International Journal of Innovative Computing, Information and Control
Volume **9**, Number **4**, April **2013**

ICIC International ©2013 ISSN 1349-4198
pp. 1737–1757

REAL WORLD SOURCE SEPARATION BY COMBINING ICA AND VD-CDWT IN TIME-FREQUENCY DOMAIN

Zhong Zhang, Yasudake Aoki, Hiroshi Toda
Takashi Imamura and Tetsuo Miyake

Instrumentation Systems Laboratory
Toyohashi University of Technology
Toyohashi 441-8580, Japan
{ zhang; miyake; ima }@is.pse.tut.ac.jp

Received February 2012; revised June 2012

Abstract. *It is well known that in real-world source separation, environment noise removal must take into account complex reverberating sound and various noises. In this study, in order to improve the voice recognition accuracy in real-world source separation, a new method that uses independent component analysis (ICA) in the time-frequency domain, the variable density complex discrete wavelet transform (VD-CDWT), and the subspace method is proposed. Through comparison of the results by their signal noise ratios (SNR), the effectiveness of the proposed method is confirmed.*
Keywords: Independent component analysis, Wavelet transform, Sound source, Time-frequency analysis

1. **Introduction.** Recently, the demand for speech signals as a user interface for robotic applications, home electric appliances, and cellular phones has increased. Accordingly, sound recognition, noise reduction and other signal processing techniques that can be applied in a real environment are necessary. However, the performance of conventional sound recognition methods decreases profoundly when other sound sources (noise) exist in the real environment besides the object sound source. Therefore, the problem now becomes how to improve the performance of voice recognition. Generally, in a real environment, the prerequisites of sound recognition are that there are two or more sound sources, the location information is unknown and the observed signals are a mixture of the sounds with complex reverberating sound and various noises. In this case, a sound source separation method called blind source separation based on Independent Component Analysis (ICA) has been used [1, 2, 3].

ICA is a statistical method which guesses an original signal from the mixture signal even if the original signal and the transfer function are unknown under the assumption of statistical independence [4]. It is well known that if the original sounds are mixed in the real environment (in the time domain) then the observed sounds are a convolution mixture between the original sounds with a delay and a reverberation. In order to simplify this convolution mixture, it is a good idea to convert the signal from the time domain into the time-frequency domain and transform the convolution mixture into a linear mixture by a time-frequency analysis method. By doing this, the drawback of poor performance with unsteady sounds of the ICA can also be improved.

The time-frequency analysis method is usually a combination of the ICA and the Short Time Fourier Transform (STFT) [5] or the ICA and the Discrete Wavelet Transform (DWT) [6], respectively. Hereafter, the former is called the ICA+STFT and the latter

is called the ICA+DWT. The STFT is probably the most common approach for time-frequency analysis. It subdivides the signal into short time segments (it is the same as using a small window to divide the signal), and a discrete Fourier transform is computed for each of these. For each frequency component, however, the window length is fixed. So it is impossible to choose an optimal window for each frequency component, that is, the STFT is unable to obtain optimal analysis results for individual frequency components [7]. On the other hand, the DWT that uses Mattal's fast algorithm also has a drawback of lacking shift invariance, although it can solve the window width problem and obtain optimal frequency resolution for each frequency component [8]. Fortunately, in order to improve this fault, a Complex Discrete Wavelet Transform (CDWT) [9] was proposed and it has been applied to blind source separation by using the ICA [10] (hereafter, it is called the ICA+CDWT). As a result, the maximum signal separation performance having SNR = 17.0[dB] (Signal to noise ratio: SNR) in the condition of 0[ms] reverberation time was obtained and the separation performance from the traditional ICA+STFT was improved about 3[dB].

In order to achieve blind source separation in the real world, Saruwatari et al. [11, 12] proposed a new algorithm by combining the ICA+STFT and beamforming, and obtained a maximum signal separation performance of about 9.5[dB] (Noise reduction rate: NRR) with a reverberation time of 0.30[sec] and the performance improved about 1.5[dB] compared with the traditional ICA+STFT. Sawada et al. [13] proposed a two-stage processing technique where a spatial filter is first employed in each frequency bin and time-frequency masking is then used to improve the performance further. To obtain the spatial filter, the ICA+STFT is used and then the component of the target source is selected. Time-frequency masks in the second stage are obtained by calculating the angle between the basis vector corresponding to the target source and a sample vector. The experiments showed good results for extracting a dominant source out from six interferences with only 2 or 3 microphones, in a room in which the reverberation time was 0.13[sec]. Generally, in the office and other real world environments the reverberation time is greater than 0.30[sec]. Additionally, it is thought that the separated target source signal having an SNR of 10[dB] or larger is necessary to obtain good sound recognition performance. However, the separation of the sound source in the real world is very difficult because of complex reverberation that occurs by various noises.

In this study, therefore, in order to improve the sound source separation effect in the real world, we propose a new sound source separation method, in which the subspace method (SSM) using principal component analysis and blind source separation techniques based on the ICA in the time-frequency domain are combined. Here, the SSM is used to decrease sound reverberation. ICA in the time-frequency domain is combined with ICA and a new Variable Density Complex Discrete Wavelet Transform (VD-CDWT) [14], which can divide sound into a variable sub-band instead of the octave band. Furthermore, in order to confirm the effectiveness of the proposed method, simulation of separation experiments is carried out and a significant finding is obtained.

2. Review of the ICA Using the CDWT(ICA+CDWT).

Here, the case of N sound sources and N microphones is considered for simplicity. First of all, the N sound signals $\vec{x}(t) = [x_1(t), x_2(t), \cdots, x_N(t)]^T$ are observed by microphones from N sound sources $\vec{s}(t) = [s_1(t), s_2(t), \cdots, s_N(t)]^T$, where t denotes time. The relation between the observed signal $\vec{x}(t)$ and the sound source $\vec{s}(t)$ is as follows:

$$\vec{x}(t) = A(t) * \vec{s}(t), \tag{1}$$

where $*$ denotes convolution, $A(t)$ is the impulse response matrix that is from the sound sources to the microphones and can be shown as the following equation:

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1N}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2N}(t) \\ \vdots & \vdots & \ddots & \\ a_{N1}(t) & \cdots & \cdots & a_{NN}(t) \end{bmatrix}. \tag{2}$$

If one transforms the signal $\vec{x}(t)$ from the time domain to the time-frequency domain by the CDWT then the (1) can be expressed as (3), in which the convolution of the sound source and the impulse response was changed to simple multiplication [10].

$$\vec{X}(\omega, T) = \hat{A}(\omega)\vec{S}(\omega, T), \tag{3}$$

where ω is the frequency and T the time in the time-frequency domain. Next, whitening of the observed signal is carried out as follows:

$$\vec{\hat{X}}(\omega, T) = Q(\omega)\vec{X}(\omega, T), \tag{4}$$

where $\vec{\hat{X}}(\omega, T)$ is a whitened signal matrix and $Q(\omega)$ a whitened mixture, which can be obtained from the observed mixture signal $\vec{X}(\omega, T)$ in each frequency. Then, the ICA is carried out by the whitened signal matrix using the Fast-ICA [1], in which the separation matrix $W(\omega)$ can be presumed. As a result, the separated signal $u(\omega, T)$ shown in (5) can be obtained.

$$\vec{U}(\omega, T) = W(\omega)\vec{\hat{X}}(\omega, T). \tag{5}$$

Finally, after solving the scaling and the permutation problems, the sound sources of each microphone, which are presumed to be in the time-frequency domain, are reconstructed by applying the inverse-complex discrete wavelet transform (ICDWT) to (5) and the final separated signals can be obtained. Through comparison of the results obtained by the method (ICA+CDWT) shown above according to the signal-to-noise ratio (SNR) with traditional methods of the ICA in the time-frequency domain by the STFT and the DWT, the effectiveness of the proposed method was confirmed [10]. However, some problems, such as the influence of reverberation in the real world, hinder the practical use of the technique.

3. New Approach for Real World Source Separation.

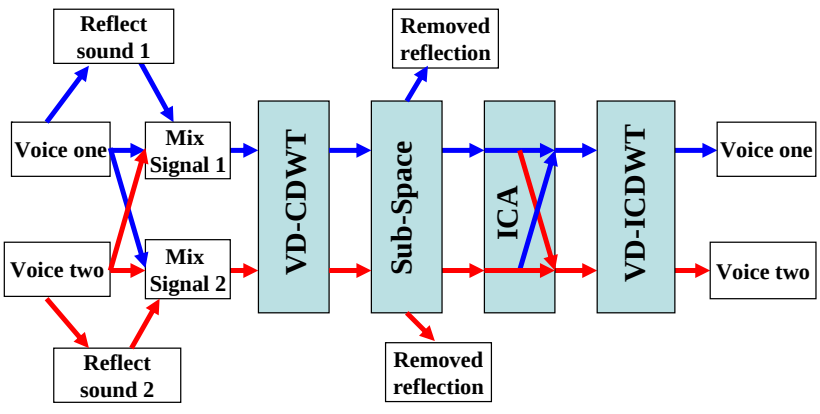
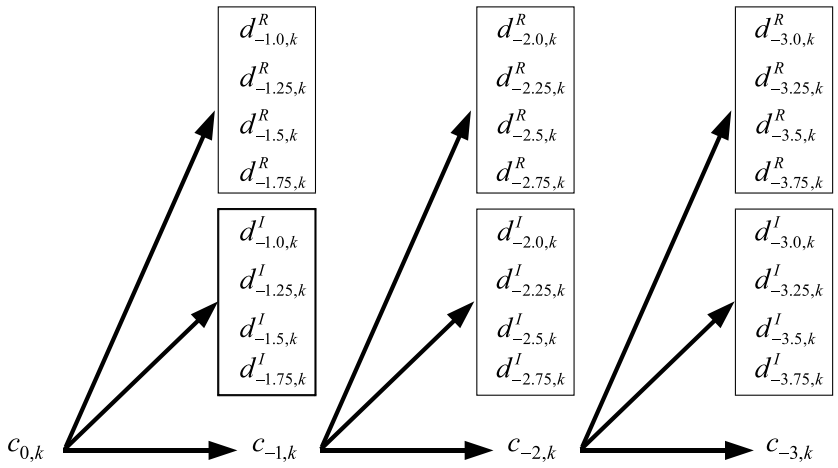
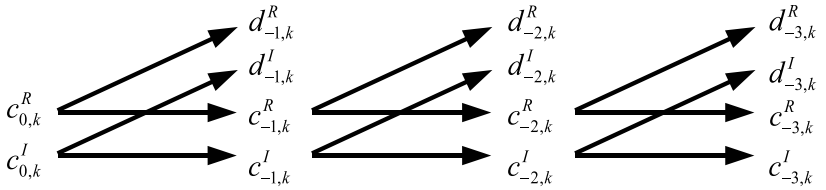


FIGURE 1. Flowchart of the new approach



(a) Decomposition algorithm of VD-CDWT



(b) Decomposition algorithm of CDWT

FIGURE 2. Decomposition tree of the VD-CDWT and the PTI-CDWT

3.1. **Processing flow.** By the ICA using the CDWT (ICA+CDWT) shown in Section 2, the environmental noises, such as all sounds except the object sound and the noises can be removed when there is no reverberation sound, but this cannot be done well when there is sound reverberation. Therefore, it is necessary to remove the reflected sound by another technique before performing this process. In this study, this problem is solved by using the SSM. Furthermore, in order to improve the voice separation effect in the real world, the VD-CDWT is used instead of the CDWT. Figure 1 shows the flowchart of the proposed new separation processing (ICA+SSM+VD-CDWT). First of all, the mixed signals $x_n(t)$ that are measured by the microphones are transformed to the time-frequency domain by using the VD-CDWT and the $X_n(\omega, T)$ can be obtained. Here, the case of two sound sources ($n = 1, 2$) and two microphones is considered for simplicity. The VD-CDWT can divide an octave frequency band into the number $N \geq 1$ of filter banks [14]. Different to this, in the case of the CDWT, the octave frequency band is used in all the frequency areas like with the traditional DWT. Then the reverberation components of the mixed signal are removed by using the SSM. Furthermore, the voice signal and environmental noise are separated by ICA. Finally, after solving some problems peculiar to the ICA, such as the solving problem and permutation problem, the original voice signal can be obtained by inverse VD-CDWT.

3.2. **Variable density complex discrete wavelet transform.** We have already proposed the variable density complex discrete wavelet transform (VD-CDWT) and have shown the details of the theory in [14, 15, 16]. Here we simply introduce the calculation algorithm and show the decomposition tree of the VD-CDWT in Figure 2(a), where as an example, an octave frequency band has been divided into the number $N = 4$ of filter banks. For comparison to the VD-CDWT, the decomposition tree of the PTI-CDWT is shown in Figure 2(b). Generally, the VD-CDWT is calculated as follows:

1) First, from the target digital signal $\{x_n\}$, the scaling coefficients $\{c_{0,k}\}$ of level 0 are obtained by

$$x(t) = \sum_k c_{0,k}\ \phi_{0,k}(t), \tag{6}$$

$$c_{0,k} = \frac{1}{2}\sum_l \overline{\phi_{0,k}(l)}\ x_l, \tag{7}$$

$$\phi_{0,k}(t) = \phi\left(t - \frac{1}{2}k\right). \tag{8}$$

Note that the scaling function $\phi(t)$ in (8) is Meyer's scaling function [8] whose details are shown in the Appendix, and $\phi_{0,k}(t)$ in (8) is the k-th scaling function of level 0 arrayed with the interval 0.5, and $\overline{\phi_{0,k}(l)}$ in (7) is the complex conjugate of $\phi_{0,k}(l)$. Additionally, $x_n = x(n)$ (n is an integer) holds, so the function $x(t)$ in (6) interpolates the target digital signal $\{x_n\}$.

2) Second, from the scaling coefficients $\{c_{0,k}\}$ of level 0 in (7), by using the decomposition sequences $\{b_k^{R_n}\}$, $\{b_k^{I_n}\}$ ($n = 1, 2, \cdots, N$: N is a positive integer), the real and imaginary parts of the wavelet coefficients $\{d^R_{-1-(n-1)/N,k}\}$ and $\{d^I_{-1-(n-1)/N,k}\}$ of level $-1-(n-1)/N$ are obtained by

$$d^R_{-1-(n-1)/N,k} = \sum_l b_l^{R_n} c_{0,\{2p_{-1-(n-1)/N}\}\times k-l}, \quad n = 1, 2, \cdots, N, \tag{9}$$

$$d^I_{-1-(n-1)/N,k} = \sum_l b_l^{I_n} c_{0,\{2p_{-1-(n-1)/N}\}\times k-l}, \quad n = 1, 2, \cdots, N. \tag{10}$$

Note that the decomposition sequences $\{b_k^{R_n}\}$, $\{b_k^{I_n}\}$ ($n = 1, 2, \cdots, N$) in (9), (10) are shown in the Appendix, and $p_{-1-(n-1)/N}$ ($n = 1, 2, \cdots, N$) are the intervals between wavelets of level $-1-(n-1)/N$, and when $N = 4$, these $p_{-1-(n-1)/N}$ are represented as

$$p_{-1} = 2, \quad p_{-1.25} = 2.5, \quad p_{-1.5} = 3, \quad p_{-1.75} = 3.5. \tag{11}$$

3) Next, from the scaling coefficients $\{c_{0,k}\}$ of level 0 in (7), with the decomposition sequence $\{a_k\}$, the scaling coefficients $\{c_{-1,k}\}$ of level -1 are obtained as follows:

$$c_{-1,k} = \sum_l a_l c_{0,2k-l}. \tag{12}$$

Note that the decomposition sequences $\{a_k\}$ are shown in the Appendix.

4) The coefficients of level -2 or later are obtained in the same manner as 1)-3), that is the decomposition calculations from the scaling coefficients $\{c_{j+1,k}\}$ of level $j+1$ to the wavelet and scaling coefficients of level $j-(n-1)/N$ (j is an integer and $n = 1, 2, \cdots, N$) are represented as

$$d^R_{j-(n-1)/N,k} = \sum_l b_l^{R_n} c_{j+1,\{2p_{-1-(n-1)/N}\}\times k-l}, \quad n = 1, 2, \cdots, N, \tag{13}$$

$$d^I_{j-(n-1)/N,\,k} = \sum_l b_l^{I_n} c_{j+1,\{2p_{-1-(n-1)/N}\}\times k-l}, \quad n = 1, 2, \cdots, N, \tag{14}$$

$$c_{j,k} = \sum_l a_l c_{j+1,2k-l}. \tag{15}$$

It is very important that $p_{-1-(n-1)/N}$ in (13) and (14) is never $p_{j-(n-1)/N}$.

5) Finally, with the redundant coefficient $r_{j-(n-1)/N}$ of level $j-(n-1)/N$ shown in (16),

$$r_{j-(n-1)/N} = p_{j-(n-1)/N} m^{N(j+1)-n}(m-1), \qquad n = 1, 2, \cdots, N, \tag{16}$$

$$m = 2^{1/N}, \tag{17}$$

the transform to level L (L is a negative integer) is represented as

$$x(t) = \sum_{j=L}^{-1} \sum_{n=1}^{N} \left[r_{j-(n-1)/N} \sum_{k} \left\{ d^R_{j-(n-1)/N,k} \psi^R_{j-(n-1)/N,k}(t) + d^I_{j-(n-1)/N,k} \psi^I_{j-(n-1)/N,k}(t) \right\} \right] + \sum_{k} c_{L,k} \phi_{L,k}(t). \tag{18}$$

Note that, when $N = 4$, from (11), the interval $p_{j-(n-1)/N}$ between wavelets of level $j-(n-1)/N$ in (16) is obtained by

$$p_{j-(n-1)/4} = 2^{-j-1} p_{-1-(n-1)/4}, \quad n = 1, 2, 3, 4, \tag{19}$$

and from (16), (17) and (19), the redundant coefficient $r_{j-(n-1)/N}$ is obtained by

$$r_{j-(n-1)/4} = r_{-1-(n-1)/4} = p_{-1-(n-1)/4} m^{-n}(m-1), \quad n = 1, 2, 3, 4, \tag{20}$$

and $\psi^R_{j-(n-1)/N,k}(t)$, $\psi^I_{j-(n-1)/N,k}(t)$ in (18) are the real and imaginary parts of the k-th wavelet of level $j-(n-1)/N$, and $\phi_{L,k}(t)$ in (18) is the k-th scaling function of level L, and these wavelet and scaling functions are shown in detail in the Appendix.

6) Next, the inverse transform is carried out as follows:

$$c_{j+1,k} = \sum_{n=1}^{N} \sum_{l} \left[r_{j-(n-1)/N} \left\{ h^{R_n}_{k-\{2p_{-1-(n-1)/N}\}\times l} \, d^R_{j-(n-1)/N,l} + h^{I_n}_{k-\{2p_{-1-(n-1)/N}\}\times l} \, d^I_{j-(n-1)/N,l} \right\} \right] + \sum_{l} g_{k-2l} \, c_{j,l}. \tag{21}$$

Note that, from the decomposition sequences $\{b^{R_n}_k\}$, $\{b^{I_n}_k\}$ and $\{a_k\}$, which are shown in the Appendix, the reconstruction sequences $\{h^{R_n}_k\}$, $\{h^{I_n}_k\}$ and $\{g_k\}$ in (21) are obtained by

$$h^{R_n}_k = \overline{b^{R_n}_{-k}}, \quad h^{I_n}_k = \overline{b^{I_n}_{-k}}, \quad g_k = \overline{a_{-k}}. \tag{22}$$

Repeating (21), the scaling coefficients $\{c_{0,k}\}$ of level 0 are obtained, and from (6), (7), (8) and $x_n = x(n)$ (n is an integer), the digital signal $\{x_n\}$ can be obtained by

$$x_n = \sum_{k} c_{0,2n-k} \phi\left(\frac{k}{2}\right). \tag{23}$$

7) Furthermore, if in some level j, octave analysis is required, the following decomposition calculation of the traditional PTI-CDWT can be carried out instead of (13) and (14).

$$d^R_{j,k} = \sum_{l} b^R_l c_{j+1,2(2k-l)}, \tag{24}$$

$$d^I_{j,k} = \sum_{l} b^I_l c_{j+1,2(2k-l)+1}. \tag{25}$$

Note that the decomposition sequences $\{b_k^R\}$ and $\{b_k^I\}$ are shown in the Appendix. Additionally, the inverse transform of (24), (25) and (15) is represented as

$$c_{j+1,k} = \begin{cases} \sum_l h^R_{k/2-2l}\, d^R_{j,l} + \sum_l g_{k-2l}\, c_{j,l}, & k = \text{even integer}, \\ \sum_l h^I_{(k-1)/2-2l}\, d^I_{j,l} + \sum_l g_{k-2l}\, c_{j,l}, & k = \text{odd integer}. \end{cases} \tag{26}$$

Note that the reconstruction sequences $\{h_k^R\}$ and $\{h_k^I\}$ in (26) are obtained by

$$h_k^R = \overline{b_{-k}^R}, \quad h_k^I = \overline{b_{-k}^I}. \tag{27}$$

The complex wavelet coefficients obtained by using the equations above can be shown as follows:

$$d^c_{j-(n-1)/N,k} = d^R_{j-(n-1)/N,k} + i d^I_{j-(n-1)/N,k}, \tag{28}$$

where $d^c_{j-(n-1)/N,k}$ denotes a complex wavelet coefficient.

In the processing shown in Section 2, the VD-CDWT is carried out on the observed signal $x(t)$ shown in (1), and the $X(\omega, T)$ shown in (3) can be obtained. Here, the relation between $X(\omega, T)$ and the complex wavelet coefficients $d^c_{j-(n-1)/N,k}$ can be shown as follows:

$$X\left(\omega = 2^{j-(n-1)/N} \times 2\pi,\ T = p_{j-(n-1)/N}\Delta\tau\right) = d^c_{j-(n-1)/N,k}, \tag{29}$$

where ω denotes the angle frequency, $j-(n-1)/N$ the frequency level, T the time, k the digital time and $\Delta\tau$ the sampling interval.

3.3. **Principle of the subspace method.** The subspace method is used as preprocessing of the BSS that uses ICA. If the magnitude of the direct sound, which reaches the microphone directly from the sound source, is larger than that of the reverberating sound, then the direct sound can be considered to be a principal component. Oppositely, the reverberating sound can be considered the non-principal components and they can be decreased by using the subspace method. Strictly, the subspace method takes out the direct sound and the initial reverberating sound, and removes other reverberating sound. Moreover, this method has a fault where if the reverberating sound came from a direction near to that of the sound source, it cannot be removed because the reverberating sounds are confused with the sound source. In order to overcome this drawback, we assume one microphone is set up near the driver's mouth, and another one is set up away from a specific sound source.

First, the observation vector is defined by the next expression.

$$\vec{X}(\omega, t) = [X_1(\omega, t), X_2(\omega, t), \ldots, X_m(\omega, t)]^T, \tag{30}$$

where $X_m(\omega, t)$ denotes the wavelet transform of the signal $x_m(t)$ observed with microphone m. Next, space correlation procession $R(\omega)$ is defined as follows:

$$R(\omega) = E[\vec{X}(\omega, t)\vec{X}^H(\omega, t)]. \tag{31}$$

This develops in the generalization eigenvalue.

$$R = K\vec{V}\Lambda\vec{V}^{-1}, \tag{32}$$

where K is a correlation coefficient of the reverberation defined by $K = E[\vec{n}(t)\vec{n}^H(t)]$ and it is unknown because it cannot observe $\vec{n}(t)$ alone. Therefore, it is assumed that the noise element is spatially white and the standard eigenvalue expansion $K = I$ (I: unit procession) is substituted generally. $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_M)$ is a procession which has the eigenvalue in the corner element, $\vec{V} = [e_1, e_2, \ldots, e_M]^T$ is a procession which has the

eigenvector corresponding to each eigenvalue in the row vector. Such processions as these are divided as follows:

$$\vec{V} = [e_1, e_2, \ldots, e_N | e_{N+1}, \ldots, e_M]^t = [\vec{V_s} | \vec{V_n}]^T, \tag{33}$$

$$\begin{aligned} \Lambda &= diag(\lambda_1, \ldots, \lambda_N | \lambda_{N+1}, \ldots, \lambda_M) \\ &= \begin{bmatrix} \Lambda_s & 0 \\ 0 & \Lambda_n \end{bmatrix}. \end{aligned} \tag{34}$$

Therefore, the sound source separation filter using the SSM can be shown by the next expression.

$$\begin{aligned} Z &= \Lambda_s^{-1/2} V_s^H, \\ \Lambda_s &= diag(\lambda_1, \lambda_2, \ldots, \lambda_N). \end{aligned} \tag{35}$$

Here, in order to remove reverberating sound, Equation (36) is used instead of (4)

$$\hat{x}_n(\omega, T) = Q(\omega) Z(\omega) x_n(\omega, T), \tag{36}$$

where $Z(\omega)$ is obtained from (35).

4. **Separation Experiment.**

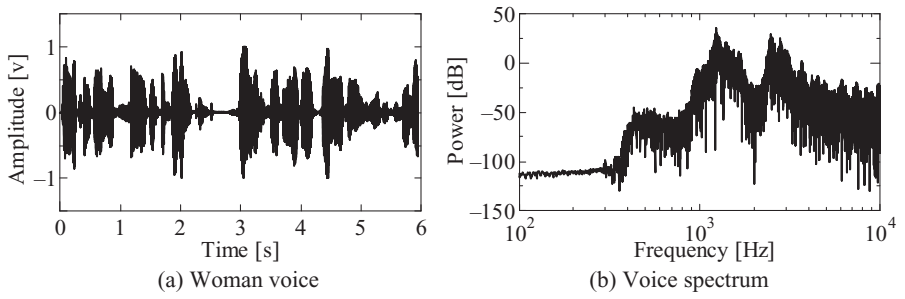


FIGURE 3. Example of a woman's voice signal [17]

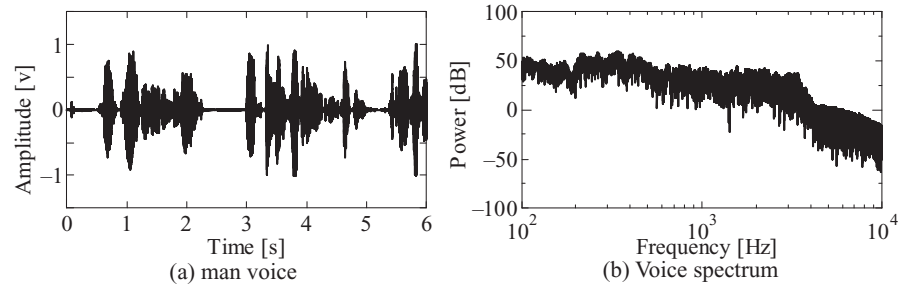


FIGURE 4. Example of a man's voice signal [17]

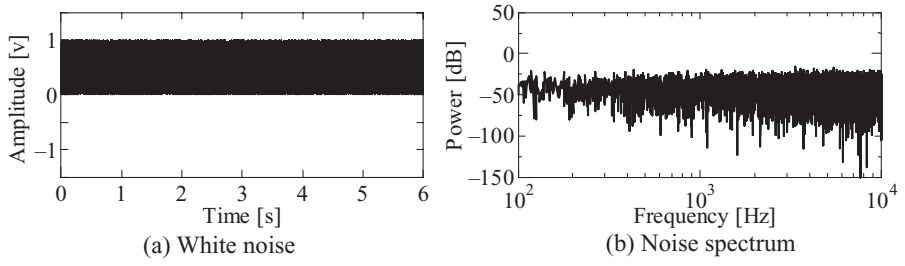


FIGURE 5. Example of a white noise signal

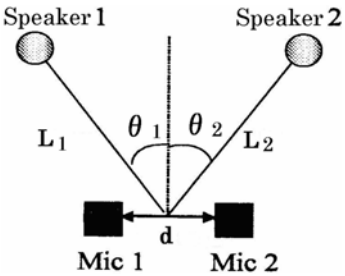


FIGURE 6. Microphone and sound source setting [17]

4.1. **Experimental conditions.** In order to confirm the effectiveness of the proposed method, echo sound removal performance of the SSM is first confirmed and then the sound source separation experiment by connecting ICA with the SSM filter and the VD-CDWT (ICA+SSM+VD-CDWT) is conducted. The separating sound is made by using the opened public sound data base RWCP, which is recorded in a real environment [17]. Figure 3 shows an example of a woman's voice and its power spectrum, Figure 4 shows an example of a man's voice and its power spectrum, and Figure 5 shows an example of a white noise signal and its power spectrum. These signals were measured in an anechoic room, with a sampling frequency of 48[kHz] and data length of 6[sec]. By comparing Figures 3(b), 4(b) and 5(b), we can see that in the case of the woman's voice, the main frequency components are almost entirely in the area from 300[Hz] to 5000[Hz] and in the case of the man's voice, the main frequency components are almost entirely in the area from 100[Hz] to 5000[Hz], but in the case of the white noise, all the frequency components are almost the same and they are distributed over the entire frequency area.

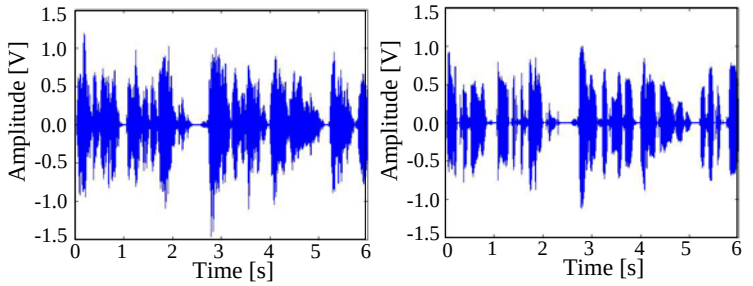
The experiment setting, such as microphone positions, the distance between the speakers and microphones are shown in Figure 6, where $L_1 = 2.11$[m], $L_2 = 2.14$, $\theta_1 = 16$[deg], $\theta_2 = 9$[deg] and $d = 0.211$[m].

4.2. **Reverberation reduction experiment.** Figure 7(a) shows an example of the reverberating sound that was made by the convolution integral between the woman's voice shown in Figure 3(a) and the impulse response between microphone 1 and speaker 1 with a reverberation time of 1.30[sec]. This simulates an echo chamber. Figure 7(b) shows the reverberation detection result that was obtained by using the SSM shown in Section 3.3. Here, in order to evaluate the effect of the reverberation detection, the following signal-noise ratio (SNR) is used.

$$SNR = 10 \log_{10} \frac{\sum s(t)}{\sum (u(t) - s(t))^2}, \tag{37}$$

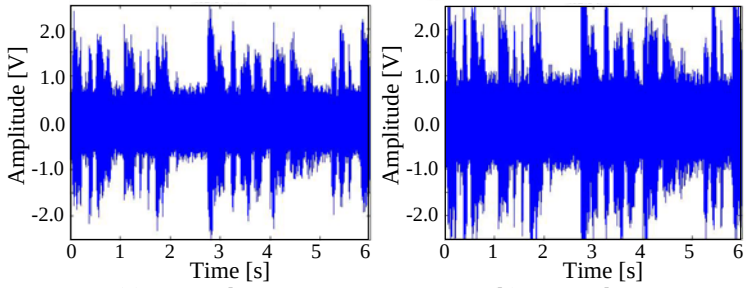
where $s(t)$ is the original woman's voice shown in Figure3(a), $u(t)$ the reverberating sound shown in Figure 7(a) or the reverberation detection signal shown in Figure 7(b). The results show that in the case of Figure 7(a), $SNR = 4.43$[dB], and in the case of Figure 7(b), $SNR = 10.35$[dB]. That is, by using the SSM, the influence of the reverberation has been improved by about 6[dB].

4.3. **Separation result of a female voice from white noise.** The voice of the woman shown in Figure 3(a) is mixed with the white noise shown in Figure 5(a), and the mixture results are shown in Figure 8, where (a) shows the mixture result in which microphone 1 is assumed to be near the sound source of the female's voice and (b) microphone 2 is assumed to be near the sound source of the white noise. Here the reverberation time of



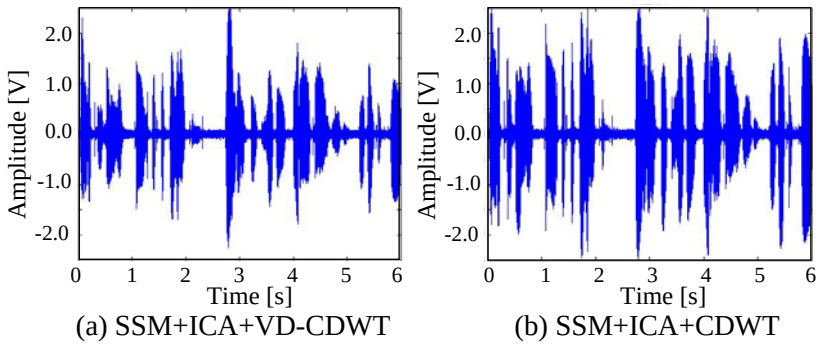
(a) Reverberation time 0.31[s] (b) Reverberation deletion result

FIGURE 7. Example of reverberation detection result by using the subspace method

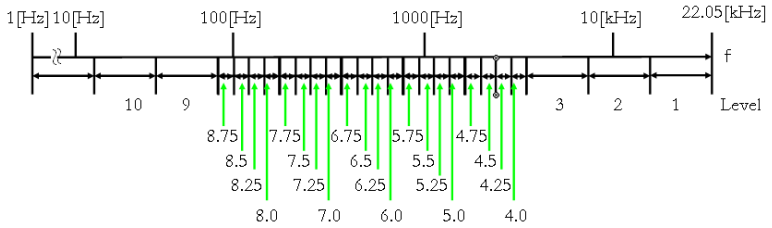


(a) Microphone 1 (b) Microphone 2

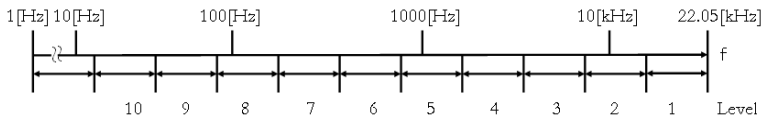
FIGURE 8. Example of observed signals made by sound signal and white noise



(a) SSM+ICA+VD-CDWT (b) SSM+ICA+CDWT



(c) Filter Band by using the VD-CDWT



(d) Filter Band by using the CDWT

FIGURE 9. Separated signal obtained by proposed method

the impulse response between the microphone and speaker is 0.31[sec]. This simulates the sound environment of an office or car interior.

Figure 9(a) shows the separation experiment result by using the proposed method that uses the ICA+SSM+VD-CDWT. For comparison, Figure 9(b) shows the separation result by using the ICA+SSM+CDWT. In addition, Figure 9(c) shows filter band division in the case of the VD-CDWT and Figure 9(d) shows the filter band division in the case of the CDWT. Because the main frequency components of the woman's voice are only in the area from 300[Hz] to 5000[Hz] shown in Figure 3(b), in the case of the VD-CDWT shown in Figure 9(c), we divided an octave frequency band into the number $N = 4$ of filter banks in the area from 100[Hz] to 5000[Hz]. Different to this, in the case of the CDWT as shown in Figure 9(d), the octave frequency band was used in the entire frequency area like in the traditional CDWT. And for both cases, the components of high frequency that are higher than 5[kHz] were removed when the inverse transform was performed.

In the case of the proposed method by using the ICA+SSM+VD-CDWT, $SNR = 12.11$[dB] was obtained. By comparison, in the case of the ICA+SSM+CDWT, $SNR = 10.21$[dB] was obtained. That is, by using the proposed method, the separation effect could be improved by about 2[dB], although we almost could not observe a difference between Figures 9(a) and 9(b) from the wave form of the separated signals. However, if we do not use SSM, that is, in the case of the ICA+CDWT, $SNR = -1.15$[dB] was obtained. And in the case of the ICA+SSM+CDWT, in which the components of high frequency that are higher than 5[kHz] were not cut off, $SNR = 5.24$[dB] was obtained. Therefore, the SSM is effective to detect reverberation.

4.4. **Separation result of a female voice from a man's voice.** The voice of the woman shown in Figure 3(a) is mixed with the man's voice shown in Figure 4(a). The mixture results are shown in Figure 10, where (a) shows the mixture result of microphone 1, which is assumed to be near the sound source of the female's voice and (b) microphone

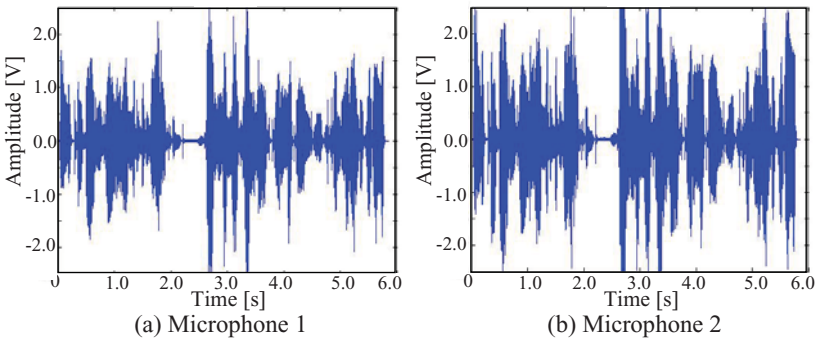


FIGURE 10. Example of observed signals made by voices of a woman and man

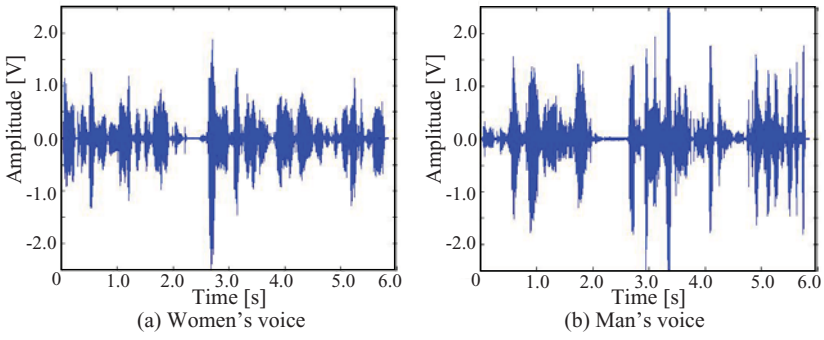


FIGURE 11. Separated signal obtained by ICA+SSM+VD-CDWT

TABLE 1. Separated results of woman's voice

Method	Separated results SNR [dB]	
	Mixd with man's voice 3.24[dB]	Mixed with white noise –0.62[dB]
ICA+SSM+CDWT	7.81	5.55
ICA+SSM+VD-CDWT	10.22	8.64
ICA+SSM+CDWT(noise cut)	10.54	10.21
ICA+SSM+VD-CDWT(noise cut)	11.74	12.11

2 is assumed to be near the sound source of the man's voice. Here the reverberation time of the impulse response between the microphone and speaker is 0.31[sec].

Figure 11 shows the separated voices obtained from the mixture signals by using the proposed method that uses the ICA+SSM+VD-CDWT, where (a) denotes the woman's voice and (b) the man's voice. Same as Figure 9, the octave frequency bands have been divided into the number $N = 4$ of filter banks in the area from 100[Hz] to 5000[Hz] and high frequency components above 5[kHz] have been removed when inverse transform is performed. The calculating condition and separation accuracy are shown in Table 1 based on the SNR. For comparison with the ICA+SSM+CDWT, the separation results obtained in Section 4.3 are also shown. From Table 1, we can see that in the case of the woman's voice, a maximum separation accuracy of $SNR = 11.74$[dB] could be obtained by ICA+SSM+VD-CDWT(noise cut). Compared with this, when the ICA+SSM+CDWT was used, the accuracy of the separated woman's voice was $SNR = 10.54$[dB]. That is, by using the proposed method, the separation effect can be improved about 1.2[dB], although this improved effectiveness does not hold in the case where the female's voice is mixed with white noise. This is because the difference of the power spectrum between the female voice and white noise is bigger than that between the female's voice and the man's voice. In addition, the separation accuracy of the man's voice is $SNR = 11.64$[dB] in the case of the ICA+SSM+VD-CDWT.

5. **Conclusions and Remarks.** In this study, we proposed a new sound source separation method, in which the subspace method (SSM) and the blind source separation technique based on the ICA in the time-frequency domain by the VD-CDWT are combined. In order to verify the effectiveness of our method, a voice in a noisy environment was synthesized by using the RWCP sound scene database [17], and the voice source separation performance of our method was verified by computer experiment. Furthermore, we applied the proposed method to voice separation in a real environment with a reverberation time of 0.31[sec] and obtained a significant finding. In the case of the woman's voice mixed with the white noise, the separated woman's voice with $SNR = 12.11$[dB] could be obtained. Additionally, in the case of the woman's voice mixed with the man's voice, the separated woman's voice with $SNR = 11.74$[dB] could be obtained.

However, it is not enough to evaluate the separation results by only using the SNR. Therefore, in the future we will try to use the index or another evaluation parameter. Furthermore, we want to apply this method for sound source separation in various real environments.

REFERENCES

[1] E. Bingham and A. Hyvarinen, A fast fixed-point algorithm for independent component analysis of complex-valued signals, *Int. J. of Neural Systems*, vol.10, no.1, pp.1-8, 2000.

[2] N. Hirai et al., A proposal of blind source separation using wavelet transform, *The Institute of Electronics, Information and Communication Engineers*, vol.104, no.456, pp.51-54, 2004.
[3] H. Sawada et al., Frequency domain-blind source separation, in *Speech Enhancement*, J. Benesty et al. (eds.), Springer, 2005.
[4] M. Khosravy, M. R. Asharif, R. Shomija and K. Yamashita, Image mosaic and semi-blind de-mosaic based on independent component analysis, *ICIC Express Letters, Part B: Applications*, vol.3, no.3, pp.689-694, 2012.
[5] H. Gotanda et al., Permutation correction and speech extraction based on split spectrum through fast-ICA, *Proc. of ICA2003*, pp.379-384, 2003.
[6] N. Hirai et al., A proposal of blind source separation using wavelet transform, *The Institute of Electronics, Information and Communication Engineers*, vol.104, no.456, pp.51-54, 2004.
[7] L. Cohen, *Time-Frequency Analysis*, Prentice-Hill PTR, 1995.
[8] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
[9] Z. Zhang, H. Toda, H. Fujiwara and F. Ren, Translation invariant RI-spline wavelet and its application on de-noising, *International Journal of Information Technology & Decision Making*, vol.5, no.2, pp.353-378, 2006.
[10] Z. Zhang, C. T. Miyake, T. Imamura, T. Enomoto and H. Toda, Blind source separation by combining independent component analysis with the complex discrete wavelet transform, *International Journal of Innovative Computing, Information and Control*, vol.6, no.9, pp.4157-4172, 2010.
[11] H. Saruwatari, T. Kawamura, T. Nishikawa, A. Lee and K. Shikano, Blind source separation based on a fast-convergence algorithm combining ICA and beamforming, *IEEE Transactions on Audio, Speech, and Language Processing*, vol.14, no.2, pp.666-678, 2006.
[12] Y. Mori, H. Saruwatari, T. Takatani, S. Ukai, K. Shikano, T. Hiekata, Y. Ikeda, H. Hashimoto and T. Morita, Blind separation of acoustic signals combining SIMO-model-based independent component analysis and Binary masking, *EURASIP Journal on Applied Signal Processing*, vol.2006, pp.1-7, 2006.
[13] H. Sawada, S. Araki, R. Mukai and S. Makino, Blind extraction of a dominant source signal from mixtures of many sources, *Proc. of ICASSP2005*, vol.III, pp.61-64, 2005.
[14] H. Toda and Z. Zhang, Variable-density complex discrete wavelet transform based on perfect translation invariance, *Proc. of the ICWAPR2008*, pp.711-716, 2008.
[15] H. Toda and Z. Zhang, Perfect translation invariance with a wide range of shapes of Hilbert transform pairs of wavelet bases, *Int. J. Wavelets Multiresolut. Inf. Processing*, vol.8, no.4, pp.501-520, 2010.
[16] H. Toda, Z. Zhang and T. Imamura, The design of complex wavelet packet transforms based on perfect translation invariance theorems, *Int. J. Wavelets Multiresolut. Inf. Processing*, vol.8, no.4, pp.537-558, 2010.
[17] *Real World Computing Partnership, RWCP Sound Scene Database in Real Acoustic Environment*, http://tosa.mri.co.jp/sounddb/indexe.htm.

Appendix. Firstly, the Fourier transform $\hat{f}(\omega)$ of the function $f(t)$ is defined by

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t} dt, \tag{38}$$

and the inverse Fourier transform is defined by

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega t} d\omega. \tag{39}$$

Next, the VD-CDWT, in which an octave frequency band has been divided into the number N (N is a positive integer) of filter banks, is based on the following Meyer scaling function [8] $\hat{\phi}(\omega)$:

$$\hat{\phi}(\omega) = \begin{cases} 1, & |\,\omega\,| \leq (2-m)\pi, \\ 0, & |\,\omega\,| \geq m\pi, \end{cases} \tag{40}$$

$$|\,\hat{\phi}(\omega - 2\pi)\,|^2 + |\,\hat{\phi}(\omega)\,|^2 = 1, \qquad (2-m)\pi < \omega < m\pi, \tag{41}$$

$$m = 2^{1/N}. \tag{42}$$

(41) is the condition for the area $(2-m)\pi <\mid \omega \mid < m\pi$, and the following function $\hat{\phi}(\omega)$ satisfies this condition:

$$\hat{\phi}(\omega) = \cos\left[\frac{\pi}{2}\,\nu\left(\frac{2-m}{2m-2}\left\{\frac{1}{(2-m)\pi}\mid\omega\mid -1\right\}\right)\right], \quad (2-m)\pi <\mid \omega \mid < m, \tag{43}$$

$$\nu(x) = x^4(35-84x+70x^2-20x^3). \tag{44}$$

The k-th scaling function $\phi_{j,k}(t)$ of level j (j is an integer) is defined by

$$\phi_{j,k}(t) = \sqrt{2}^{\,j}\,\phi(2^j t - k/2). \tag{45}$$

And now, the function values of $\phi(k/2)$ (k is an integer) are needed for the calculation of (7) and (23), so we list $\phi(k/2)$ when $N=4$ in Table 2.

Next, the real part $\psi^R(t)$ and the imaginary part $\psi^I(t)$ of the mother wavelet are defined by

$$\hat{\psi}^R(\omega) = \frac{1}{\sqrt{m-1}}\sqrt{\left|\hat{\phi}\left(\frac{\omega}{m}\right)\right|^2 - \left|\hat{\phi}(\omega)\right|^2}, \tag{46}$$

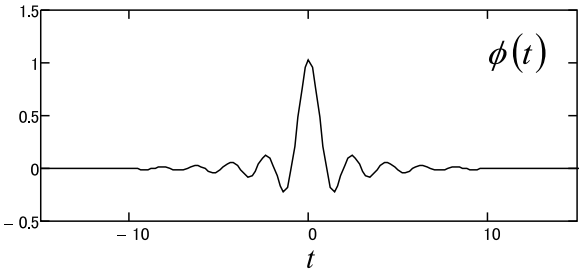
$$\hat{\psi}^I(\omega) = \begin{cases} i\hat{\psi}^R(\omega), & \omega \le 0, \\ -i\hat{\psi}^R(\omega), & \omega > 0. \end{cases} \tag{47}$$

The real part $\psi^R_{j-(n-1)/N,k}(t)$ and imaginary part $\psi^I_{j-(n-1)/N,k}(t)$ of the k-th wavelet of level $j-(n-1)/N$ (j is an integer and $n=1,2,\cdots,N$) are defined by

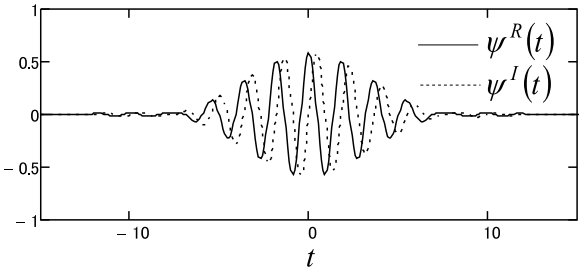
$$\psi^R_{j-(n-1)/N,k}(t) = m^{\frac{N(j+1)-n}{2}}\psi^R\left(m^{N(j+1)-n}(t-p_{j-(n-1)/N}k)\right), \tag{48}$$

$$\psi^I_{j-(n-1)/N,k}(t) = m^{\frac{N(j+1)-n}{2}}\psi^I\left(m^{N(j+1)-n}(t-p_{j-(n-1)/N}k)\right). \tag{49}$$

Note that $p_{j-(n-1)/N}$ is the interval between wavelets of level $j-(n-1)/N$, and when $N=4$, $p_{j-(n-1)/N}$ is obtained by (11) and (19).



(a) Scaling function



(b) Mother wavelet

FIGURE 12. VD-PTI complex wavelet ($N=4$)

TABLE 2. The scaling function $\phi(t)$ of VD-CDWT, $N = 4$, $\phi(t) = \phi(-t)$

t	$\phi(t)$	t	$\phi(t)$
0.0	1.02941521816	20.0	0.00000154228
0.5	0.63384587968	20.5	0.00007492424
1.0	–0.02899712778	21.0	–0.00002751131
1.5	–0.20401024503	21.5	–0.00004209741
2.0	0.02777395979	22.0	0.00003884351
2.5	0.11407026897	22.5	0.00001276248
3.0	–0.02583562017	23.0	–0.00003656956
3.5	–0.07320820224	23.5	0.00000802280
4.0	0.02332107290	24.0	0.00002550184
4.5	0.04924813784	24.5	–0.00001848897
5.0	–0.02040357585	25.0	–0.00001139661
5.5	–0.03347218600	25.5	0.00001973576
6.0	0.01727286850	26.0	–0.00000098236
6.5	0.02252579428	26.5	–0.00001468462
7.0	–0.01411662487	27.0	0.00000883244
7.5	–0.01480207004	27.5	0.00000686436
8.0	0.01110347608	28.0	–0.00001144278
8.5	0.00938610500	28.5	0.00000061904
9.0	–0.00836956176	29.0	0.00000974101
9.5	–0.00567597568	29.5	–0.00000572549
10.0	0.00600995769	30.0	–0.00000555387
10.5	0.00323038261	30.5	0.00000768416
11.0	–0.00407555954	31.0	0.00000085280
11.5	–0.00170441742	31.5	–0.00000682423
12.0	0.00257520972	32.0	0.00000282116
12.5	0.00082238014	32.5	0.00000417331
13.0	–0.00148216279	33.0	–0.00000465715
13.5	–0.00036609153	33.5	–0.00000098473
14.0	0.00074349468	34.0	0.00000459723
14.5	0.00016879142	34.5	–0.00000166818
15.0	–0.00029083055	35.0	–0.00000314987
15.5	–0.00011001857	35.5	0.00000314337
16.0	0.00005080650	36.0	0.00000109793
16.5	0.00010959959	36.5	–0.00000330033
17.0	0.00004604626	37.0	0.00000079431
17.5	–0.00012038316	37.5	0.00000241524
18.0	–0.00005887098	38.0	–0.00000199590
18.5	0.00012016567	38.5	–0.00000100144
19.0	0.00003287938	39.0	0.00000230620
19.5	–0.00010360231	39.5	–0.00000039373
		40.0	–0.00000183554

TABLE 3. The real part $\{b_k^{R_1}\}$ and imaginary part $\{b_k^{I_1}\}$ of decomposition sequences of VD-CDWT, $N = 4$, $b_k^{R_1} = b_{-k}^{R_1}$, $b_k^{I_1} = -b_{-k}^{I_1}$

k	$b_k^{R_1}$	$b_k^{I_1}$	k	$b_k^{R_1}$	$b_k^{I_1}$
0	0.86955280098	0.00000000000	40	–0.00003845040	0.00010563789
1	–0.21165674304	–0.79562184545	41	0.00012231081	0.00007248553
2	–0.60344670911	0.34323783128	42	0.00009806875	–0.00012807377
3	0.35315785330	0.36642901398	43	–0.00010298325	–0.00011144955
4	0.16465736547	–0.25518977720	44	–0.00012071558	0.00005160127
5	–0.10561963575	–0.04785291423	45	–0.00000358442	0.00012673362
6	–0.01850553700	–0.02951227943	46	0.00011940921	0.00004345686
7	–0.10676036691	0.04162700376	47	0.00006485573	–0.00009060907
8	0.07139965694	0.11980763492	48	–0.00004597748	–0.00007473681
9	0.09180890243	–0.07690106824	49	–0.00007715673	0.00000186339
10	–0.05303992986	–0.05472873382	50	–0.00002805858	0.00006901545
11	–0.03033967259	0.01442057686	51	0.00004722685	0.00004156395
12	–0.01963765743	0.02298521563	52	0.00004449623	–0.00001655152
13	0.02457029178	0.03768603330	53	0.00001171995	–0.00004165178
14	0.03978334226	–0.02481069943	54	–0.00003282354	–0.00002834464
15	–0.01876505709	–0.03309679852	55	–0.00003226164	0.00001685442
16	–0.02504168483	0.00801458075	56	–0.00000298746	0.00002853094
17	–0.00284033019	0.01913753714	57	0.00002181698	0.00001937464
18	0.01510650786	0.01026212266	58	0.00002647330	–0.00001300352
19	0.01349973422	–0.01135150963	59	–0.00000154681	–0.00002424123
20	–0.00708149225	–0.01377375695	60	–0.00001711159	–0.00001068359
21	–0.01258681850	0.00273428578	61	–0.00001923481	0.00000926127
22	–0.00083306497	0.01074751371	62	0.00000183647	0.00002067358
23	0.00846956346	0.00317903201	63	0.00001561153	0.00000562727
24	0.00443198709	–0.00590660567	64	0.00001234917	–0.00000783954
25	–0.00341382328	–0.00492758043	65	–0.00000085941	–0.00001583695
26	–0.00487267426	0.00138760262	66	–0.00001422034	–0.00000432858
27	0.00000413447	0.00433140530	67	–0.00000832048	0.00000840164
28	0.00340441140	0.00082836877	68	0.00000148714	0.00001111048
29	0.00129907508	–0.00232344992	69	0.00001147832	0.00000376375
30	–0.00135910377	–0.00152023100	70	0.00000666721	–0.00000849793
31	–0.00150804628	0.00066706321	71	–0.00000309407	–0.00000799741
32	0.00023811970	0.00128073555	72	–0.00000824917	–0.00000231343
33	0.00093043377	0.00002045214	73	–0.00000568092	0.00000697571
34	0.00017693473	–0.00059168826	74	0.00000376653	0.00000666976
35	–0.00035637071	–0.00024290125	75	0.00000619450	0.00000059500
36	–0.00021753661	0.00023058037	76	0.00000432660	–0.00000502578
37	0.00016733218	0.00013422160	77	–0.00000314042	–0.00000601017
38	0.00005412478	–0.00012720729	78	–0.00000560068	0.00000036303
39	–0.00010428640	–0.00002215407	79	–0.00000274302	0.00000407729
			80	0.00000231091	0.00000494742

All the above scaling functions and wavelets are designed based on the Meyer's scaling function, so we have

$$||\phi_{j,k}|| = ||\psi^R_{j-(n-1)/N,k}|| = ||\psi^I_{j-(n-1)/N,k}|| = 1. \tag{50}$$

TABLE 4. The real part $\{b_k^{R_2}\}$ and imaginary part $\{b_k^{I_2}\}$ of decomposition sequences of VD-CDWT, $N = 4$, $b_k^{R_2} = b_{-k}^{R_2}$, $b_k^{I_2} = -b_{-k}^{I_2}$

k	$b_k^{R_2}$	$b_k^{I_2}$	k	$b_k^{R_2}$	$b_k^{I_2}$
0	0.24529053897	0.00000000000	40	0.00057426059	–0.00060409468
1	0.08351239163	–0.22925680432	41	–0.00025115092	–0.00060391153
2	–0.18385833595	–0.15447381076	42	–0.00050533855	0.00001667523
3	-0.20263459486	0.11666951087	43	–0.00009698032	0.00037280182
4	0.03865241538	0.22186685601	44	0.00026578891	0.00012870456
5	0.21113353378	0.03795256740	45	0.00012767209	–0.00020221817
6	0.10190698659	–0.17442888617	46	–0.00016906309	–0.00012941846
7	–0.11974260340	–0.14494719944	47	–0.00014548409	0.00014136854
8	–0.16305763209	0.05730988952	48	0.00010000511	0.00016659148
9	–0.00244808731	0.15682357809	49	0.00017432290	–0.00004145177
10	0.13085568347	0.05084566724	50	0.00002240471	–0.00015425528
11	0.08232914006	–0.09248063543	51	–0.00010456878	–0.00007207123
12	–0.05003726050	–0.09510573982	52	–0.00009091868	0.00003724989
13	–0.09095156734	0.01117390190	53	–0.00002739959	0.00007325927
14	–0.01849537918	0.07433005979	54	0.00002684512	0.00006960394
15	0.05109077724	0.03613483530	55	0.00007811185	0.00003059601
16	0.04175068110	–0.02708068637	56	0.00007867805	–0.00005410798
17	–0.00697913759	–0.03765544184	57	–0.00000988079	–0.00010214113
18	–0.02753834109	–0.00642308981	58	–0.00009596021	–0.00003667028
19	–0.01241148720	0.01540810146	59	–0.00006909913	0.00006607579
20	0.00466494106	0.01204822390	60	0.00002595487	0.00007816302
21	0.00753697751	0.00251065819	61	0.00006422981	0.00000922876
22	0.00536899683	–0.00145816035	62	0.00002831733	–0.00003584006
23	0.00392312400	–0.00442728420	63	–0.00000544359	–0.00002766291
24	–0.00104164016	–0.00712648846	64	–0.00001137359	–0.00001573121
25	–0.00763400077	–0.00312988614	65	–0.00002164280	–0.00001135094
26	–0.00661989050	0.00579171608	66	–0.00003040460	0.00001290863
27	0.00251643436	0.00848648438	67	–0.00000442806	0.00003870195
28	0.00843898609	0.00107742794	68	0.00003440258	0.00002199916
29	0.00401824674	–0.00676673659	69	0.00003264827	–0.00002082731
30	–0.00413195768	–0.00570350499	70	–0.00000441802	–0.00003299826
31	–0.00597736130	0.00131078493	71	–0.00002420364	–0.00000820905
32	–0.00103719447	0.00507165245	72	–0.00001286132	0.00001084909
33	0.00345669989	0.00251532162	73	–0.00000132798	0.00000906592
34	0.00303150280	–0.00166284397	74	–0.00000026157	0.00000785015
35	–0.00012831049	–0.00274784309	75	0.00000709233	0.00001042879
36	–0.00197380754	–0.00089075517	76	0.00001706066	–0.00000054119
37	–0.00133793595	0.00104627021	77	0.00000820590	–0.00001779299
38	0.00023348673	0.00131471380	78	–0.00001295198	–0.00001514688
39	0.00100615461	0.00031566387	79	–0.00001752454	0.00000508058
			80	–0.00000235715	0.00001477488

Figure 12 shows an example of the designed VD-CDWT when $N = 4$. Finally, we list the decomposition sequences $\{b_k^{R_n}\}$, $\{b_k^{I_n}\}$, $\{a_k\}$ of VD-CDWT when $N = 4$, and the decomposition sequences $\{b_k^R\}$, $\{b_k^I\}$ of CDWT in Tables 3, 4, $\cdots$, 8.

TABLE 5. The real part $\{b_k^{R_3}\}$ and imaginary part $\{b_k^{I_3}\}$ of decomposition sequences of VD-CDWT, $N = 4$, $b_k^{R_3} = b_{-k}^{R_3}$, $b_k^{I_3} = -b_{-k}^{I_3}$

k	$b_k^{R_3}$	$b_k^{I_3}$	k	$b_k^{R_3}$	$b_k^{I_3}$
0	0.22477924019	0.00000000000	40	0.00338650392	−0.00037396455
1	0.11586627181	−0.19163663534	41	0.00138818494	−0.00243993002
2	−0.10287177640	−0.19609662005	42	−0.00091770496	−0.00209592642
3	−0.21692523925	−0.01321343107	43	−0.00179936795	−0.00044475517
4	−0.12038640202	0.17410416197	44	−0.00118487976	0.00092142933
5	0.08369158335	0.18667666469	45	−0.00001687905	0.00122182428
6	0.19468632709	0.02410794649	46	0.00077590795	0.00063633042
7	0.11562035810	−0.14646811889	47	0.00081000470	−0.00018447315
8	−0.06152881154	−0.16493263171	48	0.00027602134	−0.00063243328
9	−0.16171612777	−0.03093245816	49	−0.00030266338	−0.00048439581
10	−0.10243353205	0.11314926994	50	−0.00046851655	0.00000069026
11	0.03974512245	0.13441640991	51	−0.00018324207	0.00033361796
12	0.12314409478	0.03290181833	52	0.00018139816	0.00025029961
13	0.08314648468	−0.07901062258	53	0.00024883522	−0.00006341361
14	−0.02113033795	−0.09984588048	54	0.00001904878	−0.00022390673
15	−0.08442757973	−0.03031356694	55	−0.00019109112	−0.00008539725
16	−0.06095929662	0.04833230241	56	−0.00014497085	0.00014187989
17	0.00739414982	0.06605094910	57	0.00006676823	0.00018595965
18	0.05022884002	0.02436587997	58	0.00018681425	0.00002546298
19	0.03919603050	−0.02403963826	59	0.00010644673	−0.00013729115
20	0.00100717624	−0.03698788183	60	−0.00005186462	−0.00014426959
21	−0.02360558678	−0.01675474031	61	−0.00012428793	−0.00003434328
22	−0.02058823713	0.00735577228	62	−0.00008420129	0.00005998112
23	−0.00470813774	0.01507384643	63	−0.00001337810	0.00007939848
24	0.00568781867	0.00919110225	64	0.00003033317	0.00005808198
25	0.00678881398	0.00209065289	65	0.00005437070	0.00003062367
26	0.00503413239	−0.00097815741	66	0.00006786381	−0.00000989021
27	0.00414401570	−0.00298899128	67	0.00004655809	−0.00006189829
28	0.00177920116	−0.00580601758	68	−0.00001841047	−0.00008236916
29	−0.00351829129	−0.00614000599	69	−0.00007904084	−0.00003719928
30	−0.00770868571	−0.00116589369	70	−0.00007545990	0.00004069378
31	−0.00574609529	0.00580768646	71	−0.00001056479	0.00007906606
32	0.00148651217	0.00808443655	72	0.00005073528	0.00004851828
33	0.00731611424	0.00323043157	73	0.00005733730	−0.00000929960
34	0.00639490237	−0.00403582857	74	0.00002247466	−0.00003867640
35	0.00018851828	−0.00696965837	75	−0.00000835508	−0.00003066746
36	−0.00512230349	−0.00363941182	76	−0.00001647731	−0.00001427474
37	−0.00519068794	0.00195030395	77	−0.00001748607	−0.00000666508
38	−0.00113734503	0.00466137362	78	−0.00002237799	0.00000260221
39	0.00270498176	0.00305017779	79	−0.00001860219	0.00002101235
			80	0.00000425983	0.00003204182

TABLE 6. The real part $\{b_k^{R_4}\}$ and imaginary part $\{b_k^{I_4}\}$ of decomposition sequences of VD-CDWT, $N = 4$, $b_k^{R_4} = b_{-k}^{R_4}$, $b_k^{I_4} = -b_{-k}^{I_4}$

k	$b_k^{R_4}$	$b_k^{I_4}$	k	$b_k^{R_4}$	$b_k^{I_4}$
0	0.20612347208	0.00000000000	40	0.00694861761	–0.00145319843
1	0.13356695003	–0.15627825929	41	0.00344176825	–0.00573780348
2	–0.03176106985	–0.20146736782	42	–0.00174018293	–0.00595132070
3	–0.17149074668	–0.10536284014	43	–0.00506221673	–0.00252426047
4	–0.18799734880	0.06073197062	44	–0.00476121230	0.00178598262
5	–0.07423262083	0.17802091560	45	–0.00173107613	0.00416280663
6	0.08448520736	0.16713504108	46	0.00160460168	0.00360410752
7	0.17564700495	0.04296187911	47	0.00320948734	0.00115408346
8	0.14101857713	–0.10125965918	48	0.00262111451	–0.00127904315
9	0.01418992519	–0.16510470616	49	0.00079244822	–0.00234038330
10	–0.11015745030	–0.11219842996	50	–0.00091170282	–0.00186799944
11	–0.14794408095	0.00991924945	51	–0.00163795618	–0.00059100788
12	–0.08328822932	0.11120604519	52	–0.00133392814	0.00058777115
13	0.02791856303	0.12627645425	53	–0.00047755337	0.00112657590
14	0.10528218410	0.05662873296	54	0.00035607036	0.00096943131
15	0.10245794253	–0.03919817011	55	0.00078552179	0.00039065742
16	0.03401789871	–0.09391849048	56	0.00071406798	–0.00022647254
17	–0.04397371905	–0.07876465921	57	0.00029425098	–0.00056939931
18	–0.07903268343	–0.01654518874	58	–0.00017927093	–0.00051699922
19	–0.05711230554	0.04315545120	59	–0.00042839014	–0.00017948648
20	–0.00454724967	0.06262981254	60	–0.00034739021	0.00018018970
21	0.03813076039	0.03886087111	61	–0.00005752199	0.00032295851
22	0.04652826218	–0.00232053051	62	0.00019505842	0.00019508784
23	0.02472653864	–0.03050490902	63	0.00023064836	–0.00005177736
24	–0.00492179977	–0.03215110714	64	0.00006451384	–0.00019998365
25	–0.02184724414	–0.01480179786	65	–0.00013053966	–0.00014545492
26	–0.02040832380	0.00442468754	66	–0.00018530277	0.00003423530
27	–0.00866561030	0.01348375952	67	–0.00007222646	0.00016981564
28	0.00207778426	0.01167618802	68	0.00009341817	0.00015389180
29	0.00636337624	0.00555185012	69	0.00017157100	0.00001932121
30	0.00586204857	0.00098457439	70	0.00011593007	–0.00011242991
31	0.00453830729	–0.00100828809	71	–0.00000761419	–0.00014640136
32	0.00387989977	–0.00252906781	72	–0.00009860833	–0.00008275064
33	0.00245803761	–0.00472058481	73	–0.00010879765	0.00000899317
34	–0.00104866139	–0.00604498276	74	–0.00006201278	0.00006479112
35	–0.00534384456	–0.00422474825	75	–0.00000861944	0.00007211385
36	–0.00723525338	0.00074876334	76	0.00002522428	0.00005545306
37	–0.00467625530	0.00587792085	77	0.00004496969	0.00003490422
38	0.00103260902	0.00747047981	78	0.00005933475	0.00000841329
39	0.00603468374	0.00427136274	79	0.00005926405	–0.00002995519
			80	0.00002953860	–0.00006680494

TABLE 7. The decomposition sequence $\{a_k\}$ of VD-CDWT, $N = 4$, $a_k = a_{-k}$

k	a_k	k	a_k
0	0.36395324071	40	0.00212484092
1	0.32530988864	41	0.00230721863
2	0.22409835987	42	0.00114211272
3	0.09777070269	43	−0.00043590894
4	−0.01025203284	44	−0.00144092789
5	−0.06913654409	45	−0.00141902363
6	−0.07212851385	46	−0.00060260256
7	−0.03605909637	47	0.00036938786
8	0.00981957765	48	0.00091047413
9	0.03919677631	49	0.00081520853
10	0.04032993036	50	0.00029075529
11	0.01873327145	51	−0.00026100937
12	−0.00913427111	52	−0.00052402368
13	−0.02662904401	53	−0.00042924126
14	−0.02588300812	54	−0.00012943290
15	−0.01053479444	55	0.00015014362
16	0.00824524440	56	0.00026286507
17	0.01922333054	57	0.00020188384
18	0.01741184611	58	0.00005967678
19	0.00588879774	59	−0.00005908046
20	−0.00721375342	60	−0.00010282413
21	−0.01412720559	61	−0.00008261559
22	−0.01183420485	62	−0.00003889744
23	−0.00307458385	63	−0.00000299390
24	0.00610688122	64	0.00001796281
25	0.01034824822	65	0.00003074944
26	0.00796407094	66	0.00003874931
27	0.00135943901	67	0.00003606897
28	−0.00499098059	68	0.00001627981
29	−0.00745799547	69	−0.00001577250
30	−0.00523332205	70	−0.00004256187
31	−0.00035592592	71	−0.00004565288
32	0.00392567162	72	−0.00002081403
33	0.00523737564	73	0.00001679597
34	0.00331848925	74	0.00004248498
35	−0.00017574116	75	0.00003965297
36	−0.00295908694	76	0.00001162462
37	−0.00355287607	77	−0.00002123882
38	−0.00200676045	78	−0.00003662895
39	0.00040009385	79	−0.00002600138
		80	0.00000054528

TABLE 8. The real part $\{b_k^R\}$ and imaginary part $\{b_k^I\}$ of decomposition sequences of CDWT, $N = 4$, $b_k^R = b_{-2-k}^R$, $b_k^I = -b_{-1-k}^I$

k	b_k^R	b_k^I
–1	0.72790648142	
0	–0.44819671975	–0.65061977728
1	–0.02050406569	0.19554140538
2	0.14425702769	0.13827308818
3	0.01963915530	–0.07211819273
4	–0.08065986072	–0.07839355262
5	–0.01826854222	0.03746654289
6	0.05176601624	0.05325808802
7	0.01649048879	–0.02106958887
8	–0.03482369223	–0.03844666107
9	–0.01442750684	0.01177759549
10	0.02366840970	0.02825441117
11	0.01221376244	–0.00614916770
12	–0.01592814188	–0.02069649644
13	–0.00998196117	0.00271887801
14	0.01046664410	0.01491599095
15	0.00785134323	–0.00071185184
16	–0.00663697849	–0.01047475129
17	–0.00591817387	–0.00035148233
18	0.00401352089	0.00710575215
19	0.00424968184	0.00080018771
20	–0.00228422545	–0.00461443726
21	–0.00288185579	–0.00087181789
22	0.00120520511	0.00283804726
23	0.00182094825	0.00073877572
24	–0.00058151058	–0.00163041707
25	–0.00104804736	–0.00052201874
26	0.00025886580	0.00085848251
27	0.00052573013	0.00030028725
28	–0.00011935356	–0.00040376769
29	–0.00020564825	–0.00011816091
30	0.00007779488	0.00016523118
31	0.00003592562	–0.00000598780
32	–0.00007749861	–0.00006149888
33	0.00003255962	0.00007213794
34	0.00008512375	0.00003154500
35	–0.00004162807	–0.00009130575
36	–0.00008496996	–0.00003359194
37	0.00002324923	0.00007930595
38	0.00007325789	0.00004247763
39	0.00000109056	–0.00005200275
40		–0.00004601933