

装置操作遠隔訓練のための異音に着目した 音響特徴量と装置状態推定技能の関係

Relationship between Acoustic Features Focused on Abnormal Noise and Equipment Condition
Estimation Skills for Remote Training of Equipment Operation

藤澤 充紘[†]
Mitsuhiro Fujisawa

中平 勝子[†]
Katsuko T. Nakahira

1 はじめに

装置の操作には、装置の状況の把握が肝要であり、人は視覚、聴覚、触覚情報などから様々な要素をモニタリングしている。そして、得た情報を用いてフィードバックを行って装置を操作している。その際に普段とらない挙動や、異音といった異常事態を知覚した際に、操作の不手際や、装置の不具合が想定され、それに対して適切な対応が要求される。

しかしながら、五感や経験などに頼った装置の診断は個人に依存する部分が存在し、判断基準に個人差が生じてしまう。また、このような知覚による技能は共有することが難しいために、ベテランからの技術の継承が困難である。そして、ベテランが退職することにより熟練技の継承が断絶してしまうことを防ぐため、技能の教育は急務となっている。

また、昨今の感染症の流行により、実地で多くの研修を行うことが難しい時期があったことは記憶に新しい。そのおりに、大学をはじめとする各種教育機関では、遠隔による教育を施すことで教育の連続性に対応した。その際、問題となったのも、やはり実技を伴う実験、芸術にかかる教育の質保証であった。これまでも多くの e ラーニング実践によって遠隔による教育の可能性追求はなされていたが、いわゆる選定された教育機関による先駆的实践が多く、いざ、実際の教育機関へ落とし込もうとした際には、技術導入を含め、多くの問題があったことは否めない。

現在、感染症は収束しつつあるが、それと同時に、遠隔による教育の可能性が改めて見直されることとなった。この流れは、各種技能教育に対しても同じであろう。こうした情勢を踏まえると、装置操作の訓練についても、現地でなくても行える可能性の追求を行うことが新たな技能教育の可能性の観点から好ましいと考える。その際、装置操作となると、装置に対する視点を様々に変更できることが好ましいため、通常のビデオオンデマンド教育よりも VR、メタバース空間での教育がより有益であると考え。そこで、本稿では、VR による遠隔技能教育を行うことを念頭に置いた。

教材を作成する際に、視覚情報であれば、異常を動画に収めて、それを教育に用いることが可能である。しかし、聴覚情報は、音高やパターン性などは録音できるが、音がどこからなっているかといった空間的な情報は一般的な動画だと収めること

ができない。そのため、異音に関する教育は、異常な挙動などの視覚情報と異なり、再現をし、教育に用いることが難しい。それに加えて、聴覚情報は視覚情報に対して、聞こえる範囲内であれば、どこを向いていても情報が伝達される（無指向性）。そのため、注視していなくても異変に気づきやすい。そのため、音情報に着目した。

そこで、本稿では、装置の状態推定を音響特徴量によって訓練することを念頭に置き、技術者の異音感知の技能レベルを調べるために正常時、異常時の音の違いはどのような要素、成分にあるのかを周波数解析を行い、違いを可視化し、教育に用いるための違和感のない異音の再現はできるのかを検討する。また、仮想現実空間にて機械を映し出し、立体音響にて異音を再生する異音教育について確認する。正常時の音声と異音の聞き分け、音像定位は可能か、またそれらの教育の効果の程度に関して調べる。

2 設計

2.1 SRK モデル

SRK モデルは Rasmussen(1986)[1] が提唱した熟練度と人間の行動の制御の関係を示したモデルである。熟練度別に、知識 (K) ベース、ルール (R) ベース、スキル (S) ベースの 3 段階に分類される。知識ベースは、初心者の段階であり、ひとつひとつに注意を払い知識や経験を活用し考え操作を行う。ルールベースは、慣れてきた段階であり、すでに手順書が作成されていて、手順書通りに操作を行える。スキルベースは、無意識的にこなせる段階であり、繰り返しの作業が身にしみつき、考えずとも体が動く。

装置の異音の対処について考えると、

- 知識ベースの場合では、音を聞き、どこから鳴っているかの同定を行い、そこから原因を一から考えて、対処法を一から考える。
- ルールベースは、音を聞き、どこから鳴っているかを再認し、その発生源、音の種類から考えられるパターンから原因を絞り込み、経験則的に対処法を探る。
- スキルベースは、音を聞き、原因がわかり、対処法に直結する。

といった様に当てはめることができる。

原因の推定を行う際に、装置に対する知識がないと異常が発生しても原因の推定が行えない。さらに言えば異常を異常とし

[†] 長岡技術科学大学

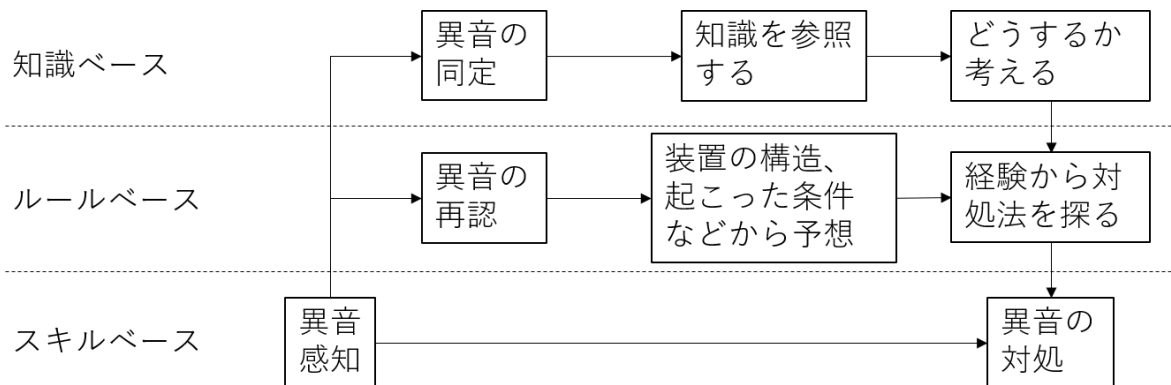


図1 SRK モデル

て認識できないことも考えられる。そこで、正常時の動きや音を知らせたうえでの異常に対する教育が必要である。

2.2 没入経験

VR(バーチャルリアリティ)とは、コンピュータによって作られた仮想的な空間内に入り込み、その中で、様々な疑似体験を行える技術のこと [4] であり、使用者は HMD を着用し、この体験をすることができる。そして、現時点で VR に実用化できているマルチモーダルインターフェースは視覚と聴覚のみである。そこで、聴覚の空間的な知覚に視覚が影響を及ぼす効果の代表例として「腹話術効果」がある。これは、人形の口の動きに合わせた腹話術師の声が、その人形から発せられているように聞こえる現象である。このような視覚と聴覚が同期している条件下では音像定位がバイアスにより向上し、音源が正面付近で効果が最大となるという報告もある [6]。そのため、VR と相性がいいと考えた。

また、VR は技術的發展により実用に耐えうようになり、様々な分野で応用されており、教育分野もその一つである。例として、防災 [5] や臨床実習 [3] などに用いられている。このように VR コンテンツの拡充も求められることも想定される。

本稿では、視覚情報である装置の動作と聴覚情報である装置音という要素に着目しているため、視覚と聴覚のクロスモーダル効果を没入感によりさらに期待できる VR を用いた実験を行うこととした [2]。

2.3 撮影

スクリーン投影用の撮影として SONY Exmor R5 を使用した。また、VR 映像用の撮影に Canon EOS R5 C に RF5.2mm F2.8 L DUAL FISHEYE を装着したものを使用した。RF5.2mm F2.8 L DUAL FISHEYE は魚眼レンズが水平に2つついており、180度のVR映像を撮影可能にするレンズである。それに加え、指向性ステレオマイクロホンの DM-E1D を装着した。

上記のカメラをそれぞれ三脚に固定し、視角の都合上、それぞれ別に撮影した。

Canon EOS R5 C で撮影した画像を図2に示す。これを EOS VR Utility にて正距円筒図法への変換を行い、視差補正、

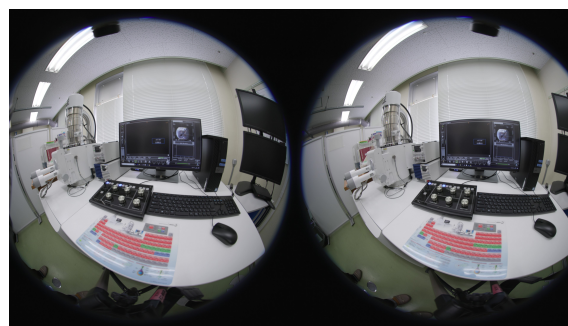


図2 DUAL FISHEYE での写真

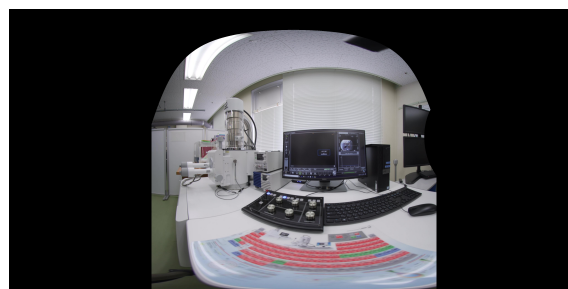


図3 正距円筒図法による変換後(視差補正, 水平補正あり)

水平補正をかけたものを図3に示す。今回録音した異音は電子顕微鏡のテーブルの X 軸の移動する際の駆動音に金属的な響きのノイズが乗るものであった。

これを meta quest2 にプリインストールされている再生ソフトにて再生する。

3 教材作成

3.1 教材設計

技術者の熟練度別に、つまり、知識ベース、ルールベース、技能ベースのそれぞれの技能レベルで必要となる教育が異なってくる。

そのため、用いる教材も技術者の技能レベルに合わせて変更する必要がある。そこで、録音した正常な音源にフィルターをかけ、誰でもわかる程度に正常音からかけ離れる具合にノイズを混ぜた音や、少しだけノイズを混ぜた音などを用いて予備実

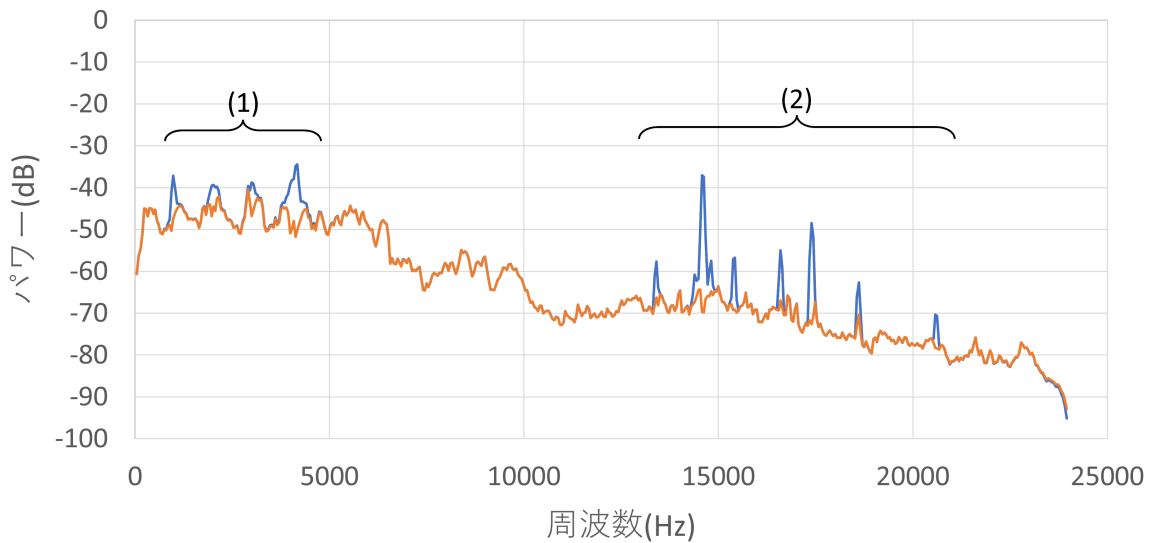


図 4 異音に対するノイズカット前後の周波数解析結果。青線がノイズカット前，オレンジ線がノイズカット後。

験を行って，作成した音源がどの程度の難易度のものになっているのかを確認する。また，高周波が聴き分けに影響する場合は，単純な聴力による影響が発生する可能性がある。そのため，実験時には被験者の聴力をあらかじめ測定する必要がある。

3.2 周波数解析

周波数解析にはフリーソフトの Audacity を使用した。

窓関数にはハンギング窓を使用し，周波数は 46.875Hz 刻みでの分解能で行った。録音した異音に対して周波数解析を行った結果の一例を図 4 の青線で示す。

FFT は周波数スペクトル推定であり，必ずしも適当な結果が得られるとは言えないため，無作為に抽出した 5 点に対して周波数解析を行った。しかし，すべてに図同様に (1)，(2) に相当する部分にピークが発生するような出力が得られた。

図中，(1) に相当する部分，すなわち 1000Hz から 5000Hz までの間で約 1000Hz 刻みでピークが生じている。また，(2) に相当する部分でも，14.6kHz から約 1000Hz 刻みでピークが生じている。

今回発生したピークは (1)，(2) と少し離れたところにどちらもおよそ 1000Hz 刻みで発生している。これについて (2) で発生したピークから何らかの影響を受けて (1) のピークが生じたといった考えをすることができる。こうしたときの原因の一つとして (2) の振動の倍音成分で装置の筐体部分，あるいは駆動部が共振して (1) のピークが生じたと考えられる。しかしながら，装置を無視して録音を行うのは現実的ではなく，また，電子顕微鏡のモデルも複雑であるため，再現して固有周波数の解析を行うのも現実的とは言えず，検証するのは難しい。

上記から推測された周波数スペクトルに対して Audacity のノッチフィルターを使用しフィルタリングを行った。ノッチフィルターとは，特定の波長帯域を非常に低いレベルまで減衰

させ，それ以外のほとんどの波長を少ない強度損失で通過するフィルターである。前述したピークをノッチフィルターで減衰させ，

- (1) 1000Hz から 5000Hz
- (2) 14.6kHz 以上
- (3) 1000Hz から 5000Hz と 14.6kHz 以上

の 3 パターンでフィルタリングを行った。(3) の結果を図 4 のオレンジ線に示す。

(3) を見ると，約 7kHz までは横ばいになっており，そこから約 10kHz までの間にパワーが低くなっている。また，そこから約 17kHz まで横ばいになり，そこからパワーが低くなっていている。そのため，この音の主成分は 7kHz 以下であると考えられる。

ここから得られた音源を聴取したところ，(3) では未編集のもの比べて，高域のピークを削ったために耳に刺さる音ではなくなった。しかし，それでも明らかに異音と認識できる音であった。

SRK モデルに基づいて考えると，教材として要求されるもののレベル分けは

- 知識ベースでは，元の音をわかるような異音
- ルールベースでは，元の音を知っていれば比較してわかるような異音
- スキルベースでは，元の音を知っていても訓練しなければ気づけないような異音

と分けて考えられる。今回録音した異音は知識ベースの教育に有用であると思われる。

4 まとめと今後の課題

現状では、異音の原因となっている要素について解明できていない。そのため、装置の修理作業が終了したのちに正常音の録音を行い、正常音と異音の差分から異音の原因となっている要素の抽出を行う。それにより得たデータから技能レベルごとに用いる教材を作成する。

また、技能レベルごとに分けて作成した教材が実際に聴取者の技能レベルに即して分けられているのかを確認するために予備実験を行う必要がある。

謝辞

本研究の一部は科研費 JSPS (22K12284, 代表: 岐阜工業高等専門学校・小川信之, 23K11334, 代表: 長岡技術科学大学・中平勝子) および経営改革促進事業の助成を受けたものである。また、収録にあたり、長岡技術科学大学技術支援センター・河原夏江様、近藤みずき様には大変お世話になりました。厚く御礼申し上げます。

参考文献

- [1] Jens Rasmussen. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-13, No. 3, pp. 257–266, 1983.
- [2] 岡嶋克典. 2. 視覚情報によって誘発されるクロスモーダル効果. *映像情報メディア学会誌*, Vol. 72, No. 1, pp. 8–11, 2018.
- [3] 進士誠一, 横堀将司, 清水哲也, 神田知洋, 林光希, 安康勝喜, 吉田寛. Virtual reality 技術を活用した外科系臨床実習. *日本医科大学医学会雑誌*, Vol. 18, No. 1, pp. 98–104, 2022.
- [4] 廣瀬通孝. 第 2 世代を迎えたバーチャルリアリティの研究開発. *日本画像学会誌*, Vol. 58, No. 3, pp. 293–299, 2019.
- [5] 石井裕剛, 半田大樹, 下田宏. 屋内環境の自動モデル化機能を備えた vr 地震体験システムの開発と評価. *日本バーチャルリアリティ学会論文誌*, Vol. 21, No. 2, pp. 345–357, 2016.
- [6] 北島律之, 山下由己男. 視覚的注意が音源定位に及ぼす影響. *心理学研究*, Vol. 69, No. 6, pp. 459–467, 1999.