

2026年6月19日

長岡技術科学大学

「声を出す仕組み」を学習した AI による音声合成手法を開発

－ニューラルネットワークにより音声の生成過程を再現－

ポイント

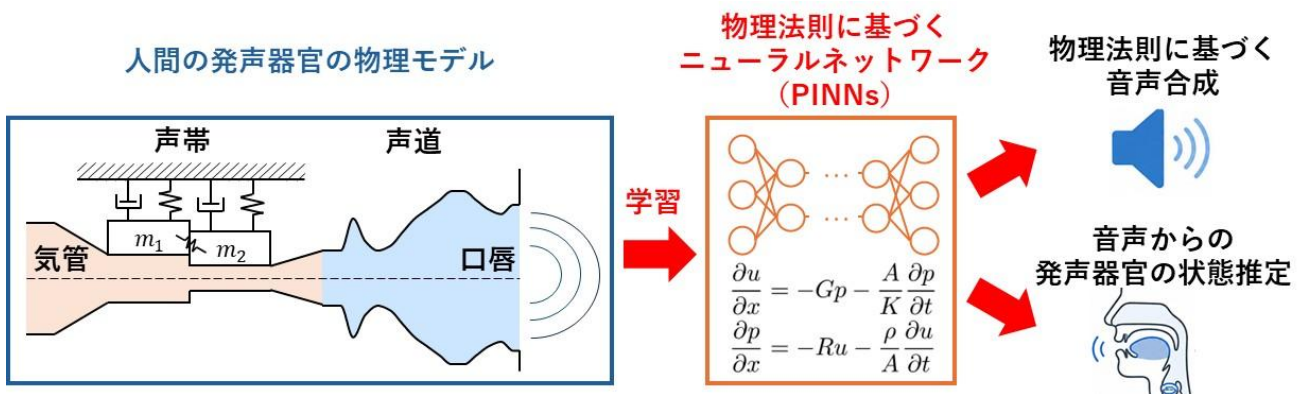
- 声帯と声道による「声を出す仕組み」をニューラルネットワークに学習させ、人の声の母音を合成することに成功
- 人の発声メカニズムの解明につながる新たな解析手法として期待
- 音声障害の診断・解析や発声訓練の支援といった医療応用への展開が見込まれる

概要

長岡技術科学大学 技学研究院 機械系の横田和哉助教、電気電子情報系の原川良介准教授、機械系の馬場将亮准教授、電気電子情報系の岩橋政宏教授の研究グループは、Physics-Informed Neural Networks (PINNs)^{*1} という物理法則に基づくニューラルネットワークを用いて、人の声の生成される過程を解析・再現する新たな音声合成手法を開発しました。

人の声は、声帯の振動によって生じた空気の流れが、声道で共鳴することによって作られます。本研究では、この「声帯の振動」と「声道での共鳴」という異なる物理現象を PINNs の中で結びつけて解析する手法を提案しました。本研究は声帯と声道の相互作用を物理モデルとして結びつけ、PINNs により音声合成を実現した初めての報告です。

さらに本手法により、音声データから声帯の動きや声門流などの発声器官の状態を推定できる可能性も示されました。本成果は、私たちが聞いている声がどのように生成されたのかを、データと物理法則の両面から理解するための新しい解析技術として、音声科学の発展や音声障害の解析などへの応用が期待されます。



詳細

【研究の背景】

音声は、声帯の振動、空気の流れ、声道での共鳴が組み合わさって生じる身近でありながら複雑な物理現象です。これらを数式で表し、計算機上で再現することは、音声科学や発声メカニズムの理解に役立ちます。さらに、音声波形から声帯の動きなどを推定できれば、発声状態の評価や音声障害の解析への応用も期待されます。

近年、物理法則に基づくニューラルネットワーク（Physics-Informed Neural Networks, PINNs）と呼ばれる数値シミュレーション手法が注目されています。これは物理法則を表す偏微分方程式をニューラルネットワークに学習させる手法であり、「声を出す仕組み」、すなわち声帯と声道の物理法則を学習したPINNsを構築することで、音声に関する様々な解析が可能になると考えられました。

【研究の内容と成果】

本研究では、PINNsを用いて声帯の振動と声道での共鳴を結びつける、新しい音声合成手法を開発しました。声帯が閉じる際の複雑な動きを学習させるための仕組みや、ニューラルネットワークの学習により肺の圧力から声の高さを推定するための手法、声帯と声道の相互作用を扱うためのネットワークの構造などを導入しました。

妥当性の確認として、順解析および逆解析を行いました。順解析では、母音「あ」および「う」の音声波形を合成することに成功し、その妥当性は従来の数値シミュレーション手法との比較により確認されました。声帯と声道の物理法則を学習したPINNsによる音声合成は初めての報告です。逆解析では、声道の形状や声帯のパラメータが既知という条件下で、シミュレーションにより合成された音声波形から声帯の動きや声門流、声門下圧などを推定できることを示しました。

【今後の展開】

本成果は、人間の音声がどのように生成されたのかを、観測データと物理法則を結びつけて理解するための新しい音声合成手法を提案するものです。将来的には、音声から声帯や声道など発声器官の状態を推定することで、発声メカニズムの解明や音声障害の解析、発声訓練支援などの分野における新たな音声合成・解析手法としての発展が期待されます。

【研究成果の公表】

本研究成果は、IEEE（米国電気電子学会）が発行する科学誌 IEEE Transactions on Audio, Speech and Language Processing に掲載されました。

論文タイトル：Physics-Informed Neural Networks for Speech Production

著者名：Kazuya Yokota, Ryosuke Harakawa, Masaaki Baba, Masahiro Iwahashi

掲載誌：IEEE Transactions on Audio, Speech and Language Processing

掲載日：2026年6月2日

DOI：https://doi.org/10.1109/TASLPRO.2026.3700036

【助成・謝辞】

本研究は JSPS 科研費 JP25K03137、JSPS 地域中核・特色ある研究大学強化促進事業 JPJS00420240017、および小野音響学研究助成基金の支援により実施されました。

【用語解説】

※1 **Physics-Informed Neural Networks (PINNs)** : 物理法則を表す偏微分方程式を学習したニューラルネットワーク。教師データに基づき学習を行うニューラルネットワークとは対照的に、PINNs は「ネットワークの出力値が偏微分方程式を満たしているか」を基準に学習を行う。