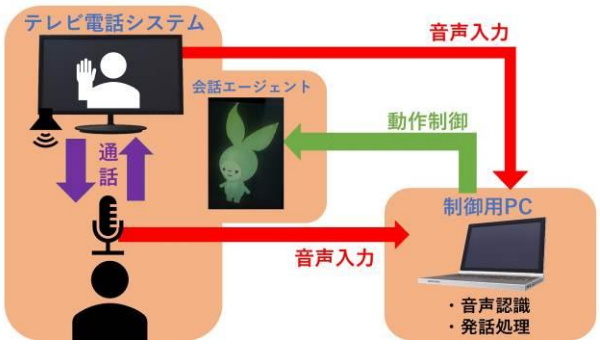

図1. システム概要

3. おわりに

本稿では，ビデオ通話における発話状況を把握して 3D モデルによる会話促進行動を与えることで対話継続を試みるシステムを提案した．現在の問題点として音声認識精度があまりよくないため，発言を繰り返す機能においても精度が期待できないことが挙げられる．

本システムにおける会話の促進効果については今後実験をして検証していく予定である．また，その際には今回使用した 3D モデルの他に実物体となるロボットによって行動を模倣し，どちらのほうが利用しやすいかについても検証していく．

本研究の一部は JSPS 科研費 JP17K18850 の助成による．

参考文献

[1] 大津耕陽，松田 成，福田悠人，小林貴訓，久野義徳，”テレビ電話と連携した人形型デバイスを用いた高齢者の遠隔見守りシステム，” 画像センシングシンポジウム，2016.

[2] Masaaki Tanaka, Akira Ishii, Emi Yamano, Hiroki Ogikubo, Masatsugu Okazaki, Kazuro Kamimura, Yasuharu Konishi, Shigeru Emoto, and Yasuyoshi Watanabe, “Effect of a human-type communication robot on cognitive function in elderly women living alone,” Medical Science Monitor 2012.

Copyright © 2018 IEICE