

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 5月31日現在

機関番号：12608  
 研究種目：若手研究（B）  
 研究期間：2010～2012  
 課題番号：22700146  
 研究課題名（和文） インタラクションマイニングによる対話コンテキスト依存型ジェスチャ認識モデルの構築  
 研究課題名（英文） Context based Gesture Modeling based on Interaction Data Mining  
 研究代表者  
 岡田 将吾（OKADA SHOGO）  
 東京工業大学・大学院総合理工学研究科・助教  
 研究者番号：00512261

### 研究成果の概要（和文）：

本研究では会話中に表出するハンドジェスチャの役割を認識するためのモデルを構築した。提案する枠組みでは、ある会話参加者がジェスチャを行った時の非言語的な会話状態（他者の顔向け状態・頷き動作・発話状態）を利用することにより他者指向の説明に用いられるジェスチャを認識する。結果として、従来の手の動きの特徴量のみで学習したモデルよりも提案モデルを用いた場合、最大で約20%認識精度が向上した。

### 研究成果の概要（英文）：

This research proposed a conversation context based gesture recognition approach. In this research, we extract features of co-occurring nonverbal patterns with gestures, i.e., speech act, head gesture, and head direction of each participant, by using pattern recognition techniques. In the experiments, we collect eight group narrative interaction datasets to evaluate the classification performance. The experimental results show that gesture phase features and nonverbal features of other participants improves the performance to discriminate communicative gestures that are used in narrative speeches and other gestures from 20 %

### 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,200,000	360,000	1,560,000
2011年度	500,000	150,000	650,000
2012年度	500,000	150,000	650,000
年度			0
年度			0
総計	2,200,000	660,000	2,860,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：知識発見とデータマイニング，マルチモーダルパターン認識モデル

### 1. 研究開始当初の背景

対話中に表出される非言語情報（ジェスチャ）は人間の意図や対話状況の認識，インタラクションの理解に必要である。従来の社会言語

学の知見より，ジェスチャは談話構造や，対話相手の表出する言語・非言語情報に関連していることが知られているため，これらの要因を踏まえてジェスチャ認識モデルを構築する必要がある。

しかし従来会話中に自然と表出されるジェスチャの機能や役割のモデルに焦点を当てているものは少ない。また従来の認識モデルでは手の動作軌跡や形状に関する特徴量が用いられており、対話コンテキスト(対話相手の非言語情報や会話状態)を利用していない。これらは、従来のジェスチャ認識システムが人・機械間のインターフェイスに利用されることを想定していたため、対話コンテキストに依存して表出されるようなコミュニケーションに用いられるジェスチャを対象としていなかったことに起因する。

## 2. 研究の目的

本研究の目的は対話中に現れる非言語パターンを認識するシステムを構築することである。目標を達成するためには、会話中に交わされるジェスチャや手の動きの役割を識別するために有用な特徴量を抽出することである。自明でない上記の特徴量を抽出するために、データから頻出する非言語パターンと共起ルールを発見可能な非言語データマイニング機構を実現する。また、データマイニングおよび分析の結果から抽出された特徴量を利用して、ジェスチャ認識モデルを構築し、このモデルを評価する。

## 3. 研究の方法

### (1) 対話データの取得とデータ分析

#### ① 会話タスク

三話者による会話タスクを対象とした。このタスクにおいて二者は事前にアニメーションを観察し、このアニメーションを一者に言語・非言語情報を交えながら説明する。このタスクでは、アニメキャラクターの空間的な位置関係の遷移を表現する際に非言語情報が説明に多用されており、過去のジェスチャ分析研究でもこのタスクは盛んに用いられた。三話者のグループでこの課題を行うことにより、説明者以外の参加者が発話権をとるために手を動かしたり、同意を示すために手を動かしたりするジェスチャが観測された。また説明者のジェスチャはいつでも情景描写に用いられるとは限らず、説明の話速を整えるためのジェスチャや説明内容を思い出そうとして手が動くような言いよどみに付随するジェスチャなどが観測された。本研究では特に2名の説明者の非言語行動と聞き手の表出する非言語行動の組み合わせに着目してジェスチャの役割機能の自動分類を試みる。

#### ② センシングする非言語行動

非言語モダリティについては頭部の動き・手の動き・姿勢を光学式モーションキャプチャから、発話の有無をマイクから取得する。頭部の縦方向動作よりうなづき行動を、横方向動作より誰だ誰を向いているかという顔向け状態を検出する。

### (2) データマイニング手法の開発

対話データから非言語パターンを抽出するために、少量のアノテーションデータを利用してデータにアノテーションを行う半教師付き学習手法、時系列データであるジェスチャパターン群からクラスタを発見するためのクラスタリング手法、連続多次元時系列データから動作の変化点を検知する手法、連続時系列データから頻出するジェスチャパターンを発見する手法、同じ時間帯に共起する非言語パターンのペアを列挙する手法をそれぞれ開発する。

### (3) ジェスチャ認識手法の開発・評価

会話中に用いられるジェスチャの機能を分類するために有効な特徴量を抽出するため、(2)で開発した方法を利用した。抽出した特徴量を組み合わせて、ジェスチャの役割の認識精度を評価し、有効な特徴量を特定する。

## 4. 研究成果

本研究の成果概要を図1に示す。

### (1) 基盤データマイニングアルゴリズム

#### ① 半教師付アルゴリズム

少数のラベル付きアノテーションデータとラベルなしデータを利用して、距離学習の一種である AML に基づいて類似度グラフを構築し、さらに正則化を導入した学習を適用することにより、多クラスのパターンを高精度に分類可能な手法:GSSCML を提案した。

#### ② 時系列パターンのクラスタリング

時系列長の異なるジェスチャパターン群を高精度にクラスタリング可能な HB-SOINN アルゴリズムを提案した。本手法は追加学習可能な自己増殖型ニューラルネットワーク SOINN と HMM を統合することで、適切なクラスタ数も決定することが可能である。約 30 クラスのジェスチャデータセットのクラスタリング実験によりスペクトラルクラスタリングや混合 HMM といった従来手法よりも

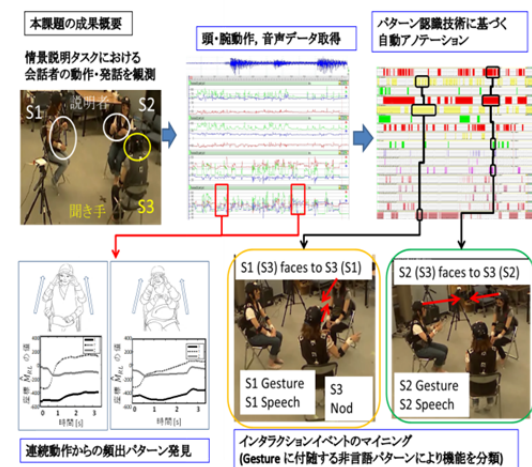


図 1: 成果概要

高精度にクラスタリング出来るだけでなく、正解に近いクラスタ数を出力できることを示した。

### ③多次元時系列の変化点検知手法

手の動きは、連続的な多次元時系列信号としてモーションセンサから観測される。1次元変化点検知アルゴリズム SWAB を多次元時系列に適用可能な形に拡張することで高精度に動作の開始点を検知するためのアルゴリズムを提案した。従来、多次元時系列データの分節化に利用されていた Online HMM Segmentation よりも高速かつ高精度に動作区間を分節化できることを実験より示した。

### ④頻出するパターンを発見する手法

時系列データ中に頻出するパターンを発見する手法を系列モチーフ発見アルゴリズムである SAX+Random Projection (SAXRP) アルゴリズムを基盤として構築した。SAXRP は一次元時系列からの頻出パターン発見アルゴリズムであるため、これを多次元時系列に適用出来るよう拡張した。重要な特徴次元を事前知識として利用することにより、各次元で共起するパターンを統合し多次元頻出パターンを抽出する手法を提案した。この手法は手の動作時系列からの頻出ジェスチャパターンの抽出・マルチモーダル非言語系列(頭部動作, 顔向け方向, 発話状態, ハンドジェスチャ)からの頻出共起イベントの抽出に利用できる。

3-(1)で取得した会話データから頻出パターンマイニングを行ったところ、聞き手の視線が向けられていない状態でのジェスチャ、言いよどみ時のジェスチャを行っている場合、人の顔を見ず、下を向くことなどのパターンが発見された。一方で説明に付随するジェスチャが行われる場合は、多くの場合で相手の視線がジェスチャを行う人に集まっていることがわかった。

### (2) 対話データにおけるジェスチャの分析

3-(1)で取得した会話データで用いられたジェスチャを分析した。話者によるジェスチャ使用の頻度を自動計算することにより、各シーンの説明に用いられたジェスチャの使用頻度、二人の説明スタイル(一方がジェスチャを使って説明する, 協調的にジェスチャの受け渡しを行いながら説明する)を自動分析できることを示した。

### (3) ジェスチャの役割の自動分類

4-(1)の成果を利用して、会話参加者の非言語行動から規定される対話コンテキストを利用してジェスチャの機能分類を試みた。

最初に相手に向けた説明のために用いられたジェスチャとそれ以外のジェ

スチャ(自己接触や自己指向性ジェスチャなど)の識別に有用な特徴量の抽出を行った。説明者の手の動きについてはジェスチャユニットのホールド・ストロークの回数やそれらのジェスチャ区間における時間占有率が有用であり、説明に利用されたジェスチャは聞き手の視線を得られている時に表出されることが示された。この結果を利用してジェスチャ役割の分類モデルを構築した。具体的にはジェスチャを行う人以外の人間の非言語情報(顔向け, うなづき, 発話状態)との共起関係をトピックモデル LDA により抽出し、ジェスチャ区間の検出は HMM と変化点検知手法を組み合わせたモデルにより行い、

これらの特徴量を統合することで、話し手のジェスチャが説明に用いられているか、それ以外のジェスチャかを 75% の精度で識別できることを示した。

パターン認識分野でジェスチャ認識に関する研究成果は多いが、会話中に表出されるジェスチャの役割の認識に焦点をあてたものは少なかった。また当初の計画通り、ジェスチャを行っている人以外の参加者の非言語情報を利用することで説明状況下でのジェスチャ役割の識別率を向上できることを示すことが出来た。

会話中の非言語の役割の認識技術は会話エージェント・ロボットなどのセンシングモジュールとして有用である。また会話中に表出される非言語情報を自動認識する技術は、大規模会話分析・モデリング・知識獲得に欠かせない技術であり、コミュニケーション行為分析・コミュニケーション行為の定量化・モデル化の基盤技術となることが期待される。

### (4) 今後の課題

対話コンテキストに依存したジェスチャ認識を行うために、非言語パターンマイニングによる対話コンテキスト抽出、コンテキストに基づくジェスチャ認識モデルの構築・評価は達成した。しかしながら、当初予定していた言語情報からのシーンの特定・韻律情報の利用・対話のダイナミクスのモデル化については満足な結果が得られなかった。これらを行うことが今後の課題となる。

### 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〕 (計 5 件)

- [1] **岡田将吾**, 坊農真弓, 角康之, 高梨克也,  
会話インタラクションにおけるジェスチャの量的分析を支援する時系列データマイニング手法の開発, 社会言語科学, Vol.15, no. 1, pp.38-56,2012 (査読あり)
- [2] **岡田 将吾**, 西田 豊明, 自己増殖型ニューラルネットワークを用いた時系列データの追加学習型クラスタリング, 日本神経回路学会論文誌 Vol.17, No.4, pp.174-186, 2010年12月(査読あり) [https://www.jstage.jst.go.jp/article/jnns/17/4/17\\_4\\_174/article/-char/ja/](https://www.jstage.jst.go.jp/article/jnns/17/4/17_4_174/article/-char/ja/)
- [3] **Shogo Okada**, Yoichi Kobayashi, Satoshi Ishibashi, Toyooki Nishida, Incremental Learning of Gestures for Human-Robot Interaction, Springer Journal of AI & Society, Vol. 25, Num. 2, pp.155-168, 2010年5月(査読あり) <http://link.springer.com/article/10.1007%2Fs00146-009-0248-8>

〔学会発表〕 (計 33 件)

- [1] Youwei Lu, **Shogo Okada** and Katsumi Nitta, Semi-supervised Latent Dirichlet Allocation for Multi-label Text Classification, The 26th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE) 2013, 2013年6月19日, Amsterdam, Netherland
- [2] **岡田将吾**, 神谷祐樹, 佐藤裕作, 藤田善弘, 山田敬嗣, 新田克己, センサ環境を利用したオフィスワーカーの行動パターン分析, 人工知能学会全国大会 2013 2013年6月5日 富山
- [3] **岡田将吾**, 多人数インタラクション分析・理解のためのマルチモーダル時系列データマイニング, 電子情報通信学会 HCG VNV 研究会 2012 年次大会, 2013年3月23日 東京, (基調講演)
- [4] Ayano Tamura, **Shogo Okada**, Tetsuya Harada, Makoto Sato, Katsumi Nitta, HAPPicom: Haptic Pad for Impressive Text Communication, ACM International Conference on Intelligent

User Interfaces (IUI2013), 2013年3月20日, Santa Monica, CA

- [5] **Shogo Okada**, Yusaku Sato, Yuki Kamiya, Keiji Yamada, Katsumi Nitta, Analysis of the Correlation between the Regularity of Work Behavior and Stress Indices Based on Longitudinal Behavioral Data, 14th ACM International Conference on Multimodal Interaction (ICMI) 2012, Santa Monica, CA, 2012年10月25日
- [6] **岡田将吾**, 人見謙太郎, チャンドラシリナイワラ, 呂有為, 新田克己, 車載センサログの時系列データマイニングに基づく運転挙動の分析, FIT2012, O-015, 2012年9月5日, 東京
- [7] **岡田将吾**, 坊農真弓, 角康之, 高梨克也, 新田克己: 非言語パターンの自動抽出による状況説明会話における言い淀みシーンの分析, インタラクション 2012年3月17日, 東京
- [8] Meng Chen, **Shogo Okada**, Katsumi Nitta: Effectiveness of haptic interaction in online negotiation between Chinese and Japanese. GrC 2011: 111-114, 2011年11月9日, Kaohsiung, Taiwan
- [9] **Shogo Okada**, Toyooki Nishida :Online Incremental Clustering with Distance Metric Learning for High Dimensional Data, IEEE International Joint Conference on Neural Networks(IJCNN)2011, 2011年8月2日 San Jose, CA,
- [10] **岡田将吾**, 伊豆蔵拓也, 西田豊明, 二段階モチーフ発見アルゴリズムに基づく連続インタラクションデータからのジェスチャパターンの抽出とロボットナビゲーションへの応用, 人工知能学会全国大会 2011年6月1日, 岩手

- [11] Sutasinee Thovuttikul, Divesh Lala, Hiroki Ohashi, **Shogo Okada**, Yoshimasa Ohmoto, and Toyoaki Nishida, Simulated Crowd: Towards a Synthetic Culture for Engaging a Learner in Culture-dependent Nonverbal Interaction, 2011年2月12日, Palo Alto, CA.
- [12] **Shogo Okada**, Osamu Hasegawa, Toyoaki Nishida, Machine Learning Approaches for Time-series Data Based on Self-Organizing Incremental Neural Network, International Conference on Artificial Neural Networks (ICANN2010) workshop, 2010年9月16日, Thessaloniki, Greece
- [13] **Shogo Okada**, Toyoaki Nishida: Multi Class Semi-supervised Classification with Graph Construction Based on Adaptive Metric Learning, ICANN2010, 2010年9月15日発表, Thessaloniki, Greece
- [14] Kae Sakamoto, **Shogo Okada**, Toyoaki Nishida, Multi-modal Processing for Acquisition of Interaction Behavior in Social Robot, International Workshop on Social Intelligence Design (SID2010), 2010年9月13日, London, UK
- [15] Hiroki Ohashi, **Shogo Okada**, Yoshimasa Ohmoto and Toyoaki Nishida, A Proposal of Novel WOZ Environment for Realizing Essence of Communication in Social Robots, International Workshop on Social Intelligence Design (SID2010), 2010年9月13日, London, UK
- [16] **岡田将吾**, 西田豊明, 適応型距離学習とグラフ学習の統合に基づく半教師付き多クラス分類, 画像の認識・理解シンポジウム(MIRU) 2010年7月28日, 釧路
- [17] **岡田将吾**, 神田賢一, 大本義正, 西田豊明, ユーザとの相互作用を通じた画像分類システムの開発, 人工知能学会全国大会 2010年6月10日 宮崎
- [18] **Shogo Okada**, Satoshi Ishibashi, Toyoaki Nishida: On-line Unsupervised Segmentation for Multidimensional Time-series Data and Application to Spatiotemporal Gesture data, IEA/AIE2010, 2010年6月3日, Córdoba, Spain
- 〔図書〕(計5件) 全て分担執筆
- [1] Masaki Sugimoto, Takahiro Ueda, **Shogo Okada**, Yukio Ohsawa, Yoshiharu Maeno and Katsumi Nitta, Discussion Analysis Using Temporal Data Crystallization, New Frontiers in Artificial Intelligence, LNCS (採録決定済) (査読あり)
- [2] Youwei Lu, **Shogo Okada**, Katsumi Nitta, Semi-supervised Latent Dirichlet Allocation for Multi-label Text Classification, LNCS, Volume 7906 (採録決定済) (査読あり)
- [3] **Shogo Okada**, Satoshi Ishibashi, Toyoaki Nishida On-Line Unsupervised Segmentation for Multidimensional Time-Series Data and Application to Spatiotemporal Gesture data, Trends in Applied Intelligent Systems, Lecture Notes in Computer Science (LNCS), Volume 6096, 2010, pp 337-347 (査読あり)
- [4] **Shogo Okada**, Osamu Hasegawa, Toyoaki Nishida Machine Learning Approaches for Time-Series Data Based on Self-Organizing Incremental Neural Network, Artificial Neural Networks – ICANN 2010, LNCS Volume 6354, 2010, pp 541-550 (査読あり)
- [5] **Shogo Okada**, Toyoaki Nishida, Multi Class Semi-Supervised Classification with Graph Construction Based on Adaptive Metric Learning, Artificial Neural Networks – ICANN 2010,

LNCS, Volume 6353, 2010, pp

468-478 (査読あり)

[その他]

ホームページ等

<http://www.ntt.dis.titech.ac.jp/~okada/index.html>

6. 研究組織

(1)研究代表者

岡田 将吾 (OKADA SHOGO)

東京工業大学・大学院総合理工学研究科・

助教

研究者番号：00512261

(2) 研究分担者

なし

(3)連携研究者

なし