Study on automatic identification method of X-ray fluorescence spectrum

Jiening Xia[1, 2*], Zhigao Chen[1, 2]

[1]*Hubei Key Laboratory of Earthquake Early Warning, Institute of Seismology, China*

[2]*Wuhan Base of Institute of Crustal Dynamics, CEA, Wuhan, China*

Received 1 July 2014, www.cmnt.lv

Abstract

The way to improve the accuracy and reliability of automatic unscrambling and identification technology on X-ray fluorescence spectrometer spectrum is studied in this essay. Accordingly, two different automatic identification methods based on Fast Fourier Transform and Wavelet Transform are presented. By the tool LabVIEW, such two methods are applied to the qualitative analysis on X-ray fluorescence spectrums, and the features of such two methods are compared. Based on the experiments and analysis on amount of samples, it can be concluded that the automatic identification method based on the Wavelet transform theory is better than the other method for the former has a better local resolution. Therefore, the characteristic values of the singular points are more clearly recognized by the method based on the Wavelet transform. Through the study in this essay, theories on automatic identification are enriched, which set a foundation for further studied in future.

Keywords: X-ray fluorescence spectrometer, spectrum identification, LabVIEW, fast Fourier transform, wavelet transform

1 Introduction

X-ray fluorescence spectroscopy technology as a widely used ray-detection technology, began in the early 1950s. Since the spectrum is less interfered, accurate, precise, fast, can detect multiple elements, has a wide range of measurement, and performs non-destructive testing, it is applied to various fields, which becomes an indispensable and means in the laboratory and no alternatives. And now it becomes an indispensable method for analysis of substance composition. X-ray fluorescence spectroscopy detection technology has overwhelming advantages such as fast analysis speed, wide range of elements, which can be tested, simple pretreatment, free of pollution, low cost and non-destructive testing, etc. Nowadays, there are more and more types of X fluorescence spectrometers. Those X fluorescence spectrometers can be divided into two types according to the method of acquiring and distinguishing the characteristics of the X-ray fluorescence spectroscopy: wavelength dispersive X-ray fluorescence spectrometers (WDXRF) and energy dispersive X-ray fluorescence spectrometers (EDXRF) [1].

An X-ray fluorescence spectrometer is generally made up of three parts in structure: a light source for exciting the sample (X-ray tube), a detection system (for dispersion, detection, and controlling the spectrometer), and a data processing system [2]. Incident X-rays are generated from the X-ray tube for exciting the sample. Accordingly, secondary X-rays can be emitted from the to-be-tested elements of the sample by the excitation, in which different secondary X-rays emitted from excitation of different elements each owns specific energy characteristics and wavelength characteristics. The energy of such secondary X-rays are collected and measured by the detection system, and through a data processing software in the detection system, the collected fluorescence information of such secondary X-rays are calculated into types and contents information of the to-be-tested elements of the sample. Spectrum characteristics of a certain element is not only related to the energy and intensity of the excitation source, but also related to the level of such element in the sample. According to the spectrum characterristics of such element, the content information thereof can be obtained by converting.

There are mainly three types of analysing methods of the X-ray fluorescence spectroscopy detection technology: qualitative analysing method, semi-quantitative analysing method and quantitative analysing method [3]. The qualitative analysing method is discussed in this essay, which indicates the aforementioned solution and identification methods on the spectrum. The Moseley's law tells that the wavelength of the secondary X-ray emitted from the to-be-tested element is one to one related to the atomic number. In qualitative analysis, characteristic peaks and common interference lines of the scanned spectrum figure should be identified, thereby to determine the types of elements contained in the sample, and thus to determine the types of materials of the sample.

Currently, the automatic recognition algorithm of the X-ray fluorescence spectrometer is being studied by various spectrometer manufacturers. According to the algorithm, by determining the peak of the scanned spectrum, calculating the net intensity of the background and peak, as well as pairing with the characteristic spectrum line database, the types of the to-be-tested elements and the types of the spectrum lines are determined. However, the actual working situation is complex and changing. For example, when the

**Corresponding author* e-mail: dr.xjning@gmail.com

content of an element is too little or there exists interference between of spectral lines of different elements, the artificial identification by experiences of the operator is essential. First, the characteristic X-ray lines of the target X-ray tube and the associated lines of the peak are identified. Afterwards, the remaining spectrum lines are marked according to energy. When unknown spectrum lines are analysed, other factors such as the origin and nature of the sample should also be taken into consideration, so as to get comprehensive judgments.

In order to improve the accuracy and reliability of the automatic recognition and reconciliation of spectrums, this essay presents a method for feature extraction and recognition based on fast Fourier transform, and a method for spectrum solution based on wavelet. And such two methods are proved and compared by specific experiments with LabVIEW.

2 RELATED WORK AND BACKGROUND

Spectrum recognition models either domestic or international are essentially divided into two types: a spectrum matching method based on spectrum reconstruction and standard spectrum similarity measurement, and an intelligent recognition method based on sample knowledge. The spectrum matching method is to compare the reconstructed spectrum with the reference spectrum, and to measure the similarity or correlation between them by a certain measurement function, so as to identify the sample. The similarity measurement function can be a distance function (Euclidean distance, Mahalanobis distance), similarity coefficient, correlation coefficient, spectral information divergence or spectrum vector angle, etc.; and the reference spectrum can generally be a standard spectrum in the spectrum database.

According to the data used during the matching process, the spectrum matching method can be divided into two types: direct matching and indirect matching. The direct matching method uses the reconstructed spectrum to match the reference spectrum data directly; whereas the indirect matching method encodes or transforms the spectrum data first, and uses the encoded and transformed data to match, such as encoded spectrum matching, amplitude and phase of the Fourier frequency of the spectrum, wavelet or fractal dimension matching, other than using the raw spectrum data to match directly.

When the X-ray fluorescence spectrum analysis method is carried out, due to the interaction of X-rays substances, such as coherent scattering, incoherent scattering, Compton scattering and other reasons, the Almighty certain peaks of the characteristic X-ray may superimpose on the background. In the qualitative and quantitative analysis, it is very necessary to subtract the background accurately in order to obtain the peaks of the spectrum and the net intensity of peak areas. It is a very critical and difficult technology to subtract the background effectively and accurately. Therefore, various digital filtering and deconvolution algorithms are powerful tools for the spectrum solution.

In recent years, more and more advanced intelligent methods have been used in such field. The intelligent identification method is on the basis of physical spectrum knowledge. According to this method, appropriate diagnostic spectrum features or spectrum parameters with identification ability are selected first, then identification rules are established, and samples are identified accordingly. Characteristics or parameters of the spectrum can be the similarity measure for the whole spectrum, such as the matching degree or matching rate compared to the standard spectrum, or the parameters of specific bands, or both. The identification rules can be generally expressed in 3 ways: uniqueness identification (existence identification), i.e., when some certain conditions are met, it can be considered that some certain element exists; negative identification, i.e., when one or more features occur, it can be considered that some element must not exists; likelyhood identification, i.e., the likelihood whether the material to be identified is determined according to the conditions.

The intelligent identification method combines the full spectrum matching method and the feature spectrum band identification method, which takes the advantages of the two methods, and establishes an identification rule referring to the expert system. As such, the reliability and automation of the identification are greatly improved. For example, someone introduces the genetic algorithm to the solution of X-ray fluorescence spectrum and achieves good results. The genetic algorithm is a kind of probabilistic algorithm based on simulation on the laws of life, which searches for a solution suitable for the environment. Some studies present a method that takes the fractal dimension as the spectrum identification feature. In this method, the fractal dimension of the spectrum signals is obtained by phase-space reconstruction. Different spectrums can be identified through comparing the fractal dimension of the spectrum signals. Someone also proposes another spectrum identification scheme based on multi-spectral features and integrated neural network, and implements system design. However, most of the aforementioned methods and corresponding implementation software are on the basis of human involvement or assistance. This paper attempts to present an automatic spectrum identification method based on traditional signal processing methods and advanced algorithms. Under the current condition, overly complex algorithms do not required for the automatically recognizes by machines, while characteristic values are needed used for the machine's independent judgment.

The basic ideas of both the Fourier transform automatic identification method and the wavelet transform automatic identification method require a standard library, which is used to compare the calculated characteristic values of the samples with the characteristic values in the library obtained by the same method. Therefore, it is very significant to create such a feature library. The general spectrums of an alloy, which includes Stainless steel, cobalt alloy, nickel base alloy, Titanium alloy and middle-low alloy steel are chosen as the spectrum library, wherein each library contains multiple spectrums. The spectrum library is characterized by a processing method based on the fast Fourier transform and characterized by a processing method based on wavelet trans-

form respectively, which will be mentioned below. The feature libraries corresponding to the two methods thus obtainned. The alloy of Stainless steel, cobalt alloy, nickel base alloy, Titanium alloy and middle-low alloy steel are selected as the testing samples. The spectrums of the samples are general random spectrums completely different from the feature libraries.

A suitable experimental environment and proper tools are in need on the basis of aforementioned works, We had chosen the NI LabVIEW environment. LabVIEW (Laboratory Virtual Instrument Engineering Workbench) is a kind of graphical programming language, which uses icons to create applications instead of text lines. In traditional text-based programmming languages, programs are executed in sequences of statements and directives. Whereas, in LabVIEW a dataflow programming mode is used, that VI (abbreviation of virtual instrument, refers to program modules in LabVIEW) and execution sequences of functions are determined by the data flow between nodes of block diagrams. With the LabVIEW programming language, flow charts or diagrams are used instead of program codes. LabVIEW is an end-user oriented tool, which adapts terms, icons and concepts familiar with technical staffs, scientists, and engineers as much as possible. LabVIEW can enhance the ability of building your own science and engineering systems, and can provide a convenient way to implement instrumental programming and data acquisition systems. The working efficiency of the instrument system can be greatly improved with LabVIEW in its principle researching, designing, testing and implementation.

3 The spectrum unscrambling and identification method based on fast Fourier transform

The Fourier transform would be firstly taken into consideration if Mentions signal analysis in transform domain. Through the discrete Fourier transform (DFT), a finite-length sequence *x(n)* can be transformed into a discrete finite-length sequence *x(k)* in its frequency domain [4].

$$x(k)=\sum_{n=0}^{N-1}x(n)W_N^{nk},\ k=0,1,2,\ldots,N-1. \tag{1}$$

Whose inverse discrete Fourier transform (IDFT) is

$$x(n)=\frac{1}{N}\sum_{k=0}^{N-1}x(k)W_N^{nk},\ n=0,1,2,\ldots N-1, \tag{2}$$

wherein $W_N=e^{-j2\pi/N}$, and *x(n)* and *x(k)* may be a real number or a complex number. It can be concluded that to calculate a sample sequence, *N* times of multiplication operation and *N*-1 times of addition operation of complex numbers should be done. Therefore, it is difficult to perform real-time processsing with DFT because of its huge amount of computations.

Since it is very important to do discrete Fourier transform in signal analysis, it is very necessary to find fast algorithms for discrete Fourier transform and corresponding inverse transform. Fast Fourier Transform (FFT), namely fast discrete Fourier transform (FFT) is a kind of fast DFT algorithm that invented by J.W.Cooley and J.W.Tukey in 1965, wherein the periodicity and symmetry of the factor of the W_N is utilized. The Discrete Fourier Transform is equivalent to multiplying a sequence *x* (as a column vector of *n* elements) by a $n\times n$ matrix F_n, which needs n^2 multiplications. Whereas the fast Fourier transform is to change the matrix F_n into a special form, which can reduce the times of multiplications into approximate $5n\log_2 n$ times. If n=1000, then by the fast Fourier Transform, the times of multiplications can be reduced from millions to about 50, 000.

In the later decades, it is further developed with the FFT algorithm, wherein the most commonly used FFT algorithms now are radix-2 algorithm and split-radix algorithm. It is known that the FFT algorithm has been introduced in detail in many other documents, so this essay will not describe it in detail.

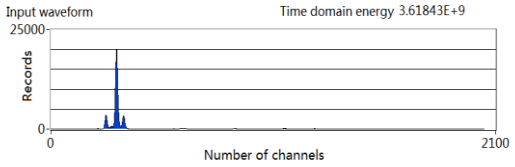

FIGURE 1 A spectrum captured by X-ray fluorescence spectrometer

Referring to Figure 1, a spectrum is shown. In physical, it is the channel value that corresponds to the horizontal axis, rather than a time-sequence. However, in a broad view, it can still be taken as a complete process. The aforementioned fast Fourier transform can be performed using the signal analysis tool in LABVIEW, as shown in Figure 2.

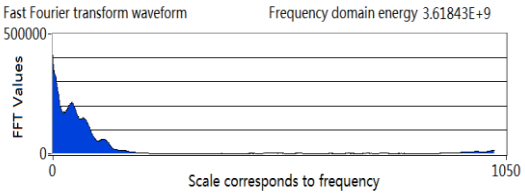

FIGURE 2 The result of fast Fourier transform

It is shown in the Figure 1 and Figure 2 that the energy in time domain calculated from the original pattern before the transformation must be equal to the energy in frequency domain calculated after the transformation, which is in line with Parseval's Theorem. It should be noted that the points in the abscissa only needs half of the original after the Fourier transform, even a quarter or less effective portion thereof with experiences. Accordingly, the amount of computation can be greatly reduced. Unfortunately, it is clear that the characteristic frequencies of almost all the general spectrum samples after the FFT in different spectrum patterns cannot be clearly distinguished by the Fourier transform. In order to make the FFT method continue, a method which combines the energy values and the boundary detection is used, considering the Parseval's Theorem. In this essay, the energy value is selected as the characteristic value, and then transformed by FFT transform, and last detected by its boundary for spectrum pattern recognition.

As shown in Figure 3, the energy value of the to-be-identified spectrum pattern is compared with that of the sample spectrum pattern in Library. The situation should be

very clear if the two energy values are not in the same magnitude. Otherwise, a threshold of range should be defined if the two energy values are close to each other. Additionally, comparison between the boundary values can be performed under the premise that the energy values are as the characteristic values. The pink line and blue line in the Figure 3 represent the top boundary value and bottom boundary value respectively. The identified spectrum pattern is transformed by FFT transform and compared with the boundary values. Points out of the boundaries are marked in red. In this figure, points that do not match occupy 0 percent of the total number of points, in some instances, this value may become large. Therefore, such percentage can also be taken as the threshold value of range for automatic identification.

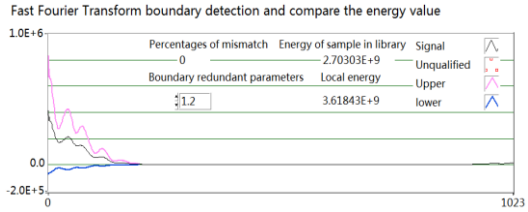

FIGURE 3 The spectrum unscrambling based on Fast Fourier transform

As can be concluded from the identification result, the material as shown in the Figure 3 is probably stainless steel.

4 The spectrum unscrambling and identification method based on wavelet transform

The basic principal of wavelet transformation is: after being dilated and translated, the original signal is decomposed into a series of sub-band signals with good local features in time domain and frequency domain, thus achieving local time and frequency analysis [5]. Wavelet transform is strongly correlated to the data, which make the energy of the signal exist only in some large wavelet coefficients in the wavelet domain, whereas the energy of the noise distributes throughout the whole wavelet domain. Thus, through the wavelet decomposition, the amplitude of the wavelet coefficients of the signal should be greater than that of the noise. It is considered that whose amplitude of the wavelet coefficients is relatively larger should be generally the signal, while whose amplitude of the wavelet coefficients is relatively smaller should probably be the noise. Therefore, wavelet transform shows better performance of detection in non-stationary signals with low SNR. When the scale of a signal to be detected changes, there will be a wavelet with an appropriate scale to match it. On the contrary, by the short time Fourier transform, the profiles of the signal can only be detected when the basic function matches the signal scaling function, while it is difficult to get a satisfactory result in other cases.

The purpose of Wavelet transform and the purpose of Fourier transform are the same: the signal is expressed as a linear combination of basic functions. The difference between the two transformations is: the basic function of the Fourier transform is the harmonic function $e^{-j2\pi/N}$, whereas the basic function of the wavelet transform is the Mother Wavelet Function $\Psi(t)$ with compact support generation. In the wavelet transform, a wavelet sequence is obtained through dilating and translating to the mother function $\Psi(t)$:

$$\psi_{a,b}(x)=\frac{1}{\sqrt{a}}\psi(\frac{x-b}{a}),\quad a,b\in R;a\neq 0 \tag{3}$$

where a is the dilating factor, b is the translating factor.

As to a given function $f(t)\in L^2(R)$, its continuous wavelet transform is:

$$W_f(a,b)=<f,\psi_{a,b}(x)>=\int_{-\infty}^{\infty}f(x)\psi_{a,b}(x)dx =\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}f(x)\psi(\frac{x-b}{a})dx \tag{4}$$

The selection of the basic wavelet function is very important, which often depends on the application [6]. The wavelet function generally has two basic characteristics geometrically: it must be an oscillating function and it must be a function with rapid convergence. The above two criteria should be followed When selecting a wavelet function or constructing a wavelet function. Different dilating factors and translating factors will bring about great changes in geometrical shapes of the wavelet function.

Since the translating factor corresponds to the time t, and the dilating factor corresponds to the frequency *f*, the time-frequency plane (*t, f*) becomes a time-scale plane (*b, a*). The principal of weak signal detection based on Wavelet transform is to put the signal with the noise on a time-scale two-dimensional plane, and then process the signal in time-division and frequency-division modes based on distinct characteristics of the signal and the noise. When the scale parameter a increases, the basic function $\psi_{a,b}(x)$ turns to a broadening wavelet, which corresponds to a low frequency function, whose frequency resolution is improved, but time resolution is reduced. Otherwise, when the scale parameter a decreases, the basic function $\psi_{a,b}(x)$ turns to a compressed wavelet, whose time resolution is improved, but frequency resolution is reduced. When the signal to be analysed has a sudden change, it is better to choose a smaller "a" to improve the resolution in the time domain. Therefore, the wavelet transform is more suitable for studying signals which are non-stationary and contains short-term transient components [7]. There is a special signal processing package for wavelet calculation in LabVIEW. It is very convenient to do experimental verifications on the idea of wavelet-based identification.

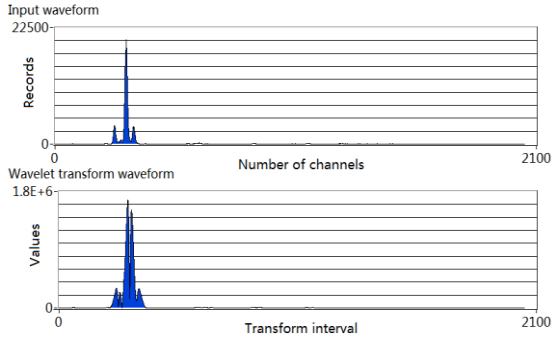

FIGURE 4 The waveform of wavelet transform

As shown in Figure 4, in the upper part, the input signal is a spectrum recorded by the X-ray fluorescence spectrometer. The abscissa of the spectrum physically represents the channel number of the spectrometer, while the ordinate of the spectrum represents values recorded by the instrument in every channel. In the lower part of the figure, there shows a waveform transformed from the input signal by broadly wavelet transform.

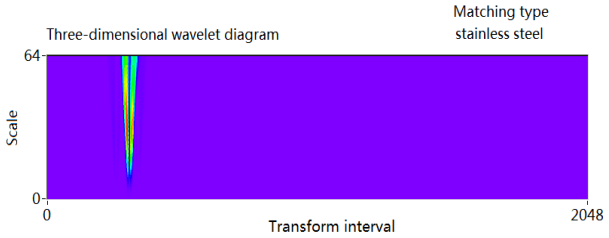

FIGURE 5 Singular points detection in a three dimensinal wavelet digram

As can be seen in Figure 5, it should be noted that the parameter SCALS here controls the scaling ratio of Wavelet transform. And the peaks with different amplitudes of the spectrum in the figure can be considered as singular points of the signal, which can be used as characteristic values of the spectrum. Thus, it can be taken as a detection process of singular points that the solution and identification s process of the spectrum [8]. For a long time, the Fourier transform is a main tool for study the singularity of functions. In detail, by studying the attenuation in the Fourier transform domain of the function of to infer whether there is singularity on the function and how serious the singularity is. However, referring to Fourier transform, its ability to distinguish details is often unsatisfactory. In contrast, wavelet transform is more spatially localized. Accordingly, it is more effective to use wavelet transform to analyse the location of singularity and the singular degree.

A multi-scale peak detection can be carried out toward the data that processed through continuous wavelet transform from the original spectrum, and the detected peaks as shown in Figure 5. Positions of singularity can be considered as characteristic values for identifying the spectrum. The magnitude of the ordinate and the resolution of the abscissa of the wavelet transform data can be used as scaling parameters of the multi-scale peak detection. Peaks less affected can be excluded by adjusting the amplitude threshold, and a proper characteristic value can be accurately found from the overlapped peaks by adjusting the resolution parameter. The characteristic value is corresponding to the position of transform interval, and it can also be corresponding to the channel number of the original input spectrum. Accurate spectrum identification can be performed by comparing that characteristic value with the characteristic library of different materials recorded in the sample library.

Figure 5 gives the automatic identification result, the material shown in the in figure is probably stainless steel.

5 Experiments and results

Two automatic recognition algorithms were designed for testing sample data in LABVIEW environment, which were based on fast Fourier transform and wavelet transform. Sample data contain general spectroscopy spectrum of Stainless steel, cobalt alloy, nickel base alloy, Titanium alloy and middle-low alloy steel five alloys. Before the experiment, we were prepared alloy feature library for each algorithm.

5.1 TESTING FOR THE AUTOMATIC IDENTIFICATION METHOD BASED ON FAST FOURIER TRANSFORM

The FFT method similar to a spectrum matching filtering method (Matched Filtering) is a kind of method to identify a specified target from a background whose spectral information is unknown. The basic idea thereof is to take the material spectrum in a standard spectrum database as a vector in the Q mean space, and to build a filtering vector detector according to the matching degree between the spectrum of the image and the reference spectrum vector, then to filter the input image, so that the spectral response of the target is maximized while the spectral response of other end of the element is suppressed. As such, the contrast between the target and the background is maximized and an optimal matching is achieved, and then an appropriate threshold should be selected, and the target information can be extracted.

Originally, the matching filtering is developed for calculating the relative abundance of the material content of less. Therefore, when the content of the target substance is not few, it should be careful and cautious to apply the method and to interpret the results of the method. The Mixture Tuned Matched Filtering (MTMF) [9] is a composite method, which combines the linear mixed decomposition method and the matched filtering method. It combines the advantages of matched filtering method that do not need other background end elements and the constraint condition that the content of each end element of the image element in the linear mixed decomposition method should be positive and the sum of the content should be 1, thereby the detection limit of the elements is reduced, and the trace elements in the samples that other methods could not detect would be detected.

In Table 1, results by artificial matching identification of 51 samples are excerpted, which are compared with the results by FFT automatic identification algorithm. As can be seen, most of the results are consistent, except that the results in No. 24 and No. 25 by automatic identification identified the stainless steel as the nickel base alloy. After careful analysis, it can be known that by the automatic identification algorithm based on FFT, for both the No. 24 and No. 25 samples, it is difficult to clearly distinguish whether they belong to stainless steel or nickel base alloy, either by energy identification condition or by boundary matching condition. As to the energy identification condition and the boundary matching condition, the libraries of the stainless steel and the nickel base alloy themselves are also very similar. This kind of situation is improved in the automatic identification algorithm based on wavelet.

TABLE 1 FFT Automatic identification test results

Sample number	Manual matching results	Automatic matching results
1	cobalt alloy	cobalt alloy
2	cobalt alloy	cobalt alloy
3	cobalt alloy	cobalt alloy
4	cobalt alloy	cobalt alloy
5	cobalt alloy	cobalt alloy
6	cobalt alloy	cobalt alloy
7	cobalt alloy	cobalt alloy
8	middle-low alloy steel	middle-low alloy steel
9	middle-low alloy steel	middle-low alloy steel
10	middle-low alloy steel	middle-low alloy steel
11	middle-low alloy steel	middle-low alloy steel
12	middle-low alloy steel	middle-low alloy steel
13	nickel base alloy	nickel base alloy
14	nickel base alloy	nickel base alloy
15	nickel base alloy	nickel base alloy
16	nickel base alloy	nickel base alloy
17	nickel base alloy	nickel base alloy
18	nickel base alloy	nickel base alloy
19	nickel base alloy	nickel base alloy
20	nickel base alloy	nickel base alloy
21	nickel base alloy	nickel base alloy
22	nickel base alloy	nickel base alloy
23	nickel base alloy	nickel base alloy
24	stainless steel	nickel base alloy
25	stainless steel	nickel base alloy
26	stainless steel	stainless steel
27	stainless steel	stainless steel
28	stainless steel	stainless steel
29	stainless steel	stainless steel
30	stainless steel	stainless steel
31	stainless steel	stainless steel
32	stainless steel	stainless steel
33	stainless steel	stainless steel
34	stainless steel	stainless steel
35	stainless steel	stainless steel
36	stainless steel	stainless steel
37	stainless steel	stainless steel
38	stainless steel	stainless steel
39	stainless steel	stainless steel
40	stainless steel	stainless steel
41	Titanium alloy	stainless steel
42	Titanium alloy	Titanium alloy
43	Titanium alloy	Titanium alloy
44	Titanium alloy	Titanium alloy
45	Titanium alloy	Titanium alloy
46	Titanium alloy	Titanium alloy
47	Titanium alloy	Titanium alloy
48	Titanium alloy	Titanium alloy
49	Titanium alloy	Titanium alloy
50	Titanium alloy	Titanium alloy
51	Titanium alloy	Titanium alloy

In Table 2, results that comparing and matching a sample with the samples in the library database of FFT algorithm are excerpted, in which the calculation process of such automatic identification algorithm can be reflected. First, it can be seen that for the energy identification condition, the sample is relatively easier to pass cobalt alloy, stainless steel and Titanium alloy. Then the boundary matching condition should be considered, and also the statistical analysis law, through which it can be determined that the sample belongs to Titanium alloy.

TABLE 2 Single sample FFT automatic recognition and matching rsults

Library types	If energy comparison passed?	No matching percentage
cobalt alloy1	T	42.5
cobalt alloy2	T	42.5
cobalt alloy3	T	42.5
cobalt alloy4	T	41
cobalt alloy5	T	66.5
cobalt alloy6	T	52
cobalt alloy7	T	46.5
cobalt alloy8	T	65
cobalt alloy9	T	50.5
cobalt alloy10	T	56.5
cobalt alloy11	T	69
middle-low alloy steel1	F	100
middle-low alloy steel2	F	98.5
middle-low alloy steel3	F	100
middle-low alloy steel4	F	100
middle-low alloy steel5	F	100
middle-low alloy steel6	F	100
middle-low alloy steel7	F	100
middle-low alloy steel8	F	98.5
nickel base alloy1	T	60
nickel base alloy2	F	57.5
nickel base alloy3	F	46
nickel base alloy4	T	44
nickel base alloy5	T	45.5
nickel base alloy6	T	44
nickel base alloy7	F	51
nickel base alloy8	F	72
nickel base alloy9	F	62.5
nickel base alloy10	T	44
nickel base alloy11	T	51
nickel base alloy12	T	42.5
nickel base alloy13	T	40.5
nickel base alloy14	T	43.5
stainless steel1	F	69.5
stainless steel2	T	67.5
stainless steel3	T	70.5
stainless steel4	T	68.5
stainless steel5	T	73.5
stainless steel6	T	72
stainless steel7	T	64
stainless steel8	T	71.5
stainless steel9	T	70.5
stainless steel10	T	71.5
stainless steel11	T	65.5
stainless steel12	T	69
stainless steel13	T	72.5
stainless steel14	T	71.5
stainless steel15	T	69
stainless steel16	T	73
stainless steel17	T	68
Titanium alloy1	T	67.5
Titanium alloy2	T	66.5
Titanium alloy3	T	39
Titanium alloy4	T	74.5
Titanium alloy5	T	73
Titanium alloy6	T	77.5
Titanium alloy7	T	77.5
Titanium alloy8	T	77
Titanium alloy9	T	77
Titanium alloy10	T	72.5
Titanium alloy11	T	77.5
Titanium alloy12	T	78.5
Titanium alloy13	T	77
Titanium alloy14	T	75

To summarize, confirmed by experiments of hundreds of samples, the accuracy rate of identification can reach 90%

by the FFT transform automatic spectrum pattern identification method based on the energy threshold of range and boundary threshold of range. However, there are still some spectrum patterns that could not be clearly distinguished by their characteristics after FFT transform, even the energy value of such spectrum patterns can be easily confused with the sample libraries of different materials. So such spectrum patterns could not effectively be identified through the Fourier transform method. In order to solve such problem and further improve the accuracy of spectrum patterns identification, a wavelet method is introduced here.

5.2 TESTING FOR THE AUTOMATIC IDENTIFICATION METHOD BASED ON WAVELET TRANSFORM

The spectrum is with high spectral resolution, multi-bands, massive data, strong correlation between the bands, high data redundancy, so it is necessary to reduce the dimensions of the data and to weaken the noises before selecting the end elements and identifying the materials. The functions for the wavelet transform are:

1) To separate the information from the noise. After the transformation, the transformation components are arranged in a descending order according to their corresponding characteristic value, whose noise components decreases, while corresponding component information of greater characteristic value is more. Corresponding component information of smaller characteristic value gradually becomes less, while the noise level increases. Corresponding components whose characteristic values are very small and tend to be constant are almost all noises. Therefore, separation the information from the noise can be achieved, and the noise of each component image can be removed on target, or components whose noises are dominant can be abandoned, so as to remove noises.

2) To reduce the dimensions. After the transformation, the correlation between the original spectral bands can be eliminated, so that the information can be concentrated in a small number of components, thus redundancy of data can be reduced. Through observing the characteristic values and corresponding images, the intrinsic dimension of the data can be determined. Selection of end elements and identification of samples can be carried out in transformed low-dimensional space, so as to shorten the duration of data processing and improve the efficiency of processing.

3) To separate features. The spectrum information of samples is isolated in the transformed feature space after wavelet transformation, wherein the spectrum converges on the class feature vector set, and some weak information is enhanced in noise removing transformation. Accordingly, the separability of the data from the sample is increased, and the reliability of the identification is improved.

In Table 3, results by artificial matching identification of 51 samples are excerpted too, which are compared with the results by wavelet automatic identification algorithm. It can be seen that only one sample No.31 cannot be identified by the automatic identification algorithm. As a matter of fact, all the experimental results are not shown in the Table 3 due to the length limitation of the table. There is only this one, which could not be identified in all the hundreds of experiments we have done. Simultaneously, the characteristic values of the samples that calculated through the automatic identification algorithm based on wavelet are shown in Table 3. It can be seen that the numbers of characteristic values that calculated from different samples are different. Value1 and Value2 of each sample (if any) are better consistent. In fact, it can be identified what material the sample belong to basically from the first and second characteristic value. For some samples with more characteristic values, it is better to continue to match with the library database, of course, the premise is that there is similar characteristic value record in the library database, and to prove that the characteristic value is not illegal. For the sample No.31, it can be seen that its characteristic value does not match with stainless steel well. Therefore, it could not get a result by the automatic identification algorithm. The reason may be there some errors occur during the data acquisition process. Interestingly, it can be identified by the automatic identification algorithm based on FFT, which is actually worth further studying.

To summarize, verified by experiments on amount of samples, the identification accuracy of the automatic spectrum identification method based on Wavelet transform can reach 99%. Quite differences were found between the identified spectrum and its corresponding material in the sample library for the unique error, there possibly was something wrong during the data collection process using the spectrometer. It can be concluded that the automatic spectrum identification method based on wavelet transform owns some advantages, but there is still space for further improvement with such method.

TABLE 3 Wavelet automatic identification test results

Sample number	**Manual matching results**	**Characteristic values matching**				**Automatic matching results**
		Value1	*Value2*	*Value3*	*Value4*	
1	cobalt alloy	253	329			cobalt alloy
2	cobalt alloy	253	329			cobalt alloy
3	cobalt alloy	253	329			cobalt alloy
4	cobalt alloy	254	331	376	860	cobalt alloy
5	cobalt alloy	253	329			cobalt alloy
6	cobalt alloy	253	330	860		cobalt alloy
7	cobalt alloy	253	330			cobalt alloy
8	middle-low alloy steel	302				middle-low alloy steel
9	middle-low alloy steel	302				middle-low alloy steel
10	middle-low alloy steel	302				middle-low alloy steel
11	middle-low alloy steel	302				middle-low alloy steel
12	middle-low alloy steel	302				middle-low alloy steel

13	nickel base alloy	357	374			nickel base alloy
14	nickel base alloy	358	375			nickel base alloy
15	nickel base alloy	358	375	860	880	nickel base alloy
16	nickel base alloy	357	374			nickel base alloy
17	nickel base alloy	357	374			nickel base alloy
18	nickel base alloy	357	374			nickel base alloy
19	nickel base alloy	358	374	860	880	nickel base alloy
20	nickel base alloy	357	374			nickel base alloy
21	nickel base alloy	357	374			nickel base alloy
22	nickel base alloy	357	374			nickel base alloy
23	nickel base alloy	357	374			nickel base alloy
24	stainless steel	302				stainless steel
25	stainless steel	302				stainless steel
26	stainless steel	302				stainless steel
27	stainless steel	302				stainless steel
28	stainless steel	252	302			stainless steel
29	stainless steel	302				stainless steel
30	stainless steel	302				stainless steel
31	stainless steel	253	319			No matching type
32	stainless steel	302				stainless steel
33	stainless steel	303	859			stainless steel
34	stainless steel	302				stainless steel
35	stainless steel	303				stainless steel
36	stainless steel	253	303			stainless steel
37	stainless steel	303				stainless steel
38	stainless steel	303				stainless steel
39	stainless steel	303				stainless steel
40	stainless steel	303				stainless steel
41	Titanium alloy	208				Titanium alloy
42	Titanium alloy	208				Titanium alloy
43	Titanium alloy	208	861			Titanium alloy
44	Titanium alloy	209	861			Titanium alloy
45	Titanium alloy	208	817			Titanium alloy
46	Titanium alloy	209	776	861		Titanium alloy
47	Titanium alloy	208				Titanium alloy
48	Titanium alloy	209	818			Titanium alloy
49	Titanium alloy	208				Titanium alloy
50	Titanium alloy	208				Titanium alloy
51	Titanium alloy	209	861			Titanium alloy

6 Conclusions

Methods of automatic unscrambling and identification on spectrums captured by X-ray fluorescence spectrometers are discussed in this essay. It is difficult to accurately distinguish the peaks in the spectrum in the time domain only with experiences, because the peaks in the spectrum are always overlapped and are easily confused. It should be noted that such time domain is a broad concept which is corresponding to the channel number of the abscissa of the spectrum. Therefore, in order to achieve fast and accurate automatic spectrum identification, a method for unscrambling and identifying characteristics based on fast Fourier Transform and a method based on wavelet transform are proposed herein.

It can be concluded from the results of experiments by the aforementioned two methods that: by the method based on fast Fourier transform, the spectrum is transformed from the time domain into the frequency domain (broadly) to find characteristics of the spectrum in the frequency domain, but actually in some spectrums it is difficult to distinguish peaks of different time points (channel numbers) in the frequency domain because the peaks are also overlapped. In contrast, by the method based on wavelet, it can be clearly and accurately to find the characteristic values of the spectrum and thus effectively to identify various elements and materials, wherein the characteristic of "being amplified in part" of non-stationary and short-term transient component accompanied signals is fully utilized.

Accordingly, a theoretical idea for automatic spectrum identification and solution for X-ray fluorescence spectrometers is provided based on the discussion on the aforementioned two methods, and a basis for further development of faster and more accurate qualitative analysis method is built.

References

[1] Zhang L, Dai T 2008 The present status of energy dispersive x-ray fluorescence spectrometry *Modern instrument* **5** 50-5 (*in Chinese*)
[2] An G 2002 Development of wavelength dispersive X-ray fluorescence analysis *Modern instrument* **4** 40-3 (*in Chinese*)
[3] Guo P, Qin D, Hu Z, Du W 2003 A Novel Method for Spectral Signal Pattern Recognition *Spectroscopy and Spectral Analysis* **l23**(4) 811-5 (*in Chinese*)
[4] Yang L, Zhang B, Ye X 2004 Fast Fourier transform and its applications *Opto-Electronic Engineering* **31** 1-7
[5] Xiong Y, Wen Z, Chen G, Huang J 2005 Spectral Signal Recognition Based on Wavelet Packet Analysis *Spectroscopy and Spectral Analysis* **25**(8) 1332-5
[6] Wu B, He L, Cai Z, Wu Y 1993 Feature Extraction and Recognition of Waveform Based on Wavelet Transform *J Huazhong Univ of Sc. & Tech* **21**(1) 77-81

[7] Guo T, Wu C, and Qu D 2004 Wavelet Transform Theory and Its Application Progress: a Review *Information and Control* **33**(1) 67-71
[8] Mallat S, Hwang W L 1992 *IEEE Trans on Information Theory* **38**(2) 618-43
[9] Wang R, Yang S, Yan B 2007 A Review Of Mineral Spectral Identification Methods and Models with Imaging Spectrometer *Remote Sensing for Land& Resources* **1** 1-9 (*in Chinese*)
[10] Song Y, Zeng W-m, Liu G-j, Chang Z-y *2012* Development of X-ray Fluorescence Spectrometer for Determination of Tc *Annual Report of China Institute of Atomic Energy* **164** (*in Chinese*)
[11] Lai MS, Xiang LW, Lin J-M, Li H-F 2013 Quantitative analysis of elements (C,N,O,Al,Si and Fe) in polyamide with wavelength dispersive X-ray fluorescence spectrometry *Science China (Chemistry)* 1164-70 (*in Chinese*)
[12] Yokoi K, Ohashi H, Tochio T, Ito Y, Shoji T, Matsuno T 2006 High resolution X-ray fluorescence spectrometry for water analysis of metals in East China Sea *Chinese Journal of Geochemistry* **S1** 152-3 (*in Chinese*)
[13] Schipp F, Wade W R 1998 Fast Fourier Transforms On Binary Fields *Approximation Theory and Its Applications* **1** 91-100
[14] Novel algorithm for complex bit reversal:employing vector permutation and branch reduction methods *Journal of Zhejiang University(Science A:An International Applied Physics & Engineering Journal)* 2009 **10** 1492-9 (*in Chinese*)
[15] Li S-h, Xian J 2013 A multiscale Galerkin method for the hypersingular integral equation reduced by the harmonic equation *Applied Mathematics:A Journal of Chinese Universities(Series B)* **1** 75-89 (*in Chinese*)
[16] Ma CM, Chen H, Yu JY, Long T 2013 A novel conflict-free parallel memory access scheme for FFT constant geometry architectures *Science China(Information Sciences)* **04** 180-8 (*in Chinese*)
[17] Han WQ, Zhou JY 2013 Acoustic emission characterization methods of damage modes identification on carbon fiber twill weave laminate *Science China(Technological Sciences)* **9** 2228-37 (*in Chinese*)
[18] Arumugam V, Kumar S.B, Santulli C, Stanley A J 2011 Effect of fiber orientation in uni-directional glass epoxy laminate using acoustic emission monitoring *Acta Metallurgica Sinica (English Letters)* **5** 351-64 (*in Chinese*)
[19] Comparison between Windowed FFT and Hilbert-Huang Transform for Analysing Time Series with Poissonian Fluctuations: A Case Study *Chinese Journal of Astronomy and Astrophysics* 2006 **4** 503-12 (*in Chinese*)
[20] Asamwar R S, Bhurchandi K M, Gandhi A S 2010 Interpolation of Images Using Discrete Wavelet Transform to Simulate Image Resizing as in Human Vision *International Journal of Automation & Computing* **1** 9-16
[21] Luo M-J 2007 Application Of Wavelet Transform On Diagnosis And Prediction Of Milling Chatter *Chinese Journal of Mechanical Engineering* **3** 67-70 (*in Chinese*)
[22] Sclabassi R J, Hsu J T, Li C C 2007 A Wavelet Transform Based Pocs Superresolution Algorithm *Journal of Electronics(China)* **5** 642-8 (*in Chinese*)
[23] Wu W 1985 Applications Of Orthogonality Theory Of Discrete Normal Modes To The Identification Of Free Oscillation Eigenfrequencies And Determina Tions Of Spectral Peak Positions And Spectral Line Widths *Science in China* **Ser.B 06** 641-54 (*in Chinese*)
[24] Guo L, Moore J B 1991 Stochastic System Identification Via Adaptive Spectral Factorization *Systems Science and Mathematical Sciences* **03** 275-88
[25] Boumaiza S, Ghannouchi F M 2007 Spectral Methods for Accurate Identification and Quantification of Memory Effects of Wideband RF Power Amplifiers *Chinese Institute of Electronics(CIE). Proceedings of 2007 5th International Conference on Microwave and Millimeter Wave Technology[C].Chinese Institute of Electronics(CIE)* **4**
[26] Kozánek J, Pust L 2011 Spectral properties and identification of aerostatic bearings *Acta Mechanica Sinica* 01 63-7 (*in Chinese*)
[27] Guo Y 2001 Parameter Identification By Mixed Spectral-Pseudospectral Approximations *Acta Mathematicae Applicatae Sinica (English Series)* **02** 218-32 (*in Chinese*)
[28] Balachandar V, Lakshman K B, Sasikala K, Manikantan P, Sangeetha R, Mohana D S 2007 Identification of a high frequency of chromosomal rearrangements in the centromeric regions of prostate cancer patients *Journal of Zhejiang University(Science B:An International Biomedicine & Biotechnology Journal)* **09** 638-46 (*in Chinese*)
[29] Zheng W, Shao J, Wang M, Huang D 2013 A Thin Cloud Removal Method from Remote Sensing Image for Water Body Identification *Chinese Geographical Science* **04** 460-9 (*in Chinese*)
[30] Titulaer M K 2013 Candidate Biomarker Discovery for Angiogenesis by Automatic Integration of Orbitrap MS1 Spectral- and X!Tandem MS2 Sequencing Information *Genomics, Proteomics & Bioinformatics* **03** 182-94
[31] Yao Z, Song Y, Yang J, Gao Q 2012 Spectrum Matching Analysis of Chaotic Binary Modulation Excitation Sequence for Sonar Systems *IEEE Beijing Section.Proceedings of 2012 IEEE 11th International Conference on Signal Processing (ICSP 2012) IEEE Beijing Section* **4**

Authors

Jiening Xia, born in June 27, 1982, NingXia, China

Current position, grades: research assistant, master Institute of Seismology.
University studies: Wuhan University, Institute of Seismology, China Earthquake Administration.
Scientific interest: electronic engineering, photoelectric technology, instrumentation and algorithms.
Publications: 6.

Zhigao Chen, born in April 15, 1975, Wuhan Hubei, China

Current position, grades: the president at the Wuhan Institute of Earthquake Science Instruments.
University studies: associate professor, master.
Scientific interest: measurement and control technology and instrument.
Publications: 5.