Aflatoxin B1 Detected by Terahertz Time-domain Spectroscopy

LIAN Fei-yu[1, a], GE Hong-yi[2, b]

[1].College of Information Science and Engineering, Henan University of Technology. Zhengzhou, 450001, China

[2].State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, 100080, Beijing, China

[a]email: lfywork@163.com, [b]email:gehongyi2014@163.com

Keywords: Aflatoxin B1;Terahertz; Spectrum Analysis; Support Vector Machine

Abstract. Developing physical methods to detect aflatoxin is a significant area of research in food safety. Because the current detection methods that use near-infrared spectroscopy are not sufficiently accurate, we focus here on using Terahertz Time-Domain Spectroscopy (THz-TDS) to detect aflatoxin B1 in food. We first divided the concentration of aflatoxin B1 preserved in acetonitrile into ten classifications which are 5ppb, 10ppb,25ppb,50ppb,100ppb,500ppb,1ppm and 10ppm respectively. Each classification corresponds to a certain THz spectroscopy. For the samples with unknown concentration, we respectively used Least-Squares Support Vector Machine (LS-SVM), Naïve Bayes and Back-Propagation Artificial Neural Network(BP-ANN) to identify their concentrations based on THz spectroscopy. Verification results from High Performance Liquid Chromatography (HPLC) show that, compared to the other two methods, LS-SVM has more precision for identifying the concentration of alatoxin B1. The proposed method may provide another effective way for rapid detection of aflatoxin in food.

Introduction

Because aflatoxin poses a serious threat to public health, it has become mandatory to measure the aflatoxin content in foodstuffs in international trade . Spurred by this necessity, methods to detect aflatoxins have quickly multiplied in recent years. At present, the main methods include thin-layer chromatography (TLC), high-performance liquid chromatography (HPLC), enzyme-linked immunosorbent assay (ELISA), and immune affinity column fluorescence spectrophotometry [1]. Although these methods have been mature technologies, and have already been applied commercially, and even some of them have already become standard detection methods, they still have many limitations. For example, methods for sample preparation are complicated, and the test results depend on the experience and technological knowledge of inspectors. At present, the main requirement for aflatoxin detection is to test quickly, precisely, and early. Unfortunately, existing detection methods are still insufficiently sensitive for early detection, so the best opportunities for preventing mildew are frequently missed. Therefore, it is extraordinarily important to find a fast, precise, and destruction-free method to detect aflatoxin.

Terahertz spectroscopy is a new technology that has developed rapidly since the 1980s. It is widely applied in areas such as security checks [2], material science [3], military communications [4], biomedicine [5], and agricultural inspection products [6]. For testing aflatoxin of grain by using THz-TDS, there is still little research. In this paper, we reported a method we developed to detect aflatoxin B1 by using THz-TDS technique and realized a multi-class classification of concentration of aflatoxin B1 by using terahertz spectrum. We respectively used Least-Squares Support Vector Machine(LS-SVM), Naïve Bayes and Back-Propagation Artificial Neural Network(BP-ANN) to identify concentration of aflatoxin B1 based on THz spectroscopy. We focused on the analysis and comparison of experimental results of the three classification methods for classifying the concentration of aflatoxin B1 using THz spectrum.

The Fundamental of Terahertz Detection

THz time-domain spectroscopy system

A schematic of the experimental apparatus for terahertz spectroscopy [12] is shown in Fig. 1. A self-mode-locked Ti-sapphire femtosecond pulsed laser generates laser pulses with a central wavelength of 800 nm, a pulse width of 100 fs, a repetition frequency of 80 MHz, and an energy of 9.0 nJ per pulse. A beam splitter divides the laser beam into a stronger pump beam and a weaker probe beam.

The pump beam is modulated by a chopper at 1.1 kHz and focused by an optical lens onto a biased microstripline antenna on a GaAs crystal. Upon illumination by a laser pulse, the rapid change in conductivity in the GaAs creates electrical transient currents that, through optical rectification, generate electromagnetic pulses in the frequency range of 0.2 to 3 THz. Because polar materials and, in particular, water strongly absorb THz radiation, humidity must be minimized in the THz beam path. For this reason, the THz beam path from the sample to the detection crystal is enclosed in a nitrogen-purged box that has a relative humidity of less than 2% and a temperature of around 293 K. The signal-to-noise ratio (SNR) of the system can be as large as 1000 and the spectral resolution is greater than 40 GHz.
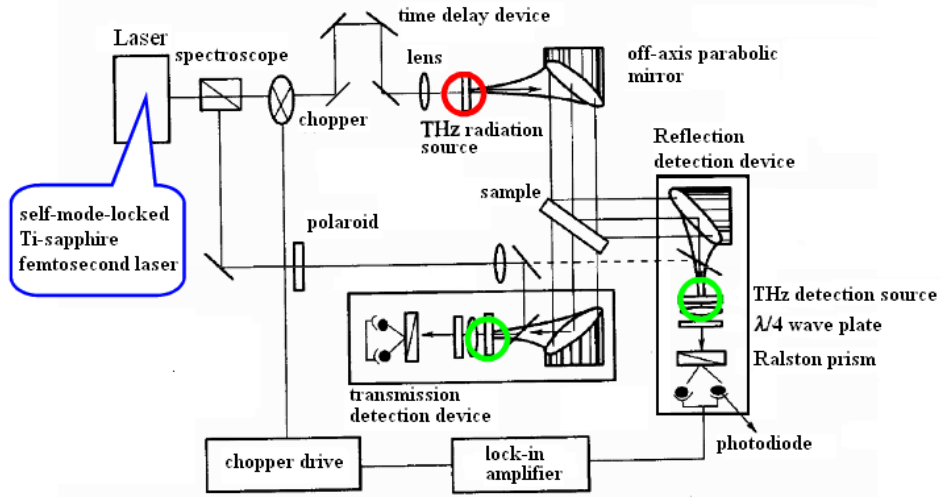

Fig.1.   Terahertz time-domain spectroscopy system.

Data Acquisition

The macro optical properties of sample can be described by a complex index of refraction $\tilde{n}$:

$$\breve{n}(\omega) = n(\omega) - ik(\omega) \tag{1}$$

Where $n(\omega)$ is the real index of refraction that reflects the dispersion characteristics of the sample, and $k(\omega)$ is the extinction coefficient that describes the absorption characteristics.

From the Fresnel equation, the reflection and transmission coefficients between media 1 and 2 can be expressed as

$$R_{12}(\omega) = \frac{\tilde{n}_1 - \tilde{n}_2}{\tilde{n}_1 + \tilde{n}_2} \tag{2}$$

$$T_{12}(\omega) = \frac{n_1 \%}{n_1 \% + n_2 \%} \tag{3}$$

Media 1 is air, media 2 is the cuvette made of SiO2, and media 3 is the solution including B1 in the cuvette. According to the Fabry-Perot (FP) effect, we get

$$FP_{2,\delta}^{13}(\omega) = \overset{\delta}{\underset{k=0}{A}} \left\{ R_{23}(\omega) P_2^2(\omega, d) R_{21}(\omega) \right\}^k \tag{4}$$

where δ is the phase difference between the two adjacent optical transmissions, and $P_2(\omega, d) = \exp(-i\tilde{n}_2 \omega d / c)$ is the propagation factor in media 2, where d is the thickness of media 2, and c is the speed of light in vacuum.

Based on the FP effect, the electric field of the THz reference signal is

$$E_{ref}(\omega) = \eta_0(\omega)E_0(\omega)T_{10}(\omega)P_0(\omega)FP_0^{13}(\omega)T_{03}(\omega) \tag{5}$$

and the electric field of the THz signal from the sample is

$$E_{sample}(\omega) = \eta_2(\omega)E_0(\omega)T_{12}(\omega)P_2(\omega)FP_2^{13}(\omega)T_{23}(\omega) \tag{6}$$

where $\eta_x(\omega) = T_{01}(\omega)P_1(\omega)FP_1^{0x}(\omega)P_3(\omega)FP_3^{x0}(\omega)T_{30}(\omega)$ is a multiple-reflection factor describing reflections between medium 1 and medium 2.

Based on Eqs (x)–(y), the transmission signal is the ratio of the reference signal to sample signal:

$$H_{thoery}(\omega) = \frac{E_{sample}^{theor}(\omega)}{E_{ref}^{theor}(\omega)} = \frac{4n_{B1}\%n_{SiO_2}\%}{(n_{B1}\% + n_{SiO_2}\%)^2} \exp[-i(n_{B1}\% - n_{air}\%)\omega d / c] \tag{7}$$

The index of refraction and absorption coefficient can be obtained by replacing the theoretical values by the observed values:

$$n_{sample}(\omega) = \frac{\arg[H_{Measure}(\omega)]c}{\omega d} + n_{air} \tag{8}$$

$$a_{sample}(\omega) = \frac{2k(\omega)}{c} = \frac{2}{L} \ln \frac{4n_{sample}(\omega)}{|H_{Measure}(\omega)|(n_{sample}(\omega) + 1)^2} \tag{9}$$

The Spectrum Classification of Standard Concentration

To determining the content of aflatoxin B1 in samples, we first used acetonitrile to dilute the standards substance of afatoxin B1 with concentration of 100µg/ml, and form different concentrations in certain Gradient using pipette with different range. The concentration classifications we made are 10ppm, 1ppm, 500ppb, 100ppb, 50ppb, 25ppb, 10ppb and 5ppb, where 1ppm=1000ppb=1µg/ml. We detected these concentrations using THz-TDS and got their THz spectrums to setup standard spectroscopy classification.

(a) Time-domain spectroscopy   (b) Frequency spectroscopy

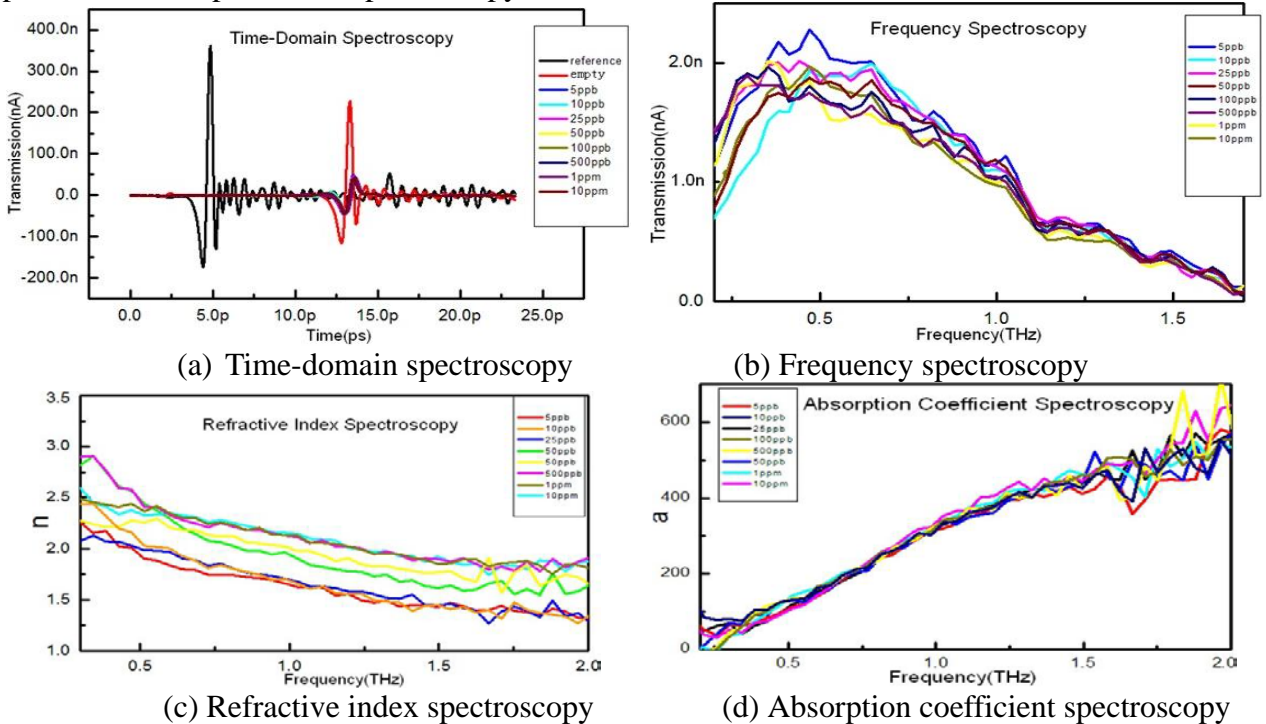(c) Refractive index spectroscopy   (d) Absorption coefficient spectroscopy

Fig.2. The terahertz spectrum for standard concentrate of Aflatoxin B1

The standard THz-TDS time-domain spectrums are shown in Fig.2(a), and by performing a fast Fourier transform (FFT) for the spectrums in Fig.2(a), we can get standard THz frequency spectrums for these different concentrate classifications of aflatoxin B1, as shown in Fig.2(b). After further calculation, we got standard refractive index spectrums and absorption spectrums of these different concentrate classification of afatoxin B1, as shown in Fig.2(c) and Fig.2(d).

To realizing application of the method in this paper, it's necessary and helpful to setup more classifications with greater range and smaller gradient. To keep things simple, the classifications we made are only to illustrate the validity of our method.

Sample Preparation

Aflatoxin B1 in corn usually occurs as solid, and it can be extracted by immunoaffinity column and preserved in nonpolar organic solvents such as acetonitrile, chloroform, and so on. In our experiment, we used aflatoxin B1 dissolved in acetonitrile as test materials. Because the concentration of aflatoxin preserved in solvent is relatively low and the cuvettes used are relatively thick, and the cuvettes can absorb a lot of THz energy, which causes problems with sensitivity, and even lead to a lack of response to THz radiation, we used chemical glassware called removable liquid cell in our experiment.

Classification Methods

Least Squares Support Vector Machines

Least Squares Support Vector Machines (LS-SVM) is presented by Suyken s J. A. K [8]. The main difference the SVM and the LS-SVM is as following, 1) The SVM solution is typically found by solving a quadratic programming problem where as the LS-SVM solution can be found by solving a set of linear equations.2) The SVM has inequality constraints where as the LS-SVM has only equality constraints. 3) LS-SVM reduces the computational cost and easy to be implemented for on-line application [9].

Given a training data set of K points $\{x_k, y_k\}_{k=1}^{N}$, with input data $x_k \in R^n$ and output data $y_k \in R^n$, LS-SVM aims at to build a classifier of the form as follow

$$y(x) = sign[\sum_{k=1}^{N} a_k y_k \psi(x, x_k) + b] \tag{10}$$

Where a_k are positive real constants, b is a real constant and $\psi(x, x_k)$ is a Kernel function. In LS-SVM for classification the following optimization problem is formulated

$$\min_{w,b,e} J_3(w,b,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{N} e_k^2 \tag{11}$$

Subject to constraint as follow

$$y_k[w^T \varphi(x_k) + b] = 1 - e_k, \qquad k = 1, \cdots, N \tag{12}$$

To solve this quadratic programming, Lagrange multiplier is used as follow

$$L_3(w,b,e;a) = J_3(w,b,e) - \sum_{k=1}^{N} a_k \left\{ y_k[w^T \varphi(x_k) + b] - 1 + e_k \right\} \tag{13}$$

Where a_k are Lagrange multipliers, the conditions for optimality are given as follow

$$\begin{cases} \dfrac{\partial \mathbf{L_3}}{\partial \mathbf{w}} = \mathbf{0} \rightarrow \mathbf{w} = \sum_{k=1}^{N} \mathbf{a_k y_k} \varphi(\mathbf{x_k}) \\ \dfrac{\partial \mathbf{L_3}}{\partial \mathbf{b}} = \mathbf{0} \rightarrow \sum_{k=1}^{N} \mathbf{a_k y_k} = \mathbf{0} \\ \dfrac{\partial \mathbf{L_3}}{\partial \mathbf{e_k}} = \mathbf{0} \rightarrow \mathbf{a_k} = \gamma \mathbf{e_k} \qquad\qquad \mathbf{k = 1, \cdots, N} \\ \dfrac{\partial \mathbf{L_3}}{\partial \mathbf{a_k}} = \mathbf{0} \rightarrow \mathbf{y_k}[\mathbf{w^T} \varphi(\mathbf{x_k}) + \mathbf{b}] - \mathbf{1} + \mathbf{e_k} = \mathbf{0}, \quad \mathbf{k = 1, \cdots, N} \end{cases} \tag{14}$$

Equ.(14) can be written as the linear equations

$$\begin{bmatrix} I & 0 & 0 & -Z^T \\ 0 & 0 & 0 & -Y^T \\ 0 & 0 & \gamma I & -I \\ Z & Y & I & 0 \end{bmatrix}\begin{bmatrix} w \\ b \\ e \\ a \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vec{1} \end{bmatrix} \tag{15}$$

Where $Z=\left[\varphi(x_1)^T y_1;\cdots;\varphi(x_N)^T y_N\right]$, $Y=\left[y_1;\cdots;y_N\right]$, $\vec{1}=[1;\cdots;1]$, $e=[e_1;\cdots;e_N], a=\left[a_1;\cdots a_N\right]$

With Mercer's condition

$$\Omega_{kl}=y_k y_l \varphi(x_k)^T \varphi(x_l)=y_k y_l \psi(x_k,x_l), k=1,\cdots N \tag{16}$$

Then select the equations related with a and b, equations(15) can be transformed into

$$\begin{bmatrix} 0 & -Y^T \\ \gamma & \Omega+\gamma^{-1}I \end{bmatrix}\begin{bmatrix} b \\ a \end{bmatrix}=\begin{bmatrix} 0 \\ \vec{1} \end{bmatrix} \tag{17}$$

Least squares method can be applied to calculate a and b from linear equations (17).

The Radial Basis Function(RBF) kernel $\psi(x_i,x_k)=\exp\left\{-\left\|x_i-x_k\right\|^2/\delta\right\}$ was employed in the experiment, and the values of the kernel parameter δ and of the regularization parameter γ were determined by a grid search and cross validation to select those values that give the smallest error on the test data set.

Naive Bayes Classifier

The Naive Bayes Classifier technique is based on the socalled Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. Naive Bayes classifier has gained popularity in solving various classification problems including microarray data analysis [10,11].

Naive Bayes classifier assuming strong independence within attributes of an instance.

$$p(C\mid A_1,\cdots,A_n)=\frac{1}{Z}p(C)\prod_{i=1}^{n}p(A_i\mid C) \tag{18}$$

Where Z is a scaling factor depends only on A_1 through A_n, C is the class variable, A_1 through A_n are independent attributes variables, and $p(C)$ is called class prior probability.

BP Artificial Neural Networks

Back Propagation Artificial Neural Network is one of the most widely used neural networks. It is a single-direction multilayer neural network that contained of input layer nodes, output layer nodes and one or more layers of hidden nodes. The information transfer from nodes to nodes of different layers, and the degrees of the connections are controlled by the connection weights. The connection weights are adjusted on the basis of data by training. BP Artificial Neural Network input information from the input layer and output the decisions to the output layer. Actually, the BP Artificial Neural Network is a kind of highly nonlinear mapping from input to output[12].

In order to achieve multi-class classification of concentration for aflatoxin B1, two different four-layer neural networks was constructed in this paper, which with 9 hidden layer neurons and 12 hidden layer neurons respectively. Trainlm was selected as the BP Artificial Neural Network training function, purelin was selected as the output layer as transfer function and tansig was selected as the hidden layer transfer function. The training goal was set to 0.001 and the maximum number of training was set to 10,000.

Experimental Results and Analysis

For this study, we measured mainly three types of corn samples, which had varying degrees of mildew due to aflatoxin B1. Based on the external feature of the samples, we classified them three

types which titled mild mildew, medium mildew, and serious mildew. Aflatoxin B1 in all samples was extracted using immunoaffinity column and preserved in acetonitrile.

The typical THz time-domain transmission spectrums of these three types concentration samples are shown in Fig.3(a). By Fourier transforming, we obtained their frequency-domain spectrums, which are shown in Fig.3(b). By using Eqs.8 and Eqs.9, we obtained their refraction and absorption coefficient spectrums, which are shown in Figs.3(c) and Fig.3(d) respectively.

(b) Time-domain spectroscopy   (b) Frequency spectroscopy

(c) Refractive index spectroscopy   (d) Absorption coefficient spectroscopy
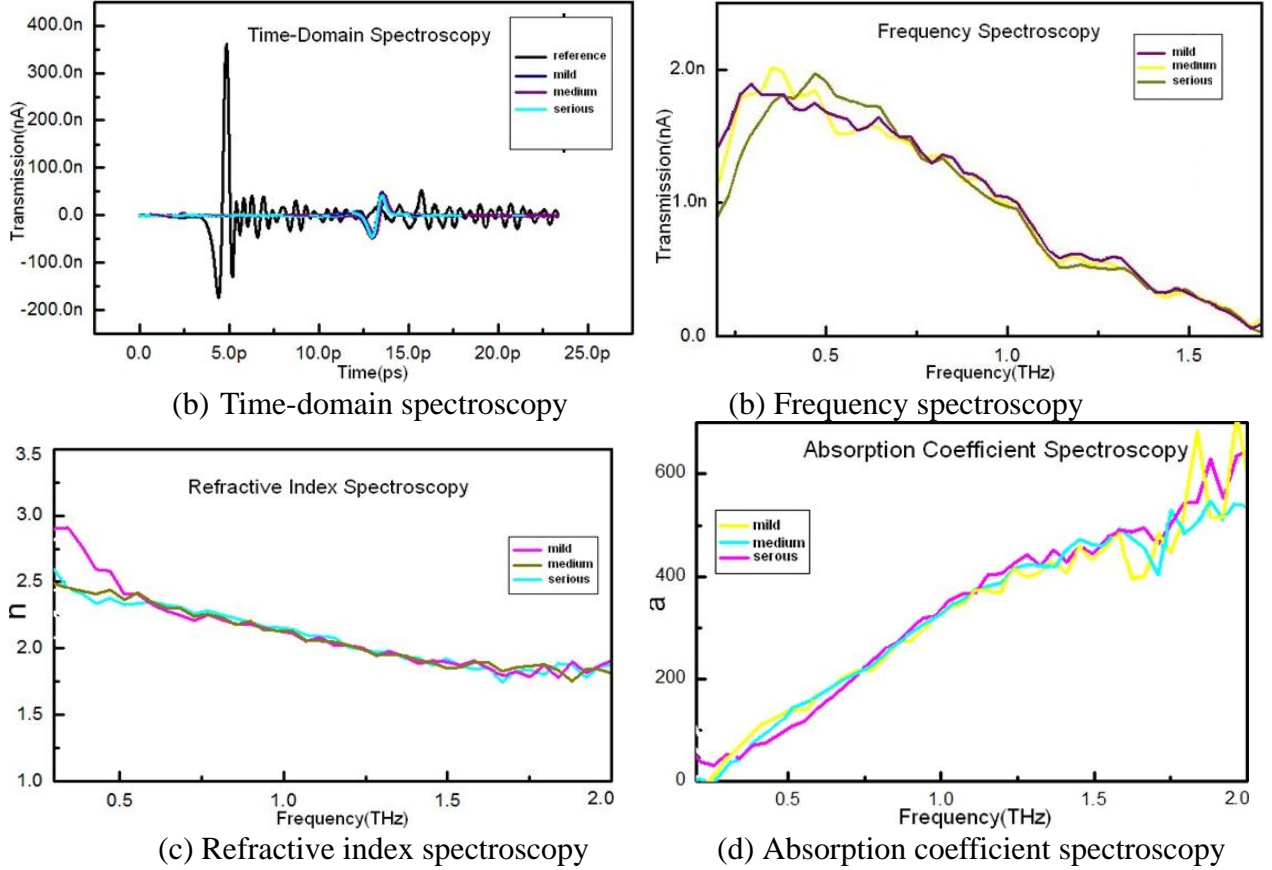
Fig.3 Terahertz spectrum of three samples.

It is clear that these three kinds of concentration samples are not significantly different in absorption and refractive THz spectrums, and there is difficult to be directly distinguished. Therefore, the refractive index spectrums and absorption spectrums from 0.1 to 2.0 THz were treated as the feature data, and Multi-class classification was used in this paper in order to distinguish them one by one. Refractive index spectrums and absorption spectrums of these three kinds of concentration samples with 100 samples were all used as the training and testing data by cross validation.

Genetic Algorithms was applied in getting the regularization parameter γ and the RBF kernel parameter δ of LS-SVM. Compared to grid search and cross validation, the speed of obtaining optimal parameters has been greatly improved by GA-LS-SVM. The regularization parameter γ was set to be 400 and the RBF kernel parameter δ was set to be 60 in LS-SVM classification Algorithm in this experiment after optimization.

Gaussian distribution and kernel density estimation was used to model the feature in Naive Bayes classification Algorithm respectively. 9 hidden layer neurons and 12 hidden layer neurons was applied in two different four-layer neural networks to test the classification results. All three classification algorithm used the same training samples data and testing samples data, which shown as TABLE I. To test the stability of these algorithms applied in concentration samples classification, 5% and 10% noise was added respectively, which showed as TABLE II and TABLE III.

TABLE I.   Classification results of concentration of aflatoxin B1

Algorithm or Method	Recognition Rate					
	<500ppb	500ppb	1ppm	10ppm	>10ppm	Parameters
Naïve Bayes	0%	85%	86%	88%	0%	Gaussian distribution
Naïve Bayes	0%	85%	88%	86%	0%	Kernel density estimation
LS-SVM	0%	93%	95%	96%	0%	RBF Kernel
BP ANN	0%	83%	86%	86%	0%	12 hidden layer neural
BP ANN	0%	80%	83%	85%	0%	9 hidden layer neural
HPLC	0%	98%	100%	100%	0%	

TABLE II   Classification results after adding 5% noise

Algorithm or Method	Recognition Rate					
	<500ppb	500ppb	1ppm	10ppm	>10ppm	Parameters
Naïve Bayes	0%	76%	80%	80%	0%	Gaussian distribution
Naïve Bayes	0%	78%	78%	82%	0%	Kernel density estimation
LS-SVM	0%	85%	88%	90%	0%	RBF Kernel
BP ANN	0%	78%	80%	80%	0%	12 hidden layer neural
BP ANN	0%	75%	78%	78%	0%	9 hidden layer neural

TABLE III   Classification result after adding 10% noise

Algorithm or Method	Recognition Rate					
	<500ppb	500ppb	1ppm	10ppm	>10ppm	Parameters
Naïve Bayes	0%	68%	70%	70%	0%	Gaussian distribution
Naïve Bayes	0%	65%	60%	68%	0%	Kernel density estimation
LS-SVM	0%	75%	80%	80%	0%	RBF Kernel
BP ANN	0%	60%	65%	65%	0%	12 hidden layer neural
BP ANN	0%	55%	63%	65%	0%	9 hidden layer neural

The experiment results show that, without the application of complex feature extraction Algorithm, concentration samples can be effectively classified. Dealing with the data of noise pollution shows that support vector machines have very strong generalization ability. The identification result of BP Artificial neural network is affected by training sample size, training errors, and the Structural complexity of network. For the number of hidden number of BP Artificial neural network, too small networks are unable to adequately learn the problem while excessively large networks tend to over fit the training data and consequently result in weak performance. The LS-SVM is an algorithm based on the statistical learning theory for small samples, which transforms the problem of searching for the optimal hyper plane and it is well suited for solving data as type of this experimental. The classification result of this experiment indict that the LS-SVM outperforms the Naive Bayes and BP Artificial neural network significantly, when the training samples are limited.

Conclusion

In this paper, we applied THz-TDS on corn samples containing aflatoxin and obtained their frequency-domain spectrums, index of refraction spectrums and absorption coefficient spectrum, then we used Least squares support vector machines, Naïve Bayes and Back propagation artificial Neural Network to realize multi-class classification of concentration of aflatoxin B1. Meanwhile the stability of the three algorithms used in concentration classification was tested by adding 5% and 10% noise respectively. The results indicated that, LS-SVM in the trial have the best classification results. The experiment showed that based on THz spectroscopy, the varieties of concentration of aflatoxin B1 can be identified effectively, which provide a new way for concentration identification.

Acknowledgments

References

[1] MA Zhi-ke, ZAN Lin-sen. Progress on Hazard, Detection Techniques and Biodegradation of Aflatoxin. Progress in Veterinary Medicine, 2009, 30(9):91-94.

[2] J. S. Melinger, N. Laman, and D. Grischkowsky, "The underlying terahertz vibrational spectrum of explosives solids," Appl. Phys. Lett., vol. 93, no. 1, Jul 7, 2008.

[3] B. Ferguson, and X. C. Zhang, "Materials for terahertz science and technology," Nat. Mater., vol. 1, no. 1, pp. 26-33, Sep, 2002.

[4] J. Federici, and L. Moeller, "Review of terahertz and subterahertz wireless communications," J Appl Phys,vol. 107, no. 11, Jun 1, 2010.

[5] P. H. Siegel, "Terahertz technology in biology and medicine," IEEE Trans.Microw. Theory Tech., vol. 52, no. 10, pp. 2438–2447, Oct. 2004.

[6] Y. F. Hua, H. J. Zhang, and H. L. Zhou, "Quantitative Determination of Cyfluthrin in N-Hexane by Terahertz Time-Domain Spectroscopy With Chemometrics Methods," IEEE Trans. Instrum. Meas., vol. 59, no. 5, pp. 1414-1423, May, 2010.

[7] Y. F. Hua, H. J. Zhang, and H. L. Zhou, "Quantitative Determination of Cyfluthrin in N-Hexane by Terahertz Time-Domain Spectroscopy With Chemometrics Methods," IEEETransactions on Instrumentation and Measurement, vol. 59, no. 5, pp. 1414-1423, May, 2010.

[8] J.A.K.Suykens, and J.Vandewalle, "Multiclass least squares support vector machines," International Joint Conference on Neural Networks, vol. 2, pp. 900-903, 1999.

[9] Anthony Kuh, "Analysis of detectors for support vector machines and least square support vector machines," International Joint Conference on Neural Networks, vol. 2, pp. 1075-1079, 2002.

[10] Eunseog Youn, Myong K. Jeong , "Class dependent feature scaling method using naive Bayes classifier for text datamining," Pattern Recognition Letters, vol. 30, pp. 477-485, April 2009.

[11] Liwei Fan, Kim-Leng Poh, Peng Zhou, "A sequential feature extraction approach for naive bayes classification of microarray data," Expert Systems with Applications, vol. 36, pp. 9919-9923, August 2009.

[12] Qin Ju, Zhongbo Yu, Zhenchun Hao, Gengxin Ou, Jian Zhao, Dedong Liu, "Division-based rainfall-runoff simulations with BP neural networks and Xinanjiang model," Neurocomputing, vol. 72, pp. 2873-2883, August 2009.