Acoust. Sci. & Tech. **35**, 1 (2014)

©2014 The Acoustical Society of Japan

ACOUSTICAL LETTER

Blind source separation and two-signal localization in time-frequency domain considering time lag information: Application to the case where one signal includes a reflected signal

Megumi Yanai*, Fumio Sasaki†, Osamu Tanaka‡ and Masahito Yasuoka§

Department of Architecture, Faculty of Engineering, Tokyo University of Science, 6–3–1 Niijuku, Katsushika-ku, Tokyo, 125–8585 Japan

(*Received 27 November 2012, Accepted for publication 22 May 2013*)

Keywords: Blind source separation, Specification of location, One reflection, Time-frequency information

PACS number: 43.55.Ka, 43.60.Ac, 43.60.Hj [doi:10.1250/ast.35.55]

1. Introduction

We previously proposed a technique for both determining the location of source signals and separating them by blind source separation [1]. Moreover, by extending this technique, we showed the separation and specification of the location of signals in the case of a source signal embedded in another signal, such as noise in the time-frequency domain [2,3]. However, in that method [1–3], the reflected sounds of the source signal could not be considered.

In this paper, for a signal embedding a source signal and its reflected sound in the time-frequency domain, we describe a method for the separation and specification of these two signals. Our method is validated through a numerical test.

2. Formulation

2.1. Assumptions

In this paper, we assume that two signals exist. Let $s_1(t)$ and $s_2(t)$ be the signal data and s_1 be embedded in s_2 in the time-frequency domain. We assume that these data are point sound sources that emanate from different locations. In addition, we assume that the number of observation points is $M (M \geq 4)$; $x_k(t)$ $(1 \leq k \leq M)$ refers to the observation data. In general, $M = 4$ is sufficiently large. In this study, there is one wall and the reflected sound of s_1 is considered. s_2 originates from the wall and s_1 originates at another point away from the wall. Therefore, two signals exist from s_1 to the observation points: the direct signal and the signal reflected by the wall. Moreover, to obtain signal data, we assume the following relations hold:

$$x_k(t) = a_{k11} s_1(t - c_{k11}) + a_{k12} s_1(t - c_{k12}) + a_{k21} s_2(t - c_{k21}). \quad (1)$$

Here, a_{kjh} represents the real-valued damping coefficient and c_{kjh} represents the real-valued time lag between $x_k(t)$ and $s_1(t)$, $s_2(t)$ $(1 \leq k \leq M, j = 1, 2, h = 1, 2)$. In Eq. (1), the first term on the right side expresses the direct signal of s_1. In this paper, a term including s_1 expressing a reflected sound is newly added as the second term. The third term expresses a direct signal of s_2. The only known data are the observation signal data $x_k(t)$, the location of x_k $(1 \leq k \leq M)$ and the number of signals $(j = 1, 2)$; $s_j(t)$, a_{kjh}, c_{kjh}, and the locations of $s_j(t)$ $(j = 1, 2)$ are all unknown data.

We specify $s_j(t)$, a_{kjh}, and c_{kjh} only using $x_k(t)$.

2.2. Time-frequency information

Let $X_k(t, \omega)$ be the time-frequency information for observation data $x_k(t)$, in the form of a complex-valued function. In this paper, the continuous wavelet transform, for which the integral kernel consists of a complex mother wavelet, which, in turn, consists of a Meyer wavelet (the real part) and the Hilbert transform of the Meyer wavelet (the imaginary part), is adopted to obtain $X_k(t, \omega)$ [4,5]. Thus, Eq. (1) is transformed into the following equation by the continuous wavelet transform.

$$X_k(t, \omega) = a_{k11} S_1(t - c_{k11}, \omega) + a_{k12} S_1(t - c_{k12}, \omega) + a_{k21} S_2(t - c_{k21}, \omega). \quad (2)$$

Here, $S_j(t, \omega)$ is the time-frequency information for $s_j(t)$. We define the set of time-frequency domains as $H_{k1h} = \{(t, \omega) : S_1(t - c_{k1h}, \omega) \neq 0\}$ $(h = 1, 2)$, $H_{k21} = \{(t, \omega)$: $S_2(t - c_{k21}, \omega) \neq 0\}$. In this paper, because we assume that s_1 is embedded in s_2 in the time-frequency domain, the following relationships hold: $H_{k11}, H_{k12} \subset H_{k21}$ and $H_{k11} \neq H_{k12}$, $E_k = H_{k21} - (H_{k11} \cup H_{k12}) \neq \{\phi\}$ & *measure* $\neq 0$.

Let $\tilde{E}_k$ be a set of time shifts of E_k to $-c_{kjh}$. In addition, let G be the set defined by

$$G = \bigcap_{k=1}^{M} \tilde{E}_k, \quad (3)$$

which has a measure in the same range as E_k.

2.3. Calculation of damping ratio and time lag of s_2

For arbitrary values of k, l $(1 \leq k, l \leq M, k \neq l)$, the complex-valued quotient function $Q(t_k, t_l, \omega)$ is introduced and defined as $Q(t_k, t_l, \omega) = X_k(t_k, \omega)/X_l(t_l, \omega)$. Because $X_k(t_k, \omega)$ and $X_l(t_l, \omega)$ are complex-valued functions, Q is generally a complex-valued function. However, if $(t_k, \omega) \in E_k$, then $X_k(t_k, \omega) = a_{k21} S_2(t_k - c_{k21}, \omega)$, and if $(t_l, \omega) \in E_l$, then $X_l(t_l, \omega) = a_{l21} S_2(t_l - c_{l21}, \omega)$. Furthermore, if

$$t_l = t_k - (c_{k21} - c_{l21}) \quad (4)$$

is established, then Q satisfies the following relationship:

*e-mail: megumi_yanai_421@yahoo.co.jp
†e-mail: fsasaki@rs.kagu.tus.ac.jp
‡e-mail: otanaka@rs.kagu.tus.ac.jp
§e-mail: yasuokawma@y3.dion.ne.jp

$$Q(t_k, t_l, \omega) = a_{k21}/a_{l21} \in \mathbf{R} \ (\mathbf{R}\text{: the set of real values}). \quad (5)$$

Thus, Q takes at least all the real values in the set G. By keeping ω fixed and varying only t_k and t_l in the time-frequency domain, we can obtain the damping coefficient ratio a_{k21}/a_{l21}, which is a real, constant value. Q may take other real values outside the domain G; however, in such cases, the existence of a corresponding measurable region is very rare.

2.4. Specification of location of s_2

When Eq. (4) holds, $t_l - t_k = c_{l21} - c_{k21}$. The right side of this equation is the relative time lag between s_2 to x_k and x_l. When we multiply the speed of sound by this time lag, the relative distance from s_2 to x_k and x_l is obtained. Generally, in a two-dimensional plane, the possible locations of s_2 form a hyperbolic curve. Thus, the points of intersection become the positions of s_2 when we consider a pair of observation points and draw several hyperbolic curves. When the location of s_2 has been decided, the distances to the observation points become clear, and c_{k21} is obtained.

2.5. Calculation of damping ratio and time lag of s_1

In Eq. (1), for arbitrary values of k, l $(1 \le k, l \le M, k \ne l)$, the next equation holds:

$$\begin{aligned} &x_l(t + c_{l21})/a_{l21} - x_k(t + c_{k21})/a_{k21} \\ &\quad = a_{l11}/a_{l21} \cdot s_1(t - c_{la}) - a_{k11}/a_{k21} \cdot s_1(t - c_{ka}) \\ &\qquad + a_{l12}/a_{l21} \cdot s_1(t - c_{lb}) - a_{k12}/a_{k21} \cdot s_1(t - c_{kb}), \end{aligned} \quad (6)$$

where $c_{la} = c_{l11} - c_{l21}$, $c_{lb} = c_{l12} - c_{l21}$ and $c_{ka} = c_{k11} - c_{k21}$, $c_{kb} = c_{k12} - c_{k21}$. We multiply a_{k21} by the statement in Eq. (6) and define $y_1(t)$ as

$$\begin{aligned} y_1(t) &= a_{la} \cdot s_1(t - c_{la}) + a_{ka} \cdot s_1(t - c_{ka}) \\ &\quad + a_{lb} \cdot s_1(t - c_{lb}) + a_{kb} \cdot s_1(t - c_{kb}), \end{aligned} \quad (7)$$

where $a_{la} = a_{l11} \cdot a_{k21}/a_{l21}$, $a_{ka} = -a_{k11}$, $a_{lb} = a_{l12} \cdot a_{k21}/a_{l21}$ and $a_{kb} = -a_{k12}$. Let l be fixed for an arbitrary value of m $(1 \le m \le M, m \ne l, m \ne k)$; we define $y_2(t)$ in the same manner as $y_1(t)$:

$$\begin{aligned} y_2(t) &= \gamma \cdot a_{la} \cdot s_1(t - c_{la}) + a_{ma} \cdot s_1(t - c_{ma}) \\ &\quad + \gamma \cdot a_{lb} \cdot s_1(t - c_{lb}) + a_{mb} \cdot s_1(t - c_{mb}), \end{aligned} \quad (8)$$

where $\gamma = (a_{m21}/a_{l21})/(a_{k21}/a_{l21})$, $a_{ma} = -a_{m11}$, $a_{mb} = -a_{m12}$ and $c_{ma} = c_{m11} - c_{m21}$, $c_{mb} = c_{m12} - c_{m21}$. In Eqs. (7) and (8), $y_1(t)$, $y_2(t)$ and γ are known. Equations (7) and (8) are transformed into the Fourier domain, then the following equations hold:

$$\begin{aligned} \hat{y}_1(\omega) &= (a_{la} \cdot e^{-i\omega c_{la}} + a_{ka} \cdot e^{-i\omega c_{ka}} \\ &\quad + a_{lb} \cdot e^{-i\omega c_{lb}} + a_{kb} \cdot e^{-i\omega c_{kb}})\hat{s}_1(\omega) \end{aligned} \quad (9)$$

$$\begin{aligned} \hat{y}_2(\omega) &= (\gamma \cdot a_{la} \cdot e^{-i\omega c_{la}} + a_{ma} \cdot e^{-i\omega c_{ma}} \\ &\quad + \gamma \cdot a_{lb} \cdot e^{-i\omega c_{lb}} + a_{mb} \cdot e^{-i\omega c_{mb}})\hat{s}_1(\omega). \end{aligned} \quad (10)$$

Here, the symbol ^ represents the Fourier domain. If we eliminate $\hat{s}_1(\omega)$ from Eqs. (9) and (10) and normalize the equations using $a_{la}e^{-i\omega c_{la}}$, we can obtain the following equation:

$$\begin{aligned} \hat{z}(\omega) &= \hat{y}_2(\omega) \cdot b_1 \cdot e^{-i\omega\tilde{d}_1} + \hat{y}_1(\omega) \cdot b_2 \cdot e^{-i\omega\tilde{d}_2} \\ &\quad + (\hat{y}_1(\omega) \cdot \gamma - \hat{y}_2(\omega)) \cdot b_3 \cdot e^{-i\omega\tilde{d}_3} \\ &\quad + \hat{y}_2(\omega) \cdot b_4 \cdot e^{-i\omega\tilde{d}_4} + \hat{y}_1(\omega) \cdot b_5 \cdot e^{-i\omega\tilde{d}_5}, \end{aligned} \quad (11)$$

where $\hat{z}(\omega) = \hat{y}_1(\omega) \cdot \gamma - \hat{y}_2(\omega)$ and $b_1 = a_{ka}/a_{la}$, $b_2 = -a_{ma}/a_{la}$, $b_3 = -a_{lb}/a_{la}$, $b_4 = a_{kb}/a_{la}$, $b_5 = -a_{mb}/a_{la}$, $\tilde{d}_1 = c_{ka} - c_{la}$, $\tilde{d}_2 = c_{ma} - c_{la}$, $\tilde{d}_3 = c_{lb} - c_{la}$, $\tilde{d}_4 = c_{kb} - c_{la}$ and $\tilde{d}_5 = c_{mb} - c_{la}$. $b_j (j = 1, \cdots, 5)$ represent the real-valued damping coefficient ratios and satisfy the following relationship:

$$b_1 < 0,\ b_2 > 0,\ b_3 < 0,\ b_4 < 0,\ b_5 > 0. \quad (12)$$

Furthermore, $\tilde{d}_j (j = 1, \cdots, 5)$ represent the discretized number of steps. In Eq. (11), $\hat{z}(\omega)$ is known, whereas b_j and d_j are unknown.

Next, we discretize Eq. (11). Let N be the total number of discretized steps and Δt and $\Delta\omega$ be the intervals of time and frequency, respectively; these indicate that

$$t = r \cdot \Delta t (r = 0, 1, 2, \cdots, N),$$
$$\omega = u \cdot \Delta\omega (u = 0, 1, 2, \cdots, N/2)$$

and $\tilde{d}_j = d_j \cdot \Delta t$. Moreover, let $z_u \overset{\leftarrow}{=} \hat{z}(u \cdot \Delta\omega)$, $y_{u1} \overset{\leftarrow}{=} \hat{y}_1(u \cdot \Delta\omega)$ and $y_{u2} \overset{\leftarrow}{=} \hat{y}_2(u \cdot \Delta\omega)$, and let p_{nj} be the value that is substituted into $u = N/2^n (n = 1, \cdots, 5)$ for the exponential part $\exp(-iu(2\pi/N)d_j)(j = 1, \cdots, 5)$. Then, the following equation holds:

$$\begin{aligned} z_{N/2^n} &= y_{N/2^n \cdot 2} \cdot b_1 \cdot p_{n1} + y_{N/2^n \cdot 1} \cdot b_2 \cdot p_{n2} \\ &\quad + (y_{N/2^n \cdot 1} \cdot \gamma - y_{N/2^n \cdot 2}) \cdot b_3 \cdot p_{n3} \\ &\quad + y_{N/2^n \cdot 2} \cdot b_4 \cdot p_{n4} + y_{N/2^n \cdot 1} \cdot b_5 \cdot p_{n5}. \end{aligned} \quad (13)$$

Then, we introduce the matrix representation

$$\boldsymbol{Pb} = z, \quad (14)$$

where $\boldsymbol{P} = [q_{nj}]$, $\boldsymbol{b} = [b_1\ b_2\ b_3\ b_4\ b_5]^{\mathrm{T}}$ and $z = [z_{N/2}\ z_{N/2^2}\ z_{N/2^3}\ z_{N/2^4}\ z_{N/2^5}]^{\mathrm{T}}$. Here, the entries of matrix $\boldsymbol{P}$ are $q_{nj} = y_{N/2^n \cdot m} p_{nj}$ $(m = 2 (j = 1, 4), m = 1 (j = 2, 5))$ and $q_{nj} = (y_{N/2^n \cdot 1}\gamma - y_{N/2^n \cdot 2}) p_{nj} (j = 3)$. z is known. Two candidates are considered as entries of $\boldsymbol{P}$ on the basis of the relations given by the argument, although their values are known. We substitute each candidate in to Eq. (14) and calculate b_j such that it satisfies the condition of a real value and Eq. (12).

Then, for $p_{nj} (n \ge 6, j = 1, \cdots, 5)$, Eq. (13) is considered to have 2^5 possible candidates when $n = N/2^6$. We substitute each candidate into Eq. (13) and calculate a combination that satisfies Eq. (13). In this way, $\mathrm{mod}(d_j, 2^6)$ is obtained. We apply the above-mentioned calculation in the same way for $n = N/2^7$, $n = N/2^8$, $\cdots$. For example, when the analytical area is 10 m in every direction and the sampling frequency is 44,100 Hz, we can assume $|d_j| \le 2^{11}$. Thus, we should calculate up to $n = N/2^{12}$; it is for this value that $\mathrm{mod}(d_j, 2^{12})$ is equivalent to d_j. If $\mathrm{mod}(d_j, 2^{12})$ is not less than 2^{11}, then $d_j = m_j - 2^{12}$ $(m_j = \mathrm{mod}(d_j, 2^{12}))$.

When k and l are fixed and m is varied, by repeating the same operations as applied to Eqs. (6) to (14), the damping coefficient ratios and relative time lags of s_1 can be obtained. Moreover, we can specify the location of s_1 in the same way as for s_2.

2.6. Separation of s_1 and s_2

Let the separation data of $s_1(t)$ and $s_2(t)$ be

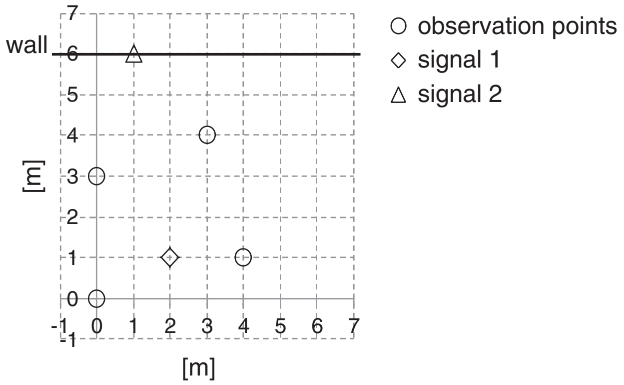

Fig. 1 Setting locations.

Table 1 Setting locations.

	coordinates [m]
s_1	$(2, 1)$
s_2	$(1, 6)$
x_1	$(0, 0)$
x_2	$(4, 1)$
x_3	$(0, 3)$
x_4	$(3, 4)$

$$\tilde{s}_1(t) = a_{l11}s_1(t),\ \tilde{s}_2(t) = a_{l21}s_2(t) \quad (1 \leq l \leq M). \quad (15)$$

In this paper, we calculate the values of $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$ instead of $s_1(t)$ and $s_2(t)$, respectively. We substitute Eq. (15) for Eq. (1) and perform a Fourier transform.

$$\hat{x}_k(\omega) = a_{k11}/a_{l11}e^{-i\omega c_{k11}}\hat{\tilde{s}}_1(\omega) + a_{k12}/a_{l11}e^{-i\omega c_{k12}}\hat{\tilde{s}}_1(\omega) + a_{k21}/a_{l21}e^{-i\omega c_{k21}}\hat{\tilde{s}}_2(\omega) \quad (16)$$

$(1 \leq k \leq M)$. $\hat{\tilde{s}}_1(\omega)$ and $\hat{\tilde{s}}_2(\omega)$ are obtained by solving Eq. (16) for all frequencies. $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$ can be calculated by performing the inverse Fourier transform.

3. Numerical test

3.1. Setting

Figure 1 and Table 1 describe the locations of the two signal sources and the observation points for $M = 4$. We derive signal data from a collection of sounds recorded in digital video disc (DVD) format [6]. Assumptions in this method are shown in Table 2. The sampling frequency is 44,100 Hz, and the total duration is approximately 2.97 s. We assumed that the damping coefficients are proportional to the reciprocal of the propagation distance and set them as such. We then collected observation data for the signal data. Here, the speed of sound is considered to be 340 m/s. Figures 2 and 3 show signal data in the time domain and the time-frequency domain, respectively. From Fig. 3, s_2 is found to have values in a wide time-frequency band, and we see that s_2 includes s_1. In the following discussion, the observed signal data $x_k(t)$, the locations of $x_k (1 \leq k \leq M)$ and the number of signals $(j = 1, 2)$ are the only known data. In this case, we can specify the locations of signal sources and separate these signals.

Table 2 Assumptions in this method.

propagation system	homogeneity invariability of time
kind of microphone	nondirectional microphone
placement of observation points	arbitrary placement

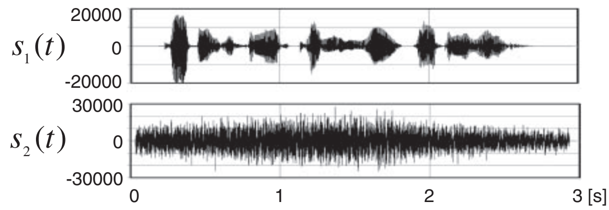

Fig. 2 Signal data in time domain.

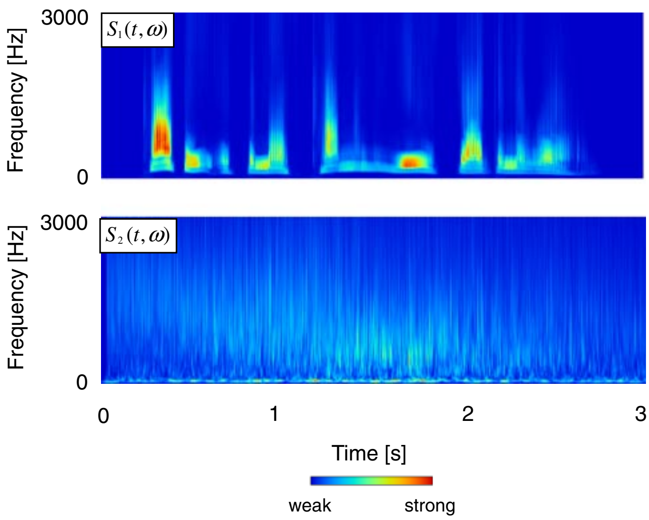

Fig. 3 Signal data in time-frequency domain.

3.2. Results for s_2

Table 3 lists the results of calculations of the damping coefficient ratios, time lags and the specification of the location of s_2. Compared with the set point, the location error is approximately 14.4 mm. In addition, when we find time lags, discretization errors may occur. In this numerical test, we normalize c_{121} and calculate c_{k21}. Errors in the time lags are 2 steps.

3.3. Results for s_1

Table 4 lists the results of calculations of damping coefficient ratios, time lags and the specification of location of s_1. Compared with the set point, the location error is approximately 2.72 mm.

3.4. Results of separation for s_1 and s_2

We separate s_1 and s_2 using their respective damping coefficient ratios and time lags that were obtained. To compare the separation signals and the original signals, we calculate errors using the following equation.

Table 3 Results for s_2.

	a_{221}/a_{121}	a_{321}/a_{121}	a_{421}/a_{121}	
calculation	1.043185174	1.923538442	2.150581335	
setting	1.043185168	1.923538406	2.150581317	
	$c_{121} - c_{221}$	$c_{121} - c_{321}$	$c_{121} - c_{421}$	
calculation [step]	33	379	422	
setting [step]	33	379	422	
location of s_2 [m]	(0.99640, 6.0139)			
	c_{121}	c_{221}	c_{321}	c_{421}
calculation [step]	791	758	412	369
setting [step]	789	756	410	367

Table 4 Results for s_1.

	a_{211}/a_{111}	a_{311}/a_{111}	a_{411}/a_{111}	
calculation	1.118034	0.7905698	0.7071054	
setting	1.118034	0.7905694	0.7071068	
	$c_{111} - c_{211}$	$c_{111} - c_{311}$	$c_{111} - c_{411}$	
calculation [step]	31	−77	−120	
setting [step]	31	−77	−120	
location of s_1 [m]	(2.0024, 0.99871)			
	c_{111}	c_{211}	c_{311}	c_{411}
calculation [step]	290	259	367	410
setting [step]	290	259	367	410

$$E_{rr} = \frac{\sum_{r=1}^{N} (s_j(r \cdot \Delta t) - \tilde{s}'_j(r \cdot \Delta t))^2}{\sum_{r=1}^{N} s_j(r \cdot \Delta t)^2} \quad (j = 1, 2) \qquad (17)$$

Here, $\tilde{s}'_j(t)$ is $1/a_{1j1}$ multiplied by $s_j(t)$. When $j = 1$, $E_{rr} = 5.2 \times 10^{-7}$ and when $j = 2$, $E_{rr} = 4.6 \times 10^{-6}$. Therefore, we can confirm that the errors are very small.

4. Conclusion

In this paper, we described a method for the separation and specification of a source signal that includes a reflected signal and another signal. Moreover, our numerical tests showed that results were highly accurate.

Future topics include the examination of multiple reflections and the establishment of experimental techniques that use real-world observation data.

Acknowledgement

This work was partially supported by JSPS. KAKENHI (C) 23560702.

References

[1] F. Sasaki, S. Ueda, M. Yasuoka and O. Tanaka, "Blind source separation and sound source localization on time-frequency domain considering time lag information: Formulation and numerical test," *J. Environ. Eng. AIJ*, **74**, 545–552 (2009) (in Japanese).

[2] F. Sasaki, M. Yanai, O. Tanaka and M. Yasuoka, "Separation and specification of location of source signal embedded in noise using BSS Part 1: Mathematical formulation," *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 1137–1140 (2011) (in Japanese).

[3] M. Yanai, F. Sasaki, O. Tanaka and M. Yasuoka, "Separation and specification of location of source signal embedded in noise using BSS Part 2: Improvement of computation time and numerical test," *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 1141–1144 (2011) (in Japanese).

[4] F. Sasaki, T. Maeda and M. Yamada, "Study of time history data using wavelet transform," *Struct. Eng.*, **38B**, 9–20 (1992) (in Japanese).

[5] I. Daubechies, *Ten Lectures on Wavelets* (CBMS 61, Society for Industrial and Applied Mathematics, 1992).

[6] Architectural Institute of Japan, *Kenchiku to Kankyou no Sound Library* (Gihodo Shuppan Co., Ltd., Tokyo, 2004) (in Japanese).