

対話型コミュニケーション状況把握のための 行動認識センサにおける発話状況収集機能の評価

伊藤 優樹^{1,a)} 上野 正義^{1,b)} 高橋 大夢^{1,c)} 千葉 慎二^{2,d)} 阿部 亨^{1,3,e)} 菅沼 拓夫^{1,3,f)}

概要：疫学分野の研究において、他者との対話型コミュニケーションと高齢者の健康寿命との関連性が示されているが、アンケート調査によるコミュニケーション状況の調査・分析は完全性や正確性に課題がある。本研究では、高齢者の健康寿命との関係性の明確化を目指し、他者との対話型コミュニケーション状況を定量的に把握するための行動認識センサを開発する。本発表では、行動認識センサにおいて複数人での対話コミュニケーション時の発話状況を収集する機能に焦点を当て、その評価について述べる。

1. はじめに

近年、日本国内において少子高齢化が問題となっている [1]。高齢者の増加によって介護・医療費の増大が予想されており [2]、改善策の一つとして「健康寿命」の延伸が期待されている。健康寿命とは「健康上の問題で日常生活が制限されることなく生活できる期間」 [3] であり、健康寿命が他者との対話コミュニケーションに関連があることがわかってきている [4], [5], [6]。実際、対話型コミュニケーションが健康寿命の延伸と相関関係があることが疫学分野において示されている [7] が、従来の疫学調査では被験者へのアンケート調査による主観的評価が用いられており、網羅性や正確性に課題がある。

これらを確保する手法として ICT の利活用による定量的データ収集が考えられ、そのデータ収集手法としてセンサデバイスを活用する例がある。既存研究では、被験者が装着したセンサデバイスを用いて発話や動作のデータを収集・分析し、コミュニケーション状況を定量的に把握する手法が提案されている [8], [9]。これらの研究では、被験者の活動への支障を低減するため、センサデバイスの小型化・軽量化が検討されているが、十分なセンシング精度の

確保、センサからのデータ収集手法に課題がある。

そこで本研究では、高齢者の対話型コミュニケーション状況の定量的な把握のための行動認識センサおよびデータを収集するシステムを提案する。行動認識センサは装着者である高齢者への負担を低減するため、必要十分なセンサにより構成し、可能な限り小型・軽量とする。センサからのデータ伝送には省電力で広域通信が可能な LPWA 通信を導入し、LPWA 通信の RSSI により装着者の大まかな位置を検出する。以上の取得データから、被験者の対話型コミュニケーションの状況として、対話グループの構成員やその中でのコミュニケーション上の役割等を検出する。

本提案でのセンサおよびシステムの実現により、前述した疫学調査における課題を解決するとともに、コミュニケーション状況の可視化やリアルタイム支援システムによるコミュニケーションの活性化が期待できる。

当該センサにおいて、複数人での対話時における発話状況を収集する対話状況収集機能について、軽量・小型化を実現するために MEMS マイクを用いた音声収集と処理手法を実装した行動認識センサを用いた実験を行った。本発表では、その評価について述べる。

2. 既存研究

既存研究では、センサデバイスを用いた音声や動きのデータを収集・分析する方法によるコミュニケーション状況把握手法が提案されている [8], [9]。これらの研究では、被験者がセンサを装着することで、定量的なデータ収集によるコミュニケーション状況把握を実現している。また、被験者の装着による負担を軽減するために、センサの小型・軽量化を図っている。しかし、既存研究では以下の 2 点が

¹ 東北大学大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University
² 仙台高等専門学校
National Institute of Technology, Sendai College
³ 東北大学サイバーサイエンスセンター
Cyberscience Center, Tohoku University
a) yuito@ci.cc.tohoku.ac.jp
b) mueno@ci.cc.tohoku.ac.jp
c) t.hiromu@ci.cc.tohoku.ac.jp
d) chiba@sendai-nct.ac.jp
e) beto@tohoku.ac.jp
f) suganuma@tohoku.ac.jp

課題として挙げられる。

(P1) 位置・近接検出のための電力消費量の増加

既存研究では、位置検出や近接検出に複数の赤外線通信モジュールまたは Bluetooth の RSSI を用いている。赤外線通信を用いる手法では精度向上のために赤外線モジュールの搭載数に応じて消費電力が増加しており、Bluetooth の RSSI を用いる手法はセンサに搭載したモジュールの受信状態が続くことから消費電力が大きくなっている。高齢者を対象とする本研究では、必要十分なセンサを装備してセンシング精度を保ちつつ、被験者の負担を低減するために消費電力の低減による必要最小限のバッテリーモジュールの採用により小型軽量化する必要がある。

(P2) データ回収のためのリスクや手間の増大

既存研究では、デバイスの稼働時間の観点から定期的な PC への USB 接続による充電時間を利用したデータ回収手法が用いられている。これは、リアルタイムでの解析が不可であること、省電力化により1度の充電で長期運用となる場合においてセンサデバイスの故障などによりデータ消失のリスクが増大するという問題がある。本研究では、最終的にはコミュニケーション状況に合わせたアドバイスや、支援システムによるコミュニケーション活性化を行いたいため、オンラインでのデータ収集・解析が必要である。

3. 行動認識センサとデータ収集システム

3.1 概要

本研究では、以下の2点を提案し、それを用いて高齢者の対話型コミュニケーション状況の定量的な把握のためのシステムを実現する。なお、高齢者が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交流する場面を想定し、その際に行われる対話型コミュニケーションを対象とする。

(S1) 小型軽量・低消費電力型行動認識センサ： マイクと加速度センサのみを用いた小型軽量・低消費電力の行動認識センサ。マイクからの音量レベルと加速度センサからの身体の動きの情報から、対面型コミュニケーションの状況を認識する。また、位置検出には特別なセンサを導入せず、センサから送信される電波の信号強度を利用して、簡易的に行う。以上によりセンサデバイスの搭載モジュールを少数にとどめ、小型軽量化、かつ低消費電力化を実現し、(P1)の課題を解決する。

(S2) LPWA による省電力な高頻度データ集約システム： 省電力で広範囲通信が可能な LPWA 通信を用いて、効率的に連続データを収集する。1回当たりの通信可能なパケットサイズを小さくすることで、高頻度での連続データ送信を可能とする。また、データを行動認

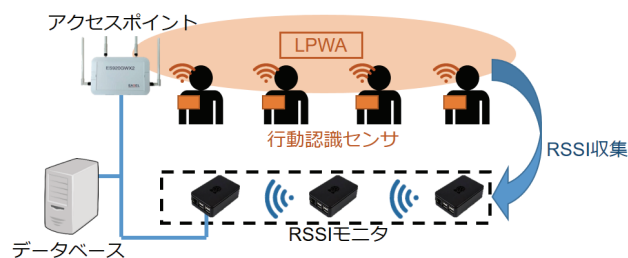


図1 全体システム構成

識センサ内に蓄積せずに連続的に基地局に送信するため、データ蓄積装置のために必要な電力を削減することが可能である。以上から、消費電力を抑えつつオンライン処理が可能なリアルタイムデータ収集を実現し、(P2)の課題を解決する。

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、RSSIモニターで構成される。

行動認識センサは被験者に装着する。行動認識センサに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LPWAパケットをHTTPパケットへ変換し、分析をおこなうためのデータベースにリアルタイムに保存する。

RSSIモニターは、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、その信号強度(RSSI)を記録してデータベースにリアルタイムに保存する。RSSIモニターは、対話型コミュニケーションが行われる可能性が高いスポット内と、それ以外の場所では見通し距離が大きくなる天井等に距離を空け設置する。

3.3 行動認識センサの設計

3.3.1 行動認識センサの概要

行動認識センサでは、高齢者の対話型コミュニケーション状況を把握するために、装着者の発話状況、および身体の動作状況を観測する。具体的には、発話状況はマイクによる100ms毎の発話パワー、身体動作状況は加速度センサによる100ms毎の加速度をそれぞれ収集し、LPWA通信によりアクセスポイントを介してデータベースに送信する。センサはできる限り小型・軽量とし、音声を集めるため被験者の胸部に装着するバッジ型とし実装した。実装した行動認識センサの内部構造を図2、センサ装着時の外観を図3、使用したセンサモジュールを表1に、それぞれ示す。また、行動認識センサによるデータ生成フローを図4に、収集したデータからコミュニケーション状況を分析するフローを図5に、それぞれ示す。

表 1 搭載モジュール

モジュール	型番	メーカー
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
LPWA	ES920LR2	株式会社 EASEL

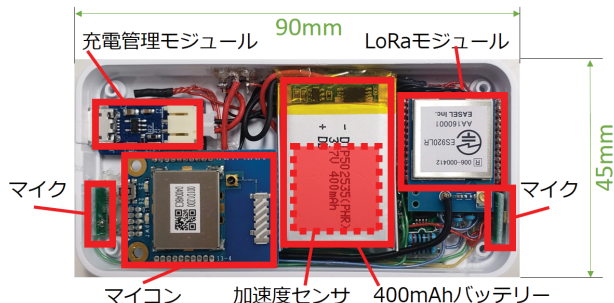


図 2 実装した行動認識センサの内部



図 3 行動認識センサの装着時の外観

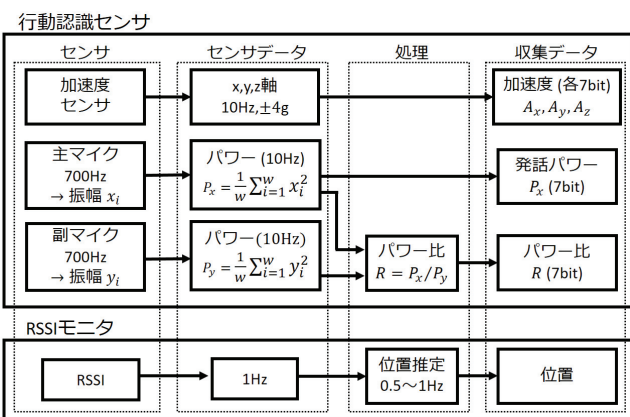


図 4 行動認識センサによるデータ生成

3.3.2 発話状況収集機能

正確な発話状況を取得するためには、環境音や他者の発話等の雑音を除去し、装着者の発話のみを取得する必要があるが、装着者の負担軽減のために可能な限り小型・軽量とする必要がある。そのため先行研究 [10] に基づき 2 つの MEMS マイクを用いて、それぞれのマイクが音源からの距離差により生じるパワー差から装着者の発話とそれ以外の音を判別することで、発話状況を取得する。パワー差は、音源である装着者の口元に近い点に主マイク、主マイ

行動認識のための分析

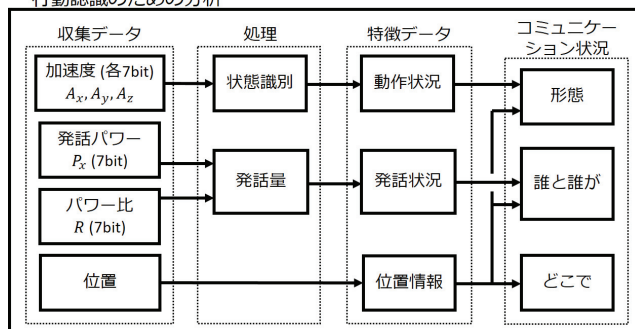


図 5 サーバによるコミュニケーション状況の分析

クより口元からさらに距離を離れた位置に副マイクを配置することで発生させる。発話状況は情報量、プライバシーの点からマイクが取得するパワー情報により収集することとする。2つのマイクで取得した音声をマイコンに搭載された A/D コンバータで 700 Hz でサンプリング、12 bit で量子化した振幅を取得する。1 フレームは 48 ms とし、1 フレーム毎のサンプル数 $w = 32$ の主マイクから得られた振幅 x_i の平均振幅 P_x を (1) 式を用いて、また副マイクから得られた振幅 y_i の平均振幅を (2) 式を用いて、それぞれ算出する。

$$P_x = \frac{1}{w} \sum_{i=1}^w x_i^2 \quad (1)$$

$$P_y = \frac{1}{w} \sum_{i=1}^w y_i^2 \quad (2)$$

また、得られた 2 つのマイクのパワーの値の比であるパワー比 R を次のように算出する。

$$R = P_x / P_y$$

得られたデータのうち主マイクの平均振幅 P_x を、発話状況における発話パワーとして、またパワー比 R を、得られた発話パワーにおける装着者の発話の成分を示す値として、それぞれ LPWA を用いて送信する。送信のためそれぞれを 7 bit へ丸め込み、1 秒間に 10 回ずつ、計 140 bit 取得する。コミュニケーションの分析では、発話パワー P_x およびパワー比 R から発話時間および強弱を算出し、装着者の発話状況を分析する。

また、発話パワー P_x とパワー比 R から得られる装着者の周囲音のパワー、および位置情報から、対話型コミュニケーションを行っているグループを検知する。さらに、発話のタイミングや長さ等から、当該コミュニケーションの状況をより詳細に分析する。例えば、図 6 に示すデータからは、被験者 A, B, C が近接してコミュニケーションしていることがわかり、また、交互に話者を入れ替えつつ会話している状況が推測可能である。

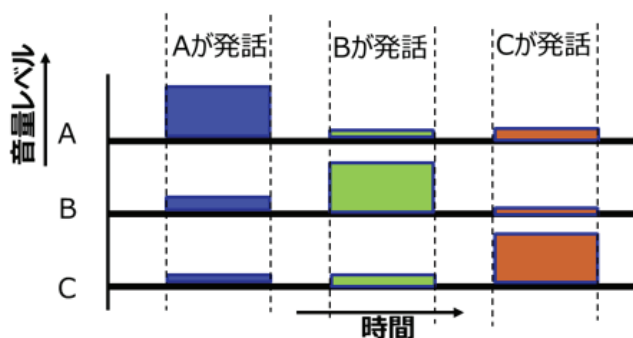


図 6 音量レベルによるコミュニケーション状況の推定

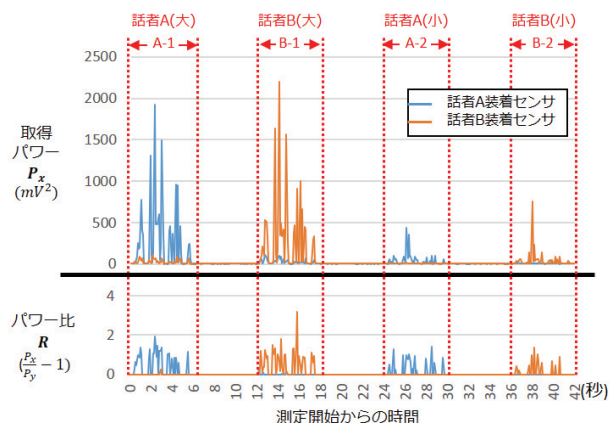


図 8 静音環境下での2台のセンサによる発話状況収集

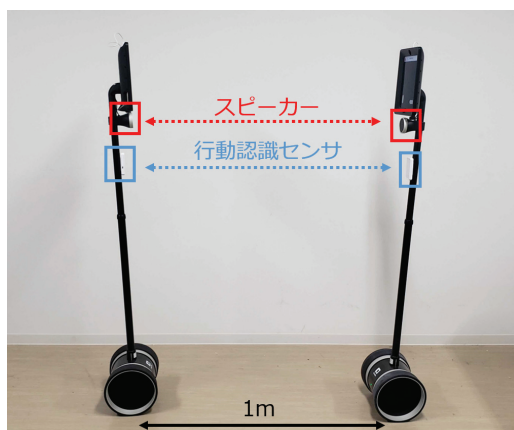


図 7 実験環境の概要

4. 実験

4.1 実験目的, 環境, 方法

実験環境の概要を図7に示す。本実験は、2名での対話を想定し、自走式ロボット「Double」[11]を2台用いて2名間での対話コミュニケーションを再現した。発話状況を収集するためにそれぞれのDoubleに行動認識センサを装着し、それぞれを「話者A」、「話者B」とし、2名間の距離を1mとした。静音環境及び騒音が含まれる環境のそれぞれにおいて、2名が声量の強弱を変化させた発話を行動認識センサにより収集し、環境音や他者の発話の影響を受けずに正確な装着者の発話が収集可能であることを確認する。

実験環境について、静音環境はマイクの収集データに影響の無い十分に静かな環境、騒音環境はスピーカーを用いてショッピングモールの環境音を再生し、2名の発話以外の音が存在する環境とした。騒音計により計測した行動認識センサ近傍における騒音レベルは、静音環境が32dBA、騒音環境が55~60dBAであった。それぞれの環境において、話者A、Bがそれぞれ順に、通常の声量(静音時においてセンサ付近で約70dBA)で約6秒間発話、次に、話者A、Bがそれぞれ順に小さい声量(約60dBA)で約6秒間発話した。

4.2 実験結果と考察

静音環境における収集結果を図8、騒音環境における収集結果を図9にそれぞれ示す。横軸は測定開始からの経過時間、縦軸は図の上部が行動認識センサに搭載されたマイクの取得パワー、図の下部が2つのマイクのパワー比を示す。

静音環境の実験結果である図8において、話者Aの発話区間であるA-1およびA-2区間では話者Aが装着したセンサの収集パワー及びパワー比の値が大きく変化しており、同区間における話者Bが装着したセンサの収集パワー及びパワー比に大きな変化が無いことがわかる。また、同様に話者Bの発話区間であるB-1およびB-2区間でも同様の結果が得られた。また、センサが同程度のパワーを収集するが話者が異なる区間、例えばB-1及びA-2区間における話者Aの装着するセンサの収集データは、話者Aが小さい声量で発話するA-2区間ではパワー比の変化が見られる一方、話者Bが通常の声量で発話するB-1区間にパワー比の値の変化が見られなかった。このことから、装着者の発話時のみパワー比の値が変化するため対話者の発話の影響を受けずに装着者の発話のみを収集可能であると言える。

次に、騒音環境の実験結果である図9において、静音環境と同様に装着者の発話区間におけるセンサの収集パワー及びパワー比の値に大きな変化が見られる一方、話者A、Bともに非発話区間における収集パワー及びパワー比の値の変化は微小であることがわかる。このことから、他者の発話および環境音の影響を受けずに装着者の発話のみを収集可能であると言える。

これらの実験結果から、行動認識センサは他者の発話・環境音の影響を受けずに装着者の発話時間および強弱の収集が可能であることを示した。

5. まとめ

高齢者の健康寿命との関係性の明確化を目指した、他者との対話型コミュニケーション状況を定量的に把握するた

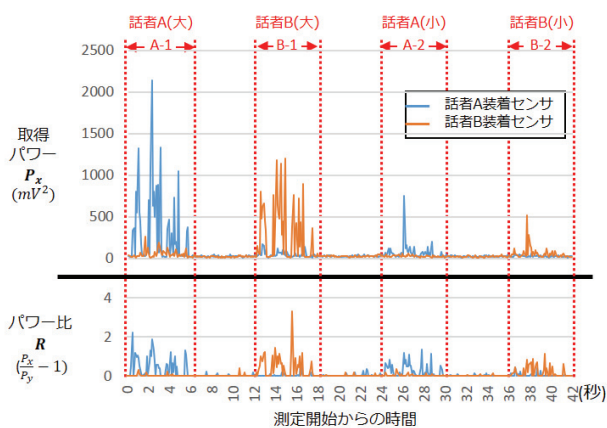


図 9 騒音環境下での 2 台のセンサによる発話状況収集

めの行動認識センサにおける、複数人対話時における装着者の発話状況収集の正確性を確認した。

今後、位置検出が可能となるよう複数の RSSI モニタを用いた位置検出手法の検討、近傍に複数の会話グループが発生する場合における発話状況の変化から会話グループを特定する手法を検討する必要がある。また、本研究での行動認識センサとデータ収集システムを、仙台市内のコミュニティセンターに設置し、実証実験を実施してその有効性を検証する予定である。

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] 内閣府：令和元年版高齢社会白書，内閣府（オンライン），入手先〈https://www8.cao.go.jp/kourei/whitepaper/w-2019/html/gaiyou/s1_1_1.html〉（参照 2019-12-01）。
- [2] 健康保険組合連合会：2025 年度に向けた国民医療費等の推移，健康保険組合連合会（オンライン），入手先〈https://www.kenporen.com/include/press/2017/20170925_1.pdf〉（参照 2019-12-01）。
- [3] WHO: Healthy life expectancy (HALE) at birth, available from 〈https://www.who.int/gho/mortality_burden_disease/life_tables/hale_text/en/〉 (accessed 2020-05-27).
- [4] Saito, T., Murata, C., Saito, M., Takeda, T. and Kondo, K.: Influence of social relationship domains and their combinations on incident dementia: a prospective cohort study, *J Epidemiol Community Health*, Vol. 72, No. 1, pp. 7–12 (2018).
- [5] Murata, C., Saito, T., Tsuji, T., Saito, M. and Kondo, K.: A 10-year follow-up study of social ties and functional health among the old: The AGES Project, *International journal of environmental research and public health*, Vol. 14, No. 7, p. 717 (2017).
- [6] Yokobayashi, K., Kawachi, I., Kondo, K., Kondo, N., Nagamine, Y., Tani, Y., Shirai, K., Tazuma, S. et al.: Association between social relationship and glycemic control among older Japanese: JAGES cross-sectional study, *PloS one*, Vol. 12, No. 1, p. e0169904 (2017).
- [7] Saito, M. et al.: Development of an instrument for

community-level health related social capital among Japanese older people: The JAGES Project, *Journal of epidemiology*, Vol. 27, No. 5, pp. 221–227 (2017).

- [8] Olguin, D. O. et al.: Sensible organizations: Technology and methodology for automatically measuring organizational behavior, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 39, No. 1, pp. 43–55 (2008).
- [9] 早川 幹ほか：ビジネス顕微鏡；実用的人間行動計測システムの開発，電子情報通信学会論文誌 D，Vol. 96, No. 10, pp. 2359–2370 (2013).
- [10] 伊藤優樹ほか：高齢者の対話型コミュニケーション状況把握のための行動認識センサの実装と評価，マルチメディア，分散協調とモバイルシンポジウム 2020 論文集，Vol. 2020, pp. 852–858 (2020).
- [11] BRULÉ: BRULÉ Japan - Double 2, available from 〈<https://www.brule.co.jp/robots/double2.shtml>〉 (accessed 2020-09-28).