

クライアントの情動変化に基づき心の見守りを行う対話 エージェントの提案

A Proposal of Dialogue Agents that Watch the Heart based on Emotional Change of Clients

高 潔[†] 佐野 睦夫[†] 足立 奈生[†] 早川 祥太郎[†] 宮脇 健三郎[†]

Jie Gao Mutsuo Sano Nao Adachi Syotaro Hayakawa Kenzaburo Miyawaki

1. はじめに

高齢化社会が進む中、一人暮らしの高齢者が増えており、日常生活の中でコミュニケーション機会は減りつつある。一方で、コミュニケーション機会の減少につれて、うつ病や認知機能の低下などの発生するリスクは他の人と比較するとかなり高いことが明らかである。また、コミュニケーション機会の減少は高齢者だけではなく、若年者にも広まっている。このような背景の中で、対話エージェントとコミュニケーションを行うことによって、クライアントの元気な生活を過ごすこと目的として、クライアントの情動変化に基づき心の見守りを行う対話エージェントを提案する。本報告では、日常生活の対話の中で、クライアントの情動変化を捉え、心の見守りを行うエージェントを提案する。具体的には、日常対話に対する情動の反応モデルをベイジアンネットワークにより体系化し、情動の変化パターンを弁別するとともに、情動変化への対処を専門家と共に解決するという方法論を提案する。

2. 情動変化について

図 1 (a) ~ (b) に、クライアントの情動変化による処理フローを示す。対話エージェントとのコミュニケーション、いわゆるシステムの開始はエージェントから「おかえり、お疲れ様」など声をかけてからと設定した (図 1 (a))。そして、クライアントの内面変化は「元気かどうか」の部分にしぼり (図 1 (b))、「元気である」場合は普通の日常会話をを行い、「元気でない」場合は異常ととらえる。その異常の原因を探すには、ベイジアンネットワークを用いてクライアントの状態推定を行う (図 1 (c))。その後、推定した状態にあった対応をする (図 (d))。図 1 (a) と (b) の時、web カメラとマイクを通じてクライアントの表情と音声の情報を取って、センシング処理を行う。

感情認識方式としては、6 感情 (喜び、驚き、嫌悪、悲しみ、怒り、恐れ) とそれらの覚醒度に基づいた感情モデルに基づき、顔画像情報、音声の韻律情報・言語情報から統計的パターン認識アルゴリズム (K-NN, SVM など) を適用する²⁾。ユーザが発話していない場合は主に表情認識を行っていき、感情をあらわす表情の特徴量から、6 感情 (喜び、驚き、嫌悪、悲しみ、怒り、恐れ)¹⁾ から有意な感情を検出する。ユーザが発話している場合は、取得した表情の分析結果と音声の韻律情報・言語情報を合わせ、現在の感情を複合的に認識する。

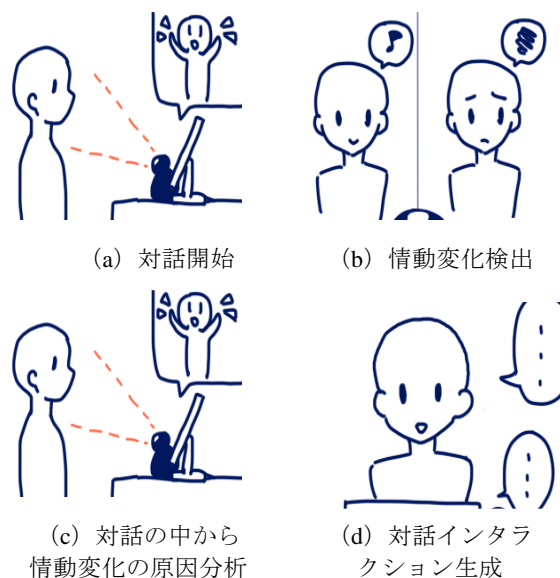


図 1 : 心の見守りを行う対話エージェント

3. 状態推定

ここでは「元気がない」の原因を推定するため、確率理論に従う推論 (確率推論) を行うネットワークモデルを構築した (図 2)。ベイジアンネットワークを用いて「元気がない」の原因を推定し、その中一番確率が高い原因に応じて、使用者への対応を変える。

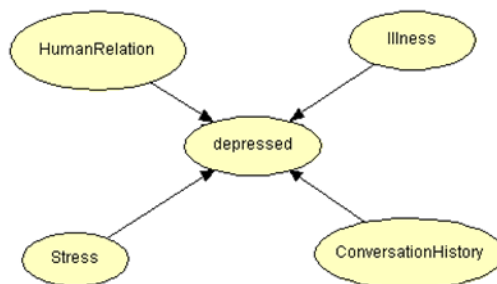


図 2 : 状態推定モデル

対話エージェントによって収集された生活エピソード履歴に対してエピソードの分類をキーワードの抽出によって行う。この分類は、形態素解析を行い、品詞分類をしてキーワードと 6 感情モデルのキーワードとの類似性 (相関関係) をもとに、言葉の類似度が高ければ、エピソードは、その感情を伴ったものとする。

[†] 大阪工業大学

コミュニケーションをさらに促進するために、クライアントの反応に応じて、エピソード展開（話題の深堀）を行う。生活エピソードデータベース（会話履歴など）から、クライアントの現在の状態に起因するエピソードを検索し、「確か、この前、***ということ話をしていたようだったけど、何かあったの?」「確か、この前も***があったね。」など、エピソード記憶をキーとしながらクライアントの元気を出せるような会話を行う。

3.1 学習について

図 3.1.1 に示すのはインタラクションモデルのフローである。エージェントとクライアントは会話するとき、エージェントの発話に対してクライアントの反応の変化（感情変化）を捉え、クライアントの反応パターンを学習する。この学習データはまた次の会話に活用する。

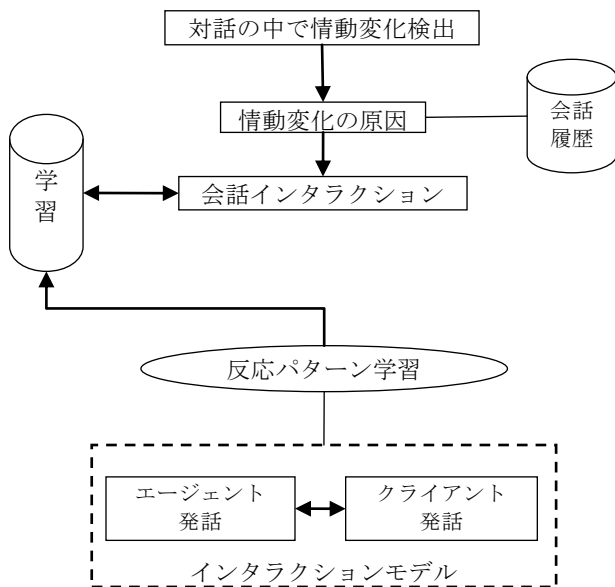


図 3.1.1 インタラクションモデル

4. 実験

提案システム全体の実装及び評価は現在進行中である。ここでは、本システムの要素技術である感情認識の基礎実験と情動変化に基づいてインタラクション生成例をする。

4.1 感情認識実験と結果

今回は表情認識のみからなる感情分析と韻律情報と同時表出した表情の感情分析との 2 種類の実験を行った。表情も音声も 6 感情+平常状態の合計 7 つが持つ特徴量を学習させたクラスに対して K-NN 法によって感情を推定する実験とする。6 感情（笑い、怒り、嫌悪、恐怖、驚き、悲しみ）をマイナスとプラス感情に分ける。

被験者に行ってもらふこととして、①については被験者 10 人にそれぞれの感情を持った表情を作ってもらった。②については被験者 5 人にそれぞれの感情を持った発話をしてもらい、同時に表情も分析した。なお、その際の表情は作ってもらふことなく自然なものを対象とした。

表情の学習対象は、個人ごとの平常時の表情+6 感情を含む表情を 10 個ずつ、合計 70 個の表情データを利用

した。表情データには 7 つの特徴変化量が含まれている。取得している特徴変化量を表 4.1.1 に示す。

表 4.1.1 表情特徴変化量

特徴量	内容
F 1	左頬の長さ
F 2	右頬の長さ
F 3	口の幅
F 4	口の開け具合
F 5	右目と右眉毛の距離
F 6	左目と左眉毛の距離
F 7	眉間の距離

音声の学習対象は振幅構造に反映している「短時間パワー」とピッチ構造に反映している「基本周波数」を用いて、音声データから一定時間ごとのそれぞれの最大値、最小値、平均値を抽出した。そして、「短時間パワー」、「基本周波数」それぞれの全体平均値との変化率を出し、それを特徴量として感情認識に使用した。音声認識で得られる特徴量のまとめを表 4.1.2 に示す。

表 4.1.2 音声特徴量変化量

特徴量	内容
F 1	Pitch の最小値の変化率
F 2	Pitch の最大値の変化率
F 3	Pitch の平均値の変化率
F 4	Power の最小値の変化率
F 5	Power の最大値の変化率
F 6	Power の平均値の変化率

認識結果は表 4.1.3 に示す。音声認識のところが 60% の認識率しか実現できない、認識率を改善中である。

表 4.1.3 各感情の正解率のまとめ

感情	表情の正解率	音声の正解率
平常	82%	56%
笑い	86%	52%
怒り	60%	56%
嫌悪	56%	42%
恐怖	70%	34%
驚き	98%	38%
悲しみ	74%	48%

4.2 情動変化に基づいてインタラクション生成例

クライアントはプラス感情の場合であれば、エージェントも笑いながらクライアントと喋り続く。マイナス感情の場合であれば、クライアントの心を見守りため、共感を行う。エージェントの表情もクライアントと同調させる。図 4.2.1 は対話エージェントとのインタラクション生成を示す。

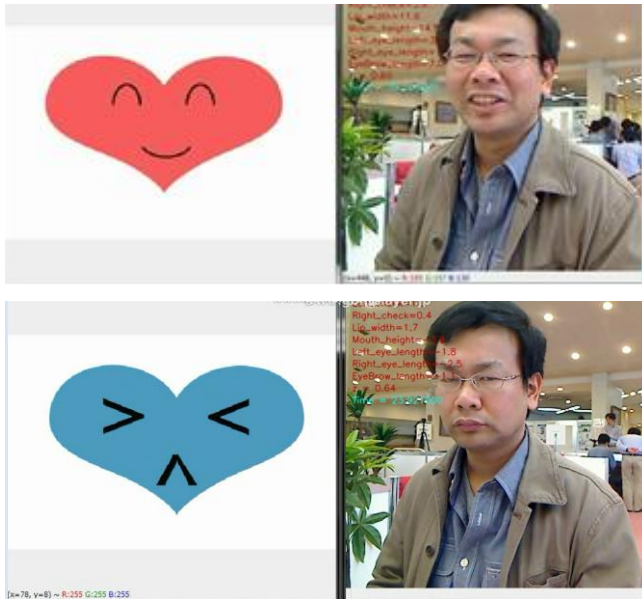


図 4.2.1 プラス感情とマイナス感情に対するエージェントの反応

5. おわりに

クライアントの情動変化を検出ことによって、心の見守りコミュニケーションが促進できると考えられる。今後の展望としては、認知障害者に対する記憶想起支援（リハビリテーション）への適用も可能であり、現在、計画中である。

参考文献

- 1) Paul Ekman, Wallace V. Friesen 著, 工藤 力 訳：“表情分析入門”, 誠信書房(1987)
- 2) 森博章, 宮脇健三郎, 佐野睦夫, 西口敏司：“コミュニケーションを円滑に進めるための表情変化検出方式の検討”, 情報処理学会研究報告. CVIM, 2008(82), 161-164 (2008)
- 3) Stuart Russell, Peter Norvig 著, 古川 康一 監訳：“エージェント アプローチ人工知能”, 共立出版株式会社 (2008)