

サウンド検出センサを用いたマイクの自動調整

池田 咲也^{1,a)} 串田 高幸¹

概要: 近年, オンライン上でのリモート会議が多くなっている. 特に IT 企業では在宅勤務がほとんどであり, オンライン上で社員同士コンタクトを取るのが当たり前になっている. リモート会議にも色々な方法があるが, 今回問題として挙げた事例は多人数で 1 つのマイクを使用する際, 発言者がマイクから遠い場合声を拾わない問題である. そのためマイク付きのサウンド検出センサ (FC-04) と ESP32 を使用して発言者の声を拾いやすくする. 実験方法としてサウンド検出センサを 8 個を円周上に等間隔で土台に取り付ける. サウンド検出センサには無指向性マイクが備えられており, 発言者の声を入力することができる. その音声を ESP32 に送り発言者の声是最も大きいセンサ以外のセンサを全て遮断することで, 無駄な音声や雑音が少なくなる. さらにセンサから読み取った音声を高速フーリエ変換を用いて雑音の除去を行い, さらに声が拾いやすく, ユーザに発言者の声が聞き取りやすくなることを実現する.

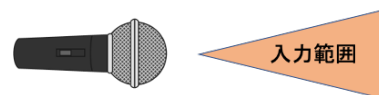
1. はじめに

背景

近年 IoT は急速な普及が進んでいる [1]. IoT 普及初期で稼働数が多いものはスマートフォンや通信機器であったが, 年々医療関連や自動車などでも多く活用されるようになった. IoT に分類されるモノはセンサー, カメラ, 無線通信などが搭載されているモノである [2]. センサーを活用した IoT ではモノの動きを検知し, 人やモノの状況を把握することができる. 例えば, 温度や湿度を測定を記録し, データの可視化を行うモノや, 室内に人がいなければ自動で照明を消す人感センサー, 声の入力や音の反応で動作する音センサーなどがある [3][4]. カメラ, 無線通信を活用した IoT では, 外出先から自宅の様子をリアルタイムで監視することができる [5]. 無線通信では他にも遠隔から操作できるモノがいくつもある. 例えばエアコンの温度を遠隔から管理し, 室内の温度をコントロールすることが可能である. IoT の開発, 研究でよく使用されるデバイスが ESP32 である. ESP32 とは Wi-Fi と Bluetooth が内蔵されたマイクロコントローラである [6]. ESP32 には数多くのセンサを接続することができ, 音の入力を行うセンサや, 温度, 湿度, 気圧を測定するセンサ, モノの検出センサなどを取り付けることのできるため, 幅広いものを測定できる. 前述にもある通り ESP32 は Wi-Fi と Bluetooth が内蔵されているため, 測定結果を離れた場所から読み取ることができるのが特徴である. 本論文では

ESP32 に FC-04 を繋ぎ, FC-04 のマイクを用いて声の入力を行う. FC-04 とはサウンド検出センサである. 特徴としてセンサに無指向性のマイクが搭載されており, ボリュームを絞ることでマイク感度も調整できる小型のマイクセンサである [7]. マイクには一般的に 2 種類あり, 単一指向性マイクと無指向性マイクがある. 単一指向性マイクとは正面の音のみを拾うマイクであり, 無指向性マイクは周囲の音を拾うマイクである. 単一指向性マイクは音を拾う範囲が狭いため雑音が少なく, 無指向性マイクは音の拾う範囲が広い雑音が多いのが特徴である [8].

単一指向性マイク



無指向性マイク

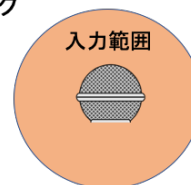


図 1 単一指向性と無指向性

¹ 東京工科大学コンピュータサイエンス学部
〒192-0982 東京都八王子市片倉町 1404-1
a) C0118014

本論文ではサウンド検出センサに搭載されている無指向性マイクを 8 つ使用し, 発言者の声を一番拾っているセンサ

以外を遮断し、高速フーリエ変換 (Fast Fourier Transform) することで無指向性マイクのデメリットである雑音を取り除く研究を行う。高速フーリエ変換とは離散フーリエ変換 (Discrete Fourier Transform) という処理を高速に行うアルゴリズムであり、ノイズの発生源を特定する際に使用されるものである [9]。本論文ではサウンド検出センサから入力した音声の雑音を、高速フーリエ変換で特定し、規定の数値以上のノイズを除去する。8 つの中で不必要な無指向性マイクの遮断、無指向性マイクのデメリットである雑音を高速フーリエ変換することで 360° の方向からでも雑音を取り除いた音声の入力が可能となる。

課題

概要でも述べたように、近年オンラインでのリモート会議が多くなっている。活用されているのは大学での講義を初め、企業でも取り入れられており、出勤する必要のないリモートでの会議は注目度が高くなっている [10]。しかし、リモート会議の問題として多人数で 1 つのマイクを使用する際に、マイクの向きによって声を拾いにくい問題と雑音が多くなってしまふ課題がある。原因として単一指向性マイクの使用である場合は声の入力範囲が狭く、無指向性マイクの使用であった場合は雑音が入りやすいためである。そのため本論文ではサウンド検出センサの FC-04 と ESP32 を採用することにより発言者の方向を特定して、発言者に最も近いサウンド検出センサのみから音声の入力をし、サーバ上で雑音を除去することを提案する。

各章の概要

第 2 章ではサウンド検出センサについて述べる。第 3 章では提案するシステムについて述べる。第 4 章では提案したシステムの実装や実験環境について述べる。第 5 章では実験の評価と分析について述べ、第 6 章では提案したシステムの議論を述べる。第 7 章で研究のまとめを行う。

2. 関連研究

FC-04 が使用されている研究としてノイズレベルを検出している論文がある [11]。これは予め閾値を設定し、その値に近いノイズの発生している場所を GPS を用いて特定するものである。本論文の研究では FC-04 のマイクで音声入力をし、音声をサーバに送信して高速フーリエ変換でノイズを検出するため、ノイズの特定は本論文の方が確実性があると言える。次に raspberry pi を用いた音の研究である [12]。この研究は raspberry pi も用いてダムや川の汚染度や騒音問題を課題とした研究で、学校や病院など騒音により支障を来す場所の近くの川の状態を監視する。システムが汚染や騒音を検出した場合、携帯電話やウェアラブルデバイスで受信することができ、川の氾濫による危険性にも対応できる。本論文と同じくこの研究では音の入力を

目的としてマイクを使用されている。本論文の研究は発言者に限りなく近いマイク以外を遮断をすることが目的ではあるが、音の大きさを常時測りノイズの特定も行えるため本論文のマイクシステムをこの研究にも役立てることが可能である。次に Arduino を用いた音の研究である [13]。この研究では Arduino を用いて大気汚染と音の騒音をモニタリングし、騒音を検出した際に当局に警報を発するものである。モニタリングはリモートで監視することが可能で、センサーから収集されたデータをクラウド上に保存し、Web ブラウザ上で可能である。音の入力は本論文と同じくサウンド検出センサを用いており、マイク部分で音の入力を行っている。主に検出するのはノイズ値で、ノイズが規定の dB を超えていると警報がなるシステムになっている。本論文と同じくサウンド検出センサを用いてノイズを扱っているが、この研究では音の入力のみであり、音声入力を可能としないため、本論文のマイクシステムを活用することにより人の声による騒音問題も解決することが可能である。次にセンサによる音声の入力の研究である [14]。この研究では騒音に焦点を当てた音の研究をしており、主に図書館や会社のオフィスなどの不必要な騒音を自動で検知し、ユーザーに通知をする研究である。騒音が声であった場合、規定の騒音レベルを超えていると自動で録音し保存され、社員への注意喚起などに使用される。この研究は本論文と同じく声の入力をメインとした研究であるが、本論文は 360° のどの方向からでも声の入力をしやすいシステムなため、図書館やオフィスであれば音を拾いやすくノイズの特定も可能なため本論文のシステムの方が優位性があると言える。

次に医療に使用される体温感知の研究である [15] この研究では本論文と同じくセンサを使用するが、この研究で開発されるのはワイヤレス体音モニタリングシステムであり、体音の感知をいつでも可能なシステムである。センサはスマートセンサを用いており、胸部に当てることで体音の測定が可能となる。体音をリアルタイムで感知できることから、非常に早い段階で病気の発生を発見できる。本論文のマイクシステムは入力された音声を一度サーバに送信しノイズの処理を行うため、この研究で本論文のノイズ処理システムを活用することで体音の感知の確実性の向上に繋げることが可能である

3. 提案

1 つのマイクを多人数で使用する場合、マイクの向きによって声を拾わない課題に対しての提案を行う。サウンド検出センサを円周上に 8 個用意し、45° ずつ等間隔で設置する。サウンド検出センサの配置を図 2 に示す。提案内容を図 3 に示す。サウンド検出センサには音声入力が可能となる無指向性のマイクが搭載されている。まず 1 で発言者の声はセンサによって入力する。2 では 1 で入力された音声を ESP32 を介す。この時点で発言者の声の値が規定より低

いセンサを全て遮断する.3で規定の値より高かったセンサから得られた音声を一度サーバに送信する.ここでサーバに送信する理由は音声のノイズを高速フーリエ変換によって除去を行うためである.4で高速フーリエ変換を用いてノイズ除去を行う.5では発言者の輪郭を出した音声をESP32に送り,6では1~5までの工程で聞き取りやすくなった音声をユーザに送信する.

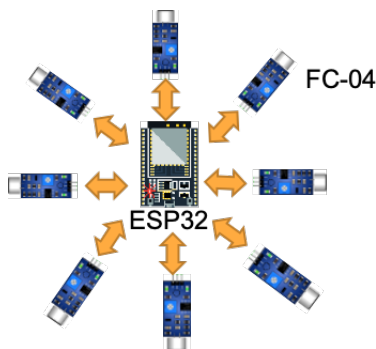


図 2 提案

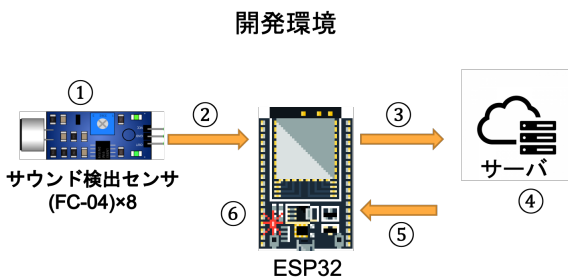


図 3 提案 2

4. 実装と実験環境

4.1 実装

本システムで使用したハードウェアを表 1 に示す.

表 1 使用ハードウェア/OS

| 表記名 | ハードウェア/OS |
|----------|-----------|
| Computer | ESP32 |
| Sensor | FC-04 |
| Server | VMware |

FC-04 は円周上に配置するため 8 個使用している. まず FC-04 に入力される信号を ESP32 内で受け取るためのプログラムを microPython 言語を用いて ESP32 に書き込む. ここで信号を受け取る理由はサーバに音声を送信する際に,

信号が Low のセンサをこの時点で遮断し, High の信号が出ている音声のみをサーバに送るためである. 続いて FC-04 に入力された音声を ESP32 を介し, サーバに送信されるプログラムを書き込む. 音声を受け取ったサーバ内ではノイズの除去をし, 除去した音声を ESP32 に送り返すプログラムを書き込む. ノイズ除去は FFT(高速フーリエ変換)の結果データにノイズとする境界値を入力し IFFT(逆高速フーリエ変換)をすることで可能となる. ノイズを除去した音声を元に, ESP32 内で FC-04 から High の信号で送られてきた音声から最も発言者に近いセンサ以外のセンサを遮断するプログラムを書き込む.

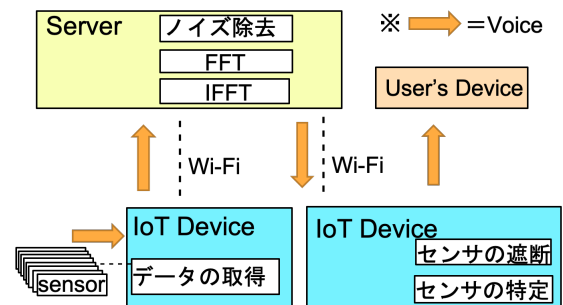


図 4 ソフトウェア構成図

4.2 実験環境

まず初めにサウンド検出センサに内蔵されているマイクで音を検出できるのか実験を行った. 実験写真を図 5 に示す. サウンド検出センサ (FC-04) は ESP32 と繋げること

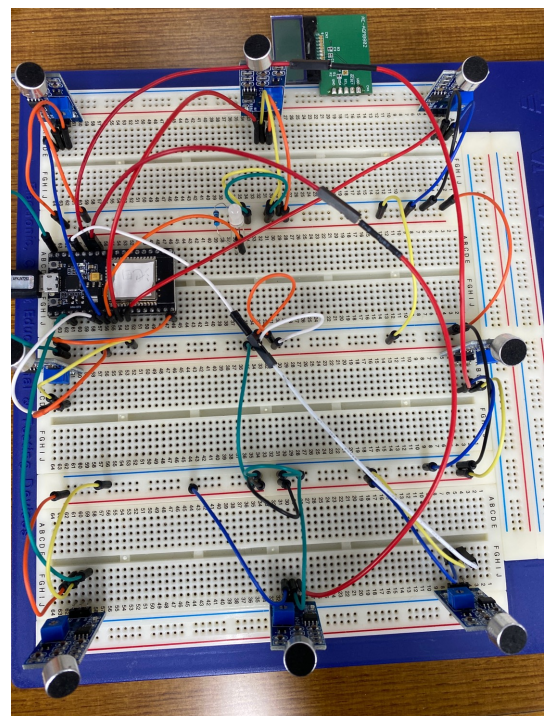


図 5 実験写真

で、自動でマイクの電源が入るため、それを可視化できるように出力するプログラムを作成する。今回は音を検知するとLEDライトが点灯するプログラムを作成して実験を行った。LEDライトの点灯が確認できたため、FC-04のマイクで検知した音の入力の大きさを可視化するプログラムを作成した。これは実装の章で述べたように、音声をサーバに送る前にLowの信号が出ているFC-04をこの時点で遮断し、Highの出ているFC-04の音声のみをサーバに送るためである。実験を行った結果、音の大きさを0.1秒ごとに0.00 1.00の大きさで出力されることに成功した。次に行ったのは音の大きさが0.70以下だった場合、センサを遮断するプログラムの作成である。作成したプログラムを実行した結果0.71以上のみの出力に成功した。

5. 評価と分析

検証方法として、まずPC(A)の入力マイクをFC-04とESP32を用いた環境で構築する。FC-04とESP32を卓上の中心に配置し、デバイスを中心に6人で囲むように着席する。PC(A)を図6に、配置図を図7に示す。



図 6 PC(A)

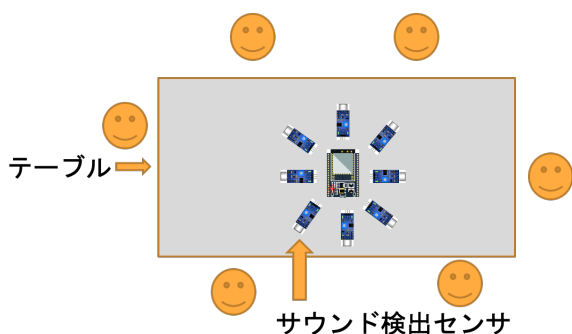


図 7 6人でセンサを囲んだ配置図

続いて別室にPC(C)を配置する。PC(A)とPC(C)を使用して実際にリモート会議を行う。それと同様にPC(B)では単一指向性マイクを用いた環境を構築し、前述と同様の環境でリモート会議を行い、どちらの環境が6人の発言者の声のスケール値をより高く出力できるのか検証する。PC(B)を図8に、PC(C)を図9に示す。評価方法として発言者の声

のスケール値が通常マイクよりFC-04とESP32を使用している方が大きいかを検証する。

単一指向性マイクを用いた環境



PC(B)

図 8 PC(B)

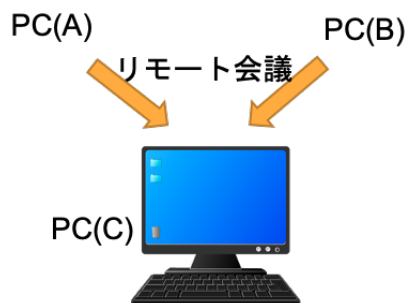


図 9 PC(C)

10秒間の音の検知の結果と、FC-04を8つ用いて実験を行った結果を図11に示す。

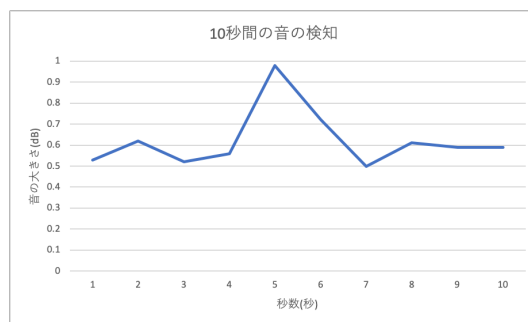


図 10 検知した音の波形

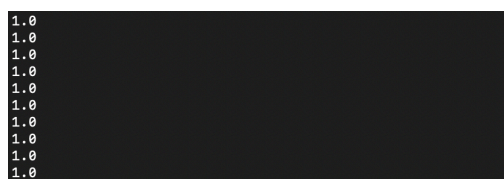


図 11 8つのセンサから検出した結果

研究全体の結果として音声の入力は成功出来なかったが、センサを使用して音を検知し、検知した音の大きさまで測定することに成功した。図 10 の結果から分かることは意図した音を出していない場合の雑音の平均値が 5~6 程度あり、開始から 5 秒後に大きな音を出したことにより数値が 0.98 まで上がった。これにより雑音と意図した音の大きさに差が出ているのがわかる。

続いて FC-04 を 8 つ用いて音の大きさが 0.70 以下だった場合に全てのセンサを遮断する実験で、図 11 の結果から 0.71 以下の音を出しているセンサを遮断することは成功しているが、0.71~0.99 までの数値を検出できておらず、1 のみ出力される結果となった。

6. 議論

本論文ではサウンド検出センサを用いて声の入力を行い、発言者の声が一番近いセンサ以外を遮断するシステムであったが、実際に検知できたのは声の入力ではなく音のみであり、課題解決には繋がらなかった。現状として課題解決に足りない要素は音声の入力、サーバへ音声を送信するプログラムである。まず音声を読み取れなかった原因としては FC-04 からアナログピンが取り出せなかったためである。FC-04 にはデジタル信号のみしか流れていないため、マイクから読み取れるのは音のみである。マイクで音声を読み取るにはアナログ信号が必要なため、アナログピンを取り出さなければいけない。アナログピンを取り出す方法として FC-04 の基盤にアナログピン用にピンを取り付ける箇所があり、VCC,GND,Signal Output の他に 4 本目としてピンを付け加える必要があった。次にサーバへ音声を送れなかった原因として、サーバの用意と高速フーリエ変換の知識が欠けていたからであった。今後の展望としてアナログピンの取り出し、サーバの用意、高速フーリエ変換の演算式の理解に励み、課題解決に望みたい。そして本研究の見直す部分として高性能のコンデンサマイクの方が周囲の音を拾いやすく雑音も少ないという課題であるコンデンサマイクとは低域から高域まで音の入力が可能で、音の拾う幅が広くさらには高音質なため、多人数で 1 つのマイクを使用する際に多く利用される。本論文で使用するサウンド検出センサは音質の悪い無指向性マイクであり、雑音も一度サーバに送信して高速フーリエ変換を使用して取り除くため手間が多い。それに対しコンデンサマイクは回線も使わず、ハードウェアもマイク単体で解決してしまうため遅延やハードウェア間での問題も少なく上位互換と言える。この問題を解決するためハードウェア間の遅延を抑え、雑音の除去も高性能なものに仕上げたい。

7. おわりに

本論文ではサウンド検出センサを 8 つ ESP32 に接続し、360° どの方向からでも発言者の声を拾い、一番近いセ

ンサ以外を遮断することで声を拾いやすく不必要な雑音を除去する研究であった。声の入力はサウンド検出センサ (FC-04) を使用し、声の大きさを検知するシステム、センサの遮断をするシステムは ESP32 で行った。結果から検知できたのは声の入力ではなく音の入力のみであり、課題解決には繋がらなかった。しかし、課題解決のため議論で挙げたように声の入力、サーバへ音声を送信するプログラムを作成し、次の研究で行いたい。

参考文献

- [1] Gubbi, J., Buyya, R., Marusic, S. and Palaniswami, M.: Internet of Things (IoT): A vision, architectural elements, and future directions, *Future generation computer systems*, Vol. 29, No. 7, pp. 1645–1660 (2013).
- [2] Lee, H., Sin, D., Park, E., Hwang, I., Hong, G. and Shin, D.: Open software platform for companion IoT devices, *2017 IEEE International Conference on Consumer Electronics (ICCE)*, IEEE, pp. 394–395 (2017).
- [3] Gondchawar, N., Kawitkar, R. et al.: IoT based smart agriculture, *International Journal of advanced research in Computer and Communication Engineering*, Vol. 5, No. 6, pp. 838–842 (2016).
- [4] Sani, A. and Beauty, P.: Smart framework for environmental pollution monitoring and control system using iot-based technology, *Sens. Transducers*, Vol. 229, pp. 84–93 (2019).
- [5] Tresanchez, M., Pujol, A., Pallejà, T., Martínez, D., Clotet, E. and Palacín, J.: An inexpensive wireless smart camera system for IoT applications based on an ARM CORTEX-M7 microcontroller, *Journal of Ubiquitous Systems and Pervasive Networks*, Vol. 11, No. 2, pp. 1–8 (2019).
- [6] Jamaludin, H. and Shah, P. S. I.: Designing ESP32 Base Shield Board for IoT Application.
- [7] Apostol, F. and Predusca, G.: Advanced Fan Controller Using Atmega 1284P, *The Scientific Bulletin of Electrical Engineering Faculty*, Vol. 18, No. 2, pp. 43–47 (2018).
- [8] Hirahara, T., Sawada, Y. and Morikawa, D.: Impact of dynamic binaural signals on three-dimensional sound reproduction, *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Vol. 2011, No. 6, Institute of Noise Control Engineering, pp. 1955–1962 (2011).
- [9] Carr, P. and Madan, D.: Option valuation using the fast Fourier transform, *Journal of computational finance*, Vol. 2, No. 4, pp. 61–73 (1999).
- [10] Sahi, P. K., Mishra, D. and Singh, T.: Medical education amid the COVID-19 pandemic, *Indian pediatrics*, Vol. 57, No. 7, pp. 652–657 (2020).
- [11] Khandelwal, S., Raghuvanshi, S., Kumar, R. and Dharpale, S.: Nuisance Monitoring (2018).
- [12] Singh, A., Pathak, D., Pandit, P., Patil, S. and Golar, P. C.: IOT based Air and Sound Pollution Monitoring System, *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, Vol. 6, No. 3 (2017).
- [13] Kaushik Vipul, R., Dabade, T. and Patil, V. N.: IOT based air and sound pollution monitoring system-a review (2019).

- [14] Kumar, M., Choudhary, R., Malik, K. and Sharma, R.: Noise Detector and Sound Recording System Using IoT, *EmErqing TrEnds in Big daTa, ioT and CyBEr sECu-riTy*, p. 49.
- [15] Dong, J., Jiang, H., Yang, K., Weng, Z., Li, F., Wei, J., Ning, Y., Chen, X. and Wang, Z.: A wireless body sound sensor with a dedicated compact chipset, *Circuits, Systems, and Signal Processing*, Vol. 36, No. 6, pp. 2341–2359 (2017).