

令和 6 年 6 月 6 日現在

機関番号：12701

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K12011

研究課題名（和文）マルチモーダル非言語行動の機能スペクトラム解析に向けた計算基盤の構築

研究課題名（英文）Building Computational Framework Toward Functional Spectrum Analysis of Multimodal Nonverbal Behaviors

研究代表者

大塚 和弘 (Otsuka, Kazuhiro)

横浜国立大学・大学院工学研究院・准教授

研究者番号：30396197

交付決定額（研究期間全体）：（直接経費） 4,700,000円

研究成果の概要（和文）：人の対話において非言語行動は重要な役割を担うものの、その曖昧性・多様性故に従来、定量的な扱いが困難であった。本研究は、そのような非言語行動を定量的に解析するため、非言語行動がもつ意味や機能の分布強度を表す『非言語機能スペクトラム』という新概念を提唱し、複数モダリティの非言語行動を統一的に分析・認識するための計算基盤を構築した。非言語モダリティとして、頭部運動や顔表情、視線、相槌に着目し、対話者の顔や身体の運動、音声信号などの観測情報から、各モダリティの各機能の現出を自動推定する種々の機械学習モデルを構築した。また推定された機能スペクトラムから対話者の主観的印象を予測・説明する方法を開発した。

研究成果の学術的意義や社会的意義

本研究では、人のコミュニケーションにおいて重要な役割を果たす、複数の非言語モダリティを介して生じる多数の機能を統一的に表現・分析・認識することが可能な計算基盤を構築した。この計算基盤は、人の繊細かつ複雑で、極めて多様性に富んだ感情や意図の表出・交換の過程を定量的に解明することに貢献すると期待される。さらに将来的には、実社会において、人の気持ちに寄り添い、人を支援し、協業することが可能な人工社会知能（Artificial Social Intelligence）を構築するための基盤技術として活用されることが考えられる。

研究成果の概要（英文）：Although nonverbal behavior plays important roles in human conversations, it has been difficult to analyze them quantitatively due to their ambiguity and diversity. To analyze such nonverbal behaviors quantitatively, this research proposed a new concept of "nonverbal functional spectrum," representing the distribution of perceived intensity of nonverbal behavior's functions, and constructed a unified computational framework. Focusing on head movement, facial expression, eye gaze, and Aizuchi as non-verbal modalities, we constructed machine-learning models that automatically estimate the occurrence of each function for each modality from observed information, such as the interlocutor's facial and body movements and speech signals. We also developed a method for predicting and explaining the interlocutor's subjective impressions based on the estimated functional spectrum.

研究分野：知能情報学

キーワード：対話 非言語コミュニケーション マルチモーダル 社会的信号処理 機械学習

1. 研究開始当初の背景

(1) SDGs や Society5.0、新型コロナウイルスなどの社会課題の解決には、社会の成員による対話を介した合意形成と集合知創出が重要であり、対面・遠隔を問わず人々の対話活動を支援し、人との高度な知的協調を可能とする人工知能(AI)が要請されている。近年、ChatGPT に代表される大規模言語モデルにより、人工知能の言語理解・運用能力が飛躍的に向上している。しかし、「知・情・意」と称される総合的な人の知能の内、「情」や「意」に対応する「社会的知能」(Social Intelligence)の実現は、言語的知能に比べて立ち遅れている。社会的知能の実現には、言語だけではなく、顔表情や頭部運動、視線、身体動作、韻律などの非言語行動に含まれる多様かつ微細な情報のやり取りが重要な役割を果たすとされ、それらの非言語行動の認識や意味理解を行う技術の開発が急務とされている。

2. 研究の目的

本研究では、人と人とのコミュニケーションにおいて重要な役割を果たす、複数の非言語モダリティを介して達成される「機能」という観点から、対話現象を質的・定量的に分析可能な計算基盤の構築を目的とした。本課題は具体的には以下の部分問題から構成される。

(1) 対話において感情や意図の表出・交換を担う、繊細かつ複雑で、極めて多様性に富んだ非言語行動を捉えるため、本研究では、各々のモダリティがどのような機能を現出・含意しうるか(機能多重性) また、それが見るものによってどのように曖昧に解釈されるか(解釈曖昧性) さらにどのように複数のモダリティが連携して相乗効果を発揮するか(相乗作用)という問いに着目した。これらの問いに対して、我々の構築した非言語機能コーパスの分析により現象の様相を明らかにするとともに、観測可能な身体運動や音声信号からこれらの機能を自動認識することが可能な機械学習モデルを構築することを目的とした。

(2) 対話に参加する人物間、特に話し手と聞き手の間にて行われる、非言語行動の交換・連鎖によりどのような相互作用上の機能や意味が生じるかを明らかにすること、及び、これらを自動認識することが可能な機械学習モデルの構築も併せて目的とした。

(3) さらに、上記の非言語行動がもつ機能と対話者の内面感情とが深い関係性をもつという仮定のもと、非言語行動の機能の認識結果から、対話者の主観的印象を予測するという問いを立て、機械学習モデルの構築を行うことも目的とした。

3. 研究の方法

(1)本研究では、対話者の非言語行動を定量的に解析するため、非言語行動がもつ意味や機能の分布強度を表す『非言語機能スペクトラム』という新概念を提唱した。非言語機能スペクトラムとは、表出行動が持ちうる意味や機能の知覚強度の分布を表したものであり、一つの行動表出が同時に複数の意味・機能を含意しうる機能多重性、並びに、受け手側の解釈の曖昧性を捉えることを意図した表現である。具体的に機能スペクトラムは、発話の強調、問い掛け、反応確認、同意要求、傾聴、思考、理解、同意/非同意、感情表出などを構成要素とし、本研究では、主要なモダリティとして、頭部運動、顔表情、視線、相槌を対象とした。

(2)各モダリティの各機能の現出を自動推定するため、対話者の顔や身体の運動、音声信号などの観測情報を入力とする機械学習モデルを構築した。また、複数モダリティが連携して発揮する機能を分析するため、頭部運動と顔表情、頭部運動と相槌、それぞれの間の機能の連関構造を調査し、複数の相乗機能を定義し、これらを自動認識するモデルを構築した。

(3)また、話し手・聞き手間の頭部運動を介した相互作用により現出する機能を調査し、これらを認識する深層転移モデルを構築した。

(4)さらに、機能スペクトラムから対話者の主観的印象を予測・説明する方法を開発した。複数の機能の出現頻度の比率などを特徴量とし、これと対話者の感じた主観的印象(楽しさなど)との関連性を解析し、非言語行動の機能スペクトラムが対話者の主観的印象を推測する手がかりとして有用かどうか検証を行った。

4. 研究成果

本研究により、対話者の非言語行動をその機能的側面から認識・理解するための計算基盤が構築された。その核として、非言語行動がもつ意味や機能の分布強度を表す『非言語機能スペクトラム』という新概念を提唱し、これに基づき前述の目的を達成した。

(1)頭部運動、顔表情、視線、相槌といった複数の非言語モダリティを対象として、それぞれのモダリティの非言語行動が発揮しうる機能集合を定義し、それらの内、頻出する機能群を分析により見出した。これらを対象として、観測可能な身体運動、音声信号から各時刻における機能の現出を認識するための深層学習モデルを構築した。各々のモダリティについて、非言語機能を認識する深層学習モデルをそれぞれ論文誌上において発表している(代表的文献、頭部運動、顔表情、視線、相槌)。

(2)複数モダリティの連携によって発揮される相乗機能を定義・分析するための枠組み、及び、機械学習モデルを構築した。具体的には、頭部運動と顔表情の連携により発揮される相乗機能を定義するため、各々のモダリティの各機能ペアの共起関係に基づく方法(文献)及び、複数のモダリティにまたがる多機能間の連携を捉えるためのスペクトラム分解に基づく方法を開発した(文献)。加えて、頭部運動と声の相槌の同時出現により発揮される機能を分析し、代表的な相乗機能を定義し、さらにこれらを頭部運動や声の韻律特徴から自動認識する方法を提案した(文献)。

(3)話し手・聞き手間の頭部運動を介した相互作用により現出する機能を調査し、話し掛けに対する相槌のような典型的な相互作用の機能を見出し、これらを認識する深層転移モデルを構築した(文献)。

(4)さらに非言語行動の機能スペクトラムから対話者の主観的印象を予測・説明するモデルを開発し、非言語機能スペクトラムが、対話者の主観的印象を推測する手がかりとして有用であることを明らかにした(文献)。また、印象形成に寄与した行動とその時刻を説明可能なモデルを構築した(文献)。

<引用文献>

武田一輝, 大塚和弘:「対話中の頭部運動機能認識のための特徴膨張収縮深層ニューラルネットワーク」, 電子情報通信学会論文誌 A, Vol. J105-A, No. 2, pp26-37, 2022

武田一輝, 今村まい, 大塚和弘:「対話中の頭部運動機能が持つ解釈の曖昧性を推定する深層回帰モデル」, 電子情報通信学会論文誌 A, Vol. J106-A, No.10, pp.255-266, 2023

城戸祐世, 杉山隼大, 大塚和弘:「言語特徴を利用した対話中の頭部運動機能の認識」, 電子情報通信学会論文誌 A, Vol. J107-A, No.1, pp.6-18, 2024

今村まい, 武田一輝, 熊野史朗, 大塚和弘:「対話中の頭部運動と顔表情の相乗機能の認識」, 電子情報通信学会論文誌 A, Vol. J106-A, No.3, pp.70-87, March 2023

田代絢子, 今村まい, 熊野史朗, 大塚和弘:「マルチモーダル非言語行動に基づく対話者の視線機能の認識」, 電子情報通信学会論文誌 A, Vol. J106-A, No.12, pp.296-311, 2023,

飯塚海斗, 大塚和弘:「対話中における聞き手の頭部運動と相槌の相乗機能の解析」, 人工知能学会論文誌, Vol. 38, No. 3, 2023

Mai Imamura, Ayane Tashiro, Shiro Kumano, and Kazuhiro Otsuka, "Analyzing Synergetic Functional Spectrum from Head Movements and Facial Expressions in Conversations," in Proceedings of 25th ACM International Conference on Multimodal Interaction (ICMI2023), pp.42-50, 2023

森 駿志, 大塚和弘:「対話中の頭部運動相互作用の機能認識のための深層転移学習」, 電子情報通信学会論文誌 A, Vol. J105-A, No.10, pp.111-124, Oct. 2022

大土隼平, 石井陽子, 中谷桃子, 大塚和弘:「複数人対話における頭部運動機能に基づく対話参加者の主観的印象の予測」, 電子情報通信学会論文誌 D, Vol. J105-D, No.09, pp.504-517, Sep. 2022

大土隼平, 三好一輝, 石井陽子, 石井 亮, 永徳真一郎, 大塚和弘:「頭部運動機能特徴に基づく対話者の主観的印象の予測・説明モデルの構築」, 人工知能学会論文誌, Vol. 38, No. 3, 2023

5. 主な発表論文等

〔雑誌論文〕 計10件（うち査読付論文 10件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 大土 隼平、石井 陽子、中谷 桃子、大塚 和弘	4. 巻 J105-D
2. 論文標題 複数人対話における頭部運動機能に基づく対話参加者の主観的印象の予測	5. 発行年 2022年
3. 雑誌名 電子情報通信学会論文誌D 情報・システム	6. 最初と最後の頁 504 ~ 517
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transinfj.2021JDP7064	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 森 駿志、大塚 和弘	4. 巻 J105-A
2. 論文標題 対話中の頭部運動相互作用の機能認識のための深層転移学習	5. 発行年 2022年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 111 ~ 124
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2021JAP1037	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 丸山 光貴、森 駿志、大塚 和弘	4. 巻 J106-A
2. 論文標題 遠隔対話者間の相互作用に対する頭部運動再現機構の機能的効果の分析	5. 発行年 2023年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 88 ~ 103
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2022HAP0009	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 今村 まい、武田 一輝、熊野 史朗、大塚 和弘	4. 巻 J106-A
2. 論文標題 対話中の頭部運動と顔表情の相乗機能の認識	5. 発行年 2023年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 70 ~ 87
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2022HAP0007	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 大土 隼平、三好 一輝、石井 陽子、石井 亮、永徳 真一郎、大塚 和弘	4. 巻 38
2. 論文標題 頭部運動機能特徴に基づく対話者の主観的印象の予測・説明モデルの構築	5. 発行年 2023年
3. 雑誌名 人工知能学会論文誌	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1527/tjsai.38-3_H-M74	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 飯塚海斗、大塚和弘	4. 巻 38
2. 論文標題 対話中における聞き手の頭部運動と相槌の相乗機能の解析	5. 発行年 2023年
3. 雑誌名 人工知能学会論文誌	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1527/tjsai.38-3_J-M91	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 武田 一輝、大塚 和弘	4. 巻 J105-A
2. 論文標題 対話中の頭部運動機能認識のための特徴膨張収縮深層ニューラルネットワーク	5. 発行年 2022年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 26～37
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2021JAP1016	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 武田 一輝、今村 まい、大塚 和弘	4. 巻 J106-A
2. 論文標題 対話中の頭部運動機能をもつ解釈の曖昧性を推定する深層回帰モデル	5. 発行年 2023年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 255～266
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2023JAP1005	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 田代 絢子、今村 まい、熊野 史朗、大塚 和弘	4. 巻 J106-A
2. 論文標題 マルチモーダル非言語行動に基づく対話者の視線機能の認識	5. 発行年 2023年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 296～311
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2023JAP1003	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 城戸 祐世、杉山 隼大、大塚 和弘	4. 巻 J107-A
2. 論文標題 言語特徴を利用した対話中の頭部運動機能の認識	5. 発行年 2024年
3. 雑誌名 電子電子情報通信学会論文誌A 基礎・境界	6. 最初と最後の頁 6～18
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transfunj.2023JAP1010	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計14件 (うち招待講演 0件 / うち国際学会 6件)

1. 発表者名 武田 一輝, 大塚 和弘
2. 発表標題 対話中の頭部運動機能の解釈の曖昧性を推定する深層回帰モデルの検討
3. 学会等名 第21回情報科学技術フォーラム(FIT2021)
4. 発表年 2022年

1. 発表者名 大土隼平, 三好一輝, 石井陽子, 石井 亮, 永徳真一郎, 大塚和弘
2. 発表標題 非言語機能特徴に基づく対話者の主観的印象の予測・説明モデルの構築
3. 学会等名 人工知能学会全国大会 (JSAI2022)
4. 発表年 2022年

1. 発表者名 今村まい, 武田一輝, 熊野史朗, 大塚和弘
2. 発表標題 対話中に現れる顔表情と頭部運動の相乗機能の分析と認識モデルの構築
3. 学会等名 電子情報通信学会HCGシンポジウム2022
4. 発表年 2022年

1. 発表者名 Shumepi Otsuchi, Yoko Ishii, Momoko Nakatani, Kazuhiro Otsuka
2. 発表標題 Prediction of Interlocutors' Subjective Impressions Based on Functional Head-Movement Features in Group Meetings
3. 学会等名 23rd ACM International Conference on Multimodal Interaction (ICMI2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Kazuki Takeda, Kazuhiro Otsuka
2. 発表標題 Inflation-Deflation Networks for Recognizing Head-Movement Functions in Face-to-Face Conversations
3. 学会等名 23rd ACM International Conference on Multimodal Interaction (ICMI2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Takashi Mori, Kazuhiro Otsuka
2. 発表標題 Deep Transfer Learning for Recognizing Functional Interactions via Head Movements in Multiparty Conversations
3. 学会等名 23rd ACM International Conference on Multimodal Interaction (ICMI2021) (国際学会)
4. 発表年 2021年

1. 発表者名 大塚和弘
2. 発表標題 対話中の非言語行動の機能スペクトラム解析に向けて～頭部運動を対象とした機能コーパスと深層学習モデルの構築事例～
3. 学会等名 電子情報通信学会 ヒューマンコミュニケーション基礎研究会 (HCS)
4. 発表年 2021年

1. 発表者名 大土隼平, 石井陽子, 中谷桃子, 大塚和弘
2. 発表標題 頭部運動機能特徴を用いた複数人対話における対話参加者の主観的印象の予測
3. 学会等名 電子情報通信学会 ヒューマンコミュニケーション基礎研究会 (HCS)
4. 発表年 2021年

1. 発表者名 武田一輝, 大塚和弘
2. 発表標題 対話中の頭部運動機能を認識する特徴膨張収縮ニューラルネットワーク
3. 学会等名 第20回情報科学技術フォーラム(FIT2021)
4. 発表年 2021年

1. 発表者名 森 駿志, 大塚和弘
2. 発表標題 複数人対話における深層転移学習による頭部運動相互作用の機能認識
3. 学会等名 電子情報通信学会HCGシンポジウム2021
4. 発表年 2021年

1. 発表者名 Shumpei Otsuchi, Koya Ito, Yoko Ishii, Ryo Ishii, Shin-ichiro Eitoku, and Kazuhiro Otsuka
2. 発表標題 Identifying Interlocutors' Behaviors and its Timings Involved with Impression Formation From Functional Head-Movement Features and Linguistic Features
3. 学会等名 25th ACM International Conference on Multimodal Interaction (ICMI2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Mai Imamura, Ayane Tashiro, Shiro Kumano, Kazuhiro Otsuk
2. 発表標題 Analyzing Synergetic Functional Spectrum from Head Movements and Facial Expressions in Conversations
3. 学会等名 25th ACM International Conference on Multimodal Interaction (ICMI2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Ayane Tashiro, Mai Imamura, Shiro Kumano, Kazuhiro Otsuka
2. 発表標題 Analyzing and Recognizing Interlocutors' Gaze Functions from Multimodal Nonverbal Cues
3. 学会等名 25th ACM International Conference on Multimodal Interaction (ICMI2023) (国際学会)
4. 発表年 2023年

1. 発表者名 田代絢子, 今村まい, 熊野史朗, 大塚和弘
2. 発表標題 複数モダリティの非言語情報に基づく対話者の視線機能の分析と認識
3. 学会等名 電子情報通信学会HCGシンポジウム2023
4. 発表年 2023年

〔図書〕 計0件

〔出願〕 計2件

産業財産権の名称 行動推定装置、行動推定方法および行動推定プログラム	発明者 大土隼平, 大塚和弘, 三好一輝, 石井陽子, 石井亮, 永徳真一郎	権利者 国立大学法人横 浜国立大学、日 本電信電話株式
産業財産権の種類、番号 特許、特願2022-080134	出願年 2022年	国内・外国の別 国内

産業財産権の名称 主観的印象予測手法	発明者 大塚和弘、大土隼 平、石井陽子、中谷 桃子、蔵内雄貴	権利者 国立大学法人横 浜国立大学、日 本電信電話株式
産業財産権の種類、番号 特許、特願2021-135890	出願年 2021年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------