

# Two Wavelet-based Algorithms for Chemical Recognition Using Transmission Terahertz Spectral Imaging Through Turbid Media

Mahmoud E. Khani and M. Hassan Arbab  
Stony Brook University, Stony Brook, NY, 11794 USA

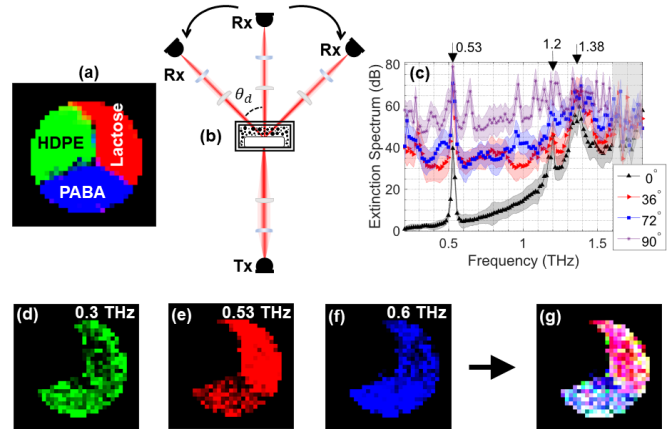
**Abstract**—Dielectric heterogeneity in addition to wavelength-scale air voids and granular particles lead to significant electromagnetic scattering in the transmission-mode THz-TDS. These scattering processes result in the distorted or obscured resonant signatures, a frequency-dependent extinction loss, and the appearance of anomalous spectral artifacts, imposing severe challenges on the real-world automatic material characterization schemes. In this work, we use a combination of the wavelet multiresolution analysis, the principal component analysis, and the bimodality coefficient to simultaneously identify all the resonant frequencies of a heterogeneous sample embedded beneath a highly scattering turbid medium.

## I. INTRODUCTION

THE low-frequency vibrational modes associated with the intra- or intermolecular collective motions or the crystalline lattice modes of molecular crystals appear as resonant signatures in the dielectric functions of many materials of interest. These resonant modes can be utilized as characteristic fingerprints for material characterization in the security, pharmaceutical, and nondestructive testing applications [1]. However, the multiple scattering events risen from the dielectric heterogeneity and the wavelength-scale granular particles can impede resolving of these resonant signatures by the diffuse THz waves. In particular, the electromagnetic scattering results in obscured resonant fingerprints, a frequency-dependent extinction loss, and the appearance of anomalous spectral artifacts, complicating the identification of the characteristic resonant frequencies. Solutions proposed in the literature to overcome the scattering effects are mainly focused on the ballistic transport of the THz waves. They also rely on some *a priori* information on the chemicals and the scattering characteristics or demand a multitude of spatial or angular averages [1], [2]. Here, we demonstrate the implementation of two techniques based on the wavelet multiresolution analysis (MRA), the principal component analysis, and the bimodality coefficient to simultaneously identify all the characteristic resonant frequencies of a heterogenous sample from diffuse THz waves recorded at detection angles as high as  $90^\circ$  with respect to the incident light. This approach demonstrates unprecedented material characterization capabilities in the presence of severe scattering effects.

## II. RESULTS

The imaging target, a heterogenous sample disk with 50 mm diameter, was composed of  $\alpha$ -lactose monohydrate, 4 aminobenzoic acid (PABA), and high-density polyethylene (HDPE), as shown in Fig. 1(a). This sample disk was embedded beneath a highly scattering turbid medium made from 5 gr LDPE with a  $180 \mu\text{m}$  mean particle size. We used the apparatus shown schematically in Fig. 1(b) to measure the diffuse THz waves at oblique detection angles. Figure 1(c) compares the



**Fig. 1.** (a) A spectral image of the sample target composed of 3 different chemicals, (b) schematic of the imaging apparatus, (c) comparison of the extinction spectra at different diffuse scattering angles with the forward-scattered beam at  $\theta_d = 0^\circ$  for  $\alpha$ -lactose monohydrate, (d-g) spectral images at 0.3, 0.53, and 0.6 THz and their combination affected by the volume scattering.

extinction spectra of the  $\alpha$ -lactose monohydrate region measured at the forward and oblique detection angles. It can be seen that at the higher detection angles, the scattering effects are dominant in the extinction spectra, and they can obscure the resonant signatures. Moreover, Figs. 1(d-g) demonstrate that the chemical maps formed from the scattered spectra will not yield the accurate area covered by each material.

In this work, we use the wavelet multiresolution analysis of the extinction spectrum  $\varepsilon(f)$  at each pixel given by,

$$\varepsilon(f) = \sum_{k=0}^{N_j-1} \tilde{g}_j(k) \tilde{v}(J, k + f \bmod N) + \sum_{j=1}^J \sum_{k=0}^{N_j-1} \tilde{h}_j(k) \tilde{w}(j, +f \bmod N) \quad (1)$$

to separate the characteristic spectral signatures from the scattering-induced artifacts. In (1),  $\tilde{g}_j(k)$  and  $\tilde{h}_j(k)$  represent the  $J$ th level scaling filter and the  $j$ th level wavelet filter, respectively. Moreover,  $\tilde{v}(J)$  and  $\tilde{w}(j)$  are the scaling and wavelet coefficients obtained using the maximal-overlap discrete wavelet transform (MODWT) pyramid algorithm [3], [4]. Accordingly, we need to identify the wavelet basis functions that better capture the resonant signatures [5], [6]. In this work, we compare the effectiveness of two techniques based on the principal component analysis (PCA) of the wavelet coefficients and an optimization over the total variation of the bimodality coefficients in addressing this problem.

In the first technique, which is more efficient computationally, because the wavelet multiresolution analysis coefficients (wavelet detail vectors) create a high-dimensional feature space, e.g. eleven new dimensions per pixel for  $J = 10$

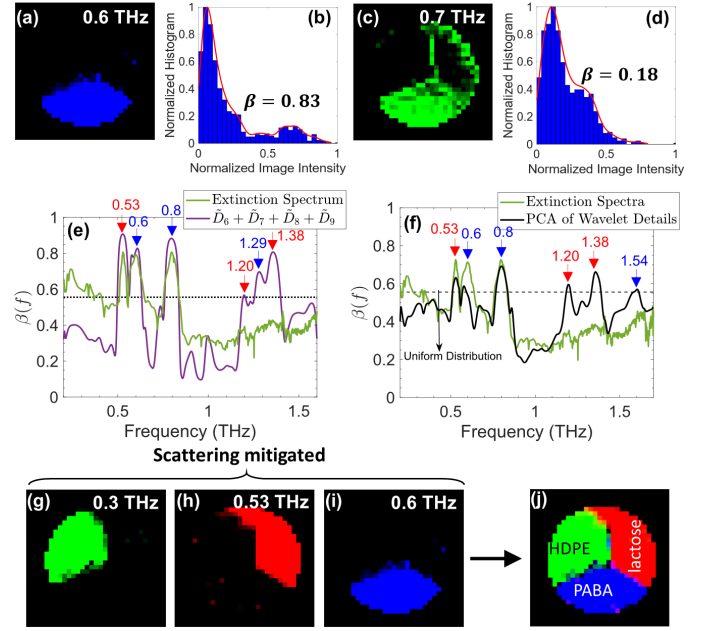
levels of decomposition, an orthogonal transformation of the wavelet-domain features using the PCA can yield a lower-dimensional wavelet feature space with linearly uncorrelated variables. Reconstruction of the extinction spectra from the principal component-transformed wavelet detail vectors maximizes the variance while minimizing the total reconstruction error. The second algorithm is also defined based on the modal characteristics of the spectral images assessed quantitatively by the bimodality coefficient given by,

$$\beta(f) = \frac{\gamma^2(f) + 1}{\kappa(f) + \frac{3(n-1)^2}{(n-2)(n-3)}}, \quad (2)$$

where  $\gamma(f)$  and  $\kappa(f)$  are the skewness and kurtosis of the  $n$ -pixel image formed at the frequency  $f$ , respectively. A bimodality coefficient greater than  $5/9$  (the bimodality coefficient of a uniform distribution) can indicate the presence of a resonant frequency. For example, Figs. 2(a-b) illustrate a spectral image formed at the PABA resonant frequency at 0.6 THz and its corresponding histogram distribution. The histogram in Fig. 2(b) demonstrates two well-separated peaks, which is also confirmed by its bimodality coefficient ( $\beta = 0.83$ ) computed using (2). Whereas Figs. 2(c-d) illustrate that the histogram of an image formed at a nonresonant frequency is unimodal, which is also confirmed by its lower bimodality coefficient,  $\beta = 0.18$ . Therefore, for finding the wavelet detail vectors that are associated with the resonant signatures, we solve an optimization problem given by,

$$\min_{\tilde{D}_1, \dots, \tilde{D}_J} \frac{\sum_i |\beta(f_i) - \beta(f_i - 1)|}{K \sum_i \beta(f_i), \forall \beta(f_i) > \frac{5}{9}} \quad (3)$$

for all possible combination of the wavelet detail vectors over  $J$  dyadic scales,  $\tau = 2^0, 2^1, \dots, 2^{J-1}$ . The numerator in (3) represents the total variation of the  $\beta(f)$  along the spectrum. The denominator accounts for the bimodalities greater than  $5/9$ , and  $K$  represents the number of included levels of wavelet details. Noteworthy here, we demonstrate that after reconstruction of the extinction spectra by including only the wavelet details set by (3), the local maxima greater than  $5/9$  in the bimodality spectrum reveal the resonant frequencies. For example, solving (3) for the extinction spectra measured at  $\theta = 0^\circ$  reveals that reconstruction of  $\varepsilon(f)$  from the sixth to ninth-level MODWT details,  $\tilde{D}_6(f) + \tilde{D}_7(f) + \tilde{D}_8(f) + \tilde{D}_9(f)$ , yields simultaneous recognition of all the resonant frequencies using the bimodality spectrum. Figure 2(e) compares the bimodality coefficients of the images formed from the extinction spectra with those of the images formed from the mentioned combination of MODWT details. The red arrows in Fig. 2(e) point out the  $\alpha$ -lactose's resonances at 0.53, 1.20, and 1.38 THz, while the blue ones indicate the PABA's resonances at 0.6, 0.8, and 1.29 THz. Clearly, in the bimodality spectrum of the MODWT details, all the local maxima with  $\beta$  greater than  $5/9$  correspond to the resonant frequencies. However, in the  $\beta(f)$  of the extinction spectra, although based on the same criterion resonances at 0.53, 0.6, and 0.8 THz remain



**Fig. 2.** (a) A spectral image at the PABA's resonance at 0.6 THz, (b) histogram distribution of the image shown at (b), (c) a spectral image at 0.7 THz, which is a nonresonant frequency, (d) histogram distribution of the image in (c), (e) the bimodality coefficients of the images formed from the extinction spectra and the bimodality-based wavelet MRA in the range 0.2 - 1.6 THz, (f) the bimodality coefficients of the images formed from the wavelet PCA, (g-j) spectral images after the scattering effects are removed using the bimodality-based wavelet MRA technique.

distinguishable, the higher-frequency resonances at 1.20, 1.29, and 1.38 THz are completely obscured. Figure 2(f) also compares the bimodality coefficients of the images formed from the wavelet PCA and the original extinction spectra. It can be seen that in the images formed from the wavelet PCA, all the local maxima over the  $5/9$  threshold correspond to the resonant frequencies. Moreover, in comparison to the bimodality-based wavelet MRA results shown in Fig. 2(e), the wavelet PCA can detect the PABA's resonance at 1.54 THz as well. However, the 1.29-THz resonance does not reach to the  $5/9$  threshold. Moreover, in the wavelet PCA results, the bimodality values are smaller, which denotes images with lower contrasts. Figures 2(g-j) also show the spectral images after the scattering effects are removed using the bimodality-based wavelet MRA technique. By comparing them to the images shown in Figs. 1(d-g), it can be seen that after mitigating the scattering effects, all the images provide a clear representation of the area covered by each chemical.

## REFERENCES

- [1]. O. B. Osman and M. H. Arbab, *PLoS one*, vol. 14, 2019.
- [2]. Y. C. Shen, P. F. Taday, and M. Pepper, *Appl. Phys. Lett.*, Vol. 92, pp. 051103, 2008.
- [3]. M. H. Arbab, D. P. Winebrenner, E. I. Thorsos, and A. Chen, *Appl. Phys. Lett.*, vol. 97, pp. 181903, 2010.
- [4]. M. H. Arbab, D. P. Winebrenner, E. I. Thorsos, and A. Chen, in *Proc. SPIE 7601*, San Francisco, CA, USA, 2010, pp. 760106-1–760106-7.
- [5]. M. Ebrahimkhani and M. H. Arbab, in *2018 43rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*, Nagoya, 2018, pp. 1-2.
- [6]. M. E. Khani, D. Winebrenner and H. Arbab, in *IEEE Trans. Terahertz Sci. Technol.*, doi: 10.1109/TTHZ.2020.2997595.