

ジェスチャ認識のための動的  
アクティブ音響センシング手法の確立

2023年3月

雨坂 宇宙

ジェスチャ認識のための動的  
アクティブ音響センシング手法の確立

雨坂 宇宙

理工情報生命学術院  
システム情報工学研究群  
筑波大学

2023年3月

## 概要

ウェアラブルデバイスの発展において、ユーザの多様な嗜好や利用環境に適応するために多様なデバイスタイプ（イヤホン型、眼鏡型、指輪型、衣服型）に様々な機能（ライフログ、入力インタフェース、バイタルセンシング）を搭載する研究・開発が進められている。これらの研究・開発の多くは、ユーザの状態・動きなどをセンシングし、ユーザにあった機能・サービスを提供することを目的としている。その実現のために、様々な種類のセンシング手法（慣性、筋電、音響、光）から、センシング対象や組み込むデバイスタイプなどの条件を基に利用されるセンシング手法が選定される。その中の一つにアクティブ音響センシング（AAS: Active Acoustic Sensing）手法がある。AASは人や物体に測定信号を伝播・反射させ、その応答を解析することで動きや状態を認識するセンシング技術である。AASは電気信号や光信号が適用できない不導体や光を通しにくい物体のセンシングが可能となるメリットがある。また、スピーカやマイクはスマートフォンやスマートウォッチ、イヤホンなどの多くのデバイスに音楽鑑賞や通話、音声アシスタント利用のために搭載されている。そのため、スピーカやマイクをセンシング技術に応用することは実装コストなどの観点からも有用といえる。

一方で、これまでのAASは一定の測定信号を再生し、信号の伝播や反射を取得・解析する手法が利用されてきたが、デバイスの位置ずれ（装着誤差）によって認識精度が低下してしまう課題があった。また、これまでのAASに関する研究はブロックや人体などの形状が安定しているセンシング対象 [1, 2, 3] に適用され、タッチインタフェース/ジェスチャ入力を実現する手法が多く研究されてきた。一方で、布などの形状が不安定なセンシング対象への調査が行われていなかった。これは、布などの形状が不安定な物体は信号伝播時の減衰が大きく、また、形状が不安定なために信号の変化も不安定であるという課題があるためである。

本研究では上記の装着誤差と不安定なセンシング対象へのAASの適用という課題に対して、それぞれ測定信号の動的補正手法とAASの動的利用手法を提案した。これらの提案手法はこれまでのAASとは異なり、測定信号とAASの利用方法が動的に変化する。我々はこれらの提案手法をまとめたセンシング手法を動的アクティブ音響センシング（DAAS: Dynamic Active Acoustic Sensing）と呼び、本研究にてその有効性を評価した。まず、測定信号の動的補正手法の有効性を評価するために、イヤホン型デバイス（ヒアラブルデバイス）のための頭部状態認識手法に本手法を適用し、調査を行った。実験の結果、測定信号の動的補正手法は装着誤差の低減に有効であることを確認し、認識システムの高精度化に貢献した。続いて、AASの動

的利用手法の有効性を評価するために、衣服上のタッチインタフェース構築手法に本手法を適用し、調査を行った。実験の結果、AASの動的利用手法はノイズ環境での認識精度の低下を改善することを確認した。その結果、これまでAASが適用されていなかった一般的な衣服に対して、ウェアラブルデバイスの機能の一部を埋め込むことに成功し、AASの適用可能範囲の拡大に貢献した。

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 研究背景	1
1.1.1 ウェアラブルデバイスの発展	1
1.1.2 センシング	2
1.1.3 アクティブ音響センシングに関する研究の現状および課題	3
1.2 研究概要	5
1.2.1 研究内容	5
1.2.2 研究調査対象	5
1.2.3 研究の貢献	6
研究テーマ1：動的信号補正による装着誤差の低減	6
研究テーマ2：ノイズ環境での AAS の動的利用	6
1.3 本論文の構成	7
<b>第2章 動的信号補正による装着誤差の低減</b>	<b>8</b>
2.1 外耳道伝達関数を用いた頭部状態認識手法	8
2.2 関連研究	9
2.2.1 人体へ AAS を適用した研究	9
2.2.2 外耳道インパルス応答の測定と応用に関する研究	10
2.2.3 頭部関連動作の認識に関する研究・製品	10
2.3 提案手法	11
2.3.1 認識原理	12
2.3.2 外耳道伝達関数測定原理	14
2.3.3 Swept-Sine 法	14
円状畳み込みの原理を用いたインパルス応答測定法	15
測定信号の修正によるインパルス応答の補正	15
2.3.4 特徴量抽出および分類モデル	17
2.4 実装	18
2.4.1 外耳道反響音取得デバイス	18
2.4.2 頭部状態認識アルゴリズム	19
2.5 評価実験	20
2.5.1 実験協力者および実験環境	20
2.5.2 分類モデルの評価方法	20
2.5.3 未補正信号による認識実験	20

	実験手順 . . . . .	21
	実験結果 . . . . .	21
2.5.4	補正信号による認識実験 . . . . .	22
	実験手順 . . . . .	22
	実験結果 . . . . .	22
2.5.5	各頭部状態のアンケート評価 . . . . .	23
2.5.6	アプリケーションを考慮した頭部状態認識実験 . . . . .	26
	頭部状態の選定 . . . . .	26
	実験結果 . . . . .	27
2.6	考察 . . . . .	28
2.6.1	動的信号補正手法に関する議論 . . . . .	28
	信号補正による装着誤差の低減 . . . . .	28
	測定信号の自動補正 . . . . .	30
	信号補正手法の定量的な性能調査 . . . . .	30
2.6.2	頭部状態認識手法に関する議論 . . . . .	30
	応用的な頭部状態の調査 . . . . .	30
	市販製品での実装 . . . . .	31
	実用環境での認識精度 . . . . .	31
2.7	まとめ . . . . .	31
<b>第3章</b>	<b>ノイズ環境での AAS の動的利用</b>	<b>32</b>
3.1	音響センシングを用いた衣服へのインタフェース構築手法 . . . . .	32
3.2	関連研究 . . . . .	33
3.2.1	音響センシングを用いたインタフェース構築に関する研究 . . . . .	33
	AAS を用いた研究 . . . . .	33
	PAS を用いた研究 . . . . .	34
3.2.2	衣服インタフェース構築に関する研究 . . . . .	34
3.2.3	衣服装着型デバイスに関する研究 . . . . .	35
3.3	本手法 . . . . .	36
3.3.1	衣服上タッチジェスチャ認識の実現可能性に関する基礎調査 . . . . .	37
	アクティブ音響センシング . . . . .	37
	パッシブ音響センシング . . . . .	39
3.3.2	ジェスチャ部分の抽出 . . . . .	39
3.3.3	特徴量抽出および分類モデル . . . . .	40
3.3.4	ピエゾ素子の選定 . . . . .	40
3.4	実装 . . . . .	41
3.4.1	信号再生・録音デバイス . . . . .	41
3.4.2	タッチジェスチャ認識アルゴリズム . . . . .	41
3.5	評価実験 . . . . .	42

3.5.1	前腕部タッチジェスチャ認識実験 . . . . .	43
	実験設計：測定ジェスチャ . . . . .	43
	実験設計：実験協力者および実験環境 . . . . .	44
	実験設計：実験手順 . . . . .	44
	ジェスチャ抽出精度 . . . . .	45
	認識システム設計のための予備調査：分類モデルの選定 . . . . .	46
	認識システム設計のための予備調査：周波数帯域の組み合わせパター ンの選定 . . . . .	46
	評価実験：衣服ごとの認識性能調査 . . . . .	50
	評価実験：データ測定時間と認識精度の関係 . . . . .	53
	評価実験：同一モデルの異なる衣服での精度評価 . . . . .	53
	評価実験：デバイス装着位置と認識精度の関係 . . . . .	55
	評価実験：一般分類モデルでの認識精度 . . . . .	56
3.5.2	ボタン/ポケット部タッチジェスチャ認識実験 . . . . .	57
3.6	考察 . . . . .	57
3.6.1	ノイズ環境での AAS の動的利用に関する議論 . . . . .	59
	AAS の動的利用に関する課題 . . . . .	59
	認識アルゴリズムの改善 . . . . .	59
3.6.2	衣服上タッチジェスチャ認識手法に関する議論 . . . . .	59
	ジェスチャ認識精度 . . . . .	59
	衣服と認識精度の関係 . . . . .	60
3.7	まとめ . . . . .	61
<b>第 4 章</b>	<b>結論</b>	<b>62</b>
4.1	本論文のまとめ . . . . .	62
4.1.1	動的信号補正による装着誤差の低減 . . . . .	62
4.1.2	ノイズ環境での AAS の動的利用 . . . . .	62
4.2	本研究の貢献 . . . . .	62
4.3	検討課題および発展 . . . . .	63
4.3.1	検討課題 . . . . .	63
	超音波利用のガイドライン制定 . . . . .	63
	モバイルデバイスでの実装および調査 . . . . .	63
4.3.2	発展 . . . . .	63
	動的信号補正手法の応用範囲の拡大 . . . . .	64
	AAS の動的利用手法の応用範囲の拡大 . . . . .	64
	謝辞	65
	参考文献	66

付録 A 動的信号補正による装着誤差の低減	77
A.1 実験後アンケート	77
付録 B ノイズ環境での AAS の動的利用	78
B.1 各ジェスチャのスペクトログラム	78
B.2 実験同意書	80
著者論文リスト	81



# 目 次

1.1 ユーザの特性とウェアラブルデバイスの発展. ユーザの特性（嗜好・利用目的・利用環境）に合わせて，デバイスと機能は多様化していく. . . . .	2
1.2 センシング手法の数値化モデル. ユーザのふるまいによって生じる物理量・化学量の変化はセンサを用いて数値化される. . . . .	3
1.3 AASの利用イメージ. a: 内部構造の異なるブロックの音響伝播特性を測定することでブロックの識別を行う手法, b: 手の動きで発生するドップラースhiftを測定し，ジェスチャとして利用する手法. . . . .	4
1.4 スピーカとマイクの利用例と利用周波数帯域. 再生・録音する周波数帯域によって用途を使い分けることが可能である. . . . .	4
1.5 AASの課題と本研究の概要. . . . .	5
2.1 頭部状態認識システム構成. . . . .	12
2.2 表情による外耳道伝達関数の変化. . . . .	12
2.3 外耳道周辺部および側頭部の骨格/筋肉モデル. . . . .	13
2.4 測定する頭部状態. . . . .	13
2.5 伝達関数測定原理. . . . .	14
2.6 円状畳み込みの原理を用いたインパルス応答測定. . . . .	16
2.7 デバイス着脱による外耳道伝達関数の変化 (未補正信号による測定時). . . . .	16
2.8 デバイス着脱による外耳道伝達関数の変化 (補正信号による測定時). . . . .	17
2.9 線形 20 分割フィルタバンク. . . . .	18
2.10 実験デバイス. . . . .	19
2.11 デバイス構成. . . . .	19
2.12 デバイス装着方法. a: 普通掛け, b: 耳掛け. . . . .	21
2.13 混同行列. a: 未補正信号による認識実験, b: 補正信号による認識実験 . . . . .	24
2.14 質問 1（再現性の困難さ）のアンケート結果. . . . .	25
2.15 質問 2（人前で再現することの印象）のアンケート結果. . . . .	26
2.16 6 種類の頭部状態での混同行列. a: 未補正信号による認識実験, b: 補正信号による認識実験. . . . .	28
2.17 デバイス着脱による振幅スペクトルの変化 (未補正信号). . . . .	29
2.18 デバイス着脱による振幅スペクトルの変化 (補正信号). . . . .	29
3.1 システムの利用想定イメージとプロトタイプデバイス a: イヤホン操作イメージ b: スマートグラス操作イメージ c: スマートホーム内での利用イメージ d: 実験に用いたプロトタイプデバイス. . . . .	33

3.2	認識原理. . . . .	36
3.3	タッチジェスチャ認識システム構成. . . . .	36
3.4	測定環境と変形用重りの配置位置. 右図の赤点の変形用重りの配置位置を示す. . . . .	37
3.5	張力テストによる周波数応答の変化. . . . .	38
3.6	変形テストによる周波数応答の変化. . . . .	38
3.7	前腕部ジェスチャ時のスペクトログラムの違い a: ねじるジェスチャ b: 円を描くジェスチャ. . . . .	39
3.8	3種類のピエゾ素子の組み合わせによる周波数特性の差異. . . . .	41
3.9	実験デバイス a: 実装デバイス b: デバイスの装着構造. . . . .	42
3.10	デバイス構成. . . . .	42
3.11	タッチジェスチャ認識の流れ. . . . .	43
3.12	測定する衣服タッチジェスチャ. . . . .	44
3.13	各分類モデルのジェスチャ認識精度の全体平均. エラーバーは参加者ごとの標準偏差を示す. . . . .	47
3.14	各周波数帯域組み合わせパターンの認識精度. エラーバーは参加者ごとのF値の標準偏差を示す. . . . .	48
3.15	各ノイズの周波数帯域組み合わせパターンの精度変化. . . . .	49
3.16	各衣服の全体認識精度. エラーバーは参加者ごとのF値の標準偏差を示す. a: 全参加者の平均 b: 全ての衣服のデータ測定に参加した6名の平均. . . . .	50
3.17	各衣服の混同行列. . . . .	51
3.18	5種類のジェスチャ認識時の各衣服の全体認識精度. エラーバーは参加者ごとのF値の標準偏差を示す. . . . .	52
3.19	5種類のジェスチャ認識時の全体の混同行列. . . . .	53
3.20	各ノイズ付加による認識精度の推移. 実線は本手法(全帯域+可聴域)による認識精度を示し, 破線は可聴域のみによる認識精度を示す. . . . .	54
3.21	データ測定の所要時間と, 認識精度およびユーザの印象度の推移. . . . .	54
3.22	別インスタンスデータを検証データとした場合の認識精度. . . . .	55
3.23	手首/肘側データを検証データとした場合の認識精度. . . . .	56
3.24	一般分類モデルによる各衣服のジェスチャ認識精度. エラーバーは参加者ごとのF値の標準偏差を示す. ×は各参加者のF値を示す. . . . .	56
3.25	t-SNE法を用いたシャツAの特徴量の分布. 青色は特定の参加者1名のジェスチャデータを示し, 赤色はその他の参加者のジェスチャデータを示す. a: 認識精度が最も高い参加者を青色で示した分布 b: 認識精度が最も低い参加者を青色で示した分布. . . . .	57
3.26	ボタン/ポケット部タッチジェスチャ a: デバイス位置とジェスチャ実行位置 b: 測定するタッチジェスチャ. . . . .	58
3.27	ボタン部とポケット部の各実験参加者の認識精度. エラーバーはジェスチャごとのF値の標準偏差を示す. . . . .	58

3.28 混同行列 a: ボタン部実験 b: ポケット部実験. . . . .	58
3.29 各衣服のスペクトログラム (スワイプ右/左). . . . .	60
3.30 デバイス距離を変化させたときのスワイプ右/左のスペクトログラム (シャツ B). . . . .	61

# 表 目 次

2.1	頭部関連動作の認識に関する研究の比較. . . . .	11
2.2	特徴量とフィルタバンクの精度に関する予備調査. . . . .	18
2.3	全ての頭部状態での各参加者の認識精度 [%]. . . . .	23
2.4	認識精度の高い頭部状態上位 7 種類 [%]. . . . .	24
2.5	認識精度の低い頭部状態下位 7 種類 [%]. . . . .	25
2.6	6 種類の頭部状態での各実験参加者の認識精度 [%]. . . . .	27
3.1	各手法の比較. . . . .	35
3.2	特徴量一覧. . . . .	40
3.3	実験測定に用いた 4 種類の衣服詳細. . . . .	45
3.4	各タッチジェスチャのジェスチャ抽出精度 [%]. . . . .	46
3.5	各分類モデルのパラメータ調査範囲 (SVM, RF, NN, GBDT) とモデル構成 (SEL). . . . .	47
3.6	各タッチジェスチャのジェスチャ認識精度 [%]. . . . .	52

# 第1章 序論

本章では研究全体の背景およびに研究概要（内容・調査対象・貢献）について述べ、最後に本論文の構成を示す。

## 1.1 研究背景

研究の背景は、ウェアラブルデバイスの発展（1.1.1項）とセンシング（1.1.2項）について説明し、本研究のテーマであるアクティブ音響センシングに関する研究の現状および課題（1.1.3項）について述べる。

### 1.1.1 ウェアラブルデバイスの発展

昨今のウェアラブルデバイスは更に発展し、スマートウォッチやイヤホン型デバイスは多機能化が進むと共に、多くのユーザに利用されるようになった [4]。ウェアラブルデバイスとは手や足などの身体に装着可能かつ、ユーザとデバイス間の情報入出力によってユーザの日常生活をサポートするデバイスを指す。例えば、ユーザの行動や生体情報などを記録（ライフログ/バイタルセンシング）する機能や、その他の接続機器を操作するための入力インタフェース機能を持つデバイスが、情報入力機能を持つウェアラブルデバイスである。一方、音楽再生やメール通知、経路案内などのユーザへ情報を提示する機能を持つデバイスが、情報出力機能を持つウェアラブルデバイスである。ウェアラブルデバイスに関する研究は、1990年代に Steve Mann らが中心となり、多目的な用途を有するデバイスの研究・開発が進められた。その後、ウェアラブルデバイスが2000年代頃から市場に出回り始めるが、カレンダーやアラーム、メモ帳などの簡易的な機能が、服や腕時計などのファッションアイテムに取り付けられたものであった。しかし、2010年代より Fitbit [5] や AppleWatch [6] などの健康管理機能や電子決済機能を備えたスマートウォッチ型デバイスが発売され、ユーザへの知名度が高まり、今では多くのユーザの生活に欠かせないデバイスとなった。

今後、更に多くのウェアラブルデバイスユーザを獲得するためには、ユーザの多様な嗜好や目的などに合ったデバイスや機能が求められる（図 1.1）。例えば、スマートウォッチやイヤホン型デバイスは装着時の機械感や装着感が、ユーザに精神的ストレスを与えるという課題がある [7]。また、デバイスが直接皮膚と触れることによる肌荒れや、イヤホン型デバイスを長時間装着することで外耳道内が高湿多湿状態になり外耳炎を発症する問題 [8] などもある。他にも、イヤホン型デバイスや眼鏡型デバイスは他の人に視認されるため、会議中などの利用が憚れる場面やデザイン性などの観点から利用しないユーザがいる。これらの課題を解消するために、機械感や装着感が少ない衣服タイプ [9, 10, 11] や視認されても目

立つことのないアクセサリタイプなど [12, 13, 14] の新しいデバイスタイプの研究・開発が現在も行われている。また、これまでのデバイスタイプに新しい機能を搭載する研究も進められている。例えば、画面が小さいスマートウォッチでは、キーボードを用いた文字入力時に誤入力が発生しやすいという課題（ファットフィンガ問題 [15]）が存在する。この問題解消のためにジェスチャを用いた文字入力手法などが現在も研究されている [16, 17]。

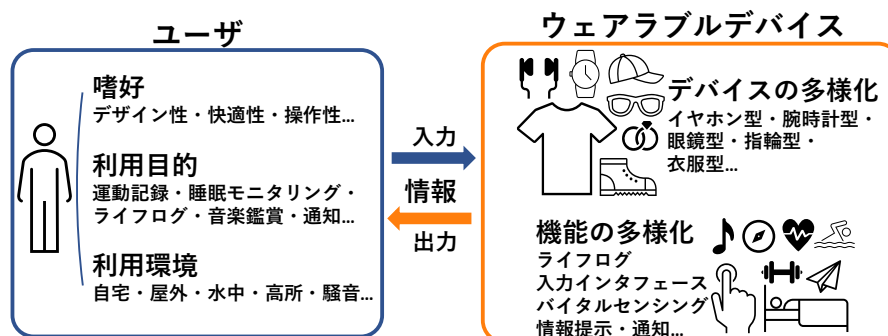


図 1.1: ユーザの特性とウェアラブルデバイスの発展。ユーザの特性（嗜好・利用目的・利用環境）に合わせて、デバイスと機能は多様化していく。

### 1.1.2 センシング

これらのウェアラブルデバイスに関する研究の内、特に情報入力機能に関する研究では、ユーザや物体の状態、動きをセンシングすることが研究の重要技術である。センシングとはセンサを用いて測定対象の定量的な情報を取得し、数値化する技術であり、センサごとに測定可能な物理量・化学量が異なる。図 1.2 にいくつかのセンシング手法の数値化モデルを示す。例えば、ユーザがデバイスへの入力のために腕を振るジェスチャを行った場合、腕には加速度や角速度の変化および筋肉の収縮が発生する。この加速度や角速度の変化を測定する場合、慣性センサが用いられ、筋収縮を測定する場合、筋電センサなどが用いられる。測定されたデータを基に各々のシステムはユーザの動作や物体の状態などを認識・識別する。したがって、センシング精度は高いほうが望ましいが、その他にデバイスに組み込むセンサの数やサイズ、システムの利用環境などを考慮して、デバイスに搭載されるセンサが決定される。例えば、慣性センサはセンサ自体の回転方向や加速度を測定できるため、センサを装着しているユーザの動作を認識する技術として利用でき、多くのウェアラブルデバイスに搭載されている。一方で、慣性センサは人体や物体内部、空間の状態を認識することには適していない。例えば、人体内部の情報である血中酸素濃度を測定する場合は、光の吸収率の変化を測定するため光センサが利用される [18]。また、ユーザの会話を記録するためにはマイクが必要であり、体温を測定するためには温度センサが必要となる。これらの例のように、多くのセンシング手法を利用し、ウェアラブルデバイスにおける様々な情報入力機能は実現されている。

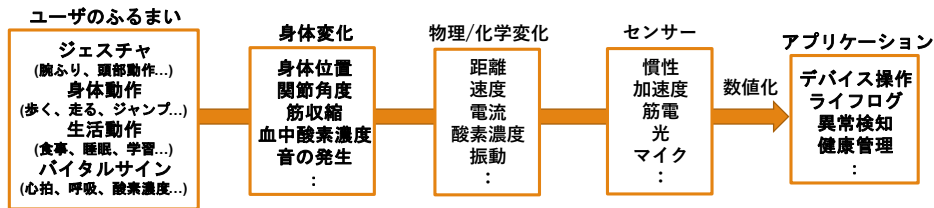


図 1.2: センシング手法の数値化モデル. ユーザのふるまいによって生じる物理量・化学量の変化はセンサを用いて数値化される.

### 1.1.3 アクティブ音響センシングに関する研究の現状および課題

様々なセンシング手法の中で、音を利用したセンシング（音響センシング）の一つにアクティブ音響センシング（AAS: Active Acoustic Sensing）がある。AASとは人や物体に主に超音波帯域（18kHz～）の測定信号を伝播・反響させ、その応答を解析することで状態や動きを認識するセンシング技術である。例えば久保ら [19] の研究では、物体内部を伝播する測定信号の周波数応答が物体の内部構造ごとに異なる性質を利用して、AASを用いた物体識別システムを提案している（図 1.3a）。Ruanら [20] の研究では、市販のPCやスマートフォンに内蔵されているマイクとスピーカを用いて、ユーザの手の動作時に発生するドップラシフトを測定し、手の動きを認識するシステムを提案している（図 1.3b）。

AASには電気信号や光信号が適用できない不導体や光を通しにくい物体のセンシングが可能である特長がある。また、AASで利用されるスピーカ・マイクはスマートウォッチ、イヤホンなどの市販されているウェアラブルデバイスの多くに搭載され、音楽再生や情報通知、通話のための音声の再生・録音に利用されている。その他に、AASで利用するマイクはパッシブ音響センシング（PAS: Passive Acoustic Sensing）にも利用できる。PASとは人や物体から自然発生する音を解析することにより、音の発生源の動作や状況を認識するセンシング技術である。例えば、皮膚上でタップやスワイプした時に発生する音を解析することで、ジェスチャとして利用する研究 [21] がある。また、睡眠時に発生するいびきや体動の音を解析し、ユーザの睡眠状況をモニタリングする研究 [22] などがある。これらの利用例のように、可聴域帯域（～18kHz）を音声の再生・録音、PASに有効利用することで幅広い情報入出力機能に応用できる（図 1.4）。したがって、AASを用いてライフログやジェスチャ入力が可能となれば、マイク・スピーカの有効活用が可能となる。これは、実装コストやデバイスの小型化などの観点からも有用といえる。

一方で、これまでのAASにはデバイスの位置ずれと不安定なセンシング対象へのAASの適用という未解決課題があった。前者に関して、ウェアラブルデバイスはデバイスの特性上、充電時や非利用時にデバイスを外すことがあり、デバイスの再装着時に元の装着状態とは異なる装着状態になることがある。また、運動中や歩行中などのユーザの動作によってデバイスに荷重がかかり装着状態が変化してしまう。これらのデバイスの装着状態の変化は装着誤差と呼ばれる。装着誤差は信号の伝播経路に影響するため、同一のセンシング状態であっても、装着誤差によって得られる音響データが変化してしまう。この変化がシステムが認識すべきジェスチャ時などに生じる音響データの変化と混在することで誤認識

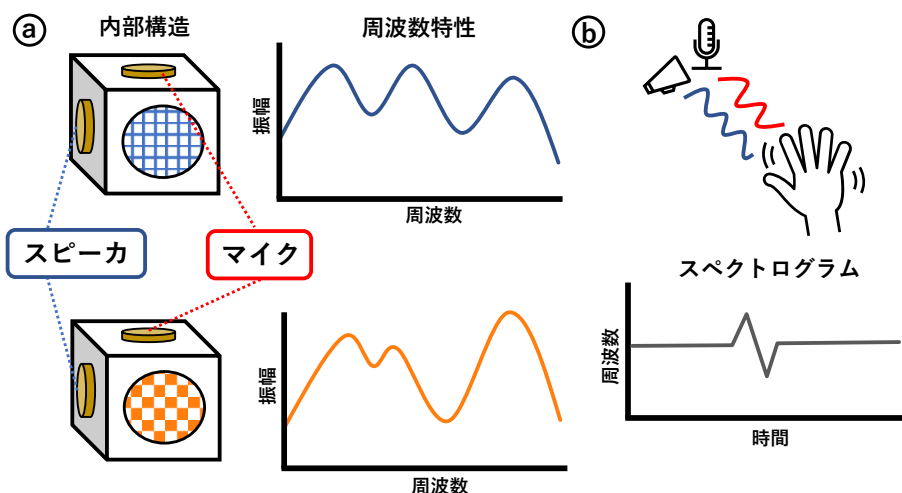


図 1.3: AAS の利用イメージ. a: 内部構造の異なるブロックの音響伝播特性を測定することでブロックの識別を行う手法, b: 手の動きで発生するドップラーシフトを測定し, ジェスチャとして利用する手法.

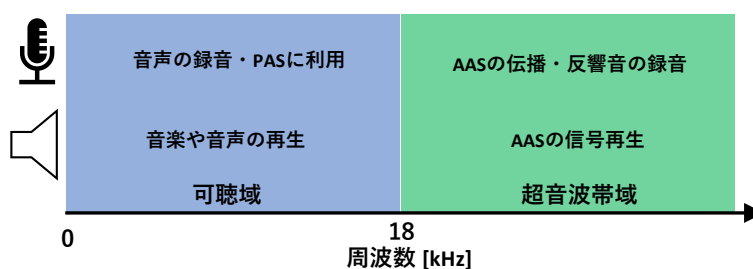


図 1.4: スピーカとマイクの利用例と利用周波数帯域. 再生・録音する周波数帯域によって用途を使い分けることが可能である.

の要因となる. これまでの AAS に関する研究では装着誤差に関する深い検討はされておらず, デバイス装着毎にデータを再測定する方法や装着誤差を含んだままのデータが認識に利用されてきた. したがって, 装着誤差を低減する手法が実現できればデータの再測定が不要になり, 認識システムの高精度化に貢献できる. 後者に関して, これまでの AAS はブロックや人体などの形状が安定しているセンシング対象 [1, 2, 3] に適用され, タッチインタフェースやジェスチャ入力を実現する手法が多く研究されてきた. 一方で, 布などの形状が不安定なセンシング対象への調査が行われていなかった. これは, 布などの形状が不安定な物体は信号伝播時の減衰が大きく, また, 形状が不安定なために信号の変化も不安定であるという課題が存在するためである. しかし, 布で構成された物体は衣服や帽子などが挙げられ, これらはユーザが日常生活で着用することが非常に多い. したがって, ウェアラブルデバイスの機能をこれらの衣服や帽子に埋め込むことができれば, ウェアラブルデバイスの多様化が期待できる.





く存在し、上記の研究分野のなかで最も研究報告例が多い。したがって、本研究にてジェスチャ入力手法での DAAS の有効性を示すことができれば、多くの既存研究に適用できる可能性が高いと考え、ジェスチャ入力手法を本研究の調査対象とした。

### 1.2.3 研究の貢献

本研究では、これまでの AAS では未解決であった装着誤差と不安定なセンシング対象への AAS の適用に関する課題を解消した DAAS を確立するために、各課題に対してそれぞれ以下の研究テーマに取り組んだ。

#### 研究テーマ 1：動的信号補正による装着誤差の低減

これまでの AAS にはデバイスの着脱やユーザ動作によって装着誤差が生じ、誤認識の要因となる課題があった。その解決方策として測定信号を動的に補正し、装着誤差を低減させる手法を考案した。本手法の有効性を評価するために、イヤホン型のウェアラブルデバイスであるヒアラブルデバイスのための頭部状態認識手法に本手法を適用し、調査を行った。実験では、11名の参加者より21種類の頭部状態のデータの収集を行った。実験の結果、参加者ごとの個人分類モデルによる6種類の頭部状態認識の結果、未補正下での平均認識精度が74.4% (F値)であったのに対して、信号補正を行うことで平均認識精度が90.0% (F値)まで改善されることを示した。

#### 研究テーマ 2：ノイズ環境での AAS の動的利用

布などの不安定なセンシング対象への AAS のみの適用は、信号の減衰や不安定な音響変化が起こるため実用的な認識性能を得られない。そのため、不安定なセンシング対象への AAS の適用は調査されていなかった。その解決方策として PAS を併用し、ノイズ環境での AAS の動的利用という新しい AAS の利用手法を考案した。例えば、布に対してタッチやスワイプを行うと衣擦れ音が発生するため、それぞれの音響データを解析することでタッチジェスチャとして認識することが可能であると考えられる。しかし、PAS は周囲のノイズの影響を受けやすいという課題がある。一方で、AAS は測定信号に超音波帯域 (18kHz~) を利用することで、日常生活で発生しやすい可聴域のノイズによる影響を避けることができる。本手法の有効性を評価するために、衣服上のタッチインタフェース構築手法に本手法を適用し、調査を行った。実験では、12名の参加者より11種類の前腕部分でのタッチジェスチャのデータの収集を行った。実験の結果、ノイズなし環境での5種類のタッチジェスチャ認識精度が95.9% (F値)であることを確認した。また、AAS の動的利用によってノイズ環境での精度が PAS のみの精度と比較して平均で3.6ポイント改善できることを示した。

上記のそれぞれの研究テーマの成果により、DAAS はこれまでの AAS に対して以下の2点の拡張に貢献した。

**研究テーマ 1 の貢献** AAS を用いた認識システムに適用可能な装着誤差低減手法を示し、認識システムの高精度化に貢献した。

研究テーマ2の貢献 ウェアラブルの多様化が期待される布への AAS の適用可能手法を示し、AAS の適用範囲の拡大に貢献した。

### 1.3 本論文の構成

本論文の構成は以下の通りである。

- 1章 研究の背景，研究内容，本研究の貢献を説明する。
- 2章 動的信号補正による装着誤差低減手法を外耳道伝達関数を用いた頭部状態認識手法に適用し，有効性，課題ならびに議論を述べる。
- 3章 ノイズ環境での AAS の動的利用手法を音響センシングを用いた衣服インタフェース構築手法に適用し，有効性，課題ならびに議論を述べる。
- 4章 本研究の結論および今後の検討課題を述べる。

## 第2章 動的信号補正による装着誤差の低減

本章では動的信号補正による装着誤差低減手法に関する研究の詳細を述べる。本研究では外耳道伝達関数を用いた頭部状態認識手法に本手法を適用し、有効性を評価した。本章の構成は以下の通りである。2.1節にて、外耳道伝達関数を用いた頭部状態認識手法の研究背景および本研究の成果を述べる。2.2節にて関連研究との比較・差異を述べる。2.3節にて本手法の詳細、2.4節にてハードウェア・ソフトウェアの実装の詳細を述べる。2.5節にて評価実験の設計・手順・結果を述べ、2.6節にて本研究の議論を述べる。最後に、2.7節にてまとめを述べる。

### 2.1 外耳道伝達関数を用いた頭部状態認識手法

近年のIT技術の進歩により、スマートウォッチなどの、身につけながら様々な機能を手軽に利用できるウェアラブルデバイスが普及してきた。その中で、注目されているウェアラブルデバイスとしてヒアラブルデバイスが挙げられる。ヒアラブルデバイスとはイヤホン型のデバイスで、イヤホン本来の機能である音楽鑑賞だけでなく、装着者の個人認証やライフログ、バイタルセンシングなどが可能となるデバイスのことを指す。例えば、スマートウォッチでしか利用できなかった機能がヒアラブルデバイスで利用できれば、スマートウォッチを利用しない選択肢をユーザに提示することが可能となる。また、ヒアラブルデバイスはスマートウォッチなどの他のウェアラブルデバイスと異なり、装着者のみに音声情報を提示できるため、自動翻訳や音声アシスタント、音声経路案内などのヒアラブルデバイス特有の機能が提供できる。実際に市販されているヒアラブルデバイスに、自動翻訳機能や音声アシスタント機能を搭載したデバイス [38, 39] も存在し、今後より多くのデバイスに様々な機能が実装されると考えられる。これらの例のように、ヒアラブルデバイスの多機能化が進むことで、これまでウェアラブルデバイスを利用していなかったユーザに対して、ヒアラブルデバイスを介してウェアラブルデバイスのサービスを提供することが可能となる。

一方で、ヒアラブルデバイスは荷物を持った状態や作業時など、手が使用できない場面においても使用することが想定される。そのため、通話への応答や話者の切り替え、様々な機能のON/OFFなどのデバイス操作をハンズフリーで実現できる機能が求められている。ヒアラブルデバイスのハンズフリー操作を実現するために、首を動かす、目を閉じるなどの頭に関わる動き（頭部ジェスチャ）を利用したデバイス操作手法が研究されてきた。これらの研究では赤外線センサ、気圧センサ、電位センサ等が使われてきた [40, 41, 42]。赤外線センサは認識できるジェスチャの種類が少なく、気圧センサは急激に気圧が変化するエレベータなどの環境下で認識精度が低下する。また、電位センサを含めたこれらのセンサは市販製品には搭載されておらず、新たにセンサを組み込む必要があるため、導入コストがかかる。

本研究では首、顎、顔などの状態（頭部状態）によって、外耳道の形状が変化することに着目し、音響信号により頭部状態を認識する手法を示す。具体的には、測定信号を用いたインパルス応答測定手法により外耳道内のインパルス応答を求め、そのフーリエ変換により外耳道伝達関数を求める。頭部状態の変化にともなう外耳道の形状変化によって、得られる外耳道伝達関数も変化する。各頭部状態の外耳道伝達関数を収集し、機械学習を用いて頭部状態の認識が可能な分類モデルを作成する。分類モデルを用いて現在の頭部状態の認識を行い、ヒアラブルデバイスへの入力コマンドとして利用することで、デバイスへのハンズフリー操作を実現する。本手法では超音波帯域の音を利用するため、外耳道内の音響信号を録音する内向きマイクが超音波帯域の録音に対応している必要がある。現在、超音波録音に対応した内向きマイクを搭載した市販製品は販売されていないが、ノイズキャンセリングなどのために可聴域の音響信号が録音可能な製品は販売されている [38, 43, 44]。したがって、超音波対応マイクを利用する本手法はヒアラブルデバイスと親和性が高く、ノイズキャンセリングなどの他の用途とも併用できる利点がある。

一方、ヒアラブルデバイスの利用が想定される場面においても、デバイスの着脱やユーザ動作によって装着誤差が生じるため、提案する頭部状態認識手法でも同様に認識精度の低下要因となると考えられる。その解決方策として、本研究では測定信号を動的に補正し、装着誤差を低減する手法を考案した。本手法の有効性を評価するために頭部状態認識手法に適用し、評価実験を行った。本研究の成果は以下のとおりである。

- ヒアラブルデバイスへのハンズフリー入力を可能とする音響信号を用いた頭部状態認識手法を考案した。
- 参加者ごとのデータを用いた個人分類モデルにて6種類の頭部状態認識実験を行った。実験の結果、認識システムの精度は未補正信号による認識実験にて平均74.4% (F値)、補正信号による認識実験にて平均90.0% (F値)であることを確認した。
- 動的信号補正手法によってデバイス装着時の装着誤差の低減を実現し、認識精度の向上に貢献した。

## 2.2 関連研究

本節では外耳道を含む人体の様々な部位にAASを適用した研究と、頭部関連動作の認識に関する研究・製品について説明し、本研究との比較・差異を述べる。

### 2.2.1 人体へAASを適用した研究

人体へ音波を伝播・反響させ、その応答を解析することでユーザの動作や個人性を認識する研究は広く行われてきた [23, 24, 25, 27, 34, 35, 26, 36, 37]。例えば Tan ら [45] は、スマートフォンから発する音響信号を用いて発話による口周辺部の動きを読み取り、口の動きによる個人認証システムを提案している。Watanabe ら [1] は腕と足にコンタクトスピーカとコンタクトマイクを装着し、人体内部を伝播する音の変化からユーザの状態や動作を

認識することに成功している。Mujibiya ら [2] は皮膚を伝播する超音波を用いて体接触やハンドジェスチャの認識を行った。Takemura ら [46] は骨伝導マイクを用いて指のタップ位置検出と肘の角度推定を行っている。また、柿澤ら [28] はコンタクトスピーカとコンタクトマイクを用いて手首部分の音響伝播特性を取得し、個人認証を行っている。

上記の研究はデバイスの装着状態に関わらず同一の測定信号を再生し、データが測定されているため、装着誤差を含んだデータを用いて機械学習が行われている。筆者の知る限り、測定信号を補正することで装着誤差を低減した後にデータ測定を行い、認識精度の改善を目指した研究は行われておらず、本研究にて初めて調査を行った。

### 2.2.2 外耳道インパルス応答の測定と応用に関する研究

外耳道のインパルス応答や伝達関数を取得する試みは外耳道の音響特性の調査と個人認証システムへの応用のために調査されてきた。Hiipakka ら [47] は外耳道の音響特性などを調査するために、マイク内蔵イヤホンを用いて Swept-Sine 信号 [48] を測定信号とした外耳道インパルス応答測定を行っている。Akkermans ら [49] はプローブ信号を用いて外耳道インパルス応答を測定し、伝達関数より外耳道の個人性を見出している。更に Arakawa ら [50, 51] は Maximum Length Sequence 法 (MLS 法) [52] を用いて外耳道の個人性を詳しく調査している。Yang ら [53] は同様の手法を用いてヒアラブルデバイスのための断続的なユーザ認証システムを提案している。また、Wang ら [54] はユーザの口の動きによって生じる外耳道の変形と、それに伴う音響特性の変化を個人認証に応用したシステムを提案している。

本研究では Hiipakka らと同じ Swept-Sine 信号を用いてインパルス応答を測定しているが、これらの既存研究では可聴域 ( $\sim 18$  kHz) のインパルス応答に着目している。本研究では、実用性のために、ユーザに聞こえない超音波領域に帯域制限したインパルス応答の測定を行った。また、既存研究は外耳道音響特性の個人性に着目しているが、頭部状態の変化に伴う音響特性の変化には着目されていなかった。本研究では、頭部状態の変化による外耳道音響特性の変化を利用し、現在の頭部状態を推定する手法を考案した。更に、インパルス応答を基に測定信号を補正することでデバイスの装着誤差を低減している。

### 2.2.3 頭部関連動作の認識に関する研究・製品

ヒアラブルデバイスを用いた静的な顔や首の状態認識 (頭部状態認識) や、一連の動作の認識 (頭部動作認識) に関する研究はこれまで多く行われてきた。表 2.1 に本手法と頭部状態および頭部動作認識に関する研究の比較をまとめた。Denys ら [42] は電位センサ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。Ando ら [41] は、顔関連の動作によ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。更に、4 段階の口の開け幅を 87.5% の精度で認識することに成功している。Verma ら [55] はイヤホン内蔵の加速度/慣性センサを用いて 32 種類の表情の認識を 89.9% の精度で認識することに成功している。また、Choi ら [56] は光センサの

1種である脈拍センサと、加速度センサを搭載したイヤホン型デバイスを用いて感情を表す7種類の表情の認識を93.5%の精度で認識することに成功している。その他に、Taniguchiら[40]はLEDとフォトランジスタを使って舌の特定の動きを外耳道の変形から認識し、音楽プレー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[57]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、DashPro[43]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ー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。

上述した既存研究・製品は電位、気圧、慣性、光センサを用いて頭部状態および頭部動作の認識を行っている。気圧センサを用いる手法では、高い精度で動作の認識を行なっているが、飛行機やエレベータなどの急激に気圧が変化する環境で認識精度が落ちる可能性がある。光センサを用いた手法では太陽光などの強い外光の影響を受けてしまう課題がある。また、電位センサを含めたこれらのセンサは一般的なヒアラブルデバイスには搭載されておらず追加のセンサが必要である。一方、慣性センサは搭載されているヒアラブルデバイスも存在するが、慣性センサ単体での利用はユーザの動きの影響を受ける課題がある。Choiら[56]は脈拍センサを用いることで、その課題を低減しているが、2つのセンサを利用するコストが生じる。これらのセンサを他用途に利用する手法や製品例は存在するため、実装コストの比較は困難であるが、本手法が利用する外耳道内部の超音波を取得できるマイクは外耳道内のノイズキャンセリング（騒音抑制）やイコライゼーション（音源調整）[58]の用途とも併用可能である。また、スピーカも音楽鑑賞や通話において必須の素子である。したがって、本手法はヒアラブルデバイスとの親和性が高く、これらの素子をジェスチャ入力に利用することでデバイス操作手法にも応用が可能となる。

表 2.1: 頭部関連動作の認識に関する研究の比較。

	本手法	Denys [42]	Ando [41]	Verma [55]	Choi [56]
利用素子	スピーカ マイク	電位センサ	気圧センサ	加速度センサ 慣性センサ	加速度センサ 脈拍センサ
認識対象	頭部状態	頭部動作	頭部動作	頭部動作	頭部状態
認識精度	90.0%	90.0%	87.6%	89.9%	93.5%
認識数	6種	5種	11種	32種	7種
他用途への 利用例	音源再生 騒音抑制 音源調整	—	—	ライフログ	ライフログ 装着検知 心拍測定

## 2.3 提案手法

本システム全体の認識の流れを図 2.1 にまとめた。ユーザは外耳道内部の音を録音できるように設計されたマイク内蔵イヤホンを装着する。システムはイヤホンから測定信号を再生し、外耳道の周波数応答である外耳道伝達関数を測定する。ユーザが顔の向きや表情を

変化させることで外耳道の形状が変化するため、外耳道側面や鼓膜からの反響特性が変化する。そのため、得られる外耳道伝達関数も変化する。図 2.2 に同じ実験参加者の閉口時と開口時の外耳道伝達関数を示す。外耳道伝達関数からそれぞれの頭部状態の特徴量を抽出し機械学習を行うことで、頭部状態を認識する分類モデルを作成する。本研究では、特徴量に LFCC (Linear-Frequency Cepstrum Coefficients) [59] を用いた。また、分類モデルにはサポートベクタマシン (SVM: Support Vector Machine) を用いた。本研究で用いる測定信号は超音波領域に帯域制限することで、ユーザの聴覚への影響を少なくする。また実環境には可聴域の音がノイズとして多く存在する。可聴域の測定信号を用いる場合、測定信号とノイズが混同されてしまい認識精度に大きく影響してしまう。一方で、超音波を測定信号とすることでノイズの影響が減らせる利点がある。

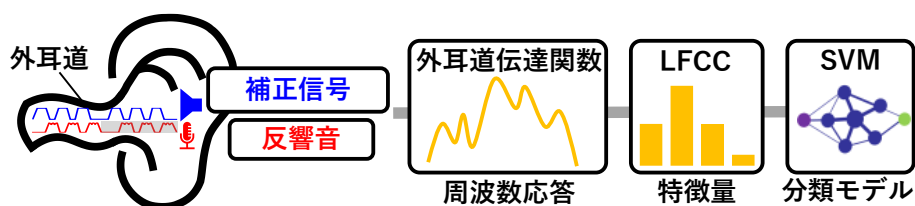


図 2.1: 頭部状態認識システム構成。

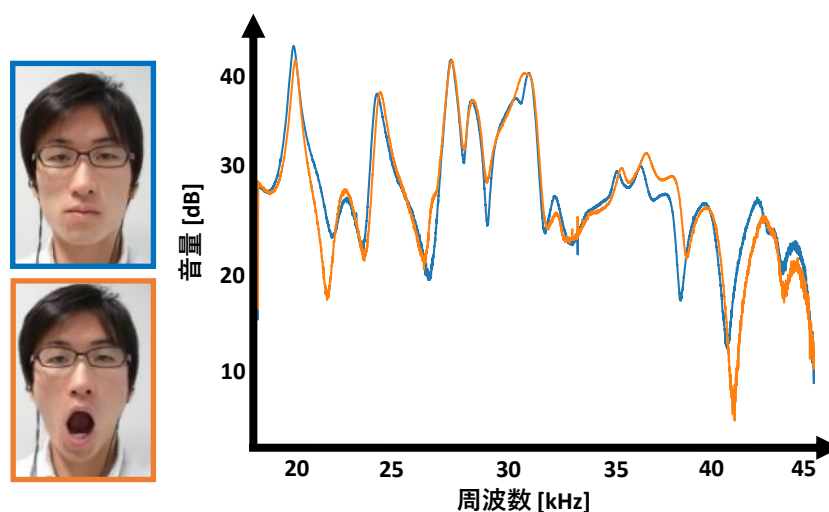


図 2.2: 表情による外耳道伝達関数の変化。

### 2.3.1 認識原理

図 2.3 に外耳道周辺部および側頭部の骨格/筋肉モデルをまとめた。図に示すように、下顎骨は外側翼突筋を介して外耳道の下側と接しており、祁ら [60] は顎運動時に下顎骨の動きと外耳道のひずみに相関関係があることを報告している。また、外耳道上部は側頭骨と接している。側頭骨に接する側頭筋周辺には表情筋が密集しているため、眼球や瞼、眉の運



動が外耳道の形状に影響すると考えられる。その他に、頭の方角を変化させる運動はデバイスの荷重方向を変化させ、デバイスと外耳道の相対位置に影響すると考えられる。よって本研究では、顎・眼・首周辺の運動を基に頭部状態を定義し、認識を行う。本研究では、どの頭部状態が本手法の認識に有効であるか調査するために、先行研究 [42] で測定された頭部状態を参考にしながら、顎・眼・首周辺の運動で外耳道の変形が大きいと考えられる 21 種類の頭部状態を定義した。図 2.4 に選出した頭部状態を示す。先行研究 [42] と共通の頭部状態は A, B, E, I, L, P-S の 9 種で、その他は新たに追加した頭部状態である。

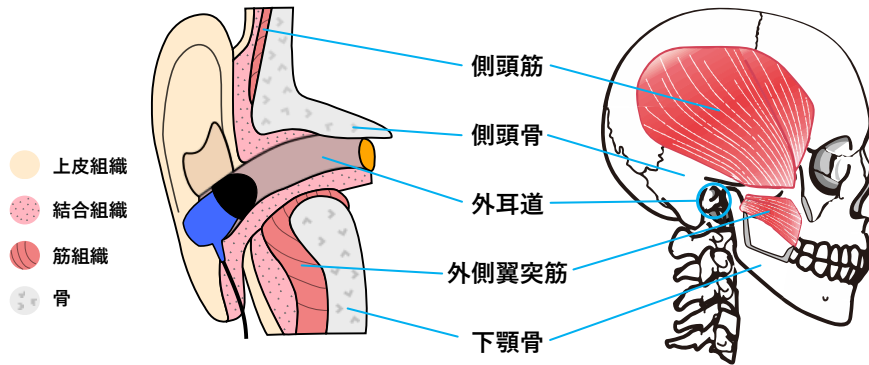


図 2.3: 外耳道周辺部および側頭部の骨格/筋肉モデル。



図 2.4: 測定する頭部状態。

### 2.3.2 外耳道伝達関数測定原理

測定信号を用いた伝達関数の測定原理を図 2.5 に示す。線形システム  $H$  の前に周波数特性  $S(k)$  ( $k$ : 離散周波数番号) を持った測定信号  $S$  と、後ろに逆特性  $1/S(k)$  を持つ逆フィルタ  $1/S$  を加えた測定系を考える。測定信号  $S(k)$  を特性  $H(k)$  を有する線形システムに入力すると、出力は  $H(k) \cdot S(k)$  となる。この出力を逆フィルタ  $1/S(k)$  に入力すると、出力は伝達関数  $H(k)$  となる。本研究では測定信号  $S$  に Swept-Sine 信号 [48] を利用し、外耳道を線形システム  $H$  と考え、外耳道伝達関数を測定する。ただし、伝達関数の測定中に頭部状態が変化しないと仮定することで、外耳道を線形システムとしている。



図 2.5: 伝達関数測定原理.

### 2.3.3 Swept-Sine 法

インパルスを時間軸上に引き伸ばした Swept-Sine 信号を測定信号として用いる手法を Swept-Sine 法 (SS 法) [48] と呼び、インパルス応答の測定に広く利用されている。周波数領域での Swept-Sine 信号  $SS(k)$  は以下のように表される。

$$SS(k) = \begin{cases} \exp(j\alpha k^2) & (0 < k \leq N/2) \\ \exp(-j\alpha(N-k)^2) & (N/2 < k \leq N) \end{cases}$$

ただし、 $\alpha = 4m\pi/N^2$ 、 $j$  は虚数、 $k$  は離散周波数番号、 $N$  は離散データ数、 $m$  はパルスの引き伸ばし係数である。

$SS(k)$  を逆フーリエ変換することで Swept-Sine 信号  $ss(n)$  を得る。また逆フィルタ  $SS^{-1}(k)$  は  $SS(k)$  の複素共役で求められる。インパルス応答は全ての周波数の音を含むため、測定信号発信時にユーザにも聞こえ、不快となりうる。そこで、本研究では測定信号の帯域制限を行う [61]。サンプリングレートを  $f_s$  として 0 から  $f_0$  Hz までの帯域制限を行うには  $0 \leq k \leq (f_0 N / f_s)$  と  $N - (f_0 N / f_s) \leq k \leq N$  の範囲の  $SS(k)$  の振幅値を 0 にすることで実現する。なお、帯域制限を行うことで逆フィルタとの畳み込みは完全なインパルスとはならず、本研究で測定する外耳道インパルス応答や外耳道伝達関数は厳密にはインパルス応答や伝達関数とはならないが、便宜上、本論文では外耳道インパルス応答・外耳道伝達関数と表記する。

## 円状畳み込みの原理を用いたインパルス応答測定法

円状畳み込みの原理を用いることで、測定信号長をインパルス応答の長さよりも長く設定すれば、インパルス応答を取得できることが知られている。Swept-Sine 信号 (図 2.6 (a)) を用いた測定手順は以下の通りである。

1. Swept-Sine 信号を同期加算の回数分だけ隙間なく並べ、1 周期分の無音区間を末尾に追加する (図 2.6 (b))。これが本手法で利用する測定信号となる。
2. 測定信号を再生し、その反響音を録音することで出力信号 (図 2.6 (c)) を取得する。
3. 出力信号を Swept-Sine 信号長で切り出し加算平均する (図 2.6 (d))。加算平均によってノイズ同士で打ち消されるため、高 SN 比の出力信号が得られる。
4. 加算平均した出力信号と逆フィルタ信号 (図 2.6 (e)) のフーリエ変換を要素ごとに掛け算する。

上記の信号処理を外耳道に対して適用することにより外耳道伝達関数を取得する (図 2.6 (f))。

### 測定信号の修正によるインパルス応答の補正

2.3.2 節の手順で得られる外耳道伝達関数はデバイスの着脱や時間経過に伴って生じる装着誤差によっても変化するため、同じ頭部状態でも測定ごとに外耳道伝達関数が異なるという問題がある (図 2.7)。これらの変化は頭部状態による外耳道伝達関数の変化と混同される可能性があり、認識精度向上のためには装着誤差は最小限にする必要がある。そこで閉口状態時に測定信号を補正することで、装着誤差を低減する手法を提案する。

外耳道伝達関数とその時の測定信号、出力信号の周波数特性をそれぞれ  $H(k)$ 、 $X(k)$ 、 $Y(k)$  ( $k$ : 離散周波数番号) とすると以下の関係が成り立つ。

$$X(k)H(k) = Y(k) \quad (2.1)$$

デバイスの装着誤差によって  $H(k)$  は変化する。装着誤差を含んだ場合の外耳道伝達関数を  $H'(k)$  とする。ここで、測定信号  $X(k)$  による新たな出力を  $Y'(k)$  とすると同様に

$$X(k)H'(k) = Y'(k) \quad (2.2)$$

と表せる。この時、測定信号に補正を加えて出力を  $Y(k)$  にするような補正測定信号  $X'(k)$  は式 (2.1) (2.2) より

$$X'(k) = \frac{Y(k)}{H'(k)} = \frac{Y(k)}{Y'(k)} X(k) \quad (2.3)$$

と表せる。 $Y(k)$  を一意に設定し、閉口状態時に式 (2.3) による信号の補正を与えることで装着誤差が低減でき、頭部状態の変化による外耳道伝達関数の変化を効率的に測定できると考えられる。本研究では事前に各実験参加者から取得した閉口状態での外耳道伝達関数の加算平均を  $Y(k)$  として利用する。図 2.8 が実際にデバイスの装着誤差を補正し、得られ

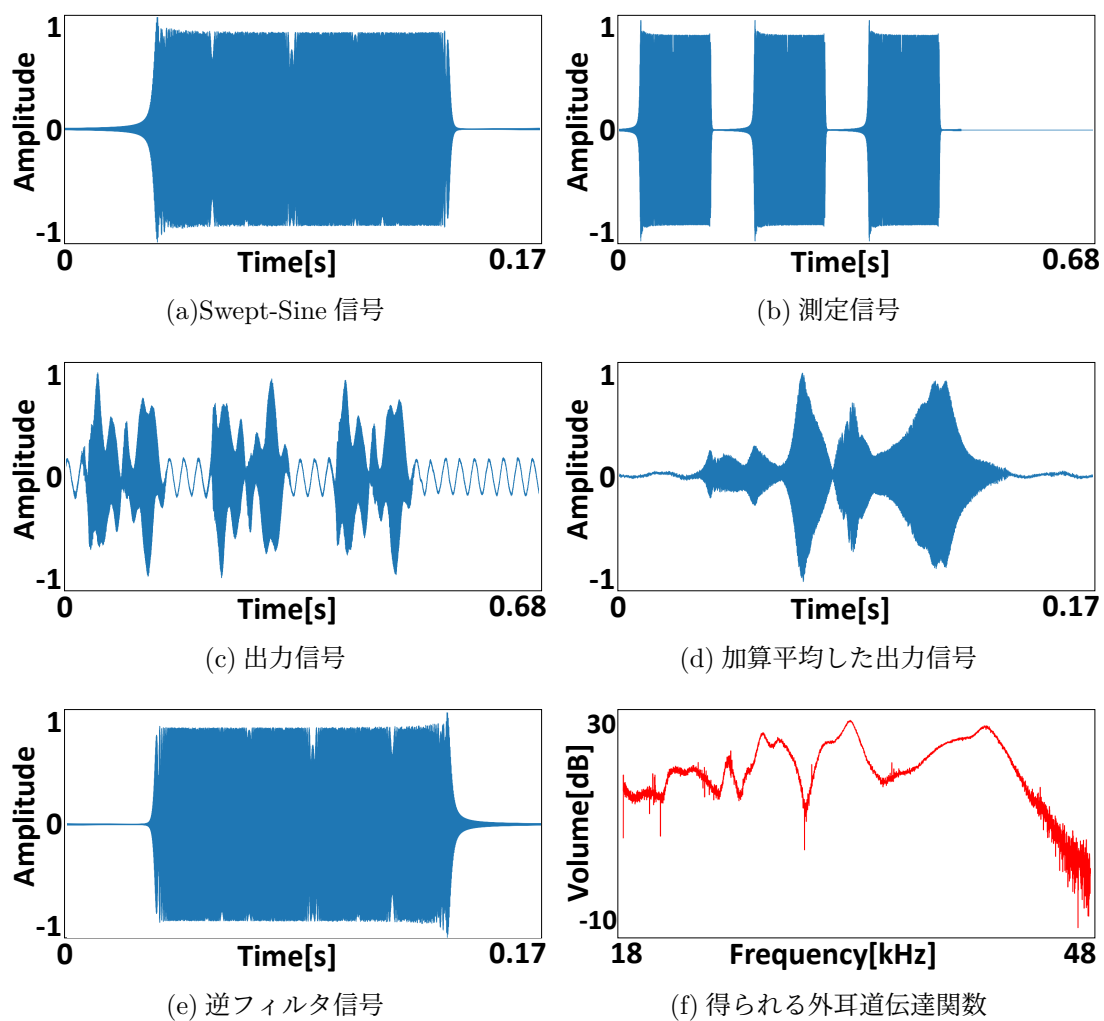


図 2.6: 円状畳み込みの原理を用いたインパルス応答測定.

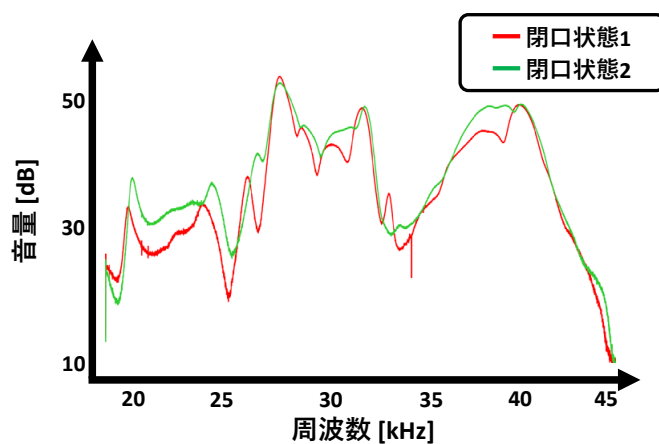


図 2.7: デバイス着脱による外耳道伝達関数の変化 (未補正信号による測定時).

た外耳道伝達関数である。図 2.7 と図 2.8 を比較すると、同じ頭部状態の時に得られる外耳道伝達関数が近くなっており、補正されていることが確認できる。

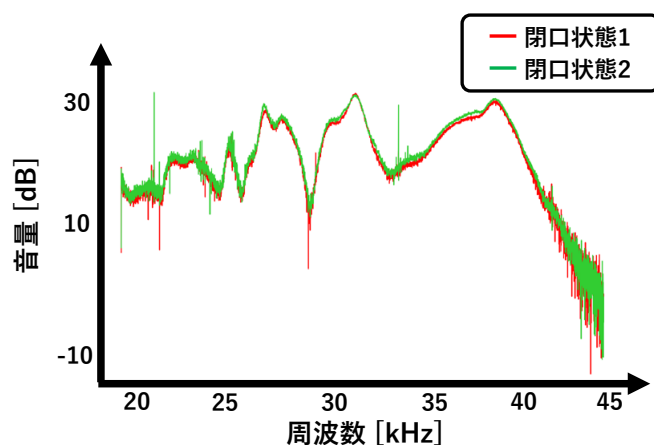


図 2.8: デバイス着脱による外耳道伝達関数の変化 (補正信号による測定時).

#### 2.3.4 特徴量抽出および分類モデル

取得した音声信号をそのまま機械学習に使用すると、特徴量の次元数が多く学習効率が悪いため、特徴量抽出を行う。本研究では特徴量として人間の音声認識によく用いられるメル周波数ケプストラム係数 (MFCC: Mel-Frequency Cepstrum Coefficients) の線形バージョンである LFCC (Linear-Frequency Cepstrum Coefficients) [59] を用いる。これらの違いは、音声データの周波数スペクトルの対数表示に対して、MFCC ではメル周波数領域において等間隔なメルフィルタバンクを適用するが、LFCC では周波数領域において等間隔な線形フィルタバンクを適用する。5 種の頭部状態の認識予備調査による精度比較の結果、LFCC の認識精度が高いことを確認したため、LFCC を採用した (表 2.2)。図 2.9 に実際に利用した線形フィルタバンクを示す。本研究では LFCC を特徴量として最大値、最小値で正規化を行い機械学習に利用し、分類モデルを作成する。次元数を決定するフィルタ数は 13 分割フィルタバンク (12 次元) と 20 分割フィルタバンク (19 次元) から精度の高かった後者を選定した (表 2.2)。フィルタ処理によって得られる 20 次元の LFCC から直流成分である 1 次元目を除いた 19 次元を両耳から抽出するため、特徴量の次元数は計 38 次元となる。

分類モデルにはサポートベクタマシン (SVM: Support Vector Machine) とランダムフォレスト (RF: Random Forest) の 2 種類から 21 種の頭部状態の認識予備調査にて精度の高かった SVM を選定した (SVM: 47.2%, RF: 27.4%)。分類モデルを基に現在の外耳道伝達関数から頭部状態を予測する。

表 2.2: 特徴量とフィルタバンクの精度に関する予備調査.

	13 分割フィルタバンク	20 分割フィルタバンク
LFCC	91.1%	92.2%
MFCC	84.1%	91.7%

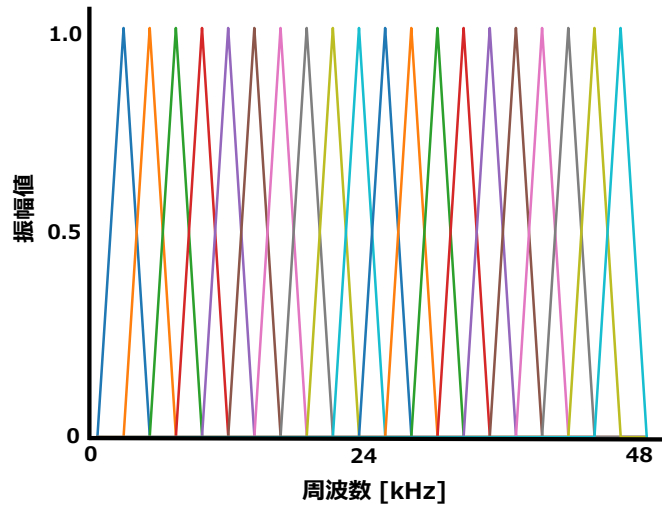


図 2.9: 線形 20 分割フィルタバンク.

## 2.4 実装

本節では提案システムの実装について述べる. 提案システムは外耳道内部に音声信号を再生し, その反響音を取得するためのハードウェアの実装と頭部状態認識アルゴリズムのソフトウェアの実装に分かれる. 以下でそれぞれの実装の詳細を述べる.

### 2.4.1 外耳道反響音取得デバイス

外耳道の反響音を取得するために市販のイヤホンと小型 MEMS マイクを用いてデバイスを作製した. 市販のマイク内蔵型イヤホン [38, 44, 62] を用いることも考えられるが, これらのデバイス内蔵マイクの周波数特性では超音波領域の音を取得できないため, 本研究では自作のデバイスを用いた. 外部の騒音の影響を小さくし, 反響音を最大限に取得するためにイヤホンは耳穴に挿入するタイプのカナル型イヤホンを使用した. また超音波を測定信号に使用するため, イヤホンとマイクの対応周波数帯域も考慮し, イヤホンに Xiaomi の Pro HD ハイレゾ対応 (再生周波数帯域:20–48 kHz), マイクに knowles の SPU0410LR5H (録音周波数帯域:100–80kHz) を用いた. マイクは反響音を効率よく録音できるようにイヤホンの音声再生部分に密着するように取り付けた (図 2.10). また, これらの作業を行うにあたってイヤホンの外耳道挿入部分を取り除いたため, 新たに 3D プリンタでパーツを作り直した. 音声入出力時の DA/AD 変換にオーディオインタフェース (Komplete Audio 6) を

使用し、マイクの電源に Sunhayato の DK-910 を使用した。また、PC は Thinkpad X270 を使用した。再生、録音共にサンプリングレートは 96 kHz で行う。再生・録音デバイスの構成を図 2.11 にまとめた。

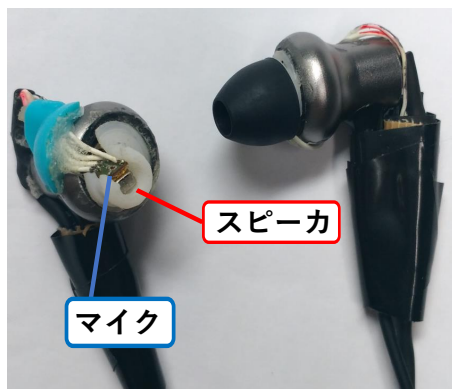


図 2.10: 実験デバイス.

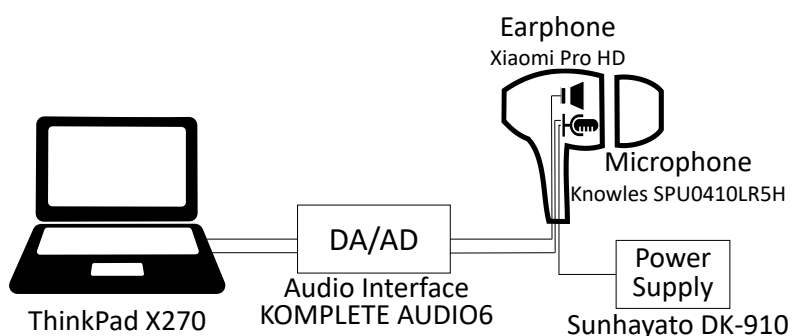


図 2.11: デバイス構成.

#### 2.4.2 頭部状態認識アルゴリズム

本研究で使用する Swept-Sine 信号は 18 kHz 以下を帯域制限した、18 kHz から 48 kHz に周波数が上昇していくアップスイープ信号 (16,384 サンプル) を使用した。また、SN 比向上のために同期加算を行う必要がある。本研究で使用する Swept-Sine 信号が 1 周期で約 0.17 秒であり、分類できる時間間隔と SN 比のバランスを考慮し、同期加算回数は 3 回とした。無音区間を追加するため、測定信号は全体で約 0.68 秒となる。Swept-Sine 信号を連続で再生する場合、信号の繋ぎ目付近での周波数の急激な変化によってクリックノイズが発生する。クリックノイズの音量を抑えるために、Swept-Sine 信号の開始・終了部の 1,638 サンプル分の音量を徐々に変化させるフェードイン・アウト処理を行った。実験では測定信号を連続で 7 回再生し、1 回の測定で 7 個の外耳道伝達関数を取得した。得られる外耳道伝達関数は Swept-Sine 信号長と同じ 16,384 サンプルとなるので LFCC 抽出時のフーリエ変換は

16,384 サンプルで行い、線形フィルタバンクは 20 分割で行った。学習に用いる特徴量は直流成分である 1 次元目を除いた 19 次元を特徴量とし、外耳道伝達関数は両耳から取得するため、特徴量の次元数は計 38 次元となる。測定信号の補正時に、一意に定める必要のある  $Y(k)$  は事前に実験参加者から取得した閉口状態での外耳道伝達関数の平均値とした。外耳道伝達関数の測定・特徴量抽出・機械学習は Python 3.6 で実装した。

## 2.5 評価実験

本節では評価実験のデータ測定手順および調査結果を述べる。信号の動的補正による装着誤差の低減および認識精度への影響を調査するために (i) 未補正信号による認識実験と (ii) 補正信号による認識実験の 2 種類の評価実験を行った。また、実用を考慮した認識頭部状態の選定と精度改善の評価実験を行った。以下で各実験の詳細を述べる。

### 2.5.1 実験協力者および実験環境

実験は 23–26 歳の 11 名のボランティア (男:10 名, 女:1 名) に参加してもらった。実験は筆者の所属する研究室の学生部屋で行い、参加者には座った状態でデータの測定を行ってもらった。空調の停止や周りの人に静かにしてもらおうなどの騒音対策は行っていない。

### 2.5.2 分類モデルの評価方法

既存研究 [53, 54] で調査されているように、ユーザの外耳道音響特性および頭部動作にともなう外耳道の変形は個人認証システムに応用される程、高い個人性を示す。そのため、他者のデータを基に作成した分類モデルによって自身のデータの認識を行う一般分類モデルによる頭部状態の認識は困難であると考えられる。したがって、本研究では実験参加者ごとの測定データを用いて、実験参加者ごとの頭部状態を認識する個人分類モデルを作成し、本手法の認識性能を評価する。分類モデルの作成時に、デバイスの装着具合が同じであるセット内データを訓練データと検証データの両データに含めてしまうと、実利用時にユーザはデバイス装着の度にデータ測定が必要となる。したがって、本研究ではセット間ごとの交差検証を行い、装着具合の影響を考慮した分類モデルでの評価を行う。訓練データによる分類モデルのチューニングは 5 分割の交差検証を行い、F 値が最も高かったパラメータをハイパーパラメータとした。カーネル関数は線形カーネルと RBF カーネルの 2 種類を使用した。

### 2.5.3 未補正信号による認識実験

筆者の知る限り、未補正の音響信号を用いた頭部状態認識に関する調査はこれまで行われていない。本手法の有効性を評価するために未補正信号による頭部状態認識実験を行う。



## 実験手順

最初に、図 2.4 に示す頭部状態を参加者に図と口頭を用いて説明し、全ての頭部状態を再現できるようにする。そしてデバイスを両耳に装着してもらう。参加者によって安定する装着方法が異なったため、装着方法は通常の掛け方（普通掛け）と耳の後ろからコードを通し耳上部から装着する方法（耳掛け）の 2 種類から、デバイスが安定している装着方法を参加者に選択してもらった（図 2.12）。耳掛けはケーブルから伝わる振動を低減できる装着方法であり、一部のユーザは日常的に利用している。したがって、音楽鑑賞などの通常利用において影響はない。

測定手順は、デバイスを装着した状態でディスプレイに表示された頭部状態の写真を再現、維持してもらう。維持状態を確認した後に未補正信号を再生し、録音を行う。この測定を全ての頭部状態でランダムな順番で行ったものを 1 セットとし、デバイスの着脱をセットごとに行った。実験は 3 日間に分けて行い、測定 1 日目に 4 セット、測定 2 日目に 4 セット、同様に測定 3 日目に 4 セットのデータ測定を行い、各参加者から 1,764 個分の外耳道伝達関数（7 外耳道伝達関数 × 21 状態 × 12 セット）を集めた。本実験で使用した測定信号の音量は、使用するイヤホンで音楽を聴いて、適切であった音量レベルに設定した。なお、本実験で利用している未補正信号とは 2.4.2 節で述べた信号長 16,384 サンプルのアップスイープ信号（18 kHz–48 kHz）である。

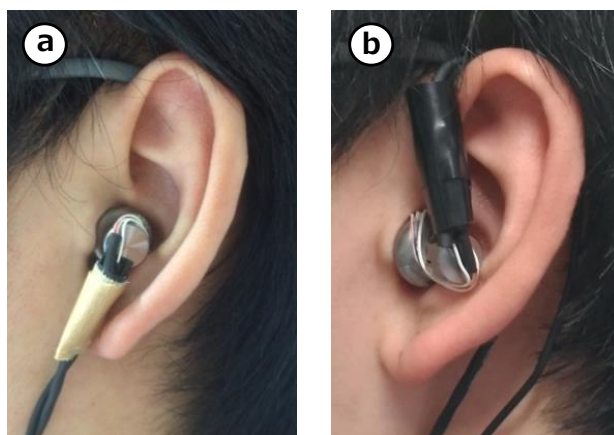


図 2.12: デバイス装着方法. a: 普通掛け, b: 耳掛け.

## 実験結果

未補正信号を用いて取得したデータを使用して認識精度を確認した。未補正信号による認識実験での実験参加者の認識精度の平均を表 2.3 (a) にまとめた。全体の平均は F 値 40.2%であった。未補正信号による認識実験で最も認識精度が高い参加者は F 値 62.0%, 低い参加者は F 値 26.4%であった。未補正信号による認識実験での各頭部状態毎の認識精度を図 2.13a にまとめた。最も認識精度の高い頭部状態は「顎を右にずらす」で F 値 71.0%, 低い頭部状態は「両目を閉じる」で F 値 19.4%となった。

## 2.5.4 補正信号による認識実験

### 実験手順

補正信号による認識実験では、デバイスを装着した状態で閉口状態を維持してもらう。その間の外耳道伝達関数を取得し、2.3.3節で説明した信号補正処理を行い補正信号を得る。その後、ディスプレイに表示された頭部状態の写真を再現、維持してもらう。維持状態を確認したら補正信号を再生し、録音を行う。この測定を全ての頭部状態でランダムな順番で行ったものを1セットとし、デバイスの着脱をセットごとに行った。実験は3日間に分けて行い、測定1日目に3セット、測定2日目に3セット、同様に測定3日目に3セットのデータ測定を行い、各参加者から1,323個分の外耳道伝達関数（7外耳道伝達関数×21状態×9セット）を集めた。その他の実験環境および実験手順は未補正信号による認識実験と同様である。

### 実験結果

補正信号を用いて取得したデータを使用して認識精度を確認した。補正信号による認識実験での実験参加者の認識精度の平均を表2.3(b)にまとめた。全体の平均はF値62.5%であった。補正信号による認識実験で最も認識精度が高い参加者はF値83.6%、低い参加者はF値40.2%であった。未補正/補正信号による認識実験の認識精度に対してt検定を行った。その結果、精度向上において有意差が見られた( $t(10) = 3.760, p = 0.0024$ )。各頭部状態毎の認識精度と認識の分布を図2.13bにまとめた。最も認識精度の高い頭部状態は「上を向く」でF値93.7%、低い頭部状態は「舌をだす(開口)」でF値37.1%となった。

表2.4に各評価実験での認識精度の高い上位7種類の頭部状態をまとめた。実験の結果、信号補正の有無に関わらず、顎や首の動作を含む頭部状態の認識精度が高かった。まず、顎を左右にずらす頭部状態は下顎骨の左右方向への変動を生じさせる。図2.3からもわかるように下顎骨の頂点部(下顎頭)は外側翼突筋を介して外耳道下側と接している。したがって、左右方向の変動が外耳道の形状変化に大きく影響したため、得られた音響データにも変化が顕著に表れたと考えられる。首を動かす動作に関して、デバイスが有線であるために顔の方向が変わるとデバイスの荷重方向が変わる。荷重方向によってデバイスと外耳道の相対位置が変化することで音響データの顕著な変化が得られたと考えられる。

表2.5に各評価実験での認識精度の低い下位7種類の頭部状態をまとめた。実験の結果、信号補正の有無に関わらず、頬、目、舌、口の動作を含む頭部状態の認識精度が低かった。頬や目、舌の動作は表情筋に属する筋肉が伸展されるが、外耳道の周辺部に属していないため、外耳道変形に大きく影響することは少なく、得られる音響データの変化も小さくなってしまったと考えられる。一方、口を開ける動作は下顎骨が上下に動くため、外耳道の変形が大きく生じると考えられる。しかし、図2.13が示すように「B-口を開ける」と「M-舌を出す(開口)」のそれぞれを誤認識する傾向がある。これは舌を出す動作が外耳道変形に与える影響が小さく、結果として「口を開ける」と似た外耳道変形が生じてしまったことが原因であると考えられる。また、口を開ける動作は左右対称であるため、左右の外耳から得られる音響データに差異が小さくなることも精度低下の原因であると考えられる。

表 2.3: 全ての頭部状態での各参加者の認識精度 [%].

参加者	(a) 未補正信号認識実験			(b) 補正信号認識実験		
	適合率	再現率	F 値	適合率	再現率	F 値
P1 <sup>N</sup>	56.1	61.6	58.7	82.5	84.7	83.6
P2 <sup>N</sup>	25.7	37.1	30.4	39.1	41.3	40.2
P3 <sup>N</sup>	48.8	54.5	51.5	59.5	63.0	61.2
P4 <sup>N</sup>	40.9	49.0	44.6	47.2	50.3	48.7
P5 <sup>N</sup>	44.4	50.1	47.1	65.5	69.2	67.3
P6 <sup>N</sup>	48.2	54.0	50.9	51.2	54.5	52.8
P7 <sup>E</sup>	24.0	29.3	26.4	59.4	63.4	61.3
P8 <sup>N</sup>	34.3	40.8	37.3	62.6	69.0	65.6
P9 <sup>N</sup>	25.2	31.4	28.0	67.4	69.2	68.3
P10 <sup>N</sup>	60.3	63.8	62.0	51.2	54.5	52.8
P11 <sup>N</sup>	24.4	29.3	26.6	61.2	64.5	62.8
平均	37.7	43.1	40.2	60.8	64.2	62.5

<sup>N</sup> 普通掛け <sup>E</sup> 耳掛け

### 2.5.5 各頭部状態のアンケート評価

筆者を除く 10 名の参加者に、各頭部状態の再現性の困難さと人前で再現することの印象について、それぞれ 5 段階（1：非常に簡単-5：非常に難しい）と 4 段階（1：気にならない-4：やりたくない）のリッカート尺度を用いて回答してもらった。各項目の詳細な情報は付録 A に示している。各質問の結果を図 2.14 と図 2.15 にまとめた。縦軸の数値は選択肢の数値と一致している。また、実験の感想を記述式で回答してもらった。主な回答を以下にまとめた。

- 連続してジェスチャを測定すると気疲れする
- 口角を上げる動作が難しかった
- 顎をずらす動作の左右が混乱してしまう
- デバイスの固定がもっと安定していれば、動かしやすと感じた

再現性の困難さについてのアンケート結果より、ほとんどの頭部状態の再現性が比較的簡単（平均スコアが 3 以下）であることがわかった。一方で、「顎をずらす」頭部状態（C, D）は普通より少し難しい（両平均スコア：3.22）という結果になった。この要因としては記述回答にもあるように左右方向を混乱してしまった実験参加者がいたためであると考えている。本実験では、ディスプレイ上に表示された頭部状態を再現してもらい、データの測定を行ったため、このような左右の混乱が起きたと考えている。実際の利用環境を想定し

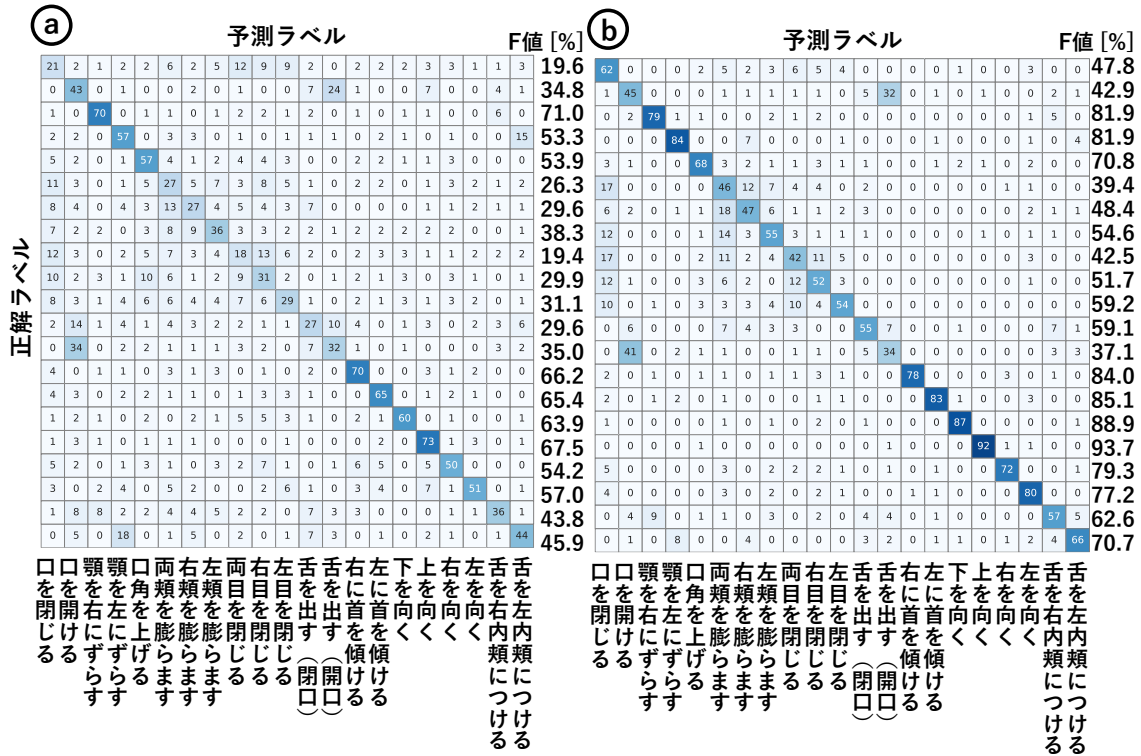


図 2.13: 混同行列. a: 未補正信号による認識実験, b: 補正信号による認識実験

表 2.4: 認識精度の高い頭部状態上位 7 種類 [%].

順位	頭部状態 (F 値)	
	(a) 未補正	(b) 補正
1	顎を右にずらす (71.0)	上向き (93.7)
2	上を向く (67.5)	下向き (88.9)
3	右に首を傾げる (66.2)	左首傾 (85.1)
4	左に首を傾げる (65.4)	右首傾 (84.0)
5	下を向く (63.9)	顎を右にずらす (81.92)
6	左を向く (57.0)	顎を左にずらす (81.91)
7	右を向く (54.2)	右を向く (79.3)

表 2.5: 認識精度の低い頭部状態下位 7 種類 [%].

順位	頭部状態 (F 値)	
	(a) 未補正	(b) 補正
15	口を開ける (34.8)	右目を閉じる (51.7)
16	右目を閉じる (29.9)	右頬を膨らます (48.4)
17	舌を出す (閉口) (29.6)	口を閉じる (47.8)
18	右頬を膨らます (29.6)	口を開ける (42.9)
19	両頬を膨らます (26.3)	両目を閉じる (42.5)
20	口を閉じる (19.6)	両頬を膨らます (39.4)
21	両目を閉じる (19.4)	舌を出す (開口) (37.1)

た場合、ユーザはコマンドの結果と対応するジェスチャを紐付けて記憶するため、ディスプレイからの情報による左右方向の混乱は減ると考えられる。

人前で再現することの印象についてのアンケート結果より、「舌を出す」動作を含む頭部状態 (K, L) はスコアが3を超えていた。アンケートは全員が日本人であり、舌を出す行為は悪意のあるイメージを持つ。そのため、舌を出す行為を人前で行う心理的ハードルは非常に高いと考えられる。

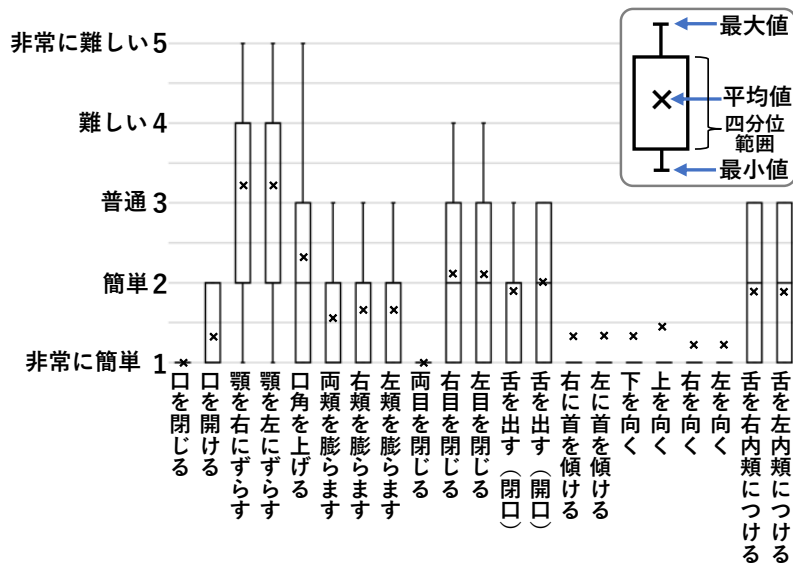


図 2.14: 質問 1 (再現性の困難さ) のアンケート結果.

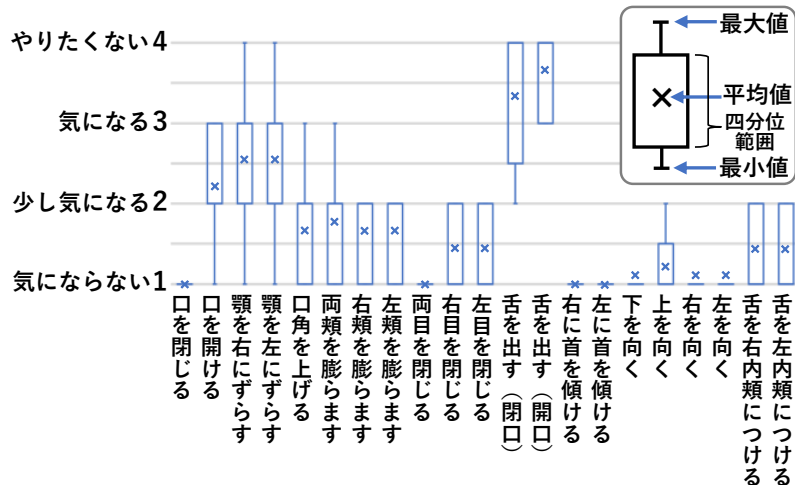


図 2.15: 質問 2 (人前で再現することの印象) のアンケート結果。

### 2.5.6 アプリケーションを考慮した頭部状態認識実験

#### 頭部状態の選定

上記の実験では、どの頭部状態が本手法に有効であるかを調査するために 21 種類の頭部状態を調査した。しかし、デバイスの操作や入力インターフェースとしての実用性を考えると認識精度が 62.5% (補正信号による認識実験) では十分とはいえない。そこで、想定する利用シーンと、その時に必要な頭部状態を絞ることで本手法の実用性を評価する。利用シーンとしては、ハンズフリーでの操作が求められる音楽プレーヤーやコンテンツリーダーの操作を想定する。5 種類の頭部状態をこれらのアプリケーションのコマンドとして選定し、認識精度の調査を行う。頭部状態の選定にあたっては以下の 3 条件を設定し、条件 1, 2 よりいくつかの頭部状態を選定候補から除外し、条件 3 を基に選定を行った。

**条件 1: 視界の安全性** 本手法が適用されるウェアラブルデバイスの利用環境は、歩行中や屋外が想定される。デバイス操作時に視線が逸れることや、視界が塞がってしまうことは歩行者や物体との衝突の危険性がある。したがって、視線が塞がってしまう目を閉じる頭部状態 (I-K) と、視線が逸れてしまう顔の向きを変える頭部状態 (P-S) を選定候補から除外した。

**条件 2: 公共の場での利用** 同様にウェアラブルデバイスの利用環境には電車内や公園などの公共の場が想定される。したがって、2.5.5 節のアンケート結果より、ユーザが人前で行うことが憚られる舌を出す頭部状態 (L, M) を選定候補から除外した。

**条件 3: 認識精度** 頭部状態を 5 種類に絞ることで認識精度が向上することは想定されるが、認識精度の低い頭部状態を選定してしまうと大幅な向上は見込めない。本実験では、表 2.4 より除外されていない頭部状態である顎をずらす頭部状態 (C, D) と、首を傾げる頭部状態 (N, O) を選定した。次に図 2.13 より、B-開口と M-舌開口はそれぞれを混同する傾向があることが分かる。この場合、一つの頭部状態のみを選定すること

で、それぞれの頭部状態を混同することがなくなるため、認識精度は向上すると考えられる。本実験では、条件2よりM-舌開口は候補から除外されているため、B-開口を5種類目の頭部状態とする。

以上の考察より、5種類「開口、右顎、左顎、右首傾、左首傾」の頭部状態をコマンドとして選定し、アプリケーションの非操作状態としてA-閉口状態を加えた6種類のデータから再度分類モデルを作成し、認識精度の調査を行った。

## 実験結果

各実験参加者毎の認識精度の平均を表2.6にまとめる。未補正信号による取得データでの認識精度は平均74.4%、補正信号による取得データでの認識精度は平均90.0%であった。未補正信号下ではP10のみ85%以上の認識精度を示した。一方で、補正信号下では多くの実験参加者(9/11名)が85%以上の認識精度を示した。未補正/補正信号による認識実験の認識精度に対してt検定を行った。その結果、精度向上において有意差が見られた( $t(10) = 4.587$ ,  $p = 0.0010$ )。また図2.16に6種類の各頭部状態の認識精度をまとめた。未補正信号下ではB-開口以外の頭部状態の認識精度が85%を下回ったのに対して、補正信号下では全ての頭部状態が85%以上の認識精度であった。以上の結果より、補正信号下では多くの実験参加者が各頭部状態をコマンドとして使える可能性が高いことが分かった。

表 2.6: 6種類の頭部状態での各実験参加者の認識精度 [%].

参加者	(a) 未補正実験			(b) 補正実験		
	P <sup>1</sup>	R <sup>2</sup>	F <sup>3</sup>	P	R	F
P1 <sup>N</sup>	73.4	79.6	76.4	96.5	96.0	96.2
P2 <sup>N</sup>	76.1	80.6	78.3	67.5	72.8	70.0
P3 <sup>N</sup>	79.3	84.7	81.9	94.2	96.0	95.1
P4 <sup>N</sup>	76.5	80.6	78.5	92.8	90.5	91.6
P5 <sup>N</sup>	70.7	74.2	72.4	88.2	90.5	89.3
P6 <sup>E</sup>	75.8	81.0	78.3	78.1	82.3	80.1
P7 <sup>N</sup>	70.5	76.0	73.1	99.5	99.5	99.5
P8 <sup>N</sup>	61.9	70.8	66.1	87.7	90.7	89.2
P9 <sup>N</sup>	60.3	68.8	64.3	92.1	92.9	92.5
P10 <sup>N</sup>	86.7	89.9	88.3	99.6	99.5	99.5
P11 <sup>N</sup>	55.8	61.5	58.5	86.1	87.3	86.7
平均	71.7	77.2	74.4	89.3	90.7	90.0

<sup>1</sup> 適合率   <sup>2</sup> 再現率   <sup>3</sup>F 値   <sup>N</sup> 普通掛け   <sup>E</sup> 耳掛け

正解ラベル	予測ラベル						F値 [%]		予測ラベル						F値 [%]
	64	5	7	8	9	5			93	1	1	0	2	1	
5	5	87	0	4	0	1	87.6	4	87	3	0	1	3	87.8	
5	5	4	83	0	5	0	83.4	1	3	90	1	2	1	90.2	
9	9	6	1	77	1	4	77.3	4	0	0	91	1	2	91.7	
10	10	1	4	2	79	0	79.8	7	2	1	0	89	0	89.0	
11	11	5	5	3	0	78	78.8	3	2	1	2	0	90	90.3	
	口を閉じる	口を開ける	顎を右にずらす	顎を左にずらす	右に首を傾げる	左に首を傾げる		口を閉じる	口を開ける	顎を右にずらす	顎を左にずらす	右に首を傾げる	左に首を傾げる		

図 2.16: 6 種類の頭部状態での混同行列. a: 未補正信号による認識実験, b: 補正信号による認識実験.

## 2.6 考察

本節では動的信号補正手法に関する議論およびに頭部状態認識手法に関する議論について述べる.

### 2.6.1 動的信号補正手法に関する議論

信号補正による装着誤差の低減

矢野らは外耳道音響特性による個人認証実験の際に音響測定に含まれる計測誤差の原因は

- 背景雑音や電氣的ノイズ等の雑音性誤差
- イヤホンの装着具合等に起因する観測揺らぎ
- ユーザの動きによるアーチファクト
- 温度や気圧等の環境変動

の4つに分類されると述べている [51]. 本研究では (c) を利用して頭部状態を認識し (a), (b), (d) が認識誤差の原因になると考えられる. (a) による誤差の低減には回路構成などのハードウェアによる対策とノイズ処理などのソフトウェアによる対策を行う必要がある. (d) による誤差に関して, 気圧の変化は音速に影響せず無視できると考えられる. 一方, 気温の変化は音速に影響するため, 今後の調査が必要であるが, 本研究で提案した信号補正手法を応用することで認識精度の低下は抑えられると考えられる. 最も大きい誤差要因は (b) によるものと考えられ, その対策として本研究では測定信号に補正を加える手法を提案し, 評価実験を行った. 実験の結果, 21 種類の頭部状態認識において 22.3 ポイント, 6 種類



の頭部状態認識において 16.6 ポイントの改善が確認されたため、本手法は装着誤差への対策として有効であると考えられる。実際に実験参加者 A の 1 セット目から 7 セット目までの閉口状態の未補正時と補正時の振幅スペクトルを、それぞれ図 2.17 と図 2.18 に示す。図 2.18 を見ると、音量の大小はあるが、図 2.17 よりもセットによらず同様の概形が得られている。特徴量抽出時には正規化を行うため、グラフの概形が同じであれば、音量の大小は問題ないといえる。また実験参加者 A の 1 セット目から 7 セット目までの閉口状態の LFCC の分散を確認した。未補正時では各次元の分散の平均が  $5.1 \times 10^{-2}$  であったが、補正時には  $6.2 \times 10^{-3}$  に減少していることが確認できた。

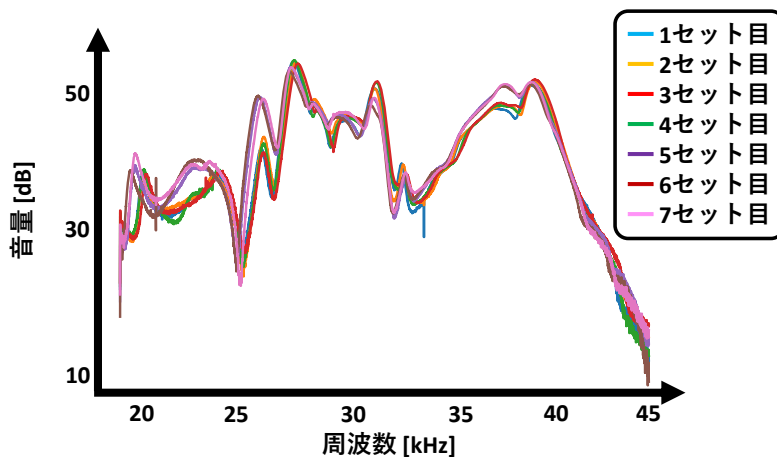


図 2.17: デバイス着脱による振幅スペクトルの変化 (未補正信号)。

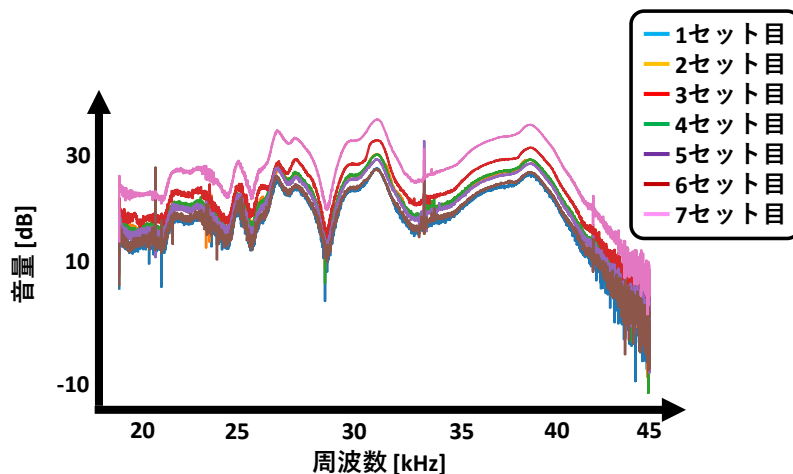


図 2.18: デバイス着脱による振幅スペクトルの変化 (補正信号)。

## 測定信号の自動補正

本研究ではデバイス装着時に測定信号の補正を行うことで認識精度の向上を実現したが、条件として一意の頭部状態（本研究では閉口状態）の時に補正を行う必要がある。本研究では閉口状態を確認して手動で補正を行ったが、実用性を考慮すると自動で補正タイミングを検出する必要がある。現在販売されているデバイスのいくつかには、音楽再生の自動停止を行うために近接センサを用いてデバイスの着脱状態の識別を行っている。したがって、ユーザがデバイスを装着したタイミングの検出は可能であるため、装着時の信号補正は自動で行えると考えられる。その他に、ユーザ動作によって生じる装着誤差を補正するためには、現在の頭部状態が閉口状態か非閉口状態であるかを認識する必要がある。その場合、システムの誤認識にはユーザが閉口状態の時に非閉口状態と誤認識してしまう場合と、非閉口状態の時に閉口状態と誤認識してしまう場合の2種類の誤認識が存在する。前者に関しての誤認識は、信号の補正が行われただけであり、システムとしては大きな問題にはならないと考えられる。一方、後者に関しての誤認識は非閉口状態時に信号の補正が行われるため、補正後の認識精度は著しく低下すると考えられる。そのため、非閉口状態の認識は100%に近い精度であることが求められる。

実際に実用を考慮した6種類の頭部状態に対して閉口状態と非閉口状態の2グループに分割し、再度認識精度の調査を行った。調査には2.5.2節と同じ学習手法を用いた。調査の結果、閉口状態の認識精度は79.7%、非閉口状態の認識精度は93.7%であった。閉口状態の認識精度は、適切な頻度で信号の補正を行うために十分な認識精度と考えられる。しかし、非閉口状態の認識精度に関しては90%以上ではあるものの、経時的に補正する場合、不適切な補正が行われる可能性が高いため、より精度の高い認識精度が求められる。また、信号の誤補正が起こった場合の適切な対処手法も考察する必要がある。

### 信号補正手法の定量的な性能調査

本研究ではデバイスの装着は毎回同じ状態になるように、実験参加者に求めた。また、装着誤差がどの程度生じているか計測していない。したがって、装着誤差は比較的小さく、データ毎にどの程度の装着誤差が含まれているのか不明である。今後は、本手法の信号補正手法がどの程度の装着誤差に対して有効であるか調査するために、装着誤差を定量的に生じさせた場合の低減性能を調査する実験が必要である。また、装着誤差の大きい状態に対して、信号補正が可能となる新たな手法を考察する必要がある。

## 2.6.2 頭部状態認識手法に関する議論

### 応用的な頭部状態の調査

本実験では複数の頭部状態を組み合わせた状態や詳細な頭部状態などを認識していない。例えば、口を開けたまま右を向くなどの状態や、多段階の口の開け方の認識も頭部状態と考え得る。これら応用的な頭部状態が認識可能となれば、例えば日常生活で再現することが非常に少ない頭部状態をコマンドとして利用でき、コマンドモードへの移行システムが

不要になる． 今後はこれらの詳細な調査を進め，本手法の更なる可能性を調査していく必要がある．

#### 市販製品での実装

荒川らは市販製品である inCore [44] を用いて外耳道音響特性から個人認証を行っている [50]．しかし，inCore を含め市販製品では超音波の再生・録音はできないため，本手法のような常時の信号再生を前提としたシステムには利用できない．そのため，本実験ではプロトタイプデバイスを自作して，認識精度の調査を行った．現在は OpenEarable [63] のような超音波帯域の外耳道音響特性が取得可能な開発者向けデバイスも販売されている．したがって，今後の機器の発展によって市販のデバイスを用いて本手法の実装が可能になると考えられる．

#### 実用環境での認識精度

本実験では座位状態にて各実験参加者の外耳道伝達関数を取得した．今後は実用を考慮し，歩行時や音楽鑑賞時に外耳道伝達関数を取得し，認識精度の調査を行う必要がある．また，頭部状態をコマンドとして使用しているデバイスは普及していない．そのため，本手法のユーザビリティ調査も行う必要がある．さらに，実環境ではデバイスのコマンドに使用する頭部状態を日常生活でも使用するため，ユーザの意図しないタイミングでのデバイス操作につながる可能性がある．そのため，デバイスの非操作モードと操作モードを切り替える必要がある．コマンドモードを切り替える専用のコマンドとして，歯同士の接触音による切替 [64] や複数の頭部状態を組み合わせたなどの日常生活では意図せず出にくいものを採用することを考えている．

## 2.7 まとめ

本研究では AAS における課題である装着誤差の低減を目的とした測定信号の動的補正手法を提案し，ヒアラブルデバイス向け頭部状態認識手法に適用した．外耳道音響特性が取得可能なプロトタイプデバイスを実装し，評価実験を行なった結果，未補正信号による認識実験で 40.2% (F 値)，補正信号による認識実験で 62.5% (F 値) の精度で認識できることを確認した．また，実際の利用シーンを考慮し，認識する頭部状態を 6 種類に限定したところ，未補正信号による認識実験で 74.4% (F 値)，補正信号による認識実験で 90.0% (F 値) の認識精度が得られた．

## 第3章 ノイズ環境でのAASの動的利用

本章ではノイズ環境でのAASの動的利用手法に関する研究の詳細を述べる。本研究では音響センシングを用いた衣服へのインタフェース構築手法に本手法を適用し、有効性を評価した。本章の構成は以下の通りである。3.1節にて、音響センシングを用いた衣服へのインタフェース構築手法の研究背景および本研究の成果を述べる。3.2節にて関連研究との比較・差異を述べる。3.3節にて本手法の詳細、3.4節にてハードウェア・ソフトウェアの実装の詳細を述べる。3.5節にて評価実験の設計・手順・結果を述べ、3.6節にて本研究の議論を述べる。最後に、3.7節にてまとめを述べる。

### 3.1 音響センシングを用いた衣服へのインタフェース構築手法

ヒアラブルデバイスやスマートウォッチなどのウェアラブルデバイスが普及してきたが、これらのデバイスは身体に装着する必要があるため、デバイスの機械感や装着感を強く感じさせるという課題 [7] がある。この課題を緩和するデバイスとして衣服型ウェアラブルデバイスが注目されている。衣服はほとんどのユーザが常に着用しているため、イヤホン型や時計型などのデバイスと比較して長時間利用する場合のユーザの心理的/肉体的負担が少ないというメリットがある。そのため、衣服にウェアラブルデバイスの機能を上手く溶け込ませることができればより多くのユーザがウェアラブルデバイスの機能を享受できる。現在、衣服型ウェアラブルデバイスの研究は電気を通す導電繊維を利用し、ジェスチャなどを認識する研究 [9, 65, 66] が主流である。しかし、導電繊維の利用には導電繊維を編み込んだ専用の衣服やスリーブ型デバイス、刺繍での衣服への埋め込みが必要であるため、デザインの選択肢が限られ、導入コストや耐久性などの課題がある。

本研究ではこれらの課題を解決するため、普段着用している一般的な衣服上をタッチインタフェース化する着脱式デバイスとタッチジェスチャの認識手法を提案する。本手法は着脱式のデバイスによって、ユーザの好みの場所にタッチインタフェースを構築することができる (図 3.1)。認識システムは一对のピエゾ素子で構成される。ピエゾ素子は電圧の印加によって振動する性質を利用してスピーカとして利用できる。また、逆の効果を利用して振動を電圧に変換するマイクとしての利用も可能となる。ユーザが衣服上をタッチやスワイプしたとき、衣服には生地の変形と衣擦れ音が生じる。この時、スピーカ側の素子から超音波の信号を衣服に再生し、マイク側で信号の伝播を録音すると、生地の変形によって信号の伝播経路が変化し、得られる信号の音響特性も変化する。また、衣擦れ音はタッチジェスチャによって、その音響特性の分布が変化する。認識システムはタッチジェスチャごとの超音波信号と衣擦れ音の変化パターンを機械学習することで現在のタッチジェスチャを認識する。本手法はAASとPASを併用し、ユーザのタッチジェスチャを認識している。

AASとPASの併用によって、ノイズ環境下においても認識精度の低下を抑えられることを確認した。本研究の成果を以下にまとめる。

- 音響センシングによる衣服上タッチジェスチャ認識手法を提案した。
- 参加者ごとのデータを用いた個人分類モデルにて5種類の前腕部タッチジェスチャ認識実験を行った。実験の結果、認識システムの精度は95.9%(F値)であることを確認した。
- AASの併用によって、PASのみによる認識と比較してノイズ環境下で平均3.6ポイントの精度向上を確認した。

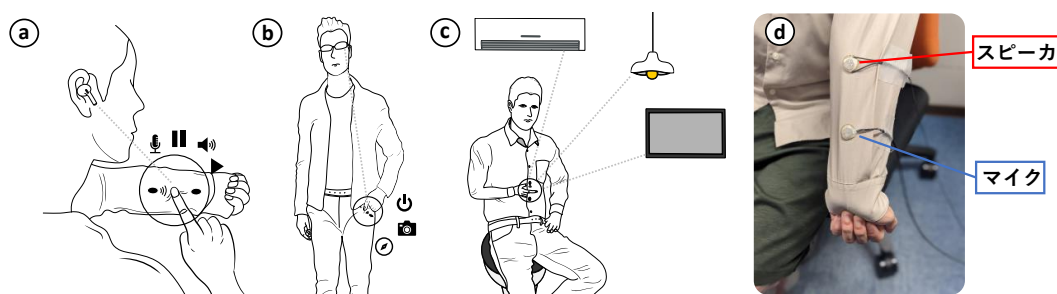


図 3.1: システムの利用想定イメージとプロトタイプデバイス a: イヤホン操作イメージ b: スマートグラス操作イメージ c: スマートホーム内での利用イメージ d: 実験に用いたプロトタイプデバイス。

## 3.2 関連研究

本節では音響センシングを用いたインタフェース構築に関する研究と、衣服インタフェース構築に関する研究およびに衣服装着型デバイスに関する研究について説明し、本研究との比較・差異を述べる。

### 3.2.1 音響センシングを用いたインタフェース構築に関する研究

音響センシングを用いたインタフェース構築に関する研究は、筆者の知る限り AAS と PAS を同時に利用した研究はこれまで行われておらず、AAS・PAS の一方が利用されてきた。各センシング手法ごとに関連研究を説明する。

#### AAS を用いた研究

既存のオブジェクトに対して AAS を適用し、タッチインターフェースや物体認識センサーとして利用することに焦点を当てた研究がいくつか行われてきた [67, 68, 3, 69, 70]。例えば Ono ら [3] はコンタクトスピーカとコンタクトマイクを用いてブロックや食器などに超音波を伝播させ、その伝播音を録音・解析することでそれらの触れ方の違いを認識する手法

を提案している。岩瀬ら [69] はアクリル板を伝播する超音波を利用して、板の上に乗っている複数の物体の識別と位置推定を行っている。

### PAS を用いた研究

人体や物体の表面上をスワイプ、タップすることでユーザの入力として利用する研究はこれまで多く行われてきた [71, 72, 73, 74]。例えば Mujibiya ら [2] はユーザが指先で行うタップ、スワイプジェスチャ時に発生する骨導音を録音・解析することでこれらのジェスチャを認識する手法を提案している。Murray-Smith ら [75] はユーザのひっかけ動作やこすり動作時に特殊な音響信号が発生する小型デバイスを開発し、デバイス入力機器としての利用手法を提案している。Chris ら [76] は木や布などの物体表面を指先で擦る動作によって生じる特有の音をコンタクトマイクで収録し、ユーザの指先の軌跡を推定する手法を提案している。Chris ら [77] はマイクアレイを用いて皮膚上のタップ位置の推定を行っている。Xu ら [21] は市販のイヤホンを用いて、デバイス周辺のタッチ、スワイプジェスチャ時に発生する音を解析し、ジェスチャ入力として利用する手法を提案している。

これらの研究は、AAS や PAS を用いて人体や物体表面をタッチインタフェース化する点で本手法と同様である。一方で、本手法は一般的な衣服のタッチインタフェース化を目指す点でこれまでの研究と異なる。衣服はこれまでの AAS に関する研究で適用されてきたブロックや机、食器などの物体と比較して音波の伝播強度が弱く、形状が不安定であったため、調査が行われていなかった。また、PAS に関する研究ではノイズの影響で認識精度が低下する課題があるが衣服に関しても同様であると考えられる。本手法では、AAS と PAS を併用することで、AAS 単体では困難であった衣服上の様々なタッチジェスチャを認識可能なシステムを提案し、PAS の課題であったノイズの影響を AAS を利用することで低減できることを確認した。

### 3.2.2 衣服インタフェース構築に関する研究

これまでの衣服インタフェース構築に関する研究は、導電繊維を衣服に埋め込み、ユーザの動作や生体反応によって変化する抵抗値や静電容量値をセンシングする手法が主流であった。多くの導電繊維埋め込み手法に関する基礎研究が実装コストや耐久性、センシング精度などの向上を目指して行われてきた [78, 66, 79, 80, 81, 82, 83, 84, 65, 85, 86, 11]。

導電繊維を既存の衣服に埋め込む手法に関する研究では、Patrick ら [65] が圧力センシングを可能とする抵抗型圧力検知繊維を開発している。また、Ivan ら [85] は静電容量型タッチ検知繊維を開発している。Roland ら [86] は、市販のミシンを用いて相互キャパシタンス方式の導電繊維を高速かつ簡易に埋め込む手法を提案している。

一方、衣服型スマートデバイスの実用的な利用手法に関する研究は主にジェスチャ認識 [87, 88, 89, 90, 9, 91]、姿勢認識 [92, 93, 94, 95, 96]、バイタルセンシング [97, 98, 99, 100] の3種類にわかれる。表 3.1 に本手法と導電繊維を利用したジェスチャ認識手法および衣服装着型デバイスの類似研究の比較をまとめた。ジェスチャ認識に関する研究では、Patrick ら [9] が導電繊維を埋め込んだスリーブ型のデバイスを開発し、デバイスへのタッチジェ

スチャの認識を行っている。評価実験の結果、9種類のジェスチャを89.5%の精度で認識できることを確認している。また、Te-Yenら[91]は導電繊維をアンテナのように利用することで、センシング部分での空中ジェスチャ時に発生するドップラーシフトを測定し、それらのジェスチャを認識する手法を提案している。評価実験の結果、11種類の空中/タッチジェスチャを85.2%の精度で認識できることを確認している。姿勢認識に関する研究では、Corinneら[95]が伸縮センサを埋め込んだ衣服型デバイスを用いて27種類の姿勢を97.0%の精度で認識できることを確認している。また、Rubioら[96]は伸縮と圧力が検知可能な市販の導電繊維を利用して肘角度推定を行っている。評価実験の結果、実験参加者間での推定において9.69°の誤差で推定可能であることを確認している。バイタルセンシングに関する研究では、Aliら[99]が2種類の導電布を組み込んだ衣服を用いて睡眠時の心拍数と呼吸数のモニタリングシステムを提案している。Jiら[100]は綿で覆われた導電繊維を用いて汗の吸収量を測定し、ユーザの発汗モニタリングシステムを提案している。

これらの導電繊維埋め込み手法は、ユーザは特殊な衣服を着用するか、自身の衣服に導電繊維を埋め込む必要がある。ユーザは天候や目的、嗜好にあわせて様々な衣服を使い分けするため、導電繊維を利用する手法は導入の煩雑性や複数回の洗濯による摩耗などの課題を解決する必要がある。一方、本手法は磁石を用いてデバイスを衣服に装着する機構を採用しているため、簡易に装着/脱着が可能であり、他の衣服にも転用可能である。また、洗濯などの影響を受けない。ただし、測定信号を伝播させるために衣服表面を張った状態にする必要があり、前腕ジェスチャの実行時には両手を必要とする。

表 3.1: 各手法の比較.

	本手法	Patrick [9]	Te-Yen [91]	Pin-Sung [101]	Artem [10]
利用素子	スピーカ マイク	導電繊維	導電繊維	赤外線センサ 電界センサ	加速度センサ 慣性センサ
認識	衣服上 ジェスチャ	衣服上 ジェスチャ	空中 ジェスチャ	タップ スワイプ ピンチ	タップ 腕の回旋
認識精度	95.9%	92.0%	92.8%	98.9%	—
認識数	5種	9種	11種	5種	2種
着脱	可	不可	不可	可	可
制限	衣服を張る	—	—	ジッパー機構 が必要	—

### 3.2.3 衣服装着型デバイスに関する研究

衣服に装着するデバイスに関する研究もいくつか行われてきた。Scottら[102]は加速度センサを搭載した小型デバイスを開発し、ポケット部分に装着することでタップ認識可能なシステムを提案している。Pin-Sungら[101]は赤外線/電界/指紋センサを搭載したジッ

パー型デバイスを開発し、ユーザインタラクションに関する幅広い調査を行っている。ジェスチャ入力に関する調査では、デバイスのタップやジッパーを上下に動かすジェスチャなどを認識している。Artemら [10] は9軸センサを内蔵したボタン型デバイスを開発し、デバイスへのタップや腕の回旋を認識するシステムを提案している。

本手法と同様にこれらの研究は衣服に着脱可能な小型デバイスである。しかし、これらの研究は衣服表面をタッチインタフェース化する手法ではなく、認識できるジェスチャやインタフェースの構築される場所に限りがある。本手法は音響センシングを用いることで衣服上のタッチ、スワイプ、つまむなどのジェスチャの認識が可能である。

### 3.3 本手法

ユーザは piezo 素子で構成されたデバイスを一對衣服に装着する。この時、スピーカ側の piezo 素子から測定信号を再生し、マイク側は衣服上を伝播する測定信号と衣擦れ音を録音する。タッチジェスチャの種類によって衣服の変形し、伝播する測定信号が変化する。また、ユーザが衣服に触れるタッチジェスチャでは衣擦れ音が発生する。これらの伝播する信号と衣擦れ音の分布はタッチジェスチャによって、得られる音響データが変化する (図 3.2)。この音響データの変化パターンを機械学習を用いて学習し、分類モデルを作成する。提案システム全体の流れを図 3.3 にまとめた。

それぞれのタッチジェスチャの音響データを2種類のフィルタを通して得られる可聴域/全帯域のデータから特徴量を抽出し機械学習を行うことで、タッチジェスチャを認識する分類モデルを作成する。本研究では、特徴量に LFCC, スペクトル重心などのスペクトル特徴量, アタックタイムなどの波形特徴量を用いた。また、分類モデルには SVM を用いて、個人ごとのジェスチャデータによる分類モデルの作成と、ジェスチャの認識を行う。

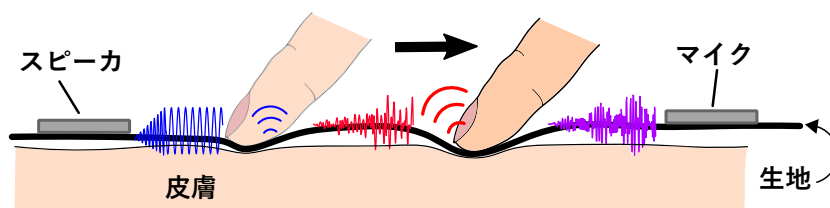


図 3.2: 認識原理.

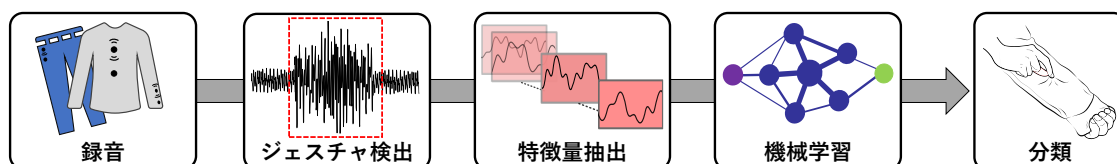


図 3.3: タッチジェスチャ認識システム構成.



### 3.3.1 衣服上タッチジェスチャ認識の実現可能性に関する基礎調査

筆者の知る限り音響センシングを用いて衣服上のタッチジェスチャ認識を行う研究はない。したがって、本手法が実現可能であることを確認するための基礎調査を行った。

#### アクティブ音響センシング

AASに関して、衣服上のタッチジェスチャ時に起こると考えられる衣服の変形や張力の変化が、伝播する超音波信号に与える影響を調査した。図 3.4 に測定器具を示す。ジーンズ（綿 100%、非導電性布、厚み：0.50 mm）の一部を長方形に切り出し、張力用重りを配置できるように縫製している。信号再生と録音に用いる piezo 素子を 10 cm の距離で固定する。piezo 素子の固定には磁石を用いている。この状態でスピーカ側の piezo 素子から 18 kHz から 48 kHz に遷移するアップスイープ信号を繰り返し再生し、伝播する信号をマイク側の piezo 素子で録音する。信号の再生・録音共にサンプリングレートは 96 kHz で行った。スピーカ側の piezo 素子は Thrive 社の OMR20F10-BP-310、マイク側の piezo 素子は Murata 社の 7BB-20-6L0 を用いた。張力用重りの重量を変化させていく張力テストと、変形用重りの重量を変化させていく変形テストを行った。張力テストでは中央に 50 g の重りを配置した状態で張力用おもりの重量を 100 g から 50 g ごとに 300 g まで変化させ、それぞれの周波数応答を測定した。変形テストでは 100 g の張力用おもりを配置し、50 g の変形用おもりをスピーカ側、中央、マイク側に配置し、それぞれの周波数応答を測定した。

図 3.5 に張力テスト時の周波数特性の変化をまとめ、図 3.6 に変形テスト時の周波数特性の変化をまとめた。それぞれのテストで張力や変形によって周波数応答が変化していることがわかる。したがって、衣服上で行われるタッチジェスチャ時に超音波信号が変化し、認識システムに利用できると思われる。

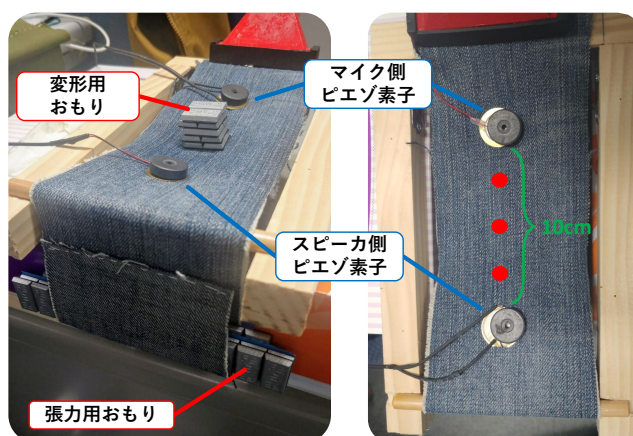


図 3.4: 測定環境と変形用重りの配置位置。右図の赤点が変形用重りの配置位置を示す。

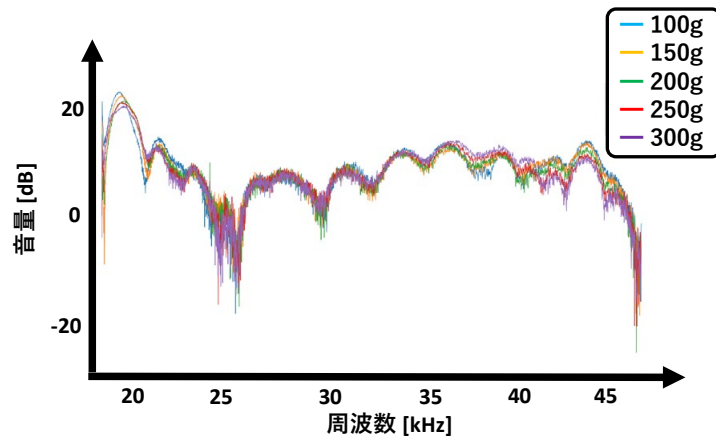


図 3.5: 張力テストによる周波数応答の変化.

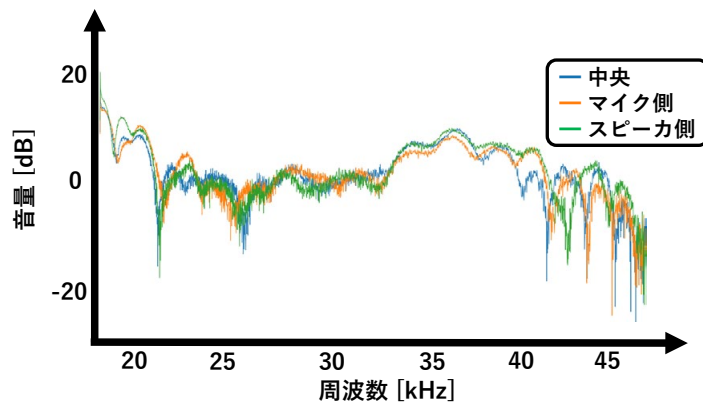


図 3.6: 変形テストによる周波数応答の変化.

## パッシブ音響センシング

PAS に関してジェスチャによって得られる衣擦れ音の分布が十分に異なることを確認するための測定テストを行った。測定は市販の長袖シャツの前腕部分にマイク用の piezo 素子を磁石で固定する。その状態で、録音を行い装着者に異なるジェスチャを piezo 素子周辺で行ってもらう。信号録音のサンプリングレートは 96 kHz で行う。piezo 素子は Murata 社の 7BB-20-6L0 を用いた。図 3.7 に服をねじるジェスチャ時と円を描くジェスチャを行ったときのスペクトログラムをまとめた。ジェスチャによってスペクトルの概形や持続時間、時間変化が異なることがわかる。したがって、PAS を用いたジェスチャ認識が可能であると考えられる。

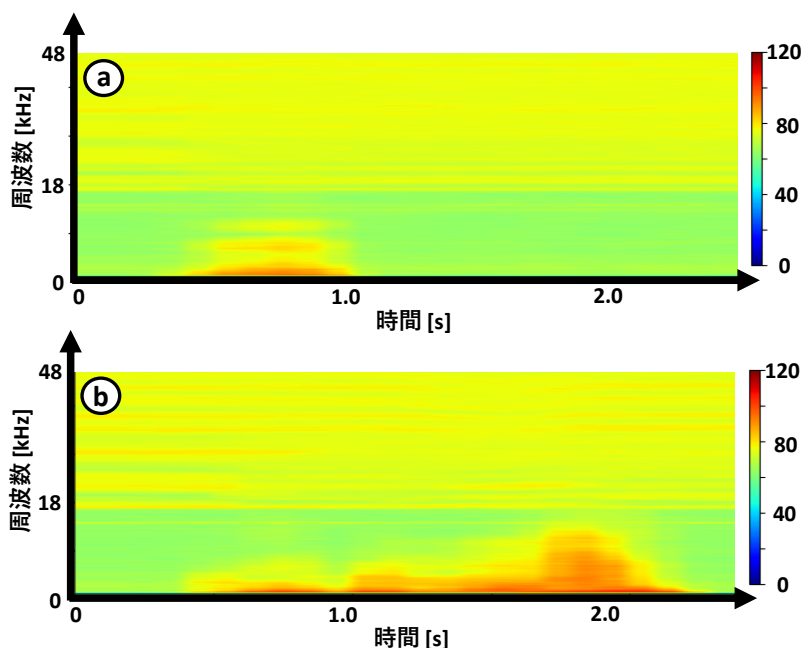


図 3.7: 前腕部ジェスチャ時のスペクトログラムの違い a: ねじるジェスチャ b: 円を描くジェスチャ。

### 3.3.2 ジェスチャ部分の抽出

認識システムはジェスチャの認識を行うためにジェスチャ部分の音響信号を抽出する必要がある。本手法ではスペクトルの移動平均を基準とした閾値ベースによるジェスチャ開始推定を行う。いくつかのジェスチャ時のスペクトログラムを確認したうえで、各ジェスチャのスペクトログラムはジェスチャ開始時の衣服と指先が触れるときの音量が最も大きく、その後、音量が高い状態がジェスチャ終了まで続くことがわかった。そこで、短期移動平均と長期移動平均の 2 つの移動平均を用いてジェスチャ検出を行う。まず短期移動平均で衣服と指先が接触した瞬間を検出し、次に長期移動平均でジェスチャーが行われたかどうかを判断する。スペクトログラムの移動平均  $MA(k)$  は以下の式で表される。

$$MA(k) = \frac{1}{n} \sum_{i=k-n+1}^k s_i.$$

ただし、 $k$ はスペクトログラムの $k$ 番目のインデックス、 $n$ は移動平均のサイズ長、 $s_i$ は $i$ 番目のスペクトルパワーの合計値である。短期移動平均のサイズ長を $n_s$ 、長期移動平均のサイズ長を $n_l$ とし、 $MA(k)$ を用いてジェスチャー発生を推定する。これら2つの移動平均値がそれぞれ設定された閾値を超えたときの $k$ をジェスチャー開始時刻とする。ジェスチャーの開始時刻の前後の音響信号データを一定長で切り出し、ジェスチャー部分の音響信号データとして機械学習に利用する。

### 3.3.3 特徴量抽出および分類モデル

機械学習の効率を上げるために、ジェスチャーの音響信号データに対して特徴量抽出を行った。本研究ではAASとPASで一般的に用いられる特徴量を利用する。具体的には、スペクトル特徴量と波形特徴量を抽出した。表3.2に特徴量の一覧と各特徴量の次元数をまとめた。

分類モデルにはSVMを利用した。ランダムフォレストやニューラルネットワークなどの、その他の分類モデルの利用も考えられるが最も精度の高かったSVMを選定した。分類モデルの選定は3.5.1項にて詳細を述べる。

表 3.2: 特徴量一覧.

特徴量	次元数
LFCC <sup>S</sup>	456
スペクトル重心 <sup>S</sup>	24
サブバンドピーク <sup>S</sup>	24
スペクトラルフラックス <sup>S</sup>	24
スペクトラルロールオフ <sup>S</sup>	24
スペクトラルフラットネス <sup>S</sup>	24
スペクトラルバンド幅 <sup>S</sup>	24
ゼロ交差率 <sup>W</sup>	24
二乗平均平方根 <sup>W</sup>	24
アタックタイム <sup>W</sup>	1
分散 <sup>W</sup>	8

<sup>S</sup> スペクトル特徴量    <sup>W</sup> 波形特徴量

### 3.3.4 ピエゾ素子の選定

ピエゾ素子はメーカーによって信号再生時の周波数特性や録音時に得られる周波数特性が異なる。本研究で用いるピエゾ素子をMurataの7BB-20-6L0 (Murata), ThriveのOMR20F10-BP-310 (Thrive1), K2512BS1 (Thrive2)の3種類から選定した。各素子をスピーカ側とマイク側に用いた場合の全ての組み合わせ(9通り)で周波数特性を調査した(図3.8)。信号の測定環境および測定方法は3.3.1節と同様である。測定の結果、スピーカにThriveの

OMR20F10-BP-310, マイクに Murata の 7BB-20-6L0 を使用した場合 (凡例赤色) の周波数応答が最も高かったため, これらの Piezo 素子の組み合わせを利用する.

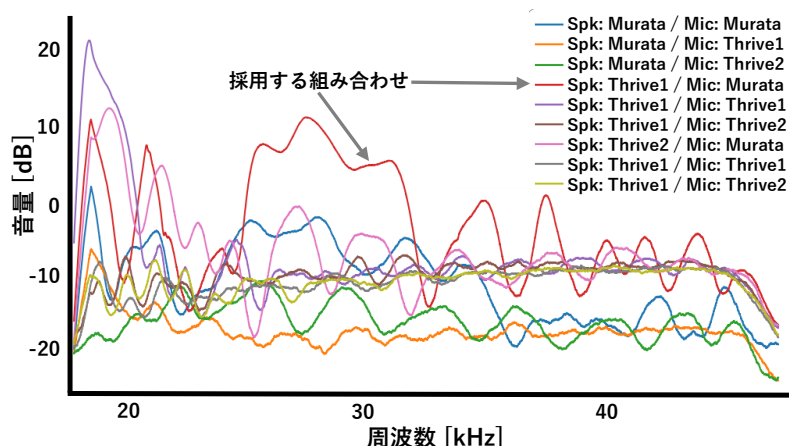


図 3.8: 3 種類の Piezo 素子の組み合わせによる周波数特性の差異.

## 3.4 実装

提案するシステムは衣服上に測定信号を伝播させ, 伝播信号と衣擦れ音を録音する着脱式デバイスのハードウェア実装とタッチジェスチャ認識アルゴリズムのソフトウェア実装にわかれる. 以下でそれぞれの実装の詳細を述べる.

### 3.4.1 信号再生・録音デバイス

図 3.9a に作製したデバイス, 図 3.9b にデバイスの構造および装着機構を示す. 信号の再生と録音には Piezo 素子を用いるが, デバイスの着脱を可能にするために Piezo 素子を磁石で挟み込む機構を採用した. 信号再生時の気導音の発生および録音時の外部からのノイズの影響を減らすために, 上部をグルーで覆った. リアルタイムで録音信号のスペクトルを確認したが, デバイス上部での空中動作によるドップラーシフトの発生はないことを確認した. また, デバイスと人体の接触は下部磁石と生地を介しているため, 人体からの信号伝播の影響は非常に小さい. 実際に, デバイスを前腕部に装着した状態で指先や手首を動かしたがスペクトルの変化は確認されなかった. 図 3.10 にデバイス構成をまとめた. スピーカ側の Piezo 素子は Thrive の OMR20F10-BP-310, マイク側は Murata の 7BB-20-6L0 を用いた. 音響信号入出力時の DA/AD 変換に Roland の OCTA-CAPTURE を使用し, PC は Mouse Computer の DAIV 4N を使用した. 再生, 録音共にサンプリングレートは 96kHz で行う.

### 3.4.2 タッチジェスチャ認識アルゴリズム

図 3.11 にタッチジェスチャ認識の流れをまとめた. ジェスチャの発生を判定する移動平均の 2 つのパラメータ  $n_s$  と  $n_l$  はそれぞれ 2 と 8 に設定した. AAS に使用する Swept-Sine 信

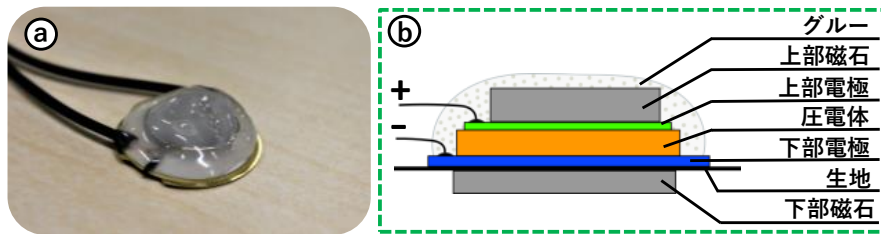


図 3.9: 実験デバイス a: 実装デバイス b: デバイスの装着構造.

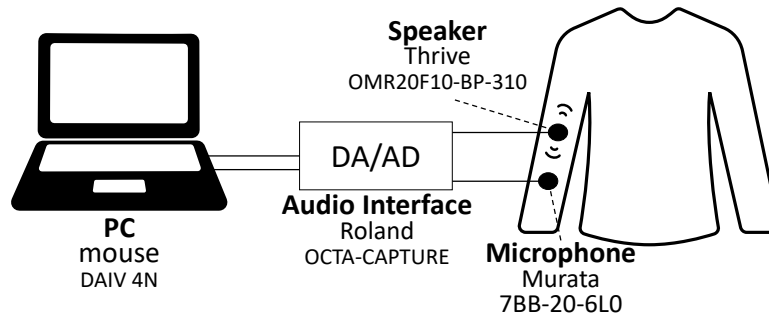


図 3.10: デバイス構成.

号は 18kHz 以下を帯域制限した, 18kHz から 48kHz のアップスイープ信号 (8,192 サンプル) を使用した. ジェスチャの抽出時間は 196,608 サンプル (約 2,048ms) とし, ジェスチャ開始時刻より 32,768 サンプル (約 341 ms) 前から抽出する. 音響データの時間変化を抽出するためにジェスチャ部分の音響データを一定長のサンプル窓で分割する. 今回は Swept-Sine 信号が最低でも 3 周期以上含まれるように, 32,768 サンプル幅のハミング窓を用いて音響データを分割した. ハミング窓のシフト幅は 8,192 サンプルとした. したがって, 1 回のタッチジェスチャの音響データから 24 フレーム ( $(196,608 - 32,768) / 8,192 + 1 = 24$ ) のデータ配列を得る. 各配列データに対して, 3.3.3 節で述べた特徴量を抽出する.

本手法は AAS と PAS を併用しているため, 全周波数帯域 (0–48 kHz), 測定信号の再生帯域 (18kHz–48 kHz), 可聴域帯域 (0–18 kHz) ごとに得られる特徴量が異なると考えられる. 実験の結果, 最も認識精度とノイズ頑健性が高い周波数帯域の組み合わせは全周波数帯域と可聴域帯域の音響データを用いた場合であることがわかった. 調査の詳細は 3.5.1 節にて述べる. したがって, 上記で述べた特徴量抽出は 2 つのフィルタによって得られる音響データからそれぞれ抽出するため, 合計で 1,346 次元 ( $673 \times 2$ ) となる. データの測定・特徴量抽出・機械学習は Python 3.7 で実装した.

### 3.5 評価実験

本手法の有効性を確認するためには, 複数の衣服や装着場所でシステムが利用できることを確認する必要がある. 本実験では大別して前腕部タッチジェスチャ認識実験とボタン/ポケット部タッチジェスチャ認識実験の 2 種類の実験を行った. 前者で収集したデータを

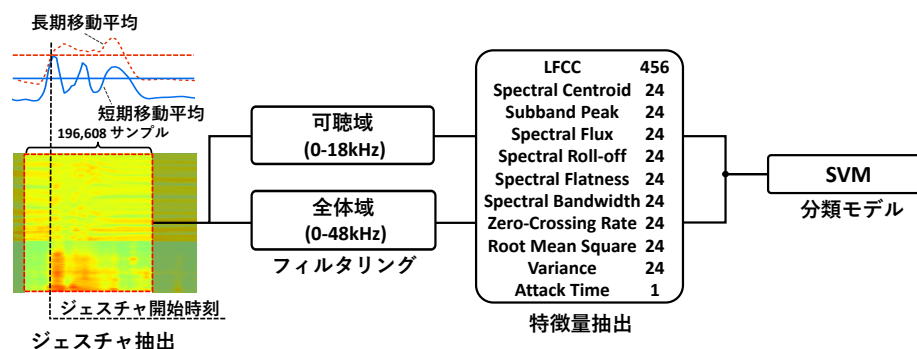


図 3.11: タッチジェスチャ認識の流れ。

用いて、認識システム設計のための予備調査や4種類の異なる衣服の認識性能などを調査した。後方で収集したデータを用いてデバイスの複数箇所での利用可能性を調査した。以下でそれぞれの測定実験の詳細を述べる。

### 3.5.1 前腕部タッチジェスチャ認識実験

本項では前腕部タッチジェスチャ認識実験に関する詳細について述べる。本項は以下の4つの内容にわかれる。

- 実験設計：本実験で測定するジェスチャ・実験参加者・実験環境・実験手順についてそれぞれ説明する
- ジェスチャ抽出精度：収集したデータのジェスチャ抽出精度をまとめる
- 認識システム設計のための予備調査：認識に用いる分類モデルおよびに利用する周波数帯域の組み合わせを決定する
- 評価実験：4種類の異なる衣服の認識性能の差異およびに認識システムの性能評価を調査する

#### 実験設計：測定ジェスチャ

前腕部は関連研究 [87, 9, 90] にて最も調査が行われている部位であり、人体表面をタッチインタフェースとして利用する場合に最も利用される頻度が高いと考えられる。したがって、前腕部でのタッチジェスチャ認識実験を行う。測定するタッチジェスチャは、認識アルゴリズムとデバイス構成の条件などを考慮し、以下の3つの基準を参考に選定する。

生地変形を発生させるジェスチャ 伝播する測定信号が変化するためには、衣服の変形が必要である。したがって、ジェスチャは衣服の変形を発生させるものとし、指先で触れるジェスチャはスピーカとマイクの間で行う。

スピーカ/マイク間に対して非対称な軌跡をもつジェスチャ スワイプの上方向と下方向などの組み合わせである，スピーカとマイクを結ぶ線分に対して，対称であるジェスチャは得られる衣擦れ音が似てしまうと考えられる．したがって，このようなジェスチャの組み合わせは選定しなかった．

デバイスに触れないジェスチャ 測定信号の減衰による情報消失を防ぐため，スピーカとマイクのデバイス距離は約 7.0cm としている（図 3.12 左破線）．測定中にデバイスに触れてしまうと，ノイズとなり認識精度に影響するため，選定するジェスチャは設定されたデバイス距離内で行えるものとした．

上記の基準を基に 11 種類のタッチジェスチャを選定した（図 3.12）．タッチジェスチャ実行時は，測定信号を伝播させるためにユーザは左手で袖を引っ張る．その状態にて，「回旋する」ジェスチャ以外は右手で行う．

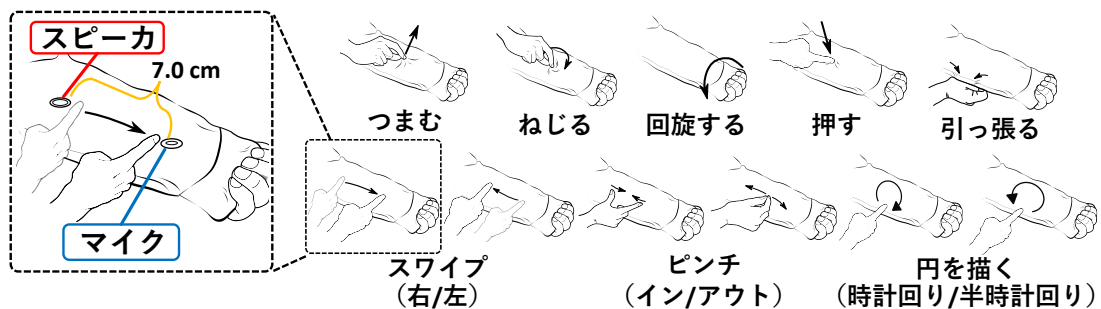


図 3.12: 測定する衣服タッチジェスチャ．

#### 実験設計：実験協力者および実験環境

実験は 20–26 歳の 12 名に参加してもらった．参加者は全員右利きである．実験は筆者の所属する研究室の学生部屋で行い，参加者には座った状態でデータの測定を行ってもらった．空調の停止や周りの人に静かにしてもらうなどの騒音対策は行っていない．

#### 実験設計：実験手順

本実験では衣服の構成要素（厚み，表面加工，素材）が異なる 4 種類の衣服（表 3.3）でデータ収集を行った．まず，参加者は半袖シャツの上から自身のサイズ（S-LL）に合う衣服を着用してもらう．その後，図 3.12 に示すタッチジェスチャを図および動画を用いて口頭で説明し，全てのタッチジェスチャを再現できるようにする．その後，デバイスを衣服に装着してもらう．デバイスの装着位置は肘と手首間の中点位置から肘/手首方向に約 3.5 cm の位置にそれぞれスピーカ/マイクを装着してもらう．測定手順は，デバイスを装着した状態で衣服の袖を引っ張り，タッチ部分を張った状態にしてもらう．この状態を確認し，測定



信号の再生と録音を開始する。再生・録音開始後、参加者に前面ディスプレイに表示されたタッチジェスチャを行ってもらい、ジェスチャ終了後に録音を停止する。この測定を全てのタッチジェスチャでランダムな順番で行ったものを1セットとし、デバイスの着脱をセットごとに行った。これを15セット測定し、各参加者から165回分（11種類×15セット）のタッチジェスチャデータを収集した。実験参加者12名の内、6名はすべての衣服の測定実験に参加してもらった。同一参加者より得られた各タッチジェスチャのスペクトログラムを各衣服ごとに付録Bの図B.1にまとめた。

表 3.3: 実験測定に用いた4種類の衣服詳細.

衣服タイプ				
	シャツ	シャツ	フリース	ジャケット
ラベル	シャツ A	シャツ B	フリース	ジャケット
素材	100% 綿	67% ポリエステル, 33% 綿	100% ポリエステル	100% ポリエステル
導電性	なし	なし	なし	なし
表面加工	なし	なし	起毛	なし
生地の厚さ [mm]	0.33	0.17	0.99	0.27
テクスチャ				

### ジェスチャ抽出精度

収集した7,920個（12名×衣服4種×11ジェスチャ×15セット×）のデータを用いて、3.3.2節で述べたスペクトログラムの移動平均によるジェスチャ抽出アルゴリズムの抽出精度を調査した。各衣服とタッチジェスチャごとの抽出精度を表3.4にまとめた。

実験の結果、全体の抽出精度は95.7%であった。「押す」や「つかむ」ジェスチャの抽出精度が低かった。「押す」に関してフリースの抽出精度（79.2%）が著しく低かった。これはフリースが厚みがあるため、衣擦れ音の音量が小さく、かつ「押す」ジェスチャは他のジェスチャと比べてジェスチャ実行時間が短い。そのため、長期移動平均の閾値を超えないデータが多くなってしまったと考えられる。「つかむ」ジェスチャはデバイス装着位置の反対側に触れる動作である。したがって、指先の衝突時の音が小さくなってしまい、短期移動平均の閾値を超えないデータが多くなってしまったと考えられる。ジェスチャ実行時にシステムに検出されなかった場合、再度ジェスチャを実行する必要がある。これはユーザビリティ低下につながるため、今後は抽出アルゴリズムの改善を行う必要がある。

表 3.4: 各タッチジェスチャのジェスチャ抽出精度 [%].

種類	シャツ A	シャツ B	フリース	ジャケット	平均
つまむ	97.8	93.3	98.3	98.3	97.1
ねじる	98.3	98.3	100.0	95.8	97.9
回旋する	95.0	94.2	91.7	95.0	94.0
押す	97.2	95.8	79.2	93.3	91.5
つかむ	91.7	93.3	90.8	85.0	90.2
スワイプ右	98.3	96.7	95.0	98.3	97.1
スワイプ左	97.2	95.8	100.0	97.5	97.5
ピンチイン	98.3	94.2	99.2	99.2	97.7
ピンチアウト	98.9	95.8	98.3	97.5	97.5
円を描く (時計回り)	99.4	92.5	95.8	95.8	95.2
円を描く (反時計回り)	97.8	95.0	98.3	97.5	96.9
平均	97.3	95.0	94.9	95.8	95.7

#### 認識システム設計のための予備調査：分類モデルの選定

機械学習に用いる分類モデルを決定するための予備調査を行った。本研究では、実験データのサイズを考慮し、SVM、ランダムフォレスト (RF: Random Forest)、ニューラルネットワーク (NN: Neural Network)、勾配ブースティング決定木 (GBDT: Gradient Boosting Decision Tree)、スタッキングアンサンブル学習 (SEL: Stacking Ensemble Learning) の 5 種類の分類モデルの性能比較を行う。各分類モデルのハイパーパラメータの調査範囲とスタッキング学習のモデル構成を表 3.5 にまとめた。ハイパーパラメータは、ベイズ最適化アルゴリズム [103] を用いて導出した。

シャツ A のデータを用いて、上記のモデルにて個人分類モデルの学習を行った。機械学習ではデバイスの着脱の影響を含むために Leave-One-Round-Out-CrossValidation で行った。訓練データの交差検証は 5 分割交差検証を行い、F 値が最も高かったパラメータをハイパーパラメータとした。図 3.13 に各モデルの認識精度の全体平均を示す。実験の結果、SVM の認識精度 (F 値: 87.4%) が最も高かった。以上より、今後の評価実験には SVM を分類モデルとして利用する。

#### 認識システム設計のための予備調査：周波数帯域の組み合わせパターンの選定

本手法は AAS と PAS を併用しているため、全周波数帯域 (0–48 kHz)、測定信号の再生帯域 (18 kHz–48 kHz)、可聴域帯域 (0–18 kHz) ごとに得られる特徴量が異なると考えられる。本実験では、それぞれの 3 種類の帯域パターンを全帯域、信号帯域、可聴域帯域とし、帯域パターンの組み合わせとジェスチャ認識性能の関係をノイズなしデータとノイズありデータを用いてそれぞれ調査した。

最初に、ノイズなしデータを用いて認識精度を調査した。3 種類の帯域パターンの全ての組み合わせは 7 通り (信号帯域、可聴域帯域、全帯域、信号帯域 + 可聴域帯域、信号帯域

表 3.5: 各分類モデルのパラメータ調査範囲 (SVM, RF, NN, GBDT) とモデル構成 (SEL).

分類モデル	パラメータ/モデル構成
SVM	カーネル (rbf/sigmoid) $C$ ( $1-10^5$ ), $\gamma$ ( $10^{-5}-10^3$ )
RF	深さ (1-100) , 葉の数 (1-100)
NN	活性化関数 (双曲線正接関数/正規化線形関数) 最適化 (準ニュートン法/Adaptive Moment Estimation) レイヤー数 (1-5), ニューロン数 (20-1000)
GBDT	深さ (2-16) , 葉の数 (1-100), 最小サンプル数 (1-5)
SEL	1 段目 (RF×6, NN×2, SVM, ロジスティック回帰, アンサンブル学習) 2 段目 (GBDT)

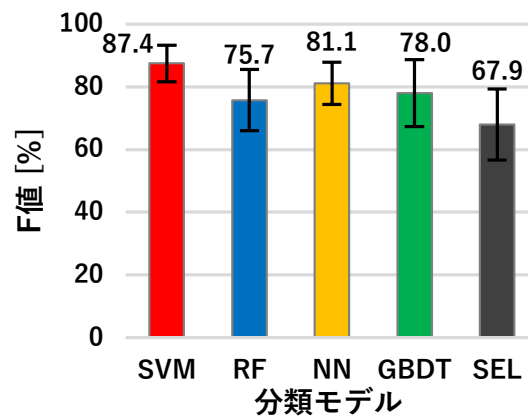


図 3.13: 各分類モデルのジェスチャ認識精度の全体平均. エラーバーは参加者ごとの標準偏差を示す.

+全帯域, 可聴域帯域 + 全帯域, 信号帯域 + 可聴域帯域 + 全帯域) となる。これらの帯域で抽出した特徴量を用いて、節と同様にシャツ A のデータを用いてジェスチャ認識精度を調査した。図 3.14 に各組み合わせの認識精度の全体平均を示す。実験の結果、可聴域帯域 + 全帯域を用いた分類モデルの精度 (F 値: 87.4%) が最も高かった。

続いて、ノイズありデータを用いて認識精度を調査した。本実験では一般的な利用環境での発生が多いと考えられる 4 種類のノイズデータ (車/音楽/音声/衣擦れノイズ) を利用する。各ノイズと信号の SNR(Signal Noise Ratio) を固定するために、事前に録音したノイズデータをノイズなしデータに追加するシミュレーション方式による調査を行った。ノイズ録音には実験と同じマイク (Murata 7BB-20-6L0) を用いた。車ノイズは屋外の道路脇にて車が走行する音を録音した。音楽ノイズはロックミュージックをスピーカから再生し、その音声を録音した。音声ノイズはマイクを実験と同様に装着した状態で、装着者が文章を読み上げる音声を録音した。衣擦れノイズはマイクを実験と同様に装着し、装着している腕を前後に振った時に生じる音を録音した。各ノイズ (車/音楽/音声/衣擦れ) と実験データの SNR はそれぞれ平均で 4.9 dB, 3.6 dB, 6.6 dB, 7.9 dB であり、全体の標準偏差が 3.4 dB であった。これらの調査の結果、SNR は主に 10 dB 1 dB であるといえる。したがって、本実験では各ノイズと実験データの SNR を 10 dB, 6 dB, 3 dB, 1 dB に調整し、認識精度の調査を行った。ノイズなしデータの評価と同様にシャツ A のデータを用いる。特徴量抽出の際に、検証データに対してノイズの付加を行う。ノイズなしデータで作成した分類モデルに対してノイズを含んだ検証データの認識精度を評価する。図 3.15 に各組み合わせのノイズなしデータの精度および SNR を変化させた場合の認識精度の全体平均の推移を示す。各ノイズのすべての SNR において、可聴域帯域 + 全帯域データによる認識精度が最も高かった。ノイズなし実験では、可聴域帯域データのみによる認識精度 (F 値: 87.3%) も高かったが、SNR が 1 dB の時の認識精度は平均で 21.0 ポイント低下することがわかった。

以上の実験より、ノイズなし/ありの両方において可聴域帯域 + 全帯域データによる認識が高い性能を示すことがわかった。今後の評価実験は可聴域帯域 + 全帯域データを用いた認識精度調査を行う。

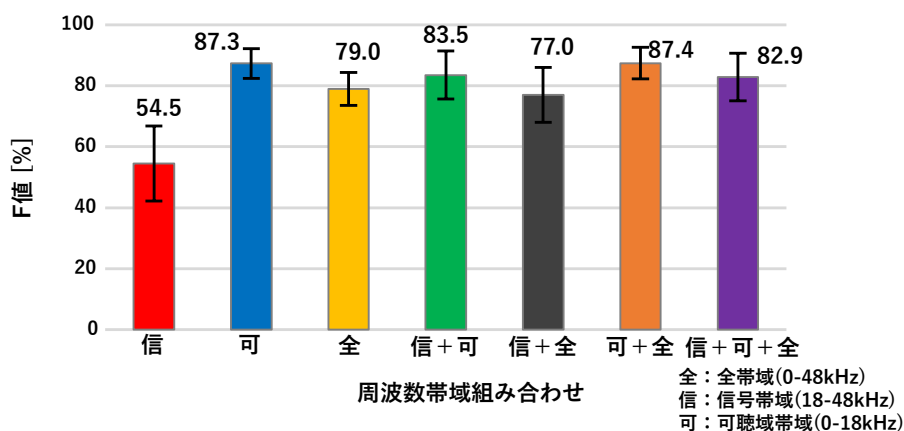


図 3.14: 各周波数帯域組み合わせパターンの認識精度。エラーバーは参加者ごとの F 値の標準偏差を示す。

✕ 信号帯域+全帯域    ▲ 可聴域帯域+全帯域    ★ 信号帯域+可聴域帯域+全帯域  
 ● 信号帯域    ◆ 可聴域帯域    ■ 全帯域    ◆ 信号帯域+全帯域

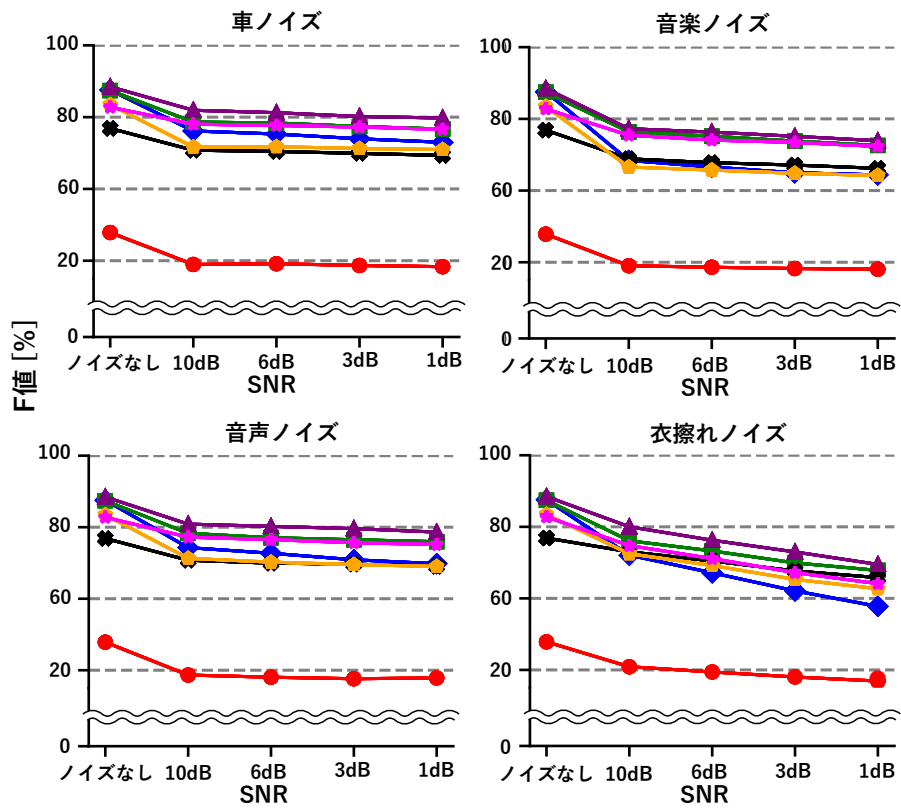


図 3.15: 各ノイズの周波数帯域組み合わせパターンの精度変化.

## 評価実験：衣服ごとの認識性能調査

上記の実験で決定した分類モデル (SVM) と周波数帯域の組み合わせ (可聴域帯域 + 全帯域データ) による本手法の認識性能の調査を行う。一般的な衣服には様々な構成要素 (素材, 厚み, 生地 of 縫い方, 表面加工) があるが, これらのすべての要素に対して本手法の認識性能を調査することは現実的ではない。したがって, 本研究では厚みや表面加工などが異なり, 上着として着用されることが多い 4 種類の衣服の認識性能調査を行った。本実験では衣服ごとの認識性能の差異を前項と同様にノイズなしデータおよびノイズありデータにて調査した。図 3.16a にノイズなしデータによる各衣服ごとの認識精度の実験参加者全員の平均を示し, 図 3.16b に全ての衣服のデータ測定実験に参加した 6 名の平均を示す。12 名の参加者の認識精度の平均はシャツ A, シャツ B, フリース, ジャケットの順にそれぞれ 87.4%, 80.2%, 85.1%, 82.6%であった。また, 衣服全体の平均認識精度は 83.9%であった。全衣服実験参加者 6 名の認識精度の平均はシャツ A, シャツ B, フリース, ジャケットの順にそれぞれ 85.9%, 81.6%, 88.3%, 79.3%であった。また, 衣服全体の平均認識精度は 83.8%であった。各衣服ごとのジェスチャの認識の分布を示す混同行列を図 3.17 にまとめた。衣服全体で最も認識精度の高いタッチジェスチャは「押す」で 95.6%, 低いタッチジェスチャは「ピンチアウト」で 77.7%であった。

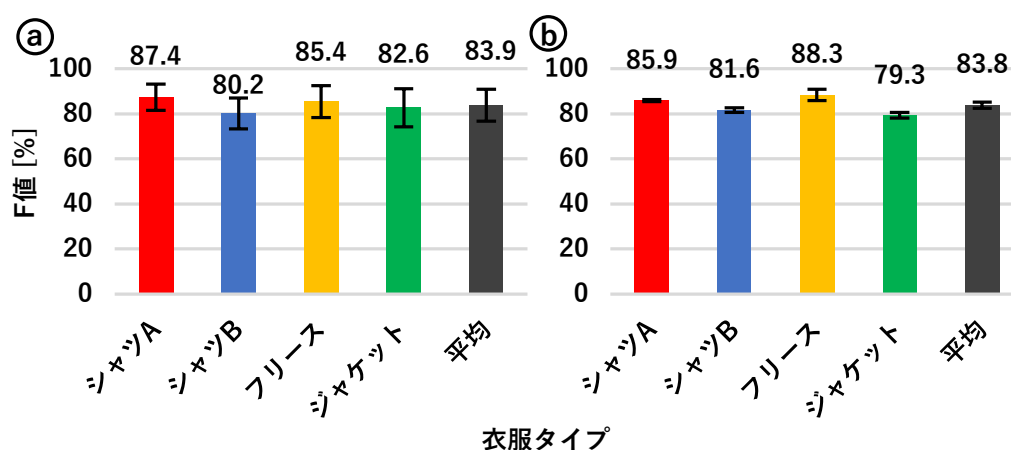


図 3.16: 各衣服の全体認識精度。エラーバーは参加者ごとの F 値の標準偏差を示す。a: 全参加者の平均 b: 全ての衣服のデータ測定に参加した 6 名の平均。

上記の実験より, どのタッチジェスチャが本手法に有効であるかを 11 種類のタッチジェスチャから調査した。しかし, デバイスの操作や入力インターフェースとしての実用性を考えると認識精度が 83.9%では十分とはいえない。本研究でも 2.5.6 節と同様に 5 種類のタッチジェスチャを選定し, 認識精度の再調査を行う。認識精度や誤認識の傾向を考慮し「ねじる, 回旋する, 押す, スワイプ右, 円を描く (半時計回り)」の 5 種類のタッチジェスチャを選定した。これらのジェスチャデータのみを用いて再度分類モデルを作成し, 認識精度の調査を行った。各衣服ごとの認識精度の平均を図 3.18 にまとめた。認識精度の平均はシャツ A, シャツ B, フリース, ジャケットの順にそれぞれ 98.5%, 94.9%, 94.6%, 95.2%であった。衣服全体の平均認識精度は 95.9%となり, 平均 95%を超えた。ジェスチャの認識の分布

		予測ラベル														F値 [%]		予測ラベル														F値 [%]																	
		シャツA							シャツB									シャツA							シャツB																								
正解ラベル	つまむ	158	7	4	0	3	1	1	1	0	0	1	90.4	132	26	3	0	7	0	1	1	0	1	3	79.0	150	16	6	0	3	0	0	1	2	0	0	85.0	147	23	5	0	3	0	0	0	0	0	0	82.4
	ねじる	11	158	0	0	4	2	0	0	0	0	1	89.8	18	145	1	0	5	2	0	2	1	0	1	80.1	12	157	1	0	3	1	0	0	0	5	1	84.6	12	150	2	1	3	0	0	1	1	2	3	83.4
	回旋する	3	0	163	1	4	0	0	0	0	0	0	94.2	5	2	131	5	8	0	3	0	0	0	0	91.1	4	1	148	4	6	1	0	0	2	0	0	88.8	9	3	140	2	15	0	1	0	1	0	0	82.0
	押す	0	0	2	171	2	0	0	0	0	0	0	97.2	1	0	1	144	1	2	0	0	1	0	0	94.9	0	0	2	134	2	2	0	0	1	0	0	95.2	0	1	1	162	4	0	0	0	0	0	0	96.1
	引っ張る	1	5	4	0	153	1	0	0	0	0	1	89.9	2	7	9	3	100	2	0	0	2	0	1	82.6	7	5	5	0	138	5	0	3	1	0	0	81.5	3	3	18	4	118	2	0	1	2	0	2	76.3
	スワイプ左	0	1	0	1	2	152	9	2	9	1	0	84.9	2	1	1	1	3	138	17	1	8	2	2	76.3	0	4	1	2	10	132	0	2	21	0	0	75.6	2	0	2	0	5	141	7	4	13	3	1	80.7
	スワイプ右	0	1	1	1	0	8	139	13	10	1	1	79.2	2	0	0	0	0	13	144	15	4	1	0	69.0	0	3	0	0	1	1	155	11	7	0	2	85.7	3	0	0	1	0	10	136	13	11	1	2	79.1
	ピンチイン	0	1	0	1	2	3	16	146	6	0	2	82.6	1	1	1	0	0	6	19	134	11	4	0	68.4	1	1	1	0	1	7	17	144	4	2	1	83.7	0	1	0	0	2	3	10	151	11	1	0	85.1
	ピンチアウト	0	1	1	1	1	12	10	14	138	0	0	80.9	2	2	0	1	2	13	8	14	133	0	2	75.8	1	2	0	0	1	27	6	3	136	1	0	76.3	0	2	0	0	3	12	11	4	140	4	1	77.8
	円を描く(時計回り)	0	0	1	1	3	3	0	0	0	155	16	85.8	0	3	3	0	8	0	2	1	1	111	29	83.7	0	3	0	0	2	0	1	1	1	148	16	83.4	1	1	2	0	2	0	1	1	1	148	16	83.4
	円を描く(半時計回り)	0	1	0	0	3	0	0	0	0	21	151	86.9	3	5	0	0	1	1	0	0	0	21	139	81.7	0	1	0	0	1	0	3	0	1	11	161	90.9	1	2	0	0	1	2	0	0	2	24	144	82.6

図 3.17: 各衣服の混同行列.

表 3.6: 各タッチジェスチャのジェスチャ認識精度 [%].

種類	シャツ A	シャツ B	フリース	ジャケット	平均
つまむ	90.4	79.0	85.0	82.4	97.1
ねじる	89.8	80.1	84.6	83.4	97.9
回旋する	94.2	91.1	88.8	82.0	94.0
押す	97.2	94.9	95.2	96.1	91.5
つかむ	89.9	82.6	81.5	76.3	90.2
スワイプ右	84.9	76.3	75.6	80.7	97.1
スワイプ左	79.2	69.0	85.7	79.1	97.5
ピンチイン	82.6	68.4	83.7	85.1	97.7
ピンチアウト	80.9	75.8	76.3	77.8	97.5
円を描く (時計回り)	85.8	83.7	88.5	83.4	95.2
円を描く (反時計回り)	86.9	81.7	90.9	82.6	96.9
平均	87.4	80.2	85.1	82.6	83.9

を示す混同行列を図 3.19 に示す. すべてのジェスチャの認識精度が 94%以上であった. 以降の実験では, 5 種類のジェスチャを用いた認識精度を評価する.

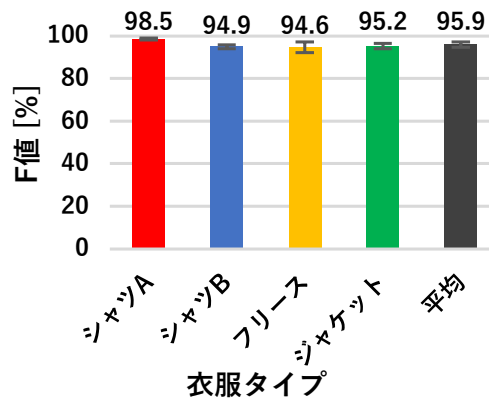


図 3.18: 5 種類のジェスチャ認識時の各衣服の全体認識精度. エラーバーは参加者ごとの F 値の標準偏差を示す.

ノイズありデータによる衣服ごとの認識精度の変化を調査した. 利用ノイズや評価手法などは 3.5.1 節のノイズシミュレーション実験と同様である. 図 3.20 に各ノイズごとに, SNR を変化させた時の衣服ごとのジェスチャ認識精度の推移をまとめた. また, 本手法は AAS を利用することでノイズ頑健性を高めている. 本手法の有効性を比較するために, 可聴域帯域のみのデータによる認識精度の変化も調査した. 実線が本手法 (可聴域帯域 + 全帯域データ) の認識精度の推移を示し, 破線が可聴域帯域データのみでの認識精度の推移を示す.

実験の結果, 本手法の認識精度はノイズなしから 1dB にかけて全体平均で 95.9% から 89.5% まで 6.4 ポイントの低下が確認された. 一方, 可聴域帯域データのみでの認識精度はノイズなしから 1dB にかけて 95.0% から 85.2% となり 9.8 ポイントの低下が確認された. 同



		予測ラベル					F値 [%]
正解ラベル	ねじる	678	8	3	7	11	95.5
	回旋する	9	629	14	10	0	95.4
	押す	1	4	625	4	0	96.7
	スワイプ (右)	15	12	16	652	8	94.4
	円を描く (反時計回り)	10	4	1	5	680	97.2

図 3.19: 5 種類のジェスチャ認識時の全体の混同行列.

じ 1 dB で比較すると本手法は 4.3 ポイントの精度向上があることがわかった。また、全てのノイズ・衣服・SNR の組み合わせ (全 64 パターン) において本手法の認識精度が可聴域帯域データのみでの認識精度よりも高いことを確認した。精度向上の平均は 3.6 ポイントであった。

#### 評価実験：データ測定時間と認識精度の関係

これまでの評価実験では、訓練データに 14 セット分のデータを用いて、残りの 1 セットを検証データとして認識精度を評価してきた。しかし、個人分類モデルの作成にはユーザに事前にデータ測定を求める必要がある。必要な訓練データ数は少ないほうがユーザビリティの向上に繋がる。本実験では訓練データ数と、認識精度の関係およびユーザの印象度評価を調査した。印象度評価の調査は、実験参加者にデータ測定に要する時間が 3/5/10/15 分であると仮定した場合の印象を 5 段階 (1: 不満-5: 満足) のリッカート尺度を用いて回答してもらった。図 3.21 に横軸をデータ測定に要する時間とし、ジェスチャ認識精度およびユーザの印象度の推移を縦軸にまとめた。訓練データの測定時間の換算は、5 ジェスチャを 1 セット測定する場合の所要時間を 50 秒 (10 秒 × 5 ジェスチャ) とする。つまり、14 セットは 700 秒 (約 11.7 分) と換算される。認識精度に関して、6 セット分 (5 分) 以上の訓練データを用いた場合、認識精度が 90% を超えている。一方、ユーザの印象度は 5 分を超えるデータ測定に対しての印象が悪い。これらの結果より、6 セット以下のデータ測定がユーザビリティの観点から望ましく、現在の認識システムでは 6 セットが最小のデータ必要数と考えられる。

#### 評価実験：同一モデルの異なる衣服での精度評価

本実験では同一モデルの異なる衣服に対して、訓練データの再測定を必要とせずにジェスチャ認識が可能であることを確認する。シャツ A の測定実験に参加した 12 名より 6 名の実験参加者に別インスタンスのシャツ A を着用してもらい、データ測定を 5 セット分を行った。このデータを検証データとし、オリジナルのシャツ A で測定したデータを用いて作成した分類モデルに適用する。図 3.22 にオリジナルデータと別インスタンスデータを検証データとした場合の認識精度をまとめた。全体平均でオリジナルデータの認識精度が 98.2%、別インスタンスデータでの認識精度が 95.5% であった。

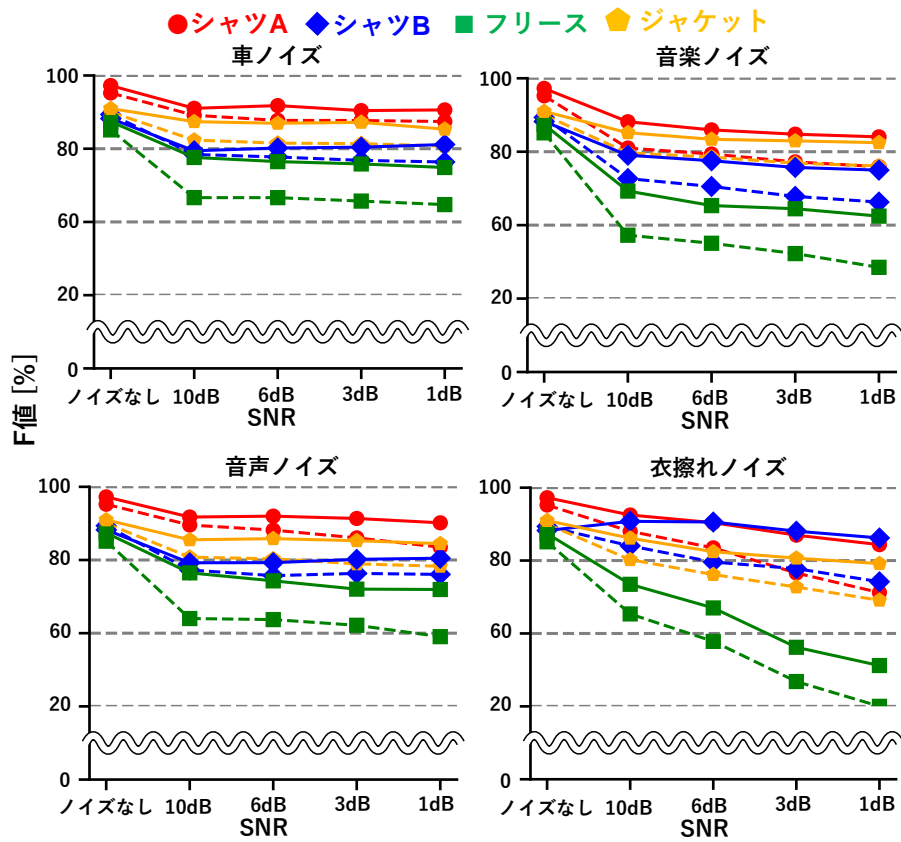


図 3.20: 各ノイズ付加による認識精度の推移. 実線は本手法 (全帯域+可聴域) による認識精度を示し, 破線は可聴域のみによる認識精度を示す.

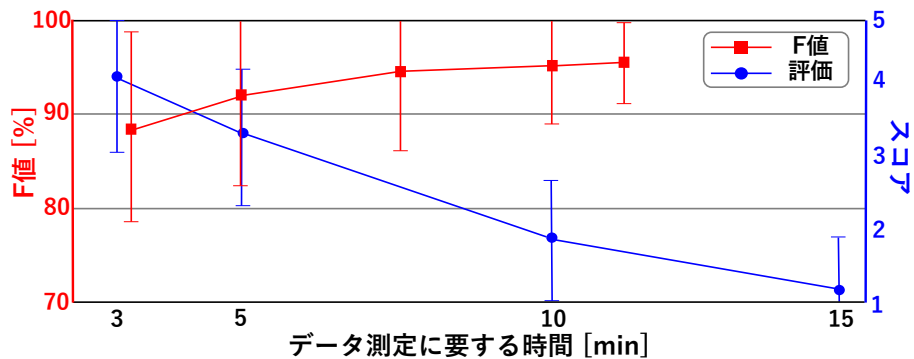


図 3.21: データ測定の所要時間と, 認識精度およびユーザの印象度の推移.

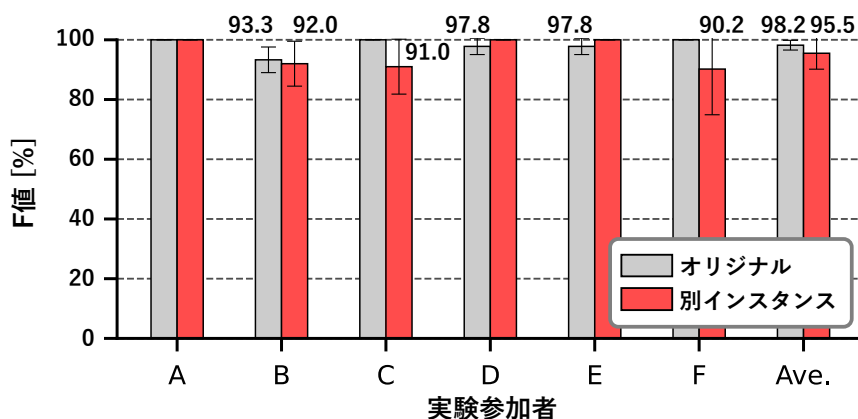


図 3.22: 別インスタンスデータを検証データとした場合の認識精度.

#### 評価実験：デバイス装着位置と認識精度の関係

本実験ではデバイスの装着位置が異なる場合のジェスチャデータによる認識精度を調査する。具体的には、これまでのデバイス装着位置（中央配置）からスピーカとマイク間の距離は保ったまま、肘方向と手首方向にデバイスをずらした場合のジェスチャデータを測定する。肘側へのデバイス配置では、スピーカが肘上に配置されるようにする。手首側へのデバイス配置では、マイクが手首上に配置されるようにする。肘方向/手首方向にデバイスを配置した時のそれぞれの状態で得られるデータを肘側データと手首側データとする。データ収集はシャツ A の測定実験に参加した 12 名より 6 名の実験参加者に対して 5 セット分のデータを測定した。肘/手首側データを検証データとし、中央配置で測定したシャツ A のジェスチャデータ（中央データ）を用いて作成した分類モデルに適用する。図 3.23 に肘側データと手首側データ、一部の中央データを検証データとした場合の認識精度をまとめた。認識精度の平均は肘側データと手首側データで 90.7% と 88.6% であり、中央データによる認識精度（F 値：98.2%）と比較して 7.5 ポイントと 9.6 ポイントの精度低下が確認された。認識精度の低下要因は、デバイスの装着位置が変わることでデバイス周辺の衣服形状が変化し、音響データにも変化が生じてしまったためであると考えられる。認識精度の改善のためには、それぞれの装着位置でのジェスチャデータを訓練データに組み込むことなどが考えられる。本実験では 5 セット分のデータ測定を行ったため、1 セット分のデータを検証データとして、残りの 4 セット分のデータを訓練データに組み込み、精度調査を行った。その結果、肘側データ/手首側データによる認識精度の平均は 94.7% ( $SD=5.44$ ) と 94.9% ( $SD=4.79$ ) まで向上した。4 セット分のデータ測定は概算で 200 秒（10 秒  $\times$  5 ジェスチャ  $\times$  4 セット）を要するため、今後はユーザビリティとの関係性などを考慮して、最適な必要データ数を決定する必要がある。

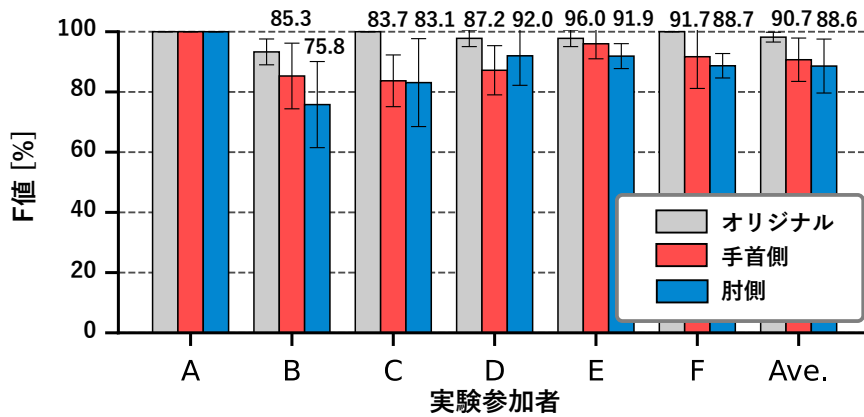


図 3.23: 手首/肘側データを検証データとした場合の認識精度.

#### 評価実験：一般分類モデルでの認識精度

これまでの実験では個人ごとのデータを用いて個人ごとのジェスチャを認識する個人分類モデルによる精度評価を行った．本実験では他者のデータを用いて作成した一般分類モデルによる精度評価を調査する．図 3.24 に各衣服の一般分類モデルによる認識精度の結果をまとめた．個人分類モデルの認識精度と比較して，各衣服（シャツ A，シャツ B，フリース，ジャケット）の平均認識精度はそれぞれ 7.1, 5.8, 6.3, 5.6 ポイント低下した．認識精度が低下した要因はジェスチャ時の指先の移動速度や力のかけ具合が参加者によって異なったためであると考えられる．図 3.25 は t 分布型確率的近傍埋込み法（t-SNE 法） [104] によるシャツ A の各特徴の分布を，対象参加者と他の参加者で色分けして示したものである．認識精度の高い参加者の特徴量分布はジェスチャごとに分離している（図 3.25a）．一方，認識精度の低い参加者の特徴量分布はジェスチャごとの分布のばらつきが少なく，対象者自身の他ジェスチャのデータとの類似性が高いため，一般的な分類器には適さないことが分かる（図 3.25b）．

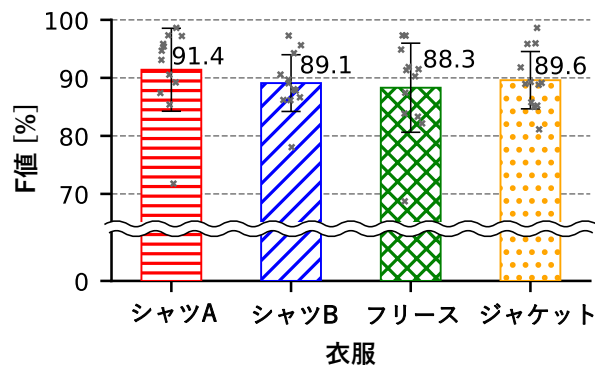


図 3.24: 一般分類モデルによる各衣服のジェスチャ認識精度．エラーバーは参加者ごとの F 値の標準偏差を示す．× は各参加者の F 値を示す．

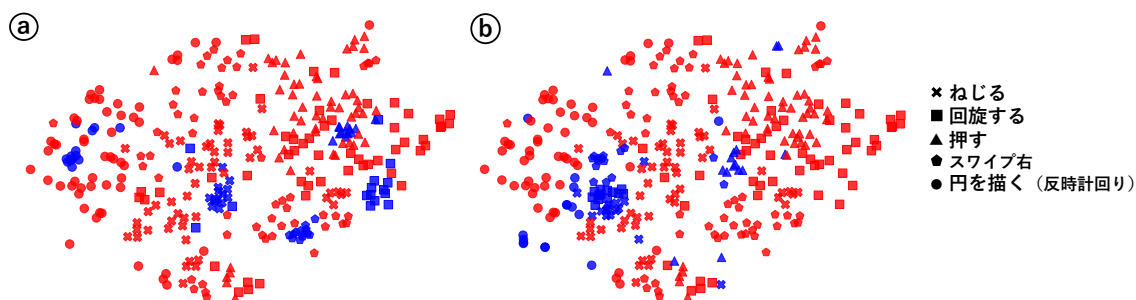


図 3.25: t-SNE 法を用いたシャツ A の特徴量の分布. 青色は特定の参加者 1 名のジェスチャデータを示し, 赤色はその他の参加者のジェスチャデータを示す. a: 認識精度が最も高い参加者を青色で示した分布 b: 認識精度が最も低い参加者を青色で示した分布.

### 3.5.2 ボタン/ポケット部タッチジェスチャ認識実験

本手法は磁石を用いてデバイスと衣服を挟み込む装着機構を採用しており, 磁石を外すことでユーザは自身の好みに合わせてジェスチャ入力の箇所を変更できることを想定している. 本実験ではシャツのボタン部分とズボンのポケット部分でのジェスチャ認識が可能であるか調査し, 本手法の複数箇所での利用可能性を評価する. 図 3.26 に各箇所のデバイス位置および測定するタッチジェスチャをまとめた. ユーザは中指, 薬指, 小指を用いて衣服を張る. その状態から同じ手の親指にてタッチジェスチャを行う. 本実験では, 片手で行えるタッチジェスチャかつ, 3.3.2 節の 3 つの条件を満たすジェスチャを選定した. また, 誤認識が多くなるスワイプ左/スワイプ右などの左右対称なジェスチャ同士の組み合わせを極力避けた. ボタン部分のジェスチャデータ測定 (ボタン部分実験) にはシャツ A を用いた. ポケット部分のジェスチャデータ測定 (ポケット部分実験) には市販のパンツ (素材: 86%綿, 9%ポリエステル, 5%ポリウレタン. 厚さ: 0.53mm) を用いた. 各箇所のジェスチャデータを 5 名の参加者から収集した.

図 3.27 に各実験の参加者毎と全体平均の認識精度をまとめた. 各実験の全体平均はそれぞれボタン部分実験の認識精度が 89.2%, ポケット部分実験の認識精度が 92.6%であった. 図 3.28 に各実験の混同行列をまとめた. 全てのジェスチャの認識精度が 85%以上を有することを確認した. 一方, これらの実験と前腕部分実験 (平均: 95.9%) を比較すると認識精度は前者の方が低かった. この要因は, 本実験ではねじるなどの前腕部分実験で精度が高かったジェスチャを選定しなかったためであると考えられる. 一方で, ねじるジェスチャを実行する場合, 両手が塞がるという課題がある. 今後はユーザビリティと認識精度のバランスを考慮したジェスチャ選定を行う必要がある.

## 3.6 考察

本節ではノイズ環境での AAS の動的利用に関する議論およびに衣服上タッチジェスチャ認識手法に関する議論について述べる.

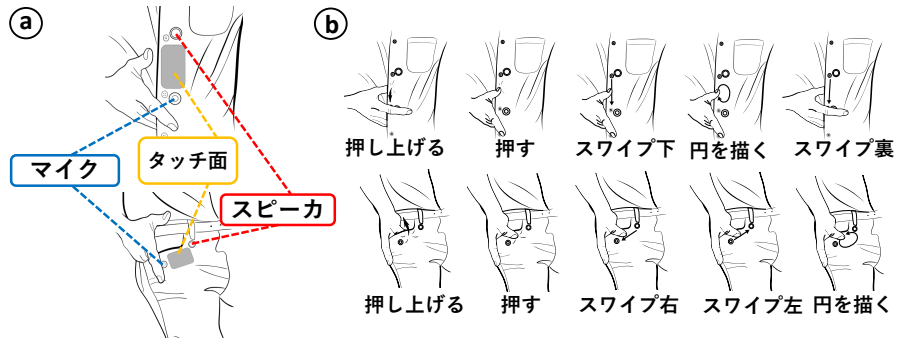


図 3.26: ボタン/ポケット部タッチジェスチャ a: デバイス位置とジェスチャ実行位置 b: 測定するタッチジェスチャ.

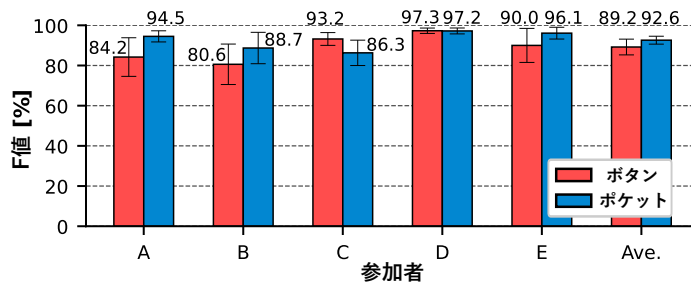


図 3.27: ボタン部とポケット部の各実験参加者の認識精度. エラーバーはジェスチャごとの F 値の標準偏差を示す.

		予測ラベル					F値 [%]
正解ラベル	押し上げる	61	4	2	0	6	85.3
	押す	1	71	0	0	0	95.9
	スワイプ下	2	1	65	3	3	87.4
	円を描く	0	1	5	66	0	91.4
	スワイプ裏	5	0	3	2	61	86.2
		予測ラベル					F値 [%]
正解ラベル	押し上げる	67	1	2	2	1	92.4
	押す	1	67	2	1	0	93.3
	スワイプ右	2	4	67	2	0	90.5
	スワイプ左	1	1	2	67	3	90.8
	円を描く	1	0	0	1	71	95.9

図 3.28: 混同行列 a: ボタン部実験 b: ポケット部実験.

### 3.6.1 ノイズ環境での AAS の動的利用に関する議論

#### AAS の動的利用に関する課題

本手法はノイズ環境での AAS の動的利用手法を考案し、これまで未調査であった衣服に対して AAS を適用する手法を提案した。5 種類のジェスチャ認識実験の結果、ノイズ環境 (1 dB) における認識精度は本手法 (可聴域帯域+全帯域) が全体平均で 89.5%であったのに対して、可聴域帯域のみによる認識精度は 85.2%であることが確認された。したがって、AAS を利用することによって 4.3 ポイント精度が向上することがわかった。また、全てのノイズ・衣服・SNR の組み合わせ (全 64 パターン) において本手法の認識精度が可聴域帯域のみによる認識精度よりも高いことを確認した (平均 3.6 ポイント)。したがって、本手法によるノイズ頑健性の強化は有効であることがわかった。

一方、ノイズなし環境での認識精度は本手法と可聴域帯域のみでの認識精度はそれぞれ 95.9%と 95.0%であり、ノイズあり環境と比較して、認識精度の向上の寄与率は低い。AAS の利用には超音波信号の再生と録音が必要であり、PAS と比較して高いサンプリングレートでの再生・録音および信号再生のための電力消費の増加がデメリットである。したがって、本手法の有効性を評価するために、これらの課題と AAS の利用による精度向上のトレードオフを考慮した調査が今後の課題であると考えられる。

上記の電力消費に関する課題に対して、本手法の認識アルゴリズムを更に改善する手法が考えられる。例えば、ノイズなし環境では AAS を利用せず信号の再生を停止し、ノイズ環境であることを検知した場合 AAS を利用する改善案が考えられる。このシステム改善によって、信号の再生時間を減らすことができ、電力消費を抑えることができると考えられる。

#### 認識アルゴリズムの改善

AAS の動的利用によってノイズ環境下での認識精度の低下を抑えられることが確認できた。今後はノイズ環境での精度低下を更に抑えるために、特徴量の信頼度を調整する手法を認識アルゴリズムに組み込むことを予定している。Mitake ら [105] は足音と足の加速度を用いた路面状況認識手法を提案しているが、ノイズ環境での認識精度の低下を防ぐために加速度により得られる特徴量の重要度を調整するアルゴリズムを提案している。本手法では、AAS の情報が含まれる全帯域の特徴量の重要度をノイズ環境にて調整することで認識精度の低下を更に抑えることが可能であるか調査する予定である。

### 3.6.2 衣服上タッチジェスチャ認識手法に関する議論

#### ジェスチャ認識精度

本研究で調査した 11 種類のジェスチャにおいて、衣服全体で最も認識精度の高いタッチジェスチャは「押す」で 95.6%、低いタッチジェスチャは「ピンチアウト」で 77.7%であった。まず、「押す」の認識精度が高かった理由は、得られる衣擦れ音がパルス波のような波形であり、他のジェスチャで得られる音響データと大きく異なったためであると考えられ

る。続いて、認識精度の低かった「ピンチアウト」は図 3.17 から示されるように「スワイプ右/左やピンチイン」と誤認識する傾向が高い。これらのジェスチャはマイクとスピーカを結ぶ線分に対して平行に指先をなぞる動作であり、得られるスペクトログラムの分布が似ていたため、誤認識が多くなったと考えられる。本研究ではこのような誤認識しやすいジェスチャの組み合わせの一方を避けることで認識精度を改善した。

### 衣服と認識精度の関係

衣服ごとのジェスチャ全体の認識精度平均を一元配置分散分析にて調査したが、有意差は示されなかった。しかし、ジェスチャ同士の混同が多かったスワイプ左右/ピンチインアウトの4種類の認識精度平均を確認すると、各衣服（シャツ A、シャツ B、フリース、ジャケット）の認識精度はそれぞれ 81.9%、72.4%、80.3%、80.7%であり、一元配置分散分析の結果、有意差が確認された ( $p = 0.0177$ )。事後多重比較として行った Tukey HSD 検定によると、シャツ A とシャツ B ( $p = 0.0211$ ) およびシャツ B とジャケット ( $p = 0.0458$ ) に有意差があることが確認された。図 3.29 に各衣服のスワイプ右とスワイプ左で得られたスペクトログラムを示す。衣服によって得られるスペクトログラムの音量が異なることが分かる。本来、マイクに近づく動作であるスワイプ右は衣擦れ音が大きくなり、逆の動作であるスワイプ左は音量が小さくなる。一方で、精度が比較的低かったシャツ B のスペクトログラムでは左右差による減衰・増幅の特性が不明瞭であった。これは、本研究では全ての衣服に対して同じデバイス距離で実験を行ったため、衣服ごとの衣擦れ音の特性を考慮していなかったことが原因であると考えられる。図 3.30 にシャツ B に対してデバイス距離を 9 cm, 11 cm, 13 cm, 15 cm に変更した場合の、各距離でのスワイプ左右のスペクトログラムをまとめた。測定の結果、デバイス距離が大きくなるほど、減衰や増幅が明瞭になっている。一方で、超音波信号の伝播強度も低下しており、ノイズ環境での認識精度の低下が懸念される。今後はデバイス距離と認識精度の関係をノイズなし/あり環境にて調査し、衣服ごとの最適なデバイス距離を決定する必要がある。

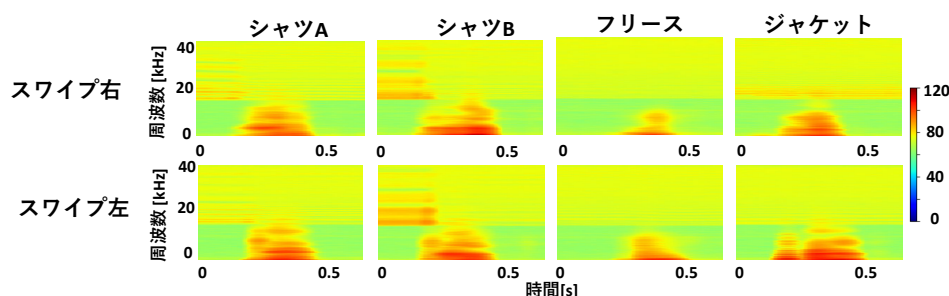


図 3.29: 各衣服のスペクトログラム (スワイプ右/左)。

また、個人分類モデルの認識精度と比較して、各衣服（シャツ A、シャツ B、フリース、ジャケット）の一般分類モデルの平均認識精度はそれぞれ 7.1, 5.8, 6.3, 5.6 ポイント低下することがわかった。一般分類モデルはユーザに事前のデータ測定を必要としなくなるため、ユーザビリティの向上が可能となるメリットがある。今後は実験参加者を更に募ることで、



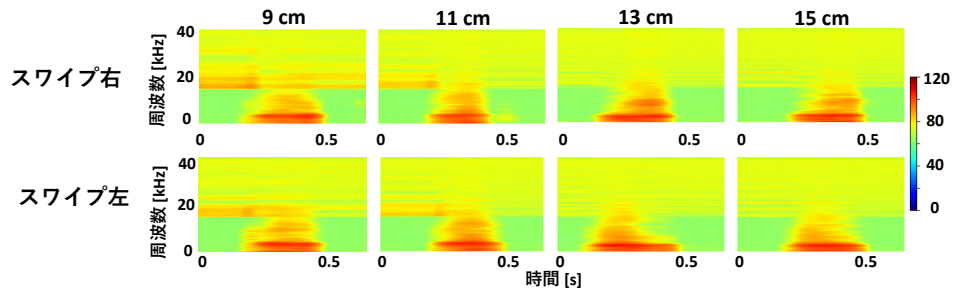


図 3.30: デバイス距離を変化させたときのスワイプ右/左のスペクトログラム (シャツ B).

より多くのジェスチャデータを収集し、深層学習のような大規模データを用いた機械学習手法を適用し、一般分類モデルの認識精度の向上を目指す。

### 3.7 まとめ

本研究では衣服上のタッチジェスチャ時に発生する生地の変形と衣擦れ音に着目し、AAS と PAS を併用することでノイズに頑健な衣服インタフェース構築手法を提案した。実験の結果、AAS はノイズ環境下で平均 3.6 ポイントの精度向上に寄与することがわかった。一方、ノイズのない環境での精度向上は 0.9 ポイントであり向上率は低かった。AAS の利用には信号の再生が必要になるため、PAS 単体の利用と比較して電力消費が大きくなる。したがって、ノイズレベルに応じて動的に AAS の利用を切り替えることで電力消費を効率的に抑えながら認識精度の維持が可能であると考えられる。本研究は形状が不安定な衣服に着目し、AAS の動的利用によるノイズに頑健な衣服インタフェース構築手法を提案した。本手法は、衣服だけでなく人形やソファなどの布生地で作られる物体に適用可能だと考えられ、AAS の利用可能範囲の拡張に貢献した。

## 第4章 結論

本章では本論文のまとめと本研究の貢献について述べ、最後に DAAS の検討課題および発展をまとめる。

### 4.1 本論文のまとめ

本研究では、これまでの AAS にて未解決であった装着誤差と不安定なセンシング対象への AAS の適用の 2 つの課題を解消した動的アクティブ音響センシング (DAAS) を提案し、ジェスチャ認識手法にて有効性を評価した。

#### 4.1.1 動的信号補正による装着誤差の低減

ジェスチャ認識の誤認識の要因である装着誤差を低減するために、測定信号を動的に補正する手法を提案した。本研究ではイヤホン型のウェアラブルデバイスであるヒアラブルデバイスのための頭部状態認識手法に本手法を適用し、調査を行った。実験では、11 名の参加者より 21 種類の頭部状態のデータの収集を行った。実験の結果、参加者ごとの個人分類モデルによる 6 種類の頭部状態認識の結果、未補正下での平均認識精度が 74.4% (F 値) であったのに対して、信号補正を行うことで平均認識精度が 90.0% (F 値) まで改善されることを示した。

#### 4.1.2 ノイズ環境での AAS の動的利用

不安定なセンシング対象へ AAS を適用は調査されていなかった課題に対して、ノイズ環境での AAS の動的利用という新しい AAS の利用手法を提案した。本研究では衣服上のタッチインタフェース構築手法に本手法を適用し、調査を行った。実験では、12 名の参加者より 11 種類の前腕部分でのタッチジェスチャのデータの収集を行った。実験の結果、ノイズなし環境での 5 種類のタッチジェスチャ認識精度が 95.9% (F 値) であることを確認した。また、AAS の動的利用によってノイズ環境での精度が PAS のみの精度と比較して平均で 3.6 ポイント改善できることを示した。

### 4.2 本研究の貢献

これらの研究にて、動的アクティブ音響センシング手法を確立した。本論文の貢献は以下のとおりである。

- 動的信号補正を用いた装着誤差の低減手法を提案し、認識システムの高精度化に貢献した。
- ノイズ環境での AAS の動的利用手法を提案し、AAS の適用可能範囲の拡大に貢献した。

### 4.3 検討課題および発展

以上にて、これまでの AAS 手法の課題を解決する DAAS の確立に関する研究成果を述べた。しかし、DAAS の実用化のためには、考慮すべき点および更なる発展的な調査が課題として残されている。本節では DAAS における今後の検討課題および発展について述べる。

#### 4.3.1 検討課題

DAAS の実用化には以下の検討課題についての調査および実験が必要である。

##### 超音波利用のガイドライン制定

超音波は医療機器および通信技術 [106] として実用されているが、それぞれの用途において人体への影響調査やガイドライン制定が行われており、筆者の知る限りウェアラブルコンピュータの用途に関する超音波利用ガイドラインは制定されていない。現在、最も近い分野としてパラメトリックスピーカ利用における超音波暴露の人体への影響が検討されている [107, 108, 109]。しかし、パラメトリックスピーカにおける検討ではデバイスの特性上、周波数帯域が一定の信号かつ、音源の位置がユーザから数十 cm 以上離れていることが前提となっている。一方、ウェアラブルデバイスにて超音波を利用する場合、周波数が変化するスイープ信号をユーザに近い距離 (0~数 cm) で再生することになる。今後、既存の調査を参考にしながら、ウェアラブルデバイスでの超音波利用を想定した場合の超音波暴露について検討し、利用ガイドラインを制定する必要がある。

##### モバイルデバイスでの実装および調査

提案手法は AAS で利用する測定信号を動的に補正、またはノイズ環境で動的に利用することでこれまでの AAS の課題を解決する手法である。一方で、これまでの AAS ではスピーカからの信号は一定であったが、動的 AAS はマイクからの收音データを基に信号の補正や再生の切り替えを行う必要がある。今後は、超音波信号の再生・録音が可能なマイクロコンピュータ [110] を用いて動的な信号補正や再生・停止の切り替えシステムを実装し、実用環境での調査を行っていく必要がある。

#### 4.3.2 発展

本研究ではジェスチャ認識手法における DAAS の確立を目指し、評価実験を行った。今後の発展として、DAAS が応用可能なシステムや研究について述べる。

#### 動的信号補正手法の応用範囲の拡大

本研究の頭部状態認識手法以外に、AASを利用したジェスチャ認識に関する関連研究 [111, 35] に提案手法は適用可能であると考えている。また、ジェスチャ認識手法にとどまらず、ユーザの運動記録や動作認識などの人体にAASを適用する研究 [34, 26] や物体の状態などを推定する研究 [3, 69] においてもデバイスの位置ずれが生じると考えられる。これらの研究に対して提案手法を適用した場合の精度改善や課題を発見し、動的信号補正手法の応用範囲の拡大を目指す。

#### AASの動的利用手法の応用範囲の拡大

本研究では布で構成された衣服に着目し、AASの動的利用手法の有効性を評価した。今後の発展として、布で構成されたその他のオブジェクト（人形やソファなど）に提案手法を適用することが可能であると考えている。人形やソファに音響センシングを適用することで、ユーザのタッチ入力が可能となるが、これまでの関連研究 [112, 113] では音響センシングは利用されていなかった。例えば、人形に提案手法が適用可能となれば、ユーザが愛用している人形とのタッチインタラクションが可能となると同時に、提案手法が利用していないスピーカの可聴帯域を音声インタラクションとして利用することが期待できる。今後は、更に多くのセンシング対象に提案手法を適用し、ジェスチャ認識手法にとどまらない応用展開を目指す。

## 謝辞

本研究を進めるにあたって、これまで多くの方々に支えられてきました。

まず、指導教員である、筑波大学 IPLAB の志築文太郎教授に感謝申し上げます。志築先生には研究生生活から研究内容まで多くのサポートをしていただきました。また、研究者や教育者としての姿勢を近くで学ぶことができました。COVID-19 の流行と共に始まったつくばでの研究生生活において、3年間で卒業できたのはなによりも志築先生のサポートのおかげです。ありがとうございました。

学士、修士課程在籍時にお世話になり、現在も研究ミーティングを続けてくださった、北海道大学情報科学研究科の杉本雅則教授、渡邊拓貴助教に心より感謝いたします。杉本先生の研究への姿勢や的確なご指導は、研究者として多くのことを学ばせていただきました。渡邊先生と出会えたことでウェアラブルコンピューティング研究の世界に出会い、多くの経験を積むことができました。研究に関する議論や新しいテーマについて話す時間はとても幸せな時間でした。以上、3名の先生方には6年間の研究生生活で論文執筆から研究方法などについて多大な時間を割いていただき、研究指導していただきました。改めてお礼申し上げます。先生方から学んだことをこれからも後輩や学生に伝えていけるよう、研究者として邁進してまいります。

また、IPLAB でお世話になりました、高橋伸准教授、ヴァシラケ シモナ助教、川口一画助教には普段の研究やゼミ活動および春合宿など、様々な場面にて、多くのご助言・ご指導をいただきました。心より感謝いたします。これまでの研究生生活でお世話になった、先輩・同期・後輩の皆様にもお礼申し上げます。実験参加や研究議論、進路相談などたくさんのサポートを受け、楽しく充実した研究生生活を過ごすことができました。特に、筑波大学の唯一の同期として一緒に頑張った磯本君には、慣れないつくばでの生活のサポートをしていただきありがとうございます。また、お互いの研究について議論したり、切磋琢磨できたことはとても楽しかったです。また、研究室秘書の新濱さん、高浜さん、鳥羽さん含め大学職員の皆様には大学生活や事務処理などの様々な場面で大変お世話になりました。心より感謝致します。

本論文の審査に際して、主査として志築文太郎教授、副査として滝沢穂高教授、三谷純教授、山田武史准教授、渡邊拓貴助教に引き受けていただき、予備審査・本審査の際には様々なご意見とアドバイスをいただきまして、深く感謝申し上げます。

最後に、大学進学をサポートしてくれ、今も研究を応援してくれる両親と家族に心より感謝致します。自身の研究者としての素養は何よりも楽しかった子供時代の食卓や日常生活での会話のおかげだと思っています。

改めて、みなさまに感謝申し上げます。ありがとうございました。

## 参考文献

- [1] Hiroki Watanabe, Tsutomu Terada, and Masahiko Tsukamoto. Gesture Recognition Method Utilizing Ultrasonic Active Acoustic Sensing. *Journal of Information Processing*, Vol. 25, pp. 331–340, 2017.
- [2] Adiyana Mujibiya, Xiang Cao, Desney S. Tan, Dan Morris, Shwetak N. Patel, and Jun Rekimoto. The Sound of Touch: On-Body Touch and Gesture Sensing Based on Transdermal Ultrasound Propagation. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces, ITS '13*, pp. 189–198. ACM, 2013.
- [3] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. Touch & Activate: Adding Interactivity to Existing Objects Using Active Acoustic Sensing. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13*, pp. 31–40. ACM, 2013.
- [4] Julia E. Mück, Barış Ünal, Haider Butt, and Ali K. Yetisen. Market and Patent Analyses of Wearables in Medicine. *Trends in Biotechnology*, Vol. 37, No. 6, pp. 563–566, 2019.
- [5] Google LLC. fitbit. <https://www.fitbit.com/global/jp/home>, 2007.
- [6] Apple Inc. . AppleWatch. <https://www.apple.com/jp/watch/>, 2015.
- [7] James Knight and Chris Baber. A Tool to Assess the Comfort of Wearable Computers. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 47, No. 1, 2005.
- [8] Hassan Mohammed, Michael W Mather, Joanna Lumb, Chris C Butler, and Janet A Wilson. Otitis Externa: What Is the Problem with Getting It Right? A Mixed-Methods Study in Primary and Secondary Care. *The Journal of Laryngology & Otolaryngology*, Vol. 136, No. 6, pp. 486–491, 2022.
- [9] Patrick Parzer, Adwait Sharma, Anita Vogl, Jürgen Steimle, Alex Olwal, and Michael Haller. SmartSleeve: Real-time Sensing of Surface and Deformation Gestures on Flexible, Interactive Textiles, Using a Hybrid Gesture Detection Pipeline. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, UIST '17*, pp. 565–577, 2017.

- [10] Artem Dementyev, Tomás Vega Gálvez, and Alex Olwal. SensorSnaps: Integrating Wireless Sensor Nodes into Fabric Snap Fasteners for Textile Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST '19*, pp. 17–28. ACM, 2019.
- [11] Tony Wu, Shiho Fukuhara, Nicholas Gillian, Kishore Sundara-Rajan, and Ivan Poupyrev. ZebraSense: A Double-Sided Textile Touch Sensor for Smart Clothing. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, UIST '20*, pp. 662–674. ACM, 2020.
- [12] Kyosuke Futami, Kohei Oyama, and Kazuya Murao. Augmenting Ear Accessories for Facial Gesture Input Using Infrared Distance Sensor Array. *Electronics*, Vol. 11, No. 9, 2022.
- [13] Damien Brun and Jonna Häkkinen. VäriWig: Interactive Coloring Wig Module. In *Proceedings of the 2021 International Symposium on Wearable Computers, ISWC '21*, pp. 166–169. ACM, 2021.
- [14] Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. CyclopsRing: Enabling Whole-Hand and Context-Aware Interactions Through a Fisheye Ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology, UIST '15*, pp. 549–556. ACM, 2015.
- [15] Katie A. Siek, Yvonne Rogers, and Kay H. Connelly. Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In *Proceedings of the 2005 IFIP TC13 International Conference on Human-Computer Interaction, INTERACT '05*, pp. 267–280. Springer-Verlag, 2005.
- [16] Jun Gong, Zheer Xu, Qifan Guo, Teddy Seyed, Xiang 'Anthony' Chen, Xiaojun Bi, and Xing-Dong Yang. WrisText: One-Handed Text Entry on Smartwatch Using Wrist Gestures. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18*, pp. 1–14. ACM, 2018.
- [17] Timo Götzelmann and Pere-Pau Vázquez. InclineType: An Accelerometer-Based Typing Approach for Smartwatches. In *Proceedings of the XVI International Conference on Human Computer Interaction, Interacción '15*, pp. 1–4. ACM, 2015.
- [18] Kirk H. Shelley. Photoplethysmography: Beyond the Calculation of Arterial Oxygen Saturation and Heart Rate. *Anesthesia & Analgesia*, Vol. 105, pp. S31–S36, 2007.
- [19] 久保勇貴, 江口佳那, 青木良輔, 近藤重邦, 東正造, 犬童拓也. 内部構造パターンの差異を利用した3Dプリントオブジェクト識別手法. 第27回インタラクティブシステムとソフトウェアに関するワークショップ論文集, WISS '19, pp. 43–48. 日本ソフトウェア科学会, 2019.

- [20] Wenjie Ruan, Quan Z. Sheng, Lei Yang, Tao Gu, Peipei Xu, and Longfei Shanguan. AudioGest: Enabling Fine-Grained Hand Gesture Detection by Decoding Echo Signal. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '16, pp. 474–485. ACM, 2016.
- [21] Xuhai Xu, Haitian Shi, Xin Yi, WenJia Liu, Yukang Yan, Yuanchun Shi, Alex Mariakakis, Jennifer Mankoff, and Anind 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.
- [22] Tian Hao, Guoliang Xing, and Gang Zhou. Isleep: Unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, SenSys '13, pp. 1–14. ACM, 2013.
- [23] Yuki Kubo, Yuto Koguchi, Buntarou Shizuki, Shin Takahashi, and Otmar Hilliges. AudioTouch: Minimally Invasive Sensing of Micro-Gestures via Active Bio-Acoustic Sensing. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '19, pp. 1–13. ACM, 2019.
- [24] Wentao Xie, Qian Zhang, and Jin Zhang. Acoustic-Based Upper Facial Action Recognition for Smart Eyewear. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 5, No. 2, 2021.
- [25] Yasha Iravantchi, Mayank Goel, and Chris Harrison. BeamBand: Hand Gesture Sensing with Ultrasonic Beamforming. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–10. ACM, 2019.
- [26] 雨坂宇宙, 渡邊拓貴, 杉本雅則, 志築文太郎. ピエゾ素子を用いた筋力トレーニング支援システムのための基礎検討. マルチメディア, 分散, 協調とモバイルシンポジウム 2022 論文集, 第 2022 巻, pp. 1610–1617. 情報処理学会, jul 2022.
- [27] Yasha Iravantchi, Yang Zhang, Evi Bernitsas, Mayank Goel, and Chris Harrison. Interferi: Gesture Sensing Using On-Body Acoustic Interferometry. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–13. ACM, 2019.
- [28] Hiroki Watanabe, Hiroaki Kakizawa, and Masanori Sugimoto. User Authentication Method Using Active Acoustic Sensing. *Journal of Information Processing*, Vol. 29, pp. 370–379, 2021.
- [29] Stefan Schneegass, Youssef Oualil, and Andreas Bulling. SkullConduct: Biometric User Identification on Eyewear Computers Using Bone Conduction Through the Skull. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pp. 1379–1384. ACM, 2016.



- [30] Gierad Laput, Robert Xiao, and Chris Harrison. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pp. 321—333. ACM, 2016.
- [31] 深堀孔明, 坂本大介, 五十嵐健夫. 靴下型圧力センサを用いた足裏ジェスチャ. コンピュータ ソフトウェア, Vol. 33, No. 2, pp. 116–124, 2016.
- [32] Kyosuke Futami, Yuki Tabuchi, Kazuya Murao, and Tsutomu Terada. Exploring Gaze Movement Gesture Recognition Method for Eye-Based Interaction Using Eyewear with Infrared Distance Sensor Array. *Electronics*, Vol. 11, No. 10, 2022.
- [33] Jonghwa Kim, Stephan Mastnik, and Elisabeth André. EMG-Based Hand Gesture Recognition for Realtime Biosignal Interfacing. In *Proceedings of the 13th International Conference on Intelligent User Interfaces*, IUI '08, pp. 30–39. ACM, 2008.
- [34] Naoki Kimura, Michinari Kono, and Jun Rekimoto. SottoVoce: An Ultrasound Imaging-Based Silent Speech Interaction Using Deep Neural Networks. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–11. ACM, 2019.
- [35] 横田智大, 橋田朋子. 人体を伝送路としたアクティブ音響センシングによるハンドジェスチャおよび肌上へのタッチの認識. 情報処理学会論文誌, Vol. 59, No. 2, pp. 703–714, 02 2018.
- [36] Yuya Okawa and Kentaro Takemura. Haptic-Enabled Active Bone-Conducted Sound Sensing. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, UIST '15 Adjunct, pp. 87—88. ACM, 2015.
- [37] Nobuhiro Funato and Kentaro Takemura. Estimating Contact Force of Fingertip and Providing Tactile Feedback Simultaneously. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16 Adjunct, pp. 195—196. ACM, 2016.
- [38] Apple Inc. . AirPodsPro. <https://www.apple.com/jp/airpods/>, 2016.
- [39] Timekettle. TIMEKETTLE M2. <https://timekettle.jp/products/timekettle-m2>, 2016.
- [40] Kazuhiro Taniguchi, Hisashi Kondo, Mami Kurosawa, and Atsushi Nishikawa. Earable TEMPO: A Novel, Hands-Free Input Device that Uses the Movement of the Tongue Measured with a Wearable Ear Sensor. *Sensors*, Vol. 18, No. 3, 2018.
- [41] Toshiyuki Ando, Yuki Kubo, Buntarou Shizuki, and Shin 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.

- [42] Denys J. C. Matthies, Bernhard A. Strecker, and Bodo 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.
- [43] Bragi. DashPro. <https://www.bragi.com/>, 2014.
- [44] NAP Enterprise Co., Ltd. . inCore . <http://www.incore.jp/>, 1983.
- [45] Jiayao Tan, Xiaoliang Wang, Cam-Tu Nguyen, and Yu Shi. SilentKey: A New Authentication Framework Through Ultrasonic-based Lip Reading. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 2, No. 1, pp. 36:1–36:18, March 2018.
- [46] Kentaro Takemura, Akihiro Ito, Jun Takamatsu, and Tsukasa Ogasawara. Active Bone-Conducted Sound Sensing for Wearable Interfaces. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology*, UIST '11 Adjunct, pp. 53–54. ACM, 2011.
- [47] Marko Hiipakka. Measurement Apparatus and Modelling Techniques of Ear Canal Acoustics. *Espoo: Helsinki University of Technology*, 2008.
- [48] 佐藤史明. Swept-Sine 法に基づく音響伝播測定. 音響学会誌, Vol. 63, No. 6, pp. 322–327, 2007.
- [49] Ton H. M. Akkermans, Tom A. M. Kevenaer, and Daniel W. E. Schobben. Acoustic Ear Recognition. In David Zhang and Anil K. Jain, editors, *Advances in Biometrics*, pp. 697–705. Springer Berlin Heidelberg, 2005.
- [50] Takayuki Arakawa, Takafumi Koshinaka, Shohei Yano, Hideki Irisawa, Ryoji Miyahara, and Hitoshi Imaoka. Fast and Accurate Personal Authentication Using Ear Acoustics. In *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, APSIPA '16, pp. 1–4. IEEE, 2016.
- [51] Shohei Yano, Takayuki Arakawa, Takafumi Koshinaka, Hitoshi Imaoka, and Hideki Irisawa. Improving Acoustic Ear Recognition Accuracy for Personal Identification by Averaging Biometric Data and Spreading Measurement Error over a Wide Frequency Range. *IEICE Transactions on Electronics*, Vol. J100-A, pp. 161–168, 2017.
- [52] Jeffrey Borish and James B. Angell. An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise. *Journal of the Audio Engineering Society*, Vol. 31, No. 7/8, pp. 478–488, 1983.
- [53] Yang Gao, Wei Wang, Vir V. Phoha, Wei Sun, and Zhanpeng Jin. EarEcho: Using Ear Canal Echo for Wearable Authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3, No. 3, 2019.

- [54] Zi Wang, Sheng Tan, Linghan Zhang, Yili Ren, Zhi Wang, and Jie Yang. EarDynamic: An Ear Canal Deformation Based Continuous User Authentication Using In-Ear Wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 5, No. 1, 2021.
- [55] Dhruv Verma, Sejal Bhalla, Dhruv Sahnan, Jainendra Shukla, and Aman Parnami. Expresrear: Sensing fine-grained facial expressions with earables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 5, No. 3, sep 2021.
- [56] Seokmin Choi, Yang Gao, Yincheng Jin, Se jun Kim, Jiyang Li, Wenyao Xu, and Zhanpeng Jin. Ppgface: Like what you are watching? earphones can "feel" your facial expressions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 6, No. 2, jul 2022.
- [57] Abdelkareem Bedri, David Byrd, Peter Presti, Himanshu Sahni, Zehua Gue, and Thad 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.
- [58] Apple Inc. . AirPodsPro. <https://www.apple.com/jp/airpods-pro/>, 2019.
- [59] Howard Lei and Eduardo López 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.
- [60] 祁君容. 顎運動時に起こる外耳道のひずみと下顎頭運動の相関関係. PhD thesis, 松本歯科大学, 2016.
- [61] 橘秀樹, 矢野博夫. 環境騒音・建築音響の測定, 5.2 インパルス応答の測定方法. コロナ社, 2004.
- [62] SONY. h.ear in NC. <https://www.sony.jp/headphone/products/MDR-EX750NA/>, 2016.
- [63] Tobias Röddiger, Tobias King, Dylan Ray Roodt, Christopher Clarke, and Michael Beigl. OpenEarable: Open Hardware Earable Sensing Platform. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp ' 22, pp. 29—34. ACM, 2022.
- [64] Wei Sun, Franklin Mingzhe Li, Benjamin Steeper, Songlin Xu, Feng Tian, and Cheng Zhang. Teethtap: Recognizing discrete teeth gestures using motion and acoustic sensing on an earpiece. IUI '21, pp. 161—169. ACM, 2021.

- [65] Patrick Parzer, Florian Perteneder, Kathrin Probst, Christian Rendl, Joanne Leong, Sarah Schuetz, Anita Vogl, Reinhard Schwoediauer, Martin Kaltenbrunner, Siegfried Bauer, and Michael Haller. RESi: A Highly Flexible, Pressure-Sensitive, Imperceptible Textile Interface Based on Resistive Yarns. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, pp. 745–756. ACM, 2018.
- [66] Md. Tahmidul Islam Molla, Crystal Compton, and Lucy E. Dunne. Launderability of Surface-Insulated Cut and Sew E-Textiles. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, ISWC '18, pp. 104–111. ACM, 2018.
- [67] Tim Collins. Active Acoustic Touch Interface. *Electronics Letters*, Vol. 45, No. 20, 2009.
- [68] John Lamme Stewart and E. C. Westerfield. A Theory of Active Sonar Detection. *Proceedings of the Institute of Radio Engineers*, Vol. 47, No. 5, 1959.
- [69] 岩瀬大輝, 伊藤雄一, 秦秀彦, 尾上孝雄. Sensesurface : アクティブ音響センシングによる物体識別と位置推定. *情報処理学会論文誌*, Vol. 60, No. 10, pp. 1869–1880, 2019.
- [70] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. Sensing Touch Force Using Active Acoustic Sensing. In *Proceeding of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '15, pp. 355–358. ACM, 2015.
- [71] Robert Xiao, Greg Lew, James Marsanico, Divya Hariharan, Scott Hudson, and Chris Harrison. Toffee: Enabling Ad Hoc, around-Device Interaction with Acoustic Time-of-Arrival Correlation. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '14, pp. 67–76. ACM, 2014.
- [72] Karsten Seipp and Kate Devlin. BackPat: One-Handed off-Screen Patting Gestures. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '14, pp. 77–80. ACM, 2014.
- [73] Mingshi Chen, Panlong Yang, Jie Xiong, Maotian Zhang, Youngki Lee, Chaocan Xiang, and Chang Tian. Your Table Can Be an Input Panel: Acoustic-Based Device-Free Interaction Recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3, No. 1, 2019.
- [74] Cheng Zhang, Anandghan Waghmare, Pranav Kundra, Yiming Pu, Scott Gilliland, Thomas Ploetz, Thad E. Starner, Omer T. Inan, and Gregory D. Abowd. Finger-Sound: Recognizing Unistroke Thumb Gestures Using a Ring. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 1, No. 3, 2017.

- [75] Roderick Murray-Smith, John Williamson, Stephen Hughes, and Torben Quaade. Stane: Synthesized Surfaces for Tactile Input. In *Proceedings of the 2008 CHI Conference on Human Factors in Computing Systems*, CHI '08, pp. 1299–1302. ACM, 2008.
- [76] Chris Harrison and Scott E. Hudson. Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile Finger Input Surfaces. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, pp. 205–208. ACM, 2008.
- [77] Chris Harrison, Desney Tan, and Dan Morris. Skininput: Appropriating the Body as an Input Surface. In *Proceedings of the 2010 CHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 453–462. ACM, 2010.
- [78] Maggie Orth, Rehmi Post, and Emily Cooper. Fabric Computing Interfaces. In *Proceedings of the 1998 CHI Conference on Human Factors in Computing Systems*, CHI '98, pp. 331–332. ACM, 1998.
- [79] Simon Ozbek, Md. Tahmidul Islam Molla, Crystal Compton, and Brad Holschuh. Novel Manufacturing of Advanced Smart Garments: Knitting with Spatially-Varying, Multi-Material Monofilament. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, ISWC '18, pp. 120–127. ACM, 2018.
- [80] Nur Al-huda Hamdan, Simon Voelker, and Jan Borchers. Sketch&Stitch: Interactive Embroidery for E-Textiles. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 1–13. ACM, 2018.
- [81] Leah Buechley, Mike Eisenberg, Jaime Catchen, and Ali Crockett. The LilyPad Arduino: Using Computational Textiles to Investigate Engagement, Aesthetics, and Diversity in Computer Science Education. In *Proceedings of the 2008 CHI Conference on Human Factors in Computing Systems*, CHI '08, pp. 423–432. ACM, 2008.
- [82] Ernest Rehmi Post, Maggie Orth, Peter R Russo, and Neil A Gershenfeld. E-Broidery: Design and Fabrication of Textile-Based Computing. *IBM Systems Journal*, Vol. 39, No. 3&4, 2000.
- [83] Alex Olwal, Jon Moeller, Greg Priest-Dorman, Thad Starner, and Ben Carroll. I/O Braid: Scalable Touch-Sensitive Lighted Cords Using Spiraling, Repeating Sensing Textiles and Fiber Optics. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*, UIST '18 Adjunct, pp. 203–207. ACM, 2018.
- [84] Yiyue Luo, Kui Wu, Tomás Palacios, and Wojciech Matusik. KnitUI: Fabricating Interactive and Sensing Textiles with Machine Knitting. In *Proceedings of the 2021*

- CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 1–12. ACM, 2021.
- [85] Ivan Poupyrev, Nan-Wei Gong, Shiho Fukuhara, Mustafa Emre Karagozler, Carsten Schwesig, and Karen E. Robinson. Project Jacquard: Interactive Digital Textiles at Scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pp. 4216–4227. ACM, 2016.
- [86] Roland Aigner, Andreas Pointner, Thomas Preindl, Rainer Danner, and Michael Haller. TexYZ: Embroidering Enameled Wires for Three Degree-of-Freedom Mutual Capacitive Sensing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 1–12. ACM, 2021.
- [87] Stefan Schneegass and Alexandra Voit. GestureSleeve: Using Touch Sensitive Fabrics for Gestural Input on the Forearm for Controlling Smartwatches. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, ISWC '16, pp. 108–115. ACM, 2016.
- [88] Thorsten Karrer, Moritz Wittenhagen, Leonhard Lichtschlag, Florian Heller, and Jan Borchers. Pinstripe: Eyes-Free Continuous Input on Interactive Clothing. In *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 1313–1322. ACM, 2011.
- [89] Florian Heller, Stefan Ivanov, Chat Wacharamanotham, and Jan Borchers. Fabri-Touch: Exploring Flexible Touch Input on Textiles. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, ISWC '14, pp. 59–62. ACM, 2014.
- [90] Kentaro Ueda, Tsutomu Terada, and Masahiko Tsukamoto. Evaluation of Input/Output Interface Using Wrinkles on Clothes. *Journal of Data Intelligence*, Vol. 1, No. 1, 2018.
- [91] Te-Yen Wu, Shutong Qi, Junchi Chen, MuJie Shang, Jun Gong, Teddy Seyed, and Xing-Dong Yang. Fabriccio: Touchless Gestural Input on Interactive Fabrics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–14. ACM, 2020.
- [92] Federico Lorussi, Walter Rocchia, Enzo Scilingo, Alessandro Tognetti, and Danilo de rossi. Wearable, Redundant Fabric-Based Sensor Arrays for Reconstruction of Body Segment Posture. *IEEE Sensors Journal*, Vol. 4, No. 6, 2004.
- [93] Guido Gioberto. Garment-Integrated Wearable Sensing for Knee Joint Monitoring. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program*, ISWC '14 Adjunct, pp. 113–118. ACM, 2014.

- [94] Guido Gioberto, James Coughlin, Kaila Bibeau, and Lucy E. Dunne. Detecting Bends and Fabric Folds Using Stitched Sensors. In *Proceedings of the 2013 International Symposium on Wearable Computers*, ISWC '13, pp. 53–56. ACM, 2013.
- [95] Corninne Mattmann, Oliver Amft, Holger Harms, Gerhard Troster, and Frank Clemens. Recognizing Upper Body Postures using Textile Strain Sensors. In *Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers*, ISWC '07, pp. 29–36. IEEE, 2007.
- [96] Ruibo Liu, Qijia Shao, Siqi Wang, Christina Ru, Devin Balkcom, and Xia Zhou. Reconstructing Human Joint Motion with Computational Fabrics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3, No. 1, pp. 19:1–19:26, 2019.
- [97] Emilio Sardini, Mauro Serpelloni, and Viviane Pasqui. Wireless Wearable T-Shirt for Posture Monitoring During Rehabilitation Exercises. *IEEE Transactions on Instrumentation and Measurement*, Vol. 64, No. 2, 2015.
- [98] Ali Kiaghadi, Morgan Baima, Jeremy Gummeson, Trisha Andrew, and Deepak Ganesan. Fabric as a Sensor: Towards Unobtrusive Sensing of Human Behavior with Triboelectric Textiles. *SenSys '18*, pp. 199–210. ACM, 2018.
- [99] Ali Kiaghadi, Seyedeh Zohreh Homayounfar, Jeremy Gummeson, Trisha Andrew, and Deepak Ganesan. Phyjama: Physiological Sensing via Fiber-Enhanced Pyjamas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3, No. 3, 2019.
- [100] Ji Jia, Chengtian Xu, Shijia Pan, Stephen Xia, Peter Wei, Hae Noh, Pei Zhang, and Xiaofan Jiang. Conductive Thread-Based Textile Sensor for Continuous Perspiration Level Monitoring. *Sensors*, Vol. 18, No. 11, 2018.
- [101] Pin-Sung Ku, Jun Gong, Te-Yen Wu, Yixin Wei, Yiwen Tang, Barrett Ens, and Xing-Dong Yang. Zippro: The Design and Implementation of An Interactive Zipper. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13. ACM, 2020.
- [102] Scott E. Hudson, Chris Harrison, Beverly L. Harrison, and Anthony LaMarca. Whack Gestures: Inexact and Inattentive Interaction with Mobile Devices. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '10, pp. 109–112. ACM, 2010.
- [103] Preferred Networks Inc. Optuna, 2018.
- [104] Laurens van der Maaten and Geoffrey Hinton. Visualizing Data Using t-SNE. *Journal of Machine Learning Research*, Vol. 9, No. 86, 2008.

- [105] Hiroto Mitake, Hiroki Watanabe, and Sugimoto Masanori. A Method for Recognizing Road Surface Condition based on Footsteps and Inertial Data. *Journal of Information Processing*, Vol. 61, No. 10, 2020.
- [106] YAMAHA CORPORATION. Infosound. <https://jp.yamaha.com/>, 2013.
- [107] Takahashi Hironobu. A Survey of Acoustic Standard for Airborne Ultrasound Especially Focused on Measurement Techniques. Technical report, National Institute of Advanced Industrial Science and Technology, 2006.
- [108] 鎌倉友男, 伊藤幹也, 野村英之. 超音波暴露の調査 : パラメトリックスピーカを話題にして (小特集 : 超音波暴露に関する研究の動向). *日本音響学会誌*, Vol. 67, No. 5, pp. 200–203, 2011.
- [109] 星貴之. 超音波曝露に関する再考. *日本音響学会誌*, Vol. 78, No. 9, pp. 508–513, 2022.
- [110] SONY. spresense. <https://developer.sony.com/ja/develop/spresense/>, 2018.
- [111] Cheng Zhang, Qiuyue Xue, Anandghan Waghmare, Sumeet Jain, Yiming Pu, Sinan Hersek, Kent Lyons, Kenneth A. Cunefare, Omer T. Inan, and Gregory D. Abowd. SoundTrak: Continuous 3D Tracking of a Finger Using Active Acoustics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 1, No. 2, 2017.
- [112] Suzanne Low, Yuta Sugiura, Kevin Fan, and Masahiko Inami. Cuddly: Enchant Your Soft Objects with a Mobile Phone. In Dennis Reidsma, Haruhiro Katayose, and Anton Nijholt, editors, *Advances in Computer Entertainment*, pp. 138–151. Springer International Publishing, 2013.
- [113] Yuri Suzuki, Kaho Kato, Naomi Furui, Daisuke Sakamoto, and Yuta Sugiura. Cushion Interface for Smart Home Control. TEI '20, pp. 467–472. ACM, 2020.



# 付録A 動的信号補正による装着誤差の低減

第2章の実験に用いた実際のアンケート用紙を付録として以下に示す。

## A.1 実験後アンケート

実験アンケートのお願い  
各頭部状態について、以下の質問に対して数字で答えてください。  
名前 ( )

1. 各頭部状態について、その状態を再現するのは簡単でしたか? ()内に、当てはまるものの数字を記入してください。






















1:非常に簡単  
2:簡単  
3:普通  
4:難しい  
5:非常に難しい

・口を閉じる( )・口を開ける( )・顎を右にずらす( )・顎を左にずらす( )  
・口角を上げる( )・両頬を膨らます( )・右頬を膨らます( )・左頬を膨らます( )  
・両目を閉じる( )・右目を閉じる( )・左目を閉じる( )・舌を出す[閉口]( )  
・舌を出す[開口]( )・右に首を傾ける( )・左に首を傾ける( )・下を向く( )  
・上を向く( )・右を向く( )・左を向く( )・舌を右内頬につける( )  
・舌を左内頬につける( )

2. 各頭部状態について、その状態を人前で再現することにどう感じますか? ()内に、当てはまるものの数字を記入してください。

1:気にならない  
2:少し気になる  
3:気になる  
4:やりたくない

・口を閉じる( )・口を開ける( )・顎を右にずらす( )・顎を左にずらす( )  
・口角を上げる( )・両頬を膨らます( )・右頬を膨らます( )・左頬を膨らます( )  
・両目を閉じる( )・右目を閉じる( )・左目を閉じる( )・舌を出す[閉口]( )  
・舌を出す[開口]( )・右に首を傾ける( )・左に首を傾ける( )・下を向く( )  
・上を向く( )・右を向く( )・左を向く( )・舌を右内頬につける( )  
・舌を左内頬につける( )

						
口を閉じる(閉口)	口を開ける(開口)	顎を右にずらす(右顎)	顎を左にずらす(左顎)	口角を上げる(口角)	両頬を膨らます(両頬)	右頬を膨らます(右頬)
						
左頬を膨らます(左頬)	両目を閉じる(両目)	右目を閉じる(右目)	左目を閉じる(左目)	舌を出す(舌閉口)	舌を出す(舌開口)	右に首を傾ける(右首傾)
						
左に首を傾ける(左首傾)	下を向く(下向き)	上を向く(上向き)	右を向く(右向き)	左を向く(左向き)	舌を右内頬につける(右内頬)	舌を左内頬につける(左内頬)

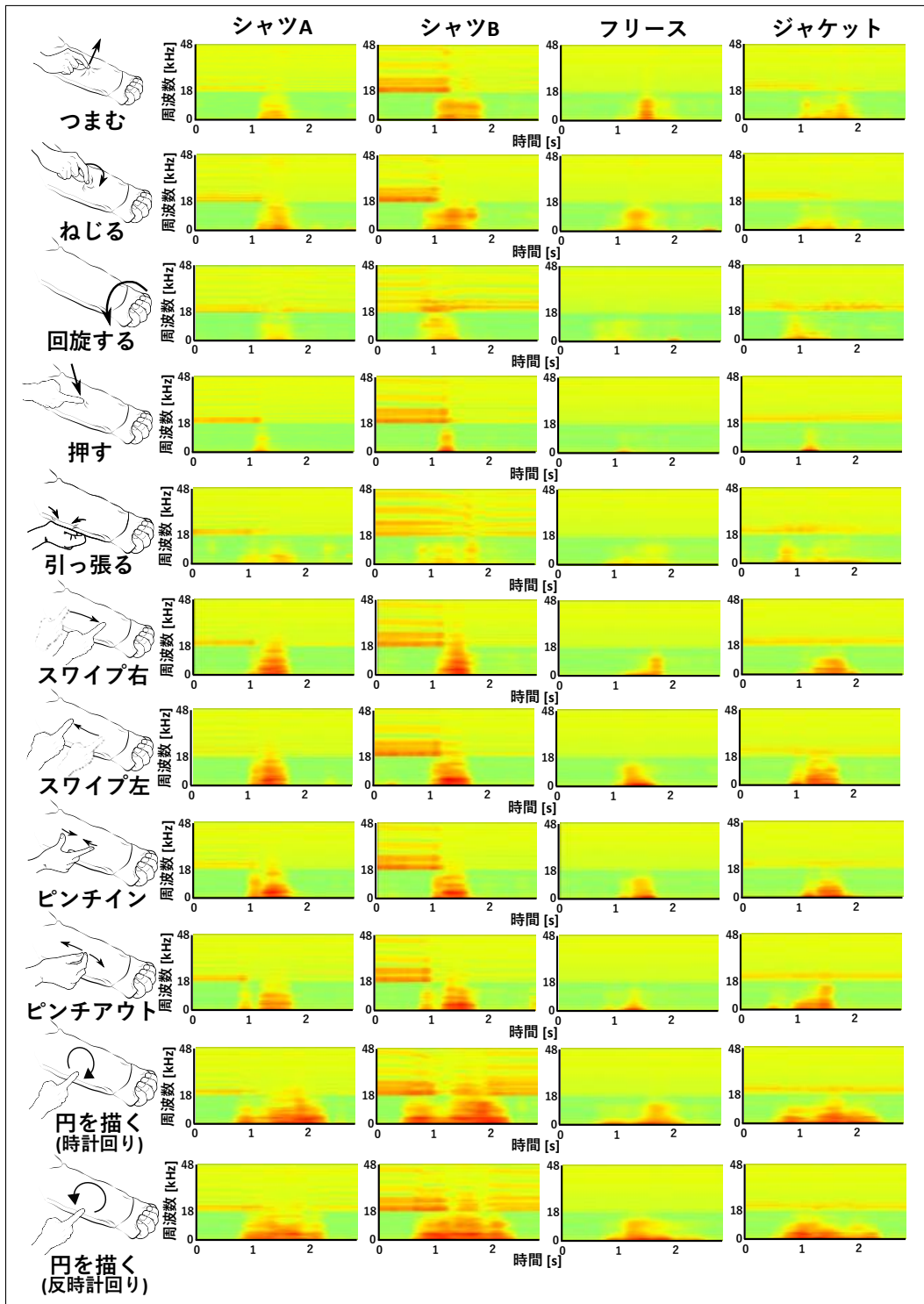
実験の感想があれば自由に記入をお願いします。

## 付 録 B ノイズ環境での AAS の動的利用

第 3 章の実験に用いた実際のアンケート用紙を付録として以下に示す.

### B.1 各ジェスチャのスペクトログラム

各衣服ごとのジェスチャ時のスペクトログラムの一覧を以下に示す.



## B.2 実験同意書

実験時には各参加者に以下の同意書に署名してもらった

<p>同 意 書</p> <p>筑波大学システム情報系長 殿</p> <p>私は、「課題名：衣服を用いたジェスチャ認識システムの評価」について、研究概要、方法、研究対象者の必要性、研究対象者に対するリスクと安全性、研究に参加する上で想定される危険の回避、ビデオの録画を含めた個人情報の保護について十分な説明を受けました。</p> <p>説明の際、本研究に協力することに同意しなくても何ら不利益を受けないこと、さらに、同意後も私自身の自由意思により不利益を受けず、参加の同意を撤回できることも理解しました。また、研究や実験の終了後であっても、枠下に指定の期日迄であれば、データ提供の同意を撤回できることも理解しました。私は、このことを理解した上で研究対象者になることに同意します。</p> <p>令和 年 月 日</p> <p style="text-align: right;">氏 名 _____ (署名)</p> <p>「課題名：衣服を用いたジェスチャ認識システムの評価」の研究について、次の内容について令和 年 月 日に説明を行い、上記のとおり同意を得ました。</p> <p style="text-align: center;">実施責任者 所 属 システム情報系 _____</p> <p style="text-align: right;">氏 名 志築 文太郎 (署名又記名押印)</p>
<p>データ提供の同意撤回の期限は同意書署名の日から_30_日後までとさせていただきます。</p> <p>研究や実験に協力した結果、不都合があった場合の連絡先 実施分担者(所属:筑波大学 情報理工学位プログラム 氏名:雨坂宇宙 TEL:029-853-5165) 実施責任者 (所属:筑波大学 システム情報系 氏名:志築文太郎 TEL:029-853-5520) 筑波大学 システム情報系 研究倫理委員会 事務局 (システム情報エリア支援室 TEL:029-853-4989)</p>

# 著者論文リスト

## 参考論文

本論文の主内容は、下記 (1)–(6) として公表済みである。

- 公表済み論文

- (1) Takashi Amesaka, Hiroki Watanabe, Masanori Sugimoto, Buntarou Shizuki “Gesture Recognition Method Using Acoustic Sensing on Usual Garment”, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), Vol. 6, No. 2, Article 41, 2022.
- (2) 雨坂 宇宙, 渡邊 拓貴, 杉本 雅則, 志築 文太郎 “外耳道伝達関数による頭部状態認識手法”, 情報処理学会論文誌, 情報処理学会, Vol. 61, No. 8, pp.1333-1342, 2020.
- (3) Takashi Amesaka, Hiroki Watanabe, and Masanori Sugimoto. “Facial Expression Recognition Using Ear Canal Transfer Function”, Proceedings of the 23rd International Symposium on Wearable Computers (ISWC), London, UK, pp.1-9, 2019.
- (4) 雨坂宇宙, 渡邊拓貴, 杉本雅則 “衣服へのアクティブ音響センシングを用いたジェスチャ認識技術の基礎検討”, 情報処理学会研究報告, Vol. 2019-UBI-64, No. 13, pp. 1-8, 2019.
- (5) 雨坂宇宙, 渡邊拓貴, 杉本雅則 “外耳道伝達関数を用いた頭部状況認識手法”, 情報処理学会研究報告, Vol. 2019-UBI-61, No. 7, pp. 1-8, 2019.
- (6) 雨坂宇宙, 渡邊拓貴, 杉本雅則 “外耳道特響特性を用いた頭部状態認識手法の提案”, 情報処理北海道シンポジウム 2018, 2018.

## その他の論文

- 公表済み論文

- (1) 雨坂宇宙, 渡邊拓貴, 杉本雅則, 志築文太郎 “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”, 第30回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2022) 予稿集, Article 12, 7 pages. ソフトウェア科学会, 宮城, 2022.

- 査読のない発表論文

- (6) 雨坂宇宙, 渡邊拓貴, 杉本雅則, 志築文太郎 “ piezo素子を用いた筋力トレーニング支援システムのための基礎検討”, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO 2022) 論文集, pp. 1610–1617, 2022.