

# Estimation Algorithm for Low Frequency Seismic Motions Based on Extended Semblance Function

Lianming Sun

Faculty of Environmental Engineering, The University of Kitakyushu

1-1 Hibikino, Wakamatsu, Kitakyushu 808-0135, Japan, sun@kitakyu-u.ac.jp

**Abstract.** Monitoring low frequency and long period ground motions is very important for earthquake detection and alarm systems, and signal processing techniques help to investigate the mechanism of deep ground motions and earthquake occurrence by detecting effective seismic information from the seismograms. In this paper a frequency domain algorithm is presented to estimate the travel time difference of seismic waves with multi-path interferences caused by the refraction effect. The spectra of the seismograms are calculated through the fast algorithm, and the multi-path parameters as well as the difference of travel time between a reference position and the seismographic stations are given by an optimization algorithm in the frequency domain. The new approach can work under conditions of refraction interference, and improve the estimation performance by using an extended semblance function. Its effectiveness is demonstrated through some numerical examples, and it is shown that the proposed algorithm is applicable to the analysis of low frequency seismograms.

**Keywords:** Spectral analysis, low frequency seismograms, semblance function, frequency domain.

## 1. Introduction

It has been revealed that the fault activities often accompany by low frequency underground motions, and long period of their accumulation may induce earthquakes. Since many geological phenomena are so complicated that the mechanisms of deep ground tremors and fault slips still remain mysteries, it is difficult to construct an effective physical model for the low frequency motions. Consequently, it is an essential task to monitor the low frequency seismic motions and then to give the alarm of earthquake occurrence. With the significant progress of instrumentation technology, several seismographic networks have been applied to directly monitor the seismic activities in the urgent earthquake detection and alarm systems. Incorporated Research Institutions for Seismology (IRIS) in America, International Data Center managed by Comprehensive Test Ban Treaty (CTBT), Hi-net, Kyoshin-net (K-net) in Japan, are the typical monitoring networks [1].

Using the monitoring networks, the low slip or low-frequency earthquakes, which are considered as the characteristic mode of moment released at a deep structure of a subduction plate interface, are detected from the seismograms [2-5]. Moreover, revealing their mechanisms may help to explore the intrinsic characteristics of seismic phenomena. Nevertheless, the detected signals of low frequency seismograms do not have obvious P-wave and S-wave arrivals; hence it is not easy to locate the epicenters through the ordinary hypocenter determination method [2]. Some array signal processing methods were proposed to maximize a semblance function using a grid search algorithm, then to maximize the cylindrical wave index using steepest descent algorithm [5, 6], or to synchronize the peak of correlation functions of the detected seismic waves in the array records [7]. These methods were performed in time domain, and had been applied in analysis of seismic tremor in Bungo channel region [3], Tokachi channel region [5], Kyoto basin [7], and the long period ground motion of deep ground structure in Yokohama city [8]. Furthermore, the large array scale is often desired in order to acquire much seismic information, whereas a fine grid is required to avoid local minimum in the existing time domain methods. As a result, they lead to considerably heavy computational load. On the other hand, the representation of semblance coefficient has been investigated in the frequency domain, and a frequency domain algorithm has been developed to reduce the computational complexity since some fast frequency algorithm can be utilized [10, 11]. It could achieve the similar accuracy as that of the time domain; moreover, the performance of the presented frequency approach can be improved by compensating the discretization error at low sampling rate. However, when the seismic waves have refraction effects, the multi-path interferences occur and they will shift the maximum of the conventional semblance function due to the side-lobe effects of multi-path. As a result, the estimation accuracy of travel time as well as epicenter location degrades significantly; therefore the effective techniques to reduce the multi-path effects are necessary in the processing algorithms.

An extended semblance function is investigated in this paper. By performing subspace decomposition of the spectral data space, the multi-path parameters can be estimated, and the optimization for the extended semblance function yields the estimation of travel time difference between a reference position and tiltmeter stations. It is seen that the multi-path effects can be handled easily in the frequency domain using the estimated parameters, and the estimation of fine travel time difference can improve the location accuracy of epicenter from the seismograms spectra in the proposed algorithm. Some numerical simulations demonstrate the effectiveness of the proposed approach.

## 2. Problem Statement

Consider the seismograms observed by the tiltmeter arrays, which are composed of several stations in a tiltmeter network. If the separation distance between these stations is smaller than the wavelength of the interested low frequency seismograms, the sampling of the seismic waves holds most of the original information with little aliasing. In the seismograms of low frequency earthquake, though the records have not explicit P-wave and S-wave arrivals, they might be analyzed by array processing techniques to detect and locate the low frequency earthquakes, and to utilize the information to reveal the occurrence mechanism and activities in subduction zones.

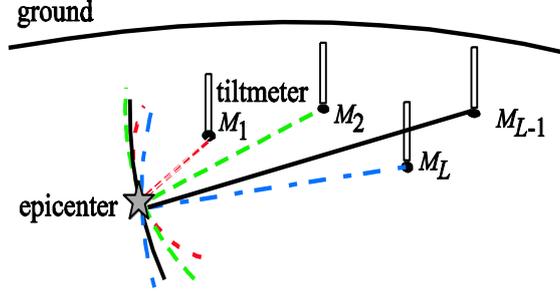


Fig.1 Tiltmeter array in monitoring network

Fig.1 illustrates a tiltmeter array in the monitoring network. It has  $L$  stations, and each station is equipped with a sensor unit at the bottom of a borehole deeper than 100m [3]. Assume that the separation between stations is smaller than the wavelength of seismic waves to reduce the affection of heterogeneity in crust and mantle.

The conventional semblance function  $C(t_{\text{ref}})$  [5, 6] is defined by

$$C(t_{\text{ref}}) = \frac{\sum_{k=1}^K \left( \sum_{l=1}^L a_l(t_{\text{ref}} + kt_{\text{smp}} + t_{l,\text{trv}}) \right)^2}{L \sum_{k=1}^K \sum_{l=1}^L \left( a_l(t_{\text{ref}} + kt_{\text{smp}} + t_{l,\text{trv}}) \right)^2}, \quad (1)$$

where  $k$  is the sample number within the time window  $1 \leq k \leq K$ ,  $t_{\text{ref}}$  indicates a reference time,  $t_{\text{smp}}$  and  $t_{l,\text{trv}}$  represent the sampling interval, the travel time difference of the  $l$ th station from a reference position, respectively, while the later is a function of the propagation path, wave velocity or the horizontal apparent slowness vector, which is a reciprocal vector of velocity. In [5], the semblance function is maximized by finding an appropriate apparent slowness vector using a grid search algorithm, where the grid is set as small spacing in latitude and longitude, and the estimate of epicenter is determined from the estimated vector of apparent slowness. When the grid search algorithm in the time domain is used simultaneously for multiple arrays, the estimation efficiency will be very poor especially for smaller spacing in both latitude and longitude. An efficient frequency domain approach has been considered for the estimation of travel time difference by using the seismic spectra with low computational load [9]. Define the relation function of the observed signal at stations  $l_1$  and  $l_2$  as follows:

$$R_{l_1, l_2}(n_{l_1}, n_{l_2}) = \frac{1}{K} \sum_{k=1}^K a_{l_1}(t_{\text{ref}} + kt_{\text{smp}} + n_{l_1} t_{\text{smp}}) a_{l_2}(t_{\text{ref}} + kt_{\text{smp}} + n_{l_2} t_{\text{smp}}). \quad (2)$$

Then the numerator and denominator of the semblance function can be rewritten as

$$C(t_{\text{ref}}) = \frac{1}{L} + \frac{2 \sum_{l_1=1}^{L-1} \sum_{l_2=l_1+1}^L R_{l_1, l_2} \left( \left[ \frac{t_{l_1, \text{trv}}}{t_{\text{smp}}} \right], \left[ \frac{t_{l_2, \text{trv}}}{t_{\text{smp}}} \right] \right)}{L \sum_{l=1}^L R_{l, l} \left( \left[ \frac{t_{l, \text{trv}}}{t_{\text{smp}}} \right], \left[ \frac{t_{l, \text{trv}}}{t_{\text{smp}}} \right] \right)}, \quad (3)$$

where  $[x]$  is the nearest integer to the real number  $x$ . It is easily seen that maximizing the semblance function (1) is equivalent to maximizing the second term of (3). On the other hand, let the Fourier transform of the observed record be defined by

$$A_l(e^{-i\omega_f}) = \sum_{k=1}^K a_l(t_{\text{ref}} + kt_{\text{smp}}) e^{-ik\omega_f}, \quad (4)$$

where  $\omega_f = f\pi/F$  is the  $f$  th frequency grid,  $-F+1 \leq f \leq F$ . Then the inverse Fourier transform of  $A_{l_1}(e^{-i\omega_f}) A_{l_2}^*(e^{-i\omega_f})$  can be given by

$$X_{l_1, l_2}(k) = \frac{1}{KF} \sum_{f=-F+1}^F A_{l_1}(e^{-i\omega_f}) A_{l_2}^*(e^{-i\omega_f}) e^{ik\omega_f} \approx R_{l_1, l_2}(k_1, k_1 - k). \quad (5)$$

Therefore, the semblance function can be approximated by

$$C(t_{\text{ref}}) \approx \frac{1}{L} + \frac{2 \sum_{l_1=1}^{L-1} \sum_{l_2=l_1+1}^L X_{l_1, l_2}(k_1, k_1 - k)}{L \sum_{l=1}^L X_{l, l}(0)}, \quad (6)$$

then  $C(t_{\text{ref}})$  in (3) reaches a maximum if the values  $k_{l1, l2} = k_{l1} - k_{l2}$  maximize the following function

$$\{ k_{1,2}, \dots, k_{L-1,L} \} = \arg \max_{k_{1,2}, \dots, k_{L-1,L}} \sum_{l_1=1}^{L-1} \sum_{l_2=l_1+1}^L X_{l_1, l_2}(k_{l_1}, k_{l_2}), \quad (7)$$

and  $t_{l_1, \text{trv}} - t_{l_2, \text{trv}} \approx k_{l_1, l_2} t_{\text{smp}}$ . If the integer  $F$  is the power of 2, both  $A_l(e^{-i\omega_f})$  and  $X_{l_1, l_2}(k_{l_1}, k_{l_2})$  can be computed through the algorithm of fast Fourier transform, therefore the algorithm can be performed efficiently [10]. Furthermore, a genetic algorithm based approach can be used to give the estimation of epicenter location [11].

### 3. Extended Semblance Function

Assume that the observed signal at tiltmeter station  $l$  can be approximated as follows:

$$a_l(t_{\text{ref}} + kt_{\text{smp}}) = g_{l,0} a(t_{\text{ref}} + kt_{\text{smp}} - \tau_{l,0}) + g_{l,1} a(t_{\text{ref}} + kt_{\text{smp}} - \tau_{l,1}) + \dots, \quad (8)$$

where  $g_{l,0}$ ,  $g_{l,1}$  and  $\tau_{l,0}$ ,  $\tau_{l,1}$  are the gain and delay time of the direct wave, delayed wave, respectively.  $a(t_{\text{ref}} + kt_{\text{smp}})$  is the source wave comes from the epicenter. Then the spectral representation of  $a_l(t_{\text{ref}} + kt_{\text{smp}})$  can be given by

$$\begin{aligned} A_l(e^{-i\omega_f}) &= \sum_{k=1}^K a_l(t_{\text{ref}} + kt_{\text{smp}}) e^{-ik\omega_f} = \sum_{m=0}^M g_{l,m} e^{-i\omega_f \tau_{l,m}} A(e^{-i\omega_f}) \\ &\triangleq G_l(e^{-i\omega_f}) A(e^{-i\omega_f}). \end{aligned} \quad (9)$$

It is seen that  $t_{l, \text{trv}} = \tau_{l,0}$  if no delay waves in the observed signal  $a_l(t_{\text{ref}} + kt_{\text{smp}})$ . Let the extended semblance function be given by

$$C(t_{\text{ref}}) \approx \frac{1}{L} + \frac{2 \sum_{l_1=1}^{L-1} \sum_{l_2=l_1+1}^L \left( \sum_{f=-F+1}^F \frac{A_{l_1}(e^{-i\omega_f}) A_{l_2}^*(e^{-i\omega_f})}{G_{l_1}(e^{-i\omega_f}) G_{l_2}^*(e^{-i\omega_f})} \right)}{L \sum_{l=1}^L \sum_{f=-F+1}^F \frac{A_l(e^{-i\omega_f}) A_l^*(e^{-i\omega_f})}{G_l(e^{-i\omega_f}) G_l^*(e^{-i\omega_f})}}, \quad (10)$$

then the parameters of  $g_{l,m}$  and  $\tau_{l,m}$  might be estimated by optimizing the semblance function in (10). Nevertheless, the direct optimization of (10) is difficult since it is a severe nonlinear problem, and the optimization may converge to local solution easily. In order to improve the processing efficiency, a simplified semblance function will be used in the estimation algorithm.

### 4. Estimation Algorithm

Without loss of the generality, let the reference position be the tiltmeter station 1, and fix  $g_{1,0}$  as  $g_{1,0}=1$ , then for  $l = 2, \dots, L$ , the travel time difference  $t_{l, \text{trv}}$  of direct waves from the reference position is  $t_{l, \text{trv}} = \tau_{l,0} - \tau_{1,0}$ . Let the sub-semblance function  $C_{1,l}(t_{\text{ref}})$  be given by

$$\begin{aligned} C_{1,l}(t_{\text{ref}}) &\approx \frac{1}{2} \\ &+ \frac{\sum_{f=-F+1}^F \left( G_l(e^{-i\omega_f}) A_l(e^{-i\omega_f}) A_l^*(e^{-i\omega_f}) G_l^*(e^{-i\omega_f}) + G_1(e^{-i\omega_f}) A_l(e^{-i\omega_f}) A_1^*(e^{-i\omega_f}) G_1^*(e^{-i\omega_f}) \right)}{2 \sum_{f=-F+1}^F \left( G_l(e^{-i\omega_f}) A_l(e^{-i\omega_f}) A_l^*(e^{-i\omega_f}) G_l^*(e^{-i\omega_f}) + G_1(e^{-i\omega_f}) A_l(e^{-i\omega_f}) A_1^*(e^{-i\omega_f}) G_1^*(e^{-i\omega_f}) \right)}. \end{aligned} \quad (11)$$

Commonly only a few frequency components are contained in the seismograms, hence the optimization of (11) will be an ill-conditioned problem, and is influenced by noise easily. Instead the direct optimization, the optimization will be performed in the following procedures.

**A.Initial Values.** Set the initial values  $g_{l,0}=1$ ,  $g_{l,1}=g_{l,2}=\dots=0$ , then the initial estimates of  $t_{l, \text{trv}} = \tau_{l,0} - \tau_{1,0}$  can be obtained by (7).

**B.Estimation of Multi-path Parameters.** These parameters will be estimated by a subspace decomposition technique in a grid research way. Construct a spectral data matrix  $\Omega_l(k_{1,1}, k_{1,2}, \dots, k_{l,1}, k_{l,2}, \dots)$  containing both the main lobe and the side-lobe of seismic waves as follows:

$$\begin{aligned}
& \mathbf{Q}_l(k_{1,1}, k_{1,2}, \dots, k_{l,1}, k_{l,2}, \dots) \\
&= \begin{bmatrix} A_l(e^{-i\omega_{-F+1}}) & \dots & A_l(e^{-i\omega_F}) \\ e^{-i\omega_{-F+1}(k_{1,1}\delta+t_{l,\text{trv}})} A_l(e^{-i\omega_{-F+1}}) & \dots & e^{-i\omega_F(k_{1,1}\delta+t_{l,\text{trv}})} A_l(e^{-i\omega_F}) \\ \vdots & & \vdots \\ e^{-i\omega_{-F+1}(k_{l,1}\delta+t_{l,\text{trv}})} A_l(e^{-i\omega_{-F+1}}) & \dots & e^{-i\omega_F(k_{l,1}\delta+t_{l,\text{trv}})} A_l(e^{-i\omega_F}) \\ e^{-i\omega_{-F+1}(k_{l,1}\delta+t_{l,\text{trv}})} A_l(e^{-i\omega_{-F+1}}) & \dots & e^{-i\omega_F(k_{l,1}\delta+t_{l,\text{trv}})} A_l(e^{-i\omega_F}) \\ \vdots & & \vdots \end{bmatrix}^T \quad (12) \\
&\triangleq \begin{bmatrix} \mathbf{Q}_{l,1} \\ \mathbf{Q}_{l,2} \end{bmatrix}^T,
\end{aligned}$$

where  $t_{l,\text{trv}}$  is the estimate given in Step A,  $\delta$  is a grid interval. In order to reduce the influence of noise, the subspace decomposition of (12) is performed as follows:

$$\mathbf{U}_l \mathbf{S}_l \mathbf{V}_l^H = \mathbf{Q}_l^H(k_{1,1}, k_{1,2}, \dots, k_{l,1}, k_{l,2}, \dots) \mathbf{Q}_l(k_{1,1}, k_{1,2}, \dots, k_{l,1}, k_{l,2}, \dots), \quad (13)$$

then the noise subspace vectors  $\mathbf{V}_\perp$ , which are the last column vectors in  $\mathbf{V}_l$ , can be obtained. Denote  $\mathbf{V}_{\perp,1}$  and  $\mathbf{V}_{\perp,2}$  as the parts corresponding to  $\mathbf{Q}_{l,1}$ ,  $\mathbf{Q}_{l,2}$ , respectively, where  $\mathbf{V}_{\perp,1}$  is associated with the parameters  $g_{l,m}$ , whereas  $\mathbf{V}_{\perp,2}$  is associated with the parameters  $k_{l,m}$ . Therefore the optimal values of  $g_{l,m}$ ,  $k_{l,m}$  will be given by

$$\begin{aligned}
& \begin{cases} g_{1,0} = 1, \\ \{g_{1,1}, \dots, g_{1,M}, k_{1,1}, \dots, k_{1,M}\} \end{cases} \quad (14) \\
& = \arg \max_{k_{1,1}, \dots, k_{1,M}} \alpha \frac{\mathbf{V}_{\perp,1}^H \mathbf{Q}_{l,1}^H \mathbf{Q}_{l,2} \mathbf{V}_{\perp,2} + \mathbf{V}_{\perp,2}^H \mathbf{Q}_{l,2}^H \mathbf{Q}_{l,1} \mathbf{V}_{\perp,1}}{\mathbf{V}_{\perp,1}^H \mathbf{Q}_{l,1}^H \mathbf{Q}_{l,1} \mathbf{V}_{\perp,1} + \mathbf{V}_{\perp,2}^H \mathbf{Q}_{l,2}^H \mathbf{Q}_{l,2} \mathbf{V}_{\perp,2}} + (1-\alpha) \frac{1}{\mathbf{V}_\perp^H \mathbf{Q}_l^H \mathbf{Q}_l \mathbf{V}_\perp}.
\end{aligned}$$

The first term in (14) indicates an approximation of the extended semblance function  $C_{1,l}(t_{\text{ref}})$ , while the second one is the criterion for multi-path parameter estimation.

On the other hand, instead of the extended semblance function, an appropriate pre-filter can be utilized to compensate the multi-path effect by using the estimates of  $g_{l,m}$  and  $k_{l,m}$  [12].

**C.Optimization of Extended Semblance Function.** Substitute the estimates of  $g_{l,m}$  and  $k_{l,m}$  into the extended semblance function  $C_{1,l}(t_{\text{ref}})$ , and leave  $t_{l,\text{trv}}$  as a variable, then the optimization of  $C_{1,l}(t_{\text{ref}})$  with respect to  $t_{l,\text{trv}}$  yields fine estimate of the difference of travel time between the tiltmeter stations and reference position.

**D.Estimation of Epicenter Location.** Using the estimates of difference travel time  $t_{l,\text{trv}}$ , the epicenter location can be estimated by the methods given in [5, 11]. Moreover, in order to decrease the influence of noise, the value of semblance function  $C_{1,l}(t_{\text{ref}})$  can be used as weight function for  $t_{l,\text{trv}}$  in the epicenter estimation.

## 5. Numerical Examples

Some simulation examples are considered to investigate the effectiveness of the proposed algorithm. The main frequency components are within 0.03~1.5Hz, and the components beyond the frequency band, the interferences caused by refraction or reflection are considered as noise. The seismic wave travels at velocity of 3km/s. 10 stations whose separation are within 18.1km~20.5km are utilized in the tiltmeter array, and 10 simulation runs be performed in the example. The true locations of epicenter in 10 runs are randomly distributed as normal distribution as  $N(39.0622', (3.2909')^2)$  west in longitude,  $N(18.3011', (2.6903')^2)$  south in latitude, as uniform distribution as 10~15km in the depth from the reference station 1. There is a refraction wave in the layer depth of 21km. The seismograms are sampled at the sampling rate 20Hz, and the records in a time window of 105s are used for the data analysis.

The simulations are performed under the signal to noise ratio of 12.5, 15, 20dB, respectively. The estimation errors of difference travel time and epicenter location are summarized in Tab.1 and 2. For comparison, the results of conventional methods without compensation of multi-path effects are also given in the tables. It is seen that the proposed algorithm can reduce the estimation errors under the situations with multi-path effects, and it works well especially if a great deal of data can be used in estimation even if the signal to noise ration is low.

Tab.1 Standard deviation of difference travel time estimation in 10 simulation runs

Method	Difference travel time( $\pm$ s)		
	12.5dB	15dB	20dB
Conventional algorithm	0.03716	0.02809	0.02998
Proposed algorithm	0.02518	0.02325	0.01158

Tab.2 Standard deviation of epicenter location in 10 simulation runs

Method	Epicenter location ( $\pm$ %)		
	12.5dB	15dB	20dB

Conventional algorithm	3.443	3.119	3.934
Proposed algorithm	2.975	1.751	1.754

## 6. Conclusions

The extended semblance function based algorithm has been proposed to estimate the difference of travel time and the epicenter location for low frequency seismograms. It has been shown that the proposed algorithm can perform optimization in the frequency domain, where the side-lobe can be treated by the shift property of signal spectra. Compared with the conventional methods, the proposed algorithm improves the estimation accuracy of the difference travel time when the seismograms have multi-path effects; therefore it can help to locate the epicenter more precisely. The approach to the case with several epicenters, and the effectiveness validation of the real seismograms will be investigated in the future work.

## References

- [1] M. Kikuchi. Realtime Seismology, University of Tokyo Press, 2003
- [2] K. Obara. "Nonvolcanic deep tremor associated with subduction in southwest Japan," *Science*, vol.269, pp.1679-1681, 2002.
- [3] H. Hirose and K. Obara. "Repeating short- and long-term slow slip events with deep tremor activity around the Bungo channel regions, southwest Japan," *Earth Planets Space*, vol.57, pp.961-972, 2005.
- [4] Committee for Earthquake Prediction Studies, Seismological Society of Japan. *The Science of Earthquake Prediction*, University of Tokyo Press, 2007.
- [5] Y. Asano, K. Obara and Y. Ito. "Spatiotemporal distribution of very-low frequency earthquakes in Tokachi-oki near the junction of the Kuril and Japan trenches revealed by using array signal processing, " *Earth Planets Space*, vol.60, pp.871-875, 2008.
- [6] N.S. Neidell and M.T.Taner. "Semblance and other coherency measures for multichannel," *Geophysics*, vol.36, pp.482-497, 1971.
- [7] C. Shirakawa and T. Iwata. "Ground motion characteristics of South-East Kyoto Basin site using three dimensional small aperture seismic array at Uji Campus, Kyoto University," *Annals of Disaster Prevention Research Institute, Kyoto University*, no.50(B), pp.251-258, 2007.
- [8] H. Miura and S. Midorikawa. "Effects of 3-D deep underground structure on characteristics of rather long-period ground motion," *Earthquake*, no.54, pp.381-395, 2001.
- [9] W. Zhu, L. Sun and X. Zhu. "Analysis of low frequency seismograms in frequency domain," *ICIC Express Letters*, vol.5, no.9(A), pp.3249-3254, 2011.
- [10] L. Sun and A.Sano. "System identification and error analysis in frequency domain," *International Journal of Innovative Computing, Information and Control*, vol.3, no.5, pp.1201-1218, 2007.
- [11] W. Zhu, L. Sun and X. Zhu. "Epicenter location via genetic algorithm for low frequency seismograms, *ICIC Express Letters*, vol.6 (to appear)
- [12] G. Jacovitti and G. Scarano. "Discrete time techniques for time delay estimation," *IEEE Trans. Signal Processing*, vol.41, no.2, pp.525-533, 1993.