



Tokyo Tech



2021年1月8日

東京工業大学
科学技術振興機構

「聞こえた音、思い出した音」を脳波から 音で再現する技術を開発

— 脳内の音声処理機構の理解に向けて —

【要点】

- 脳波信号から音声を直接再構築する手法を開発
- 音を聞いている時、思い出した時に活動する脳領域の違いを示唆
- 脳内聴覚や音声、および言語処理の客観的評価に脳波を利用できる可能性

【概要】

東京工業大学 科学技術創成研究院の吉村奈津江准教授（科学技術振興機構さきがけ研究者兼務）、明石航大学院生（研究当時）、神原裕行助教、緒方洋輔特任助教（研究当時）、小池康晴教授、ルドビコ・ミナチ特定准教授は、頭皮で記録された脳波信号（**EEG**、用語 1）から音声を直接再構築するために有望な手法を開発した。

参加者が 2 つの母音「ア」と「イ」を視聴後に思い出したときに記録された **EEG** を用いて、聞かせた音源のパラメータを**畳み込みニューラルネットワーク（CNN**、用語 2）によって推定した。推定されたパラメータを用いて復元した母音の音声は非常に明瞭で、実際に視聴した者とは別の参加者が音声の弁別を行ったところ、**85%**を超える認識率を示す音声だった。さらに、音源パラメータの推定のために **CNN** が抽出した脳波の特徴は、何の音かを特定するために使われる脳内の聴覚経路（**What** ストリーム、用語 3）に含まれる領域であり、この手法の脳科学的妥当性も示された。

この抽出された特徴を調べることで、音声を聞いている時とそれを想起している時の脳活動領域の違いがさらに示唆された。この手法はその人がどのように聞こえているか、聞こえていないのか、を客観的に把握し、さらに脳のどこを使っているのかを調べられる可能性があるため、**EEG** を利用した脳内の聴覚・音声・言語処理の評価が可能になる効果が期待される。

本研究は、ドイツ科学技術誌「*Advanced Intelligent Systems*（アドバンスド・インテリジェント・システムズ）」に1月7日（現地、ドイツ時間）に掲載された。

●研究成果

32カ所の脳波信号 (EEG) 電極から、母音の「ア」、「イ」、および白色雑音を聞いているときと、その後、聞いた音声を想起しているときの EEG を記録し、聞かせた音源のパラメータ信号を EEG から畳み込みニューラルネットワーク (CNN) を用いて推定して音源を復元した。復元された音声が耳で聞き分けられるかどうか聞き取り判別試験を行ったところ、全ての参加者の脳波データにおいておおよそ8割程度の判別が可能だった。

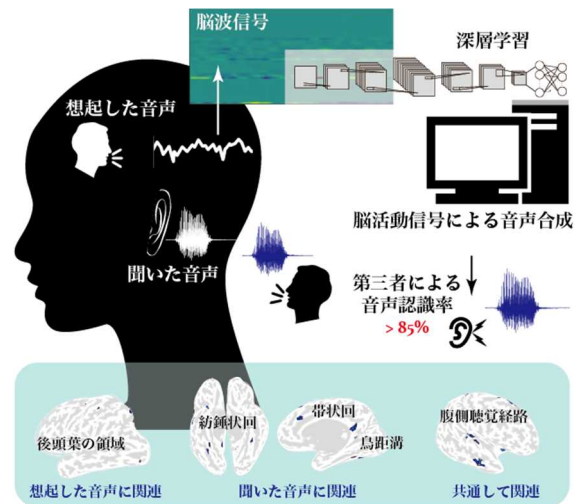


図1 研究の概念

このような高い精度で CNN が音声情報を抽出できたということは、CNN が音源推定に利用した脳の領域と信号の時間的なタイミングは、脳内の音声処理過程を間接的に反映していると考えられる。その領域を調べた結果、脳内の聴覚処理において、何の音かを検知するための信号が処理される「What ストリーム」と呼ばれる脳領域群が主に使われていることがわかった。

これはコンピュータが抽出した脳内の特徴が脳科学的にも妥当であったことを示唆している。さらに興味深いことに、音を聞いている時と音を思い出した時で CNN が抽出した脳領域に違いがあり、個人ごとの脳領域の違いもみられた。2種類の母音の違いだけを調べた研究だが、この技術をさらに進歩させることで、個人の脳内の聴覚・音声・言語処理のさらなる理解に貢献できるものと期待される。

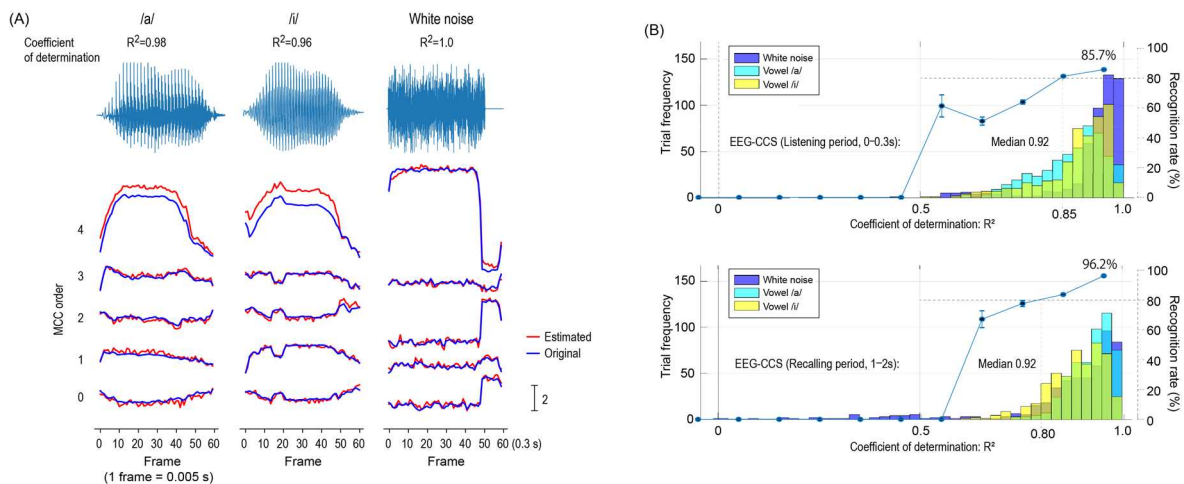


図2 EEG から推定した音源パラメータ波形と復元した音声の聞き取り精度を示した結果

(A) 元の音源パラメータ波形 (青) と EEG から推定した波形 (赤) の比較 ($R^2=1$ で、完全一致)

(B) 音を聞いた時の EEG (上) と、音を思い出した時の EEG (下)、でどの参加者も高い R^2 分布を示した。「ア」と「イ」の違いは R^2 が 0.8 を超えると 80% 以上の確率で聞き分けられた。

●研究の背景

コンピュータ処理技術の飛躍的な革新に伴い、脳活動信号から脳内の情報を読み出すブレイン・コンピュータ・インタフェース（用語 4）に関する研究が近年、盛んに進められている。この技術を用いて、脳内の情報処理過程を知るために研究者が何十年にもわたって注目してきた課題が次第に明らかになってきた。

音声情報の読み出しに関しては、電極を頭蓋骨下の脳皮質表面に埋め込む手術を伴う皮質脳波信号（ECoG、用語 5）を使用して聴覚処理に関連する脳領域から直接信号を取得し、それによって音声合成が試みられている。しかし深層学習（用語 6）など、機械学習の最近の進歩にもかかわらず、ECoG を使用しても、聞き取りやすい音声の合成は依然として困難な状況である。

●研究の経緯

吉村准教授らの研究では、EEG から脳内の神経活動を機械学習により推定し、筋活動、指の動きなど、これまで EEG からでは困難だと考えられてきた情報を抽出することに成功してきた。今回の研究では、音声認識や音声合成に用いられているメルケプストラム（用語 7）という音声情報を表現しているパラメータを CNN モデルで推定し、物理的に提示または想起された母音を EEG から合成することができた。

●今後の展開

耳で聞き分けられる聴取性能の高い音声を EEG から再構成できたため、本人がどのように聞こえているかを第三者に伝えることができる可能性があり、聴覚検査の客観的な手法として使える可能性がある。また、脳のどの領域が聴覚・音声・言語処理に関係しているのかについての理解を深め、ブレイン・コンピュータ・インターフェイスなどのさまざまな将来のアプリケーションへの道を開く可能性がある。

【用語説明】

- (1) **脳波信号 (EEG)** : 脳内の神経細胞の活動を反映した電気信号を頭皮に付着した電極から記録した信号。
- (2) **畳み込みニューラルネットワーク (CNN)** : 広義の人工知能、機械学習法のひとつで、脳内の神経細胞がネットワークを形成して情報伝達を行う事象を模して数式的なモデルとして表現し、入力された情報の中から目的とする出力を構成するために必要な情報を抽出するために利用できる。
- (3) **What ストリーム** : 脳内の聴覚経路のひとつであり、何の音であるかを特定するための役割があるとされている。耳付近の側頭部から前頭部にかけて位置する領域が含まれると考えられている。音の場所を特定する経路は **Where ストリーム** と呼ばれる。

- (4) **ブレイン・コンピュータ・インタフェース**：脳から記録した信号をコンピュータ処理し、コンピュータやロボットを操作するための指令信号として出力することで、頭で考えるだけでコンピュータやロボットを操作可能にする技術。
- (5) **皮質脳波信号 (ECoG)**：頭蓋骨の下にある脳皮質の表面に外科的手術により設置した電極（シート）から、脳内の神経細胞の活動を反映した電気信号を記録した信号。
- (6) **深層学習**：CNN のネットワーク構造は層を形成しており、これまでは 3 層が一般的であったが、近年のコンピュータ処理速度の大幅な向上により、より多くの層（深層）の構造を用いた計算処理が可能となった。
- (7) **メルケプストラム**：特に音声認識でよく用いられる技術で、音源の内容を表す特徴を時系列信号として抽出したもの。

本研究は、科学技術振興機構（JST）戦略的創造研究推進事業さきがけ「人とインタラクションの未来」（研究総括：暦本純一 東京大学大学院情報学環教授／株式会社ソニーコンピュータサイエンス研究所副所長）研究領域における「脳波を用いたセルフケアサポートシステム」（研究者：吉村奈津江）（JPMJPR17JA）、日本学術振興会科学研究費助成事業（基盤研究（C）15K01849）の支援を受けて実施された。

【論文情報】

掲載誌：*Advanced Intelligent Systems*

論文タイトル：Vowel Sound Synthesis from Electroencephalography during Listening and Recalling

著者：Wataru Akashi, Hiroyuki Kambara, Yousuke Ogata, Yasuharu Koike, Ludovico Minati, Natsue Yoshimura*

DOI：10.1002/aisy.202000164

【問い合わせ先】

東京工業大学 科学技術創成研究院 准教授

吉村奈津江

Email: yoshimura.n.ac[at]m.titech.ac.jp

TEL: 045-924-5066/5086 FAX: 045-924-5066

【取材申し込み先】

東京工業大学 総務部 広報課

Email: media[at]jim.titech.ac.jp

TEL: 03-5734-2975 FAX: 03-5734-3661

科学技術振興機構 広報課

Email : jstkoho[at]jst.go.jp

TEL : 03-5214-8404 FAX : 03-5214-8432

<JST 事業に関すること>

舘澤 博子

科学技術振興機構 戦略研究推進部 ICT グループ

Email : presto[at]jst.go.jp

TEL : 03-3512-3526 FAX : 03-3222-2066