

( 続紙 1 )

京都大学	博士 (情報学)	氏名	PARK, Jong-Yun
論文題目	Sound Reconstruction from Human Brain Activity (ヒトの脳活動からの音の再構成)		

(論文内容の要旨)

ヒトの脳活動を用いた知覚内容の再構成は、意識経験の神経表現に関する洞察を提供してきた。視覚については、比較的詳細な知覚内容を任意の画像から再構成する方法が実現しているが、聴覚については、音声の時系列の複雑さやニューロイメージング手法の限定的な解像度のため、任意の音声を再構成する方法は確立していなかった。Park氏は本博士論文で、脳の聴覚処理の階層的情報表現と音声生成モデルを利用することで、任意の音声を脳活動から再構成する方法を示した。

Park氏は音声再構成の課題を克服するために、脳の聴覚処理をディープニューラルネットワーク (DNN) を用いて理解する近年のアプローチに注目した。これまでの研究で、人間の聴覚システムとDNNモデルの間には、階層的な相同性があることが示されている。さらに、音声生成モデルの進歩により、圧縮された表現を高解像度の音に戻すことが可能になった。本論文では、聴覚特徴の脳デコーディングと音声生成モデルを組み合わせた音声再構成法を提案した。自然音に対するfMRI応答を用いて、DNNモデルの階層的音声特徴を高い精度でデコードできることを確認し、デコードされたDNN特徴の圧縮された時間情報を音声生成モデルを用いて分離することで音声信号の再構成を実現した。

本手法により、訓練データセットに含まれないカテゴリを含む任意の音刺激を再構成することが可能であることが示され、知覚内容や音質を反映した、制約のない音声再構成が初めて実現した。異なる聴覚領域からの再構成は実際の音に類似しており、聴覚表現の分散性が示された。再構成が実際の主観的知覚経験を反映しているかどうかを確認するため、重なり合う音の1つに聴覚的な注意を選択的に向ける実験も実施した。その結果、再構成された音が注意を向けた音に類似する傾向が見られた。これらの結果は、提案モデルが、主観的聴覚経験を脳活動から外在化することができることを示している。

本論文は以下のように構成されている。第1章では、人間の聴覚処理に関する包括的な紹介と、視覚と聴覚の両方における知覚経験の再構成に関する現状のレビューを行っている。第2章では、音声信号と脳イメージング信号のギャップを埋めるために設計されたモデルと情報表現について説明している。第3章では、自然音に対するfMRI反応を測定するために、本論文で使用した実験プロトコルを詳細に説明している。第4章では、さまざまな音声特徴を用いたデコーディング分析の結果を示している。第5章では、音声再構成モデルを適用し、実際の音刺激に類似した内容と音質を保った複雑なスペクトル-時間パターンを再構成できることを実証している。第6章では、提案モデルの汎化性を調べるために、トレーニングデータセットから一部のカテゴリを除外するアブレーション解析を行っている。第7章では、音声再構成における聴覚野とDNN表現の階層の役割を分析している。第8章では、「カクテルパーティー条件」を用いて、トップダウン的注意により変容した主観的知覚内容を再構成できることを示している。最後に第9章では、提案モデルの各要素の役割について議論し、聴覚知覚経験の再構成の意義と応用の可能性について展望を示している。

(続紙 2)

(論文審査の結果の要旨)

Park氏の論文に関する審査結果について、以下の通り報告する。

再構成法の進展とその重要性について：

これまでの聴覚分野における研究では、限られた刺激に対する再構成しか実現しておらず、その精度も低い場合が多かった。しかし、本研究においては、不完全であるものの認識可能なレベルの再構成が達成され、任意の刺激に対しても同じモデルが適用できることが示されている。これはこれまでに実現されていなかった新たな技術の創出であり、神経科学のみならず、ブレイン・マシン・インターフェースなど未来の情報通信や医療にも大きなインパクトをもたらす成果である。

fMRI信号の処理とDNNの有用性について：

fMRI信号の時間解像度の低さを克服する方法として、大脳聴覚野と相同性を持つ深層ニューラルネットワーク (DNN) 情報表現と音声生成モデルを組み合わせる方法が提案され、その有効性が示されている。これによって、再構成の信頼性が向上し、複雑な音の再構成が可能となった。これらの手法の特徴や導入理由についても明確に議論されている。

再構成の実証と検証の厳密性について：

本研究では、特定のカテゴリの刺激を訓練データから取り除いても、そのカテゴリの刺激を再構成できることが実証されている。その他、脳信号とモデルに対する詳細な検証が行われており、従来の研究とは異なるレベルの高い信頼性をもつ研究と言える。

実験結果と応用における可能性について：

堅固な実験計画と統計的評価により、実験結果の信頼性が確保されている。聴覚的注意の実験結果については、個人差が大きく、今後、個人差を加味した検討が期待される。応用面については、実世界でのデコーディングや再構成の可能性に関する議論が行われ、広範な応用の見通しが示されている。

これまで実現が困難と思われていたfMRI信号からの任意音声の再構成法を実現し、学術的に信頼性の高い方法で検証した点で、この博士論文は高く評価でき、博士(情報学)の学位取得に相応しいものと評価される。

令和5年8月24日に実施された論文内容と関連する口頭試問においても合格と認定された。また、本論文のインターネットでの全文公表についても支障がないことを確認した。