

研究報告書

「行動の記号化を基盤とした身振り・言語を通じて コミュニケーションするロボットの知能設計」

研究タイプ: 通常型

研究期間: 平成22年10月～平成26年3月

研究者: 高野 渉

1. 研究のねらい

人間は、他者の動きを自らの身体に置き換え、そこから湧き上がってくる感覚を共有することで、他者の行動を理解する。たとえば、ヘビの動きを観察して、その運動を自身の身体に投影して、感覚を共有することができるだろうか？人間とヘビの身体構造には大きな隔たりがあり、その間の身体運動の変換は容易にはできない。そのため、ヘビの感覚を推定することができず、結果として理解できない、親近感の湧かない存在になってしまう。一方、手足を有した人間のような姿・形をしている動物キャラクタが動き回る様子を見ると、その動きからそのキャラクタが意図していること、その動作の目的を知らず知らずのうちに推測していることはないだろうか？これは、そのキャラクタの身体が人間の身体に類似していることによって、キャラクタの動きを容易に自身の身体運動に置き換えることができるからである。そして、人間と類似した身体構造を有するヒューマノイドロボットは、このような擬人化能力を大いに刺激し、人間・ロボットの関係性を大きく変える可能性を秘めている。

また、人間は、連続情報の身体運動を離散的に情報圧縮して記号として扱っている。他者が歩く動作は、手足の動きの連続値の時系列情報として感覚される。しかし、この連続情報を分割・分類することによって「歩く」という記号として理解している。歩く動作を他者に伝達する場合には、歩くジェスチャを再現する必要なく、「歩く」という単語のみでその動きを伝えることができる。さらに、「歩く」と他の記号を組み合わせることによって、多様な動作を簡単に作り、創造することも可能になった。このように、実世界に膨大に溢れている連続情報を切り分けて、分類することによって、実世界を記号として知覚し、記号を効率的に計算処理することによって、高度な思考や効果的なコミュニケーションを実現してきたのである。

このように、実世界を自分の身体に投影して理解する擬人化能力と、実世界に溢れた連続情報を離散的に分割・分類する記号化能力が、人間の思考・コミュニケーションの高度な認知メカニズムの基礎ある。本研究では、人間の身体運動を運動パターンとして分割・分類して記号として表現する方法論、およびその運動の記号と自然言語を結びつけることによって、動作を言語して理解し、言語から身体運動を創造することができるロボットの知能構成論を確立することを目指す。

2. 研究成果

(1) 概要

本研究課題では、人間の全身運動を記号として表現し、その記号と自然言語を結びつけることによって、人間の行動を言語して理解でき、言語表現に対応する行動を生成することができるヒューマノイドロボットの知能を設計する。その基盤技術として、以下の項目に関する研究

を行った。

- (A)人間の全身運動と自然言語モデルを統合する数学的手法の開発 [2][3]
- (B)クラウドソーシングを用いた運動と自然言語の大規模化 [3]
- (C)大規模な運動の学習認識の高速化手法の開発 [3]
- (D)環境適応するための運動の記号からの全身運動生成手法の開発 [1]

項目(A)では、全身運動と単語の関係性および文章における単語の並びの関係性を統計モデルによって抽出した。この2つの関係性を用いることによって、人間の全身運動を文章として理解するための計算法、および与えられた文章に対する全身運動を生成する計算法が確立した。

項目(B)では、全身運動および言語表現を大規模化することによって、多様な動き、多様な文章を理解できる枠組みへと発展させた。特に、ウェアラブルモーションセンサスーツを着用した人間の日常データを記録することによって運動の大規模化、動きへ多様な文章を付与する方法としてクラウドソーシングを活用することで、大規模な動き・言語のデータを収集できる環境を整備した。

運動・言語データの数が増大するにつれて、運動と言語の双方向計算に掛かる時間が膨れ上がる。そこで、全身運動データから最も確率が高い運動の記号を検索する計算、およびその記号から単語を並び替えて文章を生成する計算をそれぞれ並列に処理するアルゴリズムを構築した。それをクラウドコンピュータにて実装することで、リアルタイムで行動を認識できるシステムを構築した。

また、身体を通して環境に働き掛ける機能として、上述のように記憶した人間の全身運動を環境に適応するように修正を加えて再利用する運動生成手法を構築した。ターゲットとなる物体の位置に応じて手先位置の軌道を修正する方法や環境から受ける外力を目標値となるような環境とインタラクションするための軌道生成および制御方法を開発した。

このように、多様な全身運動を言語として高速に理解できる計算から、言語入力から環境に応じた全身運動生成までを実現する計算方法を開発した、これは、身振り手振りを伴いながら言語を用いてコミュニケーションするヒューマノイドロボットの知能の基盤となる。

(2) 詳細

研究テーマ(A)「人間の全身運動と自然言語モデルを統合する数学的手法の開発」

人間の全身運動の時系列データを統計モデルのパラメータとして情報圧縮する。各モデルを運動の記号とみなし、その記号とその運動アノテーションの対応関係を統計モデルによって学習する。これによって、運動から単語が確率的に連想することが可能になる。また、運動アノテーション内における単語の並びを単語Ngramとして統計的数式モデルに落とし込む。これによって、先の運動から連想された単語を用いて、適切に並べ替えることによって文章を作成することができる。465 個の運動記号、764 個の運動アノテーションを学習したシステムを構築した。全身運動を文章として理解する計算で、88%の認識率が達成できた。



研究テーマ(B)「クラウドソーシングを用いた運動と自然言語の大規模化」

日常生活における人間の行動および言語表現は複雑かつ多様である。このような多岐に渡る行動を理解するロボットの知能を構築する方法として、大量の運動データと言語表現を活用したデータ駆動型アプローチが有効である。しかし、これまでの光学式モーションキャプチャを利用したデータ収集方法は、計測設備の条件から運動データの大規模化に限界が見えていた。そこで、どこでも人間の動きを計測できるようにウェアラブルモーションスーツを着用した被験者の運動データを蓄積できる計測環境を整備した。これによって、これまでの 20 分程度の運動データから約 20 時間の運動データを収集することに成功した。また、その運動データに文章を付与する作業において、誰でも動きを見てその動きを表現する文章を与えることができるクラウドソーシングの環境整備した。その結果として、現在約 60,000 個の文章からなるアノテーションデータを収集した。



募集中 CG動画を見ていただき、英語で文章記載 (その2)

全員ポイント進捗 53%

タスクの種類 11件

もらえるポイント 112

タスクにかかるめやす時間 5分00秒

残り件数 11件 (全340件)

タスク実施中のユーザー: 1人

95%完了

実施できる件数の上限 1件

報酬時間 10分00秒

残り日数 9日 (2014/02/01 20:00)

オーナー 野公開

このタスクって、どんな内容?

【タスク詳細】
人間の動作を計測したCG動画を確認し、動作内容を英文で記述して頂くタスクです。
(英語、例: a student sits in the desk, he operates a PC)
予め時間と生活が入力されています。その時間にどんな動きに見えたかを主題に詳しく文章として入力してください。
CG動画は何の動作が分かりにくい場合がありますが、直感で記載頂く結構です。

※作業実施画面に回答例があるので、確認の上、作業を実施してください。
※英文でご記入下さい。
※翻訳サイトへのリンクを用意しているので必要に応じてご使用ください。

【注意】
不適切な回答をされた場合、今後回答をお断りする場合がございますので、ご了承ください。

タスクのサンプル画面

研究テーマ(C)「大規模な運動の学習認識の高速化手法の開発」

運動や言語のデータが膨大になるにつれて、観察した運動を言語表現として理解する計算も膨れ上がる。特に、人間とインタラクションするヒューマノイドロボットにおいて、その計算時間による遅れは全体の性能評価に直結する。そこで、その計算の高速化として、(1)観察した運動に一番合致する運動記号の統計モデルを選択する計算、および(2)運動記号から文章を

生成する計算をそれぞれ並列処理化するアルゴリズムを構築した。(1)では、観察した運動データが各運動記号の統計モデルから生成される確率の計算を独立に処理する。(2)では、文章生成は、単語を並べて単語列を作る探索計算であり、複数の単語列から単語を追加して文章を求める計算部分を並列化する方法を考案した。考案したアルゴリズムをAmazonのクラウドコンピュータ(CPU E5-2670 8core、メモリ 60.5GB)に実装した。運動から文章を生成する計算に従来の手法では620秒掛かるところ、考案したアルゴリズムによって320プロセスの並列化を行い6.1秒に短縮できることを確認した。

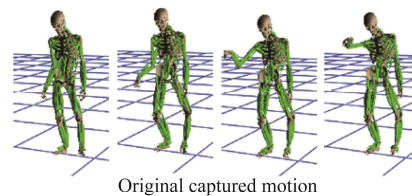
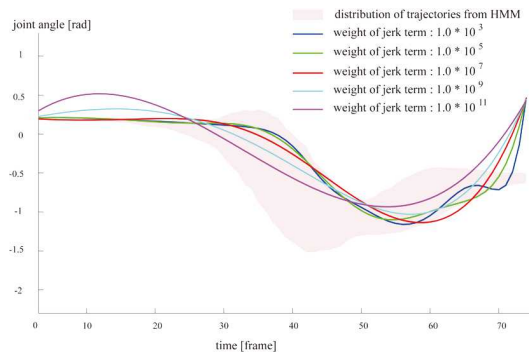
**研究テーマ(D)「環境
適応するための運動の記
号からの全身運動生成手
法の開発」**

人間の全身運動を記憶
し、その動きをロボットの

行動生成に再利用することで人間のような自然な振り舞いをするヒューマノイドロボットを実

現できる。しかし、記憶した動きをそのまま再生するだけでは、求められるような行動を実現することができない。例えば、ある位置に置かれたコップに手を伸ばさず動きを記憶し、コップの位置が異なる状況下で同じ動きを行っても手先はコップに到達することはなく、その動きは意味のないものになってしまう。コップの位置を考慮し、手先がコップに到達するように記憶した運動を修正する必要がある。そこで、生成する運動の軌道が学習した統計モデルから生成される確率と手先がターゲットの位置の誤差を足し合わせた評価関数を設計した。その評価関数を最大化する軌道を最適化問題として解くことによって環境に適応した全身運動を生成する計算方法を確立した。ターゲットに手先を伸ばさず運動を学習した運動記号のモデルを用いて、ターゲットの位置を変えても学習した動きに類似しながら手先がターゲットに到達するような運動を生成できることを確認した。さらに、この手法を力制御へ展開することによって、環境と身体的作用・反作用力を理想な状態に維持しながら動きを制御する手法へと発展させた。

The number of process	(1) motion recognition time [s]	(2) sentence generation time [s]
1	3.4492	620.895
16	0.2354	67.2654
320	0.0427	6.1809



3. 今後の展開

本研究課題では、人間の全身運動を統計モデルとして記号化し、その記号と言語の関係性を抽出することによって、ヒューマノイドロボットが人間の動作を文章として理解すること、および入力された文章から全身運動を生成することを可能とする知能を構築した。特に、運動データや言語データが増大しても、多様に動きを理解できることを確認し、運動や単語の種類に対するスケーラビリティを示すことできた。しかし、身体運動は身体の動きだけで必ずしも記述されるものではない。モノに手を伸ばし、操作し、環境に作用する動きが実世界には多様に存在する。本研究課題で開発した枠組みを、物体操作に係わる運動へ展開することでより詳細に人間の行動を正確に理解できる知能へ繋げて行く必要がある。特に、以下のような基盤技術を確認しながら、物体操作に関する行動認識や行動生成技術を向上させていく必要があるであろう。

- ・身体運動と環境のインタラクションに関する記述法の開発
- ・全身運動に指の動きを追加するための計測環境の整備
- ・環境中のモノなどを指し示す言語表現の増加に伴うスケーラビリティの確保
- ・環境中のモノをカメラや距離センサで分節化・分類する手法の開発
- ・身体と環境の作用・反作用力を運動のモデル化に組み入れる手法の開発

4. 評価

(1) 自己評価

本研究課題では、人間やロボットの身体運動に関する実世界情報と自然言語を統合することによって、多様な身振り手振りや言語表現を交えてコミュニケーションするロボット知能の基盤を確立することを研究目標に掲げている。本研究成果の位置づけおよび自己評価は以下の3点として纏められる。

これまでの行動認識では、観察された運動データを 20 程度のクラスターに分類する問題として設定されて、今も Chalearn (Challenges in Machine Learning)などでジェスチャ認識精度が競われている。しかし、本研究では数千種類の運動を認識し、数万語の単語を用いてその行動を表現する問題としてジェスチャ認識を捉えており、これまでのジェスチャ認識と比較して、カバーする運動の多様さ、言語表現の豊かさでは間違いなくトップクラスである。また、このようなデータセットを公開することで運動認識の分野を引っ張っていくことも期待できる研究成果を上げていることも評価できる。

ヒューマノイドロボットの運動制御に関する研究は古くから行われているが、歩行などの動きに関する安定性に関する制御手法にフォーカスが向けられがちである。そのため、歩行以外の動きを制御するコントローラを設計する指針があまり考案されていない。本研究成果では、多様の運動パターンを統計モデルとして表現することで、そのモデルから全身の運動軌道を発生する生成器としての利用することで、複雑多岐に渡る運動を実行するヒューマノイドロボットのコントローラの要素となる。

自然言語処理にて開発されてきた言語モデルを利用しながら、それを実世界の運動と繋が

る枠組みを開発したことは、ロボット研究だけでなく自然言語処理研究にも簡易に利用できるメリットがあり、その波及効果は大きいと評価できる。

(2) 研究総括評価

人間の身体運動を記号として解釈し、自然言語を用いてロボットの運動を生成することを目指した研究である。本研究の特徴は、理論に裏付けられた大規模な実装にある。まず、身体運動の記号化では、465 個の運動記号を学習するシステムを構築し 88%の認識率を達成している。また、その学習のために、クラウドソーシングを活用して 6 万件の運動アノテーションデータを収集している。さらに、並列計算によってリアルタイムで運動を認識できるシステムを構築し、320プロセスの実装で100倍の性能向上を達成している。こうした成果は、従来の理論研究とは一線を画すもので、実際にロボットを自然言語で制御する突破口を開くものとして高く評価できる。今後は、急速に発達するウェアラブルコンピューティングとの接点を探りつつ、人とロボットの共生社会を実現していくことを期待したい。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

- | |
|--|
| 1. 高野 渉, 濱野 聖也, 中村 仁彦, “統計的相関を用いて運動と言語の構造を結びつけたデータベースの設計”, 電気学会論文誌 C, Vol.133, No.9, pp.1680-1686, 2013 |
| 2. 高野 渉, 中村 仁彦, “全身運動から言語空間の構築と運動の認識への応用”, 人工知能学会論文誌, Vol.28, No.4, pp.361-369, 2013 |
| 3. 高野 渉, 鮎澤 光, 濱野 聖也, 梅澤 慶介, 中村 仁彦, “身体運動と言語を結ぶロボットの統計的情報処理”, 日本ロボット学会誌, Vol.30, No.7, pp.12-21, 2012 |
| 4. 高野 渉, 上段 達弘, 中村 仁彦, “運動の認識・生成計算の再帰的情報処理を通じた群集をすり抜けるヒューマノイドロボット”, 情報処理学会論文誌, Vol.53, No.11, pp.2466-2476, 2012 |
| 5. 高野 渉, 今川 洋尚, クリッチ・ダナ, 中村 仁彦, “クリスタルボール, “運動の記号推論を通じた未来の行動予測”, 日本ロボット学会誌, Vol.29, No. 8, pp.745-751, 2011 |

(2) 特許出願

研究期間累積件数: 0件

(2) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

主な国際学会

1. Wataru Takano, Yoshihiko Nakamura, “Synthesis of Whole Body Motion with Pose-Constraints from Stochastic Model,” Proceedings of IEEE International Conference on Robotics and Automation, Hong Kong, China, May.31-June.7, 2014 (to be published)
2. Yusuke Goutsu, Wataru Takano, and Yoshihiko Nakamura, “Generating Sentence from Motion by using Large-Scale and High-Order N-grams,” Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 151-156, Tokyo,

Japan, November, 2013.

3. Wataru Takano, Junichi Ishikawa, Yoshihiko Nakamura, “Recovery of 3 Dimensional Human Whole Body Motion from a Monocular Image Sequence by Using Human Behavior Database”, The 2nd IFToMM Asian Conference on Mechanism and Machine Science, Tokyo, Japan, Nov.7–10,2012
4. Wataru Takano, Yoshihiko Nakamura, “Bigram-Based Natural Language Model and Statistical Motion Symbol Model for Scalable Language of Humanoid Robots”, IEEE International Conference on Robotics and Automation, pp.1232–1237, St. Paul, USA, May 14–18, 2012
5. Wataru Takano, Hiroataka Imagawa, Yoshihiko Nakamura : “Prediction of Human Behaviors in the Future through Symbolic Inference,” IEEE International Conference on Robotics and Automation, pp.1970–1975, Shanghai, China, May 9–13, 2011

受賞

1. 郷津 優介, 小林 誠季, 小原 潤哉, 草島 育生, 武市 一成, 高野 渉, 中村 仁彦, “運動・音声・画像の特徴を用いた統合モデルによるマルチモーダルジェスチャ認識”, 第19回ロボティクスシンポジウム 優秀論文 ファイナリスト, 2014.3.13–14
2. 高野 渉, “クリスタルボール:運動の記号推論を通じた未来の行動予測”, 日本ロボット学会研究奨励賞, 2011.9.8
3. 高野 渉, 中村 仁彦, “ヒューマノイドロボットの運動の言語理解と言語からの運動生成計算”, Joint Agent Workshop and Symposium 2011 優秀論文賞, 2011.10.26–28

招待講演

4. Wataru Takano : “Towards Language from Database of Human Whole Body Motions”, Workshop on Semantics, Identification and Control of Robot-Human-Environment Interaction, IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, May 6–10, 2013
5. Wataru Takano : “Scalable Stochastic Modeling of Human Whole Body Motions”, Workshop on Behavior Learning through Body/Environment Interactions, IEEE International Conference on Biomedical Robotics and Biomechatronics, Roma, Italy, June 24–27, 2012
6. Wataru Takano, “Robotics Education with NAO –Beauty in Behaviors of Human and Humanoid Robot”, Workshop on Educating Robotics Engineers and Scientists, IEEE International Conference on Robotics and Automation, pp.1970–1975, Shanghai, China, May 9–13, 2011
7. 高野 渉, “身体運動の記号と自然言語を結び付ける数理モデルからロボットの知能へ”, 日本ロボット学会ロボット工学セミナー「意味や状況を扱うロボット技術—記号化・対話・ビッグデータ」, 東京, 2013.10.9