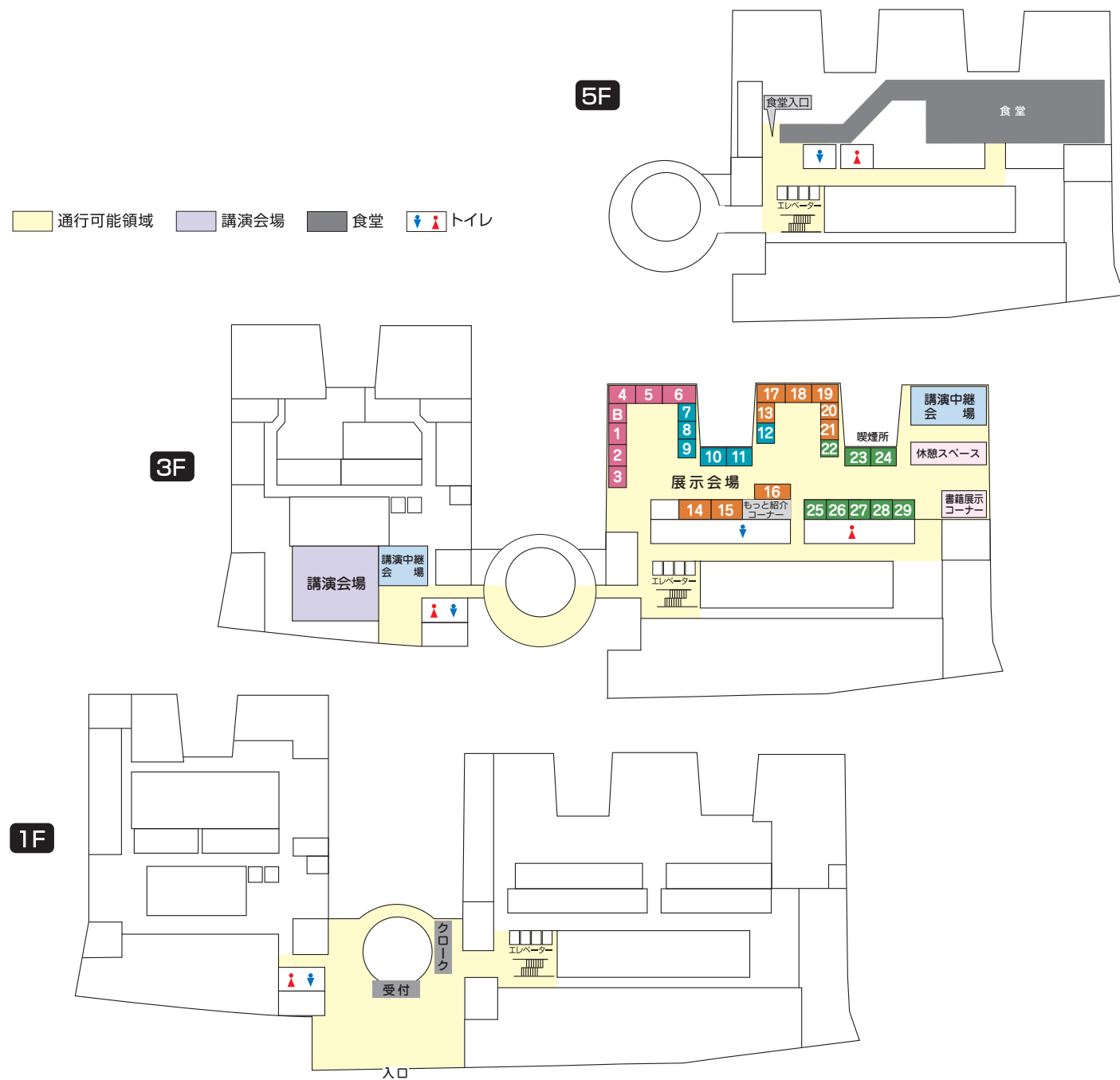




研究展示

ビッグデータ解析

- 8** ビッグデータチャレンジ ～NTT 研究所でのビッグデータ解析の取り組み～
- 1** 大事な特徴を捉えて流れるデータを把握する ～多次元時系列データからの特徴抽出～
- 2** 押し寄せる膨大な「今」を瞬時に賢く分析する ～フロー型ビッグデータを分散オンライン機械学習で分析～
- 3** 大規模なグラフデータをあっという間に分析 ～グラフデータを対象とした高速なマイニングアルゴリズム～
- 4** ビッグデータ分析によるネットワーク異常対策 ～ネットワーク故障、サイバー攻撃などの NW 異常を早期に検出～
- 5** このつぶやきを書いたのはどんなひと？ ～マイクロブログの特徴を活用したユーザ属性推定技術～
- 6** 大規模移動データの分析と可視化 ～異種複合情報を用いた移動パターン分析～

革新的コンピューティング

- 7** 効率的な環境モニタリングを実現 ～相関性を用いたセンサデータ圧縮技術～
- 8** あなたの居場所を安全に公開 ～匿名交換による位置プライバシー～
- 9** 光がつなぐ量子の情報 ～レーザー光を用いた最適な量子もつれ生成方法～
- 10** ちゃんとしたデタラメを作る ～レーザー光の高速乱雑変動を利用した物理乱数生成～
- 11** 高度なプログラミングが可能なビジュアル言語 ～プログラミング上の様々な拡張をした新ビジュアル～
- 12** 計算機による自動英文法解釈 ～統計的文法獲得技術に基づく構文解析～

メディアとコミュニケーション

- 13** 言語の壁をどう越える？ ～英語・中国語・韓国語から日本語への統計翻訳～
- 14** 「触りながら読む」読文を実現：Yu bi Yomu ～なぞり動作を利用した動的な文章表示方式～
- 15** 国際電話会議への参加を楽にする方法 ～0.2 ～ 0.4 秒の通信遅延が非母語者に及ぼす影響～
- 16** せつなの微笑みがこころを伝える ～対話者間の共感 / 反感に関する客観的解釈モデルの提案～
- 17** 録音機器を協調させて目的音声を聞き分けます ～録音ノードからの仮説統合に基づく音声強調～
- 18** 誰がどのように話しても正確に聞き取ります ～話者や発話スタイルの多様性に頑健な音声認識技術～
- 19** 知らない音も書き起こす ～ベイズアン半教師あり音響イベント認識～
- 20** 目的のものが映っているシーンを探す ～静止画クエリによる映像検索技術～
- 21** この子にぴったりな絵本を探す ～幼児単語に重みを置いたグラフベースの類似絵本検索～

ヒューマンサイエンス

- 22** こどもが単語を覚える順番の謎 ～幼児語彙学習期間の言語共通性～
- 23** 「花」と「鼻」、こどもはいつから聞き分ける？ ～幼児の日本語音声知覚の発達～
- 24** 呼吸と音楽の一期一会をめざして ～呼吸の位相と音楽的フレーズを同期させる再生システム～
- 25** それっぽくしゃべります ～非負値時空間分解法に基づく発話リズムの抽出～
- 26** 見ることで感じる疲れや手ごたえ ～映像遅延により変化する運動感覚～
- 27** 未来が過去を変える ～視覚におけるポストディクション～
- 28** 見るだけでモノの形や質感が変わる ～錯視からわかる脳の物体認識のしくみ～
- 29** 聴いて分かる腕の長さ ～身体認識における聴覚の役割～



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



オープンハウス2013



NTT コミュニケーション科学基礎研究所

オープンハウス 2013

6月6日 **木** 12:00～17:30
7日 **金** 9:30～16:00

会場：NTT京阪奈ビル（京都府相楽郡精華町光台2-4）

<http://www.kecl.ntt.co.jp/openhouse/2013/>

NTTコミュニケーション科学基礎研究所



NTT コミュニケーション科学基礎研究所
所長 前田 英作

「オープンハウス2013」へようこそ

NTT コミュニケーション科学基礎研究所「オープンハウス2013」へご来場頂き、心より御礼申し上げます。

私たちの生活は、電話によるコミュニケーションの時代から、クラウド上を飛び交う大量データを手軽な情報機器で利用する新しいコミュニケーションの時代に移りつつあります。そして、扱う情報の量と質の変貌にあわせて、「人間」と「情報」を結ぶ新しい技術基盤の構築が必要です。

そこで私たちの研究所では、情報科学と人間科学の両面からこの問題に取り組んでいます。そして、新しい着想と確かな技術を次の世代へとつなぎ、人類の将来に資することができるような「果実（み）のなる樹木（き）」を育てていきたいと考えています。

オープンハウス2013では、「ビッグデータ解析」、「革新的コンピューティング」、「メディアとコミュニケーション」、「ヒューマンサイエンス」の4領域にまたがる30項目の研究展示と最新の研究動向を紹介する4つの研究講演とにより、当研究所の活動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。

■ 所長あいさつ	1
■ 会場案内図／講演スケジュール	3
■ 所長講演	4
■ 招待講演	5
■ 研究講演	
みんなの会話を聞き取るコンピュータを目指して	6
－複数人会話音声認識研究の現状と今後の展望－	
身体のリアリティ	8
－私たちはどのように自分の身体を認識するのか－	
革新的発展期を迎えた機械翻訳	10
－統計翻訳で越える技術文献の言葉の壁－	
音や声から隠れた情報を取り出す	12
－確率的生成モデルアプローチによる音声音響信号処理－	
■ 研究展示	15

ビッグデータ解析

B	ビッグデータチャレンジ	－NTT 研究所でのビッグデータ解析の取り組み－	16
1	大事な特徴を捉えて流れるデータを把握する	－多次元時系列データからの特徴抽出－	17
2	押し寄せる膨大な「今」を瞬時に賢く分析する	－フロー型ビッグデータを分散オンライン機械学習で分析－	18
3	大規模なグラフデータをあっという間に分析	－グラフデータを対象とした高速なマイニングアルゴリズム－	19
4	ビッグデータ分析によるネットワーク異常対策	－ネットワーク故障、サイバー攻撃などの NW 異常を早期に検出－	20
5	このつぶやきを書いたのはどんなひと？	－マイクロブログの特徴を活用したユーザ属性推定技術－	21
6	大規模移動データの分析と可視化	－異種複合情報を用いた移動パターン分析－	22

革新的コンピューティング

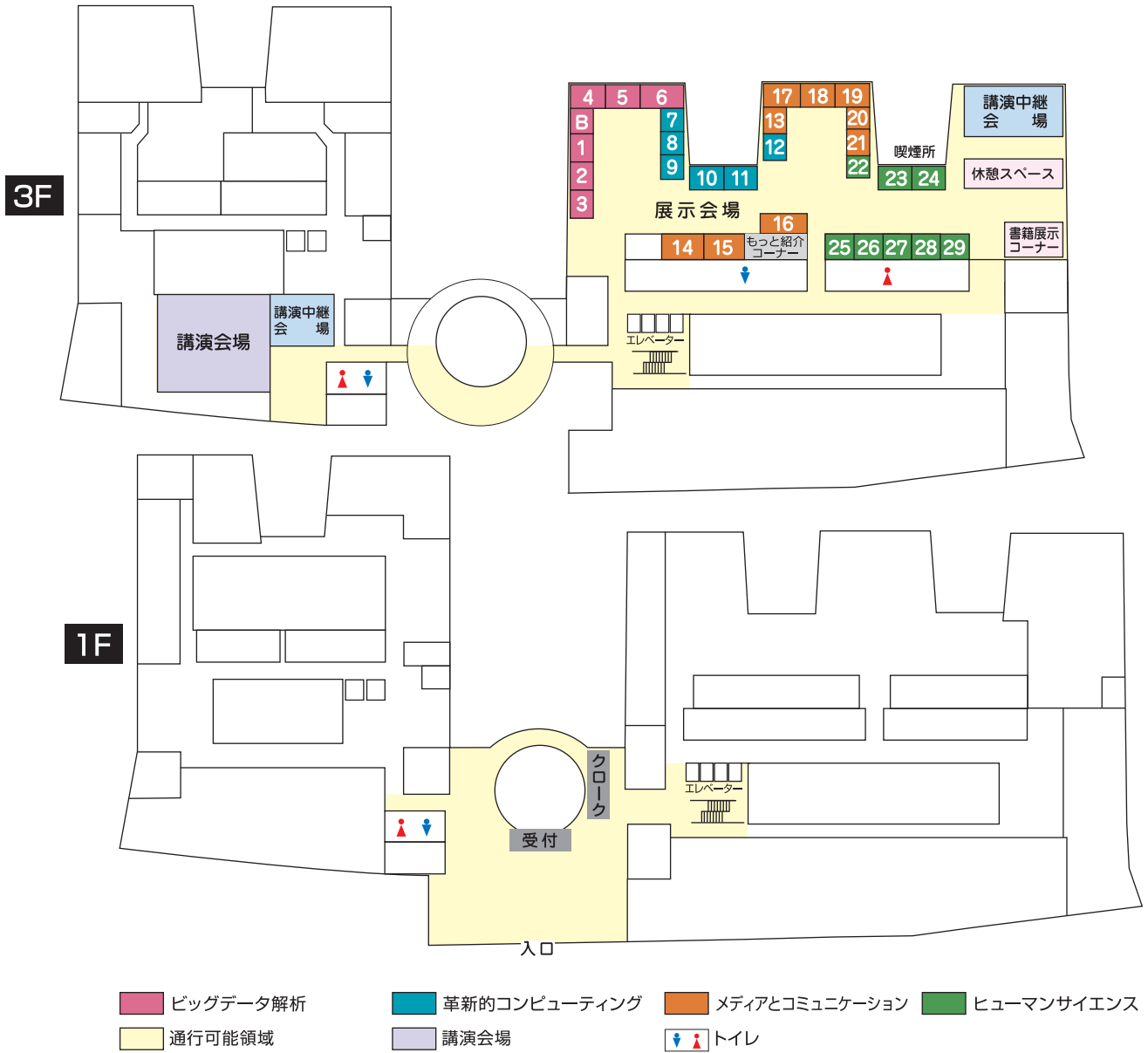
7	効率的な環境モニタリングを実現	－相関性を用いたセンサデータ圧縮技術－	23
8	あなたの居場所を安全に公開	－仮名交換による位置プライバシー	24
9	光がつなぐ量子の情報	－レーザー光を用いた最適な量子もつれ生成方法－	25
10	ちゃんとしたデタラメを作る	－レーザー光の高速乱雑変動を利用した物理乱数生成－	26
11	高度なプログラミングが可能なビジュアル言語	－プログラミング上の様々な拡張をした新ビスケット－	27
12	計算機による自動英文法解釈	－統計的文法獲得技術に基づく構文解析－	28

メディアとコミュニケーション

13	言語の壁をどう越える？	－英語・中国語・韓国語から日本語への統計翻訳－	29
14	「触りながら読む」読文を実現：Yu bi Yomu	－なぞり動作を利用した動的な文章表示方式－	30
15	国際電話会議への参加を楽にする方法	－0.2 ～ 0.4 秒の通信遅延が非母語者に及ぼす影響－	31
16	せつなの微笑みがこころを伝える	－対話者間の共感／反感に関する客観的解釈モデルの提案－	32
17	録音機器を協調させて目的声を聞き分けます	－録音ノードからの仮説統合に基づく音声強調－	33
18	誰がどのように話しても正確に聞き取ります	－話者や発話スタイルの多様性に頑健な音声認識技術－	34
19	知らない音も書き起こす	－ベイズアン半教師あり音響イベント認識－	35
20	目的のものが映っているシーンを探す	－静止画クエリによる映像検索技術－	36
21	この子にぴったりの絵本を探す	－幼児単語に重みを置いたグラフベースの類似絵本検索－	37

ヒューマンサイエンス

22	こどもが単語を覚える順番の謎	－幼児語彙学習期間の言語共通性－	38
23	「花」と「鼻」、こどもはいつから聞き分ける？	－幼児の日本語音声知覚の発達－	39
24	呼吸と音楽の一期一会をめざして	－呼吸の位相と音楽的フレーズを同期させる再生システム－	40
25	それっぽくしゃべります	－非負値時空間分解法に基づく発話リズムの抽出－	41
26	見ることで感じる疲れや手ごたえ	－映像遅延により変化する運動感覚－	42
27	未来が過去を変える	－視覚におけるポストディクション－	43
28	見るだけでモノの形や質感が変わる	－錯視からわかる脳の物体認識のしくみ－	44
29	聴いて分かる腕の長さ	－身体の認識における聴覚の役割－	45



講演スケジュール	
● 6月6日 (木) 所長講演 (13:00～13:30) 果実(み)のなる樹木(き)を育てたい －「情報」と「人間」を結ぶ技術基盤の構築に向けて－ 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 (13:40～14:40) 数学からの新技術、そして産業からの新しい数学 九州大学 副学長 マス・フォア・インダストリ研究所 所長 若山 正人 研究講演 (14:50～15:30) みんなの会話を聞き取るコンピュータを目指して －複数人会話音声認識研究の現状と今後の展望－ メディア情報研究部 堀 貴明	● 6月7日 (金) 研究講演 (11:00～11:40) 身体のリアリティ －私たちはどのように自分の身体を認識するのか－ 人間情報研究部 北川 智利 (13:00～13:40) 革新的発展期を迎えた機械翻訳 －統計翻訳で越える技術文献の言葉の壁－ 協創情報研究部 永田 昌明 (13:50～14:30) 音や声から隠れた情報を取り出す －確率的生成モデルアプローチによる音声音響信号処理－ メディア情報研究部 亀岡 弘和

果実(み)のなる樹木(き)を育てたい
～「情報」と「人間」を結ぶ技術基盤の構築に向けて～

Cultivate trees that will bear fruit
Building a technical infrastructure connecting information and human

NTT コミュニケーション科学基礎研究所 所長 前田 英作

21世紀に入って、私たちの生活を取り巻く情報環境は目まぐるしく変化しています。そして何より、その変化のスピードがとてつもなく速いことに特徴があります。ユビキタス、グリッド、センサネット、セマンティックWEB、WEB2.0、クラウド、ビッグデータなどこの10年間に喧伝されてきたキーワードの変遷からもそのことは明らかでしょう。また、ネットワークにアクセスする情報機器も携帯・デスクトップからスマホ・タブレットへと移り、その利用者層も子供からお年寄りまで広がっています。この変化に合わせて、提供するサービスも異なる様相を呈するとともに、環境の変化に合わせた研究開発が求められています。

コミュニケーション科学基礎研究所は、この情報環境の劇的な変化の中において、「情報」と「人間」を結ぶ新しい技術基盤の構築をめざしています。基礎研究は、時代のニーズに即応することが求められるサービス開発と異なり、中長期的な視点に立った技術革新を求められます。しかし、変化の速度が増している中で、基礎研究の姿・形も10年、20年前のものとは当然違ってしかるべきでしょう。

些細な発想の転換や問題の発見が新しい研究の種になります。そこに水を撒いてあげるとしばらくして芽が出てくるものがあります。さらに、肥料と陽の光を与えるといずれ論文や特許として花が咲きます。しかし花は受粉のための誘引道具であって、普通は食べることができません。やがて花は果実となり、摘み取ることができるようになります。それでも、収穫した果実がそのまま食べられるとは限りません。硬く

熟していないものもあれば、酸っぱいものや、毒のあるものさえあります。料理の素材として姿を変えて使われることも少なくありません。さらに、貯えて寝かせておく場合もあります。果実から種を採って次の世代に引き継ぐこともある一方で、敢えて種のない果実を育てることもあるでしょう。

いずれにしても、果実が何らかの形で人々の口に入り滋養となることによってはじめて、研究の成果が技術として使われ、世の中の役に立ったのだと言うことができます。この果実(み)のなる樹木(き)を育てる営みこそが、基礎研究所の役割であり、昔から連綿と行ってきた研究開発そのものに他なりません。コミュニケーション科学基礎研究所の創立以来20年の余を経て、実社会の中へ入っていった基礎研究発の技術が少しずつ生まれています。メディア探索、音声認識、残響制御、質問応答、機械翻訳、質感情報学など、そうした成功例を分析してみると、実は、種を蒔いてから10年の時間を要していることがわかります。

情報通信環境の激しい変化を見据えた時、この10年という時間を皆さんはどのようにとらえるでしょうか。果実(み)のなる樹木(き)をたくさん育てていくために、何を変え、何を守るべきなのか。将来に向けて基礎研究所が今すべきことは何なのかを考えなければなりません。本年のオープンハウスで紹介する一つ一つの研究展示からどのような果実が生まれようとしているのか、是非見極めて頂きたいと思います。

数学からの新技術、そして産業からの新しい数学

New technology from/with mathematics and new mathematics from industry



九州大学 副学長 マス・フォア・インダストリ研究所 所長

若山 正人
(わかやま まさと)

(専門分野)

リー理論、整数論

(プロフィール)

1978年 東京理科大学理学部数学科卒業
1985年 広島大学大学院理学研究科博士後期課程数学専攻修了
1994年 九州大学理学部数学科助教授
1997年 九州大学大学院数理学研究科教授
2006年～2010年 九州大学大学院数理学研究院院長
2008年より 文部科学省グローバルCOEプログラム「マス・フォア・インダストリ教育研究拠点」拠点リーダー
2010年より 九州大学副学長
2011年より 九州大学マス・フォア・インダストリ研究所所長
2011年より 文部科学省学術審議会先端科学部会数学イノベーション委員会主査
この間、Princeton大学客員研究員、Illinois大学、Bologna大学、KIAS客員／訪問教授、Indiana大学 Distinguished Lecturer等を歴任。

(主要研究テーマ)

- (1) 非可換調和振動子のスペクトルと付随するゼータ関数・楕円曲線・保型形式の研究
 - (2) 確率論・統計学に現れる α 行列式の表現論・不変式論
 - (3) 対称空間上の調和解析と統計多様体の群論的研究
 - (4) コンピュータグラフィックスのリー理論的研究
- 編・著書として『ゼータの世界』(日本評論社、1999年)、『絶対カシミール元』(岩波書店、2002年)、『技術に生きる現代数学』(岩波書店、2008年)、『Casimir Force, Casimir Operators and the Riemann Hypothesis』(De Gruyter、2010年)等。



みんなの会話を聞き取るコンピュータを目指して ～複数人会話音声認識研究の現状と今後の展望～

Towards a computer that can recognize everyone's conversations
Research on multi-speaker conversational speech recognition: past, present and future

メディア情報研究部 堀 貴明

■音声認識技術の進歩

近年、声で操作できるスマートフォンが登場するなど、私たちが音声認識技術に触れる機会が増えてきました。音声認識とは、人が話した音声をコンピュータ等で解析し、その内容を言語表現（文字データ）として取り出す処理を言います。研究が始まった1950年代より、音声認識は人の言葉を理解する技術として注目を集めてきました。しかし音声は複雑な性質を持ち、かつ膨大な言葉の組み合わせで成り立っていることから、当時のコンピュータでは扱えない課題が多く、その研究は困難を極めました。そのような困難を乗り越え、今日までにいくつかの画期的な音声モデリング手法や効率的なアルゴリズムが開発され、これにコンピュータの進歩が加わり、音声認識技術は徐々に実用レベルへと発展しました。

■会話を認識することの難しさ

スマートフォン等の情報端末に話す内容は、多くの場合、端末を操作する命令や検索のキーワードです。時々誤った答えを返しますが、丁寧に話せば成功する確率は高まります。一方、端末の操作ではなく、私たちの日常の会話を認識したい、というニーズがあります。これは会議の議事録作成など、従来、人が行ってきた高度な聞き取り作業をコンピュータで自動化するものです。ところが、

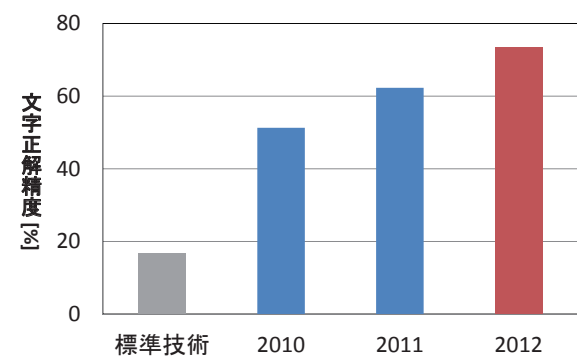


図1: 自由会話の音声認識ベンチマーク (参加者4名、オフィス環境)

現在の標準的な技術で会話を認識させると、話した言葉の半分も正しく認識できません。これは、端末を操作する発声と人に話す会話とでは、その性質が大きく異なることを意味します。そして、技術的には会話の認識の方が遥かに難しく、これには次の2つの理由があります。

①会話音声の多様性

会話の中で人は意識して丁寧に話そうとはせず、発音を怠けたり、言葉を省略する傾向が強まります。また、方言等のなまりや喜怒哀楽の感情が入ることもあります。したがって、発音表記上は同じ音であっても、話す内容、話し方、話す人によって音響的特徴は大きく異なります。さらに言語表現は話す相手、話題、状況によっても変化します。このような多様性が音声認識の性能を著しく低下させます。

②マイクを意識しない会話

スマートフォンに話す場合と異なり、対面の会話ではマイクを目の前に置いたり、身体に装着することは、会話中の行動が制限されるのであまり好まれません。よって少し離れた場所にマイクを置くことが大切です。しかし、離れたマイクでは周囲の雑音が相対的に大きくなり、かつ複数の人の声が混ざって収録されてしまいます。これも音声認識の性能を著しく低下させます。

■会話音声認識の研究

コミュニケーション科学基礎研究所では、会話の音声認識を重点課題の一つと位置づけ、精力的に研究してきました¹⁾²⁾。従来の音声認識では、音韻の普遍的な特徴を表すモデルを作り、これを入力音声と照合して解を求めています。しかし、この普遍的なモデルでは会話音声の多様な音響的特徴を捉えることは困難です。そこで私たちは、話す内容や話し方による音韻の変動を表現できる新しいモデルを考案し、会話音声の認識精度を大きく向上させました。さらに、個人の声質や話題にモデルを適応させて精度を向上させる技術も開発しました。一方、マイクを意識しない会話を実現するため、私たちは、複数のマイクにより推定される音の到来方向の情報と、予め学習した音声の音響的かつ時間的な変化パターンに基づいて、一人一人の音声と雑音とをきれいに分離する技術を開発しました。これらの成果により、会話音声認識の精度は着実に向上してきました（図1）。そして、近年、会話音声認識の可能性を示す目的で、会話中に「いつ誰が何を話したか」をリアルタイムで認識する「会話分析システム」（図2）を構築しました。このシステムは、テーブル中央のマイクで会話を収録し、一人ひとりの発言内容を即座に表示します。このような会話音声認識が可能なシステムは世界的にも例がありません。

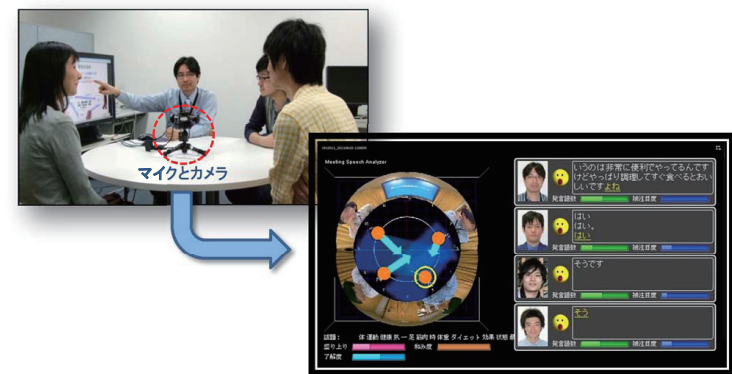


図2: リアルタイム会話分析システム



堀 貴明 (ほり たかあき)

NTT コミュニケーション科学基礎研究所 メディア情報研究部 主任研究員

1999年、山形大学大学院工学研究科システム情報工学専攻博士後期課程修了。同年、日本電信電話株式会社入社。博士（工学）。音声認識、音声言語処理の研究に従事。IEEE、電子情報通信学会、日本音響学会、各会員。



■今後の展望

会話の音声認識が可能になれば、様々な応用の可能性が広がります。将来は会議の議事録作成だけでなく、過去の記憶を呼び起こすように会話を閲覧できるかもしれません。会話中に質問に答えたり、今後の予定を登録してくれる電子秘書が登場するかもしれません。そのためには認識精度の向上は必須ですが、それ以外に「誰が誰に話しているか」「どのように話しているか」といった文字には表れない情報の認識、さらには認識された情報を統合・理解する技術が重要になります。私たちは、このような課題に従来の音声認識技術の枠を超えて取り組んでいます。

【参考文献】

- 1) T. Hori, S. Araki, T. Yoshioka, M. Fujimoto, S. Watanabe, T. Oba, A. Ogawa, K. Otsuka, D. Mikami, K. Kinoshita, T. Nakatani, A. Nakamura, J. Yamato, "Low-latency real-time meeting recognition and understanding using distant microphones and omni-directional camera," *IEEE Trans. ASLP*, Vol. 20, No. 2, pp.499-513, 2012.
- 2) 堀貴明, 小川厚徳, 藤本雅清, 大庭隆伸, 久保陽太郎, 威聖俊, 荒木章子, ソウデン・メレズ, デルクロー・マーク, 吉岡拓也, 木下慶介, 中谷智広, 中村篤, “会話分析タスクにおける複数人自由会話音声認識の改善,” 日本音響学会秋季全国大会, 1-1-19, 2012.

身体の実在性 ～私たちはどのように自分の身体を認識するのか～

The reality of the body
How we perceive our own body

人間情報研究部 北川 智利

私たちは、自分の身体がどんな大きさで、どんな形をしているか知っていますし、「自分の身体はこんなふうだ」と身体を感じることもできます。これを本稿では「身体の認識（身体感覚）」と呼びます。身体を認識していなければ、私たちは自分の身体を適切にコントロールすることができないでしょう。みなさんが感じている自分の身体の形は、いつも同じ形で安定していると思います。ところが、自分が感じている身体の形状が、わりと簡単に変化することがあると聞いたらどう思われるでしょう。これは姿勢が変わるという話ではありません。身体の高さや長さが変わってしまうという話です。ちょっとした細工をしてやると、身体そのものの形はそのままなのに、自分が感じている身体の形は変わってしまうことがあります。鼻や腕が伸びたり、ゴム手袋に詰め物をしたものをまるで自分の手のように感じたり、自分の身体を外側から眺めているように感じることもさえます。そんな身体に関わる錯覚から、私たちがどのように身体を認識しているかが分かってきます。

身体に関わる錯覚の代表格とも言えるのがラバーハンド錯覚です¹⁾。図1では、奥の実験参加者の右手は見

えないように隠されていて、右手の模型がその正面に置かれています。手前の実験者は実験参加者の手と、手の模型の同じ場所を同時に繰り返し叩きます。実験参加者は、自分の手が叩かれる感触を感じながら、手の模型が同じように叩かれるのを見ることになります。数分間この状況を経験していると、叩かれている感触が手の模型からやってくるように感じ、さらに、手の模型が自分の身体の一部であるように感じてしまいます。この錯覚が生じている時に手の模型にナイフで脅すなどの危害を加えると、自分の手がそうされた時と同じような生理反応（発汗）が生じるのだそうです。叩かれる自分の手からくる触覚情報、同じように叩かれる模型の手という視覚情報、両者が与えられた時、私たちの知覚システムは「模型の手が自分の手に違いない」と解釈して触覚情報と視覚情報の矛盾を解消するのです。私たちの身体の認識は、視覚や触覚、自己受容感覚（姿勢の感覚）などの感覚情報に基づいて、常に更新されています。これらの感覚の間に矛盾があると、私たちの「これが自分の身体だ」という認識は、いとも簡単に変化してしまうのです。

これまで、身体を認識するのに重要な感覚は、触覚や



図1：ラバーハンド錯覚

運動感覚などの体性感覚と視覚だと考えられてきました。しかし我々の研究グループでは最近、聴覚も身体の認識に貢献していることを発見しました²⁾。図2のように、床をたたく音が実際より遠くから聞こえるような状況をつくります。体性感覚情報と聴覚情報の間に矛盾をつくるわけです。このような状況を経験すると、私たちの知覚システムはこう解釈するかもしれません。「音が出ているのは自分の右手が叩いているところ」＝「自分の右手は遠くを叩いている」→「その場所を叩いているならば自分の右腕は長いはず」。実験を行って調べみると、右腕で触れた物体の長さは、この状況を経験した後では、より長く感じられることが分かりました。つまり、右腕は実際よりも長く認識されていたのです。

このような身体の認識に関わる錯覚は私たちに重要な示唆を与えます。私たちが認識している身体は常にも柔軟に更新されているということです。ですから、うまく操作することで任意の身体感覚を生じさせることができるかもしれません。任意の身体感覚を伴わせることでバーチャルリアリティ空間の中での体験をよりリアルなものにすることもできるでしょう。身体認識の変化は、使い慣れた道具を自分の身体の一部とを感じる感覚とも

関連があると言われています。身体感覚は空間的に、時間的にどこまで拡張できるのでしょうか。空間的にも時間的にも離れた月面のロボットを自分の身体の一部のように認識し操作するにはどうしたら良いのでしょうか。今は答えることができませんが、今後の研究によって答えが出てくると期待しています。

私たちが自分の身体を認識するということは、私たちが「自分を自分である」と認識する、つまり自己という概念を持つことの基礎になっています。そして、自己の概念は他者を理解するためにも必要です。我々の研究グループでは、身体を認識するメカニズムを明らかにすることで、より良く、より深い、未来のコミュニケーションの実現につなげていきたいと考えています。

【参考文献】

- 1) M. Botvinick and J. Cohen, "Rubber hands 'feel' touch that eyes see," *Nature*, Vol. 391, 756, 2008.
- 2) A. Tajadura-Jiménez, A. Vajamäe, I. Toshima, T. Kimura, M. Tsakiris and N. Kitagawa, "Action sounds recalibrate perceived tactile distance," *Current Biology*, Vol. 22, R516-R517, 2012.



図2：音のフィードバックで腕が伸びる錯覚



北川 智利 (きたがわ のりみち)

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

2000年、東京都立大学大学院人文科学研究科単位取得退学。博士（心理学）。東京都立大学人文学部助手、日本学術振興会特別研究員（SPD）、金沢工業大学情報フロンティア学部講師、NTTコミュニケーション科学基礎研究所リサーチスペシャリストなどを経て現職。感覚間相互作用や感情と知覚の相互作用の研究に従事。日本心理学会、日本基礎心理学会、日本音響学会、各会員。

革新的発展期を迎えた機械翻訳 ～統計翻訳で越える技術文献の言葉の壁～

Innovative user experiences brought by statistical machine translation
Breaking the language barriers of technical information

協創情報研究部 永田 昌明

英語と日本語は語順の対応関係が正順になる部分と逆順になる部分が複雑に入り混じっており、恐らく世界で最も翻訳が難しい言語対の一つです。我々は、主辞後置性という日本語の言語学的な特徴を利用して、英語の文の単語を日本語と同じ語順に並べ替えてから日本語へ翻訳する方法を考案し、英日翻訳の精度を劇的に改善しました¹⁾。また翻訳の精度を客観的に計測することは実はとても難しい問題なのですが、我々は、日本語と英語の翻訳において人手による評価とよく一致する新しい自動評価尺度RIBESを考案しました。本講演では、これらの統計的機械翻訳の研究成果の概要をご紹介します。

コンピュータを利用してある言語を別の言語に翻訳する技術を「機械翻訳」(machine translation)と呼びます。機械翻訳の研究はコンピュータの誕生とほぼ同時に1950年代から始まり、今日までに多数の機械翻訳システムが開発されました。

インターネットの普及により、ごく普通のユーザが英語・中国語・韓国語などの外国語で書かれたWebページに接する機会が飛躍的に増えました。世界中に市場を持つ多国籍企業はマニュアルなどの製品情報を迅速かつ正確に現地語で提供する必要があります。言葉の壁を越えて人とコミュニケーションしたい、あるいは、言葉の壁を越えて知識を交換したいという想いは人類の普遍的な要求の一つと言ってよいでしょう。しかし、正直なところ従来の機械翻訳がこれらの様々なユーザの要求を満足させてきたとは言い難いです。

従来の機械翻訳システムでは、新しい言語間の翻訳を実現するために、数人から数十人の専門家が何年もの歳月をかけて翻訳規則や対訳辞書を作成していました。このような機械翻訳へのアプローチを「ルールベース翻訳」と呼びます。これに対して「統計的機械翻訳」

(statistical machine translation, 以下、統計翻訳)は、数十万から数百万文の大規模な対訳データから翻訳規則や対訳辞書に相当する統計モデルを自動的に学習し、新しい言語対や特定の分野の機械翻訳システムを短期間に低コストで作成する技術です。

統計翻訳は1990年前後に英語とフランス語の翻訳を対象として研究が始まり、2000年代には英語とフランス語のような語順が近い言語対では実用レベルに到達しました。しかし、英語と日本語のように語順が大きく異なる言語対における統計翻訳の精度は低く、従来のルールベース翻訳の精度を上回ることができませんでした。

2010年前後から世界の主要な研究機関は、語順の違いを克服する技術として、原言語(翻訳元)の文の単語を目的言語(翻訳先)の語順に並べ替えてから統計翻訳を行う「事前並べ替え」(preordering)に注目し始めました。その中でNTTは、主辞後置性と呼ばれる日本語の言語学的な特徴に着目し、「主辞を後置する」(head finalization, 主辞後置化)という唯一の規則を使って英語の単語を日本語の語順に並べ替える方法を考案し、英日翻訳の精度を劇的に改善しました¹⁾。この方法の概要を図1に示します。

主辞(head)とは、文を構成する部品である句(phrase)において句の文法的な役割を決める単語です。例えば前置詞句なら前置詞が主辞です。別の言い方をすると、中学校の国語で習う「係り受け」において係り

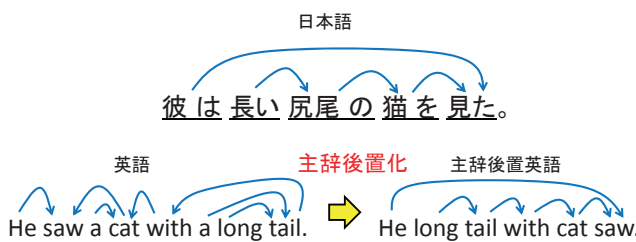


図1：日本語の主辞後置性に基づく事前並べ替え翻訳

先になる単語が主辞です。日本語には「必ず前から後ろに係る」すなわち「必ず主辞を後置する」という性質(主辞後置性)があります。世界の言語の中でこれは非常に稀な性質で、通常は図1の英語の例のように「前から後ろ」にも「後ろから前」にも係ります。したがって、英語の係り受け構造に基づいて「必ず前から後ろに係る」ように単語の順番を入れ替えると、日本語と同じ語順に変換できます。これが主辞後置化です。

我々はこの主辞後置化に基づく事前並べ替え型の統計翻訳技術により2011年に開催された特許翻訳に関する評価側ワークショップNTCIR-9の英日翻訳において史上初めて人手による評価でルールベース翻訳を上回ることができました²⁾³⁾。

日本語の主辞後置性を利用した統計翻訳技術は、原理的には英語以外の言語へも適用可能であり、我々は、特許のように百万文を超える対訳データを入手可能な分野では統計翻訳は実用レベルに到達したと考えています。今回のオープンハウスでは、研究展示「言葉の壁をどう超える? ～英語・中国語・韓国語から日本語への統計翻訳～」において特許分野における我々の技術的な到達点をご覧くださいことができます。

つぎに、翻訳精度の自動評価について簡単にご説明します。機械翻訳の精度を客観的に計測することは実は非常に難しい問題です。ある文に対する翻訳の正解は何通りもあり、訳語選択の誤りと語順の誤りのどちらを重視するかは主観的な判断です。1990年代に考案された機械翻訳の自動評価尺度BLEUは、機械翻訳の研究の活性化に大きく貢献しました。これは食味計の発明によりお米の「おいしさ」を数値化することで、産地間の競争や品種改良が進むのに似ています。しかし、BLEUは日本語と英語の翻訳では人手による評価とあまり一致しないという問題がありました。2008年に開催された評価側ワークショップNTCIR-7における自動評価尺度



永田 昌明 (ながた まさあき)

NTT コミュニケーション科学基礎研究所 協創情報研究部 主幹研究員(上席特別研究員)

1987年、京都大学大学院工学研究科修士課程修了。同年、日本電信電話株式会社入社。工学博士。(株)国際電気通信基礎技術研究所、米国AT&T研究所客員研究員などを経て現職。統計的自然言語処理の研究に従事。電子情報通信学会、情報処理学会、人工知能学会、言語処理学会、ACL、各会員。

BLEUと人手による評価の相関関係を図2に示します。

そこで我々は、翻訳結果と正解の語順の一致度を重視する自動評価尺度RIBES (Rank-based Intuitive Bilingual Evaluation Measure) を考案し、そのプログラムを一般公開しました。RIBESは、前述のNTCIR-9において自動評価尺度の一つとして採用され、主催者の評価においても、英日・日英・中英翻訳タスクにおいて人間による評価との相関がBLEUより高いことが示されました²⁾。

現在、我々は、特許、マニュアル、科学技術論文などの技術文書では、言葉の壁を意識しない情報収集が実現可能になったと考えています。今後は、言葉の壁を意識しない情報発信、すなわち日本語から他言語への翻訳へ取り組むとともに、技術文書からビジネス文書、話し言葉へと翻訳対象を広げていきたいと考えています。

【参考文献】

- 1) H. Isozaki, K. Sudoh, H. Tsukada, K. Duh, "HPSG-based Preprocessing for English-to-Japanese Translation," *ACM TALIP*, Vol. 11, No. 3, 2012.
- 2) I. Goto, B. Lu, K. P. Chow, E. Sumita, B. K. Tsou, "Overview of the Patent Machine Translation Task at the NTCIR-9 Workshop," in *Proc. NTCIR-9*, pp.559-578, 2011.
- 3) K. Sudoh, K. Duh, H. Tsukada, M. Nagata, X. Wu, T. Matsuzaki, J. Tsujii. "NTT-UT Statistical Machine Translation in NTCIR-9 PatentMT," in *Proc. NTCIR-9*, pp.585-592, 2011.

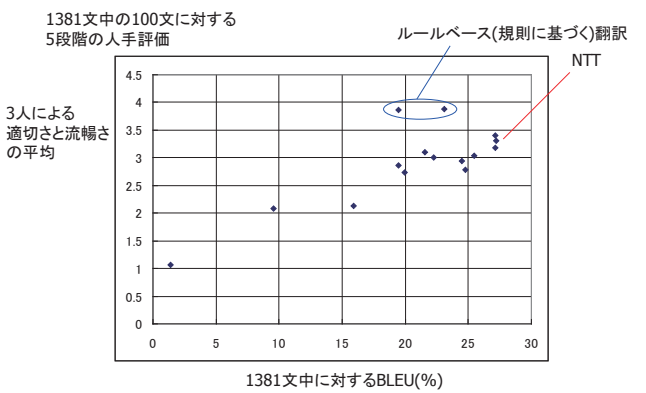


図2：NTCIR-7 日英特許翻訳タスク(2008)のBLEUと人手評価の関係

音や声から隠れた情報を取り出す ～確率的生成モデルアプローチによる音声音響信号処理～

Extracting hidden information from speech and audio signals
Generative modeling approach to speech and audio signal processing

メディア情報研究部 亀岡 弘和

日ごろ私たちが耳にしている音や声には実に多くの情報が含まれており、私たちはこれを助けにしながら日々生活をしています。例えば、カフェテリアにいと、店内がどれだけ賑わっているか、周囲の人たちがどのような会話をしているか、店内にどのような音楽が流れているか、などの状況を把握することができます。両耳に入ってくる音の波形には、混在する個々の音がどのような音であるかという情報やどの方向から到来したかという情報が含まれます。また、一人一人の話し声には、話し手の発話内容（言語情報）や言い回しの中に表れる話し手の個性や意図や思いなどの情報（パラ言語・非言語情報）が含まれます。さらに音楽には楽器構成、メロディ、リズム、和声などの情報が含まれます。私たち人間は音や声からこのようなさまざまな種類の情報を無意識的ないし意識的に認識し、理解する能力を有していますが、計算機を使って音の波形（音響信号）からこれらの情報を自動的に取り出すことができるようになれば、例えば、音声認識や合成、話者認識、自動採譜などのシステムに活かすことができるだけでなく、音の定位による屋内や屋外の安全モニタリングシステム、語学支援システム、パラ言語・非言語情報に着目した音声変換システム、ユーザの意を汲む音声対話システム、類似音楽やカバー曲の検

索システム、音楽の盗作モニタリングシステムなどの実現に道が拓け、人々の生活を大いにサポートできたり豊かにできたりするはずです。

複数の音が混合されるプロセス、話し手が意図した言葉を音声波形として発するプロセス、発話において各音節のタイミング・アクセントの有無・プロミネンス（強調の度合い）などの韻律情報を音声の基本周波数パターンとして表現するプロセス、楽器演奏において楽譜をイメージしながら演奏表情をつけて奏でるプロセス。これらは、いずれも何らかの「原因」によってある「結果」が得られる順問題と見なせます。したがって、その逆プロセス、つまり、混合音を個々の音に分離する問題、音声波形から言語情報を認識する問題、基本周波数パターンから各種韻律的特徴を抽出する問題、音楽の音響信号を個々の楽器音に分解する問題、音楽演奏から楽譜を書き起こす問題は、「結果」から「原因」を推定する逆問題と捉えることができます（図1）。「結果」から「原因」が一意に特定できる場合には逆問題の解決は比較的容易ですが、例えば $a+b=10$ となる a と b を特定することができないように、上述の逆問題においては、同じ「結果」を与える「原因」が無数に存在しうるため、正しい「原因」を特定するのは容易ではありません。ところが、私たち

人間にとっては、混合音から音を聴き分けたり、声から言葉や話者を認識したり、言い回しから話し手の癖や意図を感じ取ったり、演奏を聴いて楽譜化したりすることはさほど難しいことではありません。人間が上述の逆問題をどのような認知プロセスを経て解決しているのかは必ずしも明らかではありませんが、経験や知識を通じて得られた常識や傾向が、逆問題の解を絞り込むためのヒントとなっている可能性があります。

我々はこれまで、音や声に隠れた重要な情報を抽出することを目的とした以上のような逆問題に対し、確率モデルに基づくアプローチの検討を進めてきました。すなわち、観測量である「結果」を生成する「原因」となっている背後のプロセスを確率モデルによって巧みに記述し、そのモデルをもとに観測量を最も良く説明する「原因」を推論する、というアプローチです。図2は、音源分離問題のために文献1)で提案した確率モデルの構造の例を示したものです。グレーのノードが観測量、白いノードが確率モデルの変数に対応し、それぞれの矢印がノード間の依存関係を表しています。これはあくまで例ですが、確率モデルの構造をどう立てるかが確率モデルアプローチにおける最大の腕の見せ所で、観測量とパラメータ（変数）の因果関係を現実の生成プロセスに即した確率モデルで適切に記述することができれば、逆問題の解に対する先験的知識や傾向をヒントにした効果的な推論が可能になります。一方、確率モデルアプローチにおいてはパラメータの推論が簡潔かつ効率的なアルゴリズムで行えるかどうかという点も重要です。現実の生成プロセスに即したモデルかどうかという点とパラメータ推論が効率的に行えるモデルかどうかという点の両面を考慮して適切な確率モデルを立てることが逆問題解決の成功の鍵となります。

本講演では、音源分離¹⁾、音声強調²⁾、音声韻律特徴抽出³⁾、多重音解析^{4),5)}などの問題を題材にし、上記の方法論をベースにした研究アプローチを紹介します。

【参考文献】

- 1) H. Kameoka, M. Sato, T. Ono, N. Ono, S. Sagayama, "Blind separation of infinitely many sparse sources," in *Proc. IWAENC 2012*, 2012.
- 2) H. Kameoka, T. Nakatani, T. Yoshioka, "Robust speech dereverberation based on non-negativity and sparse nature of speech spectrograms," in *Proc. ICASSP2009*, pp.45-48, 2009.
- 3) H. Kameoka, J. Le Roux, Y. Ohishi, "A statistical model of speech F0 contours," in *Proc. SAPA 2010*, pp.43-48, 2010.
- 4) H. Kameoka, N. Ono, K. Kashino, S. Sagayama, "Complex NMF: A new sparse representation for acoustic signals," in *Proc. ICASSP2009*, pp.3437-3440, 2009.
- 5) H. Kameoka, M. Nakano, K. Ochiai, Y. Imoto, K. Kashino, S. Sagayama, "Constrained and regularized variants of non-negative matrix factorization incorporating music-specific constraints," in *Proc. ICASSP2012*, pp.5365-5368, 2012.

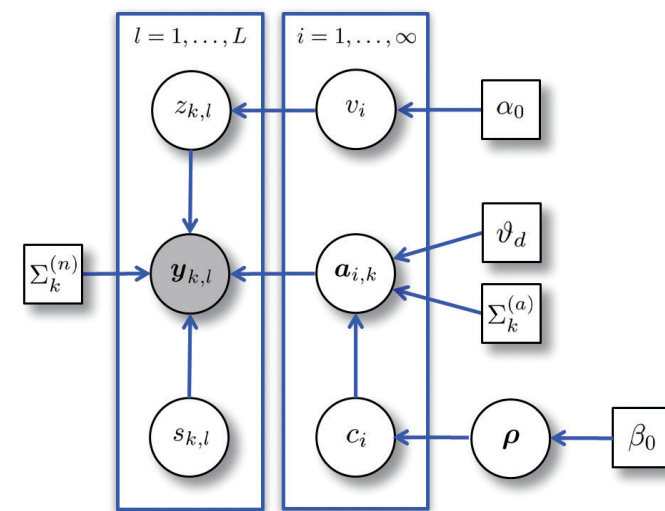


図 2：音源分離問題のための確率モデルの構造の例

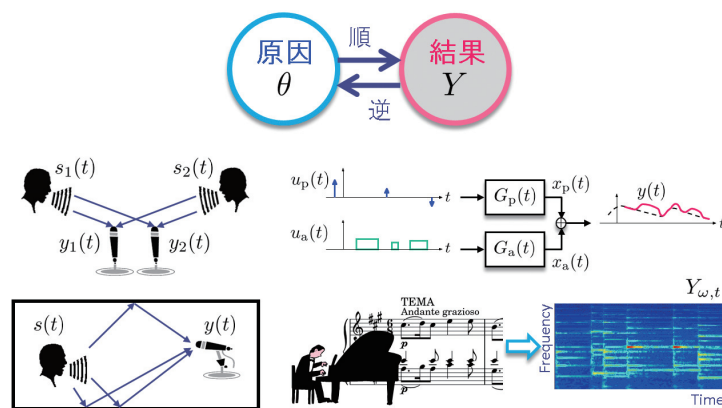


図 1：音源分離、残響除去、音声韻律特徴抽出、楽音分離、自動採譜などの音声音響信号処理問題



亀岡 弘和（かめおか ひろかず）

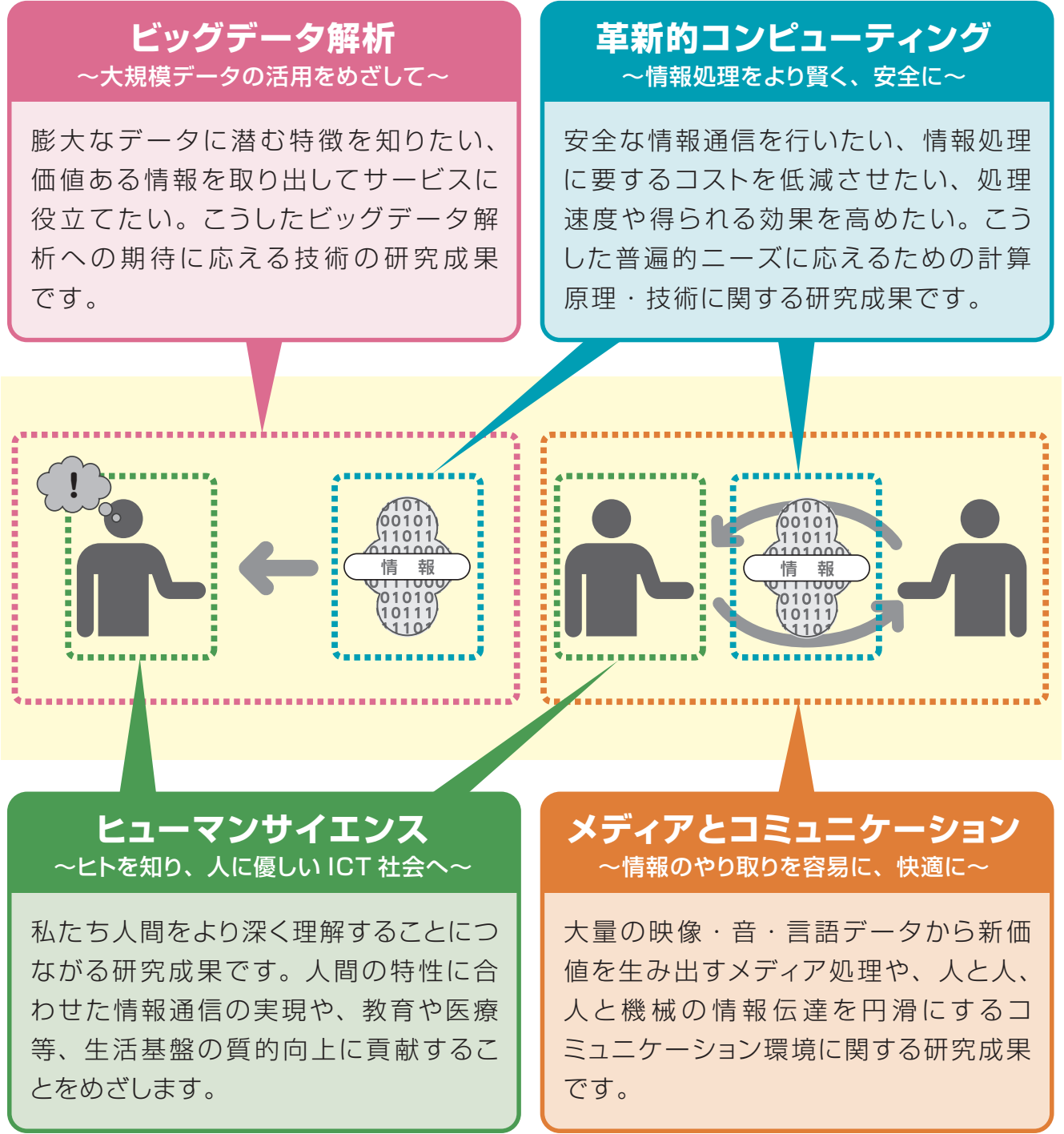
NTT コミュニケーション科学基礎研究所 メディア情報研究部 研究主任

2007年、東京大学 大学院情報理工学系研究科 システム情報学専攻 博士課程修了。同年、日本電信電話株式会社入社。情報理工学博士。2011年より、東京大学大学院情報理工学系研究科システム情報学専攻 客員准教授。音声・音楽を対象とした音響信号処理・機械学習の研究に従事。日本音響学会、情報処理学会、電子情報通信学会、IEEE、各会員。

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では、「人間」と「情報」を結ぶ新しい技術基盤の構築をめざして、情報科学と人間科学の両面から革新技術の考案と新原理の発見に取り組んでいます。

今年のオープンハウスでは、「革新的コンピューティング」、「メディアとコミュニケーション」、「ヒューマンサイエンス」の3分野に関する23展示に加え、NTT 研究所全体で取り組んでいる「ビッグデータ解析」の6展示をご紹介します。





ビッグデータチャレンジ

～NTT 研究所でのビッグデータ解析の取り組み～

背景：ビッグデータ時代の到来（クラウド、モバイル、ソーシャル、ビッグデータがICTの新潮流） ビッグデータ活用により、様々な産業分野で大きな市場を創出（MGIレポート：ヘルスケア分野では世界で年間3000億ドルの経済効果）

NTT研究所はビッグデータを活用した革新的サービスを実現し、「新たなステージ」における「バリューパートナー」への変革を実現します



各種センサ情報

ビッグデータ解析

データの背後に
潜む潜在情報を
膨大かつ多様な
情報から学習

因果関係分析
・ 予兆発見
・ 未来予測など

ビッグデータ解析により、あらゆる経済・社会活動における“現在”と“未来”のデータに基づく意思決定支援が可能

機械学習・データ科学センタ (MLC : Machine Learning・Data Science Center)

背景：関連技術者、(非ICTを含む) 応用領域の専門家との連携が必須

ミッション

- ・ビッグデータ/M2Mの革新的汎用技術の構築
- ・事業会社等からの先端的依頼案件に対する技術支援

実応用領域
(セキュリティ、トラフィック、CRM等)



バッチ型データ
処理基盤技術

Jubatus

オンラインデータ
処理基盤技術

機械学習技術

データベース技術：
多種多様なデータの蓄積・
編集・検索技術

並列分散計算技術：
大規模データの高速計算技術

機械学習技術：
潜在情報から新価値を
創出するための分析技術

連絡先

上田 功功 (Naonori Ueda) 機械学習・データ科学センタ

1

大事な特徴を捉えて流れるデータを把握する

～多次元時系列データからの特徴抽出～

どんな研究

本研究は、多次元（多くの特徴量を持つ）時系列データから時間的なパターンを抽出する技術です。時系列データの連続性の制約のもとで、各次元の特徴の有効性を推定して、時系列データに内在する特徴の共起パターンを確率モデルで推定します。

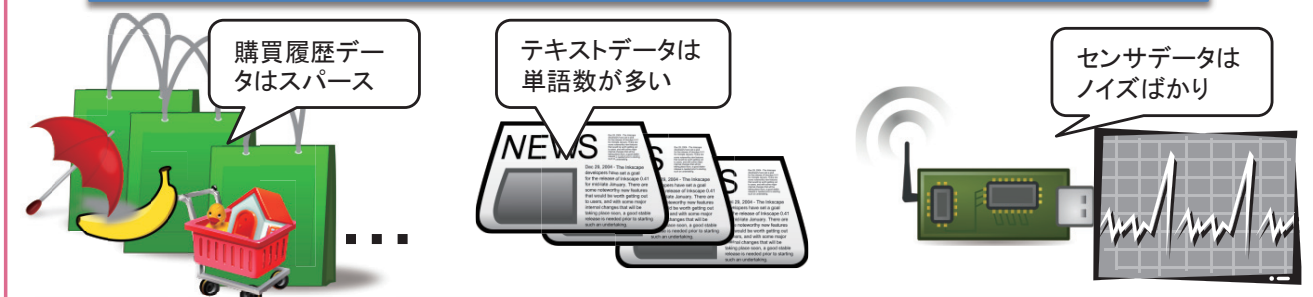
どこが凄い

本技術は、時系列データからパターンを見つけるだけでなく、「パターンを特徴づける特徴量」も自動的に抜き出します。これにより、大量のノイズが混ざった時系列データからパターンを発見できる上に、解析後のパターンの解釈などの人手の工数を減らすことに貢献できます。

めざす未来

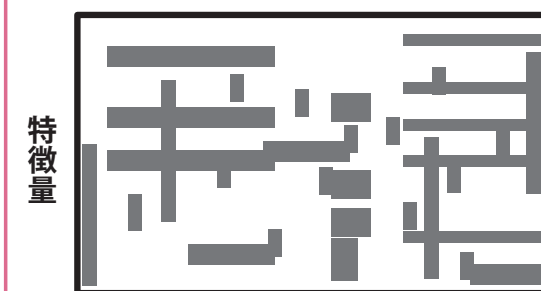
「ビッグデータ」に代表される多次元のデータから隠れた関係や知識を発見する技術は、今後ビジネスプロセスの効率化や業績向上のため益々必要となります。我々の研究技術は、実用的な精度と時間で「ビッグデータ」を活用できる社会の実現に役立ちます。

多次元の時系列データの全ての特徴量が本当に重要なのか？



この技術は時間的なパターンの違いを“特徴づける”特徴量だけを抽出

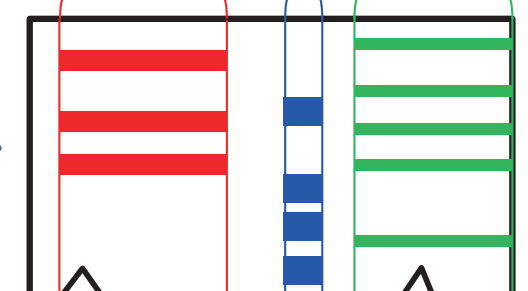
観測される時系列データ



時間

何が何やら・・・

パターン1 パターン2 パターン3



時間

ここにパターンがあったのか！

このパターンはこの特徴量が大事なのか！

- ⊙ 大量のノイズに影響されずに時間パターンを見つける
- ⊙ 各パターンに重要な特徴量が紐づけられ、解析結果の解釈が容易になる
- ⊙ パターンの数を自動的に推定

関連文献

[1] 石黒勝彦, 上田修功, 澤田宏, “動的サブセットクラスタリング,” 第14回情報論的ワークショップ(IBIS), 2011.

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

押し寄せる膨大な「今」を瞬時に賢く分析する

～フロー型ビッグデータを分散オンライン機械学習で分析～

どんな研究

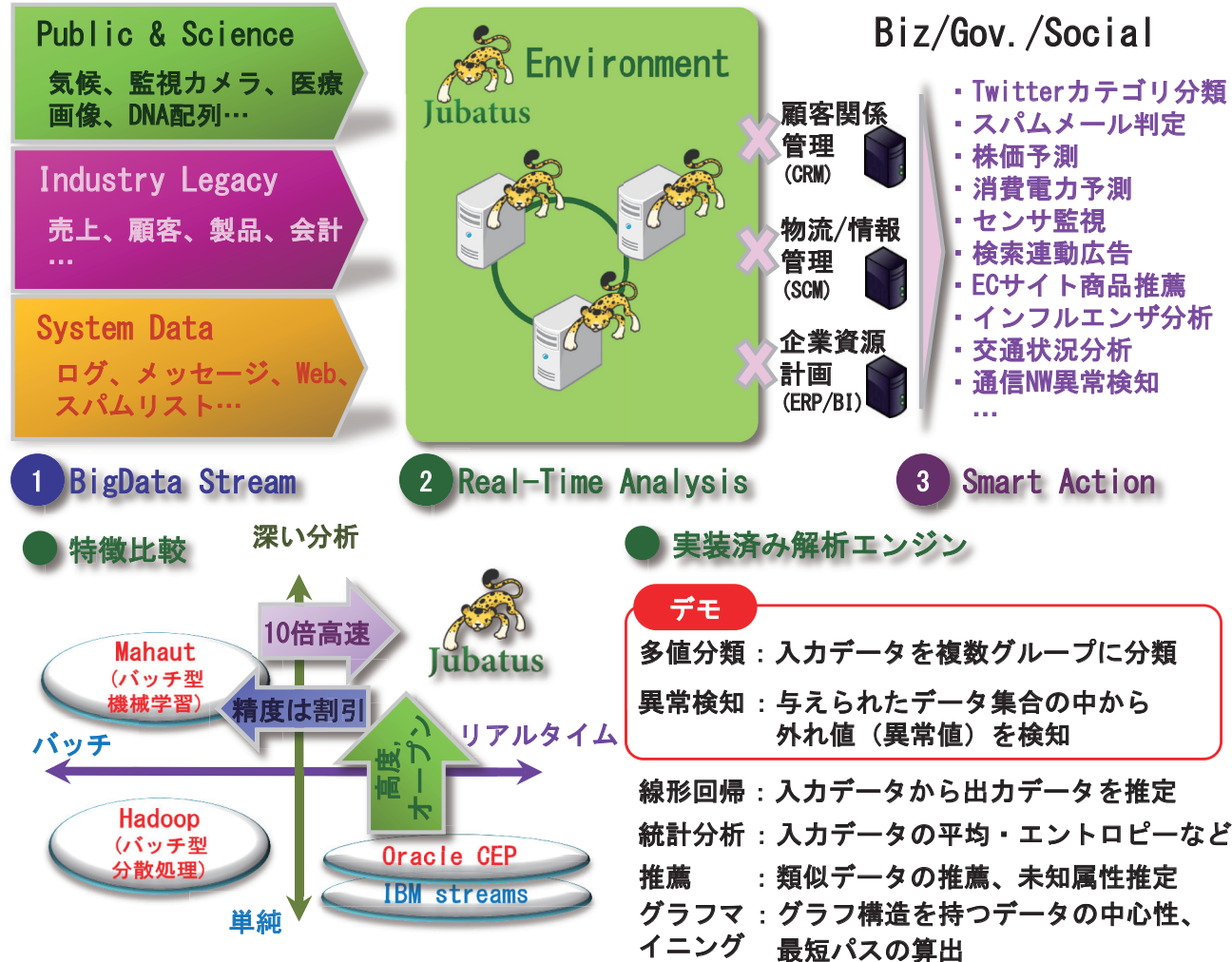
ビッグデータ分析の新たなニーズである、(1)広範囲に散在する多様な情報源のデータを統合し、(2)素早いリアルタイム分析でリスクを最小限に抑えて、(3)事前に全貌を把握しきれない多様な要求にも柔軟に対応できる処理基盤技術を開発しました。

どこが凄い

3つの要件「大量」「速い」「深い」処理はトレードオフの関係です。すべて両立することは本質的に困難ですが、これを克服し、毎秒16メガバイトのデータ流量に対し、自然言語解析と複数カテゴリへの自動分類を、遅滞なく高速実行します。またOSSとして広く公開しています。

めざす未来

世界規模でのオープンイノベーションとマーケット創造を推進し、リアルタイムマーケティングやスマートな社会インフラマネージメントへの適用拡大を図ります。



関連文献

- [1] 岡野原大輔, 海野裕也, 熊崎宏樹, 小田哲, “大規模リアルタイム解析エンジンJubatusの創り方,” 情報処理学会デジタルプラクティス, Vol. 4, No. 1, pp. 20-28, 2013.
 [2] 小田哲, 中山心太, 上西康太, 木下真吾, “Jubatus: Big Dataのリアルタイム処理を可能にする分散処理技術,” 信学技報, Vol. 111, No. 409, IN2011-126, pp. 35-40, 2012.
 [3] H. Makino, “Jubatus: Scalable Distributed Processing Framework for Realtime Analysis of Big Data,” in Proc. XLDB2012, 2012.

連絡先

堀川 桂太郎 (Keitaro Horikawa) NTT ソフトウェアイノベーションセンタ 分散処理基盤技術プロジェクト

大規模なグラフデータをあっという間に分析

～グラフデータを対象とした高速なマイニングアルゴリズム～

どんな研究

実世界にはソーシャルメディアやSNSなどの情報のつながりを表す膨大なグラフデータが存在します。従来の技術では、大量のグラフデータの分析のために、膨大な時間を要していました。今回は、このグラフデータを対象にした高速な分析アルゴリズムを紹介します。

どこが凄い

グラフデータの統計的な性質を用いて、グラフデータをよりコンパクトな形に変形していくことで、**世界最高速のグラフデータ分析アルゴリズム**を確立しました。このアルゴリズムを用いることで1億人分の友人関係を表すグラフデータも分析精度を落とさず3分程度で解析できます。

めざす未来

TwitterやSNSなどの極めて大規模なソーシャルメディアや人間の**行動履歴などを瞬時に分析し、情報推薦や情報予測、情報理解に活用**できるような未来を目指しています。本技術は、従来考えられなかった規模のデータの分析に貢献できる可能性があります。

■パーソナライズドページランク^注高速化技術 (関連文献[1])

POINT

- 人間関係や購買履歴などのグラフから重要なデータを高速に発見
- 従来技術よりも**10～100倍**程度高速に処理可能

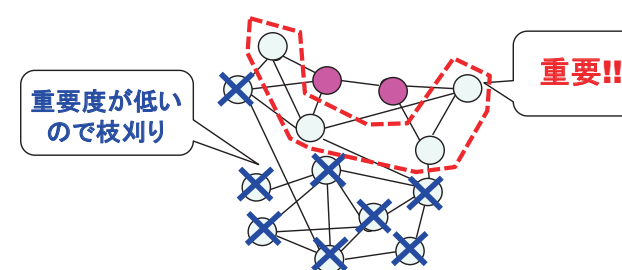
- ◎ グラフを行列にし、行列分解と並べ替えてグラフをコンパクトにして重要度を計算
- ◎ 重要度の低いデータを枝刈りし、グラフをさらにコンパクトに変換

$$S = (1-c) \begin{bmatrix} G \\ S + cN \end{bmatrix}$$

S : ノードの重要度ベクトル M : 趣向ベクトル
 G : 隣接行列 c : ランダムウォーク再スタート確率

$$S = C \begin{bmatrix} P^T \\ Q \\ R \end{bmatrix} \begin{bmatrix} P \\ N \end{bmatrix}$$

P : 置換行列 Q : QR分解における直交行列
 R : QR分解における上三角行列



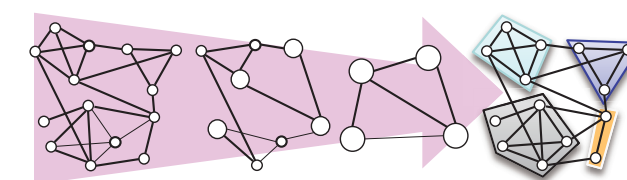
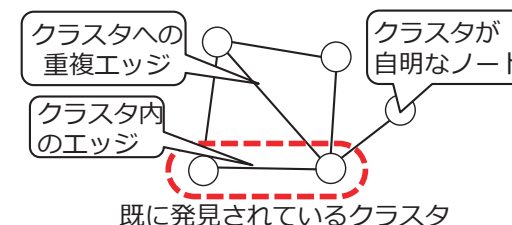
注) パーソナライズドページランク: 個人に有用な情報検索技術の1つ

■クラスタリング高速化技術 (関連文献[2])

POINT

- 人間関係や購買履歴などのグラフの中からコミュニティを高速に発見
- 1億人規模のデータも3分**で処理可能

- ◎ グラフから重複計算の生じるノードとエッジを削減
- ◎ 計算不要なデータを逐次的に集約しデータをコンパクトにすることで計算を高速化



関連文献

- [1] Y. Fujiwara, M. Nakatsuji, T. Yamamuro, H. Shiokawa, M. Onizuka, “Efficient Personalized PageRank with Accuracy Assurance,” in Proc. the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2012.
 [2] H. Shiokawa, Y. Fujiwara, M. Onizuka, “Fast Algorithm for Modularity-based Graph Clustering,” in Proc. the 27th AAAI Conference on Artificial Intelligence (AAAI), 2013.

連絡先

鬼塚 真 (Makoto Onizuka) NTT ソフトウェアイノベーションセンタ 分散処理基盤技術プロジェクト

ビッグデータ分析によるネットワーク異常対策

～ ネットワーク故障、サイバー攻撃などのNW異常を早期に検出～

どんな研究

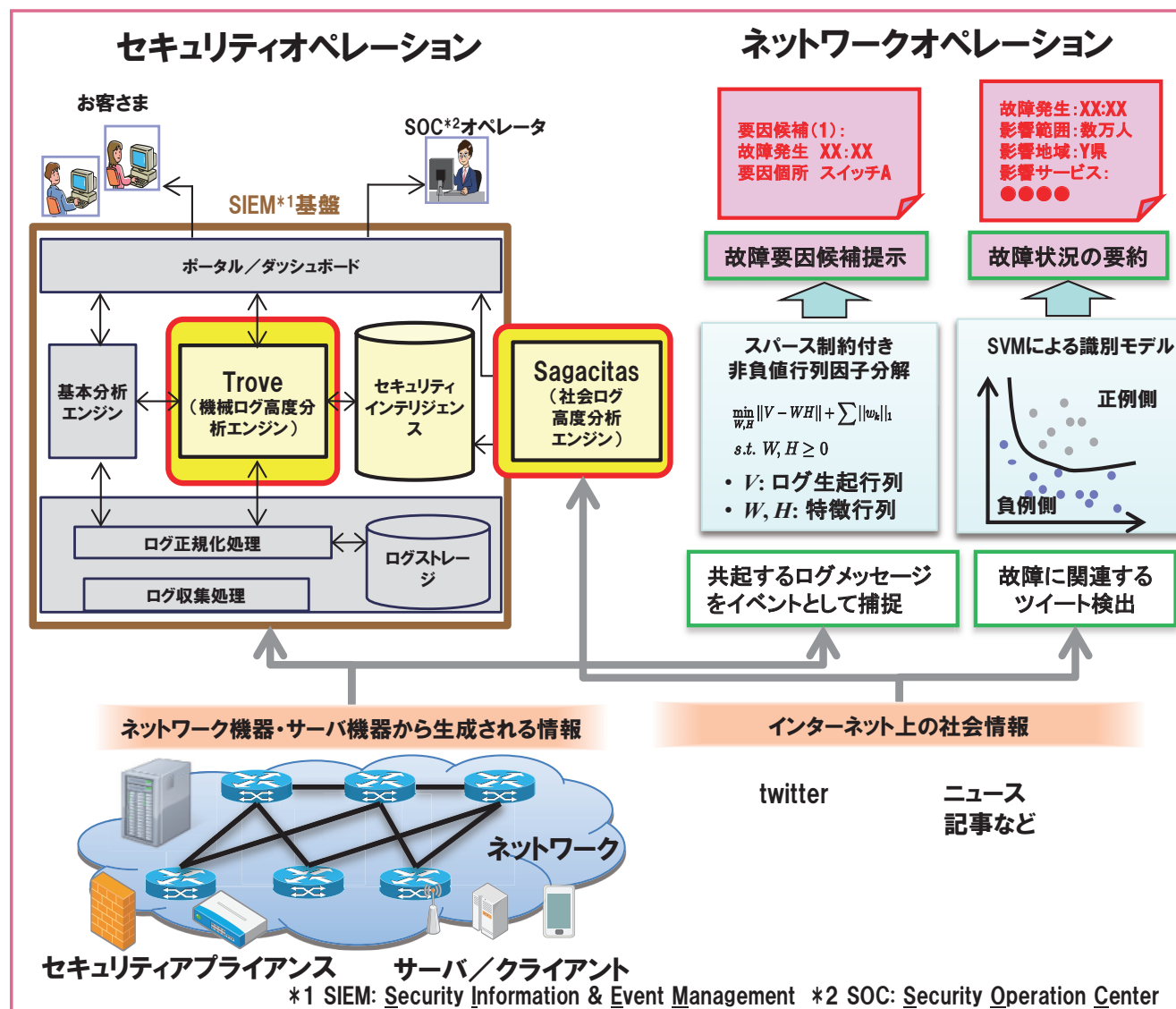
ネットワーク機器やサーバ機器からのログ情報や、SNSやニュース記事などの社会情報といった膨大、不統一、かつ非定型なデータから、機械学習等の高度分析技術を用いて、**従来検出できなかったネットワーク故障やセキュリティインシデント(サイバー攻撃等)**を早期に検出します。

どこが凄い

ネットワーク情報を**時間共起性や通信先共起性**を用いてグルーピング化し、故障イベントの可視化や未知の悪性ホスト検出を行います。また社会情報から、セキュリティや故障に関連する情報を抽出し、リアルタイムに分析し、攻撃予兆や故障発生を早期に検出します。

めざす未来

ネットワーク故障、サイバー攻撃検出技術により、それぞれNTTグループの**ネットワークオペレーション、セキュリティオペレーション**を高度化し、今までより一歩進んだ安心・安全なネットワーク環境の実現に寄与します。



関連文献

- [1] K. Sato, K. Ishibashi, H. Hasegawa, and H. Yoshino, "Extending Black Domain Name List by Using Co-occurrence Relation between DNS Queries," *IEICE Trans. Commun.*, 2012.
- [2] 針生剛男, 秋山満昭, 青木一史, 八木毅, 岩村誠, 倉上弘, "進化するマルウェア等によるサイバー攻撃の検知・解析・対策技術," NTT技術ジャーナル, 2012.
- [3] 木村達明, 森達哉, 石橋圭介, 塩本公平, "大規模ネットワーク監視情報における重要イベント抽出法," 信学技報 NS2011-225, 2012.

連絡先

石橋 圭介 (Keisuke Ishibashi)
NTT ネットワーク基盤技術研究所
通信トラヒック品質プロジェクト

綱川 光明 (Mitsuoka Tsunakawa)
NTT セキュアプラットフォーム研究所
セキュリティマネジメント推進プロジェクト

このつぶやきを書いたのはどんなひと？

～ マイクロブログの特徴を活用したユーザ属性推定技術～

どんな研究

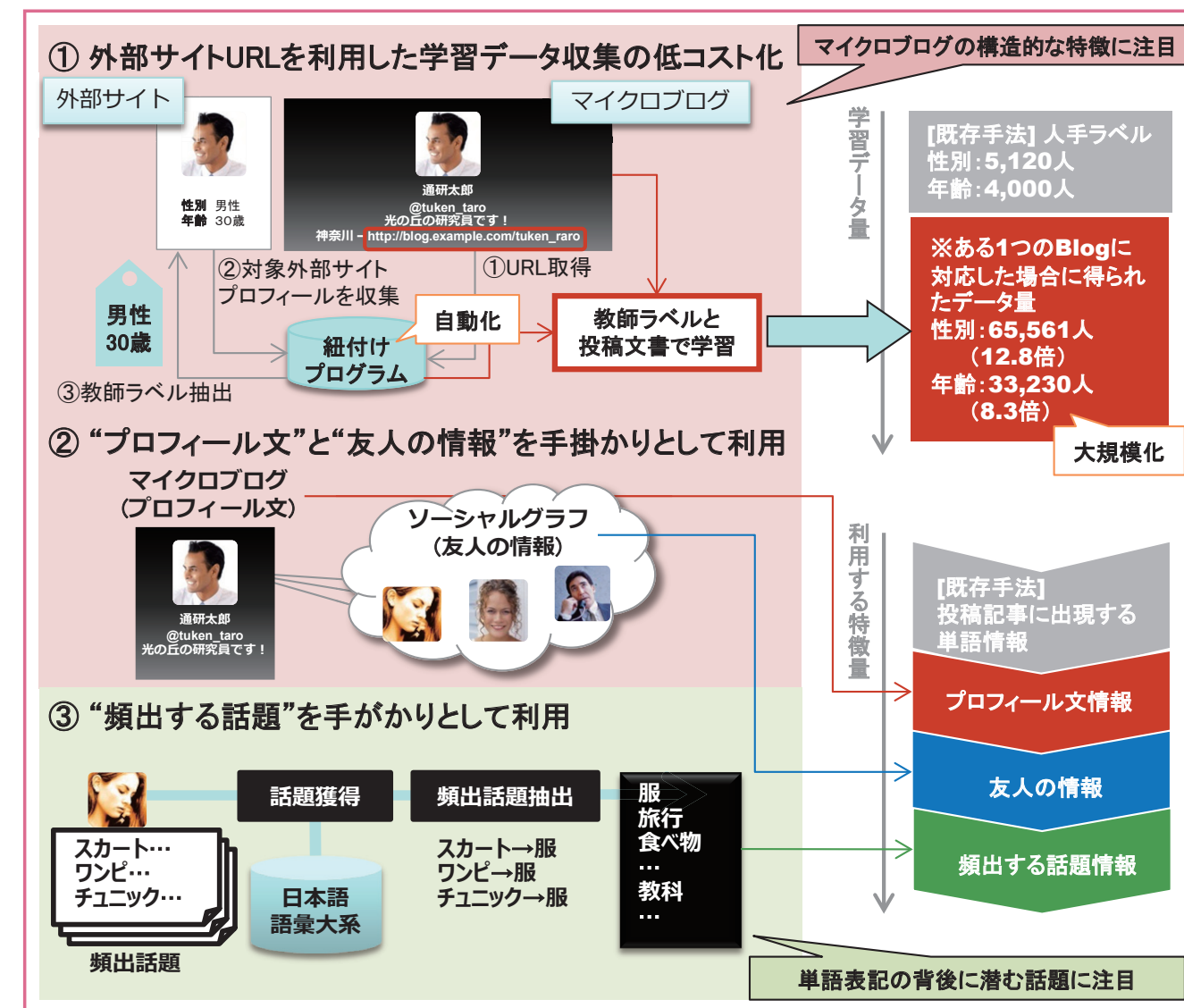
マイクロブログには個人の意見が多数存在し、マーケティング等での活用が望まれています。しかし、既存のアンケート調査と異なり、それら**書いたユーザの属性**(年齢、性別等)が明らかでないという問題があります。本研究では、それらを**自動的に推定**する技術に取り組んでいます。

どこが凄い

従来は人手で作成した学習データを利用し、投稿記事中の単語を元にユーザの属性を推定する手法が提案されていました。本研究では、**単語表記の背後に潜む話題や、マイクロブログの構造に注目**することで、性別を約90%の精度で推定する等、**主要なユーザ属性を高い精度で推定可能**としました。

めざす未来

現在の推定対象はマーケティングでよく利用される年齢、性別、職業、興味等ですが、**今後、より幅広い属性を、より深掘りした形で提供することを目指しています**。それにより、マーケティングでの活用はもとより、個人の行動支援やコミュニケーション支援での活用も期待できます。



関連文献

- [1] 伊藤淳, 西田京介, 星出高秀, 戸田浩之, 内山匡, "TwitterとBlogの共通ユーザおよび会話ユーザの同類性に着目したTwitterユーザ属性推定," 第5回データ工学と情報マネジメントに関するフォーラム, C5-3, 2013.

連絡先

伊藤 淳 (Jun Ito)
NTT サービスエボリューション研究所
サービスハーモナイズプロジェクト

平野 徹 (Tohru Hirano)
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

大規模移動データの分析と可視化

～異種複合情報を用いた移動パターン分析～

どんな研究

GPSデバイスや交通系ICカード等の普及により、人々の移動情報が大量に蓄積されています。こうした人の移動パターンは、交通機関の運行状況や、天候などの外部状況に依存して変化します。本展示では、様々な複合情報を用いた人の移動パターンの分析可視化ツールを紹介します。

どこが凄い

従来、単一情報源からのデータを用いた移動パターン分析が行われてきましたが、本研究ではソーシャル・ネットワークサービスのログや気象情報等の異種情報を組み合わせた分析により、移動パターン変化の複雑な因果関係をとらえ、多面的に可視化するツールを構築しています。

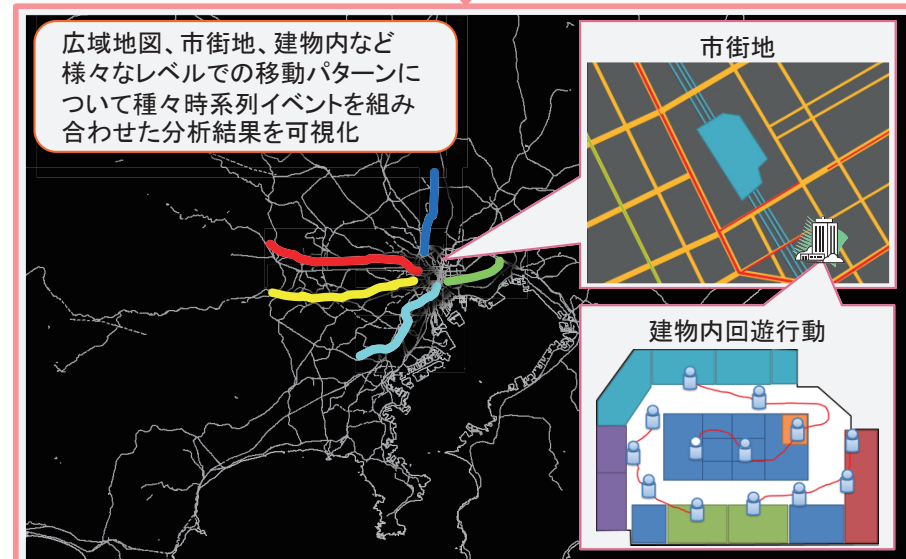
めざす未来

多様な要因に基づいて変化する人の移動パターンを多面的に分析し、様々なイベントに関連した人の行動や状況を予測する技術を確立します。これにより、災害時の輸送計画や最適な避難経路をタイムリーに提示したり、新たな都市交通計画などへの幅広い応用が期待されます。

分析可視化ツールの利用イメージ



多様な外部情報を用いて移動パターンを分析し可視化



関連文献

- [1] J. Yuan, Y. Zheng, X. Xie, G. Sun, "T-Drive: Enhancing Driving Directions with Taxi Drivers' Intelligence," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 25, No. 1, pp. 220-232, 2013.

連絡先

納谷 太 (Futoshi Naya) 協創情報研究部 知能創発環境研究グループ

効率的な環境モニタリングを実現

～相関性を用いたセンサデータ圧縮技術～

どんな研究

無線センサネットワークを構成するバッテリー駆動型のセンサ装置が観測、送信するデータの特性を活かして圧縮を行うことで、無線での送受信時にかかる電力消費量を抑え、センサネットワークの稼働時間を大幅に伸ばすことを目指します。

どこが凄い

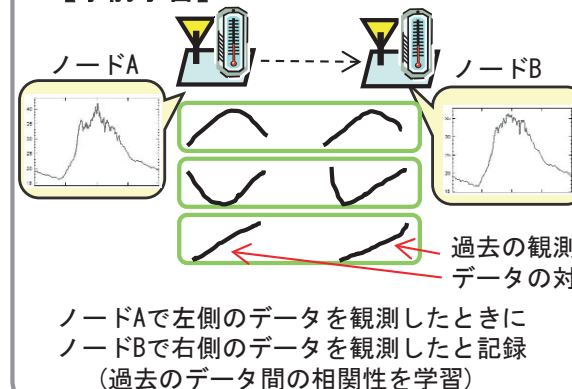
空間内に配置された複数のセンサで観測するデータ間には相関性(類似性)があります。この相関性を利用してデータを圧縮することで、単独のセンサで観測したデータを圧縮する手法に比べ、高い圧縮率でデータを可逆圧縮することができます。

めざす未来

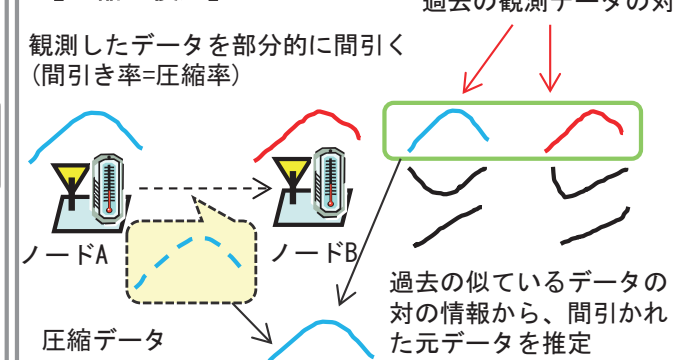
広い範囲の実世界の環境データを、センサネットワークでリアルタイムに取得し、解析技術を用いて分析することで、これまで把握することが難しかった、実世界で生じているさまざまな事象をすばやく把握し、きめこまやかな機器の制御等を実現することができます。

【圧縮の仕組み】過去のデータと、観測したデータ系列間の相関性を利用して圧縮する手法を用いることで、データ系列間の相関性が高くなればなるほど、データの圧縮率を高めることが可能

【事前学習】

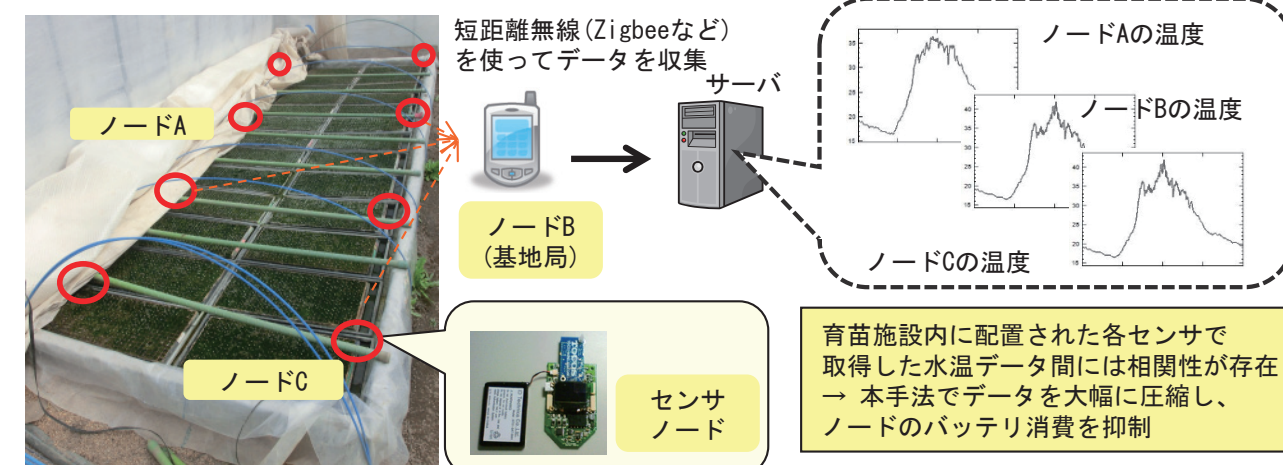


【圧縮・復元】



【用途例】

育苗環境データをセンサノードで取得して分析し、発芽率の改善に活用(農業ICT)



関連文献

- [1] J. Muramatsu, T. Uyematsu, T. Wadayama, "Low density parity check matrices for coding of correlated sources," *IEEE Transactions on Information Theory*, Vol. IT-51, No. 10, pp. 3645-3653, 2005.
- [2] 柳沢豊, 岸野泰恵, 前川卓也, 須山敬之, "相関のあるデータを観測する無線センサネットワークのためのデータ集約手法," 情報処理学会第32回ユビキタスコンピューティングシステム研究会, 講演番号(4), 2011.

連絡先

柳沢 豊 (Yutaka Yanagisawa) 協創情報研究部 知能創発環境研究グループ

あなたの居場所を安全に公開

～仮名交換による位置プライバシー～

どんな研究

位置情報を活用することで、ユーザの所在に応じたきめ細かなサービスが可能になる一方、そのトレードオフとしてプライバシーの保護が問題になります。この展示では、これらの2つのバランスを、アプリケーションに応じて適切に取るための技術を紹介합니다。

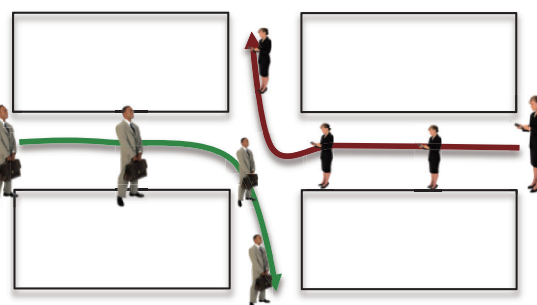
どこが凄い

同じ場所に居合わせたユーザ同士がランダムに**仮名**(仮の名前)を**交換**することで、長期的な**位置プライバシー**の保護と、短期的な位置情報の**有用性**とを**両立**させることに成功しました。また、プライバシー保護度合を厳密に見積もるための効率的なアルゴリズムを考案しました。

めざす未来

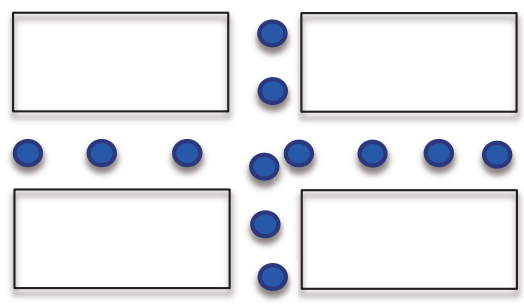
位置情報の不正利用や第三者による監視を気にすることなく、位置情報を活用したサービスを安心・安全に享受できる、ゆたかな情報化社会の構築に役立ちます。

個人が完全に識別できる位置データ



きめ細かなサービスが可能
位置プライバシーの保護なし

個人識別情報を除去した位置データ

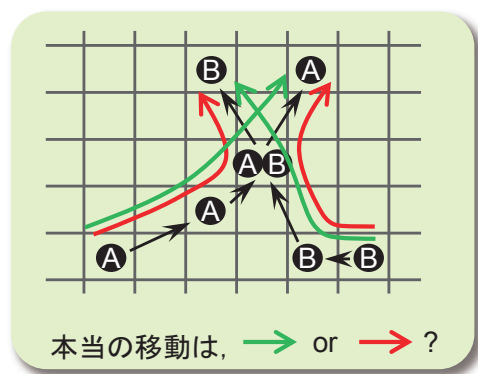


完全な位置プライバシー保護
きめ細かなサービス不可能

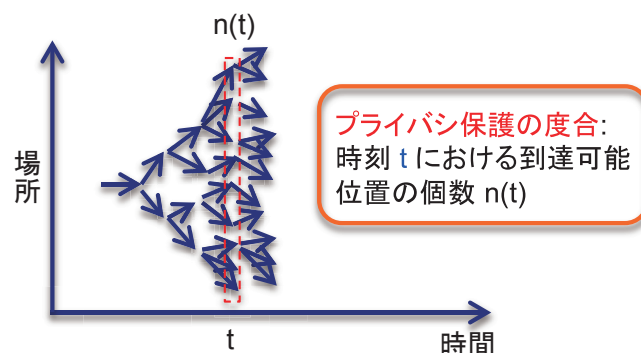
トレードオフ

バランスを取るための技術: **仮名交換**

仮名交換: 同じ場所に居合わせたユーザ同士が**ランダムに仮名を交換**



仮名交換で保護されるプライバシー:
他のユーザと居合わせると、到達可能
(だと攻撃者が考える)位置が増える



関連文献

[1] K. Mano, K. Minami, H. Maruyama, "Protecting location privacy with K-confusing paths based on dynamic pseudonyms," in *Proc. 5th IEEE International Workshop on Security and Social Networking (SESOC)*, 2013 (to appear).

連絡先

真野 健 (Ken Mano) 協創情報研究部 情報基礎理論研究グループ

光がつなぐ量子の情報

～レーザ光を用いた最適な量子もつれ生成方法～

どんな研究

量子情報通信と呼ばれる未来の通信を実現するには、微弱なレーザ光を使った方法が最も有望であると考えられています。今回、この方式が持つ可能性と限界を理論的に解析し、厳密かつ定量的な評価に成功しました。

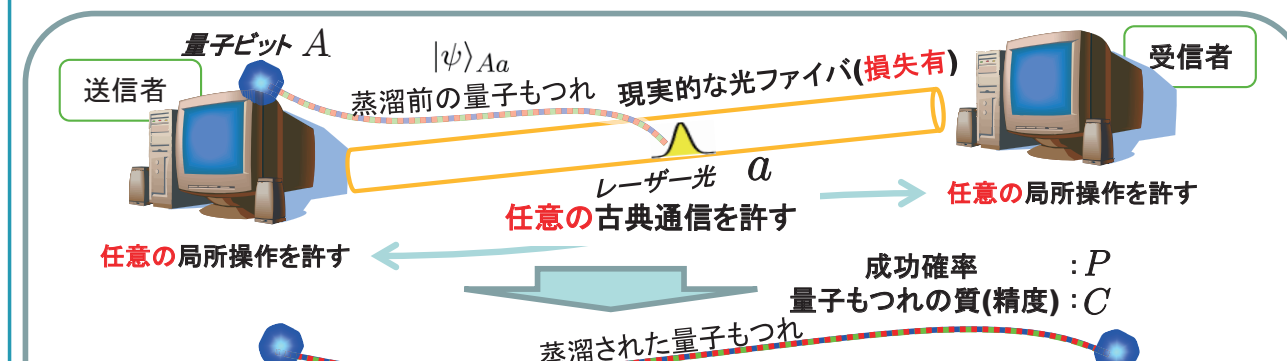
どこが凄い

従来の量子通信の性能に関する研究は、より効率的なプロトコルを順次提案するものであり、改善の余地が不明確でした。そこで、本研究では、**理論限界を定量的に示しました**。また、従来の**全ての提案プロトコルを凌駕した理論限界に近いプロトコル**の提案にも成功しました。

めざす未来

現在推し進められている量子通信の方式における理論限界が見つかったため、現状の技術の延長線上でできる事が明確化し、量子通信技術の迅速な発展が見込まれます。それによって、量子暗号・量子中継・量子分散計算・量子テレポーテーションなどの早期実現が期待されます。

問題設定: ビット値に応じたレーザ光を、現実的な光通信路(損失のある光ファイバ)で伝送すると、精度の低い(=蒸溜前の)量子もつれが生成される。これを、送受信者の任意の局所操作と無制限の古典通信によって「蒸溜」し、高精度な量子もつれ(量子通信のリソース)を高成功確率で生成する。



目的: 送受信者間に、できるだけ高い精度 C の量子もつれを、できるだけ高い成功確率 P で生成

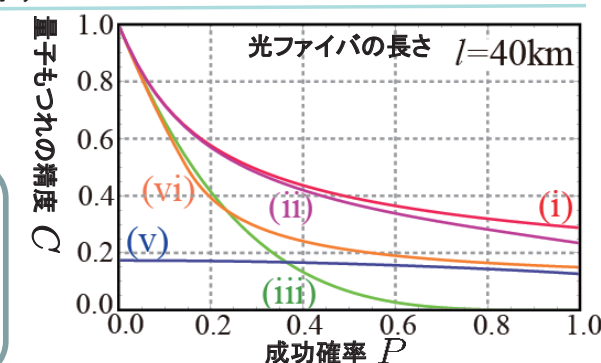
- 光ファイバが持つ損失により、成功確率と量子もつれの質には原理的な限界が存在すると予想される
- 量子通信の性能に関する理論的な既存研究は、すべてこの量子もつれ生成方法の特殊化とみなせる
- 量子もつれの精度とは、古典情報との区別のしやすさを示す

結果: 理論限界の導出に成功

成功確率が P 時の量子もつれの質の理論限界

$$C^{\text{upp}}(P) = \frac{u^{\frac{1-T}{T}} \sqrt{(1-u)(2P+u-1)}}{P} \quad \text{with} \\ u := \frac{1}{2} \left[(1-P)(2-T) + \sqrt{4P^2(1-T) + (1-P)^2 T^2} \right]$$

T : 光ファイバの透過率



(i): 本研究で得られた理論的上限
(ii): 本研究で提案した現実的に実装可能なプロトコル
(iii), (vi), (v): 既存のプロトコル

関連文献

[1] K. Azuma, G. Kato, "Optimal entanglement manipulation via coherent-state transmission," *Physical Review A (PRA)* 85, 060303(R), 2012.

連絡先

加藤 豪 (Go Kato) 協創情報研究部 情報基礎理論研究グループ

ちゃんとしたデタラメを作る

～レーザ光の高速乱雑変動を利用した物理乱数生成～

どんな研究

予測不可能性など品質が保証された高速乱数生成器は、安全な通信への応用などの際の重要な要素技術となります。これまでカオスレーザを使った高速な乱数生成器を提案してきました。今回は生成された乱数の品質保証理論およびその実験的検証に関して紹介します。

どこが凄い

カオスレーザを用いた高速な物理乱数生成器において、マイクロノイズとカオスの不安定性に基づく生成乱数列の予測不可能性を保証する理論を提唱してきました。今回、実験によりこの**乱数品質保証理論**を確かめることに世界で初めて成功しました。

めざす未来

今後益々発展が期待されるネットワーク社会において機密情報の秘匿通信の重要性は増し、高速で品質の保証された物理乱数生成技術は必要不可欠なものとなります。我々の物理乱数生成器の研究により、誰もが安全で高速な通信ができる環境が実現することをめざします。

カオスレーザによる乱数生成の理論

レーザ内部のマイクロノイズ
観測困難な不確定さ

戻り光が引き起こす
不安定なダイナミクスによって**拡大**

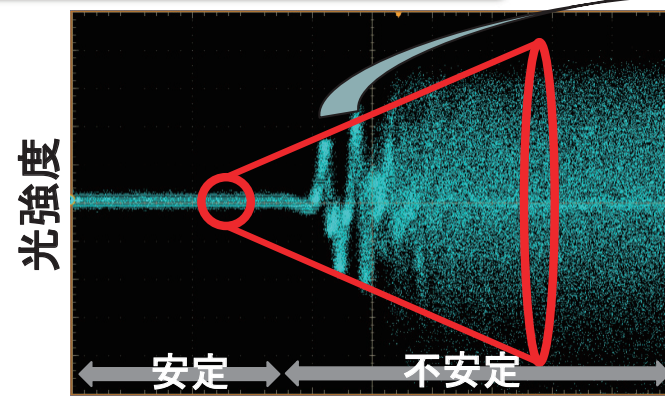
マクロな不規則信号
観測し易い不確定さ

乱数

実験により検証

予測不可能なマイクロノイズが拡大される様子

不安定ダイナミクスの効果



不安定な時間領域でマイクロノイズがマクロな領域に拡大

予測不可能

光強度の観測値はマイクロノイズに依存

光強度から抽出された乱数は予測不可能

レーザの状態を安定から不安定に切り替える試行を繰り返し、光強度の時系列を重ね書きしたもの。

関連文献

- [1] S. Sunada, T. Harayama, P. Davis, K. Tsuzuki, K. Arai, K. Yoshimura, A. Uchida, "Noise amplification by chaotic dynamics in a delayed feedback laser system and its application to nondeterministic random bit generation," *Chaos* 22, 047513, 2012.
[2] K. Arai, T. Harayama, S. Sunada, P. Davis, "Randomness in a Galton board from the viewpoint of predictability: Sensitivity and statistical bias of output states," *Physical Review E* 86, 056216, 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

高度なプログラミングが可能なビジュアル言語

～プログラミング上の様々な拡張をした新ビスケット～

どんな研究

プログラミング入門用のツールとして、ビスケットは一定の評価を得てきました。入門後にどこへ進めばよいか、といった要求も出てきました。そこで、より高度なプログラミングへ連続的に進むことができるように機能拡張を行いました。

どこが凄い

規模の大きなプログラムをビスケットで作ることができます。他人の作ったプログラムをライブラリとして使うことができます。AndroidなどのタブレットとPCとで同じアプリが使用できます。**ビジュアルだけで高度なシステムを構築**することができます。

めざす未来

コンピュータのアプリケーションを「自分で作って、自分で使う」ことを可能にします。もともとコンピュータとはそういうものでした。**誰もが情報化革命に能動的に参加**できるようになります。

■ ビスケットの広がり期待

- 2012年度、約2,000名の子どもの対象にプログラミング入門授業を実施（NTTドリームキッズ、教育スクエアxICT、公立学校、NPO）
- その中で、子どもたちが自ら教材を作成し、それで学べる、より高度なプログラミング環境への期待が高まった（遊びのツールだけではなく、使えるツールにしたい）

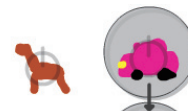
■ ビスケットの拡張機能

【拡張機能1】絵の継承[1]の導入

大規模なプログラミングを可能にする仕組みとして、オブジェクト指向なプログラムの再利用を導入

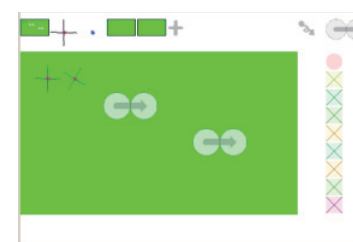


「車はすすむ」



「馬と人」は「車」を継承

ビスケットが使われている様子
NTTグループCM「君の未来篇」より



制作中のシステムの画面

【拡張機能2】分散された複数の紙によるビジュアルプログラミング

- 従来のシステム：一般的Web技術+ビスケットアプリ
複数の個別の作品をWeb技術で相互に関係を持たせているだけ
- 新バージョン：分散されてつながった「紙」と「ビジュアルな分散ポイント」
Webプログラミングをせずに、ビジュアルだけで分散プログラミング可能

ネット上に広がる巨大な模造紙で子どもたちが学び合うプラットフォーム

関連文献

- [1] 原田康徳, "ビスケットへの絵の継承の導入," 第52回プログラミングシンポジウム, 2011.

連絡先

原田 康徳 (Yasunori Harada) メディア情報研究部 コミュニケーション環境研究グループ

計算機による自動英文法解釈

～統計的文法獲得技術に基づく構文解析～

どんな研究

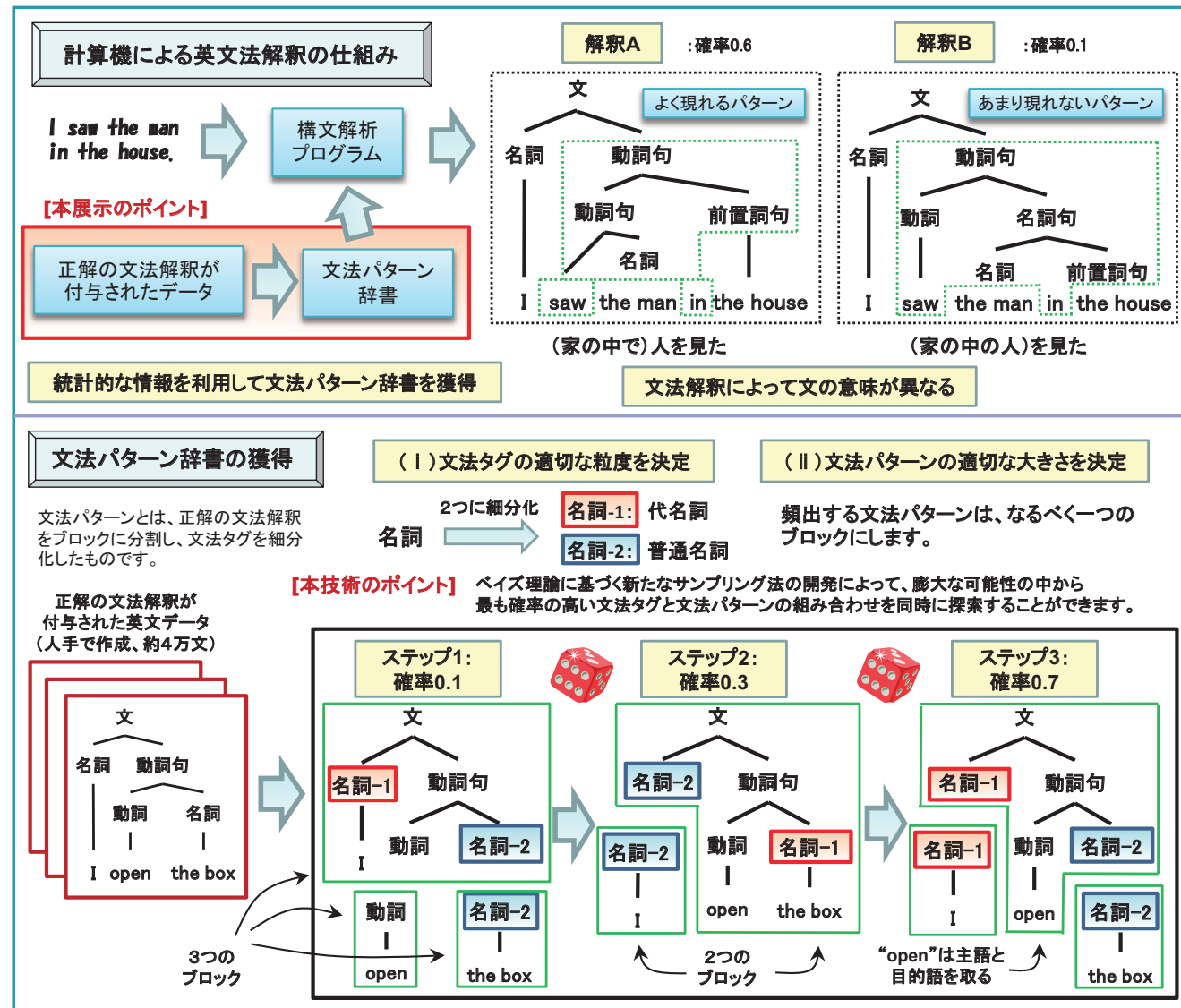
英語や日本語を理解するには、文の主語や述語などを見分けて全体の文法を把握する必要があります。しかし、これを計算機で自動的に行うことは容易ではありません。本展示では、**自動的に英文法解釈を行う**計算機プログラムの仕組みを紹介します。

どこが凄い

正解の文法解釈が付与された英文データから、文脈に応じた適切な**文法パターン**を統計的に獲得する方法を開発しました。この方法を用いて構築された英語の構文解析システムは、国際標準ベンチマークデータで**世界一の精度**を達成しました。

めざす未来

人の言葉を理解してコミュニケーションを手助けする計算機の実現を目指しています。特に構文解析は、文法的に正しく自然な文を生成する**機械翻訳**や、英語学習者の作文の文法誤りを指摘・修正する**英文校正**システムへの応用が期待できます。



関連文献

- [1] H. Shindo, Y. Miyao, A. Fujino, M. Nagata, "Bayesian Symbol-Refined Tree Substitution Grammars for Syntactic Parsing," in *Proc. ACL*, 2012. (Best Paper Award 受賞)
- [2] 進藤裕之, 宮尾祐介, 藤野昭典, 永田昌明, "シンボル細分化を適用した階層Pitman-Yor過程に基づく置換文法獲得法と構文解析への応用," 言語処理学会第18回年次大会, 2012. (最優秀賞受賞)
- [3] H. Shindo, A. Fujino, M. Nagata, "Insertion Operator for Bayesian Tree Substitution Grammars," in *Proc. ACL*, 2011.

連絡先

進藤 裕之 (Hiroyuki Shindo) 協創情報研究部 言語知能研究グループ

言語の壁をどう越える？

～英語・中国語・韓国語から日本語への統計翻訳～

どんな研究

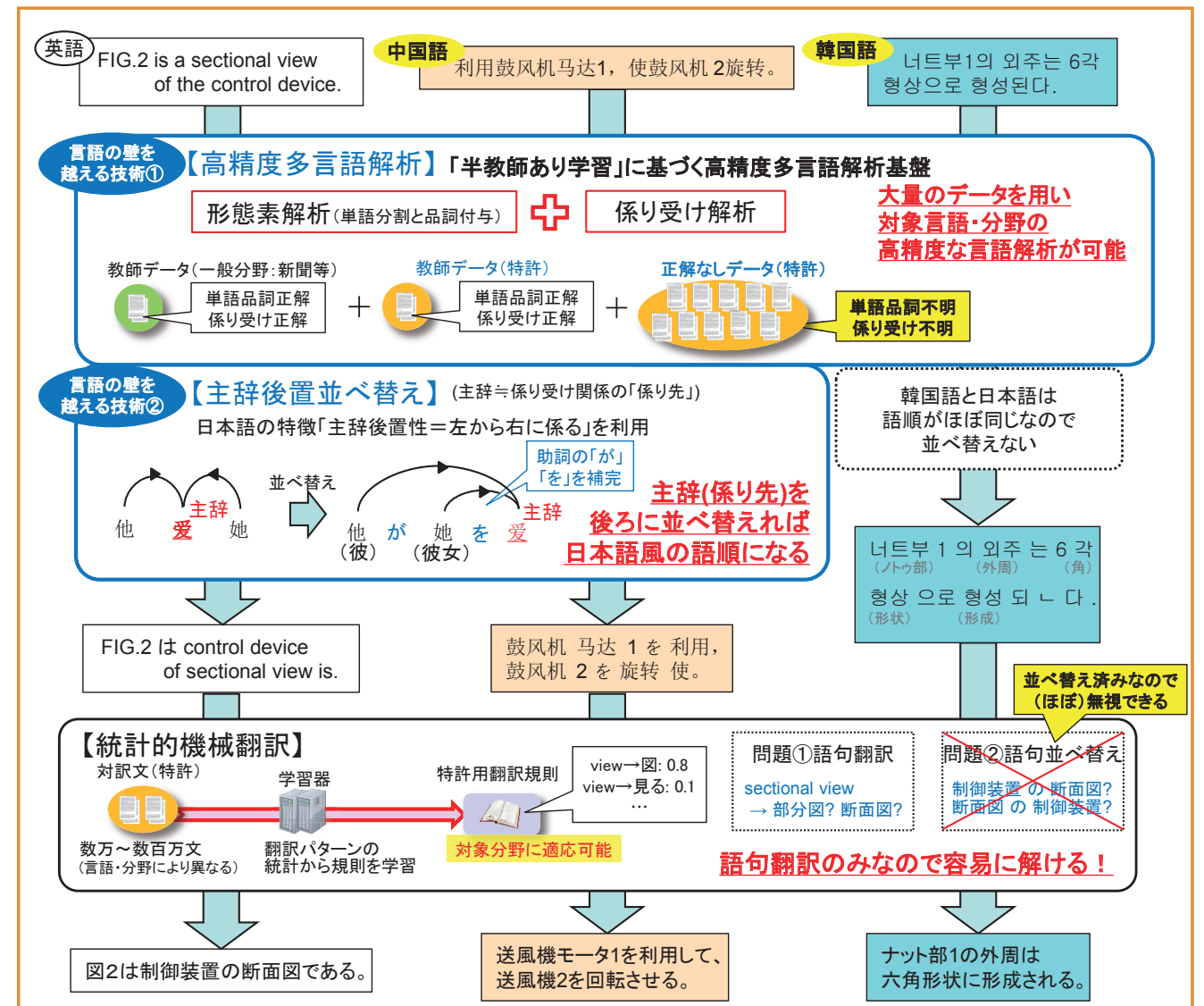
対訳例文を大量に使うことで翻訳の規則を確率的にモデル化し、機械翻訳を実現する「統計的機械翻訳」に関する研究です。この展示では、**高精度なことばの解析**と、**語順の違いを克服する並べ替え技術**とを組み合わせ、英語・中国語・韓国語を日本語に翻訳する技術を紹介します。

どこが凄い

多言語の学習データを独自に整備し、英語で培ってきた高度な言語解析技術を**多言語に拡張**しました。英語や中国語は日本語と大きく語順が異なりますが、この解析結果を用いて「翻訳前に」**日本語風に単語を並べ替える**ことで、高精度な日本語への翻訳を可能にします。

めざす未来

外国語で書かれた特許やマニュアル等の技術文書を日本語で探したり読んだりすることができるようになります。将来的にはより一般的な文書や話しことばへも対象を広げ、言語の壁を越えたコミュニケーションを実現します。



関連文献

- [1] J. Suzuki, H. Isozaki, X. Carreras, M. Collins, "An Empirical Study of Semi-supervised Structures Conditional Models for Dependency Parsing," in *Proc. EMNLP*, pp. 551-560, 2009.
- [2] 鈴木潤, Kevin Duh, 永田昌明, "拡張ラグランジュ緩和を用いた同時自然言語解析法," 言語処理学会年次大会, pp.1284-1287, 2012.
- [3] H. Isozaki, K. Sudoh, H. Tsukada, K. Duh, "HPSG-based Preprocessing for English-to-Japanese Translation," *ACM TALIP*, Vol. 11, Issue 3, Article No. 8, 2012.

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

「触りながら読む」読文を実現：Yu bi Yomu

～なぞり動作を利用した動的な文章表示方式～

どんな研究

“Yu bi Yomu”はタブレットPC端末の特徴を活かした**新しいデジタル文章表示方式**です。読者が画面を指でなぞると、その位置の文字が現れては消えてゆきます。文字表現に時間的な情報を加えることで、そこに**話し言葉が持つようなリズム、抑揚、余韻を感じる**ことができます。

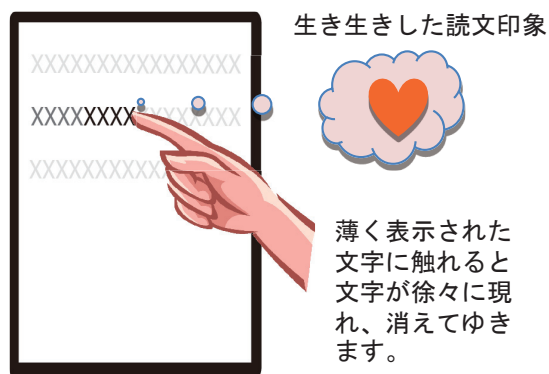
どこが凄い

“Yu bi Yomu”は文字の現れ方・消え方を変化させるだけで、**文章内容の印象**を変化させることができます。また、自分で読むだけでなく、なぞり動作で生じる**動的な文章を記録、送付**することで、個人の特徴や感情を表現する豊かな文字コミュニケーションを行うことができます。

めざす未来

文字と音声の中間的な特徴を持つ“Yu bi Yomu”は、**新たな文章表現**の可能性をひらきます。また、メール文や電報といった電子的な文章のやり取りに、書き手の**思いをより豊かに伝える**手段を提供します。将来的には、教育支援分野への適用も視野に入れています。

動的な文字表示となぞり動作を組み合わせることで話し言葉のようなリズム、抑揚を持つ文章表示方式を実現



使用法1：自分でなぞり動作を行う
自分のなぞりによって文字が現れ・消えることで、声を出して文章を味わうような、深い読文が行えます。

使用法2：作成した表示を送付する
作成者はなぞり動作で生じる動的な文章を記録して、他の人に送付します。受信者は送られた動的な文章を再生することで、作成者の抑揚まで含めた文章が伝達されます。

特徴1：時間的特性を付与した文章表示
Dynamic text (ダイナミックテキスト)

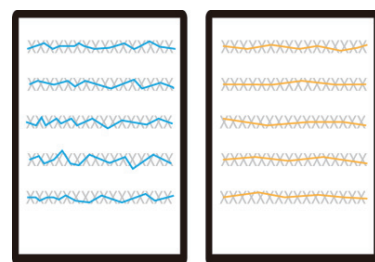
→文字の出現・消滅時間の操作によって異なる読文印象が生じることを実験で確認しました。



触れてからの時間

特徴2：なぞり動作による文字表示制御

→指先の動きの違いで、簡単に抑揚や、個人の違いを作り出すことができます。



読文時のなぞり動作の違い

関連文献

- [1] 丸谷和史, 植月美希, 安藤英由樹, 渡邊淳司, “ユーザのなぞり動作に基づく動的な文章表示形式,” 情報処理学会論文誌, Vol. 53, No. 9, pp.1-11, 2013.
[2] K. Maruya, M. Uetsuki, H. Ando, J. Watanabe, “Yu bi Yomu”: Interactive reading of dynamic text,” in *Proc. ACM MultiMedia*, 2012.
[3] 丸谷和史, 渡邊淳司, “文書表示装置, 文書表示方法及びプログラム,” 特願2011-168229, 2011年8月1日。

連絡先

丸谷 和史 (Kazushi Maruya) 人間情報研究部 感覚表現研究グループ

国際電話会議への参加を楽にする方法

～0.2～0.4秒の通信遅延が非母語者に及ぼす影響～

どんな研究

国際電話会議に非母語で参加することは大変困難です。対面会議では問題なく発言できても、電話会議になると**理解や発言頻度が低下**してしまいがちです。本研究では、電話会議中に非母語者が**過負荷な状態**に陥っていることに注目し、**負荷を軽減**するための支援方法を考案しました。

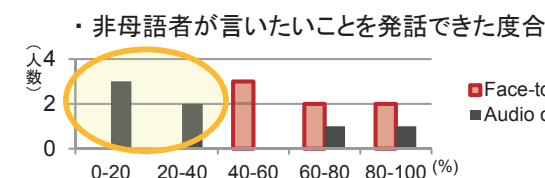
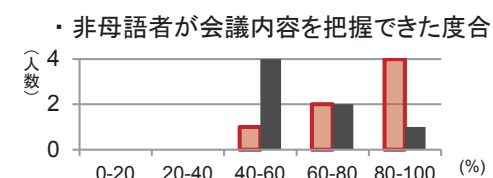
どこが凄い

母語が同じ話者同士の対話に対して、通信遅延がいかに問題であるかを示す研究は多数存在し、これまで通信遅延は除去するべきものとして扱われてきました。これに対して、本研究では、やり方によっては、**僅かな遅延が非母語者にかえって有益**になり得ることを示しました。

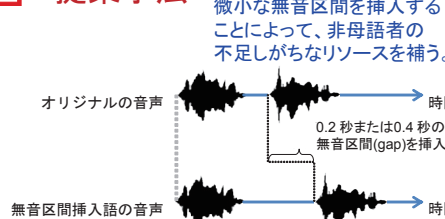
めざす未来

非母語者でも大きな負担なくリアルタイムの多人数会議に参加することができるようになります。これによって、**母語者と非母語者が対等**に意思決定プロセスに加わることができる社会の実現に貢献します。

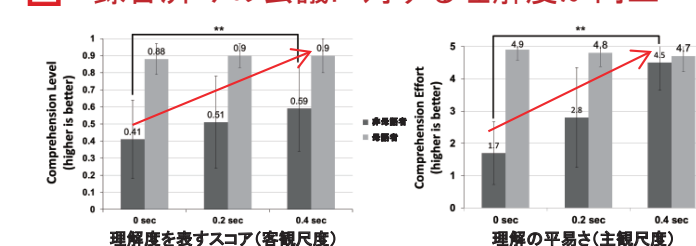
動機 日本を代表する研究者の7割が、英語の電話会議で言いたいことの半分も発言できないと感じている。



提案手法 母語者間の話者交代時に微小な無音区間を挿入することによって、非母語者の不足しがちなリソースを補う。



録音済みの会議に対する理解度が向上



リアルタイム電話会議への影響

・非母語者への影響

Gap Length	0 sec	0.2 sec	0.4 sec
Average rate of spontaneous speech	0.76	0.81 <i>best</i>	0.69 <i>worst</i>
Speaking effort (5pt Likert scale: <i>higher</i> is better)	3.1	3.4 <i>best</i>	2.3 <i>worst</i>
Comprehension effort (5pt Likert scale: <i>higher</i> is better)	4.1	4.1	3.5 <i>worst</i>
Frustration (5pt Likert scale: <i>lower</i> is better)	2.9	2.7 <i>best</i>	3.1 <i>worst</i>

・母語者への影響

Gap Length	0 sec	0.2 sec	0.4 sec
Speaking effort (5pt Likert scale: <i>higher</i> is better)	4.2	4.2	4.0 <i>worst</i>
Comprehension effort (5pt Likert scale: <i>higher</i> is better)	4.4	4.5	4.3 <i>worst</i>
Frustration (5pt Likert scale: <i>lower</i> is better)	1.7	1.8	2.1 <i>worst</i>
Percentage of conversations initiated simultaneously	22%	< 33%	< 41%
Frequency of speaker switches (# of speaker switches per min.)	7.8	> 7.3	> 6.6

録音済みの音声会議に無音区間を挿入した場合、非母語者の会話に対する理解度が有意に増加する。リアルタイムの電話会議では、無音区間の挿入によって非母語者に一定のメリットを生むものの、会話に不自然さが生じるといった負の効果も生じる。

関連文献

- [1] N. Yamashita, A. Echenique, T. Ishida, A. Hautsaari, “Lost in Transmittance: How Transmission Lag Enhances and Deteriorates Multilingual Collaboration,” in *Proc. The 16th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW’13)*, pp. 923-934, 2013.
[2] 山下直美, エチエニーケアンディ, 葛岡英明, 石田亨, ハウタサーアリ, “国際電話会議の負担を軽減する手法の提案と評価,” 情報処理学会論文誌, 2013 (to appear).

連絡先

山下 直美 (Naomi Yamashita) メディア情報研究部 コミュニケーション環境研究グループ

せつなの微笑みがこころを伝える

～対話者間の共感／反感に関する客観的解釈モデルの提案～

どんな研究

対話は社会的な生活を営む上で不可欠ですが、意思疎通の些細な齟齬から対人関係を損なうことも少なくありません。その問題解決を目指し、本研究では、対話中の**二者間の共感／反感**に関する**外部観察者の解釈**を、対話者の行動から**推定**可能なモデルを**世界で初めて提案**しました。

どこが凄い

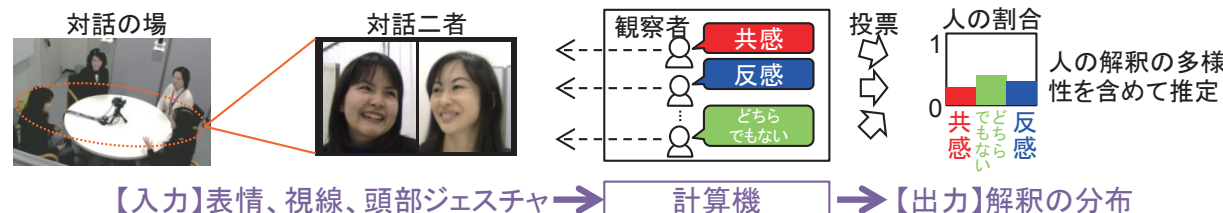
従来の研究がある**一人**の感情状態を推定しようとしていたのに対し、本研究では対話中の**二者の間**の関係性を推定するという新たな問題に取り組みました。さらに、感情という**曖昧な対象**に対し、**観察者の解釈のばらつき**の分布を推定するという新しい研究の枠組みを提案しました。

めざす未来

感情がどのような行動として**表出**され、それが他者にどのように**知覚**されるかの**現象の解明**を目指します。そこから、円滑な対話を実現するための遠隔対話システムの設計指針を提供し、また、コミュニケーション障がい現象の解明するための研究に役立てることが期待されます。

➤ 本研究の概要

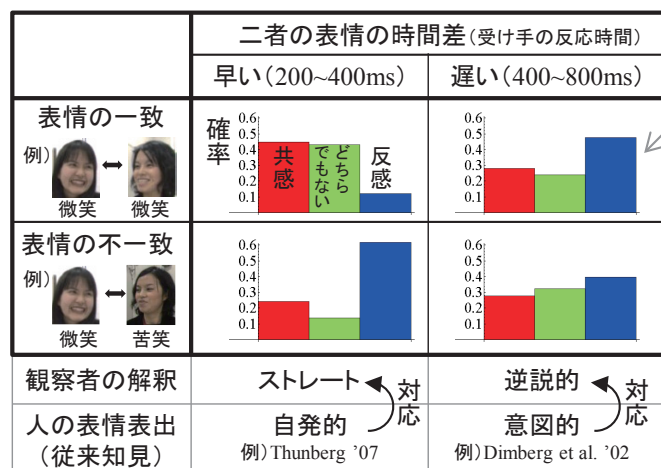
対話中の二者間の共感／反感に関して、外部観察者がどう解釈するのかを対話者の行動から推定する。



➤ 観察者の解釈を推定するカギ

(1) 二者の表情の一致性と時間差

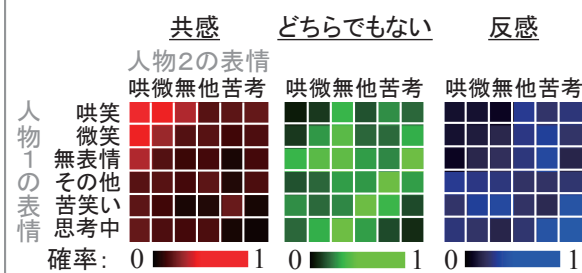
例) 笑顔に対して相手の笑顔が遅れると反感に見えやすい



(2) 二者の表情の組合せ (共起)

例) 笑顔と笑顔の組は共感に見えやすい
笑顔と苦笑いの組は反感に見えやすい

ポイント: この傾向を共起確率行列で記述した



※視線、頭部ジェスチャについても、共起の仕方では解釈の傾向が変化

➤ モデルと推定精度

- ・上記の関係性を確率モデル化 ・データ: 計70分間の対話に対する5~9名の観察者の解釈
- ・時間差情報を使用することで反感の推定精度が14ポイント向上 (66%→80%)

関連文献

- [1] S. Kumano, K. Otsuka, M. Matsuda, J. Yamato, "Analyzing perceived empathy/antipathy based on reaction time in behavioral coordination," in Proc. International Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EmoSPACE), 2013.
- [2] 熊野史朗, 大塚和弘, 松田昌史, 大和淳司, "二者間の行動の一致性と時間差に基づく外部観察者の共感解釈の分析," 電子情報通信学会HCGシンポジウム, pp. 330-337, 2012. (優秀インタラクティブ発表賞受賞)

連絡先

熊野 史朗 (Shiro Kumano) メディア情報研究部 コミュニケーション環境研究グループ

録音機器を協調させて目的音声を聞き分けます

～録音ノードからの仮説統合に基づく音声強調～

どんな研究

実環境で収録された音声には、目的音声とともに雑音（背景雑音や目的話者以外の音声）が含まれます。ここでは、**複数の録音機器** (ICレコーダやパソコン等) を**協調**させ、これらの**雑音成分を除去**し目的音声をきれいに抽出する技術※を紹介します。

(※NTT MD研と協調して研究推進)

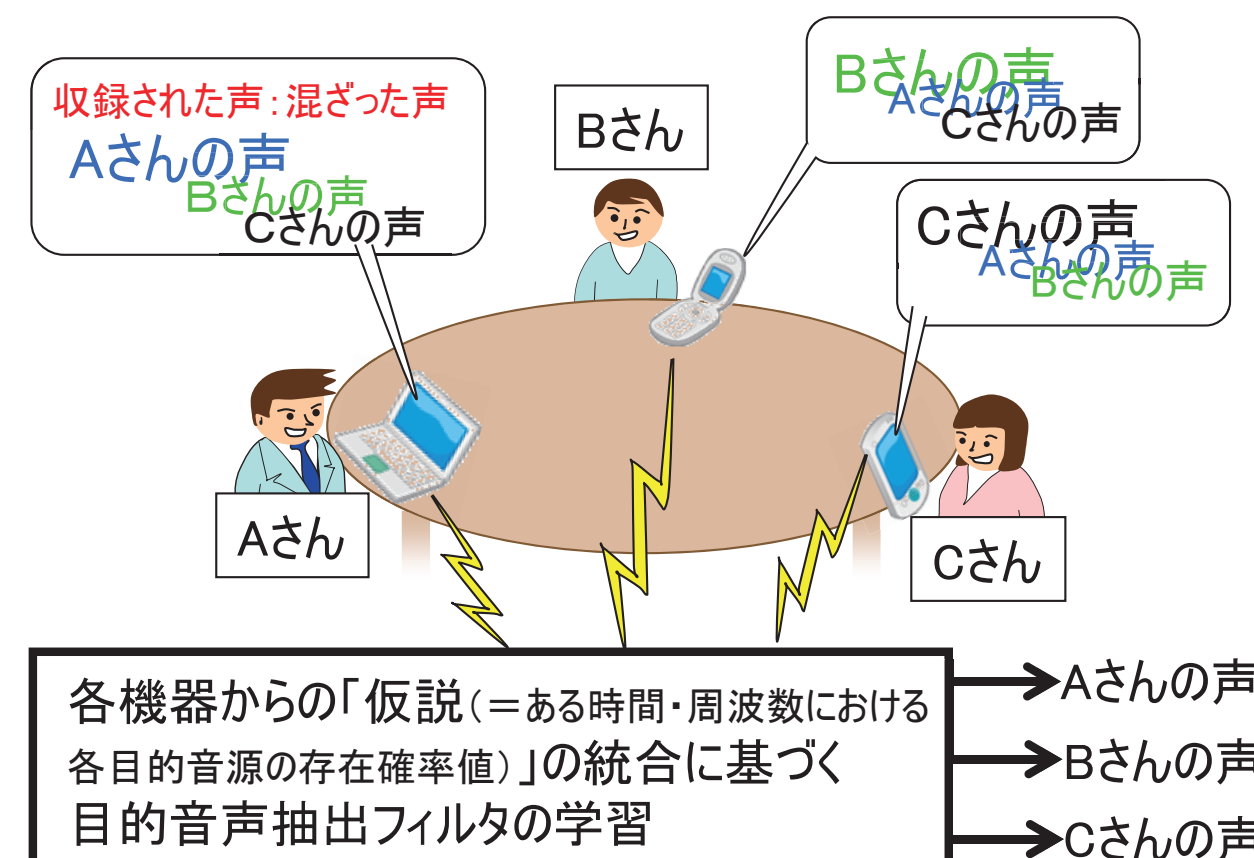
どこが凄い

下図のような環境で、全録音機器を単純に用いて従来技術を動作させても、録音機器の特性 (例: 標準化周波数) の違いのため、効果的な目的音抽出はできません。そこで、**各録音機器内で目的音抽出**を行い、それらの**仮説を統合**し、より**高精度な目的音抽出**を行う方法を開発しました。

めざす未来

身近にある録音機器を互いに協調させて様々な音声や目的音を効果的に抽出することができるになれば、専用の大きな録音機器を用いずとも、**あらゆる場面できれいな目的音の収録**や、より**精度の高い音声認識**が可能となり、音声アプリケーションの応用範囲が大きく広がります。

問題: ICレコーダやパソコン等を使って収録をすると…目的音声以外の雑音も収録される ☹️



課題解決方法: 複数の録音機器を協調させて目的音声のみを取り出す

関連文献

- [1] M. Souden, K. Kinoshita, T. Nakatani, "An integration of source location cues for speech clustering in distributed microphone arrays," in Proc. International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2013.
- [2] ソウデン メレス, 木下慶介, 中谷智広, "ノード内・ノード間情報の統合に基づく分散マイクアレイ音源分離," 日本音響学会春季研究発表会, pp. 797-798, 2013.

連絡先

木下 慶介 (Keisuke Kinoshita) メディア情報研究部 信号処理研究グループ

誰がどのように話しても正確に聞き取ります

～話者や発話スタイルの多様性に頑健な音声認識技術～

どんな研究

音声認識技術は高度な進化を遂げ、丁寧な話し方で正しい文を話せば高い精度で認識できるようになってきています。しかし、未だに話す人や発話の丁寧さ等によって性能が上下してしまいます。本展示ではこのように**多様な音声を適切に処理**する音声認識技術について紹介致します。

どこが凄い

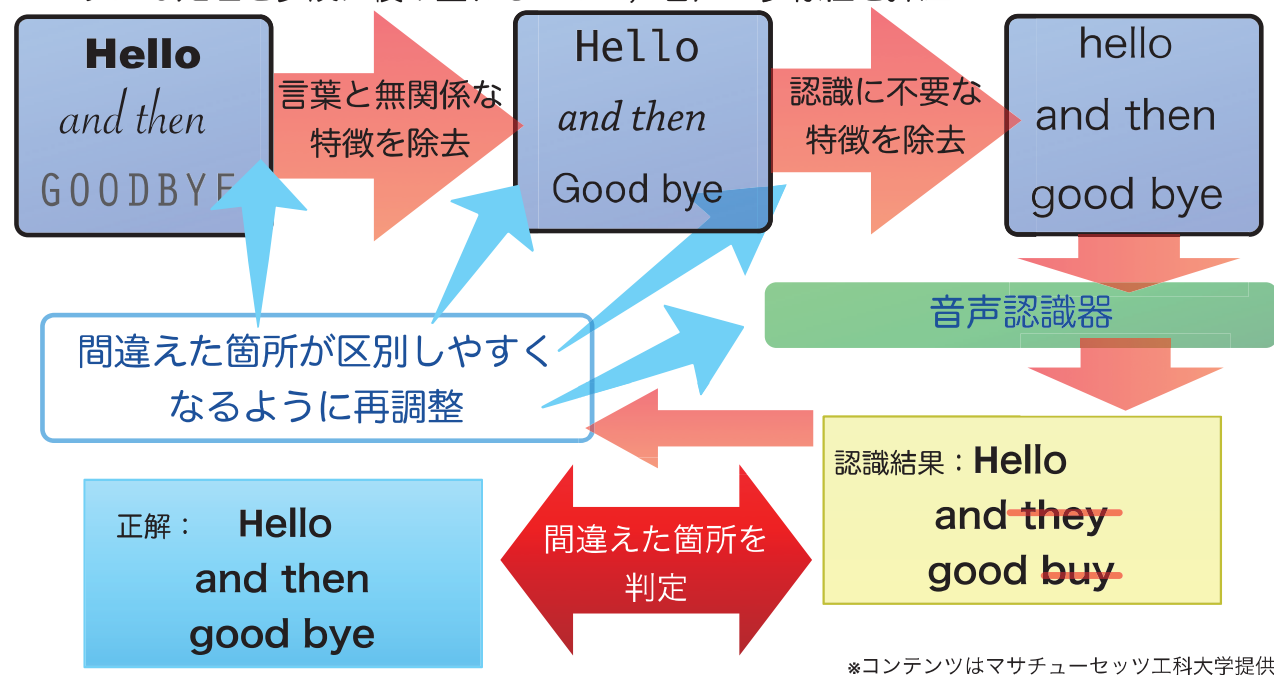
音の多様性を吸収する**多段の情報抽出方法**を大量の音声データから自動で学習し、その上で、音声と言語の対応を、音声言語一体型統計モデルにより表現する手法を考案しました。この手法で音声と言語の両面から多様性を吸収することで、世界トップクラスの性能を実現しました。

めざす未来

誰が何処で何をどのように話しても正確に聞き取ることのできる、頑健な音声認識技術を目指します。このような音声認識技術を用いることで、対話エージェント技術、音声データマイニング技術のさらなる高度化を可能にし、**様々な情報機器を自然に活用できる社会**を目指します。



シンプルな処理を多段に積み重ねることで、音声の多様性を抑圧



関連文献

- [1] Y. Kubo, T. Hori, A. Nakamura, "Integrating deep neural networks into structured classification approach based on weighted finite-state transducers," in *Proc. 13th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2012.
 [2] Y. Kubo, T. Hori, A. Nakamura, "Large vocabulary continuous speech recognition based on WFST structured classifiers and deep bottleneck features," in *Proc. The 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.

連絡先

久保 陽太郎 (Yotaro Kubo) メディア情報研究部 信号処理研究グループ

知らない音も書き起こす

～ベイジアン半教師あり音響イベント認識～

どんな研究

膨大なメディアデータを自在に検索して活用するためには、メタデータに頼るだけでなく、音や映像自体から中身の情報を自動的に引き出す技術が必要不可欠です。この展示では、音響信号に含まれる**音響イベント（音によって人に認識される事象）**を書き起こす技術を紹介し

どこが凄い

一部区間のラベルを手がかりとして、重なり合った音響イベントの全区間を特定します。さらに、音響信号に存在する音響イベントの数を学習することにより、ラベルに無い音を、「**知らない音（未知の音響イベント）**」として、書き起こすことができます。

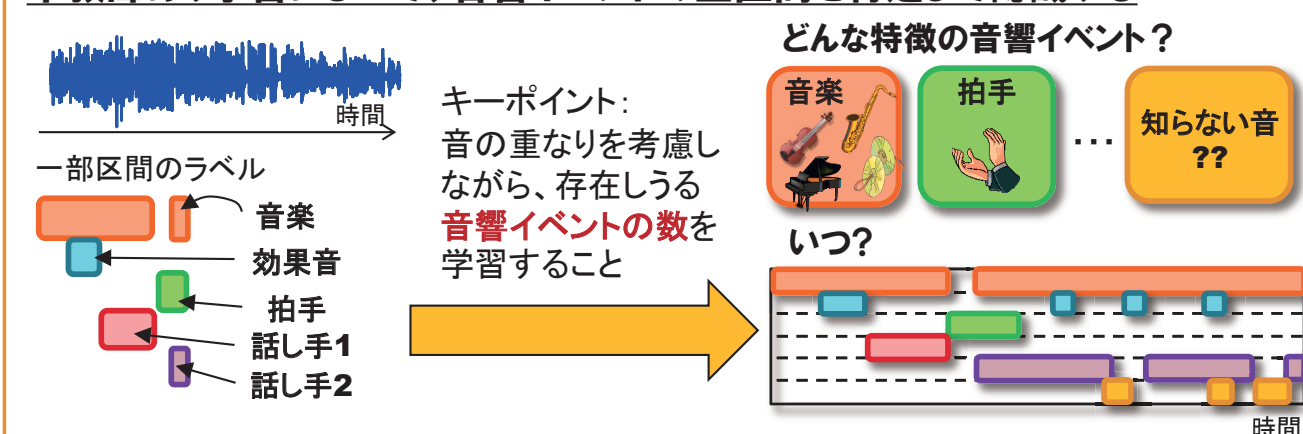
めざす未来

人間の耳のように、どんな音が、どこで、どんな風に鳴っているかを聞き分けて認識することで、クラウドにおける大量の音コンテンツの可視化を目指します。これにより、**特定の音を含む動画検索や危険な音を知らせてくれる監視・行動支援システム**など、様々な応用が期待されます。

「どんな」音響イベントが、「いつ」鳴っているか認識したい



半教師あり学習によって、音響イベントの全区間を特定して認識する



適用例(期待される効果)



関連文献

- [1] Y. Ohishi, D. Mochihashi, T. Matsui, M. Nakano, H. Kameoka, T. Izumitani, K. Kashino, "Bayesian semi-supervised audio event transcription based on Markov Indian buffet process," in *Proc. 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.
 [2] 大石康智, 持橋大地, 松井知子, 中野允裕, 亀岡弘和, 泉谷知範, 柏野邦夫, "mIBPに基づくベイジアン半教師付き音響イベントダイアライゼーション," 日本音響学会2013年春季研究発表会, 1-9-6, pp. 19-22, 2013.

連絡先

大石 康智 (Yasunori Ohishi) メディア情報研究部 メディア認識研究グループ

目的のものが映っているシーンを探す

～静止画クエリによる映像検索技術～

どんな研究

特定の人物/物体/場所などの静止画を検索キー（クエリ）として、目的のものが映っているシーンを大量の映像の中から検索する技術の研究です。このような映像検索は**インスタンスサーチ**と呼ばれ、映像検索分野の国際的なコミュニティTRECVIDの重点課題の一つになっています。

どこが凄い

新たに確率的なコンテンツベースの映像検索手法を考案しました。本技術では検索クエリであるインスタンスと各映像間の適合性を確率的に評価し、大量の映像をランキングします。これによりTRECVIDのデータを使用した実験で、**現在世界最高レベルの検索精度**を記録しています。

めざす未来

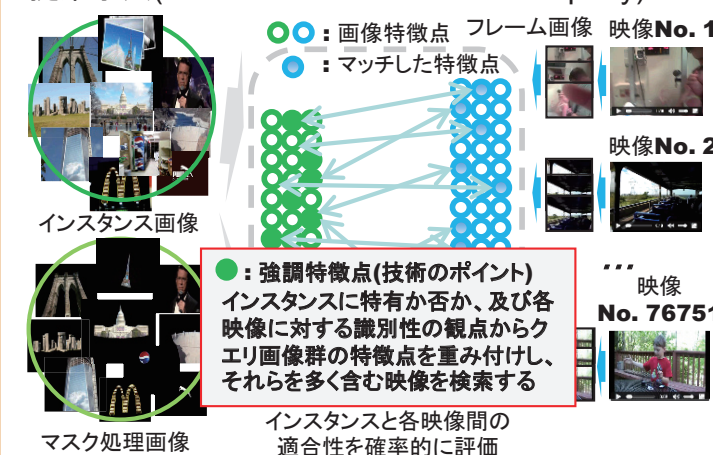
雑音があったり、目的のものの見た目が異なっても情報検索を可能にするスマートなメディア探索技術の実現をめざします。この技術は、身のまわりの音・画像・映像を用いて、大量の音や映像の中から必要な情報を瞬時に探し出せるメディア社会の実現に寄与します。

・インスタンスサーチ

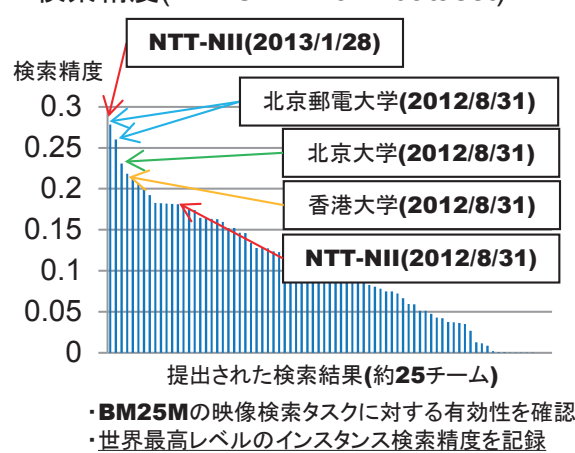


インスタンス = 特定の人物・物体・場所
「システムはインスタンスが映っている映像をデータベースから検索し、適合度順でランキングして検索結果をユーザに提示する」

・提案手法(BM25M: BM25 with masked query)



・検索精度(TRECVID 2012 dataset)



関連文献

- [1] M. Murata, T. Izumitani, H. Nagano, R. Mukai, K. Kashino, S. Satoh, "NTT Communication Science Laboratories and National Institute of Informatics at TRECVID2012: Instance Search and Multimedia Event Detection Tasks," in *Proc. TRECVID2012 Workshop*, 2012.
- [2] 村田真哉, 永野秀尚, 向井良, 柏野邦夫, 佐藤真一, "画像をクエリとしたインスタンス映像検索," 信学技報 *PRMU2012-175*, pp. 215-219, 2013.

連絡先

村田 真哉 (Masaya Murata) メディア情報研究部 メディア認識研究グループ

この子にぴったりの絵本を探す

～幼児単語に重みを置いたグラフベースの類似絵本検索～

どんな研究

子どもの言語・認知発達を促す手段として、**絵本の読み聞かせ**が注目されています。今回、たくさんの出版された絵本の中から子どもにあった絵本を容易に選べるようにするために、今読んでいる絵本と**類似の絵本を簡単に検索できるシステムを構築**しました。

どこが凄い

これまで困難だった**ひらがな・カタカナが多く含まれるテキストに強い形態素解析器**を開発し、絵本の高精度な解析が可能になりました。さらに、幼児の各発達段階での単語情報を加味するとともに、我々の開発した**グラフ探索アルゴリズム**による、高速な類似探索を実現しています。

めざす未来

絵本を介した子どもの国語力アップのための教育プログラムでの利用、子どもや養育者を対象にした絵本購入のための推薦システム、それぞれの子どもに合った新しい絵本の創作支援システムなどへの応用が考えられます。

・絵本の読み聞かせ

- ・子どもの言語・認知発達に良い影響
- ・しかし、適切な絵本選びは困難

・絵本選びのための類似絵本検索システム

絵本同士の関連性が素早くわかる



出現する単語分布の近さをもとにグラフを作成

ポイント1 ひらがな・カタカナに強い形態素解析器

- ・大規模な辞書を使い、ひらがな・カタカナ語の元の漢字を考慮し、解析器を再学習

入力文(絵本のテキスト)

ぼくらの もりに はるが きたよ

(小手鞠るい著、たかすかずみ絵「はじめてのもり」、金の星社、2010年より)

従来の形態素解析システム(かな部分で誤り)

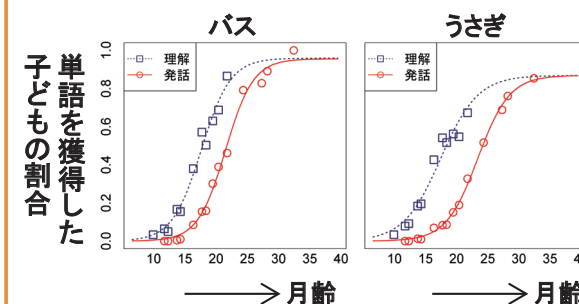
ぼく	ら	の	もり	に	はる	が	きた	よ
名詞	接尾	名詞	動詞	助詞	動詞	助詞	名詞	助詞
ぼく	ら	の	も	に	は	が	きた	よ

今回開発した形態素解析システム(正しく解析)

ぼくら	の	もり	に	はる	が	き	た	よ
名詞	助詞	名詞	助詞	名詞	助詞	動詞	助動詞	助詞
僕ら	の	森	に	春	が	来る	た	よ

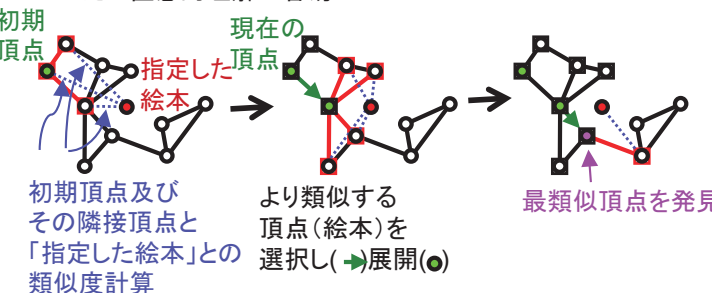
ポイント2 幼児の語彙発達研究の成果を活用

- ・幼児の各月齢で、どのような単語を理解しているかについての研究データの活用



ポイント3 グラフを索引とする類似探索アルゴリズム

- ・有力な従来法(ハッシュ法)と比べても6倍以上高速
- ・検索結果表示を従来のリストだけでなくグラフでも行えるため直感的理解が容易



関連文献

- [1] 平博順, 藤田早苗, 小林哲生, "絵本テキストに現れる語彙の分布," 日本赤ちゃん学会第12回学術集会, p. 92, 2012.
- [2] 平博順, 藤田早苗, 小林哲生, "絵本テキストにおける高頻度語彙の分析," 平成24年度情報処理学会関西支部大会, F-103, 2012.
- [3] 藤田早苗, 平博順, 小林哲生, "絵本のテキストを対象とした形態素解析," 言語処理学会第19回年次大会, A6-4, 2013.
- [4] K. Aoyama, K. Saito, H. Sawada, N. Ueda, "Fast approximate similarity search based on degree-reduced neighborhood graphs," in *Proc. 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1055-1063, 2011.

連絡先

平 博順 (Hirotoishi Taira) 協創情報研究部 言語知能研究グループ

こどもが単語を覚える順番の謎

～幼児語彙学習期間の言語共通性～

どんな研究

チェックリストを使って、いつどんな単語を覚えたかを多人数の幼児を対象に調べました。このデータから、単語を**理解する日齢**、および、**理解してから発話するまでの日数**を求めた結果、日本語、英語とスペイン語[3]で、単語を覚える順番に共通性があることを発見しました。

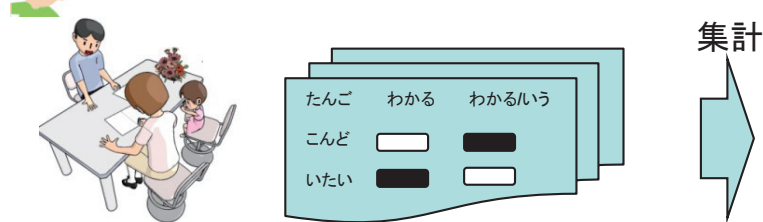
どこが凄い

これまで、幼児は動詞より名詞を早く覚えると言われていましたが、少数の幼児を調べる、あるいは、品詞の頻度を求める以外で、この証拠を示した例はありません。今回単語を覚える時期を特定することにより、カテゴリや単語の間に**言語共通の順序**があることを初めて見出しました。

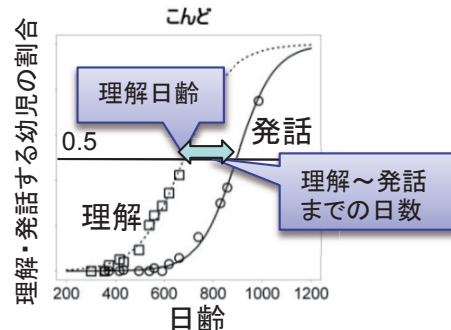
めざす未来

語彙学習原理を解明することにより、ヒトでのみ進化した言語能力の解明を実現します。この過程で得られた知見をもとに、子どもの健やかな知的発達を後押しする社会システムの実現に寄与します。また、解明された原理をもとに、ヒトをモデルとする新たな計算原理を確立します。

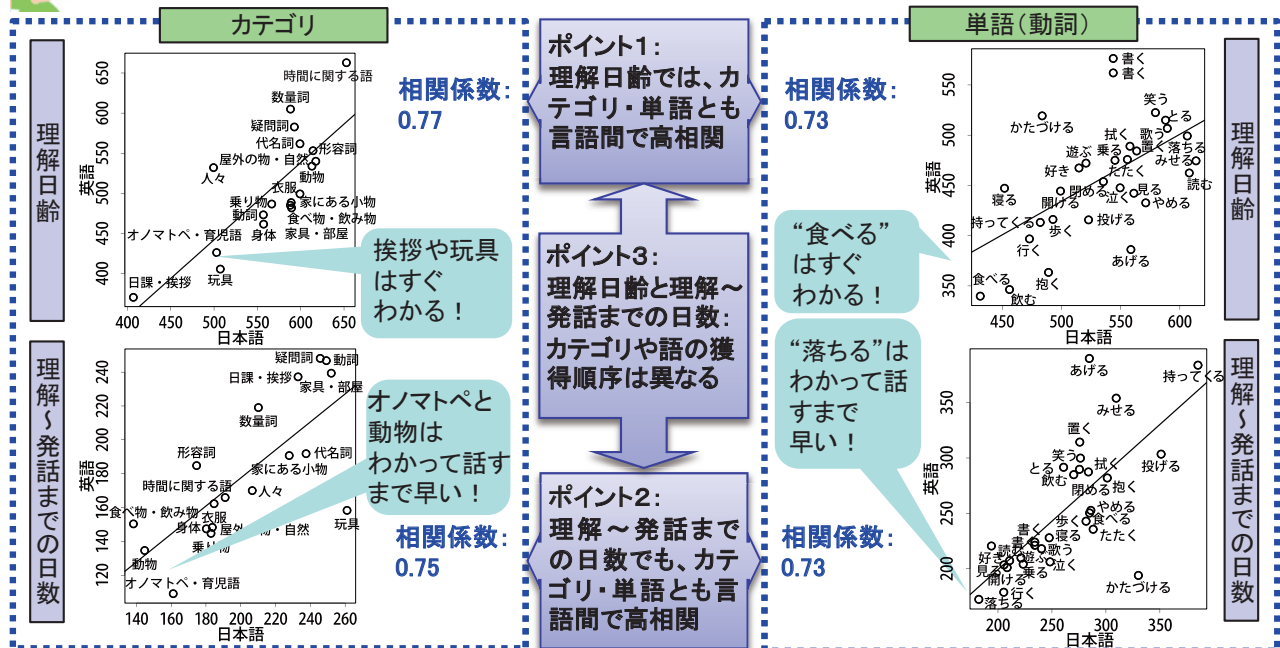
どんな言葉をいつ覚えるかをチェック



集計



複数の言語でカテゴリ、単語(動詞)毎に順序を比較



- ・こどもが単語を覚える順番: 言語によらず共通性がある。
- ・理解日齢と理解～発話までの日数の間では、単語の学習順番が大きく異なる。

関連文献

- [1] Y. Minami, T. Kobayashi, H. Sugiyama, "Cross-linguistic universality of word acquisition ages in comprehension and production," in *Proc. WILD*, 2013.
- [2] Y. Minami, T. Kobayashi, "Influence of predominance in noun learning examined by period from comprehending to producing words: A cross-linguistic statistical investigation using CDI," in *Proc. WILD*, 2013.
- [3] P. S. Dale, L. Fenson, "Lexical development norms for young children," in *Proc. BRMIC*, Vol. 28, pp. 125-127, 1996.

連絡先

南 泰浩 (Yasuhiro Minami) メディア情報研究部 コミュニケーション環境グループ

「花」と「鼻」、こどもはいつから聞き分ける？

～幼児の日本語音声知覚の発達～

どんな研究

日本語には、**他の言語とは異なる特有の音声構造**があります。この音声構造を日本語を母語とする幼児がどのように知覚しているのか、年齢によって知覚の精度がどのように変化するかを**タブレットを使った行動実験**によって捉え、**日本語音声知覚の発達過程**を詳細に検証しました。

どこが凄い

幼児に簡単に適用できる実験手法をタブレットに実装することで、**客観的なデータを効果的に取得**できるようになりました。さらに、定型発達児と**言語聴覚療法を受けている幼児を比較**することで、ことばの遅れのある子どもがつかみやすい知覚の難しい音声特徴を明らかにしました。

めざす未来

これまで未解明の日本語音声知覚の初期発達過程を比較検討することで、ことばの獲得と個々の音声特徴の知覚との関係を明らかにし、**発達・教育分野への基礎的知見の提供**を目指します。特に、言語聴覚療法・発達スクリーニングといった種々の**臨床的場面への貢献**が期待できます。

日本語特有の音声構造

英語と日本語に共通の音声対立

例) bass vs. pass ≡ バス vs. パス

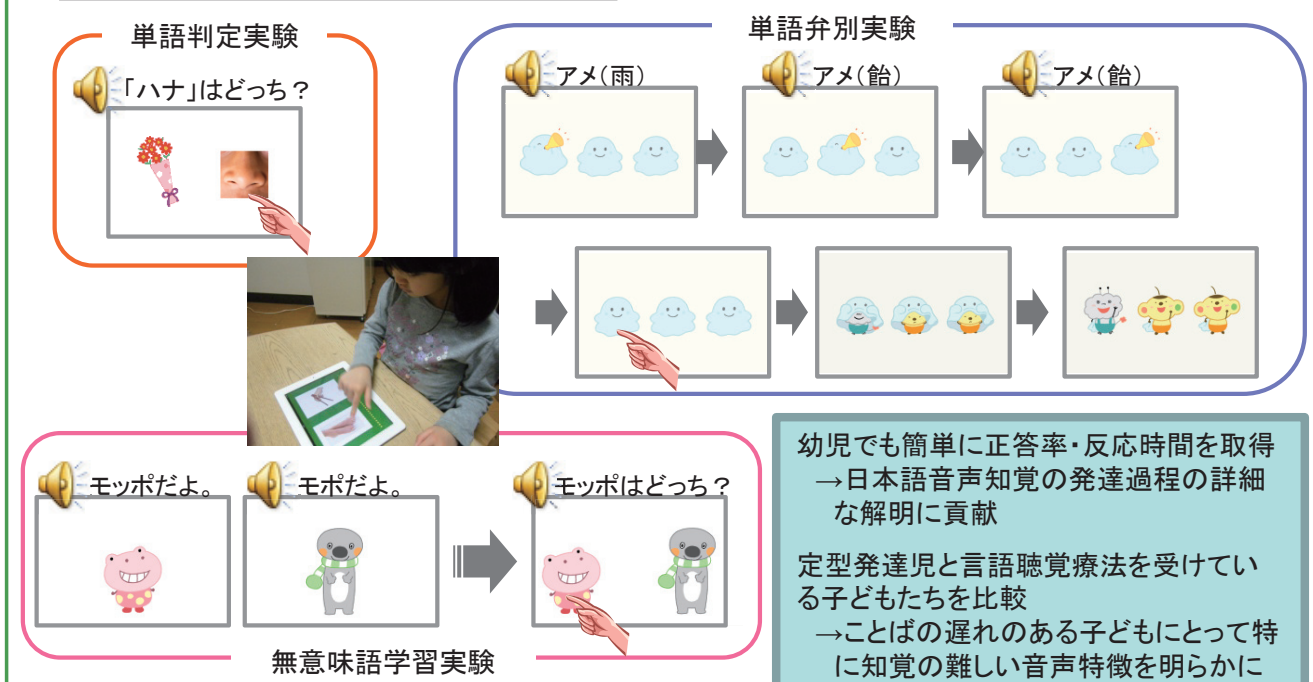
英語特有の音声対立

例) rink vs. link

日本語特有の音声対立

- アクセント 例) 時 vs. 朱鷺
- 母音長 例) 朱鷺 vs. 冬季
- 促音 例) 時 vs. 突起

タブレットを使った種々の音声知覚実験



幼児でも簡単に正答率・反応時間を取得
→日本語音声知覚の発達過程の詳細な解明に貢献
定型発達児と言語聴覚療法を受けている子どもたちを比較
→ことばの遅れのある子どもにとって特に知覚の難しい音声特徴を明らかに

関連文献

- [1] R. Mugitani, A. Hayashi, "Word-identification and discrimination of Japanese pitch accent in preschoolers," in *Proc. The 36th Annual Boston University Conference on Language Development (BUCLD)*, 2011.
- [2] 麦谷綾子, 林安紀子, "3～5歳の幼児における日本語特有音声の知覚特性," 日本音響学会春季研究発表会論文集, 2012.

連絡先

麦谷 綾子 (Ryoko Mugitani) 人間情報研究部 感覚共鳴研究グループ

呼吸と音楽の一期一会をめざして

～呼吸の位相と音楽的フレーズを同期させる再生システム～

どんな研究

同じ音楽を聴いていても、異なるように聞こえる時があると思います。音楽を聴くと呼吸が変化することが知られていますが、我々はその呼吸変化が聞こえ方に影響する可能性に着目しました。**音楽を呼吸に合わせて再生することで新たな音体験の可能性を探っています。**

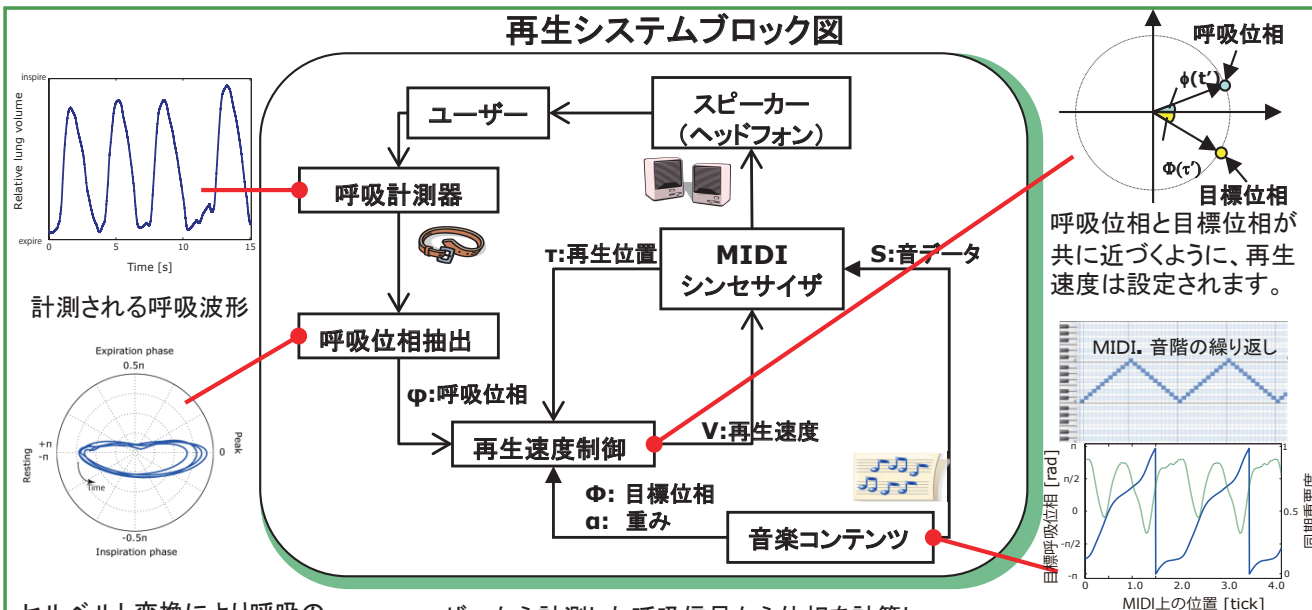
どこが凄い

あらかじめ音楽に追従するべき呼吸の**目標位相と同期重要度**を設定することで、音楽のフレーズを柔軟にユーザーの呼吸位相に一致させる再生法を新たに提案し実装しました。単純な音階を使った実験データからは、心理的な影響が生じることが示唆されています。

めざす未来

呼吸に限らず、**生体情報をフィードバックする音楽提示**を行うことで、新たな音楽再生法を研究します。従来にはない音の提示法により、既存の音楽の楽しみ方の選択肢を増やすだけでなく、新たな音楽分野を創設できる可能性があります。

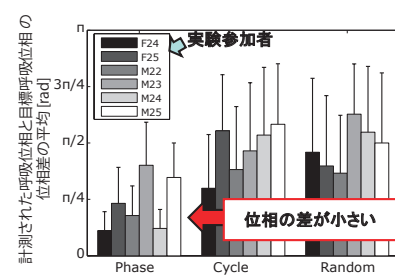
再生システムブロック図



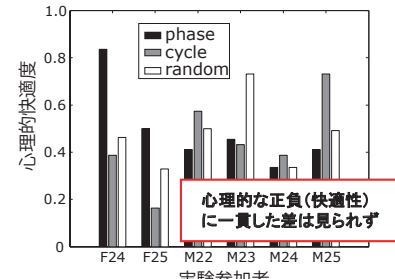
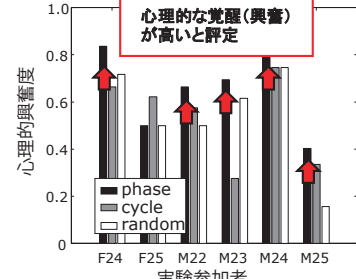
ヒルベルト変換により呼吸の位相情報を取り出します。

ユーザーから計測した呼吸信号から位相を計算し、その位相と、音楽コンテンツにあらかじめ設定した目標呼吸位相が近づくよう再生速度を調節します。

音楽データには目標とする呼吸の位相も設定しておきます。



世界初の試みとして目標位相と呼吸位相角を近づけて再生する方法を提案し、音階の繰り返し音で実験を行いました。本方式(phase)、音階周期を平均呼吸周期に合わせる場合(cycle)と、ランダムに変化させる場合(random)で再生し比較しました。提案する再生方式が目的通り動作したことのみにならず、再生方式の違いで心理的な覚醒度(arousal)に影響を与える可能性を見出しました。



関連文献

- [1] 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.
- [2] T. G. Sato, Y. Kamamoto, N. Harada, T. Moriya, "A playback system that synchronizes the musical phrases with listener's respiration phases," in *Proc. The ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2013.

連絡先

佐藤 尚 (Takashi G. Sato) 守谷特別研究室

それっぽくしゃべります

～非負値時空間分解法に基づく発話リズムの抽出～

どんな研究

外国語でのコミュニケーションの困難さは、しゃべること、聞くことの両方において存在します。我々は、言語による**発話リズム**の違いを考慮することが、外国語の正しい聞き取りや発話において重要であると考え、音声信号から発話リズムを抽出し、制御する技術を考案しました。

どこが凄い

発話時の音声信号から発話リズムを抽出するための**非負値時空間分解法**というアルゴリズムを考案しました。従来法（非負値行列因子分解）にはない音声に特化した制約を考慮しながら、高い精度で発話リズム（時間情報）を抽出でき、また発話リズムを制御できることが分かりました。

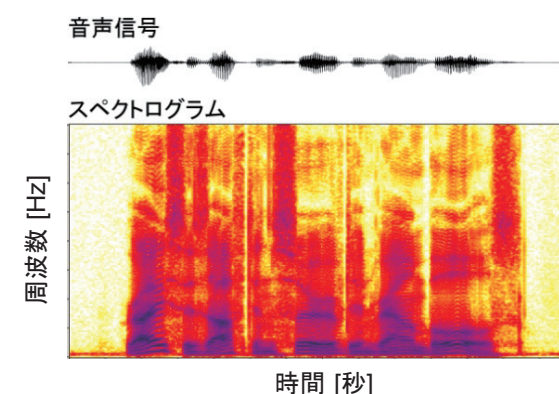
めざす未来

本技術は、外国語の聞き取りにおける発話リズムの重要性の検証のために大きく貢献できると考えられます。将来、日本語母語話者の英語の発話リズムを変換することで、英語が英語母語話者に伝わりやすくなり、英語でのコミュニケーションの苦勞が軽減されるかもしれません。

スペクトログラムを周波数情報と時間情報に分解

$$Y \cong A \Phi$$

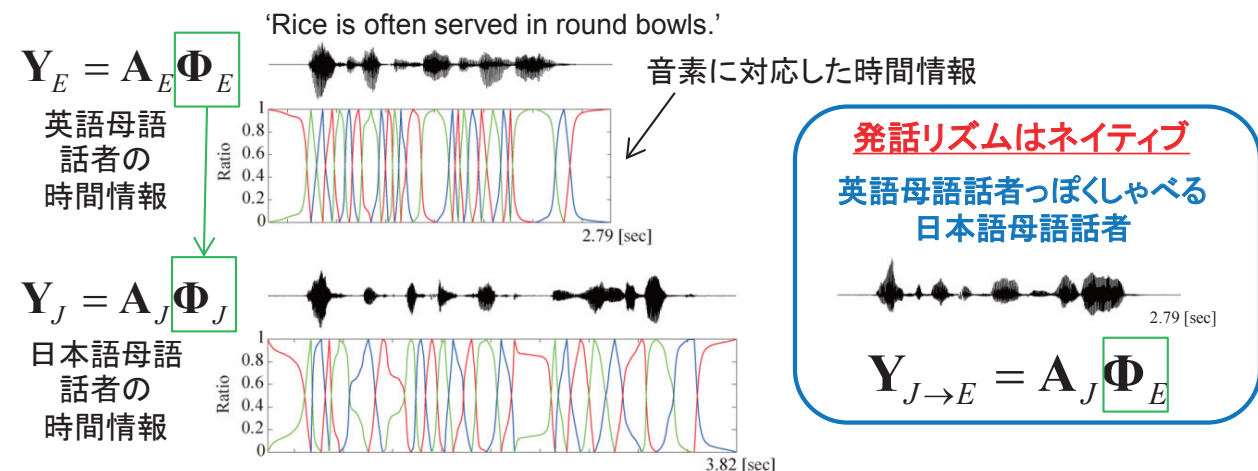
スペクトログラム 周波数情報 時間情報 = 発話リズム



音声であることを考慮した時間情報への制約

- ①発話リズム(時間情報)を音素単位で表現するモデルを導入
- ②隣り合う音素のみが時間情報に影響を及ぼすこと(調音結合)を考慮

時間情報を入れ替えることによりそれっぽくしゃべらせる



関連文献

- [1] S. Hiroya, "Non-negative temporal decomposition of speech parameters," in *Proc. The 35th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010.
- [2] S. Hiroya, T. Kitamura, "Generation of a vocal-tract movie based on sparse sampling," in *Proc. The 9th International Seminar on Speech Production (ISSP)*, 2011.

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

見ることで感じる疲れや手ごたえ

～映像遅延により変化する運動感覚～

どんな研究

ネットワークゲームなどのインタラクティブなサービスでは、**通信遅延**に起因する映像の遅れによって、抵抗感などの不快な感覚が生じることがあります。このような**操作感の変化**を引き起こす脳の情報処理のしくみを、心理物理実験によって調べました。

どこが凄い

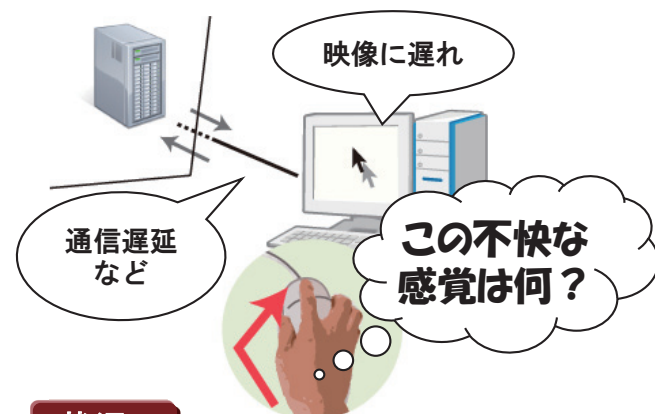
人は操作対象の映像の反応の遅れを、状況に応じて、筋疲労や重りを引く力として知覚することを明らかにしました。この発見は、運動中の主観的な体験が、視覚情報を利用した**脳の潜在的なダイナミクス推定**（状況解釈）のプロセスによって形づくられている可能性を示しています。

めざす未来

今回得られた脳内情報処理のしくみに関する知識は、通信の遅れがあっても**ストレスなく操作できるインターフェイス**を実現する指針になります。また、バーチャルリアリティにおける**リアルな感覚の提示**、リハビリテーションでの疲労感軽減などへの応用も期待できます。

操作中に映像遅延

⇒ **操作感の変化**として知覚



状況1

指の動きの映像が遅れている場合

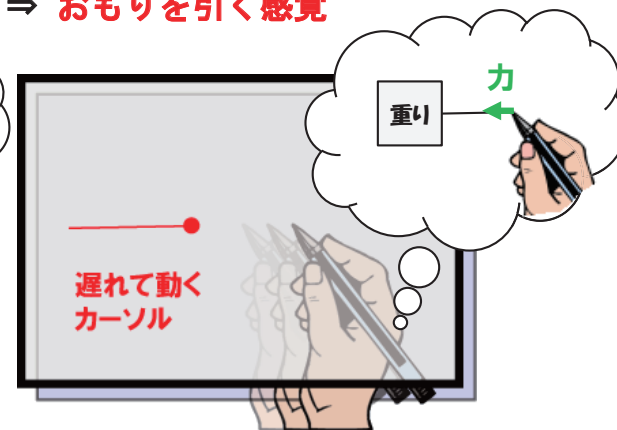
⇒ **指の疲労感**



状況2

カーソルの動きが遅れている場合

⇒ **おもりを引く感覚**



関連文献

- [1] S. Ito, T. Kimura, H. Gomi, "Enhancement of fatigue perception and phase lag of cyclic movement induced by delayed visual feedback," in *Proc. Society for Neuroscience 42nd Annual Meeting (SfN)*, 2012.
 [2] S. Takamuku, H. Gomi, "Principle factor yielding the sluggish sensation during movements with delayed visual feedback," in *Proc. Society for Neuroscience 42nd Annual Meeting (SfN)*, 2012.

連絡先

伊藤 翔 (Sho Ito) 人間情報研究部 感覚運動研究グループ

未来が過去を変える

～視覚におけるポストディクション～

どんな研究

過去のモノの見え方が、未来のモノの見え方によって変わってしまう錯覚現象を発見しました。物理的な世界はめまぐるしく変化します。変化する世界の情報を上手にまとめることで私たちの知覚を安定させる脳のしくみに一歩迫りました。

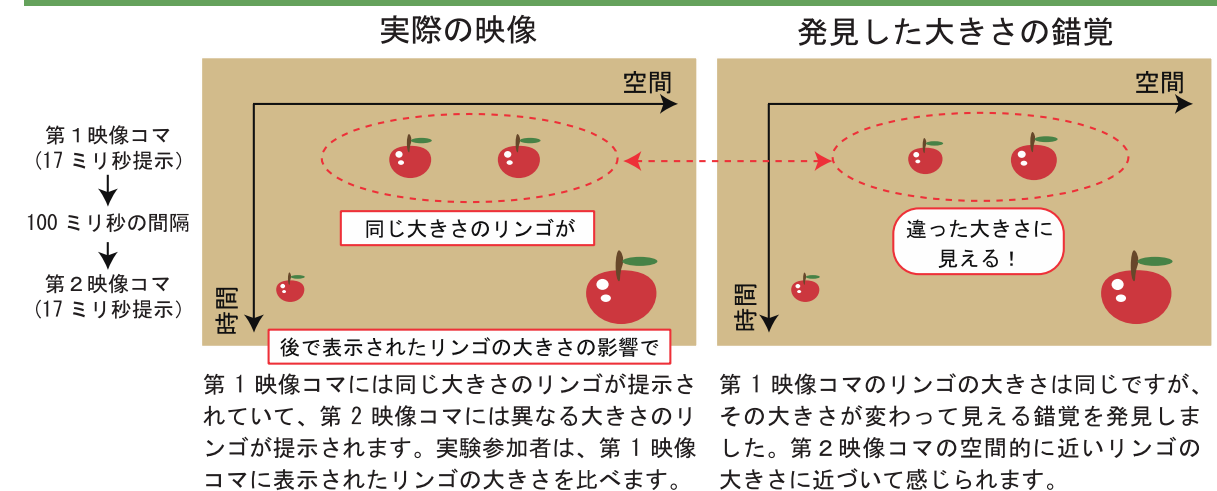
どこが凄い

二つの物体の絵が離れた位置に順番に提示されると、一つの物体が動いたように見えます。この時、脳は二つの物体を一つの物体にまとめます。二つの絵の特徴が違うと、その平均化を試みます。すると、はじめに提示された物体の絵が後から提示された絵の特徴に引っ張られるのです。

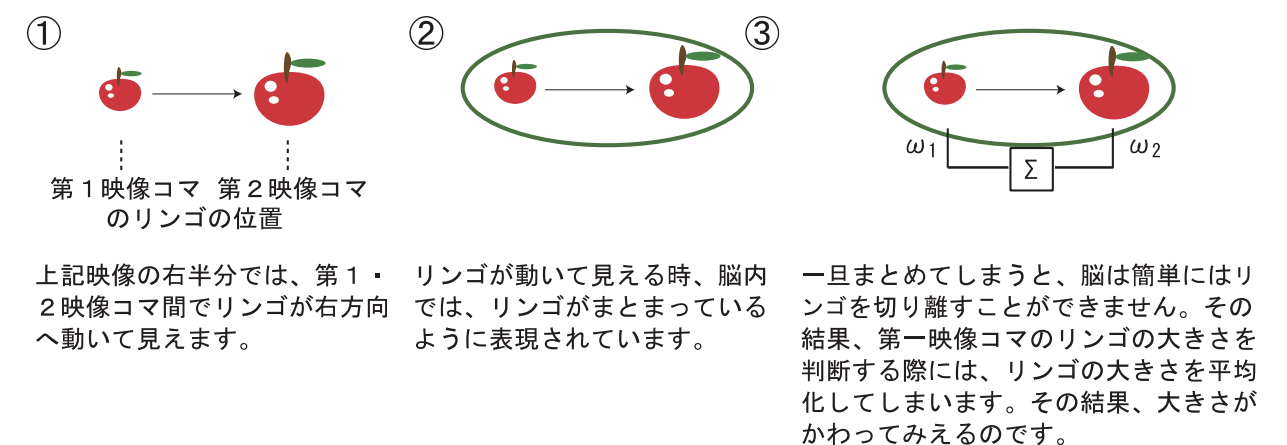
めざす未来

映像の見え方を支える脳の仕組みを解明することで、**人間の特性を利用した情報提示技術の実現**へとつながります。例えば映像中の対象の見間違えを減らす手法の提案や、逆に錯覚を積極的に利用して、映像中の対象の見え方を適切に操作する技術の提案につながります。

未来の情報によって過去の情報の見え方が変わる錯覚の例



想定されるメカニズムの説明



関連文献

- [1] T. Kawabe, "Nonretinotopic processing is related to postdictive size modulation in apparent motion," *Attention, Perception, & Psychophysics*, 73, pp.1522-1531, 2012.
 [2] T. Kawabe, "Postdictive modulation of visual orientation," *PLoS ONE*, 7(2): e32608, 2012.

連絡先

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

見るだけでモノの形や質感が変わる

～錯視からわかる脳の物体認識のしくみ～

どんな研究

ある物体の画像を見続けると、その後で見る物体の形や質感が劇的に変わって見える錯視（物体残効）を発見しました。物体残効は、二次元的な画像特徴を見ただけでも起こります。また、色の残像などと違って、網膜上の位置関係にあまり関係なく起こります。

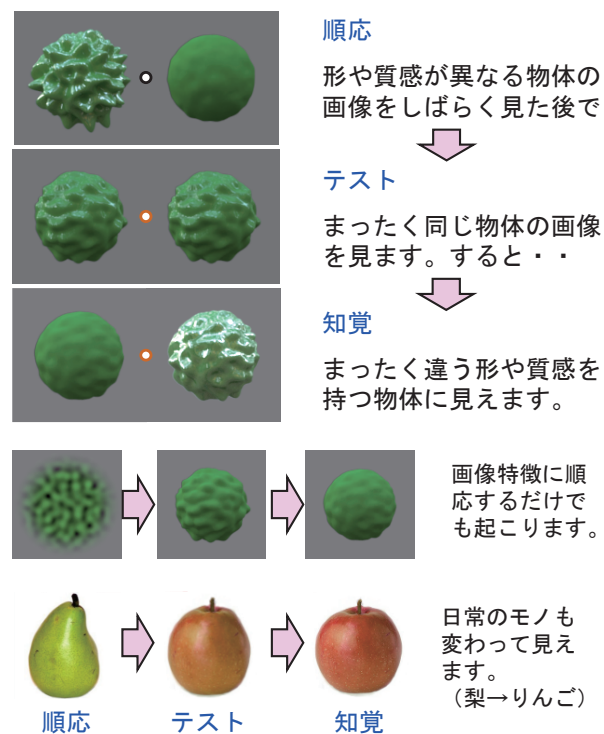
どこが凄い

これまで、脳が複雑な物体を認識するには、二次元の映像から三次元の形状を「復元」することが不可欠だと考えられてきましたが、物体残効錯視の発見により、**脳**が二次元画像の特徴の寄せ集めから物体を認識するメカニズムをもっていることを明らかにしました。

めざす未来

認識の決め手となる画像特徴の種類が理解できれば、**画像特徴**を操作するだけで物体の形や質感を変えたり、よりリアルに見せることも可能になると考えられます。また、現在の高速な**物体認識**・検索のアルゴリズムをより人間の知覚に即したものにできるかもしれません。

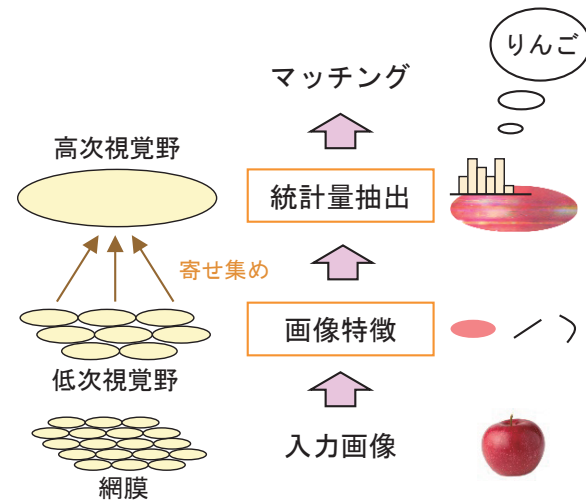
物体残効錯視 (Object aftereffect)



物体残効は、テスト画像が順応画像とまったく違う位置に現れても強く起こります。

つまり、質感残効は、視野の一部だけを担当する網膜や低次の視覚メカニズムではなく、視野の広い範囲を担当する高次のメカニズムの特性が変化したために起こる錯視といえます。しかも、その高次のメカニズムは低次の画像特徴を扱っていると考えられます。

脳内における画像特徴からの物体認識



物体残効錯視の分析結果は、脳は物体の画像の持つ特徴を寄せ集めて、その統計的な性質から物体の形や色、質感、名称までも認識している可能性を示しています。これは最近の機械学習モデルに似ていて、かなり簡単な計算で実行できます。

関連文献

- [1] I. Motoyoshi, "Visual aftereffects in 3D shape and material of a single object," *Journal of Vision*, 12(9): 229, 2012.
- [2] I. Motoyoshi, "Broad spatial tunings of the object aftereffect: Evidence for global statistical representations of 3D shape and material," *Perception*, 41, 17a, 2012.
- [3] I. Motoyoshi, "Visual aftereffects in natural object categories," in *Proc. Vision Sciences Society Annual Meeting*, 2013.

連絡先

本吉 勇 (Isamu Motoyoshi) 人間情報研究部 感覚表現研究グループ

聴いて分かる腕の長さ

～身体の認識における聴覚の役割～

どんな研究

無意識に認識している身体の形状を**身体図式**と言います。私たちは身体図式を参照することで、物のある場所に手を伸ばしたり、狭いところを通ったりすることが出来ますが、その獲得の仕組みは未解明です。この展示では、聴覚が身体図式に影響を与えるという新発見を紹介しします。

どこが凄い

自分が床をたたき時に出る音を遠くから出たように感じさせると、自分の腕の長さが長く感じられることを世界で初めて発見しました。これは、聴覚が身体図式の認識に役立つことを示しています。身体図式獲得の仕組みの解明に不可欠な発見であり、極めて学術的価値の高いものです。

めざす未来

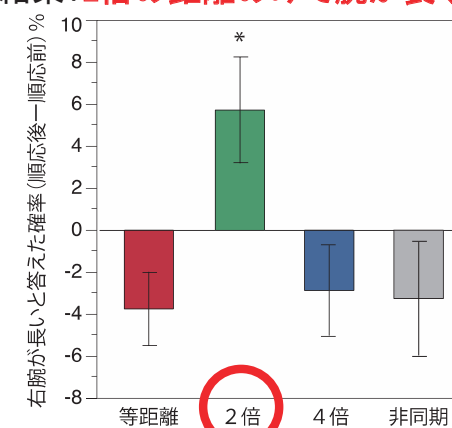
私たちが感じる「リアリティ」は、無意識の身体認識や身体感覚を伴っています。身体図式を獲得する仕組みの解明は、この**無意識の部分**を**正確に伝える方法の確立に寄与**します。これは、身体感覚まで伝える、未来の通信を実現するための基本原理の一つとなります。

- 順応：一方の腕で床をたたき、その音をスピーカから提示する。
→ しばらくたたき続けてその状態に慣らす。



- 腕の長さの認識を測定：左右の腕で、各2カ所を刺激し、長い方を答える。
床をたたいた腕を長いと感じているのか。
刺激の距離が長く感じる
＝ 腕が伸びたと感じている。

- 結果：2倍の距離のみで腕が長く認識されている！



発見の意味

1. 身体図式は聴覚で再構成される。(2倍の距離から音を出すと腕が伸びて感じる。)
2. 再構成には限界がある。(4倍の距離から音を出すと腕は伸びて感じない。)
3. 再構成には自分が動いて音を出している感覚が必要。(動きと音の出るタイミングを同期させないと腕が伸びて感じない。)

関連文献

- [1] A. Tajadura-Jiménez, A. Väljamäe, I. Toshima, T. Kimura, M. Tsakiris, N. Kitagawa, "Action sounds recalibrate perceived tactile distance," *Current Biology*, Vol. 22, No. 13, pp. R516-517, 2012.

連絡先

戸嶋 巖樹 (Iwaki Toshima) 人間情報研究部 感覚共鳴研究グループ

MEMO

MEMO

MEMO

NTT コミュニケーション科学基礎研究所
オープンハウス 2013 実行委員会

(委員長) 藤野 昭典

(副委員長) 木村 聡貴

(委員) 丸谷 和史／Marc Delcroix／杉山 弘晃