

音源可視化技術の開発

(マイクアレイを用いた音源可視化システムの構築)

基盤技術部長 兼 機械システム科長 田 口 喜 祥

機械装置の異常を音情報から検知する技術が注目されている。従来は、熟練技術者による長年の経験と勘から通常と違う音が発生したことを認識し機械装置の異常を把握していた。近年、音データを FFT (Fast Fourier Transform) 解析し、異常音の特徴から監視を実施する試みや、IoT (Internet of Things) 技術を用いて音データを収集し AI (Artificial Intelligence) 技術により異常を検出するための試みが数多く実施されている^[1-4]。このような取り組みは、熟練技術者の技術伝承としても有効であると考えられる。

そこで、本研究開発では、音と画像の情報を組合せることで音源の可視化を行うことにより、監視対象としている装置から発生している音であることを認識し、監視することを特徴とした音源可視化技術を開発することを目的とする。マイクアレイと TV カメラ、もしくは熱画像カメラを組み合わせたセンサーアレイを製作し、AI 技術を用いて音の発生方向を認識し、監視を行う音源可視化装置を試作した。

1. 緒言

近年、製造現場では IoT や AI を用いた取り組みが注目されている。機械装置の異常監視は、生産現場の自動化や省力化に向けての重要な取り組みであり、さらなる発展が期待されている。その中でも音を用いた機械装置の監視技術については、FFT などの音響信号処理の活用や、AI と IoT を応用して異常監視を実施する研究や報告が数多くなされている^[1-4]。

機械装置の異常を動作音から認識する技術開発は関心が高く、長崎県内企業からも異常音の監視を行う IoT 機器を開発したいという相談がある。これまでの音を用いた異常監視は、音をマイクロホンで収集し、AD 変換器を用いてデジタルデータに変換した後、FFT 演算処理を行うことで周波数ごとの音圧信号に変換し、音圧信号の特徴量を抽出することにより行われていた。

一方、近年 AI 技術の進歩は目覚ましく、様々な分野で AI を応用する研究開発がなされてきている。例えば、カメラ映像から人物を認識する、画像から特定の部分を切り出すなど映像処理に関して AI 技術を用いることで高度な処理が可能となってきた^[5]。さらに、異常音の監視についても AI 技術の一手法であるオートエンコーダを用いることで、正常時に発生する音のみを学習し、学習した音と違った音が発生した場合は異常音と認識するなど、これまでより高度な認識が可能となってきた。

しかし、実際の工場内では様々な機械装置が稼働しており、単純にマイクロホンで音を収集しても様々な

機械装置の音を収集することになってしまい、監視対象となる機械装置からの正常な音の収集や機械装置ごとの異常音を監視することが現実的に難しいという問題点がある。

そこで、本研究開発では、音と画像の情報を組合せることで音源の可視化を行い、まず音源の発生位置を推定し監視対象から発生している音であることを識別し監視するための音源可視化技術を開発することを目的とする。

複数のマイクロホンを用いて構築したマイクアレイを用いて音を収集すると共に、マイクアレイと TV カメラや熱画像カメラを組み合わせて構築したセンサーアレイを用いて音源方向を可視化し、監視対象物から発生している音を識別する装置を開発する。

令和3年度は市販のフィールドレコーダーや FFT アナライザを用いて音源可視化装置のハードウェア構成を検討するための予備実験を実施した^[6]。令和4年度は収集した音データを基に音源位置を推定するためのニューラルネットワークの開発を行った^[7]。

令和5年度には、実際に工場内で使用できる音源可視化装置を開発するために、3D プリンタを用いてセンサーアレイを製作すると共に、FFT アナライザと計測用マイクロホンで実現していた装置を、実際に工場内に設置可能な組み込み用マイコンボードと安価なマイクロホンなどの部品を用いて再設計した。以上により、マイクアレイで取得した音データおよび TV カメラもしくは熱画像カメラで取得した映像を組み合わせることで音源方向を可視化する装置開発を行ったので報告する。

2. システム構成

開発する音源可視化装置のシステム構成を図1に示す。複数のマイクロホンとTVカメラもしくは熱画像カメラで構成されたセンサーアレイと、制御装置(PC)により音源方向を可視化し、PC やスマートフォンなどで閲覧できる装置開発を行う。

このような装置を開発するためには、複数のマイクロホンからの音データを基に音源の方向を推定するAIプログラムを開発する必要がある。開発するAIプログラムは、事前に音源方向が異なる複数の音データを収集し、収集した音データと音源方向を関連付けたデータセットを作成した後、ディープラーニング学習用ワークステーションを用いて音源方向を推定するAIの探索を行うことで製作した。

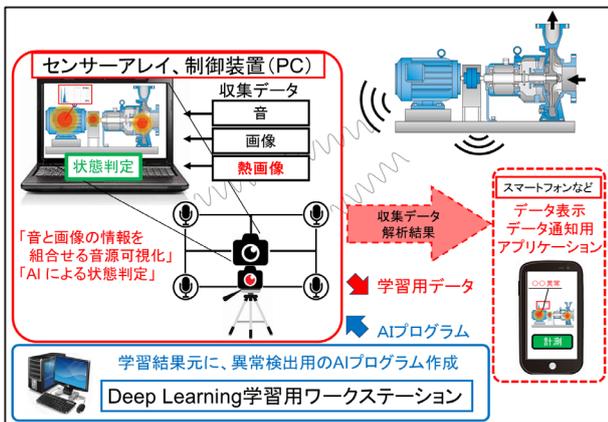


図1 システム構成

3. 音源可視化予備実験

考案したシステム構成で音源可視化が可能であるかについて予備実験を実施した。熱画像カメラを用いて、稼働中の機械装置を検出するために実施した予備実験結果を図2に示す。

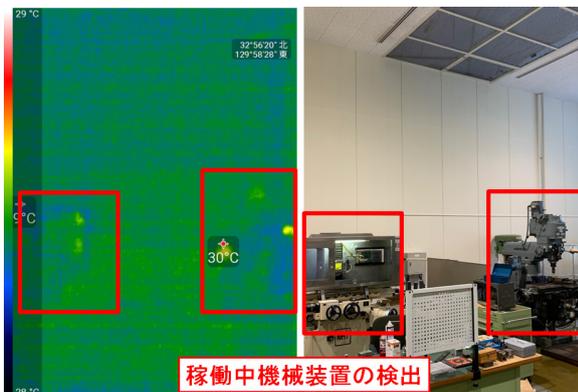


図2 熱画像カメラを用いた予備実験

この予備実験では、熱画像カメラ（Infinitegra製 OWLIFT Type-F）を使用して、稼働中の機器および可動部の判別が可能であるかについて評価を行った。実際に動作中の工作機械を撮影した結果を図2に示す。この予備実験により、稼働している箇所は周囲より温度が高くなる傾向があるため、熱画像カメラの画像で稼働中もしくは稼働箇所の判別が可能であることを確認した。

つぎに、マイクアレイを用いて音源の方向が推定できるかに関する予備実験を実施した。マイクアレイで取得した音データを用いて音源の方向が推定できれば、TVカメラの映像と組み合わせることで、複数の機械装置がある場合であっても監視対象からの音であることを識別することが可能となる。また、熱画像カメラからの映像と組み合わせることで、稼働中の装置や稼働箇所から発生している音であることを認識できると考えられる。

予備実験では、音データから音源方向を推定するため、無響室内の任意の場所にスピーカー1個を設置し、マイクアレイで音データを収集すると共に、スピーカーの位置を記録し、データセットとしてAIの学習に用いた。図3に無響室で実施した予備実験で用いたマイクアレイおよびスピーカーの写真を示す。マイクアレイのデータ記録には、6chの同期した音データをWAV形式データとして記録できる市販のフィールドレコーダー（ZOOM製、H8）を用いた。



図3 マイクアレイおよびスピーカー

図4に示した1から8までの位置にスピーカーを設置し、FFTアナライザから出力したホワイトノイズをスピーカーで再生し、それぞれの箇所で同期して収集した6chの音データをWAV形式で記録した。スピーカーの位置番号と6chの音データを用いてAI学習

用のデータセットを製作した。この予備実験に使用した機材を表1に、使用したFFTアナライザの写真を図5に示す。

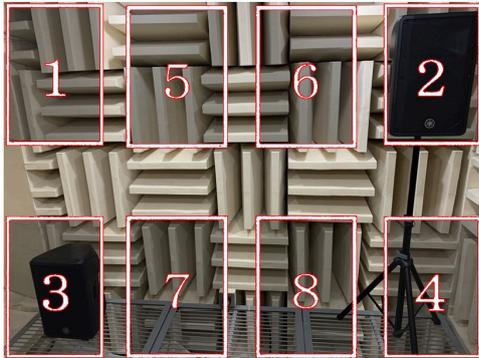


図4 データ取得時の音源配置

表1 音データ収集に用いた機器

レコーダー	H8 (ZOOM)
マイクロホン	ECM8000 (Behringer)
スピーカー	DBR12 (YAMAHA)
FFTアナライザ	DS-3000 (小野測器)
記録媒体	SDカード又はPC



図5 FFTアナライザ

表2 学習用ワークステーション構成

OS	Windows10Pro 64 bit
CPU	Core i9 12900K
GPU	GeForce RTX3090
RAM	DDR5 64 GB
ストレージ	M2-SSD 2 TB (PCI-E-GEN4)

収集したデータセットを用いて、ディープラーニング学習用ワークステーションにより、6 ch の音デー

タから、スピーカーの位置を推定する音源探査用ニューラルネットワークの探索と学習を実施した。学習用ワークステーションの構成について表2に示す。

ディープラーニングによる学習を行うために用いたソフトウェアは、Windows10 ローカル版の Neural Network Console (SONY) [8] である。バージョン 1.80 以降、WAV データを直接取り扱うことができるようになり、音を取り扱う分野でもより使用し易いものとなっており積極的に活用を進めている。実際に使用したバージョンは 2.10 で、WAV データの取り扱いについては CSV での相対ファイルパスを指定する形で利用可能である。

予備実験で探索した音源探査ニューラルネットワークの例を図6に、音源探査ニューラルネットワーク学習曲線を図7に、評価結果を図8に示す。

ニューラルネットワーク探索の結果、8か所の音源位置に対する音源探査の正解率 98.25% で判別できることを確認した。

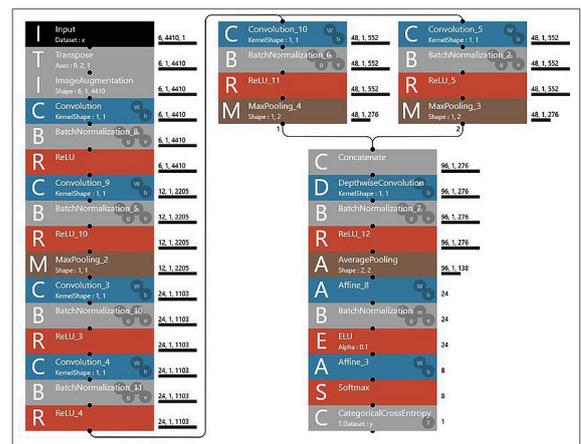


図6 音源探査ニューラルネットワーク

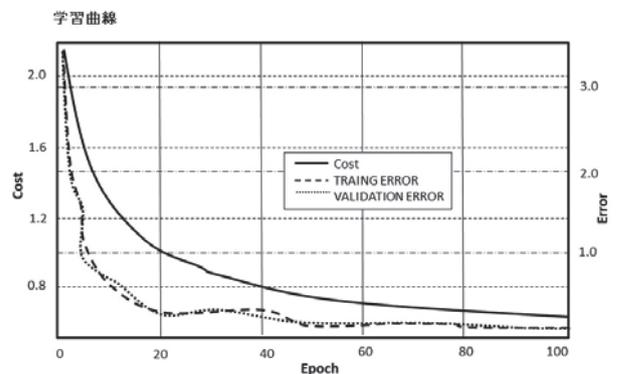


図7 音源探査ニューラルネットワーク学習曲線

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7
y=0	2000	0	0	0	0	0	0	0
y=1	0	1999	0	0	0	1	0	0
y=2	0	0	2000	0	0	0	0	0
y=3	0	0	0	2000	0	0	0	0
y=4	202	0	0	0	1798	0	0	0
y=5	0	3	0	1	39	1956	0	1
y=6	0	0	31	0	0	0	1967	2
y=7	0	0	0	0	0	0	0	2000
Precision	0.9082	0.9985	0.9847	0.9995	0.9787	0.9994	1	0.9985
F-Measures	0.9518	0.9989	0.9922	0.9997	0.9371	0.9885	0.9916	0.9992
Accuracy	0.9825	正解率 98.25%						
Avg-Precision	0.9834							

図8 音源探索ニューラルネットワーク評価

4. 音源可視化装置

予備実験により、マイクアレイを用いて同期して収集した音データとTVカメラもしくは熱画像カメラ映像を組み合わせることで音源可視化装置を開発できると推察できたので、工場内で使用することを想定した音源可視化装置の設計を行った。工場内に常時設置して使用することを前提とし、社内LANからネットワークを通して監視を行わせること、マイクアレイからの音データを同期して収集できること、および製造コストを考慮して、音データの収集をマイコンボード（SONY製、SPRESENSE^[9]）で実現することとした。FFTアナライザを用いた予備実験では、6chのマイクロホンのデータを使用した。SPRESENSEでのアナログ入力マイクロホン最大接続可能チャンネル数は4chまでである。計測原理を考えると4chでも問題ないと考え、マイクロホン4chで動作する音源可視化装置を試作することとした。

3D-CAD (Solid works office premium 2007) を用いてセンサーアレイの設計を行い (図9)、3Dプリンタを用いてセンサーアレイの機構部品を製作した。今回、TVカメラと熱画像カメラのレンズ中心位置高さの違いを考慮し、センサーアレイのカメラ取付部の高さをそれぞれのカメラに合わせた高さとした2種類のセンサーアレイを製作することとした。そのため、センサーアレイは2つの部品を組み合わせることで製作することとし、マイク1個とカメラを取り付ける箇所の部品をTVカメラ用、熱画像カメラ用の2種類製作し、マイク3個を取り付ける部品を1種類製作することとした。使用した3Dプリンタは短繊維CFRP材料であるOnyxを用いて造形が可能なMarkfoged X7である。センサーアレイの設計結果を図9に、試作に用いた3Dプリンタの写真を図10に示す。また、センサーアレイ機構部品として、マイク4個とTVカメラもしくは

熱画像カメラを取り付けて製作したセンサーアレイの写真を図11に示す。製作したセンサーアレイを用いて、音源可視化用ニューラルネットワークを製作するためにデータセットの収集を行った。

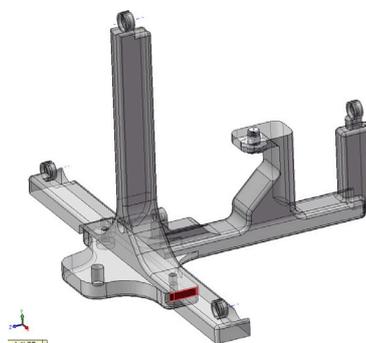


図9 センサーアレイの設計



図10 高強度3Dプリンタ



(a) TVカメラ (b) 熱画像カメラ

図11 試作センサーアレイ

今回は、USBスピーカーからホワイトノイズを発生し、図12に示した音源位置にUSBスピーカーを設置し4chの音データとスピーカー位置からなるデータセットを製作した。約8千点のデータセットのうち8割を学習用、2割を評価用としてニューラルネットワーク探索を実施した。探索の結果、9か所の音源位置に対する音源探索の正解率99.94%で判別できるニューラルネットワークを構築した。

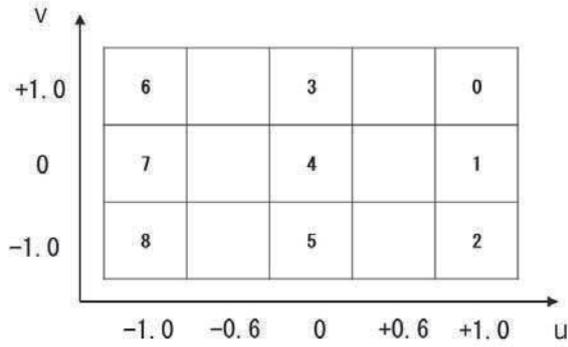


図 12 学習用音源配置

分類問題で高精度な推定が可能であったので、回帰分析の実施を試みた。画面座標を (u, v) とし、スピーカ位置を分類問題の9か所に加えて u 座標が 0.6 および -0.6 の6か所を増やしてデータの収集を行った。図 12 に示す (u, v) 座標系の場所に USB スピーカを設置し、4 ch の音データとそれぞれの (u, v) 座標をデータセットとして、4 ch の音データから (u, v) 座標を推定するニューラルネットワークの探索を試みた。学習用のデータセットを収集する実験を実施している無響室の様子を図 13 に示す。

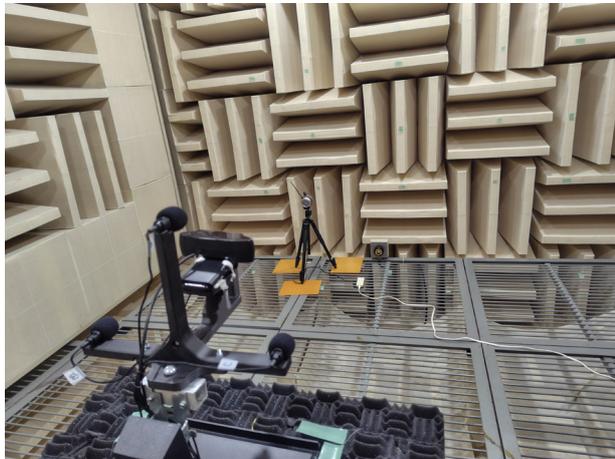


図 13 データセット収集実験

15,000 点のデータを収集し、12,000 点を学習用のデータ、3,000 点を評価用のデータとし 4 ch の音データから音源の方向 (u, v) を推定するニューラルネットワークの探索を行った。その結果、トレーニングエラー 7.35×10^{-4} 、バリデーションエラー 8.97×10^{-4} となるニューラルネットワークを探索した。

これらのことにより、ニューラルネットワークを用いた回帰推定により、4 ch の音データから音源の方向 (u, v) を推定することが可能となったが、マイコン

ボード SPRESENS のみではメモリ容量の関係で、探索したニューラルネットワークを用いた音源方向の推定、音源の可視化および可視化結果を表示するプログラムを実行することが困難であった。

そこで、4 ch の音データを同期して取り込む機能のみをマイコンボード SPRESENS で実行し、一定間隔で取り込んだ音データを制御用 PC に送信し、制御用 PC で音源方向の推定、TV カメラもしくは熱画像カメラと推定方向を組み合わせた音源の可視化画像表示、および PC やスマートフォンなどからブラウザで閲覧できるようにするための Web サーバプログラムを実行させることとした。開発した装置のシステム構成を図 14 に示す。

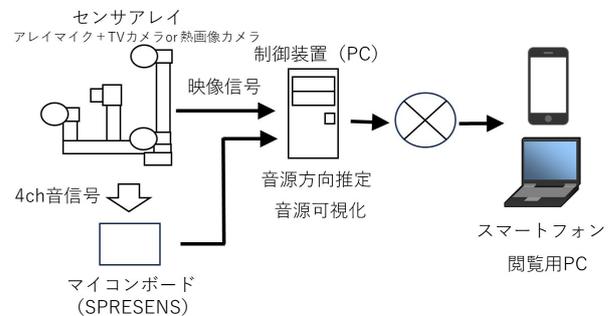


図 14 音源可視化装置システム構成

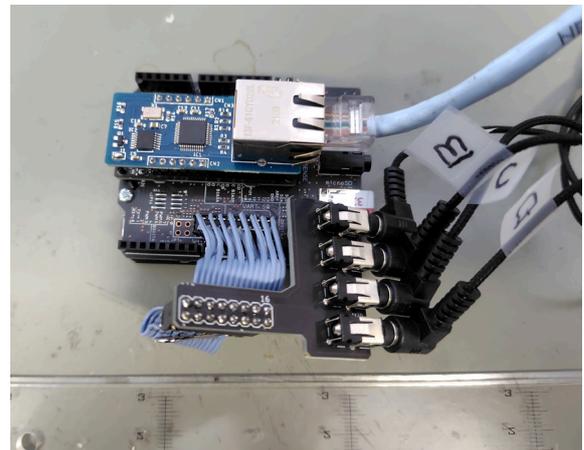


図 15 マイコンボード (SPRESENS)

表 3 制御用 PC の仕様

OS	Windows10
CPU	Intel (R) Core i7-10700F
メモリ	64GB
GPU	NVIDIA(R) GeForce RTX 3090Ti
ストレージ	SSD 2TB

同期した4 ch の音データ収集のために用いたマイコンボードSPRESENSの写真を図15に示す。マイクロホンを接続する端子とネットワーク接続のための拡張ボードが干渉したため接続用ケーブルを製作した。TVカメラや熱画像カメラの映像収集、SPRESENSEで収集した4 ch の音データからニューラルネットワークを用いて音源方向の推定、収集した映像と推定した音源方向を組み合わせて音源を表示し、スマートフォンやPCのブラウザで表示可能とするためのプログラムを制御装置(PC)で作動させた。今回使用した制御装置(PC)の仕様を表3に示す。

表3に示した制御装置(PC)で動作する音源可視化プログラムを開発した。開発したプログラムは音源可視化装置のセンサーアレイで収集した音データを基にAI処理により音源方向を推定し、TVカメラで撮影している映像に推定した音源方向を重ねて表示する機能を有する。音源可視化装置で撮影した映像はLANに接続されているPCやスマートフォンのブラウザから閲覧することが可能である。図16は、実験室内に小型スピーカーを設置し、この小型スピーカーからホワイトノイズの音を再生した際のブラウザ表示画面である。



図16 音源可視化実験 (TV カメラ)

TVカメラを用いた音源可視化実験の結果、開発した装置により音源方向を推定しTVカメラの映像に重ね合わせて可視化することで、取得した音データが監視対象の装置から発生している音データであることを確認可能となった。

5. 結言

音源方向と映像情報を組み合わせることで音源の可視化を行うことにより、監視対象としている装置から発生している音であることを確認し、監視することを特徴とする音源可視化技術を開発した。4 ch の音データによる音源位置9分類をニューラルネットワークにより99.9%以上の高精度で推定可能であることを確認した後、音データから音源方向を推定し、TVカメラ映像や熱画像カメラ映像と併せて音源方向を表示する機能を有する音源可視化装置を試作した。性能評価のための実験を行ったところ、音源方向と映像情報を基に音源位置を可視化し、対象とした装置からの音であることを認識できることを確認した。

今後、開発した音源可視化技術を用いて実際の工場内で稼働している機械装置の異常音解析や、音を用いたIoT機器等の応用開発に取り組んでいきたい。

参考文献

- [1] 植松 他：動作音から機器の異常を検知する異常音検知技術、NTT技術ジャーナル、vol. 29 No. 6, pp. 24-27, 2017. 6.
- [2] 高嶋：音響センシングによる異音検知、OKIテクニカルレビュー、第233号Vol. 86 No. 1, pp. 32-35, 2019.
- [3] 亀岡：深層学習モデルを用いた音声音響信号処理、計測と制御、計測自動制御学会、58巻、3号、pp. 195-202, 2019.
- [4] 木下：機械学習による機械稼働音監視とデータ拡張による学習効率化に関する実験的検討、公益社団法人計測自動制御学会中国四国支部、第28回学術講演会論文集、pp. 77-78, 2019.
- [5] 田口、久保田：AIを用いた監視装置の開発、長崎県工業技術センター研究報告、No51, pp. 10-15, 2022.
- [6] 久保田：音源可視化技術の開発、長崎県工業技術センター研究報告、No51, pp. 23-26, 2022.
- [7] 久保田：音源可視化技術の開発、長崎県工業技術センター研究報告、No52, pp. 19-21, 2023.
- [8] <https://dl.sony.com/ja/>, Accessed 2021.
- [9] 太田：SPRESENSEではじめるローパワーエッジAI、ISBN:4873119677, 2022.