

# 人の聴覚特性を考慮した特徴量設計による異音検知AIの確立

Establishment of Anomalous Sound Detection AI Based on Feature Design Considering Human Auditory Characteristics

佐藤 晃輔<sup>\*1</sup> 大嶽 舜<sup>\*1</sup>  
Kosuke Sato Shun Otake

\*1 ITデジタル推進部

## 1 はじめに

日本では少子高齢化が急速に進んでおり、高齢者人口が大幅に増加している。内閣府「令和4年高齢社会白書」によると15歳から64歳の人口を示す生産年齢人口は減少傾向にあり、労働力の不足・経済規模の縮小など社会的・経済的課題の深刻化が懸念されている<sup>[1]</sup>。将来的には約5人の熟練者の技能・技術を1人の若手が引き継がなければならない状況が訪れるとも言われている。当社も同様に平均年齢は上昇傾向にあり、熟練者の技術継承は喫緊の課題となっているという点に着目し、人の五感のデジタル化のように全社に適用可能な汎用分析技術の手の内化とビジネス活用拡大に取り組んでいる。現在多くの現場では、人の感性に依る定性的な異音検知が実施されており、人による判定のばらつきや感覚的な技術の継承が課題となっている。音を定量的に評価し分析技術確立することで全社の暗黙知解消・業務改善・製品の品質向上を目指す。今回はトヨタL&Fカンパニー 生産技術部と協力し、フォークリフトのモータ駆動などを制御するドライブユニットの組立工程を対象として異音検知に取り組んだ。



図1 ドライブユニットの取り付け位置  
Fig.1 Mounting Position of the Drive Unit

## 2 現状整理

現状の異音検査は、国内で異音検査を実施したのち合格品が海外へ輸送され再度検査を行う流れとなっている。国内の検査は、組立工程の作業の一環として対象の製品をダミーモータに取り付け、正転・逆転で動作させたときの動作音を人が聴覚による官能評価を行い可否を判定している。海外では実際のフォークリフトに取り付けたうえで検査担当者が実機を運転させながら異音聞こえるかどうか官能検査を行っている。

国内で異音を確認した製品は廃棄、もしくは不良箇所を組み替え再検査を行い合格したもののみを出荷している。一方、海外で異音を確認されるとドライブユニットは日本へ返送され正常品を再輸送する必要がある、輸送コストの増大が課題となる。日本へ返送されたドライブユニットを国内で再度検査すると異音を確認できることがあり、生産工程の作業途中かつ騒音下でも異音が発生するドライブユニットを発見し、海外へ輸送してしまうことを阻止することが求められている。

## 3 開発の狙い

生産工程で実施していた異音検査は各人の聴覚による定性的な判断に依存しており、定量的な評価方法が未確立である。異音を検知可能な技術確立することで、検査の省人化・異常発生原因の特定等、高度な評価・分析を実現できることを目指す。まずはドライブユニットの動作音を定量的に評価し、異音の発生するドライブユニットを検知できるシステムを開発することで、不良品流出の抑制・検査の高信頼化を目指す。

## 4 異音検知モデルの開発

本取り組みは、工場の騒音下での異音検知となるため、周囲の騒音の影響を受けにくく音の代替特性として利用可能な加速度情報を採用した。加速度センサから取得される振動データを音の情報として取り扱う。取得したデータに対し、正常音・異常音を判別可能か検証する。今回は振動データを取得する際、サンプリングレートを40,000Hz、音の取得秒数は6秒前後(人が検査を実施している時間と同等)という条件を設定している。

### 4.1 波形情報の確認

図2に正常な製品と異常な製品の音の波形を示す。検査初心者が聞いても判別が容易な異音は波形上でも顕著な特徴が出ている。一方、熟練者が聞かないと判別できないような異音については、図3のように波形の形状からは判別しにくいことが確認できた。波形形状からは正常・異常を分離可能な適切なしきい値を設定することができない一方で、熟練検査員は聴覚での官能検査により異音を感じ取ることができるという点に着目し、波形データを『人の聴覚特性を考慮した特徴量』に変換することで、機械学習手法を導入し異音を判別可能なモデルを作成できるか検証する。

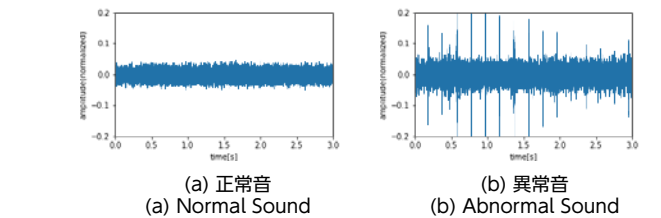


図2 官能評価で判別可能な波形データ  
Fig.2 Waveform Data Easily Distinguishable by Sensory Evaluation

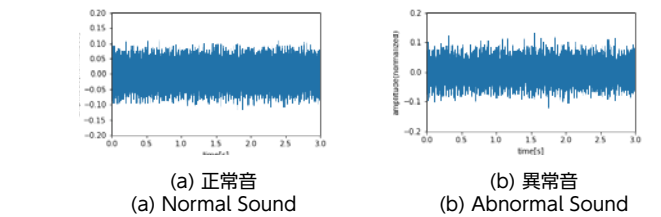


図3 官能評価で判別困難な波形データ  
Fig.3 Waveform Data Difficult to Distinguish by Sensory Evaluation

### 4.2 人の聴覚特性を考慮した特徴量

今回の異音検知では人の聴覚特性を表現可能な特徴とした18個の特徴量を設計した<sup>[2]</sup>(表1)。今回は代表的な特徴量について概要を紹介する。

表1 今回利用した特徴量  
Table1 Features Used in This Analysis

|                |                          |
|----------------|--------------------------|
| スペクトル重心        | Chroma Energy Normalized |
| スペクトル帯域幅       | 調波音                      |
| スペクトル平坦度       | 打楽器音                     |
| スペクトルコントラスト    | poly features            |
| tonnets        | mfcc                     |
| ロールオフ周波数       | データ長                     |
| メルスペクトログラム     | テンポ                      |
| クロマグラム(STFT変換) | RMS                      |
| クロマグラム(QCT変換)  | ゼロ交差率                    |

#### 4.2.1 mfcc(メル周波数ケプストラム係数)

音声信号の特徴量抽出に広く用いられ、人間の耳は低周波数域を細かく聞き分けることができる一方、高周波数域では相対的に雑音に強いものの、細かい周波数の変化は捉えにくい。この聴覚特性を定量的に表現可能な手法である。

#### 4.2.2 メルスペクトログラム

振動データの周波数成分が時間推移とともにどのように変化するか表現可能な特徴量で、縦軸を周波数、横軸を時間、各成分の強度を数値で表現することで2次元空間上で周波数の変動を表現可能な手法である(図4(a))。

#### 4.2.3 クロマグラム

振動データの音階成分(ド～シ)が時間推移とともにどのように変化するか表現可能な特徴量で、横軸を時間、各成分の強度を色で表現することで2次元空間で音階の変動を表現可能な手法である(図4(b))。

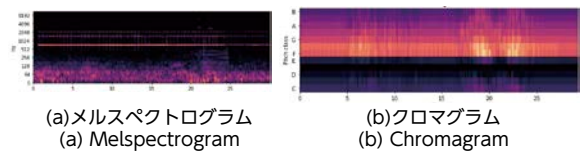


図4 音の変動の時間推移の見える化  
Fig.4 Visualization of Temporal Variations in Sound

#### 4.2.4 調波音打楽器音分離

振動データに対して、定常的に流れる音(調波音)と突発的な変動音(打楽器音)に分離可能な手法で、調波音・打楽器音それぞれに対して解析を実施できるため、各成分の傾向を評価できる(図5)。

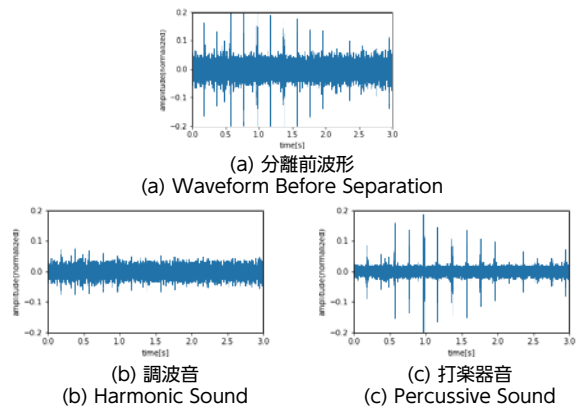


図5 調波打楽器音分離  
Fig.5 Harmonic and Percussive Sound Separation

### 4.3 機械学習モデル作成

4.2で紹介した全18個の特徴量に対して、それぞれ8個の統計量(平均・分散・最大・最小・標準偏差・中央値・尖度・歪度)を計算することで、1394個の特徴量を設計した。異音の発生する製品を十分に用意することができないため、今回は正常な動作音数千件に対して、異常が発生している音を数十件用意し、教師あり学習のモデルを生成した。異常音を真、正常音を偽としてラベル付けを行い、評価指標として全データの予測がどの程度正解したかの指標としてAccuracy(正解率)、全異常音のうちどの程度異常を検知できたかの指標としてRecall(再現率)、異常と予測したデータのうちどれだけ正解していたかの指標としてPrecision(適合率)を用いた。今回は正常音と比べて異常音の数が極端に少ないため、Accuracyのみを評価基準とすると、全てのデータを正常音と評価したものが高精度なモデルと判断され、本来確認したい異常音の検出精度を十分に評価することができない。一方で、異常音の検出精度としてRecallのみを評価基準とすると、全てのデータを異常音と評価したのも高精度なモデルと判断され、実際の現場には適用することができないものが高精度なモデルと判断されてしまう。そこで今回は全体的な予測精度を評価するAccuracyと異常音の検知率を評価するRecallのバランスが良いモデルを高精度なモデルと仮定して手法の比較を実施した。

今回設計した特徴量は数が多いため、主要な特徴を抽出することが可能な前処理を実施する手法と、前処理を実施せず設計した特徴量を直接利用する手法の2つを比較検証している。1つ目の手法では、多次元データの重要なパターンを見つけるための主成分分析(PCA)を適用する。PCAは観測

されたデータの線形変換を行い、新たな変数(主成分)を形成する。これらの主成分は、元のデータの分散を最大化するように選ばれるため、相互に直交しデータの本質的な特性を鮮明に浮かび上がらせることができる。2つ目の手法では設計した特徴量に対して正規化のみを行い評価を実施する。『主要な特徴抽出の前処理を実施した特徴量』と『主要な特徴抽出の前処理を未実施の特徴量』の各1394個の特徴量を説明変数として、モデルを設計し評価している。12個の機械学習モデルに対して、k分割交差検証を行い、Recallが最適化されるようにハイパーパラメータのチューニングを実施した(表2、表3)。

表2 前処理を実施した特徴量が説明変数のモデル  
Table2 Model with Preprocessed Features as Explanatory Variables

| モデル名                            | 前処理あり    |        |           |
|---------------------------------|----------|--------|-----------|
|                                 | Accuracy | Recall | Precision |
| Naive Bayes                     | 94.31%   | 41.7%  | 15.4%     |
| Logistic Regression             | 98.51%   | 37.5%  | 75.0%     |
| SVM - Linear Kernel             | 60.89%   | 100.0% | 4.8%      |
| Linear Discriminant Analysis    | 98.68%   | 62.5%  | 68.2%     |
| Ridge Classifier                | 30.78%   | 100.0% | 2.8%      |
| K Neighbors Classifier          | 98.10%   | 8.3%   | 66.7%     |
| Gradient Boosting Classifier    | 3.05%    | 79.2%  | 1.6%      |
| Ada Boost Classifier            | 98.18%   | 12.5%  | 75.0%     |
| Decision Tree Classifier        | 97.85%   | 8.3%   | 33.3%     |
| Light Gradient Boosting Machine | 98.18%   | 12.5%  | 75.0%     |
| Random Forest Classifier        | 95.05%   | 79.2%  | 25.7%     |
| Quadratic Discriminant Analysis | 2.06%    | 100.0% | 2.0%      |

表3 前処理を未実施の特徴量が説明変数のモデル  
Table3 Model with Unprocessed Features as Explanatory Variables

| モデル名                            | 前処理なし    |        |           |
|---------------------------------|----------|--------|-----------|
|                                 | Accuracy | Recall | Precision |
| Naive Bayes                     | 9.98%    | 100.0% | 2.2%      |
| Logistic Regression             | 98.27%   | 25.0%  | 66.7%     |
| SVM - Linear Kernel             | 59.24%   | 95.8%  | 4.5%      |
| Linear Discriminant Analysis    | 99.01%   | 66.7%  | 80.0%     |
| Ridge Classifier                | 32.18%   | 100.0% | 2.8%      |
| K Neighbors Classifier          | 98.27%   | 16.7%  | 80.0%     |
| Gradient Boosting Classifier    | 97.03%   | 12.5%  | 16.7%     |
| Ada Boost Classifier            | 98.18%   | 16.7%  | 66.7%     |
| Decision Tree Classifier        | 98.43%   | 33.3%  | 72.7%     |
| Light Gradient Boosting Machine | 98.02%   | 20.8%  | 50.0%     |
| Random Forest Classifier        | 88.04%   | 83.3%  | 12.4%     |
| Quadratic Discriminant Analysis | 1.98%    | 100.0% | 2.0%      |

計24個のモデルを評価し、最もAccuracyとRecallのバランスが良い『主要な特徴抽出の前処理を未実施の特徴量』とLDA(Linear Discriminant Analysis:線形判別分析)の組み合わせを採用した。一般的に機械学習手法を用いた解析を行う場合には、主要な特徴の抽出を実施したほうが高精度なモデルを作成できる。一方、今回はPCAを実施することでドライブユニットの動作音の基本的な特性が主に抽出されてしまい、異常音を判断できる特徴を十分抽出できなかった

め、前処理を未実施の特徴量のほうが高精度なモデルを作成できている。LDAは、統計や機械学習の分野で使用される手法で、データのクラス間の分離を最適化するためのものとなっている。学習時には、クラスごとの平均と分散情報を利用して、データを低次元空間に変換し、クラス間の分離を最大化することができるようモデルを構築する。評価時には、評価したいデータを低次元空間に写像しどちらのクラスに属しているかを確認することで、正常音・異常音の判定を行うことができる。

## 5 異音検知システムの実装

### 5.1 システム構成

4章で作成したモデルを用いて、現状の組立工程内でシステムによる異音検査を実施できるように開発を実施した(図6)。

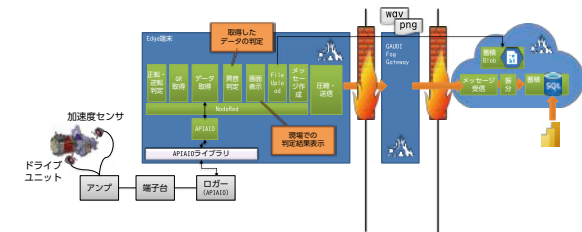


図6 システム全体図  
Fig.6 Overall System Diagram

#### 5.1.1 データ判定方法

組立工程内でドライブユニットに加速度センサーを取り付ける。ドライブユニットを動作させ、その振動データを加速度センサーからアンプを経由してEdge端末にて受信する。振動データは、特徴量に変換され4章で作成したモデルにて異音判定を実施し、現場に設置したモニタに表示する。モニタには、今回取得した各稼働音の波形、それぞれの稼働音に対する正常・異常の判定結果、各判定結果を踏まえた検査の総合判定結果が表示される(図7)。現場では作業員がモニタを視認し、製品に問題があるかどうかを判断する。本システムにおいて異常と判断された製品に関しては、作業工程の管理者の官能検査による再検査が行われ、異音と判断されたものは出荷工程から除外される。

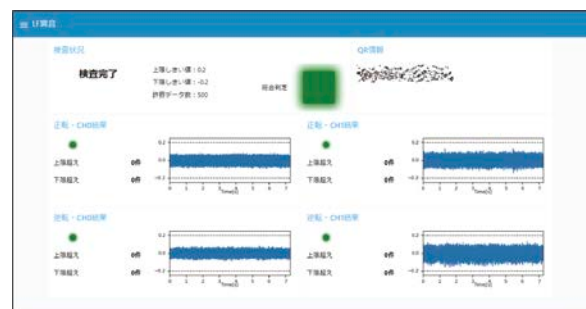


図7 判定結果表示画面  
Fig.7 Result Display Screen

#### 5.1.2 データ保管方法

判定に利用した波形データと判定結果はクラウド上に保管する構成とした。結果の振り返りと確認を行うことができるため、モデルの再学習に活用することができるようになっている。

#### 5.2 効果確認

モデルの汎用性評価のために、モデル作成後に収集した振動データを使って、異音検知モデルの異音警告率を評価した。1か月当たりの平均異常警告数は20件前後となっており、作業工程の管理者による出荷停止判断が行われた製品は数個程度となっている。海外出荷後の日本への返送は1年間0件を継続しており、異音警告のしくみとしては現場で十分に運用できるものと判断・活用されている。

## 6 まとめ

今回の事例の中で、音を収集し、異音検知に効果的な特徴量へ変換してモデルを作成し評価するまでの一連の工程の実装を達成した。人の聴覚特性に適合した特徴量設計手法を確立し、その特徴量を用いた異音検知モデルの作成に成功した。1年間後工程への異常品出荷0件を達成しており、今後は類似工程への横展開を目指す。

### 謝辞

本開発はトヨタL&Fカンパニー 生産技術部の協力のもと実施することができました。厚く御礼申し上げます。

### 参考文献

- [1] 令和4年版高齢社会白書,  
[https://www8.cao.go.jp/kourei/whitepaper/w-2022/zenbun/04pdf\\_index.html](https://www8.cao.go.jp/kourei/whitepaper/w-2022/zenbun/04pdf_index.html)
- [2] Work w/ Audio Data: Visualise, Classify, Recommend,“<https://www.kaggle.com/code/andradaolteanu/work-w-audio-data-visualise-classify-recommend/notebook>”

### 著者紹介



佐藤 晃輔

大嶽 舜