Multi-window algorithm for detecting seismic first arrivals

Zuolin Chen and Robert Stewart, CREWES, University of Calgary, Canada

2005 CSEG National Convention

Abstract

A robust and accurate multi-window algorithm for detecting the first arrivals of impulsive P phases of seismic events in low *SNR* environments has been developed. Similar to the conventional time-domain *STA/LTA* algorithm, this method acquires the averages of absolute amplitudes from a seismic trace by using three moving time windows before and after each time point (sample). When the instantaneous absolute amplitude exceeds an automatically adjusted threshold, ratios based on the averages of the windows over previous time samples provide standards to differentiate an expected event from unwanted noise. This algorithm has been examined on synthetic and real data sets with varying *SNR*. Comparison with published algorithms shows that this method has superior accuracy and is more tolerant of high background noise. The accuracy of the picking of the impulsive P phase can be as accurate as 1-2 samples even when the *SNR* is lower than 2.

Introduction

Accurate automatic detection of the first arrival of the P phase is of considerable importance for the rapid identification and location of seismic events in large volumes of digital and real-time passive seismic data. Generally, picking of the first arrival of the P phase is carried out in the time-domain (e.g. Allen, 1982; Boschetti et al., 1996; Der & Shumway, 1999). Among the approaches in the time-domain, the short-term average over long-term average ratio (*STA/LTA*) automatic picker is well-known. The absolute value, the square, or the envelope of the seismic trace is usually selected as the characteristic function for the calculation of *STA* and *LTA* (e.g., Kanasewich, 1981; Coppens, 1985). The basic idea of the algorithm detects an event when the *STA/LTA* ratio exceeds a pre-defined threshold. This method is well suited for the arrival detection of events, but not for providing accurate arrival times due to the long delay associated with the length of the *STA* window.

In this paper, we propose a modified algorithm which very effectively picks first arrivals at various levels of noise background, based on the instantaneous absolute amplitude of seismic trace and its averages over three moving time windows. The procedure of our multi-window algorithm resembles that of a human operator in detecting P arrivals of events and is comparable in accuracy to the latter.

Methodology

Our multi-window automatic P phase picker operates in the time-domain. It includes procedures for defining time windows, standards, corresponding thresholds and waveform correction for the acquisition of a more accurate corrected onset time.

Windows and standards

The definition of the three moving time windows is mainly based on measuring the averages of the signal/noise level (*SNR*) within certain time extents before, after and after a delay of an instantaneous time point (sample). The averages of absolute amplitudes within *BTA* (Before Term Average), *ATA* (After Term Average) and *DTA* (Delayed Term Average) windows are respectively defined as follows:

$$\overline{BTA(t)} = \sum_{i=1}^{m} | u(t-i) | \Big/ m \qquad (1)$$

$$\overline{ATA(t)} = \sum_{j=1}^{n} | u(t+j) | \Big/ n \qquad (2)$$

$$\overline{DTA(t)} = \sum_{k=1}^{q} | u(t+j+d) | \Big/ q \qquad (3)$$

Here, $u(t)$ represents the amplitude of seismic trace at a time point t; m, n, and q represent the lengths of the windows in samples respectively; d is the time delay for a *DTA* window.

The instantaneous absolute amplitude $|u(t)|$ is selected as the first standard to judge the arrival of possible high amplitude event. Two extraordinary standards $R_2(t)$ and $R_3(t)$ are introduced based on the ratios of averages of *ATA, DTA* over *BTA* and take the forms of

$$R_2(t) = \overline{ATA(t)} / \overline{BAT(t)} \qquad (4)$$

$$R_3(t) = \overline{DTA(t)} / \overline{BAT(t)} \qquad (5)$$

Standards $R_2(t)$ and $R_3(t)$ are mainly designed for the discrimination of high-amplitude short-duration and long-duration noise respectively. The working functions of $R_2(t)$ and $R_3(t)$ are based on the following rationale:

When the first threshold, $H_1(t)$, is exceeded at a time point t, the second standard, $R_2(t)$, is used to separate a high-amplitude short-duration noise from a high-amplitude long-duration event. From (4), it can be seen that when the *ATA* window contains noise, and the length of the window is several times longer than the noise, $R_2(t)$ will only be limitedly affected by the noise; however, when the *ATA* window contains only a part of an event, then $R_2(t)$ will be much amplified. Based on this feature of $R_2(t)$, long duration noise will be removed. When the duration of high-amplitude noise is comparable to the length of the *ATA* window, $R_2(t)$ will lose its ability to distinguish an event from noise. Equation (5) makes it obvious that an appropriate delay time will reduce $R_3(t)$ to a required level. Therefore, false triggers caused by high-amplitude long-duration noise can be avoided.

Thresholds

Thresholds pertaining to the standards can be pre-defined based on the expected *SNR*. An optimum first threshold, $H_1(t)$, should be larger than the fluctuations of most noise but lower than an expected signal. This demand can be satisfied by measuring the mean (E_m) and standard deviation (E_{sd}) of the envelope ($E(t)$) of pre-existing noise within a *BTA* window. The envelope $E(t)$ can be acquired from the absolute value of the Hilbert transform of a seismic trace $u(t)$. In practice, $H_1(t)$ is often shifted several samples backwards to avoid the unnecessarily early rising of $H_1(t)$ caused by the first arrival. In this way, an instantaneous $H_1(t)$ is expressed as

$$H_1(t) = E_m(t-p) + \alpha E_{sd}(t-p). \qquad (6)$$

Here, p is the number of shifted samples; α is the coefficient to adjust the height of the first threshold and is taken to be 3, whereby approximately 99% of the noise will not exceed $H_1(t)$. From equation (6) it is obvious that $H_1(t)$ is automatically adjusted with the variance of the background noise.

Waveform correction

As $H_1(t)$ is defined larger than most pre-existing noise levels, and the instantaneous absolute amplitude at the trigger time point is higher than $H_1(t)$, the trigger time point must be somehow later than the real onset time according to the configuration of the first arrival of an event. This belated onset time can be compensated by a waveform correction. For an impulsive first arrival, this process is easily accomplished by using the height of the absolute amplitude and the representative gradient at the trigger point.

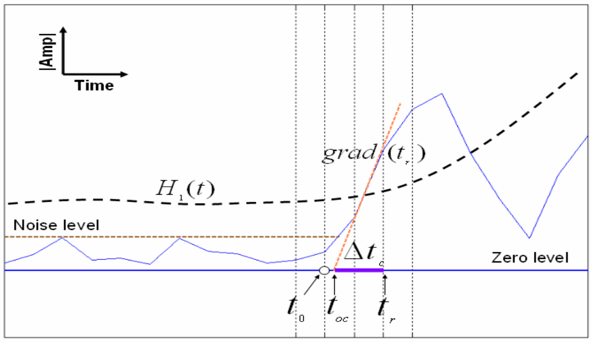

Figure 1. Schematic diagram showing the items used in the waveform correction at trigger point (***t_r***) for the corrected onset time (***t_{0c}***) of an event. The thick dashed curve represents the first threshold ***$H_1(t)$***; the red dashed oblique line represents the adapted gradient; the thick purple line marks the calculated amount of waveform correction (***Δt***); the open circle shows the real onset time (***t_0***). Vertical dotted lines mark the sample points adjacent to trigger point ***t_r*** with an interval of one sample.

Examination on synthetic and real data

The performance in accuracy and tolerance of our algorithm is firstly examined on a series of synthetic data with various *SNR*. Basically, the synthetic data is composed of three parts: ramdom noise with various gains, a high-amplitude noise and an impulsive sine wave seismic event with the onset time of the first arrival at a fixed time point. Ramdom noise still superimposes with the event after its arrival.

Figure.2 examines the accuracy of our algorithm on a synthetic seismic trace with a 20 Hz dominant frequency event. The maximum amplitude of the event and high-amplitude noise are assumed to be 1. The amplitude coefficient of random noise is assumed to be 0.25. From this figure, it is noted that the three standards at the onset time of the event exceeded their corresponding pre-defined thresholds at time point 402, and the event is triggered there. Waveform correction modifies the trigger point t_r to the corrected onset time t_{oc} of 400.91, which is only 0.09 samples earlier than the real onset time at time point 401.

The performance in accuracy and tolerance is further examined in the case of various background noise levels. As the noise part is computer-generated randomly, the time series used in the examination is slightly different each time. The examination of each noise levels is repeated 100 times to acquire a stable statistic of errors of detection between real and corrected onset times.

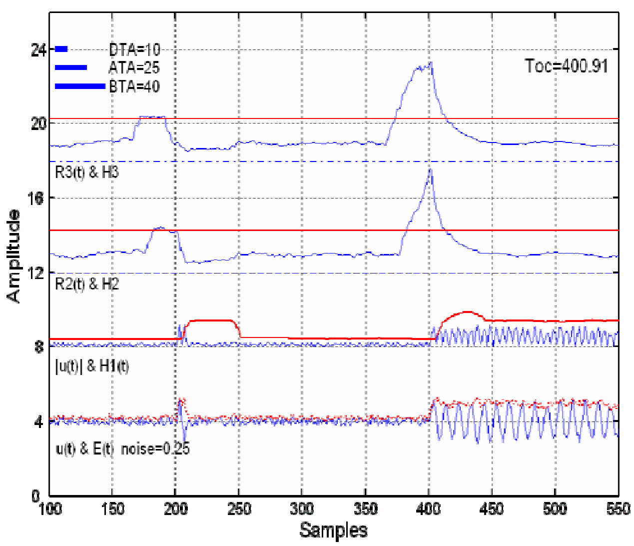

Figure 2. An example illustrates the picking of the first arrival of an event in a synthetic seismogram. Diagrams from the bottom to the top are synthetic seismogram (blue) and envelope (red dotted); absolute amplitude of seismic trace ***u(t)*** (blue) and automatically adjusted threshold ***$H_1(t)$*** (red thick); second standard ***$R_2(t)$*** (blue curve) and threshold ***H_2*** (red); third standard ***$R_3(t)$*** (blue curve) and threshold ***H_3*** (red). The dash-dotted line in each diagram represents the zero reference levels. The lengths of moving time windows are shown in the top left corner. The real onset time of an event is set at sample point 401. The corrected onset time of the event is shown in the top right corner.

Figure 3 shows the statistical results of accuracy and tolerance of our algorrithm with the increase of the amplitude of random noise, where it can be seen that the maximum onset time error ranges from -1 to +1.25 samples. When the coefficient of random noise is increased to 0.4, false triggers due to the existence of high-amplitude noise occur.

The scatter of the corrected onset time increases with noise levels. This is considered to be caused by the disturbance of the waveform of first arrivals by the superimposed random noise. The disturbance of the waveform thus reduces the reliability of the two crucial factors of height and representative gradient adopted in the waveform correction at the trigger point.

The performance of our automatic P phase picker is further examined on some real local events recorded by a small six-station vertical-component seismic array deployed in southern Alberta. (Bingham,1996). Results of examinations show that the picking accuracy of real events with impulsive first arrivals can be stably picked as accurate as less than 1-2 samples comparing to the pickings by a human operator.

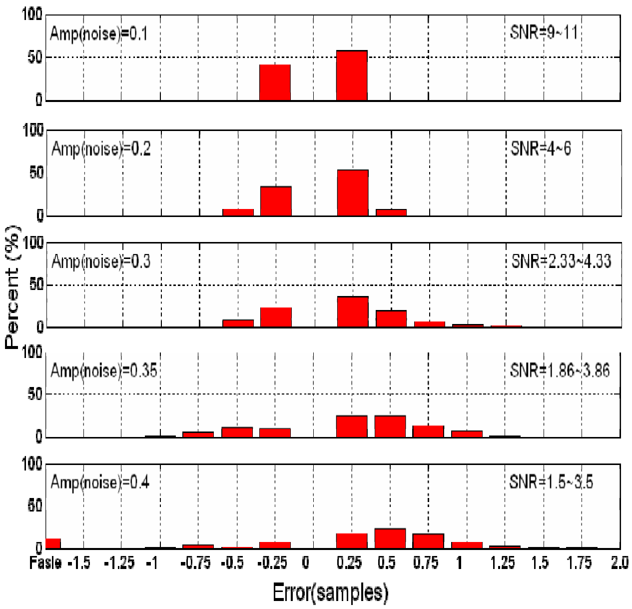

Figure 3. Error distribution of corrected onset times with different noise levels. For each diagram, picking is repeated 100 times to get stable statistics. Histograms indicate the percentage of events with corrected onset time errors within the indicated amounts in the horizontal axis. The ***SNR*** extent is estimated by equation $SNR=(Amp_{event}\pm Amp_{noise})/Amp_{noise}$, where Amp_{event} and Amp_{noise} represent the maximum amplitudes of event and random noise respectively. The threshold $H_1(t)$ is pre-defined by setting the coefficient α to 3. The lengths of ***BTA, ATA and DTA*** are set to be 40, 30 and 30 samples respectively. The time delay of ***DAT*** window is set to 10 samples. H_2, H_3 are set to 0.75 times the average ***SNR***.

Conclusions

We have developed an effective and accurate multi-window algorithm for picking onset times of the seismic first arrivals of impulsive events with accuracy comparable to that achieved by a human operator. The performance in accuracy is less than 1 sample in most cases of an impulsive first arrival. This accuracy is better than that obtained from any available methods.

Our method is more tolerant to low *SNR* than any previous methods. It works accurately and reliably even when the *SNR* ratio is below 2.

The manupulation of parameters such as thresholds and time delay interval, is easy and brief to be mastered. After the limitation of expected *SNR* and duration of impulsive noise to be discarded is indicated, all parameters could be defined directly.

Due to the the efficiency and simplicity of the method, our algorithm is widely applicable in passive microseimic monitoring in real time with only moderate computing capability.

References

Allen, R., 1982, Automatic phase-pickers: their present use and future prospects: Bull. Seism. Soc. Am., 72, S225–S242.

Bingham, D.K., 1996, Seismic monitoring of Turtle Mountain: Internal report, Alberta Environmental Protection, Government of Alberta.

Boschetti, F., Dentith, M. & List, R., 1996, A fractal based algorithm for detecting first arrivals on seismic traces: Geophysics, 61, 1095–1102.

Coppens, F., 1985, First arrival picking on common-offset trace collections for automatic estimation of static correction: Geophysical Prospecting 33, 1212-1231.

Kanasewich, E. R, 1981, Time Sequence Analysis in Geophysics, Univ. of Alberta Press, Edmonton.