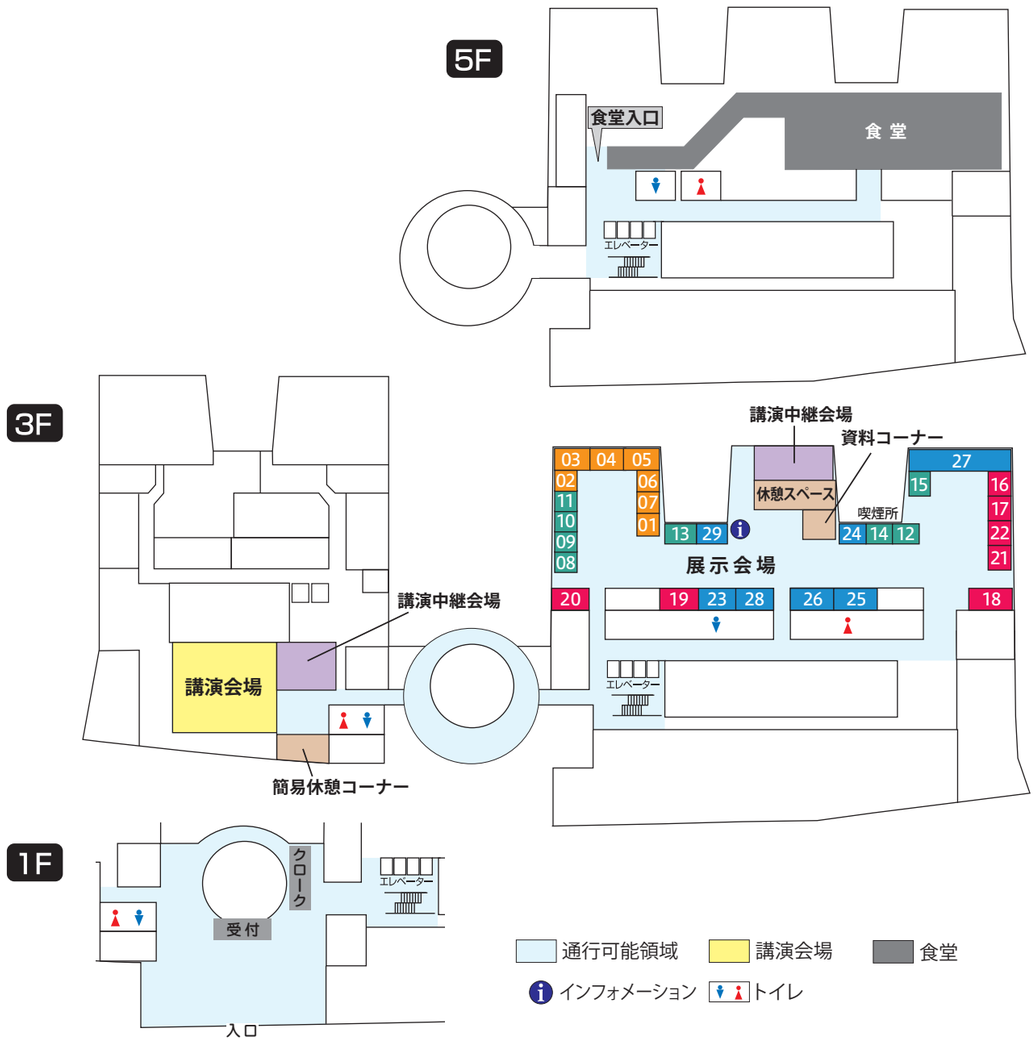


研究展示

データと学習の科学

- 01 膨大な情報の組合せから楽々学習　大域的最適性を保証する低ランク回帰学習技術：CFM
- 02 データに隠れた関係性を賢く抜き出します　無限バイクラスタリングによる特徴的部分行列の抽出
- 03 無限に広がるビッグデータの解析　無限階層の包含関係を持つ木構造の確率モデル
- 04 賢いナビで待たずに周遊　顧客満足度最大化のための動的巡回スケジューリング技術
- 05 どうやって来た？どこへ行く？　深層学習を利用した移動軌跡分析・予測技術
- 06 2020のトラヒックホットスポットはどこ？　ユーザ行動に基づく将来イベントトラヒック予測技術
- 07 高速かつ高精度な深層学習を実現します　時系列の勾配方向に着目した高効率な学習アルゴリズム

言語と計算の科学

- 08 訂正、圧縮、おてのもの　シャノン限界を達成できる多能な符号化法
- 09 巨大ネットワークに潜む構造を量子で探索　量子ウォーク技術による超グラフの高速探索
- 10 次世代ｗｅｂの安全性がより確かなものに　フォーマルメソッドによるQUIC・TLS1.3の検証
- 11 情報を守る真のデタラメをつくる　物理乱数 Gbps ストリーミングの実現
- 12 「ピタリエ」でこの子にピッタリの絵本探し　興味・発達段階を考慮したグラフ索引型類似探索技術
- 13 ニュースも会話も精読して翻訳します　書き言葉から話し言葉へ～自然言語解析と機械翻訳～
- 14 こどもがいつどんな語を覚えるのかを教えます　語彙チェックリストによる幼児語彙発達データベースの作成
- 15 ロボットが増えると話がはずむ　複数ロボットの協調動作による音声対話活性化

メディアの科学

- 16 飾り文字でも読み取れる　深層学習と系列デコーディングによる文字列認識
- 17 声の抑揚が別人に変わる　基本周波数パターン生成過程モデルに基づく韻律変換
- 18 実物鑑賞では得られない発見を得る　高精細マルチバンド画像撮影による美術品への情報付加
- 19 こんなにガヤガヤした場所でも聞き取ります　遠隔発話音声認識における雑音除去・深層学習技術の最前線
- 20 高圧縮符号化で携帯電話も放送も更に高音質に　低ビットレート符号化ＥＶＳと音響ロスレス符号化ＡＬＳ
- 21 変幻灯2.0：光をあてると踊り出す　画像検索を用いた動き情報の自動投影
- 22 乾いた風景を雨上がりに　物体表面を濡れているように見せる質感編集技術

コミュニケーションと人間の科学

- 23 英語は苦手？でもだいじょうぶ　音声認識と機械翻訳による非母語者にやさしい支援
- 24 ことばの発達、日本語と英語で何が違う？　言語間比較で見えてきた親の発話とこどもの語彙獲得の関係
- 25 離れていても目と目を合わせて会話が出来ます　キネティックアバタを用いた視線を交わせるテレプレゼンス
- 26 分かってきた耳の繊細な働き　内耳における振幅変調－周波数変調変換システム
- 27 手も心もひかれるモバイルガジェット　牽引力錯覚装置「ぶるなび」の進化と応用
- 28 「スポーツで勝てる脳」をつくる　脳科学に基づくスポーツ強化システム
- 29 「見守る」と「見守られる」をつなぎます　介護者記録用Webアプリ「みまもメイト」の開発と評価

NTT
NTT コミュニケーション科学基礎研究所
〒619-0237 京都府相楽郡精華町光台2-4
（けいはんな学研都市）
TEL: 0774-93-5020 FAX: 0774-93-5015

VEGETABLE OIL INK　VOC FREE T&K

オープンハウス2016

NTT

NTT コミュニケーション科学基礎研究所

オープンハウス 2016

6/2(木) 12:00～17:30

6/3(金) 9:30～16:00

入場料無料・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）

Innovative R&D by NTT

NTT コミュニケーション科学基礎研究所

「オープンハウス2016」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2016」へご来場頂き、心より御礼申し上げます。

ここ数年、“人工知能（AI）”に関する技術開発が急速に進み、世の中を動かす仕掛けが大きく変わろうとしています。このとき、将来に対する適切な投資を行うためには、人工知能を支える技術の根幹とその潜在的可能性を適確に把握しておくことが必要になります。

NTT コミュニケーション科学基礎研究所では、NTTグループが進める人工知能“corevo™（コレボ）”の核となることを目指し、２種類の基礎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。

オープンハウス2016では、「データと学習の科学」、「言語と計算の科学」、「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、ご来場の皆さまにとって有益で愉しいひとときとなれば幸いに存じます。

目　次

ごあいさつ ･･ 1
会場案内図／講演スケジュール ･･ 3
所長講演
　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 ･･････････････････････････ 4
招待講演
　わかりあえないことから ･･･ 5
研究講演
　音声認識が人間に近づく・人間を超える　ディープラーニングと信号処理の進化がもたらす音声認識のひろがり ･･････ 6
　きれいな音で伝えます。携帯電話も放送も。　携帯電話向け音声音響符号化ＥＶＳとハイレゾ放送向け音響ロスレス符号化ＡＬＳ ･･ 8
　視覚をだまし、質感を操作する　液体質感の科学的解明と変幻灯の開発 ･････････････････････････ 10
　雑談できるコンピュータを作る　オープンドメイン言語処理技術の進展と課題 ･････････････････････ 12
研究展示 15

データと学習の科学

01 ❙膨大な情報の組合せから楽々学習　大域的最適性を保証する低ランク回帰学習技術：CFM ･････････････ 16
02 ❙データに隠れた関係性を賢く抜き出します　無限バイクラスタリングによる特徴的部分行列の抽出 ････････ 17
03 ❙無限に広がるビッグデータの解析　無限階層の包含関係を持つ木構造の確率モデル ･･････････････････ 18
04 ❙賢いナビで待たずに周遊　顧客満足度最大化のための動的巡回スケジューリング技術 ･････････････････ 19
05 ❙どうやって来た？どこへ行く？　深層学習を利用した移動軌跡分析・予測技術 ･･･････････････････････ 20
06 ❙2020 のトラヒックホットスポットはどこ？　ユーザ行動に基づく将来イベントトラヒック予測技術 ･･･････ 21
07 ❙高速かつ高精度な深層学習を実現します　時系列の勾配方向に着目した高効率な学習アルゴリズム ･･････ 22

言語と計算の科学

08 ❙訂正、圧縮、おてのもの　シャノン限界を達成できる多能な符号化法 ･･････････････････････････ 23
09 ❙巨大ネットワークに潜む構造を量子で探索　量子ウォーク技術による超グラフの高速探索 ･･････････ 24
10 ❙次世代ｗｅｂの安全性がより確かなものに　フォーマルメソッドによる QUIC・TLS1.3 の検証 ･･････････ 25
11 ❙情報を守る真のデタラメをつくる　物理乱数 Gbps ストリーミングの実現 ･････････････････････････ 26
12 ❙「ピタリエ」でこの子にピッタリの絵本探し　興味・発達段階を考慮したグラフ索引型類似探索技術 ･････････ 27
13 ❙ニュースも会話も精読して翻訳します　書き言葉から話し言葉へ～自然言語解析と機械翻訳～ ･･････････ 28
14 ❙こどもがいつどんな語を覚えるのかを教えます　語彙チェックリストによる幼児語彙発達データベースの作成 ･･･ 29
15 ❙ロボットが増えると話がはずむ　複数ロボットの協調動作による音声対話活性化 ･････････････････････ 30

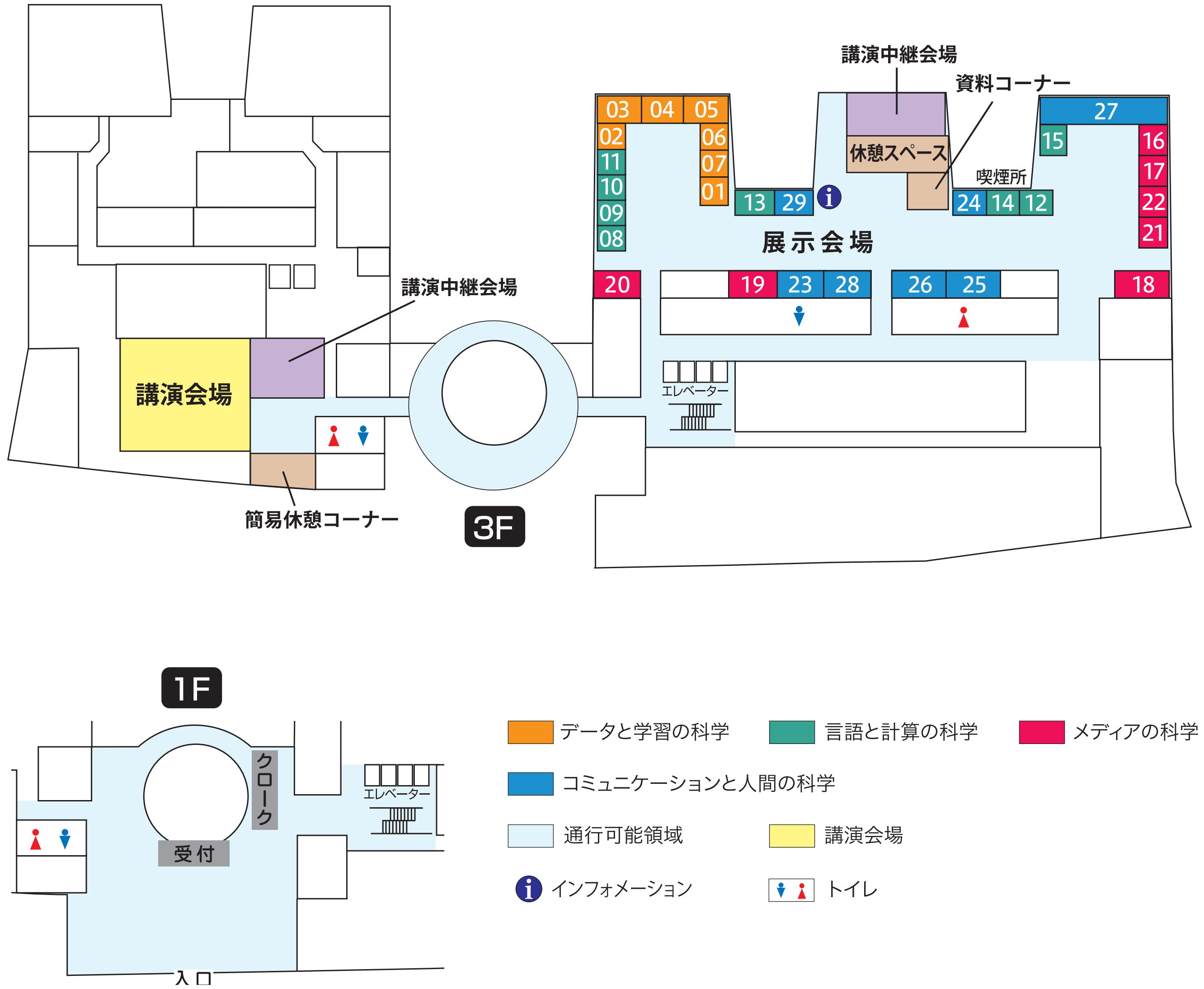
メディアの科学

16 ❙飾り文字でも読み取れる　深層学習と系列デコーディングによる文字列認識 ･････････････････････ 31
17 ❙声の抑揚が別人に変わる　基本周波数パターン生成過程モデルに基づく韻律変換 ･････････････････････ 32
18 ❙実物鑑賞では得られない発見を得る　高精細マルチバンド画像撮影による美術品への情報付加 ･･････････ 33
19 ❙こんなにガヤガヤした場所でも聞き取ります　遠隔発話音声認識における雑音除去・深層学習技術の最前線 ････ 34
20 ❙高圧縮符号化で携帯電話も放送も更に高音質に　低ビットレート符号化ＥＶＳと音響ロスレス符号化ＡＬＳ ･･･ 35
21 ❙変幻灯 2.0：光をあてると踊り出す　画像検索を用いた動き情報の自動投影 ･････････････････････ 36
22 ❙乾いた風景を雨上がりに　物体表面を濡れているように見せる質感編集技術 ･･･････････････････････ 37

コミュニケーションと人間の科学

23 ❙英語は苦手？でもだいじょうぶ　音声認識と機械翻訳による非母語話者にやさしい支援 ････････････････ 38
24 ❙ことばの発達、日本語と英語で何が違う？　言語間比較で見えてきた親の発話とこどもの語彙獲得の関係 ･････ 39
25 ❙離れていても目と目を合わせて会話が出来ます　キネティックアバタを用いた視線を交わせるテレプレゼンス ･･･ 40
26 ❙分かってきた耳の繊細な働き　内耳における振幅変調－周波数変調変換システム ･･････････････････ 41
27 ❙手も心もひかれるモバイルガジェット　牽引力錯覚装置「ぶるなび」の進化と応用 ･･････････････････ 42
28 ❙「スポーツで勝てる脳」をつくる　脳科学に基づくスポーツ強化システム ･･･････････････････････････ 43
29 ❙「見守る」と「見守られる」をつなぎます　介護者記録用 Web アプリ「みまもメイト」の開発と評価 ････････ 44

会場案内図／講演スケジュール



講演スケジュール

6月2日〈木〉

所長講演
(13:20～13:50)
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 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)
わかりあえないことから
劇作家・演出家　平田オリザ

研究講演
(15:30～16:10)
音声認識が人間に近づく・人間を超える
ディープラーニングと信号処理の進化がもたらす音声認識のひろがり
メディア情報研究部　吉岡　拓也

6月3日〈金〉

研究講演
(11:00～11:40)
きれいな音で伝えます。携帯電話も放送も。
携帯電話向け音声音響符号化EVSとハイレゾ放送向け音響ロスレス符号化ALS
守谷特別研究室　鎌本　優

(13:00～13:40)
視覚をだまし、質感を操作する
液体質感の科学的解明と変幻灯の開発
人間情報研究部　河邉　隆寛

(13:50～14:30)
雑談できるコンピュータを作る
オープンドメイン言語処理技術の進展と課題
協創情報研究部/メディアインテリジェンス研究所 音声言語メディアプロジェクト　東中　竜一郎

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

From information transmission to the mutual understandings

Paradigm shift in the age of data

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

近年、“人工知能（AI）”に対する関心が急速に高まる中で、過度の期待と醒めた批判とが錯綜しています。しかしながら少なくとも、人工知能に関する新しい発見や技術によって、世の中を動かす仕掛けが大きく変わろうとしていることは間違いありません。そして、将来に対する適切な投資を行うためには、人工知能を支える技術の根幹とその潜在的可能性を適確に把握しておくことが必要になります。

NTTのAI技術“corevo™（コレボ）”

NTTでは、NTTグループのAI技術に対して“corevo™（コレボ）”というブランド名を付けました。さまざまなプレイヤーとのcollaborationを通じて、ともにrevolutionを起こしたいという思いを込めています。“corevo™”は人の知性や思考を模倣するものではなく、人の活動を支援し、人の能力を補完、引き出すAI技術であり、社会課題の克服と産業競争力の強化を目指しています。

知能の市場経済性

実世界にある森羅万象（人・もの・環境）あらゆるものからデータが取得され、解読、探索（探究）、デザイン（実装）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、バイタル、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、知能がさまざまな形で分解、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。

コミュニケーションの本質を探る

実用領域に達した音声認識、機械翻訳から人間の心と身体を読み解くセンシング技術に至るまで、多様な形でAI技術が実生活に入り込んで来る中で、“コミュニケーション”の本質について再考すべき時期が来ています。物理学的な時間の流れは変わらない、生物としての人間の寿命も大きくは変わらない、こうした堅固な制約の中にありながら人間を取り巻く環境は大きく変貌しています。人と人、人と機械とのコミュニケーションの本質について深くそして広く追究し、未来の設計図を描いていきたいと考えています。

【関連文献】

[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）
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（[特集]未来への羅針盤としてのコミュニケーション技術）, *NTT技術ジャーナル*, Vol. 27, No. 9, pp. 8-9, 2015.（http://www.ntt.co.jp/journal/1509/special.html）
前田英作：基礎研究は「時代」とともに在り（[特集]コミュニケーション科学の新展開）, *NTT技術ジャーナル*, Vol. 26, No. 9, pp. 12-15, 2014.（http://www.ntt.co.jp/journal/1409/special.html）
[特集]こころ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, *NTT技術ジャーナル*, Vol. 25, No. 9, pp. 12-13, 2013.（http://www.ntt.co.jp/journal/1309/special.html）

わかりあえないことから

Standing together on no common ground

劇作家・演出家
平田オリザ
Oriza Hirata
http://www.seinendan.org/hirata-oriza

© T.Aoki

（プロフィール）

1962年　東京生まれ。国際基督教大学教養学部卒業。
1995年　『東京ノート』で第39回岸田國士戯曲賞受賞。
1998年　『月の岬』で第5回読売演劇大賞優秀演出家賞、最優秀作品賞受賞。
2002年　『上野動物園再々々襲撃』(脚本・構成・演出)で第9回読売演劇大賞優秀作品賞受賞。
2002年　『芸術立国論』(集英社新書)で、AICT評論家賞受賞。
2003年　『その河をこえて、五月』(2002年日韓国民交流記念事業)で、第2回朝日舞台芸術賞グランプリ受賞。
2006年　モンブラン国際文化賞受賞。
2011年　フランス国文化省より芸術文化勲章シュヴァリエ受勲。

東京藝術大学COI研究推進機構 特任教授、大阪大学コミュニケーションデザイン・センター客員教授、四国学院大学客員教授・学長特別補佐、京都文教大学客員教授、（公財）舞台芸術財団演劇人会議理事長、埼玉県富士見市民文化会館キラリ☆ふじみマネージャー、日本劇作家協会副会長、日本演劇学会理事、（財）地域創造理事、豊岡市文化政策担当参与。

音声認識が人間に近づく・人間を超える

ディープラーニングと信号処理の進化がもたらす音声認識のひろがり

Avenues toward super-human speech recognizers

Advances in deep learning and signal processing that are making speech recognition leap forward

メディア情報研究部
吉岡　拓也
Takuya Yoshioka

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　研究主任。京都大学博士(情報学)。IEEE ASRU2015 Best Paper Award Honorable Mention、日本音響学会独創研究奨励賞 板倉記念等受賞。音声認識、マイクロホンアレイの研究に従事。

音声認識のカンブリア紀

私たちが話した声をコンピュータが自動的に認識する―それを可能にする音声認識技術の実用化が今、急速にすすんでいます。AppleのSiriやGoogleのGoogle Now、NTTドコモのしゃべってコンシェルに代表される音声アシスタント、各種検索エンジンに搭載されている音声検索、カーナビやゲーム端末の音声操作、官公庁における会議録作成支援。音声認識技術は今や多様な場面で利用されており、まさにカンブリア紀を迎えています。その背景には、近年の音声認識精度の劇的な改善があります。本講演では、この革新を生み出したディープラーニングと呼ばれる技術と、音声認識の応用領域をさらに拡大する上で重要となる信号処理技術について我々の取り組みを交えて紹介し、音声認識の未来を展望します。

ディープラーニングの台頭

ディープラーニングは階層構造をもつニューラルネットワークを用いた機械学習技術です。機械学習では、大量のデータ(学習データと呼ばれます)をコンピュータに与えて、そのなかから有用な規則を自動的に抽出させます。音声認識ならマイクロホンから得られる人の声の波形から文字列に変換する規則を自動的に学習します。こうして得られた規則を用いることで、音声を文字列に自動的に変換できます。

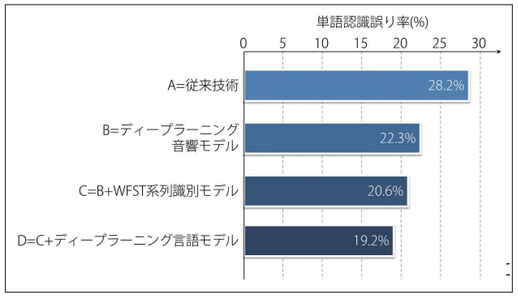


図1：ディープラーニングの発展がもたらした音声認識制度の改善(NTTでの実験結果)

ディープラーニングの音声認識への導入は2009年頃にはじまりました。2011年にはSwitchboardと呼ばれる本格的な音声認識タスクで、従来技術と比べて音声認識誤りを約3割削減できたという報告がなされました[1]。この大成功をきっかけに、多くの研究機関がディープラーニングを用いた音声認識技術の研究にいっせいに取り組み、系列識別学習、話者適応化、畳み込みニューラルネットワーク、長短期記憶ネットワーク、コネクショニスト時系列分類、大規模並列学習など、重要技術が次々と発表されました。その結果、前述のSwitchboardタスクでは、かつては考えられなかった認識誤り率8%が報告され[2]、人間の認識誤り率といわれる4%に急速に近づいています。データによっては人間並みの認識精度も報告されています。

「どこでも使える音声認識」を実現する信号処理技術

音声認識の前処理である信号処理技術もまた、音声認識の世界に劇的な変化をもたらしています。現在の音声認識技術は、声の波形から文字列への変換規則を学習データから抽出するため、学習データの性質が実際の利用シーンと異なる場合はどうしても認識精度が劣化してしまいます。このことは、マイクロホンで録音された音声に不要な雑音や残響が含まれている場合に特に問題になります。信号処理技術は、あらかじめ不要な音を取り除くことで、こうした性能劣化を防ぎます。

特に、複数のマイクロホンを用いて雑音を除去するマイクロホンアレイ信号処理技術が実際に使われるようになってきています。これによって、例えばMicrosoftのKinectやAmazonのEchoでは、手元にマイクロホンがなくても声を使った操作が可能になっています。さらに、我々の研究グループが昨年参加してベストスコアを達成したCHiME3と呼ばれる音声認識評価国際イベントでは、タブレットに取り付けられた複数のマイクを組み合わせることで騒がしい環境でも音声認識精度を劇的に向上させられることが示されています[3]。このような技術が進展すれば、音声認識がどこでも使えるようになり、応用範囲は飛躍的に広がるでしょう。

人間に近づく・人間を超える音声認識

音声認識技術は今後ますます発展し、近い将来に多くのタスクで人間並み・もしくは人間以上に正しく認識できるようになるでしょう。インターネットでの買い物、天気の確認、ビデオの視聴、ロボット掃除器の動作。こういったことが室内に端末を一台置いておくだけでできるようになり、私たちの日常生活は一層便利になると期待されます。街に出れば、ヒューマノイドが案内を務めたり、音声応答システムが券売窓口やドライブスルーで動作したりしているでしょう。我々は、そのような未来を目指してこれまで以上に研究をすすめると同時に、社内外の機関と連携して研究成果を世の中にタイムリーに提供できるよう取り組みをすすめていきます。

図2：社会に溶け込む音声認識技術

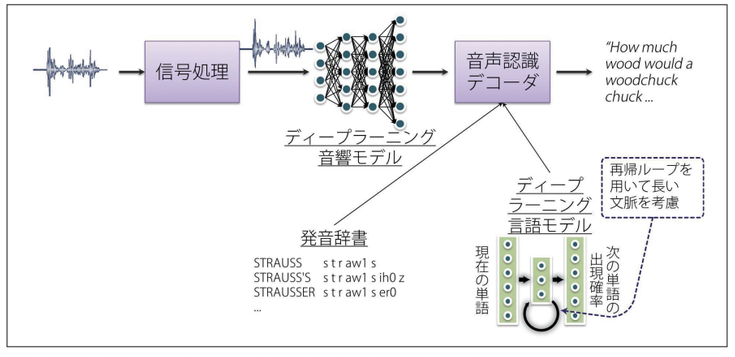


図3：信号処理とディープラーニングを統合した音声認識システム

【関連文献】

[1] F. Seide, G. Li, D. Yu, "Conversational speech transcription using context-dependent deep neural networks," in *Proc. 12th Annual Conference of the International Speech Communication Association (Interspeech)*, pp. 437-440, 2012.
[2] G. Saon, H.-K. J. Kuo, S. Rennie, M. Picheny, "The IBM 2015 English conversational telephone speech recognition system," arXiv:1505.05899, 2015.
[3] 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. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 436-443, 2015.

きれいな音で伝えます。携帯電話も放送も。

携帯電話向け音声音響符号化ＥＶＳとハイレゾ放送向け音響ロスレス符号化ＡＬＳ

How to transmit high-quality sound via network

Reseach and deployment of speech codec EVS for VoLTE and lossless audio codec ALS for broadcasting

守谷特別研究室
鎌本　優
Yutaka Kamamoto

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。情報処理学会論文賞、日本音響学会粟屋潔学術奨励賞、電気通信普及財団テレコムシステム技術奨励賞・技術学生賞、各受賞。電波産業会（ARIB）音声符号化作業班　委員。IEC MT 61937－10 Convenor。超臨場感コミュニケーション産学官フォーラム（URCF）超臨場感音響技術に関する調査WG幹事。

携帯電話向け音声音響符号化EVS

移動通信システムの国際標準規格を策定する3GPP (The 3rd Generation Partnership Project) では将来の携帯電話やVoLTE (Voice over Long-Term Evolution) 向け音声音響統合符号化方式として、Enhanced Voice Services (EVS) を新たに策定しました [1]。EVSコーデックは音声向けの時間領域の符号化方式と音楽向けの周波数領域の符号化を組み合わせたものであり、これまで実現できなかった、背景雑音や背景音楽を含む音声信号や音楽信号そのものを高音質のまま低遅延で伝送することが可能となっています [2]。

EVSコーデックは8 kHz (Narrow-band; NB), 16 kHz (Wide-band; WB), 32 kHz (Super-WB; SWB), 48 kHz (Full-band; FB)というように幅広いサンプリングレートをサポートし、1フレーム20 msごとに伝送し、アルゴリズム遅延32 msという通話可能な遅延で処理を行います。また、VoLTE向けということもあり、ビット列の誤り耐性ではなくパケット消失を隠蔽する方策がとられ、パケットベースの伝送に適した方式となっています。さらに、対応するビットレートもNBは5.9 kbpsから24.4 kbps、WBは5.9 kbpsから128 kbps、SWBは9.6 kbpsから128 kbps、FBは16.4 kbpsから128 kbpsというように幅広く対応し、任意の音声帯域・任意のビットレートにフレームごとに切り替えることができます。そして、現在普及しつつあるAMR-WB (Adaptive Multi-Rate WB)との互換性も具備するため、シームレスなコーデックの置き換えが可能です。

各社の提案する技術を正当に評価するために、標準化の過程において、様々な条件・音源・言語での大規模な主観品質評価実験が第三者機関によって行われ、実験結果が評価レポートに示されており、従来の音声符号化方式・音響符号化方式よりも高性能であることが確認されています。技術を提案した12社が協力し、統一方式を改善することにより、従来の符号化方式よりも高性能のEVSを策定することに貢献しました。

ハイレゾ放送向け音響ロスレス符号化ALS

放送の音声は、放送局内では48 kHz, 24 bitのいわゆるハイレゾ品質で制作されています。ところが、放送を家庭まで届けるために使われる電波は有限な資源であり節約しなければならないため、仕方なく歪のある圧縮を行い皆様の元へ届けられていました。これから始まる4K/8K放送では、映像で使われる情報量が大きくなるため、それに伴い、音声に割り当てられる情報量も大きくなることが期待されます。音声を伝送するための符号化方式の一つとして、日本の放送規格であるARIB STD-B32では音響ロスレス符号化のMPEG-4 ALSが使えるようになりました。ALSにより放送局のスタジオで制作された音声をそのまま家庭まで届けることができるようになります。電波で届く放送に先駆けてIPTVでの放送コンテンツ伝送においても、ALSが使われる可能性が高まっています。

ALSは圧縮しても元に戻る国際標準技術で [3]、NTTも標準化に貢献してきました [4]。圧縮率は入力信号に依存しますが、おおよそ30%～70%の大きさのビットレートへ圧縮することができます。NTTでは標準策定後も周辺分野の標準化に取り組み、より使いやすいツールに限定したMPEG-4 ALS Simple Profile規格、日本の放送規格であるARIB STD-B32、ディジタルアンプへの接続規格であるIEC 61937-10の標準化やそれらの改訂作業を進めてきました。今後、ALSが普及することにより、皆様が意識せずとも放送の音質が向上することが期待できます。

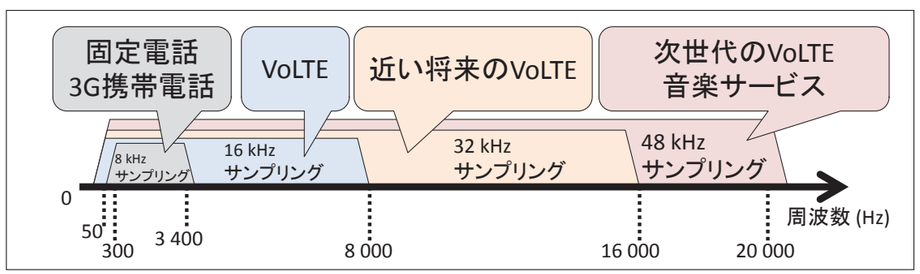


図1：電話帯域の高音質化

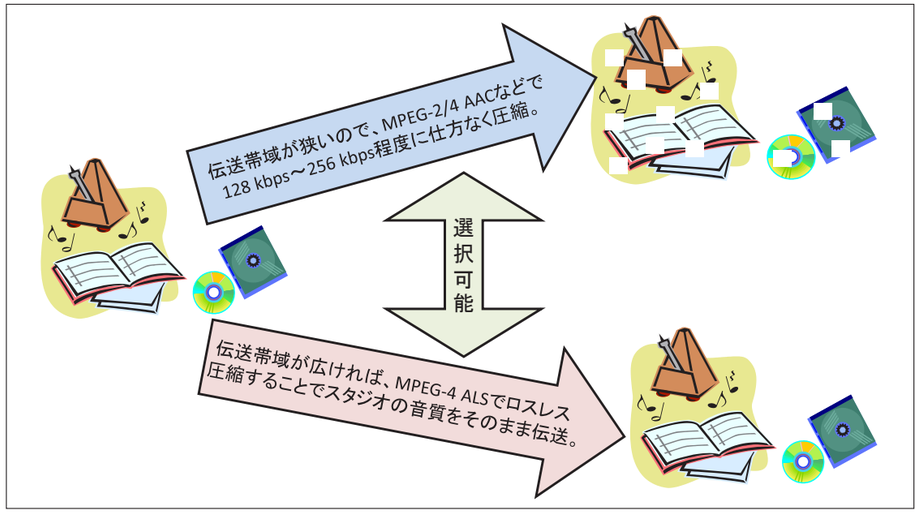


図2：放送音声の高音質化

【関連文献】

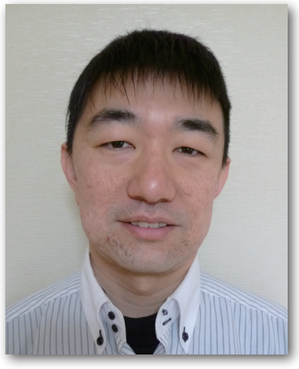
[1] 3GPP TS 26.441, “Codec for Enhanced Voice Services (EVS); General overview,”
[2] M. Dietz, M. Multrus, V. Eksler, V. Malenovsky, E. Norvell, H. Pobloth, L. Miao, Z. Wang, L. Laaksonen, A. Vasilache, Y. Kamamoto, K. Kikuiri, S. Ragot, J. Faure, H. Ehara, V. Rajendran, V. Atti, H. Sung, E. Oh, H. Yuan, C. Zhu,“Overview of the EVS codec architecture,” in *Proc. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015),* pp. 5698 - 5702, 2015.
[3] ISO/IEC 14496-3, “Information technology - Coding of audio-visual objects - Part 3: Audio,” fourth edition, 2009.
[4] T. Liebchen, T. Moriya, N. Harada, Y. Kamamoto, Y. A. Reznik, “The MPEG-4 Audio Lossless Coding (ALS) Standard -Technology and Applications,” in *Proc. AES 119th Convention, 2005.*

視覚をだまし、質感を操作する

液体質感の科学的解明と変幻灯の開発

Hack human vision, control material appearance

Liquid material impression and Hen-GenTou



人間情報研究部
河邉　隆寛
Takahiro Kawabe

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。主な研究テーマは視覚質感科学および視覚質感操作。1979年に長崎県で生まれる。2005年に九州大学大学院人間環境学府で博士号（心理学）取得。2011年にNTT コミュニケーション基礎研究所にリサーチスペシャリストとして入社、2015年より現職。2013年に日本心理学会第8回国際賞奨励賞を受賞。

はじめに

他者をだまして自らが利益を得るのはよくありませんが、脳をだましてユーザーに利益を与えることはあながち悪いこととは言えません。特に、視覚をごまかす、いわゆる錯覚を利用した情報提示は、視覚情報を送る側にも、受ける側にも恩恵を与えることができます。本講演では、画像情報処理や光投影技術によってユーザーに錯覚を与えることで、簡易に、しかしながら、劇的に、視覚質感印象を操作することに成功した研究例を紹介いたします.

質感の科学

布は柔らかく温かい、金属は固く冷たい、というように、我々は対象の材質やその特性印象、すなわち質感を感じ取ることができます。この質感を認識する脳の仕組みの探求は知覚科学で最近取り組まれ始めたホットなテーマです。一般に材質の物性は複雑であり、それが故に材質が人間にもたらす感覚情報も複雑になります。したがって、質感の科学では、材質の物性、感覚情報、および知覚内容の3者を密に検討する必要が出てきます。

画像操作による液体質感印象の制御

我々は最近、材質の中でもその挙動が大変複雑である液体を人間がどのように視覚的に認識しているか、という問題に取り組み始めました。透明な液体は空気と異なる屈折率を持っているため、液体内で反射して液体外に出る光は屈折します。その際、液体内やその底面にある対象の画像情報は変形します。

我々は、この画像の変形と液体を認識することとの間の関係を調べました[1]。その結果、画像の変形が時間的に連続して生じるときに、あたかも透明な液体が画像上を流れているように感じることを実証しました。変形が時間的に連続しない場合、透明な液体が流れている印象は生じませんでした。更に研究を深く進めると、連続した画像変形の中でも特定の時空間変形周波数帯域が、透明な液体の流れ印象を喚起する重要な要素であることをつきとめました。

そしてこの知見を利用することで、流れる透明液体の印象をつくりだす技術を開発しました。つまり、静止した画像を上記の時空間変形周波数帯域を利用して変形させることで、本来その画像には存在しなかった流れる透明液体の印象を作り出すことに成功したのです。

動きの質感を操作する光投影技術「変幻灯」の開発

上記の液体質感印象を表示する技術は、あくまで画像中の話でした。ここで我々は、画像という仮想的な世界の産物から離れ、現実世界の物体、つまり実物体へ透明液体印象を表示することが可能かどうかを検討しはじめました。

ここで光投影技術（いわゆるプロジェクションマッピング）に注目しました。光投影技術は、ビデオプロジェクタからの投影画像を利用し、実物体が実際とは異なる視覚印象をもつように投影画像を制御する技術です。既存の光投影技術は、本来実物体の表面にある模様や色味を打ち消し、新しい視覚情報を映し出す技術です。そのため既存技術では、実物体に備わっている模様や色味を残したまま、透明液体印象のみを追加するということは不可能でした。

我々はここで、錯覚を利用することで実物体に透明液体印象を与える方法を思いつきました。動きを見る際に働く視覚情報処理では、色よりも明るさが重要視されることが知られています。この特性を利用し、白黒の動きパタンを実対象へ投影することにより、色やテクスチャを含む実物体の外観がそのまま動いたり変形したりするように見えるのではないかと考えました。

この考えを基に開発したのが変幻灯です[2]。図1は変幻灯の代表的なシステム例です。まずカメラで投影対象のグレースケール画像を取得します。その画像を画像処理によって変形し、変形前後の差分画像を得ます。この差分画像は明暗のみで構成されています。その差分画像の明るさレベル調整しつつ投影対象へ投影することで、あたかもその対象が変形しているかのような印象を得ることができます。これは、投影した差分画像の明暗パタンが動きの錯覚を生じさせるためであると考えています。変形のパタンを工夫することで、液体印象、熱気印象はもとより、肖像画の表情を変えることも可能です。

まとめ

錯覚を利用することにより、豊かな視覚表現を簡易に、そして時には経済的に、ユーザーへ提供することができます。視覚を科学的に明らかにする、そして視覚を映像技術に利用する、その2つの目的を同時に満たすような研究開発に今後も取り組みたいと思います。

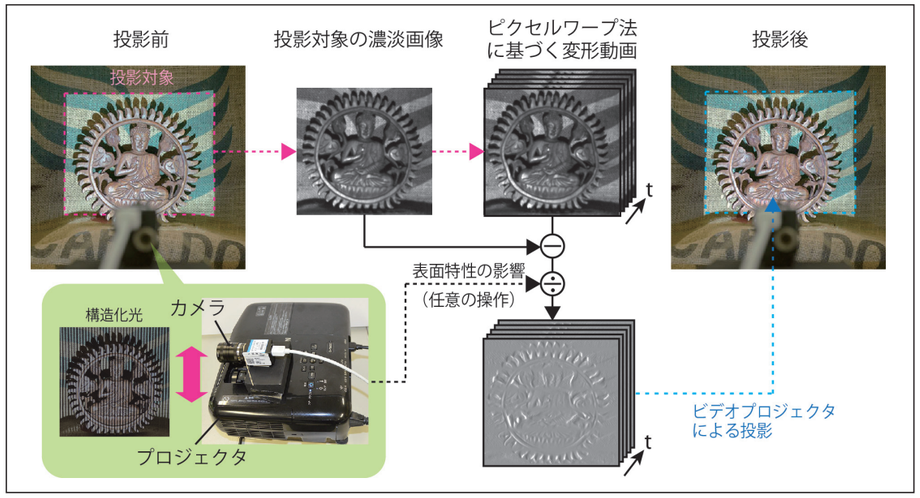


図1：変幻灯のシステム構成例

【関連文献】

[1] T. Kawabe, K. Maruya, S. Nishida, “Perceptual transparency from image deformation”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 112, No. 33, pp. E4620-E4627, 2015.
[2] T. Kawabe, T. Fukiage, M. Sawayama, S. Nishida, “Deformation lamps: A projection technique to make static objects perceptually dynamic”, *ACM Transaction on Applied Perception*, Vol. 13, No. 2, Article 10, pp. 1-17, 2016.

雑談できるコンピュータを作る

オープンドメイン言語処理技術の進展と課題

Creating a computer that can perform casual conversation

Advances and challenges in open domain natural language processing

協創情報研究部
メディアインテリジェンス研究所 音声言語メディアプロジェクト
東中　竜一郎
Ryuichirou Higashinaka

プロフィール

NTT コミュニケーション科学基礎研究所 協創情報研究部 主任研究員。NTT メディアインテリジェンス研究所 音声言語メディアプロジェクト 主任研究員（兼務）。1999年慶應義塾大学環境情報学部卒業、2001年同大学大学院政策・メディア研究科修士課程、2008年博士課程修了。2001年日本電信電話株式会社入社。質問応答システム・音声対話システムの研究開発に従事。博士（学術）。言語処理学会、人工知能学会、情報処理学会、電子情報通信学会、各会員。

オープンドメイン言語処理技術の必要性

スマートフォン上のみならず、家の中でも、街中でも対話システムが使われるようになってきました。家の中では、パーソナルロボットがいて、家の外では、ロボットが案内や販売員をしていたりします。また、今後の技術の進展に伴って、人工知能が様々な役割を担い、多くの分野で人間とやり取りをするようになるでしょう。さまざまな状況や多くの分野で、システムに入力されうる内容を、事前にすべて列挙することは困難です。そこで、任意の言葉による入力を適切に処理する「オープンドメイン言語処理技術」が必要になってきています。

雑談対話システム

ユーザの発話が幅広い内容を含む端的なシーンは日常会話、いわゆる雑談です。そこで、われわれは、雑談対話システムの研究に取り組んでいます[1]。雑談を研究することは、オープンドメイン言語処理技術を推し進めるだけでなく、人間に寄り添うインタフェースの実現にもつながります。

国立国語研究所の調査によれば、日本人の会話のうち、約60%が雑談です。雑談によって、人間は社会関係を構築していると言われており、人間の輪に入ることのできるコンピュータを実現するためには雑談は必須の機能です。さらに、人間はやり取りをする相手がコンピュータであっても、まるで相手が人間であるかのように振る舞うことが知られています[2]。つまり、人間はシステムが相手でもどうしても雑談をしてしまうのです。雑談機能を持つ必要がないと思われるシステムでも、雑談機能がないと使い勝手の悪いシステムになってしまいます。われわれは任意の入力を受け付けつつ、幅広い話題に応答できる雑談対話システムの構築を進めてきました。そして、ようやく任意の話題について、ユーザと複数回のやり取りができるまでに進展させることができました。

雑談対話システムの基本的な構成を図1に示します。発話理解部は、ユーザ発話から現在の話題、ユーザの発話意図、ユーザの発話内容を抽出します。対話管理部は、文脈を踏まえ、システムの発話意図を生成します。発話生成部は、現在の話題とシステムの発話意図を反映した発話を生成します。どのような発話にも対応できるようにするため、発話意図の表現には汎用的な類型（質問、自己開示、確認など）を用い、発話内容の表現には、述語項構造と呼ばれる文の意味を表す汎用的な表現を用いています。幅広い話題に対応するため、発話生成部にはインターネット上の大規模テキストデータから抽出した発話・知識データベースを用いています。さらに、システム発話にキャラクタとしての一貫性を持たせるため、表現変換技術も備えています。現状のシステムは、人手による対話ルールと併用し、図2に示すような自然な雑談を実現することができます。

オープンドメイン言語処理技術の実用化

オープンドメイン言語処理技術の実用化の例として、「しゃべってコンシェル」における知識Q&Aと「自然対話プラットフォーム」を紹介します。どちらも、NTT研究所が技術支援を行い、NTTドコモが開発したものです。

「しゃべってコンシェル」は音声エージェントサービスです。音声により、さまざまな機能が実行できます。その中に、知識Q&Aという機能があり、任意の質問に対して回答を得ることができます。ユーザがスマートフォンに向けて行う質問は多様であり、まさにオープンドメイン言語処理技術が必要となる場面です。われわれは、幅広い質問を網羅するような汎用的な質問分類を用いることで的確な応答を実現しました[3]。また、自然対話プラットフォームは対話システム構築のためのプラットフォームです。ここには、われわれの雑談対話システムの成果が多く盛り込まれています。自然対話プラットフォームは、店頭支援からコミュニケーショントイに至るまで、オープンドメインと呼ぶにふさわしい、多くの場面での対話を実現しています。

今後の展開

オープンドメイン言語処理技術の実現にはまだ多くの課題が残されています。雑談対話システムも、一定の日常会話ができるようになったとはいえ、質問に適切に答えられなかったり、現在の話題に関連しないことを話してしまうなど、対話が破綻してしまう状況が多々あります[4]。基礎研究を進めつつ、実用化によりユーザに実際に使ってもらいながら改良を続けていくことが必要だと考えています。

システム：こんにちは
ユーザ：コーヒーが飲みたいな
システム：コーヒー？コーヒー好きですか？
ユーザ：大好きですよ
システム：ここの豆はおいしいとか、耳寄り情報があったら知りたいです
ユーザ：スタバにはよく行きます
システム：スタバ？コーヒー専門店のコーヒーはやっぱり美味しい。
ユーザ：そうですよね
システム：コーヒーは香りが大事ですよね！
ユーザ：紅茶も好きですよ
システム：紅茶？お気に入りの紅茶の種類は何ですか
ユーザ：ダージリンかな
システム：ダージリン？ダージリンは香りが強いようです
ユーザ：その香りがよいのです
システム：香りがいいのは嬉しいですね。
ユーザ：そうそう
システム：だよね。なるほど。

図2：雑談対話システムによる対話例

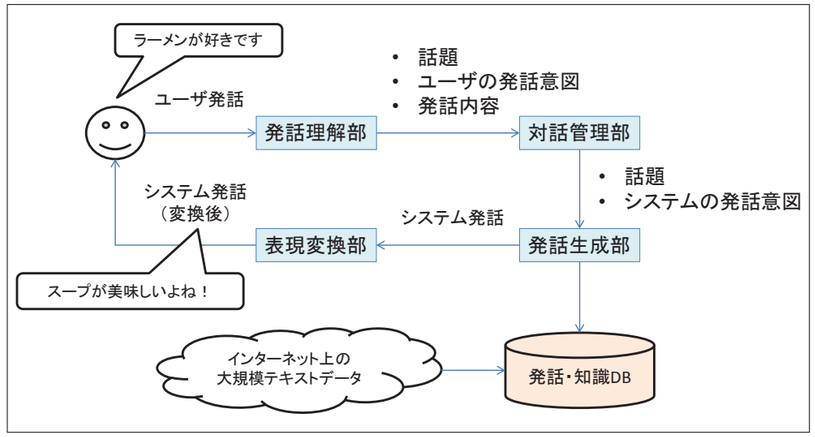


図1：雑談対話システムの基本的な構成

【関連文献】

[1] R. Higashinaka, K. Imamura, T. Meguro, C. Miyazaki, N. Kobayashi, H. Sugiyama, T. Hirano, T. Makino, Y. Matsuo, “Towards an open-domain conversational system fully based on natural language processing,” In *Proc. 25th International Conference on Computational Linguistics (COLING)*, pp. 928-939, 2014.
[2] B. Reeves, C. Nass, “How people treat computers, television, and new media like real people and places,” CSLI Publications and Cambridge university press, 1996.
[3] 東中竜一郎, 貞光九月, 内田渉, 吉村健, “しゃべってコンシェルにおける質問応答技術,” *NTT技術ジャーナル*, Vol. 25, No. 2, pp. 56-59, 2013.
[4] 東中竜一郎, 船越孝太郎, 荒木雅弘, 塚原裕史, 小林優佳, 水上雅博, “テキストチャットを用いた雑談対話コーパスの構築と対話破綻の分析,” *自然言語処理*, Vol. 23, No. 1, pp. 59-86, 2016.

NTT コミュニケーション科学基礎研究所　オープンハウス2016

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2016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」のそれぞれの研究を紹介する合計29展示をご覧頂きます。

データと学習の科学

主に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！
- 過去の事例から学習し，将来を予測したい！
- データ解析技術の活用で何ができるか事例を知りたい！

言語と計算の科学

計算機の限界に挑戦しこれまで困難であったことを可能にする研究をご紹介します

- 人と人，人とコンピュータの言葉の壁をなくしたい！
- 安全・安心に通信したい！
- 量子コンピュータを実現したい！

メディアの科学

音声，画像等のメディア情報を人間の特性に合わせて高度・高速に処理する研究をご紹介します

- どんな状況下でも音声/画像を明瞭に認識したい！
- 「もの」の質感の本質を究め，変化させたい！
- 高音質で相手に届けたい！

コミュニケーションと人間の科学

人間の他者とのコミュニケーションや情報処理の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！
- 脳の仕組みを解明し、新しいインターフェイスをデザインしたい！

CFM: low-rank regression with global optimality guarantees

01 膨大な情報の組合せから楽々学習

～大域的最適性を保証する低ランク回帰学習技術：CFM～

どんな研究

複数の特徴量から目的値を予測する回帰問題を汎用的に、かつ高精度に解く研究です。提案法「Convex Factorization Machines」(CFM)は、データを表す特徴から予測に有用な組合せを見付けて高精度なモデルを作ります。提案法は、価格推定、推薦システム、遺伝子解析などの様々な応用例に適用できます。

どこが凄い

提案法は、膨大な数の特徴の組合せを低ランク制約によって効率的に扱うことができます。さらに、初期値に依存せず大域的最適解を保証することができます。また、学習データに現れない未観測特徴の組合せの重要度（重み）も推定できるため、ユーザに新規の商品を推薦するなどの応用が可能です。

目指す未来

本技術を展開させ、より高次の特徴の組合せも考慮できれば、さらに予測精度の高いモデルを構築することができます。応用先として、推薦システムのほか、ゲノムデータから病気の有無や治療の効果を予測することで、その病気に影響する遺伝子の組合せの発見や新薬の開発技術などに利用可能です。

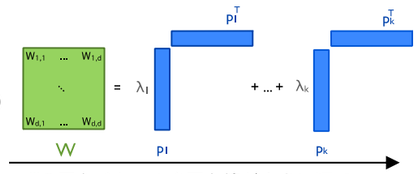
提案：重要な特徴の組合せを見つけて
高精度な回帰モデルを学習する
Convex Factorization Machines (CFM)

- 汎用的（様々な予測問題に適用可能）
- モデルのパラメタ学習が容易（解が初期値に非依存）
- 膨大な数の組合せを扱える（高次元データに適用可能）
- 未観測な特徴組合せの重要度を推定
  （新たな知識を発見可能）

$$\hat{y} = \underbrace{w^{\mathrm{T}}x}_{\text{1次重み}} + \underbrace{x^{\mathrm{T}}Wx}_{\text{2次重み}}$$

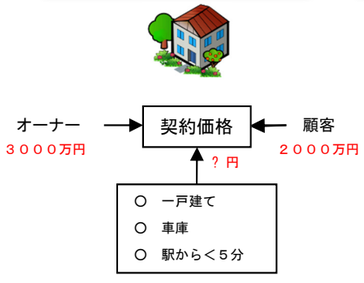
目的変数　特徴量　膨大な行列！

CFMは低ランクの制約によってWの大きさを抑制します



Wの固有ベクトルを固有値が大きい順で
逐次的に求めていくため、効率的に計算できます

応用例１
住宅の契約価格を予測



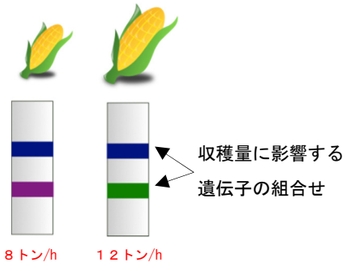
契約価格を決定する要因
（条件の組合せ）を発見

応用例２
ユーザが好む書籍を推薦

Alice	☆	☆☆☆
Bob	☆☆☆	?
Charlie	?	☆☆

ユーザと書籍の組合せから
評価点を予測することで
推薦システムを実現

応用例３
遺伝子から穀物収穫量を予測



トウモロコシの収穫量を
遺伝子の組合せから予測！

【関連文献】
[1] M. Blondel, A. Fujino, N. Ueda, “Convex factorization machines,” in *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2015.

【 連 絡 先 】
ブロンデル マチュー (Mathieu Blondel)　上田特別研究室
E-mail：mathieu.blondel(at)lab.ntt.co.jp

Infinite plaid models for infinite bi-clustering

02 データに隠れた関係性を賢く抜き出します

～無限バイクラスタリングによる特徴的部分行列の抽出～

どんな研究

行列の形式で与えられた関係データの中から、周りと大きく違う値をもつ部分（バイクラスタ）だけを自動的に見つける研究です。そのような部分は「もっとも売れる顧客層と商品群」や「特定の治験で効果のあった遺伝子群」などのように、有益な発見となる可能性が高い部分です。

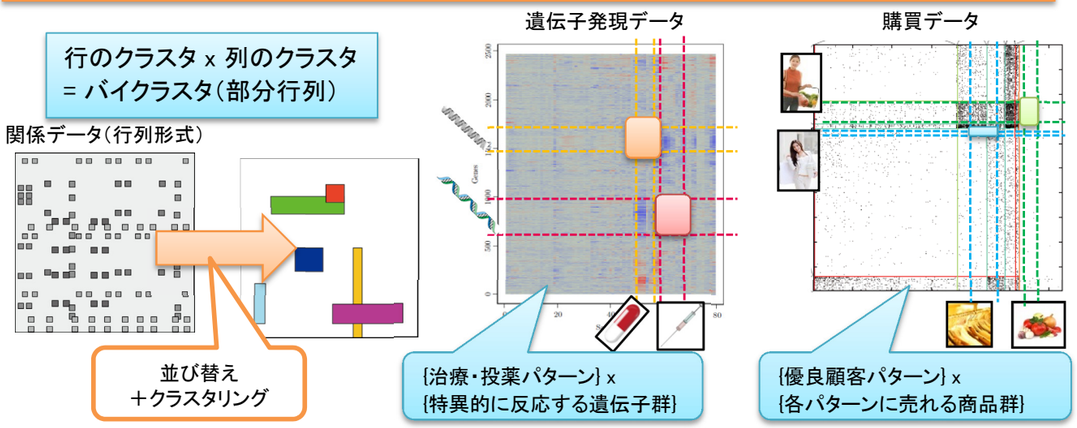
どこが凄い

これまでの技術では、バイクラスタの数を事前に指定する必要があります。しかし、事前に正確な数を知ることは困難です。そこで、任意数のバイクラスタを表現可能な機械学習モデルを用いて、各データに対して適切なバイクラスタ数を自動的に推定して抽出できます。

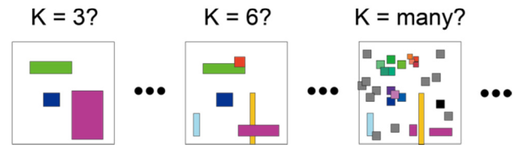
目指す未来

これまでの関係データ解析は、データを表現する隠れたパターンを発見するだけで、「どれが興味深いパターンか」は人が判断する必要がありました。今後はこの研究のように、「興味深いと思われるパターン」だけを発見する、より使いやすい関係データ解析技術を目指します。

バイクラスタリング: 与えられた関係データから
周囲と異なる特徴的な値をもつバイクラスタ(部分行列)を発見する

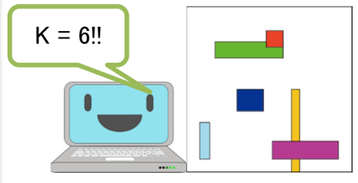


従来法: 隠れたバイクラスタの数を
指定しないといけない☹



提案法: 任意のデータからバイクラスタを個数を含めて推定し、発見する機械学習モデル

☺ 最適バイクラスタ数をデータから自動推定
→ 簡単にバイクラスタリング！



【関連文献】
[1] K. Ishiguro, I. Sato, M. Nakano, A. Kimura, N. Ueda, “Infinite plaid models for infinite bi-clustering,’’ in *Proc. 30th AAAI Conference on Artificial Intelligence (AAAI)*, 2016.
[2] 石黒勝彦，佐藤一誠，中野允裕，木村昭悟，上田修功，“Infinite bi-clusteringのための無限Plaid Model,’’ *第18回情報論的学習理論ワークショップ* (*IBIS2015*), 2015.

【 連 絡 先 】
石黒 勝彦 (Katsuhiko Ishiguro)　協創情報研究部 知能創発環境研究グループ
E-mail：ishiguro.katsuhiko(at)lab.ntt.co.jp

R-tree structure grows through stochastic process

03 無限に広がるビッグデータの解析

～無限階層の包含関係を持つ木構造の確率モデル～

どんな研究

本発表では無限のデータへの対策を考慮した機械学習技術を紹介します。従来のビッグデータ解析では、文字通り大きいながらも有限のサイズのデータを解析することを暗に仮定していましたが、本研究では無限のデータに対しても破綻することなく解析可能な機械学習の新しい技術を提案します。

どこが凄い

世界で初めてR木構造を無限データ解析へ応用する方法を創出しました。R木とは行列の包含関係の木構造を表すデータ構造であり、無限サイズの行列を根とするR木の全ての場合の数は無限通りとなります。これを計算機上で表現することの出来る確率モデルの構成に成功しました。

目指す未来

ビッグデータ解析はデータサイズの増加と解析アルゴリズムの更新のイタチごっこになる危険を抱え続けています。これを克服するため、あらゆるビッグデータ解析手法を無限データ解析へ発展すべく、その要素技術の拡大・確立していくことを目指しています。

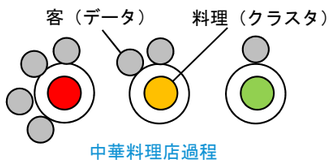
従来の無限データ解析

無限の場合の数を表現しうる確率モデル（アルゴリズム上の工夫によって無限の処理に陥らないモデル）を使い仮想的に計算機上で無限データを扱えるようにする工夫により実現。

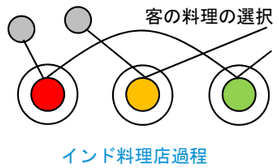
これまでに発見された無限データのための確率モデル：

- 無限のデータをクラスタリングしたい：中華料理店過程
- 無限のデータから主要構成要素を抽出したい：インド料理店過程
- 無限の行列から長方形分割模様のクラスタを抽出したい：モンドリアン過程、長方形分割過程

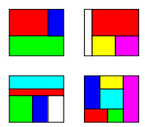
課題：無限データを扱える確率モデルは極めて少ない！

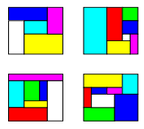


中華料理店過程



インド料理店過程

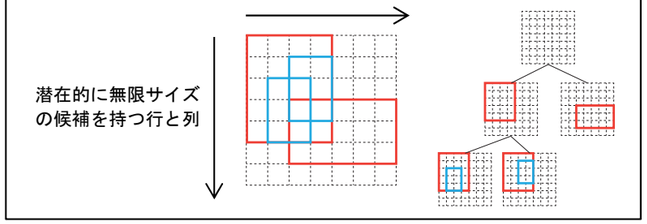

モンドリアン過程


長方形分割過程[2]

提案するモデル:R木過程

無限の行列から隠れた部分行列の無限木構造を抽出したい。

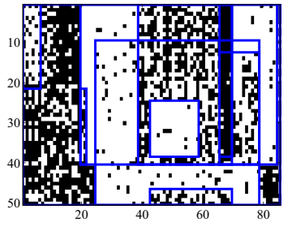
R木（各ノードに行列が対応した木構造で親ノード行列はその子ノード行列を包含するもの）と呼ばれるデータ構造に基づく、無限データのための確率モデルを初めて提案[1]。従来の無限データ解析では扱えなかった部分行列の包含関係が木構造を用いて表現可能に。



応用

関係データ解析：入力行列の行と列の並び替えとR木模様の中で、なるべく一つの領域に同じ性質のデータが集まるものを探す技術。

（例）協調フィルタリング、ネットワークの異常検知



【関連文献】

[1] 中野允裕，武小萌，森稔，木村昭悟，柏野邦夫, “R木過程,” *情報論的機械学習ワークショップ*, 2015.
[2] M. Nakano, K. Ishiguro, A. Kimura, T. Yamada, N. Ueda, “Rectangular tiling process,” in *Proc. International Conference on Machine Learning (ICML)*, 2014.

【連絡先】

中野 允裕 (Masahiro Nakano)　メディア情報研究部 メディア認識研究グループ
E-mail：nakano.masahiro(at)lab.ntt.co.jp

Dynamic migration scheduling for greater visitor satisfaction

04 賢いナビで待たずに周遊

～顧客満足度最大化のための動的巡回スケジューリング技術～

どんな研究

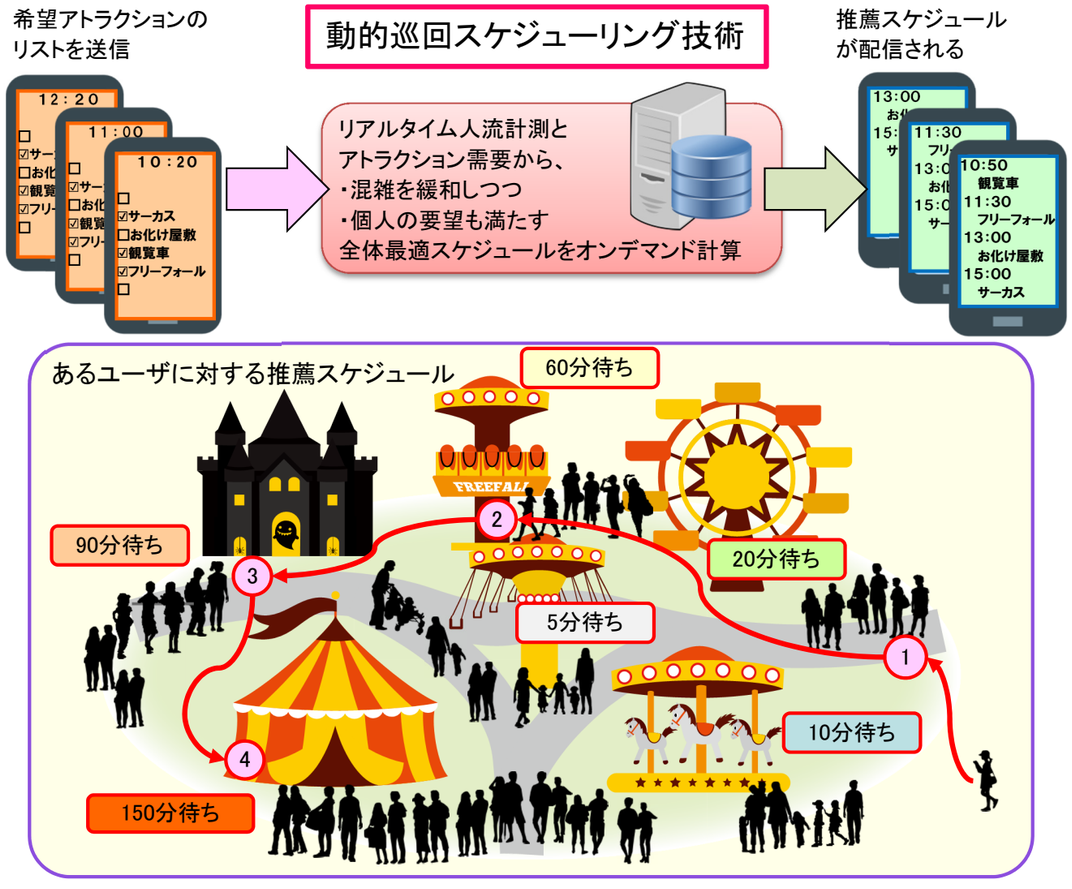
大規模イベント会場などにおいて、リアルタイムに観測された人流・交通流および、希望するアトラクションや目的地検索ログから、近未来の混雑状況やリソース需要を予測し、全てのお客様が快適に周遊できるように巡回スケジュールに基づき動的にナビゲーションする技術です。

どこが凄い

イベント会場や各アトラクションでの混雑状況を予め正確に見積ることは困難です。本技術は、リアルタイム観測に基づき、未来の混雑を予測し、個々人ではなく集団にとって最適な誘導を動的に実現します。これにより有事が生じた場合でも適応的な誘導が可能となります。

目指す未来

観光地、イベント会場、空港、商業施設などで、お客様が安心、安全に移動できる社会インフラを実現し、さらに、本技術を多方面に展開し、人、交通、モノの流れを、いつでもどこでも快適に誘導、制御できる世界の実現を目指します。



【関連文献】

[1] 上田修功, 納谷太, “集団を快適な巡回スケジューリングでナビゲーションします,” NTT R&Dフォーラム2016.
[2] G. Hernandez-Maskivker, G. Ryan, M. Blazey, M. Pàmies, “Fast lines at theme parks,” *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering,* Vol. 7, No. 6, 2013.
[3] N. Ueda, F. Naya, H. Shimizu, T. Iwata, M. Okawa, H. Sawada, “Real-time and proactive navigation via spatio-temporal prediction,” in *Proc. 2015 ACM International Symposium on Wearable Computers (UbiComp)*, pp. 1559-1566, 2015.

【連絡先】

上田 修功 (Naonori Ueda) 上田特別研究室, **納谷 太 (Futoshi Naya)** 協創情報研究部 知能創発環境研究グループ E-mail：{ueda.naonori, naya.futoshi}(at)lab.ntt.co.jp

Trajectory analysis and prediction using deep learning

05 どうやって来た？どこへ行く？

～深層学習を利用した移動軌跡分析・予測技術～

どんな研究

スマートフォン等で取得できる移動軌跡から、移動行動の潜在的なコンテキストを予測する研究です。本展示では、移動軌跡から、徒歩や鉄道、自動車等の移動手段を推定する技術と、今後訪れると考えられる目的地を予測する技術について紹介します。

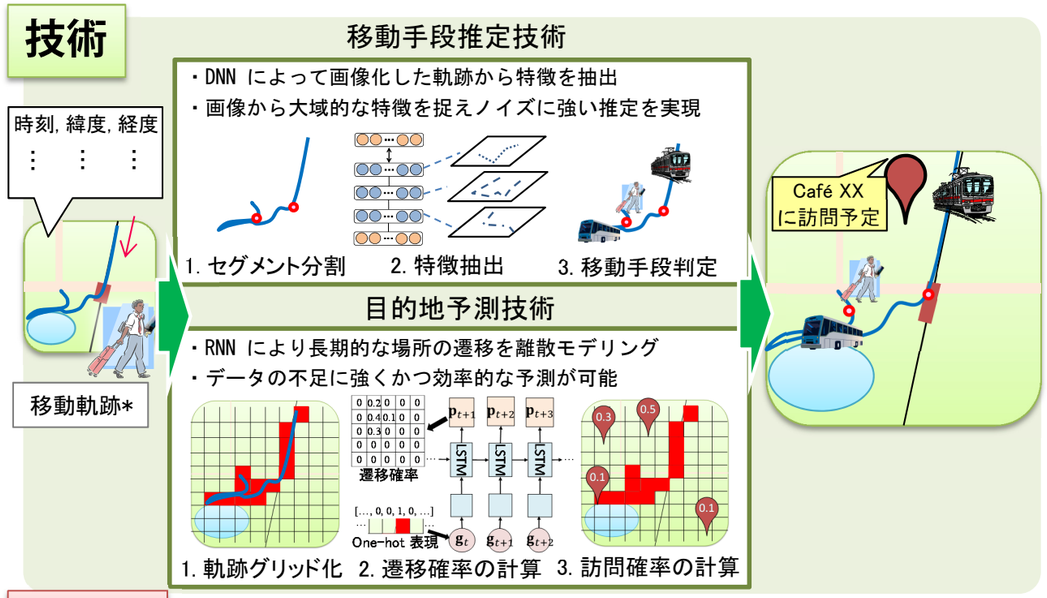
どこが凄い

(1) 移動手段推定技術は画像化した移動軌跡から深層学習により移動の特徴を自動抽出することで、従来より高精度に移動手段を推定できます。
(2) 目的地予測技術は人の移動をRecurrent Neural Network でモデリングすることで、従来よりデータ不足の影響を受けずに、長期的な動きを考慮し高精度に予測できます。

目指す未来

スマートフォン等の普及により、人の行動に関わる多様なデータが収集されています。しかしこれらのデータだけで人の行動を理解し、活用することは困難でした。本技術により人の行動を意味的に解釈すること、予測することが可能となり、移動軌跡を利用した個人向け行動支援、人流の解析・制御等につながります。

技術



応用例

コンシェルジュサービス

● 個人の移動行動分析

【車】音声で案内
【鉄道】画面で案内
渋谷に向かわれるのであれば先日オープンしたケーキ店に行ってみませんか?

観光ナビゲーション

● 観光客の移動行動集計

次のスポットまでは徒歩の方が早く到着できます

*移動軌跡は携帯電話、スマートフォン等からGPS等を活用して取得する

【関連文献】
[1] Y. Endo, H. Toda, K. Nishida, A. Kawanobe, “Deep feature extraction from trajectories for transportation mode estimation,” in *Proc. the 20th Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2016.
[2] A. Y. Xue, J. Qi, X. Xie, R. Zhang, J. Huang, Y. Li, “Solving the data sparsity problem in destination prediction,” *The VLDB Journal*, 24(2):pp. 219-243, Apr. 2015.

【 連 絡 先 】
遠藤 結城 (Yuki Endo)　サービスエボリューション研究所
E-mail：endo.yuki(at)lab.ntt.co.jp

Predict the traffic of future events based on the individual behavior

06 2020のトラヒックホットスポットはどこ？

～ユーザ行動に基づく将来イベントトラヒック予測技術～

どんな研究

モバイル端末の普及・高機能化により、巨大な通信需要を伴う過去に実施経験のない「将来イベント」の増加が見込まれます。本展示では、過去イベントにおけるモバイルユーザの行動分析に基づき、将来イベントのトラヒック発生を予測する技術についてご紹介します。

どこが凄い

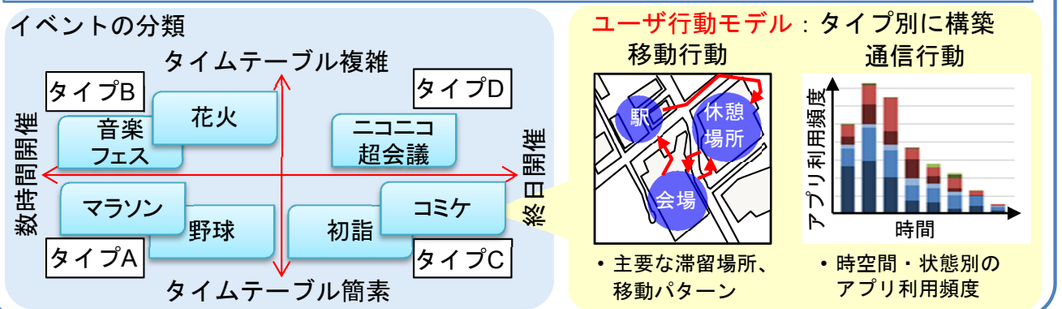
従来の予測は、定期開催のイベントを対象とし過去の経験に基づくものであり、「将来イベント」への応用が困難でした。イベント目的に則して人の行動が集約される点に着目し、イベント特有な行動メカニズムのモデル化により、「将来イベント」トラヒック予測を可能にします。

目指す未来

本技術により、必要最小限となる通信リソースの事前見積りや、トラヒックホットスポットの発生に対する通信リソースの計画的な制御が可能となり、ノウハウなしで本番に臨む２０２０年オリンピック・パラリンピックに対しても，高品質なモバイル通信環境を提供します。

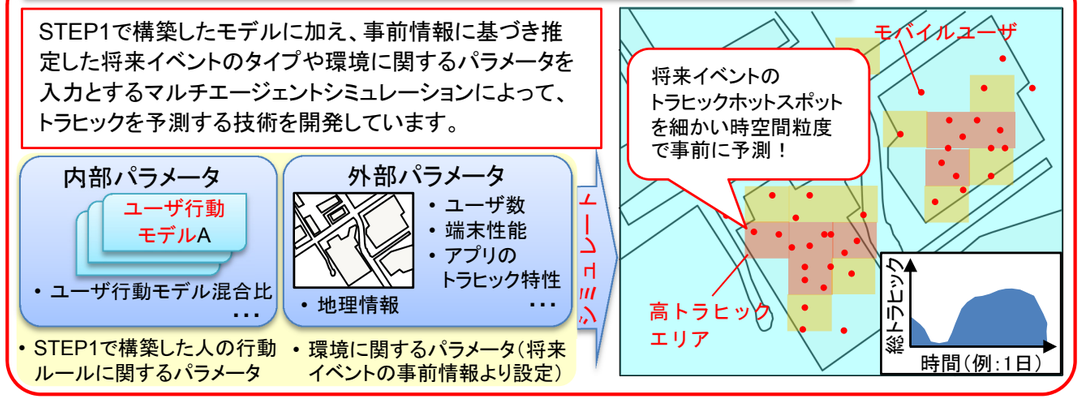
STEP1：過去データの分析からイベントタイプ別にユーザ行動モデルを構築

過去のイベントデータを通じて、モバイル端末ユーザのアプリ利用頻度、移動パターン、アプリ利用(通信発生)のトリガ等、移動・通信行動の抽出およびモデル化に取り組んでいます。



STEP2：ユーザ行動モデルを用いて将来イベントのトラヒックを予測

STEP1で構築したモデルに加え、事前情報に基づき推定した将来イベントのタイプや環境に関するパラメータを入力とするマルチエージェントシミュレーションによって、トラヒックを予測する技術を開発しています。



【関連文献】
[1] 本多泰理, 高橋洋介, 石橋圭介, “高品質･低コストの実現に向けたプロアクティブなネットワーク制御,” *NTT技術ジャーナル*, Vol. 27, No. 7, pp. 20-23, 2015.

【 連 絡 先 】
塩津 晃明 (Akihiro Shiozu)　ネットワーク基盤技術研究所
E-mail：shiozu.akihiro(at)lab.ntt.co.jp

07

Efficient learning utilizing directions of past gradients

高速かつ高精度な深層学習を実現します

～時系列の勾配方向に着目した高効率な学習アルゴリズム～

どんな研究

深層学習は音声や画像認識などの様々なタスクで高い精度を達成し、注目を浴びています。しかし、深層学習は学習に時間のかかる技術です。そこで本研究では、深層学習の学習時間を短縮するための効率のよい学習アルゴリズムの研究を進めています。

どこが凄い

過去の重みの更新方向の情報を用いて学習中の重みの更新量を適切に調整することで、深層学習の学習時間を２～５倍短縮することに成功しました。加えて、学習に必要だった手動調整コストを減らし、さらに従来よりも高い精度を達成することができました。

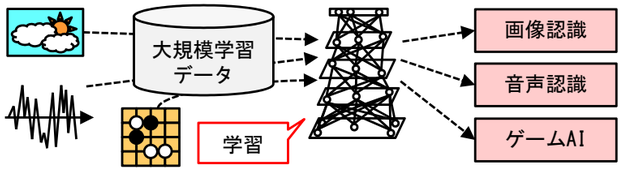
目指す未来

深層学習は学習コストが高く、気軽に使える技術ではありませんでした。我々は低コストに深層学習を実現するための研究開発を今後も進めていきます。これにより音声・画像認識、自然言語処理、AIなどがますます社会に浸透し、便利で快適な社会の実現を加速させていきます。

学習時の重みの更新量を独自の方法で自動調整することによって、
ニューラルネットの学習時間を２～５倍短縮します。

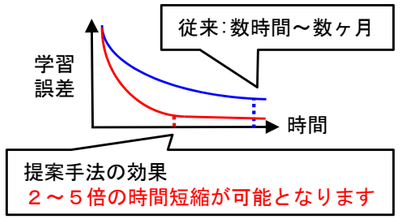
背景

深層学習はデータから深いニューラルネットを学習し、
画像認識やゲームAIで人間を超える性能を発揮しています。

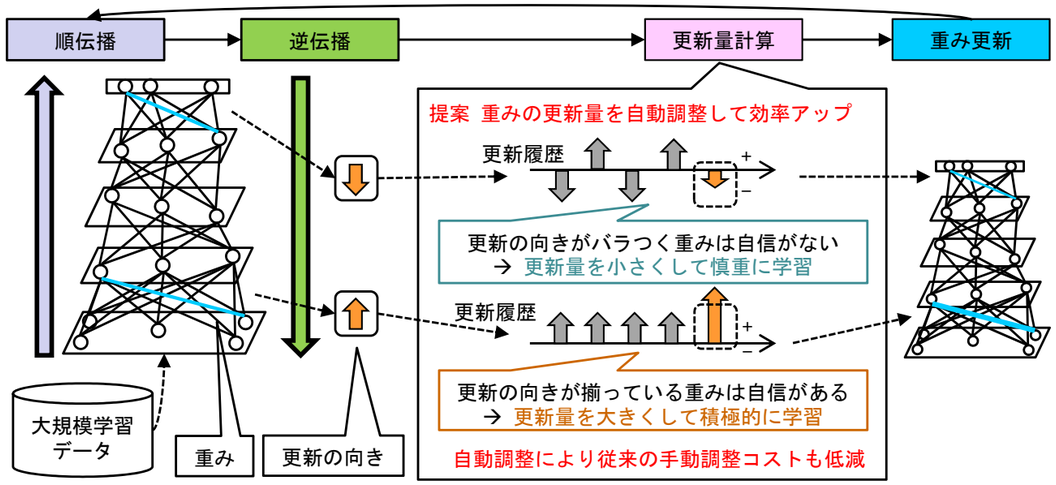


課題

学習に時間がかかります。



手法概要　学習の計算プロセスの一つである、重みの更新量計算を工夫。



【関連文献】

【連絡先】
井田 安俊 (Yasutoshi Ida)　ソフトウェアイノベーションセンタ
E-mail：ida.yasutoshi(at)lab.ntt.co.jp

08

Multipotential coding method achieving the Shannon limit

訂正、圧縮、おてのもの

～シャノン限界を達成できる多能な符号化法 ～

どんな研究

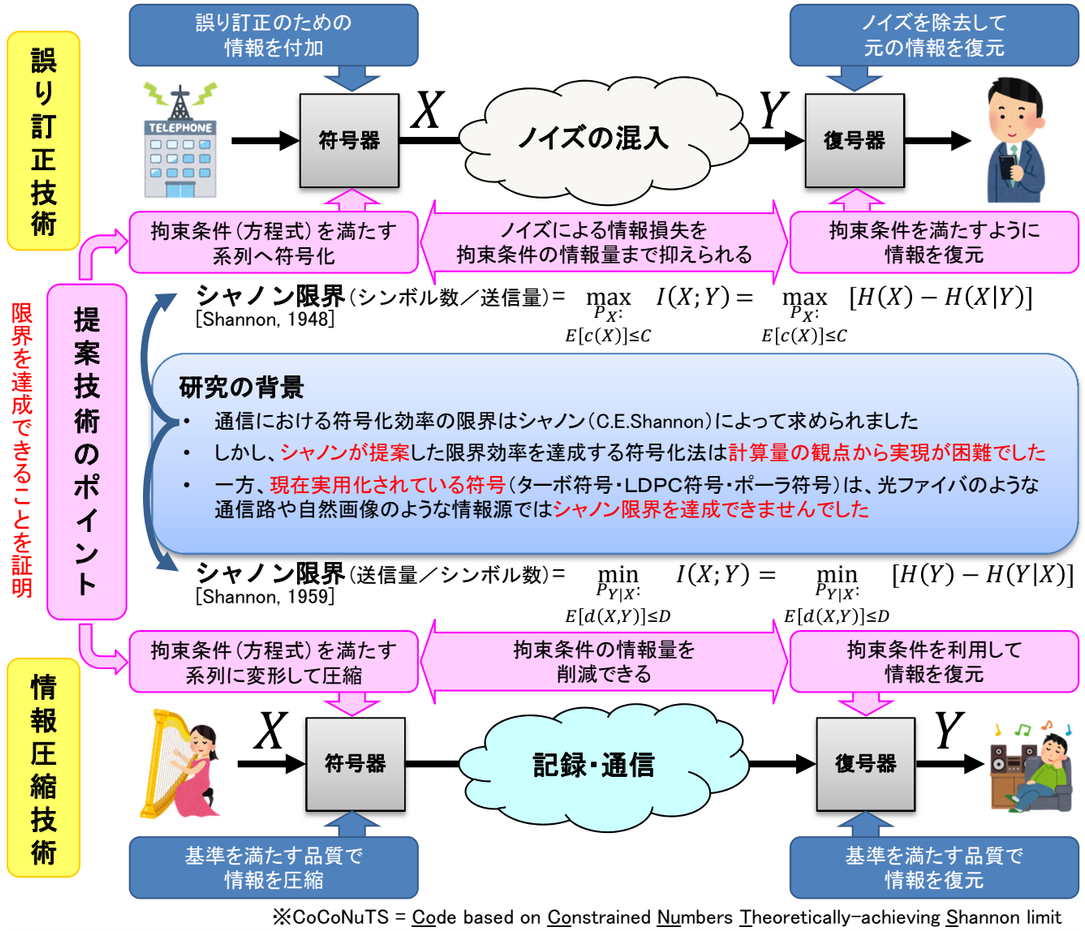
信号にノイズが混入しても正しく情報を伝える誤り訂正技術、信号を少ないビット数で表現する情報圧縮技術に関する研究です。実用化されている従来技術は、目的に特化されており、限定された通信路・情報源でしか理論限界を達成できませんでした。展示では、この課題を解決するための符号化理論を紹介します。

どこが凄い

拘束条件を満たす系列に基づく符号化法（CoCoNuTS※）を考案し、これが任意の情報源・通信路に対して理論限界を達成することを数学的に証明しました。この符号化理論により、多くの通信・記録の手段（誤り訂正や情報圧縮など）において、理論限界を達成する実用的な符号を統一的に構築することができます。

目指す未来

光ファイバ・無線ＬＡＮ・携帯電話などにおける信頼性の高い大容量通信や、ＣＤ・ＤＶＤ・ＢＤ・フラッシュメモリなどにおける高品質な大容量記録を実現できます。また、本技術により、符号化による効率を理論限界まで高めることができます。



※CoCoNuTS = Code based on Constrained Numbers Theoretically-achieving Shannon limit

【関連文献】

[1] J. Muramatsu, S. Miyake, “Hash property and coding theorems for sparse matrices and maximal-likelihood coding,” *IEEE Transactions on Information Theory*, Vol. IT-56, No. 5, pp. 2143-2167, May 2010. Corrections: Vol. IT-56, No. 9, p. 4762, Sep. 2010, Vol. IT-59, No. 10, pp. 6952-6953, Oct. 2013.
[2] J. Muramatsu, “Channel coding and lossy source coding using a generator of constrained random numbers,” *IEEE Transactions on Information Theory*, Vol. IT-60, No. 5, pp. 2667-2686, May 2014.

【連絡先】
村松 純 (Jun Muramasu)　協創情報研究部 創発知能環境研究グループ
E-mail：muramatsu.jun(at)lab.ntt.co.jp

09

High-speed search over hypergraphs via quantum walk

巨大ネットワークに潜む構造を量子で探索

～量子ウォーク技術による超グラフの高速探索～

どんな研究

巨大な超グラフの中から、与えられた条件に合う構造を高速に見つけ出す手法の研究です。例えば，ウェブアクセスログやセンサーから得た膨大なデータを超グラフとして表現することにより、様々な関係や法則を高速に発見するのに役立ちます。

どこが凄い

最新の量子ウォーク技術を駆使した新たな量子探索アルゴリズムを開発しました。既存(非量子)のアルゴリズムの限界を大きく超える高速性を実現するばかりでなく、標準的な量子探索アルゴリズムであるGroverのアルゴリズムをも凌ぎます。

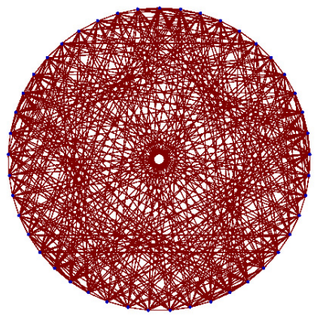
目指す未来

巨大な超グラフの中から、特定の構造を見つけ出す作業を既存のアルゴリズムで行うと、膨大な計算時間を必要とします。このため、実質的に実行不可能となります。高速な量子アルゴリズムの発見により、これが可能となり、膨大なデータ中に隠された法則や関係性を新たに発見できることが期待されます。

巨大な超グラフ全体を調べることなく、頂点と枝を量子的にサンプリングすることにより
高速(*)かつ99%以上の高確率で、条件に合う構造（部分超グラフ）を発見します
(*) 100万ノードの超グラフに対して量子を使わない場合より500万倍程度高速

技術のポイント：量子サンプリング手法
量子ウォークを用いてサンプリングを行うことにより、高速性と確実性を両立させます

- 超グラフの頂点・枝の部分集合(サンプル)をジョンソン・グラフと呼ばれるグラフの頂点に符号化
- ジョンソン・グラフ J を量子ランダム性を用いてたどる(量子ウォーク)ことにより、ほとんどランダムに J の頂点を抽出
- 抽出した頂点を復号化して得られるサンプルが，条件に合う構造を含むかチェック



ジョンソン・グラフ J(10, 2)

【関連文献】

[1] F. Le Gall, H. Nishimura, S. Tani, “Quantum algorithms for finding constant-sized sub-hypergraphs,” *Theoretical Computer Science*, Vol. 609, Part 3, pp. 569 – 582, 2016.
[2] F. Le Gall, H. Nishimura, S. Tani, “Quantum algorithms for finding constant-sized sub-hypergraphs,” in *Proc. 20th International Computing and Combinatorics Conference* (*COCOON '14*), 2014.

【 連 絡 先 】
谷 誠一郎 (Seiichiro Tani)　メディア情報研究部 情報基礎理論研究グループ
E-mail：tani.seiichiro(at)lab.ntt.co.jp

10

Formal verification of the QUIC and TLS protocols

次世代webの安全性がより確かなものに

～フォーマルメソッドによるQUIC・TLS1.3の検証～

どんな研究

インターネット上の通信の安全性は認証や暗号化の方法を定めたセキュリティ・プロトコルに従って守られていますが、その欠陥が近年多く発見されています。セキュリティ・プロトコルの安全性を、その仕様に基づいて自動的かつ厳密に検査する研究です。

どこが凄い

計算機を用いて攻撃を効率的・網羅的にチェックします。次世代web認証・暗号化プロトコルTLS1.3の安全性を確認するとともに、Googleによって策定されたweb認証・暗号化プロトコルQUICが従来研究で“証明”された安全性を満たさないことを発見しました。

目指す未来

現在はプロトコル仕様が策定されて広く利用されるようになってから欠陥が発見されるため、対策に多大な労力と時間を費やしています。また、その間は安全でない状態が続きます。本研究によって仕様を決める段階で安全性を検査し、利用前に欠陥を発見・修正できるようになります。

セキュリティ・プロトコルの策定と欠陥

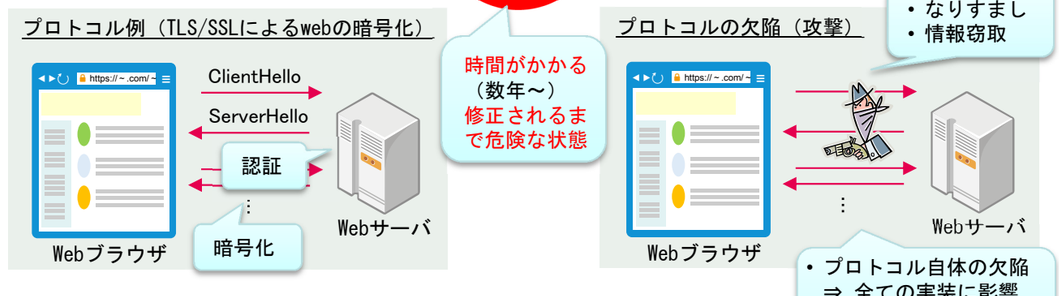
プロトコル仕様の策定

- 認証（通信相手の確認）の手順
- 部品として利用する暗号技術（暗号・署名）

実装・展開
修正

プロトコルの利用

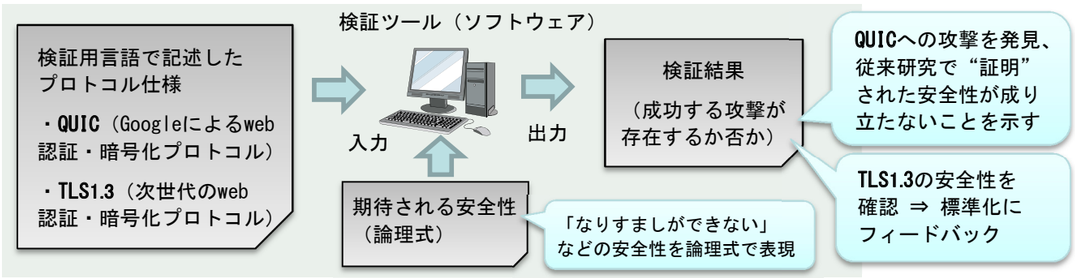
近年、広く利用されているプロトコルに欠陥が多く発見される



フォーマルメソッドによる安全性検証

目標：プロトコルの仕様を決める段階で、プロトコルの安全性を厳密に検査

- プロトコル仕様を入力として安全性を検査することで、実装・利用前の検査を可能に
- 数理論理学に基づくフォーマルメソッドにより、安全性を論理式で記述し、これに反する攻撃を網羅的に検査することで、専門家でも見逃してしまうような欠陥も発見可能に



【関連文献】

[1] 櫻田英樹, 米山一樹, 吉田真紀, 花谷嘉一, “形式検証に向けたQUICの安全性定義の検討,” *日本応用数理学会2015年度年会*, 2015.
[2] 荒井研一, 徳重佑樹, 櫻田英樹, “ProVerifによるTLS 1.3ハンドシェイクプロトコルの形式検証（その2）,” *2016年暗号と情報セキュリティシンポジウム*, 2016.

【 連 絡 先 】
櫻田 英樹 (Hideki Sakurada)　メディア情報研究部 情報基礎理論研究グループE-mail：sakurada.hideki(at)lab.ntt.co.jp

11

Gbps streaming of physical random numbers

情報を守る真のデタラメをつくる

～物理乱数 Gbpsストリーミングの実現～

どんな研究

本質的に解読が不可能な究極の安全性を保証する暗号化を実現するためには物理乱数が不可欠です。私たちは、カオスレーザから生成した物理乱数列を実用的な速度で供給する研究をおこなっています。この展示では、1秒間で1Gbit(Gbps)以上の速度で直接パソコンなどへ物理乱数列を供給するシステムを紹介します。

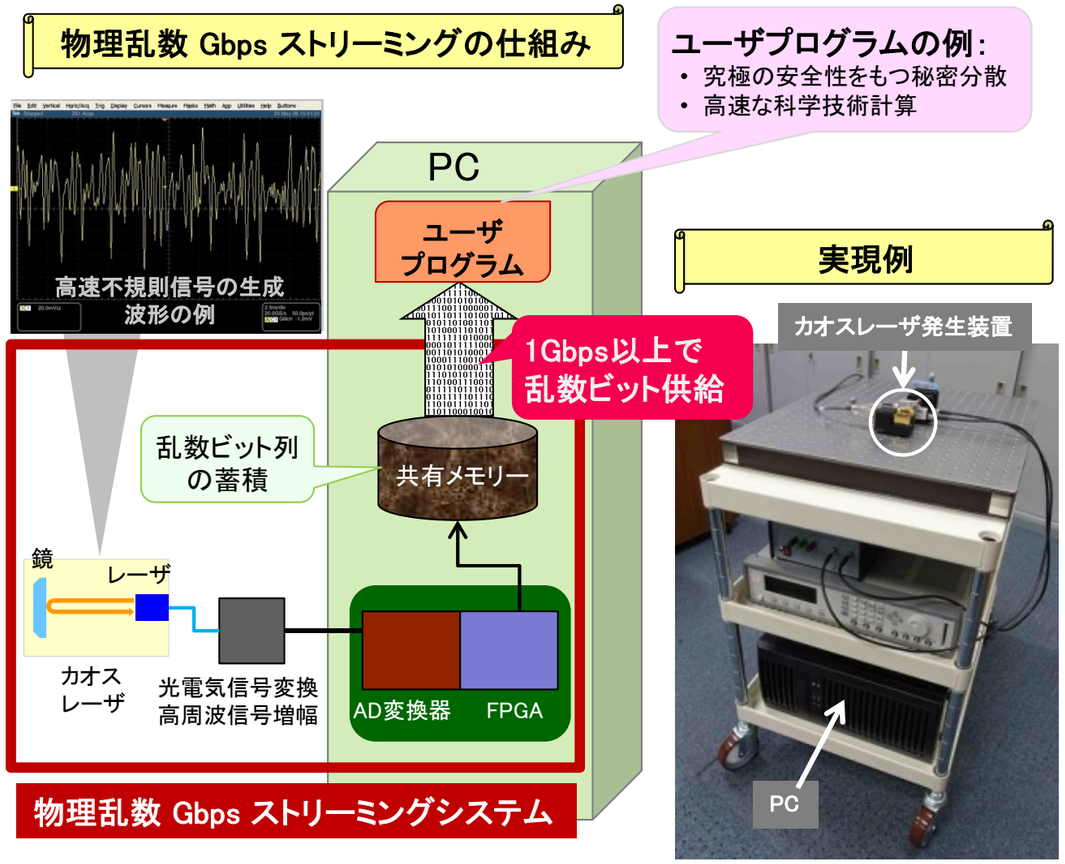
どこが凄い

生成速度が遅いことが従来の物理乱数生成器の欠点でした。研究段階で、カオスレーザの信号を記録し乱数列に事後変換することで生成速度がGbps超になることが分かりました。高速AD変換器と高速処理システムを融合し、ユーザが高速に生成される乱数列をリアルタイムで利用できるシステムを構築しました。

目指す未来

現在一般に使われているデータの秘匿化手法は、膨大な解読時間がその安全性の根拠となっています。将来、物理乱数の高速な生成システムが実用化できれば、無限に時間をかけても解読できない究極の安全性が保証された暗号化が実現可能となり、より安心安全なデータの秘匿化ができるようになります。

物理乱数 Gbps ストリーミングの仕組み



実現例

【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,” *Phys. Rev. A*, Vol. 83, 031803(R), 2011.
[2] T. Harayama, S. Sunada, K. Yoshimura, J. Muramatsu, K. Arai, A. Uchida, P. Davis, “Theory of fast nondeterministic physical random-bit generation with chaotic lasers,” *Phys. Rev. E*, Vol. 85, 046215, 2012.

【 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

12

Picture-book search with interdisciplinary approach

「ピタリエ」でこの子にピッタリの絵本探し

～興味・発達段階を考慮したグラフ索引型類似探索技術～

どんな研究

こどもの興味と発達段階に合った絵本を見つけるための研究です。従来、書誌情報やキーワードによる検索が行われており、文章の内容と絵に関する特徴は有効に利用されていませんでした。また、対象年齢に関する情報が不十分な絵本が多く、発達段階に応じた検索が困難でした。

どこが凄い

複数種類の特徴を扱える汎用検索アルゴリズムを基に、文や絵の好みを統合できる基盤を開発しました。また、こどもの語彙獲得記録と言語処理的知見を用い文の読みやすさを推定する技術を開発しました。このように、複数研究分野の知見によりこどもにピッタリの絵本を見つけます。

目指す未来

絵本は、多様な物語が様々な語彙と絵により表現されており、読み聞かせ等のコミュニケーションによる教育効果が注目されています。こどもの興味や発達段階に合う絵本を、本技術で検索し利用することにより、こどもの語彙獲得の促進と情操の育成などの効果が期待できます。

ぴたりえ でこんな絵本探しが可能に！

- あいまいな記憶しかなくても、おおまかなストーリーを入力すれば探していた絵本が見つかる！
- 大好きな絵本に絵や文章が似た絵本が見つかる！
- 借りに行った絵本が図書館にないとき、代わりに同じテーマや文章の読みやすさの絵本が見つかる！

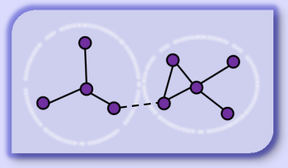
- シリーズ本の読みやすさの順序が分かって読む順の参考になる！

情報科学と人間科学の融合

類似探索

グラフ構造を索引とした高速汎用探索アルゴリズム

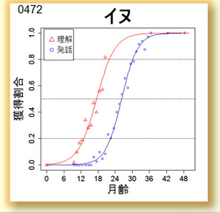
- 「知り合いの知り合いを辿ると6人で誰とでも繋がる」SNS的性質を高速探索に応用
- テキストも画像も音声も扱える高い汎用性



発達心理学

幼児語彙発達データベース

- 0−3才までの1300名分のデータを収集し「こどもがいつどんな語を理解し発話するか」についてモデル化



自然言語処理

ひらがな形態素解析と文章の読みやすさ推定

- 形態素解析の重要な手がかりである「漢字・かな・カナ」の境界がほとんどない、ほぼひらがなだけの絵本に対しても高い精度を実現
- 左記の語彙獲得記録と、文章の長さ等のテキスト特徴に基づいて読みやすさを推定

ひらがなだけだと分割困難
おひめさまになったきぶん

【関連文献】

[1] 服部正嗣, 藤田早苗, 青山一生, “テキストから得られる複数特徴量を融合する絵本類似探索法,” *人工知能学会第２９回全国大会発表論文集*, 2E4-NFC-01b-5, 2015.
[2] 藤田早苗, 小林哲生, 南泰浩, 杉山弘晃, “幼児を対象としたテキストの対象年齢推定方法,” *認知科学*. Vol. 22, No. 4, pp. 604-620, 2015.

【 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
E-mail：hattori.takashi(at)lab.ntt.co.jp

13

Written and spoken language analysis and machine translation

ニュースも会話も精読して翻訳します

～書き言葉から話し言葉へ～自然言語解析と機械翻訳～

どんな研究

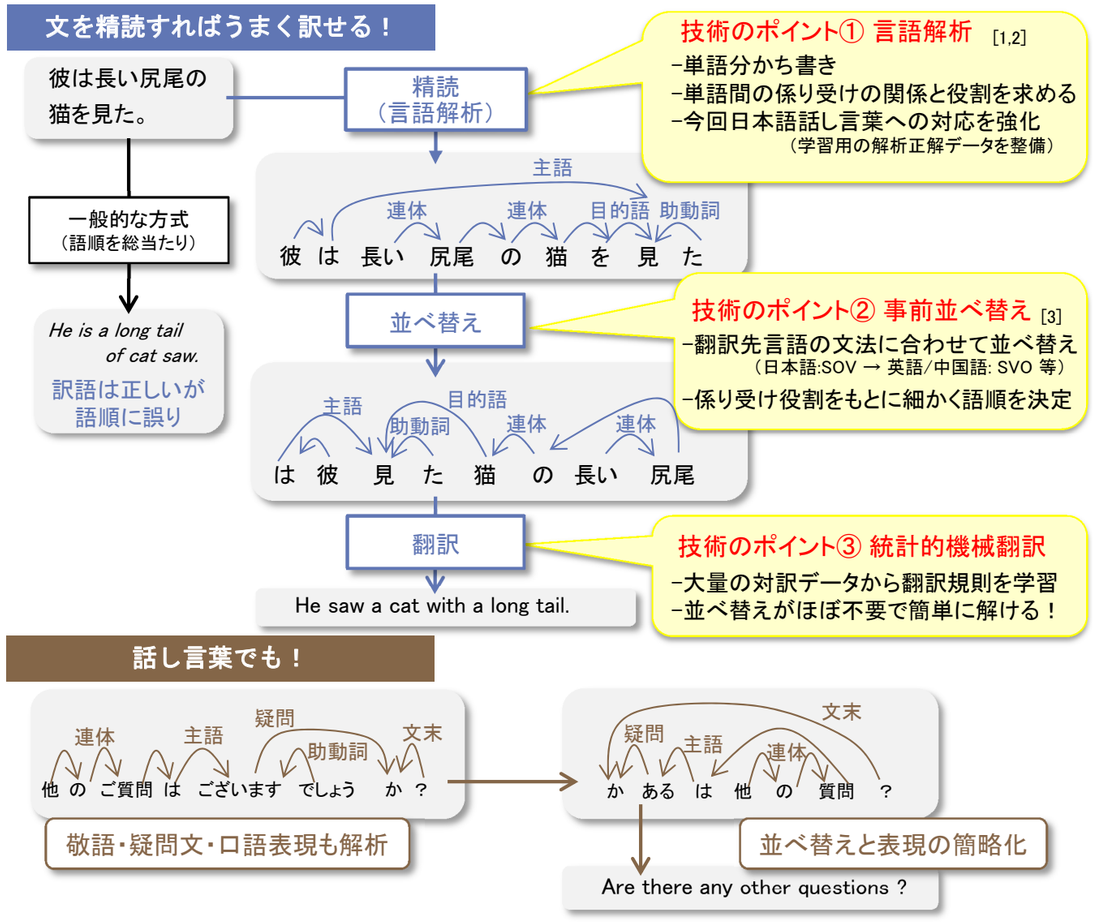
外国語は日本語と語順が異なるものが多く、機械翻訳を難しくしています。語順は文法で決まるため、文法に基づいた構文解析が重要です。この研究は、ニュースのような書き言葉だけでなく、話し言葉に対しても、構文解析を利用して語順を外国語に近づけるよう並べ替えて、翻訳しやすくするためのものです。

どこが凄い

私たちは構文解析によって得られる各単語間の係り受けとその役割に応じて、翻訳先の言語の文法に合うように語順を並べ替える方式で翻訳精度を向上させてきました。今回日本語の話し言葉についても学習データを整備し、書き言葉と同じように精度よく構文解析、並べ替え、翻訳をできるようにしました。

目指す未来

日本語と外国語の間の双方向の機械翻訳を、書き言葉・話し言葉とも高精度に行うことで、言語の壁を越えて情報収集、情報発信、コミュニケーションを円滑にします。また，本技術を拡張し，音声認識と連携した効率的な音声翻訳の実現が見込めます。



【関連文献】

[1] T. Tanaka, M. Nagata, “Word-based Japanese typed dependency parsing with grammatical function analysis,” in *Proc. ACL-IJCNLP*, 2015

[2] S. Takeno, M. Nagata, K. Yamamoto, “Empty category detection using path features and distributed case frames,” in *Proc. EMNLP*, 2015.

[3] S. Hoshino et al., “Discriminative preordering meets Kendall’s tau maximization,” in *Proc. ACL-IJCNLP*, 2015

【連絡先】

須藤 克仁 (Katsuhito Sudoh)　協創情報研究部 言語知能研究グループ

E-mail：sudoh.katsuhito(at)lab.ntt.co.jp

14

Creating child language development database by checklist method

こどもがいつどんな語を覚えるのかを教えます

～語彙チェックリストによる幼児語彙発達データベースの作成～

どんな研究

こどもがいつどんな語を習得するのかは、データ取得の困難さのため、大規模なデータベースが少なく、多くのことがわかっていませんでした。本展示では、言語習得の仕組みを科学的に解明するため、0-3才のこどもがいつどんな語を言えるようになるかを研究した成果を紹介します。

どこが凄い

0-3才のこどもがいる親にチェックリストを回答してもらい、約1300名のデータから約2700語の習得特性モデルを推定し、3才までに覚える語をほぼ網羅するデータベース構築に成功しました。また月齢と獲得割合から該当語を表示するシステムも開発しました。

目指す未来

本データベースを土台に考案した発達検査や言語訓練支援アプリを小児医療現場で実際に使ってもらっています。また教育コンテンツ作成支援や絵本推薦システムの基盤技術としても大いに期待されています。こうした研究からエビデンスに基づく新しい教育ICTを提案します。

【関連文献】

[1] 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.

[2] T. Kobayashi, Y. Minami, Y. Okumura, “Fluctuating development of common nouns and predicates in early lexical development: evidence from analysis of large-sample vocabulary checklist data in Japanese children,” in *Proc. ECDP*, 2015.

【連絡先】

小林 哲生 (Tessei Kobayashi)　協創情報研究部 インタラクション対話研究グループE-mail：kobayashi.tessei(at)lab.ntt.co.jp

15

Encouraging speech dialogue using multiple robots

ロボットが増えると話がはずむ

～複数ロボットの協調動作による音声対話活性化～

どんな研究

現在の対話システムは、音声認識のエラーなどにより、不自然な応答をしてしまうことがあります。この不自然な応答が、ユーザとって対話システムを使いにくく感じる大きな要因となっています。本展示では、人と自然に楽しく対話できる雑談対話システムに関する研究成果を紹介します。

どこが凄い

対話における人の認知特性に着目し，不自然な応答でも複数のロボットを協調動作させることで、ユーザがストレスを感じにくくなることを発見しました。また、この知見を利用することで、音声認識エラーをはじめとする様々なエラーに頑健な音声対話システムを実現しました。

目指す未来

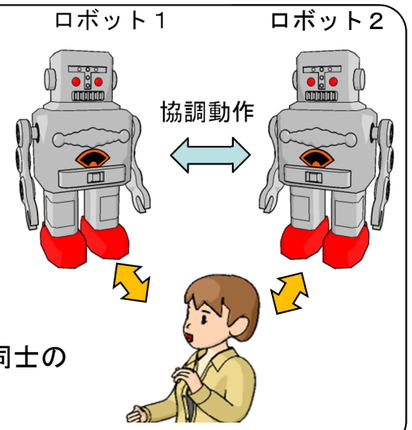
ロボットとの雑談を通して、人を様々な側面からサポートするシステムの実現を目指しています。人と気軽に雑談できるシステムは、人と情報をつなぐ自然なインターフェースとして大いに期待されています。また、人自身のコミュニケーションスキルの向上にも役立つと考えられます。

■ 従来研究：１対１の対話

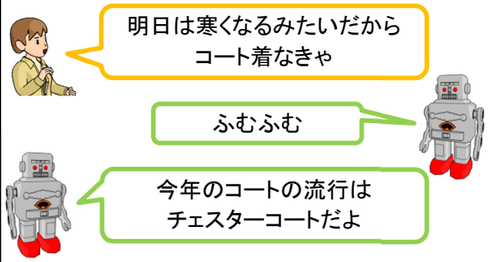
・ユーザ自身と同程度の高い対話能力をロボットにも要求してしまう
・ロボットが不自然な発話をした場合、ユーザがリカバリする必要がある
→ ユーザに大きな負担、ストレス

■ 本研究：複数ロボットとの対話

・ロボットが多数派のため，ロボットの対話能力が基準になる
・ロボットが不自然な発話をしても，ロボット同士の対話に遷移することで対話を継続できる
→ ストレス、負担を大幅軽減

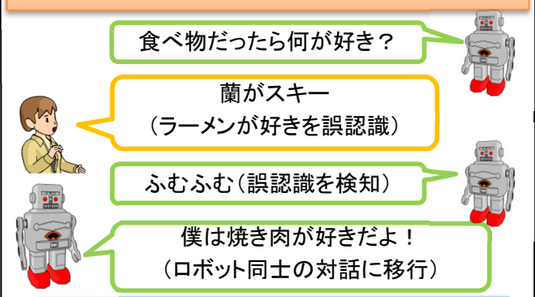


話題連続性に対する要求の低減



ロボットAが相槌などで一旦発話を受け止め、ロボットBが後続発話を行うことで、多少逸脱した発話になってもユーザの違和感が軽減される

音声認識エラー検出時の破綻回避



認識した発話が、対話の流れから逸脱した内容であった場合、音声認識の失敗と判断し、ロボット同士の対話に移行して対話破綻を回避する

【関連文献】

[1] H. Sugiyama, T. Meguro, R. Higashinaka, Y. Yasuhiro, “Open-domain utterance generation for conversational dialogue systems using Web-scale dependency structures,” in *Proc. the 14th annual SIGdial Meeting on Discourse and Dialogue (SIGDIAL)*, pp. 334-338, 2013.
[2] 杉山弘晃, 目黒豊美, 東中竜一郎, “対話システムのパーソナリティを問う質問の大規模な収集と分析,” *人工知能学会論文誌*, Vol. 31, No. 1, pp. 1-9, 2016.

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

16

Optical word recognition with CNN features and WFST decoding

飾り文字でも読み取れる

～深層学習と系列デコーディングによる文字列認識～

どんな研究

スマートフォンで撮影した街中の風景など、実環境の情景中に存在する文字を読み取ることは、その有用性が期待されながら、今なお困難な課題です。本研究は、デザインされた飾り文字も含めて、情景中に存在する様々な書体の文字列を正確に読み取ることを目指しています。

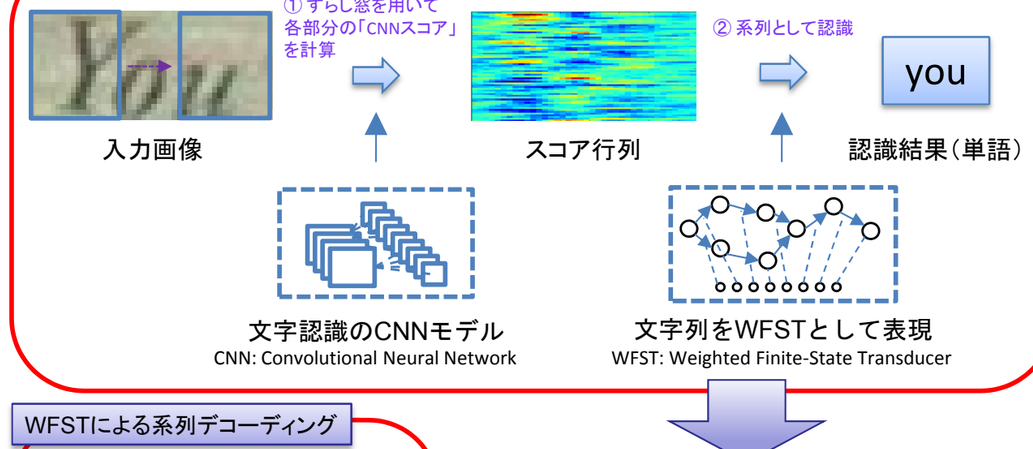
どこが凄い

同じ文字でも書体や撮影条件などにより画像には大きな違いが生まれ、認識が困難になります。これに対し、深層学習と、音声認識で有効性が示されている系列の認識手法であるWFST（重み付き有限状態トランスデューサ）とを組み合わせる方法を提案し高い認識精度を実現しました。

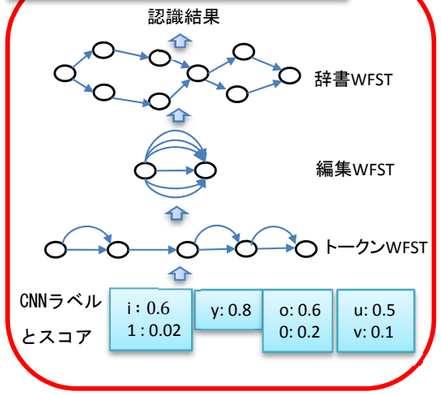
目指す未来

文字を読み取ることで、コンピュータが効率的に情報を取得することや、ユーザにとって必要な情報を的確に提示することが可能になります。効率的な画像探索や駅・店頭・市街地での歩行者用ナビゲーション、車の自動運転、画面の自動翻訳など、様々な応用の可能性が広がります。

文字列認識の流れ



WFSTによる系列デコーディング



実験結果　画像データセットを用いた単語認識精度の評価

Method \ Dataset (辞書の単語数)	ICDAR2003 (50)	ICDAR2003 (860)	SVT-WORD (50)
Jaderberg et al. [ECCV 2014]	96.2	91.5	86.1
Proposed	**96.9**	**91.8**	**93.4**

ICDAR：Intl. Conf. on Document Analysis and Recognition Dataset（画像860枚）
SVT-WORD：Street View Text Dataset（画像647枚）

【関連文献】

[1] X. Liu, T. Kawanishi, X. Wu, K. Kashino, “Scene text recognition with high performance CNN classifier and efficient word inference,” in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.

【連絡先】
劉 新豪 (Xinhao LIU)　メディア情報研究部 メディア認識研究グループ
E-mail：liu.xinhao(at)lab.ntt.co.jp

17

Prosody conversion with generative voice F0 contour model

声の抑揚が別人に変わる

～基本周波数パターン生成過程モデルに基づく韻律変換～

どんな研究

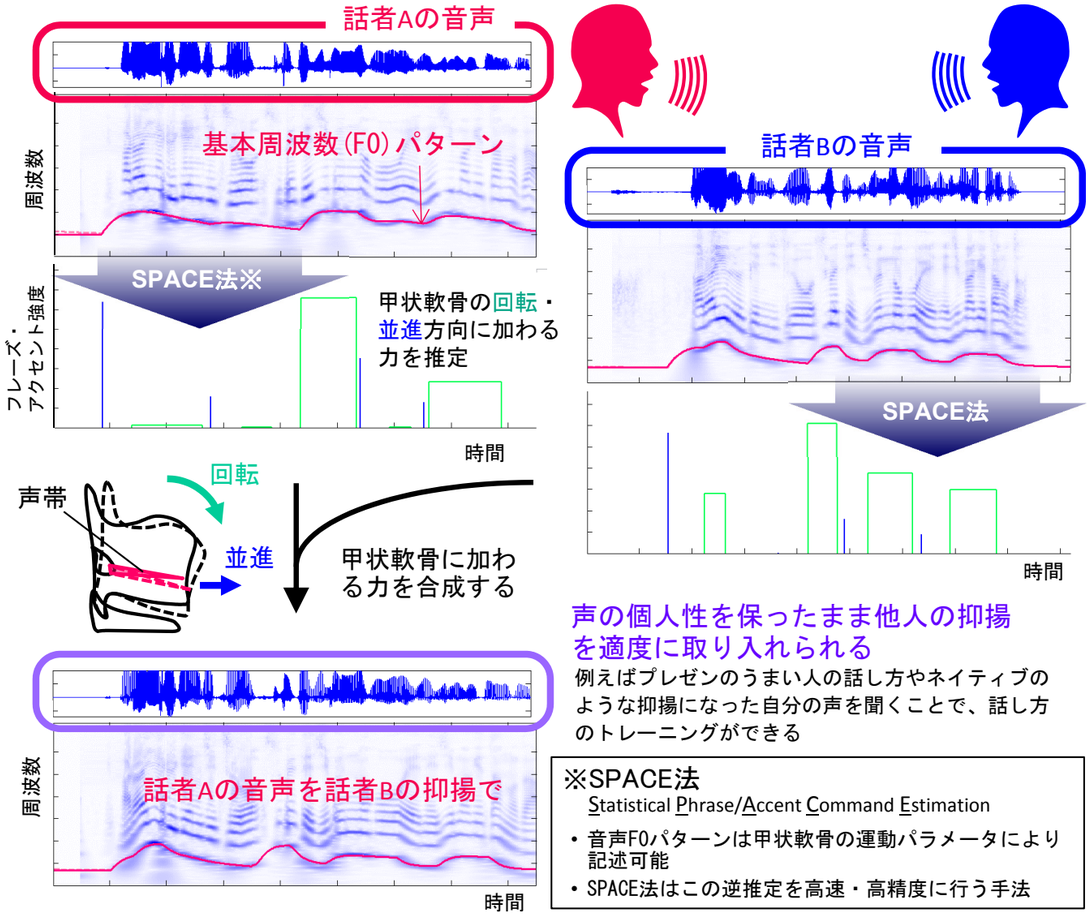
音声は抑揚（声の高さの変化）とリズムのつけ方によって大きく異なる印象の話し方になります。本展示では、ある話者と別の話者の抑揚・リズムを「混ぜる」ことによって、本人の声色は変えずに話し方だけを変えることができる高品質音声変換技術を紹介します。

どこが凄い

声の高さに相当する基本周波数(F0)パターンは声帯に張力を与える甲状軟骨の運動によって生み出されています。本技術では、二つのF0パターンの単純な伸縮や混合ではなく、甲状軟骨に加わる力そのものを合成して音声を生成するので、変換音声の自然性がいつも保たれる点がポイントです。

目指す未来

映画やアニメの俳優・声優の話し方を自分好みに変換したり、喉頭摘出者の無喉頭音声に抑揚を付与したり、非母語話者の音声の抑揚を母語話者風に付け替えたりできる音声処理技術への応用が期待されます。プレゼンの練習や語学の学習支援としても利用することができます。



声の個人性を保ったまま他人の抑揚を適度に取り入れられる

例えばプレゼンのうまい人の話し方やネイティブのような抑揚になった自分の声を聞くことで、話し方のトレーニングができる

※SPACE法
Statistical Phrase/Accent Command Estimation

- 音声F0パターンは甲状軟骨の運動パラメータにより記述可能
- SPACE法はこの逆推定を高速・高精度に行う手法

【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] H. Fujisaki, O. Fujimura, Ed., “A note on the physiological and physical basis for the phrase and accent components in the voice fundamental frequency contour,” in *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York, NY, USA: Raven, 1988.

【 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

Adding information to art using high-definition multiband image

実物鑑賞では得られない発見を得る

～高精細マルチバンド画像撮影による美術品への情報付加～

どんな研究

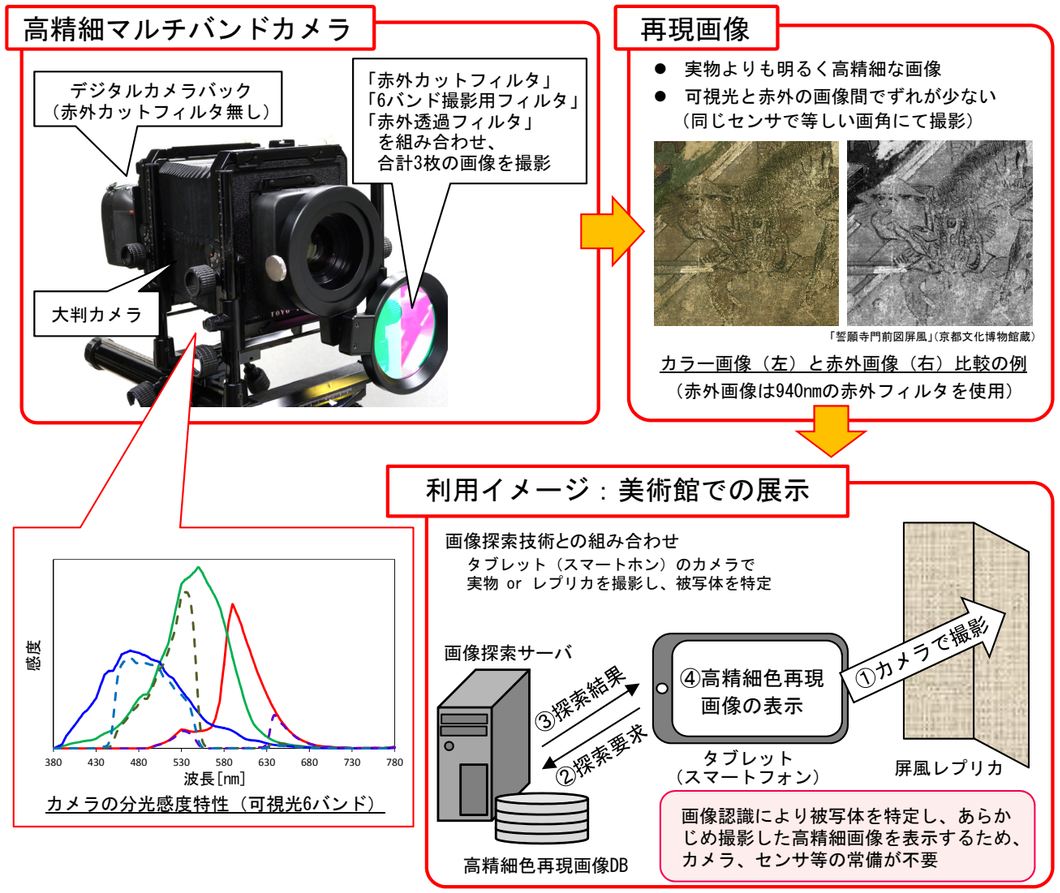
美術品の実物を直に鑑賞する際には、作品保護を目的とした展示方法の制限に加え、人間の目の性能が制約となり、様々な情報が知覚できない状態にあります。本技術による被写体の撮影により、肉眼による実物鑑賞では得られない情報を記録し、ソフトウェア上での処理により新たな知識を得ることを目的としています。

どこが凄い

赤外画像の撮影により消えかかった線画や下絵の可視化を実現しました。画像の解像度も0.1mm/pixel以上と超高精細です。被写体の正確な色再現、実際の展示環境とは異なる照明下での色の見えの再現（照明環境の変更）も可能です。さらに従来は手間が掛かったこれらの画像同士の比較を、簡易な方法で実現しました。

目指す未来

画像入力における波長方向（色）のダイナミックレンジ拡大と高分解能化、および画像処理による情報抽出技術が、肉眼では識別できない光学的な特性の可視化を実現する鍵となります。これらの技術は文化財のデジタルアーカイブに加え、医療分野や工業製品等の非破壊検査への応用が期待されます。



【関連文献】

[1] M. Tsuchida, K. Yano, and H. Tanaka, “Development of a high-definition and multispectral image capturing system for digital archiving of early modern tapestries of Kyoto Gion Festival,” in *Proc. ICPR2010*, pp. 2828-2831, 2010
[2] 土田勝, “市販のデジタルカメラを用いたマルチスペクトル画像撮影システム（特集：更なる進展が期待されるマルチスペクトル技術）,” *映像情報メディア学会誌*, Vol. 68, No. 04, pp. 286(32) - 290(35), 2014

【 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

Advances in distant speech recognition technologies

19 こんなにガヤガヤした場所でも聞き取ります

～遠隔発話音声認識における雑音除去・深層学習技術の最前線～

どんな研究

音声認識技術は近年飛躍的に向上していますが、展示会場のような騒がしい場所で口元から離れたマイクを使う場合、雑音や残響の影響を大きく受け、未だに音声認識は難しいという課題があります。この展示では、騒がしい環境でも精度のよい音声認識を実現するための基盤技術を紹介します。

どこが凄い

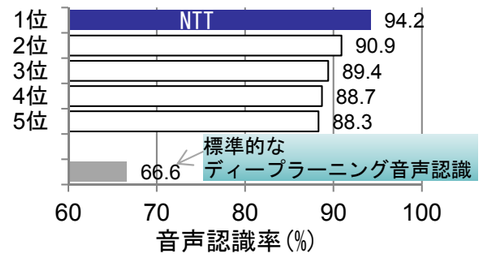
音声認識性能を劣化させる雑音や残響を、音声を歪ませることなく低減する技術や、雑音の影響を受けても精緻に音声をモデル化するディープラーニング音声認識技術を開発しました。これらの技術で、公共エリア雑音下でのモバイル音声認識国際技術評価(CHiME-3)で、世界1位の音声認識精度を達成しました。

目指す未来

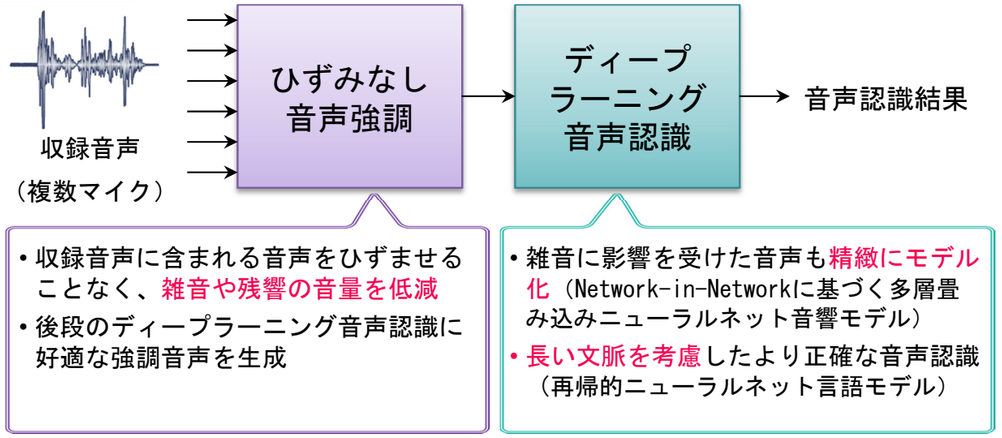
カフェや空港などの公共エリアでの音声インタフェースや、オフィスや家庭のリビングなどでの会話認識など、騒がしい所で多くの人が話す場面でも快適に動作する音声認識が実現できます。これは、スマホの音声エージェントやコミュニケーションロボットの利用シーンの拡大に大きく貢献すると期待されています。

音声認識の利用シーンの拡大へ

世界トップ性能を達成



技術の特徴



【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*, 2015.
[2] S. Araki, M. Okada, T. Higuchi, A. Ogawa, T. Nakatani, "Spatial correlation model based observation vector clustering and MVDR beamforming for meeting recognition," in *Proc. ICASSP2016*, 2016.

【連絡先】
荒木 章子 (Shoko Araki), 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　E-mail：{araki.shoko, kinoshita.k}(at)lab.ntt.co.jp

Low-bit-rate speech coding EVS and lossless audio coding ALS

20 高圧縮符号化で携帯電話も放送も更に高音質に

～低ビットレート符号化ＥＶＳと音響ロスレス符号化ＡＬＳ～

どんな研究

音声や音響信号を目的にあった条件（ビットレートや遅延）で最高の品質を維持して伝送符号に変換する（符号化する）技術の研究です。携帯電話や放送は、貴重な電波を多くのユーザーで共有する必要があり国際的な相互接続も保証するため国際標準化が必要です。私たちは標準化にも力を入れて取り組んでいます。

どこが凄い

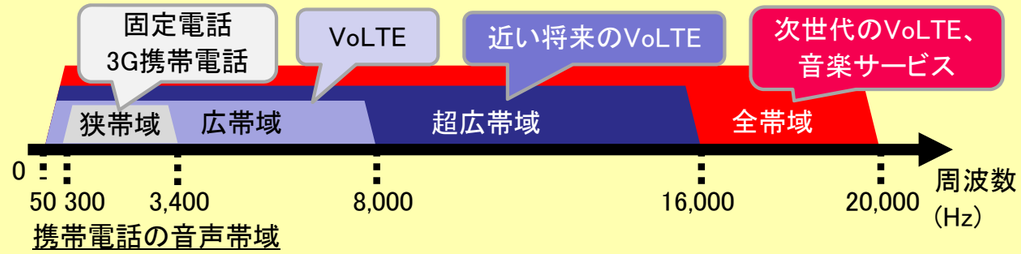
世界のトップレベルの品質、性能を達成しています。携帯電話ではほぼ目の前の肉声に近い音声、放送ではスタジオの原音をそのまま伝えることができるようになります。また，本技術は2種類の国際標準規格にそれぞれ準拠しています。これらの国際標準規格の制定にもＮＴＴ研究所は大きく貢献しました。

目指す未来

本技術は近い将来の日本の携帯電話（VoLTE)の音声通話サービス、衛星放送やIPTVの高音質サービスに利用されることが既に見込まれています。研究所は円滑なサービス導入を支援しつつ、さらなる高品質化をおこない，快適な通信環境を実現することを目指します．

携帯電話(VoLTE*)用　3GPP EVS**

- 低ビットレート
- 全帯域対応（最大48kHzサンプリングレート）
- 低遅延
- 現在のVoLTEと相互接続可能
- 背景雑音時や音楽でも高品質

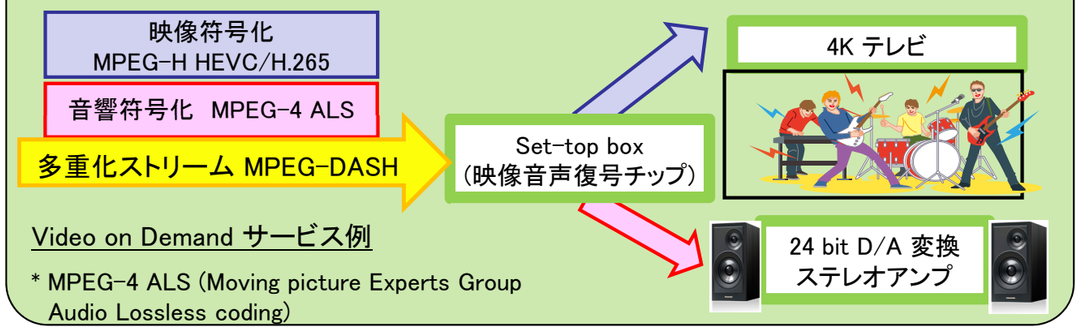


携帯電話の音声帯域

* VoLTE (Voice over Long Term Evolution)
** 3GPP EVS (3rd Generation Partnership Project Enhanced Voice Services)

放送用　MPEG-4 ALS*

- 情報が欠損しない（ロスレス）圧縮符号化
- 4K/8K放送と同期した高音質サービス（24ビット振幅、48kHzサンプリングレート）



Video on Demand サービス例

* MPEG-4 ALS (Moving picture Experts Group Audio Lossless coding)

【関連文献】

[1] 守谷健弘, "音声音響符号化技術と3GPPでの標準化," *NTT技術ジャーナル*, Vol.27, No.4, pp. 33-36, 2015.
[2] 鎌本優, 守谷健弘, 原田登, Csaba Kos, "ロスレス・オーディオ符号化MPEG-4 ALS の高性能化," *NTT技術ジャーナル*, Vol.20, No.2, pp.11-18, 2008.
[3] 原田登, 守谷健弘, 鎌本優, "MPEG-4 ALS の性能・応用と関連する標準化活動," *NTT技術ジャーナル*, Vol.20, No.2, pp.19-25, 2008.

【連絡先】
守谷 健弘 (Takehiro Moriya), 鎌本 優 (Yutaka Kamamoto)　守谷特別研究室
E-mail：{moriya.takehiro, kamamoto.yutaka}(at)lab.ntt.co.jp

21

Automatic projection of motion impression using image search

変幻灯2.0：光をあてると踊り出す

～画像検索を用いた動き情報の自動投影～

どんな研究

対象に新しい印象を与える手法として、プロジェクタを使って対象に映像を投影する「プロジェクションマッピング」が有効です。変幻灯は、プロジェクションマッピングと錯覚を組み合わせ、身の回りのものに様々な動きの印象を与える技術です。

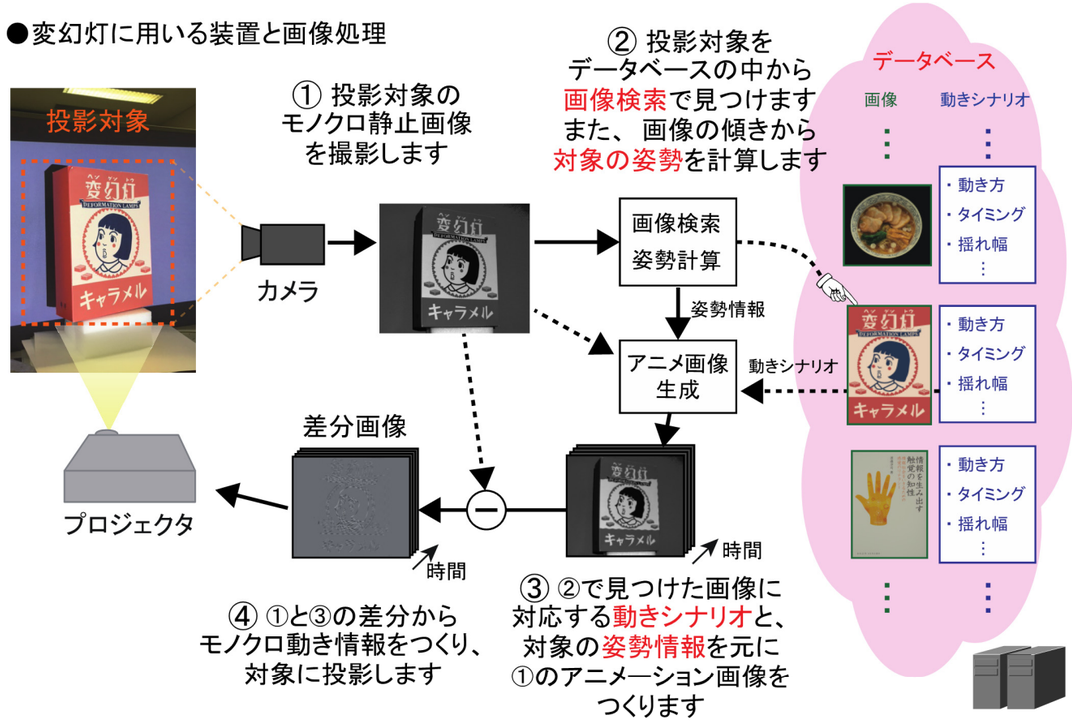
どこが凄い

対象物に映像を上書きする既存のプロジェクションマッピングと異なり、僅かな明暗の模様を投影するだけで対象そのものが動きだすような印象を生み出すことができます。さらに画像検索の技術と組み合わせることで、身の回りの対象を認識して、それに合った動きを即座に投影できるようになりました。

目指す未来

CGやアニメのような印象的な動きの情報は、これまで画面の中にしか存在し得ないものでした。変幻灯の技術を用いると、店頭の商品パッケージや部屋のインテリア等、身の回りに溢れるあらゆる物に動きがつけられます。変幻灯は現実世界と仮想世界がシームレスにつながる未来を構成する要素技術となるでしょう。

●変幻灯に用いる装置と画像処理



●変幻灯を支える脳の特性

変幻灯では白黒の動き情報だけを静止対象に投影するため、対象がもつ色や形の情報は止まったままです。しかし、人の脳は動き・色・形を別々に分析し、その間のずれを統合段階で修正します。変幻灯はこの脳の特性を利用し、対象が本当に動いているかのような印象を生み出します。

【関連文献】

[1] T. Kawabe, T. Fukiage, M. Sawayama, S. Nishida, “Deformation lamps: a projection technique to make a static object dynamic,” *ACM Transaction on Applied Perception*, 13(2), Article No. 10, 2016.
[2] T. Kawabe, M. Sawayama, S. Nishida, “Deformation lamps: a projection technique to make a static picture dynamic,” in *Proc. SIGGRAPH ‘15 Emerging Technologies*, Article No. 8, 2015.

【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

22

Material editing that makes object surfaces look wet

乾いた風景を雨上がりに

～物体表面を濡れているように見せる質感編集技術～

どんな研究

ものの質感を人間がどう認識しているかはよくわかっていません。ものの質感は雨に濡れたりすることで変化します。本研究では、湿り具合のようなものの質感の状態変化を推定する人間の知覚の方略を明らかにしました。さらに、この知見を応用して、映像の見かけのウェット感を操作する技術を開発しました。

どこが凄い

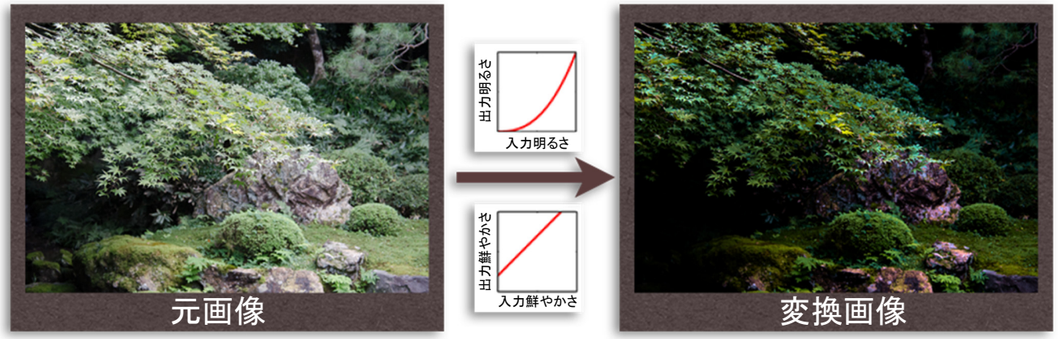
光の物理学と人間科学の知見を融合することで、ものの湿り具合を認識する際に人間が利用する明るさと色に関する画像特徴を特定しました。さらに、湿り具合の認識のように複雑な質感の認識は、映像に含まれる色の情報量の大きさに依存して行われるという知覚情報処理の新たな原理を発見しました。

目指す未来

ものの湿り具合のように複雑な質感をCG再現するためには現在とてもたくさんの計算量を必要とします。また、ものの複雑な質感を画像から認識するためには、ものの形や照明環境の情報を詳細に知る必要があります。人間の脳のしくみを利用することで、未来には簡便で高速に質感編集・質感認識が可能になります。

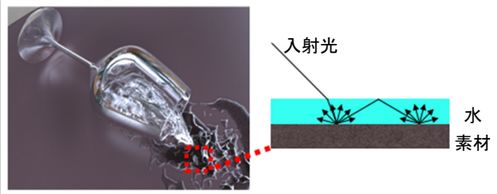
ウェット変換: 見かけのウェット感を上げる画像処理

操作1: 画像の明るさを変調します。元画像の中で明るい部分を明るいままに、暗い部分をより暗くします。
操作2: 元画像に含まれる全ての色の鮮やかさを強調します。



ウェット変換によって風景がなぜ濡れて見える？

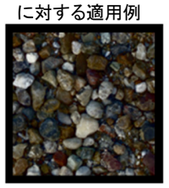
光学的な要因



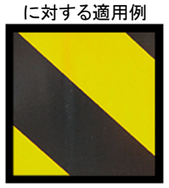
- 物理的にものが濡れた場合、水と素材との間で生じる光の多重反射によって光は暗く鮮やかになります。さらに、ものの表面を水が覆うことでハイライトが生じやすくなります。これらの色・明るさ特徴をウェット変換が定性的に捉えているために濡れて見えます。

認知情報処理的な要因

色数の多い画像に対する適用例



色数の少ない画像に対する適用例



- 色数が少ない画像より、多い画像のほうがウェット変換の効果は高くなります。
- ウェット変換が生む画像特徴はものが濡れる場合以外でも起こりますが、さまざまな表面に同時にその特徴が現れると湿りが原因である可能性が高まります。

【関連文献】

[1] 澤山 正貴, 西田 眞也, “表面画像の湿り度知覚における色・輝度処理の相互作用,” *日本視覚学会2015年冬季大会, Vision*, Vol. 27(1), p 46, 2015.
[2] M. Sawayama, S. Nishida, “Visual perception of surface wetness,’’ in *Proc. 15th Annual Meeting of Vision Sciences Society*, *Journal of vision*, Vol. 15(12), p 937, 2015.

【連絡先】

澤山 正貴 (Masataka Sawayama)　人間情報研究部 感覚表現研究グループE-mail：sawayama.masataka(at)lab.ntt.co.jp

Assistive technology for non-native speakers

23 英語は苦手？　でもだいじょうぶ

～音声認識と機械翻訳による非母語者にやさしい支援～

どんな研究

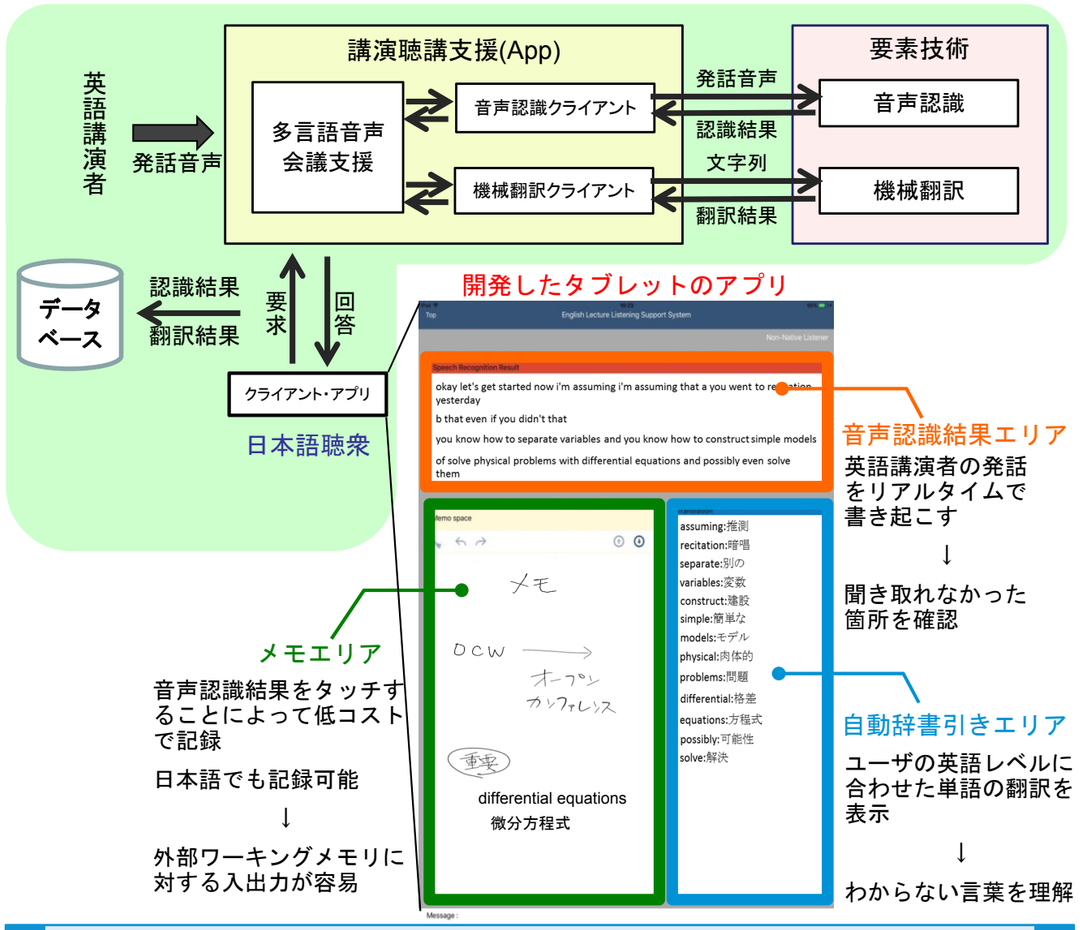
英語非母語者は、例えば英語講演の聴講中、大きな認知的負担を強いられています。情報技術の利用は非母語者支援に有用ですが、単純に使うだけでは負担を増大させ、英語理解を阻害する場合すらあります。本展示では、非母語者の認知的負荷を軽減する情報技術の効果的な支援方法について紹介します。

どこが凄い

英語講演の聴講中に、認知的負荷によってワーキングメモリ不足に陥ることが、非母語者の英語理解を阻害する大きな要因です。そこで音声認識技術や機械翻訳技術の力を活用した非母語者のワーキングメモリを拡張するシステムをデザインしました。これにより、聴講内容をより深く理解し易くなりました。

目指す未来

音声認識技術や機械翻訳技術といった最先端の情報処理技術を実際に人間が使い易い形で提供する方法を構築し、より人に優しいICT技術の実現を目指しています。また、母語者と非母語者の自然な協力関係でコミュニケーションの壁を減らすためのデザインを実現し、言語の壁が解消された未来を目指します。



【関連文献】

[1] N. Yamashita, T. Ishida, “Effects of machine translation on collaborative work,” in *Proc. CSCW'06*, pp. 515-524, 2006.
[2] N. Yamashita, A. Echenique, T. Ishida, A. Hautsaari, “Lost in transmittance: How transmission lag enhances and deteriorates multilingual collaboration,” in *Proc. CSCW'13*, pp. 923-934, 2013.
[3] G. Gao, N. Yamashita, A. Hautasaari, S. Fussell, “Improving multilingual collaboration by displaying how non-native speakers use automated transcripts and bilingual dictionaries,” in *Proc. CHI'15*, pp. 3463-3472, Honorable Mention Award, 2015.

【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

A cross-linguistic study of parental input and word acquisition

24 ことばの発達、日本語と英語で何が違う？

～言語間比較で見えてきた親の発話とこどもの語彙獲得の関係～

どんな研究

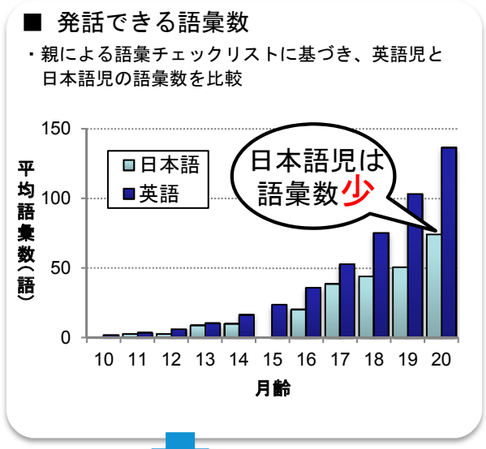
１才半ごろ、日本語児は英語児に比べて発話できる語彙数が非常に少ないのですが、新しい語の学習は早期から正確にできる能力を身につけていることがわかってきました。言語間の違いに注目しながら、親の発話スタイルなどのインタラクション解析を行うことで、言語習得メカニズムの解明を目指しています。

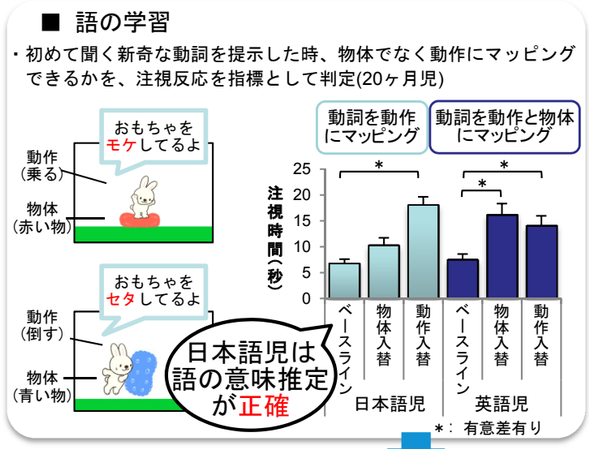
どこが凄い

母親の発話スタイルを解析すると、日本の親は英語の親に比べて育児語などの使用により、多様な語を使用する傾向が強いことがわかってきました。こうした物体や動作に対する一貫性の低い発話スタイルが、こどもの語の獲得や定着に影響を与え、発話できる語彙数の少なさと強く関連している可能性が示唆されます。

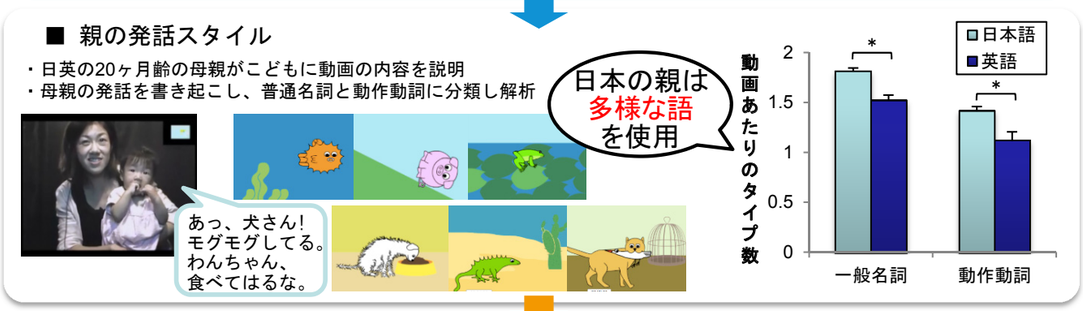
目指す未来

こどもの語彙獲得に与える親の発話スタイルの影響が明らかになれば、コミュニケーション能力の発達を支援する効果的な方法を導き出すことが可能です。こうした方法を、ペアレント・トレーニングなどに応用することにより、育児や教育をサポートするコミュニケーション環境の指針を提案していきます。





日本語児は語の意味推定が正確なのに発話できる語彙数が少ないのはなぜ？



日英比較から親の発話スタイルとこどもの語彙獲得が関連

【関連文献】

[1] Y. Okumura, Y. Oshima-Takane, T. Kobayashi, M. Ma, Y. Kayama, “Japanese children comprehend null object sentences earlier than pronominal object sentences,” in *Proc. ICIS*, 2016.
[2] Y. Oshima-Takane, T. Kobayashi, “Learning novel transitive verbs in Japanese-, French- and English-speaking infants: a cross-linguistic study,” in *Proc. IASCL*, 2014.
[3] T. Kobayashi, Y. Oshima-Takane, “How mothers describe dynamic action events to 20-month-old-childreb,” in *Proc. ICP*, 2016.

【連絡先】

奥村 優子 (Yuko Okumura)　協創情報研究部 インタラクション対話研究グループ
E-mail：okumura.yuko(at)lab.ntt.co.jp

Social telepresence with eye-contact using kinetic avatars

25 離れていても目と目を合わせて会話が出来ます

～キネティックアバタを用いた視線を交わせるテレプレゼンス～

どんな研究

長年、テレビ会議では、対面状況とは異なり、会話相手と視線が合わずコミュニケーションをとりづらいという課題がありました。この展示では、より自然な遠隔コミュニケーションの原理・方式を探るため、従来の会話場再構成システムをリアルタイム化・双方向化した新しい実験システムMMSpaceを紹介します。

どこが凄い

頭部ジェスチャなどの非言語行動をより明確に伝達するため、人の頭部と連動して動くディスプレイ（キネティックアバタ）を用いて、複数人−対−複数人の遠隔双方向・対称型の会話場を構成し、さらにアイコンタクト機構を組み込むことで、対面状況に近い自然な会話ができるようになりました。

目指す未来

将来的には、遠隔にいる人があたかもその場にいるかのように自然なコミュニケーションができる通信環境の実現が期待されます。また、より人と人との共感を醸成するよう、コミュニケーションスキル向上を促すツール、システムの開発にもつながると期待されます。

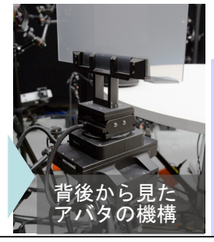
【特徴1】複数人−対−複数人の遠隔会話を対象とし、地点間で人物の位置関係を保つようにアバタを配置し、遠隔地の人の顔を表示することで、対面状況に近い形で，離れた人とも自然に会話が出来ます。

対面状況での複数人会話の場の再現を目指す

【特徴2】高機動（高精度、低遅延、静音）キネティックアバタ（人の頭と連動して動くディスプレイ）により、非言語コミュニケーションが促進されます

半透明のパネルへの投影により、相手の顔が浮かび上がります

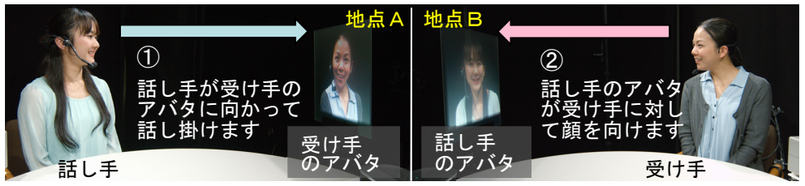
人物の頭の動きと連動して、パネルの姿勢や位置が即座に変わります。（前後左右の移動、首振り、頷きが表現可能）



パネルの動きのあり・なしを比較した結果、動きがある場合の方が

◆より相手の反応が分かりやすいです
◆より相手に伝わった気がします
◆より相手の視線や表情、ジェスチャが分かりやすいです
◆よりアイコンタクトしたと感じます
◆より相手が近くにいるように感じます

【特徴3】アバタを介した擬似的なアイコンタクト−相手の目を見て会話ができます



③
それに応答して、受け手が話し手のアバタを見返すことで、両者のアバタを介した擬似的なアイコンタクトが成立します

【関連文献】
[1] K. Otsuka, “MMSpace: Kinetically-augmented telepresence for small group-to-group conversations,” in *Proc. IEEE Virtual Reality 2016 (VR2016)*, pp. 19-28, March, 2016

【連絡先】
大塚 和弘 (Kazuhiro Otsuka)　人間情報研究部 感覚共鳴研究グループ
E-mail：otsuka.kazuhiro(at)lab.ntt.co.jp

Coding of amplitude modulation in the inner ear

26 分かってきた耳の繊細な働き

～内耳における振幅変調−周波数変調変換システム～

どんな研究

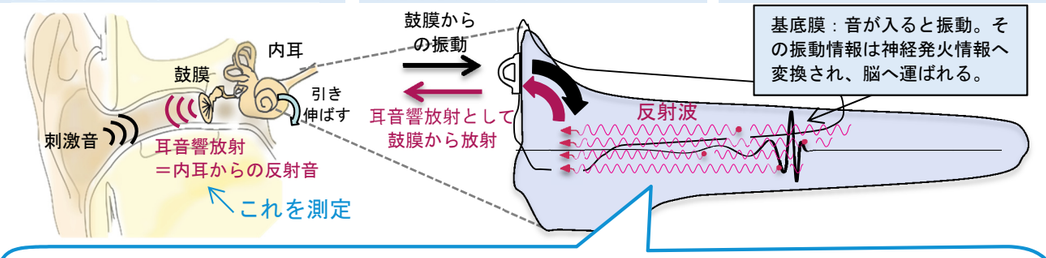
耳、は音を増幅するアンプのような役割を果たしています。私たちは、耳から出てくる音（耳音響放射）を使って、そのアンプの特性を調べてきました。本展示では、従来信じられてきた耳の特性を覆し「耳のアンプは、音を単純に増幅するだけではない」という新たに発見した耳の特性を紹介します。

どこが凄い

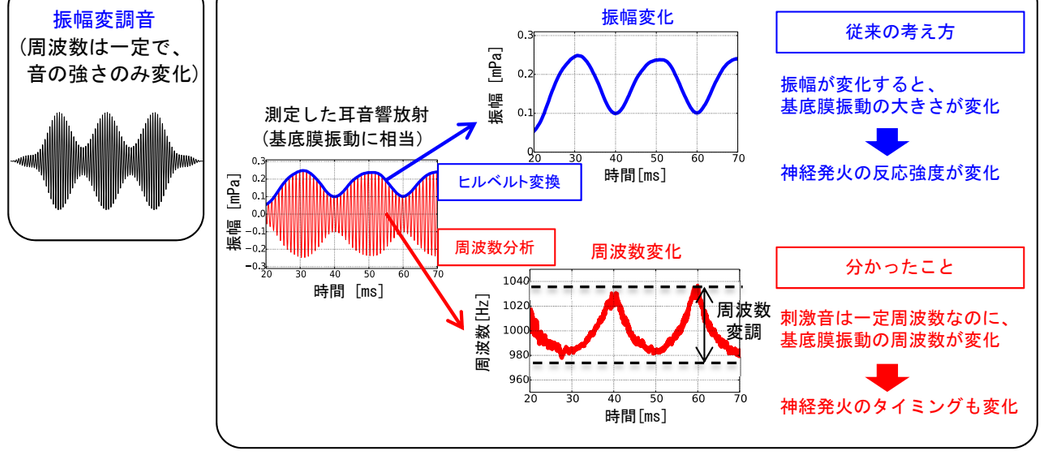
耳は、音の強さの情報を神経発火のタイミングの情報に変換して、脳に伝達することが分かりました。これまで、音の強さは神経発火の反応強度の情報として脳に伝達されていると信じられてきました。本研究の結果は、どのように音の大きさを知覚するか？という疑問に新しい知見をもたらします。

目指す未来

本研究の成果は、加齢などで耳の機能が低下した場合に、どのような聞こえの困難が生じるのかを解明することに繋がります。さらに、人工内耳や補聴器の性能向上などの工学的な分野にも波及効果が期待されています。



［従来の考え方］　音の強さの変化は、神経発火の反応強度の変化として脳に伝達される
［本研究］　音の強さの変化は、神経発火のタイミングの変化としても脳に伝達される



音の強さの変化は、神経発火のタイミングの変化としても脳に伝達されます。知覚実験の結果から、その情報が音の大きさの知覚に貢献していることが分かりました。

【関連文献】
[1] 大塚翔, 古川茂人, “聴覚末梢における振幅変調信号の位相変調による表現: 耳音響放射を用いた検討,” *日本音響学会 秋季研究発表会講演論文集*, 2015.
[2] 大塚翔, 古川茂人, “Conversion of amplitude modulation to phase modulation on the basilar membrane revealed by otoacoustic emission measurements,” *日本音響学会　聴覚研究会資料*, 2015.

【連絡先】
大塚 翔 (Sho Otsuka)　人間情報研究部 感覚共鳴研究グループ
E-mail：otsuka.s(at)lab.ntt.co.jp

27

Evolution and application of “Buru-Navi”

手も心もひかれるモバイルガジェット

～牽引力錯覚装置「ぶるなび」の進化と応用～

どんな研究

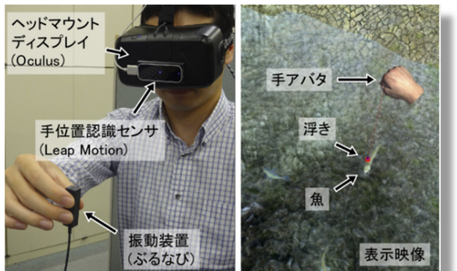
外から引っ張っていないのに引っ張られた感覚を生ずる「ぶるなび」錯覚装置を、把持（はじ）部の表面形状や制御方式を工夫したり、無線化することで、さらに進化させました。これにより、 ゲームなどのカディスプレイや簡易ナビゲーションなど、様々な応用形態が可能になりました。

どこが凄い

非対称振動によって生ずる牽引感覚を分析・研究し、表面形状や振動方法を工夫して、牽引感覚を向上させることに成功しました。また，刺激最適化を行うため、製造メーカとアクチュエータの共同開発を行っています。

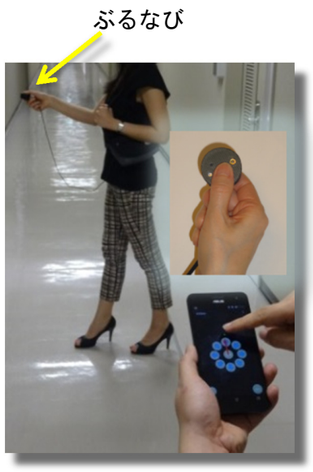
目指す未来

現在のコミュニケーションインターフェースは、音と映像がほとんどです。触覚や力覚情報を人がどのように知覚しているかを解明することにより、音・映像に加え、幅広い触力覚情報を表現する端末装置の開発につなげ、より豊かな情報伝達を実現することを目指しています。



バーチャル魚釣り

タブレットやVRゲームで実現する魚釣りでの，
生き生きとした魚の「引き」を再現



歩行ナビゲーション

手を引く歩行誘導の実現

バーチャル犬の散歩

任意の方位へのダイナミックな力覚生成

感覚感度を増す接触部形状

【関連文献】

[1] 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
[2] T. Teshima, S. Takamuku, T. Amemiya, H. Gomi, “Light touch on pillar array surface greatly improves directional perception induced by asymmetric vibration,” in *Proc. ASIA SIGGRAPH workshop on haptic media and contents design*, 2015

【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：burunavi3(at)lab.ntt.co.jp

28

Sports performance improvement system based on brain science

「スポーツで勝てる脳」をつくる

～脳科学に基づくスポーツ強化システム～

どんな研究

スポーツには心技体が大切と言われます。勝負に勝つには、強い体だけでなく、体を最適に操る技と、精神状態を自在に調節する心が伴わなければなりません。私たちは従来の研究では重視されてこなかった心と技の問題に着目し、新たなスポーツ上達支援法の開発に取り組んでいます。

どこが凄い

勝負には潜在的（自覚できない）脳活動による自動的な動作や意思決定が強く影響します。プレイ中の生体情報を包括的に収集し、抽出したエッセンスに基づき脳を最適な状態へと導くため、ウェアラブルセンサ、機械学習、五感フィードバックなどの技術の活用法を探求しています。

目指す未来

センサを備えた服を着るだけで、自動的に心身状態がモニタされ、最適なフィードバックによる潜在的で直感的な指導が受けられる、といったパーソナルコーチの実現を目指しています。子供から高齢者まで、誰もが自分に適した体の使い方を自然と身に付けられるようになります。

本研究の一部は、国立研究開発法人科学技術振興機構（JST）CREST の支援によって行われています。

SBS SPORTS BRAIN SCIENCE PROJECT http://sports-brain.ilab.ntt.co.jp/

【関連文献】

[1] T. Kimura, T. Mochida, T. Ijiri, M. Kashino, “Body-mind sonification to improve player’s actions in sports,” *NTT Technical Review*, Vol. 14, No. 1, 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, 2016.

【連絡先】

持田 岳美 (Takemi Mochida)　人間情報研究部 感覚運動研究グループ
E-mail：mochida.takemi(at)lab.ntt.co.jp

29

How tracking technologies support caregiving at home

「見守る」と「見守られる」をつなぎます

～介護者記録用Webアプリ「みまもメイト」の開発と評価～

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できる必要があります。本研究では、家族介護者の介護活動を支援する記録用アプリを開発し、2週間のホームユースを通して、その有効性を確認しました。

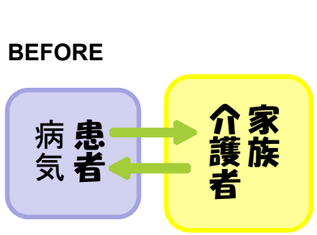
どこが凄い

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。

目指す未来

家族介護者がうつ病患者との適切な接し方を見出すことにより、家族介護者とうつ病患者の人間関係が改善し、家庭内環境が良くなります。これは、うつ病患者の症状緩和にもつながります。社会的な波及効果としては、自殺率の低下や生産活動の拡大が期待されます。

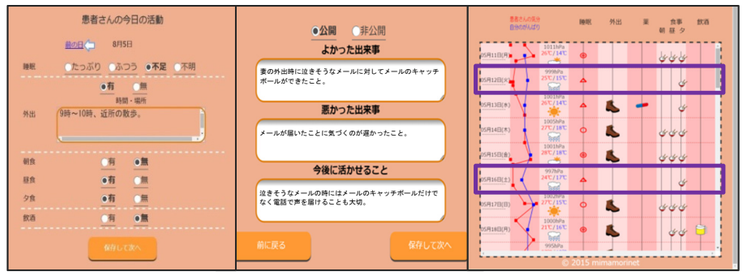
■ みまもメイトの開発



家族介護者は、病気や薬に起因する
患者の想定外の言動に困惑し
どう対応してよいかわからない

・家族介護者が、全ての項目を記入

「みまもメイト」の記録エリアと振り返りエリア



「患者の気分・言動」の記録

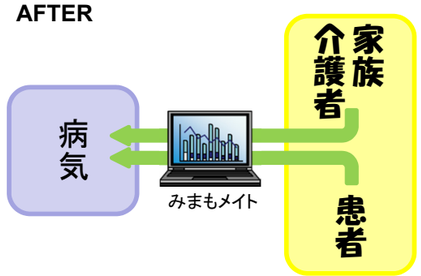
「介護活動」の記録

振り返りを促進

・選択回答方式　　・負の内容に偏らない工夫　　・1つのグラフに全て表示

■ ホームユース調査

AFTER



みまもメイトの記録が家族介護者と患者の
コミュニケーションのきっかけとなり
家族介護者と患者が一丸となって
病気と闘う構図へと変化した

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、定期的に受診

効果 1：効果的な対応方法の発見
- 患者に対する注意力の向上
- 患者の言動に対する理解
- 介護活動の患者への影響

効果 2：患者との人間関係の改善
- 患者に対する視点の変化
- 患者に寛容になること
- 気遣いのある声掛けの増加

【関連文献】

[1] N. Yamashita, H. Kuzuoka, K. Hirata, T. Kudo, “Understanding the conflicting demands of caregivers caring for depressed family members,” in *Proc. ACM Conference on Human Factors in Computing Systems (CHI'13)*, pp. 2637-2646, 2013.
[2] 山下直美, 葛岡英明, 平田圭二, 工藤喬, 荒牧英治, 服部一樹, “みまもメイト：家族介護者のための介護記録用Web アプリの開発とその効果,” *インタラクション2016*, ベストペーパー賞受賞.

【 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

NTT コミュニケーション科学基礎研究所
オープンハウス 2016 実行委員会

（委 員 長）鈴木　潤
（副委員長）北川　智利
（委　　員）石黒　勝彦／黒木　忍／土田　勝

Copyright (C) 2016 NTT Communication Science Laboratories