Signal restoration from atmospheric degradation in terahertz spectroscopy

Seong G. Kong[1,a)] and Dong H. Wu[2]
[1]*Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA*
[2]*Materials Science and Technology Division, Naval Research Laboratory, Washington, DC 20375, USA*

(Received 21 November 2007; accepted 20 March 2008; published online 5 June 2008)

This paper presents a method of restoring signals in terahertz spectroscopy by removing the distortion from the observed terahertz signals. The distortion is generated by the absorption and scattering of gas molecules and water vapor in the atmosphere during the transmission of terahertz beams through the air from the source to the spectrometer. Such atmospheric degradation causes spurious spectral dips and peaks in the terahertz spectrum, which often obscure the spectral peaks specific to the material of interest. This fact makes it challenging to measure the terahertz spectroscopic signatures of objects in a humid air environment, even at a short distance of approximately 1 m. A terahertz signal restoration filter based on a nonlinear artificial neural network model effectively removes noisy absorption peaks in terahertz spectra caused by atmospheric degradation. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2931946]

I. INTRODUCTION

The terahertz region of the electromagnetic spectrum, which lies between the microwave and the far infrared, has recently attracted a great deal of attention.[1–6] Since the typical frequencies of molecular motion of many chemicals of interest are within the terahertz region, terahertz spectroscopy can provide information that is unavailable in conventional spectroscopy. In particular, terahertz spectroscopy has great potential for detecting various chemical or biological agents through the identification of unique spectral absorption patterns of the material, often in the frequency range of 0.1–3.5 THz. In addition, terahertz beams can penetrate most nonmetallic materials, such as paper, textiles, and wood panels. Terahertz radiation does not cause harmful ionization effects unlike x rays or γ-rays because of their low photon energy. Such advantages have greatly promoted terahertz technology for the detection of concealed weapons or explosives, with no health and damage concerns for both operators and targets.[5–7] Despite these advantages, however, terahertz spectroscopy in real-world applications has not yet been realized, partly due to technical challenges such as rapid attenuation of terahertz signals in the atmosphere and spurious peaks in the spectra produced mainly by water vapor, oxygen, and carbon dioxide. These molecules create several absorption bands or frequency bands of high attenuation. The atmospheric degradation of the terahertz signal makes it difficult to identify material-specific signatures in the terahertz spectra.

Signal restoration[8] refers to a deconvolution process to recover the original signal from the observed signal distorted by a degradation process and the noise in the transmission path. The objective of terahertz signal restoration is to remove the effects of atmospheric degradation from spectroscopic measurements. Terahertz signal restoration is particularly important when the measurement is to be made to identify an object from a distance in humid open air. Many signal restoration schemes assume that the system function of the degradation process is known or estimated by linear, time-invariant models whose parameters can be found using a set of observed input-output data. The identification process then selects a model structure, computes the best model in the structure based on the information in the data, and evaluates the properties of the model. Since atmospheric degradation of terahertz radiation is highly nonlinear and time varying in nature, the dynamic characteristics of an underlying atmospheric degradation process require a nonlinear, adaptive estimation model.

This paper presents a signal restoration technique based on an artificial neural network model to remove spurious peaks in the terahertz spectra degraded by the atmosphere. Artificial neural networks,[9] inspired by the ability of the human brain to learn from observations and to generalize by abstraction, offer a model-free approximation of highly complex input-output characteristics of an underlying atmospheric degradation process. Neural networks are trained to learn any arbitrary, nonlinear input-output relationship from a set of training data. It has been proven that multilayer feed-forward artificial neural networks with one hidden layer can approximate any multivariable function to any desired degree of accuracy.[10] The training process involves adjustment of the network internal connections or weights in order to minimize the estimation error over a set of examples. Trained neural networks are used as signal restoration filters to restore terahertz signals from atmospheric degradation.

II. TERAHERTZ SPECTROSCOPY

Our time-domain terahertz spectrometer consists of a femtosecond laser, a photoconductive terahertz wave emitter, and an electro-optic (EO) detector.[1,4,11–14] Figure 1 shows an optical layout of the spectrometer used in this experiment, which is adopted from the technique of Wu *et al.*[11] The laser source is a compact mode-locked fiber laser (IMRA femtolite

[a)]Electronic mail: skong@temple.edu.

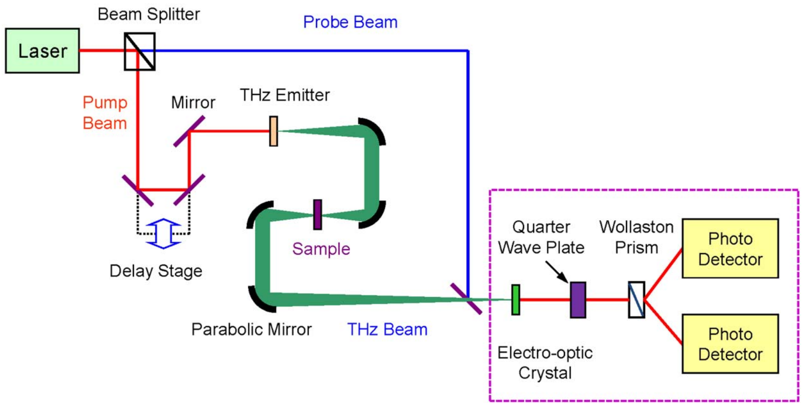

FIG. 1. (Color online) A schematic diagram of the time-domain terahertz spectrometer.

780) that emits approximately 100 fs laser pulses with a center wavelength of 780 nm, a repetition rate of 48 MHz, and an average power of 30 mW. The terahertz emitter is a photoconductive switch fabricated on a low-temperature-grown GaAs chip and modulated by the amplified reference signal of a lock-in amplifier. The modulation amplitude and frequency are typically 200 V and 9.5 kHz, respectively. The average power of our terahertz beam is estimated to be about a few microwatts. The terahertz pulses, generated by the emitter, are focused onto a sample by a pair of off-axis parabolic mirrors. The terahertz beams, transmitted through or reflected by the sample, are then sent to an EO detector by another pair of parabolic mirrors. The terahertz path length from the terahertz emitter to the ZnTe EO crystal is about 1 m. The EO detector consists of a ZnTe crystal, a quarter wave plate, a Wollaston prism, and a pair of photodiodes.[11] The probe beam, aligned by a pellicle beam splitter, travels collinearly with the terahertz beam through the ZnTe EO crystal. Due to the EO effect of the ZnTe crystal, the terahertz field results in a polarization change of the probe beam while it travels through the ZnTe crystal. The Wollaston prism separates the vertical-polarization and horizontal-polarization components of the probe beam. These vertical and horizontal components are sent to the photodetector pair, which produces the differential photocurrents. Since the polarization of the probe beam is initially set to 45°, the vertical and horizontal components are the same when the terahertz field is zero. Hence, there is no differential photocurrent. The difference between the electric currents from the photodiode pair is measured by a digital signal processing (DSP) lock-in amplifier (Stanford Research Systems SR-830) and a desktop computer.

III. MODELING OF THE TERAHERTZ SIGNAL-DEGRADATION PROCESS

A. Terahertz signal degradation by atmospheric attenuation and scattering

Terahertz beams traveling through the ambient atmosphere are absorbed and scattered by the molecules of atmospheric gases as well as pollutants. The degree of absorption and scattering depends on the size, shape, and dielectric constant of the atmospheric gases. The absorption and scattering rates vary as a function of the frequency of the electromagnetic beam.[15] In addition, the rotational motions of most molecules resonate at terahertz frequencies, which result in well-defined absorption spectra. Atmospheric constituents, such as oxygen, water vapor, and carbon dioxide, interact with terahertz signals. They not only attenuate the signal, but they also generate a number of absorption lines within the terahertz frequency band.[2,3] Such effects are observed in our experimental results. Figure 2(a) shows time-domain waveforms of our terahertz signals measured at room temperature of 65 °F in two humidity conditions: a low-humidity setting at approximately 5% relative humidity (RH) and an open-air

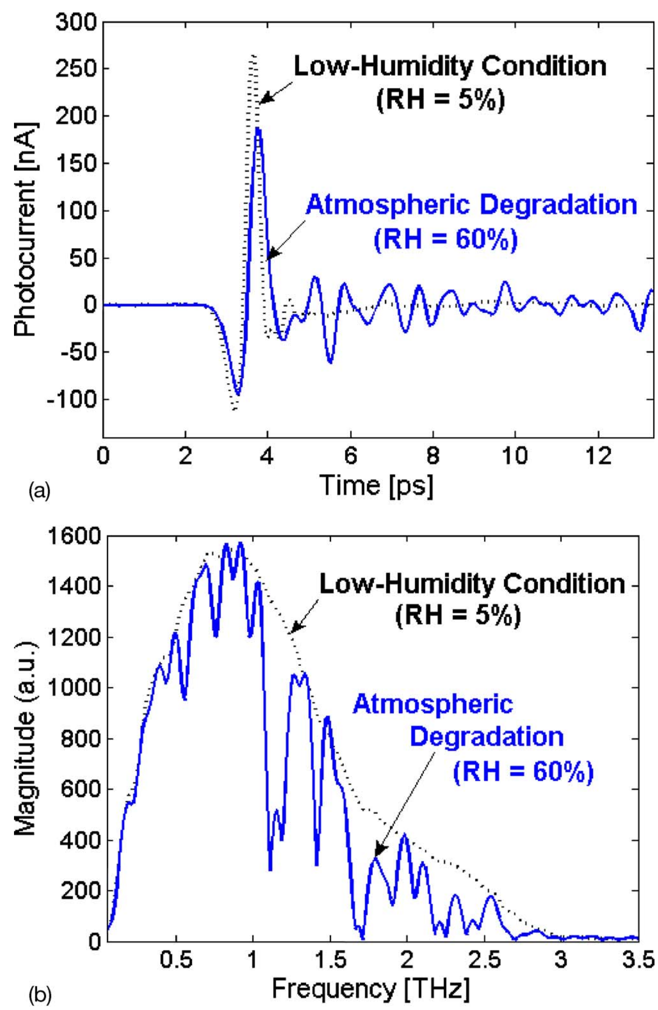

FIG. 2. (Color online) The effects of atmospheric degradation in terahertz spectroscopy signals: (a) time waveforms and (b) Fourier spectra.

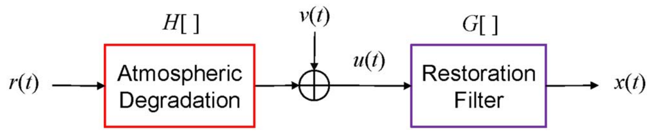

FIG. 3. (Color online) Restoration process for degraded terahertz signals.

environment at 60% RH. A low-humidity environment was created by filling the measurement chamber of the terahertz spectrometer with dry nitrogen (N_2) gas. While terahertz time-domain waveforms measured in dry nitrogen gas show a smooth tail after exhibiting a few sharp peaks in the spectrum, the waveform measured in the open-air environment reveals an attenuated main peak with a slight time delay and strong fluctuations in the tail, indicating signal attenuation and interference by water and air molecules. Indeed, as shown in Fig. 2(b), the terahertz signal taken in a dry nitrogen environment reveals a smooth spectrum, whereas the Fourier spectrum of the signal obtained in an open-air environment exhibits several strong absorption bands, such as those at 1.1 and 1.7 THz, which are indicative of water vapor absorption. In order to obtain clean, undistorted spectra, terahertz spectroscopic measurements must be carried out in low-humidity environments. The measurements in the field, however, are compromised by the environmental conditions.

B. Modeling of the atmospheric degradation process

The observed signal in terahertz spectroscopic measurements can be modeled as the output of an atmospheric degradation process in the presence of external noise for an input signal generated from a terahertz emitter. Let the atmospheric degradation process be a nonlinear system with a characteristic function H. Then, the observed terahertz signal can be represented by the model

$$u(t) = H[r(t)] + \nu(t), \tag{1}$$

where $u(t)$ denotes the observed degraded terahertz signal, $r(t)$ is the original undistorted terahertz signal, and $\nu(t)$ denotes the additive external noise, which is assumed to be zero mean and white Gaussian. Figure 3 shows a restoration process of terahertz signals from atmospheric degradation.

A signal restoration filter G takes the degraded signal $u(t)$ as an input and produces an output $x(t)$ that closely approximates the original signal $r(t)$,

$$x(t) = G[u(t)]. \tag{2}$$

Many signal restoration schemes assume that the degradation process H is known or estimated by a linear time-invariant model. Then, an approach to recover the signal $r(t)$ is to find the inverse operator such that

$$x(t) = H^{-1}[u(t)]. \tag{3}$$

In terahertz spectroscopic measurement, since the atmospheric degradation process is highly nonlinear and complex, closed-form mathematical models for such an inverse process may not be readily available. Even if the inverse of the degradation process can be approximated, the result of signal restoration will not be satisfactory due to the noise. Alternatively, one can develop parametrized nonlinear network models for characterizing the input-output relationship of an atmospheric degradation process.

The objective of terahertz signal restoration is to find a restored signal $x(t)$ that is a faithful reproduction of the original signal $r(t)$ measured in low-humidity conditions. In this paper, the absorbance is used as a metric to determine if the restored signal $x(t)$ is sufficiently close to the reference signal $r(t)$ measured in a low-humidity condition. The absorbance of a signal $x(t)$ with respect to a reference signal can be defined as

$$\text{Absorbance} = -\log_{10}\left(\frac{A^2_{\text{sample}}}{A^2_{\text{ref}}}\right), \tag{4}$$

where A_{sample} and A_{ref} denote the magnitude Fourier spectra of the sample and the reference signal, respectively.

IV. TERAHERTZ SIGNAL RESTORATION USING ARTIFICIAL NEURAL NETWORKS

Terahertz signal restoration filters can be made by employing parametrized nonlinear models, such as artificial neural networks. Artificial neural networks offer a model-free approach to the estimation of input-output characteristics of underlying nonlinear systems. Instead of theoretical analysis and development for a new model, the neural network tailors itself to the training data. The purpose of neural network modeling is to develop a parametrized nonlinear model,

$$\boldsymbol{x} = f_{\text{NN}}(\boldsymbol{u},\boldsymbol{w}), \tag{5}$$

that characterizes an unknown restoration process $G[\,]$ through a set of observed input-output data pairs. The terahertz signal restoration filter finds an undistorted signal $x(t)$ from the degraded signal $u(t)$. Here, $\boldsymbol{u}$ denotes a vector of an input signal and $\boldsymbol{w}$ is a parameter vector of the internal connections of the network. In this paper, a multilayer feedforward neural network model[16] is used as a terahertz signal restoration filter, with an input layer of $(2m+1)$ nodes and a single output node in the output layer. An input vector $\boldsymbol{u}$ consists of $(2m+1)$ delay-line elements of an input terahertz signal in a noncausal fashion,

$$\boldsymbol{u} = [u(t-mI), \ldots, u(t), \ldots, u(t+mI)]^T, \tag{6}$$

where I denotes the interval between adjacent discrete data samples. The neural network was trained to determine the terahertz signal restoration filter characteristic function $G[\,]$ from training data pairs $(\boldsymbol{u}_j, d_j)$, where d_j denotes the desired output for a given input $\boldsymbol{u}_j$. The output of each layer is computed by a nonlinear activation function of a weighted sum of inputs from the previous layer,

$$y_j(t) = g\left(\sum_{k=-m}^{m} w_{jk}u(t-kI) + w_{j0}\right), \tag{7}$$

where g denotes a nonlinear activation function, such as logistic-sigmoid and radial-basis functions. The output neuron has a linear activation function.

The training process of a neural network involves the adjustment of internal connections or weights of the network

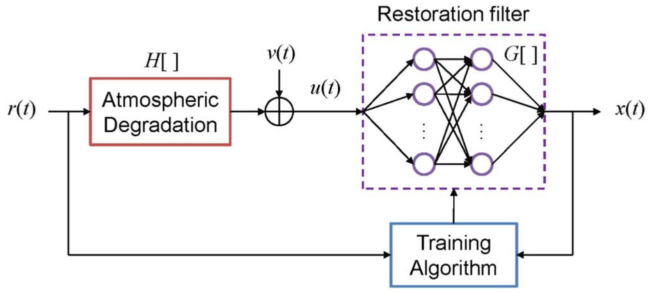

FIG. 4. (Color online) Training of the terahertz restoration filter based on artificial neural networks.

to minimize the estimation error over a set of training samples. If the training data for a neural network are general enough, spanning the entire ranges of process parameters, the resulting model will capture the complexity of the process, including nonlinearities and parameter cross couplings over the same ranges. The network can be refined at any time with the addition of new training data. Figure 4 shows the training process of the restoration filter based on a multilayer feed-forward neural network model using the back-propagation algorithm. The back-propagation training algorithm, based on a gradient descent search, determines the parameters of a multilayer feed-forward neural network model by minimizing the sum of errors $E(w)$ between the desired output and the actual output,

$$E(w) = \frac{1}{2}\sum_{j}(d_j - y_j)^2. \tag{8}$$

Here, y_j denotes the actual neural network output and d_j is the corresponding desired output. The connection weights preceding each output node are updated according to

$$w_{ij}(t+1) = w_{ij}(t) + \eta\varepsilon_j u_i, \tag{9}$$

where η is a correction gain and ε_j is the correction factor computed from the derivative of the sum of errors $E(w)$. Small values of correction gain result in slow training convergence speed, while large correction gain values increase convergence speed but with possible convergence instability. The network repeats the calculation of output values for all the input data in the training data set, compares them to the desired output values, and readjusts the network parameters of the back-propagation learning algorithm. This iterative tuning cycle is repeated until a calculated output converges sufficiently close to the desired output or until an iteration limit is reached. With large error tolerance, the network yields inaccurate network mappings, while with small error tolerance, the network requires considerably more training iterations. In network weight adjustment, the error measure tends to decrease as the network is refined by training iterations. The trained network is evaluated using the data that have not been used for training.

V. EXPERIMENTAL RESULTS

Several terahertz spectroscopic measurements were made in two different conditions: a low-humidity environment filled with dry nitrogen gas at ~5% RH and an open-air environment at ~60% RH. From these measurements, we selected 42 data sets of which we used randomly selected 31 data sets for training the signal restoration filter. The remaining 11 data sets were fed into the trained networks for evaluation. In Fig. 5(a), we compare three sets of time-domain data: (1) the dashed line represents the data measured in a dry nitrogen environment, (2) the dotted line represents the data in an open-air environment, and (3) the solid line represents the open-air data restored by our signal processing filter. Our signal restoration technique effectively removes the ringing fluctuation associated with water absorption. The restored data appear to be smooth and clean, similar to the data measured in low-humidity conditions. Most of the fluctuations due to water absorption are removed from the time-domain waveform, while the attenuation and time delay of the main peak remain the same. However, signal attenuation and time delay do not affect the location of the material-specific spectral peaks. This signal restoration technique satisfies our initial purpose, which is to discern the characteristic spectra of the object in terahertz spectroscopic measurements in humid atmosphere. The signal restoration effect becomes obvious in the Fourier transformed spectra. In Fig. 5(b), the solid line indicates the magnitude spectrum of the restored signal from atmospheric degradation, which

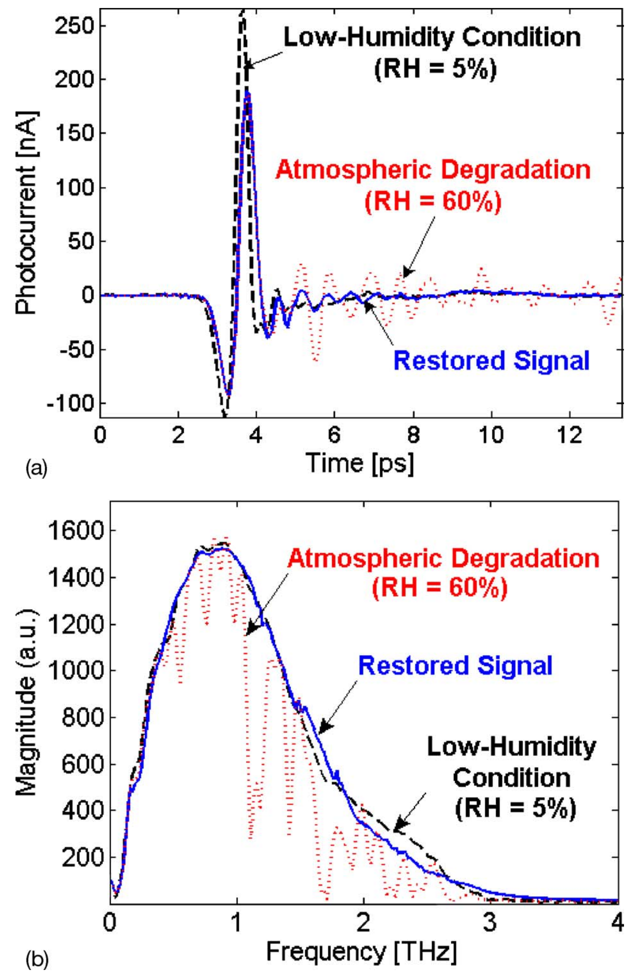

FIG. 5. (Color online) Restoration of terahertz signals degraded by atmospheric propagation: (a) time waveforms and (b) Fourier spectra magnitudes.

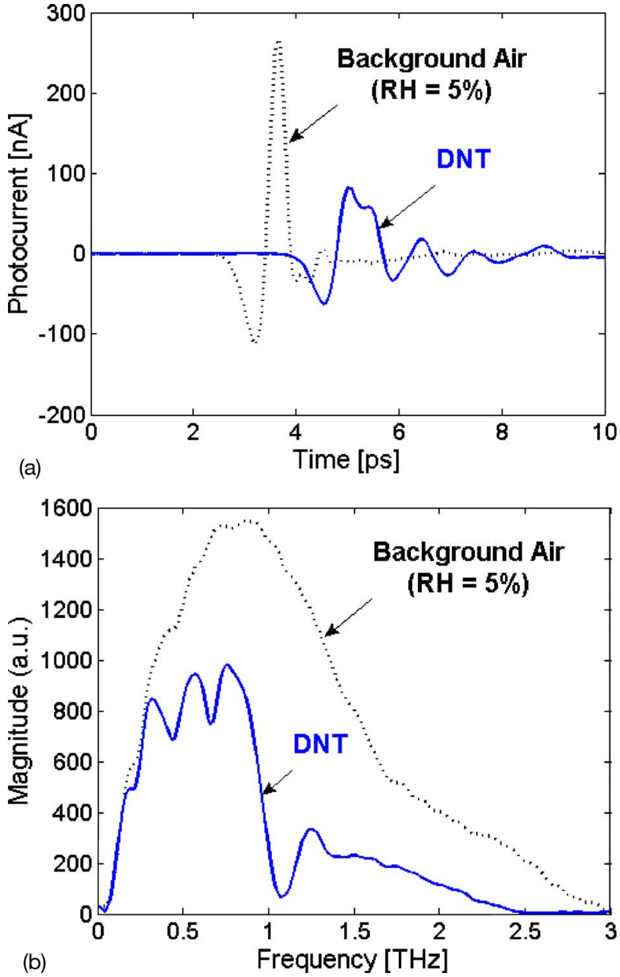

FIG. 6. (Color online) Terahertz signals of background air and a DNT sample measured in low-humidity conditions: (a) time waveforms and (b) Fourier spectra.

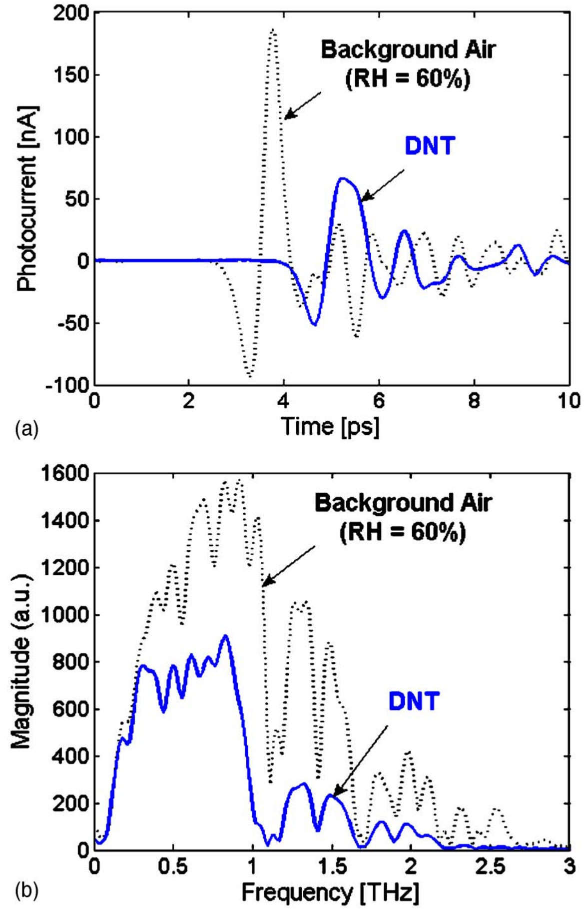

FIG. 7. (Color online) Terahertz signals of the background air and a DNT pellet, measured in an open–air environment at 60% humidity: (a) time waveforms and (b) Fourier spectra.

tends to show noticeable absorption bands. The restored signal shows no noticeable absorption bands and appears to be very similar to the data measured in dry nitrogen environments, except for an abrupt drop in intensity at around 1.7 THz.

Next, we tested the restoration filter for the restoration of terahertz signals measured from a chemical substance in the humid air. We measured the spectrum of a dinitrotoluene (DNT) sample, a surrogate explosive material, in dry nitrogen and open-air environments. The DNT sample was prepared in the form of a circular pellet with 25.4 mm diameter and 2.98 mm thickness. The terahertz beam that was focused on the sample was approximately 2 mm in diameter. Figure 6 shows typical time-domain waveforms and Fourier spectra of the reference terahertz signals measured in a dry nitrogen environment. Figure 7 shows similar measurements of the same DNT sample in an open-air environment. A number of signal fluctuations at higher frequencies in the time-domain data are observed. The Fourier spectra show many spectral dips due to the absorption of water vapor in the terahertz range.

Figure 8 shows the absorption spectra of DNT in low-humidity and open-air conditions at 60% RH. The absorption spectrum of the DNT signal measured in the air contains many spurious spectral peaks in the 1–3 THz region. Therefore, a unique spectral peak at $\sim$2.6 THz was obscured.

A multilayer neural network model is composed of an input layer, two hidden layers, and an output layer. The two hidden layers have 20 and 10 nodes, respectively. The number of input nodes is 7 ($m=3$), and the data interval is chosen to be $I=3$. Figure 9 demonstrates the performance of the

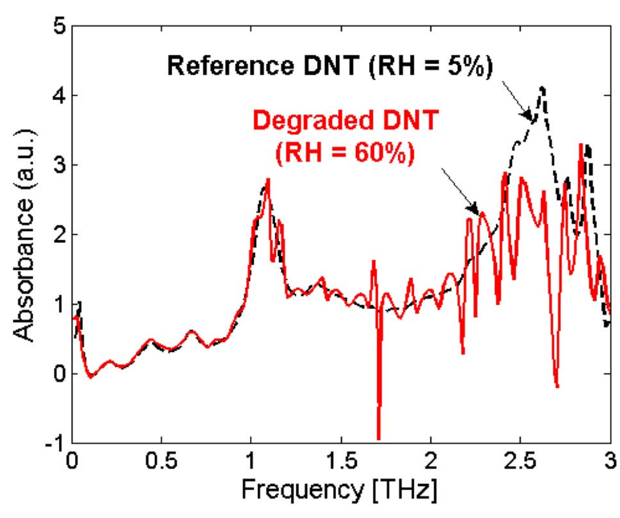

FIG. 8. (Color online) Absorption spectra of DNT measured at low-humidity and humid open-air conditions.

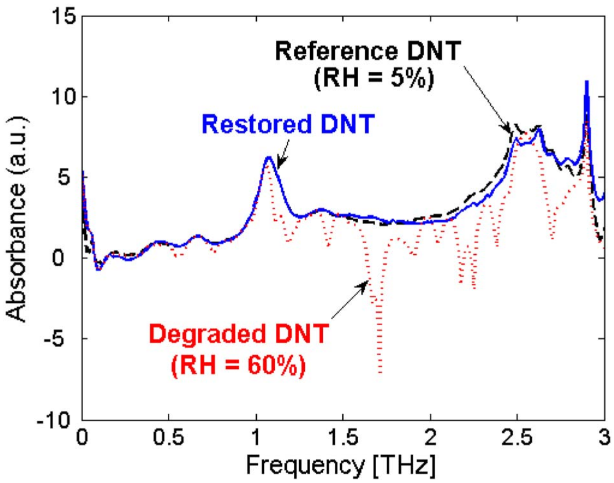

FIG. 9. (Color online) Restoration of terahertz signals from atmospheric degradation.

terahertz restoration filter. The spectrum of the DNT signal measured in the atmosphere of 60% RH is sufficiently close to that of the reference DNT signal over the frequency range of 0.1–3.0 THz. Most spurious spectral peaks were removed from the spectrum of the restored signal.

VI. CONCLUSION

Terahertz spectroscopy promises an innovative sensing technique that conventional spectroscopy may not offer since the typical frequencies corresponding to the molecular motion of many chemicals of interest are within the terahertz region. Terahertz beams can penetrate most nonmetallic materials and do not cause harmful ionization effects. However, real-world applications of terahertz spectroscopy have been limited by technical challenges, such as rapid attenuation of terahertz signals in the atmosphere and spurious peaks in the spectra produced mainly by water molecules, oxygen, and carbon dioxide. These molecules create several frequency bands of high attenuation in the terahertz range. Signal degradation by the atmosphere makes it difficult to identify unique terahertz spectral signatures of the materials even at a short distance in a humid open-air environment. In our experiments, the terahertz path length was about 1 m long and the terahertz spectra were corrupted considerably even at 60% RH, which made it difficult to discern the characteristic peaks of a target material such as DNT. Our signal restoration technique removed the noise, restored the corrupted spectra, and produced discernible spectra. The signal restoration method presented in this paper offers a cost-effective way of solving the spectral corruption problem. Experiments with DNT show that the terahertz signal restoration filter based on multilayer neural networks effectively removes spurious peaks in the terahertz spectra caused by atmospheric degradation. The spectral peaks of the restored signal are consistent with the ideal spectrum of DNT measured in low-humidity conditions. Our signal restoration technique increases the signal-to-noise ratio by removing the ringing fluctuation due to water absorption in the atmosphere. The restored spectral peaks are free from the noise and are comparable to those measured in low-humidity conditions. The current limitation in the sensing range of a terahertz spectrometer may not exceed a few tens of meters, even with employing a relatively strong terahertz source that can produce a terahertz beam of a hundred microwatts. Further development of this signal restoration technique, along with reasonably strong terahertz sources, is expected to extend the sensing range of a terahertz spectrometer to over a hundred meters.

ACKNOWLEDGMENTS

This research was supported in part by the Naval Research Laboratory.

[1] A. Nahata, D. H. Auston, T. F. Heinz, and C. Wu, Appl. Phys. Lett. **68**, 150 (1996).
[2] M. van Exter, Ch. Fattinger, and D. Grischkowsky, Opt. Lett. **14**, 1128 (1989).
[3] D. M. Mittleman, R. H. Jacobsen, R. Neelamani, R. G. Baraniuk, and M. C. Nuss, Appl. Phys. B: Lasers Opt. **67**, 379 (1998).
[4] H. Liu, Ph.D. thesis, Rensselaer Polytechnic Institute, 2005.
[5] Y. C. Shen, T. Lo, P. F. Taday, B. E. Cole, W. R. Tribe, and M. C. Kemp, Appl. Phys. Lett. **86**, 241116 (2005).
[6] D. H. Wu and J. R. Meyer, Proc. SPIE **5411**, 187 (2004).
[7] F. Huang, B. Schulkin, H. Altan, J. F. Federici, D. Gary, R. Barat, D. Zimdars, M. Chen, and D. B. Tanner, Appl. Phys. Lett. **85**, 5535 (2004).
[8] R. W. Schafer, R. M. Mersereau, and M. A. Richards, Proc. IEEE **69**, 432 (1981).
[9] T. Poggio and F. Girosi, Proc. IEEE **78**, 1481 (1990).
[10] K. Hornik, M. Stinchcombe, and H. White, Neural Networks **2**, 359 (1989).
[11] Q. Wu, M. Litz, and X.-C. Zhang, Appl. Phys. Lett. **68**, 2924 (1996).
[12] A. Garzarella, S. B. Qadri, T. J. Wieting, and D. H. Wu, J. Appl. Phys. **98**, 043113 (2005).
[13] A. Garzarella, S. B. Qadri, T. J. Wieting, and D. H. Wu, J. Appl. Phys. **97**, 113108 (2005).
[14] A. Garzarella, S. B. Qadri, T. J. Wieting, and D. H. Wu, Appl. Phys. Lett. **88**, 141106 (2006).
[15] L. A. Klein, *Millimeter-Wave and Infrared Multisensor Design and Signal Processing* (Artech House, Norwood, MA, 1997).
[16] S. W. Moon and S. G. Kong, IEEE Trans. Neural Netw. **12**, 307 (2001).