

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2012～2013

課題番号：24650091

研究課題名(和文) 言語の創造性を利用して運動の多様性を生み出すロボットの知能

研究課題名(英文) Motion Synthesis using Linguistic Creativity towards Robot Intelligence

研究代表者

高野 渉 (TAKANO, WATARU)

東京大学・情報理工学(系)研究科・講師

研究者番号：30512090

交付決定額(研究期間全体)：(直接経費) 2,900,000円、(間接経費) 870,000円

研究成果の概要(和文)：言語は有限の語彙を組み合わせることで無限の意味表現を生み出す機能と、無限の意味を文法という規則を通じて効果的に伝達する機能を兼ね備えた記号システムである。人間の高度な知能はこの記号システムにて支えられている。本研究課題では、身体運動と言語表現の連想構造を統計モデルにて表現するロボット知能の枠組みに関する研究開発を実施した。この技術は、ロボットが人間の身体運動を観察してそれを説明文として理解すること、文章に相当する全身運動を生成することを可能とする。また、運動毎に潜む重要な動きを抽出・組み合わせることによって、運動と複数の文章の対応付けも可能になり、言語理解・行動生成の多様性を向上させることができる。

研究成果の概要(英文)：Humanoid robots are required to understand the same language system as humans so that they are integrated into our daily life. In order to achieve this intelligent robot, we have made following researches in this project.

1. Designing a stochastic model for association between motions and language. Human whole body motions are encoded into the stochastic models, which can be referred to as symbols. The association between the symbols and descriptive sentences attached to the motions is represented by graphical stochastic models. 2. Understanding motions using language. The framework of association between motions and language makes it possible for robots to understand observation as sentences and to synthesis robot whole body motion corresponding to given sentences. 3. Extracting Correlated motions. Extracting featured motions and assembling them enable the robots to recognize the observations in various language expressions, and to combine multiple motions to perform complex behaviors.

研究分野：総合領域

科研費の分科・細目：情報学 知覚情報処理・知能ロボティクス

キーワード：知能ロボット 行動理解 全身運動生成

1. 研究開始当初の背景

言語は有限個の語彙を組み合わせた無限の表現性と予め決められた型を持つ文法による正確な意味の伝達性を兼ね備えた非常に創造性に富んだ記号システムである。人間は、実世界の連続情報を離散情報に変換することで言語へ発展する記号システムを獲得してきた。そして、ロボットの知能化研究においても、身体運動や物体といった実世界の事象を記号として認識するための機械学習の研究が古くから行われている。これらは、運動およびそれを取り囲む環境の連続情報を数理モデルのパラメータとして表現する低次元・記号化技術である。特に、身体運動の見真似学習のパラダイムは、運動を記号として学習し、人間の行動を記号として認識することや記号から身体運動を生成することに成功してきた。しかし、この枠組みで扱われる運動は少数・限定的であり、日常生活に溢れる人間の行動の多様性を網羅できるものではない。大規模な運動と言語表現を理解できるロボットの知能が新たな人間・ロボットコミュニケーションに必要不可欠である。

2. 研究の目的

研究代表者は、これまで身体運動を統計モデルのパラメータ表現として学習・記号化・記憶するロボットの知能の研究を行ってきた。運動の記号表現と運動に人手で付与された運動ラベルの対応関係を学習することによって、運動を運動ラベルとして理解できるロボット知能を開発してきた。そして、この枠組みを拡張することによって、運動と自然言語を結びつける枠組みを構築し、ロボットが自然言語を理解する知能、言語を活用して多様な行動を表出する知能へ接近できるのではないかという着想に至った。そこで本研究課題では、多様な人間の運動とその運動を説明する文章を接続する数理モデルを開発する。多様な運動と文章を組み合わせ、複雑な行動を理解するロボット知能の基盤となる。これを実現するために以下の項目に記す研究開発を実施した。

(1) 身体運動の記号化

人間の身体運動は、身体部位の位置や関節角度といった連続データの時系列として計測される。この連続データを離散形式の記号として表現する数理モデルを開発する。これは、言語表現と接続することを容易にする。

(2) 運動の記号と言語の接続モデル

身体運動の記号と文章の対応関係を学習する枠組みを構築する。運動と単語の関係性を学習する方法論は、これまで研究代表者が開

発してきた運動と言語ラベルを結びつける方法論を拡張する。また、言語表現を効率的に伝達するためには、文章構造を学習する必要がある。文章を単語の並びとして捉え、その並びから文章構造を表現するモデルを構築する。これら2つの統合することで運動を文章として理解できる情報処理を実現する。

(3) 言語知識を利用した運動の特徴抽出
運動に付けられた文章は、運動を人間がどのように分類しているかを示している。この言語知識を利用して、人間と同じように運動を分類するための特徴量を抽出する。同じ文章が付与されている運動データを集める。運動データは、高次元ベクトルの時系列データであり、各次元間において高い相関のある次元の抽出、および各運動データ間に渡って分散の小さい次元を抽出する。これは、ある運動において同期して変化する動きと。同じ文章が付与されている運動間に共通した不変量を発見することを目的としている。

(4) 多様な運動の理解・生成技術
観察された全身運動を運動クラスごとの特徴量に変換する。この特徴量を用いて運動を認識することで、一つの全身運動を多様な視点で認識することができる。さらにその複数の結果を文章へ変換することで多様な言語生成を実現する。複数の文章から動きを組み合わせた運動を合成する技術も開発する。

3. 研究の方法

多様な運動と言語を理解するロボットの知能を構築するために以下の要素技術を開発した。

(1) 身体運動の記号化

人間に貼り付けられたマーカの位置を計測して得られる運動データ、および外力・筋電位センサを用いて推定される体性感覚データを合わせた高次元ベクトルの時系列信号として運動を表現する。この時系列信号を統計モデルの一種である隠れマルコフモデル (Hidden Markov Model : HMM) によって学習することで、運動が HMM に記号化される。以後、HMM を運動の記号とみなすことにする。運動の認識は、観察された時系列データが各 HMM から生成される尤度を計算し、最大の尤度を出力する HMM を見つける計算となる。また、運動の生成は、HMM 内のノード遷移、ノードからの運動データの出力を繰り返すサンプリング手法を利用したモンテカルロ計算となる。

(2) 運動の記号と言語の接続モデル

身体運動はHMMを用いた認識処理によって運動の記号へ変換することができる。同じ身体運動に対して人手で説明文を付与する。運動記号と説明文の組の集合を、運動と文章の学習データとして収集する。説明文を単語に分節化することで、運動記号と単語群の関係および説明文中における単語の並びの2種類の学習データを作成する。運動記号と単語群の関係は、入力が運動記号、出力が単語、中間層に隠れ変数を介する統計モデルとして表現する。図1にその概略を示す。この統計モデルは、運動記号から単語が連想される確率パラメータと隠れ変数から単語が生成される確率パラメータから成り、これら2つのパラメータは、学習データとして与えられた運動記号から単語群が生成される確率を最大化する値として学習される。また、文章構造を単語の並びとして捉え、文章中の単語の遷移構造を表す統計モデルを学習する。図2に、単語の並びを bigram として記述したモデルの概略を示す。統計モデルの各ノードが単語に相当し、ノード間のエッジが単語の遷移確率を表す。単語の並びに関する学習データがこのモデルから生成される確率が最大となるように単語遷移確率を最適化する。

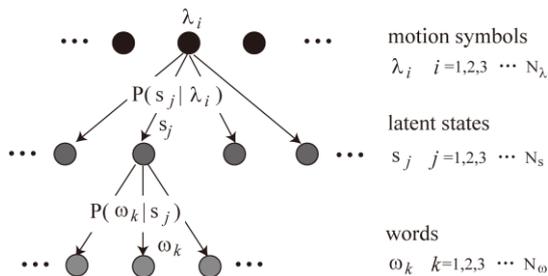


図1. 運動と単語の対応学習

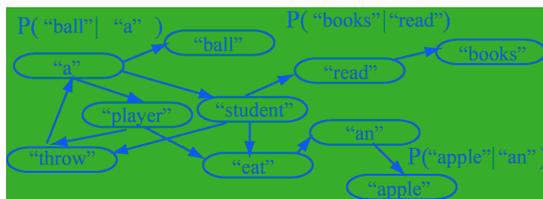


図2. 単語の遷移モデル

この2つのモデルを利用することで運動と文章を接続する。運動を文章として理解する計算では、観察された運動を運動記号として認識し、前者のモデルを用いることで、その運動記号から単語群が連想される。連想された単語を並び替えて作成される文章は、後者のモデルによって文章としての尤もらしさを示す確率値として評価される。この一連の過程を通じて、確率値が高い文章を見つけ出すことで運動を文章とし理解することが可能となる。文章から運動を生成する計算では、文章中の単語群から連想される確率が高

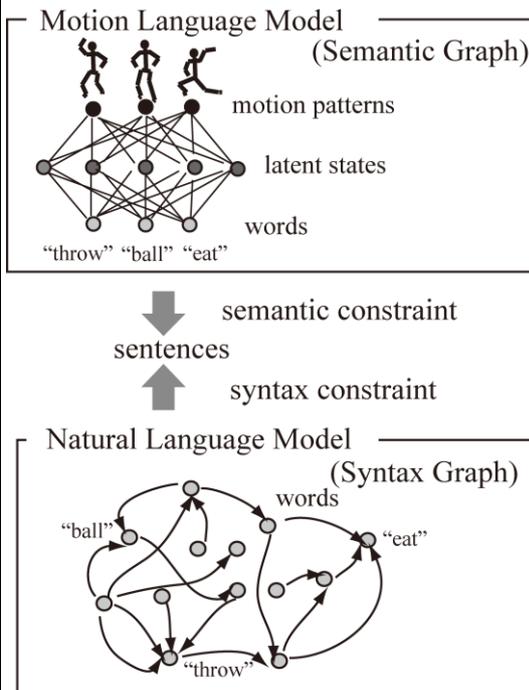


図3. 運動から文章生成の概略図

い運動を見つけることで実現する。図3にその運動と文章を接続する数理モデルの枠組みを示す。

(3) 言語知識を利用した運動の特徴抽出
運動データに付与された言語ラベルを手掛かりにして運動データをグループに分ける。各グループに属する複数の運動データに対して、相関解析を通じてそのグループにおいて重要な動きを取り出す。先述の運動記号と同様に、運動データをHMMによって学習する。このHMMの構造を利用することで、運動データの各時刻がHMMのどのノードから生成されたのかを推定できる。すなわち、時系列データの各時刻における運動データをノード毎に細分化したグループに分けることができる。同じノードとして細分化された高次元ベクトルの運動データに対して、次元間の相関を計算する。相関が高い2つの次元のデータは同期して変動していることを示しており、これは2つのデータを1変数で近似して表すことができることを意味する。このように相関解析を通じて、高次元ベクトルの時系列は代表的な低次元データとして変換され、それがHMMとして記号される。高次元運動ベクトルを低次元化するための写像は、HMMのノード毎に計算される相関行列の関数として求めることができる。

(4) 多様な運動の理解・生成技術

上述の低次元化技術は、運動にとって重要な動きを見つけることができる。そして、その重要な動きのみに着目してHMMの尤度計算を通じた運動認識ができる。全身運動に関して全ての身体部位の動きを平等に見て、動き

を分類するのではなく、各グループにとって重要な動きに着目して、動きを認識することでその分類精度を向上させる。また、これと文章がつながることで、2つの文章を入力した場合、各文章から運動記号が連想される。その運動記号が着目すべき身体部位が排他関係にあるならば、各特徴的な動きを組み合わせ、複雑な動きを生成することができる。例えば、「人が歩く」と「学生が手を挙げる」の文章から連想される各HMMにおいて、前者のHMMでは足の動きが重要であり、後者のHMMでは手の動きが重要となる。そして、前者のHMMからは足の動きを作り出し、後者のHMMからは手の動きを作り出す。そして、それらを足し合わせることで、「歩きながら手を挙げる」といった全身運動を創造することが可能となる。

4. 研究成果

467個の運動データ計測し、各運動に複数の文章を付与した学習データを収集した。付与された文章の総数は764個であり、使われた単語の種類は241個あった。図4に計測した運動とそれに付与された文章例の一部を示す。計測された運動データは、1190個の筋張力データに変換された高次元ベクトルを用いた。



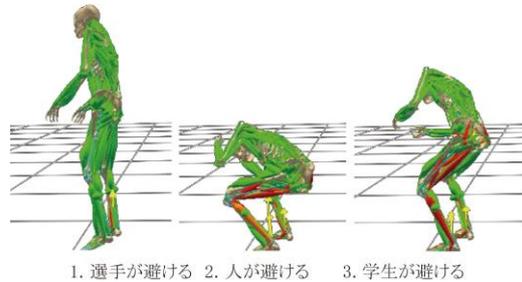
- | | |
|---|--|
| <ol style="list-style-type: none"> 1. 学生が脚を組む 2. 監督が脚を組む 3. コーチが脚を組む | <ol style="list-style-type: none"> 1. 主婦が箒で掃く 2. 主婦が掃除する |
|---|--|



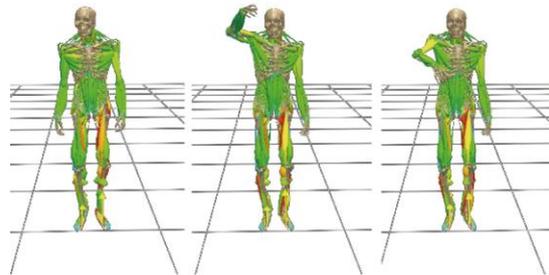
- | | |
|---|--|
| <ol style="list-style-type: none"> 1. 学生がバドミントンをする 2. 選手がバドミントンをする 3. 学生がバドミントンラケットを振る | <ol style="list-style-type: none"> 1. 選手がトスをする 2. 選手がバレーボールをする |
|---|--|

図4. 運動と文章の学習データ例

図5に、観察した運動を言語として認識した実験結果を示す。ここで、確率の高い文章を3つ提示している。屈む動作を「選手が避ける」、「人が避ける」という文章として解釈し、手招き動作を「学生が手招きをする」、「選手が手招きをする」と理解できていることが



1. 選手が避ける 2. 人が避ける 3. 学生が避ける



1. 学生が手招きをする 2. 選手が手招きをする 3. 学生が避ける
確認できる。

図5. 運動を文章として理解するロボット

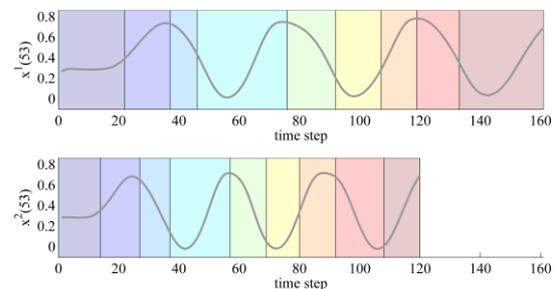


図6. 運動時系列データの細分化

同じ文章が付与されている運動データの相関解析を行った。図6は「左手を振る」動作の左手先のx座標値の時系列と各時刻がHMMのどのノードから出力されたかを示している。時系列データの長さが違うにも関わらず、類似した動きが同じノードに対応していることが確認できる。ノードの対応付けによって細分化されて時系列データに対して、相関解析を行った。図7に各ノードに細分化された時系列データから相関が高いと抽出された身体部位を示す。「左手を振る」動作では、運動間で左手先の動きが類似しており、高い相関があることが抽出されている。

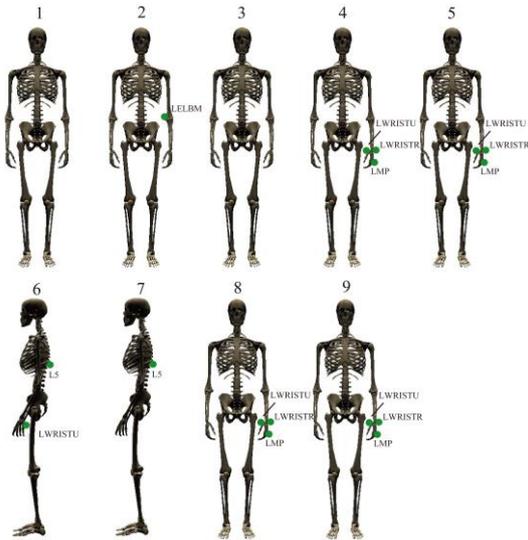


図7. 重要部位の抽出結果. マーカが付与された部位が重要と判定された.

抽出された特徴部位の運動データからHMMを学習し、複数の動きが組み合わせた動作の認識実験を行った. 動きに複数の動作文章が付与されている. 認識結果が付与されている文章の全てと一致している確率は62%, 認識結果が付与されている文章の少なくとも1つと一致している確率が85%であった.

さらに、2つのHMMから運動を合成する実験を行った. 「座る」HMMと「右手を振る」HMMから「座りながら右手を振る」動作の合成を行った. 図8に合成された動きを示す. ヒューマノイドロボットが座りながら右手を振る動作が実行できていることが確認できた. 図9は、ヒューマノイドロボットも用いた場合の、「立つ」HMMと「右手を振る」HMMから「立ちながら右手を振る」動作の合成結果である. この場合も、「立ちながら右手を振る」動作を生成できていることを確認した.



図8. 運動の合成実験結果

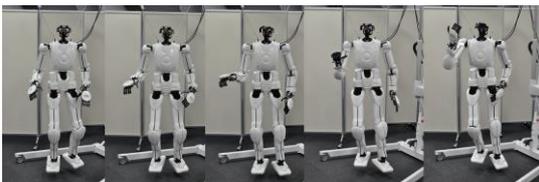


図9. ヒューマノイドロボットを使った運動合成

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 1 件)

- [1] 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

[学会発表] (計 7 件)

- [1] 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 3-8, 2013.

- [2] 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

- [3] 梅澤慶介, 高野 渉, 中村仁彦,
“相関解析を用いたタスク実現のための動作特徴量の抽出”, 日本 IFToMM 会議シンポジウム, 東京, 2014. 5. 24

- [4] 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 2A1, 有馬, 2014. 3. 13-14

- [5] 郷津優介, 高野 渉, 中村仁彦,
“大規模高次 N グラムを用いて動作生成を行う運動認識システム”, 第 26 回日本人工知能学会全国大会, 2G4-0S-19a-6, 富山市, 2013. 6. 4-7

- [6] 高野 渉, 中村仁彦,
“全身動作とその筋活動を推定して言語として理解するロボットの人工知能”, 第 26 回日本人工知能学会全国大会, 3G3-0S-12a-1, 富山市, 2013. 6. 4-7

- [7] 草島育生, 高野 渉, 中村仁彦,
“大規模データベースによる運動と言語の連想をクラウド上で行うヒューマノイドロボットの知能”, 日本機械学会ロボティクスメカトロニクス講演会 2013, 1P1-B07, つくば市, 2013. 05. 22-25

6. 研究組織

(1) 研究代表者

高野 渉 (TAKANO WATARU)

東京大学・大学院情報理工学系研究科
・講師

研究者番号 : 30512090

(2) 研究分担者

なし

(3) 連携研究者

なし