

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方が広がっており、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アレイマイク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。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。

本年度は、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行った。

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。開発する装置は、複数のマイクロフォンと可視画像カメラ・熱画像カメラを搭載したセンサーアレイと制御用 PC、学習用ワークステーションに加え、データの通知や表示を行う端末、連携するアプリケーションソフトで構成される。学習用ワークステーションでは、取得した音や画像情報を AI 技術で学習および解

析し、稼働状況や異常に関する監視結果を Web 画面で提供する機能を搭載する。

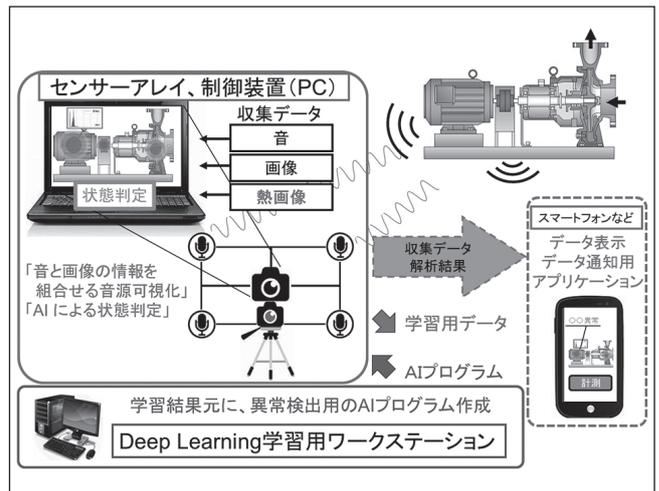


図1 システム構成

本研究で用いる技術として、熱画像カメラを用いた稼働中の機械装置の検出について、実施した予備実験の結果を図2に示す。複数の機械装置がある実験室において、通常のカメラ画像と熱画像を取得し、熱画像から稼働する機械装置の検出を目指す。今回の予備実験で、熱画像から機械装置の存在が確認できるため、今後は画角や解像度に加え、設定するレンジなどを含め、画像としての学習データの作成について検討していく。

また、音源探索の基本的な技術として、音の位相遅延を利用した音源探索の原理について図3に示す。複数のマイクロフォン間の距離と音の遅れから、音の到来方向を求めるこの技術は、各マイクロフォンにおける位相遅延を解析する事で音源位置の特定に用いることができる。

本研究では、上記の熱画像をはじめとした機械装置の検出技術と連携し、音の到来方向を指定することで、位相遅延を考慮した範囲を絞った集音の実現に利用する。以下では本年度行った予備実験におけるハードウェア構成について述べる。

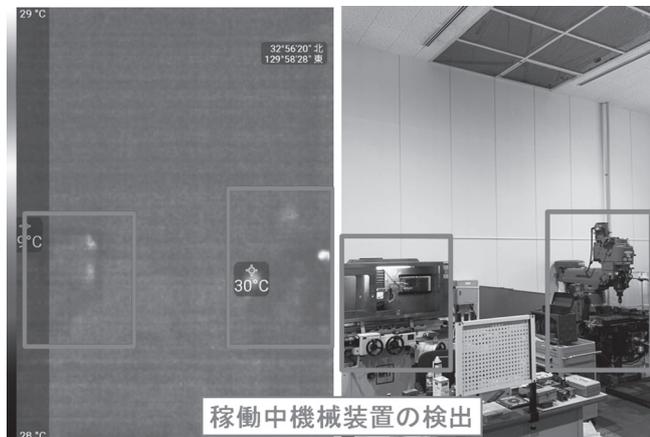


図2 熱画像カメラを用いた稼働監視

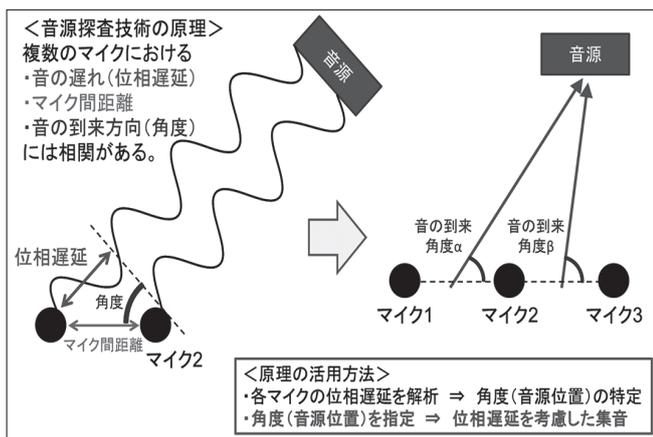


図3 音源探査の原理について

3. 実験装置の構成と試作ニューラルネットワーク

本年度実施した音源可視化装置計測原理確認実験では、大きく分けるとセンサーアレイ (図4) と学習用ワークステーション (図5) の2つで構成される装置を構築した。

センサーアレイのレコーダー部分には、6chの同期の取れた音データ (WAV形式) の録音ができる市販のフィールドレコーダーを用いた。マイクロフォンには計測用のコンデンサーマイクロフォンを使用している。構成を表1に示す。

また、学習用ワークステーションは、上記マイクロフォンで収集した6chの音データを元に、音を分割しデータセットを作成するプログラムにより作成したデータセットを用いて、音源探査の学習を実施する。学

習の際は、主にGPUでの計算となるので画像を扱うなど計算負荷によってはGPUの増強や、一度に読み込めるデータ量を増やすためにRAMの増設なども検討する必要がある。学習用ワークステーションの構成について、表2に示す。



図4 センサーアレイ



図5 学習用ワークステーション

表1 センサーアレイ構成

| | |
|---------|---------------------|
| レコーダー | H8 (ZOOM) |
| マイクロフォン | ECM8000 (Behringer) |
| スピーカー | DBR12 (YAMAHA) |
| ノイズ発生器 | DS-3000 (小野測器) |
| 記録媒体 | SDカード又はPC |

表2 学習用ワークステーション構成

| | |
|-------|----------------------------|
| OS | Windows10Pro 64bit |
| CPU | Xeon E3-1220 v3 |
| GPU | GeForce RTX3060 (メモリ 12GB) |
| RAM | 32GB |
| ストレージ | SSD 240GB (SATA3) |

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) [5] 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。試作したニューラルネットワークについて図6に示す。当初 6ch の音データを、それぞれ個別の入力データとして別々のネットワークに分けていたため学習に問題があった。そこで、6次元の1つの音データとして取り扱うネットワークを構築する事で、6ch 間の関連を含んだ学習ができた。

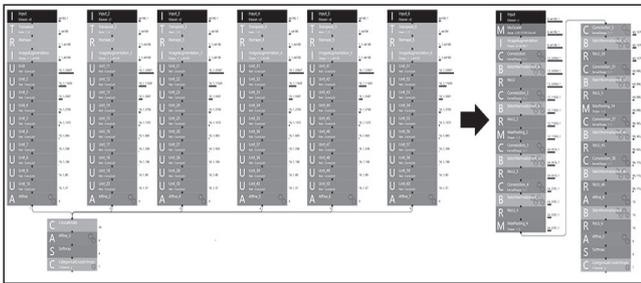


図6 試作ニューラルネットワーク

上記実験装置を用いて、表3の実験条件にて音源探査の実験を無響室で実施した。収集した音データを元に、データセットを作成し試作ニューラルネットワークによる学習を行った結果を以下に示す。スピーカー音源の位置を図7のように設置し、4箇所スピーカーに対し順番に音データを収集し、作成した学習用のデータセットにて学習を行った結果、図8の学習曲線及び図9の評価結果を得た。これにより4分類が可能である傾向が示されたので、今後は音源可視化装置の試作及び精度向上に向けて取り組むこととする。

表3 実験条件

| | |
|---------|-----------------------------------|
| 場所 | 無響室 |
| スピーカー位置 | 4箇所 (図7) |
| 出力信号種類 | 2種類 (RANDOM、 PSEUDO RANDOM) |
| 録音時間 | 600 [s] |

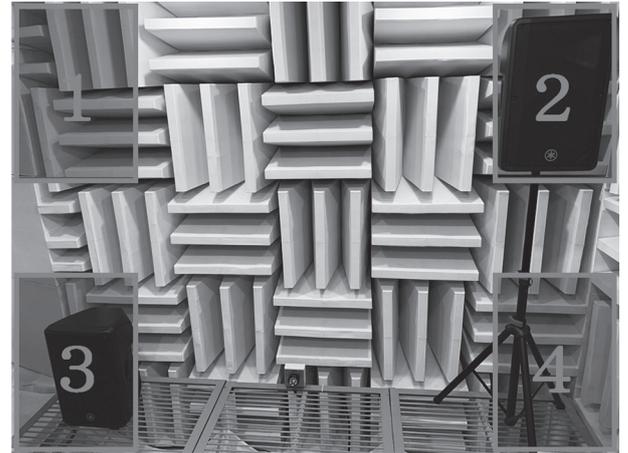


図7 スピーカー位置

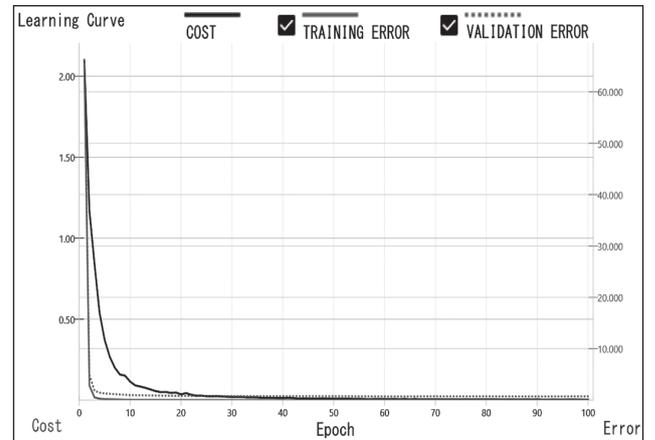


図8 学習曲線

| | y'_0 | y'_1 | y'_2 | y'_3 | Recall |
|----------------|--------|--------|--------|--------|--------|
| y=0 | 67 | 1 | 32 | 0 | 0.67 |
| y=1 | 4 | 55 | 5 | 36 | 0.55 |
| y=2 | 25 | 6 | 57 | 12 | 0.57 |
| y=3 | 0 | 27 | 4 | 69 | 0.69 |
| Precision | 0.6979 | 0.6179 | 0.5816 | 0.5897 | |
| F-Measures | 0.6836 | 0.5819 | 0.5757 | 0.6359 | |
| Accuracy | 0.62 | | | | |
| Avg.Precision | 0.6218 | | | | |
| Avg.Recall | 0.62 | | | | |
| Avg.F-Measures | 0.6193 | | | | |

図9 評価結果

4. 結言

音源可視化装置を提案した。本年度は、監視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、収集した音データからデータセットを作成する基本プログラム及び、6ch の WAV データを学習する基本的なニューラルネットワークを試作した。無響室における実験の結果、6ch の音データによる音源位置 4 分類が可能であることを確認した。

今後、音源可視化装置の試作と、音源探査の精度向上に向け、画像と併せて学習するプログラムの開発を行い、学習データの規模に応じたワークステーションの改良を実施する。また、対象に絞った音抽出に向けてプログラムの開発を実施する。

参考文献

- [1] 亀岡弘和：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.
- [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] 木下大：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] 木下大：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] <https://dl.sony.com/ja/>, Accessed 2021.