# Perceptual Spectrum Waterfall of Pattern Shape Recognition Algorithm

Wencheng Chen*, Xiangming Wen*
*School of Information and Communication Engineering,
Beijing University of Posts and Telecommunications, Beijing, 100876, China
**Email: {wenchengchen, xiangmw}@bupt.edu.cn**

*Abstract*— **This article presents a novel recognizable spectrum sensing algorithm. This algorithm is the perception spectrum waterfall features for image pattern recognition algorithm, it referred to the pattern shape recognition algorithm. This algorithms has occurred band spectrum waterfall uncertain state of the spectrum width and the length of time by machine self-learning, and to establish the pattern shape recognition library. The experiments showed that is algorithms has higher recognition rate of the primary user than the image pattern recognition algorithm, and more applicable spectrum sensing.**

***Keywords*— Spectrum Waterfall; Pattern Recognition; Spectrum Division, Spectrum Fracture Reconnection, Spectrum Deletion Completion**

## I. Introduction

All the sensing methods are predicting the future spectrum occupancy by real-time sensing current spectrum status. And the main goal of spectrum sensing is to provide as many as reliable and stable spectrum resources for Sensing Users (SU) by guaranteeing a reliable communication for Primary Users(PU). Cognitive radio (CR) is designed for the SUs opportunistically access to the unused spectrum without causing interference to the PUs [1]. Effective spectrum sensing can not only find the idle spectrum segments for SUs, but also reduce the interference to the PUs [2]. Usually, there are three sensing methods presented to settle it, i.e., energy detection [3], matched filtering (coherent) detection [4], and cyclostationary feature detection [5].This article has different idea to presents a novel way for sensing spectrum status that is pattern shape recognition algorithms by spectrum waterfall.

## II. Data Processing Flow

All spectrum data status and relationship as: time, space, frequency and power can be presented from image vision by analyzing the spectrum waterfall plot. The ability of predicting spectrum can be enhanced by analyzing the pattern of the spectrum waterfall plot and by building the machine learning library.

There are two key points of sampling the spectrum sensing waterfall plot: (1) Controlling sampling resolution: choosing the time unit and frequency unit. (2) Controlling range of sampling image: Select perception waterfall spectrum, time width on spectrum sensing holes forecast.

Based on the above points, we modify and improve the pattern recognition algorithm for images to the algorithm for analyzing the spectrum waterfall plot. Figure1 gives the processing flow of our improved pattern recognition algorithm for spectrum waterfall plot.

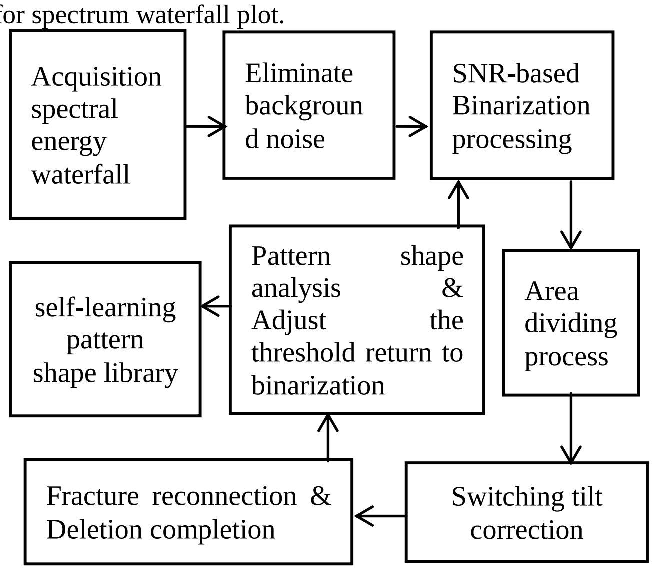


**Figure 1.** Data processing flow

## III. Algorithm Analysis

### A. *Sampling Spectrum Waterfall Plot*

The data used in this paper is from the test receiver by real-time sampling process. Then the data is transformed to the spectrum waterfall plot.

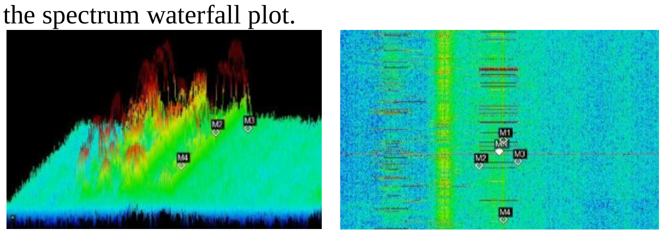

**Figure 2.** 3D & 2D Spectrum energy waterfall

### *B. Binary Processing of Spectrum Waterfall Plot*

**1)** ***Noise Reduction process:*** Signal-to-noise ratio (SNR) of the sampling image can be increased by removing noise of the energy waterfall plot in order to strengthen the recognition accuracy. In fact, increasing the image SNR is to increase the resolution of some frequency points. Normally, methods of reducing noise are: Median filtering, Gaussian filtering, Wiener filter and wavelet de-noising. In this paper, we use the classical Gaussian filtering to reduce the noise. As the color elements in the waterfall plot are according to the temperature and the energy (shown in Figure 2) which have the RGB color model property as usual colorful images, the RGB color system luminance meter equation is used to get the grayscale image (see Figure 3).

$$I(i,j) = 0.299R(i,j) + 0.587G(i,j) + 0.114B(i,j) \quad (1)$$

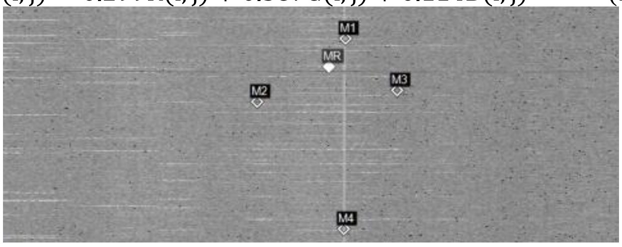


**Figure 3.** Spectrum waterfall after luminance equation

**2)** ***Binarization Threshold:*** The grayscale image from Eq. (1) still needs binary transform which is to set the grayscale maximum and minimum threshold to avoid the value that even beyond the maximum and minimum value. These thresholds are concerned to the PU tolerance and the SU initial control power. The key point here is to pick the threshold values which are classified as static threshold and dynamic threshold. Although there are adaptive threshold methods based on windows that can improve the binarization image details, the noise could be introduced when these methods are applied to image pattern recognition especially the noise that from the test receiver or the computer we use. In our work, we use the maximum similarity thresholding[6] to determine the overall binarization threshold.

Also we should notice the binarization value is related to the redundancy that the PU can bear. As the requirement for the SU's communication is to guarantee the PU's communication, the control amount for the SU signal strength is related to the SNR intensity at former time. Normally, this quantized value is layered and changeable. The binarization spectrum waterfall plots in different layer are different which means the corresponding environment for SU is different. The binary value can be between 0.0 and 1.0 which can lead to different waterfall plot (as Figure 4).

$$f(\boldsymbol{i},\boldsymbol{j}) = \begin{cases} 0.0; & I(\boldsymbol{i},\boldsymbol{j}) \le T \\ 1.0; & I(\boldsymbol{i},\boldsymbol{j}) > T \end{cases} \quad 0 \le T \le 1; \quad (2)$$

The $f(\boldsymbol{i},\boldsymbol{j})$ is the value after the binarization of point at $(\boldsymbol{i},\boldsymbol{j})$ and $I(\boldsymbol{i},\boldsymbol{j})$ is the value at the same point before binarization.

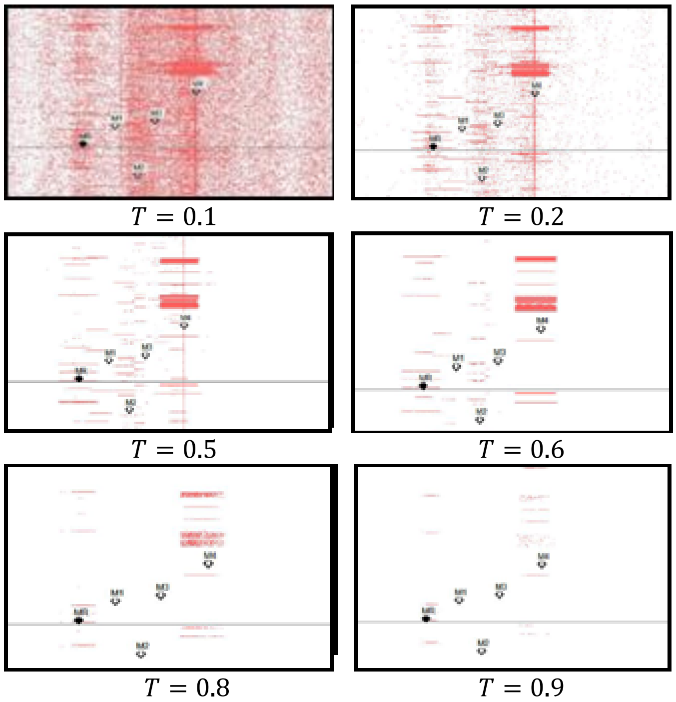


**Figure 4.** Waterfall of different threshold

### *C. Area Division process*

In pattern recognition system, there are many methods for page segmentation. E.g. projection method can get good effect for high quality signal input: firstly calculate all effective pixels in each row and column to build a vector then thresholding this vector can get the property of rows and columns of original image. However, this method is based on statistics so performance for anti-noise is quite bad. Moreover, for the image that exceeds the tilt tolerance, this methods require a preprocess step called tilt correction.

**1) Tilt Process for Waterfall Plot:** The sensing image is a non-continuously and tilt one as there is a delay caused by analyzing, predicting, processing and determining between sensing spectrum status and the beginning of SU (see Figure 5). In addition, the switching times between spectrum sensing and spectrum holes predicting will also affect the sensing redundancy time and the waterfall plot will have the tilt property. Common pattern recognition algorithm will firstly correct the tilt then process area division. If an algorithm wants to jump over the correction step and process area division directly, two factors must be concerned: (1) the sampled waterfall plot must have a significant time line features, for those have no line features, we can assume it is approximate rectilinear. (2) As long as the tilt angle of the sampling plot is within a tolerable range (the tilt is caused by the time delay among the sensing time, predicting time and SU's using time), we can correct precisely by correcting the datum feature.

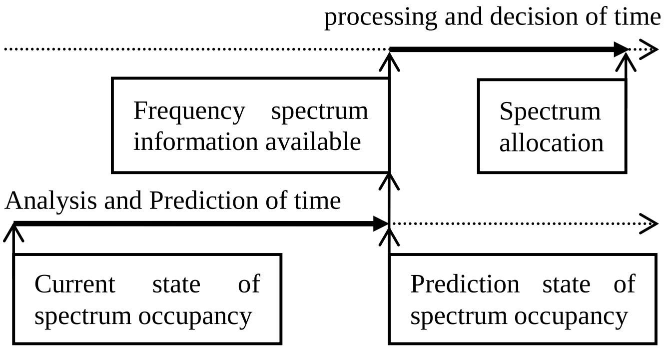


**Figure 5.** Analysis and Prediction of time and processing and decision time due to the spectrum waterfall tilt

By utilizing the area information, we can correct the tilt precisely. The slope can be calculated by linear regression for the frequency signal of each time line axis. Then the arctangent value of the slope is the tilt angle we need. And for the different regression results, the correlation coefficient R's largest slope will be chosen as the final result in order to get the corresponding regression result. Then use the following equations to rotate every pixel in the connected domain contour.

$$x' = (x - x_m) * \cos(\mathbf{Angle}) + (y - y_m) * \sin(\mathbf{Angle}) + x_m \quad (3)$$

$$y' = (y - y_m) * \cos(\mathbf{Angle}) + (x - x_m) * \sin(\mathbf{Angle}) + y_m \quad (4)$$

The $\mathbf{Angle}$ is the arctangent value of the slope, $x_m$, $y_m$ is the central point of the connected domain, $x$、$y$ is the pixel of the points in the connected domain.

**2)** ***Spectrum waterfall area dividing process:*** The so-called contour is the pixel set fit the following relationship in the binarizaion image:

$$\mathbf{Edge} = \{x_{i,j} | x_{i,j} = 0 \wedge x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1} > 0 \quad (5)$$

The $\mathbf{x_{i,j}}$ means the pixel value of the **i**th column and **j**th row.

It is easy to get an image formed by the edge pixel points by edge detection using the above standard. Then we can feature and divide the connected domain to get a rough PU contour information.

A connected domain $C$ is defined as a pixel set formed by several neighbor pixel containing same pixel value. The process of distinguishing different connected domain is called connected component labeling. This process, in fact, is to use the same grade to label the pixel belonging to the same connected domain(method of Region Growing).

Step 1: Input the unlabeled Bitmap, initialize a label matrix with the same size with the input Bitmap, a Queue and an Index;

Step 2: Scan the Bitmap by from left to right element by element and from top to down line by line, once an unlabeled foreground pixel $\boldsymbol{p_0}$ is found, $\text{index} = \text{index} + 1$ and mark in the Labelmap (assignment value is Index);

Step 3: Scan $\boldsymbol{p_0}$'s neighbor pixels, if there is an unlabeled foreground pixel, we mark it in the Labelmap and put it into the Queue as the seed for region growing, if no, continue scanning next unlabeled foreground pixel;

Step 4: Pick seed $p_i$, scan the neighbour points, scan the eight neighbour points of $p_i$, if there is an unlabelled foreground pixel, label it in the Labelmap and put it in the Queue. If not, take next seed point $p_{i+1}$;

Step 5: Repeat the Step 4 until the Queue is empty, then a connected domain is labelled in the Labelmap;

Step 6: Go to Step 2, until the whole image is scanned and we get the label matrix Labelmap and Index of connected domain.

In the worst case, the algorithm needs to scan every pixels with a searching of eight neighbours and the complexity is O(n). After all, by the label process of the connected domain, a set of several connected domains is get. However, in order to separate the connected domain formed by the regional edge from this set, several features needs to be analysed.

Firstly, we give some definitions:

Definition 1: Define the minimum enclosing rectangle as a rectangle of minimum area with all pixel of a connected domain in it: $\text{Rect}(x, y, \text{width}, \text{height})$. The $x, y$ is the coordinate of the left corner of the rectangle, the $\text{width}, \text{height}$ is width and height of rectangle.

Definition 2: Define the coordinate $x_m$、$y_m$ in the connect domain:

$$x_m = x + \frac{1}{2} \times \text{width} \quad (6)$$

$$y_m = y + \frac{1}{2} \times \text{height} \quad (7)$$

Definition 3: Define the area of the closed geometry of $\mathbf{S}$ as the area of the pixels in the connected domain.

Definition 4: Define $\boldsymbol{R_{wh} = width/height}$ (8)

Definition 5: Define the minimum area of the closed rectangle: $S_r = \text{width} \times \text{height}$ (9)

Definition 6: Define the fullness of the connected domain:

$$R_{fill} = S/S_r \quad (10)$$

Based on the above definitions, a connected domain set $\{\boldsymbol{c_i}\}$ is the connected domain formed by only edges of some selected special threshold. This set is normally a set of the specific time and frequency in a communication system. By the binarization analysis of the experiments, once choosing （$\boldsymbol{S_r} > 20 \wedge 0.3 < \boldsymbol{R_{wh}} < 1 \wedge \boldsymbol{R_{fill}} < 0.7$）as the standard, a system set can be judged. This set is different with the previous one, which is related to the SNR tolerance of different PU communication system.

However, the algorithm above is not only suitable to the case with high performance of the sensing receiver and sensing frequency waterfall plot. For the low quality waterfall plot, it cannot detect the change of the whole system and will have bad effects of building the reference library of the machine self-learning later. If there is a fault in the spectrum or wrong value in the binarization leading to the adhesions of the recognition shape and noise background, this algorithm will not be able to get the correct spectrum information.

The regional fracture in the same category of communication system. In order to deal with this case, we provide the improved regional blur method. This method means the blur reprocess of the segmentation of the above method. And this method can be divided into two parts which are reconnection of the fracture spectrum and filling the absent digits.

### D. Spectrum fracture reconnection

Normally in a communication system, the PU makes the spectrum related continuity instead of shown point by point. So the PU is thought as has the feature of broad continuous. In our algorithm, we need to find the continuous images of a specific pattern to connect some fracture spectrum parts and recover some parts abandoned by the basic algorithm. Here we give the algorithm:

(1) For every line of the connected domain, calculate the minimum value of $\boldsymbol{y_i}$ and the maximum value of $\boldsymbol{height_i} + \boldsymbol{y_i}$. Then we rebuilt a set of blur connected domain of closed rectangle $\boldsymbol{r_i}$ in order to make the top left coordinate of $\boldsymbol{r_i}$ as( $\frac{1}{2}(\boldsymbol{x_i} + \boldsymbol{x_{i-1}})$ ,min $(\boldsymbol{y_i})$ **)** with width as $\boldsymbol{width_i} + \frac{1}{2}(2\boldsymbol{x_i} + \boldsymbol{x_{i+i}} + \boldsymbol{x_{i-1}})$ and height as max$(\boldsymbol{height_i} + y_i) -$ min $(y_i)$.

(2) For each $\boldsymbol{r_i}$, traversal all the abandon connected domain $\boldsymbol{c_j}$, if the ratio $\boldsymbol{R1_{i,j}}$ of $\boldsymbol{c_j}$ and $\boldsymbol{r_i}$ (ratio of width and height) reaches a threshold ( set as $\mathrm{T} = 0.1$ in this article), $\boldsymbol{c_j}$ will be seen as a part of PU's spectrum and be added.

### E. Deletion completion process

This part is used to find the occupancy of the PU spectrum in the approximate lost in the background noise. Usually this case only happens when the SNR of PU is very weak. When T is very small and approaching the background noise, in order to protect the PU can work normally under a low SNR, its spectrum space will be filled to avoid the wrong occupancy from PU in spectrum.

The absence completion algorithm is designed for this case. Here we give the algorithm:

(1) For every line of the connected domain, calculate the minimum value of $\boldsymbol{y_i}$ and the maximum value of $\boldsymbol{height_i} + \boldsymbol{y_i}$. Then we rebuilt a set of slit rectangle $\boldsymbol{s_i}$ in order to make the top left coordinate of $\boldsymbol{s_i}$ as $(\boldsymbol{x_i} + \boldsymbol{width_i}, \boldsymbol{min}(\boldsymbol{y_i}))$ with width as $(\boldsymbol{x_{i+1}} - \boldsymbol{width_i} - \boldsymbol{x_i})$ and height as $\boldsymbol{max}(\boldsymbol{height_i} + \boldsymbol{y_i}) - \boldsymbol{min}\,(\boldsymbol{y_i})$.

(2) In each $\boldsymbol{s_i}$, try to find its sub connected domain. If the ratio of the area of the minimum surround rectangle of this sub connected domain and $\boldsymbol{R2_i}$ reaches a threshold (0.3 in this article), a PU is judged as exist here. This rectangle ROI(religion of interest) is used to extract the connected domain in order to check the area of extracted connected domain and check the intersect degree of minimum surround rectangle and $s_i$.

Above all, we can see the pattern shape recognition algorithms can effectively process the missing digits and fracture of spectrum. The reconnection and filling for missing spectrum is aiming to perfect the function of distribution of PU's spectrum.

### F. Pattern shape process

Pattern shape cognition algorithm is basically two types: matching based on the template structure features, and matching based on the template statistic features. However, different communication systems have different spectrum occupation. Different spectrum occupation may have different statistic features, so matching algorithm based on statistic features could reduce the matching correctness. Therefore, this paper considers the cognition algorithm based on template structure as the main matching method. The detailed cognition process needs to firstly compute and input the correlation coefficients of the template, then find the most matching template based on these correlation coefficients. These correlation coefficients are generated by the autonomous matching learning.

**1) Definition of correlation coefficients:** For convenience of the computation of correlation coefficients, we modify the pixels of "0" into "-1". The definition of the correlation matrix in R^2 is: Correlation matrix $D = A \otimes B$, $A$ stands for the aiming matrix under matching; $B$ represents the convolution kernel(template matrix); $\otimes$ represents the convolution operation; each value $d_{i,j}$ in $D$ represents the sum of the products of corresponding overlapped elements, which is Eq.11:

$$\mathrm{d_{i,j}} = \sum\nolimits_{\mathrm{i,j \in overlap\ region}} \mathrm{b_{i,j}} \times \mathrm{a_{m-i+1,n-j+1}} \qquad (11)$$

The correlation matrix is a $0 < i < 2m - 1, 0 < j < 2n - 1$ matrix and $0 < i < 2m - 1, 0 < j < 2n - 1$, as shown in Figure 5.

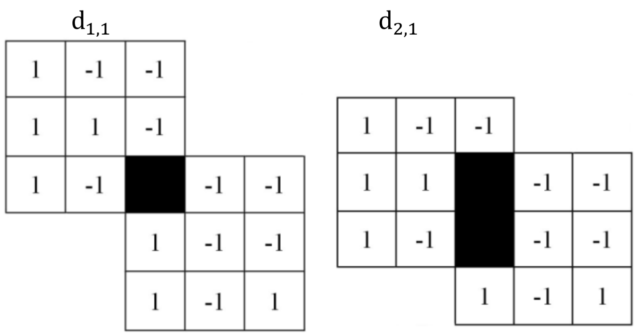


**Figure 6.** Overlap region schematic diagram

The methods of obtaining the correlation coefficients are two types: matching filter method and correlation matrix method. Both of two methods use the template matrix as the kernel of the filter, but they use different values as outputs.

**2) Matching filter method:** Matching filter method uses the value of center element of the correlation matrix as the output (correlation coefficient). According to the definition of matching filter in communication system , the center element represents the situation when matrix under matching and the convolution kernel are completely overlapped. The value of this center element is chosen as the output of the matching filter, which is the sum of the products of matrix under matching and template matrix point by point. The definition of the normalized correlation coefficient using matching filter method is: Eq.12.

$$r_1 = \frac{d_{xm,ym}}{\sum {b_{m,n}}^2} \qquad (12)$$

Where $xm, ym$ is the coordinates of the center element of correlation matrix. $b_{m,n}$ is an arbitrary element in the convolution kernel. The denominator in Eq.12 is also the area of the convolution kernel. The correlation coefficient after normalizing can be compared universally.

Matching filter uses only one element of the output correlation matrix. Hence, we can simplify the algorithm to reduce the time complexity. However, the matching filter uses less information during computation, and it is very sensitive to the translation transformation of matrix. So, the effect of matching filter method is worse than the correlation matrix method.

**3)** ***Correlation matrix method:*** Correlation matrix method finds the maximal value the whole correlation matrix as the output. In fact, the output value of matching filter is usually not the maximal one in correlation matrix. But it is almost random to regard the points at the edge of a picture as the preview or background during binarization. Even after segmentation and tilt correction, the image still has certain shifting and transforming. Correlation matrix method can handle these problems well.

Correlation matrix method needs to consider the offset distance at the same time with binarization threshold. Based on that, we give some "punishments" to the maximal value which is not appeared at the center of the convolution kernel. In this paper, the punishment is to deduce the Euclidean distance between the maximal value and the center point of the matrix, which is the adjusted as in Eq. 13.

$$d_{max}' = \begin{cases} \max(d_{i,j}) \ , \\ \max(d_{i,j}) \text{ in center of connected domain} \\ \max(d_{i,j}) - \sqrt{(i-x_m)^2 + (j-y_m)^2}, \\ \max(d_{i,j}) \text{ not in center of connected domain} \end{cases} \qquad (13)$$

The obtained maximal value should also be normalized. So the correlation coefficient obtained from matching filter method and correlation matrix method can be compared together. The definition of normalized correlation coefficient using correlation matrix method is: Eq.14.

$$r_2 = \frac{d_{max}'}{\sum {b_{m,n}}^2} \qquad (14)$$

**4)** ***Matching process of pattern shape:*** Before pattern shape cognition, we should fill all obtained connected domain to facilitate the pattern shape cognition based on template. To match the input with template, bilinear interpolation[7] is also needed to zoom the input for matching the pattern shape template. Hence, this paper proposes the adaptive matching algorithm of multi-pattern shape template based on confidence.

Common matching algorithm employs simple comparing algorithm to match existing pattern shape template, and select the one with highest correlation coefficient as output. But the matching result is independent to each other using this method. They don't use the special commons of communication systems of the same type, and this will increase the system computation.

Therefore, it needs to design an algorithm to fit different pattern shapes by adjusting matching template according to different pattern shapes. We propose the adaptive matching algorithm of multi-pattern shape template based on confidence.

The confidence of some type of PU spectrum to some input matrix is defined as $r_{con}$, which has the largest normalized correlation coefficient $r$ for input matrix under this type of pattern shape (the matrix under matching A is the pattern shape template of the same type of communication system, and the convolution kernel B is the input matrix). Besides, to combine matching filter method and correlation matrix method, we propose the conversion standard of matching algorithm, which is to use matching filter method first, and then use correlation matrix method if we cannot obtain reasonable correlation coefficient. At first, we compute each $r_{(1)}$ using matching filter method, and compute the confidence $r_{con(1)}$of pattern shape of some communication system. If $r_{(con)} \geqslant$mean($r_{(1)}$)+3×std($r_{(1)}$), the coefficient is said to be applicable for this pattern. Otherwise, we use correlation matrix method to re-compute $r_{(2)}$ and $r_{con(2)}$. (where mean() and std() are averaging and standard error function respectively.)

According to above definition, the adaptive algorithm using confidence can be described as follow:

Step1: re-order the input pattern shape, put the pattern shape closer to image center into the front. This is for increasing the credibility of the algorithm;

Step2: pick a pattern shape and compare it using $n$ types of pattern shape templates. Compute the $r_{con}$ of each PU spectrum, and obtain the cognition result of this pattern shape;

Step3: order the pattern shapes according to $r_{con}$ from high to low;

Step4: reduce $n$, repeat step 2 to obtain the final cognition result of pattern shape with largest average $r_{con}$.

In above steps, each iteration is a weighted average of all results of all experienced iterations and current iteration. This is for better statistic features. Pattern shape cognition and matching process are preceded at the same time. Besides, the reduction algorithm of n (the convergence of this algorithm) can also affect the cognition accuracy. In this paper, we apply the following sequence to reduce n: {0,0,0,1...}. The element of this sequence is the reduced amount of n for each iteration, and the number of "0" is the protection length. The correct pattern shape can be selected after several iterations, and the correctness of matching for the remaining pattern shape would be highly increased. This algorithm can not only increase the cognition correctness effectively but also significantly reduce the computation cost.

### *G. Post-processing of pattern recognition*

The post-processing is formed by several steps: firstly recover the missing area; then rearrange the output pattern shape according to former area architecture in order to get the final output. Here the recover step is by using the following standard: when the difference $y_m$ of two spectrum connected domains is less than the average height of connected spectrum, they are judged as in the same line; in the same line, the connected domain with smaller $x_m$ is in the front and in different line, the connected domain with smaller $y_m$ is in the front.

The final recognition result will have some difference with the real pattern shape. Considering a specific user case, the law of communication system needed to identify is limited. Assuming all the former spectrum distribution shapes of the communication system are already built in a database by the machine’s self-learning, we can use this database to do a reprocess to our results which is to pick the most likelihood one as the identified result. Here the Levenshtein algorithm [8] in introduced: define the distance ld of two areas is the minimum operation time by editing one to another (here the edit is defined as delete, add and replace and weight is always. Then the likelihood value Re of recognition result A and B inside the self-learning database is:

$$\mathrm{Re} = 1 - \frac{\mathrm{ld}}{\max\left(\mathrm{length}(A), \mathrm{length}(B)\right)} \qquad (15)$$

To all the Re, if max(Re)⩾mean(Re)+std(Re), return the shape of max(Re) in the database; otherwise this will be seen as the abnormal result which means the error of input waterfall plot is too large.

## IV. Experiments and Analysis

### *A. Sampling Spectrum Waterfall Plot*

For test the algorithm, we need to select a group of known PU's position of spectral band, such as: 2.4G ISM, the frequency range is 2.4GHz-2.4835GHz, a total bandwidth of 83.5MHz, planned 13 subchannels, each subchannel contain a bandwidth of 22MHz. Spectrum planning in Figure 7.

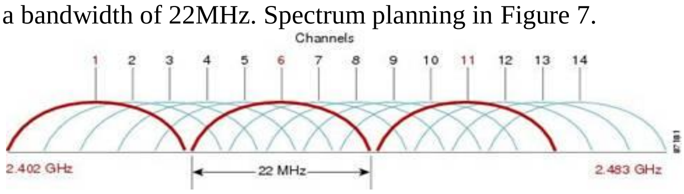


**Figure 7.** 2.4G ISM band planning

The experiments (scan interval 1.12ms, the threshold T is 0.7) showed spectrum results(Figure 8), we select the eight PU 's samples spectrum position ("TiroFinale" & "angela" have same frequency overlap with the channel 1, the center frequency is 2.412GHz; "gehua01143112" is low SNR spectrum , occupied channel 6, the center frequency is 2.437GHz; "CU_VRFY" is high SNR spectrum and has same frequency as a "guest" and "Tenda_5BD8D0" overlapping spectrum, the same as the channel 8, the center frequency is 2.447GHz; "HJX @ 1301" low SNR spectrum as "gehua01143112",occopied channel 10, the center frequency is 2.457GHz ; "HUAWEI-LECHGD" is high SNR spectrum of channel 11, the center frequency is 2.462GHz) for the experimental analysis of samples;

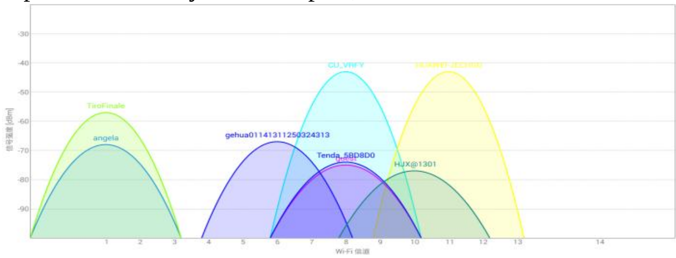


**Figure 8.** PU instantaneous measurement 2.4G ISM band spectrum

### *B. Spectrum waterfall area dividing process*

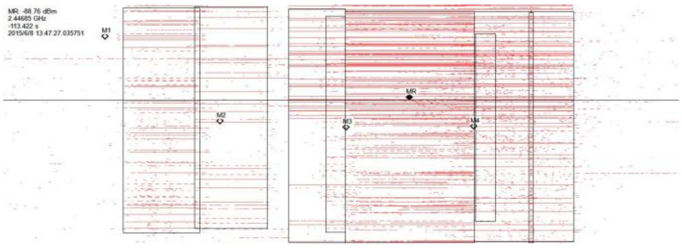

**Figure 9.** Area dividing process

Based on the spectral characteristics of the spectral characteristics of the region appear divided, according to the time linear characteristic, split box to box, when have slightly tilted, it is not obvious or within the tolerance range in the experiment can be considered, but we do not need to tilt adjustment; Usually we can see the area divided by two types: region segmentation and regional cross-independent segmentation, as shown in Figure 9; each divided region of their direct response spectrum occupancy PU maximum width.

### *C. Spectrum fracture reconnection and Deletion completion process*

**Figure 10.** Fracture reconnection process

After the area dividing completes in the region of the spectrum, we can find intermittent phenomenon, according to the PU features, to reconnect complete, shown in Figure 10.

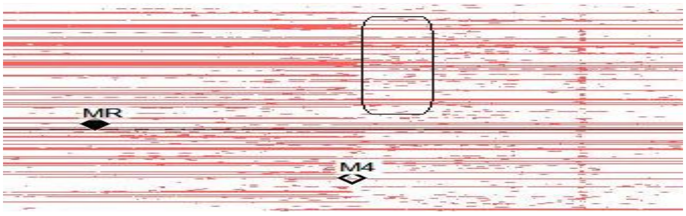


**Figure 11.**Deletion completion process

The divided region in an area maybe a low SNR ,we can find it drowned in a continuous spectrum of other phenomena, according to the PU features, which is replenished complete, as shown in Figure 11.

Absence fill and reconnection process can be carried out simultaneously and the whole process off the spectrum, in no particular order.

### *D. Spectrum waterfall pattern shape process*

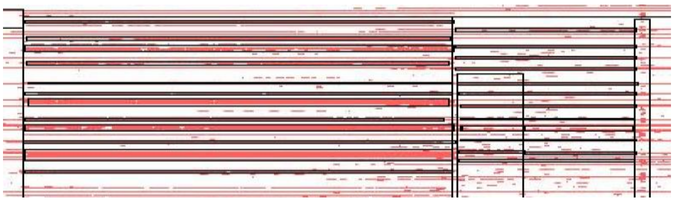

**Figure 12.**Area dividing process

According to results for area divided process, fracture reconnection, deletion completion and spectrum reconnection process, we will have complementary spectrum waterfall what can be refined after the split for the establishment of self-learning machine identification database data sources （shown in Figure 12）. These data sources will be made comparison reference for spectrum predicted at the late stage.

### *E. Perception analysis and comparison of the spectrum*

The spectral shape of the eight known PU for normalization process, statistics 10 times data for between spectrum waterfall pattern shape recognition algorithms "*" and the image pattern recognition algorithm "#" to comparison, as Figure 11; We find what same for high SNR spectrum recognition ability is differences, severe aliasing spectrum recognition rate is higher, such as "CU_VRFY" above "HUAWEI-LECHGD"(see Figure 13); Due to the customization features of this algorithm, serious aliasing spectral region is provided with high recognition rate than pattern recognition algorithms, such as: "CU_VRFY", "guest" and "Tenda_5BD8D0" is overlap with the frequency spectrum.

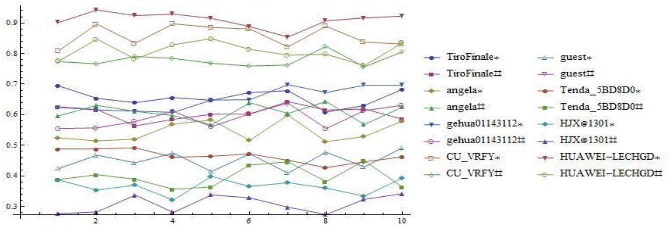


**Figure 13.**Algorithm comparison

## V. CONCLUSIONS

This article on the existing pattern recognition algorithm, according to the spectrum waterfall features, make certain innovation and improvement, made applicable to the shape of the spectrum waterfall in pattern recognition algorithm, greatly improve the spectrum sensing prediction accuracy and efficiency of the algorithm, the main :

(1) Based on the general characteristics of the communication system as a reference area segmentation algorithm is proposed fuzzy recognition algorithm has broken even and absence of spectrum recompletion of the area, making the spectrum area more perfect for division;

(2) Using adaptive pattern shape recognition will be based on the correlation coefficient and cross-correlation matrix of template matching;

(3) Combined with multi-mode shape stencil, difference in shape features based on different communication systems, and proposes an adaptive algorithm based on multi-pattern shape stencil degree of confidence.

In spite of the flaws, recognition algorithm make most cases the absence of reconnection and completion, when binarization threshold is too low to recognition, because a lot of background noise stacked, and the edge of the pattern shape is not re-identified. The algorithm should also have a lot of improving , such as image recognition algorithms based on neural networks can improve accuracy, with gigantic databases of large data mining algorithms, and so on，follow-up study to be.

## REFERENCES

[1] B. Wang, Liu, K.J.R, Advances in cognitive radio networks: A survey.Selected Topics in Signal Processing, IEEE Journal, 5(1), pp.5-23, 2011.

[2] K. Seshukumar, R. Saravanan, M.S. Suraj, Spectrum sensing review in cognitive radio. Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT): Tiruvannamalai, pp.1-4, 2013.

[3] G. C. Yu, C. Z. Long, Research on energy detection algorithm in cognitive radio systems. Computer Science and Service System (CSSS): Nanjing, pp.3460-3463, 2011.

[4] D. Bhargavi, C. R. Murthy, Performance comparison of energy, matched-filter and cyclostationarity-based spectrum sensing. Signal Processing Advances in Wireless Communications (SPAWC): Marrakech, pp.1-5, 2010.

[5] A. Bagwari, G. S. Tomar, Comparison between Adaptive Double-Threshold Based Energy Detection and Cyclostationary Detection Technique for Cognitive Radio Networks. Computational Intelligence and Communication Networks (CICN): Mathura, pp.182-185, 2013.

[6] Zou, Yaobin; Dong, Fangmin; Lei, Bangjun. Maximum similarity thresholding[J]. DIGITAL SIGNAL PROCESSING, 2013, (28): 120-135.

[7] Wang, Qunming; Shi, Wenzhong. Utilizing Multiple Subpixel Shifted Images in Subpixel Mapping With Image Interpolation[J]. IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, 2011, 11(4): 798-802.

[8] Ristad, ES; Yianilos, PN. Learning string-edit distance[J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 1992, 20(5): 522-532.

**Chen Wencheng**, is the student of school of information and communication engineering in Beijing University of Posts and Telecommunications(BUPT), where he has managed several projects related to wireless product development. He is also the teacher of BUPT-ROHDE & SCHWARZ Joint Laboratory. He received his M.Sc. in information and communication engineering from Beijing University of Posts and Telecommunications. His current research interests focus on radio resource and mobility management, software defined wireless networks, and broadband multimedia transmission technology.

**Wen Xiangming**, is the director of Beijing Key Laboratory of Network System Architecture and Convergence, where he has managed several projects related to open wireless networking. He is also the vice president of Beijing University of Posts and Telecommunications. He received both his M.Sc. and Ph.D in information and communication engineering from Beijing University of Posts and Telecommunications. His current research interests focus on radio resource and mobility management, software defined wireless networks, and broadband multimedia transmission technology.