First arrival time picking for microseismic data based on shearlet transform

To cite this article: Yao Cheng et al 2017 J. Geophys. Eng. 14 262

View the article online for updates and enhancements.

Related content
- SNR enhancement for downhole microseismic data based on scale classification shearlet transform
  Juan Li, Shuo Ji, Yue Li et al.
- Downhole microseismic monitoring for low signal-to-noise ratio events
  Hang Zhou, Wei Zhang and Jie Zhang
- Noise reduction by support vector regression with a Ricker wavelet kernel
  Xiaoying Deng, Dinghui Yang and Jing Xie
First arrival time picking for microseismic data based on shearlet transform

Yao Cheng, Yue Li and Chao Zhang

College of Communication Engineering, Jilin University, Changchun, Jilin, People’s Republic of China

E-mail: cyy2855@163.com, liyue@jlu.edu.cn and 570918942@qq.com

Received 21 June 2016, revised 24 November 2016
Accepted for publication 6 January 2017
Published 13 February 2017

Abstract

Automatic identification and first arrival time picking of microseismic data play an important role in microseismic monitoring technology, and it is the precondition for real-time microseismic hypocenter location. This paper presents a novel first arrival time picking method based on shearlet transform (ST), which aims to get satisfactory results in low signal-to-noise ratio data. The ST is used to decompose noisy microseismic data. By the coefficient differences between the signal and noise at fine scales, the signal points can be preliminarily selected from the noise. To further improve the accuracy of the signal recognition, a correction to the selected signal points is made by utilizing the scale correlation between adjacent scales. The realization of the correction depends on the distances between the signal points at one scale and those at its adjacent scale. After the correction, the moment of the first identified signal point is the first arrival time. This proposed method can produce a superior performance in the accuracy of the first arrival time picking, compared with the other methods, as demonstrated using synthetic and field microseismic data. The actual picking performance of the method is further verified by receiver operating characteristic curves.

Keywords: first arrival, time picking, shearlet transform, scale correlation, fine scale

1. Introduction

Microseismic monitoring as an unconventional oil and gas exploitation technology has been developing rapidly in recent years. In the data processing, a critical step is the first arrival time picking (Maxwell et al 2010). One of the main research directions is efficient auto-picking, because there is a large amount of microseismic data, which is also recorded by long duration.

Various methods for first arrival time picking have been put forward. The short-term average and long-term average (STA/LTA) method is widely used. This method’s first arrival is determined by the difference in energy characteristics between the short window and long window before and after the first break (Ambuter and Solomon 1974, Allen 1978, Earle and Shearer 1994). The accuracy of the STA/LTA method is greatly influenced by the window lengths (Akram and Eaton 2016). Another common method is the Akaike information criterion (AIC), which makes use of the different statistical properties between the microseismic signal and the noise (Leonard 2000, Zhang et al 2003). The AIC values of the time series are calculated, and the point with the minimum AIC value is chosen as the first arrival (St-Onge 2011). However, the picking results are unsatisfactory at low signal-to-noise ratio (SNR). In addition, artificial neural network (McCormack et al 1993, Dai and MacBeth 1995) and fractal dimension (Boschetti et al 1996) have been introduced, but they are complex and inefficient.

We propose a novel first arrival time picking method that adapted to microseismic data with low SNR. The noisy records are decomposed by shearlet transform (ST), and we can distinguish the signal points from the noise, because of the larger amplitudes of the signals at fine scales. To increase the recognition accuracy, the distances between the signal points at one fine scale and those at its adjacent scale are calculated. When the distances are all larger than a specific...
value, the points should be classified as noise points according to the scale correlation. After the above process, we regard the first signal point as the first arrival.

We demonstrate the availability of this method through the experiments on synthetic and real microseismic data. Moreover, receiver operating characteristic (ROC) curves are used to analyze the picking performance (McNeil and Hanley 1984, Yonelinas and Parks 2007).

2. First arrival time picking based on ST

ST is a multidirectional multiscale transformation, which is used for decomposing the microseismic record into various frequency scales and directions. Since microseismic data are a kind of high-frequency signal, a lot of effective information of the signal is expressed at fine scales. Hence, we only use fine scales to complete the recognition of the signal.

The signal and noise at fine scales present different characteristics. Signals are mostly concentrated on a minority of coefficients with large amplitudes, while the noise has decentralized coefficients with small amplitudes. Each scale has multiple directions, the energy of different directional subbands are distinct, and we can calculate it by

\[ E(j, k) = \sum_n \mu(j, k, n)^2, \]

where \( E(j, k) \) is the energy of a subband with \( j \)th scale and \( k \)th direction, \( \mu_{j,k,n} \) is the shearlet coefficient, \( j, k, \) and \( n \) denote the scale, direction, and location, respectively. The directional subband with large energy indicates that the signal is mainly focused in this direction. As the scale becomes finer, more energy is generated in the signal directions. In other words, there is a strong correlation between the same direction of adjacent scales (Zhao et al. 2016). Finally, we use the correlation energy to get the directions of the signal. The correlation energy between adjacent scales can be described as

\[ E_\text{Corr}_{j,j+1}(k) = E(j, k)E(j + 1, k). \]

Then, we add the subband coefficients of the signal directions together and regard them as the coefficients of the fine scale. The coefficient matrix of the fine scale is denoted as \( \text{Coef} = \{ \mu_{j}(m, n) \}, m = 1, 2, \ldots N, n = 1, 2, \ldots N. \)

Because of different coefficient characteristics of signal and noise at fine scales, we may use a threshold as a judgment standard to distinguish the signal from noise (Donoho 1995, Lin et al. 2015). We set the threshold as

\[ Th = \alpha \sigma_j \sqrt{\frac{2 \log(N)}{N}}, \]

where \( Th \) represents the threshold, \( \alpha \) is an adjustment factor, and different values are selected at different fine scales. \( N \) is the number of sample points, and \( \sigma_j \) denotes the standard deviation of the noise.

Since a few centralized coefficients of the signal can hardly affect the median of the noise, the standard deviation of the noise can be estimated as

\[ \sigma_j = \text{median}(\{|\mu_j(m, n)|\})/0.6745, \]

where the constant 0.6745 is derived from Donoho (1995).

The data point, whose absolute value of the coefficient is larger than \( Th \) is regarded as the signal point. Conversely, the data point, whose absolute value of the coefficient is lower than \( Th \) is regarded as the noise point.

The signal and noise have similar coefficient values in low SNR, so that some noise points are judged to be signal points through the above process, which will have an effect on the final first arrival time picking. Therefore, we take the scale correlation between adjacent scales into account. Signals have a correlation between adjacent scales, but the noise is irrelevant. If the distances between the signal points at one scale and those at its adjacent scale are larger than a specific value, the points should be regarded as noise. The correction eliminates the error signal points effectively and enhances the accuracy of the signal recognition, which improves the precision of the final first arrival time picking. After completing the above procedures, we take the first signal point as the first arrival.

Figure 1. (a) Pure Ricker wavelet. (b) Noisy Ricker wavelet. (c) Statistical results of arrival time picking under different specific values.
3. Parameter selection

Now, we give an explanation of the parameters selected in this paper.

3.1. Adjustment factor $\alpha$ of fine scales

We decompose the microseismic record for four scales and divide each scale into four directions, where the third and the fourth scales are fine scales, and they contain lots of useful information. The adjustment factor affects the retention of signal points at the two scales. To reserve signal points as much as possible, we select $\alpha = 0.5$ for the third scale. The coefficients of the signals become more significant when the scale becomes finer. Through lots of experiments, we found that the pre-judgment of the signal points can get satisfactory results when the adjustment factors of two adjacent scales

Figure 2. (a) Pure synthetic record. (b) Noisy record.

Figure 3. (a)–(d) Coefficients of the fine scale from the first to the fourth direction ($j = 4$). (e) The correlation energy between adjacent scales ($j = 3$ and $j = 4$). (f) The coefficients of the fine scale ($j = 4$).
meet the following equation:

$$a_4 = a_3 \times \sqrt{E_4/E_3},$$

(5)

where $a_3$ and $a_4$ are the adjustment factors of the third and the fourth scales, and $E_3$ and $E_4$ are the energy of the third and the fourth scales.

### 3.2. Specific value

After the pre-judgment of the signals, we make a correction for the selected signal points. The correction relies on the distances between the signal points at one scale and those at its adjacent scale. The points are regarded as noise points when the distances are all larger than a specific value. We can see that the accuracy of the signal recognition and the arrival time picking is affected by the selected value.

In order to determine the selection of the specific value, we establish a Ricker wavelet model (figure 1(a)) with a peak frequency of 300 Hz. This model trace contains 500 samples, and the sampling rate is 0.001 s. The real first arrival time is 0.247 s. The model is contaminated by white Gaussian noise (WGN) with the SNR of $-6$ dB. We define the SNR as:

$$SNR = 10\log_{10} \frac{\sum_i |s(t)|^2}{\sum_i |y(t) - s(t)|^2},$$

(6)

where $y(t)$ is the noisy record and $s(t)$ is the original record.

We carry out 1000 arrival time picking experiments. Different WGN is added in each test, and figure 1(b) gives one waveform data example after the addition of WGN. We make the specific value range from 1–8. The absolute errors between the automatic picks and the true arrivals are calculated, we give the percentage of the picks whose absolute errors are zero (picking with no error) and not more than two samplings (picking within 0.002 s), respectively, as shown in figure 1(c). It is observed that when the specific value is 4, the proposed method has the best picking performance. We picked about 75% of the arrival times correctly, among which 84% are within 0.002 s. Thus, the specific value is chosen as 4 in this paper.

### 4. Testing on synthetic record

In this section, to testify the availability of the method, we generate a synthetic record containing 24 channels as shown in figure 4. (a) The waveform in the third channel. (b) Shearlet coefficients of the third channel at the fine scale ($j = 4$).
Figure 2(a). The dominant frequency is 300 Hz and the sampling time is 0.001 s. Figure 2(b) is the noisy record added by WGN with the SNR of −6 dB, where the vertical components for the geophones are represented by the red lines, and the horizontal components by the blue lines.

The record is decomposed by ST for four scales and each scale is divided into four directions. Figures 3(a)–(d) give the coefficients in different directions of the fourth scale, respectively. We can observe that the first two directions contain abundant information of signals, whereas the last two are almost noise. The correlation energy between adjacent scales is shown in figure 3(e). Through the histogram, we see that the correlation energy in the first two directions is obviously larger than that in the other directions. This shows we can get accurate signal directions by the correlation energy. Then we add the coefficients of signal directions together and regard the results as the coefficients of the fourth scale, which are shown in figure 3(f). We deal with its adjacent scale \((j = 3)\) in the same way.

Figure 3(e) shows the correlation energy between adjacent scales. Through the histogram, we see that the correlation energy in the first two scales is obviously larger than that in the other scales. Thus, we can get accurate signal directions by the correlation energy. Then we add the coefficients of signal directions together and regard the results as the coefficients of the fourth scale, which are shown in figure 3(f). We deal with its adjacent scale \((j = 3)\) in the same way.

Figure 4(a) shows the waveform of the third channel. It can be seen that the signal is difficult to be identified. However, after the above process, the signal generates more significant coefficients than the noise at fine scales, which is shown in figure 4(b) (in ellipses). Thus, the signal points can be identified approximately by a threshold. The thresholds for fine scales are selected according to the introduction of the previous section.

Figure 5 shows the recognition results, which are marked with the symbol x. We can see that some noise points are judged as the signal points. We further correct the results in accordance with the scale correlation between adjacent scales. Red x represent the corrected signal points. The accuracy of signal recognition is improved by the correction. Finally, the first identified signal point is the first break. We pick the first arrival time for the record shown in figure 2(b).

Figure 6(a) gives a comparison with the STA/LTA and the AIC algorithms, and the picking results are denoted by the symbols in green, black, and magenta, respectively.
Figure 6(b) is the amplified view of the period 0.18–0.3 s, from which we can see that the arrival times received by the STA/LTA method are obviously wrong for some channels, such as the 3rd, 8th, 12th, and 21st channels. The AIC method performs unsatisfactorily when the arrivals are not evident. It is sensitive to noise and provides low-reliability arrival times at low SNR. By contrast, most of the picks obtained by the proposed method agree well with the real arrival times, even under a low SNR condition.

We further demonstrate the performance of our algorithm using a synthetic microseismic record containing three events, which is shown in figure 7(a). The vertical and horizontal components are denoted by red and blue lines, respectively. Figures 7(b) and (c) are the records embedded in WGN and real noise, respectively. The SNR of the two records is −6 dB. The arrival time picking results of noisy records obtained by different methods are shown in figure 8.

Comparison results show that when the signal is buried in strong noise, such as the horizontal component, whose amplitude of P-wave is relatively low, the first arrival obtained by the STA/LTA and AIC methods has a great deviation with the real arrival. One drawback of these methods is the poor anti-noise capability, and thus it is hard for them to get desirable picks for weak events. The arrival times of low SNR microseismic data provided by the STA/LTA and AIC methods are frequently false or missed. In contrast, our method outperforms other algorithms, and it attains good picks for the data with low SNR. The accuracy of the arrival times is well improved by the proposed method.

In addition, the arrival time picking experiment for the record with real noise illustrates the validity and feasibility of our method in the application of real microseismic data, to a certain extent. Compared with other methods, the proposed method possesses stronger applicability as it reduces the false or missed arrival times in low SNR cases.

The aforementioned experiments demonstrated the proposed method has better first arrival time picking performance. In this part, we use ROC curves to analyze the actual picking statistic characteristics of our method under different SNRs. We pick the arrival times for a 24-channel microseismic record. We add WGN and real noise to the record, respectively. We do experiments 50 times, and the noise of
each test is different, which means that we conduct the first
arrival time picking experiments 1200 times in total.

Figures 9(a) and (b) show the ROC curves of our pro-
posed method when signals are contaminated by WGN and
real noise with different SNRs, respectively. The ROC
curves closer to the top left corner indicate a higher picking
accuracy. Moreover, the performance can be evaluated by
the area under the ROC curve (AUC). The AUC value is
closer to one that represents a better time picking capability.
Therefore, from figures 9(a) and (b), we can observe that
with the decrease of the SNR, the picking accuracy is gra-
dually decreased. The detailed statistical data and AUCs are
given in table 1. The statistics show that when the SNR
is low to $-6$ dB, the proposed method still has high picking
accuracy.

Figure 9. (a) and (b) ROC curves of the ST when signals are buried in WGN and real noise with different SNRs. (c) and (d) ROC curves of different methods when signals are embedded in WGN and real noise with the SNR of $-6$ dB.

Table 1. Statistical data of ST at different SNRs.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>Number of tests</th>
<th>Picking with error</th>
<th>Picking within 0.002 s</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1200</td>
<td>1056</td>
<td>1185</td>
<td>0.9973</td>
</tr>
<tr>
<td>0</td>
<td>1200</td>
<td>1044</td>
<td>1164</td>
<td>0.9821</td>
</tr>
<tr>
<td>WGN</td>
<td>$-6$</td>
<td>1200</td>
<td>864</td>
<td>994</td>
</tr>
<tr>
<td>$-8$</td>
<td>1200</td>
<td>584</td>
<td>746</td>
<td>0.8356</td>
</tr>
<tr>
<td>$-10$</td>
<td>1200</td>
<td>360</td>
<td>510</td>
<td>0.7168</td>
</tr>
<tr>
<td>5</td>
<td>1200</td>
<td>1068</td>
<td>1180</td>
<td>0.9962</td>
</tr>
<tr>
<td>0</td>
<td>1200</td>
<td>1040</td>
<td>1158</td>
<td>0.9702</td>
</tr>
<tr>
<td>Real noise</td>
<td>$-6$</td>
<td>1200</td>
<td>855</td>
<td>987</td>
</tr>
<tr>
<td>$-8$</td>
<td>1200</td>
<td>579</td>
<td>741</td>
<td>0.8317</td>
</tr>
<tr>
<td>$-10$</td>
<td>1200</td>
<td>352</td>
<td>501</td>
<td>0.6923</td>
</tr>
</tbody>
</table>
We further compare the ROC curves of the ST, the STA/LTA, and the AIC methods at the SNR of −6 dB, which are shown in figures 9(c) and (d). Through observation, it is clear to see that the ST method has a much better picking performance at low SNR. Table 2 shows the specific statistics. Through the data given in the two tables, we see that for the tests of WGN at the SNR of −6 dB, a total of 17% arrival times within 0.002 s and 35% within 0.004 s were found by the STA/LTA method. As for the AIC method, 12.83% first-break times are picked within 0.002 s and 20.75% are within 0.004 s. In contrast, our method picked 72% arrival times accurately, of which 82.83% are within 0.002 s. The picking accuracy of our method is improved by 65.83% and 70% compared with that of the STA/LTA and the AIC methods, respectively. For the situation of the real noise, our method reduces the wrong picks, and the picked arrival times are all reasonable, even for the horizontal components whose energy is weaker.

Since the true first arrivals of the real microseismic data are unknown, here the manual picks, which are looked upon as the ‘correct’ picks, are used to estimate the accuracy of the automatic picks (Tan and He, 2016). Comparing the automatic picks with the hand-pickings, the average absolute deviations of the STA/LTA, AIC, and proposed methods are $3.1 \times 10^{-3}$s, $1.11 \times 10^{-2}$s and $1.4 \times 10^{-3}$s, respectively. Their standard deviations are $3.7 \times 10^{-3}$s, $2.05 \times 10^{-2}$s, and $1.7 \times 10^{-3}$s, respectively. The statistical data illustrates that our method is feasible and valid to improve the picking accuracy for low SNR microseismic data.

6. Conclusions

We have presented a new arrival time picking method based on ST. This method takes the coefficient difference between

<table>
<thead>
<tr>
<th>Methods</th>
<th>Number of tests</th>
<th>Picking within 0.002 s</th>
<th>Picking within 0.004 s</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>STA/LTA</td>
<td>1200</td>
<td>204</td>
<td>420</td>
<td>0.8052</td>
</tr>
<tr>
<td>WGN AIC</td>
<td>1200</td>
<td>154</td>
<td>249</td>
<td>0.7341</td>
</tr>
<tr>
<td>Real noise</td>
<td>STA/LTA</td>
<td>1200</td>
<td>199</td>
<td>352</td>
</tr>
<tr>
<td>AIC</td>
<td>1200</td>
<td>150</td>
<td>300</td>
<td>0.7272</td>
</tr>
</tbody>
</table>

Figure 10. (a) Real noisy microseismic record. (b) Picking results of the ST (• in green), STA/LTA (• in black) and AIC (• in magenta) methods.
the signal and the noise at fine scales into account. First, different criteria are selected at two adjacent fine scales to realize the prediction of signal points. Second, we use the scale correlation between adjacent scales to correct the signal points, which improves the accuracy of the signal recognition. Finally, the first identified signal point is regarded as the first arrival.

This paper aims to improve the arrival time picking accuracy for low SNR microseismic data. The effectiveness and accuracy of the proposed approach have been verified by experiments on synthetic and real microseismic records. Analysis of the actual picking performance provided by ROC curves also shows the better performance of our method. However, in future research, it is required to ensure the consistency of first arrivals for three-component microseismic recordings as far as possible, to make the algorithm have advantages in terms of accuracy and consistency.

Acknowledgments

This research is financially supported by the National Natural Science Foundation of China (Grant nos. 41574096, 41130421, and 41274118).

Appendix. Shearlet transform

ST is the combination of composite wavelet theory and multiscale geometric analysis. It is an optimal multi-dimensional sparse representation function constructed through the affine system with composite dilations (Guo and Labate 2007, Easley et al. 2009, Lim 2010). The shearlet is expressed as

$$\left\{ \psi_{a,s,t}(x) = |\det M_{at}|^{-\frac{1}{2}} \psi(M_{at}^{-1}(x-t)) : a \in R^+, s \in R, t \in R^2 \right\},$$

(A1)

where parameter $a$, $s$, and $t$ indicate the scale, shear, and translation, respectively. $R^+$, $R$, and $R^2$ are sets of positive real numbers, real numbers, and 2D real numbers, respectively. $M_{at}$ is a matrix that can be factorized as $M_{at} = B_a A_t = \left( \begin{array}{cc} 1 & 0 \\ 0 & \sqrt{a} \end{array} \right). A_t$ and $B_a$ denote the anisotropic dilation matrix and shear matrix, respectively.

The shearlet generating function $\psi$ should satisfy appropriate admissibility conditions (Kutyniok and Labate 2009). For the point $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, $\xi_1 \neq 0$, which is defined in the 2D frequency domain, let $\psi(x)$ be given by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \hat{\psi}_2(\xi_2/\xi_1),$$

(A2)

where $\hat{\psi}(\xi)$ indicates the Fourier transform of $\psi(x)$, $\hat{\psi}_1(\xi_1)$ indicates the Fourier transform of $\psi$ in $C^{\infty}(R)$, which is a continuous wavelet and an infinitely differentiable function with the support $\text{sup} \hat{\psi}_1 \subset [-1/2a, -1/a]$, $\hat{\psi}_2(\xi_2/\xi_1) \in C^{\infty}(R)$, $\text{sup} \hat{\psi}_2 \subset [-1, 1]$, with $\hat{\psi}_2 > 0$, $\|\hat{\psi}_2\| = 1$ on $(-1, 1)$.

If $\psi \in L^2(R^2)$, where $L^2$ indicates a twice continuously differentiable function, satisfies all the above assumptions, the definition of 2D ST is as follows

$$SH_{\psi_0}(f) = (f, \psi_{a,s,t}),$$

(A3)

where $SH_{\psi_0}(\cdot)$ is ST, and $(\cdot)$ defines the inner product. The function $f \in L^2(R^2)$ can be recovered by

$$f(x) = \int_{\mathbb{R}^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (f, \psi_{a,s,t})\psi_{a,s,t}(x) da/\alpha d\xi_1 d\xi_2.$$  

(A4)

The characteristics of shearlets are reflected more obviously in frequency domain:

$$\hat{\psi}_{a,s,t}(\xi_1, \xi_2) = \alpha^{3/4}e^{-2\pi i\xi_1\xi_2} \hat{\psi}_1(a\xi_1) \hat{\psi}_2(a^{-1/2}(\xi_2/\xi_1 - s)),$$

(A5)

where $\hat{\psi}_{a,s,t}$ is the Fourier transform of $\psi_{a,s,t}$. The frequency support of each function $\hat{\psi}_{a,s,t}$ can be written as

$$\text{sup} \hat{\psi}_{a,s,t} \subset \{ (\xi_1, \xi_2) : \xi_1 \in [-2/a, -1/2a] \cup [1/2a, 2/a], |\xi_2/\xi_1 - s| \leq \sqrt{a} \}.$$  

(A6)

Shearlets have good characteristics of localization and compact support. For each element $\hat{\psi}_{a,s,t}$, it is supported on origin-symmetric trapezoids, and its orientation is along a straight line with the slope equal to $s$. When $a$ tends to zero, the support becomes increasingly elongated. Parameter $a$, $s$, and $t$ control the scale, orientations, and locations, respectively (Yi et al. 2009, Philipp 2011, Gao et al. 2013). The above characteristics of shearlets are illustrated in figure A1 (Guo and Labate 2009).

References


Lin H, Li Y, Zhang C and Ma H 2015 Curvelet domain denoising based on kurtosis characteristics *J. Geophys. Eng.* **12** 419–26


St-Onge A 2011 Akaike information criterion applied to detecting first arrival times on microseismic data *SEG Annual Int. Meeting, Expanded Abstracts* pp 1658–62

Tan Y and He C 2016 Improved methods for detection and arrival picking of microseismic events with low signal-to-noise ratios *Geophysics* **81** KS89–107


Yonelinas A P and Parks C M 2007 Receiver operating characteristics (ROCs) in recognition memory: a review *Psychol. Bull.* **133** 800–32
