

Detection and mitigation of RFI in SBRS observation data

Zhen-Ping Qiang^{1,2}, Jun Cheng², Zhen-Hong Shang³, Kai-Fan Ji⁴, Fei Dai¹ and Hui Liu^{4*}

¹ College of Big Data and Intelligent Engineering, Southwest Forestry University, Kunming 650224, China

² CAS Key Laboratory of Solar Activity, National Astronomical Observatories, Beijing 100101, China

³ Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650504, China

⁴ Yunnan Observatories, Chinese Academy of Sciences, Kunming 650216, China; liuhui@ynao.ac.cn

Received 2020 November 7; accepted 2021 February 11

Abstract In view of the inconsistency of channel gains and a large amount of interference noise in Solar Broadband Radio Spectrometer (SBRS) observation data, they will seriously affect the analysis of SBRS data. In this paper, a method of Radio Frequency Interference (RFI) detection and mitigation for SBRS observation data is reported. Firstly, the SBRS observation data are preprocessed, a part of the observation data was selected to calculate the mean and variance to achieve the normalization of the entire observation data, which can avoid the influence of strong noise on the normalization result. Furthermore, we proposed an adaptive threshold RFI detection method based on fusion wavelet transform reconstruction and an RFI elimination method based on neighborhood weighted filling. It is worth mentioning that to detect RFI interference signals of different magnitudes, we adopted an iterative approach to the RFI detection and mitigation process. Through qualitative analysis of real observation data and quantitative analysis of simulated data, it is shown that the method proposed in this paper can effectively eliminate RFI in SBRS observation data, and improve the quality of observation data for further scientific analysis.

Key words: spectrographs: SBRS — techniques: image processing and spectroscopic — Sun: radio radiation

1 INTRODUCTION

Radio radiation can reflect the important characteristics and status of the radiator. For example, the radiation frequency of the radiator is closely related to its environmental parameters, and the radio waves in different wavebands can reflect the different characteristics and the status of the radiator. Therefore, by analyzing and studying the spectral characteristics of radio radiation, physical information such as the electron density or magnetic field strength of the radiator can be derived.

Solar radio-frequency radiation is closely related to the density, temperature, and magnetic field of the solar atmosphere. By analyzing and studying the solar radio-frequency radiation, especially the radio spectrum during the solar explosions, the physical mechanism of solar bursts, various physical parameters of source regions and propagation areas, energy release, transfer during bursts, material movement characteristics, and other very important scientific contents can be revealed. For example, in recent years, scientists have shown that the solar

microwave radio radiation flux has a good correlation with the relative number of sunspots; meanwhile, it has a certain relationship with meter wavelength radio bursts, solar flare bursts, X-ray bursts, high-energy electron and proton events.

The Solar Broadband Radio Spectrometer (SBRS) is designed to acquire dynamic spectrograms of solar microwave bursts with characteristics of high time resolution, high-frequency resolution, high sensitivity, and wide frequency coverage in the microwave region (Fu et al. 2004). Its function is to monitor solar radio bursts in the frequency range of 0.7 – 7.6 GHz with time resolution of 1 – 10 ms. It consists of five ‘component spectrometers’ which work in five different wavebands (0.7–1.5, 1.1–2.1, 2.6–3.8, 4.5–7.5, and 5.2–7.6 GHz, respectively). There are three ‘component spectrometers’ located at Huairou Solar Observing Station, as shown in Figure 1. The SBRS monitors the solar radio bursts all day long producing a mass of data for researchers to analyze.

The fine structure of various spectrums can be observed through SBRS, so that it can provide data on the solar corona magnetic field, energy release mechanism,

* Corresponding author

Band1: 1.1-2.1GHz
 Time Cadence: 5ms
 Freq. Res.: 2.788MHz
 Circular Polarization
 Diameter: 7.4m

Band2: 2.6-3.8GHz
 Time Cadence: 8ms
 Freq. Res.: 10MHz
 Circular Polarization
 Diameter: 3.2m

Band3: 5.2-7.6GHz
 Time Cadence: 5ms
 Freq. Res.: 20MHz
 Circular Polarization
 Diameter: 3.2m



Fig. 1 The Equipment of SBRS located at Huairou Solar Observing Station.

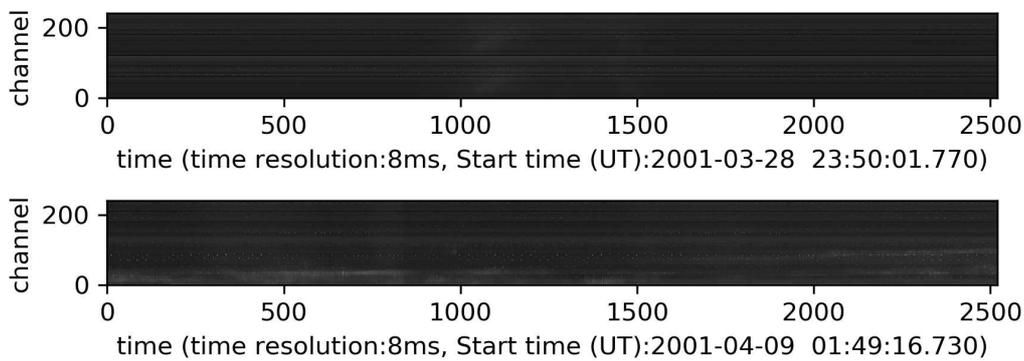


Fig. 2 Uncalibrated radio spectrum data observed by SBRS. The observation data includes a total of 120 channels. In this figure, channel numbers 1 to 120 are channel left rotation data, and 121 to 240 are channel right rotation data.

generation and propagation of high-energy particles, etc. Since 1996, SBRS has accumulated abundant observation data. Then scientists have carried out a lot of research by using these data. In particular, in the research of microwave zebra pattern (Chernov et al. 2008; Tan et al. 2012, 2014), quasi-periodic pulsating structure (Tan 2008; Tan et al. 2010; Wang et al. 2012), microwave bursts (Chen & Yan 2008; Xue et al. 2014), dynamics of flare processes (Chernov et al. 2015), etc., rich scientific research results have been achieved.

As we all know, in radio observation, Radio Frequency Interference (RFI) is inevitable. These interferences are caused by the receiving instrument itself or by other radiation sources in space. Although the designers of the SBRS spectrum analyzer adopt a series of optimization

methods to eliminate interferences of the receiving instrument, such as interference current elimination, common-mode voltages interference elimination, etc., the channel effect of the instrument is relatively stubborn. Meanwhile, the radio signal interference caused by human production and living activities is still inevitable. Therefore, various types of noise still exist in the actual SBRS observation signal. The most obvious noise is the interference noise of nearby routes. Furthermore, during the SBRS observation process, there are different channel effects, abnormal channel observations, channel saturation, and spectrometer overflowed caused by large solar bursts (Yan et al. 2002) which let the data must need normalization. Therefore, some special processes need to be adopted to reduce the impact of interferences on the observation data.

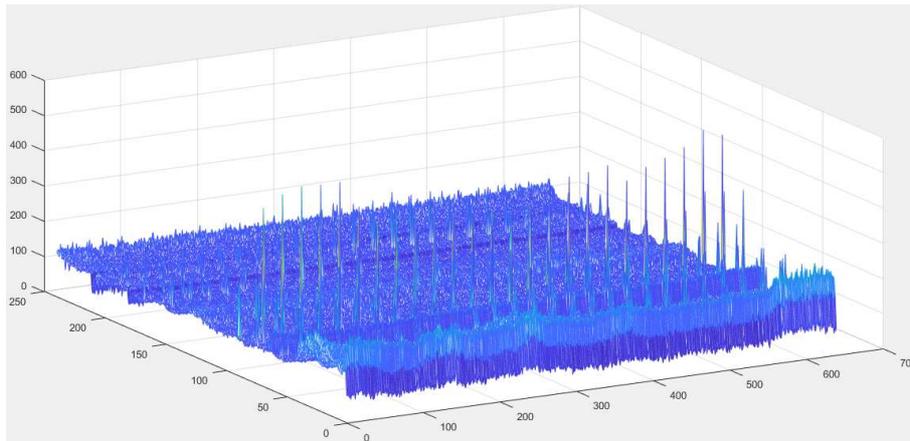


Fig. 3 3D display of uncalibrated observation data (part of the second spectrum of Fig. 2).

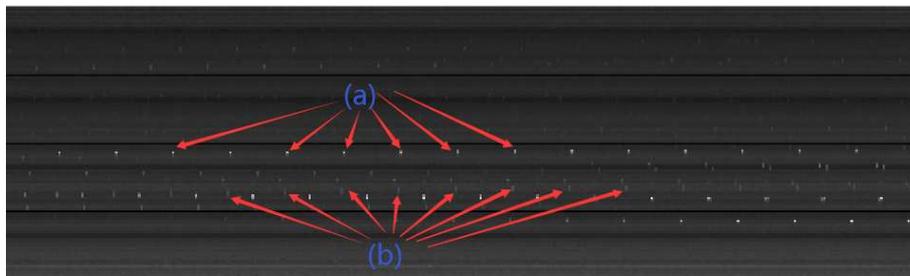


Fig. 4 Two types of RFI in the SBRS observation data (part of the first spectrum in Fig. 2).

Figure 2 shows two sets of data observed by SBRS Band2. Due to channel problems and RFI, it is very difficult to directly observe and analyze the solar burst event in Figure 2. Figure 3 is a three-dimensional display of SBRS observation data. It can be seen that there are a large number of obvious interference signals. Besides, it can be seen that the gains of different channels are very different. We analyzed the RFI in the SBRS observations and found that it mainly includes two types of RFI. One type is a large amount of periodic isolated point noise (They are most likely the stronger airport radar signal interference), indicates in Figure 4(a), and the other type is AC electrical equipment interference, which is mainly represented by a large number of vertical stripes, indicated in Figure 4(b).

These interferences make it difficult to display the information of the solar burst, which will further seriously affect the precise scientific research work and the authenticity of scientific conclusions. Therefore, the detection and mitigation of RFI are the first and key steps in radio observation data processing (Hu et al. 2016). In view of the problems in the SBRS observation data, we proposed a series of processing methods in this paper to solve the channel problems, detect the RFI, and mitigate the RFI.

The organizational structure of this paper is as follows. We will introduce the existing related processing methods

in Section 2. The observation data description will be presented in Section 3. Our methods will be described in Section 4. In Section 5, we will present the process experimental results and analysis of the proposed method for SBRS observation data. We give our discussion and conclusions in Section 6.

2 RELATED WORK

To analyze the SBRS observation data, there are two main processing tasks at this stage. One aspect is channel normalization processing, and the other aspect is denoising processing.

2.1 Channel Normalization

For the problem of different channel gains in radio observation, it can be normalized by directly adjusting the channel difference. Currently, there are three types of solutions commonly used for this problem in radio observations: (1) Normalization processing method for each channel according to the maximum and minimum values; (2) Processing method based on nonlinear transformation and channel means (Yan et al. 2002); (3) Fitting processing method based on the column median (observed values of different channels at the same time).

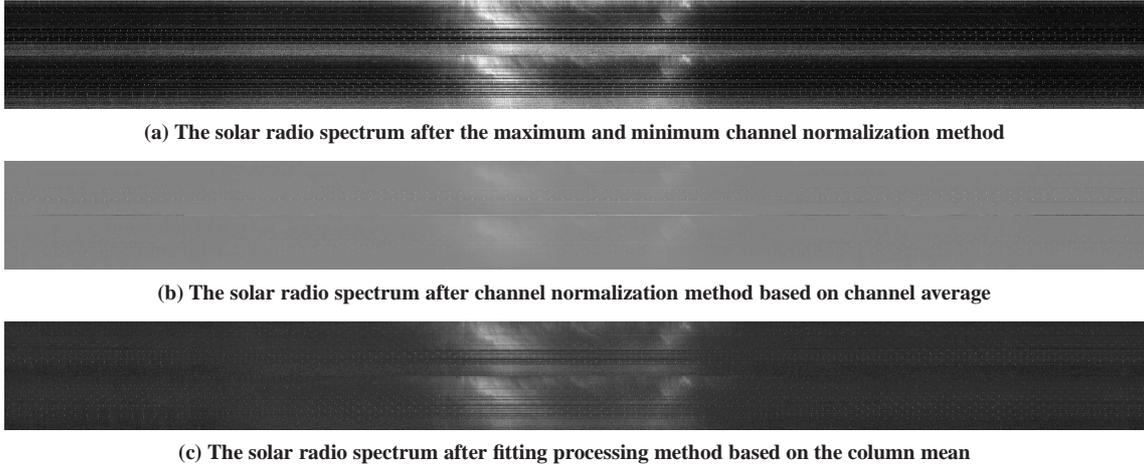


Fig. 5 For the first spectrum in Fig. 2, the results of different channel normalization processing methods.

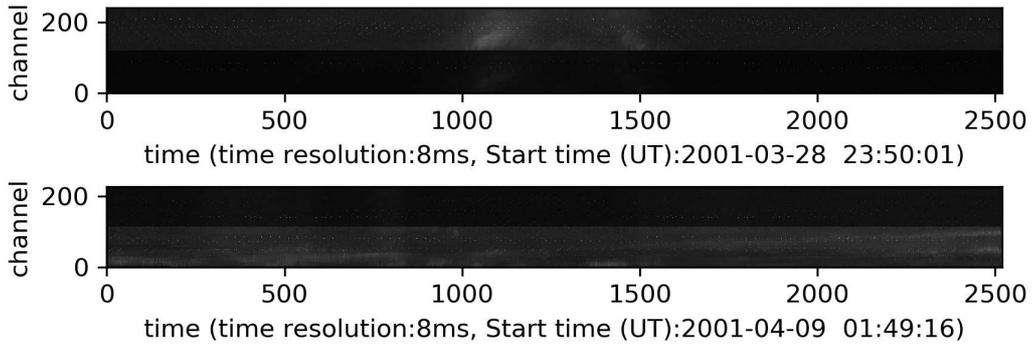


Fig. 6 The processing results of the observation data in Fig. 2 by our proposed preprocessing method.

We set $f(x, y)$ as the original spectrum image, and $g(x, y)$ as the channel normalized spectrum image. Below, we describe the processing of these three types of methods.

2.1.1 Channel normalization method based on the maximum and minimum value of each channel

Let X_i represent the i -th row vector of $f(x, y)$ (observation values of the i -th channel), $X_{i_{\min}}$ represents the minimum value of the i -th row, $X_{i_{\max}}$ represents the maximum value of the i -th row, then the channel normalization method based on the maximum value and minimum value is to calculate each row in $f(x, y)$ according to the following formula:

$$X_i' = (X_i - X_{i_{\min}}) / (X_{i_{\max}} - X_{i_{\min}}). \quad (1)$$

The channel normalized spectrum image $g(x, y)$ can be obtained after processing each line in $f(x, y)$ according to Equation (1).

2.1.2 Channel normalization method based on nonlinear transformation and channel means

The channel normalization method based on nonlinear transformation and mean of the channel includes two steps, the first step is nonlinear transformation, the second step is channel gain normalization.

Let f_{\max} represent the maximum value in $f(x, y)$ and f_{\min} represent the minimum value in $f(x, y)$, then the nonlinear transformation method of the spectrogram is:

$$t(x, y) = a * \ln(1 + 255 * [f(x, y) - b * f_{\min}(x, y)] / [c * f_{\max}(x, y) - d * f_{\min}(x, y)]), \quad (2)$$

where a, b, c, d are adjustment parameters, $a > 1, b > 1, c < 1, d > 1$.

After the non-linear transformation, each channel is divided by its average value to normalize the spectrum image. The formula is as follows:

$$g(x_i, y) = t(x_i, y) / Me_{x_i}, \quad (3)$$

where Me_{x_i} is the average value of the i -th row of the $t(x, y)$.

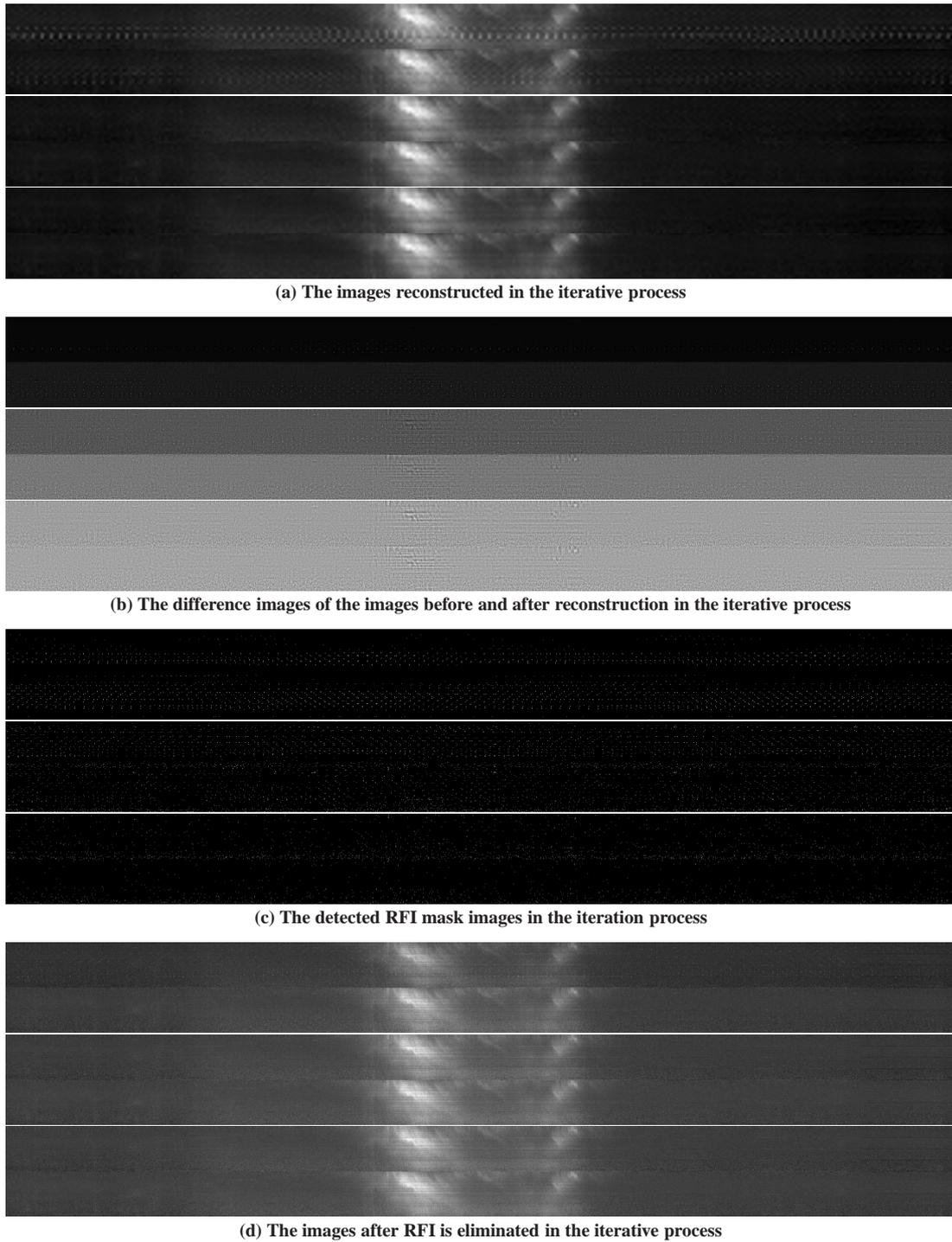


Fig. 7 An example of our algorithm to iterative process the first observation data in Fig. 2.

2.1.3 Channel normalization method based on column median fitting

Although the spectrogram has burst signals, noise, etc., the observed numerical difference is very large, the median of each column should be at the same level. In this way, we can first calculate the median of each column, then we linearly fit the observed vector of each channel to

the calculated median vector, so that all channels can be adjusted to the same gain level.

Let Mc be the vector consisting of the median of each column of the spectrum image, then the linear fitting parameters can be calculated according to the following

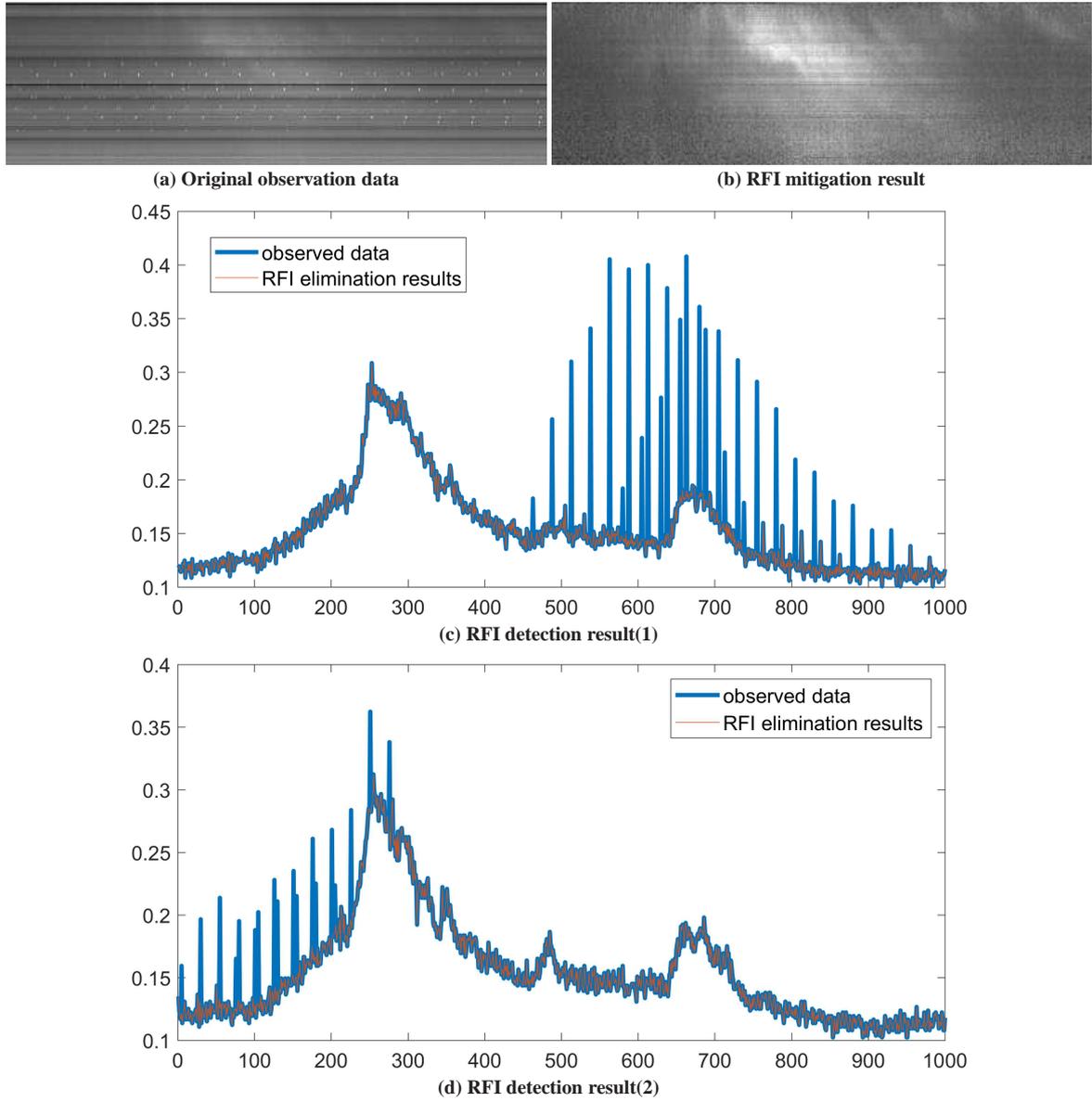


Fig. 8 Comparison of radio burst observation data before and after RFI elimination.

formula:

$$Mc_j = \alpha + \beta * Xi_j ,$$

$$\min l(\alpha, \beta) = \sum_{j=1}^n [Mc_j - (\alpha + \beta * Xi_j)]^2 , \quad (4)$$

where n is the number of signals observed in each channel in the spectrogram (corresponding to the observation duration), Xi_j is the j -th observation value of the i -th channel of the spectrogram. According to Equation (4), the optimal linear fitting parameters α' and β' of each channel can be obtained, then the gain adjustment result of each channel is calculated according to α' and β' :

$$Ti = \alpha' + \beta' * Xi . \quad (5)$$

Further, by using Equation (5), each channel after adjusting the gain is divided by its average value to realize channel normalization processing.

These three types of methods are relatively simple to use. Figure 5 shows the results of these three methods. In the figure, the vertical direction is the channel left rotation and channel right rotation data of 120 channel observation data, and the horizontal direction is the time axis of the observation.

From Figure 5, we can see that normalizing the SBRS observation data directly is susceptible to noise, which makes the further analysis of the observation signal very difficult. In this paper, to mitigate the interference of noise, etc., we propose a method to filter the data of each channel first and then calculate the mean and standard deviation of

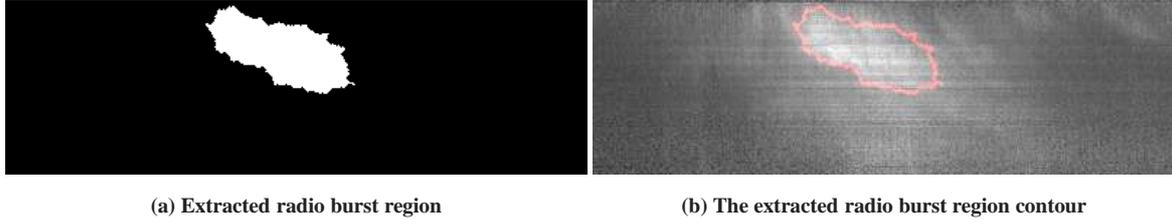


Fig. 9 Radio burst area and contour extraction results after RFI elimination.

Table 1 Algorithm 1. RFI Detection and Mitigation Method for SBRS

Input: spectrogram $f(x,y)$	
Initialize: parameters $n, len, k1, k2, k3, k4, i_n, j = 0$ and matrix M	
Step1: Data preprocessing	
1:	Filter $f(x,y)$: remove bad data to get $f'(x,y)$;
2:	Perform gain adjustment processing on $f'(x,y)$ to obtain the adjusted spectrum image $g(x,y)$;
2.1:	Sort the data of each row in $f'(x,y)$;
2.2:	Select the data which sorted sequence number in the interval $[(len/2) * n, (1 - len/2) * n]$ to calculate the mean and standard deviation;
2.3:	Process the current processing row according to formula (6) to obtain a row of $g(x,y)$;
Step2 & Step3: RFI detection and mitigation	
3:	Based on $g(x,y)$ and iteration number parameter i_n to process(while $j < i_n$):
3.1:	Perform wavelet decomposition on $g(x,y)$ to obtain the decomposition image set;
3.2:	Filter the high-frequency coefficients of the decomposed images according to the adaptive threshold which related to $k1$;
3.3:	Reconstruct the image $g^r(x,y)$ based on the filtered high-frequency coefficient image set;
3.4:	Segment the image $g^r(x,y)$ according to the global adaptive threshold which is related to $k2$ to obtain the mask image M_s (where the element value of 1 indicates that the corresponding position in the image $g^r(x,y)$ has a significant change, otherwise, 0 means that the corresponding position has no significant change);
3.5:	M_s is further morphologically processed to filter the small region labels;
3.6:	Calculate the difference image $d(x,y)$ between the image $g(x,y)$ and the reconstructed image $g^r(x,y)$;
3.7:	Detect the RFI area mask image M_g according to $d(x,y)$, M_s and the global adaptive threshold which is related to $k3$;
3.8:	Detect the RFI area mask image M_l according to $d(x,y)$, M_s and the local adaptive threshold which is related to $k4$;
3.9:	Combine the two-step detection mask of 3.7 and 3.8 to obtain the RFI area mask: $M_j = M_g \cup M_l$;
3.10:	According to M_j and $g(x,y)$, and use formulas (7) and (8) to eliminate RFI to obtain image $g'(x,y)$
3.11:	Let $o(x,y) = g'(x,y)$, $M = M \cup M_j$;
Out: Let $o(x,y) = g(x,y)$ to get a spectrum image that mitigates RFI	

the filtered data to normalize each channel. When filtering data, this method sorts the data of each channel first, and only uses the intermediate region data to calculate

the mean and standard deviation of each channel, which can effectively reduce the impact of strong noise on the normalization results.

2.2 RFI Mitigation

The RFI suppression algorithms in radio observation mainly include time-frequency domain threshold method (Gelu et al. 2007), adaptive noise cancellation method (ANC) (Mitchell & Robertson 2005), spatial domain filter method (Tian et al. 2018) and post-coherence cancellation method (Offringa et al. 2010; Geoffrey 2005). There are also many denoising methods in other research fields, such as wavelet decomposition method, Empirical Mode Decomposition (EMD) method, high-order spectrum algorithm, Singular Value Decomposition (SVD) method, Independent Component Analysis (ICA) method, Wiener filter, Kalman filter, neural network, and deep learning methods.

Among them, the time-frequency domain threshold methods are simple and efficient, so that they are often used to mitigate RFI (Akeret et al. 2017), the equipment such as the LOw-Frequency ARray (LOFAR) uses this method to process RFI (Cendes et al. 2014). The biggest problem with such methods is: how to determine the threshold based on the RFI source and the observed signal? In particular, the solar observation signal is a time-varying signal, and the selection of the threshold is particularly critical. Sometimes the weak instantaneous signal may be regarded as RFI and deleted. Besides, wavelet transform has the ability to characterize the local characteristics of signals in both the time and frequency domains, and the distribution characteristics of signal and noise in the wavelet domain are different, this type of method is often used for RFI removal (Forte et al. 2012). For SBRS observation data, there are RFI mitigation and spectral image enhancement methods based on wavelet transform in the early stage (Xu 2002). However, for RFI from different sources, the difference in time-frequency characteristics makes it very complicated and difficult to analyze all types of RFI signals through wavelet transform methods. Based on the advantages of wavelet analysis, this paper proposes an adaptive RFI elimination method.

Table 2 Description of Synthetic Data

No.	data source file(.fits)	RFI source file(.fits)
1	hsos_sbrs_2600mhz3800mhz_20010328_235001	hsos_sbrs_2600mhz3800mhz_20010325_045952
2	hsos_sbrs_2600mhz3800mhz_20010328_235001	hsos_sbrs_2600mhz3800mhz_20010329_094520
3	hsos_sbrs_2600mhz3800mhz_20010331_234816	hsos_sbrs_2600mhz3800mhz_20010325_045952
4	hsos_sbrs_2600mhz3800mhz_20010324_015703	hsos_sbrs_2600mhz3800mhz_20010120_003815
5	hsos_sbrs_2600mhz3800mhz_20010409_014916	hsos_sbrs_2600mhz3800mhz_20010119_052124
6	hsos_sbrs_2600mhz3800mhz_20010410_045757	hsos_sbrs_2600mhz3800mhz_20010202_235138

Table 3 Qualitative Analysis of Synthetic Data

No.	PSNR		MSE		SSIM		Precision		Accuracy	
	ATSM	ours	ATSM	ours	ATSM	ours	ATSM	ours	ATSM	ours
1	26.2091	28.1741	0.0024	0.0015	0.9999	0.9998	1	1	0.9830	0.9962
2	24.6470	32.9386	0.0034	0.0005	0.9999	0.9999	1	1	0.9809	0.9954
3	19.0269	28.3681	0.0125	0.0015	0.9996	0.9998	1	1	0.9922	0.9960
4	20.8564	33.0205	0.0082	0.0005	0.9998	0.9999	0.9987	1	0.9998	0.9991
5	23.4827	26.3919	0.0045	0.0023	0.9998	0.9997	1	0.9665	0.9988	0.9982
6	23.8984	33.6278	0.0041	0.0004	0.9998	0.9999	0.9756	0.9916	0.9869	0.9970
AVG	23.0201	30.4202	0.0059	0.0011	0.9998	0.9998	0.9957	0.99302	0.9903	0.9970

Wavelet transform (WT) is a signal transform analysis method, which inherits and develops the idea of localization of short-time Fourier transform, providing a “time-frequency” window that changes with frequency. It is an ideal tool for signal time-frequency analysis and processing. RFI filtering based on wavelet transform is to construct different filtering rules based on the observation target signal and interference signal having different morphological representations at different scales of wavelet transform to filter the interference signal. The essence is to reduce or even eliminate the wavelet coefficients corresponding to the noise, while keeping the coefficients corresponding to the effective signal to the maximum extent, then reconstruct the signal based on the processed wavelet coefficients. Therefore, traditional wavelet denoising includes the following three steps:

(1) Wavelet decomposition of the observed signal S : select a wavelet basis function and the wavelet decomposition level N , then calculate the decomposition coefficients from the first layer to the N -th layer of the signal S ;

(2) Filtering the high-frequency coefficients obtained by decomposition: For each layer obtained by decomposition, select a threshold, and filter the high-frequency coefficients of this layer according to the selected threshold;

(3) Wavelet reconstruction: According to the low-frequency coefficients of the N -th layer and the filtered high-frequency coefficients from the first layer to the N -th layer, the resulted signal S' of wavelet denoising is reconstructed iteratively.

Considering that the result of wavelet denoising depends on the selection of the filtering threshold, the RFI in the SBRS observation data is mainly vertical stripes and periodic outliers, and their observation values are very

different. So that we propose an iterative adaptive threshold RFI detecting and mitigation method based on wavelet transform reconstruction.

3 OBSERVATIONS

This article deals with the data observed by the SBRS Band2 radio spectrum analyzer, which has 120 channels, each channel corresponds to a frequency, from 2.6 GHz to 3.8 GHz, the frequency interval (frequency resolution) is 10MHz, and the highest time resolution is 8 milliseconds. The second-level file records two (corresponding to left and right polarization) 0.2-second integral observations per channel, and the millisecond-level file records two (also corresponding to left and right polarization) 0.05-second integral observations per channel.

At noon every day of the observation, a 0.2-second calibration file is recorded for the day, the values of the quiet sun, background sun, noise, and terminal of 120 channels are recorded respectively, and they are also divided into left and right rotations. These data are prepared for future calibration work. The daily observation also records the left and right rotations gain characteristic curves of the 120 channels. All recorded observations are only the readings of the receiver, not the radio radiation flux. To convert the readings into radio radiation flux, calibration is required. The data processed by the proposed algorithm are all spectrum images after calibration. In the experiments in this article, we have used the data after calibration in 2001 as the processing data.

4 OUR METHOD

The method that we proposed includes three steps. The first step is the preprocessing of SBRS observation data,

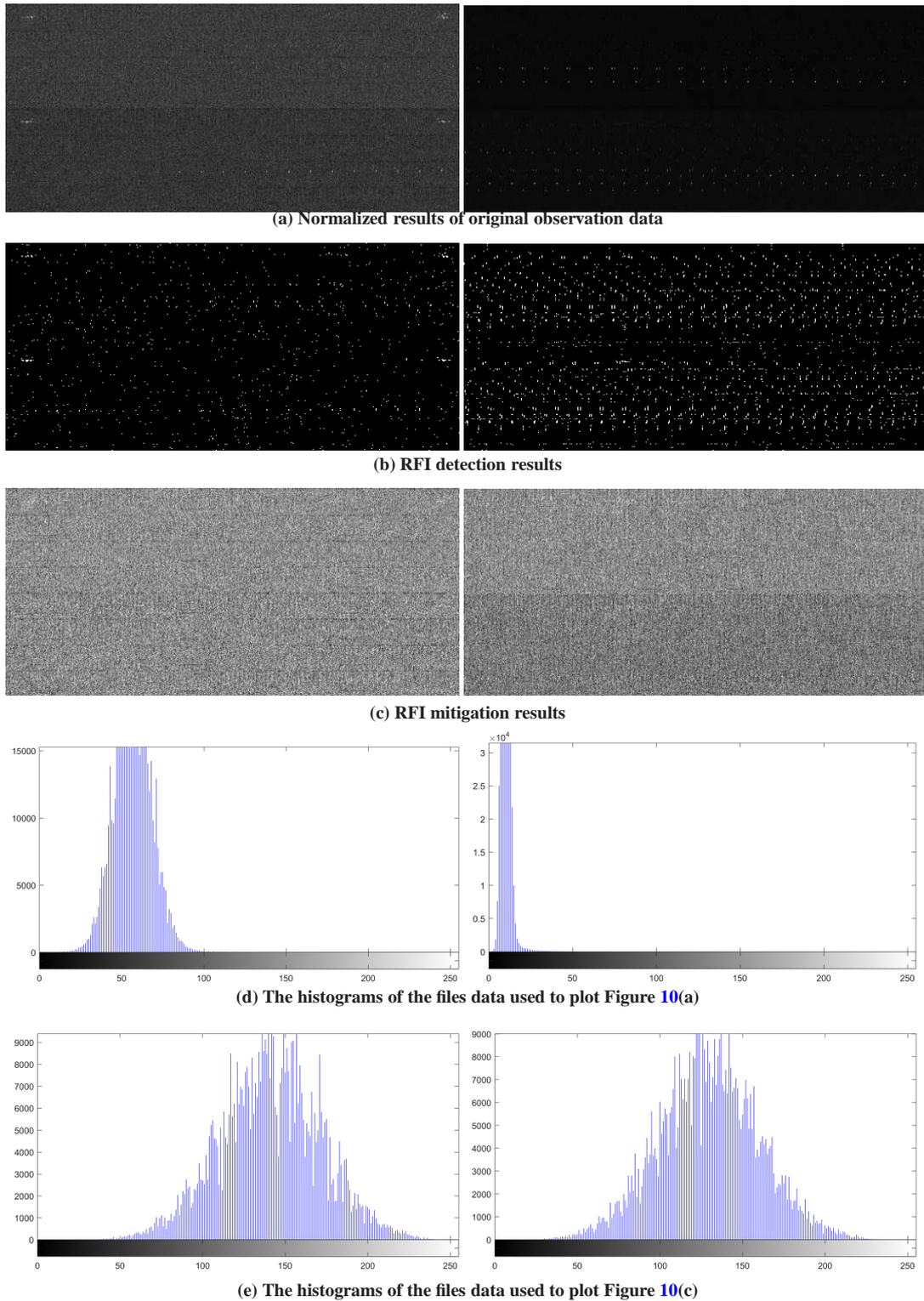


Fig. 10 The experimental results of RFI detection and elimination for two sets of data without outbreaks.

the second step is the detection of RFI, and the third step is the mitigation of RFI.

4.1 Data Preprocessing

The RFIs have high values compare to the observed burst signal and the observation signal without bursts. We have calculated the numerical distribution of a large number of

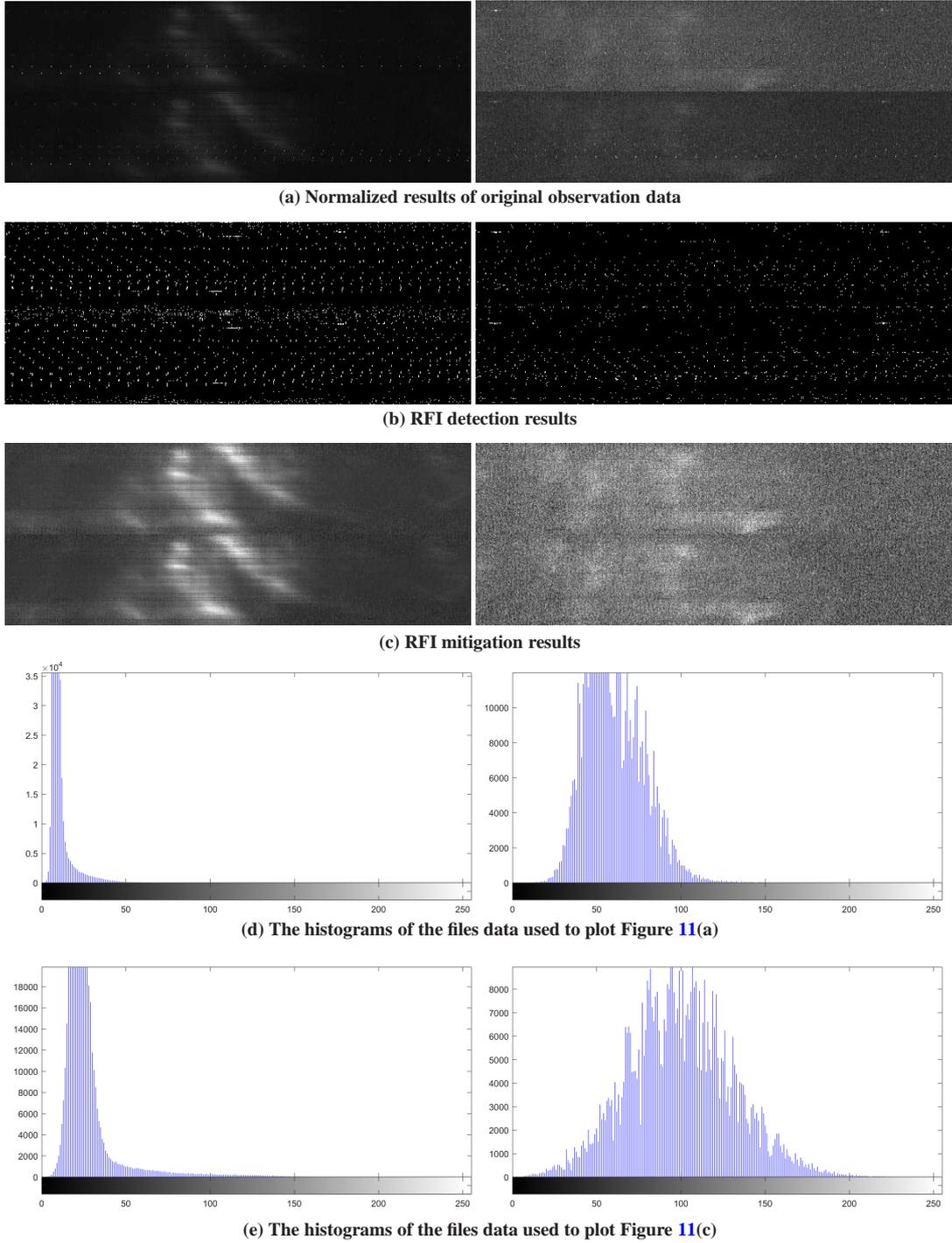


Fig. 11 The experimental results of RFI detection and elimination for two sets of data with solar radio bursts.

observational data, for each observation file (For SBRS Band 2, the time resolution is 8 ms; and X size of image (time) is 2520). In our statistics, the proportion of RFI signal in the whole observation channel signal will not exceed ten percent. According to this, we sort the observation data of each channel and then select the data in the middle part to calculate the mean and variance for channel gain adjustment processing. The adjustment

processing formula is:

$$g(x_i, y) = (f(x_i, y) - \mu_{i,p}) / \sigma_{i,p}, \quad (6)$$

where $\mu_{i,p}$ and $\sigma_{i,p}$ are the mean value and standard deviation of the selected data in the i -th row (the i -th channel) of the spectrum image $f(x, y)$.

Figure 6 shows the processing results of Figure 2 observation data by the channel gain adjustment preprocessing method. After data preprocessing, the RFI in the

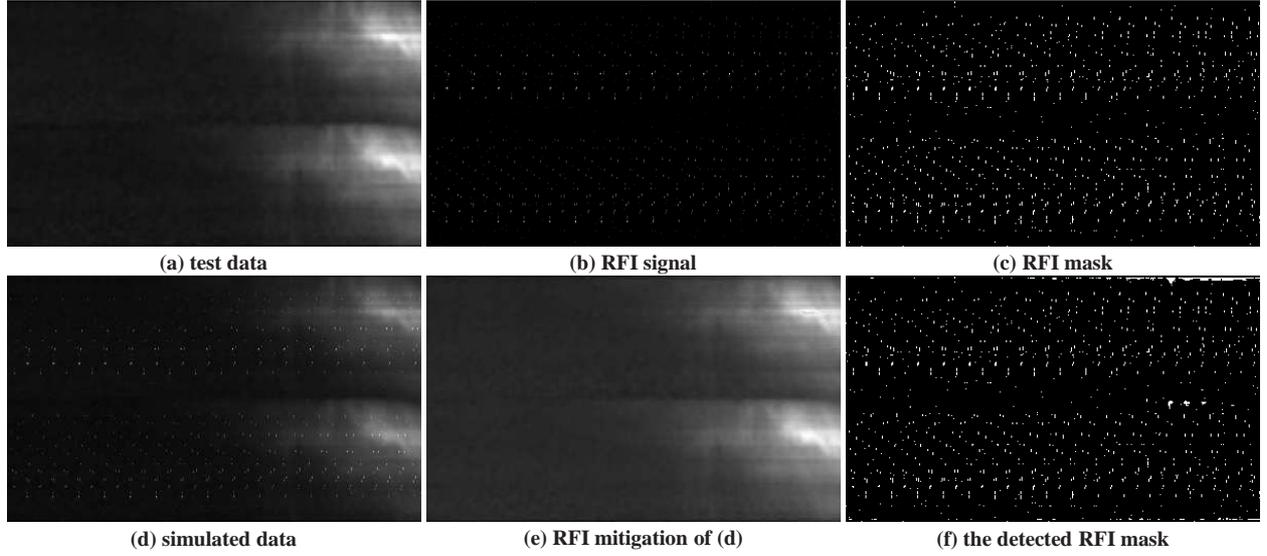


Fig. 12 An example of RFI mitigation to the simulation data.

spectrogram is more significant, which is the basis for RFI detection. We find that in Figure 6, the left and right polarization images have different brightness levels. The RFI values in the left and right polarization spectra are different, so we will take further RFI mitigation steps to solve this problem.

4.2 RFI Detection

In this paper, we mainly analyze vertical stripes and periodic outlier noise. Based on this, we propose an adaptive threshold RFI detection method based on fusion wavelet transform reconstruction. Different from the traditional wavelet denoising method, the main idea of our method is to use the wavelet transform to filter the high-frequency information of the spectrum image first, and then perform the regional type segmentation on the filtered image through the global adaptive threshold (divided into burst areas and non-burst areas), and further perform RFI detection on different types of areas based on the regional adaptive threshold.

The process of RFI detection is as follows:

- (1) The low-frequency component image f_l of the original signal f is obtained by wavelet decomposition and high-frequency coefficient filtering. Then calculate the variance and mean value of f_l ;
- (2) Based on the principle of the triple standard deviation method, f_l is divided into burst areas and non-burst areas to obtain representation mask image M of different areas, and perform morphological processing on M ;
- (3) Calculate the difference image D_f between the original signal f and the low-frequency component f_l ;

- (4) According to mask image M , calculate the variances and means of the outbreak areas and non-outburst areas of the difference image D_f ;

- (5) Based on the principle of the triple standard deviation method, detect the RFI signals in the outbreak area and non-outbreak area respectively.

At the same time, we found that the values in the vertical stripe interference in the SBRS observation data are different. It often happens that the observed values of the middle pixels in the vertical stripe are large, and the observation values of the two ends in the vertical stripe are small. Thus it is difficult to detect all RFI signals by performing the method proposed above at once. So we execute this method iteratively to detect RFI in the image until most of the RFI signals can be detected.

4.3 RFI Mitigation

The accurate detection of RFI is the key to RFI mitigation, and the RFI in the SBRS observation data is relatively isolated noise points (or thin vertical stripes). So we use a weighted filling method based on the neighborhood of each RFI pixel to eliminate RFI.

Let M be the mask matrix of RFI, the element value of which is 1 means the observation value of the corresponding position in the spectrogram is RFI, and 0 means the normal observation value. Suppose G_k is a Gaussian kernel whose length and width are both $2 * k + 1$, then the RFI elimination weight matrix is:

$$W_{p(i,j)}^k = (1 - M_{p(i,j)}^k) \otimes \left[G_k + \frac{\sum (G_k \otimes M_{p(i,j)}^k)}{\sum M_{p(i,j)}^k} \right], \quad (7)$$

where $M_{p(i,j)}^k$ is a patch with $p(i, j)$ as the center, and both length and width are $2 * k + 1$, \otimes represents the element-wise multiplication of the corresponding matrix.

According to formula (7), we calculate the weight matrix $M_{p(i,j)}^k$ for each detected RFI pixel at point $p(i,j)$ in the spectrogram, then multiply the matrix $M_{p(i,j)}^k$ and spectrum image block $g_{p(i,j)}^k$ element by element, finally the matrix elements obtained by the multiplication are summed to fill $N_{p(i,j)}$. The calculation formula is as follows:

$$g'(i,j) = \sum (W_{p(i,j)}^k \otimes g_{p(i,j)}^k). \quad (8)$$

In this way, the whole process of our proposed algorithm is shown in Table 1.

In our proposed method, four threshold-related parameters $k1$, $k2$, $k3$ and $k4$ are used. In the experiment, their values are all equal to 3, the adaptive thresholds are all taken to the interval determined by the mean μ and 3 times the standard deviation $3*\sigma$. The parameter n is the number of signals observed in each channel in the spectrogram. The parameter len can effectively avoid the influence of strong RFI on the normalized results. We set $len = 0.6$ in the experiment, that is, we calculate the mean and standard deviation through the middle 60% of the data after sorting. Through this processing method, RFI detection can be performed on areas with and without solar radio bursts, avoiding the interference of different observation levels on RFI detection. The parameter i_n in the algorithm is the number of iterations, that is after the RFI signal is detected and mitigated, the algorithm must continue to iteratively detect and eliminate RFI, so that different levels of RFI in the SBRS observation data can be eliminated. In the experiment, the value of i_n is 3.

In Figure 7 we show an example of this method. From Figure 7 we can see that with the iterative processing, the RFI of different observation levels is gradually detected and alleviated, and the quality of the observation data has also been gradually improved. To compare the experimental results more clearly, in Figure 8, we show the comparison of the solar radio burst region in Figure 7 before and after RFI elimination.

At the same time, in Figure 8, we compare some frequency observation curves in the solar radio burst area before and after RFI eliminating. The blue line is the original observation data after normalization, the orange line is the data after eliminating RFI. It is obvious that RFI can be effectively eliminated by the method in this paper.

After the RFI is eliminated, the radio burst area in the observation data can be extracted very easily, as shown in Figure 9, which is extracted by the graph cut method for Figure 8(b), the contour of the radio burst area can be extracted conveniently, as shown in Figure 9(b). The extraction of radio burst region and contour is often the basis of solar radio burst analysis. Therefore, it is very important to eliminate RFI in solar radio research.

5 EXPERIMENTAL ANALYSIS

To verify the effectiveness of our proposed method, we have completed experiments on observation data and simulated data respectively.

5.1 Observation Data Experiment

We have completed RFI detection and mitigation experiments for images with and without solar radio bursts.

Figure 10 shows the experimental results of RFI detection and elimination for two sets of data without outbreaks. It can be seen that after the processing of our proposed method, the contrast of the observation data is significantly improved. Furthermore, in Figure 10(c) and Figure 10(d) we show the histograms before and after the data processing. After processing by our proposed method, the distribution of the observation data is closer to the Gaussian distribution, which is very consistent with the characteristics of the random noise distribution.

Similarly, we have completed RFI detection and elimination experiments on the two sets of data with solar radio bursts, the results are shown in Figure 11. The experimental results show that after the RFI mitigation by the proposed method, the spectrum distribution and time range of burst emission can be studied. The histogram distribution of the processed spectrogram is more uniform, indicating that the quality of the processed data has been significantly improved.

5.2 Simulated Data Experiment

To verify the effectiveness of our proposed method, we conduct a quantitative analysis of the method through simulated data. First of all, we use the method in this paper to mitigate the RFI signal for the observation data and smooth the processed images as the test data (as shown in Fig. 12(a)). Furthermore, to avoid the influence of the burst signal on the RFI detection, we detect the RFI signal in the observation data without solar bursts (Fig. 12(b), which shows the RFI signal image and Fig. 12(c), which shows the RFI mask image). Finally, the detected RFI signal is added to the test data as simulated data (as shown in Fig. 12(d)). The result after RFI mitigation of Figure 12(d) is shown in Figure 12(e), and the detected RFI mask is shown in Figure 12(f).

For different simulated data, we used the method in this article to detect and mitigate RFI. The Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), and Structural Similarity Index Measure (SSIM) of the simulation data and the RFI mitigation result data were calculated. Further, we calculated the precision rate and accuracy rate of the detection of the added RFI signal. Precision = $TP/(TP + FP)$, where TP represents the number of pixels detected as RFI, which is the added RFI;

FP represents the number of pixels detected as RFI, but not the added RFI. $Accuracy = (TP + TN)/P_{all}$, where TN represents the number of pixels detected as non-RFI, which is not the added RFI; P_{all} represents the number of all pixels in the test data.

In Table 2, we give the composition files of six sets of synthesized data. The selection of observation data in the experiment is random, including the cases with and without solar eruptions. We use two sets of observation data to synthesize a set of simulation data. Firstly, the RFI of observation data S (corresponding to the data source file column in Table 2) is mitigated and then used as signal data D_s before synthesis. The RFI of other observation data S' (corresponding to the data source file column in Table 2) is extracted as interference signal $RFI_{S'}$; further add $RFI_{S'}$ to D_s as the simulation data S_{sim} to be processed. The algorithm proposed in this manuscript is applied to S_{sim} performed RFI elimination to obtain the data D'_{sim} after RFI mitigation. Finally, calculate the PSNR, SSIM, and MSE of D_s and D'_{sim} to verify the effectiveness of this algorithm. As shown in Table 3, the quantitative comparison results of six sets of data are given. Compared with the Adaptive Threshold Selection Method (ATSM) using the principle of Stein's Unbiased Risk Estimate (SURE), our method is better than the adaptive threshold denoising algorithm in PSNR and MSE. The mean value of SSIM is the same as that of ATSM, they are very high. This is mainly due to the very low content of RFI signals in the entire observation data. Because there are fewer real RFI signals detected by ATSM, the probability of error is lower, ATSM has better RFI detection precision. Correspondingly, our method has achieved good results. Therefore, the experimental comparison of synthetic data shows that the method we proposed can recover the original data well for the simulation data, and can detect the added RFI signal very accurately.

6 CONCLUSIONS

In this paper, we propose a method for RFI detection and mitigation in SBRS observation data. According to the characteristics of SBRS observation data, this method uses three steps which include data preprocessing, RFI detection, and RFI mitigation to perform RFI elimination on the observed data. Moreover, qualitative and quantitative analysis has been carried out on actual observation data and simulation data. The experimental results show that the method proposed in this paper can effectively eliminate RFI in SBRS observation data, and significantly improve the quality of observation data. This provides high-quality spectrum data for further scientific analysis on SBRS observation data.

Acknowledgements This work was funded by the Open Research Program of CAS Key Laboratory of Solar Activity, National Astronomical Observatories (KLSA201909), the National Natural Science Foundation of China (Nos. 11773072 and 11873027), and Yunnan Fundamental Research Projects (202001AT070135).

References

- Akeret, J., Chang, C., Lucchi, A., & Refregier, A. 2017, *Astronomy and Computing*, 18, 35
- Cendes, Y., Wijers, R. A. M. J., Swinbank, J. D., et al. 2014, arXiv e-prints, arXiv:1412.3986
- Chen, B., & Yan, Y. 2008, *ApJ*, 689, 1412
- Chernov, G., Fomichev, V., Tan, B., et al. 2015, *Sol. Phys.*, 290, 95
- Chernov, G. P., Yan, Y., Fu, Q., Tan, C., & Wang, S. 2008, *Sol. Phys.*, 250, 115
- Forte, G., Camps, A., Tarongi, J., & Vall-Ilossera, M. 2012, in 2012 IEEE International Geoscience and Remote Sensing Symposium, 1069
- Fu, Q., Ji, H., Qin, Z., et al. 2004, *Sol. Phys.*, 222, 167
- Gelu, M., Dale, E., & Liu, Z. W. 2007, *The Publications of the Astronomical Society of the Pacific*, 119, 805
- Geoffrey, C. 2005, *Radio Science*, 40, 1
- Hu, J.-F., Cao, X.-W., Chen, L., & You, B. 2016, *RAA (Research in Astronomy and Astrophysics)*, 16, 136
- Mitchell, D. A., & Robertson, J. G. 2005, *Radio Science*, 40, RS5S11
- Offringa, A. R., de Bruyn, A. G., Biehl, M., et al. 2010, *MNRAS*, 405, 155
- Tan, B. 2008, *Sol. Phys.*, 253, 117
- Tan, B. L., Tan, C. M., Zhang, Y., Mészárosóvá, H., & Karlický, M. 2014, *ApJ*, 780, 386
- Tan, B., Yan, Y., Tan, C., Sych, R., & Gao, G. 2012, *ApJ*, 744, 166
- Tan, B., Zhang, Y., Tan, C., & Liu, Y. 2010, *ApJ*, 723, 25
- Tian, Z., Wen, B., Jin, L., & Tian, Y. 2018, *IEEE Geoscience and Remote Sensing Letters*, 15, 102
- Wang, R., Tan, B., Tan, C., & Yan, Y. 2012, *Sol. Phys.*, 278, 411
- Xu, L. 2002, *Wavelet Analysis and Its Application to the Processing of Solar Radio Data Observed (Xi'an University of Electronic Science and Technology)*
- Xue, Y., Wang, S. J., & Yan, Y. H. 2014, in In 2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS) (IEEE)
- Yan, Y., Tan, C., Xu, L., et al. 2002, *Science in China A: Mathematics*, 45, 89