

### 研究展示

- データと学習の科学**
- 01 混み具合を学習して空いている経路を見つけます 二分決定グラフを用いたオンライン最短経路アルゴリズム
  - 02 AIによる空調制御でもっと省エネ・もっと快適 深層強化学習による環境再現・最適制御技術
  - 03 都市の人口分布変動から人の移動を再現 時間別エリア人口データからの移動傾向推定技術
  - 04 深層学習のボトルネック解消で精度を向上 深層学習における、より高い表現能力を持つ出力関数
  - 05 データから学ぶ: どれが原因? どれが結果? 教師あり学習に基づく時系列の因果推論
  - 06 過去データの無い地点でも未来を予測 時空間データ解析のための時空間回帰テンソル分解法
  - 07 いろいろな観点で欲しいものを見つけます グラフ索引型探索技術による絵本検索システム「びたりえ」
- コミュニケーションと計算の科学**
- 08 限界まで効率よくメッセージを送れます シヤノン限界を達成する誤訂正符号
  - 09 新たな秘密がこれまでの秘密を脅かす 「量子情報を用いた秘密分散」の脆弱性の検証
  - 10 文や段落の修飾関係を賢く判断 ニューラルネットによる階層的トプダウン修飾構造解析
  - 11 ひらがなはいつからわかる? 幼児の文字習得メカニズムを探る
  - 12 「ところが動く、通う」をどう測る? 主観・生理・行動からみた共感的コミュニケーションの分析
  - 13 「触れて深まる体感と共感」を集団から測る 共感的コミュニケーションの触覚的促進と多数同時測定
  - 14 ロボットと話そう「いつ、どこで、何をしたら」 ユーザ発話中のイベント理解に基づく雑談対話システム
- メディアの科学**
- 15 うるさい車内でも音声操作や会話をサポート 世界トップの集音技術・音声認識技術を実用化
  - 16 少量の入出力ペアから高精度に音声認識を学習 音声合成を活用した半教師ありEnd-to-End学習
  - 17 いつ、誰が、何を話した? 全部で何人いた? 何人の会話でも聞き分けられる深層学習モデル
  - 18 声と話方を好みのスタイルに変える 系列変換モデルに基づく声質と韻律の同時変換
  - 19 顔に合わせて声を作り、声に合わせて顔を作る 深層生成モデルによるクロスモーダル音声変換
  - 20 声と画像から知らないモノを学びとるAI 音声と画像によるクロスモーダル概念獲得
  - 21 どんな音? 物音を言葉で説明しよう 系列変換モデルに基づく音響信号からの説明文生成
  - 22 音からものとかたちを認識 音と画像のクロスメディア情報処理によるシーン理解
- 人間の科学**
- 23 鳥の声で喋り、水の泡で演奏する 聴覚信号処理モデルを用いた音のテキスト変換
  - 24 踊る紙人形 紙に動きの印象を与える錯覚
  - 25 あなたの目の機能を気軽に楽しく測ります ミニゲームやタブレットを使った視覚能力のセルフチェック
  - 26 勝者のメンタルとは? 実戦環境での生理状態とパフォーマンスの関連性
  - 27 速いボールを捉える脳のしくみ 視覚と身体の瞬時的な連携
  - 28 「見守る」と「見守られる」をつなぎます 介護者記録用アプリ「みまもメイト」の記録共有による効果
  - 29 身体が「見る」実世界の動きとは? 自然統計で紐解く視覚運動制御の秘密
  - 30 座っているけど歩行感 疑似歩行感が身体近傍空間を拡張する



NTT コミュニケーション科学基礎研究所

# オープンハウス 2019

5月 **30** (木) 12:00~17:30

**31** (金) 9:30~16:00

入場料無料・事前登録不要

会場: NTT 京阪奈ビル B棟 3F・大会議室  
 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2019」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



昨今のAI技術の著しい進歩は人々に夢を与えると同時に、それがもたらす予測し難い変化の兆しは時として人々を不安にします。折しも、元号が新たに「令和」となり、オリンピック・パラリンピックを間近に控えた2019年は時代の変化点と言えるかもしれません。

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「ここまで伝わる」コミュニケーションの実現を目指し、時代を先取りした基礎研究に取り組んでいます。

「オープンハウス2019」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、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致します。特に今年は、元陸上選手の為末大氏と、弊社フェローの柏野牧夫氏が、スポーツと脳科学について語り合う、特別対談もご用意しております。

本イベントが皆様にとって、最新技術の体験を通じて、確かな未来の手応えを感じ取っていただく何らかの機会となれば幸いです。所員一同、皆様のご来場を心よりお待ちしております。



ごあいさつ ..... 1

会場案内図／講演スケジュール ..... 3

所長講演  
人に迫り、人を究め、人に寄り添う ～人とAIが共生し共創する未来へ～ ..... 4

特別対談  
スポーツの未来と人間の可能性 ..... 6

研究講演  
画像や音を見聞きするだけで賢くなるAI ～クロスモーダル情報処理がひらく未来～ ..... 8  
「見る力」を簡単に測る ～日常環境での視覚能力の簡易測定とそのヘルスケア応用に向けて～ ..... 10  
いろいろな観点で似たものを探す、似たものつながり検索法 ～絵本検索システム「びたりえ」とそれを支えるグラフ索引型探索技術～ ..... 12

研究展示 ..... 15

**データと学習の科学**

01 | 混み具合を学習して空いてる経路を見つけます 二分決定グラフを用いたオンライン最短経路アルゴリズム ..... 16  
02 | AIによる空調制御でもっと省エネ・もっと快適 深層強化学習による環境再現・最適制御技術 ..... 17  
03 | 都市の人口分布変動から人の移動を再現 時間別エリア人口データからの移動傾向推定技術 ..... 18  
04 | 深層学習のボトルネック解消で精度を向上 深層学習における、より高い表現能力を持つ出力関数 ..... 19  
05 | データから学ぶ：どれが原因？どれが結果？ 教師あり学習に基づく時系列の因果推論 ..... 20  
06 | 過去データの無い地点でも未来を予測 時空間データ解析のための時空間回帰テンソル分解法 ..... 21  
07 | いろいろな観点で欲しいものを見つけます グラフ索引型探索技術による絵本検索システム「びたりえ」 ..... 22

**コミュニケーションと計算の科学**

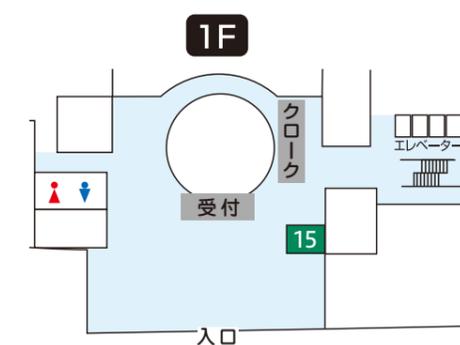
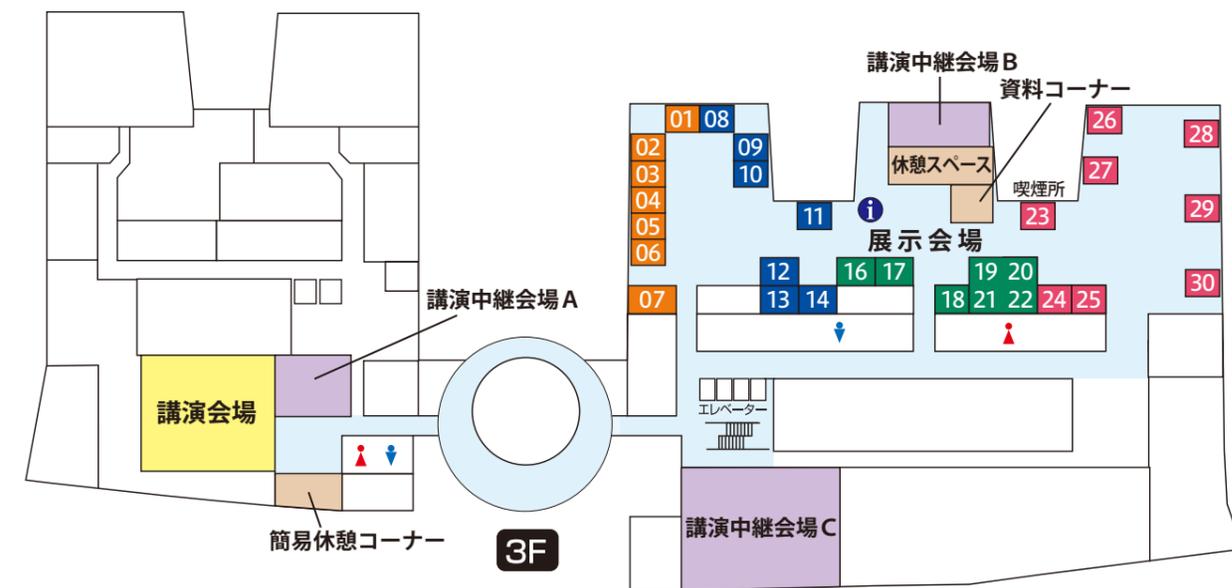
08 | 限界まで効率よくメッセージを送れます シャノン限界を達成する誤り訂正符号 ..... 23  
09 | 新たな秘密がこれまでの秘密を脅かす 「量子情報を用いた秘密分散」の脆弱性の検証 ..... 24  
10 | 文や段落の修飾関係を賢く判断 ニューラルネットによる階層的トップダウン修飾構造解析 ..... 25  
11 | ひらがなはいつからわかる？ 幼児の文字習得メカニズムを探る ..... 26  
12 | 「こころが動く、通う」をどう測る？ 主観・生理・行動からみた共感的コミュニケーションの分析 ..... 27  
13 | 「触れて深まる体感と共感」を集団から測る 共感的コミュニケーションの触覚的促進と多人数同時測定 ..... 28  
14 | ロボットと話そう「いつ、どこで、何をした」 ユーザ発話中のイベント理解に基づく雑談対話システム ..... 29

**メディアの科学**

15 | うるさい車内でも音声操作や会話をサポート 世界トップの集音技術・音声認識技術を実用化 ..... 30  
16 | 少量の入出力ペアから高精度に音声認識を学習 音声合成を活用した半教師あり End-to-End 学習 ..... 31  
17 | いつ、誰が、何を話した？全部で何人いた？ 何人の会話でも聞き分けられる深層学習モデル ..... 32  
18 | 声と話し方を好みのスタイルに変える 系列変換モデルに基づく声質と韻律の同時変換 ..... 33  
19 | 顔に合わせて声を作り、声に合わせて顔を作る 深層生成モデルによるクロスモーダル音声変換 ..... 34  
20 | 声と画像から知らないモノを学びとる AI 音声と画像によるクロスモーダル概念獲得 ..... 35  
21 | どんな音？物音を言葉で説明しよう 系列変換モデルに基づく音響信号からの説明文生成 ..... 36  
22 | 音からものとかたちを認識 音と画像のクロスメディア情報処理によるシーン理解 ..... 37

**人間の科学**

23 | 鳥の声で喋り、水の泡で演奏する 聴覚信号処理モデルを用いた音のテキスト変換 ..... 38  
24 | 踊る紙人形 紙に動きの印象を与える錯覚 ..... 39  
25 | あなたの目の機能を気軽に楽しく測ります ミニゲームやタブレットを使った視覚能力のセルフチェック ..... 40  
26 | 勝者のメンタルとは？ 実戦環境での生理状態とパフォーマンスの関連性 ..... 41  
27 | 速いボールを捉える脳のしくみ 視覚と身体の瞬間的な連携 ..... 42  
28 | 「見守る」と「見守られる」をつなぎます 介護者記録用アプリ「みまもメイト」の記録共有による効果 ..... 43  
29 | 身体が「見る」実世界の動きとは？ 自然統計で紐解く視覚運動制御の秘密 ..... 44  
30 | 座っているけど歩行感 疑似歩行感が身体近傍空間を拡張する ..... 45



- 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
- 通行可能領域
- 講演会場
- 講演中継会場A・B・C
- インフォメーション
- トイレ

### 講演スケジュール

**5月30日(木)**

**所長講演**  
(13:20-13:50)  
人に迫り、人を究め、人に寄り添う  
～人とAIが共生し共創する未来へ～  
コミュニケーション科学基礎研究所 所長 山田 武士

**特別対談**  
(15:00-16:00)  
スポーツの未来と人間の可能性  
元陸上選手 為末 大  
NTTフェロー・スポーツ脳科学プロジェクト統括 柏野 牧夫

**5月31日(金)**

**研究講演**  
(11:00-11:40)  
画像や音を見聞きするだけで賢くなるAI  
～クロスモーダル情報処理がひらく未来～  
メディア情報研究部 柏野 邦夫

**研究講演**  
(13:00-13:40)  
「見る力」を簡単に測る  
～日常環境での視覚能力の簡易測定と  
そのヘルスケア応用に向けて～  
人間情報研究部 丸谷 和史

**研究講演**  
(13:50-14:30)  
いろいろな観点で似たものを探す、  
似たものつながり検索法  
～絵本検索システム「びたりえ」と  
それを支えるグラフ索引型探索技術～  
協創情報研究部 服部 正嗣

# 人に迫り、人を究め、人に寄り添う

～人とAIが共生し共創する未来へ～

*Processing like people, understanding people, helping people*

*Toward the future where humans and AI will cohabit and co-create*

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

コミュニケーション科学基礎研究所（以下、CS研）は、人と人、あるいはコンピュータと人の間の「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います。最近のAI（人工知能）技術の発展にはめざましいものがあります。もともとコンピュータは人間が処理できない大量のデータを一度に処理し、人間が苦手な処理を人間に代わって高速に処理するのが得意です。しかし特に深層学習の発展のおかげで、本来人間が得意で、なかなかコンピュータが追いつけなかった音声や画像の認識や自然言語処理などにおいても、人間の能力に迫り、場合によっては凌駕する性能を実現しつつあります。今後さらにAIの進歩は加速すると期待されますが、とはいえAIの性能が複雑な人間の脳を超えるほどに進歩するのはまだ先と言えます。一方で人間は認知バイアスにとらわれ間違いを犯したり、実際にはありもしない錯覚をリアルに感じてしまうなど、複雑であるが故に一見すると不完全な存在でもあります。このように、限定された範囲で急速に発展を続けるコンピュータと、複雑であるがゆえに不完全でもある人間とをつなぎ、両者のギャップを埋めることが「心まで伝わる」を目指すCS研の使命と言えます。そのためには、人間の能力に迫る技術をさらに追求するのはもちろんのこと、人間の機能、特性を解明し、人間のことをよく理解すること、その上で人間に寄り添う技術の実現をめざすこと、そのための基礎研究が一層重要であると考えています[1]。

### 人間の能力に迫る技術

世の中にはまだまだ、人間は得意でも、コンピュータには苦手な処理が多数存在します。確かに機械翻訳の精度が上が

り、大学入試の英語穴埋め問題をある程度正解できるようにはなりましたが[2]、文章の意味を深く理解したり、常識を身につけたり、というレベルにはまだ達していません。一方で、深層学習技術を駆使することで、画像認識や音声認識など、特定の面では人間の能力に迫ってきたことも事実です。例えば、会議やパーティでの歓談などにおいて、複数の人が同時に話したり、背景に音楽が流れているとします。人間はこのような状況においても「聞きたい」人の声の特徴を選び分けて、話す内容を聞き取ることができます。これは人間の聴覚の優れた能力の一つで、選択的聴取と呼ばれます。従来、このような選択的聴取を、コンピュータは苦手でしたが、CS研では独自の深層学習技術により、人間同様、コンピュータが目的話者の声の特徴に基づき、その声だけを聞き取る技術を実現し、さらにそれを発展させています。

これらの技術が今後さらに進歩し、人間に近づく、その手掛かりの一つがクロスメディアです。従来、「音声」「映像」「テキスト」などのメディアはそれぞれ解析手法も異なり、別々に研究されてきました。しかしここに来て、深層学習といういわば「共通言語」が整備されたおかげで、メディアの種類を越えた（＝クロスメディア）「認識」「生成」「変換」が可能になりつつあります。人間は、例えば、音声を聞いただけでその場の情景をある程度頭の中に思い浮かべることができるなど、このような処理を日常生活の中で当たり前のように行っています。CS研はこれをコンピュータで実現する、言わば「音から画像認識」するクロスメディア情景分析技術に取り組んでいます。実現するとカメラでは死角になってしまうような箇所の情報も音を使って予測できるようになるでしょう。また、人間はTVの視

聴を通じて日々新たな知識を身につけています。コンピュータも将来はTV放送のようなメディアデータから、音と映像の共起を見つけながら自律的にモノやコトの概念を学習し賢くなる、それを目指した基礎研究にも取り組んでいます。

### 人間を深く理解し究める技術

このように特定の場面ではコンピュータの能力は人間に近づき、凌駕しつつあります。しかし、AIの性能が複雑な人間の脳を超えるほどに進歩するのはまだ先でしょう。一方で人間は、「振り込め詐欺」にも簡単に騙されるなど、時として「認知バイアス」に支配され、思わぬ誤りを犯します。また、人間の脳は錯覚します。CS研が運営するウェブサイト「イリュージョンフォーラム」には様々な錯覚が取り上げられています[3]。

複雑かつ不完全な人間と今のところ限定的なAI、この両者のギャップを埋めるためには、人間をもっと深く知ることが必要です。そのために、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。錯覚もその重要な手掛かりです。また、一流のアスリートを題材として、その優れた能力を脳科学の視点から解明し、人間の「心・技・体」の関係の本質に迫る、スポーツ脳科学にも取り組んでいます。例えば、優れた打者がわずか0.1秒という短い時間で、如何に遅い球と速い球を見極めて球種に応じたタイミングで動いているか、その仕組みの解明などに挑んでいます。スポーツ脳科学は、最新技術で主に体を鍛える、従来のスポーツ科学やスポーツ分析手法とは一線を画した、野心的な取り組みと言えます。

### 人間に寄り添う技術

スポーツ脳科学で得られる成果はスポーツに限らず、人間が普段の生活の中で心身の潜在能力を最大限に発揮する、すなわち、ウェルビーイングのための知見として生かすことができます。この、人間のウェルビーイングという、一見定性的でとらえどころのない課題を人間科学の立場から定量的に扱い、これを向上させるための設計指針を確立することにもチャレン

ジしています。例えば、複数の人間が、場を一緒に共有することによって生じる共感的コミュニケーションの効果測定などがその例です。

一方、錯覚は「潜在的な脳の働き」解明の手掛かりとして重要なのもちろんのこと、人間とAIとのギャップを埋め、人間に寄り添うインターフェースやフィードバックのための鍵でもあります。CS研ではこれまで人間の錯覚を利用したインターフェースとして、引っ張られる錯覚を生じさせるデバイス「ぶるなび」を考案しました。さらには、座っているのにあたかも歩いているような感覚の生成にも取り組んでいます。また、印刷した絵や写真に光を当てただけで動き出して見える「変幻灯」、3Dメガネを掛けると3D映像に、メガネを外すと鮮明な2D映像を楽しめる「Hidden Stereo」、印刷物などの2次元平面上の対象に対して影に見えるパターンを投影することで、その対象があたかも3次元的に浮き上がって見える光投影技術「浮像（うくぞう）」などを次々と生み出してきました。これからも、錯覚を利用した新たなインターフェースの提案と同時に、錯覚を通して物理的には生じ得ない体験を生み出すことによる、斬新な知覚表現の可能性も追究していきます。

### おわりに

昨今のAI技術の著しい進歩は人々に夢を与えると同時に、それがもたらす予測し難い変化の兆しは、例えばAIが職を奪うのでは、など、時として人々を不安にします。それ以外にも、「地球温暖化」には科学的根拠がないとか、果ては、「地球平面説」を信じる人まで現れる始末です。日本においては、2019年は折しも、元号が新しくなり、オリンピックを間近に控えた年です。我々はいろいろな意味で、時代の変化点にいるのかも知れません。今後ますます技術の進歩のスピードが速くなり、競争が厳しくなる中で、CS研は、人に迫り、人を究め、人に寄り添う技術を中心に、これからも新たなチャレンジに大胆かつ粘り強く取り組んでまいります。本オープンハウスがみなさまにとって、最新技術の体験を通じて、確かな未来の手応えを感じ取っていただく何らかの機会となれば幸いです。

### 関連文献

- [1]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, NTT技術ジャーナル 2018.9, <http://www.ntt.co.jp/journal/1809/files/JN20180908.pdf>
- [2]「ロボットは東大に入れるか」における英語問題の回答手法, NTT技術ジャーナル 2015.4, <http://www.ntt.co.jp/journal/1504/files/jn201504063.pdf>
- [3]錯視と錯聴を体験! Illusion Forum イリュージョンフォーラム, <http://www.kecl.ntt.co.jp/IllusionForum/>



# スポーツの未来と人間の可能性

## Sports in the future and human potentiality

スポーツはそれ自体大いなる楽しみであるとともに、人間の本質を垣間見ることのできる窓でもある。可能性と限界、成長と破壊、多様性とコンプレックス、才能と鍛錬、共感と断絶、信頼と盲信、自覚と無自覚、精神と肉体、…。スポーツの実践や分析を通じて見えてくる人間の本質とは。そして、ICT、AI、認知科学、脳科学といった科学技術が、スポーツを、そして人間をどのように変えていくか。アスリートと研究者の立場から縦横に語り合う。



元陸上選手・Deportare Partners 代表

**為末 大**

*Dai Tamesue*

**(プロフィール)**

1978年広島県生まれ。スプリント種目の世界大会で日本人として初のメダル獲得者。男子400メートルハードルの日本記録保持者(2019年3月現在)。現在は、Sports × Technologyに関するプロジェクトを行う株式会社Deportare Partnersの代表を務める。新豊洲Brilliaランニングスタジアム館長。主な著作に『走る哲学』、『諦める力』など。

NTTフェロー・スポーツ脳科学プロジェクト統括

**柏野 牧夫**

*Makio Kashino*

**(プロフィール)**

1964年岡山県生まれ。1989年東京大学大学院修士課程修了。博士(心理学)。1989年NTTに入社、コミュニケーション科学基礎研究所人間情報研究部長などを経て、2018年NTTフェロー。Wisconsin大学客員研究員、東京工業大学工学院情報通信系 特任教授、JST CREST 研究代表者等を歴任。専門は心理物理学・認知神経科学。平成28年度文部科学大臣表彰科学技術賞(研究部門) 受賞。著書に『音のイリュージョン～知覚を生み出す脳の戦略～』(岩波書店、2010)、『空耳の科学-だまされる耳、聞き分ける脳』(ヤマハミュージックメディア、2012) 他。



Series of horizontal dashed lines for taking notes.

# 画像や音を見聞きするだけで賢くなるAI

～クロスモーダル情報処理がひらく未来～

See, Hear and Learn to Describe

AI to sketch the world for semantic media search and conversion



メディア情報研究部  
**柏野 邦夫**  
Kunio Kashino

### (プロフィール)

NTTコミュニケーション科学基礎研究所 メディア情報研究部長、上席特別研究員。1995年東京大学大学院工学系研究科博士課程修了。博士(工学)。同年、日本電信電話株式会社に入社、現在に至る。東京大学大学院情報理工学系研究科客員教授、国立情報学研究所客員教授を兼任。IEEE、ACM、電子情報通信学会、情報処理学会、人工知能学会、日本音響学会の会員。

近年の人工知能(AI)の発展を支える立役者は、何と言っても深層学習の技術でしょう。例えば、様々な物体を撮影した画像と「りんご」「みかん」といった物体の名前(クラスラベル)とを組(ペア)にしたデータを大量に用意して深層学習を行うと、画像中の物体が何であるかを高い精度で認識できるようになることが知られています。その優れた特性のために様々な分野で研究や活用が進む深層学習ですが、私たちが特に着目している能力の一つは、異種のメディア情報(例えば、画像と音)の対応づけができることです。画像、音、テキストといった情報の種類のことをモダリティ(modality)と言いますので、異なるモダリティにまたがる情報の対応づけをクロスモーダル(crossmodal)情報処理と呼ぶことにします。では、このクロスモーダル情報処理とはどのようなもので、どんなメリットがあるのでしょうか。

### 新しい情報変換—音から画像を作る

クロスモーダル情報処理のメリットの一つは、異種のメディア情報が対応づけられた共通の場所である「共通空間」を介

することで、従来は考えられなかったような情報の変換が可能なことです(図1)。その一つとして、私たちの研究チームでは、音から画像を推定する課題に取り組んでいます。私たち人間は、目を閉じていても周囲の音からその場の情景を思い浮かべることができます。そこで、マイクで拾った音からその場の情景を表す画像を作ってみようというわけです。現在、一定の条件下で、どこに何がありそうかといった大まかな情報を画像として示すことが可能になっています[1](展示22)。この技術が発展すると、カメラを置くことが望まれない場所やカメラがとらえきれない状況(物陰や暗闇など)での安全確認などに応用できると考えています。

### 新しい情報変換—物音を言葉で説明する

異種情報の変換のもう一つの例は、音からテキストへの変換です。音声認識システムを用いると話し言葉をテキストに変換できますが、これまでは、音声認識システムを用いても話し言葉以外の物音などを適切なテキストに変換することはできませんでした。これに対し私たちは、マイクで拾った音から、

その音を表現する擬音語や、その音を記述する説明文を生成する技術を開発しました[2](展示21)。この技術を用いると、「車」「風」などといったクラスラベルだけではなく、音の高さや大きさ、変化の様子なども含めて、文字によって素早く把握することが可能になります。

### 概念獲得—未知の概念を自ら学習する

クロスモーダル情報処理のもう一つのメリットは、「共通空間」において異種情報の対応を見出すことで概念獲得が可能になることです。深層学習に必要とされる大量のデータの準備には、手間がかかったり、データの入手自体が難しかったり、クラスラベルのつけ方を事前に設計することが難しかったりといった困難さを伴うことが少なくありません。そこで私たちは、メディア情報の中に含まれるひとまとまりのもの、つまり概念を自動的に獲得し、認識や検索に活用することを目指した研究に取り組んでいます[3]。

異種のメディア情報の共起、つまり現実世界の中で、同じものに端を発する異種のメディア情報がランダムにではなく特定の関係性をもって現れることなどを利用すると、人手でメディアデータ同士をペアリングすること無しに、共通空間を介したメディアデータのペアリングが可能(展示20)。これを用

いると、事前に「りんご」の画像とクラスラベルのペアを与えなくても、「皆がこの物体を指して「りんご」と言っているようだ。これは「りんご」というものなのだ」といった方式での学習が可能になるのです(図2)。画像や音を見聞きするだけで賢くなる、というわけです。しかも周囲が「りんご」と言えばりんご、「apple」と言えばappleであると学習するといったように、周囲の人間の感じ方や振る舞い方を習得していきます。これは、私たち人間が、生まれてから成長するにつれて日常生活の中で様々なことを学んでいく過程にたとえることができるでしょう。

これらの一連の研究に共通する考え方は、音、画像、テキストといった様々なメディア情報に対して、私たちの目や耳に触れる表層の表現形式と、その背後にある共通空間、つまり特定の表現形式には依存しない本来的な情報とを、分離して取り出してそれぞれを活用しようということです。このような研究が発展すれば、私たち人間と共に暮らし、感じ方や振る舞い方を共有しながら、自ら学習していくAIも実現できそうに思われます。そのようなAIは、今よりもっと親しみを感じられるパートナーになり得るのではないのでしょうか。

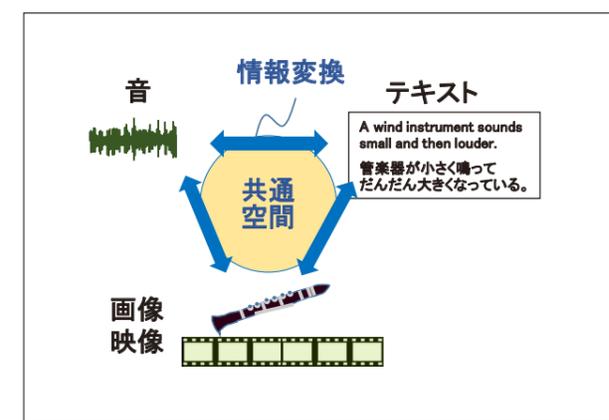


図1 クロスモーダル情報処理による情報変換(概念図)

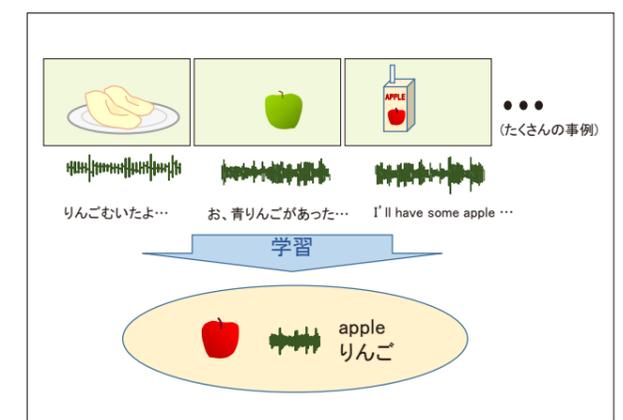


図2 クロスモーダル情報からの概念獲得

### 関連文献

[1] G. Irie, M. Ostrek, H. Wang, H. Kameoka, A. Kimura, T. Kawanishi, K. Kashino, "Seeing through sounds: Predicting visual semantic segmentation results from multichannel audio signals," in Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2019.  
[2] S. Ikawa, K. Kashino, "Generating sound words from audio signals of acoustic events with sequence-to-sequence model," in Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2018.  
[3] 大石康智, 木村昭悟, 川西隆仁, 柏野邦夫, D. Harwath, J. Glass, "画像を説明する多言語音声データを利用したクロスモーダル探索," 電子情報通信学会技術研究報告 パターン認識・メディア理解 (PRMU), 2019.

# 「見る力」を簡単に測る

～日常環境での視覚能力の簡易測定とそのヘルスケア応用に向けて～

## Measuring multiple visual abilities in daily circumstances

Towards establishment of daily selfcheck for eye health



人間情報研究部  
**丸谷 和史**  
Kazushi Maruya

### (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。2004年東京大学大学院人文社会系研究科博士課程修了。博士(心理学)。専門は知覚・認知心理学。錯視を含む視知覚のメカニズムの研究、また、人間の知覚・認知特性を利用したインタフェース技術の研究などに従事。日本心理学会、日本基礎心理学会、日本視覚学会、日本図学会、米国視覚科学協会(Vision Sciences Society)各会員。

近年、医療や教育の現場では、それぞれの個人が抱える問題に個別にフォーカスを当てたきめ細かい対応が、より強く求められています。それらの社会的要請に答えるために、コミュニケーション科学基礎研究所で継続的に取り組んできた人間科学も大きく姿を変えつつあります。実験室という整った環境で比較的少数の人から得た精密で深いデータを用いた検討だけでは、環境や課題、個人に応じて複雑にその働き方が変化する人間の視覚システムの理解に限界があります。そこで私たちは、より日常生活に近い環境で、様々な人から得られる「広いデータ」を用いた多面的な検討に取り組んでいます(図1)。

「広いデータ」の解析によって、視覚能力の個人差や多様性の実態が明らかになれば、背後にある視覚プロセスの相互作用を理解する手掛かりとなるからです。

私たちの取り組みの一つとして、簡易的な視覚能力のセルフチェックの実現があります。緑内障、白内障といった目の病気を軽度のうちに発見し、対処することは重要です。例えばこれらの病気では、視野の一部での感度低下などの異常が起こります。そのような異常を自覚症状が現れる前に知ることができ

れば、未病と呼ばれる状態のうちに対処が行えるかもしれません。自分で視覚能力を短時間に測定できるツールと、それを日常的に使ってもらえるような仕掛けがあれば、この問題の解決に向けた大きなステップになります。

しかし、「広いデータ」の検討には、これまでの実験では顕在化しなかった様々な問題を解決する必要があります。例えば、上で述べたセルフチェックの実現には、以下に挙げるような課題があります。まず①測定時間の大幅な短縮(数分以内の測定時間が望ましい)が必要です。次に、一般の電気店で市販されている②汎用的な機器で安価にデータ収集する方法が必要です。さらに、インターネットを通じたデータ収集の仕組みも備えていることが望ましいでしょう。また、日常的な反復利用が望ましく、③自発的に取り組みたいくなるような楽しく、課題が理解しやすい測定法である必要があります。一方で、④これまでの測定法とデータが比較可能であることは、検査の信頼性やデータに異常が見られたときに病院での本格的な検査へとつなげるために重要です。

私たちは、神戸市立神戸アイセンター病院と共同で、タブ

レット、あるいはタッチパネルを備えたPCで動作する2種類の視覚テストを作成して、これらの課題解決に取り組みました(図2)。一つは、短い時間でプレイできるゲームの中に視覚テストを埋め込んだ視覚測定ミニゲームです。これは、普段の生活の中での反復利用を通して、日常生活の中で視覚能力の異常をセルフチェックするためのツールです。従来の手続きを一部変更・簡略化したものをベースに、デザインやカバーストーリーを付加して作成された複数のミニゲームを通して、視覚能力を楽しくチェックできます。もう一つは従来の視覚テストを踏襲しつつ簡略化した簡易テストツールです。従来のテスト方法を踏襲しているため、データの信頼度についてはゲーム形式のものより高く、さらに、ゲーム形式からいくつかのテスト項目が追加されているので、より自分の状態を細かく把握することができます。

これらのツールでは、共通した複数の視覚能力の測定が可能であり、それらのパフォーマンスのデータを比較できます。また、双方とも汎用的なウェブブラウザで動作し、インターネット上での測定にも利用できるように設計をしています[1]。

従来方法の簡略化・一部改変により、従来の実験・検査データより精度が低下する種目もあります。しかし、そのような場合でも、スクリーニングや簡易チェックに必要な測定精度は達成しています。例えば、ソフトウェア的手法による輝度の解像度の拡張やWebカメラによる視距離の簡易測定などで、実験室環境のエミュレーションを行い、測定精度の低下を小さくしています[2], [3]。

これらのツールから得られるデータは、日常的なセルフチェックに役立つだけでなく、医療現場での診断精度の向上や、病院外でのリハビリテーション支援への貢献も期待できます。このことは、科学研究の立場からみると、健康者、未病者、患者を含めた多様な群での個人差を示す広いデータと、医療のプロフェッショナルが持つ患者を中心とした群の精密なデータの関係を明らかにすることにつながります。この取り組みによって、視覚能力の多様性、個人差の実態解明と、その背後にある人間の複雑な視覚プロセスの相互作用の理解を加速したいと考えています。

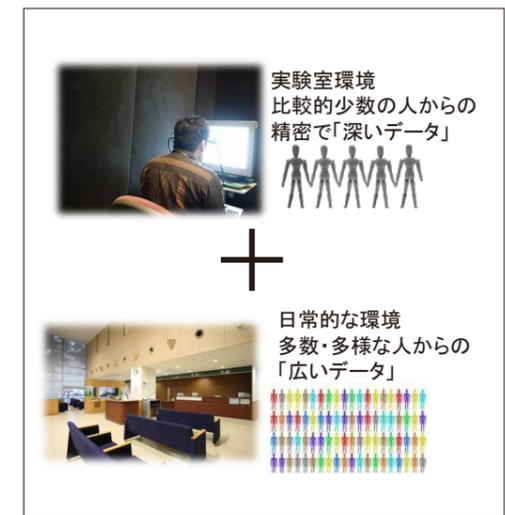


図1 人間科学で用いられる2種のデータ



図2 視覚能力のセルフチェックシステム概要

### 関連文献

[1] K. Hosokawa, K. Maruya, S. Nishida, "Testing a novel tool for vision experiments over the Internet," *Journal of Vision*, Vol. 16, 2016.  
[2] R. Allard, J. Faubert, "The noisy-bit method for digital displays: Converting a 256 luminance resolution into a continuous resolution," *Behavior Research Methods*, Vol. 40, No. 3, pp. 735-743, 2008.  
[3] X. P. Burgos-Artizzu, M. R. Ronchi, P. Perona, "Distance estimation of an unknown person from a portrait," in *Proc. European Conference on Computer Vision*, pp. 313-327, 2014.

# いろいろな観点で似たものを探す、似たものつながり検索法

## ～絵本検索システム「ぴたりえ」とそれを支えるグラフ索引型探索技術～

Like likes like strategy: search suitable for various viewpoints

Picture book search system "Pitarie" with graph index based search



協創情報研究部  
**服部 正嗣**  
Takashi Hattori

### (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。2004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。複合的メディアを対象とした類似探索の研究に従事。情報処理学会、人工知能学会各会員。最も思い出深い絵本は新美南吉著「ごんぎつね」。

今や多くの方が一日一回は何らかの検索サービスを利用していることでしょう。キーワード検索で様々な目的に適した情報が載ったウェブページを簡単に見つけることができます。特定の目的に特化したウェブ検索サービスも年々充実してきており、指定地点近辺のレストランなどを探せる地図検索や、電車やバスの乗換案内検索など、枚挙に暇がありません。

このような目的特化型の検索サービスが多数存在する中で、絵本に特化した初めての検索システムとして、コミュニケーション科学基礎研究所 (CS研) で研究開発した「ぴたりえ」を紹介します。絵本の検索なんて簡単にできるのでは?と思われる方もいらっしゃるかもしれませんが、実はその実現には解決すべき絵本検索固有の課題があり、様々な工夫が必要でした。それらの工夫と、「ぴたりえ」を支える検索の仕組みが絵本に限らず様々な検索対象に応用できる汎用性に優れたものであることをご説明します。

### 絵本検索システム「ぴたりえ」

「ぴたりえ」は一人一人の子どもに「**ぴったりのえほん**」を探すための検索システムを目指して研究開発しました。従来は購入者・専門家による情報発信や大量の購買履歴などが絵本選びにおける重要な情報として活用されていましたが、子どもの成長が早く個人差も大きいことを考えると、それらの情報は個々の子どもの特定の時点の成長度合いによっては必ずしも

有用でない可能性があります。そこで我々は、以下に述べる3つの基礎研究分野の技術・知見を応用して直接絵本そのものが持つ情報を深く分析し、従来できなかった検索を実現しました。

1つ目の研究分野は自然言語処理で、文章をコンピュータに上手く処理させるための工学的な分野です。文章をコンピュータが扱う際には、扱いやすい単語単位に分解する「形態素解析」という処理を行います。絵本の文章はひらがなばかりで言葉の切れ目が分かりにくいので、従来の技術では形態素解析精度が低いという問題がありました。この問題を解決するため、我々はまず正しい解析結果が分かっている大量の大人向け文章を絵本同様ひらがなばかりに自動的に書き換えて、大量の正解データ付きのひらがな文章を準備しました。このひらがな文章を学習データとして機械学習を行うことで、絵本の文章に対しても高精度な形態素解析を低コストで実現しました。

2つ目の研究分野は、子どもの言語発達メカニズムを明らかにする言語習得科学です。研究過程で得られた、1500名以上もの子どもがいつごろどんな言葉を理解・発話できるようになったかを集約した幼児語彙獲得データベースを参考に、絵本に出現する単語の「子どもにとっての難易度」を推定できます。単語の難易度情報と前述の自然言語処理によって得られた絵本の文章の特徴とを基に機械学習を行うことで、対象の

絵本が何歳向けとして適しているかを高精度で自動推定できるようになりました。さらに、同じ対象年齢に分類されている絵本同士のどちらがより簡単/難しいか、などの詳細な比較もできるようになったため、平均的な発達の仕方と比較して子どもの発達が早くても遅くても、その子どもの絵本選びに役立つと期待しています。

3つ目の研究分野は類似探索で、検索の中でも特に入力と似たものを探すことを目的にした領域です。「ぴたりえ」は検索の基盤としてグラフ索引型類似探索法というCS研独自の方法を用いています。グラフ索引型類似探索法は、高速性と様々なメディアに対応できる汎用性が特長です。絵本に適用することで、文章の内容がお気に入りの絵本と類似の本を探したり、お気に入りの絵本と画風が似た絵本作家を探したり(図1)、曖昧な記憶に基づいた、間違いを含むあらすじで絵本を探したりなど、様々な観点に基づく絵本検索を実現できます。

### グラフ索引型類似探索法

グラフ索引型類似探索法は、メディアの特性によらず、検索対象の一つ一つを頂点とするグラフ(ネットワーク)を索引とすることで高速な探索を実現しています。似たもの同士が繋がるようにグラフを構成すると、そのグラフはどんなに無関係そうなもの同士も少数の「知り合い」を経由することでつながるスモールワールド性という性質を示します。たとえば「人間関係」はスモールワールド性を持つことが知られており、世界中の人々の中から無作為に抽出した2人は平均して6人の知り

合いを介することで繋がっていることが「6次の隔たり」として知られています(図2)。数十億人の中で6人だけ経由すれば誰にでも繋がると考えると、とても検索向きの性質です。グラフ索引型類似探索法では、検索対象の似たもの繋がりやグラフ構造を構築し、そのグラフ上をより似ている知り合いに移動する手順で、高速かつメディアによらない検索を実現しています。もちろん、すべての頂点が互いに接続されたグラフを構築すればどの頂点からでも他の頂点に直接到達できますが、それでは結局すべての頂点を調べるのと同じで高速な検索は実現できません。一つ一つの頂点と隣接している頂点の数が人間関係の知り合いの数くらいに限られているにも関わらず、「6次の隔たり」でどの頂点にも到達できるということがグラフ索引のポイントとなっています。

### 今後の展開

絵本検索システム「ぴたりえ」は自治体図書館や大学の教育関連学部などでの実証実験やワークショップを通じて有効性を検証し改善を重ねた結果、NTTデータ九州社により図書館・幼児関連施設などを対象に2019年4月から商用化されています。現在、絵本を活用した病院での発音訓練支援や英語絵本を活用した小学校での英語教育支援などを目指し、各所との共同研究を進めています。グラフ索引型類似探索単独でも一般書籍、技術文書、音声などへの活用範囲拡大を目指して研究に取り組んでいきます。



図1 「ぴたりえ」で、一冊の絵本(左上「しろくまちゃんとほっとけーき」)を入力として類似の表紙画像を検索した例。上位に複数の絵本がランクインする絵本作家を表示することで、お気に入りの絵本作家と画風の似た別の絵本作家を見つけることができる。

引用させていただいた絵本:  
しろくまちゃんのほっとけーき、わかやまけん、こくま社、1972年  
うさこちゃんと うみ、ディック・ブルーナ、福音館書店、1964年  
うさこちゃんと あかちゃん、ディック・ブルーナ、福音館書店、2005年  
うさこちゃんの だんす、ディック・ブルーナ、福音館書店、2009年  
うさこちゃんと どうぶつえん、ディック・ブルーナ、福音館書店、1964年  
うさこちゃんのはたけ、ディック・ブルーナ、福音館書店、2005年  
わたし ほんが よめるの、ディック・ブルーナ、福音館書店、1972年

### 関連文献

[1] 服部正嗣, 小林哲生, 藤田早苗, 奥村優子, 青山一生, "ピタリエ: 興味・発達段階にピッタリの絵本を見つけます," *NTT 技術ジャーナル*, Vol. 28, No. 6, pp. 54-59, 2016.  
[2] K. Aoyama, K. Saito, H. Sawada, N. Ueda, "Fast approximate similarity search based on degree-reduced neighborhood graphs," in *Proc. The 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1055-1063, 2011.

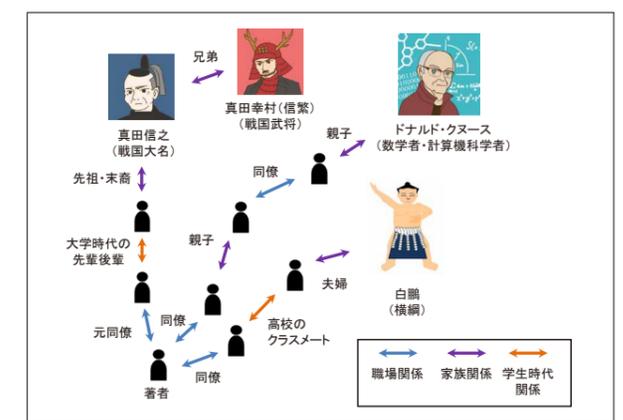


図2 人間関係の一例。職場関係、家族関係、学生時代の関係など、複数の異なる関係を通ることで、数人経由しただけでバラエティ豊かな著名人に行き当たる。世界中からこれら著名人を探す際、一見関係がなさそうな著者からスタートしたとしても、「より近そうなり合いを辿る」という方法ですぐに検索が終了する。

# 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目指して、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2019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30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の研究を基礎から活用事例まで包括的に紹介します。

- 膨大な数の経路や、複雑な制御の候補から最適な解を見つけたい
- 人間行動の特性や時空間の性質、因果関係の発見により、未来を予測したい
- 深層学習モデルの表現能力を高めたい

## 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・計算の限界に挑戦する研究を紹介します。

- 人と人の共感を深め、AIと人の対話を豊かにしたい
- 文書の修辞構造や幼児の言語獲得の仕組みを明らかにしたい
- 限界まで効率的かつ安全に情報を伝達したい

## メディアの科学

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

- うるさい環境でも会話を楽しみたい
- 多人数会話を聞き分け、高精度に音声を認識したい
- 各種メディアの本質を捉え、自在に変換したい

## 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
- 多様性を許容する社会をサポートしたい
- トップアスリートの秘密を解明したい



# 03 都市の人口分布変動から人の移動を再現

## ～時間別エリア人口データからの移動傾向推定技術～

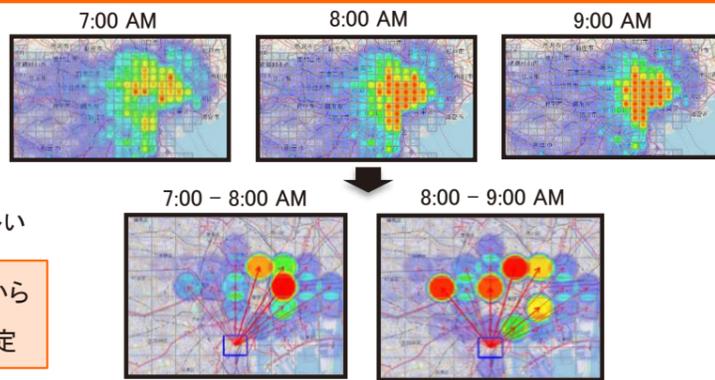
**どんな研究** 群衆の位置情報を解析する上で、時間別エリア人口データ(いつ・どこに人が何人いたか)が注目を集めています。本技術はそのような時間別エリア人口データから、人の移動の主な傾向(いつ・どこから・どこへ何人移動したか)を推定します。

**どこが凄い** 移動傾向を推定するためには、たくさんの「観測された人口データの変化をうまく説明する群衆の移動パターン」の中から正しい解を絞り込む必要があります。本技術では、「都市における人間行動らしさ」を判断基準に組み込むことで、高精度な推定を実現しました。

**めざす未来** 本技術を用いることで、扱いの難しい個人の移動情報を利用せずに群衆の移動傾向を知ることが可能です。タクシー配車やバスの運行計画などの交通系サービスにおいて、「エリア毎の人数」に加えて「エリア間の移動人数」も考慮した、高度な運用計画の立案が可能になります。

### 何に取り組んでいるか？

- ✓ 群衆の移動を記録したデータを企業間を跨いで様々なサービスで利用することはプライバシーの問題で難しい
- ✓ 実用上は、時間別のエリア人口の形などに集計したものしか利用可能でない場合が多い



**本技術** 時間別エリア人口データから  
エリア間の移動人数を推定

### 技術ポイント

「都市における人間行動らしさ」を考慮した判断基準により高精度な推定を実現

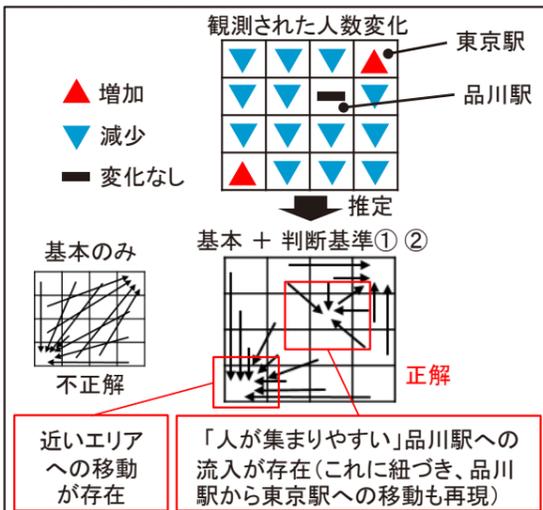
- 基本 : 減少エリアから増加エリアへ移動を割り当て
- 判断基準①: エリアごとの特徴 (人が集まりやすい・出ていきやすい)を考慮
- 判断基準②: エリア間の距離が移動しやすさに影響

エリア  $i$  から  $j$  への移動確率  $\theta_{ij}$

$$\theta_{ij} \propto \pi_i \times s_j \times \exp(-\beta \cdot \text{dist}(i, j))$$

判断基準①      判断基準②

- エリア  $i$  にいる人は 確率  $\pi_i$  で他のエリアへ移動
  - 次の行き先  $j$  は  $(s_j: \text{エリア} j \text{の人の集まりやすさ}) \times (i \text{と} j \text{の距離})$  に比例した確率で決定
- ※  $\pi_i, s_j, \beta$  はデータから教師なし推定



### 関連文献

[1] Y. Akagi, T. Nishimura, T. Kurashima, H. Toda, "A fast and accurate method for estimating people flow from spatiotemporal population data," in Proc. the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence (IJCAI-ECAI-18), 2018.

[2] 赤木康紀, 西村拓哉, 倉島健, 戸田浩之, "人の移動の特徴を考慮した人口統計情報からの移動人数推定," 日本データベース学会和文論文誌, Vol. 17-J, No. 9, 2019.

[3] 赤木康紀, 西村拓哉, 倉島健, 戸田浩之, "人口データからの移動傾向推定の高速化について," 第21回情報論的学習理論ワークショップ, 2018.

### 連絡先

赤木 康紀 (Yasumori Akagi) サービスエボリューション研究所  
Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 04 深層学習のボトルネック解消で精度を向上

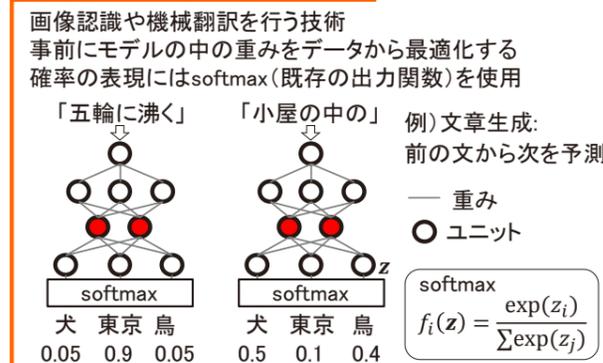
## ～深層学習における、より高い表現能力を持つ出力関数～

**どんな研究** 深層学習による画像認識や機械翻訳において予測精度を改良するための研究です。深層学習は大量のデータをモデルが学習することで、例えば画像が入力されたときにそれが動物、車、人であるという確率を出力します。モデルの最後に使用する出力関数を改良して高精度化します。

**どこが凄い** 高精度なモデルを構築するためには、様々な確率のパターンを出力できる表現力が必要となります。既存の出力関数は表現力のボトルネックとなり、精度向上に大量のパラメータを必要とするため、計算コストが高くなります。そこで私達はパラメータを増やすことなく表現力を改善する技術を作りました。

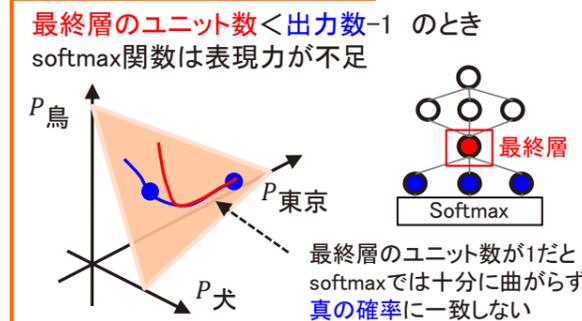
**めざす未来** 本技術を用いることで、高精度なモデルを既存のモデルの一箇所を変更するだけで構築できるようになります。本研究のように表現力のボトルネックを解消していくことで、より少ないパラメータ数、計算コストで高精度な予測が可能な深層学習モデルの実現に寄与します。

### 深層学習とは



高精度な予測のためには様々な確率のパターンを出力できる能力(表現力)が必要

### 表現力のボトルネック



言語処理の場合、出力数 = 語彙数 = 数万でありユニット数を出力数と合わせると膨大な重み(パラメータ)が必要となり計算コストが高い

### 関連文献

[1] S. Kanai, Y. Fujiwara, Y. Yamanaka, S. Adachi, "Sigsoftmax: Reanalysis of the softmax bottleneck," in Proc. 32nd Conference on Neural Information Processing Systems (NeurIPS), 2018.

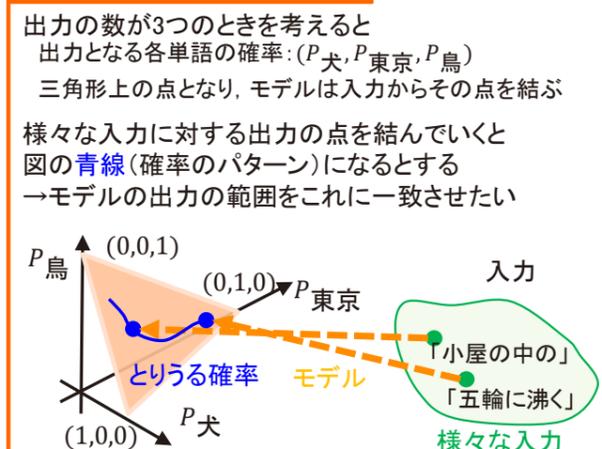
### 連絡先

金井 関利 (Sekitoshi Kanai) ソフトウェアイノベーションセンタ  
Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

### 問題設定



### 新しい出力関数 sigsoftmax

sigmoidをかけてより曲がる(表現力が向上する)ような関数を提案

sigsoftmax

$$f_i(z) = \frac{\sigma(z_i) \exp(z_i)}{\sum \sigma(z_j) \exp(z_j)}$$

$$\sigma(z) = \frac{1}{1 + \exp(-z)} : \text{sigmoid}$$

0から1に近づく関数

softmaxより高い表現力[1]  
追加のパラメータなく計算コストが低い

# 05 データから学ぶ：どれが原因？どれが結果？

～教師あり学習に基づく時系列の因果推論～

## どんな研究

時々刻々と変化する時系列データを対象に、その間の原因と結果の関係(因果関係)を自動的に発見する研究です。例えば「研究費は売上に影響を与える」という因果関係は、企業が意思決定をする際に有用な知見です。このほか、遺伝子制御ネットワークの推定など、多種多様な応用が期待できます。

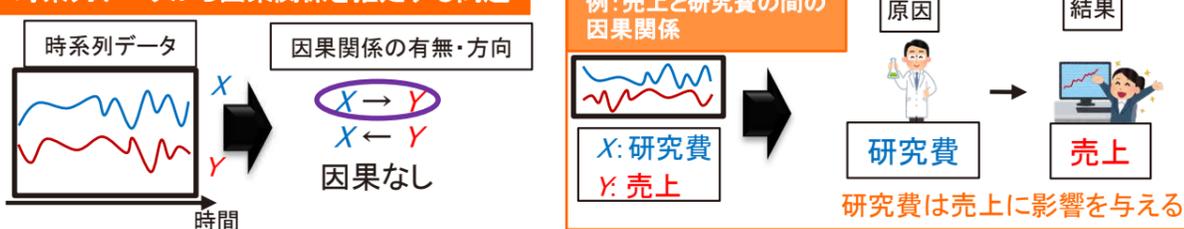
## どこが凄い

因果関係を正しく推定するには、各データの性質に即して数式表現(自己回帰モデル)を選ぶ必要がありますが、データ分析の専門家でなければ困難でした。そこで多様な性質を持つデータを用いて機械学習モデルを学習することで、データ分析の専門知識なしに因果関係を推定できる技術を構築しました。

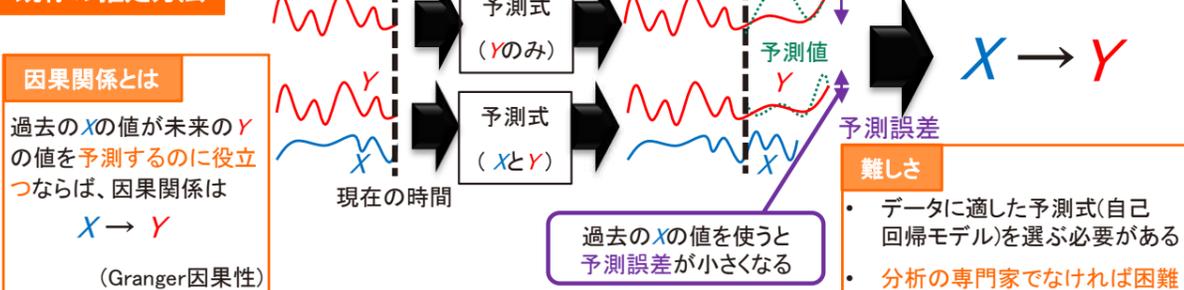
## めざす未来

因果関係に関する知見によって、例えば売上を上げるために研究費を上げるといった重要な意思決定を効果的に行うことができる未来をめざしています。そのために、データの間の相関関係がどのように生じたものかを、因果関係の言葉でより正確に説明できる技術の実現をめざしています。

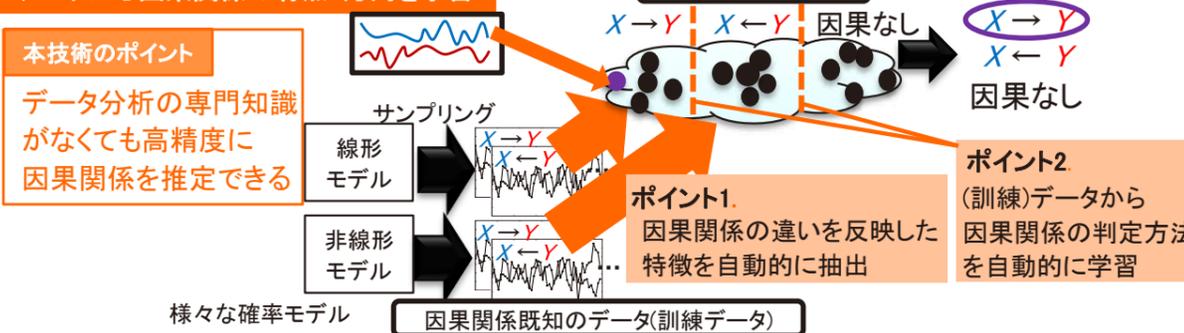
### 時系列データから因果関係を推定する問題



### 既存の推定方法



### データから因果関係の有無・方向を学習



### 関連文献

- [1] Y. Chikahara, A. Fujino, "Causal Inference in Time Series via Supervised Learning," in *Proc. 27th International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.
- [2] 近原鷹一, 藤野昭典, "教師あり学習に基づくGranger causalityの推定," *情報処理学会論文誌: 数理モデル化と応用(TOM)*, Vol. 11, No. 3, pp. 58-73, 2018.

### 連絡先

近原 鷹一 (Yoichi Chikahara) 協創情報研究部 知能創発環境研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 06 過去データの無い地点でも未来を予測

～時空間データ解析のための時空間回帰テンソル分解法～

## どんな研究

時空間データ解析において、過去のデータが未観測な地点での未来予測技術の研究をしています。過去に得られた膨大な時空間データから、データに隠された少数かつ重要な時空間パターンおよびパターン毎に現れる時間・空間相関を同時に学習することで過去データの無い地点の未来を予測します。

## どこが凄い

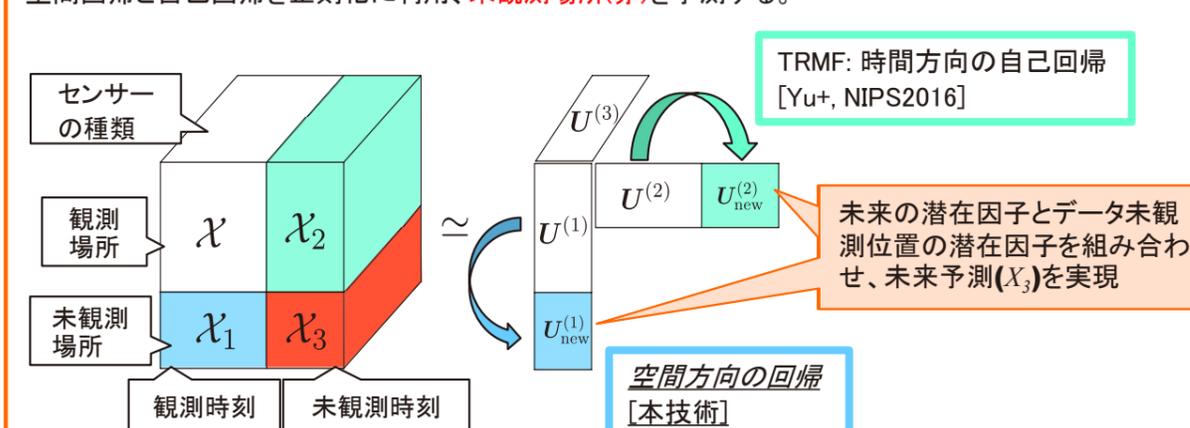
時空間データには、データの観測位置が規則正しく格子状に置かれた場合と、そうでない場合が存在します。どのような観測位置の形状にも対応可能な空間回帰正則化を新たに提案し、自己回帰テンソル分解に適用することで、潜在空間における空間・時間相関を同時モデル化を実現しました。

## めざす未来

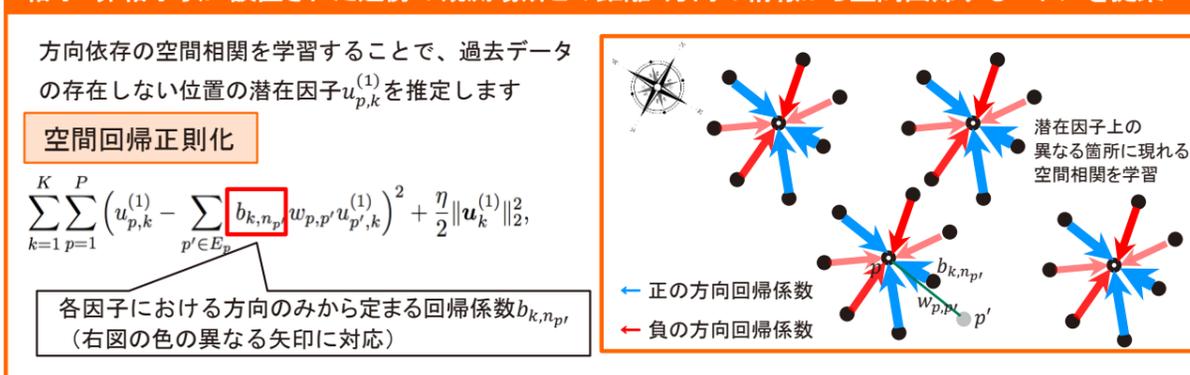
本技術を用いることで、既存の方法よりも高精度な時空間データの予測ができるようになります。今後は企業、行政機関、人工衛星、センサーネットワークなどから得られる異なる種類の時空間データに対応した、より柔軟かつ複雑な相関関係を捉える時空間分析技術の実現を目指します。

### 時空間回帰問題

未観測場所の因子(青)を空間回帰により推定。自己回帰テンソル分解と組み合わせ、空間回帰と自己回帰を正則化に利用、未観測場所(赤)を予測する。



### 格子・非格子状に設置された近傍の観測場所との距離・方向の情報から空間回帰するモデルを提案



### 関連文献

- [1] K. Takeuchi, H. Kashima and N. Ueda, "Autoregressive Tensor Factorization for Spatio-Temporal Predictions," in *Proc. of 2017 IEEE International Conference on Data Mining (ICDM)*, 2017.
- [2] 竹内孝, 鹿島久嗣, 上田修功, "自己回帰テンソル分解による時空間データ予測," *2018年度人工知能学会全国大会(第32回)*, 2018.

### 連絡先

竹内 孝 (Koh Takeuchi) 上田特別研究室  
 Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 07

## いろいろな観点で欲しいものを見つけます

～グラフ索引型探索技術による絵本検索システム「ぴたりえ」～

### どんな研究

「間違い混じりのあらすじを入力して絵本を探す」「好みの絵本と似た画風の作家による絵本を探す」など、**いろいろな観点で絵本を検索できるシステム**について紹介します。独自の検索技術に加え、ひらがな文の自然言語処理技術、子どもの語彙獲得についての知見を取り入れた学際的研究です。

### どこが凄い

提案手法は、検索対象となる大量のデータを事前に整理して、似たもの同士が接続されたグラフ(ネットワーク)を構成しておくことで高速な検索を実現します。従来手法と異なり、メディア(文章、画像、音声など)の特性に依存しないため、**幅広いメディア**から抽出する特徴に対応して**高速化**を実現します。

### めざす未来

病院での発音訓練に適した絵本探し、小学校での英語教育に適した英語絵本探しなど、**絵本を活用した様々なサービス実現**をめざすとともに、一般書籍や技術文書など**絵本以外のメディアについても多様な観点からの検索を実現**するための基盤となるような研究を行います。

#### 絵本検索システム「ぴたりえ」

いろいろな観点で絵本を探せます。

観点1: あらすじ  
観点2: お気に入りの絵本と内容が似ている  
観点3: 画風が似ている (下記例)

#### 検索に先立ちグラフ索引を用意(オフライン処理)

##### グラフ索引構築アルゴリズム

似たもの同士が接続されたグラフを構成し、索引として活用。

- 「スモールワールド性」: 任意の2頂点が少数ホップ以内に存在し、検索に適している。
- メディアの特性に依存しない工夫なので、文章画像音声など多様な対象に適用可能。

#### グラフ索引を用いて高速な検索を実現

##### 検索例

入力した絵本と同作者の絵本はもちろん、似た画風の作家の絵本も見つけられます。

クエリ → グラフ探索アルゴリズム → クエリ絵本と異なる作者の絵本

索引上を移動して類似物を探索

引用させていただいた絵本:  
しろくまちゃんのほっとけき、わかやまけん作、こぐま社、1972年  
うさこちゃんとうみ、ディックブルーナ作、福音館書店、1964年  
('雪の女王'は本スライドのため内製したものです)

#### 様々な応用

##### 応用例

構音(発音)障害の訓練に適した絵本の検索システム「ぴたりえメディカル」

わたち(私)は、おひめさまが好き!

サ行の発音をタ行と混同しがちな女の子

サ行タ行の発音を含む3音以上の言葉が多くて、30ページくらいで…

言語聴覚士(構音訓練を行う専門家)

両者からの複合的な条件とよく一致する絵本を出力

竹田総合病院と共同研究中

67種類

#### 関連文献

[1] 服部正嗣, 小林哲生, 藤田早苗, 奥村優子, 青山一生, “ピタリエ: 興味・発達段階にピッタリの絵本を見つけます,” *NTT 技術ジャーナル*, Vol. 28, No. 6, pp. 54-59, 2016.

[2] K. Aoyama, K. Saito, H. Sawada, N. Ueda, “Fast approximate similarity search based on degree-reduced neighborhood graphs,” in *Proc. The 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1055-1063, 2011.

**連絡先**  
**服部 正嗣 (Takashi Hattori)** 協創情報研究部 知能創発環境研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp

Innovative R&D by NTT  
 オープンハウス 2019  
 Copyright © 2019 NTT. All Rights Reserved.

# 08

## 限界まで効率よくメッセージを送れます

～シャノン限界を達成する誤り訂正符号～

### どんな研究

高速な通信環境を実現するためには、**雑音のある環境下で効率よく正しい情報を送受信**する必要があります。これを実現する技術は「誤り訂正符号」と呼ばれています。本技術は、計算科学者シャノンによって求められた、**符号化効率の限界(シャノン限界)を達成する誤り訂正符号を実現**します。

### どこが凄い

第5世代移动通信システム(5G)に実装されている誤り訂正符号(LDPC符号・ポーラ符号)は、特殊な通信路ではシャノン限界を達成できますが、一般の通信路では達成できません。本展示の技術を用いることにより、**一般の通信路でもシャノン限界を達成**することができます。

### めざす未来

本技術(CoCoNuTS\*)を用いることで、既存の方法よりも**効率のよい通信を実現**できます。これは光通信や無線通信の帯域が貴重となる環境で効果を発揮します。今後は実環境に近い通信路に適用するための周辺技術を確立して、**将来の通信技術の実現**をめざします。



#### 実現の難しさと提案技術のポイント

##### 幾何学的なイメージ

通信路出力(雑音を含む)

#### 符号化・復号化の難しさ

- 限界を達成するためには、通信路入力  $\{x_i\}$  を効率よく(左図ではできる限り離して)配置する必要があります。
- LDPC符号やポーラ符号は、最適な通信路入力  $X$  が一様分布になる通信路では限界を達成しますが、これを満たさない通信路では限界を達成する配置になっていません。
- 素朴な復号法では、雑音を含む通信路出力から総当たりで通信路入力を推定する(左図では  $y$  の最近点  $x_4$  を求める) 必要があり、実時間で計算できませんでした。

#### 提案技術のポイント

- 符号器に拘束条件を満たす乱数生成器を用いることにより、理想的な通信路入力の配置を実現できます。
- 復号器に拘束条件を満たす乱数生成器を用いて総当たりを回避することにより、実時間の計算が可能になりました。

\* CoCoNuTS = Code based on Constrained Numbers Theoretically-achieving the Shannon limit

#### 関連文献

[1] J. Muramatsu, “Channel coding and lossy source coding using a generator of constrained random numbers,” *IEEE Transactions on Information Theory*, Vol. 60, No. 5, pp. 2667-2686, 2014.

[2] J. Muramatsu, S. Miyake, “Construction of a channel code from an arbitrary source code with decoder side information,” *IEEE Transactions on Information Theory*, Vol. 65, No. 1, pp. 500-508, 2019.

**連絡先**  
**村松 純 (Jun Muramatsu)** 協創情報研究部 知能創発環境研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp

Innovative R&D by NTT  
 オープンハウス 2019  
 Copyright © 2019 NTT. All Rights Reserved.

# 09 新たな秘密がこれまでの秘密を脅かす

～「量子情報を用いた秘密分散」の脆弱性の検証～

**どんな研究** 秘密情報を安全に保管する方法として「量子情報を用いた秘密分散」が提案されていますが、その安全性を確認するには、量子状態の推定可能性を解明することが不可欠です。この展示では、**未知の量子状態が増えれば増えるほど、全状態を推定できるようになるという現象**の解析結果を紹介します。

**どこが凄い** 上記の現象が起こる必要十分条件を世界で初めて示し、その現象を引き起こす**実装可能な推定方法**を発見しました。この結果は、量子状態の推定可能性の解明に大きく貢献し、同時に、「量子情報を用いた秘密分散」の決定的な**脆弱性**を顕在化させました。

**めざす未来** 量子状態の推定可能性の解明は、**量子情報を用いた未来の様々な情報処理技術の実現**に貢献します。これらが実現することで、飛躍的に安全な情報処理が可能となります。また、今の技術では原理的に実現できない機能を持った情報処理が可能となることも期待されています。

### 量子状態の推定タスク

- ランダムに選ばれた量子状態を分割してユーザに配る
- ユーザは通常の通信路を用いて、配られた量子状態を推定する

### ベル状態の一斉推定タスクの解析

- ランダムに選ばれた多数のベル状態をユーザに配る
- ユーザは通常の通信路を用いて、配られたベル状態の種類  $x_1, x_2, \dots, x_N$  を**全て推定**する

**解析結果**

ランダムさのエントロピー > 1 の時  
全推定に成功する確率の上限

ランダムさのエントロピー < 1 の時  
全推定に成功する確率の下限

量子状態が4種類のベル状態(量子情報処理における特有用な量子状態)からランダムに選ばれる時

ベル状態1, 2, 3, 4 からベル状態x

推定成功確率 < 1

### 「量子情報を用いた秘密分散」の脆弱性の検証

- 秘密情報  $x$  を量子状態に符号化し、通常の通信路で結ばれたユーザに配る
- 攻撃ユーザは、配られた量子状態を推定して秘密情報  $x_1, x_2, \dots, x_N$  を読み出そうとする

本解析結果 → 秘密情報の数  $N$  が増えれば増えるほど、全ての秘密情報を読み出されてしまう危険性がある

**関連文献**

[1] S. Akiue, G. Kato, "Bipartite discrimination of independently prepared quantum states as a counterexample to a parallel repetition conjecture," *Physical Review A*, Vol. 97, No. 10, 042309, 2018.

**連絡先**

秋笛 清石 (Seiseki Akiue) メディア情報研究部 情報基礎理論研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp

Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

# 10 文や段落の修飾関係を賢く判断

～ニューラルネットによる階層的トップダウン修飾構造解析～

**どんな研究** 文書の背後にある談話構造を木として表現する研究を紹介します。文書を段落系列と文系列で構成された入れ子の階層構造としてとらえ、それぞれの階層での修飾構造をニューラルネットを用いて推定します。そして、それらをつなぎ合わせることで文書全体をあらわす修飾構造木を構築します。

**どこが凄い** 段落、文系列を独立かつ再帰的に二分割することで木を構築するため、並列化が容易であり長い文書であっても高速に動作することが期待できます。ニューラルネットを用いてテキストデータを自動的に特徴ベクトルに変換するため、人手で複雑な特徴を考える必要がありません。

**めざす未来** 自然言語処理技術は「文」だけではなく「文章」も処理対象として扱う必要があります。例えば、文書を翻訳する際には文脈に応じて適切な訳語を選ぶ、要約する場合には原文書の話題の一貫性を損なわぬようにすることが重要です。本研究はこうした自然言語処理の基盤として必要不可欠です。

### テキスト系列の分割によるトップダウンパーシング

Point

- 入れ子の階層構造
- 文書は段落の系列 → 段落を葉とする木を構築
- 段落は文の系列 → 文を葉とする木を構築

系列を二分割することで上から木を構築

赤字: 関係ラベル

### 修飾構造木の構築

### 修飾構造解析のベンチマークセットによる評価結果

	木の形	+主/従ラベル	+関係ラベル
提案法	72.0	58.6	46.7
提案法 階層化なし	66.5	53.4	43.3
従来法	68.6	55.9	45.8
Human	78.3	66.8	57.1

Bracket F値: 正解の修飾構造木との一致の割合

本展示は東京工業大学 奥村研究室との共同研究の成果です

**関連文献**

[1] 小林尚輝, 平尾努, 奥村学, 永田昌明, "階層構造を考慮したトップダウン談話構造解析," *言語処理学会第25回年次大会*, pp. 1002-1005, 2019.

**連絡先**

平尾 努 (Tutomu Hirao) 協創情報研究部 言語知能研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp

Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

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

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

## 11

## ひらがなはいつからわかる？

～幼児の文字習得メカニズムを探る～

## どんな研究

小学生以降の文字の読み書きはこれまで多くの研究が行われてきましたが、**就学前の幼児が文字をいつ、どのように習得していくか**についてはよくわかっていませんでした。本研究では、ひらがなの習得に注目して、幼児がひらがなを理解し始める時期や習得しやすい文字の特徴等を調べています。

## どこが凄い

幼児がひらがなを見ている時の目の動きを測定することで、**3歳の誕生日前後から音と文字の対応を正しく理解し始めている**ことを発見しました。加えて、大規模絵本コーパスなどを用いた解析により、ひらがなの読みと書きの習得の間で異なる要因が関連していることを明らかにしました。

## めざす未来

幼児のひらがな習得プロセスとその仕組みを科学的に明らかにすることにより、ディスレクシアなどの**読み書きが苦手なお子さんをできるだけ早期に発見する手法**を確立するとともに、各個人の発達段階に応じたテーラメード型の文字学習法の提案に貢献することを目指しています。

## 1 目の動きを用いたひらがな理解の測定



- 対象児: 80名の2-3歳児 (24-48ヶ月児)
- 方法: 幼児がモニターを見ている時の目の動きをアイトラッカーで記録
- 分析: 標的文字の注視時間を計算

3歳の誕生日前後の幼児(32-39ヶ月)は、ひらがなをほぼ読めない(図1)にも関わらず、音と文字の対応を理解し始めている可能性(図2)を示唆

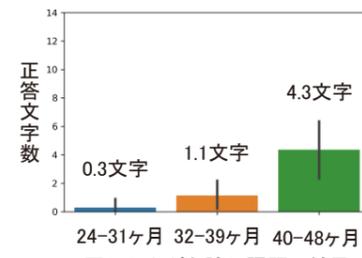


図1 ひらがな読み課題の結果

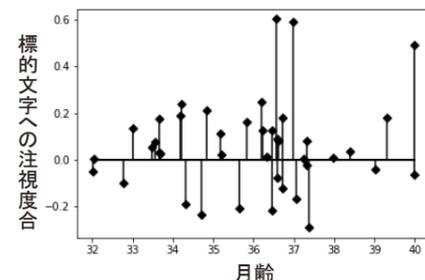


図2 音と文字の対応の理解 (32-39ヶ月)

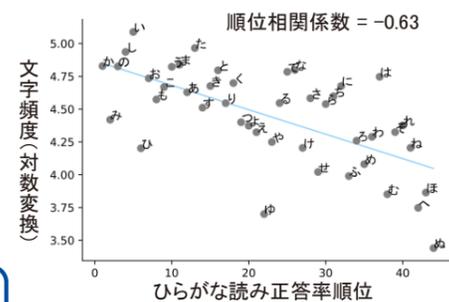
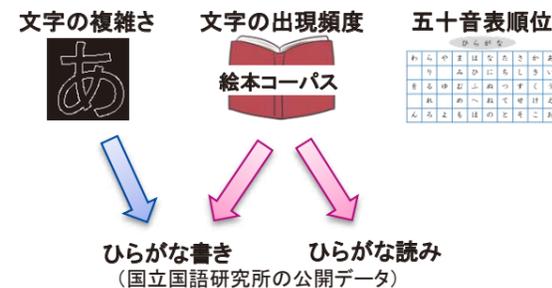


図3 絵本の文字出現頻度と読み正答率順位

## 2 ひらがな読みと書きの習得に与える影響の分析



ひらがなの読み習得は絵本に出現する文字頻度と関連(図3)書き習得はその要因に加えて文字の複雑さも影響

## 関連文献

- [1] H. Higuchi, Y. Okumura, T. Kobayashi, "Acquisition of letter-sound correspondence in Japanese-speaking 2-year-olds: An eye-tracking study," in *Proc. Biennial Meeting of Society for Research in Child Development (SRCD)*, 2019.
- [2] 樋口大樹, 奥村優子, 小林哲生, "幼児のひらがな読み・書き習得に及ぼす文字特性の影響," *音声言語医学*, (in press).
- [3] 樋口大樹, 奥村優子, 小林哲生, "幼児のカタカナ読み書き習得に関する文字特性の検討," *音声言語医学*, (in press).

## 連絡先

樋口 大樹 (Hiroki Higuchi) 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-liaison-ml at hco.ntt.co.jp



Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

## 12

## 「こころが動く、通う」をどう測る？

～主観・生理・行動からみた共感的コミュニケーションの分析～

## どんな研究

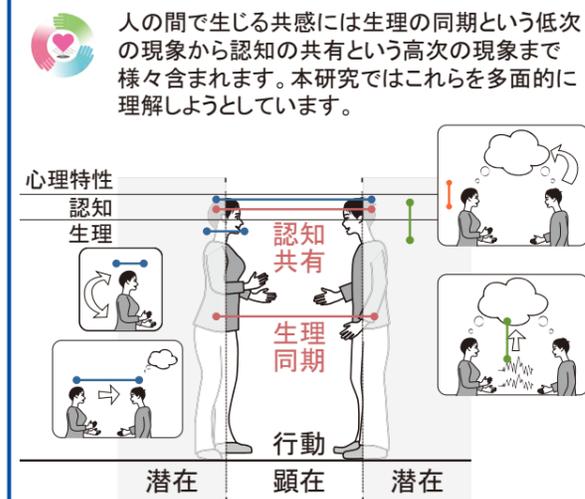
共感とは人が社会的な生活を営む基盤ですが、主観・生理・行動から構成される複雑な現象であるために、その仕組みはまだ十分に解明されていません。本研究では、**共感を多面的に理解するために定量化し**、そこに含まれる個人差を考慮した方法論を提案していきます。

## どこが凄い

個人の中で**生理と行動**がどのような関係にあるのか、複数の個人の間で**情動がどのように伝わるのか**を明らかにしました。さらに、個人差が大きい主観の分析において、複数人のデータを集約する集合知のアプローチと、心理特性にもとづく**個人差を含むモデル構築**の両方に取り組んでいます。

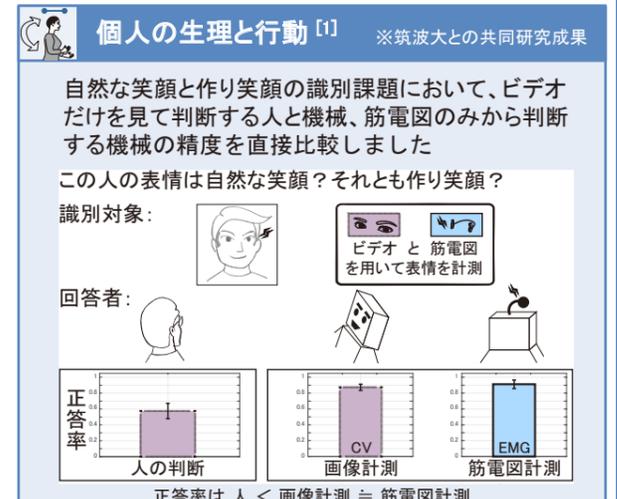
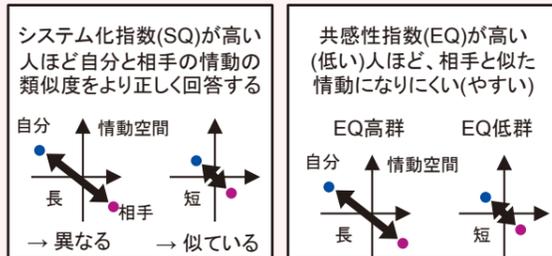
## めざす未来

共感を個人差を含めて定量計測する仕組みを整えば、一人一人にあわせた**共感を促す介入やその効果の評価・予測**が可能になります。他者との関わりがより豊かになれば、ウェルビーイング(心身の潜在能力を発揮し、いきいきと満足して生きること)の向上にもつながると期待されます。



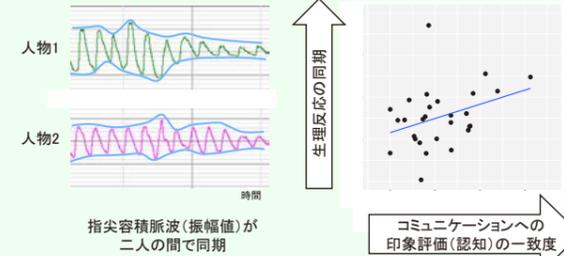
## 個人間の情動伝染の認知と心理特性 [2]

どのような人が他者と似た情動になりやすく、また、似た情動になっていることを正しく認知しやすいかを類似性判断理論にもとづき解明しました



## 個人間での生理の同期と認知の共有

協力課題を実行中に、生理反応が同期すること、生理的同期が高いほど「場」に対する印象(認知)も一致しやすいことを明らかにしました



## 関連文献

- [1] M. Perusquia-Hernandez, S. Ayabe-Kanamura, K. Suzuki, S. Kumano, "The invisible potential of facial electromyography: A comparison of EMG and computer vision when distinguishing posed from spontaneous smiles," in *Proc. CHI Conf. Human Factors in Computing Systems (CHI)*, 2019.
- [2] L. Antakiet, M. Matsuda, K. Otsuka, S. Kumano, "Analyzing generation and cognition of emotional congruence using empathizing-systemizing quotient," *International Journal of Affective Engineering*, Vol. 17, No. 3, pp. 183-192, 2018.

## 連絡先

熊野 史朗 (Shiro Kumano) 人間情報研究部 感覚共鳴研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp



Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

## 13

## 「触れて深まる体感と共感」を集団から測る

～共感的コミュニケーションの触覚的促進と多人数同時測定～

## どんな研究

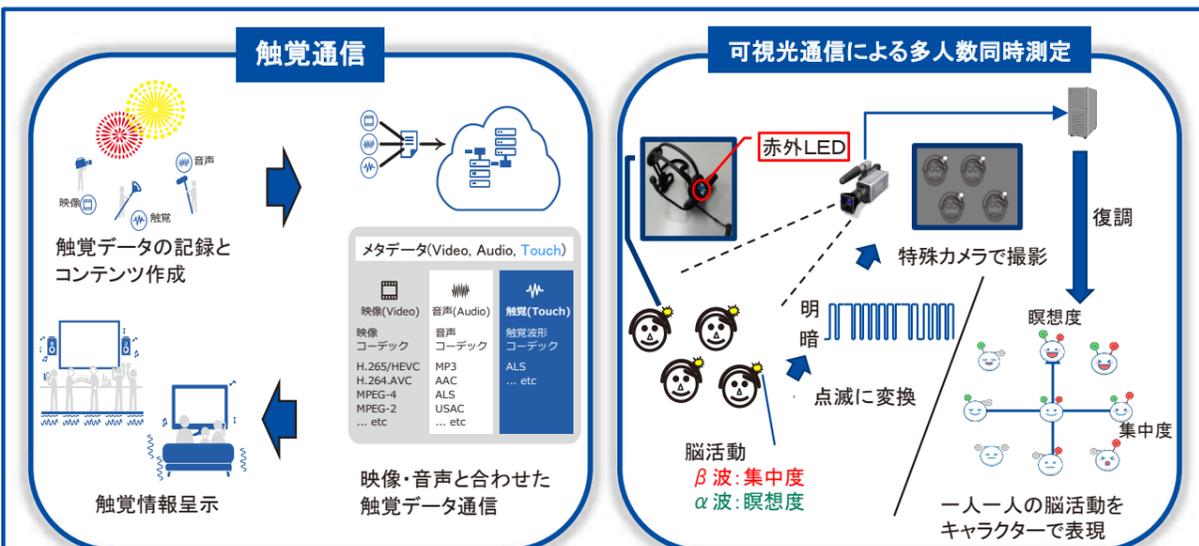
触覚体験によって配信コンテンツの視聴効果を向上させると共に、体験の場を共有することで生じる共感的コミュニケーション(一緒に見る効果)を、可視光通信を使った多人数同時測定技術によって客観的な生理指標として測定・評価していく研究です。

## どこが凄い

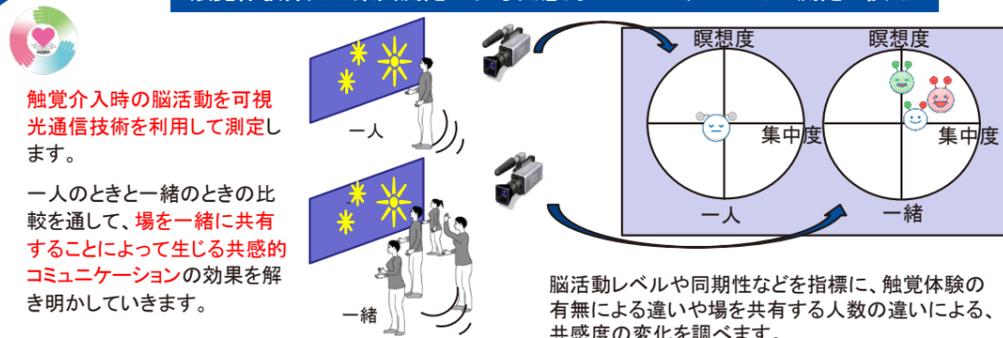
触覚情報を映像音と共に記録、配信、呈示する触覚通信の仕組みにより、体感を深めることにつながります。可視光通信により触覚体感を生理的な反応として多人数から同時に測定することが出来ます。異なる技術分野を組み合わせ共感的コミュニケーションに関する新たな研究を進めることが出来ます。

## めざす未来

触覚呈示技術と多人数測定技術を統合的に用いることによって、五感介入、測定、評価、要因特定といった循環的プロセスに基づく研究を行い、従来は直感的に語られていた共感的コミュニケーションを定量化するとともに、人々がいきいきと満足できる場を設計する指針や方法論を提案していきます。



## 触覚体験介入と集団測定による共感的コミュニケーション測定の試み



## 関連文献

- [1] 渡邊淳司, “触/身体感覚の記録・伝送・再生の原理とその社会での価値創生,” 電気通信, 2019年2月号.
- [2] Y. Shiraki, T. G. Sato, T. Moriya, “Flexible synchronization in optical camera communication with on-off keying,” in Proc. IEEE GLOBECOM Workshops, pp. 1-6, 2017.
- [3] 佐藤尚, 白木善史, 守谷健弘, “可視光通信による群集Well-Beingの測定の試み,” 人工知能学会全国大会(第32回)論文集, 2F2-OS-4a-03, 2018.

## 連絡先

佐藤 尚 (Takashi G Sato) 守谷特別研究室  
 Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

## 14

## ロボットと話そう「いつ、どこで、何をした」

～ユーザ発話中のイベント理解に基づく雑談対話システム～

## どんな研究

従来の対話システムに欠けていた、ユーザの発話内容をシステムが理解しているとユーザに感じさせる「理解してくれている感」のある対話を目的とした研究です。ユーザの話したイベントを構造化された情報(いつ、どこで、等)として理解することで、イベントの内容に即した共感や話題展開を実現します。

## どこが凄い

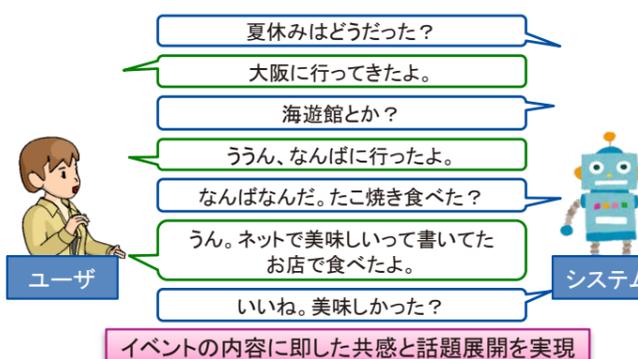
ユーザの発話からイベントを正しく理解するためには、発話に含まれる様々な表現の抽出が必要でした。これに対し、従来手法では抽出できなかった雑談中に現れやすい単語やフレーズを抽出可能にし、対話からの構造化されたイベント理解を実現します。

## めざす未来

本技術により、ユーザの発話からシステムが理解したい情報を整理して抽出できるようになります。今後は、抽出した結果を、外界の情報やシステムの知識にグラウンディングさせることで、ユーザとシステムの共通理解を実現し、相手が人であるときと同じように安心して対話ができる世界を目指します。

## イベント理解に基づくシステム発話生成

ユーザの発話に含まれるイベント(出来事)を構造化された情報として理解し、システムの知識と照合することで、イベントの内容に即した発話生成を行います。



## 【理解されたイベントの構造化例】

いつ	夏休み
どこで	大阪/なんば
誰と	
何を	たこ焼き食べた
感想	

## 【類似イベントの例(システムの知識)】

いつ	連休に	いつ	9月に
どこで	大阪/海遊館	どこで	大阪/なんば
誰と	家族で	誰と	友達と
何を	イルカを見た	何を	たこ焼き食べた
感想	かわいかった	感想	美味しかった

## ユーザ発話からのフレーズ抽出

ユーザ発話から正しくイベントを理解するためには、発話に含まれる様々な表現の抽出が求められます。雑談中に現れやすい単語やフレーズ表現を分析し、これまで検出が困難であった一般名詞やフレーズの抽出を可能にしました。

【本技術で抽出できた場所フレーズの例と従来技術の結果】

ユーザ発話(赤字:場所フレーズ)	従来技術の抽出結果	成否
夏休みにイタリアに行きました。	イタリア	○
京都駅近くの公園でお花見したよ。	京都駅	×(不足)
暇なときは電気屋に行きます。	なし	×(未抽出)

従来の固有表現抽出技術では、雑談対話に登場するフレーズ全体の約7割が抽出不可(場所フレーズを対象とした場合)



一般名詞やフレーズの抽出を実現

## 関連文献

- [1] H. Narimatsu, H. Sugiyama, M. Mizukami, “Detecting Location-Indicating Phrases in User Utterances for Chat-Oriented Dialogue Systems,” in Proc. The Fourth Linguistic and Cognitive Approaches to Dialog Agents Workshop (LACATODA), 2018.
- [2] 杉山弘晃, 成松宏美, 水上雅博, 有本庸浩, “文脈に沿った発話理解・生成を行うドメイン特化型雑談対話システムの実験的検討,” 人工知能学会 言語・音声理解と対話処理研究会 (SLUD) 第84回研究会(第9回対話システムシンポジウム), 2018.
- [3] M. Mizukami, H. Sugiyama, H. Narimatsu, “Event Data Collection for Recent Personal Questions,” in Proc. LACATODA, 2018.

## 連絡先

成松 宏美 (Hiromi Narimatsu) 協創情報研究部 インタクション対話研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 15 うるさい車内でも音声操作や会話をサポート

～世界トップの集音技術・音声認識技術を実用化～

## どんな研究

走行音などによりうるさい車内においても、**音声操作や会話をストレスなく行えるようサポートします**。音声操作を実現するため、音声の歪みを少なく保ちながら混入したノイズを除去しました。また、車内の離れた座席間や電話相手との会話の際にストレスとなる遅延をできる限り低減しました。

## どこが凄い

音質を保ちながら雑音だけを除去するためには、膨大なメモリと演算量が必要でした。**これまで培ってきた音響関連のノウハウ活用**により、この問題を解決、かつ、低遅延な処理を実現できました。また、複数マイクアレイ連携により、**ハウリングの兆しを迅速に見つけることができます**。

## めざす未来

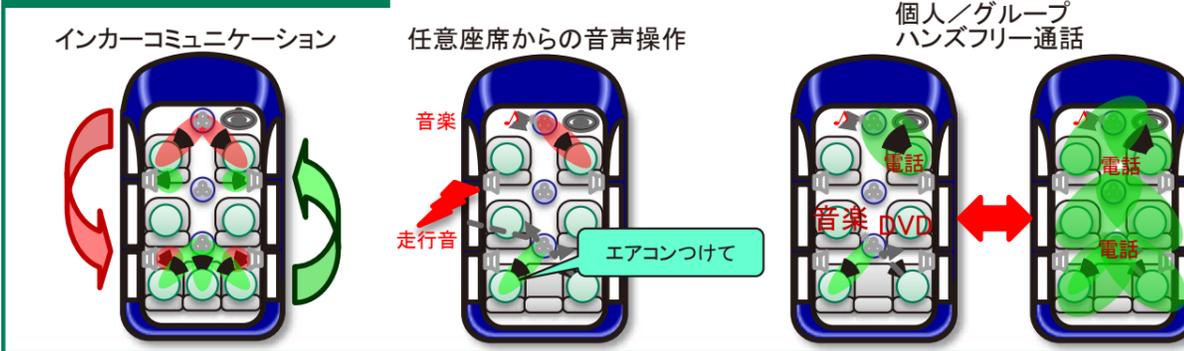
今後は、自動車内に混入する自車のエンジン音やロードノイズ、他車からのノイズを低減させることにより、**自車内の音環境の改善**を目指します。また、緊急車両の接近や故障などによる異常音などのイベント検知を行い、**運転支援や早期のメンテナンスに役立つ技術の確立**を目指します。

## 従来技術との違い

IM: Intelligent Microphone  
 ASTER: Anti-distortion Suppression of noise with mask-based TransER function estimation

	従来NTT技術1 (IM*)	従来NTT技術2 (ASTER*)	本技術 (IM-ASTER)
概要	線形処理と非線形処理の組み合わせにより、目的方向の音とそれ以外の音を分離	音声と雑音が含まれる信号から、音声歪みを抑えつつ最大限に雑音を除去	インテリジェントマイクとASTERの強みを組み合わせ、低演算量・少メモリを実現
処理結果イメージ	 雑音(入力信号) 音声 雑音を除去する際に <b>音声も削ってしまう</b>	 音声歪みを抑えつつ雑音を除去	 音声歪みを抑えつつ雑音を除去
音声歪みの抑制	△	○	○
高騒音への対応	○	△	○
演算量/メモリ量	○	×	○

## デモ内容



## 関連文献

[1] Y. Hioka, K. Furuya, K. Kobayashi, K. Niwa, Y. Haneda, "Underdetermined sound source separation using power spectrum density estimated by combination of directivity gain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 6, pp. 1240-1250, 2013.  
 [2] T. Yoshioka, N. Ito, M. Delcroix, A. Ogawa, K. Kinoshita, M. Fujimoto, C. Yu, W. H. Fabian, M. Espi, T. Higuchi, S. Araki, T. Nakatani, "The NTT CHiME-3 system: Advances in speech enhancement and recognition for mobile multi-microphone devices," in *Proc. IEEE ASRU*, pp. 436-443, Dec. 2015.

## 連絡先

原田 登 (Noboru Harada) メディアインテリジェンス研究所  
 Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 16 少量の入出力ペアから高精度に音声認識を学習

～音声合成を活用した半教師ありEnd-to-End学習～

## どんな研究

音声認識は音声を書き起こした文字列へ変換する仕組みです。音声認識モデルの学習用に人手で用意する対応付いた音声と文字列の**ペアデータが少ない**場合、高精度なモデルの実現は困難でした。この研究ではペアではない**音声のみ・文字列のみのデータも活用**できる学習方法を提案します。

## どこが凄い

音声認識モデルと音声合成モデルの構造が似ていることに注目し、二つを組み合わせ**音声・テキストのみで学んだ特徴量と音声認識の特徴量が近づく**。音声認識に活かせる半教師あり学習を実現しました。実験では少量のペアデータのみで学習する場合と比べて、文字誤り率を半分に削減しました。

## めざす未来

既存の方法よりも**少ない音声と文字列の学習用ペアデータで高精度な音声認識モデル**を学習できます。将来的には、マイナーな言語や大量に準備しにくい音声(子供など)といったペアデータがほとんど得られない環境の音声認識など、より挑戦的な場面で活用できる技術の実現を目指します。

## 音声認識の仕組み

- 音声と文字列の学習用ペアを準備
- 音声を受け取るエンコーダと文字列を出力するデコーダを学習
- 未知の入力音声から正しく文字列が出力できるか評価



## 音声認識の課題

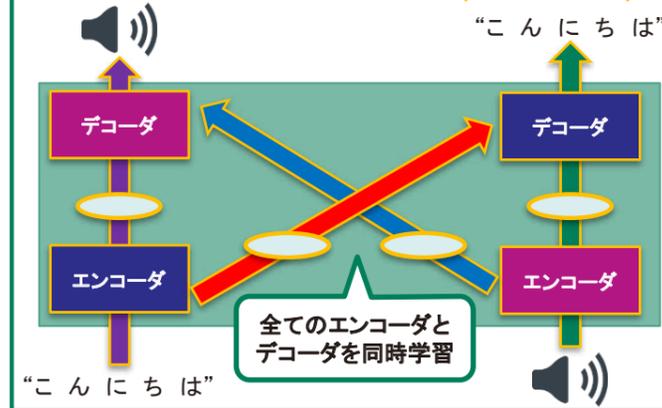
- エンコーダ・デコーダ型モデルの学習は、沢山の音声と書き起こし文字列のペアデータが必要
- 大量のペアデータの準備にはコストと時間がかかってしまう
- もし音声だけでエンコーダを、文字列だけでデコーダを学習できればデータの準備が簡単に



## 半教師あり学習への拡張

ポイント1: 音声認識と音声合成を組み合わせる

- ペアデータが必要なタスク(音声認識↑、音声合成↑)
- ペアデータが不要なタスク(音声復元↓、文字列復元↓)



学習タスク					音声認識エラー率
音声認識	文字列復元	音声復元	音声合成	音声認識	
✓					15.0%
✓	✓				9.0%
✓		✓			8.7%
✓	✓	✓	✓		8.4%

ポイント2: それぞれの中間特徴量が似るように学習

## 関連文献

[1] S. Karita, S. Watanabe, T. Iwata, A. Ogawa, M. Delcroix, "Semi-supervised end-to-end speech recognition," in *Proc. Interspeech*, 2018.  
 [2] S. Karita, S. Watanabe, T. Iwata, M. Delcroix, A. Ogawa, T. Nakatani, "Semi-supervised end-to-end speech recognition using text-to-speech and autoencoders," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

## 連絡先

苅田 成樹 (Shigeki Karita) メディア情報研究部 信号処理研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 17 いつ、誰が、何を話した？全部で何人いた？

～何人の会話でも聞き分けられる深層学習モデル～

## どんな研究

会話収録音から、「いつ、誰が、何を話したか」という情報を推定します。従来の類似技術は、収録状況に関して様々な条件(話者は移動不可、話者数は既知、等)を仮定し、その条件が満たされた時のみうまく動作するものでした。しかし、実データではこれらの条件が満たされないことも多々あります。

## どこが凄い

提案法は、複数人の声が重なっていても、話者数を数え上げながら、「いつ、誰が話したか」という情報を、話者の声の特徴に基づき精度良く推定します。深層学習に基づく、任意の会話状況を表現できるモデルであり、適切な学習データがあれば、あらゆる実会話データに対応できることが期待されます。

## めざす未来

人と人との会話から「いつ、誰が、何を話したか」という情報を自動獲得する技術は、会話を自動分析する技術の基盤となり、議事録自動作成や、私たちのコミュニケーションを助けるロボットの実現に寄与します。今後は、より実際のデータを用い、提案原理の有効性を検証していきます。

### 会話分析(「いつ、誰が、何を話したか」の推定)の難しさ

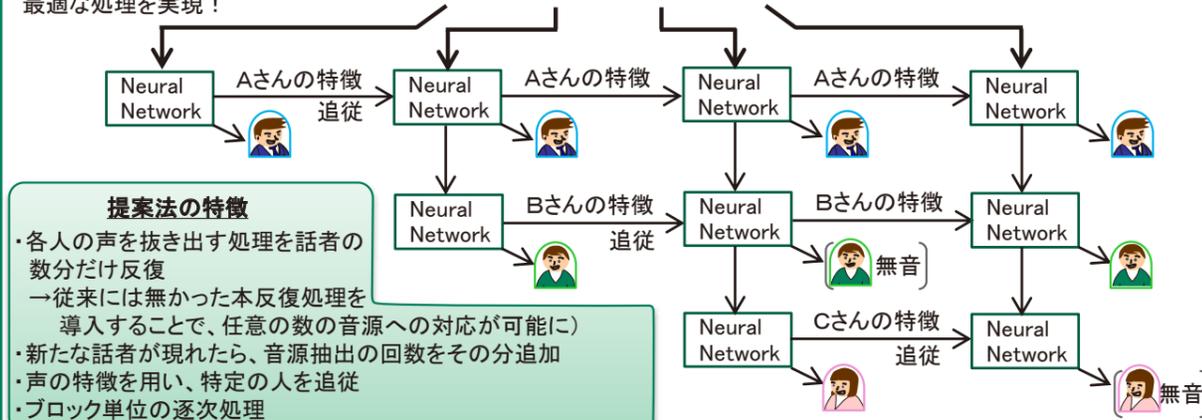


### 提案法

深層学習を用いて、収録音から「いつ、誰が、何を話したか」を推定！学習データに基づき、最適な処理を実現！

マイク収録音 (何人いて、誰がいつ、何を喋っているか不明)

第1処理ブロック 第2処理ブロック 第3処理ブロック 第4処理ブロック



### 提案法の優位性(従来の音源分離手法との比較)

- 提案法は、音源分離と話者数推定を同時に実現。一方、従来技術は、話者数を既知としていた。
- 提案法は、話者の声の特徴に注目し、目的話者の声を聴き続けられる。喋っていない時に、話者が移動しても対応可能。一方、従来技術は、方向情報を基にある話者に注目し続けるものが多く、喋っていない話者が移動すると対応が困難。

### 関連文献

[1] K. Kinoshita, L. Drude, M. Delcroix, T. Nakatani, "Listening to each speaker one by one with recurrent selective hearing networks," in *Proc. IEEE International Conference on Acoustics, Speech & Signal Processing (ICASSP)*, pp. 5064-5068, 2018.  
 [2] T. von Neuman, K. Kinoshita, M. Delcroix, S. Araki, T. Nakatani, R. Haeb-Umbach, "All-neural online source separation, counting, and diarization for meeting analysis," in *Proc. IEEE International Conference on Acoustics, Speech & Signal Processing (ICASSP)*, 2019. (to appear)

### 連絡先

木下 慶介 (Keisuke Kinoshita) メディア情報研究部 信号処理研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



# 18 声と話し方を好みのスタイルに変える

～系列変換モデルに基づく声質と韻律の同時変換～

## どんな研究

話し声は抑揚や声質、リズムによって雰囲気が大きく変化します。系列変換モデルを用い、音声の様々な変動要素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、ユーザーが望んだ人の音声に自動変換したり、聞き取りやすい母語話者の発音とリズムに変換することが可能です。

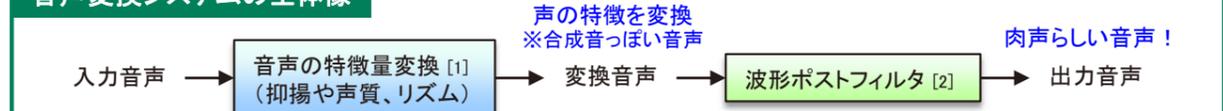
## どこが凄い

従来の音声変換技術の研究では声質の変換のみに着目したものが多かったのですが、本技術では声質だけでなく抑揚やリズムも自動変換することが可能です。また、機械的に感じる合成音声の波形を心地よく聞き取りやすい音声へと直接変換する深層学習アプローチを世界で初めて実現しました。

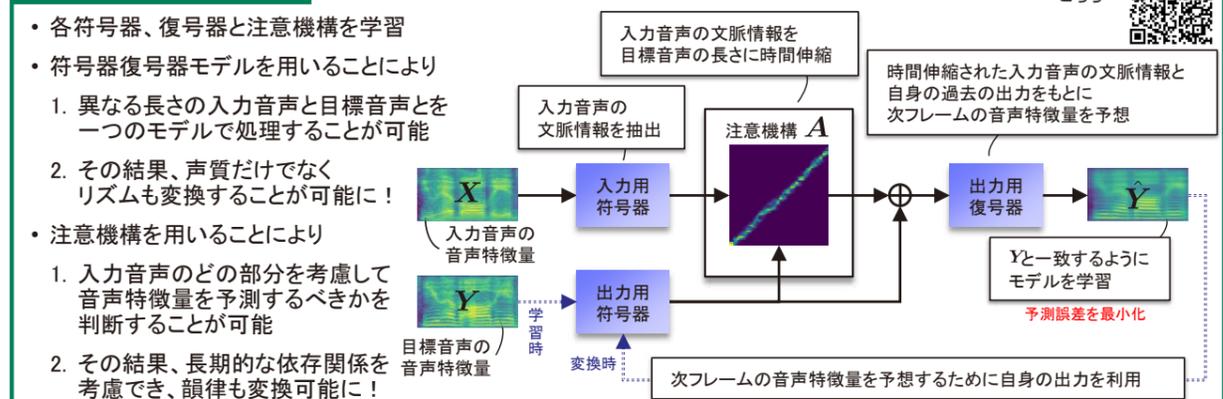
## めざす未来

本技術を用いることで、既存の音声変換で実現し得なかった高品質な音声変換が可能になります。今後は、感情音声変換や非母語話者の聞き取りにくい音声を母語話者風の音声に変換する研究、発声障がい者支援への応用などを行い、多様で表情豊かな音声変換技術の実現を目指しています。

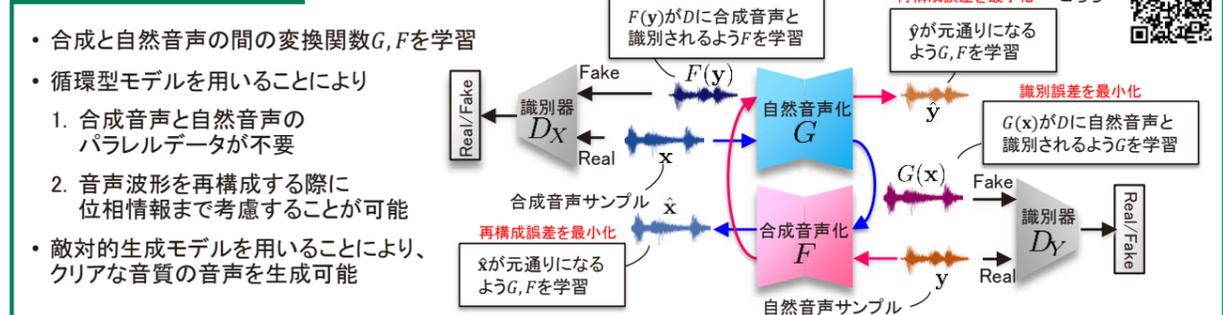
### 音声変換システムの全体像



### 音声の特徴量変換 [1]



### 波形ポストフィルタ [2]



### 関連文献

[1] 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)*, May 2019.  
 [2] K. Tanaka, H. Kameoka, T. Kaneko, N. Hojo, "WaveCycleGAN2: Time-domain Neural Post-filter for Speech Waveform Generation," *arXiv:1904.02892*, Apr. 2019. (submitted to *Interspeech2019*).

### 連絡先

田中 宏 (Kou Tanaka) メディア情報研究部 メディア認識研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



# 19 顔に合わせて声を作り、声に合わせて顔を作る

～深層生成モデルによるクロスモーダル音声変換～

## どんな研究

私たちは、声の印象からその人がどんな顔か、また、顔の印象からその人がどんな声かある程度想像できます。これは、声と顔には何らかの相関があるからだと考えられます。本研究では、**与えられた顔画像の印象に合った声を作り出すクロスモーダル音声合成**の問題に初めて取り組みました。

## どこが凄い

音声変換器を**深層生成モデル**で表し、出力音声と入力顔画像との相互情報量を規準として音声変換器を学習する**情報論的アプローチ**を考案しました。これにより、入力顔画像に合った声質に入力音声を変換する**クロスモーダル声質変換技術**を実現することに初めて成功しました。

## めざす未来

私たち人間は、異なる感覚器官から得られる情報(視覚情報や聴覚情報など)を無矛盾に関連付けてモノや出来事を認識しています。本研究では、**人間のこの知的な認識機能を実現することを究極の目標**としています。

### クロスモーダル声質変換問題の提案

声のみから顔を想像したり顔のみから声の雰囲気や特徴を想像したりできる⇒声質と容貌には相関があることを示唆  
 → 音声特徴量系列と顔画像との間の相関を捉え、入力音声を入力顔画像に適合した声質に変換する  
 ・ さらに入力音声に適合した容貌の顔画像を生成する

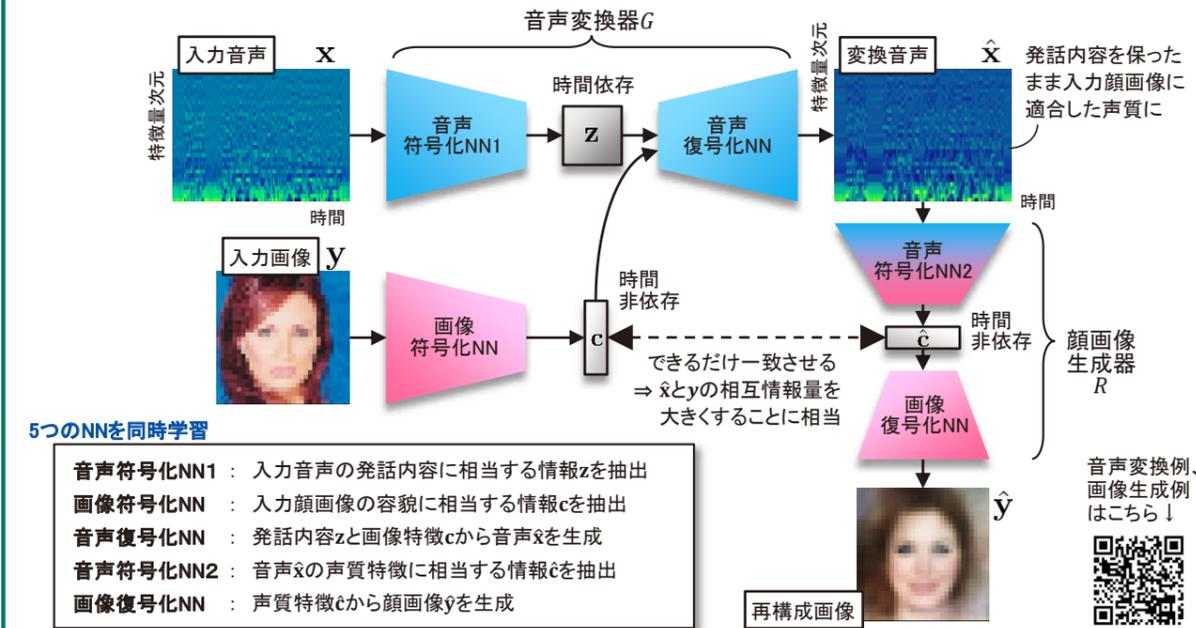
### 情報論的アプローチによる問題の定式化

**変換器G**: 音声xと顔画像yを入力として変換音声 $\hat{x} = G(x, y)$ を出力するニューラルネットワーク(NN)

**変換器Gの学習**:  $\hat{x} = G(x, y)$ とyの相互情報量を規準としてGを学習する問題として定式化

$$I[G(x, y) \| y] \geq \mathbb{E}_{(x, y) \sim p(x, y)} [\log R(y | G(x, y))] \rightarrow \text{下界をGとRに関して最大化}$$

音声と顔画像のペアデータ (音声G(x, y)から予測される顔画像yの条件付分布を近似するNN)



### 関連文献

[1] H. Kameoka, T. Kaneko, K. Tanaka, N. Hojo, "StarGAN-VC: Non-parallel many-to-many voice conversion using star generative adversarial networks," in *Proc. 2018 IEEE Workshop on Spoken Language Technology (SLT 2018)*, pp. 266-273, 2018.  
 [2] H. Kameoka, T. Kaneko, K. Tanaka, N. Hojo, "ACVAE-VC: Non-parallel voice conversion with auxiliary classifier variational autoencoder," *arXiv:1808.05092 [Stat.ML]*, 2018.  
 [3] H. Kameoka, K. Tanaka, A. Valero Puche, Y. Ohishi, T. Kaneko, "Crossmodal Voice Conversion," *arXiv:1904.04540 [cs.SD]*, 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-liaison-ml at hco.ntt.co.jp



# 20 声と画像から知らないモノを学びとるAI

～音声と画像によるクロスモーダル概念獲得～

## どんな研究

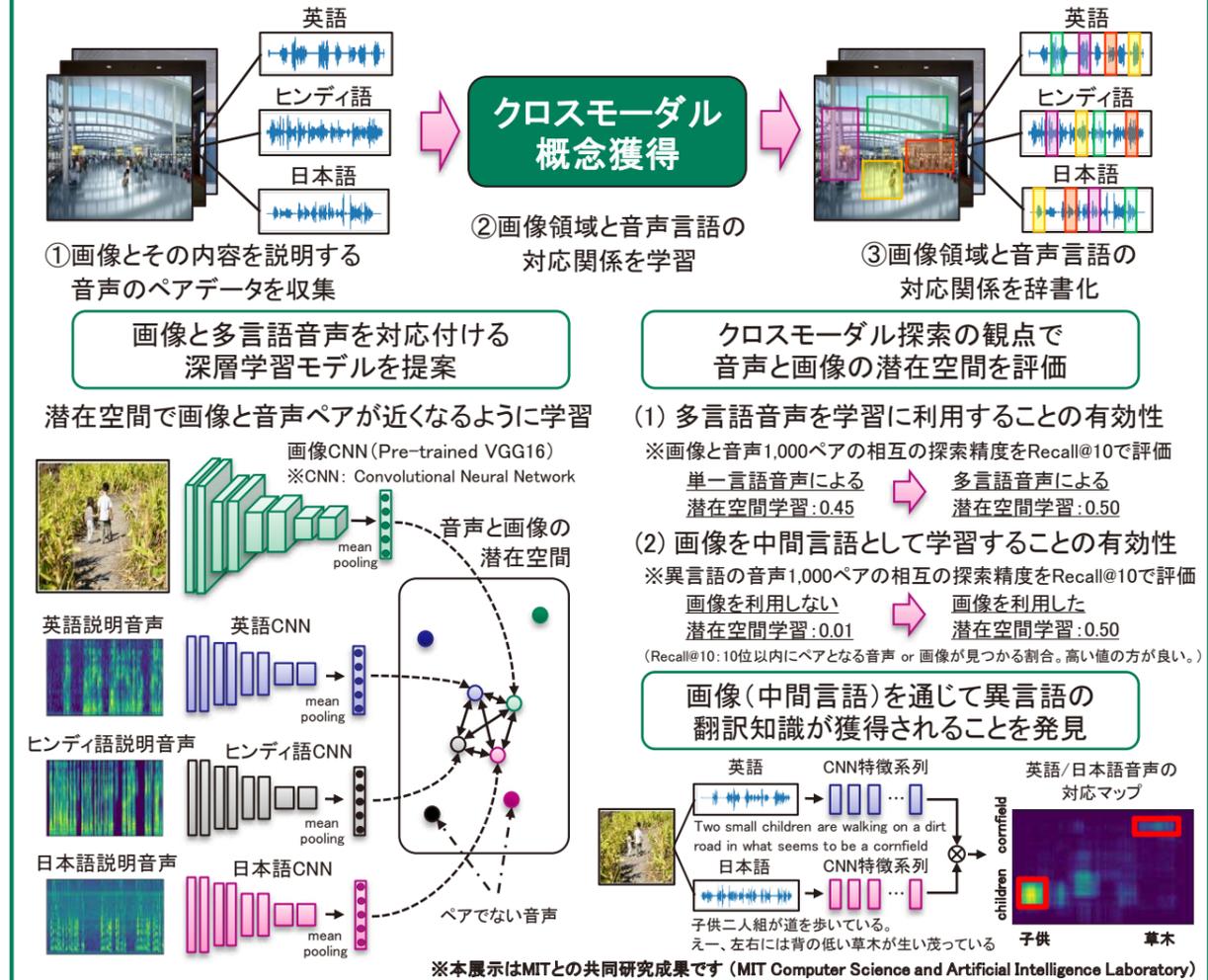
AIがモノを認識するためには、モノの見え方とその言語表現(名前)を紐付けた辞書が必要です。この展示では、**画像を説明する音声データ**だけから、モノが映る画像領域と声による言語表現を教師ラベルなしで対応付けて、**AIが知らないモノを自ら学び取る(辞書を自動作成する)技術**を紹介します。

## どこが凄い

日本語音声データセットを新たに構築し、**既存の英語/ヒンディ語音声と合わせて、画像との潜在空間を深層学習**することで、従来よりも精度良く、画像領域と言語表現が紐付けられることを確認しました。この手法により、**画像を通じて異なる言語の単語翻訳知識**が学習されることも大きな特徴です。

## めざす未来

TV放送のような世の中に多く存在するメディアデータを与えるだけで、AIが音と映像の共起(時空間的な偏り)を見つけながら自律的にモノやコトの概念を学び、賢くなる未来を目指しています。音や映像、言語を自在に横断する**超大規模アーカイブ検索や自動アノテーション**などへの応用を検討しています。



### 関連文献

[1] 大石康智, 木村昭悟, 川西隆仁, 柏野邦夫, D. Harwath, J. Glass, "画像を説明する多言語音声データを利用したクロスモーダル探索," *電子情報通信学会技術報告, パターン認識・メディア理解(PRMU)研究会, (to appear)*  
 [2] D. Harwath, G. Chuang, J. Glass, "Vision as an Interlingua: Learning Multilingual Semantic Embeddings of Untranscribed Speech," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2018)*, 2018.

### 連絡先

大石 康智 (Yasunori Ohishi) メディア情報研究部 メディア認識研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



# 21

## どんな音？物音を言葉で説明しよう

～系列変換モデルに基づく音響信号からの説明文生成～

### どんな研究

これまで、音声認識システムを用いても、人の話し声以外の音を的確にテキストに変換することはできませんでした。本展示では、**任意の音**に対して、それが**どんな音か**を説明するテキスト(擬音語や説明文)を生成する研究を紹介します。本研究により物音などを文字で伝えることが可能になりました。

### どこが凄い

音を文字にすることで、**一見するだけでどのような音かを把握**できます。どの程度の**詳しく**で音を説明して欲しいかを数値で指定することで**目的に沿った文を生成**できます。擬音語や説明文が表現するような**主観的な音の近さ**に基づいて、イメージした音に最も近い音を**検索**することなどにも応用できます。

### めざす未来

効果音や異常音などの**音の検索**がより便利になります。**動画中の音**を文字で表現することで、公共の場所や騒々しい場所などでの**動画視聴**も便利になることでしょう。また、**AIが人間に近い音の感覚を身につける**ことで、宅内ロボットなどの日常のコミュニケーションが円滑になることも期待されます。

#### 音の説明文生成とは

- 任意の音に対して **どんな音か**を言葉で説明する深層学習システム
- 「ベル」「車」といった単なる音源種類の認識 (=従来研究あり) だけでなく **変化の様子**なども含む詳しい記述も可能 **NEW**

環境音など → 説明文生成 → (この方向) 金属の音車が進む音が聞こえるような甲高い摩擦音が鳴っています

#### 実験結果 (説明文モードの例)

説明文が音に...

- よく当てはまる
- 当てはまる
- 部分的に当てはまる
- 当てはまらない

約 3/4

説明文の適切さ (詳細度制御なしの場合)

#### 系列変換モデルによる実現

甲高い摩擦音が～

デコーダ (LSTM) → z (潜在空間) ← エンコーダ (LSTM)

(詳細度) c

#### 詳細度制御は...

- 制御ありのほうはずっとよい
- 制御ありのほうがよい
- どちらも言えない
- 制御なしのほうがよい
- 制御なしのほうはずっとよい

約 2/3

説明文生成における詳細度制御の効果

生成された説明文の例 (バスドラムの音)

c	生成された文
なし	低い音が一瞬だけ響く
20	低い音が一瞬だけ鳴る
50	畳に打ち付けるような低い打撃音が一瞬だけ響く
80	つまらなそうにとても低い音でドラムが一度だけ奏でられている
110	つまらなそうにとても低い音で何かを叩くような軽くて少し低い音が小さな音で一瞬鳴ってすぐに消える

※本展示は東京大学との共同研究成果です (大学院情報理工学系研究科連携講座)。

#### 関連文献

[1] Shota Ikawa, Kunio Kashino, "Generating sound words from audio signals of acoustic events with sequence-to-sequence model," In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)*, April 2018.

[2] Shota Ikawa, Kunio Kashino, "Acoustic event search with an onomatopoeic query: measuring distance between onomatopoeic words and sounds," In *Proc. Detection and Classification of Acoustic Scenes and Events (DCASE 2018)*, November 2018.

#### 連絡先

柏野 邦夫 (Kunio Kashino) メディア情報研究部  
Email: cs-liaison-ml at hco.ntt.co.jp

Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

# 22

## 音からものとかたちを認識

～音と画像のクロスメディア情報処理によるシーン理解～

### どんな研究

**音から物体や人物を認識する研究**です。具体的には、マイクで収録した音だけを分析して、そこにある物体や人物などの種類や形状を含む画像認識結果を予測します。生活空間や公共空間での見守り・防犯等への応用に向けて研究開発を進めています。

### どこが凄い

音のみから物体や人物の種類・形状を予測するには、音の方向と音色の時空間的な関係を捉える必要があります。本研究では、この複雑な関係を効率的にモデル化できる特徴統合層を備えた深層ニューラルネットワークを設計し、より**多様な種類の物体の認識を可能**にしました。

### めざす未来

**安心・安全な見守り・監視システム**の実現が期待できます。本技術は、カメラが利用できないようなプライバシー性の高い家庭や公共空間への適用が可能です。また、音ならではの伝搬特性を活かし、カメラでは死角になってしまうような箇所の情報も予測できるようになる可能性があります。

#### “音から画像認識”するクロスメディア情景分析

複数のマイクで収録した音の方向と音色を基に、音を発している物体や人物のクラス(種類)と形状を予測

セットアップ: マイクフォンアレイ (マルチチャネル音声信号) と カメラ (RGB画像) を同期して学習時のみ利用(認識時は無し)。

予測: 特徴抽出 (方向情報, 音色情報) → 深層ニューラルネットワーク → 画像認識 (画像認識結果: クラス+形状)

カメラでは写らない・写したくない空間の様子を、音だけからでも描写可能に

暗い部屋の様子も... 音声から描写可能

これまで、少数のクラスの物体や人物の認識が可能であることを確認

実験装置: アクリルケース、マイクフォンアレイ、おもちゃの電車

より多様なクラスの物体や人物の認識、実音声からの認識を実現

ここがポイント

方向と音色の関係を効率的に捉えることのできる特徴統合層を導入することにより、パラメータ数を抑えつつ精度を向上

全結合層による特徴統合: 方向特徴, 音色特徴 → 統合特徴 (パラメータ数が膨大!)

低次元化特徴の重み付き外積を取ることでパラメータ数を抑えつつ高次関係を捉える

低次元化: 方向特徴, 音色特徴 → 重み付き外積 → 統合特徴

※本展示は東京大学との共同研究成果です (大学院情報理工学系研究科連携講座)。

#### 関連文献

[1] G. Irie, M. Ostrek, H. Wang, H. Kameoka, A. Kimura, T. Kawanishi, K. Kashino, "Seeing through sounds: predicting visual semantic segmentation results from multichannel audio signals," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

[2] 王昊辰, 入江豪, 亀岡弘和, 木村昭悟, 平松薫, 柏野邦夫, "双線形特徴統合を用いた音からのセマンティックセグメンテーション," 第21回 画像の認識・理解シンポジウム (MIRU), 2018.

#### 連絡先

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  
Email: cs-liaison-ml at hco.ntt.co.jp

Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

データの学習の科学  
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

データの学習の科学  
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

# 23

## 鳥の声で喋り、水の泡で演奏する

～聴覚信号処理モデルを用いた音のテクスチャ変換～

### どんな研究

ブクブク泡立つ水音や森の木々のざわめきのように、自然な環境音からは、それぞれ**特有の質感**が感じられます。音声や音楽に、このような**質感を人工的に持たせる**方法を開発しました。画像の質感を操作する研究から着想を得て、その方法を音に適用できるように発展させました。

### どこが凄い

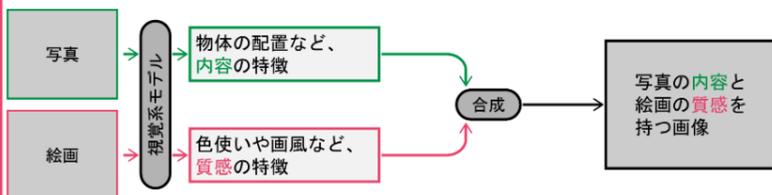
音が聴こえるメカニズムを考慮した計算モデルを用いて、聴感上効果的な**質感操作**が実現できました。この質感操作の仕組みは、画像の質感操作と同じ枠組みで構成されています。これにより、見た物と聴いた物の質感が、脳内で類似した機構で処理される可能性が強く示されました。

### めざす未来

学術的には、モデルのなかで生じている状態を、実際に人が音を聴いているときの脳活動と比較することで、**音の質感知覚メカニズムの理解**につながります。応用の観点からは、実際には存在しない声で喋ったり、実際には存在しない楽器で演奏したりすることができる可能性があります。

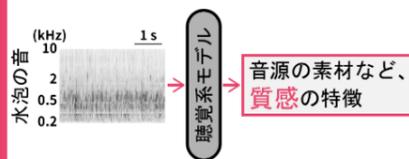
#### 既存研究：画像の質感変換

写真の**内容**と絵画の**質感**を組み合わせ、それらの両方を持つ音を合成。  
(写真の**内容**を保持したまま**質感**を絵画のものに変換)



#### 既存研究：音の質感表現

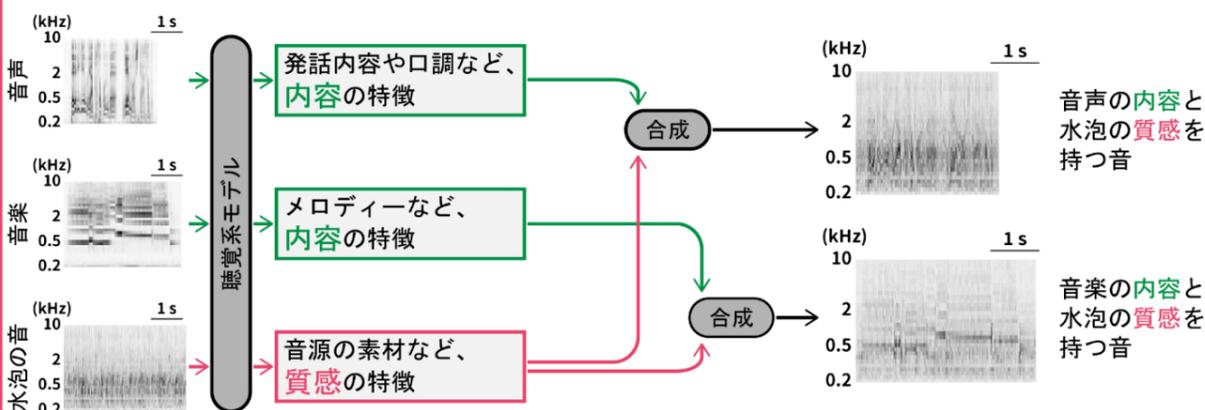
聴覚系モデルを用いて、環境音の**質感**を表す特徴を計算。



#### 本研究：音の質感変換

聴覚系モデルを用いて、音声や音楽の**内容**を表す特徴と、環境音の**質感**を表す特徴を計算。

音声や音楽の**内容**と環境音の**質感**を組み合わせ、それらの両方を持つ音を合成。  
(音声や音楽の**内容**を保持したまま**質感**を水泡のものに変換)



#### 関連文献

- [1] T. Koumura, H. Terashima, S. Furukawa, "Chimeric sounds with shuffled "texture" and "content" synthesized by a model of the auditory system," in Proc. International Symposium on Universal Acoustical Communication, 2018.
- [2] 上村卓也, 寺島裕貴, 古川茂人, "聴覚系モデルを用いた音のテクスチャ変換," 第31回人工知能学会全国大会, 2017.

#### 連絡先

上村卓也 (Takuya Koumura) 人間情報研究部 感覚共鳴研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 24

## 踊る紙人形

～紙に動きの印象を与える錯覚～

### どんな研究

**実物体**である紙人形に**動きの錯覚を与える技術**です。先行研究では、画面上の仮想的な対象が輪郭の明暗と背景の輝度変化との相互作用で動いて見えることが報告されていましたが、背景の輝度変化を利用して実物体に動きの印象を与える手法については提案されてきませんでした。

### どこが凄い

本研究では先行研究の成果を応用し、**紙人形の輪郭に明暗のエッジを加えつつその紙人形を明るさの変化する背景上に置くだけで、紙人形が動いて見える錯覚を発見**しました。また、明暗エッジパターンを工夫することで、拡大縮小や回転も表現できることがわかりました。

### めざす未来

Danswing papersを利用することで、**本来動きを与えることが難しい紙媒体のオブジェクトに面白い、注意を引くような動きの印象を付与することができます**。錯覚を情報提示に利用することにより、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編集することが可能となります。

#### Danswing papersの制作方法

※Danswing papersの「Danswing」は、DanceとSwingの2語に由来する造語です。

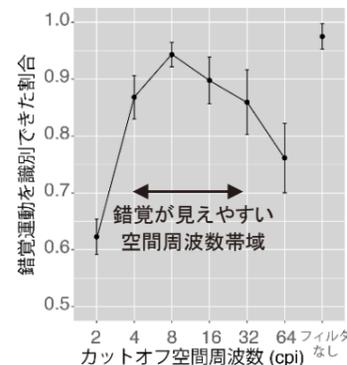
●例えば少し回転して見えるハートを作りたいときは、以下のように紙人形を制作します。

- ①グレイのハート、少し回転させた白いハート、
- ②①で作成した3つハートをグレイのハートが一番逆方向へ回転させた黒いハートの画像を準備します。



#### Danswing papersの視覚原理

●本錯視には特定の空間周波数帯域の動き情報が関わっていることを明らかにしました。



映像に含まれる空間周波数の中から、所望の範囲の空間周波数成分を選択的に取り出す空間周波数フィルタをかけた映像を使って錯覚運動を識別する実験を行ったところ、左グラフのように特定の空間周波数帯域で錯覚が見えやすいことがわかりました。

右のQRコードからDanswing papersのYouTube動画をご覧ください。



#### 関連文献

- [1] T. Kawabe, "Danswing papers," in Proc. SIGGRAPH Asia 2018 (SA '18) Posters Article No. 4.
- [2] T. Kawabe, "Danswing papers," Top 10 finalist of Best illusion of the year contest. <http://illusionoftheyear.com/2018/10/danswing-papers/>

#### 連絡先

河邊隆寛 (Takahiro Kawabe) 人間情報研究部 感覚表現研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 25

## あなたの目の機能を気軽に楽しく測ります

～ミニゲームやタブレットを使った視覚能力のセルフチェック～

### どんな研究

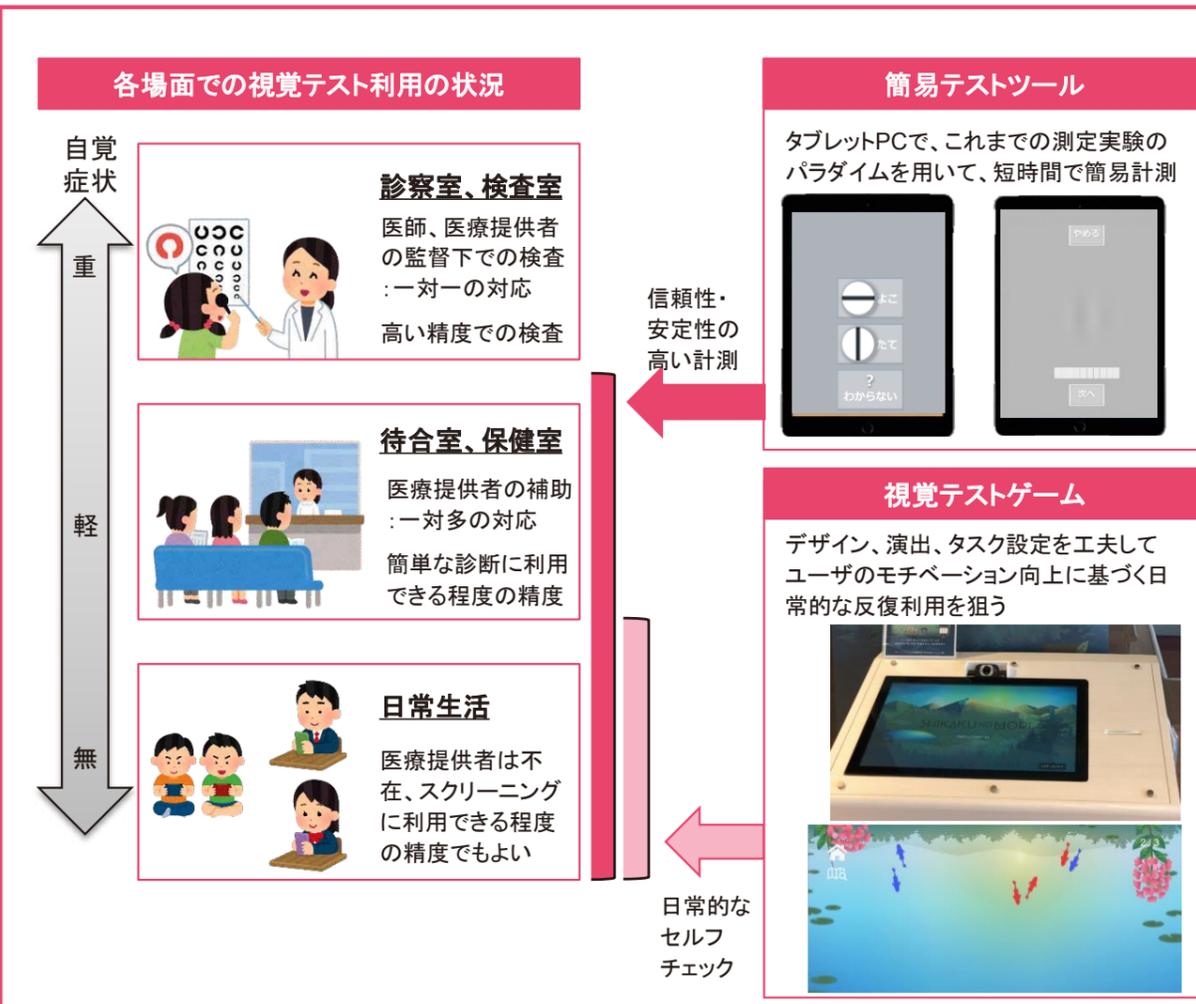
目の機能を気軽に、楽しく測るための方法についての研究です。汎用的なタブレット機器向けに視覚テストを作成しました。ゲーム形式で楽しく、安全に、目の機能のセルフチェックができます。これまでに、コントラスト感度、複数の動く物体を追跡する視覚能力を含むいくつかのテストを作成しています。

### どこが凄い

これまでの視覚機能のテストでは、時間がかかったり、エキスパートの同席が必要でした。これまでの視覚科学実験のノウハウと、ウェブブラウザ向けの精確描画技術と組み合わせて、汎用的なタブレットPCでも、およそ3分でスクリーニングテストとして利用可能な精度で各機能の測定を可能にしました。

### めざす未来

提案するシステムを使うと、日常的に目の状態をセルフチェックすることができるようになります。また、簡易で短時間の測定で、眼病患者を含む多くの人のデータを通して知見を蓄積することで、眼病の早期発見や、リハビリテーション応用、そして、複雑な視覚プロセスの科学的解明が期待できます。



### 関連文献

[1] K. Hosokawa, K. Maruya, S. Nishida, "Testing a novel tool for vision experiments over the internet," *Journal of Vision*, Vol. 16, p. 967, 2016.

### 連絡先

丸谷 和史 (Kazushi Maruya) 人間情報研究部 感覚表現研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 26

## 勝者のメンタルとは？

～実戦環境での生理状態とパフォーマンスの関連性～

### どんな研究

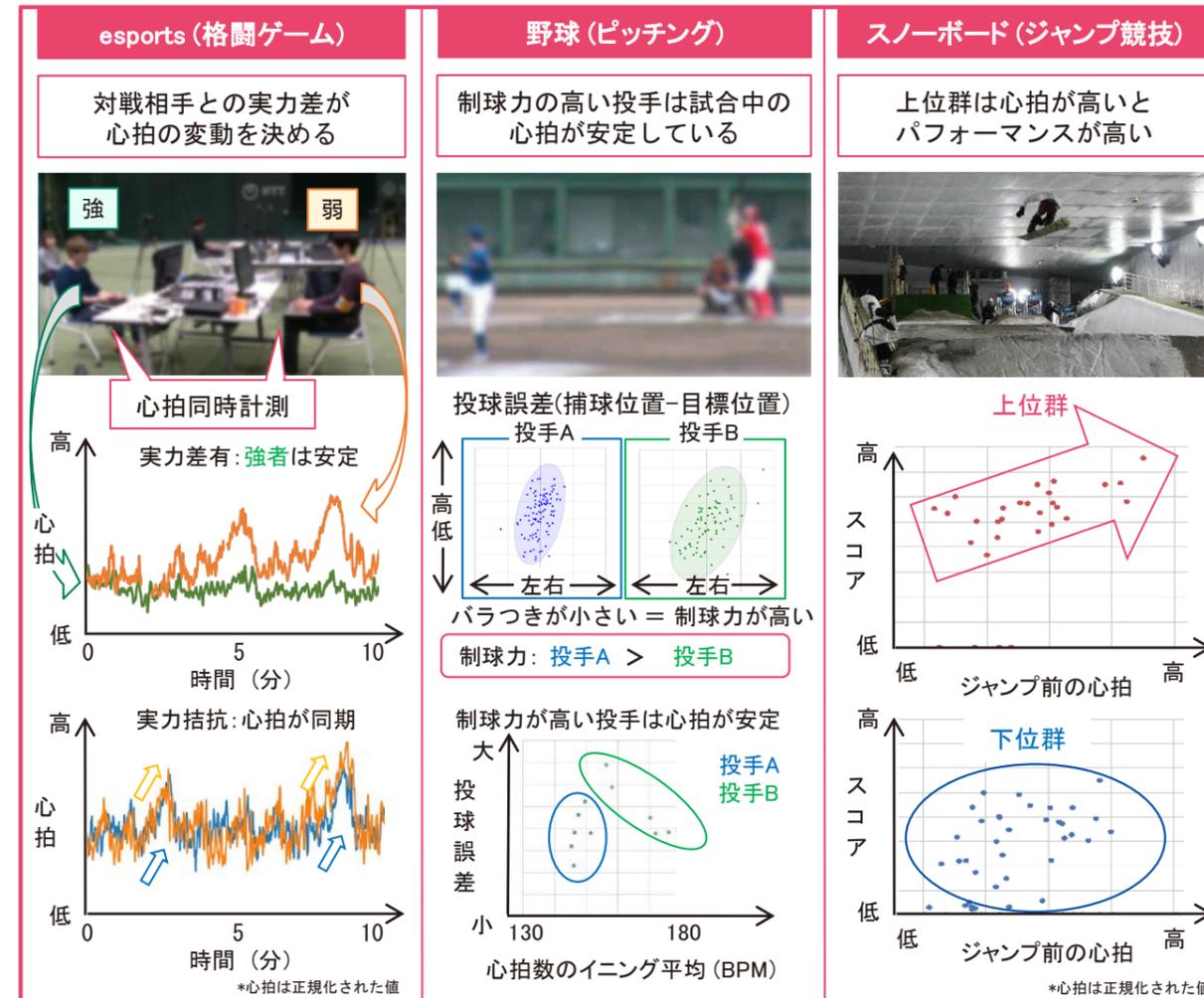
スポーツにおいてメンタルは試合の勝敗を左右する重要な要素とされていますが、その根拠となる客観的データは多くありません。本研究では、トップアスリートの実戦中の生理状態の変化(メンタルの指標)とパフォーマンスの関係性を定量的に評価しました。

### どこが凄い

これまで困難だった実戦におけるトップアスリートの生理状態の計測に成功しました。その結果、トップアスリートには、心拍増加に伴うパフォーマンスの向上や、直近の状況によらない安定した生理状態と高いパフォーマンスの維持などの特徴があることがわかりました。

### めざす未来

生理状態とパフォーマンスの定量的な関係から“メンタル”の構成要素とメカニズムを科学的に理解し、それに基づいて、アスリートの生理状態を最適に整えてパフォーマンスを向上させるための体系的な方法論を確立します。



### 関連文献

[1] K. Watanabe, N. Saijo, M. Kashino, "The physiological change reflecting the fight-or-flight response of an esports player correlates strongly with that of the opponent," in *Proc. NEURO2019*, 2019.  
[2] 福田岳洋, 持田岳美, 西條直樹, 柏野牧夫, "投手の制球力は状況に影響される - 実戦中の投球分布から -," *日本野球科学研究会 第5回大会*, 2017.  
[3] 松村聖司, 渡辺謙, 木村聡貴, 柏野牧夫, "スノーボードのジャンプ競技大会における選手の生理状態計測," *日本スキー学会 第29回学会大会*, 2019.

### 連絡先

渡辺 謙 (Ken Watanabe) スポーツ脳科学プロジェクト  
Email: cs-liaison-ml at hco.ntt.co.jp



Copyright © 2019 NTT. All Rights Reserved.

# 27 速いボールを捉える脳のしくみ

～視覚と身体の瞬間的な連携～

## どんな研究

野球等の球技では、高速で動くボールに対して適切に体を動かす必要がありますが、そのメカニズムは分かっていません。本研究では、錯覚を利用した基礎実験と、プロ野球打者の実打中の眼球運動及び身体運動の計測により、限られた時間の中で**速いボールを捉える脳の仕組み**を調べました。

## どこが凄い

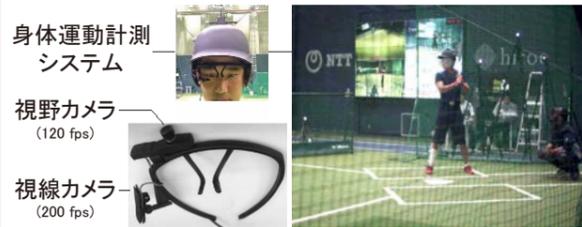
実戦に近い場面で眼球運動と身体運動を詳細に計測することで、**トップアスリートの洗練された技を捉えることに成功しました**。錯覚を利用した基礎実験では、**脳が視覚情報をもとに腕の動きを修正するメカニズム**を明らかにしました。

## めざす未来

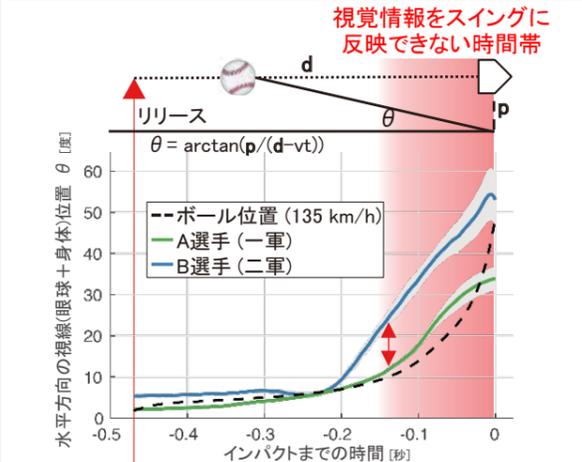
「見て、動くための潜在的な脳の働き」を解き明かし、**状況に応じて身体を最適に操るための「技」を鍛える新しいトレーニング法**の確立を目指します。これはトップアスリートに限らず、子どもからお年寄りまで幅広い人の運動技能向上に役立ちます。

### プロ野球選手はどうやって速いボールを捉えているのか？

**実験** 実戦に近い対戦場面で眼球運動及び身体運動をウェアラブルセンサによって計測



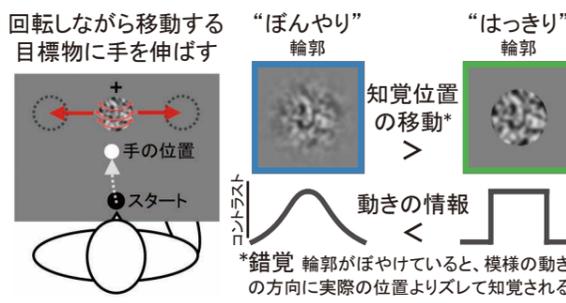
**結果** 良い打者は視覚情報をスイングに反映できない時間まで目と身体を連携させてボールを捉えている



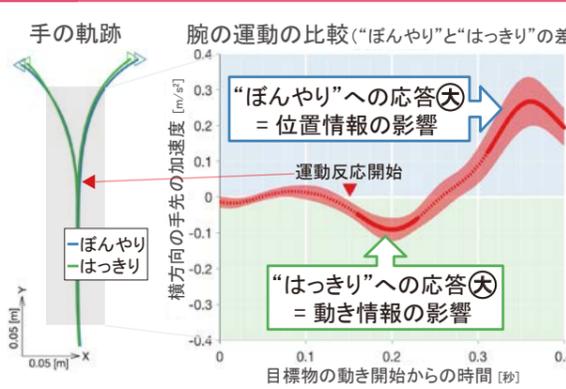
視覚情報だけでなく脳内の運動指令を利用している?!

### 限られた時間でなぜ速いボールが捉えられるのか？

**実験** 目標物の脳内情報のうち、「位置」と「動き」のどちらを基に腕を動かしているのかを検証



**結果** 腕の運動は、まず目標物の動きの情報によって駆動され、その後位置の情報によって調整される



ボールの位置を特定する前にバットの軌道を変えている?!

## 関連文献

[1] Y. Kishita, H. Ueda, M. Kashino, “Eye movements in real baseball batting by elite players,” in *Proc. The 48th Annual Meeting of the Society for Neuroscience*, 2018.  
[2] H. Ueda, N. Abekawa, S. Ito, H. Gomi, “Temporal development of an interaction effect between internal motion and contour signals of drifting target on reaching adjustment,” in *Proc. The 47th Annual Meeting of the Society for Neuroscience*, 2017.

## 連絡先

上田 大志 (Hiroshi Ueda) スポーツ脳科学プロジェクト  
Email: cs-liaison-ml at hco.ntt.co.jp



# 28 「見守る」と「見守られる」をつなぎます

～介護者記録用アプリ「みまもメイト」の記録共有による効果～

## どんな研究

うつ病の早期改善には、良好な家庭環境が重要です。そのためには、家族介護者がうつ病患者の言動に上手く対処する必要があります。本研究では、**家族介護者が介護記録を付けて介護者同士で情報共有することができるアプリ**を開発し、その有効性を確認しました。

## どこが凄い

情報分野におけるヘルスケアに関する研究の大半は患者の治療や介護を効果的に行う方法に焦点を当て、家庭内環境の改善を目指したものは殆ど存在しません。本研究は、家族介護者に働きかけることによって**家庭内環境の改善を実現**しました。

## めざす未来

単にハンディキャップを持つ人をサポートするだけでなく、ハンディキャップを持つ人が抱える問題をマジョリティと相互に理解し合い、マジョリティからも思いやりを引き出せるようなICTを開発することによって、**多様性を許容して活用する社会の実現**を目指します。

### 目的

家庭内介護を改善することにより、患者の早期回復を支援するICTツールを開発すること

### 開発

第1フェーズ  
介護記録機能を実装  
各家庭で家族介護者が記録を取る  
・患者の状況(気分・症状)  
・介護活動

第2フェーズ  
異なる家庭間の情報共有機能を追加  
・他の家族介護者の記録が閲覧可能になる

家庭内介護 (刺激/ストレス) → うつ病患者 ↔ 悪循環 ↔ 家族介護者 (結果/反応・言動)

**みまもメイト**

### 検証

目的: 第2フェーズの情報共有機能の影響を検証  
期待されたこと:  
- 他者の記録から**学び**を得る  
- 類似他者から**精神的な支え**を得る  
懸念されたこと:  
- 他の家族介護者への**ねたみ**  
- 患者と家族介護者の**衝突**

参加者(家族介護者):  
- 家族のうつ病患者を介護している成人14名  
- 女性11名、男性3名、平均年齢43歳  
- 専業主婦6名、フルタイム雇用6名、自営業2名

患者の病状:  
- 発症時期は平均8年前、8名が再発を経験  
- 全員が抗鬱剤を服用、定期的を受診

### 結果

第2フェーズの効果:  
他者にも理解しやすい文章の増加  
主語の補足や背景・状況の説明

期待されたこと: 他者からの学び、精神的な支えとなった。  
懸念されたこと: ねたみは一部報告されたが、衝突は報告されなかった。

**患者-介護者間のコミュニケーションが増加**

- 他者の記録を材料とした患者への話し掛け
- 患者から他の家族への間接的な情報発信
- 介護者から患者への間接的な情報発信

→ **間接的なコミュニケーション**を支援することの重要性が明らかになった。

入力文字数

第1フェーズ 第2フェーズ

試用日数

## 関連文献

[1] N. Yamashita, H. Kuzuoka, T. Kudo, K. Hirata, E. Aramaki, K. Hattori, “How Information Sharing about Care Recipients by Family Caregivers Impacts Family Communication,” in *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, 2018.  
[2] 山下直美, 葛岡英明, 平田圭二, 荒牧英治, 工藤喬, 服部一樹, “家族介護者同士の情報共有が家庭内コミュニケーションに及ぼす影響,” *インタラクシオン*, 2018.

## 連絡先

山下 直美 (Naomi Yamashita) 協創情報研究部 インタラクシオン対話研究グループ  
Email: cs-liaison-ml at hco.ntt.co.jp



# 29 身体が“見る”実世界の動きとは？

～自然統計で紐解く視覚運動制御の秘密～

## どんな研究

ヒトは体性感覚に加えて、視覚情報にも頼って運動制御を行っています。どのような情報を利用して視覚運動制御を行うのか、日常動作の一人称映像から得られる視覚的な動き情報や頭の動きの統計的性質からその秘密に迫ります。

## どこが凄い

低空間周波数情報のみを取り出した映像(低空間周波数映像)は輪郭がボケた映像となり、何が映っているかよく分かりません。しかし、ボケた映像の方が元映像よりも算出される自己運動の推定精度が高く、また、ヒトの無意識な視覚—運動制御情報処理と類似しているという結果が得られました。

## めざす未来

様々な実験やモデルシミュレーションを通して、ヒトが無意識に行っている視覚運動制御の仕組みを解明します。また、ヒトの仕組みを理解することで、実環境における視覚情報を使って巧みな動作ができるロボットの実現や映像加工技術への応用に役立つことが期待されます。

### MT細胞モデル[1]を用いた自己運動推定

・ヒトの視覚特性を再現する数理モデルを計算に利用



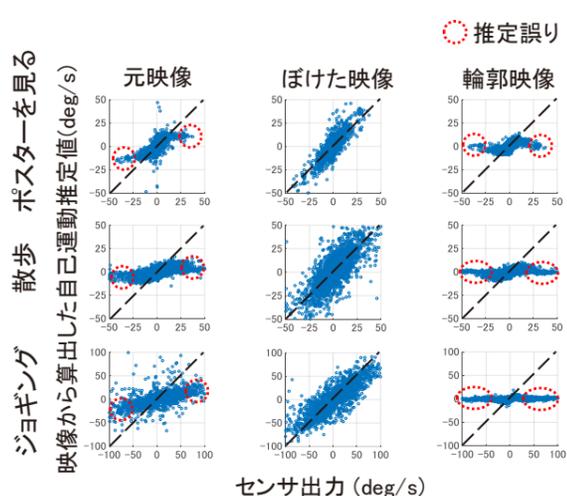
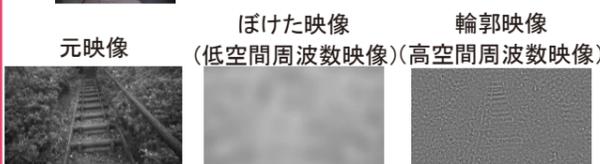
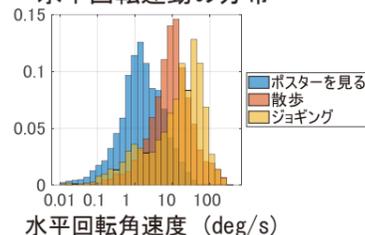
### 動きの特徴とモデルの自己運動推定精度

- ・動作によっては速い動きが多く含まれる
- ・ボケた映像を用いた方が自己運動推定精度が高い(散布図中列)
- ・速い動きを抽出するには低空間周波数情報が重要

ポスターを見る

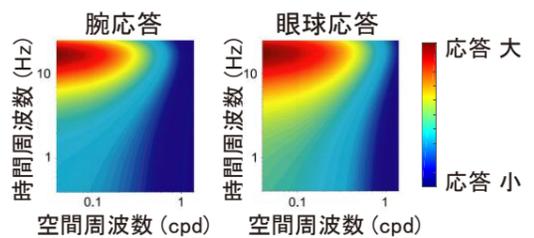


水平回転運動の分布



### 行動実験: 視覚運動で誘発される運動応答[2]

・ヒトも低空間周波数刺激に対して強い運動応答を示す(視覚刺激に対する無意識な運動応答に関して)



速い自己運動情報を抽出するために体はボケた映像を見ている

## 関連文献

[1] 中村大樹, 佐藤俊治, “計算論的に最適な速度推定器によってMT野細胞の複雑な反応特性を説明する,” 第27回日本神経回路学会全国大会, 2017.  
 [2] H. Gomi, N. Abekawa, S. Nishida, “Spatiotemporal tuning of rapid interactions between visual-motion analysis and reaching movement,” *The Journal of Neuroscience*, vol. 26, No. 20, pp. 5301–5308, 2006.  
 [3] D. Nakamura, H. Gomi, “Statistical analysis of optic flow induced by body motion characterizing OFR and MFR,” in *JNNS satellite meeting*, 2018.

## 連絡先

中村 大樹 (Daiki Nakamura) 人間情報研究部 感覚運動研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp



# 30 座っているけど歩行感

～疑似歩行感が身体近傍空間を拡張する～

## どんな研究

視覚や聴覚に加えて触覚や身体感覚の刺激を与えることにより、座っている体験者にあたかも歩いているような感覚を生成させます。さらにこのとき、われわれの脳内表現である自己の身体を取り囲む「身体近傍空間」が前方に拡張していることを明らかにしました。

## どこが凄い

バーチャリアリティ(VR)空間内などで歩く感覚を作るためには体験者を実際に歩かせる必要があります。本技術は実際に歩くことなく歩行したような感覚を生み出すことを目指しています。また、歩行が困難な利用者でも体験することができます。

## めざす未来

本技術を用いることで、VR空間などにおいて多様な移動表現を実現できるようになります。今後はVR酔いとの関連を明らかにし、より臨場感の高い旅行体験や一緒に歩きまわるといった体験など、より幅広い場面で活用できる技術の実現を目指します。

### 多感覚刺激による疑似歩行感の生成

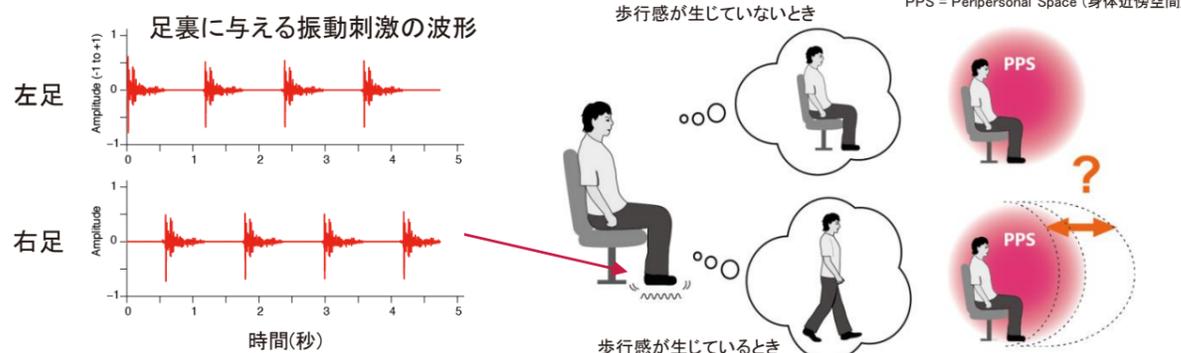
これまでに「着座状態での歩行感覚の生成」を目指し、上肢や下肢を外から動かす機構、身体動揺を実現する椅子、風・匂いの提示といった五感情報を組み合わせるとして疑似的な歩行感を生成する手法を開発してきました。

本研究では足裏に着目し、振動刺激を提示することで歩行感を生成する手法について開発を進めています。

### 疑似歩行感覚による身体近傍空間の拡張

足裏に与える振動刺激に応じて身体に接近する刺激(たとえば体に向かってくる物体)に対する反応時間が変化することを明らかにしました。

→この結果は、本手法により「身体近傍空間」が実際に歩いているかのように広がることを意味します。



## 関連文献

[1] 雨宮智浩, 池井寧, 広田光一, 北崎亮晃, “歩行を模した足底振動刺激による身体近傍空間の拡張,” *日本バーチャリアリティ学会論文誌*, Vol. 21, No. 4, pp. 627–633, 2016.  
 [2] 雨宮智浩, “触覚・身体感覚の錯覚を活用した感覚運動情報の提示技術,” *基礎心理学研究*, Vol. 36, No.1, pp.135–141, 2017.  
 [3] Tomohiro Amemiya, “Haptic Interface Technologies Using Perceptual Illusions,” in *Proc. 20th International Conference on Human-Computer Interaction (HCI International 2018)*, pp.168–174, 2018.

## 連絡先

雨宮 智浩 (Tomohiro Amemiya) 人間情報研究部 感覚運動研究グループ  
 Email: cs-liaison-ml at hco.ntt.co.jp

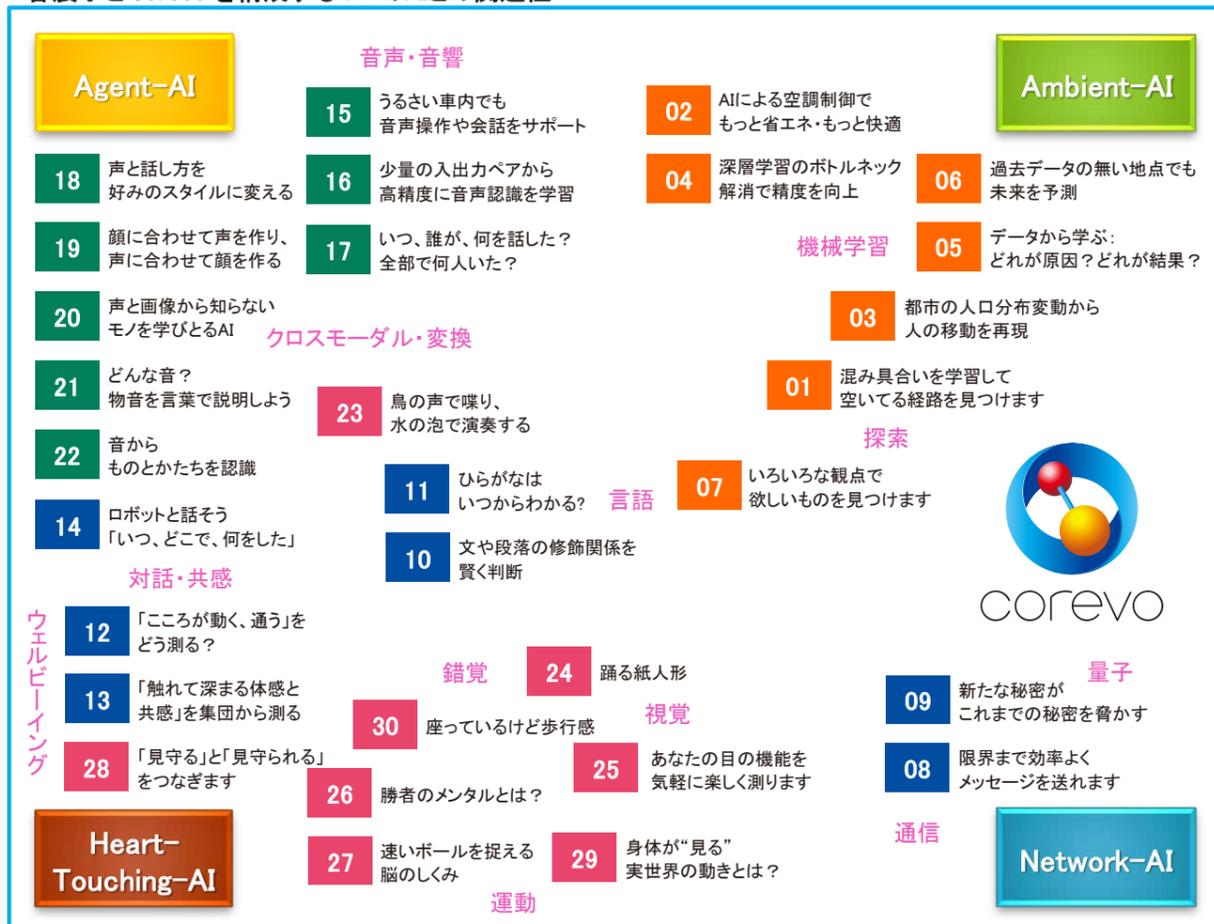




**corevo®**  
～NTTグループのAI技術～

- Agent-AI** 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-AI** 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-AI** 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-AI** 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



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™(コロボ)」,” *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/>

**連絡先**

納谷 太 (Futoshi Naya) 企画担当  
Email: cs-liaison-ml at hco.ntt.co.jp



Innovative R&D by NTT  
オープンハウス 2019  
Copyright © 2019 NTT. All Rights Reserved.

**NTT コミュニケーション科学基礎研究所  
オープンハウス 2019 実行委員会**

(委員長) 小川 厚徳  
(副委員長) 武 小萌  
(委員) 西野 正彬 / Mathieu Blondel / 持田 岳美