



Data-oriented Approaches for Improving Neural Dialogue Generation

♡ 3



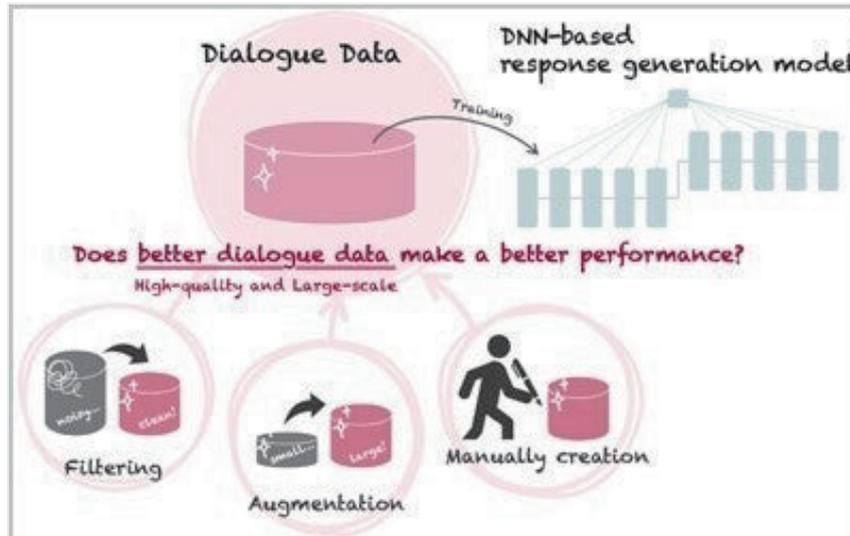
情報処理学会・学会誌「情報処理」
2021年8月15日 08:50



邦訳：ニューラル対話応答生成の性能向上のためのデータ駆動アプローチ

赤間怜奈

(東北大学 助教)



----- keyword -----

自然言語処理

深層学習

対話応答生成

【背景】 深層ニューラルネットワーク技術が高性能な応答生成モデルを実現

【問題】 訓練データ改良を通してモデルのさらなる性能向上を図る

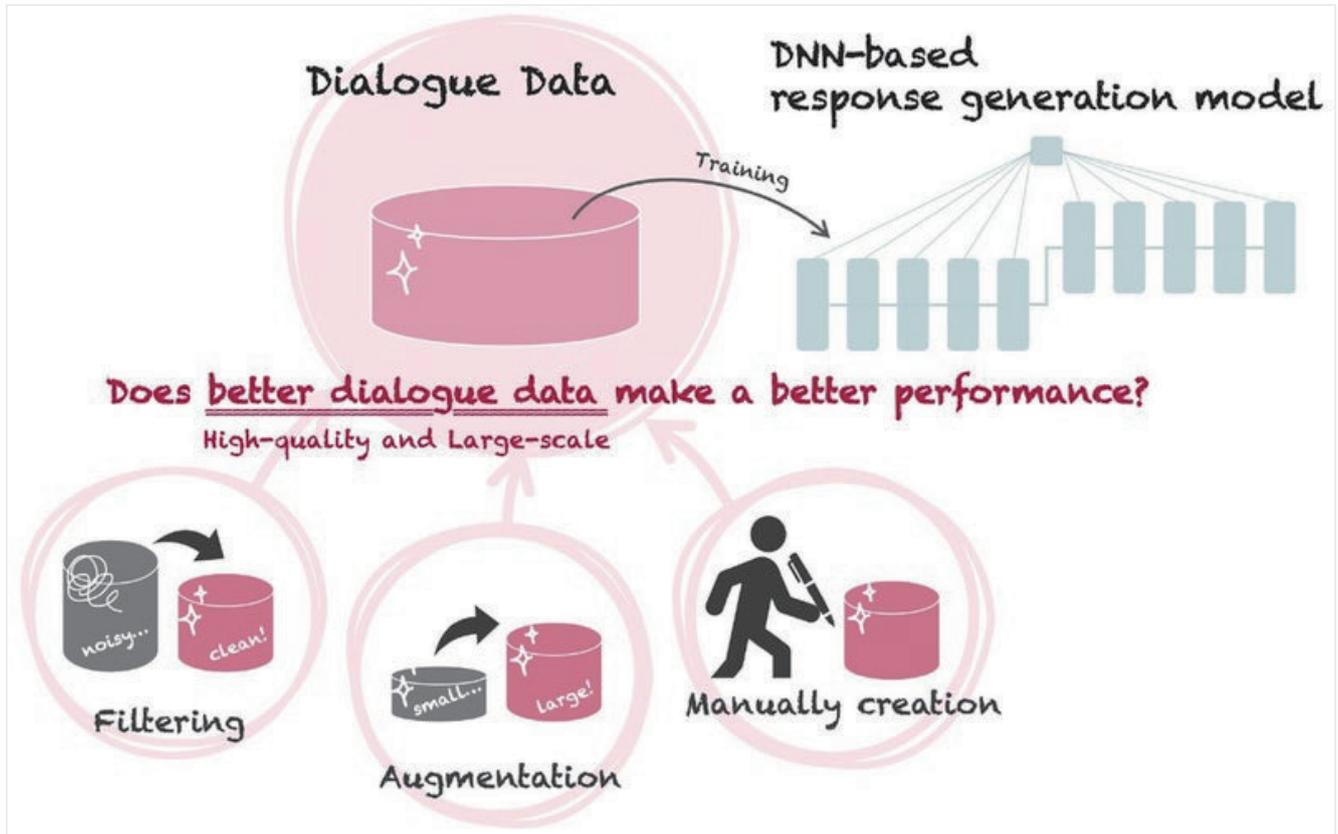
【貢献】 訓練データの品質等を改善する方法論を提案，データ改良の有効性を実証

人間のことばを用いて、人間と同じように対話をする——このようなコンピュータの実現は、古くから考えられてきた人工知能応用の理想のひとつです。「自然言語処理」と呼ばれる私たちの研究分野では、人間が日常的に使っていることば（自然言語）を上手に扱うための計算機構の構築やそれを支える基礎技術の開発に取り組んでいます。人間のように対話するコンピュータ、いわゆる対話システムに関する研究も、当分野の主要な研究領域のひとつです。

自然言語処理における文生成技術は、機械学習とりわけ深層学習の台頭とその進展に伴い、急速な発展を遂げてきました。一般論として、深層ニューラルネットワークを用いた生成モデルがそのポテンシャルを十分に発揮するためには大規模かつ高品質な訓練データが不可欠であるということが、これまでの研究で明らかになっています。対話応答生成技術の研究領域でも、深層ニューラルネットワークを用いた応答生成モデルは、未知の対話に対する汎化能力や生成応答の流暢性を大きく向上させました。近年の研究では、対話応答生成技術のさらなる向上に向けて、より洗練されたネットワークの適応やモデルアーキテクチャの改良などが盛んに行われています。しかしながら、その一方で、訓練データの規模および品質の改善により応答生成技術の向上を図る方向性については、ほとんど議論が進んでいないのが現状です。

本論文では、深層ニューラルネットワークを用いた対話応答生成モデルの性能向上を、モデル構築のための訓練データの改良を通じて実現することに取り組みまし

た。訓練データの高品質化および大規模化に着目してそのための方法論を複数確立し、実験を通して訓練データの改良が応答生成モデルの性能向上に貢献するということを経験的に示しました。本研究で確立した方法論のひとつは、大規模対話データを高品質化するデータフィルタリングのための方法論です。これまでの対話応答生成研究で頻繁に利用されている大規模な対話データには、低品質な、つまり対話として明らかに成立していないような発話と応答の連なり（発話ペア）が含まれていることを明らかにしたうえで、このような低品質な発話ペアを特定し訓練データから自動で除去するために、各発話ペアの「対話としての良さ」をスコアとして算出する方法論を考案しました。提案法によって獲得した大規模かつ高品質な対話データで訓練した応答生成モデルは、訓練データを改良しない場合に比べて、より良い品質の応答を生成しました。対話応答生成モデルの性能向上におけるデータ戦略の有効性を実証した本研究は、これからの深層ニューラルネットワークを用いた対話応答生成技術のさらなる発展に向けて、引き続きモデル自身の改良を進めていくことに加え、いかにして良い訓練データを効率良く獲得するかという問いについても積極的に議論していくことの重要性を示唆しています。



(2021年6月5日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（情報科学）

大学：東北大学

推薦文：（自然言語処理研究会）

本論文は、深層ニューラルネットワーク技術に基づく対話応答生成モデルに対し、訓練データの改良によって性能向上を図るデータ指向的な方法論の確立に取り組んでいる。従来のモデル指向的研究とは相対する立場から、対話応答生成技術の発展に繋がる新たな方向性を提示した点に、当該分野における大きな学術的貢献がある。

赤間怜奈

研究生活：「人工知能」という存在に興味をもったきっかけは、小学生のころに出会った児童書でした。そこで描かれていた（世界一の怪盗の友人、否、仕事上のパートナーである）人工知能は、人間のことばを流暢に扱い、人間のような感情と思考回路を持ち、そして人間よりも遥かに物知りで頭の回転が早い「知的な」存在でした。そんな人工知能に憧れを抱き、そのようなシステムが実現する未来を夢見、自然言語処理の研究に従事しています。私が卒業後も大学で自然言語処理の研究者を続ける道を選んだのは、修士課程・博士課程で研究活動に明け暮れた日々がとにかく愉しかったから、という理由に尽きます。同じ学問に興味を持つ仲間に出会われ、支えられ、思う存分自分の興味関心に向き合う。面白い発見や嬉しい成果が

あれば仲間とわいわい盛り上がる。そんな心地の良い充実した毎日でした。お世話になったすべての方にこの場をお借りしてお礼申し上げます。ありがとうございました。



Grammatical and Semantic Biases in Representation Learning from Raw Datasets

♡ 2

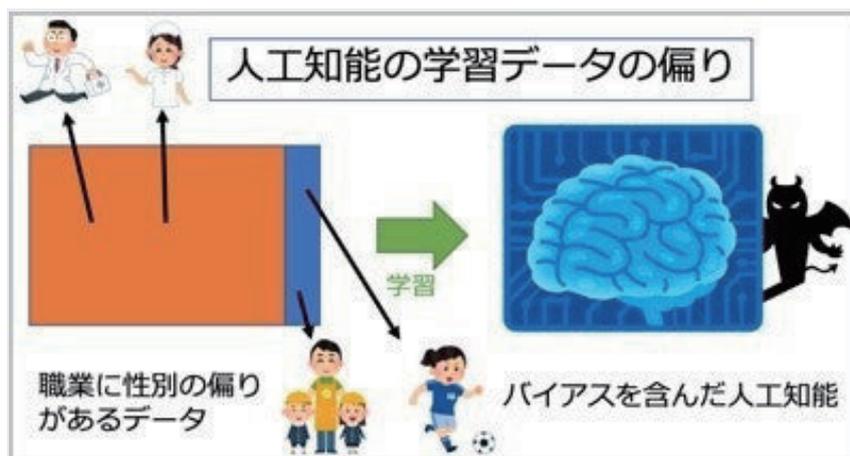
 情報処理学会・学会誌「情報処理」
2021年8月15日 08:51



邦訳：生データを用いた表現学習における文法・意味バイアス

金子正弘

(東京工業大学 研究員)



----- keyword -----

自然言語処理

分散表現

データの偏り

【背景】 大規模なデータで分散表現を学習

【問題】 データの偏りが分散表現に悪影響を与える

【貢献】 学習データの偏りの影響を低減する手法の提案

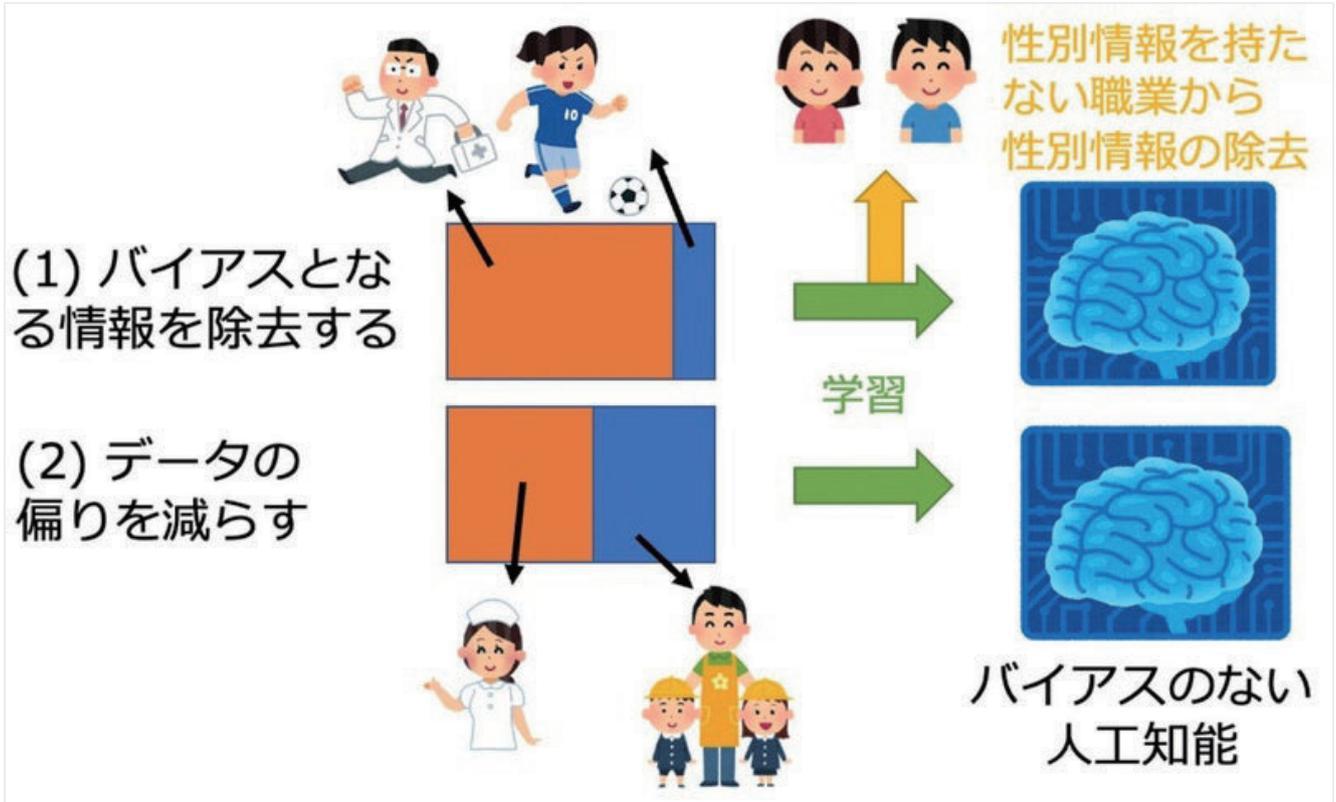
自然言語処理はテキストを理解する人工知能を作る研究分野である。テキストを理解可能にするには、まずテキストの情報をコンピュータが読み込めるように数値に置き換える必要がある。テキストの文法や意味に関する情報を表現した数値を分散表現と呼ぶ。単語間の関係や文の意味を正しく捉えた質の高い分散表現を獲得することはテキストを扱うシステムの性能改善につながる。近年の自然言語処理ではWikipediaやWebニュースデータなどの大規模なデータを使い、単語や文の共起から分散表現を学習することが主流となっている。

一方で、大規模データでは文法や意味に関して偏った分布になっていることが知られている。このようなデータを用いると少数の素性より多数の素性を優先し、バイアスを含んだ分散表現が学習されてしまう。文法に関する文法バイアスであればデータセットに含まれるテキストの大部分は「私は医者である」のように文法的に正しく、「私に医者である」のように文法的に誤ったテキストはほとんど含まれていない。このように文法的に偏ったデータで学習された分散表現は文法誤りを考慮できていない。そのため、文法誤り検出や文法誤り訂正のように文法が誤ったテキストを扱うタスクに最適な情報を学習できていない。意味に関する意味バイアスでは、単語同士の共起頻度の違いにより意味に関する偏った情報を分散表現が学習してしまう。たとえば、データセットに「彼女は看護師である」のように「彼女」と「看護師」が共起した文は「彼女は医者である」のように「彼女」と「医者」が共

起した文より高頻度に含まれている。一方で、「彼」と「医者」が共起した文は「彼」と「看護師」が共起した文より高頻度である。このようなデータを使った場合、分散表現は共起の偏りから「看護師」は女性であり「医者」は男性であるというバイアスを学習してしまう。

そこで、本研究ではこれらのバイアスの影響を低減するために、(1) 学習されたバイアスの除去する手法と(2) データセットの偏りをなくす手法を提案した。

(1) のバイアスの除去では、学習されたバイアスを分散表現から取り除く。たとえば、意味バイアスの性別に関するバイアスであれば「医者」や「看護師」のように性別情報を保持する必要がない分散表現から性別情報を取り除く。(2) の偏りをなくす手法では、少数のデータを増やすまたは重点的に学習することで分散表現がそれらのデータを考慮して学習できるようにする。たとえば、文法バイアスであればテキストデータに対して、擬似的な誤りを発生されることで文法的に正しいテキストと誤ったテキストの両方を考慮できるようにする。図は意味バイアスにおける2つの手法を示している。これらの手法はテキストを扱う人工知能のバイアスを低減することができる。



■研究ブログ

<https://masahiro-kaneko.com/>

(2021年5月31日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（情報科学）

大学：東京都立大学

推薦文：（自然言語処理研究会）

本研究は自然言語処理において文や単語をどのように表現すればいいか、という問題に対して、人手でメタ情報を付与していない生のデータから表現を学習する際に生じる文法や意味のバイアス（たとえば性別バイアス）についての研究を行ない、さまざまな深層学習の手法を駆使して、人工知能分野のホットな話題に切り込んでいます。



金子正弘

研究生活：自分は研究の問題を解決するためにアイデアを出すところが好きなた

め、特に研究テーマに対してこだわりはなく、そのときに面白そうだと思ったことに取り組んできました。そのため、最終的に異なるテーマの研究に取り組むことになり、博士論文としてまとめるのに苦労しました。そして、このような苦労以外にも博士課程では進捗がでないことや論文に採択されないなどのつらい部分があります。博士課程を楽しみながらサバイブするコツは研究に対して気負いすぎないことと自分の心がときめくことを優先することだと思います。



A Conversational System for Interactive Image Editing

♡ 2

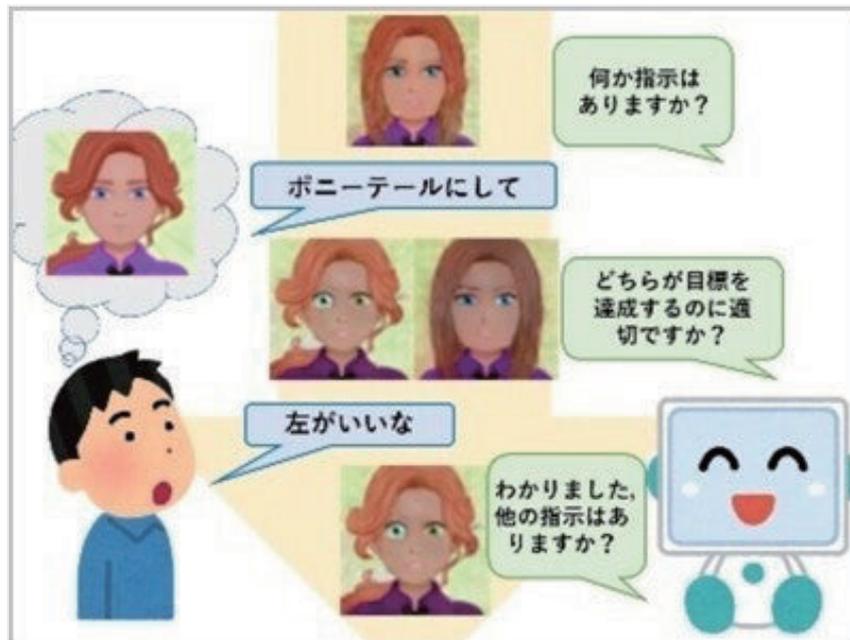
 情報処理学会・学会誌「情報処理」
2021年8月15日 08:53



邦訳：自然言語を用いた対話型画像編集システム

品川政太郎

(奈良先端科学技術大学院大学 助教)



----- keyword -----

画像編集

自然言語インターフェース

深層学習

【背景】 広告やイラストの作成は高コストなので自動で生成したい

【問題】 言語情報を使うのは手軽だが多様な表現が許されるので難しい

【貢献】 言語情報で段階的に画像を編集できる深層画像編集モデル

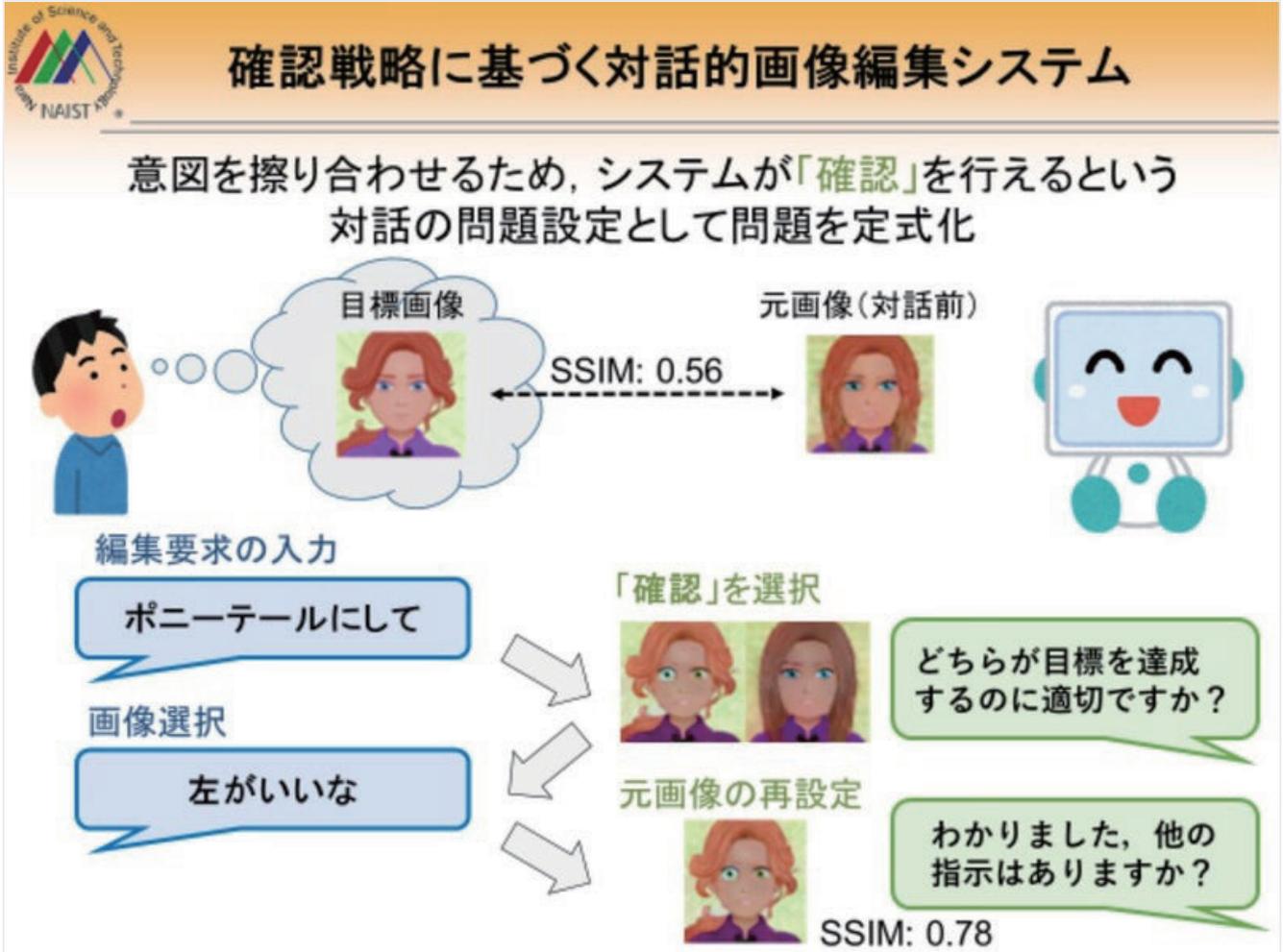
本研究は、Vision and Language (V&L) と呼ばれる、視覚的情報と言語情報（ことば）を融合させて種々の問題解決を行う研究の1つです。技術の発展に伴い、私たちの生活は、どんどん便利になる一方で日々大量の情報を処理する必要性に迫られています。V&Lの研究は、人の代わりにその場で何が起きているかを見て必要な情報だけ説明したり、人が「かわいい犬のイラストが欲しい」と言ったときに人の代わりに画像を作ったりと、日常的に使っている言語情報と視覚的情報を組合わせてさまざまな問題を解決する可能性を秘めており、近年、盛んに研究が進められています。

本研究が題材としているのは、犯罪者の目撃証言から似顔絵を作成するような、言語情報によって画像を編集する課題です。私たちは手に入れた画像があった時に、画像検索で探すことができます。しかし、画像検索で得られる画像はすでに存在する画像だけであり、「ちょっと思ったのと違う」画像しか得られないことがあります。かといって、自分で描くのは大変ですし、描く技術が必要です。そこで本研究では、私たちが日常的に使っている言語情報を用いて画像を編集することで、目標に近い画像を得られるようなシステムを作れないかと研究を進めてきました。

言語情報を入力として扱う難しさは、さまざまな言い方ができるという言語情報の多様性をうまく扱う必要がある点です。本研究は、この多様性の問題に対して2

つの方法でアプローチしました。1つ目の方法は、言語情報を入力とする、深層学習ベースの対話型画像編集モデルを提案した点です。深層学習ベースにすることで、多様に表現される言語情報を、目標とする特定の編集操作に柔軟に対応させることができます。しかし、多様な言語情報を特定の画像編集操作に結び付けることは難しく、これだけでは編集要求と関係ない領域も一緒に編集されてしまう問題に悩まされました。そこで本研究ではさらに、元画像のどの部分が編集すべき領域なのかを明示的に分けるマスク機構を提案し、マスク機構なしの場合より優れた編集が可能であることを明らかにしました。

2つ目の方法は、システムが編集要求に沿って編集を行えるかどうか自信がないときには、複数のモデルで編集した画像からユーザの意図に合うものを選んでもらえる確認戦略をシステムに導入した点です。機械学習モデルは、限られたデータセットを用いて学習されるため、1つのモデルで多様な入力のすべてに対応することは困難です。たとえばマスクありモデルは、髪の毛など大きな領域の編集を苦手としています。そのような場合、複数のモデル（マスクあり・なし）による出力をユーザに見せて選んでもらい、より目標を達成できそうな方の画像を選んでもらうという対応策が考えられます。かといって、毎回確認するのはユーザの負担となります。そこで本研究では、生成されたマスクのエントロピーを基準として必要な時だけユーザに確認することで、冗長な対話を削減できることを明らかにしました。



■WEBサイト

<https://seitaroshinagawa.github.io/>

■動画URL (YouTubeチャンネル用)

<https://youtu.be/Xz8tfC7MngQ>

(2021年5月31日受付)

(2021年8月15日note公開)

取得年月日：2020年9月

学位種別：博士（工学）

大学：奈良先端科学技術大学院大学

推薦文：（自然言語処理研究会）

本研究は、画像と自然言語の情報を結び付け、ユーザが自然言語で対話しながら画像を編集できる深層学習ベースのシステムの研究です。自然言語は自由度が高く、システムが想定していない入力は問題となります。そこで本研究では、入力文が不確実なときはシステムから確認要求を行うことで目標の画像の生成を図る点が特徴です。



品川政太郎

研究生活：「人が人をお願いごとをするみたいに、雑にお願いしてもちゃんと仕事をしてくれるシステムを作りたい。特に画像生成のような、候補が無数にあるような難しいタスクをうまく解くには、対話をしてユーザとの意図をすり合わせながら問題解決できるシステムがこれから重要だ」と考えてこの研究を始めました。実際には画像生成部分だけでも難しく対話的な方法論を深掘りできなかつたのが心残りとなりました。今後はこれまでのモデルを基礎として、自然言語による対話的な最適化により力を入れたいと思っています。他のV&L研究でも、対話的なインタラクションの要素が重要だと報告する論文が徐々に出てきており、今後のこの分野の発展がますます楽しみです。博士課程の間はしんどいこともたくさんありましたが、日々新しい発見があったり、国内外のたくさんの人に出会えたり、とても刺激的な毎日でした。博論を終えた後の空の青さは一生忘れない思い出になるのでぜひ体験してみてください。



3D Reconstruction in Scattering Media

 情報処理学会・学会誌「情報処理」
2021年8月15日 08:54



邦訳：散乱媒体下での三次元復元

藤村友貴

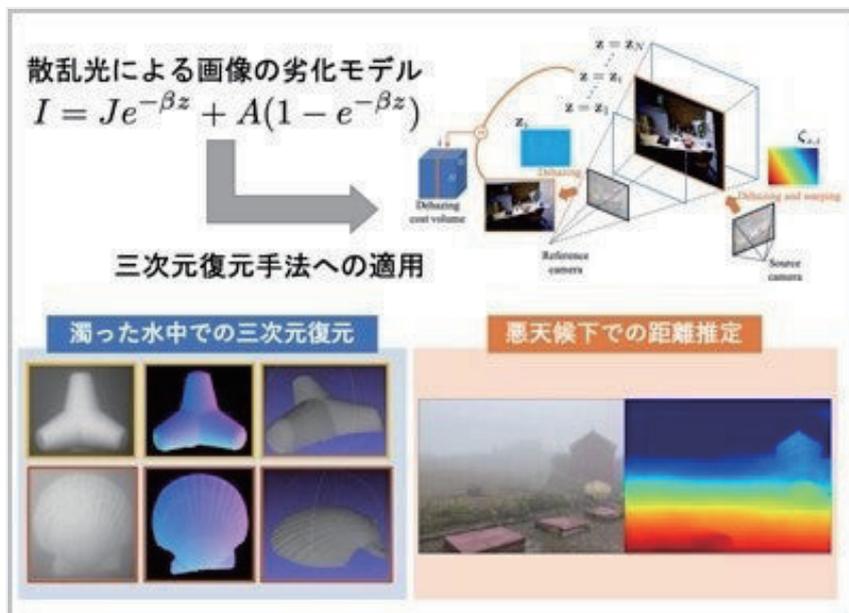
(奈良先端科学技術大学院大学 先端科学技術研究科情報科学領域 助教)

----- keyword -----

コンピュータビジョン

三次元復元

散乱媒体



【背景】 散乱媒体下では光の散乱で画像が劣化

【問題】 劣化した画像を用いてシーンの三次元を復元

【貢献】 物理的な画像劣化モデルの三次元復元手法への適用

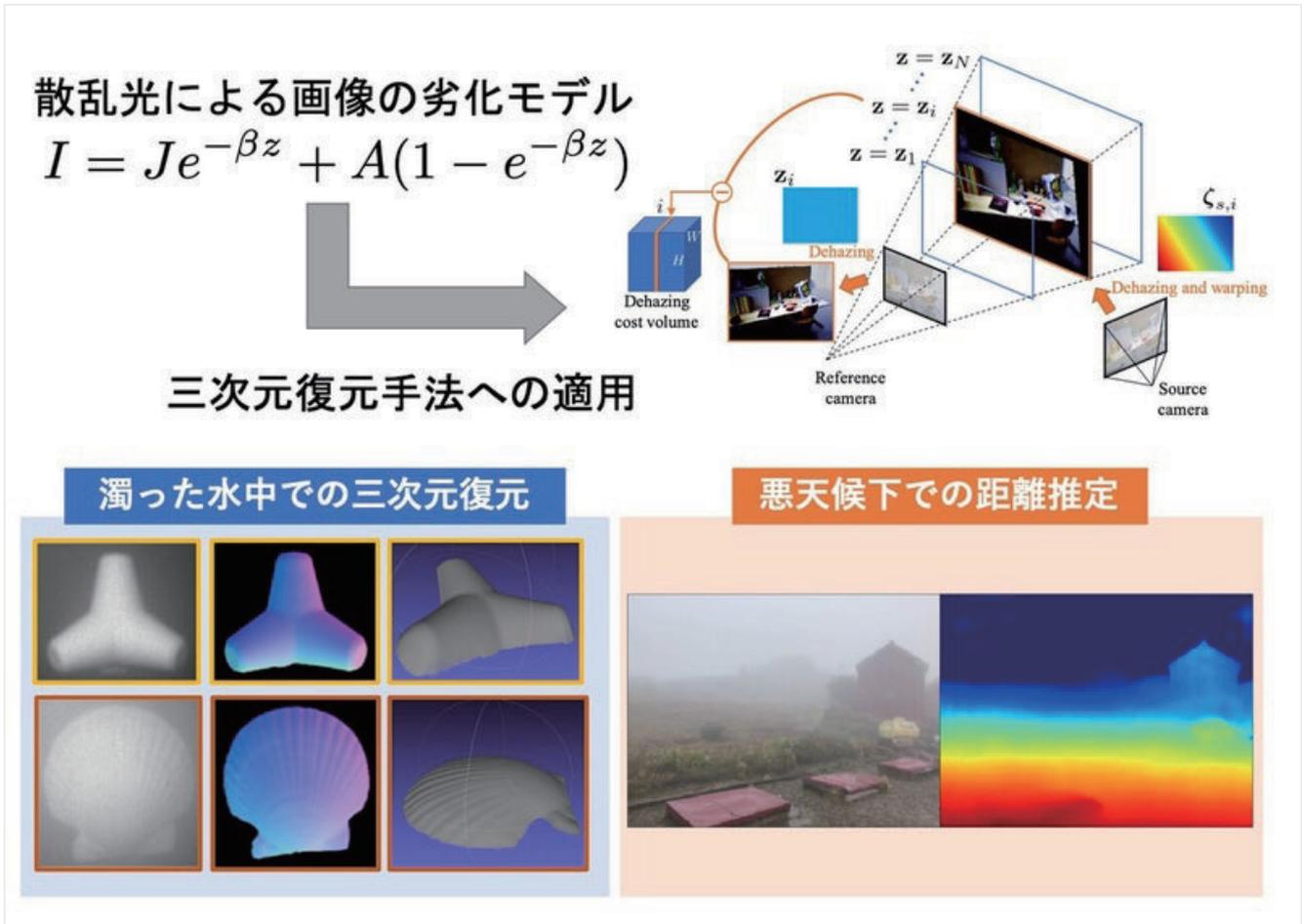
三次元復元とは、カメラで撮影された二次元画像やセンサから取得した信号を用いて、被写体の形状や法線ベクトル、被写体までの距離といったシーンの三次元情報を推定する技術のことである。近年このような技術は自動運転やVR・ARなどに広く応用されており、コンピュータビジョン分野においては依然重要な研究対象となっている。

三次元復元手法はその原理によって3つに大別できる。1つ目が「視差に基づく手法」である。人間の目と同じように2台のカメラが存在する場合は、各カメラで撮影された画像同士の間で対応点をとることで、三角測量の原理によりシーンの三次元位置を計算することができる。2つ目が「陰影に基づく手法」である。物体に光を当てると物体表面の傾きと光の入射角度から物体表面に陰影が現れる。陰影に基づく手法の1つである照度差ステレオ法では、別々の光源下で撮影された複数枚の画像を入力として、物体表面の法線ベクトルを推定する。3つ目が「光の伝搬時間に基づく手法」である。光を照射すると、微小な時間の後に物体表面で反射した光が返ってくる。Time-of-Flight (ToF) カメラではこの微小な時間から対象までの距離を計測する。

本研究は、このような三次元復元技術を散乱媒体下で実現しようという試みである。我々の暮らす実世界では、濁った水中や、霧・煙が拡散したような環境が存在

する。このような環境では、空間中の微粒子（散乱媒体）によって、光の散乱・吸収が引き起こされる。したがって、散乱媒体下でカメラを用いて画像を撮影すると、本来観測すべきでない散乱光を観測したり、物体で反射した光や信号が減衰したりすることにより、撮影された画像にはコントラストが低下するなどの劣化が生じてしまう。多くの三次元復元手法は観測した画像や信号を直接用いるため、このような散乱媒体による劣化の影響を受けてしまう。本研究では物理モデルを用いて劣化を定式化し、従来の三次元復元手法と組み合わせることで散乱媒体下での三次元復元を実現する。濁った水中における水中探査ロボットや、悪天候下での自動運転などがアプリケーションとして期待できる。

本研究では最初に述べた3つの三次元復元手法、すなわち「視差に基づく手法」、「陰影に基づく手法」、「光の伝搬時間に基づく手法」のそれぞれについて、散乱媒体下へ適用するための手法を提案している。視差に基づく手法ではカメラの間の対応点を計算する必要があるが、散乱媒体下では画像が劣化してしまうため対応点の計算が難しくなる。提案手法では対応点を計算する際、画像の劣化モデルを考え、対応点の計算と画像の劣化の復元を同時にモデル化できる手法を提案した。陰影に基づく手法では、光源から出た光が散乱を介してカメラでどのように観測されるかを詳細にモデル化し、物体表面の法線ベクトルを計算する際はその影響を考慮する手法を提案した。光の伝搬時間に基づく手法では、正弦波照射型ToFカメラを用い、散乱による影響をその観測である振幅と位相の空間でモデル化することで、散乱媒体下で生じる距離計測の誤りを修正する手法を提案した。



(2021年5月28日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（情報学）

大学：京都大学

推薦文：（コンピュータビジョンとイメージメディア研究会）

本論文は、散乱媒体下における光の散乱や吸収を物理的にモデル化し、散乱媒体下での三次元復元として多眼ステレオ・照度差ステレオ・ToFの3つのアプローチを議論している。一連の研究成果は、トップカンファレンスのオーラル発表やトップジャーナルに掲載されるなど高く評価されており、推薦に値する。

藤村友貴（正会員）

研究生活： 博士課程での研究は大変なことも多々ありましたが、指導して下さった先生方や研究室のメンバのおかげで、大いに成長することができたと感じています。これから博士課程に進学する予定、あるいは進学するかどうか迷っている人の中には、学力面で不安を抱いている方もおられるかと思います。博士課程という環境は想像以上に自分自身を成長させてくれるので、自身を持って進学してもらえればと思います。



Behavior-based DNN Compression: Pruning and Facilitation Methods

♡ 2

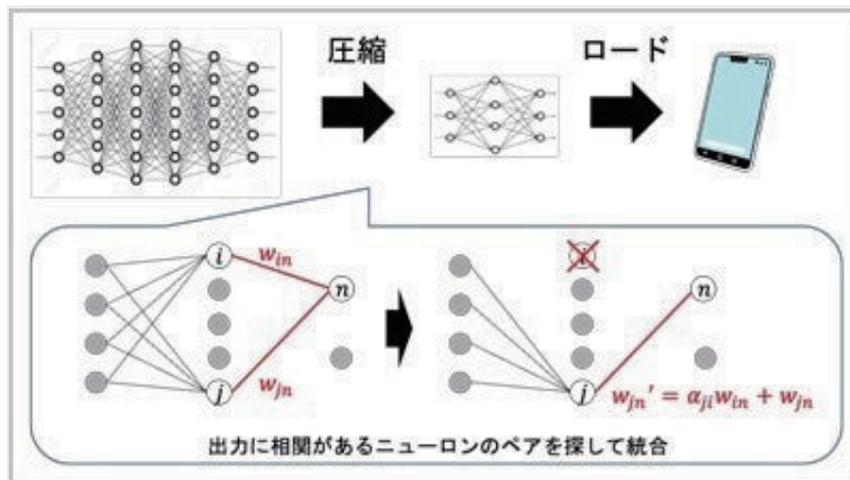
 情報処理学会・学会誌「情報処理」
2021年8月15日 08:55



邦訳：ニューロンの振舞いに基づくニューラルネットワークの圧縮:プルーニング
手法およびより効果的な圧縮のための補助的手法

菅間幸司

(和歌山大学システム工学部 特別研究員)



----- keyword -----

深層学習

高速化

双直交基底

【背景】ディープニューラルネット（DNN）は有用な機械学習モデルであるが、
計算コストが高い

【問題】 スマホや監視カメラ等のエッジデバイスではDNNを使用しにくい

【貢献】 DNNの精度を保ちつつ圧縮できる手法を開発した

ディープニューラルネット（DNN）は、画像を見てそれが何であることを識別する画像分類や、画像中の人や車の位置を特定する物体検出など、さまざまな分野で使われるようになった。DNNは性能が良いものの、計算コストが高く、大規模な設備がなければリアルタイムで動かすことが難しい。そのため、スマートフォンなどの計算能力が限られたデバイスにおいてDNNを動かすためには、性能を保ちつつ、圧縮することが必要である。

本研究は、DNN内部において、挙動が似たニューロンのペアを探して統合する、という発想からスタートした。入力データに対して、いつも近い値を出力するニューロンのペアが存在するならば、一方を削除し、もう一方の重みを更新して、削除された方の出力を再構成する。これにより、次の層にはほとんど誤差を生じさせることなくニューロン数を減らすことができる。

上記の手法の発展形として、削除されたニューロンの出力を、その他すべてのニューロンの出力から、最小二乗法を用いて再構成する手法を考えた。これにより、再構成後の誤差をより小さくできる。なお、ここでは、最小二乗法を用いるということは連立方程式を解くようなものだという理解で構わない。

問題は、削除するニューロンを選ぶために、膨大な量の計算を要することである。再構成を行うことが前提であるから、再構成後の誤差が最小になるようにニューロンを選ぶべきである。しかし、再構成後の誤差を比較するためには、各ニューロンについて、試しに削除と再構成を試してみる必要がある。これは、ニューロンが10,000個あるならば、9,999個の変数がある連立方程式を10,000回解くようなものである。

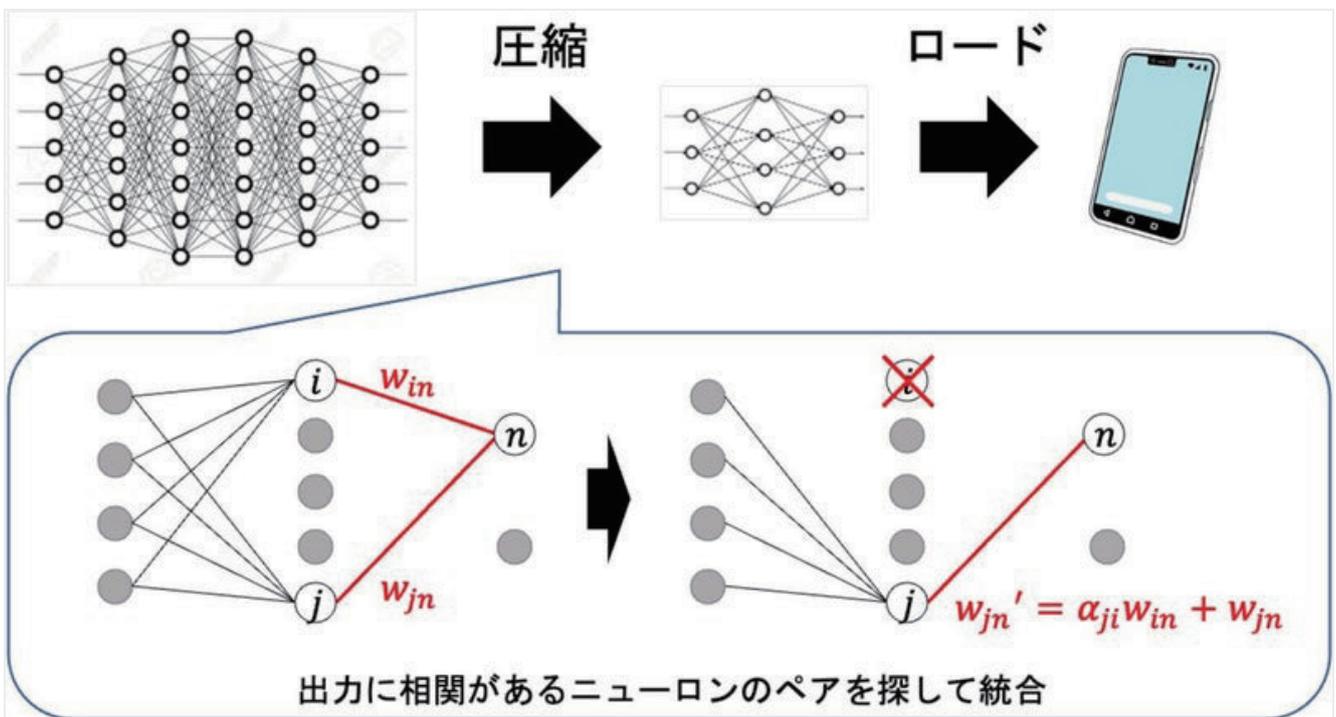
そこで、双直交基底というものをを用いて再構成後の誤差を高速に計算できるアルゴリズムを開発した。先の例においてこのアルゴリズムを用いると、10,000変数の連立方程式をたった一度解くだけで済むため、ほぼ10,000倍の速さで計算できる。これにより、再構成後の誤差を基準にニューロン選択が行うことが可能になる。

実は、最小二乗法を用いて再構成を行う手法は以前から存在している。しかし、再構成後の誤差を効率的に計算する方法がなかったため、従来手法では再構成“前”の誤差を基準にニューロン選択を行っていた。一方、本研究の提案手法では、再構成“後”の誤差を用いることができるため、より適切なニューロン選択が可能である。

実験の一例を紹介する。VGG16という画像分類を行うDNNがある。これを、計算量が元の1/5になるまで圧縮し、ImageNetというデータセットを用いて予測精度

を評価したところ、従来手法では、精度が20%以上も低下した。一方、本研究の提案手法を用いると、精度の低下は10%未満に抑えることができた。

また、本研究では、上記の圧縮手法をより効果的に用いるための補助的な手法として、DNNの各層における圧縮率（削除されるニューロンの割合）を最適化する手法や、ResNetという分岐構造を持つタイプのDNNを効率よく圧縮するために、分岐構造を直列化する手法も提案している。



(2021年5月26日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（工学）

大学：和歌山大学

推薦文：（コンピュータビジョンとイメージメディア研究会）

本論文は、ニューラルネットの圧縮法に関するものである。ある層のニューロンを削除した後、次層への影響を緩和する手法において、（1）削除すべきニューロンを効率良く求める方法、（2）全体への影響を小さくする各層の圧縮率を求める方法、（3）分岐のある場合も効率良く圧縮する方法、が骨子であり、強力な実用的圧縮法である。

菅間幸司

研究生活： 私は、修士課程までは化学系、就職してモータの技術営業、転職してマーケティング、そして情報系の博士課程と、ヘンテコな経歴の持ち主です。博士課程に進学したばかりのころは、周りに追いつかなければと焦りを感じました。で

すが、数式をいじってあれこれ考える時間が楽しく、すぐに研究に没頭するようになりました。特に1年目は毎晩遅くまで研究室に残り（以前は残業があんなに嫌いだったのに）、家族に心配されてしまったほどです。その努力の甲斐もあり、また先生や後輩たちのサポートもあり、2年目に参加した国際会議では賞をとることができましたし、3年目には博士論文を書き上げることができました。

博士課程進学に対しては、なんとなく二の足を踏んでしまうという方もいると思います。ですが、自分の興味のために時間とエネルギーを好きなだけ使えるというのは、博士課程ならではだと思えます。進学を考えている方は、思い切って飛び込んでみてください。なんとかなりますよ。



Robust and Fast Eulerian Video Magnification for Practical Applications

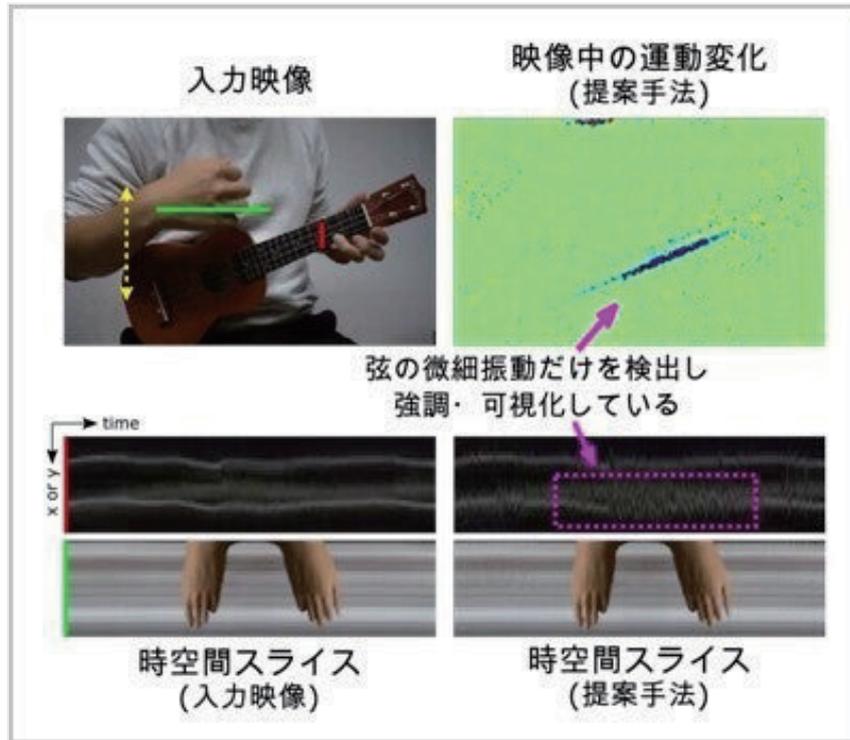
 情報処理学会・学会誌「情報処理」
2021年8月15日 08:56



邦訳：実用化に向けた頑健で高速なオイラービデオマグニフィケーション

武田翔一郎

(NTTメディアインテリジェンス研究所 研究員)



----- keyword -----

映像強調 (ビデオマグニフィケーション)

映像合成

信号処理

【背景】 映像中の小さな色／運動変化を強調し，可視化する

【問題】 映像中には大きな色／運動変化や撮像ノイズが混入する。加えて、高速処理が困難

【貢献】 神経科学知見の応用や処理過程の冗長性削減による問題解決

我々の世界には、とても小さな自然／物理現象が数多く存在している。たとえば、ヒトの皮膚は血液循環に伴って色が微かに変化しているし、風などの影響で建造物はわずかに揺れている。また、飛行ドローンは自身の安定飛行のためにホバリング（微小な揺れ）を絶えず行っている。これらは状況理解や異常検知の手がかりとして有用だと考えられるが、とても小さい色／運動の変化であり、ヒトの視覚能力で捉えることは困難である。

このような小さい色／運動の変化を映像中から検出し、強調・可視化する「映像強調（Video Magnification）」と呼ばれる映像合成技術が近年注目されている。しかしながら、従来の映像強調技術は、（1）映像中の大きい色／運動変化や（2）撮像ノイズを誤検出するため、映像強調技術を適用した際に大きなアーチファクトが発生するという課題を抱えていた。また、（3）処理が低速でリアルタイム性が要求される分野（たとえば、医療や工場監視など）への適用が困難であるという課題も存在していた。

本研究では、上記（1）、（2）、（3）の問題を克服することで、映像強調技術の実応用範囲を拡大させることを研究目的とした。具体的には、それぞれの問題に

対して以下のような解決手法を提案した.

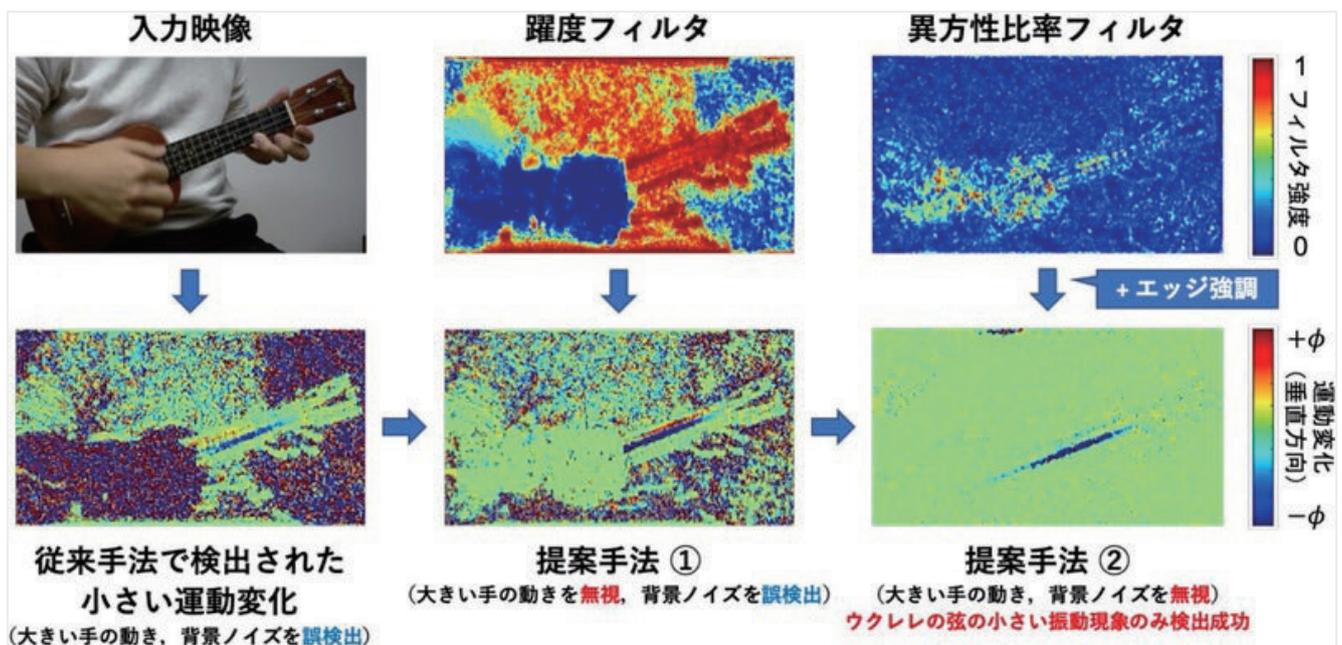
(1) 小さい色／運動変化は大きいものと比較して時間的に滑らかに変動しているという観察結果に着目し、医療分野において時系列データの滑らかさを評価する指標として広く使用されている躍度という指標を応用した新しい時空間フィルタ（躍度フィルタ）を設計した. この躍度フィルタを従来の映像強調技術に適用することで、大きい色／運動変化の影響を無視しつつ、小さい色／運動変化のみを選択的に検出し、強調・可視化することが可能になった.

(2) 自然／物理現象由来の小さい色／運動変化は、ある規則によって1方向に変化していることが多く、その時間分布は異方性（偏って広がる）を持っているが、撮像ノイズはそれが持つランダム性によって等方性（均一に広がる）の時間分布形状を持っている. 本研究では、これら時間分布形状の違いに着目し、神経科学の分野で脳内水分子形状の異方性を評価するために使われている異方性比率という指標を応用した時空間フィルタ（異方性比率フィルタ）を新たに設計した. この異方性比率フィルタを従来の映像強調技術に適用することで、撮像ノイズが存在する中でも、小さい色／運動変化のみを選択的に検出し、強調・可視化することが可能になった.

(3) 映像強調技術は他の映像処理／合成技術と同様に、映像解像度や時間フレーム数が増加すると計算量が増大し、処理速度が著しく低下するという課題を抱え

ていた。そこで本研究では、映像強調技術の処理過程における冗長性を効率的に削減するために、映像強調技術を適用すべき重要な映像領域を自動的かつ適応的に絞り込む、新しい映像処理技術を提案した。本技術は、従来の映像処理技術と比較して、約2倍の高速化を達成することに成功した。

博士論文では、実映像による定性的評価とシミュレーション映像を用いた定量的評価の両面で、提案手法の有効性を確認している。



(2021年5月20日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（工学）

大学：筑波大学

推薦文：（コンピュータグラフィックスとビジュアル情報学研究会）

動画中の、たとえば人物の拍動に伴う振動や顔色の変化など、微細な色変化や運動を強調して可視化する技術が注目されている。本論文では、既存手法の問題点であった、（1）素早く大きく動く物体に対しては検出・強調に失敗する、（2）ノイズに弱い、（3）計算に時間がかかる、の3つに対して解決策を提案し、実用性を大きく向上させた。

武田翔一郎

研究生活：私は修士過程で大学院を卒業し企業研究所に就職しましたが、企業での研究を進める中で博士という学位に改めて尊敬と強い憧れを抱き、社会人博士課程進学という道を選びました。仕事と学業の両立（どちらも研究活動ではありません

が) は想像以上に大変でしたが、両者の面から研究を捉え直すことで新たな課題や解決法が思いつくケースにも遭遇し、非常に有意義な経験をすることができました。これからは博士学位を持った企業研究者として責任を持ち、身につけた技術、知識、そして経験を広く社会に還元し、より良い未来を創出していきたいと考えています。最後に、以前までまったく交流のなかったこの私を快く受け入れてご指導してくださった金森由博先生、遠藤結城先生、そして三谷純先生にはこの場を借りて深くお礼を申し上げます。



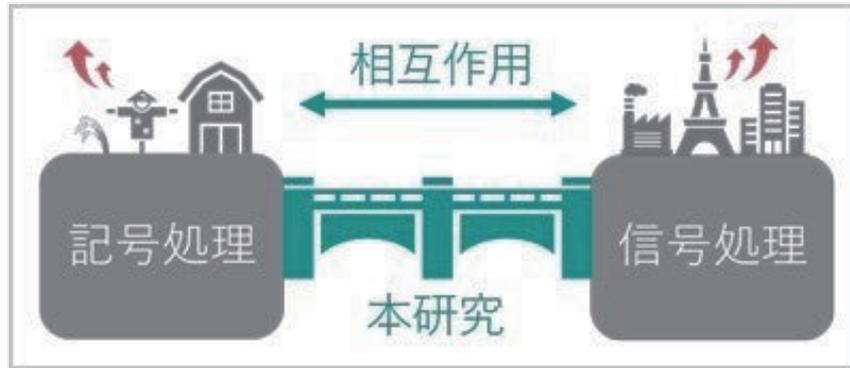
音楽の記号創発に向けた信号と記号の相互最適化フレームワークの構築

 情報処理学会・学会誌「情報処理」
2021年8月15日 08:57



澤田 隼

(東京理科大学工学部情報科学科 嘱託助教)



----- keyword -----

音楽情報処理

記号創発

記号と信号

【背景】 記号処理と信号処理を区別し，独立して発展

【問題】 記号と信号の相互的な修正が必要

【貢献】 音楽情報処理の認識制度の向上，人間の認知や知性を解明する手がかり

私はこれまで時系列データの1つである音楽を対象として，コンピュータを用いて情報処理を行うことにより，その音楽自体の根底に迫ると同時に，人間の認知や知性を探求することを目指してきました。人工知能の知識表現において，そこで使われる記号を実世界の実体が持つ意味に結び付けられるかという記号接地問題があります。私はこれまで音楽における記号接地問題に取り組んできました。音楽情報

処理の分野では、楽譜などの記号を対象とした研究を記号処理、音響信号を対象とした研究を信号処理、といった区別をしてそれぞれが独立して発展してきました。しかし、システムがより人間らしく振る舞うためには、記号の領域と信号の領域が相互に作用する枠組みを構築し、適切な記号接地を実現させる必要があると考えています。そこで、記号接地問題の本質的な解決とも言える記号創発問題に取り組みました。記号創発問題を解決することができれば、記号の領域と信号の領域が相互に作用しながら音楽音響信号が組織化されていく過程を捉えることができ、本質的な記号接地問題の解決につながると考えています。

本研究では、音楽において重要である拍、リズム、メロディーという3つの要素を対象として、性質の異なる3つの問題を、創発的な記号と信号の相互最適化フレームワークという共通の観点から解決を試みました。以下の3つの課題を遂行し、音楽情報処理における信号と記号の相互最適化フレームワークの有用性を検証しました。

1. 拍位置の修正と創発的なグルーピング構造の獲得

音楽情報処理では、音響信号を対象として拍毎に抽出した特徴量（ビート同期特徴量）を用いることがありますが、人間が演奏した音楽音響信号はテンポが一定でないことや、わずかなテンポや発音タイミングの揺れなどが含まれています。そのような音響信号の拍の推定位置のズレによって特徴量が変わってしまうという問題があります。この問題を解決するために、最適なビート位置を求める最適化問題を

解く必要があり、これは音響信号の拍レベルでの分節問題と解釈することもできます。

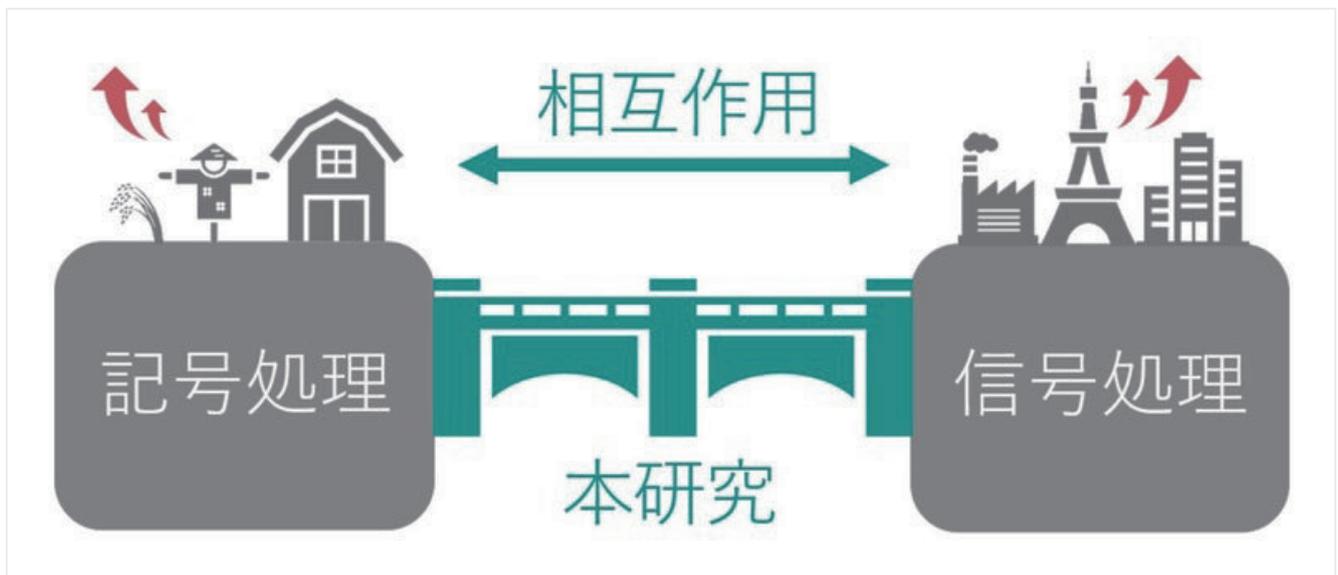
2. 小節レベルのリズムパターンに関する音響パラメタの調節による創発的なリズムパターンの獲得

楽曲中のリズムを一段階抽象化したパターンとして抽出することは、楽曲の構造を理解するうえでも重要です。そこで、リズムのパターンを分類することを考えますが、信号からパターンや特徴量を抽出する際に、たとえば、楽曲のジャンルや使用されている楽器等によって最適な音響特徴量のパラメタが異なります。そのため、音響特徴量のハイパーパラメタのチューニングが必要になります。しかし、曲ごとに人手でチューニングするのは現実的ではありません。この問題を解決するために、最適な特徴量パラメタを求める最適化問題を解く必要があり、これは音響特徴パラメタを調節することで適切な特徴空間を得る、潜在空間での分節問題と解釈することもできます。

3. 旋律からの創発的なモチーフの獲得

モチーフなど、旋律を意味のある単位に分割することは、音楽情報処理において重要な課題の1つです。一方で、モチーフの正解データの作成には専門的な知識が必要であり、十分な質と量のデータを作成するコストは膨大です。さらに、モチーフの解釈は人によって異なることもあり、一意に定まらず、システムとしてそれが正解であるという保証もありません。この問題を解決するために、教師なし（正解

データを与えない) でモチーフレベルの記号を獲得する枠組みが必要です。これは、旋律のモチーフレベルでの分節問題と解釈することもできます。



(2021年6月7日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（システム情報科学）

大学：公立はこだて未来大学

推薦文：（音楽情報科学研究会）

本学位論文では、非言語メディアの1つである音楽の共時的かつ通時的な意味を数理的に理解し表現するために、音楽における記号創発を実現する枠組みとして、信号と記号の相互最適化フレームワークを提案した。当該枠組みに基づく3種類の音楽識別システムの事例を示した。



澤田 隼（正会員）

研究生活：私が音楽情報処理の研究に取り組もうと思ったのは、大学に講演にいらっしゃった先生の音楽に関する研究のお話を拝聴したのが最初のきっかけであったと思います。講演を拝聴し、音楽情報処理という研究分野を知り、音楽を研究の対象とすることが可能であるということや、音楽を研究の対象とすることの意義、奥深さに胸を高鳴らせたのを今でも覚えています。

私が博士課程へ進学したきっかけもいくつかありますが、学会などに参加しさまざまな研究者と出会い、楽しそうに研究について語る姿を目にすることや、刺激的な議論などを通して、次第に研究者という職業への憧れや、この世界に飛び込んでみたいと思う気持ちが強まっていったと感じております。

近年の情勢では、国内外問わず、私の周りの学会のほとんどがオンラインでの開催となっています。オンライン学会のメリットとデメリットを肌で感じている今日この頃です。これは私の感想ですが、もし近年のこのようなスタイルの学会では、研究者の方々との交流も希薄に感じ、博士に進学しようとは思わなかったかもしれません。ただ、近年の世相を嘆いているのではなく、目まぐるしく変化する環境に適応していくことが大事なのかもしれません。

そうは言うものの、私は環境には大変恵まれていたと思います。多くの方々にご支援いただきました。博士号取得のためにご協力いただいたみなさまに心から感謝いたします。博士課程に進学しようと思う学生を1人でも増やせるような研究者を目指して精進してまいります。



Generative, Discriminative, and Hybrid Approaches to Audio-to-Score Automatic Singing Transcription

♡ 1

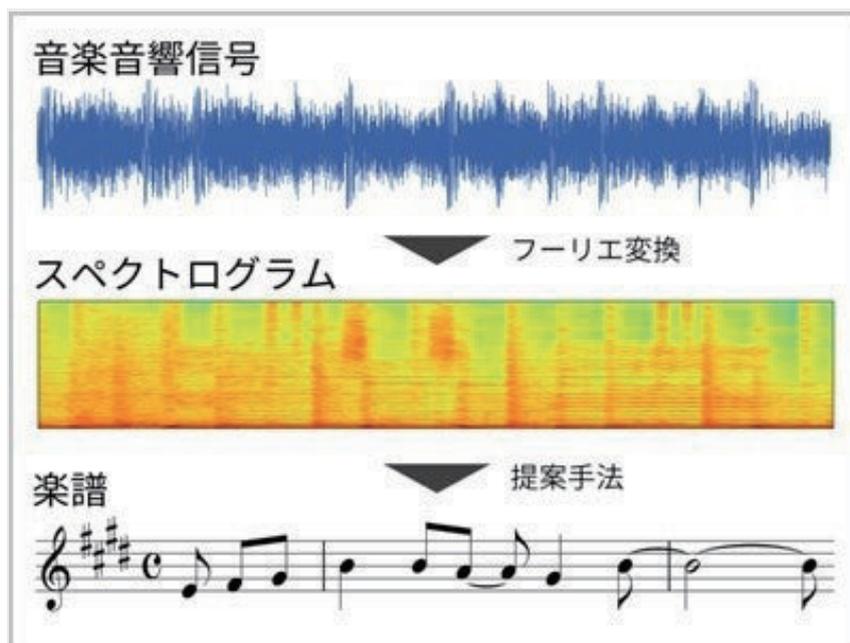
 情報処理学会・学会誌「情報処理」
2021年8月15日 08:59



邦訳：自動歌声採譜のための生成的・識別的・混成アプローチ

錦見 亮

(NTTコミュニケーション科学基礎研究所 研究員)



----- keyword -----

音楽情報処理

自動歌声採譜

機械学習

【背景】 音楽情報処理の基本的認識タスクである自動採譜

【問題】 歌声の大きな変動に起因する音符の推定誤り

【貢献】 言語モデルと音響モデルの統合アプローチ

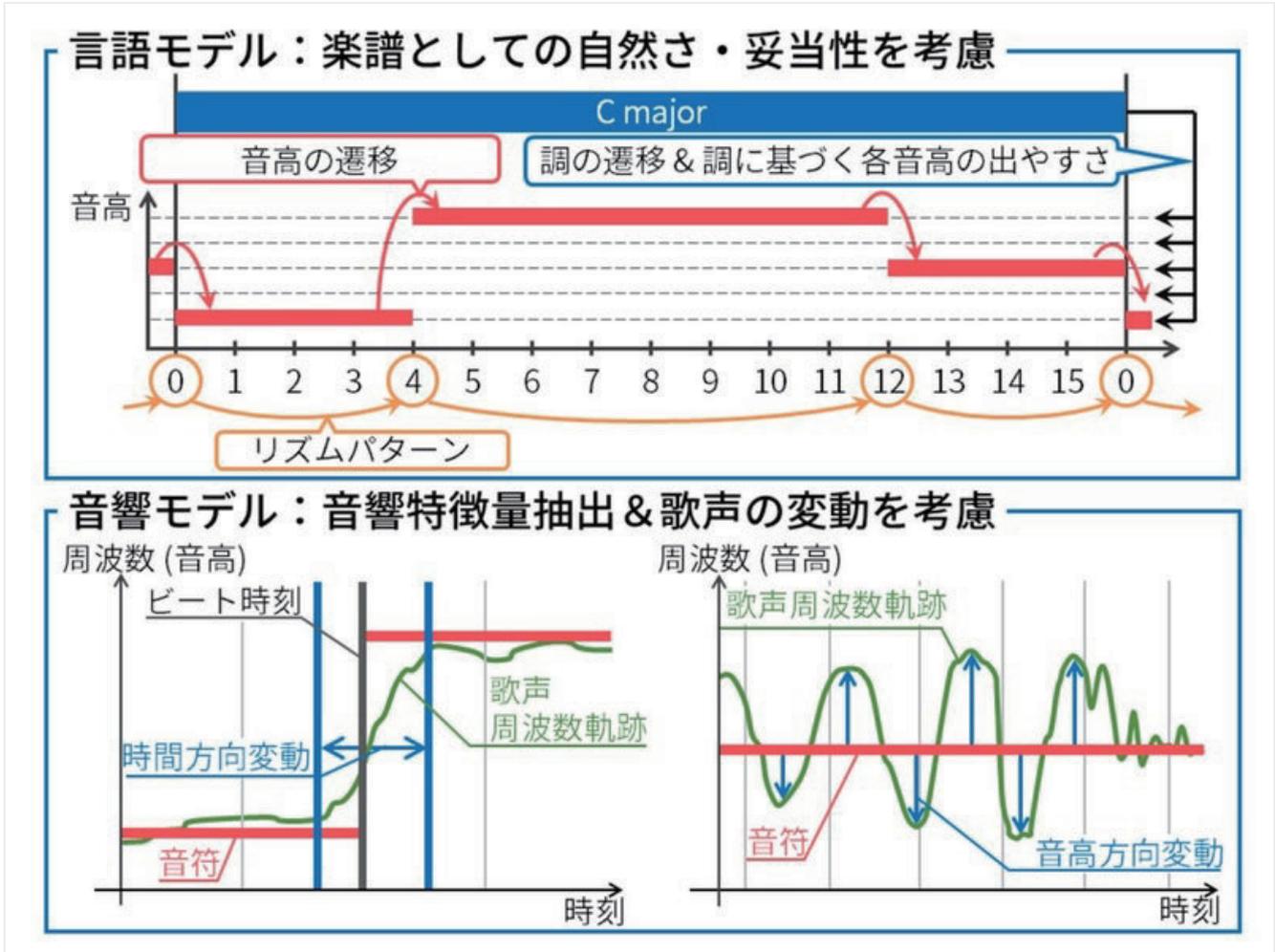
音楽情報処理分野において最も基本的で長年取り組まれている問題として音楽音響信号から楽譜を推定する自動採譜があります。楽譜はポピュラー音楽を含む多くの音楽を記述する方法として最も一般的な形式ですが、音楽に堪能な人でも音楽を聴いて楽譜に書きおこす作業には時間がかかります。自動採譜が実現すれば、演奏したい楽曲の楽譜が即時入手できるだけでなく、音楽のアーカイブ保存・配布・研究用データの収集などさまざまな場面での実応用が考えられます。また、人間の音認知メカニズムを構成論的に解明するという学術的側面からも自動採譜は重要です。音楽は複数種類の音がさまざまな音高・音量・音色・継続長で重なりあっている複雑な音響信号です。計算機にとってはただの波にすぎませんが、人間は音楽に含まれる個別の音を聞き分けることができます。この聞き分ける処理を計算機でどのように再現するかは大変興味深い課題です。

本研究ではポピュラー音楽中の歌声から主旋律（メロディ）の楽譜を推定することに取り組んでいます。本研究がポピュラー音楽中の歌声を対象とした理由は、ポピュラー音楽が最も身近な音楽ジャンルの1つであり、歌声は楽曲の印象に密接に関連する主旋律を担うことが多く中心的な役割を果たすからです。また、歌声は最も原始的な音楽表現の1つであり、その認知機構を解明できれば、他のさまざまな

楽器音の採譜技術の礎になると考えられます。実際に歌声分離や歌声音高（周波数）軌跡推定など、歌声を対象とした研究は音楽情報処理分野において多く行われています。また、採譜された楽譜は、ハミング検索、歌声合成、音楽文法解析、能動的音楽鑑賞などでさまざまな場面で応用できます。

歌声採譜をするためには、連続的な信号である歌声の周波数軌跡を音高・時間方向に量子化して、半音単位の音高やビート単位の音価（長さ）を持つ音符の系列を推定する必要があります。この歌声周波数軌跡には歌唱表現（ビブラートやポルタメント）や喉の物理的な制約に起因する多くの変動が含まれていますが、この変動が原因となり何の制約もなく歌声音高軌跡を量子化するだけでは多くの音符推定誤りが発生してしまいます。

本研究では、言語モデルと音響モデルの統合という観点から歌声採譜の高精度化を行いました。言語モデルでは、曲中で使われやすい半音の種類やリズムパターンを考慮しながら楽譜としての妥当性を表現することで、不自然な音符系列が推定されることを抑制します。音響モデルでは、採譜に必要な音響特徴量の抽出や、歌声の変動を考慮した頑健な採譜を実現します。本研究では、言語モデルと音響モデルの設計方法や統合方法を変化させて、3つのアプローチ（生成的・識別的・混成）を提案し、言語モデル・音響モデルの有効性の検証や各アプローチの精度の評価を行っています。



(2021年5月28日受付)

(2021年8月15日note公開)

取得年月日：2021年3月

学位種別：博士（情報学）

大学：京都大学

推薦文：（音楽情報科学研究会）

音楽音響信号に対する歌声の楽譜推定という国際的にも研究例が少ない課題に取り組み、確率的な枠組みのもとで生成的・識別的な言語モデル・音響モデルを統合するという一貫したアプローチを提案している点で、きわめて高く評価できる。

錦見 亮

研究生活：最初は企業に就職するか博士課程に進学するかを決めきれずに悩んでいましたが、スキルアップできる機会や周囲の環境に恵まれていると感じて博士課程への進学を決めました。博士課程では、3年間という自由な時間を使って次々に登場する新技術や研究活動への取り組み方を学びました。また、国際会議への参加や海外の研究所へインターンを通じて、世界中の人たちと活発に交流・議論できたのも良い経験でした。研究テーマは学部時代に研究室配属時のものを続投しました。必ずしも同じテーマを続けるのが良いとは限りませんが、私の場合は博士課程を含めた計6年間も専念できたことで最終的には良い成果を残せたと感じています。途中、なかなか研究成果が上がらず苦労したこともありましたが、こうして博士号を取

得できましたのも、指導教員をはじめ研究室の皆さんや家族の支えのおかげです。
この場を借りてお礼申し上げます。



楽器演奏における知識の構築と実践的な活用

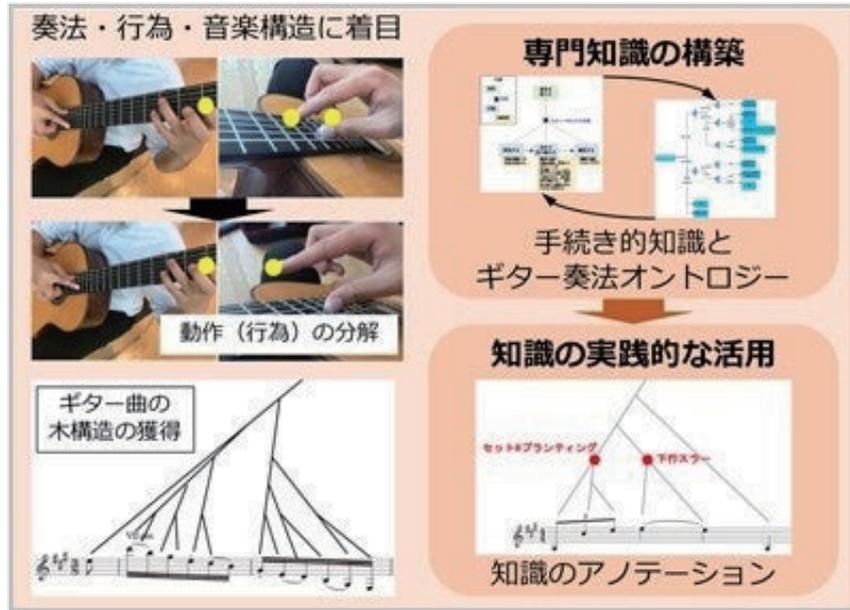
♡ 1

 情報処理学会・学会誌「情報処理」
2021年8月15日 09:00



飯野なみ

(国立情報学研究所情報学プリンシプル研究系 特任研究員)



----- keyword -----

知識の利用と共有
ドメインオントロジー
演奏行為

【背景】 楽器演奏における指導・学習方法の確立

【問題】 知識の共有と実践的な活用の不足

【貢献】 専門知識の構築と実践的な分析、楽曲および演奏者の新しい分析方法の提

案

本研究では、知識を必要とする指導や学習に関する基盤技術を提供するために、

楽器演奏を対象として知識の構築と活用を行いました。

楽器演奏は個人の経験値や身体性といったさまざまな情報を含むために、個人に合った楽曲選びや指導・学習方法の確立が困難です。楽器によっては奏法が発展しているにもかかわらず、指導者間の情報共有が不足しているために学習者が正しい情報を得られていないことがあります。個人に適応した指導、効率的な学習、分野の変化に応じた深い理解を実現するためには、知識の共有や活用が必要です。

このような背景を共有する知識工学では、知識獲得のためのモデルや枠組みが提案されていますが、詳細な実施方法や手順については論じられていません。また、オントロジーなど多くの知識が構築されていますが、実践的に活用するという観点で不足しているという問題があります。

本研究では、知識ベースによる楽器の指導・学習支援を目指して、次の2つの研究項目を行いました。

- (1) 楽器演奏における知識の構築と手順化
- (2) 実践的な知識の収集と分析

(1) では、奏法の種類が多いクラシックギターに着目し、各奏法の動き（行為）を形式的に記述することを試みました。具体的には、人間可読性の高い「手続き的知識」と、機械処理可能な形式を持つ「ドメインオントロジー」を併用しながら

ら相互に構築を繰り返すプロセスを実施し、知識の変化やアンケートを通じて有用性や効果を検証しました。その結果、ドメインオントロジーとして『ギター奏法オントロジー』を構築し、行為の形式的かつ構造的な記述を実現しました。手続き的知識の再構築ではオントロジーを併用することで形式性が高まり、知識の理解が深まることを確認しました。さらにこのプロセスを支援するシステムを設計・開発し、手続き的知識の再構築を行なったところ、行為関係の明示化、語彙の統制において効果を示しました。

(2) では、演奏者や楽曲ごとに異なる“実践的な”知識を収集するために、認知的音楽理論GTTMに基づく楽曲分析で得られた木構造に対して上記(1)の知識をアノテーションする方法を提案し、音楽構造と知識の関係を調査しました。複数のギター曲に奏法をアノテーションした結果、約80%が音楽構造に対応していることを確認しました。さらに、楽譜に書かれている知識や上級の演奏者が持っている知識の特徴を明らかにするために、国際ギターコンクールで頻繁に演奏される難易度の高い曲の奏法を調査しました。その結果、時代ごとに奏法が増加していること、実際の演奏では楽譜情報の約2倍もの奏法を行っていること、演奏者が感じる難しさが奏法の種類の多さや密度と対応していることを明らかにしました。

これらの成果は、楽器演奏領域だけにとどまらず、身体動作を伴う活動全般に対しても適用できると考えています。本研究をきっかけに、技能のような機能的な領域の指導・学習において、知識を獲得、構築、活用していくための基盤技術の研究

に興味を持っていただければ幸いです。



■個人ページ

<https://nami-iino-guitar.wixsite.com/mypage>

■Github (ギター奏法オントロジー)

<https://github.com/guitar-san/Guitar-Rendition-Ontology>

(2021年6月1日受付)

(2021年8月15日note公開)

取得年月日：2020年9月

学位種別：博士（情報学）

大学：総合研究大学院大学

推薦文：（音楽情報科学研究会）

演奏の指導は教授者のスキルに依存していて、クオリティの確保が難しいという問題があった。本研究では、標準化されクオリティが確保された指導を円滑に進めることを支援するシステムの構築を目指しており、その挑戦的な試みの第一歩として、楽器演奏における演奏知識、指導知識のモデル化を行った。



飯野なみ（正会員）

研究生活：私は3歳からクラシックギターを始め、コンクールやコンサート、指導の経験を通して楽器演奏の素晴らしさや難しさを感じてきました。音楽と情報という融合型研究に取り組むきっかけとなったのは、研究者である父の存在です。学部時代に音楽を数学・情報で見ることの面白さを教えてくれ、博士前期過程では音楽情報処理の分野で研究を行いました。その過程の中でさまざまな分野の研究者と出会い、博士後期課程では音楽を知識工学から見る価値を知ることができました。新しい領域を開拓していく研究は、独自性や受容などの面で困難なことが多いですが、音楽を軸にすることで幅広く展開していけることも学びました。指導教員である武田英明先生をはじめ、これまでご指導いただいた研究者の方々や音楽仲間、そして家族に心より感謝申し上げます。



Visualizing and Understanding Computer Go

♡ 1

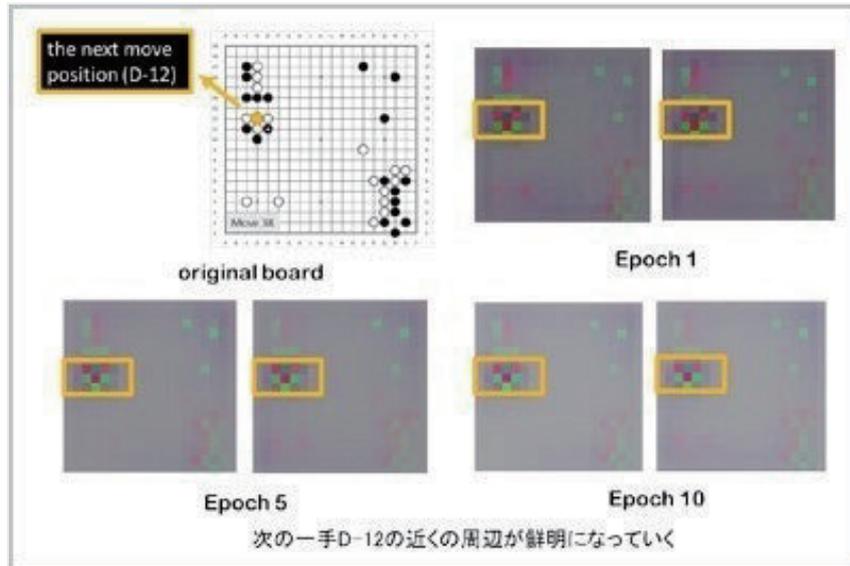
 情報処理学会・学会誌「情報処理」
2021年8月15日 09:18



邦訳：コンピュータ囲碁を用いた思考の可視化に関する研究

龐 遠豊

((株) 日立製作所 研究開発グループ)



----- keyword -----

コンピュータ囲碁

深層学習

可視化

【背景】 コンピュータ囲碁は人間を超えるという目標が達成してきた

【問題】 コンピュータ囲碁と人間棋士の思考過程には大きな違いがある

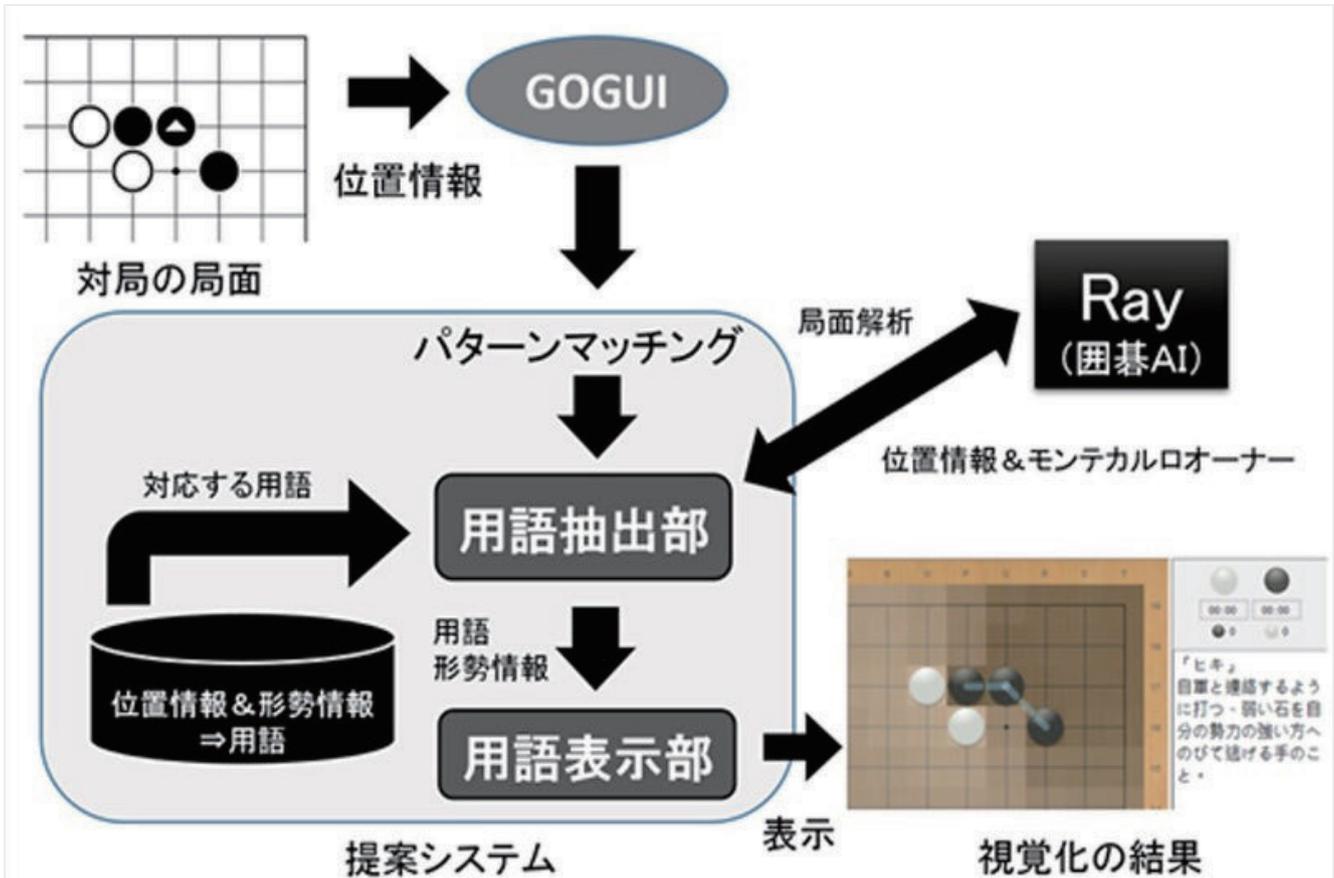
【貢献】 可視化実験によって、コンピュータ囲碁の学習プロセスを考察した

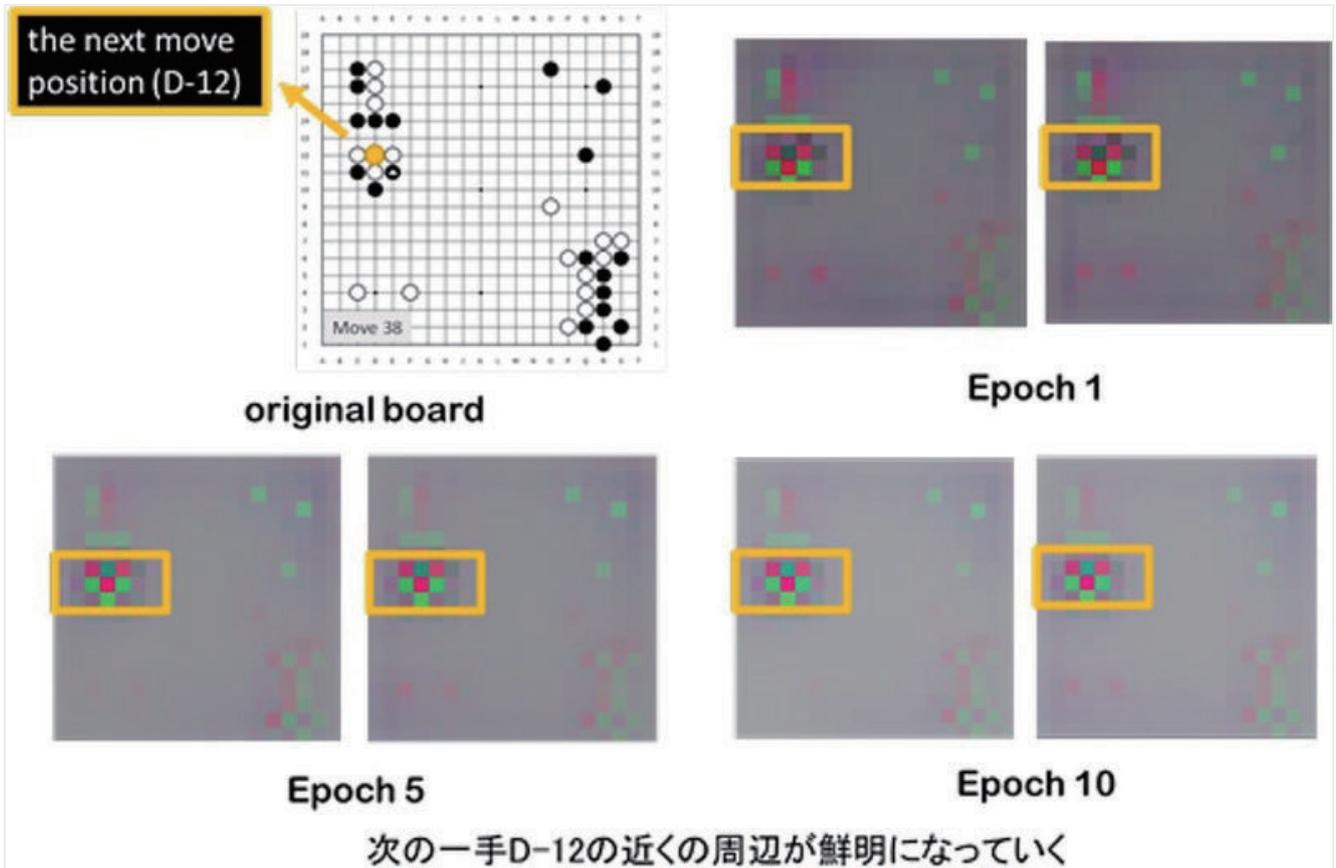
ゲームプログラムと人間の思考過程には大きな違いがあり、ゲームプログラムの考えていることを人間が簡単には理解できないという問題が指摘されている。そこで、本研究では、コンピュータ囲碁を題材にして、2つの異なる可視化研究を行った。

1つ目は、人間の理解を助けることに着目した可視化研究である。本研究は、着手に対応して囲碁用語を表示するシステムを提案した。囲碁用語の判定には、石の単なる「位置関係」だけでなく、局面判断を必要とするものがある。本研究では、囲碁用語の分類からルールベース条件を構築し、比較的強い囲碁プログラムによる局面解析を行うことで、プロ棋士の判断に近い囲碁用語を抽出するシステムを提案した。また、囲碁プログラム「Ray」の形勢分析情報を活かし、形勢や地を可視化した。既存手法に比べて、多くの用語で本システムのほうが一致率の上昇が確認された。特に局面解析を用いた用語では、多くの用語で改善が見られた。また、本研究の実用化について、2019年に、囲碁用語の抽出部分で本研究を利用して「入神の囲碁」という製品が発売されている。

2つ目は、人工知能の思考を理解するための可視化研究である。本研究は、画像認識と自然言語処理分野において深層学習を解釈するための可視化手法を転用して、コンピュータ囲碁におけるポリシーネットワークと呼ばれる深層学習モデルの可視化実験を行った。そして、その可視化結果を分析することによって、ポリシーネットワークの学習プロセスを考察し、3つ特性を指摘した。1) 囲碁のポリシーネ

ットワークは碁盤全体の幅広い情報を対象とするだけでなく、碁盤上の特定の石のパターンにも非常に敏感であること。2) 学習プロセスが進むにつれて、可視化結果に次の一手周辺は鮮明になっていく。3) 囲碁のポリシーネットワークは碁盤の隅と端を重視している。これらの結果は、囲碁のポリシーネットワークに関する重要な性質を可視化しており、難解なコンピュータ囲碁の深層学習の一端を説明する可能性を示したものと言える。最後に、可視化結果の応用例としてポリシーネットワークを改良し、評価実験を行った。改良前と改良後の可視化結果とモデル精度を比較することで、モデルの可視化結果はポリシーネットワークの設計と訓練によって有用であることを示した。





(2021年5月5日)

取得年月日：2021年3月

学位種別：博士（工学）

大学：電気通信大学

推薦文：（ゲーム情報学研究会）

本論文は、コンピュータ囲碁を用いた2つの可視化研究を行っている。1つ目は、専門的な囲碁用語を自動判別し、打ち手に対してリアルタイムで可視化する技術であり、市販の囲碁AIにも応用された。2つ目は、深層学習における学習内容を5つの手法で可視化し、囲碁AIの理解や改良に役立つ技術を提案している。



龐 遠豊

研究生活：博士課程での最大の問題は孤独感ではないかと思います。新たな研究の道を歩み始め、その独特の不安感を理解してくれる人が少ないため、博士課程が病みやすい環境になっています。私の見解では、できるだけ幅広い年齢層、それぞれ研究分野の方と交流し、定期的に自らの学問研究の位置と意味について考えること

が大事と思います。

最後に、これまでの指導教官の伊藤毅志先生をはじめ、研究相談に付き合っていた浅田麻菜さん、仲道隆史さんおよび杵渕哲彦さんに、心より感謝します。



ターン制戦略ゲームへの深層学習の適用

♡ 4



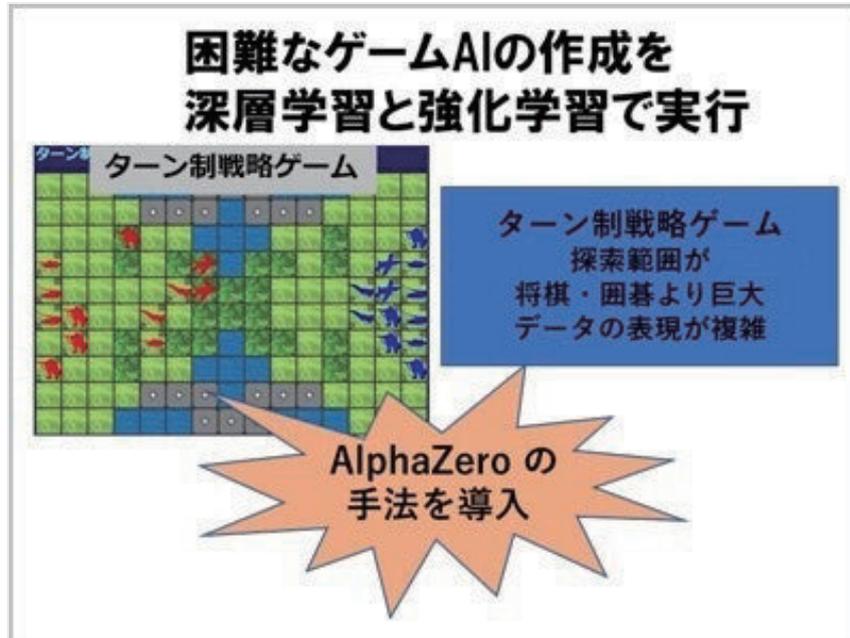
情報処理学会・学会誌「情報処理」

2021年8月15日 09:19



木村富宏

(フリーエンジニア)



----- keyword -----

ゲームAI

ターン制戦略ゲーム

深層学習

アルファゼロ

MCTS

【背景】 ゲームAIへの深層学習の適用の拡大

【問題】 実装が困難なターン制戦略ゲームに最新アルゴリズムを適用する

【貢献】 扱う問題の分析と課題を明確化し実際に動作・検証

ゲームAIの研究は人工知能研究の題材として長年続けられてきたものであるが、近年に至ってからハードウェアとソフトウェアの進化やアルゴリズム研究の進歩とあいまって人間のレベルを超えるプレーができるゲームAIが、将棋や囲碁で開発されるまでになってきた。また、AlphaGoの登場で話題となった深層学習技術を使用した深層強化学習の研究がブームとなり世界的に活発に研究が行われるようになった。そうした中で多くのゲームAIの強さが人間のプレーヤーを超えたというニュースが流れるようになった。しかしながら、まだ多くの他のゲームにおいては人間と同等レベルのプレーができるAIが開発されていない。

この研究で取り上げたターン制戦略ゲームは、人間と同等レベルのプレーをすることのできるAI開発が現状では困難な分野である。困難である要因として挙げられるのは、ゲームシステムが複雑で解析が容易ではなく、旧来の手法をそのまま適用することが難しい場合が多いということが言えるだろう。本研究では、ゲームのシステムの困難さについて実際に解析を進め、そして、近年新しくゲームAIで注目されている深層強化学習の技術を適用し、データを取得しながら動作できるゲームAIを作成し評価を行った。

ターン制戦略ゲームの難しさの要因として言えることは、ゲームに登場する駒や地形の複雑さ、ルールの複雑さ、1ターンで複数の駒が動作するような設定、初期設定であるマップが固定されておらず多種多様なものがある、扱うデータ構造が複雑で多層的になっているなどといったことが挙げられる。本研究ではこのような要

因がありながらも、個々の要素について解決を図りながら動作するAIを試作し、評価した。

研究の最初の段階では、マップの少なさを補うためとAIの性能評価に使用できるようにするためにベンチマークマップ問題を多数作成し、性能を評価して発表した。このベンチマーク問題集は一般の人たちも活用できるように公開している。

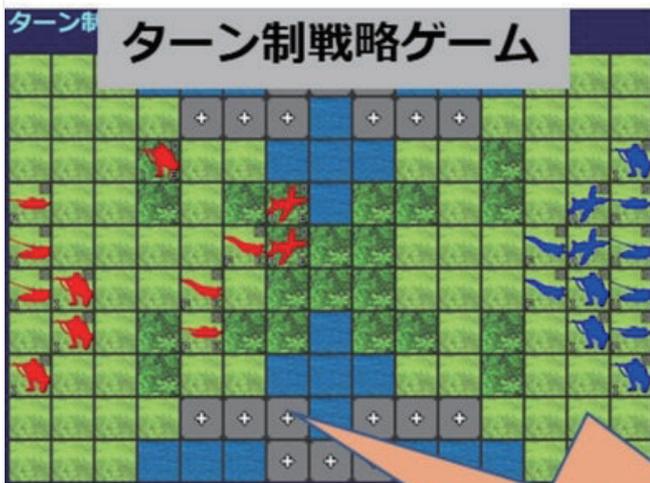
次に複雑で階層的なデータ構造をリカレントネットワークを使用して時分割で出力することで、単純化し深層強化学習の手法としてDQNとProfit Sharing法を採用し、性能を評価した。

次に、AlphaZeroの手法を導入する前に大量の対戦データに基づいた教師付き学習によるPolicy networkの作成に取り組んだ。対戦データはMCTSアルゴリズムによる実際に対戦を行わせることで作成した。作成したPolicy networkの性能は作成に使用したMCTSアルゴリズムとほぼ互角に対戦できるか場合によっては勝ち越すレベルの性能に達した。

最終的にはAlphaZeroアルゴリズムを導入し、かつターン制戦略ゲームに深層強化学習を導入するために必要な変更も含めて全体を設計した。Policy networkに加え新たに必要となるValue networkを設計し、ニューラルネットワークのアーキテクチャとしてはAlphaZeroで採用されたResidualブロックを使用して効率化を図っ

た。結果的な性能は対戦実験で勝率80%以上を達成した。

困難なゲームAIの作成を 深層学習と強化学習で実行



ターン制戦略ゲーム
探索範囲が
将棋・囲碁より巨大
データの表現が複雑

AlphaZero の
手法を導入

(2021年5月26日受付)

(2021年8月15日note公開)

取得年月日：2020年 12月

学位種別：博士（情報科学）

大学：北陸先端科学技術大学院大学

推薦文：（ゲーム情報学研究会）

ターン制戦略ゲームは多数の駒を任意の順で動かせるため行動空間が非常に大きく、AlphaZero等の手法を直接用いても高い性能が得られない。本研究は駒の選択をニューラルネットへの出力ではなく入力として扱うという新発想により、出力の次元を大幅に削減し、高い性能を持つプレイヤーの作成に成功したので推薦する。

木村富宏

研究生活：近年のゲームAIの進化は目覚ましく、将棋や囲碁のAIでは人間のプロ棋士を超える強さのソフトも開発されるレベルになりました。特にAlphaGoの登場に刺激を受けて深層強化学習の研究は世界的にブームとなりました。その後多くのゲームの分野でAIによるゲームプレイのレベルが人間を超えたことが話題になるようになりました。

このような中で人工知能研究の題材としてのゲームAIは、もうふさわしくないのではないかと思われることもあるかもしれません。しかし、まだまだゲームAIの研究で強くしていくのが困難なゲームは多数あります。たとえば、戦略シミュレーションゲームは大変人気のあるゲーム分野ですが、しかしながら、人間と同等の複雑な動きのできるAIを開発することは簡単ではありません。これはゲームの持つシステムが複雑であることとか扱うマップが多数であること、一度に多数の駒を同時に動作させることができる設定などといった事情によっています。

このような分野で研究を進めていくことが人工知能の研究をさらに促進させていくと考えています。