Optik 123 (2012) 1129–1132

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Identifying Radix Curcumae by using terahertz spectroscopy

Hong Li [a], ShaoQing Du [a], Le Xie [a], Lin Chen [a], Yan Peng [a], YiMing Zhu [a,*], Tao Wu [b], Han Li [c], Ping Dong [c], JiaTai Wang [c]

[a] *Engineering Research Center of Optical Instrument and System, Ministry of Education and Shanghai Key Lab of Modern Optical System, University of Shanghai for Science and Technology, Shanghai 200093, China*
[b] *Institute of Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China*
[c] *Shanghai YangPu Geriatric Hospital, Shanghai 200093, China*

ARTICLE INFO

Article history:
Received 20 January 2011
Accepted 28 July 2011

Keywords:
Terahertz time-domain spectroscopy
Support vector machines
Huangyujin
Lvyujin
Guiyujin
Wenyujin

ABSTRACT

The absorption spectra from 0.2 THz to 1.6 THz of four kinds of similar Chinese herbs, including *huangyujin*, *lvyujin*, *guiyujin* and *wenyujin*, have been investigated by terahertz time-domain spectroscopy (THz-TDS). Furthermore, by using support vector machines (SVM) method, the linear kernel function, the polynomial kernel function, and the radial basis kernal function are employed for separating four kinds of Radix Curcumae. The calculated results show that the accuracy of discrimination for these four kinds of Chinese herbs is 100%.

1. Introduction

As a big buried treasure of China, Chinese herbs, which can treat much chronic illness and consumptive sickness, have an irreplaceable role of western medicine. Especially, *huangyujin*, *lvyujin*, *guiyujin* and *wenyujin*, which are the same class, play an important role in daily life, i.e., for the treatment of common gynecological diseases. Because of its similar morphologic and different habitat, it is extremely difficult to distinguish them by using general approaches, such as infrared spectroscopy, gas chromatography, and capillary electrophoresis, which are destructive and low efficient for Chinese herbs detection.

Terahertz radiation, generally defined in the frequency range of 0.1–10.0 THz is between microwave and infrared range. Nowadays, terahertz time-domain spectroscopy (THz-TDS) has been utilized in a wide range of research fields such as biological molecules identification, materials investigation, explosives inspection, and medicine analysis [1–5]. Compared with general approaches, THz-TDS can provide absorption spectra of samples with high signal-to-noise (S/N) ratio and is a non-destructive and efficient detecting technique. However, THz absorption spectra of Radix Curcumae obtained by THz-TDS cannot be used as fingerprint spectra because they have no obvious characteristic absorption peaks. Therefore, SVM is employed, together with THz spectroscopy, to fast and effectively identify Radix Curcuma, even for all of Chinese herbs.

In this paper, the 0.2–1.6 THz absorption spectra of *huangyujin*, *lvyujin*, *guiyujin*, and *wenyujin* were obtained by THz-TDS. It is found that they have no characteristic absorption peaks in 0.2–1.6 THz range. Furthermore, with the help of SVM, four kinds of Radix Curcumae can be separated properly, and the accuracy of discrimination is 100%. The method and conclusion is significant to research and identify Chinese herbs.

2. Experiment

In this paper, as shown in Fig. 1, the THz-TDS was used. A mode-locked Ti:sapphire laser with the central wavelength at 800 nm, the full width at half maximum (FWHM) of spectral bandwidth at ∼20 meV, pulse duration around 80 fs, repeat frequency at 76 MHz, and output power at 1.1 W was applied. The laser was split into pump beam and probe beam by beam splitter. Pump beam modulated by an optical chopper was focused on Gallium Arsenide (GaAs) crystal. The purpose for the optical chopper, whose operation frequency is 3 kHz, is to reduce background noise and improve signal-to-noise ratio. At such condition, electrons were created near the bottom of the conduction band, as well as holes near the top of the valence band. The photoexcited electrons were accelerated by the applied electrical field and THz wave emitted out [6,7] from gallium arsenide crystal. As shown in Fig. 1, THz was focused on

* Corresponding author.
E-mail address: ymzhu@usst.edu.cn (Y. Zhu).

0030-4026/$ – see front matter © 2011 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ijleo.2011.08.005

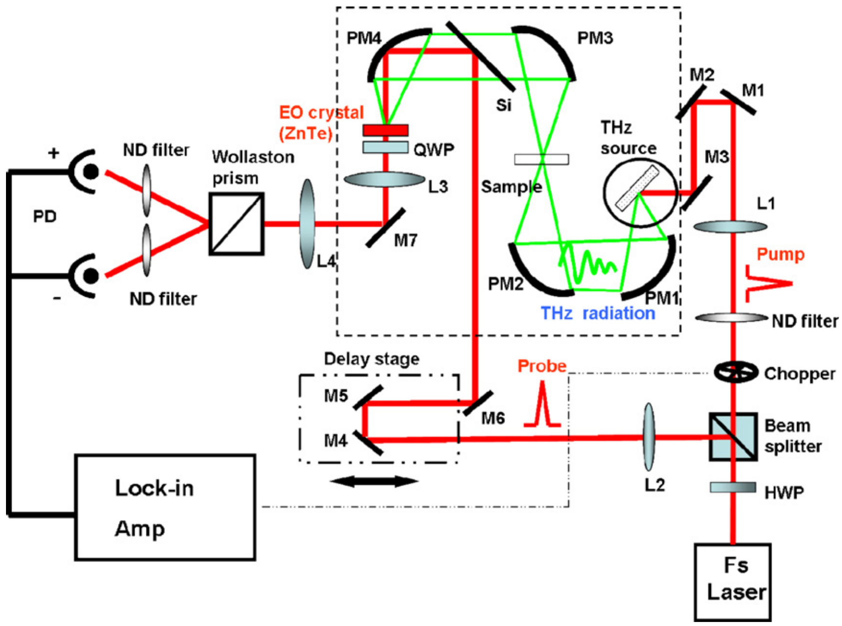

Fig. 1. THz-TDS system with reflecting mirrors (M_1–M_7), lenses (L_1–L_4), off-axis parabolic mirrors (PM_1–PM_4), beam splitter, high resistant silicon wafer (Si), electro-optic crystal (EO), half-wave plate (HWP), quarter-wave plate (QWP), wollaston prism.

sample by off-axis parabolic mirrors, PM1 and PM2. After passing the sample, the THz wave with the information of sample was focused on zinc telluride electro-optic crystal (EO crystal) by PM3 and PM4, together with probe beam. The free space EO sampling technique was used to record temporal waveforms of THz electric fields emitted from the samples [8,9]. The EO sensor used in this experiment was a 700 μm-thick ⟨1 1 0⟩-oriented ZnTe crystal. The high cutoff frequency (~3 THz) was determined by the thickness of the EO crystal and the pulse width of femtosecond laser [10,11].

To avoid the strong absorption of vapor, nitrogen gas was infused into a covered box as indicated as dashed line in Fig. 1. The humidity of covered box was kept as less than 8% and temperature was ranging from 23.5 °C to 24.5 °C.

In order to accurately control the thickness of samples, as well as the absorption coefficient, Radix Curcumae powders, which provided by Shanghai University of Traditional Chinese Medicine, were placed between two 2 mm-thick polyethylene slabs and the thickness of powders was about 0.5 mm. Fig. 2 shows the schematic diagram of the reference-polyethylene and sample-polyethylene box. $E_{ref}(\omega)$ and $E_{sam}(\omega)$ represent the THz signals from reference and sample box, respectively. Four groups of time-domain signals including the information of *huangyujin*, *lvyujin*, *guiyujin* and *wenyujin* were measured in the experiment.

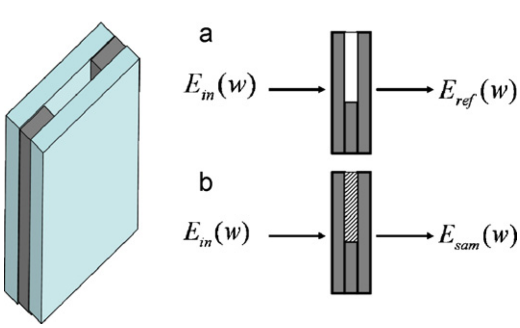

Fig. 2. Schematic diagram of the structure of sample box. The powder of Radix Curcumae is loaded in the slot; (a) reference polyethylene cell with the structure of polyethylene–air–polyethylene; (b) sample polyethylene cell with the same structure filled with the power of Radix Curcumae.

3. Results and discussion

Fig. 3 shows the THz waveforms of four kinds of Radix Curcumae in time domain. As a result of the refractive index between the sample and air, the time delay between the sample and the reference signal is about 0.8 ps. Furthermore, we also find that these Chinese herbs show similar terahertz transmission property.

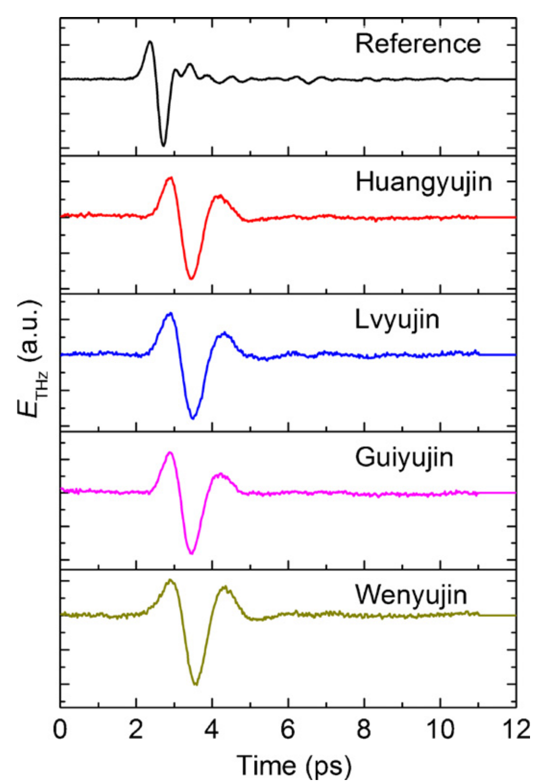

Fig. 3. Time-domain THz transmission waveforms of reference and four kinds of Radix Curcumae.

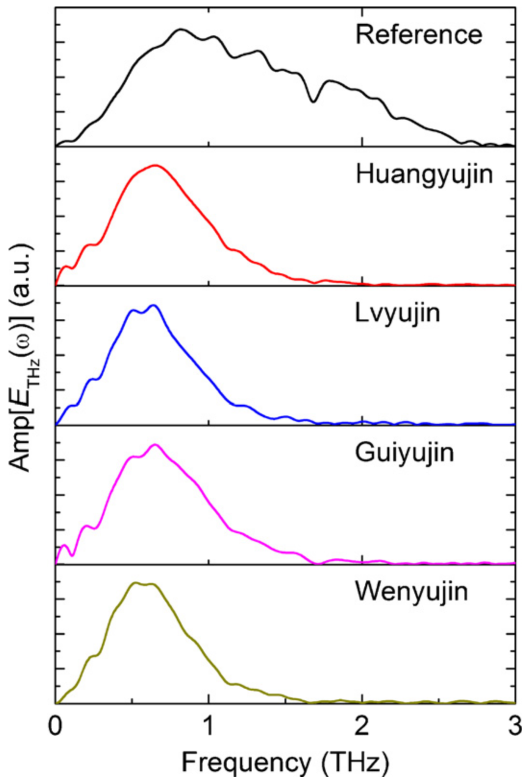

Fig. 4. THz transmission spectra of reference and four kinds of Radix Curcumae obtained from the Fourier transform of time domain THz transmission traces.

From the Fourier spectrum of the reference signal, as shown in Fig. 4, we can find the effective detected range is 0.2–2.8 THz and the signal-to-noise ratio of the experimental system is $\sim 10^3$. Furthermore, the transmission spectra of four kinds of Chinese herbs plotted in Fig. 4 show narrow bandwidth (0.2–1.6 THz), indicating that Radix Curcumae with about 0. 5 mm thickness are nearly opaque for THz wave whose frequency is above 1.6 THz.

Absorption spectra can be obtained from

$$\alpha = -[\ln(I_{\text{sample}}/I_{\text{reference}})]/\text{d}, \tag{1}$$

where d is the thickness of sample, I_{sample} is the intensity of THz pulses from samples and $I_{\text{reference}}$ is the intensity of reference THz pulse. The absorption spectra of four kinds of Radix Curcumae, plotted in Fig. 5, are smooth, which indicates that these Chinese herbs have no obvious absorption peaks in this frequency range. Furthermore, for *lvyujin* and *wenyujin*, the absorption peak around 1.4 THz was observed and, for comparison, this peak does not appear in the absorption spectra of *huangyujin* and *guiyujin*.

The main ingredients of *huangyujin*, *lvyujin*, *guiyujin* and *wenyujin*, which belong to Radix Curcume, are curcumin, demethoxycurcumin, bisdemethoxycurcumin, ar-turmerone, germacrone, and bisacurone B, etc. Because of the slightly different habitat and ingredient, the absorption coefficients show slight difference. To separate these same class of Chinese herbs, the measured absorption spectra from 0.2 to 1.6 THz, which is effective and sensitive, are used with the help of SVM method.

Support vector machine (SVM) is a kind of supervised learning methods for analyzing data and recognizing patterns, used for classification and regression analysis and destiny estimation [12]. In this work, a multi-class SVM method, which uses "one-against-one" approach [13,14], is employed. In this method, $k(k-1)/2$ classifiers are constructed, where k is the total classes of data. For training data

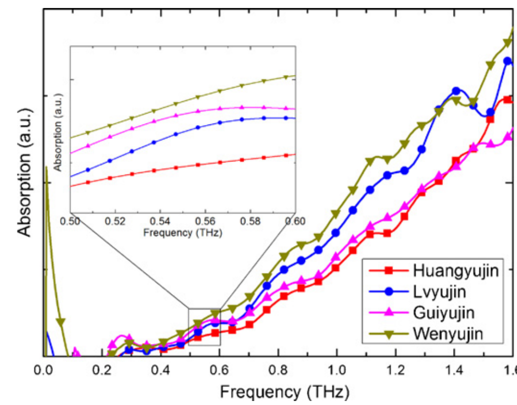

Fig. 5. Absorption spectra of *huangyujin*, *lvyujin*, *guiyujin*, *wenyujin* calculated by using Eq. (1).

from the ith and the jth classes, we solve the following two-class classification problem:

$$\min_{w^{ij},b^{ij},\xi^{ij}} \quad \frac{1}{2}(w^{ij})^T w^{ij} + C\left(\sum_t (\xi^{ij})_t\right) \tag{2}$$

$$\text{Subject to} \quad \begin{aligned} &(w^{ij})^T\phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, \quad \text{if } x_t \text{ in } i\text{th class} \\ &(w^{ij})^T\phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij}, \quad \text{if } x_t \text{ in } j\text{th class} \\ &\xi_t^{ij} \geq 0. \end{aligned}$$

where $\phi(x)$ is the sample set, w and b define the optimal hyperplane, C is the penalty factor, ξ is the slack variable. For the one-against-one approach, classification is done by a max-wins voting strategy, in which every classifier assigns the instance to one of the two classes, then the vote for the assigned class is increased by one vote, and finally the class with most votes determines the instance classification [13,14].

For identifying these Radix Curcumae, three kernel functions were tested: linear kernel function, polynomial kernel function, and radial basis kernel function. Three typical kernel functions are described as following:

1. Linear kernel function $k(x, y) = x \cdot y$ (3)
2. Polynomial kernel function $k(x, y) = [(x \cdot y)+]^d$ (4)
3. Radial basis kernel function $k(x, y) = \exp\{-|x-y|^2/\delta^2\}$ (5)

Penalty factor C and $g = \delta^2$, which are the parameters of three kernel functions, should be defined in suitable values, too small or too large values of g or C result wrong classification. We select the value of g and C ($g = 0.25$, $C = 1.0$) in the calculation.

In the process of calculation, the same parameters for all the kernel functions are used. We choose the absorption spectra of four kinds of Radix Curcumae from 0.5 THz to 0.6 THz with the highest signal-to-noise ratio, as the classified characteristic data as shown in the inset of Fig. 5, because the data in this region is effective and can reflect the nuances among these Chinese herbs.

In the experiment, each sample was measured four times, which means there are 16 sets of testing data for 4 kinds of Chinese herbs. Fig. 6 shows the distribution of the classified characteristic data (0.5–0.6 THz) from the absorption spectra of *huangyujin*, *lvyujin*, *guiyujin*, and *wenyujin*. The data from three groups (random choice) are chosen as the training data. The data from the remaining group were assumed as unknown materials and were classified. The accuracies of classification are listed on the Table 1. The correct rate for identification of these four samples is essentially 100%.

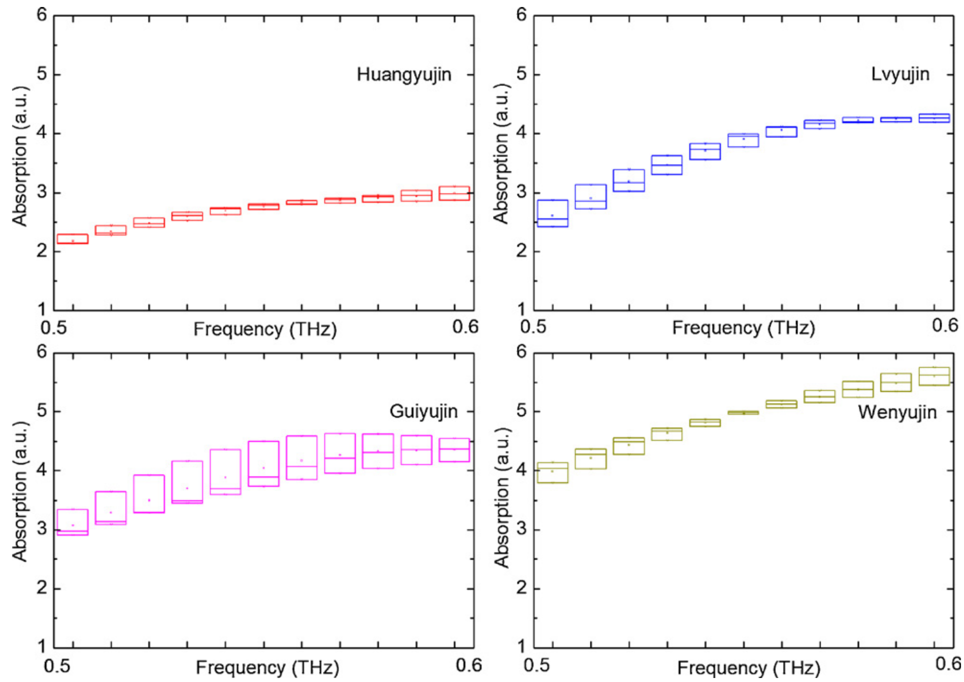

Fig. 6. Distribution of the absorption spectra (0.5–0.6 THz) of *huangyujin*, *lvyujin*, *guiyujin*, and *wenyujin*.

Table 1
Result of classification with three analyses.

Kernel type	Prediction accuracy
Linear	100%
Polynomial	100%
RBF	100%

4. Conclusion

THz-TDS was used to measure the absorption coefficients of *huangyujin*, *lvyujin*, *guiyujin* and *wenyujin* in THz frequency range. We found that they have no strong characteristic absorption peaks in this frequency region. Furthermore SVM is employed to distinguish four kinds of Radix Curcumae. The accuracy of identification basically achieves 100% in the 0.2–1.6 THz region. Therefore, we conclude 0.2–1.6 THz is an effective frequency band which can distinguish these four kinds of Radix Curcumae.

Acknowledgements

This work was partly supported by Dawn Project from Shanghai Municipal Education Commission and Shanghai Education Development Foundation (08SG48), Nano-technology Special Project of Science and Technology Commission of Shanghai (0952nm02400), Shanghai PuJiang Program from Science and Technology Commission of Shanghai (09PJ1407800), Innovation Program of Shanghai Municipal Education Commission (11YZ117), Ministry of Education Doctoral Fund of new teachers of China (20093120120007), Ministry of Education Scientific Research Foundation for Returned, and National Natural Science Foundation of China (61007059).

References

[1] H.B. Liu, P. George, E. Sarah, Y.Q. Chen, F. Bradley, X.-C. Zhang, Sensing minute changes in biological cell monolayers with THz differential time-domain spectroscopy, Biosens. Bioelectron. 22 (2007) 1075–1080.
[2] P.U. Jepsen, U. Moller, H. Merbold, Investigation of aqueous alcohol and sugar solutions with reflection terahertz time-domain spectroscopy, Opt. Express. 15 (2007) 14717–14737.
[3] J.S. Melinger, N. Laman, D. Grischkowsky, The underlying terahertz vibrational spectrum of explosives solids, Appl. Phys. Lett. 93 (2008) 011102.
[4] G.F. Liu, X.J. Ma, S.H. Ma, H.W. Zhao, M.W. Ma, M. Ge, W.F. Wen, Identification of explosives 2,4-DNT and 2,6-DNT using terahertz time-domain spectroscopy, Chin. J. Chem. 26 (2008) 1257–1261.
[5] S.H. Ma, G.F. Liu, P. Zhang, X.Y. Song, T. Ji, W.F. Wen, The THz fingerprint spectra of the active ingredients of a TCM – Herba Ephedrae, Proc. SPIE 7277 727711 (2009) 1–10.
[6] Y.M. Zhu, T. Unuma, K. Shibata, K. Hirakawa, Power dissipation spectra and terahertz intervalley transfer gain in bulk GaAs under high electric fields, Appl. Phys. Lett. 93 (2008) 232102–232104.
[7] Y.M. Zhu, T. Unuma, K. Shibata, K. Hirakawa, Femtosecond acceleration of electrons under high electric fields in bulk GaAs investigated by time-domain terahertz spectroscopy, Appl. Phys. Lett. 93 (2008) 042116–042118.
[8] P.C.M. Planken, H.K. Nienhuys, H.J. Bakker, T. Wenckebach, Measurement and calculation of the orientation dependence of terahertz pulse detection in ZnTe, J. Opt. Soc. Am. B 18 (2001) 313–317.
[9] Q. Wu, X.-C. Zhang, Free-space electro-optics sampling of mid-infrared pulses, Appl. Phys. Lett. 71 (1997) 1285–1286.
[10] Q. Wu, X.-C. Zhang, 7 Terahertz broadband GaP electro-optic sensor, Appl. Phys. Lett. 70 (1997) 1784–1786.
[11] A. Leitenstorfer, S. Hunsche, J. Shah, M.C. Nuss, W.H. Knox, Detectors and sources for ultrabroadband electro-optic sampling: experiment and theory, Appl. Phys. Lett. 74 (1999) 1516–1518.
[12] X.-K. Ruan, Z.-Y.\ Zhang, Signal Blind Recovery in Communication Systems Using Support Vector Machines, Journal of Nanjing University of Posts and Telecommunications (Natural Science) 30 (2010) 1–5.
[13] S. Knerr, L. Personnaz, G. Dreyfus, Single-layer learning revisited: a stepwise procedure for building and training a neural network, in: J. Fogelman (Ed.), Neurocomputing: Algorithms, Architectures and Applications, Springer-Verlag, 1990, pp. 41–50.
[14] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge University Press, 2000, pp. 26–51.