AN INTEGRATED REAL TIME SEISMIC SIGNAL PROCESSOR

N. Magotra, J. W. McCoy, L. Tan, J. Jiang

TEL: (505) 277-0808, FAX: (505) 277-1413, E-mail: magotra@houdini.eece.unm.edu

Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, New Mexico 87131

Abstract

This paper presents the results of research involving the development of techniques for detecting and locating the sources of seismic events as well as discriminating phases in the coda of these events. The work represents a continuation of the research effort to detect, locate and discriminate regional seismic events. These algorithms were originally developed on SUN workstations and are currently being implemented on a real-time system based on Texas Instrument's TMS320C30 DSP chip. The original motivation for this research was the idea of using seismic information to verify compliance with respect to a Comprehensive Test Ban Treaty. However, the current world political situation means that these algorithms could be used for monitoring nonproliferation treaties. The signal processing algorithms developed in this research can also be applied in the analysis of seismic data in general.

1. Introduction

Since the original motivation for this research was the idea of using seismic information to verify compliance with respect to a Comprehensive Test Ban Treaty, it is quite possible that countries party to such a treaty may not permit large spatially distributed arrays to be set up within their respective territories. Hence, one concept which has been proposed involves the deployment of a small number of regionally spaced three-channel seismic stations within countries party to such a treaty. These stations could be used to monitor regional seismic events. The data gathered at these unattended stations would be transmitted back to monitoring countries for analysis. The individual stations forming such a network would be designed to acquire seismic data along three orthogonal directions.

Some of the prominent phases observed in the coda of regional events (epicentral distance, $\Delta \leq 20$') [1] are $P_s$ phase, $P_a$ phase, $S_n$ phase, $L_n$ phase. Exactly which of the phases described above are observable on the seismogram is often determined by the nature of the seismic source and the geology of the source-to-receiver path. However, the order of appearance on the seismogram is always $P_n$ followed by $P_s$ then $S_n$ and finally $L_n$. It is also important to note that the phase arrivals are more spread out for more distant events.

The field-acquired data used in this research consisted of three-channel data obtained from three orthogonally placed seismometers - Vertical ($z(m)$), North ($n(m)$) and East ($e(m)$), provided by Sandia National Laboratories (SNL), Albuquerque, New Mexico. The seismic data were gathered at three regional sites located in the southwestern United States. The original data sampling rate was 200 samples per second (sps). Based on a spectral and polarization analysis of the signals the data were low-pass filtered with a cut-off frequency of 15 Hz using a fourth order Butterworth filter and resampled to 40 sps.

2. Algorithm Development

The tasks performed by the seismic signal processor have been outlined in the flowgraph shown in Figure 1. The first task is that of event detection. Once an event is detected, a correlated energy level-threshold is set and as long as this threshold is exceeded it is assumed that the same event is on. During this time the detector switches to a phase detection mode. As each phase is detected the phase classifier is activated to identify the $P_s$ and $L_n$ phases. If the initial detection is correctly identified as a $P_n$ $P_n$ phase, then the three-component data during that phase is used to perform event classification (explosion...
or earthquake) and source bearing estimation. Finally the time difference between the $P_s$ and $L_p$ phases is converted (via standardised look-up tables) to a source to receiver distance estimate thus completing the location task.

### Event Detection

To detect events, three-component time series data from the three channels are combined to provide a measure of polarized energy in the data. We have shown that from the point of view of detecting low signal-to-noise ratio signals it is preferable to use a three-component detection parameter $\delta$. The detection parameter used to obtain the results presented in this paper is based on the estimated autocorrelation coefficients $\hat{R}_{xx}(m,l)$ of each channel $s(m)$. Note that $c(m)$ can be set equal to the channel $z(m)$ whose autocorrelation function needs to be estimated. The autocorrelation coefficients are updated recursively as shown below

$$\hat{R}_{xx}(m,l) = \beta \hat{R}_{xx}(m-1,l) + (1-\beta) c(m)c(m+l)$$  \hspace{1cm} (1)

where 'm' is the time index, 'l' is the lag index in the time window $[L, -L]$ and $0<\beta<1$ is the smoothing constant.

Equation (1) represents a recursive estimator which corresponds to sliding an exponential window over the data with a time constant $\tau = 1/((1-\beta) f_s)$ where $f_s$ represents the sampling frequency (sps).

The detection parameter $s(m)$ is then defined as,

$$s(m) = \sum_{l=-L}^{L} (\hat{R}_{xx}^2(m,l) + \hat{R}_{xx}^2(m,l) + \hat{R}_{xx}^2(m,l)) \quad 1\neq 0$$  \hspace{1cm} (2)

The signal $s(m)$ is then passed through a sliding window detector implemented via the following three equations.

$$w_1(m) = \beta_1 * w_1(m-1) + (1-\beta_1) * s(m)$$  \hspace{1cm} (3a)

$$w_2(m) = \beta_2 * w_2(m-1) + (1-\beta_2) * s(m-\Delta)$$  \hspace{1cm} (3b)

$$d(m) = w_1(m) - k_s w_2(m) > 0.0$$

In the equations above, $\beta_1$ and $\beta_2$ are chosen so that $w_1(m)$ represents a short-time average of $s(m)$ and $w_2(m)$ represents a delayed long-time average of $s(m)$. The constant $k_s$ represents a threshold setting parameter. The signal $d(m)$ results from the comparison of $w_1(m)$ with its past history represented by $w_2(m)$ to look for a sudden increase in the correlated energy level in the input signal $s(m)$, indicating the presence of an event of interest.

### Phase Discrimination

As each phase in the coda of the event is detected, it is processed through the phase classifier.

In designing a classifier, the first critical task is to make a decision regarding the features to be used as discriminants. The features that we chose were coherence functions derived from the 3x3 data covariance matrix. Five different coherence functions were defined using the eigenvalues of the 3x3 $(Z, E, N)$ data covariance matrix as well its diagonal elements. These functions are listed below:

$$1 - (\lambda_{max} + \lambda_{mid})/(2 \lambda_{min})$$  \hspace{1cm} (4a)

$$1 - (2\lambda_{min})/(\lambda_{mid} + \lambda_{max})$$  \hspace{1cm} (4b)

$$1 - (\lambda_{min}/\lambda_{max})$$  \hspace{1cm} (4c)

$$1 - (\lambda_{med}/\lambda_{max})$$  \hspace{1cm} (4d)

and

$$\sigma_N^2/(\sigma_E^2 + \sigma_N^2 + \sigma_Z^2)$$  \hspace{1cm} (4e)

where $\lambda_{max}$, $\lambda_{med}$, and $\lambda_{min}$ are the maximum, medium, and minimum eigenvalues of the covariance matrix and $\sigma_E^2$, $\sigma_N^2$, $\sigma_Z^2$ are the variances of the $E$, $N$, and $Z$ channel data respectively. The discriminant set consisted of five numbers obtained by averaging each of the five coherence functions over a 4 second window.

In estimating the covariance matrix of the $Z$, $N$ and $E$ data channels, the data were assumed to be zero mean. The estimated covariance matrix is defined by

$$\hat{C}_{zz}(m) = \begin{pmatrix}$\mathbf{d}_1(m)$ & $\mathbf{d}_2(m)$ & $\mathbf{d}_3(m)$ \\
$\mathbf{d}_4(m)$ & $\mathbf{d}_5(m)$ & $\mathbf{d}_6(m)$ \\
$\mathbf{d}_7(m)$ & $\mathbf{d}_8(m)$ & $\mathbf{d}_9(m)$ \end{pmatrix}$$  \hspace{1cm} (5)
where the elements of $\hat{C}(m)$ were obtained recursively using equations similar to equation (1).

To perform the classification itself, we chose to use a neural net classifier (NNC) since the complex nature of the data seemed to indicate that it would not reside in geometrically shaped regions which could be accurately represented by a small number of prototypes, but could be spread out in the pattern space such that some sort of simple nonlinear decision boundary could more easily separate the classes.

Instead of starting with a relatively large neural net classifier (NNC) we chose to start with the simplest possible NNC, analyze its performance and increase its complexity only if the results warranted such an increase. In keeping with this philosophy we started with a single-layer NNC, capable only of creating a linear decision boundary. Our results indicated that a linear decision boundary did not suffice, hence we switched to a two-layer NNC (two nodes in "input layer and one node in output"). In both cases we used the back-propagation algorithm to update the NNC weights.

**Event Discrimination**

Normally, for regional seismic events, the initial detection should correspond to the $P_s$ phase. Once this phase has been identified, the estimated autocorrelation function of the Z channel is used to perform event classification (explosion versus earthquake). The autocorrelation functions are estimated over a 4 second window straddling the onset of the $P_s$ phases. The autocorrelation estimation is performed by using a recursive estimator similar to the one defined in equation (1).

The idea of using correlation coefficients was based upon the fact that the frequency content of seismic signals varies depending on the type of source. This frequency difference shows up in the power density spectrum (PDS) of the signals and by the Wiener-Khinchin theorem, also appears through the inverse Fourier transform of the PDS in the correlation function of the signals.

To perform the classification, we once again used a NNC with 2 layers (two nodes in the input layer and one node in the output layer) utilizing the back propagation algorithm for weight updates. Note that this neural net is not the same as the one used to perform phase classification.

**Source Location**

The source location task consists of two sub-tasks - bearing estimation and source-to-receiver distance estimation. The bearing estimate is obtained from the eigenvector corresponding to the maximum eigenvalue of the 2x2 N and E channel data co-variance matrix. The direction of this vector ($\phi(m)$), during the onset of the $P_s$ phase, yields an estimate of the line of travel of the propagating signal, while the signs of $\hat{\delta}_{NW}(m)$, $\hat{\delta}_{SE}(m)$, $\hat{\delta}_{NE}(m)$, resolve the quadrant ambiguity as shown in Table 1 below.

<table>
<thead>
<tr>
<th>$\delta_{NW}(m)$</th>
<th>$\delta_{SE}(m)$</th>
<th>$\delta_{NE}(m)$</th>
<th>Quadrant</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>SW</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>NW</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>NE</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>SE</td>
</tr>
</tbody>
</table>

Finally, the source-to-receiver distance estimate is obtained by converting the difference in the arrival times of the $P_s$ and $L_s$ phases to a distance estimate via standard look-up tables [1].

**3. Results and Conclusions**

The algorithms described have been tested extensively using field acquired data on Sun/Dec stations and fixed point real-time systems. Currently they are being implemented on a TMS 320C30 floating-point digital signal processor manufactured by Texas Instruments.

Figure 2 displays some plots relating to the detection and location algorithms. The top trace in the figure displays the Z channel data for a magnitude 3.2 earthquake in Southern California, the next two traces in the figure display $d(m)$ (equation (3c)) and the detector on/off output respectively. The lower most trace displays $\phi(m)$. Note that this azimuth estimate tends to stabilize during the onset of the $P_s$ and $L_s$ phases. The detection and location algorithm was able to locate sources to within 50 km of their correct location for a test data set where the source-to-receiver distance was 1200 km.

The algorithm used to perform event discrimination has been described in detail in [4]. It was able to classify the training set with a 100% success rate and was able to classify 87% of a test data set correctly. Figure 3 shows 3D plots of the discriminants (autocorrelation functions) for an earthquake.
and a nuclear explosion.

The phase discrimination algorithm described above also tested out successfully. Using the covariance matrix parameters as discriminants we achieved 96% correct classification with both, the training and test, data sets. Figure 4 shows a scatter plot for the data set indicating a fairly clear delineation between the $P_a$ and $L_g$ phases.

A block diagram of the experimental set-up linking the real-time system to the workstations is shown in Figure 5.

At the present time we are continuing the process of refining the algorithms and increasing the size of the test data set.

References


