

# 人間の大規模行動認識のための社会実装技術に関する研究開発 (142103011)

## Recognition of Large Scale Human Motion for Social Implementation

### 研究代表者

高野 渉 大阪大学

Wataru Takano Osaka University

研究期間 平成 26 年度～平成 28 年度

### 概要

本研究課題では、人間の全身運動、手指運動・力反力・全身筋張力・環境データの詳細な行動データおよび言語のヘテロビッグデータセットの作成、それらの統計的機械学習から多様な人間行動を認識するシステムの基盤技術を開発した。運動と言語の行動ビッグデータベースから、運動を統計モデルによって記号化する技術、運動記号と言語表現を結び付ける統計的情報処理技術を開発し、日常の多様な運動を認識するシステムの実現性を検証した。

### 1. まえがき

ICTは、WEBなどに代表されるようにさまざまな言語情報を結びつけることによって膨大な集合知を形成してきている。これら情報と実世界の結びつきはまだ乏しいのが現実であり、コンピュータが実世界を理解するには至っていない。実世界を理解するような計算機システムを構築するには、これまで言語中心で設計されてきた知識ネットワークを実世界に拡張する技術が必要不可欠である。本研究課題では、実世界情報として人間の身体運動に着目し、実世界運動と自然言語を結びつけるための膨大なデータベース、および運動から言語を連想する計算システムを開発した。これらは、日常生活における多様な行動を言語として認識するための計算システムの基盤技術となる。

### 2. 研究開発内容及び成果

#### 2.1 全身運動の識別技術

全身に貼られたマーカーの位置を光学式モーションキャプチャにて計測し、逆運動学計算を用いて全身の関節角の時系列として人間の運動を表現する。この時系列データを運動パターン毎に対応した隠れマルコフモデル (Hidden Markov Model : HMM) にて学習する。計測された人間の運動は、その時系列データが生成される確率が最大の HMM として識別することができる。

全身の全ての関節角の時系列から運動を識別することは、計算量や学習データの収集に高いコストがかかる。「手を振る」動作なら「手」の動き、「歩く」動作なら「足」の動きが各動作を特徴付けることが予想される。運動データの特徴は、HMM による統計量として表現されてきた。統計的距離に着目し、運動データが HMM から生成される対数尤度を HMM のパラメータにて偏微分して得られる特徴ベクトルを採用する。この特徴ベクトル間の L2 ノルムは、フィッシャ情報行列を計量とした統計モデルのパラメータ間の距離に相当し、統計的整合性が良いデータ間の距離を算出することができる。さらに、この特徴ベクトルに重みづけをしたカーネル関数を利用したサポートベクターマシンによって、特徴ベクトルの要素毎の重みを最適化した識別機を設計する。20 種類の運動 (学習データ 6380 個、テストデータ 3200 個) に対して、HMM を用いた運動の識別率が 38%であったのに対して、開発した手法では識別率 74%の高い識別性能を達成した。

全身運動に手指の運動情報を付加することによって環境中の物体とのインタラクションを詳細に識別する方法を開発した。手指の運動は、ひずみセンサーを内蔵したグローブによって指の関節角度を計測することができる。全

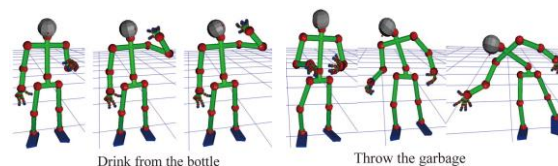


図 1. 全身および手指運動の計測

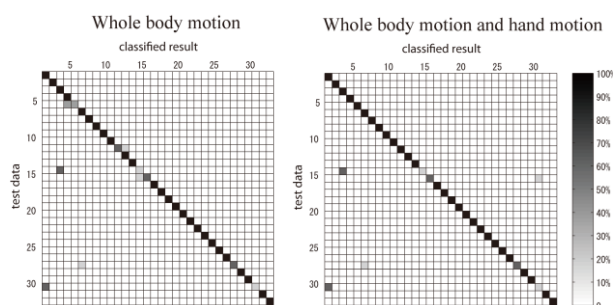


図 2. 全身運動のみを活用した運動識別 (左図)、全身と手指運動を活用した運動識別 (右図) の実験結果

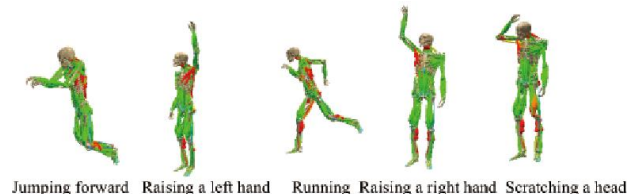


図 3. 全身筋張力の推定

身の関節角と指の関節角度を並べたベクトルを特徴量として扱うことによって HMM の数学モデルを利用することができる。32 種類の運動に対して全身運動のみを用いた場合、識別率が 91%であった。手指運動も用いることによって識別率 94%の運動識別方法を開発した。図 1 に全身および手指運動の計測結果、図 2 に識別結果の混同行列を示す。

全身運動に加えて身体に働く外力および筋電位を計測することによって、運動中における関節角と 1190 個の筋張力をロボットの運動学・動力学計算を通じて推定することができる。筋張力からの運動識別について、筋張力を並べたベクトルは高次元であること、筋張力がパルス状に変化することを勘案して、各筋肉に分散型の HMM を割り当て、それらを統合することで運動を識別する方法を開発した。154 種類 (図 3 はその一例) の運動に対して、従来の連続型 HMM を用いた場合の識別率は 9.8%であった。開発した方法によって識別率を 36%に高めることを実現した。

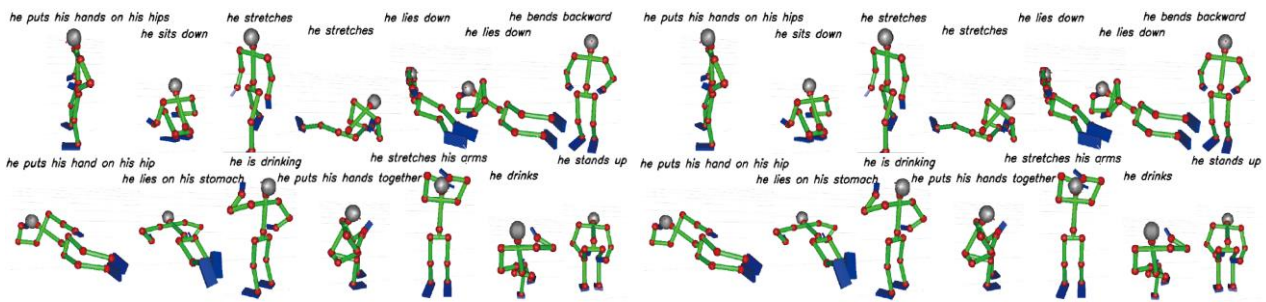


図 4. 運動から文章を生成する実験の結果

## 2. 2 運動と言語の大規模データベース

日常の運動は多様である。多様な運動を識別するには膨大な運動を記憶・学習するためのデータセットが必要となる。光学式モーションキャプチャでは計測できる環境が限定されること、計測したデータに対して人手によって修正する作業が必要であることなどの理由により、日常生活の膨大な運動を計測・蓄積することは困難である。そこで、17 個の IMU センサーを装備したウェアラブルモーションセンサーズーツを利用して日常動作の計測を行った。蓄積した運動データセットの規模は、総時間は 813,600 秒（被験者数 10 名、計測日数 93 日）である。

計測した運動データをクラスとして識別するだけでなく、言語として理解する運動認識システムを構築するには、運動とそれを説明する文章のデータセットが必要である。運動に対して東京大学の学生が付与した説明文のデータセット（28,000 文章）とクラウドソーシングを利用して収集した説明文章のデータセット（423,000 文章）を構築した。クラウドソーシングでは、ユーザーは Youtube に掲載された運動を閲覧し、指定された時刻に対して文章をつけるマイクロタスクを設計した。

## 2. 3 運動と言語の統計的機械学習

前節の運動および説明文のデータセットを用いて運動と言語を結びつける統計モデルを開発した。運動データは HMM によって分類される。分類結果とその運動データに付与された文章中の単語を統計的に結びつけるモデルと文章中において単語の出現規則を統計的に表現するモデルを学習する。この 2 つのモデルを利用することによって、運動の分類結果に強く結びついている単語が前者のモデルから連想され、後者のモデルによってそれら単語の適切な並びを計算することによって運動から文章が生成することができる。運動から文章を生成する処理を探索問題として表現することによって効率的に計算することができる。

図 4 に運動から文章を生成する実験の一例を示す。各運動を適切に説明する文章が生成されていることが確認できる。運動から文章の生成に関して、定量的に評価するために、動きと生成された文章を見て、その文章が説明文として正しいかどうかをチェックするマイクロタスクをクラウドソーシングに掲載した。236,400 の回答を収集し、生成された文章が正しいと回答した割合は 48%であった。

運動の学習、識別、文章生成の並列計算アルゴリズムの考案およびクラウド計算機への実装を行った。15,429 台（123,432 プロセス）の計算機を活用して、運動の学習および識別計算を 9,145 倍、7,481 倍高速化することや 20 台（320 プロセス）の計算機を用いて文章生成を 100 倍高速化することを実現した。

## 3. 今後の研究開発成果の展開及び波及効果創出への取り組み

運動データの大規模データベースおよび運動と言語を結びつける機械学習法が、本研究課題の基盤技術である。これら開発成果は、(A)カメラからの全身運動計測、(B)実社会の人間行動の認識・言語化、(C)言語を用いた人間・ロボットインタラクションの新たな展開を見せようとしている。データセットのカメラ映像と 3 次元運動のビッグデータと深層学習の組み合わせはカメラから 3 次元運動を計測する技術を革新する。行動認識と言語化技術を、スポーツのスキル・集団行動戦略の解析、審判によって採点していた競技の自動スコアリングといったスポーツ分野へ展開している。さらに、言語と運動を繋げる人工知能は、言語指令に応じた行動を生成・制御するヒューマノイドロボット技術へ応用されようとしている。

## 4. むすび

全身運動から各運動に応じて識別に適した特徴量を抽出する手法、手指の動きや筋張力といった情報へ運動データを拡張する方法論を開発し、運動の識別を向上する基盤技術を確立した。日常の多様な運動を言語として認識するために、膨大な運動データと文章データを蓄積したデータベースを構築し、運動と文章を統計的に結びつける機械学習の枠組みを開発した。運動を言語として認識するための基盤となりうる技術であることを実験によって確認した。

### 【誌上発表リスト】

- [1]Wataru Takano, Yoshihiko Yamada, Yoshihiko Nakamura, "Generation of Action Description from Classification of Motion and Object," Robotics and Autonomous Systems, Vol.91, pp.247-257, (2017.2.9)
- [2]Wataru Takano, Ikuo Kusajima, Yoshihiko Nakamura, "Generating Action Descriptions from Statistically Integrated Representations of Human Motions and Sentences," Neural Networks, Vol.80, pp.1-8, (2016.3.16)
- [3]Wataru Takano, Junya Obara, Yoshihiko Nakamura, "Action Recognition from only Somatosensory Information Using Spectral Learning in a Hidden Markov Model," Robotics and Autonomous Systems, Vol.78, pp.29-35, (2016.1.19)

### 【本研究開発課題を掲載したホームページ】

<http://webpark1757.sakura.ne.jp/motiondb/>