

脳波とAIを用いた音声合成と データ拡張手法

東京工業大学 科学技術創成研究院
准教授 吉村 奈津江

2021年11月2日

従来技術とその問題点(音声合成)

欧米を中心に実用化されつつあるものには、脳内に電極を埋め込む方式による脳内情報の解読法があるが、

- ・外科的手術が必要
- ・一般人に適用可能な認可が下りていない

等の問題があり、広く利用されるまでには至っていない。

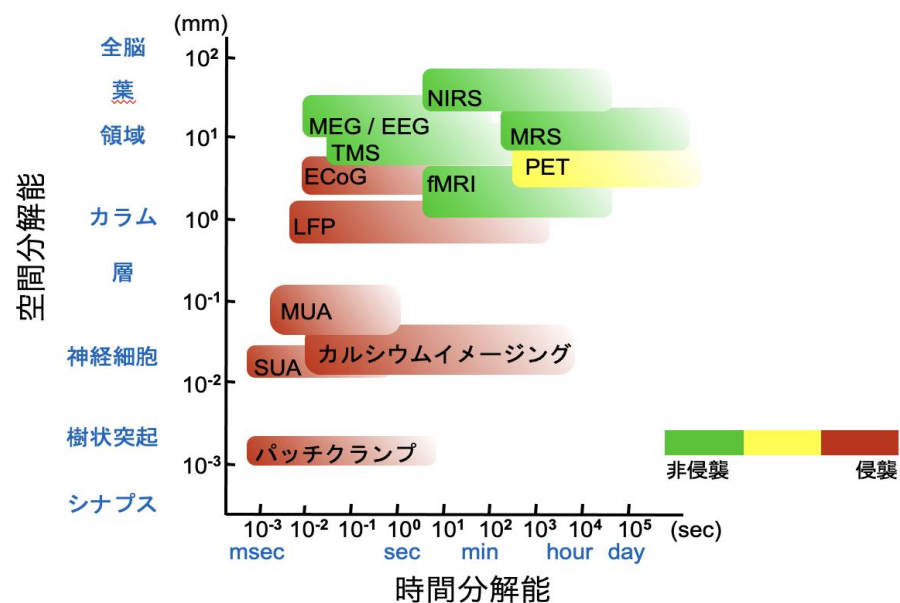
Brain Chip by Neuralink Could Improve Communication and Create Universal Language



(Photo : Neuralink)

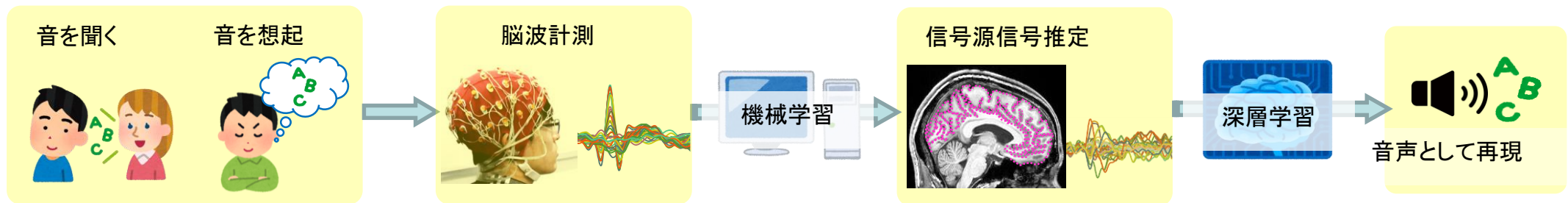
一方、外科的手術が不要な脳波の場合は、

- ・詳細な情報を解読できるほどの抽出精度がない



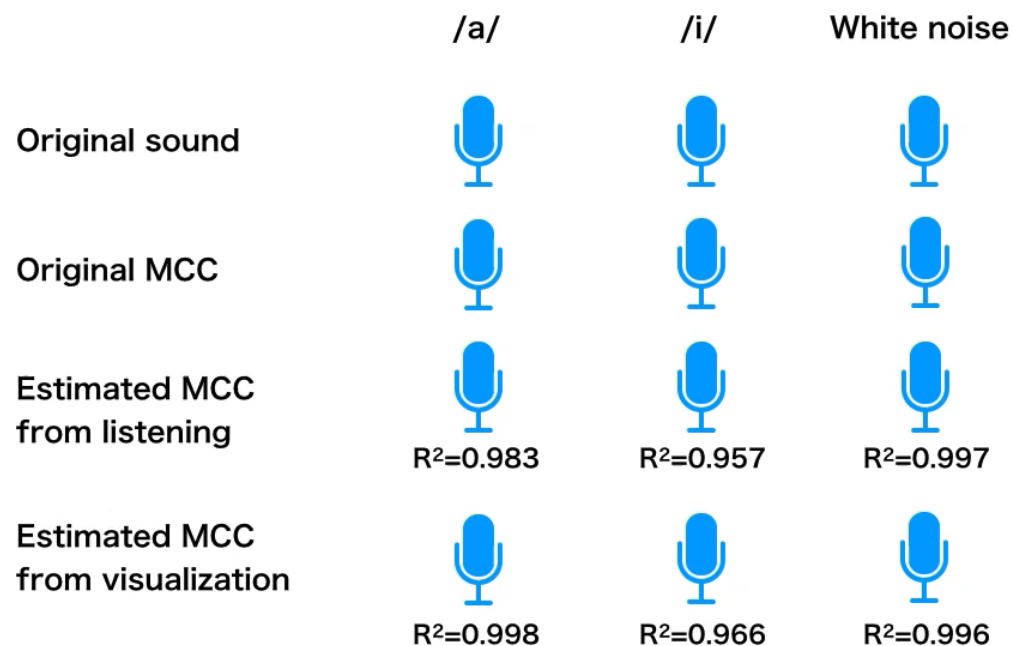
新技術の特徴・従来技術との比較(音声合成)

- 脳波による情報抽出の問題点であった、空間分解能を計算处理的に改善し、聞いた音、思い出した音を脳波から再現することに成功した。
- 従来は何の音であるかを「判別」する手法が主流であったが、音声波形を再現し、明瞭な音声になる可能性が確認できたため、個人ごとに異なる「頭の中の音声」を再現できる可能性がある。
- さらに、脳のどこで音声を聞いたり思い出したりしていたかを調べられるため、聴覚から認知に渡る脳領域の機能を調べられることが期待される。



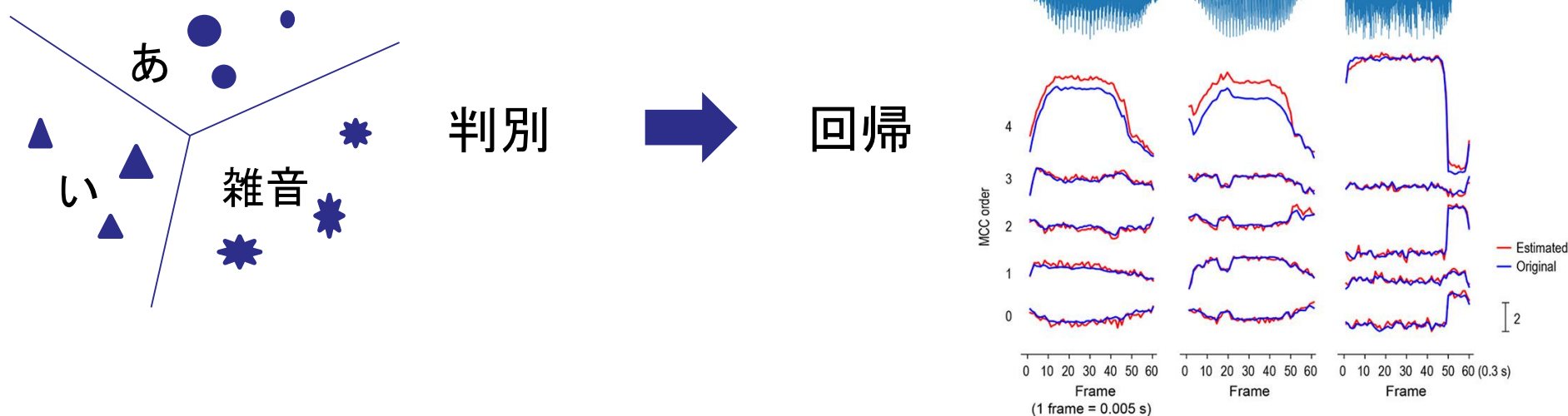
新技術の特徴・従来技術との比較(音声合成)

- 脳波による情報抽出の問題点であった、空間分解能を計算处理的に改善し、聞いた音、思い出した音を脳波から再現することに成功した。
- 従来は何の音であるかを「判別」する手法が主流であったが、音声波形を再現し、明瞭な音声になる可能性が確認できたため、個人ごとに異なる「頭の中の音声」を再現できる可能性がある。
- さらに、脳のどこで音声を聞いたり思い出したりしていたかを調べられるため、聴覚から認知に渡る脳領域の機能を調べられることが期待される。



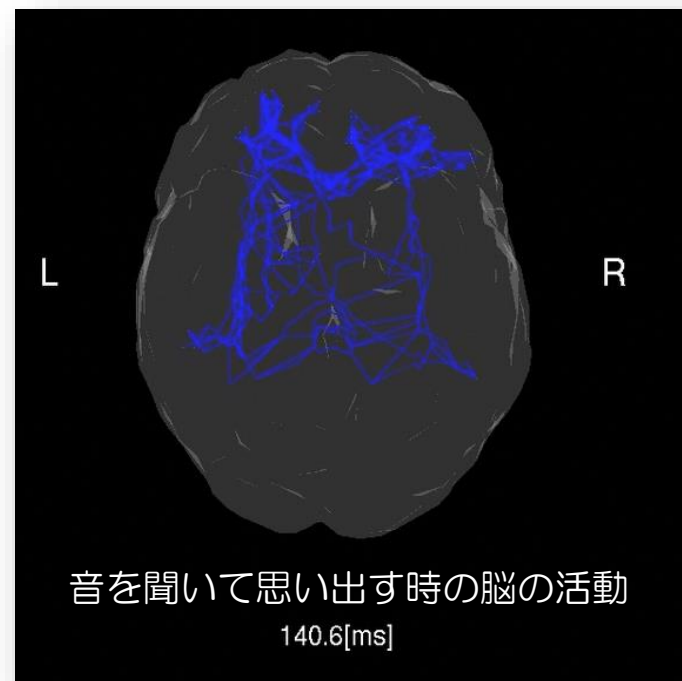
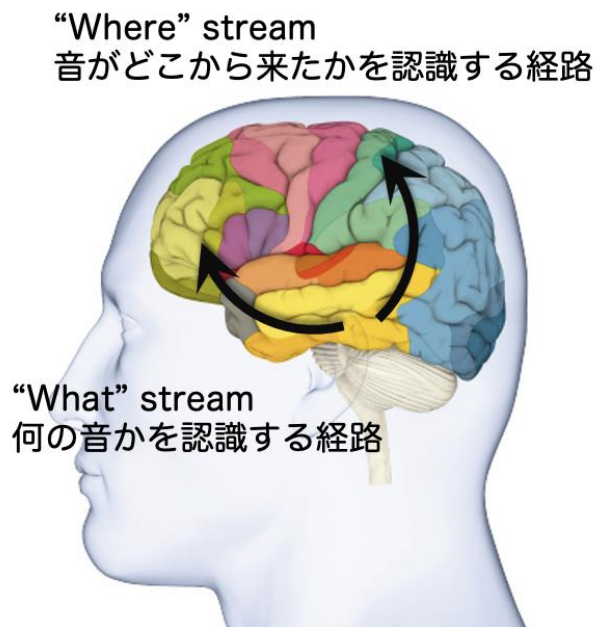
新技術の特徴・従来技術との比較(音声合成)

- 脳波による情報抽出の問題点であった、空間分解能を計算处理的に改善し、聞いた音、思い出した音を脳波から再現することに成功した。
- 従来は何の音であるかを「判別」する手法が主流であったが、**音声波形を再現し、明瞭な音声になる可能性が確認できたため、個人ごとに異なる「頭の中の音声」を再現できる可能性がある。**
- さらに、脳のどこで音声を聞いたり思い出したりしていたかを調べられるため、聴覚から認知に渡る脳領域の機能を調べられることが期待される。



新技術の特徴・従来技術との比較(音声合成)

- 脳波による情報抽出の問題点であった、空間分解能を計算处理的に改善し、聞いた音、思い出した音を脳波から再現することに成功した。
- 従来は何の音であるかを「判別」する手法が主流であったが、音声波形を再現し、明瞭な音声になる可能性が確認できたため、個人ごとに異なる「頭の中の音声」を再現できる可能性がある。
- さらに、脳のどこで音声を聞いたり思い出したりしていたかを調べられるため、聴覚から認知に渡る脳領域の機能を調べられることが期待される。



従来技術とその問題点(データ拡張)

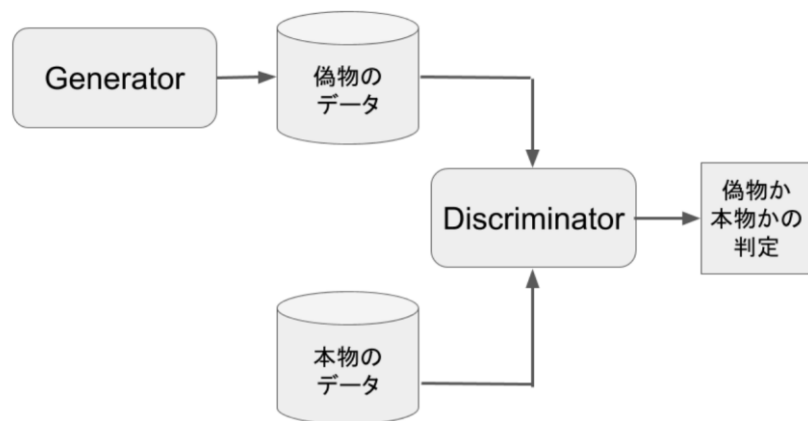
深層学習によりデータ推定精度の向上が実証されているが、

- ・生体信号など計測できるデータ数に限界がある(ex. < 100)場合には
深層学習のメリットが活かせず精度が上がらない

等の問題があり、実社会で広く利用されるまでには至っていない。

その問題を解消するために既存のデータを拡張してデータを嵩増しする手法が提案されているが、

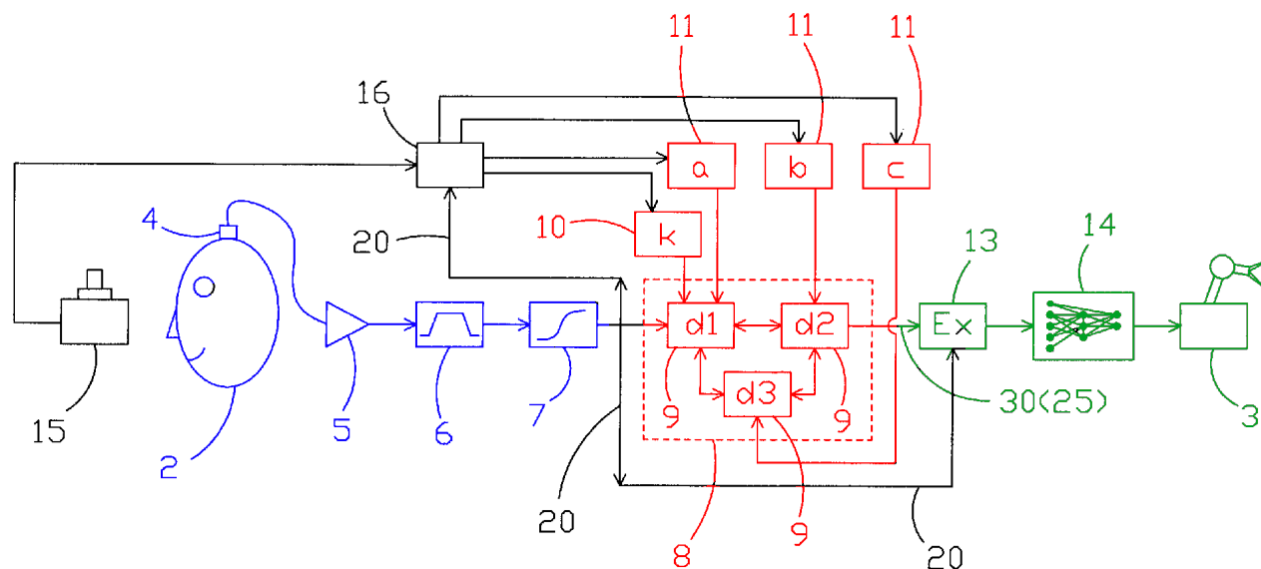
- ・拡張データの作成に時間を要するためリアルタイム処理に適した手法がない



データ拡張の代表例
GAN(敵対的生成ネットワーク)
→ データラベルは不要だが、学習には時間を要する

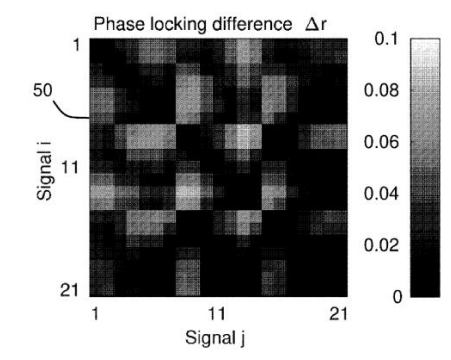
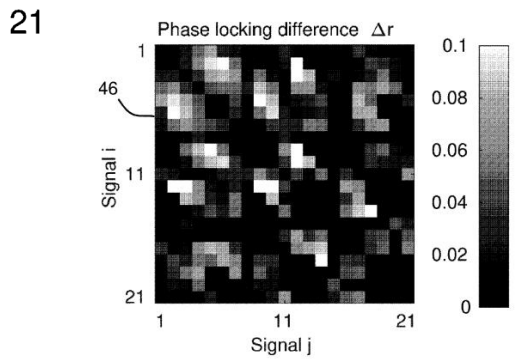
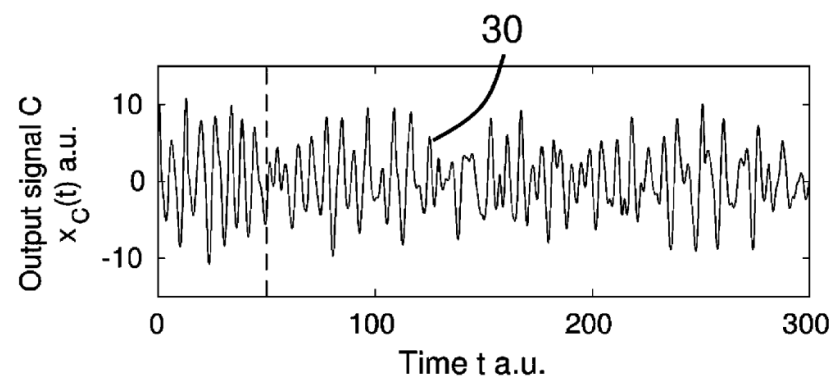
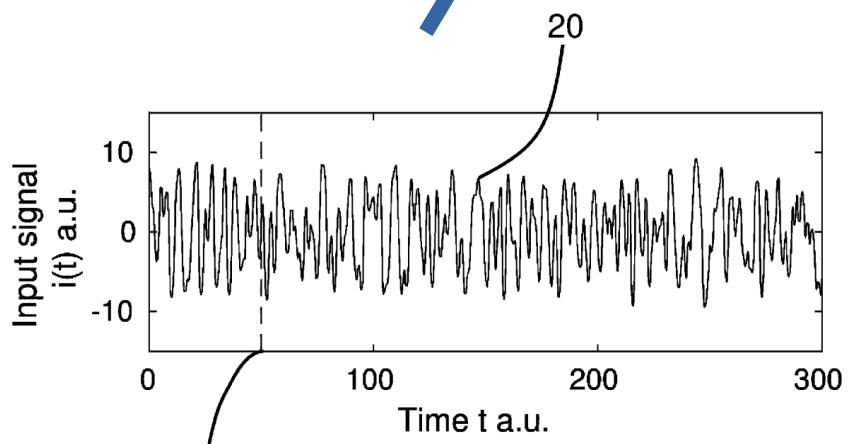
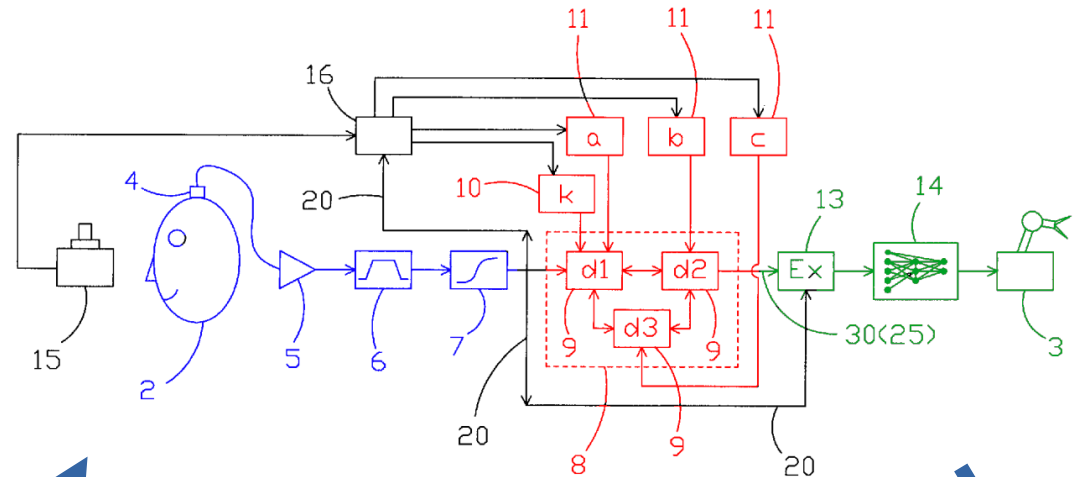
新技術の特徴・従来技術との比較(データ拡張)

- 深層学習に用いるデータ数不足を補うために、敵対的生成ネットワーク(GAN)に代表されるデータ拡張技術が利用されるが、**全く別のコンセプトによるデータ拡張技術**を提案した。
- 従来手法では既存データの学習によりデータを拡張するためある程度のデータが必要であり、学習用の時間が必要であるが、**新技術では1つの元データの揺らぎを生成するため、リアルタイムにデータが生成可能**となった。
- GANなどの手法は特徴を鮮明化したデータ生成を目標とするため画像に適しているが、**揺らぎデータの再現をコンセプトにした新技術は、脳波のように複数の情報が混在した信号に適している**ことが期待される。



非線形システムに脳波を通すことで信号が生成される

新技術の特徴・従来技術との比較(データ拡張)



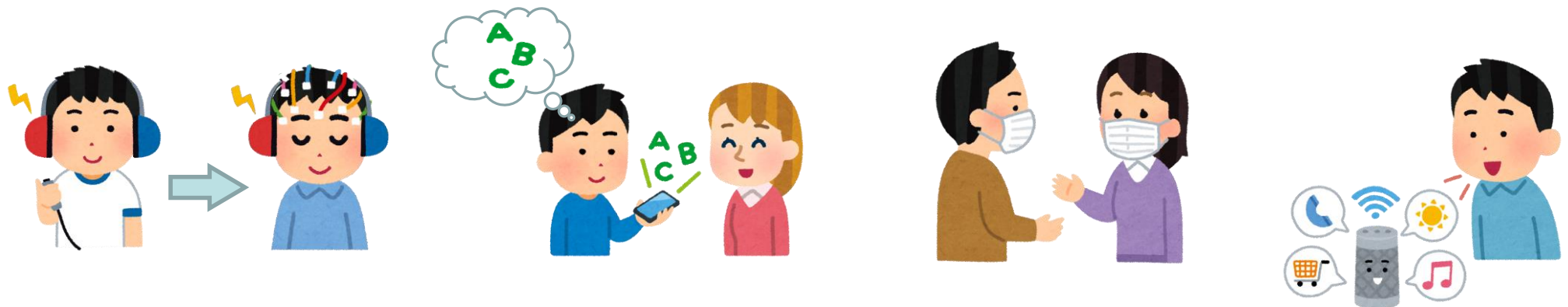
想定される用途

音声合成

- 聴覚検査（脳のどこで聞いているかが客観的にわかるため、万人に適用可能）
- 言語機能検査（言語能力と脳機能領域の関係を調べるなど）
- 意思伝達システム（発話せずに意思伝達が可能なシステムなど）

データ拡張

- 脳波を用いた情報抽出全般（音声情報に限定されない）
- その他の時系列信号を用いた深層学習



実用化に向けた課題

音声合成

- 母音の音声合成が可能なところまで開発済み。現在、子音やフレーズの合成を実施中。
- 用途によっては、リアルタイムに高精度な音声合成を可能にするアルゴリズムの開発が必要。

データ拡張

- 実際の計測信号による効果は現在検討中。
- 扱う信号や抽出する情報によって、データ生成に用いる最適な数理システムやパラメータの設定調整が必要。

企業への期待

- 企業に保有の技術との融合による新製品開発。
- 脳波計測やブレインマシンインタフェースのためのシステムやハードウェア開発を行う企業との共同研究。
- クラウドシステムとIoT技術を扱う企業において、ブレインマシンインタフェース分野への展開に有効となる可能性あり。

本技術に関する知的財産権

- 発明の名称 : 情報処理装置および情報処理方法
- 出願番号 : 特願2020-092110,
PCT/JP2021/17180
- 出願人 : 東京工業大学
- 発明者 : 吉村奈津江、小池康晴

本技術に関する知的財産権

- **発明の名称** : Brain-computer interface system suitable for synchronizing one or more nonlinear dynamical systems with the brain activity of a person
- **出願番号** : PCT/JP2019/030295
- **出願人** : 東京工業大学
- **発明者** : Minati Ludovico, 吉村奈津江、小池康晴

お問い合わせ先

東京工業大学
研究・産学連携本部

TEL 03-5734-2445

FAX 03-5734-2482

e-mail sangaku@sangaku.titech.ac.jp