

## 腕時計型ウェアラブルデバイスを用いた会話時間計測手法の構築に向けて

小松勇輝<sup>†</sup> 下条和暉<sup>‡</sup> 西山勇毅<sup>‡</sup> 瀬崎薫<sup>§</sup>立命館慶祥高等学校<sup>†</sup> 東京大学 生産技術研究所<sup>‡</sup>・空間情報科学研究センター<sup>§</sup>

## 1. はじめに

社会的接触の頻度は、自己評価の身体的健康度と正の相関があると報告されている[1]。社会的接触の頻度とは、会う・話す・挨拶する・共に過ごす・友人や家族、同僚と交流を行うことであり、これらの情報の客観的な計測は、人々の生活環境調査において重要な情報である。例えば、新型コロナウイルス感染症（COVID-19）の世界的な蔓延により、人々の生活様式は COVID-19 の感染リスクを下げるため、人の接触機会をできるだけ少なくする生活スタイルに大きく変容した[2]。COVID-19 は、人が密集・密接した場所での滞在や、対面での会話により感染する可能性が高くなるため、外出の自粛やいわゆる三密（密集・密閉・密室）回避など行動制限が実施された。このような生活スタイルは、感染リスクを下げる可能性が高い一方で、50.4%の人が「人と会えないこと」を主な理由に閉塞感を抱いている[3]など、感染症対策による二次的な健康被害も危険視されている。社会的接触の頻度を継続的に低負荷に記録することで、迅速な状況の把握と対策が実現できると考えられる。そこで本研究では、社会的接触のうち「会話」に焦点を当て、日常生活での会話時間を自動計測することを目的とする。本稿では、市販の腕時計型ウェアラブルデバイスのマイクより収集した環境音と機械学習を組み合わせる手法を提案する。収集した環境音と会話音から会

話の判定モデルを実装し、その精度と可能性を実験室環境と実際のユーザ環境で評価する。

## 2. 関連研究

先行研究では、市販のスマートフォンなどのモバイルデバイスに搭載されたモーションセンサや位置情報センサ、マイクなど様々なセンサを用いて人の行動傾向を把握する研究[4]が数多く行われている。本研究と同様にマイクを用いた会話時間計測[5]の計測手法も提案されているが、24時間肌身離さずスマートフォンを持ち計測し続けるには負担や精度上の問題が考えられる。またスマートフォンは、ポケットや鞆、机の上など、時と場合によって計測環境が大きく異なるため、計測精度に不安が残る。

## 3. 会話時間計測

本研究では、日常生活での会話時間の定量的な計測を目的とする。それらの実現のためには、マイクより収集した環境音から雑音と会話を区別し、さらにウェアラブルデバイス上で継続的に動作できる必要がある。

## 3.1 会話判別モデル

会話を判別する機械学習モデルを作成した。機械学習モデルは macOS の CreateML を使用して作成した。学習データは Common Voice の多言語データセット [6]より会話を識別するための 470 種類の話し声、ESC: Dataset for Environmental Sound Classification [7]より雑音を識別するための 2010

「A Method for Measuring Conversation Duration Using a Wristwatch Type Wearable Device」  
Yuki KOMATSU<sup>†</sup> Kazuki SHIMOJO<sup>‡</sup> Yuuki NISHIYAMA<sup>‡</sup>  
Kaoru SEZAKI<sup>§</sup>  
Ritsumeikan Keisho Senior High School<sup>†</sup>  
Institute of Industrial Science<sup>‡</sup>・Center for Spatial  
Information Science in The University of Tokyo<sup>§</sup>

種類の環境音を用いた。ただし、会話のみを正確に計測するため、笑い声・悲鳴などの発声や小さな吹きは雑音として学習した。

### 3.2 OHANASHI : 会話時間計測アプリ

前述の機械学習モデルを搭載したウェアラブルデバイス用アプリケーションのOHANASHIを開発した。OHANASHIはApple社のApple Watch上で動作し、Swift言語を用いてXcode上で開発した。本アプリを起動すると、AVAudioEngineを用いて常時マイクから周囲音を収集し、前述の機械学習モデルを用いて、毎秒「会話」か「雑音」を判断する。機械学習モデルが会話と判断し、その判定の信頼度合いが50%以上であれば、会話時間とする。会話と雑音の判定は、全て端末内でエッジ処理され、判定結果は端末上のデータベース(SQLite)に保存される。

## 4 評価実験

### 4.1 会話・雑音の認識精度

機械学習モデルの精度確認のために予備実験を行った。静かな空間で会話を10秒、その後に雑音を10秒間記録し、機械学習モデルの判別と正解データとを比較し精度を検証した。結果はAccuracyが0.91、F値が0.92であった。また、静かな環境でのテレビ音(ニュース番組)、静かな環境でのラジオ音、雑音の中での会話音、地下鉄構内とスーパーマーケットでのアナウンス音で精度テストを行ったが、全てAccuracyが0.85以上の精度であった。

### 4.2 日常生活での会話時間計測

実際の人を対象に一日の会話時間計測実験を実施した。被験者7名に対してApple Watchを配布し被験者は24時間Apple Watchを装着し、OHANASHIを起動した状態で二週間生活した。

図1にOHANASHIから得られた時刻毎の平均会話時間(秒)を被験者毎に示す。全体的な会話時間は被験者によって異なるがどの被験者も8時から会話時間が増加し、深夜は減少していることが確認できた。また、実験終了後に行ったアンケート調査から、会話時間を数値として可視化することで、会話を意識的に促すような行動変容を促進できる可能性が示唆された。

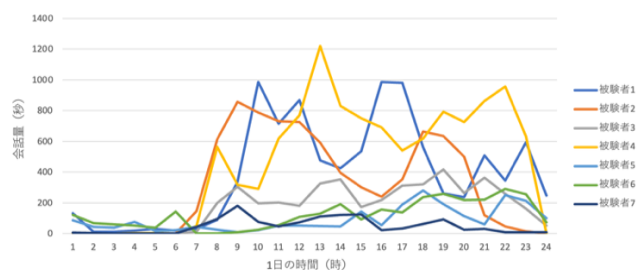


図1 被験者毎の平均会話時間

## 5 おわりに

本稿では、腕時計型のウェアラブルデバイスを用いた会話時間の継続的な計測手法の提案・実装・評価を行った。実装した機械学習モデルを用いることで、コントロール環境では85%以上の精度で会話と雑音を判断でき、実環境においてもある程度の会話時間の計測を行える可能性が示された。しかし、実環境における会話時間の検知精度については、実際の日常生活中の音に対して会話と雑音のラベル付けを行った上でその精度を評価する必要があるため、今後の課題である。また会話時間の計測とその可視化が、自身の社会的接触の頻度の気づきと行動変容に与える効果の検証も今後の課題である。

## 謝辞

本研究は、JST グローバルサイエンスキャンパスの一環として行われたものである。研究進行にあたって、関根康介先生はじめ立命館慶祥高等学校教職員の皆様に多くのサポートをいただいた。

[1] Stavrova, O. et al. (2021). Social Psychological and Personality Science, 12(6), 1058-1070.

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/YDEPUT> (last accessed on 2021-12-5)

[2] アメリカ疾病予防管理センター(Centers for Disease Control and Prevention: CDC). <https://www.cdc.gov/coronavirus/2019-ncov/index.html>. (last accessed on 2021-12-5)

[3] 日本財団「18歳意識調査」: 「コロナ禍とストレス」  
<https://www.nippon-foundation.or.jp/who/news/pr/2021/20210325-55393.html> (last accessed on 2021-12-5)

[4] Wang, R. et al. (2014). StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14). 3-14.

[5] Lu, H. et al. (2009). SoundSense: scalable sound sensing for people-centric applications on mobile phones. In Proceedings of the 7th international conference on Mobile systems, applications, and services (MobiSys '09). 165-178.

[6] Common Voice 多言語データセット  
<https://commonvoice.mozilla.org/ja> (last accessed on 2021-12-5)

[7] Karol J. Piczak, ESC: Dataset for environmental Sound Classification, Harvard Dataverse, 2015,  
<https://doi.org/10.7910/DVN/YDEPUT>