# Identification of maize seeds by terahertz scanning imaging

**Meihong Lu (逯美红), Yan zhang (张 岩), Jinhai Sun (孙金海), Sijia Chen (陈思嘉), Ning Li (李 宁), Guozhong Zhao (赵国忠), and Jingling Shen (沈京玲)**

*Department of Physics, Capital Normal University, Beijing 10003*

Varietal purity is the most important quality parameter of maize seeds, which has direct and prominent influence on the output and quality of maize. For the first time, to our knowledge, we present a new kind of terahertz (THz) scanning imaging technology for identification of maize seeds. Terahertz images of DNA samples are obtained by point-by-point scanning imaging technology. Inspection and identification of specific kinds of seeds are realized successfully by using the method of component pattern analysis. In this method, what we need are only data of image and absorption spectral information of samples; no specific features of samples are required. This technology provides a new approach for the detection and identification in biology and it can also be extended to poison inspection.

*OCIS codes:* 100.1160, 100.2960, 100.5010, 110.0110, 110.2960.

The terahertz (THz) spectral region, which occupies an extremely large portion between the infrared (IR) and microwave bands, is a scientifically rich frequency band. THz time-domain spectroscopy (THz-TDS)[1] is an extremely promising technique for biomedical applications since rotational and vibrational transitions of many proteins and DNA molecules lie within the THz frequency region. Consequently, the time-resolved THz spectroscopy applied to examine and identify some biological samples is possible[2−5]. However, to date, there are few terahertz spectroscopy investigations on DNA molecules of maize seeds to implement the inspection and identification since characteristic "fingerprint" absorption spectra is not observed in the THz region. The traditional identification techniques of maize seeds, such as direct field inspection, electrophoretic technique, DNA (RAPD) molecule labeling technique and DNA fingerprint technique have some disadvantages, making the identification process intricate.

In contrast, the THz imaging technique[6−15] presented in this letter is more facilitated and feasible. We separated the spatial patterns of each DNA sample from their THz images successfully by using component pattern analysis method. The THz imaging technique is a promising application to determine and identify the quality and specific kind of inspection for different maize seeds. It resolved the identification problem of electrophoretic technique widely used in identification of maize seed purity. In addition, it can be used to realize the true-false identification that is impossible to be solved in the past using traditional identify technique. With this technique, agricultural experts can complete the identification, quality determination and inspection conveniently and rapidly. It is very helpful to the division of hybrid vigor, the quality innovation of seeds, and the selection of new species.

The schematic experimental setup used for two-dimensional (2D) THz scanning imaging is shown in Fig.1, which is similar to a typical THz-TDS system. A MaiTai laser beam with 100 fs pulse duration and an 80 MHz repetition rate at a central wavelength of 810 nm is used. The laser beam is separated into two, the pump light to illuminate THz emitter, and the probe light to drive the THz detector. THz pulses across 0.2–2.6 THz (3–100 $cm^{-1}$) are generated from the semiconductor surface of ⟨100⟩ InAs crystal. After the emitter, the first pair of parabolic mirrors in the path of THz beam collimate and focuse the THz beam onto the sample to be imaged, producing a nearly frequency-independent focal spot size of 1.5 mm. The sample is mounted on a translation stage that can be moved in the $X$-$Y$ plane perpendicular to the THz beam and be controlled by computer during the data acquisition. The other pair of parabolic mirrors collect and recollimate the transmitted pulse and focuses it onto the ⟨110⟩ ZnTe crystal. The output signal is fed to a lock-in amplifier. The time constant of lock-in amplifier is 100 ms. After passing through an analog-to-digital conversion board (AD), the output signal of the lock-in amplifier is recorded and stored in the computer during the continuous motion of the sample. The setup has a signal-to-noise ratio(SNR) of up to 600 and a spectral resolution of 50 GHz. In the experiment, the atmospheric moisture near the sample is less than 30% and the temperature is around 22 °C.

The object imaged in the experiment contains three different DNA (108DNA, 958DNA, and fishDNA) and an embryo (958pei) samples extracted from different kinds of maize seeds respectively. The extracted DNA samples are crushed and ground, and further dried in low temperature and vacuum condition. At last, the sample materials in the form of powder are carefully mixed with polyethylene powder in different proportion and pressed into discs at room temperature. Polyethylene has low absorption (<5 $cm^{-1}$ below 4 THz) and constant refraction index (almost is 1) in this range[16]. It serves as an additive for our samples. The diameters of the four

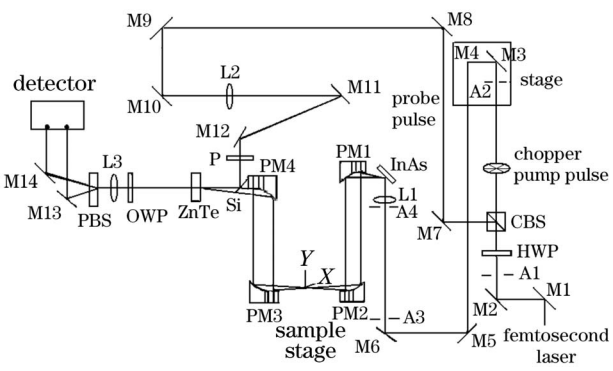


Fig. 1. Schematic diagram of the THz scanning imaging system. A: aperture; L: lens; M: mirror; P: polarizer.

samples are all 13 mm and the thickness is approximately identical. The four samples used in our experiment are packed together with a piece of lens paper and are placed in a 30×30 (mm) square orifice.

In the implementation of the THz point-by-point scanning imaging, the sample rather than the THz beam is moved over 25×25 (mm), which corresponds to 25×25=625 pixels. The THz pulse is scanned across the sample in a transmission mode and the THz time domain waveforms with sample information at every spatial point of the $X$-$Y$ plane are obtained. Each scan requires 65 s to collect time domain waveforms, leading to 10 hours for the image which contains 625 points. The time domain data obtained is converted to frequency domain by Fourier transform. Thus a great deal of spatial and spectral information about the samples is collected. The amplitude information of the spectrum at each point makes THz image, as shown in Fig. 2. The scale of the image is the transmitted intensity of the THz wave in frequency domain, which means that the smaller the absorption, the brighter the image.

To separate the spatial patterns of the samples, the component spatial pattern analysis methodis utilized in the data processing[17]. The principle of the method is described as follows. The studied object consisting of $M$ substances that have different spectral characteristics, is imaged in the whole useful range from 0.1 to 2.6 THz. Assume that the imaging system is linear one, the transmitted intensity can be described as

$$[I] = [S][P], \tag{1}$$

where $[I]$ is a $N \times L$ matrix of the observed image in which the row vectors $i_1, \cdots, i_N$ denote the image at the whole frequency band, the two-dimensional(2D) pixels are rearranged into one-dimensional(1D) vectors. $[S]$ is a $N \times M$matrix of the measured spectra of $M$ samples in which the column vectors $s_1, \cdots, s_M$ denote the spectrum data set of each sample; $[P]$ is the $M \times L$ matrix composed of $M$substances spatial patterns and the row vectors $p_1, \cdots, p_M$ denote the spatial pattern of the corresponding substance. Dimensions and elements of these three matrices are shown in Fig. 3. In the experiment, the data information is taken from the whole measured frequency band, thus the frequencies number $N$ is much larger than the samples number$M$. Therefore the matrix $[P]$ should be solved by use of the least-squares method as

$$[P] = ([S]^T[S])^{-1}[S]^T[I], \tag{2}$$

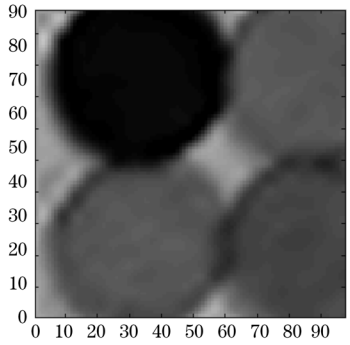


Fig. 2. A scanning transmitted THz image generating a matrix $[I]$.

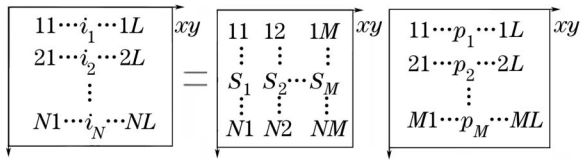


Fig. 3. Schematic illustration of Eq. (1).

where $T$ denotes a transposition operate. Therefore, if the matrices $[S]$ and $[I]$ are measured, the spatial distribution of each sample can be obtained.

As the incident THz waves are mainly attenuated by the absorption of samples, the transmitted THz waves satisfy the Lambert-Beer law. We take the ratio of the detected intensity divided by the reference intensity of THz frequency spectrum for the elements of matrix $[I]$. The absorption spectra of the four samples measured with the same experimental system are obtained as shown in Fig. 4. The corresponding absorption intensity in the whole frequency band is extracted to obtain the matrix $[S]$ of dimensions $N \times M = 256 \times 4$. In Fig. 4, it can be seen that the spectra of 108DNA, 958DNA and fishDNA have not obvious "fingerprint" absorption characteristic and there is little difference among them. Even this negligible difference can enable us to separate and identify the different samples using the method of component pattern analysis.

After obtaining the matrix $[I]$ from observed image and $[S]$ from the absorption spectra, the spatial pattern $[P]$with dimensions of $M \times L = 4 \times 625$ can be calculated by substituting known $[S]$ and $[I]$ into Eq. (2). Figure 5 shows the calculated results of the four different samples extracted from the matrix $[P]$. Since each row vector of matrix $[P]$ contains the information of each sample, the corresponding image of each sample should be obtained. The obvious images of Fig. 5 reveal that the corresponding spatial pattern of each sample is obtained and the four different samples have been separated. The detection and identification of DNA samples extracted from maize seeds are possible. The overlap of the images in Fig. 5 results from the non-uniform of mixture with polyethylene and the impurity of DNA in the samples. Moreover, in Fig. 5(d), the free space around the four samples becomes bright which is different from the brightness in Figs. 5(a), (b), and (c). The phenomenon is attributed to the low absorption of the fishDNA, which makes a low contrast with free space.

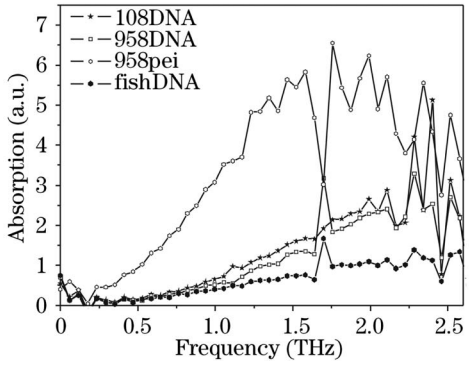


Fig. 4. Absorption spectra of the 108DNA, 958DNA, 958pei, and fishDNA.

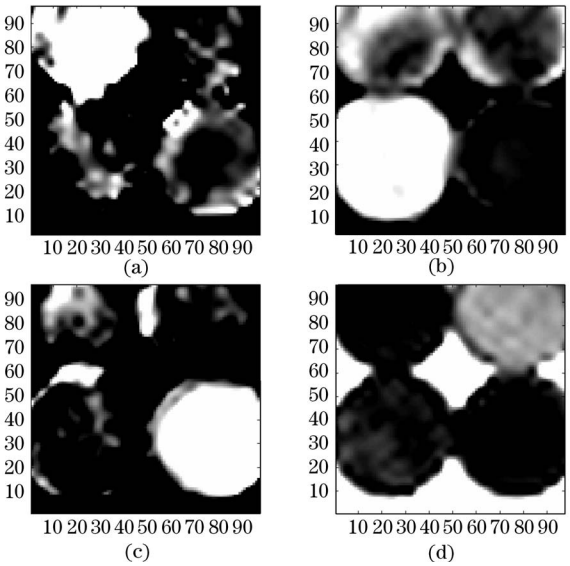


Fig. 5. Extracted spatial patterns of (a) 108DNA, (b) 958DNA, (c) 958pei, and (d) fishDNA. Four samples are basically distinguished and corresponding spatial patterns are obtained.

In conclusion, the THz-TDS technology is used to achieve THz scanning image of different DNA samples in a wide frequency region. Spatial patterns of DNA samples extracted from maize seeds are recognized basically by the method of component pattern analysis. The identification of the quality and species of maize seeds are feasible. Along with the demonstration of the imaging ability of the identification, some considerations that affect the results are deserved to be explored. Since the imaging process takes a relatively long time, the stability of the THz system during the experiments becomes important. The relative changes of the laser pulses and consequently THz pulse have affected the results. Furthermore, the water absorption resulting from the samples and/or environment in the experiment is quite strong. We believe that better results can be obtained if the whole system is placed in a closed box purged with nitrogen. The samples should be prepared more subtly and carefully to keep the DNA uniform and purity. This approach is expected to be explored widely for the identification of maize seeds.

This work was supported by the Beijing Natural Science Foundation of China (No. 6032006), the National Natural Science Foundation of China (No. 10390160), and the Science Foundation of Education Commission of Beijing, China (No. KM200310028115). We acknowledge the supply of the samples from Agriculture & Forestry Academy of Sciences and the help of Professor Jinglun Guo for the extraction of DNA. We thank Zhenwei Zhang for his help with the experiment. J. Shen is the author to whom the correspondence should be addressed (e-mail: sjl_phy@mail.snu.edu.cn)

## References

1. D. H. Auston, K. P. Cheung, J. A. Valdmanis, and D. A. Kleinman, Phys. Rev. Lett. **53**, 1555 (1984).
2. A. Wittlin, L. Genzel, F. Kremer, S. Häseler, A. Poglitsch, and A. Rupprecht, Phys. Rev. A **34**, 493 (1986).
3. A. G. Markelz, A. Roitberg, E. J. Heilweil,.Chem. Phys. Lett., **42**, 320 (2000).
4. T. R. Globus, D. L. Woolard, A. C. Samuels, B. L. Gelmont, J. Hesler, T. W. Crowe, and M. Bykhovskaia, Appl. Phys. **91,** 6105 (2002).
5. T. R. Globus, D. L. Woolard, T. Khromova, T. W. Crowe, M. Bykhovskaia, B. L. Gelmont, J. Hesler, and A. C. Samuels, Bio. Phys. **29,** 89 (2003).
6. T. S. Hartwick, D. T. Hodges, D. H. Barker, and F. B. Foote, Appl. Opt. **15,** 1919 (1976).
7. T. S. Hartwick, Proc. SPIE **108,** 139 (1977).
8. A. A. Lash and D. N. Yundev, Int. J. Infrared Millim. Waves **5,** 489 (1984).
9. B. B. Hu and M. C. Nuss, Opt. Lett. **20,** 1716 (1995).
10. D. M. Mittleman, S. Hunsche, L. Boivin, and M. C. Nuss, Opt. Lett. **22,** 904 (1997).
11. S. Hunsche, M. Koch, I. Brener, and M. C. Nuss, Opt. Commun. **150,** 22 (1998)
12. D. M. Mittleman, M. Gupta, R. Neelamani, R. G. Baraniuk, J. V. Rudd, and M. Koch, Appl. Phys. B **68,** 1085 (1999).
13. T. Löffler, T. Bauer, K. Siebert, H. G. Roskos, A. Fitzgerald, and S. Czasch, Opt. Express **9,** 616 (2001).
14. K. J. Siebert, H. Quast, R. Leonhardt, T. Loffler, M. Thomson, T. Bauer, H. G. Roskos, and S. Czasch, Appl. Phys. Lett. **80,** 3003 (2002).
15. B. Ferguson, S. Wang, D. Gray, D. Abbot, and X.-C. Zhang, Opt. Lett. **27,** 1312 (2002).
16. G. Zhao, R. N. Schouten, N. van der Valk, W. Th. Wenckenbach, and P. C. M. Planken, Rev. Sci. Instr. **73,** 1715 (2002)
17. Kawata, K. Sasaki, and S. Minami, J. Opt. Soc. Am. A **4,** 2101 (1987).