

MimiSense: 外耳道内の気圧変化を利用した下顎運動認識システム

安藤 宗孝* 久保 勇貴† 志築 文太郎‡ 高橋 伸‡

概要. 種々の下顎運動を認識するシステム「MimiSense」を示す。MimiSenseは、口を開ける、口を閉じる、及び顎を左に動かすなどの下顎運動を認識することが可能であり、認識結果を情報端末に対する操作に利用する。下顎運動による操作はハンズフリーかつアイズフリーであり、タッチ操作などの手を使った操作が行えないような状況においても使用することが可能である。MimiSenseを使用するにはユーザはイヤホン型センサ部を装着しさえすればよい。センサ部は市販の外耳道挿入型イヤホンと同様の形状をしており容易に装着できる。本稿にて、MimiSenseの動作原理、実装方法、下顎運動認識精度を確かめるために行った著者実験、及びアプリケーション例を述べる。また、本システムの有用性を示すため、本システムのさらなる発展性を議論する。

1 はじめに

日常生活において手が使えず、スマートフォンなどの情報端末（以降、端末）を操作できない状況がしばしば発生する。例えば、満員電車において手を動かさずイヤホンを使って聴いている音楽の操作ができない場合、手袋をつけているため操作ができない場合、及び両手に荷物を持っており端末への通知に対応できない場合などである。

このような状況において端末を操作するために手を用いない、すなわちハンズフリーな操作手法が提案されている。例として、足 [18]、視線 [5]、口及び顎 [1] を用いた操作手法が挙げられる。我々はこれらの中でも顎、特に下顎運動をハンズフリー操作手法に用いる。下顎は上下、前後、左右、及び回転など、多様に動かせ、かつ見ずとも動かせるという特徴がある。

本稿では下顎運動による外耳道内の気圧変化を利用した下顎運動認識システム MimiSense を示す。図 1a に MimiSense 全体を示す。MimiSense は気圧センサを埋め込んだ外耳道挿入型イヤホン（以下、イヤホン型気圧センサ）（図 1b）を用いて、外耳道をイヤホンにて密閉した状態における外耳道内の気圧（以降、気圧値）を取得する。これにより、本システムは、口を開ける、口を閉じる、及び顎を左に動かす等の下顎運動を認識できる。また、これらを連続的に組み合わせた下顎運動も認識できる。

MimiSense を使用するにはユーザはイヤホン型気圧センサを装着しさえすればよい（図 1c）。本システムのイヤホン型気圧センサは市販されている外耳道挿入型イヤホンを利用して作成可能であり、か

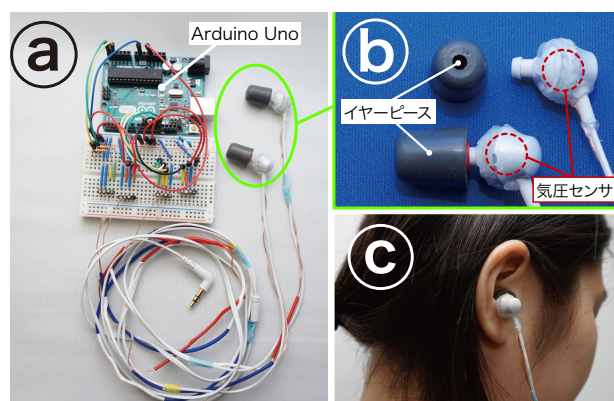


図 1. MimiSense: a) MimiSense 全体, b) イヤホン型気圧センサ (破線内部イヤホン内に気圧センサ), c) イヤホン型気圧センサを装着した様子.

つその形状を変更する事がないため、容易に装着する事ができ、装着感も良い。本稿においては本システムの動作確認及び音を鳴らした状態で使用できるかどうかを検証するために行った著者実験も示す。

2 関連研究

MimiSense と同様に、外耳装着インタフェース、顎周辺動作の識別、及び気圧を用いた状況や動作の認識の研究は様々になされている。

2.1 外耳装着インタフェース

外耳装着インタフェースの研究がいくつか提案されている。それらの研究では、心拍の取得 [12, 19]、イヤホンを装着している耳の認識 [10]、ヘッドフォン [20] 及び耳 [8] へのタッチ入力への認識、顎動作に基づく食事動作認識 [3]、眼電図による目の動きの認識 [9]、黙声認識 [13]、頭がどれだけ回転させているかの認識 [6]、表情の認識 [21] を行っている。

Copyright is held by the author(s).

* 筑波大学情報学群情報メディア創成学類

† 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻

‡ 筑波大学システム情報系

MimiSense は機械学習によって下顎運動ごとの外耳道内の気圧値の変化を分類することにより、下顎の状態ではなく下顎運動を認識する。さらに、認識した下顎運動を端末に対する操作に利用する。

2.2 顎周辺動作の認識

顎周辺動作の認識を試みた研究がある。例えば、Bitey [1] は噛んだ時の音を用いて噛む動作を認識し、ハンズフリー入力手法に利用した。On the Tip of My Tongue [4] は頬の外側に装着した布製圧力センサアレイを用いて舌の動きを認識し、その動きを入力として利用した。加藤 [17] は、外耳道に挿入したチューブが顎運動に伴って変形することを利用して、咀嚼運動の検出を試みた。一方、MimiSense はイヤホン型気圧センサを外耳に取り付け気圧値を取得し、これを用いて下顎運動を認識する。

2.3 気圧を用いた状況や動作の認識

気圧を状況や動作の認識に用いた研究がいくつかある。それらの研究では、例えば建物屋内のドアの開閉 [14]、屋内における場所検出 [7, 11]、建物の階数識別 [15]、及び風船へのタッチ検出 [16] を気圧を用いる事により行う。本研究は気圧を用いて下顎運動の認識を行う。

3 MimiSense

MimiSense は、装着したユーザーの下顎運動の認識を行い、端末に対する操作に利用するシステムである。MimiSense による操作はハンズフリーかつアイズフリーである。本節では、外耳道内の気圧変化を用いて下顎運動を認識する原理及び、認識対象とする下顎運動を述べる。

3.1 認識原理

本システムでは下顎運動に伴う外耳道の形状変化を利用する。

下顎が動くとき、下顎骨の位置の変化によって外耳道が圧迫され外耳道の形状が変化する。下顎運動による骨の位置変化を図 2 に示す。下顎骨は下顎の骨全体を指す。下顎骨の後方に位置する突起を関節突起と呼び、下顎頭は関節突起の上端を指す。図 2a は閉口時、図 2b は開口時の下顎骨の位置を示している。図 2a においては下顎頭が外耳道に近い位置にあるが、図 2b では図 2a と比べ下顎頭と外耳道の位置関係が変化している。この位置変化は開口に伴う下顎運動によるものである。下顎頭と外耳道の位置関係の変化に伴い外耳道の形状が変化する。外耳道の形状変化により外耳道内の体積も変化する。この時、外耳道が密閉されている場合、外耳道内の気圧も変化する。この変化は下顎運動ごとに特徴的である。

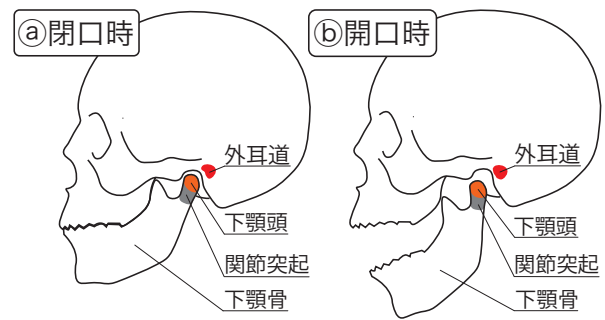


図 2. 顎の構造.

MimiSense はイヤホン型気圧センサを用いて外耳道を密閉しつつ気圧値の変化を読み取ることにより下顎運動を認識する。

3.2 認識対象とする下顎運動

MimiSense は、口を開ける (図 3a)、口を閉じる (図 3b)、顎を左に動かす (図 3c)、及び顎を右に動かす (図 3d) という 4 つの下顎運動とこれらを連続的に組み合わせた下顎運動を認識対象とする。例えば、a 及び b を組み合わせた下顎運動、つまり、口を閉じた状態から口を開けその後、口を閉じる下顎運動 (図 3e) を新たな下顎運動として認識する。

下顎運動ごとの左右の気圧値の変化を図 4 及び図 5 に示す。図 4 及び図 5 より、下顎運動ごとに気圧値の変化が異なることがわかる。具体的には、縦方向の下顎運動においては左右の気圧値が同時に上昇または下降している。これに対して横方向の下顎運動においては左右の気圧値の変化が異なる形に変化している (例: 顎を右に動かすと、左の気圧値は上昇し始め、右の気圧値は下降し始める)。口を開けて閉じる下顎運動の気圧値の変化は口を開ける下顎運動及び口を閉じる下顎運動の組み合わせにて成り立っている。

4 MimiSense の実装

MimiSense はイヤホン型気圧センサ及び下顎運動認識プログラムから成る。

4.1 イヤホン型気圧センサ

下顎運動に伴う外耳道内の気圧変化を取得するために、イヤホン型気圧センサを作成した。図 1b に作成したイヤホン型気圧センサを示す。これは市販の外耳道挿入型イヤホン (Panasonic 社製 RP-HJE260) の内部に気圧センサ (Bosch 社製 BMP280) を埋め込んだものである。

外耳道内の気圧の変化を正確に把握するためには、外耳道を密閉する必要がある。そのためにイヤホン型気圧センサによって外耳道を密閉する。イヤ

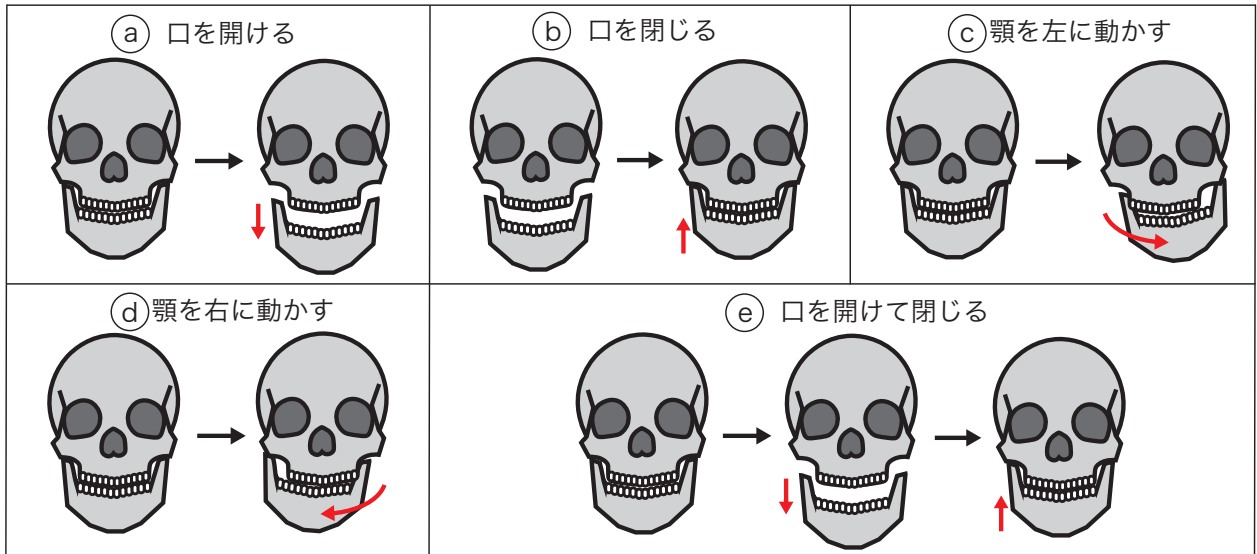


図 3. 認識対象とする下顎運動.

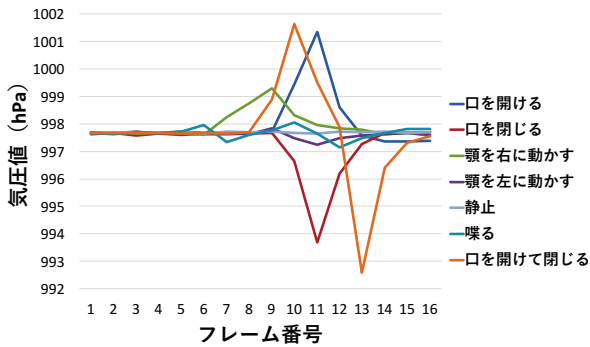


図 4. 下顎運動ごとの気圧値 (左).

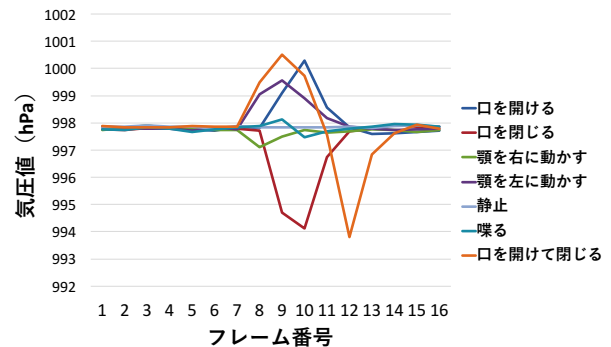


図 5. 下顎運動ごとの気圧値 (右).

ホン型気圧センサの密閉性を高めるため、イヤホンの空気孔をホットグルーを用いて塞いだ。また、外耳道の形状には個人差があるため、外耳道の密閉性を高めるためには個々人に適合したイヤピースを利用することが望ましい。そのため、今回作成したイヤホン型気圧センサにおいては、押しつぶした状態にて外耳道に挿入し、その後外耳道の形状に沿って膨らむフォームタイプのイヤピースを用いる。フォームタイプのイヤピースは、形状を変化させることが可能であるため、密閉性を高めつつ多くの人が装着可能である。

4.2 下顎運動認識プログラム

下顎運動認識プログラムは学習フェーズ及び分類フェーズの2つの処理にて構成される。これらの実装にはプログラミング言語 Python 3.5.1 及び機械学習ライブラリ scikit-learn 0.17.1 を使用した。

4.2.1 学習フェーズ

学習フェーズでは、機械学習に使用する学習データを生成するため下顎運動ごとに以下の処理を行う。

まず、気圧値を片耳 16 フレーム、左右合わせて 32 フレーム (以降、気圧変化パターン) 随時取得する。

次に以下の関数 $F(i, j)$ を用いて、 $F(4, 8)$ 及び $F(8, 12)$ を両耳に対して計算する。

$$F(i, j) = \frac{\sum_{k=i+1}^j (g_k - g_{k-1})}{j - i} \quad (1)$$

g_k は直近 16 フレームの気圧データであり、最も古いデータが g_0 、最も新しいデータが g_{15} である。いずれかの耳の $F(4, 8)$ 及び $F(8, 12)$ が式 (2) 及び (3) を満たす場合、特徴的な気圧変化パターンが生じたとして保存する。

$$F(4, 8) > a \wedge F(8, 12) < -a \quad (2)$$

$$F(4, 8) < -a \wedge F(8, 12) > a \quad (3)$$

なお、本実装においては閾値 a を 0.1 とした。

また、前述の関数では対応しきれない気圧変化も存在するため（例：4 フレーム以内に上下に変化するようなデータ）式 (4) を満たす気圧変化パターンも、保存する。

$$\max(g_k) - \min(g_k) > 2a \quad (4)$$

$\max(g_k)$ は g_k の最大値であり、 $\min(g_k)$ は g_k の最小値である。

最後に、保存した気圧変化パターンから機械学習に利用する特徴量を生成する。特徴量として、各フレームに対する直前のフレームとの差分（左右各 15 個、合計 30 個）、左右の気圧変化パターンの差分値に対してフーリエ変換を行った振幅スペクトル（8 個）、標準偏差（左右各 1 個、合計 2 個）、最大最小値幅（左右各 1 個、合計 2 個）、第 0 フレームから第 15 フレームを引いた値（左右各 1 個、合計 2 個）、最大値から第 0 フレームを引いた値（左右各 1 個、合計 2 個）、及び、最小値から第 0 フレームを引いた値（左右各 1 個、合計 2 個）の計 48 個を用いている。

4.2.2 分類フェーズ

学習フェーズにて生成した特徴量を利用して下顎運動の分類を行う。機械学習には Random Forest を利用する。前述した特徴量を用いて作成した分類器を利用して直近 16 フレームのデータを随時分類する。最新 5 回分の分類結果のうち最も判定回数が多い下顎運動を最終的な分類結果として出力する。

5 下顎運動認識実験

本システムの動作確認及び音を鳴らした状態で使用できるかどうかを検証するために、著者 1 名を被験者とした下顎運動認識実験を行った。実験は同じ場所及び時間帯にて 3 日間連続して行われた。

5.1 実験機器

実験には作成したイヤホン型気圧センサを使用した。イヤピースにはフォームタイプのもの（Westone 社製 True-Fit Foam Eartips）を用いた。

5.2 実験内容

本実験は被験者は椅子に深く腰掛けた状態にて行われた。被験者はイヤホン型気圧センサを装着した状態にて、口を開ける、口を閉じる、顎を左に動かす、顎を右に動かす、口を開けて閉じる、及び喋るという合計 6 種類の下顎運動を行った。また、下顎運動を行っていない状態（静止）も取った（被験者

に何も行わないよう指示し、実験者がその状態の気圧変化パターンを記録した）。喋る下顎運動に関しては被験者の前に設置したディスプレイに文章を表示した上で、その文章を音読させた。文章は「我輩は猫である」（夏目漱石）の冒頭 3 文である。各下顎運動はランダムな順番にて 1 つずつ提示される。このとき、提示された下顎運動を実行することを 1 試行とする。また、6 種類の下顎運動に静止を加えた 7 種類を 1 試行づつ行うことを 1 セッションとする。なお、静止については 4.2.1 節のパターンに該当しないため手動にて記録した。イヤホンから音を鳴らしている条件（以降、音あり条件）及び音を鳴らしていない条件（以降、音なし条件）をそれぞれ 12 セッションずつ行い、計 168 試行を行った。音を鳴らしている条件の際、使用した音源は Ludwig van Beethoven 作曲交響曲第 5 番-第 1 楽章ハ単調作品 67 である。また、音量は [2] において音量調整の目安としてあげられている、「ヘッドホン使用状態にて普通に話した自分の声が聞こえる程度」の範囲において最大の音量とした。

5.3 結果

実験によって得られた気圧変化パターンに対して交差検定を行った。分類には機械学習ライブラリ Weka が提供する Random Forest を用いた。2, 3, 4, 6, 12-fold 交差検定を行った際の認識率を図 6 に示す。図 6 より学習回数が 2（すなわち、3-fold）よりも上の場合、認識率が低下しているものがあることがわかる。次に最も高い認識率を示した、3-fold 交差検定を行った際の結果を述べる。1 日目、音なし条件の認識率は 96.4% であり、音あり条件の認識率は 89.3% であった。2 日目、音なし条件の認識率は 100% であり、音あり条件の認識率は 100% であった。3 日目、音なし条件の認識率は 98.8% であり、音あり条件の認識率は 100% であった。

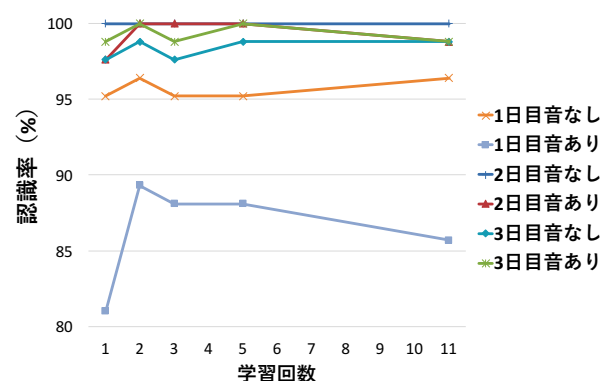


図 6. 著者実験における認識精度。

5.4 考察

1日目の認識率が2日目、3日目に比べて大きく低下している。この原因として実験1日目に実験実施場所付近を通過した台風(2016年10号、ライオンロック)による気圧値への影響、特に台風の移動及び風に伴う気圧変化が気圧値に与える影響が大きかったと考えられる。

また、学習回数が3回以上の場合に認識率が低下している。これは過学習によるものと考えられる。よって本システムにおいては学習回数が2回で十分である可能性が示唆される。

音なし条件と音あり条件の認識率を比較するために、学習回数を2回とした時の3日分の認識率を対応のあるt検定を用いて比較した。結果、音なし条件と音あり条件に有意な差はなかった($t(2) = .759, p = .527 > .005$)。よって、音による認識率への影響はないと考えられる。

6 アプリケーション例: 音楽プレイヤー

MimiSenseを用いた下顎運動認識によって操作を行う音楽プレイヤーを作成した。再生/停止には顎を開けて閉じる動作、曲送りには顎を右に動かす動作、曲戻しには顎を左に動かす動作を割り当てた(図7)。顎を開ける動作及び顎を閉じる運動は、その後顎を開ける運動または顎を閉じる運動が付随的に発生するため顎を開けて閉じる運動を再生及び停止に割り当てた。ユーザはこの音楽プレイヤーを、両手に荷物を持っているなど手が塞がっている状況においても操作することが可能である。



図 7. 音楽プレイヤーの操作割り当て。

7 議論と今後の課題

7.1 気圧に関する条件

MimiSenseは気圧センサを利用しているため、気圧が変化する状況において下顎運動が認識できない可能性がある。気圧が変化する状況とは、建物の内外を行き来する状況、地表面からの高さが変化する状況、及び天候の変化が激しい状況などである。今回、センサから取得した気圧値をそのまま利用したが、ローパスフィルタを用いることによってノイズを低減することにより、より正確に下顎運動が認識できると考えられるため、今後実装し、検証を行う。

また、既存の気圧センサを利用した状況認識技術と組み合わせることによって、周囲の状況に合わせ

た操作方法の提案が可能であると考えられる。

7.2 外耳道の密閉性に関する条件

外耳道の形状が人によって異なるため、多くの人が同じイヤピースにて本システムを利用することは難しい。また、密閉性の低いイヤピースを使用した場合、外耳道の形状変化による外耳道内の気圧値の変化量が低下するため、下顎運動の認識率が低下する。しかし、ユーザが密閉性の高いイヤピースを使用する場合、顎を動かした際の気圧変化が顕著になるため、下顎運動をより高い精度にて認識できると考える。さらに、耳型を取り作成するオーダーメイドイヤホンを用いればより密閉性が確保出来る。このような場合、田山ら[19]の研究に基づいて心拍数を取得することができる可能性がある。

7.3 より複雑な下顎運動の認識

下顎運動の運動量認識

下顎運動による外耳道内の気圧値の変化は、下顎運動をどれだけ大きさにて行うかによって気圧値の変化量は変わる。したがって、この変化量から、下顎運動の運動量を推定することが可能であると考えられる。今後、外耳道内の気圧値の変化量と下顎運動の運動量の関係性を明らかにすることによって下顎運動の運動量の推定を試みる。

黙声母音認識

顎の動きを元に黙声母音認識を行っている研究[13]がある。本研究においても母音を発する際の下顎運動を学習することにより母音認識ができる可能性がある。黙声認識を行うことができれば、様々な事情で声を出すことができない人が実際に声を出すことなく擬似的な音声出力を行うことができると考えられる。

7.4 下顎運動以外の認識

イヤホン型気圧センサを装着した状態にて首を曲げると外耳道内の気圧値が変化する。これを用いて下顎運動以外の動作を認識することができると考えられる。下顎運動認識プログラムを用いて、首を左右に傾けた際の気圧変化パターンを学習すると、首の動きと下顎運動を分類できることを確認した。今後、この分類精度を検証するとともに、その他の体の部位による動作を検出することを検討する。

7.5 MimiSenseの利用可能性

MimiSenseはヘッドマウントディスプレイやスマートグラスへの入力装置として利用できると考えられる。これらのデバイスは装着位置がずれることによる使用者への影響が大きいため、本体に触れる操作が難しい問題がある。また、非透過型デバイスの場合、映し出されている空間以外を見ることがで

きないため、アイズフリー操作手法であることが望ましい。MimiSense はハンズフリーかつアイズフリー操作が行えるため、これらの問題に対して有効である。また、MimiSense はハンズフリー及びアイズフリーであることから身体障害者向け入力装置としても利用することも可能だと考えられる。

8 おわりに

本稿では、外耳道内の気圧変化を利用した下顎運動認識システムを提案し、その実装方法を示した。実装においては、市販のイヤホンに気圧センサを埋め込んだイヤホン型気圧センサを作成した。これを用いることによって、下顎運動に対応する気圧変化パターンを計測し、気圧変化パターンから導き出した特徴量を利用した機械学習を行うことにより下顎運動の認識を行った。また、下顎運動の認識精度を示すために著者実験を3日間行った。実験では下顎運動ごとの気圧値の変化を記録し、交差検定を行った。3-fold 交差検定の結果、3日間の平均認識率は音なし条件にて98.4%、音あり条件にて96.4%となった。また、MimiSense を利用したアプリケーション例を示した。今後、本システムが著者以外にも使用可能であることを検証するため、被験者実験を行う。

参考文献

- [1] D. Ashbrook, C. Tejada, D. Mehta, A. Jiminez, G. Muralitharam, S. Gajendra, and R. Tallents. Bitey: An Exploration of Tooth Click Gestures for Hands-free User Interface Control. In *Proc. of MobileHCI '16*, pp. 158–169.
- [2] AVandIT 機器標準化専門委員会. 電子情報技術産業協会規格 JEITA RC-8140A ヘッドホンおよびイヤホン. 1998年3月制定, 2010年3月改正.
- [3] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner. Detecting Mastication: A Wearable Approach. In *Proc. of ICMI '15*, pp. 247–250.
- [4] J. Cheng, A. Okoso, K. Kunze, N. Henze, A. Schmidt, P. Lukowicz, and K. Kise. On the Tip of My Tongue: A Non-invasive Pressure-based Tongue Interface. In *Proc. of AH '14*, pp. 12:1–12:4.
- [5] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proc. of UIST '15*, pp. 457–466.
- [6] D. Hong, B. Zhang, Q. Li, S. Nirjon, R. Dickerson, G. Shen, X. Jiang, and J. Stankovic. SEPTIMU: Continuous In-situ Human Wellness Monitoring and Feedback Using Sensors Embedded in Earphones. In *Proc. of IPSN '12*, pp. 159–160.
- [7] S. Hyuga, M. Ito, M. Iwai, and K. Sezaki. Estimate a User's Location Using Smartphone's Barometer on a Subway. In *Proc. of MELT '15*, pp. 2:1–2:4.
- [8] R. Lissermann, J. Huber, A. Hadjakos, S. Nanayakkara, and M. Mühlhäuser. EarPut: Augmenting Ear-worn Devices for Ear-based Interaction. In *Proc. of OzCHI '14*, pp. 300–307.
- [9] H. Manabe, M. Fukumoto, and T. Yagi. Conductive Rubber Electrodes for Earphone-based Eye Gesture Input Interface. In *Proc. of ISWC '13*, pp. 33–40.
- [10] K. Matsumura, D. Sakamoto, M. Inami, and T. Igarashi. Universal Earphones: Earphones with Automatic Side and Shared Use Detection. In *Proc. of IUI '12*, pp. 305–306.
- [11] K. Muralidharan, A. J. Khan, A. Misra, R. K. Balan, and S. Agarwal. Barometric Phone Sensors: More Hype Than Hope! In *Proc. of Hot-Mobile '14*, pp. 12:1–12:6.
- [12] M.-Z. Poh, K. Kim, A. D. Goessling, N. C. Swenson, and R. W. Picard. Heartphones: Sensor Earphones and Mobile Application for Non-obtrusive Health Monitoring. In *Proc. of ISWC '09*, pp. 153–154.
- [13] H. Sahni, A. Bedri, G. Reyes, P. Thukral, Z. Guo, T. Starner, and M. Ghovanloo. The Tongue and Ear Interface: A Wearable System for Silent Speech Recognition. In *Proc. of ISWC '14*, pp. 47–54.
- [14] M. Wu, P. H. Pathak, and P. Mohapatra. Monitoring Building Door Events Using Barometer Sensor in Smartphones. In *Proc. of UbiComp '15*, pp. 319–323.
- [15] H. Ye, T. Gu, X. Tao, and J. Lu. B-Loc: Scalable Floor Localization Using Barometer on Smartphone. In *Proc. of MASS'14*, pp. 127–135, 2014.
- [16] 中島康祐, 伊藤雄一, 林勇介, 池田和章, 藤田和之, 尾上孝雄. Emoballoon: ソーシャルタッチインタラクションのための柔らかな風船型インタフェース. インタラクション 2013, pp. 95–102.
- [17] 加藤隆史. 咀嚼動作検出装置. 特開 2008-48791, 2008-03-06.
- [18] 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, 33(2):2.116–2.124, 2016.
- [19] 田山愛, 山下紘正, 佐藤智夫, 北角権太郎, 千葉敏雄, 土岐彰. イヤホン式小型生体情報センサの開発と精度に関する検討. 昭和学生会雑誌, 74(1):60–66.
- [20] 真鍋宏幸, 福本雅朗. Headphone Taps: 通常のヘッドホンへのタップ入力. 情報処理学会論文誌, 55(4):1334–1343.
- [21] 谷口和弘, 西川敦, 小林英津子, 宮崎文夫, 佐久間一郎. みみスイッチ: 外耳の動きを入力情報とする常時着用型入力装置. インタラクション 2010, pp. 243–246.