

メディア研究から人の活動を支援・代替するAI技術の研究開発へ

近年、デジタルトランスフォーメーション（DX）に大きく期待が寄せられており、昨今の新型コロナ禍によって取り組みが加速していくことが想定されます。また、AI（人工知能）の技術競争が激化しており、学習データ量においてはプラットフォームによる莫大なデータの獲得がなされています。このような中で、NTTメディアインテリジェンス研究所では培った技術やノウハウを強みとして、人の活動を支援・代替するAI技術の研究開発を推進しています。本特集ではその取り組みについて紹介します。

たなか	ひでのり	きたはら	まさき
田中	秀典	北原	正樹
くさち	よしのり		
草地	良規		

NTTメディアインテリジェンス研究所

はじめに

NTTメディアインテリジェンス研究所では、これまで音声・音響・言語・画像・映像等のメディアを処理する技術の研究開発に取り組み、さまざまな技術を実用化してきました。近年では、コンタクトセンタにおけるオペレータ支援⁽¹⁾やAIエージェントの実現^{(2),(3)}、緊急通報システムにおける集音⁽⁴⁾、4K・8K放送における映像圧縮装置⁽⁵⁾などの事業貢献を行っています。

しかし、昨今市場環境は大きく変化するとともに、競争も激化しています。各産業において、デジタル化によって既存の仕組みを変革するデジタルトランスフォーメーション（DX）が進みつつあり、昨今の新型コロナ禍によってさらに取り組みが加速していくことが想定されています。また、深層学習（ディープラーニング）の登場によって第三次AIブームが起これ、AI技術の基本アルゴ

リズムは誰でも活用できるようになっています。さらに、性能に寄与する学習データはGAFA（Google, Apple, Facebook, Amazon）を代表するプラットフォームによって大規模に収集されており、AIの性能が日々向上する世界が実現されています。これらの外部環境に対して、NTTグループでは、中期経営戦略においてSmart World実現に向けB2B2XやDXの取り組みを推進しています。さらには、革新的な技術によってスマートな世界を実現するIOWN（Innovative Optical and Wireless Network）構想を提唱し推進しています。

このような背景を踏まえ、NTTメディアインテリジェンス研究所では、これまでメディア処理における研究開発で培った技術やノウハウを活かして、価値の源泉となる人の活動を支援・代替するためのAI技術の研究開発に取り組むとともに、中長期的な新しい価値の創出をめざしたデジタルツインコンピュー

ティングの研究開発に取り組んでいます⁽⁶⁾。
本特集では、人の活動を支援・代替する領域
に向けたAI技術の研究開発について紹介し
ます。

人の活動を支援・代替するAI技術の 概要

人の活動を支援・代替するAI技術の適用
領域については、いくつかのシーンが考えら
れます。例えば、効率化として、これまで私
たちが取り組んできたコンタクトセンタにお
けるオペレータの生産性向上やAIエージェ
ントにとどまらず、オフィスの業務プロセス
改善や生産性向上、また新しい価値として生
活の質の向上などがあります。さらには、昨
今の新型コロナ禍によって、在宅勤務やオン
ライン会議等が浸透していくとともにその在
り方が変容していくことも考えられます。

こういったシーンにおいて、現行のAI技
術を適用するだけでは、実現が難しいこと
があります。例えば、より個人やその環境に即
した支援・代替を実現しようとする、個人
や環境にかかわるデータを取得する必要性が
ありますが、多量のデータが取得できない場
合があり、そのような条件下で性能を出すの
は容易ではありません。また、音声認識技術
1つをとっても、電話と会議では、音声をテ
キスト化するだけで事足りるのか、誰が話し

ているのかまで認識する必要があるのかなど
の違いが出てきます。オンライン会議とな
るとさらに求められる性能や要件の違いが出て
くる可能性もあります。

そこで、NTTメディアインテリジェンス
研究所では、少量データから効率的に学習を
行う技術、新しい効果を生み出す技術、既存
技術の性能にブレークスルーをもたらす技術
に着目して取り組みを始めています。

人の活動を支援・代替するAI技術の 取り組み状況

本特集記事では、現在取り組みを進めてい
る技術群について紹介します。『究極のプ
ライベート音空間を実現するメディア処理技
術』では、在宅勤務などでの応用が期待でき
る究極のプライベート空間を実現するために、
重要な要素の1つである音に着目し、周囲の
状況を音から理解する技術、聞きたい人にだ
け聞かせる技術、および聞きたくない音を消
す技術といった、新しい効果を生み出す技術
を確立することをめざしています。

『多様なユースケースに適用可能な音声合
成エンジン「Saxe」』では、バーチャルア
ナウンサーやAIエージェントの声を生成す
る音声合成技術に関して、文脈に応じて同形
異音語の高精度な読み分けを可能とする技
術、低コストで多様な話者性を再現する

DNN 音声合成技術といった、既存技術の性能にブレークスルーをもたらす技術および少量データから効率的に学習を行う技術について紹介します。

『コミュニケーションの知識源化を実現する音声認識技術』では、会議や対面接客における音声の認識を想定し、従来の音声をテキスト化する技術の向上に加えて、音声から話者の性別や感情を抽出するといった新しい効果を生み出す技術についても紹介します。

『顧客接点業務を支援・代替する知識・言語処理技術』では、適用シーンに応じて長さを指定して文書を要約する文書要約技術、インサイドセールスにおけるオペレータの生産性向上を実現する応対分析技術といった、既存技術の性能にブレークスルーをもたらす技術および新しい効果を生み出す技術について紹介します。

そして最後に、『4D デジタル基盤の実現に向けた空間情報処理技術』では、多様なセンシングデータをリアルタイムに統合しさまざまな未来予測を可能とする4D デジタル基盤⁽⁷⁾の実現に向けて、実空間を構造化する技術、時間変化を含む3D データを効率的に保存・活用する点群符号化技術といった、少量データから効率的に学習を行う技術および既存技術の性能にブレークスルーをもたらす技術について紹介します。

想定するユースケース

昨今の新型コロナ禍もかんがみ、人の活動を支援・代替するAI技術のユースケースを紹介します(図)。個人空間の創出では、在宅勤務において疑似的に個人空間を構築し、プライバシーの流出がない空間を個人宅内につくり上げます。オンライン会議では、進行をテキスト化・要約・翻訳し、従来の人の働き方について時間・空間の制約を緩和することでイノベティブな共同作業を支援します。また、アナウンサー等、従来は人にしかできなかった業務をAIが代替することで、人どうしの接触を不要とした速やかな業務を実現します。物流改革の観点では、不足しているモノ・場所をSNS等から特定し、都市の3次元構造物を認識して自動で届ける(必要な物を必要な人に迅速に届ける)ことが可能になると想定しています。

今後の展望

昨今取り巻く環境は目まぐるしく変化しています。人の活動を支援・代替するには、技術もこうした変化に対応していく必要があります。マクロとミクロの変化の両面をとらえつつ柔軟に研究開発を推進していきたいと考えています。

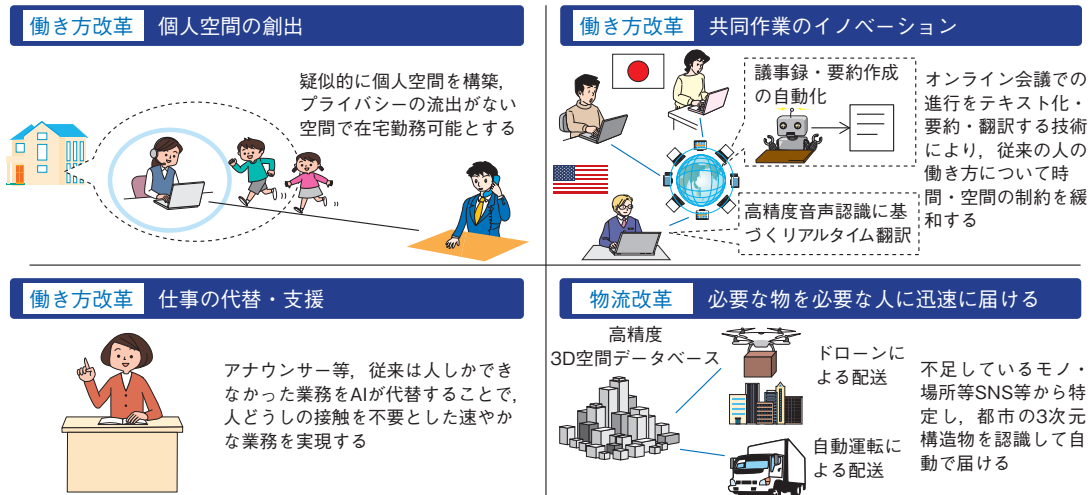


図 人の活動を支援・代替するAI技術のユースケース

■参考文献

- (1) https://www.ntt-tx.co.jp/products/foresight_vm/
- (2) <https://www.ntt.com/business/services/application.html#ai>
- (3) <https://www.nttdocomo.co.jp/service/mydaiz/>
- (4) <https://www.ntt.co.jp/news2018/1802/180219c.html>
- (5) <https://www.ntt.co.jp/news2016/1602/160215b.html>
- (6) <https://www.ntt.co.jp/svlab/DTC/whitepaper.html>
- (7) <https://www.ntt.co.jp/news2020/2003/200326c.html>



(左から) 田中 秀典 / 北原 正樹 / 草地 良規

新技術は既存の手段よりも扱いづらいことがありますが、うまく使いこなした企業は競争優位性を獲得することが可能です。ぜひNTTメディアインテリジェンス研究所の技術をご活用ください。

◆問い合わせ先

NTTメディアインテリジェンス研究所
 企画部
 TEL 046-859-2497
 FAX 046-855-1149
 E-mail hidenori.tanaka.ba@hco.ntt.co.jp