

# 行動認識センサによる対話型コミュニケーション検出の実験と評価 Experiment and Evaluation of Interactive Communication Detection Using Behavioral Recognition Sensor

伊藤 優樹<sup>†</sup>

Yuki Ito

千葉 慎二<sup>‡</sup>

Shinji Chiba

上野 正義<sup>†</sup>

Masayoshi Ueno

阿部 亨<sup>†,§</sup>

Toru Abe

高橋 大夢<sup>†</sup>

Hiromu Takahashi

菅沼 拓夫<sup>†,§</sup>

Takuo Suganuma

## 1. はじめに

Face-to-Face の対話型コミュニケーションの状況把握は、特定応用における人同士のインタクシオンの効果検証や、最近では感染症予防などで注目されている。

対話型コミュニケーション状況を把握し分析する方法として、センサを活用したデータ収集・分析手法が検討されている。センサは大別すると設置型と装着型に分けられるが、設置型センサを用いる手法は個人識別のためのコストがかかり [1]、また俯瞰的にデータ収集を継続するためプライバシーに課題がある。一方、装着型センサを用いる手法は、設置型センサで挙げた課題を解決する一方、既存研究においては、十分なセンシング精度を確保した上での小型化・省電力化と効率的なデータ集約手法に課題がある。

本研究では、装着型センサを用いたコミュニケーション状況の効果的な把握を目的として、行動認識センサおよびデータ収集システムを提案している。これまで、当該センサの設計・実装や、消費電力等の動作性能の評価を行ってきた [2]。本発表では、行動認識センサによる対話型コミュニケーション検出の精度に関する実験と評価について述べる。

## 2. 関連研究

対面型コミュニケーション状況把握のための既存研究として、被験者の胸部に装着するネームプレート型のセンサデバイスを用いる手法が提案されている [3]。この研究では、センサに搭載した 1 つの MEMS マイクを用いて発話状況の収集を行っており、外部からの音を拾いやすいことから、対象者の発話状況収集の正確性に課題がある。また、対象者の位置検出や近接検出に Bluetooth 等の通信を用いており、消費電力が大きく、また連続動作時間のために必要なバッテリーが大型化することから軽量のセンサの実現に課題がある。さらに、センサデータの回収方法として有線接続またはセンサに搭載した SD カードを取り出すなどの手動によるデータ収集が行われており、リアルタイムでの解析が不可であること、データ回収作業が煩雑であることなどの課題がある。

## 3. 行動認識センサとデータ収集システム

### 3.1. 提案センサ・システムの概要

本研究では、装着型センサを用いたコミュニケーション状況の効果的に把握するための、行動認識センサおよびデータ収集システムを提案している。想定環境として、コミュニティセンター等の室内のパブリックスペースでの人同士の交流を想定し、その際に行われる対話型コミュニケーションを認識対象とする。

行動認識センサは、胸部に装着するネームプレート型のセンサデバイスとし、発話状況、身体の動き、室内でのおおまかな位置を検出するとともに、効率的なデータ収集のため、省電力で広範囲通信が可能な LPWA 通信を用いた高頻度での連続データ送信を行う。

### 3.2. 全体システム構成

図 1 に全体システム構成を示す。本システムは行動認識センサ、アクセスポイント、および RSSI モニタから構成される。行動認識センサは被験者に装着し、内蔵されたセンサから対話型コミュニケーション状況を把握するためデータを収集し、LPWA 通信によりアクセスポイントに送信する。アクセスポイントは対象施設に 1 つ設置し、行動認識センサから受信したデータをデータベースに保存する。RSSI モニタは位置検出のため、対象施設の対話型コミュニケーションが行われやすいポイントごとに 1 つ以上設置し、行動認識センサがデータ通信のために発する電波を受信して、その信号強度 (RSSI 値) をデータベースに記録する。

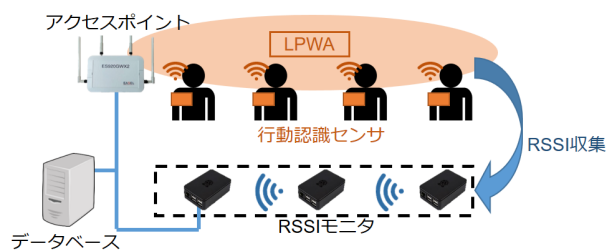


図 1: 全体システム構成

### 3.3. 行動認識センサの設計・実装

行動認識センサは、2 つの MEMS マイクと加速度センサを搭載し、音源である口元からの距離により生じる音圧差により他者の発話や環境音などの雑音の影響を除去した装着者の正確な発話状況を収集する。加えて、加速度センサにより身体の動きの情報を収集する。また、センサが送信する無線の RSSI 値から、RSSI モニタからの大まかな距離を推定する。行動認識センサに使用するセンサモジュールを表 1 に、実装したセン

<sup>†</sup>東北大学 大学院情報科学研究科, Graduate School of Information Sciences, Tohoku University

<sup>‡</sup>仙台高等専門学校, National Institute of Technology, Sendai College

<sup>§</sup>東北大学 サイバーサイエンスセンター, Cyberscience Center, Tohoku University

表 1: 搭載モジュール

モジュール	型番	製造会社
加速度センサ	KX224-1053	Kionix Inc.
マイク	SPU0414HR5H-SB	Knowles Electronics
マイコン	Lazurite 920J	ラピスセミコンダクタ
LoRa	ES920LR	株式会社 EASEL

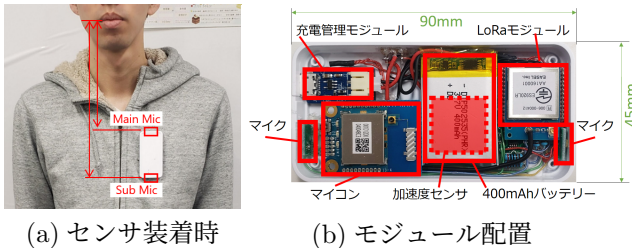


図 2: 実装した行動認識センサ

サを図 2 に示す。

実装したセンサ全体の大きさは H 90 mm × W 45 mm × D 12 mm, 重量は 46 g となった。センサの平均消費電流は 5.17 mA であるため, 使用予定の 480 mAh のバッテリーでの稼働の場合, 連続稼働時間は 93 時間と見積もられる。

#### 4. 実験

実装したセンサを用いて, 対話型コミュニケーションにおける発話検出の精度を実験により確認する。本実験では, 3 名での対話を想定し, 話者 B と C を感染症拡大防止の観点から自走式ロボット「Double」2 台としてデータ収集を行った。各話者は胸に当たる位置にセンサを装着し, 反響の影響が無い十分な広さの屋内において, それぞれが 1 m の距離をとり, 立ち姿勢で中央を向き 3 名が対面した状態とした。音声スクリプトは事前に設定し, 実験開始前に録音した音声を Double で再生し, タイミングを合わせて発話することで 30 秒間の対話を再現した。

実験中に録音した記録からラベル付けした正解の対話状態を図 3 に, センサにより収集した対話状態を図 4 に示す。発話状態のウィンドウは 0.1 秒ごとに判定し, 発話を有色で示している。図 3 と図 4 を比較すると, 話者 C は概ね正しく対話状態を示している。一方, 話者 A, B においては, 未発話時における誤った発話検出, 発話終端部での未検出, または短い発話時間での未検出が見られる。

これらの検出結果について, 正しく検出された発話時を True Positive(TP), 正しく検出された非発話時を True Negative(TN), 誤って検出された発話時を False

表 2: 検出結果の正確性と F1 スコア

センサ	TP	TN	FP	FN	正解率 (%)	F1 スコア
話者 A	86	146	28	40	77.3	0.717
話者 B	60	206	29	5	88.7	0.779
話者 C	45	246	0	9	97.0	0.909
全体	191	598	57	54	91.3	0.775

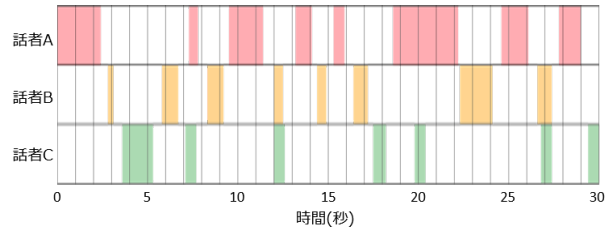


図 3: 正解の対話状態

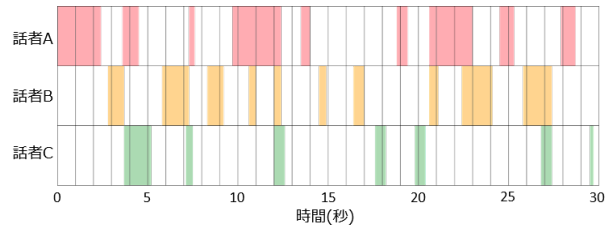


図 4: センサにより収集した対話状態

Positive(FP), 誤って検出された非発話時を False Negative (FN) とし, 混同行列と算出した正解率, および F1 スコアを表 2 に示す。

各センサの正解率はそれぞれ 77.3 %, 88.7 %, 97.0 % となり全体として 91.3 % の正解率となった。また, F1 スコアはそれぞれ 0.717, 0.779, 0.909 となり全体として 0.775 となった。特に話者 C が装着したセンサにおいては正解率が 97.0 %, F1 スコアが 0.909 とセンサの検出精度として良い結果となった一方, 話者 A は TP に対する FN の割合が大きくなっている。これは, 発話終端部の未検出, および短い発話の未検出部において, コミュニケーションの検出に十分な声量を得られなかったことから精度が低下したと考えられる。

以上から, 話者の声量によっては語尾や短い発話検出時に精度低下が見られるが, 全体としては正確にコミュニケーション検出が可能であると言える。

#### 5. まとめ

コミュニケーション状況の効果的な把握を目的とした行動認識センサにおける, 対話型コミュニケーションにおける発話検出精度を実験により評価した。

今後, 加速度データを用いた姿勢等の検出によるコミュニケーション検出の高度化を目指す。

#### 謝辞

本研究の一部は東北大学人工知能エレクトロニクス卓越大学院プログラムの支援を得て行われたものである。

#### 参考文献

- [1] X. Alameda-Pineda, et al., “Salsa: A novel dataset for multimodal group behavior analysis,” IEEE transactions on pattern analysis and machine intelligence, Vol. 38, No. 8, pp. 1707–1720, 2015.
- [2] 伊藤優樹ほか, “対話型コミュニケーションの状況を把握する行動認識センサの性能評価,” 情報処理学会第 83 回全国大会講演論文集 pp. 3-43-44, 2021.
- [3] O. Lederman, et al., “Rhythm: A unified measurement platform for human organizations,” IEEE MultiMedia, vol. 25, no. 1, pp. 26-38, 2018.