

Title	音質評価指標を用いた異常音検知の研究
Author(s)	水野, 滉介
Citation	
Issue Date	2020-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/16422
Rights	
Description	Supervisor: 鷗木 祐史, 先端科学技術研究科, 修士 (情報科学)

修士論文

音質評価指標を用いた異常音検知の研究

1810176 水野 滉介

主指導教員 鵜木 祐史
審査委員主査 鵜木 祐史
審査委員 赤木 正人
党 建武
吉高 淳夫

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和2年2月

Abstract

Machines require high reliability operating in factories, power stations, and transportation requires high reliability. If a machine or system malfunctions, it can lead to social disruption and economic loss. Furthermore, there is a risk of causing disasters and accidents due to aging of products and systems. In order to prevent these risks, daily maintenance and inspection work on the machines that make up to the mark the system is indispensable.

Skilled maintenance technicians can determine whether the machine is normal or abnormal based on the slight difference in the operation noise emitted from the machine. However, there is an individual difference in hearing, and even a skilled technician may make an incorrect judgment or disagree. In addition, no systematic training method has been established to bring engineers up to skill level. For this reason, training engineers are expected to take a lot of time and cost. There is a concern that maintenance know-how will be lost due to a decrease in the working population and a shortage of successors. From the background of these labors, practical application of abnormal sound detection by machines without human intervention is being studied. In recent years, with the improvement of computer performance, the study of abnormal sound detection by machine learning has been studied. However, their performance is inferior to humans, so there is still room for consideration.

It is considered that skilled engineers detect abnormal sounds based on timbre information, especially auditory impressions such as roughness, fluctuation, and sharpness. Among the timbres, there are sound quality metrics (roughness, fluctuation strength, sharpness) as objective evaluation indices related to the roughness, fluctuation, and metal factors of the sound. It is thought that abnormal sounds can be detected based on this sound quality metrics. The purpose of this research is to examine what features appear in auditory impressions such as roughness, fluctuation, and instep height of machine operation sounds, and to clarify whether abnormal sounds can be detected based on the characteristics. First, in order to clarify the auditory characteristics of normal or abnormal machine operation sound, the sound quality metrics of the database of normal sound and abnormal sound was analyzed. A toy motion sound database (ToyADMOS) was used in this analysis. The toy to be analyzed was a railway model (HO scale). This database contains the sounds of broken wheels, broken rails, foreign objects on the rails, and loosened rail connections. In this analysis, the operation sound that matches the conditions when running on a damaged rail, running on a rail on which foreign matter is placed, or running on a rail with loose connections is used with broken wheels. There was a total of 14 operation sounds that matched this condition. The same number of normal sounds as the abnormal sounds were also used. The

sound quality metrics were used to analyze the toy train sounds. These metrics were derived by using OROS FFT analyzer (OR-34). The high-order statistics (mean, variance, skewness, and kurtosis) of the sound quality metrics were used to investigate whether differences between the characteristics of normal sound and abnormal sound can be appeared in the results of the sound quality metrics.

As a result of analysis, the average and variance of roughness and the variance of fluctuation strength for abnormal sounds were higher than those for normal sounds. In contrast, no difference between the normal and abnormal average sharpness could be confirmed. From these results, it was found that abnormal sounds have characteristics in the index related to sound fluctuation.

Next, the classification of normal sounds and abnormal sounds is performed using the statistics of roughness and fluctuation strength as thresholds, and the classification accuracy was evaluated. In addition to ToyADMOS, the classification accuracy was evaluated for a database of machine operation sounds (motor sounds) determined by a skilled engineer. A new database of machine operation sounds (motor sounds) determined by a skilled technician was prepared, and the classification accuracy was also evaluated for this database. There was a total of 37 motor sounds, of which 30 were determined to be normal and 7 were determined to be abnormal.

In the ToyADMOS, the normal sound and the abnormal sound were classified using the average and variance of the roughness and the variance of the fluctuation strength as the threshold values. This result is equivalent to the result of classifying the same database by machine learning, and high accuracy was obtained. On the other hand, the classification results for the motor sound showed that the F-measures of mean, variance of the roughness, the F-measures of variance of the fluctuation strength were 0.600, 0.632, and 0.667, respectively, which was lower accuracy than ToyADMOS.

From these results, it was suggested that abnormal sound detection can be realized by using the roughness and fluctuation strength, which are the sound quality evaluation indices, for the machine operation sound that is easy to discriminate.

目 次

第 1 章	序論	1
1.1	はじめに	1
1.2	研究背景	2
1.3	研究の目的	3
1.4	本論文の構成	3
第 2 章	本研究の方略	5
2.1	異常音検知に関する研究	5
2.1.1	決定論的アプローチ	5
2.1.2	統計的アプローチ	5
2.1.3	問題点	6
2.2	本研究の着目点	6
2.3	音質評価指標	7
2.4	研究の方法論	7
第 3 章	正常 / 異常音についての聴覚的特徴の分析	9
3.1	目的	9
3.2	分析に利用したデータベース	9
3.3	分析方法	17
3.4	音質評価指標における高次統計量の結果	17
3.5	考察	19
第 4 章	分類精度の評価	20
4.1	目的	20
4.2	機械動作音のデータベース	20
4.3	分類方法と精度の評価	20
4.4	ToyADMOS に対する検知精度	20
4.5	機械動作音に対する検知精度	26
4.6	考察	31
4.6.1	ToyADMOS	31
4.6.2	機械動作音	31
第 5 章	全体考察	33

第 6 章	結論	34
6.1	明らかにしたこと	34
6.2	残された課題	34
6.2.1	臨海帯域ごとの音質評価指標の分析	34
6.2.2	異常音検知の聴知覚メカニズムの解明	35
参考文献		35
研究業績		38
謝辞		40

目 次

1.1	本論文の構成	4
3.1	分析対象の鉄道模型 (画像上部の車両が HO スケール) [21]	10
3.2	正常状態の車輪 [21]	11
3.3	異常状態の車輪 [21]	12
3.4	異常状態の線路 (ラジオペンチによる線路破損) [21]	13
3.5	異常状態の線路 (異物の配置) [21]	14
3.6	異常状態の線路 (連結部分を緩めた場合) [21]	15
3.7	模型電車における音質評価指標の高次統計量の結果 (赤:異常音, 青:正常音, 行は統計量, 列は音質評価指標に対応)	18
4.1	ラフネスの平均を閾値とした場合の ROC 曲線	22
4.2	ラフネスの分散を閾値とした場合の ROC 曲線	23
4.3	変動強度の分散を閾値とした場合の ROC 曲線	24
4.4	ラフネスの平均を閾値とした場合の ROC 曲線	27
4.5	ラフネスの分散を閾値とした場合の ROC 曲線	28
4.6	変動強度の分散を閾値とした場合の ROC 曲線	29
4.7	モーター音における音質評価指標の高次統計量の結果 (赤:異常音, 青:正常音, 行は統計量, 列は音質評価指標に対応)	32

表 目 次

3.1	分析に用いた音源と異常状態	16
4.1	ROC 曲線から得られた閾値をテスト用データに適用した場合の検知精度	25
4.2	ROC 曲線から得られた閾値をテスト用データに適用した場合の検知精度	30

第1章 序論

1.1 はじめに

18世紀から19世紀にかけて、イギリスでは人口増加、年収の上昇、都市化を背景に年間経済成長率が0.5%から3.5%と急激な増加が見られた [1]。これに伴い工業の形態が、生産者が仕入れ、生産、販売を一貫して行い、人の手により作業が行われる家内制手工業から、仕入れ、生産、販売を分業する工場制手工業へと変化した。さらに同時期に起こった産業革命により、従来の手作業から機械による自動化がなされ、機械工業が発達した [1]。機械工業の発達により、大量の原料・製品などを大量、安価かつ迅速に輸送するための交通輸送機関も発達し、水運から鉄道に置き換わるなどインフラストラクチャにおいても機械化がなされた。その後、化学、電気、石油、鉄鋼の分野で技術革新が進んだ第二次産業革命、コンピュータによる生産の自動化、原子力エネルギーの技術革新が進んだ第三次産業革命など、時代とともに工業の高度化が進んだ。今日では、自動車や家電製品など、一般家庭でも機械を扱う機会が増加し、機械は日常生活においても必要不可欠の存在となっている。製品の生産から電力、水道、交通機関といったインフラストラクチャの整備、維持に至るまで、機械の存在なしに人々の生活は成り立たないといえるだろう。

機械を動作させる際に、必ず動作音が生じる。この動作音は場合によって騒音となり、人々に不快感を与えてしまうが、機械の調子や状態を判断する手がかりとして、機械動作音は重要視されている。例えば、手動変速車で発進の際、エンジンの回転数に伴う動作音を手がかりにクラッチペダルを操作する。また、機械の故障診断の際には、通常の稼働音とは異なる音、すなわち異音の種類を手がかりにしている。例えば、古いフィルムカメラでは、シャッターを切った際に異音が生じる所謂「シャッター鳴き」と称される症状がある。この症状の原因は、シャッターの機構にあるギアが経年劣化により固まってしまい、スムーズに回転しなくなるためである。このように、機械動作音は騒音といった負の側面も有するものの、機械の動作状況や調子を知る上で欠かせない存在である。

1.2 研究背景

工場などで稼働している機械，システムが機能不全を起こすこと，今まで安定して供給できた製品，インフラストラクチャなどが正常に働かなくなり，社会的混乱や経済的損失に直結する．また，機械やシステムの経年劣化によって，災害や事故の発生を招く恐れがある．このため，工場や発電所，交通機関で稼働する機械には高い信頼性が求められており，それを実現，維持するために機械に対する保守・点検業務が行われている．

ヒトが保守・点検業務を行う際に，ヒトの五感を基本として行われている．視覚を用いて機械の様子を観察する方法，聴覚を用いて機械の動作音を聞く方法，触覚を用いて機械に直接触れる方法，嗅覚を用いて臭いから機械を診断する方法が挙げられる．保守・点検業務において最も用いられている感覚は視覚であり，全発見数の41.6%を占めている [2]．次点で用いられている感覚は聴覚であり，全発見数の17.7%である．聴覚は，視覚では賄えない場面，例えばブラックボックスとして機械の構造を目視できない場合や，人が立ち入ることが困難な機械の点検などに対して用いられる．また，触覚や嗅覚はそれぞれ7.6%，0%であり，保守・点検業務では視覚と聴覚が重要視されていることがわかる．

保守・点検の熟練技術者は，機械の動作音の僅かな違いを聞き取ることができ，音から機械の正常／異常の判別ができる [2]．しかし，聴覚を始めとした感覚には個人差があり，熟練技術者でも判別を誤ったり，見解が分かれる場合が考えられる．また，熟練レベルまでに技術者を育てるための体系的なトレーニング方法が確立されていないため [3]，技術者育成には多くの時間やコストがかかることが予想される．

加えて，少子高齢化や産業構造の変化，雇用形態の変化などに伴う労働者不足が問題となっている [4]．保守・点検において高度な技術を保有する労働者が，定年退職などによって人口が減少している．産業別労働分布の変化により，製造業などの二次産業から，サービス業などの三次産業へ労働人口が移り，後継者不足が指摘される [4]．労働人口の減少，後継者不足により，保守点検のノウハウの喪失が懸念される．この他，保守・点検作業の自動化，無休稼働，遠隔操作といった要求もある．

これら人間の能力や労働環境などを背景として，人を介さずに，機械による異常音検知の実用化が検討されている．近年では，計算機の性能が向上したことにより，機械学習による異常音検知の検討が行われている [5]．しかし，正常／異常を判別するモデルの作成が容易でないこと，またそのモデルを作成するためのデータセットが数少ないことから，機械による異常音検知の性能は人のそれに遠く及ばない状況にある．機械による異常音検知を実現するためには，人の異常音検知のメカニズムを知り，それを模倣する必要がある．

1.3 研究の目的

熟練技術者は音色の情報，特に音の粗さや甲高さといった聴感印象を手がかりに異常音を検知しているものと考えられる．実際に聴能形成訓練でも，音の大きさや高さだけでなく，音色の弁別訓練も重要視されている [3]．音色の中でも音の粗さ感，変動感，金属因子に関係する客観評価指標として，音質評価指標（ラフネス，変動強度，シャープネス）があり [6]，音質評価指標を手がかりに異常音を検知できるのではないかと考えられる．本研究の目的は，機械動作音の粗さ感，変動感，甲高さといった聴感印象にどのような特徴が現れるかを検討し，それらの特徴を手がかりに異常音検知が可能であることを明らかにすることである．

1.4 本論文の構成

本論文は，6章で構成される．図 1.1 に本論部の構成を示す．

第 1 章は序論であり，本論文で対象とする異常音検知に関する背景を述べる．これにより研究の目的を明らかにする．

第 2 章は本研究の方略について述べる．まず，異常音検知に関する先行研究を紹介し，その手法と問題点を述べる．次に，本研究での着想点を述べ，新規性である音質評価指標についての説明をする．最後に本研究の方法論を述べる．

第 3 章では，正常 / 異常と判別された機械動作音についての聴覚的特徴の分析について述べる．分析では，おもちゃの動作音データベースに対し音質評価指標を分析し，その高次統計量を求める．その結果より，正常音 / 異常音にどのような特徴が現れるか検討する．

第 4 章では，第 3 章で得られた特徴を用いて，正常音 / 異常音の分類精度について評価する．また，おもちゃの動作音のデータベースに加え，実際の機械動作音のデータベースについても分類精度の評価を行う．

第 5 章では，正常音 / 異常音の聴覚的特徴と聴覚的特徴を用いた分類精度について全体的な考察を述べる．

第 6 章は結論であり，本研究で明らかになったこと，残された課題について述べる．

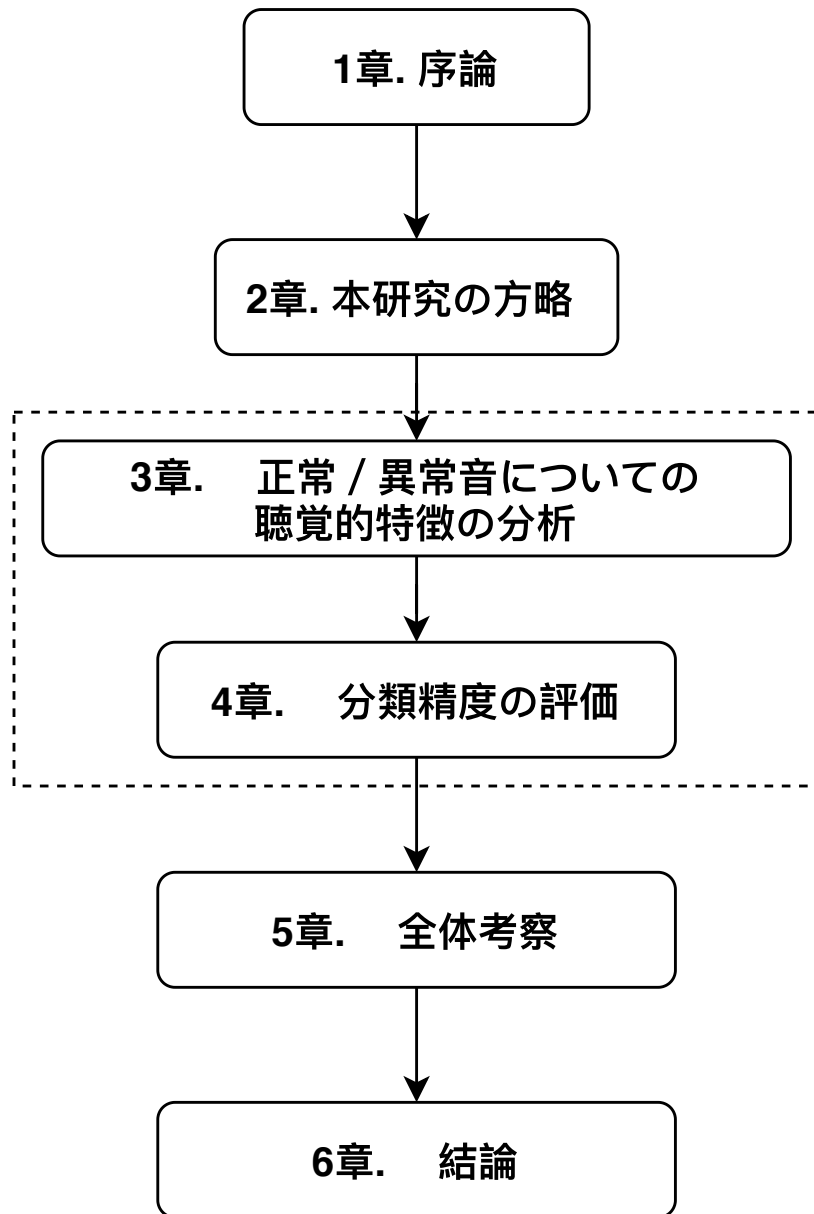


図 1.1: 本論文の構成

第2章 本研究の方略

2.1 異常音検知に関する研究

2.1.1 決定論的アプローチ

機械の音から異常を判別する異常音検知に関して、これまで大きく分けて二つのアプローチで行われてきた。一つは決定論的アプローチであり、正常時と異常時の音響特徴の差異を利用した判別法を提案するものである。鈴木は、火力発電プラントにおける異常検知方法として、個々の異常に対する音響特徴を、正常／異常とで比較することで異常検知を行っている [7]。具体的には、蒸気リーク音はある特定の周波数帯域に特徴が現れることから、特定の周波数のレベルを監視することで異常を検知するものである。また、ハンマリング音のように、部品などが接触した際に発せられる音に対しては、その音圧レベルや周期性から正常／異常の判定を行っている。安面らは、連続ウェーブレット変換および離散ウェーブレット変換を用いて、車載用エアコンにおける摩擦異音の検査を行う報告をしている [8]。この報告では、異常音が発する時間と、その時間帯における周波数特性を正常／異常とで比較し異常音検知を行っている。この他、機械動作音の振幅の二乗平均に対して閾値を設定し、異常検知する方法もある [9]

2.1.2 統計的アプローチ

もう一つは、統計的アプローチによる方法であり、機械学習を用いて異常検知を行っている [10–13]。統計的アプローチによる方法での異常音検知は、モデル学習の段階とそのモデルを用いた検出の段階に分けることができる [14,15]。モデル学習では、予め正常／異常と判別された機械音データベースを用意し、それを学習に用いる。検出の段階では、検知対象の音響特徴を抽出し、学習データを用いて正常／異常の検知を行う。機械学習における学習データの最適化の手法として、オートエンコーダを用いることで次元削減を行っている [16,17]。使用する音響特徴や分類手法は様々挙げられる。Clavel らは、パワーやスペクトル重心、MFCC(Mel-Frequency Cepstrum Coefficient) を音響特徴量とし、GMM(Gaussian Mixture Model) を用いて識別している。[18]。Ito らは、パワーや MFCC, Δ MFCC, Δ パワー を特徴量として、GMM によって異常音検知を行っていた [19]。Suzuki らは、特徴量として、

動作音の波形，スペクトルフラックス，MFCC を用いており，階層的クラスタリングによって異常音検知を行っている [20]．

また Koizumi らは，おもちゃの動作音の異常検知に関する報告を行っている [21]．正常に動作しているおもちゃの動作音 1000 個を用いて正常音を学習しモデルを作成し，このモデルを用いて正常／異常の判別を行っている．モデルの生成にオートエンコーダを用いて次元削減を行い [22]，正常／異常の分類に用いる特徴量として，動作音の 64 次元の対数メル振幅スペクトルを用いている．

2.1.3 問題点

二つのアプローチにおいて，未学習・未経験の異常音に対して対応可能かの問題が挙げられる．例えば，鈴木 [7] が報告した異常音検知システムにおいて，蒸気リーク音の特徴が現れる周波数帯域を 6～12 kHz と予め決定してる．機械学習による異常音検知においても，対象とする機械動作音のモデルを作成している．しかし，稼働する機械の変更や，別な工場への移転などが想定されるなど，新しい環境に対応するために，監視すべき音響特徴を見直したり，機械動作音のモデルを再度作成し直す必要が考えられる．分類精度の向上や様々な環境への対応を実現するために，大量のデータを用いる方法もあるが，機械の異常が滅多に発生しないなどの要因でデータの収集が困難である．このような背景もあり，異常音検知用のデータセットが少ないという問題もある [21]．

また，いずれのアプローチとも，異常音検知のための手がかりとして，安易に典型的な特徴，例えば MFCC といった特徴やスペクトル，基本周波数といった音響特徴を利用するだけであり，どのような特徴が適切であるかなど深く検討されていないことが考えられる．

2.2 本研究の着目点

異常音検知に関しては，これまでにパワーやスペクトルといった音響特徴を用いて検討されてきた．しかし，人がどのようにして異常音検知を行っているかを明らかにするためには，聴知覚の観点からの検討も必要である．

聴能形成訓練において，音色の弁別訓練が行われており，保守・点検技術者は音色を手がかりに正常／異常判別を行っているものと考えられる．本研究では，機械動作音の音色に着目し，音色知覚に係る音質評価指標を手がかりに異常音検知を行えるか検討する．

2.3 音質評価指標

音質評価指標とは，人間の音に対する感覚（心理量）と音を規定する物理量を結びつけた指標である [6]．音質評価指標は，音の質的な評価に用いられており，自動車の車内音，車外音の対策として用いられている [27, 28]．また，自動車の他に電車における車内騒音対策 [29] や，ミシン [30]，掃除機 [31]，といった家電製品の付加価値向上のために，音質評価指標が用いられている．

代表的な音質評価指標として，音の大きさを表すラウドネス，音の鋭さ，甲高さを表すシャープネス，粗さを表すラフネス，音の変動感を表す変動強度がある．

音の鋭さ，甲高さを表すシャープネスは，音圧レベルが 60 dB で，1 kHz を中心周波数とする臨海帯域幅の狭帯域雑音のときに 1 acum と定義されており，式 (2.1) のように示される [6, 23]．

$$S = 0.11 \frac{\int_0^{24 \text{ Bark}} N'(z) g'(z) z dz}{\int_0^{24 \text{ Bark}} N'(z) dz} \quad (2.1)$$

ここで， $N'(z)$ は臨界帯域ごとのラウドネス， z は臨界帯域番号， c は定数， $g(z)$ はシャープネスの重みを表す．

音の粗さを表すラフネスは，音圧レベルが 60 dB，搬送波周波数が 1 kHz，変調周波数が 70 Hz，変調度が 1 のときに 1 asper と定義されており，式 (2.2) のように示される [6, 23]．

$$R = 0.3 f_{\text{mod}} \int_0^{24 \text{ Bark}} \Delta L(z) dz \quad (2.2)$$

ここで， f_{mod} は変調周波数， $\Delta L(z)$ は臨界帯域ごとの興奮レベルの山と谷の差， z は臨界帯域番号である．

音の変動感を表す変動強度は，音圧レベルが 60 dB，搬送波周波数が 1 kHz，変調周波数が 4 Hz，変調度が 1 のときに 1 vacil と定義されており，式 (2.3) のように表される [6, 23, 24]．

$$F = \frac{0.008 \int_0^{24 \text{ Bark}} \Delta L(z) dz}{\frac{f_{\text{mod}}}{4} + \frac{4}{f_{\text{mod}}}} \quad (2.3)$$

ここで， f_{mod} は変調周波数， $\Delta L(z)$ は臨界帯域ごとの興奮レベルの山と谷の差， z は臨界帯域番号である．

2.4 研究の方法論

第 3 章において，正常音 / 異常音データベースに対し，それらの音質評価指標を分析する．分析した音質評価指標の高次統計量（平均，分散，歪度，尖度）より，異常音検知に有効な特徴を検討する．

第4章では、第3章において異常音検知に有効であると判断した特徴を用いて異常音検知を行い、その分類精度を検討する。分類精度については、先行研究で用いられている音響特徴を用いた場合との比較を行い、音質評価指標による異常音検知法の評価をする。

第3章 正常 / 異常音についての聴覚的特徴の分析

3.1 目的

本章では，正常または異常な機械動作音がどのような聴覚的特徴を有するかを明らかにするため，正常音・異常音のデータベースを利用して，対象音の音質評価指標による分析を行う．

3.2 分析に利用したデータベース

分析には，おもちゃの正常 / 異常動作音データベース (ToyADMOS) を利用した [21, 32]．ToyADMOS には，おもちゃの自動車，鉄道模型，ベルトコンベアの動作音が収録されており，本章では図 3.1 に示す鉄道模型 (HO スケール，新幹線車両) とした．

異常の種類には，先頭車両と後方車両の車輪をペンチで破損させたもの (図 3.2，図 3.3)，直線，曲線レールの破損させたもの (図 3.4)，レールへ異物を配置したもの (図 3.5)，レール連結部分を解除したもの (図 3.6) がある．表 3.1 に，分析に利用した稼働音と，その稼働音における異常状態を示す．分析では，車輪を破損させた場合，車輪とレールを破損させた場合，車輪の破損とレールへの異物を配置した場合，車輪の破損とレールの連結部分を解除させた場合の条件に一致する異常音 14 個を選択した．また，異常音と同数の正常音も用意した．



図 3.1: 分析対象の鉄道模型 (画像上部の車両が HO スケール) [21]



図 3.2: 正常状態の車輪 [21]

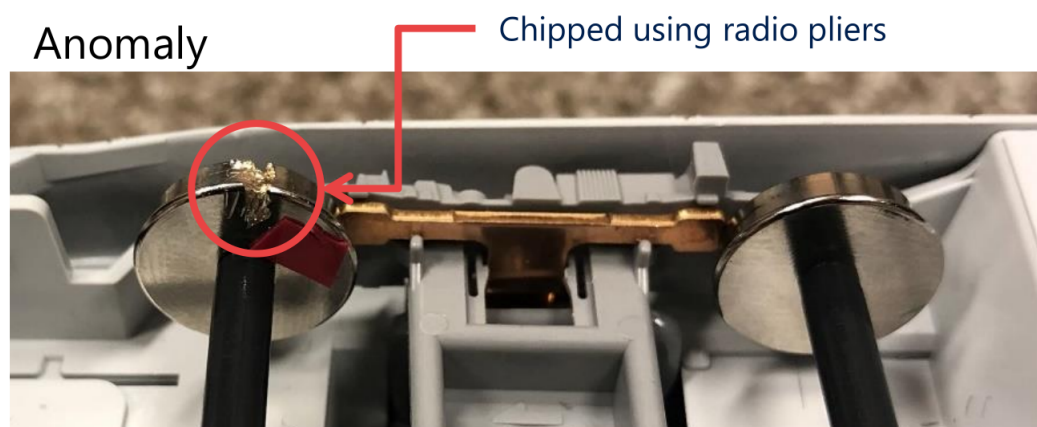


図 3.3: 異常状態の車輪 [21]

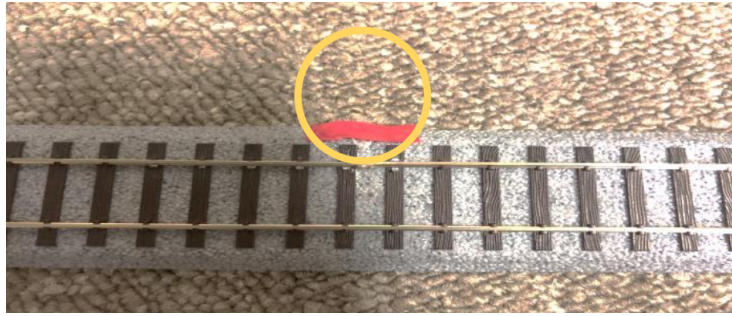


図 3.4: 異常状態の線路 (ラジOPENチによる線路破損) [21]

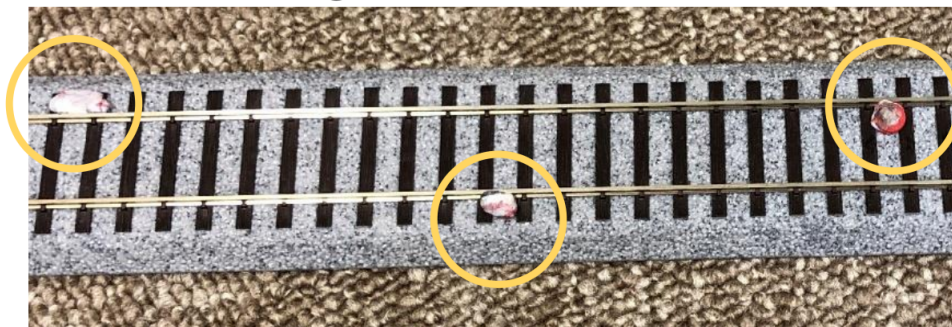


図 3.5: 異常状態の線路 (異物の配置) [21]

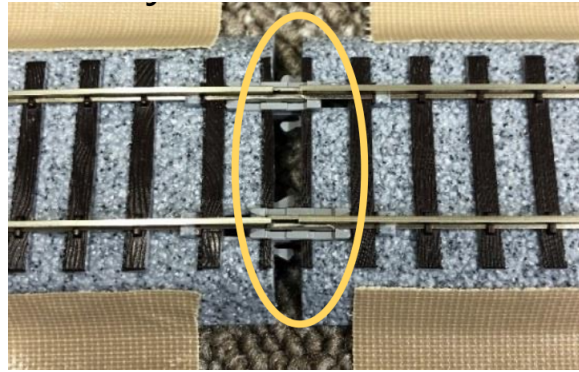


図 3.6: 異常状態の線路 (連結部分を緩めた場合) [21]

表 3.1: 分析に用いた音源と異常状態

Name	Straight railway track	Curved railway track	First carriage	Last carriage
ab01	Normal	Normal	Normal	Chipped wheel axle
ab03	Normal	Normal	Chipped wheel axle	Chipped wheel axle
ab05	Normal	Disjointed	Normal	Chipped wheel axle
ab07	Normal	Disjointed	Chipped wheel axle	Chipped wheel axle
ab09	Normal	Obstructing stone	Normal	Chipped wheel axle
ab13	Normal	Broken	Normal	Chipped wheel axle
ab15	Normal	Broken	Chipped wheel axle	Chipped wheel axle
ab17	Disjointed	Disjointed	Normal	Normal
ab18	Disjointed	Disjointed	Normal	Chipped wheel axle
ab20	Disjointed	Disjointed	Chipped wheel axle	Chipped wheel axle
ab22	Disjointed	Obstructing stone	Normal	Chipped wheel axle
ab24	Disjointed	Obstructing stone	Chipped wheel axle	Chipped wheel axle
ab26	Disjointed	Broken	Normal	Chipped wheel axle
ab28	Disjointed	Broken	Chipped wheel axle	Chipped wheel axle

3.3 分析方法

分析対象の音源に対し音圧レベルを揃えた。音圧レベルは、工場や発電所などで稼働している機械を想定し、80 dB とした。音質評価指標（ラフネス、変動強度、シャープネス）の算出には、FFT アナライザである OROS 社「OR-34」および解析ソフトウェアの「Sound Quality」を用いた。

音質評価指標の分析結果に対し、正常音 / 異常音に特徴が現れるかを検討するため、音質評価指標の高次統計量（平均、分散、歪度、尖度）を求めた。高次統計量を算出するに当たり、音質評価指標の分析結果における立ち上がり部分、データ終端の不連続値を含まないように調整した。

3.4 音質評価指標における高次統計量の結果

図 3.7 に、音質評価指標の高次統計量（平均、分散、歪度、尖度）の結果を示す。図 3.7 の行は高次統計量に対応しており、最上段から平均、分散、歪度、尖度を示している。列は音質評価指標に対応しており、最左列からラフネス、変動強度、シャープネスを示している。また、赤い棒グラフは異常音の結果を、青い棒グラフは正常音の結果を示している。

図 3.7(a),(b) より、ラフネスと変動強度の平均は、異常音の場合に高いことがわかる。また、図 3.7(d),(e) より、ラフネス、変動強度の分散も、異常音の方が正常音よりも高いことがわかる。一方、図 3.7(c),(f) よりシャープネスの平均と分散において、正常音 / 異常音に差異は確認できない。

図 3.7(g),(h),(i) より、各音質評価指標の歪度は正常 / 異常ともに負数の結果となっており、歪度から正常 / 異常の差異は確認できなかった。同様に図 3.7(j),(k),(l) に示す尖度においても、正常 / 異常に差異は確認できなかった。

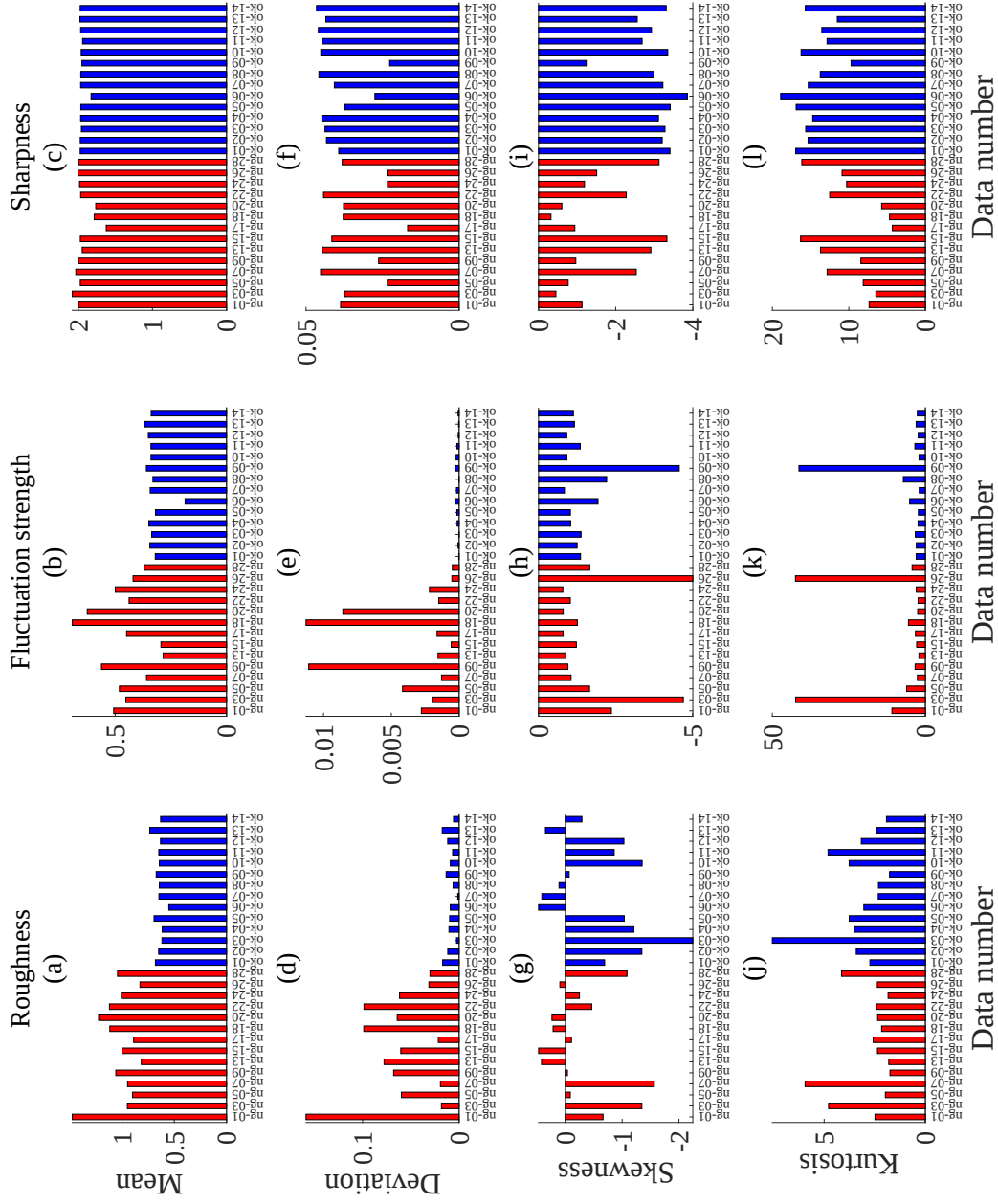


図 3.7: 模型電車における音質評価指標の高次統計量の結果 (赤:異常音, 青:正常音, 行は統計量, 列は音質評価指標に対応)

3.5 考察

図 3.7(a),(b) より，異常音におけるラフネス，変動強度の平均値は，正常音よりも高くなっている．車輪の破損箇所がレールに接触する瞬間，レールの異物や連結部分を通過する瞬間に発せられる音が，ラフネス，変動強度に影響を及ぼしているものと考えられる．また，異常状態が車輪のみの場合 (ab03) と，車輪と線路の両方である場合 (ab07,ab20) を比較すると，異常状態が複数の場合においてラフネスの平均値が高くなっていることがわかる．このことから，異常状態が複数であるほどラフネス，変動強度の平均が高くなると考えられる．

図 3.7(d),(e) より，平均と同様に異常音におけるラフネス，変動強度の分散が正常音よりも高いことが示されている．図 3.7(e) において，異常音 ab-05,ab-09,ab-18 の分散が特に高い結果であった．これらの異常状態を見ると，車輪の破損箇所が後方車両のみである点が共通している．このことから変動強度の分散は，異常状態が少ないほど高くなることが考えられる．

ラフネス，変動強度といった音の変動感に起因する音質評価指標では正常／異常に差異が確認できた．その一方で，図 3.7(c),(f) に示すシャープネスの平均，分散には，正常／異常の差異は確認できなかった．特に図 3.7(c) に示すシャープネスの平均は，正常／異常ともにほとんど差異が見られない．シャープネスに差異が現れなかった要因として，分析対象の鉄道模型に鋭い音を発する部品が少ないためと考えられる．鉄道模型において，車輪や線路に金属が用いられているが，そこから発する音よりも車体や動力部分から発せられる音が支配的であったためと予想される．また，異常状態によってラウドネスが変化しても，シャープネスへの影響は 20～30%程度にとどまるのみであるため，正常／異常に差異が現れなかったものと考えられる．

第4章 分類精度の評価

4.1 目的

第3章で，異常音はラフネス，変動強度に影響することがわかった．本章では，ラフネス，変動強度の統計量を用いて正常／異常の分類を行い，その分類精度を評価する．また ToyADMOS に加えて，熟練技術者が実際に判別した機械動作音（モーター音）のデータベースに対しても分類精度の評価を行う．

4.2 機械動作音のデータベース

実際の機械動作音として，回転機器（モーター）動作音を利用した．モーター音は合計 37 個あり，このうち 30 個は正常と判定されたもの，7 個は異常と判定されたものである．

4.3 分類方法と精度の評価

第3章で，ラフネスの平均，分散と変動強度の分散の値において，正常／異常に差異があることがわかった．そこで，これらの指標および統計量を閾値として正常／異常の分類を行った．

分類精度の評価を行うに当たり，データベースを分析用とテスト用に分割した．閾値の算出には，分析用データに対し ROC 曲線 (Receiver Operatorating Characteristic curve) を用いて最適な閾値を求めた．ROC 曲線で求めた閾値をテスト用データに適用し，そのときの，再現率，適合率，特異度，正確度，F 値，偽陰性率，偽陽性率を求めた．

4.4 ToyADMOS に対する検知精度

ラフネスの平均，分散，変動強度の分散から描かれる ROC 曲線と閾値をそれぞれ図 4.1，図 4.2，図 4.3 に示す．図 4.1，図 4.2，図 4.3 より，ラフネスの平均の場合，閾値は 0.7 asper，分散の場合 0.018，変動強度の分散の場合，閾値は 0.0004 であった．

ROC 曲線から得られた閾値を，テスト用データに適用した場合の検査精度を表 4.1 に示す．表 4.2 より，F 値は 80%を超えており，高い分類精度であることがわかった．

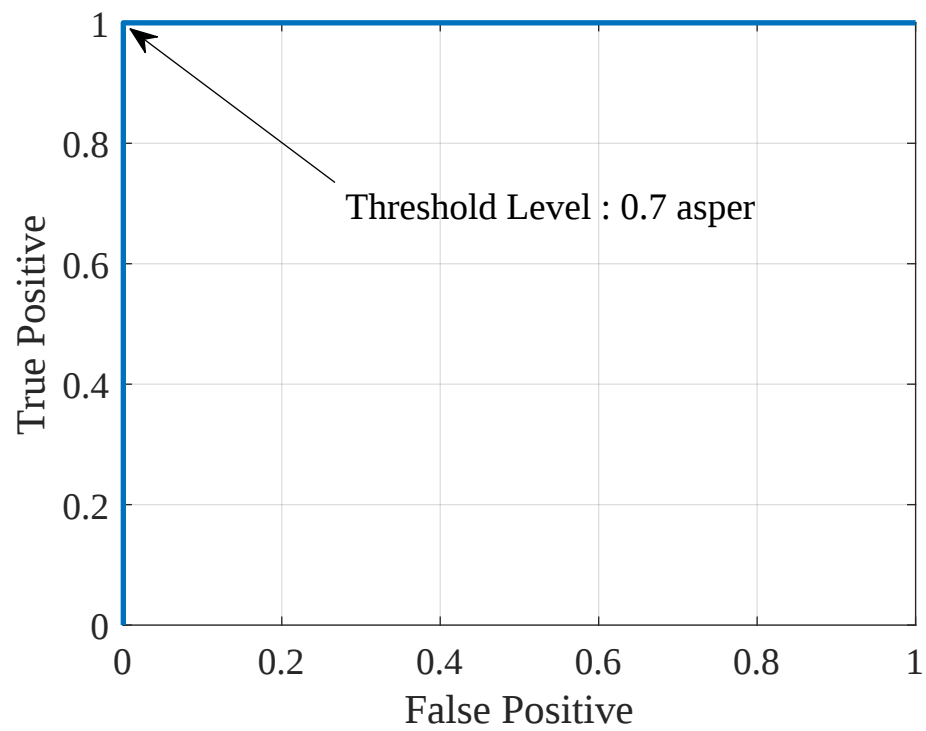


図 4.1: ラフネスの平均を閾値とした場合の ROC 曲線

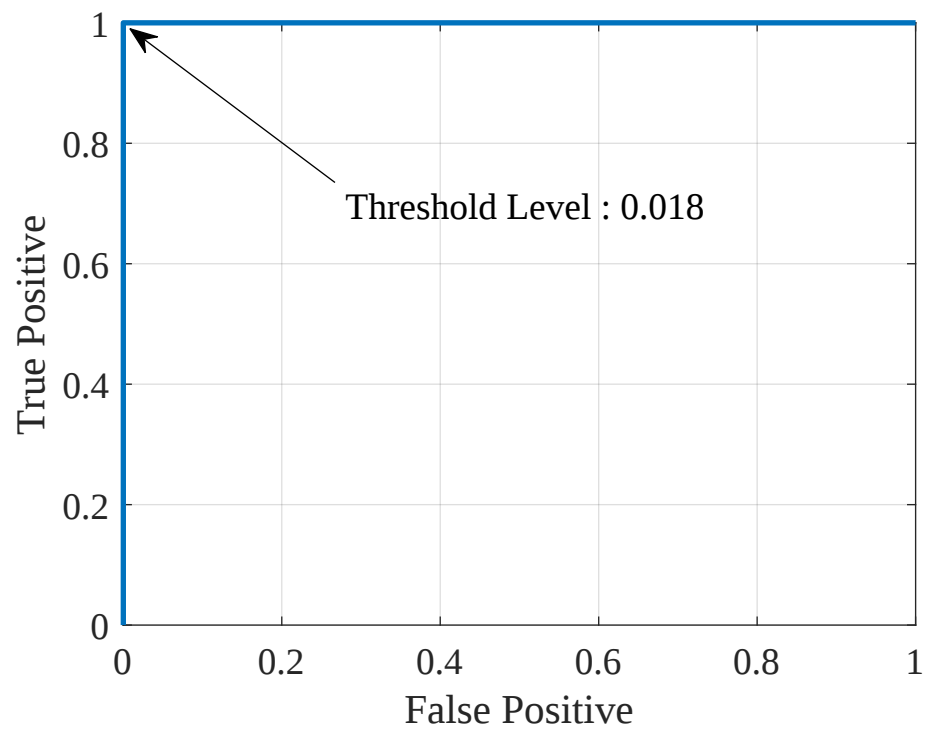


図 4.2: ラフネスの分散を閾値とした場合の ROC 曲線

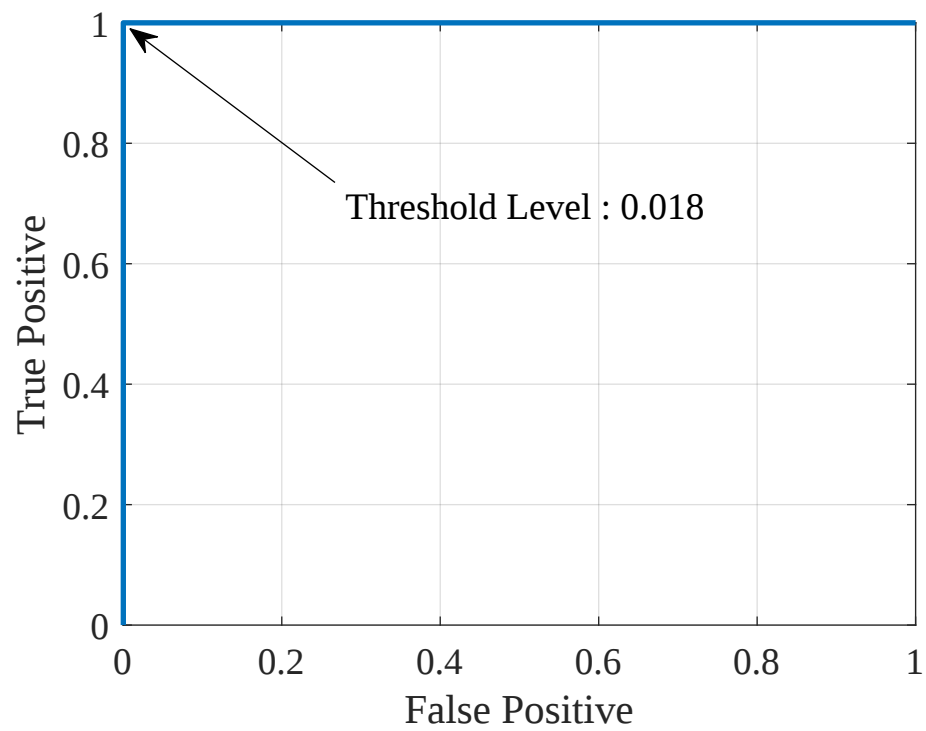


図 4.3: 変動強度の分散を閾値とした場合の ROC 曲線

表 4.1: ROC 曲線から得られた閾値をテスト用データに適用した場合の検知精度

閾値	感度 [%]	適合率 [%]	特異度 [%]	正確度 [%]	偽陰性率 [%]	偽陽性率 [%]	F 値
ラフネス 平均	85.7	85.7	85.7	85.7	14.3	14.3	0.857
ラフネス 分散	100	100	100	100	0	0	1.000
変動強度 分散	100	87.5	88.9	93.8	0	11.1	0.933

4.5 機械動作音に対する検知精度

ラフネスの平均，分散，変動強度の分散から描かれる ROC 曲線と閾値をそれぞれ図 4.4，図 4.5，図 4.6 に示す．また ROC 曲線から得られた閾値を，テスト用データに適用した場合の検査精度を表 4.2 に示す．

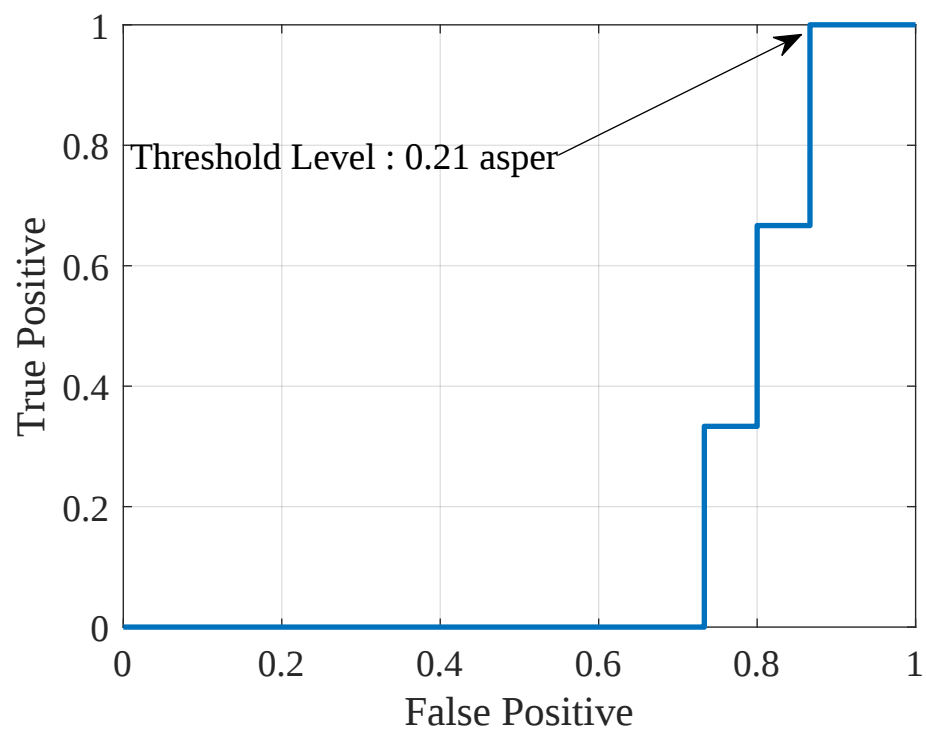


図 4.4: ラフネスの平均を閾値とした場合の ROC 曲線

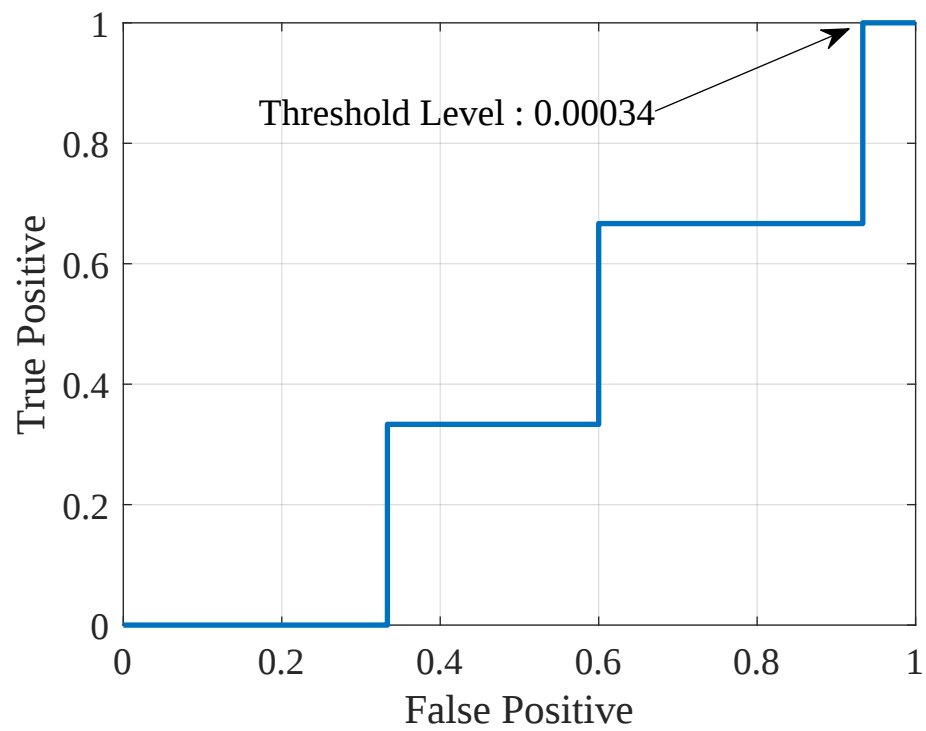


図 4.5: ラフネスの分散を閾値とした場合の ROC 曲線

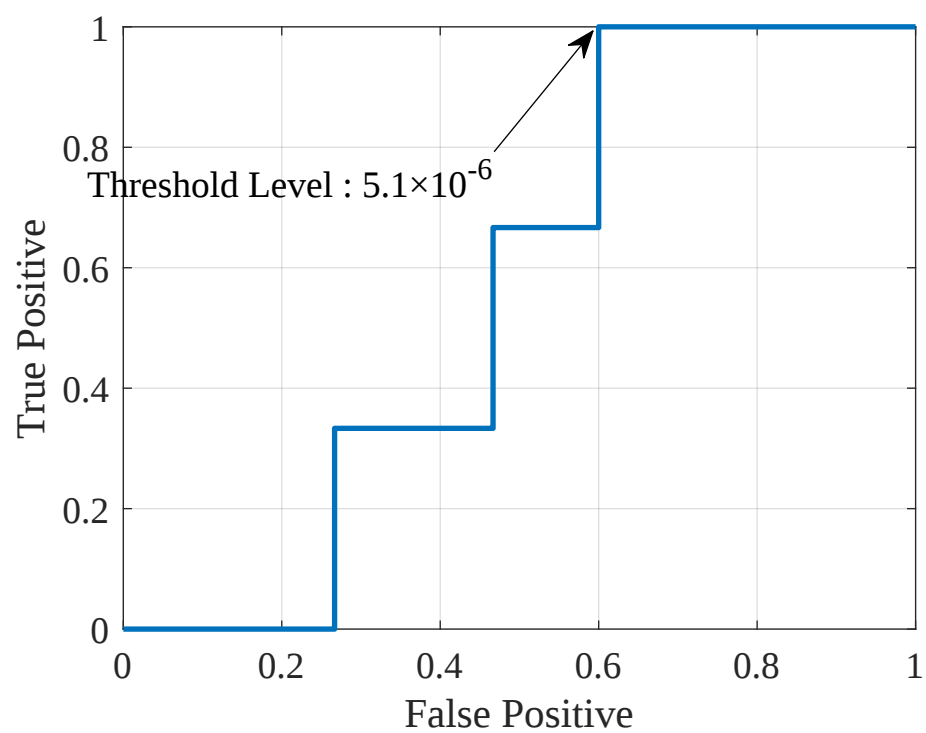


図 4.6: 変動強度の分散を閾値とした場合の ROC 曲線

表 4.2: ROC 曲線から得られた閾値をテスト用データに適用した場合の検知精度

閾値	感度 [%]	適合率 [%]	特異度 [%]	正確度 [%]	偽陰性率 [%]	偽陽性率 [%]	F 値
ラフネス 平均	46.2	85.7	0	42.9	81.3	0	0.600
ラフネス 分散	50.0	85.7	50.0	50.0	50.0	50.0	0.632
変動強度 分散	20.0	100	100	33.3	80.0	0	0.667

4.6 考察

4.6.1 ToyADMOS

ラフネスの平均，分散と変動強度の分散を閾値とした場合，図 4.1 から図 4.3 より 100%の精度で分類できた．ROC 曲線から得られた閾値をテスト用データに適用した場合，表 4.1 よりラフネスの平均，分散，変動強度の分散における F 値はそれぞれ 0.857，1.000，0.933 であった．同じデータに対し，MFCC を特徴量とした機械学習を用いた場合，F 値は 0.834 であり，本提案と同程度の結果を得られた．これらの結果を踏まえると，異常音は変動感に係る指標であるラフネス，変動強度に特徴が現れるものと考えられる．

4.6.2 機械動作音

実際の機械動作音であるモーター音に対しても，ラフネスの平均，分散と変動強度の分散を閾値として正常 / 異常の分類を行った．表 4.2 より，F 値はそれぞれ 0.600，0.632,0.667 と，ToyADMOS の結果と比較すると著しく精度が低くなった．特に，偽陰性率がラフネスの平均，分散ではそれぞれ 81.3%，50.0%，変動強度の分散では 80.0%と 50%以上の確率で誤った陰性判定を行っている．

図 4.7 に示す，モーター音における音質評価指標の高次統計量を見ると，ラフネスの平均，分散，変動強度の分散は正常 / 異常に関わらず同程度であることがわかる．この他，シャープネスの平均，分散や各音質評価指標の歪度，尖度を見ても，正常 / 異常との間に顕著な差異が確認できない．

正常音と異常音の違いを考えた際に，異常と判別された音は，通常稼働している音に機械の異常状態によって発せられる音が加わったものであると考えられる．ToyADMOS では，破損した車輪や線路，線路上の異物が，通常稼働音に加わり，それらがラフネスや変動強度の値に影響していた．しかし，モーター音の場合，通常稼働音が，異常状態によって発せられる音（擦れ音やサーサー音）よりも支配的であり，ラフネス，変動強度の値に差異が現れなかったと考えられる．このため，異常状態によって発せられる音が極端に小さい場合は，音質評価指標による異常音検知が困難であるといえる．

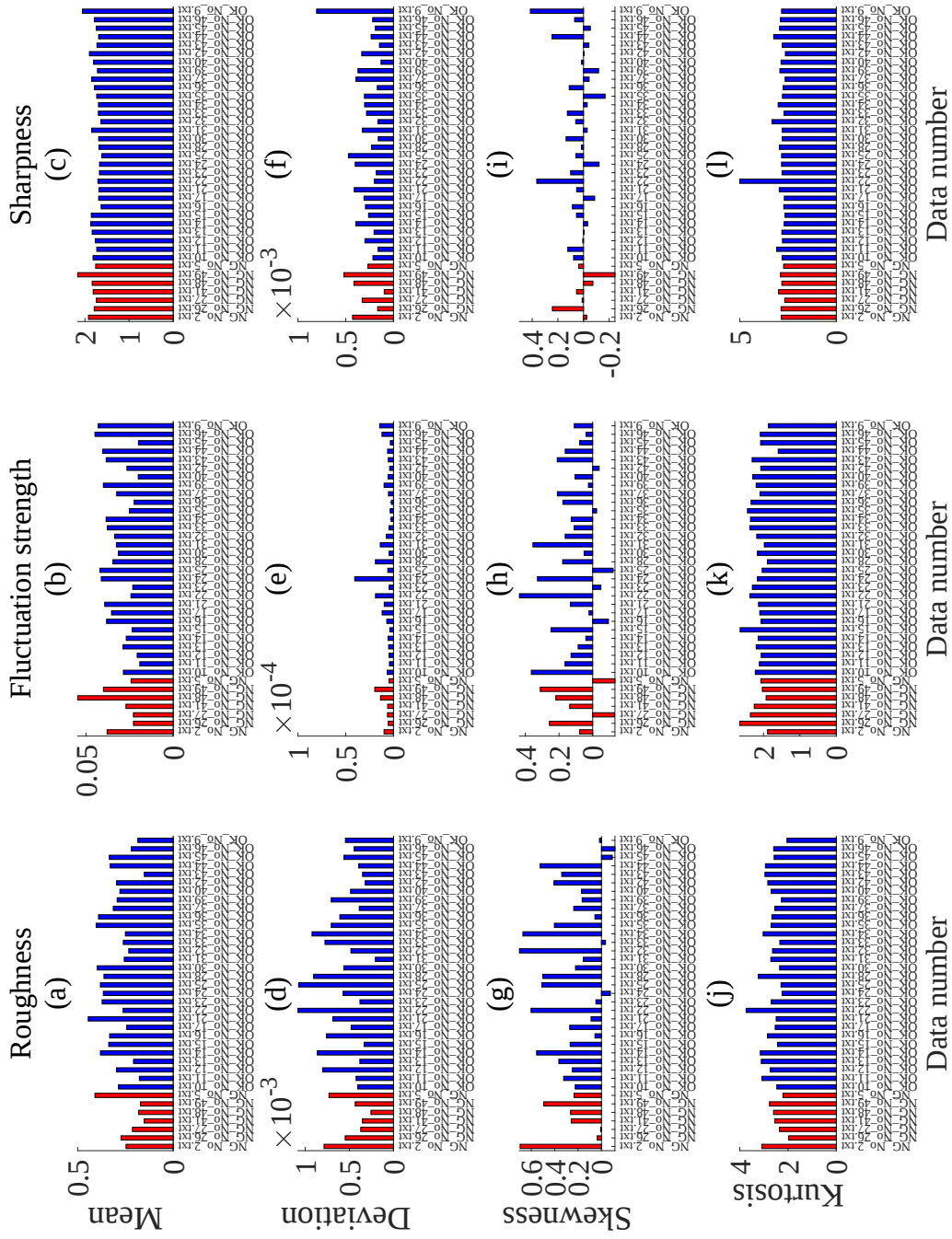


図 4.7: モーター音における音質評価指標の高次統計量の結果 (赤:異常音, 青:正常音, 行は統計量, 列は音質評価指標に対応)

第5章 全体考察

正常 / 異常と判別された機械動作音のデータベース (ToyADMOS, 模型電車) の音質評価指標 (ラフネス, 変動強度, シャープネス) を分析した結果, 異常音において, ラフネスと変動強度の平均が正常音のそれよりも高い結果が得られた. 一方, シャープネスの平均は正常 / 異常に関わらず差異はなかった. ラフネス, 変動強度共に, 音の変動に関連する指標であることから, 異常音は正常音に比べて音の変動感, 粗さ感が強いといった特徴を有していると考えられる. ラフネスの平均においては, 変動強度の平均よりも正常 / 異常との間に顕著な差異があるといえる. 今回分析した模型電車そのものが小型であり, 車輪の回転数や収録マイクを通過する速度が高速であることから, 異常音が変動感ではなく粗さ感を有するものとなったと考えられる. ラフネス及び変動強度の分散は, 平均と同様正常 / 異常との間に顕著な差が確認できた. この結果より, 異常音におけるラフネス, 変動強度は時々刻々と変化していることが示された. 異常状態別に比較すると, 車輪が1両分のみ破損した場合と2両分破損した場合では, 1両分のみ破損した場合においてラフネス, 変動強度の分散が高くなっていることがわかった. また, 車輪に加え線路も異常状態にした場合, 分散は小さくなる傾向になった. 異常の要因が複数存在する場合, 各々の異音 (車輪や線路から発する音) が混合することで, 変動感や粗さ感の時間変動が小さくなるものと考えられる. このことから, 異常箇所が多い機械の場合, ラフネス, 変動強度の分散の観点から正常 / 異常検知が困難であることが考えられる.

実際の機械動作音として, モーター音の音質評価指標を分析した結果, ToyADMOS に現れたようなラフネス, 変動強度に顕著な差異が確認できなかった. 正常 / 異常に差異が現れなかった要因として, 正常なモーター音または収録環境雑音の音圧レベルが異音よりも大きいことが考えられる. ToyADMOS では, 破損した車輪, 線路を通過する際の異音が, 通常動作している音よりも大きいため, 熟練レベルに達していない人でも正常 / 異常の判別が容易であった. 一方モーター音は, 熟練技術者が正常 / 異常を判別したものであり, 熟練レベルに達していない人は判別が難しいものとなっている. 熟練技術者がどのようにして判別したかを考えたとき, 異音の全周波数帯域を聴いているのではなく, 特定の帯域に絞って聴いていると予想される. その特定の帯域における僅かな音の違い (変動感, 粗さ感, 鋭さ) を手がかりに正常 / 異常を判別しているものではないかと考えられる.

第6章 結論

6.1 明らかにしたこと

本研究の目的は、機械動作音の粗さ感、変動感、甲高さといった聴感印象にどのような特徴が現れるかを検討し、それらの特徴を手がかりに異常音検知が可能であるかを明らかにすることである。そのために、おもちゃの動作音データベースに対し音質評価指標を分析し、正常／異常における特徴を検討した。その結果、異常音はラフネスと変動強度の平均、分散が正常音よりも高いという結果が得られた。また、シャープネスは正常／異常との間に差異が確認できなかった。これらの結果より、異常音は音の変動に関する指標（ラフネス、変動強度）の平均と分散に特徴があることを明らかにした。

ラフネス、変動強度の平均、分散を用いて、おもちゃの動作音の正常／異常の分類精度を評価した。その結果、F 値が 0.85 以上という結果が得られた。同じデータベースを MFCC を特徴量として機械学習によって分類した精度と同等であった。しかし、熟練技術者が判別した機械動作音（モーター音）に対しては、F 値は 0.60 ～ 0.67 と精度は低下した結果が得られた。

これらの結果から、熟練レベルに達していない者でも判別できる機械動作音に対しては、音の粗さ感、変動感に対応する音質評価指標である、ラフネス、変動強度を手がかりに異常音検知を実現可能であることが示唆された。

6.2 残された課題

6.2.1 臨海帯域ごとの音質評価指標の分析

熟練レベルに達していないような人でも、判別が容易な機械動作音に対しては、高い分類精度を出すことができた。しかし熟練技術者が判別した機械動作音に対しては、分類精度が低い結果となった。本研究で行った音質評価指標の分析は、指標の算出にはすべての臨界帯域を用いている。もし、熟練技術者が異常の手がかりとしている音が、ある特定の帯域に現れているのならば、他の帯域の音によって、手がかりとなる音が隠されている可能性がある。このため、臨界帯域ごとにラフネス、変動強度を調べることによって、熟練技術者が判別した音に対しても音の変動に関する指標の平均、分散に特徴が現れると考えられる。

6.2.2 異常音検知の聴知覚メカニズムの解明

本研究では，異常音検知の手がかりとして，ラフネスや変動強度といった音の変動感を用いている可能性が示唆された．この結果から，異常音検知では聴知覚の観点からも考慮しなければならないことが考えられる．本研究で扱ったラフネス，変動強度，シャープネス以外の音色についても調べることで，異常音検知の聴知覚メカニズムの解明につながると考えられる．

参考文献

- [1] 佐藤建吉, “産業革命,” 日本機械学会誌, Vol. 109, No. 1054, pp. 767–770, 2006.
- [2] 寺島真介, “音響信号による設備監視 音響診断技術とその適用事例,” 紙パルプ技術協会誌, Vol. 57, pp. 90–97, 2003.
- [3] 日本音響エンジニアリング, “音を聞き分けるトレーニング「聴能形成」,” <https://www.noe.co.jp/technology/43/43news3.html> (2020 年 1 月 27 日アクセス)
- [4] 総務省統計局, “平成 27 年国勢調査: 就業状態等基本集計結果,” <https://www.stat.go.jp/data/kokusei/2015/kekka.html> (2020 年 2 月 3 日アクセス)
- [5] 小泉悠馬, 齊藤翔一郎, 植松尚, “深層学習を用いた機器動作音の異常音検知,” 日本音響学会春季研究発表会講演論文集, pp. 473–476, 2017
- [6] 岩宮眞一郎, 小澤賢司, 小坂直敏, 山内勝也, 高田正幸, 藤沢望, 音色の感性学, コロナ社, 東京, 2010.
- [7] 鈴木伸洋, “音と画像で発電設備の異常を検知,” 電気学会誌, Vol.112, No. 3, pp. 157–161, 1992.
- [8] T. Anmen, T. Murosaki, N. Takaha, H. Ito, Y. Kondo, “Fricative inspection with wavelet analysis,” *SICE Annual Conference*, pp. 115–122, 2004.
- [9] 大瀧淳司, 小菅太, 三澤勉, 植原孝之, 藤江龍一, “異音を定量的に検知する手法の開発,” 富士通テン技報, Vol. 23, No. 2, pp. 37–46, 2005.
- [10] 渡辺嘉二郎, 豊田成人, “異常音による機械の診断 -軸受の呼び番号の同定と故障の診断-, ” 計測自動制御学会論文集, Vol. 29, pp. 94–101, 1993.
- [11] Y. Koizumi, S. Saito, H. Uematsu and N. Harada, “Optimizing acoustic feature extractor for anomalous sound detection based on neyman-pearson lemma,” *25th European Signal Processing Conference*, pp. 698–702, 2017.
- [12] 河内祐太, 小泉悠馬, 原田登, “Lp ノルム回帰を用いた異常音検知の検討,” 日本音響学会秋季研究発表会講演論文集, pp. 533–534, 2017.

- [13] 近藤稔, “振動の音色で機器の異常を検知する,” 鉄道総合技術研究所, Vol. 74, pp. 16–19, 2017.
- [14] 伊藤彰則, “環境音から異常を検知する統計的手法,” 日本音響学会誌, Vol. 75, No. 9, pp. 1–6, 2019.
- [15] 井本桂右, “音響イベントと音響シーンの分析,” 日本音響学会誌, Vol. 74, No. 4, pp. 198–207, 2018.
- [16] K. Yohei and E. Takashi, “How can we detect anomalies from subsampled audio signals?,” *IEEE International Workshop on Machine Learning for Signal Processing*, pp. 1–6, 2017.
- [17] E. Marchi, F. Vesperini, F. Eyben, S. Squartini and B. Schuller, “A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional LSTM neural networks,” *IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 1996–2000, 2015.
- [18] C. Clavel, T. Ehrette and G. Richard, “Events detection for an audio-Based surveillance system,” *IEEE International Conference on Multimedia and Expo*, pp. 1306–1309, 2005.
- [19] A. Ito, A. Aiba, M. Ito and S. Makino, “Detection of abnormal sound using multi-stage GMM for surveillance microphone,” *5th International Conference on Information Assurance and Security*, Vol. 1, pp. 733–736, 2009.
- [20] K. Suzuki, M. Tanaka and H. Inoue: “Clustering and modeling of hazardous or non-hazardous sound in daily life,” *SICE Annual Conference*, pp.1991–1995, 2012.
- [21] Y. Koizumi, S. Saito, H. Uematsu, N. Harada and K.Imoto, “ToyADMOS: A Dataset of miniature-machine operating sounds for anomalous sound detection,” in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2019.
- [22] Y. Koizumi, S. Saito, Y. Masatakau, M. Shin and H.Noboru, “Batch uniformization for minimizing maximum anomaly score of DNN-based anomaly detection in sounds,” *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2019.
- [23] H. Fastl, “Fluctuation strenght and temporal masking patterns of amplitude-modulated broadband noise,” *Hearing Research*, 8, pp. 59–69, 1982.

- [24] H. Fastl., E. Zwicker, “Psychoacoustics : Facts and models, third edition” Springer, pp. 257-264, 2006.
- [25] 高田正幸 , “音質評価に基づいた騒音対策と音のデザイン ,” 日本騒音制御工学会 , Vol.10 , No.10, pp.582–589, 2019.
- [26] 高田正幸 , “音質評価指標の計算法と 適用事例 ,” 日本音響学会誌 , Vol.29 , No.4, pp. 247–252, 2005.
- [27] 前田修, “音質評価のためのエンジン音合成技術,” 日本機械学会第 11 回環境工学総合シンポジウム 2001 講演論文集, pp. 49–52, 2001.
- [28] 岡本宜久 , 古郡了, “電子合成自動車音による音質評価,” 自動車技術, Vol.45, No.12, pp. 29–36, 1991.
- [29] M. Klemenz, “Sound quality of starting electric railbound vehicles - dimensions of hearing and relevant psychoacoustic quantities,” *Acta Acustica united with Acustica*, Vol. 91, pp. 779–788, 2005.
- [30] 多喜健司 , 大久保伸行 , 戸井武司 , 大川成樹 , “ミシンの音質評価に基づく音質改善手法の開発 ,” 日本音響学会春季研究発表会講演論文集, pp.753–754, 2002.
- [31] M. Takada, S. Arase, K.Tanaka and S. Iwamiya, “Economic valuation of sound quality of noise emitted from vacuum cleaners and hairdryers by conjoint analysis ,” *Noise Control Engineering Journal*, Vol. 57, pp. 263–278, 2009.
- [32] Y. Koizumi, S. Saito, H. Uematsu, N. Harada and K. Imoto, “ToyADMOS dataset”, <https://github.com/YumaKoizumi/ToyADMOS-dataset>(2020 年 1 月 27 日アクセス)

研究業績

国内発表

1. 水野滉介, 鵜木祐史, “音質評価指標を用いた異常音検知の研究,” 2019 年度電気・情報関係学会北陸支部連合大会, G-1 ,2019.

謝辞

研究活動から私生活まで厳しくも非常に丁寧な指導をしていただいた，主指導教員である鵜木祐史教授に深く感謝致します．

研究室会議やミーティングなどの場において，赤木正人教授，木谷俊介助教は研究のあり方や考察のご指導を数多くいただきました．個々に感謝致します．

本研究に関して多大な助言をしていただいた鳥谷さん，磯山さん，関谷さん，高橋さんに心より感謝申し上げます．

並河さん，畠山さん，川村さん，大谷さん，松本さんを始めとする先輩の皆様方には，研究や研究室運営など様々なお助言をいただきました．心より感謝いたします．

竇坂さん，西垣さん，宮川さん，森田さん，黄さんとともに過ごした時間は，研究室生活や人生において大変有意義で，精神的にも支えられました．深く感謝致します．

研究で行き詰まったときに相談に乗っていただいた，鵜木・赤木研究室のスタッフを含めた全メンバーの皆様に感謝致します．

最後に，これまで私の学生・研究生活を様々な面で支えてくださった家族の皆様に心から感謝致します．