



2024年10月11日

報道関係者各位

慶應義塾大学
公立はこだて未来大学
筑波大学

イヤホンにタッチせず空中ジェスチャで操作する新手法を開発

-イヤホンに装備されているスピーカと

收音マイクのみで実現する音を使った新技術-

慶應義塾大学 理工学部の杉浦裕太准教授、雨坂宇宙訪問研究員、大学院理工学研究科の鈴木俊汰（修士課程1年）、公立はこだて未来大学システム情報科学部の渡邊拓貴准教授、筑波大学システム情報系の志築文太郎教授の研究グループは、高機能なイヤホン型デバイスであるヒアラブルデバイス（※1）から発生する音漏れの信号源への転用可能性に着目し、機械学習（※2）を用いて機器付近で実行される空中ジェスチャの検出および種類分類を行う新手法「EarHover」を開発しました。

多くのヒアラブルデバイスは、タッチセンサなどを利用した機器への接触による操作が主流ですが、機器付近上空で実行される空中ジェスチャが認識可能になることで、機器に直接触れなくてもヒアラブルデバイスを操作することが可能になります。

研究グループは、ヒアラブルデバイスから外部に音が漏れる現象である音漏れがデバイス周囲のセンシングに活用できることに着目しました。本手法では、ヒアラブルデバイスから人間には聞こえない超音波信号を再生し、その音漏れが空中ジェスチャを行う手に反射した時に発生するドップラー効果（※3）を活用しています。また、本手法は従来手法では利用が困難であった、手が汚れているなどの特定の環境下でも利用できます。加えて、ヒアラブルデバイス内蔵のスピーカと收音用マイクで実装可能であるため、市販製品への低コストでの導入が期待されます。

本研究の一部はJSPS 科研費（JP23KJ1884, JP21H03485）ならびにJST 戦略的創造研究推進事業さきがけ（JPMJPR2134, JPMJPR2138）の支援のもとで行われたものです。また、本研究成果は、人間と計算機の相互作用を扱うヒューマン・コンピュータ・インタラクション分野で最重要とされる国際会議の1つ「UIST '24: The ACM Symposium on User Interface Software and Technology」に採択され、Best Paper Awardの受賞が決まっています。

1. 本研究のポイント

- ・音漏れ現象とドップラー効果を利用することで、機器上空で実行される空中ジェスチャの認識が可能になり、非接触な機器操作が可能になる。
- ・超音波信号の音漏れと、その反射によって発生するドップラー効果を利用して空中ジェスチャによる操作を可能にする新手法を開発した。
- ・従来の手法では操作ができなかった、手が汚れている、手を清潔に保ちたい等の場合でも、機器を操作することが可能である。
- ・ヒアラブルデバイス内蔵スピーカと收音マイクで実装可能で、市販製品への低コストでの導入が期待される。

2. 研究背景

近年、ヒアラブルデバイスは、通話や音楽再生だけでなく、音声アシスタントやヘルストラッキング機能などを備え、我々の日常生活で重要な役割を担うようになりつつあります。これらの機能を日常的に利用するためには、ヒアラブルデバイス本体を用いた機器操作が重要となります。多くのヒアラブルデバイスは、タッチセンサを利用した操作が主流ですが、機器を直接手で触れる必要があり、手が汚れている場合や、清潔を保つ必要がある場合に、利用が困難であるという課題がありました。また、機器に触れることで耳と機器の間に圧力が生じ、痛みや押し込みノイズなどの不快感に繋がることがありました。

これらの課題に対処するため、機器付近上空で手を動かす・形状を再現する空中ジェスチャの導入が期待されています。しかし、これまで提案されてきた空中ジェスチャ認識技術は、カメラや赤外線センサを用いるため、実装コストが高く、デザインの複雑化が課題となっていました。

3. 研究内容・成果

本研究では、ヒアラブルデバイスの外部で発生する音漏れ現象に着目し、空中ジェスチャを認識する新手法を開発しました。ヒアラブルデバイス内蔵のスピーカから再生された信号が耳外側に漏れ、空中ジェスチャ時の手に反射し、收音マイクで測定されます（図 1）。測定された信号には、動いている手（空中ジェスチャ）に反射することで発生するドップラー効果が含まれるため、信号をスペクトログラム画像（※4）に変換し、機械学習を使ってジェスチャの検出・分類が可能となります。この手法は、市販製品の多くのヒアラブルデバイスに搭載されているスピーカと收音マイクで実装できるというメリットがあります。

本研究では提案した 27 種類の空中ジェスチャ（図 2）から、予備実験にて認識に最適な空中ジェスチャを 7 種類選定し（図 3）、その後、13 名の実験参加者を対象に、4 種類の利用環境下での空中ジェスチャ検出・分類性能を評価しました。深層学習を用いて推定モデルを作成した結果、ジェスチャではない日常生活の動作をジェスチャとして誤認識してしまう確率を 1.8%に抑えつつ、ジェスチャの種類を 80%程度で認識できることを確認しました。

4. 今後の展開

今後は、騒音の大きい環境などの実環境に近い状況での精度検証を行うほか、認識システムの改良および市販デバイスのみでの実装を目指します。また、音漏れ信号の利用可能な応用事例の探索と音漏れ信号を安全に利用するための利用ガイドラインの制定を進めていきます。

<参考図>

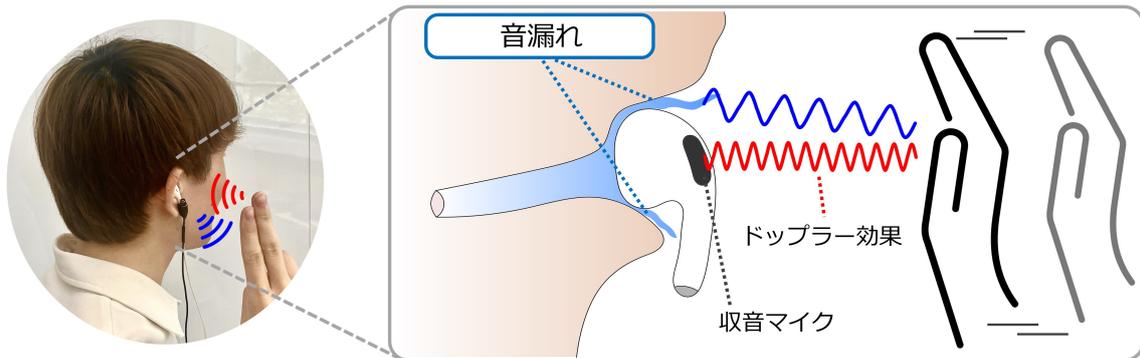


図1 新手法「EarHover」の音漏れ現象とドップラー効果発生の様子

本手法ではヒアブルデバイスから外部に漏れる音漏れと、音漏れが空中ジェスチャを実行する手に反射することで発生するドップラー効果を利用しています。

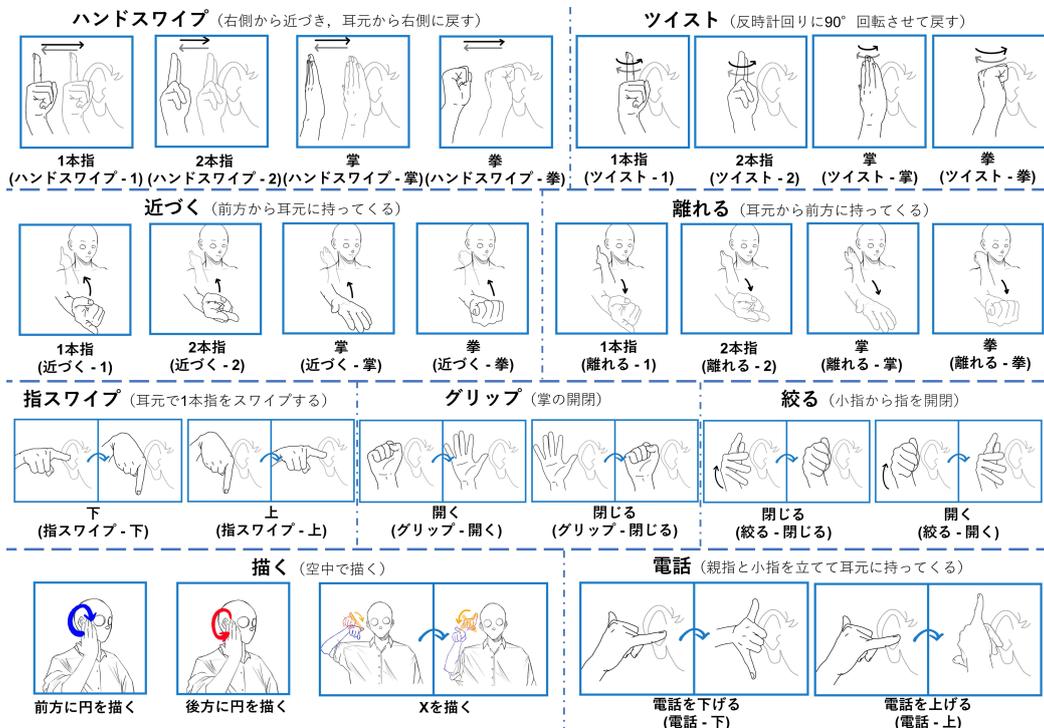


図2 本研究で提案した27種類の空中ジェスチャ

図2に示す27種類の空中ジェスチャから本手法に最適なジェスチャを選定しました。

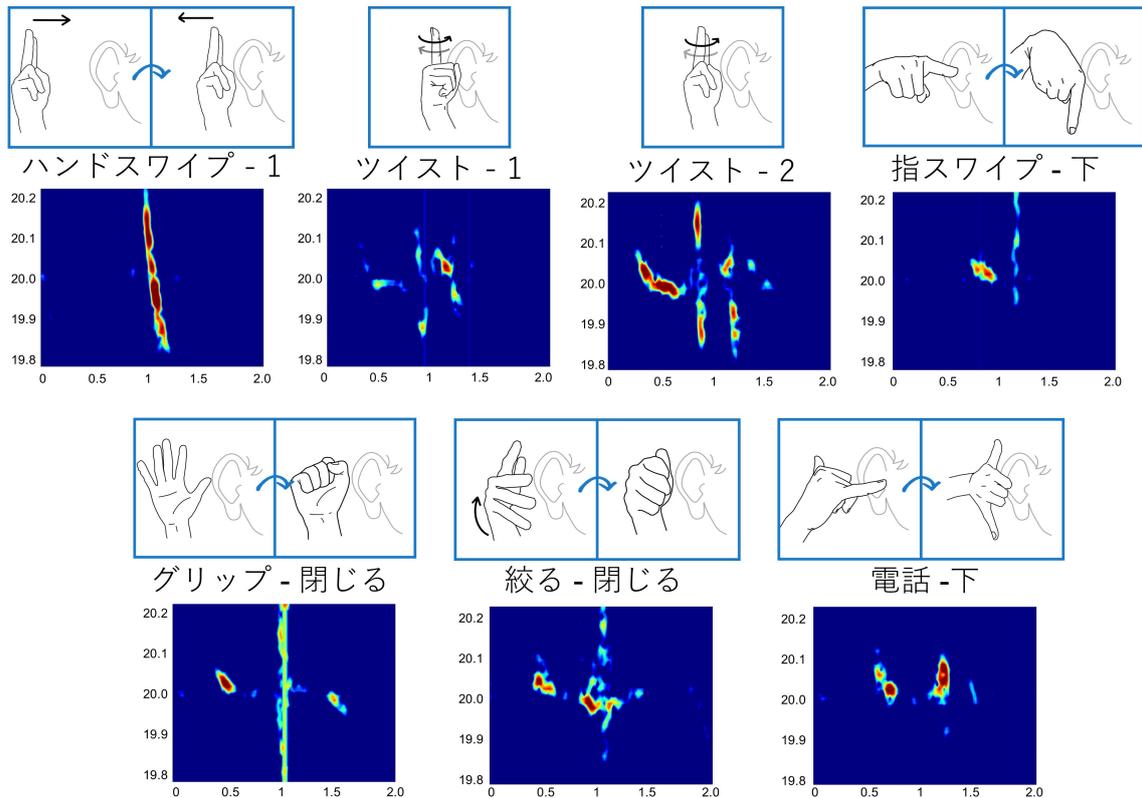


図3 最終的に選定された7種類の空中ジェスチャと、それらのスペクトログラム画像
 空中ジェスチャの識別しやすさと使いやすさの観点から、7種類のジェスチャを選定しました。それぞれの空中ジェスチャは発生するドップラー効果に違いがあることがわかります（スペクトログラム画像の縦軸は周波数で、横軸は時間を表しています）。

<原論文情報>

Shunta Suzuki, Takashi Amesaka, Hiroki Watanabe, Buntarou Shizuki, and Yuta Sugiura. 2024. EarHover: Mid-Air Gesture Recognition for Hearables Using Sound Leakage Signals (ヒアラブルデバイスにおける音漏れ信号を用いた空中ジェスチャ認識). In The 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24), October 13-16, 2024, Pittsburgh, PA, USA. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3654777.3676367>

<用語説明>

※1 ヒアラブルデバイス: 音声アシスタントやヘルストラッキング機能などを搭載した、従来のイヤホンと比較して高機能なイヤホン型デバイスのことです。

※2 機械学習: 与えられたデータをコンピュータが学習し、学習した成果に基づいて予測や判断を行う仕組みのことです。

※3 ドップラー効果：音声などの波の発生源と観測者の間に速度があるときに、観測される音の周波数が増える現象のことです。

※4 スペクトログラム画像：音声などの周波数情報の時間変化を可視化した画像です。

※ご取材の際には、事前に下記までご一報くださいますようお願い申し上げます。

※本リリースは文部科学記者会、科学記者会、各社科学部等に送信させていただいております。

・研究内容についてのお問い合わせ先

慶應義塾大学 理工学部 情報工学科 准教授 杉浦 裕太

TEL：045-566-1769 E-mail：sugiura@keio.jp <https://lclab.org/projects/earhover>

・本リリースの配信元

慶應義塾広報室（増田）TEL：03-5427-1541 FAX：03-5441-7640

E-mail：m-pr@adst.keio.ac.jp <https://www.keio.ac.jp/>

筑波大学広報局 TEL：029-853-2040 FAX：029-853-2014

Email：kohositu@un.tsukuba.ac.jp <https://www.tsukuba.ac.jp/index.html>

公立はこだて未来大学 事務局企画総務課 企画総務担当

TEL：0138-34-6448 FAX：0138-34-6470

Email a-dm@fun.ac.jp