

入場無料·事前登録不要

会場:NTT京阪奈ビル

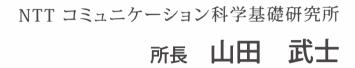
京都府相楽郡精華町光台2-4(けいはんな学研都市)







「オープンハウス2018」へようこそ





NTT コミュニケーション科学基礎研究所「オープンハウス2018」へご来場頂き、心より 御礼申し上げます。

昨今、人工知能(AI)やビッグデータ、IoTなどが日々、注目を浴びています。NTTグループ では、コンピュータと人間が協創する人工知能であるcorevo®に取り組んでいます。 corevo®がめざすのは、人の活動の一部を代替・支援して、人と共存、共創することによ って、人に寄り添い、人の生活を豊かにする人工知能の実現です。

NTT コミュニケーション科学基礎研究所は、人と人、あるいはコンピュータと人の間の 「こころまで伝わる」コミュニケーションの実現をめざし、時代を先取りした基礎研究に 取り組んでいます。ここで生み出される、人間と情報の本質に迫る基礎理論や革新技術は、 未来のための地図や羅針盤であり、corevo®を支える基盤となるものです。

「オープンハウス2018」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」の4領域について、最新の研究成果をわかり やすくご紹介致します。この二日間のイベントが、ご来場の皆様にとって、いわば未踏の地を めざして大海原に船出する、ワクワクする知的冒険の体験となれば幸いです。



ごあいさつ	
会 場 案 内 図 / 講 演 スケ ジュール	03
所長講演	
****	04
招待講演	
ロ 15 時	06
	00
研究講演	
基本演算を操る量子コンピュータの真価 ~ゲート型量子コンピュータの計算能力の分析~ ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	
ウェルビーイングにおける触覚の役割 〜触れることの科学とデザインが人の心を豊かにする〜 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	
「組合せ爆発」を乗り超える ~二分決定グラフを用いた膨大な量の組合せの数え上げと最適化~ ···································	12
脳からみた聞くと話すの共通性 〜音声変換技術と脳機能計測による人間の音声コミュニケーションの仕組みの解明〜 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	14
研 究 展 示	17
(データと学習の科学	
01 膨大なデータから似た音声を見つけます! グラフ索引に基づく類似音声探索 ······	
02 複数の問題に共通して重要な情報の組合せを発見 共通因子を効率的に学習する低ランク回帰技術: MOFM · · · · · · · · · · · · · · · · · · ·	19
03 人はどこから来て、どこへ行くのか? 人流データ同化と学習型誘導 ·······	20
04. 都市の今を知る 環境センシングと異種データ融合分析によるイベント解析 · · · · · · · · · · · · · · · · · · ·	21
(05) 深層学習をモバイル向けに小さくします量子化による深層学習のモデル圧縮技術・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	22
07 ネットワーク構造から深層学習のしくみを知る ニューラルネットの理解に向けたコミュニティ抽出技術・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	2
リア・インドノーノ 博足がり休信 子目のし、のを知る ニューブルネットの建解に同じたコミュニア 1 油田収削	2
コミュニケーションと計算の科学	
○	25
09 こわれにくいネットワークをデザインします 二分決定グラフを用いたネットワーク信頼性最大化	
10 人工知能は文脈を読んで翻訳できるか? ニューラル翻訳の文脈理解度をテストする ····································	
11 ことばの発達がゆっくりなお子さんの特徴を探る 小児医療現場で収集したデータの解析からみえてきたこと・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	
12 物知りロボットとおしゃべりしながら賢くなろう 複数ロボット対話制御に基づく雑談と質問応答の融合・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	
13 いつでもどこでもそれっぽくしゃべれます! スマホで音声リズムを英語母語話者っぽく変換 ······	30
14 離れていても盛り上がりを共有できる 双方向性の臨場感・一体感の向上をねらう拍手音符号化	31
(メディアの科学	
15 照明光で色の鮮やかさを操る 彩度強調成分を用いた分光スペクトルの制御 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	
16 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 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	33
17 二択問題にして解くことでAIは賢くなる 深層学習による仮説比較と音声認識結果選択への応用 ······	34
18 音だけから情景を推定音から画像認識結果を予測するクロスメディア情景分析・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	35
19 浮像(うくぞう)影を駆使して絵に奥行きを与える光投影技術・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	37
21 選んで創るお気に入りの画像 DTLC-GANを用いた画像の階層的理解・生成制御 ····································	38
ZI ZIV (A) OU XIVC) () OU DIEC GIMENO REBOVEREDAM EMBE	
(人間の科学	
22 音への注意を眼で測る 瞳孔反応に現れる聴覚空間注意 ····································	39
23 ウェルビーイングを測る、知る、育む "いきいきとした心的状態"の実現を科学する領域横断研究 ····································	40
24 人工知能で人の聴こえの仕組みを理解する 機械学習モデルによる聴覚神経機構の分析 ····································	
25 一流打者はボールをどのように打っているか? 打撃中の身体運動計測から認知過程を探る ······	
26 一流打者はボールをどのように見ているか? 打撃中の眼球運動計測から認知過程を探る ······	
27 急ぐ方が正確? 視知覚と運動に潜む情報処理のからくり	44
28 力で感じるかたちとうごき ぶるなび4による環境・状況の呈示 · · · · · · · · · · · · · · · · · · ·	
29 平らなシートなのに凹凸感?! 磁性シートに凹凸感を"書き込む"磁性触覚印刷技術 · · · · · · · · · · · · · · · · · · ·	46
corevo® NTTグループのAI技術・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	48

会場案内図/講演スケジュール







講演スケジュール

5月31日(木)

所長講演

13:20~13:50

新たな次元へとシフトする

~さらに深化するコミュニケーション科学の取り組み~

コミュニケーション科学基礎研究所 所長 山田 武士

招待講演

14:00~15:00

AIと倫理および社会的影響

理化学研究所 革新知能統合研究センター 社会における人工知能研究グループ グループディレクター 中川 裕志

研究講演

15:30~16:10

基本演算を操る量子コンピュータの真価

~ゲート型量子コンピュータの計算能力の分析~

メディア情報研究部 高橋 康博

6月1日(金)

研究講演

11:00~11:40

ウェルビーイングにおける触覚の役割

~触れることの科学とデザインが人の心を豊かにする~

人間情報研究部 渡邊 淳司

研究講演

13:00~13:40

「組合せ爆発」を乗り超える

~二分決定グラフを用いた膨大な量の組合せの数え上げと最適化~

協創情報研究部 西野 正彬

研究講演

13:50~14:30

脳からみた聞くと話すの共通性

~音声変換技術と脳機能計測による 人間の音声コミュニケーションの仕組みの解明~

人間情報研究部 廣谷 定男



新たな次元へとシフトする

~さらに深化するコミュニケーション科学の取り組み~

Shift to new dimensions

Deepening and evolving Communication Science

NTT コミュニケーション科学基礎研究所 所長 武士 Takeshi Yamada 山田

コミュニケーション科学基礎研究所(CS研)は、1991年、米 国独立記念日と同じ日の7月4日に、NTTの情報系の基礎研究 を担う研究所として、けいはんな学研都市内に設立されました。 NTTの他の研究所がみな関東に集中する中、小さな研究所が ひとつ関西に飛び出し、いわば独立した瞬間でした。設立当初 から―貫して既存の研究領域にとらわれず、コミュニケーション を科学するとはどういうことか?にこだわり、情報科学から人間 科学、社会科学、人文科学をも巻き込んだ学際的な取り組み を進めてきました。特に音声音響処理、メディア処理、自然言 語処理、機械学習に代表されるいわゆるAI(人工知能)技術を 中心とした情報科学と人間の脳内における感覚運動処理や幼 児の言語獲得の仕組みを解明する人間科学の両方に取り組 むことで研究テーマ間の相乗効果を高めています。

深層学習に代表される、最近のAI技術の発展には目覚まし いものがあります。研究所としては、これらの最先端の技術を 高度に使いこなし、直面する課題にうまく適用することも重要 ですし、求められてもいます。しかし情報系の基礎研究を担う CS研としては、それ以上に、これまでの地の利やノウハウを活 かし更に技術を深め、いわば足場を固めつつ、これまでの延長 ではない、新たな次元を切り拓く取り組みに常に挑戦し、研究 テーマを大胆にシフトさせていくことが一層重要だと考えてい ます。

例えば音声認識の研究では、これまでの接話マイクに向 かって一人が話す音声認識から、テーブルを囲んで複数人が マイクから離れて自由に話す状況下での音声認識へとシフトし ています。そのためには音声認識そのものの性能向上も必要 ですが、それ以上に騒がしい環境でも周囲の雑音を取り除く雑 音抑圧技術や、室内での壁や床などへの反射による残響を除 去する、残響除去技術等の音声強調技術との組み合わせが 重要となります。CS研の音声認識は、これらの組み合わせによ り2015年に行われた国際技術評価(CHiME-3)で、参加25機 関中で第1位の精度を実現しました[1]。さらには、複数の発話 の中から聞きたい人の声を選んで聞き取る技術等にも取り組 んでいます。

機械学習技術に関して、CS研は、大量のデータから専門家 でも気づかない特徴的なパターンを自動的に見つけ出す技術を 研究開発してきました。現在はさらにこれを時空間に拡張した、 時空間多次元集合データ解析技術へとシフトさせています[2]。 さらに、単なる未来の予測だけでなく、予測の結果どうすべきか (例えばどう人々の流れを誘導すべきか)の最適な方策をリア ルタイムに提案してフィードバックする学習型シミュレーション 技術にも力を入れています。また、無数の組み合わせの中から 与えられた条件を満たす最適なものを見つけ出す、組合せ最 適化問題についても、BDDやZDDと呼ばれる、組み合わせ爆



発を回避するためのデータ構造を工夫することでこれまで不可 能だった大規模な問題が解けるようになっています。

ロボットとの対話処理の研究においては、これまでの一台の ロボットとの対話から、複数台(2台)のロボットとの対話処理へと シフトしています。一人の人間と対話する場合、ロボットは一台 で十分に思えますが、複数台のロボットを用いてうまく役割分 担することで、一台の場合より、たとえ音声認識に失敗したり、 対話の文脈が破綻したりしたとしても、人間にとって自然な対 話をより長く続けることができます。また、単なる雑談ではなく、 質問応答と雑談の融合により、対話への興味をユーザから引 き出しながらユーザに知識を伝える対話システムの実現にも取 り組んでいます[3]。

また、メディア処理においては、単一のメディアの処理から複 数のメディアを組み合わせたり、相補的に扱ったりするクロス モーダル情景分析へとシフトしています。この技術が完成すれ ば、例えばマイクで収録した音声情報から室内のビジュアルな 情景を再現することが可能になります。

人間科学研究にも変化があります。CS研では「視覚」「聴

覚」「運動感覚」といった人間の基本的な感覚に関する「潜在 的な脳の働き」の解明に注力しています。これらは今後も引き 続き取り組みますが、さらに新たなテーマとして、スポーツ脳科 学への取り組みが進んでいます[4]。脳科学の知見とICTを活 かし、「勝つための脳をきたえる」研究です。また、最近新たに、 人間の心の豊かさや満足に関する概念であるウェルビーイン グに関する研究が立ち上がりました。これらは触覚研究などと あわせて、人間の「心と身体」の両面に焦点を当てた新たな取 り組みです。

これら、新たな取り組みにはリスクが伴います。すぐに期待通 りの学術的な成果が得られるとは限りません。AI研究の分野で はExploration-Exploitation Dilemmaがしばしば問題となり ます[5]。これは限られたリソースの中で、うまくいくとわかってい る周辺をより詳細に探索するExploitationを行うべきか、目先 を変えてより広範囲に探索するExplorationを優先すべきか、 選択を迫られる、あるいは、バランスをとる必要がある、というも のです。基礎研究を担うCS研としては京阪奈も厚木もその独 立精神を胸に刻み、未踏の地を目指して大海原に船出する Explorationに、今後も大胆かつ粘り強く取り組んでいきたい と思います。

関連文献

- [1] 公共エリア雑音下でのモバイル音声認識の国際技術評価で、世界1位の精度を達成: http://www.ntt.co.jp/news2015/1512/151214a.html
- [2] 機械学習・データ科学センタ: http://www.kecl.ntt.co.jp/mlc/index-j.html
- [3] 世界初、雑談を交えながら知識を伝える新感覚対話AIを開発:http://www.ntt.co.jp/news2018/1801/180131b.html
- [4] スポーツ脳科学プロジェクト: http://sports-brain.ilab.ntt.co.jp/
- [5] 例えば: https://en.wikipedia.org/wiki/Reinforcement_learning



AIと倫理および社会的影響

Al, ethics and social impact



理化学研究所 革新知能統合研究センター(AIP) 社会における人工知能研究グループ グループディレクター

中川 裕志 Hiroshi Nakagawa

プロフィール

1975年 東京大学 工学部 卒業

1980年 東京大学 工学系研究科博士課程 修了(工学博士)

1980年 横浜国立大学 工学部 講師

1981年 横浜国立大学 助教授

1994年 横浜国立大学 教授

1999年 東京大学 情報基盤センター 教授

2001年 東京大学 大学院 学際情報学府 兼坦

2003年 東京大学 情報理工学系研究科 数理情報学専攻 兼坦

2016年 理化学研究所 革新知能統合研究センター(AIP) プライバシーと社会制度チーム チームリーダ(非常勤)

2017年 理化学研究所 革新知能統合研究センター(AIP) 社会における人工知能研究グループ グループディレクター

専門分野は人工知能、自然言語処理、情報検索、機械学習、プライバシー保護。

著書に『プライバシー保護入門:法制度と数理的基礎』(勁草書房、2016)、『東京大学工学教程情報工学「機械学習」』 (丸善出版、2015)など。

MEMO



基本演算を操る量子コンピュータの真価

~ゲート型量子コンピュータの計算能力の分析~

The real worth of quantum computers with elementary operations

Analysis of computational power of gate-based quantum computers



メディア情報研究部

高橋 康博 Yasuhiro Takahashi

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 主任研究 員。1998年東北大学理学部数学科卒業。2000年同大学院理学研究科 数学専攻博士前期課程修了。同年日本電信電話株式会社入社。2008 年電気通信大学大学院電気通信学研究科情報通信工学専攻博士後期 課程修了。博士(工学)。量子計算理論の研究に従事。情報処理学会、電 子情報通信学会各会員。

現在のコンピュータでは実行不可能な超高速計算が、量子 コンピュータでは可能になると期待されています。量子コン ピュータと現在のコンピュータの大きな違いの一つは情報の表 現にあります。現在のコンピュータは、0または1のどちらか一方 の値をもつ「ビット」により情報を表現します。一方で、量子コン ピュータにおいて情報を表現する「量子ビット」は、0または1だけ でなく、例えば0が30%で1が70%というような0と1の重ね合わ せを表現できます(図1)。量子ビットの個数が増えれば、より多 くの状態の重ね合わせが表現できます。この重ね合わされた状 態を用いた並列処理が量子コンピュータの特徴の一つです。

量子ビットを利用した情報の表現については共通ですが、 様々な方式の量子コンピュータが提案されています。これまで 世界で最も研究されてきた量子コンピュータをゲート型と呼びま す。現在のコンピュータと同様に、解きたい問題に対し、基本演 算(ゲート)の組み合わせで表現されるアルゴリズムを人間が構 築し、これを実行します(図2)。言わば、基本演算を積極的に操 る量子コンピュータです。一方で、近年活発に議論されているア ニーリング型は、解きたい問題に依存して量子状態の遷移のお よその方向性を指定するだけであり、その後は量子状態の自然 な状態遷移に任せることで、最終的に得られる状態が問題の解 になることを期待します。

我々の目標は、計算能力について、現在のコンピュータに対 する量子コンピュータの優位性を理論的に証明し、そのような 優位性を示す量子コンピュータを実現することです。ゲート型に ついては計算能力の理論的な分析が可能であり、現在のコン ピュータより計算能力が高いという証拠が示されているため、 我々はゲート型に焦点を当てています。しかし、ゲート型は他の 方式と比較して、実現に要求される技術水準が高いという問題 があります。本講演では、ゲート型量子コンピュータによる超高 速計算の実現へのハードルを大きく下げる二つの成果を紹介 します。



ゲート型の実現を難しくしている原因の一つは、複雑な重ね 合わせ状態を利用することです。このような状態は外界からの 影響を受けやすく、現在の技術では短時間で状態が崩壊してし まいます。この問題に対処する一つの方法は、アルゴリズムの ステップ数を制限し、量子状態の崩壊前にアルゴリズムの実行 を完了させることです。しかし、このような制限はゲート型量子コ ンピュータの計算能力を著しく低下させる可能性があるため、ス テップ数を制限したゲート型量子コンピュータの計算能力の分 析は重要な研究テーマになっています。

我々は論理和関数(入力ビット中に1があれば1を、そうでない 場合は0を出力する関数)の計算に着目しました。これは論理和 関数が少ないステップで計算できれば、様々な量子アルゴリズ ムが少ないステップで実行できるからです。この関数を計算す る従来方法では、入力ビット数が増えるにつれてステップ数も 増えます。一方で、我々は、入力ビット数に依存せず一定のス テップ数でこの関数が計算できることを示しました[1]。この成果 は10年来の未解決問題の解決であり、暗号の安全性の基礎と なる離散対数問題を従来より少ないステップで解くことに応用 できます。ステップ数を制限した状況においても、ゲート型量子 コンピュータによる超高速計算が可能であることを示した成果 です。

ゲート型の実現を難しくしているもう一つの原因は、0に初期 化された多数の量子ビットを必要とすることです。現在の技術 では、少数の初期化量子ビットしか用意できず、このような状況 では意味のある問題はほとんど解けません。我々は、少数の初 期化量子ビットに加え、現実的に用意しやすい未初期化量子 ビット(どんな初期状態でも良い量子ビット)を使うことで計算能 力が大幅に向上することを示しました[2]。初期化量子ビット数 を制限した状況においても、ゲート型量子コンピュータの高い 計算能力を引き出すことに繋がる成果です。

利用できる基本演算を現実的なものに制限する等、ゲート型 量子コンピュータに関する検討課題は残っています。このような 課題を解決し、ハードウェア研究と融合することで、ゲート型量 子コンピュータによる超高速計算が実現されると考えています。 また、ゲート型と他の方式との融合が、さらなる高速化を生み出 すと期待しています。我々は今後もゲート型量子コンピュータの 真価を明らかにする研究を続けていきます。

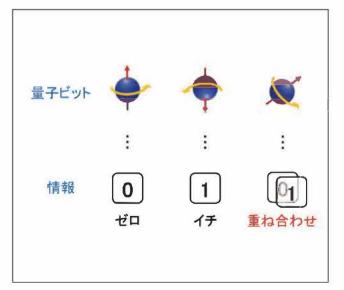


図1:電子スピン(ある軸についての回転)の向きを 利用した量子ビット

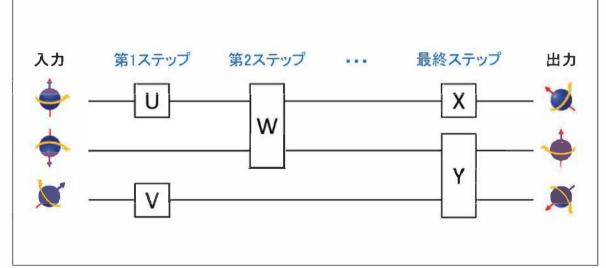


図2:ゲート型量子コンピュータにおける計算の模式図。入力量子状態に対し、 基本演算(ゲート)を適用して状態を遷移させ、出力量子状態を得る

関連文献

- [1] Y. Takahashi, S. Tani, "Collapse of the hierarchy of constant-depth exact quantum circuits," Computational Complexity, Vol. 25, Issue 4, pp. 849-881, 2016.
- [2] Y. Takahashi, S. Tani, "Power of uninitialized qubits in shallow quantum circuits," in Proc. 35th International Symposium on Theoretical Aspects of Computer Science (STACS), pp. 57:1-57:13, 2018.



ウェルビーイングにおける触覚の役割

~触れることの科学とデザインが人の心を豊かにする~

Role of haptics in improving wellbeing

Science and design of touch can enhance human flourishing



人間情報研究部

渡邊 淳司 Junji Watanabe

プロフィール

NTT コミュニケーション科学基礎研究所 人間情報研究部 主任研究員 (特別研究員)。人間の知覚特性を利用したインタフェース技術を開発・展 示公開するなかで、人間の感覚と環境との関係性を理論と応用の両面か ら研究している。平成20年度 文化庁メディア芸術祭 アート部門 優秀賞 受賞。平成24年Ars Electronica Prix 審査員。著書『情報を生み出す触 覚の知性』(化学同人、平成27年毎日出版文化賞受賞)。監訳『ウェル ビーイングの設計論』(BNN新社)。

携帯型情報端末や無線通信環境の普及をはじめ、情報通 信技術(ICT)は著しいスピードで発展してきました。しかし、その 一方で、ストレスや依存など心身への負の影響も指摘されてい ます。この背後には、ICTのコモディティ化(一般化)だけでなく、 技術の根源的な目標である「人間をしあわせにする」という事に 関する設計指針や評価原理が確立されていないという問題が あります。このような背景のなかで、近年、ウェルビーイング (Wellbeing、以下WB)という、人間の心の豊かさや満足に関 する概念が世界的に注目され、効率性・生産性とは異なる視点 からのICTの設計論が期待されています。例えば、2015年に国 連が採択した「2030年までの持続可能な開発目標(SDGs) | においても、WBは達成目標の一つとして挙げられ、また、IT企 業での瞑想の実践やウェブ・雑誌での言及もWBへの関心が 一般に広まりつつあることを示しています。情報系の研究分野 でも"Positive Computing"(2014)[1]をはじめ、ICTと心につ いての研究が増加しつつあります。

そもそもWBとは何か。それは主に三つの視点から定義されて います。一つ目は医学的WBで、心身の機能が不全でないかを 問う医学の定義といえます。二つ目は快楽的WBで、その瞬間 の気分の良し悪しに関する定義です。三つ目は持続的WBで、 人間が心身の潜在能力を発揮し、意義を感じ、周囲の人との関 係の中でいきいきと活動している状態を指します。英語ではフ ローリシング(flourishing)=開花という言葉で表現されます。 特に、持続的WBは、「景気」や「天気」といった概念がそうであ るように、構成概念(メカニズムを説明するために人為的に構成 された概念)であり、直接観察したり測定したりすることはできま せん。しかし、その構成概念の存在を仮定した場合、影響を受 けるであろう行動や心的状態といった要因を規定し、計測する ことで把握することができます。例えば、天気は温度、湿度、気 圧、風速など定量評価できる要因を設けることで「良い天気」や 「悪い天気」を示すことができます。同様にWB もその構成要因 を特定し、具体的に把握することができれば、向上のための設 計指針も明らかになるでしょう。



これまで、欧米で主潮となっている「主観的幸福」に着目した WBの研究では、例えば、図1にあるような構成要因がWBの実 現には重要であるといわれています[2]。しかし、その多くは個人 内の要因であり、人間同士の関係性やプロセスに高い価値を 置く日本など集産的文化においては、人間関係の要因の割合 がより多くなるかもしれません。さらに、WB向上に向けた情報提 示や介入では、触覚・身体感覚によるインタラクションが必要に なってくるでしょう。現在、NTT コミュニケーション科学基礎研究 所では、WBの設計論に資するための人間科学の観点からのメ カニズムの解明と、具体的なケースに適用可能な「WBの構成 要因の特定、計測、モデル化、介入、評価」の循環プロセスの 検討に取り組んでいます。具体的には、社会心理や生理計測 に基づくWBの指標づくり、母子間や対人コミュニケーション、 集団における身体的・心的な結びつきのメカニズムの解明、さら には、触覚を通した介入といった研究が行われています。

特に私自身は、触覚・身体感覚による情報提示[3]の研究を 行っていたこともあり、WB実現における触覚・身体感覚の役割 について検討しています。触覚・身体感覚を通して働きかけるこ とは、意識的な思考の様式だけでなく、情動や意識下の行動を 大きく変化させるということが示されています。つまり、触覚・身 体感覚はWBと関連した感情へ直接的に働きかけ、さらには、 WBと関連の強い行動を引き起こすきっかけとして機能するとい うことです。また、触覚・身体感覚に働きかける刺激は人間と人 間の関係を変容させます。日常生活でも温泉や祭などの身体 的共同体験は、他者を身近にするなど、適切な距離を設定する よいきっかけとなります。

これまでの触覚・身体感覚の研究は、主に一対一の関係の 中での情報伝達それ自体を目的としてきました。しかし、それが 現在は産業の対象となり、さらには、ここ数年の計測・伝送・提 示技術の発展により、ネットワークを介して多人数に触覚・身体 感覚情報を送ることも可能になってきています(図2)。そうであ るならば、現在、取り組むべきは、触覚・身体感覚の提示だけで なく、持続的WB実現の鍵となる触覚・身体感覚を通じて、どの ように人と人の間で生じるWBを向上させ、社会に貢献するかで あるといえます。

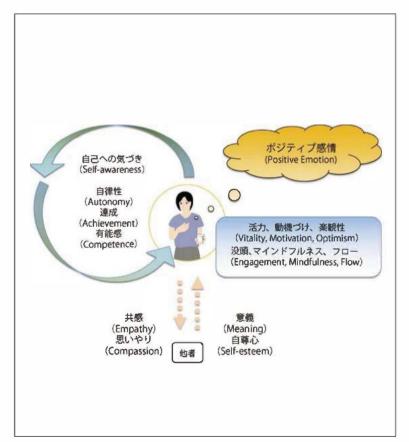


図1:ウェルビーイングを構成する主な要因



図2:触覚・身体感覚の情報を伝送することによる応用例

関連文献

- [1] R, Calvo, D. Peters, "Positive computing," MIT Press, 2014.(邦書 監訳:渡邊淳司, ドミニク・チェン, 翻訳:木村千里, 北川智利, 河邉隆寛, 横坂拓巳, 藤野正寛, 村田藍子, "ウェルビーイングの設計論," BNN新社, 2017.)
- [2] 渡邊淳司、村田藍子、安藤英由樹、"持続的ウェルビーイングを実現する心理要因、"日本バーチャルリアリティ学会誌, Vol. 22, No. 1, pp. 12-18, 2018.
- [3] 渡邊淳司, 情報を生み出す触覚の知性 化学同人, 2014.



「組合せ爆発」を乗り超える

~二分決定グラフを用いた膨大な量の組合せの数え上げと最適化~

Beyond combinatorial explosion

Enumeration and optimization with Binary Decision Diagrams



協創情報研究部

西野 正彬 Masaaki Nishino

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 研究主任。 2008年京都大学大学院情報学研究科修士課程修了。同年日本電信電 話株式会社入社。博士(情報学)。アルゴリズム、自然言語処理、組合せ 最適化の研究に従事。2017年度山下記念研究賞受賞。言語処理学会、 人工知能学会、情報処理学会、各会員。

10人の従業員を5人ずつ2つのグループに分けたいとしま す。このときに可能なグループ分けの種類は何通りあるでしょう か。高校数学で習う組合せの公式を適用すると、可能なグルー プ分けの数は252通りあることがわかります。では、従業員の数 が50人、100人と増えていくと、可能なグループ分けの種類はど うなるでしょう。同様に組合せの公式を適用してみると、50人の ときには約126兆通り、100人のときには10穣(1穣は100兆の 100兆倍)通りものグループ分けが存在することがわかります。 このように問題のサイズが大きくなるにつれて答えとなる組合せ の数が爆発的に増えていく現象は「組合せ爆発」と呼ばれま す。グループ分けの例のほかにも、ある駅から出発して別の駅に 到達するための鉄道の乗り継ぎの方法や、あるいは将棋の局 面の数など、組合せ爆発は様々な場面で見られる普遍的な現 象です。

計算機を使って問題を解く際には組合せ爆発にうまく対処す る必要があります。いま、可能なグループ分けの中から、業務が 滞りなく進むような良いグループ分けを探したいとします。この 問題を解くための最も素朴な方法として、全ての可能なグルー プ分けを順番に調べ上げるという方法が考えられます。しかしな がら、従業員数が多くなるとそのような解法は現代の高性能な 計算機をもってしても困難です。最新のパソコンは1秒間の間 に1兆回近くの計算を行うことができるといわれています。仮に1 つのグループ分けを1回の計算で見つけられたとしても,従業員 100人の可能なグループ分けを全て調べるのには1億年以上か かることになるため、組合せ爆発を起こす対象を全て数え上げ るのは現実的ではありません。そこで、これまでは組合せ爆発を 全て数え上げずに、いわば組合せ爆発を「避けて」問題を解く 方法が広く検討されてきました。



一方で、近年のアルゴリズム研究の発展の成果として、これ までは困難だと考えられていた組合せ爆発の数え上げが可能 なケースも数多くあることがわかってきました。組合せ爆発を正 確に数えることができると、可能な組合せの中からもっともよい ものを選び出すことや、問題の性質を詳しく分析することなどが できるようになります。私達は、二分決定グラフ(Binary Decision Diagram: BDD)とよばれるデータ構造を用いること で膨大な数の組合せを数え上げて問題を解く方法を研究して います。ここまで,膨大な数の組合せといってきたものは,数学的 には何らかの要素からなる組合せの集まり(集合)として表現で きます.例えば従業員の可能なグループ分けは、各従業員を要 素とする組合せの集合として表現できます.BDDは組合せの集 合を、図1に示すような丸や四角と矢印との組合せ(グラフ)とし てあらわす表現方法です。組合せ爆発に含まれる一つ一つの 組合せは似通っているものが多いことが知られています。従業 員のグループ分けの例では、ほとんど同じグループ分けがいく つもあることは容易に想像がつくでしょう。BDDはこうした似 通った組合せをまとめあげて表現することで、膨大な数の組合 せの集合を表現することができます(図1)。講演では、ポリオミ

ノとよばれる、四角い盤面にすき間なくピースを敷き詰めるパズ ルの全ての解を高速に見つけるための方法や[2]、通信ネット ワークの故障確率を正確に計算することで、故障する確率が最 小となる通信ネットワークの設計方法 [3] などを紹介します。い ずれの応用課題においても、BDDを用いて組合せ爆発を起こ す対象を圧縮して表現することで、これまでよりもサイズの大き な問題を、より正確に解くことに成功しています(図2)。

もちろん、BDDを用いて世の中の組合せ爆発を起こす対象 をすべて数え上げられるわけではありません。しかしながら、近年 の計算機の性能の向上と、BDDに代表される最新のアルゴリ ズムとを組み合わせると、さまざまな場面で正確な数え上げがで きるようになりつつあります。組合せ爆発を避けるのでなく、真 面目に数え上げて問題を解く、いわば組合せ爆発を「乗り超え る」アプローチというのはこれまであまり取られてこなかった新し いものであり、多大な可能性を秘めていると考えています.今後 もBDDの数え上げの威力を活かしたアルゴリズムの研究に取 り組むことを通じて世の中の問題の解決に取り組んでいく予定 です。

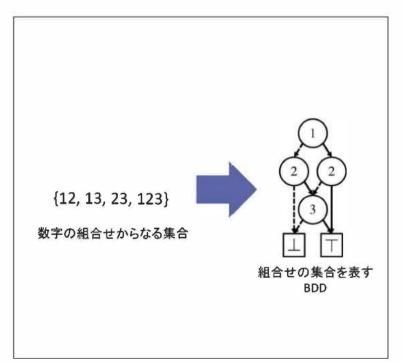


図1:組合せの集合と、それを表現する 二分決定グラフ(BDD)の例

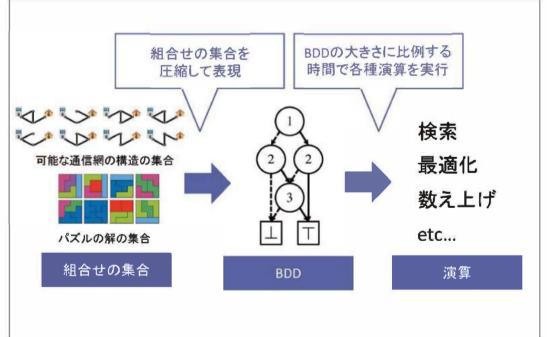


図2:二分決定グラフを用いて組合せ爆発を扱う処理の流れ

関連文献

- [1] M. Nishino, N. Yasuda, S. Minato, M. Nagata, "Dancing with decision diagrams: a combined approach to exact cover," in Proc. the 31st AAAI Conference on Artificial Intelligence (AAAI), 2017.
- [2] M. Nishino, T. Inoue, N. Yasuda, S. Minato, M. Nagata, "Optimizing network reliability via best-first search over decision diagrams," in Proc. IEEE International Conference on Computer Communications (INFOCOM), 2018.



脳からみた聞くと話すの共通性

~音声変換技術と脳機能計測による人間の音声コミュニケーションの仕組みの解明~

Speech production and perception share common brain pathways

Investigation of the mechanisms of speech communication by speech conversion and brain imaging



人間情報研究部

廣谷 定男 Sadao Hiroya

プロフィール

NTT コミュニケーション科学基礎研究所 人間情報研究部 主任研究員 (特別研究員)。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。音声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。IEEE、日本音響学会、電子情報通信学会、北米神 経科学学会、言語神経生物学学会各会員。

聞くことと話すことによる音声コミュニケーションは、日常生活 において欠かすことのできない重要なものです。「話す」過程で は、脳でイメージした伝えたい意図が調音器官を通じて音声とし て放射され、「聞く」過程では、音声が聴覚器官を経て脳で理解 されます。聞くと話すの目的はそれぞれ異なることから、これまで ほとんどの音声研究は聞くと話すを別々に扱ってきました。

人間の音声脳科学研究においても、「聞く」に関与する脳部 位はウェルニッケ野、「話す」はブローカ野というように別々に議 論が行われてきました。ところが、近年の脳機能計測の発展に 伴い、聞くと話すに共通した仕組みが脳の中に存在することが 分かってきました。この共通した仕組みは、話すときに「聞く」に 関与する脳部位の活動、聞くときに「話す」に関与する脳部位 の活動を調べることで明らかになりました。つまり、聞くと話すを 別々に扱うのではなく、一緒に扱うことが大切なのです[1]。

話すときに「聞く」に関与する脳部位が活動する

自らが発話した音声を200ミリ秒程度遅らせ、それをヘッドホ ンで聞きながら発話を行うと、どもるようになる、間延びするなど 発話が困難になる現象が起こります。また、音声のタイミングを 変換するのではなく、「え」と発話した音声が「あ」と聞こえるよう に音声の共振周波数をリアルタイムに変換した音声を聞きなが ら発話を行っても発話に変化が見られます。聞こえた音声に応 じて発話が変化するということは、人間は自らの音声をリアルタ イムにモニタリングしながら話しをしているということです。これら 変換音声を聞きながら発話を行うときの脳活動を調べますと、 聞くに関与する脳部位(上側頭回)の活動が増加することが分 かりました。つまり、モニタリングで発話に誤りが見つかったとき に、脳は発話を正しく修正しようとするのです。この誤りに対する 発話の修正量の大きさについて様々な検討が行われています が、最近私たちは、変換音声の音質が良くなると、修正量が大き くなることを発見しました[2]。



聞くときに「話す」に関与する脳部位が活動する

話すときに「聞く」に関与する脳部位が活動することはモニタ リングを考えれば理解しやすいかと思います。しかし反対に、聞 くときに「話す」に関与する脳部位が活動するということは起こり うるのでしょうか。もしそうであれば、音声を聞くと、つられて自ら の口が動いてしまうことになります。ところが、口が動かないまで も、話すに関与する脳部位(運動前野)が音声を聞くときに活動 することが明らかになっています。このことは、調音器官の運動 をシミュレーションしながら音声を聞いているという可能性を示 しています。人間がなぜこのようなことをしながら音声を聞いて いるのかについてはまだ分かっていませんが、音声が聞き取りに くい場合に運動前野の活動が大きくなることから、音声に足りな い情報を補うために調音運動のシミュレーションを行うことで音 声理解(what)を助ける、また音声の自然性(how:合成音声ら しさ、たどたどしさ)を判断するためでないかと考えられています。

音声リズムを変換する

さて、自然性を変換した音声を聞くときに、話すに関与する脳 部位の活動がどのように変化するかは分かっていません。私た ちは、日本語母語話者の英語音声のリズムを変換する技術[3] により作成した刺激音声を英語母語話者が聞くときの脳活動 計測を行いました(図1)。実験の結果、英語母語話者にとって 自然でないリズムの音声を聞くときに、発話に関与する補足運

動野の活動が増加することが分かりました[4]。つまり、聞くにお ける発話のシミュレーションは音声の自然性を判断するために 行われている可能性があります。

まとめと今後の展望

近年、音声変換技術の発展に伴い、音声変換技術と脳機能 計測を組み合わせることで、これまで知られていなかった音声脳 情報処理の仕組みが解明されつつあります。音声変換技術は、 今後も音声脳科学研究において有効な手段であると考えられ ます。また、この技術を応用して、ネイティブに聞き取りやすい英 語音声に変換することで、日本語母語話者の英語での音声コ ミュニケーションを支援できると期待されます。

話すときのモニタリングに関わる脳情報処理および聞くとき のシミュレーションに関わる脳情報処理を紹介してきましたが、 実は、モニタリングもシミュレーションも、聴覚野から上側頭回と 角回、縁上回を経由して運動前野に音声が送られるという共通 した経路です(図2)。しかも、これらは人間の聞くにおける頑健 性や話すにおける安定性に深く関わっていると考えられていま す。しかし、現在の音声認識および音声合成技術ではこの経路 に相当する処理が考慮されていません。もしかすると、これら経 路における音声脳情報処理の理解が今後音声認識の性能や 音声合成の品質を向上させる鍵となるかもしれません。

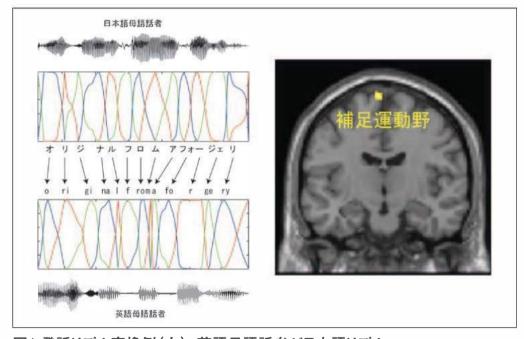


図1:発話リズム変換例(左)。英語母語話者が日本語リズム 英語音声を聞いているときの脳活動(右)

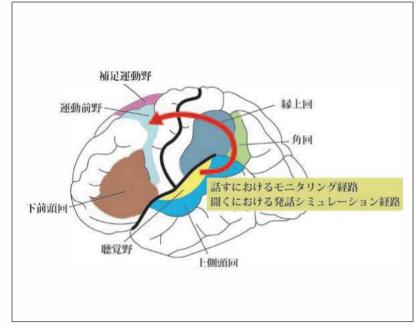


図2:「話す」におけるモニタリング経路と「聞く」における発話 シミュレーション経路。「聞く」と「話す」によらず共通した経路

関連文献

- [1] 廣谷定男編著, 筧一彦, 辰巳格, 皆川泰代, 持田岳美, 渡辺眞澄著, "聞くと話すの脳科学," コロナ社, 2017.
- [2] S. Hiroya, T. Mochida, "Speech sound naturalness alters compensation in response to transformed auditory feedback," in Proc. J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, p. 3228, 2016.
- [3] S. Hiroya, "Non-negative temporal decomposition of speech parameters by multiplicative update rules," IEEE Trans. Audio, Speech and Lang. Process., Vol. 21, No. 10, pp. 2108-2117, 2013.
- [4] S. Hiroya, K. Jasmin, S. Evans, S. Krishnan, M. Ostarek, D. Boebinger, S.K. Scott, "Effects of speaking rhythm naturalness on the neural basis of speech perception," in Proc. Society for Neuroscience, 2015.



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、あるいはコンピュータと人の間の「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、情報科学と人間科学の両側面から基礎研 究に取り組んでいます。オープンハウス2018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展示をご紹介いたしま す。

データと学習の科学

データ解析と機械学習技術の研究をご紹介します

- ■膨大なデータから価値ある知見を見つけたい
- ■過去の事例に学び、未来を予測したい
- ■データ解析の応用事例を究めたい



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を ご紹介します

- ■AIとの自由な会話を実現したい
- ■言語の本質を深く理解したい
- ■膨大な組み合わせから適した解を見つけたい



メディアの科学

音声・画像などのメディア情報を高度・高速に処理する研究を ご紹介します

- ■最新のAIを駆使し、メディア処理を高度化したい
- ■メディアの本質を知り、望みどおりに変換したい
- ■人間の知覚特性をメディア表現に活かしたい



人間の科学

人間の情報処理メカニズムを科学的に解明する研究を ご紹介します

- ■心豊かなライフスタイルをデザインしたい
- ■トップアスリートの脳のメカニズムを解明したい
- ■脳の仕組みを理解し、新しいICTを提案したい



膨大なデータから似た音声を見つけます!

~グラフ索引に基づく類似音声探索~



どんな研究

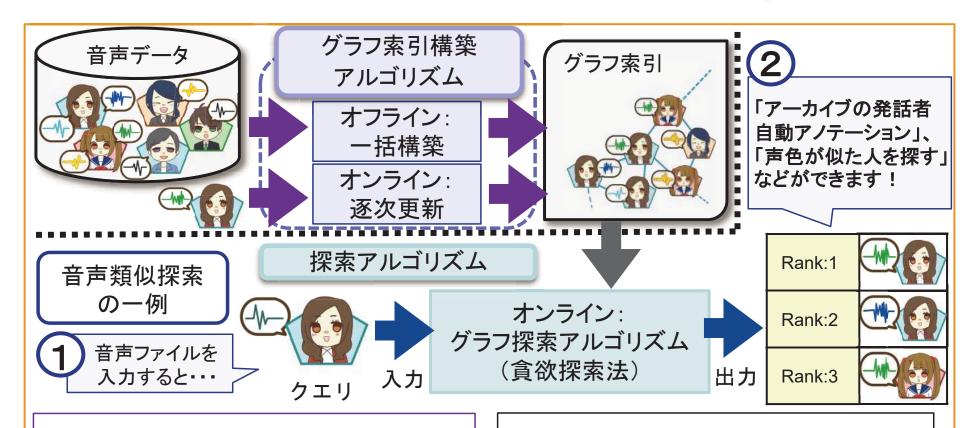
膨大なデータの中から即座に類似の データを見つけるための研究です。 例えば、多数の話者がいろいろな発 話をしている膨大な音声データの中 から、ある特定の話者の発話した音 声データや、その話者に類似の声色 の別の話者を探すことができます。

どこが凄い

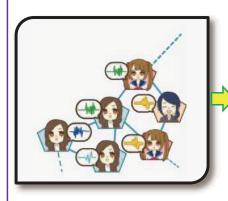
提案法は、検索対象間の類似度さえ 定義できれば構築可能なグラフ索引 に基づく、多様なメディアに利用で きる汎用性に優れた方法です。グラ フ索引は、検索の度に随時更新でき、 新しい音声を即座に追加することで 最新のデータで検索できます。

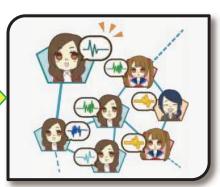
めざす未来

あらゆるメディアを検索可能なプ ラットフォームを構築し、なんでも 気軽に検索できる未来をめざします。 さらには、複数のメディアを統合的 に検索可能なクロスメディア検索シ ステムを実現し、文書、画像、音声 など何を入力しても役立つ情報が得 られる将来をめざします。



グラフ索引の逐次更新





入力した音声もグラフ索引に即座に追加可能!

提案法の特徴

- グラフ索引の持つスモールワールド性により、 高速性を実現
- グラフ索引は、検索対象の関係性のみで構 成できるため、様々な対象に適用可能
- グラフ索引は、一括更新と逐次更新により常 に最新の状態で運用可能

関連文献

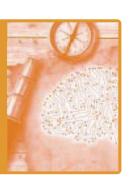
- [1] K. Aoyama, A. Ogawa, T. Hattori, T. Hori, "Double-layer neighborhood graph based similarity search for fast query-by-example spoken term detection," in Proc. Int. Conf. Acoustics, Speech, Signal Process. IEEE, April 2015, pp. 5216-5220.
- [2] K. Aoyama, S. Watanabe, H. Sawada, Y. Minami, N. Ueda, K. Saito, "Fast similarity search on a large speech data set with neighborhood graph indexing," in Proc. Int. Conf. Acoustics, Speech, Signal Process. IEEE, March 2010, pp. 5358-5361.

担当者

服部 正嗣 (Takashi Hattori)、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

複数の問題に共通して重要な情報の組合せを発見

~共通因子を効率的に学習する低ランク回帰技術: MOFM~



どんな研究

今まで一つの目的変数の予測に限ら れていた「Factorization

Machines」を複数の目的変数の予測 に拡張した研究です。提案法はデー タを表す特徴から各目的変数に対し て有用な組合せを見つけて高精度な モデルを作ります。

どこが凄い

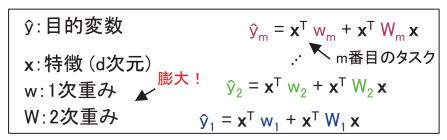
提案法は、各目的変数に共通の基底 を用いて特徴の組合せを表現し、重 みを求めるため、次元数が膨大に なっても効率的に学習できます。ま た、初期値に依存せず大域的最適解 を得られることが理論的に保証され ています。

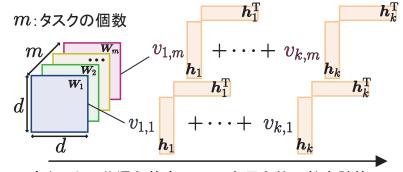
めざす未来

応用先として、医療診断、推薦シス テム、遺伝子解析などに利用可能で す。今後は、理論的性質(表現力、 汎化能力など)に関する理解を深め、 提案法を再帰的に適用することで深 層学習での効率的な計算に活用して いきます。

Multi-Output Factorization Machines (MOFM): 複数タスクのモデルを同時に 学習し重要な特徴の組合せを発見

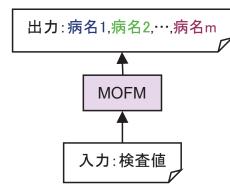
- 汎用的(様々な予測問題に適用可能)
- タスク間の相関関係を活用 (各重み行列を共通の基底で低ランクに分解)
- モデルのパラメータ学習が容易 (解が初期値に依存せず一意に求まる)
- 膨大な数の組合せを扱える (高次元データに適用可能)





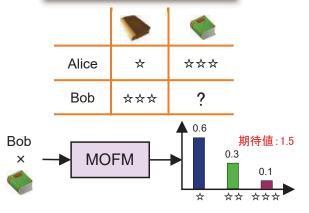
全タスクに共通な基底 ħ,...,ħ,を逐次的に効率計算

応用例1: 医療診断



複数の病名(1… m)を診断し共通 な要因(検査値の組合せ)を発見

応用例2: 商品推薦



評価点(1… m)に対して確率を予測し

応用例3: 収穫量予測



各環境(1… m)に対する 遺伝子の組合せから予測

関連文献

- [1] M. Blondel, A. Fujino, N. Ueda, "Convex Factorization Machines," in Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2015.
- [2] M. Blondel, V. Niculae, T. Otsuka, N. Ueda, "Multi-output Polynomial Networks and Factorization Machines," in Proc. Neural Information Processing Systems, 2017.

担当者

ブロンデル マチュー (Mathieu Blondel) 上田特別研究室

人はどこから来て、どこへ行くのか?

~人流データ同化と学習型誘導 ~



どんな研究

多くの人が集まるスポーツイベント 等で人の流れをスムーズにすること をめざした研究です。そのためには 事前や当日に誘導計画を策定するこ とが重要な課題です。この展示では、 実世界の観測値を人流のシミュレー ションに反映して、誘導計画策定に 活用する技術を紹介します。

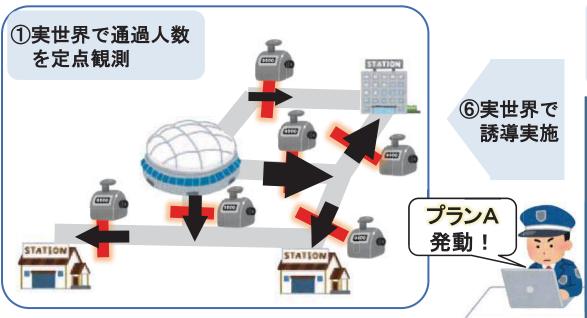
②観測データを活用

どこが凄い

実世界で定点観測した通過人数を基 に、経路ごとの人の流量を推定し、 正確に人流をシミュレーションでき ます。さらに膨大な組み合わせの誘 導策の中から、シミュレーション上 で効率的に最適誘導策を探索するこ とができます。昨年、実際のイベン トで誘導計画の策定に活用しました。

めざす未来

混雑予測や適切な誘導策の導出によ り、より安全で効率的にイベントを 開催することができるようになりま す。来場者は安心して快適にイベン トに参加できるようになります。自 動車の効率的なナビゲーションや、 店舗内での人流分析への応用展開も 可能です。



⑤不測の事故などの多様な状況を 考慮したシミュレーションに より最適誘導計画の策定を支援



人流を再現するシミュレータ用の 入力データを自動生成

③機械学習により 経路毎の人の 流量を推定

STATION

膨大な組み合わせの誘導策



④シミュレータへ 入力し、人流を再現

関連文献

STATION

- [1] 清水仁, 松林達史, 田中悠介, 岩田具治, 澤田宏, "滞在人数を考慮した経路別人数の推定," 第32回人工知能学会全国大会, 2018.
- [2] 清武寛, 幸島匡宏, 松林達史, 戸田浩之, "時間遅れを考慮した経路別通行人数の推定," 第32回人工知能学会全国大会, 2018.
- [3] N. Ueda, F. Naya, "Spatio-temporal multidimensional collective data analysis for providing comfortable living anytime and anywhere," APSIPA Transactions on Signal and Information Processing, Vol. 7, No. 4, 2018.

担当者

清水 仁 (Hitoshi Shimizu) 協創情報研究部 知能創発環境研究グループ

都市の今を知る

~環境センシングと異種データ融合分析によるイベント解析~



どんな研究

都市で発生したイベントを収集し、 分析する研究に取り組み、藤沢市で 6年に渡って実証実験を行ってきま した。ゴミ収集車などの公共車両に 搭載したセンサによる環境センシン **グと、そのセンサデータからのイベ** ント抽出、抽出したイベントの解析 に取り組んでいます。

どこが凄い

変化する実環境に応じてシステムを 変更できるプログラム動的更新技術 や、通信コストやサーバでの処理を 軽減するためのデータ縮約技術を実 現しました。さらに異種データの融 合分析を行い、職員の直感をデータ によって裏付け、今までにない観点 からの新たな気づきが得られます。

めざす未来

都市では、さまざまなイベントが起 きていますが、多くは情報化されず、 都市マネジメントや、市民生活に活 用されていませんでした。提案する 技術を用いてイベントを収集・解析 することにより、より効率的な都市 運用、詳細な予測に基づく都市計画 立案が可能となります。

目的:

スマートシティ実現に向けた 時空間イベント解析による



大気状態モニタリング

定化を実現

□ ゴミ収集車に大気センサを搭載

して街中の環境をセンシング

ロ プログラム動的更新技術により、

試行錯誤しながらシステムの安

騒音レベルマップの表示や 大気汚染物質の分布を元にした

綿密な大気汚染対策を可能に

地区別ゴミ量推定

- □ エッジセンサにおけるデータ縮約技術を実現
- エンジン振動をセンシングしてゴミ収集区間を抽出
- イベント抽出により、送信データ量を大幅削減



ゴミ収集イベント

ゴミ収集状況

地区別ゴミ量の長期変動や地域差を考慮した ゴミ収集計画立案を支援可能に

焼却場で計測した重量を 地区毎の収集時間で案分

ゴミと地域の特徴の異種データ分析

残渣ゴミが少ない地区の特徴

* 残渣ゴミ:捨て間違い、未回収のゴミ

残渣ゴミ 人口構成 不動産情報 E

正則化 線形回帰

- 特徴
- 1 分譲賃貸 2 バルコニー
- 3 インターネット対応
- 4 70歳代
- 5 ルームシェア不可
- □ 地域の特徴を入力、残渣ゴミ*を出力とする回帰により、 ゴミの出し方と地域の特徴の関係を異種データ融合分析
- □ 抽出された特徴は清掃員の直感と一致

地域の特徴を考慮したゴミ減量講習会などの対応が可能に

- 本研究成果の一部はNICT委託研究「ソーシャル・ビックデータ利活用・基盤技術の研究開発」によるものです。
- 地図データは、国土地理院 電子国土基本図を利用。街区の区切りは、政府統計の総合窓口 統計GIS 境界データを使用。

関連文献

- [1] Yasue Kishino, Koh Takeuchi, Yoshinari Shirai, Futoshi Naya, Naonori Ueda, "Datafying city: Detecting and accumulating spatio-temporal events by vehicle-mounted sensors," in Proc. of IEEE BigData, International Workshops on Smart Citied: People, Technology and Data (IWSC'17), 2017.
- [2] 竹内 孝, 岸野泰恵, 白井良成, 須山敬之, 納谷 太, 上田修功, "空間情報を用いた局所化 Lasso よる都市データ解析," 電子情報通信学会総合 *大会講演論文集*, pp.SS-103 - SS-104, 2018.

担当者

岸野 泰恵 (Yasue Kishino) 協創情報研究部 知能創発環境研究グループ

深層学習をモバイル向けに小さくします

~量子化による深層学習のモデル圧縮技術~



どんな研究

画像や音声などの認識に深層学習が 盛んに用いられています。しかし、 重みを保持するパラメータの数が多 いため、メモリを大量に消費します。 本研究ではパラメータの分布を考慮 した新たな量子化により、深層学習 のメモリ消費量を32分の1に低減 できる学習法を提案します。

どこが凄い

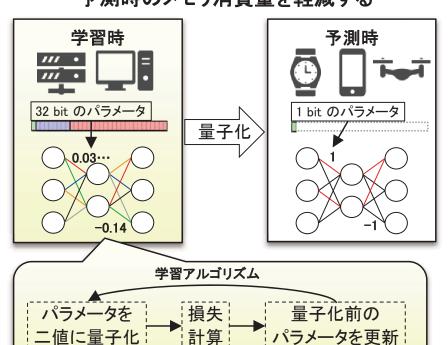
深層学習の連続値であるパラメータ を ± 1に限定する量子化の影響を最 小限にすると同時に効率的な学習ア ルゴリズムとすることで、従来手法 よりも高精度かつ高速な量子化が可 能となります。また、学習後には ビット演算で計算できるため、高速 に予測できるようになります。

めざす未来

本技術は要求性能の低い深層学習を 作ります。そのため、パラメータが 多いほど高精度な予測の行える深層 学習を性能の低い小型端末上で動か せるようになります。これにより、 これまで深層学習が使えなかった 様々な端末で画像や音声などの高度 な認識が行えるようになります。

背景

パラメータの値を ±1 に限定して 予測時のメモリ消費量を軽減する



問題点 認識性能の低下

-0.14 → -1

+1への変換は近似誤差が大きい 原因①

原因② 適切な更新量が求まらないため 学習がうまくいかない

更新量 = 学習率 × 関数の微分値

±1に変換する関数 +1 微分値 ○ → 更新量の消失

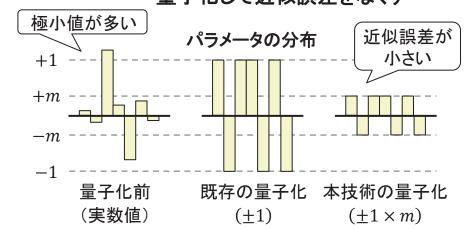
更新量の消失対策

-0.14 → -0.09

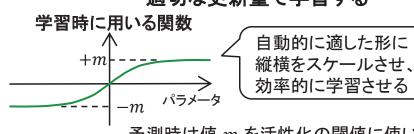
- 学習率を高くする ・微分値を強制的に
- 0から1にする
- ⇒ 学習が困難

提案手法

原因①の対策 パラメータをノルムの平均値 m に 量子化して近似誤差をなくす

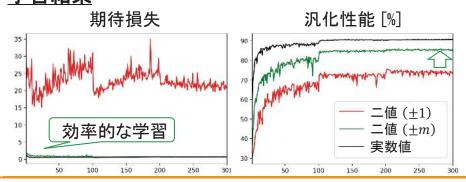


原因②の対策 近似した連続関数を用いて 適切な更新量で学習する



予測時は値 m を活性化の閾値に使い回す ことでパラメータを±1にして推論する

学習結果



関連文献

- [1] 大屋優, 井田安俊, 藤原靖宏, 岩村相哲, "正則化による深層学習の重み量子化手法の検討," 電子情報通信学会技術研究報告, Vol. 117, No. 211, pp. 51-52, 2017.
- [2] 大屋優, 井田安俊, 藤原靖宏, 岩村相哲, "バイナリニューラルネットにおける非活性化手法の検討," 電子情報通信学会技術研究報告, Vol. 117, No. 238, pp. 119-120, 2017.

担当者

大屋優(Yu Oya) ソフトウェアイノベーションセンタ 分散処理基盤技術プロジェクト 分散システムアーキテクチャ基盤グループ

光で機械学習をスピードアップ

~光リザーバーコンピューティングによる高速機械学習~



どんな研究

光演算で機械学習を加速します。莫 大な数の行列演算を光の干渉で処理 することで、並列かつ光速で演算が できます。この性質を最大限活かせ る構成として、小脳を模擬したリザ ーバーコンピューティングに着目し その光回路実装を進めています。

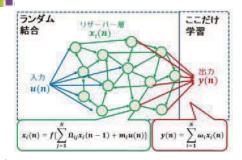
どこが凄い

一つの回路で入力信号が1秒あたり ギガビットを越える信号を処理する ことができます。さらに、光の位相、 波長、偏波を用いることで演算量を 数百倍にすることも期待できます。 これによって、大規模なニューロ演 算を高速に行うことが可能です。

めざす未来

機械学習のアクセラレータとしての 適用をめざします。超高速性を活か して、通信や高速な動画処理に機械 学習の情報処理技術を適用すること も検討しています。光の有する性質 を最大限に活用することで究極に速 い計算機の実現をめざします。

<u>-バーコンピュー</u>ティング 📭 リザ



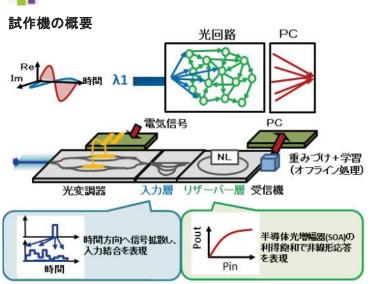
(概要)

(応用)

- 入力層と中間層をランダム結合で表現した ニューラルネットワーク(小脳的な構成)
- 学習変数が少ない上、線形回帰の ため高速で収束
- 時系列信号処理、音声認識、画像処理など

光波を用いたリザーバーコンピューティングの特徴 特長②: 特長①: 波長、位相、偏波を利用した 光速伝搬、超短パルスによる 光演算による超低電力 大規模並列化が可能 高速演算を実現 なニューロ演算が可能 特長④: ~10ピコ秒 光通信で培った技術による 大規模小型集積が可能 λ2 時間 光導波路 空間光学系 線形回帰 光変調器 入力層

🜉 試作機の実装と原理検証



ベンチマークタスク(NARMA10*1)での動作検証

X1 NARMA10: 10次の時間遅れ+2次の非線形関数の近似タスク (b) Imaginary valued output (a) Real valued output 2 RC output RC output Output (a.u) -2 -2 -3 -3 0 10 30 40 20 Time (nsec) Time (nsec)

最小ブロックである単一波長で動作する試作機を構築し、以下を検証

- ① 光演算で高速処理の可能性(本試作機では、入力層・リザーバー層までの光実装を検討)
 - ⇒ナノ秒(GHz)オーダのスループット(CPUに比べ3-4桁程度高速な処理を実現)
- ② 光の波動性を利用した複素空間での情報処理は可能性
 - ⇒ 実部·虚部の両方で所望の応答を確認し、位相空間への並列化を実現

波長・偏波・空間等を利用したネットワークの大規模化と更なる高スループット化 オンライン処理化と音声認識や画像処理、光通信の信号処理等をはじめとした応用問題への展開

関連文献

[1] M. Nakajima, M. Inubushi, T. Goh, T. Hashimoto, "Coherently Driven Ultrafast Complex-Valued Photonic Reservoir Computing," CLEO 2018,

担当者

中島 光雅(Mitsumasa Nakajima)

先端集積デバイス研究所 光電子融合研究部 光電子複合機能集積グループ

ネットワーク構造から深層学習のしくみを知る

ーユーラルネットの理解に向けたコミュニティ抽出技術~



どんな研究

深層ニューラルネット(DNN)は様々 な実データに対し高精度な予測を実 現していますが、その複雑な学習結 果を人間が理解することは困難です。 本研究では、ネットワーク解析によ りDNNの中に潜むコミュニティ構造 を自動的に発見し、可視化する方法 を提案します。

どこが凄い

データから学習されたDNNに対して コミュニティ抽出と呼ばれる手法を 適用することで、DNNの大局的な構 造を捉えることを世界で初めて可能 としました。これは、DNN内部の各 部分が予測や認識において果たす役 割を解明するための、新たな糸口と なりえます。

めざす未来

DNNの内部の仕組みを人間が解釈し やすい形で表現することで、自動運 転や医療など、機械学習技術が導き 出した予測結果に対して根拠を説明 できることが必要となる応用分野に おいても、安心してDNNが使えるよ うな未来をめざしています。

①入出力データの組から ニューラルネットを学習

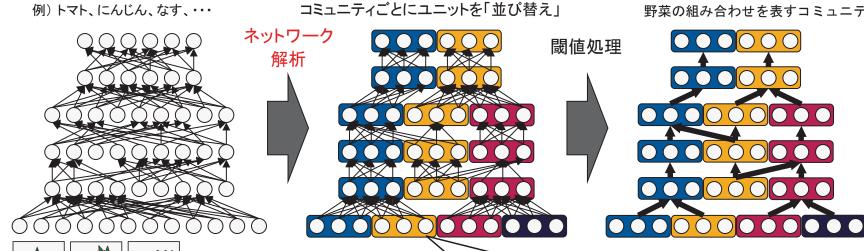
画像に写った野菜を識別する問題の場合

②隣接する層のユニットに対し、 似た結合パターンを持つユニット のコミュニティを推定

複数の結合をまとめて表示

例)類似した推論情報を受け取る 野菜の組み合わせを表すコミュニティ

③各コミュニティ間に存在する



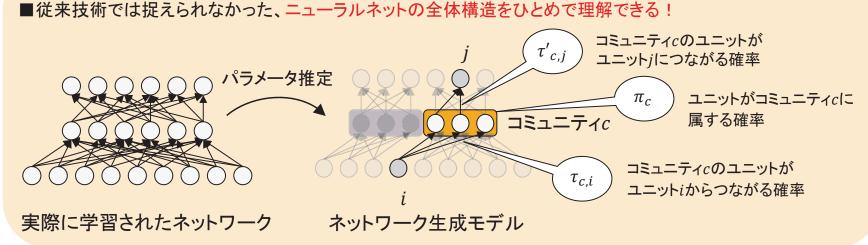
例)





推定されたコミュニティ

- 例)類似した推論情報を受け渡す入力 画素の組み合わせを表すコミュニティ
- ■ネットワーク構造の生成モデルに基づくコミュニティ抽出法を、学習済ニューラルネットに適用



関連文献

- [1] C. Watanabe, K. Hiramatsu, K. Kashino, "Modular representation of layered neural networks," Neural Networks, Vol. 97, pp. 62-73, 2018.
- [2] 渡邊千紘, 平松薫, 柏野邦夫, "深層ニューラルネットにおける正負の結合重みに基づく大局構造抽出, " *情報科学技術フォーラム*, 2017.
- [3] C. Watanabe, K. Hiramatsu, K. Kashino, "Recursive extraction of modular structure from layered neural networks using variational Bayes method," in Proc. Discovery Science, Lecture Notes in Computer Science, Vol. 10558, pp. 207-222, 2017.

担当者

渡邊 千紘(Chihiro Watanabe)メディア情報研究部 メディア認識研究グループ

08

あなたの量子メモリをちょっと拝借!

~未初期化量子ビットを利用した高速量子計算~



どんな研究

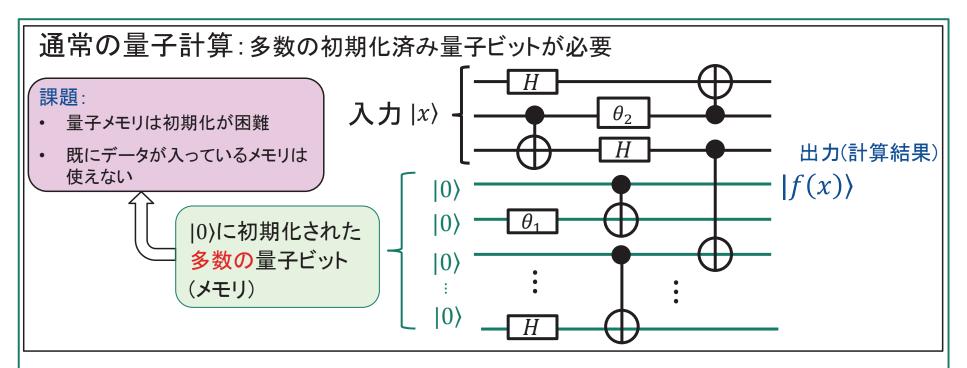
計算を行う前に、通常、メモリを初期化する必要があります。しかし、量子コンピュータの場合、この初期化が容易ではありません。本研究は、初期化を省略して、初期状態が不明である量子メモリを使っても、複雑な計算を間違うことなく実行することを可能にします。

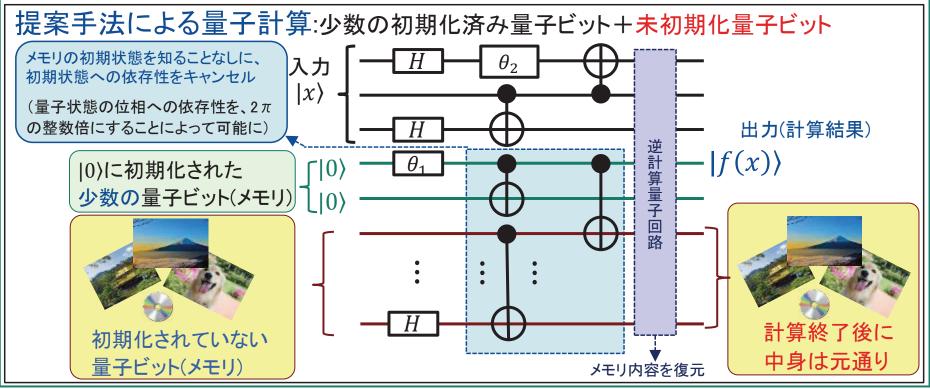
どこが凄い

通常の量子アルゴリズムでは、出力がメモリの初期状態に依存するため、正しく計算できません。本研究成果は、メモリの初期状態を知ることなしに、初期状態への依存性をキャンセルする技術です。計算終了後、量子メモリは元の状態に復帰するため、メモリ内データを再び利用できます。

めざす未来

量子メモリを複数の計算タスクで共 有することが可能になるため、実質 的に、より多くの量子メモリを使用 した量子計算が可能になります。こ の結果、量子アルゴリズムの並列性 が増し、量子コンピュータのさらな る高速化、適用領域の拡大や早期実 用に貢献します。





関連文献

- [1] Y. Takahashi, S. Tani, "Power of uninitialized qubits in shallow quantum circuits," in *Proc. 35th Symposium on Theoretical Aspects of Computer Science (STACS 2018)*, pp. 57:1–57:13, 2018.
- [2] Y. Takahashi, S. Tani, "Power of uninitialized qubits in shallow quantum circuits," arXiv:1608.07020v3.

担当者

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

09

こわれにくいネットワークをデザインします

~二分決定グラフを用いたネットワーク信頼性最大化~



どんな研究

電線や道路などを増強する場合に、 最も信頼性が高くなるようなネットワークの増強方法を自動的に見つけます。信頼性とは各電線や道路が寸断されたとしても通信や行き来ができる度合いを表します。信頼性の高い増強方法を見つけることは膨大な計算が必用な難しい問題でした。

どこが凄い

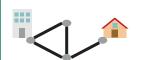
最適な増強方法を見つけるためには、 膨大な数の故障パターンを調べ上げ る必用があります。私達は二分決定 グラフとよばれるデータ構造を用い て故障パターン集合を小さく表現し で、既存方法よりも十倍した。 大きい実用的な通信ネットワーした。 最適解を求めることに成功しました。

めざす未来

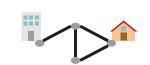
本技術を用いることで、既存の方法 よりも信頼性の高いネットワークを より安価に設計できるようになりま す。今後は広域の道路網などの大規 模なネットワークの設計や、自然の 書による断線が発生するケースへの 対応などを行い、より幅広が 活用できる技術の実現をめざします。

ネットワーク信頼性最大化問題

リンク故障時に通信できる確率(<mark>信頼性</mark>)が高くなるように 増強費用が予算内に収まる範囲でリンクを配置する問題







いちばん故障に強い通信ネットワークは?

問題の難しさ

信頼性が最大となるリンクの配置方法を計算するのに 膨大な計算時間がかかる

難しさ1. 信頼性を計算するために指数 (2^{リンク数}) 通り の故障パターンを数え上げなければならない

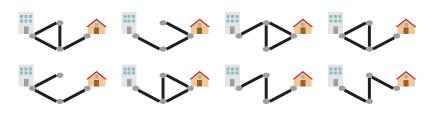




通信可能な故障

通信不能な故障

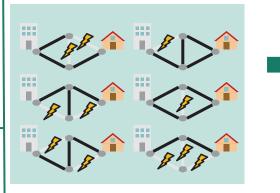
難しさ2.リンクの配置方法の候補が指数通り存在

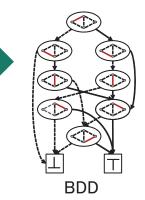


二分決定グラフを用いた解法

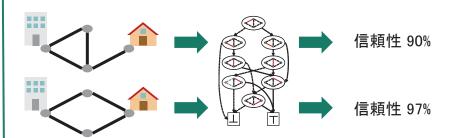
二分決定グラフ (BDD) を用いて故障パターンの共通 部分をまとめ上げて小さく表現することで計算を高速化

→ 約150頂点からなる大規模なネットワークにおいて 最適解を求めることに初めて成功

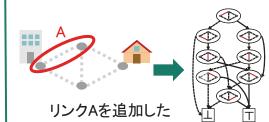




ポイント1. 高速に信頼性を計算



ポイント2. 信頼性を推定することで最適なリンク配置方法の 候補を絞り込む



ほうがよい?

追加する: 最大99%

追加しない: 最大93% リンクAを追加する配置

方法を優先的に調べる

関連文献

[1] M. Nishino, T. Inoue, N. Yasuda, S. Minato, M. Nagata, "Optimizing network reliability via best-first search over decision diagrams," in *Proc. IEEE Conference on Computer Communications (INFOCOM 2018)*. April 2018.

担当者

西野 正彬 (Masaaki Nishino) 協創情報研究部 言語知能研究グループ

人工知能は文脈を読んで翻訳できるか?

ューラル翻訳の文脈理解度をテストする~



どんな研究

複数の文を一度に入力した場合に、 機械翻訳システムが文脈を理解して 適切な訳文を生成できるかを評価す る方法の研究です。先行する文に よって入力文に対する正解訳が変化 する例を集め、訳し分けの鍵となる 言語現象ごとに分類して、文脈理解 度テストとしてまとめました。

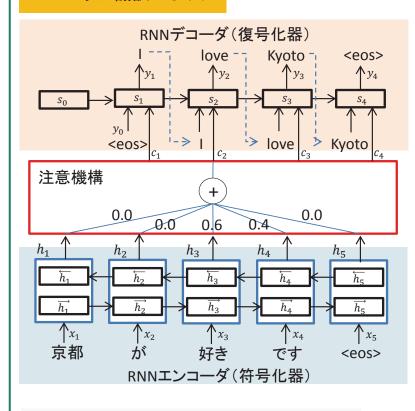
どこが凄い

日本語から英語への翻訳において文 脈の理解が必要となる事例を系統的 に収集した初めてのデータベースで す。主語や目的語の省略が多く、語 義の違いを漢字で表現するという日 本語の性質から、英仏翻訳を対象と した先行研究とは異なる独自のアプ ローチが必要でした。

めざす未来

深層学習の発展により、一つの文の 意味を大まかに伝える機械翻訳技術 は確立しました。今後は、文脈や状 況を理解して、省略を補完しながら 一貫性のある文を生成するだけでは なく、個人性・感情や文体などまで 伝達可能なコミュニケーションツー ルとしての機械翻訳をめざします。

ニューラル翻訳のしくみ



曖昧性解消:日英辞書から多義語を探索

文脈(直前の文)を利用したニューラル翻訳の予備実験

✓ 翻訳精度は上がるが何が起こっているか分からない

翻訳の単位	社説	講演	字幕
1文ずつ(最大50単語)	15.17	10.04	11.00
2文ずつ(最大100単語)	16.36	10.30	11.37

ニューラル翻訳の文脈理解度テスト

- ✓ 直前の原文と訳文、現在の原文と訳文の4つ組を作成
- ✓ 正解の確率が誤りの確率より大きくなる割合でモデルの 能力を評価(言語現象ごとに誤り分析できる)

	照応•省略		結束性•一貫性		
	代名詞	冠詞	対応	繰り返し	曖昧性解消
手法1	00%	♦♦%			
手法2	ΔΔ%	□□%			

対応・繰り返し:作業者の内省により構造的な制約を探索

	原言語	目的言語		原言語	目的言語
先行文	昨日、渋谷へ行った。	I went to Shibuya yesterday.	先行文	いい時計ですね。	It's a nice clock.
入力文	すごい人だった。	正解:There were a lot of people. 誤り:He was a great man.	入力文	1	正解:This clock is a memento of my father. 誤り:This watch is a memento of my father.

代名詞: 日本語の主語や目的語の省略を分析したデータとその機械翻訳による英語訳から作成

	原言語	目的言語
先行文	申し訳ありませんが、 <mark>先生</mark> は午後少し遅れているんです。	I'm afraid that the doctor is running a bit late this afternoon.
入力文		正解:It might be about 20 minutes before <mark>he</mark> can see you. 誤り:It might be about 20 minutes before we can see you.

関連文献

- [1] M. Morishita, J. Suzuki, M. Nagata, "NTT Neural Machine Translation Systems at WAT 2017," in Proc. The 4th Workshop on Asian Translation (WAT-2017), 2017.
- [2] R, Bawden, R. Wennrich, A. Birch, B. Haddow, "Evaluating Discourse Phenomena in Neural Machine Translation," in Proc. The 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-2018), 2018.

担当者

永田 昌明 (Masaaki Nagata) 協創情報研究部 言語知能研究グループ

ことばの発達がゆっくりなお子さんの特徴を探る

~小児医療現場で収集したデータの解析からみえてきたこと~



39.5 語

どんな研究

ことばの発達は個人差が大きく、発 達が早い子もいれば、ゆっくりな子 もいます。本研究では、ことばの発 達がゆっくりなお子さんに焦点を当 てて、語彙発達データを小児医療現 場で取得することにより、幼児の言 語習得メカニズムや個人差を生み出 す仕組みの解明をめざしています。

どこが凄い

ことばが遅いという理由で来院した 子を追跡調査し、1) 定型発達児と同 様の語彙爆発が起こること、2)定型 発達児が獲得していく語とほぼ同様 の語を獲得していくことが分かりま した。これは、ことばに遅れが見ら れても定型発達児とほぼ同様のプロ セスをたどることを示しています。

定型発達児

めざす未来

ことばの発達がゆっくりなお子さん の特徴とそのプロセスを詳細に把握 し定型発達児との違いを知ることで、 ヒトが言語をどのように習得するか に関する理論確立へのヒントになる とともに、小児医療現場での言語訓 練法に関する新しい技術の指針提案 にも貢献することができます。

日誌法による縦断データ解析



- 対象児: 言語発達遅滞児2名
- 方法:新しく言えるようになった語を親 が家庭で語彙成長アプリに随時記録
- 分析: 折れ線近似による特徴推定

B児(初診:2才0ヶ月) A児(初診:2才7ヶ月) 得 開始時 語 彙 数

語彙学習速度 語彙爆発開始時期 語彙爆発後 語彙爆発前 月齢 累積語彙数 1.21 語/日 A児 44.1 ヶ月 136 語 0.36 語/日 1.84 語/日 B児 0.27 語/日 33.7 ヶ月 94 語

0.83 語/日

範囲:0.18-1.85

20.1 ヶ月

0.18 語/日

範囲:0.05-0.30

結果

急激な語彙学習速度の変化あり

→ ことばの発達がゆっくりな子でも**語彙爆発**が存在 (ただし、爆発までに多くの語彙数を獲得する必要あり)

語彙チェックリストデータの解析



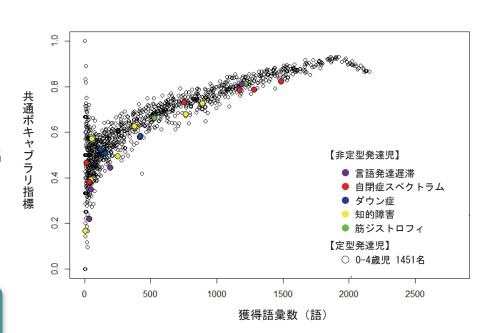
- 対象児: 非定型発達児 20名
- 方法: 来院時に, 現時点で言える語を チェックリストアプリで親に回答
- 分析: 定型発達児の多くが獲得していく 語とどの程度共通しているかを算出

 $|setVoc(i) \cap setAoA(i)|$ VocIndex(i) =|setAoA(i)|

VocIndex(i): 共通ボキャブラリ指標 setVoc(i): 幼児iの発話できる語の集合

setAoA(i): setVoc(i)の単語と同数分、獲得月齢(AoA)の 早い単語から順に選んできた集合

ことばの発達に遅れのある子も 定型発達児と同様の語を覚えていく傾向を明確化



本研究は竹田綜合病院と電気通信大学との共同研究成果です

関連文献

- [1] Y. Cao, Y. Minami, Y. Okumura, and T. Kobayashi, "Analyzing Vocabulary Commonality Index Using Large-scaled Database of Child Language Development," 11th edition of Language Resources & Evaluation Conference (LREC), 2018.
- [2] 阿久津由紀子, 小林哲生, 渡辺佐和, 齋藤貴美子, 南泰浩, "語彙成長記録アプリ活用による言語発達遅滞児2例の語彙発達の分析," 第18回日本 *言語聴覚学会*, 2017.

担当者

小林 哲生 (Tessei Kobayashi) 協創情報研究部 インタラクション対話研究グループ

物知りロボットとおしゃべりしながら賢くなろう

~複数ロボット対話制御に基づく雑談と質問応答の融合~



どんな研究

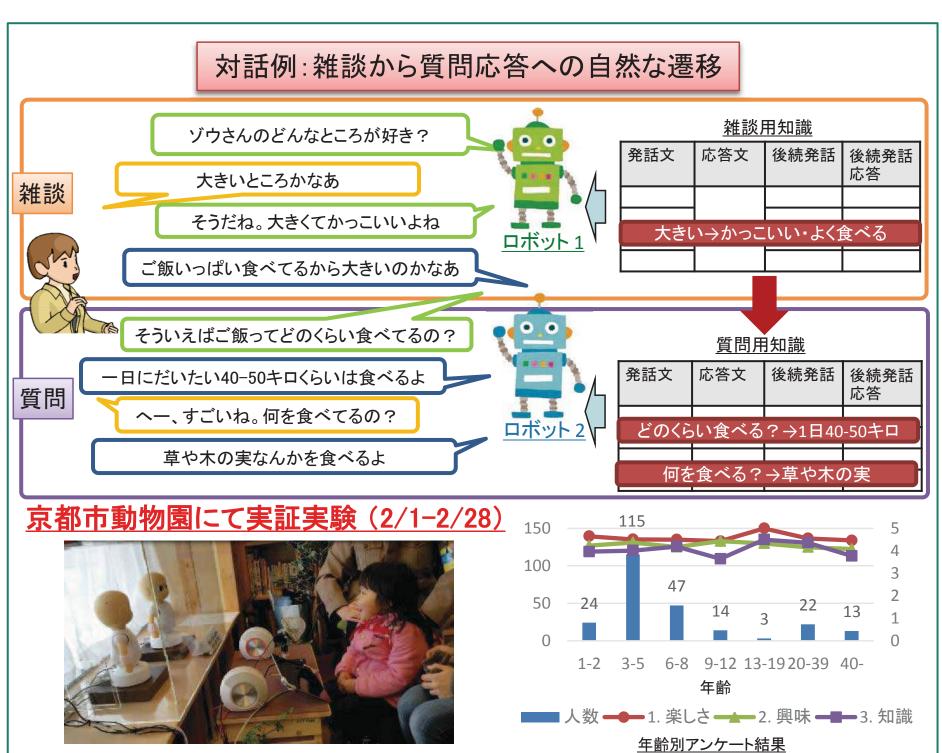
ユーザの対話への興味を引き出しな がら知識を伝える対話システムを世 界で初めて実現しました。従来は 別々に作られていた、質問応答など の知識を伝える機能と、雑談を行う 機能とを融合することで、質問応答 と雑談を違和感なく相互に行き来で きる対話を実現しています。

どこが凄い

従来の雑談用の発話知識は話題の多 様さ重視のため、質問応答用の知識 に比べ詳細さが不足していました。 本研究では複数ロボットの連携によ り雑談の話題を自然に制約すること で、一定の詳細さを維持した雑談知 識での対話を可能にし、質問応答と 雑談の融合を実現します。

めざす未来

ロボットがより身の回りに普及し社 会に受け入れられていくには、ロ ボットとユーザが信頼関係を構築で きることが重要です。私たちは、対 話を通してユーザと信頼関係を構築 し、ユーザを様々な側面からサポー トするパートナーとしてのロボット の実現をめざしています。



関連文献

- [1] H. Sugiyama, T. Meguro, Y. Yoshikawa, J. Yamato, "Avoiding breakdown of conversational dialogue through inter-robots coordination," in Proc. International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2018), 2018 (to appear).
- [2] 杉山弘晃, 水上雅博, 成松宏美, "複数ロボット協調による一問一答型雑談対話からの脱却", 人工知能学会全国大会, 2018 (to appear).

担当者

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

13

いつでもどこでもそれっぽくしゃべれます!

~スマホで音声リズムを英語母語話者っぽく変換~



どんな研究

これまで我々は、音声の発話リズムを変換する技術 "それっぽくしゃべります"を提案し、知覚実験や脳機能計測実験によりその有効性を示ってきました。この展示では、スマホを使って、日本語母語話者が発話りた音声を英語母語話者っぽい発話りたるに変換する技術を紹介します。

どこが凄い

日本語母語話者および英語母語話者が発話した大量の英語音声から機械学習することで、日本語母語話者の発話リズムを英語母語話者のそれに変換するシステムを構築しましたに意文で表す。 音声の発話リズムをその場で英語母語話者っぽく変換できます。

めざす未来

日本語母語話者が発話した英語音声は、ネイティブに伝わりにくいことがよくあります。この技術を応用して、ネイティブに聞き取りやすい英語音声に変換することで、日本語母語話者の英語での音声コミュニケーションを支援できると期待されます。

日本語母語話者が発話した英語音声は ネイティブに伝わりにくい

日本語母語話者はモーラ拍リズム^(注1)で 英語を発話する傾向がある



注1:単調なリズム

発話リズムを変換すると伝わりやすくなる

発話リズム変換は日本語母語話者の英語音声の 発話リズムを強勢拍リズム(注2)に変換する技術



注2:強弱のあるリズム

発話リズム変換のための3つの技術

頑健な声道スペクトル推定

PEAR(位相等化線形予測)法[1]

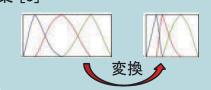
声道スペクトル系列から 時間関数を抽出

NTD(非負値時空間分解)法[2]



リズム変換則の学習

リズムを自動変換するために、 大量の英語音声(日本語母語 話者 350名、英語母語話者 700名)からNTD法により抽出 した時間関数を機械学習する ことで発話リズム変換則を構 築[3]



スマホを使って英語母語 話者っぽい発話リズムに変換



日本語母語話者がスマホに向かって 発話した英語音声を無線LANで飛ばし、 サーバでネイティブっぽく変換

英語母語話者っぽい発話リズムを スマホによりその場で確認できる

関連文献

- [1] S. Hiroya, T. Mochida, "Phase equalization-based autoregressive model of speech signals," in Proc. Interspeech, pp. 42-45, 2010.
- [2] S. Hiroya, "Non-negative temporal decomposition of speech parameters by multiplicative update rules," *IEEE Trans. Audio, Speech, and Lang. Process.*, Vol. 21, No. 10, pp. 2108–2117, 2013.
- [3] 廣谷定男, "非負値時空間分解法と機械学習を用いた任意音声の発話リズム変換," 日本音響学会秋季研究発表会講演論文集, pp. 423-424, 2017.

担当者

廣谷 定男 (Sadao Hiroya) 人間情報研究部 感覚運動研究グループ

離れていても盛り上がりを共有できる

~双方向性の臨場感・一体感の向上をねらう拍手音符号化~



どんな研究

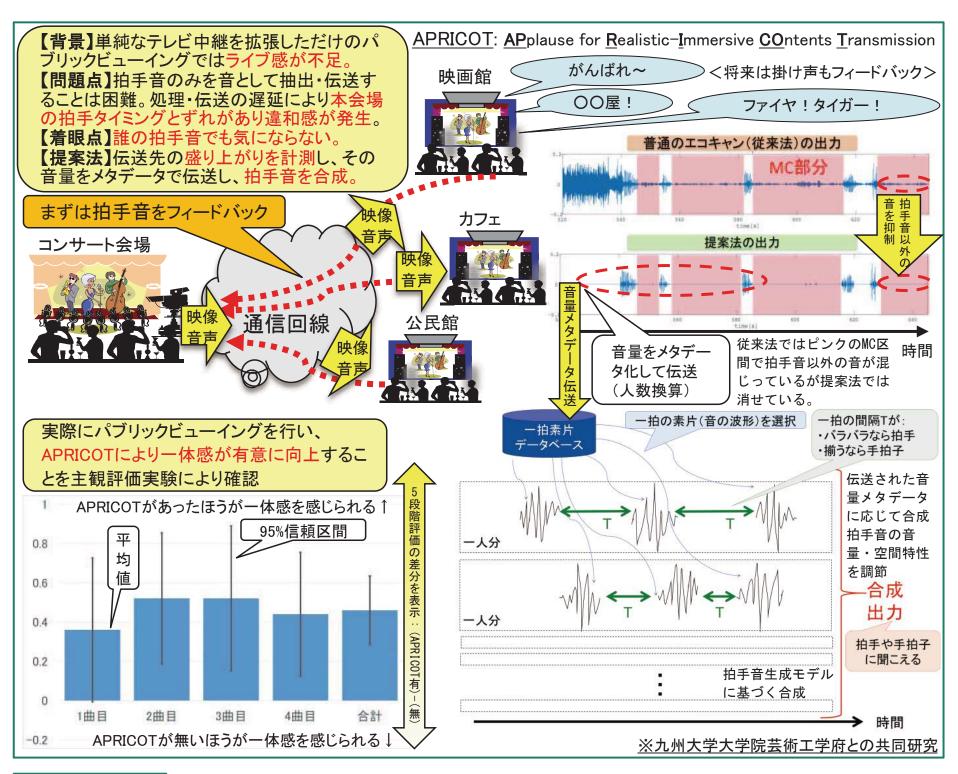
現在のパブリックビューイングはほ とんどが片方向通信(一方通行)の ため、伝送先の盛り上がりを本会場 に伝えることができません。この展 示では伝送先の拍手音を本会場に伝 えることで、双方向性のある臨場 感・一体感を向上させることができ るAPRICOTシステムを紹介します。

どこが凄い

自然な拍手音を合成することが可能 になりました。伝送先の環境音から 拍手音の音量を抽出することで、ど れだけ盛り上がっているかを計測で きるようになりました。遅延を感じ させないフィードバックにより、臨 場感・一体感の向上に寄与できるこ とを確認しました。

めざす未来

パブリックビューイング会場からで も、演者への賞賛や選手への激励が 届くようになります。通信の力で離 れていてもこちらの想いをあちらへ 伝えることができるようになり、実 際に会場に行けない場合でも参加し ている体験を味わうことができるよ うになります。



関連文献

- [1] 鎌本優, 河原一彦, 尾本章, 守谷健弘, "音楽鑑賞時に励起される拍手音の低遅延伝送に向けた検討," 2014年秋季日本音響学会講演論文集
- [2] K. Kawahara, A. Fujimori, Y. Kamamoto, A. Omoto, T. Moriya, "Implementation and demonstration of applause and hand-clapping feedback system for live viewing," in Proc. 141st Audio Engineering Society (AES) Convention, 2016.
- [3] M. Nishikawa, K. Kawahara, Y. Kamamoto, A. Fujimori, A. Omoto, T. Moriya, "Extraction of applause from a sound field for ambient transmission in a live viewing system," in Proc. IEEE Global Conference on Consumer Electronics (GCCE), 2017.

担当者

鎌本 優(Yutaka Kamamoto) 守谷特別研究室

15

照明光で色の鮮やかさを操る

~彩度強調成分を用いた分光スペクトルの制御~



どんな研究

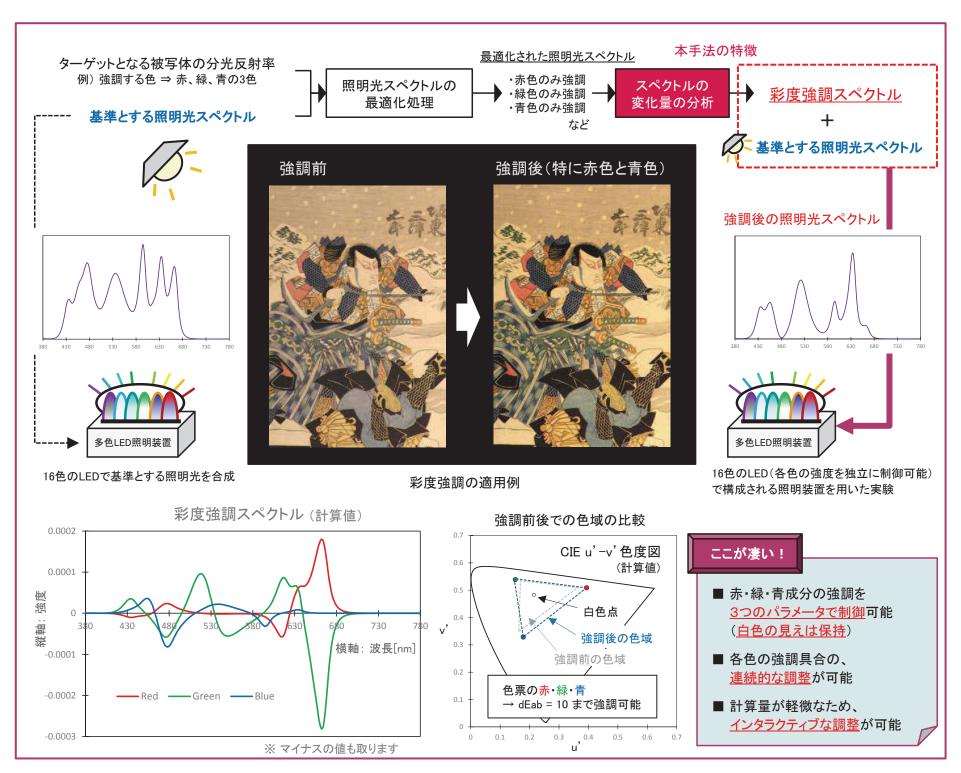
照明光の分光スペクトルを変化させることで、被写体の見た目の色の鮮やかさを制御する技術です。劣化により薄くなった色や、人間の目では区別しにくい僅かな色の差などを強調します。

どこが凄い

お肉や野菜などの食材を鮮やかに見せる照明が市販されていますが、特定の色に対してのみ有効で、また白色にも色がつき全体のカラーバランスが崩れる問題があります。本技術はホワイトバランスを保ちつつ様々な色の被写体の彩度を同時に強調し、その度合いの調整が可能です。

めざす未来

経年劣化等により退色した文化財の 仮想復元、医療分野での診断支援、 食品や工業製品の検査支援(変化の 状態や混入異物の強調)、店舗内で 探したい/強調したい物品への注意 喚起等を、照明光スペクトルの変調 のみで実現します。



関連文献

- [1] M. Tsuchida, K. Hiramatsu, K. Kashino, "Designing Spectral Power Distribution of Illumination with Color Chart to Enhance Color Saturation," in *Proc. IS&T 24th Color and Imaging conference (CIC24)*, pp. 278–282, 2016.
- [2] M. Tsuchida, K. Hiramatsu, K. Kashino, "Color Enhancement by Optimizing the Illumination Spectrum," NTT Technical Review, Vol. 16 No. 1, 2018, https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201801ra1.html

担当者

土田 勝(Masaru Tsuchida) メディア情報研究部 メディア認識研究グループ

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~



どんな研究

会話の中で複数の人が同時に話して いる時でも、人間は聞きたい人の声 に集中し、聞き分けること(選択的 聴取)ができます。一方、コン ピュータにはその能力がなく、聞き たい人の声だけをうまく聞き取るこ とはできません。コンピュータによる 選択的聴取の研究を進めています。

どこが凄い

事前に収録した聞きたい人の声を補 助情報として利用し、複数人が同時 に話している時に、その人の声だけ を聞き取ることができる技術『適応 型ニューラルネットワーク』を実現 しました。これにより、聞きたい人 の声に耳を傾けることができるコン ピュータを実現しました。

めざす未来

ロボット・ホームアシスタント・ス マートスピーカなどの遠隔音声収音 装置が、注目すべき話者の声だけを 聞き取ることができるようになりま す。それにより、例えば、ロボット が特定の人の声にのみ反応するなど、 人とより自然に会話できるようにな ります。

課題

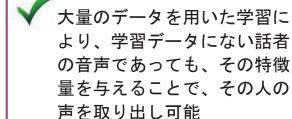
- 複数人での会話や、テレビの音が背景で流れて いる時など、日常の様々な場面において、 人の声同士が混ざることはよく起こる
- 従来のコンピュータ・音声入力装置は、混ざった 音声が入力されてくると、その中から、聞きたい 人(目的話者)の声だけに集中して聞き取るとい うことはできなかった



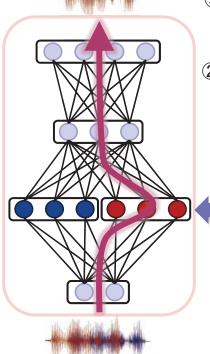
Picture designed with Sweet Home 3D. Includes 3D models created by Reallusion, Pencilart, Scopia and eTeks.

深層学習に基づく選択的聴取

聞きたい話者の声の特徴に基づき、 混合音声からその話者の音声のみ を取り出すようにニューラルネット ワークを学習する



話者数や目的話者の方向など、 従来技術で必要とされる情報 が不要な、汎用性の高い手法



混合音声

③ 聞きたい話者の声のみが出力される

② 聞きたい話者の特徴量に適応して、ニューラル ネットワークがその振る舞いを自動的に変え、 その話者を取り出すようになる

話者特徴量 抽出

① 事前に収録した聞きたい話者 の音声データ(数秒程度の事前 録音)から、その話者の声を表 す特徴量を抽出する



Picture designed with Sweet Home 3D. Includes 3D models created by Reallusion

関連文献

- [1] K. Zmolikova, M. Delcroix, K. Kinoshita, T. Higuchi, A. Ogawa, T. Nakatani, "Speaker-aware neural network based beamformer for speaker extraction in speech mixtures," in Proc. Interspeech, 2017.
- [2] M. Delcroix, K. Zmolikova, K. Kinoshita, A. Ogawa, T. Nakatani, "Single channel speaker extraction and recognition with SpeakerBeam," in Proc. of 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP' 18), 2018.

担当者

デルクロア マーク (Marc Delcroix) メディア情報研究部 信号処理研究グループ

17

二択問題にして解くことでAIは賢くなる

~深層学習による仮説比較と音声認識結果選択への応用~



どんな研究

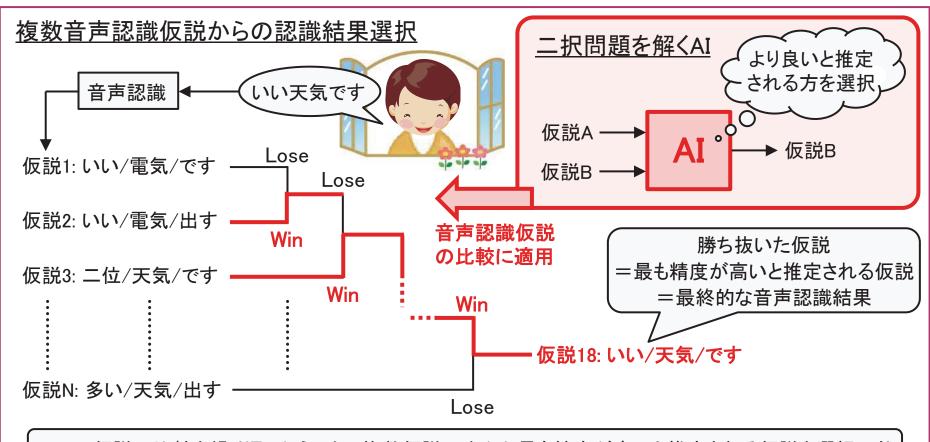
本展示で紹介するAI(深層学習モデル)は、ある問題に対する解の候補(仮説)が二つ与えられたときに、より良いと推定される方を選択します。本AIを複数の音声認識仮説から最終的な認識結果を選択する問題に応用することで、認識精度を大きく向上させることができます。

どこが凄い

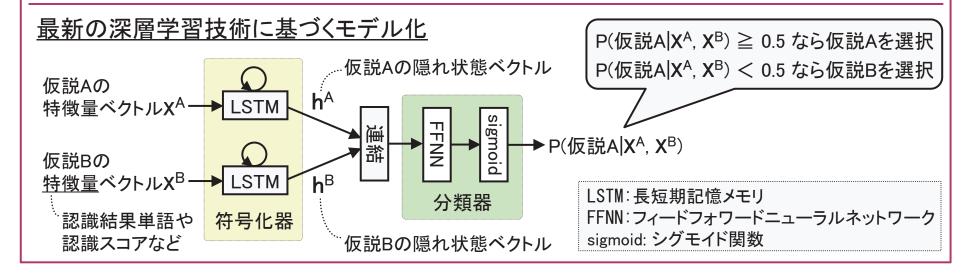
本AIは最新の深層学習技術に基づき モデル化されています。二択問題を 解くことに特化することで、高精度 化・コンパクト化を実現しています。 現在最も広く用いられている仮説選 択モデルと比較して、1/10以下のパ ラメータ数で、より高精度に仮説の 選択を行うことができます。

めざす未来

本AIは音声認識に限らず機械翻訳や文書要約など、複数の仮説が出力されるタスクに幅広く応用可能です。また我々の普段の生活においても二択問題は頻繁に生じます。将来的には本AIが我々の重要な決定においてより良い選択ができる上期待できます。



二つの仮説の比較を繰り返し行うことで複数仮説の中から最も精度が高いと推定される仮説を選択可能



関連文献

[1] A. Ogawa, M. Delcroix, S. Karita, T. Nakatani, "Rescoring n-best speech recognition list based on one-on-one hypothesis comparison using encoder-classifier model," in *Proc. of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)*, April 2018. [2] 小川厚徳, デルクロア・マーク, 苅田成樹, 中谷智広, "一対一の仮説比較を行うencoder-classifierモデルを用いたNベスト音声認識仮説のリスコアリング," 日本音響学会2018年春季研究発表会講演論文集, 1-8-9, March 2018.

担当者

小川 厚徳 (Atsunori Ogawa) メディア情報研究部 信号処理研究グループ

音だけから情景を推定

~音から画像認識結果を予測するクロスメディア情景分析~



どんな研究

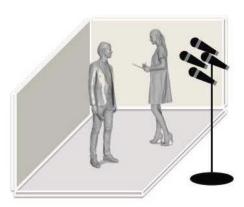
部屋などの屋内空間に配置したマイ クロホンアレイにより収音した音だ けを分析することによって、まるで 画像認識したかのように「どこにど んな物体があるのか」を推定する研 究です。生活や公共空間での見守り や防犯等への応用に向けての研究開 発を進めています。

どこが凄い

『音から画像認識結果を予測する』 という試み自体、これまで取り組み 例のない新しい技術課題です。本研 究では、入力が音響特徴、出力が画 像認識結果となるような、クロスメ ディアな深層学習モデルを設計し、 技術的な実現可能性を初めて示しま した。

めざす未来

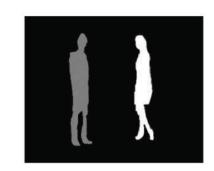
現在の見守り・防犯技術はカメラの 利用を前提としており、プライバ シー性の高い家庭や公共空間には適 用しにくい場合がありました。本技 術により、カメラの設置が好ましく ない空間の様子もわかりやすく確認 できるようになるため、適用領域の 拡大が期待できます。





クロスメディア 情景分析

②音声から物体の種類・形状を予測

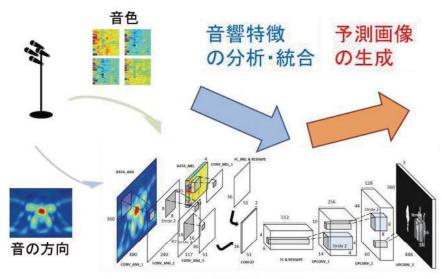


①マイクアレイで音声収録

③予測結果を画像として直接出力

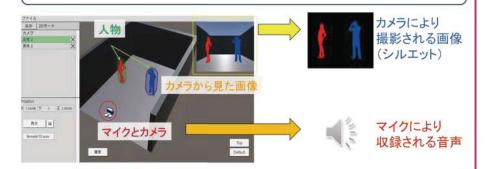
音響特徴から、空間にある物体の種類や位置 を予測する深層学習モデルを提案

異なる音響特徴を分析・統合し、出力画像を生成する処理を 一つのニューラルネットにより実現



ニューラルネットのアーキテクチャ

多様な状況をシミュレーションにより再現する ことで学習を効率化し、実現性を確認



カメラでは写らない・写したくない空間の様子 も、音だけから情景を描写可能

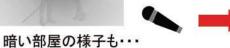
見にくい暗がりや写したくない空間の様子も音から描写







カメラでは 写らない





音声から

関連文献

[1] オストレクミレラ, 入江豪, 亀岡弘和, 木村昭悟, 平松薫, 柏野邦夫, "Seeing through Sounds: 音響情報からの視覚的情景理解に向けて," 画像の

担当者

入江 豪 (Go Irie) メディア情報研究部 メディア認識研究グループ

浮像(うくぞう)

~影を駆使して絵に奥行きを与える光投影技術~



どんな研究

影を使った錯覚を利用して、印刷された対象や手書き文字が浮いてみえるように錯覚させる光投影技術です。 画面上の対象に影を書き加えると、 その対象が浮いているかのように見 えることは知られていますが、現実 の物体を浮いているかのように見せ る技術は報告されていませんでした。

どこが凄い

本技術を利用することで、前もって 準備した絵だけではなく、手書を た文字や絵に対しても、浮き上がっているかのような視覚効果を即座に 与えることができます。また本技術による影の効果により、紙なのに透明フィルムであるかのような質感を 実物体へ与えることができます。

めざす未来

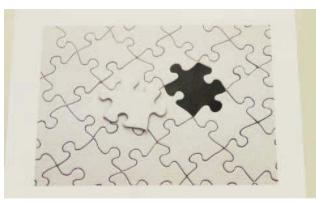
錯覚を利用することで、現実場面でのモノの知覚や認知をより便利に、そしてより楽しい方向へ変えていくことが可能です。普段私たちが何気なく行なっている見ること、聞くことがあることをより一層楽しむ開発をめざしています。

●本技術による奥行き錯覚

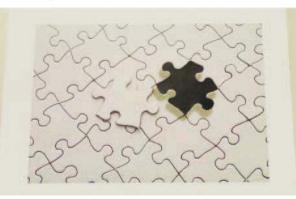
①印刷した対象に対して

②影パタンを投影すると

③影パタンの効果により、印刷された対象 (ここではパズルの形をした暗い空間)が 浮いて見えます







●本技術による質感錯覚

影のパターンを工夫することで、印刷された対象に透明感を与えることができます。

オリジナルの印刷物







透明感をもたないように影をつけた印刷物



関連文献

[1] T. Kawabe, "The illusion of floating objects caused by light projection of cast shadow," in *Proc. the Eighteenth Annual Meeting of the Vision Sciences Society (VSS2018)*, 2018.

担当者

河邉 隆寛 (Takahiro Kawabe) 人間情報研究部 感覚表現研究グループ

声の雰囲気や聞き取りやすさを変換する

~深層生成モデルを用いた音声属性変換~



どんな研究

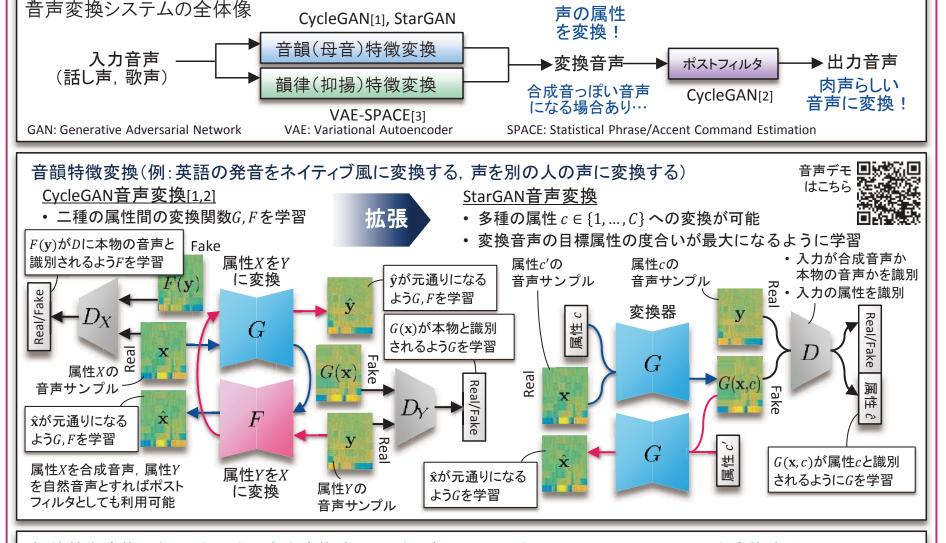
深層生成モデルを用い、音声の属性 や非言語情報を柔軟かつ高品質に変 換・生成する技術の研究です。例え ば非母語話者の音声を母語話者風の 発音の音声に自動変換したり、好き な歌手の歌い方で任意の楽譜を歌わ せたりすることが可能です。

どこが凄い

音声という現象には言語情報以外に 様々な変動要因が複雑にもつれ合っ て混在しています。本研究では、敵 対的生成ネットワーク(GAN)や変分 自己符号化器(VAE)などの深層生成 モデルを用いてこれら変動要因のも つれを解き、属性の個別操作を可能 にする音声変換法を検討しています。

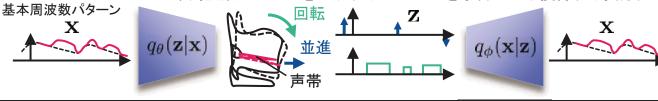
めざす未来

聞き取りにくい音声を聞き取りやす い音声に変換することでコミュニ ケーションを円滑化することができ ます。本研究では、音声変換技術を 通して非母語話者音声や喉頭摘出者 などの音声を聞き取りやすい発音や 抑揚の音声に変換するシステムなど の実現をめざしています。



韻律特徴変換(例:歌の歌い方を変換する,話し声のイントネーションやアクセントを変換する)

VAE-SPACE [3] · 基本周波数パターン(声の高さの時間変化)から甲状軟骨の動きを推論するプロセスと甲状軟骨の動き から基本周波数パターンを生成するプロセスを学習により獲得する変分自己符号化器(VAE)



歌声合成への応用:

歌手の歌い方を学習し、 任意の楽譜を入力して 歌わせることが可能

関連文献

- [1] T. Kaneko, H. Kameoka, "Non-Parallel voice conversion using cycle-consistent adversarial networks," submitted to EUSIPCO2018, 2018.
- [2] K. Tanaka, T. Kaneko, N. Hojo, H. Kameoka, "Synthetic-to-natural speech conversion using cycle-consistent adversarial networks," submitted to EUSIPCO2018, 2018.
- [3] K. Tanaka, H. Kameoka, Kazuho Morikawa, "VAE-SPACE: Deep generative model of voice fundamental frequency contours," in Proc. ICASSP,

担当者

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

選んで創るお気に入りの画像

~DTLC-GANを用いた画像の階層的理解・生成制御~



どんな研究

深層学習の発展により低次元データ から高精細な画像を生成できるよう になりつつあります。しかし、生成 する際にコントロールできることは 限定的で好みの画像を創ることは困 難でした。本研究では属性を階層的 に選びながら画像生成できる機構を 考案し、この問題に対処しました。

どこが凄い

属性の階層的表現を人手で与えよう とすると、詳細なラベル付けが必要 で大変です。そこで、本研究では、 教師データなしまたは最上位層の教 師データだけで下位層の表現を自動 発見できるアルゴリズムを考案しま した。これにより、手間を増やすこ となく高い表現力を獲得できます。

めざす未来

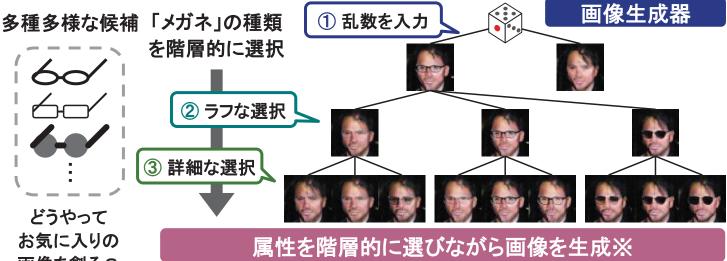
本技術は、深層生成モデルの代表的 な一つであるGANの自然な拡張であ り、画像生成だけにとどまらず、音 声などの他のメディアデータの生成 への応用も期待されます。また、得 られた階層的表現は生成だけではな く、検索・編集・変換などの他タス クで活用することも可能です。

目的:お気に入りの画像を「選んで」 創る



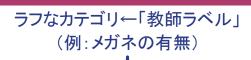


画像を創る?



※同様にして「表情」の選択や「髪型」の選択も可能

方法:DTLC-GAN「詳細なカテゴリはデータから学習(教師ラベル不要)」



※具体的な意味はデータ から見つけ出すことが必要



DTLC-GAN





階層的 カテゴリ分類

詳細なカテゴリ←「データから学習」 (例:サングラス. 細縁メガネ…)

Decision Tree Latent Controller Generative Adversarial Networks CVPR 2018 (to appear) [2]

展開:ベーシックな技術で様々な応用が可能

展開①:音声などの他のメディアデータの生成

展開②:検索•編集•変換

関連文献

[1] T. Kaneko, K. Hiramatsu, K. Kashino, "Generative Attribute Controller with Conditional Filtered Generative Adversarial Networks," in Proc. The 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[2] T. Kaneko, K. Hiramatsu, K. Kashino, "Generative Adversarial Image Synthesis with Decision Tree Latent Controller," in Proc. The 31st IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018 (to appear).



http://www.kecl.ntt.co.jp/people/ kaneko.takuhiro/projects/dtlc-gan/

担当者

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

音への注意を眼で測る

~瞳孔反応に現れる聴覚空間注意~



どんな研究

複数の音源が混在する日常環境の中、 適切な空間(方向)に注意を向ける ことで、私たちはその時々に聞きた い音を適切に聞き取ることができま す。この研究では、聴覚的注意を向 けている方向の明るさが、あたかも 対光反射のように瞳孔の開き具合に 反映されることを発見しました。

どこが凄い

聴覚的注意を向けている方向を知る には、これまでは頭や眼の動きと いった行動に頼るしかありませんで した。本研究の結果は、眼が動かな くても、視覚入力の明るさの空間分 布と、瞳孔径との対応を調べること により、聴覚的な注意方向をある程 度推定できることを意味します。

めざす未来

この注意方向推定の原理を利用すれ ば、どんな音がどんな条件で注意を ひきつけやすいかを知ることもでき るでしょう。人間が機器を安全に運 転する際のサポートや、注意機能障 害の診断・リハビリといった様々な 分野への応用も期待されます。

見ているものの明るさによって、瞳孔の大きさが変わることは良く知られてい ます。しかし、瞳孔の大きさは、目に入ってくる画像の明るさだけで決まるわけ ではなさそうです。

目を動かさず、視覚刺激を一定にした条件であっても、注意を向けて いる耳の方向に応じて、あたかもその方向を見ているかのように 瞳孔径が反応する現象を見つけました。



瞳孔は聴覚の注意が向いている方向の 明るさを反映する

実験方法

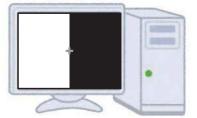
- 左右の明暗が異なる画像を呈示
- 眼を動かさないまま、右または左の耳に注意 を向けるように指示



- 注意を向けた方向に対応する視覚刺激の 明るさに応じて瞳孔径が変化
 - → 空間的注意のメカニズムは視覚と聴覚で共通?
- 無意識の反応

左右の明暗(輝度)が異なる画像

視覚刺激



耳への注意





瞳孔反応





縮瞳

散瞳

注意を向けた側の輝度を反映

関連文献

[1] H.-I. Liao, H. Fujihira, S. Furukawa, "The pupillary light response reveals the focus of auditory spatial attention," Association for Research in Otolaryngology (ARO) 41st Annual MidWinter Meeting. San Diego, California, USA. 2018.

担当者

リャオ シンイ(Hsin-I Liao) 人間情報研究部 感覚共鳴研究グループ

ウェルビーイングを測る、知る、育む

~"いきいきとした心的状態"の実現を科学する領域横断研究~



どんな研究

人や組織が持続的にいきいきと活動 している状態(ウェルビーイング) を実現するための設計論を人間科学 の視点から構築する研究です。科学 的な解明とともに、具体的なケース に適用可能な、要因の特定、計測、 モデル化、介入、評価からなる循環 プロセスに取り組みます。

どこが凄い

ウェルビーイングの要因は個人内に 留まらず、家族、友人、集団など 様々な段階の人間関係によって規定 されます。私たちはそれぞれの段階 の要因を主観だけでなく、行動解析 や生体計測など総合的に分析するこ とが可能であり、ウェルビーイング を包括的な視点から検討します。

めざす未来

情報技術によって人や組織のウェル ビーイングを実現するためのエビデ ンスを提供します。個人毎や年齢・ 状況による変化、さらには集団の ウェルビーイングを定量可能な指標 を見出し、それぞれの特性にあった 循環プロセスを設計する指針や方法 論を提案します。

情報化社会において、「どのように情報技術が心の豊かさをサポートできるか」が問われています。CS研の 人間科学の研究、心的状態の計測技術、感覚情報の提示技術は、この問い(現代社会におけるウェルビー イング*のあり方)をエビデンスに基づいて検討するための重要な素地となります。

*ウェルビーイング(Wellbeing): 人間が心身の潜在能力を発揮し、意義を感じて「いきいきと」満足して生きている状態。 主観、行動、生理の多次元的な計測によって定量化され、ライフステージ、対人関係、組織などによっても変化します。

ウェルビーイングの設計

人間科学の研究

ウェルビーイング

の要因を特定

(Specification)

心的状態の計測技術

要因と関連する 情動•行動•関係 の計測とモデル化 (Measurement)

感覚情報の提示技術

五感を通した 継続的介入 (Intervention) 介入による状態 の変容を評価 (Assessment)



持続的にいきいきと 活動している状態

ウェルビーイングの研究対象



関連文献

- [1] R, Calvo & D, Peters, "Positive computing," MIT Press, 2014. (邦訳『ウェルビーイングの設計論』監訳:渡邊淳司,ドミニク・チェン 翻訳:木村千里, 北川智利,河邉隆寬,横坂拓巳,藤野正寬,村田藍子)
- [2] 渡邊淳司, 村田藍子, 安藤英由樹, "持続的ウェルビーイングを実現する心理要因," *日本バーチャルリアリティ学会学会誌*, pp. 12-18, 2018.

担当者

渡邊 淳司 (Junji Watanabe) 人間情報研究部 感覚共鳴研究グループ

人工知能で人の聴こえの仕組みを理解する

~機械学習モデルによる聴覚神経機構の分析~



どんな研究

我々が日常出会う音は非常に複雑で す。そして、耳や脳は膨大な数の神 経細胞からなるネットワークで音の 情報を処理しています。私たちは、 現代の「人工知能」を構成する機械 学習技術の力を借りて、これまでに ない形で、耳と脳と音との複雑な関 係性を理解しようと試みています。

どこが凄い

典型的な仮説検証型の認知神経科学 研究では、設定される仮説の良し悪 しが、研究の進展を左右します。 我々は、複雑なデータに含まれる関 係性を抽出する機械学習技術を導入 することで、強い仮説を前提とせず に、聴覚のメカニズムに関するヒン トを得ることができました。

めざす未来

このアプローチにより、研究者がこ れまで見過ごしてきた課題や仮説を 発見・提起することができ、認知神 経メカニズムの本質的な理解につな がる可能性があります。また、高齢 者や難聴者も含む、一人ひとりに とって聴きやすい音を設計するのに も役立つでしょう。

音は環境や音源に関する様々な情報を反映しています。 耳や脳はその情報を複雑な神経メカニズムで処理して います。

私たちは、人工知能にも使われる機械学習の手法を データに適用することで、聴覚における音の表現につい て、これまでの研究では説明できなかった、より本質的 な仕組みを明らかにしました。

が 4.4-0 類 0 4.20

0 000

音の認識精度[%]

183 4.0 30 35 40



音の時間変化の表現[1]

- ▶ 時間変化の表現は、音の認識に 最適となるように進化してきた?
- ▶ 聴覚系の神経は音の時間変化の パターンに選択性を持ち、選択性 は脳の領野によって異なる
- ▶ 音認識のために訓練されたニュー ラルネットワークも、聴覚系と似た 選択性を持つ
- ▶ 認識精度が高いと、聴覚系との類 似性も高い

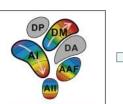
100

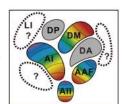
変調周波数 [Hz]

▶ 音処理の脳領域はどう分かれて いて、どのような特性か?

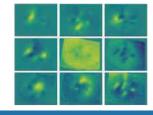
脳領域のデータ駆動解析 [2]

- ▶ 合成音と異なる自然音聴取時の マウス脳活動を広範囲で記録し、 機械学習で脳領域を自動分割
- ▶ 未知領域の可視化・機能同定へ (新潟大学との共同研究)





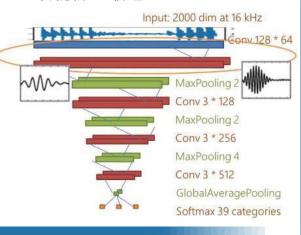
大脳皮質聴覚野の新領域同定へ



大脳皮質 聴覚野の9つ の活動単位

周波数選択性の意義[3]

- 聴神経の周波数チューニング特 性はなぜ獲得されたのか?
- ▶ 残響のある自然音を考えると、既 存の教師なし学習モデルよりも、 音認識のための教師あり学習(深 層学習)でよく説明できる
- ▶ 周波数選択性は自然環境下での 音認識に最適か



人工知能を使って、音を聴くための仕組みを理解するという新しい試み

関連文献

<u>₩</u> 0.8

|活動強|

モデルコ:平均活動

- [1] T. Koumura, H. Terashima, S. Furukawa, "Representation of amplitude modulation in a deep neural network optimized for sound classification," in Proc. 41st Annual Midwinter Meeting of the Association for Research in Otolaryngology, 2018.
- [2] 寺島裕貴, 塚野浩明, 古川茂人, "自然音刺激を用いたマウス聴覚野構造解析の試み," in 「脳と心のメカニズム」冬のワークショップ, 2018.
- [3] H. Terashima, S. Furukawa, "Reconsidering the efficient coding model of the auditory periphery under reverberations," in Proc. 41st Annual Midwinter Meeting of the Association for Research in Otolaryngology, 2018.

担当者

上村 卓也 (Takuya Koumura) 人間情報研究部 感覚共鳴研究グループ

一流打者はボールをどのように打っているか?

~打撃中の身体運動計測から認知過程を探る~



どんな研究

巧みな打撃を実現するためには、正 確なバット操作だけでなく、瞬時に 球筋を見極めたり、打つ・打たない を判断したりする優れた認知機能も 不可欠です。本研究では、ソフト ボールー流打者の潜在的(無自覚 的) 認知情報処理の特徴を明らかに することをめざしています。

どこが凄い

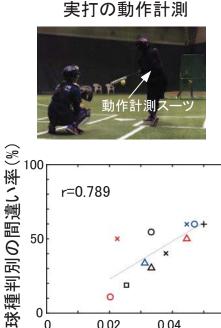
一流打者の認知機能を理解するには、 リアリティのある状況下で行動を観 測することが肝要です。私たちは、 実戦(投打対戦)場面やバーチャル リアリティ環境での身体運動計測を 通して、打撃のタイミング調節に関 わる球種判断能力や投球フォーム由 来の予測能力を測りました。

めざす未来

本研究のデータを蓄積していくこと で、認知機能に関する打者の特性や タイプを評価・判別することに役立 てます。また、こうした認知機能を 向上させるためのフィードバック手 法やトレーニング・メソッドを開発 し、選手のスキル上達を支援してい きます。

速球/チェンジアップボールへの運動反応から、 打者の認知情報処理過程を探る

1) 投打対戦型の実験



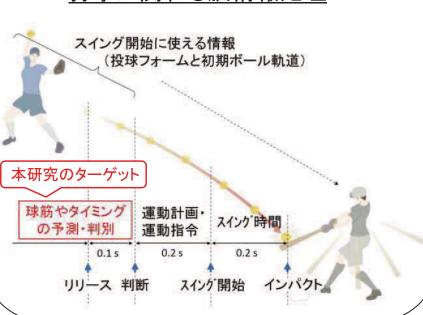
r=0.789

球種判別の計測



球種判別が正確 なほど、スイング 開始のタイミング 調節が正確

打撃に関わる脳情報処理



2) バーチャルリアリティ(VR) 打撃の実験

0.06

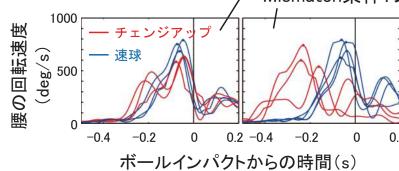
0.04

スイング開始のズレ(s)

Match条件: 正しいフォームと軌道の組合わせ Mismatch条件:フォームと軌道を入れ替え



0.02



フォームを入れ替えると スイングタイミングが狂う

関連文献

- [1] D, Nasu, M. Yamaguchi, T. Fukuda, N. Saijo, M. Kashino, T. Kimura, "Perception-action linkage in top athletes during batting," in Proc. Society for Neuroscience 47th Annual Meeting, 2017.
- [2] T, Kimura, D. Nasu, M. Kashino, "Utilizing virtual reality to understand athletic performance and underlying sensorimotor processing," in Proc. the 12th Biennial conference on the Engineering of Sport on behalf of the International Sports Engineering Association, 2018.

担当者

木村 聡貴(Toshitaka Kimura)スポーツ脳科学プロジェクト

·流打者はボールをどのように見ているか?

~打撃中の眼球運動計測から認知過程を探る~



どんな研究

野球やソフトボールのバッティング では、ボールの軌道を正確に捉える ことが不可欠です。0.5秒に満たな いごく短時間で、さまざまな球種、 球速、コースを見分け、適切に身体 を動かすという離れ業は、どのよう に実現されているのでしょうか。眼 球運動の計測でその秘密に迫ります。

どこが凄い

打者用のヘルメットに高速アイカメ ラを装着し、実際に投手が投げる ボールを打つ際の眼球運動を精度よ く計測することに成功しました。 トッププロや、それに準じるレベル の選手の眼球運動を解析することに より、優れたパフォーマンスの背後 にある認知過程が見えてきました。

めざす未来

バッティングフォームやスイングス ピードのような見えやすい特徴と異 なり、目の使い方は本人も自覚でき ない部分が多いので、技術レベルの 評価も、技術のコーチングも勘頼み でした。本研究が発展すれば、客観 的データに基づいた技術の評価や コーチングが可能になります。

バッティングでは

- 打者から見えるボールの速度は、リリース~0.3秒まではほとんど変化しないが(約0°/s)、 その後、インパクトまでに急激に速くなる(約0~700°/s)ため、その領域での視線に寄る 追従は困難
- ・ 視覚・運動情報処理の時間遅れがある以上、予測が不可欠
- ・ 眼球運動と動作の関係は未解明

実打時の打者の視線の動きを計測

投手: 元プロ野球、社会人野球; ストレート(130~140km/h)、カーブ(100~110km/h)

打者: プロ野球、社会人野球、元東京六大学野球

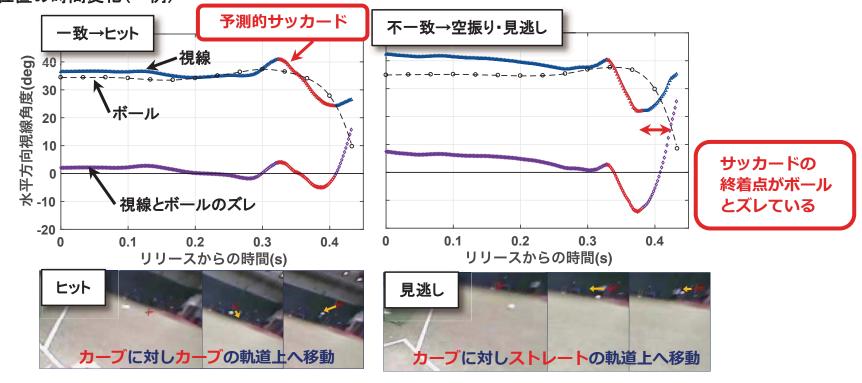
視線計測装置

視野カメラ

(30fps)

アイカメラ (500fps)

視線位置の時間変化(一例)



- ・ボールのリリースから約0.2秒の間に球筋を予測。予測の精度が打撃の成否を左右。
- ・リリース後約0.3秒以降、高速な眼球運動(サッカード)により、予測した球筋に視線をあらかじめ移動。
- 予測的サッカードは、ボールを見るためではなく、身体の動きを制御するため?

関連文献

[1] 木下祐輝, 柏野牧夫, "バッターはいかにボールを見ているのか? 一流打者の眼球運動戦略の解明を目指して," 日本野球科学研究会, 2017.

担当者

柏野 牧夫 (Makio Kashino) スポーツ脳科学プロジェクト

急ぐ方が正確?

~視知覚と運動に潜む情報処理のからくり~



どんな研究

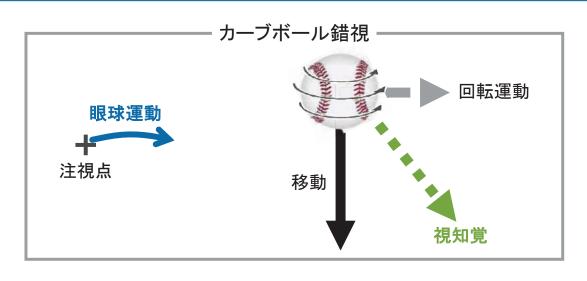
物体が回転しながら移動する場合、 物体の位置は回転運動の方向にズレ て見えます(カーブボール錯視)。こ の錯覚は、視知覚では生じるものの、 視線移動(眼球運動)では生じないと されてきました。本研究では、なぜ 視知覚はだまされて、運動はだまさ れないのかを調べました。

どこが凄い

視知覚や運動に関わらず、物体を見てから位置を判断するまでの時間が短ければだまされず、長ければだまされることが分かりました。これは回転する移動物体の定位では、いつ判断するかが重要であり、時間をかけることが必ずしも良いとは限らないことを意味しています。

めざす未来

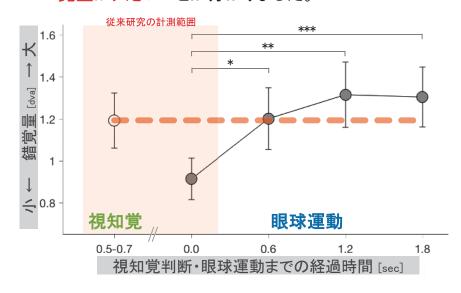
今回の研究から、視覚運動情報の入力から意思決定までにかける時間の重要性が分かりました。これらの知見から、ものを見てからから適切な時間で視知覚判断や運動反応を促すことを意識したスポーツ選手のトレーニングや、UIデザインの設計などへの応用を考えています。

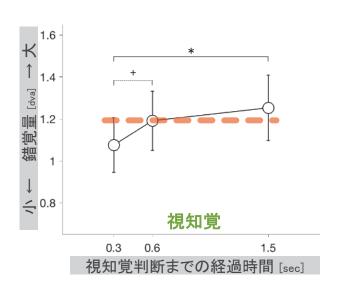


なぜ視知覚はだまされて、運動はだまされないのでしょうか?

実験

回転しながら移動する物体の映像が消えた位置を**視知覚判断**と 眼球運動で答える場合、経過時間に応じた錯覚量に違いが見られるかを調べました。





従来、視覚運動を含む移動物体の脳内表現は、視知覚と運動では異なっており、それぞれの位置表現は時間的に不変であると考えられてきました。本研究は、それらの位置表現が視覚情報処理の過程で同じように変化していることを明らかにしました。

関連文献

- [1] H. Ueda, N. Abekawa, H. Gomi, "Temporal integration of sensory evidence for position representation of a moving object containing motion signal in perceptual and motor decision making," in *Proc. The 44th Naito Conference on Decision Making in the Brain—Motivation, Prediction, and Learning*, 2017.
- [2] H. Ueda, N. Abekawa, H. Gomi, "Temporal change of position representation of a moving object filled with Gabor motion for saccadic eye movements" in *Proc. The 46th Annual Meeting of the Society for Neuroscience*, 2016.

担当者

上田 大志 (Hiroshi Ueda) スポーツ脳科学プロジェクト

力で感じるかたちとうごき

~ぶるなび4による環境・状況の呈示~



どんな研究

我々は、目で見るだけでなく、触っ たり動いたりして得られる体性感覚 情報も使って外界状況を理解してい ます。本研究では、触力覚の知覚原 理に基づいて開発した力覚呈示ガ ジェットで、視覚や音情報だけでは 伝わりにくい「体感」を伝えること をめざしています。

どこが凄い

映像に整合した力覚が与えられると、 映像だけでは得られない没入感・一 体感が得られ、視覚体験が強化され ます。また、視覚的には不自由な状 況でも、壁や障害物に直接触れずに その位置や形状を「体感」として捉 えることができるようになります。

めざす未来

触覚や力覚などの体性感覚を呈示す る技術により、視覚や聴覚に加えて、 あるいはそれらの感覚に不自由があ る場合に、体感を伴ったより豊かな 情報伝達や、親しい人同士の遠隔体 感コミュニケーションなどの実現が 期待されます。



(イメージ) ガジェット「ぶるなび4」(B4SF-P)により、

視覚情報がなくても、壁の空間的情報を 力覚で与えることにより、壁に当たらず に歩行できるようにすることが目標です。

力覚呈示ガジェットで映像に適合する力覚を与えることにより、没入感や一体感が得られます。 また、映像だけでは伝えにくいスポーツ動作における力の入れ具合のタイミングを伝えやすくす るなどの効果が期待されます。

ハンドル操作や車の加減速に応じた

「体感」が与えられている様子

関連文献

- [1] 五味裕章, 高椋慎也, 雨宮智浩, 伊藤翔, "力感覚の大きさ・方向をダイナミックに変化させる携帯デバイス," *「狙いどおりの触覚・触感をつくる技* 術」, 第4章, 第3節, (6) サイエンス&テクノロジー株式会社, 2017.
- [2] 五味裕章, "牽引感覚を与えるモバイルデバイスの開発と応用," スマートインフォメディアシステム研究会, 電子情報通信学会, 2017.
- [3] 五味裕章, "モバイルガジェット「ぶるなび」 ~引っ張って方向を伝える触覚デバイスの開発と応用~," 電波技術協会報, pp.22-25, 2016.

担当者

五味 裕章 (Hiroaki Gomi) 人間情報研究部 感覚運動研究グループ

平らなシートなのに凹凸感?!

~磁性シートに凹凸感を"書き込む"磁性触覚印刷技術~



どんな研究

従来の触覚提示技術は電気的に振動や力覚を提示するものが主でしたが、本研究では2枚の磁性ゴムシートのみでボコボコとした凹凸触覚の提示を行います。シート上の磁性パタンを編集可能にすることで、提示凹凸感を制御する手法を確立しました。

どこが凄い

磁性パタンの組み合わせによって凹 凸感の強度や粒度が設計できるため、 非常に簡素な構成ながら多様な触覚 表現への応用が可能です。触れ合わ せる箇所や対象を変えることで凹凸 感の強度を変えたり、特定の対象の みに凹凸感を提示させたりといった 表現もできます。

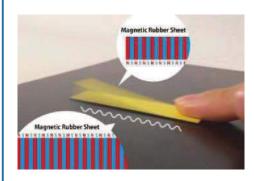
めざす未来

触覚提示には一切電源を必要とせず、 また一度書き込まれた磁性パタンは 長期間保持されるため、おもちゃや 本、床や壁面へと、触覚コンテンツ の幅を広げる展開が期待されます。 触覚の表現領域、応用領域を広げき より多くの方々に届けることができ ればと考えています。

磁性ゴムシートを媒体とした磁性触覚印刷技術を考案

【ポイント1】簡易かつ詳細着磁可能な 着磁手法を考案

パタン着磁した2枚の 磁性ゴムシート片をこ すり合わせるだけで、 シート間にあたかも凹 凸があるように感じる ことができる

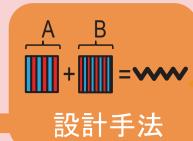


【ポイント3】応用として触覚絵本など様々な触覚コンテンツを実現

磁性触覚印刷技術

着磁装置

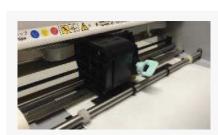




凹凸感提示物

- •触覚絵本
- ・触覚インタフェース など





磁性シートおよび 着磁に必要な装置 のすべてが汎用、 家庭用のものを用 いて構成可能

【ポイント2】提示 される凹凸感の設 計手法を確立

生じる凹凸感の強度、凹凸間隔はパタンの組み合わせによって異なり、設計も可能

関連文献

- [1] K. Yasu. "Magnetic Plotter: A Macrotexture Design Method Using Magnetic Rubber Sheets," in *Proc. of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 2017.
- [2] K. Yasu. "Magnetic Plotter: A Macrotexture Design Method Using Magnetic Rubber Sheets," in *ACM SIGGRAPH 2017 Studio (SIGGRAPH '17)*, 2017.

担当者

安 謙太郎 (Kentaro Yasu) 人間情報研究部 感覚表現研究グループ

MEMO



corevo®

~ NTTグループのAI技術~



Agent-Al

人らしさの追求、 人との共存

人の発する情報を読み 解き、意図・感情を理解

Heart-Touching-Al

潜在的な心身の 状態を理解

心と体を読み解き、深層 心理・知性・本能を理解

Ambient-Al

森羅万象から 未来を予測

人間・モノ・環境を読み 解き、瞬時に予測・制御

Network-Al

通信ネットワークの 知能化

複数のAIがつながり、 社会全体を最適化

各展示と corevo を構成する4つのAIとの関連性

Agent-Al

10

12

16

人工知能は文脈を

しながら賢くなろう

聞きたい人の声に

耳を傾けるコンピュータ

読んで翻訳できるか?

物知りロボットとおしゃべり

- いつでもどこでも それっぽくしゃべれます!
- 声の雰囲気や聞き取りやすさを

ネットワーク構造から 07 深層学習のしくみを知る

- 照明光で色の鮮やかさを
- 二択問題にして 解くことでAIは賢くなる
- 音だけから情景を推定

離れていても盛り上がりを

共有できる

Ambient-Al

- 膨大なデータから似た音声を 見つけます!
- 複数の問題に共通して重要な 情報の組合せを発見
- 人はどこから来て、 どこへ行くのか?
- 都市の今を知る

- ことばの発達がゆっくりな
- 浮像(うくぞう)

選んで創る

お気に入りの画像

- 11 お子さんの特徴を探る
 - 音への注意を眼で測る -流打者はボールを 25 どのように打っているか?
- ウェルビーイングを 23 測る、知る、育む
 - 人工知能で人の聴こえの 仕組みを理解する 27
 - 急ぐ方が正確?

26

力で感じるかたちとうごき

·流打者はボールを

どのように見ているか?

平らなシートなのに

- 深層学習をモバイル向けに 小さくします
- 光で機械学習を スピードアップ
- あなたの量子メモリを ちょっと拝借!
- こわれにくいネットワークを デザインします

Network-Al

Heart-Touching-Al

凹凸感?!

※「corevo®」は日本電信電話株式会社の登録商標です。

関連文献

- [1] "[特集] NTTグループにおけるAIの取り組み," NTT 技術ジャーナル, Vol. 28, No. 2, pp. 6-37, 2016. (http://www.ntt.co.jp/journal/1602/special.html)
- [2] "[特集] コラボレーションを通じて一緒に革新(レボリューション)を起こすNTTグループのAI技術「corevo™(コレボ)」," ビジネスコミュニケーション, Vol. 53, No. 8, pp. 6-21, 2016
- [3] http://www.ntt.co.jp/corevo/

担当者

納谷 太 (Futoshi Naya)

NTT コミュニケーション科学基礎研究所 オープンハウス 2018 実行委員会

(委員長) 岩田 具治

(副委員長) 雨宮 智浩

(委員) 白井 良成/中嶋 秀治/白木 善史



研究展示

(データと学習の科学

- **III 膨大なデータから似た音声を見つけます!** グラフ索引に基づく類似音声探索
- 複数の問題に共通して重要な情報の組合せを発見 共通因子を効率的に学習する低ランク回帰技術:MOFM
- 人はどこから来て、どこへ行くのか? 人流データ同化と学習型誘導
- 04 都市の今を知る 環境センシングと異種データ融合分析によるイベント解析
- ○15 深層学習をモバイル向けに小さくします 量子化による深層学習のモデル圧縮技術
- 06 光で機械学習をスピードアップ 光リザーバーコンピューティングによる高速機械学習
- ◎ ネットワーク構造から深層学習のしくみを知る ニューラルネットの理解に向けたコミュニティ抽出技術

コミュニケーションと計算の科学

- 08 あなたの量子メモリをちょっと拝借! 未初期化量子ビットを利用した高速量子計算
- ① こわれにくいネットワークをデザインします 二分決定グラフを用いたネットワーク信頼性最大化 **③ いつでもどこでもそれっぽくしゃべれます!** スマホで音声リズムを英語母語話者っぽく変換
- 10 人工知能は文脈を読んで翻訳できるか? ニューラル翻訳の文脈理解度をテストする
- 111 ことばの発達がゆっくりなお子さんの特徴を探る 小児医療現場で収集したデータの解析からみえてきたこと
- 12 物知りロボットとおしゃべりしながら賢くなろう 複数ロボット対話制御に基づく雑談と質問応答の融合
- 14 離れていても盛り上がりを共有できる 双方向性の臨場感・一体感の向上をねらう拍手音符号化

メディアの科学

- **15 照明光で色の鮮やかさを操る** 彩度強調成分を用いた分光スペクトルの制御
- 16 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
- 17 二択問題にして解くことでAIは賢くなる 深層学習による仮説比較と音声認識結果選択への応用 21 選んで創るお気に入りの画像 DTLC-GANを用いた画像の階層的理解・生成制御
- 18 音だけから情景を推定 音から画像認識結果を予測するクロスメディア情景分析
- 19 浮像(うくぞう) 影を駆使して絵に奥行きを与える光投影技術
- 20 声の雰囲気や聞き取りやすさを変換する 深層生成モデルを用いた音声属性変換

人間の科学

- 22 音への注意を眼で測る 瞳孔反応に現れる聴覚空間注意
- 23 ウェルビーイングを測る、知る、育む "いきいきとした心的状態"の実現を科学する領域横断研究
- 24 人工知能で人の聴こえの仕組みを理解する 機械学習モデルによる聴覚神経機構の分析
- **25** 一流打者はボールをどのように打っているか? 打撃中の身体運動計測から認知過程を探る
- 26 一流打者はボールをどのように見ているか? 打撃中の眼球運動計測から認知過程を探る
- 27 急ぐ方が正確? 視知覚と運動に潜む情報処理のからくり
- 28 力で感じるかたちとうごき ぶるなび4による環境・状況の呈示
- 29 平らなシートなのに凹凸感?! 磁性シートに凹凸感を"書き込む"磁性触覚印刷技術



〒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: 077 4-93-5020 FAX: 077 4-93-50 15

