

NTT コミュニケーション科学基礎研究所

オープンハウス 2020

on the Web



6 / 4 (木) 12:00より
web 公開

本年度はNTT京阪奈ビルでの開催は行わず
最新の研究成果をwebにて公開いたします。
<http://www.kecl.ntt.co.jp/openhouse/2020/>

随時最新情報を公開いたしますので、是非ご覧ください。



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2020」へようこそ



NTT コミュニケーション科学基礎研究所
所長 山田 武士

新型コロナウイルス感染症への対策として、「オープンハウス2020」を実会場でのイベントとして開催することは断念し、最新成果のウェブ公開、という形をとることにいたしました。

1985年に電電公社が民営化されたNTTの発足から35年、1991年に当研究所が「NTT コミュニケーション科学研究所」として京阪奈に設立されてから30年が経とうとしています。

私達は設立当初から、コミュニケーションの本質は、「情報を正確かつ効率良く伝達すること」のみならず、「お互いに理解を深め、感動を共有し、心のふれあいを実現すること」との理念のもと、時代を先取りした基礎研究に取り組んで参りました。

人との接触を大幅に減らさざるを得ない今日状況において、コミュニケーションも希薄になりがちなか中、コミュニケーションにおける心のふれあいの本質を見極めることはなおさら重要な課題と言えます。

私達はこの理念を引き継ぎ、「ここまで伝わる」コミュニケーションの実現と、今後の「新しい生活様式」を見据えた心豊かな社会のデザインのために、これからも更に技術の糸を紡いでいく所存です。

「オープンハウス2020」では、機械学習や音声音響処理から人間の錯覚やスポーツ脳科学まで、ウェブ上の講演と展示によって、最新の研究成果をわかりやすくご紹介いたします。私達のめざす未来の手応えを、ぜひ間近に体感していただければ幸いです。

ごあいさつ 1
所長講演
あなたを・もっと・知りたくて ~AIで人に迫り脳科学で人を究める~ 3
招待講演
<個性>を科学するためのチャレンジ 5
研究講演
ご所望の声でコミュニケーション ~深層生成モデルが切り拓く音声変換の可能性~ 7
知覚心理学で探る皮膚感覚のしくみ ~指先の時空間情報処理~ 9
言葉の難しさを測る ~テキストの難易度と人の語彙数の推定~ 11
研究展示 13
データと学習の科学
01 WWW 上のみんな、オラに力を分けてくれ! WWW 上のリソースを活用した機械学習用データ作成手法 15
02 システム障害を早期に解決する方法を見つけます ニューラルネットを用いた障害復旧コマンドの生成 16
03 都市における空間集約データの高解像度化 空間集約データを補間する多変量ガウス過程 17
04 データに適した異常検知器を高速に生成します 未知データセットのための転移異常検知法 18
05 低い誤検知率で異常を検知 部分 AUC 最大化のための半教師あり学習 19
06 そのデータ、本当に偏ってますか? 決定グラフを用いた組合せ的相関検定 20
コミュニケーションと計算の科学
07 みんなが協力せず勝手に急ぐとどうなる? 混雑ゲームの均衡計算 21
08 小さな窓から量子世界の全てをコントロール 量子系の間接的制御が持つ可能性を探る 22
09 少量の追加データで作るカスタム機械翻訳 汎用対訳コーパス JParaCrawl を用いた機械翻訳の領域適応 23
10 こどもの感情発達レベルを測ります 表情・文脈・音声テストによる感情発達プロセスの解明 24
11 こどもの興味と発達に合わせて絵本を作ります パーソナル知育絵本を用いた親子の絵本読み活動支援の試み 25
12 あなたの語彙数測ります 令和版語彙数推定テスト 26
13 京町セイカがご案内! 地域連携で作るなりきり AI 27
14 こんなとき、あの人だったらどう思う? 人の個性を考慮した体験に紐づく感想生成技術 28
メディアの科学
15 この声、何歳? 話者クラスタリングを用いた深層話者属性推定 29
16 ワイヤレスマイクを同時により多く使えます ビット誤りに頑健で低遅延な音声音響符号化方式 BRAVE 30
17 聞きたい人の声に耳を傾けるコンピュータ (II) 音声と映像を手がかりとしたマルチモーダル選択的聴取 31
18 顔で声の表情を制御する クロスモーダル音声表情変換 32
19 探し方を学びながら探す 適応的スポットティング法による効率的な物体探索 33
20 データを端末から漏洩させない分散深層学習 分散 NW 上で機械学習をするための非同期合意形成技術 34
21 心臓らしい心臓モデル 物理法則拘束付きガウス過程回帰を用いた心臓のモデル 35
22 あなたの鼓動に耳を澄ます 音響観測に基づく血流動態の解析 36
人間の科学
23 ヒト知覚モデルで「自然な」錯覚をつくる 「不自然さ」予測に基づく「変幻灯」の視覚運動量最適化 37
24 微小な眼球運動から垣間見る認知状態 眼球運動の動特性と認知タスク・瞳孔径の関係 38
25 触ると似てしまうテクスチャ 3D プリンタを用いた触り心地を変えないテクスチャ変調 39
26 情動はいつ変化するのか? どうやって測るのか? 実験室環境と日常生活における情動変化の計測 40
27 e スポーツ達人の脳力 パフォーマンス、身体の生理状態、脳活動の相互関係 41
28 ラグビースクラムのハーモニーを紡ぐ ウェアラブルセンサを用いた選手間協調の簡便な評価 42
29 ストレートは“まっすぐ”か? 物理計測と知覚計測からピッチングを捉えなおす 43
30 巧みで素早い運動を支える脳の中の身体表現 手の位置推定の不確かさは伸張反射を調節する 44
31 意識より賢い無意識 環境に応じた顕在・潜在的視覚運動応答の調節 45

あなたを・もっと・知りたくて

～AIで人に迫り脳科学で人を究める～

I want to know more about you

— Getting closer to humans with AI and brain science —

NTT コミュニケーション科学基礎研究所 所長

山田 武士

Takeshi Yamada

1985年に電電公社が民営化され、NTTが発足してからすでに35年が経過しました。一方で、NTT コミュニケーション科学基礎研究所（以降、CS研）が京阪奈に設立されたのが1991年ですから、来年で設立30年を迎えようとしています。

私達は設立当初から、コミュニケーションの本質は、「情報を正確かつ効率良く伝達すること」のみならず、「お互いに理解を深め、感動を共有し、心のふれあいを実現すること」であるとの理念のもと、時代を先取りした基礎研究に取り組んできました。当初は人と人とのコミュニケーションが主題でしたが、現在は、人と人のみならず、人とコンピュータとの間の「ここまで伝わる」コミュニケーションの実現をめざし、基礎理論の構築と革新技術の創出に取り組んでおります。

1985年当時、電話は、当時の流行歌の歌詞にあるように、「親しい人が「いまどこにいるのか」とか、「今何をしているのか」などを知るためのコミュニケーションツールの主役でした[1]。現在はソーシャルメディアが新たな主役として発達し、あまり親しくない人であっても、それがある程度分かっています。そもそも、個人が日々利用するスマートフォンは、これらの情報をすべて把握しており、むしろ使用者本人よりも詳しいかもしれません。一方、黒電話には不思議な存在感とぬくもりがありました。今後さらに技術が発達すると、コミュニケーションはどう変化するのか？「心まで伝わる」コミュニケーションとは何か？私達はこの問いに答えるべく日々研究を続け、パートナーのみならず、世の中にも問いかけています。

CS研の研究領域

CS研の研究領域は、「メディア処理」「データ・機械学習」など、「人間の能力に迫り凌駕する」ための技術の追

求と、「人間科学」「多様脳科学」など、「人間を深く理解する」ためのサイエンスの追求からなっています[2]。以下に、いくつかの研究例をご紹介します。

人間の能力に迫り凌駕する技術

コミュニケーションの基本はまず話し言葉を認識し理解することです。人間は、複数人が同時に話す状況でも、聞きたい人の声に集中し、その声を聞き取ることができます。CS研では、人間が持つそうした能力をコンピュータ上で実現することをめざしています。最近では声の特徴に加え、唇の動きを手がかりとすることで、似通った声の人でも聞き分けられるようになってきました。また、最新の音声変換技術では、人が話した音声の内容は保持したまま、声質や抑揚などの特徴を自由に変えることができます。これらの技術が発展すれば、発声・聴覚機能の障がいや加齢による衰えを克服した自然なコミュニケーションや、更には、不慣れな外国語での会話のサポートなどが実現します。

人間は、街なかでふと流れてきた音楽の短い断片から、知っている曲であれば、その曲名を当てることができます。CS研ではこれまで、音や映像の信号の断片を手がかりに、楽曲や映像の膨大なデータベースの中から一致するものを高速に見つけ出す「ロバストメディア探索」技術を研究開発してきました。この技術はNTTデータを通じて実用化され、放送番組で使用された楽曲を自動検出し、楽曲権利処理に必要な使用楽曲リストを生成するサービスとして、数多くの放送局で利用されています[3]。最近では実空間での物体の探索にも取り組んでいます。例えば「適応的スポットティング法」は、実空間の3次元点群データから所望の形状の物体を素早く探し出す技術であり、人間が行うのと同じように、効率的な探し方を自ら学習することができます。

私達は数年前から国立情報学研究所の人工知能プロジェクト「ロボットは東大に入れるか」の一環として、人間が実際に解く問題を人工知能がどこまで解けるのかを明らかにする研究を進めています。特にCS研は英語を担当し、大学入試センター試験の英語筆記科目に挑戦してきました。その結果、2019年センター試験の英語筆記本試験において、185点（偏差値64.1）という極めて高い成績を達成できました[4]。英語問題は、自然言語処理、知識処理の統合的な問題を多く含んでおり、得られた知見は、AIと人間のより自然な、よりわかり合える対話の実現に活かすことができます。

人間を深く理解し究めるサイエンス

一方で、AIの発展により、人間のことを深く知り、理解することは、ますます重要になってきました。例えば、インターネットで調べものをしている最中に、検索ワードにマッチした商品広告がさりげなく提示されると、いつのまにかクリックして購入してしまう、という事例は良く起こります。その際ユーザは、あくまでも自分の意志でその商品を購入したのであって、第三者に操られて商品を買わされたとは、あまり自覚しません。今後AI技術が発展すると、より巧みな、いわばAI版「サブリミナル効果」のリスクが高まると考えられます。

このようなリスクを防ぐためにも、人間がいつ、どのような思い込みを持つのか、そしてその思い込みがどう行動につながるかを、深く理解することが重要です。CS研では、特に卓越したスキルを持つアスリートに着目して、アスリートの身体に現れる様々な生体情報から、彼らはどのように情報を捉えて判断しているのかなど、脳情報処理の解明に取り組んでいます。例えば野球選手において、打てる打者と打てない打者は何が違うのか？「ボールを良く見て打つ」というのは本当か？ストレートは本当に「まっすぐ」か？などです。これらの知見に基づき、脳機能を鍛えるトレーニング手法としてアスリートにフィードバックすること

もめざしています。

幼児の言語獲得過程の解明も進めています。人間の幼児は親とのコミュニケーションを通じて言語を習得します。人類は長い時間をかけて言語とそれを用いたコミュニケーションを進化させてきました。しかしながら、人類が文字を使うようになったのは比較的最近のことで、「読む」能力はもともと脳に備わった機能ではありません。「視覚」「聴覚」「言語」「認知」など脳の基本機能を柔軟に組み合わせることで実現しているのです[5]。

私達は、言語獲得の仕組みを理解するために、こどもがいつどんな語を理解し、発話できるかを大規模調査し、モデル化した「幼児語彙発達データベース」を構築しています。このデータベースは、こどもの「読む」能力の育成にも役立つと考えています。そこでNTT印刷を通じてサービス化されたのが、お子さま一人ひとりのことばの成長に合わせて内容をカスタマイズする、「パーソナル知育絵本」です。私達とNTT印刷は、沖縄県恩納村と協力して、検診に来た親子に、パーソナル知育絵本を作成してもらうことをきっかけに子どもを図書館に誘導することを始めています。これは、こどもが早い段階から図書館に通い、絵本を読むことを習慣化させようという取り組みです[6]。

おわりに

CS研は、人の能力に迫る研究、人を究める研究に取り組み、「ここまで伝わる」理想のコミュニケーションとは何かを探求しています。ちなみに日本語の「幸せ」の語源は、「為合わせ」や「仕合わせ」であって、これはまさに、他者とのインタラクションやコミュニケーションが「うまく行った」状況をさしているとも言えるそうです。私達はこれからも人々の幸福、最近の言葉で言えばウェルビーイングの向上に資する技術を創出し、パートナーの皆さまとのコラボレーションを通じて、心豊かな社会の実現へとつなげたいと考えています。

▼関連文献

- [1] 「あなたを・もっと・知りたくて（作詞：松本隆）」は、1985年、NTT民営化直後の「TALK ON THE PHONE」イメージソング
- [2] 人に迫り、人を究め、人に寄り添うーデジタルとナチュラルの共生・共創に向けて NTT技術ジャーナル 2019.9, <https://www.ntt.co.jp/journal/1909/files/pdf/JN20190906.pdf>
- [3] 第9回 ASPICクラウドアワード2015において、NTTデータが「全曲報告サービス」で総務大臣賞を受賞 <https://www.nttdata.com/jp/ja/news/information/2015/100901/>
- [4] 2019年大学入試センター試験英語筆記科目においてAIが185点を獲得! <https://www.ntt.co.jp/news/2019/1911/191118a.html>
- [5] Reading in the Brain: The New Science of How We Read, Stanislas Dehaene, Penguin Putnam Inc, 2010
- [6] パーソナル知育絵本を用いた3者共同実験を開始しました! <http://www.nttprint.com/company/itemid419-000048.html>

<個性>を科学するためのチャレンジ

A challenge to scientifically understand "individuality"



東北大学 副学長・大学院医学系研究科 教授

大隅 典子

Noriko Osumi

プロフィール

1985年 東京医科歯科大学歯学部卒
 1989年 同大学院歯学研究科修了 歯学博士
 1989年 同大学歯学部助手
 1996年 国立精神・神経センター神経研究所室長
 1998年 東北大学大学院医学系研究科教授
 2018年 東北大学副学長・附属図書館長
 この間、CREST研究代表者(2004～2009)、東北大学グローバルCOE拠点リーダー(2007～2012)など歴任。現在、新学術領域「個性」創発脳(2016～)、領域代表。

専門の内容

発生発達神経科学、脳の発生発達維持の分子機構、神経新生低下と精神疾患発症の関わり、脳の健やかな発生発達維持のための栄養、自閉症の病因病態に関する神経生物学的検証

著書

著書に、『脳の誕生 -発生・発達・進化の謎を解く』(ちくま新書、2017)、
 『脳からみた自閉症 -「障害」と「個性」のあいだ』(ブルーバックス、2016)など

2003年にヒットしたSMAPの「世界に一つだけの花」のサビの部分は、以下のような歌詞となっている。

『世界に一つだけの花
 一人一人違う種を持つ
 その花を咲かせることだけに
 一生懸命になればいい』

私たちは皆、それぞれ異なる<個性>を持つ。それはいったい、どのようにして獲得されるのだろうか？

命は、たった1個の受精卵として始まる。父親の精子と母親の卵子が受精したときに、「個」の情報が規定される。それは、DNAのACTGの文字として書き込まれた「ゲノム」という情報、そしてゲノムの“上書き”として付け加えられた「エピゲノム」という情報だ(「エピ」は「上」や「後」という意味の接頭語)。

エピゲノムは、DNAそのものや、DNAの鎖が巻き取られるヒストンというタンパク質の化学修飾に基づくが、ちょうど句読点等の“記述記号”のように、その有りなしでゲノムの言葉の意味が変わってしまうような働きがある。ゲノム情報は基本的に生涯、変わらないが、エピゲノム情報は、生活習慣や経験により変わりうるものであり、遺伝子の“働き方”を変化させる。

この問題に実証的に迫ろうと、5年前に「多様な<個性>を創発する脳システムの統合的理解」という研究グループ

を立ち上げた。私たちのグループはとくに、天才や自閉スペクトラム症などの神経発達障がいの方が「非定型発達」を示すということに着目している。そのような「非定型」さ(それは「個性」と捉えることもできる)がどうやって生じるのかについて、ヒトの疫学的研究より示唆される「父親の加齢」による次世代への影響について、精子エピゲノム情報の変化による影響を明らかにしつつある。

私たちはマウスをモデル動物として研究している。実験動物であるマウスは、“遺伝的に均一”であることが大きなメリットだ。つまり、ゲノム情報が同じであるにも関わらず、<個性>が認められるということは、種々のエピゲノム情報が異なることを示している。

これまでの研究成果により、私たちは確かに父加齢により非定型発達を示す仔マウスが多くなることを見出し、精巣内で精子が形成される過程においてどのようなエピゲノム変化(これを私たちはこれを「エピ変異」と名付けている)が生じるかを示しつつある。

このような<個性>を科学するためのチャレンジは、生物学的には有性生殖における卵子側(=命自体の継承)と精子側(=多様性の付与)の役割を浮かび上がらせるものであるとともに、近年、なぜ発達障がいが増えているのかについての考察として、社会科学的に大きなインパクトがあると考えられる。

▼関連文献

『』内:出典「世界に一つだけの花」(作詞:横原敬之)
 大隅典子『脳からみた自閉症 「障害」と「個性」のあいだ』(講談社ブルーバックス、2016)
 大隅典子『脳の誕生 発生・発達・進化の謎を解く』(ちくま新書、2017)

ご希望の声でコミュニケーション

～深層生成モデルが切り拓く音声変換の可能性～

Communication with desired voice

— Deep generative model opens the way to innovative speech transformation —



メディア情報研究部

田中 宏 Kou Tanaka

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 研究員。2017年奈良先端科学技術大学院大学 情報科学研究科 博士課程修了。博士(工学)。同年NTTに入社以来、音声合成・音声変換の研究に従事。深層生成モデルを用いた音声信号処理に特に興味を持つ。奈良先端科学技術大学院大学優秀学生賞や日本音響学会第47回栗屋潔学術奨励賞を受賞。日本音響学会の会員。

音声は、言語情報だけでなく話者性などの非言語情報も伝達できるという大きな特徴を有しており、人々がお互いにコミュニケーションを取るうえで特に重要なツールの1つとなっています。発話することで、自分/相手の意図や感情を、伝える/理解することができるため、音声の特徴(例えば、抑揚や声質・リズム)をその時々で変化させることで、相手に与える印象を変えることもできます。しかしながら、一個人の生成できる音声の表現力は身体的・能力的・心理的制約により制限されてしまいます。この制約を超え、発話者が所望の音声で思いのままに表現できるよう能力の拡張を行う技術が音声変換です。その適用先は、話者性の変換や発声障がい者補助、感情などの発話スタイル変換、語学学習のための発音/アクセント変換など、多岐にわたります。これらの利用シーン

に応じて、変換したい音声特徴・学習データ・リアルタイム性に関する要件など、様々な前提条件が想定されます。私たちは、高品質であること、少量データ・非パラレルデータ*で学習可能であり効率的であること、リアルタイムに音声変換が動作すること、声質だけでなく超分節の特徴などの柔軟な変換が可能であること、上記の4点が音声変換において重要な要件であると考えています。

従来技術において代表的なものは、混合ガウス分布に基づく統計的声質変換 [1] です。事前に時間整合をとった、入力音声と目標音声の同一発話文(パラレルデータ)を用意し、この両者の特徴量の同時確率を最大化することで、前者から後者への変換関数を求めます。また、近年では、上述のパラレルデータを必要とする枠組みにおいて、性能改善のため、

ニューラルネットワークを用いた手法や非負値行列因子分解などを用いた事例ベースの手法の検討も進められています。しかしながら、これら従来技術には、1) 学習データとして同一発話内容の音声ペアが必要であること、2) 変換可能な音声特徴が声質に限定されること、と技術的制約があります。さらに、音声の特徴量から波形を合成する際に古典的なボコーダを用いているため、不自然でいかにも合成音といった音が生成されてしまいます。

一方、深層学習界隈において、2014年を皮切りに、画像処理や自然言語処理の分野から、非常に興味深いモデルが台頭してきました。変分自己符号化器 (Variational Auto-Encoder: VAE) や敵対的生成モデル (Generative Adversarial Networks: GANs)、系列変換モデル (Sequence-to-Sequence model: Seq2Seq) といったモデルです。Seq2Seqのモジュールの一つである自己再帰型モデル (Auto-Regressive model: AR) とVAE、GANsを総称して三大深層生成モデルと呼ばれることもあり、画像処理や自然言語処理など様々な分野・タスクでその有効性が確認されています。また、2015年中期の機械翻訳タスクにて、注意機構 (Attention mechanism) がニューラルネット

に導入され、その高い有効性が瞬刻間に注目を浴びました。

音声変換の入力・出力はどちらも音声信号(連続値で構成された時系列データ)であることを意識しつつこれらの深層学習技術をうまく拡張することで、従来音声変換技術の課題を克服し様々な利用シーンに柔軟に対応可能な多用途音声変換システムを実現します。私たちの音声変換システムでは、1) 非パラレルデータを用いて声質変換を行える「非パラレル声質変換機能 [2]」、2) 声質だけでなく長期依存特徴である韻律やアクセントの変換を行える「音声系列変換機能 [3]」などの新機能を創出しました。また、合成音声波形から実音声波形へ波形空間上で深層学習を用いて変換を行い、出力音声の高音質化を実現する「波形ポストフィルター機能」を世界で初めて実現しました。結果として、人の声と聴き分けられないほど高品質の音声変換や任意の目標話者へ変換を行うためのモデルの多対多化、非パラレルデータを用いることができるという高効率化、リアルタイム化を実現しています。また、近年盛んに研究されている分野横断型の研究として、目標話者の顔画像を用いてその顔らしい音声に変換する「クロスモーダル音声変換機能」も実現しています。

*入力音声と目標音声とで発話内容が異なるデータ(非同一発話文)を示す。

▼関連文献

- [1] T. Toda, A. W. Black, K. Tokuda, "Voice conversion based on maximum likelihood estimation of spectral parameter trajectory," *IEEE Trans TASLP*, Vol. 15, No. 8, pp. 2222-2235, 2007.
- [2] T. Kaneko, H. Kameoka, K. Tanaka, N. Hojo, "CycleGAN-VC2: Improved CycleGAN-based non-parallel voice conversion," in *Proc. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
- [3] K. Tanaka, H. Kameoka, T. Kaneko, N. Hojo, "ATTS2S-VC: Sequence-to-sequence voice conversion with attention and context preservation mechanisms," in *Proc. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.

知覚心理学で探る皮膚感覚のしくみ

～指先の時空間情報処理～

Towards understanding human skin sensations

— How the brain integrates spatio-temporal information in touch —



人間情報研究部

黒木 忍 Scinob Kuroki

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。2011年東京大学大学院 情報理工学系 博士課程修了。博士(情報理工学)。同年NTTに入社以来、触覚におけるタイミング、周波数、位相、運動、粗さなど、基礎的な知覚について研究。

触覚 × 知覚心理学

手は道具でありセンサです。手指を巧みに使う事で、人間は上手に進化を遂げてきたと思います。道具をうまく使うには速くて正しい計測と処理が有効です。私はこの、皮膚を通じた情報処理に興味を持っています。「手に取るようにわかる」「肌で感じる」など、触覚に関わる比喩表現には、ものごとを深く理解する様子をイメージさせるものが多くあります。ところが、触覚が実際に何をやっているのか、その情報処理に関する研究は意外なほど疎で、その処理のメカニズムはあまり良く解っていません。自己と世界の境界から得られる感覚であるため、その接触面で起きる現象を直接計測したり、人工的に再現したりすることが難しいなど、技術的な課題もありました。近年の技術進歩に伴い、視覚や聴覚の研究を追いかける形で、少しずつ触覚についても解明が進んでいます。

知覚心理学とは、人間をシステムとして扱い、その内部処理を同定する学問です。人間に対して、何か画像や音のような入力を加えたときに、出力としてどのような知覚が得られるのかを調

べます。例えば、明かりをゆっくり点けたり消したりすると明るさの変化を知覚することが出来ますが、切り替え速度を上げていくと明滅は徐々に見えにくくなり、蛍光灯ほど高速になると変化は検出出来ません。このように、少しずつ入力を変化させながら、私たちの知覚に影響を及ぼす物理量や、その範囲を調べていきます。

時空間情報処理の大切さ

皮膚にはセンサである機械受容器がアレイ状に敷き詰められています。指先でモノに触れると、これらのセンサが局所的に反応してそれぞれ神経信号を脳へ送ります。例えば硬い物体に触れた場合、皮膚の変形は一定量で止まりますが、指よりも柔らかい物体に触れた場合には、接触後に指が物体にめり込むため、指が包み込まれるかたちで接触面積が増えてゆきます(図1)。こうした時空間的な変化を捉えることが、接触対象を正しく推定するためには必要です。

形・応答・処理の違い

網膜に複数のセンサがあり、それぞれ異なる色に対して感度を持つように、皮膚にあるセンサにも幾つかの種類があります。それぞれのセンサは、ある種類は小粒で皮膚表面に密集し、また別の大きなものは皮膚深くにぼつぼつと点在するなど、異なる形と分布を持っています。その結果、ゆっくりした変形に反応するセンサや、着信パイプレーションのような高周波の振動に良く反応するセンサといった具合に、種類ごとに個性が生まれます(図2)。こうした、異なるセンサ系からの信号は、脳において個別に処理されるのでしょうか、まとめて処理されるのでしょうか?指先で実際の物体に触れる日常的な状況では複数のセンサ系が応答しているため、脳内の処理についてセンサ系ごとに切り分けて調べることは困難でした。この講演では、正弦波振動などの非常にシンプルな入力を用いることで、センサ系の応答割合を制御し、各系の時間情報が統合される過程について調べた研究をご紹介します[1]。

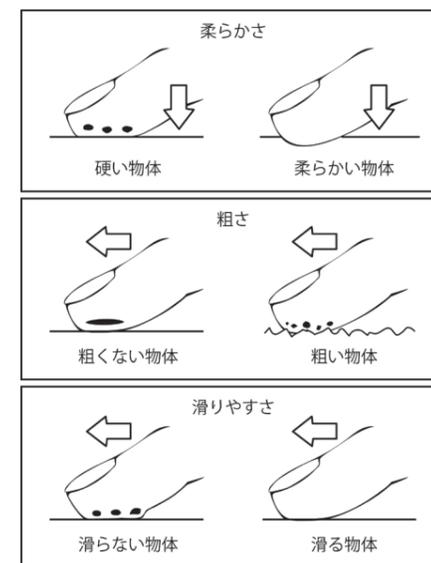


図1 接触対象の物性が皮膚表面を時空間的に変形させる

より複雑な入力・処理・知覚へ

触覚の知覚研究では、振動などの素朴な入力だけでなく、布や金属、毛皮やスポンジといった日常素材もしばしば入力として使われます。日頃触れる素材が引き起こす感覚について調べるわけですから、いかにも多彩な結果が得られそうです。しかし一方で、入力のバリエーションが複雑になるにつれ、知覚変化の鍵を理解することは難しくなってきます。例えば「布Aと石Bの区別はつきのに、布Aと金属Cは区別がつかない」という結果が得られた場合、石と金属の物理的特徴の違いが多岐に渡るため、原因を絞りこむのが難しいのです。なお、最近では複雑な入力と出力の関係をビッグデータ解析することで紐解く研究が増えています。とても頼もしいのですが、触覚については大きな触覚データセットを用意するのが困難なこともあり、未だ大きな成果は出ていません。

我々のグループでは、単純な実験刺激と複雑な現実世界の間を繋ぐ試みとして、レーザーカッターや3Dプリンタを使い、複雑な空間パターンを持つ表面テクスチャを作成しています。中心周波数(紙やすりで言うところの番手)のような、単純な統計量だけでなく、より複雑な統計量の操作を手軽に行うことが可能になりました。脳内で、どの程度まで複雑な統計量が計算されているのかを調べた研究[2]について、本講演でご紹介するとともに、展示25番「触ると似てしまうテクスチャ」にて、詳細にご説明いたします。こうした時空間情報処理やセンサ間の情報統合の仕組みに関する理解を積み重ねていくことが、触覚の理解につながると期待されます。

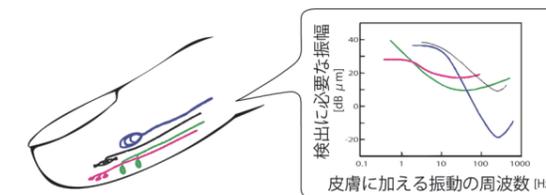


図2 指先皮膚に存在する複数のセンサと、それらに繋がる神経の応答特性

▼ 関連文献

[1] S. Kuroki, J. Watanabe, S. Nishida, "Integration of vibrotactile frequency information beyond the mechanoreceptor channel and somatotopy," *Scientific Reports*, Vol. 7, 2758, 2017.
 [2] S. Kuroki, S. Sawayama, S. Nishida, "Haptic metameric textures," *bioRxiv*, 2019. doi: <https://doi.org/10.1101/653550>.

言葉の難しさを測る

～テキストの難易度と人の語彙数の推定～

Which word is more difficult for you, “car” or “vehicle” ?

— Estimation of text readability and human vocabulary size —



協創情報研究部

藤田 早苗 Sanae Fujita

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 主任研究員。1999年奈良先端科学技術大学院大学 情報科学研究科 修士課程修了。同年、NTTに入社。博士(工学)。自然言語処理研究に従事。言語処理学会、情報処理学会各会員。3人の子どもの成長とともに、絵本の難易度推定、学齢期の語彙数調査、英語学習支援と研究テーマを広げています。

文字を覚えてたの子どもが自分で選んだ絵本が読めず、読んであげることになったことはありませんか。中1の時とても苦労して読んだ英文が、大学生になる頃にはとても簡単に感じられたことはありませんか。同じ文を読もうとしても、難しいと感じるか易しいと感じるかは、読み手の知識量に依存します。もし、読み手にとってちょうど読めるくらいの、あるいは少し頑張れば読めるくらいの絵本や本、英文を薦めることができれば、読み手の知識を無理なく増やしていけるかもしれません。しかし、「ちょうど良い難しさ」を判断するのは簡単ではありません。文(テキスト)側の難易度と、人側の知識量の両方を適切に推定する必要があるからです。

人の語彙数調査と推定方法

人側の知識の一つとして語彙力があります。NTTでは20年以上前から、人の語彙数の調査や推定に取り組んできました。

幼児を対象とする場合、理解/発話できるすべての語彙を調査することも不可能ではありません。実際私たちは、1500組以上

の親子モニターの皆さんにご協力をいただき、子どもがいつごろどのような語を覚えるか、発話できるかというデータを蓄積し、幼児語彙発達データベースを構築してきました。しかし小学生以上となると、知っているすべての語彙を調査することは困難です。そこで、提示した語を知っているか回答してもらうことにより、語彙数を推定します。提示する語は多いほど正確な推定ができますが、数十語でも推定可能です。この方法では、ある語を知っていると回答したときに、何語知っているか推定するかがポイントとなります。例えば「銀行」と「地歩」だと、「地歩」の方が知っている人は少ないでしょう。そのため、「銀行」だけ知っている人より「地歩」も知っている人の方が語彙数は多いと思われるのでしょうか？

その推定の根拠となるのが、語のなじみ深さを数値化した「単語親密度」です。NTTでは20年以上前に約7万7千語の単語親密度を調査、それを元に公開した語彙数推定テストは多くの方に利用されてきました。しかし、調査から20年以上が経過したこと

から、新しい語の追加と再調査を実施、16万3千語というより大規模な単語親密度データベースを再構築しました[2]。

データベースを元に令和版語彙数推定テストを作成、4千人以上を対象に調査を行ったところ、小学6年生で約2万語、大人で約5万語の語彙数があることがわかりました(図1)。また、同じ学年でも生徒によって語彙数に大きなばらつきがあることから、支援が必要な生徒を見つけることにも役立つと考えています[3]。

テキストの難易度推定

テキスト側の難易度については、まず絵本の難易度推定から取り組みました。本研究所の主要テーマでもある幼児の語彙発達の解明、発達支援などにも寄与できると考えたからです。しかし、絵本には電子データ化されたデータベース(コーパス)自体が存在せず、その構築からスタートしました。人手でほとんどの本文テキストを入力するという地道な作業の結果、NTTの絵本コーパスは、日本語6000冊・英語2000冊を超える、世界に類をみない規模になりました。しかも今でも拡張中です。

さて、いよいよ難易度推定です。テキストの難易度には、語彙の難しさ、文構造の難しさの両方が影響しますので、それぞれを

適切に反映する特徴量を見出しました。加えて、絵本の場合、ひらがなを正しく解析することも必要です。例えば、「とうさん」が「父さん」か「倒産」かによって、難易度は大きく異なります。こうした特徴量選択やひらがな解析の工夫により、高精度な難易度推定を可能にしました[1]。また、この精度向上の工夫は、教科書など他分野のテキストの難易度推定にも有効であることが分かってきました。

今後の展開

今回ご紹介した「人の語彙数推定」と「テキストの難易度推定」を組み合わせ、定期的な語彙数の確認をすれば、一人ひとりがその時、「ちょうど読める・少し頑張れば読める」テキストを推薦することが可能になります。実際我々は、英語の語彙数推定と難易度推定の研究も進めており、語彙力にあった英語絵本を推薦して学校の英語教育に活かす取り組みも始めています(図2、[4])。

我々は今後も、日本語でも英語でも、幼児でも小中高校生でも、大人に対しても、エビデンスを積み重ねながら、一人ひとりにあった育児・教育支援の実現をめざしていきます。

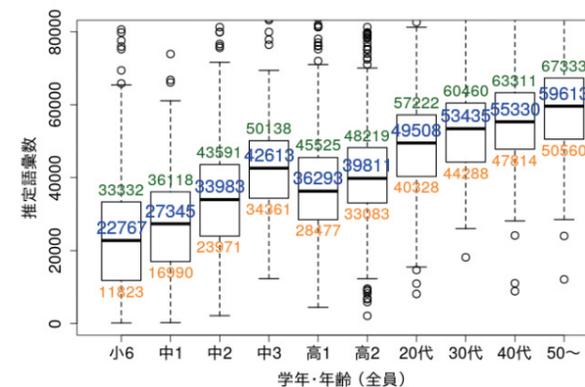


図1 小学6年生から成人を対象とした語彙数推定結果(約4600名)



図2 英語の語彙数推定画面

▼関連文献

- [1] 藤田早苗, 小林哲生, 南泰浩, 杉山弘晃, “幼児を対象としたテキストの対象年齢推定方法,” *認知科学*, Vol. 22, No. 4, pp. 604-620, 2015.
- [2] 藤田早苗, 小林哲生, “単語親密度の再調査と過去のデータとの比較,” *言語処理学会第26回年次大会 (NLP)*, 2020.
- [3] 藤田早苗, 小林哲生, 山田武士, 菅原真悟, 新井庭子, 新井紀子, “小・中・高校生の語彙数調査および単語親密度との関係分析,” *言語処理学会第26回年次大会 (NLP)*, 2020.
- [4] 藤田早苗, 服部正嗣, 小林哲生, 納谷太, “日本人初学者の語彙数推定方法の検討,” *2020年度人工知能学会全国大会 (JSAI)*, 2020.

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2020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31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。

- 膨大なデータを正確かつ詳細に分析したい
- 多様なリソースを活用して、データの解像度を高めたい
- 稀に発生する異常でも見逃さずに高速に検知したい

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・計算の限界に挑戦する研究を紹介します。

- 効率のよいネットワークやノイズに強い量子計算機を設計したい
- 子供の語彙獲得や感情発達に合わせた支援をしたい
- 対話システムや機械翻訳を状況に合わせて柔軟に対応させたい

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。

- 多様な環境で会話を楽しみ活用したい
- 各種メディアの本質をとらえ、自在に変換したい
- 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。

- 人の知覚特性を理解し、斬新で創造的な表現様式を提案したい
- 人の知覚と運動の関係を明らかにし、包括的な人間理解をめざしたい
- アスリート達がつもつ秀でた能力の秘密を解明したい

概要図

音声・音響

- 16 ワイヤレスマイクを同時により多く使えます
- 15 この声、何歳？
- 17 聞きたい人の声に耳を傾けるコンピュータ(II)
- 18 顔で声の表情を制御する

機械学習

- 02 システム障害を早期に解決する方法を見つけます
- 04 データに適した異常検知器を高速に生成します
- 05 低い誤検知率で異常を検知
- 03 都市における空間集約データの高精度化

異常検知

深層学習

- 20 データを端末から漏洩させない分散深層学習

探索

- 19 探し方を学びながら探す
- 07 みんなが協力せず勝手に急ぐとどうなる？

検定

- 06 そのデータ、本当に偏ってますか？

量子

- 08 小さな窓から量子世界の全てをコントロール

翻訳

- 09 少量の追加データで作るカスタム機械翻訳
- 12 あなたの語彙数測ります
- 13 京町セイカがご案内！

言語

- 11 こどもの興味と発達にあった絵本を作ります
- 10 こどもの感情発達レベルを測ります
- 14 こんなとき、あの人だったらどう思う？

対話・共感

- 26 情動はいつ変化するのか？ どうやって測るのか？

錯覚

- 23 ヒト知覚モデルで「自然な」錯覚をつくる

触覚

- 25 触ると似てしまうテクスチャ

視覚

- 24 微小な眼球運動から垣間見る認知状態
- 30 巧みで素早い運動を支える脳の中の身体表現

無意識

- 31 意識より賢い無意識

運動

- 27 eスポーツ達人の脳力
- 28 ラグビースクラムのハーモニーを紡ぐ
- 29 ストレートは“まっすぐ”か？

クラウドソーシング

- 01 WWW上のみんな、オラに力を分けてくれ！

決定グラフ

銀行
 担当
⋮
 親疎
 改錫

01 WWW上のみんな、オラに力を分けてくれ!

WWW上のリソースを活用した機械学習用データ作成手法

どんな研究

WWW上の多種多様、大量のデータに、WWW利用者の力を借りて機械学習用の正解ラベルを付与する仕組みを提供します。既存のWebブラウザにラベリング機能を組み込むことで簡単にラベリングができます。また、サーバと連携して様々な報酬を提供することで、利用者のラベリングを促進します。

どこが凄い

教師あり機械学習で利用する大量かつ高品質な正解ラベル付きデータを作成するには、通常膨大な人的、金銭的リソースが必要です。WWW上のデータ素材に対してWWWの利用者に自発的にラベリングしてもらうというアプローチをとることで、データの準備からラベリングまでを低コストに実現します。

めざす未来

WWW上のデータをWWWの利用者自身が整理していくことができる本技術は、いわばクラウドソーシングの新たな形態を実現します。本技術によって機械学習用の学習データを作成できるようになれば、機械学習を利用した多種多様なサービスをより安価に創出できるようになるでしょう。

AIを支えるのは大量かつ高品質な学習データ

素材の準備

- 良質な素材の選択には時間がかかる。購入は、高額なことも。

素材に対するラベリング

- 作業への報酬が必要。性能の維持には定期的なラベルの見直しも必要。

WWWを利用している人がWWW上のデータを整理(学習データ化)してくれないかなあ

World Wide Web (WWW)

世界最大のデータジェネレータ

- さまざまなデータが未整理状態で存在し、また日々更新されている。

第二の生活空間

- 多くの人が日常的にWWW上で活動している(潜在的な作業員)。

WWW横断型ヒューマンコンピューテーション
WWW上での活動にちょっとしたプラスアルファを組み込むことで、WWW上で人が活動していると様々なデータが自動的に整理されていく仕組み

まとめサポーター

複数人で協力して特定のトピックに関する画像を収集していくと、正解ラベルの付与された画像データセットができていく

画像を右クリックし、まとめサイトのトピックを選択

各画像の評価

トピックごとに画像をまとめたWebページが生成される

テキストモンスター

Webページに隠れている単語(テキモン)を捕まえ、Webサイトを奪い合うゲームをしていくと、単語に単語親密度が付与されていく

課題に沿ってページ中の単語を選択し、テキモンを捕獲

Webページ上に出現するテキモン(単語付きキャラクター)を選んで捕獲ゲームを開始

マルチボイスラベラー

小説などを複数の音声合成ボイスで読み上げるためのラベルを付与していくと、談話構造解析用正解データが作成されていく

話者情報(登場人物名や年齢、性別など)を入力

選択した話者に対応する文章を指定

すべての発話文を選択したら保存

関連文献

[1] Y. Shirai, Y. Kishino, Y. Yanagisawa, S. Mizutani, T. Suyama, "WWW横断型ヒューマンコンピューテーション," 第27回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2019), 2019.

[2] Y. Shirai, Y. Kishino, Y. Yanagisawa, S. Mizutani, T. Suyama, "Building human computation space on the WWW: labeling web contents through web browsers," in Proc. The seventh AAAI Conference on Human Computation and Crowdsourcing (HCOMP2019), 2019.

連絡先

白井 良成 (Yoshinari Shirai) 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



オープンハウス 2020

02 システム障害を早期に解決する方法を見つけます

ニューラルネットを用いた障害復旧コマンドの生成

どんな研究

システム障害の発生時に、その障害を復旧させるコマンド列を自動で推定し、その信頼度とともに運用者へ提示します。系列変換を行うニューラルネットワークの一種であるSeq2Seqを用い、障害時に取得できるアラームやログを、機械実行可能な復旧コマンド列に変換します。

どこが凄い

従来、復旧コマンドの検討は、運用者の知見に依存することが多く、運用品質のばらつきや復旧の長時間化が問題となっていました。運用者は、本技術が推定したコマンドの実行可否を、同時に提示される信頼度に基づき判断するだけで、多くの復旧作業を半自動で完了させることができます。

めざす未来

本技術を用いることで、障害復旧を迅速化し、お客様へのサービス影響を最小限に抑えることができます。今後は、推定された復旧コマンドが対象システムの性能や可用性に及ぼす影響も考慮し、コマンド実行可否の判断も自動化することで、システムの完全自動復旧の実現をめざします。

課題: 復旧シナリオ作成の困難性

- システム障害の復旧は運用者の知見に依存
- 復旧シナリオ(=ログ/アラームと復旧コマンドを紐づけるルール)の事前作成も多大な稼働が必要

IF case1 IDs={1,2} THEN []

IF case2 IDs={1,4}&{5} THEN []

IF case3 IDs={1}&{3,4} THEN []

THEN []

どんなログをトリガとするか

どんなコマンドを設定するか

技術イメージ: 復旧コマンド自動生成

入力: ログ/アラーム列

出力: 復旧コマンド列 & 信頼度

候補① (信頼度: 93%)

候補② 候補③

・Seq2Seqでログ列・アラーム列に対する復旧コマンド列を学習
→推定時は複数の候補を提示
・推定の信頼度 (= 実行の指標)も同時提示

手法: Seq2Seqによるログ-コマンド変換

Seq2Seq(系列間の関係性を学習するニューラルネットワーク)で、過去の復旧履歴を学習し、ログ/アラーム列を、コマンドの単語列に変換するモデルを構築

発生したログ

各ログをID化

ログID列 [... , 5, 12, 3, 7]

【出力】 復旧コマンド列

systemctl restart nova-api <Enter>

【入力】 ログ/アラーム

人手なしでコマンドを獲得

「Enter押下」等も一単語として、復旧完了まで一連のコマンドを出力

単語の生成確率の積として、信頼度を算出

ビームサーチにより、信頼度上位N位まで提示

関連文献

[1] H. Ikeuchi, A. Watanabe, T. Hirao, M. Morishita, M. Nishino, Y. Matsuo, K. Watanabe, "Recovery command generation towards automatic recovery in ICT systems by Seq2Seq learning," in Proc. of IEEE/IFIP Network Operations and Management Symposium (NOMS), 2020, to appear.

[2] T. Kimura, A. Watanabe, T. Toyono, K. Ishibashi, "Proactive failure detection learning generation patterns of large-scale network logs," IEICE Transactions on Communications, Vol. E102-B, No. 2, pp. 306-316, 2019.

連絡先

池内 光希 (Hiroki Ikeuchi) ネットワーク基盤技術研究所
Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

03 都市における空間集約データの高解像度化

空間集約データを補間する多変量ガウス過程

どんな研究

都市において取得されたデータ(貧困度や病気の感染率等)は、多くの場合、特定の領域分割(行政区画や管轄区域等)に従って集約されます。我々の技術は、**都市における様々な種類の空間集約データを同時に活用することにより、低解像度の空間集約データを高解像度化**します。

どこが凄い

空間集約データは集約がなされた領域分割の違いにより、様々な解像度を持ちます。**多変量ガウス過程に基づいてデータ集約処理をモデル化**することで、それぞれの解像度を持つ多様な空間集約データを有効活用し、低解像度の空間集約データから、任意の解像度のデータを予測できます。

めざす未来

例えば、貧困地域など重要な地域をピンポイントで知ることができるようになり、**都市環境改善に向けた効果的なフィードバック**につながります。今後は、大規模なデータ取得が難しい都市への対応等を行い、より幅広い場面で活用できる技術の実現をめざします。

問題設定：空間集約データの高解像度化

- ◆ タスク
多種多様なデータ群を活用して、低解像度データから高解像度データを推定
- ◆ 基本的なアイデア
空間パターンが類似した高解像度データを活用して、低解像度データを空間補間する
- ◆ 難しさ
データの種類によって紐づく領域分割が異なるので、単純にはデータセット間の類似度を評価できない

提案技術：データ集約処理を組み込んだ多変量ガウス過程

複数の空間集約データの生成プロセスをモデル化し、観測に基づいて高解像度データを自動的に推定

- ◆ ポイント①: 空間補間
データの背後に滑らかな関数(ガウス過程)を想定
- ◆ ポイント②: データセット間の関係性の導入
空間パターンの類似度(矢印の太さ)に基づいて、データセット間において関数 $g_i(x)$ を共有
- ◆ ポイント③: データ集約処理のモデル化
高解像度データ(関数)を各領域において平均化

関連文献

[1] Y. Tanaka, T. Tanaka, T. Iwata, T. Kurashima, M. Okawa, Y. Akagi, H. Toda, "Spatially aggregated Gaussian processes with multivariate areal outputs," in *Proc. 33rd Conference on Neural Information Processing Systems (NeurIPS)*, pp. 3000-3010, 2019.

連絡先

田中 佑典 (Yusuke Tanaka) サービスエポリューション研究所
Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

04 データに適した異常検知器を高速に生成します

未知データセットのための転移異常検知法

どんな研究

異常検知は正常なデータとは性質が異なるデータ(異常)を見つけ出すタスクです。高精度な異常検知器を作るためには正常/異常の両方のデータが必要です。本研究では、**正常データしか得られない状況でも、関連するデータセットを用いることで適切な異常検知器を高速に生成**する技術を作りました。

どこが凄い

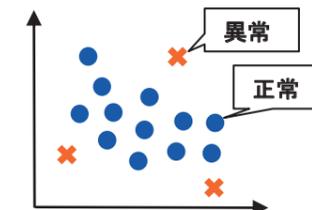
正常データのみから高精度な異常検知器を作ることは一般に困難です。本技術では、**関連データセットの情報を活用することで高精度な異常検知器を生成**します。とくに、特殊なニューラルネットモデルを導入することで、**計算コストの高い処理(再学習)なしに適切な異常検知器を生成**できます。

めざす未来

IoTの発展により多種多様なデータが得られるようになってきています。これらのデータをAI(人工知能)に利活用することで、**特別なドメイン知識がなくとも高精度な異常検知器を自動生成**することが期待できます。今後は本技術を含むAI技術を発展させ、**より安心・安全な社会の実現**をめざします。

異常検知とは

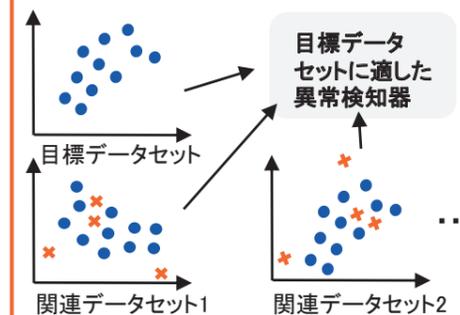
正常データと、性質が異なるデータ(異常)を見つけ出すタスク



- 異常/正常データを用いれば高精度な異常検知器を学習可
- しかし**異常データは希少なため学習に利用できない**ケースが多い

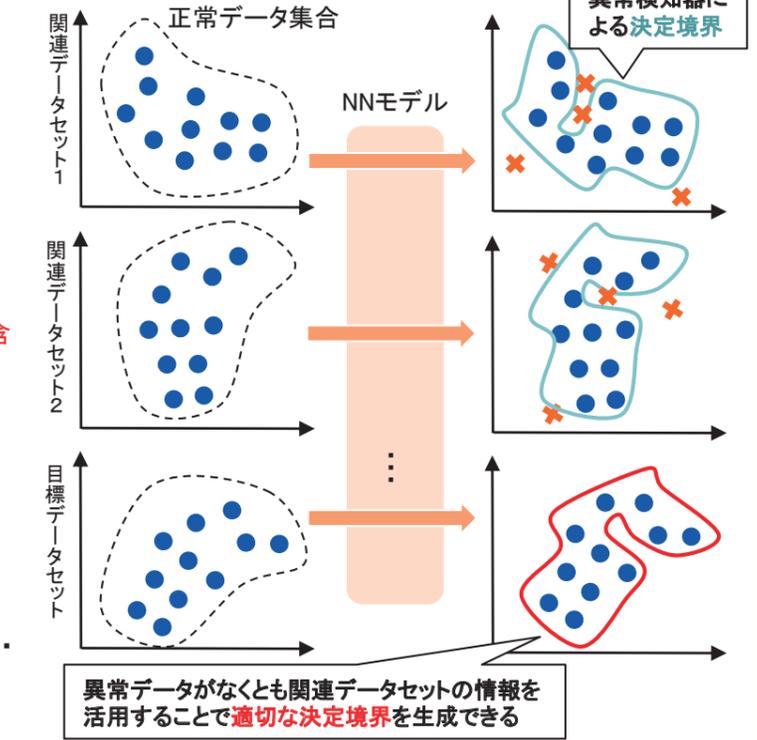
問題へのアプローチ

正常データセットに加えて**異常/正常を含む複数の関連データセットも活用**
例) 画像データ: 撮影環境が異なる画像など



提案法

- 正常データの集合を入力すると対応する異常検知器を出力する**ニューラルネットワークモデル(NNモデル)**を導入
- 複数の関連データセットでNNモデルを事前学習
⇒ **正常データのみ**のデータセットから**適切な異常検知器を再学習なしで生成可能**に



関連文献

[1] A. Kumagai, T. Iwata, Y. Fujiwara, "Transfer anomaly detection by inferring latent domain representations," in *Proc. 33rd Conference on Neural Information Processing Systems (NeurIPS)*, 2019.

連絡先

熊谷 充敏 (Atsutoshi Kumagai) ソフトウェアイノベーションセンタ
Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

05 低い誤検知率で異常を検知

部分AUC最大化のための半教師あり学習

どんな研究

機械学習における二値分類問題において、ラベルなしデータを活用することによって、部分AUC(偽陽性率が特定の範囲での真陽性率)を高める分類器の学習方法に関する研究です。例えば異常検知において、誤検知率を低く抑えた状態で真の異常を見逃さなくすることが可能になります。

どこが凄い

従来技術ではラベルなしデータを活用することができませんでした。私たちはラベルありデータとラベルなしデータを用いて近似的な部分AUCを計算する方法を考案しました。そして、その近似部分AUCを最大化させるように分類器を学習することによって、高精度の分類を可能にしました。

めざす未来

本技術を用いることで、コストが高いラベルを付ける作業を減らしても高い分類性能を達成することができます。これまで少数の学習データしか得られない場合には機械学習技術を適用できませんでした。が、本技術を発展させることにより、機械学習の適用範囲を広げることがめざします。

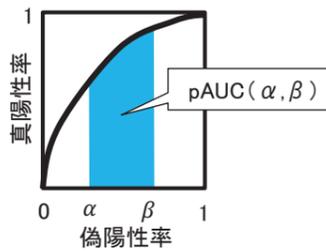
部分AUC PAUC: Partial Area Under the Receiver Operating Characteristic Curve

PAUC: $\alpha \leq$ 偽陽性率 $\leq \beta$ のときの真陽性率曲線(ROC)の下の面積

正例の割合と負例の割合が大きく異なる二値分類問題の評価尺度として注目

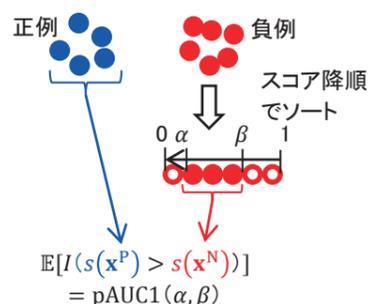
応用例1: 異常検知: 異常と誤って検知するとオペレータの負担増

応用例2: 病気診断: 病気と誤って診断すると不必要な検査によりコスト増



従来技術(教師あり学習)

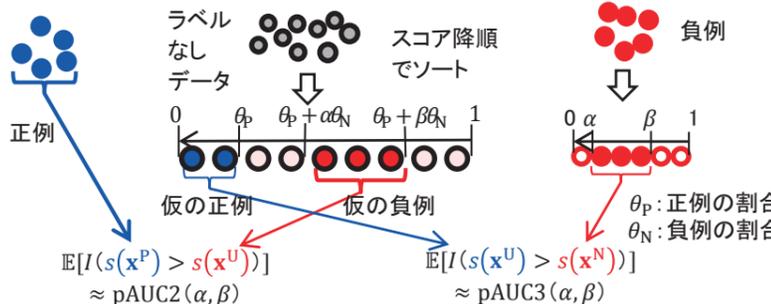
ラベルありデータ(正例、負例)を入力とし、部分AUC最大化により正例らしさを出力するスコア関数 s を学習



提案技術(半教師あり学習)

ラベルなしデータから計算できる2つの部分AUCの近似を導出し、スコア関数の学習のために利用

正例の割合 θ_p を境に、スコアの良い(悪い)ラベルなしデータは仮の正例(負例)とみなす



正例と負例で計算される部分AUC

目的関数 $L = \text{pAUC1}(\alpha, \beta)$

正例とラベルなしデータで計算される近似部分AUC

負例とラベルなしデータで計算できる近似部分AUC

目的関数: 3つの部分AUCの重み付き和 $L = \lambda_1 \text{pAUC1}(\alpha, \beta) + \lambda_2 \text{pAUC2}(\alpha, \beta) + \lambda_3 \text{pAUC3}(\alpha, \beta)$

関連文献

[1] T. Iwata, A. Fujino, N. Ueda, "Semi-supervised learning for maximizing the partial AUC," in *Proc AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

連絡先

岩田 具治 (Tomoharu Iwata) 上田特別研究室
Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

06 そのデータ、本当に偏ってますか?

決定グラフを用いた組合せ的相関検定

どんな研究

病気が流行したとき、その流行に地域性があるかないかを知ることは重要です。組合せ的相関検定では観測データである地域別病気発症率と地域の隣接関係から流行の地域性の有無を検定します。具体的には観測データの偏りとその稀さを計算し、病気の地域性の有無を判定します。

どこが凄い

偏りはScan統計量として、その稀さはP値(その統計量が考えられる全ての観測パターンのうち上位何%であるか)として計算されます。これらを愚直に計算すると47都道府県別の検定には1億年以上かかります。私達はこの検定時間を決定グラフと呼ばれるデータ構造を用いることで1日に短縮します。

めざす未来

本手法は流行の地域性のみでなく、センサ網上の侵入者の検知や購買履歴中の顕著な組合せの発見など、様々な組合せ構造に適用可能な一般的な手法です。本手法により、日々蓄積される様々な組合せデータに潜む重要な情報を自動的に抽出可能になります。

組合せ的相関検定 「観測は構造に依存している？」に回答



(例1) 病気の流行
赤: 病気が流行している都道府県
白: 病気が流行していない都道府県
※この観測は説明のための例です。
この病気の流行は
1. 地域性がある? (帰無仮説)
2. 偶然偏っている? (対立仮説)

(例2) センサ網
センサの反応は
1. 侵入者?
2. ノイズ?

(例3) 購買履歴
商品 A B の組合せは
1. 人気?
2. 偶然?

購買履歴	頻度
A B C	3
A B	5
B C	1
A	1
B	1

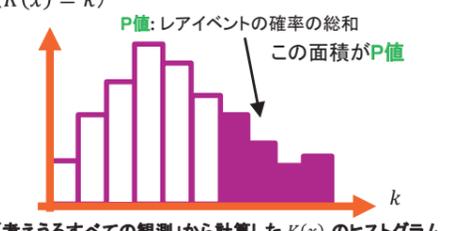
この問題の難しさ 観測の統計量が「考えられる全ての観測」のうち上位何%であるかの計算が必要

組合せ的相関検定の手順

- 観測 x と 仮説パターン集合 \mathcal{F} (例: 全隣接都道府県の集合) を受け取る
- 観測 x の Scan 統計量 (= 偏り) $K(x)$ を計算
- 統計量 $K(x)$ の P 値 (= 稀さ) を レアイベント集合 \mathcal{W} から計算
- P 値 \leq 有意水準 (例 0.05) なら対立仮説を棄却 \Rightarrow 地域性有り

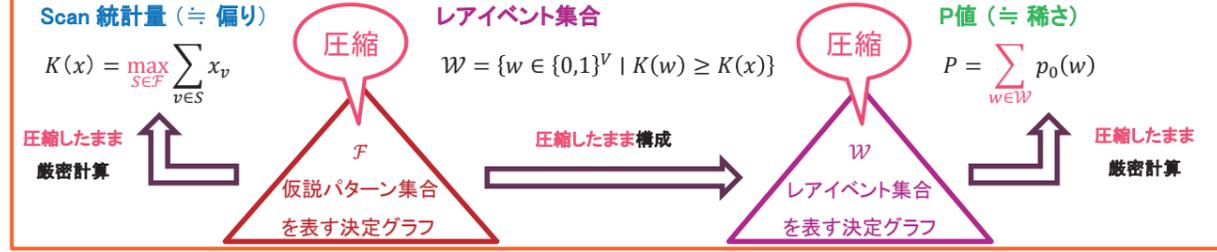
\mathcal{F} と \mathcal{W} は 指数的に巨大な集合 \Rightarrow 愚直に計算すると(例1)で1億年以上かかる
以下の提案技術を用いれば1日で計算可能!

「考えるすべての観測」から計算した $K(x)$ のヒストグラム



$p(K(x) = k)$ レアイベント: 現在の観測と同等以上に偏った観測
P値: レアイベントの確率の総和
この面積がP値

提案技術 仮説パターン集合とレアイベント集合を決定グラフで圧縮してP値を効率的に計算



関連文献

[1] M. Ishihata, T. Maehara, "Exact Bernoulli scan statistics by binary decision diagrams," *The 28th International Joint Conference on Artificial Intelligence (IJCAI-2019)*, 2019.

連絡先

石島 正和 (Masakazu Ishihata) 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



オープンハウス 2020

07

みんなが協力せず勝手に急ぐとどうなる？

混雑ゲームの均衡計算

どんな研究

道路や通信などのネットワークでは、ある辺(リンク)を使う人が多いほどその辺は混雑し、コスト(所要時間)が大きくなります。ネットワークの各利用者が経路(辺の組合せ)を選ぶ際に、誘導や制御がなく各々自分勝手にコストが小さい経路を選ぶ場合、どの辺がどの程度混雑するか計算できます。

どこが凄い

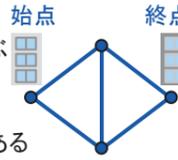
素朴な方法では、全ての経路について確率の計算をすることになりますが、経路の数が膨大なため実用上は不可能でした。私たちは二分決定グラフとよばれるデータ構造を用いて経路の集合を小さく表現することで、計算を高速化し、実用的な大きさのネットワークでの計算を可能にしました。

めざす未来

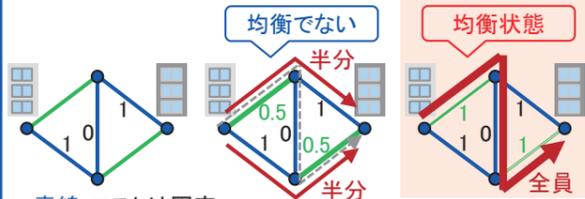
本技術により、道路や通信のネットワークの混雑状況を簡単に予測できるようになり、ネットワークの設計に役立ちます。また、本技術における「辺」は「アイテム」、「辺の組合せ」は「アイテムの組合せ」と一般化でき、経済学などの一般的な場合におけるゲーム理論的解析にも役立つと考えられます。

混雑ゲームと均衡状態

- ・ 無数のプレイヤー
- ・ プレイヤーは経路(辺の組合せ)を選ぶ
- ・ なるべく総コストが小さくなるよう選ぶ
- ・ 選べる経路は決まっている
- ・ 使う人が多いほどコストが増える辺もある



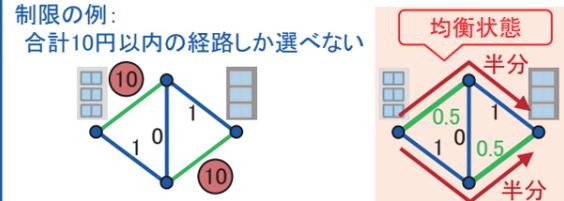
→ 各プレイヤーはどのような経路を選ぶ？
全員が、選べる中で今コスト最小の経路を選んでいる状態
= プレイヤーの不満が出ない状態 … 均衡状態



- *青線: コストは固定
- *緑線: 使うプレイヤーの割合と同じだけのコストがかかる

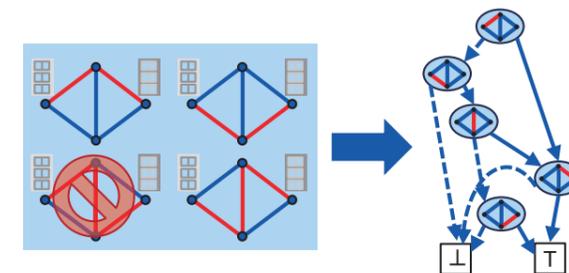
従来技術の課題

選べる経路に制限がないとき: 均衡状態は簡単に求める
制限があるとき: 選べる経路を一つ一つ見るしかない
→ 膨大な数の経路があり均衡状態の計算は困難

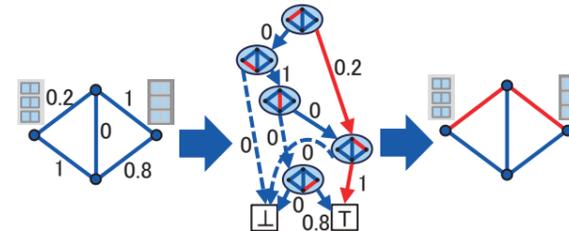


二分決定グラフを用いた解法

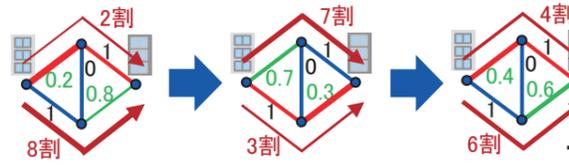
ゼロサプレス型二分決定グラフ (ZDD) を用いて、選べる経路の集合のうち共通部分をまとめてコンパクトに表現
・ 例えば8,000兆個の経路を1MB未満で表現できることも



ポイント1. 選べる経路の集合の中でコスト最小の経路は ZDDを使って簡単に求められる



ポイント2. 現時点のコストの値でコスト最小の経路を求め、その経路を選ぶプレイヤーを増やすことの繰り返し(反復)で均衡状態が求められる



関連文献

[1] K. Nakamura, S. Sakaue, N. Yasuda, "Practical Frank-Wolfe method with decision diagrams for computing Wardrop equilibrium of combinatorial congestion games," in Proc. 34th AAAI Conference on Artificial Intelligence (AAAI), 2020.

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



オープンハウス 2020

08

小さな窓から量子世界の全てをコントロール

量子系の間接的制御が持つ可能性を探る

どんな研究

量子的な系の全てを直接制御できることは稀です。仮に直接制御をしようとすると、ノイズがそのシステムに混入してしまいます。本研究では、ノイズを減らすために、あえて直接制御を避けて間接的に量子系を制御をした場合の制御上の影響を数学的に探究しています。

どこが凄い

間接的に量子系を制御しようとするとき、一般にどのような制御が可能になるのかを明らかにしました。その結果、直接制御可能な量子系の自由度が3以上あればそれ以外の直接制御不可能な量子系の自由度がどれほど大きくても実質的には任意に制御できるなどの普遍的な事実を発見しました。

めざす未来

量子系を間接的に自在に制御できるようにすることで、量子計算機をはじめとする量子情報処理におけるノイズ問題のブレイクスルーを起こします。これによって、遠隔での完全秘密乱数共有や巨大な因数分解といった、量子情報処理によって初めて可能になる機能を実際に実現します。

現在の量子計算機

- ・ I社やG社の量子計算機は53量子ビット使えるようになったが、ノイズも混入
- ・ スケールが大きくなるとノイズは増加
- ・ 数千ビット以上の大きい数の因数分解には、ノイズを小さく抑えて、数百万量子ビット必要

ノイズの入りかた

ノイズは操作可能な部分から侵入する

直接操作可能 / 直接操作不可能

新しい方針: あえて制御の容易性を犠牲にしてノイズの侵入を防ぐ

ノイズを減らすための理論的手段

直接制御できないがノイズの影響を受けにくい量子系を、多少ノイズの影響は受けるが直接制御できる小さな量子系と相互作用させて、全体を量子計算機として利用

一疑問一
制御可能な自由度※を減らせば制御回路からのノイズは減少するが、ちゃんと制御できるのか?
※自由度: 3量子ビットならば8

例: 超電導量子ビット

発見した事実: 自由度が2つの場合はダメだが、3つ以上あれば自在に量子系全体を制御可能
いかなる相互作用でも成り立つ普遍的な事実 (詳細は右に)

➡ ノイズを低く抑えつつ量子ビットの数を飛躍的に増やせる可能性が!!!

数学的背景

間接制御の場合、実行可能な操作(行列)の集合が下記の形式で必ず書けることを発見

$$L \approx \mathcal{L}(su(d_S) \otimes J \cup I \otimes [J, J])$$

Jという集合が満たす条件(特定の操作に関して閉じている)から、背後の量子的なシステムが存在する空間を持つべき構造を発見

直接制御できる一部の量子系の次元が2次元か3次元以上かによって、Jに要求される条件が変わることが左記の普遍的事実を保証

関連文献

[1] G. Kato, M. Owari, K. Maruyama, "Algebra and Hilbert space structures induced by quantum probes," *Annals of Physics*, Vol. 412, 168046, 2020.
[2] G. Kato, M. Owari, K. Maruyama, "Hilbert space structure induced by quantum probes," in Proc. 11th Italian Quantum Information Science conference (IQIS2018), 2019.

連絡先

加藤 豪 (Go Kato) 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



オープンハウス 2020

09 少量の追加データで作るカスタム機械翻訳

汎用対訳コーパスJParaCrawlを用いた機械翻訳の領域適応

どんな研究

機械翻訳では、対訳コーパスと呼ばれる学習データから自動で翻訳器を学習します。そのため、特定の領域(分野)に特化した翻訳器を作成するためには、その領域の**学習データが大量に必要**となっていました。この展示では、**少量の追加データだけで翻訳器を特定領域に特化**させる技術を紹介します。

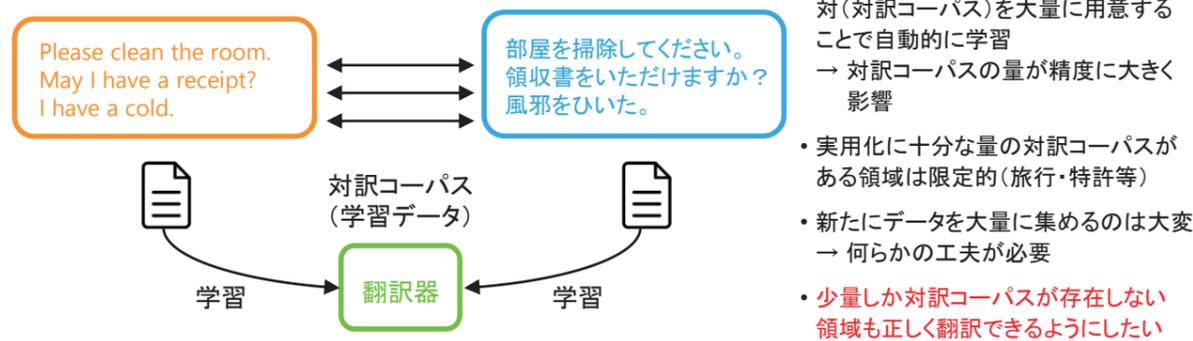
どこが凄い

Webデータを大量に収集し、自動的に対訳になっている文を見つけることで**大規模な日本語-英語の学習データを作成**しました。この学習データは様々な領域を網羅しているため、これを併用することで少量の学習データだけで特定領域への翻訳器を特化させることが可能になりました。

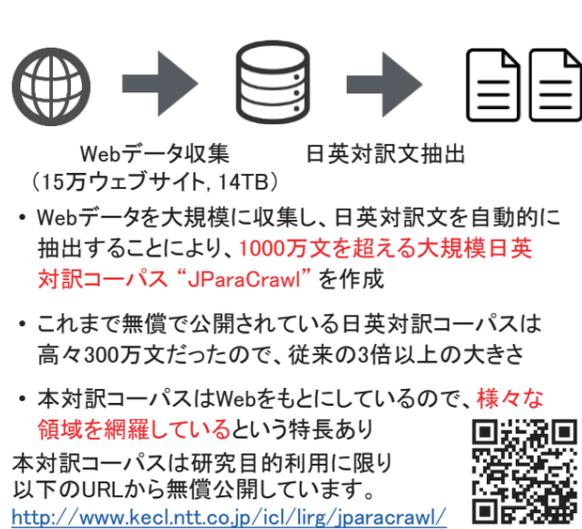
めざす未来

本技術を用いることで、これまで学習データが乏しかったため翻訳精度が低かった領域に対しても、少量の追加学習データで翻訳精度を飛躍的に向上させることが可能になります。将来的には、**どの領域に対しても高精度な機械翻訳の実現**をめざします。

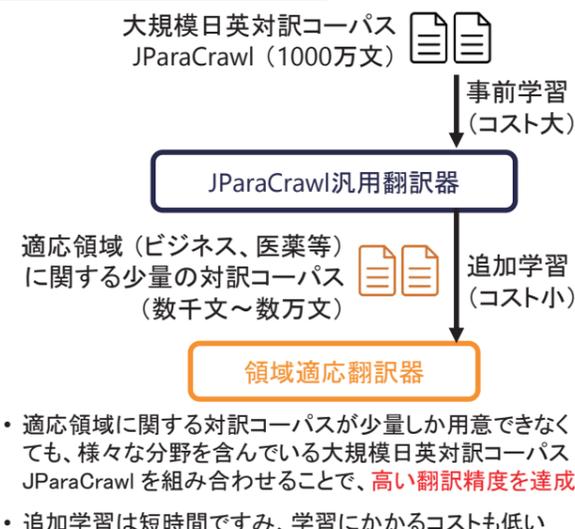
機械翻訳器の学習



大規模日英対訳コーパスの作成



少量追加データでの領域適応



関連文献

[1] M. Morishita, J. Suzuki, M. Nagata, "JParaCrawl: A large scale web-based Japanese-English parallel corpus," in *Proc. 12th International Conference on Language Resources and Evaluation (LREC)*, 2020.

連絡先

森下 睦 (Makoto Morishita) 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



オープンハウス 2020

10 こどもの感情発達レベルを測ります

表情・文脈・音声テストによる感情発達プロセスの解明

どんな研究

自分や他人の気持ちを理解する能力は対人コミュニケーションにおいて不可欠です。本研究では、その能力がどのように発達していくのかを明らかにするために、**こどもの感情発達を多角的視点から定量的に測定し、発達レベルを見える化**します。

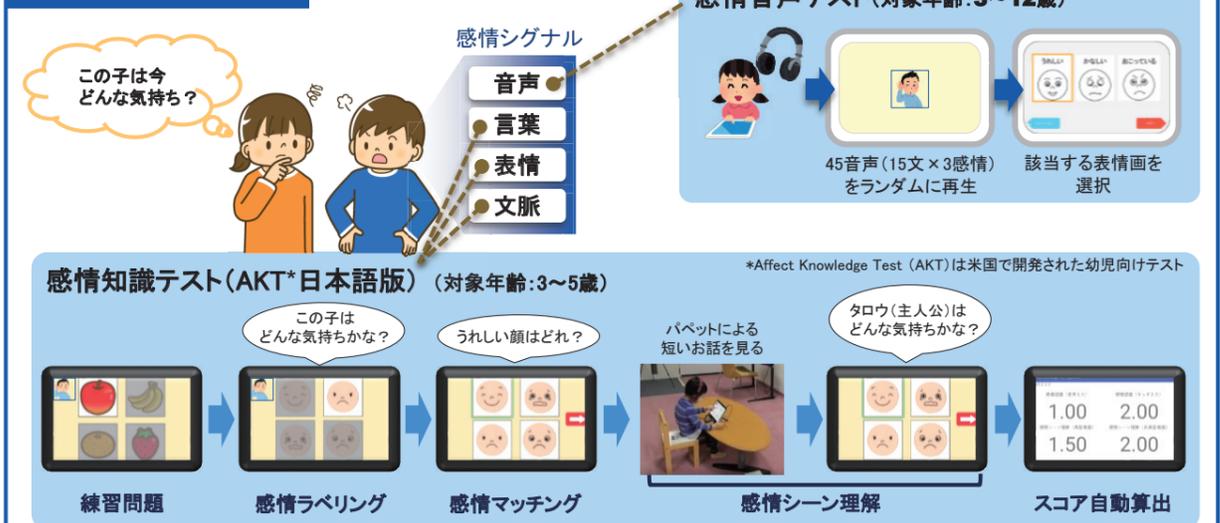
どこが凄い

従来手法では研修を受けたエキスパートのみが実施可能であったテストを、**タブレットアプリ化することにより場所の制約なくどこでも手軽に実施できるようにしました**。また、本テストを用いることで、どの感情シグナルの理解が難しいのかや、発達がゆっくりな子を容易に見出すことができます。

めざす未来

本テストを保育・教育機関に導入することで、こどもの感情発達を定期的に測定し、**先生が気になる子や人とのコミュニケーションに困難を抱えている子に適切な支援ができるようになります**。早期介入により、こどもの学校生活への適応や円滑な人間関係構築を支援することが期待されます。

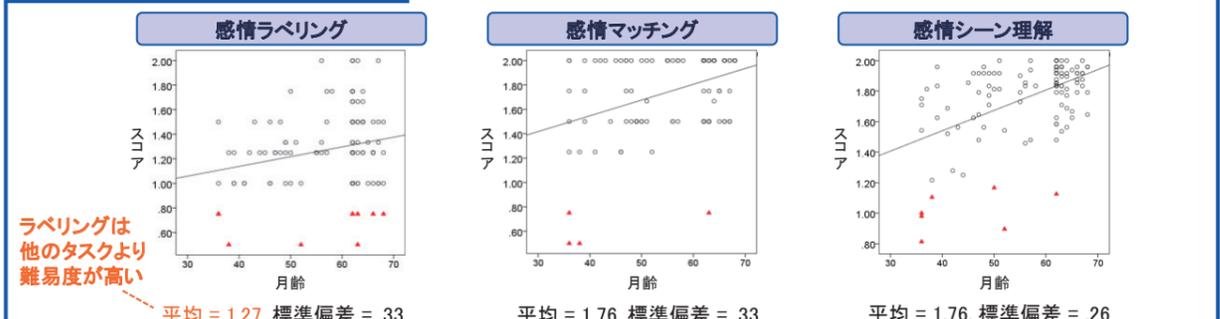
感情理解テストの開発



3-5歳児における感情理解発達

【調査対象児】3-5歳児116名(男児65名)

○ 定型発達群
▲ 発達遅延群
感情発達がゆっくりな子を容易に見出し可能



関連文献

[1] N. Watanabe, S. A. Denham, N. M. Jones, T. Kobayashi, H. H. Bassett, D. E. Ferrier, "Working toward cross-cultural adaptation: Preliminary psychometric evaluation of the Affect Knowledge Test in Japanese preschoolers," *SAGE Open*, 2019.

[2] N. Watanabe, T. Kobayashi, "Computerization of an emotion knowledge assessment for preschoolers: Supporting their school readiness," in *Proc. International School Psychology Association 41st Annual Conference*, 2019.

連絡先

渡邊 直美 (Naomi Watanabe) 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



オープンハウス 2020

11 こどもの興味と発達に合わせて絵本を作ります

パーソナル知育絵本を用いた親子の絵本読み活動支援の試み

どんな研究

絵本の読み聞かせがこどもの言語発達を後押しすることはよく知られています。特に、興味や発達にあった絵本を読んであげることが重要です。本研究では、こどもの興味と発達にあった絵本をテラメイドに作成する方法を提案します。読み聞かせを多くの親子に楽しんでもらうスキームも検証しています。

どこが凄い

こどもが言える語と興味があるものをウェブ上で親に回答してもらい、その情報を「幼児語彙発達データベース（幼児がいつどんな語を理解/発話できるかをモデル化）」と照合します。これから発話できそうな語をそれぞれの子に合わせて推定し、絵とともに挿入した絵本（パーソナル知育絵本）を作成します。

めざす未来

現在、沖縄県恩納村で乳幼児健診受診者に対し図書館でパーソナル知育絵本の受付と配布を行う実証実験を行っています。こうした取り組みからこの絵本の効果を探るとともに、図書館への来訪を促して絵本と触れ合う機会を増やすことにより、親子の絵本読み活動を後押しできるかどうかを検証しています。

■ パーソナル知育絵本

かしわらあきお作、NTT印刷(株)



対象年齢: 1-2さい

申し込みサイトから情報投入
<https://ehon.nttprint.com>

- ・なまえ、性別
- ・すきなもの
- ・語彙チェックリスト (現在発話できる語)

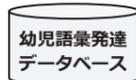
① 登場する主人公の名前と性別を本人に合わせる



② お子さんそれぞれの興味にあわせてストーリーをカスタマイズ



③ これから覚える語を推定し該当する絵を挿入



■ 沖縄県恩納村での実証実験

- ・期間: 2020年2月~2021年3月
- ・NTT印刷(株)、恩納村と共同で実施



乳幼児健診で全受診者へ案内 (恩納村総合保健福祉センター)



図書館で申し込みと受け取り (恩納村文化情報センター)



約2週間でお渡し 家庭での絵本読み活動

- ・親子を図書館へ誘導
- ・絵本に触れ合う機会を提供

おなまえの入った親しみやすい知育絵本で絵本を大好きに!

関連文献

- [1] NTTグループウェブサイト TOPICS: 「パーソナル知育絵本を用いた親子の絵本読み活動の支援に関する共同実験を沖縄県恩納村で開始」 <https://www.ntt.co.jp/topics/oyakoehon/index.html>
- [2] 絵本ナビ: 「赤ちゃん研究から誕生した絵本で楽しく親子のコミュニケーション: その子の言語発達段階にあった絵本を一人ひとりに届けたい! パーソナルちいこえほんが誕生するまで」 https://style.ehonnai.net/ehon/2020/03/17_358.html

連絡先

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



オープンハウス 2020

12 あなたの語彙数測ります

令和版語彙数推定テスト

どんな研究

単語親密度とは、語のなじみ深さを評定実験によって数値化したものです。NTTが20年以上前に調査した約7万7千語の単語親密度は、語彙数推定テストをはじめ、今でも多くの分野で活用されています。我々は経年変化の調査や、新語追加のため、令和版単語親密度調査に取り組みました。

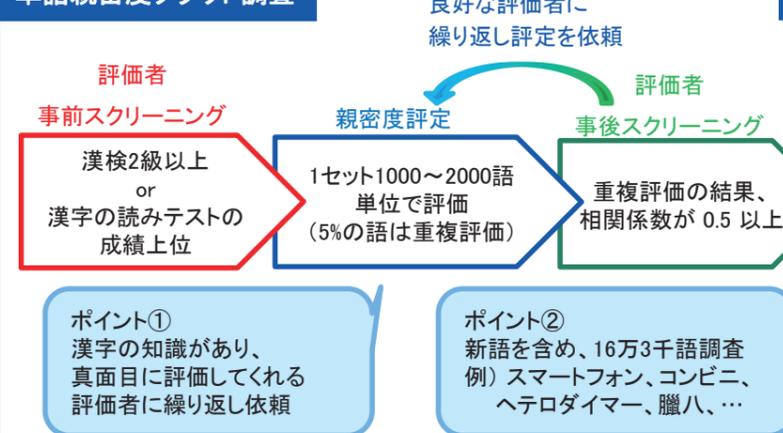
どこが凄い

単語親密度の調査を16万3千語という類を見ない規模で実施しました。クラウドソーシングを用いながら、スクリーニングを丁寧に実施することで、信頼度の高い調査結果を得ることに成功しました。これにより、20年前との比較が可能になりました。これを基に、令和版語彙数推定テストを作成しました。

めざす未来

本調査結果を用いることで、現代に適した語彙数の推定が簡単に行えるようになります。小学生~高校生を含む幅広い年代の語彙数調査に取り組んでいます。今後、語彙数と読解力や学力全般との関係の調査・分析を実施、効果的な学習支援の実現をめざします。

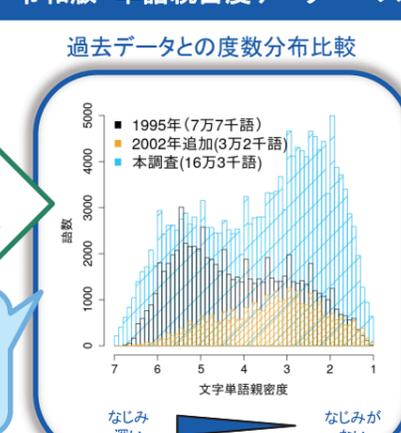
単語親密度クラウド調査



ポイント①
漢字の知識があり、真面目に評価してくれる評価者に繰り返し依頼

ポイント②
新語を含め、16万3千語調査例) スマートフォン、コンビニ、ヘテロダイマー、臘八、...

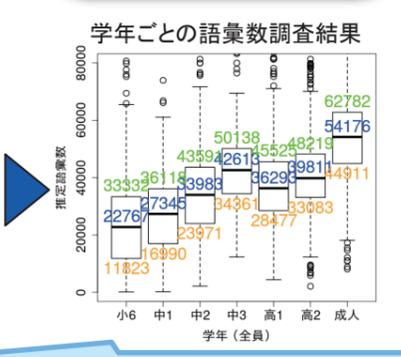
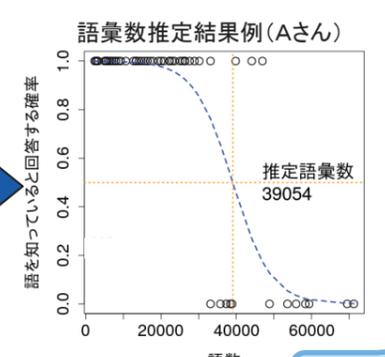
令和版 単語親密度データベース



語彙数推定方法

知っている語をチェックしてください。

- 銀行
- 担当
- ...
- 親疎
- 改鑄



ポイント③
50語程度の簡単チェック

ポイント④
小・中・高・成人 約3600人で語彙数調査を実施

語彙数推定テストはこちらから: <http://www.kecl.ntt.co.jp/icl/lirg/resurces/goitokusei/>

関連文献

- [1] S. Fujita, T. Kobayashi, "単語親密度の再調査と過去のデータとの比較," 言語処理学会第26回年次大会, pp.1037-1040, 2020.
- [2] S. Fujita, T. Kobayashi, T. Yamada, S. Sugawara, T. Arai, N. Arai, "小・中・高校生の語彙数調査および単語親密度との関係分析," 言語処理学会第26回年次大会, pp.355-358, 2020.

連絡先

藤田 早苗 (Sanae Fujita) 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



オープンハウス 2020

13 京町セイカがご案内！

地域連携で作るなりきりAI

どんな研究

市役所等での業務に関するQAや受付対応、観光案内などを行う“なりきりAI”を学習します。従来のQAや受付対応、案内などを行うAIの学習には、多くの誤りのない学習データが必要であり、それらを集めることは非常に高コストで困難な課題でした。この課題を地域連携を活用することで解決します。

どこが凄い

データ収集作業を地域連携の活動として行うことで、正確なデータを低コストで収集します。その地域に住む人と、その地域に興味のある人を結びつけ、来たい人が質問を、来て欲しい人が回答をすることで非常に高品質なAIの学習データを収集し、**地域と密接に連携したなりきりAIの学習を実現**しました。

めざす未来

本技術を用いることで、**それぞれの地域の需要に合わせた“なりきりAI”を提供できるようになります**。今後は自治体での事例を発展させて、企業や団体等での利用も見据えて、より幅広い場面で活用できる技術の実現をめざします。

AIの発展に伴い、「AIによる業務の効率化」が進んでいます。特に、**受付対応、観光案内といった対話を通してユーザに情報を提供するAIの実現**は、人の負担の軽減や、人とAIの協働に向けて注目されています。この研究では**京都府精華町との共同実験を通して実社会で役立つAIの実現**に取り組んでいます。

なりきりAIによるタスク対話の実現

“なりきりAI”は、アニメキャラクターや有名人といった特定のキャラクターを模した対話システムです。企業の公式キャラクターや有名人が受付や案内をすることで**お客様の満足度向上や、エンターテインメント性の付与を行いつつ、受付担当者の負担を軽減する効果**が期待されています。

住民票はどこで発行できますか？



学習データ

2階の総合窓口で受付をしています。このフロアの先を左に曲がってください。

従来は(少数の)人間が手作り
→課題:データ収集コストが高い

地域連携による学習データ収集

公式キャラクター

課題であった**正確な学習データ収集を「興味のある人」と「知識のある人」を結びつけることで解決**→知識のある人(精華町民のみなさま)による**学習データ作成は地域連携活動の一環として実施**

精華町に住みたい・行ってみたい人が質問
「春だと、精華町では何が有名ですか？」



精華町に住んでいる人が回答
「冬から春にかけて、いちご狩りが有名ですよ！」

受付

- ポイント
- なりきりAI 学習データの収集
- ・精華町に興味のある人が質問
= **需要の高い情報**
- ・精華町の人が回答
= **正確性の高い情報**

地域連携によるデータ作成
= **非常に低コスト、高品質**

関連文献

[1] R. Higashinaka, M. Mizukami, H. Kawabata, E. Yamaguchi, N. Adachi, J. Tomita, “Role play-based question-answering by real users for building chatbots with consistent personalities,” in *Proc. 19th Annual SIGdial Meeting on Discourse and Dialogue*, 2018.

連絡先

水上 雅博 (Masahiro Mizukami) 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



オープンハウス 2020

14 こんなとき、あの人だったらどう思う？

人の個性を考慮した体験に紐づく感想生成技術

どんな研究

対話システムが人間と良い関係を築いていくには、人間の感情を理解し、表現することが必要不可欠です。この研究は、「**こういう出来事が起きたとき、あの人だったらこういうことを思いそうだな**」という**人に紐づいた感想の予測を行う**ことで、より人間の感情を理解、表現できる対話システムを実現します。

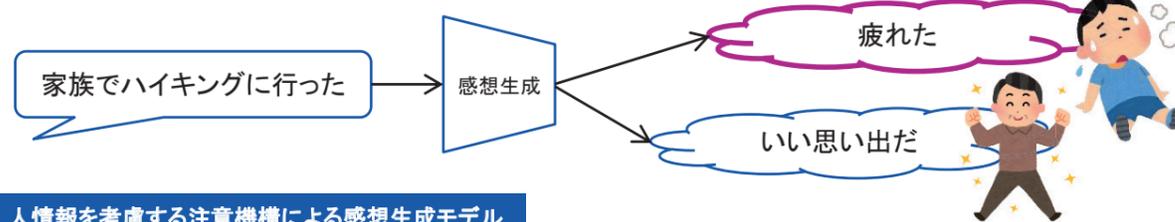
どこが凄い

従来の研究では、感情の表現は喜怒哀楽のような類型で行われていました。これに対して、私たちは「**誰が」「何をしたとき**」を入力として、「**どのような感想を抱くか**」という出力を特定の「人」に応じて推定します。これにより、対話システムがより柔軟な感想の表現を、人に応じて使い分けすることができます。

めざす未来

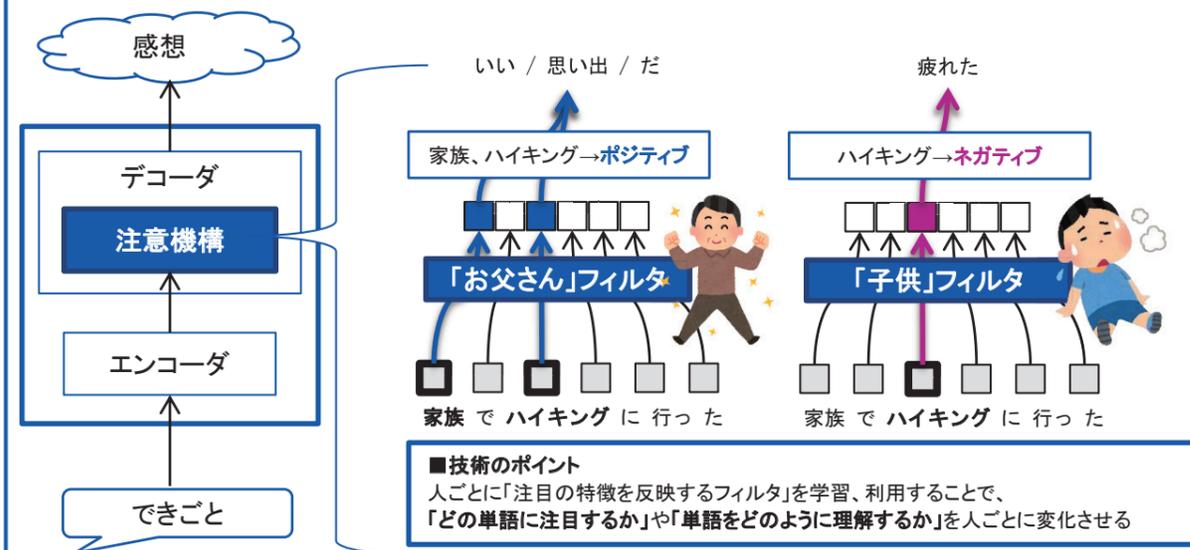
本技術を用いることで、対話システムが人間の感情をより柔軟に表現できるようになります。今後は、事前に学習した人のみでなく、**対話中のユーザが、どのような「感想の表現の特徴」を持っているかをリアルタイムに学習**したり、ユーザの述べた感想からどのような人物かを推定する手法を実現します。

感想生成とは ある「できごと」が起きたときに、「どのような気持ちになるか」を推定、文として生成する技術
先行研究→一般的に・多数派がポジティブ・ネガティブもしくは喜怒哀楽のどれだと思ふかをクラス推定
課題→同じ出来事でも、経験する人によって感想は千差万別＝「人」に応じて出力を変化させる必要



人情報を考慮する注意機構による感想生成モデル

注意機構に着目→人の注目の特徴を反映するフィルタを導入し、人ごとに異なる出力を実現



■技術のポイント
人ごとに「注目の特徴を反映するフィルタ」を学習、利用することで、「どの単語に注目するか」や「単語をどのように理解するか」を人ごとに変化させる

関連文献

[1] M. Mizukami, H. Sugiyama, H. Narimatsu, “Event data collection for Recent Personal Questions,” in *Proc. LACATODA*, 2018.
[2] M. Mizukami, H. Sugiyama, H. Narimatsu, T. Arimoto, R. Higashinaka, “話者情報を考慮する注意機構を用いた応答生成手法の検討,” *人工知能学会*, 2020.

連絡先

水上 雅博 (Masahiro Mizukami) 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



オープンハウス 2020

15

この声、何歳？

話者クラスタリングを用いた深層話者属性推定

どんな研究

音声から発話者の年齢や性別といった話者に関する情報を推定する研究です。顔画像や動画といった視覚的な情報からこれら情報を高い精度で推定する手法は既にいくつか知られていますが、音声のみしか利用できない場合、最新の深層学習技術をもってしても未だ解決が困難な問題です。

どこが凄い

高精度な年齢推定を行うためには各年代の話者の膨大な学習データが必要です。しかし実際には年代毎にデータ量の違いがあり、特にデータが少ない年代の推定が困難でした。そこで、**声**が似た他の話者の推定結果を用いて補正することで、従来よりも高い精度で年齢推定できる技術を考案しました。

めざす未来

本技術は年齢のみならず感情など話者に関する様々な属性推定へ応用できます。今後は、各属性推定のための深層学習モデルと共に更なる性能改善を行い、話者属性を推定する汎用的な枠組を実現し、**ユーザに特化した新たな音声インタフェース開発やマーケティングへの応用**をめざします。

深層話者属性推定

問題設定

音声から**発話者の年齢や性別などの話者属性情報**を深層学習モデルにより推定

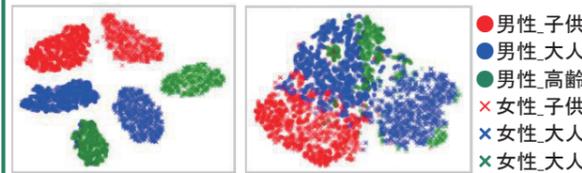
応用分野

コールセンタの応答決定やマーケティングの支援、ユーザの属性により挙動を変える音声対話システムの実現など



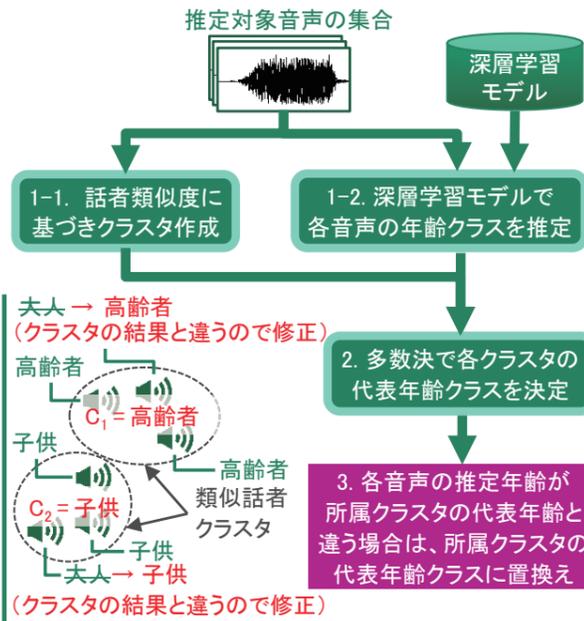
問題の難しさ

- 年代ごとに学習データ量の偏りが大きい(図1)
- 特定の話者・年代に対し過学習してしまう(図2)



話者クラスタリングに基づく性能改善法

深層学習モデルによる年齢推定の推定結果を類似話者の推定結果の多数決により補正する



●: 推定対象音声とその年齢推定結果
○: クラスタ毎に多数決で決定した代表年齢クラス

評価尺度	従来手法	提案手法
年齢クラス分類精度 (子供・大人・高齢者クラス)	59 %	72 %
年齢推定誤差	± 10.9 歳	± 8.7 歳

関連文献

- [1] N. Tawara, H. Kamiyama, S. Kobashikawa, A. Ogawa, "Improving speaker-attribute estimation by voting based on speaker cluster information," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020 to appear.
- [2] N. Tawara, H. Kamiyama, S. Kobashikawa, A. Ogawa, "話者クラスタリングに基づく話者年齢・性別推定精度改善法," *日本音響学会研究発表会講演論文集*, pp. 815-816, 2019.

連絡先

俵 直弘 (Naohiro Tawara) 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



16

ワイヤレスマイクを同時により多く使えます

ビット誤りに頑健で低遅延な音声音響符号化方式BRAVE

どんな研究

マイクとスピーカの間で音声を無線で送る際、**音声データを圧縮して送ることで同時に多くのマイクを使えます**。その際短い時間で圧縮する必要があり、時には圧縮データが正しく伝わらない場合があります。このような条件下で音質を保ったまま音声を圧縮することは難しい課題でした。

どこが凄い

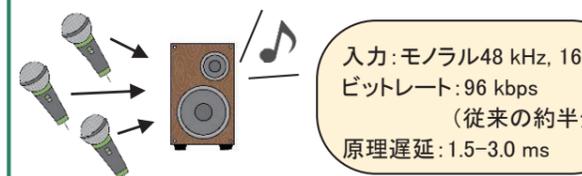
音声データの圧縮では、入力をそれぞれ何ビットで表現するかを適応的に決めなければなりません。しかし、その割当ビット数がデコーダに誤って伝わると大きなノイズが生じる可能性があります。私達は入力の整形法を工夫することで、**短時間で圧縮しつつ誤りに頑健な表現**を実現しました。

めざす未来

本技術を用いることで、ワイヤレスマイク以外にも**より多くの機器の間で音声をリアルタイムに高音質で送りあうことができるようになります**。拡声の用途だけでなく、より豊富な音声の情報を活用できるようにすることで、IoT (Internet of Things) 技術の発展に貢献できると期待しています。

音声音響符号化方式BRAVE

ワイヤレスマイク等、機器間のリアルタイムな音声通信に用いるための音声音響データの圧縮符号化方式を開発



ビットレートを低くできれば同時に使えるマイク数が増やせる

問題の難しさ

伝送中に符号がビット単位で誤る場合がある

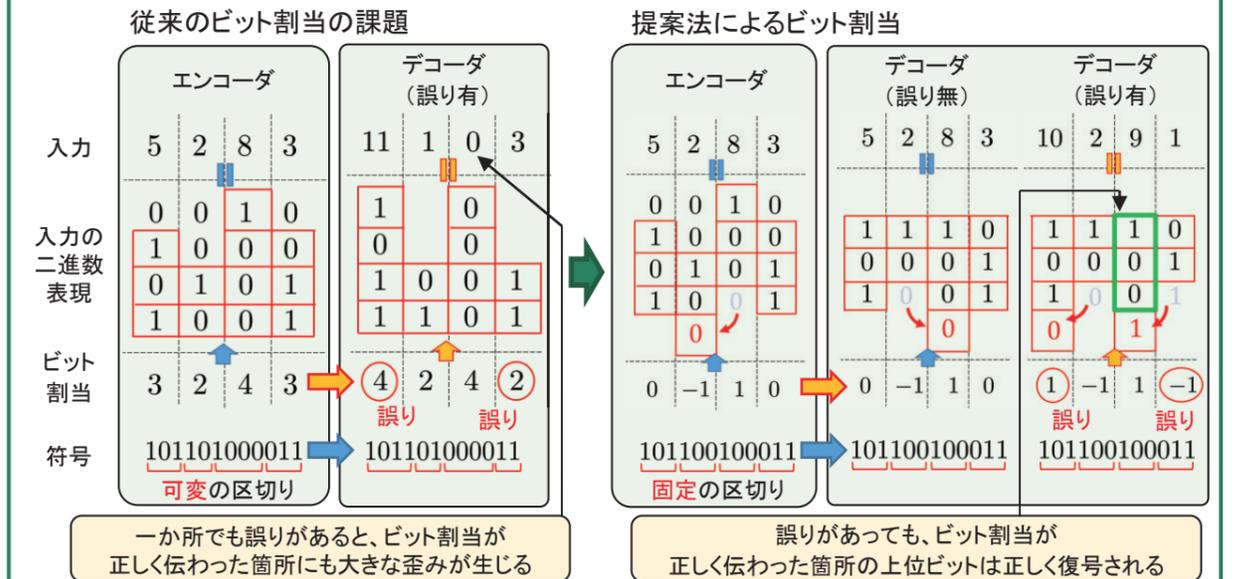
- 従来手法の多くはパケット等で保護している
- 保護には単位時間毎に付加情報が必要
- 低遅延の条件では付加情報の割合が大きくなりすぎる

→ パケット等を用いずに符号誤りへの対策をしたい



提案技術: ビットプレーン置換による整形

符号が間違っただけによる音質への影響を小さくする



関連文献

- [1] R. Sugiura, Y. Kamamoto, T. Moriya, "Spectral-envelope-based least significant bit management for low-delay bit-error-robust speech coding," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.

連絡先

杉浦 亮介 (Ryosuke Sugiura) 守谷特別研究室
Email: cs-openhouse-ml@hco.ntt.co.jp



17 聞きたい人の声に耳を傾けるコンピュータ(Ⅱ)

音声と映像を手がかりとしたマルチモーダル選択的聴取

どんな研究

人間は、複数の人が同時に話している状況においても、聞きたい人の声に集中し、聞きたい人の声を聞き取る能力(= 選択的聴取)を持っています。本研究は、そうした人間が持つ**選択的聴取の機能をコンピュータ上で実現**することをめざしたものです。

どこが凄い

音声情報に加え、映像情報を手がかりとして利用する、**マルチモーダル選択的聴取の技術を実現**しました。人間のように**複数の情報源を適切に活用**することで、声の性質が似た話者の会話といった、音声情報だけでは困難であった状況でも安定して動作可能な技術へと発展しました。

めざす未来

複数の人の声が入った音声から「聞きたい人の声のみを抽出する技術」は、人の音声を入力とする様々なデバイスの基盤となる技術です。人を認識して対応を変えるロボットやスマートスピーカーの実現といった、**人とより自然に対話するコンピュータの実現**に寄与します。

音声情報に基づく選択的聴取

□ 選択的聴取とは

- 複数の人の声が入った状況においても、聞きたい人の声に注意し聞くことができる能力
- 日常的な会話シーンでは、複数の人の声が入った状況は自然に起こる
- ⇒ 選択的聴取は人間には容易であるが、従来のコンピュータには困難な問題であった
- ⇒ 音声情報に基づく選択的聴取の初提案 (オープンハウス2018)

□ 課題

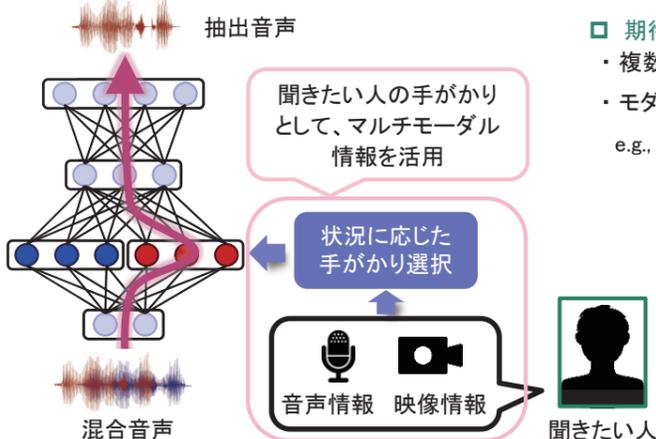
音声情報のみを手がかりとした選択的聴取は、**声の性質が似た話者同士の場合には困難**



音声と映像双方に由来する話者手がかりの利用

□ SpeakerBeam (= 深層学習に基づく選択的聴取モデル)

聞きたい人に関する手がかりを与えることで、混合音声からその話者の音声のみを取り出す深層学習モデル



□ 解決の方針: Multimodal SpeakerBeam の提案
声の特徴(音声情報)に加え、唇の動き(映像情報)を手がかりとして利用

⇒ 人間のように**マルチモーダル情報を活用**

□ 期待される効果

- 複数モダリティの活用による性能向上
- モダリティの劣化や欠損に対する頑健性向上
- e.g., 声の特徴が役に立たない(**声の性質が似た話者**)
映像データが欠損した(**唇が映らない**)
状況でも**抽出**できる

* 聞きたい人の手がかりについて

- 事前に録音された聞きたい人の音声データ
- 混合音声と同時に録画された聞きたい人の映像データ(唇周り)

関連文献

[1] T. Ochiai, M. Delcroix, K. Kinoshita, A. Ogawa, T. Nakatani, "Multimodal SpeakerBeam: Single channel target speech extraction with audio-visual speaker clues," in *Proc. Interspeech*, 2019.
[2] K. Zmolikova, M. Delcroix, K. Kinoshita, T. Ochiai, T. Nakatani, L. Burget, J. Cernocky, "SpeakerBeam: Speaker aware neural network for target speaker extraction in speech mixtures," *IEEE Journal of Selected Topics in Signal Processing*, 2019.

連絡先

落合 翼 (Tsubasa Ochiai) 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



オープンハウス 2020

18 顔で声の表情を制御する

クロスモーダル音声表情変換

どんな研究

対話や講演において、心理的な緊張状態や能力的な限界などにより思うような話し方で話せない場合があります。本研究では、**入力音声の話し方の雰囲気(表情)を、顔の表情や体の動作などの手段により制御**することを目的としたクロスモーダル音声表情変換の問題に初めて取り組みました。

どこが凄い

話し方の表情は声質・抑揚・リズムによって決まります。従来技術の多くは声質のみの変換を扱っていましたが、我々の音声変換技術は、**声質とともに抑揚やリズムの変換も可能**にします。この技術と顔表情識別技術を組み合わせることで、**顔画像を用いて声の表情を変換する技術を実現**しました。

めざす未来

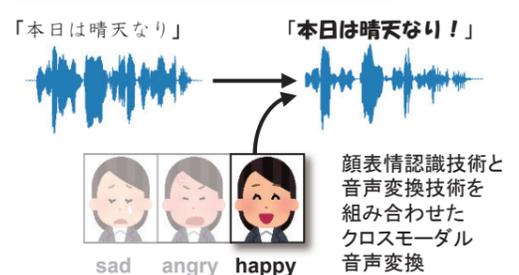
人と人のコミュニケーションには、物理的・能力的・心理的な状態に起因する様々な形の制約が存在します。本研究では、このような制約を取り除き、**あらゆる人が不自由なく快適にコミュニケーションを行える環境を実現**することをめざしています。

音声変換によるコミュニケーション機能拡張

音声変換技術を通じて多様なシーンにおいて人が不自由なくコミュニケーションできる手段を創出

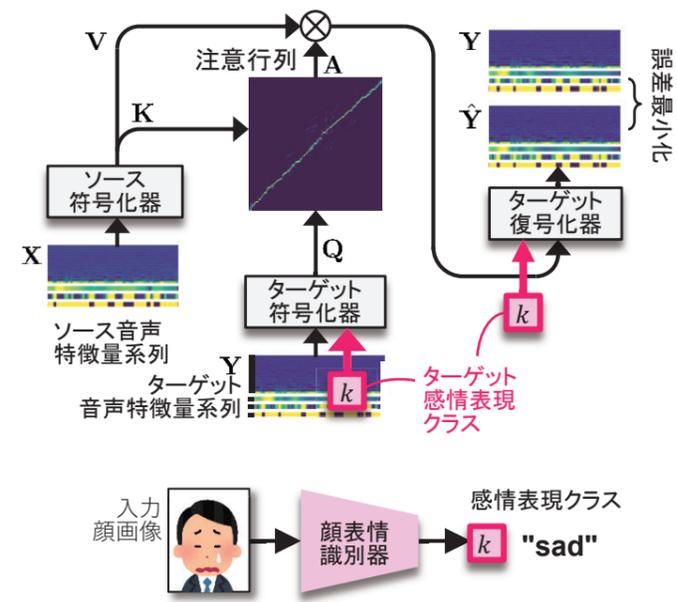


顔画像による音声表情の制御



系列変換モデルを用いた音声変換

- 系列変換(S2S)モデル
- 系列から系列への変換則を学習する**深層学習**の枠組(機械翻訳や音声認識など)での有効性が知られる
- エンコーダ/デコーダ構造と注意機構で**長期依存関係を捉えた系列変換**を扱える
- 変換元と変換先の系列の**要素間の対応づけ規則を学習**できる
- **通常は大規模な学習データが必要**
- S2Sモデルを用いた音声変換[1,2]
- 話し方の雰囲気や感情表現は、**抑揚やリズム**に特に色濃く現れる
- S2Sモデルにより**声質だけでなく抑揚、リズム、話速の変換**を扱える
- **少ない学習データでも学習可能な方式**を考案



顔表情認識

- 全層畳み込みネットワークを用いたクラス識別
- 顔表情識別結果を音声変換器に入力

関連文献

[1] H. Kameoka, K. Tanaka, T. Kaneko, N. Hojo, "ConvS2S-VC: Fully convolutional sequence-to-sequence voice conversion," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, under review.
[2] K. Tanaka, H. Kameoka, T. Kaneko, N. Hojo, "AttS2S-VC: Sequence-to-sequence voice conversion with attention and context preservation mechanisms," in *Proc. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2019)*, pp. 6805-6809, 2019.

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



オープンハウス 2020

19 探し方を学びながら探す

適応的スポッティング法による効率的な物体探索

どんな研究

点群データにより3次元表現された実空間の中から、**所望の形状の物体を素早く探し出す研究**です。従来の物体探索技術は、総当たり探索を基本に、見込みのない場所だけをスキップする発想で高速化が図られてきましたが、それでもたくさんの場所を探す必要がありました。

どこが凄い

効率的な探し方を自ら学習する物体探索法「適応的スポッティング法」を考案しました。より早く正確に所望の物体を探し出すことができた場合に高い報酬を与える深層強化学習により、物体がありそうなところだけを注意深く探す、人と同じような探し方ができるようになりました。

めざす未来

物流・製造・交通など、**3次元空間中の物体を高速・正確に探し出す**機能が求められる領域での活用が期待できます。身の回りを取り巻くセンサやロボットへの搭載を進め、世界中のあらゆる空間にある物体を自在に探し出して活用・操作可能にする仕組みの実現をめざします。

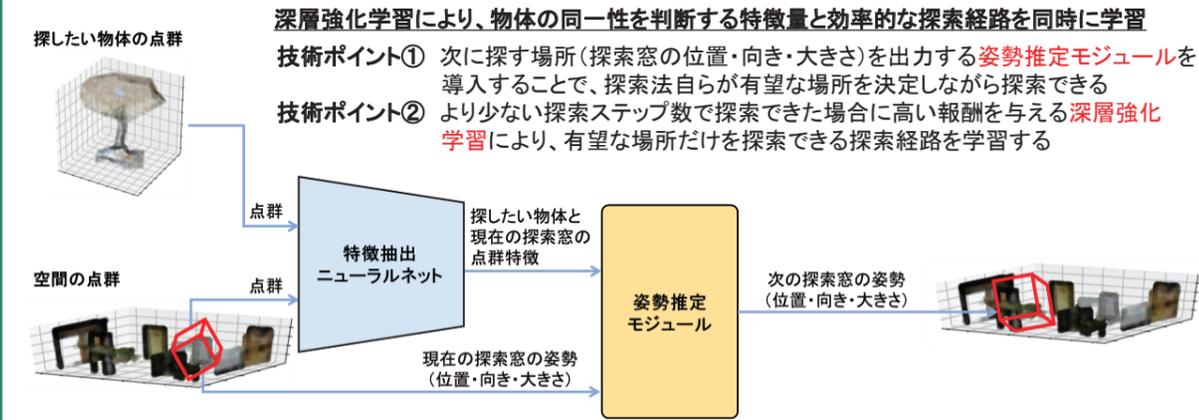
点群物体探索問題

3次元センサ(LiDARなど)で計測した空間の点群マップの中から、**所望の形状の物体**を探し出す

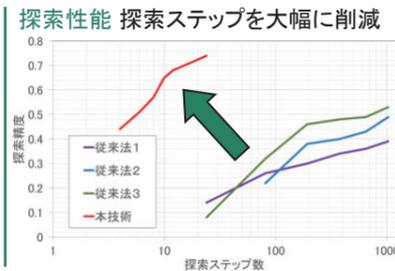
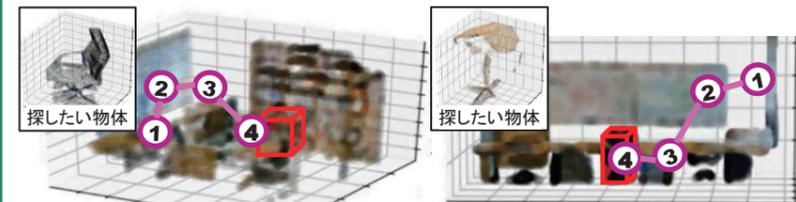
空間の形はさまざま、かつ、物体に対して空間が非常に大きい
⇒ **総当たり探索では時間がかかる**



適応的スポッティング法



探索経路の例(オフィスの椅子を探索した場合)
少ないステップで探したい物体にアプローチ可能



点群データの出典: Stanford 2D-3D-Semantics Dataset <http://buildingparser.stanford.edu/dataset.html>

関連文献

- [1] O. Krishna, G. Irie, X. Wu, T. Kawanishi, K. Kashino, "Learning search path for region-level image matching," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
- [2] O. Krishna, G. Irie, X. Wu, T. Kawanishi, K. Kashino, "Deep reinforcement template matching," in *第22回 画像の認識・理解シンポジウム (MIRU)*, 2019.
- [3] O. Krishna, G. Irie, X. Wu, T. Kawanishi, K. Kashino, "適応的スポッティング: 深層強化学習に基づく3D点群物体探索," in *第26回 画像センシングシンポジウム (SSII)*, 2020.

連絡先

クリシュナ オンカー (Onkar Krishna) 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



オープンハウス 2020

20 データを端末から漏洩させない分散深層学習

分散NW上で機械学習をするための非同期合意形成技術

どんな研究

現在の深層学習では、1か所に集約したデータを使ってモデルを学習することが一般的です。しかし、データ量の激増や**プライバシー保護**の観点から、近い将来データは**分散蓄積**されるようになるでしょう。多端末に分散蓄積された**データを外に出すことなく、機械学習モデルを最適化**する手法を提案します。

どこが凄い

多端末に蓄積されたデータは、**統計的に偏っている**と仮定することが自然です(例:各端末には一部クラスのデータしか存在しない)。その状況で、端末同士がモデル等の変数を**非同期に交換(通信)**しながら、全データを使って獲得したかのような**グローバルモデル**を学習するアルゴリズムを開発しました。

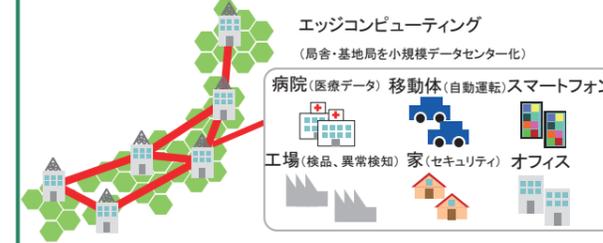
めざす未来

現在は、一部のアプリケーションプラットフォームがデータを集約/独占することにより、高度なサービスを提供しています。**データの所有権を個人に帰属**させ、プライバシーを保護しながら、多様なサービスに利用できる社会をめざしています。

目的・アプリケーション

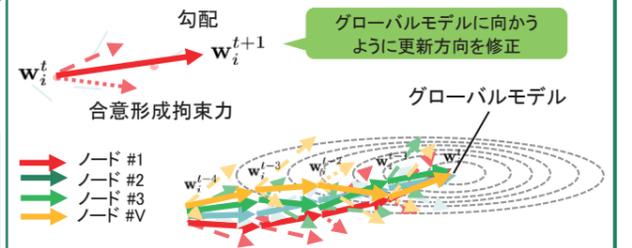
背景: データ量、プライバシー保護、法的規制(e.g. GDPR)の観点でデータを分散蓄積する時代になる。

目的: **分散蓄積されたデータを外に出すことなく、深層学習モデルを最適化**したい(ただし、モデルなどの変数を非同期で交換(通信)することは許容)。



非同期分散型深層学習アルゴリズム

提案方式: モデル主変数とラグランジュ双対変数をノード間で非同期に交換しながら、**グローバルモデル**を得るための学習アルゴリズムを構築(任意のネットワーク構造で動作可)。



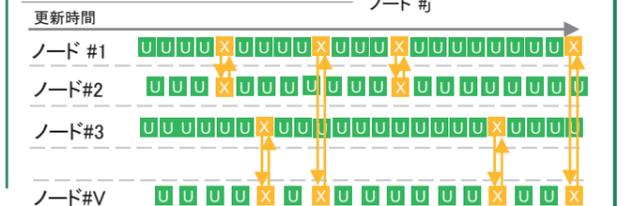
構築したアルゴリズム

Algorithm 1 PDMM SGD/ADMM SGD

```

1: Initialization of  $\hat{x}_{ij}^0, w_i^0$ 
2: for  $k \in \{0, \dots, K-1\}$  do
3:   Step 1: Update model for each node
4:   for  $i \in \mathcal{V}$  do
5:      $w_i^{k+1} \leftarrow (\mu w_i^k - \nabla F_i(w_i^k; x_{i,\cdot}^k)) / (\mu + \alpha |\mathcal{N}(i)| + \gamma |\mathcal{N}(i)|)$ 
6:     for  $j \in \mathcal{N}(i)$  do
7:        $\hat{x}_{ij}^{k+1} \leftarrow \hat{x}_{ij}^k - 2\lambda_{ij} w_i^{k+1}$ 
8:     end for
9:   end for
10:  Step 2: Exchange and update variables at random time  $k$ 
11:  for  $i \in \mathcal{V}$  do
12:    Select  $j \in \mathcal{N}(i)$  at random
13:    Transmit  $(w_i^{k+1}, \hat{x}_{ij}^{k+1})$ 
14:     $\hat{x}_{ij}^{k+1} \leftarrow \hat{x}_{ij}^{k+1} + (1-\theta)\hat{x}_{ij}^k$  (PDMM SGD)
15:     $\hat{x}_{ij}^{k+1} \leftarrow \hat{x}_{ij}^{k+1} + (1-\theta)\hat{x}_{ij}^k$  (ADMM SGD)
16:  end for

```



問題の難しさ

難しいポイント: 各ノードにある**データが統計的に偏っている**とき、各ノードの評価関数を最小化するだけでは、**グローバルモデルは得られない**。

アプローチ: 全ノードの**モデルが一致する制約条件下で、評価関数**を最小化問題を解く。

$$\inf_{\{w_i | i \in \mathcal{V}\}} \sum_{i \in \mathcal{V}} F_i(w_i; x_i) \quad \text{合意形成}$$

$$\text{s.t. } A_{ij} w_i + A_{ji} w_j = 0 \quad A_{ij} = \begin{cases} \mathbf{I} & (i > j, j \in \mathcal{N}(i)) \\ -\mathbf{I} & (j > i, i \in \mathcal{N}(j)) \end{cases}$$

分散型深層学習: データ(x_i)がVノードに渡って**分散蓄積**している。最終的にノード間でモデルが一致する制約下(s.t. ...)で、**評価関数**を最小化するようにモデル変数(w_i)を更新

関連文献

- [1] T. Sherson, R. Heusdens, B. Kleijn, "Derivation and analysis of the primal-dual method of multipliers based on monotone operator theory," *IEEE transactions on signal and information processing over networks*, Vol. 5, 2, pp. 334-347, 2018.
- [2] K. Niwa, N. Harada, G. Zhang, B. Kleijn, "Edge-consensus learning: deep learning on P2P networks with nonhomogeneous data," submitted to KDD 2020

連絡先

丹羽 健太 (Kenta Niwa) コミュニケーション科学基礎研究所/メディアインテリジェンス研究所
Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

21 心臓らしい心臓モデル

物理法則拘束付きガウス過程回帰を用いた心臓のモデル

どんな研究

心臓の核磁気共鳴画像(Magnetic Resonance Imaging: MRI)から心臓の3次元形状とその時間的な形状変化を推定します。この心臓形状推定問題は(1)画像のセグメント分割問題と(2)セグメントからの3次元形状の回帰問題に分けることができます。本研究では特に後者の回帰問題に注力しました。

どこが凄い

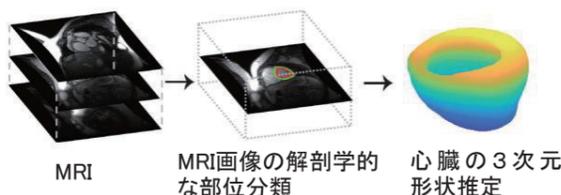
心臓の形状推定問題において、MRI画像から得られる見た目の手がかりとは別に心臓の形状変化は全身に血液を送るポンプのような動きをするという物理法則を隠れた手がかりとして新たに活用するアイデアを導入し、これを拘束付き3次元形状回帰問題として解く方法を考案しました。

めざす未来

心臓病は重篤な症状を引き起こす疾患であり、世界の成人人口の約1~2%が何らかの心疾患の影響を受け、特に70歳以上の罹患率は10%以上にもなるといわれています。病気の予防・早期診断に応える医療技術の今後の発展の一要素として、機械学習の活用を図っていきたく考えています。

心臓の三次元形状推定問題

核磁気共鳴画像からの心臓の三次元形状の推定



セグメント分割を解く深層学習

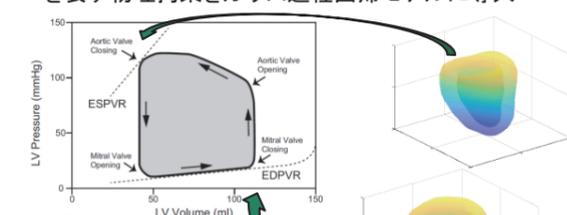
形状推定問題を解く回帰モデル学習

最新の深層学習手法を利用 [2]

本研究の焦点: ガウス過程回帰モデルに対して新たに心臓の物理的拘束を導入 [1]

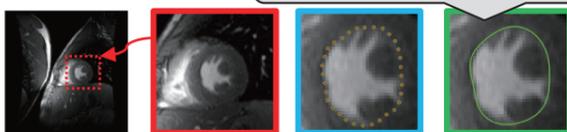
物理法則による拘束付き回帰モデル

Frank-Starling法則と呼ばれる「心臓らしい心臓の動き」を表す物理拘束をガウス過程回帰モデルに導入



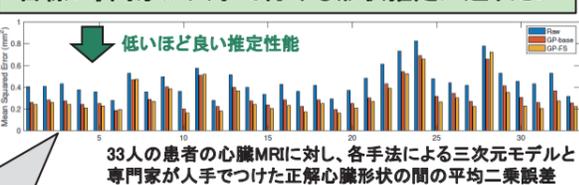
隠れた(モデルが推定すべき)心臓の圧力と体積の間の物理法則を隠れマルコフモデルで表現

提案モデル(右)は専門家による人手ラベル(中央右)に近い形状を推定



左: 観測(入力)MRI。中央左: 心内膜部拡大図。中央右: 専門家が人手で描いた心内膜。右: 提案モデルによって推定された心内膜。

目標: 専門家が人手で付ける形状推定に迫りたい



33名の患者中29名で形状推定精度が改善。全患者平均二乗誤差で0.173→0.158mm²の性能改善。

- Raw: 回帰モデルを介さないセグメント分割
- GP-base: 物理法則拘束なしの従来のガウス過程回帰
- GP-FS: 提案法、物理法則拘束付きガウス過程回帰

問題の難しさと本研究の貢献

- 真の心臓形状を知ること自体が容易ではない(専門家が膨大な手間をかけ人手で画像にラベル付け)。教師なし回帰モデルに物理法則拘束を付けることで代替
- 心臓の詳細な細部を捉えることが本質的に難しい(核磁気共鳴画像自体がもともと心臓の時間的・空間的に局所的な統計の手がかりしか与えてくれない)。時系列モデルを使って、ある時刻の形状手がかりを別の時刻へも伝播

関連文献

[1] M. Nakano, R. Shibue, K. Kashino, S. Tsukada, H. Tomoike, "Gaussian process with physical laws for 3D cardiac modeling," under review.
 [2] T. Ngo, Z. Lu, G. Carneiro, "Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance," *Medical Image Analysis*, Vol. 35, pp. 159-171, 2019.

連絡先

中野 允裕 (Masahiro Nakano) 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ー
 Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

22 あなたの鼓動に耳を澄ます

音響観測に基づく血流動態の解析

どんな研究

人体からは生命活動に伴い常に様々な音が発生しています。“聴診”と呼ばれるように、その音を聞き取り、解析することにより、身体の機能や状態に関する有用な情報を得ることができます。本研究では特に心音に着目し、音響信号の観測に基づく心臓・血管の機能や状態の推定に取り組んでいます。

どこが凄い

マイクロホン胸部など数か所に装着して心臓の活動をとらえます。とらえた音をもとに、異常の兆候の無い度合い(正常度)を推定するとともに、音の特徴を説明する文章(説明文)を生成します。正常度の推定と指定した詳しさでの説明文生成がテストデータに対し有効に機能することを確認しました。

めざす未来

専門病院では高度な検査機器や様々な検査手法により心臓や血管の状態の詳細な把握が可能になっていますが、熟練した医師は、聴診から的確に病態を把握できるといいます。機械学習と対象モデルの両面から、多くの人の病気の予防や早期発見に資する“AI聴診器”の実現をめざします。

AI聴診器のコンセプト

- 複数の小型マイクロホンを胸部に装着すれば、有用な情報を聞き取って様々に可視化(見える化)
- 異常の有無や病名の推定(従来のパターン認識)だけではなく、深層学習技術の活用により、例えば心音の特徴を表す説明文の生成が可能 [1]
- 今後物理モデル [2]【展示21】・信号モデルとの統合を進め、私たちの“デジタルツイン”における心臓や血流動態の可視化・解析ツールとして有効性を追求

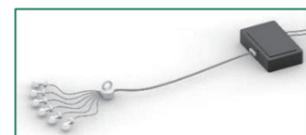


図1: 心音収集部のプロトタイプ

系列変換モデルによる説明文やスコアの生成

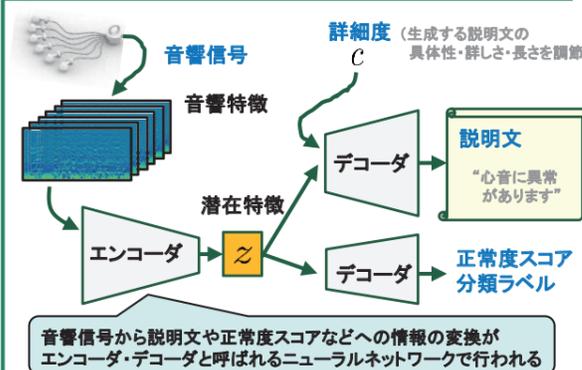


図2: 本研究における系列変換モデルの構成例 [1]

関連文献

[1] S. Ikawa, K. Kashino, "Neural audio captioning based on conditional sequence-to-sequence model," In *Proc. DCASE 2019 Workshop*, 2019.
 [2] M. Nakano, R. Shibue, K. Kashino, S. Tsukada, H. Tomoike, "Gaussian process with physical laws for 3D cardiac modeling," under review.
 [3] The PhysioNet Computing in Cardiology Challenge, <http://physionet.org/content/challenge-2016/1.0.0/>, 2016.

連絡先

柏野 邦夫 (Kunio Kashino) 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ー
 Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

AI聴診器プロトタイプ動作例

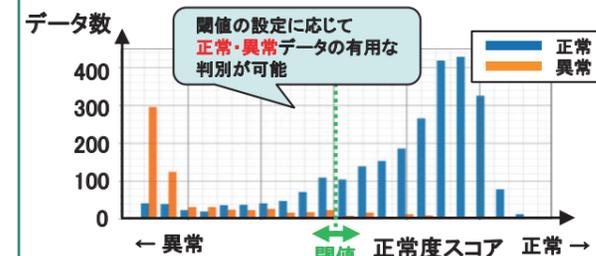


図3: テストデータ [3] に対する正常度スコアの分布

表1: 説明文の生成例

詳細度	説明文の生成例	詳細度に応じて有用な説明文が生成されている
20	心音に異常があります。	
40	心音に収縮期雑音があります。	
60	心音に収縮期雑音があり、心筋症の兆候の可能性があります。	

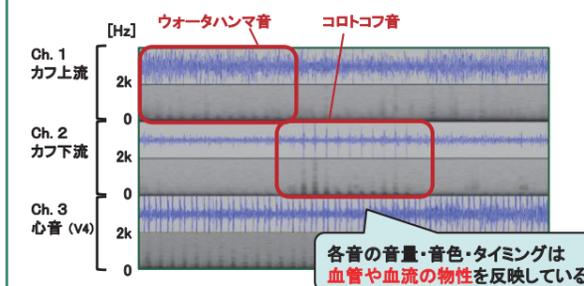


図4: 左上腕にて血圧測定中の心音・血流音の測定例 (各 Ch. において上: 音響信号波形 下: スペクトログラム)

本研究の一部は、NTT (バイオメディカル情報科学研究センター: BMC) から NTT Research, Inc. (生体情報処理研究所: MEI Lab.) への委託研究として実施しています。

23 ヒト知覚モデルで「自然な」錯覚をつくる

「不自然さ」予測に基づく「変幻灯」の視覚運動量最適化

どんな研究

CS研が考案した「変幻灯」は実物に動き的印象を与える錯覚ベースのプロジェクションマッピングです。しかし、人の視覚を上手に「だます」ためには、変幻灯を設置する環境毎に与える動きの量を適切に調節する必要があります。本研究では、この動きの調節作業を自動化する方法を提案しました。

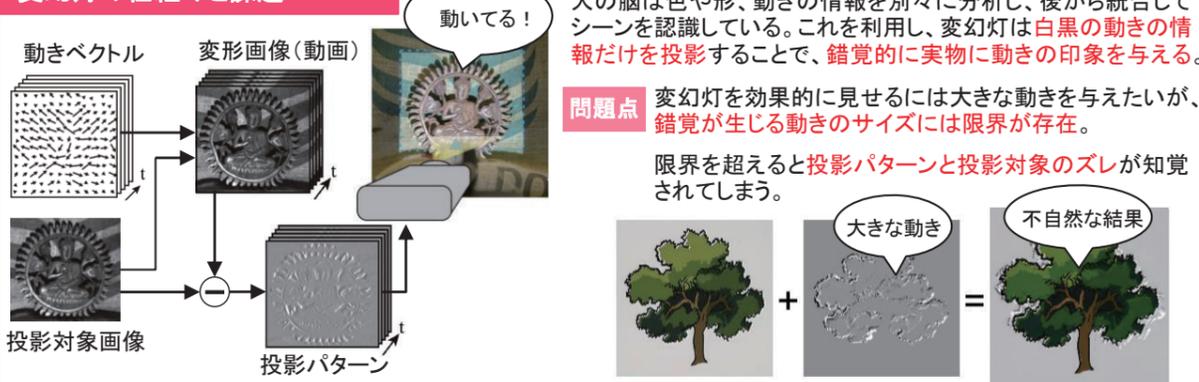
どこが凄い

人の視覚情報処理の計算モデルを用いて投影結果の「不自然さ」を予測できることを示しました。この予測モデルを用いて、動きの量をリアルタイムに最適化する手法を考案しました。これにより、インタラクティブなアプリケーションへの応用など、変幻灯の活用が広がります。

めざす未来

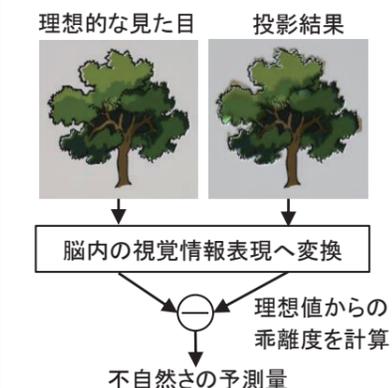
変幻灯のように、錯覚を利用することでこれまでは実現不可能であった表現を可能にするメディア技術は他にも考えられます。今後は、「人の知覚のモデルで錯覚を制御する」という枠組みをさらに発展させ、錯覚ベースの表示手法の普及に貢献していきます。

変幻灯の仕組みと課題



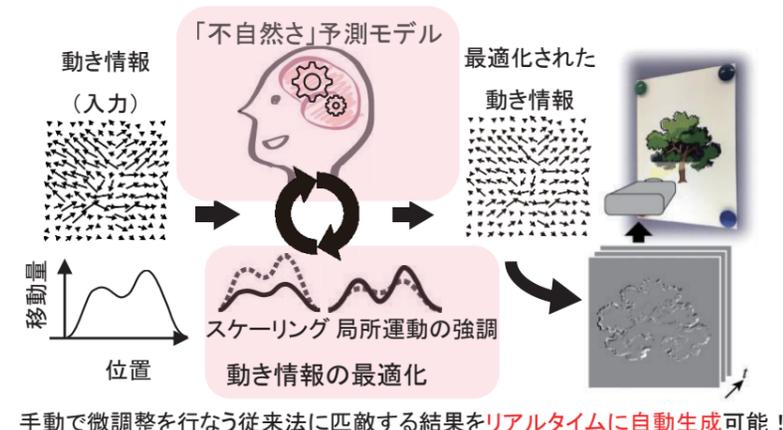
投影結果の「不自然さ」の予測

脳内の視覚情報表現内で投影結果が理想的な見た目からどれくらい乖離しているかを計算し、不自然さの指標とする。



「不自然さ」予測モデルに基づく動き情報の最適化

投影パターンと投影対象のズレが不自然に感じられない範囲で最大の動きが得られるように動き情報を最適化。



関連文献

[1] T. Fukiage, T. Kawabe, S. Nishida, "Perceptually based adaptive motion retargeting to animate real objects by light projection," *IEEE Transaction on Visualization and Computer Graphics*, Vol. 25, No. 5, pp. 2061-2071, 2019.
 [2] T. Kawabe, T. Fukiage, M. Sawayama, S. Nishida, "Deformation Lamps: A projection technique to make static objects perceptually dynamic," *ACM Transactions on Applied Perception*, Vol. 13, No. 2, pp. 1-17, 2016.

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



オープンハウス 2020

24 微小な眼球運動から垣間見る認知状態

眼球運動の動特性と認知タスク・瞳孔径の関係

どんな研究

我々の眼球は、視覚刺激の情報を取り入れるセンサーであるだけではありません。眼球の運動特性はその時々々の認知状態を反映します。本研究では、視線移動タスク中に生じる眼球運動の詳細な動特性が、タスクの難しさや瞳孔径の変化とどのように関連するかを調べました。

どこが凄い

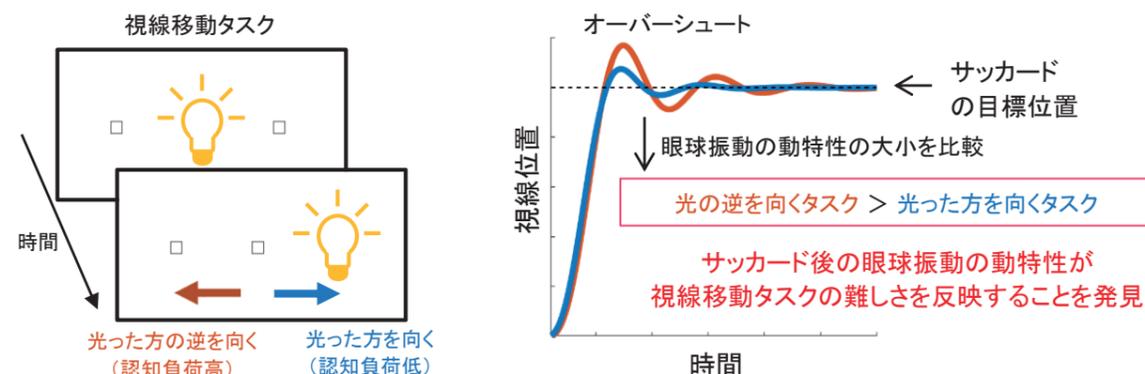
これまで、視線や瞳孔の変化から認知状態を知ろうとする試みは多く行われてきました。本研究では、これまで見過ごされてきた微小な眼球運動を詳細に解析することにより、その動特性が認知状態の影響を受けることを発見しました。

めざす未来

無意識に生じる微小な眼球運動の計測に基づいて、本人も意識していないようなその時々々の認知状態(認知的負荷、注意の方向など)を読み取る技術(マインドリーディング技術)の発展をめざします。作業者の集中度合いの評価やマーケティングなどにも応用できる可能性があります。

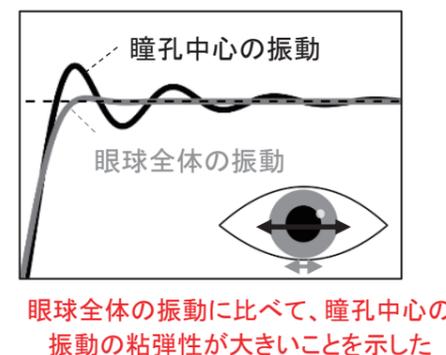
眼球運動(サッカド)の動特性と認知タスクの関係

- サッカド後の視線は目標位置でぴたり止まるわけではなく、目標位置周辺で振動する(オーバーシュート)
- 単なる機械的振動とされてきたオーバーシュートと認知状態との関連を検証



動特性から眼球の物性(粘弾性)がわかる

- 画像処理により瞳孔中心と虹彩中心の位置を抽出



関連文献

[1] S. Yamagishi, M. Yoneya, S. Furukawa, "Relationship of postsaccadic oscillation with the state of the pupil inside the iris and with cognitive processing," *J Neurophysiol*, Vol. 123, pp. 484-495, 2020.

連絡先

山岸 慎平 (Shimpei Yamagishi) 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



オープンハウス 2020

25 触ると似てしまうテクスチャ

3Dプリンタを用いた触り心地を変えないテクスチャ変調

どんな研究

異なる凹凸パターンを持ち、目で見ると区別がつくのに、指で触ると手触りの違いが判らなくなるような表面テクスチャを設計しました。素材表面の凹凸パターンについて、深さを輝度値に変換してモノクロ画像として扱うことで、3次元のテクスチャを2次元の画像のように解析・変調しています。

どこが凄い

複雑な統計量を画像上で操作し、作成した画像を元に3Dプリントすることで、これまで計測・実現が困難だった実物体の統計量を簡易に操作できます。その結果、触覚による知覚に影響しない統計量を発見しました。視覚と触覚では知覚に影響する統計量が異なることも明らかになりました。

めざす未来

テクスチャ知覚の決め手となる表面形状の特徴が理解できれば、画像特徴を操作するだけで簡単に物体表面の質感を変えることが可能になるかもしれません。また、見た目と触り心地を分けて設計できるため、より自由なテクスチャ設計が可能になります。

3次元の表面テクスチャを2次元画像のように扱い 画像処理によって表面の手触りを操作

- 素材表面の凹凸パターンについて、彫りの深さを輝度値に変換してモノクロ画像として扱うことで、画像処理の技術を実物体に対して適用可能に
- 従来の粗さ指標よりも複雑な統計量を簡易に操作することで、触り心地に影響する低次統計量(振幅スペクトル)・影響しない高次統計量(位相スペクトル)を発見



似ているけれど 区別はつく

5枚見比べると はっきり異なる 見分けがつく

輝度の平均・分散を揃えた異なる画像
低次統計量 **不一致** 高次統計量 **不一致**

明るい点: 凸 暗い点: 凹

↓ 画像処理によって 低次統計量を一致させる ↓

低次統計量 **一致** 高次統計量 **不一致**

画像処理後も 見分けがつく

触覚は視覚に比べ低次統計量の細かな違いに感度が高く 視覚は触覚より高次の統計量、位相などに感度を持つ

▶ 手触りを一定に保ちながら見た目のデザインを変更することが可能に

関連文献

[1] S. Kuroki, M. Sawayama, S. Nishida, "Haptic metameric textures," *bioRxiv*, 653550, 2019.

連絡先

黒木 忍 (Scinob Kuroki) 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



オープンハウス 2020

26 情動はいつ変化するのか？どうやって測るのか？

実験室環境と日常生活における情動変化の計測

どんな研究

私たちの情動は、他者とのインタラクションをはじめとする外的環境と自分自身の内的変化による影響を受けて変化します。本研究では、実験室環境においてインタラクションが情動変化に及ぼす影響を調べるとともに、日常生活における情動の内的変化を評価する新しい枠組みを提案します。

どこが凄い

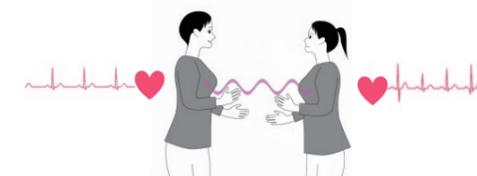
共同体験中の自律神経系の生体反応を計測する実験により、他者の情動から受ける影響がネガティブ情動とポジティブ情動で異なることを明らかにしました。日常生活のなかで起こる情動の内的変化を感性表現語(感嘆詞やオノマトペ)を用いて継続的に記録するセルフトラッキング方法を考案しました。

めざす未来

人と人のインタラクションによる情動変化の機序を解明することで集団パニックや盛り上がりのような社会現象のモデル化に貢献します。日常生活における情動の内的変化のパターンやそれに関わる要因を明らかにすることは、ウェルビーイング向上をサポートする仕組みづくりに役立つと期待されます。

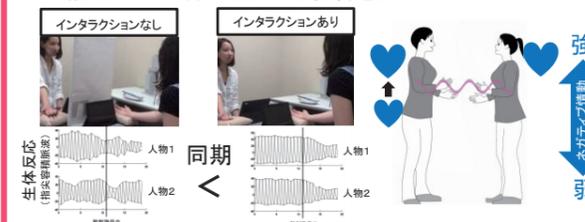
実験室における共同体験中の情動反応の計測

他者とのインタラクションによってどう変化するのか？



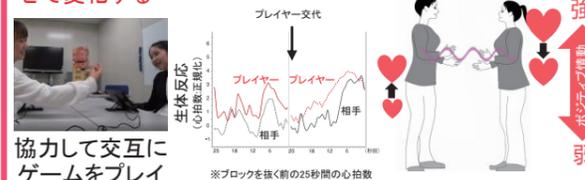
実験1 ネガティブ情動 熱刺激による痛みの共同体験

ネガティブ情動はインタラクションによって同期し、自分よりも相手の情動反応が強いときはそれに合わせて強まるが、相手の情動反応が弱いときには影響を受けない



実験2 ポジティブ情動 ブロックゲームによる興奮の共同体験

ポジティブ情動はインタラクションによって同期し、自分よりも相手の情動反応が強いときも弱いときも相手の反応に合わせて変化する



関連文献

[1] A. Murata, H. Nishida, K. Watanabe, T. Kameda, "Convergence of physiological responses to pain during face-to-face interaction," *Scientific Reports*, Vol. 10, 450, 2020.
[2] A. Murata, S. Kumano, J. Watanabe, "協力場面における対人インタラクションの当事者評価と客観評価," *電子情報通信学会技術研究報告*, Vol. 118, No. 487, pp. 111-114, 2019.

連絡先

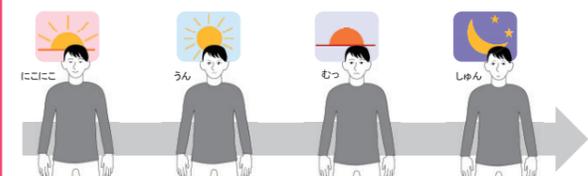
村田 藍子 (Aiko Murata) 人間情報研究部 感覚共鳴研究グループ
Email: cs-openhouse-ml@hco.ntt.co.jp



オープンハウス 2020

日常生活における情動のセルフトラッキング

日常生活の中でどのような時間的変化を辿るのか？

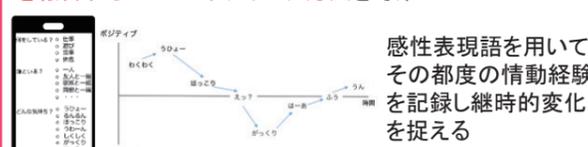


従来手法 感情を表す形容詞ごとの数値評価

- 嬉しい 0:全く感じない~10:とても感じる
- 悲しい 0:全く感じない~10:とても感じる
- ・ 認知的負荷が高い(直感的に回答しにくい)
- ・ 身体的な情動経験を捉えることが難しい
- ・ 正直に回答しにくいことがある(否定的感情など)

提案手法 感性表現語を使った報告

従来手法の問題を解消するため、身体的感覚の表現に適した感嘆詞やオノマトペなどの感性表現語により情動経験を報告するセルフトラッキング方法を考案



感性表現語を用いてその都度の情動経験を記録し継続的な変化を捉える

感性表現語に基づく情動評価を可能にするため、約1万4千人を対象とした大規模調査により情動の種類・強度と感性表現語の対応表を作成

喜び	悲しみ	予感	驚き	怒り	恐怖	嫌悪	好感
強 うひょー	弱 うひょー						
弱 うひょー							

関連文献

[1] A. Murata, H. Nishida, K. Watanabe, T. Kameda, "Convergence of physiological responses to pain during face-to-face interaction," *Scientific Reports*, Vol. 10, 450, 2020.
[2] A. Murata, S. Kumano, J. Watanabe, "協力場面における対人インタラクションの当事者評価と客観評価," *電子情報通信学会技術研究報告*, Vol. 118, No. 487, pp. 111-114, 2019.

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

27 eスポーツ達人の脳力

パフォーマンス、身体の生理状態、脳活動の相互関係

どんな研究

肉体的要因で結果が左右されにくいeスポーツでは、**選手の実力・コンディション・基礎能力(視覚能力)**を把握することが困難でした。本研究は、脳科学的アプローチによって、**eスポーツ選手のこれらの特性を客観的に評価しました。**

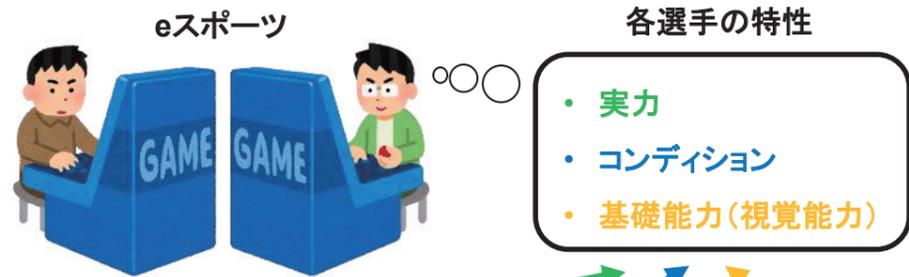
どこが凄い

eスポーツ選手のパフォーマンスと身体の生理状態・脳活動の関係を調べることで、**試合中の心拍同期が選手の実力を反映することが分かり、未来の勝敗と関連する脳の状態が見つかりました。**また、視覚科学の研究成果に基づき、**視覚情報処理能力と選手の強さの関係にも注目しています。**

めざす未来

eスポーツ競技に適した認知機能と生理状態を、多角的なアプローチから明らかにし、その知見に基づいて、**eスポーツ選手が実戦環境でパフォーマンスを向上させるための脳科学的な方法論を確立します。**

eスポーツと選手特性



アプローチ

心拍計測

心拍同期が実力を反映[1]

脳波計測

未来の勝敗と関連する脳の状態を発見

視覚テスト

強さにつながる視覚情報処理能力の追究

関連文献

[1] K. Watanabe, N. Saijo, M. Kashino, "The across-player correlation of the physiological change reflecting the fight-or-flight response in esports," in *Proc. The Society for Neuroscience 49th Annual Meeting*, 2019.
 [2] K. Hosokawa, K. Maruya, S. Nishida, M. Takahashi, S. Nakadomari, "Gamified vision test system for daily self-check," in *Proc. 2019 IEEE Games, Entertainment, Media Conference (GEM)*, 2019.

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



オープンハウス 2020

28 ラグビースクラムのハーモニーを紡ぐ

ウェアラブルセンサを用いた選手間協調の簡便な評価

どんな研究

チームスポーツでは選手たちが協調してプレーすることが重要ですが、そうした協調性を評価することは容易ではありません。本研究では、ラグビーのスクラムを対象に、選手たちがいかに力を結集して相手を押し上げているか、その**協調関係(ハーモニー)**をウェアラブルセンサを用いて分析します。

どこが凄い

練習中の選手の腰に無線慣性センサを装着し、得られる加速度データから、**各選手がいつ、どの位の強さで、どの方向に押そうとしたのか、そして全体としてどのように出力したかを簡便に評価します。**その結果は、手元のPCアプリにデータベース化され、その場で確認することができます。

めざす未来

日々の実践的なデータを蓄積することで、各選手の特徴や再現性などを把握できるだけでなく、**どのような選手の組み合わせが適切かを見極めることにもつながります。**また、この技術は、スポーツに限らず、様々な集団の行動(動きやリズムなど)を評価することにも応用できます。

簡便なウェアラブルセンシング

選手間協調をフィードバック

分析・データベース化

← 日々のトレーニング

加速度データの例

フィードバック画面の例

- ウェアラブルセンサで簡便に運動計測
- その場で協調関係(ハーモニー)を見える化
- データを蓄積し、選手の特徴や組み合わせの評価に活用
- スポーツ以外にも転用可能

ピークタイミングを音にするフィードバックも

良いスクラムは、低い音から高い音に聞こえる

* NTTコミュニケーションズ シャイニングアークスとの共同実験です

関連文献

[1] T. Kimura, N. Saito, H. Okamoto, K. Ohta, "IMUを利用したラグビースクラム時の選手間の協調関係の評価," in *Proc. スポーツ工学・ヒューマンダイナミクス2019*, 2019.

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



オープンハウス 2020

29 ストレートは“まっすぐ”か？

物理計測と知覚計測からピッチングを捉えなおす

どんな研究

スポーツ選手は感覚(知覚)は並外れて敏感ですが、物理世界をそのまま捉えているとは限りません。本研究では、野球の投球軌道の物理特性を正確に定量化した上で、その物理量を打者がどのように知覚しているのかを明らかにしました。

どこが凄い

物理計測では、ボールの回転特性をカメラ一台から簡便に計測する技術を考案しました。それを元に知覚計測では、軌道の変化量に対して、打者は細かな違いを判別できるにも関わらず、系統的にズレて(錯覚して)知覚すること、また錯覚の方向が投球フォームによって変化することを発見しました。

めざす未来

従来のスポーツ科学では、物理計測が重視される傾向にあり、選手の知覚とのズレが大きな問題となっています。物理計測と知覚計測を組み合わせることで、このズレの原因を解明することで、スポーツトレーニングや指導、戦略などに劇的な変化をもたらすことができます。

物理計測

「投球の回転軸の向き(投手1名の例)」
打者側からみたボールの平面図。回転軸の向き(各点)を計測し、1球ごとの軌道の変化を詳細に解析。

ほとんどのストレートは物理的にはシュートする

本計測技術は、プロ野球球団、ソフトボール女子日本代表、NTT東・西野球部などと実用に向けたトライアルを実施中です。

知覚計測

打者の知覚の例

物理的には曲がっている球を“まっすぐ”と錯覚

右投手か左投手かで、錯覚の向きが逆転

「ストレート」が“まっすぐ”に見える方向に知覚がバイアスされる。

関連文献

[1] D. Nasu, T. Kimura, M. Kashino, "Do baseball batters perceive straight ball trajectory as straight?" in *Proc. 2020 Conference on North American Society for Psychology of Sport and Physical Activity*, 2020.

連絡先
那須 大毅 (Daiki Nasu) 柏野多様脳特別研究室
Email: cs-openhouse-ml@hco.ntt.co.jp

オープンハウス 2020

30 巧みで素早い運動を支える脳の中の身体表現

手の位置推定の不確かさは伸張反射を調節する

どんな研究

人間には、目や手足などからの感覚入力によって運動目標や姿勢の変化をとらえ、その影響を無意識的な反応によって小さくする反射的な運動修正の仕組みが備わっています。このような反射的応答を状況に応じて調節するために、脳内でどのような情報処理が行われているかを調べています。

どこが凄い

急激な姿勢変化を修正するように働く「伸張反射」が、自分の体の動きが見えない状況では小さくなることを発見しました。このことから、反射的応答の調節が、従来考えられていたように単一の感覚情報にのみ基づくのではなく、複数の感覚情報を統合した身体表現を介し行われていると考えられます。

めざす未来

反射的な応答が生成・調整される仕組みを解明し、人間が巧みな運動を行う上で、意識に上らない運動制御がどのように役立っているかを明らかにしていきます。将来的には、アスリートの能力の分析・解明に役立てることや、スポーツトレーニングへの応用をめざします。

意識を介さない運動修正

※筋の急激な伸びに対して、筋肉を収縮させることで姿勢を修正する反射応答

意識・思考

意識を介した運動修正 (200[ms]以降)

伸張反射 (=体性感覚性の応答, 約50[ms])

筋への運動指令

体性感覚入力

- 姿勢の変化
- 外界との接触

反射的な応答によって、外乱や外界の変化の影響を素早く補償できる

反射的な応答を調整する情報処理

従来仮説: 単一の感覚情報による調整

新規仮説: 多感覚情報の統合による調整

視覚情報

体性感覚情報

多感覚統合

伸張反射の大きさ

実験: 視覚情報によって伸張反射が調整される

- 手首運動中に外力による姿勢変化を与え、伸張反射を計測
- 手の位置を示す視覚情報(カーソル)を消し、反射応答への影響を検証
- カーソルの消去時間が長いほど反射が小さくなることを発見
- 多感覚統合に基づく手の位置表現の不確かさが反射応答を調節することを示唆

実験設定

画面表示: 外力を加える位置、カーソル、目標

手首の動き: 運動方向、外力

筋活動を計測

結果

外力を加えた前後の筋活動

長潜時伸張反射

伸張反射の大きさ

p<0.05

関連文献

[1] S. Ito, H. Gomi, "Visually-updated hand state estimates modulate the proprioceptive reflex independently of motor task requirements," *eLife*, 9:e52380, 2020.

[2] S. Ito, H. Gomi, "Online modulation of proprioceptive reflex gain depending on uncertainty in multisensory state estimation," in *Proc. The Society for Neuroscience 49th Annual Meeting*, 2019.

連絡先
伊藤 翔 (Sho Ito) 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

オープンハウス 2020

31

意識より賢い無意識

環境に応じた顕在・潜在的視覚運動応答の調節

どんな研究

人間が複雑で多彩な動きを行うことができるのは、外界の状況を理解し、意識的に運動を制御する情報処理能力が高いからであると考えがちです。しかし実は、**巧みな運動は無意識的に制御されている**部分も多く、どのような情報処理が無意識に行われているかを明らかにしていくを試みています。

どこが凄い

姿勢安定性や視覚ノイズが異なる環境下で、視覚運動刺激に対して素早く対応する能力を随意・反射応答について調べたところ、**状況に適した調節ができるのは反射応答のみであることを発見しました。**このことは、**無意識的処理が意識的処理よりも状況適応能力が高い場合があることを示唆しています。**

めざす未来

感覚運動系における意識的情報処理と無意識的情報処理の異なる点を明らかにすることで、脳の情報処理のさらなる解明、通信・マンマシンインタフェースのデザイン、アスリートのトレーニングなどに活かしていくことをめざします。

日常における身体の動きと腕の制御

- 人は日常生活の中で、常に体を動かしながら、外界とのインタラクションを行っています。
- コンピュータに比べて伝達が遅い脳神経系で、こうした巧みな動きを実現するためには、どのような情報処理が行われているのでしょうか。



運動のための意識的・無意識的情報処理



⇔ 意識的情報処理
⇔ 無意識的情報処理

従来、運動生成に使われる無意識的処理は、意識的処理よりも、素早く反応できるものの、環境・状況に合わせた調節機能は低いと考えられてきました。

➡ **本研究は、無意識的処理の方が、意識的処理と比較して、より賢く環境・状況を認識し、適切に応答を調節できる場合があることを明らかにしました[1]。**

実験：視覚運動刺激に対する無意識な腕応答と意識的な応答における環境の文脈依存性

実験

- 姿勢を揺らす・視覚運動ノイズを入れる、という「環境の文脈」を実験的に与える。
- 視覚運動刺激に対する2種の応答を計測。
1. 反射応答 (Manual Following Response: MFR)
2. 随意応答 (Motion Direction Discrimination)

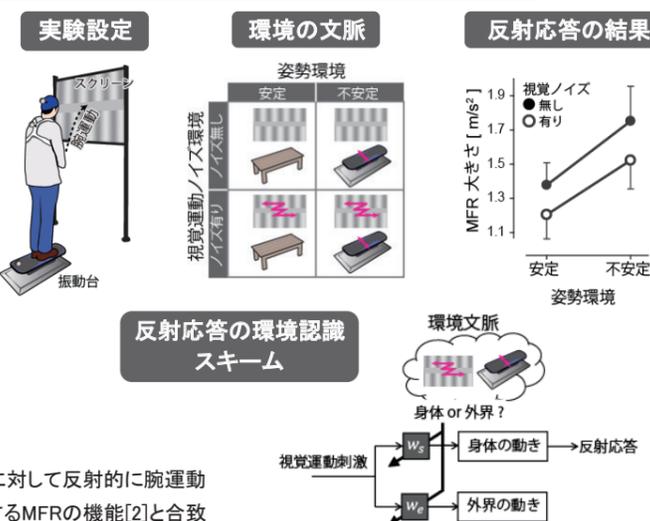
結果

1. 反射: 状況依存・合理的な調節を観察
2. 随意: 状況を無視した調節を観察

解釈

反射応答は、姿勢と視覚の動きを結び付けた賢い調節がされる一方、随意応答ではその結び付けができないことを示唆

身体が揺れやすい環境 ⇒ MFR大
外界が動きやすい環境 ⇒ MFR小
➡ 身体の揺れに対して反射的に腕運動軌道を修正するMFRの機能[2]と合致



関連文献

- [1] N. Abekawa, H. Gomi, "Modulation difference in visuomotor responses in implicit and explicit motor tasks depending on postural stability," in *Proc. The Society for Neuroscience 47th Annual Meeting*, 2017.
 [2] H. Gomi, K. Kadota, N. Abekawa, "Dynamic reaching adjustment during continuous body perturbation is markedly improved by visual motion," in *Proc. The Society for Neuroscience 40th Annual Meeting*, 2010.

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



NTT コミュニケーション科学基礎研究所 オープンハウス 2020 実行委員会

- (委員長) 亀井 剛次
 (副委員長) 安部川 直稔
 (委員) 大國 智樹/村松 純/上田 大志/澤山 正貴