音漏れ信号を用いたヒアラブルデバイス向け手形状認識手法の検討

雨坂 宇宙 * 渡邉 拓貴 † 杉本 雅則 † 志築 文太郎 *

概要. 次世代のウェアラブルデバイスとしてヒアラブルデバイスが注目されている．ヒアラブルデバイスは，これまでの音楽鑑賞/通話機能だけでなく，音声ガイドやメッセージ読み上げなど様々な機能が搭載されると考えられる．これらの種々の機能の切り替え，ON/OFF 操作をヒアラブルデバイスから実行するために，多くのデバイス操作手法が研究されてきた．本研究ではヒアラブルデバイスの音漏れが手と耳介で形成される空間内を反響することに着目し，耳介反響音を測定，解析することで現在の手形状を認識する手法を開発した．提案手法はヒアラブルデバイスにて一般的に搭載されているスピーカと外側マイクで実装できる点でヒアラブルデバイスとの親和性が高い．本研究ではプロトタイプデバイスを用いて，5 名の実験参加者から 5 種類の手形状のデータを収集した．実験の結果，参加者ごとの個人分類モデルによる認識精度が平均 92.7%（F 値）であることを確認した．

1 はじめに

近年の研究/技術開発に伴い，イヤホン型のウェアラブルデバイス（ヒアラブルデバイス）が更に注目されている．ヒアラブルデバイスとは従来の音楽鑑賞や通話としての機能だけでなく，様々な機能が付加された高機能なイヤホン型デバイスを指す．現在も多くの研究や開発が進められており，今後はより多くのアプリケーションがヒアラブルデバイスに搭載されると考えられる．一方，現在市販されているヒアラブルデバイスの多くはスマートフォンと連携して使用することを想定しているが，デバイスを操作するために，まずはスマートフォンの画面をタッチする必要があり，ユーザビリティを損う．そのため，ヒアラブルデバイスのみを用いたデバイス操作手法が求められている．市販されているヒアラブルデバイスでは，デバイスに感圧センサや静電容量センサ，物理ボタンを埋め込み，タッチや押し込みを認識する手法が採用されている．これらのセンサの利用は，感圧センサではデバイスデザインの制約が生じ，静電容量センサでは手袋をしたままの利用ができない．また，物理ボタンではボタンの押し込み動作によって外耳への押し込みが生じるため身体負担がある．その他に，これらのセンサ・ボタン押下によるデバイス操作手法は，押下回数や押下時間を変えることでコマンドの指定を行うが，指定できるコマンドに限りがある．そのため，市販されている AirPods [3] などのヒアラブルデバイスの多くは，1〜3 回の押下回数と長押し，上下スワイプの 6 種類程度のコマンドの利用に限られている．したがって，例えば音楽鑑賞中に受信メッセージを読み上げる機能や，時刻/予定を確認する機能を実装するためには，タッチ入力とは異なるデバイス操作手法が追加で必要となる．

ヒアラブルデバイスのための操作手法は，主に手を使わない入力（ハンズフリー入力）と手を使って行う入力（ハンズ入力）にわかれる．ハンズフリー入力手法は，音声入力手法や慣性/気圧/筋電/赤外線/音響センシング [1,2,4,9,10] を用いた頭部ジェスチャ手法などがこれまで研究されてきた．音声入力の利用は公共の場での発声や騒音環境下での認識精度の低下などの課題がある．頭部ジェスチャの利用は，頭を振るなどの大きな頭部の動きを伴うジェスチャは周囲に与える印象の懸念や視線が外れてしまう危険性の課題がある．一方，顎をずらすなどの小さな頭部の動きで実現できるジェスチャは日常生活で行う動作と混同する課題がある．ハンズ入力手法ではヘッドホンに回路を追加することでデバイスのタップ検知を行う手法 [13] や，デバイス内蔵マイクを用いて頬や耳に触れながら行うスワイプジェスチャを認識する手法 [12] がある．前者の手法は追加の回路が必要であるという課題があり，後者の手法はマスクなどの装着中に一部のジェスチャが利用できない．

本研究ではヒアラブルデバイスが音響信号を再生する時に発生する外部へと漏れる音響信号（音漏れ信号）に着目し，装着者の手形状をジェスチャとして認識する手法を開発した．具体的には，ヒアラブルデバイスから周波数が遷移するチャープ信号を再生することで発生する音漏れ信号を，デバイス外側に搭載されたマイクにて録音する．この時，ユーザがデバイスを手で覆うことにより音漏れ信号は手と耳介で形成された空間内で反響する．我々はこの反響音を耳介反響音と定義する．この時，手の覆い方を変えることで得られる耳介反響音が変化すると考えられる．この変化パターンを機械学習手法を用い

* 筑波大学
† 北海道大学

て認識することで手の覆い方をジェスチャとして利用する．提案手法はヒアラブルデバイスに一般的に搭載されているスピーカと外側マイクを利用する点で実装コストが抑えられる．本研究では提案手法の実現可能性を検討するためにプロトタイプデバイスにてデータの収集を行い，ジェスチャ認識性能の評価を行った．本研究の成果を以下にまとめる．

- ヒアラブルデバイスのスピーカと外側マイクで実装可能な手形状認識手法を提案した．
- 5 名の参加者データより，5 種類の手形状を 92.7%（F 値）の精度で認識できることを示した．

2 関連研究

ヒアラブルデバイスの操作を目的としたジェスチャ認識手法は，ハンズフリー/ハンズ入力手法にわかれる．

2.1 ハンズフリー入力手法

ハンズフリー入力手法は，主に顔の動きや表情の変化などを頭部ジェスチャとしてデバイスを操作する．Ando [2] らは顔関連の動作時に生じる外耳道内部の気圧変化を気圧センサで取得し，ユーザごとに 11 種類の顔関連の動作を 87.6%の精度で認識することに成功している．Matthies ら [9] は電極を貼付したイヤーピースを用いて外耳道内部から筋肉の動きを読み取り 5 種類の頭部状態を精度 90.0%（座位状態），85.2%（歩行状態）で認識することに成功している．Taniguchi ら [10] は LED とフォトトランジスタを用いて舌の特定の動きを外耳道の変形から認識している．また，音楽プレイヤの操作を対象にユーザビリティの調査も行っている．Bedri ら [4] は近接センサを用いて外耳道の変形を読み取り，心拍数，舌・顎の動作，まばたきを認識している．製品では，Bragi の DashPro が慣性センサを用いて頭部ジェスチャを認識し，電話の応答や音楽プレイヤの曲変更などの簡単なハンズフリー操作を実現している．Amesaka ら [1] は外耳道内部の音を取得できるマイクを用いて，首，顎，顔の状態（頭部状態）にともなって外耳道が変形することに着目し，スピーカから流した測定信号をもとに，外耳道伝達関数を測定，解析することで頭部状態を認識している．

これらの手法は，ハンズフリーでジェスチャが行えるというメリットがある．一方で，頭を振るなどの大きな動きを必要とするジェスチャは視線が外れる危険性や公共の場での利用が憚られるという課題がある．顎をずらすなどの小さな頭部の動きで実現できるジェスチャは日常生活で行う動作と混同する課題がある．また，これらの手法は追加のセンサや内向きマイクなどが必要なためデバイスコストや導入コストなどの問題も生じやすい．提案手法が目指すハンズ入力手法は片手を使う必要があるが，視認性や社会的受容性が高い動作でデバイス操作が可能となる．また，提案手法はヒアラブルデバイスに一般的に搭載されているスピーカと外側マイクで実装可能である．

2.2 ハンズ入力手法

ハンズ入力手法に関する研究では，真鍋ら [13] が市販されているヘッドホンに簡単な回路を組み合わせることでヘッドホンのタップ動作を認識することに成功している．また，菊池ら [14] は反射型リフレクタを用いて 5 つの方向ジェスチャの認識に成功している．Xiu ら [12] はデバイス内蔵マイクを用いて頬や耳上で行うスワイプジェスチャ時に発生する音を認識し，8 種類のジェスチャを 95.3%の精度で認識することに成功している．

真鍋らや菊池らの手法はデバイスへの回路の追加やセンサを追加する必要があり，追加のデバイスコストが生じる．また，Xiu らの手法はマスクの装着中に頬上のジェスチャなどの一部が利用できないという課題がある．提案手法はヒアラブルデバイスで一般的に搭載されているスピーカと外側マイクで実装可能であり，デバイスコストを抑えることができる．また，ユーザはマスクや眼鏡を装着したまま利用できる．

3 提案手法

提案手法の全体の流れを図 1 にまとめた．ユーザはデバイスを装着し，耳介を手で覆う．この状態でデバイスから測定信号を再生することでデバイスからの信号が耳介側へと漏れ，手と耳介で形成された空間内で反響する．この反響音をデバイスの外側に搭載されたマイクで録音する．得られる耳介反響音はユーザの手の形状によって変化する．これらの手形状の違いをジェスチャとして利用することでデバイスの操作手法に用いる．認識システムは得られた反響音の特徴量を抽出し，機械学習を行うことで手形状認識モデルを作成する．

3.1 特徴量＆機械学習モデル

取得した音響信号からジェスチャの特徴を示す特徴量の抽出を行う．本研究ではフーリエ変換による周波数成分（FFT: Fast Fourier Transform)/メル周波数ケプストラム係数（MFCC: Mel-Frequency Cepstrum Coefficients）/線形周波数ケプストラム係数（LFCC: Linear-Frequency Cepstrum Coefficients）[8] の 3 種類から最も精度の高かった MFCC を特徴量に用いる．選定過程は 5.1 節に後述する．機械学習モデルには勾配ブースティング決定木アルゴリズム（GBM：Gradient Boosting Machine）を用いる．分類モデルを用いて現在の耳介反響音から手形状の予測を行う．

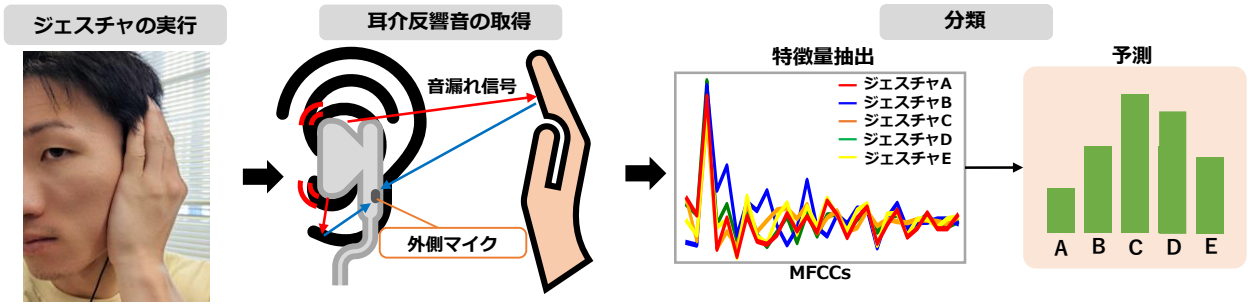

図 1. 提案手法の流れ.

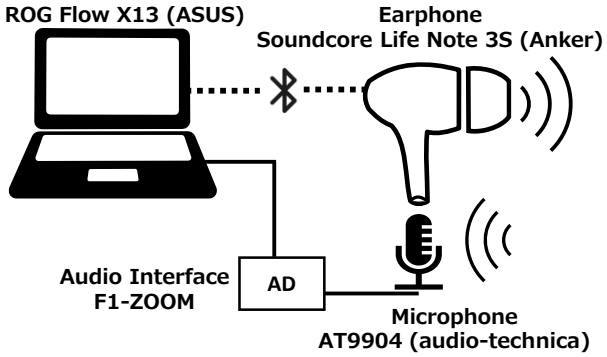

図 2. デバイス構成.

図 3. 実験デバイス.

4 実装

4.1 ハードウェア

提案手法はデバイスの外側マイクを利用することを想定しているが，市販のイヤホン型デバイスに搭載されているマイクはローパスフィルタが内部的にかけられており，高周波帯域の周波数の利用が制限されている．提案手法の本来の性能を調査するために，本実験では市販のイヤホンに別のマイクを固定する．また，予備調査よりカナル型イヤホンの音漏れは小さいことがわかったため，今回の実験ではインナーイヤー型のイヤホンを用いる．デバイス構成を図 2 にまとめた．図 3 に実験に用いたデバイスを示す．本デバイスでは，イヤホンに Anker 社の Soundcore Life Note 3S，マイクに audio-technica 社の AT9904 を用いた．マイクから取得した耳介反響音の AD 変換には ZOOM 社の F1 オーディオインタフェースを使用し，PC（ASUS 社: ROG Flow X13）へとデータを保存した．

4.2 ソフトウェア

本研究で使用するチャープ信号は 0 から 22 kHz のアップチャープ信号 (4,096 サンプル) を用いた．信号の再生/録音のサンプリングレートは 44.1 kHz である．信号の再生は PC と Bluetooth 接続したイヤホンより Realtek(R) Audio ドライバを用いて再生する．信号の音量は音漏れを発生させるために通常より大きく設定する必要がある．今回の信号音量は外耳形状を模した模型と騒音計（SANWA SUPPLY CHESD1）を用いて測定したところ，鼓膜部分にて約 80 dB であった．これは地下鉄車内や航空機内と同程度の騒音レベルであり，短時間での使用を想定すると聴覚健康上問題ないといえる [11]．MFCC 抽出のためのフーリエ変換のウィンドウサイズは 16,384 サンプルとし，8,192 サンプル分のスライドを行った．また，30 分割のメルフィルタバンクを用いて MFCC を 30 次元分抽出し，直流成分である 1 次元目を除いた．機械学習では訓練データを 5 分割の交差検証を行い，F 値が最も高いパラメータをハイパーパラメータとした．データの測定，特徴量抽出，機械学習のプログラムは Python3.9 で実装した．

5 評価実験

本実験では 5 名（20〜27 歳）の参加者から耳介反響音を取得した．手形状の種類は図 4 で示す 5 種類である．実験は研究室内にて座位状態で行った．研究室の空調の電源は切っていない．実験参加者は測定デバイスを図 5 のように右耳に装着し，右手で手形状を再現してもらう．その間に，チャープ信号を再生し，耳介反響音を 5 秒間録音する．これを 5 種類のすべての手形状で測定したものを 1 ラウンドと

図 4. 評価した手形状.

図 5. 実験デバイス装着時の様子.

する．ラウンド終了後にデバイスの脱着を挟み，合計で 6 ラウンド分のデータを取得した．参加者はマスクを装着し，眼鏡装着者は装着したまま行ってもらった．

5.1 特徴量と認識精度の関係

提案手法に有効な特徴量の選定を行う．音響信号の反響・伝播を解析し，物体・人体の状態などを認識するアクティブ音響センシングで一般的に利用される特徴量の FFT/MFCC/LFCC の 3 種類の特徴量を調査した．各特徴量の抽出に用いられるウィンドウサイズは 16,384 サンプルとし，8,192 サンプル分のスライドを行った．そのため，FFT で得られる周波数成分は 8,192 次元となるがデータ数に対して次元数が大きい．したがって，本研究では 10/40 次元ごとの間引きを行い，819/204 次元まで次元圧縮を行った．また，MFCC/LFCC の抽出に用いられるフィルタバンクの分割数は 30 分割と 50 分割で調査する．なお，直流成分である 1 次元目を除くため次元数はそれぞれ 29 次元と 49 次元となる．

取得した 5 名のデータを用いて特徴量ごとの精度の差異を評価する．精度の調査には参加者ごとのデータを用いて個人ごとの分類モデルを作成する leave-one-round-out 交差検証を行った．図 6 に各特徴量での認識精度（F 値）の平均をまとめた．実験の結果，MFCC の 2 パターン共に F 値が 88.6% で最も精度が高かった．次元数は少ないほうが機械学習の効率が良くなるため，以下の調査では 30 分割フィルタバンクによる MFCC を特徴量として用いる．

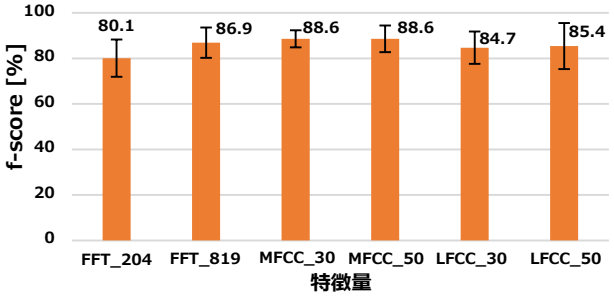

図 6. 特徴量ごとの認識精度.

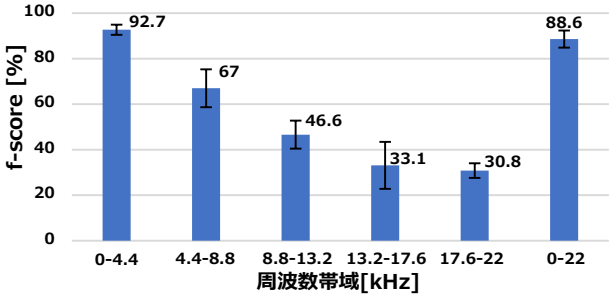

図 7. 利用周波数帯域ごとの認識精度.

5.2 利用周波数帯域と認識精度の関係

実験で利用したチャープ信号の帯域は 22 kHz 分である．得られたデータより認識性能に有効な帯域を調査するために全体の 20%にあたる 4.4 kHz ごとに信号を帯域制限した場合の各認識精度を調査した．精度の調査には参加者ごとのデータを用いて個人ごとの分類モデルを作成する leave-one-round-out 交差検証を行った．各参加者の認識精度の全体平均を周波数帯域ごとに図 7 にまとめた．実験の結果，0 から 4.4 kHz までの周波数帯域を利用した特徴量による認識精度が最も高く，F 値は 92.7%であった．図 8 はジェスチャの予測分布を示す混同行列である．最も認識精度の高かった手形状は「覆う（空間なし）」で，最も認識精度の低い手形状は「覆う（空間あり）」であった．

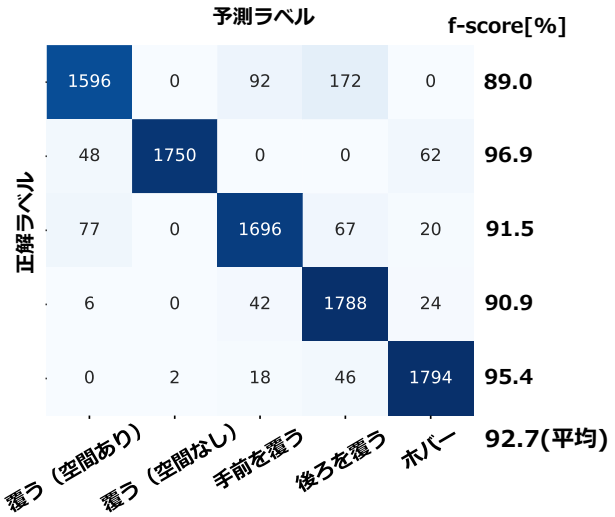

図 8. 混同行列 (0–4.4kHz).

5.3 一般分類モデルの評価

上記の実験により最も精度の高かった 0 から 4.4 kHz の帯域の信号を用いて，一般分類モデルの評価を行う．テストデータ以外の参加者のデータを訓練データとして作成した分類モデルでテストデータを評価する leave-one-user-out 交差検証を行った．調査の結果，各参加者の F 値の平均は 65.7% (SD=11.7) であった．個人分類モデルの認識精度に比べて，27 ポイントの精度低下が確認された．

6 議論および今後の課題

6.1 ジェスチャ認識性能

5.2 節より，0 から 4.4 kHz に帯域制限したデータによる認識精度が最も高いことがわかった．一方，高周波の信号帯域の認識精度が低いことがわかった．実験参加者 1 名の各手形状の周波数特性と MFCC の t-SNE 分布を図 9 にまとめた．周波数特性から分かるように，4 kHz 付近までの周波数応答が顕著であり，かつ手形状ごとの特性の変化も大きい．一方で，15 kHz 付近以降の周波数応答は著しく低く，手形状ごとの変化も小さい．周波数応答の不均衡の原因は，信号の減衰率やマイク特性などが影響していると考えられる．今後は，高周波帯域の信号音量を上げることで，高周波数帯域の応答を高くする改善を行う．高周波帯域の応答改善によって，認識性能が向上するか調査する．また，特徴量に用いた MFCC は，人間の聴覚の特徴である低周波帯域の音に敏感な聴覚特性に合わせたフィルタ設計が行われている．したがって，高周波帯域の音響データによる認識には不適であったと考えられる．今後は，周波数帯域に依存しない特徴量を用いた調査を進める予定である．これらの調査をもとに，信号の最適な周波数遷移帯域や周波数ごとの音量を決定する必要がある．

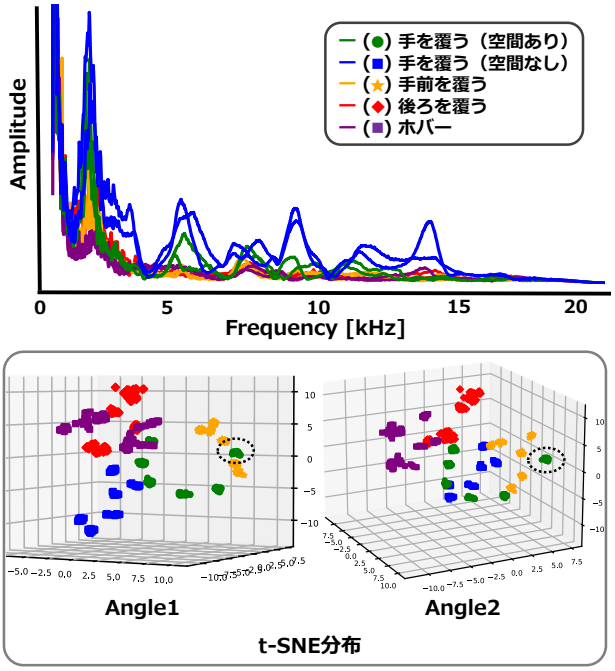

図 9. 各手形状の周波数特性（上）と t-SNE 分布（下）.

各手形状の認識精度について，最も認識精度の低い形状は「覆う（空間あり）」であったが，「手前/後ろを覆う」のそれぞれと混同する傾向にあった（図 8）．図 9 の t-SNE 分布からも，黒破線で囲まれた「手を覆う（空間あり）」のデータ群は同ラベルのデータ群よりも「手前を覆う」のデータ群に近いことがわかる．これらの手形状は共に手の形は同じで，開放する部分のみが変わる．そのため，開放部分が小さかった場合，得られる周波数特性が似てしまうと考えられる．今回の実験では，実験参加者に手前や後ろを覆う時の開放部分の幅などを指定しなかった．今後は，開放部分の幅と誤認識率の関係を調査することで，適切な手形状の指示を行うなどの改善を行う予定である．

6.2 ジェスチャ検出/ユーザビリティ

提案手法は可聴域の信号を利用する必要があり，常に信号を再生し続けるのはユーザビリティを大きく損なう．また，ジェスチャ認識を常に行う場合，生活動作をジェスチャとして誤認識し，意図しないジェスチャが発生する問題もある．したがって，提案手法を実環境で利用する場合，ユーザが手形状をジェスチャとして行ったことを検出し，ジェスチャ認識モードに認識システムを切り替える必要がある．ジェスチャの検出には，例えば搭載されているタッチセンサをトリガーとする方法などが考えられ，これは実装が容易である．しかし，タッチセンサが必要なうえに入力コマンドを一つ減らすことになる．提案手法を応用することで，タッチセンサなどを利

用せずに，ジェスチャ検出を行う方法も今後調査する必要がある．例えば，音楽などのユーザが気になりにくい信号や超音波帯域（18 kHz–22 kHz）のみのチャープ信号を利用することで，耳介反響音によるジェスチャ検出を検討している．今後は，これらの調査も行っていく予定である．

また，今回の実験では信号の音量が鼓膜部分でおよそ 80 dB であった．信号音量は大きいほど，音漏れが大きくなり得られる周波数応答も良好になるが，ユーザへの聴覚負担も大きくなる．一般的に，85 dB 以上の音を長時間聴き続けると，騒音性難聴のリスクが高まるとされている [11]．今後は，信号音量と認識精度の関係やアンケートによる聴覚への負担などを調査し，最適な信号音量を決める必要がある．また，音楽再生時に提案手法を利用するためには，音楽再生を中断する必要があり，ユーザビリティの低下につながる．例えば，測定信号の再生帯域を超音波帯域のみにする手法が考えられるが，超音波信号の再生/録音に対応したデバイスが必要である．他には，ユーザが聴いている音楽の応答を利用して周波数特性を測定する手法 [7] などもある．今後は，これらの手法を提案手法に組み込むことで音楽再生時にも利用可能な手法に改善する予定である．

6.3 実用環境/オープンイヤーデバイスでの調査

本研究は提案手法の実現可能性の調査が目的であったため，安定環境（静かな部屋/座位状態）で収集したデータを用いた基礎的な性能調査に留まっている．今後の実利用を考慮すると，ノイズ環境やモバイル環境での調査を進めていく必要がある．また，本実験はインイヤータイプのデバイスによる実験のみであったが，オープンイヤータイプ [5] のデバイスでも耳介反響音は取得できると考えており，このようなデバイスタイプでは信号音量も小さくできる可能性が高い．今後は，このようなデバイスタイプの手形状認識性能も調査していく予定である．

6.4 認識モデルの一般化と個人認証への応用

5.3 節より，一般分類モデルによる認識精度は個人分類モデルに比べて 27 ポイントも低下することがわかった．これは参加者ごとに耳介や手のサイズ，形状が大きく異なるため，得られる音響データも異なったことが原因だと考えられる．一般分類モデルの精度を向上させるためには，個人の耳介/手形状に依存しない特徴量を見つけ出す特徴量エンジニアリングなどの認識アルゴリズムの改善が考えられる．また，動きや手の形状変化を伴うジェスチャを採用することでドップラー効果による周波数変化を利用し，個人に依存しない手の動作を認識に利用する手法などが考えられる．

一方で，これらの個人の差異を利用することで個人認証が実現できる可能性がある．ヒアラブルデバイスによる個人認証手法も多くの研究 [6] が進められており，デバイスから秘匿性の高い情報にアクセスするための認証や決済アプリケーションなどの応用が考えられる．我々は，現在保持している 5 名の音響データより「手を覆う（空間あり）」のデータを用いて簡易的な個人識別モデルを作成した．その結果，5 名の参加者の個人識別率は 94.5%であることが確認された．したがって，今回我々が測定した耳介反響音は個人識別/認証システムへの応用も期待できる．

7 まとめ

本研究ではイヤホン型デバイスの音漏れ信号を利用した手形状認識手法を提案し，認識システムの実装を行った．利用する特徴量や周波数帯域と提案手法の認識性能の関係などを調査した．実験の結果，5 名の参加者による 5 種類の手形状認識において 0 から 4.4 kHz の MFCC を用いることで 92.7%（F 値）の認識精度が得られることを確認した．

謝辞

本研究は JSPS 科研費 21J10706/18H04104, JST さきがけ JPMJPR2138 の助成を受けたものです．

参考文献

[1] T. Amesaka, H. Watanabe, and M. Sugimoto. Facial Expression Recognition Using Ear Canal Transfer Function. In *Proceedings of the 23rd International Symposium on Wearable Computers*, ISWC ’19, pp. 1–9. ACM, 2019.

[2] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi. CanalSense:Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST ’17, pp. 679–689. ACM, 2017.

[3] Apple Inc. AirPods, (2022.09.22. accessed). `https://www.apple.com/jp/airpods-pro/`.

[4] A. Bedri, D. Byrd, P. Presti, H. Sahni, Z. Gue, and T. Starner. Stick It in Your Ear: Building an In-ear Jaw Movement Sensor. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, UbiComp/ISWC ’15 Adjunct, pp. 1333–1338. ACM, 2015.

[5] BOSE. WHAT ARE OPEN-EAR HEADPHONES?, (2022.09.22. accessed). `https://www.bose.com/en_us/better_with_bose/open-ear-headphones.html`.

[6] Y. Gao, W. Wang, V. V. Phoha, W. Sun, and Z. Jin. EarEcho: Using Ear Canal Echo for Wearable Authentication. *Proceedings of*

the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 3(3), 2019.

[7] H. Kawahara and K. Yatabe. Safeguarding test signals for acoustic measurement using arbitrary sounds: Measuring impulse response by playing music. *Acoustical Science and Technology*, 43(3), 2022.

[8] H. Lei and E. L. Gonzalo. Mel, linear, and antimel frequency cepstral coefficients in broad phonetic regions for telephone speaker recognition. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association*, INTERSPEECH '09, pp. 2323–2326. ISCA, 2009.

[9] D. J. C. Matthies, B. A. Strecker, and B. Urban. EarFieldSensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input Through Facial Expressions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 1911–1922. ACM, 2017.

[10] K. Taniguchi, H. Kondo, M. Kurosawa, and A. Nishikawa. Earable TEMPO: A Novel, Hands-Free Input Device that Uses the Movement of the Tongue Measured with a Wearable Ear Sensor. *Sensors*, 18(3), 2018.

[11] World Health Organization. Safe Listening Devices and Systems, 2019.

[12] X. Xu, H. Shi, X. Yi, W. Liu, Y. Yan, Y. Shi, A. Mariakakis, J. Mankoff, and A. K. Dey. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–14. ACM, 2020.

[13] 真鍋 宏幸, 福本 雅朗. Headphone Taps：通常のヘッドホンへのタップ入力. 情報処理学会論文誌, 55(4):1334–1343, 2014.

[14] 菊地 高史, 正井 克俊, 杉本 麻樹, B. H. Thomas, 杉浦 裕太. EarTouch: 耳の入力インタフェース化. 日本バーチャルリアリティ学会論文誌, 22(4):477–483, 2017.