NTT

変化する現在（いま）
持続する未来（あす）

NTT コミュニケーション科学基礎研究所

オープンハウス2022

6/2 thu ◎ 3 fri

6/2 12:00よりWeb公開

On the Web

今年度は、特設ウェブサイトにて最新の研究成果を公開いたします。
※NTT京阪奈ビルでの現地開催は行いません。
https://www.kecl.ntt.co.jp/openhouse/2022/

こちらのQRコードからも
ホームページへアクセスできます。

「オープンハウス2022」へようこそ

NTT コミュニケーション科学基礎研究所
所長　**納谷　太**

昨今の新型コロナウイルス感染症のパンデミックは私たちの生活様式や社会活動を激変させました。こうした中、多様な価値観を持つ人と人、人と機械、人と社会を繋ぐコミュニケーションと、これを支える技術はますますその重要性を増しています。

私どもNTT コミュニケーション科学基礎研究所では、「こころまで伝わる」コミュニケーションの実現に向けて、人間と情報の本質に迫る基礎理論の構築と、社会に変革をもたらす革新技術の創出をミッションとして、世界トップレベルの基礎研究に持続的に挑戦してまいりました。

当研究所は昨年2021年7月に創立30周年の節目を迎え、2021年10月にはNTT R&Dにおける「知の泉」の源を一層強化すべく、新たに現代数学の基礎理論を研究するバーチャルな組織「NTT基礎数学研究センタ」を研究所内に設立しました。

「オープンハウス2022」では、上記、基礎数学研究センタにおける取り組みをはじめ、メディア処理、データと機械学習（AI）、人間科学や脳科学に関する最新の成果について、講演動画・ポスター展示・オンラインデモ等を交えてわかりやすくご紹介いたします。

新型コロナウイルス感染症対策のため、今年も残念ながらオンライン開催となりますが、激変する人・社会・環境に寄り添いつつ、未知なる真理の探求と、より良い未来社会の創造に向けて、みなさまとの議論と交流のきっかけの場となることを祈念しております。所員一同、多数のみなさまのご来場をお待ちしております。

目　次

ごあいさつ …… 2

●所長講演

変化する現在(いま)に適応し、持続する未来(あす)を切り拓くコミュニケーション科学　～人・社会・環境との調和と共生をもたらす技術の創出～ …… 5

●招待講演

コミュニケーション行動から探る動物の心　⊕『パネルディスカッション』 …… 9

●研究講演

あなたの声を「すぐそば」品質で聴くAI　～遠くからでも近接マイク品質で混ざった音を聞き分ける革新的音響処理技術～ …… 11

あなたの心臓をデジタル化　～モバイルセンシングを活用したパーソナル心臓モデリングとその応用～ …… 13

デジタルツインでモビリティ群を賢く制御する　～分散深層学習がもたらす未来の可能性～ …… 15

数学はどこから来たのか 数学はどう進むのか 数学はどこへ行くのか

～対称性を駆使した解析学と幾何学的手法による整数論の未解決問題と量子計算への挑戦～ …… 17

●研究展示 …… 19

データと学習の科学

01 信号機を使わないリアルタイム分散交通制御　デジタルツインを介した集合知学習技術 …… 21

02 光回路で構成されるAIに適した学習法　細層構造を持つ光ニューラルネットワークの高速学習 …… 22

03 複数のAIがどう寄れば文殊の知恵？　汎化性能を向上させるためのベイズ的アンサンブル学習 …… 23

04 ニューラルネットを軽量・高速化する学習技術　ニューラルネットの枝刈りにおける反復ランダム化技術 …… 24

05 その出来事、どのように拡散する？　時変Hawkes過程を用いて状態の変化を推定 …… 25

コミュニケーションと計算の科学

06 災害復旧を加速する効率的な避難所運営　被災者の順次帰宅を考慮した避難所計画の最適化 …… 26

07 あなたの好きな表現で翻訳します　指定した語句をかならず使うニューラル機械翻訳 …… 27

08 ユーザに平等かつ混雑しないインフラをめざして　組合せ混雑ゲームの均衡最適化 …… 28

09 機械翻訳の間違いを探します　原文と訳文の単語対応に基づく機械翻訳の後編集支援 …… 29

10 高齢者にとって聞いて分かりやすい話し方　分かりやすさにつながる声と言葉の使い方の解明 …… 30

11 移りゆく車窓の風景についてロボットと話そう　移動中に得られる画像・周辺情報を話題とする雑談対話技術 …… 31

12 量子技術による強い攻撃でもその暗号安全ですか　ランダムハッシュ関数の衝突を発見する量子アルゴリズム …… 32

13 数の不思議はどこから生まれる？　一般化モチーフ理論を用いた新たな数論的現象の解明 …… 33

メディアの科学

14 「え、なに？どういうこと？」長い話まとめます　音声認識誤りに頑健な音声要約 …… 34

15 写真の奥行を推定し、ボケ度合を自在に変えられます　写真群のみから未知の奥行とボケ効果を学習する深層生成モデル …… 35

16 音と電気の信号で心臓の働きを見守ります　マルチチャンネルの心音と心電からの心臓内部の活動推定 …… 36

17 声で顔画像の表情を動かす　クロスモーダルアクションユニット系列推定と顔画像変換 …… 37

18 いつでもどこでも快適な「見えやすさ」を　コンテンツ適応的な視認性予測に基づく画像ブレンディング …… 38

人間の科学

19 まなざしや声かけが赤ちゃんの学びを助けます　社会的手がかりが乳児期の物体学習に及ぼす影響の解明 …… 39

20 なぜ人は接触確認アプリの利用をためらうのか　COCOAの普及に影響を与える社会的要因 …… 40

21 ノビのある速球は錯覚？　VRで投球の質感を操作する …… 41

22 脳計測で見えたeスポーツ熟練者の勝負術　前頭脳領域の脳波とパフォーマンスの相互関係 …… 42

23 人工ニューラルネットワークで紐解く聴覚システム　人工内耳と両耳情報処理へのより深い理解のために …… 43

24 マインドフルネス瞑想はストレスをどう減らす？　二つの瞑想法がもたらす自律神経系・内分泌系の状態変化 …… 44

25 ウェルビーイングをいろいろな側面から測ります　心身の状態変化・価値観・自己観から捉えたウェルビーイング …… 45

26 壁が動くと速く歩く？　視野の動きに基づく歩行速度制御の仕組みの解明 …… 46

27 指先で感じる“こっち”　引かれる方向感覚を生成する脳の情報処理 …… 47

28 みんな何を触りたいの？　大規模なTwitterデータを活用した触りたさの理解 …… 48

29 瞳が語るこころの状態　瞳孔径から読み取る知覚の変化 …… 49

●講　演

6月2日［木］

所長講演
（13:00～13:30）

変化する現在（いま）に適応し、持続する未来（あす）を切り拓くコミュニケーション科学
～人・社会・環境との調和と共生をもたらす技術の創出～

NTT コミュニケーション科学基礎研究所 所長　**納谷 太**

招待講演
（13:40～15:10）

コミュニケーション行動から探る動物の心
帝京大学 先端総合研究機構 教授　**岡ノ谷 一夫**

＋『パネルディスカッション』
パネリスト：**岡ノ谷 一夫**（帝京大学）、**柏野 牧夫**、**小林 哲生**、**廣谷 定男**（NTT）

研究講演
（15:30～16:10）

あなたの声を「すぐそば」品質で聴くAI
～遠くからでも近接マイク品質で混ざった音を聞き分ける革新的音響処理技術～

メディア情報研究部　**中谷 智広**

6月3日［金］

研究講演
（11:00～11:40）

あなたの心臓をデジタル化
～モバイルセンシングを活用したパーソナル心臓モデリングとその応用～

メディア情報研究部/バイオメディカル情報科学研究センタ　**柏野 邦夫**

研究講演
（13:00～13:40）

デジタルツインでモビリティ群を賢く制御する
～分散深層学習がもたらす未来の可能性～

協創情報研究部　**丹羽 健太**

研究講演
（14:00～14:40）

数学はどこから来たのか 数学はどう進むのか 数学はどこへ行くのか
～対称性を駆使した解析学と幾何学的手法による整数論の未解決問題と量子計算への挑戦～

基礎数学研究センタ　**若山 正人**

変化する現在（いま）に適応し、持続する未来（あす）を切り拓くコミュニケーション科学

～人・社会・環境との調和と共生をもたらす技術の創出～

Communication science that adapts to the changing present and creates a sustainable future
– Aiming to create technology that brings harmony and symbiosis with people, society, and the environment –

NTT コミュニケーション科学基礎研究所 所長
納谷　太
Futoshi Naya

はじめに

CS研は1991年7月に京阪奈に設立されました。当初、国際電気通信基礎技術研究所（ATR）内に間借りし、2つの研究グループからスタートしたCS研は、現在、京阪奈と厚木の2拠点において総勢約150名の所員が集う組織となりました。設立から現在に至るまで、さまざまな形で支えていただきました皆さまにこの場を借りて御礼申し上げます。また、この30年間のあゆみを振り返り、昨年のオープンハウスの開催に合わせて、創立30周年記念ウェブサイトを開設しました[1]。設立当初からの主な研究成果をまとめていますので、是非ご覧いただければ幸いです。

この30年間の研究テーマの変遷を振り返りますと、人と人とのコミュニケーションを理解することを中心として始まったCS研の研究は、見る・聞く・話すといった人と同等の能力をコンピュータに持たせるメディア処理や、量子情報理論や機械学習・データ科学などの「人の能力に迫り凌駕する研究」と、人のさまざまな感覚運動能力のメカニズムを追求する人間科学および、トップアスリートなどの優れた認知能力の解明をめざした多様脳科学などの「人を深く理解する研究」として発展してきました。これらの研究は、それぞれの年代における技術的・社会的な背景を反映して変化し発展しながらも、コミュニケーションの本質を理解するというCS研設立当初からの研究理念が貫かれ、人や社会に寄り添う技術の創出に向けて継続されてきました。以下では、それぞれにおける最近の研究事例のいくつかをご紹介します。

人の能力に迫り凌駕する

深層学習技術の登場で、見る・聞く・話すといったいわゆるメディア処理に関するAI技術は格段に進歩しました。しかし、一般的に深層学習では膨大な学習データが必要です。一方で、人間は少ない情報しか得られない場合においても、これまでの経験からの類推や、欠けている情報を他の情報から補完するなどにより、柔軟かつ高度な推論を行うことができます。

例えば、写真は三次元の空間情報を二次元画像に変換したものですが、人間はこれまでの経験に基づき、写真を見ただけで被写体の持つ形状や奥行きなどの三次元情報をある程度推測できます。通常、このような能力をコンピュータで学習するには、三次元情報を深度センサーやステレオカメラなどの特殊な装置で計測した入力データと、写真として撮影された二次元画像としての出力データをペアとして大量に取得し用意せねばなりませんが、これには膨大なコストを要します。CS研では、このようなデータ取得上の課題を解決するため、カメラの持つ光学的な制約として、絞りとボケの関係を考慮した新しい深層学習技術「Aperture Rendering GAN」[2]を提案しました。本技術により、インターネット上にある公開画像などの一般的な写真群（二次元情報）のみから、奥行きやボケ効果といった三次元情報を学習することが可能です。

人間は、複数の人が同時に話している場合においても、聞きたい人の声に集中して聞き分ける能力（選択的聴取）を持っています。CS研では、これまで音声のみを用いて聞きたい人の声を聞き分ける技術SpeakerBeamを提案してきましたが、最近では、音声に加えて話者の映像を組み合わせることにより、人間のように複数の手がかりを活用して選択的聴取を実現する「マルチモーダルSpeakerBeam」[3]を提案しました。本技術により、声質が似通った複数話者が存在する場合には唇の動きを主な手がかりとし、逆に唇の映像が得られない場合には音声を主な手がかりとすることにより、頑健かつ高精度な話者の特定が可能になります。本研究は、音声だけでなく、消防車の音や犬の鳴き声など、注目すべき音だけを聞

き分けるユニバーサル音抽出技術へと発展しつつあります。

CS研では、1990年代から人と自然に会話する対話システムの研究を続けてきました。当初は予約や検索などの特定の目的に特化した対話システムの研究が中心でしたが、最近では目的によらず広い話題を扱いながら、自然な応答ができる雑談対話システムの研究を進めています[4]。昨年は、京阪奈CS研の所在する京都府精華町役場の協力の下、窓口案内や観光案内などの業務を行いつつ、雑談対話も楽しめるAIの実証実験[5]を開始したほか、Web等で収集した超大規模対話データと深層学習を組み合わせた日本語最大規模のTransformer対話モデルを無償公開しました[6]。今年のオープンハウスでは、車を運転中の車窓から見た画像や周辺情報を話題とした対話システムを展示しています。今後は、対話相手の嗜好などを記憶し、対話内容の一貫性を保ちながら、人に寄り添った対話を継続できるシステムの構築に向けた研究に取り組んでいます。

人に寄り添うという点では、機械学習の分野においても進展があります。例えば、融資承認や人材採用など、人を対象とした意思決定を機械学習によって行う場合、単純に予測精度のみを優先する従来の機械学習技術では、性別・人種・障がいなど、人間が持つ機微な特徴に関して不公平な予測になってしまう可能性があります。因果関係に基づく公平・高精度な機械学習[7]は、不公平さに関する事前知識を、特徴・予測結果間の因果関係（因果グラフ）としてモデル化することにより、個々人に対して公平かつ高精度な予測を実現しました。このように、人の置かれている状況を考慮し、人の求める性能により近い出力を行うAIのニーズは高まっています。今年の展示においても、被災者の順次帰宅を考慮した避難所計画の最適化や、指定した語句を必ず使うニューラル機械翻訳などの研究を紹介しています。

人を深く理解する

CS研では、人の感覚知覚運動メカニズムを探る上で、脳の潜在機能が引き起こすさまざまな錯覚現象を手掛かりに研究を進めてきました。視覚や聴覚に関する錯覚を体験できるWebサイト：イリュージョンフォーラム[8]も公開しています。今年の展示においても、「指先で感じる“こっち”」「ノビのある速球は錯覚？」「壁が動くと速く歩く？」などは、引っ張られる感覚や、視覚および運動との相互作用が引き起こす興味深い錯覚現象から脳の多様な潜在機能を探る研究です。

また、2021年10月には、静岡県立総合病院と人工内耳使用者の音声・言語認知の共同研究を開始しました[9]。難聴児であっても、早期に人工内耳を装用することで健聴児と同程度の音声言語獲得ができることが実証されていますが、このような脳における音声知覚や言語発達のメカニズムは未だ明らかになっていません。医学と脳科学の両面からのアプローチにより、高齢者などの難聴者における聴覚機構の態様を明らかにし、音声知覚・言語発達の個人差の背後にあるメカニズムの解明とこれに基づく支援などの研究に着手しています。

言語獲得に関する研究では、CS研では1999年に日本語約8万語の単語のなじみ度合いを7段階で評価した単語親密度データベースを公開してきましたが、2021年には新しく出現した単語を加えた16万語以上について再調査した「令和版単語親密度データベース」をNTT印刷から提供を開始しました[10]。単語親密度データベースを用いることにより、数十の単語を知っているか否かを答えることで、おおよその語彙数を推定する技術も構築しており、「令和版語彙数推定」としてインターネットから利用できるWebサイト[11]も公開しています。

人・社会・環境との調和と共生をもたらす技術の創出に向けて

CS研の研究は各分野の専門性を究めることで継続・発展してきました。これまで、「見る」や「聞く」などの単一のモダリティに特化したメディア処理研究は、前述した音と映像を用いた話者特定などのように、複数のメディアにまたがったクロスモーダル処理へと発展してきています。人間科学の分野においても多感覚統合のメカニズム解明などに研究がシフトしつつあります。さらに昨年では、CS研における脳科学・人間科学・メディア処理の知見と技術を総動員した研究成果である「投手シミュレータマシン」が、東京オリンピック2020におけるソフトボール日本代表の金メダル獲得に貢献し、新聞やテレビなどで「秘密兵器」として取り上げられました[12]。

昨今の技術の進歩と普及スピードは凄まじく、分野内での研究を研ぎ澄ますだけではなく、周辺分野の研究成果を組み合わせて新たな価値を創造する研究や、異分野と融合させた新たな学際分野を切り拓く気運は高まっています。一方で、新型コロナウイルス感染症のパンデミックや、地球規模の気候変動や自然災害の増大、国際紛争の勃発などの社会情勢の変化により、日々の生活様式や価値観が激変している時代です。ますます複雑化・多様化する社会課題の解決には、これ

までとは抜本的に異なる長期的視野に立った研究アプローチが必要です。このような問題意識のもと、現代数学の基礎理論を研究する組織「基礎数学研究センタ」を2021年10月にCS研内に設立しました[13]。今後、さまざまな分野で生まれる課題について、最新の現代数学手法を駆使しながらアプローチして解決していくとともに、多様な価値観を持つ人と、社会、環境との調和と共生をもたらし、未来へと持続させていくコミュニケーション科学研究に取り組んでまいりたいと思います。

納谷　太：1992年慶応義塾大学理工学部電気工学科卒、1994年同大学院理工学研究科計算機科学専攻修士課程修了。同年、日本電信電話株式会社入社。2003年より2009年までATRメディア情報科学研究所、同知識科学研究所出向。コミュニケーションロボット、センサネットワークによる実世界センシング、人流の時空間予測・最適誘導技術、テーラーメイド学習支援技術等の研究に従事。博士（工学）。2022年4月より現職。

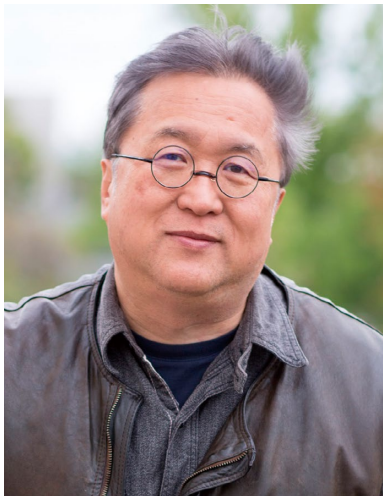
●参考文献

[1] NTT コミュニケーション科学基礎研究所 創立30周年記念ウェブサイト　https://www.kecl.ntt.co.jp/30th/
[2] 世界初、一般的な写真群から未知の奥行きとボケ効果の学習を実現～限られたデータから三次元世界の理解が可能なAIの実現に向けて前進～
https://group.ntt/jp/newsrelease/2021/06/25/210625a.html
[3] 聞きたい人の声に耳を傾けるAI ――深層学習に基づく音声の選択的聴取技術SpeakerBeam, Marc Delcroix, 落合翼, 佐藤宏, 大石康智, 木下慶介, 中谷智広, 荒木章子, NTT技術ジャーナル 2021.7
[4] 機械が会話のパートナーになる日 ――大規模深層学習で拓く雑談対話システムの新境地, 杉山弘晃, 水上雅博, 有本庸浩, 成松宏美, 千葉祐弥, 中嶋秀治, NTT技術ジャーナル 2021.7
[5]「なりきりAI京町セイカ」を用いた実証実験を精華町役場で開始　https://group.ntt/jp/topics/2021/11/12/ai_seika.html
[6] 日本語最大規模のTransformer対話モデルを無償公開　https://group.ntt/jp/topics/2021/09/30/transformer.html
[7] 因果関係に基づく公平・高精度な機械学習予測を実現～どんな予測が差別的かを指定しながら、人を対象とした効果的な予測が可能に～
https://group.ntt/jp/newsrelease/2021/06/01/210601b.html
[8] 錯視と錯聴を体験 イリュージョンフォーラム　https://illusion-forum.ilab.ntt.co.jp/
[9] 難聴者・人工内耳装用者の音声・言語認知の共同研究に関する協定の締結について
～医療と脳科学の両面で難聴のメカニズム解明に貢献することを期待～
https://group.ntt/jp/newsrelease/2021/10/19/211019a.html
[10] 令和版単語親密度データベースの構築と語彙数推定テストの作成～語彙数推定から学習支援へ～
https://group.ntt/jp/newsrelease/2020/06/03/200603a.html
[11] 令和版語彙数推定テスト　https://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/vocabulary_test/php/login.php
[12] 女子ソフトボール × スポーツ脳科学　https://www.rd.ntt/research/JN202112_16471.html
[13] NTT基礎数学研究センタを設立～未知なる量子コンピューティングの原理解明などに挑戦する研究を加速～
https://group.ntt/jp/newsrelease/2021/10/01/211001a.html

MEMO

コミュニケーション行動から探る動物の心

Animal mind and communicative behavior



帝京大学
先端総合研究機構 教授
岡ノ谷　一夫
Kazuo Okanoya

●プロフィール

1983年 慶応義塾大学文学部卒業
1989年 メリーランド大学心理学研究科修了、Ph.D.
1994年 千葉大学文学部行動科学科助教授
2004年 理化学研究所脳科学総合研究センターチームリーダー
2010年 東京大学大学院総合文化研究科教授
2022年 帝京大学先端総合研究機構複雑系認知部門教授（現職）

●受賞歴

2001年 米国心理学会最優秀比較心理学論文賞
2009年 中山人間科学振興財団・中山賞大賞
　同年 日本科学ジャーナリスト会議・科学ジャーナリスト賞
2011年 日本動物行動学会・日高敏隆賞
2012年 ニューロクリアティブ研究会・創造性褒章
2013年 国際進化言語学会夏の学校・最優秀講師賞
など

●専門の内容

鳥のさえずり、ネズミの超音波発声、人間の音楽から、
言語と意識の起源と進化を知ろうとする研究を進めている

●著書

『言葉はなぜ生まれたのか』（文藝春秋,2010）、『つながりの進化生物学』（朝日出版社,2013）、
『脳に心が読めるか』（青土社,2017）、『ことばと心』（玉川大学出版部,2019）（編著）、
『本棚から読む平成史』（河出書房新社,2019）（共著） など

心という言葉を、人はそれぞれ微妙に異なる意味で使っている。知覚や認知、記憶など計測できる部分を指す場合もあれば、それに伴う情報処理過程や神経科学的相関を指す場合もある。しかし、もっとも重要なのは、心とは個人の主観的な体験であるということだ。自分の心は体験できるが、他者の心は体験できない。だから、他者の心の科学とは、自己と他者との類似性にもとづく推測にすぎず、「擬人主義」に立脚している。これまでの心理学は、人間以外の動物の心について考えることは擬人主義として厳しく戒めてきたが、他者の心については比較的寛容に擬人主義を受けいれてきた。現在、人工知能の心的体験についても、擬人主義が許容されているようである。いっそのこと、動物の心的体験についても擬人主義を許容して研究してみるという方法も必要なのではないだろうか。

このような考えのもと、私の研究室ではあえて擬人的解釈を許容して動物の行動を観察し、その行動から推測できる心的体験と神経相関を探る研究を積み重ねてきた。コミュニケーション行動では、ある信号を送信することで受信者の行動が変化し、それによって送信者がなんらかの利益を得る。この際、受信者は送信者の心的状態を推測することによって、より的確な応答が可能になる。以下、動物の場合でも他個体のことを他者と表現する。括弧内に擬人的な解釈を入れる。

コミュニケーションにはいろいろなレベルがある。他者の存在により行動が促進される現象を「社会的促進」という。他者の情動状態が伝染し行動が変化することを「情動伝染」という。他者のために自己の利益を一時的に減らすことを「利他行動」という。他者を操作し自己の利益を増大させることを「他者操作」という。信号を交代で出しあうことを「話者交代」という。

社会的促進の例として、小鳥の採餌行動をあげる。小鳥は新奇物を警戒する。餌入れに新奇物を入れておくと、なかなか食べにこない。しかしある1羽が（勇気をもって）餌を食べ始めると、他の鳥たちも（安心して）接近してくる。最初の1羽はリスクをとるわけだが、続く何羽かがいることで、自分自身が捕食されるリスクを軽減しているともいえる。情動伝染の例として、ラットの超音波発声をあげる。ラットは快を感じると 50kHz の、不快を感じると 22kHz の鳴き声を出す。これらを録音して放鳴すると、50kHz には（良いことを期待して）接近し、22kHz では（危険を警戒して）不動化する。利他行動の例としてラットの救援行動をあげる。ラットは他個体が水に溺れている様子をみると、（その個体が陸地にあがれるように）手を差し伸べる。この際、何が手掛かりとなっているかは明瞭ではないが、メスどうしでこのような行動がよくみられる。ラットが共同保育することと関連するかも知れない。他者操作の例として、小鳥のさえずり行動をあげる。小鳥のオスは、多様な音を配列したさえずりを発する。さえずりは、その個体の適応度をあらわす正直な信号である。メスはオスのさえずりを評価して交尾するかどうかを決める。オスはさえずりによってメスを操作しているといえる。話者交代の例として、ハダカデバネズミの鳴き返し行動をあげる。この動物は地下のトンネルに住んでいる。2匹がトンネル内ですれちがうとき、お互いに鳴きあう。大きな個体のほうが低い声でなき、上を通ってすれちがう。声の低さが序列に対応しているのである。

このような行動の観察から、動物の心的体験について推測し、行動と相関する神経活動を計測することで、心的体験の進化と神経機構を知ることができると期待している。講演ではこれらの行動の動画をみてもらいながら解説してゆく。

⊕『パネルディスカッション』

帝京大学 先端総合研究機構 複雑系認知部門 教授
岡ノ谷 一夫 Kazuo Okanoya

協創情報研究部 インタラクション対話研究グループリーダー・上席特別研究員
小林 哲生 Tessei Kobayashi

2004年東京大学大学院総合文化研究科博士課程修了。博士（学術）。2005年日本電信電話株式会社入社。
専門は発達心理学。乳幼児における言語習得や認知発達の研究に従事。
著書に『0－3さいはじめてのことば』（小学館,2008）、監修絵本に『あかちゃんごおしゃべりえほん』（主婦の友社,2017）など。

柏野多様脳特別研究室長・NTTフェロー
柏野 牧夫 Makio Kashino

1989年東京大学大学院修士課程修了。博士（心理学）。同年日本電信電話株式会社入社。コミュニケーション科学基礎研究所人間情報研究部長などを経て、2018年NTTフェロー。現在、東京大学大学院教育学研究科客員教授、日本学術会議連携会員。
専門は心理物理学・認知神経科学。平成28年度文部科学大臣表彰科学技術賞（研究部門）受賞。
著書に『音のイリュージョン～知覚を生み出す脳の戦略～』（岩波書店,2010）、『空耳の科学-だまされる耳，聞き分ける脳』（ヤマハミュージックメディア,2012）など。

人間情報研究部 感覚運動研究グループ・主任研究員
廣谷 定男 Sadao Hiroya

2001年東京工業大学大学院総合理工学研究科修士課程修了。同年日本電信電話株式会社入社。2006年東京工業大学大学院総合理工学研究科博士後期課程修了。博士（工学）。2007年から2008年ボストン大学客員研究員。
専門は音声コミュニケーション。音声の知覚生成メカニズムの解明および音声信号処理研究に従事。
著書に『聞くと話すの脳科学』（コロナ社,2017）、『こどもの音声』（コロナ社,2019）。

あなたの声を「すぐそば」品質で聴くAI

～遠くからでも近接マイク品質で混ざった音を聞き分ける革新的音響処理技術～

AI hears your voice as if it were "right next to you" – Audio processing framework for separating distant sounds with close microphone quality –

メディア情報研究部 信号処理研究グループ

中谷　智広

Tomohiro Nakatani

●プロフィール

ＮＴＴコミュニケーション科学基礎研究所　メディア情報研究部　上席特別研究員。１９９１年　京都大学　情報学研究科　応用システム科学専攻　修士課程修了。博士（情報学）。１９９１年にＮＴＴに入社以来、音源分離、残響抑圧、雑音抑圧、音声特徴量抽出、ロバスト音声認識など、統計的信号処理を用いた音声処理の研究に従事。電子情報通信学会、日本音響学会の各会員。ＩＥＥＥフェロー。

近年、スマートフォンによる音声認識やヘッドセットを用いたリモート会議など、話者の口元のマイク（近接マイク）を用いた音声アプリケーションが広く利用されています。一方、今後、AIが、より深く私たちの生活に溶け込み、身近（＝「すぐそば」）な存在になるためには、日常生活の中で、必ずしもマイクの「すぐそば」で話されていない音声をも扱うことが求められるようになります。しかし、話者から離れたマイクでは、壁や天井からの反射である残響、複数の話者の音声、背景雑音などが混ざってしまうため、音声アプリケーションの性能は大きく劣化します。これを解決するために、マイクから離れていても近接マイク品質で複数の音声を聞き分ける音声強調技術の研究を進めています。本講演では、単一マイクよりも高品質な処理が可能な複数マイクを用いる音声強調の最新技術を紹介します。

近接マイク品質の実現のための課題

収録音から近接マイク品質の音声を抽出するには、残響抑圧、音源分離、雑音抑圧の3つの処理が必要です。まず、残響抑圧により、遠くで響いている印象のぼやけた音声を、すぐ近くにいる印象のはっきりした音声に変えます。さらに、複数の音声や背景雑音が混在している場合は、音源分離や雑音抑圧により個々の音に分解します。

複数マイク音声強調では、この残響抑圧、音源分離、雑音抑圧の各課題に対し、音が音源からマイクに伝播し混合する各過程（収録音の生成過程）を推定し、その逆変換を適用することで、近接マイク品質の音声を求めます（図1（a））。具体的には、残響が壁や天井に反射してマイクに到達する過程、複数の音声が各方向から到来し混合する過程、雑音が全方向から到来し重畳する過程、をそれぞれ推定し、逆変換を行います。

しかし、従来は、この３つの処理を同時に行うことは困難でした。このため、各処理を順番に適用するしかなく、全体として最適な処理を行えませんでした。例えば、最初に、混ざったままの音に対し残響抑圧を高精度に行うのは、困難な課題でした。全体最適な処理の実現は、音響処理における重要な未解決課題でした。

残響抑圧、雑音抑圧、音源分離の統一モデル

これに対し、私たちは、3つの処理を全体最適な形で実現で

きる統一モデルを考案(世界初)し、さらに、実時間処理が可能な高速アルゴリズムを構築しました［1,2］。統一モデルでは、まず、近接マイク品質の音声や雑音が満たすべき一般的な性質(音声はスパースな信号、背景雑音は定常信号など)を数理的にモデル化します。そして、3つの処理を組み合わせた結果の音がこの性質を最もよく満たすようにするという「統一基準」に基づき各処理を最適化することで、全体最適な処理を実現します(図1（b))。その結果、例えば、離れたマイクで収録した音声の認識性能を大幅に改善できるようになりました(図2（a)-(c))。

さらに、この統一モデルを応用して、比較的少数のマイクでも高精度な推定が可能なスイッチ機構を考案しました［3,4］。従来の複数マイク音声強調では、収録音に含まれる音源数に対し十分な数のマイクがないと推定精度は大きく低下します。一方、収録音を各短時間区間でみると、多くの場合、同時に音を出している音源数は少なくなります。この性質を利用して、統一モデルに基づくスイッチ機構では、収録音を、比較的少数の同じ音源のみからなる時間区間に分類し、時間区間ごとに個別に音声強調を適用します。そして、上記の統一基準に従い、時間区間の分類を含めた全体の最適化を行うことで、少数マイクでも高精度な推定が可能になりました。

統一モデルは、これまで試行錯誤的に組み合わされてきた各処理に対し、理論的にも実用的にも優れた統合指針を与えます。今後、音響処理のデファクト技術として広く利用されていくことが期待されます。

今後の発展：深層学習との最適な統合

音声強調のもう一つの重要なアプローチに深層学習があります。深層学習では、SpeakerBeam［5］が実現した声の特徴に基づく選択的聴取や単一マイクによる音声強調など、複数マイク音声強調では困難な処理が実現できる一方で、残響があると処理音質が悪くなる、音声認識性能の改善は限定的であるなどの課題があります。これらの特徴は、複数マイク音声強調と相補的であり、例えば、深層学習で注目した音声を、複数マイク音声強調で近接マイク品質にすることが可能です(図2（d))。今後、両者の最適な統合方法を構築することで、より高機能で高品質な音声処理が実現されると考えています。

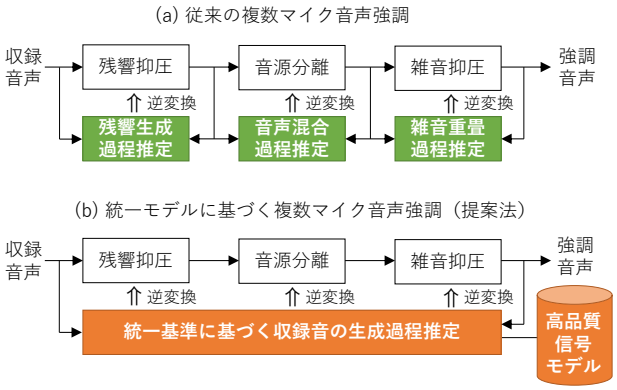


図 1：複数マイク音声強調の従来法と提案法

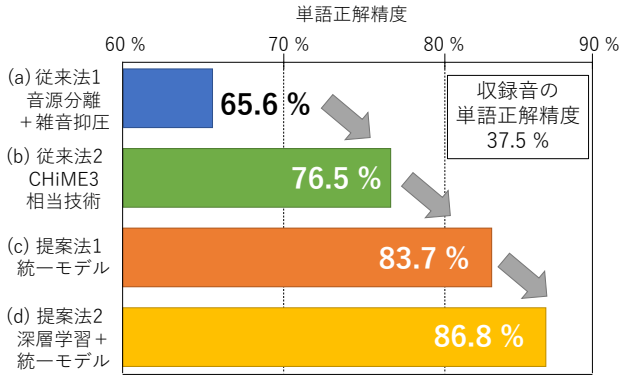


図 2：複数マイク音声強調による音声認識率の改善

●参考文献

[1] T. Nakatani, C. Boeddeker, K. Kinoshita, R. Ikeshita, M. Delcroix, R. Haeb-Umbach, “Jointly optimal denoising, dereverberation, and source separation,” IEEE/ACM Trans. Audio, Speech, and Language Processing, vol. 28, pp. 2267-2282, 2020.

[2] R. Ikeshita and T. Nakatani, “Independent vector extraction for fast joint blind source separation and dereverberation,” IEEE Signal Processing Letters, vol. 28, pp. 972-976, 2021.

[3] R. Ikeshita, N. Kamo, T. Nakatani, “Blind signal dereverberation based on mixture of weighted prediction error models,” IEEE Signal Processing Letters, vol. 28, pp. 399-403, 2021.

[4] T. Nakatani, R. Ikeshita, K. Kinoshita, H. Sawada, N. Kamo, S. Araki, “Switching independent vector analysis and its extension to blind and spatially guided convolutional beamforming algorithms,” IEEE/ACM Trans. Audio, Speech, and Language Processing, 2022.

[5] M. Delcroix, K. Zmolikova, 木下慶介, 荒木章子, 小川厚徳, 中谷智広, “SpeakerBeam: 聞きたい人の声に耳を傾けるコンピュータ――深層学習に基づく音声の選択的聴取,” NTT技術ジャーナル, vol. 30, no. 9, pp. 12-15, 2018.

あなたの心臓をデジタル化

～モバイルセンシングを活用したパーソナル心臓モデリングとその応用～

Making Digital Twin of Your Heart – Personal Heart Modeling with Mobile Sensing and Its Potential Application –

メディア情報研究部 / バイオメディカル情報科学研究センタ

柏野　邦夫

Kunio Kashino

●プロフィール

NTTコミュニケーション科学基礎研究所・物性科学基礎研究所 上席特別研究員。1995年東京大学大学院 工学系研究科 博士課程修了。博士（工学）。同年NTTに入社以来、メディア情報の認識・探索・情報変換、及びバイオメディカル情報科学の研究に従事。文部科学大臣表彰科学技術賞、高柳健次郎業績賞、IEEE Transactions on Multimedia論文賞などを受賞。IEEE、ACM、電子情報通信学会、情報処理学会、人工知能学会、日本音響学会の各会員。

生体のモバイルセンシング

近年、生体情報を計測できる腕時計型の機器が普及しています。腕時計型の機器は小型軽量であることから、日常生活の中で負担なく計測できるという特長があります。私たちの研究チームでは、その長所をなるべく生かしながら、より詳しい生体情報を得ることを目的として、生体情報の新しい計測方法と、得られた情報を活用する情報処理技術の研究を進めています[1]。

その一つがウェアラブル心電計です。医療機関で一般的に行われる心電図検査では、手足と胸部に設置した合計10個の電極から12種類の電位差を得て、その波形を診断に役立てます。また腕時計型の端末では両手の電位差から簡易的な心電図を得ることができます。これらに対し私たちは、心臓の活動を立体的にとらえる目的で、心臓が胸郭と最も近接する心尖部領域を基準点とし、ほぼ直交する3方向に対極を備えたウェアラブル心電計を提案しています（図1左）。試作品では、電極と配線が伸縮性のベルトと一体化しており、肩ベルトとウエストベルトを締めるだけで簡単に装着できるように工夫されています[2]。

また、電気信号と同時に多チャンネルの音響信号を計測できる、AIテレ聴診器と呼ぶ装置を提案しています。ウェアラブル型（図1中）と手持ち型（図1右）を試作しており、いずれも、計測した情報をリアルタイムで遠隔地に送り、遠隔地において端末画面上で体表面上の聴きたい位置を選びながら音を聴ける機能が備わっています。一般的な無線式の電子聴診器とは異なり、心臓の活動を立体的にとらえられるように、複数の箇所で同時に音響信号をとらえる点がAIテレ聴診器の特徴です。

信号から生体内部を探る

このように、立体的に情報をとらえようとするのにはわけがあります。心電計では、心臓を構成する多数の心筋細胞の活動電位の集合体が電位差として観測されます。心音は、心臓の力学的作用、特に心臓内部に4つある弁の開閉がその主な要因で、もし心臓にある種の異常があれば、血流の乱れなどからも音が生じて心雑音として観測されることが知られています。電気も音も、観測できるのは数多くの要因が混ざり合って生じた結果だけであり、その物理的機序を逆にたどってそ

の原因、つまり心臓の状態を推定するのは容易なことではありません。そこで私たちは、場所の手がかりを得るために立体的にとらえた電気や音の情報を活用し、さらに機械学習への物理的制約の導入[3]をはじめとする様々な情報処理の工夫を行うことで、この問題にアプローチしています。

心臓の異常を検知しようとする場合、一般的な心電図では波形の異常がわずかなものにとどまる場合も少なくありません。これに対し活動電位の情報を可視化すれば(テンソル心電図)[4]、非定型的な異常が明瞭に表現されやすく、心不全や虚血性心疾患などの発見に役立つと期待されます。また、心音を立体的にとらえることで、心臓のどこからどのような音が出ているかを推定できるようになると考えられます。図2は、大動脈弁狭窄症に罹患された方の体表面で観測した4チャンネルの音響信号から、生成モデルに基づく心音の振動子分解技術［5］により、8つの体内音源波形を推定したものです。将来的には、これらの音源の位置や音の内容から、大動脈弁に由来する音だけを自動的に取り出して聴くことも可能になるかも知れません。このようにして、一般的な心電図や電子聴診器を用いた場合よりも詳細に心臓の活動を把握できると考えられます。私たちは、このような現在の観測情報に基づく推定結果と、その人の過去の受診履歴や検査結果などとを総合して、個人ごとの心臓のモデル化とシミュレーション(パーソナル心臓モデリング)を実現することをめざしています。

活用に向けて

NTTでは、バイオデジタルツインと呼ぶシミュレータにより心身の状態の未来を予測し、リスクのコントロールやウェルビーイングの向上を支援しようという「医療健康ビジョン」を掲げています[6]。パーソナル心臓モデリングもその試みの一つです。私たちが健康な生活を送るためには、医療機関の中だけではなく、日常生活の中でのヘルスケアを充実させることが重要であると指摘されています。このような社会的要請に応えるためにも、心不全の予防やリハビリテーションへの活用などを端緒として、医療機関等とも連携しながら段階的にパーソナル心臓モデリングの検証と実用化を進めていきます。

図1：ウェアラブル心電計（左）とAIテレ聴診器（中＝ウェアラブル型、右＝手持ち型）（研究用試作品）

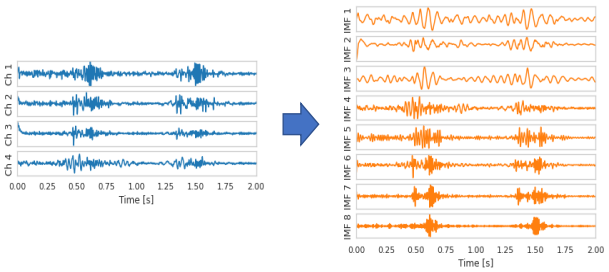


図2：体表面でとらえた音（左）から推定した体内音源波形（右）の例

●参考文献

[1] 中野允裕, 渋江遼平, 柏野邦夫, 塚田信吾, 小笠原隆行,“生体音と心電信号の新たな計測と解析の技術 ――パーソナル心臓モデリングによる心疾患の早期発見・リハビリ応用に向けて,” NTT技術ジャーナル, Vol. 33, No. 5, pp. 27-31, 2021.
[2] Y. Tsukada, M. Tokita, H. Murata, Y. Hirasawa, K. Yodogawa, Y. Iwasaki, K. Asai, W. Shimizu, N. Kasai, H. Nakashima, and S. Tsukada,“Validation of wearable textile electrodes for ECG monitoring, ” Heart Vessels, Vol. 34, No. 7, pp.1203–1211, 2019.
[3] M. Nakano, R. Shibue, K. Kashino, S. Tsukada, and H. Tomoike,“Gaussian process with physical laws for 3D cardiac modeling, ” in *Proc. EUSIPCO 2020*, pp. 1452-1456, 2020.
[4] S. Tsukada, “Wearable textile electrodes for long-term vector ECG monitoring “Tensor Cardiography”, ” in *Proc. ISMICT 2020*, 2020.
[5] R. Shibue, M. Nakano, T. Iwata, K. Kashino, H. Tomoike, “Unsupervised heart sound decomposition and state estimation with generative oscillation models,” *in Proc. EMBC 2021*, pp. 5481-5487, 2021.
[6] NTT報道発表,“医療健康ビジョン: バイオデジタルツインの実現,”2020年11月17日.

デジタルツインでモビリティ群を賢く制御する

～分散深層学習がもたらす未来の可能性～

Smart traffic coordination via digital twins – A future concept of distributed deep learning –

協創情報研究部 知能創発環境研究グループ

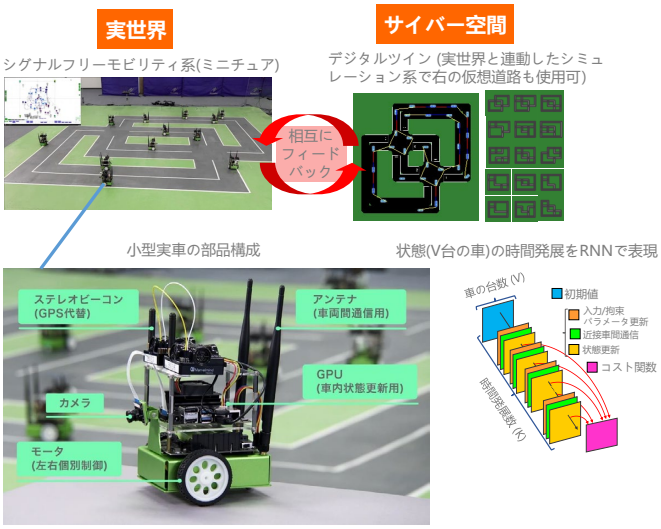
丹羽　健太

Kenta Niwa

●プロフィール

NTTコミュニケーション科学基礎研究所 協創情報研究部 知能創発環境研究グループ 特別研究員、コンピュータ＆データサイエンス研究所 兼務。2008年名古屋大学大学院修士課程修了。同年、NTTに入社。2014年 名古屋大学大学院博士課程修了。博士（情報科学）。入社より音響信号処理（主にマイクロホンアレイ信号処理）の研究に従事。2017-2018年 Victoria University of Wellington客員研究員。分散最適化（機械学習の一分野）の研究に興味を持ち、ネットワークを介した集合知形成の研究に従事。IEEEシニアメンバー。

スマートフォンやスマートスピーカを通じた音声による情報操作等の機械学習の恩恵を一般の人が利用することが当たり前になってきました。次世代の機械学習に関する大きなイノベーションの可能性として、「個」ではなく、通信で連結された「群」を協調して制御/推論することで大規模な系（例えば街全体）を最適化したり、分散蓄積されたデータから集合知を形成することが期待されています。本講演では、私たちNTT・CS研が行ってきた学習可能なデジタルツインを介してモビリティ群を賢く制御する研究について報告します。



(1) シグナルフリーモビリティ プロジェクト

将来、ICTの高度化により、ヒト・クルマ・インフラが高度に協調し、安全・効率的な移動を提供する高度協調型モビリティ社会の実現が期待されています。IOWN構想により究極に高度化されたICTがもたらすモビリティ社会のコンセプトとして、2019年12月に無信号の交通調整（シグナルフリーモビリティ）についてコンセプト動画（Mobility by IOWN）が公開されています[1, 2]。シグナルフリーモビリティの実現に向けた第一歩として、私たちは、車が衝突せずに、移動・輸送時間を限界まで短縮するための分散制御/推論問題に取り組んでいます[3]。

シグナルフリーモビリティを構成する技術について、図1を用いて概説します。まず、実際の車の代わりに小型実車群を用いて、ミニチュア実世界にシグナルフリーモビリティを体現することを目標としています。図1（左下）に示すように、小型実

車には通信モジュール、車輪の駆動用モーター、GPS代替となる位置センサー、演算用小型GPU等が搭載されていて、近隣の車群と通信しながら、状態（速度、位置、モーター駆動等）を逐次更新することが可能です。実世界にある車とサイバー空間上にあるデジタルツインは連動し、シグナルフリーモビリティ（ぶつからずに輸送時間を極小化）を達成するための最適状態をデジタルツイン上で予測し、現実世界にフィードバックします。

デジタルツインを構成する演算の一つとして、状態の逐次更新処理があります。例えば、車の状態の更新は、近接した車群だけではなく、道路との相互作用によって決まります。例えば、交差点の通過優先順位、車速規制の順守等が含まれます。なお、サイバー空間上にある車や道路は、必ずしも実世界の模擬に限定されるわけではありません。実世界には存在しない仮想道路（図1右上）を構築し、多様な道路環境で、車の台数や初期位置などを変えて、あたかも走行したかのようなデータ系列を取得することが可能です。この大量かつ多様なデータを使って学習することにより、効率的に汎化性の高い運転制御モデルを得ることができるでしょう。学習されたモデルを通じて、実世界へフィードバックすることで、大規模かつ高精度な交通制御を可能にします。

サイバー空間上で、デジタルツインの状態変数を時間発展させることで、交通シミュレーションを実施できます。このシミュレーションは、ニューラルネットを用いて表現することが可能です。図1（右下）に示すように、具体的には、(i)各車の状態変数更新、(ii)近接車間の通信による衝突防止トークンの交換、(iii)入力画像や衝突防止パラメータの更新を交互に繰り返す、という処理で構成されています[3]。図1（右下）の各方形の幅が車の台数、奥行きが時間（更新回数）に対応する巨大なリカレントニューラルネットワーク（RNN）ですが、演算の分散並列性が高く、リアルタイムに状態変数を更新することができます。車群が衝突しないような制約を状態変数に課しながら、平均速度を向上する運転制御モデルを学習しています。なお、詳細は研究展示[4]にて紹介しますが、学習可能なデジタルツインを介した集合知形成により、初期状態（ランダム値でモデル化）と比較して約30%の平均速度の向上を実現することを確認しています。

(2)非同期分散型の連合学習 プロジェクト

シグナルフリーモビリティでは、シミュレーションで得られたデータを1つのサーバー上に集約して運転制御モデルの学習を行っていました。しかし、今後、ヒトやモノのデジタルツインを介して全体の系を協調していく世界では、データを一か所に集約するのではなく、デジタルツインごとに分散蓄積されたデータを用いて補助情報（トークン）を交換するだけで、集合知が形成される未来になるでしょう。

ノード数やエッジ数（デジタルツインの数やそれらの通信接続数）が大きく、巨大なネットワークグラフで演算が行われることを想定して、非同期分散型の連合学習の研究を進めています。従来研究では、接続しているノード間でモデル変数を交換して平均化しながら合意形成をとる連合学習則[5,6]が多用されていますが、私たちが提案した方法[7,8]は、(i)非同期分散通信を許容でき、(ii)計算サーバー間のモデルの合意形成に関する制約を課すなど、計算サーバー間の統計的なデータの偏りに耐性のあるアルゴリズムを構築したことが特徴です。

今後、基盤となる数理をさらに発展させるとともに、応用事例（交通網、エネルギー網、物流網など）を増やし、群全体の最適制御によるデジタルツインコンピューティング社会の早期実現に貢献します。

●参考文献

[1] IONWコンセプトビデオ, “Mobility by IOWN”, https://www.youtube.com/watch?v=4fo_kEYrY6E
[2] Digital twin computing white paper (ver. 2.0.0), https://www.rd.ntt/dtc/DTC_Whitepaper_jp_2_0_0.pdf, 2019.
[3] K. Niwa, N. Ueda, H. Sawada, A. Fujino, S. Takeda, B. Kleijn, and G. Zhang, “CoordiNet: Constrained dynamics learning for state coordination over graph,” in *Proc. the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2022)*, (under review).
[4] 丹羽 健太, “信号機を使わないリアルタイム分散車流制御　～デジタルツインを介した集合知学習技術”, NTTコミュニケーション科学基礎研究所 オープンハウス2022　（展示番号01） https://www.kecl.ntt.co.jp/openhouse/2022/exhibition_01.html
[5] J. Chen, A. H. Sayed, “Diffusion adaptation strategies for distributed optimization and learning over networks,” *IEEE Transactions on Signal Processing*, Vol. 60, No. 8, pp. 4289–4305, 2012.
[6] B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y Arcas, “Communication–efficient learning of deep networks from decentralized data,” in *Proc. Artificial Intelligence and Statistics (AISTATS 2017)*, pp. 1273–1282, 2017.
[7] K. Niwa, N. Harada, G. Zhang, and B. Kleijn, “Edge-consensus learning: deep learning on P2P networks with nonhomogeneous data,” in *Proc. the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2020)*, pp. 668–678, 2020.
[8] K. Niwa, G. Zhang, B. Kleijn, N. Harada, H. Sawada, and A. Fujino, “Asynchronous decentralized optimization with implicit stochastic variance reduction,” in *Proc. of the 38th International Conference on Machine Learning (ICML 2021)*, 2021.

数学はどこから来たのか 数学はどう進むのか 数学はどこへ行くのか

～対称性を駆使した解析学と幾何学的手法による整数論の未解決問題と量子計算への挑戦～

Where does mathematics come from? How does it proceed? Where it goes?
– Challenges to unsolved problems in number theory and quantum computation by symmetry-based analysis and geometry –

基礎数学研究センタ
若山　正人
Masato Wakayama

●プロフィール

1978年東京理科大学理学部卒業。1985年広島大学大学院理学研究科博士課程修了。鳥取大学助教授、九州大学数理学研究院教授、同院長、九州大学主幹教授、同マス・フォア・インダストリ研究所長、理事・副学長を経て、2020年東京理科大学副学長・理学部教授、国立研究開発法人科学技術振興機構研究開発戦略センター上席フェロー。2021年10月NTT基礎数学研究センタ数学研究プリンシパル、同上席を兼務。九州大学名誉教授。専門は表現論・数論。

数学研究の目的は未だ見ぬ数学世界に光をあてることです。研究は既存の未解決問題の解決と新しい問題・テーマの発掘により進みます。予見し難いものの、進歩する数学がいつか役立つことは歴史が示しています。たとえば、素数の研究が現代のセキュリティを支え、非ユークリッド幾何学が相対論を通して全地球測位システムGPSを生みました。

数学研究の特質のひとつは、いろいろと散らばっているものをひとまとめにしていく点にあります。数学は多様な科学から刺激を受けつつ究極の抽象化も行ないます。数学内部に限らず、実験科学等からの題材においても、まったく離れ離れにみえていた事実を統一的・普遍的に理解しようとすることは数学のもつ本質的なはたらきです。長い間、互いに無縁なことであると信じられてきたものの間に、疑いようのない親近性や等価性が明らかにされていくことには価値があります。数学の研究は、種々の実験的事実が物理法則へと私たちを導いてくれることに似ていると述べたPoincaré の言葉もあります。そうしたことから数学は、世界の理解の要となる問題の『発見』や体系化による『視座の提供』にさえ貢献しうるものです。

今春、宮﨑、Reyes-Bustos、堀永の3名が基礎数学研究センタに集いました。専門は順に、数論幾何学（Grothendieckのモチーフ理論に不足する困難部分の整備と開拓）、グラフ理論・表現論（暗号分野等で待たれるRamanujan グラフの無限族の新しい構成と量子相互模型の数論）、保型表現論（Deligne 予想に迫るL-関数の特殊値の幾何構造）です。

さて、日本のフィールズ賞はいずれも代数幾何学から生まれました。数論幾何（図形による整数論研究）はWeil予想解決のために代数幾何を再構築した Grothendieck が創始しました。Weil 予想の解決はDeligneです。また、それに連なる20世紀後半からの数学の巨大研究課題としてLanglands哲学（予想群）があります。その原点は表現論の解析的側面にあります（Weil への手紙[1]）。背景にあったのはHarish Chandra のリー群の無限次元表現論とSelberg の跡公式の研究です。なお、表現論とは対称性の学問であり、保型表現論とは、楕円曲線に関する志村・谷山予想なども含む、表現論による保型形式の研究です。Fermatの最終定理の証明もLanglands哲学の一角をなす志村・谷山予想が要でした。プリンストン高級研究所など世界有数の大学等ではこの予想群の研究が活発です。また、その幾何学的変形版は物理学の究極理論といわれる弦理論の定式化にも使われています。さらにABC予想攻略に向け構築された望月のIUT理論と量子計算の課題には類似点があるとの指摘もあります。ただし、量子計算の研究では、未だ現代の代数幾何は使われていないようです。

講演者の関心は、対称性の観点からの数学の開拓です。今回は、量子コンピュータの構築の土台ともなり、理論・実験双方で活発な研究が進む量子光学の模型と数論を繋ぐ話をします[2]。具体的には、非可換調和振動子（NCHO [3]）なる数学モデルと光と物質の相互作用を記述する量子ラビ模型（QRM）及びその非対称模型（AQRM）のエネルギースペクトルとそのゼータ関数を通じた数論がテーマです。実際、ゼータ関数の研究は物理でいえば系の分配関数（状態和）の研究と同等です。なお、それぞれのHeun微分方程式描像により、NCHOがQRMの被覆を定めることがわかります[4]。また、(A) QRMの理論は、超伝導人工原子と電磁場（光）の相互作用を測る実験結果[5]を見事に予言し、量子もつれの実現を強く示唆しました。量子もつれにより物質と光は情報の交換を行います。これにより量子コンピュータの基本素子ができます。

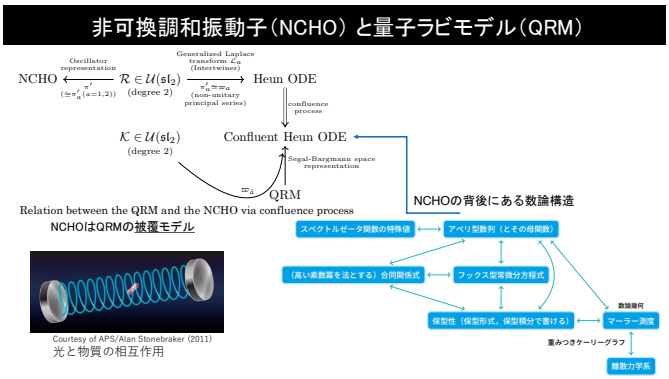
定理は発見によるものです。発明ではありません。したがって、紡ぎ方は数学者個人によりますが、そこにある事実（定理）は発見以前から存在していると思うのが妥当です。ところで近年、地球外高等生命の存在が示唆されています。正しければ、地球では未発見の定理が既に知られている可能性もあります。一方で、量子情報や量子計算などの研究からは、一見抽象的にみえる計算も物理法則から逃れられないように映ります。つまり“計算⊂物理法則⊂数学”なる包含関係があり、数学は宇宙より前にあったのかもしれません。地球上ではバビロニア数学には遡れます。その後も数学は脈々と進歩しています。講演では、現実の数学研究がどのように進むのかを、実際の動機や問題の立て方、アプローチなどを、メンバーの数学にも触れながら以下を例にお話しします：

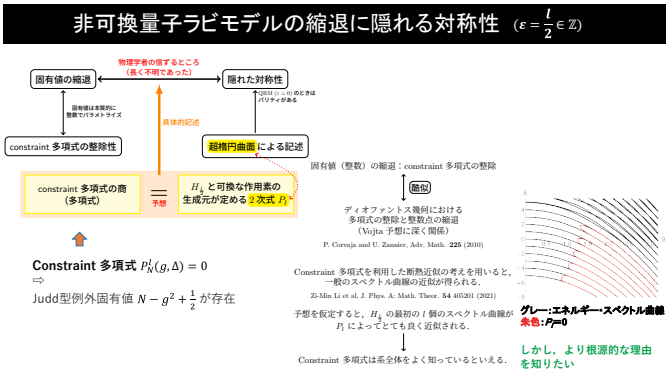
①NCHOのスペクトルゼータ関数の特殊値が生む保型形式等の数論と楕円曲線[6,7]。

②QRMの伝搬関数の解析的公式と無限対称群、経路積分、ゼータ関数[8,9]。

③AQRMの縮退[10,11]と隠れた対称性[12]。ディオファンタス幾何に繋がる課題や予想[13]。

将来、銀河系の知的生命との通信が実現されたとき、自信をもって交流ができるように地球の数学研究は向かうのでしょう。





●参考文献

[1] Letter to André Weil, https://publications.ias.edu/rpl/paper/43

[2] 若山正人, “光とゼータ関数の特殊値,” 日本数学会誌「数学通信」第25巻 第4号pp. 24-52, 2021.

[3] A. Parmeggiani, M. Wakayama, “Oscillator representations and systems of ordinary differential equations,” *PNAS*, 98, pp. 26-30, 2001.

[4] M. Wakayama, “Equivalence between the eigenvalue problem of non-commutative harmonic oscillators and existence of holomorphic solutions of Heun differential equations, eigenstates degeneration and Rabi model,” *Int. Math. Res. Not.*, 2016:145, pp. 759-794.

[5] 布施智子, 吉原文樹, 角柳孝輔, 仙場浩一, “超伝導人工原子と電磁場の相互作用～強結合のその先へ～,”「最近の研究から」日本物理学会誌 73, pp. 21-26, 2018.

[6] T. Ichinose, M. Wakayama, “Zeta functions for the spectrum of the non-commutative harmonic oscillators,” *Comm. Math. Phys.*, 258, pp. 697-739, 2005.

[7] K. Kimoto, M. Wakayama, “Apéry-like numbers for non-commutative harmonic oscillators and automorphic integrals,” *Ann. Inst. Henri Poincaré - D*, (印刷中), 2020, - arXiv:1905.01775.

[8] C. Reyes-Bustos, M. Wakayama, “Heat kernel for the quantum Rabi model,” - arXiv:1906.09597, 2020.

[9] C. Reyes-Bustos, M. Wakayama, “Heat kernel for the quantum Rabi model: II. Propagators and spectral determinants,” *J. Phys. A: Math. Theor. 54*, pp.115202-115231, 2021.

[10] M. Wakayama, “Symmetry of asymmetric quantum Rabi models,” *J. Phys. A: Math. Theor.*, 50, pp. 174001-174020, 2017.

[11] K. Kimoto, C. Reyes-Bustos, M. Wakayama, “Determinant expressions of constraint polynomials and the spectrum of the asymmetric quantum Rabi model,” *Int. Math. Res. Not.*, 2021:12, pp. 9458–9544.

[12] C. Reyes-Bustos, D. Braak, M. Wakayama, “Remarks on the hidden symmetry of the asymmetric quantum Rabi model,” *J. Phys. A: Math. Theor.* 54, pp. 285202-285215, 2021.

[13] C. Reyes-Bustos, M. Wakayama, “Degeneracy and hidden symmetry -- an asymmetric quantum Rabi model with an integer bias,”- arXiv:2106.08916, 2021.

研究展示のカテゴリ

NTTコミュニケーション科学基礎研究所は、人と人、あるいは人とコンピュータの間の「こころまで伝わる」コミュニケーションの実現をめざして、人間と情報の本質に迫る基礎理論の構築と社会に変革をもたらす革新技術の創出に取り組んでいます。オープンハウス2022では、「データと学習の科学」、「コミュニケーションと計算の科学」、「メディアの科学」、「人間の科学」に関する合計29展示をご紹介します。

データと学習の科学

機械学習技術に基づくデータ解析を基礎から活用事例まで包括的に紹介します。

- 巨大な交通網を分散並列計算でリアルタイムに最適制御したい
- オールフォトニクス・ネットワーク時代の機械学習を実現したい
- 柔軟・高速・正確な学習と推論を機械にさせたい

コミュニケーションと計算の科学

人と人、人とコンピュータのやりとりを円滑にする研究や、数学の不思議を解き明かし計算の限界に挑戦する研究を紹介します。

- 資源を最適に配置して多くの人に快適な社会デザインを設計したい
- コンピュータ・人とのやりとりを、楽しくかつ円滑に行いたい
- 数の本質を解明することで、より安全な暗号開発や数の真理探究に貢献したい

メディアの科学

音声・画像・テキストなどのメディア情報を個別に、または組み合わせて高度に処理する研究を紹介します。

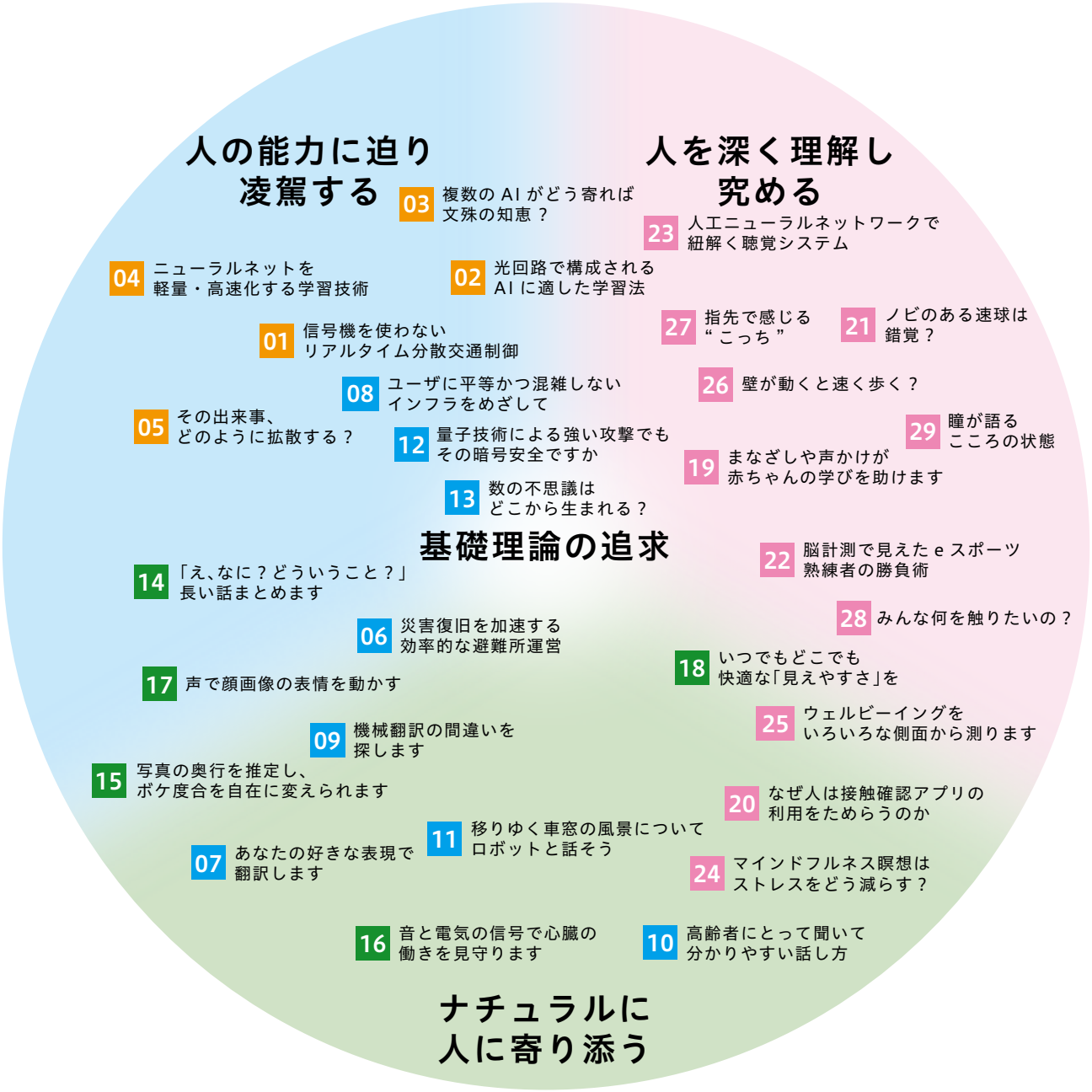
- メディアの内容や構造を理解して高度な画像加工や音響処理を実現したい
- 各種メディアの本質をとらえ、相互に変換したり、高精度な要約をしたい
- メディア情報処理をもとに医療への応用をしたい

人間の科学

人の情報処理メカニズムを科学的に解明し、豊かなライフスタイルを提案・実現する研究を紹介します。

- 人の知覚・運動特性を理解し、心理の解明や新たなインタフェース開発に活かしたい
- 触れ合いやコミュニケーションの役割を解明し、豊かで安全な社会の実現に役立てたい
- アスリート達がもつ秀でた能力の秘密を解明したい

トピックマップ


人の能力に迫り凌駕する
人を深く理解し究める
基礎理論の追求
ナチュラルに人に寄り添う
03 複数の AI がどう寄れば文殊の知恵？
23 人工ニューラルネットワークで紐解く聴覚システム
04 ニューラルネットを軽量・高速化する学習技術
02 光回路で構成されるAI に適した学習法
27 指先で感じる“こっち”
21 ノビのある速球は錯覚？
01 信号機を使わないリアルタイム分散交通制御
08 ユーザに平等かつ混雑しないインフラをめざして
26 壁が動くと速く歩く？
05 その出来事、どのように拡散する？
12 量子技術による強い攻撃でもその暗号安全ですか
29 瞳が語るこころの状態
19 まなざしや声かけが赤ちゃんの学びを助けます
13 数の不思議はどこから生まれる？
22 脳計測で見えたｅスポーツ熟練者の勝負術
14 「え、なに？どういうこと？」長い話まとめます
28 みんな何を触りたいの？
06 災害復旧を加速する効率的な避難所運営
18 いつでもどこでも快適な「見えやすさ」を
17 声で顔画像の表情を動かす
25 ウェルビーイングをいろいろな側面から測ります
09 機械翻訳の間違いを探します
15 写真の奥行を推定し、ボケ度合を自在に変えられます
20 なぜ人は接触確認アプリの利用をためらうのか
11 移りゆく車窓の風景についてロボットと話そう
07 あなたの好きな表現で翻訳します
24 マインドフルネス瞑想はストレスをどう減らす？
16 音と電気の信号で心臓の働きを見守ります
10 高齢者にとって聞いて分かりやすい話し方

Learning of collective intelligence via digital twins

01 デジタルツインを介した集合知学習技術
信号機を使わないリアルタイム分散交通制御

どんな研究	信号機のない街を自動運転車が走行する時代を見据え、通信を介して各車が自律的に速度等の状態を制御すること(シグナルフリーモビリティ)がIOWNで提唱されています。その実現のための(a)基盤数理、(b)デジタルツイン(実世界と連動したシミュレーション系)を用いた制御モデルの学習、(c)実世界交通制御へのフィードバックについて紹介します。
どこが凄い	車が衝突しない制約下で平均速度を最大化する学習法を考案しました。分散並列計算によりリアルタイム処理を実現できる交通流を常微分方程式によって表現しました。その離散化により、(i)近接車間の通信(衝突防止のためのトークン交換)と(ii)各車内の状態更新を交互に繰り返すニューラルネットワークを構築しました。
めざす未来	デジタルツインを介して、IoT機器群が分散協調することにより、単一のIoT機器では成し得なかった高度な全体効率性や集合知モデルを手に入れることができるでしょう。リアルタイム性の高い分散計算を通じて巨大な系(例えば、交通網、エネルギー網、物流網)を最適制御することにより、社会に貢献することをめざします。

目的

自動運転車が無信号の街を自律的に走行するシグナルフリーモビリティのコンセプトが示されています[1]。その実現に向けて、車が衝突しない中で、移動・輸送時間を限界まで短縮するための分散学習問題に取り組んでいます[2]。

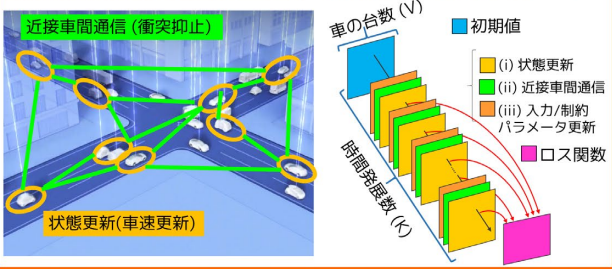
提案法: 制約付きダイナミクス学習

- 車が衝突しないように状態変数(速度等)に制約を課しながら各車が自律的に状態変数を更新する交通流を常微分方程式で表現

$$\frac{d\boldsymbol{x}}{dt} = \underbrace{M_1(\boldsymbol{x}, t, \theta, \boldsymbol{A}, \boldsymbol{b})}_{\text{状態更新}} + \underbrace{M_2(\boldsymbol{x}, t, \boldsymbol{A}, \boldsymbol{b})}_{\text{近接車間通信}}$$

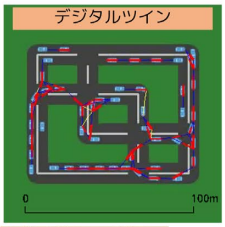
- 上記の常微分方程式を離散化することで、V台の車の状態変数がK回時間発展するリカレントニューラルネットワーク(RNN)を構成。その結果、下図のように(i)各車の状態更新と(ii)衝突防止制約を満たすための近接車間通信(トークン交換)、(iii)入力/制約パラメータ更新を交互に繰り返す処理で構成
- この巨大なRNN(幅V×奥行K)は、分散並列性の高い演算セットで構成されるため、状態変数の時間発展をリアルタイムに実行可能
- 平均速度を向上させるロス関数を小さくするように運転ダイナミクスモデルを最適化

時刻tにおける処理のイメージ　時間発展した処理のイメージ

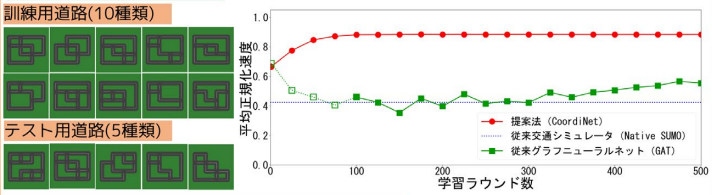


デジタルツインを使ったモデルの学習

- 制約付きダイナミクスを効率的に学習するために実世界と連動した交通のデジタルツイン(車や道路で構成)を構築
- 実世界にある道路だけではなく、仮想世界の道路で車を走行させることで、モデルを汎化させるために十分なデータを得ることが可能(下図)

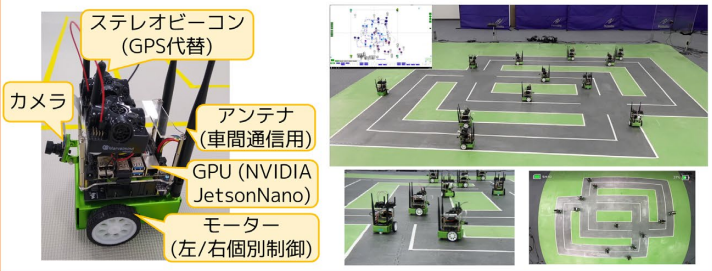


- 提案法(制約付きダイナミクス学習)は、初期(ランダム)と比較して30%程度の平均速度の向上を確認（赤線)。制約のないグラフニューラルネットワーク(緑線, GAT[3])や交通シミュレータ(青線, SUMO[4])を使った場合と比較しても平均速度が高く、優位性を確認。



実世界へのフィードバック

- 12台の小型実車を使ってシグナルフリーモビリティの実世界系を構築し、学習後の運転ダイナミクスを実世界にフィードバックする実験を実施。各車が互いに譲り合いながら衝突することなく自律的に走行することを確認。



関連文献

[1] IONWコンセプトビデオ, “Mobility by IOWN,” YouTube, 2019
[2] K. Niwa, N. Ueda, H. Sawada, A. Fujino, S. Takeda, B. Kleijn, G. Zhang, “CoordiNet: Constrained dynamics learning for state coordination over graph,” in *Proc. the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2022)*, 2022 (under review).
[3] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.
[4] Simulation of Urban MObility (SUMO), https://www.eclipse.org/sumo/

連絡先

丹羽 健太 （Kenta Niwa）協創情報研究部 知能創発環境研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Accelerated method of fine-layered optical neural networks

02 細層構造を持つ光ニューラルネットワークの高速学習

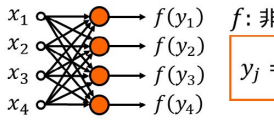
光回路で構成されるＡＩに適した学習法

どんな研究	光回路を用いたニューラルネットワーク（ＯＮＮ）は超低消費電力と高速動作により注目されています。ＯＮＮは干渉計の多層構造で実現されるため、その数理モデルは干渉計表現行列の多段の積になります。このため、既存の自動微分による学習は多くの時間を要するという問題がありました。この展示では、ＯＮＮの高速学習法を紹介します。
どこが凄い	干渉計表現行列の多段積を細層構造ネットワークとしてモジュール化し、誤差逆伝播法で各細層の干渉計パラメータを更新するための専用偏微分式を用いる学習法を構築しました。回帰型ニューラルネットワークにおける文字認識実験で、本学習法は既存自動微分法に比べ、50倍以上の高速学習を達成しました。
めざす未来	本学習法は、実光回路上のオンチップ学習や新しいデバイスへの応用という発展に加え、従来の膨大な電力を消費するＡＩではなく、超低消費電力で動作する光ニューラルネットワークによるグリーンＩＴの思想を実践するＡＩの実現に寄与します。

光ニューラルネットワーク (ONN) とは？

光干渉計アナログ回路で光信号と重み行列の乗算を行うNN

ニューラルネットワーク



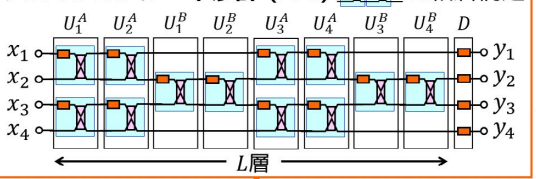
f: 非線型関数

$$y_j = \sum_{i=1}^{4} w_{ji} x_i$$

物理層

入力信号 x（光・振幅：位相）→ 重みと入力信号の**アナログ乗算器** → 出力信号 y（光・振幅：位相）

マッハツェンダー干渉計 (MZI) の細層構造



PS 位相変調器 可変 φ

DC 方向性結合器 固定分岐比

数理モデル

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = D \cdots \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & U^B_{1[1]} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} U^A_{2[1]} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & U^A_{2[2]} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} U^A_{1[1]} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & U^A_{1[2]} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad dim$$

$$U^R_{p[q]} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\varphi^R_{p[q]}} & i \\ ie^{i\varphi^R_{p[q]}} & 1 \end{pmatrix}$$

U^R_p: ユニタリ行列　$\varphi^R_{p[q]}$: パラメータ

$R = A, B$　$1 \le p \le L/2$　$1 \le q \le dim/2$

学習における問題点

物理的要請：小規模光回路でデータ処理の必要有
→ 回帰型ニューラルネットワーク(RNN)の適用

細層構造： 各MZI層を線型回路と見なすことによる深層化
→ **超深層ニューラルネットワークの学習**

既存自動微分法 (AD) では
計算グラフの大規模化により**多大な計算時間が必要**

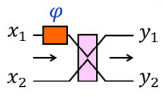
高速学習法

ポイント1 専用偏微分式 (Customized derivatives: CD)

パラメータ更新式：$\varphi \leftarrow \varphi - \eta \boxed{(\partial L/\partial \varphi)}$　L: 損失関数　η: 学習率

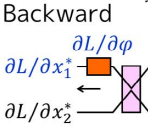
1回の乗算で $(\partial L/\partial \varphi)$ を表す**CD**により**高速化**

Forward



$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\varphi} & i \\ ie^{i\varphi} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Backward　↕ 共役転置



$$\begin{pmatrix} \partial L/\partial x_1^* \\ \partial L/\partial x_2^* \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\varphi} & -ie^{-i\varphi} \\ -i & 1 \end{pmatrix} \begin{pmatrix} \partial L/\partial y_1^* \\ \partial L/\partial y_2^* \end{pmatrix}$$

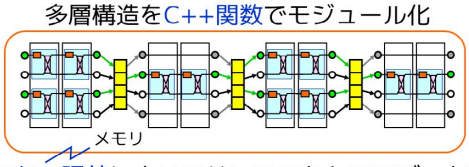
2つの値 $(x_1^*, \partial L/\partial x_1^*)$ の乗算で更新可能

$$\frac{\partial L}{\partial \varphi} = 2 \cdot \mathrm{Im}\left(x_1^* \frac{\partial L}{\partial x_1^*} \right)$$

ポイント2 細層間高速アクセス (Pointer rewiring: PR)

細層間データ転送を**コピー無し**で実行することで**高速化**

多層構造をC++関数でモジュール化

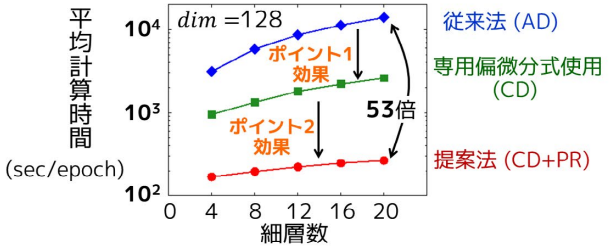


ポインター張替によるアドレスアクセスでデータ取得

効果

RNNの隠れ層を光回路で構成
Pixel-by-pixel MNISTタスク(784時系列)

従来法と提案法の学習時間比較



関連文献

[1] K. Aoyama, H. Sawada, “Accelerated method for learning fine-layered optical neural networks,” in *Proc. of IEEE/ACM the 40th International Conference on Computer-Aided Design*, 2021.

連絡先

青山 一生（Kazuo Aoyama） 協創情報研究部 知能創発環境研究グループ
Email: cs-openhous-ml@hco.ntt.co.jp

Bayesian ensemble learning for better generalization performance

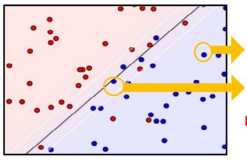
03 汎化性能を向上させるためのベイズ的アンサンブル学習

複数のAIがどう寄れば文殊の知恵？

どんな研究	機械学習を用いた予測の際、予測値に加えてその**確からしさ**の評価が重要になる場合があります。確からしさの評価は、例えば予測が信頼できるのかの判断基準や、意思決定問題などで広く活用されています。本研究では**複数の学習モデル**を使って、ニューラルネットワークなどの複雑なモデルで確からしさを効率良く計算する手法を考案しました。
どこが凄い	これまでも複数のモデルを用いた予測の確からしさの評価は機械学習で広く行われてきましたが、**どのように複数のモデルを用意すればよいのか**は理論的に不明でした。本研究では、特に未知のデータに対する予測の確からしさを複数のモデルを用いて表現するための学習方法を理論的に導出しました。
めざす未来	予測の確からしさは、**機械学習を安全に用いる指針**を与える他、**実験計画法**でのより有用なデータ収集に活用されるなど自然科学分野等で重要な役割を果たします。本研究を発展させることで、複数のモデルによる確からしさの評価をより簡単に実現できるようにし、より機械学習の適用範囲を広げることをめざします。

機械学習における予測の「確からしさ」

機械学習の応用時に、予測の「**値**」だけでなくその予測がどれぐらい正しそうかという「**確からしさ**」が役に立つ場合があります。



確からしさ高 → 予測を信頼し採用
確からしさ低 → 予測を信頼しない
→（例）医療診断など、確かな予測が求められる場合には、代わりに専門家に依頼

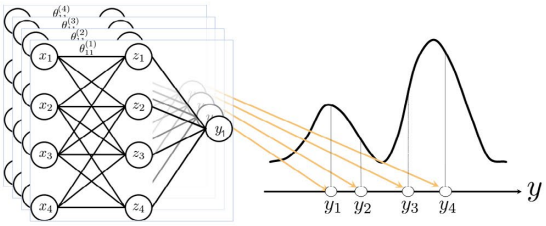
確からしさが使われる応用例
- 学習データを集める際に学習に特に有用なデータを集める、実験計画法
- 次にどのような行動を行うかを決定する意思決定問題

✓確からしさを得る手法として「**ベイズ推論**」により予測値の**分布**を得るということが広く行われます。
✓ニューラルネットのような大きなモデルではベイズ推論の実行には**近似**が必要になります。

既存研究：複数のモデルのばらつき

複数のニューラルネットを用意して、その出力のばらつきを分布とみなして確からしさを表現します。

複数のニューラルネットの予測のばらつき → 予測の分布 → 確からしさとして活用

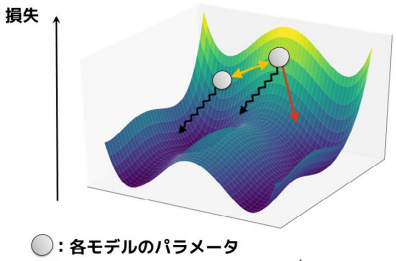


課題
✓同じ学習のさせ方をすると、同じ学習結果になる可能性があります。
✓どうやって複数のニューラルネットを用意すればよいのでしょうか？

提案手法：複数のモデルの用意の仕方

複数のモデルを用意し、**モデルのパラメータが互いに異なるようになる「反発項」**を学習のための目的関数に取り入れます。

勾配降下の際の学習のイメージ



貢献１：反発項は予測の重み付き分散として計算可能な形で理論より導出できます。

貢献２：得られた複数のモデルは、反発項なしの学習より優れた予測の分布になることが理論的に保証されます。

：各モデルのパラメータ

パラメータの更新方向 ＝ 損失の勾配 ＋ 反発項

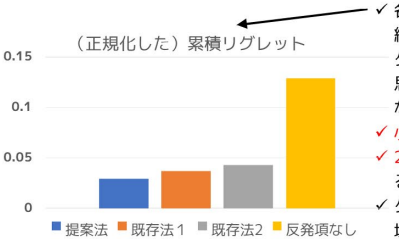
実験：確からしさを活用する意思決定問題

✓ 各時刻で与えられるキノコの特徴量をもとに、食べるかどうかの**意思決定**を繰り返し行います。
✓ 毒の有無によって得られる報酬が異なり、**トータル報酬の最大化が目標**です。報酬の情報のみが学習の手がかりです（分類問題のようにラベルは与えられません）。

特徴量22次元

毒無し→ 報酬 ＋５ポイント
毒有り→ 報酬－３５ポイント（確率５０％）
報酬 ＋５ポイント（確率５０％）

✓ 確からしさの情報から、色々と**試行錯誤して情報を集めつつ**も、より報酬をもらえそうな行動をとっていくトレードオフの制御が必要です。



✓ 各時刻で正しく毒の有無を区別し続けた場合の報酬との**差**で評価。グラフでは、完全にランダムな意思決定をした場合の報酬と理想的な報酬の差を１として正規化。
✓ 小さいほどよい性能を意味します。
✓ 20のニューラルネットワークによる分布の近似。
✓ グラフは意思決定を5万回し続けた場合の結果。

関連文献

[1] F. Futami, T. Iwata, N. Ueda, I. Sato, M. Sugiyama, “Loss function based second-order Jensen inequality and its application to particle variational inference,” in *Proc. Neural Information Processing Systems (NeurIPS)*, 2021.

連絡先

二見 太（Futoshi Futami） 上田特別研究室
Email: cs-openhouse-ml@hco.ntt.co.jp

Pruning neural networks with iterative randomization

04

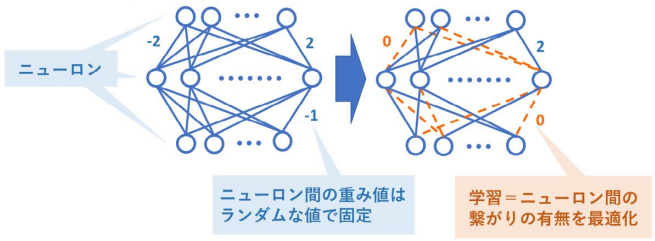
ニューラルネットの枝刈りにおける反復ランダム化技術

ニューラルネットを軽量・高速化する学習技術

どんな研究	深層学習の新たな学習メカニズムとして枝刈り最適化が注目されています。枝刈り最適化で学習されたニューラルネットはスパース＆軽量である一方、学習時に従来より大きなネットワークが必要なことが課題でした。本研究ではこの課題を解決する新たな原理を考案しました。
どこが凄い	枝刈り最適化が従来に比べて大きなネットワークを必要とすることは理論的にも証明されており、実用上の課題となっていました。本研究ではパラメータの反復ランダム化によりこの課題が解決されることを理論的に証明し、また画像・テキスト分類において実験的にも有効性を確認しました。
めざす未来	近年の深層学習の大規模化に伴い、その運用コストやエネルギー消費量は無視できないものとなっています。本研究を発展させることで、ニューラルネットを更に軽量・高速化できる学習技術を実現し、AIの運用コスト削減や省エネルギー化に貢献することをめざします。

背景：ニューラルネットの枝刈り最適化

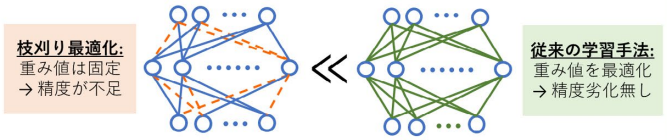
- 近年のニューラルネット研究 [Ramanujan+,2020] により、**ニューロン間の繋がりの有無のみを最適化(枝刈り最適化)**することでニューラルネットを訓練できることが発見されました。



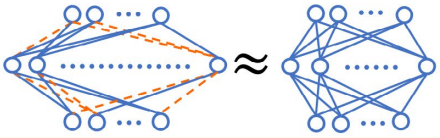
- ニューロン間の重み値を二値化・量子化したまま訓練可能で、訓練後のネットワークはスパース構造を持つため、**訓練と同時に軽量・高速化を実現可能**です。

課題：大きなネットワークサイズが必要

- **枝刈り最適化**では重み値を固定したまま学習が進むため、重み値を連続的に最適化する**従来の学習手法**に比べて、**ネットワーク出力の精度が不足**してしまいます。



- そのため従来と同程度の精度を達成するためには**ネットワークサイズを大きくする必要があり、訓練時のメモリ・計算効率が課題**となります。

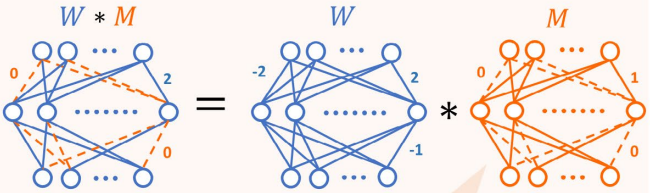


提案手法：重み値の反復ランダム化

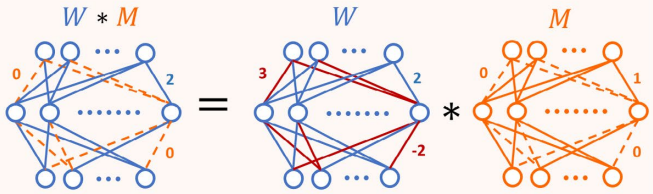
アイデア　枝刈り最適化の過程で**不要となった重み値をランダム化し、より高精度な解を探索します。**

アルゴリズムの概略

① **ニューラルネットW**の枝刈りを表す**マスクネットMを最適化**します。



② 対応する M の重み値が0となっている**Wの重み値を、ランダムな値で置き換えます。**



③ ①に戻り、再び M を最適化します。（繰り返し）

- 一度不要とみなされたニューロン間の繋がりに対してランダムな重み値で復活できるようにすることで、**枝刈り最適化の精度を高めることに成功しました。**
- 本手法を用いることで、ネットワークサイズを大きくする必要性が無くなることを**理論的にも証明しました。**従って本手法により**枝刈り最適化のメモリ・計算効率の課題が解決されました。**
- GPU 環境では**追加の計算コストもほとんど不要です。**

関連文献

[1] D. Chijiwa, S. Yamaguchi, Y. Ida, K. Umakoshi, T. Inoue, “Pruning randomly initialized neural networks with iterative randomization,” *Advances in Neural Information Processing Systems 34*, 2021.
[2] V. Ramanujan, M. Wortsman, A. Kembhavi, A. Farhadi, M. Rastegari, “What's hidden in a randomly weighted neural network?,” in *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11893–11902, 2020.

連絡先

千々和 大輝 （Daiki Chijiwa） コンピュータ＆データサイエンス研究所
Email: cs-openhouse-ml@hco.ntt.co.jp

Learning time-evolving states via dynamic Hawkes processes

05 時変Hawkes過程を用いて状態の変化を推定

その出来事、どのように拡散する？

どんな研究	SNSにおける情報の伝搬、デモや伝染病の広がりなど種々の社会現象はコミュニティ（国や地域、オンラインコミュニティ）間の拡散過程として記述されます。拡散過程は、過去の事象の発生だけでなく、予測時刻における各コミュニティの「状態」（例えば、各地域の人々のデモ参加へのモチベーションやオンラインコミュニティにおけるユーザの関心の変化）によって変化します。拡散過程のモデル化に広く用いられるHawkes過程に深層学習モデルを導入することで、イベント履歴の背後にあるコミュニティの状態の時間変化を自動で学習する手法を構築しました。
どこが凄い	拡散過程のモデル化に広く用いられている既存のHawkes過程では、イベント発生の背後にある各コミュニティの「状態」を捉えることができません。本研究では、Hawkes過程においてイベントの発生確率を表す強度関数をニューラルネットとその微分を使って設計することで、コミュニティの状態の変化を柔軟にモデル化しつつ、強度関数のパラメータを効率的に学習する手法を構築しました。
めざす未来	提案手法によって、近未来のイベントの高精度な予測が可能になります。近未来のイベント予測は予測警備、マーケティング、防疫など様々な応用において重要なタスクです。例えば、暴力的なデモが起きそうな場所を正確に予測できれば事前の警備配置に役立てることができます。

拡散過程の予測

SNSにおける情報の伝搬、デモや伝染病の広がりなど様々な社会現象はコミュニティ間の拡散過程として記述できる

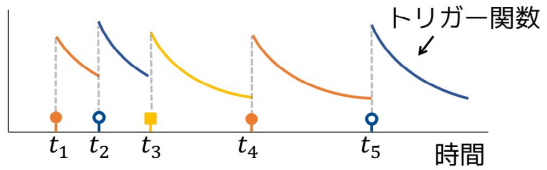
→ イベント間の関係 ● 過去のイベント ✖ イベント予測

全米における抗議デモの発生履歴

拡散過程の背後にあるメカニズムの理解、近未来のイベント予測は予測警備・マーケティング・防疫など様々な応用で重要

既存手法：Hawkes過程

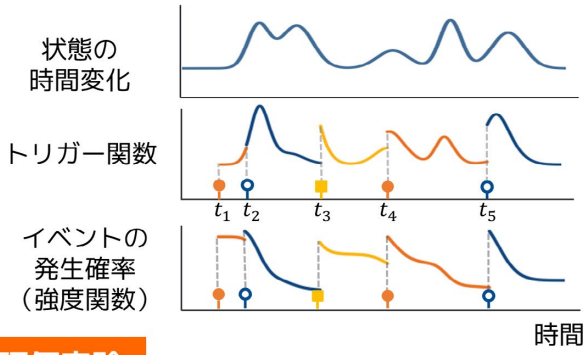
イベントの発生確率を表す強度関数を過去のイベントの発生時刻を入力とする「トリガー関数」を用いて記述することで過去のイベントの影響を考慮



課題：コミュニティの"状態"の変化を考慮できない
例) COVID-19流行による抗議デモの減少
　　ロックダウンによる情報拡散の加速

提案手法：時変Hawkes過程

- 各コミュニティの「状態」の時間変化を表すニューラルネットを導入
- ニューラルネットとその微分を用いてトリガー関数を記述することで、コミュニティの状態変化の柔軟なモデル化と効率的なパラメータ学習が可能に



評価実験

- 4つの実イベントデータを用いて5つの既存手法と提案手法のイベント予測精度を比較
- イベント数の予測値と観測値の間の誤差（MAPE）を評価指標として使用

	Reddit	ニュース	抗議デモ	犯罪
一様ポアソン過程	0.553	0.6	0.345	0.144
Hawkes過程	0.458	0.471	0.415	0.179
Reinforced点過程	0.595	0.481	0.581	0.175
SelfCorrecting過程	0.475	0.452	0.524	0.123
RMTPP	0.311	0.446	0.639	0.302
提案手法	**0.305**	**0.442**	**0.318**	**0.117**

提案手法は全てのデータで既存手法を上回るイベント予測精度を達成

関連文献

[1] M. Okawa, T. Iwata, Y. Tanaka, H. Toda, T. Kurashima, H. Kashima, “Dynamic Hawkes processes for discovering time-evolving communities' states behind diffusion processes,” in *Proc. of the 27th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD21)*, pp. 1276–1286, 2021.

連絡先

大川 真耶（Maya Okawa） 人間情報研究所
Email: cs-openhouse-ml@hco.ntt.co.jp

Optimal shelter operation with sequential return of evacuees

06 被災者の順次帰宅を考慮した避難所計画の最適化

災害復旧を加速する効率的な避難所運営

どんな研究	災害により自宅が倒壊した被災者には避難所が提供されますが、施設を本来の用途に復旧するためには避難所の効率的な運営が必要です。本研究では、被災者の帰宅時刻を考慮して避難所閉鎖時の被災者の移転先を決定することで、避難所を運営するコストと被災者が避難所を移転する負担の合計を最小化します。これにより災害後早期に避難所を本来の用途に使えるようになり、迅速な復旧が可能となります。
どこが凄い	帰宅時刻が同じ被災者を同じグループとして扱うことにより、被災者数が多数の場合でも効率的に計算する手法を考案して、シミュレーション実験で既存手法より約１２倍高速になることを確認しました。また、被災者が避難所間を移転する際の負担の大きさを推定することで、避難所の運営コストと被災者の移転コストの二者間のバランス調整を可能にしました。
めざす未来	災害シミュレーションを用いて対応策を検討する際に、対応策のパターンを網羅的に実行するだけでは、効率的に最適解を見つけることができません。本技術を発展させて最適な対応策を効率的に発見できるシミュレーション基盤を確立し、災害に限らず様々な社会課題のシミュレーションによる解決をめざします。

研究の目的

災害時の避難所として、小学校や中学校が指定されることが多いが、学校教育の再開には避難所を閉鎖する必要がある

避難所に宿泊する被災者の減少に合わせて少数の避難所に集約するように移転してもらえば早期に避難所を閉鎖できる

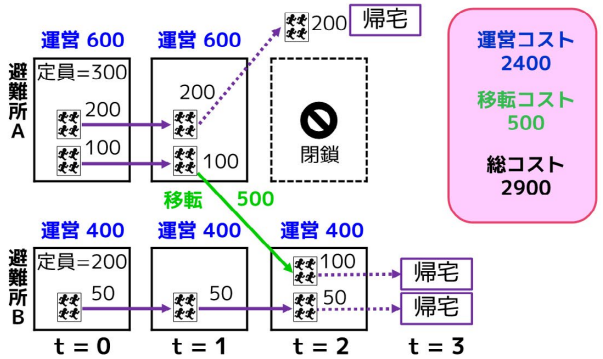
被災者の移転コスト
「近所の避難所に滞在し続けたい」

避難所の運営コスト
「避難所を閉鎖して活動を早期に復旧したい」

コストの合計を最小化する手法を開発

提案手法のポイント①

帰宅時刻が同じ被災者をグループ化して、人数を表現する変数を導入することで、被災者数に依存しない計算量を実現



運営コスト 2400
移転コスト 500
総コスト 2900

提案手法のポイント②

過去の災害データ（阪神・淡路大震災）に以下の仮定をおき、現実の運営日数との誤差が最小となるパラメータを採用

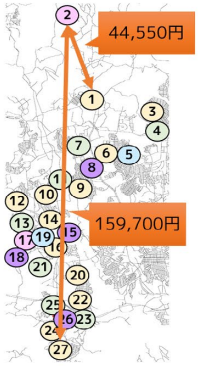
仮定

1. 被災者は区をまたいで移転しない
2. 移転コストが移動距離に比例する
3. 各時刻でコスト最小の運営をした

過去災害データ
- 被災者数推移
- 避難所運営日数

→ 推定

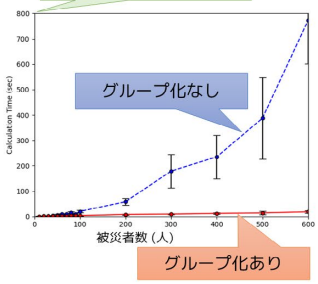
移転コスト：１人１ｋｍ当たり **１万円相当**の負担感

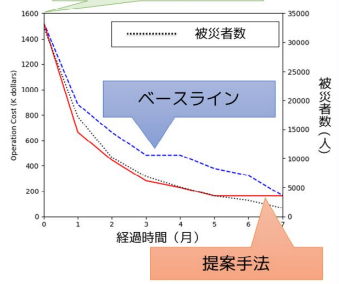


評価実験

奈良県生駒市での震災を想定したシミュレーション実験で性能評価

提案手法は短い計算時間で、運営コストと移転コストを低減





手法	ベースライン	提案手法	
運営コスト/8か月	4.9億円	**3.5億円**	運営コストを **29%削減！**
移転回数/32707人	8259回	3611回	移転の負担も削減
移転コスト	8707万円	5383万円	

関連文献

[1] H. Shimizu, H. Suwa, T. Iwata, A. Fujino, H. Sawada, K. Yasumoto, “Evacuation shelter scheduling problem,” in *Proc. the 55th Hawaii International Conference on System Sciences (HICSS 2022)*, pp. 5705–5714, 2022.
[2] 清水仁, 諏訪博彦, 岩田具治, 藤野昭典, 澤田宏, 安本慶一, “避難所計画問題の定式化と効果推定,” マルチメディア、分散、協調とモバイル(DICOMO2021)シンポジウム, pp. 607–614, 2021.

連絡先

清水 仁（Hitoshi Shimizu） 協創情報研究部 知能創発環境研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Lexically constrained neural machine translation

07 指定した語句をかならず使うニューラル機械翻訳

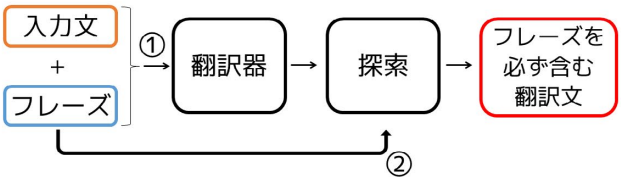
あなたの好きな表現で翻訳します

どんな研究	近年、機械翻訳技術が発展し高精度で流暢な翻訳結果が得られるようになりました。しかし従来の方法ではどの訳語を用いるかを指定しながら高い精度を実現することができないため、翻訳文が意図しない表現となることがありました。本展示では指定した表現を用いた翻訳を行うために機械翻訳の操作性を向上させる技術を紹介します。
どこが凄い	私たちは翻訳文に指定した語句を必ず含む機械翻訳手法を新たに開発し、従来の方法に比べてより高速に高い精度で翻訳できるようになりました。また、この技術を用いて国際ワークショップWAT2021で開催された日英翻訳のコンペティションに参加し、世界一位の成績を収めました。
めざす未来	法務や特許、科学技術などの分野の文書を翻訳する際には、専門用語や固有名詞などの語句が文書内で一貫して同じ表現で翻訳されることが強く求められます。本技術を用いることで、それらの語句に対して訳語を指定できるようになり、表現の一貫性を保つことができるようになります。将来的にはユーザとの対話的な操作による翻訳を行うことで、ユーザにとって望ましい表現を用いた翻訳の実現をめざします。

指定した語句を含む機械翻訳

- ユーザの好きな表現を翻訳で用いたい
- 特許や法務の翻訳では専門用語などの訳語が文書内で一貫して同じであることが求められる
- 一般的な機械翻訳では出力される翻訳文に対して表現などの指示を与えるのは難しい
  → 機械翻訳の操作性を向上させる必要がある
- フレーズを指定するとそのフレーズを**必ず含む**翻訳ができる機械翻訳手法を提案
  - 高速に動作して高い翻訳精度を達成
  - WAT2021 語彙制約付き翻訳コンペで世界1位

提案手法のポイント



① 翻訳器に 翻訳したい文と指定したフレーズを一緒に入力して 翻訳器がフレーズを**できるだけ含む**翻訳をするように学習する
② 翻訳器の出力をもとにフレーズを**必ず含む**翻訳文を探索する

ポイント:
翻訳器がフレーズを出力するように学習する (①) ことにより後段の翻訳文の探索 (②) が**効率化**
- ①だけではフレーズが必ず含まれる保証がない
- ②だけでは翻訳精度が低く計算時間もかかる

提案手法の効果

科学技術論文データによる実験

- 専門用語に対して人間が正解の表現を指定した時の英日/日英翻訳の性能評価を実施

翻訳精度 BLEU (高い方が良い)

手法	英語→日本語	日本語→英語
一般的なNMT†	44.64*	29.30*
①のみ	53.79*	41.88*
②のみ	45.38	23.22
提案手法	**55.49**	**43.33**

†フレーズの情報を利用していない
*翻訳結果が指定したすべてのフレーズを含んでいない

- ①と②を組み合わせることで **フレーズを指定できる**ことに加えて **翻訳精度を大きく改善**
  → 適切に正しいフレーズを指定できれば 人間と同等の精度での翻訳が実現可能
- 指定したフレーズを必ず含む翻訳では 提案手法が②のみよりも**3倍以上高速**に動作することを確認

関連文献

[1] K. Chousa, M. Morishita, “Input augmentation improves constrained beam search for neural machine translation: NTT at WAT 2021,” in *Proc. of the 8th Workshop on Asian Translation (WAT2021)*, pp. 53–61, 2021.

連絡先

帖佐 克己（Katsuki Chousa） 協創情報研究部 言語知能研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Equilibrium optimization of combinatorial congestion games

08 組合せ混雑ゲームの均衡最適化

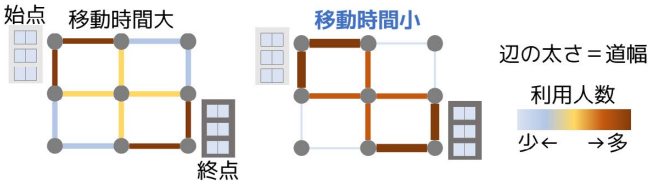
ユーザに平等かつ混雑しないインフラをめざして

どんな研究	道路や通信網などの社会インフラは、あるリンクに利用が集中すると混雑し所要時間といったコストが増えます。本展示では、各ユーザが自分勝手に経路やリンクの組合せを選択しても、結果的に混雑せずにユーザのコストが低くなるような、最適な社会デザインを効率よく求める技術を紹介します。
どこが凄い	私たちは、社会デザインをどう変化させればよりコストが下がるかを微分可能計算という方法を用いて求める新手法を開発しました。また、膨大な数の経路を二分決定グラフとよばれるデータ構造で圧縮し計算に用いることで、経路設定に限らない広い設定の問題を効率的に扱えるようになりました。
めざす未来	本技術により、例えば道路の改良や通信帯域の拡充などのネットワーク設計を通じて、人流や通信網の混雑を減らすことに貢献できる可能性があります。また、我々の混雑状況計算手法は汎用的であり、組合せ最適化を含む機械学習の問題など、より広範な分野への適用が期待されます。

社会デザイン

道路の道幅・速度制限、通信路の通信帯域など、インフラ設計時に調整が利く部分
→ 調整してインフラの混雑を抑えたい

(例) 道幅を調整して移動時間を抑える



組合せ混雑ゲームと均衡状態

人々が自分勝手に道路ネットワーク等のインフラを利用する状況をモデル化
- 無数の人々が各々コストの小さい経路やリンクの組合せを選択
- リンクのコストは通る人が多いほど増加

均衡状態

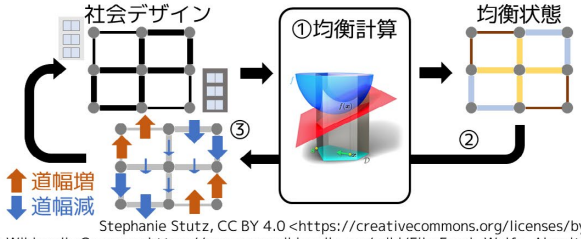
プレイヤーが各々自分のコストを下げようとして至る状態
…どのプレイヤーのコストも等しく最小な状態

均衡状態の最適化

予算制約などの下で、どう社会デザインを決めれば平等性を確保した上でもプレイヤーのコストが小さくなる？
- 経路の数は膨大→ある社会デザインの下で均衡状態を計算するのも難しい
- さらに、良い社会デザインを見つけ出す必要がある

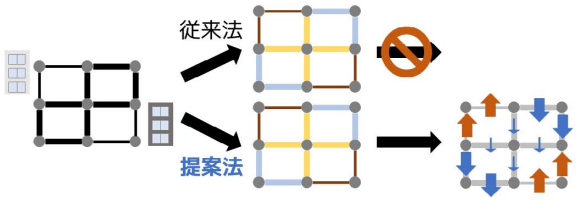
提案法のポイント

① ある社会デザインの下で均衡状態を計算
② 社会デザインの変化に対するコストの増減（**微分**）を均衡計算の過程を使って計算
③ 微分をもとに社会デザインの変化を定めて更新
を繰り返すことで良い社会デザインを求める

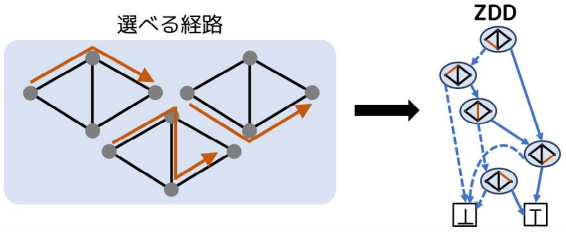


Stephanie Stutz, CC BY 4.0 <https://creativecommons.org/licenses/by/4.0>, via Wikimedia Commons https://commons.wikimedia.org/wiki/File:Frank-Wolfe_Algorithm.png

ポイント1：社会デザインに対する微分の情報を計算できる均衡計算手法を開発し、**広い問題に適用可能**に



ポイント2：ゼロサプレス型二分決定グラフ(ZDD)を用いて選べる経路の共通部分をまとめて表現、圧縮したまま均衡計算に利用し、**高速計算が可能**に
(例) 8,000兆個の経路を**1MB未満**で表現



関連文献

[1] S. Sakaue, K. Nakamura, “Differentiable equilibrium computation with decision diagrams for Stackelberg models of combinatorial congestion games,” in *Proc. 35th Conference on Neural Information Processing Systems (NeurIPS)*, 2021.

連絡先

中村 健吾（Kengo Nakamura）協創情報研究部 言語知能研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Post-editing support based on source-target word alignment

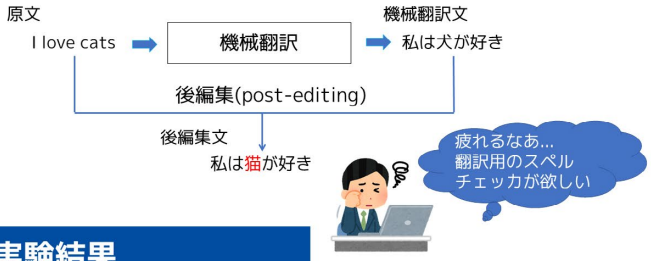
09 原文と訳文の単語対応に基づく機械翻訳の後編集支援

機械翻訳の間違いを探します

どんな研究	ニューラル機械翻訳には、流暢ではあるが必ずしも原文の内容とは一致しない訳文が生成されるという問題があります。この展示では、人間と機械が協調して機械翻訳の誤りを検出し、訂正する「後編集」作業を支援する技術を紹介します。
どこが凄い	正しく翻訳された訳文と原文に対して単語対応を求めるだけでなく、誤って翻訳されて必ずしも意味的に等しくない訳文と原文に対して単語対応を求めることができる技術を実現しました。さらに単語ごとに翻訳品質を推定する既存技術と本技術を組み合わせることにより、訳文の誤りを訂正するためにどこをどう修正すべきかをユーザに提示できます。
めざす未来	従来の文法誤り訂正技術では、文法的に正しいかはチェックできても、意図した内容を表現しているかはチェックできませんでした。この研究は、原文と訳文を見比べながら、スペルチェッカーのように簡単に使える対話的な機械翻訳技術の実現をめざしています。

機械翻訳の後編集支援の必要性

ニューラルネットによって機械翻訳の精度は大きく向上したが、誤りがゼロになることはない。
医療や特許など誤りが許されない分野では、後編集（人間による確認と修正）が不可欠である。



実験結果

国際会議WMT-2020の翻訳品質推定タスクのデータ（英独と英中）
- 単語ごとに翻訳の正(OK)/誤(BAD)が付与された原文、機械翻訳文、後編集文の3つ組 8000件
- うち1000件に対して独自に人手で単語対応を付与
- 残り7000件+800件を訓練データ、200件をテストデータとして編集タグと単語対応の精度(F1)を評価

	英語からドイツ語への翻訳			英語から中国語への翻訳		
	原文編集タグ	機械翻訳文編集タグ	単語対応	原文編集タグ	機械翻訳文編集タグ	単語対応
ベースライン	0.626	0.767	0.828	0.360	0.733	0.739
提案法	0.755	0.827	0.916	0.849	0.897	0.888

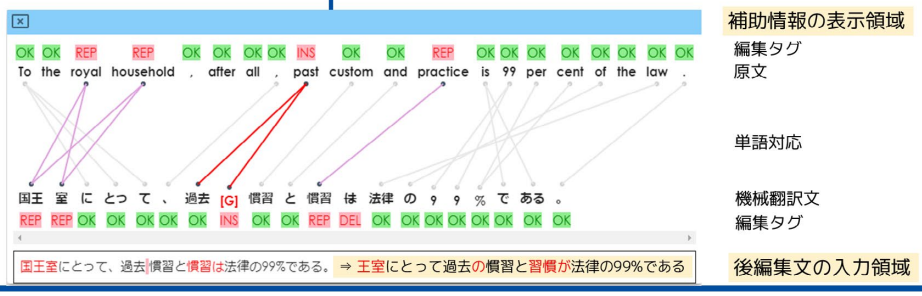
アプローチ

単語ごとに翻訳品質(OK/BAD)を推定する技術と誤りに頑健な単語対応技術を組み合わせて、後編集で行うべき編集操作を予測する

挿入(INS): 原文BAD, 単語対応なし
削除(DEL): 機械翻訳文BAD. 単語対応なし
置換(REP): 原文BAD, 機械翻訳文BAD, 単語対応あり

ユーザインタフェース

後編集の作業者は、支援システムが予測した単語対応と編集操作を参考にして、機械翻訳文を編集



関連文献

[1] M. Nagata, K. Chousa, M. Nishino, “A supervised word alignment method based on cross-language span prediction using multilingual BERT,” in *Proc. the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020.
[2] Y. Wei, T. Utsuro, M. Nagata, “Word-level quality estimation for machine translation based on source-MT Word alignment,” 言語処理学会第27回年次大会予稿集, 2021. (筑波大学との共同研究)

連絡先

永田 昌明（Masaaki Nagata） 協創情報研究部 言語知能研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Selecting voice and words for better understandability

10 分かりやすさにつながる声と言葉の使い方の解明

高齢者にとって聞いて分かりやすい話し方

どんな研究	高齢者にとって聞いて分かりやすい話し方を明らかにするために、ことば選びと声の使い方の両面から研究を進めています。これまでは「大きな声でゆっくりと話す」以上の文献記述がなく、どこでどのように話し方を変えれば良いのかが解明されていませんでした。
どこが凄い	高齢者応対業務従事者や有資格者の中から、聞いて分かりやすいと考えられる高齢者向けの模範話者を高齢者に選んでもらい、その模範話者へのインタビューや音声分析から、高齢者向けの話し方の詳しい特徴を一部解明しました。これは話し方の実践に必要不可欠な新しい知見です。
めざす未来	高齢者向けの模範話者の中で暗黙知となっている分かりやすい話し方を明確な知識にすることで、その話し方を誰もが実践できるようにします。その話し方で集められた音声で学習したAIと高齢者との豊かなコミュニケーションもめざしています。

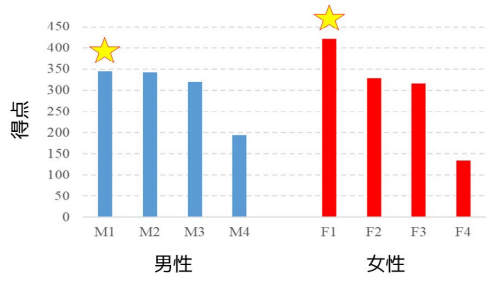
高齢者による音声情報取得上の課題

・認知機能が衰え始めた高齢者への音声系支援技術は限定的：
　補聴器や話速の一律伸長など
・音声系での支援拡大のために、高齢者応対業務従事者の話し方を文献で探しても
　「大きな声で ゆっくりと 話す」
に留まり、未整理または暗黙知の状況

【本研究の目的】
「高齢者にとって聞いて分かりやすい話し方」を定式化し、具体的で実践できる形にすること

アプローチ

1）高齢者が、高齢者応対業務従事者や有資格者の音声を対比較で聞いて、聞いて分かりやすい話者を選び、選ばれた話者に得点を与え、男女の最高得点者を高齢者が聞いて分かりやすいと考える模範話者に選定（下図の☆のM1とF1）
2）模範話者へのインタビューを通して意識されているコツの聞き出し
3）無意識の話し方のコツを探るため文間ポーズ長比較など音声分析を実施



模範話者選びでの各話者の得点

これまでの到達点と課題

高齢者にとって聞いて分かりやすい話し方を徐々に明確化（青字が従来、下線が新規）

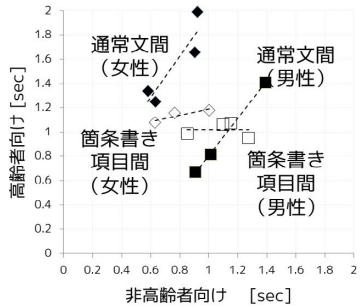
【インタビューから】

言語面
馴染みがあり曖昧性のない言葉で簡潔な構文に言い換え

音声面
大きな声：相手に聞こえる程度の大きさで
声の高さ：ほぼ全般的には普通程度に低く、感情伝達時は高く
ゆっくり：遅め、かつ、一定が基本で、重要箇所では更にゆっくり
区切って：難聴者向けや雑音下ほどはブツブツには区切らない
⇒「どこで？」「どの程度？」の多くは無意識

だから分析

【音声分析から】
ポーズ長：文書構造と表現に応じた差異



高齢者向け音声で箇条書き項目を読む際の項目間のポーズ長（◇□）が縦軸の1秒付近に揃ってくる様子
（破線は回帰直線、◆■は通常文間のポーズ長）

【今後の課題】
・その他の「どこで？」「どの程度？」を具体化
・高齢者以外の聞き手（乳幼児や外国人など）や話し手（アナウンサなど）との共通点・相違点の明確化

関連文献

[1] H. Nakajima, Y. Aono, “Collection and analyses of exemplary speech data to establish easy-to-understand speech synthesis for Japanese elderly adults,” in *Proc. 23rd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, pp. 145–150, 2020 (https://ieeexplore.ieee.org/document/9295000).
[2] 中嶋秀治, 宮崎昇, 阪内澄宇, “高齢者への語りかけ音声におけるポーズ長の分析,” 日本音響学会 秋季研究発表会 講演論文集, pp. 399–400, 2015.

連絡先

中嶋 秀治（Hideharu Nakajima） 協創情報研究部 インタラクション対話研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Casual-dialog system based on scenery and nearby information

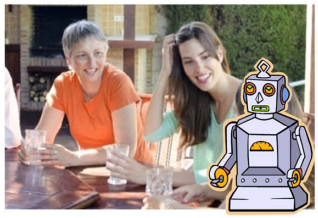
11 移動中に得られる画像・周辺情報を話題とする雑談対話技術

移りゆく車窓の風景についてロボットと話そう

どんな研究	時々刻々とリアルタイムに変化する車からの景色を話題とする、移動体向けパートナー対話システムの研究です。自車位置周辺の情報も取り込むことで、物知りな対話システムと「今」を共有する楽しみを感じながらドライブする、新しい体験を実現しています。
どこが凄い	NTTが構築した、日本語最大規模の「対話データ」を用いた深層学習に基づく対話モデルを利用することで、自然な雑談を実現しています。本システムはさらに、車窓画像の情報と自車位置周辺のスポット情報を取り込むことで、車から見える風景やそれに紐づく情報を話題とした対話を世界で初めて実現しています。
めざす未来	対話システムは機械であるからこそ、いつでも気兼ねなく、どんな話題についても、私たちそれぞれに合ったスタイルで話すことができます。人生に楽しみをもたらす話し相手として、対話システムが当たり前のパートナーになる未来をめざしています。

人の雑談相手となる対話システム

対話システムが単なるタスク解決のツールではなく、
雑談を楽しむ相手として期待され始めている



- **いつでも**話せる
  （時間や場所の制約を受けない）
- **気楽に**話せる
  （遠慮がいらない・相手からの評価を気にしなくてよい）
- **深い話題も**話せる
  （細分化した趣味にも対応できる・プライベートな話も相談できる）

事前学習を利用した対話システム

大規模事前学習に基づく手法で雑談システムの性能が急速に向上

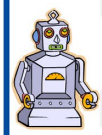
※事前学習：文の自然さや大まかな応答パターンなどを、あらかじめ大規模データを用いてシステムに学習させる方法。

超大規模対話データ
（数十億発話ペア）

大まかな応答関係や文の自然さを事前学習
(=Pre-train)

汎用対話モデル
文理解・文生成・常識的知識を内包

望ましい応答関係を少量データで調整
(=Finetune)



システム

こんにちは。よろしくお願いします。早速ですけど、何か趣味ってありますか？

こんにちは。水曜どうでしょうを見るのが趣味です。

水曜どうでしょう！大好きです。どのシリーズが好きですか？

対決列島が好きです。

対決列島は名作ですね！僕はヨーロッパリベンジが好きです。好きな回とかありますか？

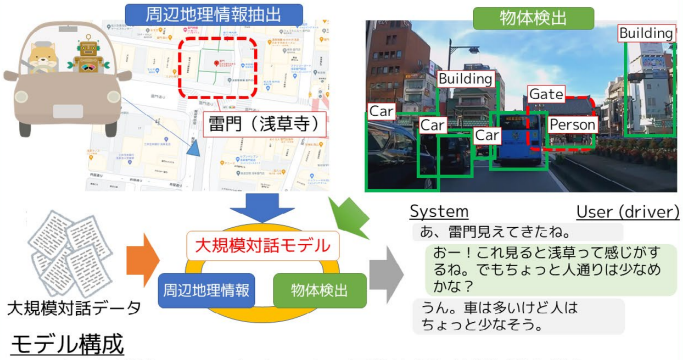
ユーザ

テキストに閉じた対話では非常に自然な雑談を実現

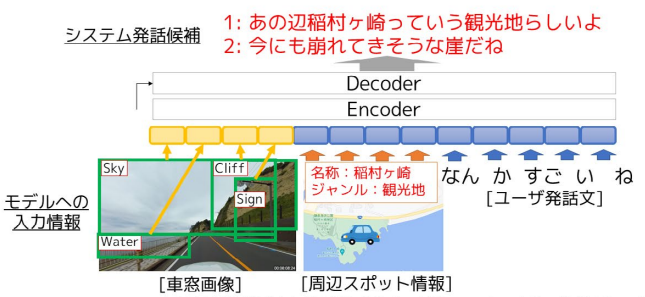
風景を「見ながら」雑談するロボット

テキストに閉じた雑談の課題
- システム自身の周辺の状況を踏まえた対話ができない

本展示
- ドライブ中のクルマから見える、時々刻々と変化する風景や地理情報を認識し話題とする雑談ロボットを実現
  （トヨタ自動車と連携・一人称視点動画が対象の雑談は世界初）



モデル構成
- テキスト雑談システムをベースに、画像情報と外部知識を導入
  - 画像から物体を検出する技術を利用し、物体が検出された位置の画像特徴をモデルへ入力
  - 自車位置周辺のスポット情報をテキストで入力
  - 映像から得られる連続画像に対し、印象を表現する発話を逐次的に生成＆過去数秒間で最も話題性の高い発話を選択



本研究は科研費「人と社会的に共生する対話システムのための行動決定モデル基盤技術の確立」（課題番号19H05693）の支援を受けた

"Artificial Neural Network with Chip" by Ch'enMeng is licensed under CC BY 2.0

関連文献

[1] 杉山弘晃, 有本庸浩, 水上雅博, 千葉祐弥, 中嶋秀治, 成松宏美, “学習データの違いに対するTransformer Encoder-decoder対話モデルの応答変化の分析,” *SIG-SLUD*, Vol. 93, pp. 107–112, 2021.（対話システムライブコンペティション4優秀賞）
[2] H. Sugiyama, M. Mizukami, T. Arimoto, H. Narimatsu, Y. Chiba, H. Nakajima, T. Meguro, “Empirical analysis of training strategies of Transformer-based Japanese chit-chat systems,” *arxiv:2109.05217*, 2021.
[3] 杉山弘晃, 成松宏美, 水上雅博, 有本庸浩, 千葉祐弥, 目黒豊美, 中嶋秀治, “Transformer encoder-decoder モデルベース雑談対話システムの学習方法に対する主観評価の変動分析,” 人工知能学会第35回全国大会論文集, 2021.

連絡先

杉山 弘晃（Hiroaki Sugiyama） 協創情報研究部 インタラクション対話研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Quantum algorithm for finding collisions of hash functions

12 ランダムハッシュ関数の衝突を発見する量子アルゴリズム

量子技術による強い攻撃でもその暗号安全ですか

どんな研究	近年、量子計算機の研究開発の進展を背景に、量子計算機を用いた強力な攻撃を考慮に入れた、暗号の安全性評価が必要になってきています。この安全性評価のためには、具体的な攻撃手法が必要です。本展示では、暗号の中核技術であるハッシュ関数に対する、量子計算機を用いた攻撃手法をご紹介します。
どこが凄い	暗号の中核技術であるハッシュ関数に対して、量子計算機を用いた最高速の攻撃手法を発見しました。本手法の攻撃速度は理論限界を達成しており、理論上可能な最強の攻撃手法と言えます。また、汎用的な手法であるため、様々なハッシュ関数の安全性評価に利用可能です。
めざす未来	ハッシュ関数は、日常生活で使われている電子署名やユーザ認証に加え、最近では、ブロックチェーンなどにも利用される、中核的な暗号技術です。本成果は、現在進んでいる耐量子計算機暗号の議論における、ハッシュ関数の安全性解析に重要な知見を与えることにより、将来においても安全な暗号の開発に貢献します。

背景と成果の概要

- **暗号技術の安全性は、攻撃にかかる時間の膨大さで評価**
  例）「現在最高速の計算機でも解読に数百億年必要」
- 量子計算機の研究開発の進展を背景に、**量子計算機を用いた強力な攻撃を考慮に入れた、安全性**(攻撃に必要な時間)**の評価が必要**に

暗号の中核技術である**ハッシュ関数**に関して、**多重衝突**を発見(＝攻撃)するための、**量子計算機を用いた最高速アルゴリズムを発見**
⇨様々なハッシュ関数の安全性評価に利用可能

ハッシュ関数

入力の推測が難しい短いデータを出力

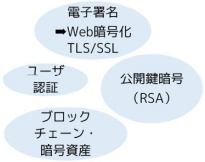
入力 d: Quantum information science is expected to be an important field in information processing in the near future.

ハッシュ関数 f

出力 $f(d)$: rbxXx....zp7kMY

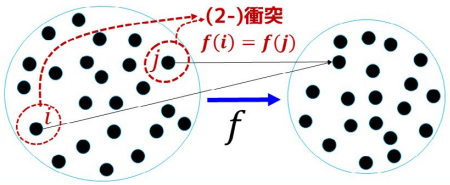
利用

改ざん防止等が必要な、非常に多くの場面で利用



ハッシュ関数の衝突

要素のペア(i,j) が関数fにより同じ値に移されるとき($f(i)=f(j)$)、ペア(i,j) を(2-)衝突と言う。同様に、fにより同じ値に移されるℓ個の要素をℓ-衝突と言う。(例えば、3-衝突の場合、$f(i)=f(j)=f(k)$)



衝突が発見できると、文書の改ざんなどができてしまう（例：同じ電子署名を持つ異なる文書に差し替える）
⇨衝突発見の困難性(=膨大な時間を要するかどうか)を評価する必要
⇨この評価のためには，**衝突を発見するアルゴリズムが必要**

ハッシュ関数の安全性向上の原動力＝衝突発見アルゴリズムの考案
標準的に利用されているハッシュ関数の世代交代は、
安全性を脅かす「衝突発見アルゴリズム」の考案により起こってきた

MD5 ['91] ➡ SHA-1 ['95] ➡ SHA-2 ['02] ➡ SHA-3 ['15]

アルゴリズムの詳細

以下では，ハッシュ関数のℓ-衝突発見問題に対する量子アルゴリズムの「理論的に保証された性能」と「動作の概略」を説明する

一様ランダムに選ばれたハッシュ関数*
$f:\{1,\cdots,M\}\to\{1,\cdots,N\}\ (M\geq N)$
に対する ℓ-衝突発見問題を解く量子アルゴリズムを構築し、
$N^{\frac{1}{2}\left(1-\frac{1}{2^{\ell}-1}\right)}$**時間***で 動作することを証明。 *$f$の計算は十分高速であると仮定

◎理論的に可能な速度上限を達成([LZ20]による下界と一致)

既存アルゴリズム[HSX17]との計算速度比較

各多重度(ℓ)に対して、ハッシュ値の総数(N)と多重衝突を発見するために要する時間の関係を表す

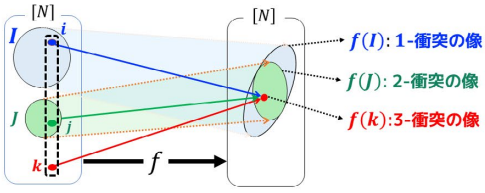
ℓ (多重度)	2	3	4	5	・・・	ℓ
既存結果[HSX17]	$N^{\frac{1}{3}}$	$N^{\frac{4}{9}}$	$N^{\frac{13}{27}}$	$N^{\frac{40}{81}}$	・・・	$N^{\frac{1}{2}\left(1-\frac{1}{3^{\ell}-1}\right)}$
提案アルゴリズム	$N^{\frac{1}{3}}$	$N^{\frac{3}{7}}$	$N^{\frac{7}{15}}$	$N^{\frac{15}{31}}$	・・・	$N^{\frac{1}{2}\left(1-\frac{1}{2^{\ell}-1}\right)}$

例：多重度$\ell=3$、ハッシュ値の総数Nが2000bit値の場合、
既存手法の計算時間に比べ、提案手法の計算時間は１０億倍程度高速
N:2000bit値→ $N^{\frac{4}{9}}:N^{\frac{3}{7}}\approx 1,000,000,000:1$

動作の概略（ 3-衝突発見問題を例に）

1. 部分集合$I\subset[N]$をサンプルし、Iの像$f(I)$を計算($[N]\equiv\{1,\cdots,N\}$)
2. Iの要素と2-衝突を構成する部分集合$J\subset[N]\setminus I$ を探索し、像$f(J)$を計算
3. IおよびJ の要素と3-衝突を構成する要素$k\in[N]\setminus(I\cup J)$を探索
4. 得られた3-衝突の３つ組み (i,j,k) を出力

1.部分集合Iをサンプル
2.$J\subset[N]\setminus I$を量子的に探索
3.$k\in[N]\setminus(I\cup J)$を量子的に探索



関連文献

[1] A. Hosoyamada, Y. Sasaki, S. Tani, K. Xagawa, “Improved quantum multicollision-finding algorithm,” in *Proc. 10th International Conference on Post-Quantum Cryptography (PQCrypto 2019)*, pp. 350–367, vol. 11505, 2019.
[2] A. Hosoyamada, Y. Sasaki, S. Tani, K. Xagawa, “Quantum algorithm for the multicollision problem,” *Theoretical Computer Science*, vol. 842, pp. 100–117, 2020.

連絡先

谷 誠一郎（Seiichiro Tani）メディア情報研究部 情報基礎理論研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Finding new arithmetic phenomena via generalized motives

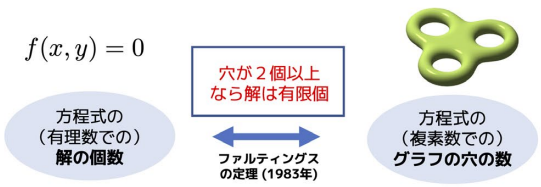
13 一般化モチーフ理論を用いた新たな数論的現象の解明

数の不思議はどこから生まれる？

どんな研究	数学では、一見まったく異なる２つの事柄が結びつくことがよくありますが、それは単なる偶然ではないはずです。本研究では、発表者がこれまでに開発した一般化モチーフ理論を用いて、数の世界の不思議な結びつきが生み出される本質的なメカニズムを明らかにします。
どこが凄い	数の研究は、代数多様体という図形の研究に置き換えられます。一般化モチーフ理論は、代数多様体をあらゆる視点から一挙に観測する数学的装置です。別々の視点（＝コホモロジー）から眺めた代数多様体の姿は一見、全く異なりますが、一般化モチーフ理論で隙間なく観測することで、異なる視点を結びつけることができます。
めざす未来	代数多様体の究極の観測装置である一般化モチーフ理論を用いれば、表面的には全く異なる２つの事柄を、偶然に頼らず系統的に結びつけることができます。本研究を推進することにより、人類の活動をいたるところで支える数の真理の探究を爆発的に加速させることができると期待されます。

数の世界の不思議な結びつき

方程式の有理数の解を調べることは非常に難しい問題です。しかし、驚くべきことに、グラフの穴の数を数えることで、方程式の有理数の解の個数の情報を得ることができます。

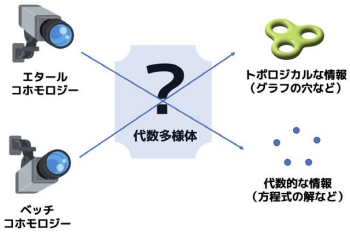


このように表面的には異なる２つの事象の結びつきを探り、難問への新しい解決策を見出すことは、数学の得意技です。

不思議を生み出す代数多様体

このような隠れた結びつきを探る数学的装置として、数学者はコホモロジーを作り上げました。コホモロジーは、複雑な図形から数値的な情報を取り出す優れた観測装置です。

解の個数や穴の数といった異なる情報は、代数多様体という図形を異なるコホモロジーで観測した結果として得られます。他にも多様な情報を代数多様体から引き出すことができます。（代数多様体は暗号理論などの応用分野でも重要な研究対象です。）

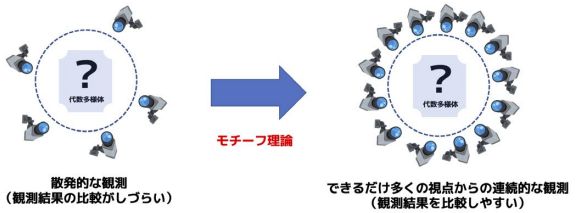


代数多様体は目に見えないが、コホモロジーを使うと情報を引き出せる。

引き出したい情報の種類に応じて、別々のコホモロジーが用いられる。

代数多様体を一挙に観測

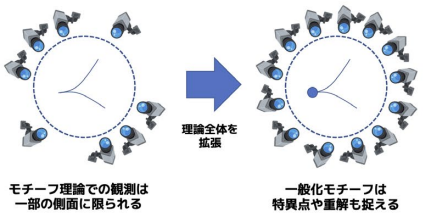
コホモロジーには、図形の着目する側面に応じて様々な種類がありますが、これら一つ一つを個別に眺めているだけでは代数多様体の異なる姿をうまく比べることができません。そこで、多くの側面を連続的に、一挙に観測してしまおう、という理論（モチーフ理論）が２０世紀後半に作られました。



モチーフ理論により異なる側面を比較しやすくなったことで、新たな事象の結びつきが次々に発見されています。

究極の観測装置：一般化モチーフ

しかし、実際にはモチーフ理論で観測できる範囲は限られており、特異点などの重要な情報が抜け落ちてしまいます。これまでの研究では、従来のモチーフ理論の欠点を克服する上位互換として、一般化モチーフ理論の開発に成功しました。



図形の滑らかでない点を特異点といい、図形の重要な特徴を含んでいる。従来のモチーフ理論では、特異点が持つ情報をほとんど観測できない。

今後は、一般化モチーフを用いた精度の高い観測を通して、従来のモチーフ理論では捉えられないさらなる未知の不思議なつながりを探究していきます。

関連文献

[1] B. Kahn, H. Miyazaki, S. Saito, T. Yamazaki, “Motives with modulus, III,” *Annals of K-theory* (to appear).
[2] B. Kahn, H. Miyazaki, S. Saito, T. Yamazaki, “Motives with modulus, II,” *Épijournal de Géométrie Algébrique*, Vol. 5, epiga:7115, 2021.
[3] B. Kahn, H. Miyazaki, S. Saito, T. Yamazaki, “Motives with modulus, I,” *Épijournal de Géométrie Algébrique*, Vol. 5, epiga:7114, 2021.
[4] B. Kahn, H. Miyazaki, “Topologies on schemes and modulus pairs,” *Nagoya Mathematical Journal*, Vol. 244, pp. 283–313, 2021.
[5] H. Miyazaki, “Nisnevich topology with modulus,” *Annals of K-theory*, Vol. 5, pp. 581–604, 2019.

連絡先

宮﨑 弘安（Hiroyasu Miyazaki） 基礎数学研究センタ
Email: cs-openhouse-ml@hco.ntt.co.jp

Robust speech summarization against speech recognition errors

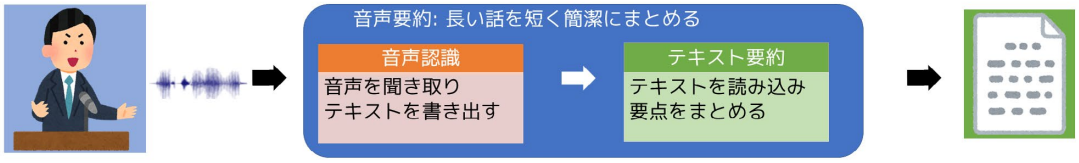
14 音声認識誤りに頑健な音声要約

「え、なに？どういうこと？」長い話まとめます

どんな研究	人の話の意図を正しく汲み取る機械の実現をめざして、長い話の要点を的確にまとめる音声要約の研究をしています。音声要約は、人の話を聞く音声認識と、聞いた内容をまとめるテキスト要約で構成されますが、前段の音声認識誤りを完全に無くすことは困難です。本展示では、音声認識誤りに頑健な音声要約技術を提案します。
どこが凄い	音声認識器がいくつかの単語を誤認識しても、話全体の文脈や単語間の関連性を考慮しながら正しく要約する音声要約を実現しました。複数の音声認識器を使うことで、それぞれの認識結果に誤りが含まれていても、精度の高い音声要約を実現します。また本技術は、音声要約のみならず音声翻訳などにも応用できることを確認しています。
めざす未来	人と人、人と機械がお互いに深く理解しあえる未来をめざします。人の言っていることを一言一句聞き取るだけではなく、人が何を意図しているのか、感じているのか、を理解し、人と人、人と機械のより豊かなコミュニケーションを支える音声言語処理技術を研究していきます。

目的：聞き間違えても正しく理解する

音声要約：講演などの長い音声から要点をまとめる技術
高度な音声理解を音声認識とテキスト要約を組み合わせて実現



しかし、音声認識の聞き間違いがテキスト要約に大きな影響を与えることも！

発話内容: 当社のサーモン大好き弁当の、内容量を再考します。検討結果はWEBで発表します。
認識結果: 当社のさぁ問題　好き弁当の、内容量は最高です。　検討結果はWEBで発売します。
⬇
要約結果：　大容量の「さぁ問題好き弁当」をオンラインで発売！

提案：音声認識誤りに頑健な要約

聞き間違いが無い音声認識の実現は困難
・複数の音声認識器を使い、正解単語がいずれかの結果に含まれるように様々な認識結果を出力

聞き間違いがあるかもしれないという前提でテキスト要約
・最先端の言語処理器（BERT*）を活用して、単語の意味や関係性を解析
・複数の認識結果を比較し、文脈に沿った単語に着目して正確に要約
* Bidirectional Encoder Representations from Transformers

音声認識器A	当社のさぁ問題　好き弁当の、内容量は最高です。　検討結果はWEBで発売します。
音声認識器B	当社のサーモン_好き弁当＿＿＿＿を再考します。検討結果はWEBで発表します。
音声認識器C	当社のサーモン大好き弁当は、内容量を最高します。検討結果はWEBで発布します。

⬇

文脈を考慮した仮説統合　（BERTを利用した音声要約器を構築）

１つ１つの認識は間違っても、認識器A,B,Cの単語の意味や内容を比較すると
- 「さぁ問題好き弁当」より「**サーモン大好き弁当**」の方が文として自然
- 話の後ろに、「検討結果」が続くということは**「再考」**であるべき
- 「発表」と「発布」も似た意味で、「検討結果」なら「**発表**」が自然

⬇

要約結果：サーモン大好き弁当の内容量を変更

提案法により、
認識器の聞き間違いによる性能劣化を最大30%低減することができます！

関連文献

[1] T. Kano, A. Ogawa, M. Delcroix, S. Watanabe, “Attention-based multi-hypothesis fusion for speech summarization,” in *Proc. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 487–494, 2021.
[2] 叶高朋, 小川厚徳, マーク・デルクロア, 渡部晋治, “音声要約のためのBERTを用いた認識仮説統合,” 日本音響学会2022年春季研究発表会
[3] T. Kano, A. Ogawa, M. Delcroix, S. Watanabe, “Integrating multiple ASR systems into NLP backend with attention fusion,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2022.

連絡先

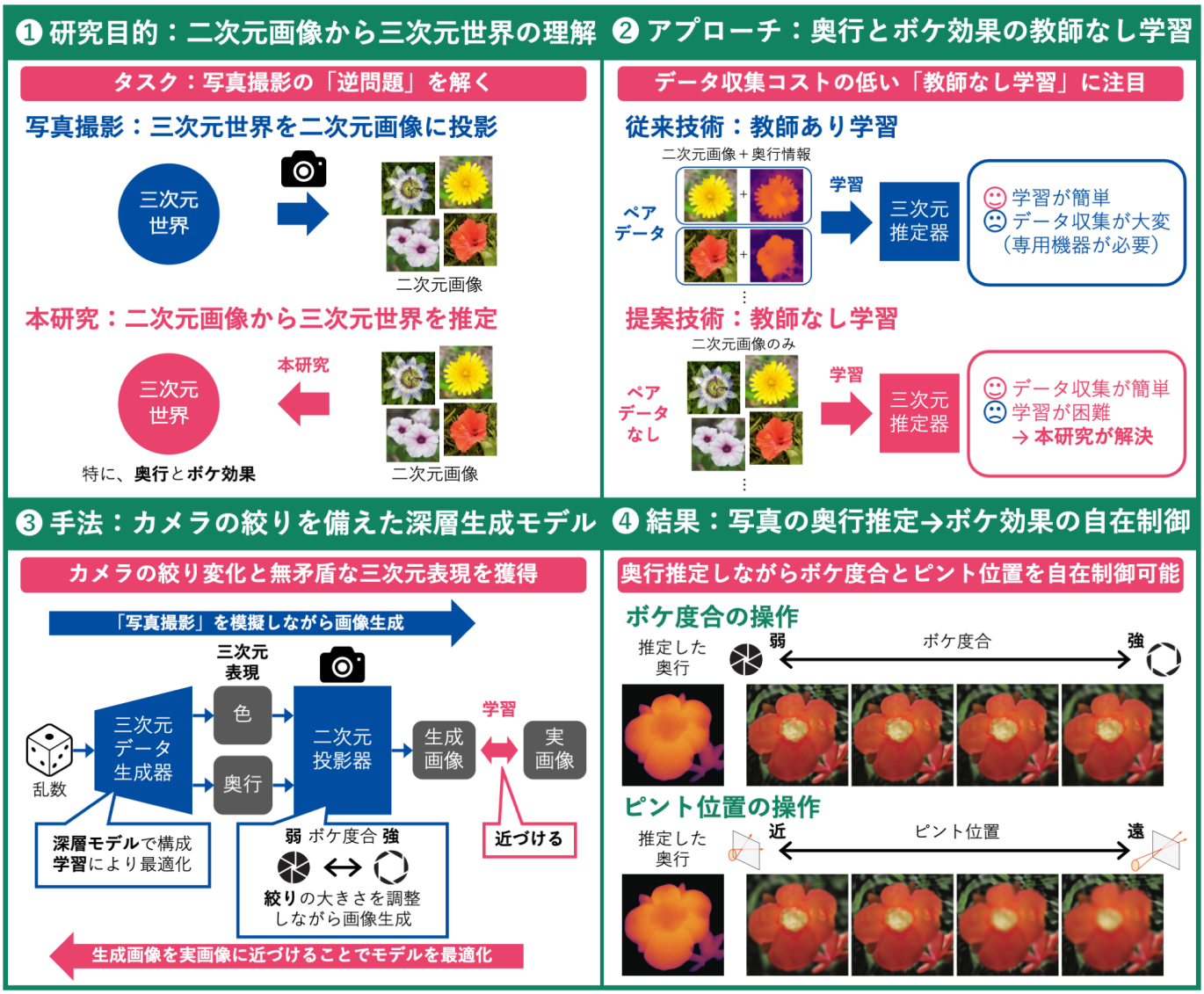
叶 高朋（Takatomo Kano） メディア情報研究部 信号処理研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Deep generative model for learning depth and bokeh effects only from natural images

15 写真群のみから未知の奥行とボケ効果を学習する深層生成モデル

写真の奥行を推定し、ボケ度合を自在に変えられます

どんな研究	ヒトは写真を見れば経験や知識から奥行や写真のどこがボケているかを推定することができますが、コンピュータにはそうした経験や知識がないため容易ではありません。これに対して、本研究では、カメラの絞りにおける光学的制約を組み込んだ深層生成モデルを新たに構築することで写真の奥行を推定しボケ効果を制御することを可能にしました。
どこが凄い	一般に三次元推定器を学習するためには、特殊な機器で三次元情報を収集する必要がありました。一方、本技術で学習に必要なデータはWeb上で見かけるような一般的な写真だけであり、奥行やボケ効果の情報は必要ありません。本技術は世界で初めて写真のみから未知の奥行とボケ効果の学習を可能にしました。
めざす未来	私たちは三次元世界で生活をしており、ヒトと親和性の高いコンピュータを実現するためには三次元世界を理解可能なコンピュータの実現が不可欠です。本技術は学習に必要なデータの収集コストを削減し、三次元情報を活用した様々な応用上の障壁を低減できるものであり、コンピュータによる三次元世界の理解において新たな可能性を切り拓く技術であると期待しています。



関連文献

[1] T. Kaneko, “Unsupervised learning of depth and depth-of-field effect from natural images with aperture rendering generative adversarial networks,” in *Proc. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR2021)*, pp. 15679–15688, 2021.
[2] T. Kaneko, “AR-NeRF: Unsupervised learning of depth and defocus effects from natural images with aperture rendering neural radiance fields,” in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR2022)*, 2022 (to appear).

連絡先

金子 卓弘（Takuhiro Kaneko） メディア情報研究部 メディア認識研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Estimating heart activities from multichannel sounds and ECG signals

16 マルチチャンネルの心音と心電からの心臓内部の活動推定

音と電気の信号で心臓の働きを見守ります

どんな研究	心臓の異常を早期に発見するには、日常的に容易に計測できる情報に基づいて心臓の活動を推定する必要があります。そこで、体表面でとらえた音響信号や電気信号から、体内の音の発生源を推定したり、心筋細胞の活動電位を反映して電気的活動を可視化したりする技術を研究しています。
どこが凄い	心音の生成モデルを取り入れた振動子分解技術により、通常の聴診器では直接には聴き取れない心臓内部の複数の弁の開閉などの音源波形を推定可能になりました。さらにテンソル心電図と呼ぶ電気的活動の可視化により、一般的な心電図には表れにくい異常を明瞭に検出できる可能性を示唆しました。
めざす未来	手軽に自分で計測でき循環器の状態や機能が詳しく分かるパーソナルなシステムを実現し、心不全、虚血性心疾患、突然死と関連するような不整脈などの心疾患の早期発見や、疾病治療後のリハビリテーション、および健康な方のトレーニングのサポートなどへの貢献をめざします。

心臓内部の活動推定

タスク：生体情報の観測 ➡ 心臓内部の活動

【手がかりとする情報】 侵襲なく容易に計測できる生体情報

心音（心臓の力学的活動に由来する音響信号）
心臓内部の弁の開閉や血流によって生じる振動成分を体表面で観測したもの
専門医は心音の聴診から心臓の健全性や病態を推定できる

心電（心臓の電気的活動に由来する電気信号）
心臓を構成する多数の心筋細胞の活動電位の集合体を体表面で電位差として観測したもの
これを図示した心電図は医療機関等で広く利用されている

【技術的ハードル】
心音や心電は体内の多数の発生源から生じた混合信号のため観測した信号だけから個々の発生源を突き止めることは困難

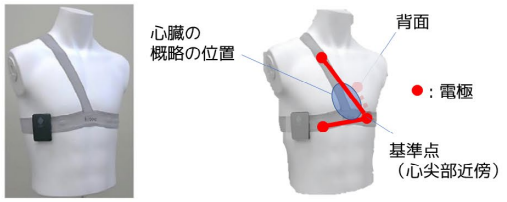
【アプローチ】
① 統計的／物理的モデルを活用した情報処理を工夫
② 空間的な手がかりを得るためにマルチチャンネルで観測

心電の技術† ➡ 心音の技術† ➡

† 研究段階（薬機未認証）

テンソル心電図

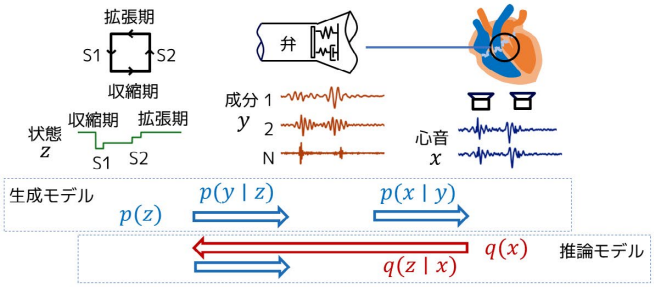
① 心筋細胞の集団の電位変化（脱分極と再分極）のタイミングをガウス分布で統計的にモデル化 [1]
② 心臓と体表面が最も近接する点（心尖部近傍）を基準点としほぼ直交する３軸上に対電極を配置することにより空間的な情報を取得（3チャンネル）



ウェアラブル心電計の外観と電極配置

生成モデルによる心音の振動子分解

① 心臓の動きの周期性と心音が発生する仕組みに着目し心音の発生機序を表現する確率的生成モデルを定義
【仮定１】弁の物理モデルに基づいた振動成分（振動子）が複数存在
【仮定２】それらの振幅が心周期の状態に応じて変化

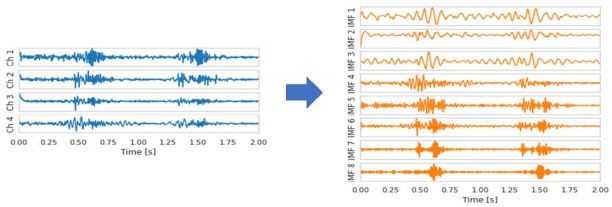


提案法の有効性を示す実験結果の例

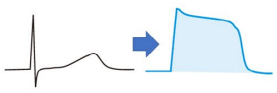
F1スコア	提案法	従来法
S1	**96.1**	86.5
S2	**96.4**	85.7

S1、S2 の 時間推定精度を従来法（心音の生成モデルを用いない振動子分解）と比較したもの [2]。

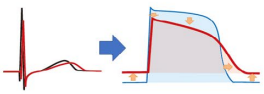
② 多チャンネルの観測信号への適用
体表面で観測した4チャンネルの音響信号から8つの振動成分を推定した例



観測信号（左）と 推定された振動成分（右）



体表面で観測される電位差（左）からの心筋の活動電位（右）の推定



テンソル心電図による心臓の異常の可視化（模式図）

関連文献

[1] S. Tsukada, “Wearable textile electrodes for long-term vector ECG monitoring 'Tensor Cardiography',” in *Proc. ISMICT 2020*, 2020.
[2] R. Shibue, M. Nakano, T. Iwata, K. Kashino, H. Tomoike, “Unsupervised heart sound decomposition and state estimation with generative oscillation models,” in *Proc. EMBC 2021*, pp. 5481–5487, 2021.

連絡先

渋江 遼平（Ryohei Shibue） 物性科学基礎研究所
塚田 信吾（Shingo Tsukada） 物性科学基礎研究所
Email: cs-openhouse-ml@hco.ntt.co.jp

Crossmodal action unit sequence estimation and image-to-image mapping

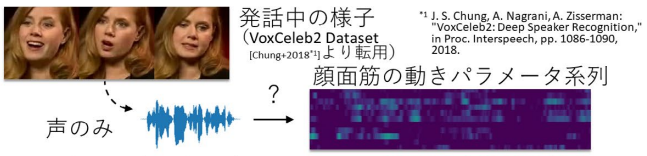
17 クロスモーダルアクションユニット系列推定と顔画像変換

声で顔画像の表情を動かす

どんな研究	音声には発話内容に相当する言語情報だけでなく、感情表現やムードに相当する非言語情報が含まれ、音声対話において重要な役割を担っています。本研究では、音声の非言語情報は話者の顔表情に表出されていると仮定し、音声のみから話者のアクションユニット（顔面筋パラメータ）を推定することを初めて試みた研究です。
どこが凄い	これまで音声のみからアクションユニットを推定する試みはなく、どの程度の精度を達成できるかは未知数でしたが、本研究ではこれを初めて明らかにしました。また、音声から推定したアクションユニットと画像変換器を用いることで、声に合わせて静止顔画像の表情を動かすシステムを実装し、声の表情や雰囲気を可視化することを可能にしました。
めざす未来	感情表現やムードは、従来、主観に基づく大まかなラベルにより記号的に扱われることが主流でした。これに対し、アクションユニットは感情表現やムードを細やかに表現する連続量として好適であり、本研究で音声からアクションユニットを推定できることを示しました。今後、顔表情に合った音声合成、音声に合った顔画像生成など、音声と顔画像を同時利用した様々な応用技術が拓けると期待しています。

音声から顔の動きを推定できるか？

✓ 声から顔面筋の動きを予測できれば・・・



*1 J. S. Chung, A. Nagrani, A. Zisserman: "VoxCeleb2: Deep Speaker Recognition," in Proc. Interspeech, pp. 1086-1090, 2018.

・予測結果を顔画像に転写することで音声の非言語情報を可視化可能に
・音声の非言語情報を表現する有用な特徴量として利用可能に
（音声合成や音声変換に応用可能）

✓ どれほど難しいのか？そのようなことが可能なのか？
➡ 本研究はこの問いに答えるもの

深層学習と音声付き動画像を用いた解法

✓ 顔面筋の動きを表したパラメータとして
アクションユニット(Action Unit; AU)[*2]に着目

*2 **顔の表情動作の最小単位で、眉毛・瞼・口・唇の動きを数値化したもの**

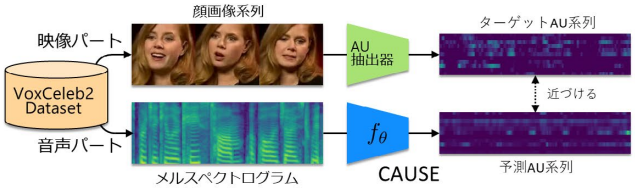
✓ 声からAUを予測するニューラルネットワークを学習
「Crossmodal Action Unit Sequence Estimator (CAUSE)」

アプローチ

大量の音声付き動画を用いて
- 映像パートからOpenFace[*3]により抽出されるAU系列
- 音声パートからCAUSEにより予測されるAU系列

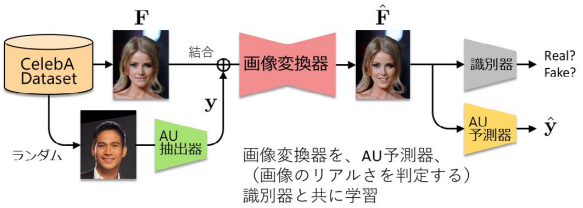
が近くなるようにCAUSEを学習

*3オープンソースの顔動作解析ツールキット



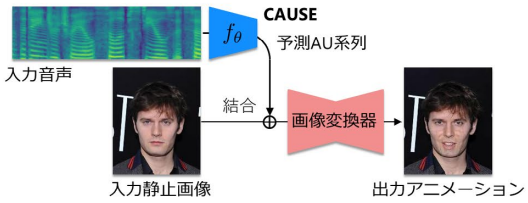
クロスモーダル顔画像制御

✓ 顔画像変換器の学習：GANimation [Pumarola+2018]



画像変換器を、AU予測器、（画像のリアルさを判定する）識別器と共に学習

✓ 声から予測したAU系列を用いて静止顔画像を変換



➡ 声を用いて顔画像の表情を制御することが可能

※顔画像はいずれもCelebA Dataset [Liu+2015*4]のものを使用

*4 Z. Liu, P. Luo, X. Wang, X. Tang: "Deep Learning Face Attributes in the Wild," in Proc. ICCV, pp. 3730-3738, 2015.

顔画像制御実験

その他の例はこちら

✓ 同一音声で異なる顔画像を制御した例

✓ AUの代わりに感情状態（平常、喜び、驚き、悲しみ、怒り、嫌悪、恐怖、軽蔑）の確率ベクトルを介して制御を行った場合に比べ、自然な顔アニメーションを生成できることを確認

関連文献

[1] H. Kameoka, T. Kaneko, S. Seki, K. Tanaka, “CAUSE: Crossmodal action unit sequence estimation from speech,” submitted to The 23rd Annual Conference of the International Speech Communication Association (Interspeech 2022).

連絡先

亀岡 弘和（Hirokazu Kameoka） メディア情報研究部 メディア認識研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Image blending with content-adaptive visibility predictor

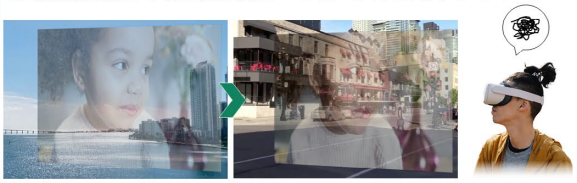
18 コンテンツ適応的な視認性予測に基づく画像ブレンディング

いつでもどこでも快適な「見えやすさ」を

どんな研究	画像を背景に対し半透明にブレンドする時、その見えやすさ（視認性）は背景の影響を受けて変化するため、望んだ視認性を保つことは困難です。本研究では、任意の画像の組をブレンドした際の視認性を予測できる知覚モデルを用い、どんな画像同士でも一定の視認性を保てるブレンド技術を実現しました。
どこが凄い	ブレンドする画像コンテンツがもつ模様の細かさや色等の特徴の偏り方によって、全体的な視認性に対する各特徴の寄与度が変わってくることを明らかにしました。この効果を取り入れた視認性予測モデルを用いることで、ブレンド画像の視認性を従来技術より正確に制御できるようになりました。
めざす未来	今後はAR技術等の成熟に伴って、視界内に別の情報を半透明表示する場面が増えていくと考えられます。本技術により、このような情報の視認性を快適に保つことができるようになります。また、映像編集ソフト等で、ブレンド画像の視認性をこれまでより直感的に操作可能になります。

半透明画像の視認性問題

視野全体を覆うようなメディア(ex. AR/VR)では、画像情報の半透明表示がしばしば必要とされる

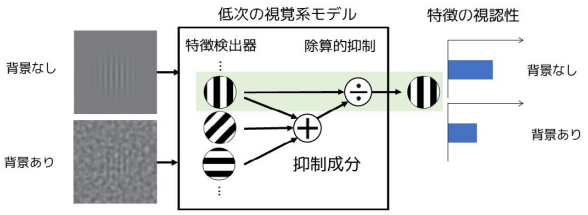


半透明表示するコンテンツや背景が変化する状況で見えやすさ（視認性）を一定に保つのは困難

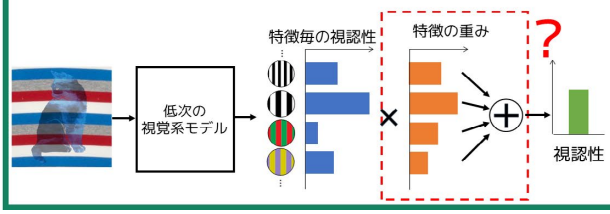
➡ 視認性を正確に予測できる知覚モデルが必要

視認性に関わる視覚メカニズム

画像に含まれる各特徴（パターンの細かさや色など）の視認性が背景画像によって低下する現象は視覚系の抑制メカニズムで説明可能



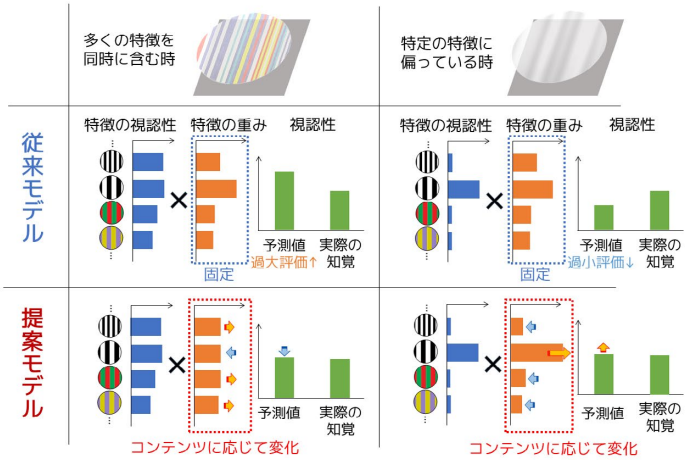
自然画像のように多くの特徴を同時に含むような場合に、特徴毎の視認性がどのように最終的な視認性に寄与するかは、よくわかっていなかった



技術ポイント① コンテンツ適応的な特徴集約

従来モデル 各特徴を予め決められた値で重み付け

提案モデル 表示する画像コンテンツがもつ特徴の偏り方に応じて各特徴の重み付けを適応的に調節

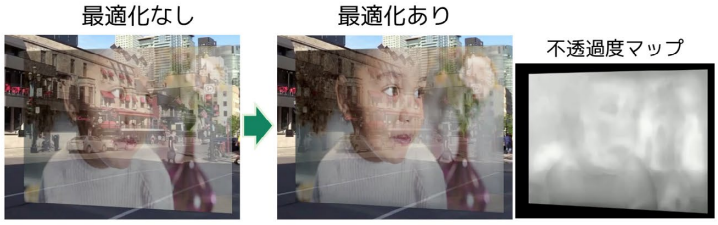


従来モデルは元々特徴が多い画像は視認性を過大評価、特徴が少ない画像は視認性を過小評価する傾向

提案モデルはより高い精度で視認性予測が可能に！

技術ポイント② 視認性ベースの画像ブレンディング

ユーザーが指定した視認性を保つように透過度を自動最適化するブレンディング技術を実現



本技術は東京大学との共同研究成果です

関連文献

[1] T. Fukiage, T. Oishi, “A computational model to predict the visibility of alpha-blended images,” *Vision Sciences Society Annual Meeting 2021* (Abstract published in: Journal of Vision, Vol. 21, No. 2493).

[2] 吹上大樹, 大石岳史, “コンテンツ適応的な視認性予測モデルに基づく知覚ベースの画像合成,” コンピュータビジョンとイメージメディア研究会 (CVIM), Vol. 229, No. 45, pp. 1–8, 2022.

連絡先

吹上大樹（Taiki Fukiage） 人間情報研究部 感覚表現研究グループ

Email: cs-openhouse-ml@hco.ntt.co.jp

Elucidating effects of social cues on infants' object learning

19 社会的手がかりが乳児期の物体学習に及ぼす影響の解明

まなざしや声かけが赤ちゃんの学びを助けます

どんな研究	乳児は環境から得られる情報に基づき様々な知識を学んでいますが、その詳細な学習プロセスは明らかにされていません。本研究では、乳児の物体学習や言語獲得が、声かけや視線など養育者からの社会的手がかりによって促されることを実験心理学的アプローチにより明らかにしました。
どこが凄い	乳児の学習に社会的手がかりと注意手がかりのどちらが寄与するか議論がありましたが、学習プロセスを明確にする2種類のテストを行い社会的手がかりの重要性を実証しました。この結果は、人間の学習メカニズムをモデル化した「ナチュラル・ペダゴジー理論」を支持し発展させます。
めざす未来	効果的な学習に及ぼす要因やメカニズムを明らかにすることにより、ヒトが言語や知識をどのように習得するかに関する理論を確立し、ペアレントトレーニングや乳児の学習環境考案などの育児や教育支援に貢献することをめざしています。

乳児の学習理論

乳児の学習を助ける手がかりについて**２つの学習理論が対立**

ナチュラル・ペダゴジー理論

コミュニケーションシグナルである
社会的手がかりが、学習に効果的

- アイコンタクト
- 声かけ
- ジェスチャ

注意覚醒理論

乳児の注意を喚起する**注意手がかり**により学習が成立

- 機械音
- 点灯
- 頭振り

本研究のアプローチ

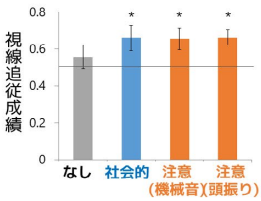
- 社会的手がかりと注意手がかりによる学習への影響を実験心理学的アプローチにより検証
- 各手がかりが付与された条件を設けて比較

対象児：140名の9ヶ月児

成果１：社会的手がかりの役割の明確化

■ 視線追従テスト

女性モデルの視線方向の物体（ブタの人形）を見るかを測定

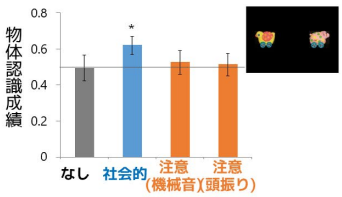


- 社会的・注意手がかりのいずれも物体学習の基礎となる視線追従を促進するが、両者で有意差なし

■ 物体学習テスト

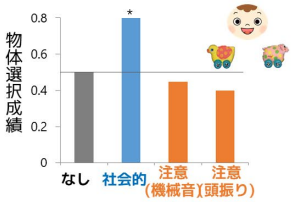
（1）物体認識テスト

注視時間に基づき物体の記憶認識を評価



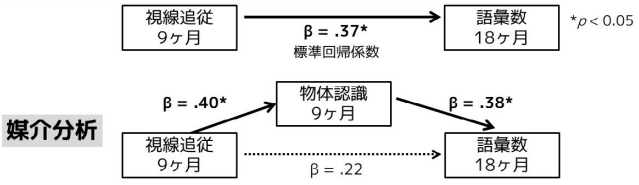
（2）物体選択テスト

選択行動に基づき物体選好を評価



- 社会的手がかりは物体学習場面において効果的
  ➡ **ナチュラル・ペダゴジー理論を支持**

成果２：語彙習得プロセスの解明



- 9ヶ月の視線追従頻度が物体認識を促し、18ヶ月の語彙数に影響
  ➡ **社会的手がかり（視線）と語彙獲得の関連性を発見**

関連文献

[1] Y. Okumura, Y. Kanakogi, T. Kobayashi, S. Itakura, “Individual differences in object-processing explain the relationship between early gaze-following and later language development,” *Cognition,* Vol. 166, pp. 418–424, 2017.
[2] Y. Okumura, Y. Kanakogi, T. Kobayashi, S. Itakura, “Ostension affects infant learning more than attention,” *Cognition*, Vol. 195, 104082, 2020.
[3] 奥村優子, “乳児期における社会的学習: 誰からどのように学ぶのか,” 東京大学出版会, 2020.

連絡先

奥村 優子（Yuko Okumura） 協創情報研究部 インタラクション対話研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Social factors influencing adoption of COCOA

20 COCOAの普及に影響を与える社会的要因
なぜ人は接触確認アプリの利用をためらうのか

どんな研究	COCOAはコロナ感染症対策の一つとして期待されますが、その普及率は低迷しています。本研究でCOCOAのインストールや利用実態を調査した結果、人々は社会的規範や社会便益を意識して行動する一方で、偏見などを恐れて陽性登録を行わない傾向があることがわかりました。
どこが凄い	本研究結果は、たとえアプリのインストール数が増えたとしても、アプリ本来の効果が得られない可能性があることを示唆しています。アプリの普及に努めるだけではなく、正しい利用を促す対策の必要性、特に人々の不安を解消することの重要性を指摘しました。
めざす未来	本研究は、個人の権利や利益を守りつつ社会的便益を実現する情報技術の設計方法に示唆を与えるものです。この研究をさらに発展させることで、SDGsや環境問題など、個々の行動を結集して社会的便益を実現するような問題の克服に貢献することが期待されます。

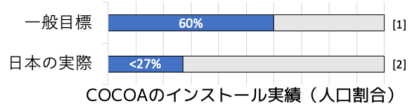
新型コロナウイルス接触確認アプリ（COCOA）

COCOA は、日本政府が提供する新型コロナウィルス接触確認アプリです。

濃厚接触の疑いのある利用者間の接触をBluetoothによって検知・記録し、その旨を通知します。

利用者が増えることで、感染拡大の防止につながることが期待されています。

日本国内にとどまらず、多くの国で、接触確認アプリの普及率が低いことが報告されています。



COCOAのインストール実績（人口割合）

→ アプリの利用意向を促進・阻害する要因を理解する必要性

[1] Hinch, R., et al., "Effective configurations of a digital contact tracing app: A report to NHSX," 2020.
[2] 厚生労働省, "新型コロナウイルス接触確認アプリ," 2022. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/cocoa_00138.html

研究背景と調査設計

既存研究により、利用者が接触アプリをインストールするか否かは、アプリの有効性、使いやすさ、社会的な影響、プライバシー侵害への懸念などに影響されることが分かっています。

一般に、感染拡大を防止するために次の行動を取ることが期待されます。

新型コロナウィルス感染症の陽性と診断されると、アプリに陽性者登録を行う。

接触通知を受けると、なるべく自己隔離をして体調の様子をみたり、職場や関係者に連絡するなど、適切に対応する。

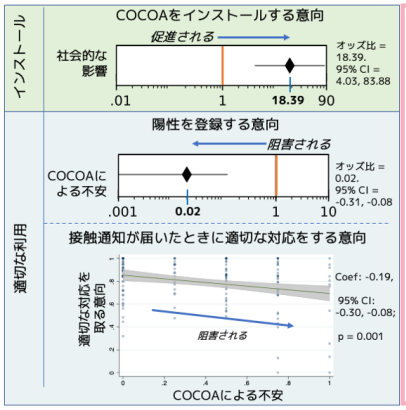
本研究では、接触確認アプリをインストールした後の利用状況や日常生活への適合性を調査します。

調査方法： アンケート（153人）、インタビュー（15人）

従属変数
- *COCOAインストールの有無*
- *陽性登録に対する意向*
- *接触通知に対して適切に対応する意向*

独立変数
- *社会的な影響（アプリをインストールしている知人の有無）*
- *アプリ利用による不安（アプリを利用することによって不安が増幅するか否か）*

アプリの利用意向に影響する因子



社会的な影響
= インストールの促進要因

COCOAを利用することによって生じる不安
= 感染拡大防止に役立たない行動を取る傾向:
a) 陽性を登録しない
b) 接触通知が来ても適切な対応をしない

インストール数が増えても感染拡大防止に役立つとは限らない

29% の人がCOCOAの利用によって不安が増幅されると回答した理由

社会的偏見にさらされる恐れ	経済的損失に対する恐れ
"田舎に住んでいるので、匿名でもすぐに誰が感染者かわかってしまうし、感染後に色々な噂が立つのも怖い。"	*"職場に伝えたら職を失ったというニュースを見るので、通知が来たことを考えると少し心配です。"*

COCOAを取り巻く現状
- 接触通知が来てもそのことを隠す
- 感染しても陽性登録をしない

今後必要な取り組み

アプリの利用によってネガティブな事柄が想起されやすい→
社会的報酬が得られる仕組みをアプリに導入

社会的偏見などの不安によって適切にアプリを利用できない→
社会的偏見や経済的損失に対する補償など、ユーザを取り巻く環境を変えるための組織的な取り組みの重要性

接触確認アプリを超えて→
集団の利益に役立つ情報技術を研究開発することの必要性

関連文献

[1] J. Jamieson, N. Yamashita, D.A. Epstein, Y. Chen, "Deciding if and how to use a COVID-19 contact tracing app: Influences of social factors on individual use in Japan," in *Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 481, (CSCW'21)*, pp. 1–30, 2021.
[2] J. Jamieson, D.A. Epstein, Y. Chen, N. Yamashita, "Unpacking intention and behavior: Explaining contact tracing app adoption and hesitancy in the United States," in *Proc. the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*, pp. 1–14, 2022.

連絡先

ジャック ジェーミソン（Jack Jamieson） 協創情報研究部 インタラクション対話研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

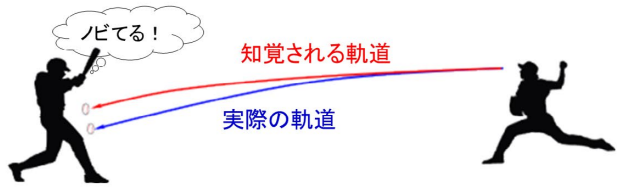
Modifying pitched ball perception by VR

21 VRで投球の質感を操作する
ノビのある速球は錯覚？

どんな研究	投手の投球の質は、球速や軌道などのボールの物理的特徴量だけでなく、「ノビ」といった打者の主観で評価されることがあります。本展示では、バーチャルリアリティ（VR）を用いて「ノビ（ボールが浮き上がる知覚）」の脳情報処理メカニズムを明らかにしました。
どこが凄い	VR環境で投手が投げたボール（速球）を野球選手が観察しているときに、投手の投球動作時間を人工的に操作すると、ボール挙動は同じにも関わらず、「ノビ」の知覚は変化しました。この結果から、打者は投球動作パターンからボールの挙動を予測し、その予測がボールが「ノビ」る“錯覚”を生み出していることがわかりました。
めざす未来	VRを用いることで、本研究のように実環境では観測が難しいアスリートの知覚メカニズムを明らかにできるだけでなく、意思決定（判断力）など様々な認知トレーニングやこれまでにないコーチングなどへの応用も期待されます。

投球の知覚

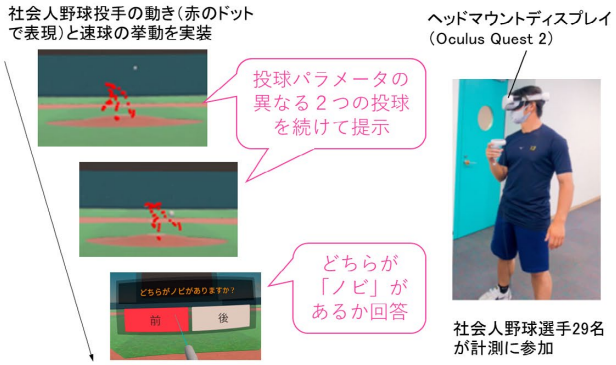
投球の質は、球速や軌道などのボールの物理的特徴量だけでは決まらない



「ノビ（ボールが浮き上がる知覚）」のように、打者の主観で評価されることがあるが、その知覚メカニズムは未解明

VRを用いて「ノビ」知覚を評価

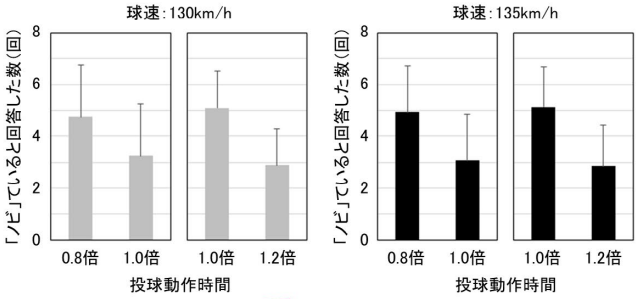
バーチャルリアリティ(VR)を用いて、(ボール軌道は一定のまま)投手の投球動作時間(動作刺激の再生時間)や球速などのパラメータを操作して「ノビ」知覚を評価



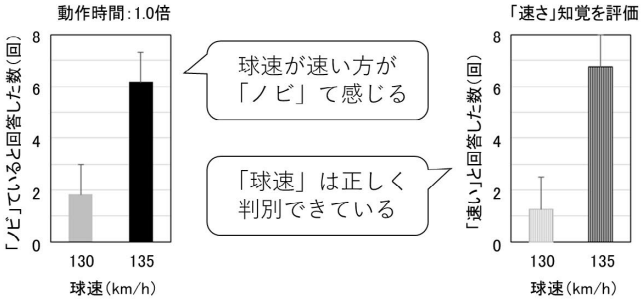
＊本研究は，慶應義塾大学との共同研究です．

投球動作時間を操作すると「ノビ」が変わる

異なる投球動作時間(0.8/1.0/1.2倍)で、「ノビ」ていると回答した試行数の比較



ボールの物理的な挙動（球速や軌道）が同じでも、投球動作時間が短い方が「ノビ」て知覚される



◆投球の「ノビ」知覚は、球速などの物理的特徴量だけでなく、投手の動作との関係から生成される。
◆「ノビ」は、投球動作に基づいたボール挙動の予測が生み出した“錯覚”を含んでいる。
◆VRを用いた視覚情報の操作は、アスリートの認知メカニズムの解明だけなく、認知トレーニングにも応用できると期待される。

関連文献

[1] T. Fukuda, A. Endo, M. Sugimoto, T. Kimura, “How do elite baseball batters perceive a "rising" fast ball?,” in *Proc. North American Society for the Psychology of Sport and Physical Activity 2022 Annual Conference*, 2022.
[2] 遠藤愛奈, 福田岳洋, 三浦礼二, 杉本麻樹, 木村聡貴, “バーチャルリアリティを用いた野球の投球の質感に影響を及ぼす投球パラメータの解析,” スポーツ工学・ヒューマンダイナミクス2021, 2021.

連絡先

木村聡貴（Toshitaka Kimura） 柏野多様脳特別研究室
Email: cs-openhouse-ml@hco.ntt.co.jp

The relationship between frontal neural oscillation and performance

22 前頭脳領域の脳波とパフォーマンスの相互関係

脳計測で見えたeスポーツ熟練者の勝負術

どんな研究	eスポーツ選手は優れた**戦略判断能力**や**感情制御能力**を持ちますが、その能力が試合成績にどのように影響するのかはあまり知られていません。この展示では、eスポーツの実戦における選手の脳活動に着目することで、上記の能力と試合成績の関係を明らかにしました。
どこが凄い	eスポーツ上級者を対象とした試合中の脳波計測と試合後のアンケートによって、**試合の序盤では戦略判断能力が重要**になり、**終盤では感情制御能力が重要**になることが分かりました。さらに勝利した試合の序盤と終盤では、それぞれの能力と関連して変化する脳活動が**前頭葉に由来**することが世界で初めて観察されました。
めざす未来	eスポーツ選手の精神状態を試合の局面に応じた理想的な状態へ近づける**新しいトレーニング手法の確立**をめざします。また、選手の能力拡張だけでなく、選手が抱える課題の克服につながる特定の脳機能の解明も期待できます。

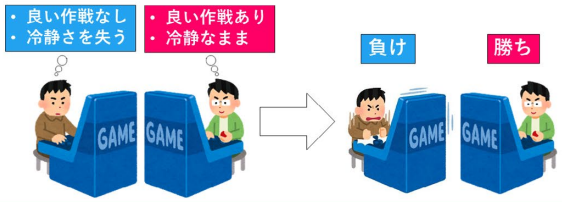
eスポーツで重要な精神的準備

- eスポーツの試合では**精神的準備**が重要
- 特に上級者は対戦直前の**戦略判断**と**感情制御**に優れる

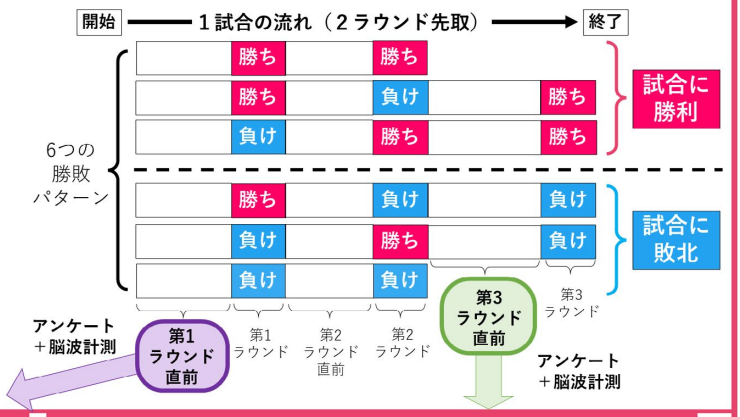
戦略判断：相手の特徴から対戦中に取るべき行動パターンを事前に推測する試み

感情制御：大事な局面で対戦前に感じる精神的動揺を意識的に抑制する試み

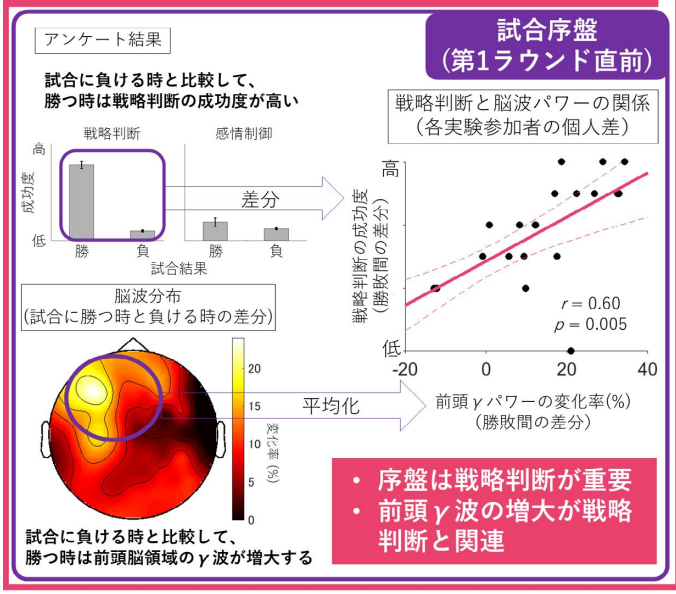
2つの能力が試合成績にどのように影響するのかを、局面ごとに必要な能力が変わる**格闘ゲーム**に着目し調査



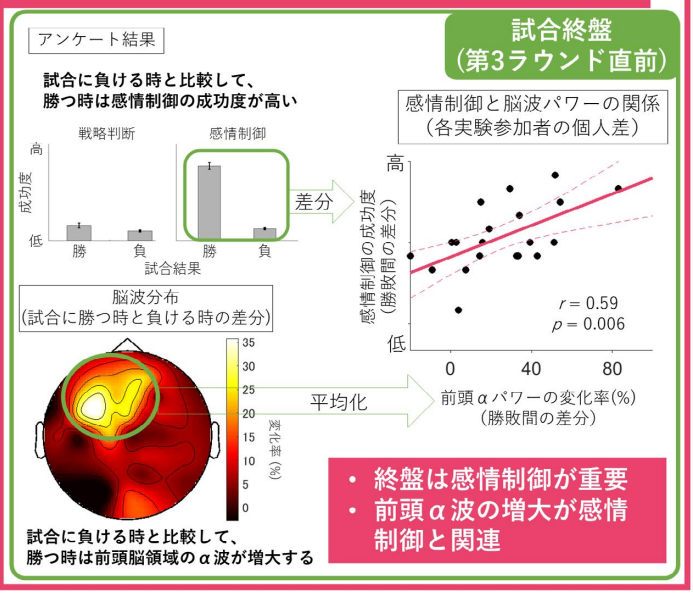
格闘ゲームの試合形式と勝敗パターン



試合序盤に重要な戦略判断と関連する脳活動



試合終盤に重要な感情制御と関連する脳活動



関連文献

[1] S. Minami, K. Watanabe, N. Saijo, M. Kashino, “Amplitude of neural oscillations in the parietal area is associated with the results of esports competitions,” in *Proc. IEEE Conference on Games (CoG)*, 2021 (in press).

連絡先

南 宇人（Sorato Minami） 柏野多様脳特別研究室
Email: cs-openhouse-ml@hco.ntt.co.jp

Approaches to cochlear implant and binaural processing

23 人工内耳と両耳情報処理へのより深い理解のために
人工ニューラルネットワークで紐解く聴覚システム

どんな研究	従来、人の獲得した聴覚特性そのものについては多く語られてきました。しかしその特性がどのような入力と学習を介して獲得されたのかという問題に直接取り組むのは非常に困難です。そこで本展示では、臨床的・学術的な側面の問題に人工ニューラルネットワークを用いて取り組み、それぞれ新たな知見が得られました。
どこが凄い	①人工内耳を装用した難聴者はピッチ知覚が苦手と言われていますが、人工内耳信号の中にある程度十分なピッチ情報が含まれていることを発見し、主に生理学的な要因に起因することを明らかにしました。②自然音を認識するように訓練した人工ニューラルネットワークの中身を調べたところ、生物の聴覚系にみられるものと同等の両耳情報処理特性を示す素子（ニューロン）が発現することを発見しました。
めざす未来	人工内耳装用後も適切なリハビリテーションによって、静音環境下であれば健常者と大差ないピッチ知覚を獲得出来るようになるかもしれません。また、聴覚神経系に合致した聴覚情報処理技術をより高度にしていくことで、人らしい振る舞いをするAI技術や人工内耳のような装置の更なる発展をめざします。

人工ニューラルネットワーク（ANN）を用いた聴覚の理解

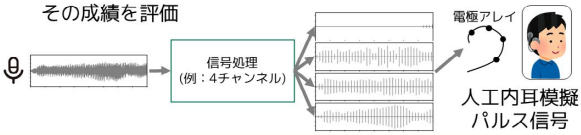
テーマ①：人工内耳※が伝達するピッチ情報

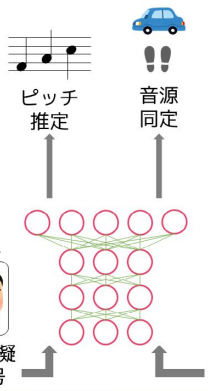
※蝸牛の損傷等による難聴を回復させる人工臓器の1つ

背景：
- 人工内耳装用により、音声知覚は大きく回復
- しかしピッチ知覚は困難でその個人差も大きい

問い：
- 人工内耳が伝達する信号にそもそもピッチ推定に必要な情報は含まれているのか？

アプローチ：
- 人工内耳模擬パルス信号からANNを用い直接ピッチ推定し、その成績を評価



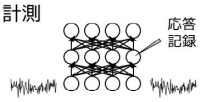


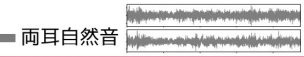
テーマ②：両耳音への神経細胞の応答特性

背景：
- 音が左右の耳に届くまでの時間差（両耳間時間差、ITD）は音源の定位や同定に重要
- 神経細胞の応答特性が調べられてきた

問い：
- ITDへの応答特性はどのように獲得される？

アプローチ：
- 両耳自然音に含まれる音源を同定するANNを訓練しITDを持つ刺激への単一細胞の応答を計測

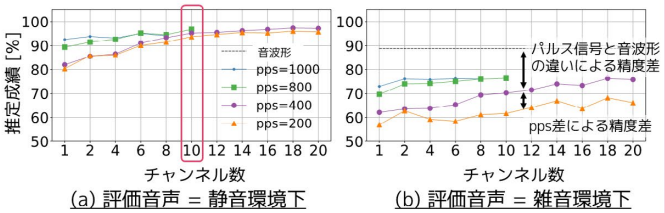




①ピッチ情報は人工内耳信号に一定量包含

■ 対象となる音が単独で提示された場合(静音環境下(a))では、十分なチャンネル数(約10チャンネル)があれば音波形入力と同程度の精度
▶ 人工内耳信号は音波形に匹敵するピッチ情報を含む

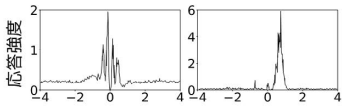
■ 雑音が同時に存在する場合(雑音環境下(b))では、音波形入力との差は大きいが、秒間パルス頻度(pps)が増加すると精度改善
▶ 雑音環境下ではピッチ知覚は困難
▶ 細かい時間解像度が重要な役割を果たす



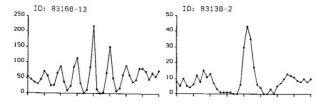
■ ピッチ知覚が困難である事は、人工内耳デバイスが伝達する信号そのものより、生理学的な要因が大きい事を示唆

②自然音認識で発現するITD応答特性

■ ANN細胞の応答強度が刺激のITDによって変化
▶ 生物と定性的に類似したITD応答特性がANNにも発現

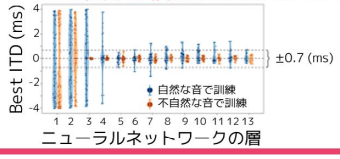


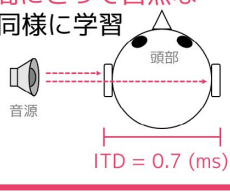
発現したITD応答特性の例



動物のITD応答特性の例
(Yin+ 1986 J. Neurophysiol.)

■ 人間が自然に経験するITDの範囲で応答強度が高い
▶ モデル構築には人間の体型の情報を不使用
▶ 不自然な音で訓練すると、その範囲が狭くなる
▶ 音に含まれる情報のみから人間にとって自然な環境の構造がANNにおいても同様に学習





関連文献

[1] 芦原孝典, 古川茂人, 柏野牧夫, “DNNモデルによる人工内耳模擬パルス信号からのF0推定,” 日本音響学会春季研究発表会, 2022.
[2] T. Koumura, H. Terashima, S. Furukawa, “Emergence of ITD selectivity in a deep neural network trained for binaural natural sound detection,” in *Proc. 42nd Association for Research in Otolaryngology (ARO) MidWinter Meeting*, 2019.
[3] TC. Yin, JC. Chan, DR. Irvine, “Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise,” *Journal of Neurophysiology*, pp. 280–300, 1986.

連絡先

芦原 孝典（Takanori Ashihara） 人間情報研究所
上村 卓也（Takuya Koumura） 人間情報研究部 感覚表現研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Autonomic and endocrinological variation by meditation style

24 二つの瞑想法がもたらす自律神経系・内分泌系の状態変化

マインドフルネス瞑想はストレスをどう減らす？

どんな研究	マインドフルネス瞑想は注意を制御することでストレスを低減することが知られていますが、そのメカニズムは明らかにされていません。本研究では、ストレスに関連している自律神経活動やホルモン分泌に着目し、瞑想がこれらの生理活動をどのように変化させているかを明らかにしました。
どこが凄い	マインドフルネス瞑想を構成する集中瞑想と洞察瞑想の技法に対し、それぞれの音声インストラクションを開発しました。これにより、集中瞑想はリラックスの指標である副交感神経を賦活化し、洞察瞑想は活力の指標である交感神経を賦活化しながらストレスホルモン濃度を下げることを発見しました。
めざす未来	瞑想の生理・心理・神経メカニズムの解明を進め、科学的知見に基づいた新しい瞑想法の開発をめざします。開発した瞑想法を社会に普及させることで、人々のwell-being実現に貢献できると考えています。

1.マインドフルネス瞑想は“集中瞑想”と”洞察瞑想“の２技法から構成される

マインドフルネス：
次々と生じている今この瞬間の感覚や感情や思考に受容的な注意でありのままに気づいている状態

マインドフルネス瞑想：

集中瞑想
呼吸に一点集中

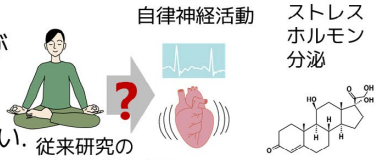
洞察瞑想
様々な感覚や感情をありのままに観察

マインドフルネス瞑想
ストレス低減の効果が有名. うつ再発予防やwell-being向上にも効果あり

マインドフルネス瞑想の主要な構成要素

2.従来研究では２技法それぞれの生理学的機序が未解明

背景：
マインドフルネス瞑想がストレスを低減させる背後の生理学的機序が未だ不明. 統一見解が無い.



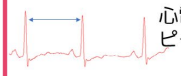
原因：
集中瞑想と洞察瞑想を分けて、それぞれの生理学的機序を検討していないため

解決法：
集中瞑想、洞察瞑想それぞれについて、30分間の音声インストラクションを開発

集中瞑想：「これから集中力のエクササイズを始めます。集中力とは・・・」という誘導から始まり、雑念が生じた時に呼吸に注意を向け続けるエクササイズをする.
洞察瞑想：「これから気づきのエクササイズを始めます。気づきとは・・・」という誘導から始まり、思考や感情を価値判断せずありのままに観察するエクササイズをする。

3.本研究で２技法が自律神経活動とコルチゾール値に異なる影響を与えることを解明

自律神経活動の解析：

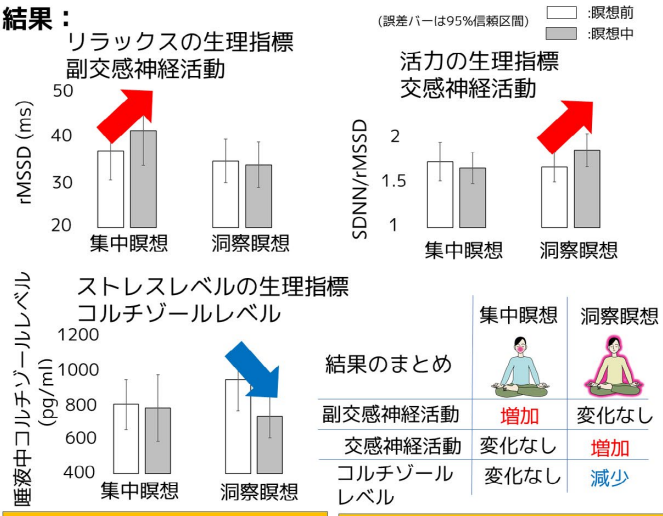


時間領域心拍変動解析で以下を算出
・SDNN
(standard deviation of normal-to-normal interval)
・rMSSD
(root mean square of successive differences)

ストレスホルモン分泌の解析：

唾液中のコルチゾール濃度を分析

結果：



結果のまとめ

	集中瞑想	洞察瞑想
副交感神経活動	増加	変化なし
交感神経活動	変化なし	増加
コルチゾールレベル	変化なし	減少

集中瞑想をすると
呼吸に一点集中することで他の情報に振り回されにくくリラックスできている可能性

洞察瞑想をすると
様々な情報を観察するため覚醒度は高く、ありのままに観察するためストレスは低い可能性

関連文献

[1] 大石悠貴, 藤野正寛, 井上ウィマラ, 野村理朗, 北川智利, “マインドフルネス研究：集中瞑想と洞察瞑想が自律神経系・内分泌系にもたらす効果,” 第40回日本神経科学大会, 2017.
[2] 藤野正寛, 上田祥行, 井上ウィマラ, イエットG. サンダーズ, スティーブンマーフィ重松, 野村理朗, “心理学実験のための集中・洞察・慈悲瞑想の短期介入インストラクションの開発,” マインドフルネス研究, 第 4 巻, 第 1 号, pp. 10–33, 2019.
[3] Y. Ooishi, M. Fujino, V.Inoue, M. Nomura, N.Kitagawa, “Differential effects of focused attention and open monitoring meditation on autonomic cardiac modulation and cortisol secretion,” *Front. Physiol.*, Vol. 12, 675899, 2021.

連絡先

大石 悠貴（Yuuki Ooishi） 人間情報研究部 感覚共鳴研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Well-being in terms of mental states, values, and idea of self

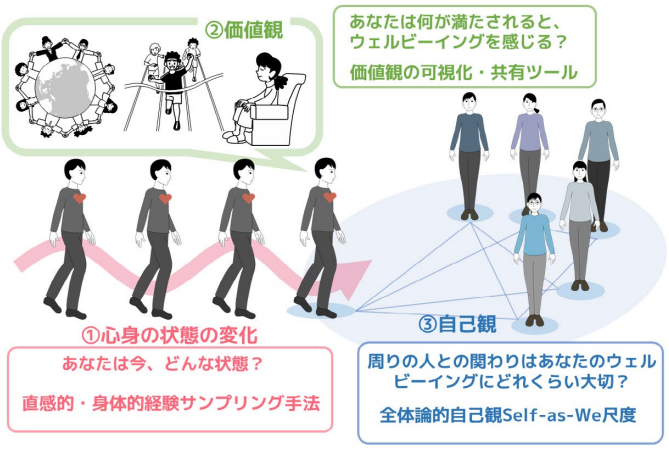
25 心身の状態変化・価値観・自己観から捉えたウェルビーイング

ウェルビーイングをいろいろな側面から測ります

どんな研究	わたしたちはどんな時にウェルビーイング（身体的・精神的・社会的にいきいきとした状態）を感じるでしょうか。それを知るためには ①心や体の状態 ②自分にとって大切なこと ③他者との関わりを総合的に捉えることが必要です。本研究ではこれら３つを計測する独自の手法を考案しました。
どこが凄い	①心身の状態を身体感覚を反映しつつ直感的に記録できる感性表現語を用いた新たな経験サンプリング手法を考案しました。②各個人で異なるウェルビーイングの価値観を簡易に可視化し共有できるツールを考案しました。 ③自己と他者を切り離さない東アジアの思想伝統に依る全体論的自己観の程度を定量化するSelf-as-We尺度を開発しました。
めざす未来	それぞれの人が暮らしの中でウェルビーイングを感じながら生きていくためには、自身の心身の状態や価値観・自己観の在り方に主体的に気がつき、評価し、他者とともに行動していく必要があります。本研究は、心理学や哲学、工学、デザインといった包括的な視点から、このプロセスを支援します。

概要

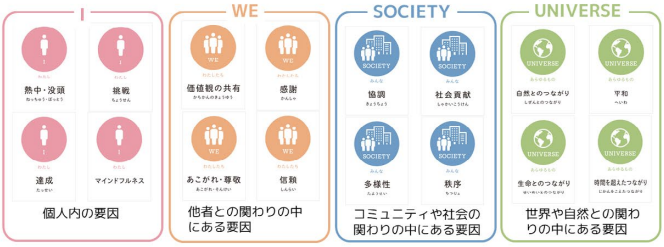
多様な人々が関わり支え合う社会の中でひとりひとりのウェルビーイングを実現するためには①②③を包括的に捉えることがその第一歩です。



② 価値観の可視化

自分や他者が何を大切にしているかを可視化し、気づきを促すためのツール

「自分の大切なこと」を表すカードを３枚選び、他の人と共有
➡自己の価値観への気づき・価値観の多様性の相互認識を促進

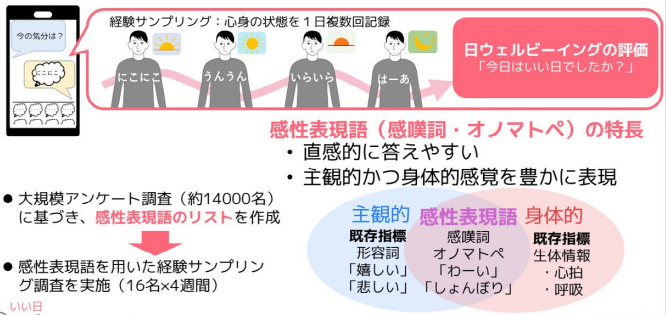


カードセットの一部を抜粋

- さまざまなウェルビーイングの要因を表すカードを作成
- 約1300名に「 あなたのウェルビーイングに大切なこと」を３つ挙げてもらい、約3900のウェルビーイングの要因をカテゴリ化

① 心身の状態の記録

心身の状態を直感的に記録できる感性表現語を活用した経験サンプリング手法

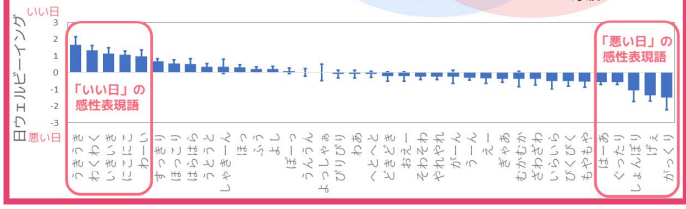


感性表現語（感嘆詞・オノマトペ）の特長
- 直感的に答えやすい
- 主観的かつ身体的感覚を豊かに表現

- 大規模アンケート調査（約14000名）に基づき、感性表現語のリストを作成
- 感性表現語を用いた経験サンプリング調査を実施（16名×4週間）

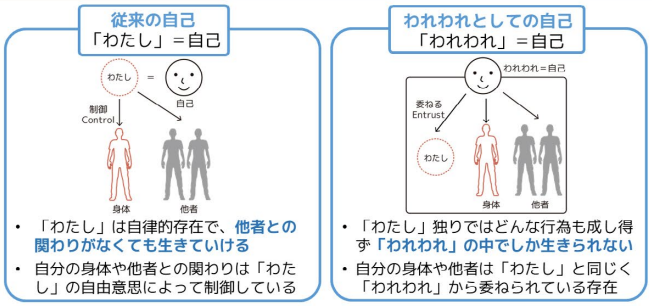
主観的 感性表現語 身体的

既存指標	感嘆詞	既存指標
形容詞	オノマトペ	生体情報
「嬉しい」	「わーい」	・心拍
「悲しい」	「しょんぼり」	・呼吸



③ 自己観の評価

他者との関わりの中で自己をどう認識しているか（＝自己観）を測るための心理尺度



- 「わたし」は自律的存在で、他者との関わりがなくても生きていける
- 自分の身体や他者との関わりは「わたし」の自由意思によって制御している

- 「わたし」独りではどんな行為も成し得ず「われわれ」の中でしか生きられない
- 自分の身体や他者は「わたし」と同じく「われわれ」から委ねられている存在

東アジアの思想伝統を踏まえた全体論的自己観「われわれとしての自己」※の程度を評価するSelf-as-We尺度を心理学の手法に基づき開発

※京都大学出口康夫教授が提唱

Self-as-We尺度項目例
- ✓『チームの取り組みで得られた成果はチームの成果であって、特定の誰かの手柄であるとは言えない』
- ✓『チームの活動に参加するときには、自ら行動している能動的な感覚と、チームの要請に従う受動的な感覚の両方を同時に感じる』

関連文献

[1] 渡邊淳司, 村田藍子, “ポジティブ・コンピューティングを自分事とするために：ウェルビーイングへの身体性からのアプローチ,” 感性工学, Vol. 18, No. 2, pp. 63–67, 2020.
[2] 渡邊淳司, ドミニク・チェン監修・編著, 安藤英由樹, 坂倉杏介, 村田藍子編著, “わたしたちのウェルビーイングをつくりあうために その思想、実践、技術,” ビー・エヌ・エヌ新社, 2020.
[3] 渡邊淳司, 村田藍子, 高山千尋, 中谷桃子, 出口康夫, “「われわれとしての自己」を評価する—Self-as-We尺度の開発—,” PROSPECTUS, Vol. 20, pp. 1–14, 2020.

連絡先

村田 藍子（Aiko Murata） 人間情報研究部 感覚共鳴研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Vision-based speedometer regulates human walking

26 視野の動きに基づく歩行速度制御の仕組みの解明

壁が動くと速く歩く？

どんな研究	歩いている時の視覚情報は、通路上や脇の障害物を検出するというだけでなく、歩行を制御するためにも使われています。本研究では、バーチャルリアリティ(VR)で与える視覚情報を操作し、どのような情報を脳が検出して歩行運動の制御に使っているかを調べました。
どこが凄い	VR環境の中で壁のある通路を歩行している最中に、壁を前後に動かすと、歩行速度の変化が生じます。様々な環境条件における歩行速度の変化から、脳は、網膜速度や時間周波数情報ではなく、壁までの推定距離に基づき歩行速度を計算して歩行を調節していることを明らかにしました。
めざす未来	人は、体の動きに伴って時間変化する視覚情報を、様々な形で無意識のうちに四肢や歩行の運動制御に利用しています。このような脳の潜在的な情報処理メカニズムを明らかにしていくことは、人に使いやすい安全なバーチャルリアリティ環境のデザインなどに役立つことが期待されます。

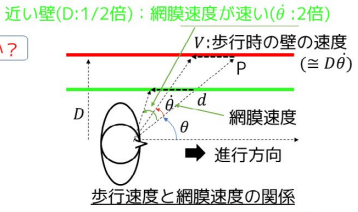
歩く速度はどのように調節されているか？

- 人は、巧みに足を動かして一定の速さで歩くことができる。
- 歩行動作を行うためには、筋・関節の状態や頭部の動きをコードする感覚情報に加え、目の前の障害物や路面の状態を知るための「視覚情報」も重要。
- また視覚情報は、歩行速度を実時間で調節する機能においても重要な役割を持つことが知られていた。

問題点：脳が何の情報を計算して歩行速度調節に結び付けているかは理解されていなかった。

脳は視覚情報から何を計算しているか？

仮説１：網膜速度を計算
仮説２：時間周波数を計算
仮説３：歩行速度を計算

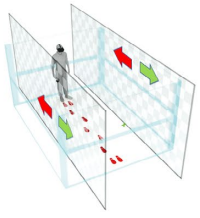


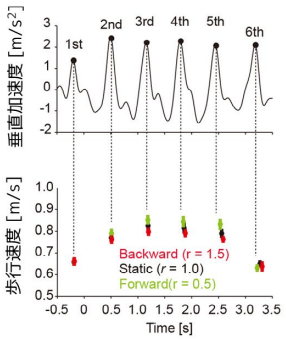
歩行速度と網膜速度の関係

歩行中に壁を動かすと歩行速度が変化する

ヘッドマウントディスプレイを使って、仮想的な壁を設けた通路を見せて歩行。

歩行中壁を前に動かすと歩行速度が上がり、後ろに動かすと下がる現象が起こる（無意識な歩行速度調節）。





壁の見え方の変化から脳情報処理を探る

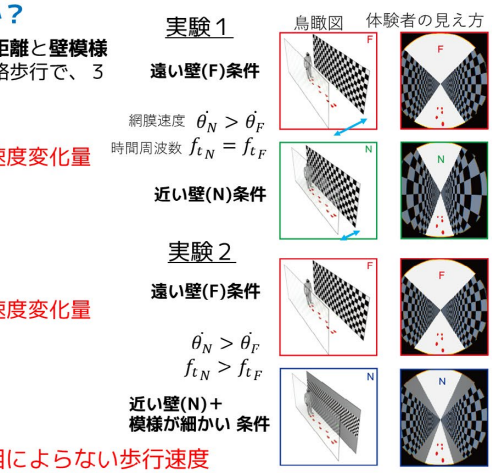
網膜速度・時間周波数の計算により歩行速度が調節されているか？

実験：壁までの距離と壁模様の粗さの違う通路歩行で、３つの仮説を検証

実験１の結果
2条件で、歩行速度変化量は不変
⇒仮説１は棄却

実験２の結果
2条件で、歩行速度変化量は不変
⇒仮説２も棄却

結論：壁の様相によらない歩行速度が計算されている（仮説3を支持）

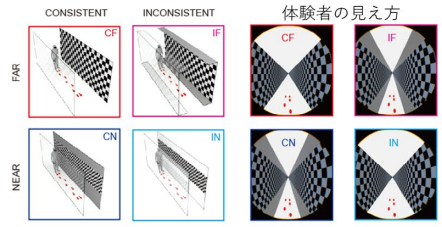


歩行速度計算に必要な深度（壁距離）推定に、両眼・単眼奥行き手がかりが寄与するか？

実験：両眼・単眼奥行き手がかりで距離推定を誤る条件での通路歩行で、各奥行き手がかりの重要性を検討

IF: 遠い壁に庇や縁をつけると、体験者の見え方はCNと酷似

IN: 近い壁の上下を無地にすると、体験者の見え方はCFと酷似



結果：壁距離推定を誤る条件で、歩行速度変化量に変化が見られた

脳は、両眼・単眼奥行き手がかりを使って歩行速度を計算し、歩行制御に使っていることが明らかになった　$V \cong D\dot{\theta}$

関連文献

[1] S. Takamuku, H. Gomi, “Vision-based speedometer regulates human walking,” *iScience 24:103390*, 2021. https://doi.org/10.1016/j.isci.2021.103390.

連絡先

五味 裕章（Hiroaki Gomi） 人間情報研究部 感覚運動研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

How the brain decodes pulling sensations

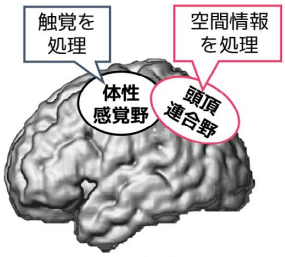
27 引かれる方向感覚を生成する脳の情報処理

指先で感じる“こっち”

どんな研究	手に持ったものの重さを感じる・手をつないだ相手の意図を汲み取るなど、私たちは指先で感じる力の情報によって様々な現象を知覚しています。本研究では、脳活動計測を通して、このような手指が引かれる感覚を生み出している脳の情報処理の仕組みを探っています。
どこが凄い	指先の振動によって手が引かれて感じる錯覚現象を利用することで、力の方向を感じる処理に関わる脳活動を計測しました。その結果、手指が引かれる感覚の生成には、従来知られていた触覚を処理する脳部位だけでなく、頭頂連合野と呼ばれる空間や方向の情報処理を司る脳部位の活動が関わっていることを明らかにしました。
めざす未来	指先で力の方向を感じる脳の仕組みを明らかにすることで、触覚を用いた新しい情報提示手法への応用が期待されます。例えば、目の不自由な方の手を引いて道案内するモバイルデバイスや、義手で指先の感覚を再現する技術の研究開発に役立つ可能性が考えられます。

手が引かれる力の感覚

人間は手指に加わる力の感覚によって様々な情報を知覚



このような感覚は脳のどこで生じている?

仮説：手が引かれる力の感覚には**頭頂連合野**（空間処理の脳部位）が関与するのでは？

手が引かれる感覚に関わる脳活動の計測

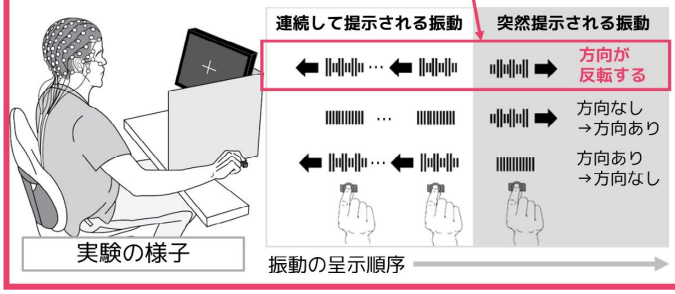
“指先の非対称振動によって手を引かれて感じる錯覚”を利用

3種類の振動を呈示:
右向き、左向き、方向なし

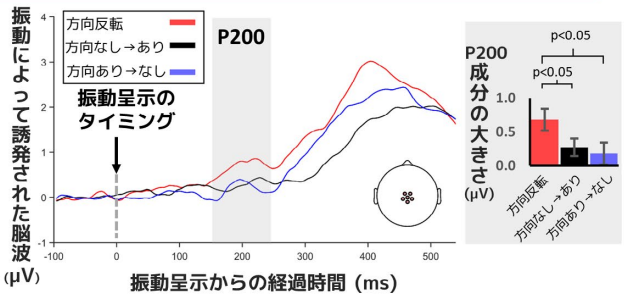
力の方向に関するoddball課題*1を実施し、脳波を計測

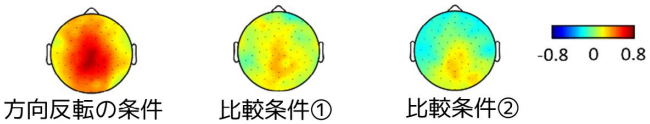
*1: 連続した同一の刺激の後で、突然、異なる刺激を呈示

方向の反転が起こったときの脳活動に着目



結果：空間処理の脳部位の関与が示された



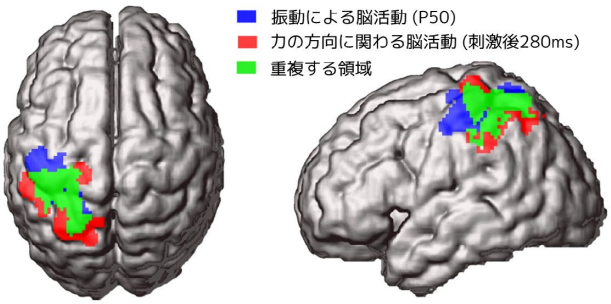


方向反転の条件
(右→左 or 左→右)

比較条件①
方向なし→あり

比較条件②
方向あり→なし

脳波の**P200成分** (空間情報の処理に関与) の大きさ:
力の方向が反転する条件 > 他の実験条件
→ この脳活動が**手が引かれる方向の処理に関与**



力の方向に関する処理 (赤)：体性感覚野（青：触覚を処理）に加え、**頭頂連合野 (空間処理) が重要**

関連文献

[1] J. De Havas, S. Ito, S. Bestmann, H. Gomi, “Neural dynamics of illusory tactile pulling sensations,” *bioRxiv*, 2021. https://doi.org/10.1101/2021.10.12.464029

連絡先

ジャック デ ハバス（Jack De Havas） 人間情報研究部 感覚運動研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Understanding of desire to touch using large-scale Twitter data

28 大規模なTwitterデータを活用した触りたさの理解

みんな何を触りたいの？

どんな研究	人々は日常の暮らしの中でどんなものに触りたいと思っているのでしょうか。本研究では、人々がソーシャルネットワーク上で「○○を触りたい」とつぶやいた膨大なテキストデータを収集・解析することで見えてきた、日常における触りたさについてご紹介します。
どこが凄い	解析を通して、触りたい身体部位と触り方との間には一定の関連性があることが示されました。また、肌の温もりを求める欲求であるスキンハンガーの慢性化や、非生物への接触回避願望の変動が新型コロナウイルス感染拡大によって生じたことも明らかにしました。
めざす未来	本研究で得られた触りたさの知見は、幅広い実世界の問題に応用可能です。日常における触りたさのメカニズムの解明、消費者が自然と触りたくなる製品の設計、感染拡大のような実際の社会問題が人々の意識に与える影響の監視など、様々な場面での貢献が期待されます。

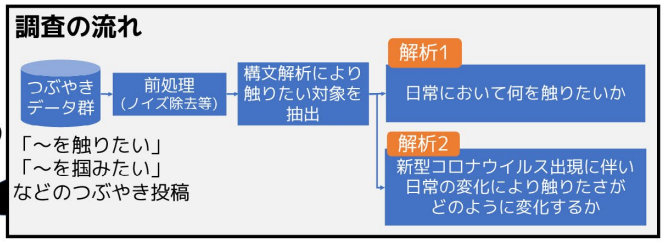
触りたさの理解

問い： 日常において人はどのようなものを触りたいと思っているのでしょうか？

従来研究： 3Dプリンティングした物体など、日常で触る機会の少ない物体への触りたさを調べており日常における触りたさは明らかにしていません

本研究： 「○○を触りたい」というTwitterに投稿された膨大なテキストデータを解析し、日常における触りたさを理解します

調査の流れ



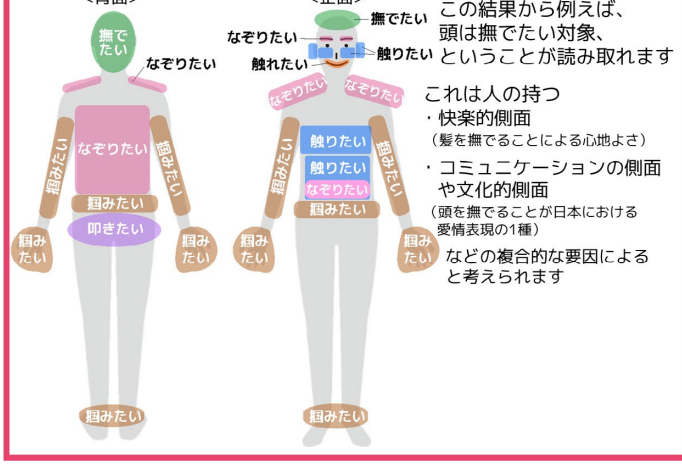
解析1: 日常 × 触りたさ

つぶやきの構文解析により**触りたい対象と触り方の関連性**を示しました

触り方	1位	2位	3位	4位
触る	胸	髪	お尻	猫
触れる	あなた	人々	肌	猫
撫でる	頭	猫	犬	お腹
掴む	腰	手	お尻	しっぽ
押す	ボタン	スタンプ	カート	お腹
叩く	ドラム	お尻	キーボード	頭
なぞる	体のライン	腹筋	まつげ	筋肉

身体部位に着目した触りたい対象と触り方の関係の可視化

身体部位-触りたい触り方マップ

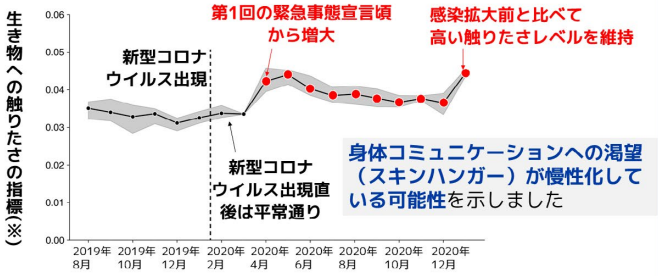


この結果から例えば、頭は撫でたい対象、ということが読み取れます

これは人の持つ
・快楽的側面
（髪を撫でることによる心地よさ）
・コミュニケーションの側面や文化的側面
（頭を撫でることが日本における愛情表現の1種）
などの複合的な要因によると考えられます

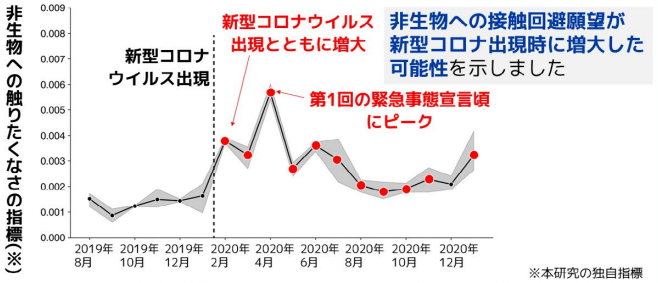
解析2:日常の変化(新型コロナウイルス) × 触りたさ

新型コロナウイルス出現後に、人や動物など生き物への触りたさがどのように変化したかを解析しました



身体コミュニケーションへの渇望（スキンハンガー）が慢性化している可能性を示しました

新型コロナウイルス出現後に、ドアノブなど非生物への触りたくなさがどのように変化したかを解析しました



非生物への接触回避願望が新型コロナ出現時に増大した可能性を示しました

※本研究の独自指標

上記の2つには**時間特性の違い**があります。**スキンハンガーは感染拡大直後は見られず、緊急事態宣言頃から見られ始めました。一方、接触回避願望は感染拡大直後から見られ、緊急事態宣言頃に更に高まりました**

関連文献

[1] Y. Ujitoko, Y. Ban, T. Yokosaka, “Getting insights from Twitter: What people want to touch in daily life,” *IEEE Transactions on Haptics*, Vol. 15, No. 1, pp. 142–153, 2022.
[2] Y. Ujitoko, T. Yokosaka, Y. Ban, H. Ho, “Tracking changes in touch desire and touch avoidance before and after the COVID-19 outbreak,” *PsyArXiv*, 2021.

連絡先

宇治土公 雄介（Yusuke Ujitoko） 人間情報研究部 感覚表現研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

Pupil size tracks subjective perceptual changes

29 瞳孔径から読み取る知覚の変化
瞳が語るこころの状態

どんな研究	同じ音声を聞いていても、その「聞こえ方」が時間経過に伴って変化する場合があり、この現象は多義的知覚と呼ばれます。この聞こえ方の変化は知覚交替と呼ばれます。本研究では、聞こえ方に対する知覚交替の頻度と瞳孔径の大きさが関係していることを明らかにしました。
どこが凄い	知覚交替は脳内神経回路の確率的な状態変化を反映しています。本研究では、このような脳内における見えない状態変化を瞳孔径の大きさを通じて客観的に捉えることを可能にしました。
めざす未来	本研究成果を活かして、瞳孔径のような眼球運動計測から得られる情報から脳神経回路の内部状態などのリッチな情報取得技術が発展すると期待されます。更には、本人すら自覚していない心の情報を読み取る技術（マインド・リーディング）の実現を目指しています。

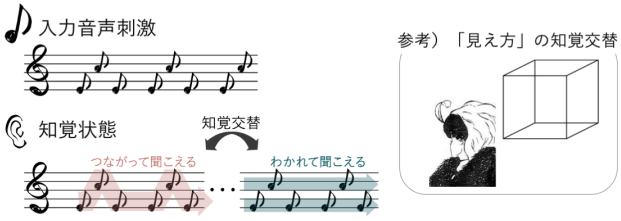
瞳孔径から読み取る知覚の変化

同じ音を聞いていても、その時々で「聞こえ方」が違い、個人ごとに異なるタイミングで変化していく

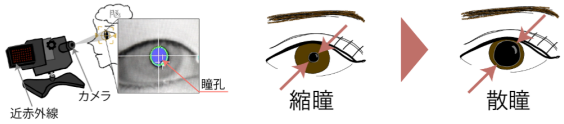
このような音の解釈を生じさせる脳内の神経回路の変化を **瞳孔径が事前に反映している** 可能性を示した

多義的な音に対する知覚（知覚交替）とは？

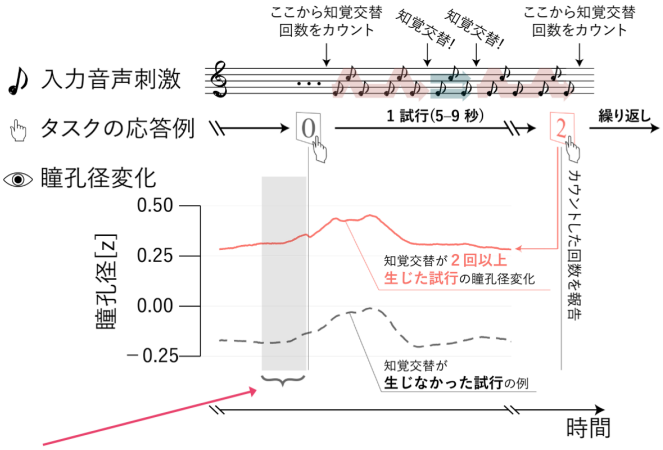
一定の入力信号に対して、
「聞こえ方」（知覚）が **時間的に変化（交替）**する



瞳孔径計測:
自律神経系（ノルアドレナリン）と関連して**覚醒度の指標**になる

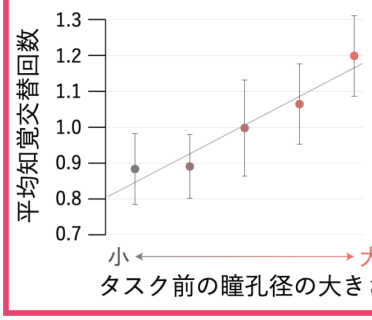


知覚交替のカウントタスクと瞳孔径計測実験



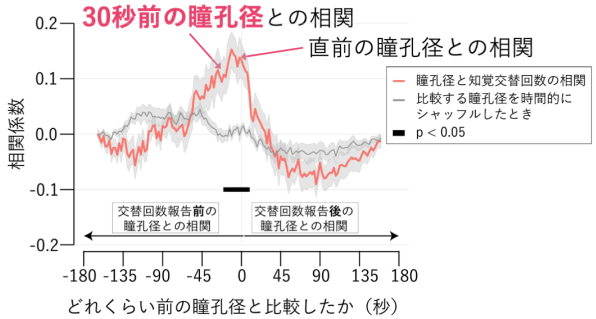
タスク（知覚交替が生じた回数のカウント）前の瞳孔径が大きいときに**知覚交替がより多く**生じることを発見

ノルアドレナリンレベルの変化が知覚交替（脳神経回路の遷移のしやすさ）と関連している可能性



知覚交替カウントタスクの開始前の
瞳孔径が大きいときに
知覚交替回数が多い

瞳孔径と知覚交替回数の相関は数十秒持続する



関連文献

[1] Y. Suzuki, H. Liao, S. Furukawa, “Temporal dynamics of auditory bistable perception correlated with fluctuation of baseline pupil size,” *Psychophysiology*, 2022. doi:10.1111/psyp.14028

連絡先

鈴木雄太（Yuta Suzuki） 人間情報研究部 感覚表現研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp

NTT コミュニケーション科学基礎研究所
オープンハウス 2022 実行委員会

（委 員 長） 田中　貴秋
（副委員長） 廣谷　定男
（委　　員） 石畠　正和／大國　智樹／俵　直弘／千葉　祐弥／成松　宏美／武　小萌／目黒　豊美

NTT
NTT コミュニケーション科学基礎研究所
〒619-0237 京都府相楽郡精華町光台2-4 (けいはんな学研都市)
TEL : 0774-93-5020 FAX: 0774-93-5026