

新たな次元へとシフトする

～さらに深化するコミュニケーション科学の取り組み～

Shift to new dimensions

Deepening and evolving Communication Science

NTT コミュニケーション科学基礎研究所 所長

山田 武士 Takeshi Yamada

コミュニケーション科学基礎研究所(CS研)は、1991年、米
国独立記念日と同じ日の7月4日に、NTTの情報系の基礎研究
を担う研究所として、けいはんな学研都市内に設立されました。
NTTの他の研究所がみな関東に集中する中、小さな研究所が
ひとつ関西に飛び出し、いわば独立した瞬間でした。設立当初
から一貫して既存の研究領域にとらわれず、コミュニケーション
を科学するとはどういうことか?にこだわり、情報科学から人間
科学、社会科学、人文科学をも巻き込んだ学際的な取り組み
を進めてきました。特に音声音響処理、メディア処理、自然言
語処理、機械学習に代表されるいわゆるAI(人工知能)技術を
中心とした情報科学と人間の脳内における感覚運動処理や幼
児の言語獲得の仕組みを解明する人間科学の両方に取り組
むことで研究テーマ間の相乗効果を高めています。

深層学習に代表される、最近のAI技術の発展には目覚まし
いものがあります。研究所としては、これらの最先端の技術を
高度に使いこなし、直面する課題にうまく適用することも重要
です。求められてもいます。しかし情報系の基礎研究を担う
CS研としては、それ以上に、これまでの地の利やノウハウを活
かし更に技術を深め、いわば足場を固めつつ、これまでの延長
ではない、新たな次元を切り拓く取り組みに常に挑戦し、研究
テーマを大胆にシフトさせていくことが一層重要だと考えてい
ます。

例えば音声認識の研究では、これまでの接話マイクに向
かって一人が話す音声認識から、テーブルを囲んで複数人が
マイクから離れて自由に話す状況下での音声認識へとシフトし
ています。そのためには音声認識そのものの性能向上も必要
ですが、それ以上に騒がしい環境でも周囲の雑音を取り除く雑
音抑圧技術や、室内での壁や床などへの反射による残響を除
去する、残響除去技術等の音声強調技術との組み合わせが
重要となります。CS研の音声認識は、これらの組み合わせによ
り2015年に行われた国際技術評価(CHIME-3)で、参加25機
関中で第1位の精度を実現しました[1]。さらには、複数の発話
の中から聞きたい人の声を選んで聞き取る技術等にも取り組
んでいます。

機械学習技術に関して、CS研は、大量のデータから専門家
でも気づかない特徴的なパターンを自動的に見つけ出す技術を
研究開発してきました。現在はさらにこれを時空間に拡張した、
時空間多次元集合データ解析技術へとシフトさせています[2]。
さらに、単なる未来の予測だけでなく、予測の結果どうすべきか
(例えばどう人々の流れを誘導すべきか)の最適な方策をリア
ルタイムに提案してフィードバックする学習型シミュレーション
技術にも力を入れています。また、無数の組み合わせの中から
与えられた条件を満たす最適なものを見つけて、組合せ最
適化問題についても、BDDやZDDと呼ばれる、組み合わせ爆

発を回避するためのデータ構造を工夫することでこれまで不可
能だった大規模な問題が解けるようになっていきます。

ロボットとの対話処理の研究においては、これまでの一台の
ロボットとの対話から、複数台(2台)のロボットとの対話処理へと
シフトしています。一人の人間と対話する場合、ロボットは一台
で十分に思えますが、複数台のロボットを用いてうまく役割分
担することで、一台の場合より、たとえ音声認識に失敗したり、
対話の文脈が破綻したりしたとしても、人間にとって自然な対
話をより長く続けることができます。また、単なる雑談ではなく、
質問応答と雑談の融合により、対話への興味をユーザから引
き出しながらユーザに知識を伝える対話システムの実現にも取
り組んでいます[3]。

また、メディア処理においては、単一のメディアの処理から複
数のメディアを組み合わせたり、相補的に扱ったりするクロス
モーダル情景分析へとシフトしています。この技術が完成すれ
ば、例えばマイクで収録した音声情報から室内のビジュアルな
情景を再現することが可能になります。

人間科学研究にも変化があります。CS研では「視覚」「聴

覚」「運動感覚」といった人間の基本的な感覚に関する「潜在
的な脳の働き」の解明に注力しています。これらは今後も引き
続き取り組みますが、さらに新たなテーマとして、スポーツ脳科
学への取り組みが進んでいます[4]。脳科学の知見とICTを活
かし、「勝つための脳をきたえる」研究です。また、最近新たに、
人間の心の豊かさや満足に関する概念であるウェルビーイング
に関する研究が立ち上がりました。これらは触覚研究などと
あわせて、人間の「心と身体」の両面に焦点を当てた新たな取
り組みです。

これら、新たな取り組みにはリスクが伴います。すぐに期待通
りの学術的な成果が得られるとは限りません。AI研究の分野で
はExploration-Exploitation Dilemmaがしばしば問題となり
ます[5]。これは限られたリソースの中で、うまくいくとわかってい
る周辺をより詳細に探索するExploitationを行うべきか、目先
を変えてより広範囲に探索するExplorationを優先すべきか、
選択を迫られる、あるいは、バランスをとる必要がある、というも
のです。基礎研究を担うCS研としては京阪奈も厚木もその独
立精神を胸に刻み、未踏の地を目指して大草原に船出する
Explorationに、今後も大胆かつ粘り強く取り組んでいきたい
と思います。

関連文献

- [1] 公共エリア雑音下でのモバイル音声認識の国際技術評価で、世界1位の精度を達成:
<http://www.ntt.co.jp/news2015/1512/151214a.html>
- [2] 機械学習・データ科学センタ: <http://www.kecl.ntt.co.jp/mlc/index-j.html>
- [3] 世界初、雑談を交えながら知識を伝える新感覚対話AIを開発:<http://www.ntt.co.jp/news2018/1801/180131b.html>
- [4] スポーツ脳科学プロジェクト: <http://sports-brain.ilab.ntt.co.jp/>
- [5] 例えば: https://en.wikipedia.org/wiki/Reinforcement_learning