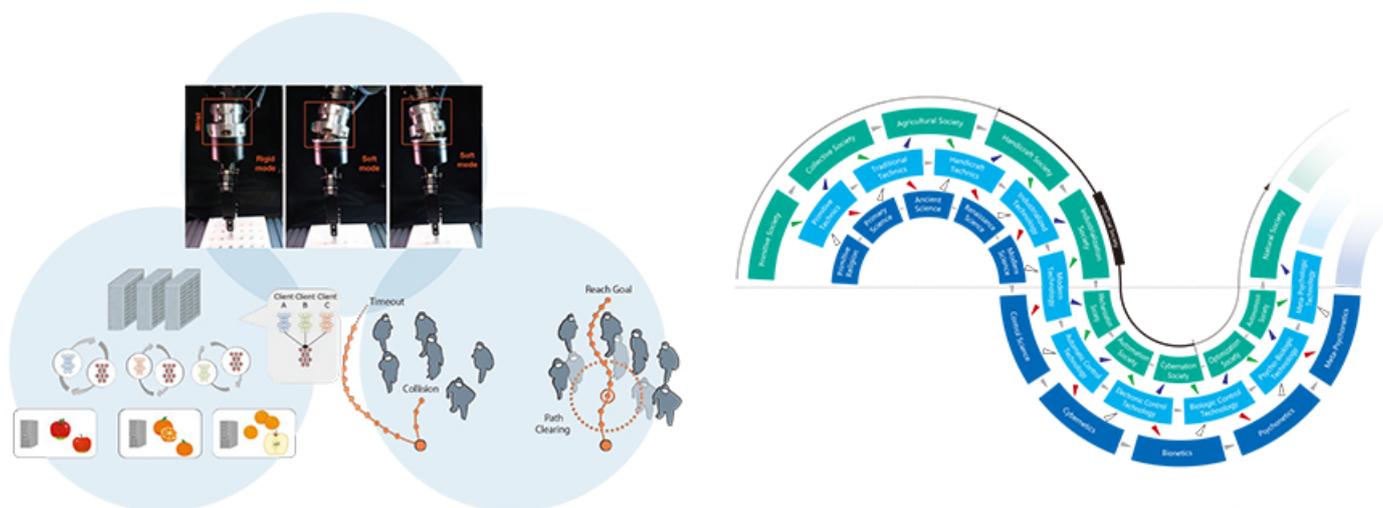


# OMRON 165

Vol.53 No.2  
2021

# TECHNICS®

オムロングループ技術論文誌



特集

## OMRON SINIC X as a Place for Innovation

Pursuing the Forefront of Harmonizing  
Humans and Machines

OMRON

OMRON SINIC X Corporation  
オムロン サイニックエックス株式会社



# オムロンのコア技術

## Sensing & Control +Think

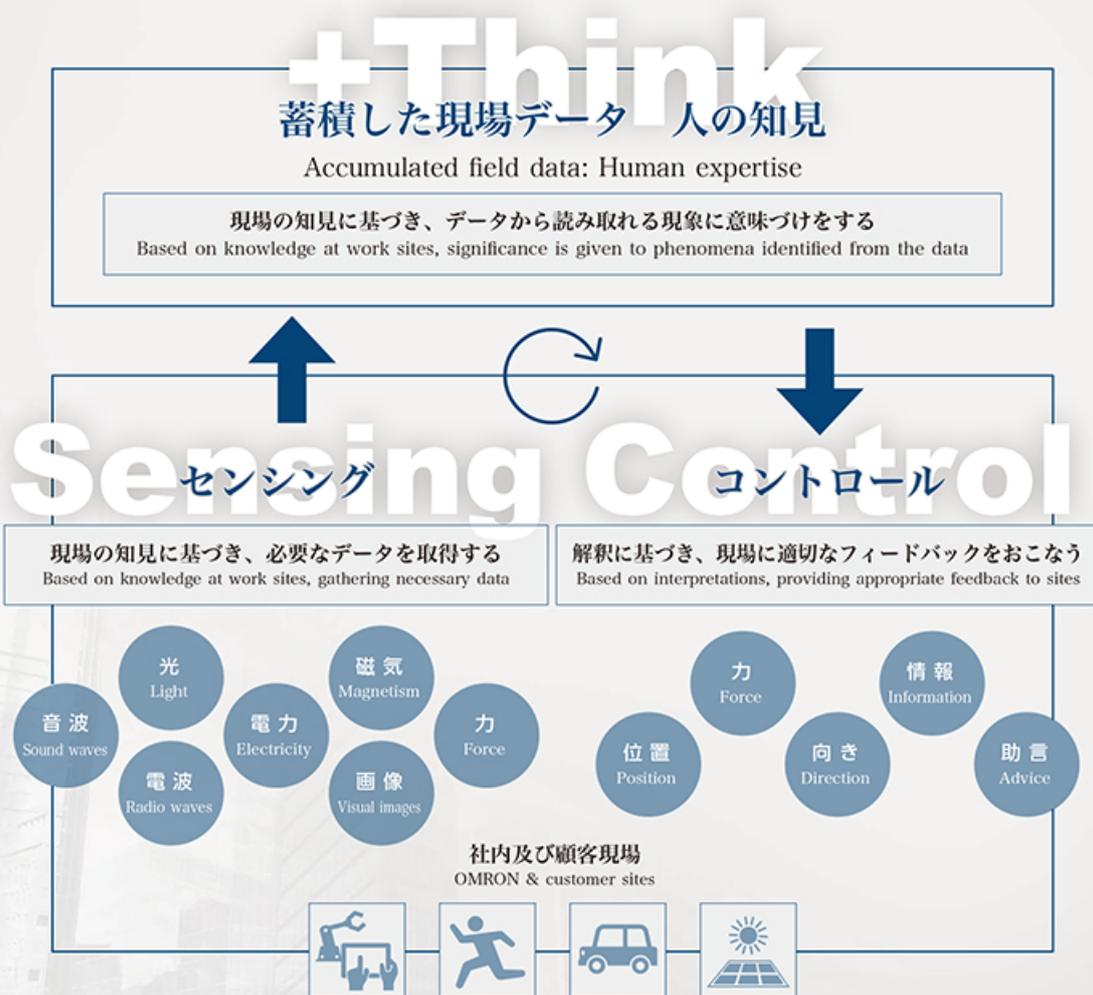
オムロンは創業以来、時代に先駆けた価値を生み出すために、独自のコア技術を進化させてきました。「Sensing & Control + Think」はオムロンが誇る全社共通のコア技術です。

「Sensing」とは現場の知見に基づき、人やモノの状態・情報から必要なデータを取得することです。「Control」とは、「Sensing」によって得られた情報をもとに、現場に適切なソリューションを提供することです。

2011年からこの「Sensing & Control」に「人の知恵」を表す「+ Think」を加えました。AI、IoTなどの技術革新が進む中で、より賢く価値に変換させる必要があるからです。

オムロンでは、「ファクトリーオートメーション」、「ヘルスケア」、「ソーシャルソリューション」を注力する事業領域として、これからも社会に貢献できる価値を最大限に生み出すべくコア技術の強化と進化に取り組んでいきます。

### Sensing & Control + Think



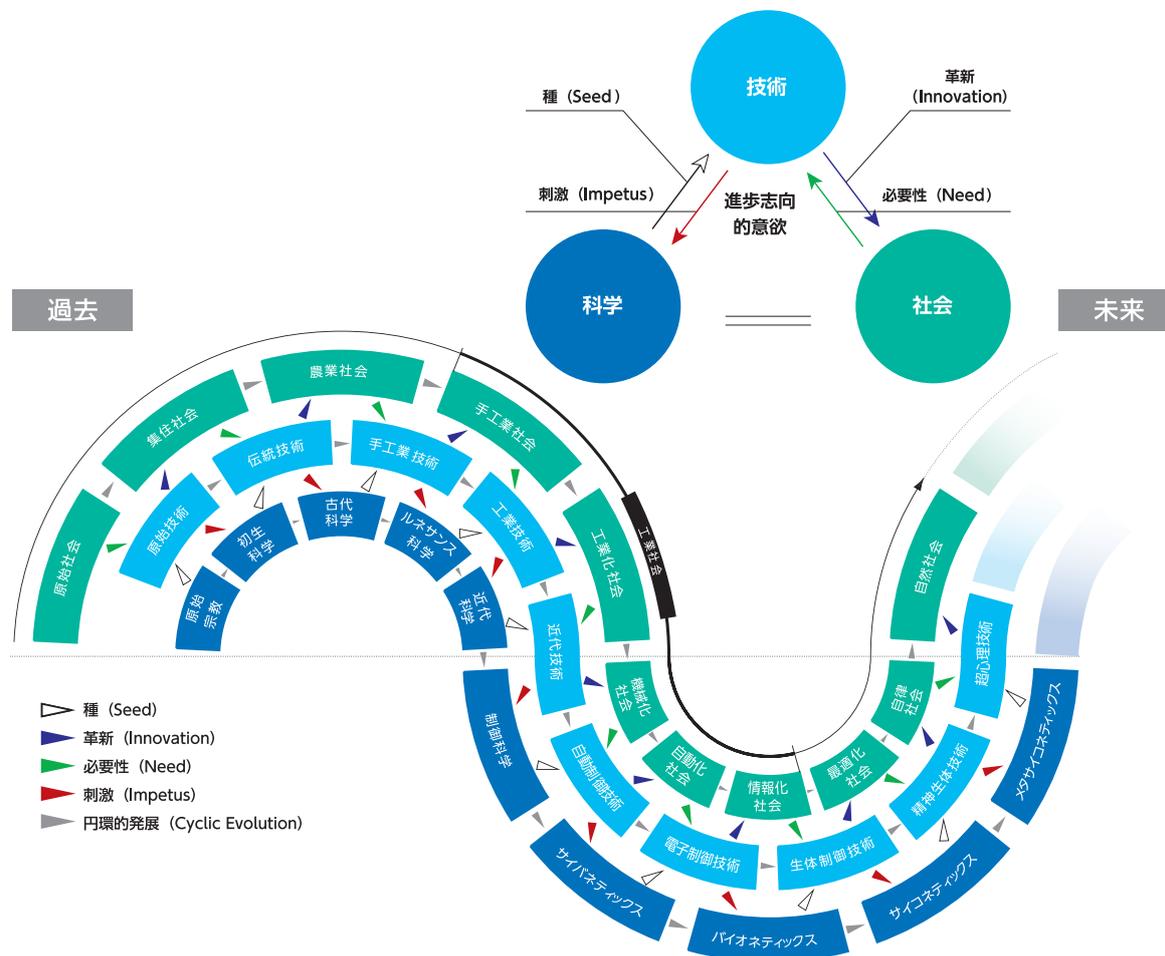
# 科学・技術・社会の相互作用から 未来を予測する「SINIC\*理論」

オムロンの創業者・立石一真は「事業を通じて社会的課題を解決し、よりよい社会をつくるにはソーシャルニーズを世に先駆けて創造することが不可欠になる、そのためには未来をみる羅針盤が必要だ」と考えました。そこで、科学・技術・社会それぞれの円環的な相互関係から未来を予測するSINIC理論を1970年に構築し、国際未来学会で発表しました。以降、オムロンはこれを未来シナリオとし、経営の羅針盤としています。

SINIC理論の基本的な考え方は、科学・技術・社会が相互に作用しながら発展していくというものです。情報化社会を例にとると、1940年代に勃興した総合科学であるサイバネティクスやコンピュータ科学の発展が、新しい電子制御技術、プログラミングなどの種（シーズ）となり、パーソナルコンピュータやインターネットの普及を通じて、情報化社会が実現しました。一方、情報化社会が発達し、より多くのデータを正確に素早く分析・解析したいという社会の必要性（ニーズ）が、CPUやGPUなどの処理装置の性能を向上させ、ディープラーニングなどの人工知能（AI）技術の進化を促し、脳科学や認知科学の新たな展開を刺激していることも挙げられます。

最適化社会の現在は、工業社会から自律社会へと向かうパラダイムシフトの過渡期です。社会の変化が大きく、将来の見通しがつかないように見える今、SINIC理論をオープンにし、さまざまな人々との議論を通して未来創造のための社会の知として活用を進めています。

\*SINIC：Seed-Innovation to Need-Impetus Cyclic Evolutionの略。



科学・技術・社会の相互作用から未来を予測する「SINIC理論」	1 (149)
目次	2 (150)
巻頭言：イノベーションの場としてのオムロン サイニックエックス オムロン株式会社 技術・知財本部 副本部長 兼 技術・知財戦略室長 兼 オムロン サイニックエックス株式会社 代表取締役社長 諏訪 正樹	4 (152)
寄稿文：人とAIとの共生創発をもたらすHumanity2.0 慶應義塾大学・理工学部 兼 慶應義塾大学・共生知能創発社会研究センター 兼 電気通信大学・人工知能先端研究センター教授 (センター長) 兼 オムロン サイニックエックス 取締役 栗原 聡 (工学博士)	7 (155)
寄稿文：SINIC X-formation —パラダイム・シフトの先を拓く技術への期待— 株式会社ヒューマンルネッサンス研究所 代表取締役社長 中間 真一	9 (157)
本特集号の章立てについて	13 (161)

## 1. 人がやっていることを自動化する

### 人のようにしなやかなロボット～ソフトロボット

田中 一敏	16 (164)
-------	----------

### しなやかなロボットをうまく制御する～モデルベース学習

濱屋 政志	22 (170)
-------	----------

### Advancing Automated Assembly-Team O2AC at the World Robot Summit 2020

FELIX von drigalski, NAKASHIMA Chisato, KASAURA Kazumi and CRISTIAN beltran-hernandez	29 (177)
---	----------

## 2. 人から学ぶ 人のように学ぶ

### 機械学習を活用した経路計画技術

谷合 竜典, 米谷 竜	36 (184)
-------------	----------

### A Machine Learning System that Adaptively Aggregates Knowledge from Multiple Models

馬 家昕	46 (194)
------	----------

### 過去の学習経験を組み合わせて適応するロボット

田中 一敏	54 (202)
-------	----------

## 3. 人と協調する 人とともに働く

### 新しい現場で簡単に立ち上げる

米谷 竜	62 (210)
------	----------

### ヒトの教示やインタラクションを活用したロボット学習

濱屋 政志	69 (217)
-------	----------

### 視覚と言語の共有潜在表現獲得によるモノづくり手順理解

橋本 敦史	76 (224)
-------	----------

### 雑踏の中で動き回る自律走行ロボット～社会ジレンマを解決する強化学習の活用

西村 真衣	82 (230)
-------	----------

### 動的に変化する環境の中で自己位置を推定する自律走行ロボット

西村 真衣	89 (237)
-------	----------

## 4. 技術解説・サーベイ

### 3次元画像計測におけるステレオマッチングの基礎から最先端まで

谷合 竜典	98 (246)
-------	----------

### 連合学習入門

米谷 竜	113 (261)
------	-----------

### クロスモーダルな理解～サーベイ

牛久 祥孝	120 (268)
-------	-----------

オムロン サイニックエックスが発表した主な論文・講演など	129 (277)
------------------------------	-----------

SINIC Theory: Predicting the Future Through the Interrelationships of Science, Technology, and Society ..... 1 (149)

Table of Contents ..... 2 (150)

Introduction: OMRON SINIC X as a Place for Innovation  
 SUWA Masaki, Deputy Senior General Manager, Technology and Intellectual Property H.Q.,  
 OMRON Corporation, President, OMRON SINIC X Corporation ..... 4 (152)

Contribution: Symbiotic Emergence between Humans and AI in Humanity 2.0  
 KURIHARA Satoshi, Ph.D., Professor, Faculty of Science and Technology, Keio University,  
 Director, Center of Advanced Research for Human-AI Symbiosis Society, Keio University,  
 Specially Appointed Professor, Artificial Intelligence eXploration Research Center, The University of  
 Electro-Communications, Outside Director, OMRON SINIC X Corporation ..... 7 (155)

Contribution: SINIC “X-formation” Expectations for Future Technology beyond the “Paradigm shift”  
 NAKAMA Shinichi, President, Human Renaissance Institute Co., Ltd. .... 9 (157)

Chapters in this special issue ..... 13 (161)

**1. Automate What Humans do**

**Soft Robots: Soft Touch Like Humans**

TANAKA Kazutoshi ..... 16 (164)

**Learning Control for Soft Robots: Efficient Model-based Learning Approaches**

HAMAYA Masashi ..... 22 (170)

**Advancing Automated Assembly-Team O2AC at the World Robot Summit 2020**

FELIX von drigalski, NAKASHIMA Chisato, KASAURA Kazumi and CRISTIAN beltran-hernandez ..... 29 (177)

**2. Learn from Humans, Learn like Humans**

**Path Planning using Machine Learning**

TANIAI Tatsunori and YONETANI Ryo ..... 36 (184)

**A Machine Learning System that Adaptively Aggregates Knowledge from Multiple Models**

MA Jiaxin ..... 46 (194)

**Adaptation to a New Setup by Combining Skills**

TANAKA Kazutoshi ..... 54 (202)

**3. Cooperate with Humans, Work with Humans**

**Adapting Machine Learning Models to New Environments**

YONETANI Ryo ..... 62 (210)

**Robotic Learning from Human Demonstrations and Interactions**

HAMAYA Masashi ..... 69 (217)

**Cross-modal Representation Learning for Understanding Manufacturing Procedures**

HASHIMOTO Atsushi ..... 76 (224)

**Towards Interactive Crowd-aware Robot Navigation Inspired by Social Dilemmas**

NISHIMURA Mai ..... 82 (230)

**Mobile Robot Navigation in Densely Crowded Environment**

NISHIMURA Mai ..... 89 (237)

**4. Advanced Technology Descriptions and Survey**

**Binocular Stereo: From the Basics to the State of the Art**

TANIAI Tatsunori ..... 98 (246)

**Introduction to Federated Learning**

YONETANI Ryo ..... 113 (261)

**Cross-modal Understanding: A Survey**

USHIKU Yoshitaka ..... 120 (268)

Recent Publications ..... 129 (277)

# 巻頭言

## イノベーションの場としての オムロン サイニックエックス



- ・企業理念を、より現代的なスタイルで体現する場として創設
- ・そのために、研究のみならずそのためのマネジメントも刷新し新しいスタイルを築き上げることを目指した。

### 1. 近未来を作る会社 ～オムロン サイニックエックスに込めた思い～

オムロンでは、創業者・立石一真の時代からオムロングループに受け継がれた技術経営スタイル「SINIC理論」という土壌で育まれた近未来デザインからのバックキャストを実践しています。オムロン サイニックエックス (OMRON SINIC X 以下OSX) は、この最先端を担い、近未来の社会から必要とされる革新的技術を世の中に先駆けて創出し、社会実装を具現化するための技術アーキテクチャをデザインするミッションを背負い、2018年に東京・本郷の地に誕生しました。会社名に入っている「サイニック」は冒頭にも登場した“SINIC”すなわちオムロンの創業者・立石一真が提唱した未来予測理論そのものです。そこに未知なる技術“X”を掛け合わせ、技術革新にもとづく近未来デザイン作りを推進していくという思いを会社名に込めています。OSXでは、世の中の変化の兆しと、そこから生まれる社会的課題をいち早く捉え、解決のボトルネックとなる技術的課題との重なりから、本質的な研究上の「問い」、魅力的な技術創出の「お題」を作り出すことを目指しています。社会的課題に根ざした研究課題を共有して、アカデミアとビジネスが本気で英知を結集することで、社会を動かし人々のライフスタイルを革新する大きなインパクトが実現できるとの信念の下、日々の活動を進めています。

OSXはもともと「AIの研究拠点を新たに設立する」という構想で準備が進められていました。しかし、オムロン社内でのディスカッションを深めていく中で拠点構想が日夜ブラッシュアップされていき、最終的には「近未来をデザインし、そのために必要なテクノロジー群を構想し、中でもコアとなる革新技術を創出するための近未来デザイン研究拠点」というコンセプトへと発展していきました。しかしながら、オムロンのメンバーだけでこの非連続的であり複雑さを増す十数年先の技術進化を予測し近未来をデザインすることは極めて困難です。ゆえに、このOSXという「出島」をオープンイノベーションの場として、国内外から優秀な研究者や専門家などを招いてディスカッションしながら革新技術起点の近未来をデザインしたいという思いをもつようになりました。その結果として生まれたのがOSXです。

もともとオムロンは、創業者の描く近未来デザインをベースに、世界に先駆けて信号制御システムや自動改札機など数々のイノベーションを生み出してきました。創業者はこれを「ソーシャルニーズの創造」と呼

### Our Mission (社憲)

われわれの働きで われわれの生活を向上し よりよい社会をつくりましょう

### Our Values 私たちが大切にしている価値観

- ・**ソーシャルニーズの創造**  
私たちは、世に先駆けて新たな価値を創造し続けます。
- ・**絶えざるチャレンジ**  
私たちは、失敗を恐れず情熱をもって挑戦し続けます。
- ・**人間性の尊重**  
私たちは、誠実であることを誇りとし、人間の可能性を信じ続けます。

図1 オムロンの企業理念

んでおり、今でもオムロンの企業理念のなかに根付いています (図1)。

この文脈でOSXという組織を表現するならば、創業者・立石一真の頭の中に存在したであろう「近未来デザイン創出のメカニズム」を、オープンイノベーションと人財の多様性にもとづいて組織実装するという挑戦の証とも言えます。人財の多様性については、これまでに10か国以上の研究者がOSXに来て「企業理念 (図1) の実践は下から」をモットーに互いに切磋琢磨しながら共に活動してきました。今後もSINIC理論という未来予測論を新時代に即したものと見据えながら、革新技術を起点にしたソーシャルニーズを創造するユニークな研究拠点としての存在感を高めていきたいと考えています。

### 2. 本特集によせて

「近未来デザイン」といっても、単なる未来の夢ではありません。OSXでは具体的な未来の姿をクリアな社会実装のイメージにまで落とし込みます。現在はAIやロボティクス、センシングなどが着目している技術の中心ではありますが、SINIC理論でいうところの機械が人の可能性や創造性を広げる「人と機械の融和」を具現化していくためには、人と機械のコミュニケーションや、その間を流れる知識や情報などの多様な

データのハンドリング、機械の身体性など、さらに深掘りしていくべき技術領域はどんどん広がっていています。この度発刊されたOMRON TECHNICSでは、OSXが誕生して3年あまりで生み出されてきたこれら技術領域における「近未来デザインの種」の代表事例をご紹介します。その多くは、著名な国際学会などで発表してきた論文の和訳をベースに、それぞれの技術について解説を加えた構成になっています。

本特集で紹介した技術の数々は、近いうちに社会実装され世の中に登場してくるでしょう。OSXでは引き続きグローバルレベルでのスピード感に遅れることなく、技術革新と社会実装とを世に先駆けて高度に繋げていく活動に邁進していきます。

オムロン株式会社 技術・知財本部 副本部長 兼 技術・知財戦略室長  
兼 オムロン サイニックエックス株式会社 代表取締役社長

諏訪 正樹

# 人と AI との共生創発がもたらす Humanity2.0

慶應義塾大学・理工学部 兼 慶應義塾大学・共生知能創発社会研究センター 兼  
電気通信大学・人工知能先端研究センター教授（センター長）兼  
オムロン サイニックエックス株式会社 取締役

博士（工学） 栗原 聡



現在、国が進める IT 政策である Society5.0 は、現在の情報化社会（Society4.0 と定義される）をさらに推し進め、実世界のあらゆる情報をサイバー空間にビッグデータとして集約（デジタルツイン化）し、そのビッグデータに対する AI による分析や学習・予測を通して生成される高付加価値な情報を、実世界にフィードバックさせようというものである。現在においてもすでにサイバー空間の利用は DX が叫ばれる昨今ではあるが、現実にはコロナ禍もあり、なかなか導入が進まない現状となっている。結果的に、まだまだ実環境からの情報流入は質・量共に少なく、サイバー空間内においても情報の連携などにおいてまだまだ対応が遅れている。

5G の本格導入や深層学習を軸とする機械学習を基盤として、DX やスーパーシティといったキーワードが新聞雑誌に多く登場する状況を見ても、デジタル後進国と揶揄される日本がその汚名を返上するためにも、現在の動きは決して間違っていないと「は」思う。ただし、「は」としたのは理由がある。何か重要な点を見落としていると感じるのである。

「社会（Society）」は実に使い勝手のよい言葉である。人々の行動の総体として創発される「実体なき質感」に「社会」というラベルを付けたものにすぎない。「よりよい社会を作る！」、ということは当然ながら社会を構成する個人にとってよりよい状況を目指す、という意味だと思いたいが、昨今の世界情勢やサイバー空間で起きている諸問題や AI に対する懸念等を見るに、肝心の個人が置き去りになってしまっている。社会という総体として捉えることが、逆に個人をオブラートに包んで見えにくくしてしまうことになってしまっている。そもそも一人一人が基本単位であるにも関わらず、そのような危惧が実際に感じられることの現れとして「人間中心」という強調を取って

する状況も散見される。

情報過多・情報洪水の問題が指摘されてすでに 20 年は経過していると思うが、人（ホモサピエンス）としての認知能力は 8 万年変わらないまま、インターネットやコンピュータの高性能化を背景として、より多くのデータ、より何でも繋ぐことが追求され、SNS におけるデマ、フェイク、フィルターバブルなど、もはやサイバー空間に人が翻弄される状況が深刻化している。

このまま Society5.0 を目指すことが、表面的には利便性が向上するものの、これがもたらす負の面の影響は計り知れない。まだ手遅れ確定ではないと思いたいが、実世界とサイバー空間における「社会」を人々の行動が創発する現象として捉えた上での取り組みを開始する必要がある。

産業革命にてテクノロジーが急速発展を開始するまでの人類を Humanity0.0 とすれば、工業時代に入って人は Humanity1.0 にアップグレードしたと言えよう（工業社会は Society3.0 である）。狩猟時代、農耕時代を通して人類は道具を生み出し、道具による効率化により発展してきた。それは、Society4.0 に区分けされる情報社会となり、特にインターネットの発明により、道具の性能は飛躍的に向上し、我々の生活は劇的に変わった。しかし、テクノロジーは人の使う道具である、という立ち位置には変わりはない。そもそもテクノロジーとはそういうものだ、という割り切った考え方もある。しかし、その立ち位置は来る Society5.0 においても変わらない、と安易に決めつけてしまってもよいのであろうか。

現在の AI はほぼ機械学習のことを指しており、データから学び動作する。与えられるデータから学ぶことで事足りるのであれば、我々はもっとスマートに生きられている

はず。実環境は複雑系であり、学習したことも利用しつつ高い適応能力と汎用性を持ち、何より我々は受動的な道具ではなく、能動的に動作する自律性を有する。

深層学習が飛躍的に画像・言語処理を中心としてその性能を大きく向上させたことで、次世代型 AI の実現に向けた研究を進める土壌がおおかた整った。深層学習分野においても、現在が学習されたモデルによる即応的反応が主たる能力であり、これを System1 と呼び、今後は熟考できる System2 を目指す動きがある。どの道、相変わらず学習に頼ることから、System1 であっても適応性には大きな期待はできないのも、当然の流れであろう。深層学習（学習）のみを基盤とする限りは、System2 であっても道具の位置づけからは脱却できない。

では、Humanity1.0 はどのタイミングで 2.0 にアップグレードするのであろうか。タイミングとしては Society5.0 が実現される頃だと思うが、ただし、個人を主役とする捉え方でのコンセプトとして大きく見直す必要がある。

Society5.0 が実現される日常には、自律汎用型 AI が社会浸透しつつあり、人と AI との共生が実現され始めているだけでなく、人自体のマシンとの融合も加速していると想像される。人が AI と共生することで、新たな能力が創発されることから（これを共生創発と呼びたい）、それは人としてのアップグレードを意味する。すなわち Humanity2.0 へのアップグレードである。テクノロジーが道具的立ち位置である限りは、道具を使うレベルは人次第となってしまうが、人が自律汎用型 AI と共生する状況では、その壁を越えることが期待できる。実際、筆者が関わった TEZUKA2020 プロジェクト (<https://tezuka2020.kioxia.com>) では、人と AI との共生に向けた取り組みが行われ、共生創発への可能性を感じる事ができた。

ちなみに、著書「サピエンス全史」や「ホモ・デウス」で有名な歴史学者のユヴァル・ノア・ハラリが言うところの人間のアップグレードは、人の脳とサイバー空間との接続までを視野に入れていることから、これが実現された際は Humanity3.0 ということになるだろうか。我々はまずは Humanity2.0 の実現を目指すことになるが、2.0 の次は 3.0 が待っており、一連の Humanity のアップグレードに向けた取り組みを『Project HumanityX.0』と呼称したいと思う。

さらに、共生創発に関しては、多段階の創発が複雑に絡み合うことが想像される。人体を例にすれば、乱暴な言い回しであるが、細胞レベルが臓器レベルを創発し、臓器レベルが人体レベルを創発し、人体レベルがローカルコミュニティレベルを創発し、ローカルコミュニティレベルが社会レベルを創発し・・・、というダイナミクスを工学として活用するための方法論が必要となる。自然界に存在する多段階創発系においては階層と階層との間には壁が存在し、これがあることで系としての安定が維持されていると考えているが、インターネットの功罪は、階層間に縦串を通してしまったことなのではと推察している。共生創発の仕組みをデザインすることは、このような多段階創発現象の理解・構築・制御方法の確立を意味し、まさに「複雑系社会工学」と呼ぶべき分野の確立にも繋がるのだと思う。

いろいろ妄想を述べてきたが、最後に主張したいことは、上記の考え方は今更であるが SINIC 理論と本質的に同じである、ということである。オムロンサイニックエックスも本格稼働中であり、2021年4月1日に筆者は、慶應義塾大学に Humanity2.0 実現を目指すことを柱とする共生知能創発社会研究センターを設立させた (<https://sites.google.com/keio.jp/hass/>)。同じ野望を持つ者同士、徒党を組んで進んで行きたい。

# SINIC X-formation —パラダイム・シフトの 先を拓く技術への期待—

株式会社ヒューマンルネッサンス研究所 代表取締役社長

中間 真一



はるか彼方の学生時代、分厚い就職案内のページを繰る中で、「機械にできることは機械に任せ、人間はより創造的な分野での活動を楽しむべきである」というオムロンの企業哲学が目にとまった。当時、私は管理工学という学際分野を専攻し、作業疲労、習熟、工程設計、生産管理、R&D戦略、組織文化等、ものづくり現場の問題解決を研究テーマとする、広い意味でのIE（インダストリアル・エンジニアリング）研究室に所属していた。その恩師が、いつも私たち学生に投げかける挨拶は、「みんな、ハッピーしてる？」だった。そういう背景、文化、環境も影響してか、人間がより創造的な分野での活動を楽しむためのオートメーションという、オムロンの創業者・立石一真の鮮明なメッセージに惹かれた。

そのIEとは、狭く見ると、現場の作業や動作を観測して、ムリ・ムダ・ムラを徹底的に削ぎ落とし、生産効率を最大化しようという科学的管理法である。これは、使い方によって、生産現場で働く人々を、厳しく辛い作業環境に追い込んだ。学生時代の現場実習先の自動車会社のトイレの壁には「IE消えてなくなれ！」と殴り書きされていた。就職先の新人時代には、写真フィルム製造設備の向こう側から「なんで、あんな若造に、俺の動きにムダが多いと言われなきゃいけないんだ。やってられねえ！」というベテランの方の怒声が聞こえてきた。

もちろん、それはIEが目指すところではない。働く人も経営者も、そして使い手となる顧客も、共にハッピーになることが目指される。立石一真も、F.W.テイラーの科学的管理法を「能率学」<sup>1)</sup>として紹介した上野陽一先生に、いち早く師事し、熱心に研究して活用していたと聞く。動作研究の権威であったM.E.マンデル博士を60年代初頭にアメリカから招聘した際、日本での随行を恩師が担い、東京から関西に向かう車中で、隣の座席の初老の紳士か

ら話しかけられ、互いに素性がわかって、ますます話しが弾み、京都で下車して、そのまま鳴滝の工場まで連れていかれ、夜通し熱心に工程改善や工場経営の議論をしたという逸話を聞いたこともある。まさに「ハッピーな」工場、社会、世界をつくろうと、一生懸命に考えていたのだそうだ。より幸せな未来社会と暮らしをつくるという目的をぶらさず、技術と事業で実現する。それがオムロンの原点にある。

さらに、もう一つ惹きつけられたのがSINIC理論である。現場観察から人間観察、社会観察へとスコープを拡げ、事業づくり、組織改革など問題解決の範囲も拡がり、隠れた課題も洞察できることを感得した私は、文化人類学の世界に興味を持ち、その大御所であり探検家でもあった梅棹忠夫先生の思考と創造の世界、未来探検の世界に引き込まれた。ちょうど、60年代初頭に情報産業論<sup>2)</sup>を発表し、情報社会という未来を展望していた時期である。さらに、当時の気鋭の若手社会学者の加藤秀俊先生やSF作家の小松左京氏など、未来発想力豊かな面々が集まり、繰り広げられた議論は、1970年の大阪万博に結実し、人々に未来を見せて夢と希望を与えた。その中で、未来学なる新学問領域も立ち上がっていたが、当時の経営者としては珍しく「未来予測」の重要性を感じとり、梅棹忠夫先生を中心とした、言わば京都の「梁山泊」の面々とやりとりしていた中に、他でもない立石一真もいたのだ。

到達すべき未来社会のビジョンを「自然」と留めると共に、一気に原始社会まで遡って人類史全体を俯瞰し、人間社会の価値観の変遷と、科学・技術・社会の円環的な相互作用の発展過程から未来予測する理論、それがSINIC理論である。情報化社会以降の未来として、最適化社会、自律社会、自然社会を設定した背景には、京都の自由な学者の叡智や東洋思想的な観点も注ぎ込まれていた。さらに、そ

これらの智慧に、当時の中央研究所に設置したコンピュータを使ったシミュレーションも加え、2033年に至る未来予測の理論が、創業者・立石一真と若いスタッフらの情熱で、まとめあげられた<sup>3)</sup>。

SINIC理論は、不易の未来予測「理論」なのだ。だから、VUCAの時代の今こそ、SINIC理論を神棚に祀って拝むだけでなく、手元に取り戻し、未来への航路の羅針盤として、針路を定めるために使うべき時なのだ。

では、これから到来する「自律社会」とは、どんな社会であろうか。これを考えてきた結果、私は「自立」「連携」「創造」の3つの構成要件が重なる社会像にたどり着いた。そして、この社会は「モノ中心・個人中心」という近代工業社会の価値観の対極に位置し、「こころ中心・集団中心」の価値観に基づいた世界と予測されている。だから、この社会の到来を引き寄せる技術は「精神生体技術」、こころの技術なのである。また、「自律社会とは、ノー・コントロールの理想社会に近づく社会」だと、立石一真が言い残していることも興味深い。確かに、SINIC理論で予測される技術を見ると、最適化社会までの新技術には「制御」という文字が入るが、自律社会への革新を促す「精神生体技術」以降には「制御」の二文字が消える。

人の肉体機能を強化し、拡張して機械に代替するオートメーション、人の一定の判断機能と情報処理を代替するサイバネーション、一人ひとりの健康を支えるバイオネーション、これらのイノベーション展開予測は、N.ウィーナーが提唱したサイバネティクスの思想<sup>4)</sup>に基づくものであり、「制御」の思想である。

そのN.ウィーナーの著作には、興味深い指摘が見つかる。「人間の奴隷労働を機械に行わせたとしても、機械的奴隷をつくっているだけで、それは良いこととは言えない。人と機械が奴隷レベルで競い合っても、結局は人の価値を下落させるだけで、それは目的ではない。人間の価値を尊重する社会をつくることが目指すべきゴールだ」とか、「世界は決して機械仕掛けではない。生きた生命体なのだ」とも、数学的に説いている。これらの言及からは、N.ウィーナーが見通していた制御論の進化の先にも、自律、自然社会への道筋を重ねられることがわかる。

もちろん、いきなりノー・コントロールにして、理想社会ができるはずはない。そこには、自転車の補助輪のように、最初は無いと困る支えだが、自走できるようになると、次第に邪魔になり消えていくような、一人ひとりの個人の価値観や行動の変容を助ける技術が必要となるだろう。しかし、これはマインド・コントロールと紙一重の危

うさも孕んでいる。やはり、精神生体技術に、コントロールは禁物なのだ。生き方、暮らし、社会に融け込み、散りばめられたカム・テクノロジーとも言える。立石一真は意味深いコメントを残している。「自律社会の到来は、人間の真の変容なしには実現されない」、真の変容を遂げて、自律社会に向かうために必要な技術、それが現代社会の大きなソーシャルニーズのはずだ。

このように考えてくると、これからの技術開発には、これまでにない「倫理」のメカニズムが重要になる。自律社会の新たな規範を目指した「倫理」である。AIの深層学習、自律型のロボティクス、ライフログによる行動最適化など、その適用機会は広がる。そして、従来からのトロッコ問題に立ち止まったままのAI倫理ではなく、近年、急速に研究が進みつつある認識論の一分野である「徳認識論」<sup>5)</sup>のような知的な徳や、文化の地域性、デジタル・エシックスと呼べるような新しい人文社会的な研究フロンティアに挑戦していく必要性も高い。

さらに、自然社会とは何か。私はこの「自然」を、Naturalの訳語としての「しぜん」というよりは、仏教用語としての「じねん」に近いもの、すなわち、無為自然であり、ありのままという状態を意味するものだと考えている。地球上の生態系では、多様な生きものが、何者もコントロールしているわけではないのに、40億年という持続性の中で、時々の環境変化に適応して約870万種の生物が生きている。つまり、「制御」を超えた自律的な地球環境への「適応」であり、固定したシステムではなく、構造を融通無碍に変えて、環境に最適に対応する、人も自然も人工物も含めた地球生態系の「ホメオスタシス」への発展である。S.カウフマンやI.プリゴジンらの天才科学者たちが構想した「自己組織化」<sup>6)</sup>や「カオス理論」<sup>7)</sup>等が、SINIC理論の自然社会以降の次なる螺旋的未来予測の中心的概念となるはずだ。それは、機械仕掛けではなく、まさに生命体としての地球の未来だ。ハイパー原始社会と呼べるかもしれない。

そこに向かう技術は、人と人、人と自然、人と人工物、人工物と人工物の間で互いに働きかけ合って、折り合いをつける機能が重要になるだろう。このようにして、人間と技術と自然が、共によい塩梅で作用し合う世界とは、I.イリイチが提唱した、豊かでハッピーな自立共生社会「コンヴィヴィアルな社会」<sup>8)</sup>に通じる。私は、自然社会とはコンヴィヴィアル社会だと考えている。そして、そこにはAIの発展領域として、池上高志氏らのAlife(人工生命)の発想から生まれる技術の社会実装も現実となるだろう。

1970年、人類の進歩と調和をテーマとして開催された大阪万博の同年に発表されたのが、2033年の自然社会か

らバックキャストした SINIC 理論である。奇しくも、SINIC 理論では自律社会の始まる 2025 年、再び大阪・関西万博が「いのち輝く未来社会のデザイン」をテーマとして開催される。SINIC 理論も、2 周期目の未来を見据えた、新たな進化の準備の時を迎えている。

今こそ SINIC 理論を羅針盤として、渾沌としたパラダイム・シフトを超えた地平、新しい社会への変容を先駆ける時である。まさに、「人間がより創造的な分野での活動を楽しむ」、みんながハッピーになれる未来可能性を高める SINIC X-formation (サイニック・トランスフォーメーション) を、オムロン サイニックエックスの革新技術が拓く時だ。

そして、「SINIC 理論」を活用して、社会・技術・科学の未来を描くヒューマンルネッサンス研究所の未来社会論と相互円環的に高め合い、よりよい社会への進化を拓いていきたい。

- 1) F. W. テーラー／上野陽一訳『科学的管理法』技報堂, 1957 年
- 2) 梅棹忠夫「情報産業論」『中央公論』中央公論社, 1963 年 3 月
- 3) 立石一真ほか. 未来接近への SINIC 理論. OMRON TECHNICS, 1970. Vol. 10, No. 3, (通巻 34 号), P. 26-39.
- 4) ノーバート・ウィーナー／池原止戈夫他訳『サイバネティックス—動物と機械における制御と通信—』岩波書店, 1957 年.
- 5) 飯塚理恵, 「倫理的徳と認識的徳」, 『ワードマップ心の哲学—新時代の心の科学をめぐる哲学の問い』信原幸弘編, 新曜社, 2017 年
- 6) スチュアート・カウフマン／米沢富美子訳『自己組織化と進化の論理』日本経済新聞社, 1999 年
- 7) イリヤ. プリゴジン, イザベル・スタンジェール／伏見康治他訳『渾沌からの秩序』みすず書房, 1987 年
- 8) イヴァン・イリイチ／渡辺京二・渡辺梨佐訳『コンヴィヴィアリティのための道具』日本エディタースクール出版部, 1989 年



## 本特集号の章立てについて

本特集号は、自律走行ロボットや自動組み立てロボット、ファクトリーオートメーションにおける品質管理システム、ヘルスケアにおける医療システムのデータマネジメント機能など、いま挙げた事例やそこに留まることのない「多様な形で社会実装をイメージすることのできる」基盤技術に関するオムロンサイニックス研究者が執筆した論文を集めています。

様々なアプリケーションに適用できるポテンシャルを持つ技術であるが故に、各々の技術は一見すると雑多な話題の寄せ集めにも見えてしまうかもしれません。しかしながら、オムロン創業者 立石一真の言葉でもあり、オムロンの企業哲学ともいえる「機械にできることは機械に任せ、人間はより創造的な分野で活動を楽しむべきである」を3つの切り口で分解して得た研究成果であり、各切り口毎に第1章から第3章を構成しました。

加えて、AI技術やセンシング技術の中から3つの技術を取り上げ、その技術トレンドや技術展望について述べた、技術解説・サーベイ論文を第4章に構成しました。

人と機械の未来の関係性に思いを馳せつつ、本特集号の論文をご高覧いただけますと幸いです。

本誌構成

第1章：人がやっていることを自動化する

第2章：人から学ぶ 人のように学ぶ

第3章：人と協調する 人とともに働く

第4章：技術解説・サーベイ



機械にできることは機械に任せ、  
人間はより創造的な分野での活動を楽しむべきである

オムロン創業者 立石一真

創業者物語：<https://www.omron.com/jp/ja/about/corporate/history/founder/>



# 1

---

人がやっていることを  
自動化する

# 人のようにしなやかなロボット～ソフトロボット

田中 一敏

ロボットは製品生産ラインに組み込まれ、製品の大量生産に使われてきた。同一動作の反復しかできないロボットが、臨機応変に作業を行って人手の作業を代わられるようになれば、製品を利用する一人ひとりに合った製品を提供できるようになる。部品の相対位置を正確に合わせて部品同士をはめあわせる製品組立作業は、ロボットにとって困難な作業の代表である。臨機応変に製品組立作業を行うロボットのアーキテクチャを構築するため、身体を構成する部品の変形によって物体と柔らかく接触するソフトロボットに着目し、ソフトロボット研究のプラットフォームとして、市販の硬いロボットに柔軟さを加えられるロボット手首を開発した。開発した手首は、物体と接触する時にはバネの変形によって柔らかさを生む一方、物体を運ぶ時には振動して物体を落とさないように動かない状態を取れる。ベースに置かれたシリンダがワイヤを介して手首を駆動するため、手先の限られた空間を変位に活用できる。実験において、柔らかい手首が衝撃の緩和や位置誤差の補償に寄与することを確認した。筆者らの提案するロボットの手首は、身体の柔らかさを駆使して作業する新しいロボットの構成法の研究を可能にした。

## Soft Robots: Soft Touch Like Humans

TANAKA Kazutoshi

Robots have been integrated into product production lines and used for the mass production of products. If robots, which can only repeat the same motions, can perform a task flexibly and replace manual labor, it will be possible to provide products that suit each product user. Parts assembly tasks are a typical example of tasks that are difficult to be replaced by robots. In parts assembly tasks, the relative positions of parts must be precisely aligned to fit them together. To solve this problem, we focused on soft robots that can make soft contact with objects by deformation of parts in their body and developed a robotic wrist that can add softness to a commercially available hard robot. The developed wrist can take a fixed state to avoid vibrating and dropping an object when carrying this object. A cylinder placed on the base drives the wrist via a wire. Thus, the limited space at the end of the arm can be used for displacement. With the developed wrist, we can explore new architectural methods of robots that work by leveraging the softness of the body.

### 1. まえがき

本稿ではまず、ロボットが臨機応変に製品組立作業を行う必要性について解説する。続いて、臨機応変に製品組立作業を行うロボットの実現に貢献すると期待できる、ソフトロボットを用いた新しいロボットアーキテクチャを紹介する。最後に、ソフトロボット研究のプラットフォームとして開発した柔らかいロボット手首を紹介する。

### 2. 多様な製品組立作業を行うロボットの必要性

ロボットは、人間を幸せにする賢い機械として、人間の仕事を代替してきた。時代が進み、社会の求めるものが変われば、ロボットの果たすべき役割も変わる。大量生産・

大量消費の時代は経て、現在は、多様な価値観を持つ個人ひとりひとりが自分に適した製品を手にし、各自の苦しみを解消し、幸福を追求できるよう、ロボット技術を含む情報技術に対して、人間を補完、支援、能力拡張することへの期待が高まっている。ロボットが人手の作業を代替できれば、コストが下がり、個人に最適な製品を安価に提供できるようになる。これまで数十年に渡って生産に使われてきたロボットは、技術の発展によって、臨機応変に作業し、人手の作業を代替するようになったであろうか。以下では、過去の産業用ロボットの解説について総括する。

産業用ロボットは多品種少量生産にも使えるであろうか。大量消費社会において、同一の製品を大量に生産するために使われた産業用ロボットは、(1) マニュアル・マニピュレータ、(2) 固定シーケンス・ロボット、(3) 可変

Contact : TANAKA Kazutoshi kazutoshi.tanaka@sinicx.com

シーケンス・ロボット、(4) プレイバック・ロボット、(5) 数値制御ロボット、(6) 知能ロボットに分類され、後者ほど技術的に高度だとされている<sup>1)</sup>。これらのうち、(1) から (5) のロボットに共通する基本特性は、あらかじめ定められた通りの動きをそのまま繰り返すことにあり、この繰り返し型のロボットを中野は第一世代の産業用ロボットと呼んだ<sup>2)</sup>。中野は、感覚器の情報に基づいて行動を修正できるロボットを第二世代、学習を通じて事前にプログラミングされていない行動を獲得するロボットを第三世代、学習手法が成熟した後に現れる新しい未知概念の、より高度な基本的手法が特徴づけるロボットを第四世代のロボットであると予見した。1982年の時点で、使用されている産業用ロボットのほとんどは繰り返し型の第一世代のものであった。米本が指摘したように、その後、知能ロボットの割合は増加した<sup>3)</sup>。水川と小山は、産業用ロボットへの教示は Manual Data Input やティーチングプレイバックがほとんどであると指摘した<sup>4)</sup>。米本が指摘したように、1980年を境に大量生産の時代から、多品種少量生産の時代に移ったが<sup>5)</sup>、大量生産にロボットは使われ続けた。

産業用ロボットを多品種少量生産に使うことを阻む要因は何であろうか。野田は産業用ロボットを用いた生産システム全体の主要コストは、エンジニアリングコストであると指摘した<sup>6)</sup>。すなわち、典型的な組立作業系のロボットシステムの見積り内訳のうち、全体の約半分以上が設計、据付、調整のためのエンジニアリングコストで占められており、当該案件ごとのカスタマイズに費やされていると報告した。野田らは次世代ロボット知能化技術開発プロジェクトにおいて、機種切り替えが迅速で、長時間連続創業可能な、生産現場へ導入できるシステムを開発した。この経験を踏まえて、産業用ロボットの設計の中心には、高速高精度にロボット手首のフランジ面が目標位置姿勢を再現する機能に特化して最適化されており、さらなる発展のためには、ロボットを利用する目的を今一度再考し、ロボットの成り立ちを材料と制御からゼロ・ベースで組み立て直す、新たなイノベーションにチャレンジすべきだと述べている<sup>6)</sup>。

近年の急速な技術発展によって、産業用ロボットを多品種少量生産に使えるようになったであろうか。堂前が報告したように、機械学習を用いた認識によるばら積みピッキングなど、発展した基礎技術が知能ロボットの作業を可能にした<sup>7)</sup>。しかし、原田と万が指摘するように、現在でも、製品組立作業は、人手での単純作業を要求している<sup>8)</sup>。原田と万は、産業用ロボットに組み立て作業を行わせるのが困難である原因として、人間に簡単にできることはロボットにも簡単にできると考えがちであること、操作物体の特性を理解する必要があることを指摘した。

以上のように、現在でもロボットは同一動作の反復しかできず、多品種少量生産に使う経済性がほとんどない。も

し今後ロボットが、臨機応変に作業を行って人手の作業を代わられるようになれば、製品を利用する一人ひとりに合った製品を提供できるようになる。野田が述べたように、製品組立作業などの生産を臨機応変に行うロボットを実現するには、全く新しいロボットアーキテクチャが必要である。そのようなロボットアーキテクチャは、どうすれば作れるであろうか。

### 3. 製品組立作業におけるソフトロボットの必要性

#### 3.1 柔らかく物体に触れるソフトロボット

従来の電子計算機による演算と正確なアクチュエーションを基礎とするロボットアーキテクチャとは異なる先進的なアーキテクチャとして、形態による演算を用いたアーキテクチャが挙げられる。形態による演算では、身体と物体の相互作用を電子計算機による演算の代わりに用いる。形態による演算として、受動歩行器が挙げられる。受動歩行器は、計算機による計算もアクチュエータもなしに、下り坂を安定して歩行できる<sup>9)</sup>。

身体の柔らかさを生かしてタスクを遂行するロボットとして、近年、ソフトロボットが注目を集めている<sup>10)</sup>。ソフトロボットとして、物体の形に変形し、形を固定することで物体を把持する、ジャミンググリッパが開発されている<sup>11)</sup>。シリコンゴムを材料にしており、空気の流入によって曲がる柔軟なロボット指が開発されている<sup>12)</sup>。

ソフトロボットの一つとして、空気圧筋骨格ロボットが挙げられる。空気圧筋骨格ロボットは、ヒトを始めとする柔らかさを備えた筋骨格系を規範として設計されている。新山らは跳躍ロボットの Mowgli を開発した<sup>13)</sup>。このロボットは、二関節筋の働きによる複数関節の機械的な協調によって、外乱に対して安定した跳躍を行える。池本らは、空気圧人工筋で駆動される柔軟なロボットアームによって、環境に拘束されるドア開け動作を簡単な制御で行えることを示した<sup>14)</sup>。同様に、製品組立作業においても、製品同士の相対位置を合わせるのに、柔軟な身体構成要素の変形が寄与すると期待できる。

#### 3.2 機械的柔軟性を活用したロボット作業

製品組立作業はロボットに代替させるのが困難な作業の代表である。製品組立作業においては部品同士をはめあわせるため、部品の相対位置を正確に合わせる必要がある。

製品組立において、受動的可動性をもつ、柔らかい手首が有用であることが知られている。その一例として Remote center compliance (RCC) と呼ばれる機構が Whitney らにより提案された<sup>15)</sup>。RCC では、手首内のバネの変形によって、軸と穴の位置ズレを補償する。同様に、バネや、受動的可動部を有する手首機構は数多く提案されている<sup>16)</sup>。

### 3.3 ソフトロボットによる製品組立作業

従来のロボットアーキテクチャは、手先の位置と姿勢を高い精度で制御できることを前提とする。さらに、人間が所与の作業について考察し、治具やロボットなどを決め、経由姿勢と姿勢切り替えの規則をプラグラミングし、変数を調整する。対して、ソフトロボットは機械的な柔軟性が位置の不確かさと応答の遅れを生むため、正確に位置と姿勢を制御できない。すなわち、既存のアーキテクチャの前提を満たさない。さらに、筆者らが目指す、臨機応変に作業するロボットは、作業内容が代わるたびに動作をプログラムするのではなく、専門知識なしに動作を自動で生成し、必要に応じて自動で調整される必要がある。しかし、そのような技術は未だ実現には至っていない。機械的な柔軟性を利用すれば、簡単な制御で柔らかい接触を実現し、結果的に物体同士の相対位置を高い精度で合わせられる。しかし、その動作は自動で生成され、かつ必要に応じて自動で調整される必要がある。さらに、ソフトロボットは柔らかく、材料が複雑に変形する。このため、従来の硬いロボットと比べて、幾何的、機構的な自由度が高い。加えて、全身の状態を十全に取得することが困難である。かつ、状態と状態遷移のモデル化も困難である。このモデル化と計測の困難さも制御を難しくする要因のひとつである。

ソフトロボットを用いた新しいアーキテクチャはどのようなものであろうか。まず、機械学習の技術を用いて、作業を自律的に獲得することを目指す。作業に応じて、複雑なモデルを作り、制御器を設計すると、手間がかかり、臨機応変に作業を行えない。ゆえに、多点分布接触センサ、力覚センサ、近接覚センサなど、多数、多種のセンサを搭載して全身の状態を十全に取得する。柔らかい身体は、積極的な物体への接触や衝突を可能にする。結果として、より多くの情報の取得を可能にするとともに、学習の過程における不完全な行動で破損する危険を減らせる。以上のように、柔軟な機構、多種、多数のセンサ、機械学習の3要素は互いの弱点を補い、長所を引き出し合う。これら3要素を基礎として、臨機応変に作業を行うロボットの新しいアーキテクチャを構築する。アーキテクチャを実現するためには、これら3要素からなる構成を前提として、それぞれの構成要素を開発する必要がある。

## 4. 可変剛性手首の提案

### 4.1 作業のための柔軟なロボット手首の要求機能

臨機応変に作業を行うロボットの柔軟手首には、5つの機能が求められる。

筆者らは、与えられた作業を臨機応変に行うロボットの実現を目指す。そのようなロボットは、製品組立において、事前に想定されない新しいやり方で部品を組み付ける

ことが要求される可能性がある。これを実現するための要求機能としては、以下の5つが挙げられる。

#### (1) 6自由度大変位

未知の変位方向を必要とする可能性があるため、6自由度が必要であり、可能な限り大きく変位するのが望ましい。この多自由度と大変位は、始めから作業とロボットの運動が決まっている従来のアーキテクチャではほとんど必要とされない、新しいロボットアーキテクチャのための機構的な特徴である。

#### (2) 小型性

未知の作業として干渉の多い場所や対象への作業を行う可能性を考慮した時、手首は部品同士、部品とロボットの物質的な干渉を避ける必要があるため、できるだけ小型であるのが望ましい。

#### (3) 後付け機能

どのようなロボットアームにも柔らかい手首を後付けできるようにする。すなわち、手首の取り付けによって既存の硬いロボットをソフトロボット化できる。

#### (4) 軽量性

手首が重い場合、可搬重量の大半を手首の重量に費やすことになるため、軽い必要がある。

#### (5) 切替機能

手首が柔軟なままでは、物体を運ぶ時に振動して物体を落とすおそれがあるため、硬い状態と柔らかい状態の両方を取る必要がある。すなわち、固定と解放の切り替え機構が必要である。

### 4.2 ケーブル駆動型6自由度小型可変剛性手首

ここでは筆者らが開発した、前節で述べた機能を備える手首について述べる(図1)。

開発した手首は、上部と下部の接続部品からなる。上部の部品はアームの先端に固定され、下部の部品はグリッパに固定される。上部と下部の部品は3つのバネで接続されている。グリッパや把持した物体が環境中の物体に触れるとバネが変形するため、柔らかく接触できる。

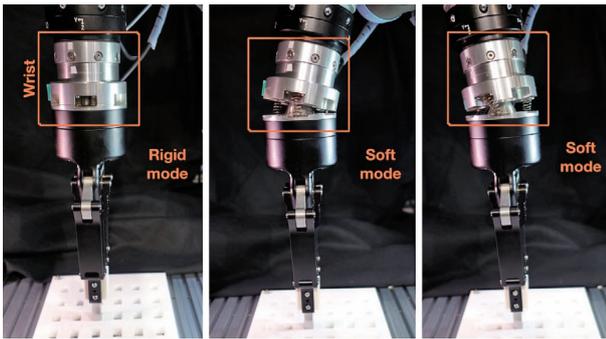


図1 開発した可変剛性手首

開発した手首は、ワイヤケーブルでアームのベースから駆動している（図2）。ワイヤケーブルにも弾性があるが、さらに弾性を与え、簡単な制御でワイヤの牽引を行えるよう、空気圧シリンダで駆動する。ワイヤシリンダは5ポートバルブ1個で駆動する。ワイヤで牽引した時、バネが縮んだ状態で上部と下部の部品、ピンと穴がはめ合せて固定される（図1左）。

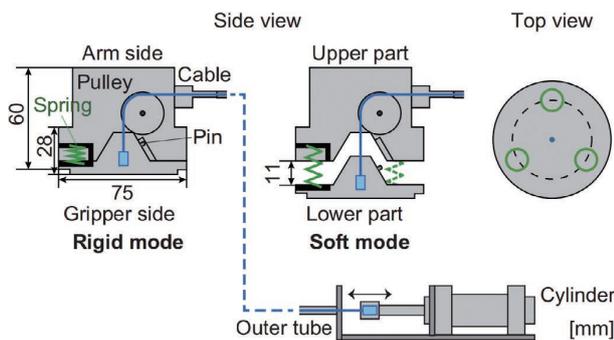


図2 手首の構成

現在、手首にセンサは入っていない。このため、本稿で取り上げる実験以外の、手首を用いたロボット実験では、モーションキャプチャを用いて位置と姿勢を計測した。ワイヤの張力は張力センサで計測でき、シリンダの内部圧力は圧力センサで計測できる。6つのヒンジ関節で上部と下部を接続すれば、下部の部品の位置姿勢を計測できるが、容積と重量が増加してしまう。このため、手の姿勢を計測する場合には、カメラや近接覚センサによる計測が適しているであろう。

### 4.3 評価実験<sup>17)</sup>

本節では開発したロボット手首の評価実験とその結果について述べる。

手首が柔らかい場合、環境に固定された物体に接触する時、接触を検知した後、停止が遅れても、過剰な力が生じにくい。手首の設計と過剰な力の生起を防ぐ機能との関係性を調べるため、接触して停止する実験を行った。ロボットに速度を設定してテーブルに向けて移動させ、接触させて

停止した。手首の上部と下部の間の距離を、1, 6, 11 mm と変え、バネを取り替えて弾性を変え、移動する速度を変え、接触する傾きを変え、緊急保護停止の生じる確率を計測した（図3）。手首の柔軟性の効果を確認するため、手首を固定した「リジッドモード」も試し、比較した。

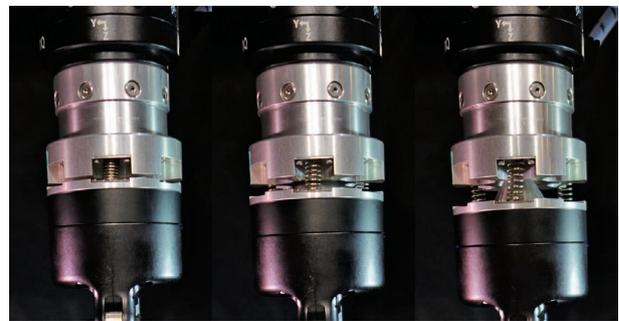


図3 上下部品間距離 1, 6, 11 mm の手首

高速な位置での成功率が高まった（図4）。1-6 mm 程度の上部和下部の距離、硬いバネを選択することで、衝撃吸収機能を高められると分かった。

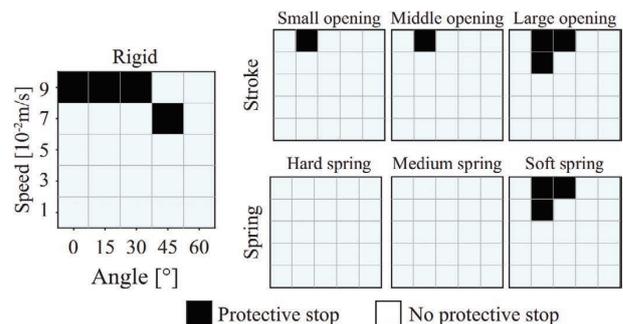


図4 緊急停止の有無

速度と力の関係について述べる。速度が大きくなるほど、力は大きくなった（図5）。リジッドモードの方が力は大きかった。これは、手首の柔軟性が過剰な力を避ける機能があることを示す。上下部品の間隔や弾性を比較すると、硬いバネ、小さな開きの条件において力は小さかった。これは上で述べた、緊急停止の有無から見た時の、適切な設計と共通する。

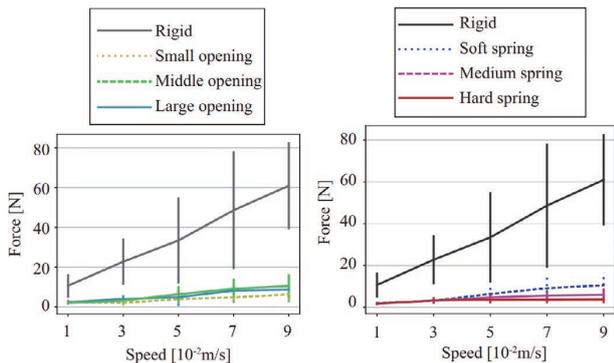


図5 速度と力の関係

手首の変位が部品間の位置のズレを補償する。この効果を調べるため、ペグインホールタスクの成功率を計測した。実験では、柔らかい手首が位置決め誤差の影響を受けたときに、ペグインホールタスクのパフォーマンスに及ぼす影響を評価した。ロボットには、ナイロン製の3Dプリントされた、テーブルに固定された長方形の穴に、辺の長さが10mm×15mm、長さが80mmのステンレス製の角型ピンを挿入した(図6)。このときのクリアランスは、0.1mmに設定した。

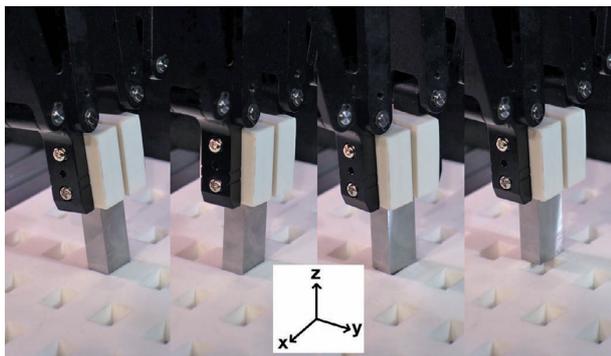


図6 実験の様子と実験における座標系

グリップの軸と穴の中心との間に異なるオフセットを与えることで、計測ノイズを模擬した。XとYのオフセットは0, 1, 3, 5, 7 mm、角度のオフセットは $\theta = 0, 1, 3, 5, 7^\circ$ であった。角度オフセットは個別に評価し、X, Y, Zの各軸に適用した。

グリップを穴の前に置き、オフセットを加えた後、挿入試行を実施した。柔軟なモードでは、 $r = 3\text{ mm}$ の螺旋状の動作を行った。位置を固定したリジッドモードでは、まず穴を見つけるために螺旋状の動きを行い、その後、ロボットの先端が穴の近くにあると仮定し、 $315\ \mu\text{m}$ の公差を持つカスタムルーチンを使用してペグを挿入した。このルーチンでは、力を加え、ペグが動かなくなったときにX軸またはY軸で1 mm前後に移動する。このように、各モードが比較的同等の速度で動作し、一方の方法が過度に不利に

ならないような挿入方法を使用するように注意した。挿入タスクで可能な限り最高のパフォーマンスを実現するために、ほぼ同じ時間をかけてチューニングを行った。

タスク終了後、挿入が成功したかどうか、ペグが1cm以上深く挿入されたかどうか、実行中に緊急停止が発生したかどうかを確認した。ロボットは各条件で5回タスクを実行し、成功、失敗、緊急停止の回数を記録した。

挿入の成功率は、位置ずれが大きくなるにつれて小さくなった(図7)。角度のずれが大きくなるにつれ、物理的な柔らかさによる補正効果が特に顕著に現れている。X/Y, Zの角度がそれぞれ3度と5度ずれた時、ソフトリストなしで、ロボットはペグの挿入に失敗した。

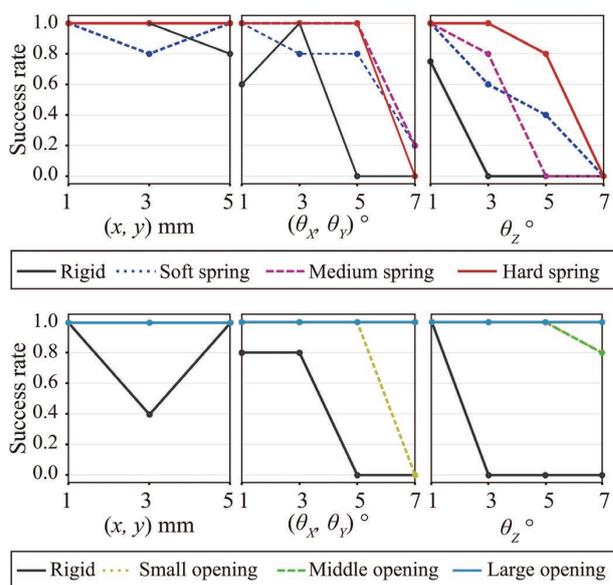


図7 緊急停止の有無

純粹に並進方向の誤差であれば、手首がなくても成功率は高い。これは、穴を螺旋状に探索することで、よく成功したためであると考えられる。しかし、円形のペグとベアリングを使った予備実験では、穴の周囲の隆起によって力信号のノイズが生じるため、スパイラルサーチの成功率は低くなった。リジッドモードでの失敗の主な原因は、スパイラルサーチ動作中の緊急停止や、力センサが検知していないのにグリップが穴に入ってしまった場合の誤認であった。

手首の開き幅が大きくなると、手首の可動域が大きくなり、成功率が向上することがわかったが、一方で、開き幅が最小の場合でも成功率が向上することがわかった。これは、手首の開き幅を制御するシリンダが反力で動くため、大きな力がかかると、手首の開き幅が小さくても、大きく変位できるためだと考えられる。

## 5. むすび

本稿では、産業用ロボットによる組み立て作業の現状について解説するとともに、組立作業に筆者らの提案するソフトロボットの適用事例について紹介した。

筆者らの提案する柔らかい手首の開発によって、機械学習を用いたロボットによる製品組立作業の研究が可能になった。それら学習に関する研究については、別稿に譲る。現在では、身体の柔軟さを活かして作業を行うロボットアーム全体の設計へと、設計理論の拡張を進めている。

## 参考文献

- 1) JIS B 0134:1979. ロボット及びロボティックデバイス—用語.
- 2) 中野栄二. 産業用ロボットの将来. 日本機械学会誌. 1982, Vol.85, No.766, p.1078-1082.
- 3) 米本完二. 産業用ロボットの今後の技術動向. 日本ロボット学会誌. 1993, Vol.11, No.1, p.7-13.
- 4) 水川真, 小山俊彦. 作業の教示とプログラミング 産業用ロボットの教示方法の現状と展望. 日本ロボット学会誌. 1999, Vol.17, No.2, p.180-185.
- 5) 米本完二. 産業用ロボットの現状と展望. 農業機械学会誌. 1985, Vol.47, No.2, p.219-224.
- 6) 野田哲男. 生産分野に向けた作業知能：産業用ロボットを進化させる知能化技術. 日本ロボット学会誌. 2013, Vol.31, No.1, p.10-13.
- 7) 堂前幸康. 産業用ロボットマニピュレーション：パタン認識・機械学習編. 計測と制御. 2017, Vol.56, No.10, p.776-781.
- 8) 原田研介, 万偉偉. 産業用ロボットの次世代技術について. 日本ロボット学会誌. 2019, Vol.37, No.10, p.915-918.
- 9) Collins, S. H.; Wisse, M.; Ruina, A. A three-dimensional passive-dynamic walking robot with two legs and knees. *International Journal of Robotics Research*. 2001, Vol.20, No.7, p.607-615.
- 10) Kim, S.; Laschi, C.; Trimmer, B. Soft robotics: a bioinspired evolution in robotics. *Trends in Biotechnology*. 2013, Vol.31, No.5, p.287-294.
- 11) Brown, E.; Rodenberg, N.; Amend, J.; Mozeika, A.; Steltz, E.; Zakin, M. R.; Lipson, H.; Jaeger, H. M. Universal robotic gripper based on the jamming of granular material. *Proceedings of the National Academy of Sciences*. 2010, Vol.107, No.44, p.18809-18814.
- 12) Thuruthel, T. G.; Shih, B.; Laschi, C.; Tolley, M. T. Soft robot perception using embedded soft sensors and recurrent neural networks. *Science Robotics*. 2019, Vol.4, No.26.
- 13) Niiyama, R.; Nagakubo, A.; Kuniyoshi, Y. "Mowgli: A bipedal jumping and landing robot with an artificial musculoskeletal system". *IEEE International Conference on Robotics and Automation*. 2007, p.2546-2551.
- 14) Ikemoto, S.; Nishigori, Y.; Hosoda, K. Adaptive motion of a musculoskeletal robot arm utilizing physical constraint. *IEEE Transactions on Robotics*. 2009, Vol.56, No.10, p.3975-3984.
- 15) Whitney, D. E.; Rourke, J. M. *Mechanical Behavior and Design*

Equations for Elastomer Shear Pad Remote Center Compliances. *ASME. J. Dyn. Sys., Meas., Control*. 1986, Vol.108, No.3, p.223-232.

- 16) Nishimura, T.; Suzuki, Y.; Tsuji, T.; Watanabe, T. "Peg-in-hole under state uncertainties via a passive wrist joint with push-activate-rotation function". *IEEE-RAS International Conference on Humanoid Robotics*. 2017, p.67-74.
- 17) Tanaka, K.; Drigalski, F.; Hamaya, M.; Lee, R.; Nakashima, C.; Shibata, Y.; Ijiri, Y. "A Compact, Cable-driven, Activatable Soft Wrist with Six Degrees of Freedom for Assembly Tasks". *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2020, p.8752-8757.

## 執筆者紹介



田中 一敏 TANAKA Kazutoshi  
オムロン サイニクエクス株式会社  
リサーチアドミニストレイティブディビジョン  
専門：知能ロボット工学  
所属学会：日本ロボット学会  
博士（情報理工学）

本文に掲載の商品の名称は、各社が商標としている場合があります。

# しなやかなロボットをうまく制御する ～モデルベース学習

濱屋 政志

我々は、柔軟要素を持つしなやかなロボットを使用して、より容易に実現できる組立自動化を目指している。柔らかさによって対象物体との馴染みが生じ、部品の破損やロボットの緊急停止が起きにくくなり、その結果、本来必要であったセンシングや制御計画に対する労力を大幅に削減できると期待される。一方で、柔らかさによってダイナミクスが複雑となるため、学習技術の活用が必要不可欠となるが、生産現場の早い立ち上げのためなるべく少ない試行回数で学習したい。そこで我々は、モデルベース学習に基づく効率的なロボット学習制御手法を提案した。本稿では、我々の最新の研究成果である、組立作業の区分化と柔らかさの活用による学習効率化手法と、柔らかいロボットに应用が期待できる Sim-to-Real モデルベース転移学習手法について解説する。

## Learning Control for Soft Robots: Efficient Model-based Learning Approaches

HAMAYA Masashi

We develop an automatic assembly system using a soft robot. Thanks to the physical softness, the robot can interact with the environment safely, resulting in a significantly easier setup for the robotic system. Meanwhile, as the dynamics model becomes complex due to the softness, we need to use learning-based approaches. However, a large amount of taring data might not be preferable for real robot applications. To deal with this problem, we developed model-based learning approaches, which are efficient and suitable for real robot applications. This article introduces our recent studies using model-based learning approaches: 1) lower-dimensional dynamics models by leveraging subtask segmentation and physical softness and 2) model-based Sim-to-Real transfer learning for multiple environments.

### 1. まえがき

ロボットによる組立自動化は、産業的価値が極めて高く、これまで数多くの研究開発が行われてきた<sup>1)</sup>。組立の自動化研究が今日まで広く行われてきた理由として作業の難しさが考えられる。一般的に、組立は部品同士の接触を伴い、部品同士の公差も小さい。もしセンシングの不確実性などから制御の誤差が少しでも生じた場合、部品同士に過剰な力が生じ破損などの致命的な失敗につながる可能性がある。

現在の産業用ロボットの多くは位置制御により溶接や塗装を行うために使用されている。このような位置制御に基づくロボットを組立に適用する際には、可能な限り誤差や不確実性を排除するため、入念なセンシング設定と動作計画が必要となる。一方で、力制御を用いた手法も数多くの

研究開発がされてきており、接触力を制御することで安全な組立作業を実現することができる。しかし、一般的に高精度な力・トルクセンサや、高制御周期のコントローラなど、高性能なハードウェアが必要となる上に、作業に応じて適切なパラメータ設定を行なければならない<sup>2)</sup>。総じて、ロボットの組立自動化は、莫大な労力が必要となるため、依然として工場制手工業に依存している<sup>3)</sup>。

この問題に対し、我々は物理的な柔らかさをロボットに持たせることに着目する。柔軟要素が対象部品になじむことで、誤差が生じても致命的な失敗を引き起こす可能性が軽減される。さらに、力・トルクセンサや、高制御周期のコントローラを導入することなく組立作業が実現でき、結果ロボットの立ち上げにかかる労力を大幅に削減できると期待される。そこで我々は、柔軟要素を持つ手首を開発した<sup>4)</sup>。この手首は従来の柔軟手首に比べてより軽量で大きな変位量を持つため組立に適しており、様々な市販の汎用

Contact : HAMAYA Masashi mailto:masashi.hamaya@sinicx.com

ロボットに搭載することができる。

しかし、物理的な柔らかさと接触を含むダイナミクスは非線形性を持ちモデル化が極めて難しいことが懸念される。複雑なダイナミクスに対処するために、機械学習のアプローチを適用することが考慮されるが、訓練データ取得時における部品の摩耗や、学習環境の初期化などの労力も無視できないため、なるべく少ない実環境とのインタラクションで組立作業動作を学習できるようにしたい。

この問題に対し、我々はモデルベースによる学習手法を利用することで活用することで非線形性を有する複雑なダイナミクスを記述することを試みる。モデルベースの手法は、一般的にモデルフリーの手法と比較してサンプル効率がよいことが知られており、様々なロボット制御に応用されている<sup>5)</sup>。

我々は、このモデルベース学習に基づいて、さらに学習の効率化を試みており、その最新の研究成果について本稿で紹介する。第2節では、モデルベース学習における背景を説明する。第3節では、組立作業を区分化し、各小区分におけるモデルを単純化することで学習を効率化させる手法<sup>6)</sup>を紹介する。第4節では、柔らかいロボットや組立作業にも十分応用が期待できる、シミュレーションでダイナミクスモデルを事前学習し、実環境に転移させることで、実環境でのデータ収集の労力を削減する Sim-to-Real 転移学習手法<sup>7)</sup>について解説する。第5節では、まとめと今後の課題について紹介する。

## 2. モデルベース学習

本節では、モデルベース学習における問題設定を簡単に説明する。一般的には、対象のロボットや操作される物体のダイナミクスは、現在の時間の状態と行動から次時間の状態に遷移するという、マルコフ決定過程 (Markov decision process: MDP) によって以下のようにモデル化される。

$$s_{t+1} = f(s_t, a_t) + \omega \quad (1)$$

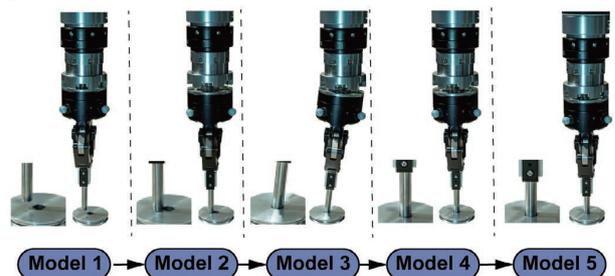
ここで、 $s$  は状態 (ロボットまたは対象物体の手先位置、関節角、速度など)、 $a$  はロボットの行動 (ロボットの速度、またはトルク指令など)、 $\omega$  はノイズである。モデルベース学習は、このモデルに基づき、累積期待コスト  $\sum_{t=0}^T \mathbb{E}[c(s_t, a_t)]$  を最小にするロボットの行動を決定することを目的とする。最適行動はモデルを使用して将来の状態を予測しながら求められる。

対象の環境が、極めて単純な場合は事前にモデルを設計することができるが、実世界におけるロボットの環境では、厳密な接触や摩擦などのモデルを解析的に求めることは極めて困難であり、柔軟要素が加わった場合は、その非線形性によりダイナミクスがさらに複雑になるのでより問題が難しくなる。このため、ダイナミクスモデルを機械学習に基づく手法で学習し、ロボット制御に応用するアプ

ローチが存在する<sup>5)8)9)</sup>。これらの手法は、モデルをガウス過程やニューラルネットワークで学習し、複雑なダイナミクスを表現しつつ、効率的にロボットの制御器を学習することを示した。

しかし、ダイナミクスの自由度が大きくなる、あるいはモデルがより複雑になるなどの場合、モデル学習に必要なデータ数も増加する問題点が生じ、実ロボットのへの応用が困難になる。この問題を解決する手段の一例として、ロボットが作業を遂行可能な範囲で、モデルをできる限り単純に表現するものや、過去に得られたモデルを新しい環境に適用する手段が考えられる。そこで、我々は、組立作業の環境を区分化し、そこに柔らかさによる物体のなじみを考慮することで、ダイナミクスモデルを低次元で表現する手法を開発し、各小区分をモデルベース強化学習で学習する手法を提案した (図 1a)。また、シミュレーション上で大量のデータを収集し、シミュレーション上でモデルを自己教師あり学習で学習し、多様な実世界の環境に追加の訓練なしに適応できる Sim-to-Real 転移学習手法を提案した (図 1b)。次節以降で、これらについて解説する。

(a) Lower dimensional dynamics using environment and softness



(b) Model-based multiple environment sim-to-real transfer

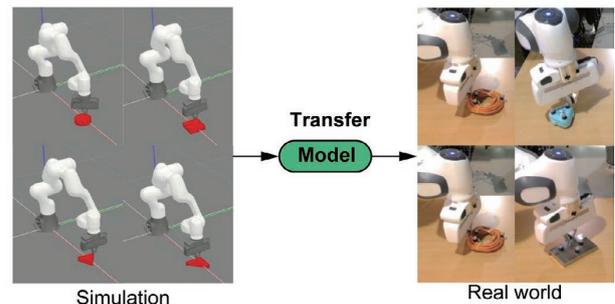


図 1 本稿で紹介する研究事例：(a) 環境拘束と物理的柔らかさを活用した低次元ダイナミクス表現<sup>6)</sup>と (b) 多様な環境に即座に適応可能なモデルベース Sim-to-real 転移学習<sup>7)</sup>。

## 3. 環境と柔らかさを活用した低次元ダイナミクス表現による組立学習の効率化

本節では、組立作業の中で代表的なものの1つであるベグ挿入作業に着目し、作業区分化と柔らかさを活用することで、よりシンプルなモデル表現ができ、その結果として学習が効率的化できることを紹介する。

### 3.1 低次元ダイナミクス表現

ペグ挿入作業では、たとえばペグが地面と接触する、穴の縁や底に触れるなど、挿入完了に至るまでの環境の物理的な相互作用から、複数の Manipulation primitives (MP) という小課題に区分化できることが示されている<sup>10)</sup>。この区分化によって、各区分における状態・行動空間を低次元で示すことができる。また、柔らかいロボットを活用することで、状態・行動空間をより低次元で示せることが知られている。Hangらは、柔軟なグリッパを使用して、薄い物体を押し付けてスライドさせる制御手法を提案している<sup>11)</sup>。柔らかさのおかげで、物体と地面は安定した接触を保障することができ、押し付け方向に対する制御を陽に行う必要がなく、平面上のみを考慮した動作計画が可能となる。また、柔らかさによる物体のなじみのおかげで、低次元の制御入力で、多様な物体の把持が可能になることも報告されている<sup>12)</sup>。そこで、本研究では、MPと柔らかさを活用することで、より状態・行動空間を低次元で表現できるのではないかと考えた。

図2に示すように、MPによってペグ挿入作業は5種類の小課題に区分化できる。単純のため、矢状面の空間を考慮する。ロボットの状態は、y・z方向におけるグリッパの手先の位置とα方向の回転角及びそれぞれの速度であり、行動はアーム先端の位置もしくは速度指令とする。

- n1 Approach: ロボットが、穴表面近くに接近する。この際の状態・行動空間は  $\mathbf{s}^{(\text{approach})} = [y, z]^T$ ,  $\mathbf{a}^{(\text{approach})} = [y_d, z_d]^T$  となる。Approachでは、事前設計しておいた位置制御器を使用する。
- n2 Contact: ロボットが地面と接触する。この際の状態・行動空間は  $\mathbf{s}^{(\text{contact})} = z$ ,  $\mathbf{a}^{(\text{contact})} = z_d$  となる。柔軟要素が衝撃を緩和するので、この小課題においても事前設計した位置制御器を使用することができる。
- n3 Fit: ロボットが把持したペグをスライドさせ、穴の縁に当てる。この際の状態行動空間は、 $\mathbf{s}^{(\text{Fit})} = [y, \dot{y}]^T$ ,  $\mathbf{a}^{(\text{Fit})} = \dot{y}_d$  で表現される。柔らかさのおかげで、地面

とペグが安定な接触が保障されているので、z方向の状態・行動空間を考慮する必要がない。ここから、接触を伴う作業になりダイナミクスが複雑になるので、最適な速度指令値をモデルベース強化学習によって学習する。

- n4 Align: 穴の縁との接触を保ったまま、ペグの姿勢を修正する。状態・行動空間は  $\mathbf{s}^{(\text{Align})} = [\alpha, \dot{\alpha}]^T$ ,  $\mathbf{a}^{(\text{Align})} = \dot{\alpha}_d$  となる。ここでも、柔らかさにより安定した接触を保障できるため、z方向の状態行動空間は考慮しなくてもよい。さらに、ペグ先端は穴の縁で拘束されており、柔軟要素が6自由度の変位が許容できるため、並進運動のみで、ペグの姿勢を修正できる。
- n5 Insertion: ペグを穴の底まで挿入する。この際の状態・行動空間は  $\mathbf{s}^{(\text{Insertion})} = [z, \dot{z}]^T$ ,  $\mathbf{a}^{(\text{Insertion})} = [z_d, \dot{z}_d]^T$  となる。入りにy方向も考慮するのは、ジャミングを回避するためである。

以上により、各区分は状態・行動空間ともに1~2次元で表現できることが示された。

### 3.2 ガウス過程に基づくモデルベース強化学習

区分化された、fit、align、insertionの各区分をモデルベース強化学習で学習する。学習の目的は、Tタイムステップ分の累積期待コストを最小化させるような速度制御器のパラメータを取得することである。

本研究では、サンプル効率の高いモデルベース強化学習 Probabilistic Inference for Learning Control (PILCO)<sup>8)</sup>を使用する。PILCOはガウス過程回帰を使用し、モデルを学習する。ガウス過程を使用することで、データ不足による不確実性を予測分散として明示的に表すことができる。Tステップ先の状態予測を行うとき、PILCOは解析的モーメントマッチングによる近似推論手法を適用し、過去の状態の不確実性を伝播することができ、その結果効率的な学習を実現している。解析的モーメントマッチングの詳細な説明は松原の解説記事<sup>13)</sup>を参照されたい。

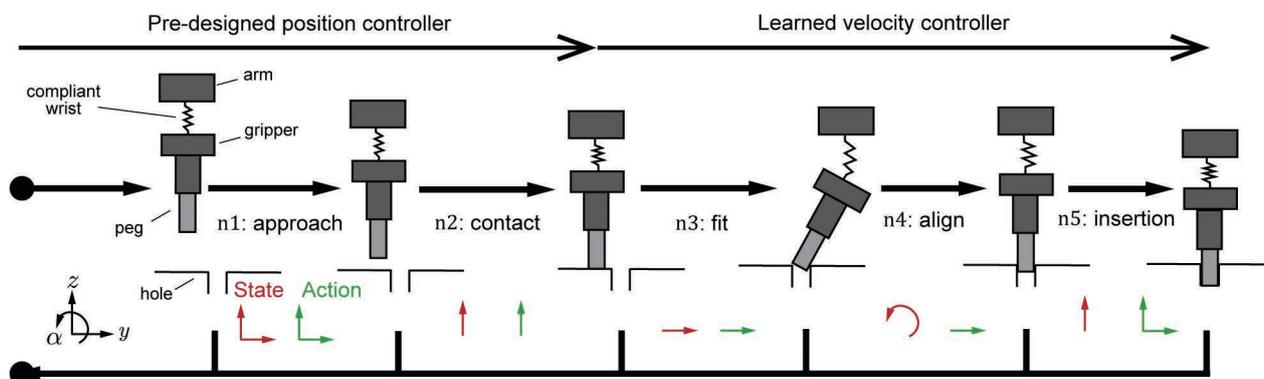


図2 環境と柔らかさを活用した組立作業におけるダイナミクスの低次元表現 (文献6)より引用。

### 3.3 シミュレーション

提案手法の有効性を確認するため、シミュレーションを行った。シミュレーションの目的は、提案手法が、低次元空間表現されていないシステムと比較して、より効率的に学習できるかを検証することにある。

セットアップ：本研究では、Box2Dという物理エンジンを使用し、2次元平面上のシミュレータを開発した。図2と同様に、 $y \cdot z \cdot \alpha$ 方向の変位が可能である。比較対象は、低次元空間表現を持たないダイナミクスで、状態・行動空間は  $s = [y, z, \alpha, \dot{y}, \dot{z}, \dot{\alpha}]^T$ ,  $a = [\dot{y}_d, \dot{z}_d]^T$  で表現される。

コスト関数は、それぞれの小課題における目標位置・姿勢の誤差である。1学習実験における試行回数は20回とした。

結果：図3にFit、Align、Insertionの各学習試行における、目標位置・姿勢の平均絶対誤差を示す。青線は提案手法、赤線は低次元空間表現を用いない手法である。エラーバーは、30回の学習実験の標準偏差を示す。提案手法を使用した場合、より小さい誤差を示し、3回の試行で収束していることを確認した。この結果から、提案手法を使用することでより効率的に学習できることを示した。

### 3.3 実機実験

提案手法が実環境でも動作するかを検証するために、実

機実験を行った。柔軟手首を協働ロボット UR5 に搭載した。ロボットのグリップの位置・姿勢はモーションキャプチャで計測した。実験で使用したベグと穴は、直径10mm、はめあい公差はH7/h7であった。

学習後のロボットの動作のスナップショットを図4に示す。シミュレーション同様、ロボットは、Fit、Align、Insertionの小課題を2～3回程度の試行回数で学習でき、学習終了時における成功率はそれぞれ8/10、9/10、10/10であった。以上より、実環境でも少数の試行回数で学習できることを示した。

## 4. 多様な環境に即座に適応する Sim-to-Real 転移学習

本節では、シミュレータで事前に学習したモデルを多様な実環境に転移させる、Sim-to-Real 学習について解説する。

### 4.1 多様な環境への Sim-to-Real 転移学習

近年の学習技術の発展によって、ロボットがより効率的に複雑なタスクを学習できるようになってきた。しかし、実環境でのデータ取得にかかる労力は依然として大きい。そのため、シミュレーション学習した事前知識を、実世界に転移させる Sim-to-Real の研究が広く行われている。我々

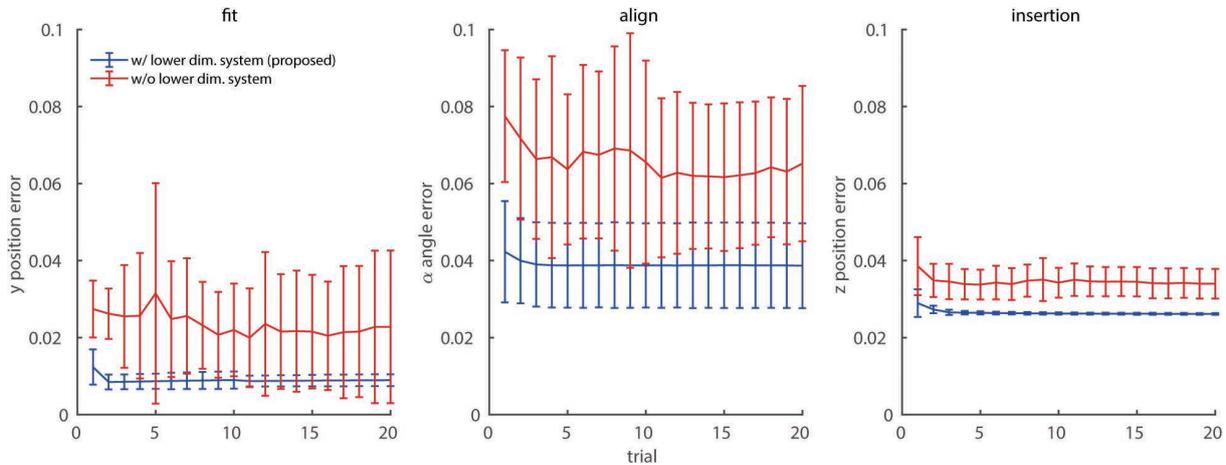


図3 シミュレーション結果 (文献6) より引用)。



図4 学習後のロボットの組立動作 (文献6) より引用)。

は、事前学習されたモデルが、多様な環境に対していかに即座に適応できるかについて興味がある。Sim-to-Realの一般的なアプローチとして、Domain Adaptation (DA)<sup>14)</sup> と Domain Randomization (DR)<sup>15)</sup> が挙げられる。しかしながら、DAは未知環境に適応するために追加で訓練する必要があり、DRは環境のパラメータ数が多い時に獲得された制御器もしくはモデルが保守的になることが指摘されている<sup>16)</sup>。

この問題に対し、我々は、Explicitly and Implicitly Conditioned Network (EXI-Net) というネットワークを提案した<sup>14)</sup>。図5にEXI-Netの概要図を示す。EXI-Netはダイナミクスモデルを、明示的 (Explicit)・暗示的 (Implicit) パラメータで条件付けされたニューラルネットワークで学習する。明示的パラメータは、質量や摩擦係数など、定量的に表現できるものを表現し、暗示的パラメータは、形状など定量的に示せないものを表現する。

シミュレーションにおける訓練時は、明示的パラメータを与えて、順ダイナミクスモデルのニューラルネットワークのパラメータと暗示的パラメータを学習する。テスト時には、未知環境における明示的・暗示的パラメータを誤差順・逆伝搬を反復的に行いながら推定し、最適な行動を計算する。この反復的計算による推定によって、未知環境にオンラインで適応することができる。

本研究では、ロボットが物体を目標位置まで押す課題に着目する。この課題は押す物体によって、明示的・暗示的パラメータが大きく異なるので、多様な環境に適応できるかを検証するのに適している。

### 4.2 EXI-Net を用いたモデル学習

まず、シミュレーション上でモデルを訓練させる。EXI-Netによるダイナミクスモデルは  $s_{t+1} = f(s_t, a_t, d_e, d_i)$  で与えられる。ここで、 $d_e$ は明示的パラメータ、 $d_i$ は暗示的パラメータである。 $d_e$ は質量・摩擦・重心位置で構成されており、訓練中ランダムに与えられる。一方、同時に複数の異なる形状の物体も与えられ、 $d_i$ はパラメトリックバイアス<sup>17)</sup> という手法で学習される。また、ランダムに行動を与えて、自己教師あり学習によってモデルを学習する。

次に、テスト時において、EXI-Netが未知物体にどのように適応するかを説明する。前節と同様、EXI-Netも、順モデル予測によって、コスト関数を最小化させる行動を計算する。この手法では、最適な行動は、ニューラルネットワークの誤差順・逆伝搬によって以下のように求められる。

$$\begin{aligned} g_a &= \partial c(s_{t+1}^{goal}, s_{t+1}^{pred}) / \partial a_t, \\ a_t &\leftarrow a_t - \varepsilon_a g_a / \|g_a\|_2, \end{aligned} \tag{2}$$

ここで  $\varepsilon_a$  は学習率である。 $c$  はコスト関数であり、目標状態と予測状態の2乗誤差で定義される。また、未知環境における  $d_e, d_i (d = [d_e^T, d_i^T]^T)$  とする) も同様に誤差順・逆伝搬によってオンラインで計算することができる。

$$\begin{aligned} g_d &= \partial c(s_{t+1}^{obs}, s_{t+1}^{pred}) / \partial d, \\ d &\leftarrow d - \varepsilon_d g_d / \|g_d\|_2, \end{aligned} \tag{3}$$

ここで、 $\varepsilon_d$  は学習率である。勾配は過去の時間  $\tau$  に獲得された状態・行動・次の状態のペアのデータをランダムに

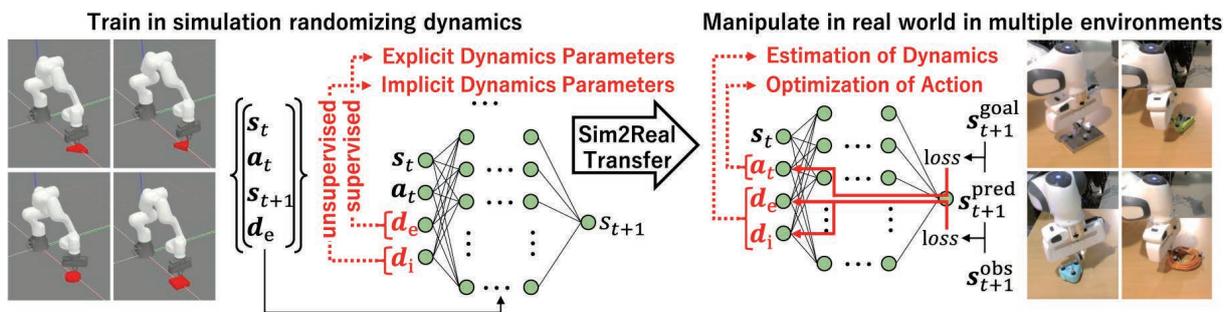


図5 EXI-Netの概要 (文献7)より引用。

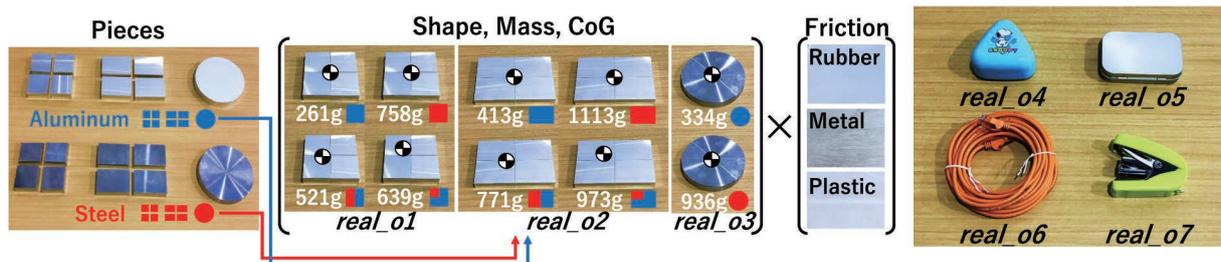


図6 テストに使用した物体 (文献7)より引用。

サンプルして計算する。これにより、未知の環境においても、最適な行動とパラメータ  $d$  をオンラインで計算することができる。

### 4.3 実験

提案手法の有効性を検証するために、実機実験を行った。実験の目的は、提案手法が、一般的に使用されている DR の手法と比較して高い成功率を示すかを確認することである。

セットアップ: シミュレータ上では、長方形 2 種類・三角形 3 種類・円の 1 種類、計 6 種類の形状を持つ物体で訓練を行った。実験には Gazebo というシミュレータを使用した。順モデルには、3 層、ニューロン数 100 個の全結合ネットワークを使用した。活性化関数には ReLU を使用した。各物体にそれぞれ、60～120 種類の明示的パラメータ  $d_e$  を与えて、ネットワークパラメータと暗示的パラメータ  $d_i$  を学習した。

テスト時における実環境の物体を図 6 に示す。real\_o1～o3 では多様な質量・重心・摩擦係数を設定するために、アルミと鉄のプレートを組み合わせ、底面にはゴムやプラスチックのシートを貼り合わせたものを設計した。また、真のダイナミクスが未知である物体 real\_o4～o7 も検証した。比較対象は、提案手法 (EXI) と EXI-Net に真の  $d_e$  が与えられた場合 (EXI w/ True  $d_e$ , real\_o1～3 のみ対象)、従来の DR で使用されている、LSTM を使用したネットワーク (LSTM) である。real\_o1～3 では、各物体にそれぞれ 12 条件、12 条件、6 条件あり、各条件で 5 試行 (計 150 回) した。また、real\_o4～7 では、各物体 10 試行した。テスト中は、ロボットが物体を 15 回押す間に指定された目標位置に到達するかどうかを確認した。

実験結果: 実験結果を表 1 に示す。各物体における、平均の成功率・目標到達のためのステップ数、真の  $d_e$  との誤差を示す。この表より、すべての物体で提案手法が最も高い成功率を示し、少ないステップ数で目標に到達していることがわかる。興味深いことに、真の  $d_e$  が与えられた場合と比較しても、提案手法の方が高い性能を示している。これは、シミュレーションと実環境とのモデル化誤差が残っているためであり、提案手法はオンラインパラメータ推定によってこの誤差を減少させていることが原因であると考えられる。

表 1 テスト環境における結果 (成功率/到達ステップ数/真の  $d_e$  との誤差)

Object	EXI w/ True $d_e$	LSTM	EXI
real_o1	0.89/11.3/0.0	0.72/12.8	1.0/8.3/0.52
real_o2	0.95/13.2/0.0	0.69/13.1	0.98/6.2/0.33
real_o3	0.98/10.8/0.0	0.92/13.4	1.0/7.9/0.38
real_o4	—	0.5/13.8	1.0/7.0
real_o5	—	0.6/14.5	0.9/9.5
real_o6	—	0.5/13.3	1.0/7.2
real_o7	—	0.5/13.6	1.0/6.8

以上の結果より、提案手法が従来手法と比較してより高い成功率で物体を操作することができた。この手法は、硬いロボットを使用しているが、現在我々は柔軟手首を持つロボットのシミュレータを開発しており<sup>18)</sup>、柔らかいロボットへの応用や、様々な部品の組立への応用も期待できる。

## 5. むすび

本稿では、組立自動化に向けて柔らかいロボットをより容易に制御するために、モデルベース学習を活用した研究事例について紹介した。モデルベース学習の手法は、モデルフリーの手法と比較してサンプル効率が高く、実ロボットの応用に適している。我々はさらに学習を効率化させるために、作業を区分化し柔らかさを活用することで、よりシンプルなダイナミクスにすることで学習を効率化させることができた。また、柔らかいロボットにも十分応用ができる、多様な環境に適應できる Sim-to-Real 転移学習に関する手法も紹介した。我々は、EXI-Net という、シミュレータ上でモデルを事前学習し、未知環境のダイナミクスパラメータをオンラインで推定できる手法を提案し、従来の一般的な Domain Randomization の手法より高い成功率を示した。

本稿を読み、モデルベース学習に基づく柔らかいロボット制御や、Sim-to-Real 転移学習に興味を持っていただければ幸いです。

## 参考文献

- 1) Li, R.; Qiao, H. A survey of methods and strategies for high-precision robotic grasping and assembly tasks - some new trends. IEEE/ASME Transactions on Mechatronics. 2019, Vol.24, No.6, p.2718-2732.
- 2) Beltran-Hernandez, C. C.; Petit, D.; Ramirez-Alpizar, I. G.; Nishi, T.; Kikuchi, S.; Matsubara, T.; Harada, K. Learning force control for contact-rich manipulation tasks with rigid position-controlled robots. IEEE Robotics and Automation Letters. 2020, Vol.5, No.4, p.5709-5716.
- 3) 井尻善久, フェリクス フォンドリガルスキ, 産業用ロボッ

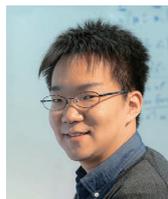
トの進化によるものづくりの近未来. 日本ロボット学会誌. 2019, Vol.37, No.8, p.675-678.

- 4) Tanaka, K.; Von Drigalski, F.; Hamaya, M.; Lee, R.; Nakashima, C.; Shibata, Y.; Ijiri, Y. "A compact, cable-driven, activatable soft wrist with six degrees of freedom for assembly tasks". IEEE/RSJ International Conference on Intelligent Robots and Systems. 2020, p.8752-8757.
- 5) Ibarz, J.; Tan, J.; Finn, C.; Kalakrishnan, M.; Pastor, P.; Levine, S. How to train your robot with deep reinforcement learning: lessons we have learned. The International Journal of Robotics Research. 2021, Vol.40, No.4-5, p.698-721.
- 6) Hamaya, M.; Lee, R.; Tanaka, K.; von Drigalski, F.; Nakashima, C.; Shibata, Y.; Ijiri, Y. "Learning robotic assembly tasks with lower dimensional systems by leveraging physical softness and environmental constraints", IEEE International Conference on Robotics and Automation. 2020, p.7747-7753.
- 7) Murooka, T.; Hamaya, M.; von Drigalski, F.; Tanaka, K.; Ijiri, Y. "EXI-Net: explicitly/implicitly conditioned network for multiple environment sim-to-real transfer". Conference on Robot Learning, 2020, [https://corlconf.github.io/corl2021/paper\\_268/](https://corlconf.github.io/corl2021/paper_268/), (参照 2021-06-30).
- 8) Deisenroth, M. P.; Fox, D.; Rasmussen, C. E. Gaussian processes for data-efficient learning in robotics and control. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015, Vol.37, No.2, p.408-423.
- 9) Lenz, I.; Knepper, R. A.; Saxena, A. DeepMPC: Learning deep latent features for model predictive control". Robotics: Science and Systems. 2015.
- 10) Johannsmeier, L.; Gerchow, M.; Haddadin, S. "A framework for robot manipulation: Skill formalism, meta learning and adaptive control". IEEE International Conference on Robotics and Automation. 2019, p.5844-5850.
- 11) Hang, K.; Morgan, A. S.; Dollar, A. M. Pre-grasp sliding manipulation of thin objects using soft, compliant, or underactuated hands. IEEE Robotics and Automation Letters. 2019, Vol.4, No.2, p.662-669.
- 12) Deimel, R.; Brock, O. A novel type of compliant and underactuated robotic hand for dexterous grasping. The International Journal of Robotics Research. 2016, Vol.35, No.1-3, p.161-185.
- 13) 松原崇充. ガウス過程に基づくロボットの運動制御・学習—解析的モーメントマッチングによる近似推論. システム/制御/情報. 2016, Vol.60, No.12, p.515-520.
- 14) Patel, V. M.; Gopalan, R.; Li, R.; Chellappa, R. Visual domain adaptation: A survey of recent advances. Signal Processing Magazine. 2015, Vol.32, No.3, p.53-69.
- 15) Peng, X. B.; rychowicz, M.; Zaremba, W.; Abbeel, P. "Sim-to-real transfer of robotic control with dynamics randomization". IEEE International Conference on Robotics and Automation, 2018, p.3803-3810.
- 16) Nachum, O.; Ahn, M.; Ponte, H.; Gu, S.; Kumar, V. Multi agent manipulation via locomotion using hierarchical sim2real, arXiv preprint arXiv:1908.05224, 2019, (参照 2021-06-30).
- 17) Ogata, T.; Ohba, H.; Tani, J.; Komatani, K.; Okuno, H. G. "Extract-

ing multi-modal dynamics of objects using RNNpb". IEEE/RSJ International Conference on Intelligent Robots and Systems. 2005, p.966-971.

- 18) Hamaya, M.; Tanaka, K.; Shibata, Y.; von Drigalski, F.; Nakashima, C.; Ijiri, Y. Robotic learning from advisory and adversarial interactions using a soft wrist. IEEE Robotics and Automation Letters. 2021, Vol.6, No.2, p.3878-3885.

## 執筆者紹介



**濱屋 政志 HAMAYA Masashi**  
 オムロン サイニクエクス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門：ロボティクス、機械学習、強化学習  
 所属学会：日本ロボット学会、IEEE  
 博士（工学）

本文に掲載の商品の名称は、各社が商標としている場合があります。

# Advancing Automated Assembly–Team O2AC at the World Robot Summit 2020

*FELIX von drigalski, NAKASHIMA Chisato, KASAURA Kazumi and CRISTIAN beltran-hernandez*

When installing robot systems in manufacturing applications, only 20 to 30% of the cost is due to the hardware, while more than 50% of the cost comes from system integration, such as setting up, programming and connecting peripheral devices into a unified system. Given the labor shortages both in Japan and worldwide, robot systems that can handle a variety of products with low production volumes are strongly desired. We participated in the World Robot Summit (WRS) Industrial Category to move towards its proclaimed goal of “agile one-off manufacturing”, and aimed to build a production system that can respond quickly, leanly and without waste to various changing production requirements, thereby realizing the “manufacturing of the future”.

Together with other organizations, we formed a team and have been working since 2018 on a system that is robust to changes in parts and assembly positions, without using parts-specific jigs or teaching. We have continuously improved the system by tackling issues that came to light at the pre-convention (WRS2018) and other events, where we obtained the SICE Special Award. In addition, we participated in the WRS2020 (postponed to 2021 due to Covid-19) as Team O2AC and obtained 3<sup>rd</sup> place as well as the Artificial Intelligence Society Award.

In this article, we describe our approach to solving the assembly problem, the hardware and software technologies we used to implement our solution, and the lessons we learned during the development.

## 自動組み立て

フェリックス フォン ドリガルスキ, 中島 千智, 笠浦 一海, クリスティアン ベルトラン

ものづくり現場で導入されているロボットシステムのロボット本体のコストは全体の2、3割に過ぎず、周辺装置や周辺機器を1つのシステムとして組み上げるシステムインテグレーションに5割以上のコストが割かれており、かつ時間もかかるため変更が容易でない。一方で近年では国内外での労働者不足や製品のライフサイクルも変わってきていることから、変種変量生産に対応可能なロボットシステムの実現が期待されている。World Robot Summit (以下 WRS) のものづくり分野では “toward agile one-off manufacturing=迅速な一品ものづくりを目指して” を目標に掲げ、様々に変化する生産要求に迅速かつ無駄なくスリムに対応できる生産システムを、競技会を通じ構築することで、「未来のものづくり」の実現を目指している。

筆者らは他組織とともにチームを結成し、極力治具を使わず部品や組付け位置の変更にロバストなシステムの研究開発に、プレ大会である WRS2018 が開催された 2018 年より取り組んできた (この年は計測自動制御学会賞を受賞)。そしてプレ大会などを通じて見えてきた課題の解決に挑戦し続けシステムをブラッシュアップしてきた。また、コロナの影響で 2021 年に延期されていた WRS2020 にチーム O2AC として参加し、総合 3 位に入賞したことに加え人工知能学会賞を受賞した。

本稿では、この競技会で使用したシステムのハードウェアおよびソフトウェア面における創意工夫について紹介する。

## 1. Introduction

Automated assembly with robotic arms is highly challenging, as it places high requirements on the physical robot system, as well as the software implementation. Due to the tight part tolerances, it requires high mechanical precision and fine control. Because of the many interlocking steps in the assembly process which can change according to part measurements and positions, it is laborious to program the motions (“teach”) manually.

Commonly, assembly tasks are realized by building parts-specific jigs, which ensure that each part is in a known position, so that the process can be repeated by the robot in an identical manner. However, the cost and size of these jigs makes it intractable to use the same robot system for many different types of products, or to change from one product to another.

In order to advance towards jigless and teachless assembly, we built a system for the World Robot Summit 2020 Assembly Challenge with Team O2AC (OMRON SINIC X, OMRON, Osaka University, AIST, Chukyo University). We obtained the 3<sup>rd</sup> place (with only 2 points to the 2<sup>nd</sup> place) and an award from the Japanese Society for Artificial Intelligence. In this article, we describe our approach and lessons drawn from the development.

## 2. System Design

As our goal was to test the limits of and challenges in current jigless and teachless assembly, we teamed up with universities and research institutions which did not have access to and/or experience with industrial hardware. To collaborate efficiently and immediately start working productively, we decided to build our prototype 1) using ROS-based open software, and 2) as a lean robot manipulation system that will stay useful beyond the WRS, with little to no hardware customization. This minimized the time spent on hardware implementation, allowed the use of public, reviewed code and made collaboration significantly more efficient.

The code running our system is public and can be found at <https://github.com/o2ac/o2ac-ur>

### 2.1 Hardware—Two arms with hand-held tools

After having built a prototype with 3 arms in 2018<sup>1,2)</sup>, for the WRS2020 we decided to make the system as lean as possible. We used only two generic robot arms equipped with a parallel gripper and wrist cameras, together with hand-held tools and parts feeders/holders. Fig. 1 shows the complete robot system.



Fig. 1 The O2AC robot system at the World Robot Summit 2020 in Aichi-ken.

The system was controlled by two PCs, running an Ubuntu real-time kernel and Nvidia CUDA drivers respectively. This allowed both high-speed robot control as well as neural-network based computer vision algorithms.

We also used hand-held tools which are described in section 5.

### 2.2 Software—Rapid prototyping and collaboration

We used ROS, MoveIt and other open-source software to drive our solution. The system is encapsulated in a Docker container, which can be downloaded and run on any Ubuntu PC (18.04+). This allowed us not only to collaborate quickly and efficiently with our external collaborators, but also to use existing open code and drivers rather than writing custom ones, which accelerated our development and prototyping speed.

Among the packages and code we developed or used:

- An **assembly database** which generates target part positions based on their CAD models, so that motions are updated automatically according to part specifications
- A **symbolic planner** that **tracks uncertainty** (see section 3) and generates actions that reliably position the part in known positions, even in environments with high levels of noise
- A **computer vision** system (see section 4) that uses neural networks, image processing and CAD matching to determine the position of parts as well as the success/failure of assembly operations
- A **hybrid force-position** controller applicable to any 6-DOF robot arm, which succeeded on all the insertion tasks in the competition

If you use our framework, you can define how parts are assembly and display how parts like Fig. 2.

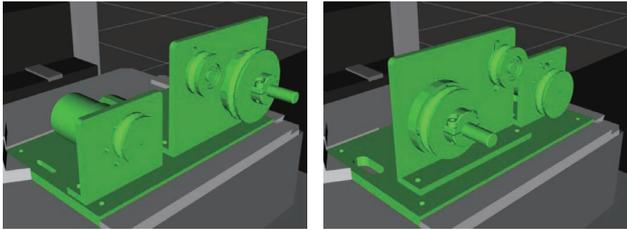


Fig. 2 Different assemblies defined by their CAD models, generated dynamically. Robot motions are adjusted automatically. **Left:** Standard. **Right:** Surprise Plus.

Notably, we developed a scheduler for MoveIt which allows simultaneous planning and execution with multiple robot arms, and which ensures safety even for overlapping workspaces. This reduced overall execution time by 20–40% and opened up new avenues for operating even more efficiently in multiple tasks.

### 3. Uncertainty-aware Manipulation

When detecting an object using computer vision, the result is always subject to an unavoidable amount of noise and uncertainty. Conventional approaches attempt to work around this by using expensive sensors, very rigid and precisely manufactured robot bases, and complicated calibration procedures.

By contrast, we accept that some uncertainty cannot be avoided, and integrate it into our planning and design process. As shown in Fig. 3, we associate every object pose result obtained from vision with an uncertainty, which is reduced only in the course of manipulation. Actions such as grasping and pushing the object limit the possible positions, and gradually reduce the uncertainty until the position is known with high precision and certainty.

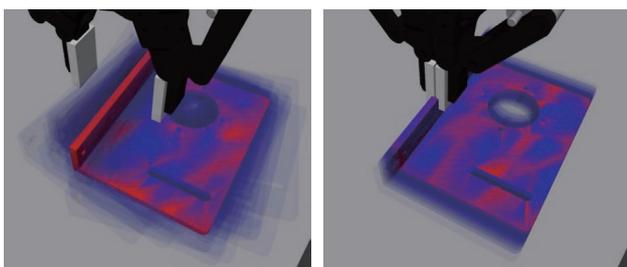


Fig. 3 An L-plate recognized by the vision system, at the assumed position (red) with uncertainty (blue). **Left:** Before the grasp, many positions are possible. **Right:** After the grasp, the possible positions are restricted.

#### 3.1 Planning with uncertainty and the environment

In our previous papers [3], we have proposed different actions to reduce the uncertainty of grasped objects. In a paper that we are currently preparing, we extend these actions to include extrinsic manipulation, such as releasing the object from the gripper, pushing it with the gripper, or lightly grasping the object while

pushing it into an obstacle. These actions are highly effective and have been used in multiple occasions in the competition.

Based on these actions, we have built a planner which generates action sequences that succeed reliably even if the object position has an error of multiple millimeters. Rather than fine-tuning each position, the system generates full motion sequences to perform assembly procedures. These sequences can then be inspected and optimized by the user, which makes programming and teaching significantly easier. Fig. 4 shows example planning results.

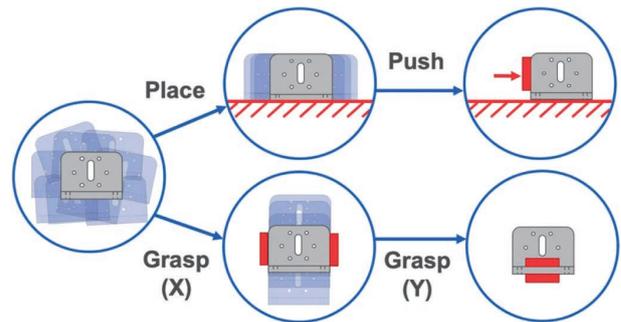


Fig. 4 Different action sequences generated by the planner to determine the object position with high certainty.

### 4. Computer Vision

We are using a combination of computer vision techniques, based on both 2D and 3D image data obtained from the robots’ wrist-mounted cameras. The cameras are consumer-grade (Intel Realsense SR305), which limits the maximum precision of the vision pipeline. However, as described in the previous section, our approach accepts this uncertainty and achieves precise positioning even with very affordable sensors.

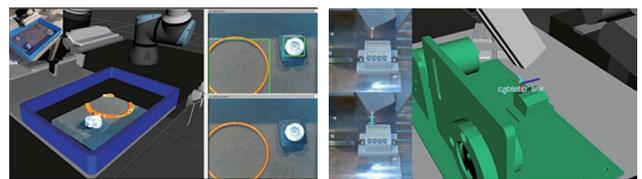


Fig. 5 The vision system developed with Prof. Akizuki from Chukyo University and Toshio Ueshiba from AIST.

Fig5-1: Parts being recognized in the tray (using neural networks)

Fig5-2: Cable tip detection (using background subtraction)

#### 4.1 Tray picking

We used a combination Single-Shot MultiBox Detector[4] and 3D Localization (Fig. 5–1). (TODO: Cite the paper) A common difficulty with neural-network based approaches is the sensitivity to lighting conditions. Due to both COVID-19 and the distributed nature of our team, we had to tackle the robustness of our vision system early on, as the lighting in the OMRON SINIC X office and the laboratory in Chukyo University was very different. As a result, our system was very robust to differ-

ent lighting conditions, and we experienced practically no difficulty with the conditions at the competition site.

In fact, the RGB-based parts detection was so robust that it often outperformed the 3D-point-cloud-based recognition routines. In retrospect, it would have been very worthwhile to encode the rotation angle of parts in the network, to make use of the high speed of the network.

#### 4.2 Bearing orientation

Unlike scene-mounted cameras mounted on the work cell's ceiling, the wrist-mounted cameras allow close-up views of assembled objects. As shown in Fig. 6, the bearing is inspected after insertion, to ensure that the holes align with the screws' target positions. If they do not match, the bearing is rotated in-place by the required amount.

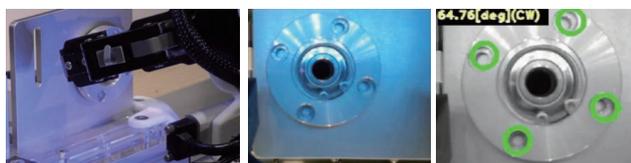


Fig. 6 LED-equipped wrist cameras allow parts inspection from up close, to determine assembly success.  
Left: Camera facing the bearing. Middle: Camera view. Right: Vision result (angle offset from edge detection)

#### 4.3 Cable detection

As the cable is highly deformable, the position of the tip cannot be predicted reliably, and detecting it in an unstructured environment is highly difficult. We solve this problem by first executing a synchronized collaborative motion with both robot arms, which cages the cable and places the tip in one of the robots' grippers. Then, we use background subtraction to find the tip of the cable in the image and project it back into 3D space (Fig. 5-2).

### 5. Compliant, hand-held tools

Instead of using expensive and specialized tool-changers, we used tools that were grasped by the gripper, as a human hand would. These tools were of three different types, as shown in Fig. 7:

- Suction-equipped screw tools, with compliance (using precision parts from SAWA Ltd.)
- 3D-printed screw tools with compliance, but without suction
- Passive tools (e.g. hooks, plungers)



Fig. 7 Different tool types. Left: Screw tool with suction and compliance. Middle: Screw tool with compliance and without suction. Right: Passive tools.

The screw tools were driven by small Dynamixel motors with torque control. Combined with the suction pressure sensor, this allowed the detection of both motor stalling and screw contact, which we used to determine successful fastening of screws.

However, some tools that were used for specific manipulations, just like crochets used by humans, did not need to be actuated. For example, we used a hook tool to thread the belt around the pulleys, and a plunger tool to separate items that were packed too tightly to be picked.

The screw tools and plunger used physical compliance in order to apply light pressure to the environment. As a comparable level of compliance would require a very high control frequency if generated by the robot's joint actuators, physical compliance allows the use of affordable robots, while ensuring high reliability and performance.

### 6. Lessons

One significant difficulty that we had hardly encountered during testing occurred frequently in the competition: screws becoming stuck in the thread before the fastening is finished. We determined the success of the fastening procedure on the basis of the pressure sensor and torque control in our screw tools, but this ended up being insufficient.

To ensure that the screw was flush after fastening, we should have used the wrist camera to confirm screwing success. The wrist camera could have also been used to track the tool tip and implement a fallback procedure. This cost us a significant amount of points during the final days of the competition and would surely have obtained the 2 points separating us from the 2<sup>nd</sup> place in the competition.

Another takeaway was the importance of developing in a distributed manner and in simulation. Enabling development in simulation early on allows more people to be productive without the robots and allows quicker and safer testing.

Finally, we found that it is important to visualize the system state as early as possible. This not only helps in debugging and understanding errors, but also in explaining the system to new users and interested parties, such as the audience and jury at the WRS2020.

## 7. Conclusion

We presented a complete robot system with the capability to solve the World Robot Summit 2020 Assembly Task as well as the Surprise Plus task. It will continue to be used as a research platform at OMRON SINIC X, now that it has proved its performance and usability. The open release of its source code should make waves and leave a favorable impression on the open-source community: <https://github.com/o2ac/o2ac-ur>

Finally, we wish to thank our team members at Osaka University, AIST and Chukyo University for their hard work.

## References

- 1) von Drigalski, F.; Nakashima, C. et al. Team O2AS at the world robot summit 2018: an approach to robotic kitting and assembly tasks using general purpose grippers and tools. *Advanced Robotics*. 2020, Vol. 34, p. 514–530.
- 2) von Drigalski, F. et al. Robots assembling machines: learning from the World Robot Summit 2018 Assembly Challenge. *Advanced Robotics/Francis & Taylor*. 2020, Vol. 34, p. 408–421.
- 3) von Drigalski, F. et al., “Contact-based in-hand pose estimation using Bayesian state estimation and particle filtering”, *IEEE International Conference on Robotics and Automation (ICRA)*. 2020, p. 7294–7299.
- 4) Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C. Y.; Berg, A. C. Ssd: Single shot multibox detector. In *European conference on computer vision*. Springer, 2016, p. 21–37.

## About the Authors



*FELIX von drigalski, Ph.D.*  
(Engineering)

OMRON SINIC X Corporation  
Robotics Group  
Specialty: Robot Manipulation  
Affiliated Academic Society: IEEE



*NAKASHIMA Chisato*

Technology and Intellectual Property H.Q.  
Core Technology Center,  
Robotics Department  
Specialty: System Engineering



*KASAURA Kazumi*

OMRON SINIC X Corporation  
Robotics Group  
Specialty: Mathematics (Number Theory) /  
Algorithm / Programming  
Affiliated Academic Society: IEEE



*CRISTIAN  
beltran-hernandez*

Osaka University  
Graduate School of Engineering Science  
Specialty: Robot Manipulation

---

The names of products in the text may be trademarks of each company.



## 2

---

人から学ぶ 人のように学ぶ

# 機械学習を活用した経路計画技術

谷合 竜典, 米谷 竜

経路計画は、環境中において与えられたスタートからゴールまでの障害物と衝突しない経路を出力する問題であり、モバイルロボットの自律移動におけるグローバルプランニングやマニピュレータの動作計画など、様々な分野において広く取り組まれている。このような経路計画問題に対して、機械学習を活用することで性能の向上を図る取り組みが近年多く報告されている。本稿では機械学習ベースの経路計画技術についていくつかの代表的アプローチを紹介するとともに、著者らの最近の取り組みである深層学習ベースのA\*探索技術 (Neural A\*) について詳細に解説する。

## Path Planning using Machine Learning

TANIAI Tatsunori and YONETANI Ryo

Path planning refers to a problem of finding a low-cost and valid path between start and goal in a given environment map. This problem has been studied widely for a variety of robotics applications such as the global planning for autonomous mobile robots and motion planning for robot manipulators. Recently, some work has proposed to leverage machine learning techniques for improving path planning performances. In this paper, we overview several popular approaches of such machine-learning-based path planning, and furthermore, introduce our recent achievement called Neural A\*, a deep-learning-based A\* search for path planning problems.

### 1. まえがき

経路計画 (path planning, pathfinding) は人工知能分野やロボティクス分野において古くより取り組まれる基本的問題である。いま、図 1a のように障害物のある環境を考える。経路計画は、この環境において与えられたスタート地点からゴール地点に至るまでの、障害物に衝突しない最短の移動経路 (path) を出力する問題である。たとえばモバイルロボットや自動運転車両の自律移動といったシナリオでは、環境としてSLAM等で獲得された障害物の地図、スタート・ゴール地点としてはその地図上における自律移動体の現在・目的地を考えることになる。あるいはロボットマニピュレーションの分野では、環境としてロボットの関節の状態等で定義されるコンフィギュレーション空間、スタート・ゴール地点としてロボットの開始・目標状態が与えられることになる (このような問題はしばしば動作計画、motion planning と呼ばれ経路計画と区別される)。いずれの状況においても、経路問題を解くことによって、我々はロボットを安全かつ短時間で所望の状態に移動させることが可能となる。

経路計画技術は、環境をどのように表現するか、またそ

の表現された環境において経路をどのように探索するかという観点で、いくつかの種類に分類することができる。たとえばモバイルロボットの経路計画においては、図 1b のように環境を格子地図に離散化したうえで4近傍や8近傍のグラフとして表現し、A\*探索や最良優先探索といったグラフ探索アルゴリズムによってスタート地点からゴール地点への最短経路を探索することが一例として挙げられる。ロボットの移動する環境が非常に広域である場合、あるいはロボットマニピュレーションのように環境が高次元空間である場合は、図 1c のように環境中でランダムに選択された地点をつないだグラフ (ロードマップ) を構築したうえで経路計画をすることや、スタート地点からゴール地点に向かって逐次的に次の移動候補地点を選択しつつ経路計画をすることがある。いずれの場合も、環境を細かい格子地図や多数地点から構成されるロードマップで表現することによって経路計画の成功率や得られる経路長を改善することができるが、それにもなって経路計画の計算コストも増大する。すなわち、経路計画の性能とその計算コストはトレードオフの関係にある。

これに対して近年、機械学習技術を活用することによってこのトレードオフを改善する研究が活発に進められている。直感的には、経路計画の問題インスタンス (環境、ス

Contact : TANIAI Tatsunori tatsunori.taniai@sinicx.com

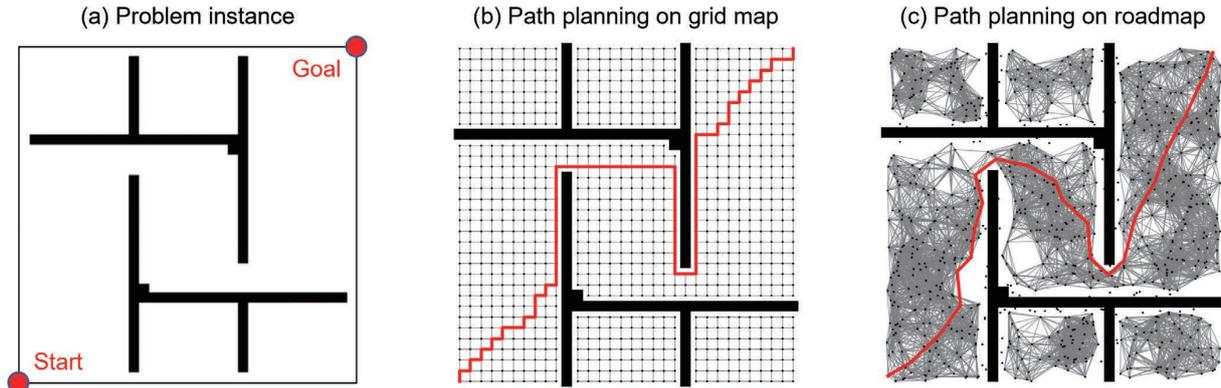


図 1 経路計画の例

スタート、ゴールの組合せ)とその解(スタートからゴールまでの経路)の組を学習データとして用意し、その関係を何らかの機械学習モデルにより学習する。これにより、新たな問題インスタンスに対して学習済みモデルを適用することで、解となる経路を効率的に計画するための手がかりを得ることができる。あるいは、典型的な経路計画技術では環境中における障害物の位置は既知である必要があるが、未知の環境に対して機械学習モデルを適用することで障害物の配置を推定し、それに基づいて経路計画を実行することもできる。本稿ではこのような機械学習を活用した経路計画に関していくつかの代表的な技術を紹介するとともに、著者らの近年の成果である深層学習ベースのA\*探索について詳しく解説する。

## 2. 機械学習を活用した経路計画技術

前節で述べた通り、経路計画技術を分類する一つの基準として環境の表現方法がある。以下では、環境を格子地図に離散化した探索ベースの経路計画およびランダムに選択された地点をつないだロードマップを用いるサンプリングベースの経路計画について、機械学習を活用した事例を紹介する。

### 2.1 機械学習を活用した探索ベース経路計画

格子地図上での経路計画は、しばしばA\*探索や最良優先探索が利用される。いずれのアプローチも、格子地図を4近傍や8近傍のグラフとして表現し、スタート地点からゴール地点まで最短経路を逐次的に探索する。いま、グラフを $G=(V, E)$ と表す。 $V$ は格子地図上の各地点を表すノードの集合であり、 $E$ は地点間で移動可能な経路を表すエッジの集合である。また、あるノード $v$ に対して接続された近傍ノードの集合を $N(v)=\{v' | (v, v') \in E, v \neq v'\}$ という関数で表現する。各ノードに対しては、そのノードに移動する際に発生する非負のコスト $c(v) \in \mathbb{R}_+$ が与えられる。このような条件下でスタート地点とゴール地点がそれぞれノード $v_s, v_g \in V$ として与えられたとき、それらを結ぶ経

路を $P=(v_1, \dots, v_T), v_1=v_s, v_T=v_g$ と表し、コストの総和 $\sum_{i=1}^{T-1} c(v_{i+1})$ が最小となる経路を探索することを目指す。

上記の問題に対して、探索ベース経路計画は一般に以下のアルゴリズムに基づいて行われる。

Algorithm 1: Search ( $G, c, v_s, v_g$ )

- 1: Initialize  $O \leftarrow v_s, C \leftarrow \emptyset, \text{Parent}(v_s) \leftarrow \emptyset$
- 2: While  $v_g \notin C$  do
- 3:   Select  $v^* = \arg \min_{v \in O} f(v)$  based on the criterion  $f(v)$
- 4:   Update  $O \leftarrow O \setminus \{v^*\}, C \leftarrow C \cup \{v^*\}$
- 5:   Extract  $V_{\text{nbr}} = N(v^*) \setminus (O \cup C)$
- 6:   Update  $O \leftarrow O \cup V_{\text{nbr}}$
- 7:   Foreach  $v \in V_{\text{nbr}}$  do
- 8:       Update  $\text{Parent}(v) \leftarrow v^*$
- 9:   Endfor
- 10: Endwhile
- 11: Compute  $P \leftarrow \text{Backtrack}(\text{Parent}, v_g)$
- 12: Return  $P$

ここで、探索の進捗はオープンリスト $O$ とクローズドリスト $C$ と呼ばれる二種類のリストによって管理される。オープンリストには解経路の候補となるノードが蓄積され、そこからある評価値 $f(v)$ が最小となるノードが選択され(3行目)、クローズドリストに移動される(4行目)。 $f(v)$ は手法によって異なるが、たとえばA\*では以下の形式で与えられる。

$$f(v) = g(v) + h(v) \tag{1}$$

ただし、 $g(v)$ はスタートノードから $v$ までのコスト総和、 $h(v)$ は $v$ からゴールまでのコスト総和の予測値を与えるヒューリスティック関数である。具体的には、スタートから $v$ までの経路が $P(v)=(v_1, \dots, v_T), v_1=v_s, v_T=v$ として与えられたとき、 $g(v)=\sum_{i=1}^{T-1} c(v_{i+1})$ である。一方、最良優先探索では $f(v)=h(v)$ となり、 $g(v)$ が与えるスタートノードから $v$ までどのような経路がとられてきたかという情報を

利用することなく、 $h(v)$ に基づいて貪欲にゴールまでの経路が探索されることになる。5行目では選択されたノード  $v^*$  の周辺ノードが展開され、6行目においてオープンリストに追加される。また、この追加にともなって、探索で発見された経路の接続関係を Parent 関数に記憶させる (7~9行目)。その結果 11行目では、ゴールノードからスタートノードまで Parent が示すノードを辿る (backtrack) ことにより、最終的な経路が構築されることになる。

このような探索ベースの経路計画において、機械学習を用いたアプローチは主にヒューリスティック関数  $h$  の学習とコスト関数  $c$  の学習という2種類に大別される。たとえば Choudhury らの手法<sup>1)</sup> では、探索時に選択されたノード  $v$  についてゴールへの最短経路長  $h^*(v)$  を実際に計算し、 $v$  からゴールまでのユークリッド距離や周囲の障害物への距離といった特徴量から  $h^*(v)$  を回帰するモデルを学習する。このとき、学習データとしては異なる障害物の配置からなる経路計画問題インスタンスと、その問題に対して最短経路を計算可能な経路計画アルゴリズム (たとえばダイクストラ法や A\* 探索) が与えられる。いったん学習が行われると、新たな問題インスタンスにおいてモデルから回帰されたヒューリスティック関数の値を利用しつつ最良優先探索をすることで、従来手法と比較してより少ない探索数で経路を発見することができる。Takahashi ら<sup>2)</sup> はこの手法をさらに発展させ、手で設計されていた特徴量の代わりに畳み込みニューラルネット (具体的には U-Net<sup>3)</sup>) を用いて  $h^*(v)$  の回帰を行なっている。一般にこのような深層学習ベースの手法は、手動での特徴量設計と比較して性能向上につながる良い特徴量表現をデータから獲得できる利点があるが、経路計画に関する本問題に関しては、学習データの生成にあたって各問題インスタンスの全ノードについてゴールへの最短距離  $h^*(v)$  を計算する必要があり、計算コスト面で課題がある。

一方 Vlastelica ら<sup>4)</sup> は、障害物配置が未知の環境画像からコストマップ  $c(v)$  を推定するアプローチを提案している。具体的には畳み込みニューラルネットを用いて環境画像のピクセルごとにコストを出力する。そして、その出力されたコストに基づくダイクストラ法により、スタートからゴールへの最短経路を算出する。このとき学習データとして、同環境画像には人手等により真の経路が与えられている状況を考えると、モデルの出力した経路と真の経路の間で損失 (ここではハミング損失) を計算することが可能になる。この損失を最小化するようにニューラルネットをブラックボックス最適化することにより、「学習データにおいて与えられた真の経路を計画するために必要なコスト関数」をニューラルネットが学習できる。このようなアプローチは経路探索にかかる計算・時間的コストを直接的に短縮するものではないが、「経路計画を実行するために必要なコストマップの獲得」というプロセスを単一のニュー

ラルネットワークに置き換えられる点で、モバイルロボットの自律移動のような経路計画を含むパイプライン処理の一部を効率化しているとみられることもできる。

Vlastelica らのアプローチは経路計画アルゴリズムをブラックボックス関数とみなすことで、ブラックボックス最適化によるコスト関数の学習を実現している。これに対して、我々は経路計画を微分可能な形で陽に書き直すことにより、通常の誤差逆伝播を用いたコスト関数の学習を実現している。このことにより、探索の最終結果のみ活用できる Vlastelica らの手法と異なり、我々の手法は探索の各ステップにおけるノード選択結果に基づいてコスト関数を学習できる。3節ではその詳細について解説する。

## 2.2 機械学習を活用したサンプリングベース経路計画

サンプリングベース経路計画においては、ロードマップの構築がしばしば手続き全体を律速する。これは、より良い経路を探索するためには環境全体に対して密にサンプルをまく必要がある点、そのサンプル集合からロードマップを構築するために大量の衝突判定が必要である点に起因する。これに対して、もし与えられた問題インスタンスに対して、実際にロードマップ構築をする前に最終的な解経路を予測できれば、その経路周辺を重点的にサンプリングすることにより、より少数のサンプルを用いた効率的なロードマップ構築と経路計画が可能となる。

このような着想のもと、いくつかの研究ではロードマップ構築のためのサンプリングモデル (サンプラー) の学習法が提案されている。たとえば文献5) では条件付き変分オートエンコーダを用いて、環境の障害物レイアウトおよび経路のスタート・ゴール地点を入力として解経路周辺の地点をサンプリング可能なモデルを学習している。このとき、学習されたモデルから得られるサンプルとランダムに抽出されたサンプルを混合してロードマップを構築することで、学習データと異なる環境でモデルからのサンプリングが実際の解経路から外れる場合であっても経路計画が可能となる。ほかに文献6) では、環境中において多様な経路が共通して通りうる地点を重要な (critical) 点と定め、周囲の障害物配置から地点の重要性を推定するモデルを学習している。このようなモデルはいったん学習されると新たな環境の各地点に対する重要度を推定可能であり、推定された重要度に比例した確率で同地点をサンプリングすることで、より少数のサンプルから効率的なロードマップ構築が可能となる。

なお、上記のアプローチでは経路探索に先立ってロードマップの構築が行われるが、環境を逐次的に探索するサンプリングベース経路計画に機械学習を活用する取り組みもある。たとえば Motion Planning Network と呼ばれる手法<sup>7)</sup> では、移動エージェントの現在位置とスタート・ゴール地点、障害物の情報を入力として、次ステップで移動すべき

地点を直接出力するニューラルネットワークを学習する。そして現在位置からその地点への移動が可能であるか判定し、移動不可能であったときには既存のリプランニングアルゴリズムを用いることで経路を生成する。このようなアプローチは環境全体をカバーするロードマップを構築しないためしばしば効率的に経路計画が可能であるが、得られた経路が最適（あるいは準最適）である保証はない。これに対してNeuralEXTと呼ばれる手法<sup>8)</sup>では、次ステップでの移動先に加えて、各地点におけるゴールまでの予測経路長（2.2節における $h^*(v)$ に対応する値）を予測するネットワークを学習する。これにより、2.2節で紹介した評価値 $f(v)$ に基づいて実際の移動先を決定し、より良い経路を探索することが可能となる。

### 3. Neural A\* による経路計画

ここでは、著者らの近年の成果である Neural A\*<sup>9)</sup> について解説する。Neural A\* は2.2節で紹介した探索ベースの経路計画に機械学習を活用するものである。先に紹介した通り、近年の機械学習を活用した経路計画の研究においては、いくつかの関心の高い問題があるが、本研究ではとくに、1) 最短経路問題における探索の効率性と最適性のトレードオフの向上、および、2) 障害物配置が未知の環境画像（生画像）に対する経路計画、という2つの問題に対して統一的なアプローチを提案する。

以下では、これら2種類の問題に対する提案アプローチの基本的な考え方、Neural A\* の学習方法の概要、そして Neural A\* の中核となる微分可能 A\* の詳細について、順々

に説明する。なお、2.2節で紹介した手法と同様、Neural A\* は主に格子地図上での経路計画を対象とする。

#### 3.1 ガイドマップを介した経路計画

1つ目の問題である最短経路探索について、実例を用いながらより詳しく説明したい。Neural A\* は A\* 探索アルゴリズムをベースにその探索性能を向上させる手法であるため、まずは通常の A\* 探索にどのような問題があるのかを見てみる。図 2a に示す経路計画の問題インスタンス（即ち、環境マップおよび開始・目標点の組）の例に対して、A\* による実際の経路探索の様子を図 2b に示す。図 2a を人間がみれば、開始・目標点を結ぶ最短経路はおよそ一瞬で見当がつきそうである。しかし、A\* 探索では手前にあるポケット状の行き止まりで探索が大きく滞留していることがわかる。

2つ目の問題である生画像に対する経路計画についても、通常の A\* 探索が有する問題点と見ることがができる。例として、図 3a に示すような屋外シーン画像を環境マップとし、マップ中の2点間を結ぶ移動可能な経路の計画問題を考える。図 3a の画像を人間が見れば、これが円形状の交差点であり、交差点中心部の領域は通行不可に見えるのでそこを避けた経路を思いつくだらう。実際に、ある歩行者の経路は図 3b のようなものであった。しかし、通常の A\* 探索では、どこにそのような障害物が存在するかのラベル付けがない生画像に対して、このような経路計画を遂行することはできない。

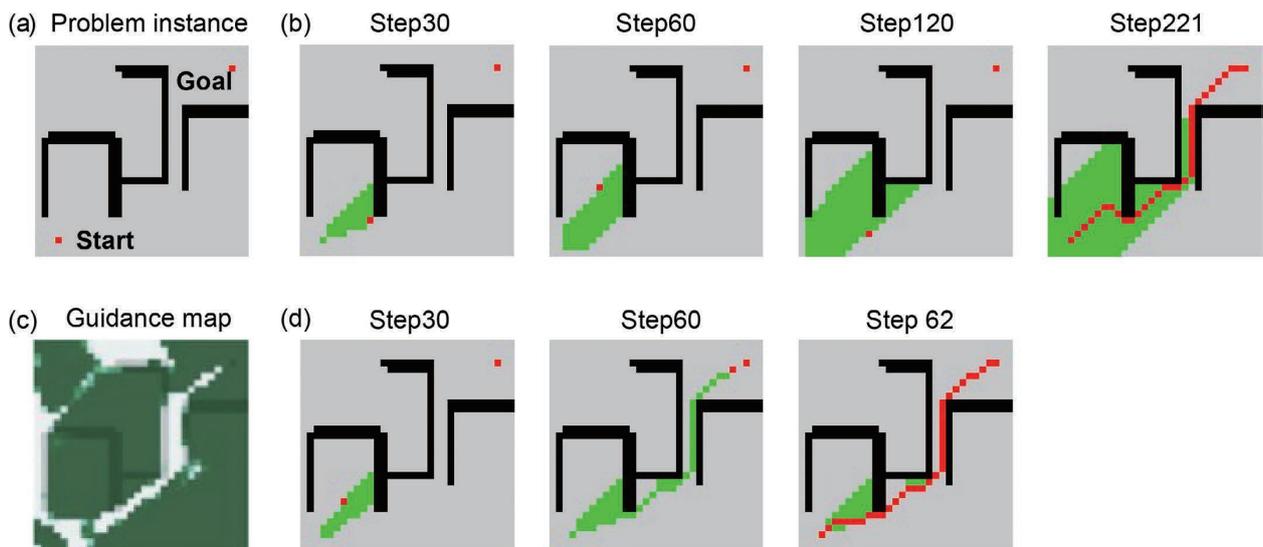


図 2 最短経路問題における A\* と Neural A\* の動作比較

(a) 経路計画の問題インスタンスについて、(b) 通常の A\* 探索の動作過程および、(d) Neural A\* の動作過程を示す。各マップにおいて、黒は障害物、赤は開始・目標点や探索点、または探索結果を示し、緑は探索済みの領域を表す。Neural A\* は問題インスタンスを (d) ガイドマップに変換して探索を行う。ガイドマップ中の白と緑は、優先的に探索すべき低コスト領域および回避すべき高コスト領域を表す。

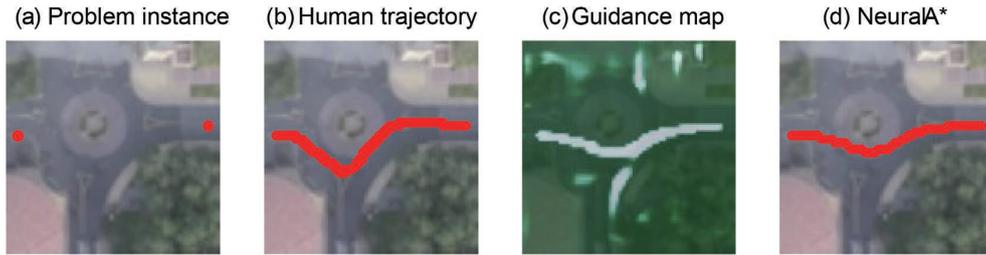


図3 生画像に対する経路計画

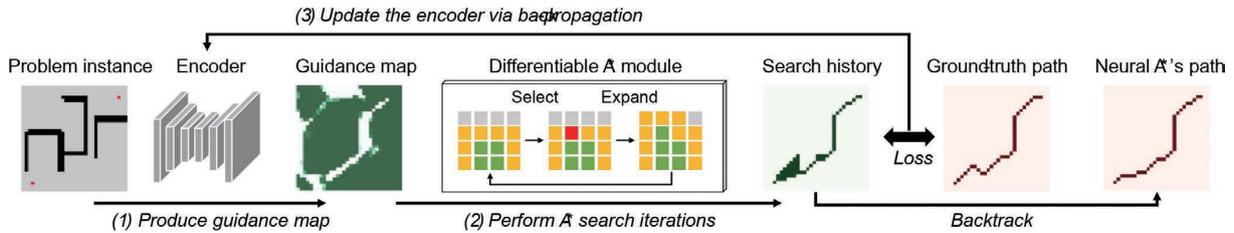


図4 Neural A\*のネットワーク構成

これら2つの問題は、一見異なるようであるが、本質的には、環境マップ中の行き止まりや障害物などの視覚的手がかりを探索アルゴリズムが認識・利用できていないという問題として、両者をひと括りに捉えることができる。そこで Neural A\* では、それぞれの形式の問題インスタンス (図 2a および図 3a) を畳み込みニューラルネットワーク (以下、エンコーダと呼ぶ) を介してガイドマップ (guidance map) と呼ぶ中間表現 (図 2c および図 3c) に変換することで、これら2つの問題に対して統一的なアプローチを可能にする。

ガイドマップは、環境マップ中の各地点  $v$  に対してガイドコストと呼ぶ移動コストを割り当てるもので、ここではノードからコストへのマッピング関数  $\phi(v)$  で表す。Neural A\* は、このガイドマップ上で A\* 探索を行い、ガイドコストの総和が最小となる経路を探索する。ガイドマップの役割は、その名の通り、経路探索を各問題の目的に応じて適切に誘導することである。具体的には、障害物位置を既知とした環境マップ上での最短経路問題では、ガイドマップはマップ中の“行き止まり”や“通行路”などの視覚的手がかりを反映し、探索中どのノードを優先して探索すべきか、あるいは避けるべきかの優先順位をノードに与え、探索効率と最適性のトレードオフの改善を図る。一方、障害物位置を未知とした生画像上での経路計画問題では、ガイドマップは、通行可能な領域とそうでない領域を入力画像中の色や模様から認識し、それらを低コストおよび高コスト・ノードとして探索アルゴリズムに伝える。一度このようなガイドマップが学習されれば、どちらの問題もガイドマップ上での最短コスト経路計画問題として A\* 探索によって解くことができる。

### 3.2 微分可能 A\* 探索によるガイドマップの End-to-end 学習

上述のアプローチの課題は、問題インスタンスが適切なガイドマップに変換されるように、どのようにしてエンコーダを学習するかという点である。エンコーダからのガイドマップ出力に対して直接損失を与えて学習する方法も考えられるが、どのような損失を定義すれば経路計画に対して効果的なのかは自明ではない。

この問題に対し我々は、図 4 に示すような、ガイドマップ変換から A\* 探索の実行までを包括的に扱う End-to-end 学習アプローチをとる。即ち、A\* アルゴリズムを微分可能な手続きとしてニューラルネットワーク内に組み込むことで、探索結果に対する正解経路との差 (損失) の勾配が、誤差逆伝播法により、A\* アルゴリズムの一連の探索ステップを介してガイドマップ、そしてエンコーダへと逆伝播可能にする。したがって学習の際には、問題インスタンスと正解経路例の組の集合からなる学習データセットを作成し、この正解経路に探索結果が近づくようエンコーダを学習する。

なお、学習の際の損失関数として、以下のような探索履歴マップ  $C$  と正解経路マップ  $\bar{P}$  の間の平均 L1 損失 (図 4 の“Loss”参照) を用い、これを確率的勾配法により最小化する。

$$L = \|C - \bar{P}\|_1 / |V| \tag{2}$$

ここで、探索履歴マップ  $C \in \{0, 1\}^V$  は A\* アルゴリズムのクローズドリスト (2.2 節) をマップ  $V$  と同じサイズの 2 値行列で表現したもので、正解経路マップ  $\bar{P} \in \{0, 1\}^V$  は正解経路上のノードを 1、それ以外を 0 で表した正解経路行列、 $|V|$  はマップサイズを表す。正解経路  $\bar{P}$  として、障害

物位置が既知の最短経路問題においてはダイクストラ法などにより求めた最短経路を、生画像上の経路計画問題においては人間が与えた経路例などを用いる。

この損失は、次の2種類のノード選択誤りに対してペナルティを与える。1) false-negative エラー：本来  $\bar{P}$  を見つけるためには  $C$  に含まれなければいけなかったノード。2) false-positive エラー： $\bar{P}$  と比べて  $C$  の中で余計に選択されたノード。言い換えれば、この損失は探索アルゴリズムに対し、1) なるべく得られる経路を正解のものへと近づける一方、2) より少ないノード探索で実現するよう働きかける。とくに最短経路問題の場合では、正解経路  $\bar{P}$  は最短経路であるため、1) なるべく最短経路に近い解を、2) より少ないノード探索で見つけるよう学習を促すことになる。これにより Neural A\* は、単に探索効率のみを改善するのではなく、探索の最適性とトレードオフを改善するように学習する。

### 3.3 微分可能 A\* の詳細

Neural A\* の中核となる微分可能 A\* の詳細を解説する。微分可能 A\* は、順伝搬時においては、2.2 節で述べた A\* 探索の手続き（即ち、Algorithm 1 において評価関数に  $f(v) = g(v) + h(v)$  を用いたもの）を基本的に1対1の対応でそのままニューラルネットワーク上で実行し、逆伝搬時においては、一部の微分不可能な手続きの勾配を近似値で置き換えながら誤差逆伝播法を適用する。

微分可能 A\* では、Algorithm 1 中の各変数や関数をマップと同じサイズの行列で表し、各演算を行列に対する基本的な演算の組み合わせで表現する。具体的には、開始点  $v_s$  や目標点  $v_g$  などの単一ノードの位置は、該当ノードで1、それ以外で0をとる2値の one-hot 行列  $V_s, V_g \in \{0, 1\}^V$  で表す。ガイドマップ  $\phi(v)$  や累積コスト  $g(v)$ 、ヒューリスティック関数  $h(v)$  などの環境ノードに対するマッピング関数は、実数行列  $\Phi, G, G \in \mathbf{R}^V$  で表す。また、クローズドリスト  $C$  と同様に、オープンリストも2値行列  $O \in \{0, 1\}^V$  で表す。また、入力マップは変数  $X$  で表し、最短経路問題では、 $X \in \{0, 1\}^V$  は通行可ノードで1、障害物ノードで0をとる2値行列、生画像上の経路探索問題では、 $X \in \mathbf{R}^{3*V}$  は実数値のカラー画像とする。各問題インスタンスは、マップ  $X$  と開始点  $v_s$  および目標点  $v_g$  の組  $(X, V_s, V_g)$  で表される。

Algorithm 1 を微分可能な手続きとして書き直すにあたり、最も重要なステップはノード選択（Algorithm 1 の3行目）である。これは評価関数  $f$  に対する離散的な操作、arg-min を含むため微分可能ではない。この操作に対して適切な近似勾配を与えるため、評価関数  $f(v) = g(v) + h(v)$  を行列変数  $F = G + H$  で表し、ノード選択を以下の式で行う。

$$V^* = I_{\max} \left( \frac{\exp(-F/\tau) \odot O}{\exp(-F/\tau), O} \right) \quad (3)$$

ここで関数  $I_{\max}(A)$  は、順伝搬時では行列  $A$  の最大値要素  $\operatorname{argmax}(A)$  を2値の one-hot 行列として返し、逆伝搬時には恒等写像  $I_{\max}(A) = A$  として振る舞うよう定義する。上式での行列  $A$  に相当する部分は、基本的に、負のコスト  $(-F)$  に対する softmax 関数を温度パラメータ  $\tau$  付きで計算するもので、オープンリスト行列  $O$  は、ノードがオープンリストから選ばれるように制限するための2値マスクとして機能する。言い換えれば、式(3)は順伝搬時には負のコストに対するマスク付きの hardmax 関数として動作し、逆伝搬時には、hardmax 関数があたかも softmax 関数であったかのように動作する。これは、本来勾配が常に零になってしまう離散化された活性化関数に対して、離散関数を“ソフト化”して擬似的な勾配を与えるという Hubara et al.<sup>10)</sup> によるアイデアに基づく。

選択ノード  $V^*$  が求めれば、オープンリスト  $O$  およびクローズドリスト  $C$  の更新（4行目）は、単純な行列同士の加減算  $O \leftarrow O - V^*$  および  $C \leftarrow C + V^*$  により行える。なお、 $V^*$  は常に one-hot 行列で、1回の探索中に同じノードが複数回選択されることはないため、 $O$  および  $C$  はアルゴリズムの実行中は常に2値行列として保たれる。

選択ノード  $V^*$  の近傍を抽出する工程（5行目）も、一般的な微分可能操作の組み合わせにより、以下のような2値行列で表現できる。

$$V_{\text{nbr}} = (V^* \star K) \odot (1-O) \odot (1-C) \odot X \quad (4)$$

ここで、 $(V^* \star K)$  は選択ノード行列  $V^*$  と固定カーネル  $K = \begin{bmatrix} [1, 1, 1]^T, [1, 0, 1]^T, [1, 1, 1]^T \end{bmatrix}$  との2次元畳み込み演算で、その後の2値行列  $(1-O)$  および  $(1-C)$  との積は抽出ノードを限定するマスク処理とみなせる。なお、障害物位置が既知で2値行列  $X$  により表されるとき、ノード選択を通行可能領域に限定するため  $X$  をさらに乗じる。この処理は、 $X$  が生画像である場合は省略する。

その後の近傍ノードのオープンリストへの追加（6行目）も、 $O \leftarrow V_{\text{nbr}} + O$  で行える。なお、7~9行目の親子関係の記録は、損失計算に関わる  $C$  の更新には影響しないため、微分可能にする必要はない。

これらの処理を目標点を選択されるまで繰り返すと、探索完了となり、それまでに選択されたノードの履歴が2値行列  $C$  に蓄積される。 $C$  に対して3.3節で述べた損失を計算することで、その勾配を微分可能 A\* の入力変数であるガイドマップ  $\Phi$  に逆伝搬し、エンコーダを学習することが可能になる。推論時には、通常の A\* 探索同様に親子関係のツリーを辿る（11行目）ことで発見された経路を出力する。

### 3.4 その他の実装の詳細

以上が単一問題インスタンスに対する微分可能 A\* の基本動作である。実際には、上記内容に加え、学習の効率化

や安定化などのための実装上の工夫や詳細があるので、最後にこれらについて軽く補足する。より完全な情報については論文<sup>9)</sup>を参照されたい。

**累積コストの更新** 2.2 節の説明および Algorithm 1 では詳細を割愛したが、ノード選択の評価関数  $f$  が累積コスト  $g$  を含む場合、 $g$  も探索中に逐次更新する必要がある。これは、基本的には毎回のノード選択後に、その近傍ノード  $v'$  に対する累積コスト  $g(v')$  の値を、選択ノードまでの累積コスト  $g(v^*)$  と選択ノードから近傍ノードまでの移動コスト  $\phi(v')$  の和で更新することになる。ただし厳密には、ここでの  $v^*$  の近傍は、初めて開かれる近傍ノード  $V_{nbr}$  とすでにオープンリストに存在する近傍ノード  $\bar{V}_{nbr}$  に場合分けされ、後者については前回計算された累積コストと新たな経路による累積コストを比べて小さい方を選ぶといった処理が必要になる。このような処理も基本的な行列操作の組み合わせで表現可能であるが、式がやや煩雑になるのでここでは割愛する。

**バッチ学習** ニューラルネットワークの学習においては、複数の入力サンプル  $\left\{ \left( X^{(i)}, V_s^{(i)}, V_g^{(i)} \right) \right\}_{i=1}^B$  に対するバッチ平均勾配を用いて学習を効率化・安定化させることが大事である。しかし、経路探索は各問題インスタンスによって探索が終了するタイミングが異なるため、このようなバッチ処理には工夫が必要である。バッチ処理版の微分可能 A\* では、ノード選択後の各リスト更新 (4 行目) において、バッチ中の各問題インスタンスに対する終了フラグ変数  $\eta^{(i)} = 1 - \left( V_g^{(i)}, V^{*(i)} \right)$  を計算し、 $O^{(i)}$  および  $C^{(i)}$  を以下のように更新する。

$$O^{(i)} \leftarrow O^{(i)} - \eta^{(i)} V^{*(i)}, C^{(i)} \leftarrow C^{(i)} + \eta^{(i)} V^{*(i)} \quad (5)$$

これは直感的には、ゴール  $V_g^{(i)}$  が選択された問題インスタンスについては、それ以降、 $O^{(i)}$  および  $C^{(i)}$  の更新を停止するように機能する。これにより、各探索の進捗状況に関係なく、全てのバッチインスタンスについて同じ行列演算を同時に実行でき、全てのバッチインスタンスで目標点が発見された時点で探索ループ終了とすることができる。

**エンコーダ設計** エンコーダとなるニューラルネットワークには、セマンティックセグメンテーションなどで用いられる U-Net<sup>3)</sup> を用い、ガイドマップの出力解像度が入力マップの解像度と同じになるようにした。エンコーダの入力には、 $X$  と  $V_s + V_g$  を連結した 2 次元画像を用いた。ガイドマップの変換を開始・目標点により条件付けることで、行き止まり内に目標点がある場合などでも適切なガイドマップの学習ができるようになる。

**一部勾配の無効化** 実際の学習では、式(3)の  $O$  や  $G$  の計

算内での一部の変数を計算グラフから切り離して定数として扱うことで、再帰的な逆伝搬の計算チェーンを単純化し、学習の安定化や消費メモリの削減を行った。

## 4. 評価実験

本節では、経路探索問題における Neural A\* の探索性能の評価実験について述べる。

### 4.1 ベースライン手法

Neural A\* の性能評価のため、以下の 2 つの経路探索手法と比較した。これらはいずれも機械学習を活用した手法で、かつ、内部で Neural A\* と同様の A\* 探索を用いた探索ベースの手法である。

- SAIL<sup>1)</sup>: ヒューリスティック関数を事例データから学習することで高い探索効率を達成した最良優先探索手法。本実験では、学習サンプルを学習したヒューリスティック関数から導く場合 (SAIL) と、理想的なオラクルから導く場合 (SAIL-SL) の 2 手法を試した。
- BB-A\*<sup>4)</sup>: A\* などの組合せ最適化アルゴリズムに対して、ブラックボックス微分と呼ばれる汎用的な近似勾配計算法を適用して学習。この手法は提案手法と似ている面があるが、A\* 探索を完全なブラックボックスとして扱う点で異なる。

また、非学習型の一般的な探索手法として、最良優先探索 (BF) と重み付き A\* (WA\*) とも比較した。さらに提案手法の変形として、式(3)で累積コストを無視して常に  $F = \Phi + H$  を用いることで、最良優先探索のような挙動をする Neural BF とも比較した。

### 4.2 評価指標

探索性能の定量的な評価指標として、以下の 3 つを用いた。

- Opt: 手法がどれくらいの頻度で最適 (最短) な経路を見つけられたかを 0-100 (%) の値で表す。
- Exp: 通常の A\* 探索と比べて削減できた探索ステップ数を 0-100 (%) の値で表す。
- Hmean: Opt と Exp の調和平均。これは探索の効率と最適性のトレードオフがどれだけ改善したかを示すもので、本研究の主たる評価指標として用いる。

### 4.3 最短経路問題における性能評価

評価実験では、Motion Planning (MP)<sup>11)</sup>、Tiled MP、City/Street Map (CSM)<sup>12)</sup> の 3 つのデータセットを用いた。MP は 32×32 サイズの人工的なマップによる経路探索データセットで、Tiled MP では MP のマップをランダムに 2×2 枚並べて 64×64 サイズに拡張した。CSM は実際の都市

の2値画像マップを64×64サイズに切り出したものである。それぞれ1,000個以上のマップデータを学習用、検証用、テスト用として80%、10%、10%に分割して用いた。各データセットにおけるOpt、Exp、Hmeanスコアの比

較を表1に示す。Neural A\*が最も高いHmeanスコアを達成し、探索の効率と最適性のトレードオフ改善効果が最も大きかったことがわかる。MPデータセット中のマップに対する実際の探索結果の例を図5に示す。

表1 最短経路問題における定量評価  
各評価手法に対して、経路最適性 (Opt) とノード探索の削減率 (Exp)、そしてそれらの調和平均 (Hmean) に対するブートストラップ平均および95%信頼区間。

MP Dataset						
	Opt		Exp		Hmean	
BF	65.8	(63.8, 68.0)	44.1	(42.8, 45.5)	44.8	(43.4, 46.3)
WA*	68.4	(66.5, 70.4)	35.8	(34.5, 37.1)	40.4	(39.0, 41.8)
SAIL	34.6	(32.1, 37.0)	48.6	(47.2, 50.2)	26.3	(24.6, 28.0)
SAIL-SL	37.2	(34.8, 39.5)	46.3	(44.8, 47.8)	28.3	(26.6, 29.9)
BB-A*	62.7	(60.6, 64.9)	42.0	(40.6, 43.4)	42.1	(40.5, 43.6)
Neural BF	75.5	(73.8, 77.1)	45.9	(44.6, 47.2)	52.0	(50.7, 53.4)
Neural A*	87.7	(86.6, 88.9)	40.1	(38.9, 41.3)	52.0	(50.7, 53.3)

Tile MP Dataset						
	Opt		Exp		Hmean	
BF	32.3	(30.0, 34.6)	58.9	(57.1, 60.8)	34.0	(32.1, 36.0)
WA*	35.3	(32.9, 37.7)	52.6	(50.8, 54.5)	34.3	(32.5, 36.1)
SAIL	5.3	(4.3, 6.1)	58.4	(56.6, 60.3)	7.5	(6.3, 8.6)
SAIL-SL	6.6	(5.6, 7.6)	54.6	(52.7, 56.5)	9.1	(7.9, 10.3)
BB-A*	31.2	(28.8, 33.5)	52.0	(50.2, 53.9)	31.1	(29.2, 33.0)
Neural BF	43.7	(41.4, 46.1)	61.5	(59.7, 63.3)	44.4	(42.5, 46.2)
Neural A*	63.0	(60.7, 65.2)	55.8	(54.1, 57.5)	54.2	(52.6, 55.8)

CSM Dataset						
	Opt		Exp		Hmean	
BF	54.4	(51.8, 57.0)	39.9	(37.6, 42.2)	35.7	(33.9, 37.6)
WA*	55.7	(53.1, 58.3)	37.1	(34.8, 39.3)	34.4	(32.6, 36.3)
SAIL	20.6	(18.6, 22.6)	41.0	(38.8, 43.3)	18.3	(16.7, 19.9)
SAIL-SL	21.4	(19.4, 23.3)	39.3	(37.1, 41.6)	17.6	(16.1, 19.1)
BB-A*	54.4	(51.8, 57.1)	40.0	(37.7, 42.3)	35.6	(33.8, 37.4)
Neural BF	60.9	(58.5, 63.3)	42.1	(39.8, 44.3)	40.6	(38.7, 42.6)
Neural A*	73.5	(71.5, 75.5)	37.6	(35.5, 39.7)	43.6	(41.7, 45.5)

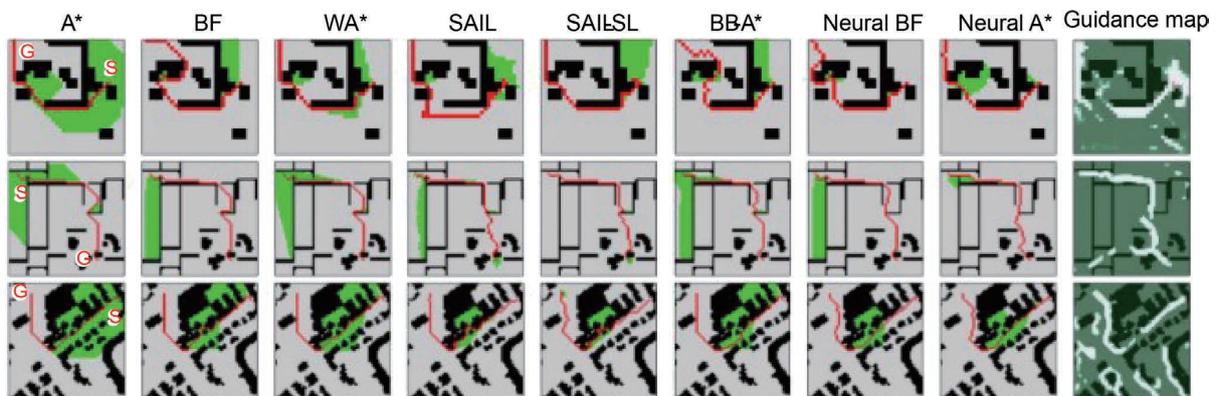


図5 経路探索結果例の比較

### 5. 生画像マップに対する経路計画

障害物位置のラベル付けがない生画像マップに対する経路計画問題について、Neural A\* と BB-A\* の性能比較実験を行った。実験では、Stanford Drone (SD) データセット<sup>13)</sup>に含まれる屋外シーンの歩行者軌跡データを用いて、歩行者の経路予測問題のデータセットを作成した。そして、シーン画像と開始・目標点が与えられた際に、各手法がどれだけ正確に実際の歩行者経路を予測できるか評価した。

表2では予測経路と正解経路の間の chamfer 距離を比較している。Intra-scenes では、学習データとテストデータに同じ撮影場所の画像を用いることで、未知の歩行者に対する予測精度を測る。Inter-scenes では、学習データとテストデータで異なる撮影場所の画像を用いる leave-one-out 交差検定を行い、未知の場所に対する汎化性能を測る。どちらの場合においても Neural A\* が、より正解に近い経路を予測していることがわかる。

実際の結果の例を図6に示す。最初2つの例では、Neural A\* が定性的にも正解に近い経路を予測していることがわかる。一方、3つ目の例では、開始・目標点の間に複数の妥当な経路があるため、予測に失敗している。このような場合に対する拡張として、複数の経路例を確率的にサンプリングできるようにする生成的アプローチ<sup>14)</sup>などが考えられる。

表2 SD データセットにおける定量評価  
予測された経路と正解経路の間の chamfer 距離に対するブートストラップ平均と 95%信頼区間。

	Intra-scenes	Inter-scenes
BB-A*	152.2 (144.9, 159.1)	134.3 (132.1, 136.4)
Neural A*	16.1 ( 13.2, 18.8)	37.4 ( 35.8, 39.0)

### 6. むすび

本論文では、「機械学習を活用した経路計画手法」と題して、この分野における最新の研究動向および、我々の研究成果である Neural A\* について解説した。同分野は様々な問題意識のもと研究がされているが、そのなかでも、経路探索の最適性と効率性のトレードオフ向上と、障害物位置を未知とした生画像マップに対する経路計画という2つの問題に対して、Neural A\* は統一的な解決アプローチを提案した。評価実験では、従来の機械学習を用いない経路探索アルゴリズムと比べ、Neural A\* は探索の最適性を大きく損なうことなく効率性を飛躍的に向上できることが確かめられた。また、固定カメラ画像に対する歩行者の経路予測問題では、実際の歩行者の移動経路を従来手法より正確に予測できることが確かめられた。同技術は、ロボットマニピュレータの高次元状態空間における動作計画をはじめ、様々なシーンへの適用が期待される。

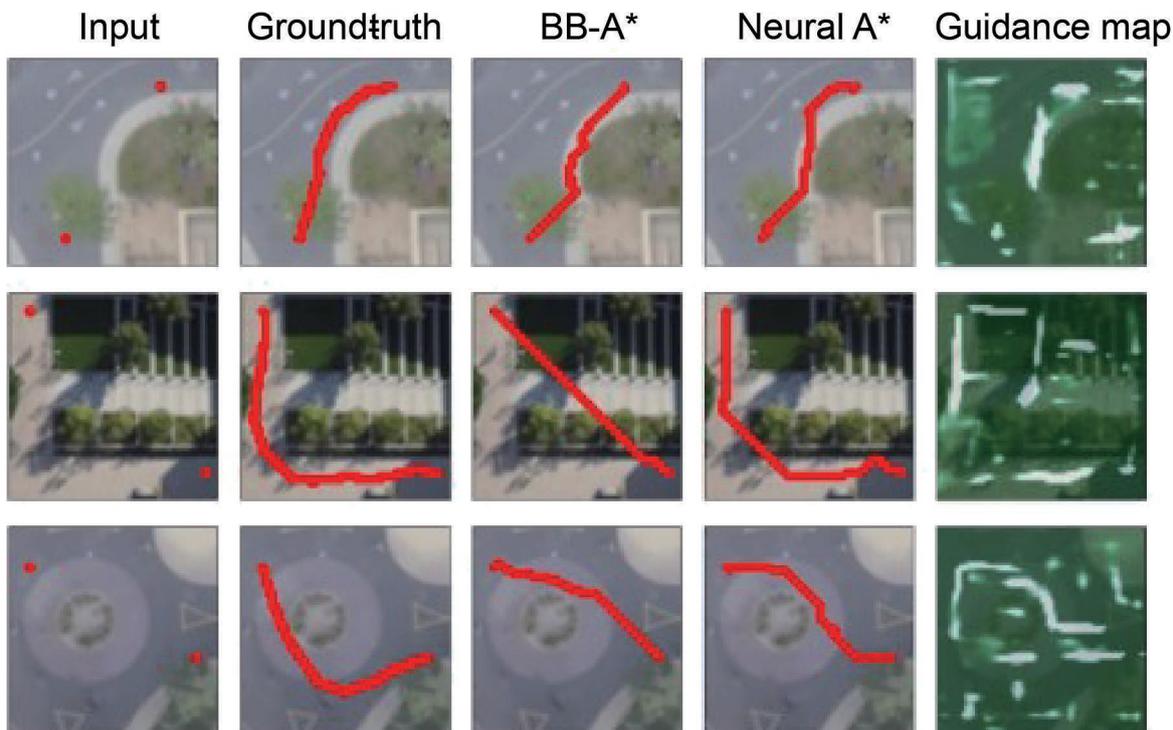


図6 SSD データセットにおける経路計画結果の例

## 参考文献

- 1) Choudhury, S.; Bhardwaj, M.; Arora, S.; Kapoor, A.; Ranade, G.; Scherer, S.; Dey, D. Data-driven planning via imitation learning. *International Journal of Robotics Research (IJRR)*. 2018, Vol.37, No.13-14, p.1632-1672.
- 2) Takahashi, T.; Sun, H.; Tian, D.; Wang, Y. "Learning heuristic functions for mobile robot path planning using deep neural networks". *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS)*. 2019, p.764-772.
- 3) Ronneberger, O.; Fischer, P.; Brox, T. "U-net: Convolutional networks for biomedical image segmentation". *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*. 2015, p.234-241.
- 4) Vlastelica, M.; Paulus, A.; Musil, V.; Martius, G.; Rolinek, M. "Differentiation of blackbox combinatorial solvers". *Proceedings of the International Conference on Learning Representations (ICLR)*. 2020, p.1-19.
- 5) Ichter, B.; Harrison, J.; Pavone, M. "Learning sampling distributions for robot motion planning". *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2018, p.7087-7094.
- 6) Ichter, B.; Schmerling, E.; Lee, T. W. E.; Faust, A. "Learned critical probabilistic roadmaps for robotic motion planning". *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2020, p.9535-9541.
- 7) Qureshi, A. H.; Simeonov, A.; Bency, M. J.; Yip, M. C. "Motion planning networks". *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2019, p.2118-2124.
- 8) Chen, B.; Dai, B.; Lin, Q.; Ye, G.; Liu, H.; Song, L. "Learning to plan in high dimensions via neural exploration-exploitation trees". *Proceedings of the International Conference on Learning Representations (ICLR)*. 2020, p.1-15.
- 9) Yonetani, R.; Taniai, T.; Barekatin, M.; Nishimura, M.; Kanazaki, A. "Path planning using neural A\* search". *Proceedings of International Conference on Machine Learning (ICML)*. 2021, p.12029-12039.
- 10) Hubara, I.; Courbariaux, M.; Soudry, D.; El-Yaniv, R.; Bengio, Y. "Binarized neural networks". *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*. 2016, p.4107-4115.
- 11) Bhardwaj, M.; Choudhury, S.; Scherer, S. "Learning heuristic search via imitation". *Proceedings of the Conference on Robot Learning (CoRL)*. 2017, p.271-280.
- 12) Sturtevant, N. Benchmarks for grid-based pathfinding. *IEEE Transactions on Computational Intelligence and AI in Game*. 2012, Vol.4, No.2, p.144-148.
- 13) Robicquet, A.; Sadeghian, A.; Alahi, A.; Savarese, S. "Learning social etiquette: Human trajectory understanding in crowded scenes". *Proceedings of the European Conference on Computer Vision (ECCV)*. 2016, p.549-565.
- 14) Salzman, T.; Ivanovic, B.; Chakravarty, P.; Pavone, M. "Trajectron++: Multi-agent generative trajectory forecasting with heterogeneous data". *Proceedings of the European Conference on Computer Vision (ECCV)*. 2020, p.1-17.

## 執筆者紹介



谷合 竜典 TANIAI Tatsunori  
 オムロン サイニックエックス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門：コンピュータビジョン、3次元計測、密対応点推定  
 所属学会：情報処理学会、IEEE  
 博士（情報理工学）



米谷 竜 YONETANI Ryo  
 オムロン サイニックエックス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門：コンピュータビジョン、機械学習  
 所属学会：情報処理学会、電子情報通信学会、IEEE  
 博士（情報学）

---

本文に掲載の商品の名称は、各社が商標としている場合があります。

# A Machine Learning System that Adaptively Aggregates Knowledge from Multiple Models

MA Jiaxin

As a machine learning framework, decentralized learning aims to address the difficulties of data collection and annotation by breaking down and assigning these tasks to a group of clients to utilize their own data resources. Federated learning is a conventional approach of decentralized learning, but it is not suitable to deal with cases when the client model architectures or data distributions are diverse. This article introduces one of our published research results, which is a method called Decentralized Learning via Adaptive Distillation (DLAD). As a method based on knowledge distillation, it learns a model by aggregating and imitating the client models' outputs, without requiring identical client-model architecture. Moreover, this method casts adaptive aggregation weights to different clients, to give priorities to learn from client models with higher confidence. This approach is especially useful for the non-IID (Independent and Identically Distributed) data. We have conducted evaluation experiments on multiple public datasets and demonstrated the effectiveness of this method.

## 複数のモデルから適応的に知識を統合する 新たな機械学習スキーム

馬 家昕

非集中学習は、データ収集やアノテーションのコストの高さという課題に対し、これらのタスクを分解してクライアント（ローカル）に割り当てることで、クライアント独自のデータリソースを活用する機械学習の枠組みである。連合学習は従来からの非集中学習の方策の一つだが、各クライアントのモデル構造やデータ分布が多様な場合には適さない。本稿では、適応的蒸留による非集中学習（Decentralized Learning via Adaptive Distillation: DLAD）という手法を解説する。この手法では、知識蒸留に基づいてクライアントモデルの出力を集約し、模倣することにより、不均一なクライアントモデルに対する非集中学習を可能にする。この際、各クライアントのデータが独立同分布とならない状況にも対応するため、適応的に学習の重みを求める。我々は、多数の公開データセットによって評価実験を行い、提案手法の効果を確認した。

### 1. Introduction

The content of this article is based on our recent paper named “Adaptive Distillation for Decentralized Learning from Heterogeneous Clients”<sup>1)</sup>. In this article, the author would like to provide a reader-friendly explanation of the original paper. The purpose is not only to explain the details of technology, but more importantly, to give readers an insight that how to use the related technology to solve real-world problems, since knowledge deserves to be understood and utilized, to contribute to the business, and the world.

In current days, machine learning is undoubtedly a promising

technology not only in academic research but also in enterprises. The most common difficulty to deploy machine learning in a real-world project is related to **data**, which is usually two-fold. First, deploying machine learning can be costly due to data collection. Some sensitive data, such as life logging videos, and medical data, that their owners wish to keep private, are hardly accessible. Second, deploying machine learning can be costly due to data annotation. Supervised machine learning (the most commonly used machine learning algorithm) requires that to learn a model, training data must be annotated with ground truth labels. Depending on the difficulties of tasks, annotation sometimes needs certain levels of professional know-how (e.g., a doctor's diagnosis), and thus can be extremely expensive.

Contact : MA Jiaxin jiaxin.ma@sinix.com

So, is there a method to alleviate the cost of data collection or data annotation? One promising solution is **Decentralized Learning**, which means to put the data collection and data annotation processes on the client side. Here, “client” usually means some institutions, companies, or end-users who are supposed to conduct data collection and data annotation as their daily behaviors (for comparison, we call the other side “server side”). For example, Google developed a machine learning model to predict the next word of the keyboard input<sup>2)</sup>, and the data collection and data annotation are all performed by smartphone users’ daily keyboard input. It is important to note that, during the above process, end-users train their own prediction models on their devices, and only transfer **model weights** with Google, but not any sensitive data they have input using their smartphone keyboards. This decentralized learning framework is called Federated Learning (FL)<sup>3,4)</sup>.

To better understand the advantages of decentralized learning, we give another example. Let us assume that OMRON is about to develop a new cardiac diagnostic device. It can help early detection of heart diseases by reading and analyzing users’ vital signs. Usually, the development of such a device would face a high hurdle because vital signs are sensitive and private data, and annotations need expertise from doctors. With FL, OMRON just needs to deploy the copies of their machine learning model to different hospitals, and get the models trained with the daily medical data. At the hospital (client) side, operators only need to input patient data (vital signs, profiles, etc.) and corresponding diagnosis into the model. Since the model neither uploads any private data to the OMRON side nor requires doctors to make additional diagnoses beyond their daily work, the difficulty of development is greatly reduced.

From the above example, FL is a promising decentralized learning framework and should be encouraged to use in practice. However, some limitations of FL still exist, for example:

1. FL requires client model architectures to be identical. Usually, it is applicable to just deploy the same model to all the clients. However, this requirement is inconvenient under some practical circumstances, such as that, clients may have needs of model customization (due to limitation of computational resources, privacy policies, performance bias, etc.); clients may already have their own trained models and we want to directly use them, and so on.
2. FL requires frequent data communication during the model training process. Although such data communication does not involve private data, the data communication itself also brings limitations and concerns, such as network qualities, securities, and so on. Also, if different clients have different

communication conditions, it will be a problem to balance the training process among all the clients. Extra efforts are needed to improve the communication efficiency<sup>5,6)</sup>.

In this article, the author would like to introduce another decentralized learning framework that is based on knowledge distillation (KD). It can solve the above issues that FL is not good at. Moreover, the proposed method uses an improved weight aggregation strategy to deal with the non-IIDness problem, which will be explained later. One should note that the KD-based decentralized learning framework is not necessarily superior to FL-based. One should be able to identify which framework is more suitable for their practical applications.

## 2. Method

### 2.1 What is knowledge distillation?

Knowledge distillation<sup>7)</sup> is a method that allows one trained model (the source model) to “teach” another new model (the target model). In other words, it allows a new model to imitate the output of an existed model, without significant loss of performance. The original purpose of KD is mainly on model compression, which means that, usually, the source model is a large (deep) model, and the target model is a small (shallow) model which is less expensive to be deployed in practice.

The reason a small model can approximate a large model in its performance is that, firstly, a large model usually has some excessive capacity or power which is not fully utilized; secondly and more importantly, a target model can benefit from learning “soft labels” from a source model. We will use an example to explain the latter one.

Consider an image classification task, where one of the images illustrates a cat playing with a mouse (see Fig. 1). Usually, the label (ground truth) will be “cat” only, since the cat occupies the main body of the image (it is theoretically possible but too ambiguous and inefficient to annotate this image as “cat and mouse”). In this way, the information of the mouse will be missing from the true label, and thus the machine learning model that learns from the true label will only learn “this image is showing a cat but nothing else”, which is actually not ideal.

On the contrary, in a KD scenario, the source model provides “soft labels” rather than true labels. Assume that the source model is well-trained (i.e., it can at least recognize cat and mouse precisely). Then in this case the model may produce a classification output like “70% cat, 30% mouse”. Unlike a true label that only represents one possibility, a soft label will represent all the possibilities in ratios. As a result, it can handle the cases such as “A and B are in the same image” or “this object looks like both A and B”, where the true label cannot. In KD, while the target model learns from soft labels (and from true

labels, at the same time), it has been proved that it can perform better compared with only learning from true labels.

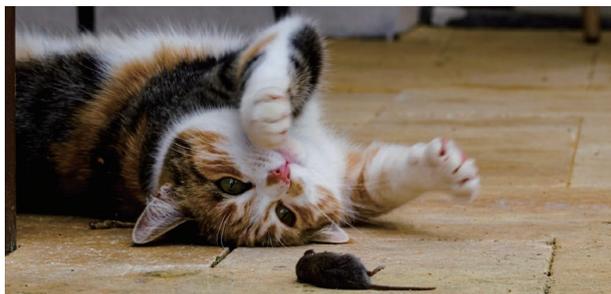


Fig. 1 Should a machine learning model classify this image as a cat or a mouse?

## 2.2 How does knowledge distillation benefit decentralized learning?

As mentioned previously, the original usage of KD is mainly about model compression, but KD can also benefit decentralized learning. In a KD-based decentralized learning framework, the client side owns source models, and the server side owns a target model. The client side trains source models with their private data and annotations, while the server side needs to collect its own data and input them to source models to get output (soft labels). After that, the server side trains the target model with its own data, and the corresponding soft labels which are aggregated from all the clients.

KD-based frameworks do not have the limitations of FL-based frameworks mentioned in the previous section. First, KD does not aggregate model weights but soft labels, so there is no need to keep model architectures identical. Any client or server can have a unique model. Second, there is no frequent data communication during the training process. Actually, the data communication only occurs twice: once for the server sending data to the clients, and the other for the clients sending soft labels to the server. The training processes (both on the client side and the server side) can be totally off-line.

KD also has its own limitations. During the above process, there is no data transfer from the client side to the server side, so the data privacy of the clients is protected, however, there is data transfer from the server side to the client side, which means the server side still need to collect enough data. Also, there is no need to annotate the server-side data by human labor, instead, the annotation is done by client-side models. Asking the clients to run their models may still incur costs. Compared with FL, requiring data collection may be a main limitation of KD, however, for some types of tasks, it is not so difficult to get unlabeled data, while sometimes the model heterogeneity can be a critical advantage.

## 2.3 How does our work differ from traditional knowledge distillation?

When we distill knowledge from multiple sources, there is no guarantee that all the sources provide outputs of the same quality. Different sources are likely to have different confidences towards different categories of samples. Here, the confidence may be due to many factors, for example, model architecture, annotation qualities, number of training data samples, and so on. Among them, the number of training data samples (on certain categories) is a very common factor caused by data distribution.

Traditional decentralized learnings, both FL-based and KD-based, aggregate the output (model weights or soft labels) evenly from multiple clients. It means that it does not distinguish which client gives a more trustable output and train with a bias accordingly. It is fine for IID (Independent and Identically Distributed) data. But in most cases, real-world data are non-IID. For example, patient data distribution will be diverse depending on regions, seasons, hospital categories, and so on. For the case of non-IID data, it is highly possible that some clients have never seen some categories of data samples during their training processes, and thus cannot provide confident outputs.

In our work, we proposed an improved KD method, Decentralized Learning via Adaptive Distillation (DLAD). It allows the target model to selectively learn from source models, which is an effective solution to non-IID data.

## 2.4 The details of DLAD

In our distillation process, for each data sample sent from the server to the clients, we not only expect to get an output from each client model but also want each client model can provide a confidence estimation, which represents “how confident am I to correctly classify this sample”. In our scenario, this condition is simplified to “how familiar is this sample to me” to address the issue of non-IIDness. It is rather tricky to let client models report their familiarity or confidence. Here we introduce our implementation as follows.

Fig. 2 shows the overview of the proposed DLAD framework. In the figure,  $D_1 \cdots D_N$  represents the data owned by clients,  $M_1 \cdots M_N$  (in black) represent the client-side models and orange ones are their binary copies (explained later),  $D_{dist}$  represents the data collected by the server side, and  $M$  represents the model owned by the server side.

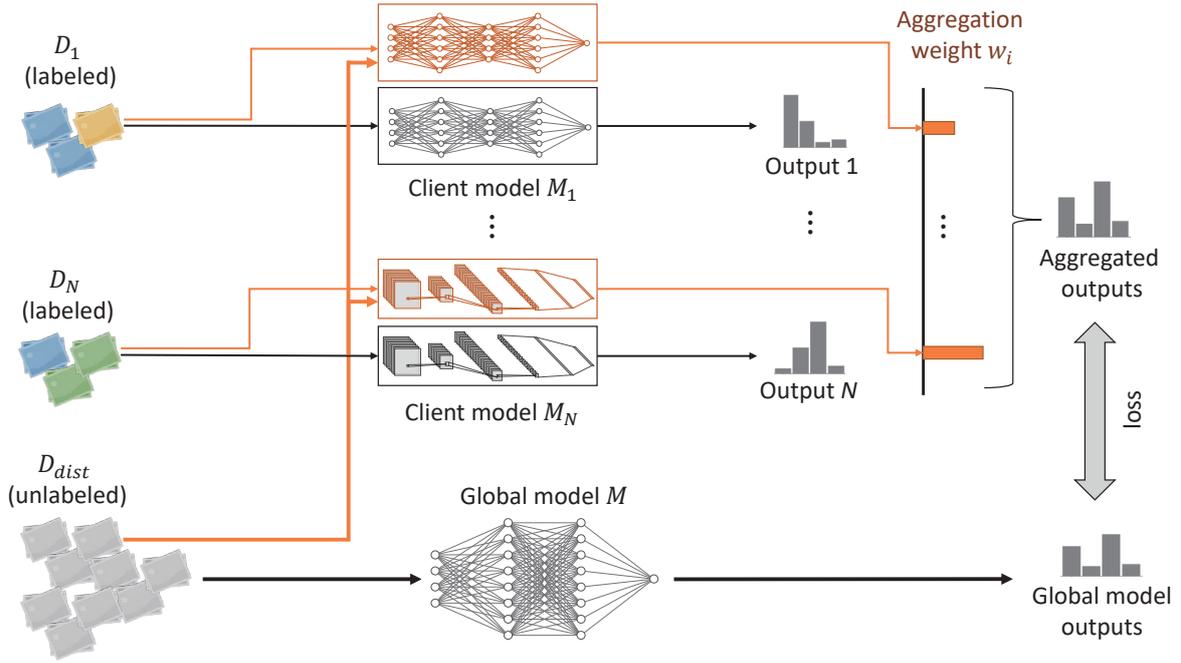


Fig. 2 The overview of DLAD

The training of DLAD has three steps.

**Step 1, training client models.** Clients train models  $M_1 \dots M_N$  with their own data  $D_1 \dots D_N$ . (The models can be either provided by the server or owned by the clients themselves.)

**Step 2, training binary models.** The server sends  $D_{dist}$  to all the clients. Each client duplicates its trained client model (including model weights) and swaps the model’s final classification layer with a binary classification layer (with sigmoid activation to ensure the output is  $[0,1]$ ). We call these new models binary models ( $M_{b1} \dots M_{bN}$ ). Then the binary models are trained with both  $D_{dist}$  and  $D_i$  ( $i = 1 \dots N$ ), while  $D_{dist}$  will have labels of 0s, and  $D_i$  will have labels of 1s.

**Step 3, training the server model.** The clients run their models (both  $M_i$  and  $M_{bi}$ ) with  $D_{dist}$  and send the outputs to the server (where the output of  $M_i$  becomes soft labels, and the output of  $M_{bi}$  becomes aggregation weights). The server aggregates these outputs (into aggregated soft labels) and then uses these aggregated outputs and  $D_{dist}$  to train model  $M$ .

We can see that the data communication between the server and the clients only occurs at the beginning of each step. The model training processes can be totally off-line.

The most important part of DLAD is how to design the confidence of each model towards a certain sample, as well as the aggregation method. In our implementation, we define the confidence of client  $i$  towards sample  $x$  as  $C_i(x) = M_{bi}(x)$ . The concept of designing  $C_i(x)$  is that  $C_i(x)$  should become larger if the

model  $M_{bi}$  recognizes sample  $x$  similar to its own data  $D_i$  and becomes smaller otherwise.

Then, we aggregate  $C_i(x)$  from all the clients, to calculate a confidence weight for each client, which is

$$w_i(x) = \frac{\exp C_i \left( \frac{x}{T} \right)}{\sum_j \exp C_j \left( \frac{x}{T} \right)} \quad (1)$$

$w_i(x)$  needs to be normalized across all the clients to ensure that it has the same scale. The above equation is equivalent to a softmax normalization, with a hyper-parameter  $T$  that adjusts the smoothness of output. Then the final aggregation result which is also the label to train the model  $M$  would be

$$\sum_i w_i(x) M_i(x) \quad (2)$$

There are also some limitations of the design of DLAD. First, it requires each client to additionally train a model of  $M_{bi}$ . Second, it is not always true that the more overlapping between  $D_{dist}$  and  $D_i$ , the higher  $M_{bi}(x)$  is, also, since the structure of  $M_{bi}$  is inherited from  $M_i$ , it can be biased according to the difference of model architectures. Anyway, to improve this idea, for example, we may properly define a distance function to represent the distance between any new sample  $x$  and the dataset  $D_i$ , and then aggregate using Eq. 1. There are many possibilities to define this distance function. However, this exploration is not included in the current study.

### 3. Experiments

#### 3.1 Problem setting

To evaluate our method, we use image classification as our task. This is a very common task of computer vision problems. In our study, we choose three datasets, namely MNIST, CIFAR-10, and CINIC-10 for evaluation. MNIST is a handwritten digit database, which includes  $28 \times 28$  pixel grayscale images of single digits from 0 to 9. CIFAR-10 and CINIC-10 are both real-world photo databases, and both include  $32 \times 32$  color images of 10 classes. All of them are commonly used and publicly available.

Note that image classification is a very popular problem in machine learning studies. The abundance of public datasets and baselines also attracts researchers to evaluate their machine learning models on it. However, our proposed method can be applied to any classification problem but not limited to computer vision. We hope the readers can have their own idea that how to apply this method to the real-world problems that they are facing.

In order to simulate a decentralized learning environment, we divided the datasets into client-side data and server-side data. Among them, the client-side data is paired with ground truth labels, and the server-side data is unlabeled. As in the real world, unlabeled data is always much easier to get, we assign a larger part of data to server side. Specifically, for MNIST and CIFAR-10, among their training sets (60,000 and 50,000 samples, respectively) we assign 80% samples to server side and 20% samples to client side. For CINIC-10, the whole validation set (90,000 samples) is assigned to server side, and the whole training set (90,000 samples) is assigned to client side.

After that, for each client, its own data  $D_i$  is created by repeatedly and randomly sampling (allowing duplicates) from the data assigned to client side, until  $D_i$  reaches a predetermined number of data samples (which is arbitrarily determined as 6,000 for MNIST, 5,000 for CIFAR-10, and 20,000 for CINIC-10). If the data distribution is IID,  $D_i$  will include data of all ten classes. If the data distribution is non-IID,  $D_i$  will only include data of a part of classes. For simplicity,  $D_i$  is a balanced dataset. For example, if  $D_i$  includes data of six classes, the probability of each class being sampled should be  $1/6$ . In a real-world problem,  $D_i$  might be unbalanced, but it should not affect the performance of DLAD.

In order to simulate the different levels of non-IIDness existed in the real world, we defined one type of IID and three types of non-IID data distributions. They are shown in Table 1. As either MNIST, CIFAR-10, or CINIC-10 has 10 classes of data (noted as  $c_0 \sim c_9$  in Table 1), for simplicity, when we define the data distribution, we assume that the number of clients is a

multiple of five (client 1~5 have the same data distribution as client 6~10, and so on).

- IID: all ten classes are accessible to all clients.
- Non-IID #1: every two classes are exclusively accessible to only one client, e.g.,  $c_0$  and  $c_1$  are accessible to client 1;  $c_2$  and  $c_3$  are accessible to client 2; and so on.
- Non-IID #2:  $c_0 \sim c_4$  are accessible to all clients, and  $c_5, c_6, c_7, c_8, c_9$  are exclusively accessible to only one client each.
- Non-IID #3: every class is accessible to only two clients among five, e.g.,  $c_0$  is accessible to clients 1 and 2;  $c_1$  is accessible to clients 1 and 3; and so on.

Table 1 The data distribution setting in our experiment: one IID case and three non-IID cases

Client Index	5n+1	5n+2	5n+3	5n+4	5n+5
IID	$c_0 \sim c_9$				
Non-IID #1	$c_0, c_1$	$c_2, c_3$	$c_4, c_5$	$c_6, c_7$	$c_8, c_9$
Non-IID #2	$c_0 \sim c_4, c_5$	$c_0 \sim c_4, c_6$	$c_0 \sim c_4, c_7$	$c_0 \sim c_4, c_8$	$c_0 \sim c_4, c_9$
Non-IID #3	$c_0, c_1, c_2, c_3$	$c_0, c_4, c_5, c_6$	$c_1, c_4, c_7, c_8$	$c_2, c_5, c_7, c_9$	$c_3, c_6, c_8, c_9$

#### 3.2 Experiment setting

The experiment involves all the three steps of the training process that were mentioned in Section 2.4.

**Step 1, training client models.** Theoretically speaking, it is possible to assign any type of supervised machine learning model to each client, e.g., support vector machine, decision tree, and so on. In our experiment, though, we tested two deep learning models: ResNet18<sup>8)</sup> and DenseNet<sup>9)</sup>. The reason for adopting them is that both models are typical deep learning models that are usually seen in papers. We use pre-trained weights (on ImageNet) on both ResNet and DenseNet to reduce the necessary training time. Each client model  $M_i$  is trained for 50 epochs with a batch size of 250. Adam optimizer with the learning rate of 0.001 is applied (the same below).

**Step 2, training binary models.** After client models are trained, we duplicate each client model and swap the final layer with a binary-output layer to get the binary model  $M_{bi}$  and train them for 20 epochs. If the training sample is from  $D_i$ , we additionally apply a sample weight of 1.5 to alleviate the effect of data imbalance since in our problem setting  $D_i$  has much fewer samples than  $D_{dist}$  (see Section 3.1).

**Step 3, training the server model.** The server model is also chosen from ResNet and DenseNet, and its initial weights are also pre-trained weights on ImageNet. It is trained for 100 epochs. A temperature parameter  $T$  of 0.05 is used for calculat-

ing the weight aggregation as in Eq. 1.

During all the three training steps, data augmentation is applied to the input data to increase the robustness, where the following operations are used: rotation (within  $\pm 20^\circ$ ), shift in width, height, and color (within 20%), and horizontal flip.

We compare the result of standard DLAD with the other two baselines. One is **simple averaging**, where the aggregation weight  $w_i(x)$  is fixed as  $1/N$  ( $N$  is the number of clients). The other is **labeled confidence**, which calculates aggregation weight in the same manner as DLAD, but instead of letting the confidence  $C_i(x) = M_{b_i}(x)$ , it uses a ground truth label to express  $C_i(x)$ . This ground truth label is equal to the class distribution probability. For example, for the four distribution cases (IID and non-IID #1~3) in Table 1, if sample  $x$  belongs to  $c_1$ ,  $C_i(x)$  is equal to  $1/10$ ,  $1/2$ ,  $1/6$ , and  $1/4$ , respectively, when  $i = 1$ , and equal to  $1/10$ ,  $0$ ,  $1/6$ , and  $0$ , respectively, when  $i = 2$ . The former baseline is the traditional strategy used in most (even recently) decentralized learning frameworks<sup>10,11</sup>, while the latter one can be treated as DLAD with ideal values of confidence, which is also a theoretical upper bond of DLAD.

### 3.3 Experiment results

To evaluate the performance of our proposed method in various situations, we conducted three experiments. Their details are listed in Table 2. We control the variables of dataset, distribution, client model architecture, global model architecture, and the number of clients. The experimental results are discussed as follows.

Table 2 The variables of experiment 1~3

Variable	Dataset	Distribution	Client model	Global model	No. of clients
Exp1	MNIST, CIFAR-10, CINIC-10	IID, NIID 1~3	ResNet	ResNet	10
Exp2	CIFAR-10	NIID 1	ResNet, DensNet, both	ResNet, DenseNet	10
Exp3	CIFAR-10	NIID 1~3	ResNet, DensNet, both	ResNet	5, 10, 20, 30

Due to space limitations, we do not quote the complete results of experiments 1~3 in this article (they are described in the original paper). We will use Fig. 3 as an example (experiment 1 on CINIC-10) to show how the experiment results look like. In Fig. 3, the first 50 epoch is Step 1 which represents the training process of 10 client models, and the last 100 epoch is Step 3 which represents the training (knowledge distillation) process of the server model (noted as “global” in the figure). Step 2 is

unrelated to the image classification task, so it is not shown. In Step 1, the black line represents the mean validation accuracy while the gray shade represents the area between max and min accuracy of all client models. In Step 3, the red line is the accuracy of standard DLAD; the green line is the method of simple-average aggregation; the blue line is DLAD with labeled confidence (they were explained in Section 3.2).

From Fig. 3, we can find the following facts:

- For the IID case, since there is no difference in data distribution and model architecture among clients and server, the server model is unlikely to be benefited or disturbed by any decentralized learning method. Still, the sever model converges faster compared with client modes, and its final accuracy is a litter higher, which is probably due to the effect of soft labels and a larger number of training samples.
- For the three types of non-IID cases, client models’ accuracies are obviously low because they only have access to two, six, and four classes of training samples, respectively (in other words, their performance would be capped at 0.2, 0.6, and 0.4, respectively). In this situation, simple-average aggregation (green) played a limited role where it helped boost accuracies in non-IID #1 and #3 but lost accuracy in non-IID #2. In other words, simple average is not suitable for all the non-IID cases. Compared to that, DLAD (red) showed its effectiveness as well as stableness in all the three non-IID cases. With labeled confidence (blue), the performance can be further improved.

The above results have already shown the usefulness of our proposed DLAD method compared to the commonly used simple-average aggregation. From our other experiment results described in the original paper but not shown here, we have other observations as follows.

- **About datasets:** The difficulty of tasks is like  $MNIST < CIFAR-10 < CINIC-10$ , so their accuracies decrease accordingly. Anyway, DLAD showed no abnormal behavior on all three datasets.
- **About model architectures:** In about half experiments, using a combination of ResNet and DenseNet as client models gave better results in server model performance, compared with using identical architecture (ResNet or DenseNet only). Our experiments only tested two model architectures, so it might be not enough to prove that diversity in client model architectures necessarily benefits DLAD results. But still, allowing customization of client

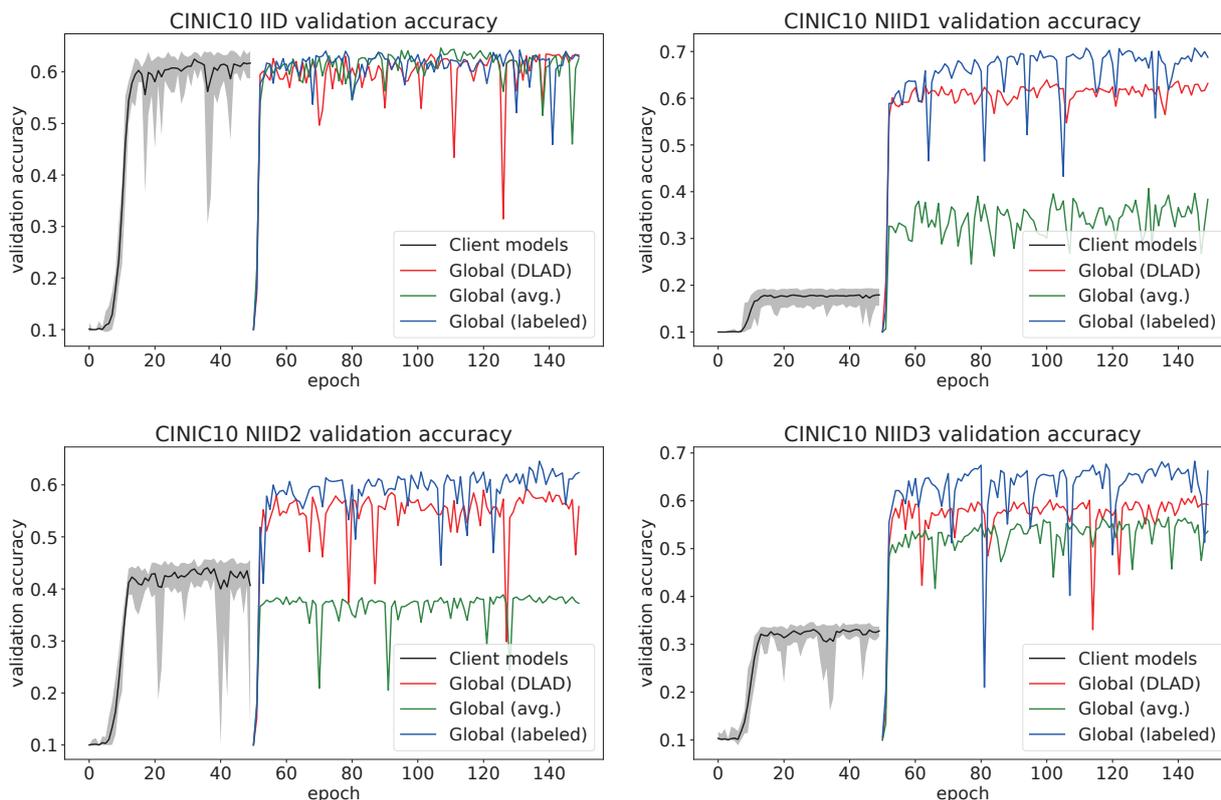


Fig. 3 The result of experiment 1 on CINIC-10

models without harming the overall performance will be a great plus for real-world problems.

- **About the number of clients:** The performance of DLAD showed a generally increasing trend with the number of clients. It indicates that DLAD is potentialized for large-scale usage.

To recap, the effectiveness of DLAD mainly attributes to the mechanism of aggregation weights. If we can precisely estimate the confidence of client models when doing aggregation, the performance of DLAD will be enhanced to approach a high level (as DLAD with confidence labels). On the other hand, if we cannot estimate the confidence due to some reasons, the performance of DLAD will be downgraded to approach the simple-average method. One of the possible reasons is the domain difference between  $D_i$  and  $D_{dist}$ . For example, assume  $D_i$  is the data of Asian patients and  $D_{dist}$  is the data of American patients, then a binary classifier can easily distinguish between the two groups no matter their samples have the same label or not. We should prevent such a case because no client model will show predominant confidence and thus the aggregation weights will not work as intended.

#### 4. Conclusion

Starting from the background of decentralized learning, this article introduced the details of DLAD, which is an original decentralized learning approach based on knowledge distillation. The article mainly answered the following questions:

- Why is decentralized learning useful?
- What are the features of federated learning and knowledge distillation?
- How does our method solve the non-IID issue?
- How is our method implemented?
- How do the experimental results of our method compare with baselines?

The author believes that decentralized learning, either FL-based or KD-based, is a very promising and applicable technology for practical use. Hopefully, this article can stimulate readers' interest and bring fresh ideas even new business chances to their domains.

#### References

1) J. Ma, R. Yonetani, and Z. Iqbal, "Adaptive distillation for decentralized learning from heterogeneous clients," in *2020 25th Int. Conf. Pattern Recognit. (ICPR)*, 2021, pp.7486-7492.

- 2) A. Hard *et al.*, “Federated learning for mobile keyboard prediction,” *arXiv*. preprint arXiv:1811.03604, 2018.
- 3) H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. “Communication-efficient learning of deep networks from decentralized data,” in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics*, 2017, pp.1273–1282.
- 4) K. Bonawitz *et al.*, “Towards federated learning at scale: System design,” *arXiv*. preprint arXiv:1902.01046, 2019.
- 5) T. Nishio and R. Yonetani, “Client selection for federated learning with heterogeneous resources in mobile edge,” in *ICC 2019 - 2019 IEEE Int. Conf. Commun.*, 2019, pp.1–7.
- 6) J. Konečný *et al.*, “Federated learning: Strategies for improving communication efficiency,” *arXiv*. preprint arXiv:1610.05492, 2016.
- 7) G. Hinton, O. Vinyals, and J. Dean, “Distilling the knowledge in a neural network,” *arXiv*. preprint arXiv:1503.02531, 2015.
- 8) K. He, X. Zhang, S. Ren, J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, 2016, pp.770–778.
- 9) G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, 2017, pp.4700–4708.
- 10) J. H. Ahn, O. Simeone, and J. Kang, “Wireless federated distillation for distributed edge learning with heterogeneous data,” in *2019 IEEE 30th Annu. Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC)*, 2019, pp.1–6.
- 11) D. Li, and J. Wang, “Fedmd: Heterogenous federated learning via model distillation,” *arXiv*. preprint arXiv:1910.03581, 2019.

## About the Authors



*MA Jiaxin, Ph.D. (Engineering)*

Research Administrative Division

OMRON SINIC X Corporation

Specialty: Biomedical Engineering and  
Machine Learning

---

The names of products in the text may be trademarks of each company.

# 過去の学習経験を組み合わせることで適応するロボット

田中 一敏

システムインテグレータは顧客の要望する作業を自動化するためにロボットをプログラミングする。ロボットシステムが構築される時だけではなく生産内容に軽微な変更が生じる時にも都度、プログラミングに大きな労力が必要とされている。もしロボットが自律的に新しい作業を学習できれば、この労力を省ける。このため、筆者らはロボットが新しい作業を効率的に学習する手法を新たに提案した。提案手法を用いたロボットは、過去の経験を活かして新しい作業を短時間で学習する。挿入方向の異なるペグインホールの実験を行い、提案手法の有効性を確認した。この手法によって、ロボットが新しい作業を短時間で学習できるようになり、ロボットを製造やサービスの新しい用途に利用できるようになる。

## Adaptation to a New Setup by Combining Skills

TANAKA Kazutoshi

The robotic system integrator programs the robot to automate the work requested by the customer. Robotic work for new parts that require minor changes in production requires a large amount of effort to program the robot. If the robot can learn new tasks autonomously, this effort can be eliminated. For this reason, we have proposed a method for robots to learn new tasks. The robot using the method learns new tasks in a short period of time by utilizing its experience. To evaluate the usefulness of the method, we conducted experiments on a peg-in-hole task with different insertion directions. The method enables the robot to perform new tasks in a short time, which will lead to new ways of using robots for manufacturing and service.

### 1. まえがき

本稿では、ロボットが新しい作業を学習する必要性について解説し(2章)、ロボットが新しい作業を学習する既存手法について検討し(3章)、筆者らが開発した過去の経験を活用してロボットが新しい作業を高速に学習する手法(4章)と実験結果(5章)を紹介する。本稿は、筆者らが発表した論文<sup>1)</sup>を参照し、過去の学習経験を組み合わせることで適応するロボットについて解説したものである。

### 2. ロボットによる新規作業学習の必要性

本章では製品生産を例として、産業用ロボットがどのように使われるかを紹介し、ロボットが新しい作業を学習する必要性について述べる。

#### 2.1 ロボットシステムの構築方法<sup>2)</sup>

産業用ロボットは、半製品と呼ばれる。すなわち、産業用ロボットのみ購入しただけでは、使えない。ロボットエンジニア、あるいはロボットのシステムインテグレータ

(以後、SIerと表記)と呼ばれる人間が、ロボットを他の機械と組み合わせることで、はじめてロボットシステムは完成し、利用できるようになる。

ロボットシステムを導入する手順として、RIPS (Robot system integration process standard) が定義されている。この手順における各フェーズ(各作業工程)とSIerの仕事は以下のとおりである。

1. 準備フェーズ(引合、企画構想): 顧客からの自動化要望の取りまとめ
2. 設計フェーズ(仕様定義、基本設計、詳細設計): 要件化、システム仕様の定義、設計
3. 製造フェーズ(製造、内部テスト): 加工、組立、配線、プログラミング
4. テストフェーズ(統合テスト、ユーザテスト): テスト
5. 稼働後(運用保守サポート): 保守

設計フェーズにおいてSIerは、操作対象物体と物体に対する操作を定義して、ロボット、エンドエフェクタ、セン

Contact: TANAKA Kazutoshi kazutoshi.tanaka@sinicx.com

サなどを選択し、ロボット用の架台、スライダ、治具、ポジショナーを設計する。その後、ロボットの動作をプログラミングする。

ロボットへのプログラミングは主に以下の2方式がある。

- ティーチングプレイバック方式
- マニュアル数値入力方式

ティーチングプレイバック方式では、移動させたい位置をロボットに教示し、記憶させ、その位置へどのように移動するかを指示する。一般的に産業用ロボットはティーチングプレイバック方式（教示再生方式）でプログラミングする。マニュアル数値入力方式では、ロボットの位置姿勢を表す数値を直接入力する。

ティーチングプレイバック方式は、さらに以下の3方式に細分される。

- ダイレクトティーチング
- オンラインティーチング
- オフラインティーチング

ダイレクトティーチングでは、ロボットを直接掴んで動かしながらプログラミングする。オンラインティーチングでは、ティーチングペンダントを使ってロボットを動かしながらプログラミングする。オフラインティーチングでは、ロボットを動かさず、部品の幾何的情報を用いたシミュレーションによってプログラミングする。これら3つの方式はいずれもロボットプログラミングにおいて多くの労力を要する。

## 2.2 新規の物体や作業に対するロボット自動適応

生産ラインでは、製品のマイナーチェンジなどがあり、それにとまらなロボットの作業をプログラミングし直す必要がある<sup>2)</sup>。加えて、現在行われているような単一品種大量生産ではなく、ロボットによる多品種少量生産を目指す場合、新しい部品の作業動作を、短時間でロボットに実装する必要がある。

新しい部品を扱う新しい作業におけるロボットの動作を、オフラインティーチングによって自動で生成する方向も検討されている<sup>3)</sup>。オフラインティーチングは部品や作業の変化のうち幾何的な変更に対応するのに使われる。他の、部品組立作業などの接触を多く含む動的な作業などに使える方法として、機械学習を用い、試行錯誤を通じて、ロボットの動作を獲得する方法もある。以下では、このロボットによる作業学習について検討する。

## 3. ロボットによる新しい作業の学習

代表的なロボット作業である部品組立作業を例として、ロボットによる新しい作業の学習について考察する。特に、部品組立作業における基礎的な動作であるペグインホールに着目し、新しい傾きの穴への挿入を学習することを考える。続いて、そのようなロボットによる新しい作業の学習を可能にする手法について既存研究を概観するとともに、筆者らの提案する新たな学習手法について解説する。

### 3.1 ロボット学習課題としての製品組立作業の新規物体操作

ロボットやエージェントが試行錯誤を通じてタスクの解き方を学ぶ技術として、強化学習<sup>4)</sup>がある。強化学習は、ロボットに作業を自動で獲得させることを可能にする。現在も強化学習を用いたロボットの作業学習に関する研究は活発に行われている。しかし、ロボットが既に学習した内容（経験）を活用して新しいタスクを短時間で学ぶことは未だ挑戦的な課題である。

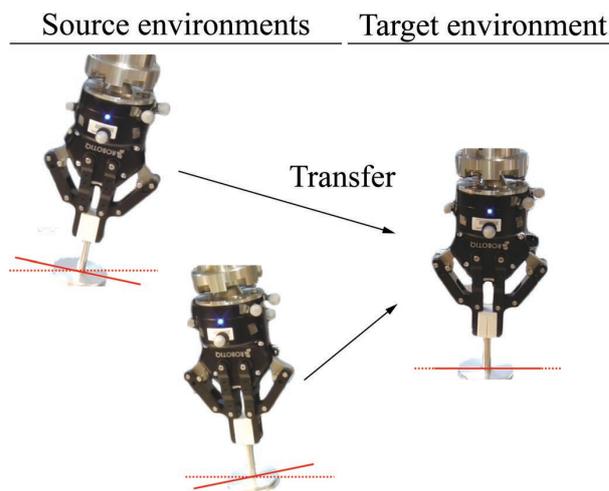


図1 本論文における問題設定：過去の経験を活用した新しい傾きの穴へのペグインホールの学習

ロボット作業の基礎的な動作であるペグインホールについて考える。特に既にペグの挿入傾きが異なる場合のペグインホールを学習したロボットが、同じ部品（ペグ）を未知の挿入傾きでペグインホールを学習することを想定する（図1）。これはペグ挿入の傾きが異なる一連のペグインホールタスクであり、実際の製造において生じる製品の軽微な変更を想定している。

この新しい傾きの穴へのペグインホール学習は、転移強化学習の課題として捉えられる。新しい傾きの穴へのペグインホールを転移強化学習と捉える場合、過去実施した傾きの穴へのペグインホールはソース環境と呼ばれ、新しい傾きの穴へのペグインホールはターゲット環境と呼ばれる。ペグインホールの転移強化学習は具体的には、以下の

問題として定式化できる。

1. ソース環境とターゲット環境は同じタスクだが、状態遷移ダイナミクスは異なる。
2. 正確な環境のダイナミクスは未知である。
3. ターゲット環境における学習時に、ソース環境との通信は不可能である。

ソース環境とターゲット環境の状態遷移ダイナミクスが異なる場合、ペグと穴が接触する位置も異なるため、穴の方向の小さな変化がロボットの状態遷移に大きな影響を及ぼす。すなわち、同じ状態で同じ行動を取っても、次の状態であるペグの位置や姿勢も変わる。ゆえに、穴の向きに応じた制御器が必要である。状態遷移ダイナミクスの完全なパラメータは未知か、不正確か、動的に変化するものである。ここでいうパラメータは、例えばリンクの重量、慣性テンソル、関節の粘性摩擦などがある。

さらに、実際の製造において、工場のラインにロボットは分散して配置されており、ロボット間の通信チャンネルはほとんどない。このため、ロボット同士が情報を大規模に収集、交換するのは難しい。

### 3.2 製品組立作業のロボット学習

強化学習のうち、モデルベース強化学習は、サンプル効率が高く、ロボットの学習に適している。近年では、複雑なダイナミクスを扱うために、深層ニューラルネットワークが用いられている<sup>5)</sup>。

転移強化学習<sup>6)</sup>も、他のタスクで得られた経験を使うことでサンプル効率を高められる。転移強化学習は、報酬関数や、状態遷移ダイナミクスや、状態行動空間が既知のソース環境と新しい環境が異なる強化学習として分類される。方策、Q 関数、ダイナミクスモデルを転用する。転移強化学習に関する多くの成果があるが、状態遷移ダイナミクスの異なる環境間での転用に関する研究は限られている<sup>7-11)</sup>。特に、モデルフリー強化学習の手法が多く、サンプル効率の高いモデルベース強化学習には使えない。

他の関連する研究領域はメタ学習である。メタ学習は、未知のタスクにロボットエージェントが高速に適應する方法である<sup>12-16)</sup>。メタ学習の多くはモデルフリー強化学習であるため、実用的なロボットの製品組立作業に使うのは難しい。

### 4. ダイナミクスモデルを集約するモデルベース転用強化学習<sup>1)</sup>

上述した問題を標準的な強化学習の問題として、あらためて定式化する。ロボットの部品組立作業をマルコフ決定過程としてモデル化する。状態と行動の空間  $S$  と  $A$ 、状態遷移  $T$ 、報酬関数  $R$  からなるタプルでマルコフ決定過程は

定義される。

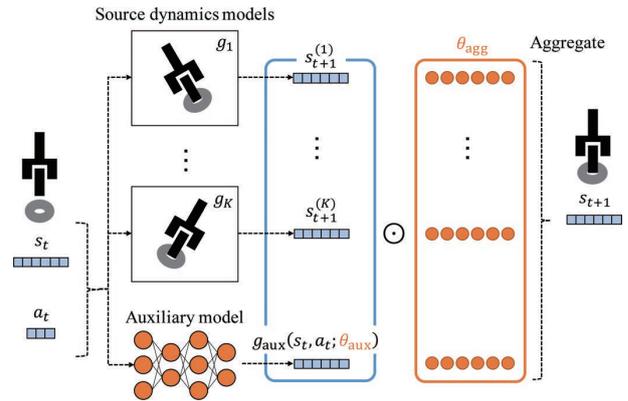


図2 提案手法。既知モデルの集約による新規モデルの学習。

強化学習における目標は、報酬関数の時間総和である収益を最大化する最適な制御器を得ることである。提案手法には、サンプル効率の優れたモデルベース強化学習を用いた。提案手法は、実際のダイナミクスから収集したデータを用いて、状態遷移  $T$  を関数  $g$  のモデルで近似する。モデルは行動の系列から状態の系列を予測するため、予測に基づいて最適な行動を選択できる。

転移強化学習の設定を用いた。ダイナミクスモデル  $g_k$  で表現された  $K$  個の関連する既知環境 (source environment) を考える。このモデルは、パラメタライズされてなくても、訓練できなくても良い。例えば、学習したニューラルネットワークでも、システムを近似したヒューリスティックな規則を手動で設計したシミュレータでも良い。モデルは、固定された訓練するパラメータのないブラックボックス関数でも良い。新しい環境における目標は、未知の状態遷移を近似したダイナミクスモデル  $g^{target}$  を短時間で学べるよう  $G = \{g_k\}$  を活用してモデルベース強化学習のサンプル効率でさらに改善することである。

提案手法である TRANS-AM は、過去のモデルの出力の線形和として、新しいモデル出力を計算する。それに加えて、線形和の残差を補償するため、残差モデルを学習する (図2)。

すなわち、提案手法は既知のダイナミクスモデルの出力を適應的に集約し、同時に、集約結果の残差を予測し、補償する補助モデル (auxiliary model) を学習する。既知モデルの出力  $s_{t+1}^{(k)} = g_k(s_t, a_t)$  は以下のように結合される。

$$S_{t+1} = \left[ \left( s_{t+1}^{(1)} \right)^T, \dots, \left( s_{t+1}^{(K)} \right)^T \right] \in \mathbb{R}^{K \times D}$$

ここで、 $D$  は状態空間  $S$  の次元である。新規環境におけるダイナミクスモデルを以下のように表現する。

$$g^{target} = \mathbf{1} \left( \theta_{agg} \odot S_{t+1}, g_{aux}(s_t, a_t; \theta_{aux})^T \right)$$

ここで、 $\theta_{agg} \in \mathbb{R}^{(K+1) \times D}$  は訓練可能な補助重みパラメータ

の行列、 $\odot$  は要素ごとの積、 $\mathbf{1}$  は  $K+1$  の大きさの全ての要素が  $\mathbf{1}$  の列ベクトルである。この  $\theta_{agg}$  は正則化しない。

提案手法は、モデルベース強化学習のアルゴリズムで、ネットワークのパラメータ  $\theta = (\theta_{agg}, \theta_{aux})$  を訓練する。Cross entropy method を用いて、モデルに基づく行動を生成する。

提案手法の利点をまとめる。提案手法はモデルベース強化学習であり、サンプル効率が非常に高い。提案手法では収集した過去のダイナミクスモデルを利用する。この過去のダイナミクスモデルとして、ニューラルネットワークだけでなく、他の非線形関数による収集したデータの近似関数も利用できる。他にも、あらゆるブラックボックスが利用できる。例えばヒューリスティックに状態遷移が決まるシミュレータ、物体の先験的知識を用いた概形とパラメータの同定によって作ったモデルなどが利用できる。

### 5. 実験<sup>1)</sup>

提案手法の実用性を評価するため、異なる穴の傾きのペグインホールタスクを行った (図3)。実験には、ロボットアーム UR5 (Universal Robots) を用いた。ロボットアームには筆者らが開発した柔軟手首<sup>17)</sup>、グリッパ (2F-85, ROBOTIQ)、6軸力覚センサ (FT300, ROBOTIQ) を搭載した。グリッパの位置と姿勢は6個のモーションキャプチャカメラ (FLEX13, OptiTrack) で計測した。ペグとして、ステンレス製の、直径 10 mm のものを用いて、グリッパに固定した。ペグと穴の公差は H7/h7 であった。異なるペグの傾きを試した。  $\phi \in \{-10^\circ, 0^\circ, 10^\circ\}$  と変更した。

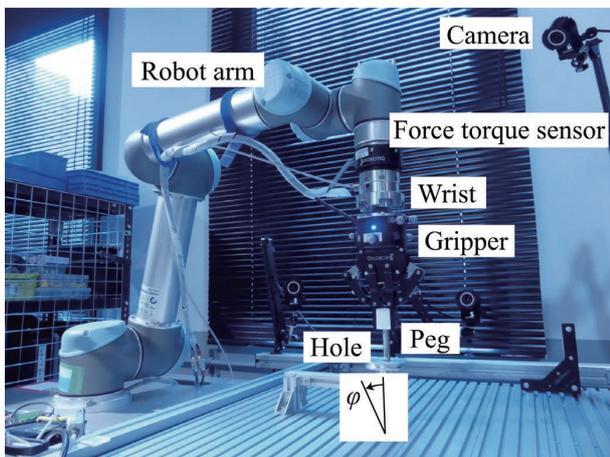


図3 実験に用いたロボットシステムの概観

モデルベース強化学習における行動として、ロボットの手先 (アームの先端) 速度の水平方向成分を用いた。ただし手先の傾きは一定に維持すると仮定する。行動として決められた目標手先速度から、ロボット制御のフレームワークである MoveIt<sup>16)</sup> を用いて目標関節角速度を計算し、5 Hz でロボットに送った。手先の位置、接触力の並行成分、

グリッパの位置と姿勢を用いて状態を定義した。水平成分の位置と力は3変数で、グリッパの姿勢はオイラー角の正弦と余弦からなる6変数で表現した。この際に、グリッパの位置は目標位置を原点とする相対位置で表現して正規化し、力も特定の大きさを正規化した。また報酬は、目標位置までの距離と接触力の線形和で定義する。

ダイナミクスモデルはモデル予測制御で制御し、取得したデータから学習した。ネットワークは Adam で最適化し、学習率は、0.01 に設定した。  $\phi = 10^\circ$  から  $\phi = 0^\circ$  への転移 ( $K=1$ ) と、  $\phi = \{-10^\circ, 10^\circ\}$  から  $\phi = 0^\circ$  への転移 ( $K=2$ ) の TRANS-AM を試した。

学習の1セッションは20エピソード、1エピソードは100タイムステップとした。過剰な接触力が生じるか、穴からペグが過剰に離れた場合、ロボットや部品の破損を防ぐため、エピソードを失敗したとして、終了した。最初の2エピソードでは、ランダムに行動を決定した。異なる5つのランダムシードを試し、5, 10, 15, 20エピソードでの成功率を計算した。

表1 初回成功エピソードの平均と標準偏差。各エピソードにおける成功率。

	First Success	e=5	10	15	20
Baseline	9.25 ± 5.49	0.2	0.4	0.6	0.8
TRANS-AM (K=1)	8.20 ± 5.34	0.4	0.6	0.8	1.0
TRANS-AM (K=2)	7.20 ± 3.37	0.4	0.6	1.0	1.0

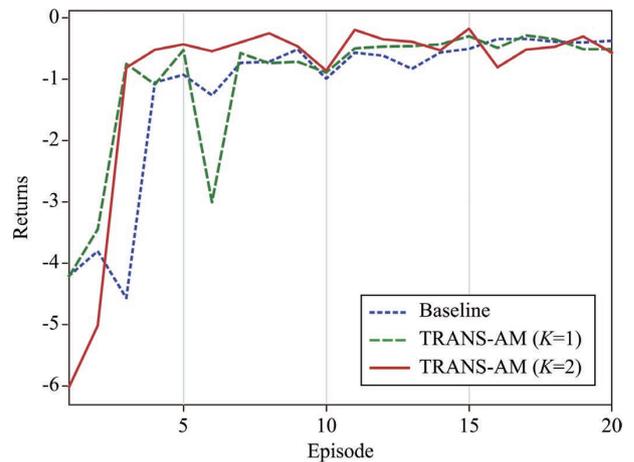


図4 収益比較。

表1に成功率と初回成功の平均と標準偏差を示す。TRANS-AMがソースダイナミクスモデルを用いて、先見の知識なしに学習するベースラインよりも、初回の成功を早めたことが分かる。図4から、TRANS-AMの方が、早いエピソードで高い報酬を得たことが分かる。図5に成功したエピソードのスナップショットを示した。以上のように

に、TRANS-AM が短期間での適応を可能にすることが分かった。

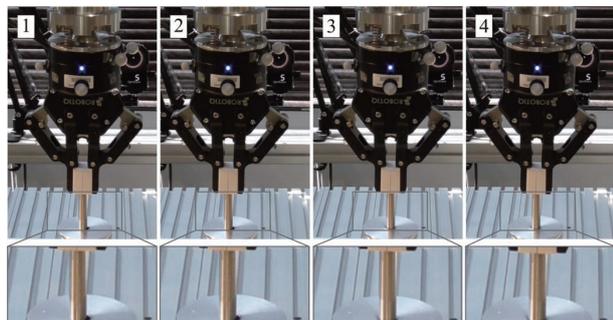


図5 学習したペグインホールの外観。

## 6. むすび

本稿では、新しい作業をロボットが学習する必要性について解説し、筆者らの提案するロボットの学習方法を紹介した。

提案手法によって、ロボットが過去の経験を活用することで、新しいタスクに対して高速に作業を学習することが可能になった。この手法を発展させて使うことで、生産する製品の軽微な変更に対して自動で新しいロボット動作を作成できる。加えて、ロボットによる多品種少量生産が可能になり、一人ひとりに合った製品を届けることが可能になる。さらに、ロボットプログラミングのコストを下げることで、製品組立作業にとどまらず、様々な単純手作業のロボットによる代替が現実的な選択肢になる。

多数のモデルを既に学習していた際に、どのモデルを学習するのか、また、将来においてダイナミクスを再利用することを考慮して、使いやすい形でモデルを学習する方法は、この研究において残された課題であり、筆者たちが今取り組んでいる課題である。

## 参考文献

- 1) Tanaka, K.; Yonetani, R.; Hamaya, M.; Lee, R.; Drigalski, F.; Ijiri, Y. TRANS-AM: “Transfer Learning by Aggregating Dynamics Models for Soft Robotic Assembly”. IEEE International Conference on Robotics and Automation. 2021.
- 2) 経済産業省, 日本ロボット工業会. ロボットインテグレータのスキル読本. 2018, 87p.
- 3) 小沢邦昭, 熊本健二郎, 明石吉三, 中田英樹. オフラインロボット教示における高速干渉チェックの一方式. 日本ロボット学会誌. 1986, Vol.4, No.2, p.79-88.
- 4) Sutton, R. S.; Barto, A. G. Reinforcement Learning: An Introduction, MIT Press, 1998, 344p.
- 5) Chua, K.; Calandra, R.; McAllister, R.; Levine, S. “Deep reinforcement learning in a handful of trials using probabilistic dynamics models”. Advances in Neural Information Processing Systems, 2018, p.4754-4765.
- 6) Taylor, M. E.; Stone, P. Transfer learning for reinforcement learning domains: A survey. J. Mach. Learn. Res. 2009, Vol.10, No.7, p.1633-1685.
- 7) Chen, T.; Murali, A.; Gupta, A. “Hardware conditioned policies for multi-robot transfer learning”. Advances in Neural Information Processing Systems. 2018, p.9333-9344.
- 8) Barekatin, M.; Yonetani, R.; Hamaya, M. “Multipolar: Multi-source policy aggregation for transfer reinforcement learning between diverse environmental dynamics”. International Joint Conference on Artificial Intelligence. 2020, p.3108-3116.
- 9) Song, J.; Gao, Y.; Wang, H.; An, B. “Measuring the distance between finite markov decision processes”. International Conference on Autonomous Agents and Multiagent Systems. 2016, p.468-476.
- 10) Yu, W.; Liu, C. K.; Turk, G. “Policy transfer with strategy optimization”. International Conference on Learning Representations. 2019.
- 11) Rajendran, J.; Lakshminarayanan, A. S.; Khapra, M. M.; Prasanna, P.; Ravindran, B. “Attend, adapt and transfer: Attentive deep architecture for adaptive transfer from multiple sources in the same domain”. International Conference on Learning Representations. 2017.
- 12) Vanschoren, J. Meta-learning: A survey. arXiv:1810.03548, 2018, (参照 2021-10-01).
- 13) Nagabandi, A.; Clavera, I.; Liu, S.; Fearing, R. S.; Abbeel, P.; Levine, S.; Finn, C. “Learning to adapt in dynamic, real-world environments through meta-reinforcement learning”. International Conference on Learning Representations. 2019.
- 14) Sæmundsson, S.; Hofmann, K.; Deisenroth, M. “Meta reinforcement learning with latent variable gaussian processes”. Conference on Uncertainty in Artificial Intelligence. 2018, Vol.34, p.642-652.
- 15) Clavera, I.; Rothfuss, J.; Schulman, J.; Fujita, Y.; Asfour, T.; Abbeel, P. “Model-based reinforcement learning via meta-policy optimization”. Conference on Robot Learning. 2018, p.617-629.
- 16) Schoettler, G.; Nair, A.; Ojea, J. A.; Levine, S.; Solowjow, E. “Meta-reinforcement learning for robotic industrial insertion tasks”. International Conference on Intelligent Robots and Systems, 2020. p.9728-9735.
- 17) Tanaka, K.; Drigalski, F.; Hamaya, M.; Lee, R.; Nakashima, C.; Shibata, Y.; Ijiri, Y. “A Compact, Cable-driven, Activatable Soft Wrist with Six Degrees of Freedom for Assembly Tasks”. IEEE/RSJ International Conference on Intelligent Robots and Systems. 2020, p.8752-8757.
- 18) Chitta, S.; Sukan, I.; Cousins, S. Moveit! IEEE Robotics & Automation Magazine. 2012, Vol.19, No.1, p.18-19.

## 執筆者紹介



**田中 一敏 TANAKA Kazutoshi**  
オムロン サイニックエックス株式会社  
リサーチアドミニストレイティブディビジョン  
専門：知能ロボット工学  
所属学会：日本ロボット学会  
博士（情報理工学）

---

本文に掲載の商品の名称は、各社が商標としている場合があります。



# 3

---

人と協調する 人とともに働く

# 新しい現場で簡単に立ち上げる

米谷 竜

ある環境におけるタスクに関して獲得された知識を活用し、別の環境におけるタスクの学習を効率化するアプローチとして転移学習がある。本稿では特に転移強化学習をとりあげ、その概要を紹介するとともに、著者らの最近の成果である MULTIPOLAR (MULTI-source POLIcy AggRegation) を解説する。

MULTIPOLAR は、転移元・転移先環境の間で状態遷移確率分布が未知かつ異なる状況を対象とした転移強化学習手法である。同様の問題を扱う先行研究の多くと異なり、MULTIPOLAR は転移元環境において獲得された方策関数を適応的に統合して学習に利用するため、環境において学習サンプルを収集する必要がないという利点がある。OpenAI Gym 環境におけるシミュレーション評価によって、MULTIPOLAR の有効性を確認している。

## Adapting Machine Learning Models to New Environments

YONETANI Ryo

Transfer learning is a machine-learning approach that leverages knowledge acquired from a certain environment to efficiently train a model in a new environment. In this paper, we particularly focus on transfer “reinforcement” learning and introduce a brief abstract and our recent achievement called MULTIPOLAR (MULTI-source POLIcy AggRegation). MULTIPOLAR is a method for transfer reinforcement learning between diverse environmental dynamics. It enables transfer by adaptively aggregating policy functions acquired from source environments, thus does not require many training samples to be collected from each of the source environments. Through extensive simulation evaluations using OpenAI Gym, we confirmed the effectiveness of the proposed approach over existing transfer reinforcement learning methods.

### 1. まえがき

機械学習においてモデルが識別や生成、制御といったタスクを実現できるようにするためには、そのタスクにおいて期待される入出力を表現するデータ（学習データ）を収集する必要がある。このデータ収集が機械学習を成功させる鍵の一つとなる。ある現場において、モデルに高いパフォーマンスを発揮させるためには、その環境に現れうる現象を広くカバーした学習データを用意できていることが望ましい。なぜならば、学習されたモデルを同現場において運用する際、学習時に含まれなかった入力に対して期待される出力を得ることは必ずしも保証されないためである。

一方で、現実世界において機械学習モデルの運用を期待する現場は数多く存在するものの、また現場によってデータの計測条件やタスクの細かな種別はしばしば異なる。「監視カメラの設置場所ごとに光源環境が異なる」「製造現場においてラインごとに扱う部品が異なる」などがその例

である。これに対して、現場ごとにデータを大量に収集・学習することで、その現場に適応した高パフォーマンスのモデルが獲得できると考えられる。しかしながら、データ収集・学習それぞれにかかる時間・資源的コストの観点から、このようなアプローチは必ずしも現実的ではない。

このような課題を解決する一つのアプローチとして、**転移学習**と呼ばれる技術がある。転移学習は「ある問題（転移先）を効果的かつ効率的に解くために、別の関連した問題（転移元）のデータや学習結果を再利用する」技術を広く指す<sup>1)</sup>（括弧内は著者による注記）。「ある問題」を「新しい現場におけるモデルの運用」、「別の関連した問題のデータ」を「そのモデルを学習するために別の現場で収集したデータ」と読み替えると、先の課題との対応が見える。すなわち転移学習を用いることで、多様な現場のそれぞれにおいて機械学習モデルを一から構築するのではなく、過去の現場から得られた何らかのデータを活用して学習を効率的にすることができる。より具体的には、たとえば扱う部品が同種ではあるが完全には一致しないような製造ラインが多数存在するとき、あるラインで収集したデー

Contact : YONETANI Ryo ryoyonetani@sinicx.com

タや獲得したモデルを用いて、別のラインにおける機械学習モデルの立ち上げを早期化・省力化することが期待される。

このような転移学習は、機械学習およびその関連分野において長らく取り組まれてきた課題である。深層学習の発展とともに近年研究が活発であるドメイン適応やマルチタスク学習、メタ学習などのトピックとも関連が深い。さらに、転移元・転移先においてどのようなデータが利用可能かによって、様々な派生問題やその解決手段が存在する。本稿では、転移強化学習一すなわち強化学習のための転移学習を取り上げ、その概要に加えて「転移元と転移先で環境の従う状態遷移確率分布が異なる」「転移元における生データにアクセスできない」といった状況を扱うことが可能な著者らの研究成果<sup>2)</sup>を紹介する。

## 2. 研究の背景

### 2.1 強化学習

転移強化学習の前に、まずは基本的な強化学習<sup>3)</sup>の問題設定と応用について簡単に述べる。強化学習では、マルコフ決定過程 (Markov Decision Process; MDP) でモデル化された環境において、ある決められたルール (方策; policy) に基づいて行動するエージェントを考える (図1)。MDPは  $M = (\rho_0, \gamma, S, A, R, T)$  というタプルで表現される。 $S$  はエージェントのとりうる状態が定義される空間 (状態空間)、 $A$  はエージェントがとりうる行動の定義される空間 (行動空間) を表す。エージェントは (1) 観測された現在の状態  $s_t \in S$  において、方策  $\pi(a_t | s_t, \theta)$  に基づいて (2) 行動  $a_t \in A$  を選択し、(3) あらかじめ決められた報酬関数  $R$  に基づいた報酬  $r_t = R(s_t, a_t)$  を受け取る。その後、エージェントは (4) 状態遷移確率分布  $T(s_{t+1} | s_t, a_t)$  に基づいて次の状態へと遷移する。 $\rho_0$  は初期状態を定義する分布である。このような枠組みにおいて、強化学習は累積報酬  $\sum_{k=0}^{\infty} \gamma^k r_{t+k}$  を最大化する方策のパラメタ  $\theta$  の推定を目指す ( $\gamma$  は割引率)。

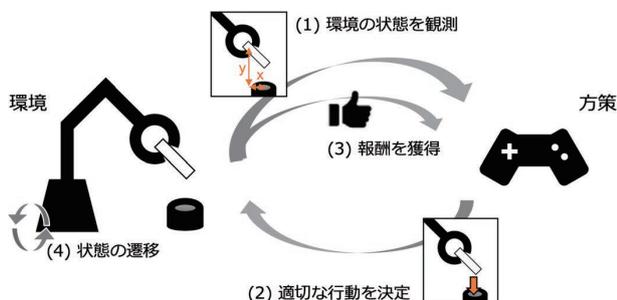


図1 強化学習の枠組み

このように強化学習は、エージェントによる個別の行動に対して何かしらの明確な正解が与えられるのではなく、

行動の“良さ”を示す報酬という量のみが教師信号として与えられる点、そのようななかで単一時刻における適切な行動ではなく、長期にわたる行動の“系列”を推定する必要がある点などにおいて、一般によく知られる教師あり学習とは大きく異なる。その結果、強化学習はロボットマニピュレータの制御や自動走行ロボットの制御<sup>4)</sup>、ヘルスケア<sup>5)</sup>、ゲームAI<sup>6)</sup>など、「ある環境において、長期的に効果のある意思決定をしたい」という要求のある応用シーンにおいて、幅広い活用が期待されている。

一方で、強化学習では固定された学習データセットが与えられるのではなく、エージェント自身の行動に基づいて学習サンプルが逐次的に収集されることになる。そのため、エージェントの行動に対して何かしらのリスク (たとえば誤ったマニピュレーションによってロボットが緊急停止、あるいは破損してしまうなど) が無視できないケースでは、「いかに少ないデータから良い方策を学習できるか」が重要な課題となる。これを解決する一つのアプローチが、次節で紹介する転移強化学習である。

### 2.2 転移強化学習

転移強化学習では、転移元の環境で得られた知識を活用することにより、転移先における強化学習をなるべく少ないサンプルで実現することを目指す<sup>7)</sup>。それぞれの環境はMDPによってモデル化され、転移元・転移先のMDPがどのように異なるか、転移元環境の“知識”として具体的に何が与えられるかといった条件に基づいて、いくつかの問題に派生する。たとえば、Successor Features と呼ばれる有名な研究<sup>8)</sup>では、転移元と転移先で報酬関数が異なる状況を扱っている。転移元からは状態遷移サンプル  $(s_t, a_t, s_{t+1})$  が収集され、そのサンプルを用いて新たな環境における未知の報酬関数を推定するという形で方策の学習を効率化する。より挑戦的な課題として状態空間、行動空間、および状態遷移確率分布までもが環境間で異なるという状況を扱う研究も存在する<sup>9)</sup>。同研究はこのようなケースに対して、両環境で共通に利用できる特徴空間を学習し、そこで方策を学習することによる転移を実現している。

### 2.3 状態遷移確率分布の異なる環境間の転移

とりわけ本稿では、環境間で状態遷移確率分布  $T$  のみが異なるという問題設定を考える。直感的には、「エージェントが同一の状態  $s_t$  で同一の行動  $a_t$  を取った後の、次の状態  $s_{t+1}$  (より正確には次にある特定の状態に遷移する確率) が異なる」ということを示す。たとえば、ロボットマニピュレータにおいてリンクの長さや重さが異なるケースが挙げられる<sup>10)</sup>。また、自動走行ロボットの制御においては、車載物の重量や路面状況が状態遷移確率分布に影響を与えるであろう。文献10) で取り上げられているように、このような課題の難しさは、実際の状態遷移確率分布が転

移元・転移先それぞれの環境においてしばしば未知であるという点にある。したがって、未知の状態遷移確率分布をどのように推定するか、あるいは環境間の状態遷移確率分布の違いをどのように定量化するかが、同問題を扱う転移強化学習において重要な課題となる。

このような課題に取り組んだ関連文献を紹介したい。転移強化学習における先述の課題を解決するために、既存研究の多くは転移元の環境から学習サンプル  $(s_t, a_t, s_{t+1})$  あるいは  $(s_t, a_t, s_{t+1}, r_t)$  を大量に収集し活用する。たとえば文献<sup>11)12)</sup>では、転移元から収集したサンプルを利用して転移元・転移先における MDP の違いを定量化している。このような定量化に基づいて、転移先に最も類似する転移元環境を特定し、転移学習に活用することができる。一方、文献<sup>10)</sup>では転移元環境が多数与えられる問題設定を取り上げ、環境を表現する条件パラメタ (explicit encoding や implicit encoding と呼ばれる) に応じて異なる行動を出力可能な方策を学習する。これにより、未知の転移先環境においても、その環境に対応する条件パラメタから適切な行動を選択可能となっている。同文献ではさらに環境間で状態空間や行動空間が異なる状況においても同様の手法が適用できる可能性を示している。

なお、多数の転移元環境を想定した設定は、メタ学習やマルチタスク学習等類似するトピックの研究でも見られる。たとえば文献<sup>13)</sup>では、それぞれ異なるルールに従うサブ方策と、与えられたタスクに対してどのサブ方策を利用するかを決定するマスター方策を、多様な環境から収集されたサンプルを用いてメタ学習するアプローチを提案している。類似するアプローチは Actor-mimic と呼ばれる手法<sup>14)</sup>でも採用されており、ルールの異なるゲームを複数用いた方策のマルチタスク学習にもとづいて、新たなゲームにおいても効率的に方策を転移できることを示している。

### 3. MULTIPOLAR: Multi-source Policy Aggregation

以降では、環境間で状態遷移確率分布の異なる状況を想定した転移強化学習に関する、著者らの近年の取り組み<sup>2)</sup>を紹介する。まず本節では研究のモチベーションと具体的なアプローチを解説し、次節で評価実験の結果を報告する。

#### 3.1 モチベーション

2.3 節で紹介したとおり、環境間における状態遷移確率分布の違いを扱う既存研究は、しばしば転移元における学習サンプルが大量に必要となる。しかしながら、このような前提に基づく学習は現実の場面において必ずしも容易ではない。たとえば、さまざまな製造現場においてロボットがマニピュレーションタスクに取り組むシナリオを考える。たとえそれらのタスクが同一であったとしても、ロ

ボットの持つキネマティクスやダイナミクスが異なれば、それらは異なった MDP としてモデル化されることとなり、結果として最適となる方策も異なる<sup>10)</sup>。各環境においてサンプルを大量に収集し学習に活用できるかどうかは、現場それぞれのストレージや通信環境に依存する。たとえばセキュリティ等の都合から閉じたネットワークの中でロボットが稼働するケースであれば、転移元環境から逐次的に学習サンプルを集める必要のある既存手法<sup>10)13)14)</sup>の適用は難しい。

これに対して本研究では、(複数の) **転移元環境から何らかの方法で獲得された方策関数が与えられる**という別の問題設定を考える。製造現場における上記の例であれば、各現場において個別のロボットを動かす制御アルゴリズム (= 方策関数) を集約し活用することで、新たな現場における方策の獲得を効率化することを目指す。このような方策関数は、ヒューリスティックに設計されたルールの集合であろうと、強化学習を含む機械学習でデータドリブンに獲得されたモデルであろうと、既存研究で必要とされる学習サンプル集合と比較するとデータサイズが大幅に小さく、また転移元・転移先間での情報のやりとりも一度ですむため、先に述べたストレージや通信環境に関する課題に対処することができる。一方で、各転移元環境から共有された方策関数自体は同環境の状態遷移を記述するものではない。すなわち、転移先環境においてその方策がどの程度有効に働くかは未知という新たな問題が生まれることになる。

これに対して本研究では、**MULTIPOLAR (MULTI-source POLicy AggRegation)** と呼ばれる新たな転移強化学習手法により、この問題の解決を目指す。MULTIPOLAR は、複数の転移元環境から収集された方策関数の出力を、転移先環境における累積報酬が最大化されるように、強化学習の枠組みの中で“適応的に統合”する。以降の節で解説するように、この統合は転移元・転移先環境それぞれの状態遷移確率分布が未知であっても実現できる。さらに、同じく方策の統合を行う他の手法<sup>15)</sup>と異なり、行動空間が連続・離散いずれの場合も扱うことができるという応用範囲の広さを利点として持つ。

#### 3.2 準備

MULTIPOLAR の解説にあたって、与えられた複数の転移元各環境をインデクス付きの MDP  $M_i = (\rho_0, \gamma, S, A, R, T_i)$  として表現する。ある 2 環境において  $T_i \neq T_j$ 、かつ全ての環境において  $T_i$  は未知である。さらに、既存研究と異なり  $T_i$  に基づいてサンプルを収集することもできないものと仮定する。

このような状況において本研究では、ある環境  $M_i$  において方策関数  $\mu_i: S \rightarrow A$  が与えられるものとする。この方策関数  $\mu_i$  は決定的であり、ニューラルネットワークのよう

に陽にパラメタ化されたモデル、あるいは人によってヒューリスティックに設計されたアルゴリズムなど、さまざまな形態をとりうる。いずれにしても、それぞれの方策関数がどのような状態遷移確率分布を持つ環境でどのように設計・獲得されており、転移先環境においてどの程度の性能を発揮するかは、学習時に未知である。

いま、合計  $K$  種類の転移元環境から方策関数の集合  $L = \{\mu_1, \dots, \mu_K\}$  が与えられたとする。このとき、本研究のゴールは、新たな環境において累積報酬を最大化する方策  $\pi_{\text{target}}(a_t | s_t; L, \theta)$  のパラメタ  $\theta$  を、強化学習のなかで効率的に推定することとなる。

### 3.3 方策関数の適応的統合

以下では、MULTIPOLAR の一番のポイントである方策関数の統合について述べる (図2も参照)。転移先環境において観測された状態  $s_t$  に対して、ある転移元の方策関数  $\mu_i$  を用いて決定的に得られる行動を  $a_t^{(i)} = \mu_i(s_t)$  とおく。連続行動空間では、 $a_t^{(i)}$  は  $D$  次元の実数ベクトルであり、ある時刻  $t$  において同時に実行される  $D$  種類の行動を表現している。一方行動空間が離散である場合、 $a_t^{(i)}$  は  $D$  種類の行動から一つを選ぶワンホットベクトルとなる。いずれの場合も、方策関数集合  $L$  のそれぞれの要素に対して同様の行動を考えることで、以下のような行動行列を定義することができる。

$$A_t = \left[ (a_t^{(1)})^\top, \dots, (a_t^{(K)})^\top \right] \in \mathbb{R}^{K \times D} \quad (1)$$

MULTIPOLAR では、この  $A_t$  を強化学習のループにおいて累積報酬を最大化させるように適応的に統合する。この統合結果は転移先での方策の学習において「およそこのように行動すれば良い」という帰納バイアスを与えるものであり、転移元の状態遷移分布  $T_i$  を知ることなく推定可能である。より具体的には、以下のような適応的統合関数  $F_{\text{agg}}: S \rightarrow A$  を定義する。

$$F_{\text{agg}}(s_t; L, \theta_{\text{agg}}) = \frac{1}{K} \mathbf{1}^K (\theta_{\text{agg}} \odot A_t) \quad (2)$$

ただし、 $\theta_{\text{agg}} \in \mathbb{R}^{K \times D}$  は学習可能な統合重み、 $\odot$  は行列の要素ごとの積、 $\mathbf{1}^K$  は全要素が1である  $K$  次元ベクトルを表す。 $\theta_{\text{agg}}$  は状態  $s_t$  によらず学習の中で一つに決まり、また正規化・正則化されずに推定される量である。その結果 MULTIPOLAR は転移元から得られる行動を単純に内挿するのではなく、外挿することや、重要と思われる行動を強調すること、あるいは逆に有効でない行動を抑制することも可能である。

### 3.4 補助ネットワーク

図2に示すように、MULTIPOLAR では  $F_{\text{agg}}$  に加え、以下のような補助ネットワーク  $F_{\text{aux}}: S \rightarrow A$  を導入する。

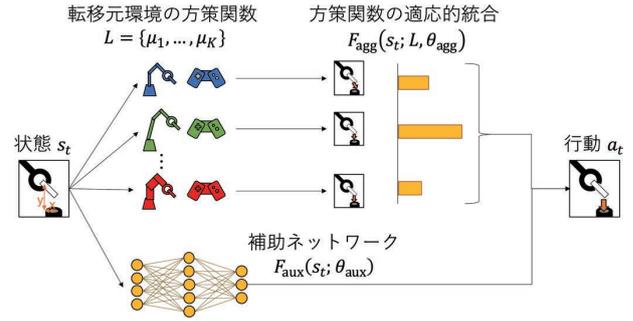


図2 MULTIPOLAR の全体図

$$F(s_t; L, \theta_{\text{agg}}, \theta_{\text{aux}}) = F_{\text{agg}}(s_t; L, \theta_{\text{agg}}) + F_{\text{aux}}(s_t; \theta_{\text{aux}}) \quad (3)$$

この補助ネットワーク  $F_{\text{aux}}$  は2種類の機能を持つ。

- 1) もし  $F_{\text{agg}}$  が転移先環境において既に最適に近い行動を出力できている場合、 $F_{\text{aux}}$  はさらに高い累積報酬を獲得するための「補正項」としての役割を果たす。
- 2) 一方で  $F_{\text{agg}}$  が転移先環境において有効でない場合、 $F_{\text{aux}}$  は累積報酬を最大化するための行動を学習することができる。なお、類似したアプローチは残差方策学習 (residual policy learning) と呼ばれる手法<sup>16)17)</sup> にも見られる。ただしこれらの既存研究は、転移元の方策が単一、かつ行動空間が連続という状況のみを扱うという点で本研究の取り組みとは大きく異なる。

### 3.5 方策の構成

上記の関数  $F(s_t; L, \theta_{\text{agg}}, \theta_{\text{aux}})$  をリパラメタライズすることで、転移先環境の方策  $\pi_{\text{target}}(a_t | s_t; L, \theta)$  として用いることができる。たとえば連続行動空間を扱う場合、 $F(s_t; L, \theta_{\text{agg}}, \theta_{\text{aux}})$  を平均とするガウシアン方策として  $\pi_{\text{target}}$  を与えればよい。一方、行動空間が離散である場合、 $F(s_t; L, \theta_{\text{agg}}, \theta_{\text{aux}})$  をソフトマックス関数によって正規化することで方策として扱うことができる。いずれの場合も、方策  $\pi_{\text{target}}(a_t | s_t; L, \theta)$  のパラメタは  $\theta = [\theta_{\text{agg}}, \theta_{\text{aux}}]$  のみであり、方策を直接更新可能な標準的な強化学習アルゴリズムによって学習可能である。

## 4. MULTIPOLAR の性能評価

本節では、文献<sup>2)</sup> において実施した OpenAI Gym による MULTIPOLAR の評価結果を紹介する。なお、本評価実験に用いたプログラムは、著者らの GitHub リポジトリ (<https://github.com/omron-sinix/multipolar>) において公開されている。

### 4.1 評価設定

評価環境として、OpenAI Gym から CartPole, Acrobot, LunarLander, RoboschoolHopper, Roboschool Ant, Roboschool

InvertedPendulumSwingUp という 6 種類のシミュレーション環境を選択した。それぞれの環境は、ロボットのリンク長や重量、摩擦など、MDP の状態遷移確率分布を特徴づけるいくつかのキネマティクス、ダイナミクスパラメータが含まれる。本評価ではこのパラメータをランダムに変化させることにより、環境ごとに状態遷移分布の異なる 100 種類の環境インスタンスを作成した。いくつかの例を図3に示す。

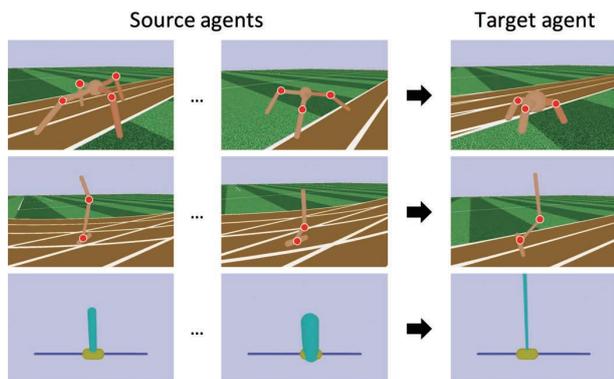


図3 状態遷移確率分布の異なる転移元環境 (source agents) と転移先環境 (target agents)。 (文献<sup>2)</sup>より引用。Copyright(c)2020 International Joint Conferences on Artificial Intelligence, All rights reserved.)

さらに、各環境インスタンスに対して、多層パーセプトロン (MLP) でモデル化された方策を学習し、転移元における方策関数として利用した。MULTIPOLAR の評価にあたっては、1) まず 100 種類のうちそれぞれのインスタンスを転移先環境として設定し、2) 自身以外でランダムに選択された  $K=4$  種類のインスタンスを転移元環境としてさらに選択、対応する方策関数の集合  $L = \{\mu_1, \dots, \mu_K\}$  を利用することで、3) 転移先環境におけるエピソード報酬

(episodic reward) がタイムステップとともにどのように改善されていくかを記録、他手法と比較した。

### 4.2 実装と比較手法

全ての実験は Stable Baselines<sup>18)</sup> をベースとした実装により実施した。LunarLander 環境では Soft Actor Critic<sup>19)</sup>、その他の環境では Proximal Policy Optimization<sup>20)</sup> を用いて方策を学習した。

また MULTIPOLAR の比較手法として、以下を評価した。

- **多層パーセプトロン (MLP)** : 転移先環境で新規に方策を学習するベースラインであり、転移元環境の情報の有無が与える影響を評価できる。
- **残差方策学習 (Residual policy learning; RPL)** : MULTIPOLAR において  $K=1$  の場合に相当するベースラインであり、転移元環境が複数与えられることの利点を評価できる。
- **Attend, adapt, and transfer (A2T)<sup>15)</sup>** : MULTIPOLAR において  $F_{agg}$  のパラメータが、 $s_t$  を入力とした別モデルによって与えられる場合に対応する。もともと A2T は離散行動空間のみを扱うが、ここでは MULTIPOLAR と同様に統合結果をガウシアン方策の平均として利用することで、連続行動空間における学習に対応した。

### 4.3 実験結果

図4に環境・手法ごとのエピソード報酬の変化をプロットした。さらに、学習終了時における、全エピソードにわたるエピソード報酬の平均値 (average episodic reward) を表1に示した。

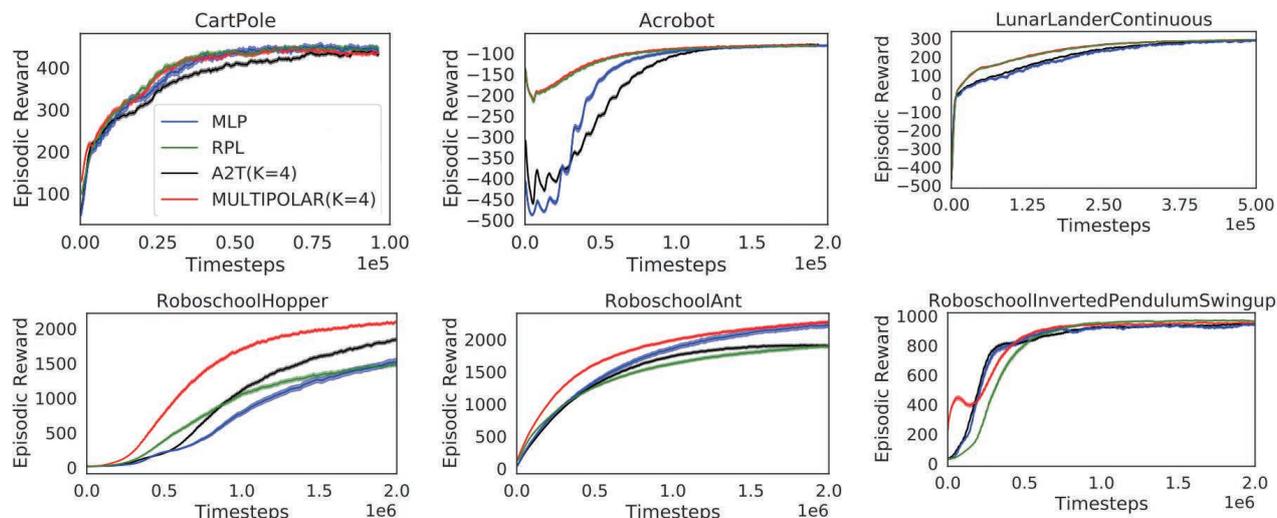


図4 エピソード報酬の変化 (文献<sup>2)</sup>より引用。Copyright(c)2020 International Joint Conferences on Artificial Intelligence, All rights reserved.)

表1 平均エピソード報酬の比較 (文献<sup>2)</sup> より引用。Copyright (c)2020 International Joint Conferences on Artificial Intelligence, All rights reserved.)

	CartPole	Acrobot	LunarLander
MLP	291	-111	216
RPL	289	-98	246
A2T	281	-120	226
MULTIPOLAR	299	-96	246
	Hopper	Ant	InvertPendulum
MLP	92	1500	409
RPL	152	1432	322
A2T	126	1361	458
MULTIPOLAR	283	1744	588

CartPoleのような簡単な環境を除いて、転移元環境の方策を活用するRPL, A2T, MULTIPOLARはより少ないタイムステップで高いエピソード報酬を獲得できていることが分かる。これらの3手法間の優劣は環境によって異なるが、いずれの場合もMULTIPOLARが最も高い性能を示している。

MULTIPOLARの挙動をさらに分析するため、図5では4種類の方策関数のうち2種類を高性能な方策( $\mu_1, \mu_2$ )、残り2種類を低性能な方策( $\mu_3, \mu_4$ )とした際の重み $\theta_{agg}$ を示した。タイムステップの増加にともない、高性能な方策関数により大きな重みが与えられるように学習が進んでいることが分かる。

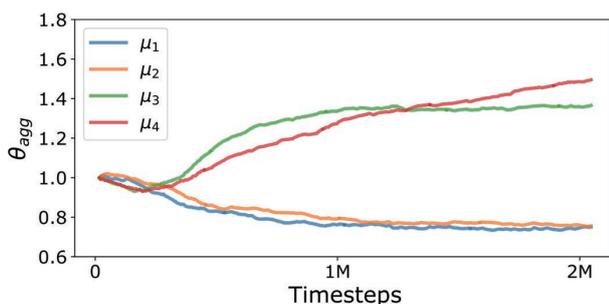


図5 性能の異なる方策関数に対する重みの変化 (文献<sup>2)</sup> より引用。Copyright(c)2020 International Joint Conferences on Artificial Intelligence, All rights reserved.)

## 5. むすび

本稿では、既存の環境で得られた知識を活用して新たな環境における強化学習を効率化するアプローチとして転移強化学習を取りあげ、その概要を解説した。また、環境間で状態遷移確率分布が未知かつ異なる状況に適した転移強化学習手法として、MULTIPOLARについて紹介した。MULTIPOLARは転移元環境で獲得された方策関数を活用

するため、同環境において学習サンプルを大量に集める必要がない。また、行動空間が連続・離散いずれの場合においても、一貫した方法によって方策の転移を実現できる。

MULTIPOLARは、強化学習の中でも特にモデルフリー学習を対象とした手法である。一方、転移元で獲得されたモデルを転移先環境において適応的に統合するというアプローチ自体は汎用的であり、我々はモデルベース強化学習<sup>21)</sup>、教師あり学習<sup>22)</sup>、教師なし学習<sup>23)</sup>においてもその有効性を確認している。興味のある読者は参照されたい。

## 参考文献

- 1) 神島敏弘. 転移学習. 人工知能学会誌. 2010, Vol.25, No.4, p.572-580.
- 2) Barekatin, M. et al. "MULTIPOLAR: Multi-Source Policy Aggregation for Transfer Reinforcement Learning between Diverse Environmental Dynamics". Proceedings of the International Joint Conference on Artificial Intelligence. 2020, p.3108-3116. doi:10.24963/ijcai.2020/430.
- 3) Sutton, R. S.; Barto, A. G. Introduction to Reinforcement Learning. 1st ed., MIT Press, 1998.
- 4) Kober, J. et al. Reinforcement Learning in Robotics: A Survey. The International Journal of Robotics Research. 2013, Vol.32, No.11, p.1238-1274.
- 5) Gottesman, O. et al. Guidelines for Reinforcement Learning in Healthcare. Nature Medicine. 2019, Vol.25, No.1, p.16-18.
- 6) Mnih, V. et al. Human-Level Control through Deep Reinforcement Learning. Nature. 2015, Vol.518, No.7540, p.529-533.
- 7) Taylor, M. E.; Stone, P. Transfer Learning for Reinforcement Learning Domains: A Survey. Journal of Machine Learning Research. 2009, Vol.10, No.7, p.1633-1685.
- 8) Barreto, A. et al. "Successor Features for Transfer in Reinforcement Learning". Proceedings of the Annual Conference on Neural Information Processing Systems. 2017, p.4058-4068.
- 9) Gupta, A. et al. "Learning Invariant Feature Spaces to Transfer Skills with Reinforcement Learning". Proceedings of International Conference on Learning Representations. 2017, p.1-14.
- 10) Chen, T. et al. "Hardware Conditioned Policies for Multi-Robot Transfer Learning". Proceedings of Annual Conference on Neural Information Processing Systems. 2018, p.9355-9366.
- 11) Ammar, H. B. et al. "An Automated Measure of MDP Similarity for Transfer in Reinforcement Learning". Workshops at the AAAI Conference on Artificial Intelligence. 2014, p.31-37.
- 12) Song, J. et al. "Measuring the Distance between Finite Markov Decision Processes". Proceedings of International Conference on Autonomous Agents & Multiagent Systems. 2016, p.468-476.
- 13) Frans, K. et al. "Meta Learning Shared Hierarchies". Proceedings of International Conference on Learning Representations. 2018.
- 14) Parisotto, E. et al. "Actor-Mimic: Deep Multitask and Transfer Reinforcement Learning". Proceedings of International Conference on Learning Representations. 2016, p.1-16.
- 15) Rajendran, J. et al. "Attend, Adapt and Transfer: Attentive Deep

- Architecture for Adaptive Transfer from Multiple Sources in the Same Domain”. Proceedings of International Conference on Learning Representations. 2017, p.1-18.
- 16) Silver, T. et al. Residual Policy Learning. arXiv preprint arXiv:1812.06298. 2018.
  - 17) Johannink, T. et al. “Residual Reinforcement Learning for Robot Control”. Proceedings of International Conference on Robotics and Automation. 2019, p.6023-6029.
  - 18) Hill, A. et al. Stable Baselines. <https://github.com/hill-a/stable-baselines>. 2018.
  - 19) Haarnoja, T. et al. “Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor”. Proceedings of International Conference on Machine Learning. 2018, Vol.80, p.1861-1870.
  - 20) Schulman, J. et al. Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347. 2017.
  - 21) Tanaka, K. et al. “TRANS-AM: Transfer Learning by Aggregating Dynamics Models for Soft Robotic Assembly”. Proceedings of International Conference on Robotics and Automation. 2021, <https://ieeexplore.ieee.org/xpl/conhome/1000639/all-proceedings>, to appear.
  - 22) Ma, J. et al. “Adaptive Distillation for Decentralized Learning from Heterogeneous Clients”. Proceedings of International Conference on Pattern Recognition. 2020, p.7486-7492.
  - 23) Yonetani, R. et al. Decentralized Learning of Generative Adversarial Networks from Non-iid Data. arXiv preprint arXiv:1905.09684. 2019.

## 執筆者紹介



**米谷 竜 YONETANI Ryo**  
 オムロン サイニックエックス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門：コンピュータビジョン、機械学習  
 所属学会：情報処理学会、電子情報通信学会、  
 IEEE  
 博士（情報学）

本文に掲載の商品の名称は、各社が商標としている場合があります。

# ヒトの教示やインタラクションを活用した ロボット学習

濱屋 政志

近年の目覚ましい学習技術の発展により、ロボットが複雑な作業を自律的に達成できる可能性を示してきた。しかし、その背景には、ロボットが適切に動作するまでにパラメータや制御目標（報酬関数）の設計をはじめとする、無視できない手間や労力がある。さらに、たとえ制御器（方策）を適切に学習できても、外乱など環境の変化に関して脆弱であることが知られている。この問題に対して、ヒトの知識を活用し、ロボット学習をより容易にさせる研究が行われている。ヒトがロボットに教示を与える、あるいは物理的に誘導することによって学習を加速できる。我々は、失敗の教示や、物理的な外乱など、「矛盾した」情報をあえてヒトが与えることによって、さらに学習の性能を向上させる手法を開発した。本稿では、我々の最新の成果である「矛盾からのロボット学習手法」に関して解説する。

## Robotic Learning from Human Demonstrations and Interactions

HAMAYA Masashi

Thanks to recent learning techniques, robots can obtain skills for complex tasks automatically. Meanwhile, engineering efforts of tuning hyperparameters and designing appropriate reward functions are not trivial. Some researchers have leveraged human domain knowledge such as demonstrations and interactions to facilitate robotic learning. In this article, we introduce our recent works using contradictory information provided by humans. Concretely, we present learning from<sup>1)</sup> successful and failed demonstrations and<sup>2)</sup> advisory and adversarial interactions.

### 1. まえがき

近年の機械学習・強化学習の目覚ましい発展に伴い、ロボットに学習技術を適用し、自律的に作業戦略を獲得する研究が盛んになってきている。特に、深層学習による優れた汎化性や表現能力により、画像からロボットの行動を直接決定することができるなど、多感覚統合を実現することができる。また、解析的な表現が困難であるダイナミクスを持つシステムの制御に関しても学習技術は極めて有効である。

しかしながら、ロボットが適切に動作する背景には無視できない労力がある。たとえば、適切なハイパーパラメータや、報酬関数の設定などが挙げられる。これらの設定を実ロボットで試行錯誤的に検証することは、データを取得するための労力が大きい。特に、組立作業など対象物体とロボット間の接触を多く含む作業では、学習初期において

ロボットが予期せぬ動作によって、部品の破損やロボット緊急停止が生じる可能性がある。さらに、たとえロボットが無事に作業を学習できたとしても、学習された制御方策は、外力、ノイズ、モデル化誤差などの外乱に対して脆弱であることが知られている。

この問題を解決するためには、いくつかの例が挙げられる。一つ目は、本誌の別章で紹介した「サンプル効率の良い学習手法」を適用することである。たとえば、モデルベース強化学習などのアプローチは、複数の試行回数で所望のスキルを学習することができるため、設定したパラメータや報酬関数が理にかなっているかどうかを早い段階で評価できる。次に、シミュレータを使用し、学習した方策あるいはモデルを実世界に転移させることも考えられる。シミュレータにより、実世界におけるデータ収集の手間を大幅に削減できることが期待される。そして、これらに加えて、ヒトのドメイン知識を活用する方法が考えられる。たとえば、ヒトが教示を与える、動作の良し悪しを評価し、

Contact : HAMAYA Masashi masashi.hamaya@sinicx.com

ロボットにフィードバックを与える、あるいは、ロボットと物理的にインタラクションしながら、適切な動作を学習させる、などの方法が挙げられる。

このような作業を成功に導くヒトの教示やインタラクションに加えて、本稿では失敗の教示や敵対的なインタラクションをあえて与える「矛盾からの学習」を活用した我々の最新の研究成果を紹介する。さらに、物理的柔軟さを持つロボットを使用することで、教示やインタラクションが容易になることも併せて紹介する。第2節では、成功と失敗の教示を与えることで、失敗を回避し作業の成功率を向上させる学習手法<sup>1)</sup>について紹介する。第3節では、ロボットの学習中に、作業を成功に導く外力を与えることで、学習を促進させる誘導インタラクションと、作業を阻害する外力を与えることで、方策を頑健にさせる敵対インタラクションによる学習手法<sup>2)</sup>について紹介する。

## 2. 成功・失敗教示からの学習

本節では、図1に示すように成功と失敗の教示からロボットを学習させる手法について紹介する。ロボットが作業を学習する際、制御目標（報酬関数）の設計は極めて重要である。適切でない報酬関数を設計した場合、ロボットの学習性能が著しく低下する可能性がある。そこで、ヒトからの教示を活用して報酬関数を学習することを考える。

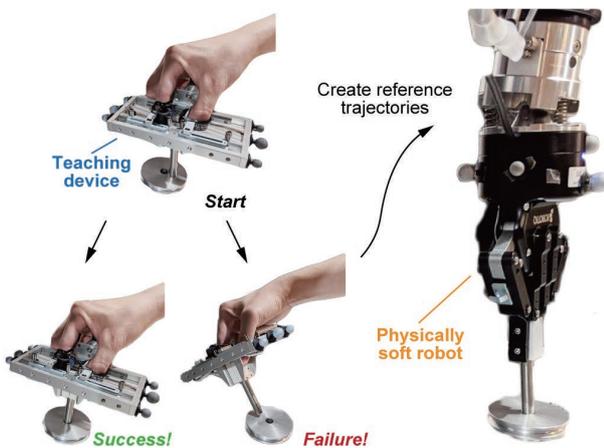


図1 成功・失敗教示からの学習 (文献<sup>1)</sup> から引用)

教示からの学習は、Learning from Demonstration と呼ばれており、2000年代から研究が盛んに行われている<sup>3)</sup>。教示の手法は、主にロボットに直接接触して教示させるダイレクトティーチングや、特殊なデバイスで遠隔で操作する手法が挙げられる。しかし、ダイレクトティーチングは、ロボットの可動域を考慮する必要があるため、精密な教示は困難である。また、ゲームパッドなどのデバイスを使用してロボットと同期的に教示を与えるためには、教示者の熟練が必要になる。

そこで、ヒトの動作を直接計測して、ロボットに与える

ことができれば、より直感的な教示が可能となるが、ロボットとヒトの身体（キネマティクス・ダイナミクスの）差が問題となる。この問題点に対して、我々はハードウェアからのアプローチを考えた。我々は、直感的に教示ができ、かつロボットとの身体差を軽減させることができる新しい教示デバイスを開発した<sup>1)</sup> (図2)。このデバイスの動作部はロボットのグリッパと形状が近いため、身体の違いを軽減させることができるのが特徴である。

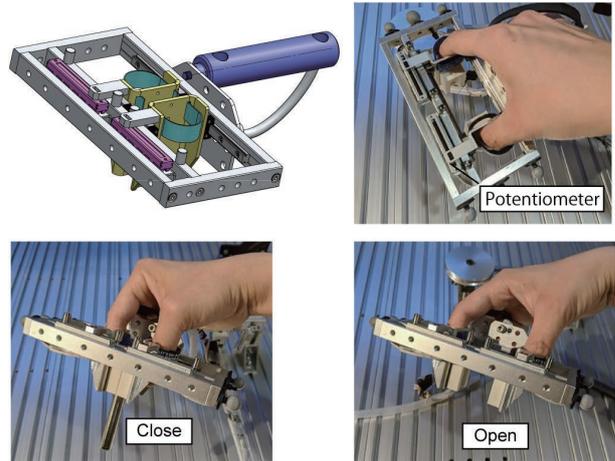


図2 教示デバイス (文献<sup>1)</sup> から引用)

さらに本研究では、柔軟手首を持つロボットを使用する。柔軟要素は環境との安全な接触を許容するので、教示デバイスから位置・姿勢の情報のみを与えるだけで、組立動作など接触を多く含む作業の教示が可能となる。つまり、教示デバイスと対象物体の接触力を計測する必要がなく、デバイスの軽量化・簡素化が期待できる。

提案した教示デバイスと柔軟なロボットにより、ヒトがより直感的に教示を行うことができるため、成功だけでなく、失敗の教示を与えることを考えた。ロボットの学習においては、過去の失敗から遠ざかるように拘束をかける手法<sup>4)</sup> や、失敗や不完全な教示から学習する手法が提案されている<sup>5)</sup>。これらの手法に対し、我々は、成功・失敗の両方の教示を与えて、失敗教示軌道なるべく避けるように、成功軌道を生成する手法を提案する。

### 2.1 目標軌道生成・ロボット動作学習

図3に本提案手法の手順を示す。本手法は、ロボットの強化学習問題と考え、報酬関数となる目標軌道の生成を、失敗・成功軌道の教示によって生成する。

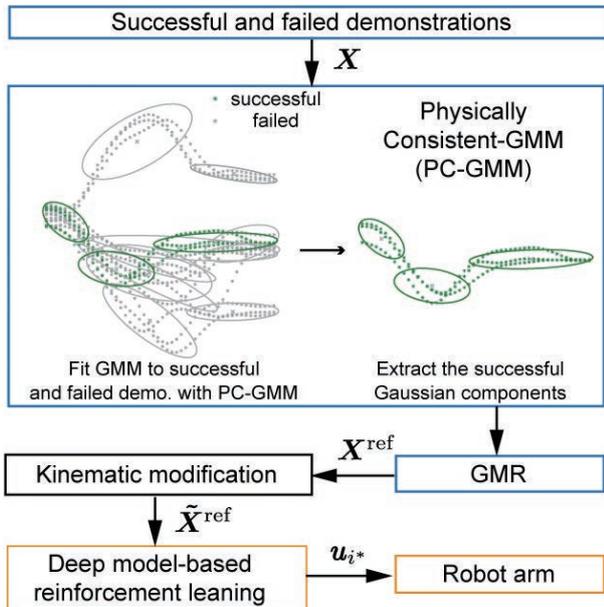


図3 提案手法の手順 (文献<sup>1)</sup> から引用)

まず、教示デバイスで、成功・失敗の教示データを収集し、これらのデータから目標軌道を生成する。目標軌道を生成するために、Physically-consisted Gaussian Mixture Model (PC-GMM)<sup>6)</sup> を適用する。GMM はガウス分布の線形重ね合わせで表現されるモデルであり、PC-GMM は、データにガウス分布を適合させるときに、データ同士の類似度を計算し、類似度が近いデータに同じガウス分布を割り当てるようにする。また、最適なガウス分布の数は、ノンパラメトリックベイズによって自動的に決定される。本研究では、失敗と成功データの類似度計算を行うときに、類似度を下げる項を導入し、成功・失敗データがそれぞれ離れるようにガウス分布が割り当てられるようにする。結果、失敗軌道を避けた状態で、成功軌道にガウス分布が適合される。そして適合された GMM から、連続軌道を生成するため、Gaussian Mixture Regression (GMR) を適用する。GMR によって、時間依存の目標の位置・姿勢の軌道を生成する。

次に、得られた軌道とロボットの動作軌道との誤差を報酬関数として、強化学習を行う。本研究では、ニューラルネットワークを使用したモデルベース強化学習<sup>7)</sup> を適用する。この手法は、ニューラルネットワークで学習された複数の順モデルを同時に学習し、モデルをランダムに選択しながら、動作予測することで、不確実性に対処でき、結果としてサンプル効率が向上する。また、行動選択には、Cross Entropy Method (CEM) を使用する。CEM は反復的に最適行動を計算する手法である。ランダムに生成した行動をモデルに入力し、数ステップ間の期待収益が最も高くなった行動上位数種類からガウス分布を生成し、次の反復からはそのガウス分布から行動の候補をサンプルする。

CEM は一様乱数的に行動の候補をサンプルするより効率的である。

## 2.2 実機実験

提案手法の有効性を検証するために、実機実験を行った。本実験の目的は、成功・失敗教示を両方使用することで、より高い作業成功率を示すことである。本研究では、ペグ挿入作業を課題とした。教示軌道は、成功軌道と4種類の失敗軌道をそれぞれ3回ずつ与えた (図4)。ロボットには柔軟手首<sup>8)</sup> を搭載した。教示デバイスと、ロボットの手先の位置姿勢を計測するために、モーションキャプチャシステムを使用した。比較手法は、成功・失敗教示を与えた場合 (success and failure)、成功教示のみ与えた場合 (success only)、教示を与えず、挿入穴の位置のみを与えた場合 (no demo) とする。



図4 成功・失敗軌道の教示例 (文献<sup>1)</sup> から引用)

図5に、学習時における、穴の位置との誤差を示す。左はxyz方向に、右図はz方向 (穴の挿入方向) における絶対平均誤差を示す。エラーバーは10回の学習実験における標準偏差を示す。この結果から、提案手法が最も小さい誤差を示した。また、学習後の作業成功率も、提案手法8/10、成功教示のみ5/10、教示無し0/10となり、提案手法が最も高い成功率を示した。成功軌道教示のみの場合は、ペグを穴の縁に当てる動作の際に、勢いが余り頻繁に穴を通り過ぎた。これに対し、提案手法は、この失敗を避けるように軌道が生成され、結果として高い成功率が示されたことが考えられる。

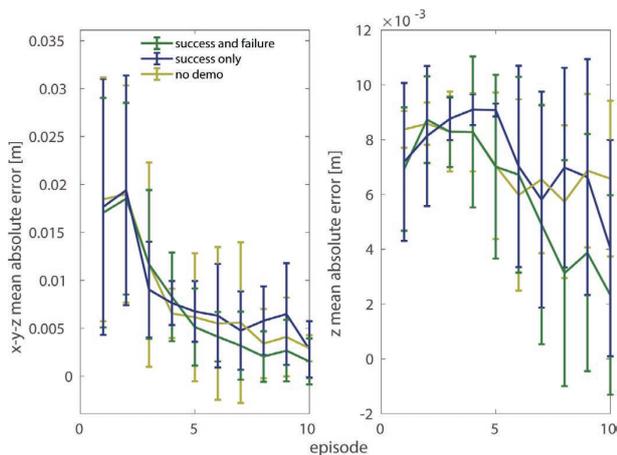


図5 学習時における位置誤差 (文献<sup>1)</sup> から引用)

以上により、成功と失敗の教示を活用することで、ベグ挿入作業において高い成功率を示すことが確認できた。今後は、提案手法の理論保証や、どのように失敗教示を与えるかという課題について取り組む予定である。

### 3. 物理インタラクションからの学習

本節では、誘導・敵対インタラクションからロボットを学習させる手法について紹介する。前節の教示からの学習で、失敗・成功教示を与えて学習させた手法は、教示を与えない場合より高い成功率を示した。しかし、適切な教示を与えるためには、熟練が必要となる。

熟練者でない場合においても、ヒトの知識を活用してロボット学習を促進させる方法はあるだろうか？一つの例として、ヒトのフィードバックを活用する方法が挙げられる。ロボットがある作業を遂行した後に、ヒトがロボットの作業の良し悪しを与える。ロボットは与えられたフィードバックをもとに、作業を改善するように学習する。方策を直接改善する手法<sup>9)</sup>、あるいは報酬関数を改善する手法<sup>10)</sup>が提案されている。しかしながら、作業後のフィードバックは、作業中のどの行動が良いか悪いかに与えることができないため、フィードバックの質が課題となる。

一方、たとえロボットが適切に作業を達成できたとしても、学習された方策・モデルは外乱、モデル化誤差に関して極めて脆弱であることが知られている<sup>11)</sup>。この問題に対して、学習中に敵対的な外力を加えて、より頑健な方策を学習する手法が提案されている。実ロボットに適用する際には、作業を行うロボットと外乱を加えるロボットを同時に学習させる<sup>12)</sup>手法があるが、複数台のロボットのセットアップのために大きな労力が生じる可能性がある。

これらの二つの問題を同時に解決するために、本研究では、ヒトの物理インタラクションを活用する手法を提案する。ヒトとロボットが常に接触している状態であるため、任意のタイミングでヒトがフィードバックとしてのインタラクションを発生させることができる。そして、本研究で

は、誘導 (Advisory)・敵対 (Adversarial) インタラクションを提案する (図6)。誘導インタラクションは、学習中の目標誤差を減らす役割を持ち、敵対インタラクションは学習方策を頑健にさせる役割を持つ。また、ヒトとロボットの安全な接触を保障するため、本研究においても柔軟要素を持つロボットを使用することを提案する。

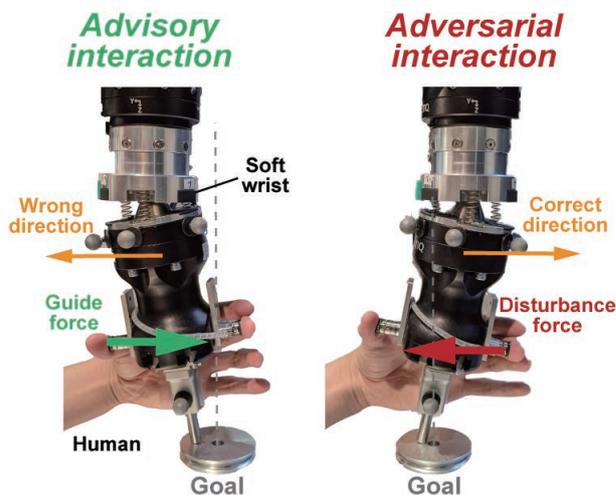


図6 物理インタラクションからの学習 (文献<sup>2)</sup> から引用)

#### 3.1 誘導・敵対インタラクションからの学習

本研究では、誘導・敵対インタラクションから学習する問題をモデルベース強化学習として定式化する。ヒトとロボットのインタラクションを考慮したダイナミクスモデルは  $s_{t+1} = f(s_t, a_t)$  で示される。ここで、 $s = [s^R, s^H]^T$  は状態、 $s^R$  はロボットの状態 (関節角度や手先位置など)、 $s^H$  はヒトの状態 (ヒトがロボットに与える外力)、 $a_t$  はロボットの行動 (関節トルクや、手先位置の目標速度指令など) である。本研究では、ヒトからのインタラクションをロボットに与えながらモデルを学習させる。報酬は  $r(s_t, a_t)$  で与えられ、累積期待収益を最大化させる行動を決定することがこの学習の目的となる。

誘導インタラクション：ロボットが作業中に、目標から離れるような間違った方向に動作したとき、ヒトが正しい方向に外力を加える。誘導インタラクションにおける報酬関数は、 $r(s_t, a_t) = r^R(s^R, a^R) - \alpha \|s^H\|^2$  で与えられる。ここで、 $r^R$  は、ロボットにおける報酬であり目標の位置誤差であり、 $\alpha$  は重み係数である。ヒトからの誘導外力が大きければ大きいほど、罰則が与えられるため、ロボットはヒトからの外力をなるべく受けないように動作することが期待される。つまり、ヒトから誘導インタラクションは、作業に使用される報酬の補助的な情報となっており、報酬の表現力が向上されると考えられる。その結果、学習中の目標状態との誤差を軽減できると期待される。

敵対インタラクション：誘導インタラクションによって

ロボットがある程度学習された後に、敵対インタラクションを適用する。ここでは、ロボットが目標に正しい方向に動作した場合、その方向とは逆にヒトが外力を加える。報酬は  $r(s_t, a_t) = r^R(s^R, a^R) - \beta \|s^H\|^2$  で与えられる。ここで、 $\beta$  は重み係数である。大きな敵対外力を受けるほど、報酬も大きくなり、結果外乱に対して頑健な方策が学習される。

モデルベース強化学習は前節と同じアルゴリズムを使用する<sup>7)</sup>。ヒトとロボットのインタラクションを含む複雑なモデルに対しても、少ない試行回数で方策を学習できると期待される。

### 3.2 シミュレーション実機実験

提案手法の有効性を検証するため、シミュレーションを行った。シミュレーションの目的は、提案手法がインタラクションを用いない場合と比較して、学習中の誤差を減少させ、頑健な方策を獲得できるかを確認することである。本研究においても検証対象としてペグ挿入課題を扱う。

本実験には6人の被験者が参加した。被験者はモニターに表示されるロボットに、ゲームパッドで外力を与えた(図7)。ロボットシミュレータは、PyBullet という動力学シミュレータを使用して設計された。ロボットには、柔軟手首<sup>8)</sup>が搭載されており、柔軟要素は6自由度のPID制御によって再現されている。



図7 シミュレーションの概要 (文献<sup>2)</sup> から引用)

本シミュレーションでは、1回の学習実験で、7回の誘導インタラクション、7回の敵対インタラクション、計14回行った。評価として、インタラクションなし (no interaction)、ランダム外力からインタラクション (random interaction)、経験則で設計した誘導・敵対インタラクション (heuristic interaction)、提案手法の4種類の状況で性能を比較した。

図8に学習中の目標位置との平均誤差を示す。橙色の線はインタラクションなし、黄色はランダムインタラクション、緑は経験則インタラクション、青は提案手法である。

提案手法を使用した結果、誘導インタラクションにおいて、学習中の誤差を減少させることができた。また、敵対インタラクションにおいても、提案手法の誤差が一度増加するが、最終試行においては、インタラクションを与えないあるいはランダムなインタラクションを与えた場合よりも小さい誤差を示した。

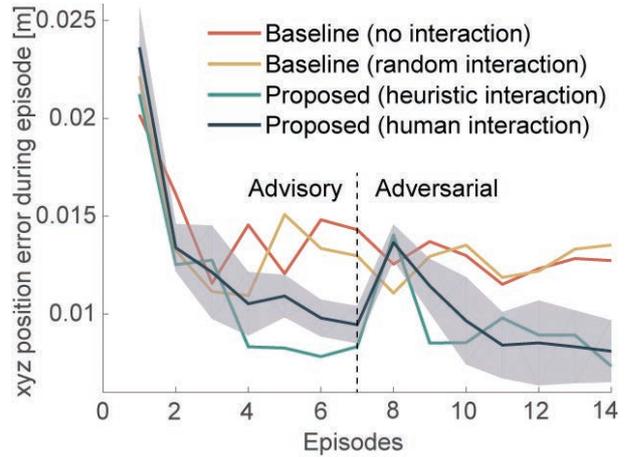


図8 目標位置の誤差 (文献<sup>2)</sup> から引用)

次に、学習された方策の頑健性を検証するため、ロボットの質量や、穴の摩擦係数を変更してテストを行った。図9に未知環境におけるペグ挿入作業の成功回数を示す。この結果から、提案手法はすべての場合において、最も高い成功率を示した。以上の実験から、提案手法の有効性が示された。

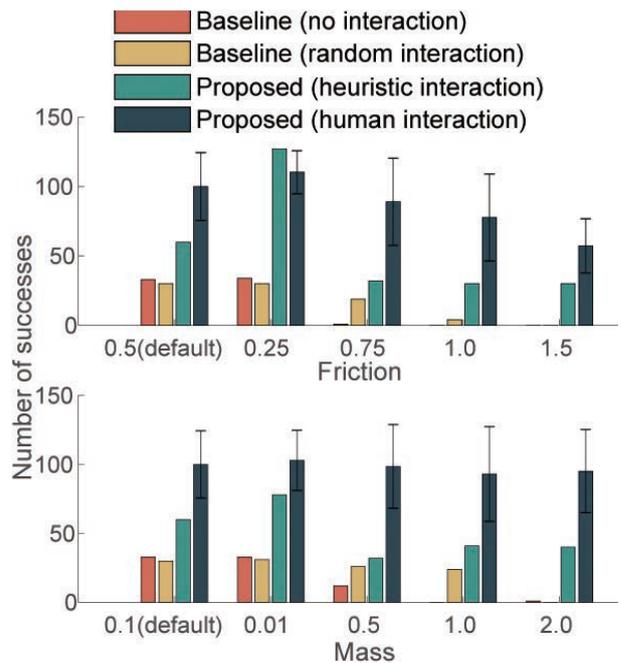


図9 未知環境での作業成功回数 (文献<sup>2)</sup> から引用)

### 3.3 実機実験

最後に、実機を使用した実験を行った。提案手法が実世界においても適用可能かを検証するのが目的であった。

実機実験においても柔軟手首<sup>8)</sup>を使用した。ロボット及び学習のセットアップは前節と同じであるが、ヒトの外力を計測するために、触覚センサを搭載した。ロボットを様々な方向から学習させた場合と、異なる素材のペグを適用したときの作業成功率を、インタラクションを与えない場合と比較した。

図 10 に、穴の 3 方向からペグ挿入作業を行った場合と、ある 1 方向から、様々な素材のペグで作業を行った時の成功回数を示す。橙色の棒グラフはインタラクションを与えない場合で、青は提案手法である。この結果より、提案手法はすべての場合において高い成功回数を示した。以上より、実機実験においても提案手法の有効性を示した。

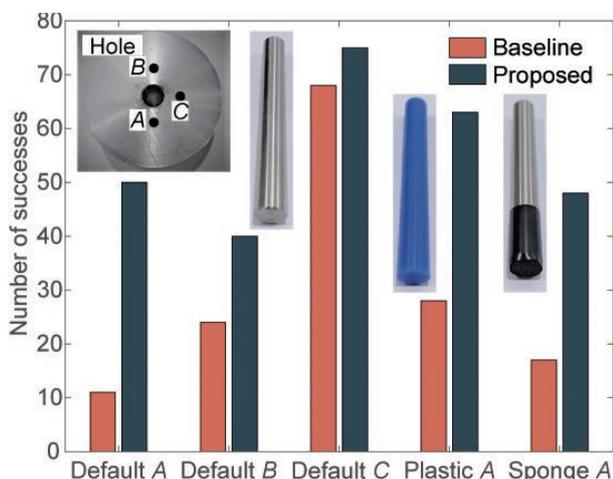


図 10 実機実験における作業成功回数 (文献<sup>2)</sup> から引用)

今後の課題として、更なる性能向上のために、修正・敵対インタラクションのより適切な順序やユーザビリティを向上させる手法を検討する予定である。

## 4. むすび

本稿では、ヒトの知識を活用して、ロボット学習をより容易にする我々の最新の研究成果について紹介した。核心となるアイデアは、「矛盾した」情報を与えることで、ロボット学習をより促進させるという点である。1つ目の研究事例として、ヒトからの教示を利用する際に、成功軌道だけでなく、失敗軌道も陽に活用することで失敗を避けるような目標軌道生成手法を紹介した。2つ目の事例として、ヒトがロボットに直接触れて学習する際、誘導インタラクションを活用することで、学習中の誤差を減少させ、学習後の方策を頑健にすることを紹介した。また、これらの作業を行う際に、対象物体やヒトとの安全な接触のために、柔らかさを持ったロボットが重要な役割を持つことも

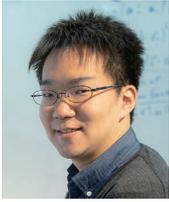
注目されたい。

本稿を読み、ヒトの知識を活用して学習を促進させる研究に興味を持っていただければ幸いである。

## 参考文献

- 1) Hamaya, M.; von Drigalski, F.; Matsubara, T.; Tanaka, K.; Lee, R.; Nakashima, C.; Shibata, Y.; Ijiri, Y. "Learning Soft Robotic Assembly Strategies from Successful and Failed Demonstration". IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020, p. 8309-8315.
- 2) Hamaya, M.; Tanaka, K.; Shibata, Y.; von Drigalski, F.; Nakashima, C.; Ijiri, Y. Robotic Learning From Advisory and Adversarial Interactions Using a Soft Wrist. IEEE Robotics and Automation Letters. 2021, Vol.6, No.2, p. 3878-3885.
- 3) Billard, A.; Calinon, S.; Dillmann, R.; Schaal, S. Survey: Robot programming by demonstration, Handbook of Robotics. Springer, 2008, Chapter 59, p.1371-1394.
- 4) Esteban, D.; Rozo, L.; Caldwell, D. G. "Learning deep robot controllers by exploiting successful and failed executions". IEEE RAS International Conference on Humanoid Robots, 2018, p. 1-9.
- 5) Gao, Y.; Xu, H.; Lin, J.; Yu, F.; Levine, S.; Darrell, T. Reinforcement learning from imperfect demonstrations. International Conference on Learning and Representation. arXiv preprint arXiv:1802.05313.
- 6) Figueroa, N.; Billard, A. "A physically-consistent bayesian non-parametric mixture model for dynamical system learning". Conference on Robot Learning, 2018, p. 927-946.
- 7) Chua, K.; Calandra, R.; McAllister, R.; Levine, S. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. Advances in Neural Information Processing Systems. 2018, Vol.31, p. 4754-4765
- 8) Tanaka, K.; von Drigalski, F.; Hamaya, M.; Lee, R.; Nakashima, C. Shibata, Y.; Ijiri, Y. "A Compact, Cable-driven, Activatable Soft Wrist with Six Degrees of Freedom for Assembly Tasks". IEEE/RSJ International Conference on Intelligent Robots and Systems. 2020, p. 8752-8757.
- 9) Celemin, C. E.; Maeda, G.; Ruiz-del Solar, J.; Peters, J.; Kober, J. Reinforcement learning of motor skills using policy search and human corrective advice. International Journal of Robotics Research. 2019, Vol.38, No.14, p. 1560-1580.
- 10) Biyik, E.; Palan, M.; Landolfi, N. C.; Losey, D. P.; Sadigh, D. "Asking easy questions: A user-friendly approach to active reward learning". Conference on Robot Learning, 2020, p. 1177-1190.
- 11) Ilahi, I.; Usama, M.; Qadir, J.; Janjua, M. U.; Al-Fuqaha, A.; Hoang, D. T.; Niyato, D. Challenges and countermeasures for adversarial attacks on deep reinforcement learning. arXiv preprint arXiv:2001.09684, 2020.
- 12) Pinto, L.; Davidson, J.; Gupta, A. "Supervision via competition: Robot adversaries for learning tasks". IEEE International Conference on Robotics and Automation, 2017, p. 1601-1608.

## 執筆者紹介



濱屋 政志 HAMAYA Masashi  
オムロン サイニックエックス株式会社  
リサーチアドミニストレイティブディビジョン  
専門：ロボティクス、機械学習、強化学習  
所属学会：日本ロボット学会、IEEE  
博士（工学）

---

本文に掲載の商品の名称は、各社が商標としている場合があります。

# 視覚と言語の共有潜在表現獲得による モノづくり手順理解

橋本 敦史

組立作業や創薬実験、生化学実験を始め、複数のものを組み合わせ、様々な工程を経て価値の高いプロダクトを作り上げる作業は、わかりやすい価値創出の方法であろう。このような作業を機械に理解させることができれば、該当のシーン検索から手順書の生成、果てはロボットによる作業の代替まで、機械との多様な協働方法が実現できる。このような理解の1つの形として、作業を観測した視覚情報データを手順書中の教示テキストと対応付ける視覚言語統合技術がある。特に視覚情報データとして画像列を入力として教示テキストを生成する課題において、筆者らが提案した一連の手法は State-of-the-Art となっている。本稿ではこれらの成果を中心に視覚情報と言語情報の統合による作業理解のための機械学習処理を解説する。

## Cross-modal Representation Learning for Understanding Manufacturing Procedures

HASHIMOTO Atsushi

Creating high-value products through combining materials and processing them, including assembly work, drug discovery experiments, and biochemical experiments, would be one of the primary methods of value creation. If computers can understand such activities, we can collaborate with machines in various directions, such as relevant video scene retrieval, procedural text generation, and eventually substituting work by robots. As one form of such understanding, there are the vision and language arts to extract shared latent representation between visual observation and textural description of manufacturing. In particular, a series of our methods for procedural text generation from image sequence are state-of-the-art. With the central focus on these methods, this paper explains the recently-developed vision and language technologies for understanding manufacturing procedures.

### 1. はじめに

本稿では機械に手順を理解させるための視覚言語統合 (Vision&Language) 技術について解説する。まず「手順を理解する」とはどういう状態だろうか？ 1980年代に提起された「中国人の部屋問題<sup>\*1</sup>」は記号論理の世界に閉じた情報処理が「理解」として十分であるかを哲学的な側面から問いかけた。例えば、もし、中国語の部屋の中の人に言葉で回答できる質問ではなく、行動することを要求したらどうだろうか？ 辞書だけでは行動を起こすことはできないだろう。行動をするということは、言葉に対応する行為を観測し、対応関係を得る必要がある。視覚言語統合はこのような対応関係を得るための技術である。我々は、この技術を用いることで「手順の理解」、すなわち、実世界で生じる事象と自然言語で記述される記号世界を結びつけるこ

とを目指している。自然言語と視覚情報の対応付けの試みとしては、言語指示をアニメーションでシミュレートするというものが1988年に既に行われていたようである<sup>9)</sup>。残念ながらこの時代のシステムは基本的にルールベースであり、言語や視覚情報に多くの制約を課す箱庭的なシステムとなっている。このような制約なしに機械が視覚情報と言語情報を対応付けられるようになると何ができるようになるだろうか？ まず、作業を観測したデータの中から、言語で表現された手順書の各工程に対応する箇所を自由に検索することができるようになる。多様な作業員に対して、ある工程がどのように行われるかを横断的に解析することができれば、誤った方法を取っている作業員に働きかけて効率的に教育をすることができるようになる。また、作業を観測したデータから、その自然言語表現や明確なフォーマットに従ったスクリプトを生成できるようになる。これにより、個人差なく作業手順を一定の品質で書き

Contact : HASHIMOTO Atsushi atsushi.hashimoto@sinicx.com

出すことができるようになるだろう。例えば、現在研究ノートに手作業で書き留められるような生命科学分野の実験手順を正規化された手順書として書き出せれば、それは最新研究の再現性の問題に寄与する。同様に、個々の製品の組立状況を書き出すことは新たなトレーサビリティにもなり、品質保証の向上にも資するだろう。究極的には、言葉だけで誰でもロボットに作業を依頼できるようになるかもしれない。

このような未来を実現するには、作業と、その作業内容に対応する言語表現が同じ点に射影されるような共通潜在特徴空間を得ることが重要になる (図1)。このような空間を得ることができれば、空間中の近傍点を探すことで、テキストをクエリにして映像を (または映像をクエリにしてテキストを) 検索できる。また、映像を射影した点から、その点に対応するであろうテキストを復号 (デコード) することで手順書を生成できる (またはテキストから映像を生成できる)。本稿では、このような共有潜在空間の獲得と、この空間を用いた応用研究について解説する。

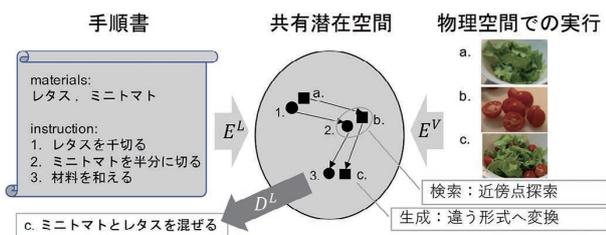


図1 手順書と観測データの共有潜在空間への埋め込み

## 2. 関連研究

### 2.1 視覚言語統合処理

静止画や動画と言語の対応を取る研究は深層学習の登場により急速に発展した分野であり、一般には Vision & Language と呼ばれる。まだ定着した邦訳は存在しないが、本稿では筆者の過去の著作<sup>25)</sup>に合わせ、これを視覚言語統合\*と呼ぶ<sup>2)</sup>。視覚言語統合は深層学習に伴い発展した分野ではあるが、その源流は深層学習以前から連綿と研究されてきた Web 上での画像検索課題に遡ることができる。

画像検索課題はテキスト  $x^L$  を  $D$  次元潜在特徴  $z^L$  へ写像する関数  $E^L: x^L \rightarrow z^L$  (言語符号化器) と、画像 (または映像)  $x^V$  を  $D$  次元潜在特徴  $z^V$  へ写像する関数  $E^V: x^V \rightarrow z^V$  (視覚符号化器) を得ることで実現される (図1)。学習は同じ内容を示す  $\{x^L, x^V\}$  の対からなるデータセットを用いて、対照性損失\* (Contrastive Loss) や三重項損失\* (Triplet Loss) といった距離学習用の損失関数を最小化することで実現される。

検索課題は言語や画像といったモダリティを超えた共通の情報 (key feature) を見つけることができれば解くこと

ができた。これに対して、生成課題は共通部分だけでなく、生成対象のモダリティに含まれる情報すべてを予測しなければならないという点で検索よりも挑戦的な課題であり、特に画像からテキストを生成する「キャプション」は深層学習による視覚言語統合の代名詞ともなっている。キャプションの学習データは検索課題同様、 $\{x^L, x^V\}$  の対となるが、 $E^V$  に加えて、潜在特徴からテキストを復号する関数  $D^L: z^V \rightarrow x^L$  を学習する (図1)。単純にはテキスト中の単語をカテゴリとして、カテゴリ系列の一致度最大化により学習を行う。

上述の課題において、潜在特徴  $z^{L/V}$  は  $x^L$  と  $x^V$  のいずれとも対応づくことから、(複数のモダリティの) 共有潜在特徴と呼ばれる (図1の●や■の点に相当する)。また、このような  $z^{L/V}$  の集合が張る空間が図1に示した共有潜在空間となる。

### 2.2 手順理解に関する研究

手順とは目的を達成するために実行すべき工程の内容だけでなく、その実行順序の依存関係を記述したものとされる。したがって、手順は前節で述べたような単一の点ではなく、点と点の関係を示す枝を含む構造で表現される。

テキストや動画は基本的には同一軸上で順序付けられた、つまりシリアルライズされたデータである。このため、これらは単純には図1に示したように直線的な遷移により表現される。

視覚情報は伴わないモノモーダルな自然言語においては、このような遷移を追跡し、工程ごとに引き起こされる材料の状態変化を潜在空間中の点の移動として表現し、それを追跡可能にする Neural process network (NPN)<sup>1)</sup> という研究が知られている。この研究では、例えば「トマトを洗って、串切りにする」といった指示に対して「トマト」、「洗ったトマト」、「串切りにされたトマト」のそれぞれを潜在空間中の異なる点として表現する。また、「トマト」に対して「洗う」という工程を行うことで、潜在空間中のどこへ移動するかも予測する。これは、いわば潜在空間中の作業のシミュレーションとみなすことができる。

NPN はあくまでも材料の特徴変化を追跡する、つまり手順書に示された順序に従った図1のような直線的な遷移を予測するものである。一方で、手順書は本質的には手順の入れ替えが一部許された半順序構造を持つ。例えば図中のレシピではトマトとレタスのどちらから取り組んでも差し支えない (工程1と工程2には順序の依存性がない)。一方で、工程3は工程1、2の結果得られるものを混ぜる工程であり、両方終了していないと取りかかれない。このような「材料を混ぜる」ことによって生じる順序の依存関係は図2のようなグラフで表現される。自然言語処理による手順の理解には、このようなグラフ構造を抽出するものも存在する。

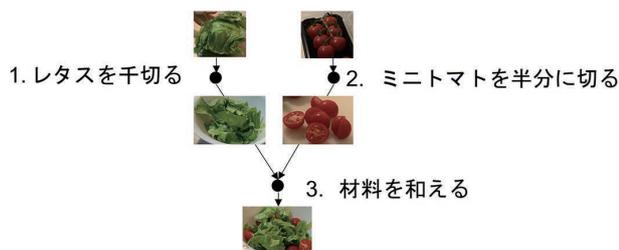


図2 順序の依存関係のグラフ表記

このような手順に関する研究は、Web上で大量に入手可能な調理レシピを対象に行われるのが主流である。これは調理レシピのユーザ数が、化学実験の手順書のユーザに対して圧倒的に多いこと、および、工業製品の組立手順書に対して機密性が低いことに起因している。

例えば、グラフ構造の抽出については、その構造の定義が粗い順から、SIMMR<sup>8)</sup>、Action Graph<sup>10)</sup>、Flow Graph Corpus<sup>12,23)</sup>がある。いずれも、構造の定義だけでなく、そのような構造がアノテーションされたレシピのコーパスが公開されているが、このうち、特にFlow Graph Corpusは筆者を含む日本の研究グループによるもので日本語<sup>12)</sup>と英語<sup>23)</sup>のコーパスがそれぞれ公開されている。

### 2.3 クロスモーダル処理による手順理解研究

クロスモーダル処理による手順理解研究の先駆けは、深層学習が登場する以前の2011年に行われた道満らによるレシピと調理映像の整合\* (alignment)<sup>26)</sup>であろう。この研究では物体認識や動作認識結果を駆使してテキストに書かれたレシピの各工程に対応する映像区間を特定している。

深層学習以前であるため、共有潜在空間は図1のような連続的なものではなく、物体認識結果や動作認識結果のone-hot-vectorによる離散空間が用いられている。

これに対して、料理に関連したクロスモーダル処理で近年もっとも活発に研究されているのはim2recipeと呼ばれる、完成した料理の画像からそのレシピを特定する研究であろう。この研究の端緒は画像から対応するレシピを検索するもの<sup>18)</sup>であったが、その精度向上を主な目的として、レシピに応じた料理画像の生成<sup>24)</sup>を行う手法などが登場している。さらにTransformerの登場によって、これらの研究はさらにレシピ検索からレシピ生成へと拡張されてきている<sup>17)</sup>。この中では、画像のみから前節で述べたようなグラフ構造の推定<sup>22)</sup>を行うことにより精度向上を目指した手法なども提案されている。

しかし、完成した料理を見るだけでレシピを推定する問題は人間にも難しく、実応用を想定するにはやや不良設定な問題に思える。これらの研究に対して、我々は調理の過程を撮影した画像列や動画を入力としたレシピ生成に取り組んでおり、本稿では、それら2つの入力を対象とした最近の研究を紹介する。

### 3. 画像列からの手順書生成

一般的な視覚言語統合において、画像列を入力とした文章生成は視覚的叙述生成\* (Visual Storytelling)<sup>7)</sup>として知られている<sup>11)</sup>。しかし、視覚的叙述生成は、対象とする画像列が物語であり、どこまで逸脱した内容を正解として良いか定めることが難しい。また、複数の材料を追跡するモノづくりと異なり、物語の軸となる登場人物は絞られている。これに対してChanduらは調理を対象とし、手順書と画像列からそれぞれ作業の状態遷移を有限オートマトン上でシミュレートし、その状態が一致するような損失を導入することでレシピ生成の精度向上を測る手法を提案した<sup>2)</sup>。これは状態の遷移をモデル化し、その一致度を高めることで工程の進行度を反映した共有潜在空間を得ることを目指したものと解釈できる。

これとは独立に筆者を含む研究グループは大量のレシピ・画像列の対があるCookpad Image Dataset<sup>4)</sup>を利用し、検索課題による事前学習を利用するRetrieval Attention (RetAttn)<sup>14)</sup>という手法を提案した(図3)。この手法では、工程の違いまで含めて正確に検索できるようにモデルを予め訓練することで、工程の進行度を反映した共有潜在空間を得る。同時に、得られた検索機能を利用し、画像に対応する文書の特徴を参照することで、画像と文書という異なる入力に起因する潜在特徴分布の偏り(モダリティギャップ)による悪影響を低減することで高い性能を達成している(詳細は次に紹介する手法と併せて表1で述べる)。

さらに、同グループは中間状態としてグラフ構造を推定し、そのグラフ構造をテキスト生成に積極的に用いる手法<sup>15)</sup>を提案し、さらなる精度向上を実現した(図4)。このモデルでは画像列から図2に示す依存関係を推定する。この推定結果を二値化して木構造を得たのち、デコーダにTreeLSTM<sup>21)</sup>を採用し、推定された木構造を組み込んだテキスト生成を計算する。木構造の学習は半教師あり学習で行った。具体的にはSIMMRを参考に、そのクロスモーダル版であるVisual SIMMRデータセットを作成し、全体の1%程度のデータには木構造の正解を付与することで学習を行った。正解が存在しないデータに対しては、画像列から推定された木構造と生成されたテキストから推定される木構造が一致するような自己教師学習(Tree Re-prediction)を適用した。

表1に、これらの手法の定量的評価を掲載する。テキスト生成課題では、明確に1つの正解が定まらず、また全てのありえる正解の列挙も難しいことから単語の重複などを考慮した複数の指標(ワードオーバーラップ指標)によって近似的な評価が行われており、この表もその方法に従って得られたものとなっている。なお、実際に生成されたテキストを含む定性的評価を含む詳細な議論について我々の論文<sup>14,15)</sup>を参照いただきたい。

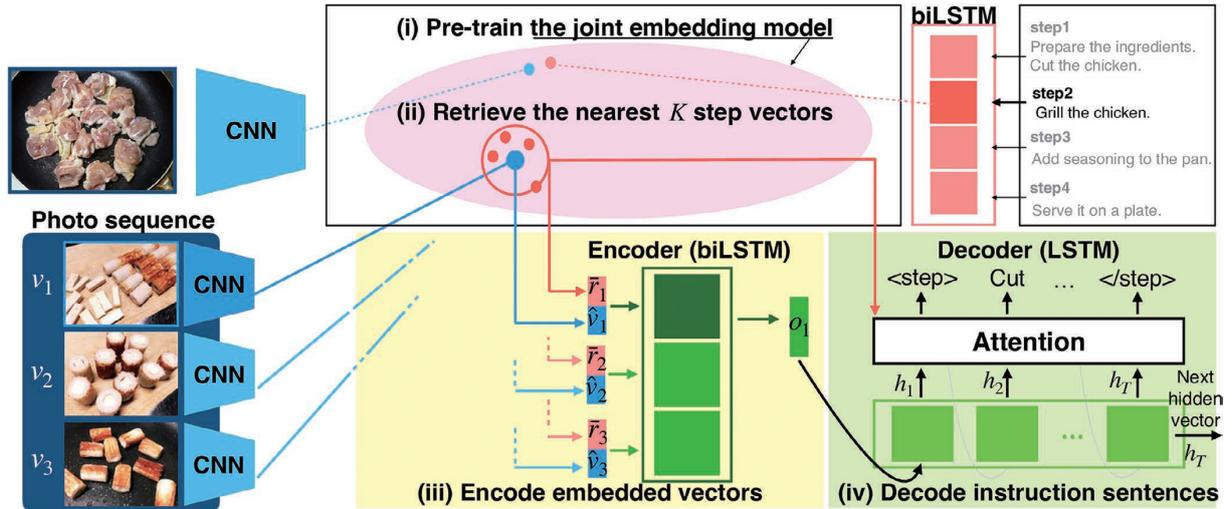


図3 Retrieval Attentionの概要。各画像から得られる潜在特徴と手順書から得られる各工程の潜在特徴が近くなるような距離学習を行い、共有潜在空間を学習する。テキスト生成では、各画像の潜在特徴に加えて、その画像の近傍にある（学習データ中の）手順書の特徴を平均したものを入力としてテキストを生成する（文献14）より引用

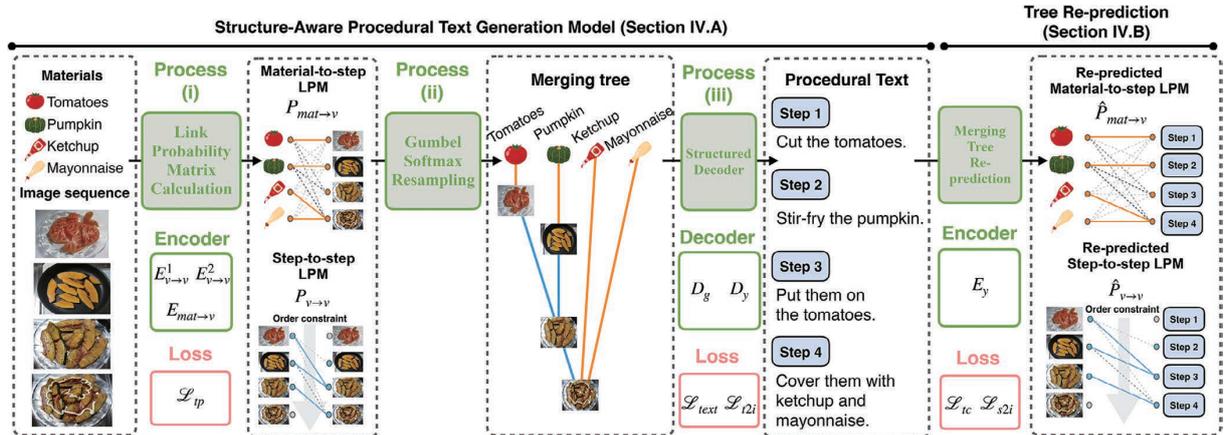


図4 依存関係を考慮したテキスト生成モデルの概要。入力材料リスト（記号の集合）と画像列となる。Process (i) では材料-画像、および、画像-画像間の依存関係を推定する。Process (ii) ではGumbel Softmax Resamplingというテクニックを用いて微分可能な形で推定結果を離散化し、木構造を得る。Process (iii) ではTree LSTMにより木構造を反映したテキスト生成計算を行う。最後にTree Re-predictionでは生成テキストからも木構造を推定し、画像から推定された木構造との一致度が高くなるような自己教示型の木構造推定を行う（文献15）より引用

表1 3つのベースライン手法と、それぞれを<sup>15)</sup>による構造考慮型モデルに拡張したモデルに対するワードオーバーラップ指標による評価（+Halfと+Fullは図4のTree Re-prediction部の有無の違い）。B\$=\$BLEU, RL\$=\$ROUGE-L, D\$=\$Distinctであり、食材や動作に相当する単語のみを抜き出した単語列による評価をさらにI\$=\$Ingredient, Ac\$=\$Actionを付して表記している。また手法別の比較で最良のスコアを太字で、手法内の比較で最良のスコアを下線で表している（文献15）より抜粋

	B1	B4	RL	D1	D2	I-B1	I-B4	I-RL	Ac-B1	Ac-B4	Ac-RL
GLAC Net [11]	28.5	5.9	21.4	46.6	69.0	9.5	0.3	13.2	21.2	2.1	22.8
+Half	28.5	5.9	<b>21.8</b>	46.7	68.8	10.5	<b>0.4</b>	15.6	21.0	2.3	<b>23.1</b>
+Full	28.9	6.1	21.3	47.2	69.9	11.8	<b>0.4</b>	16.3	21.0	2.3	22.5
SSiD [2]	28.7	6.0	20.9	45.5	66.6	8.6	0.2	13.1	20.1	2.1	21.6
+Half	30.3	6.2	20.8	43.9	65.1	11.4	<b>0.4</b>	16.4	19.8	<b>2.2</b>	22.1
+Full	<b>31.1</b>	<b>6.4</b>	<b>21.6</b>	<b>48.3</b>	<b>71.0</b>	<b>12.9</b>	<b>0.4</b>	<b>16.7</b>	<b>21.5</b>	<b>2.2</b>	<b>23.5</b>
RetAttn [15] (Ours)	32.2	6.5	21.6	40.2	60.3	11.2	<b>0.3</b>	14.5	22.1	9.0	21.2
+Half	32.2	6.5	21.8	52.4	77.8	11.9	<b>0.3</b>	<b>14.8</b>	21.8	9.4	22.9
+Full	<b>33.2</b>	<b>7.1</b>	<b>22.1</b>	<b>52.7</b>	<b>78.6</b>	<b>12.1</b>	<b>0.3</b>	<b>14.8</b>	<b>22.3</b>	<b>9.5</b>	<b>23.1</b>

各手法単独では RetAttn が最も性能が高く、この手法が最高性能を達成していることがわかる。また、構造考慮型モデルへの拡張がモデルを問わず性能向上に寄与することも読み取れる。

この他、同グループではデータ構造という観点では画像列中の食材矩形と、Flow Graph 中の要素（つまり調理途中状態の食材）との対応関係をアノテーションした Recipe FlowGraph Bounding Box (RFG-BB) データセット<sup>16)</sup>も公開しており、今後の活用を検討している。

#### 4. 動画と手順書のクロスモーダル処理

前節までで紹介してきた画像列を入力とした手順理解では「材料の変化」を追跡することにその主眼が置かれてきた。これに対して、動画があれば、「材料の変化」を引き起こす作業者の動作そのものも観測対象となり、「動作」と「材料の変化」の関係性獲得が技術的な挑戦となる。

この問題に対して Naim らは化学実験のように手順が厳格に定義され、図 2 に示した順序の曖昧性は存在しない状況を仮定し、教師なしでの手順書と動画中のシーンの対応付けを目指した<sup>13)</sup>。この研究では、把持された物体と文書中の名詞との対応を隠れマルコフモデルにより確率的に定式化し、各シーンでの名詞と物体の対応関係の一貫性と、映像全体を通した手順の再現性の双方が最尤となる対応付けを EM アルゴリズムによって推定する手法を提案している。また、調理を対象として、深層学習を用いながらも基本的な発想としては類似した取り組みが後年に Stanford 大の研究グループによって行われている<sup>3,5,6)</sup>。残念ながら精度という観点では、これらの取り組みはあくまで実験的なものと言わざるをえない結果となっている。これは、おそらく教師情報やデータ量の不足が根本的な原因であろう。

これらは手順書と作業映像との対応付けに関するものであったが、編集済みの動画から手順書を生成する取り組みも発表されている<sup>19,20)</sup>。ただし、これらの研究では、手順書を生成するために、映像中の音声書き起こしも入力に加えており、この書き起こしから得られる言語的な情報に大きく依存したモデルとなっている。このような外部情報を削減しながら、映像から抽出できる情報を増やしていくことが、モノづくりに対するクロスモーダルな手順理解研究にとっての次の課題となっている。

#### 5. むすび

本稿では視覚言語統合技術を用いて、手順の共有潜在表現を抽出するための諸技術を概観した。現状ではデータセットの不足や重要シーンのみを抽出することの困難さから、人が編集し、重要シーンだけに分割されたような動画に対する技術しか研究が進んでいない。また、組み合わせによる特徴変化を追跡するモデルは構築できているものの、単独で材料の状態を変化させるような作業に対して、

その状態変化を追跡する技術も未解決である。これらについては現在、問題解決に向けての取り組みをそれぞれ進めている。

#### 参考文献

- 1) Bosselut, A.; Ennis, C.; Levy, O.; Holtzman, A.; Fox, D.; Choi, Y. "Simulating action dynamics with neural process networks". ICLR. 2018.
- 2) Chandu, K.; Nyberg, E.; Black, A. W. "Storyboarding of recipes: grounded contextual generation". Proc. ACL. 2019, p.6040-6046.
- 3) Chang, C.-Y.; Huang, D.-A.; Sui, Y.; Li, F.-F.; Niebles, J. C. "D3TW: Discriminative differentiable dynamic time warping for weakly supervised action alignment and segmentation". CVPR. 2019.
- 4) Harashima, J.; Someya, Y.; Kikuta, Y. "Cookpad image dataset: An image collection as infrastructure for food research". SIGIR. 2017, p.1229-1232.
- 5) Huang, D.-A.; Buch, S.; Dery, L.; Garg, A.; Li, F.-F.; Niebles, J.C. "Finding "it": Weakly-supervised reference-aware visual grounding in instructional videos". CVPR. 2018.
- 6) Huang, D.-A.; Lim, J. J.; Li, F.-F.; Niebles, J. C. "Unsupervised visual-linguistic reference resolution in instructional videos". CVPR. 2017.
- 7) Huang, T.-H. K.; Ferraro, F.; Mostafazadeh, N.; Misra, I.; Agrawal, A.; Devlin, J.; Girshick, R.; He, X.; Kohli, P.; Batra, D.; Zitnick, C. L.; Parikh, D.; Vanderwende, L.; Galley, M.; Mitchell, M. "Visual storytelling". NAACL-HLT. 2016, p.1233-1239.
- 8) Jermurawong, J.; Habash, N. "Predicting the structure of cooking recipes". EMNLP. 2015, p.781-786.
- 9) R. F. Karlin. "Defining the semantics of verbal modifiers in the domain of cooking tasks." ACL. 1988, p.61-67.
- 10) Kiddon, C.; Ponnuraj, G. T.; Zettlemoyer, L.; Choi, Y. "Mise en place: Unsupervised interpretation of instructional recipes". EMNLP. 2015, p.982-992.
- 11) Kim, T.; Heo, M.-O.; Son, S.; Park, K.-W.; Zhang, B.-T. GLAC Net: Glocal attention cascading networks for multi-image cued story generation. arXiv preprint arXiv:1805.10973, 2018, (参照 2019-2-13).
- 12) Mori, S.; Maeta, H.; Yamakata, Y.; Sasada, T. "Flow graph corpus from recipe texts". LREC. 2014, p.2370-2377.
- 13) Naim, I.; Song, Y. C.; Liu, Q.; Kautz, H.; Luo, J.; Gildea, D. "Unsupervised alignment of natural language instructions with video segments". AAAI. 2014, p.164-174.
- 14) Nishimura, T.; Hashimoto, A.; Mori, S. "Procedural text generation from a photo sequence". INLG. 2019, p.409-414.
- 15) Nishimura, T.; Hashimoto, A.; Ushiku, Y.; Kameko, H.; Yamakata, Y.; Mori, S. Structure-aware procedural text generation from an image sequence. IEEE Access. 2020, Vol.9, p.2125-2141.
- 16) Nishimura, T.; Tomori, S.; Hashimoto, H.; Hashimoto, A.; Yamakata, Y.; Harashima, J.; Ushiku, Y.; Mori, S. "Visual grounding annotation of recipe flow graph". Proceedings of the 12th Language Resources and Evaluation Conference. European Language

- Resources Association, 2020, pp.4275-4284.
- 17) Salvador, A.; Drozdal, M.; Nieto, X. G.; Romero, A. "Inverse cooking: Recipe generation from food images". CVPR. 2019.
  - 18) Salvador, A.; Hynes, N.; Aytar, Y.; Marin, J.; Ofli, F.; Weber, I.; Torralba, A. "Learning cross-modal embeddings for cooking recipes and food images". CVPR. 2017
  - 19) Shi, B.; Ji, L.; Liang, Y.; Duan, N.; Chen, P.; Niu, Z.; Zhou, M. "Dense procedure captioning in narrated instructional videos". ACL. 2019, p.6382-6391.
  - 20) Shi, B.; Ji, L.; Niu, Z.; Duan, N.; Zhou, M.; Chen, X. "Learning semantic concepts and temporal alignment for narrated video procedural captioning". ACMMM. 2020, p.4355-4363.
  - 21) Tai, K. S.; Socher, R.; Christopher D. "Manning. Improved semantic representations from treestructured long short-term memory networks". Proc. ACL-IJCNLP. 2015, p.1556-1566.
  - 22) Wang, H.; Lin, G.; Hoi, S. C.; Miao, C. "Structure-aware generation network for recipe generation from images". ECCV. Springer, 2020, p.359-374.
  - 23) Yamakata, Y.; Mori, S.; Carroll, J. "English recipe flow graph corpus". Proceedings of the 12th Language Resources and Evaluation Conference. European Language Resources Association, 2020, p.5187-5194.
  - 24) Zhu, B.; Ngo, C.-W.; Chen, J.; Hao, Y. "R2GAN: Cross-modal recipe retrieval with generative adversarial network". CVPR. 2019.
  - 25) 原島純, 橋本敦史. キッチン・インフォマティクス—料理を支える自然言語処理と画像処理—. オーム社, March 2021.
  - 26) 道満恵介, 高橋友和, 井手一郎, 村瀬洋ほか. マルチメディア料理レシピ作成のための料理レシピテキストと料理番組映像との対応付け. 電子情報通信学会論文誌 A. 2011, Vol.94, No.7, p.540-543.

- \*1 漢字の意味を全く知らない人が中国語の辞書が山のように積まれた部屋の中にいるとする(字が同じかどうかはなんとか判定できる)。部屋には手紙をやり取りするための穴があり、そこには中国語で書かれた質問が書かれている。部屋の中の人言葉の意味がわからないながらも辞書を参照して質問に答えることができるかも知れない。部屋の中の人質問に答えられるとして、これは果たして質問の意味を理解したといえるだろうか?という問い。
- \*2 この分野はまだ邦訳が定まっていないものが多いため、本稿では25)での訳語を用いる際には初出時に\*を付与する。

## 執筆者紹介



**橋本 敦史** HASHIMOTO Atsushi  
 オムロン サイニクエックス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門: 画像処理、機械学習、HCI  
 所属学会: IEEE、電子情報通信学会、情報処理学会  
 博士(情報学)

本文に掲載の商品の名称は、各社が商標としている場合があります。

# 雑踏の中で動き回る自律走行ロボット ～社会ジレンマを解決する強化学習の活用

西村 真衣

私たちの日常にロボットが進出するにつれ、人が介在する環境下におけるロボットナビゲーション技術はますます重要なものとなってきている。我々人間は混雑した駅構内やショッピングモールなど様々な場所、状況において社会規範を守りながら目的地に到達するために高度な意思決定を行っているが、同等の能力をロボットで実現することは容易ではない。人が介在する環境でのロボットナビゲーションでは安全性と効率性2つの側面で問題を捉える必要があり、この両者はトレードオフの関係にあるためである。本稿では、混雑環境下におけるナビゲーションの主なアプローチを俯瞰すると共に、我々が公共空間で暗黙のうちに意識している社会規範のモデルとして社会ジレンマに着想を得たナビゲーション手法について解説する。

## Towards Interactive Crowd-aware Robot Navigation Inspired by Social Dilemmas

NISHIMURA Mai

The interactive robot navigation system in crowded environment has been receiving increasing attention for diversified applications, such as automated delivery service, security service at public spaces, and guidance at the airport. We as human beings can easily navigate in congested environment following the social norms, however, empowering such sophisticated skills for mobile robots is still a challenging problem. In crowded spaces, robots are required to consider not only efficiency of the planned path but also the safety to avoid potential collisions to nearby pedestrians. That is, interactive crowd-aware navigation involves the problem of safety-efficiency trade-offs. In this document, we briefly review the recent approaches for robot navigation in crowded environments and introduce a novel approach to balance the trade-offs inspired by sequential social dilemmas.

### 1. まえがき

物流、病院、ショッピングモール、そして駅構内での案内に至るまで、我々の日常さまざまな場面で移動ロボットが導入され始めている。それに伴い、ヒューマンセントリックな環境下におけるロボットナビゲーション技術はますます重要なものとなっているといえよう。

ロボットナビゲーションでは環境マップ中において障害物を回避しながら目的地へ至る経路の計画を行うが、混雑環境下における経路計画は未だ困難な課題である。第一に人が介在する環境では環境マップが刻一刻と変化するため、取るべき経路を事前に計画しておくというアプローチは取れず、環境の変化に合わせて計画を逐次更新していく必要がある。更に、公共空間、特に人が介在する環境下において、社会規範を意識した振る舞いを（我々人間が自然

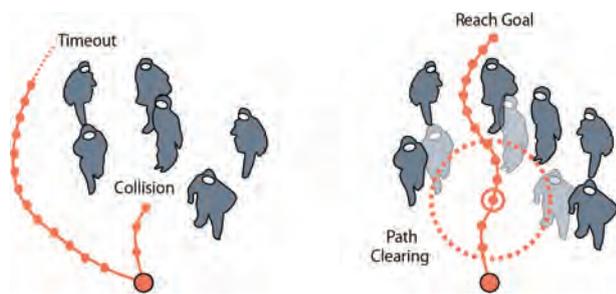
とそうしているように）ロボットに行わせることは極めて難しい。より具体的には我々は雑踏の中を歩く際、周辺歩行者との衝突回避だけでなく目的地へ効率的にたどり着けるよう、つまり安全性と効率性のトレードオフを意識しながら非常に高度な意思決定を適応的に行なっているが、これと同等の機能をロボットで実現する必要がある。そこで本稿では、「人が介在する環境下における適応的なナビゲーションをどう実現するか？」及び「社会規範をロボットにどう認識させるべきか」という2つの技術課題について解説する。

動的な環境下におけるナビゲーションにおいて、従来では予測と計画を別々に取り扱うアプローチが取られてきた<sup>1,2)</sup>。まず環境マップの変化を予測し、その予測結果をもとに安全なナビゲーションを行うというものである。一方、これらの2段階に分かれたナビゲーションに対し、適応的に環境の予測と取り得る行動の決定を強化学習によつ

Contact : NISHIMURA Mai mai.nishimura@sinicx.com

て同時に行う手法が研究されている<sup>4-6)</sup>。しかしながら、いずれのアプローチも環境の複雑さが増加するほど、つまり多数の動的オブジェクトで囲まれた混雑環境下では Freezing Robot Problem<sup>7,8)</sup> と呼ばれる問題に陥ることが知られている。Freezing Robot Problem とは、周辺環境がある混雑度に到達すると、ロボットは安全な経路を導き出すことができず停止 (Collision) してしまうか、不必要な迂回路を取ることで目的地への到達が大幅に遅れてしまう状態 (Timeout) を指す (図1左)。周囲に十分な走行可能領域がない限り迂回路を取り続けることはできず、また搭載されているバッテリー容量も有限であるため、結果としてロボットは高い確率で目的地へ到達することができなくなってしまう。我々人間はなぜ極めて混雑した状況でもこの問題に陥らず確実に目的地へ辿り着くことができるのだろうか？そこで、移動ロボットによる経路計画とは異なり、人間は衝突を回避して迂回路を取るほか、環境に介入して自ら経路を作り出すことができるという点に着目した。この環境への介入行動をロボットに模倣させることを考える。

まずロボットが行うことができる環境への介入行動を以下のように定義する。従来のナビゲーションではロボットが取れる行動として移動及び衝突回避のみが想定されていたが、更に自動車のクラクションのように警告音を鳴らして介入することができるとする。つまり、環境介入によって目的地へ至る経路を自ら作り出す (Path Clearing) 能力を持たせるとする (図1右)。このようにロボットが取得する行動パタンの拡張によって、完全にロボットが停止してしまう状態 (図1左) から脱し、確実に目的地へ到達する確率を高めることができると考えられる。しかし、この介入行動の実装には「この環境への介入どのような判断基準及び頻度で行うべきか」という新たな問題が伴う。



左) 混雑環境において安全な経路を見つけられず、立ち止まってしまうか必要以上の迂回路を取ってしまう。

右) 介入行動によって、確実に目的地へ到達することができる。

図1 Freezing Robot Problem

ロボットが環境に介入する機構については、古くから提示、誘導光を用いる方法<sup>9)</sup>、警告音を用いる方法<sup>10)</sup> が提案されてきた。しかしながら、環境介入は歩行者が接近した際の局所的な判断によってのみ行われ、介入による環境へ

の長期的な影響はこれまで議論されてこなかった。十分な走行可能領域が確保できないような混雑環境下においては、環境介入を頻繁に行うことが安全かつ効率的な経路計画に繋がるわけではない。何故ならば、環境に頻繁に介入すると自然な人の流れを阻害してしまい、結果として目的地への到達が遅れてしまうからである。また逆に、衝突を避け迂回路をとりすぎると目的地への到達が遅れてしまう。つまり、能動的に介入行動を駆使しながら目的地へ効率的に辿り着こうとすること、消極的に周辺の歩行者にぶつからないよう安全性に配慮した経路を選択すること、この効率性と安全性はトレードオフの関係にある。このトレードオフをどのように調節して環境への介入タイミングを制御するかが介入行動を伴うナビゲーションの主要な技術課題となる。

本稿では、この衝突回避と環境への介入頻度の制御を社会ジレンマの問題として捉えることを考える。社会ジレンマとは、個人と全体が資源を共有している状況において、個人にとって合理的な選択が必ずしも社会全体の利益と一致しない葛藤 (ジレンマ) がある状態を指す。つまり、個人にとって最適な行動が必ずしも社会全体にとって最適な行動とならないということである。経済学の分野で有名な共有地の悲劇<sup>11)</sup> を例に挙げる。まず複数の酪農家が共有地である牧草地に牛を放牧するとする。個々の酪農家としては放牧するほどに牛を肥やすことができ短期的には恩恵を受けられるが、全ての酪農家が同様に牛を放牧すると共有資源である牧草の枯渇が起り、長期的には全体の利益を損ねることになる。このような社会ジレンマは牧草地における過放牧のほか、環境汚染や乱獲、無線通信のリソース制御など我々の身近な社会の中で当て嵌まる例が非常に多くある。

ロボットナビゲーションの例を前述の社会ジレンマの枠組みで捉えてみると、公共空間における限られた走行可能領域を複数の周辺歩行者と移動ロボットが共有し、毎時取り得る経路を分け合っている状況であると考えられる。ここでの個人の利益とは目的地に可能な限り早く到達する状態に対応し、個人の利益を追求し頻繁に環境への介入を行うことにより、自然な人の流れを阻害し自身も含めた全体の目的地への到達を阻害する状態が全体の利益を損ねることに対応するであろう。この着想を元に、本稿では我々の提案する環境介入を含む適応的なロボットナビゲーションを、社会ジレンマを伴う意思決定問題としてモデル化し、深層強化学習によって最適な行動方策を獲得する手法 L2B (Learning to Balance) について解説する。

## 2. 関連研究

### 2.1 混雑環境下におけるロボットナビゲーション

混雑環境下におけるナビゲーションでは、まず群衆の自

然な衝突回避行動を模倣するアプローチが考えられてきた。一般に広く用いられるモデルとして RVO (reciprocal velocity obstacles) 及びその拡張手法である ORCA (optimal reciprocal collision avoidance)<sup>1)</sup> があり、これらは Velocity Obstacles として定義される衝突範囲を利用した衝突回避手法の安定な拡張手法として広く利用されている。また実際の人流の振る舞いと類似する動きを学習する模倣学習を用いる方法もある<sup>16)</sup>。更に、近年深層強化学習によって衝突回避と経路計画を同時に行う方法が提案され、混雑環境下におけるロボットナビゲーションにおいて優れた成果を上げてきた<sup>4-6)</sup>。

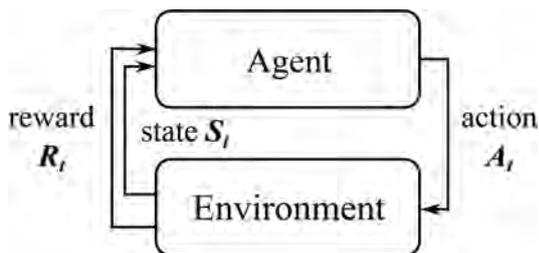


図2 強化学習におけるマルコフ決定過程

ここではまず、ロボットナビゲーションを強化学習の問題として解く基本的なアプローチについて解説する。混雑環境下における経路計画は各時刻で障害物を避けながら目的地へ到達するためのロボットの進行方向及び速度を決定する、連続的な意思決定問題と捉えることができる。この連続的な意思決定はマルコフ決定過程 (Markov Decision Process, MDP) によってモデル化することができる。その概念図を図2に示す。MDPでは環境 (Environment) から状態 (State) が行動主体 (Agent) から観測可能であると、エージェントが方策  $\pi(s_t, a_t)$  に基づいて最適な行動 (Action) を選択するとことにより環境が次のステップへ遷移し、エージェントは報酬関数に基づいて報酬  $r_t = R(s_t, a_t)$  を得る。

混雑環境下のナビゲーションにおいて状態  $s_t$  はロボット及び周辺歩行者の状態によって構成されるが、移動ロボット自体の状態は完全に観測可能である一方で、目的地情報を含む周辺歩行者の状態については、外部から観測可能な速度情報等の情報のみ観測可能 (部分観測可能) であるとする。つまり、部分観測マルコフ決定過程 (Partially Observable Markov Decision Process, POMDP) としてロボットナビゲーションの問題を定式化することになる。強化学習ではこのような枠組みの中で観測と行動を繰り返しながら累積報酬を最大化するような最適方策  $\pi^*$  を学習で獲得することを目的とする。

深層強化学習を複数ロボットの衝突回避制御に先駆的に取り入れた CADRL (Collision Avoidance with Deep Reinforcement Learning)<sup>4)</sup> では、2エージェント間の衝突に関す

るペナルティを報酬関数に組み込み、獲得した方策を各エージェントで独立に実行させることで、中央集権的な制御を必要としない衝突回避手法を実現した。CADRLはロボット同士の衝突回避において優れた性能を示したが、歩行者による混雑環境下においては歩行者同士、及び歩行者-ロボット間の複雑なインタラクションを十分に考慮できていないために積極的に危険な経路を取ってしまうケースが散見された。これに対し、後続研究である SARL<sup>12)</sup> では、強化学習における価値関数を表現するネットワークに周辺歩行者のインタラクションをエンコードするモジュールを組み込むことで、より安全な経路を確実に選択することができるナビゲーション手法を提案した。

しかしながら、これらのアプローチは周辺環境が5名以下といった比較的混雑していない環境を対象としているほか、周辺歩行者に対する衝突回避のみに重点がおかれている。環境がある混雑度に達するとプランナは全てのパスが走行不能と判断してしまい、ロボットは完全に停止してしまうか衝突を回避するため不必要な迂回路をとってしまう。深層強化学習を始めとしたこれらの手法はいずれも、Freezing Robot Problem を完全には解決できない。我々のアプローチは、衝突回避だけでなく能動的に群衆に働きかけ経路を作り出すことができる方策を学習によって獲得することにより、ロボットが停止せず確実に目的地へ辿り着くようにするものである。

## 2.2 社会規範のモデル化

周囲の歩行者と適切にインタラクションを取るには社会規範を意識する必要がある、ロボットナビゲーションの文脈においても社会規範の学習が盛んに研究されてきた。群衆の自然な流れから社会規範に従うルールを学習し、その枠組みをロボットナビゲーションに応用する試みもある<sup>14,15)</sup>。また別のアプローチでは歩行者の社会規範を表現する方法としてゲーム理論に影響を受けたモデルが提案されている<sup>3)</sup> が、中でも興味深いのはマルチエージェントシミュレーションの分野で提案された Sequential Social Dilemmas (SSD)<sup>13)</sup> である。

まずゲーム理論について概説するが、単目的最適化や多目的最適化は全体合理性を追求する目的関数が設定されている一方で、「ゲーム」は個々の行動主体の合理性、つまり個々のエージェントが他者とは独立に自己の目的関数を最大化することを前提とした最適化問題である。ゲーム的状况とは、環境に複数の意思決定主体または行動主体が存在し、それぞれが個別の目的関数の最適化を目指して相互に依存している状況を意味する。ゲーム的状况の中で行動主体がどのような意思決定をするのかを数理モデルによって表現し、分析手法を理論化したものがゲーム理論である。

社会におけるさまざまなゲーム的状况の中で、特に我々

の身近に存在する問題としては共有地の悲劇や囚人のジレンマに代表される社会的ジレンマがある。社会的ジレンマとは、個人と全体が資源を共有している状況において、個人にとって合理的な選択が必ずしも社会全体の利益と一致しない葛藤（ジレンマ）がある状態を指す。表 1 に社会的ジレンマがある状況において、エージェント A、B がそれぞれ協力（ $\pi_c$ ）・非協力（ $\pi_d$ ）行動を選択した際にエージェント A が受け取る利得構造を示す。

表 1 社会ジレンマにおける協調・非協調方策の報酬構造

	協力（ $\pi_c$ ）	非協力（ $\pi_d$ ）
協力（ $\pi_c$ ）	R 相互協力	S 他方に裏切られる
非協力（ $\pi_d$ ）	T 自身が裏切る	P 両者が裏切る

この利得行列（Pay-off Matrix）を持つマトリックスゲームが以下の条件を満たすケースが社会的ジレンマである。

- R>P: 相互協力が相互敵対より利得が高い
- R>S: 相互協力が一方に裏切られるよりも利得が高い
- 2R>T+S: 相互協力が全体的に最も利得が高い
- T>R: 相手を裏切る方が相互協力よりも利得が高い
- または P>S: 相互に裏切る方が一方的に裏切られるよりも利得が高い（罰則が軽い）

しかしながら、実世界の問題では時間軸が存在し、エージェントの行動空間に対し協調・非協調は二値で常に分離可能ではなく、段階的な量である。また、エージェントは他エージェントの状態を完全に観測することはできず、部分的な観測のみで意思決定を行う必要がある。これらを考慮し、従来の Matrix Game Social Dilemma (MGSD) を時間軸に拡張し、POMDP として扱えるようにした枠組みが Sequential Social Dilemmas (SSD) である。

時間軸での状態遷移を伴うマルコフゲームは、状態空間  $S$ 、行動空間  $A_1, A_2$ 、状態遷移関数  $T$ 、報酬関数  $R \in S \times A_1 \times A_2$  から構成される。各エージェントの行動はそれぞれ独立に決定される一方で、報酬は状態とエージェントの行動でなく、自己と他者両方の行動に依存する。SSD はこのマルコフゲームにおける報酬構造に対し、上述の社会的ジレンマの制約を課したものである。本稿では、混雑環境下におけるロボットナビゲーションを社会ジレンマを伴うマルコフゲームである、SSD に類似した状況であるということに着想を得、強化学習によって最適な方策を獲得する方法について解説する。

### 3. 強化学習を用いたナビゲーション

介入行動を伴うロボットナビゲーションを POMDP として定式化を行う。ここで、ロボットと群衆（周辺歩行者の

グループ）をそれぞれロボットエージェントと群衆エージェント 2つのタイプのエージェントとして取り扱い、独立した異なる行動方策を持っているものとする。特に、今回はロボットと個々の周辺歩行者ではなく、ロボットと群衆全体との間での意思決定に伴う社会的ジレンマを扱うため、群衆エージェントは複数の歩行者を 1つのエージェントとして簡易的に取り扱い、ロボットエージェントから未知の方策に従っているものとする。

各タイムステップ  $t$  において、群衆エージェントの状態  $s$  は部分的に観測可能であり、目的地の情報は外部からは観測不可能であるとする。群衆エージェントの状態が  $s = \{s_o, s_h\}$  のように観測可能 (Observable) な状態  $s_o$  と観測不能 (hidden) な状態  $s_h$  から構成されているとすると、ロボットエージェントから観測できる状態は自身の状態と群衆エージェントの部分観測を含め  $s^{jn} = \{s_t, \tilde{s}_t^o\}$  となる。部分観測可能な群衆エージェントの状態は以下で表現される。

$$s_t = [d_t^{(g)}, v_t, v_{pref}, r^{(c)}, r^{(b)}] \\ \tilde{s}_t^o = [\tilde{p}_t, \tilde{v}_t, \tilde{d}_t, \tilde{r}] \quad (1)$$

ここで、 $d^{(g)}$  は目的地までの距離、 $v_t, v_{pref}$  はそれぞれ現在速度及び理想とする速度、 $r^{(c)}, r^{(b)}$  は衝突を考慮する半径及び介入行動によって周辺歩行者が影響を受ける距離半径である。強化学習では、累積報酬の期待値を最大化するような最適な方策  $\pi^*: s_t^{jn} \mapsto a_t$  を学習によって獲得することである。

$$\pi^*(s_t^{jn}) \\ = \arg \max_{a_t} R(s_t^{jn}, a_t) \\ + \gamma^{\Delta t \cdot v_{pref}} \int T(s_t^{jn}, s_{t+\Delta t}^{jn} | a_t) V^*(s_{t+\Delta t}^{jn}) ds_{t+\Delta t}^{jn} \quad (2)$$

また、累積報酬の期待値をエンコードする最適な価値観数  $V^*$  は以下の式で表される。

$$V^*(s_t^{jn}) = \mathbb{E} \left[ \sum_{t'=t}^T \gamma^{t'-v_{pref}} R(s_{t'}^{jn}, \pi^*(s_{t'}^{jn})) \right], \quad (3)$$

ここで、 $\gamma \in [0,1)$  は割引率であり、エージェントの現在速度によって割引率が増減しないようタイムステップ及び理想とされる速度  $v_{pref}$  によって正規化されている。高次元の状態空間を扱うため、価値関数を SARL<sup>12)</sup> に準拠したニューラルネットワークで近似することとする。また、学習するのはロボットエージェントの方策  $\pi$  のみであり、群衆エージェントの方策  $\pi$  としてはルールベースの衝突回避手法である ORCA<sup>1)</sup> を用いる。

### 4. 社会的ジレンマを考慮した報酬設計

本研究では群衆への介入行動を伴うロボットナビゲーションを社会ジレンマを伴う意思決定問題として捉える。Sequential Social Dilemmas (SSD) では、個々のエージェン

トの行動によって毎時状況が遷移するマルコフゲームにおいて、個人と全体が資源を共有している状況において、個人にとって合理的な選択が必ずしも社会全体の利益と一致しない葛藤（ジレンマ）がある状態を指す。ロボットナビゲーションの文脈における個人の利益とは目的地に可能な限り早く到達することであるとすると、個人の利益を追求し頻繁に環境への介入を行うことは自然な人の流れを阻害し自身も含めた全体の目的地への到達を阻害するため、自身を含む全体の長期的な利益を損ねることに対応する。 $\pi_c, \pi_d$ をそれぞれ協調的、非協調的方策であるとすると、SSDの枠組みにおいて、R、P、S、Tそれぞれの利得を強化学習の報酬として設計することを考える。ここで、前章で解説した通り R は相互協力、S は自分のみが協力、T は相手のみが協力、P は双方とも協調しなかった場合の利得である。

相互協力による報酬  $R = V(\pi_c, \pi_c)$  をロボットと群衆がそれぞれ衝突回避し道を譲り合う状態、 $S = V(\pi_c, \pi_d)$  を群衆の経路を優先してロボットが迂回路を取るケースであるとする。 $T = V(\pi_d, \pi_c)$  は介入動作や自己経路優先により群衆の通行をロボットが阻害するケースであり、 $P = V(\pi_d, \pi_d)$  はロボット、群衆が両者とも自己経路を優先し、渋滞或いは衝突してしまうケースと考える。また、各エピソードは目的地に到達、周辺歩行者に衝突、所定の所要時間を超過する何れかの条件を満たすとその結果に紐づく報酬を得て終了するものとする。

以上を踏まえて社会ジレンマを考慮した報酬関数の設計を行う。環境から得られる報酬  $R_e$  及び群衆エージェントの行動と自己の行動の組み合わせによって得られる報酬を  $R_s$  とすると、報酬関数  $R(s_t, a_t)$  を以下のように表すことができる。

$$R(s_t^m, a_t) = R_e(s_t^m, a_t) + R_s(s_t^m, a_t), \quad (4)$$

$R_e$  は一般的な衝突回避に関する報酬である。ただし、ロボットエージェントの現在位置  $p_t$  が目的地  $p^{(g)}$  に到達すると得られる成功報酬は、所定時間に対する所要時間の割合とその係数  $\alpha$  に応じて減衰するものとする。また、周辺歩行者との最短距離  $d_t$  が衝突判定距離  $d_{min}$  を下回ると衝突とみなし、それ以外の状態については報酬が発生しないものとする。エージェントが目的地に到達または衝突するとエピソードは終了する。

$$R_e(s_t^m, a_t) = \begin{cases} 1.0 - \alpha \frac{t}{t_{lim}} & \text{if } p_t = p^{(g)} \\ -0.25 & \text{elseif } d_t < d_{min} \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

一方、周辺エージェントとのインタラクションにより発生する社会的な報酬  $R_s$  は介入行動及び接近行動によって生じるペナルティによって形成される。2 値符号である  $b_t$  は介入行動の有無（有=1、無=0）を、 $r(b)$  を警告音など

の介入行動の有効半径とし、 $d_{disc}$  を周辺歩行者が不快と感じる近接距離とすると、報酬関数  $R_s$  は以下のようにかける。

$$R_s(s_t^m, a_t) = \begin{cases} \beta(d_t - r^{(b)}) & \text{if } d_t < r^{(b)} \wedge b_t = 1 \\ \eta(d_t - d_{disc}) & \text{elseif } d_t < d_{disc} \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

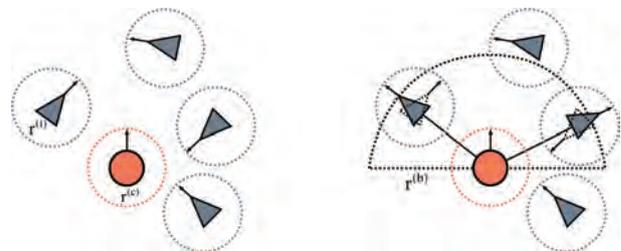
ここで、 $\alpha, \beta, \gamma$  はそれぞれ目的地へ早く到達したいとする積極性の制御項、積極的な介入行動によって生じるペナルティの重み、消極的に衝突を避けようとする傾向の調整項である。提案法では、このような社会ジレンマを考慮した報酬系の上で累積報酬の期待値を最大化することにより、安全性と効率性のトレードオフを調停（L2B: Learning to Balance）することを目指す。

### 5. 実験

社会ジレンマを考慮した提案フレームワーク L2B の効果を検証するため、群衆環境を模したシミュレーションにおいて提案手法の有効性の評価を行なった。

#### 5.1 シミュレーション

前章までで述べたように混雑環境を対象とした既存研究は衝突回避のみを扱っているため、介入行動を伴うナビゲーションの評価環境は存在しない。そこで、脅威に対して人の回避行動が距離に応じてどう変化するかを RVO の拡張として取り入れた ERVO（Emotional Reciprocal Velocity Obstacles）を模したシミュレーションシステムを実装し、介入行動に対する人の反応を ERVO により生成することとした。図3に示すように、ロボットエージェントが半径  $r^{(b)}$  を影響半径とする警告音で介入行動を選択した際、その距離  $r^{(b)}$  に応じて影響半径から退避する回避行動を取るものとする。また、群衆が目的地へ向かう際の動きは同様に RVO の拡張である ORCA に従うものとする。4.0 [m<sup>2</sup>] の空間に歩行者人数  $N = \{5, 10, 15, 20\}$  複数の混雑度のシナリオを作成し、各混雑度における提案法の効果を確認した。



（橙●：ロボットエージェント 灰▲：群衆エージェント）

図3 ERVO を元にした介入行動に対するシミュレーション

## 5.2 評価方法

混雑度に応じて円周上にランダムにエージェントを配置し、対岸の目的値にお互い交差しながら進行するシナリオ (Circle Crossing Scenario) をベンチマークとして用いる。エージェントの配置には正規分布に従うランダムノイズを加えて500回のテストケースを作成し、パフォーマンスを目的地までの到達率 (Success)、衝突率 (Collision)、目標時間内に目的地に到達せずタイムアウトした割合 (Timeout)、目的値までの平均所要時間 (Time) 4つのメトリクスで評価を行なった。比較手法として、介入方策を学習によって獲得する手法はこれまで存在していないため、SARLのモデルをベースに介入行動を導入した提案法を今回L2B-SARLと呼称するものとし、介入方策を用いないオリジナルのSARLを各混雑度のシナリオに適用した結果と比較を行った。

表2に各メトリクスでの評価結果を示す。SARLはより混雑した環境下では到達率が下がってしまう結果となった。これは介入方策を持たないオリジナルのSARLでは消極的に衝突回避を行うことしかできないので、迂回路を取りすぎてタイムアウトになるケースが頻繁に起こるためである。一方、L2B-SARLは $N > 15$ の混雑した環境においてもタイムアウト率が非常に低い結果となっており、目的地へ確実に到達できていることがわかる。また、環境への介入は常により効率的に目的地へ到達する結果には至らず、状況によっては介入を用いない方策よりも目的地への到着が遅れるケースが散見された。これは前章で述べたように環境介入は群衆の動きを乱すことにより、多用するこ

とによって自己及び全体に不利益をもたらす効果があるという仮説と一致している。

表2 定量的な評価結果

Methods	$N$	Success	Collision	Timeout	Time
L2B-SARL	20	0.906	0.094	0.000	13.85
SARL [12]		0.700	0.184	0.116	13.50
L2B-SARL	15	0.880	0.118	0.002	11.60
SARL [12]		0.778	0.064	0.158	12.43
L2B-SARL	10	0.904	0.086	0.004	11.31
SARL [12]		0.922	0.046	0.032	11.91
L2B-SARL	5	0.978	0.020	0.002	10.14
SARL [12]		0.966	0.032	0.002	10.09
L2B-SARL	average	0.917	0.079	0.002	11.72
SARL [12]		0.841	0.081	0.077	11.98

更に、図4に定性的な評価結果を示す。橙マークがロボットエージェントが取った軌跡及び介入行動を選択した地点、灰マークが群衆エージェント介入方策を用いた上段 (提案法) では、衝突回避方策の下段 (SARL) よりも迂回路に偏らず効率的な経路を選択でき、かつ適切な頻度で介入行動を行えていることが確認できた。

## 6. むすび

本稿では、適応的なロボットナビゲーションの枠組み及び混雑環境下における課題を紹介し、介入行動を伴う新た

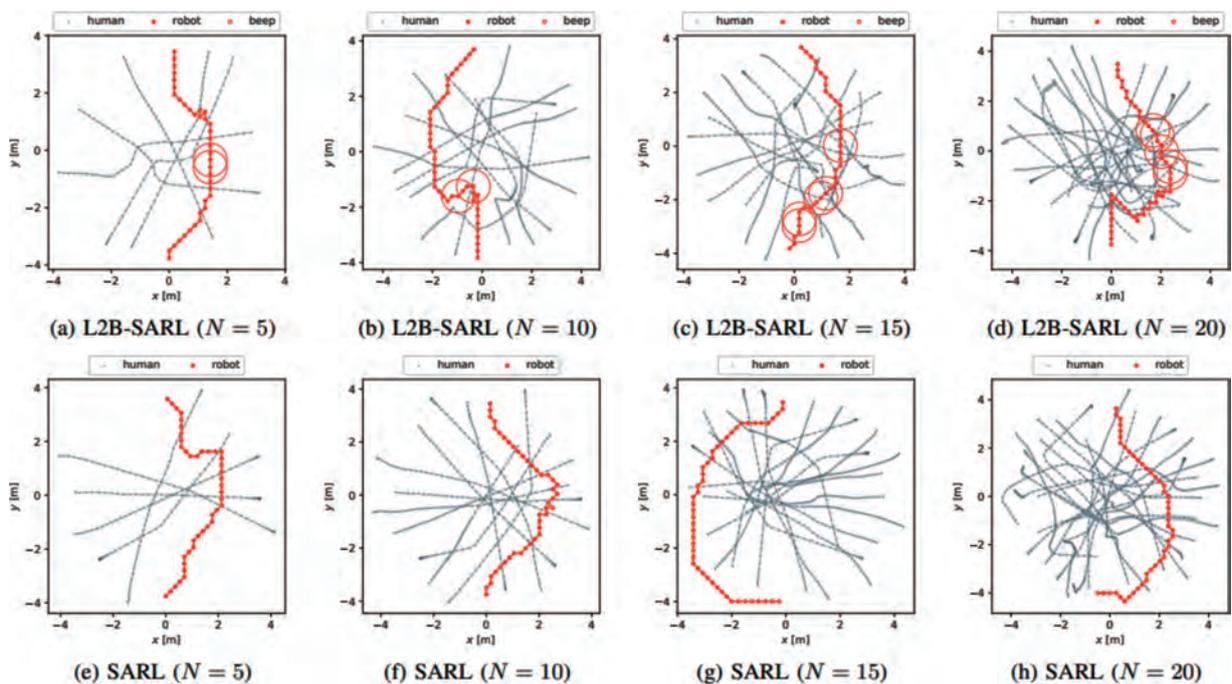


図4 定性的な評価結果

なナビゲーション技術の枠組みについて解説した。現状はシミュレーション環境のみでの評価のみであるが、提案法は Non-Holonomic なロボットのキネマティクスにも対応することができるため、移動ロボット実機で訓練した方策を利用することも可能である。今回は単一エージェントのみを学習対象としたが、社会ジレンマの介在する環境下において、マルチエージェント強化学習の枠組みにより複数ロボットの協調行動を引き出すような発展も考えられる。

## 参考文献

- 1) Van Den Berg, J.; Guy, S. J.; Lin, M.; Manocha, D. "Reciprocal n-body collision avoidance". *Robotics Research*. Springer, 2011, p.3-19.
- 2) Fox, D.; Burgard, W.; Thrun, S. The dynamic window approach to collision avoidance. *IEEE Robotics & Automation Magazine*. 1997, Vol.4, No.1, p.23-33.
- 3) Fisac, J. F.; Bronstein, E.; Stefansson, E.; Sadigh, D.; Sastry, S. S.; Dragan, A. D. "Hierarchical game-theoretic planning for autonomous vehicles". *IEEE International Conference on Robotics and Automation*. 2019, p.9590-9596.
- 4) Chen, Y. F.; Liu, M.; Everett, M.; How, J. P. "Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning". *IEEE International Conference on Robotics and Automation*. 2017, p.285-292.
- 5) Chen, Y. F.; Everett, M.; Liu, M.; How, J. P. "Socially aware motion planning with deep reinforcement learning". *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2017, p.1343-1350.
- 6) Everett, M.; Chen, Y. F.; How, J. P. "Motion planning among dynamic, decision-making agents with deep reinforcement learning". *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2018, p.3052-3059.
- 7) Trautman, P.; Krause, A. "Unfreezing the robot: Navigation in dense, interacting crowds". *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2010. p.797-803.
- 8) Trautman, P.; Ma, J.; Murray, R. M.; Krause, A. Robot navigation in dense human crowds: Statistical models and experimental studies of human-robot cooperation. *International Journal of Robotics Research*. 2015, Vol.34, No.3, p.335-356.
- 9) Matsumaru, T.; Kusada, T.; Iwase, K. "Mobile robot with preliminary-announcement function of forthcoming motion using light-ray". *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, p.1516-1523.
- 10) Kayukawa, S.; Higuchi, K.; Guerreiro, J.; Morishima, S.; Sato, Y.; Kitani, K.; Asakawa, C. "Bbep: A sonic collision avoidance system for blind travellers and nearby pedestrians". *CHI Conference on Human Factors in Computing Systems*, 2019, p.1-12.
- 11) Hardin, G. The tragedy of the commons. *Science*. 1968, Vol.162, No.3859, p.1243-1248.
- 12) Chen, C.; Liu, Y.; Kreiss, S.; Alahi, A. "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning". *IEEE International Conference on Robotics and Automation*. 2019, p.6015-6022.
- 13) Leibo, J. Z.; Zambaldi, V.; Lanctot, M.; Marecki, J.; Graepel, T. "Multi-agent reinforcement learning in sequential social dilemmas". *International Conference on Autonomous Agents and Multi-Agent Systems*. 2017, p.464-473.
- 14) Robicquet, A.; Sadeghian, A.; Alahi, A.; Savarese, S. "Learning social etiquette: Human trajectory understanding in crowded scenes". *European Conference on Computer Vision*. 2016, p.549-565
- 15) Chen, Y. F.; Everett, M.; Liu, M.; How, J. P. "Socially aware motion planning with deep reinforcement learning". *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2017, p.1343-1350.
- 16) Tai, L.; Zhang, J.; Liu, M.; Burgard, W. "Socially compliant navigation through raw depth inputs with generative adversarial imitation learning". *IEEE International Conference on Robotics and Automation*. 2018, p.1111-1117.

## 執筆者紹介



西村 真衣 NISHIMURA Mai

オムロン サイニックエックス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門：コンピュータビジョン, GPGPU  
 所属学会：IEEE

本文に掲載の商品の名称は、各社が商標としている場合があります。

# 動的に変化する環境の中で自己位置を推定する 自律走行ロボット

西村 真衣

移動ロボットの自律移動システムは搭載センサによる環境認識と地図構築、その地図上での自己位置認識に基づく経路計画によって構成される。静的な環境においては事前に地図構築を行い、環境センサ情報を登録地図と照合することにより比較的安定して移動システムを動作させることができるが、人が介在するなど動的な環境においては地図構築の観点からはノイズとなる移動体を如何に処理するかが課題となる。本稿では自律移動システムのうち特に動的環境を対象とした自己位置推定に焦点をあて、移動体が存在する環境下における主なアプローチを俯瞰すると共に、我々が提案する移動体のみを利用する新たな移動軌跡復元手法の枠組みについて述べる。

## Mobile Robot Navigation in Densely Crowded Environment

NISHIMURA Mai

The autonomous mobile robot (AMR) system consists of environment mapping using equipped sensors, self-localization and path planning on the top of the map. In static environment, where the layout of objects is fixed and no dynamic objects appear in the map, the system can build the whole environment map in advance. However, when it comes to dynamic environments, the AMR system is required to build the map sequentially by handling dynamic obstacles in real-time. In this document, we especially focus on the self-localization system in the dynamic environment. We first briefly review fundamentals of localization system based on the multi-view geometry and introduce its extensions to incorporate dynamic points and obstacles. Moreover, towards navigating in highly congested scenarios, we propose a novel self-localization framework that depend only on dynamic objects in the observed images.

### 1. まえがき

移動ロボット（モバイルロボット）、またその土台となる自律移動システムは工場内や公共施設における運搬、警備、物流、清掃など多様なシーン、用途で活用されている。一方で工場内のように経路、障害物を固定できる環境と比較して、人が介在するなど特に動的に大きく変化する環境へのモバイルロボットの導入は未だ課題が多い。自律移動システムは搭載センサによる周辺環境の認識と地図構築、構築した地図上での自己位置認識と経路計画によって構成されるが、その全てのプロセスにおいて移動体の存在を考慮する必要がある。第一に地図構築において、静的な環境では事前に地図構築を行うことが可能であるが、移動体を含む環境地図は刻一刻と変化するため、固定の静的な背景と移動体による動的な前景を分離可能にした環境地図

を逐次更新しながら構築し、走行可能領域を算出する必要がある。更に、その構築した地図上で自律走行を行うには、変化する環境地図上での自己位置及び障害物となる移動体の認識が不可欠となる。本稿では、このような動的環境下を対象とした自律移動システムの中で特に自己位置推定に着目し、既存のアプローチを俯瞰すると共に、移動体のみを利用した新たな自己軌跡復元の枠組みについて述べる。

移動ロボットでは、IMU（Internal Measurement Unit）、LiDAR等多様なセンサが利用されているが、本稿では特に単眼RGB画像センサで構成される自律移動システムについて解説する。RGBカメラを用いた地図構築及び自己位置推定は静的なランドマークを複数の視点から観測し、多視点幾何による幾何学的な拘束式を解くことによりカメラの位置パラメータを逐次算出することが基本となる。そのため、環境地図及び自己位置推定は静的なシーンの観測

Contact : NISHIMURA Mai mai.nishimura@sinicx.com

が前提とされ、歩行者や自転車、自動車など移動体が含まれる環境ではそれらを外れ値として、又は意味レベルで障害物として認識し、除外する方法が取られてきた<sup>1,2)</sup>。更に、移動体が多く存在するシーンでは安定した自己位置推定を実現するため、移動体であるオブジェクトも同時にランドマークとして複合的に最適化対象とするアプローチも提案されている<sup>3,4)</sup>。しかしながら、これらはいずれも静的なランドマークの観測をベースとした従来の手法の拡張として提案されており、静的なランドマークが安定してトラッキングし続けられることが前提となる。

以上で述べた通り、従来の自律移動システムは周辺環境が静的で変化しないことが前提であり、動的な環境の変化に対応するための拡張手法群もシーン全体の中で静的なランドマークが支配的である環境を対象としている。では、雑踏など移動体の存在がシーンの内で支配的であり、遮蔽により静的なランドマークの安定した追跡が不可能な状況下でも尚自律走行ロボットを運用し続けるにはどのような方法が取れるだろうか。本稿の後半では更に、そのような従来のアプローチが機能しない“超”動的な環境における自己位置推定の新しい枠組みについて紹介する。

## 2. 多視点幾何の基礎

### 2.1 多視点撮影による幾何パラメータの推定

移動ロボットでは以下 i)～iv) のプロセスの反復によって移動しながら環境地図と自己位置を逐次更新していく。

- i) 2 視点でのランドマークの検出・マッチング
- ii) ランドマークの 3 次元位置及びカメラ位置姿勢の復元
- iii) 復元したランドマークを環境地図へ登録
- iv) 新たな観測点において再度ランドマークを検出、登録されたランドマークとの対応関係から、カメラ位置姿勢を推定
- v) 新規に観測されたランドマークの 3 次元位置を復元し、環境地図へ登録

このように地図構築 (Mapping) と自己位置推定 (Localization) は相互に依存しているため、自律走行におけるロボットの自己位置推定技術は Simultaneous Localization and Mapping (SLAM) と呼称される。従来の SLAM システムでは、観測する環境が静的で変化しないランドマーク (特徴点) が存在することを前提とし、複数フレームで共通のランドマークが追跡可能である必要がある。

図 1 に多視点幾何の概念図を示す。3 次元空間中のランドマーク P は各視点 (View1, View2) において画像平面上の点 p, p' として投影される。この画像平面上における対

応点を用いて、視点間のカメラ位置姿勢を回転行列 R, 並進移動ベクトル t を算出する。

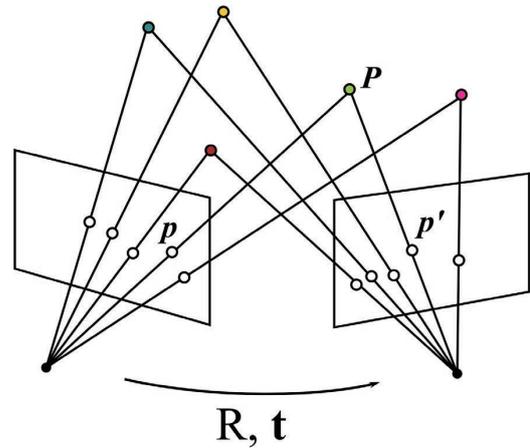


図 1 共通のランドマークの観測によるカメラ幾何パラメータ推定

### 2.2 バンドル調整

バンドル調整<sup>17)</sup>とは、図 1 におけるランドマーク P の 3 次元位置とカメラを結ぶ光線束 (Bundle) を対応する撮影画像上の特徴点 p の投影関係を利用して調整し、その背後にある幾何パラメータを推定・最適化するための方法である。実用上、計算過程において画像平面上における特徴点 p の検出座標には観測誤差が含まれるため、ランドマーク P の 3 次元位置及びカメラ位置姿勢パラメータの推定結果はその誤差の影響を受けたものとなる。バンドル調整ではこのような観測誤差の混入を前提とし、推定された各カメラ姿勢  $R_i, t_i$  について元の観測された点  $p_i$  と 3 次元点  $P_i$  の投影関数  $f(P_i | R_i, t_i)$  による再投影誤差

$$\min_{R_i, t_i} \sum |p_i - f(P_i | R_i, t_i)|_2^2 \quad (1)$$

を最小二乗法によって最小化することで最適なパラメータを求める。なお、最小二乗法は誤差に対し正規分布のノイズを仮定した場合の最尤推定に相当する。多視点幾何を用いたカメラの位置姿勢推定は観測を増やすごとに各処理の誤差が蓄積していくため、過去に推定された位置姿勢パラメータ群を局所的なグループ毎 (local Bundle Adjustment, local BA) またはループが検出されたタイミングで過去に推定された全ての位置姿勢に対して最適化を行う (global Bundle Adjustment, global BA)。

## 3. 動的な環境下における自己位置推定

### 3.1 ロバスト推定手法

前章までで述べたように従来の自己位置推定技術は静的なランドマークを多視点で観測することを基本として構成されている。従って、観測された特徴点の一部に動的な点

群が含まれた場合、それらをまず外れ値として処理するアプローチが取られてきた。その最も代表的な手法がロバスト推定手法である。ロバスト推定とは与えられた観測値、予測値に外れ値が含まれていることを前提とし、その誤差影響を抑えることを目的とした手法群である。代表的な手法として Random Sample Consensus (RANSAC) や Least Median of Squares (LMedS)、M-estimator 等がある。2.2 節で触れた最小二乗法では、目的値  $y$  に対する予測値  $f(x)$  の二乗誤差最小化を行うが、観測に含まれる誤差の分布が正規分布に従うことを仮定していた。

$$LMS = \min_x \sum \epsilon^2, \quad \epsilon = |y - f(x)| \quad (2)$$

これに対し、M-estimator は誤差基準  $\epsilon$  に対して誤差重み関数  $\rho$  を設定し、重み付き誤差を最小化するアプローチである。

$$M = \min_x \rho(\epsilon) \quad (3)$$

重み関数としては外れ値に対してより小さな重みを与えるような偶関数が選ばれることが多く、特徴点ベースで代表的な Visual SLAM 手法である ORB-SLAM<sup>5,6)</sup> では Huber のコスト関数を使用している。また、一般にはデータに対するノイズの分布は未知であることから、学習データに対して最適なこのロバスト推定手法における誤差関数  $\rho$  を学習させる手法<sup>7)</sup> も近年提案されている。このようなロバスト推定に基づくアプローチは主に誤検出や僅かな移動体によって生じる外れ値を対象としており、移動体が多く存在するようなシナリオには適用できない。

### 3.2 Dynamic SLAM

近年の移動体を含む動的なシーンを対象とした SLAM システムの研究群 (Dynamic SLAM) では、大きく二つの方向性において従来の静的シーンを前提とした SLAM システムを拡張している。まず 1 つは、環境マップを離散的な点群として扱うのではなく、Semantics を考慮して Object レベルでの処理を行う Object-aware なシステムを構築している点である。最も単純な実装として、Semantic レベルで検出した移動オブジェクトを個別にトラッキングしながらフィルタリング処理によって環境地図を構成するランドマークから取り除く方法が提案されている<sup>1,2)</sup>。更にもう 1 つは、静的なランドマークと同時に移動体であるオブジェクトとその動きを外れ値として除外せず、最適化対象として含めるという点である。

移動体を同時に最適化対象に含める Dynamic SLAM では、静的なランドマークの検出と同時にオブジェクトレベルでの検出・追跡処理が実行される。つまり、静的シーンを扱う SLAM システムが観測フレーム間でマッチング可能な静的な特徴点のみを対象とするのに対し、Dynamic SLAM では特徴点及びオブジェクトのフレーム間対応付け

を同時に行い、因子グラフとしてバンドル調整のパイプラインに組み込むアプローチをとっている<sup>3,4)</sup>。

## 4. 移動体のみを利用する自己位置推定

既存の静的なランドマーク観測に基づく SLAM の拡張として記述される Dynamic SLAM では、移動体を含むシーン全体の復元を実現している一方で、移動体と同時に常に安定して静的なランドマークを追跡し続ける必要があった。従って歩行者で混雑した路上など移動体が非常に多く含まれ、相互遮蔽により静的なランドマークの安定した追跡が不可能な状況下において自己位置推定を行うことは極めて難しい。では見方を変えて、静的なランドマークの観測に依存せず、移動体のみを用いてカメラの自己運動を含めた周辺環境を復元することはできないだろうか。つまり、図 2 に示すように一人称視点での周辺歩行者の動きの 2D 観測から、俯瞰視点でのカメラ自己運動及び周辺歩行者の移動軌跡を復元する問題を考える。

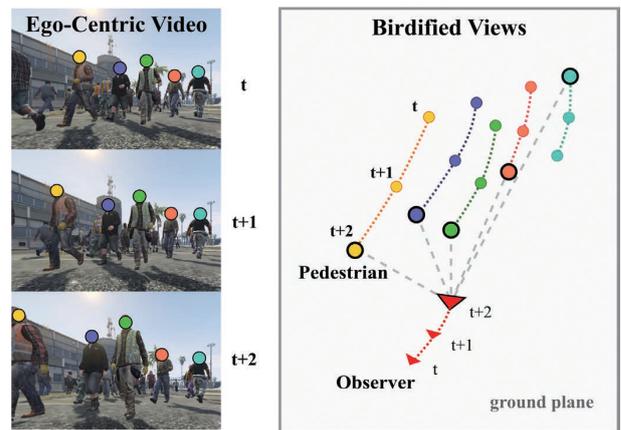


図 2 一人称観測に基づく自己及び周辺歩行者の軌跡復元

各タイムステップ  $t$  において、歩行者  $1, \dots, K$  の一人称映像に基づく相対的な観測  $Z_{t,K}^i = z_1^i, z_2^i, \dots, z_K^i$  が与えられたとき、その俯瞰視点上での位置  $X_{t,K}^i = \{x_1^i, x_2^i, \dots, x_K^i\}$  及び観測カメラの位置  $x_0^i$  を同時復元するとする。移動体が大きさ (身長)

$$\begin{bmatrix} \tilde{x}_k \\ \tilde{y}_k \\ \tilde{z}_k \end{bmatrix} = \frac{fh_k}{l_k} R_x \left( -\frac{\pi}{2} \right) A^{-1} \begin{bmatrix} u_k \\ v_k \\ 1 \end{bmatrix} \quad (4)$$

に限られる人物であると仮定すると、画像平面上における人物検出位置  $[u_k, v_k]^T$  に対し、観測カメラに対する相対的な 3D 位置  $[\tilde{x}_k, \tilde{y}_k, \tilde{z}_k]^T$  は、カメラ内部行列  $A$ , 焦点  $f$  を用いて以下に示す逆投影で表現される。ここで、 $R_x(\theta)$  は  $x$  軸周りで角度  $\theta$  の回転、 $h_k, l_k$  はそれぞれ人物の 3次元空間での身長、画像平面における観測上の身長である。

実環境での人物の身長が平均  $\mu_h$ 、分散  $\sigma_h$  のガウス分布

に従う<sup>9)</sup>とすると、観測モデルは逆投影線上のガウス分布として記述される。

$$z_k^i \sim p(z_k^i | x_k^i; h_k) = N(\mu_h, \sigma_h^2) \quad (5)$$

また、観測される歩行者が群として共通の移動モデルに沿って移動していると仮定すると、カメラ及び歩行者の移動軌跡  $X_{0:k}^i$  は観測モデルである尤度分布  $p(z_k | x_k)$  及び移動モデルについての事前分布  $p(x_k)$  を用い、以下の事後分布を最大化させることによって求められる。

$$p(X_{0:k}^i | Z_{1:k}^i, X_{0:k}^{i-1}) \propto p(X_{0:k}^i | X_{0:k}^{i-1}) p(Z_{1:k}^i | X_{0:k}^i, X_{0:k}^{i-1}) \quad (6)$$

ここで、周辺歩行者の推定位置  $X_{1:k}^i$  を固定すると、観測カメラの移動量  $\Delta x_0^i = [\Delta x_0^i, \Delta y_0^i, \Delta \theta]$  は以下の MAP 推定により求められる。

$$\Delta \hat{x}_0^i = \arg \max_{\Delta x_0^i \in \mathbb{R}^3} p(x_0^i | \chi_0^{i-\tau+1}) \prod_k p(x_k^i | \hat{\chi}_k^{i-\tau+1}, \Delta x_0^i) p(z_k^i | x_k^i, \Delta x_0^i), \quad (7)$$

ここで、 $p(x_0^i | \chi_0^{i-\tau+1})$ 、 $p(x_k^i | \chi_k^{i-\tau+1})$  は  $\tau$  フレーム分の観測におけるそれぞれカメラ、周辺歩行者の移動モデルである。周辺歩行者の移動軌跡は上記で推定されたカメラ位置  $\Delta \hat{x}_0^i$  を固定することにより、同様に以下式により求められる。

$$\hat{\chi}_{1:k}^i = \arg \max_{x_k^i \in \chi_{1:k}^i} \prod_k p(x_k^i | \chi_k^{i-\tau+1}, \Delta \hat{x}_0^i) p(z_k^i | x_k^i, \Delta \hat{x}_0^i) \quad (8)$$

この観測カメラ移動量の推定と歩行者位置推定は相互に依存しているため、交互に推定量を固定しパラメータの逐次更新を繰り返すことにより、カメラ及び歩行者軌跡の復元を行う。

## 5. データセット

動的環境での自己位置推定の評価を行うデータセットは 3次元シーンをキャプチャしたデータに 2次元の動きを含む移動体を撮影画像に 2D 合成したデータセット<sup>8)</sup>等が提案されているが、歩行者により混雑した環境を対象とし、その移動軌跡とカメラ自己運動の復元を目的としたデータセットは存在しない。また、リアルな歩行者の軌跡データは俯瞰視点映像で撮影されたものが殆どであり<sup>10-12)</sup>、一人称視点での群衆の動きを捉えたデータセットは希少である。そこで、今回は人工的に生成又は群衆の俯瞰視点映像から抽出した軌跡データを仮想カメラモデルによって一人称映像に投影し、疑似的に生成した一人称視点の観測と俯瞰視点での軌跡データをペアとしたデータを作成した。

### 5.1 シミュレーションデータセット

仮定する移動モデルが既知のケースにおける提案法の性能を検証するため、Social Force Model<sup>13)</sup>によって人工的に生成した軌跡データを 10 ~ 50 人の群衆内に含まれる人数毎に作成した。仮想カメラモデルの内部パラメー

タは既知とし、歩行者の身長はガウス分布  $h_k = N(\mu_h, \sigma_h^2)$ 、 $\mu_h = 1.70[m]$ 、 $\sigma_h \in [0.00, 0.07][m]$  に従ってサンプリングしたデータを生成した。

### 5.2 リアルな歩行者軌跡を利用したデータセット

Motion Model が未知の軌跡データに対する提案法の有効性を確認するため、公開の群衆データセットである ETH<sup>10)</sup>、UCY<sup>11)</sup> から映像中に含まれる群衆人数に応じて軌跡データを抽出し、歩行者の 1 人に仮想カメラをマウント・その視点における疑似的な投影データを俯瞰視点軌跡とペアとしたデータを作成した。

### 5.3 写実的なゲーム映像を利用したデータセット

5.1、5.2 節で作成したデータはいずれも仮想的なカメラモデルを用いた疑似投影であり、歩行者間の相互遮蔽は考慮されていない。そこで、コンピュータビジョン分野での CG データセットとして広く利用されている Grand Theft Auto V (GTAV)<sup>14)</sup> を用い、人物軌跡データとペアとなる写実的な一人称映像を生成した。GTAV では Script Hook V<sup>15)</sup> というライブラリを用いて GTAV 内のネイティブ関数を呼び出す任意のプラグインを記述することができ、指定のシーンにおいて人物配置や深度画像を逐次取得しながら、各エージェントについてプログラムされた軌跡通りの行動を実行することが可能である。図 3 に GTAV によって作成した軌跡データのサンプルを示す。

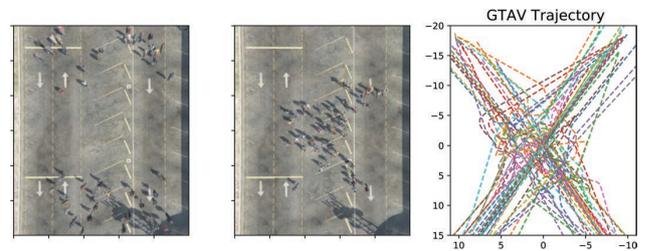


図 3 GTAV で作成した軌跡データサンプル

## 6. 実験

### 6.1 シミュレーションデータでの評価

5.1 節で作成した既知の移動モデルを利用して構築したデータセットにおいて、提案手法の評価を行った。図 4 にシミュレーション結果を示す。 $\Delta r$ 、 $\Delta t$  はそれぞれ自己運動による時刻  $t-1 \sim t$  での回転角及び移動量、 $\Delta \hat{x}$ 、 $\Delta x$  は周辺歩行者のカメラに対する相対位置、観測カメラの運動を含めた絶対位置の推定結果である。歩行者の身長<sup>10)</sup>の分散  $\sigma_h$  が大きいほど推定位置のエラー率は上昇するが、群衆の人数が増加するほどカメラの自己運動の推定精度は高くなり、その観測カメラ位置に基づく周辺歩行者の位置推定精度も高くなる傾向が認められた。

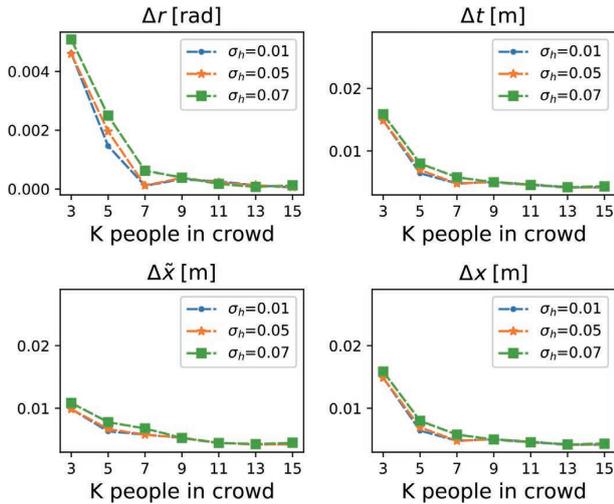


図4 シミュレーションデータにおける提案法の推定結果

### 6.2 リアルな移動軌跡データセットでの評価

5.2節で作成した未知の移動モデルによる群衆の軌跡データセットを利用して構築したデータセットにおいて、提案手法の評価を行った。図5に提案法によって推定されたカメラ及び歩行者位置の確率分布の可視化結果を示す。

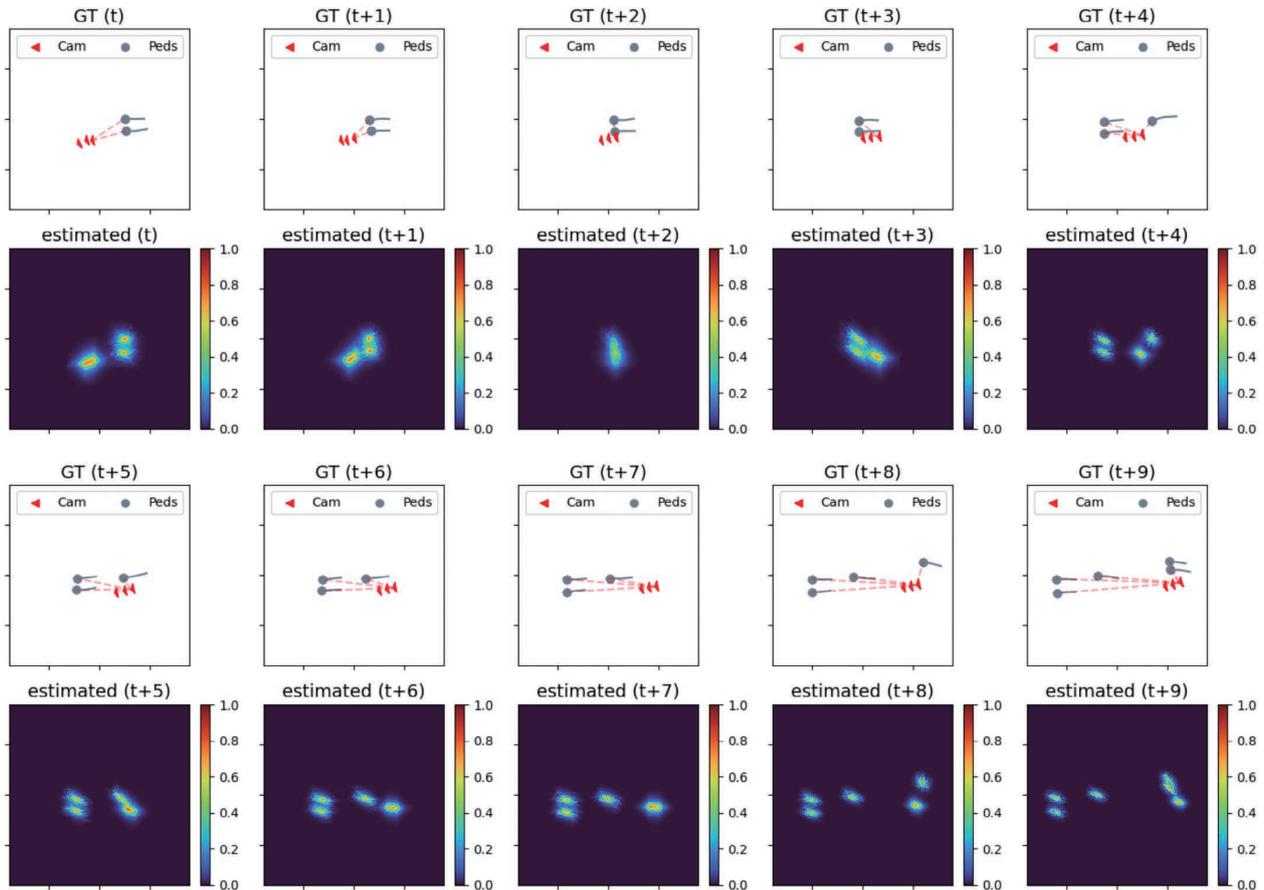


図5 リアルな軌跡データに対するカメラ及び歩行者位置の確率分布推定結果

カメラ軌跡は赤、歩行者軌跡はグレーで描画され、上段 (GT) が正解データ、下段 (estimated) は推定分布に相当する。数値実験結果は割愛するが、 $t \sim t+3$  のように歩行者数が少人数の際は推定される位置分布の裾が広がるため曖昧性が高くなり、 $t+4 \sim t+9$  のように歩行者人数が一定以上のケースでは分布の尖度が向上し、観測する移動体が多くなるに従って位置推定は安定して機能する特性が認められた。

### 6.3 GTAV により生成した一人称映像データによる評価

GTAV により生成した一人称映像に対し、MOT-16 によって事前学習された既存の Multi-Object Tracker (MOT)<sup>16)</sup> を用いて画像平面上での人物位置に相当する Bounding Box を抽出し、そのトラッキング結果に基づいて俯瞰視点上でのカメラ・歩行者移動軌跡復元を行った。図6に提案法の推定結果を示す。より実用的な場面を意識し、MOT の出力を軌跡復元の入力として用いた場合においても、混雑した環境下において提案法は安定したトラッキング結果を示した。

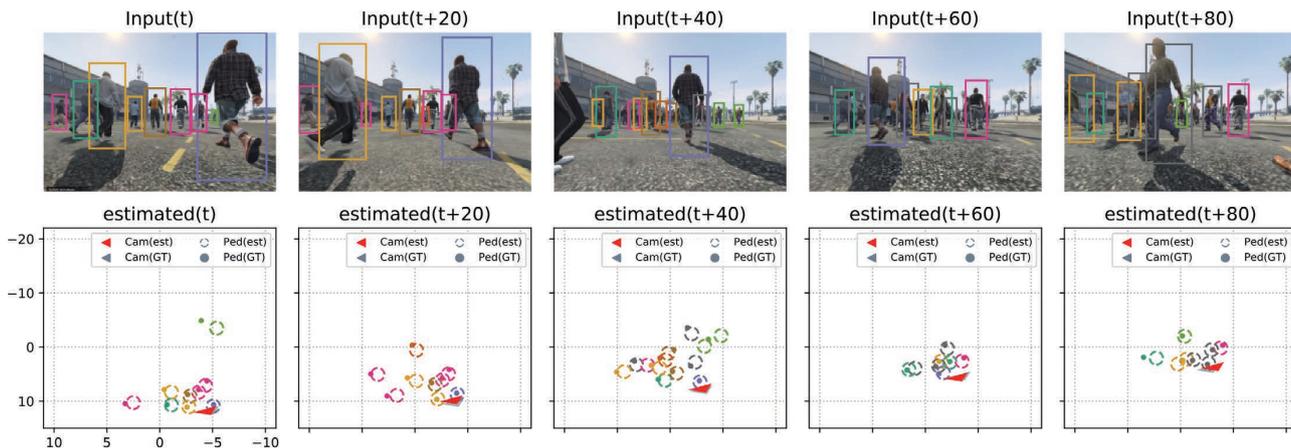


図 6 CG により生成した一人称映像 (上段) に対する俯瞰視点移動軌跡推定結果 (下段)

### 7. むすび

本稿では、多視点幾何に基づく移動ロボットの自己位置推定の基礎を俯瞰するとともに、動的な環境を対象とした拡張手法について紹介した。また後半では更に、静的なランドマークの観測を基本とした従来手法では扱うことができなかった混雑環境下において、移動体のみをランドマークとして用いることにより移動ロボット及び歩行者軌跡を復元する取り組みを提案した。移動体のみを利用するアプローチはテクスチャが乏しい背景環境においても適用することが可能であり、これまで従来法の枠組みで扱うことが不可能であった様々なシーンへ応用を広げることが期待される。

#### 参考文献

- 1) Bescos, B. et al. DynaSLAM: Tracking, mapping, and inpainting in dynamic scenes. *IEEE Robotics and Automation Letters*. 2018, Vol. 3, No.4, p.4076-4083.
- 2) Yu, C. et al. "DS-SLAM: A semantic visual SLAM towards dynamic environments". 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2018, p.1168-1174.
- 3) Henein, M. et al. "Dynamic SLAM: The need for speed". 2020 IEEE International Conference on Robotics and Automation (ICRA). 2020, p.2123-2129.
- 4) Huang, J. et al. "Clustervo: Clustering moving instances and estimating visual odometry for self and surroundings". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, p.2168-2177.
- 5) Mur-Artal, R.; Jose, M. M. M.; Tardos, J. D. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*. 2015, Vol.31, No.5, p.1147-1163.
- 6) Mur-Artal, R.; Tardós, J. D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Transactions on Robotics*. 2017, Vol.33, No.5, p.1255-1262.
- 7) Lv, Z. et al. "Taking a deeper look at the inverse compositional

- algorithm". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, p.4581-4590.
- 8) Lv, Z. et al. "Learning rigidity in dynamic scenes with a moving camera for 3d motion field estimation". *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, p.468-484.
- 9) Luo, Y. et al. "Where, What, Whether: Multi-modal learning meets pedestrian detection". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, p.14065-14073.
- 10) Pellegrini, S. et al. "You'll never walk alone: Modeling social behavior for multi-target tracking". 2009 IEEE 12th International Conference on Computer Vision. 2009, p.261-268.
- 11) Lerner, A.; Chrysanthou, Y.; Lischinski, D. Crowds by example. *Computer Graphics Forum*. 2007, Vol.26, No.3, p.655-664.
- 12) Wen, L. et al. "Detection, Tracking, and Counting Meets Drones in Crowds: A Benchmark". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, p.7812-7821.
- 13) Mehran, R.; Oyama, A.; Shah, M. Abnormal crowd behavior detection using social force model. 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009, p.935-942.
- 14) Rockstar Games. "Rockstar Games". <https://www.rockstargames.com>, (参照 2021-09-17).
- 15) Script Hook V. "Script Hook V". <http://www.dev-c.com/gtav/>, (参照 2021-09-17).
- 16) Wang, Z. et al. "Towards real-time multi-object tracking". *Computer Vision-ECCV 2020: 16th European Conference. Proceedings, Part XI 16*. Springer International Publishing, 2020, p.107-122.
- 17) 岡谷貴之. バンドルアドジャストメント. *情報処理学会研究報告*. 2009, Vol.2009-CVIM-167, No.37, p.1-16.

## 執筆者紹介



**西村 真衣 NISHIMURA Mai**  
オムロン サイニックエックス株式会社  
リサーチアドミニストレイティブディビジョン  
専門：コンピュータビジョン、GPGPU  
所属学会：IEEE

---

本文に掲載の商品の名称は、各社が商標としている場合があります。



# 4

---

## 技術解説・サーベイ

# 3次元画像計測におけるステレオマッチングの基礎から最先端まで

谷合 竜典

ステレオマッチングは、自動運転における環境認識や自律走行ロボットのSLAM、Augmented Reality (AR)、3次元スキャンなど、さまざまな分野において基盤的技術要素となっている3次元画像計測技術の一つである。コンピュータビジョン分野の歴史の中で、最も古くから取り組まれている問題の一つでもあるが、物理的拘束や幾何学的拘束に基づいて技術構築してきたこの古典的領域に、いま、AI（深層学習技術）が侵食し、新たな展望が見え始めた。本論では、ステレオマッチングの基本的な問題設定や課題、深層学習以前のアプローチ、そして深層学習以降の最新の研究動向を解説し、さらにオプティカルフローやマルチビューステレオなどの周辺分野の紹介も交えながら、この分野を包括的に俯瞰する。

## Binocular Stereo: From the Basics to the State of the Art

TANIAI Tatsunori

Binocular stereo is getting better, faster, stronger. This fundamental 3D sensing technique not only makes key building blocks in various technologies such as autonomous driving, augmented reality, and digital 3D scanning, but also provides a basis for other relevant problems in computer vision such as multiview stereo and optical flow. As the field of computer vision has experienced a major turning point since the emergence of deep learning, the field of stereo vision was not an exception, being strongly influenced by deep learning. This article reviews the basics of binocular stereo from its concept, challenges, and formulations, and further provides overviews of the past and current state of the art in the literature before and after the advent of deep learning.

### 1. まえがき

ステレオマッチングとは、同一の静止シーンを別視点からとらえた2枚の画像を用いて、画像中に写るシーンの奥行きを推定する問題であり、人間の両眼奥行き知覚を計算的に模倣した3次元画像計測技術の一つである。ステレオマッチングは、コンピュータビジョン分野において最も古くから取り組まれている問題の一つであり、同時に、現在でも国際会議などで盛んに研究が発表されている問題でもある。ステレオマッチング技術の応用は幅広く、自動運転における環境認識や自律走行ロボットのSLAM、Augmented Reality (AR)、3次元スキャンなど、さまざまな分野における基盤的技術要素となっている一方、コンピュータビジョン分野内でも、オプティカルフローやマルチビューステレオなどの他の問題に対する基礎的な問題と位置づけることができる。近年の深層学習の台頭は、コンピュータビジョン分野に大きな転換期をもたらしたが、ス

テレオマッチング分野もその例に漏れることなく深層学習の影響を強く受けた。本論では、ステレオマッチングの基本的な問題設定や課題、深層学習以前のアプローチ、そして深層学習以降の最新の研究動向を解説し、さらにステレオマッチングの周辺分野の解説も交えながら、この分野を俯瞰する。本論は、池内克史編纂のComputer Vision: A Reference Guideへ筆者が寄稿した“Binocular Stereo”の章<sup>1)</sup>の内容に対して、和訳および最新動向等を踏まえた加筆・修正をしたものである。

### 2. ステレオマッチングの基礎

#### 2.1 基本知識

ステレオマッチングにおける基本的な想定として、対象は静止シーン、つまり2枚の画像間で被写体が動いていないシーンとし、さらに2枚の画像の内容は互いに十分な視覚的重なりがあるとする。

ステレオマッチングの撮影システムは、通常は、向きが

Contact : TANIAI Tatsunori tatsunori.taniai@sinicx.com

揃えられた2台の同種のカメラを水平方向に並べて設置する。この際のカメラ間の距離を“ベースライン長”と呼ぶ。このとき、ステレオマッチング問題は、一方の視点画像の各画素について、それに対応する被写体上の点がある一方の視点画像のどの位置に写っているかを推定する問題、即ち、画像間の密対応点推定問題（dense correspondence estimation）に帰着する。特にステレオマッチングの対応点ペアは、画像上で同一の高さの水平スキャンライン上に存在するため、2点間の対応関係を通常は水平方向の座標の変位、すなわち視差（disparity）によって表す。したがって、ステレオマッチングは、2枚の画像間での視差を各画素について求める問題とも言える。

ステレオマッチングによって得られる出力の形式として、奥行きマップ（depth map）や視差マップ（disparity map）などがある。奥行きマップは、メートルなどを単位とした各画素の奥行き値を画像で表したもので、視差マップは画素数を単位とした視差値を画像で表したものである。これら2つの表現は、撮影システムのベースライン長とカメラの焦点距離などの情報があれば相互に変換できる。

## 2.2 背景および周辺知識

人間の両眼奥行き知覚のように、ステレオマッチングも三角測量の原理を用いており、その数学的原理はエピポーラ幾何学（epipolar geometry）により基礎づけられている<sup>2)</sup>。ただし、エピポーラ幾何学自体は、基本的に、画像間の対応点を既知としてカメラや物体間の3次元的な位置関係について数学的理解を与えたり、逆に、カメラ同士の位置関係を既知として画像座標上での被写体の位置関係について数学的理解を与えたりするもので、実際にどのようにして画像中に対応点を見つけるかまではその範疇にない。これはステレオマッチング問題の要となる。

我々が普段全く意識することなく行っている両眼奥行き知覚であるが、改めて少し考えると、実に高度な処理をし

ていることがうかがえる。例えば、左視点のある点が右視点のどこにあるかを探すとき、その一点のみの色で判断しても対応点を一意に定めることは難しい。したがって、ある点の対応点を探すにしても、その周辺の視覚的コンテキストを見ながら、広域的な手掛かりと局所的な手掛かりを総合してマッチングしていることが推察される。これは、局所的な手掛かりのみでは残る曖昧性を、広域的な手掛かりで補っているとも言える。似たようなことはステレオマッチングでも行うが、しかし、図1の上図に示すような、全くテクスチャがない領域や繰り返しパターンが存在するシーンだと、広域的な手掛かりをもってしても対応関係に曖昧性が生じて、不良設定問題となる。これらテクスチャなし物体や繰り返しパターン問題は、ステレオマッチングの基本的な課題に数えられる。また、鏡面反射する物体や、光沢のある物体など、視点によって見た目や明るさが変化する物体も、ステレオマッチングが原理的に難しい対象である。

このほか、二眼のステレオマッチングにおいて避けて通れない問題が、遮蔽（occlusion）である。これは、図1の下図に示すように、一方の視点画像で写っている被写体領域が、もう一方の視点画像において視界に入らない、或いは、他の被写体で遮蔽されることで画像中に写らない問題を指す。遮蔽が生じている領域では、真の意味での対応点は存在しないので、原理上、対応点/奥行き推定は失敗する。ステレオマッチング問題では、このような遮蔽領域は多かれ少なかれ必ず存在するため、基本的に遮蔽の取り扱い（occlusion handling）が必要になる。

ワイドベースライン・ステレオマッチングは、ステレオマッチングにおいて特にベースライン長が大きい場合を指す。ベースライン長の大小は対象シーンの奥行きスケールとの相対関係で定まるが、ベースライン長を大きくすることで、奥行きを何センチ単位まで測れるかといった意味での計測の限界精度を上げられるメリットがある。一方で、デメリットとして、視差の取りうる範囲が大きくな



図1 ステレオマッチングにおける繰り返しパターンとテクスチャ無し領域および遮蔽問題

上：背景に繰り返しパターンとテクスチャ無し領域が存在する。下：左右の視点画像においてマッチングが取れない2種類の遮蔽領域が存在する。

る、画像間の視覚的重なりが少なくなる、被写体の遮蔽領域が増える、画像間での被写体の見た目の幾何学的変化が大きくなる、などが挙げられ、奥行き推定の難易度は増す。

ステレオマッチングは、一方でまた、コンピュータビジョン分野の他の問題を特殊化、或いは、簡単化したものと見なせる。

例えば、マルチビューステレオ (multiview stereo) は 2 視点以上からの多視点画像を用いるもので、通常はカメラの位置関係や撮像モデルは既知とする一方、カメラ配置や画像枚数は任意とする。複数の視点画像を用いることで、マッチングの曖昧性を減らせ、また、より多くの物体表面が最低 2 視点以上から観測できるようになるため、遮蔽問題も軽減できる。このため、原理上、マルチビューステレオは二眼のステレオマッチングよりも高い精度が期待される。一方で、不規則なカメラ配置では、例えば物体正面側と裏面側など、視覚的重なりが少ない画像ペアが存在する場合があります、これらをマッチング対象から適切に除外するための視点選択 (view selection) という新たな問題をほらむ。また、不規則なカメラ配置下では、ある視点画像上でカメラと正面平行な小平面 (パッチ) が、別視点では大きく歪んで写る場合があります、単純な正方形パッチ同士の類似度評価では精度を落とす。このように、マルチビューステレオは、二眼のステレオマッチングを一般化した、より複雑なタスクであると言える。

一方、オプティカルフロー (optical flow) は、ステレオマッチングと同様に 2 枚の画像間での対応点を推定する問題であるが、奥行きではなく、被写体の動きそのものを推定する。対象となる 2 枚の画像は、動的なシーンを 1 台の可動カメラにより異なる時刻に撮影したものである。このときフレーム間の対応点の動き (フロー/flow という) は、被写体の動きによるものや視点カメラ自身の動き (ego motion) によるもの、或いは、その両方の複合効果かもしれない。もし被写体が全く動かず、カメラモーションが正面平行な左右方向であれば、オプティカルフロー問題はステレオマッチング問題と一致する。ステレオマッチングにおける対応点の動き (視差) は、被写体の奥行き方向の位置のみを変数として説明可能だが、オプティカルフローにおける対応点の動き (フロー) は、物体の位置や動き、未知の視点モーションと、説明変数が複雑化する。このため、二眼ステレオマッチングやマルチビューステレオでは視差や奥行きといった 1 次元の探索空間で推定可能であったが、オプティカルフローの推定では 2 次元の探索空間が必要となる。遮蔽問題も依然として存在し、静的シーン下での幾何学的解釈が可能だったステレオマッチングの遮蔽問題と比べ、動的物体の存在が遮蔽の発生原理をより複雑にしている。

ステレオ・シーンフローは、二眼のステレオマッチング

とオプティカルフローを組み合わせた問題で、移動可能な二眼カメラシステムにより撮影された 4 枚の画像 (2 視点×2 フレーム) から、シーンの奥行きと時間方向のフローを同時に求める問題である。これらの 2 種類の対応点情報は、フレーム間のカメラモーション情報と併せることで、3 次元空間での被写体表面の 3 次元的な動き (3 次元シーンフロー) を表すことができる。ステレオ・シーンフローは、ステレオマッチングやオプティカルフロー単体よりも扱える情報が増えるため、複数画像からの手がかりを適切に利用すれば精度向上や高速化が見込める<sup>3)</sup>。

このように、二眼のステレオマッチング問題は、コンピュータビジョン分野に存在する様々な密対応点推定問題の中で最も根本的な問題と考えることができ、また同時に、次節以降で述べるように、カメラキャリブレーションや、画像フィルタ、組合せ最適化など、同分野の様々な技術の上に成り立っている技術でもある。

### 2.3 数学的原理

ここでは、エビポラ幾何学にもとづくステレオマッチングの数学的原理を説明する。ここでの目標は、物体表面上の点の 3 次元座標が、どのようにしてステレオ画像間の対応点情報と紐づけられるかを数学的に理解することである。

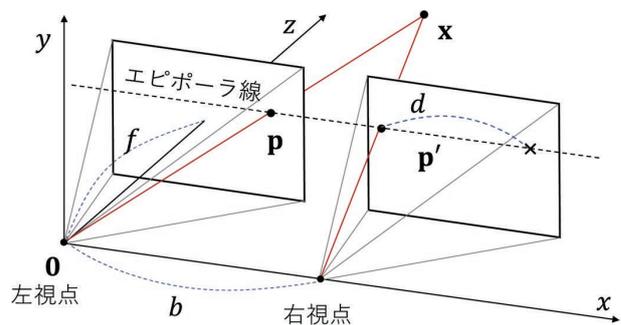


図 2 平行化されたステレオマッチングの設定  
同じ向きに左右に設置されたカメラで撮影されたシーンにおいて、3 次元点  $x$  が左右の視点画像に投影される状況を考える。このとき投影点の画像座標  $p$  および  $p'$  は、同じ高さのライン (エビポラ線) 上に位置する。カメラのベースライン長を  $b$ 、焦点距離を  $f$  とするとき、この対応点間の水平方向の座標の変位  $d$  は、3 次元点の奥行き  $z$  に対応して  $d = fb/z$  と表される。

ステレオマッチングでは一般に、図 2 に示すような、平行化された (rectified) 設定を考える。ここでは、2 台の同一のピンホールカメラが、どちらも 3 次元座標系の  $z$  軸方向の向きで、それぞれ原点  $(0, 0, 0)^T$  と、そこから水平方向にシフトした位置  $(b, 0, 0)^T$  に設置されているとする。このとき、原点の視点を左視点、もう一方を右視点と呼び、その間隔  $b$  をベースライン長とする。カメラは両方ともキャリブレーション済みとし、以下の同一の内部パラメータ行列を持つとする。

$$K = \begin{bmatrix} f & 0 & c_u \\ 0 & f & c_v \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

ここで、 $(c_u, c_v)$  はカメラの光学的な主点 (principal point) を表し、 $f$  は焦点距離 (focal length) を表す。なお、現実にはこのような理想的に平行化されたステレオ撮影システムを構築することは困難であるが、カメラシステムがキャリブレーション済み (即ち、2 台のカメラの相対的姿勢と内部パラメータが既知) であれば、ステレオ画像平行化 (stereo image rectification)<sup>2)</sup> と呼ばれる技術により、撮影画像をあたかも図 2 の平行化されたシステムで撮影したかのように変換できる。よって、以降はこの平行化されたステレオを前提に議論を進める。

原点に座する左視点を参照視点 (reference view) として、この視点画像に写る物体表面上のある点に注目し、その画素の座標を  $\mathbf{p} = (u, v)^T$  とする。我々が知りたいのは、参照カメラ座標系における、この対象点の 3 次元座標  $\mathbf{x} = (x, y, z)^T$  であり、これはピンホールカメラの投影モデルから以下のように表せる。

$$\mathbf{x} = K^{-1}(z\bar{\mathbf{p}}) \quad (2)$$

ここで、 $\bar{\mathbf{p}} = (u, v, 1)^T$  は  $\mathbf{p}$  の同次座標表現である。この 3 次元点  $\mathbf{x}$  が右視点画像上で写る位置、即ち、 $\mathbf{p}$  の対応点  $\mathbf{p}'$  の 2 次元座標は、以下の式で表される。

$$\mathbf{p}' = \pi(K[R \ \mathbf{t}]\bar{\mathbf{x}}) \quad (3)$$

ここで、 $[R \ \mathbf{t}]$  は一般にカメラ間の相対的姿勢を表す、回転行列 (rotation matrix)  $R$  と並進ベクトル  $\mathbf{t}$  (translation vector) から成る  $3 \times 4$  の姿勢行列で、今、カメラは同一方向に間隔  $b$  で設置されているため、 $R$  は 3 次元の単位行列、 $\mathbf{t} = (-b, 0, 0)^T$  である。姿勢行列は、 $\mathbf{x}$  の同次座標  $\bar{\mathbf{x}} = (x, y, z, 1)^T$  に対して作用する。関数  $\pi(x, y, z) = (x/z, y/z)^T$  は、透視投影変換に基づき 3 次元ベクトルを画像上の 2 次元座標に変換する。式(3)に対し、式(2)を代入して  $\mathbf{x}$  を消去すると、対応点ペア  $\mathbf{p}$  と  $\mathbf{p}'$  の関係式として以下が得られる。

$$\mathbf{p}' = \mathbf{p} - \begin{bmatrix} fb/z \\ 0 \end{bmatrix} \quad (4)$$

式(4)は、つまり、左画像上の各画素  $\mathbf{p}$  について、その右画像上での対応点の座標  $\mathbf{p}'$  は、座標  $\mathbf{p}$  から左方向に水平に  $(fb/z)$  画素分だけずらした位置にあることを意味する。この時の水平方向への座標の変位

$$d = fb/z \quad (5)$$

のことを視差 (disparity) と呼ぶ。式(5)からわかるように、点  $\mathbf{p}$  の視差  $d$ 、即ち、対応点  $\mathbf{p}'$  が判明すれば、 $\mathbf{p}$  にお

ける奥行き  $z$  は、ステレオシステムの焦点距離  $f$  とベースライン長  $b$  を用いて計算することができる。さらに、このようにして  $z$  が分かれば、式(2)により対象点の 3 次元座標  $\mathbf{x}$  も推定可能となる。

### 3. 非学習型のステレオマッチング

これまでの内容により、ステレオマッチングの基本的な定義や難しさ、他の問題との関連性、そして数学的な原理について理解できたはずである。以降では、より具体的な方法論について紹介することとし、ここではまず、深層学習登場以前の非学習型の古典的アプローチについて包括的に解説する。

#### 3.1 定式

古典的なステレオマッチング手法の俯瞰的な理解として、Scharstein と Szeliski<sup>4)</sup> による分類が有名である。彼らはステレオマッチングのアルゴリズムを、マッチングコストの計算 (matching cost computation)、コスト集約 (cost aggregation)、視差計算と最適化 (disparity computation and optimization)、視差微調整 (disparity refinement) の 4 つのステップに分割することで、ステレオマッチング手法の分類を与えた。本解説では、古典的なステレオマッチング手法を、以下の式で表される目的関数の定義および最小化方法の組み合わせととらえる新たな観点を導入して俯瞰的理解を試みる。

$$E(\mathbf{D}) = \sum_{\mathbf{p}} C_{\mathbf{p}}(D_{\mathbf{p}}) + R(\mathbf{D}) \quad (6)$$

ここで、変数  $\mathbf{D}$  は入力 of ステレオ画像ペアに対する視差マップを表し、 $D_{\mathbf{p}} \in \mathbb{R}_+$  は画素  $\mathbf{p}$  に対する視差を表す。 $C_{\mathbf{p}}(D_{\mathbf{p}})$  はマッチングコスト項と呼ばれるもので、画像の各画素  $\mathbf{p}$  に対する視差の推定値  $D_{\mathbf{p}}$  の妥当性を、 $\mathbf{p}$  とその対応点  $\mathbf{p}'$  における左右の画像の見た目の整合性 (photo-consistency) の観点で評価する。 $R(\mathbf{D})$  は正則化項 (regularization term)、或いは、平滑化項 (smoothness term) とも呼ばれ、視差マップ  $\mathbf{D}$  に対してある種の滑らかさを促す。通常、視差は離散的な視差値のラベリング  $D_{\mathbf{p}} \in \{d_1, d_2, \dots, d_K\}$  として定義され、式(6)の  $E(\mathbf{D})$  は、組合せ最適化 (離散最適化) アルゴリズムによって最小化される。これは、ステレオマッチングで用いられる目的関数  $E(\mathbf{D})$  は、一般的に非常に非凸な形をしており、勾配法などの連続最適化手法では容易に悪い局所解に陥ってしまうからである\*1。

この目的関数を主眼とした観点をを用いると、古典的なステレオマッチング手法は、ローカルモデル手法とグローバルモデル手法の 2 種類に分別できる。ローカルモデル手法は、式(6)において、マッチングコスト項のみによる目的関数  $E(\mathbf{D}) = \sum_{\mathbf{p}} C_{\mathbf{p}}(D_{\mathbf{p}})$  の最小化により視差マップを推定する。このとき、各画素の視差  $D_{\mathbf{p}}$  はその 1 変数関数

(unary function) であるマッチングコスト関数  $C_p(D_p)$  のみによって決定づけられ、その最適解は、視差の候補ラベル  $\{d_1, d_2, \dots, d_k\}$  の中で関数  $C_p$  の値が最も小さなラベルとなる。このような最適化方法を指して、しばしば勝者総取り (winner-takes-all) 方式と呼ぶ。

グローバルモデル手法は、正則項  $R(D)$  を明示的に持つ式(6)の目的関数  $E(D)$  を用いる。ローカルモデル手法と違い、各画素の視差変数  $D_p$  は  $R(D)$  を介して互いに影響し合うようになるため、 $E(D)$  の最適化にはより複雑な計算を要する。一部の特殊な場合を除いて、 $E(D)$  の最適化は一般に NP 困難となるため、グローバルモデル手法では近似解を得ることが目的となる。

以下では、これらローカルモデル手法およびグローバルモデル手法において重要となる要素や技術について解説する。

### 3.2 見た目の整合性尺度 (photo-consistency measure)

マッチングコスト項の中で、おそらくもっとも重要な要素は、左右の画像での画素間の類似度 (正確には非類似度) を測る尺度、即ち、photo-consistency 関数  $\rho(\mathbf{p}, \mathbf{p}')$  の定義である。この関数は、左画像  $I$  における画素  $\mathbf{p}$  と右画像  $I'$  における画素  $\mathbf{p}'$  の間の非類似度、或いは、それらの画素を中心とする小領域 (パッチ) の間の非類似度をスカラー値で評価する。最も単純な尺度は AD (absolute difference) と呼ばれるもので、 $\rho(\mathbf{p}, \mathbf{p}') = |I(\mathbf{p}) - I'(\mathbf{p}')|$  と定義される。しかし、画素輝度を直接比較することは 2 枚の画像間で照明変動があった場合に頑健でない。このため、輝度勾配  $\nabla I$  の差も考慮して

$$\rho(\mathbf{p}, \mathbf{p}') = \alpha |I(\mathbf{p}) - I'(\mathbf{p}')| + (1 - \alpha) |\nabla I(\mathbf{p}) - \nabla I'(\mathbf{p}')| \quad (7)$$

という定義を用いる場合もある。

より頑健な尺度として、2つの画像パッチに対する、以下のような NCC (normalized cross correlation, 日本語では正規化相互相関) を用いる場合もある。

$$\text{NCC}(\mathbf{p}, \mathbf{p}') = \sum_{i \in W} \frac{(I(\mathbf{p}_i) - \mu_p)(I'(\mathbf{p}'_i) - \mu_{p'})}{\sigma_p \cdot \sigma_{p'}} \quad (8)$$

ここで、 $\mathbf{p}_i, i \in W$  は左画像  $I$  の座標  $\mathbf{p}$  を中心とする矩形領域 (パッチ) の座標を表し、 $\mu_p$  と  $\sigma_p$  はその領域内でのグレースケール輝度  $I(\cdot)$  の平均と標準偏差を表す\*2。 $\mu_{p'}$  と  $\sigma_{p'}$  は右画像  $I'$  の座標  $\mathbf{p}'$  を中心とする矩形領域に対して同様に計算される。パッチサイズは  $3 \times 3$  から  $7 \times 7$  程度の比較的小きなパッチを用いることが多い。NCC は  $-1$  から  $1$  の範囲で正規化された類似度の尺度であるから、マッチングコストとして用いる場合、例えば関数  $\rho(\mathbf{p}, \mathbf{p}')$  を以下のように定義して用いる。

$$\rho(\mathbf{p}, \mathbf{p}') = \max \{1 - \text{NCC}(\mathbf{p}, \mathbf{p}'), \tau\} \quad (9)$$

ここで、閾値  $\tau$  を  $0 < \tau < 2$  の範囲の値 (典型的には  $\tau=1$ ) に設定することで、類似度が一定値以下の場合について NCC の値の変動を無視することができ、遮蔽領域などにおける外れ値に対して頑健な尺度を与えることができる。一方で、NCC は比較的計算コストが高いという問題がある。

Zabih と Woodfill<sup>6)</sup> は、画像パッチを 2 値のベクトルに変換することで、パッチ間の非類似度をハミング距離により高速に計算する手法、CENSUS 変換を提案している。CENSUS 変換は、あるパッチに対し、パッチ内の中央画素  $\mathbf{p}$  とその他の画素  $\mathbf{q}$  の輝度差の正負  $\text{sign}(I(\mathbf{p}) - I(\mathbf{q}))$  を 0 と 1 で符号化する。近傍画素間の輝度差に基づく尺度であるから、NCC と同様に照明変動に対して頑健であるが、量子化の度合いが強いため尺度の表現力は NCC より低い。CENSUS 変換は、屋外シーンに対するリアルタイム・ステレオ推定システムなどによく用いられる。

### 3.3 コスト集約 (cost aggregation)

photo-consistency 関数は上述の通り、1 画素、或いは、小さなパッチ単位で非類似度を評価するため、それ単体ではマッチングの曖昧性が高く、ノイズに弱い傾向がある。そこでコスト集約では、マッチングコスト  $C_p(d)$  を計算する際に、 $\rho(\mathbf{p}, \mathbf{p}')$  だけではなく、 $\mathbf{p}$  の周辺のサポート画素  $\mathbf{s} \in W_p$  におけるコスト  $\rho(\mathbf{s}, \mathbf{s}')$  も以下のように足し合わせて考慮することで、マッチングコストの推定の安定化を図る。

$$C_p(d) = \sum_{\mathbf{s} \in W_p} \omega(\mathbf{p}, \mathbf{s}) \rho(\mathbf{s}, \mathbf{s}') \quad (10)$$

ここで、 $W_p$  は対象画素  $\mathbf{p}$  を中心とするサポート窓 (support window) と呼ばれる領域であり、 $\omega(\mathbf{p}, \mathbf{s})$  は後述する何らかの重み関数、 $\mathbf{s}'_d = \mathbf{s} - (d, 0)^T$  は画素  $\mathbf{s}$  に対する視差  $d$  による対応点を表す。

コスト集約は、しばしばコストボリューム・フィルタ (cost volume filtering) とも呼ばれる。その理由を理解するために、式(10)の計算を、生のマッチングコスト値  $\rho(\mathbf{p}, \mathbf{p}'(d))$  の事前計算と、それらの集約の二段階で考えてみる。このとき、 $\rho(\mathbf{p}, \mathbf{p}'(d))$  を  $H \times W$  サイズの画像の全ての画素  $\mathbf{p}$  および全ての候補視差ラベル  $d \in \{d_1, d_2, \dots, d_k\}$  について事前計算すると、これらの値は  $H \times W \times K$  サイズの、コストボリューム (cost volume) と呼ばれる 3 次元ボリューム  $V(\mathbf{p}, d)$  を成す。コストボリューム  $V$  に対して式(10)の集約計算をすることは、 $V(\mathbf{p}, d)$  の同一  $d$  に沿った各 2 次元スライス  $V_d(\mathbf{p})$  (コストマップ/cost map という) に対して、カーネル  $\omega(\mathbf{p}, \mathbf{s})$  による画像フィルタリングを施すことに他ならない。通常、式(10)の計算を安易に実装すれば、各マッチングコスト  $C_p(d)$  の計算には  $O(|W_p|)$  の計算量がかかる。しかし、このコストボリューム・フィル

タの考えを導入し、フィルタ  $\omega(\mathbf{p}, \mathbf{s})$  に対して定数時間フィルタを用いれば、各  $C_p(d)$  の計算は窓サイズに依存しない  $O(|I|)$  の計算量で実現できる。

コスト集約操作の妥当性は、サポート窓  $W_p$  内の画素  $\mathbf{s}$  は中心画素  $\mathbf{p}$  と同じ視差（奥行き）を持つ、という仮定に基づいている。しかし、Blayer et al.<sup>7)</sup> が議論する通り、この仮定は次の2つの場合において頻繁に破られる。1) 窓内に物体の境界が存在するとき、および、2) 窓領域が正面並行ではない、大きく傾いた被写体表面を写しており、視差が窓領域内で大きく変動するときである。前者は、推定された視差マップにおいて、物体境界を滑らかにするアーティファクト（boundary flattening）を生じさせ、後者は、傾いた表面に対して階段状のアーティファクト（staircase artifact）を招く。これらの問題は大きなサポート窓を使うほど顕著になる一方で、マッチングコストの信頼性をあげるためには窓サイズを大きくする必要があり、このトレードオフが大きな足かせになっていた。

1つ目の問題は、適応的なサポート窓（adaptive support-window）を用いるアプローチ<sup>8)</sup>により効果的に対処可能である。適応的サポート窓は、対象画素  $\mathbf{p}$  とそのサポート画素  $\mathbf{s}$  の類似度により重み付け  $\omega(\mathbf{p}, \mathbf{s})$  を与え、サポート窓  $W_p$  の形を画像内容に応じて実質的に変形させるものである（図3）。Yoon と Kweon<sup>8)</sup> は、ジョイント・バイラテラルフィルタをコスト集約に適用した。これに対し Hosni et al.<sup>9)</sup> は、He et al.<sup>10)</sup> が考案したエッジ保持定数時間フィルタである Guided filter を用いることで、高速なコストボリューム・フィルタリングを提案した。

2つ目の問題の原因は、サポート窓に対して常に正面平行な面を仮定することで、視差マップ全体に正面平行バイアス（fronto-parallel bias）が生じることである。この問題に対して Bleyer et al.<sup>7)</sup> は、サポート窓に対して単一の視差  $d$  を推定するのではなく、以下のように、視差を平面式  $d=au+bv+c$  で表し、各画素  $\mathbf{p}$  に対して平面パラメータ  $(a, b, c)$  を推定した。

$$C_p(a, b, c) = \sum_{\mathbf{s} \in W_p} \omega(\mathbf{p}, \mathbf{s}) \rho(\mathbf{s}, \mathbf{s}'(as_u + bs_v + c)) \quad (11)$$

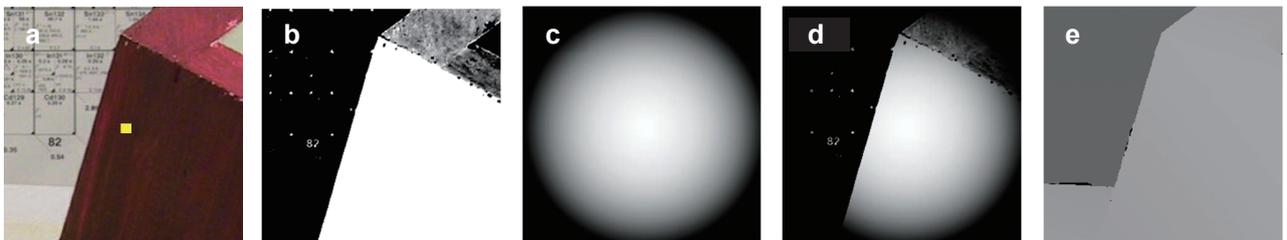


図3 適応的サポート窓

Yoon と Kweon<sup>8)</sup> による適応的サポート窓は、対象画素（aの中心点）のサポート窓（a）に対して、輝度類似度ベースの重み（b）と距離ベースの重み（c）を組み合わせたバイラテラル・フィルタカーネル（d）を計算し、マッチングコストを集約する。実際の奥行きマップ（e）と比べると、効果的に対象画素と異なる物体表面を集約計算から除外できている。

これは、窓領域で視差が線形に変化することを許容し、マッチングコスト計算で生じる正面平行バイアスを効果的に軽減した。一方で、視差の表現は1次元の離散値から、3次元の連続値  $(a, b, c)$  でパラメータ化された平面に代わり、この推定には後述の連続値視差推定のアプローチが必要となる。

### 3.4 正則化（regularization）

ローカルモデル手法では、これまでに述べた photo-consistency 関数とコスト集約を用いて、マッチングコスト関数をいかにして設計するかが中心となる。しかしながら、マッチングコスト関数だけではテクスチャが弱い領域、照明変動、画像ノイズなどの存在によって、推定精度が低下することがある。そこでグローバルモデル手法は、正則項  $R(\mathbf{D})$  を追加した目的関数  $E(\mathbf{D})$  の最小化を通じて、視差マップ  $\mathbf{D}$  全体を最適化し、モデルの精度向上を図る。

最も単純な正則化モデルは、以下のような線形モデル（linear model）である。

$$R(\mathbf{D}) = \lambda \sum_{(\mathbf{p}, \mathbf{q}) \in N} \omega(\mathbf{p}, \mathbf{q}) |D_p - D_q| \quad (12)$$

ここで、 $(\mathbf{p}, \mathbf{q}) \in N$  は画像グリッドの8近傍や4近傍などの近傍画素ペアで、 $\omega(\mathbf{p}, \mathbf{q})$  は近傍画素間の類似度を表したエッジ重み、 $\lambda$  は平滑化項全体の重みパラメータである。この線形モデルを用いた  $E(\mathbf{D})$  の最小化は、多項式時間で厳密解を求めることが可能である<sup>11)</sup>。一方、線形モデルは、色が近い近傍画素であれば近い視差をとるという仮定を用いているが、物体境界において過大なペナルティが発生し、境界付近の推定精度が悪くなる傾向がある。

このため、実用的には線形モデルに閾値  $\tau$  を導入した、以下の閾値付き線形モデル（truncated linear model）が最も広く用いられる。

$$R(\mathbf{D}) = \lambda \sum_{(\mathbf{p}, \mathbf{q}) \in N} \omega(\mathbf{p}, \mathbf{q}) \max\{|D_p - D_q|, \tau\} \quad (13)$$

このモデルを用いた  $E(\mathbf{D})$  の最小化は一般に NP 困難であるが、比較的シンプルな2変数関数（pairwise function）の

和で表せられるため、良い近似解が得られる最適化アルゴリズムが知られている(後述)。しかし、これら2つの線形モデルはどちらも近傍画素に対して同じ視差を好む正面平行バイアスがあり、傾いた被写体表面に対して階段状のアーティファクト(図4左)を生じさせる傾向がある<sup>12)</sup>。

このステレオマッチングの正則化における正面平行バイアスを軽減させるために、これまでに様々な正則化モデルが提案されている。Woodford et al.<sup>12)</sup> が提案した2階平滑化モデル(second-order smoothness model)は、 $D_p$  の1回微分値 $|D_p - D_q|$ の代わりに、3画素の組に対して計算される2階微分値 $|D_{q_1} - 2D_p + D_{q_2}|$ を評価する。図4に示すように、この正則化モデルは正面平行バイアスを大きく軽減できるが、3変数関数を含む目的関数の最適化は2変数の場合よりも複雑になる\*3。Olsson et al.<sup>14)</sup> は曲率ベースの正則化モデルを提案した。これは、画素ごとに視差平面を推定するモデルを用いることで、2変数関数形式によって定式化され、効率的な最適化が可能になった。Scharstein et al.<sup>15)</sup> は、事前に推定した物体表面の傾きを正則化モデルに埋め込むことで、最適化にかかる計算量の増加なしに正面平行バイアスを軽減する方法を提案した。

### 3.5 最適化 (optimization)

最適化はグローバルモデル手法において必須の工程で、2変数関数、或いは多変数関数の正則項を含む目的関数 $E(D)$ の最小化により、視差マップ $D$ 全体を最適化する。

視差が離散変数であり、かつ、目的関数が高々2変数関数しか含まない場合、離散最適化手法(組合せ最適化手法とも呼ばれる)を直接適用して近似解を求めることが可能である。この場合の最適化方法には確立されたものがあり<sup>16)</sup>、アルファ拡張/拡張移動法(alpha expansion/expansion move)に代表されるグラフカット法(graph cuts)ベースのアプローチと、信頼度伝搬法(belief propagation)

や Sequential Tree Reweighted Message Passing (TRW-S) 法に代表されるメッセージ伝搬法(message passing)ベースのアプローチに大別される。グラフカット法ベースの手法は、現在の視差マップの推定値に対して、画素ごとに現在ラベルと提案ラベルのどちらが良いかを選択する2値問題を反復的に解くことで、視差マップを更新していく。一方、メッセージ伝搬法は、各画素とその近傍画素の間で、ラベルに関する信頼度メッセージを交換しながら、コストボリュームを更新していく。これらの最適化アプローチは比較的高い計算量を要求するものが多く、リアルタイム・システムなどの実用場面においては、Semi-Global Matching (SGM) 法<sup>17)</sup> が用いられることが多い。SGM法は、メッセージ伝搬法の近似アルゴリズムであることが知られている<sup>18)</sup>。

### 3.6 連続値の視差の推定

視差は本来連続値であるため、より正確な3次元シーンの表現と推定には、連続値の視差マップ $D$ を推定する必要がある。

連続値の視差推定の一つのアプローチは、最適化の中で連続値の視差を求めることである。この場合、離散最適化手法を直接適用することはできないが、離散最適化手法を利用して連続値変数を最適化する、離散連続最適化アプローチが取られることが多い。例えば、セグメントベースの手法<sup>19)</sup> は、RANSAC法などを利用して事前計算したシーン中の平面を候補ラベルとして、スーパーピクセル領域に平面ラベルを割り当てる組合せ最適化問題として解く。融合ベースの手法<sup>12)</sup> は、多数の連続値表現の視差マップの解候補(proposal)を融合してより良い解を求めるもので、グラフカット法を用いた組合せ最適化問題として定式化される。融合ベースの手法では、解候補の生成方法が鍵となるが、例えば、様々な粒度のスーパーピクセル

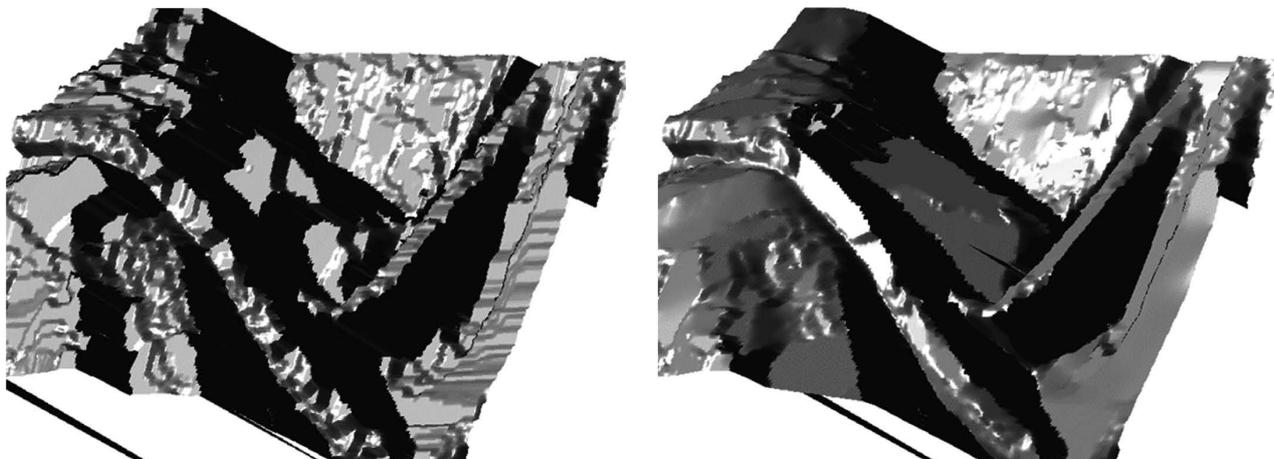


図4 正則化モデルによる正面平行バイアスの影響

左の線形モデル(1階平滑化モデル)には正面平行バイアスがあり階段状のアーティファクトが生じるが、右の2階平滑化モデルでは傾いた面もなめらかに推定されている。出典: Woodford et al.<sup>12)</sup> の発表資料より。

で推定したセグメントベース手法の結果などが用いられる(図5)。

PatchMatch ステレオ法<sup>7)</sup>は、PatchMatch と呼ばれる空間伝搬とランダム探索を組み合わせた対応点推定アルゴリズム<sup>20)</sup>を用いることで、各画素に対して連続値3次元ベクトルで表される視差平面を推定した。この手法は融合移動法と比べ、解候補の事前生成が不要である一方、正則化を考慮しないローカルモデル手法であった。PatchMatch のようなランダム探索を正則化付きのグローバルモデル手法として実現する研究として、Besse et al.<sup>21)</sup>による信頼度伝搬法ベースの手法や、筆者(Taniai et al.<sup>22)</sup>によるグラフカット法ベースの手法がある。

連続値の視差推定のもう一つのアプローチは、離散的な視差マップを推定した後にそれを後処理として精細化するものである。通常は、整数値の視差を精細化するため、そのような処理のことをしばしばサブピクセル精細化(sub-pixel refinement)と呼ぶ。具体的な方法として、メッセージ伝搬法によって得られる各画素の各候補ラベルに対する信頼度の値を用いて曲面フィッティングをする方法<sup>17)</sup>や、勾配法<sup>23)</sup>により目的関数  $E(\mathbf{D})$  を初期解付近で連続値最適化する方法などがある。

### 3.7 遮蔽対処 (occlusion handling)

遮蔽問題は、推定済みの視差マップに対して後処理として対処するアプローチ<sup>7,17)</sup>と、目的関数  $E(\mathbf{D})$  に遮蔽モデルを組み込むことで最適化の最中に対処するアプローチ<sup>24,25)</sup>がある。後処理ベースのアプローチは、左右の視差マップを用いての左右一貫性チェック(left right consistency check)による遮蔽領域検知と穴埋め(hall filling)によって行われる。この後処理は、ローカルモデル手法とグローバル手法のどちらにも用いることが可能である。一

方、最適化ベースのアプローチは、通常1変数関数であるマッチングコスト項を多変数関数に拡張する必要があるため、グローバルモデル手法でしか取り扱うことはできず、最適化の計算コストも高いが、より正確に遮蔽問題を扱うことができる。

## 4. 学習型のステレオマッチング

深層学習の台頭により、学習型のステレオマッチング手法が近年盛んに研究されている。とりわけ、ニューラルネットワークを用いて一貫学習するアプローチが主流で、これは以下のようなステレオ画像ペアから視差マップへのマッピング関数  $f$  を直接学習する。

$$\mathbf{D} = f(I, I'; \Theta) \tag{13}$$

関数  $f$  は畳み込みニューラルネットワーク(以下CNN)により実装され、そのパラメータ  $\Theta$  は大量の学習データに対して損失関数  $\ell(\mathbf{D})$  が最小になるように、確率的勾配降下法により最適化される。学習データセットによって正解の視差マップが与えられている際は、教師あり学習として、損失  $\ell(\mathbf{D})$  は推定視差マップ  $\mathbf{D}$  と正解視差マップとの差を平均L1損失やHuber損失(Smooth L1損失とも呼ばれる)などを用いて評価する。正解の視差マップの存在を仮定しない場合、損失  $\ell(\mathbf{D})$  には、式(6)の古典手法で用いられた目的関数  $E(\mathbf{D})$  に類似したものが用いられ、これは自己教師あり学習と見なされる。

このような学習型手法の利点として、これまで非学習型の古典手法においては、マッチングコスト計算における画像バッチの歪みや照明変動の問題、遮蔽問題、ステレオ平行化やキャリブレーションの誤差などの様々な問題に対して個別に対処が必要だったが、学習型のデータ駆動アプローチではこれらがほぼ自動的に解決できる。また、

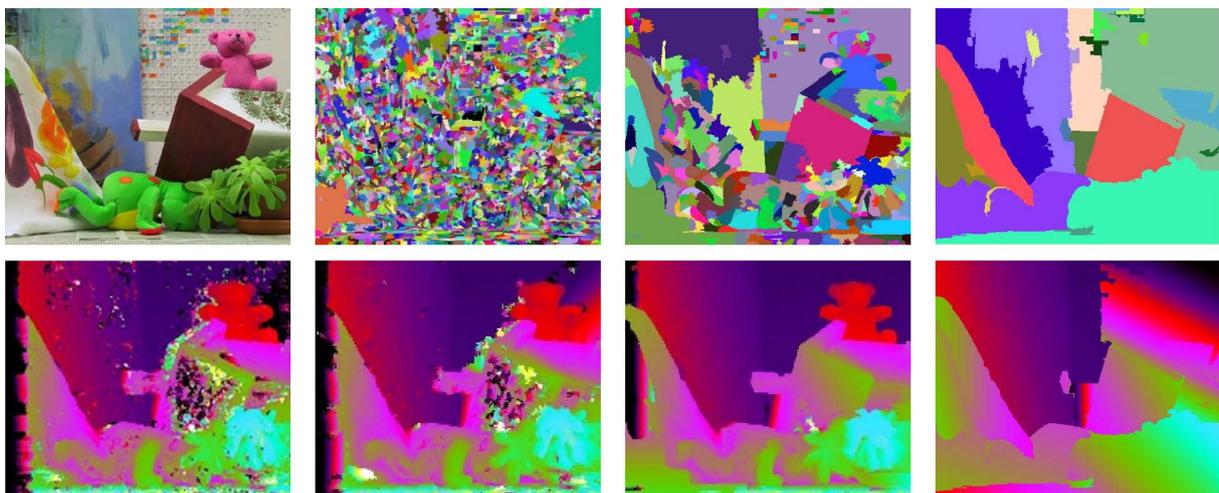


図5 融合ベースのステレオマッチング

上:参照画像に対し、様々な粒度のスーパーピクセルを推定。下:それぞれについてセグメントベース手法により視差マップ候補を生成する。視差マップ候補はグラフカットベースの最適化により融合される。出典:Woodford et al.<sup>12)</sup>より。

CNN の推論計算は GPU 上で大規模に並列化することが可能で、従来の古典手法では計算に時間がかかった連続値視差の推定も、高速に実行することができる。

既存の学習型手法を俯瞰的に理解するならば、これらはコストボリュームにもとづく古典ステレオ手法の模倣とみることができる。そこで本論では、Scharstein と Szeliski<sup>4)</sup> による古典ステレオ手法の分類に倣い、学習型手法のニューラルネットワークの構成を、特徴抽出 (feature extraction)、ボリューム構築 (volume construction)、コストボリューム学習 (cost volume learning)、視差回帰と精細化 (disparity regression and refinement) の 4 つの工程に分解した。図 6 は、このようなニューラルネットワークの典型的な例を図示したものである。以下では、学習型手法に関する既存研究の解説も交えながら、この 4 つの工程の内容を解説する。

#### 4.1 特徴抽出 (feature extraction)

初期の学習型手法は、ニューラルネットワークをマッチングコストの計算に特化して用いるものだった。Zbontar と LeCun<sup>26)</sup> による MC-CNN は、画像パッチから特徴ベクトルを抽出し、特徴ベクトル間のマッチングコストをコサイン距離 (fast 構成) や全結合層 (accurate 構成) によって計算した。学習されたマッチングコスト関数は古典手法に

おける photo-consistency 関数  $\rho(\mathbf{p}, \mathbf{p}')$  として用いられ、それ以外は古典的なステレオの処理パイプライン (ここでは SGM<sup>17)</sup>) にしたがって視差マップを推定する。

この特徴抽出は、後に、ステレオ一貫学習のアプローチ<sup>27-31)</sup> における種々の工程の最初のステージとしてネットワークに組み込まれた。一貫学習アプローチにおける特徴抽出の役割は、各入力画像を特徴マップに変換することであり、フィードフォワード型 CNN<sup>27)</sup> や ResNet 型 CNN<sup>28)</sup> の他、マルチスケール情報を効果的に利用する取り組みとして空間ピラミッド・プーリング (以下 SSP 層<sup>29)</sup> や U-Net 型 CNN<sup>31)</sup> を用いた実装がある (表 1 参照)。

#### 4.2 ボリューム構築 (volume construction)

学習型ステレオは、マッチングコスト関数の学習に特化したアプローチから直接視差マップの推定まで行う一貫学習アプローチへシフトしていった。その基本的な方針は、古典手法で用いられた 3 次元コストボリュームをネットワーク内で計算し、そこから視差マップを回帰出力するというものである。

このアプローチに代表される初期の事例が Mayer et al.<sup>27)</sup> による DispNetCorr (以下 DispNet) である。DispNet のコストボリューム構築は、左視点の特徴マップ上で右視点の特徴マップを横断させながら、特徴ベクトル間の相関を計

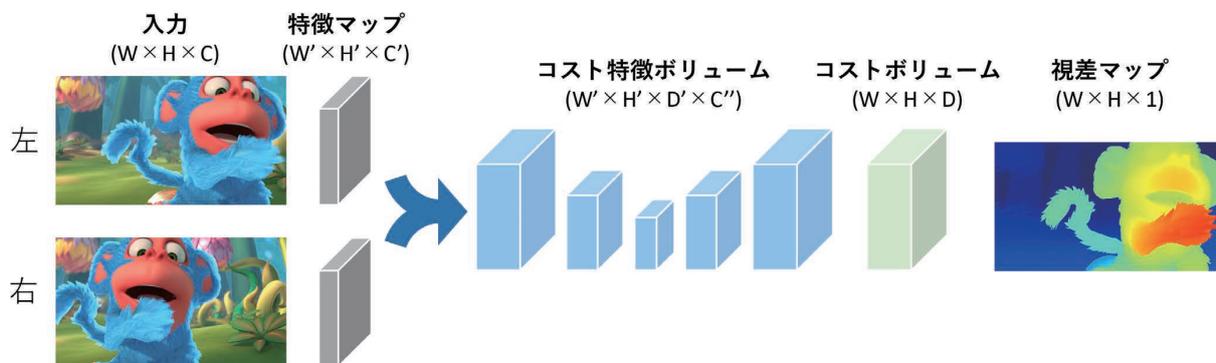


図 6 一貫学習アプローチのニューラルネットワークの構成例

学習型のステレオマッチングでは、通常、各入力画像を特徴マップに変換した後、右特徴マップを左特徴マップ上で横断させながら連結または相関計算することでコスト特徴ボリュームを構築する。コスト特徴ボリュームは、3 次元畳み込みネットワーク等によりコストボリュームに変換された後、soft-argmin 演算により視差マップが出力される。

表 1 一貫学習型のステレオマッチング手法の比較

手法	特徴抽出		ボリューム構築			コストボリューム学習	視差マップの回帰出力
	タイプ	特徴次元	方式	解像度	特徴次元		
DispNet <sup>27)</sup>	CNN	128	相関	1/4	1	なし	2D CNN
GC-Net <sup>28)</sup>	ResNet	32	連結	1/2	64	3D U-Net	soft-argmin
PSMNet <sup>29)</sup>	ResNet + SPP	32	連結	1/4	64	3D U-Net × 3	soft-argmin
GwcNet <sup>30)</sup>	ResNet	320	相関 + 連結	1/4	64	3D U-Net × 3	soft-argmin
GA-Net <sup>31)</sup>	U-Net	32	連結	1/3	64	SGA × 3 + LGA	soft-argmin

算するもので、以下のように計算される。

$$V(\mathbf{p}, d) = \text{dot}(\mathbf{F}(\mathbf{p}), \mathbf{F}(\mathbf{p} - (d, 0)^T)) \quad (14)$$

これは相関ベース (correlation-based) のボリューム生成と呼ばれ、強い帰納バイアスの恩恵を受けられる一方、ボリュームの各点をスカラー表現に落とし込むため、視覚的なコンテキスト情報などが欠落する問題がある。そこで、相関を明示的に計算せずに特徴ベクトルの連結で表現する、次式のような連結ベース (concatenation-based) の拡張<sup>28)</sup>も提案されている。

$$V(\mathbf{p}, d) = \text{concat}(\mathbf{F}(\mathbf{p}), \mathbf{F}(\mathbf{p} - (d, 0)^T)) \quad (15)$$

さらに Guo et al.<sup>30)</sup> は、相関計算を多チャンネル化したグループ毎相関 (group-wise correlation) と、連結ベースのボリューム生成を組み合わせ、相関計算の帰納バイアスと特徴ベクトルに含まれるコンテキスト情報の両方を利用する手法を提案している。

連結ベースやグループ毎相関ベースによって生成されるボリュームは、各マッチングコスト値を多次元の特徴ベクトルで抽象的に表したコスト特徴ボリュームと見なせ、4次元テンソルで表現される。このコスト特徴ボリュームはメモリ消費が多いため、通常は特徴抽出時において  $W/2 \times H/2$  や  $W/4 \times H/4$  のサイズにダウンサンプリングされた特徴マップを用いてボリューム構築する (表1参照)。画像サイズを各辺  $1/s$  に縮小すると、視差範囲も  $1/s$  になるため、特徴ボリューム全体のメモリサイズは  $(1/s)^3$  になる。生成された特徴ボリュームは、次のステップにより、明示的にマッチングコストを表す3次元テンソル表現のコストボリュームに変換される。

### 4.3 コストボリューム学習 (cost volume learning)

古典手法では、コストボリュームに対してコスト集約やSGMなどの最適化処理を施すことで、生のマッチングコストに対して正規化を適用し、推定精度を向上させた。学習型手法でも同様の効果をねらった方法が提案されている。

DispNet<sup>27)</sup> とならんで初期の一貫学習アプローチの代表例である Kendall et al.<sup>28)</sup> による GC-Net は、連結ベースのコスト特徴ボリューム (4次元テンソル) に対して、3次元畳み込み層による U-Net 型の3次元 CNN を適用することで正規化効果を学習する。Chang et al.<sup>29)</sup> はこの3次元 U-Net をカスケード化させた構成を提案した。また、Zhang et al.<sup>31)</sup> の GA-Net は、SGM 最適化処理を模倣した SGM 層と、適応的なサポート窓によるコスト集約を模倣した Local Guided Aggregation (LGA) 層を提案した。SGM の最適化処理は、元々微分可能な演算の組み合わせであったが、GA-Net では、max 関数を softmax 重みとの内積に置き換えるなど、勾配を効果的に伝搬する工夫がされている。

### 4.4 視差回帰と精細化 (disparity regression and refinement)

視差回帰は、コストボリュームから連続値の視差マップを出力する工程である。初期の学習型手法の例では、古典手法における離散視差推定のように分類ベースで行われたり、或いは、DispNet<sup>27)</sup> では、3次元のコストボリュームを2次元の特徴マップとして2次元 CNN により出力したりする方法がとられた。

これに対し、Kendall et al.<sup>28)</sup> が提案した soft-argmin 演算は、コストボリュームから効果的に連続値の視差マップを出力することを可能とし、以後の手法に標準的に用いられるようになった。soft-argmin は、3次元コストボリューム  $V(\mathbf{p}, d)$  に対して、視差  $d$  軸方向に softmin 関数を適用して確率分布

$$P_p(d) = e^{-V(\mathbf{p}, d)} / \sum_{d'=0}^{K-1} e^{-V(\mathbf{p}, d')} \quad (16)$$

を計算し、これにより視差の期待値

$$D(\mathbf{p}) = \sum_{d=0}^{K-1} d \cdot P_p(d) \quad (17)$$

を求めるものである。

このように出力された視差マップは、最終的に損失  $\ell$  と勾配が計算され、誤差逆伝搬により3次元 CNN や特徴抽出の CNN などのパラメータが学習される。コストボリューム学習にカスケード構成を用いる手法<sup>24,29,31)</sup> では、中間層から出力された視差マップに対する補助的な損失が追加される。また、最終的に出力された視差マップを、さらに2次元 CNN に通すことで精細化する場合もある<sup>32)</sup>。

### 4.5 データセット

学習型ステレオマッチングにおいては、学習データとなるデータセットや性能評価用のベンチマークの存在が重要になる。学習型ステレオマッチングの紹介の締めくくりとして、以下では、論文等において頻繁に用いられる代表的なデータセットを紹介する。

#### SceneFlow データセット

Mayer et al.<sup>27)</sup> は、ステレオマッチング、オプティカルフロー、およびステレオ・シーンフロー用の大規模な CG によるデータセットを構築した。このデータセットは、学習用データとして3万5,454組、テスト用データとして4,370組のステレオ画像ペアを提供する。各画像は、 $960 \times 540$  (0.5メガピクセル) のサイズで、正解となる視差マップのほか、前後のフレーム間のフローマップなども与えられている。Mayer et al. による DispNet<sup>27)</sup> 以降の学習型ステレオマッチング手法は、このデータセットを用いて事前学習するのが一般的となっている。

## KITTI 2012・2015 ベンチマーク

KITTI 2012 ベンチマークは、実際の車載カメラ画像を用いて、学習用およびテスト用に対してそれぞれ 200 組のステレオ画像ペアを提供している。各画像は、 $376 \times 1242$  (0.5 メガピクセル) のサイズで、学習用画像中の道路や壁や標識などの背景領域については LiDAR センサーによって計測された正解視差が与えられる。KITTI 2015 も同様の内容であるが、背景領域のみならず、車の領域に対しても 3D CAD モデルをフィッティングして推定した正解視差が与えられている。既存の論文では、SceneFlow データセットで事前学習したモデルで ablation study などの評価実験を行い、KITTI データセットで追加学習したモデルで実画像シーンに対する評価実験を行う場合が多い。

## Middlebury Stereo (バージョン 3) ベンチマーク

Middlebury ベンチマークは、2002 年の初期バージョン<sup>4)</sup>から始まる、比較的早くからあるステレオのベンチマークで、実際の屋内シーン画像に対し、structured light 技術で 3 次元計測した高精度な正解視差マップを提供している。現在のバージョン 3 では、学習用として 23 シーン、テスト用として 15 シーンあり、他のデータセットと比べシーンは少ない。もう一つの実画像によるベンチマークである KITTI と比べ、学習データが少ない点、大きな照明変動やテクスチャなし領域があるシーンを含む点、高解像度で視差の探索範囲が大きい点などの難しさがあり、リーダーボードでは、完全な学習型手法よりも、学習型マッチングコストと古典手法を組み合わせた手法<sup>22)</sup>が依然として強い傾向がある。

## 5. 他の問題への広がり

これまでに紹介した学習型のステレオマッチングのアプローチは、拡張される形で、オプティカルフローやマルチビューステレオなどの他の類似問題にも利用されている。以下では、そのような周辺分野への広がりについて簡単に紹介する。

### 5.1 オプティカルフロー (optical flow)

ステレオマッチングとオプティカルフローとの決定的な違いは、探索範囲の広さである。ステレオマッチングは、各画素について一定の深さまでの視差の探索範囲を定め、これをカバーするようなコストボリュームを構築する、いわば全探索アプローチがとられた。これは、ネットワーク上ではコスト特徴ボリューム (4 次元テンソル) に対する 3 次元 CNN 等により実装される。しかし、オプティカルフローでは探索空間が 2 次元に広がるため、同様の全探索をするには、5 次元テンソルに対する 4 次元 CNN が必要となり、メモリおよび計算コストが大幅に増大する。特に学習型手法では、誤差逆伝搬のために中間層の特徴マップや特

徴ボリュームの内容を全て保持しながら計算しなければならないため、メモリ上の制約が厳しい。

似たような問題は古典手法においても存在した。そこで、解決の糸口として古典手法のアプローチを振り返ってみる。古典手法の初期の代表例である Lucas-Kanade 法<sup>23,33)</sup>は、ステレオのような離散最適化アプローチではなく、勾配法に基づく連続最適化アプローチであった。その登場の背景には、当時のコンピュータの性能では、大量の計算とメモリを必要とするコストボリューム生成型の離散最適化アプローチが困難であった点も一因に考えられる。勾配法ベースの手法は、フローの初期値が正解値から大きく離れていると悪い局所解に陥りやすい。このため Lucas-Kanade 法では、画像ピラミッドの低解像度画像からフロー推定を行い、その結果をより高解像度の画像におけるフロー推定の初期値にしながら、順々に高解像度へと結果を伝搬、更新していく、いわゆる coarse-to-fine 推定のアプローチがとられた。その後、コストボリュームの一部を効率的にサンプリングして離散最適化するアプローチとして、融合ベースの手法と勾配法による精細化の組み合わせ<sup>34)</sup>や Patch-Match 法<sup>20)</sup>などが現れ、近年では、最適化処理に高速化の工夫を凝らしたフル・コストボリュームによる全探索アプローチ<sup>35,36)</sup>が登場するに至っている。

一貫学習アプローチの初期の代表例である FlowNet<sup>37)</sup>は、コストボリューム学習を提案した先駆的手法<sup>\*4)</sup>であったが、メモリ制約のため極度にダウンサンプリングしたコストボリュームを使用しており、精度は古典手法にも劣った。その後、解像度と使用メモリのトレードオフを改善するアプローチとして、フル・コストボリュームを用いない Lucas-Kanade 法を模倣した手法が多く登場した。例えば PWC-Net<sup>38)</sup>は、低中高の 3 つの解像度を持つ画像ピラミッドに対して、各解像度に対応したフロー推定ネットワークを用意し、低・中・高解像度と順々にフロー推定しながら結果を次のレベルでの初期値として伝搬していく。

中でも特筆すべきは、ECCV 2020 で Best Paper に選ばれた Teed と Deng<sup>39)</sup>による RAFT (Recurrent All-pairs Field Transforms) と呼ばれる手法だ。RAFT は、 $\mathbf{f}_0 = \mathbf{0}$  で初期化されたフローマップに対して、ニューラルネットワーク  $g$  を用いてフローの更新量  $\Delta \mathbf{f}_t = g(\mathbf{f}_t, \mathbf{I}, \mathbf{I}')$  を推定し、これを反復的に繰り返してフローマップを  $\mathbf{f}_{t+1} = \mathbf{f}_t + \Delta \mathbf{f}_t$  と更新していく (図 7)。これは、Lucas-Kanade 法による勾配ベースの更新をニューラルネットワークに置き換えたものと解釈できるが、他の模倣手法と大きく異なるのは、画像ピラミッドは用いず、最初から高解像度画像に対して推定できる点だ。また、反復の途中でネットワーク  $g$  を切り替える必要もなく、常に同じネットワーク  $g$  を使用して更新量  $\Delta \mathbf{f}_t$  を推定できる。

ネットワーク  $g$  の内部では、各画素について現在の推定対応点に基づいてマッチングコストをサンプリングする工

程 (図7のブロック“L”) と、マッチングコストと参照画像の特徴マップを入力として更新量を推定する工程 (図7のブロック“□”) がある。サンプリング工程では、単に現在の推定対応点のマッチングコストを計算するのではなく、その近傍を含めた複数のマッチングコストを、マルチスケール化された複数のコストボリュームからサンプリングする。このように広域的な情報を集めることで、高解像度画像が持つ大きなフローに対しても効果的に推定できる。また、更新量推定の工程では、GRU ベースの再帰レイヤーを用いている。これは、精細化用ネットワークなどの導入を不要にし、全ての反復工程を同一のネットワークで実行可能にする工夫と考えられる。

RAFT は従来のモデルと比べ、パラメータ数、推論速度、学習効率のすべての観点において、推定精度とのトレードオフを大幅に向上させた。また、CG による人工データセットによる事前学習だけでも、実画像に対する高い汎化性能を有することを示した。同様のアプローチは、ステレオマッチングにおいても、特にメモリコストが高い高解像度画像ステレオに対して有効であると筆者は考えている。

### 5.2 マルチビューステレオ (multiview stereo)

ステレオマッチングは参照画像  $I_0$  と目標画像  $I_1$  のペアに対する対応点推定問題であるが、マルチビューステレオは、目標画像が複数枚  $I_i (i=1, 2, \dots, N)$  に拡張された、 $N$  組の画像ペア  $\{I_0, I_i\}$  に対する対応点推定問題とみることができる。ただし、これらの画像は全て同一の静的シーンをとらえたもので、各視点のカメラ姿勢は既知であることを考慮すると、 $N$  組の画像ペアに対して個別に対応点マップ (視差マップ) を推定する必要はなく、参照画像に対して1枚の奥行きマップを推定すれば、それが全ての画像ペア  $\{I_0, I_i\}$  に対する対応点マップを表すことになる。

学習型のマルチビューステレオにおいて要となるのは、目標画像の数および視点位置が任意であるときに、それらに対する参照画像とのマッチング情報を、いかにして1枚の奥行きマップ推定の手がかりとして集約するかという点である。これをニューラルネットワークで実現するには、多少の工夫が必要となる。例えば、 $N$  組の画像ペア  $\{I_0, I_i\}$  について、特徴ベクトル同士のマッチングコスト  $C_i(\mathbf{p}) = |\mathbf{F}_0(\mathbf{p}) - \mathbf{F}_i(\mathbf{p}_i)|$  が  $N$  個計算できるとき、これらを単純に  $\mathbf{c} = (C_1, C_2, \dots, C_N)^T$  と連結し、学習可能な重み  $\mathbf{w}$  により  $C = \mathbf{w}^T \mathbf{c}$  と集約する方法が考えられるかもしれない。しかしこれは、学習時と推論時で一定の画像枚数  $N$  しか扱えず、また、入力の順序を  $\mathbf{c} = (C_N, C_{N-1}, \dots, C_1)^T$  のように変えただけで結果が変動してしまうといった問題がある。

学習型手法の代表例である MVSNNet<sup>40)</sup> (図8) は、古典的アプローチにおける平面スイープ法に倣い、参照画像に対して複数の候補平面を仮定し、各平面のホモグラフィ変換により、目標画像の特徴マップを参照画像上に投影してマッチングを行う。

奥行き  $z$  の候補平面によって投影された目標画像の特徴マップを  $\mathbf{F}_i^z$  とすると、各画素  $\mathbf{p}$  においては、参照画像と合わせて  $N+1$  個の特徴ベクトル  $\{\mathbf{F}_i^z(\mathbf{p})\} (i=0, 1, \dots, N)$  が重ね合わさる。これらの特徴ベクトルに対して、次式のように、分散を特徴次元ごとに計算することで、全体として1つのコスト特徴ボリューム  $\mathbf{V}(\mathbf{p}, z)$  (固定特徴次元を持つ4次元テンソル) に集約する。

$$\mathbf{V}(\mathbf{p}, z) = \frac{1}{N+1} \sum_{i=0}^N |\mathbf{F}_i^z(\mathbf{p}) - \bar{\mathbf{F}}^z(\mathbf{p})|^2 \quad (17)$$

ここで、 $\bar{\mathbf{F}}^z(\mathbf{p})$  は視点間の平均特徴ベクトルである。その後は、Kendall et al.<sup>28)</sup> の GC-Net とほぼ同様の構成で奥行きマップを回帰出力する。

式(17)からわかるように、MVSNNet は、学習後において

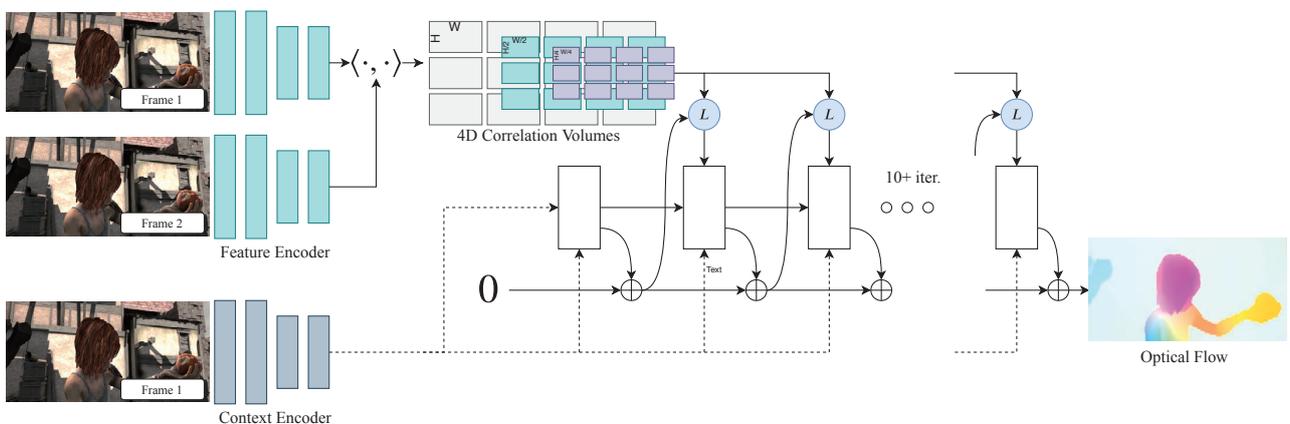


図7 RAFTのネットワーク構成

各画像を特徴マップにエンコードして4次元コストボリュームを構成した後、反復プロセスでは、現在の推定対応点周辺で4次元ボリュームを参照 (“L”ook Up) しながら、GRU が組み込まれた再帰レイヤーでフローの更新量の推定を繰り返す。出典: Teed と Deng<sup>39)</sup> より。

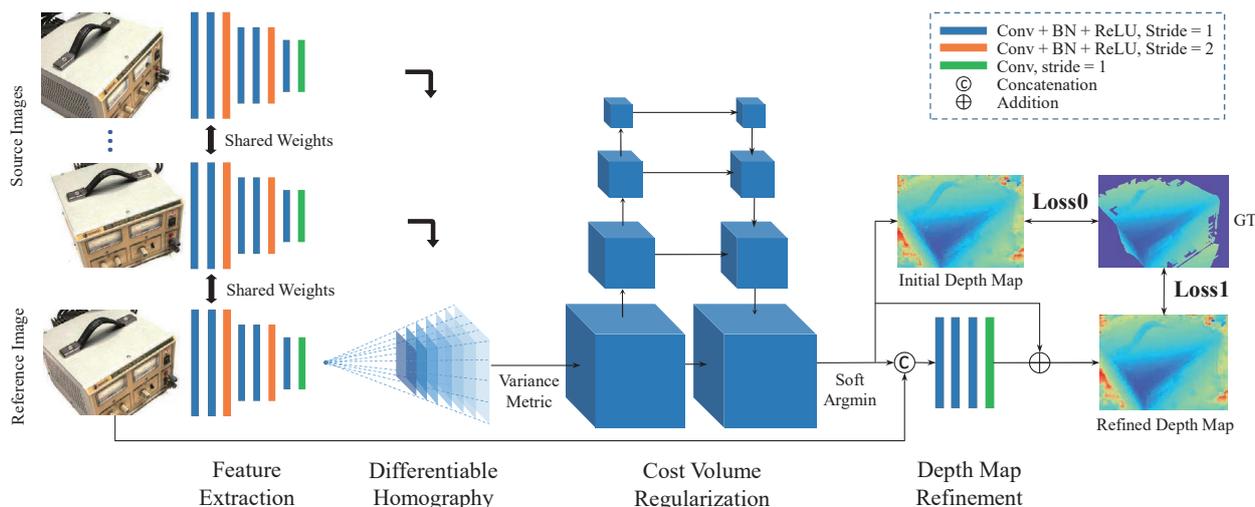


図8 MVSNetのネットワーク構成

参照画像に対して複数の奥行き平面を仮定。各視点画像の特徴マップをホモグラフィ変換で参照画像上に投影し、それらを分散ベースの集約層に通すことで3次元ボリュームを生成する。出典：Yao et al.<sup>40)</sup> より。

も目標画像の枚数  $N$  を自由に変えることができ、また、これらの画像の順番によって計算結果が変わることもない。実験では、目標画像枚数  $N = 2$  で学習したモデルにおいて、推論時に  $N$  を増やすことで推定精度が向上すること、さらに、分散ベースの集約が既存の平均ベースの方法よりも優れていることなどが確認された。

### 8. むすび

本論では、3次元画像計測におけるステレオマッチングの基礎から最先端までと題し、ステレオマッチングの基本的な定義や課題や定式の解説に始まり、深層学習登場前後での非学習型および学習型のステレオマッチング手法のアプローチを俯瞰的に解説した。さらに、深層学習ベースのステレオマッチング手法が、オプティカルフローやマルチビューステレオといった類似問題とどのように関連しているかを、実例を交えて紹介した。本論の執筆にあたり、いかに深層学習以前の古典的アプローチにおける種々のアイデアが、その後の深層学習ベースのアプローチに直接あるいは再解釈される形で取り入れられたかが分かるように書いたつもりであり、その点が読者に伝われば幸いである。

### 参考文献

- 1) Tani, T. "Binocular Stereo". Computer Vision. Ikeuchi, K., eds. Springer, 2020, p.83-92.
- 2) Hartley, R.; Zisserman, A. Multiple View Geometry in Computer Vision. 2nd. ed. Cambridge University Press, 2003, 655p.
- 3) Tani, T.; Sinha, S.; Sato, Y. "Fast multi-frame stereo scene flow with motion segmentation". Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2017, p.6891-6900.
- 4) Scharstein, D.; Szeliski, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. J. Comp. Vis. (IJCV). 2002, Vol.47, No.1/2/3, p.7-42.
- 5) Möllenhoff, T.; Laude, E.; Moeller, M.; Lellmann, J.; Cremers, D. "Sublabel-accurate relaxation of nonconvex energies". Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2016, p.3948-3956.
- 6) Zabih, R.; Woodfill, J. "Non-parametric local transforms for computing visual correspondence". Proc. Europ. Conf. Comp. Vis. (ECCV). 1994, p.151-158.
- 7) Bleyer, M.; Rhemann, C.; Rother, C. "PatchMatch stereo - Stereo matching with slanted support windows". Proc. British Mach. Vis. Conf. (BMVC). 2011, p.1-11.
- 8) Yoon, K.; Kweon, I. Adaptive support-weight approach for correspondence search. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI). 2006, Vol.28, No.4, p.650-656.
- 9) Hosni, A.; Rhemann, C.; Bleyer, M.; Rother, C.; Gelautz, M. Fast Cost-Volume Filtering for Visual Correspondence and Beyond. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI). 2012, Vol.35, No.2, p.504-511.
- 10) He, K.; Sun, J.; Tang, X. Guided Image Filtering. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI). 2012, Vol.35, No.6, p.1397-1409.
- 11) Ishikawa, H. Exact optimization for Markov random fields with convex priors. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI). 2003, Vol.25, No.10, p.1333-1336.
- 12) Woodford, O.; Torr, P.; Reid, I.; Fitzgibbon, A. Global stereo reconstruction under second-order smoothness priors. IEEE Trans. Pattern Anal. Mach. Intell. (PAMI). 2009, Vol.31, No.12, p.2115-2128.
- 13) Lempitsky, V.; Rother, C.; Blake, A. "LogCut - Efficient graph cut optimization for Markov random fields". Proc. Int. Conf. Comp. Vis. (ICCV). 2007, p.1-8.
- 14) Olsson, C.; Ulen, J.; Boykov, Y. "In defense of 3d-label stereo".

- Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2013, p.1730–1737.
- 15) Scharstein, D.; Taniai, T.; Sinha, S. N. “Semi-global stereo matching with surface orientation priors”. Proc. 2017 Int. Conf. 3D Vis. (3DV). 2017, p.215–224.
  - 16) Szeliski, R.; Zabih, R.; Scharstein, D.; Veksler, O.; Kolmogorov, V.; Agarwala, A.; Tappen, M.; Rother, C. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. IEEE Trans. Pattern Anal. Mach. Intell. (PAMI). 2008, Vol.30, No.6, 1068–1080.
  - 17) Hirschmuller, H. Stereo processing by semiglobal matching and mutual information. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI). 2008, Vol.30, No.2, p.328–341.
  - 18) Drory, A.; Haubold, C.; Avidan, S.; Hamprecht, F. A. “Semi-global matching: a principled derivation in terms of message passing”. Pattern Recognition. Springer, 2014, p.43–53.
  - 19) Birchfield, S.; Tomasi, C. “Multiway cut for stereo and motion with slanted surfaces”. Proc. Int. Conf. Comp. Vis. (ICCV). 1999, p.489–495.
  - 20) Barnes, C.; Shechtman, E.; Finkelstein, A.; Goldman, D. Patch-Match: a randomized correspondence algorithm for structural image editing. ACM Trans. Graph. 2009, Vol.28, No.3, Article 24.
  - 21) Besse, F.; Rother, C.; Fitzgibbon, A. W.; Kautz, J. PMBP: Patch-Match belief propagation for correspondence field estimation. Int. J. Comp. Vis. (IJCV). 2014, Vol.110, No.1, p.2–13.
  - 22) Taniai, T.; Matsushita, Y.; Sato, Y.; Naemura, T. Continuous 3d label stereo matching using local expansion moves. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI). 2018, Vol.40, No.11, p.2725–2739.
  - 23) Lucas, B.; Kanade, T. “An iterative image registration technique with an application to stereo vision”. Proc. Imaging Understanding Workshop, 1981, p.121–130.
  - 24) Kolmogorov, V.; Zabih, R. “Computing visual correspondence with occlusions using graph cuts”. Proc. Int. Conf. Comp. Vis. (ICCV). 2001, p.508–515.
  - 25) Wei, Y.; Quan, L. “Asymmetrical occlusion handling using graph cut for multi-view stereo”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2005, p.902–909.
  - 26) Zbontar, J.; LeCun, Y. Stereo matching by training a convolutional neural network to compare image patches. J. Mach. Learn. Res. 2016, Vol.17, p.1–32.
  - 27) Mayer, N.; Ilg, E.; Hausser, P.; Fischer, P.; Cremers, D.; Dosovitskiy, A.; Brox, T. “A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2016, p.4040–4048.
  - 28) Kendall, A.; Martirosyan, H.; Dasgupta, S.; Henry, P.; Kennedy, R.; Bachrach, A.; Bry, A. “End-to-end learning of geometry and context for deep stereo regression”. Proc. Int’l Conf. Comp. Vis. (ICCV). 2017, p.66–75.
  - 29) Chang, J.; Chen, Y. “Pyramid Stereo Matching Network”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2018, p.5410–5418.
  - 30) Guo, X.; Yang, K.; Yang, W.; Wang, X.; Li, H. “Group-Wise Correlation Stereo Network”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2019, p.3268–3277.
  - 31) Zhang, F.; Prisacariu, V.; Yang, R.; Torr, P. H. “GA-Net: Guided Aggregation Net for End-To-End Stereo Matching”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2019, p.184–195.
  - 32) Chabra, R.; Straub, J.; Sweeney, C.; Newcombe, R.; Fuchs, H. “StereoDRNet: Dilated Residual StereoNet”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2019, p.11778–11787.
  - 33) Baker, S.; Matthews, I. Lucas-Kanade 20 Years On: A Unifying Framework. Int. J. Comp. Vis. (IJCV). 2004, Vol.56, p.221–255.
  - 34) Lempitsky, V.; Roth, S.; Rother, C. “FusionFlow: Discrete-continuous optimization for optical flow estimation”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2008, p.8934–8943.
  - 35) Chen, Q.; Koltun, V. “Full Flow: Optical Flow Estimation by Global Optimization over Regular Grids”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2016, p.4706–4714.
  - 36) Xu, J.; Ranftl, R.; Koltun, V. “Accurate Optical Flow via Direct Cost Volume Processing”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2017, p.5807–5815.
  - 37) Dosovitskiy et al. “FlowNet: Learning Optical Flow with Convolutional Networks”. Proc. Int’l Conf. Comp. Vis. (ICCV). 2015, p.2758–2766.
  - 38) Sun, D.; Yang, X.; Liu, M.; Kautz, J. “PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume”. Proc. IEEE Conf. Comp. Vis. Pattern Recog. (CVPR). 2018, p.8934–8943.
  - 39) Teed, Z.; Deng, J. “RAFT: Recurrent All-Pairs Field Transforms for Optical Flow”. Proc. Europ. Conf. Comp. Vis. (ECCV). 2020, p.402–419.
  - 40) Yao, Y.; Luo, Z.; Li, S.; Fang, T.; Quan, L. “MVSNet: Depth Inference for Unstructured Multi-view Stereo”. Proc. Europ. Conf. Comp. Vis. (ECCV). 2018, p.785–801.

## 執筆者紹介



谷合 竜典 TANIAI Tatsunori

オムロン サイニクエックス株式会社

リサーチアドミニストレイティブディビジョン

専門：コンピュータビジョン、3次元計測、密対応点推定

所属学会：情報処理学会、IEEE

博士（情報理工学）

本文に掲載の商品の名称は、各社が商標としている場合があります。

- \*1 近年では、このようなステレオマッチングのような非凸最適化問題に対して、連続最適化アプローチの限界を克服しようとする試みもある<sup>5)</sup>。
- \*2  $\sigma_p$  や  $\sigma'_p$  は、実際にはゼロ除算による発散を防ぐため、小さな定数  $\epsilon$  を用いて  $\max\{\epsilon, \sigma_p\}$  などとして実装される。
- \*3 Woodford et al.<sup>12)</sup> は、さらに連続値の視差を仮定した、従来よりも複雑なモデルを用いており、これらの最適化アプローチを含めた内容が評価されて CVPR 2008 の Best Paper 賞を受賞している。ここで用いられた最適化法である融合移動 (fusion move) 法 (後述) は、同グループの別論文<sup>13)</sup> として既に発表済みであったが、一連の研究を契機に融合移動ベースの最適化法がその後広く取り入れられるようになった。
- \*4 FlowNe<sup>37)</sup> は DispNet<sup>27)</sup> と同じ研究グループから発表された手法だが、実は、FlowNet のほうが先行研究であり、DispNet は後に FlowNet を応用する形で、ステレオ・シーンフローを推定するネットワークの構成要素として提案された。

# 連合学習入門

米谷 竜

機械学習、とりわけ深層ニューラルネットワークのような複雑なモデルを学習する場合において、大量かつ多様な学習データを収集することは重要である。これに対して連合学習は、通信ネットワークに接続された多数のクライアント計算機がローカルに保持するデータを各々個別に活用することによって、結果的に大規模な学習を実現するアプローチである。連合学習において、クライアントのデータは一箇所に集約されることはない。その代わりに、クライアントは学習をコーディネートするサーバが配布したモデルを自身のデータによって学習し、その学習済みモデルをサーバに共有する。そしてサーバは収集したクライアントのモデルを統合することによって、より高性能なモデルを獲得する。本稿ではこのような連合学習に関するモチベーションや基本的なアプローチ、そしていくつかの代表的な研究について紹介する。

## Introduction to Federated Learning

YONETANI Ryo

Collecting a large-scale and diverse data collection is critical for machine learning especially when one needs to train a complex model like deep neural networks. Federated learning is an approach to enable a large-scale training by leveraging decentralized data stored locally by a population of client machines connected to the coordination server. Importantly, federated learning does not require clients to share their own data. Instead, clients will train a global model distributed from the coordination server using their own data and share the trained model with the server. A collection of the trained models is then aggregated to produce a higher-performance model. This article introduces the motivation, typical approaches, as well as some popular studies of the federated learning.

### 1. まえがき

スタンフォード大学の提供する Coursera によると、機械学習とは “the science of getting computers to act without being explicitly programmed” = 明示的にプログラミングすることなく計算機に (知的な) 行動をさせるための科学 (著者訳) である<sup>1)</sup>。いま、機械に何らかの意味で知的な行動をさせたいと考える。たとえば、ある写真にりんごが写っているかどうかを判断する機能を計算機上で実現したいとする。これを明示的なプログラミングにより行う場合、「画像中において RGB 値がこの範囲に入っている赤色のピクセルがこのような円形状の領域の一部として現れており、さらにその円領域の一部は光源の影響を受けてこのような RGB 値の範囲で色合いが変化し……」といったように、色や形状に関する特徴の定義とそのパラメタをすべて手動で指定することになる。りんごの品種やりんごの撮影された環境 (たとえば屋外か屋内か)、あるいは撮影された状

況 (木になっているのか手にもたれているのか) が多様になるほど、このプログラミングが困難になることは容易に想像できる。これに対して機械学習は、このようなプログラミングを「大量の事例 (学習データ) からの学習」に置き換える。すなわち、りんごがどのような色・形状をしており、それらが撮影環境や状況によってどのように変化するかという特徴の定義やそのパラメタを、大量の「りんごが映った画像」に基づいて自動的に獲得する。ニューラルネットや SVM といった機械学習モデルがこのようにしていったん “学習” されると、新たな画像にりんごが含まれるかどうかを自動的に判断できるようになる。深層学習がコモディティ化する以前は、ユーザ側があらかじめ特徴の抽出方法を定義し、そのパラメタを機械が学習するパラダイムが主流であった。一方深層学習では、特徴の抽出方法 (特徴表現、feature representation と呼ばれる) 自体もデータから学習可能となる。上では機械学習の中でも特に学習データにおいてモデルの入出力関係が陽に与えられる教師あり学習 (より具体的には物体検出タスク) を例に挙げた

Contact : YONETANI Ryo ryo.yonetani@sinicx.com

が、ほかにも教師なし学習や強化学習など、機械学習にはさまざまなバリエーションがある。

それでは機械学習によって、我々は労することなく機械による知的な行動を実現できるようになったであろうか。一実の所問題の難しさの種類が変わっただけであり、明示的なプログラミングの代わりに大量の学習データ（上記の例であれば大量のりんご画像）を収集することが必要となっている。機械学習の実応用においてしばしば問題となるように、学習データの質と量は、機械学習モデルの性能（何%の確率でりんごを見落とし・あるいは見過ごしてしまうか）に直結する。その結果、機械学習分野では「いかに大量・多様なデータを省コストで収集するか」「いかに少数のデータから高性能の機械学習モデルを学習するか」といった課題が継続的に取り組まれている。

本稿で取り上げる連合学習<sup>2)</sup>は、大量・多様なデータを省コストで活用するアプローチの一つである。既存の機械学習の多くは、学習データがあらかじめ一箇所に集約されている、あるいはある単一の計算機から全学習データに容易にアクセスが可能である状況を想定する。これに対して本稿では、なんらかのネットワーク上でサーバと接続された多数のクライアント計算機が、それぞれローカルにデータを分散保持しているという状況を考える。連合学習ではこのような状況において、(1) サーバがランダムなクライアントに学習をしたいモデル（グローバルモデル、と呼ばれる）を配布し、(2) クライアントは配布されたモデルを自身のデータによって学習、サーバに返送する。その後、(3) サーバはクライアントから送られてきたモデルを統合

することによってグローバルモデルを更新し、それをまた別のクライアントに配布する（図1も参照）。これらの手続きを繰り返すことにより、サーバは最終的にあたかも多数のクライアントのデータを集約して学習されたかのような、高性能なグローバルモデルを獲得することができるようになる。このような連合学習の枠組みにおいて、サーバは自身で大量のデータを集める必要がない。すなわち、「多数の他者にデータ収集と学習を依頼する」という形でデータ収集の問題を解決していると見ることができる。さらに連合学習は、クライアントは収集したデータそのものをサーバに提供する必要がないため、通信やサーバサイドのストレージ、さらにはプライバシーやセキュリティといった観点でも有望な手段である。

連合学習はさまざまな分野において実用に向けた検証が進められている。もともとはモバイルデバイスのキーボード入力における予測変換モデルの学習に利用されていた<sup>3)</sup>。これは、デバイスのハードウェアやOSが全クライアントで概ね共通である点で、連合学習を実施する比較的理想的な問題設定と見ることができる。また医療分野においては、複数の医療機関が患者のデータを直接共有することなく医療診断のための機械学習モデルを学習する手段として、連合学習が注目を集めている<sup>4)</sup>。

本稿では、このような連合学習についてその基礎を紹介するとともに、近年の研究で取り組まれている典型的な課題、さらには我々の研究事例も含め、実応用を見据えた取り組みのいくつかを紹介する。

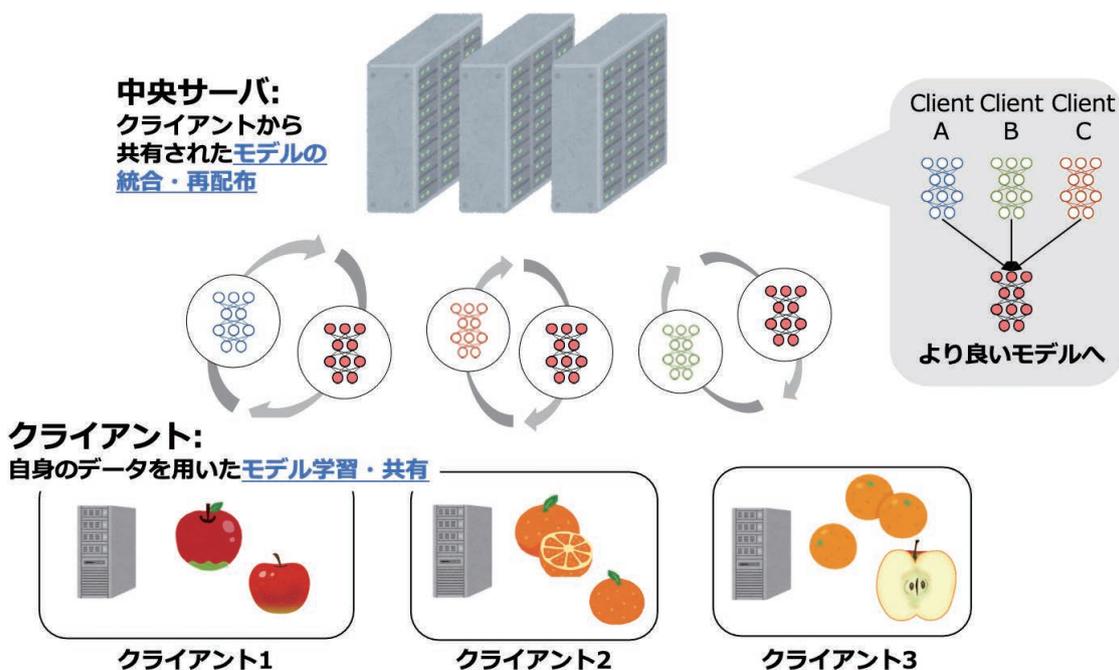


図1 連合学習

## 2. 連合学習の基礎

### 2.1 アルゴリズムの概要

連合学習では、サーバと通信が可能な多数のクライアントデバイス（以降単純に“クライアント”と呼ぶ）を考える。それぞれのクライアントはローカルにデータを保持しており、また機械学習が可能な計算資源を備えていることを想定する。一方サーバは何かしら学習したいモデルを保持しており、これをグローバルモデルと呼ぶ。サーバはランダムに選択されたクライアントに対してグローバルモデルを配布し、クライアントは自身のデータを用いてそのモデルを学習する。その後クライアントは学習済みのモデルをサーバに返送する。ここでサーバは選択されたクライアントと同数のモデルを保持することになる。サーバはこれらのモデルを統合し、新たなグローバルモデルとする。これが連合学習の1ラウンドであり、このような手続きを複数ラウンド繰り返すことで、グローバルモデルの性能を向上させることを目指す。

以下では、連合学習のもっとも基本的なアプローチである Federated Averaging (FedAvg)<sup>2)</sup> の具体的なアルゴリズムを示す。

---

#### アルゴリズム 1 Federated Averaging

$K$ : クライアント数       $C$ : クライアント選択の割合  
 $B$ : ミニバッチサイズ       $E$ : エポック数  
 $\eta$ : 学習率  
 $P_k$ :  $k$  番目のクライアントが持つデータ  
 $n_k$ :  $k$  番目のクライアントが持つデータのサンプル数  
 $\nabla L(w; b)$ : バッチ  $b$  についての損失  $L(w; b)$  の勾配

---

サーバ:

1. Initialize  $w_1$
2. for each round  $t=1, 2, \dots$  do
3.  $m \leftarrow \max(C \cdot K, 1)$
4.  $S_t \leftarrow$  (random set of  $m = \max(C \cdot K, 1)$  clients)  
    #  $K$  クライアントのうち  $C$  割をランダムに選択
5. for each client  $k \in S_t$ ; in parallel do
6.  $w_{t+1}^{(k)} \leftarrow \text{ClientUpdate}(k, w_t)$
7.  $w_{t+1} \leftarrow \sum_{k \in S_t} \frac{n_k}{n} w_{t+1}^{(k)}$      $n = \sum_{k \in S_t} n_k$     # モデルの統合

クライアント:

$\text{ClientUpdate}(k, w)$ :

8.  $B \leftarrow$  (split  $P_k$  into batches of size  $B$ )
  9. for each local epoch  $i=1, 2, \dots, E$  do
  10. for batch  $b \in B$  do
  11.  $w \leftarrow w - \eta \nabla L(w; b)$
  12. return  $w$  to server
- 

ここでは、合計  $K$  人のクライアントのうち、 $C$  割のクライアントが毎ラウンドランダムに選択されることになる。モデルはラウンド  $t$  のインデクス付きのパラメタ  $w_t$  で表現される。選ばれたクライアントは ClientUpdate サブルーチン

においてモデルを更新する。その結果得られるモデルはクライアントのデータに依存することになるため、クライアントのインデックスもついたパラメタ  $w_t^{(k)}$  で表現されることになる。11 行目にあるとおり、この更新はいわゆるミニバッチ勾配降下法を用いて、 $B$  エポックだけ行われる。最後に、サーバはクライアントが更新したモデルを、クライアントそれぞれが持つデータサンプル数 ( $n_k$ ) で重み付けされた平均処理によって統合する (7 行目)。

### 2.2 連合学習の特徴

上記のアルゴリズムは、同一のモデルを複数の計算資源にコピーして学習する点で、いわゆるデータ並列を用いたモデルの分散学習と類似している。ただし、連合学習は以下の点で分散学習と大きく異なる。

アプローチの観点から見ると、単一のミニバッチを複数の計算資源に分配する分散学習と異なり、連合学習では各クライアントがミニバッチ勾配降下を複数エポック実行する。すなわち、各クライアントがサーバに返送するモデルは、分散学習の例と比較してよりクライアントのデータに対してより適合したものとなっている。実際、文献<sup>2)</sup>ではこのエポック数を大きくすることがグローバルモデルの最終的なパフォーマンスに大きく寄与することが報告されている。一方で問題設定の観点から見ると、分散学習においてデータはあらかじめ一箇所に集約されており、学習において複数の計算資源に分散されるに過ぎない。一方で連合学習では、データはあらかじめ別々のクライアントによって独立に収集・保存されていることを想定する。このとき、クライアントによってデータの収集環境が大きく異なれば、クライアントごとのデータセットが持つ統計的性質もそれに依りて異なり、上記の FedAvg アルゴリズムを用いても効率的な学習が難しいことが知られている。これが連合学習における典型的な課題の一つであり、data non-iidness (データ分布の非独立・同一性) などと呼ばれ、その解決策が積極的に研究されている。

その他にも、クライアントによって計算資源や通信環境が異なる状況を想定することがある。連合学習では基本的に全クライアントのモデルがサーバに収集されたのち統合される (アルゴリズムの7行目) ため、学習やモデル送信により多くの時間を必要とするクライアントが含まれる場合、それが全体の連合学習を律速することになる。

## 3. 連合学習の最先端

2.2 節に述べた項目のみならず、連合学習にはいくつかの典型的な課題があり、それぞれ機械学習分野や通信分野において積極的に研究が進められている。本節ではそのいくつかを紹介する。

### 3.1 Data non-iidness への対応

連合学習ではクライアントがあらかじめ独立に収集したデータを活用してグローバルモデルを学習する。クライアントのデータ収集環境が多様であれば、最終的に学習されるグローバルモデルもより高い汎化性能を得ることが期待できる。しかしながら実際のところ、クライアントごとに学習データの分布が大きく異なると、学習の効率に悪影響を及ぼすことが知られている。例えば図1のように与えられた画像に映るりんごおよびみかんを検出するタスクにおいて、クライアント1はりんごの画像のみ、クライアント2はみかんの画像のみを保持していた状況を考える。このとき、FedAvgによってクライアント1が学習したモデルはりんごの検出に特化し、一方クライアント2のモデルはみかんの検出に特化することになる。それでは、これらのモデルのパラメータを単純に平均することで、りんごとみかんの両方を検出できるようになるであろうか。答えは否であり、両クライアントともりんごとみかんの画像を保持している状況と比較して学習に必要なステップ数が増加する。

これを解決するアプローチの一つが FedProx<sup>5)</sup> と呼ばれる手法である。FedProx における基本的な問題意識は、各クライアントが学習したモデルが、そのクライアントのデータに過適合してしまう点にある。これを防ぐために、同手法ではクライアントによるミニバッチ勾配降下におい

て、降下後のパラメータが元のグローバルモデルのパラメータから変化しすぎないように制約を与えている。具体的には、アルゴリズム 1 における 11 行目の更新を以下のように変更する。

$$w \leftarrow w - \eta(\nabla F(w; b) + \mu(w - w_i)) \quad (1)$$

ただし、 $w_i$  はこのラウンドにおいてクライアントが受け取ったグローバルモデルのパラメータであり、 $\mu$  は制約の強さを示すハイパーパラメータである。図2に FedAvg と FedProx の幾何学的な比較を示す。FedProx の方がクライアント1に過適合しない形でグローバルモデルを更新していることが分かる。

### 3.2 効率的な通信

連合学習では、サーバと選択されたクライアントが、学習対象となるモデルを送り合う。このモデルのサイズは学習に必要なデータよりは小さいことが多いものの、たとえばコンピュータビジョン分野でしばしば使われる ResNet101 モデルでは 171MB 程度となり、通信の効率化は重要である。これに対して、連合学習においてモデルのパフォーマンスを保ったままできるだけクライアントからサーバへの通信量を減らすアプローチが多く研究されている。本節では、Konečný らによる代表的な手法<sup>6)</sup> を紹介する (図3も参照)。

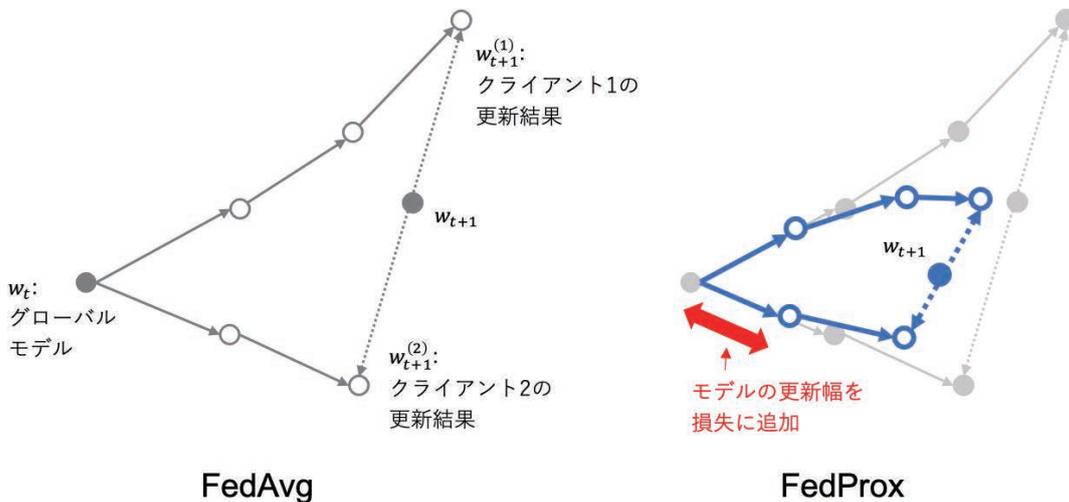


図2 FedAvg と FedProx の比較

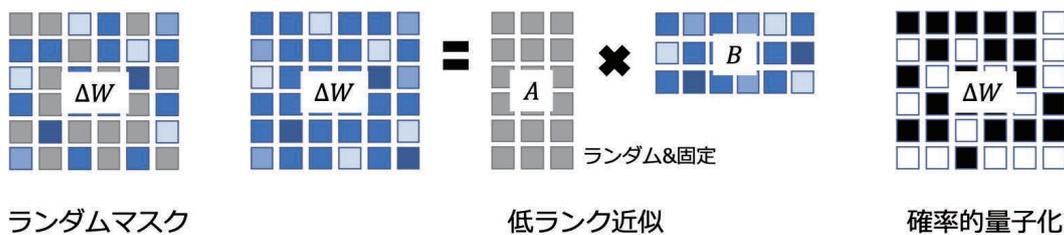


図3 効率的な通信法

FedAvg アルゴリズムでは、クライアントはグローバルモデル  $w_t$  を受け取り、それを更新したモデル  $w_t^{(k)}$  をサーバに返送する。このとき、サーバにとって  $w_t$  は既知であるため、サーバが受け取る新たな情報は、実質  $\Delta W = w_t^{(k)} - w_t$  という差分になる。したがって、クライアントはこの差分  $\Delta W$  をなんらかの手段で圧縮して送れば良いことになる。Konečný らのアプローチのうち最も単純なものはランダムマスク (図3左) であり、 $\Delta W$  と同じサイズで少量のランダムな要素が1、その他の要素がすべて0であるスパースなパターンを用意し、 $\Delta W$  と掛け合わせたものをサーバに送信する。これは  $\Delta W$  のうちランダムな数個の要素以外をマスクし、限られた要素のみをサーバに送信することと同義であり、転送量を大幅に低減できる。とはいえ、クライアントの学習結果のうち重要な情報が欠落する可能性もある。

より洗練されたアプローチとして、同論文では  $\Delta W$  をよりサイズの小さい行列同士の積で近似する (低ランク近似) 方法も提案されている (図3中)。いま、 $\Delta W$  がサイズ  $d_1 \times d_2$  の行列であるとする (たとえば  $d_1$  次元の特徴ベクトルから  $d_2$  次元の特徴ベクトルへ変換する全結合層を想像すれば良い)。このとき、適当な  $k < d_1$  を用いて、 $\Delta W = AB$ ,  $A \in \mathbb{R}^{d_1 \times k}$ ,  $B \in \mathbb{R}^{k \times d_2}$  という形に  $\Delta W$  を分解する。さらに、行列  $A$  はランダムに生成し、学習中は固定するものとする。これにより、各クライアントがサーバに対して  $B$  のみを送信すれば良いこととなり、その転送量も  $k/d_1$  まで削減されることになる。差分  $\Delta W$  における多くの要素を完全に捨ててしまう先のランダムマスクを用いたアプローチと異なり、本手法は  $\Delta W$  の全ての要素を (近似しつつ) サーバに共有できるメリットがある。

$\Delta W$  の全ての要素を近似することで転送量を低減させるもう一つのアプローチとして、確率的量子化 (図3右) がある。直感的には、深層学習フレームワークの中で通常32ビットの浮動小数点数型で表現される  $\Delta W$  を1ビット (2値) に2値化する。ただし、この際の手続きを確率的かつ各要素に対して適応的に行う。具体的には、 $\Delta W$  の最大値を  $\Delta w_{\max}$ 、最小値を  $\Delta w_{\min}$  としたとき、 $\Delta W$  中のある要素  $\Delta w$  を、 $\frac{\Delta w - \Delta w_{\min}}{\Delta w_{\max} - \Delta w_{\min}}$  の確率で  $\Delta w_{\max}$ 、 $\frac{\Delta w_{\min} - \Delta w}{\Delta w_{\max} - \Delta w_{\min}}$  の確率で  $\Delta w_{\min}$  に変換する。このとき、変換後の要素の期待値は元の  $\Delta w$  であることが簡単に確認できる。

### 3.3 クライアント選択と公平性

連合学習の各ラウンドにおいて、サーバは (1) ランダムなクライアントを選択し、グローバルモデルを配布する。そして、(2) 選択されたクライアントが学習したモデルを全て収集し、平均統合することにより、新たなグローバルモデルを獲得する。このとき、(1) において選択したクライアントの一部に限られた計算資源しか保持しない場合や不安定な通信環境にいる場合、サーバへのモデル返送が遅れることがある。その結果、サーバは (2) の平均統

合処理を実施できず、学習全体が遅れることとなる。

このような問題を解決する一つの手段が、適応的なクライアント選択<sup>7)</sup> である。このアプローチにおいて、サーバはあらかじめ全クライアントについてのモデル学習および送信にかかる時間の見積もりを知ることができるものとする。そして、たとえば10分間など限られた時間の中で、できるだけ多くのクライアントがモデルを送信できるように、クライアントの組合せを適応的に選択する。これにより、ランダムにクライアントを選択する場合と比較して、モデルが所望の性能に到達するまでの時間を大幅に短縮できることが実験的に示されている。

一方で、連合学習に参加するクライアントが偏ることや、公平性上の課題があることや、データに偏りがある際にモデルの性能が悪化することも知られている。Agnostic Federated Learning (AFL)<sup>8)</sup> はこのような課題に対応するためのアプローチである。いま、あるクライアントによるモデル  $w$  の更新を、損失関数  $L_k(w)$  の最小化として定義する。すなわち、 $L_k(w)$  に基づいてミニバッチ  $b$  で勾配降下する式がアルゴリズム1における11行目の手続きである。すると、複数クライアントが学習したモデルの統合は、以下の最小化によって表現することができる。

$$\min_w L(w) = \min_w \frac{1}{N} \sum_k L_k(w) \quad (2)$$

これに対して、AFL では以下のような minmax 問題を考える。

$$\min_w \max_\lambda L(w, \lambda) = \min_w \max_\lambda \frac{1}{N} \sum_k \lambda_k L_k(w) \quad (3)$$

上式において  $\lambda$  に関する最大化は、「現状のモデルで損失が大きいクライアントについての損失をより大きく重み付けする」という効果がある。そのようにして重み付けされた損失を  $w$  に関して最小化することにより、どのクライアントに対しても公平にモデルが学習されることになる。

### 3.4 クライアント特化モデル

上記の公平性とは少し異なった観点で、「個々のクライアントごとに特化 (パーソナライズ) されたモデルを学習したい」という要求がある。連合学習は通常、単一のグローバルモデルを学習することが最終的な目標となるため、それとは異なるアプローチが必要である。

これに対して文献<sup>9)</sup> は、Model-agnostic meta learning (MAML) と呼ばれるメタ学習のアプローチに着想したクライアント特化モデルの連合学習手法を提案している。同手法では、連合学習における最小化問題を以下のように変更する。

$$\min_w L(w) = \min_w \frac{1}{N} \sum_k L_k(w - \alpha \nabla L_k(w)) \quad (4)$$

この式は以下のように解釈できる: 各クライアントは損失関数  $L_k$  を「最終的なグローバルモデルのパラメタ  $w$  から

1ステップ  $\alpha \nabla L_k(w)$  だけ勾配降下すると最小化される」ように最小化する。このようにして得られたパラメタ  $w$  は全クライアントにわたって共通であるものの、各クライアントはローカルでモデルを  $w - \alpha \nabla L_k(w)$  という形で各自更新することで、自身のデータに関して最もよく当てはまる(つまり、 $L_k$  を最小化する)モデルを獲得することができる。これは言い換えれば、クライアントに特化したモデルを容易に獲得できるということを意味している。

### 3.5 暗号化

連合学習の持つ特徴の一つは、クライアントがサーバに対して自身のデータを直接共有しないでいいという点である。しかし、これは「クライアントのデータの機密が完全に守られる」ということは必ずしも意味しない。また、サーバ側はクライアントがどのようなデータを使って学習を行なったのかは必ずしも分からない、というのも重要な点である。これらの点について連合学習に対する攻撃を検証する研究も取り組まれている。たとえば、Inference attack は学習済みモデルから学習データの典型例を推測・生成したり、ある特定のサンプルが学習データに含まれていたかを推測したりする攻撃である。

このような攻撃を防ぐ方法の一つは、サーバに対してクライアントがモデルを共有する際に、モデルパラメタの具体的値を隠蔽することである。クライアントごとのサンプル数が同一であるという単純化をすると、連合学習におけるモデル統合は、結局のところ  $w_{t+1} = \frac{1}{N} \sum_k w_{t+1}^{(k)}$  という平均処理である。これに対して、 $b = \sum_k a_k$  という加算

を  $a_k$  の値を隠蔽したままで実現するセキュアな統合プロトコルが多く研究されており、連合学習のシナリオにも適用可能である。たとえば Bonawitzら<sup>10)</sup> が紹介している **masking with one-time pads** と呼ばれるアプローチでは、モデルの平均統合に先立って、各クライアント  $k$  は他クライアント  $j$  に対してランダムなベクトル  $s_{k,j} \in [0, R]^k$  を生成、交換する。そして、クライアントは  $w_{t+1}^{(k)}$  そのものではなく、 $x_{t+1}^{(k)} = w_{t+1}^{(k)} + \sum_j (s_{k,j} - s_{j,k})$  というランダムベクトルを加算したものを送信する。この時点でサーバは  $x_{t+1}^{(k)}$  から元の値  $w_{t+1}^{(k)}$  を知ることはできない。しかしながら、 $x_{t+1}^{(k)}$  を全クライアントにわたって加算すると、

$$\begin{aligned} \sum_k x_{t+1}^{(k)} &= \sum_k w_{t+1}^{(k)} + \sum_k \sum_j (s_{k,j} - s_{j,k}) = \\ &= \sum_k w_{t+1}^{(k)} + \sum_k \sum_j s_{k,j} - \sum_k \sum_j s_{j,k} = \sum_k w_{t+1}^{(k)} \end{aligned} \quad (5)$$

となり、 $\sum_k w_{t+1}^{(k)}$  という正しい統合結果を得ることができる。

### 3.6 教師なし学習への応用

連合学習に関する既存研究のほとんどは、ここまでに紹介したような教師あり学習タスクを想定している。一方冒頭に挙げた通り、機械学習には教師なし学習や強化学習などさまざまなタスクのバリエーションが存在する。ここでは我々の研究事例の一つとして、敵対的生成ネットワーク (Generative Adversarial Networks; GAN) を連合学習の枠組みで学習するアプローチ<sup>11)</sup> を紹介する。

いま、 $N$  人のクライアントがそれぞれ確率分布  $p_n(x)$  に従うデータセット  $X_n$  を保持している状況を考える (図4)。

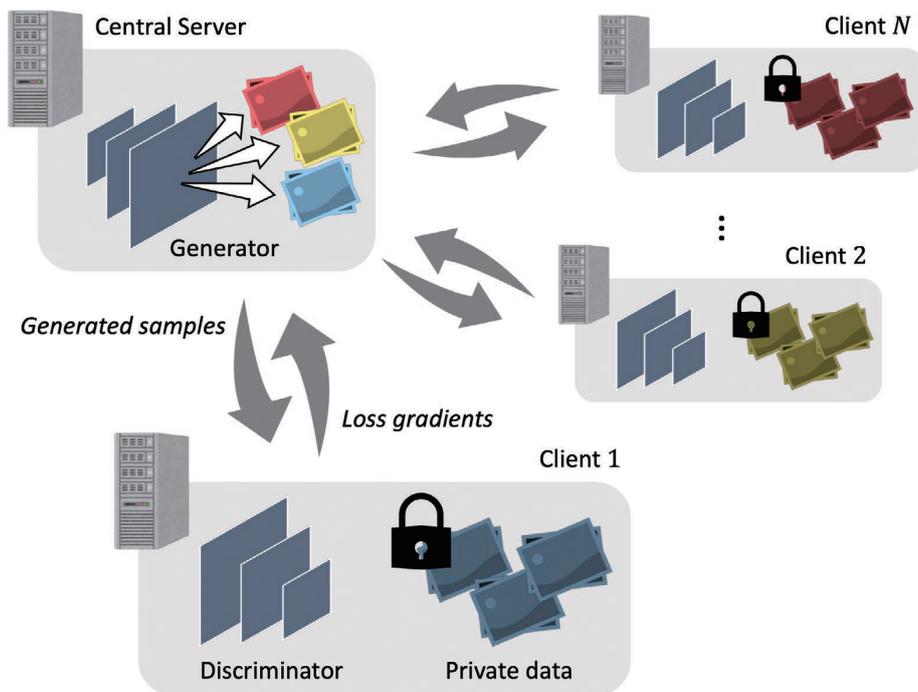


図4 連合学習の枠組みでのGANの学習

このとき、これらのデータセットに直接アクセスすることなく、全データを生成可能な分布  $p_{\max}(x) = \frac{1}{z} \max_n p_n(x)$   $z = \int_x \max_n p_n(x) dx$  に対応する生成モデル (generator)  $G(x)$  を学習することが本研究の目的である。提案するアプローチにおいてクライアントは、 $X_n$  のかわりに、そのデータと生成モデルが生成したデータを区別する識別器 (discriminator)  $D_n(x)$  を学習する。具体的には、 $D_n(x)$  は  $X_n$  からサンプリングされたデータについては 1、 $G(x)$  からのデータについては 0 を出力するように勾配法で学習される。サーバは各クライアントから  $D_n(x)$  を受け取り、 $D_{\max}(x) = \max_n D_n(x)$  を計算し、その出力が 1 に近づくように  $G(x)$  を更新する。これを繰り返すことにより、 $G(x)$  の大域的最適解が  $p_{\max}(x)$  となることが理論的に証明されている。

我々の提案するこのようなアプローチは、例えばファクトリーオートメーションの外観検査において、各現場間で不良品データに偏りがあり、個々の現場のみでは性能向上が困難な場合においても「各現場間でデータを直接やりとりすることなく」、検査アルゴリズムの性能をアップデートすることを可能にする。

#### 4. むすび

ネットワークで接続された多数クライアントが保持するデータを活用した機械学習の一つとして連合学習を紹介した。連合学習ではクライアントはサーバに対して学習済みモデルを共有するのみであり、自身のデータをローカルに留めたままにできるという利点がある。一方、クライアントによって保持するデータの統計的性質や計算・通信資源が異なる場合に学習が非効率化するという課題から、さまざまなアプローチが研究されている。

連合学習は機械学習および通信分野で近年とりわけ研究の活発なトピックの一つであり、ICML, NeurIPS, ICLR, AACL, IJCAI, ICC, GLOBECOM といった国際会議において多くの研究論文が発表されている。近年の研究動向は Kairouz らのサーベイ<sup>12)</sup> によくまとめられているので、興味のある読者は参照されたい。

- 5) Li, T. et al. "Federated Optimization in Heterogeneous Networks". Conference on Machine Learning and Systems (MLSys). 2020, p.1-16.
- 6) Konečný, J. et al. "Federated Learning: Strategies for Improving Communication Efficiency". NIPS Workshop on Private Multi-Party Machine Learning. 2016, p.1-10.
- 7) Nishio, T.; Yonetani, R. "Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge". IEEE International Conference on Communications (ICC). 2019, p.1-7.
- 8) Mohri, M.; Sivek, G.; Suresh, A. T. "Agnostic Federated Learning", International Conference on Machine Learning (ICML), PMLR. 2019, 97, p.4615-4625.
- 9) Fallah, A.; Mokhtari, A.; Ozdaglar, A. "Personalized Federated Learning with Theoretical Guarantees: A Model-Agnostic Meta Learning Approach", Annual Conference on Neural Information Processing Systems (NeurIPS). 2020, p.1-12
- 10) Bonawitz, K. et al. "Practical Secure Aggregation for Federated Learning on User-Held Data", NIPS Workshop on Private Multi-Party Machine Learning. 2016, p.1-5.
- 11) Yonetani, R.; Takahashi, T.; Hashimoto, A.; Ushiku, Y. Decentralized Learning of Generative Adversarial Networks from Non-iid Data, arXiv. 2019, preprint arXiv:1905.09684.
- 12) Kairouz, P. et al. "Advances and Open Problems in Federated Learning", Foundations and Trends in Machine Learning. 2021, Vol.14, No.1-2, p.1-210.

#### 執筆者紹介



米谷 竜 YONETANI Ryo

オムロンサイニクエックス株式会社  
 リサーチアドミニストレイティブディビジョン  
 専門：コンピュータビジョン、機械学習  
 所属学会：情報処理学会、電子情報通信学会、IEEE  
 博士 (情報学)

本文に掲載の商品の名称は、各社が商標としている場合があります。

#### 参考文献

- 1) Stanford Online. "機械学習", <https://www.coursera.org/learn/machine-learning>. (参照 2021-07-21).
- 2) McMahan, H. B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B. A. y. "Communication-Efficient Learning of Deep Networks from Decentralized Data". Proceedings of International Conference on Artificial Intelligence and Statistics (AISTATS), 2017, p.1273-1282.
- 3) Hard, A. et al. Federated learning for mobile keyboard prediction. arXiv. 2018, preprint arXiv:1811.03604.
- 4) Rieke, N. et al. The future of digital health with federated learning. NPJ Digital Medicine. 2020, Vol.3, 119.

# クロスモーダルな理解～サーベイ

牛久 祥孝

深層学習の恩恵として、各種データでの既存のタスクの精度が大きく向上したという点以外に、それぞれ隣同士だったはずなのに専門性の強いパイプラインを構築していた結果として生じていた相互参入障壁が大きく緩和された点も特筆すべき点である。例えば、今まで画像認識を専門として研究開発を進めていた研究者や技術者にとってみると、音声認識も導入しようとした際には音声信号処理の独自のパイプラインを習得する必要があった。その様にお互い名前や中身が大きく異なるパイプラインが、畳込みニューラルネットワーク (Convolutional Neural Network; CNN) や再帰ニューラルネットワーク (Recurrent Neural Network; RNN)、最近であれば Transformer などによって統一化されてきているためにお互いを理解しやすい状態になって来た。結果として、画像や音声信号、自然言語など多様なデータを行き来したり同時に理解したりするようなクロスモーダル理解の研究が大きく前進してきた。本稿では、そのようなクロスモーダルな形でデータを理解する種々のタスクについて概説する。

## Cross-modal Understanding: A Survey

USHIKU Yoshitaka

In addition to the fact that deep learning has dramatically improved the accuracy of existing tasks with various types of data, it is also worth noting that it has dramatically eased the barriers to mutual entry into neighboring fields. Such barriers had arisen as a result of the construction of highly specialized pipelines for each problem. For example, researchers and engineers specializing in image recognition would have had to learn their pipeline for speech signal processing if they wanted to introduce a speech recognition module into their image recognition system. These pipelines, which differ greatly in name and content, have been unified into Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and more recently, Transformer. As a result, it has become easier to understand the state-of-the-art pipelines in each other fields. Thus, research on cross-modal understanding has made significant progress in which diverse data, such as images, speech signals, and natural language can be traversed and understood simultaneously. In this article, the author outlines various tasks for understanding data in such a cross-modal manner.

### 1. まえがき

2021 年は、音声認識で深層学習が精度の大幅な向上を果たした<sup>1)</sup> 国際コンペティションからちょうど 10 年になる。その次の年には画像認識で同様に深層学習が劇的な精度向上を達成<sup>2)</sup> し、さらに 2 年後の 2014 年には機械翻訳で深層学習が非常に複雑なそれまでのシステムと同等の精度を達成する<sup>3)</sup> という事件が起きた。

これらの課題はそれぞれ、音声信号処理やコンピュータビジョン、自然言語処理と言われるコンピュータサイエンスを支える各領域で研究されていたものである。特に 2000 年頃から課題を解くためのお手本となる教師データを含めたデータセットを構築し、統計的機械学習手法を適

用したデータドリブンな解法がそれぞれで注目を集めていた。

その頃の各課題で主流となっていた手法は、基礎的な機械学習手法やいくつかのアイデアが共有されていたものの、基本的には各課題で独自に発達したデータの preprocessing や特徴量設計、そして後処理などが多段に組み合わされていた。つまり、例えば自然言語処理の研究者が画像も含めた新たな研究を始めようとしても、コンピュータビジョン分野の技術を習得して自身の研究に組み込むための導入コストが大きな壁となっていた。

一方で、それでも各分野の研究者が他の分野と融合したデータ理解を行おうとする取り組みも見られた。例えば動画認識として各時刻の画像情報と音声情報をそれぞれの特

Contact : USHIKU Yoshitaka yoshitaka.ushiku@sinicx.com

画像の内容を自然言語で説明する為に当時の画像認識と自然言語生成を参考にしながらパイプラインを構築したり<sup>5)</sup>といった研究が見られる。

その中で、冒頭に述べたような深層学習の研究が各領域でセンセーションを巻き起こした。音声認識で多層パーセプトロン (Multi-Layer Perceptron; MLP) によって、画像認識が CNN によって、そして機械翻訳が RNN によって塗り替えられたが、それら MLP/CNN/RNN と言うモジュールはすぐにそれぞれの分野で統一的に利用されるようになった。例えば音声認識や動画認識の時系列性は CNN や RNN で、そして一方機械翻訳分野でも CNN による系列理解が試みられるようになるといった具合である。既によく知られているようにこれらの深層学習は職人芸的な特徴量設計を必要としないことも相まって、お互いの分野での異なる課題へのパイプラインの見通しが極めて立ちやすくなった。

結果として、先ほど述べたように深層学習の流行前に徐々に取り組まれ始めていた複数のモダリティを同時に理解するような研究は、深層学習というロケットブースターを得て非連続な加速を果たした。例えば画像キャプション生成は 2010 年にその先駆けとなる研究<sup>6)</sup> が生まれ、2012 年には機械学習手法を含めたパイプライン<sup>5)</sup> が確立し、その後の精度向上が徐々に進んでいた。そして 2012 年の画像認識と 2014 年の機械翻訳の深層学習化<sup>2,3)</sup> によって直ちに深層学習の導入が進み、2015 年 6 月に開催された CVPR (Conference on Computer Vision and Pattern Recognition) というコンピュータビジョン分野の最高峰の国際会議では、CNN で画像を理解して RNN で文を生成するというアプローチによる画像キャプション生成が世界中の大学や企業から同時多発的に提案されるに至った。

このような形で、現在では音声/画像/自然言語といった多様なモダリティの自由自在な組み合わせの入力を受けつけ、その内容を統合理解して自由自在な組み合わせの出力を得るような深層学習の研究が広く行われている。本稿では、そうした複数のモダリティを入出する研究をクロスモーダルな理解と称して、その概観となるサーベイを提供することを目的とする。紙幅の都合上各分野に深く入った解説を与えるのは困難であるが、現在のクロスモーダルな理解の索引となることを目指す。

2 節では、クロスモーダルな理解としての諸研究を概観するにあたってキーとなるエンコーダとデコーダという概念について解説する。3 節から 5 節では、画像/自然言語/音声という代表的な 3 つのモダリティから 2 つを組み合わせたクロスモーダル理解の研究について述べる。すなわち、3 節で画像と自然言語、4 節で自然言語と音声、5 節で画像と音声を繋ぐ研究についてそれぞれ解説する。6 節はより多くの、またはより多様なモダリティを扱った研究についても触れながら、本稿の結びとして今後の展望について述べる。

## 2. エンコーダとデコーダ

まえがきではクロスモーダルな「理解」とあえてぼかした表現を与えたが、基本的には入力として与えられたモダリティのデータに含まれる情報を抽出し、それを既定のモダリティのデータに変換して出力することになる。

例えば画像認識課題では、入力が画像というモダリティであり、出力がクラスラベルという離散化されたシンボルに相当するモダリティである。音声認識は全体としては入力が音声、出力は自然言語というモダリティであり、機械翻訳は入力出力どちらも自然言語である。

さらに画像認識を例にとってもう少し中身を考えてみる。入力された画像から典型的には畳込みニューラルネットワークによって処理された途中のデータはベクトルとなって、単層以上の MLP を通していずれかのクラスに分類される。

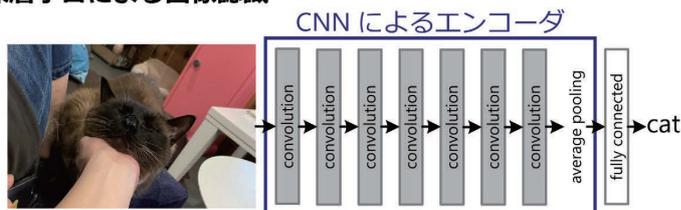
また機械翻訳で RNN を用いる場合、入力データの単語を 1 単語ずつ RNN に入力していき、翻訳したい文が全て入力されたときの RNN の隠れ変数ベクトルが得られる。これをまた別の RNN に入力し、1 単語ずつ訳文の単語を出力しつつ RNN に直前の出力を入力していくということを文の終わりという意味の出力が得られるまで繰り返す。

ここまでの 2 つの例で、途中に何らかのベクトルが出現していることに注意してほしい。つまり、深層学習においては多くの画素からなる画像や単語の系列である自然言語といった高次元だったり構造化されたりしているデータを、連続値からなる数列データに変換している。この数列データと言うのが典型的にはベクトルの形をとっている場合が多い、という訳である。このように、入力データをベクトル等の数列データに変換する機構をしばしばエンコーダ (符号化器) と呼ぶ。画像認識では CNN のエンコーダが出力したベクトルを MLP で分類しているし、機械翻訳では最初の RNN がエンコーダとして出力したベクトルが、次の RNN に与えられている。

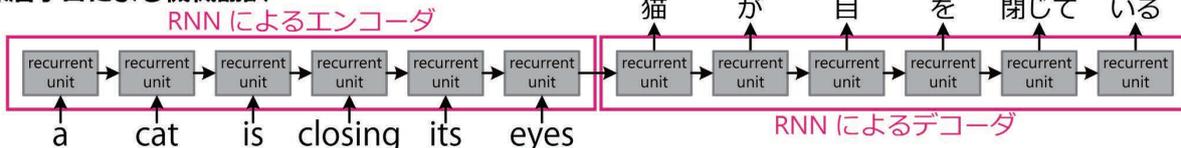
では、この機械翻訳の 2 つめの RNN は何か。このように研究課題によっては出力として自然言語や画像などの高次元ないし構造化されたデータを出力することが求められる。他によく知られたものとしては自己回帰を用いた音声合成や、Generative Adversarial Net (GAN) または Variational Auto-Encoder (VAE) を用いた画像生成などが挙げられるだろう。そしてこれらの自己回帰や GAN/VAE によるデータ生成の際にも、機械翻訳の 2 つ目の RNN と同様にやはり何らかのベクトルが入力されている。このように、ベクトルなどの数列データを入力として受け取り、音声や画像、自然言語のようなモダリティのデータを出力する機構をしばしばデコーダ (復号化器) と呼ぶ。

このように、画像認識はエンコーダと分類器の、機械翻訳はエンコーダとデコーダの組合せとみることができる。他にエンコーダやデコーダという単語が良く出てくる分野

### 深層学習による画像認識



### 深層学習による機械翻訳



### 深層学習による画像キャプション生成

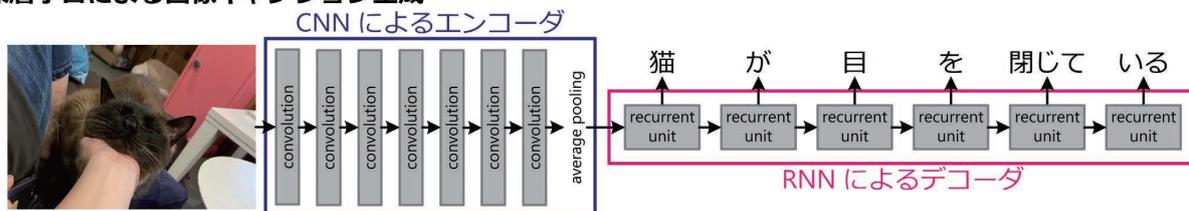


図1 深層学習による画像認識、機械翻訳、画像キャプション生成。このように各課題は入力データをベクトルなどの数列データに変換するエンコーダ（符号化器）と、数列データから出力データを生成するデコーダ（復号化器）からなる。このようなエンコーダ・デコーダの活用はクロスモーダルな理解を概観するにあたって必須の観点となる。

としては、映像などを符号化して再度映像に戻す際にもやはりエンコーダとデコーダが登場する。このように、情報を符号化するという点がエンコーダに、復号化するという点がデコーダには期待されている。画像認識の分類器がデコーダと言われないのは、この点が理由であろう。

ここまでの議論をまとめたのが、図1の「深層学習による画像認識」および「深層学習による機械翻訳」である。ここまでの議論を見ていると、次のようなことを考えないだろうか。画像認識のエンコーダは何らかのベクトルを出力している。これは画像を分類する為の情報が符号化されていると期待されるベクトルである。次に機械翻訳のデコーダはベクトルから文を出力している。文を出力するための情報がベクトルにちゃんと入っていれば、あとはデコーダがちゃんと自然言語文を出力してくれる。であれば、CNNによるエンコーダとRNNによるデコーダを組み合わせれば、入力画像を説明できるようなキャプション生成のパイプラインが実現できるのではないだろうか。

このようにしてCNNによるエンコーダとRNNによるデコーダを世界中の機関が同時に提案してきたのが、1節で述べた2015年のCVPRである。もちろん画像認識を学習しただけのCNNエンコーダと機械翻訳を学習しただけのRNNデコーダを直列してそのまま画像のキャプションが生成できるという訳ではなく、画像キャプション生成用の画像とお手本キャプションのペアからなるデータセット上で学習する必要がある。また深層学習による画像キャプ

ション生成のパイプラインが確立する前でも、実際には画像認識のモジュールと機械翻訳などの自然言語生成モジュールが組み合わされたパイプラインが提案されており、このエンコーダ・デコーダ構造を深層学習が初めてもたらしたわけではない。このように多様なモダリティのデータを解釈したり生成したりと言う課題がそのままエンコーダとデコーダとして解釈できる、といういわば観点のようなものが見えてきたのが重要である。3節以降で解説するクロスモーダルな理解の各分野でも、このように複数のエンコーダとデコーダが登場してくる。

### 3. 画像と自然言語を繋ぐ研究

画像や自然言語処理でCNNやRNN、はたまた最近のTransformerの表現学習が成功しているのは、データセットとして大量の画像やテキストを収集可能であることも大きい。そのように大量なモデルで、帰納バイアスの少ない巨大パラメータによるネットワークを学習することの恩恵を受けているのがこれらのモダリティのデータであるともいえる。そのようなデータとして、YouTubeやFacebookといったwebサービスを中心として、画像や動画像と関連するテキストが大量に投稿されている。画像ないし動画像と自然言語というペアのデータは、ドメインさえ絞らなければ大量に集めやすいマルチモーダルデータでもある。

本節では、こうした画像と自然言語を繋ぐクロスモーダル理解について述べる。この二つのモダリティの組合せは

特にビジョン&ランゲージと言う名前がついていて、コンピュータビジョンの国際会議でも自然言語処理の国際会議でも数年前から一定の存在感を示すようになってきている。

なお、画像から自然言語を出力するタスクと言うと光学的文字認識 (Optical Character Recognition; OCR) を想像するかもしれない。確かに、深層学習以前から種々の方法で文字認識がさかんに試みられている。深層学習の歴史の中でも、Yann LeCun による LeNet<sup>7)</sup> は学習可能な畳込みニューラルネットワークとして前世紀から非常に有名なモデルで、文字認識でその効果が確かめられている。OCR は画像から自然言語へのクロスモーダルな変換を行う研究と言えなくもないが、直接的な言語情報である文字以外の視覚情報を言語と結びつけるビジョン&ランゲージの議論で、OCR そのものを含めることは少ない。

### 3.1 キャプション生成

画像キャプション生成 (Image Captioning) は、画像の内容を示すテキストを生成するタスクであり、他の自然言語処理の研究と同様に英語を対象としたものが多い。ビジョン&ランゲージ分野の中でも歴史の長い課題であり、深層学習の流行を迎える前の 2010 年頃から取り組みが徐々に増えてきたテーマである。そのバリエーションについても枚挙に暇がなく、画像列や動画からのキャプション生成の研究も存在する。

最初にこの問題に取り組んだ論文<sup>6)</sup> では、条件付き確率 (Conditional Random Field; CRF) を用いて、画像の 3 種類のラベル (トリプレット): 「object」、「action」、「scene」を推定し、トリプレットが付加されたキャプションの集合から、類似のトリプレットを持つ文を検索している。つまり、システムは新たな文を生成するのではなく、既存の文の内容と一致する文を検索するのであり、画像やキャプションにトリプレットを付加する必要がある。その後、画像認識モジュールと文生成モジュールをそれぞれコンピュータビジョンや自然言語処理の分野から取り込み、画像とキャプションのペアデータのみから新たな文を生成して画像キャプションを生成することを旨とした研究が提案された。

その後、前節で述べた通り CVPR 2015 において深層学習によるキャプション生成が話題になり、複数の論文が同時に同じバイブラインでキャプション生成を提案した。その中でも、Google の Vinyals らによる方法<sup>8)</sup> は、Google の CNN 画像認識モデルである Inception と、LSTM 機械翻訳モデルを組み合わせたもので、仕組みが理解しやすい。CNN 画像エンコーダと LSTM 自然言語デコーダの組み合わせという非常にシンプルな構造でありながら、生成されたキャプションは精度、流暢性ともに大きく向上している。ちなみに深層学習の近年の手法ではしばしば用いられる注意 (アテンション) 機構は、機械翻訳<sup>9)</sup>・キャプショ

ン生成<sup>10)</sup> に導入されたのが生い立ちである。生成中のキャプションで次の単語を推定する際に、画像のどの領域に注意を払うべきかを、画像とキャプションのペアのみからデータドリブンで学習できる。

### 3.2 ビジュアル質問応答

ビジュアル質問応答 (Visual Question Answering; VQA) は、与えられた画像に関する質問に答えるタスクである。ビジュアル質問応答では、質問に答えるための情報は基本的に入力画像に含まれている。画像と質問文というマルチモーダルな入力を用いて、回答候補をいずれかの単語に分類するというクロスモーダル理解の研究である。

この研究課題はユーザーインターフェースの分野で最初にアプリケーションとして提案されたもの<sup>11)</sup> で、視覚障がい者が旅行先で何か分からなくて困っているもの撮影し、回答してほしい質問を入力するという使い方を想定していた。そしてこの論文では、アプリに接続されたクラウドソーシングの人々が手動で回答していた。今回紹介するビジュアル質問応答システムは、このプロセスを自動化する試みともいえる。

ビジュアル質問応答は、視覚と言語の分野で最も広く取り組まれている問題のひとつである。その理由は様々だが、一つにはデータセットとその評価方法が早くから確立されていたことが挙げられる。その中でも貢献の大きいベンチマーク論文<sup>12)</sup> が 2015 年に発表されているが、画像を見ないと答えられない質問をクラウドソーシングで収集し、さらに 1 つの質問に対して 10 個の回答をクラウドソーシングで様々な人から収集して新規データセットを構築している。そして、3 人以上の人が一致して回答した場合に質問が正しいと評価するプロトコルを提案し、画像と質問文をエンコーダで特徴量に変換して直列した後にクラス分類するというシンプルなベースラインで評価を行った。このように整えられたデータセットと評価方法、理解しやすいベースラインがその後の大量の新規参加者をもたらした要素であると考えられる。

また、ビジュアル質問応答もキャプション生成と同様に注意機構と親和性の高い課題である。Shih らによる手法<sup>13)</sup> は注意機構をビジュアル質問応答に適用した先駆的な研究で、階層的な質問・画像の共起を活用している。Anderson らが提案した注意メカニズム<sup>14)</sup> は、画像上で規則的に配置したグリッドに対する注意と、物体らしき領域を別の手法で複数生成したものへの注意を組み合わせたモデルである。

### 3.3 マルチモーダル機械翻訳

マルチモーダル機械翻訳とは、与えられた画像とそのキャプションを用いて、キャプションを別の言語に翻訳する作業である。概念的には、画像を使って入力文の曖昧な

部分を解消し、翻訳の精度を向上させることが期待できる。例えば、Seal という単語は、ものを貼る「シール」もあれば、ラッコにちかい海洋生物もある。実際にものを貼る「シール」が写っている画像があれば、キャプションに Seal という単語が入っていても、どちらの Seal を指しているのかはすぐにわかる<sup>15)</sup>。これは、画像と自然言語が入力されると、自然言語が出力されるという形のクロスモーダル理解の研究になる。

Hitschler らの手法<sup>16)</sup> は、まず、入力された画像とキャプションのペアから、キャプションのみに対して通常の機械翻訳を行い、翻訳候補文を複数生成している。次に、ターゲット言語の画像・キャプションペアデータセットを検索して、入力画像と翻訳候補文からなるペアを複数探し、検索結果に基づいて翻訳候補文のスコアを更新（リランク）することで画像による精度向上を促している。

マルチモーダル機械翻訳は、一部のモダリティでしか実行できない問題の精度を、モダリティの数を増やすことで向上させようとする典型的な例の一つである。一方で実は、このようにモダリティを増やして精度を向上させる問題はクロスモーダル理解のなかでも難しいものの一つになってしまう。この点については最後の節でも簡単に述べる。

### 3.4 テキストからの画像生成

テキストからの画像生成は、画像キャプション生成の逆問題で、画像の内容を示すキャプションから同じ内容を示す画像を生成するタスクである。テキストから画像へ変換するクロスモーダル理解の研究と解釈できる課題で、前節でも触れた GAN や VAE などの深層学習生成モデルが普及した結果、この複雑なタスクへの挑戦が増えている。

Mansimov<sup>17)</sup> は、入力されたキャプションを双方向 LSTM でエンコードした後、RNN ベースの画像デコーダで画像をより鮮明にするために複数回更新するというパイプラインを提案している。Reed<sup>18)</sup> は、この問題に初めて GAN を導入しており、Stack GAN<sup>19)</sup> はこのような GAN を積み上げて画像の高精細化を達成している。Stack GAN では Reed らとほぼ同じアプローチで生成した 64 ピクセル四方の画像を、入力キャプションとともに別の GAN に再度入力し、256 ピクセル四方の画像を出力させている。このように GAN を繋げて高解像度な画像を生成するアプローチが続いていくことになる。

この問題のブレークスルーは Transformer によるアプローチで、OpenAI が 2021 年 1 月 5 日に、キャプションからこれまでよりもはるかに多様な自然な画像／イラストを生成できる手法である「DALL-E」についてのブログ記事<sup>20)</sup> を公開した。キャプションから極めて多様かつ自然な画像やイラストを生成できるという例を示し、限界に衝撃を与えた。またこの記事では、後述する表現学習手法の

CLIP<sup>21)</sup> を活用しているということが述べられており、CLIP と各種画像生成手法を組み合わせた DALL-E の再現が様々なところで試みられている。

### 3.5 視覚的対話

視覚的対話は、これまで言語だけで行われていた対話の研究に、画像や映像を加えたものである。まさに Visual Dialog とするタイトルを冠した 2017 年の論文<sup>22)</sup> では、画像に関連した対話を自然言語で行うデータを収集して対話の学習を行っている。自然言語での対話履歴と関連する視覚情報から、次の対話行為として自然言語などを生成するという意味で、連続的なクロスモーダル理解の研究と解釈できる。

対話生成は人工知能の分野で古くから取り組まれてきた研究テーマで、視覚的対話でもこのような人工知能タスクとしての対話を扱う研究が多く存在する。代表的な設定としては、2 人のペアに画像を使った会話をしてもらい、その自然言語データを収集している。例えば片方が画像を見ていてもう片方が画像を見られない状態の中で、Q&A 形式の対話を通じて画像の内容を口頭で伝えるという対話<sup>22)</sup> や、2 人とも画像を見ている中で、一方が特に見ている領域をもう片方が当てるゲームを Q&A 形式の自然言語対話で進めるもの<sup>23)</sup> などがある。店舗での店員と客<sup>24)</sup> や、市街地での電話ナビゲーターと旅行者<sup>25)</sup> などより社会での活用シーンに近い設定での対話を扱う研究も存在する。

また、広い意味では、エージェントやロボットによるクロスモーダル理解の研究もある。VLN (Vision and Language Navigation) というタスク<sup>26)</sup> はこの中でも有名で、仮想・現実環境にいるエージェントやロボットが、与えられた言語による指示と現在の視覚情報を頼りに目的地に到達するタスクである。その時々々の視覚情報と最初の言語指示に基づいて、回転や移動などの次の動作を対話行為として出力するという意味では、これもまた視覚的対話に関連する研究と言える。

### 3.6 表現学習

表現学習とは、画像（または動画）とそれに対応する自然言語から、ここまで述べてきた諸タスクのための特徴空間を学習することで、様々なタスクの精度向上と学習データ量の削減を目指すものである。画像と自然言語のペアデータをそれぞれのモダリティのためのエンコーダに入力し、何らかの自己教師あり学習を経て特徴量の初期学習を行うものが典型的なパイプラインである。

VILBERT (Vision-and-Language BERT) はこのような研究の草分け的存在<sup>27)</sup> で、BERT<sup>28)</sup> のマルチモーダル版である。動画の時々々の画像フレームの特徴表現の集合と、キャプションの各単語の分散表現のベクトルを Transformer に入力して、エンコーダを学習させる。また、テキ

ストから画像を生成する手法である DALL-E とともに紹介した CLIP<sup>21)</sup> は、現時点で大きな注目を集めている手法の一つである。Web から 4 億対の画像・キャプションペアを収集し、対になっている画像・キャプションペアが最も似た値の特徴量にエンコードされるように、画像・キャプションそれぞれのエンコーダを学習する。

#### 4. 自然言語と音声を繋ぐ研究

音と自然言語を扱ったクロスモーダル理解としては、音声認識や音声合成がまさに合致するタスクである。音声認識 (Acoustic Speech Recognition; ASR) は特に OCR と類似している。OCR が画像メディアの中に埋め込まれている自然言語を抽出するものであるのに対して、ASR は音声メディアの中に埋め込まれている自然言語を抽出するもので、最近出てきたクロスモーダル理解の研究と言うには過去の歴史が長い。深層学習によるアプローチが導入される直前までは手動で設計された時系列データ処理によるローカルな特徴量設計と、混合ガウス分布や隠れマルコフモデルによるモデル化を行い、音素と呼ばれる文字のような単位をまず認識していた。ここが深層学習、具体的には MLP によって代替されたというのが最初 2011 年に起きた事件であり、その後 CNN や RNN、近年では Transformer という形で他のモダリティデータと同様にネットワーク構造の変遷が起きている。音声合成も深層学習によって研究が大きく進んだタスクで、画像生成に似たアプローチとして自己回帰モデルを利用したもの<sup>29)</sup> や GAN による敵対的学習を取り入れたもの<sup>30)</sup> などがある。

もう一つクロスモーダル理解の研究として挙げられるのが音声翻訳である。与えられた音声を自然言語として理解するのみならず、別の言語に翻訳するタスクで、旅行先での翻訳など応用への期待も高い課題である。単純に考えるとまず音声認識を経て同じ言語の音声から自然言語への変換を済ませ、その次に自然言語データ上で別の言語に更に変換するという 2 段階の変換アプローチが思いつく。実際そのような手法を提案している論文<sup>31)</sup> が歴史的に大半である中で、ある言語の音声から直接別の言語の自然言語を出力するようなネットワークを学習するアプローチ<sup>32)</sup> も出現している。

#### 5. 画像と音声を繋ぐ研究

特に音声コミュニケーションは人々の間のやり取りにおいて重要な役割を果たしており、その視覚との関係性を分析するために心理学や情報学など様々なアプローチからの研究が行われている。人間は視覚と聴覚を組み合わせることで事象を認識しているという心理学的知見があり、有名な例としてはマガーク効果がある。例えば「ババ」という音と「ガガ」と発音している口の動きの人の動画を組み合わせると、被験者には「ダダ」に聞こえるという現象である。

音だけを聞いていれば正しく認識できても、音と矛盾する視覚情報が加わると、その矛盾を解消するために認識が変わってしまうのである。

##### 5.1 動画像認識

動画像認識とは、動画に含まれる動作、物体、シーンなどの内容に基づいて、動画の特定の時間間隔や領域を何らかのクラスに分類するタスクである。動画像というと時々刻々のフレームで構成される時系列データだけを指し、音を含まない場合もある。特に人が映った動画での動作認識などは、画像データだけでもうまく識別出来てしまう。逆に画像と音声とが相補的に精度向上に貢献する課題としては、この節の冒頭のように音声認識を音声だけでなく喋っている顔の動画も組み合わせるものがある<sup>33)</sup>。

##### 5.2 視覚補助つき音源定位

視覚補助つき音源定位 (Visually Guided Sound Source Localization) とは、主に動画像中の音源の位置を特定する作業である<sup>34)</sup>。もともと伝統的に音響データで行う音源定位の研究が続けられており、複数の場所から観測された音の時系列データを用いて音源の位置を推定することが目的となっている。視覚的音源定位の出力は、映像と音が時系列的に同期して記録されている中で、映像の各フレームの中でそれぞれの音が発生している領域を切り出すという意味で、画像上のアテンションマップやセグメンテーションに似ていると言える。

##### 5.3 視覚補助つき音源分離

視覚補助つき音源分離 (Visually Guided Sound Source Separation) とは、主に映像を対象として音の複雑な組み合わせをそれぞれ音の要素に分解する作業である<sup>35)</sup>。例えば、複数の楽器が同時に演奏されている音のデータから、各楽器の音のデータを復元するのが目的となる。これを音のデータだけで行うものは音源分離と呼ばれ、統計的なアプローチや機械学習によるアプローチなど、長い間の発展の歴史をもつ。視覚補助つき音源分離は、視覚的な情報である映像を参照しながら音声データの音源分離を行う課題と捉えることができる。

##### 5.4 音と映像の相互変換

音と映像の相互変換とは、音から映像あるいは映像から音を推定するクロスモーダル変換の研究である。人間の音声を対象としたもの<sup>36)</sup> と、楽器の演奏を対象としたもの<sup>37)</sup> に大別される。

#### 6. むすび〜より多様なモダリティを繋ぐ研究へ

ここまで見てきた 2 つのモダリティを組み合わせる研究以外にも、多様なモダリティの組み合わせの研究が存在す

る。まずはここまで扱ってきたモダリティを全て組み合わせる課題である。例えばテキストによる動画検索で、画像の時系列情報以外に音声も活用してより精緻に動画を検索しようとする研究<sup>38)</sup>や、動画の音声などを活用しながらその内容を記述する研究<sup>39)</sup>が挙げられる。

また一言で表現すると例えば「視覚」などの単一のモダリティになってしまうものを、複数の形態の画像など取えてマルチモーダル表現されたデータとして同時に扱うことでクロスモーダル理解を実現する研究もある。たとえば RGB 画像データと赤外光を組み合わせた物体領域分割<sup>40)</sup>や、3次元の点群データと(2次元の)RGB画像データを組み合わせた物体検出<sup>41)</sup>などがある。

ただ注意したいのが、単純にモダリティを増やせば増やすほどよいことばかりではない点である。一つ目の問題は、データの収集の困難さである。当然ではあるが、扱うモダリティが揃ったデータを収集しようとする、モダリティの種類が増えるほどそのコストも増える。画像キャプション生成であれば画像とキャプションのペアデータを収集すればよいが、テキストによる動画検索で音声も扱おうとすると、テキスト・動画・音声の3つのモダリティが同時に揃ったトリプレットデータが必要になる。

もう一つの問題が、モダリティを組み合わせたときに単一のモダリティよりもタスクの精度が上がる場合だけとは限らない、と言う点である。例えば動画認識であれば画像列データだけの方がやや高精度だったり、マルチモーダル機械翻訳であれば自然言語データだけの方が高精度だったりすることが挙げられる。単一モダリティでもある程度の精度が出てしまうようなタスクの入力をマルチモーダルにした場合には、この問題が付きまとう。最近はこの対応するために勾配ブレンディングと言う手法が提案されている<sup>42)</sup>。簡単に言えば、収束が進んだモダリティの損失関数はどんどん重みを下げて、まだ収束していないモダリティの損失関数を下げる様な働きをする手法である。種々のタスクでの効果が確かめられており、上記の点群とRGB画像から物体を検出する研究<sup>41)</sup>でも採用されている。クロスモーダル理解に対する機械学習の解析的な議論を含めた研究はまだまだ開拓の余地があるように思われるので、発見的なアプローチだけにとらわれず、数式を操作しながらクロスモーダル理解を解き明かす研究が進むことにも期待している。

本稿ではクロスモーダルな理解の研究として、自然言語処理/画像/音声を繋ぐ研究課題について概観した。エンコーダ・デコーダと言う基本的な概念について触れたあと、上記の3つのモダリティのうち2つを繋ぐ研究についてそれぞれ解説した。画像と自然言語を繋ぐ Vision & Language と呼ばれる分野の研究が一番多く、この分野のサーベイは日本語のものも散見されるようになってきた(英語は言わずもがなである)。一方で音声や他のモダリ

ティまで含めたサーベイはあまり見かけない。本稿がより多くの研究者にとって、クロスモーダルな理解に興味をもつきっかけとなれば幸甚である。

## 参考文献

- 1) Seide, F. et al. "Conversational Speech Transcription Using Context-Dependent Deep Neural Networks". Conference of the International Speech Communication Association. 2011, p.437-440.
- 2) Krizhevsky, A. et al. "ImageNet Classification with Deep Convolutional Neural Networks". Advances in Neural Information Processing Systems. 2012, p.1097-1105.
- 3) Sutskever, I. et al. "Sequence to Sequence Learning with Neural Networks". Advances in Neural Information Processing Systems. 2014, p.3104-3112.
- 4) Inoue, N.; Shinoda, K. "A Fast MAP Adaptation Technique for GMM-supervector-based Video Semantic Indexing Systems". ACM Multimedia. 2011, p.1357-1360.
- 5) Ushiku, Y. et al. "Efficient Image Annotation for Automatic Sentence Generation". ACM Multimedia. 2012, p.549-558.
- 6) Farhadi, A. et al. "Every Picture Tells a Story: Generating Sentences from Images". European Conference on Computer Vision. 2010, p.15-29.
- 7) LeCun, Y. et al. Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE. 1998, Vol.86, No.11, p.2278-2324.
- 8) Vinyals, O. et al. "Show and Tell: A Neural Image Caption Generator". IEEE Conference on Computer Vision and Pattern Recognition. 2015, p.3156-3164.
- 9) Bahdanau, D. et al. "Neural Machine Translation by Jointly Learning to Align and Translate". International Conference on Learning Representations. 2015.
- 10) Xu, K. et al. "Show, Attend and Tell: Neural Image Caption Generation with Visual Attention". International Conference on Machine Learning. 2015, p.2048-2057.
- 11) Bigham, J. P. et al. "VizWiz: Nearly Real-time Answers to Visual Questions". ACM Symposium on User Interface Software and Technology. 2010, p.333-342.
- 12) Antol, S. et al. "VQA: Visual Question Answering". IEEE International Conference on Computer Vision. 2015, p.2425-2433.
- 13) Shih, K. J. et al. "Where To Look: Focus Regions for Visual Question Answering". IEEE Conference on Computer Vision and Pattern Recognition. 2016, p.4613-4621.
- 14) Anderson, P. et al. "Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering". IEEE Conference on Computer Vision and Pattern Recognition. 2018, p.6077-6086.
- 15) Calixto, I. et al. "Images as Context in Statistical Machine Translation". Workshop on Vision and Language. 2012.
- 16) Hitschler, J. et al. "Multimodal Pivots for Image Caption Translation". Meeting of the Association for Computational Linguistics. 2016, p.2399-2409.
- 17) Mansimov, E. et al. "Generating Images from Captions with Attention". International Conference on Learning Representations.

- 2016.
- 18) Reed, S. et al. "Generative Adversarial Text to Image Synthesis". International Conference on Machine Learning. 2016, p.1060-1069.
  - 19) Zhang, H. et al. "StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks". IEEE International Conference on Computer Vision. 2017, p.5907-5915.
  - 20) Ramesh, A. et al. "DALL·E: Creating Images from Text". <https://openai.com/blog/dall-e/>, (参照 2021-09-01).
  - 21) Radford, A. et al. "Learning Transferable Visual Models From Natural Language Supervision". International Conference on Machine Learning. 2021, p.8748-8763.
  - 22) Das, A. et al. "Visual Dialog". IEEE Conference on Computer Vision and Pattern Recognition. 2017, p.326-335.
  - 23) de Vries, H. et al. "GuessWhat?! Visual Object Discovery Through Multi-Modal Dialogue". IEEE Conference on Computer Vision and Pattern Recognition. 2017, p.5503-5512.
  - 24) Saha, A. et al. "Towards Building Large Scale Multimodal Domain-Aware Conversation Systems". AAAI Conference on Artificial Intelligence. 2018, p.696-704.
  - 25) de Vries, H. et al. Talk the Walk: Navigating New York City through Grounded Dialogue. arXiv. 2018, 1807.03367v3.
  - 26) Anderson, P. et al. "Vision-and-Language Navigation: Interpreting Visually-Grounded Navigation Instructions in Real Environments". IEEE Conference on Computer Vision and Pattern Recognition. 2018, p.3674-3683.
  - 27) Lu, J. et al. "ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks". Advances in Neural Information Processing Systems. 2019, p.13-23.
  - 28) Devlin, J. et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". Conference of the North American Chapter of the Association for Computational Linguistics. 2019, p.4171-4186.
  - 29) van den Oord, A. et al. WaveNet: A Generative Model for Raw Audio. arXiv. 2016, 1609.03499v2.
  - 30) Kumar, K. et al. "MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis". Advances in Neural Information Processing Systems. 2019, p.14910-14921.
  - 31) Vidal, E. "Finite-state speech-to-speech translation". IEEE International Conference on Acoustics, Speech, and Signal Processing. 1997, p.111-114.
  - 32) Jia, Y. et al. "Direct Speech-to-Speech Translation with a Sequence-to-Sequence Model". Conference of the International Speech Communication Association. 2019, p.1123-1127.
  - 33) Ngiam, J. et al. "Multimodal Deep Learning". International Conference on Machine Learning. 2011, p.689-696.
  - 34) Senocak, A. et al. "Learning to Localize Sound Source in Visual Scenes". IEEE Conference on Computer Vision and Pattern Recognition. 2018, p.4358-4366.
  - 35) Hershey, J. et al. "Deep clustering: Discriminative embeddings for segmentation and separation". IEEE International Conference on Acoustics, Speech and Signal Processing. 2016, p.31-35.
  - 36) Prajwal, K. R. et al. "Learning Individual Speaking Styles for Accurate Lip to Speech Synthesis". IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020, p.13796-13805.
  - 37) Owens, A. et al. "Visually Indicated Sounds". IEEE Conference on Computer Vision and Pattern Recognition. 2016, p.2405-2413.
  - 38) Gabeur, V. et al. "Multi-modal Transformer for Video Retrieval". European Conference on Computer Vision. 2020.
  - 39) Hori, C. et al. "Attention-Based Multimodal Fusion for Video Description". IEEE International Conference on Computer Vision. 2017, p.4193-4202.
  - 40) Ha, Q. et al. "MFNet: Towards real-time semantic segmentation for autonomous vehicles with multi-spectral scenes". IEEE/RSJ International Conference on Intelligent Robots and Systems. 2017, p.5108-5115.
  - 41) Qi, C. R. et al. "ImVoteNet: Boosting 3D Object Detection in Point Clouds with Image Votes". IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020, p.4404-4413.
  - 42) Wang, W. et al. "What Makes Training Multi-modal Classification Networks Hard?". IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020, p.12695-12705.

## 執筆者紹介



牛久 祥孝 USHIKU Yoshitaka

オムロン サイニクエックス株式会社

リサーチアドミニストレイティブディビジョン

専門: コンピュータビジョン、自然言語処理、  
パターン認識、機械学習

所属学会: ACM、IEEE、電子情報通信学会、  
情報処理学会、日本ロボット学会、人工知能学  
会、応用物理学会、建築情報学会

博士 (情報理工学)

本文に掲載の商品の名称は、各社が商標としている場合があります。



### 2021

- Ryo Yonetani, Tatsunori Taniai, Mohammadamin Barekattain, Mai Nishimura, and Asako Kanezaki, “Path Planning using Neural A\* Search”, International Conference on Machine Learning (ICML), PMLR 139, pp.12029–12039, 2021.
- Qing Yu, Atsushi Hashimoto, and Yoshitaka Ushiku, “Divergence Optimization for Noisy Universal Domain Adaptation”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2515–2524, 2021.
- Masashi Hamaya, Kazutoshi Tanaka, Yoshiya Shibata, Felix von Drigalski, Chisato Nakashima, and Yoshihisa Ijiri, “Robotic Learning From Advisory and Adversarial Interactions Using a Soft Wrist”, IEEE Robotics and Automation Letters (RA-L) with RoboSoft option, vol.6, no.2, pp.3878–3885, 2021.
- Felix von Drigalski, Devwrat Omkar Joshi, Takayuki Murooka, Kazutoshi Tanaka, Masashi Hamaya, and Yoshihisa Ijiri, “An analytical diabolito model for robotic learning and control”, IEEE International Conference on Robotics and Automation (ICRA), 2021.
- Felix von Drigalski, Kennosuke Hayashi, Yifei Huang, Ryo Yonetani, Masashi Hamaya, Kazutoshi Tanaka, and Yoshihisa Ijiri, “Precise Multi-Modal In-Hand Pose Estimation using Low-Precision Sensors for Robotic Assembly”, IEEE International Conference on Robotics and Automation (ICRA), 2021.
- Kazutoshi Tanaka, Ryo Yonetani, Masashi Hamaya, Robert Lee, Felix von Drigalski, and Yoshihisa Ijiri, “TRANS-AM: Transfer Learning by Aggregating Dynamics Models for Soft Robotic Assembly”, IEEE International Conference on Robotics and Automation (ICRA), 2021.
- Ukyo Honda, Yoshitaka Ushiku, Atsushi Hashimoto, Taro Watanabe, and Yuji Matsumoto, “Removing Word-Level Spurious Alignment between Images and Pseudo-Captions in Unsupervised Image Captioning”, Conference of the European Chapter of the Association for Computational Linguistics (EACL), pp.3692–3702, 2021.
- Kazutoshi Tanaka, Shih-Hsin Yang, Yuji Tokudome, Yuna Minami, Yuyao Lu, Takayuki Arie, Seiji Akita, Kuniharu Takei, and Kohei Nakajima, “Flapping-Wing Dynamics as a Natural Detector of Wind Direction”, Advanced Intelligent Systems, Vol.3, No.2, 2021.
- Yuta Kamikawa, Atsushi Hashimoto, Motoharu Sonogashira, and Masaaki Iiyama, “Curiosity Guided Fine-tuning for Encoder-Decoder-based Visual Forecasting”, IEICE Transactions on Information and Systems, Vol.E104-D, Issue 5, pp.752–761, 2021.
- 原島純, 橋本敦史. “キッチン・インフォマティクス—料理を支える自然言語処理と画像処理”. オーム社. P264, 2021.
- 米谷竜. “連合学習入門—基本的なアプローチと典型的な課題—”. 精密工学会誌, 87巻, 8号, pp.662–665, 2021.
- 濱屋政志, 田中一敏, Felix von Drigalski. “工場現場の組立応用に向けたソフトロボット運動学習”. 日本ロボット学会誌, 39巻, 7号, pp.609–612, 2021.

### 2020

- Felix von Drigalski, Christian Schlette, Martin Rudorfer, Nikolaus Correll, Joshua Triyonoputro, Weiwei Wan, Tokuo Tsuji, and Tetsuyou Watanabe, “Robots assembling machines: Learning from the World Robot Summit 2018 Assembly Challenge”, Advanced Robotics, vol.34, issue 7–8, pp.408–421, 2020.
- Nie Kaidi, Felix von Drigalski, et al., “Team O2AS Approach for Task-board Task of WRC 2018”, Advanced Robotics, vol.34, pp.477–498, 2020.
- Hiroaki Minoura, Ryo Yonetani, Mai Nishimura, and Yoshitaka Ushiku, “Crowd Density Forecasting by Modeling Patch-based Dynamics”, IEEE Robotics and Automation Letters (RA-L), vol.6, no.2, pp.287–294, 2020.

- Taichi Nishimura, Atsushi Hashimoto, Yoshitaka Ushiku, Hiroataka Kameko, Yoko Yamakata, and Shinsuke Mori, “Structure-Aware Procedural Text Generation from an Image Sequence”, *IEEE Access*, vol.9, pp.2125–2141, 2020.
- Taichi Nishimura, Suzushi Tomori, Hayato Hashimoto, Atsushi Hashimoto, Yoko Yamakata, Jun Harashima, Yoshitaka Ushiku, and Shinsuke Mori, “Visual Grounding Annotation of Recipe Flow Graph”, *Language Resources and Evaluation Conference (LREC)*, pp.4275–4284, 2020.
- Kazutoshi Tanaka, Yuji Tokudome, Yuna Minami, Satoko Honda, Kuniharu Takei, and Kohei Nakajima, “Self-Organized Soft Reservoirs”, *ICRA2020 Workshop*.
- Ryuhei Takahashi, Atsushi Hashimoto, Motoharu Sonogashira, and Masaaki Iiyama, “Partially-shared Variational Auto-encoders for Unsupervised Domain Adaptation with Target Shift”, *European Conference on Computer Vision (ECCV)*, p.1–17, 2020.
- Takayuki Murooka, Masashi Hamaya, Felix von Drigalski, Kazutoshi Tanaka, and Yoshihisa Ijiri, “EXI-Net: EXplicitly/Implicitly Conditioned Network for Multiple Environment Sim-to-Real Transfer”, *Conference on Robot Learning (CoRL)*, pp.1–10, 2020.
- Mai Nishimura and Ryo Yonetani, “L2B: Learning to Balance the Safety-Efficiency Trade-Off in Interactive Crowd-Aware Robot Navigation”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp.11004–11010, 2020.
- Felix von Drigalski, Kazutoshi Tanaka, Masashi Hamaya, Robert Lee, Chisato Nakashima, Yoshiya Shibata, and Yoshihisa Ijiri, “A Compact, Cable-Driven, Activatable Soft Wrist with Six Degrees of Freedom for Assembly Tasks”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp.8752–8757, 2020.
- Masashi Hamaya, Felix von Drigalski, Takamitsu Matsubara, Kazutoshi Tanaka, Robert Lee, Chisato Nakashima, Yoshiya Shibata, and Yoshihisa Ijiri, “Learning soft-robotic assembly strategies from successful and failed demonstrations”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp.8309–8315, 2020.
- Matthew Ishige, Takuya Umedachi, Yoshihisa Ijiri, Tadahiro Taniguchi, and Yoshihiro Kawahara, “Blind Bin Picking of Small Screw through In-finger Manipulation with Compliant Robotic Fingers”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp.9337–9344, 2020.
- Shinya Michibata, Katsufumi Inoue, Michifumi Yoshioka, and Atsushi Hashimoto, “Cooking Activity Recognition in Egocentric Videos with a Hand Mask Image Branch in the Multi-stream CNN”, *Workshop on Multimedia for Cooking and Eating Activities (CEA)*, 2020.
- Mohammadamin Barekatin, Ryo Yonetani, and Masashi Hamaya, “MULTIPOLAR: Multi-Source Policy Aggregation for Transfer Reinforcement Learning between Diverse Environmental Dynamics”, *International Joint Conference on Artificial Intelligence (IJCAI)*, pp.3108–3116, 2020.
- Jiaxin Ma, Ryo Yonetani, and Zahid Iqbal, “Adaptive Distillation for Decentralized Learning from Heterogeneous Clients”, *International Conference on Pattern Recognition (ICPR)*, pp.7486–7492, 2020.
- Takayuki Murooka, Artur Istvan Karoly, Felix von Drigalski, and Yoshihisa Ijiri, “Simultaneous Planning of Grasp and Motion using Sample Regions and Gradient-Based Optimization”, *IEEE International Conference on Automation Science and Engineering (CASE)*, pp.1102–1109, 2020.
- Matthew Ishige, Takuya Umedachi, Yoshihisa Ijiri, and Yoshihiro Kawahara, “In-Hand Small-Object Counting from Tactile Sensor Arrays Installed on Soft Fingertips”, *IEEE International Conference on Soft Robotics (RoboSoft)*, pp.272–277, 2020.
- 西村太一, 橋本敦史, 森信介. “重要語に着目した写真列からのレシピの自動生成”. *自然言語処理*, 27 卷, 2 号, pp.257–279, 2020.

- ・室岡貴之, 濱屋政志, フェリクッス フォンドリガルスキ, 田中一敏, 井尻善久. “明示・暗示ダイナミクス変数を用いた多様な環境間での動作学習”. 日本ロボット学会誌レター, 27 巻, 2 号, pp.257-279, 2020.
- ・千葉直也, 伊東秀謹, 橋本浩一. “LTM スパース推定による三次元計測の白飛び対応とばら積みピッキングロボットへの応用の紹介”. 画像ラボ, 31 巻, 11 号, pp.41-47, 2020.
- ・上瀧剛, 井尻善久. “サブピクセル物体検出とベンチマーク”. 画像ラボ, 31 巻, 8 号, pp.6-13, 2020.

## 2019

- ・Masashi Hamaya, Takamitsu Matsubara, Tatsuya Teramae, Tomoyuki Noda, and Jun Morimoto, “Design of physical user-robot interactions for model identification of soft actuators on exoskeleton robots”, *The International Journal of Robotics Research (IJRR)*, Vol.40, Issue 1, pp.397-410, 2019.
- ・井尻善久, Felix von Drigalski. “産業用ロボットの進化によるものづくりの近未来”. 日本ロボット学会誌, 37 巻, 8 号, pp.675-678, 2019.
- ・古賀達也, 井尻善久. “IOT の側面から見た、モバイルロボット・モバイルマニピュレータの現状と今後の展望”. 日本ロボット学会誌, 37 巻, 8 号, pp.707-710, 2019.
- ・牛久祥孝. “画像に関連した言語生成の取組み”. 人工知能学会誌, Vol.34, No.4, pp.483-491, 2019.
- ・牛久祥孝. “マルチモーダル 画像キャプション生成とその派生タスク”. 画像ラボ, 30 巻, 4 号, pp.28-31, 2019.

## 2018

- ・Yuki Fujimura, Masaaki Iiyama, Atsushi Hashimoto, and Michihiko Minoh, “Photometric Stereo in Participating Media Considering Shape-Dependent Forward Scatter”, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.7445-7453, 2018.

# OMRON TECHNICS

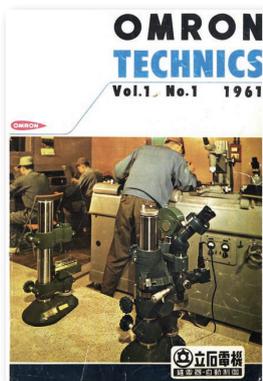
## のご紹介

技術論文誌「OMRON TECHNICS」は研究開発成果を広く公開することで様々な社会課題を解決し、

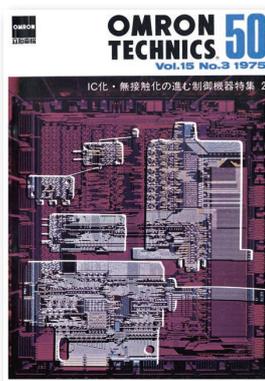
より良い社会の実現に寄与することを目的に発行しております。

オムロングループのコア技術“Sensing & Control + Think”を活用した新たな価値を是非ご覧ください。

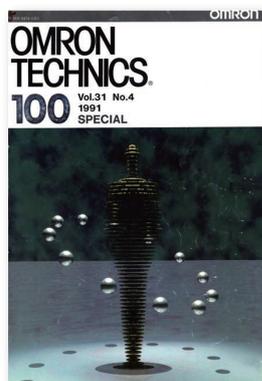
創刊よりこれまで通巻164号発行し、約1,770編以上の論文を公開しております。



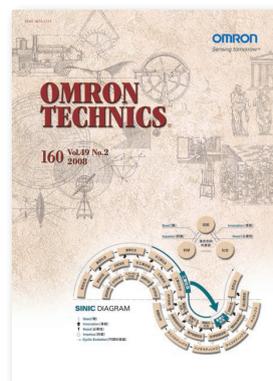
1961年5月 創刊号



1975年 50号



1991年 100号



2008年 160号



2018年 161号



2019年 162号



2020年 163号



2021年 164号

オムロンテクニクス

検索

161号以降の論文は全て WEB 公開しております  
(冊子の販売は行っておりませんのでご了承ください)



日本語



English

OMRON TECHNICS  
推進委員会

オムロン サイニクエクス  
特別推進委員

推進委員長	大平 真琴			
推進委員	岡部 浩史	河合 武宏	坂瀬 忠之	澤野井 幸哉
	高木 徳生	浪江 正樹	馬淵 雅夫	山崎 琢也 (50音順)
事務局	西崎 修	荻野 裕貴	行衛 満智子	
編集協力	オムロン エキスパートリンク株式会社 三宅 慶幸			

諏訪 正樹

## OMRON TECHNICS 非売品

第53巻 第2号  
通巻165号  
(禁無断転載)  
ISSN 0474-1315

2021年11月1日 発行

編集兼発行人 宮田 喜一郎  
発行 オムロン株式会社 技術・知財本部  
〒619-0283 京都府木津川市木津川台9-1  
電話 0774-74-2000 (代)

Published by OMRON Corporation  
9-1, Kizugawadai, Kizugawa-city,  
Kyoto, 619-0283 Japan  
TEL 0774-74-2000

本誌に記載している会社名・製品名などは、それぞれの会社の商標または登録商標です。

©OMRON Corporation 2021 All Rights Reserved.

**OMRON**