

特集 「ニューラルネットワーク研究のフロンティア」

ロボティクスと深層学習

Robotics and Deep Learning

尾形 哲也
Tetsuya Ogata

早稲田大学
Waseda University.
ogata@waseda.jp, <http://ogata-lab.jp/>

Keywords: robotics, deep learning, multimodal learning, end to end learning, sequence to sequence learning.

1. はじめに

近年、メディアなどで着目されている深層学習・ディープラーニング (Deep Learning: DL) の技術は、画像認識、音声認識、言語処理など多分野においてその有効性が確認されるとともに、さまざまな応用手法が毎日のように提案されている。DLは、一般に数千程度の入力をもった10層程度の階層型神経回路モデルを指す。その最大の特徴は、画像や音声など、多次元かつ大量のデータの処理において、これまで専門家のみがその知見と経験によって設計、選択可能だった高次特徴量を、ほとんど考慮する必要なく高い性能を得られる点にある。具体的には、教師信号さえ与えられていれば、対象の高次元データとそれに対応する教師信号を、神経回路モデルに“ほぼそのまま”与えることで、必要な出力を得るための特徴量が、神経回路モデル内に“自己組織化”される。この特徴により、画像、音声、言語などこれまで全く異なった研究領域で扱われていたモダリティーを、DLという“同様の枠組み”によって扱うことができる。そして、そこで得られる性能は従来法に迫る、課題によってはしのごの結果が得られている。

DLについては、今後も各モダリティー別に应用研究が盛んに行われると予想されるが、この“ほとんど同様の枠組みで多様なモダリティーを扱える”という特徴から、自然に“DLを利用したマルチモーダル情報の統合”という応用が考えられる。特に実世界で行動するロボットシステムは通常、カメラ、マイクロフォン、距離センサ、接触センサ、そしてアクチュエータなどを備えたマルチモーダルシステムであり、DLの活用が大きく期待されている。また著者らはすでにこの点に着目した複数の研究を行ってきた。本稿では、特に「認識」、「運動」、そして「言語」をキーワードとして、これまでのDL技術のロボットへの応用研究を概観しつつ、著者らがこれまで行ってきたマルチモーダル研究の一連の事例を紹介する。そしてその可能性、またDLに関する将来展望について述べていきたい。

まず導入として、2章で「認識」の研究例、3章で、「運動」の研究例、さらに4章では、「言語」の研究例を紹介していく。最後に5章で全体をまとめる。

2. 認識と状態評価

序章で述べたように、ロボットはマルチモーダル処理が不可欠であり、多くの認識システムを併用している。特にロボットビジョンは、重要な研究分野であり、近年DLを用いた研究例がいくつか報告されている。例えば、Ian Lenzらは、DLにより対象物の距離画像から把持のためのハンドの位置と方向(4次元)を出力させる手法を提案した [Lenz 14] (図1)。またY. Yangらは、DLの位置手法であるCNN (Convolutional Neural Network) を用いて、YouTubeにある人間が調理するビデオから、直接オブジェクト48種類と6種類の把持タイプを識別しロボットの動作に適用している [Yang 15]。

このようにこれまでロボットビジョン分野の問題とし

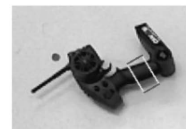
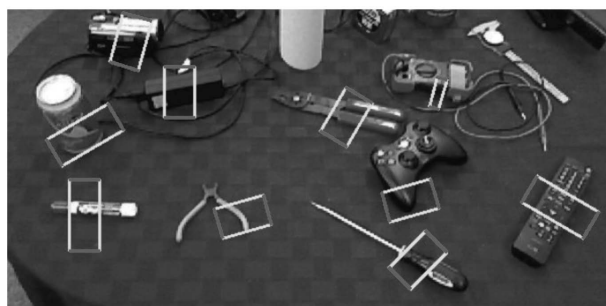


図1 Baxterによる物体把持 [Lenz 14]

て扱われてきた諸問題が、DLにより大きな発展を遂げつつある。

また特にロボットとDLとの関連で注目を集めているのが、Deep Q-Learningの手法である [Mnih 13]。本来、強化学習は、状態 s と有限数の行動パターン a の組合せからなる Q 値を、環境から受け取る報酬値から学習し行動選択に利用する手法である。コンピュータ内のシミュレーションでは十分に有効機能し、その特性と生物行動学習の対応など興味深い研究事例も多くある。しかし実世界で行動するロボットに利用する場合、扱うべき次元数が非常に多く Q-Learning を直接適用することは困難であった。

Deep Q-Learning では、DLが高い入力次元を扱うことができることを利用している。具体的には Atari 2600 の7種類のゲームを対象にして、プレイ中の画面 (84 × 84 pixel, 4 フレーム分) を高次元のまま直接入力し、4 ~ 18 種類の動作コマンドを選択、出力した際の Q 値を学習させる。動作中常に報酬は与えられないうえに、すべての動作画面を学習に利用できるわけではないが、DLの汎化性能によってのみ学習画面における動作選択での Q 値も推定ができるようになる。学習の結果、三つのゲームで人間のプレーヤーに勝利するまでになっている。このようにDLは環境の認識のみならず、その状態評価においても高い性能を示すことがわかる。

3. 運動の生成

Deep Q-Learning は“高次元入力”をもつ自律システムの学習の可能性を示しているが、ロボットへの応用には、さらにもう一つ問題がある。例えば前章で述べた例では、自律ロボットが画像などのセンサ情報から、動作に必要な少次元情報 (4次元の物体の位置、方向)、もしくは動作パターン (例えば6種類の把持パターン) の選択が行われる。ここでは入力である状況 s は高次元であるが、出力である行動 a は、あらかじめ少自由度の空間、もしくは有限の状態 (クラス) に明確に分類できるという暗黙の仮定がある。この仮定を受け入れられれば、DLの認識、もしくは状態評価の技術は、そのままロボットの有限動作パターンへのマッピングや評価の技術として利用でき、前章での研究例のように強化学習などの枠組みも比較的簡単に適用できる。

しかしながら本来、実世界におけるロボットの動作出力は、多自由度関節の連続時系列データである。運動や行動の状態の分節化は、それ自体がこれまでの知能ロボティクスにおける一大問題であり、テレビゲームの押しボタンの組合せのように簡単には定義できない。そしてこれまでこの“動作パターン”の設計・選択は、熟練した設計者がその知見と経験によって行う特殊な作業であったのである。

しかし序章で述べたDLの本来の特徴を考えれば、こ

の“シーン”や“動作パターン”さえも、DLに学習 (自己組織化) させることができないか、という発想に至る。つまり高次の入力 (画像や音声の時系列入力) から、必要な高次の出力 (複数の関節時系列出力) を直接得る、という発想に至ることになる。このように得られる入力から欲しい出力を直接学習する、という方法論は“End to End Learning”と呼ばれ、DLをシステムに適用する際に強力かつ本質的なアプローチだといえる。

例えば S. Levine らは、PR2 という人間型ロボットを利用し、現時刻の1枚の視野画像入力から次時刻のロボットの複数関節を直接 CNN で出力させることで、一連の動作を実現する手法を提案している。ハンガー掛けやボトルの蓋閉めなど、複数の動作を学習させ、物体位置変化などにも極めてロバスタかつスムーズに動作が行えることを示した [Levine 15]。図2にデモの様子を示す。強化学習による行動探索を行ってはいるが、画像と関節を“直接”つなぐという End to End Learning の方法論を取ったため、動作パターンの設計という困難な問題から解放され、効果的な動作獲得を実現している。しかし入力は1枚の画像であるため、時間的な予測などは難しい。

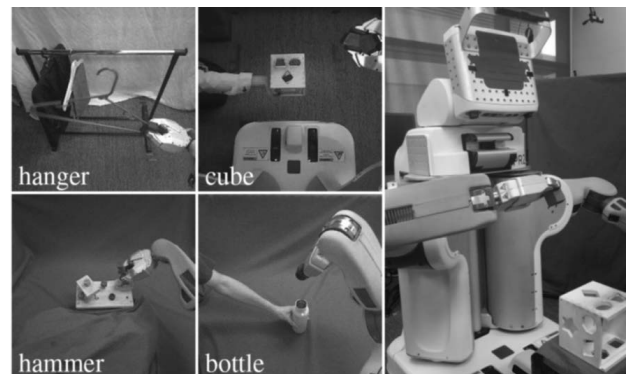


図2 PR2による物体操作タスク [Levine 15]

著者らは S. Levine の仕事に先んじて、自律ロボットの行動生成の枠組みに利用するフレームワークを提案している [Noda 14]。

ここで重要なのは、目的は物体や動作の認識ではなく、生成にある点である。そこで我々は CNN ではなく、Time-delay 型の Deep Autoencoder (DA) を、ロボットによる複数物体の操り動作学習 (視聴覚運動時系列の統合学習) に適用することを考えた。全体のモデルを図3に示す。具体的には、画像、音響信号、モータの各モダリティを複数の DA によって統合する。

Aldebaran Robotics 社の NAO を用いて直接教示により生成した、6種類の物体操作タスク (図4) に関する学習実験を行った。各動作パターンにつき6~10回教示を繰り返してデータを取得した。

画像、音響信号などの多次元 (3000 ~ 4000 次元) のデータを DA で、それぞれ30次元にまで圧縮する。

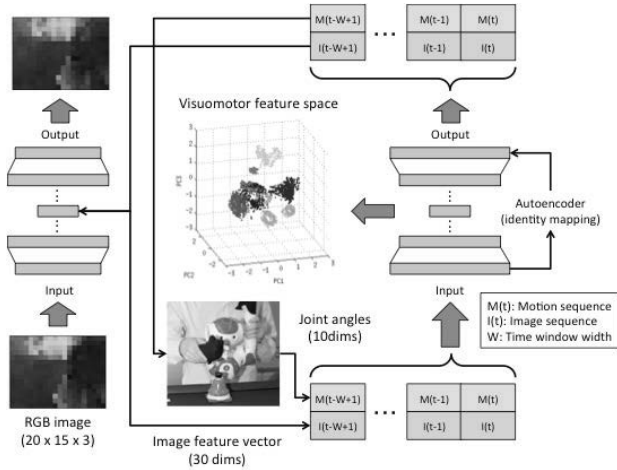


図3 DLによるロボットのマルチモーダル統合システム

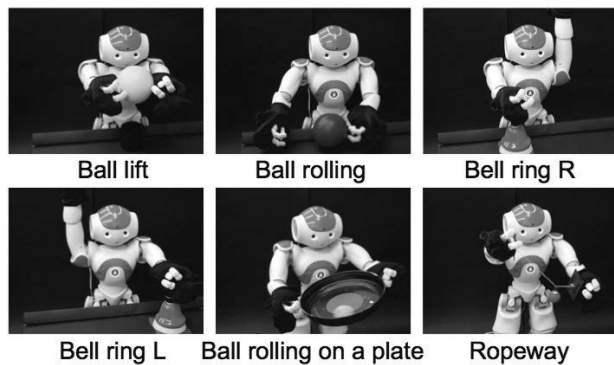


図4 NAOによる物体操作

この次元圧縮の結果得られた画像と音声の特徴ベクトルとロボットから得られた関節角度の統合についても、HMMではなくDAを用いて統合学習させた。具体的には、30ステップの時間窓で切り出したものを学習器への1回の入力とし、1ステップずつ窓をスライドさせながら学習データを作成した (Time delay 型 NN)。

このモデルにより、ロボットの複数の動作に対応する、感覚・動作の時系列データに対して、感覚特徴量だけでなく運動パターンを内部に独自に自己組織化することが確認された。獲得された内部表現を主成分分析により三次元化したものを図5に示す。この図中、各動作が色分けされているが、これらは人間の教示者がイメージした6種類のクラスタに明確に対応しているわけではないことが確認できる。ある動作はその経験数に応じて複数のクラスタを有し、またある動作とある動作は空間的な重なりをもつことで、互いを関連付けている。このように、外部の物理世界、身体構造、学習器 (DL) の連続的なインタラクションが、認識とカップリングした動作表現を自己組織化することがわかる。

さらにこの内部表現を利用することで、クロスモーダルな記憶想起を実現した。例えば、画像や音響データからのロボットの運動が生成できること、さらに身体運動情報、もしくはそれに伴う音響情報のみから、3000次

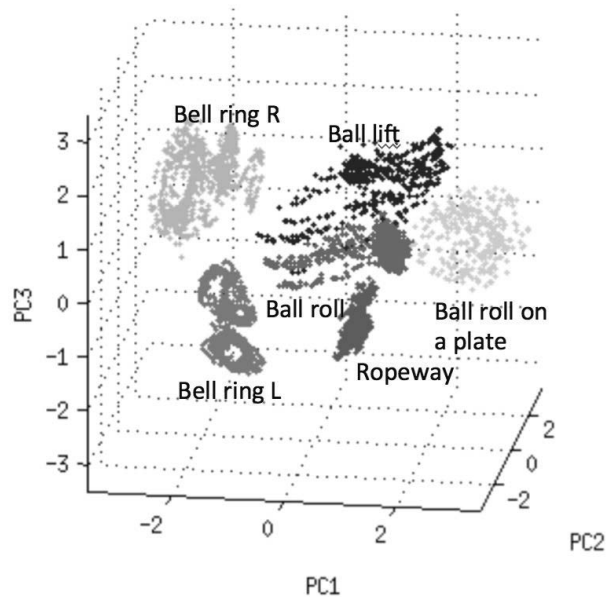


図5 獲得された行動表現

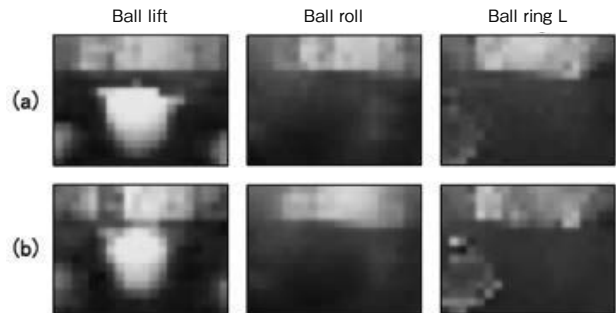


図6 関節角度 (運動) からの画像の予測。
(a) が正解画像, (b) が予測生成画像



図7 PR2によるタオル折り畳み
(右図は連想視野画像)

元の映像情報の連想, 復元 (図6) できることを確認した。このように“非常に大きな次元の表現を直接生成できる”点がDLの特長であり、今後、実ロボットへ応用する際のポイントだといえる。

獲得された内部表現の解析から、提案モデルは視聴覚運動情報をその共起性ととも自己組織的にクラスタとして構造化し学習していることも確認した。

本手法をロボットPR2に適用し、7自由度アームとハンドを利用したタオル折畳み動作生成を行った例を図7に示す [鈴木 16]。テーブル上にある 27×19 cmのタオルを操作する。五つの位置で学習を行い、その後、四つ

の未学習位置で動作を行わせたところ、動作時間に伴って実際のタオル位置に対応し、タオルが折り畳めた。

以上のように、DLを高次元の入力を扱う認識系としてだけでなく、高次元の出力が可能な生成系として利用することで、ロボット動作の生成、状況の推定予測が可能になる。今後の応用の可能性は大きいと期待できる。

4. 言語と動作

ロボットが動作を生成する際、人間とのコミュニケーションが不可欠である。その方法はさまざまなものが考えられるが、自然言語が利用できれば、その有効性は大きい。例えば、2章で紹介したY. Yangらの料理タスクの学習の例では、DLの入力に画像だけでなく、状況を説明する言語情報を加えることで作業シーンの分節化を助け、効率的な学習を行っている。

近年、自然言語処理で盛んに利用されるようになったRecurrent Neural Network (RNN)は、内部ニューロンにおいて再帰結合を有する神経回路モデルである。外部からの入力がない状態であっても自律的な出力を生成可能な時系列データの学習器である。RNNの学習では通常、ネットワークを時間方向に展開して、仮想的な多層神経回路モデル(DL)とみなして学習するBack Propagation Through Time (BPTT)が利用される。そのため近年のDLの発達の文脈の中で改めて着目を浴びている。

特に、異なった時間スケールの学習を行うニューロンを混合させたRNNとして、MTRNN (Multi Timescale RNN [Yamashita 14])やLSTM (Long Short Term Memory [Hochreiter 97])などが提案されており、従来のRNNを超える性能が報告されている。

RNNは、言語処理分野での応用研究が盛んであり、例えば言語翻訳や対話などの学習例がある。ここでは“Sequence to Sequence Learning [Sutskever 14]”と呼ばれる、対象とする2種類の文章をRNNにおける連続した系列として学習させる方法が利用されている。対象となる高次元入力を直接マッピングする考え方は、前述したEnd to End Learningに非常に近い。

ここで重要なことは、RNNはDLと同様の神経回路モデルであるため、互いをそのまま結合して別目的に再利用できる点である。例えば、GoogleのVinyalsらは物体認識用に学習されたCNNの内部出力を、文章生成用のRNNの入力として利用し“転移学習”することで、画像を文章で説明するImage Caption Generatorを開発している [Levine 15, Vinyals 15]。

著者らは、これまで言語間の変換にのみ用いられてきたSequence to Sequence Learningを、ロボットのマルチモーダル学習に応用することを考えた。具体的には、ロボットの運動と言語の統合学習に応用する手法を提案した [Yamada 15]。手法の概略を図8に示す。ここでは

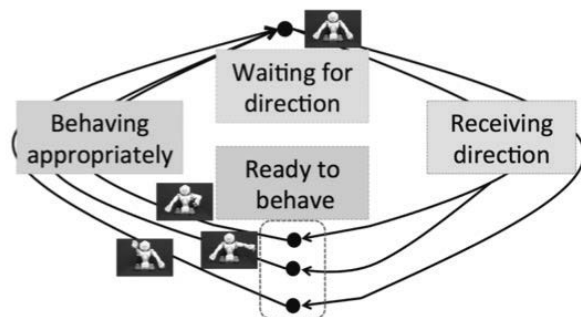


図8 Sequence to Sequence Learningによる運動と言語の連結

一つのRNNに、言語(単語)、感覚(画像)、運動(関節)に関する入出力ノードを与える。そして一連の言語入力にそのまま引き続いて続いて、対応した運動が環境(感覚)に整合した形で生成されるように学習を行わせる。

具体的なタスクのイメージを図9に示す。(1)ロボットNAOの指差しおよびベルたたきの動作を設定した。NAOの前に設置された、(2)二つのベル(左から、Red, Green, Blue)に対する動作を「動詞+左右」, 「動詞+色」で指示する。またベルが同色の場合、「動詞+色+左右」の3単語で指示する。(3)この指示に対してロボットがベルの指差しもしくはたたき動作を行う。このタスクでは、言語の“多義性”を扱っている点が特に重要である。つまり同一動作を表現する複数の文章があり、また同一文章に対応する複数の動作があり得る。またいくつかのベル配置では、色でのみ動作を指示し、ほかのいくつかのベル配置では、場所でのみ指示を行う、といった教示データを作成した。

いくつかのベルを組み合わせた環境で、言語と動作のインタラクション学習を繰り返し行った結果、多義的な言語入力に対して適切に解釈を行い動作が生成可能であること、さらに未学習の言語指示に対しても適切な動作生成が可能であることが確認された(汎化)。さらには人間側の言語入力が終了するまで、動作を生成しない、など交替(ターンテイキング)の動作が、RNNの中の固定点アトラクタとして獲得されることを確認した。

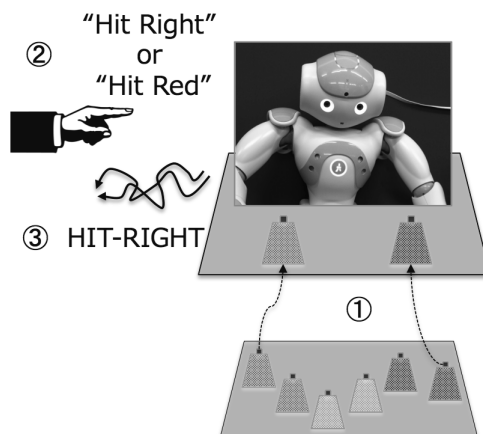


図9 ベルたたきタスク

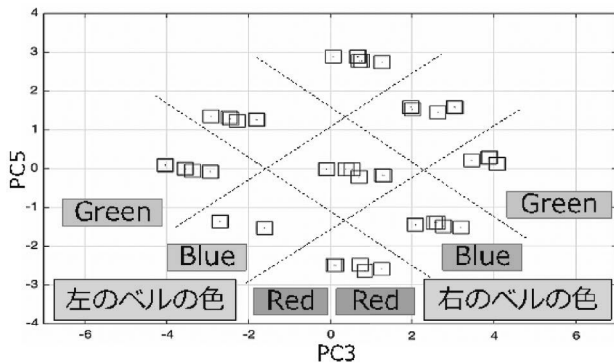


図 10 言語入力後の発火状態 (第 3, 5 主成分)

RNN 内のニューロンの発火状態を主成分解析した結果, 言語と動作のシーケンスが, 大きな周期アトラクタとして表現されていることを確認した. この周期アトラクタは, 入力される動詞, 指示 (方向や色) に応じて分岐し, その後動作を生成した後に待機状態の固定点アトラクタに収束する.

言語入力直後の発火状態の第 1, 2 主成分を可視化すると, 生成する動作パターンに対応したクラスタが形成される. ただ各クラスタは決して一点にはならない. それぞれの動作は類似しているものの, それがどのような言語で指示されたのか, 何回目の動作なのか, といった文脈情報が空間内の微妙な位置の差として表現される.

そして図 10 に言語入力直後の発火状態の第 3, 5 主成分の解析結果を示す. 図中, Red, Green, Blue はそれぞれベルの色を表しており, ここではベル配置を表現していることが確認できる. このように第 1, 2 主成分のような大きな構造だけでなく, 第 3, 5 主成分のような小さな構造にも階層的に情報を埋め込むことができる. 結果としてたった一つの一様な再起結合型神経回路に動作と言語の表現を内包させることが可能となるのである.

現状では, まだ極めて少次元の入出力, 少数語彙に限定した事例モデルではあるが, さらに高機能なロボットモデルへの利用, また上述した MTRNN もしくは LSTM の利用などによって, さらに複雑なタスクと言語に拡張していくことが可能であると思われる.

5. ま と め

本稿では, DL のロボット学習応用について, 特にロボット行動生成の視点から, 著者らがこれまで行ってきた「認識」, 「運動」, 「言語」へ至る, 一連の研究例の概略を紹介した. 1 章で述べたとおり, DL は同じ枠組みで, 異なったモーダル特有の専門的知識を利用せずに適用でき, 高い性能が得られる非常に強力な手法である. 今後, これまでは異分野で扱われてきたさまざまな問題に対して, 新しい可能性を提供すると思われる. また DL 技術の進展から生まれた RNN の各種の新しい手法と応用は, 今後, 言語処理の研究において興味深い拡張がな

れると期待される. このように DL, RNN に代表される人工神経回路モデルの応用は, 今後の新しいロボット知能を実現するうえで重要な一つの「必要条件」だと思われる.

しかし DL が今後のロボット知能の研究における「十分条件」である, とは当然ながらまだとてもいえない. 基本的に DL 自体は, Batch 学習器, そして教師あり学習器であり, ロボットが備えるべき動的なオンライン学習は実現できない. 上記したように DL を利用した強化学習では教師信号出力を事前に Batch 学習しておき, 行動のオンライン学習に利用している. また RNN による言語学習や推論学習においても, 文法や推論を近似したダイナミクスを内部に獲得するだけであり, 人間のように数理計算や論理演算を生み出すことや, 実世界に関する新しい問題を自ら提起する, などの高次の認知機能を実現できるわけではない. DL は人間の一部分の実データ処理に関する機能を超越するのみであり, その能力には少なくとも現在のところ多くの限界があることを認識することが重要であろう.

認識タスクにおいては, ロボットビジョンや音声対話などの分野で導入されていくと見込まれるが, 運動制御などの中心的問題設定に応用されるには, まだ課題が多く残されている. DL において現在のところ大きな成果を上げているのは, 多次元データの認識処理のみであり, 物理的な非線形ダイナミクス (ロボット身体) の制御などには, 直接適用できないためである. 4 章で示した RNN は, 非線形ダイナミクスを学習可能であり, 著者らも現在実ロボット制御への応用基本モデルを提案中であるが [Takahashi 15], 研究はまだ萌芽段階だといえる.

またロボットの運動学習に DL を応用する際に問題となると予想されるのは学習データである. 画像, 音声の学習であれば, すでにネットにあるデータを利用できるが, 3 章の運動学習や 4 章の言語学習を行うためには, 実ロボットが環境において能動的に行動し, そこで得られる感覚と運動に関する一貫した時系列データを (強化学習を利用する場合はその評価・報酬値も), 大量に手に入れる必要がある. しかしながら大規模なロボットにおいて大量の行動データを得るための実験は, ハードウェアの耐久性などの問題を含めて大きなコストがかかると思われ. 例えば, Pinto らは Baxter を用いて 5 万回の物体把持動作を 700 時間もの学習により獲得させている [Pinto 15]. しかしこの手法が他のロボット, タスクにおいて現実的である保証はない. 多様なタスクに共通して利用できる, 安定したロボットプラットフォーム, また多機能な仮想空間の効率的な利用などが重要となるだろう. 多様な研究者との連携を行いながら解決していきたい [JST CREST 15]. 例えば, 著者らの例のように数種類の成功動作を事前学習しておき, 仮想空間において強化学習によりファインチューニングする, といった手法は有効に働くと思われる. 例えば,

DeepMind 社の AlphaGo では、プロ棋士の差し手から事前学習した CNN を利用して、盤面評価推定を行う CNN を自己対戦シミュレーションにより強化学習させる手法で大きな成果を上げている [Silver 16].

DL を実システムに応用する際の最大の課題は、その内部メカニズムの理解が極めて困難である点にある。たとえ高い性能が得られても、誤動作が起こった場合にその原因を特定することが難しい。現在、主に応用されている認識課題や翻訳などでは、大きな問題となっていないが、本稿で示したようなロボットもしくは自動運転技術などの実世界で行動するシステムの学習では、深刻な問題になり得る。これは過去の神経回路モデルでもあった問題である。

4 章の主成分分析で示したように、DL や RNN では内部表現がその大きな構造（例えば、第 1, 2 主成分に埋め込まれた動作情報）だけでなく、小さな構造（例えば、第 3, 5 主成分に埋め込まれたパルの情報）にも階層的に埋め込まれる。また過去の著者らの研究から RNN ではフラクタル構造もその情報表現に多く利用されていると考えられる。これらのメカニズムにより、非常に長い時系列の文脈や空間の複雑な構造を自己組織化可能となり、高い性能が得られていると考えられるのである。よって単純な統計的な解析だけでは、そのメカニズムの説明に十分ではない可能性が高い。階層化サイズの拡張、そして複雑系、つまりダイナミカルシステムとして DL や RNN を数理的に理解することが、今後の重要な研究課題になると考えられる。

謝 辞

本稿で述べた著者の研究は、JST さきがけ領域「情報環境と人」および新学術領域構成論的発達科学の支援によって行われた。ここに謝意を表す。また研究を遂行するうえで、有江浩明君、菅 佑樹君、野田邦昭君、村田真悟君、佐々木一磨君、山田竜郎君など著者の研究室の多くの方々の貢献があった。改めてここに謝意を表す。

◇ 参 考 文 献 ◇

- [Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long short-term memory, *Neural Computation*, Vol. 9 No. 8, pp. 1735-1780 (1997)
- [JST CREST 15] JST CREST 研究, 記号創発ロボティクスによる人間機械コラボレーション基盤創成 (代表: 長井隆行)
- [Lenz 14] Lenz, I., Lee, H. and Saxena, A.: Deep learning for detecting robotic grasps, *Int. J. Robotics Research (IJRR)* (2014)
- [Levine 15] Levine, S., Finn, C., Darrell, T. and Abbeel, P.: *End-to-End Training of Deep Visuomotor Policies*, arxiv:1504.00702 (2015)

- [Mnih 13] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M.: Playing atari with deep reinforcement learning, *Deep Learning Workshop NIPS 2013* (2013)
- [Noda 14] Noda, K., Arie, H., Suga, Y. and Ogata, T.: Multimodal integration learning of robot behavior using deep neural networks, *Robotics and Autonomous Systems*, Vol. 62, No. 6, pp. 721-736 (2014)
- [Pinto 15] Pinto, L. and Gupta, A.: Supersizing Self-supervision: Learning to Grasp from 50K Tries and 700 Robot Hours, arXiv:1509.06825 (2015)
- [Silver 16] Silver, D., et al.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, Issue 7587, pp. 484-489 (2016)
- [Sutskever 14] Sutskever, I., Vinyals, O. and Le, Q. V.: Sequence to sequence learning with neural networks, *NIPS 2014*, pp. 3104-3112 (2014)
- [鈴木 16] 鈴木彼方, 高橋城志, Gordon Cheng, 尾形哲也: 深層学習を用いた多自由度ロボットによる柔軟物の折り畳み動作生成, 第 78 回情処全大 (2016)
- [Takahashi 15] Takahashi, K., Ogata, T., Yamada, H., Tjandra, H. and Sugano, S.: Effective motion learning for a flexible-joint robot using motor babbling, *Proc. 2015 IEEE/RAS Int. Conf. on Intelligent Robots and Systems (IROS 2015)* (2015)
- [Vinyals 15] Vinyals, O., Toshev, A., Bengio, S. and Erhan, D.: *Show and Tell: A Neural Image Caption Generator*, arXiv:1411.4555 (2015)
- [Yamada 15a] Yamada, T., Murata, S., Arie, H. and Ogata, T.: Attractor representations of language-behavior structure in a recurrent neural network for human-robot interaction, *Proc. 2015 IEEE/RAS Int. Conf. on Intelligent Robots and Systems (IROS 2015)* (2015)
- [Yamashita 08] Yamashita, Y. and Tani, J.: Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment, *PLoS Computational Biology*, Vol. 4, Issue. 11, e1000220 (2008)
- [Yang 15] Yang, Y., Li, Y., Fermüller, C. and Aloimonos, Y.: Robot learning manipulation action plans by "Watching" unconstrained videos from the world wide web, *28th AAAI Conf. on Artificial Intelligence* (2015)

2016 年 1 月 18 日 受理

著 者 紹 介



尾形 哲也 (正会員)

1993 年早稲田大学理工学部機械工学科卒業。1997 年日本学術振興会特別研究員 (DC2), 1999 年早稲田大学理工学部助手, 2001 年理化学研究所脳科学総合研究センター研究員, 2003 年京都大学大学院情報科学研究科講師, 2005 年同助教授 (2007 年同准教授) を経て, 2012 年より早稲田大学基幹理工学部表現工学科教授。博士(工学)。2009 ~ 15 年 JST さきがけ「情報環境と人」領域研究員, 2015 年より産業総合技術研究所人工知能研究センター客員研究員を兼務。人工神経回路および人間とロボットのコミュニケーション発達に関する研究に従事。日本ロボット学会, 情報処理学会, 日本機械学会, ヒューマンインタフェース学会, バイオメカニクス学会, IEEE などの各会員。