

外耳道の気圧変化を利用したカナル型イヤホン用 タッチ入力インタフェースの設計と実装

村尾 和哉¹ 渡邊 拓貴²

概要: 本研究ではカナル型イヤホンに気圧センサを搭載して外耳道の気圧を計測するデバイスを実装し、イヤホンを指で押して離す操作に伴う外耳道の気圧変化のパターンから3種類の操作方法を区別して認識する手法を提案する。操作方法は、「素早く押して素早く離す」、「素早く押すが素早く離さない(ゆっくり離すか、押したままにしておく)」、「(押した状態から)素早く離す」の3種類である。

1. はじめに

ウェアラブルコンピューティングの発展にともない、ヒアラブルデバイス (hearable device) (またはイヤラブルデバイス (earable device)) [1] が着目されている。ヒアラブルデバイスとは耳に装着するイヤホン型のウェアラブルコンピュータである。従来のイヤホンの用途である音楽鑑賞をはじめ、スマートフォンと連携することによる音声アシスタントの利用や、搭載された各種センサを用いてユーザーの生体情報や行動を認識できる。現在市販されているヒアラブルデバイスの多くはスマートフォンと連携して使用することを前提としており、デバイス进行操作する際にスマートフォンの画面に着目し、手で操作する必要がある。

ヒアラブルデバイスにセンサを搭載して直接手でタッチしてコマンド操作できるデバイスとして、感圧センサを使用する AirPods*¹や静電容量センサを使用する Google Pixel Buds*²が販売されている。感圧センサはイヤホンの軸に搭載するためイヤホンの形状に制約が生じる。静電容量センサは意図せず軽く触れた場合も検出し、また手袋を装着していると反応しない。

また、ハンズフリー入力方式として音声アシスタントによるデバイス操作手法も存在するが、公共の場での発声や騒音による影響などの課題がある。加速度センサを用いて頭部の動きによってコマンド操作できる DashPro*³も販売されているが、周囲の人に奇妙な印象を与える社会的受容

性の課題がある。このほかのハンズフリー操作を実現する手法として筋電 [3], 赤外線 [4], [5] を用いた手法がある。しかし、赤外線は現時点では十分な認識精度を得られていない。また、電極は意図しない顔の動きも検出するためインタフェースとしての利用には適さない。

気圧センサを用いて表情の変化を検出する方式 [2] も存在するが、飛行機やエレベータ内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。一方で、手でイヤホンをタッチしたときに外耳道の気圧が変化する。これは、カナル型イヤホン装着時には外耳道は鼓膜とイヤピースで囲われた空間となり、イヤピースが素早く押されると外耳道の容積が小さくなり、気圧が瞬間的に上昇する。同様に、押した状態から手が素早く離れるとイヤピースは元の位置に戻ろうとするため外耳道の容積が大きくなり、気圧が瞬間的に低下する。本研究ではこの気圧変化に着目して手でイヤホンを押すことで入力を実現するインタフェースを提案する。手でイヤホンを押したときに発生する気圧変化は飛行機やエレベータなどの高度変化によって発生する気圧変化よりも極めて大きく容易に区別可能である。また、人の動きによってイヤピースも動く場合があるが、動きが遅いため大きな気圧変化は生じず誤動作は発生しない。静電容量方式と異なり、手袋をしていても動作する。

2. 関連研究

ヒアラブルデバイスにセンサを搭載して直接手でタッチしてコマンド操作できるデバイスとして、感圧センサで軸部分を握る操作を検出する AirPods が販売されている。イヤホンの軸部分に感圧センサを埋め込むため、AirPods のように必ず軸が必要となり、デザイン面での制約があ

¹ 立命館大学情報理工学部

² 北海道大学大学院情報科学研究院

*¹ Apple AirPodsPro, <https://www.apple.com/jp/airpods-pro/>

*² Google Pixel Buds, <https://store.google.com/jp/category/earbuds>

*³ Bragi DashPro, <https://www.bragi.com/>

る。このほか、静電容量センサを用いてタッチを検出する Google の Google Pixel Buds が販売されている。イヤホンのハウジング部にタッチセンサを搭載し、タップやスワイプを検出するが、静電容量方式だと手がイヤホンに触れたときに反応するため、意図的なタッチと偶発的なタッチの区別が困難である。また、手袋などを装着していると反応しない。このほか、加速度センサを用いる方式も考えられるが、手でイヤホンに触れたときの衝撃を検出するため、手が触れていなくてもタッチを誤検出する可能性がある。また、Google Pixel Buds は音声入力にも対応しているが、音声入力は利用シーンが限定される。

センサを用いてヒアラブルデバイスで頭部状態やジェスチャを認識する研究が行われている。Ando ら [2] 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。さらに、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。Matthies ら [3] は電極を用いて外耳道内部から筋肉の動きを読み取り 5 種類の頭部状態を精度 90.0% (座位状態), 85.2% (歩行状態) で認識することに成功している。Taniguchi ら [4] は LED とフォト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、音楽プレイヤー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 [5] 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。真鍋ら [6] は市販されているヘッドホンに簡単な回路を組み合わせるだけでヘッドホンをタップする動作を認識することに成功している。Laput ら [7] はイヤホンに組み込まれたスピーカとマイクから反響音を取得し、イヤホンの着脱状態を認識している。製品では、Bragi の DashPro が慣性センサを用いて頭部ジェスチャを認識し、電話の応答や音楽プレイヤーの曲変更などの簡単なハンズフリー操作を実現している。Amesaka ら [8] は外耳道内部の音を取得できるマイクを利用して、首、顎、顔の状態 (頭部状態) にともなって外耳道が変形することに着目し、スピーカから流した測定信号をもとに、外耳道伝達関数を測定、解析することで頭部状態を認識する手法を提案した。

上述した既存研究・製品は気圧、電極、光、慣性センサを用いて頭部状態認識を行っている。気圧センサを用いる手法では、高い精度で顔の動きの認識を行っているが、顔の動きにともなって発生する気圧変化は小さく、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、慣性センサを用いた手法では現段階において認識できる頭部状態の数や認識精度が十分とはいえない。また、電極を含めた既存研究では頭部の動作に伴う電位の変化を利用するため意図した動作か否かの判断が困難である。マイクを利用する手法では、外耳道内部の超音波を取得できるマイクを追加する必要がある。

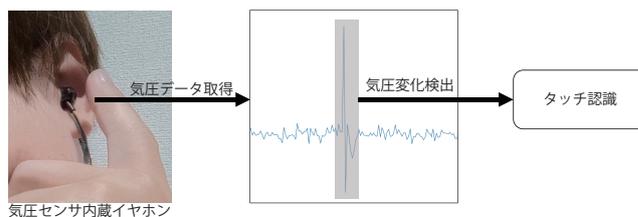


図 1 提案手法の概要

また、頭部の動作は日常生活で装着者の意図しない状況でも発生するため、頭部動作認識の結果をインタフェースとして利用すると誤入力の問題となる。

3. 提案手法

3.1 概要

提案手法の処理の流れを図 2 に示す。はじめに、イヤホンのハウジング部に格納された気圧センサからユーザの外耳道の気圧を取得する。イヤホンはカナル型であり、イヤピースによって外耳道への空気の入出は制限されているが、完全に密閉はされていない状態である。そして、ユーザがイヤホンのハウジング部を耳の中に軽く押し込むように指でタッチすると、イヤピースが外耳道に押し込まれるため外耳道の容積が小さくなり、空気が出ていく隙間が狭いため、瞬間的に気圧が上がる。しかし、完全に密閉されているわけではないため、すぐに空気が抜けて気圧は元に戻る。同様に、押した状態から素早く指を離すと外耳道の容積が大きくなり、瞬間的に気圧が下がる。このようにタッチのやり方によって気圧の変化が異なるため、取得した気圧値からタッチの種類を区別し、入力インタフェースとして動作する。

3.2 ハードウェア

本研究では CanalSense を参考にしてハードウェアを実装した。カナル型イヤホン (Panasonic 製 RP-HJE260)^{*4} を使用し、ハウジング部とノズル部が分離できる形状であるため、中を開けてスピーカを取り出し、空いた空間に気圧センサ (TDK Invensense 製 ICP-10125)^{*5} を搭載した。気圧センサの大きさは $3.55 \times 3.55 \times 1.45$ [mm] であるためイヤホンのハウジング部に余裕をもって収まる。気圧センサに接続するケーブルは本来のイヤホンのケーブルが通る穴を通してイヤホンの外に出て、Arduino Uno に接続される。圧力センサと Arduino Uno の通信は I2C である。気圧センサの計測範囲は $30 \sim 110$ [kPa], 1 気圧は 101325 Pa, 指でイヤホンを押したときに発生する気圧変化は $100 \sim 101$ [kPa] の範囲であったため、計測範囲は適切である。

^{*4} Panasonic Rp-HJE260, <https://panasonic.jp/headphone/p-db/RP-HJE260.html>

^{*5} TDK Invensense ICP-10125, <https://invensense.tdk.com/products/1-axis/icp-10125/>



図 2 実装した気圧センサ搭載カナル型イヤホン

3.3 気圧データ取得

時刻 t に取得した気圧データを $p(t)$ [Pa] とする。日時や場所によって何もしていない状態での気圧が異なるため、サイズ 100 サンプル (約 1 秒) のウィンドウ $[p(t-99), \dots, p(t)]$ の移動平均 $\bar{p}(t)$ [Pa] を計算し、 $p(t)$ との差分 $p_d(t)$ [Pa] を計算する。

$$\bar{p}(t) = \frac{1}{100} \sum_{i=0}^{99} p(t-i) \quad (1)$$

$$p_d(t) = p(t) - \bar{p}(t) \quad (2)$$

3.4 タッチ検出

イヤホンを素早く押したとき、図 3 中央に示すように、外耳道の圧力は瞬間的に上昇しすぐに元の気圧に戻る。提案手法は閾値 p_{th} [Pa] を設定し、 $p_d(t) > p_{th}$ を連続して満たす区間を検出し、その区間の長さが t_{th} [s] 以下であればタップであると認識する。一方で、イヤホンを押した状態から素早く離れたとき、図 3 右に示すように、外耳道の圧力は瞬間的に下降しすぐに元の気圧に戻る。 $p_d(t) < p_{th}$ を連続して満たす区間を検出し、その区間の長さが t_{th} 以下であればリリースであると認識する。これにより図 3 左に示すように、指でイヤホンを素早く瞬間的に押すとタップの直後にリリースが出現する。また、図 3 中央に示すように、指でイヤホンを素早く押してゆっくりと指を離す (または離さずに押した状態を維持する) とタップのみが出現する。さらに、図 3 右に示すように指でイヤホンを素早く押して、時間をあけから指を素早く離すとタップとリリースが時間差で出現する。本研究では $p_{th} = 10$ [Pa]、 $t_{th} = 0.05$ [s] とした。

気圧の変化の閾値に加えて、閾値を超えた区間の長さにも閾値を設けている理由として、気圧の変化は気候、高度変化、振動があるが、これらの変化は緩やかであり、気圧の変化が閾値を超えたとしてもその区間は長いいため、除去で

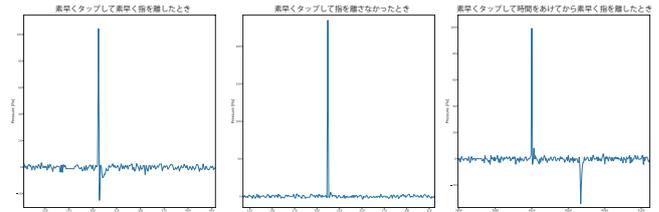


図 3 指でイヤホンを押したときの気圧センサの波形

きると考えたためである。

4. 評価

本節では提案手法の有効性を評価するために実施した実験について述べる。

4.1 評価環境

1名の被験者 (男性, 30代) の右耳に実装したカナル型イヤホンを装着して以下の状況で外耳道の気圧を記録した。

- 状況 1: タップ・リリース
- 状況 2: ロングタップ
- 状況 3: ロングタップ・リリース
- 状況 4: 発話する
- 状況 5: ナッツを食べる
- 状況 6: 水を飲む
- 状況 7: 首を左右に振る
- 状況 8: 立って座ってを繰り返す
- 状況 9: 座ったまま動かさず何もしない

状況 1 では素早く押して素早く離す動作を約 50 回行った。状況 2 では素早く押して、押した状態を維持したのち、ゆっくり離す動作を約 50 回行った。状況 3 では素早く押して、押した状態を維持したのち、素早く離す動作を 50 回行った。状況 4 では「あいうえお」と発話する動作を連続で 10 回 (約 30 秒) 行った。状況 5 ではクルミを 5 個 (約 60 秒) 行った。状況 6 ではペットボトルの水を一口ずつ (約 25 秒) 飲んだ。状況 7 では左右に首を振る動作を連続で 10 回 (約 30 秒) 行った。状況 8 では椅子に座った状態から立って再び座る動作を 10 回 (約 30 秒) 行った。状況 9 では座ったまま動かさず (約 80 秒) 何もしなかった。サンプリング周波数は約 115Hz である。データ採取日は 2022 年 5 月 22 日午前 1 時頃であり、その時の気圧 Weathernews 社^{*6}によると 999hPa であった。また、実験地点は海拔 17m の位置にある建物の 2 階である。

4.2 結果

状況 1, 2, 3, 4, 5, 6, 7, 8, 9 のデータに対する検出結果と気圧の時系列波形をそれぞれ図 4, 図 5, 図 6, 図 7, 図 8, 図 9, 図 10, 図 11, 図 12 に示す。図中の青色の波形はで気圧 $p_d(t)$ 、橙色の波形は提案システムの検出結

*6 Weathernews Inc., <https://weathernews.jp/>

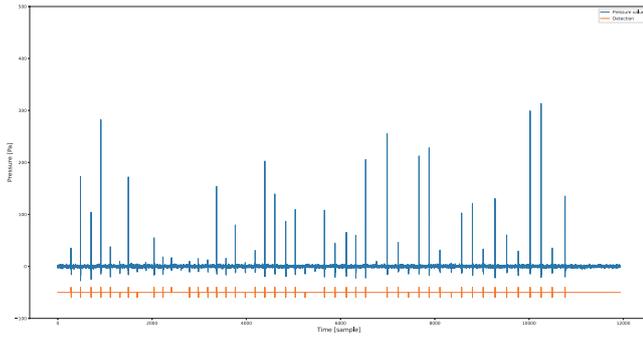


図 4 気圧データと検出結果 (状況 1: タップ・リリース)

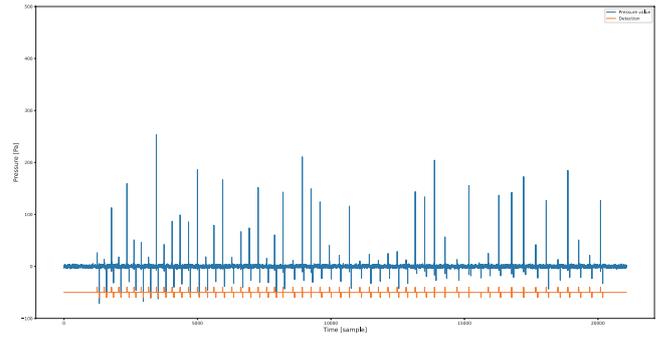


図 6 気圧データと検出結果 (状況 3: ロングタップ・リリース)

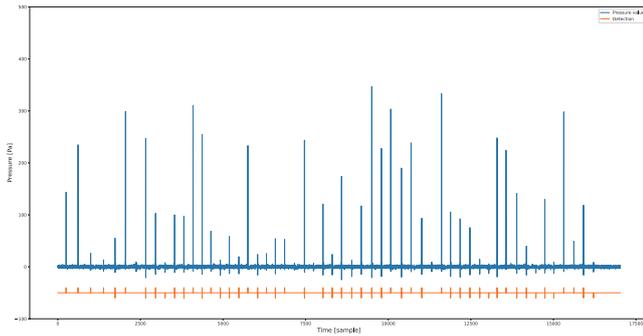


図 5 気圧データと検出結果 (状況 2: ロングタップ)

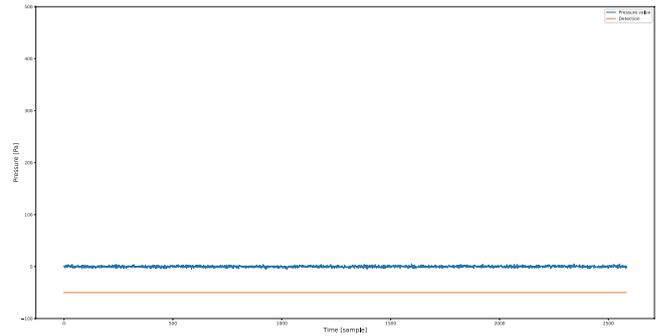


図 7 気圧データと検出結果 (状況 4: 発話する)

果であり、タップ検出時は-40 (y 軸正方向へのスパイク)、リリース検出時は-60 (y 軸負方向へのスパイク)、それ以外の場合は-50 を示す。

図 4 より、状況 1: タップ・リリースでは 50 回の試行のうちタップを 40 回、リリースを 45 回正しく検出している。また、タップとリリースを行っていないところでの誤検出は発生しなかった。

図 5 より、状況 2: ロングタップでは 55 回の試行のうちタップを 46 回正しく検出している。しかし、リリースが 35 回誤って検出されている。

図 6 より、状況 3: ロングタップ・リリースでは 55 回の試行のうちタップを 54 回、リリースを 55 回すべて正しく検出している。タップとリリースを行っていないところでの誤検出は発生しなかった。

イヤホンに触れていない状況 4 から状況 9 において気圧に瞬間的な大きな変化は現れず、タップとリリースを誤って検出することはなかった。図 11 より状況 8: 立って座ってを繰り返すでは被験者の顔の高さが変化するため、気圧の波形が波打っているように見えるが、タップやリリースを行ったときに発生する気圧変化よりは極めて小さく、誤検出はなかった。

5. まとめ

本研究では、手でイヤホンをタッチしたときに外耳道の気圧が変化することに着目して、手でイヤホンを押すことで入力を実現するインタフェースを提案した。カナル型イヤホン装着時には外耳道は鼓膜とイヤピースで囲われ

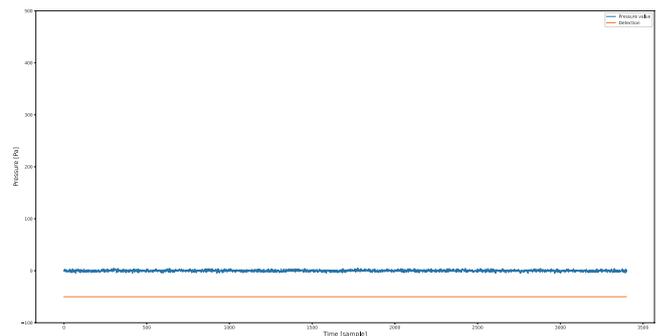


図 8 気圧データと検出結果 (状況 5: ナッツを食べる)

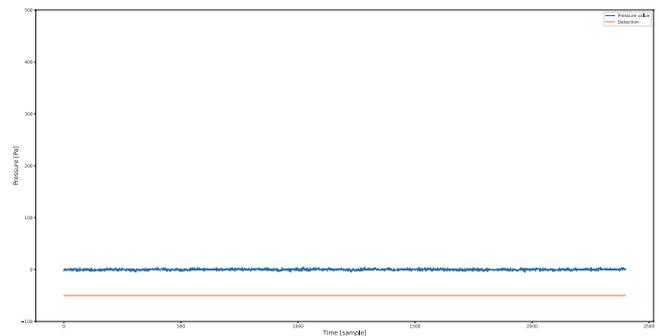


図 9 気圧データと検出結果 (状況 6: 水を飲む)

た空間となり、イヤピースが素早く押されると外耳道の容積が小さくなり、気圧が瞬間的に上昇する。同様に、押した状態から手が素早く離れるとイヤピースは元の位置に戻ろうとするため外耳道の容積が大きくなり、気圧が瞬間的に低下する。評価実験の結果、手でイヤホンを押すタップは 160 回中 140 回 (87.5%) 成功し、タップしてい

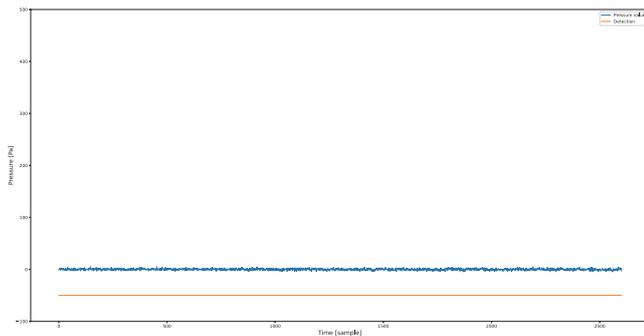


図 10 気圧データと検出結果 (状況 7: 首を左右に振る)

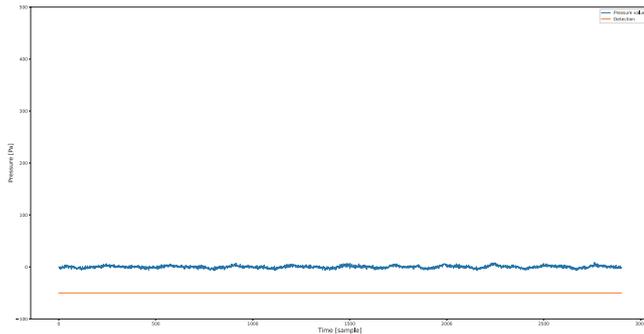


図 11 気圧データと検出結果 (状況 8: 立って座ってを繰り返す)

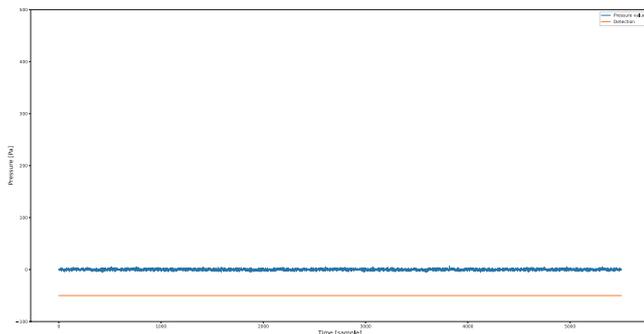


図 12 気圧データと検出結果 (状況 9: 座ったまま動かず何もしない)

ないのにタップが検出されることはなかった。また、手でイヤホンを押した状態から離すリリースは 105 回中 90 回 (85.7%) 成功した。一方で、55 回ゆっくいと手を離したにもかかわらず、そのうち 35 回で誤ってリリースが検出された。そのため、リリースの検出方式については改善が必要である。このほか、発話、食事、飲水、首振り、上下運動、安静時には 1 回もタップとリリースが検出されることはなかった。

手でイヤホンを押したときに発生する気圧変化は飛行機やエレベータなどの高度変化によって発生する気圧変化よりも極めて大きく容易に区別可能であると考えており、今後検証する。

謝辞

本研究の一部は科学技術振興機構さきがけ (JP-

MJPR1937) (JPMJPR2138) の支援を受けた。ここに記して謝意を表す。

参考文献

- [1] 古谷 聡, 越仲孝文, 大杉孝司: ヒアラブル技術によるヒューマン系 IoT ソリューションの取り組みと展望 (デジタルビジネスを支える IoT 特集) (お客様に価値を提供する IoT ソリューション), NEC 技報 = NEC technical journal, Vol. 70, No. 1, pp. 47–51 (2017).
- [2] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi, CanalSense: Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals, In Proc. of 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17), pp. 679–689 (2017).
- [3] D.J.C. Matthies, B.A. Strecker and B. Urban, EarField-Sensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input Through Facial Expressions, In Proc. of 2017 CHI Conference on Human Factors in Computing Systems (CHI '17), pp. 1911–1922 (2017).
- [4] K. Taniguchi, H. Kondo, M. Kurosawa, and A. Nishikawa, Earable TEMPO: A Novel, Hands-Free Input Device that Uses the Movement of the Tongue Measured with a Wearable Ear Sensor, Sensors, Vol. 18, No. 3 (2018).
- [5] A. Bedri, D. Byrd, P. Presti, H. Sahni, Z. Gue, and T. Starner, Stick It in Your Ear: Building an In-ear Jaw Movement Sensor, In adjunct Proc. of 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proc. 2015 ACM International Symposium on Wearable Computers, pp. 1333–1338 (2015).
- [6] 真鍋宏幸, 福本雅朗, Headphone Taps: 通常のヘッドホンへのタップ入力, 情報処理学会論文誌, Vol. 55, No. 4, pp. 1334–1343 (2014).
- [7] G. Laput, X.A. Chen, and C. Harrison, SweepSense: Ad Hoc Configuration Sensing Using Reflected Swept-Frequency Ultrasonics, In Proc. of 21st International Conference on Intelligent User Interfaces (IUI '16), pp. 332–335 (2016).
- [8] T. Amesaka, H. Watanabe, and M. Sugimoto, Facial Expression Recognition Using Ear Canal Transfer Function, In Proc. of the 23rd International Symposium on Wearable Computers (ISWC 2019), pp. 1–9 (2019).