**XVI IMEKO World Congress**
Measurement - Supports Science - Improves Technology - Protects Environment ... and Provides Employment - Now and in the Future
Vienna, AUSTRIA, 2000, September 25-28

# METHODS TO AUTOMATICALLY SEPARATE ABNORMAL SIGNAL

***X. Yu, J. Yu and H. Wang***

Precision Project Research Institute, Precision Instrument Department
Hefei University Of Technology, Hefei, Anhui, China 230009

*Abstract: Recently, the method of updating process of dynamic testing data is a research focus in measuring field. A method of processing abnormal signal by computer is described in this paper. It can automatically identify and separate abnormal components in dynamic measuring data, and repair original signal properly, to eliminate the effect of abnormal signals to the base character analysis of dynamic testing data.*
*Keywords: dynamic testing, data processing, abnormal component, wavelet analysis*

## 1 INTRODUCTION

In recent years, the applications of photoelectric technology, digital technology, microprocessing technology, image display technology and automation technology are growing wider in the measuring field, so the requirements of updating process of dynamic measuring data and updating correct of measuring error are more and more high. Though there are many mature methods in theory and technique to process static testing data, most of them are unable to fit the requirement of processing testing data. For example, there are several feasible methods available to reject the abnormal values in static testing data. However, it is still a very difficult work so far to separate abnormal values from dynamic measuring signals and repair them, so the methods of updating process of dynamic testing data become a research focus in the measuring field. In this paper, a method to process abnormal components automatically in dynamic testing data by computer is introduced. It can automatically identify and separate them, and repair original testing data properly, thus reduce the influence of abnormal signals to analyze the base character of dynamic testing signals.

Abnormal signals in dynamic measuring data are those, which only appear in a little time and their distributing regulation (amplitude, frequency) has rather large, distinguish with whole data. Statisticians found that there are generally 1% to 20% abnormal values in industrial and engineering measurements. Even in precise measurements, their ratio is often 1% to 5%.

The causes of abnormal signals are in many aspects, for example, [1] sudden change of measuring or processing environment. The common ones are the fluctuation of current source, the vibration of the surroundings and so on; [2] the measuring system or processing system appear recessive accidents which make measuring or processing state change partially; [3] there are defects in the measurand or on its surface, such as cracks, flaws, and so on.

The cause of abnormal values is different, then their appearing form in data sequence is different. Some of them appear in pulse form, and some in plaque form. These abnormal values can cause distortion of signal fitting model, and lead large error to signal analyzing result; the polluted signals can also cause wrong action if they are used to control.

Sometimes it is only required to eliminate the effect of abnormal signals to the analysis of total signal regulation; sometimes it is required to analyze them specifically and study their causes. These are all based on the identification and the separation of abnormal signals in dynamic measuring data.

Nowadays, in the process of dynamic signals, the methods to reduce the influence of abnormal signals are: the first one is to restrain their effects to signal analysis by wave filtering. This method inevitably filters some useful signals and causes some degree of distortion. Because the spectral of impulse signals is continuous, wave filtering can only eliminate part of the effect of abnormal signals. The second one is to analysis methods with stability (abnormal resistance) to limit the effect to analysis result of small percentage. But so far, only robust analysis but real stable analysis is realized because the absence of a set of rather complete statistic theory. The third one is to reconstruct polluted measuring signal, that is , to make cleaning treatment. The third method not only can eliminate the influence of abnormal signals, but can make further analysis of the case of them and obtain some useful abnormal information, according to abnormal components separated. Therefore, the object studied in this paper is the third one.

The valid identification methods of abnormal signals of amplitude and those of frequency are different. In the separation, separate the abnormal signals of amplitude at first, then separate those of frequency. In this paper, only the former is studied.

## 2 THE SEPARATION OF ABNORMAL COMPONENTS AND THE REPAIR OF POLLUTED SIGNAL BY METHOD OF SLIDING–MEAN

The method of sliding –mean can only separate the abnormal signal components of amplitude. In general speaking, abnormal signals of amplitude have the character of amplitude accident. It change amplitude relevant to adjacent signals must be much larger than that between normal signals. Therefore, it is possible to roughly reject abnormal signals from the data sequence according to the difference between the sampled value and the average of forward and afterward samplings.

Suppose the sequence of dynamic measuring data is $x(i)$ , $i = 1,2,...,n,$ compute

$$z(i) = x(i) - [x(i+1) + x(i-1)]/2 ,$$

and order the absolute values of z(i) from small to big. According to the conclusion from statistic experiment, abnormal values will not exceed 20% of total quantity in industrial measuring. So it is reasonable to regard x(i) relevant to former 80% z(i) as normal values, then calculate the average x¯ and standard deviation á of x(i) form these 80% data.

At last, calculate: $w(i) = \left|x(i)\right| - \bar{x} \quad i = 1,2,...,n,$

If it is appropriate that $w(k_1) > 3\ s$ , $w(k_2) > 3s$ , ... $w(k_m) > 3s$ , relative to $k_1$, $k_2$, ... $k_m$, then it is reasonable to regard as abnormal values. Suppose

$$yc(i) = \begin{cases} x(i) & i = k_1, k_2, \cdots, k_m \\ 0 & i = \text{others} \end{cases}$$

then the series of $yc$ is that of abnormal signals separated from $x$ series.

The rejection of abnormal values, especially parapsoriasis ones, will damage the continuum of non-periodic function and periodic one of the data sequence, and effect the fitting accuracy of determined components; and its effect is small to the analysis of random signal components, especially the character of stable random procession (it only reduce the quantity of data really analyzed ). Accordingly, the major task of repairing data is to repair the breakpoint of determined components.

To repair breakpoint must determine its breadth at first. Then adopt asymmetric weighing center algorithm to repair it according to the information of data blocks whose breakpoints are adjacent and whose breaths are equal. The weighting coefficients are determined through five-point one-time end effect method. In order to improve repair accuracy, we can make sliding average between the fore and after data blocks, and then select the average as repaired data.

Take the example of repairing the breakpoint of the breakpoint of $k_1$ and support it has the breadth of $a$ data. The adjacent data block before it is

$$\mathrm{x}_Q = \{x(k_1 - a - 3) \quad x(k_1 - a - 2) \quad x(k_1 - a - 1) \quad x(k_1 - a)\},$$

The adjacent data block after it is

$$\mathrm{x}_H = \{x(k_1 + a) \quad x(k_1 + a + 1) \quad x(k_1 + a + 2) \quad x(k_1 + a + 3)\},$$

Take fore and after weighting coefficient individually as $\mathrm{w}_Q$ and $\mathrm{w}_H$ , then the repairing data is:

$$\mathrm{x}_{XF} = (\mathrm{w}_Q \mathrm{x}_Q^H + \mathrm{w}_H \mathrm{x}_H^H)/2$$

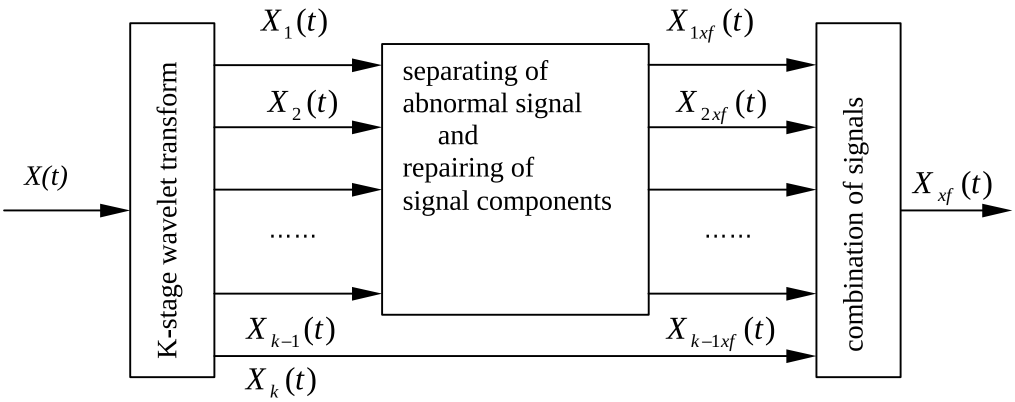
However, from the computer simulation results, we can find that the above-mentioned method can effectively process those signals without periodic components. The abnormal components, either of pulse forms or plaque form, can be separated accurately (shown in example 1). It also has the same separating effect on the signals disturbed by whitenoise (show in example 2). But to the signals containing periodic components, only part of abnormal values can be separated and the effect of abnormal components can not be rejected completely.

## 3 WAVELET ANALYSIS USED TO THE SEPARATION OF ABNORMAL COMPONENTS AND THE REPAIR OF SIGNALS

As introduced above, if we identify the abnormal components of amplitude in periodic signals by classical methods, we often can not identify the abnormal signals, which are at the bottom of wave, or which have same amplitude with periodic change of signal. It is hard to separate periodic signals at first in the case of abnormal signal components existing. Then without changing the character of signals, which containing abnormal components, is it possible to “strentch” them? Computer proves that wavelet transform has this ability.

From figure 2, we can find, when dynamic measuring signals processed by wavelet transform, the amplitudes of abnormal signal components are mainly decompose to high frequency signal components. They have little effect on the medium and low frequency signal constituent. Meanwhile, the midlines of high frequency signal components assume “line” form (figure3-b. 3-c. 3-d). Therefore, we can repair the high and minor high frequency signal components by the method of sliding-mean, then compose frequency domain signal constituents, and obtain the repaired measuring signal sequence x $_{xf}(t)$ (figure3-f). The following shows the flow-process block diagram of processing course.

When having done such processing, it is possible to separate abnormal signal components of amplitude completely, and it has better effect of signal repairing. Moreover, the method of decomposing and rejecting abnormal signal effect by wavelet can be used to process periodic and arbitrary signals. The above-mentioned method can also separate abnormal signal components of frequency.




## 4 COMPUTER DATA SIMULATION

The measured of dynamic measurement changes with time, so the dynamic measuring data are always a column of time series with discreteness, orderliness and randomness. In dynamic measurement of geometric senses, sometimes the information carrier of the measured is a determined function, sometimes it is a random function, and sometimes are the two. The part of systematic errors expressed as determined function and that of the random errors expresses as random one. Therefore, dynamic measuring data can be indicated as:

$$Y(t) = [d(t) + p(t)] + X(t)$$

where, $y(t)$ is measuring data; $d(t)$ is non-periodic function; $p(t)$ is periodic function, $x(t)$ is random function.

Because the character of measured signals is often unknown, the method, which can purify and process dynamic measuring signals effectively, must be fit for random combinations of such different components as non-periodic signals, periodic signals, smooth random signals and measuring noises. Therefore, when having computer simulation, three types of normal signal components were fitted:

Non-periodic signal components $x_1(t) = -1.12 + 0.003095t$

Periodic signal components $x_2(t) = 3\sin(0.02pt + 0.785) + \sin(0.006pt + 0.524)$

Components of relevant random signals

$$\begin{aligned} x_3(t) = &-0.5201x_3(t-1) + 0.76561x_3(t-2) - 0.5201x_3(t-1) \\ &+ 0.2803x_3(t-3) - 0.2670x_3(t-4) - 0.1751x_3(t-5) \\ &+ 0.0864x_3(t-6) \end{aligned}$$

White noise $x(4)$ and abnormal signal components $x(5)$ are also fitted. The data length n=100.

Example 1 is the processed data sequence: $x = x_1 + x_3 + x_5$. In figure 1, a is the curve diagram of polluted signals, b is the abnormal components separated from signal in a.

Example 2 is the processed data sequence: $x = x_1 + x_3 + x_4 + x_5$. In figure 1, c is the curve diagram of polluted signals; d is the abnormal components separated from signal in c.

Example 3 is the processed data sequence: $x = x_1 + x_2 + x_3 + x_5$. In figure 1, e is the curve diagram of polluted signals; f is the abnormal components separated from signal in e.

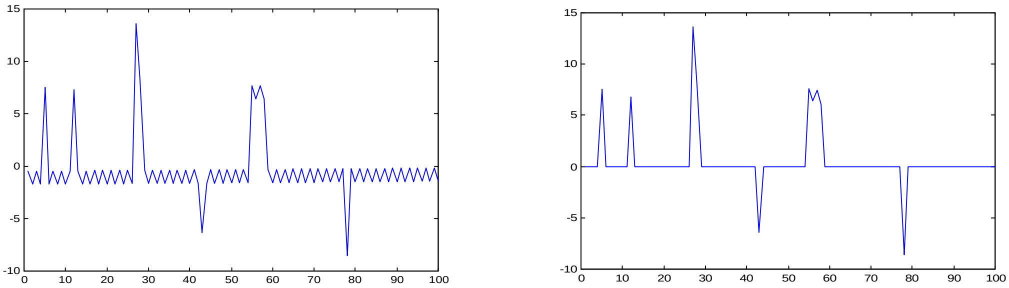
Figure 2 shows the analysis of repairing effect of polluted signals in example 2. Figure 2-a is the curve diagram of polluted signals, Figure 2-d is the polynomial fitting curve of figure 2-a.. Figure 2-b is the curve diagram of non-polluted signals, Figure 2-e is the polynomial fitting curve of figure 2-b. Figure 2-c is the curve diagram of repaired signal, Figure 2-f is the polynomial fitting curve of figure 2-c.

Figure 3 shows the three-stage wavelet-decomposing curve of a measuring data sequence, and the scheme of repairing effect after separating abnormal components. Graph 3-a is the curve of decomposed original signals X(t), it have same constituents with example 3 in figure 1,and the data length n=1000. Graph 3-b is the signal constituent of the highest frequency $X_1$(t). Suppose the cut-off frequency of sampled signals is $f_c$, then the frequency span of this part of signal is $\left[ f_c/2 , f_c \right]$;graph 3-c is the constituent of minor high frequency $X_2(t)$, frequency span is $\left[ f_c/4 , f_c/2 \right]$;graph 3-d is the medium frequency constituent $X_3(t)$ the span is $\left[ f_c/8 , f_c/4 \right]$;graph 3-e is the low frequency constituent $X_4(t)$, the span is $\left[ 0, f_c/8 \right]$;graph 3-f is the curve of repaired testing signal.

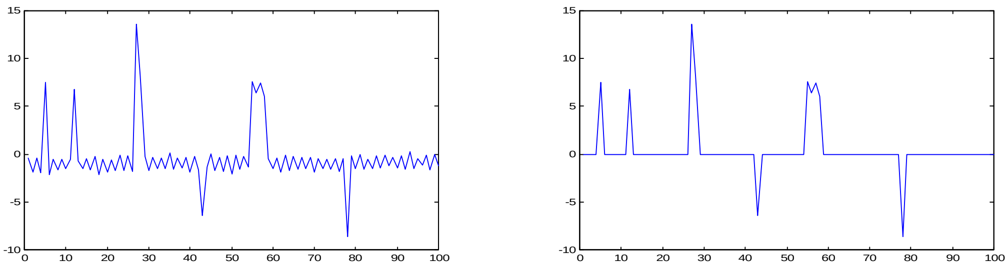
## 5 SUMMARY

The method introduced in the paper, can automatically identify and separate the abnormal components in signal sequence of dynamic measuring and repair the polluted original signals to some degree, thus greatly raise the credibility of measuring data. However, computer simulation proves that, the repairing accuracy of present repairing method is still effected by the complexion of signals and the character of abnormal components. As far as a complex signal sequence or a large “abnormal dot” is concerned, the character of repaired part still has obvious distortion sometimes compared with total character of signals.

Example 1



Example 2



Example 3

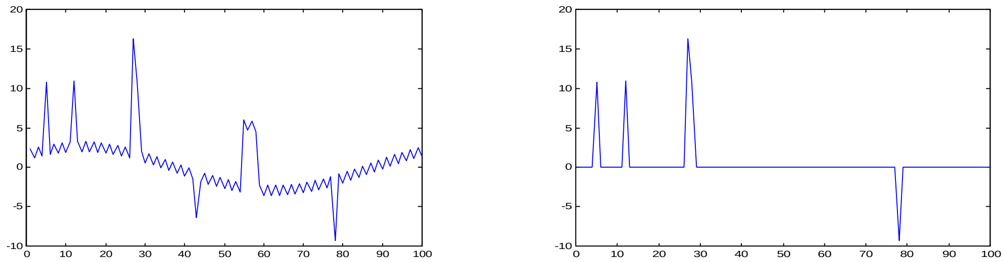
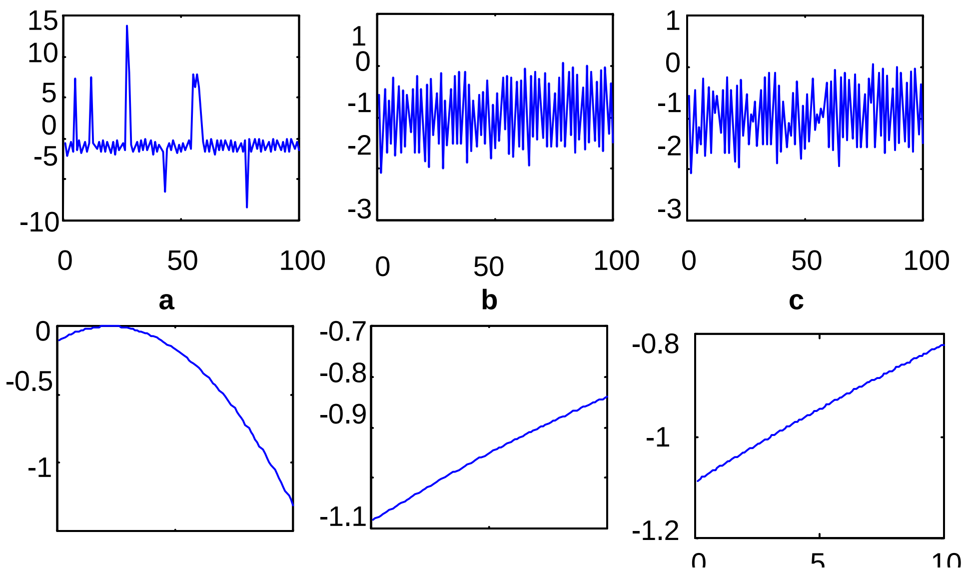


Figure 1 effect scheme of separating abnormal components by traditional methods



-1

-1.5

0 50 100 0 50 100

e f

Figure 2 scheme of repairing effect of polluted signals

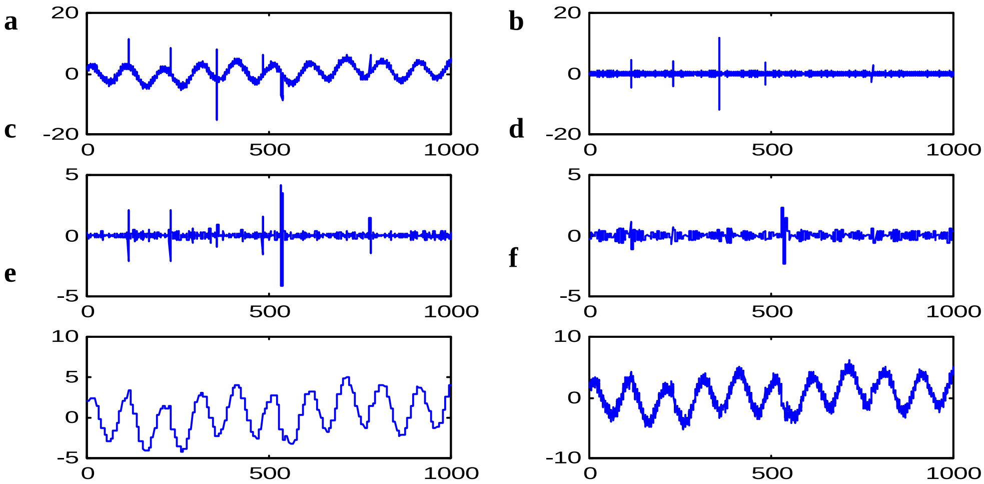



Figure 3 effect scheme of abnormal components separating and signal repairing by wavelet analysis

## REFERENCES

[1] Lin Hong-Hua [Data Processing of Dynamic Measurement] Publishing House of Beijing University of Science and Technology 1995 pp600-602 pp614-627

[2] Zhou Li-Ming, “A Method for Processing High Density Sample Data Based on Feature Points”, ACTA METROLOGICA SINICA Vol.18 pp284-288

[3] Wang Fu-Yuan, “Outlier Detection with Eigenvalue Criterion in Linear Regression and Its Application”, CHINESE JOURNAL OF SCIENTIFIC INSTRUMENT Vol.17, pp273-277

[4] The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. on Information and Theory,1990;36:961005

[5] R.A.Gopinath and C.S.Burrus, Wavelet Transform and Filter Banks, Wavelets-A Tutorial in Theory and Application, 603-654