

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)

機械システム科 主任研究員 久保田 慎一

音関連技術は非接触・非破壊検査技術としても注目され活用の場が広がっており、機器の異常監視やノイズキャンセリング、対象音の明瞭化などをはじめ AI・IoT などと連携し必要な情報を取り出すための情報選別技術へのニーズが高い。県内企業からも IoT (Internet of Things) 技術をはじめ、音を用いた機械装置の監視に関する相談が増えている。そのため、長崎県工業技術センターでは、新たに導入した音響計測機器を用いた技術支援を実施している。一方、近年では AI (Artificial Intelligence) を用いた研究が注目されており、音信号への応用が数多く報告されている^[1-4]。

そこで、本研究開発では、音と画像の情報を組合せる可視化機能に特長を有するマイクアレイや可視画像カメラに加え熱画像カメラを併用した音源可視化装置及び、連携するアプリケーションを開発することを目的とした。

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。機械装置の監視技術は、生産現場の自動化や省力化に向けて重要な技術であり、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、音響信号処理など活用するための研究や報告が多数なされている^[1-4]。

また、長崎県内の企業から、異常音の監視を行う IoT 機器を開発したいという相談が増えている。一方、近年 AI 技術を用いた画像処理の研究開発が数多くなされてきている。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。

そこで、本研究では複数のマイクロフォンを搭載したアレイマイクで音を収集するとともに、可視画像カメラに加え熱画像カメラを併用して稼働状況に関する情報を音と併せて収集することができる音源可視化装置を開発することを目的とする。

令和4年度は、音源位置を判別する音源探査の精度向上に向けたニューラルネットワークの開発と、対象に絞った音を抽出する AI を開発するために必要な学習用の音源データの収集を実施した。

2. 音源探査ニューラルネットワーク

令和3年度に計測用のコンデンサーマイクロフォン (Behringer 製 : ECM8000) 6本を使用して製作した図1のマイクアレイ^[5]を用いて、音源探査性能向上のためのニューラルネットワークの再構築を行った。こ

れまで4箇所のスピーカーに対し実施していた音源探査を8箇所のスピーカーに拡張する取り組みを実施した。音源探査に使用するデータを取得した際のスピーカー配置を図2に、製作した音源探査ニューラルネットワークを図3に示す。学習用ワークステーションにて、音源探査用に作成したデータセットを用いて、音源探査の学習を実施した。学習の結果、得られた学習曲線を図4に、評価結果を図5に示す。



図1 マイクアレイ (音源探査)

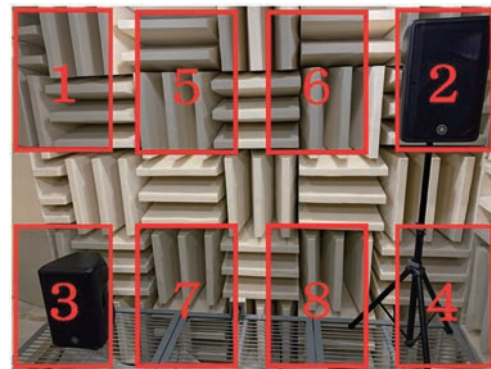


図2 データ取得時の音源配置 (音源探査)

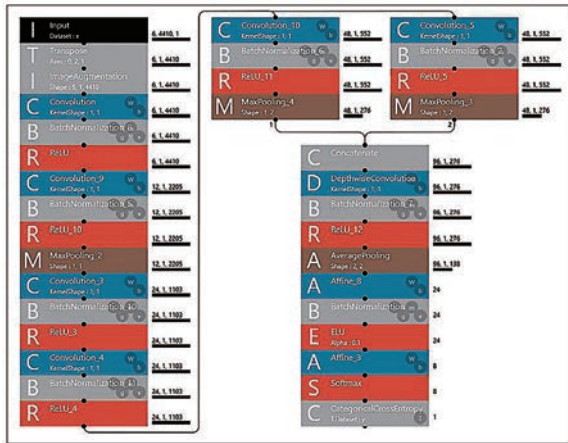


図3 音源探索ニューラルネットワーク

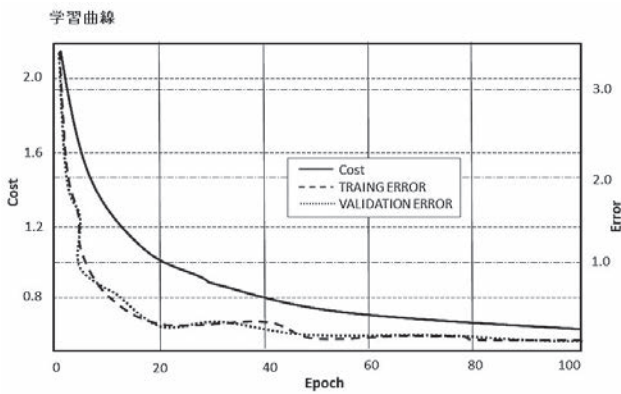


図4 音源探索ニューラルネットワーク学習曲線

	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	0.9085	0.9047	0.9095	0.9787	0.9094	1	0.9085
F-Measures	0.9518	0.9089	0.9022	0.9097	0.9371	0.9085	0.9916	0.9992
Accuracy	0.9825	正解率 98.25%						
Avg.Precision	0.9834							

図5 音源探索ニューラルネットワーク評価

ニューラルネットワーク探索の結果、8箇所音源位置に対する音源探索の正解率98%を達成した。

3. 対象音抽出AI 開発用学習データの収集

本来、目的の対象に絞った音を抽出するためには対象へセンサーを取り付ける必要があるが、それを不要とするための対象音抽出AIの開発に向けて、必要となる学習用音源データの収集を行った。

データ収集時の音源配置を図6に示す。4箇所のうち1箇所のスピーカーについては、出力する信号をFFTアナライザーへ入力信号として返し、学習データにおける抽出すべき対象音の教師データとしている。6chの同期の取れた音データの録音を実現するために、音源探査で使用したマイクアレイの配置を基に対象音抽出用マイクアレイ(図7)を製作し、FFTアナライザー(図8)を用いて、マイクアレイからの音声データと抽出する対象音を合わせた合計7chの音データの収録を行った。



図6 データ取得時の音源配置(対象音抽出)

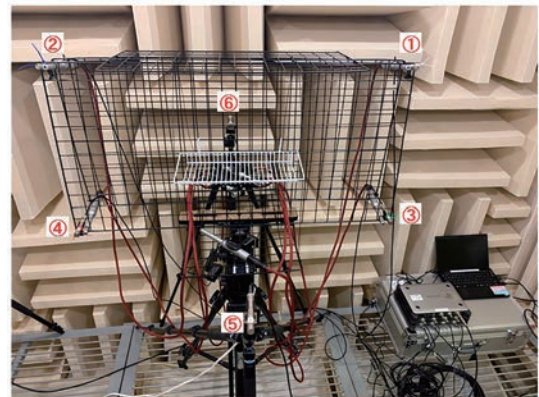


図7 マイクアレイ(対象音抽出)



図8 FFTアナライザー

本実験に用いた実験機器構成を表1に、実験条件を表2に示す。これらの音源データを活用することで、複数の音源がある環境の中で、対象の位置を指定することで、対象から出る音に絞った音データを生成するAIの開発が可能となると考えられる。

表1 実験機器構成

レコーダー (FFTアナライザー)	DS-3000(小野測器)
マイクロフォン	MI-1271(小野測器)
スピーカー (対象音、7chへ入力)	DBR12(YAMAHA)
スピーカー (その他)	SoundCore(Anker)
ノイズ発生器	DS-3000(小野測器)

表2 実験条件

場所	無響室
スピーカー位置	4箇所(図6参照)
入力信号種類	1ch~6ch: 各マイクロフォン入力信号 7ch: スピーカー出力信号 (出力元を分岐させ入力)
録音時間	600[s]

4. 学習用ワークステーションおよびソフト

本研究を通して利用した学習用ワークステーションの構成を表3に示す。学習用ワークステーションは、収集した複数チャンネルの音データを元に、音を分割するプログラムにより作成した学習用データセットを用いて、音源探査および対象音抽出の学習を実施することができる。学習の際は、主にGPUでの計算となる。

表3 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)

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)^[6]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している。

5. 結言

高精度な音源探査ニューラルネットワークの提案と、対象に絞った音を抽出するAIの開発について音源データの収集を実施した。実験の結果、6chの音データによる音源位置8分類を高精度に実施可能であることを確認した。

今後、対象音抽出に向けた学習の実施と、実際の生産現場における機械装置の異常発生音の検出に活用を進めたい。

参考文献

- [1] 亀岡：深層学習モデルを用いた音響信号処理、計測と制御、計測自動制御学会、58巻、3号、pp.195-202, 2019.
- [2] 株式会社アイ・エヌ・シー・エンジニアリング：振動のモニタリングサービス、IHI 技報 Vol.61, No.3, pp.20-23, 2021.
- [3] 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.
- [4] 木下：計測用マイクを用いたリアルタイム音信号AI処理システムの開発、鳥取県産業技術センター研究報告、No.23, pp.55-56, 2020.
- [5] 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、No.51, pp.23-26, 2022.
- [6] <https://dl.sony.com/ja/>, Accessed 2021.