

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

○木下 大*

* 地方独立行政法人鳥取県産業技術センター

Experimental study for machines operation sound monitoring by the machine learning scheme and learning efficiency improvement by data augmentation

○Dai Kinoshita*

*Tottori Institute of Industrial Technology

Abstract: Operation state monitoring system using sound information is an important tool for manufacturing technology and maintenance, but it is difficult even for an expert to distinguish the target sound under complex environmental sounds, which is a problem in the development of automatic monitoring tools. The machine learning method can solve the above problem in terms of learning rules from a complex data group. However, for learning, it is necessary to collect and label sounds that can occur in the environment. In this study, we examine data pre-processing for acquiring decision rules using machine learning methods from multiple machine operation sound data. Furthermore, the effectiveness of learning efficiency when data expansion is performed by synthesis from the target machine sound and several environmental sound databases will be verified.

1. はじめに

近年の人材不足による自動化ニーズに合わせて、製造プロセスの自動化・監視技術が注目されている。中でも聴音による監視は非接触に不可視な現象を把握できる重要な手段であるが、従来の周波数パターンマッチングを用いた方法では、周囲の環境騒音に由来する状況認識率の低さが問題となる¹⁾。これに対し、機械学習手法では、収集したデータから機械的に特徴を学習し判別規則を探索・獲得するため、人が判断しきれない複雑で大量の判別規則を扱うことができ、特にその中の手法のひとつであるニューラルネットワーク(以下NN)は音声や画像をデータに用いた様々な判別問題に対して、高い解決能力があることが示されている^{2),3)}。

本研究では、一般的に入手しやすく比較的安価なリニア PMC レコーダーを用いて工場内を想定した環境音の収集を実施し、フーリエ変換で得られた周波数特徴量から種々の環境音のもつ特性を調査した。さらに、NN により工場内の環境音下で特定の機械の稼働状況を判断できる判別規則の獲得が可能であるかどうかを検討した。

2. 実験方法と結果

2.1 測定条件

音の測定には、リニアPCMレコーダー DR07-MKII (TASCAM)を使用した。サンプリング周波数は48kHz、量子化ビット幅は16bitで測定し、データはwav形式で保存した。Fig. 1 に今回測定した機械素材研究所機械工作室の音源の例を示す。作業音については、ワークをテーブル上の万力に隙間なく固定するため、ハンマーで叩く作業で代表した。

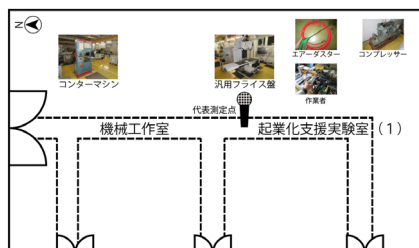


Fig.1 Measurement environment

2.2 計算処理方法

測定した音信号を人が判断しやすい情報量に変換するため、高速フーリエ変換 (以下FFT) 処理と短時間フーリエ変換 (以下STFT) 処理を行った。FFTは、窓関数をハミング窓とし、1秒間(48000サンプル)のデータを使って計算した (周波数分解能 1[Hz])。STFTは、FFTの計算設定からデータ幅を約 0.01 秒間(512 サンプル)、オーバーラップなしで計算した (周波数分解能 93.75[Hz], 時間分解能 約 0.01 秒)。Fig. 1 の代表測定点から、それぞれの対象の音を収録した結果をFig. 2 に示す。ここで、特徴を比較しやすくするため、測定信号は1秒間の最大振幅で除して規格化しており、STFTのカラーマップはlogスケールで大きさを表示している。計算処理はpython言語の数値計算ライブラリscipyを用いて計算した。

2.3 計算処理結果の比較

FFTの結果を比較すると、エアダスターの音以外は 5000[Hz]以下の周波数帯に音の成分が集中しており、物理的な回転駆動系をもつコンターマシンやフライス盤、コンプレッサーでは 50-2500Hzの範囲で顕著なピークが観測できる。対してエアダスターの音は顕著なピークを持たず広い周波数範囲に分布している

STFTの結果では、FFTからさらに時間変化の情報が付与されており、作業音のような不連続な音も発生周期や特徴が把握しやすい。また、作業音以外の時間変化の少ない音についても、音の高低で判断できない低周波数の音の大きさの変化などが観察でき、例えばコンターマシンでは、6900[Hz]周辺の音が約 0.2 秒間隔で発生していることがわかる。

2.4 稼働音の判別

前節で収集した音データより、コンターマシンと汎用フライス盤の起動・停止時音も含めたそれぞれの機械稼働音と、両機械とも稼働していない場合の工場内環境音を用い、それぞれの機械の稼働音を判別する規則の獲得を試みた。学習には、STFT処理データを画像として扱い、画像認識に汎用的に用いられる畳み込みニューラルネットワ

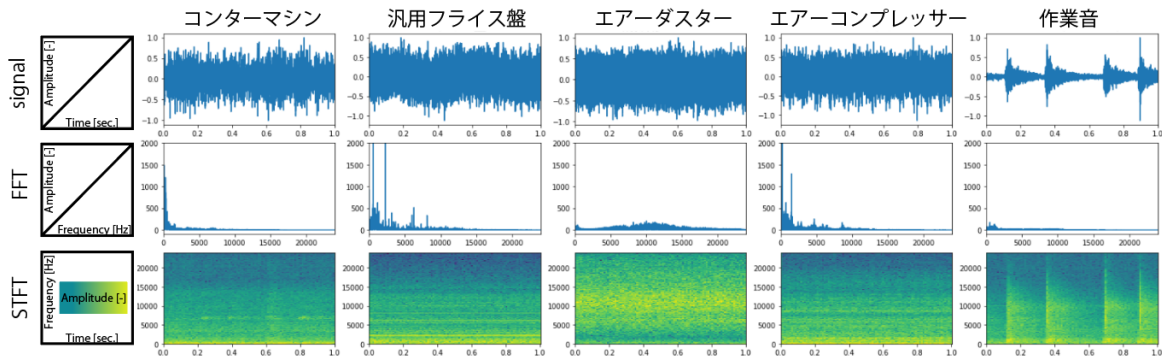


Fig.2 FFT/STFT calculation result of measured sounds

ーク（以下CNN）モデルに学習させる方法を使用した⁴⁾。教師データは、収集した音データと白色雑音とを相互デジタル合成し、Fig. 2のSTFT処理を実施して、現在時刻から前後128サンプルをまとめた257×257ピクセルの画像データを1つの教師データとし計25000枚の教師データを生成した。各画像には、コンターマシンと汎用フライス盤が稼働しているか否かのラベル付けがされている。学習計算はディープラーニング用ライブラリchainerを使用し、4層CNN+全結合層モデルを構築した。ハイパーパラメータと最適化手法は試行錯誤的にFig. 3のように設定した。学習結果の評価のため、新たに取得した汎用フライス盤とコンターマシンの稼働音が混在する音声データに、学習によって獲得した判別規則を適用した結果をFig. 4に示す。

モデル構成：CNN+全結合層

	conv1	conv2	pool1	conv3	conv4	pool2
function	relu	relu	max	relu	relu	max
ksize	5	5	2	3	3	2
pad	2	2	0	1	1	0
stride	1	1	2	1	1	2
no bias	True	True		True	True	

*conv x = chainer.links.Convolution2D
pool x = chainer.functions.max_pooling_2d

損失関数：sigmoid_cross_entropy

最適化手法:SGD(確率勾配降下) 学習率 lr=0.01

Fig.3 Details of CNN model

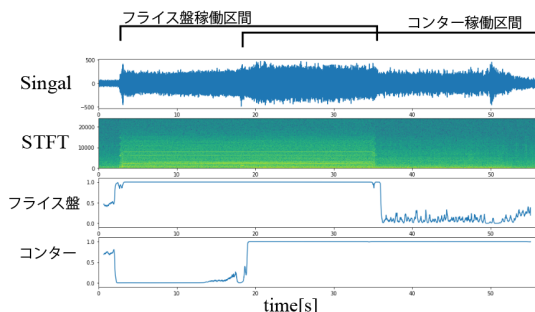


Fig.4 Identification results using new data

判別結果は稼働か否かの確率変数として扱い、実験では0.8以上で稼働と判断すれば汎用フライス盤とコンターマシンの稼働判別が可能であった。開始数

秒のコンターマシンで0.6程度とやや高く稼働確率が得られているが、これは機械学習用データ合成の結果、両機械停止時のデータが全体の5%程度と低い割合になっていたことに起因し、これを修正したデータを用いることで判別率の向上が望める。比較のため、同条件で合成した音データを使用しなかった場合の学習も実施したが、両方の機械が動いている範囲で誤検知が発生した。

3. おわりに

機械学習手法を用いて一般的な回転系を持つ駆動機械の動作音の判定が可能であることを確認した。同方法で、別機械の稼働音や直接の加工音をタグ付けすることで種々の機械の聞き分けが可能となる。また、想定される環境音のデータを蓄積することでさらなる精度向上が見込める。

今回利用した音のデジタル合成は機械学習のデータ生成で汎用的に用いられる手法であるが、新たな判別対象や環境音が追加された場合に組み合わせた音が不要となるため、データ収集量を削減する有効な手段である

参考文献

- 1) 山口智美、樋口誠宏、矢野章成、機械音スペクトルの出現頻度を用いた加工工程監視システムの開発（第1報、16進数化スペクトルパターン（HSP）表示による工具状態監視システムの構築）、日本機械学会論文集（C編）、68巻、668号、pp.1251-1256、(2002)
- 2) 三好由記博、適応フィルタとニューラルネットワークを用いた切削異常状態の検知、精密工学会誌、57(10)、pp.1774-1779、(1991)
- 3) 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、公益社団法人計測自動制御学会、58巻、3号、pp.195-202、(2019)
- 4) A. Ephrat, I. Mosseri, O. Lang, T. Dekel, K. Wilson, A. Hassidim, W. T. Freeman, M. Rubinstein, "Looking to Listen at the Cocktail Party: A Speaker-Independent Audio-Visual Model for Speech Separation", ACM Trans. Graph., Vol. 37, No. 4, Article 112. (2018)