

Exploring Cognitive Plausibility of Neural NLP Models: Cross-Linguistic and Discourse-Level Studies

♡ 2

情報処理学会・学会誌「情報処理」
2022年8月15日 11:14



2021年度研究会推薦博士論文速報

[自然言語処理研究会]



栗林 樹生
(東北大学 研究員)

邦訳：ニューラル自然言語処理モデルの認知的妥当性：言語横断的分析と談話処理

■キーワード

自然言語処理／計算心理言語学／深層学習

【背景】 ことばを処理する人工知能の性能が著しく向上

【問題】 人工知能の言語処理が人間に近づいているのか

【貢献】両者の乖離を明確化し、その原因を探求

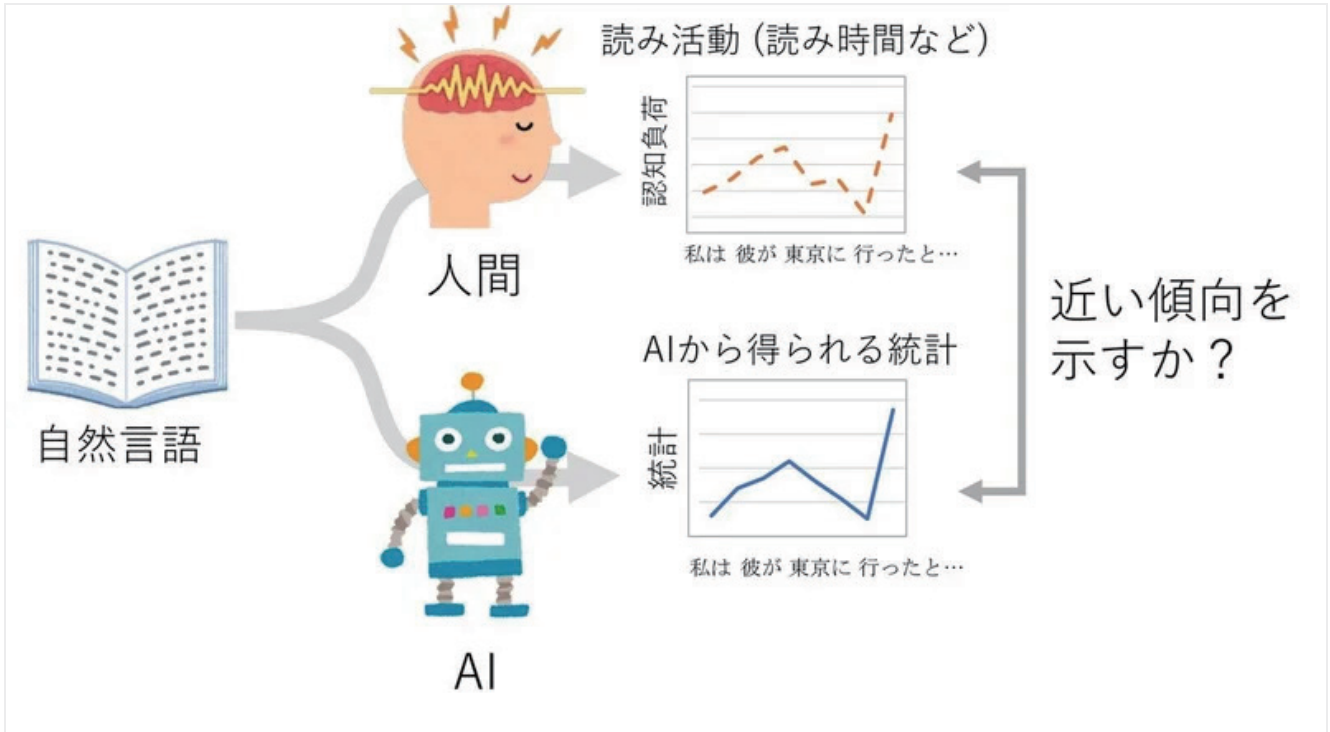
機械に人間のことば（自然言語）を処理させる自然言語処理分野の進展は近年著しい。たとえば人間のように流暢な文章を自動生成することが可能になってきた。本研究では、近年進展した人工知能の言語処理と、人間の言語処理を対比させることで、自然言語処理分野のみならず、言語学や心理学といった周辺基礎分野に新たな知見を還元することを狙っている。

たとえば、本研究の先で答えたい問いは、「どのような計算によって人間の言語処理が実現されているのか」といったものである。認知科学や心理言語学分野では、人間の読み活動（視線移動や脳活動）を計測し、コンピュータでモデリング・再現することでその仕組みを説明しようとしてきた。このようなモデリングを行うには、ことばの意味や特徴をコンピュータで扱いやすい形式で定量化しておく必要がある（たとえば「甘い」という単語の特徴をコンピュータに直接教えるのは難しい）。本研究では、そのような定量化に言語処理分野の大規模言語モデルを用い、人間の言語処理について数理的なモデリングを行った。具体的には、人間が文を読んでいる最中の「先読み」について、さまざまな言語モデルが計算する単語の予測確率と人間の読み活動を比較し、工学的に高性能な言語モデルで必ずしも人間をうまく説明できないことや、得られる観察の言語依存性などを明らかにした。

また、人工知能と人間の対照という試みは、「人間特有の言語能力」についても

示唆を与え得る。人間の「特有さ」を語る上では比較対象が必要であり、典型的には他の動物が引き合いに出されるだろう。本研究では、そのような比較対象として人工知能・自然言語処理システムを活用しているという見方もできる。言語処理システムを比較対象として持ち込む主な利点は、システムの内部を直接制御しながら人間と比較できる（直接観察できない人間の内部についても手がかりが得られるかもしれない）ことや、生物を扱った場合には非現実的な検証（たとえばWeb上の莫大なテキストをただひたすら覚えてもらうと人間らしくなるのか）が実現できることである。このような視点のもと、本研究の後半では人間の談話処理（文章の処理）について示唆を与えている。言語処理システムの一部を制御して人間の談話処理理論に従う／従わないシステムを作成し、理論に従うシステムがより人間らしい談話解析を実現することから、間接的ではあるが理論に対して概念実証的な支持を与えている。

最後に、本研究はやや探索的に行った自身の研究^{1) ~5)}を、計算心理言語学的な視点から位置づけ直したものである。自然言語処理分野と心理言語学といった周辺基礎分野を繋ぐ学際的な探求は限定的であり、本研究では、多言語を対象にする方向や、複数文からなる談話の処理を対象とする方向に焦点を拡大している。また、人間の言語能力という語りづらいものを扱う以上、方法論や何をどこまで「分かった」と言ってよいのかについて、多くの考え方・議論があると考えられる。そのような議論の余地も含め、本研究を通して、多角的な関心が自然言語処理に寄せられるきっかけになれば幸いである。



■Webサイト／動画／アプリなどのURL

<https://kuribayashi4.github.io/>

参考文献

- 1) Kuribayashi, T., Ouchi, H., Inoue, N., Reisert, P., Miyoshi, T., Suzuki, J. and Inui, K. : An Empirical Study of Span Representations in Argumentation Structure Parsing, In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL-2019), pp.4691-4698 (2019/07).
- 2) 栗林樹生, 大内啓樹, 井之上直也, 鈴木 潤, Paul Reisert, 三好利昇, 乾健太郎 : 論述構造解析におけるスパン分散表現, 自然言語処理 (domestic

journal) , Vol. 27, Number 4, pp.753-780 (2020/12).

3) Kuribayashi, T., Ito, T., Suzuki, J. and Inui, K. : Language Models as an Alternative Evaluator of Word Order Hypotheses: A Case Study in Japanese, In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020), pp.6452-6459 (2020/07).

4) Kuribayashi, T., Oseki, Y., Ito, T., Yoshida, R., Asahara, M. and Inui, K. : Lower Perplexity is Not Always Human-Like, In proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL/IJCNLP 2021), pp.5203-5217 (2021/08).

5) Kuribayashi, T., Oseki, Y., Brassard, A. and Inui, K. : Context Limitations Make Neural Language Models More Human-Like, arXiv (2022/05).

(2022年5月31日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（情報科学）

大学：東北大学

推薦文【メディア知能情報領域】 自然言語処理研究会

コンピュータが人間と同じようにことばを捉えているのか？という問いに対して、この研究ではコンピュータが深層学習を使ってことばを扱うときの特性を詳細に分析して人間とコンピュータで似ているところ、異なるところを明らかにしました。人間と同じようにことばを扱えるコンピュータの実現に向けたとても重要な研究です。

研究生生活 日本語や英語などの人間のことばを機械に処理・理解させる分野が、自然言語処理分野です。「処理・理解とは何か？」「何が達成されたらゴールなのか？」など、そもそも答えるべき問題が自明でなく、「何を問うか」自体が大きな問いでした。自分はというと、心惹かれる対象は漠然とあるものの、何を問いたいかを明確に言語化し、興味を学問の大きな流れに位置づけることに苦労しました。したがって私の博士課程は、「これはきっと面白い」というある種の直観を信じて、がむしゃらに小さな研究のサイクルを回しつつ、「つまり大きな意味で自分は何を知りたいのか」をゆっくり鮮明にしていく日々でした。自身の興味に従って自由に研究に向き合うことを容認いただいた、指導教員の先生方や、研究室のメンバー、またともに議論・研究いただいた学外の先生方に深く感謝申し上げます。



AIが読みやすいテキストの区切り方を自動で獲得

Task-Oriented Word Segmentation



情報処理学会・学会誌「情報処理」
2022年8月15日 11:15



2021年度研究会推薦博士論文速報

[自然言語処理研究会]

平岡 達也

(富士通 (株) 富士通研究所 研究員)

邦訳：タスクに応じた単語分割

■キーワード

自然言語処理／単語分割／機械学習

【背景】 AIにテキストを入力するときは短い単位に区切る**【問題】 AIがタスクを解きやすいようなテキストの区切り方は不明****【貢献】 AIに扱いやすいことばの単位を発見できる**

日本語や英語のように、わたしたちが普段使う言語のことを自然言語と呼ぶ。そして、自然言語をコンピュータで処理し、作業を効率化したり有益な知見を発見したりする技術を自然言語処理と呼ぶ。たとえば、メールやコメントがスパムであるかを分類する技術や、日本語のテキストを英語に自動で翻訳する技術などが、一般的に馴染み深い自然言語処理の応用例である。

近年の自然言語処理では、ニューラルネットワークで構築したAIにテキストを入力し、分類や翻訳などの結果を出力させる。たとえば、メールのテキストをAIに入力してスパムメールであるかの分類結果を出力させたり、日本語のテキストをAIに

入力して英訳テキストを出力させたりといった具合である。本記事では、テキストの分類や翻訳のように、特定のタスクを解くための仕組みをまとめて「AI」と呼ぶことにする。

多くの自然言語処理のシステムでは、テキストを適当な単位に区切ってからAIに入力する。たとえば「外国人参政権」というテキストを「外国人／参政／権」のような小さい単位（トークン）に区切ってからAIへと入力する。テキストの区切り方は1つではなく、「外国／人／参政／権」や「外国人／参政権」のようなパターンもあり得る。どのようにテキストを区切るかによってAIの性能は変わるため、区切り方は慎重に選択する必要がある。特に日本語のように、スペース記号で単語の境界を明示しない自然言語では、区切り方のパターンが膨大になるため、区切り方の選択は重要な研究テーマになる。

テキストの適切な区切り方は、AIが解くタスクによって異なる。たとえば、テキストの内容が政治的であるかを分類するようなタスクを解くのであれば、政治にかかわる「外国」や「参政権」といったトークンにアクセスしやすい「外国／人／参政権」という区切り方が適切かもしれない。一方で、食物の名前を含むテキストを検索したいという場合は、「外国／人／参／政権」のような区切り方が適切かもしれない。「かもしれない」とことばを濁しているのは、実際にこれらの区切り方を使ってAIを学習し、性能を評価してみるまで、その区切り方が適切かどうか分からないためである。

これまでの自然言語処理では、「新聞記事を扱うときは辞書に収録されたトークンを優先して使うとよい」というように、テキストの性質に応じた適切な区切り方の傾向が分かっている。しかし、大まかな傾向が分かっているだけであり、実際にテキストの区切り方を決める際には経験や職人の勘が必要である。

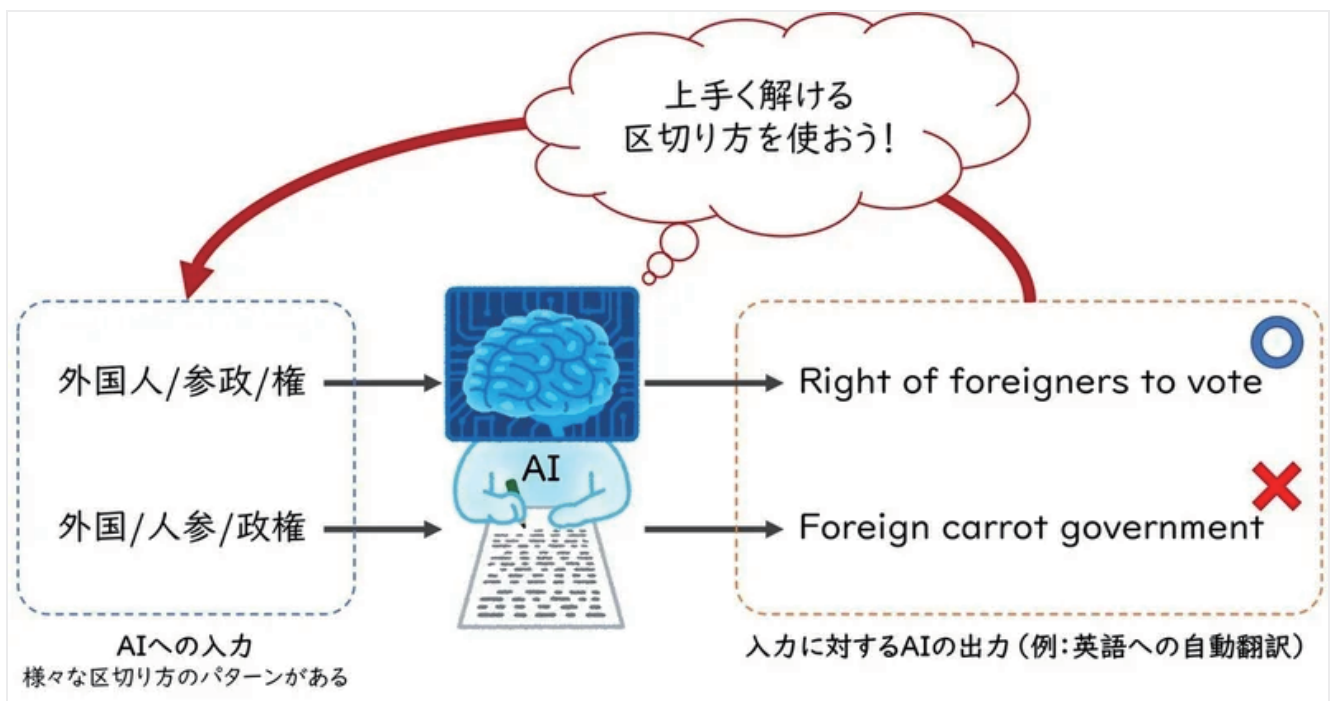
そこで本研究では、テキストやタスクの性質に応じて、AIがタスクを解きやすい適切なテキストの区切り方を自動で発見する手法を開発した。本研究によって、職人の勘がなくとも適切なテキストの区切り方を発見でき、AIの分類性能や翻訳性能の向上が得られる。たとえば日本語の分類タスクを解くときは、英語の場合よりも細かくテキストを区切ることでAIの性能が向上する。また、翻訳タスクを解くときは、翻訳される側の言語（日英翻訳の場合は日本語側）のテキストを細かく区切ることで性能の向上が得られることが分かった。このように本研究内容は、AIがタスクを解く上で扱いやすいことばの単位を調べることができる。そのため、本研究はAIの性能向上に寄与するだけでなく、AIがどのような単位でテキストを認識しているか（AIがどのように文字の世界を見ているのか）を知るための手がかりにもなる。

なお本論文の内容は英語で記述されているが、以下の国内学会論文および論文誌では日本語で詳細を確認できる。

- 平岡達也, 高瀬 翔, 内海 慶, 櫻 惇志, 岡崎直観: 後段モデルの損失値を

用いた単語分割のタスクへの最適化, 言語処理学会第27回年次大会 (NLP2021), pp.486–491 (2021年3月).

- 平岡達也, 高瀬 翔, 内海 慶, 樺 惇志, 岡崎直観: テキストベクトルの重みづけを用いたタスクに対する単語分割の最適化, 自然言語処理, 28(2):479–507 (2021年6月).
- 平岡達也, 高瀬 翔, 内海 慶, 樺 惇志, 岡崎直観: 単語分割と後段モデルの損失値を用いた同時最適化, 自然言語処理, 29(1):112–143 (2022年3月).



(2022年5月30日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（工学）

大学：東京工業大学

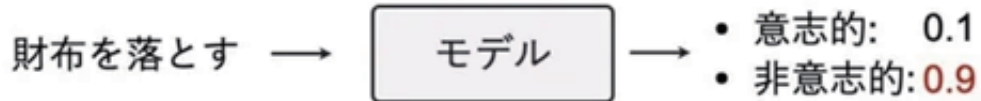
推薦文【メディア知能情報領域】 自然言語処理研究会

コンピュータがことばを扱う際には、単語等の細かい単位に区切り、ことばがどう組み合わさり何を伝えているかを解析します。この研究では、扱う問題に応じてこの区切り方を変える技術を開発し、問題を解く性能が向上することを明らかにしました。ことばの自動処理の基本的な問題を掘り下げた、実用性の高い重要な研究です。

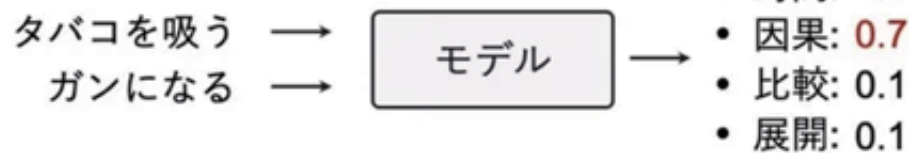
研究生活 研究を進めていく上で最も重要なことは、研究仲間との会話だと思っています。多くの先輩方が書かれているように、博士課程での研究は自分自身との戦いになります。そのため、（研究を応援してくれる人が周りにもかかわらず）研究がソロプレイであると錯覚し、行き詰まったときに誰にもSOSを出せなくなってしまうがちです。他者にSOSを出すというのは勇気が必要な行動であるため、日頃から研究仲間と会話を重ねることで、SOSを出す心理的ハードルを下げておくことが大切です。また、研究仲間の幅を広げておくことで、新しい視点からの

アドバイスが得られます。実際に、分野外からの斬新なコメントが本研究のピンチを何度も救ってくれました。同期や指導教員だけではなく、研究室の先輩や後輩、学会や勉強会で出会った他所属の仲間、授業で出会った他分野の仲間など、より広い範囲の仲間と関係を築くことで、あなたの研究はより楽しく、魅力的なものになると思います。

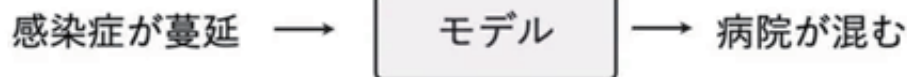
イベント分類



イベント間関係認識



イベント予測



Studies on Fundamental Problems in Event-Level Language Analysis



情報処理学会・学会誌「情報処理」
2022年8月15日 11:16



2021年度研究会推薦博士論文速報

[自然言語処理研究会]

清丸 寛一

(京都大学大学院 特定研究員)

邦訳：イベントレベルの言語解析における基礎的課題に関する研究

■キーワード

人工知能／自然言語処理／言語解析

【背景】 テキスト情報を整理・分析する基盤技術として言語解析が有用

【問題】 イベント（5W1H相当の情報単位）を処理単位とする言語解析は未熟

【貢献】 イベントレベルの言語解析における基礎的課題の解法を提案

テキスト情報を整理することはさまざまな価値を生み出す。商品・サービスのレビューを整理すれば、それは消費の判断基準となる。社会問題の因果関係に関する記述を整理すれば、それは解決の糸口となる知見や気づきに繋がる。アクセス可能なテキスト情報は日々増え続けており、テキスト情報を整理することはますます多くの場面で有用となっている。

テキスト情報を整理する上で重要な役割を果たすのが言語解析である。言語解析は、テキストを人が理解可能な意味単位——語や句、文など——の集まりとして

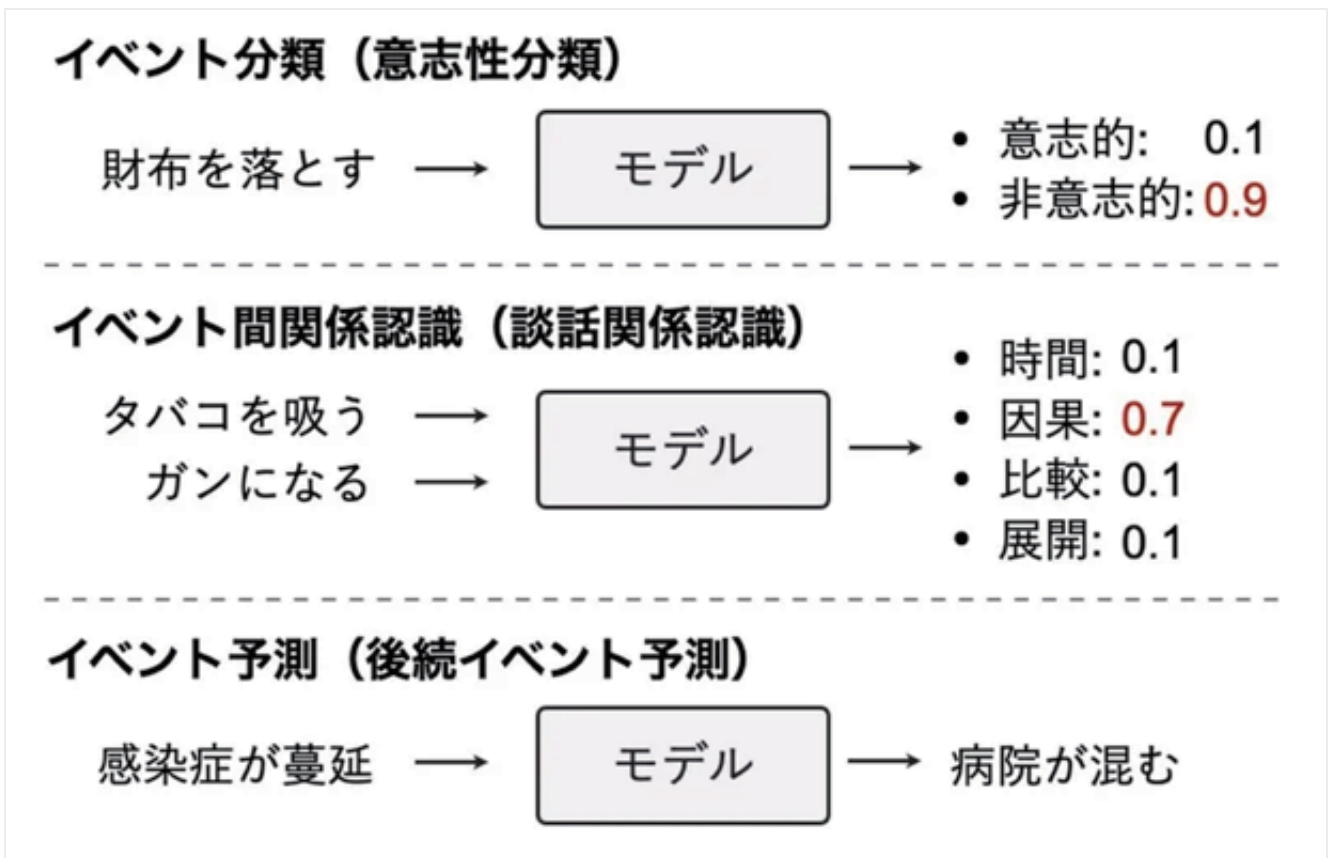
捉え、それらの意味と関係を解析するものである。

本研究では、状態や行為を表す意味単位、イベントを処理単位とする言語解析に取り組んでいる。イベントはおよそ言語学でいうところの節に対応する。たとえば「感染症が心配なのでマスクを購入した」という文は「感染症が心配だ」と「マスクを購入した」の2つのイベントからなる。イベントは実世界のできごとや人々の意見を整理するのに丁度良い意味単位であり、冒頭で述べた顧客フィードバックの分析や因果関係の整理といった応用はイベントレベルの言語解析によって実現できる。

イベントレベルの言語解析ではイベントの統計的性質を捉えることが求められる。イベントは複数の語句から構成され、そのパターン数は膨大であるため、素朴に表層表現に基づき頻度や共起を数えても意味のある結果は得られず、長らく単純な構造表現（述語と目的語だけを取り出したタプル等）として表現されてきた。しかし、ここ数年でニューラルネットワークに基づく言語処理技術が進展し、イベントを連続的なベクトル表現として上手く扱えるようになったことでこの問題は大きく緩和され、イベントレベルの言語解析に本格的に取り組むための下地が整った。

イベントレベルの言語解析はイベント分類、イベント間関係解析、イベント予測の3クラスに大別でき、本論文では、各クラスの基礎的課題である意志性分類、談話関係解析、後続イベント予測に取り組んでいる（[図参照](#)）。意志性分類はイベン

トが意志的な行為かどうか認識するタスク，談話関係解析はイベント間の意味的關係（因果関係など）を認識するタスク，後続イベント予測はあるイベントの次に起こるイベントを予測するタスクである．それぞれの詳細については参考文献 1) ， 2) ， 3) を参照されたい．



基礎的な言語解析は組合せによって真価を発揮する．たとえば，意志性分類単体で実現できる魅力的な応用はあまりない．しかし，感情極性分類と談話関係解析を組み合わせれば，冒頭の例文において「マスクを購入した」のは「感染症が心配だ」というネガティブな状況を理由にとった意志的な行為だと分かり，これらの情

報から両者が問題と解決策の関係にあると推測できる。このように基礎的な言語解析はその組合せによってテキスト情報を整理する軸を無数に生み出すことができる。その意味で、本研究はイベントレベルの言語解析に基づくテキスト分析に広く資するものとする。

■Webサイト／動画／アプリなどのURL

<https://hirokazukiyomaru.com>

参考文献

- 1) Kiyomaru, H. and Kurohashi, S. : Minimally-Supervised Joint Learning of Event Volitionality and Subject Animacy Classification, In Proceedings of the AAAI Conference on Artificial Intelligence (2022).
- 2) Kiyomaru, H. and Kurohashi, S. : Contextualized and Generalized Sentence Representations by Contrastive Self-Supervised Learning: A Case Study on Discourse Relation Analysis, In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (2021).
- 3) Kiyomaru, H., Omura, K., Murawaki, Y., Kawahara, D. and Kurohashi, S. : Diversity-aware Event Prediction based on a Conditional Variational Autoencoder with Reconstruction, In Proceedings of the First Workshop on Commonsense Inference in Natural Language Processing (2019).

(2022年5月31日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（情報学）

大学：京都大学

推薦文【メディア知能情報領域】 自然言語処理研究会

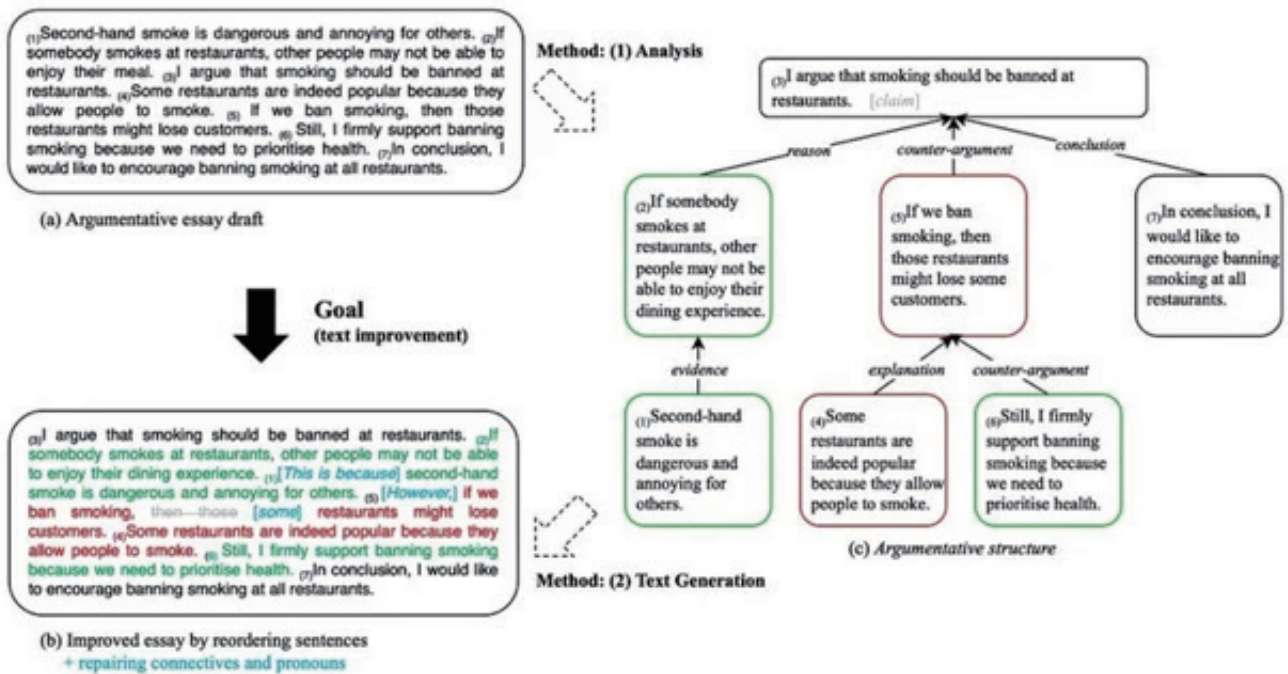
コンピュータのことばの処理はしばしば「字面」だけの表面的なものですが、この研究ではことばで語られる「できごと」に着目して、できごとの情報やできごと同士の間をことばから読み取る技術を開発し、ことばの解析性能が向上することを明らかにしました。ことばの「表現する」役割を深く考えた、とても重要な研究です。

研究生活 研究生活は楽しいものでした。特に楽しかったのは議論と雑談です。研究室には同じ興味を持ったトップクラスの専門家とその卵がいます。彼ら・彼女らとの議論・雑談はいつも学びや気付きに溢れていて、時間を忘れるほどに楽しいものでした。

苦勞したのは研究に行き詰まり、途方に暮れていた時期です。その時期は研究のサイクルを一通り経験したことで少し自信がついて、あまり人に相談をしていませんでした。しかし、いよいよ行き詰まって先生や先輩、後輩に相談すると、言語化の過程で思考が整理され、フィードバックも的確なので、それまでの停滞が嘘のように研究が進み始めました。議論や雑談に積極的になったのはそれからです。今では何も思いつかないときだけでなく、何か思いついたときですら近くの人を捕まえて雑談しています。

博士進学を悩んでいる方は近くの誰かに相談してみましょう。そこで恥ずかしかったり遠慮したりせずに相談する力は、研究生活にもきっと活きます。

最後に熱心にご指導を賜りました黒橋禎夫教授に心より感謝申し上げます。



Parsing Argumentative Structure in English-as-a-Foreign-Language Learner Essay



情報処理学会・学会誌「情報処理」
2022年8月15日 11:19



2021年度研究会推薦博士論文速報

[自然言語処理研究会]

Jan Wira Gotama Putra

(スマートニュース (株) ソフトウェアエンジニア)

邦訳：英語学習者エッセイの議論構造解析

■キーワード

言語学習支援／議論構造解析／テキストの一貫性

【背景】 英語学習者のための作文支援システム

【問題】 多くの作文支援システムはスペルミスや文法誤りなどの表層的な誤りにしか対応できない

【貢献】 テキストの議論構造を分析して、それに基づく適切な文の並べ方の助言する

Intelligent Argumentative Writing Assistant

邦訳：知能的論説作文支援システム


深層学習 (Deep learning) の有効性が幅広い分野で実証され、人工知能の応用範囲はますます広がっている。なかでも機械に人間の言葉を理解させることを目指

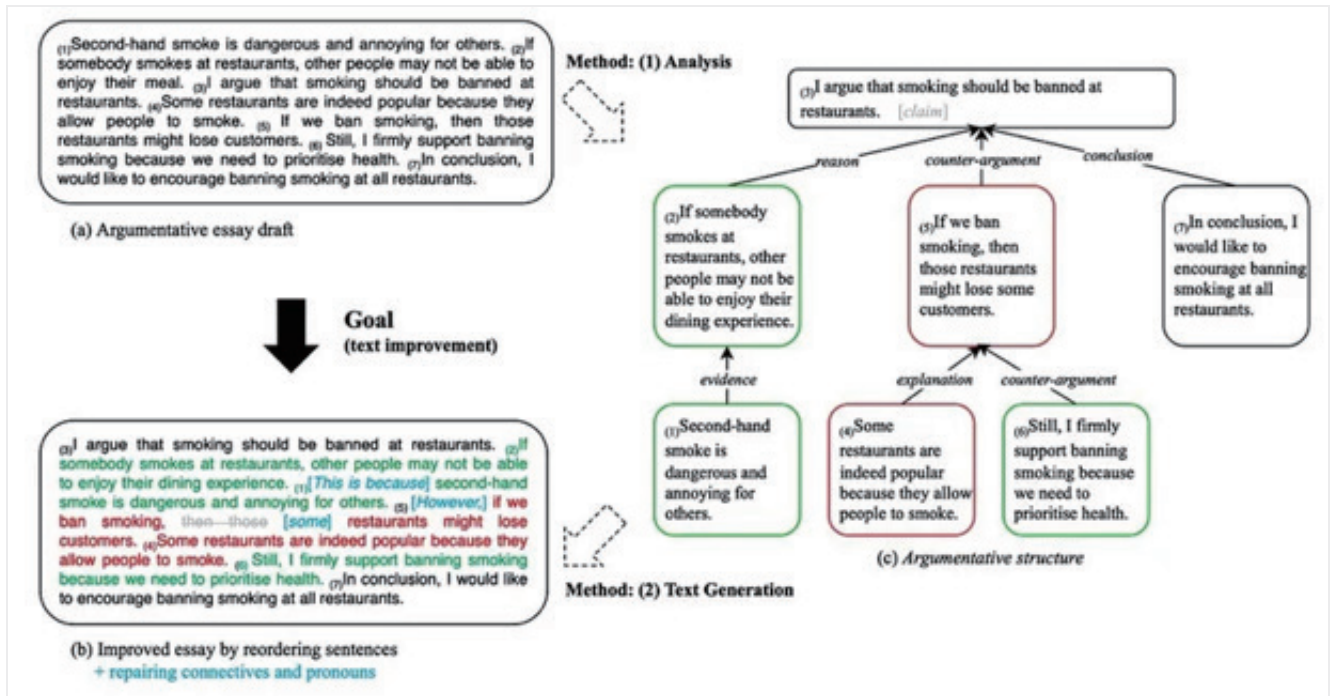
す自然言語処理の研究への期待は高まっている。その応用のひとつとして言語学習者 (language learner) の支援が研究されている。具体的には、人工知能技術によって学習者が書いたテキストを採点したり、テキスト中のスペルミスや文法誤りを指摘・訂正することで、教師の負担を減らすことが期待されている。また、言語学習者にとってもリアルタイムで適切なフィードバックを得ることによって効率的な学習が可能となる。

言語学習支援において主に研究されているのは、スペルミス、前置詞の誤用、主語と動詞の一致の間違いなどのテキストの表層的な誤りの指摘や訂正にとどまっている⁵⁾。テキスト全体の構成や意味内容まで踏み込んで修正を助言するような研究はほとんどない。この研究では、英語を対象とし、学習者が書いたテキストに対して、文を並べ替えることによって、テキストの構成をよりよくすることを目指している。

図に示すように、英語の議論的エッセイは、(1) 序論 (introduction)、(2) 主張 (claim)、(3) (主張を支える) 理由・根拠 (reason, evidence)、反論 (counter argument)、(5) 結論 (conclusion) という構成になっていることが多い。これは英語の例であるが、テキストの構成は母語言語によって異なる場合がある¹⁾。たとえば、日本語では背景から始めて最後に主張を置くことも珍しくない。英語を母語としない学習者が英語のテキストを書く場合は、英語テキストで一般的な議論の展開に沿って文を並べ替える必要がある。文を適切に並べ替えるこ

とによって、英語母語話者にも理解しやすい一貫性のあるテキストが実現できる
1) , 2) .

この研究では、学習者の書いたテキストを解析して議論構造（議論全体の流れ・
理路のマップ、例：c) を抽出し³⁾ , ⁴⁾ , 議論構造中の各文の役割を考慮して
文を並び替える手法を提案している。最初に議論構造を解析することにより、主張
であると同定した文はテキストの先頭に移動し、その根拠は主張に続けるなどの整
列を学習者に助言すること可能になる。また、教師の立場からすると、学習者の書
いたテキストの議論構造を可視化し、分析することによって学習者に適切な説明・
指導をすることが可能になる⁶⁾ . この研究をさらに発展させることにより、将来
的には、自動化による学習者ひとりひとりに適合した学習環境を実現し、言語学習
がより楽しく効率的になることが期待できる。



■Webサイト／動画／アプリなどのURL

• <https://www.americanprogress.org/article/future-testing-education-artificial-intelligence/#:~:text=challenges%20of%20AI-,Artificial%20intelligence%20can%20help%20students%20learn%20better%20and%20faster%20when,th,e%20naked%20eye%20cannot%20see.>

• 本研究で開発されたソフトウェア

<https://sites.google.com/view/tokyotechcl-tiara/>

• 招待公演, インドネシア語のみ

<https://www.youtube.com/watch?v=41grGqgJUM8>

• 雑誌・学会発表

<https://wiragotama.github.io/publication.html>

参考文献

- 1) Putra, J. W. G., Teufel, S. and Tokunaga, T. : Annotating Argumentative Structure in English-as-a-Foreign-Language Learner Essays, In Natural Language Engineering, pp.1-27 (2021).
- 2) Putra, J. W. G., Matsumura, K., Teufel, S. and Tokunaga, T. : TIARA 2.0: An Interactive Tool for Annotating Discourse Structure and Text Improvement, In Language Resources and Evaluation, pp.1-25 (2021).
- 3) Putra, J. W. G., Teufel, S. and Tokunaga, T. : Multi-Task and Multi-Corpora Training Strategies to Enhance Argumentative Sentence Linking Performance, In Proceedings of the Eighth Workshop on Argument Mining, Association for Computational Linguistics, pp.12-23 (2021).
- 4) Putra, J. W. G., Teufel, S. and Tokunaga, T. : Parsing Argumentative Structure in English-as-a-Foreign-Language Learner Essays, In Proceedings of the Sixteenth Workshop on Innovative Use of NLP for Building Educational Applications, Association for Computational Linguistics, pp.97-109 (2021) .
- 5) Strobl, C., Aihaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A. and Rapp, C. : Digital Support for Academic Writing: A Review of Technologies and Pedagogies, Computers & Education, 131, pp.33-48 (2019).
- 6) Cullen, S., Fan, J., Brugge, E. van der and Elga, A. : Improving Analytical

Reasoning and Argument Understanding: A Quasi-experimental Field Study of Argument Visualization, NPJ Science of Learning, 3(21), pp.1-6 (2018).

(2022年5月11日受付)

(2022年8月15日note公開)

取得年月日：2021年9月

学位種別：博士（工学）

大学：東京工業大学

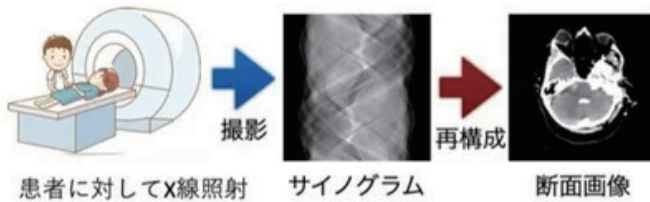
推薦文【メディア知能情報領域】自然言語処理研究会

小論文等の文章では議論の流れや構造が重要で、書くのも読むのも簡単ではありません。この研究では、英語を母語としない人が英語で書いた小論文を対象に、データセットを作成して議論の解析や順序の提案をする技術を開発し、有効性を明らかにしました。整った文章を書くサポートをするという、実用性の高い重要な研究です。

研究生活 私は小さな田舎町で生まれ育てました。大都市に比べると勉強の困難がたくさん経験し、「将来に子供たちが皆な質の高い教育を受けられるようにしたい」という関心を高めました。こうして、教育に関する研究テーマを取り組んでいました。博士研究員として専門知識はもちろん「life skill」も必要となる。最先端の論文や本を読むため英語能力が重要となる。研究にも実験の失敗も多いし、自分の生活が壊さないようにメンタルを強化しなければならない。留学生の私には日本語や日本の生活にも慣れしむ必要がある。つまり、知識のみじゃ足りません。幸いなことは、指導教員の徳永先生やSimone先生、ラボメンバ、多くの方に色々な意味で支えてもらって、博士号が取得できました。お世話になった皆様方に大変感謝いたします。

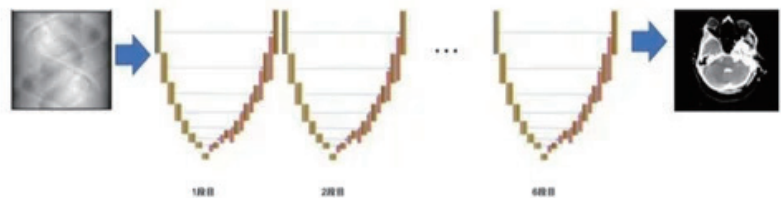
もし博士号の経験から一番重要なものをピックアップすれば「謙虚」のことだと考えます。「僕はまだまだだ」、「実際何も分からない」という経験が多かったです。こうして、多くの質問することが必要となる。質問（問題探し）がなければ、答え（発明）も出ませんのでぜひ質問してください。

CT画像再構成処理とは



提案手法

CT画像再構成にU-Netを多段(6段)で適用



Stacked U-NetによるCT再構成

♡ 1

情報処理学会・学会誌「情報処理」
2022年8月15日 11:20



2021年度研究会推薦博士論文速報

[知能システム研究会]

水澤 悟

■キーワード**Deeplearning/U-Net/CT再構成****【背景】** 患者をX線で撮影し、体の断面を再構成し可視化するCT再構成**【問題】** CT再構成の速度向上, ノイズ低減, 低被ばく化のバランス**【貢献】** Stacked U-Netによる高速なかつ高精度なCT再構成の実現

X線CT画像再構成は、X線検出器をさまざまな角度で対象物に照射して撮影したサイノグラム画像から、断面画像を再構成する処理のことである。検出器に観測されたX線の強度は、対象物の性質に応じて減衰する。減衰の程度は、X線の経路上にある物体の特性に依存する。たとえば、X線の通り道に骨や臓器があると、検出器で検出されるX線の強度はそれに応じて減衰する。つまり検出器は、X線の通り道にある物体のX線吸収の結果を積算して観測する。このようにして得られたサイノグラム画像をもとに、X線CT画像再構成により、対象物の分布を表す断面画像を得ることができる。

X線CT画像再構成では、再構成速度が早ければ素早く診断ができる、再構成精度が良くノイズが少なければ、疾病を見逃しづらくなる、X線照射数が減れば患者の負担が減る、といった点がある。そのため再構成速度を早くし、精度を良くし、X線照射量を減らす研究が行われている。

X線CT画像再構成には、大きく分けて「直接法」と「逐次近似法」の2つの方法がある。直接法は計算コストが低いが、再構成時の撮影角度が不足すると、得られる画像のノイズやアーチファクトが増えてしまう。逐次近似法は低ノイズ、低アーチファクトの結果を得ることができるが、計算コストが直接法に比べ高い。

上記直接法や逐次近似法の課題を克服するために、ディープラーニング手法と組み合わせたCTの再構成手法が提案されている。それはディープラーニングにより直接法や逐次近似法の課題をカバーする手法だが、既存手法に加えてディープラーニングの計算を実施するという点において、既存手法よりも再構成の速度が遅くなってしまうのが課題である。またディープラーニングの医療適用上の課題として患者のプライバシーの問題から学習用画像が少ないという課題がある。ディープラーニングは単に画像を多くするだけで精度が上がっていくという利点がある一方、学習用画像がすくなくと精度を上げることができず、既存手法に劣ってしまう。

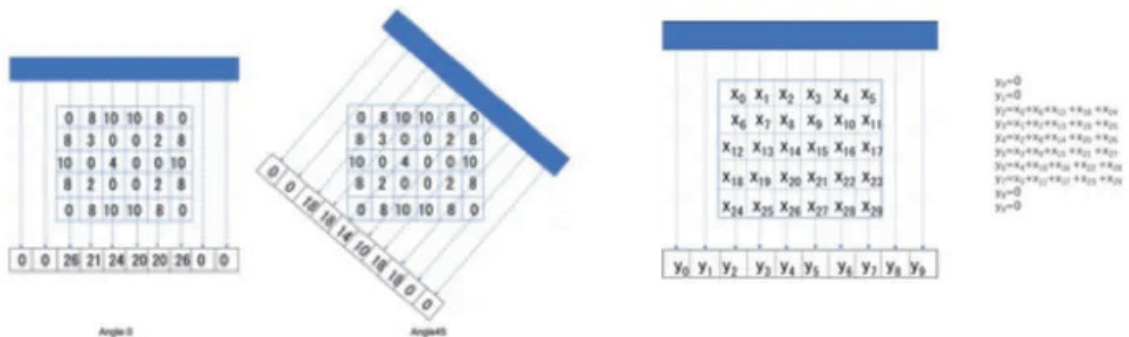
本研究ではディープラーニングのモデル (Stacked U-Net) だけで直接再構成画像を復元できることを示した。ディープラーニングは線形演算の組合せ1回で解を出すため、繰り返し演算を必要とする従来手法に比べて速度が速い。またディープラーニングの医療適用上の課題である患者のプライバシーの問題から学習用画像が少ないという課題に対しては多くデータが存在する自然画像を用いて学習することを提案した。自然画像で学習を行ったモデルにおいても、医療用画像が復元できる

ことを示し、X線照射量が少ない場合において、定量手法評価や診断用AIを用いた評価により既存手法より優れた結果を示した。

本研究で示した手法により、少ないX線被ばく量で迅速に診断ができるCT機器の開発が期待できる。

CT画像再構成処理とは

- ・撮影：対象の周りを回転し、X線により対象の積算情報を得ること



- ・再構成とは：積算情報を計算し対象の情報を得ること

CT再構成処理の難しさ：積算されているので陽にとくことがむずかしい

(2022年5月30日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（工学）

大学：電気通信大学

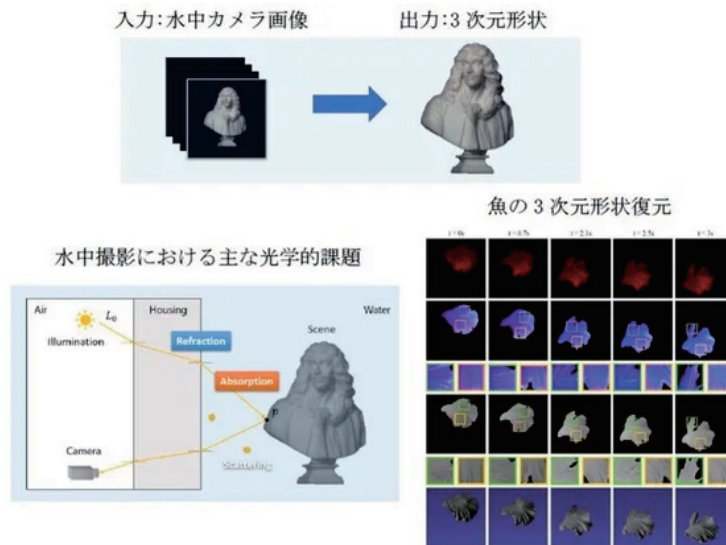
推薦文【メディア知能情報領域】知能システム研究会

X線CT検査では、X線の投影データを基に人体断面の再構成を行う必要がある。本論文では、少ないX線照射数で短時間かつ高精度に再構成を行う、深層学習モデルを提案した。この成果により、少ない被ばく量で迅速に診断ができるCT機器の開発が期待できる。

研究生生活 入学試験のときに「テーマがこの状態で進学決めるなんて勇気あるね」と言われたのを思い出します。私は社会人博士として入学し、結局4年かかって博士を取得しました。社会人として働きながら、家庭環境の変化もあり（子供が生まれた！）時間をやりくりしながら取り組みましたが、最初の3年はほぼ成果らしい成果もなく、悶々としていました。4年目に成果がドドドっと出て博士を無事

取得でき、"石の上にも三年"という言葉の正しさを知ることになりました。

上記を鑑みると、苦しい博士生活だったかのように思えますが、私としては結構楽しんでいました。一つひとつ自分の考えた課題について手を動かすと、さらに新しい課題や解決策が浮かんできます。その課題や解決策に対してまた取り組むというサイクルを繰り返すと着実に進んでいるのが実感でき、楽しくなります。またその結果が論文や学会などで認められるともうちょっと楽しくなります。博士進学を考えている方にはぜひ、"研究すること自体を楽しんで"いただければと思います。



Refraction and Absorption for Underwater Shape Recovery



情報処理学会・学会誌「情報処理」
2022年8月15日 11:23



2021年度研究会推薦博士論文速報

[コンピュータビジョンとイメージメディア研究会]

Meng-Yu Jennifer Kuo

(Postdoctoral Researcher, University of Minnesota Twin Cities)

邦訳：屈折と吸収のモデル化による水中物体の3次元形状復元

■キーワード

コンピュータビジョン／視覚情報処理／水中3次元形状復元

【背景】 光学系の工夫でコンピュータの視覚はさらに拡張可能**【問題】** 水中での光の屈折，吸収，散乱を扱える計算モデルが必要**【貢献】** 水中物体の詳細な表面形状をカメラ画像から復元

コンピュータビジョンを人間と同等，またはより優れた視覚様式として実現することは，人工知能にとってきわめて重要であり，人間の視覚的知能の理解にも貢献すると考えられる。コンピュータビジョン分野において，実物体の観測からその3次元形状を復元することは主要な研究テーマである。これまでも，さまざまな視覚的手がかりを利用した3次元形状復元法が数多く発表されており，ロボット工学，拡張現実をはじめとした応用に大きく貢献している。しかしながら，ほとんどの研究は，空气中環境での3次元形状復元のためのアルゴリズム設計にのみ焦点を当てている。

本研究は，環境のメディアとして「水」に着目する。水は身近なメディアであることに加え，地球表面の約70%は水で覆われている。水中物体の3次元形状を推定することは，水中生物の生物学的研究，内視鏡や胃カメラ等による内臓モデリング，水中インフラ（ダム壁，発電所のパイプ，貯水池など）の3Dマッピングや検査など，幅広い分野での応用に大きな関心を集めている。

水中の物体を観察する場合，透過光は主に屈折，吸収，散乱の影響を受ける。このような光学現象を明示的にモデル化することで，透過光が受ける影

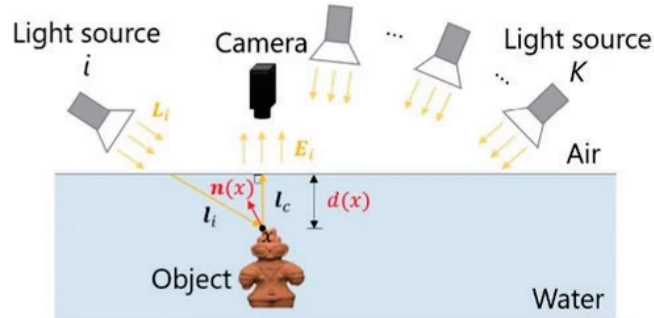
響は水中物体の形状を明らかにする手がかりとして考えることができる。水中物体の形状を復元するために従来の技術を直接適用すると、光の屈折による画像の歪みから生じる誤差が避けられない。また、水中では光の吸収や散乱により視認性が低下する。そこで私は、水中景観や物体を画像から理解するための理論的基盤の構築と効率的な計算手法の実装を行い、光学現象の影響を受けた画像からその背後にある物体の形状を明らかにする、つまり「見えることの向こう側を見る」ことを研究テーマとしている。

水の特性は課題をもたらすだけでなく、解決の一端を担っている。水の透過率は距離と波長の両方に依存し、水は近赤外波長域の光を大きく吸収する。本研究はこの効果を利用し、パッシブな水中3Dセンシングのための3つの方法を提案した。第一の手法Surface Normals and Shape From Water¹⁾は(図-1)、水中物体のピクセル単位の高密度な表面法線と形状を、互いに人為的な制約(閾値, 滑らかさ)なしに推定し、従来の手法では困難であった複雑な表面形状の復元を可能にするものである。2つ目の手法であるNon-Rigid Shape From Water²⁾は(図-2)、動的な非剛体物体の全体的かつ統合された形状を復元する。水中物体の変形や移動に伴い、異なる時間での形状を復元・統合・精緻化することにより、時間の経過とともにより緻密で完全な対象物の復元が可能となる。そして、最後のDense Underwater SLAMは(図-3)、水中の剛体の法線と形状を含む詳細な表面形状を復元する手法である。本研究ではこれらの手法の理論を提案するだけでなく、水中の実世界の物体に適用し、実験結果によって理論の有効性を実証した。

これら3つのアプローチにより、幅広い領域において実世界の水中物体やシーンの3次元解析の可能性が広がり、実用的な受動的な水中3Dセンシング手法の確立に大きく近づいたと考えられる。

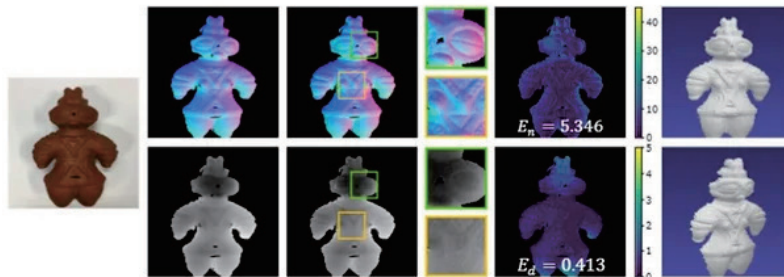
Surface Normals and Shape From Water

Near-Infrared Multi-Wavelength Imaging

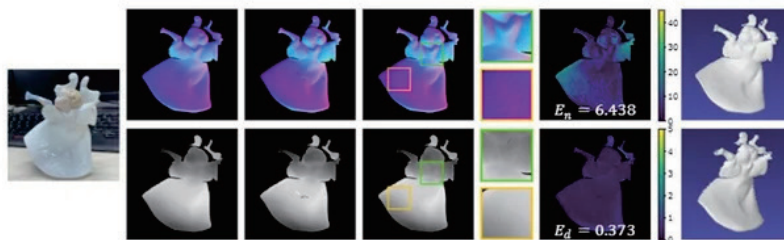


Experimental Results

Static Lambertian objects.



Static non-Lambertian objects.



Video-rate reconstruction of a swimming goldfish.

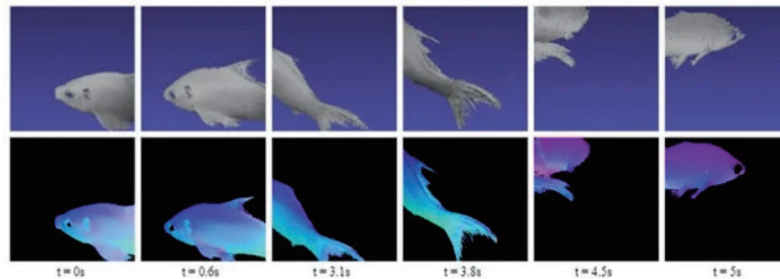
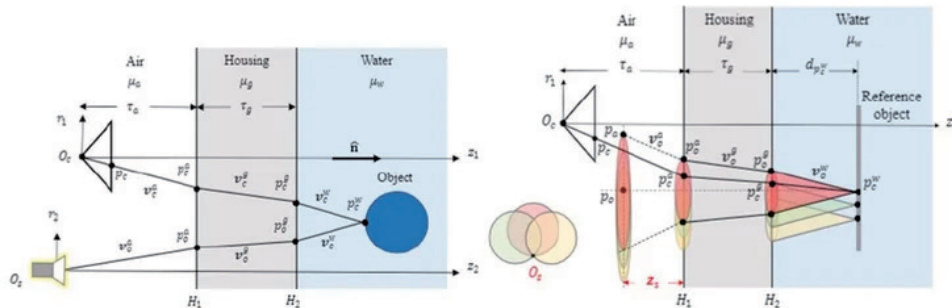


図-1

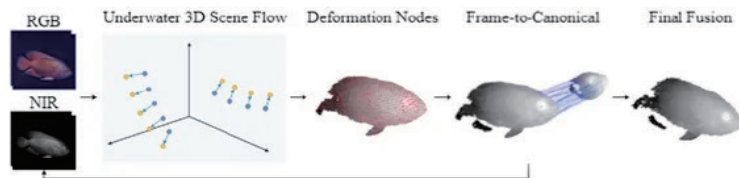
Non-Rigid Shape From Water

Perspective Shape From Water

We derive a practical calibration method that estimates the point source position with respect to the camera by using the absorption difference of two near-infrared wavelengths along the same refractive underwater light path.



Non-Rigid Shape From Water



Experimental Results

Results of swimming fish in various real-water samples that were taken from an actual lake, river, and ocean.

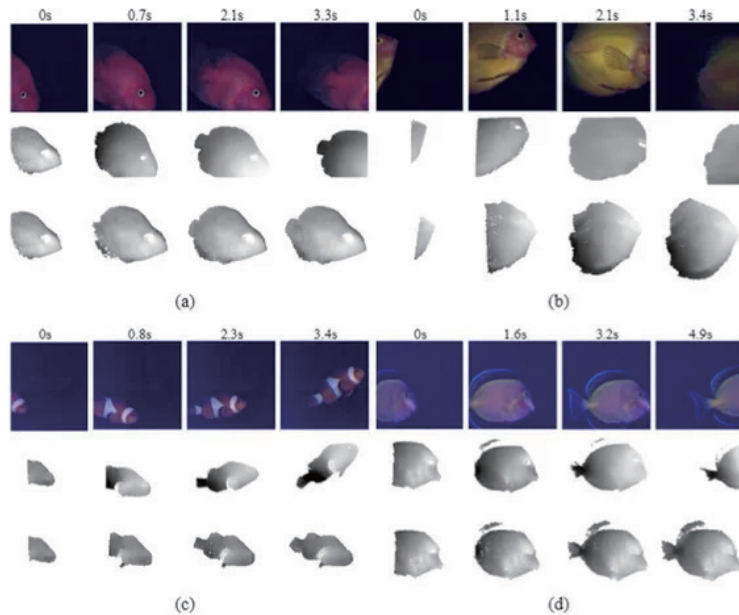


図-2

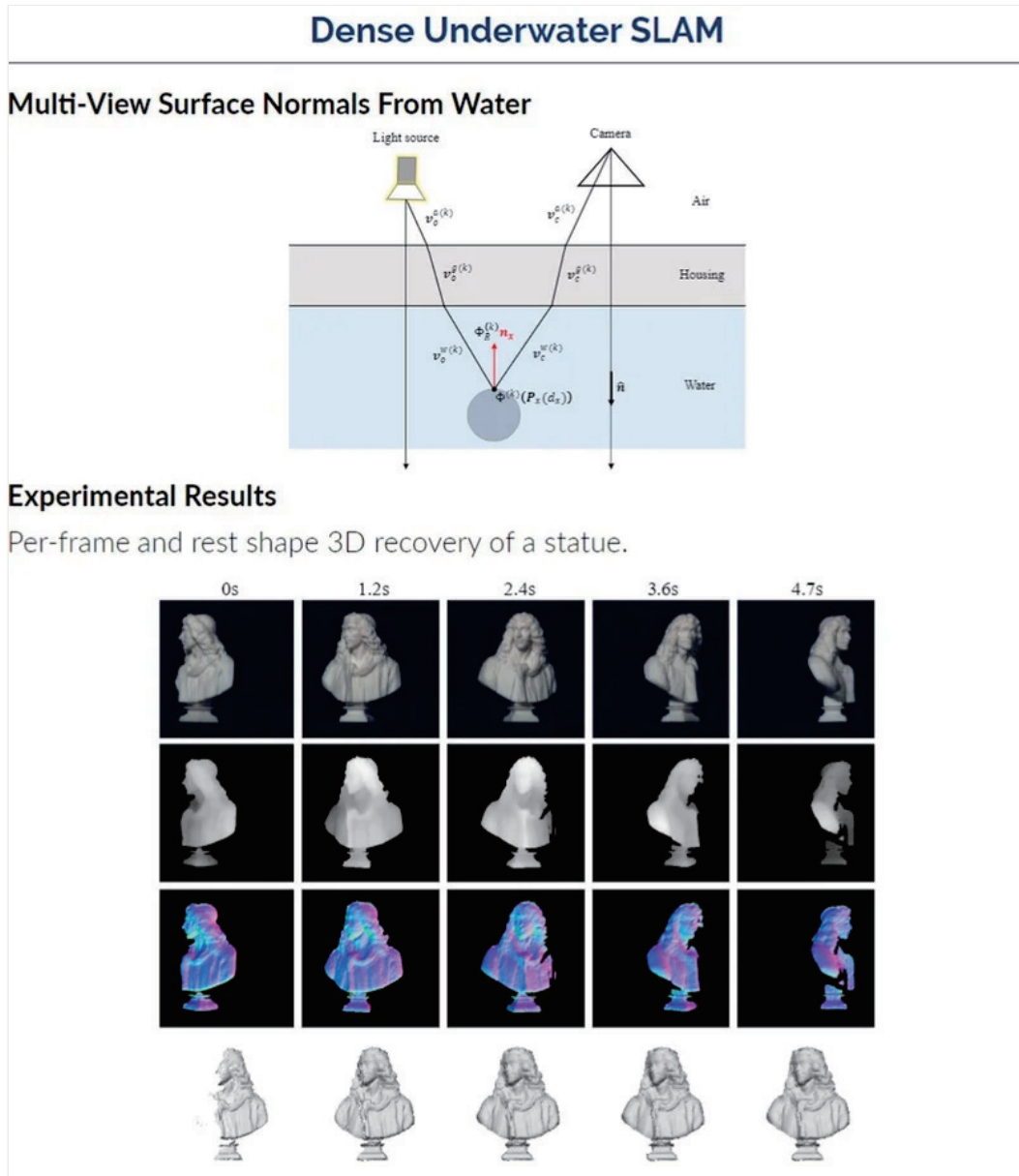


図-3

■Webサイト／動画／アプリなどのURL

<https://myjk1992.github.io/>

参考文献

- 1) Meng-Yu Jennifer, K., Murai, S., Kawahara, R., Nobuhara, S. and Nishino, K. : Surface Normals and Shape From Water, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021).
- 2) Meng-Yu Jennifer, K., Kawahara, R., Nobuhara, S. and Nishino, K. : Non-rigid Shape from Water, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, No.7 (2021): 2220-2232.

(2022年5月30日受付)

(2022年8月15日note公開)

取得年月日：2021年9月

学位種別：博士（情報学）

大学：京都大学

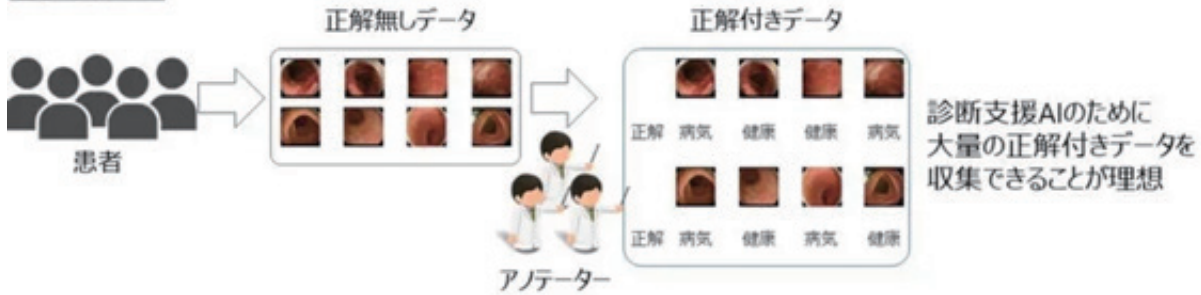
**推薦文 [メディア知能情報領域] コンピュータビジョンとイメージメディア
研究会**

本論文は、水中物体の3次元形状計測において、異なる波長による観測がもたらす輝度の変化から奥行きと法線方向を画素単位で計測する手法を提案するものであり、被写体の動きに伴う形状変化をもあわせて推定することで高精度な全周囲3次元形状計測が可能であることを実証している。一連の研究

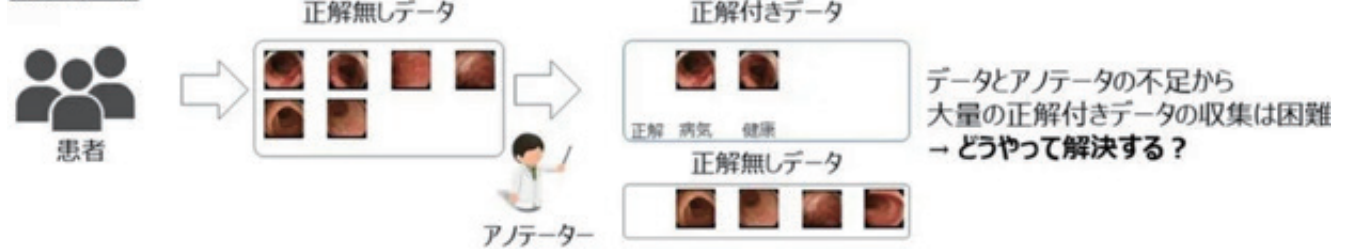
成果はトップカンファレンス・ジャーナルに採録されており、推薦に値する。

研究生生活 振り返ってみると、日本での大学院生活がこれほどまでに素晴らしい旅になるとは思ってもみませんでした。同僚、友人、家族から受けたすべてのサポートに心から感謝しています。大学院生活で、水中に沈んだターゲットを正確に計測する方法について理解が深まりました。しかし、世界の複雑さ故に、私たちには学ぶべきこと、探求すべきことがたくさん残されています。

理想的なケース



実際のケース



Bio-Medical Data Classification Approaches with Limited Annotation



情報処理学会・学会誌「情報処理」

2022年8月15日 11:30



2021年度研究会推薦博士論文速報

[コンピュータビジョンとイメージメディア研究会]

原田 翔太

(九州大学 大学院システム情報科学研究所 情報知能工学部門 /

日本学術振興会 特別研究員 (PD))

邦訳：アノテーションが限定された生体・医学データのための分類手法の提案

■キーワード

医用データ分類 / 機械学習 / 正解付きデータ不足

【背景】 高性能な自動診断支援AIの実現には大量の「正解」付きデータが必要

【問題】 医師の協力が必要なため、大量の正解付き医用データの収集は困難

【貢献】 正解付きデータ収集の手間を減らす3つの方法を提案

AIとも呼ばれて注目されている深層学習技術を活用することで、医用画像（レントゲン画像や内視鏡画像）や生体信号（脳波や心電図）といった医用データを用いた診断を自動かつ高精度に行える診断支援AIの構築が実現可能となっている。診断支援AIの導入は医師の誤診断の防止や医師不足の解消などにつながるため、非常に重要である。

高精度な診断支援AIを構築するためには、「この画像の患者は健康、この画像の

患者は病気です」といったような正解が付いたデータを大量に用意する必要あるが、それはとても困難である。たとえば、「この画像には犬（もしくは猫）が写っている」という判断は誰もが容易に行える。そのため、犬と猫を分類するAIを学習するための正解付き画像を大量に用意することも容易である。一方、診断支援AIの場合、健康もしくは病気の患者のデータを大量に収集すること、それらに正解を与えることは医師にしかできない。そのため、一般的な手段で診断支援AIのための大量の正解付きデータを用意することは困難である。そこで本研究では、診断支援AIのための正解付きデータを用意する手間を減らす3つの方法を提案した。

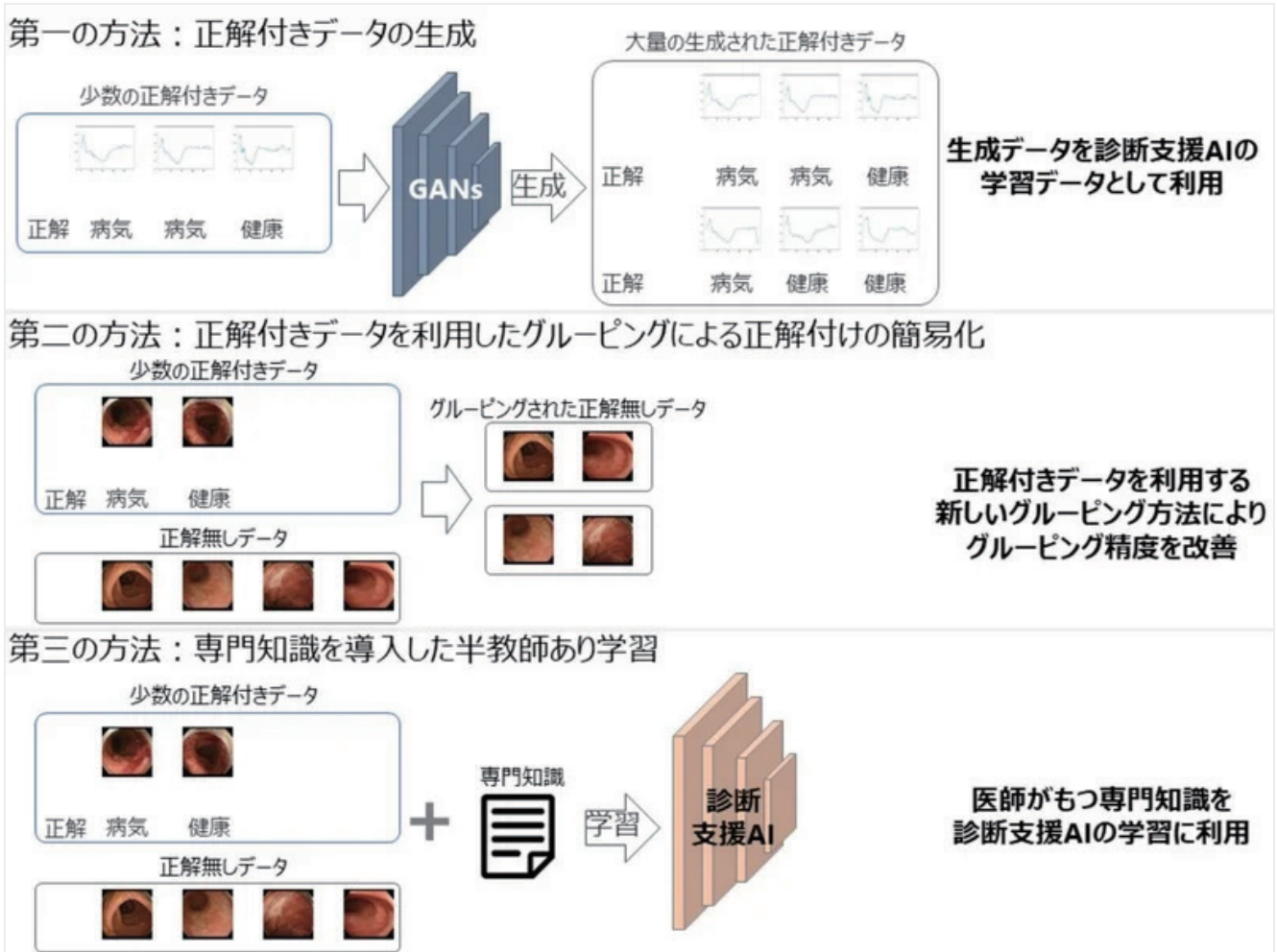
第一の方法は、少数の本物の正解付きデータから「新しい正解付きデータ」を生成することで、正解付きデータを用意する手間を削減する。本研究では、深層学習技術の1つであるGenerative Adversarial Networks (GANs) に着目した。GANs は本物のデータに似た新しいデータを生成することができるフレームワークで、画像生成の分野で多くの成果が報告されている。本研究ではGANsを生体信号生成に応用することで、脳波や心電図の少ない正解付きデータから新たな正解付きデータを生成することに成功した¹⁾。

第二の方法では、正解が付いていないデータをグルーピングし、そのグループごとに正解を与えることで、正解付きデータを用意する手間を削減する。ここで重要となるのは、正解を知らない状態で同じ正解を持つデータの集まったグループをどのように作るのかである。本研究では正解の付いていないデータのグルーピングに

少数の正解付きデータの情報と医師の専門知識を利用する新しいグルーピング方法を提案した²⁾。

第三の方法では、診断支援AIのための半教師あり学習手法（少数の正解付きデータと大量の正解が付いていないデータでAIを学習する方法）を提案し、大量の正解付きデータを用意する必要をなくす。これまでも多くの半教師あり学習手法が提案されているが、それらの多くは医用診断AIのための手法ではなかった。本研究では、内視鏡画像からの潰瘍性大腸炎診断を実現するために、医師の専門知識を導入した半教師あり学習手法を提案した。その結果、既存手法と比較して優れた診断支援AIを構築できた³⁾。

本研究では最先端の深層学習技術の応用や、医師の専門知識を活用した新たな方法の提案することで、診断支援AIのための正解付きデータを用意する手間を削減することに成功しました。



参考文献

- 1) Harada, S., Hayashi, H. and Uchida, S. : Biosignal Generation and Latent Variable Analysis with Recurrent Generative Adversarial Networks, IEEE Access (Impact Factor = 4.098), Vol.7, pp.144292-144302 (2019).
- 2) Harada, S., Bise, R., Hayashi, H., Tanaka, K. and Uchida, S. : Soft and Self Constrained Clustering for Group-Based Labeling, Medical Image Analysis, Vol.72 (2021).

3) Harada, S., Bise, R., Hayashi, H., Tanaka, K. and Uchida, S. : Order-Guided Disentangled Representation Learning for Ulcerative Colitis Classification with Limited Labels, Proceedings of the 24nd International Conference on Medical Image Computing and Computer Assisted Intervention (2021).

(2022年5月25日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（工学）

大学：九州大学

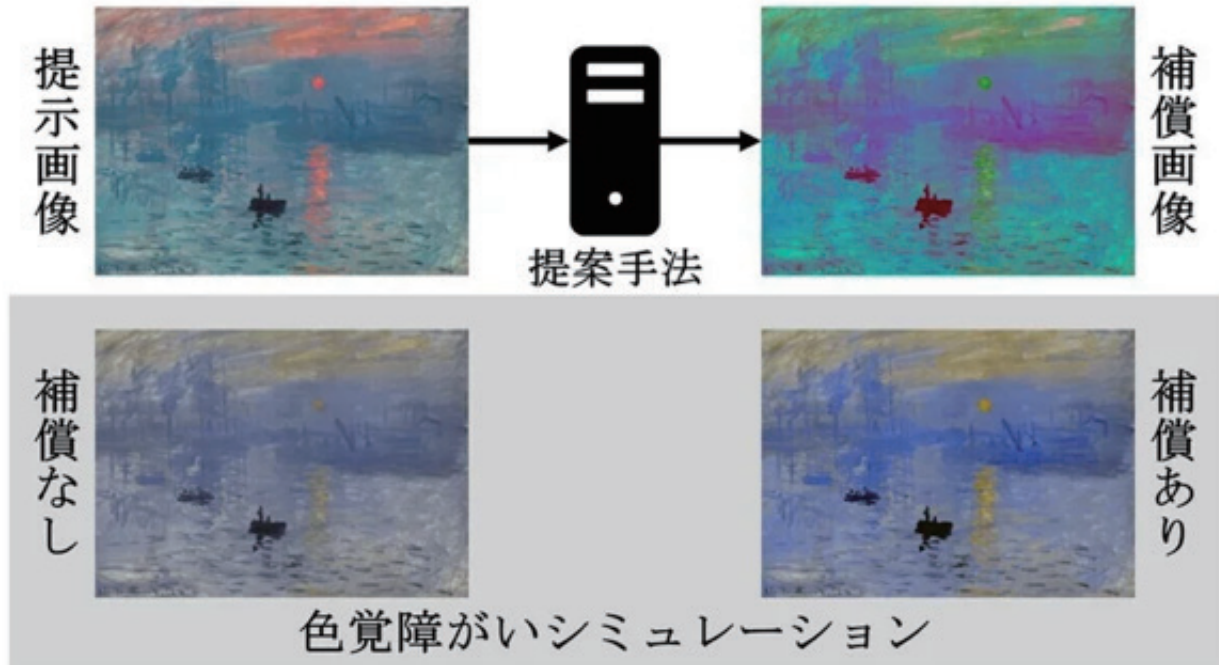
正会員

推薦文【メディア知能情報領域】 コンピュータビジョンとイメージメディア研究会
深層学習はその高い性能でAIの進展を支えています。しかし、実際には「正解」付きのデータが大量に必要という問題があります。実際、画像を用いた医療診断の場合、忙しい医師に「この画像は病気、これは健康」と「正解」をいちいち付けてもらうのは大変です。この論文では正解付けの手間を減らす3つの方法を提案してい

ます。

研究生活 修士・博士課程では生物学や医学を専門とする研究者の方との共同研究に携わる機会を数多くいただきました。いわゆる学際的研究と呼ばれるものです。この学際的研究を進めていく上で最も重要かつ難しい問題は、各研究者のバックボーンがまったく異なるという点です。専門用語1つをとっても、分野が異なるとまったく違う意味として伝わる可能性があります。そのため、「相手が本当に伝えたいことは何か」、「相手に正しく自分の意図が伝わっているか」を常に意識する必要があります。このような意識を身に付けることは容易ではありませんが、社会で生きていく上では非常に重要な能力です。研究活動を通して、この意識を育む機会を得たことは今後の大きな糧になったと思います。

博士課程は研究遂行能力以外の能力も伸ばすことができる、多様な経験が詰める場所です。近年では、博士課程の学生を金銭的に援助する仕組みも増えてきているため、博士進学ハードルが下がってきているように感じます。もし少しでも博士進学に悩んでいるようでしたら、進学することをお勧めします。



Naturalness- and Contrast-Preserving Image Recoloring for Color Vision Deficiency



情報処理学会・学会誌「情報処理」
2022年8月15日 11:32



2021年度研究会推薦博士論文速報

[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]



朱 臻陽

(山梨大学工学部コンピュータ理工学科 助教)

邦訳：コントラスト強調および自然さ保存を考慮した色覚障がい支援用色変換技術に関する研究

■キーワード

色覚障がい支援／色変換／コントラスト強調

【背景】 色覚障がいによる色識別困難でコントラストが損失

【問題】 過度のコントラスト強調で画像の自然さが失う

【貢献】 自然さとコントラストの最適化モデルを提案

人間の網膜には3種類の錐体細胞が存在し、それぞれ赤、緑、青に対応する波長の光に最も反応する。これらの細胞の反応の強さによりさまざまな色を識別することができる。しかし、錐体細胞に異常が生じると、色を正しく識別できなくなり、色覚障がい（Color Vision Deficiency, CVD）が発生する。CVDのほとんどは遺伝性であり、根治の手段はまだ発見されていない。国内のCVD発生率は、男性が約5%、女性が約0.2%である。CVDによって色差が分かりにくくなり、コントラスト損失問題（例：黒い背景から赤い文字が読み取れない、周囲とのコミュニケーションが困難）が発生する。CVDは患者の日常生活に支障を来していて、Quality of Life (QoL) を低下させている。市販されている色覚補償用偏光メガネでは、一部の波長の光を遮り、利用者の色差認識能力を向上させるが、その効果は限定的である。一方で、高機能の携帯端末（スマートフォン、タブレット等）の普及に伴い、CVD支援を目指した画像処理技術がここ数年注目されている。

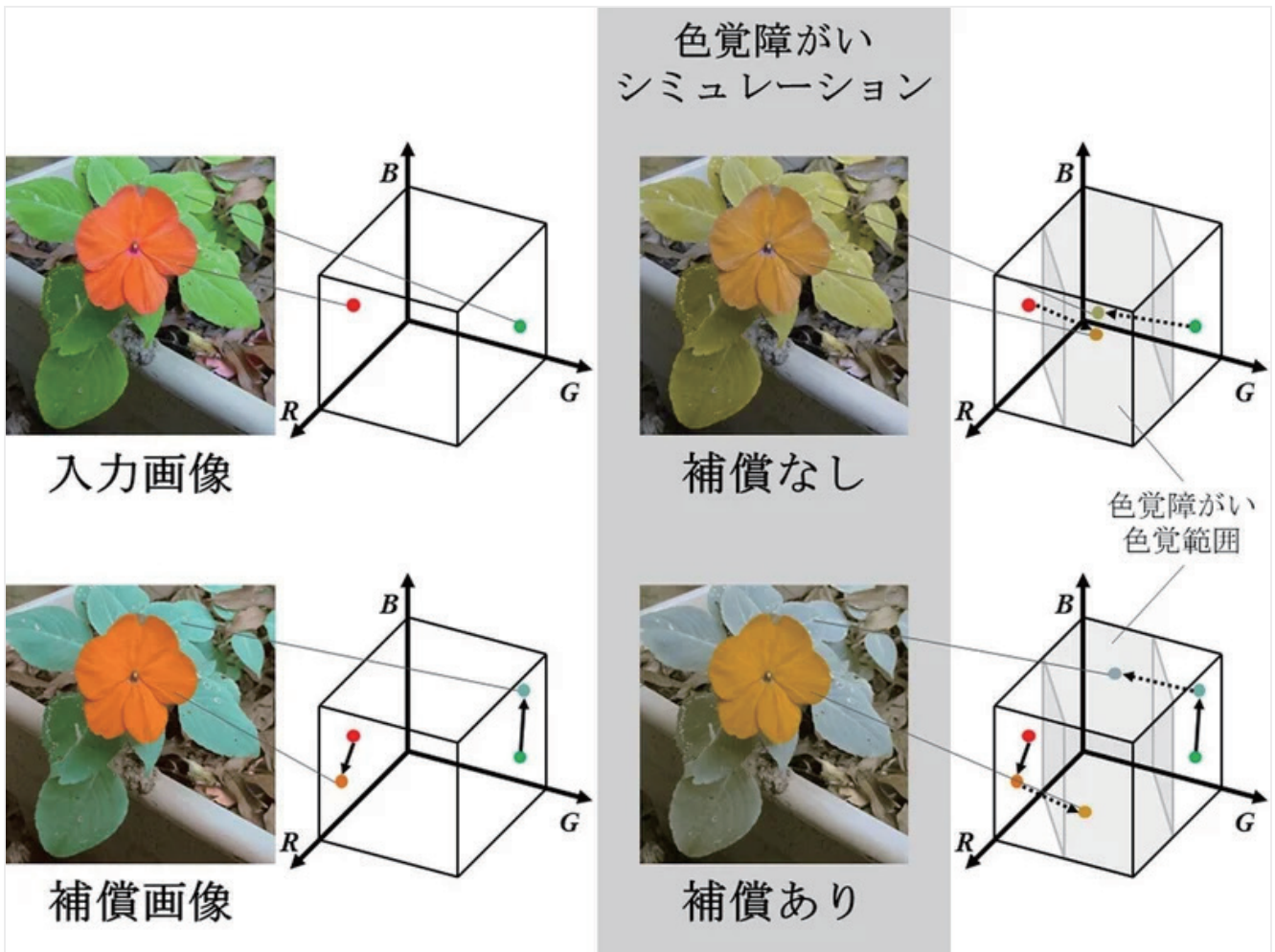
健常者の色覚範囲は三次元空間であるが、重度のCVDの色覚範囲は三次元空間内の二次元平面として近似できる。知覚できない色成分を含む2色が二次元平面に投影されると、両点間の距離が縮まり、コントラスト損失が生じる。既存の色覚補償に関する研究では、コントラストを強調するために、変える必要のない色まで変え

られるため、CVD患者に違和感を与える（自然さ損失）。また、既存研究では個人の障がいの度合に適應することができない。

そこで、本研究はCVDの色覚範囲において、コントラスト強調と自然さの保存の両立を可能とする色変換技術を開発する。提案技術では、コントラスト制約と自然さ制約を取り入れた最適化モデルを用いて、支援画像を生成する。コントラスト制約では、画像をコントラストが強調される方向に変換させるため、支援画像が本来の画像から引き離されてしまうが、自然さ制約では、支援画像を本来の画像へ引き戻す。提案モデルでは、2つの制約がそれぞれ生じるエネルギーの和が最小となるような支援画像を生成する。

本研究では、まず、錐体細胞の反応値に基づいたLMS色空間で支援画像を生成する色変換技術¹⁾を提案した。提案技術の性能を向上させるために、知覚的均等性をもつCIE L*a*b*色空間での色変換技術²⁾も開発した。評価実験では、標準ベンチマークに加え、CVD患者による主観的評価においても、提案手法が既存研究よりも良い結果を得ている。しかし、これらの手法は画像1枚を作成するのには1分以上かかるため、自然さを保つコントラスト補償を三次元色空間からCVD色域への最適線形投影として定式化し、高速色変換技術³⁾を開発した。さらに、CVDの個人差に着眼し、異なるレベルのCVDをシミュレーションモデルに基づいて、個人のCVD度合に適應した色変換技術⁴⁾を開発した。定量評価およびCVD患者による主観評価の結果によって、提案技術の有効性が示されている。

本研究は医学と情報工学の共同研究であり，いまだに医学的な根治手段が見つからない色覚障がいに対して情報技術による支援策を提供する．さらに，医工融合研究の重要性に関する発信にもつながる．



■Webサイト／動画／アプリなどのURL

<http://www.vc.media.yamanashi.ac.jp/ja/computational-ophthalmology-anomalous-trichromacy-compensation/>

参考文献

- 1) Zhu, Z., Toyoura, M., Go, K., Fujishiro, I., Kashiwagi, K. and Mao, X. : Processing Images for Red–green Dichromats Compensation Via Naturalness and Information-preservation Considered Recoloring, The Visual Computer, Springer-Nature, Vol.35, No.6-8, pp.1053-1066 (2019-6)(IF:2.835(2021))
- 2) Zhu, Z., Toyoura, M., Go, K., Fujishiro, I., Kashiwagi, K. and Mao, X. : Naturalness- and Information-preserving Image Recoloring for Red-green Dichromats, Signal Processing: Image Communication, Elsevier, Vol.76, pp.68-80 (2019-8) (IF:3.453(2021))
- 3) Wang, X.*, Zhu, Z.*, Chen, X., Go, K., Toyoura, M. and Mao, X. : Fast Contrast and Naturalness Preserving Image Recolouring for Dichromats, Computers & Graphics, Elsevier, Vol.98, pp.19-28 (2021-8) (IF:1.821(2021), *:equally contribution)
- 4) Zhu, Z., Toyoura, M., Go, K., Kashiwagi, K., Fujishiro, I., Wong, T.-T. and Mao, X. : Personalized Image Recoloring for Color Vision Deficiency Compensation, IEEE Transactions on Multimedia, IEEE, Vol.24, pp.1721-1734 (2021-3) (IF:8.182(2020-2021))

(2022年5月23日受付)

(2022年8月15日note公開)

取得年月日：2021年9月

学位種別：博士（工学）

大学：山梨大学

**推薦文 [メディア知能情報領域] 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**

世界中で一億人以上の人々が持つ色覚障がいによる色の識別困難（コントラスト損失）を補償するために、本論文では、最先端の色覚障がいシミュレーション手法に加え、生成支援画像がなるべく自然に見えるような色変換最適化技術を提案した。これまで困難であった色覚障がい度の計測についても、解決策を提供している。

研究生活 私は大学1年次に「画像処理」という分野に興味を湧いて、専門の先生とコンタクトをとり、研究室に仮配属してもらいました。当時取り組んでいた「カラー画像のグレイスケール化に関する研究」からの影響もありますが、自分が「色覚障がい支援」を博士課程の研究テーマにした大きな理由のひとつは、私自身も色覚障がい者の一人で、この研究課題によって私は「色覚障がい」と戦う機会を与えられたと思ったからです。主観的評価実験で提案手法の支援画像が最も見やす

いと評価されたとき、自分が従事してきた研究が有意義であると強く感じて、これからも同様の研究を続けていきたいと考えるようになりました。研究活動とともに、自分の視野を広げることができて、博士課程に進学して本当によかったと思います。

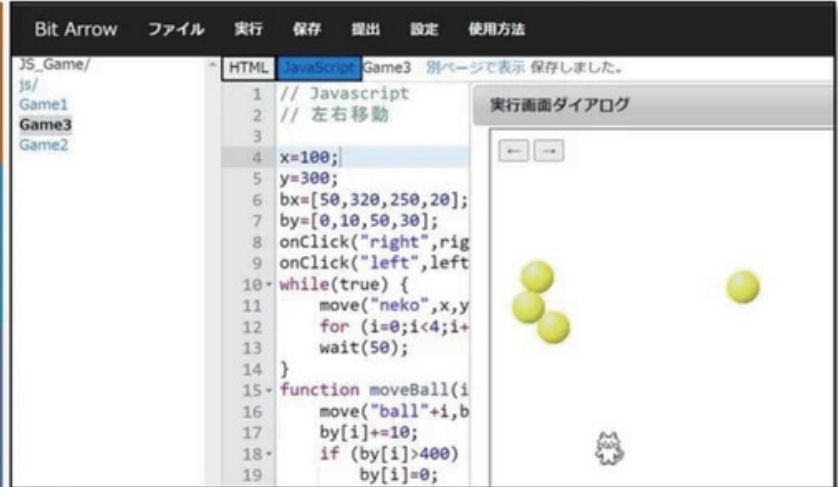
本研究の推進にあたり、指導教員の茅暁陽教授をはじめ、たくさん助言をしていただいた先生方に深く感謝を申し上げます。また、実験に協力していただいた方にもお礼を申し上げます。

Webブラウザで利用可能な プログラミング環境Bit Arrowの開発

学習者自身による
エラー修正支援


教員による学習者
の状況把握支援

高校教科「情報I」
の学習内容対応



オンラインプログラミング学習環境Bit Arrow の研究と開発

♡ 1

 情報処理学会・学会誌「情報処理」
2022年8月15日 11:34



2021年度研究会推薦博士論文速報

[コンピュータと教育研究会]

長島 和平

(東京農工大学 総合情報メディアセンター 特任助教)

■キーワード

プログラミング教育／学習環境／高等学校

【背景】 高等学校におけるプログラミングの必修修化

【問題】 インストールなしで利用可能な学習支援環境の不足

【貢献】 Webブラウザで動作するプログラミング教育支援環境の提供

近年、情報技術の発展に伴い、多くの人々が日常的にスマートフォンなどの情報機器を所有したり情報システムを利用したりするようになってきている。このような機器やシステムを利用するにあたり、それらがどのような仕組みで動作しているのか理解することが大事になる。こうした背景から、2022年度に適用された高等学校の学習指導要領では、必修科目「情報I」にプログラミングが組み込まれることとなった。これまでプログラミングを扱ってこなかった教員が初学者の生徒に授業を行うため、これを支援する環境が必要となる。PCにインストールして利用するプログラミング教育支援環境を開発した先行研究もあるが、高等学校では演習PCを自治体が管理していることも多く、環境をインストールすることが難しいという課題がある。また、学習内容に応じたプログラミング言語や内容を扱える必要がある。

そして授業での利用を考えたときには、学習者と教員双方の支援を行う必要がある。こうした課題を解消するために、本研究ではWebブラウザで動作するプログラミング学習支援環境「Bit Arrow」を開発した。

Bit Arrowは、インターネットに接続されていれば利用できるため、事前に環境を演習PCにインストールするといった作業の必要がない。Python, ドリトル, JavaScript, Cという高校や大学でも利用されているプログラミング言語をライブラリとともに提供しており、学習内容や習熟度に応じて適したプログラミング言語を選択することができる。授業内容への対応の検証では、学習指導要領を元に文部科学省から出された教員研修用教材に掲載されたプログラムと同等の内容がBit Arrowで扱えることが分かった¹⁾。

学習者への支援という観点では、エラーの自力での修正に着目した。初学者にとっては、エラーメッセージからエラーが発生している場所と原因を特定することが難しい。Bit Arrowでは独自の処理系を実装することで、エラーメッセージとともにアドバイスを提供し、エラーが発生した個所にマークをつけて通知する。また、エディタにはエラーの発生そのものを減少させるための入力補助などの機能も提供する。実際に高等学校の授業で、教科書に掲載されていた従来の環境とBit Arrowをそれぞれ使って比較すると、Bit Arrowを使ったクラスではエラーの発生率が低下し、エラーの修正にかかる時間も短くなったことが確認できた²⁾。

教員への支援という観点では、学習者の状況を把握することに着目している。学習者がプログラムを実行するたびに、Bit Arrowではログを収集している。ログには、ソースコードと実行結果、エラーが出た場合はエラーメッセージ、時刻などの情報が含まれている。この情報を教員に提供することで、エラーが続いていたり、手が止まっていたりという学習者の状況を把握させることができ、困っている学習者を指導することを支援できる³⁾。また、このログを使って、支援が必要な学習者を自動で検出するための研究にも発展している。

Bit Arrowは2016年から公開しており、約430名の高校の先生方を中心に、小学校から大学まで合わせて約550名の先生方に利用されている。博士論文では提案手法について、実際の教育現場で利用した効果や、Bit Arrowで扱える学習内容と学習指導要領との比較といった検証を通じてその有用性を示した。

Webブラウザでインストールなしで利用可能

学習者への支援

- エラー発生場所やアドバイス提示による自力でエラー修正
- 入力補助によるエラーの発生防止

エラー

'X'は定義されていません。大文字小文字は区別されます。 'x'の間違いかもしれません。 場所 : print01.c:4:19

```
#include<stdio.h>
int main(){
  int x=10;
  printf("%d\n", !x);
  return 0;
}
```

エラーメッセージとアドバイス提示

エラー発生場所通知

教員への支援

- サポートが必要な学習者を把握
 - エラーが多い
 - 手が止まっている
 - 他の学習者より遅れている

2018/05/11 17:00:00 から 2018/05/11 17:30:00までの実行状況 集計...

ユーザID	エラー/実行	実行からの経過時間	今実行しているファイル	直近の実行履歴
user20	1/4(100%)	00:22:53	0511/P0511_5.c	3213 ESESESES
user10	2/2(100%)	00:21:18	0511/P0511_5.c	1144 ESES
user62	3/8(100%)	00:01:05	0511/P0511_5.c	998 SSSSESESESSSESE
user55	1/7(14%)	00:00:51	0511/P0511_5.c	5 ESESESESESESES
user11	5/9(56%)	00:04:41	0511/P0511_5.c	5 ERSERRRRSESESE
user14	1/1(100%)	00:00:36	0511/P0511_5.c	998 RRRRSRRRSRRSESRR

エラー発生率

手が止まっている時間

■Webサイト/動画/アプリなどのURL

<https://bitarrow.eplang.jp/>

参考文献

1) 長島和平, 長 慎也, 兼宗 進, 並木美太郎 : Bit Arrowにおける高校教科「情報」教員研修資料との対応評価, 情報処理学会論文誌 教育とコンピュータ (TCE) , Vol.8, No.2, pp.91-106 (June 2022).

2) 長島和平, 長 慎也, 間辺広樹, 兼宗 進, 並木美太郎 : Webブラウザを用いたプログラミング学習支援環境Bit Arrowの設計と評価, 情報処理学会論文誌 教育とコンピュータ (TCE) , Vol.4, No.1, pp.57-69 (Feb. 2018).

3) Nagashima, K., Cho, S., Horikoshi, M., Manabe, H., Kanemune, S. and Namiki, M. : Design and Development of Bit Arrow: A Web-Based Programming Learning Environment, Proceedings of the 10th International Conference on Education Technology and Computers, pp.85-91 (2018/10).

(2022年5月31日受付)

(2022年8月15日note公開)

取得年月日 : 2022年3月

学位種別 : 博士 (工学)

大学 : 東京農工大学

正会員

推薦文 [メディア知能情報領域] コンピュータと教育研究会

オンラインプログラミング環境「Bit Arrow」を提案した研究です。ブラウザがあ


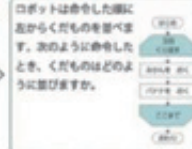
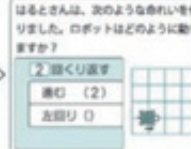






れば動作することや、複数のプログラミング言語を選択して利用できること、授業用の管理機能などが用意されていることが特徴です。システムはすでに公開され、全国の数百カ所の高校や大学で利用されている実績があります。

研究生生活 プログラミング教育に興味を持ったのは、学部生時代にプログラミング演習を受けたときです。周りに入門のような内容でも苦労している学生がいることに気づき、研究室配属後は初学者が楽しく学習できるプログラミング環境をテーマにしました。そんな折、学習指導要領の改訂が発表され、高校でプログラミングが必修になることが分かりました。これまでの研究が活かせると思い、博士課程でこのテーマを扱いました。

教育現場で実際に利用するため、不具合で授業ができないということのないように実装には注意を払いました。Bit Arrowは、本稿執筆時点で約550名の先生方に利用いただいています。特に2022年に入ってから5月末までの間には高校を中心に約140名の先生方から利用申請があり、プログラミング必修化による影響の大きさを感じています。

本研究は、多くの方々のサポートを受けて行うことができました。主指導教員の並木美太郎先生、共同研究者の兼宗進先生、長慎也先生、間辺広樹先生、そして各研

研究室の学生さんと、ご利用いただいている先生や生徒の皆様に、この場を借りてお礼申し上げます。

コーディングによる学習時の課題を解決する学習教材の検討	
課題：プログラムの入力や動作に意識が向くことで、本来の目的であるプログラムの考え方が学習できない	
小・中学校	<p>基本的な制御構造とデータ構造</p> <ul style="list-style-type: none"> ・コーディングから離れて学習 ・体系的かつ段階的にプログラムの考え方を学ぶドリル教材 <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>次のメモにかかれたとおり、左から花をならべるとき、どのようにならびますか。</p>  <p>4回くり返す</p> <p>身近な物を題材にした問題</p> </div> <div style="text-align: center;"> <p>ロボットは命令した順に忘からくたものを書べます。次のように命令したとき、くたものはどのように並びますか。</p>  <p>フローチャートによる問題</p> </div> <div style="text-align: center;"> <p>はるとさんは、次のようなおれいを作りました。ロボットはどのように動きまますか？</p>  <p>ビジュアル言語による問題</p> </div> </div>
高等学校	<p>アルゴリズム</p> <ul style="list-style-type: none"> ・コーディングから離れて学習 ・CSアンプラグド*1の学習に着目 ・既存教材の課題をAR技術で解決 <p>*1) https://csunplugged.jp/</p> <div style="display: flex; justify-content: space-around;">    </div>
	<p>実用的なアプリケーション開発</p> <ul style="list-style-type: none"> ・スマートスピーカーのアプリ開発を題材にコーディングと併用し学習 ・汎用言語の約1/10の行数で開発可能 <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>1 起動＝「クローバー！あいさつを英語で返します。」話す。』。</p> <p>2 こんにちは＝「クローバー！"Hello" 話す。』。</p>  </div> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> </div>

初等中等段階におけるプログラミング教育教材の研究



情報処理学会・学会誌「情報処理」

2022年8月15日 11:36



2021年度研究会推薦博士論文速報

[コンピュータと教育研究会]

島袋 舞子

(大阪電気通信大学 メディアコミュニケーションセンター 特任講師)

■キーワード

プログラミング教育／アルゴリズム学習／プログラミング学習環境

【背景】 初等中等段階におけるプログラミング教育

【問題】 コーディングによる学習時の課題

【貢献】 学習内容の本質を学ぶための学習教材の開発

本研究では、コーディングによるプログラミング学習時の課題を解決するために「コーディングから離れて考え方を学習する方法」と「コーディングを併用することで学習効果を高める方法」の2つのアプローチによる「学習内容の本質を学ぶための学習教材」を提案し、提案した教材の有用性を明らかにした。

日本におけるプログラミング教育は、小学校から高等学校まで段階的に行われる。小学校では主に各教科でのプログラミング活動や作品制作によるプログラミング体験を通して基本的な制御構造について触れる。中学校では技術・家庭科（技術分野）で基本的な制御構造と変数等の基本的なデータ構造を学習後、実用的なプログラムへの応用として外部機器を用いた計測・制御やネットワーク通信に関するプ

プログラムを学習する。高等学校では情報科において数値計算等におけるプログラムの活用やアルゴリズム、関数や外部ライブラリの利用について学習し、その後データベース管理・操作やデータサイエンス、人工知能等を用いた実用的なシステムや機器を活用するプログラムについて学習する。多くの場合、プログラミングの学習はコーディングを通して行われるが、その一方で児童や生徒はプログラムの動作に意識が向くことで本来の目的であるプログラムの考え方を学習できない場合がある。

各段階でのプログラミング学習におけるコーディングによる学習時の課題解決を試みるため、小学校から高等学校までのプログラミング教育の学習内容の中から、プログラミングの基礎となる「基本的な制御構造やデータ構造」と、それらを理解した上で学習が行われる「アルゴリズム」、さらに応用的な「実用的なアプリケーション開発」の3つの学習内容に着目し、研究を進めた。

まずは、児童・生徒が基本的な制御構造やデータ構造を学習する際に課題となる「コード以外のことやエラーを解消することに夢中になったり、偶然実現した動きに満足してしまいプログラミングの考え方を学習できない」点を解決するために、コーディングから離れて考え方を学習する方法として、問題を解くことでプログラミングの考え方を思考し、体系的に学ぶことができるドリル教材を提案した

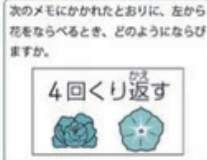
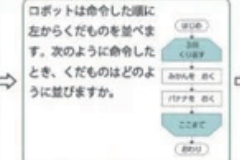
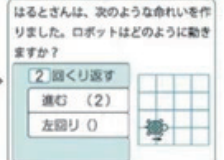




1) ~3) . ドリル教材の学習効果について小学校低学年を対象に検討を行ったところ、低学年の児童が順次処理の概念を理解した上でタートルグラフィックスの概念

やプログラミング的思考の要素の1つである「命令を組み合わせてプログラムを作成する」ことを理解できることが分かった⁴⁾。

次にアルゴリズム学習時の課題である「プログラムを間違えずに入力することに集中し、考え方を学習できない」点を解決するために、コーディングから離れて考え方を学習する教材の1つであるCSアンプラグドの学習法のソートアルゴリズム学習に着目した。既存のソートアルゴリズム学習教材に存在した課題である「作業と同時に比較回数を数えることが難しい」点と「教具の準備に時間がかかる」点を解決するために、実世界と仮想世界を端末の画面上で統合したAR天秤を提案した⁵⁾。AR天秤の学習効果について大学生と高校生を対象に検証を行ったところ、AR天秤が既存の天秤とおもりを使用する場合と同等に選択ソートの手順を理解できること、AR天秤を使用した場合にクイックソートの手順の理解が向上することが明らかとなった。また、授業実践を通してAR天秤が授業で活用できることが分かった。

最後に実用的なアプリケーション開発を通じた学習時の課題となる「実用的なアプリケーション開発は、プログラミングの初学者が行うのは容易ではなく、プログラムが複雑で理解が難しいためにプログラムを入力することに集中してしまい学習ができない」点に対して、スマートスピーカー専用のアプリケーションを開発するプログラム開発環境を提案した⁶⁾。高校生を対象に授業実践を行い、実施後の生徒へのアンケート調査と教員への聞きとり調査により検証したところ、高校生が提

案環境でアプリ開発が可能であり、人工知能を題材とする外部ライブラリやWeb APIを扱うプログラムを授業で扱うことが可能であることが分かった。

コーディングによる学習時の課題を解決する学習教材の検討	
課題：プログラムの入力や動作に意識が向くことで、本来の目的であるプログラムの考え方が学習できない	
小・中学校	<p>基本的な制御構造とデータ構造</p> <ul style="list-style-type: none"> ・コーディングから離れて学習 ・体系的かつ段階的にプログラムの考え方を学ぶドリル教材 <div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;"> <p>次のメモにかかれたとおり、左から花をならべるとき、どのようにならびますか。</p>  <p>4回くり返す</p> <p>身近な物を題材にした問題</p> </div> <div style="text-align: center;"> <p>ロボットは命令した順に左からくだものを並べます。次のように命令したとき、くだものはどのように並びますか。</p>  <p>フローチャートによる問題</p> </div> <div style="text-align: center;"> <p>はるとさんは、次のような命れいを作りました。ロボットはどのように動きますか？</p>  <p>2回くり返す 進む (2) 左回り (1)</p> <p>ビジュアル言語による問題</p> </div> </div>
高等学校	<p>アルゴリズム</p> <ul style="list-style-type: none"> ・コーディングから離れて学習 ・CSアンプラグド^{*1}の学習に着目 ・既存教材の課題をAR技術で解決 <p>^{*1}) https://csunplugged.jp/</p> 
	<p>実用的なアプリケーション開発</p> <ul style="list-style-type: none"> ・スマートスピーカーのアプリ開発を題材にコーディングと併用し学習 ・汎用言語の約1/10の行数で開発可能 <div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;"> <p>1 起動＝「クローバー！」あいさつを英語で返します。」話す。」。</p> <p>2 こんにちは＝「クローバー！」"Hello" 話す。」。</p>  </div> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> </div>

参考文献

1) 島袋舞子, 兼宗 進：ドリルの王様1, 2年のたのしいプログラミング, 新興出版社啓林館 (2019).

2) 島袋舞子, 兼宗 進：ドリルの王様3, 4年の楽しいプログラミング, 新興出版社啓林館 (2019).

3) 島袋舞子, 兼宗 進：ドリルの王様5, 6年の楽しいプログラミング, 新興出版社啓林館 (2019).

- 4) 島袋舞子, 井戸坂幸男, 兼宗 進: 小学生を対象にしたプログラミング学習用ドリル教材の開発と低学年での評価, 日本産業技術教育学会誌, Vol.62, No.2, pp.141-149 (2020).
- 5) 島袋舞子, 林 康平, 兼宗 進: 拡張現実感を用いたソートアルゴリズム学習教材の提案, 日本バーチャルリアリティ学会論文誌, Vol.22, No.2, pp.199-208 (2017).
- 6) 島袋舞子, 本多佑希, 兼宗 進: スマートスピーカーを題材にした高等学校におけるプログラミング学習環境の提案, 情報処理学会論文誌教育とコンピュータ (TCE) , Vol.6, No.3, pp.21-30 (2020).

(2022年6月1日受付)

(2022年8月15日note公開)

取得年月日：2021年12月

学位種別：博士（工学）

大学：大阪電気通信大学

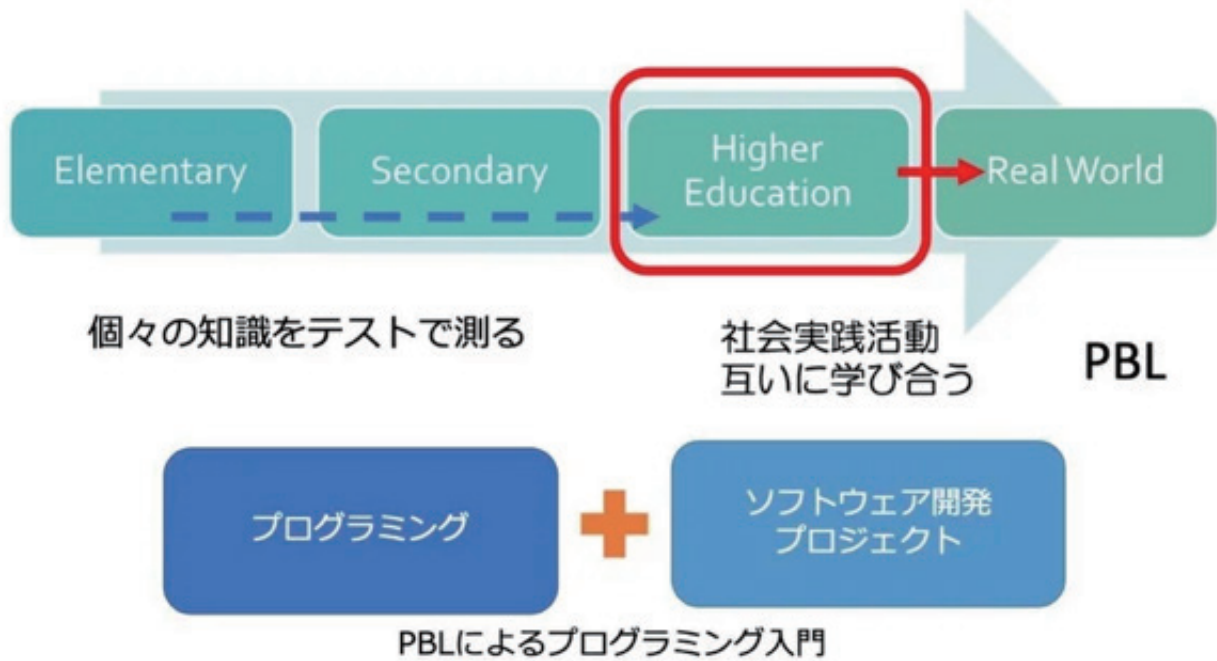
正会員

推薦文【メディア知能情報領域】 コンピュータと教育研究会

小学校からのプログラミング教育が進められる中で、プログラミングや情報科学の内容を児童・生徒が理解するための教材と授業方法の研究です。小学生に適したプログラミングの考え方を学習できるドリル教材と、高校生に適したアルゴリズム学習教材を提案し、実際に教育効果を明らかにしています。

研究生活 私は高校教員を目指して大学に入学した経緯もあり、以前からコンピュータの仕組みを楽しく分かりやすく学ぶ方法に興味を持っていました。修士課程を修了後、少し期間が空いて社会人大学院生として本研究を進めました。研究時間の確保に少し苦労しましたが、多くの共同研究者の先生方のご協力とご支援により成果を得ることができました。この場をお借りして感謝申し上げます。

今後もコンピュータの仕組みを楽しく学ぶ方法について研究を進めていきたいと思っています。



一般情報教育におけるプログラミング入門教育に関する研究

♡ 7

情報処理学会・学会誌「情報処理」
2022年8月15日 11:38



2021年度研究会推薦博士論文速報
[コンピュータと教育研究会]

内田 奈津子

(フェリス女学院大学 講師)

■キーワード

プログラミング／PBL (Project-based Learning) ／描画教材とソフトウェア
開発

【背景】 情報化社会でより良く生きていくための知識の育成

【問題】 誰もが学んでほしいプログラミングの知識

【貢献】 PBLによるプログラミング入門カリキュラム

これまで専門家を中心に、限られた人々のみで使われていた情報通信技術 (ICT) は、私たちの暮らしの中にまで浸透してきた。このような社会において、人々の生活をより良くするためには、ICTのスペシャリスト集団のみで対応できるわけではなく、さまざまな知識を持った多様な人々が協力してかかわる必要がある。そのため、これまで求められていた専門分野に特化した人材育成にとどまらず、誰もが基本的なICTスキルを身につけ、社会全体の底上げをはかる必要がある。

学習指導要領の改訂により、2020年4月から小学校でのプログラミング教育が始

まった。これにより、プログラミングは、義務教育のひとつとなり、誰もが学ぶべきものとなった。しかし、高等教育においては、必須化された科目がなく、情報を専門としない学生は、誰もが学ぶ環境には至っていない。そのため、学習指導要領改訂以前に大学に入学した中には、プログラミングの知識を学ばないまま、社会に出る者もいる。特に、著者の在籍する文系女子大においては、プログラミングを学ぶ機会を創出することは困難であった。また、これまでの入門レベルのプログラミング教育は、個々の知識を高めることに重点が置かれ、チームで学ぶような構成にはなっていない。そのため、ソフトウェア開発は、情報を専門とする専攻の大学で、プログラミングの知識を一通り学んだ後に学ぶことになる。

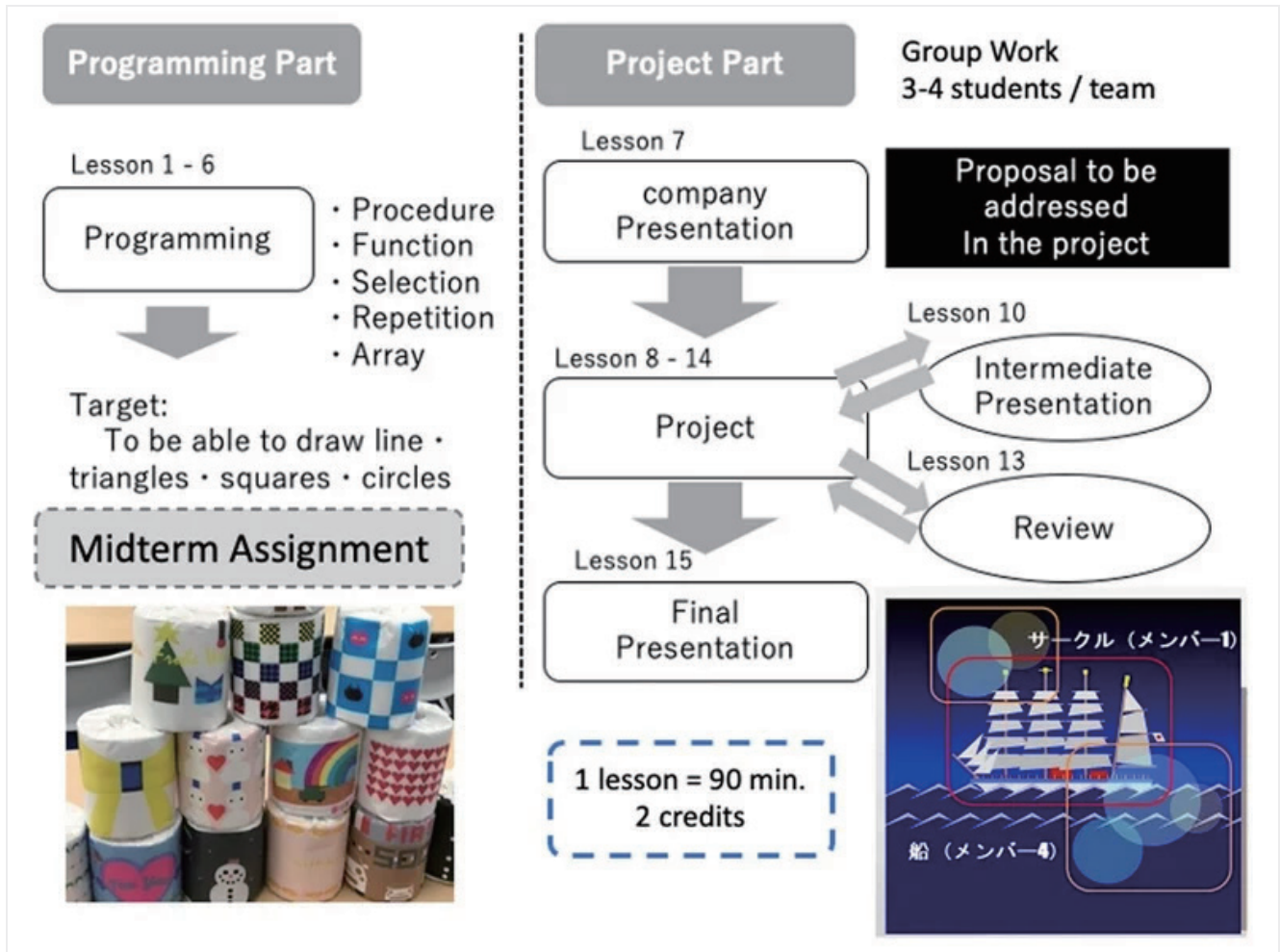
教育面では、これからの社会における知識の活用を考えると、市民が社会実践活動の中で互いに学び合うという側面を重視し、他者とのかかわりをもって学ぶ教育への変化が求められており、アクティブラーニングやPBL (Project-based Learning) による教育を取り入れる必要がある。

本研究では、入門レベルのプログラミング教育にPBLを組み合わせた方法により、情報を専門としない学生が、プログラミングの概念とソフトウェア開発プロジェクトの知識を、同時に学ぶカリキュラムの構築に取り組んだ。単に、プログラミング言語を学び、コードが書けるようになることを目的とするのではなく、プログラミングの原理を理解することに加え、ソフトウェア開発プロジェクトの知識を伴ってチームで学ぶ機会を作った。高等教育において確保できる最小の単位は、90

分15回であることから、この時間枠の中に収まるようカリキュラムを工夫した。教材は、理系科目が苦手な学生が対象となることから、誰もが取り組みやすく親しみやすい、丸や三角形などから構成される描画教材を用い、初回からコードを書くことを目指した。授業を2部構成とし、後半では、1つの絵をソフトウェアに見立てて、それをいくつかのオブジェクトに分割し、チームで1つの絵を完成させるプロセスをプロジェクトとして組み込むことにより、ソフトウェア開発プロジェクトを体験させることができた。

構築したカリキュラムは、2017年より現在まで、著者の所属する文系女子大で実施しており、プロジェクトでは、単にチームで1つの絵を完成させるだけでなく、ネイルシールやステッカーなど、リアルなものづくりを課題にして取り組んでいる。

視覚的に分かりやすく、誰にでも教えやすい容易な教材により、誰もが教えやすく、学びやすい環境が構築でき、入門レベルの限られた時間の中でも、ソフトウェア開発を伴って学べることを示した。



参考文献

- 1) 内田奈津子：ぺた語義：プログラミング入門をプロジェクトでやってみた，情報処理学会学会誌「情報処理」，Vol.59, No.3, pp.268-271 (2018).
- 2) 内田奈津子，久野 靖，中山泰一：PBLによるプログラミング入門科目の提案：一般情報教育における入門カリキュラムの構築，情報処理学会論文誌，Vol.62, No.7, pp.1393-1414 (2021).
- 3) 伊地知咲希，永井絵梨奈，内田奈津子：プロジェクトにおける活動分析ーお絵

かきプログラミングによる入門教育とPBL一, 2020年度 情報処理学会関西支部 支部大会 講演論文集 (2020).

4) 米屋美雪, 内田奈津子: 分岐を学ぶために効果的な教材を考える ーお絵かきプログラミングによる入門科目を振り返ってー, 情報処理学会 研究報告コンピュータと教育 (CE), 2021-CE-159, No.9, pp.1-6 (2021).

(2022年5月31日受付)

(2022年8月15日note公開)

取得年月日：2021年9月

学位種別：博士（工学）

大学：電気通信大学

正会員

推薦文 [メディア知能情報領域] コンピュータと教育研究会

情報を専門としない文系を含む大学生が、プログラム入門から始めて役に立つソフトウェア開発に入門することを、2単位の時間数で可能とするような新しいカリキュラムを開発した。前半はお絵描きプログラムを題材に自分の思うコードを書くこ

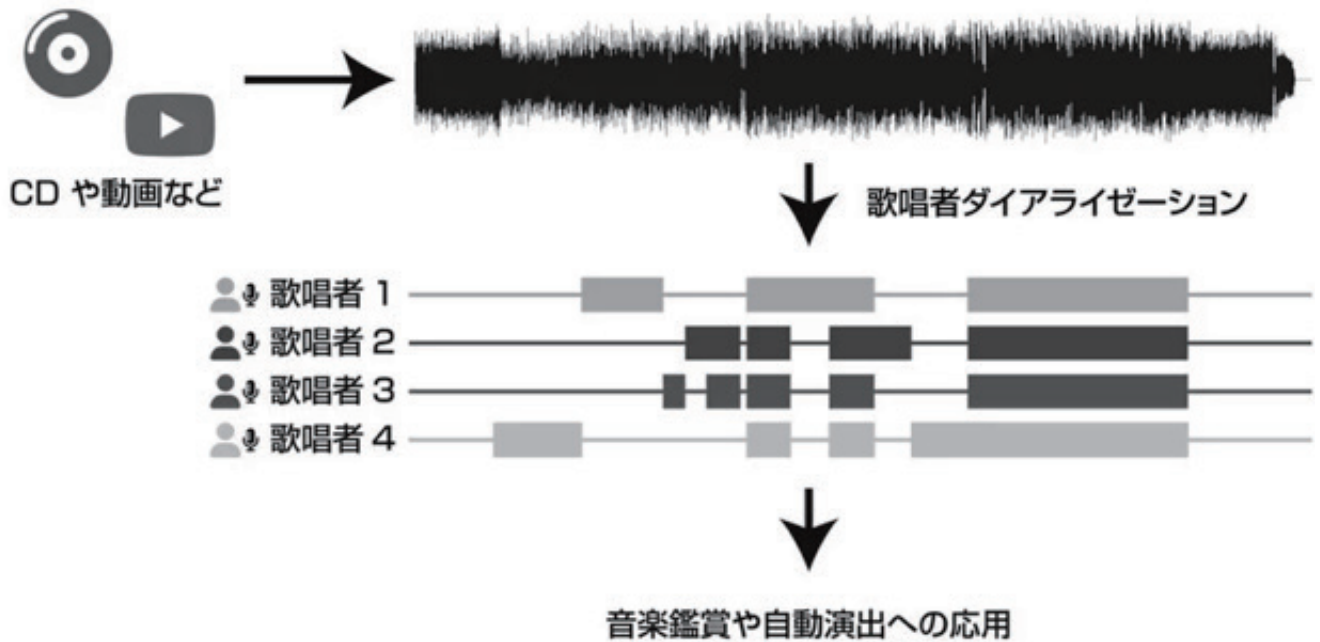
とを学び、後半はPBLでチーム開発を学ぶ。実践において高い有効性を示しており、高く評価できる。

研究生生活 長年、文系女子大で情報教育に携わってきた。多くの時間は、WordやPowerPointに関する内容の相談に費やされた。あるとき、地域団体から小学生へのプログラミング教育について相談されたことを機に、ロボット教材を活用したチームによるものづくりに取り組むことになった。

この経験から、情報を専門としない学生にも、ものづくりを通じたプログラミングを教えることを考えたいと思うようになった。子育てもひと段落したタイミングで、背中を押された。そのときは、大きなシステムリプレイスが控えていたこともあり、1年早いと思い一瞬躊躇した。しかし、今となれば思い切って飛び込んだのは正解だったかもしれない。


仕事と家庭と研究との両立はなかなか難しく、ToDoリストではなく、どれをやめるかリストを作ったことは大きな思い出である。

大学院進学の門戸を開き、指導をしてくださった久野靖先生、中山泰一先生、その他の方々の協力に大変感謝しております。



複数人歌唱楽曲に対する音楽情報処理に関する研究

♡ 1

 情報処理学会・学会誌「情報処理」
2022年8月15日 11:39



2021年度研究会推薦博士論文速報

【音楽情報科学研究会】

須田 仁志

(LINE株式会社 AIカンパニー AI開発室)

■キーワード

音楽情報処理／複数人歌唱／機械学習

【背景】 「みんなで歌う」をうまく扱えない音楽情報処理

【問題】 複数人歌唱の自動分析と鑑賞技術への応用

【貢献】 音響的に扱いにくい同時発声へのアプローチの提案

本研究は、音楽を題材とした情報処理である音楽情報処理とよばれる分野に属するものである。音楽情報処理は非常に多岐に渡る分野であるが、そのうち本研究では複数人が歌う楽曲にまつわる情報処理を扱っている。

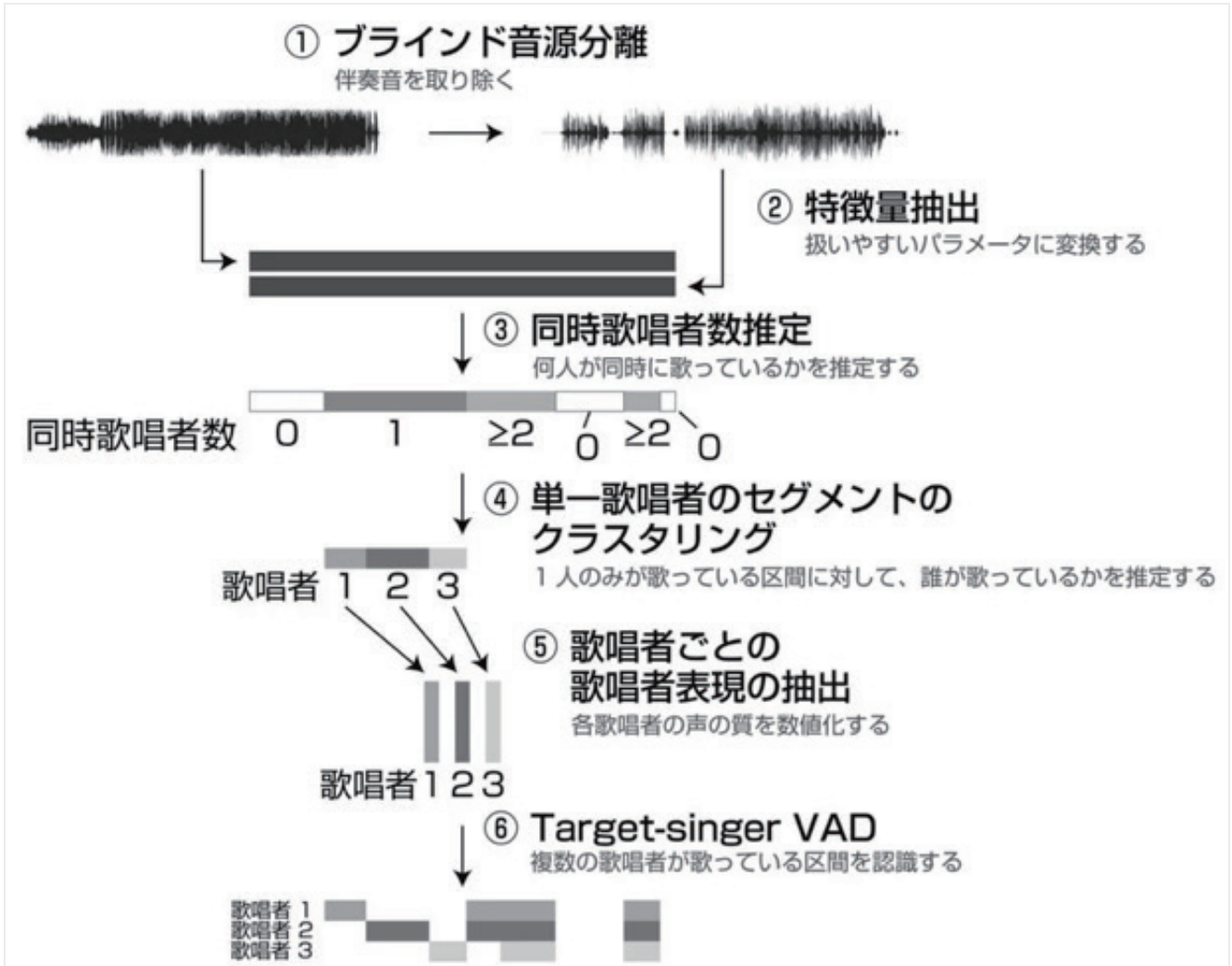
複数人が歌うといっても、合唱曲やアイドルソング、オペラなど、世界にはさまざまなスタイルの楽曲がある。本研究では特に、歌手が入れ代わり立ち代わりながら歌うような、歌い分けのあるアイドルソングを題材とした。実験には、典型的なアイドルソングを模しているものとして、「アイドルマスター」シリーズや「ラブライブ！」シリーズの楽曲を採用した。

本研究の目標は、①こうした楽曲を自動的に分析すること、そして②分析結果を音楽鑑賞に応用すること、の2点である。

①分析技術においては、「歌唱者ダイアライゼーション」とよばれる問題を扱う。これは、歌い分けのあるアイドルソングのような楽曲から「いつ誰が歌っているか？」を、歌手や楽曲の知識なくコンピュータに認識させる技術である。歌唱者ダイアライゼーションが実現できれば、どんな声の歌手が歌っているのか、何人の歌手が歌っているのか、などさまざまな情報を音楽から抽出できる。この問題は、会話音声から「いつ誰が話しているか？」を解く「話者ダイアライゼーション」という問題に似た問題である。しかし、アイドルソングは、伴奏が演奏されていたり、各歌手の声質が近かったり、全員で長時間同じ音高で歌っていたり（ユニゾン）など、会話音声とは大きく異なる点が多くある。したがって、会話音声に用いられる技術をそのまま適用することは難しい。また、なによりこうした会話音声に用いられる技術がアイドルソングに適用できるかを実験的に検討した例は存在しない。

本研究では、歌声と会話音声がそもそもまったく異なるというところから始まり、さまざまな技術や独自の分析指標（特徴量）を導入することで、既存の技術をそのまま適用するよりもはるかに高い精度で歌唱者ダイアライゼーションを行うことのできる技術を提案した（図）。

②鑑賞技術においては、「VocalRemixer」という音楽鑑賞アプリケーションを提案した。これは、複数人が歌うアイドルソングの歌い分けを可視化し、さらにその歌い分けを自由に操作できるというアプリケーションである。好きなアイドルユニットの楽曲を入力すると、特定のアイドルに単独で歌唱させたり、まったく異なる歌い分けで歌わせたりすることができる。各歌手がいつ歌っているかを歌唱者ダイアライゼーションによって与えて、各歌手の声質を声質変換とよばれる技術によって別の歌手の声質に変換することで、このアプリケーションを構築できる。アイドルユニットの楽曲は、CDなどに収録されている歌い分けそのままでは楽しむことのできないものであったが、VocalRemixerによって1つの楽曲を何倍にも楽しむことができるようになる。



参考文献

1) Suda, H., Saito, D., Fukayama, S., Nakano, T. and Goto, M. : Singer Diarization for Polyphonic Music with Unison Singing, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol.30 (2022).

(詳細は <https://ieeexplore.ieee.org/document/9767568> をご確認ください)

(2022年5月29日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（工学）

大学：東京大学

推薦文【メディア知能情報領域】音楽情報科学研究会

本博士論文は複数人が歌う楽曲において「誰がいつ歌っているか」の情報を取り出す歌唱者ダイアライゼーション技術と歌唱者を変更する声質変換技術を精錬させた上で、VoiceRemixerという新たな歌唱インタフェースを通してこれらの技術を結合し、複数人歌唱楽曲の分析、加工、提示技術を実現している。斬新なアイディアで高く評価される論文である。

研究生生活 私は音楽を収集したり聴いたりすることが趣味である。この研究テーマは、「いつも音楽を聴いていて思うことはない？」との先生方の問いかけに、「アイドルソングはいつ誰が歌っているのか聴き分けるのが難しい。自動で分析し

て可視化できたら楽しいのに」と答えたところがはじまりであった。好きなこと、そして世界で自分にしかできないことを研究テーマに決めたことで、なんとか研究のモチベーションを保つことができたと思う。そして「大好きなもので博士号とったぞ!」と言えることに誇りを持っている。

博士課程に進学したことで3年という時間を与えられ、研究だけでなく趣味の開発などでも数多くの経験を積むことができた。これは、修士時代に進学を選択しなければ得られなかったものだと思っている。博士課程は、博士号が得られるだけではない、想像以上にたくさんの経験ができる場所であると感じた。あなた色の3年間を過ごして、色んなことに挑戦してみてはいかがだろうか。

遠隔ピアノレッスンを多面的に支援するシステムを構築



遠隔ピアノレッスンのための教授支援基盤技術の構築

♡ 3

情報処理学会・学会誌「情報処理」
2022年8月15日 11:41



2021年度研究会推薦博士論文速報

【音楽情報科学研究会】



松井遼太

(MPLUSPLUS (株) 研究開発部 研究員)

■キーワード

HCI/エンタテインメントコンピューティング/音楽情報処理

【背景】 遠隔ピアノレッスンでは効率的な指導や学習が困難

【問題】 遠隔レッスン時のコミュニケーション支援と独習時の学習支援

【貢献】 遠隔と対面を越境できるピアノ学習支援システムの構築

本研究では、物理的に離れた2地点間における遠隔ピアノレッスンの実時間支援と、レッスン外の独習支援を統合する遠隔ピアノレッスン支援システムを構築し

た.

ヒューマンコンピュータインタラクション (HCI) の研究分野において、楽器やスポーツなど技能習得の学習支援に関する研究はこれまでも実施されてきた。一方、本研究はレッスンの支援に特化したシステムである点、遠隔と対面の学習環境を越境する学習モデルを提案している点において、新規性と独自性がある。

近年、ICT技術の発展に伴い、サテライト授業や会議システムなどの分野において遠隔・通信教育が一般化している。特に2020年以降は多くの会社や学校でリモートワークや遠隔授業が推進されてきた。ピアノ演奏においても、遠隔演奏試験や遠隔レッスンなどが音楽大学等の教育期間で実施されている。しかし、遠隔レッスンでは、楽譜が直接見えない、相手の身体動作を直接見ることができない、教師と生徒が対面している感覚が薄れる、という課題が生じる。そのため、対面レッスンのように円滑な指導や効率的な学習が難しい。これらの課題に対し、本研究では実時間の遠隔レッスンを支援する「遠隔レッスン時支援」と、レッスンから次のレッスンまでの期間における、生徒の自宅での独習時に利用することを想定した「独習時支援」の2つの側面からの支援を実現するシステムを構築した。

「遠隔レッスン時支援」に関する機能として、ピアノ教師の視線移動を模した複数カメラの自動スイッチング機能や、生徒の悪癖を楽譜上にアノテーションする悪癖アノテーション機能などが存在する^{1) ~3)}。これらを使用することで、レッス

ン中の指示出しやコミュニケーションが簡易となる。ユーザスタディでは、これらの機能を用いた提案群と、従来のビデオ通話のみを用いた比較群それぞれに遠隔レッスンを実施した。その結果、提案群はレッスンにおける課題曲の進行度合いが高く、また、指示語を用いた会話が有意に多かった。したがって、提案システムにより遠隔レッスンが効率化されたといえる。

「独習時支援」に関する機能として、課題曲の完成日を予測する機能や、記憶の定着方法に着目した学習支援機能などが存在する⁴⁾、⁵⁾。ユーザスタディの際に実施したアンケートでは、「楽しく学習ができた」などと肯定的な意見が多くみられた。そのため、本機能はレッスン間の独習時におけるモチベーション維持に有効であることが示唆される。

2つの機能を組み合わせて利用することで、レッスンの質の向上と、独習時の効果を高めることが期待できる。したがって、これら2つの支援側面は相乗効果的に作用するため、本研究が提案する各機能には、必然性があるといえる。

提案システムの登場によって、遠隔ピアノレッスンが実用的となり、学習者にとっては対面と等価な選択肢となり得る。つまり、提案システムはピアノレッスンを遠隔に限定するというものではなく、「遠隔と対面の境界を取り除く」ことができる。本研究では、遠隔と対面の境界を取り除く（遠隔と対面が等価である）ことを、「学習環境の越境」と定義している。したがって、提案システムは遠隔と対面

を越境した学習環境を提供できるといえる。また、提案システムは過去の学習を記録し、振り返ることで過去と現在を、現在の演奏から未来の完成日を予測することで、現在と未来を、それぞれ越境できる。学習環境が対面か遠隔か、という問題はどちらにも利点と欠点が存在する。本研究による遠隔を主とした学習環境の整備は、遠隔環境におけるピアノ学習の地位向上に寄与したといえる。



参考文献

- 1) 松井遼太, 竹川佳成, 平田圭二, : Tel-Gerich: 共同注視およびカメラスイッチングに着目した遠隔ピアノレッスン支援システム, ヒューマンインターフェース学会論文誌特集号, Vol.20, No.3, pp.321-332 (2018) (査読あり) .

- 2) Matsui, R., Takegawa, Y., Hirata, K. and Yanagisawa, Y. : Remote Piano Lesson System Considering Camera Switching, International Computer Music Conference, Paper ID:540 (2019) (査読あり) .
- 3) 松井遼太, 長谷川麻美, 竹川佳成, 平田圭二, 柳沢 豊 : ピアノ教師向け悪癖発見支援システムの設計と実装および評価, 情報処理学会論文誌特集号, Vol.61, No.4, pp.789-797 (2020) (査読あり) .
- 4) 松井遼太, 竹川佳成, 平田圭二, 柳沢 豊 : ピアノ学習における課題曲合格時期予測システムの構築, 情報処理学会論文誌特集号, Vol.63, No.3, pp.761-772 (2022) (査読あり) .
- 5) Matsui, R., Kumaki, M., Takegawa, Y., Hirata, K. and Yanagisawa, Y. : Proposal of a Piano Learning Support System Considering Gamification Based on Reconsolidation, 24th International Conference on Human-Computer Interaction (2022) (採録決定, 査読あり) .

(2022年5月31日受付)

(2022年8月15日note公開)

取得年月日 : 2022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

正会員

推薦文 [メディア知能情報領域] 音楽情報科学研究会

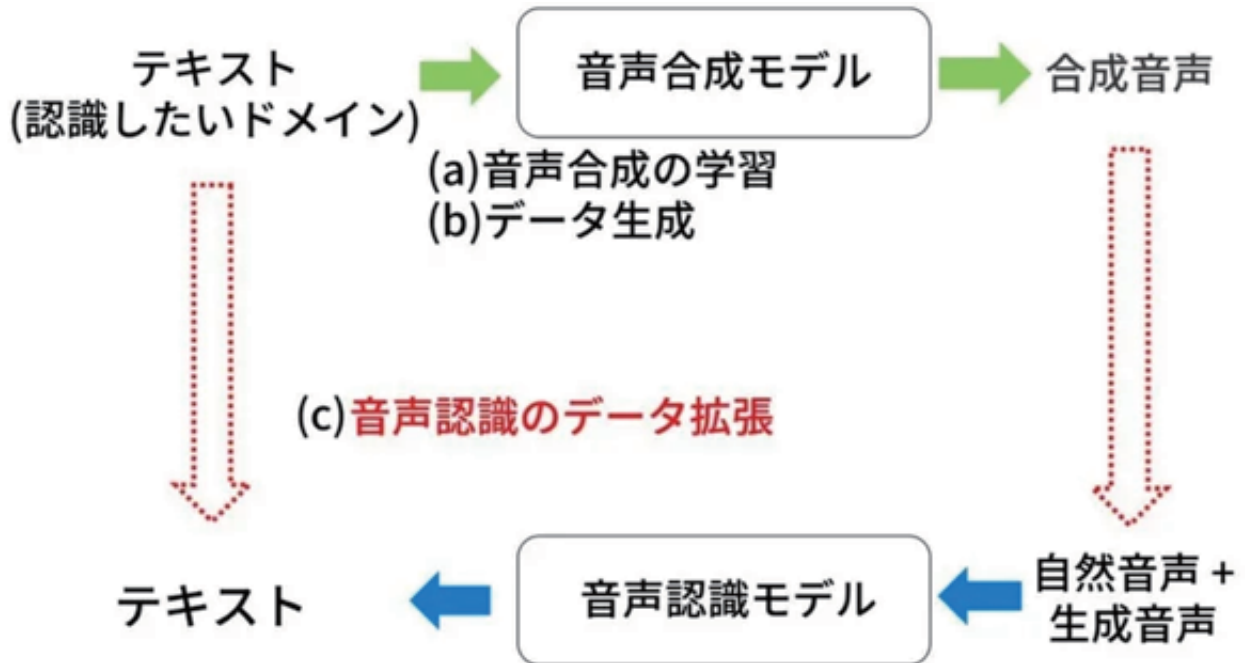
本論文では、遠隔地同士のピアノレッスンを効率的に支援するシステムを構築した。提案システムは、対面と遠隔の学習効率を等しくし、対面と遠隔を越境した学習環境を提供できる。また、過去のレッスンを振り返ることで過去と現在を、現在の演奏から未来の完成度を予測することで、現在と未来を、それぞれ越境できる。

研究生活 私は修士課程修了後に一度就職したのち、別の会社へ転職すると同時に博士課程へ進学しました。進学後は現職（MPLUSPLUS（株））の研究員を務めつつ、社会人博士としての生活も送っていました。修士時代から博士進学を考えてはいましたが、さまざまな事情により少々遠回りをしました。しかし、その間のすべての経験が貴重で、今でも自分の糧となっています。一見自分の研究とは関係がなさそうに思うことでも、さまざまな形で役に立つことはたくさんあると実感しました。

自分が心から楽しく、ワクワクするテーマを研究できたことを心から幸せに思います。また、これまで二足のわらじ生活を両立することができたのは、MPLUSPLUS

の皆様と研究室のメンバのおかげです。多くの議論を重ねることで、研究をブラッシュアップでき、また何よりも楽しい時間を過ごすことができました。

最後に、長年にわたりご指導をいただいた竹川佳成教授、平田圭二教授、研究遂行にあたり多くのアドバイスと技術的な指導をいただきました柳沢豊博士に、深く感謝いたします。



Data Augmentation Approaches for Automatic Speech Recognition Using Text-to-Speech



情報処理学会・学会誌「情報処理」
2022年8月15日 11:42



2021年度研究会推薦博士論文速報

[音声言語情報処理研究会]



上乃 聖

(名古屋工業大学大学院工学研究科 助教)

邦訳：音声認識のための音声合成を用いたデータ拡張手法

■キーワード

機械学習 / 音声認識 / 音声合成

【背景】 End-to-End音声認識は高精度な反面、音声と書き起こしのペアデータを

必要とする

【問題】 音声合成を用いたデータ拡張があるが、自然音声獲得可能時と性能の差が見られる

【貢献】 より効率的な音声合成を用いた音声認識のデータ拡張法の提案

音声認識は人の音声をテキストにマッピングする問題であり、深層学習の発展によって、実現するためのモデルが大きく変わった。従来は音声と音素をマッピングしたのちに音素とテキストを対応づけるというステップを個別で行っていたが、それらの役割を統合してモデル化するEnd-to-End音声認識と呼ばれる手法が主流となってきた。End-to-End音声認識はさまざまなモデルの提案により非常に高精度を達成し、また一括でモデル化できるため、高速に認識が可能になる。しかし、学習には大量の音声とその書き起こしのペアデータが少なくとも数百時間必要である。End-to-End音声認識の学習のために人間が書き起こしを用意することは膨大な時間とコストがかかる作業であり、特に対話の音声認識など認識したいドメイン（話題など）が決まっている場合にそのドメインにあったデータを用意することは難しい。この課題を解決するために、本研究では音声合成を用いた音声認識のデータ生成法について着目する。

音声合成を用いた音声認識のデータ生成には3つのステップが存在する。まず音声合成を学習するステップと、任意のテキストから学習した音声合成を用いて音声を生成するステップ、そして合成したデータと、一定量の人間の音声（自然音声）

を用いて音声認識を学習するステップがある。この枠組みは音声認識の、特に自然音声のデータに含まれていないドメインに対して、性能改善が見られる。しかし、その性能の改善幅は自然音声を対象のドメインに対して用意できた場合と比較し限定的であるという問題がある。

そこで、本研究ではより音声合成を用いた音声認識のデータ生成法について、より改善幅を自然音声を用意できた場合に近づけるような手法の提案・実験を行った。

本研究では以下の3つの手法を提案している。

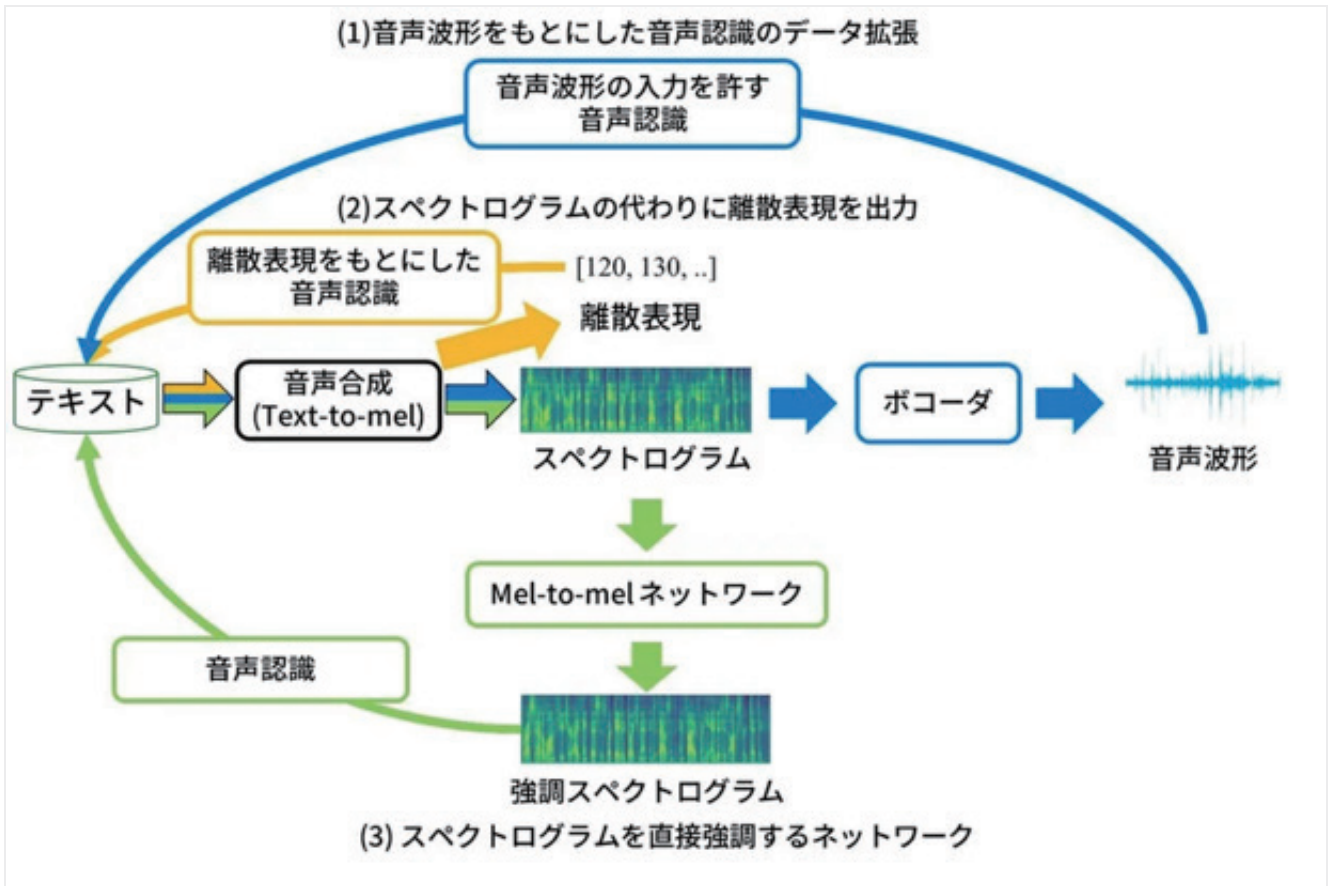
- (1) 音声波形をもとにした音声認識のデータ拡張
- (2) スペクトログラムの代わりに離散表現を出力
- (3) スペクトログラムを直接強調するネットワークの構築

(図表参照)

(1) の手法では、人間が聞くことのできる音声波形をベースにデータ拡張を行った。音声認識は通常音声波形からスペクトログラムと呼ばれる特徴を用い、テキストへと変換する。音声合成はテキストからスペクトログラムへ変換するText-to-melネットワークとスペクトログラムから音声波形へ変換するボコーダと呼ばれるネットワークの2つをもって実現される。合成された音声波形を活用するためにスペクトログラムではなく音声波形を入力とする音声認識を構築することでより効率的な拡張を行う。

(2) の手法では、音声合成がスペクトログラムを推定するのではなく、より簡略化された離散表現を予測する手法である。スペクトログラムや音声波形を推定することは、無限に近い値から推定する必要があるため非常に難しい問題となる。そのため、音声合成がその値自身を推定するのではなく、そこから抽出された離散表現を用い、限られた候補の中から推定するという手法を提案した。

(3) の手法では、生成されたスペクトログラムを直接強調するようなMel-to-melネットワークを構成することで生成スピードを高速にし、効果の高いデータ拡張を実現する。(1) の手法では、ボコーダを用いて音声波形を生成することでその性能が上がることを示したが、その生成に余計な時間がかかってしまうという問題がある。それは音声波形が非常に長く、音声認識のデータ拡張を行う際には音声波形は必ずしも必要ではないためである。そこで、音声認識モデルにとって必要なスペクトログラムを直接強調する手法を提案した。



参考文献

- 1) Ueno, S. and Kawahara T. : Phone-informed Refinement of Synthesized Mel Spectrogram for Data Augmentation in Speech Recognition, In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.8572-8576 (2022).
- 2) 上乃 聖, 河原達也 : 音声認識のデータ拡張のための合成音声の周波数スペクトログラム強調, 情報処理学会研究報告, SLP-139-28 (2021).
- 3) Ueno, S., Mimura, M., Sakai, S. and Kawahara, T. : Synthesizing Waveform Sequence-to-sequence to Augment Training Data for Sequence-to-

sequence Speech Recognition, Acoustical Science and Technology, Vol.42, Issue 6 pp.333-343 (2021).

4) Ueno, S., Mimura, M., Sakai, S. and Kawahara, T. : Data Augmentation for ASR Using TTS Via a Discrete Representation, In Proc. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), pp.68-75 (2021).

5) 上乃 聖, 三村正人, 河原達也 : vq-wav2vecによる離散idを扱う音声認識のデータ拡張, 日本音響学会研究発表会講演論文集, 1-2-16 (春季 2021).

6) Ueno, S., Mimura, M., Sakai, S. and Kawahara, T. : Multi-speaker Sequence-to-sequence Speech Synthesis for Data Augmentation in Acoustic-to-word Speech Recognition, In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.6161-6165 (2019).

(2022年5月31日受付)

(2022年8月15日note公開)

取得年月日 : 2022年3月

学位種別 : 博士 (情報学)

大学 : 京都大学

正会員

推薦文【メディア知能情報領域】音声言語情報処理研究会

深層学習に基づく音声認識では、モデルの訓練に膨大な量の声と書き起こしテキストのペアを要する。本論文は、テキストからの音声合成によって学習用ペアデータを大量に人工的に作成して音声認識の精度を向上する手法に焦点をあて、合成側と認識側を統合して情報の損失を防ぎ頑健な認識を実現するための種々の手法を提案・比較している。

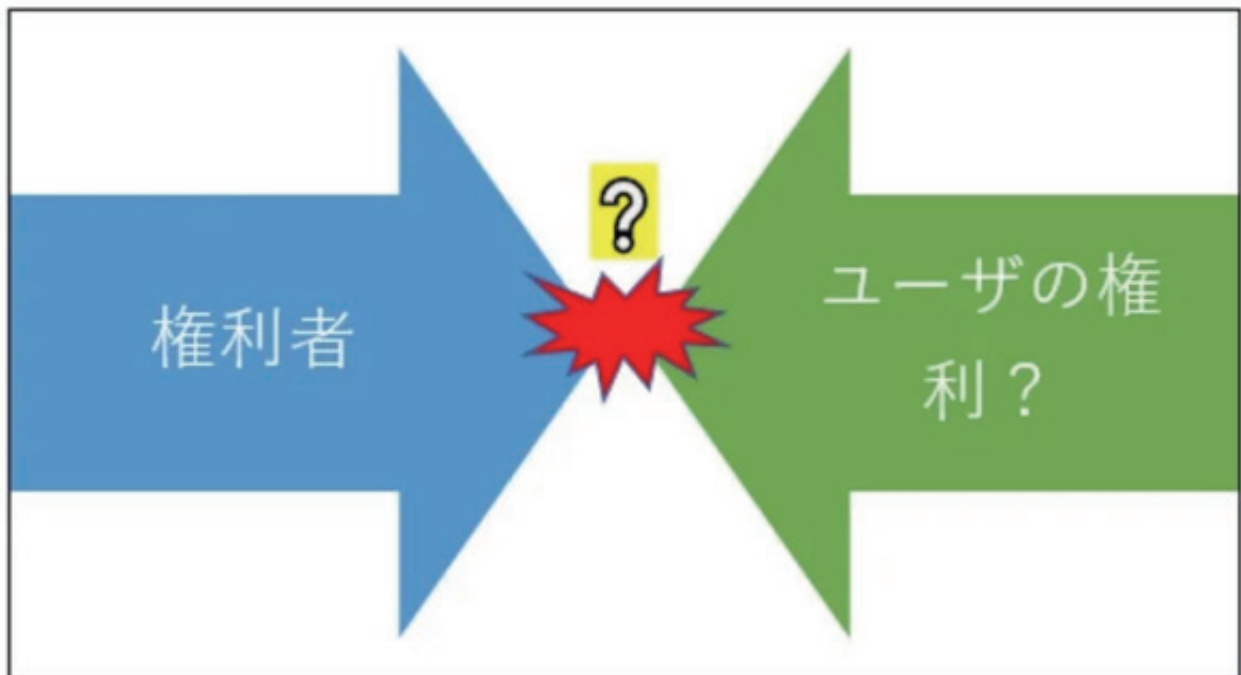
研究生活 私が博士課程に入学するあたりにディープラーニングで構成される音声認識や音声合成システムが提案されてきました。その中で音声合成が音声認識のデータの拡張として役に立つことを知り、これをもっと深めたいと思い博士課程に進学しました。

現状博士課程を考えていないという方も多くいらっしゃると思います（というよりも博士課程進学者の人数を見ればそちらの方が多いでしょう）。そのため修士課程から博士課程へ、社会人の方が博士課程へ進学しようとする方にとって、たいてい博士課程は前例が少ない事例になると思います。就職すること、働き続けることは大体先輩の話や、自身の経験から大体こんなものであるという想像はつくと思います。しかし、博士課程は意識的に集めないと解像度の高い想像を行えません。博士課程にはしっかりとした計画があった方が有利です。もうすでに進学を決めている方も含めて、博士課程に進んだ方の話を聞いたり、自身のしたい研究の価値など

を考えて選択の一助としたり，博士課程を有利に進めるための準備をしてみてください。

▲ **新型コロナウイルスに関する内容の可能性のある記事です。**

新型コロナウイルス感染症やコロナワクチンについては、必ず1次情報として厚生労働省や首相官邸のウェブサイトなど公的機関で発表されている発生状況やQ&A、相談窓口の情報もご確認ください。*非常時のため、すべての関連記事に本注意書きを一時的に出しています。



A Justification for the Limitation of Copyright Rights in Digital Archives: Research introduction for the Theory of "User's Rights"



情報処理学会・学会誌「情報処理」

2022年8月15日 11:43



2021年度研究会推薦博士論文速報

[電子化知的財産・社会基盤研究会]

栗原 佑介

(慶應義塾大学SFC研究所 上席所員)

邦訳：デジタルアーカイブにおける著作権の権利制限：「ユーザの権利」論序説

■キーワード

デジタルアーカイブ／著作権／ユーザの権利

【背景】 デジタルアーカイブ政策の法的正当性が必ずしも明らかでない

【問題】 デジタルアーカイブという行為の法的意義

【貢献】 記憶する権利としてのデジタルアーカイブ権の確立を提案

1. デジタルアーカイブを法的に検討する意義

文化資源のデジタルアーカイブ（以下「DA」という。また、この定義自体、重要な論点だが、本研究では、「（デジタル化されているか問わず）文化資源を対象に、これをデジタル形式で収集、保存、組織化し、公開・提供行為まで含む一連の行為」とした）が国内外で進められている。最近では、新型コロナウイルスの影響で美術館を含む博物館が閉鎖するなど、文化資源へのアクセスが緊急時には困難になることが顕在化した。

DAは、地理的、時間的距離を超えて、文化資源にアクセスする手段を提供する。そのため、世界中でデジタルアーカイブ政策が進められている。日本でも、2020年8月に、ジャパンサーチが正式公開された。

しかし、このDAによる恩恵は国民が受けるが、この権利は保障されているのだろうか。仮に、国立国会図書館が、パンデミック以外の理由で、DAされたデータを公表しなくなったとき、どうなるか。こうした疑問から、「DA権」はあり得るのか、あり得るとしたら、何が根拠になるのか、本研究は我が国著作権法を対象に、検証した。

2. 検討方法・対象

具体的なアプローチとして、「デジタル化」の権利と「アーカイブ」の2つに分け、それぞれから検討した。また、デジタル化された情報を享受する利益を法解釈によって明らかにすることを試みた。

アーカイブに関しては、上記のDAのメリットを前提に、近年、著作権法の領域で論じられ始めた著作物の「ユーザの権利」（端的には、他人の著作物を利用する権利。もっとも、この見解でも他人の著作物利用を無制限に認めるものではなく、著作権は、ユーザの権利との調整が必要であると考えられる立場一般を指す）に着目し、諸外国の知的財産法制や裁判例による権利制限、パブリックドメインの設計を分析し、文化資源に対するアクセスを法的に保障すべき局面を横断的に展望した。

そして、従来は「権利の空白」とされてきた領域について、「ユーザの権利」の表徴として再評価を試みた。

3. 結論と課題

DAを享受する利益は、法解釈によって、さらに2つのアプローチによった。1つは、知的財産法に関する法解釈である。著作権法の権利制限規定の中に読み込むことができるだけでなく、著作権を始めとする知的財産法によって「保護されない」領域を確保し、自由利用を促すこともまた、知的財産法の役割であり、その自由利用領域の利活用こそが、ユーザの権利として、DAを享受する利益に繋がることを明らかにした。

もう1つは、憲法25条を根拠として展開されてきた「文化権」の一内容としてのDAを享受する利益が認められることを示した。

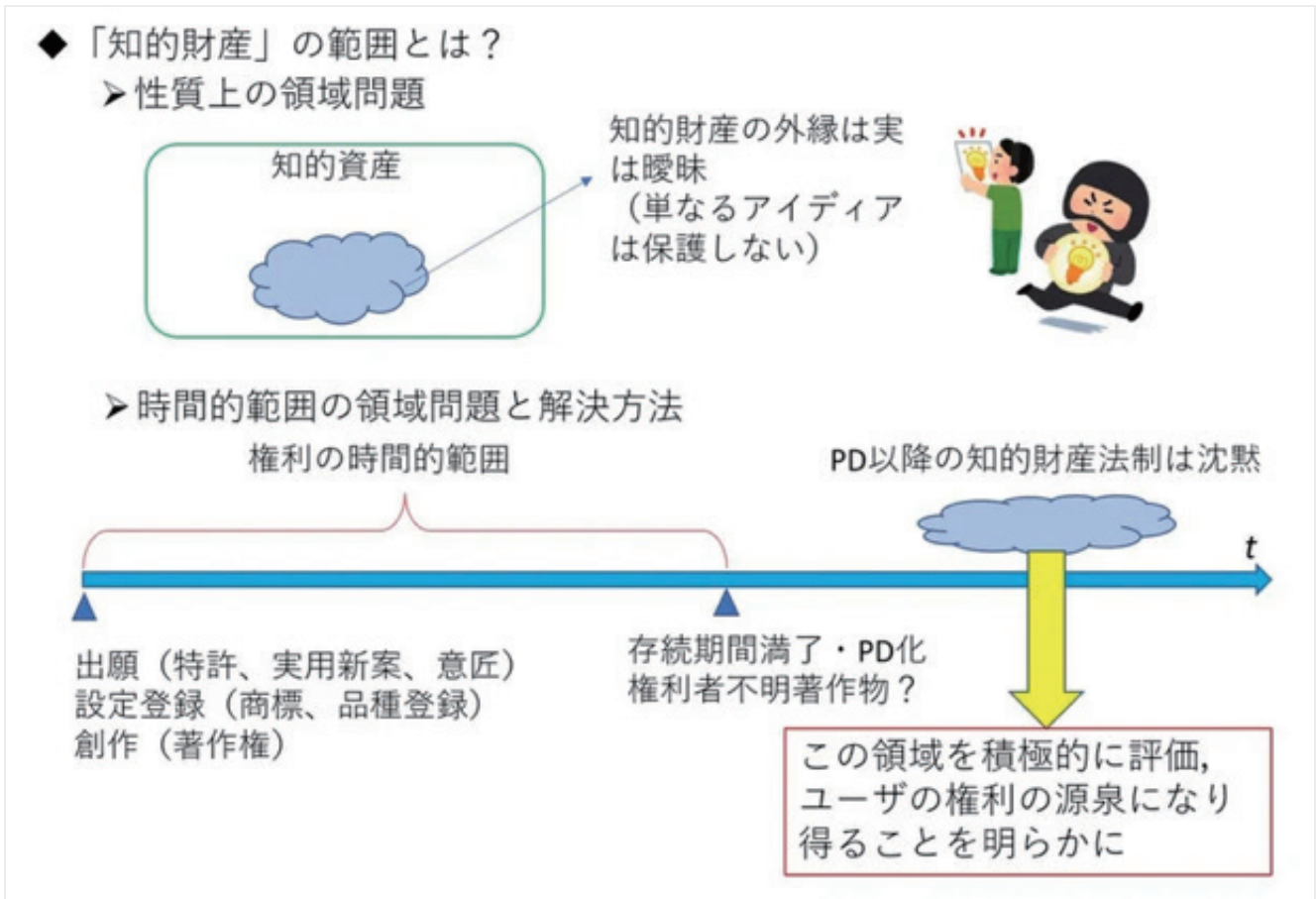
結論として、条約との関係でデジタル化に関する明文規定はあるものの、具体的な請求権としての性質を持つに至っていないこと、他方でアーカイブ（保存）という法的保護の必要性が明らかになり、解釈上、ユーザの権利や文化権の一内容として、DA権を認め得るといえる。それと同時に、DA権は、現代の文化的、社会的背景から確立されるべき権利である。たとえば、保存状態の悪いコンテンツのダークアーカイブは、公開は別として、その保存は、正当化し得る。

もっとも、この具体的内容として文化助成としての権利の側面も認められるものの、その正当化は今後の課題とした。

4. 今後の展開

本論でも強調したが、新型コロナウイルスは、DAを加速させ、すでに論文の一部は、情報が古くなっている。たとえば、博物館法は改正により、博物館事業にDAが明記された（同法3条1項3号）。また、Cultech（culture technology）の用語も普及しつつある中、DAは、メタバースでの展開にも親和性がある。さらに、プライベートアーカイブでは、NFT化も選択肢になり得る。

さらに、博物館の持続可能性の観点からは、ある程度のマネタイズも検討が必要となる。DA権の文化助成含むマネタイズは、今後の課題であるが、その根拠となり得る、DA権の普及に努めることが、今後の使命であると思っている。



(2022年5月29日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士 (学術)

大学：慶應義塾大学

学生会員

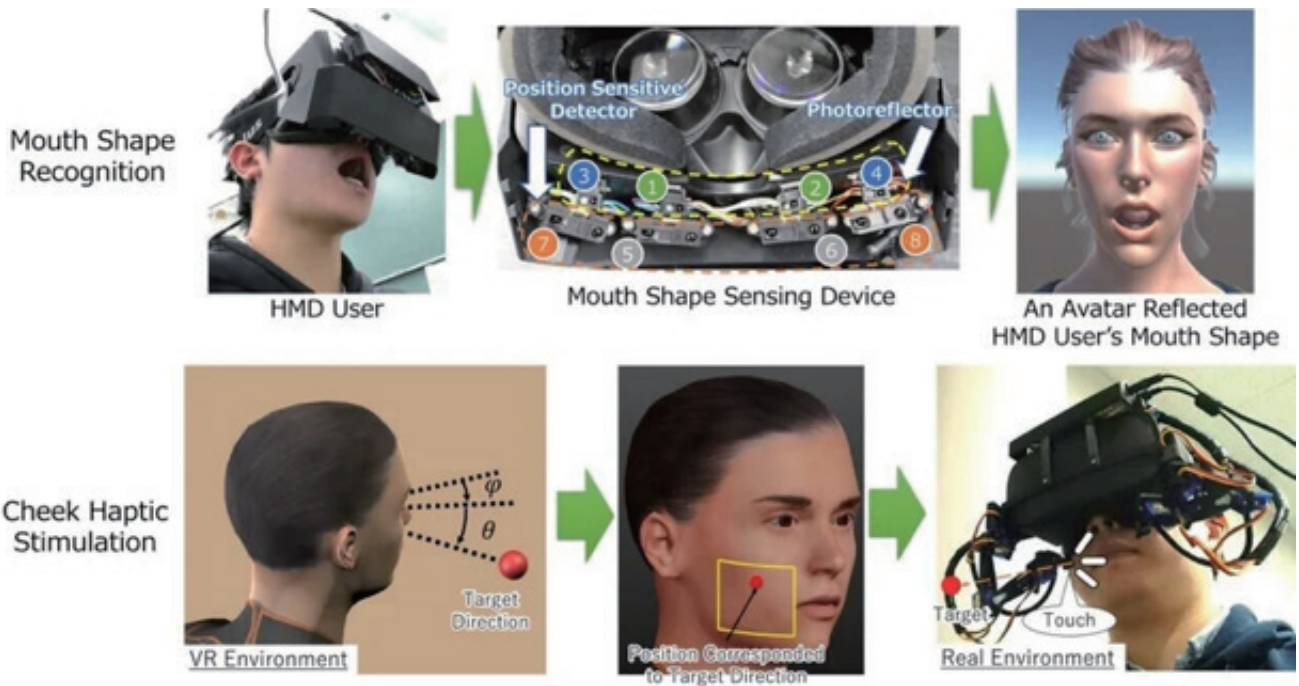
推薦文 [メディア知能情報領域] 電子化知的財産・社会基盤研究会

デジタルアーカイブとは、幅広い知的資源をデジタル化してネットワーク公開する取り組みです。これを推進するにあたっては、知的資源の著作権が障害になっているという指摘があります。この論文では、国民や市民の文化を享受する権利を起点として、保護と利活用を両立する制度の在り方を、多面的に検討しています。

研究生活 研究テーマへの関心は、2012年頃の東日本大震災の災害（デジタル）アーカイブの立ち上がりと、同時期になされたインターネット等を活用した著作物利用の円滑化を図るための措置を規定した2010年改正が契機になっている。このころから、法学者ではなくアーカイブ研究者から、DAと著作権の問題が提起され、私は法学サイドからこの論点を追うようになった。

私は、最初の修士課程から博士号取得までの約10年間、すべて社会人学生としていわば片手間で研究・論文執筆を行ってきた。これを推奨するつもりはないが、専業学生だけが選択肢ではないこと、特に博士課程は、特定の時間に拘束されることはほばないので、進学にあたって必要なものは「覚悟」だけである点は強調したい。むしろ文献研究を中心とした研究は、この数年でデータベースが整備され自宅でも

比較的容易にできるようになり、きわめて社会人にとって有利になっている
(2021年著作権法改正の国立国会図書館の個人送信により、研究環境はますます
改善された)。社会人の博士課程進学の参考になれば僥倖である。



Embedded Facial Surface Sensing and Stimulation: Toward Facial Surface Interaction in Virtual Environment

♡ 3

情報処理学会・学会誌「情報処理」
2022年8月15日 11:44



2021年度研究会推薦博士論文速報

[エンタテインメントコンピューティング研究会]

中村 文彦

(日本学術振興会 特別研究員 (PD) /

慶應義塾大学大学院 理工学研究科 特任助教)

邦訳：バーチャル環境におけるフェイシャルサーフェスインタラクションに向けた組込型顔表面計測・刺激に関する研究

■キーワード

バーチャルリアリティ／顔面インタラクション／HMD組込式システム

【背景】 高品質な視覚・聴覚ディスプレイを持つVR HMDが普及

【問題】 HMDユーザの入力や触覚フィードバック提示が限定的

【貢献】 低次元のデータを用いた口周辺の形状の認識と認識した形状に応じた触覚提示

頭部運動に応じて映像を提示できるヘッドマウントディスプレイ (HMD) の登場によって、ユーザが簡単にバーチャルリアリティ (VR) 空間に没入してインタラクションできるようになった。インタラクションにおいては、入力のためにユー

ザの行動を計測装置と、出力のためにユーザの感覚器官に対して情報提示装置が用いられる。典型的なHMDを用いたシステムでは、頭部と手の位置・姿勢を計測し、その情報に応じた視覚・聴覚情報を提示するシステムを備えている。しかし、HMDを装着すると、ディスプレイによって実環境が見えなくなるため、キーボードやマウスといった入力インタフェースが使い辛い。また、多くのHMDは高品質な視覚・聴覚ディスプレイを備えているが、触覚に関する品質は高くない。マルチモーダル情報（視覚・聴覚・触覚といった複数の感覚情報の提示）の提示は体験の質の向上につながるため、触覚の機能の統合は重要である。顔面への触覚提示技術によって、バーチャル環境でのアバターとの顔を用いたインタラクション（例：頬をつつく、頬を撫でる等）が可能になる。

そこで、適切な入力とインタラクションの質の向上のために、顔面を入出力インタフェースとして用いることが考えられる。顔面は細部まで制御することが可能であり、また、触覚刺激に対して敏感である。しかし、顔面は目や耳や鼻といった重要な感覚器官が集中しているだけでなく、皮膚が傷つきやすい。特に、口周辺や唇は皮膚の層が薄く、わずかな力でも傷つく恐れがある。さらに、顔面は多くの筋肉で制御され、複雑かつ繊細にその形状を変えられる。そのため、顔面を用いたインタラクションを実現するためには、顔面形状の高速な計測・認識と、認識した形状に基づく触覚刺激の提示が重要である。

本研究では、VR環境におけるHMDユーザの顔面を用いたインタラクションのた

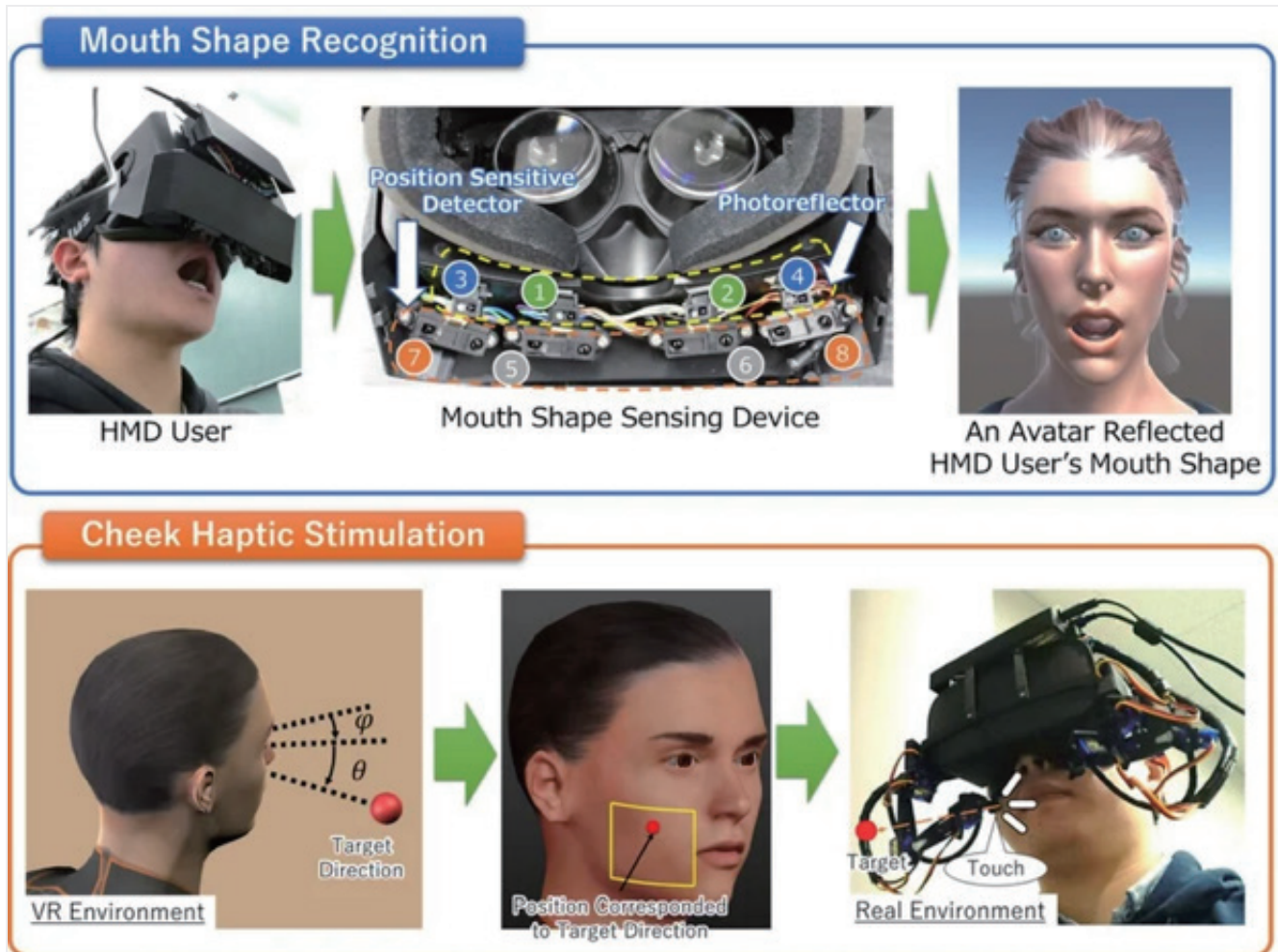
めに、1) 口元形状認識技術と、2) 頬への触覚提示技術を提案した。

1) HMD組込式光センサを用いた口元形状認識

HMDに光センサを組み込むことで、ユーザの口元形状を計測し、低次元のセンサ値を機械学習することで口元形状を認識する手法を提案した。また、口元形状を計測したセンサデータにユーザの音声から認識した母音をラベルとして付与することで、訓練データの自動ラベリング手法を提案した。実験から、本手法は高い精度で日本語5母音と口を閉じた状態の6個の口元形状を認識できることが示された。さらに、口周辺の認識結果に基づいて口形状を合成し、その結果をアバターに反映するアプリケーションを開発した。

2) 頬への触覚刺激によるVR空間での空間方向情報提示

HMDに取り付けた小型ロボットアームを用いて、空間方向情報に対応した触覚刺激を頬に提示することで空間方向誘導を行うシステム「Virtual Whiskers」を提案した。ロボットアームの先端に取り付けた反射型光センサを用いて、低次元のセンサ情報から頬の表面形状を計測・推定した。推定した頬の表面上の空間の方向情報と対応した位置を計算し、その位置をロボットアームで刺激することで、頬上に方向情報を提示した。実験から、本手法はVR空間で直感的にユーザを誘導できることを示した。



参考文献

- 1) 中村文彦, 鈴木克洋, 正井克俊, 伊藤勇太, 杉浦裕太, 杉本麻樹: HMD組み込み式光センサによる口元形状認識と母音認識を用いた訓練データのラベリング手法, 日本バーチャルリアリティ学会論文誌, Vol.22, No.4, pp.523-534 (2017), DOI: https://doi.org/10.18974/tvrsj.22.4_523
- 2) Nakamura, F., Suzuki, K., Masai, K., Itoh, Y., Sugiura, Y. and Sugimoto, M. : Automatic Labeling of Training Data by Vowel Recognition for Mouth Shape

Recognition with Optical Sensors Embedded in Head-Mounted Display, In Proceedings of ICAT-EGVE 2019 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, pp.9-16, Tokyo, Japan (2019), DOI:

<https://doi.org/10.2312/egve.20191274>

3) Nakamura, F., Verhulst, A., Sakurada, K., Sugimoto, M. : Virtual Whiskers: Spatial Directional Guidance using Cheek Haptic Stimulation in a Virtual Environment, In Proceedings of Augmented Humans Conference 2021, pp.141-151, Rovaniemi, Finland (2021), DOI:

<https://doi.org/10.1145/3458709.3458987>

4) 中村文彦, Adrien Verhulst, 櫻田国治, 福岡正彬, 杉本麻樹 : Virtual Whiskers: 頬への触覚刺激によるVR環境での方向情報の提示, 日本バーチャルリアリティ学会論文誌, Vol.26, No.4, pp.298-309 (2021). DOI:

https://doi.org/10.18974/tvrsj.26.4_298

5) Nakamura, F., Verhulst, A., Sakurada, K., Fukuoka, M., Sugimoto, M. : Virtual Whiskers: Cheek Haptic-Based Spatial Directional Guidance in a Virtual Space, ACM SIGGRAPH Asia 2021 XR, Article 17, pp.1-2, Tokyo, Japan (2021), DOI: <https://doi.org/10.1145/3478514.3487625>

6) Nakamura, F., Verhulst, A., Sakurada, K., Fukuoka, M., Sugimoto, M. : Evaluation of Spatial Directional Guidance Using Cheek Haptic Stimulation in a Virtual Environment. Frontiers in Computer Science, Vol.4 (2022). DOI:

<https://doi.org/10.3389/fcomp.2022.733844>

(2022年5月30日受付)

(2022年8月15日note公開)

取得年月日：2022年3月

学位種別：博士（工学）

大学：慶應義塾大学

推薦文 [メディア知能情報領域] エンタテインメントコンピューティング研究会

没入型バーチャルリアリティ環境を構成する頭部装着型ディスプレイを拡張し、口元形状の識別・再構成をする組み込み型計測と、ロボットアームによる頬への触覚提示を行うことで、顔の皮膚表面を介したバーチャル環境とユーザのインタラクションを実現する枠組みを提案しており、今後の研究の発展が期待される。

研究生活 私は、修士課程修了後に2年間ほど会社に勤めていましたが、学識を深めたいと考え、退職して博士課程に入学しました。博士課程では研究テーマとし

て、HMDを着用して没入したVR環境における顔面を用いてインタラクションするために必要な技術の開発を行いました。研究では、人の顔面形状の計測を基にしたものであったのですが、人の顔面形状の違いに苦労させられました。この問題に取り組むために、多くの研究者と議論しました。この議論のおかげで、楽しく飽きることなく研究を進められたのかと思います。

最後に、本研究を推進するにあたって、熱心で丁寧なご指導、ご鞭撻をいただきました慶應義塾大学工学部情報工学科 杉本麻樹先生に心よりお礼申し上げます。