

機械学習アルゴリズムを用いた桜島の火山噴火によるインフラサウンド波形の識別

Identification of infrasound waveforms from Sakurajima volcanic eruption using a machine learning algorithm

*岡本 大輝¹、山本 真行¹

*Daiki Okamoto¹, Masa-yuki Yamamoto¹

1. 高知工科大学

1. Kochi University of technology

1. はじめに

人間の可聴周波数下限の20 Hz以下で定義されるインフラサウンドは火山噴火や津波といった大規模自然災害から発生することが知られている。大規模自然災害を発生源としたインフラサウンドは長距離伝播できる特性があるため、リモートセンシングによる防災・減災への活用が期待される。高知工科大学宇宙地球探査システム研究室では株式会社SAYAと複合型インフラサウンド津波センサーを共同開発し、高知県沿岸部に現在までに計16台のセンサーを設置し、到来方向探知をするための多重アレイ(面的)配置がなされている。アレイ配置されたセンサー群は特に頻繁に噴火を繰り返す桜島の火山噴火波形を同時観測しており、現在までに多くの火山噴火波形データを蓄積している。先行研究[1]では、大規模自然災害だけではなく、花火や屋内設置の場合にはドアの開閉などの様々な要因で発生するN形波形の自動検出システムが開発されているが、そのN形波形が発生した現象が何か特定できない。そのため、観測開始から現在までに多地点観測によって計5500回以上の桜島の火山噴火波形の到達が予測されることを利用し、教師あり学習の機械学習アルゴリズムを用いて火山噴火波形とノイズ波形を識別するシステムを提案した。本研究は同アルゴリズムを用いた大規模自然災害検出をトリガーとする自動警報システムの基礎開発を研究目的とする。

2. MFCCによる特徴量抽出とSVMによる2クラス分類

本研究ではサンプリングレート2 Hz、データ長1024 (約8.5分間)の波形データから特徴ベクトルを構成し、火山噴火波形とノイズ波形を識別するための2クラス分類を行った。特徴ベクトルはメル周波数ケプストラム係数(以下、MFCC)とデルタケプストラムの2種類の計24次元で構成した。MFCCは音声認識において声道成分を密、基本周波数成分を疎にとり、対数パワースペクトルの情報を圧縮する手法である。またデルタケプストラムは非定常な周波数特性を特徴量とする。特徴量抽出後、機械学習ライブラリの一つのscikit-learnを用いて2クラス線形識別関数の学習法であるサポートベクトルマシン(以下、SVM)で2クラス分類を行った。ただし、本研究の火山噴火波形データセットは線形分離可能であると仮定した。

3. 実験結果

本研究では教師あり学習のためのデータセットは16台のインフラサウンドセンサーそれぞれに構成した。そのうち、データセットのスコアが最も高いと判断された観測点は高知県黒潮町浮鞭にて蓄積された火山噴火波形データセットであり、正解率、再現率、適合率、F-値はそれぞれ95±2%、96±3%、93±3%、95±2%となった。ただし、高知県黒潮町浮鞭のデータセット数はN=58となった。これは火山噴火到達時刻にも関わらずノイズレベルが高かったことやデータの欠損により、火山噴火波形が入手できなかった等が主な要因である。データセット数の評価については、統計量が十分にある英単語の音声収録されたSpeech Commands Dataset(©TensorFlow team, AIY team, 2017)を用いて本研究に対する比較を行った。その結果、データセット数として少なくともN=146は必要であることが示唆された。上述のデータセット数N=58はその値に達しておらず、Nをさらに増やすことが課題となった。

4. 結論

本研究ではリモートセンシングによる火山噴火波形のリアルタイム検出の基礎研究を機械学習アルゴリズムを用いて行った。その際、特徴ベクトルはMFCCおよびデルタケプストラムの計24次元で構成し、2クラス分類のための機械学習アルゴリズムは線形SVMを選定した。その結果、最もスコアの高い高知県黒潮町浮鞭の

データセットにおける4種類の性能評価値はそれぞれ95%程度となり、良質なデータセットを作成する手法を確立し、リモートセンシングされた火山噴火波形およびノイズ波形との2クラス分類が可能となった。このときのデータセット数はN=58であったが、本研究では2016年8月から2018年12月までの桜島の火山噴火波形でデータセットを構築したため、2019年1月以降の火山噴火波形でさらにデータセット数を増加させることができる。現在も観測しているセンサー群のインフラサウンドデータに対して、今後は本研究の結果を用いたりリモートセンシングによる火山噴火の自動警報システムの構築を目指す。

参考文献

[1] 反町玲聖, インフラサウンドセンサの面的配置における波源位置推定方法の最適化, 高知工科大学2017年度修士課程卒業論文, 2018

キーワード：インフラサウンド、機械学習、MFCC

Keywords: infrasound, machine learning, MFCC

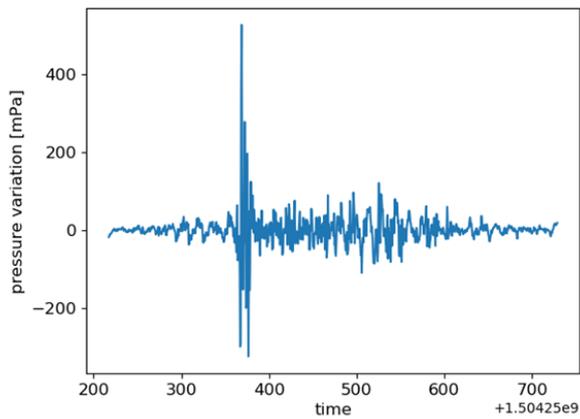


Fig. 1 One of the positive data in volcanic eruption waveform dataset

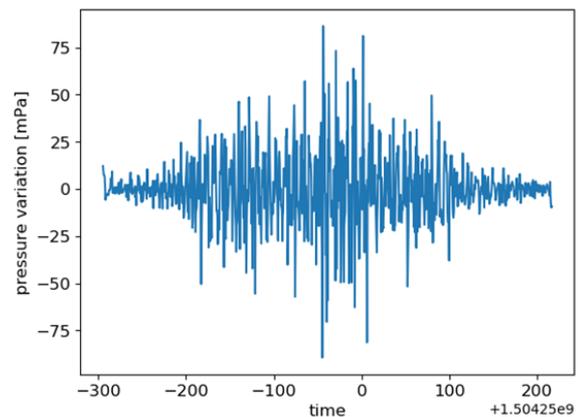


Fig. 2 One of the negative data in volcanic eruption waveform dataset