



## Automatic Recognition of the Lightning Whistler waves from the Wave Data of SCM Boarded on ZH-1 satellite

Jing Yuan<sup>1</sup>, Zijie Wang<sup>1</sup>, Dehe Yang<sup>2</sup>, Qiao Wang<sup>2</sup>, Zeren Zima<sup>2</sup>, Ying Han<sup>1</sup>, Le Zhou<sup>1</sup>, Xuhui Shen<sup>2</sup>, and Qihang Guo<sup>1</sup>

<sup>1</sup>Institute of Disaster Prevention, Department of Information Engineering, China (yuanjing20110824@sina.com)

<sup>2</sup>National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing, China  
(yuanjing20110824@sina.com)

Lightning whistlers, found frequently in electromagnetic satellite observation, are the important tool to study electromagnetic environment of the earth space. With the increasing data from electromagnetic satellites, a considerable amount of time and human efforts are needed to detect lightning whistlers from these tremendous data. In recent years, algorithms for lightning whistlers automatic detection have been conducted. However, these methods can only work in the time-frequency profile (image) of the electromagnetic satellites data with two major limitations: vast storage memory for the time-frequency profile (image) and expensive computation for employing the methods to detect automatically the whistler from the time-frequency profile. These limitations hinder the methods work efficiently on ZH-1 satellite. To overcome the limitations and realize the real-time whistler detection automatically on board satellite, we propose a novel algorithm for detecting lightning whistler from the original observed data without transforming it to the time-frequency profile (image).

The motivation is that the frequency of lightning whistler is in the audio frequency range. It encourages us to utilize the speech recognition techniques to recognize the whistler in the original data \of SCM VLF Boarded on ZH-1. Firstly, we averagely move a 0.16 seconds window on the original data to obtain the patch data as the audio clip. Secondly, we extract the Mel-frequency cepstral coefficients (MFCCs) of the patch data as a type of cepstral representation of the audio clip. Thirdly, the MFCCs are input to the Long Short-Term Memory (LSTM) recurrent neural networks to classification. To evaluate the proposed method, we construct the dataset composed of 10000 segments of SCM wave data observed from ZH-1 satellite(5000 segments which involving whistler and 5000 segments without any whistler). The proposed method can achieve 84% accuracy, 87% in recall, 85.6% in F1 score. Furthermore, it can save more than 126.7MB and 0.82 seconds compared to the method employing the YOLOv3 neural network for detecting whistler on each time-frequency profile.

Key words: ZH-1 satellite, SCM, lightning whistler, MFCC, LSTM

