

Wavelet Power Spectrum Estimation for High-resolution Terahertz Time-domain Spectroscopy

Youngchan Kim^{1,3}, Kyung Hwan Jin², Jong Chul Ye², Jaewook Ahn³, and Dae-Su Yee^{1*}

¹*Center for Safety Measurements, Korea Research Institute of Standards and Science, Daejeon 305-340, Korea*

²*Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea*

³*Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea*

(Received January 12, 2011 : revised February 7, 2011 : accepted February 9, 2011)

Recently reported asynchronous-optical-sampling terahertz (THz) time-domain spectroscopy enables high-resolution spectroscopy due to a long time-delay window. However, a long-lasting tail signal following the main pulse is often measured in a time-domain waveform, resulting in spectral fluctuation above a background noise level on a high-resolution THz amplitude spectrum. Here, we adopt the wavelet power spectrum estimation technique (WPSET) to effectively remove the spectral fluctuation without sacrificing spectral features. Effectiveness of the WPSET is verified by investigating a transmission spectrum of water vapor.

Keywords : Terahertz time-domain spectroscopy, Wavelet, Asynchronous optical sampling, High spectral resolution

OCIS codes : (300.6495) Spectroscopy, terahertz; (300.6320) Spectroscopy, high-resolution; (100.7410) Wavelets

I. INTRODUCTION

Over the past two decades, the field of terahertz (THz) waves has drawn much attention due to the development of THz science and technologies [1]. It is noteworthy that THz time-domain spectroscopy (THz-TDS) has been pioneered as a basic measurement tool for characterizing optical and electrical properties of materials [2, 3]. THz-TDS also provides a potentially attractive method for probing fundamental physical interactions as well as practical applications [4, 5]. However, conventional THz-TDS using a single mode-locked femtosecond laser and a mechanical delay line has an inherent tradeoff between frequency resolution and measurement time since the frequency resolution is inversely proportional to an overall time delay window. Translation stages can commonly offer a time delay window up to hundreds of picoseconds while they need a measurement time of several minutes. On the contrary, vibrating mirrors can give a scan rate of up to 100 Hz but the measured time delay window is

limited to less than 100 ps. Different types of optical delay line have also been devised for rapid data acquisition or high spectral resolution [5, 6].

Asynchronous optical sampling (ASOPS) THz-TDS has recently been demonstrated to overcome the tradeoff [7-10]. The ASOPS THz-TDS can serve for high-resolution spectroscopy as well as high-speed monitoring of dynamic processes [11]. It should be noted, however, that spectral fluctuation can significantly appear on high-resolution THz spectra measured in ASOPS THz-TDS. It is caused by a long-lasting tail signal following a main THz pulse in time-domain waveforms. The tail signal is considered to be due to multiple reflections within and between optical components. Since the tail signal causes spectral fluctuation on the spectrum above the background noise level, signal processing techniques are needed to remove the fluctuations.

Signal processing based on wavelets is expected to enhance signal quality without losing information as compared to filtering or smoothing. Actually, there have been a few

*Corresponding author: dsyee@kriss.re.kr

Color versions of one or more of the figures in this paper are available online.

reports in which the wavelet denoising technique is applied to time-domain waveforms from THz-TDS for gas sensing [12], imaging [13, 14], and local tomography [15]. In this paper, we intend to directly apply the wavelet denoising technique to high-resolution THz amplitude spectra to enhance the spectrum quality. Our ASOPS THz-TDS setup with a frequency resolution of 100 MHz is described along with typical experimental results. Then, the wavelet power spectrum estimation technique (WPSET) is described for fluctuation reduction in high-resolution THz amplitude spectra measured from ASOPS THz-TDS. A transmission spectrum of water vapor is investigated to verify the performance of the WPSET.

II. EXPERIMENT

Our experimental setup for ASOPS THz-TDS is illustrated in Fig. 1. We employ a laser system where two Ti:Sapphire femtosecond lasers are pumped by a diode-pumped solid state laser at 532 nm wavelength. The whole laser system is placed on a temperature-controlled base-plate to reduce thermal fluctuation. Both the femtosecond lasers 1 and 2 have a center wavelength of 800 nm and their pulse durations are 10 and 20 fs, respectively. Two phase-locked loops are used to stabilize the repetition frequencies of the femtosecond lasers. The repetition frequencies of femtosecond lasers 1 (f_{r1}) and 2 (f_{r2}) are stabilized at 100 MHz and at a variable frequency, respectively, so that the difference frequency ($\Delta f_r = f_{r1} - f_{r2}$) can be adjusted. By using double-balanced mixers (DBM), the tenth harmonics of f_{r1} and f_{r2} are compared with the outputs of a 1 GHz dielectric resonator oscillator (DRO) and a signal generator, respectively. The DRO and signal generator share a 10 MHz reference oscillator to reduce the relative timing jitter between the optical pulses from the two femtosecond lasers. The phase error signal output from the DBM is amplified by a proportional-integral amplifier and a high-voltage amplifier and is supplied to a piezoelectric transducer (PZT) to which a

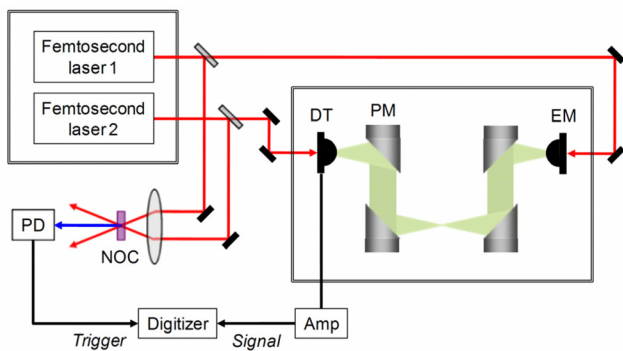


FIG. 1. Schematic diagram of the experimental setup for ASOPS THz-TDS. PD: photodetector, EM: THz emitter, DT: THz detector, PM: off-axis parabolic mirror, Amp: variable-gain current amplifier, NOC: nonlinear optical crystal.

cavity mirror is attached. The repetition frequencies are stabilized by controlling the cavity lengths via the PZTs.

Two low-temperature-grown GaAs photoconductive antennas are used for THz emission and detection. The optical pulses from the femtosecond laser 1 are incident on the THz emitter to generate THz pulses, which are guided into the THz detector by using four off-axis parabolic mirrors. The optical pulses from the femtosecond laser 2 are used to optically sample the THz pulses incident upon the THz detector. The photocurrent output from the DT represents the magnitude of the electric field of the THz pulses temporally sampled. It is amplified by a variable-gain current amplifier and then is input to a 24-bit flexible resolution digitizer. The digitizer is triggered by a cross-correlation signal generated from a nonlinear optical crystal (2-mm-thick BBO crystal) on which the optical pulses from the two femtosecond lasers are non-collinearly focused. The trigger signal has a repetition frequency equal to the difference frequency. The digitizer acquires time-domain data when triggered by the cross-correlation signal, and we can average over consecutive traces to reduce the noise. The time axis is converted from a real time to a time delay by

$$\tau(\text{time delay}) = t(\text{real time}) \times \frac{\Delta f_r}{f_{r1}}. \quad (1)$$

A typical time-domain signal with a 10 ns time delay window measured from the ASOPS THz-TDS is displayed in Fig. 2(a). It is acquired by averaging over 1,000 traces during 100 s at a difference frequency of 20 Hz. It can be seen that the tail signal following the main pulse lasts for several nanoseconds. The tail signal is considered to arise from multiple reflections within and between optical components. Fig. 2(b) shows a typical THz amplitude spectrum with a frequency resolution of 100 MHz given by the fast Fourier transform (FFT) of the time-domain waveform in Fig. 2(a). The long-lasting tail signal leads to the spectral fluctuation on the spectrum above the background noise level. The fluctuation is not observed in the conventional THz-TDS using an optical delay stage since it has a low spectral resolution due to a much shorter time delay window. The time-domain waveform around the main pulse is displayed on a time delay window of 100 ps in the inset of Fig. 2(a). Its spectrum with a frequency resolution of 10 GHz has almost no fluctuation, as shown in the inset of Fig. 2(b). Hence, a signal processing technique is needed to remove the fluctuation on high-resolution THz amplitude spectra measured from the ASOPS THz-TDS.

III. THEORETICAL DESCRIPTION OF WPSET

Wavelet based signal processing appears a natural choice for denoising THz pulses since wavelet analysis is good for capturing singularities both in the time and frequency

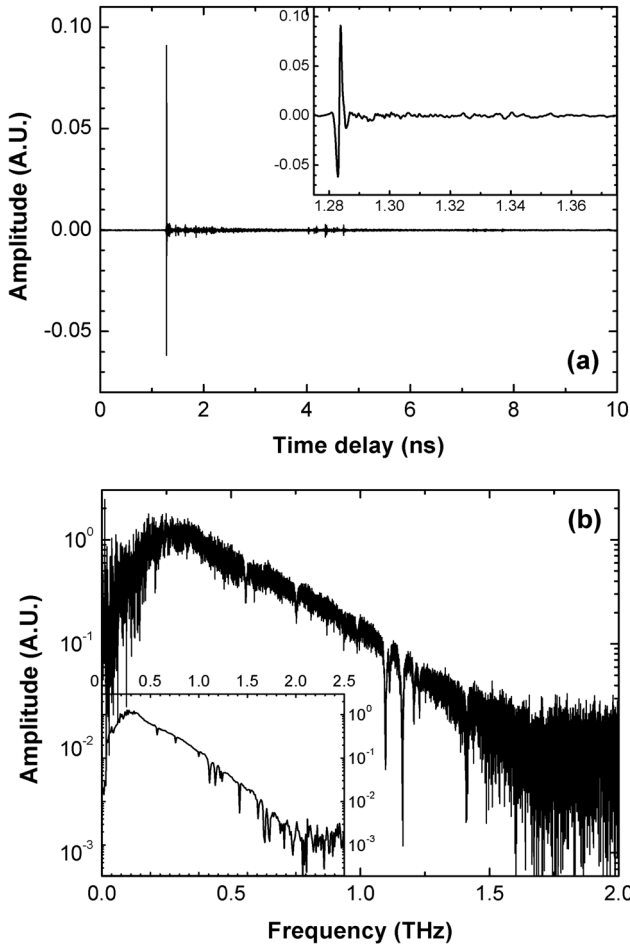


FIG. 2. (a) Typical time-domain waveform on a 10 ns time delay window measured from ASOPS THz-TDS at relative humidity of 28%. (b) THz amplitude spectrum obtained by FFT of the time-domain waveform in (a).

domains [12]. Signals can be analyzed by wavelet transform using a set of basis shifted and dilated from a mother wavelet $\psi(t)$, defined as

$$\left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2}} \psi\left(\frac{t-2^j n}{2^j}\right) \right\}_{(j,n) \in \mathbb{Z}^2}, \quad (2)$$

where j and n mean the scale and location indices, respectively, and \mathbb{Z} is the set of integers. A wavelet basis is an orthogonal basis of $L^2(\mathbb{R})$, where \mathbb{R} is the set of real numbers and $L^2(\mathbb{R})$ corresponds to the Hilbert space for functions $f(t)$ with finite energy such that $\int |f(t)|^2 dt < +\infty$. It is known that any signal $f(t) \in L^2(\mathbb{R})$ can be represented as

$$f = \sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}, \quad (3)$$

where $\langle f, g \rangle = \int f(t)g(t)dt$ is the inner product. By introducing a scaling function ϕ , we can also represent Eq. (3) as

$$f = \sum_{j=-\infty}^J \sum_{n=-\infty}^{\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} + \sum_{n=-\infty}^{\infty} \langle f, \phi_{J,n} \rangle \phi_{J,n}, \quad (4)$$

where the scaling function ϕ can be readily derived from a mother wavelet ψ and J is the maximum scale index for dilation [16]. Then, the approximation coefficients $\{a_j[n]\}_{j,n}$ and detail coefficients $\{d_j[n]\}_{j,n}$ are defined by

$$a_j[n] = \langle f, \phi_{j,n} \rangle, \quad d_j[n] = \langle f, \psi_{j,n} \rangle. \quad (5)$$

The approximation and detail coefficients of the discrete signal are processed using a shrinkage operator $\rho_r(\cdot)$:

$$\tilde{f} = \sum_{j=0}^J \sum_{n=0}^{N-1} \rho_r(d_j[n]) \psi_{j,n} + \sum_{n=0}^{N-1} \rho_r(a_j[n]) \phi_{j,n}. \quad (6)$$

Many different types of shrinkage operator $\rho_r(\cdot)$ have been proposed depending on applications. For example, the hard-thresholding defined as

$$\rho_r(x) = \begin{cases} x, & |x| \geq r; \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

is one of popular methods for denoising [16]. Note that an appropriate choice of wavelet is especially important in wavelet analysis. A mother wavelet is usually designed to have appropriate vanishing moments, support size, symmetry, regularity, etc. It is known that the Daubechies wavelet has the maximum vanishing moment for a given support size [16].

A time-domain waveform from conventional THz-TDS can be directly processed using a wavelet denoising algorithm. For example, Mittleman *et al.* demonstrated that the noises in THz time-domain signals can be effectively suppressed by appropriate choice of wavelets [12]. Ferguson *et al.* further demonstrated that T-ray images are also greatly improved by wavelet denoising followed by Wiener deconvolution [13]. However, the time-domain wavelet denoising cannot effectively remove the fluctuation on a spectrum while retaining spectral information since the fluctuation does not come from noise, but rather from the long-lasting signal of the time-domain waveform. Fig. 3 shows results of the time-domain wavelet denoising where the Daubechies wavelet with 10 vanishing moments was used at the scale of 5. For a low threshold value, the spectral features remain unchanged but the fluctuation is hardly reduced, as shown in Fig. 3(a). On the contrary, a high threshold value results in distortion at the singularities as well as

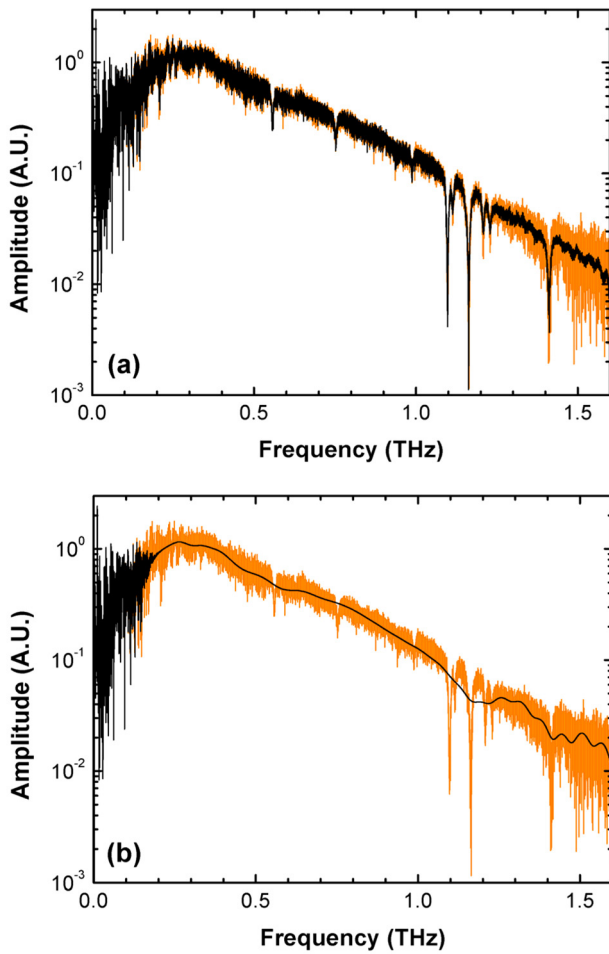


FIG. 3. THz amplitude spectra of wavelet-denoised time-domain waveforms. The orange lines show the raw spectrum in Fig. 2(b) and the black lines the spectra of the time-domain waveforms that are denoised for a low (a) and high (b) threshold value, respectively.

reduction of the fluctuation, as shown in Fig. 3(b).

Therefore, we employ a direct power spectrum estimation technique to denoise the spectrum [17]. Here, the estimation problem of the power spectrum of THz-TDS data is regarded as a nonparametric statistical estimation problem of a stationary random process. More specifically, time-domain data $x[n]$ is assumed to be a real wide-stationary Gaussian random process with zero mean and covariance $R_x[n] = E[x[m]x[m+n]]$ where the $E[\cdot]$ is the expectation of an inside parameter. The periodogram of $x[n]$ is then described by

$$I_N(\nu) = \left| \sum_{n=0}^{2N-1} x[n] e^{-i2\pi\nu n} \right|^2, \quad -\frac{1}{2} < \nu \leq \frac{1}{2}. \quad (8)$$

The periodogram is often used as an estimate of the power spectrum $S(\nu)$ which is defined as

$$S(\nu) = \sum_{n=-\infty}^{\infty} R_X[n] e^{-i2\pi\nu n}, \quad -\frac{1}{2} < \nu \leq \frac{1}{2}. \quad (9)$$

If $S(\cdot)$ is sufficiently smooth and N is large enough, the periodogram and power spectrum satisfy the following asymptotic relation [17]:

$$\ln I_N\left(\frac{k}{2N}\right) - \gamma = \ln S\left(\frac{k}{2N}\right) + \varepsilon(k), \quad k = 0, 1, \dots, N, \quad (10)$$

where $\gamma \approx 0.5772$ denotes the Euler-Mascheroni constant, and $\{\varepsilon(k); k=0, \dots, N\}$ are independent and identically distributed (i.i.d.) with zero mean and variance of $\pi^2/6$. It is known that the periodogram is asymptotically unbiased [17]. Hence, the periodogram is quite often used as an estimate of the power spectrum. In wavelet power spectrum estimation, the wavelet transform is directly applied to the log periodogram in Eq. (10). We can define the detailed coefficients for $\ln S(\cdot)$ and $\varepsilon(\cdot)$:

$$\begin{aligned} d_j^S[n] &= \langle \ln S, \psi_{j,n} \rangle \\ d_j^\varepsilon[n] &= \langle \varepsilon, \psi_{j,n} \rangle \end{aligned} \quad (11)$$

Hence, Eq. (10) can be equivalently transformed into

$$d_j[n] = d_j^S[n] + d_j^\varepsilon[n]. \quad (12)$$

If $\max_n |\psi(\nu)| = M < \infty$, then the probability density function of $d_j^\varepsilon[\cdot]$ converges to the Gaussian probability density function as N grows. Since N is sufficiently large in high-resolution THz-TDS, the model Eq. (12) is accurate and the noise has an i.i.d. Gaussian distribution with zero mean [17]. Under this condition, the result of wavelet denoising is a nearly minimax optimal estimate of the underlying true log power spectrum [18], and the estimate can be represented by

$$\ln S = \sum_{j=0}^J \sum_{n=0}^{N-1} \rho_T(\langle \ln I_N, \psi_{j,n} \rangle) \psi_{j,n} + \sum_{n=0}^{N-1} \rho_T(\langle \ln I_N - \gamma, \phi_{j,n} \rangle) \phi_{j,n}, \quad (13)$$

where the shrinkage operator ρ_T is given by Eq. (7).

IV. WAVELET POWER SPECTRUM ESTIMATION

We denoised the THz amplitude spectrum in Fig. 2(b) using the WPSET as described above. It is presented in Fig. 4 that the WPSET can effectively remove the fluctuations above the background noise level on the THz amplitude spectrum without altering the water vapor absorption

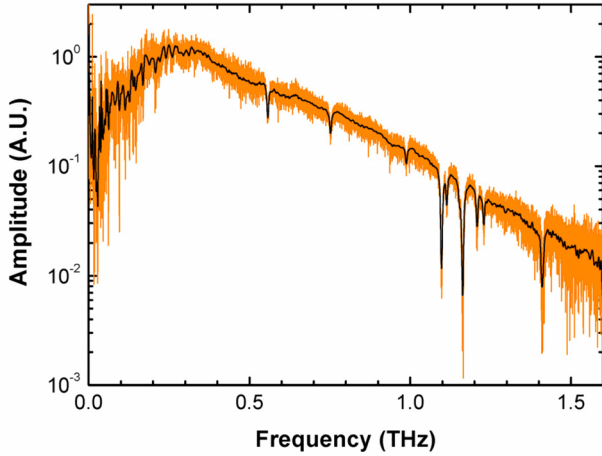


FIG. 4. THz amplitude spectrum denoised by using the WPSET. The orange and black lines display the raw and denoised spectra, respectively.

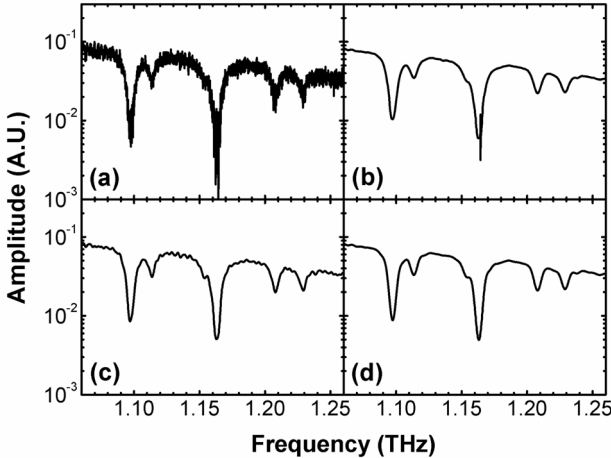


FIG. 5. Magnified views at around 1.16 THz of the raw spectrum (a) and the spectra denoised using the Haar wavelet at the scale of 5 (b), the Daubechies wavelet with 10 vanishing moments at the scale of 4 (c), and the Daubechies wavelet with 10 vanishing moments at the scale of 5 (d), respectively.

lines. The Daubechies wavelet with 10 vanishing moments was used at the scale of 5.

To find optimal parameters for wavelet power spectrum estimation, we compared the spectra denoised with different choices of the wavelet and scale. Fig. 5 shows magnified views around 1.16 THz of the raw spectrum and denoised spectra. The Haar wavelet overestimated the noise at 1.1645 GHz as a real singularity, as shown in Fig. 5(b). The Daubechies wavelet offers the minimum support for a given number of vanishing moments [13], which means that it is an optimal wavelet for denoising and detecting singularities simultaneously. Figs. 5(c) and (d) show the spectra that were denoised using the Daubechies wavelet with 10 vanishing moments at the scale of 4 and 5, respectively. It is apparently shown that the spectrum is more effectively denoised at

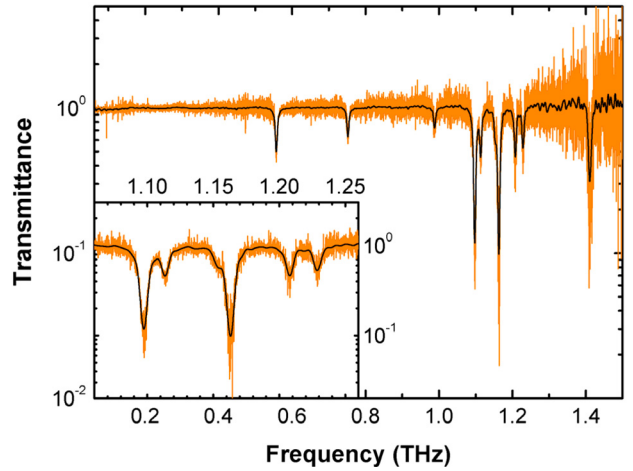


FIG. 6. THz transmission spectra of water vapor resulting from the raw (orange line) and denoised (black line) amplitude spectra. The inset is an enlarged view around 1.16 THz.

the scale of 5, compared to the scale of 4. We also observed that the scales over 6 caused distortion at singularities (not shown here). Therefore, we chose the Daubechies wavelet with 10 vanishing moments and the scale of 5 for the following studies.

To verify the performance of the WPSET, we investigated a THz transmission spectrum of water vapor. The THz transmission spectrum of water vapor was obtained from THz amplitude spectra measured at relative humidity of 28% and 2% and at 21.5°C. Fig. 6 shows the transmission spectra of water vapor before and after applying the WPSET with the optimized parameters, which are indicated by the orange and black lines, respectively. It is clearly demonstrated that the WPSET can effectively remove the fluctuations without distortion in the water vapor absorption lines. We extracted the center frequencies, linewidths, and transmittances of the 10 strongest absorption lines from the raw and denoised transmission spectra over the frequency range up to 1.5 THz by fitting them to Lorentzian line shapes. Table 1 lists these values along with the error values of the fitting results from the raw spectrum. The fitting results from the raw spectrum are in good agreement with those from the denoised spectrum within the error values.

V. CONCLUSION

We clearly demonstrated that the wavelet power spectrum estimation technique (WPSET) was successfully applied to high-resolution THz-TDS. Through the wavelet power spectrum estimation, the fluctuations on high-resolution THz amplitude spectra were removed while maintaining original spectral information. By investigating a THz transmission spectrum of water vapor, the parameters of the spectral singularities were shown to remain unchanged after the processing.

TABLE 1. Parameters of the absorption lines extracted from the raw and denoised transmission spectra of water vapor.

	Center frequency (GHz)			Linewidth (GHz)			Transmittance		
	Raw spectrum	Error	Denoised spectrum	Raw spectrum	Error	Denoised spectrum	Raw spectrum	Error	Denoised spectrum
1	557.09	0.09	557.08	7.99	0.28	8.06	0.4879	0.0353	0.4905
2	752.17	0.11	752.16	7.11	0.36	7.22	0.6183	0.0382	0.6220
3	988.00	0.26	987.96	7.87	0.85	7.74	0.7221	0.0644	0.7251
4	1097.30	0.12	1097.30	10.64	0.39	10.52	0.1241	0.0676	0.1178
5	1113.65	0.18	1113.64	7.85	0.60	7.75	0.5495	0.0746	0.5492
6	1152.78	0.44	1152.79	7.00	1.51	6.73	0.7971	0.1129	0.8078
7	1162.98	0.16	1162.98	11.08	0.50	10.85	0.1190	0.0891	0.0988
8	1207.71	0.16	1207.72	8.28	0.51	8.19	0.4884	0.0661	0.4806
9	1228.74	0.17	1228.76	7.39	0.53	7.47	0.5390	0.0697	0.5367
10	1410.98	0.49	1410.99	11.13	1.61	10.98	0.2761	0.2529	0.2617

This study shows that the WPSET is ideally suitable for high-resolution THz-TDS. The WPSET will make high-resolution THz-TDS more practical by enhancing the spectrum quality. Even if this work was performed in the THz frequency regime, the WPSET can also be applied to high-resolution spectroscopy in other frequency regions.

ACKNOWLEDGMENT

This work was supported in part by the Industrial Strategic Technology Program of the Ministry of Knowledge Economy (KI001889) and in part by the Ministry of Education, Science, and Technology through project KRISS-10011131.

REFERENCES

1. Y.-S. Lee, *Principles of Terahertz Science and Technology* (Springer, New York, USA, 2009), Chapter 3.
2. M. van Exter, Ch. Fattinger, and D. Grischkowsky, "Terahertz time-domain spectroscopy of water vapor," *Opt. Lett.* **14**, 1128-1130 (1989).
3. B. E. Cole, J. B. Williams, B. T. King, M. S. Sherwin, and C. R. Stanley, "Coherent manipulation of semiconductor quantum bits with terahertz radiation," *Nature* **410**, 60-63 (2001).
4. E. Pickwell and V. P. Wallace, "Biomedical applications of terahertz technology," *J. Phys. D: Appl. Phys.* **39**, R301-R310 (2006).
5. H. Hoshina, T. Seta, T. Iwamoto, I. Hosako, C. Otani, and Y. Kasai, "Precise measurement of pressure broadening parameters for water vapor with a terahertz time-domain spectrometer," *J. Quant. Spectrosc. Radiat. Transfer.* **109**, 2303-2314 (2008).
6. G. J. Kim, S. G. Jeon, J. I. Kim, and Y. S. Jin, "High speed scanning of terahertz pulse by a rotary optical delay line," *Rev. Sci. Instrum.* **79**, 106102-1~106102-3 (2008).
7. C. Janke, M. Först, M. Nagel, and H. Kurz, "Asynchronous optical sampling for high-speed characterization of integrated resonant terahertz sensors," *Opt. Lett.* **30**, 1405-1407 (2005).
8. T. Yasui, E. Saneyoshi, and T. Araki, "Asynchronous optical sampling terahertz time-domain spectroscopy for ultrahigh spectral resolution and rapid data acquisition," *Appl. Phys. Lett.* **87**, 061101-1~061101-3 (2005).
9. A. Bartels, A. Thoma, C. Janke, T. Dekorsy, A. Dreyhaupt, S. Winnerl, and M. Helm, "High-resolution THz spectrometer with kHz scan rates," *Opt. Express* **14**, 430-437 (2006).
10. A. Bartels, R. Cerna, C. Kistner, A. Thoma, F. Hudert, C. Janke, and T. Dekorsy, "Ultrafast time-domain spectroscopy based on high-speed asynchronous optical sampling," *Rev. Sci. Instrum.* **78**, 035107-1~035107-8 (2007).
11. C. Kistner, A. Andre, T. Fischer, A. Thoma, C. Janke, A. Bartels, T. Gisler, G. Maret, and T. Dekorsy, "Hydration dynamics of oriented DNA films investigated by time-domain terahertz spectroscopy," *Appl. Phys. Lett.* **90**, 233908-1~233908-3 (2007).
12. D. M. Mittleman, R. H. Jacobsen, R. Neelamani, R. G. Baraniuk, and M. C. Nuss, "Gas sensing using terahertz time-domain spectroscopy," *Appl. Phys. B* **67**, 379-390 (1998).
13. B. Ferguson and D. Abbott, "De-noising techniques for terahertz responses of biological samples," *Microelectronics Journal* **32**, 943-953 (2001).
14. B. Ferguson and D. Abbott, "Wavelet de-noising of optical terahertz pulse imaging data," *Fluctuation and Noise Letters* **1**, L65-L69 (2001).
15. D. M. Mittleman, M. Gupta, R. Neelamani, R. G. Baraniuk, J. V. Rudd, and M. Koch, "Recent advances in terahertz imaging," *Appl. Phys. B* **68**, 1085-1094 (1999).
16. S. G. Mallat, *A Wavelet Tour of Signal Processing* (Elsevier Academic Press, Burlington, MA, USA, 1999), Chapter 7, 9, 10.
17. P. Moulin, "Wavelet thresholding techniques for power spectrum estimation," *IEEE Trans. Signal Processing* **42**, 3126-3136 (1994).
18. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* **81**, 425-455 (1994).