

Chapter 15

Feature Extraction and Analysis of Earthquake Motion



Masumitsu Kuse and Nobuoto Nojima

Abstract The evaluation of earthquake motion is basic problem in the earthquake engineering. Deep understanding of the earthquake motion characteristics of the amplitude, period and duration included in observation records is essential. In this chapter, the feature quantity that can explain the engineering characteristics included in the earthquake motion, makes proposal. First, the 99-dimensional feature vector representing the temporal characteristics of a strong motion on the basis of Husid plot was proposed. The Husid plot is defined as the time history of cumulative power of earthquake motion normalized to the total power. In addition, the Gaussian mixture model was applied to the 99-dimensional feature vector was proposed. The Gaussian mixture model approximates the original data in terms of mixture of multiple Gaussian distributions. Next, the time-frequency analysis of earthquake motion, the period-dependent feature vector based on the evolutionary power spectrum, which is proportional to displacement response envelope, was defined. The period-dependent feature vector was calculated from the time history of cumulative power of evolutionary power spectrum normalized to the total power with reference to Husid plot. Furthermore, the Gaussian mixture model was applied to the period-dependent feature vector.

Keywords Feature vector · Gaussian mixture model · Husid plot

M. Kuse (✉)
River Basin Research Center, Gifu University, Gifu, Japan
e-mail: kuse@gifu-u.ac.jp

N. Nojima
Department of Civil Engineering, Gifu University, Gifu, Japan
e-mail: nojima@gifu-u.ac.jp

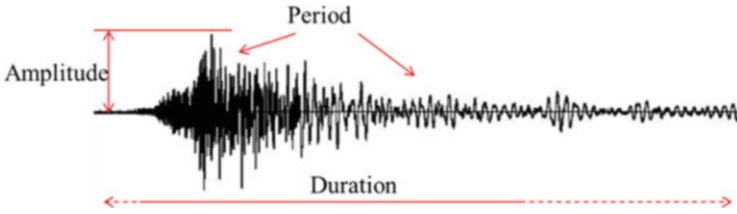


Fig. 15.1 Earthquake motion characteristics

15.1 Introduction

The evaluation of earthquake motion is basic problem in the earthquake engineering. Figure 15.1 shows the outline about the earthquake motion characteristics. As shown Fig. 15.1, deep understanding of the earthquake motion characteristics of the amplitude, period and duration included in observation records is essential. For example, the response spectrum gives useful information about earthquake motion representing period-dependent response amplitude. The group delay time representing the characteristics of duration is the evaluation of the phase characteristics of the earthquake motion. The average of the group delay time represents the position of the center of gravity on time axis, and the variance of the group delay time has the information of the duration characteristics. Using the group delay time, the prediction model of earthquake motion was developed (Sato et al. 1996). Nojima (2014) proposed the prediction equations for the JMA seismic intensity duration considering the shortest distance of the fault, shallow and deep ground data, and the earthquake type. Miyamoto and Honda (2009) proposed of the feature indices of earthquake motion by response value of nonlinear structural model. In the evaluation, the maximum response displacement and the historical absorbed energy were used. Ishii (2012) developed the response duration spectrum as period and duration characteristics of earthquake motion.

The wavelet transform (Iyama and Kawamura 1998) and the nonstationary spectrum (Kamiyama 1979) can describe the characteristics of the amplitude, period, and duration. These parameters show the time variation for each period. The amount of data expressing the earthquake motion is the same as the earthquake time history and the discrete wavelet transform. On the other hand, the amount of data for the continuous wavelet transform and the nonstationary spectrum drastically increases compared to that for the earthquake time history.

In this chapter, for the analysis of earthquake motion, the feature quantity that can explain engineering characteristics included in the earthquake motion, the 99-dimensional feature vector representing the temporal characteristics of strong motion on the basis of the Husid plot was proposed by the authors (Nojima et al. 2017, Kuse et al. 2017). The Husid plot is defined as the time history of cumulative power of earthquake motion normalized to the total power. In addition, the Gaussian mixture model was applied to the 99-dimensional feature vector was proposed. The Gaussian mixture model approximates the original data in terms of mixture of

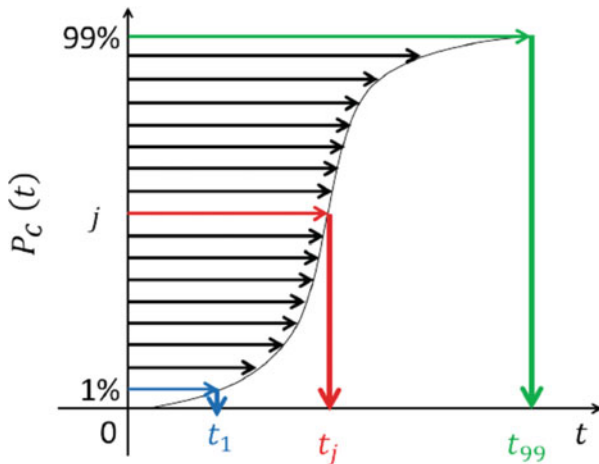
multiple Gaussian distributions. The Gaussian mixture model approximates the 99-dimensional feature vector with much less amount of data. Next, for the time-frequency analysis of earthquake motion, the period-dependent feature vector based on the evolutionary power spectrum, which is proportional to displacement response envelope, was defined. The period-dependent feature vector was calculated from the time history of cumulative power of evolutionary power spectrum normalized to the total power with reference to Husid plot. Furthermore, the Gaussian mixture model was applied to the period-dependent feature vector.

15.2 Outline of Feature Vector and Gaussian Mixture Model

In this chapter, the 99-dimensional feature vector representing the temporal characteristics of a strong motion is used. The feature vector is calculated from the Husid plot that is defined as the time history of cumulative power of earthquake motion normalized to the total power. Figure 15.2 shows the conceptual diagram of the Husid plot $P_c(t)$ and the feature vector \mathbf{t} . $P_c(t)$ is calculated from the Eq. (15.1) using the acceleration time history $A(t)$. The 99-dimensional feature vector \mathbf{t} defines from the percentile values t_i ($i = 1, \dots, 99$) discretized in every 1% of $P_c(t)$.

$$P_c(t) = \frac{100 \int_0^t A^2(\tau) d\tau}{\int_0^{t_{\text{end}}} A^2(\tau) d\tau} \tag{15.1}$$

Fig. 15.2 Husid plot and the feature vector \mathbf{t}



Next, in order to represent of earthquake motion with more less the amount of data, the Gaussian mixture model was applied. The Gaussian mixture model approximates the original data in terms of mixture of multiple Gaussian distributions. The Gaussian distributions $\phi(t)$ with mean μ and standard deviation σ is shown Eq. (15.2).

$$\phi(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(t-\mu)^2}{2\sigma^2} \right\} \quad (15.2)$$

The probability density function represented by the Gaussian mixture model of M elements is shown Eq. (15.3).

$$p(t; \theta) = \sum_{m=1}^M \pi_m \phi(t; \mu_m, \sigma_m) \quad (15.3)$$

where $\pi_m(1, \dots, M)$ is the mixture fraction of element model m , and π_m satisfies Eq. (15.4). As shown Eq. (15.5), the vector θ represents the all parameters.

$$\pi_m \geq 0, \quad \sum_{m=1}^M \pi_m = 1 \quad (15.4)$$

$$\theta = \{\pi_m, \mu_m, \sigma_m; m = 1, \dots, M\} \quad (15.5)$$

The number of element model M is defined from the information criterion. In this paper, BIC(Bayesian Information Criterion) is used.

$$\text{BIC}(\hat{\theta}) = -2 \sum_{i=1}^I \ln p(t_i; \hat{\theta}) + k \ln I \quad (15.6)$$

where I is number of data and k is degree of freedom with this model. In this study, $I = 99$ because the 99-dimensional feature vector is used, and k is $k = 3M - 1$ from Eq. (15.4) and Eq. (15.5).

In addition, for the time-frequency analysis of earthquake motion, the period-dependent feature vector based on the evolutionary power spectrum (Kameda 1975), which is proportional to displacement response envelope, is defined. The evolutionally power spectrum $G(t, \omega)$ is defined as Eq. (15.7).

$$G(t, \omega) = \frac{2h\omega^3}{\pi} \left\{ y^2(t) + \frac{\dot{y}^2(t)}{\omega^2} \right\} \quad (15.7)$$

where h is the damping constant (5%), $y(t)$ is the relative displacement of the single degree of freedom system by acceleration time history $A(t)$, ω is the natural circular frequency.

The period-dependent feature vector is defined from the time history of cumulative power of evolutionary power spectrum normalized to the total power with reference to Husid plot. The time history of period-dependent $P_e(t, \omega)$ can be calculated from Eq. (15.7) based on Fig. 15.2 and Eq. (15.1).

$$P_e(t, \omega) = \frac{100 \int_0^t G(\tau, \omega) d\tau}{\int_0^{t_{\text{end}}} G(\tau, \omega) d\tau} \quad (15.8)$$

The period-dependent feature vector $\mathbf{t}_e = \{t_{\omega j}\}$ is defined from the percentile values $t_{\omega j}$ ($j = 1, \dots, 99$) discretized in every 1% steps of $P_e(t, \omega)$. The natural period T is calculated $N = 101$ components by Eq. (15.9).

$$T_i = 0.1 \times 10^{\left(\frac{2}{100}\right)(i-1)} \quad (i = 1, \dots, 101) \quad (15.9)$$

15.3 Application to Earthquake Motion and Consideration

15.3.1 Consideration of Feature Vector and Gaussian Mixture Model

For numerical example, the acceleration records observed from the 2011 Off the Pacific Coast of Tohoku Earthquake were applied the 99-dimensional feature vector. Figure 15.3 shows the acceleration time histories observed at three K-NET stations (National Research Institute for Earth Science and Disaster Resilience 2021a) which are the observation network operated by National Research Institute for Earth Science and Disaster Resilience, and the envelope calculated by Gaussian mixture model. In Fig. 15.3, the envelope calculated by Gaussian mixture model is the approximate result with the optimal number of element models calculated from Eq. (15.6). In these acceleration records, AKT008 is the long duration motion and the standard shape of acceleration envelope. CHB003 has the single peak with sharp amplitude, and MYG004 has two peaks with large amplitude. As compared to the acceleration time history and shape of envelopes calculated by proposed method, the shape of envelope of AKT008 and CHB003 are similar to the earthquake motion. The shape of envelope at MYG004 have significant emphasis on the peak of amplitude, but the shape of envelope is sufficiently approximated.

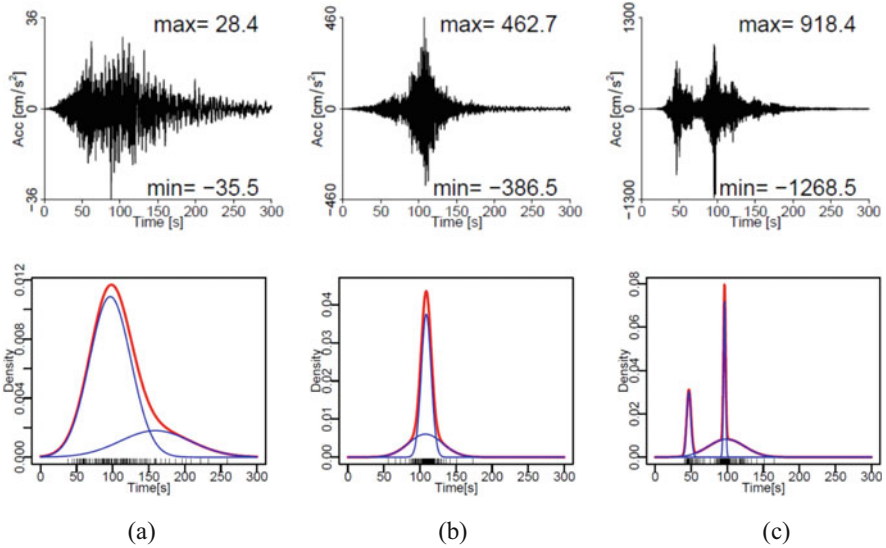


Fig. 15.3 Comparison of acceleration time history and the envelope calculated by Gaussian mixture model (Kuse and Nojima 2021, The red line is the envelope by Gaussian mixture model. The Blue line is the element model. The black line at the bottom is time distribution of the 99-dimensional feature vector.) (a) AKT008 (b) CHB003 (c) MYG004

15.3.2 Consideration of Period-Dependent Feature Vector and Gaussian Mixture Model

The period-dependent feature vector is applied to the earthquake motions observed from the Tokaido-oki earthquake that occurred at 23:57 on September 5, 2004. In this study, the earthquake motion observed in around the Nobi Plain are used. Figure 15.4 shows the location of the observation stations. The earthquake motion records were obtained at K-NET and KiK-net stations (National Research Institute for Earth Science and Disaster Resilience 2021a). Figure 15.5 shows the observed earthquake motion converted into three-dimensional components (radial, transverse and vertical). These earthquake motions have relatively long period characteristics. Figure 15.6 shows the dispersion curve of surface wave calculated using the deep ground model released by Japan Seismic Hazard Information Station(J-SHIS) (National Research Institute for Earth Science and Disaster Resilience 2021b). Figure 15.6 is used in the discussion of surface wave described later in 15.3.3. These three stations shown in Fig. 15.6 are in the Noubi Plain composed of thick sedimentary layers.

Figures 15.7, 15.8, and 15.9 show the evolutionally power spectrum, the period-dependent feature vector, and the envelope applied the Gaussian mixtre model, respectively, for AIC001. The blue and red horizontal lines in Figs. 15.8 and 15.9 are minimum value of group velocity for Love wave and Rayleigh wave. The blue

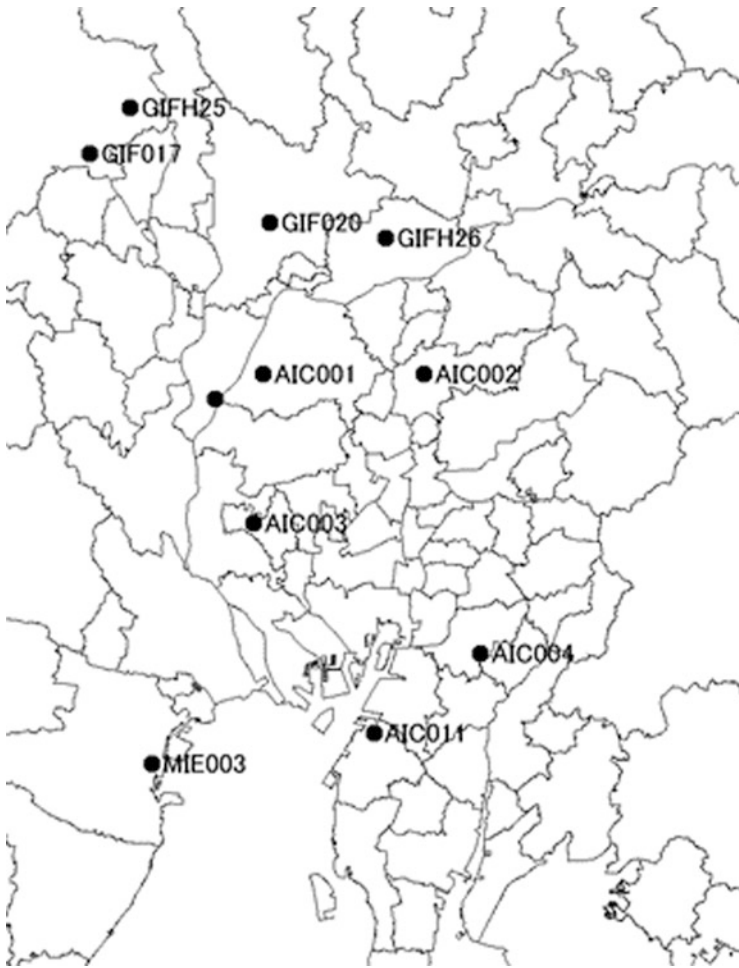


Fig. 15.4 Location of observation stations around the Noubi Plain

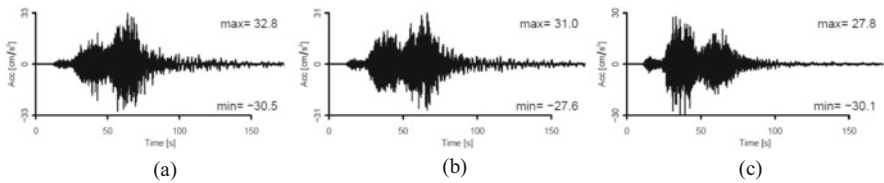


Fig. 15.5 Comparison of acceleration time history (Kuse and Nojima 2021). (a) Radial (b) Transverse (c) UD

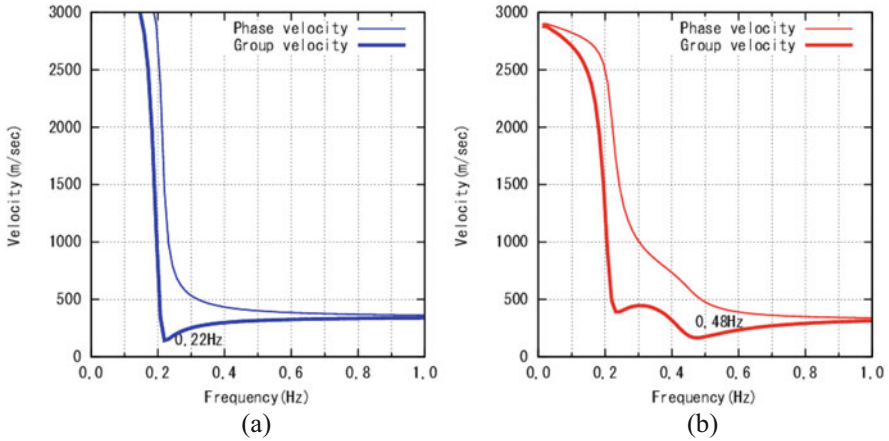


Fig. 15.6 Comparison of dispersion curve of surface wave. (a) Love wave (b) Rayleigh wave

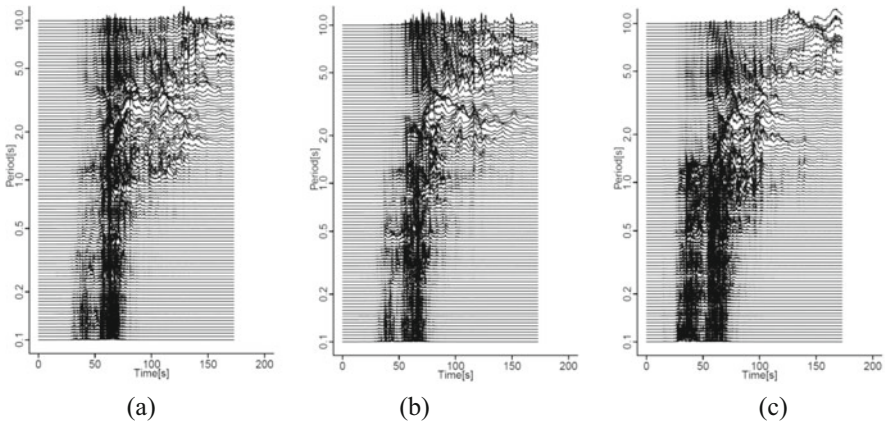


Fig. 15.7 Comparison of evolutionary power spectrum at AIC001 (Kuse and Nojima 2021). (a) Radial (b) Transverse (c) UD

and red curves on Fig. 15.8 are group delay time about Love wave and Rayleigh wave calculated using the group velocity in Fig. 15.6. The group delay time in Fig. 15.8 does not take account of the propagation of earthquake motions from epicenter to observation station. The group delay time calculated using the Fig. 15.6 assumes that the earthquake motion propagated in the Noubi Plain, and the propagation distance in the Noubi Plain is assumed as 10 km, 20 km and 30 km.

In Fig. 15.7, the large amplitude around 30 s to 50 s that seems to be the primary wave is shown in the band of period shorter than about 0.5 s of radial and transverse components, and about 1.2 s in UD component. The large amplitude that appears around 50 s to 80 s at all periods in each component of Fig. 15.7 is principal motion as compared with the acceleration time histories in Fig. 15.5. In addition, the

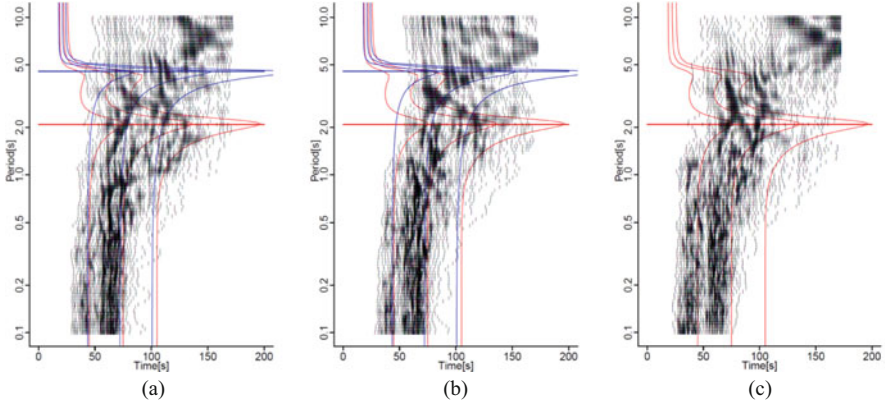


Fig. 15.8 Comparison of the period-dependent feature vector at AIC001 (Kuse and Nojima 2021). (a) Radial (b) Transverse (c) UD

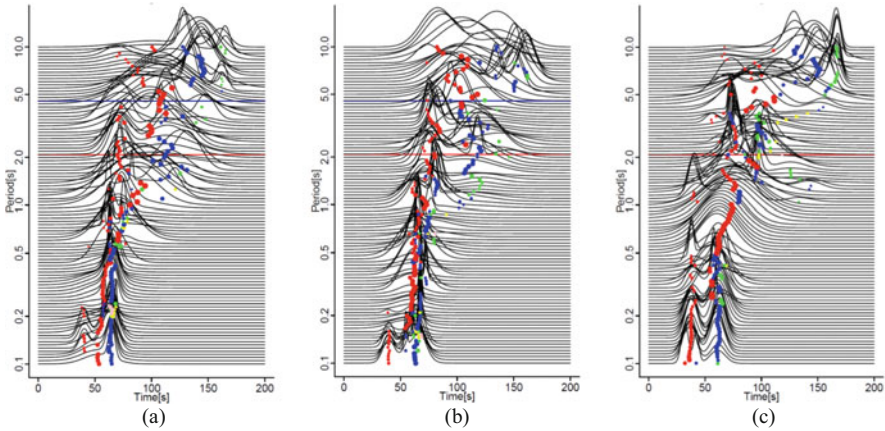


Fig. 15.9 Comparison of the envelope applied the Gaussian mixture model at AIC001 (Kuse and Nojima 2021). (a) Radial (b) Transverse (c) UD

amplitude which is considered to be the surface wave is confirmed in the long period band.

The period-dependent feature vector shown in Fig. 15.8 is calculated from the time history of cumulative power of evolutionary power spectrum normalized to the total power with reference to Husid plot. As previously explained, The period-dependent feature vector is defined by the time of every 1% steps of $P_e(t, \omega)$. Therefore, the time when the feature vector is concentrated corresponds to the time of the large amplitude of the evolutionally power spectrum.

Figure 15.9 shows the approximate envelopes from the Gaussian mixture model calculated with the optimal number of element models. In this figure, the points are the mean of the element model μ_m , and the size of points are the mixture fraction of

the element model π_m . As shown Fig. 15.9, the shape of envelopes calculated by Gaussian mixture model are similar to those by the original evolutionally power spectrum. In Fig. 15.9, the points around 40 s that seen in the short period is considered to be the primary wave. The points around 60 s that seen in the band of period shorter than about 1 s are considered to be principal motion. In the long period range, the μ_m are longer than the 60 s due to the surface wave.

15.3.3 Examination of the Dispersibility of Surface Wave

Next, the proposed period-dependent feature vector was applied to examination of the dispersibility of surface wave. Kamiyama (1979) showed that the dispersibility of surface wave can explain by use of the nonstationary spectrum. Sato et al. (1996) compared the theoretical group delay time of Love wave, and the nonstationary spectrum. Sato et al. showed long-period earthquake motion with dispersibility can be simulated by using the average and standard deviation of group delay time. Sugito et al. (1984) focused on the dispersibility of surface wave, and proposed the method to detect the surface wave using by the evolutionary power spectrum. Figure 15.10 shows the example of evolutionary power spectra. As shown Fig. 15.10a, in the earthquake motion containing surface wave, the peak of amplitude around long period appears later than principal motion. On the contrary, Fig. 15.10b shows the example

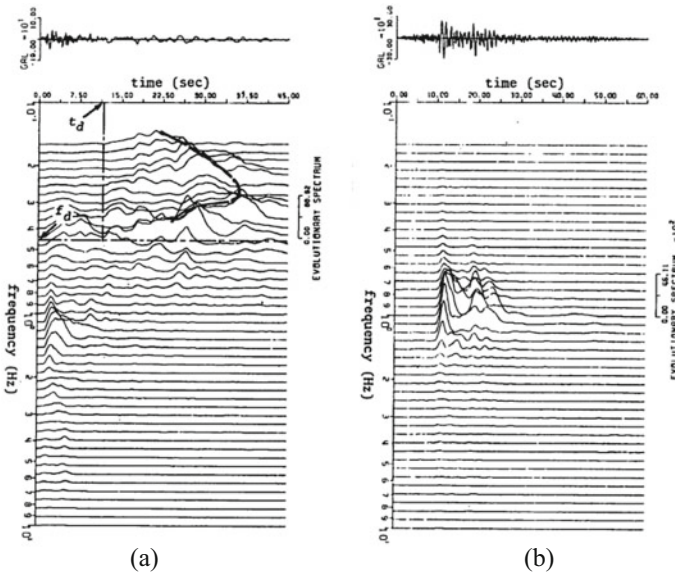


Fig. 15.10 Example evolutionary power spectra (Sugito et al. 1984) (Radial component). (a) Izu-Oshima Kinkai Earthquake (1978) (b) Miyagiken-oki Earthquake (1978). Shimizumiho site. Shiogama kojyo site

for the earthquake motion that contains no surface wave component, in which the peak of amplitude around long period does not appear.

Figure 15.7 is considered to contain surface wave from the discussion in Sect. 15.3.2. In Fig. 15.7 showing the radial component, the dispersibility characteristics of surface wave is seen at around 80 s of period 3 s. In the same way, the dispersibility characteristics are seen at around 80 s of period 3.5 s in transverse components, and around 80 s of period 3 s, around 100 s of period 4 s in UD components. In the period-dependent feature vector shown in Fig. 15.8, the dispersibility characteristics of surface wave appears the triangle-shaped distribution that the feature vector is dense. Comparing Figs. 15.7 and 15.8, the triangle-shaped distribution is seen at the time and period when the dispersibility characteristics of surface wave. The Rayleigh wave is contained in the radial and UD components, and the Love wave is contained in the transverse component. Figure 15.6 showed the dispersion curve of surface wave, and Fig. 15.8 showed the group delay time. In Figs. 15.7 and 15.8, the dispersibility of surface wave was found, but the component of surface wave is unclear whether Rayleigh wave or Love wave that compared with Figs. 15.6, 15.7, and 15.8. Figure 15.9 shows the envelope from Gaussian mixture model and the element model μ_m . In Fig. 15.9, the dispersibility characteristics of surface wave are seen in the Transverse component, but The Radial and UD component cannot be seen.

15.4 Conclusion

In this chapter, the feature quantity that can explain the engineering characteristics included in the earthquake motion was proposed. The major results derived in this study can be summarized as follows.

1. The 99-dimensional feature vector representing the temporal characteristics of a strong motion on the basis of Husid plot was proposed. The amount of data for proposed feature quantity is far less than that for earthquake time history.
2. The Gaussian mixture model was applied to the 99-dimensional feature vector. As numerical example, the acceleration records observed in the 2011 Off the Pacific Coast of Tohoku Earthquake were used. The shape of envelopes calculated by proposed method were similar to those by the original earthquake motion.
3. For the time-frequency analysis of earthquake motion, the period-dependent feature vector based on the evolutionary power spectrum, which is proportional to displacement response envelope, was defined. The period-dependent feature vector was calculated from the time history of cumulative power of evolutionary power spectrum normalized to the total power with reference to Husid plot. Furthermore, the Gaussian mixture model was applied to the period-dependent feature vector.

4. As the case study of the earthquake motion analysis, the proposed period-dependent feature vector was applied to examination of the dispersibility of surface wave. The feature vector, the Gaussian mixture model and the dispersion curve calculated from the soil profiles were compared. As a result of comparison of these parameters, the dispersibility of surface wave observed in the evolutionally power spectrum was confirmed.

References

- Ishii T (2012) Response duration time spectra of earthquake ground motions. *J Struct Constr Eng AIJ* 77(676):843–850. (in Japanese)
- Iyama J, Kawamura H (1998) Time-frequency analysis of earthquake ground motions by wavelet transform. *J Struct Constr Eng AIJ* 514:59–64. (in Japanese)
- Kameda H (1975) On a method of computing evolutionary power spectra of strong motion seismograms. *Proceedings of the Japan Society of Civil Engineers* 235:55–62. (in Japanese)
- Kamiyama M (1979) Nonstationary characteristics and wave interpretation of strong earthquake ground motions. *Proceedings of the Japan Society of Civil Engineers*, No 284:35–48. (in Japanese)
- Kuse M, Nojima N, Takashima T (2017) Application of the self-organizing maps for duration characteristics of earthquake motion. *J Jpn Soc Civil Eng, Ser. A1*, 73 (4), I_558-I_567. (in Japanese)
- Kuse M, Nojima N (2021) Analysis of earthquake motion and surface motion using the Gaussian Mixture Model, *Proc. of the 17th World Conference on Earthquake Engineering*, ID: 1d-0099 (On-line)
- Miyamoto T, Honda R (2009) Similarity estimation of seismic waves from the viewpoint of the effects on structural systems based on nonlinear structural response as feature indices. *JSCE J Earthquake Eng* 30:88–96. (in Japanese)
- National Research Institute for Earth Science and Disaster Resilience (2021a) Strong-motion Seismograph Networks (K-NET, KiK-net). <http://www.kyoshin.bosai.go.jp/>. Accessed December 10th, 2021
- National Research Institute for Earth Science and Disaster Resilience (2021b) Japan Seismic Hazard Information Station (J-SHIS). <http://www.j-shis.bosai.go.jp/map/?lang=en>. Accessed December 10th, 2021
- Nojima N (2014) Conditional prediction equations for duration of specified intensity levels under given predicted or observed JMA seismic intensity. *J Jpn Assoc Earthquake Eng* 14(5):50–67. (in Japanese)
- Nojima N, Kuse M, Takashima T (2017) Feature extraction and classification of temporal characteristics of earthquake motions. *J Jpn Assoc Earthquake Eng* 17(2):2_128–2_141. (in Japanese)
- Sato T, Sato T, Uetake T, Sugawara Y (1996) A fundamental study for envelope characteristics of long period strong motions by using group delay time. *J Struct Constr Eng, AIJ* 480:57–65. (in Japanese)
- Sugito M, Goto H, Aikawa F (1984) Simplified separation technique of body and surface waves in strong motion accelerograms. *Proc of JSCE Structural Eng/Earthquake Eng* 1(2):71–76