



NTT コミュニケーション科学基礎研究所

オープンハウス 2017

6/1 木

12:00~17:30

6/2 金

9:30~16:00

入場無料・事前登録不要

会場：NTT京阪奈ビル

京都府相楽郡精華町光台2-4 (けいはんな学研都市)



「オープンハウス2017」へようこそ

NTT コミュニケーション科学基礎研究所

所長 前田 英作



NTT コミュニケーション科学基礎研究所「オープンハウス2017」へご来場頂き、心より御礼申し上げます。

情報分野の基礎研究を担う組織としてNTT コミュニケーション科学基礎研究所が設立されてから、26年が経ちました。この間に多くの研究成果が生まれ、少しずつではありますが実社会で活躍している技術も増えています。しかしながら、夜が明け朝日が昇る頃には、新しい次の地を探し求め彷徨していくことも基礎研究の使命であります。

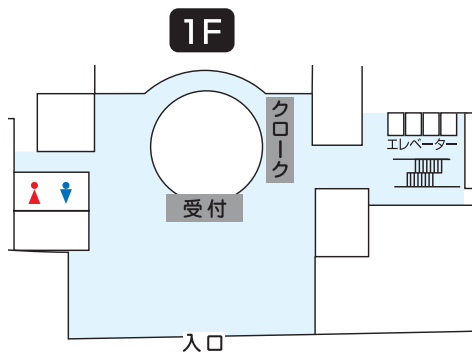
こうした視点に立ち、私たちコミュニケーション科学基礎研究所では、2種類の研究に取り組んでいます。一つは、時代を創る研究です。5年後、10年後に実となる技術を育て、知見を積み上げていくことにより、新たな時代への準備を世界に先駆けて行います。もう一つは、時代の要請に応える研究です。これまで培ってきた技術と知見を利用して、私たちが今まさに直面している社会課題にも取り組んでいます。

NTTグループでは、人工知能(AI)のブランドcorevo®を立ち上げました。当研究所の成果はcorevo®を支える基盤となります。オープンハウス2017では、「データと学習の科学」、「計算と言語の科学」、「メディアの科学」、「コミュニケーションと人間の科学」の4領域について、最新の研究成果を幅広くご紹介いたします。二日間のささやかな企画ではありますが、ご来場の皆様にとって有益で楽しいひとときとなれば幸いに存じます。



目 次

ごあいさつ	01
会場案内図／講演スケジュール	03
所長講演	
AI×IoT×BigData時代における基礎研究 科学と工学の融合による研究プロセスの新しいデザイン	04
招待講演	
あしたのまちの100の風景 生活者発想によるシナリオプランニング	06
研究講演	
ビッグデータに隠れた構造を見つけ出す 確率的潜在変数モデルに基づく知識発見	08
あなた専用のお手本映像で上達支援 深層学習によるメディア生成の可能性	10
打てるバッターは何が違うのか? 潜在脳機能からみた技の神髄	12
感じ方まで変わる? 息と音の不思議な関係 呼吸と音の相互関係とその心理的效果	14
研究展示	17
データと学習の科学	
01 車両によるスマートシティセンシング 車載型センサノードによる時空間都市イベント検知	18
02 都市における人の活動パターンとリズムを可視化 時空間統計に基づく都市データの解析	19
03 集団最適なナビをリアルタイムに実現 ベイズ的最適化に基づく効率的な誘導計画探索	20
04 プライバシーを保護しつつ人の動線を抽出 人流推定のための確率的行動モデル	21
05 時間変化するデータののための安定した深層学習 Gated recurrent unitにおける学習安定化技術	22
06 グラフを用いてデータの間接関係を素早く発見 L_1 -グラフ構成法のための高速化アルゴリズム	23
計算と言語の科学	
07 絶対に安全な共有鍵を作れるか 今あるハードウェアで量子暗号の実現を図る	24
08 漏れなくデタラメです 物理乱数源のランダムさを損なわずに乱数ビットを安定供給	25
09 敷き詰めかたをすべて見つけます 厳密被覆問題の高速な全解列挙法	26
10 絵本はこどもの言葉の教科書 絵本コーパスと幼児の語彙獲得の関係を探る	27
11 ノンネイティブのためのリスニング支援 効率的な自動字幕表示法	28
12 あなた好みに翻訳する 接頭辞制約によるニューラル翻訳の制御	29
13 単語の意味や関係を賢く学習 開発者利便性が高い単語埋め込みベクトルの獲得法	30
メディアの科学	
14 照明光で目的の色だけをくっきり見せる スペクトル最適化による色彩強調	31
15 2Dと3D、どちらも同時に楽しめます メガネなしでも二重に見えないステレオ映像生成法	32
16 あなた専用のお手本映像で上達支援 深層学習によるメディア生成の可能性	33
17 英語の発音をネイティブのように綺麗に変換 声道モデルと深層生成モデルを用いた統計的音声変換	34
18 現場の生音、高音質電話でクリアにお届け リアルタイム音源伝送向け低遅延・高音質符号化CLEAR	35
19 世界中の言葉を聞き分けます 多言語音声識別・認識技術	36
20 あなたの声に即座にあわせて音声認識 音声認識精度を高めるニューラルネットの高速適応技術	37
21 会話の流れを読んでみんなの声をうまく認識 頑健な音声処理を可能にする複数人発話区間推定技術	38
コミュニケーションと人間の科学	
22 それっぽくしゃべるとは 発話リズム変換技術を用いた脳における音声情報処理の解明	39
23 「自然な」音を調べて聴覚を理解する バイノーラル録音や空間残響に着目した聴覚再考	40
24 無意識の目の動きからココロを読む 微小眼球運動から潜在的な心を解読する技術	41
25 触り方と触り心地の深い関係 手の動きから触り心地を推定する	42
26 災害時の安否確認通話をつながりやすく 通話時間の自発的短縮を促す発信規制法	43
27 脳科学とICTでアスリート脳を解明 Body-mind readingで推進するスポーツ脳科学	44
28 私を「ぶるなび」で連れてって 屋内・屋外・仮想空間での牽引感覚デバイス応用に向けて	45
29 AIと議論:都市と田舎、永住するならどっち? 議論構造に基づき論拠提示や反論を行う議論対話システム	46
corevo NTTグループのAI技術	48



- データと学習の科学
- 計算と言語の科学
- メディアの科学
- コミュニケーションと人間の科学
- 通行可能領域
- 講演会場
- インフォメーション
- トイレ

講演スケジュール

6月1日(木)

所長講演

13:20~13:50

AI×IoT×BigData時代における基礎研究

科学と工学の融合による研究プロセスの新しいデザイン

コミュニケーション科学基礎研究所 所長 前田 英作

招待講演

14:00~15:00

あしたのまちの100の風景

生活者発想によるシナリオプランニング

博報堂生活総合研究所 上席研究員 酒井 崇匡

研究講演

15:30~16:10

ビッグデータに隠れた構造を見つけ出す

確率的潜在変数モデルに基づく知識発見

上田特別研究室 岩田 具治

6月2日(金)

研究講演

11:00~11:40

あなた専用のお手本映像で上達支援

深層学習によるメディア生成の可能性

メディア情報研究部 金子 卓弘

研究講演

13:00~13:40

打てるバッターは何が違うのか?

潜在脳機能からみた技の神髄

スポーツ脳科学プロジェクト 柏野 牧夫

研究講演

13:50~14:30

感じ方まで変わる? 息と音の不思議な関係

呼吸と音の相互関係とその心理的効果

守谷特別研究室 佐藤 尚



AI×IoT×BigData時代における基礎研究

～科学と工学の融合による研究プロセスの新しいデザイン～

Basic Research in the Age of AI×IoT×BigData

～ New Design of R&D Processes by the Blending of Science and Engineering ～

NTT コミュニケーション科学基礎研究所 所長

前田 英作 Eisaku Maeda

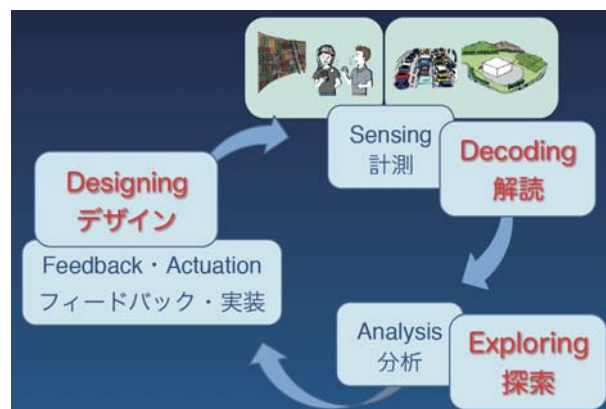
AI×IoT×BigDataが生みだす未来

AI(人工知能)として総称される新しい情報技術、IoT(Internet of Things)という新しい情報インフラ、BigDataと呼ばれるこれまでにない量と質を有するデータ群、この三つが重なることにより、世界が大きく変わろうとしています。しかしながら、これからどのような技術とサービスが現れるのか、今私たちが予見できているのは恐らくそのごく一部に過ぎません。その可能性と社会に与える影響について、さらにアイデアと知恵を絞ることが求められています。AI×IoT×BDが生みだす世界の技術基盤は、情報を“正確に”複製することを可能にした「情報のデジタル化」と、情報を“容易に”集積・拡散することを可能にした「世界のネットワーク化」の二つにあります。この基盤の上に、機械学習をはじめとした情報を処理するアルゴリズムが載ることにより、新しい未来が生まれようとしています。



解読・探索・デザイン

実世界にある森羅万象(人・もの・環境)のあらゆるものからデータが取得され、解読、探索、デザインという情報循環の過程を経て、実社会にフィードバックされます。人が書き記した文字、行動、発話、画像、音声、生体反応、筋活動、脳活動など、人に関わるあらゆるデータから知能のモデル化が行われます。その過程で、知能が様々な形で分解、再構成されて知能の部品化が進み、市場経済の中で商品として扱われる時代が早晩到来するに違いありません。



今まで純粋科学と工学は独立な学問として研究が行われてきました。しかし、先に述べた変化の中でこの構図も大きく変わってきています。センシングの低廉化と知能化により、様々な分野の基礎科学データが短時間で大量に取得できるようになり、処理・解析の信頼度が増すことにより工学への活用が容易にな



りました。一方、信号処理・データ解析などの工学領域では、生データを読み取る力や実験計画のデザインなど実験科学のスキルが必要になってきています。前者は科学の工学化、後者は工学の科学化と呼べるかも知れません。科学から工学、工学から社会実装、さらにそこから科学へのフィードバックに至るループが、太くかつ短時間で回るようになりました。そのため、牧歌的に行われていた一部の科学研究でさえ全体の経済性を視野にいた研究戦略の立案が必要になりつつあります。

音やカメラの画像を見るのではなく、解読された様々な情報をここで見ることができ、さらにその情報を操作、制御することも可能になります。今すでに、2~3歳の子どもたちがタブレットをいじってYouTubeなどを見るようになっていきます。それと同じように、ごく普通の市井の人がタブレットを使って実世界を“操作”できる時代が来るかも知れません。その良し悪しはおくとして、情報の民主化とはそういうことも指し示しています。

実世界を操作するのは誰か

ブームとしてのAIはいずれ沈静化するでしょう。しかし、技術開発そのものは緩やかにではあるが確実に進みます。その究極の将来像はどのようなもののでしょうか。人工知能の父と呼ばれるマービン・ミンスキー (Marvin Minsky) がかつて知能の秘密を知りたくて脳の中をのぞこうとしたように、専門家でない普通の人が世界の情報を手軽に見ることができるようになるでしょう。



関連文献

- [1] 前田英作, “通信からコミュニケーションへ データの時代におけるパラダイムの変容,” *NTT技術ジャーナル*, Vol. 28, No. 9, pp. 6-8, 2016. (【特集】コミュニケーションの未来を描く基礎研究, <http://www.ntt.co.jp/journal/1609/special.html>)
- [2] 前田英作, “情報科学技術を抱きしめて 世界の解読、探索とデザイン,” *NTT技術ジャーナル*, Vol. 27, No. 9, pp. 8-9, 2015. (【特集】未来への羅針盤としてのコミュニケーション技術, <http://www.ntt.co.jp/journal/1509/special.html>)
- [3] 前田英作, “基礎研究は「時代」とともに在り,” *NTT技術ジャーナル*, Vol. 26, No. 9, pp. 12-15, 2014. (【特集】コミュニケーション科学の新展開, <http://www.ntt.co.jp/journal/1409/special.html>)
- [4] 前田英作, “音声・言語・聴覚の最先端研究が拓くコミュニケーションの未来,” *NTT技術ジャーナル*, Vol. 25, No. 9, pp. 12-13, 2013. (【特集】ここまで伝わるコミュニケーションを支える音声言語と聴覚研究の最前線, <http://www.ntt.co.jp/journal/1309/special.html>)



あしたのまちの100の風景

生活者発想によるシナリオプランニング

Four scenarios of future city



博報堂生活総合研究所 上席研究員
Hakuhodo Institute of Life and Living R&D Director

酒 井 崇 匡 Takamasa Sakai

プ ロ フ ィ ー ル

2005年 早稲田大学政治経済学部経済学科卒業

2005年 株式会社博報堂 第7プランニング局 ストラテジック・プランナー

2008年 株式会社博報堂 教育コミュニケーション推進室参画

2010年 株式会社博報堂 第二カンパニー マーケティング・プランナー

2012年 博報堂生活総合研究所 主任研究員

2016年 博報堂生活総合研究所 上席研究員

2005年よりマーケティングプランナーとして、教育、通信、外食、自動車、エンターテインメントなど諸分野でのブランディング、商品開発、コミュニケーションプランニングに従事。

2012年より博報堂生活総合研究所に所属し、日本およびアジア圏における生活者のライフスタイル、価値観変化の研究に従事。

専門分野はバイタルデータや遺伝情報など生体情報の可視化が生活者に与える変化の研究。

著書に『自分のデータは自分で使う マイビッグデータの衝撃』（星海社新書）。

MEMO



ビッグデータに隠れた構造を見つけ出す

確率的潜在変数モデルに基づく知識発見

Discovering hidden structure in big data

Knowledge discovery based on probabilistic latent variable models



上田特別研究室

岩田 具治 Tomoharu Iwata

プロフィール

2001年慶應義塾大学環境情報学部卒業。2003年東京大学大学院総合文化研究科修士課程修了。同年日本電信電話株式会社入社。2008年京都大学大学院情報学研究科博士課程修了 博士(情報学)。2012年から2013年にかけてケンブリッジ大学客員研究員。現在、NTT コミュニケーション科学基礎研究所 主任研究員(特別研究員)。データマイニング、機械学習の研究に従事。2014年長尾真記念特別賞等受賞。情報処理学会、電子情報通信学会各会員。

インターネットやセンサの普及とともに膨大なデータを容易に入手、蓄積できるようになってきています。しかしながら、得られたデータをそのまま眺めたとしても、どのような性質をもつのか、どのような価値があるのかを知ることは困難です。データから有用な知識を自動的に発見することはビッグデータ解析の大きな課題となっています。

様々な場面で得られたデータを見てみると、隠れた構造を持っていることが多々あります。例えば、新聞記事では同じ記事中に同時に現れやすい単語の集まりがあります。政治記事では「首相」と「国会」という単語が同時に現れやすいですし、スポーツ記事では「サッカー」と「ゴール」が同時に現れやすいです。また、画像の場合、包丁が写っていたら、まな板も写っている可能性が高いです。ソーシャルネットワークでは友達の友達は友達であることが多いですし、SF小説を買った人は次もSF小説を買う可能性が高いです。

このような隠れた構造は有用な知識となります。例えば、たくさんの研究論文から、同時に現れやすい代表的な単語のまとまりをトピックとして抽出できれば、どのような研究が話題となっているかを知ることができます。またそのトピックを利用することで、単純なキーワード検索では得られない関連論文を探し出すことができます。購買データから同じような商品を購入している顧客グループを見つけ出せれば、グループ毎に特化したマーケティング戦略を練ることができます。

隠れた構造を自動的に見つける方法として、潜在変数モデルが注目されています。潜在変数モデルでは、観測データが生成される過程において隠れた構造が重要な役割を果たしていると考えます。具体的には、隠れた構造を観測することができない変数である「潜在変数」として表します。そして、その潜在変数に依存してデータが生成されるモデルを考えます。例えば、新聞記事が生成されるモデルの場合、まず新聞記事の潜在変数であるトピックが決められ(政治トピック)、そのトピックに従ってその記事の単語が生成される(首相、国会、政策)、と考えます。



潜在変数モデルを使うことによって潜在変数からデータが生成される過程を表現できますが、実際に与えられるのはデータなので、データから潜在変数を推定する逆問題を解く必要があります。この逆問題を解くためにベイズ則という確率の基本的な公式が利用できます。ベイズ則を利用することにより、潜在変数からデータを生成するモデルをもとに、与えられたデータに関する潜在変数を推定できます。ここで確率を基盤にすることによって、不確実性を扱うことが可能になり、データに含まれるノイズにも頑健な推定が可能になります。

機械学習は「教師あり学習」と「教師なし学習」に大別できます。確率的潜在変数モデルによる隠れた構造の推定は「教師なし学習」になります。教師あり学習では、入力と出力のペアを大量に教師データとして用意して、ある入力に対応する出力を予測できるように学習します。学習データ作成のコストが高い場合や、プライバシー保護等の理由で学習データが得られない場合は、教師あり学習は利用できません。一方、教師なし学習は、教師データなしで学習できるという利点があります。その反面、タスクに適したモデルを使わなければ、望ましい有用な結果を得ることができないため、モデルの設計が重要となります。

本講演では、まず上記の隠れた構造を発見するための確率的潜在変数モデルに基づく技術の基礎を説明します。そして、具体的なタスクにおけるモデルを紹介し、どのような場面でどのようなモデルを用いればよいかの指針を示します。本技術は様々な課題に適用できます。例えば、大量の文書から話題を見つけ出す「トピック抽出」[1]、異なる店の顧客グループを対応付けクロスドメイン推薦を可能にする「オブジェクトマッチング」[2]、超高次元データをなるべく元データに含まれる本質的な構造を保ったまま2次元空間に埋め込む「可視化」[3]、ユーザのインターネット上での活動履歴から影響力の強いユーザやユーザ間の影響関係を推定する「影響ネットワーク推定」[4]に適用できます。

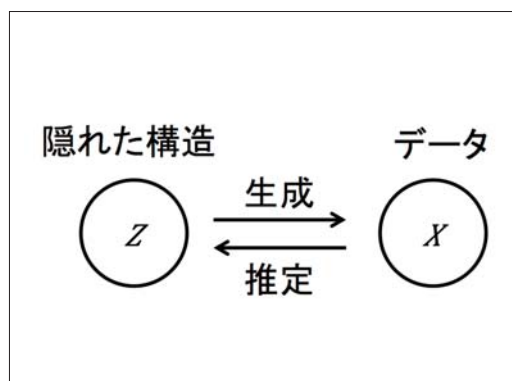


図1: 隠れた構造に依存したデータの生成と、データに基づく隠れた構造の推定

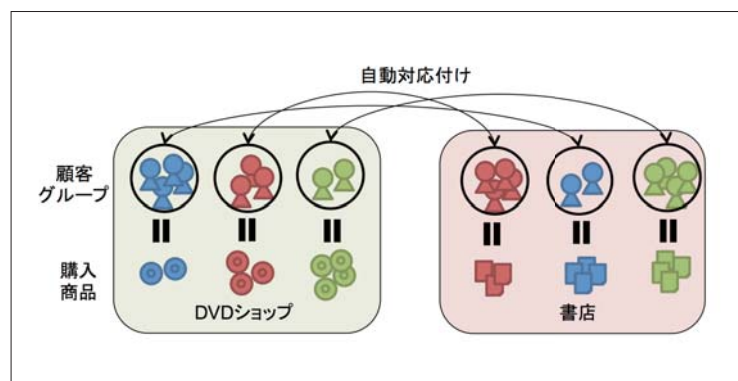


図2: 異なる店の顧客グループを対応付けるオブジェクトマッチング

関連文献

- [1] T. Iwata, T. Yamada, Y. Sakurai, N. Ueda, "Online multiscale dynamic topic models," in *Proc. 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 664-672, 2010.
- [2] T. Iwata, J. Lloyd, Z. Ghahramani, "Unsupervised many-to-many object matching for relational data," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 38, No. 3, pp. 607-619, 2016.
- [3] T. Iwata, T. Yamada, N. Ueda, "Probabilistic latent semantic visualization: Topic model for visualizing documents," in *Proc. 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 363-371, 2008.
- [4] T. Iwata, A. Shah, Z. Ghahramani, "Discovering latent influence in online social activities via shared cascade Poisson processes," in *Proc. 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 266-274, 2013.



あなた専用のお手本映像で上達支援

深層学習によるメディア生成の可能性

Generative personal assistance with audio and visual examples

Deep learning opens the way to innovative media generation



メディア情報研究部

金子 卓弘 Takuhiro Kaneko

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。
2012年東京大学工学部卒業。2014年同大学大学院情報理工学系研究
科修士課程修了、2017年博士課程入学。2014年日本電信電話株式会
社入社。コンピュータビジョン、画像処理、音声信号処理、機械学習の研
究に従事。2012年日本機械学会畠山賞受賞。同年ICPR2012 Best
Student Paper Award受賞。電子情報通信学会会員。

深層学習の広がり

近年、深層学習と呼ばれる技術の著しい発展により、少し前までは想像もつかなかったことが実現されつつあります。メディア生成はその代表例であり、例えば、任意の乱数から本物と見分けのつかないような画像を生成する技術など、世界中で驚くような研究成果が報告され始めています。同じ深層学習でも、数年前までは、画像識別や音声認識など比較的正解が明確な問題に対して精度を向上させる研究が主でしたが、近年は、深層学習ならではの表現能力の高さを活かし、新たな概念・モデル構造を導入することで、より複雑な問題も解決できるようになってきています。

上達支援のためのメディア生成

「ボールを速く投げたい」「英語をきれいに発音したい」「印象の良い表情にしたい」など「何かをしたい」という時に、「どうすれば良いのかやり方が分からない…」といった経験は誰しもがしたことがあるのではないのでしょうか。このような時には、詳しい人に教えを乞うたり、自身で書物やインターネットを頼りに情報を探

したりといった方法がよくとられます。前者は個人に合った具体的な指示が得られる点、後者は自己解決できる点で有用ですが、一方で、前者は詳しい人を探し出すことが難しく、後者は特定の個人を対象にしたものではないため自身のケースにうまく当てはまる情報を見つけ出すことが難しいという問題があります。このような問題に対し、私たちは、前者の利点である「個人性」「具体性」と後者の利点である「自己解決性」の両者を満たすものとして、「あなた専用のお手本映像の生成」システムの実現を目指しています。

あなた専用のお手本映像の生成

図1に提案システムのコンセプトを示します。このシステムでは、ユーザが与えたメディア情報(画像や音声)を元に解析を行うことで「個人性」を実現し、具体的な指示や理想像を提示することで「具体性」を実現し、さらにこれらの処理を自動的に行うことで「自己解決性」を実現します。図2に、実際に私たちのグループが考案した表情改善のためのフィードバックシステム[1]



の動作例を示します。本システムでは、まず、ユーザの顔表情をカメラで撮影して現状把握を行います。次に、この情報を元に解析を行うことで各ユーザの現状に最適な理想状態を求め、どの顔パーツをどのくらい修正すればよいかということを示印で提示します。具体的な情報が与えられているので、ユーザはこの指示に従って顔パーツを動かすことで理想状態に近づくことができます。同じような方式は、上手な話し方の練習など、音声に対しても適用することができます。

多様な入力状態への対応と具体的な情報提示

上記のようなシステムを実現するためには、二つの点で技術的な難しさがありました。一つ目は多様な入力状態への対応です。このシステムが対象としている人は千差万別であり、最適なフィードバックも個人ごとに異なります。従来のフィードバックシステムはルールベースで入出力関係を定めるものが主で、多様な入力状態に対応するためには数多くのルールを手で作成する必要がありました。二つ目の難しさは、入力に対して単に認識するだけでなく、それを元に具体的な情報提示を行う点です。従来の画像認識技術では、クラス（例えば表情）の識別や特徴的な点（例えば顔パーツ）の検出などを行うことはできましたが、提示すべき情報の生成は困難でした。私たちは、この問

題を解決するために、深層学習における新たな情報伝播方法を考案し、多様な入力状態の理解と具体的な情報提示とを一体的なモデルで実現することを可能にしました。

深層学習によるメディア生成の今後の展望

上記のアプローチの鍵は、システムがメディア生成を通してユーザに寄り添えるか、ということです。そのための必須技術として、私たちは、インタラクティブに操作しながら理想像を探索・提示する方法[2]、肉声と見分けのつかない音声の合成方法[3]なども考案してきています。これらは、上達支援だけでなく、ユーザが思い描くようなメディアを生成するために無くてはならない技術です。私たちは、どのようなものが生成できたら嬉しいかという想像力と、深層学習を用いたメディア生成に関する知見と経験を積み重ねながら、将来的には、ユーザのあらゆる要望に応えられる、極めて高品質なメディア情報を生成できる技術の確立を目指しています。

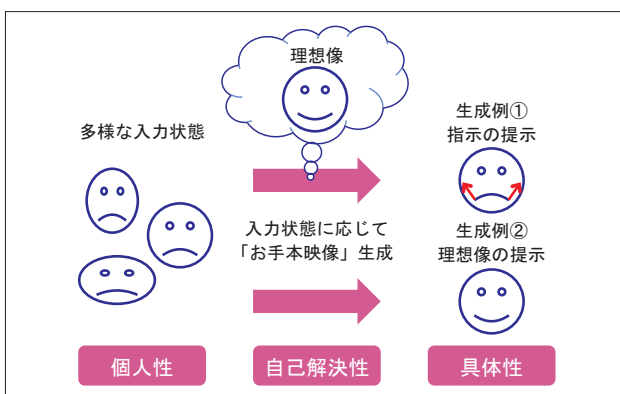


図1:「あなた専用のお手本映像の生成」システム
(例:顔の表情改善の場合)

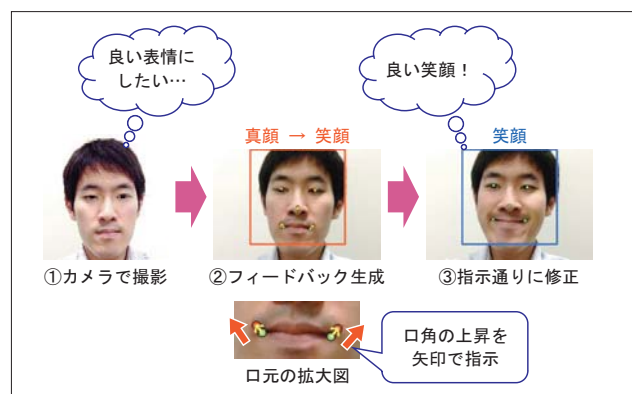


図2:表情改善のためのフィードバックシステムの動作例

関連文献

- [1] T. Kaneko, K. Hiramatsu, K. Kashino, "Adaptive visual feedback generation for facial expression improvement with multi-task deep neural networks," in *Proc. 24th ACM International Conference on Multimedia (ACMMM)*, pp. 327-331, 2016.
- [2] T. Kaneko, K. Hiramatsu, K. Kashino, "Generative attribute controller with conditional filtered generative adversarial networks," in *Proc. 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017 (to appear).
- [3] T. Kaneko, H. Kameoka, N. Hojo, Y. Ijima, K. Hiramatsu, K. Kashino, "Generative adversarial network-based postfilter for statistical parametric speech synthesis," in *Proc. 42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4910-4914, 2017.



打てるバッターは何が違うのか？

潜在脳機能からみた技の神髄

What is special about excellent batters?

The essence of fine skills in the light of implicit brain functions



スポーツ脳科学プロジェクト

柏野 牧夫 Makio Kashino

プロフィール

NTT コミュニケーション科学基礎研究所 上席特別研究員・スポーツ脳科学プロジェクトPM。東京工業大学工学院情報通信系特任教授。東京大学大学院修士課程修了、博士(心理学)。専門は心理物理学・認知神経科学。スポーツ脳科学の他に、聴覚のメカニズム、自閉スペクトラム症の感覚特性、body-mind readingなどの研究に従事。日本音響学会理事。文部科学大臣表彰科学技術賞(研究部門) 受賞。著書に「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。

高校通算本塁打60本という抜群の実績をひっさげて、期待の大型新人としてプロ球界に入ったA選手。恵まれた体格から繰り出すスイングスピードは既に一軍トップクラス。しかし試合になるとさっぱり打てず、二軍でくすぶる毎日。よくある話です。レベルが高くなればなるほど、いわゆる身体能力だけでは通用しなくなります。逆に、体格や筋力では見劣りしても、素晴らしい活躍を見せる選手もたくさんいます。では、何が両者を分けるのでしょうか。我々スポーツ脳科学プロジェクトでは、脳にその鍵があると考えています[1]。

野球のバッティングというのは、脳の情報処理という観点からみると、きわめて高度な課題です。ハイレベルの野球であれば、投手の手からボールが離れてからホームベースを通過するまで、速球ならば約0.4秒、遅い変化球なら約0.6秒。このわずかな時間内に、様々な球速や軌道で投げられるボールに対して、ホームベース上での到達点を予測し、打つ・打たないという意思決定をし、強くかつ正確にバットスイングをしなければなりません。しかも、当たり所が悪ければ命にも関わるかもしれないデッ

ドボールの恐怖や、チャンスに凡退して大ブーイングを浴びる不安などの強烈な精神的プレッシャーにもさらされています。まさに、知覚、意思決定、運動制御、情動といった脳科学の主要問題が、従来の単純化された実験室環境とはかけ離れた形で複雑に絡み合っているのです。

まず、バッティングでは、目でボールの軌道を正確に捉えることが不可欠です。バッティング以前に、そもそもモノを見るというのは思いの外複雑な情報処理です。眼球の網膜では中心部と周辺部で特性が違って、形の細部は中心のごく狭い部分でしか捉えられませんが、動きに関しては周辺部の方がむしろ鋭敏です。日常生活では、視野全体で形も動きもはっきり見えているように感じていますが、実はこれはある種の錯覚です。ある時点はこの位置、次の時点はこの位置というように眼球を素早く動かしてサンプリングした、時間的にも空間的にも断片的な情報を脳が統合して、奥行きも含めた外界の全体像を再構成したものが「見えて」いるのです。その過程で、外界の物理的特性と知覚される内容とが多かれ少なかれ乖離することは避けられません。



バッティングのようにきわめて時間的余裕のない状況では、上記の情報処理は一層困難になります。ボールの物理的軌道自体は、球速や回転数、回転軸などによって決まりますが、打者の感じるボールの変化や「キレ」というものは、物理的軌道だけでは決まりません。打者がどの時点でどこに眼球を向けてどのような情報をサンプリングしているか、それを脳がどのように統合しているかに強く依存します。スピードガンの数字は平凡なのにプロで活躍できる投手がいますが、彼らは、フォームや配球で、打者の眼や身体の動きをうまくコントロールし、脳がボールの軌道を正しく計算するのを困難にしていると考えられます。逆に、打てる打者は、その手に乗らず、ボールの軌道を正確に捉える戦略を身につけている、あるいは処理能力を備えていると言えます。

次に、ボールの軌道を正確に捉えたとしても、打つべきでないボールに手を出してはいけません。バットスイングに要する時間を考慮すると、打つか打たないかという意思決定に許される時間は高々0.2秒ほどしかありません。意思決定といっても、意識が生成されるのにはもっと時間がかかりますから、無自覚的なものということになります。身体が勝手に反応するという感覚に近いでしょうし、逆に、こうしようと思っても、実際その通りには身体は動きません。強い情動にさらされている場合にはなおさらです。打てる打者は、振り遅れないぎりぎりまでボールの軌道に関する情報を集め、情動に攪乱されることもなく、適切な意思決定ができていると考えられます。

こうしてみると、バットを振る以前に、打者の優劣を決める要因はいろいろあることがわかります。冒頭のA選手に欠けているのは、まさにそのような部分なのではないでしょうか。スポーツ脳科学プロジェクトでは、大学野球や社会人野球、女子ソフトボールなどの選手の協力を得て、投手と打者の脳や身体で何が起きているか、優れた選手とそうでない選手はどこが違うのかを詳しく調べています。その際、従来型の実験室での脳機能計測に加え、計測環境を整えたスマートブルペン、ヴァーチャルリアリティ環境、さらには実際の試合と、目的に応じた多様な環境で、身体動作、眼球運動、心拍、筋活動などの生体計測を行い、それらの背後にある脳機能を、我々の培ってきたbody-mind reading技術を駆使して解読する研究を行っています(図1)。その結果、打者の優劣に応じて、いくつかの興味深い差違が浮かび上がってきました。講演ではその一端を紹介します。

こうして得られた知見は、野球のほかにも、例えばテニスやバドミントン、さらには格闘技など、瞬間的なインタラクションが勝敗を分ける様々な競技にも展開可能です。また、長期的に計測すれば、選手の能力発達、才能の早期発見など、選手の育成に役立つ知見も得られます。これらの知見をもとに、アスリートの脳を鍛えパフォーマンス向上を支援する新しいトレーニング手法の確立をめざします。

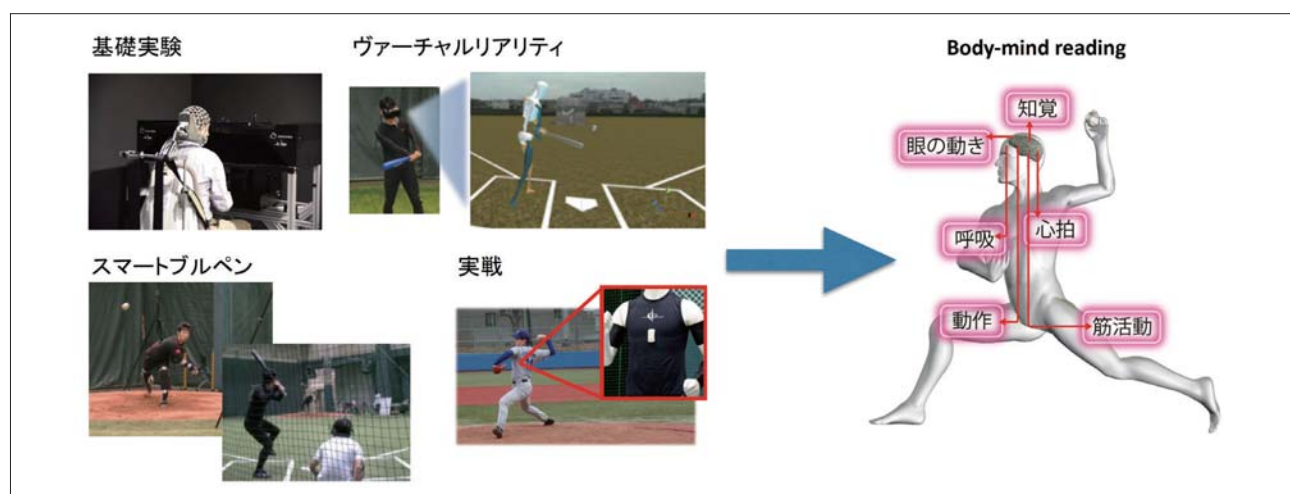


図1:スポーツ脳科学プロジェクトの研究アプローチ。様々な環境で計測されたアスリートの生体情報から、body-mind reading 技術を駆使して脳機能を解読、優れたパフォーマンスのメカニズムを探る

関連文献

- [1] スポーツ脳科学プロジェクトウェブサイト:<http://sports-brain.ilab.ntt.co.jp/index.html>



感じ方まで変わる? 息と音の不思議な関係

呼吸と音の相互関係とその心理的効果

Does breathing change the impression of a sound?

Interactive relationship between respiration and sound, and its effect on emotion



守谷特別研究室

佐藤 尚 Takashi G. Sato

プロフィール

NTT コミュニケーション科学基礎研究所 守谷特別研究室 研究主任。
東京大学博士(情報理工学)。神経工学、生体生理工学を専門。近年は
特に「音と呼吸の関係の解明と応用」、「触覚による呼吸誘導」、「可視光
通信による集団の状態計測」の研究に従事。IEEE、電子情報通信学会、
計測自動制御学会、日本神経科学学会、各会員。

呼吸、心臓、睡眠、歩調など、私達が何気なくすごしていると
きも体は常に多種多様なリズムを作り出しています。ではそのよ
うな周期的な活動が、知覚に与える影響はないのでしょうか?
私達の感情(情動)が呼吸に深く影響することは、息が詰まる、
息を潜めるといった言葉にも見られるように、よく知られていま
す。一方で、呼吸の状態が知覚体験に与える影響はまだ研究
の途に着いたばかりです。我々のグループでは呼吸と音の関係
を探ることにより、人間理解を深めるだけでなく、より深い感動
を生むインタフェース(音呈示装置)の設計に生かすことをめざ
し、研究を行っています。

呼吸のタイミングと自律神経系の反応

呼吸は一時的に止めることはできても息を止め続けることは
できません。実際、呼吸の制御経路は自分の意思で働く大脳皮
質由来の経路と、自分の意思とは関係なく働く脳幹を中心とし
た自律神経系の経路の2つが大きく関わっています。この意思
とは関係なく働く自律神経系は、外からの刺激に常に反応しな
がら感情の基礎を形作ります。例えば音が突然聞こえたりする

と、定位反応と呼ばれる心拍、血流量、脳波等の一時的変化が
現れ、来るべき状況に対する心と体の準備を行うのです。今ま
での多くの研究は、この定位反応が呼吸といった短い周期の変
化に対しては変わらないという前提で行われてきました。我々は
独自の視点から実験を行い、この音に対する定位反応が吸う吐
くといった呼吸のタイミングによって異なることを発見しました
[1]。この反応の差は知覚的に認識できるようなものではありませんが、
体の反応が異なるということは、それが積み重なること
によって認知的な体験が異なってくる可能性があることを示し
ています(図1)。

呼吸と合った音の印象

我々は呼吸と音のタイミングの関係が音の印象に影響を与
える可能性を探るために、新たに呼吸と音の関係をコントロー
ルできるシステムを作成しました。このシステムを実装するに当
たっては、呼吸の位相を解析する信号処理技術や引き込みモ
デルなどが応用されています。このシステムを使うと、周期的変
化のある音のデータに対して、楽譜のようにどこで息を吸うか、



どこで息を吐くかの情報を追加することで、呼吸の動作と音の周期的変化を同調させることができます。すなわち、このシステムを使うことにより、継続的に息を吸うと音が高くなる、吐くと音が低くなる、といった状況を作り出すことができるようになります。我々はこの装置を使った実験を通して、呼吸と音の変化を一致させると、まったく同じ音に対しても感じ方が変わるということを示すことに成功しました[2]。つまり、呼吸をするタイミングと音の変化のタイミングをコントロールすることで、体の微弱な反応を変えるだけでなく知覚的な印象をもある程度左右できることを、示すことができたことになります(図2)。

呼吸と音の関係を極めた先の 新たな体験の提供を目指して

ではこのような呼吸変化が起きることが、音楽を聴くという知覚体験に何らかの意味を与えるのでしょうか?実は、ここまでお話しした呼吸と音のタイミングが体と感情に影響を与えるという現象を、我々は無意識のうちに活用している可能性があります。例えば一定のリズムで打たれる太鼓の音を聴いていると呼吸周期が音に引っ張られて速くなることが知られています。では呈示

される音が複雑な音楽になった場合はどうでしょうか?直感的には音楽と呼吸は関係すると思うかもしれませんが、学術的には呼吸が速くなる傾向があるという以上の知見はありませんでした。そこで、我々は同じ音楽を繰り返し同じ人に聞かせる実験を行いました。すると、同じ曲を聞かせても人それぞれに特有の呼吸パターンが浮かび上がってくることを発見しました[3]。つまり、その曲の経験を積むほど、どの音の場面で息を吸うか、吐くかといったパターンが決まってくるということです。我々が得た知見(息のタイミングと音の変化が一致すると知覚的印象が変わる)とあわせてみると、より音楽の印象を深く感じるために無意識的に呼吸を変化させているとも言えます。これらの知見は将来的により音楽を楽しむためのインタフェース等を作る上で大きなヒントにもなります。

我々のグループでは呼吸と音の関係の知見を深めるとともに、音の信号処理、触覚刺激などの様々な技術を組み合わせ[4]、より楽しめる呼吸へ誘導したり、より心に響くタイミングで音を出したりといった、新たな感動を生むインタフェースへの応用も模索して研究を行っています。

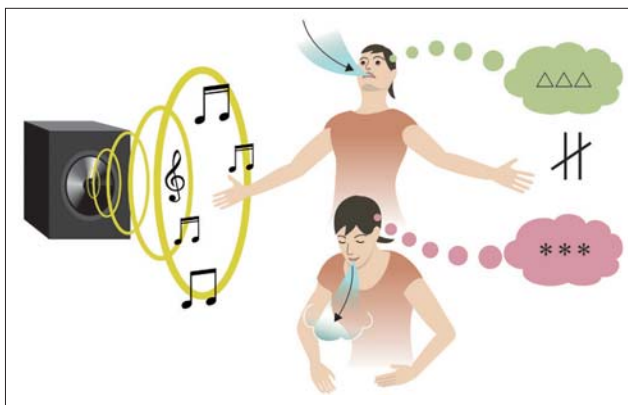


図1:物理的にまったく同じ音を聞いたとしても呼吸のタイミングによって身体の反応が異なります

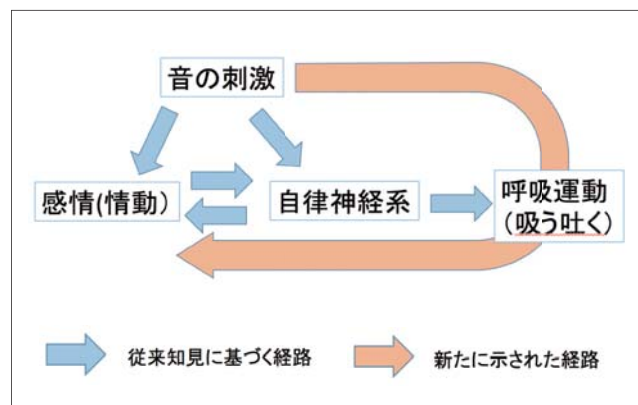


図2:我々は呼吸の運動と音の刺激タイミングを合わせることによって、感情(情動)に影響を与えることを発見しました

関連文献

- [1] T. G. Sato, Y. Ooishi, "Sound presentation during different respiration phases alters the sound-induced vasoconstriction mediated by the sympathetic nerve," *Neuroscience Letters*, Vol. 521, pp. 67-70, 2012.
- [2] T. G. Sato, J. Watanabe, T. Moriya, "Presenting changes in acoustic features synchronously to respiration alters the affective evaluation of sound," *Int. J. Psychophys*, pp. 179-186, 2016.
- [3] T. G. Sato, M. Ohsuga, T. Moriya, "Increase in the timing coincidence of a respiration event induced by listening repeatedly to the same music track," *Acoustical Science and Technology*, Vol. 33, No. 4, pp. 255-261, 2012.
- [4] T. G. Sato, M. Ohsuga, H. Boutani, T. Moriya, "Tactile phantom sensation for coaching respiration timing," *IEEE Trans. Haptic*, Vol. 8, No. 1, pp. 119-125, 2015.



研究展示のカテゴリ

NTT コミュニケーション科学基礎研究所は、「新たな時代を創る」、「時代の要請に応える」をミッションに、情報科学と人間科学の両側面から基礎研究に取り組んでいます。オープンハウス2017では、「データと学習の科学」、「計算と言語の科学」、「メディアの科学」、「コミュニケーションと人間の科学」に関する合計29展示をご紹介します。

データと学習の科学

機械学習技術とデータ解析の研究をご紹介します

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



計算と言語の科学

計算機により困難を克服する研究をご紹介します

- 膨大な組み合わせから適した解を見つけたい
- 安心・安全に通信したい
- 言葉の意味を数値で測りたい



メディアの科学

音声・画像などのメディア情報を高度・高速に処理する研究をご紹介します

- 知的に音声を認識したい
- 音を高品質で相手に届けたい
- 映像と音の本質に迫り、理解したい



コミュニケーションと人間の科学

人間の情報処理メカニズムを科学的に解明する研究をご紹介します

- 心豊かなライフスタイルをデザインしたい
- コミュニケーションの本質を深く理解したい
- 脳の仕組みを究明し、新しいICTを提案したい



01

車両によるスマートシティセンシング

～車載型センサノードによる時空間都市イベント検知～

どんな研究

ゴミ収集車のような市内を定期的に巡回する車両にセンサを設置することで、都市で発生する大小さまざまなイベントを検知する技術の研究です。たとえば、ゴミ収集車のゴミ回収イベントを検知することで、地区ごとのゴミ排出量を把握できます。

どこが凄い

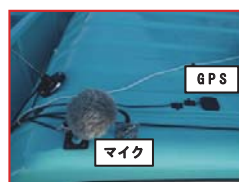
従来大まかにしか把握できなかった都市で起きているイベントを時間と空間の両面で詳細に検知できます。例えば、大気汚染のホットスポットや騒音エリア、ゴミ排出量の急激な増加など、都市で起きている様々なイベントを検知・蓄積することで、詳細な分析を可能にします。

めざす未来

都市で発生している様々なイベントを車載型センサノードを利用して詳細に検知できれば、柔軟な都市マネジメントに貢献できます。たとえば、将来予測に基づく効率的なゴミ収集などのスマートシティサービスの提供や都市計画の策定に役立ちます。

ゴミ収集車を用いた都市センシング

- ・ ゴミ収集車が定期的に市内全域を巡回
- ・ 車載型センサノードで周囲の状況をセンシング



神奈川県藤沢市のゴミ収集車にセンサノードを設置、多様なセンサデータを収集

大気状態(温度、湿度、UV、CO、O₃、PM2.5...)、振動(モーション)、音声、位置情報(GPS)



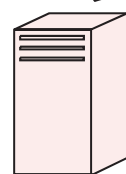
センサノード搭載ゴミ収集車

大気状態データ

加速度データ

騒音データ

位置情報



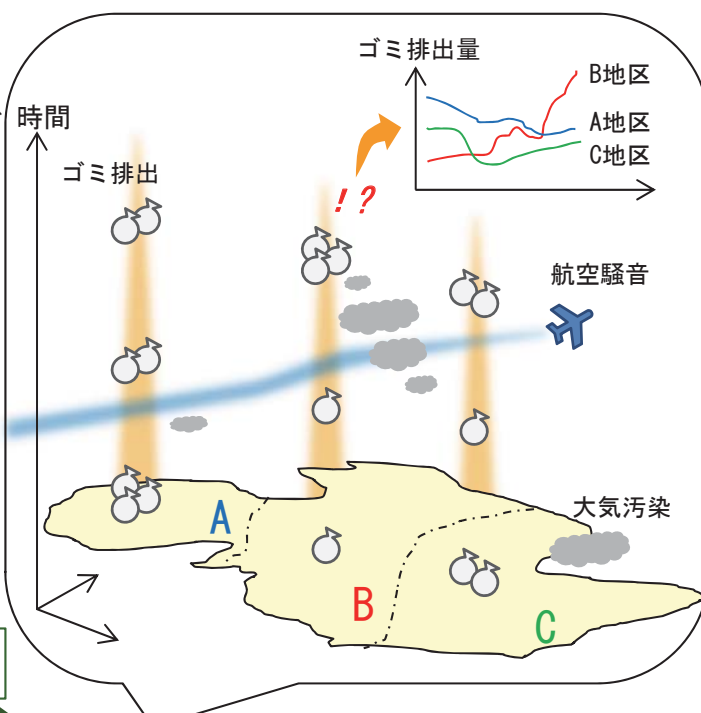
データ

マイニングサーバ

時空間都市イベント検知

車載型センサノードで計測されたセンサデータに基づくイベント推定

- ・ ゴミ収集区間推定(音、振動)
- ・ 大気汚染ホットスポット(CO、PM2.5...)



* 本研究成果の一部はNICT委託研究「ソーシャル・ビックデータ利活用・基盤技術の研究開発」によるものです。

関連文献

- [1] Y. Kishino, Y. Yanagisawa, Y. Shirai, S. Mizutani, T. Suyama, F. Naya, "Agile Environmental Monitoring Exploits Rapid Prototyping and In Situ Adaptation," *IEEE Pervasive Computing Magazine*, Vol. 16, Issue 2, pp. 61-71, 2017.
- [2] Y. Shirai, Y. Kishino, F. Naya, Y. Yanagisawa, "Toward On-Demand Urban Air Quality Monitoring using Public Vehicles," in *Proc. 2nd International Workshop on Smart Cities: People, Technology and Data (IWSC'16)*, pp. 1-6, 2016.

連絡先

岸野 泰恵 (Yasue Kishino)、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 E-mail: s-room(at)lab.ntt.co.jp

02

都市における人の活動パターンとリズムを可視化

～時空間統計に基づく都市データの解析～

どんな研究

都市から観測されるデータには人間の活動に起因した偏りが生じます。このような偏在を空間的な活動パターンと時間的なリズムとして抽出し、**解釈の容易な形に可視化**します。例えば、交通量データに適用することで、交通量の傾向を理解する際の手助けをすることができます。

どこが凄い

いつ・どこで観測が得られたかというデータと道路網などの位置に関する情報を用意すれば、**自動で活動パターンとリズムの抽出ができます**。また、センサーの故障、ネットワークの不具合、サーバー異常によってデータに欠損値が存在する場合には、観測値の推定・復元ができます。

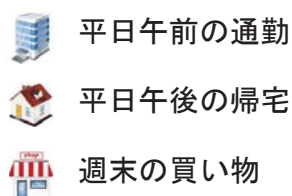
めざす未来

交通流データに限らず、人の活動に伴う多種多様なデータが今後ますます観測されるようになります。そのようなデータを**リアルタイムで解析し、活動パターンとリズムを即座に可視化**することで、緊急災害などにおける、迅速な状況把握のサポートなどを行っていきます。

都市から観測される時空間データには、**いつとどこに関する情報が含まれています**。



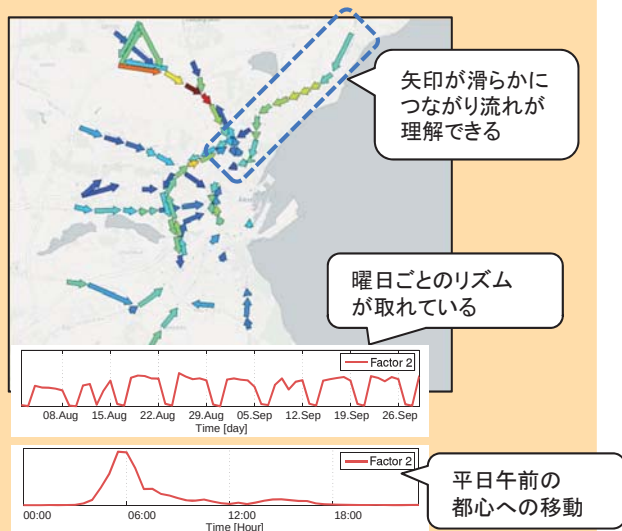
時空間データには、人間の活動に起因する**空間的な偏り**と**周期的な繰り返し**が現れます。



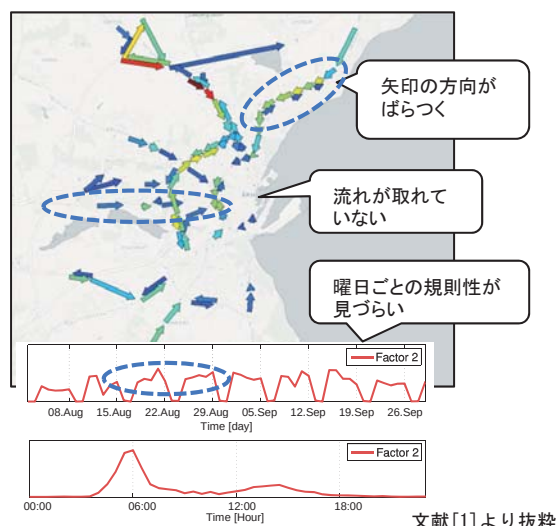
データ内の特定の曜日、時刻、位置に繰り返し発生

テンソル分解によって抽出された**少数の活動パターンとその発生リズム**を用いて、**与えられた時空間データの大部分を説明**できます。本研究では、テンソル分解に地点間の距離、時刻や日付の前後関係などを利用し、時空間的に滑らかで解釈しやすいパターン抽出を可能にしました。

提案法



既存法



関連文献

- [1] K. Takeuchi, N. Ueda, "Graph regularized Non-negative Tensor Completion for spatio-temporal data analysis," *2nd International Workshop on Smart Cities: People, Technology and Data*, 2016.
- [2] 竹内孝, 納谷太, 上田修功, "一般化KLダイバージェンスを用いた非負値テンソル補完法と都市交通流データへの応用," *第8回データ工学と情報マネジメントに関するフォーラム*, 2016.

連絡先

竹内 孝 (Koh Takeuchi) 上田特別研究室
E-mail: takeuchi.koh(at)lab.ntt.co.jp

03

集団最適なナビをリアルタイムに実現

～ベイズ的最適化に基づく効率的な誘導計画探索～

どんな研究

人や車などの集団の流れから近未来に生じる混雑・渋滞リスクを予測し、これを回避する集団最適な誘導をオンラインで自動的に導出する技術の研究です。集団の動きのシミュレーションを通じて混雑の予測や誘導策を生成し、機械学習技術を用いて効果的な誘導を効率よく探索します。

どこが凄い

不測の事態にも対応可能な集団最適誘導策をオンラインで自動導出できます。例えば「突発的なアクシデントによりこの道が封鎖されたとき」でも、どのような誘導策が混雑の緩和に寄与するかをベイズ的最適化により自動的かつ効率的に見つけ出します。

めざす未来

人や車の集団が行き交う場所でのリアルタイム観測技術と組み合わせることによってその場その時の状況に適した集団の誘導を可能にします。事前の誘導計画とは異なる不測の事態にも対応可能な誘導技術を確立し、快適かつ安全なイベント運営やインフラ安定化支援に貢献します。

めざす技術: 集団の安全・快適な移動を可能にする ICT & 機械学習技術



- 誘導策やWhat-ifシナリオのシミュレーションを用いた学習
→ 最適誘導策をオンラインで自動導出
- シミュレーション環境のパラメータを観測データに基づき自動学習

利用シーン例

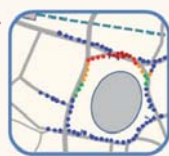
- ✓ 駅とイベント会場の人々の誘導
- 空港やテーマパーク等の待ち行列制御
- ショッピングモール周辺の車両誘導

本展示: スタジアム周辺の人流に対するベイズ的最適化による誘導策の探索

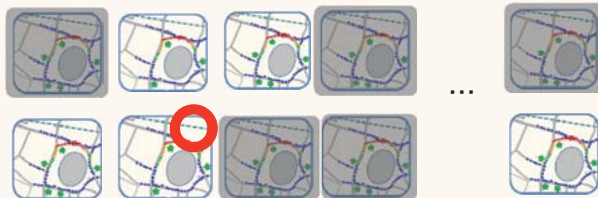
シミュレーションによる
近未来混雑リスク予測

現在

30分後



ベイズ的最適化を用いた
良い誘導策 (所要時間少) を効率的に探索



関連文献

[1] “スタジアムの誘導技術確立,” 日刊工業新聞, 3/1 全国版, 朝刊12面, 2017.

[2] “集団最適なナビを自動で学習します,” NTT R&Dフォーラム2017. https://labevent.ecl.ntt.co.jp/forum2017/elements/pdf_jpn/03/C-13_j.pdf, 2017.

連絡先

大塚 琢馬 (Takuma Otsuka) 協創情報研究部 知能創発環境研究グループ
E-mail: otsuka.takuma@lab.ntt.co.jp

04

プライバシーを保護しつつ人の動線を抽出

～人流推定のための確率的行動モデル～

どんな研究

展示会場やショッピングモール等においてスマートフォン等により取得された人の位置情報は、プライバシー保護のため集計化されることがよくあります。本展示では、人が“どこからどこへ移動したか”を追跡できない集計データから、人の動線を抽出する技術を提案します。

どこが凄い

大規模な会場では、コストや空間的な制約によって、未観測領域（通路など）が頻繁に存在します。そのような状況で観測点間の移動をより適切に捉えるために、観測点間の移動時間を組み込んだ確率モデルを提案します。これにより、従来よりも高精度な人流推定が可能となります。

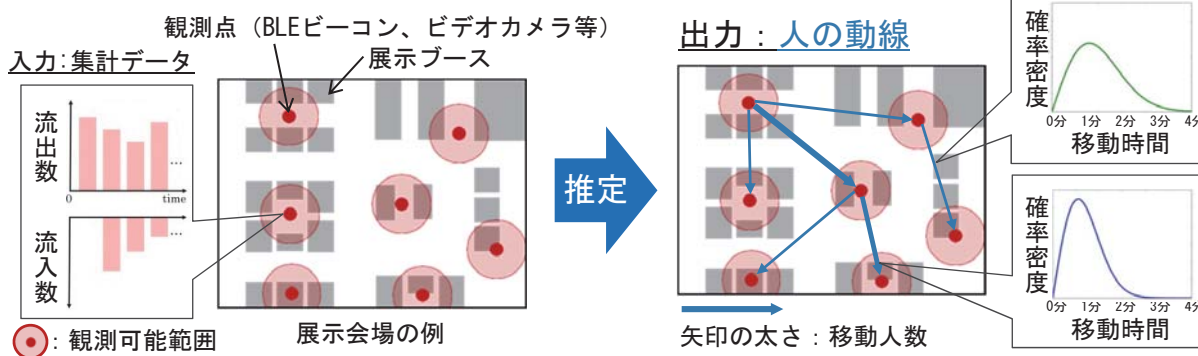
めざす未来

ユーザに関する多様なデータが観測可能になると同時に、プライバシー保護への関心が高まっています。プライバシーが保護された状況において、ユーザ行動の理解・分析を可能とする技術の確立を目指します。これにより、最適なユーザナビゲーション等の実現につながります。

観測点が少ない状況における集計データから人の動線を抽出

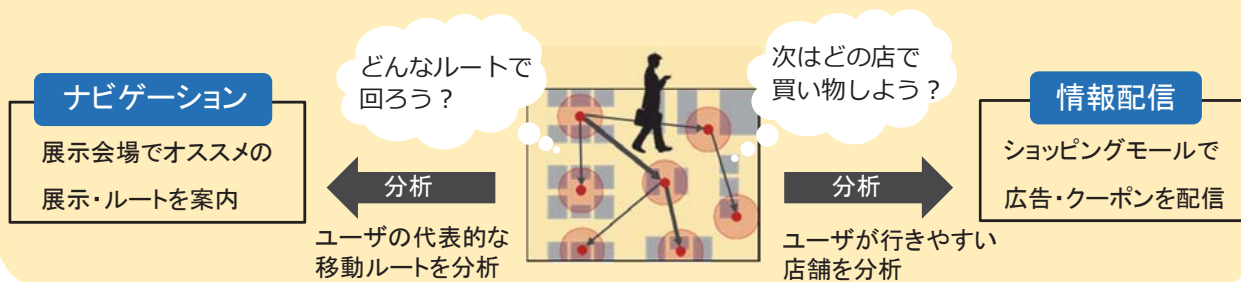
問題1：未観測領域が生じる！

問題2：個人を追跡できない！



- 提案技術：複数の観測点における流入・流出量の関係を解析し、人の動線を抽出
- 特徴：ブース間の移動時間の分布と観測点間の移動人数を同時に推定可能

応用例：ユーザのプライバシーを保護したままで…



関連文献

- [1] 田中佑典, 岩田具治, 倉島健, 戸田浩之, 上田修功, “集計データに基づく人流グラフ推定技術の検討,” 第19回情報論的学習理論ワークショップ (IBIS2016), 2016.

連絡先

田中 佑典 (Yusuke Tanaka) サービスエボリューション研究所
E-mail: tanaka.y(at)lab.ntt.co.jp

05

時間変化するデータのための安定した深層学習

～Gated recurrent unitにおける学習安定化技術～

どんな研究

音声認識や機械翻訳などにはリカレントニューラルネットワーク (RNN) が使用されています。RNNは時系列データの高精度な解析に適した深層学習技術の一つです。しかし RNN には学習が困難であるという問題があります。本研究では失敗しない RNNの学習法を提案します。

どこが凄い

RNNを使用したデータ解析では、RNNの学習が失敗してしまうため、試行錯誤やノウハウを必要とします。本技術では、RNNの一つである Gated Recurrent Unit (GRU)の学習が失敗する原因を解明し、試行錯誤なく学習する方法を提案します。

めざす未来

RNNなどを含む深層学習を用いたデータ分析は分析者の経験や勘に基づいて調整すべきパラメータが多数あります。こうしたパラメータをより解釈のしやすいもの、もしくは調整不要とすることで、深層学習を用いた解析を手軽に誰にでも利用できるようになります。

背景 Recurrent Neural Network (RNN)

様々な時系列データを表現するモデル

1. 収集したデータを用いて RNN を学習
2. 学習後の RNN を用いて予測、認識

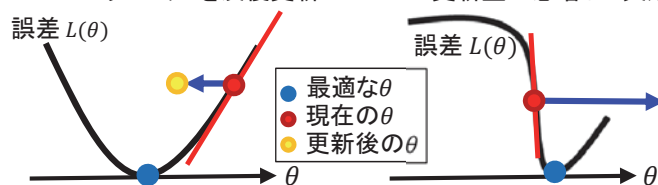


学習では誤差を下げるようにパラメータを反復更新

課題 パラメータが急激に更新されて学習が失敗

傾き(勾配)をもとに誤差の減少方向へ RNN のパラメータ θ を反復更新

勾配が急峻すぎるため更新量が急増して失敗



従来は勾配の大きさの上限を試行錯誤的に設定

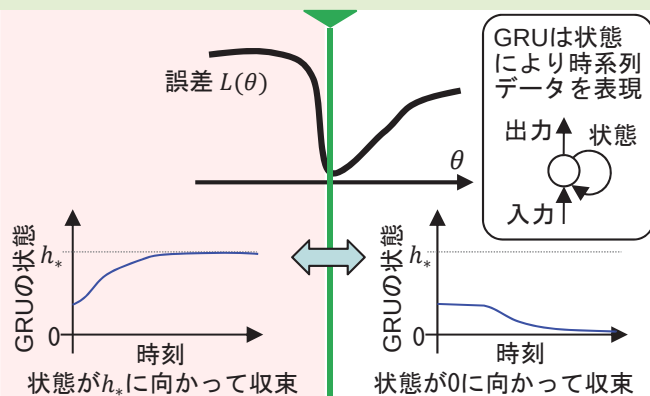
提案手法

更新量が急増する点を回避しつつ学習

(1) 更新量が急増する点を特定

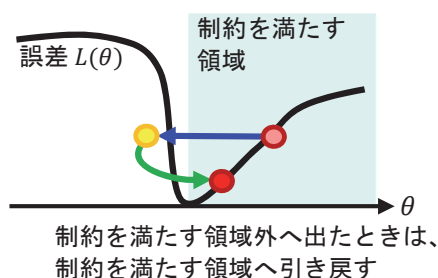
RNN の性質の切り替わる点で更新量(勾配)が急増
GRU (最新のRNN) の性質(ダイナミクス)を解析

更新量の急増する点 = 状態の軌道が切り替わる点



(2) 急増する点を回避

急増する点にパラメータを進ませない
制約つき最適化



提案する技術によって勾配の急増を抑制
試行錯誤削減による時間短縮
調整のノウハウ不要

関連文献

- [1] 金井関利, 藤原靖宏, 岩村相哲, "GRU学習時の勾配爆発の抑制方法の提案," 第19回情報論的学習理論ワークショップ (IBIS2016), 2016.

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
E-mail: kanai.sekitoshi(at)lab.ntt.co.jp

06

グラフを用いてデータの関係を素早く発見

～ L_1 -グラフ構成法のための高速化アルゴリズム～

どんな研究

ICTの進歩に伴い大規模なデータが利用可能になりつつありますが、データ間の疎なつながりを表現できるグラフ構造の研究が盛んに行われています。本研究では代表的なグラフ構造である L_1 -グラフを大規模な多次元データから高速に計算するためのアルゴリズムの開発に取り組みました。

どこが凄い

グラフにおいてノードがエッジを持つ条件を解析することにより、高速なアルゴリズムを確立しました。従来のアルゴリズムよりも大規模なデータを短時間で処理することが可能となったため、データに潜んだ関係を素早く見つけ出すことができるようになりました。

めざす未来

本技術は従来技術では考えられなかった規模のデータが分析可能であり、ビッグデータに基づいた推薦・予測・理解などに活用できます。IoTにより将来的に世界の至る所からデータが得られるようになりますが、それらを瞬時に解析・利用できる社会を実現することができます。

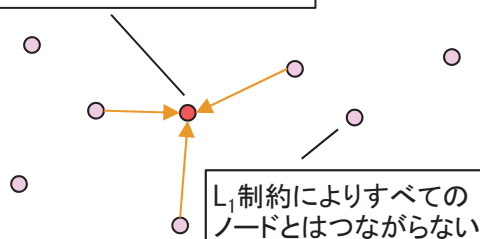
背景： L_1 -グラフ

- L_1 制約によりグラフ構造を計算しデータ間の疎な関係を適切に表現

目的関数を最小化する回帰分析からエッジの重みを計算

$$\min_{w_p} \frac{1}{2M} \|x_p - w_p X\|_2^2 + \lambda \|w_p\|_1$$

w_p : エッジの重み
 M : データの次元数
 x_p : データポイント
 X : データ行列
 λ : 正則化パラメータ



課題：高い計算コスト

- 回帰分析の係数をすべてのノードに対して計算しなければならない

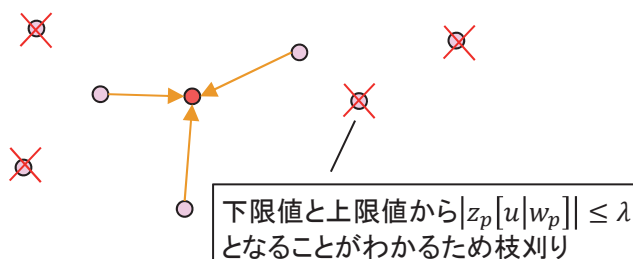
$$w_p[u] \leftarrow \begin{cases} z_p[u|w_p] - \lambda & z_p[u|w_p] > 0 \text{ かつ } |z_p[u|w_p]| > \lambda \\ z_p[u|w_p] + \lambda & z_p[u|w_p] < 0 \text{ かつ } |z_p[u|w_p]| > \lambda \\ 0 & |z_p[u|w_p]| \leq \lambda \end{cases}$$

$$\text{ここで } z_p[u|w_p] = \frac{1}{M} \sum_{i=1}^M x_u[i](x_p[i] - \bar{x}_p[i])$$

重み $w_p[u]$ を計算するためにはパラメータ $z_p[u|w_p]$ が必要になる

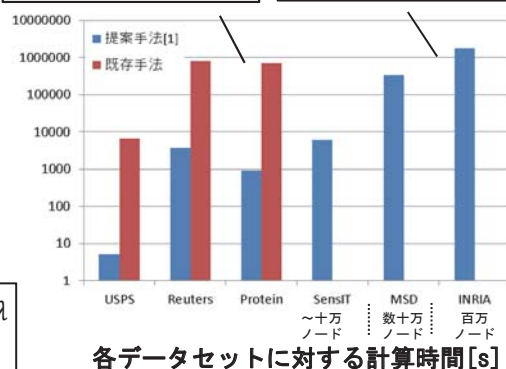
手法：ノードの枝刈り

- 特異値分解からパラメータ $z_p[u|w_p]$ の下限值 $\underline{z}_p[u|w_p]$ と上限値 $\bar{z}_p[u|w_p]$ を高速に計算



枝刈りしない手法より最大1300倍高速

提案手法は大規模データへ対応可能



関連文献

[1] Y. Fujiwara, Y. Ida, J. Arai, M. Nishimura, S. Iwamura, "Fast Algorithm for the Lasso based L_1 -Graph Construction," *PVLDB*, 10(3), pp. 229-240, 2016

連絡先

藤原 靖宏 (Yasuhiro Fujiwara) ソフトウェアイノベーションセンタ
 E-mail: fujiwara.yasuhiro(at)lab.ntt.co.jp

07

絶対に安全な共有鍵を作れるか

～今あるハードウェアで量子暗号の実現を図る～

どんな研究

絶対に盗聴されることなくランダムなビット列を二者間で共有する方法を見つける研究です。量子暗号と呼ばれる方法で、原理的には実現可能なことが保証されています。実際の装置でも実現できるように、緩やかな前提条件でも可能であると示すことを目的としています。

どこが凄い

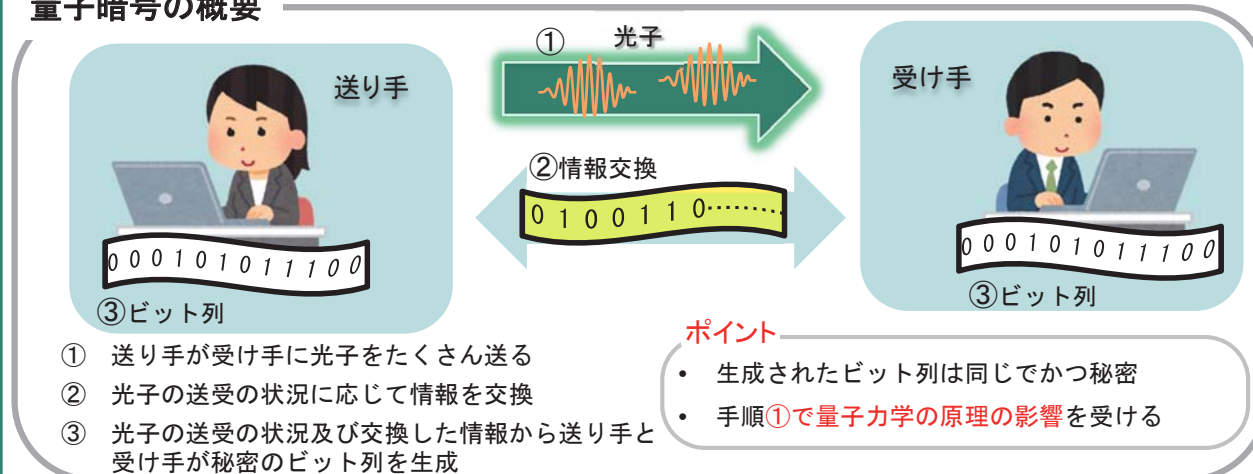
これまでの量子暗号の安全性を保証する理論では、装置が無限大の精度で動作することが仮定されていました。しかし、新しい解析手法を創出したことで、**一定の精度で動作することが保証できれば**、同一の手順で実施したとしても**安全性を担保**できることを証明しました。

めざす未来

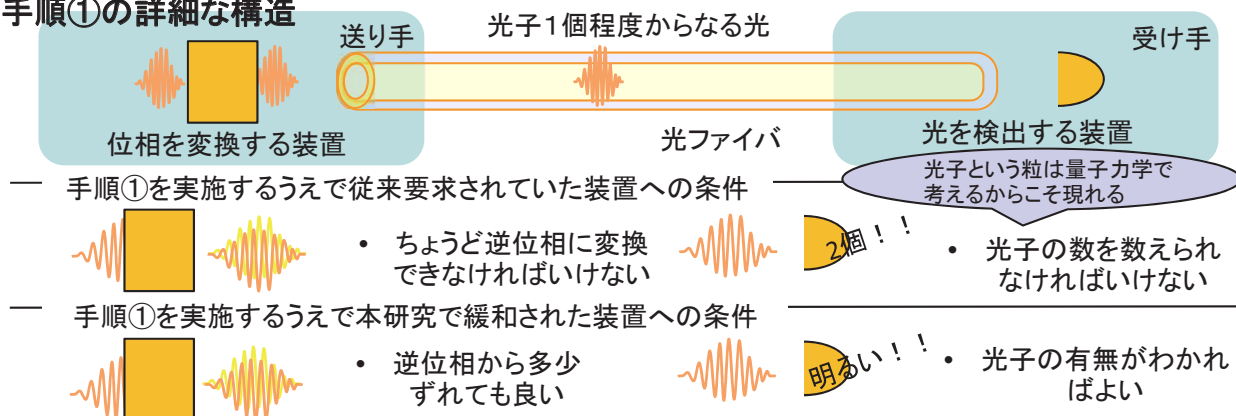
これまでの原理の実現を目指して開発された量子暗号装置では、製造者を信頼することでしか安全性を担保することができない部分が存在します。しかし、この研究の先には、利用者が**直接幾つかの項目をチェックすることによって安全性が担保**できる世界が待っています。

安全な秘密通信はランダムな秘密のビット列（秘密鍵）の共有で実現
このようなビット列の共有は通常の公開通信では原理的に不可能

量子暗号の概要



手順①の詳細な構造



関連文献

- [1] K. Tamaki, M. Curty, G. Kato, H. K. Lo, K. Azuma, "Loss-tolerant quantum cryptography with imperfect sources," *Physical Review A*, 90, 052314, 2014.
- [2] G. Kato, K. Tamaki, "Security of six-state quantum key distribution protocol with threshold detector," *Scientific Reports*, 6, Article number: 30044, 2016.

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
E-mail: kato.go(at)lab.ntt.co.jp

08

漏れなくデタラメです

～物理乱数源のランダムさを損なわずに乱数ビットを安定供給～

どんな研究

予測不可能な物理乱数生成は究極に安全な情報秘匿のための必須技術です。予測不可能性が担保された乱数をユーザが安心して使うには、**乱数が生成される仕組み**が明確にされている必要があります。この展示では安心して使える物理乱数ストリーミング装置を紹介します。

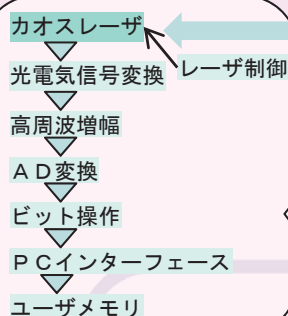
どこが凄い

物理乱数の根源であるエントロピー源の不確定性から観測、量子化、事後ビット処理を経てユーザが使う乱数ビットまでをトータルで考え、**不要なノイズの影響やビットロスがないエンドツーエンドモデルの実装**としての物理乱数ストリーミング装置を実現しました。

めざす未来

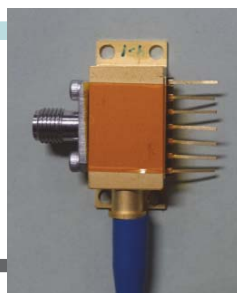
今後、大量データ秘匿化のニーズがますます増大・一般化すると考えられます。本研究は、**高速に生成された物理乱数を誰でも簡単に使える**ようにし、秘密分散や暗号化などの技術とともに用いることで、膨大なデータや計算に関して**究極の安全性**を実現することを目指しています。

エンドツーエンド物理乱数生成モデル



トータルシステムとしての乱数生成能力を確認

カオスレーザデバイス



光集積回路製作 (2011)

レーザ内の微小な不確定性をもとに不規則信号を生成

予測不可能な物理乱数を小型素子で高速に生成可能な潜在力

各処理過程において物理エントロピー源の不確定性の形は変わるが本質は変えないよう次の処理へ引き渡し、最後は乱数ビットになるよう設計

物理乱数生成装置
プロトタイプシステム

エンドツーエンド物理乱数生成モデルを実装

1Gbps 以上でフルエントロピーの物理乱数をPCのユーザメモリへ供給

初代物理乱数ストリーミング装置 (2016)



小型物理乱数ストリーミング装置 (2017)

レーザー制御、光電気信号変換、高周波増幅を、実験室レベルから一般レベルの性能にして小型化



応用例:

- ・究極の安全性をもつ秘密分散
- ・高速な科学技術計算

関連文献

- [1] T. Harayama, S. Sunada, K. Yoshimura, P. Davis, K. Tsuzuki, A. Uchida, "Fast nondeterministic random-bit generation using on-chip chaos lasers", *Physical Review A*, Vol. 83, 031803(R), 2011.
- [2] S. Shinohara, K. Arai, P. Davis, S. Sunada, T. Harayama, "Chaotic laser based physical random bit streaming system with a computer application interface", *Optics Express*, Vol. 25, pp.6461-6474, 2017.

連絡先

新井 賢一 (Kenichi Arai) メディア情報研究部 信号処理研究グループ
E-mail: arai.k(at)lab.ntt.co.jp

09

敷き詰めかたをすべて見つけます

～厳密被覆問題の高速な全解列挙法～

どんな研究

厳密被覆問題（げんみつひふくもんだい）のすべての答えを高速に見つける方法を紹介します。例えば、電子回路基盤のモジュール配置を決める問題や、マンションの間取りを決める問題などが厳密被覆問題を解くことに相当します。

どこが凄い

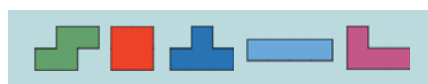
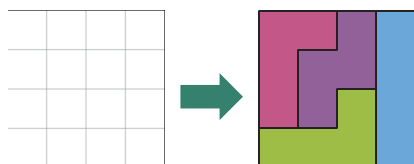
既存手法よりも**1万倍高速**に答えをすべて見つけることができます。また、見つけた膨大な数の答えを全て覚えておいて、あとから条件を満たす解を**高速に取り出す**ことができます。内部では組合せをグラフとして表現するデータ構造が用いられています。

めざす未来

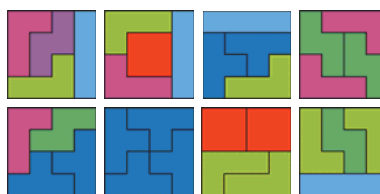
この技術によってよい敷き詰めかたを見つけることができるようになります。例えば基盤上の無駄のないモジュール配置を見つけることで消費電力を削減したり、利用者の好みに合わせたマンションの間取りを見つけることができるようになったりといった応用が考えられます。

厳密被覆（げんみつひふく）問題とは

隙間なくピースを敷き詰める方法を発見する問題



利用可能なピースの集合



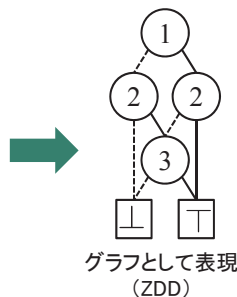
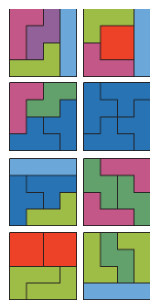
問題の特徴

膨大な数の敷き詰めかたが存在する

盤面サイズ	解の総数
4x4	117 通り
8x8	19,077,209,438 通り
12x12	13,664,822,582,333,502,156,627,512 通り

↓
解の数が急増

すべての敷き詰めかたを高速に見つける

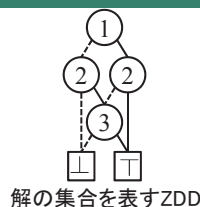


グラフとして表現 (ZDD)

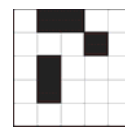
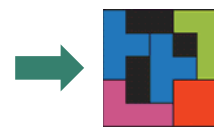
- 探索を行いつつ、すべての敷き詰めかたの集合を表すグラフ(ZDD)を構築することで、同じ部分問題の解の繰り返し計算が不要に
- 既存法よりも最大**10,000倍**高速にすべての敷き詰めかたを発見できる

※本展示は北海道大学との共同研究成果です。

条件を満たす敷き詰めかたを素早く取り出す



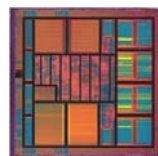
解の集合を表すZDD

障害物を避ける
解を問い合わせ

素早く解を出力

- 解の集合をZDDとして圧縮して表現することで、条件を満たす解を素早く取り出すことができる

応用例



電子回路のモジュール配置

Photo by PeterJohnBishop



マンションの間取り決め

Photo by 箱坂

関連文献

- [1] M. Nishino, N. Yasuda, S. Minato, M. Nagata, "Dancing with decision diagrams: a combined approach to exact cover," in *Proc. the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pp. 868-874, 2017.

連絡先

西野 正彬 (Masaaki Nishino) 協創情報研究部 言語知能研究グループ
E-mail: nishino.masaaki(at)lab.ntt.co.jp

10

絵本はこどもの言葉の教科書

～絵本コーパスと幼児の語彙獲得の関係を探る～

どんな研究

絵本はこどもが言語を習得する上で重要な情報源の一つだと言われていますが、絵本がどの程度語彙獲得に寄与しているかについて大規模データをを用いた検証はありませんでした。本展示では、絵本に登場する語とこどもがその語を獲得する時期の関係について分析した結果を紹介します。

どこが凄い

絵本での出現頻度と、こどもが語を言えるようになる時期の関係を調べたところ、動物名などでは両者に極めて高い相関があることを発見しました。この関係性はまた、語の学習の難しさの指標となりうることもわかりました。これは言語習得メカニズム解明につながる重要な発見です。

めざす未来

こどもにとって学習しやすい語/しにくい語を明らかにすることにより、効果的な教育支援の実現に寄与します。また、こども一人ひとりの語彙発達レベルにあった絵本の推薦を実現することで、言葉を楽しく学べる絵本を教科書として語彙獲得を後押しします。

絵本コーパス

- 絵本の本文を電子化し、解析
- ベストセラーなど2661冊、3455話



出現頻度

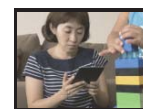
歩く: 937話 2423回
イヌ: 412話 1652回
サイ: 34話 202回
...

多く出る順

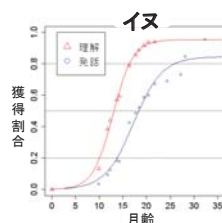


幼児語彙発達データベース

- 0-4歳の約1300名分のデータを収集
- 50%の子が言えるようになる日齢を推定



親が回答

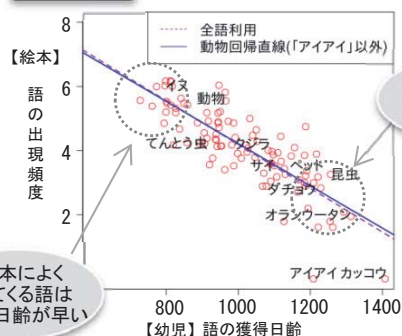


獲得日齢

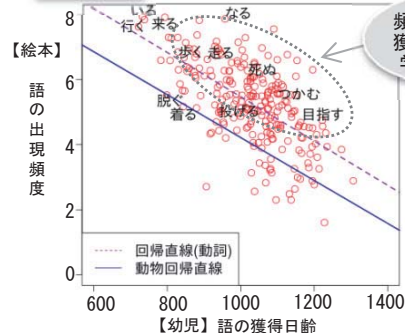
イヌ: 796日齢
歩く: 829日齢
サイ: 1100日齢
...

覚える順

絵本と語彙獲得の関係を解析

動物名 順位相関関係 $\rho = -0.83$ 

動物名は特に絵本中の頻度と獲得日齢に強い相関
→ 絵本はこどもの語彙獲得と密接に関係

動詞 順位相関関係 $\rho = -0.58$ 

動詞は動物名と比べ頻度のわりに獲得日齢が遅い
→ 頻度と獲得日齢の関係は語の学習しやすさの指標

関連文献

- [1] S. Fujita, T. Kobayashi, Y. Okumura, T. Hattori, "How do words in picture books affect child vocabulary acquisition? – An analysis of large-scale corpus in Japanese picture books –," in *Proc. Workshop on Infant Language Development (WILD-2017)*, 2017.
- [2] 藤田早苗, 小林哲生, 奥村優子, 服部正嗣, "幼児の語彙獲得と絵本コーパスの関係を探る," *言語処理学会第23回全国大会*, pp. 899-902, 2017.
- [3] 小林哲生, 奥村優子, 南泰浩, "語彙チェックリストアプリによる幼児語彙発達データ収集の試み," *電子情報通信学会技術研究報告*, Vol. 115, No. 418, pp. 1-6, 2016.

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
E-mail: fujita.sanae(at)lab.ntt.co.jp

11

ノンネイティブのためのリスニング支援

～効果的な自動字幕表示手法～

どんな研究

自動字幕はノンネイティブのリスニング能力を向上させる一方で、負担も増大させてしまうことが報告されています。本研究では、ノンネイティブがリスニング中にどのように自動字幕を利活用するかを調査し、自動字幕が解決できる問題と解決できない問題を明確にしました。

どこが凄い

ノンネイティブの特性や問題に応じて支援方法を変える必要があることを示しました。これまで、自動字幕を表示し続ける方法が取られてきましたが、自動字幕によって解決できない問題も多く存在し、状況によってはノンネイティブの負荷になってしまうことがわかりました。

めざす未来

人の特性や状況に応じて最も使いやすい形で支援技術を提供することによって、より人に優しいICTの実現を目指しています。ノンネイティブの状況をモニタリングし、適切な支援を最適なタイミングで行うことによって、言語の壁が解消された未来を目指します。

調査項目

- 1) ノンネイティブが自動字幕を利用する方法
- 2) 自動字幕を使用することの利点と欠点

実験概要

参加者: ノンネイティブ20名

条件: 音声のみ vs. 音声+字幕 (アイトラッキング)

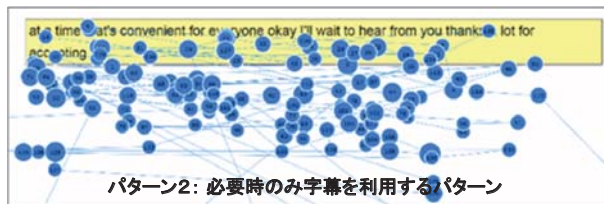
手順:

- 1) リスニング課題: 参加者はリスニング中にわからないこと(問題)があると、ボタンを押す
- 2) インタビュー: 参加者は直面した問題の内容と持続時間について説明

分析結果

ノンネイティブの自動字幕の活用法 [2]

パターン1: 字幕を追うことに注力するパターン

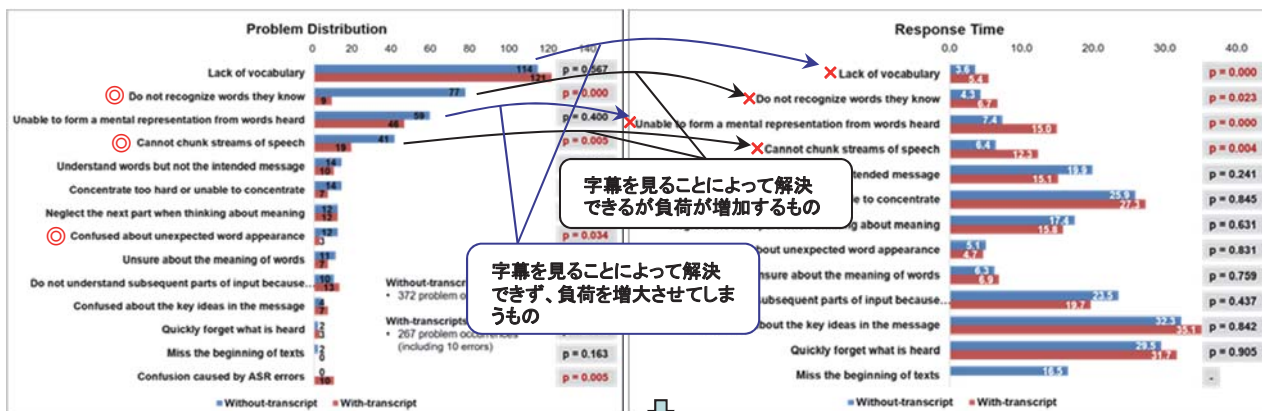


パターン2: 必要時のみ字幕を利用するパターン

↓
ノンネイティブの特性に応じて支援を変える必要性を示唆

利点: 自動字幕によって解決できた問題 [1,2]

欠点: 負荷の増加 [2]



問題に応じて支援を変える必要性を示唆

関連文献

- [1] X. Cao, N. Yamashita, T. Ishida, "How Non-Native Speakers Perceive Listening Comprehension Problems: Implications for Adaptive Support Technologies," in *Proc. the 8th International Conference on Collaboration Technologies (CollabTech 2016)*, pp. 89–104, Best Paper Award, 2016.
- [2] X. Cao, N. Yamashita, T. Ishida, "Investigating the Impact of Automated Transcripts on Non-native Speakers' Listening Comprehension," in *Proc. the 18th ACM International Conference on Multimodal Interaction (ICMI 2016)*, pp. 121–128, ACM, 2016.

連絡先

山下 直美 (Naomi Yamashita) 協創情報研究部 インタラクション対話研究グループ
E-mail: yamashita.naomi(at)lab.ntt.co.jp

12

あなた好みに翻訳する

～接頭辞制約によるニューラル翻訳の制御～

どんな研究

ニューラル機械翻訳は、原文と訳文の関係を表す確率モデルを対訳データだけから学習するので、翻訳システムが出力する文をユーザが細かく制御することが難しいという問題があります。この展示では、ユーザがニューラル機械翻訳の出力を**カスタマイズする方法**を紹介します。

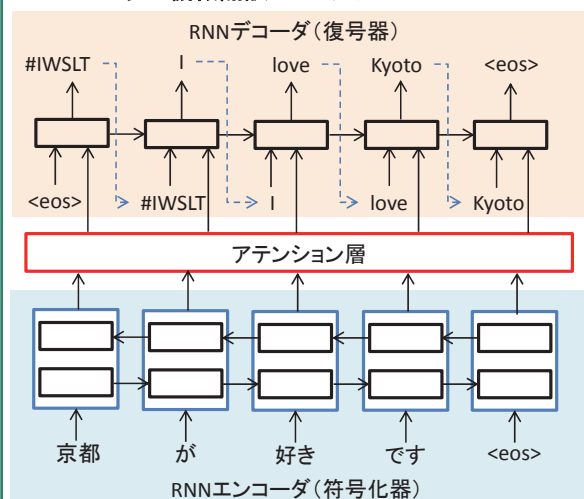
どこが凄い

本研究では、翻訳システムが出力する文の特徴をタグで表現し、入力文から**タグと出力文を同時に予測**することで、タグで表現された特徴を反映した出力文を生成します。**ユーザがタグを指定**することによりそのタグの特徴を反映した出力文を生成することも可能です。

めざす未来

翻訳の専門家は、文脈や状況、文化の違いなど様々な**言外の意味**を考慮して翻訳文を作成します。我々は、**文脈や状況を理解して**言外の意味を補いながら、最適な主語、構文、文体などを**自動的に選択**したり、ユーザへ適切な**選択肢を提示**するような機械翻訳システムを目指します。

ニューラル機械翻訳のしくみ



接頭辞タグ(出力文の特徴)の例

京都が好きです → **#3** I love Kyoto (文の長さ)
 京都が好きです → **#L2R** I love Kyoto (左から右へ生成)
 京都が好きです → **#R2L** Kyoto love I (右から左へ生成)
 京都が好きです → **#IWSLT** I love Kyoto (旅行会話分野)

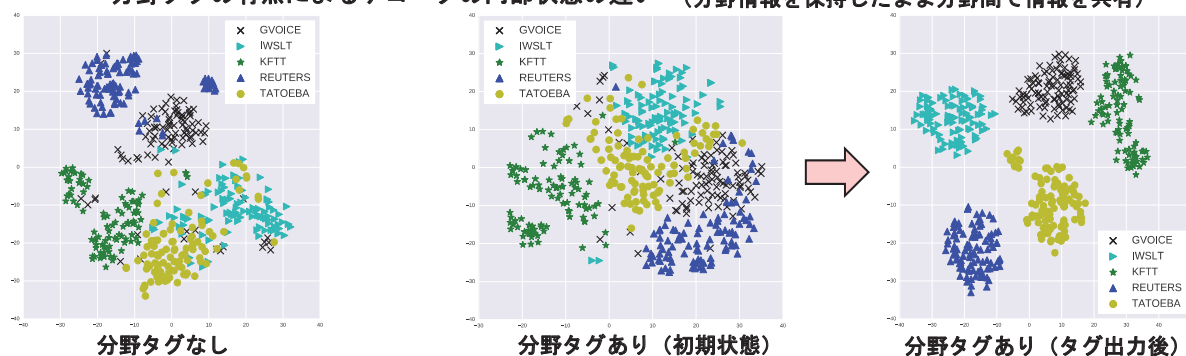
日英翻訳の精度 (BLEU)

生成方向	IWSLT	KFTT	REUTERS
順 (#L2R)	34.8	20.9	19.7
逆 (#R2L)	32.8	20.1	19.6
方向予測	35.6	21.5	20.6

タグと出力文の同時予測の例

どういたしまして
 → **#IWSLT** you're welcome
 → **#TATOEBA** don't mention it

分野タグの有無によるデコーダの内部状態の違い (分野情報を保持したまま分野間で情報を共有)



関連文献

- [1] 竹野峻輔, 永田昌明, 山本和英, “単語対応を利用した欠落語の投射による機械翻訳向きオラクル入力文の生成,” 電子情報通信学会技術報告, Vol.116, No.379, NLC2016-38, pp.135-140, 2016.
 [2] 竹野峻輔, 永田昌明, 山本和英, “ニューラル日英機械翻訳における欠落語に関する分析,” 言語処理学会第23回年次大会, pp. 659-662, 2017.

連絡先

永田 昌明 (Masaaki Nagata) 協創情報研究部 言語知能研究グループ
 E-mail: nagata.masaaki(at)lab.ntt.co.jp

13

単語の意味や関係を賢く学習

～開発者利便性が高い単語埋め込みベクトルの獲得法～

どんな研究

単語間の意味的な関連性の演算を可能とする「単語埋め込みベクトル」は、計算機による人間の直感に即した**自然言語処理/理解の実現に必須な要素技術**です。この技術を用いてシステムを構築する際に、システム開発者が不便に感じる欠点を機械学習技術を駆使して解消します。

どこが凄い

従来法と同等の性能を100分の1程度の必要メモリ量で実現する単語埋め込みベクトルです。これにより計算リソースが限定的な携帯端末のような環境でも利用可能となります。また、計算結果の再現性を担保し、利用者が安心して使える環境を提供することができます。

めざす未来

開発者利便性を高めることにより、コンピュータによる自動翻訳／質問応答／対話エージェントといった自然言語を介する**あらゆる人工知能関連システムで使われる基盤技術**の構築を目指しています。また、本技術を発展させることで、実システムの性能の底上げを実現します。

単語埋め込みベクトル

- 単語を（通常50から1000次元程度の）ベクトルで表現したもの
- 意味的な処理に基づく高度な自然言語処理/理解を実現する重要な要素技術

大規模な言語データ

例：新聞記事、web全体

	今日	明日	日本	アメリカ	中国	韓国	2016	円	ドル	元	車	自転車
今日	0.045866	-0.00371	-0.00497	0.070959	-0.06238	0.0114	-0.02119					
明日	-0.3533	-0.03058	0.08013	0.149368	0.297801	-0.09778	-0.14325					
日本	-0.01023	0.075763	0.021459	0.049479	-0.15017	-0.15543	0.222465					
アメリカ	-0.1734	0.003588	-0.09084	0.307067	0.085198	-0.16883	-0.18823					
中国	0.078726	-0.11543	0.02265	-0.06419	0.032716	0.001905	-0.20058	-0.03091	0.080452	-0.02351		
韓国	-0.02387	-0.05605	-0.20171	-0.15812	-0.14642	-0.04322	-0.36656	-0.31666	-0.14884	-0.1488		
2016	0.149431	0.079334	0.293562	-0.02928	0.262186	-0.20184	-0.01984	-0.04255	0.111553	0.095227		
円	0.163471	0.032219	-0.03764	-0.00735	0.057619	-0.1497	0.055826	-0.11493	-0.0324	-0.05825		
ドル	0.029654	0.105802	-0.32926	-0.10276	-0.27228	-0.11189	0.097063	0.220544	-0.02437	-0.06406		
元	-0.16673	0.022848	-0.08524	0.11159	0.093743	0.062614	-0.00602	0.061345	0.102776	-0.03525		
車	0.044333	-0.12678	-0.00515	0.111257	-0.26432	-0.07705	0.277751	-0.10717	-0.2606	0.201264		
自転車	-0.08889	-0.15497	-0.0867	0.196286	0.337721	-0.24721	0.198682	-0.01153	-0.00471	0.046678		

単語埋め込みベクトル

ニューラルネット/深層学習



Hello! 機械翻訳 こんにちは

人工知能言語理解



質問応答/情報検索



自然言語処理のあらゆる応用システムで活用

従来法により生成された単語埋め込みベクトルは、利用者（研究者/開発者）にとって使いにくい欠点があることが知られている

- 例1：計算結果が毎回違う
- 例2：膨大なメモリ量が必要

単語間の意味的な「関連性」

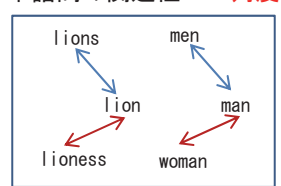
⇒ ベクトル間の「距離」「角度」で表現
また、関連性を「演算」することが可能

単語間の類似度 ⇒ 距離



例：意味的に類似した単語は空間内で近い位置に配置される

単語間の関連性 ⇒ 角度



例：単数複数の関係(青) 性別(男女)の関係(赤)

[再現性の向上] 本成果：いつ、誰が、どこで実行しても同じ結果を獲得可能

- 毎回同じ結果が安定的に得られることで、システム開発を安定して行うことが可能に

[省メモリ化] 単語埋め込みベクトル利用時の必要メモリ量を10～100分の1に削減

- 携帯端末のような計算リソースが少ない環境でも単語埋め込みベクトルが利用可能に

関連文献

- [1] J. Suzuki, M. Nagata, "Right-truncatable neural word embeddings," in *Proc. of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pp. 1145-1151, 2016.
- [2] J. Suzuki, M. Nagata, "Learning compact neural word embeddings by parameter space sharing," in *Proc. of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 2046-2052, 2016.

連絡先

鈴木 潤 (Jun Suzuki) 協創情報研究部 言語知能研究グループ
E-mail: suzuki.jun(at)lab.ntt.co.jp

14

照明光で目的の色だけをくっきり見せる

～スペクトル最適化による色彩強調～

どんな研究

照明光スペクトルを目的に合わせて最適化することで、対象物を肉眼で直接観察した際の色彩を強調する技術の研究です。全体のカラーバランスを保ちながらの彩度強調や僅かな色の差を強調する事が可能です。博物館などでの展示照明や工場などでの品質管理への適用が考えられます。

どこが凄い

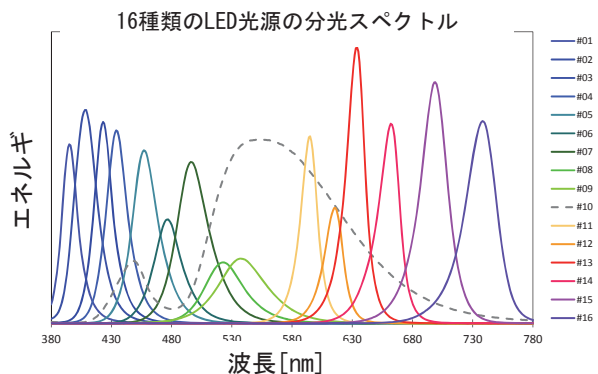
人間は白色などを基準に対象物の色を認識し、その基準が変化すると印象が大きく変わります。本技術では、白色の見え方は維持しつつ特定の色の鮮やかさを強調する、といった目標を実現します。被写体の分光反射率を考慮した照明光スペクトルの最適化により可能となりました。

めざす未来

カメラやモニタ越しではなく、実物を直接観察するので、物体の質感が保たれます。またLED光源の組み合わせで実現でき、導入コストも抑えられます。将来的には、劣化・損傷した文化財の仮想的な色彩復元や、産業・医療分野における異常領域の可視化への応用が期待されています。

課題

白色の測色値を維持しつつ、赤・緑・青の各カラーパッチの彩度が最大となる各LEDの明るさに相当する重み係数を決定



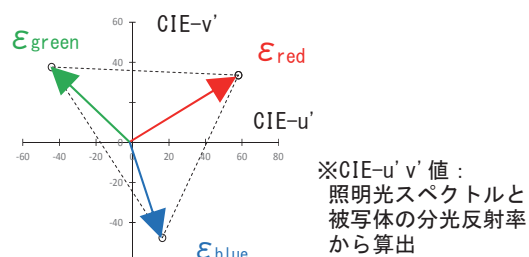
提案手法

以下の評価値 ε を最大化する重み係数を決定

$$\varepsilon = (\varepsilon_{red} + \varepsilon_{green} + \varepsilon_{blue})$$

拘束条件

白色に関し、色差 $\Delta u'v' \leq 0.01$, CIE-Yの差を1%以下

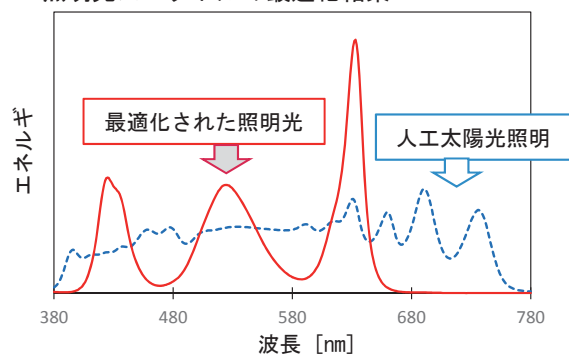


実験結果

対象物：経年劣化で退色した浮世絵



照明光スペクトルの最適化結果



X-rite色票 (赤色、緑色、青色) の鮮やかさの変化

CIE- a^*b^* 空間上で、赤：30%、緑：59%、青：13%の拡大

関連文献

- [1] M. Tsuchida, K. Hiramatsu, K. Kashino, "Designing Spectral Power Distribution of Illumination with Color Chart to Enhance Color Saturation," in Proc. IS&T 24th Color and Imaging conference (CIC24), pp. 278-282, 2016.

連絡先

土田 勝 (Masaru Tsuchida) メディア情報研究部 メディア認識研究グループ
E-mail: tsuchida.masaru(at)lab.ntt.co.jp

15

2Dと3D、どちらも同時に楽しめます

～メガネなしでも二重に見えないステレオ映像生成法～

どんな研究

従来のステレオ映像は、専用の眼鏡をかけずに見ると左右画像同士が重なり合ってしまうため、眼鏡をかけない視聴者は映像を十分に楽しめません。そこで私たちは、眼鏡をかけた視聴者とかけない視聴者、両者が同時に楽しめるステレオ映像を生成できる技術を考案しました。

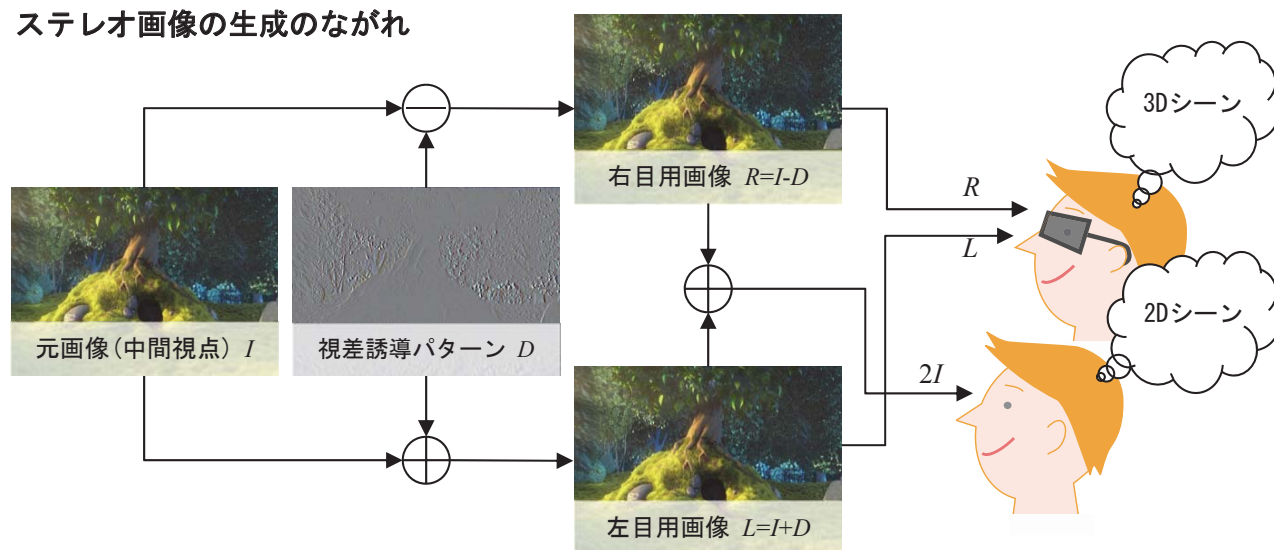
どこが凄い

眼鏡をかけた視聴者には3D映像を、眼鏡をかけない視聴者には鮮明な2D映像を提供します。本手法で生成したステレオ画像ペアは、左右画像同士が重なり合うと、両目のちょうど中間視点から見た時の鮮明な画像ができます。眼鏡をかけない視聴者はこの中間視点から見た鮮明な映像を楽しめます。

めざす未来

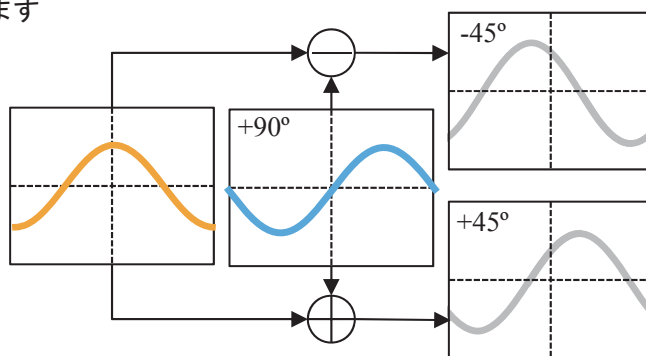
これまでのステレオ映像は、コンテンツ側が「眼鏡をかける」という楽しみ方を強制する側面がありました。私たちの技術は、コンテンツはそのままに、その場にいる視聴者1人1人が楽しみ方を自由に選択できるという「人にやさしい情報伝達」を実現します。

ステレオ画像の生成のながれ



基本原理 位相の異なる2つの正弦波を足し合わせると、その中間の位相の正弦波になるという単振動合成の性質を応用しています

- ・ 元画像から90度分位相をシフトした成分を加算/減算することで、互いに逆方向に位相がシフトした2枚の画像ができます
- ・ 90度位相シフト成分の振幅を変えることで、位相シフト(＝視差)の量を調節できます
- ・ 与えられる視差量に限界はありますが、十分自然な立体感が知覚できる映像が作れます



関連文献

- [1] T. Fukiage, T. Kawabe, S. Nishida, "Hiding of phase-based stereo disparity for ghost-free viewing without glasses," SIGGRAPH2017 (to appear).

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
E-mail: fukiage.taiki(at)lab.ntt.co.jp

16

あなた専用のお手本映像で上達支援

～深層学習によるメディア生成の可能性～

どんな研究

「したいことがあるけど、やり方が分からない…」そのような時に、ユーザが与えた音声や画像などのメディア情報を元に**具体的な指示や理想像を提示**するシステムの実現を目指しています。本展示では、このシステム実現のために行った二つの取組について紹介します。

どこが凄い

従来のフィードバックシステムは、ルールベースのものが主で、ユーザの多様な入力状態に対して対応することは困難でした。これに対し、本研究では、**深層学習をベースに新たな情報伝播方法・ネットワーク構造**を考案し、「あなた専用のお手本」を提示することを可能にしました。

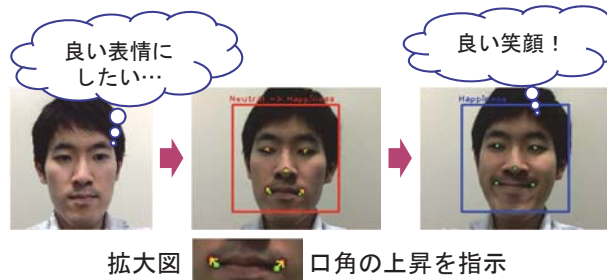
めざす未来

本研究の鍵は、システムが**メディア生成を通してユーザに寄り添えるか**ということです。これは、上達支援だけにとどまらず、ユーザが思い描くメディアを生成するために必要な技術です。将来的にはユーザの**あらゆる要望に応えられるメディア生成技術**の確立を目指しています。

「あなた専用のお手本映像」の生成（例：顔の表情改善・顔の属性変換の場合）



具体的な指示の提示（例：顔の表情改善の場合）

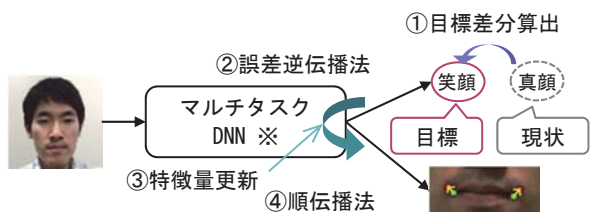


理想像の提示（例：顔の属性変換の場合）



Feedback Propagation [1]

フィードバック生成のための情報伝播方法を提案

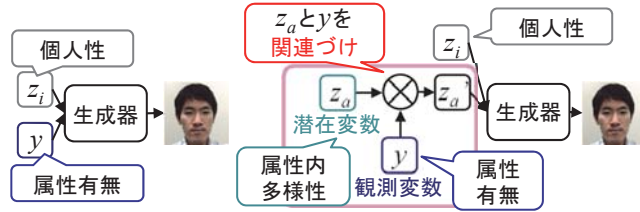


⑤フィードバック生成

※表情認識とキーポイント検出について学習済みモデルを利用

Conditional Filtered GAN [2]

属性内多様性の学習のためのネットワーク構造を提案

属性の有無変換のみ
従来法: CGAN属性の多様な変換が可能
提案法: CFGAN

関連文献

- [1] T. Kaneko, K. Hiramatsu, K. Kashino, "Adaptive visual feedback generation for facial expression improvement with multi-task deep neural networks," in *Proc. The 24th ACM International Conference on Multimedia (ACMMM)*, pp. 327-331, 2016.
- [2] T. Kaneko, K. Hiramatsu, K. Kashino, "Generative attribute controller with conditional filtered generative adversarial networks," in *Proc. The 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017 (to appear).
- [3] T. Kaneko, H. Kameoka, N. Hojo, Y. Ijima, K. Hiramatsu, K. Kashino, "Generative adversarial network-based postfilter for statistical parametric speech synthesis," in *Proc. The 42nd IEEE International Conference on Acoustic, Speech and Signal Processing*, 2017.

連絡先

金子 卓弘 (Takuhiro Kaneko) メディア情報研究部 メディア認識研究グループ
E-mail: kaneko.takuhiro(at)lab.ntt.co.jp

17

英語の発音をネイティブのように綺麗に変換

～声道モデルと深層生成モデルを用いた統計的音声変換～

どんな研究

非母語話者の聞き取りにくい音声を**母語話者風の発音の音声に自動変換**する研究です。本技術では、複合線スペクトル対 (LSP) 表現と呼ぶ独自の声道モデルを用いた信号処理アプローチと、敵対的生成ネットワーク (GAN) を用いた深層学習アプローチをとっています。

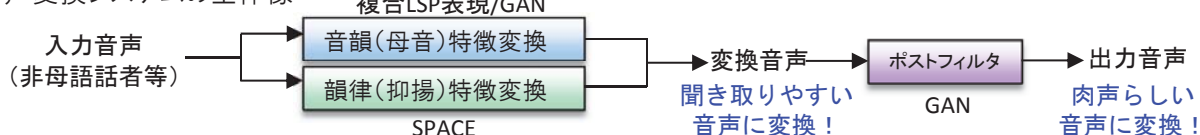
どこが凄い

本技術は**変換音声の自然さ**を多面的に追求しています。複合LSP表現では音声の生成過程のモデルを用いており、発声器官から生成される範囲の自然さを保ったまま各母音の特徴を個別に調整できます。GANポストフィルタでは合成音声をより肉声らしくなるように変換・合成します。

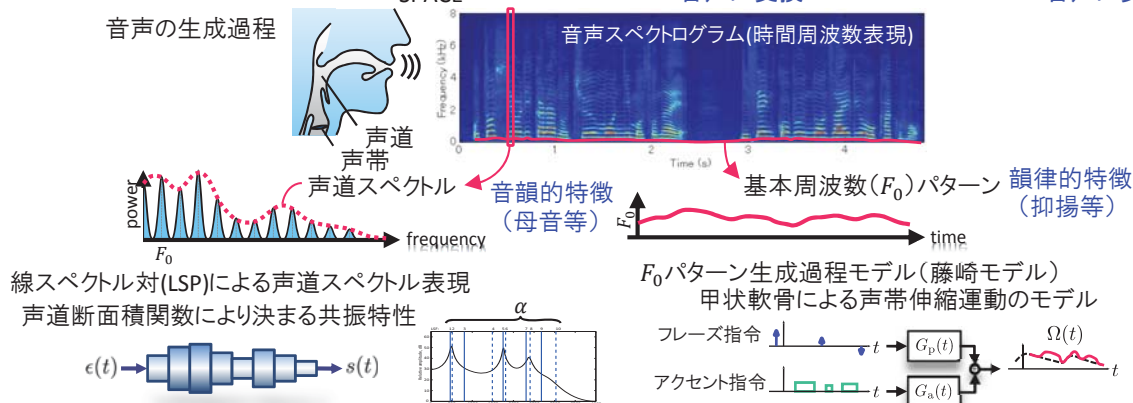
めざす未来

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

音声変換システムの全体像



音声の生成過程



線スペクトル対 (LSP) による声道スペクトル表現
声道断面積関数により決まる共振特性

$$\epsilon(t) \rightarrow \text{声道断面積関数} \rightarrow s(t)$$

複合LSP表現

- $(F_1, \dots, F_p)$ 空間上の代表点の凸結合で各時刻のLSPパラメータ(フォルマント)を表現
- 声道スペクトログラム全体をモデル化
⇒フォルマント代表点の操作で各母音の特徴を個別に操作可能

$$X(\omega_k, t_n) = \frac{c_n 2^{-P}}{A(\omega_k, t_n; \alpha)} \sin^2\left(\frac{\omega}{2}\right) \prod_{p \in \text{even}} (\cos \omega - \cos \alpha_{p,n})^2 + \cos^2\left(\frac{\omega}{2}\right) \prod_{p \in \text{odd}} (\cos \omega - \cos \alpha_{p,n})^2$$



自然性の高い音韻制御の実現

SPACE (Statistical Phrase/Accent Command Estimation)

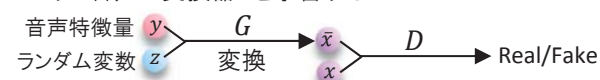
- 離散時間確率過程による藤崎モデルの確率モデル化と統計的手法を駆使した効率的なフレーズ・アクセント指令推定アルゴリズム

HMMIによる指令列モデル

自然性の高い韻律制御の実現

GAN (Generative Adversarial Net) による特徴量変換

- 本物の音声か合成音声かを識別する識別器Dをだますように音声の変換器Gを学習する



肉声感のある音声に変換可能

関連文献

- [1] H. Kameoka, K. Yoshizato, T. Ishihara, K. Kadowaki, Y. Ohishi, K. Kashino, "Generative modeling of voice fundamental frequency contours," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, Vol. 23, No. 6, pp. 1042-1053, Jun. 2015.
- [2] T. Kaneko, H. Kameoka, N. Hojo, Y. Ijima, K. Hiramatsu, K. Kashino, "Generative adversarial network-based postfilter for statistical parametric speech synthesis," in *Proc. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2017)*, pp. 4910-4914, Mar. 2017.

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
E-mail: kameoka.hirokazu(at)lab.ntt.co.jp

18

現場の生音、高音質電話でクリアにお届け

～リアルタイム音源伝送向け低遅延・高音質符号化CLEAR～

どんな研究

音声や音響信号を、決められたビットレートや処理遅延の中で**最大限に音質を保ったまま圧縮し、伝送する技術**に関する研究です。実況中継のように音声のリアルタイム伝送が必要な場合に、ビットレートに制限のある低遅延・高信頼ネットワーク（ひかり電話網）を活用できます。

どこが凄い

NTT研究所の開発したCLEARは、NTTグループの持つ低遅延・高信頼・低価格なひかり電話網に適した、リアルタイム音声通信向け音声音響符号化方式です。音声通信に許されるビットレートを最大限利用し、多くの音声信号を歪みなく圧縮、伝送することができます。

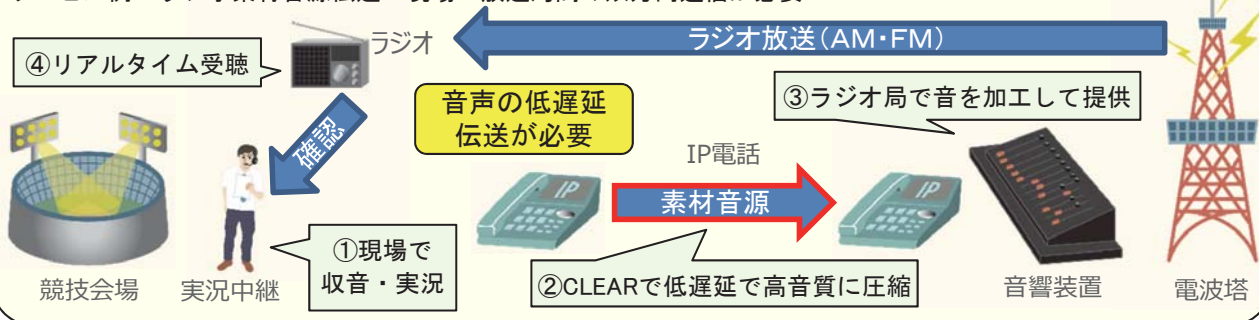
めざす未来

NTT研究所では、高い臨場感をお届けする低遅延高音質音声伝送技術を提案しています。皆様のご家庭でも使えるネットワーク（ひかり電話網）を活用し、その価値を高める技術を提供することで、従来のネットワーク（ISDN網）よりも高品質な音声通信の提供を目指します。

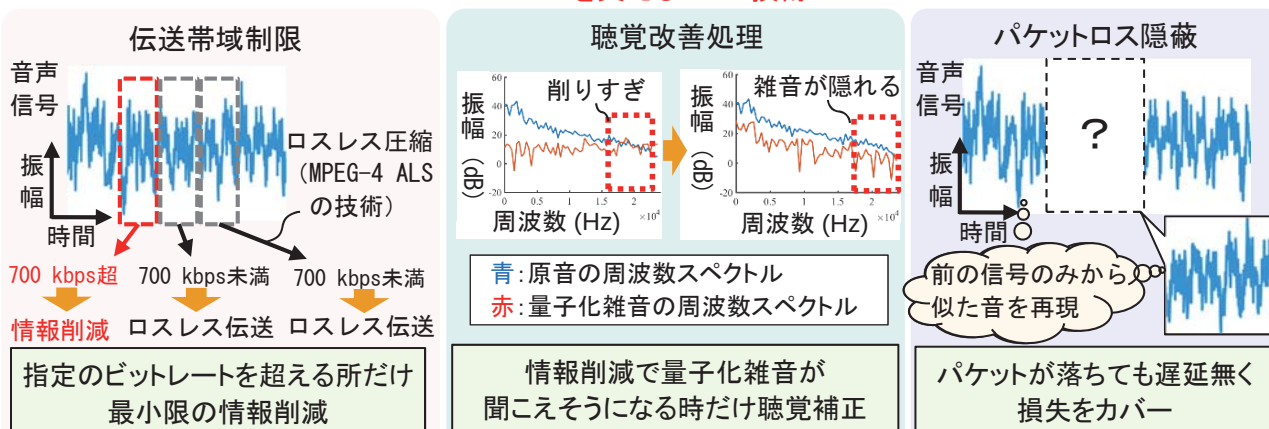
CLEAR(Conditionally Lossless Encoding under Allowed Rates)

- ・低遅延・高音質（16ビット振幅、48 kHzサンプリングレート、ステレオ、700 kbps、原理遅延 1.3 ms）
- ・指定したビットレート内で**可能な限りロスレス**で音声を圧縮

サービス例：ラジオ素材音源伝送…現場⇄放送局間の双方向通信が必要



CLEARを支える3つの技術



※本研究成果はNTTサービスエボリューション研究所（EV研）との共同開発

関連文献

連絡先

杉浦 亮介 (Ryosuke Sugiura) 守谷特別研究室
E-mail: sugiura.ryosuke(at)lab.ntt.co.jp, als(at)lab.ntt.co.jp

19 世界中の言葉を聞き分けます

～多言語音声識別・認識技術～

どんな研究

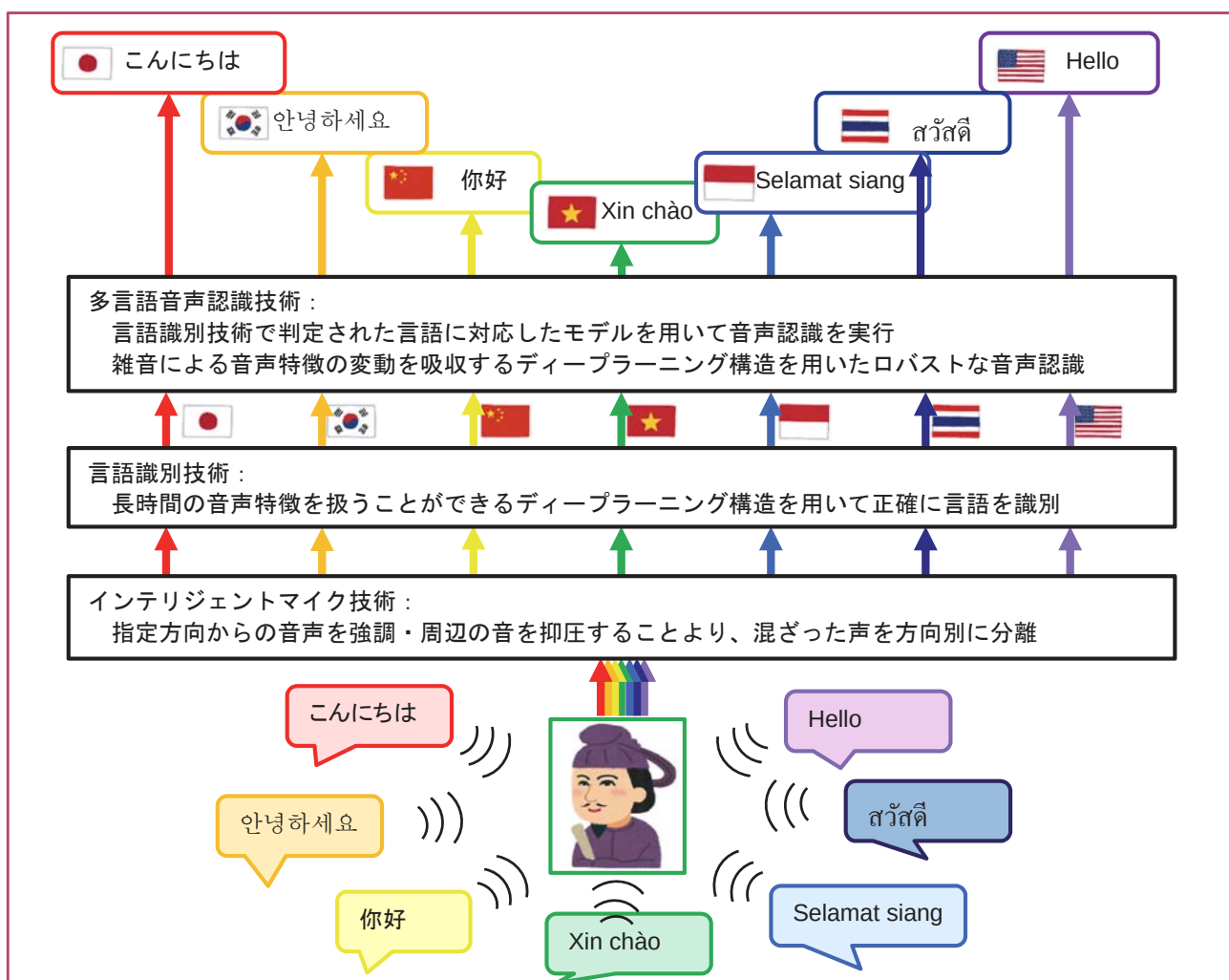
同時に話された**世界中の言葉を瞬時に聞き分けて文字に変換する研究**です。声を分離して聞き取るインテリジェントマイク技術、音声がどの言語で話されたかを判定する言語識別技術、音声を文字化する多言語音声認識技術から構成されています。

どこが凄い

インテリジェントマイクは、**最大7人が同時に話した声を分離**することができる技術です。また、雑音下での音声認識の国際技術評価で世界1位の精度を達成した音声認識技術を多言語対応し、言語識別技術と組み合わせて**アジア・パシフィック10言語の認識**を可能としました。

めざす未来

今後は対応言語を増やし、**世界のあらゆる言葉の認識**を目指します。また、対話技術・感情認識技術などを組み合わせることにより、日本を訪れる外国からの観光客に対して、**自然で気の利いた応対**ができるような技術の実現を目指します。



関連文献

- [1] T. Yoshioka, N. Ito, M. Delcroix, A. Ogawa, K. Kinoshita, M. Fujimoto, C. Yu, W. J. 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. ASRU2015*, pp. 436-443, 2015.
- [2] R. Masumura, T. Asami, H. Masataki, Y. Aono, "Parallel Phonetically Aware DNNs and LSTM-RNNs for Frame-by-Frame Discriminative Modeling of Spoken Identification," in *Proc. ICASSP2017*, 2017.

連絡先

政瀧 浩和 (Hirokazu Masataki) メディアインテリジェンス研究所
E-mail: masataki.hirokazu(at)lab.ntt.co.jp

20

あなたの声に即座にあわせて音声認識

～音声認識精度を高めるニューラルネットの高速適応技術～

どんな研究

近年、様々な場面で音声認識が利用されるようになってきましたが、現在の技術でも、**誰の声でも同じように正確に認識できるわけではありません**。本展示では、システムが話者の声の特徴に即座に適応することで、常に高い音声認識精度を実現するモデル適応技術を紹介します。

どこが凄い

音声認識のための**ニューラルネット**を、**数秒間の音声データ**だけを用いて、その話者の声を最も認識しやすいように**即座に適応させる**ことができます。ニューラルネットワークの膨大なモデルパラメータを話者の特徴に合わせるには、従来は大量の音声データでの学習が必要でした。

めざす未来

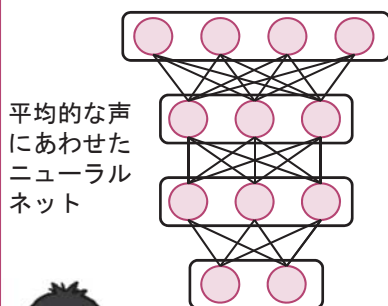
話者の声の特徴だけではなく、収録している部屋の音響特性や、雑音の性質に応じて、**常に高い精度の音声認識**が実現できるようになります。さらに、音声認識だけではなく**入力データの特徴に合わせて振る舞いを即座に変更・最適化**する人工知能技術の実現を目指します。

【従来の音声認識】

平均的な声にあわせたモデルを使用するため、話者の声の特徴によっては、高い精度は得られない。

Can you understand my **choice**? 

↑ 単語誤り率: 5.6%※



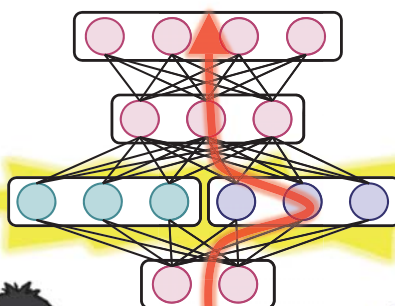
Can you understand my **voice**?

【話者の特徴に即座に適応する音声認識】

話者の声の特徴を抽出し、その特徴に応じて認識しやすいように、ニューラルネットワークの情報の流れを制御することで、各話者にあわせた音声認識を実現。

Can you understand my **voice**? 

↑ 単語誤り率: 4.8%※
※英語新聞読み上げ音声認識での評価



声の特徴に応じて、ニューラルネット内の情報の流れを認識精度が上がるように最適化



Can you understand my **voice**?

発話からその話者の音響的特徴を抽出

数秒程度の短い音声データからでもOK!

関連文献

- [1] M. Delcroix, K. Kinoshita, T. Hori, T. Nakatani, "Context adaptive deep neural networks for fast acoustic model adaptation," in *Proc. ICASSP*, pp. 5270-5274, 2015.
- [2] M. Delcroix, 木下慶介, Yu Chengzhu, 小川厚徳, 吉岡拓也, 中谷智広, "音響コンテキスト適応型DNNに基づく高速音響モデル適応," *日本音響学会春季研究発表*, 2016.
- [3] M. Delcroix, K. Kinoshita, A. Ogawa, T. Yoshioka, D. Tran, T. Nakatani, "Context adaptive neural network for rapid adaptation of deep CNN based acoustic models," in *Proc. Interspeech*, 2016.

連絡先

デルクロア マーク (Marc Delcroix)
E-mail: marc.delcroix(at)lab.ntt.co.jp

メディア情報研究部 信号処理研究グループ

21

会話の流れを読んでみんなの声をうまく認識

～頑健な音声処理を可能にする複数人発話区間推定技術～

どんな研究

複数人の会話音声から、各話者が話した時間区間（発話区間）を精度よく推定する技術を紹介します。発話区間が分かると、例えば議事録に発言者を併記したり、対話ロボットにタイミングよく発言させたりできるほか、発話区間情報を活用した各話者の音声認識率の改善も可能です。

どこが凄い

各話者の位置特徴量を精度よく表現する統計モデルに基づき高精度な発話区間推定が可能です。また、発話区間情報を使った安定的な各話者音声の強調技術や、発話区間や話者交替の情報をを使った言語モデルにより、複数人会話における各話者の音声認識率を改善できます。

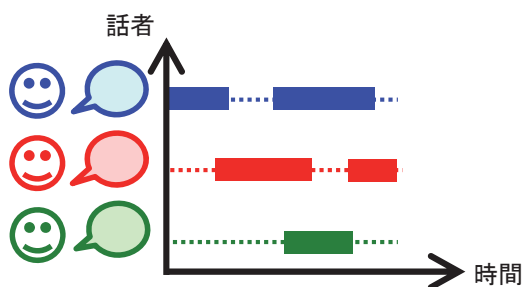
めざす未来

現状の音声認識は、一人の人がスマホやロボットに意識して話しかける利用の仕方がほとんどです。これに対し「いつ誰が何を話したか」を高精度に自動推定する本技術で、人と人が話す自然な会話の中にコンピュータが入って来られるような音声入力インタフェースを実現します。

◆発話区間推定とは：

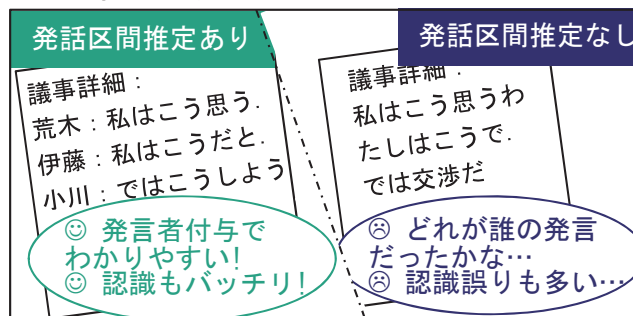
各話者が話した時間区間を推定する技術

（ダイアライゼーションと呼ぶことも）



◆応用例：

・会議録の自動作成



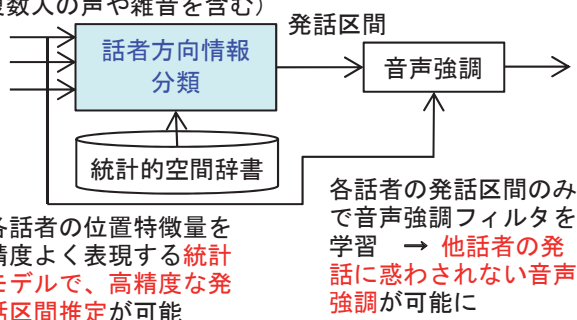
・会話シーンでの難聴者支援

・コミュニケーションロボットの制御

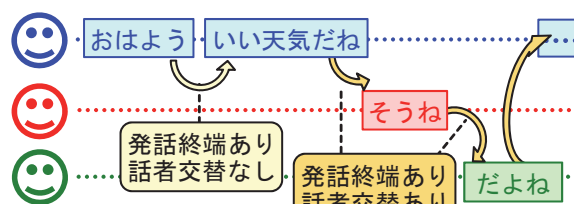
◆発話区間推定技術と、その音声処理技術への応用：

複数人会話の発話区間推定と音声強調

複数マイクでの会話収録音
（複数人の声や雑音を含む）



発話区間情報を用いた会話認識用言語モデル



発話区間情報（発話終端や話者交替など）を用いることで、会話の流れを正確に捉えて、次の発話を予測しながらより正確な音声認識が可能に

関連文献

- [1] N. Ito, A. Araki, M. Delcroix, T. Nakatani, "Probabilistic spatial dictionary based online adaptive beamforming for meeting recognition in noisy and reverberant environments," in *Proc. ICASSP2017*, 2017.
- [2] 芦川博人, 森岡幹, 小川厚徳, 岩田具治, 俵直弘, 小川哲司, 小林哲則, "複数人対話のための話者情報を用いたRNN言語モデル," *日本音響学会2016年秋季研究発表会講演論文集*, 3-Q-2, pp. 85-88, Sept. 2016.

連絡先

荒木 章子 (Shoko Araki) メディア情報研究部 信号処理研究グループ
E-mail: araki.shoko(at)lab.ntt.co.jp

22

それっぽくしゃべるとは

～発話リズム変換技術を用いた脳における音声情報処理の解明～

どんな研究

音声を聞く・話す際の人間の脳情報処理を解明する研究です。これまで、**発話リズム変換技術“それっぽくしゃべります”**を開発し、知覚実験により音声知覚における**発話リズム**の重要性を示してきましたが、発話リズムと脳情報処理の関係については分かっていませんでした。

どこが凄い

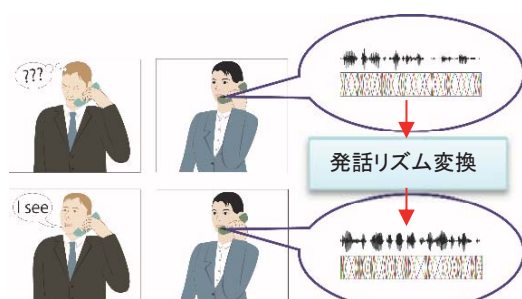
日本語母語話者の英語音声の発話リズムをネイティブっぽく変換し、英語母語話者に聞かせ、その際の脳活動をfMRI（機能的磁気共鳴画像法）により計測しました。その結果、発話の時間タイミング生成に関与するとされる**補足運動野の活動が変換により小さくなる**ことが分かりました。

めざす未来

発話リズムをネイティブに近づけるように変換することで、英語母語話者の脳の負担を軽減できる可能性が示されました。つまり、外国語音声の発話リズム変換を行うことで、**聞き取りやすく、脳にやさしい音声**を作り出すことができると期待されます。

発話リズム変換技術“それっぽくしゃべります”

※発話リズムの抽出にはNTD法[1]を使用



- 日本語母語話者と英語母語話者の発話リズムをそれぞれ抽出
- 発話リズムの入れ替えにより変換

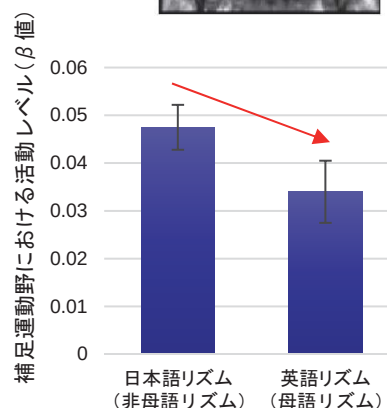
脳機能計測実験

日本語発話リズム、英語発話リズムの日本語母語話者の英語音声、英語母語話者が聞いている際の脳活動をfMRIで計測

結果

- 日本語発話リズムより、英語発話リズムの音声を聞く際に**補足運動野の活動が小さくなる**ことを発見
- 本来、補足運動野は発話における時間タイミングの生成に関与するとされていることから、**発話のシミュレーションにより、発話リズムの不自然さの検出/修正が行われている可能性**を示唆

- ネイティブっぽい発話リズムに変換したことで、脳への負荷が軽減された
- 音声知覚における発話リズムの重要性を脳の観点から示した



関連文献

- [1] S. Hiroya, "Non-negative temporal decomposition of speech parameters by multiplicative update rules," *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 10, pp. 2108-2117, 2013.
- [2] S. Hiroya, K. Jasmin, S. Evans, S. Krishnan, M. Ostarek, D. Boebinger, S.K. Scott, "Effects of speaking rhythm naturalness on the neural basis of speech perception," *Society for Neuroscience Abstracts*, 2015.

連絡先

廣谷 定男 (Sadao Hiroya) 人間情報研究部 感覚運動研究グループ
E-mail: hiroya.sadao(at)lab.ntt.co.jp

23

「自然な」音を調べて聴覚を理解する

～バイノーラル録音や空間残響に着目した聴覚再考～

どんな研究

日常生活で私たちが感じる音の豊かな質感を支える仕組みは謎に満ちており、実験室で単純な人工音を使うだけでは解明できません。この展示では、実験室外の自然な環境で私たちが経験している音に着目することで聴覚の理解を深める試みを紹介します。

どこが凄い

自然な頭の動きを含むバイノーラル録音や、壁や木々の反射由来の空間残響という見過ごされてきた特徴を持つデータベースを構築し、解析しています。空間残響に順応していく仕組みが世界で初めて明らかになったり、聴覚の定説が必ずしも成り立たないことが分かってきました。

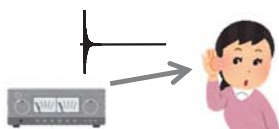
めざす未来

音の質感知覚を支える聴覚機構が理解できれば、単に音の内容を伝えるだけではなく、質感制御までもが可能になるでしょう。これにより、従来では考えられなかった心地良い音経験を実現できる可能性が見えてきます。

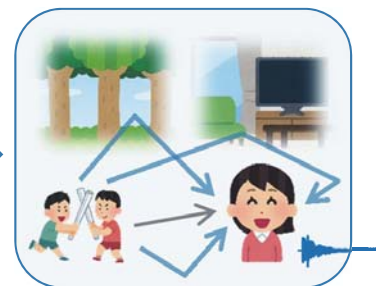
聴覚は自然音に最適か？

聴覚研究では無響室や単純な人工音が長く使われてきましたが、近年、私たちの聴覚は自然音の処理に最適化されているのではないかと考えが有力になってきています。しかし既存研究における「自然音」はまだ単純で、頭の動きや壁の反射の影響を受けて実際に私たちの両耳に届く音の質感には遠く及びません。私たちは両耳に届く音を詳しく調べ、その音が脳内でどのように処理されているか理解しようとしています。

人工音・片耳・直接音



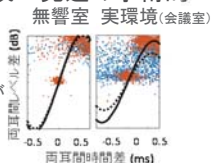
自然音・両耳・残響



その手がかり使えてる？ [1]

- 無響室での定説：音源定位では複数の手がかり（両耳間時間差・レベル差）を統合
- 自然な環境では手がかりの信頼性が低下することを確認
- 両耳機構の形成・発達の学術的議論に一石を投じる

自然環境における手がかり間の対応関係は、無響室（理想的環境）に比べて弱まる



この打撃音の材質は？ [2]

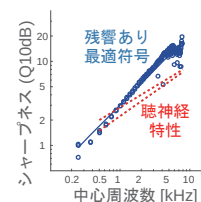
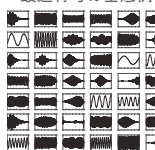
- 打撃音の減衰特性は材質判断の手がかりとして知られる
- 室内残響は減衰特性を変調 → 材質知覚にどう影響する？
- 室内音響に順応して残響の影響を低減する機構の存在を発見



最適化は教師なし学習？ [3]

- 定説：聴神経特性は自然音の最適符号と一致（教師なし学習）
- 残響のある現実的な音については、一致しないことを発見
- 単純な教師なし学習とは異なる

最適符号の基底例



自然音の知覚は、定説よりも複雑な最適化過程で実現されていることを解明

関連文献

- [1] S. Furukawa, "Natural Combinations of Interaural Time and Level Differences in Realistic Auditory Scenes," in *Proc. 39th Annual Midwinter Meeting of the Association for Research in Otolaryngology*, 2016.
- [2] T. Koumura, S. Furukawa, "Effect of Reverberation and Its Presentation Context on Material Perception Based on Impact Sounds," in *Proc. 40th Annual Midwinter Meeting of the Association for Research in Otolaryngology*, 2017.
- [3] 寺島裕貴, 古川茂人, "自然残響の考慮による聴覚音声スเปース符号化再考," 第31回人工知能学会全国大会, 2017.

連絡先

寺島 裕貴 (Hiroki Terashima) 人間情報研究部 感覚共鳴研究グループ
E-mail: terashima.hiroki(at)lab.ntt.co.jp

24

無意識の目の動きからココロを読む

～微小眼球運動から潜在的な心を解読する技術～

どんな研究

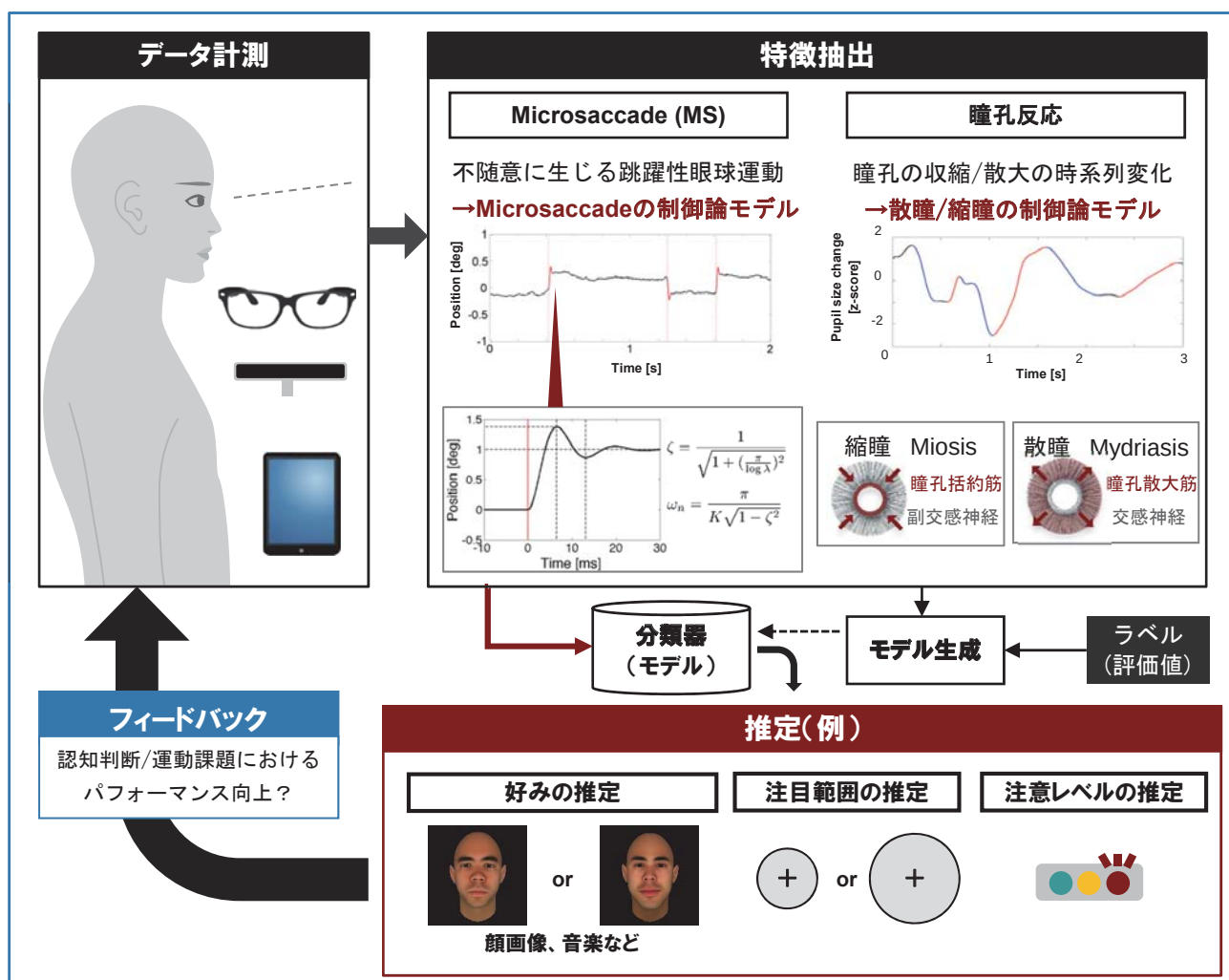
無意識に生じる**眼球の動きから、人の認知/知覚内容を推定する**研究です。マイクロサッカド（無意識に生じる微小な眼球運動）や瞳孔径変化の波形を詳細に解析し、眼球計測データから新たな特徴量を抽出することで、快不快や注意状態などの推定が一定精度で可能になりました。

どこが凄い

眼球の動きを表現する独自のモデルを考案することで、**眼球データに隠れる潜在的な情報を抽出可能**になりました。眼球計測は、脳を直接計測する手法（脳波やfMRIなど）と比較して、計測の安定性やユーザへの拘束性の面で優位性をもち、日常生活での応用が期待されます。

めざす未来

コンピュータでしか認識できない人間の潜在的な心の動きを読み取ることで、「**人間を超える**」感情認識能力をもったAIの実現を目指しています。また、眼球から推定されたメンタル状態をフィードバックすることで、個人のメンタルコントロールに利用することも考えられます。



関連文献

- [1] 古川茂人, 米家惇, Hsin-I LIAO, 柏野牧夫, “眼から読み取る心の動き ～Heart-Touching AIのキー技術～,” *NTT技術ジャーナル*, Vol. 28, No. 2, pp. 22-25, 2016.
- [2] M. Yoneya, S. Furukawa, M. Kashino, “Potential Use of Microsaccade in Personal Identification,” in *Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015.

連絡先

米家 惇 (Makoto Yoneya) 人間情報研究部 感覚共鳴研究グループ
E-mail: yoneya.makoto(at)lab.ntt.co.jp

25

触り方と触り心地の深い関係

～手の動きから触り心地を推定する～

どんな研究

物体に触っている人が感じている触り心地を推定するにはどうしたら良いでしょうか。触る可能性のある全ての物体の物理特性を計測しておくのは非常に大変です。本展示では、触るときの手や目の動きに着目し、**触り方による触り心地の推定可能性を検討**しました。

どこが凄い

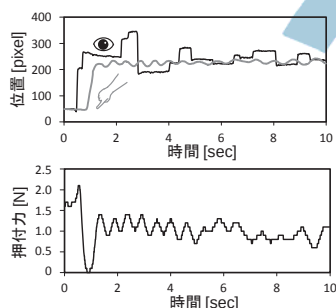
心理物理実験と多変量解析を組み合わせ、**触っているときの手や目の動きによって触った後の触り心地の評価を説明**できること、またなでる手の動きと大きな目の動きの組み合わせによって**硬さ、粗さ、粘つき、温かさの触り心地を区別して説明**できる可能性を明らかにしました。

めざす未来

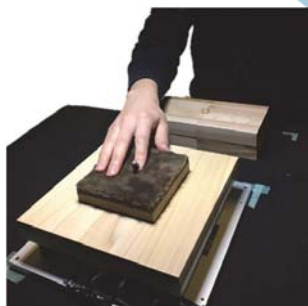
身の回りのものの複雑な物理特性を計測しなくても、**触っている人の手の動きさえ計測すればどのような触り心地を感じているかを推定**できることが期待されます。また、**触知覚メカニズム解明**のための仮説や、手に取り付ける**触覚提示装置**の設計指針につながることも期待されます。

触り方と触り心地を測る

①触っているときの手や目の動き（なぞりや押し込み）や目の動きを計測

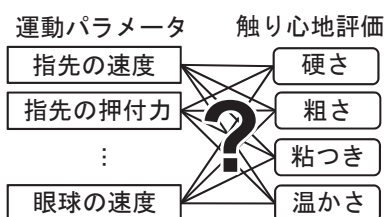


②手の動きや目の動きから運動パラメータを抽出



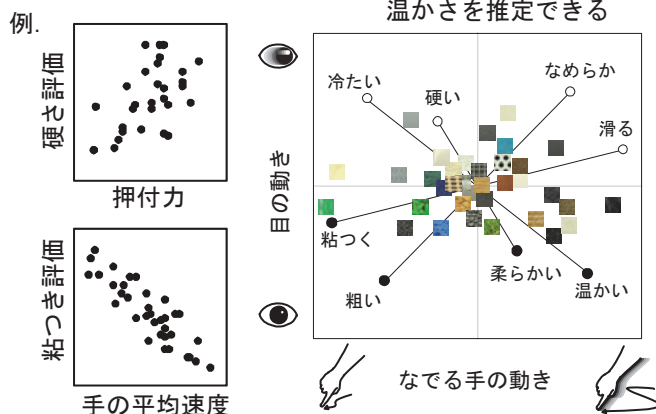
温かくて
柔らかい

③触ったあとに評価された硬さ、粗さ、粘つき、温かさとの関係を統計解析



触り方と触り心地の相関

手と目の動きさえみれば硬さ、粗さ、粘つき、温かさを推定できる



期待される応用例

身体運動を測るだけでどんな触り心地を感じているのかを推定できれば、プロダクトなどの触り心地の客観的な評価に応用可能



関連文献

- [1] T. Yokosaka, S. Kuroki, J. Watanabe, S. Nishida, "Linkage between free exploratory movements and subjective tactile ratings," *IEEE Transactions on Haptics*, (in press), DOI: 10.1109/TOH.2016.2613055.

連絡先

横坂 拓巳 (Takumi Yokosaka) 人間情報研究部 感覚表現研究グループ
E-mail: yokosaka.takumi(at)lab.ntt.co.jp

26

災害時の安否確認通話をつながりやすく

～通話時間の自発的短縮を促す発信規制法～

どんな研究

音声通話は災害時に最も重要な通信手段の一つですが、思うように繋がらない問題があります。発信可能時間帯を規制することで従来よりもつながり易くと同時に、その情報を提示することで人の行動を変化させ需要そのもの（通話時間）を減少させます。

どこが凄い

制御の対象を電話網だけでなく人間の行動にまで拡大した点と強制無しでユーザの自発的行動を促す仕組みを作った点です。強制的な規制は反発を招き輻輳緩和という観点では逆効果になりがちですがユーザ自らの意思による行動のためストレスも少なく輻輳緩和に寄与します。

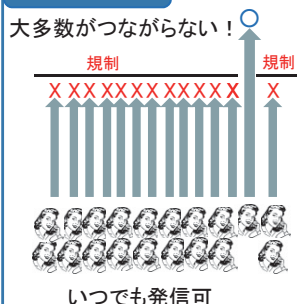
めざす未来

人間の行動を利用することで災害時に従来よりも繋がりがやすい状況を作り出すことを目指しています。電話以外の手段への誘導にも研究対象を広げ、早期の安否確認実現だけでなく、通話サービスを人命救助や復旧のために、より多く割り当てることを目指しています。

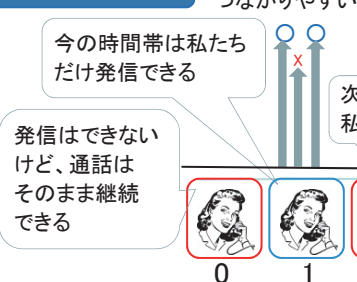
提案手法（整理券法）

- 電話番号の末尾によりユーザのグループ分けを行います。（各グループの人数はほぼ均等です [1]）
- グループごとに発信可能時間帯を設けて、ある時間帯にはあるグループのみ発信可能とします。
- 電話はいつでも受けられます。
- 通話時間の制限はありません。

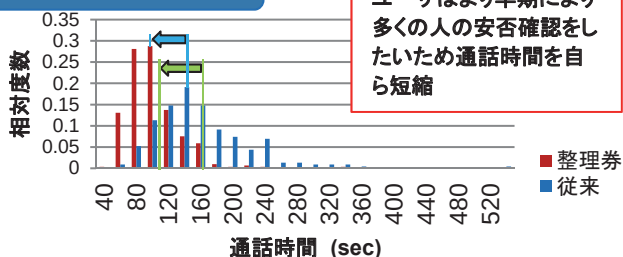
従来法



整理券法



通話時間減少



ユーザはより早期により多くの人の安否確認をしたいため通話時間を自ら短縮

全体負荷

$$0.1 (\text{人数}) \times 1.33 (\text{一人当たり単位時間発信数比})$$

$$= 0.133 > 1/10$$

$$\text{呼量} = \text{一人当たり単位時間発信数} \times \text{平均通話時間}$$

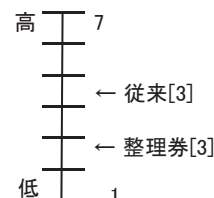
$$= 0.133 \times 0.691$$

$$= 0.092 < 1/10$$

一人当たり単位時間発信数はやや上昇。(1→1.33)

しかし、全体負荷(呼量)は1/分割数よりも削減！

発信者ストレス



関連文献

- [1] D. Satoh, H. Kawano, Y. Chiba, "Effect of Load-balancing against Disaster Congestion with Actual Subscriber Extension Telephone Numbers," *IEICE Transactions*, E98-A(8), pp. 1637-1646, 2015.
- [2] 佐藤大輔, 高野裕治, 須藤竜之介, 持田岳美, "災害輻輳自己抑制促進制御法の効果," *電子情報通信学会技術研究報告*, Vol. 115, No. 130, CQ2015-36, pp. 87-92, 2015.
- [3] 須藤竜之介, 佐藤大輔, 高野裕治, 持田岳美, "災害時電話安否確認における発信者のストレスと対策法の検討," *日本心理学会第79回大会発表論文集*, p. 292, 2015.

連絡先

佐藤 大輔 (Daisuke Satoh) ネットワーク基盤技術研究所
E-mail: satoh.daisuke(at)lab.ntt.co.jp

27

脳科学とICTでアスリート脳を解明

～Body-mind readingで推進するスポーツ脳科学～

どんな研究

スポーツで勝つには、頑健な「体」だけでなく、無自覚的（潜在的）な脳の情報処理に支えられている巧みな「技」や動じない「心」を備えることが必要です。私たちは、脳科学とICTを組み合わせ、アスリートの潜在脳機能の仕組みを解明することを目指しています。

どこが凄い

野球などの瞬時の対人インタラクションを伴うスポーツについて、実験室や実戦場面などの多角的なアプローチで計測を行い、アスリートの身体に現れる種々の生体情報から脳情報処理の仕組みを解読する「Body-mind reading」技術を駆使して、アスリート脳を理解します。

めざす未来

得られた脳科学的知見に基づいて、アスリートの競技レベルやタイプ、適正を評価・判別することに役立てます。また、五感特性をふまえたフィードバック技術と合わせて、フィジカル・メンタルの実践的なトレーニング方法を開発し、潜在脳機能の調整に活用します。



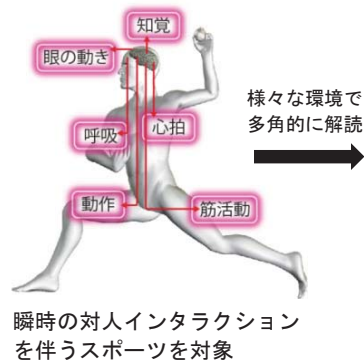
本研究のターゲット



脳科学とICTを組み合わせ、アスリートの“技”と“心”の脳情報処理の仕組みを解明

Body-mind reading

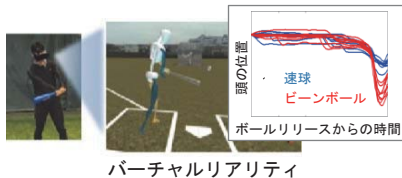
アスリートの身体に現れる様々な生体情報から脳情報処理の仕組みを解読



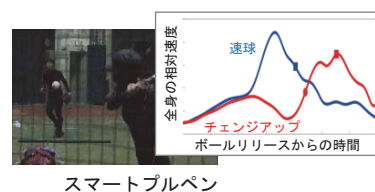
メンタル状態の推定



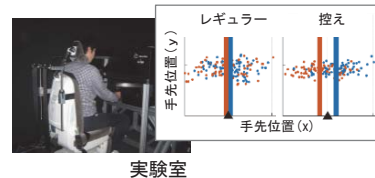
意思決定や情動の影響



瞬時の判断と運動調節



感覚運動特性の評価



関連文献

- [1] 木村聡貴, 持田岳美, 井尻哲也, 柏野牧夫, “小特集 ICTで強くなる・健康になる, 情報科学でスポーツパフォーマンス向上を支援する,” 電子情報通信学会 通信ソサイエティマガジン, Vol. 37, pp. 23-28, 2016.
- [2] 柏野牧夫, 持田岳美, 井尻哲也, 木村聡貴, “ウェアラブルセンサを用いたスポーツ中の心身状態の解読と調整-潜在脳機能に基づくスポーツ上達支援を目指して-,” バイオメカニクス研究, Vol. 19, No. 4, pp. 230-239, 2016.
- [3] スポーツ脳科学プロジェクト ウェブサイト, <http://sports-brain.ilab.ntt.co.jp/index.html>

連絡先

木村 聡貴 (Toshitaka Kimura) スポーツ脳科学プロジェクト
E-mail: kimura.toshitaka(at)lab.ntt.co.jp

28

私を「ぶるなび」で連れてって

～屋内・屋外・仮想空間での牽引感覚デバイス応用に向けて～

どんな研究

触力覚情報を振動モバイルデバイスで伝達可能にする方法について研究しています。引っ張られた感覚を生ずるデバイス「ぶるなび」は、**スマホシェル型やキュービック型など多彩な形態で力覚呈示を実現することができるようになり、様々なシーンでの応用可能性**を検討しています。

どこが凄い

非対称振動が皮膚に伝わる特性と人が力を感じる特性を研究し、装置重量が増しても多自由度の牽引感覚提示が可能な**新しい設計原理**を考案しました。この手法により世界初の**モバイル型6自由度力覚呈示デバイス**を使ったVRナビや、**スマホと一体化した経路ナビ**が実現できます。

めざす未来

人の触覚や力覚の知覚原理に基づいて触力覚情報を表現する端末装置を開発し、またその新たに生まれた手段を利用するシーンを提案していくことにより、**身体情報を伴ったより豊かなコミュニケーション**が実現できるようになることが期待されます。



T型（並進2自由度）



P型（並進1、回転1自由度）

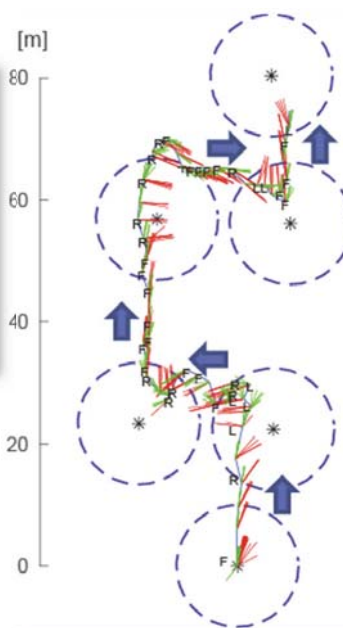
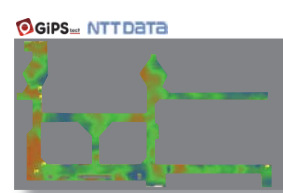
スマホケース型ぶるなび（Shell Force）

6自由度型ぶるなび
（Cubic force 6D）

VR力覚ナビ



地磁気を使った屋内触力覚歩行ナビ

屋外触力覚歩行ナビ
（GPS利用）

関連文献

- [1] 五味裕章, “モバイルガジェット「ぶるなび」～引っ張って方向を伝える触覚デバイスの開発と応用～,” 電場技術協会報, No.313, pp. 22-25, 2016.
- [2] 五味裕章, 雨宮智浩, 高椋慎也, 伊藤翔, “力覚呈示ガジェットの研究開発: 指でつまむと引っ張られる装置「ぶるなび3」,” 画像ラボ, Vol. 26(7), pp. 41-44: 日本工業出版. 2015.
- [3] 雨宮智浩, 五味裕章, “牽引方向知覚における能動的探索の有効性を活用した屋内歩行ナビゲーションシステムの開発,” 電子情報通信学会論文誌. D, 情報・システム, 97(2), pp. 260-269, 2014.

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
E-mail: gomi.hiroaki(at)lab.ntt.co.jp

29

AIと議論：都市と田舎、永住するならどっち？

～議論構造に基づき論拠提示や反論を行う議論対話システム～

どんな研究

人と対話を行う対話システムについて研究しています。有意義な対話を実現するには、多様な話題について話せることに加え、特定のテーマについて深く掘り下げて対話できることも重要な要素です。そうしたシステムを目指し、人と議論を行う議論対話システムを構築しました。

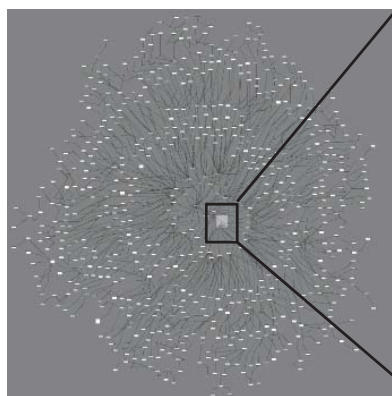
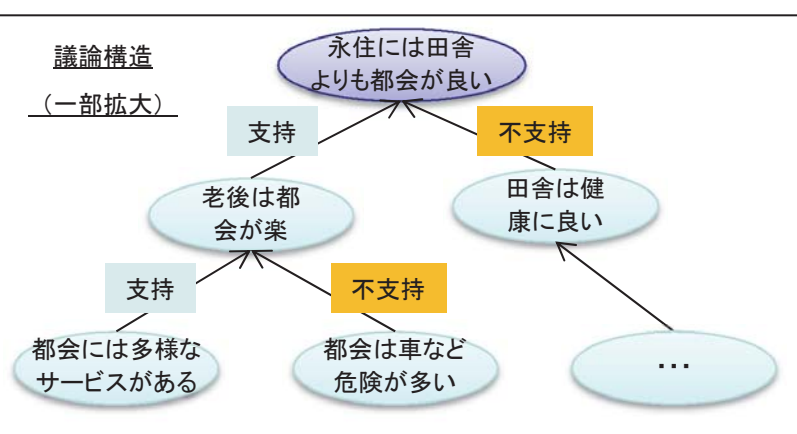
どこが凄い

議論を行うためには、システムは一貫した主張を持った上で、論拠の提示やユーザ主張への反論をする必要があります。私たちは、特定の話題について議論をするための議論構造を設計し、それに基づいて、議論の状況を適切に管理し、議論を進める議論対話技術を開発しました。

めざす未来

人と言葉でやりとりできる対話システムは、人とシステムをつなぐ重要なインターフェースです。議論を通じて人間とシステムがより理解し合うことで、人とシステムが助け合う社会の実現を目指しています。本技術は人間の議論スキルの向上や思考の活性化にも有用です。

議論構造（全体）

議論構造
(一部拡大)

不支持の知識を利用

でも、都会は車など危険が多いわよ

そうそう
老後は都会が楽だね

支持する知識を利用

僕は永住するなら都会かなあ

議論構造に基づく対話

※本展示は、大阪大学 石黒研究室との共同研究に基づく成果です。

関連文献

- [1] 杉山弘晃, 東中竜一郎, 目黒豊美, “気軽に雑談できるシステムの実現をめざして,” *NTT 技術ジャーナル*, Vol. 28, No. 9, pp. 16-20, 2016.
- [2] 杉山弘晃, 目黒豊美, 吉川 雄一郎, 大和 淳司, “複数ロボット間連携による対話破綻回避効果の分析,” *人工知能学会全国大会*, 2017.
- [3] 有本庸浩, 吉川雄一郎, 杉山弘晃, 目黒豊美, 大和淳司, 石黒浩, “対話焦点の不連続性を隠蔽する複数ロボット間連携の生成,” *言語・音声理解と対話処理研究会*, Vol. B5, No. 02, pp. 51-52, 2016.

連絡先

杉山 弘晃 (Hiroaki Sugiyama) 協創情報研究部 インタラクシオン対話研究グループ
E-mail: sugiyama.hiroaki(at)lab.ntt.co.jp

MEMO



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

意識されない人のこころ
と身体を読み解き、深層
心理・知性・本能を理解

Ambient-AI

森羅万象から
未来を予測

森羅万象(物・人・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が有機的につ
ながり成長し、社会シス
テム全体を最適化
通信ネットワークにAI技
術を適用

各展示と corevo を構成する4つのAIとの関連性

Agent-AI

- 17 英語の発音をネイティブの
ように綺麗に変換
- 18 現場の生音、高音質電話で
クリアにお届け
- 19 世界中の言葉を聞き分けます
- 20 あなたの声に即座にあわせて
音声認識
- 21 会話の流れを読んで
みんなの声をうまく認識

- 22 それっぽくしゃべるとは
- 23 「自然な」音を調べて
聴覚を理解する
- 24 無意識の目の動きから
ココロを読む
- 25 触り方と触り心地の深い関係
- 27 脳科学とICTで
アスリート脳を解明

Heart- Touching-AI

- 12 あなた好みに翻訳する
- 13 単語の意味や
関係を賢く学習
- 29 AIと議論：都市と田舎、
永住するならどっち？

- 28 私を「ぶるなび」で連れてって
- 15 2Dと3D、どちらも同時に
楽しめます
- 16 あなた専用のお手本映像で
上達支援
- 10 絵本はこどもの言葉の教科書
- 11 ノンネイティブのための
リスニング支援

- 06 グラフを用いてデータの
関係を素早く発見
- 14 照明光で目的の色だけを
くっきり見せる

Ambient-AI

- 01 車両によるスマート
シティセンシング
- 02 都市における人の活動
パターンとリズムを可視化
- 03 集団最適なナビを
リアルタイムに実現
- 04 プライバシーを保護しつつ
人の動線を抽出
- 05 時間変化するデータのための
安定した深層学習

- 09 敷き詰めかたをすべて見つけます
- 07 絶対に安全な共有鍵を作れるか
- 08 漏れなくデタラメです
- 26 災害時の安否確認通話を
つながりやすく

Network-AI



関連文献

※「corevo®」は日本電信電話株式会社の商標です。

[1] 「特集」NTTグループにおけるAIの取り組み,"NTT 技術ジャーナル, Vol. 28, No. 2, pp. 6-37, 2016.

(<http://www.ntt.co.jp/journal/1602/special.html>)

[2] 「特集」コラボレーションを通じて一緒に革新(レボリューション)を起こすNTTグループのAI技術「corevo™(コレボ)」,"ビジネスコミュニケーション, Vol. 53, No. 8, pp. 6-21, 2016.

[3] <http://www.ntt.co.jp/corevo/>

連絡先

山田 武士 (Takeshi Yamada) 企画担当
E-mail: [yamada.tak\(at\)lab.ntt.co.jp](mailto:yamada.tak(at)lab.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 2017 実行委員会**

(委 員 長) 平尾 努

(副委員長) 高橋 康博

(委 員) 松田 昌史／岸野 泰恵／高棕 慎也



研究展示

データと学習の科学

- 01 車両によるスマートシティセンシング 車載型センサーによる時空間都市イベント検知
- 02 都市における人の活動パターンとリズムを可視化 時空間統計に基づく都市データの解析
- 03 集団最適なナビをリアルタイムに実現 ベイズ的最適化に基づく効率的な誘導計画探索
- 04 プライバシーを保護しつつ人の動線を抽出 人流推定のための確率的行動モデル
- 05 時間変化するデータののための安定した深層学習 Gated recurrent unitにおける学習安定化技術
- 06 グラフを用いてデータの関係を素早く発見 L-グラフ構成法のための高速化アルゴリズム

計算と言語の科学

- 07 絶対に安全な共有鍵を作れるか 今あるハードウェアで量子暗号の実現を図る
- 08 漏れなくデータラメです 物理乱数源のランダムさを損わずに乱数ビットを安定供給
- 09 敷き詰めかたをすべて見つけます 厳密被覆問題の高速な全解列挙法
- 10 絵本はこどもの言葉の教科書 絵本コーパスと幼児の語彙獲得の関係を探索
- 11 ノンネイティブのためのリスニング支援 効果的な自動字幕表示法
- 12 あなた好みに翻訳する 接頭辞制約によるニューラル翻訳の制御
- 13 単語の意味や関係を賢く学習 開発者利便性が高い単語埋め込みベクトルの獲得法

メディアの科学

- 14 照明光で目的の色だけをくっきり見せる スペクトル最適化による色彩強調
- 15 2Dと3D、どちらも同時に楽しめます メガネなしでも二重に見えないステレオ映像生成法
- 16 あなた専用のお手本映像で上達支援 深層学習によるメディア生成の可能性
- 17 英語の発音をネイティブのように綺麗に変換 声道モデルと深層生成モデルを用いた統計的音声変換
- 18 現場の生音、高音質電話でクリアにお届け リアルタイム音源伝送向け低遅延・高音質符号化CLEAR
- 19 世界中の言葉を聞き分けます 多言語音声識別・認識技術
- 20 あなたの声に即座にあわせて音声認識 音声認識精度を高めるニューラルネットの高速適応技術
- 21 会話の流れを読んでみんなの声をうまく認識 頑健な音声処理を可能にする複数人発話区間推定技術

コミュニケーションと人間の科学

- 22 それっぽくしゃべるとは 発話リズム変換技術を用いた脳における音声情報処理の解明
- 23 「自然な」音を調べて聴覚を理解する バイノーラル録音や空間残響に着目した聴覚再考
- 24 無意識の目の動きからココロを読む 微小眼球運動から潜在的な心を解読する技術
- 25 触り方と触り心地の深い関係 手の動きから触り心地を推定する
- 26 災害時の安否確認通話をつながりやすく 通話時間の自発的短縮を促す発信規制法
- 27 脳科学とICTでアスリート脳を解明 Body-mind readingで推進するスポーツ脳科学
- 28 私を「ふるなび」で連れてって 屋内・屋外・仮想空間での牽引感覚デバイス応用に向けて
- 29 AIと議論：都市と田舎、永住するならどっち？ 議論構造に基づき論拠提示や反論を行う議論対話システム



NTT コミュニケーション科学基礎研究所

〒619-0237 京都府相楽郡精華町光台2-4
(けいはんな学研都市)
TEL: 0774-93-5020 FAX: 0774-93-5015

