

人物対話行動における応答時間分布抽出

佐藤研究室 木村綾平

1 緒論

近年、システムによって人の意図や興味を推測するため、対話行動を認識・理解するための研究が行われている [1, 2]. しかし、対話行動認識は音声・画像認識と異なり、マルチモーダルデータの統合的利用や対話意図を考慮する必要があるなど難しい面が多く、実用化に至っているものはほとんどない. 本研究ではセンサデータから自動で抽出される行動素と、対話意図が影響を与え易い対話者間の応答時間を用いた対話行動構造化モデルを提案し、対話行動認識システムの実現を目指す.

2 対話行動構造化モデル

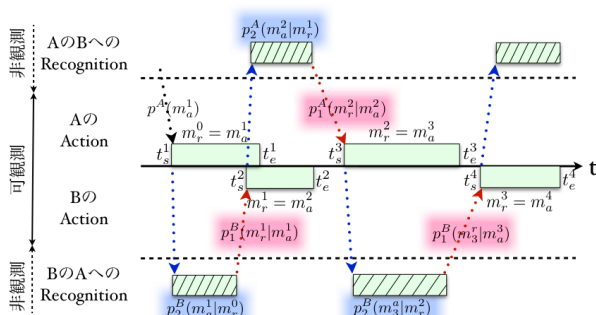


図 1: 対話行動構造化モデル

本研究で提案する対話行動構造化モデルを図 1 に示す. 人物対話行動を (A の Action → B の Recognition → B の Reaction → A の Recognition) といった部分対話行動の繰り返しとして、以下の三つの確率を用いて記述する、

- 初期行動素確率 $p_0^A(m_a^1|m_r^0)$
- 行動素遷移確率 $p_1^B(m_r^i, t^i|m_a^i, t^i)$
時刻 t^i において A による先行行動素 m_a^i が観測されたときに、ある時刻 t^i において B による後続行動素 m_r^i が観測される確率.
- 対話継続確率 $p_2^B(m_a^{i+1}, t^{i+1}|m_r^i, t^i)$
状態 i で B が先行行動素 m_a^i を受けた際、B の意図や感情、社会的立場、周囲の状況などから影響を受けた後続行動素 $m_r^i = m_a^{i+1}$ の出現確率.

これら対話遷移・継続確率の繰り返しによって、観測対話行動系列は以下の確率連鎖モデルに変換される.

$$p^A(m_a^1|m_r^0)p_2^B(m_a^1|m_r^0)p_1^B(m_r^1|m_a^1) \cdots p_2^B(m_r^N|m_a^{N-1}) \quad (1)$$

学習データを全て式 (2.4) で表される状態に変換し、対話行動毎に

$$D^{IM} = \arg \max_{\{t^i\}} \prod_{i=1}^N p_1^B(m_r^i, t^i|m_a^i, t^i)p_2^B(m_a^i, t^i|m_r^{i-1}, t^i) p_1^A(m_r^{i-1}, t^i|m_a^{i-1}, t^i)p_2^A(m_a^{i-1}, t^i|m_r^{i-2}, t^i)$$

となる $\{t^i\}$ を抽出する.

3 実験と結果

提案モデルに用いる行動素遷移確率の学習のために受付対話行動からの応答時間分布抽出を行った. 図 2 は抽出された行動素と応答時間分布の一例である.

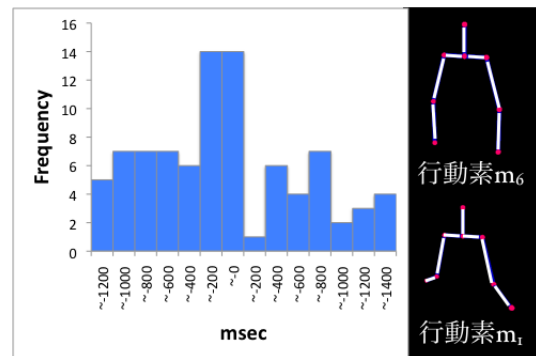


図 2: 行動素間応答時間ヒストグラム

4 結論

行動素と応答時間分布を用いた対話行動構造化手法を提案し、実験により行動素遷移確率の抽出を行った. 今後の課題として、提案モデルへの音声データの利用と、提案モデルにおける対話者の意図や興味といった暗示的情報による対話行動遷移確率の学習方法の考案が挙げられる.

参考文献

[1] 福間良平, 角康之, 西田豊明: “人のインタラクションに関するマルチモーダルデータからの時間構造発見,” 電子情報通信学会技術研究報告, vol.109, pp.119-126, 2009.

[2] 川嶋宏彰, 西川猛司, 松山隆司: ”落語の役柄交替における視覚的「間合い」の解析,” 情報処理学会論文誌, vol.48, No.12, pp.3715-3728, 2007.