

平成 30 年 6 月 11 日現在

機関番号：13302

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00300

研究課題名(和文) マルチモーダル会話モデリングに基づくグループ意思決定プロセスの解析

研究課題名(英文) Analyzing group decision making process based on multimodal conversation modeling

研究代表者

岡田 将吾 (Okada, Shogo)

北陸先端科学技術大学院大学・先端科学技術研究科・准教授

研究者番号：00512261

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究の目的は、意思決定を行うためのグループ会議に焦点を当て、会議中に交わされる会話者の言語・非言語情報から、意思決定が行われるまでのプロセスを会話データより客観的に解析出来る技術を新規に開発することである。言語・非言語情報(発話、音声、動作)と、アノテーションされたグループディスカッションの質を紐づけることで、良質なディスカッションに特有に現れる非言語情報や、コミュニケーション能力の高い人に見られる特有の言語パターンを抽出出来るフレームワークを構築した。機械学習の Support Vector Machine を用いた結果、最大82%の精度で推定するモデルを構築した。

研究成果の概要(英文)：This research project focuses on group discussion for a problem solving and develops a framework for analyzing the group decision making process based on verbal and nonverbal (multimodal) information which observed from group members. We defined the quality of group output as an index set of social science ("product dimension"), which proposed by Hackman. The annotation data of the quality of group output has been collected. The machine learning model it developed to predict the product dimension from multimodal information including dialog transcription, head motion, speech prosody and turn taking. Novel co-occurrence data mining is proposed to capture the group interaction and multimodal patterns. Through the machine learning modeling and data mining, the specific multimodal features observed in group discussion process with high/low quality can be discovered automatically. Best prediction accuracy of product dimension is 82% in binary classification task (high or low of quality).

研究分野：知能情報処理

キーワード：マルチモーダルインタラクション 社会的信号処理 機械学習 データマイニング グループディスカッション

1. 研究開始当初の背景

グループで協調し、議論を通じて意思決定・問題解決を行う能力の向上は重要な課題であり、これを支援する情報技術の開発が重要である。能力向上のためには、自身のグループの会話と他者の会話を客観的に比較し、他者との違いを理解できる技術が必要である。

社会学における会話分析の知見を基盤に会話中に交わされる発話内容の言語情報、非言語情報、会話構造を会話データから自動または半自動で抽出し・モデル化する研究が行われてきた。合意形成過程の解明を試みた研究<2>もあったが、分析の多くは人手で行われた。会話内容・非言語情報を簡易に記録するツール<3><4>も開発されていたが、会話プロセス解析自体はユーザに委ねられており、自動解析技術の提案はなされていなかった。

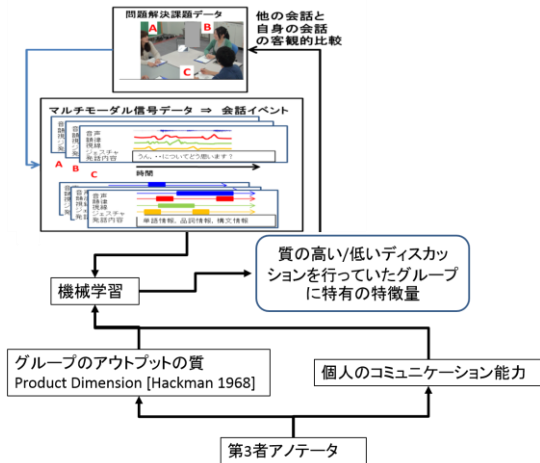


図1：本研究の概要

2. 研究の目的

本研究の目的は、意思決定を行うためのグループ会議に焦点を当て、会議中に交わされる会話者の言語・非言語情報から、意思決定が行われるまでのプロセスを会話データより客観的に解析出来る技術を、社会的信号処理技術<1>に基づき新規に開発することである。

グループのアウトプットと、グループディスカッションに参加する個人のコミュニケーション能力を社会学の指標より定義し、ディスカッションデータにアノテーションを行う。またグループディスカッションで交わされる言語・非言語情報（発話、音声、動作）と、アノテーションされたグループディスカッションの質やコミュニケーション能力を紐づけることで、良質なディスカッションに特有に現れる非言語情報や、コミュニケーション能力の高い人に見られる特有の言語パターンを抽出出来るフレームワークを構築する。この技術を用いることで、良質なミーティングプロセスで観測される議論シーンや非言語行動を自身の行動と比較することが可能となり、ミーティングの振り返りや、自身のミーティングと効率的かつ・満足いく結論が生まれた良質なミーティングプロセスとの違いを把握出来る。

3. 研究の方法

機械学習・データマイニングを用いたモデリングを通じて、会話者の言語・非言語情報から、意思決定が行われるまでのプロセスを会話データより解析するモデルの構築した。

・会話プロセスの比較解析モデル

入力 (X) としてマイク・モーションセンサを含む各種センサから取得される会話者の言語・非言語情報を、出力 (Y) として会話者のコミュニケーション能力、グループのコミュニケーションの質を定義し、X から Y を出力する機械学習モデルを構築する。このモデリングを通じて、コミュニケーション能力の高い/低い人、グループディスカッションのアウトプットの高い/低いグループを言語・非言語情報のパターンを可視化することで理

解できる。

・データ収集とアノテーション

4名による合意形成型ミーティングを30セッション行い、マイク・カメラ・モーションセンサを利用してマルチモーダル時系列データコーパスを収録した。このコーパスはMATRIXコーパスと呼ばれ、申請者が共同研究者として参画した基盤B(研究課題25280076)のプロジェクトの予算で収集された。本研究はこのデータコーパスに基づき研究を行った。本研究費で、グループアウトプットのアノテーションと、ディスカッションのプロセスを捉えるための言語行為のアノテーションを実施した。

4. 研究成果

4.1 グループディスカッション参加者のコミュニケーション能力の推定モデルの構築と分析

本研究では、会話参加者の表出するマルチモーダル情報から「コミュニケーション能力」を推定するモデルの構築・評価を行った。発話内容に含まれる単語の品詞、発話ターン、韻律情報、頭部動作量をマルチモーダル特徴量として抽出し、人事採用経験者より評定された各参加者のコミュニケーション能力値を推定するモデルを機械学習により構築した。評価実験の結果、総合的なコミュニケーション能力に関する評定値の推定に関して、回帰タスクで最大0.62の決定係数 R^2 、高群・低群の2クラスの分類タスクで最大0.93の精度を得た。また、各要素項目と総合的なコミュニケーション能力を識別するために有効な特徴量を明らかにし、コミュニケーション能力の高・低に分類され

る参加者に見られる特徴を明らかにした。

さらに、ディスカッションタスク、コミュニケーション能力、言語・非言語情報の関係を明らかにするために、タスクの種類別にコミュニケーション能力の推定モデルを構築・評価した。最初に、相関分析の結果、評定されたコミュニケーション能力値はタスクの種類に対して一定であり、能力の高い/低いと評価される人はタスクによって変わらないことを示した。様々な組み合わせの言語・非言語特徴量から、各タスクにおいて評定されたコミュニケーション能力の推定を行った結果、インバケット型課題であり、事前情報・タスク設定に制約を受けやすいタスク1では、発話ターンの特徴量のみを用いたモデルが最大の精度だった。一方、事前情報が少なく、議論進行の自由度の高いケーススタディ型のタスク2、3では言語特徴量、とりわけ談話行為タグの特徴量が有効であることが明らかになった。

4.2 グループディスカッションのアウトプットの推定モデル構築と分析

ディスカッションのアウトプットの評価方法を確立するため、社会心理学の知見をサーベイし、Hackmanにより提案されたProduct dimensions (PD)と呼ばれるプロダクトの評価指標を選定した。PDは建設的、アイデアの新規性、綿密さ、肯定的な内容か否か、といった19個の評価軸が定義されており、1-7のリッカード尺度で評定される。客観的な評定を行うために、グループディスカッションの参加者以外の第三者4名にアノテーションを依頼した。アノテータ間の評定一致率を計算した結果、クローンバック α 値は0.42

～0.72であった。この4名の評定したスコアの総和をグループディスカッションのアウトプットの数値と定義した。会話参加者の言語情報、非言語情報（頭部動作、音声、発話ターン）、コミュニケーション能力のスコアを入力として、各グループのディスカッションのアウトプットの質（PD）を推定するための機械学習モデルの構築・評価を行った。

結果として、アウトプットの 1. 創造性, 2. 理解可能かつ明快な度合い, 3. 扱っている問題に対処できていない程度, 4. 考え方, 物, 事に難色を示す度合い, 5. 綿密さの欠如, 特定の信念や意見への偏り, 6. ありふれたアイデアである度合いの6つの項目の評定値を 0.7 (70%)以上の精度で推定するモデルを構築した。線形モデルの各特徴量の重みを分析することにより、コミュニケーション能力の値、発話ターン、談話行為の遷移が重要であることが示された。

4.3 会話参加者のインタラクション・参加者の非言語マルチモーダルパターンを自動抽出するためのデータマイニング手法の構築

グループアウトプットの質やコミュニケーション能力に関わる多種多様な言語・非言語パターンを抽出するためには、人間同士のインタラクション・人間が表出するマルチモーダルパターンをボトムアップに抽出するモデルが有用である。多人数のマルチモーダル時系列データから、効率的にインタラクション・マルチモーダルパターンを抽出するグラフクラスタリングに基づくデータマイニングの手法を提案し、このマイニング手法と latest fusion と呼ばれるデータ統合方法を併用することで、高精度にリーダーシップを

含むコミュニケーション能力を推定する手法を提案した。この手法を用いることで、インタラクションを詳細に分析できることを示した。例えば、コミュニケーション能力の高い人が話した際、他の人が視線を向けやすい。重要でない発言が行われた際、他の人は説明者を向かない等の特徴量をマイニング可能である。

(参考文献)

<1>Pentland, Alex. "Social signal processing." *IEEE Signal Processing Magazine* 24.4 (2007): 108.

<2>片桐恭弘 他, 会話における合意形成と相互信頼感形成. SIG-SLUD-B001: 49-54. 2010

<3> 対面式会議コンテンツの作成と議論中におけるメタデータの可視化, 土田 貴裕 他 情報処理学会論文誌, Vol. 51, No. 2, pp. 404-416, 2010

<4>角 康之, 他, : 協調的なインタラクションの記録と解釈, 情報処理学会論文誌, Vol. 44, No. 11, pp. 2628-2637, 2003年11月

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文 (全て査読付き)] (計 14 件)

[1] Youwei Lu, Shogo Okada, Katsumi Nitta, Modeling Content Structures of Domain-specific Texts with RUP-HDP-HSMM and Its Applications, IEICE Transactions, Vol. E100-D, No. 9, pp. 2126-2137 (2017)

[2] 岡田将吾, 松儀良広, 中野有紀子, 林佑樹, 黄宏軒, 高瀬裕, 新田克己 マルチ

モーダル情報に基づくグループ会話におけるコミュニケーション能力の推定
人工知能学会論文誌, Vol131, No. 6
AI30-E [人工知能学会創設 30 周年記念
特集論文最優秀賞]

- [3] Shogo Okada, Mi Hang, Katsumi Nitta, ”
Predicting performance of
collaborative storytelling using
multimodal analysis”, IEICE
Transactions, 6ED-55, pp. 429-439
(2016)

[学会発表] (計 28 件)

- [1] Shogo Okada, Kazuhiro Otsuka:
Recognizing Words from Gestures:
Discovering Gesture Descriptors
Associated with Spoken Utterances. FG
2017: 430-437
- [2] Youwei Lu, Shogo Okada, Katsumi Nitta:
Weibull partition models with
applications to hidden semi-Markov
models. IEEE IJCNN 2017: 162-169
- [3] Shogo Okada, Yoshihiko Ohtake, Yukiko
I. Nakano, Yuki Hayashi, Hung-Hsuan
Huang, Yutaka Takase, Katsumi Nitta:
Estimating communication skills using
dialogue acts and nonverbal features
in multiple discussion datasets, 18th
ACM International Conference on
Multimodal Interaction (ICMI), Tokyo,
Nov. 2016.
- [4] Shogo Okada, Oya Aran and Daniel
Gatica-Perez, Personality Trait
Classification via Co-Occurrent
Multiparty Multimodal Event Discovery,
17th ACM International Conference on

Multimodal Interaction (ICMI),
Seattle, Nov. 2015.

- [5] Yukiko I. Nakano, Sakiko Nihonyanagi,
Yutaka Takase, Yuki Hayashi, Shogo
Okada, Predicting Participation
Styles using Co-occurrence Patterns
of Nonverbal Behaviors in
Collaborative Learning, 17th ACM
International Conference on
Multimodal Interaction (ICMI),
Seattle, Nov. 2015.

- [6] 岡田 将吾, 大竹 圭彦, 中野 有紀子,
林 佑樹, 黄 宏軒, 高瀬 裕, 新田 克
己, マルチモーダルモデリングに基づ
くコミュニケーション能力とグループ
ディスカッションタスクの関連性分析,
HCG シンポジウム 2016 [優秀インタラク
ティブ発表賞]

- [7] 岡田将吾, 米航, 新田克己, 説明行為の
質の推定に向けた会話者のマルチモー
ダル情報モデリング, 人工知能学会 言
語・音声理解と対話処理研究会 (SLUD)
第 74 回研究会 [2015 年度人工知能学会
研究会優秀賞]

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 1 件)

名称: コミュニケーション能力推定装置、
方法およびプログラム

発明者: 中野有紀子, 岡田将吾, 松儀良
広, 林佑樹, 黄宏軒, 高瀬裕, 新田克己
権利者: 学校法人成蹊学園

種類: 特許

番号: 特願 2016-128904

出願年月日: 2016 年 6 月 29 日

国内外の別：国内

〔その他〕

ホームページ等

<http://www.jaist.ac.jp/~okada-s/SocialActivity.html>

6. 研究組織

(1) 研究代表者

岡田 将吾 (Shogo Okada)

北陸先端科学技術大学院大学・先端科学技術

研究科・准教授

研究者番号：00512261