

The Method for Outliers Excluding of Shipborne Radar Data Based on Wavelet Transform

Zhiheng Zhou^{*}, Jianjun Zhao, Deyi Sang and Libin Yang

Naval Aeronautical Engineering Institute, Department of Ordnance Science and Technology, Yantai, China

^{*}Corresponding author

Abstract—In the dynamic process of warship combat system alignment, the data of shipboard radar measurement at sea calibration object contains outliers with trend item, it can not simply be excluded based on some statistical algorithms. In this paper, an outliers excluding algorithm was proposed based on wavelet transform combine with Laiyite criterion, in the algorithm trend item of data was removed based on wavelet transform first, then exclude the remained outliers in the data according to Laiyite criterion, then merge the trend items in the last, so the outliers excluding of shipborne radar data with trend items was achieved.

Keywords—component; shipborne radar, wavelet transform, 3σ criterion, outliers excluding

I. INTRODUCTION

In the dynamic process of warship combat system alignment, the distance and angle of the calibration object at sea are need to be measured. Due to poor working conditions, in the actual target tracking process, radar measurement data will create deviation interfered by external factors such as fog and rain reflection interference and sea clutter[1]. Especially when shipborne radar is in complex electromagnetic environment, some measurement data may serious deviate from the true value, and becomes a serious anomaly data[2], for this anomaly data, a definition was given by Barnett in statistical research field is: based on a species measure concerned, an exception data point have significant difference with other data[3]. In the process of radar data measurement, which is also known as outliers[4]. The presence of outliers cause serious distortion of the radar measurement data, thereby affect the alignment accuracy of warship combat system. Therefore, excluding the outliers is a very necessary work before processing data.

Classic method of outliers excluding are based on the statistical characteristics of observation data, a threshold is set by calculating the mean and standard deviation, by which to judge and reject the outliers. There some common outliers excluding guidelines such as Dixon criteria, Globus criteria, Romanowski criteria, Chauvenet criteria and Laiyite Criteria, etc. With the development of research on database and data mining technology, outliers detection is gaining more and more attention in recent years and there have been many new algorithm[5], such as extrapolated forecasts identification, differential detection method, likelihood ratio detection, clustering-based algorithm, density-based algorithm, and distance-based algorithm[6]. These algorithms can detect and eliminate data outliers basically in some certain conditions, but

they are mostly process unordered data sets. For the ordered data set such as radar measurement data, they do not take full advantage of observation information; false judgment can be taken easily. In view of this, a outliers excluding algorithm was proposed based on wavelet transform combine with Laiyite Criteria, with the consideration of both the trend item in radar data and the fully retained under the premise of its volatility, provides a guarantee to the next radar data processing, and the measured data with a radar proved the reliability of the method.

II. PRINCIPLE OF OUTLIERS EXCLUDING

A. Laiyite Criteria

According to the theory of error, when measured data follow a normal distribution, the probability of falling within $[-3\sigma, 3\sigma]$ is more than 99.73%, the probability of falling outside this interval is less than 0.3%, so the measured data which falls outside of the range can be considered as outliers and should be excluded[7]. This is the basic idea of Laiyite Criteria, which is also known as method.

Suppose there is a series of measurements with the accuracy of independence, its mean is \bar{X} , residuals is V_i , standard deviation is σ , then:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (1)$$

$$V_i = X_i - \bar{X}, (i = 1, 2, \dots, n) \quad (2)$$

$$\begin{aligned} \sigma &= \sqrt{\frac{\sum_{i=1}^n V_i^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \\ &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \end{aligned} \quad (3)$$

According to Laiyite Criteria, if the residuals $|V_a| = |X_a - \bar{X}| > 3\sigma$, then $X_a, (1 \leq a \leq n)$ can be determined as outliers of the observation data and should be discarded. Recalculate the mean and standard deviation of the data modified after outliers discarded, then distinguish individual measurements in accordance with 3σ criteria to see whether there are new outliers. Repeat the above steps until no new outliers appear, and then all measurement data fall within the 3σ range.

B. Wavelet Transform Theory

Suppose that $f(t)$, $\psi(t)$ are square integrable functions, and

$$\int_R \psi(t) dt = 0 \quad (4)$$

Then call the following integral transformation the continuous wavelet transform[8]:

$$\begin{aligned} (T_\psi f)(a,b) &= \langle f, \Psi_{a,b} \rangle \\ &= \frac{1}{\sqrt{|a|}} \int_R f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \end{aligned} \quad (5)$$

Among them, Ψ is called mother wavelet, $\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right)$ is called wavelet function, wavelet function is a family function obtained by the mother wavelet scaling and translation, the continuous wavelet transform, is a binary function, by converting the function $f(t)$ to the binary function $(T_\psi f)(a,b)$ [9] on time and frequency domain. The wavelet transform in the frequency domain can be obtained by Parseval identities:

$$\begin{aligned} (T_\psi f)(a,b) &= \frac{1}{2\pi} \langle \hat{f}, \hat{\Psi}_{a,b} \rangle \\ &= \frac{\sqrt{a}}{2\pi} \int_R \hat{f}(\omega) \overline{\hat{\Psi}(a\omega)} e^{j\omega b} dt \end{aligned} \quad (6)$$

In the continuous wavelet transform, the scale factor a , time t and offsets b are continuous, but since the computer can only handle discrete data, the scale factor a , time t and offsets b must be discrete, resulting in discrete wavelet transform[10].

In the above formula to take, $a = a_0^j, b = ka_0^j b_0$, then

$$\Psi_{j,k}(t) = a_0^{-\frac{j}{2}} \Psi\left(\frac{t - ka_0^j b_0}{a_0^j}\right) = a_0^{-\frac{j}{2}} \Psi(a_0^j t - kb_0) \quad (7)$$

Let $\Psi(t) \in L^2(R)$, $a_0 > 0$, $\Psi_{j,k}(t)$, calling:

$$(Tf)(j,k) = \int_R f(t) \overline{\Psi_{j,k}(t)} dt \quad (8)$$

as discrete wavelet transform of $f(t)$.

Iterative the wavelet decomposition process, we can get a multi-layer wavelet decomposition, which is also called multi-resolution analysis or multi-scale analysis. Broadly speaking, the functions f within the space $L^2(R)$ of which square

integrable in real number field R that can be described as the limit of a series of approximate function, these approximate functions are obtained at different scales, each function is smooth approximation to the original function.

C. Statistical Characteristics of Radar Data

Consider the discrete systems, state equation is as follows:

$$X_h = F_{h,h-1} X_{h-1} + G_{h-1} W_{h-1} \quad (9)$$

The measurement to X_h satisfied the linear relationship, then the measurement equation is:

$$Z_h = H_h X_h + V_h \quad (10)$$

Among them, the system noise W_h and measurement noise V_h are both white noise, the basic discrete Kalman filtering equation is as follows:

State step prediction equation:

$$\hat{X}_{h/h-1} = \Phi_{h,h-1} \hat{X}_{h-1} \quad (11)$$

State estimation equation:

$$\hat{X}_h = \hat{X}_{h/h-1} + K_h (Z_h - H_h \hat{X}_{h/h-1}) \quad (12)$$

Filter gain:

$$K_h = P_{h/h-1} H_h^T (H_h P_{h/h-1} H_h^T + R_h)^{-1} \quad (13)$$

Step prediction mean square error:

$$P_{h/h-1} = \Phi_{h,h-1} P_{h-1} \Phi_{h,h-1}^T + \Gamma_{h-1} Q_{h-1} \Gamma_{h-1}^T \quad (14)$$

Estimated mean square error:

$$P_h = (I - K_h H_h) P_{h/h-1} (I - K_h H_h)^T + K_h R_h K_h^T \quad (15)$$

And,

$$r_h = Z_h - H_h \hat{X}_{h/h-1} \quad (16)$$

Among them, r_h is residual. When the filter system is optimal, r_h is white noise. The classic outliers detection algorithm mentioned before, assuming that the aforementioned filtering system is the best, that is, r_h is white noise, but in fact due to various noise, radar data obtained by measuring not only includes random error but also contains a variety of system error, Thus, r_h is not necessarily white noise. The difference between optimal and sub-optimal Filtering is shown in Figure 1 and Figure 2. As can be seen from Figure 2, when the filter is not optimal, the residuals are not white noises, they generally does not follow a normal distribution. Therefore, the

assumption that residuals are white noise does not hold in many practical applications.

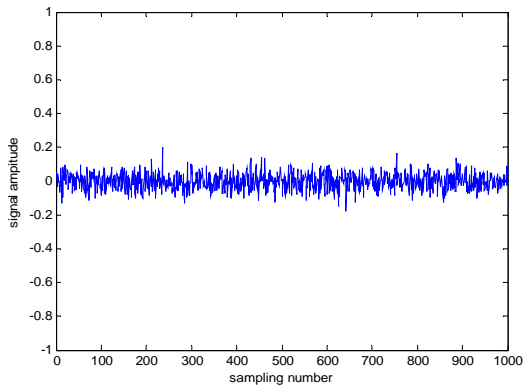


FIGURE I. RESIDUALS ARE WHITE NOISE

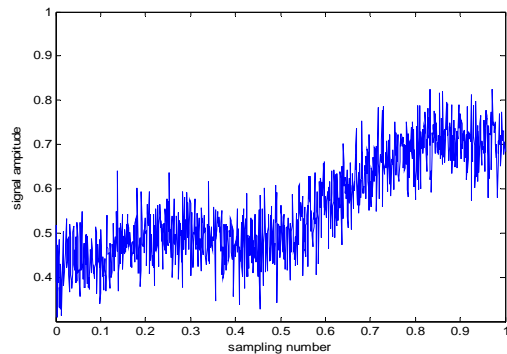


FIGURE II. RESIDUALS ARE NOT WHITE NOISE

III. THE OUTLIERS EXCLUDING ALGORITHM BASED ON WAVELET TRANSFORM AND LAIYTE CRITERIA

Whether an isolated outliers or spot outliers, they exhibit mutated jumping point in the signal, corresponding to the high-frequency part in the frequency domain, so outliers are included in the wavelet detail component of raw data. If you do not take the criterion and discard the component contained the details of outliers directly, and then reconstruct the signal, in order to complete the purpose of outliers excluding, it will cause the loss of high frequency information in the signal. To solve this problem, outliers of the detail component were identified and removed based on the guidelines in this paper, and then the details of component and trend components were reconstructed, so the outliers in the data were excluded. In this way, not only excluded the outliers in the data component, but also retained the high-frequency information of data.

The classic 3σ method is applicable only if the residuals follow a normal distribution, but there is not only random error but also variety of system errors in the actual radar measurement data. To construct the applicable conditions for 3σ method, first wavelet transform was used to extract the

high frequency component of the measured data in this paper, then outliers of high-frequency data were excluded by 3σ method, and finally superimposed the high frequency data after outliers excluding and the original trend term, to get the final results of the radar data.

IV. MEASURED DATA VALIDATION

In order to verify the feasibility and accuracy of the method, the elevation angle tracking data of a certain period of radar was intercepted, there were spot outliers nearby the vicinity of 200 sampling points and 530 sampling point of the segment data, and between 800 and 900 sampling points there were many isolated outliers. First, excluded the outliers of the segment data according to Laiyite Criteria directly, results are as follows:

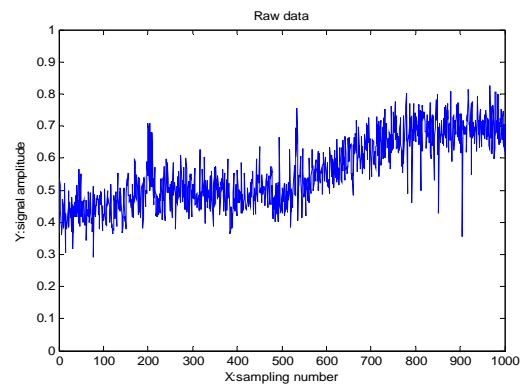


FIGURE III. RAW DATA

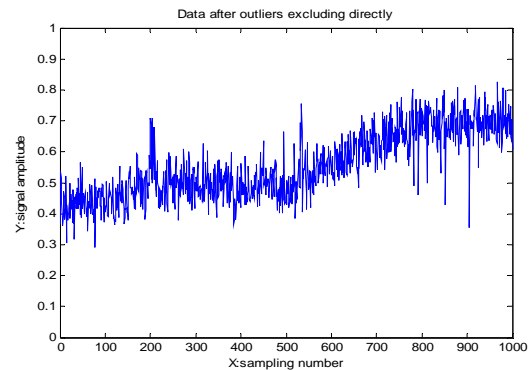


FIGURE IV. DATA AFTER OUTLIERS EXCLUDING DIRECTLY

As can be seen from the results, spot outliers and isolated outliers were basically not excluded, because the trend item contained in the data expand the standard deviation of the data, making the outliers be contained within $[-3\sigma, 3\sigma]$, so as to be used as the normal data.

Then use the method provided in this paper, first transform the data by wavelet, extract the high-frequency component, and then exclude the outliers in rest of the high-frequency data by Laiyite Criteria, results are as follows:

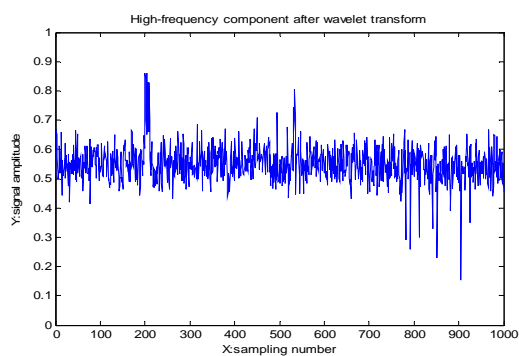


FIGURE 5. HIGH-FREQUENCY COMPONENT AFTER WAVELET TRANSFORM

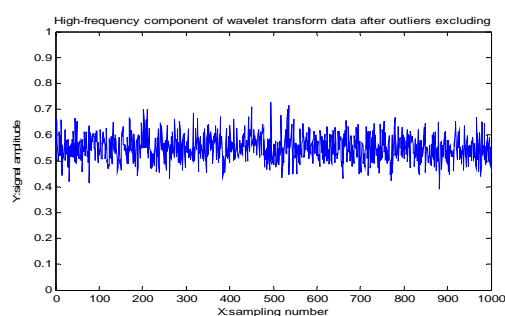


FIGURE 6. HIGH-FREQUENCY COMPONENT OF WAVELET TRANSFORM DATA AFTER OUTLIERS EXCLUDING

As can be seen from the results, after being extracted of the raw data trend term, outliers were more obvious in the data, because the high-frequency component data is more in line with the normal distribution than the original data, so in this case outliers could be excluded by Laiyite Criteria, It is bound to get better results.

At last, the data can be obtained by superposing the high-frequency data after excluding the outliers with the trend term after wavelet transformation, as is shown in Figure 7.

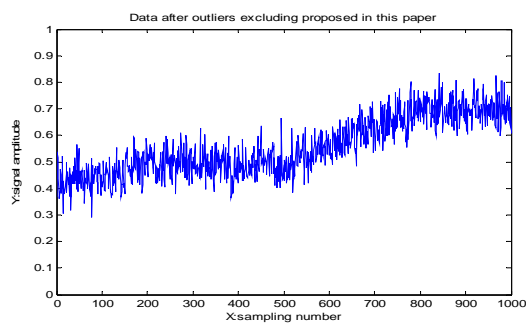


FIGURE 7. DATA AFTER OUTLIERS EXCLUDING PROPOSED IN THIS PAPER

As can be seen from the results, the isolated outliers and spot outliers have been excluded according to the methods proposed in this paper, high-frequency component of the data has been well protected, showing that the method can distinguish outliers components and high-frequency components, thereby retain the high-frequency component and exclude outliers accurately.

V. CONCLUSION

In this paper, the wavelet transform theory and statistical characteristics of the data were analyzed, the outliers excluding method was presented based on wavelet transform and Laiyite Criteria according to the analysis of radar data, the realization step of the method was given, and the method was analyzed and validated by measured data. The results show that the high-frequency components and outliers can be distinguished by this method, and the outliers were accurately excluded and high-frequency data was reserved, which is more accurate and reliable than simply using Laiyite Criteria.

REFERENCES

- [1] Yang Juan, Chi Jian-jun, Wang Zhong-wei. Simulation of Radar Operational Capability in Complex Electromagnetic Environment[J]. Ordnance Industry Automation, 2012, 31(2):1-4
- [2] Barnett V, Lewis T. Outliers in Statistical Data: 2nd[M]. New York: John Wiley & Sons, 1999
- [3] Hu Kui. Radar Data Outliers Elimination Method Based on Soft-Weighted k-Mean Distance Outliers Factor [J]. Ordnance Industry Automation, 2013, 32(4):67-68.
- [4] Zhou Ning. Study on Outliers Eliminating Method for Data Processing of Exterior Trajectory [J]. Journal of Test and Measurement Technology, 2008, 22(4):313-316.
- [5] Zhu Zhuan-min, Qiu Hong-xing, Li Ji-sheng, etc. Identification and elimination of Outliers in dynamic measurement data [J]. Systems Engineering and Electronics, 2004, 26(2):147-149.
- [6] Zhang Ting, Wang Bo. New judgment method of continuous Outliers [J]. Journal of Southeast University (Natural Science Edition), 2005, 35(S2):225-227.
- [7] Liu Ai-dong, Ni Yong-qiang, Wang Jian-guo. A Method of Eliminate the Abnormal Data in the Real-time Radar Tracking [J]. Fire Control & Command Control, 2004, 29(S1):17-19.
- [8] Zhu Xin-yan, Shi Zhong-ke. An Improved 3σ Method for Outliers Online Pre-processing [J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2008, 28(6):63-65.
- [9] Lu Yuan-lei, He Jia-zhou, An Jin. Research on Rules for Eliminating Outliers and Its Application to Target Prediction [J]. Command Control & Simulation, 2011, 33(4):26-33.
- [10] Li Zhen-Xing, Zhang Hui-Juan. Application of Wavelet Transform in Telemetry Data Outliers Eliminating [J]. Aero Weaponry, 2008, 5(6).