

ユーザの頭部動作に基づく会話参加態度の推定

大古 亮太[†] 中野 有紀子[†]

成蹊大学 理工学部 情報科学科[†]

1. はじめに

対面会話では、話し手は、聞き手の動作や視線から、聞き手が会話に積極的に参加しているか否かを察知し、積極的に参加していない様子であれば話題を変える等、会話の内容を調整している。石井ら [1] は、ユーザの注視行動の分析に基づき、理想的な会話参加態度から逸脱している視線遷移パターンを同定し、注視行動による会話参加態度推定方式を提案している。しかし、ユーザの視線を厳密に計測するためにアイトラッカを用いているため、ユーザの頭部がやや大きく移動すると視線が計測できず、その結果、会話参加態度も推定できない点が問題であった。

一方、ヘッドトラッカを用いると、視線は厳密に測定できないが、頭部全体の動きをよりロバストに計測することができる。そこで本研究では、頭部動作に着目し、エージェントとの会話中のユーザの頭部の位置、回転角度のデータをヘッドトラッカにより計測し、頭部の移動速度、角度変化率、頭部の揺れの振幅、周波数等を算出し、これらと会話参加態度との関係を分析した結果を述べる。さらに、これらの分析結果に基づき、ヘッドトラッカにより計測されたユーザの頭部動作から、会話参加態度を推定する方式を提案することを目的とする。

2. 会話参加態度と頭部動作データの分析

2.1 分析データ

先行研究[1]では、アニメーションキャラクタを用いた会話エージェントが、被験者に携帯電話の説明をする想定でコーパス収集実験を行っている(図1)。その際に被験者とエージェントを記録したビデオと、アイトラッカによる被験者の視線データ、およびヘッドトラッカによる頭部動作データが収集された。本研究ではヘッドトラッカから得た頭部の位置と回転角度(x, y, z, roll, yaw, pitch)のデータ7人分を分析データとする。頭部動作計測にはオムロンOKAO Visionを用いた岡ら[2]による頭部動作計測システムを使用した。図2にヘッドトラッキングの



図1: 会話エージェントの画面



図2: ヘッドトラッキング動作例

様子を示す。本システムにより、頭部動作は30fpsで記録された。各ユーザのビデオの記録時間は9分から19分程度であり、合計時間は約115分、ヘッドトラッキングデータは合計約19万個計測された。

2.2 ヘッドトラッキングデータの整形

頭部の位置と回転角度のデータをグラフ化したところ、上下方向の細かいノイズを除去するために、データのスムージングを行った。具体的には、データ10個幅で移動平均を算出し、さらに、データ5個の幅で、細かい上下動をならす処理を行った。

2.3 会話参加態度のラベリング

被験者の映像をコーパス収集には参加していない、別の被験者10人(観察者と呼ぶ)に見てもらい、それぞれ被験者が会話に飽きていると思った部分に、ビデオアノテーションツールを用いてラベル付けをしてもらった。その際、判断の参考として、(1)口元が動く、(2)ため息・深呼吸をする、(3)あくびをかみ殺す、(4)首がかたむく、(5)体重が移動する、(6)説明対象外のところを見る、などの変化に注目するよう指示した。

以下の節では、10人中3人以上の観察者が会話に飽きていると判断した場合と、4人以上の観察者が飽

Estimating the Conversational Engagement based on the User's Head Motion in Human-Agent Interaction

[†]Ryota OKO, Yukiko NAKANO,

Seikei University, Faculty of Science and Technology

きていると判断した場合について頭部動作との関連性を分析する、

2.4 頭部動作と会話参加態度との関連性

ヘッドトラッカから得られた頭部動作データを以下の指標に関して分析した。

- 位置情報：ヘッドトラッカから得た頭部の位置 (x, y, z)
- 回転角度：ヘッドトラッカから得た頭部の回転角度 (roll, yaw, pitch)
- 振幅：位置, 回転角度の各次元の変化を波の振幅とみなし, 各波の振幅を算出
- 周波数：位置, 回転角度の各次元の波の周波数

2.3 での会話参加態度のラベリングにおいて, 会話に飽きていると判定された時間と, 上記の指標により算出された値との関連性を分析したところ, 飽きていると判定されている時間では, x 軸 (左右) 方向の位置移動や roll (首の傾げ) の角度変化のグラフに大きな増加または減少が見られた。

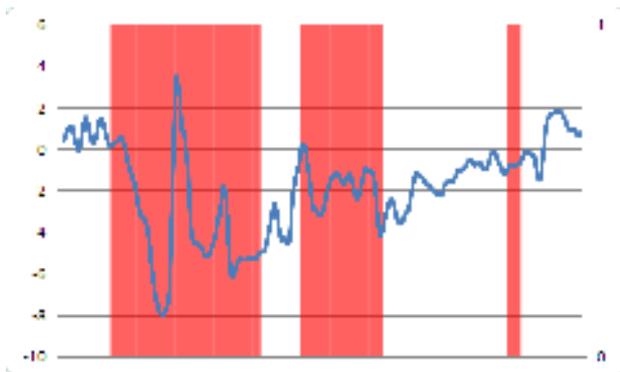


図3：roll の回転角度グラフ

図3は roll の回転角度グラフであり, 2.3 の会話参加態度のラベリングにおいて, 4 人以上の観察者が飽きていると判定した時間が影で示されている。観察者により飽きていると判断されている時間帯やその時間帯の冒頭部分で特に角度の変化が大きくなっていることが分かる。振幅や周波数についても同様にグラフ化したところ, 同じような特徴がみられた。

3. 頭部動作データからの会話参加態度の推定

2.4 で用いた指標を特徴値として, 観察者 3 人による判定, および 4 人による判定を予測する決定木学習を行った。エージェントの説明が画面上の次の説明対象に移行すると, 被験者は新たな説明対象の方に顔を向けるため, 頭部が大きく動く。これが飽き判定に影響してしまうと考えたため, エージェントによる説明対象が切り替わったか否かを特徴値に加えた。頭部の位置, 回転角度の値とそれぞれの振幅, 周波数, さらに説明対象が変わったタイミングの 19 個の特徴量を用いて会話参加態度を予測する

表 1：決定木学習の結果

	分類精度	F-Measure (True)	F-Measure (False)
3 人以上の判定	77.2%	77.7%	76.7%
4 人以上の判定	77.8%	77%	78.61%

決定木学習を行った結果を表 1 に示す。決定木学習には Weka による J48 で 10 回の交差検定を用いた。F-Measure による, 評価を行ったところ, 会話に飽きている (True) の予測, 会話に飽きていない (False) の予測の両方において, 全て 80%近い値が得られた。判別に最も大きく寄与していたのは回転角度の pitch であり, roll や x, z などの情報量も大きいことが分かった。

また, アイトラッカで視線が計測できていない時のデータをテストデータとして同様に J48 の決定木学習を行った場合も, 同程度の 80%近い値が得られた。

3. まとめ

本研究では, 頭部の動作と会話参加態度との関連性を分析し, 頭部の位置・回転角度, これらの振幅・周波数から, 会話中のユーザの飽きを判定する決定木を作成したところ, 80%近い精度が得られた。今後は本稿で提案した頭部動作に基づく参加会話態度推定方法と, [1]による注視行動に基づいた推定方法とを統合することにより, リアルタイムにより頑健性の高い会話参加態度推定機構を構築する予定である。その際, アイトラッカの計測状況によりどの情報をより重視すべきか等の検討や, 個人差を考慮したモデルへの拡張についても検討してゆく。

謝辞

本研究で使用した頭部姿勢推定システムの一部には, オムロン株式会社の OKA0 Vision 技術を利用している。ここに記して感謝する。

本研究は, 科研費特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」, 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] 石井亮, 中野有紀子: ユーザの注視行動に基づく会話参加態度の推定-会話エージェントにおける適応的会話制御に向けて-, 情報処理学会論文誌, Vol. 49, No. 12, pp. 3835-3846, 2008.
 [2] 岡 兼司, 佐藤 洋一, 中西 泰人, 小池 英樹, "適応的拡散制御を伴うパーティクルフィルタを用いた頭部姿勢推定システム", 電子情報通信学会論文誌 D-II, vol. J88-D-II, no. 8, August 2005.