

CNN を用いた船舶航行音認識のための最適なスペクトログラムの検討

A-5 Consideration of suitable spectrogram parameter for boat recognition to train convolutional network

山口 晴己[†] 武藤 憲司[†] 小林 洋介^{††}

Haruki YAMAGUCHI[†] Kenji MUTO[†] Yosuke KOBAYASHI^{††}

[†] 芝浦工業大学大学院理工学研究科 ^{††} 室蘭工業大学大学院工学研究科

[†] Graduate School of Engineering and Science, Shibaura Institute of Technology

^{††} Graduate School of Engineering, Muroran Institute of Technology

1. はじめに

ある運河では、小型船舶が頻繁に航行しており、その航行音は運河沿いの住民に騒音影響を及ぼしている。

我々は航行音による騒音印象を軽減するために船舶航行音認識を行い、運河沿い住民に通知するシステムを検討してきた。甲斐らによってシステムの有効性が検討され、船舶騒音を感じる前に船舶情報を通知することで騒音への印象軽減を確認している[1]。これより、船舶通知システムのために船舶を認識する手法が検討され、映像を利用した認識手法が検討された[2]。しかし、視認性や天候の影響など課題があり、音響信号を利用した認識手法が必要とされた。

本研究の目的は機械学習による船舶航行音認識手法の確立である。航行音認識には CNN モデルにスペクトログラムを入力特徴量とした学習によって行う。本報告はこれまでに船舶航行音を認識するために行った最適なパラメータの検討結果を報告する。

2. スペクトログラムパラメータの検討

船舶航行音を認識するために我々は既存の環境音認識モデルを利用した航行音認識を行った。しかし、評価データを利用した分析において、正解率が 7 割ほどであり、システムとしての正確性を持たせることは困難であった。そのため、CNN モデルに入力特徴量として利用するスペクトログラムを船舶航行音に最適化させることで航行音認識精度の向上を図った。航行音認識に最適化させるため、スペクトログラム作成におけるパラメータの分析時間長、時間分解能、分析周波数帯域、周波数分解能の 4 項目を検討した。

分析に用いた学習データはすべて同一の 24 時間の録音データであり、評価用のデータも同一の 30 分の録音データである。これらのデータは 1 秒ごとに我々が船舶航行音の有無を分類したラベルを付与している。録音は運河沿いのマンションで行われ、サンプリング周波数は 24 kHz、スペクトログラムは 4096 点 FFT で作成した。

分析時間長の検討では、多くの環境音認識モデルは分析時間長を 1 秒としているため、船舶航行音の特徴である数秒から数分にわたりゆるやかな音を認識できる最適な分析時間長を調査した。航行音の特徴を認識することは難しいと考えた。検討の結果、分析時間長が 7 秒のとき、正解率が最大であった[3]。

時間分解能の検討では、時間分解能が高いほど細かい音情報を学習できるが、過学習を引き起こす場合があるため、スペクトログラムに利用する FFT の最適なオーバーラップ率を調査した。検討の結果、オーバーラッ

率が 90%のとき、正解率が最大であった[4]。

分析周波数帯域の検討では、船舶航行音の騒音源であるディーゼルエンジンの低い周波数の音を認識する際に、高調波が及ぼす影響を調査した。検討は 10 Hz を下限とし、上限周波数を 50 - 12000 Hz の範囲で変更した。検討の結果、低い周波数帯のみで学習を行うことで 90%の正解率に達したが、サンプリング定理上の上限値までを入力特徴量としたところ正解率が最大となった[5]。

周波数分解能の検討では、分解能が高いほど周波数ごとの音圧を詳しく学習可能であるが、過学習を引き起こす場合があるため、1/N オクターブ法を利用し調査した。検討は N = 1, 3, 6, 12 の 4 種類で行った結果、1/3 オクターブ法が最適な周波数分解能であった[6]。

3. スペクトログラムパラメータの検討結果

航行音認識のために CNN モデルに入力特徴量としてスペクトログラムを利用しそのパラメータについて検討した結果、分析時間長 : 7 秒、時間分解能 : 90%の FFT オーバラップ、分析周波数帯域 : 10 - 12000 Hz、周波数分解能 : 1/3 オクターブ法によって作成されたスペクトログラムが航行音認識に最適であった。航行音認識における正解率は 93.6%に達し、このときのスペクトログラムを図 1 に示す。これより、船舶通知システムにおいて高精度な船舶通知が行うことができると考える。また、この条件を満たすスペクトログラムを利用することで、CNN モデルに限らず他の機械学習モデルにおいても同条件の入力特徴量とすることで、高い航行音認識精度を持つと考えられる。

参考文献

- [1] 甲斐他, 音講論 (春), 2-P-22, 2018.
- [2] Akiyama *et al.*, Proc. IEEE GCCE2018, 2018.
- [3] 山口他, 人間工学会関東支大, pp. 30-31, 2019.
- [4] 山口他, 信学総大, 2020.
- [5] Yamaguchi *et al.*, Proc. ICA2019, 2019.
- [6] 山口他, 電情関北海道支大, pp. 201-202, 2019.

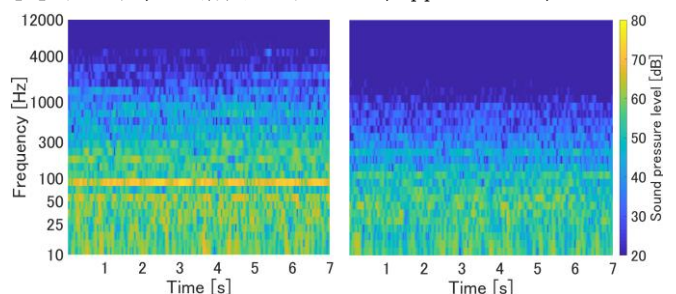


図 1. 航行音認識に最適なパラメータによるスペクトログラム