

機関番号：14603

研究種目：基盤研究（B）一般

研究期間：2008～2010

課題番号：20300071

研究課題名（和文） 運動学習支援の強化学習システム

研究課題名（英文） Reinforcement Learning System for Assistive Motor Learning

研究代表者

柴田 智広 (SHIBATA TOMOHIRO)

奈良先端科学技術大学院大学・情報科学研究科・准教授

研究者番号：40359873

研究成果の概要（和文）：ヒトの運動学習を効果的に支援する知能システムを構築し実証実験を推進した。このシステムは、モーション等学習者の生体情報を観察しながら、ロボット等で物理的・認知的に学習を支援する。また、異なる身体特性や学習能力を持つ学習者に適応しつつ、学習者が支援装置に頼り切りにならない支援方法を、強化学習法を用いて能動的に探索する。基礎的な学習課題として、ダーツ投擲運動を選んで実証実験を行い、本システムの有効性を確認した。

研究成果の概要（英文）：In this study, an intelligent system for effectively assisting human motor learning has been developed and evaluated through assisting experiments. This system physically and cognitively assists motor learning with such devices as robots, observing the trainee's biological information such as motion. The system actively explores adaptive ways that are adaptive to the trainee's physical characteristics and learning ability, and that prevent the trainee from learning to completely depend on the system. The feasibility and plausibility of the system has been validated through experiments for which we chose dart throwing task as one of the simplest throwing tasks.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	7,600,000	2,280,000	9,880,000
2009年度	4,600,000	1,380,000	5,980,000
2010年度	1,300,000	390,000	1,690,000
年度			
年度			
総計	13,500,000	4,050,000	17,550,000

研究分野：工学

科研費の分科・細目：知能機械学・機械システム

キーワード：人間機械系，運動学習支援，強化学習，ロボティクス，生体計測

## 1. 研究開始当初の背景

(1)近年、運動学習やリハビリテーションに対する情報科学の応用に注目が集まってきている。そのアプローチとしては大きく2つに分けて俯瞰できる。一方はデータ解析側からのアプローチであり、他方はVRやロボットなどを利用した支援に重点を置いた研究であ

る。前者については、モーションや筋電位などの情報を直接用いて熟達度や健常者と障害者の比較を行って分析するものから、大雑把な、もしくは精緻なモデルを導入して学習者の内部情報を推定するもの（StanfordのDelpのグループや、東京大学の中村仁彦のグループが著名）までである。しかし、比較はできてもそれではどのように目標タスクを達

成させれば良いかという方法論は確立されていなかった。

(2) 一方、VRやロボットなどを利用した支援側のアプローチとしては、MITのHoldenらや東工大の小池らのものをはじめいろいろと存在するが、リハビリテーションを目的とした研究では、ゆっくりとした運動学習（軌道追従）が目標であり、健常者であってもロボットを直接ガイダンスに使うようなアプローチでは、大自由度の細かな軌道の直接伝達は今のところ困難であるし、安全面からゆっくりとした運動学的情報しか教えられなかった。また、一般に、学習者には多様な身体特性があり、かつ習熟につれてその特性も変化すると考えられる。さらに、常に同じ指導方法（例えば実現軌道の視覚的提示）を続けるのではなく、適宜切り替える必要がある可能性もある。一方、一般に、学習者は支援装置に過適応してしまうため、装置を外すと追加学習が必要となる。

(3) 従って、運動学習支援装置は、学習者に適応的であり、かつ学習者の支持装置への過適応を予防的である必要があると考えられた。そこで研究代表者は、運動学習支援という課題は一般に、学習者を環境、指導方法を方策、タスクの達成度を報酬関数、に対応付けた指導者の強化学習問題と定式化すれば良いのではないかと着想するに至った。従来研究は学習者がどう学習するかという観点に重点が置かれていたが、本研究は、指導者が学習者と関わりながら、学習者に適した指導方法をオンラインで能動的に探索するという、独創的なものである。

## 2. 研究の目的

大自由度かつ冗長なヒト（学習者）の身体運動学習の動的過程において、学習者から得られる生体情報と運動パフォーマンスの計測データに基づいて、学習者の支持装置への過適応を予防しつつ、その学習者に適した指導方法をオンラインで能動的に探索する運動学習支援システムを、強化学習と特徴抽出に注目して構築し、行動実験によりその有効性を検証すること、とした。

## 3. 研究の方法

### (1) 運動学習支援システムの開発

運動学習支援システムは大きく、計測部、支援部、学習部で構成した。計測部はモーションキャプチャ、力覚センサ、表面筋電位センサ、眼球運動計測装置で構成した。モーションキャプチャは近赤外反射マーカを使い、また表面筋電位センサは小型で無線伝送できるものとし、被験者に極力拘束感を与えずに

済むようにした。

支援部は、ロボットアーム、視覚情報提示用モニター、音声提示用スピーカ、機能的電気刺激（Functional Electrical Stimulation; FES）装置で構成することとした。ロボットアームとしては、電気モータ駆動型のロボットの他、空気圧駆動型のロボットアームも準備することとした。前者は従来研究のように、目標軌道を学習者に直接低速で伝達する際に有用であるが、安全面に更に配慮するには空気圧駆動型がより望ましいと考えた。FES装置は時間応答性の良い刺激を与えるために準備することとした。

学習部へは、適応的な運動学習支援を行うため、学習者を環境、指導方法を方策、タスクの達成度を報酬関数と定式化して強化学習アルゴリズムを実装した。学習者と学習部がタスクを共有できているため、学習者と学習部によりマルチエージェント系が構成されているのではなく、変動する環境（学習者）においてタスク達成度（報酬）を最大化するシングルエージェント問題と定式化した。また、高次元空間における強化学習は困難であることが知られているため、大自由度の運動学習課題を対象とする本研究では、状態空間の低次元化を図る必要がある。本研究では、熟達者と非熟達者の運動比較を行い、報酬に相関した特徴量の抽出を行った。

### (2) 運動学習支援システムの検証

本研究では、学習対象となる運動タスクとして、素早い運動の必要な投擲運動のうちダーツ投擲運動を選定し、支援システムの検証を行った。ダーツ投擲運動は次のような意味で最も単純な投擲運動の一つであるため、基礎研究対象として適していると考えた。まず、タスクのスキルが、ダーツボード上の位置で容易に判定できる。次に、熟達者は、投擲に必要な片腕以外の身体は固定している（我々の計測実験からもこれを確認できた）。最後に、ダーツの質量は軽く、大きな運動量を与える必要もないので、ダーツが投擲運動に与える力学的影響は比較的無視できる。このように、単純な運動ではあるが、初心者にとってダーツボードの中央に高い確率でダーツを当てるのは困難であることも事前実験で確認した。

## 4. 研究成果

### (1) 熟達者と非熟達者の比較

どのような方法でダーツ投擲運動学習の支援をするか、また強化学習で探索する状態空間を低次元化するかを決定するために、本研究ではまず、熟達者と非熟達者のモーションや表面筋電位信号を計測し、特徴比較を行っ

た。なお、本研究においては、ダーツボードの中央に精度良くダーツを刺せることをもって、熟達度が高いと定義した。本研究における被験者の熟達度分布は図1に示す。横軸は正規化スコアであり、中央にダーツが刺さった場合を1としたものである。縦軸はその標準偏差である。図1から、本研究に参加した被験者を容易に熟達者と非熟達者に二分可能であることが分かる。

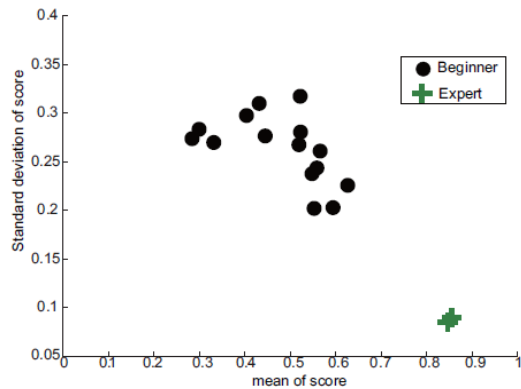


図1

そして、熟達度の違いは、運動学的に明確に現れることが分かった。一例として、図2に投擲運動中の手の高さの軌道を、50試行分重ね書きして示した。左のパネルは典型的な熟達者のもの、右のパネルは典型的な非熟達者のものである。投擲にはaiming, take-back, throwingの3つのフェーズがあるので、各パネルの下部に、時間とフェーズとの対応をイラストで示した。この図から、熟達者の手の高さは試行によらず常にほぼ同じ軌道を描くが、非熟達者の軌道はばらつきが大きいことが分かる。

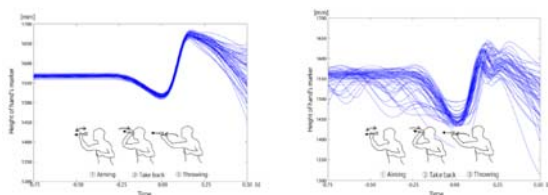


図2

手の高さのばらつきは、肘の高さのばらつきに起因する。さらに分析を進めた結果、投擲中の肩位置の投擲方向へのばらつきにも大きな差があることが分かった(学会発表2, 5, 6, 8)。

その他、動力学的な差異およびも見出した。手首(1自由度)、肘(1自由度)、肩(3自由度)関節からなる3リンクモデルを用いて、被験者のダーツ投擲運動を解析したところ、

二つのことが分かった。第一に、投擲運動の軌道が最適制御の結果だと仮定した場合、作業座標系で定義された躍度最小規範および、関節座標系で定義されたトルク変化最小規範とを比較したところ、後者の方がデータを良く説明することが分かった(学会発表5, 6, 8)。第二に、肘関節や肩関節まわりのトルクを筋肉が発生するトルク、重力が引き起こすトルク、相互作用トルクの3つの成分に分解し、スコアとの相関を調べたところ、相互作用トルクと正の相関が見られた(EMBC'11投稿中)。

このように動力学的な知見も得られたものの、本研究ではまず最も基礎的なアプローチをとって、前述の運動学的差異に着目して運動学習支援の研究を推進した。

## (2) 非熟達者の支援実験と結果

非熟達者の投擲動作中に肘や肩位置のブレが生ずることを防ぐため、上腕肘近くを固い杖的な物で下部から支持(半固定)することによる効果を見る実験を行ったところ、図3に示すように劇的な効果が見られた。左のパネルには初心者の投擲運動を、右のパネルには、その初心者の上腕をロボットで半固定した時の投擲運動を示した。半固定無しでは、肩や上体が投擲方向に前のめりになるため、バランスを取るため片足が上がって身体バランスが崩れている。それに対し、半固定ありでは、熟達者に近い投擲姿勢が、装着直後に実現された。



図3

本研究では、支持装置の個人適応と、個人の支持装置への過適応防止に着目している。そこで、ロボット無、ロボット有・適応無、ロボット有・適応有、の3種類の実験条件を準備して、提案する運動学習支援システムの有効性の検証を行った。ここで、適応有条件では、スコアと腕がロボットに与える鉛直下向き方向の力積との和を報酬とし、肘を支えるインピーダンス制御則のうちバネ定数を方策パラメータとする強化学習を適用し、支持装置への過適応防止の実現を狙った。直感的には、最初硬かった支持部が、スコアが上昇するにつれて柔らかくなるので、上腕固定に頼れなくなる。ロボットに、学習速度や記憶を確認するため、実験は2日間行った。図4

に示すように、ロボット無条件ではスコアの上昇は見られなかったが、ロボット有・適応無条件では、初日実験前スコアと二日目の練習中（ロボット有）のスコアの間に差が見られた ( $p < 0.05$ )。ロボット有・適応有では初日実験前スコアと二日目の実験後のテスト（ロボット無）との間に差が見られた。これは、支持装置への過適応を防止するという提案システムの目的が達成されたことを示唆する。

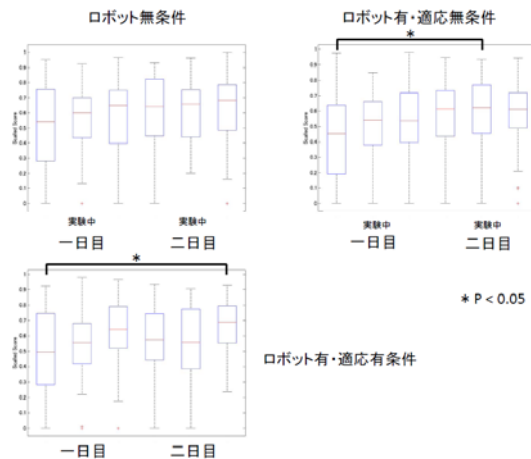


図 4

本研究期間中、機能的電気刺激（FES）装置の運動支援利用まで到達できなかったが、前述のように、熟達者と非熟達者の投擲運動の間の動力学的差異についての研究成果も出ており、今後それらに基づいたよりタイミング精度を重視した支援をFESを用いて行うことを計画している。

## 5. 主な発表論文等

〔雑誌論文〕（計 4 件）

1. Tamei, T. and Shibata, T. Fast Reinforcement Learning for Three-Dimensional Kinetic Human-Robot Cooperation with an EMG-to-activation Model. *Advanced Robotics*, 25(5), 563-580, 2011, 査読有
2. 稲邑哲也, 柴田智広. 動作とシンボル表現の相互変換のための原始シンボル空間を用いた全身動作パターンの内挿・外挿. *日本ロボット学会誌*, 21, 144-153, 2010, 査読有
3. 池田聖, 柴田智広, 和田佳郎. 複合現実環境内での輻輳性眼球運動の計測：直線加速感を測る客観的指標を求めて. *画像ラボ*, 21, 2010, 34-37, 査読無
4. 池田聖, 柴田智広, 和田佳郎. 輻輳性眼球運動に基づく直線加速感評価のための複合現実環境の構築. 64(2), 244-247, 2010, 査読有

〔学会発表〕（計 10 件）

1. Hayashi, K. Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection. *IEEE International Conference on Data Mining (ICDM 2010)*, 17 Dec 2010, Sydney, Australia
2. Shibata, T. Adaptive Robotic Training for Darts based on Comparisons of Experts and Non-experts. *The 28th Annual Conference of the Robotics Society of Japan*, 22 Sept 2010, Nagoya, Japan.
3. 和田佳郎. 健常人の傾斜感覚に対する耳石器代行装置の効果. *ニューロコンピューティング研究会*. 2010年7月27日, 京都市.
4. Tamei, T. Three-Dimensional Virtual Force Sensing for Kinetic Interaction with a Robot. *5th IEEE International Conference on Systems of Systems Engineering (SoSE 2010)*, 23 Jun 2010, Loughborough, UK
5. 大林千尋. ダーツ投擲動作における熟達者と非熟達者の比較. *第 24 回生体・生理工学シンポジウム*, 2009.9.25, 仙台市.
6. Obayashi, C. Comparison of experts and beginners in throwing darts based on optimization criteria. *The 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2 Sep 2009, Minnesota, USA.
7. 為井智也. ユーザーの生体情報を用いたロボットとの力覚的協調の強化学習. *電子情報通信学会総合大会*, 2009.3.18, 松山市
8. 今井暁. 熟達者と非熟達者によるダーツ投擲の筋活動と精度. *電子情報通信学会ニューロコンピューティング研究会*, 2009.3.12, 町田市
9. Tamei, T. Policy Gradient Learning of Cooperative Interaction with a Robot Using User's Biological Signals. *The 15th International Conference on Neural Information Processing*, 27 Nov 2008, Auckland, NZ
10. Shibata, T. Reinforcement Learning for Assisting Humans. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 22 Sep 2008, Nice, France

〔図書〕（計 1 件）

1. 柴田智広. ナノオプトニクスエナジー ロボット情報学ハンドブック, 2010, 678-684

〔産業財産権〕

○出願状況（計 0 件）

○取得状況（計 0 件）

〔その他〕

ホームページ等

<http://hawaii.naist.jp/~tom/kaken/>

## 6. 研究組織

### (1) 研究代表者

柴田 智広 (SHIBATA TOMOHIRO)  
奈良先端科学技術大学院大学・情報科学研究科・准教授  
研究者番号：40359873

### (2) 研究分担者

和田 佳郎 (WADA YOSHIROU)  
奈良県立医科大学・医学部・講師  
研究者番号：80240810

### (3) 研究協力者

為井 智也 (TAMEI TOMOYA)  
奈良先端科学技術大学院大学・情報科学研究科・研究員  
研究者番号：40548434