Large-scale matched-filter analysis with similarity search

*Naoto Mizuno[1], Satoshi Ide[1]

1. Department of Earth and Planetary Science, The University of Tokyo

The matched-filter analysis (Gibbons and Ringdal, 2006) is one of the essential tools in modern seismology. It detects events by calculating the cross-correlation between the template waveforms and the observed waveforms. In both ordinal and low-frequency earthquakes, this algorithm improves the number of detections of small events (e.g., Shelly et al., 2007a; Peng and Zhao, 2009; Kato et al., 2012; Ross et al., 2019). The matched-filter analysis is based on the assumption that the cross-correlation of two events should be high when their locations and mechanisms are similar. Because of this assumption, the detections are limited to similar events to the templates. To prevent missing events, we need to use a large number of templates to cover an analysis region. Ross et al. (2019) used nearly 284,000 earthquakes listed in the Southern California Seismic Network catalog as templates to detect small earthquakes and created a new catalog, which is almost complete for earthquakes of $M > 0.3$. Previous studies used parallel processing to manage massive computation, which is efficient to perform template matching (e.g., Chamberlain et al., 2018; Beaucé et al., 2018; Senobari et al., 2019; Zimmerman et al., 2019; Liu et al., 2020). Even with these improvements in algorithms, large-scale computation has been necessary.

In this study, we develop a new algorithm for the matched-filter analysis with a large number of template waveforms. Instead of using massive computational resources, our method improves the computation complexity with similarity search and efficiently handles over 100,000 templates by a single workstation. We reduce the formulation of the matched-filter analysis to the nearest neighbor search and introduce approximations to reduce the computational cost. Our algorithm is efficient for parallel processing on the graphics processing unit and manages the increase of templates using the approximations. The computation speed is remarkably improved and achieves over 100 times faster than the ordinary matched-filter analysis.

Keywords: Matched-filter analysis, Similarity search, Machine learning

©2020. Japan Geoscience Union. All Right Reserved.