Geophysical Research Abstracts
Vol. 14, EGU2012-6821, 2012
EGU General Assembly 2012

The ear, the eye, earthquakes and feature selection: listening to automatically generated seismic bulletins for clues as to the differences between true and false events.

H.A. Kuzma (1), E. Arehart (1), J.N. Louie (2), and J.L. Witzleben (3)

(1) East Donner Research, Truckee, CA, USA, (2) Nevada Seismological Lab, Univ. of Nevada, Reno, NV, USA, (3) School of Music, University of Maryland, College Park, MD, USA

Listening to the waveforms generated by earthquakes is not new. The recordings of seismometers have been sped up and played to generations of introductory seismology students, published on educational websites and even included in the occasional symphony. The modern twist on earthquakes as music is an interest in using state-of-the-art computer algorithms for seismic data processing and evaluation. Algorithms such as such as Hidden Markov Models, Bayesian Network models and Support Vector Machines have been highly developed for applications in speech recognition, and might also be adapted for automatic seismic data analysis. Over the last three years, the International Data Centre (IDC) of the Comprehensive Test Ban Treaty Organization (CTBTO) has supported an effort to apply computer learning and data mining algorithms to IDC data processing, particularly to the problem of weeding through automatically generated event bulletins to find events which are non-physical and would otherwise have to be eliminated by the hand of highly trained human analysts. Analysts are able to evaluate events, distinguish between phases, pick new phases and build new events by looking at waveforms displayed on a computer screen. Human ears, however, are much better suited to waveform processing than are the eyes. Our hypothesis is that combining an auditory representation of seismic events with visual waveforms would reduce the time it takes to train an analyst and the time they need to evaluate an event. Since it takes almost two years for a person of extraordinary diligence to become a professional analyst and IDC contracts are limited to seven years by Treaty, faster training would significantly improve IDC operations. Furthermore, once a person learns to distinguish between true and false events by ear, various forms of audio compression can be applied to the data. The compression scheme which yields the smallest data set in which relevant signals can still be heard is likely an excellent candidate from which to draw features that can be fed into machine learning algorithms since it contains a compact numerical representation of the information that humans need to evaluate events. The challenge in this work is that, although it is relatively easy to pick out earthquake arrivals in waveform data from a single station, when stations are combined the addition of background noise tends to confuse and overwhelm the listener. To solve this problem, we rely on techniques such as the slowing down of recordings without altering the pitch which are used by ethnomusicologists to understand highly complex rhythms and sounds. We work with professional musicians and recorders to mix the data from different seismic stations in a way which reduces noise and preserves the uniqueness of each station.