

Article

Specific Emitter Identification Based on Synchrosqueezing Transform for Civil Radar

Mingzhe Zhu ¹, Zhenpeng Feng ^{1,*}, Xianda Zhou ¹, Rui Xiao ¹ and Yue Qi ²
and Xinliang Zhang ²

¹ School of Electronic Engineering, Xidian University, Xi'an 710121, China; zhumz@mail.xidian.edu.cn (M.Z.); zhouxiana999@gmail.com (X.Z.); xiaorui@stu.xidian.edu.cn (R.X.)

² Department of Electrical and Computer Engineering, Villanova University, 800 Lancaster Ave, Villanova, PA 19085, USA; yqi@villanova.edu (Y.Q.); xzhang4@villanova.edu (X.Z.)

* Correspondence: zpfeng6790@protonmail.com

Received: 24 February 2020; Accepted: 15 April 2020; Published: 16 April 2020



Abstract: Time-frequency (TF) signal features are widely used in specific emitter identification (SEI) which commonly arises in many applications, especially for radar signals. Due to data scale and algorithm complexity, it is difficult to obtain an informative representation for SEI with existing TF features. In this paper, a feature extraction method is proposed based on synchrosqueezing transform (SST). The SST feature has an equivalent dimension to Fourier transform, and retains the most relevant information of the signal, leading to on average approximately 20 percent improvement in SEI for complex frequency modulation signals compared with existing handcrafted features. Numerous results demonstrate that the synchrosqueezing TF feature can offer a better recognition accuracy, especially for the signals with intricate time-varying information. Moreover, a linear relevance propagation algorithm is employed to attest to the SST feature importance from the perspective of deep learning.

Keywords: linear relevance propagation; specific emitter identification; synchrosqueezing transform; time-frequency analysis

1. Introduction

Specific emitter identification (SEI) has always been a challenging problem in signal recognition. With the support of big data and high computing capability, technologies such as deep neural networks (DNN) can perform the identification task efficiently. However, the insufficiency of the training data set for DNN limits the depth of network and hinders performance yielding inaccurate classifications. Furthermore, understanding and interpreting network learning is always a challenging problem. Therefore, an effective feature extraction strategy is needed for SEI tasks. The ideal feature, before the classifier, should have a complete but low dimensional representation of the signal. Neither time or frequency alone can adequately fulfill this requirement when dealing with non-stationary signals. As a powerful non-stationary signal processing tool, TF analysis can reveal the signal local frequency characteristics. However, the mapping from one-dimensional signal to two-dimensional representation can introduce a large dimensionality problem. Attempts to reduce the redundancy in time-frequency (TF) feature have become an important subject [1–3]. Towards this end, some works [4–6] treat the TF spectrum as images. Then, downsampling of both variables is applied, and the result is sent to the classifier for identification. It is noted, however, that fine details of the emitter signature and modulation profile can be missed by classification confusion between closely resembling emitters.

Two effective methods have been recently introduced to tackle the above problem [7,8], namely the ambiguity function representative slice (AF-RS) and the compressed sensing mask

(CS-MASK). These two features not only reveal physical characteristics of measured radar signal, but also alleviate the computational costs. However, when it comes to complex frequency modulation signals, these two methods may perform unstably. For AF-RS, taking a single AF slice near the origin may only capture a part of the modulation information. For CS-MASK, it is complicated, and rather an art, to set the parameter of mask. Both methods face challenges when dealing with complex modulation signals, giving inadequate identification.

To overcome the limitations above, the synchrosqueezing transform (SST) [9], as an empirical mode decomposition-like tool, can be applied since it squeezes the energy from a wide frequency range into a narrow bandwidth near instantaneous frequency points, which can greatly reduce the scale of feature, and retains most modulation information. Hence, in this paper, a feature extraction method based on SST is proposed. This feature shows a good recognition rate, meanwhile, due to flavor and philosophy of feature extraction, it can probably facilitate the interpretation of the deep networks with linear relevance propagation (LRP) algorithm [10,11]. The data from civil aviation meteorological radar and radar signal generators are used to verify the validity of the proposed feature.

The paper is arranged as follows. Two TF feature extraction methods are discussed in the next section. Section 3 illustrates LRP algorithm. The SST method and the corresponding properties are introduced in Section 4; Section 5 proposes the SST feature extraction method. Numerous simulation results based on measured data and the interpretation of deep neural networks are given in Section 6. Finally, we conclude the paper in Section 7.

2. Time-Frequency Feature Extraction Methods

In this section, we briefly introduce the two TF feature extraction methods: Ambiguity Function Representative Slice (AF-RS) and Compressed Sensing Mask (CS-MASK) [7,8]. Ambiguity function (AF) concentrates most signal energy at and near the range-Doppler origin. Because of this property, it is possible to extract a feature with a low dimension by organizing the information near the origin in the ambiguity domain (AD).

2.1. AF-RS

The AF is the inverse Fourier transform of the instantaneous autocorrelation function of a signal over the time variable. In essence, it transforms signals to delay-frequency offset domain, whereas common methods depict signatures in joint time-frequency domain (TFD). Owing to the fact that typical targets do not rapidly transient over a short time period, most of the backscattering energy is not highly offset away from the AD origin. The AF of signal f is given by:

$$A_f(\tau, \nu) = \int_{-\infty}^{\infty} f\left(t + \frac{\tau}{2}\right) f^*\left(t - \frac{\tau}{2}\right) e^{-j\nu\tau} dt, \quad (1)$$

in which τ is time delay and ν is frequency offset. It is clear from (1) that when frequency offset is 0, the AF degenerates into the integral of autocorrelation function of the signal with respect to time. According to the definition of autocorrelation function, the maximum of AF is near the origin.

For illustration, consider a single-frequency signal and an LFM signal. The AF of the former is shown in Figure 1. At zero frequency offset, the signal energy reaches the maximum value. The cross-terms arising from the two Euler representations of a real sinusoid are shown parallel to frequency axis. These terms are considered artifacts and do not aid signal identification. On the other hand, the AF of LFM signal is shown in Figure 2. It can be seen that although the energy is mainly concentrated at zero frequency offset, the time-varying frequency of the signal cannot be fully captured in only one slice of zero frequency offset.

Since the energy of AF exhibits a peak at shift frequency being zero, the slices of ambiguity function with shift frequency near zero (including zero-slice) could be considered to be the major representative feature sets of radar emitter pulse which is named representative slice. This kind of feature not only coincides with the characteristics of real radar data but also can reduce the computation complexity.

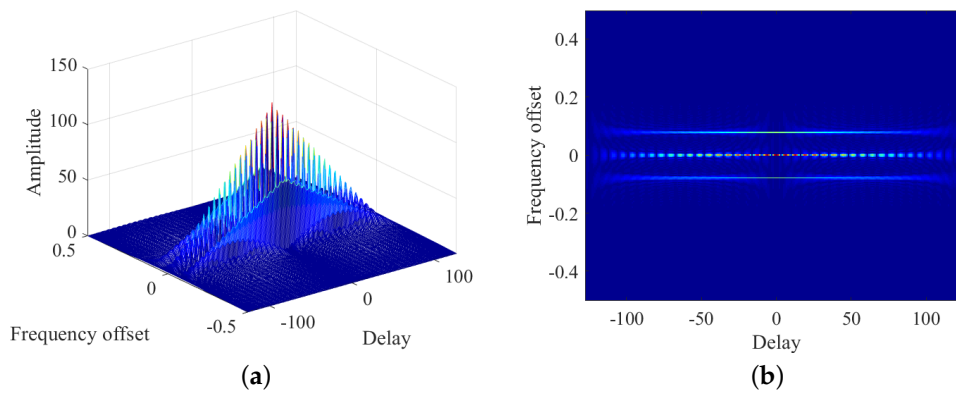


Figure 1. AF of single-frequency signal. (a) Three-dimension map of AF (b) Two-dimension map of AF.

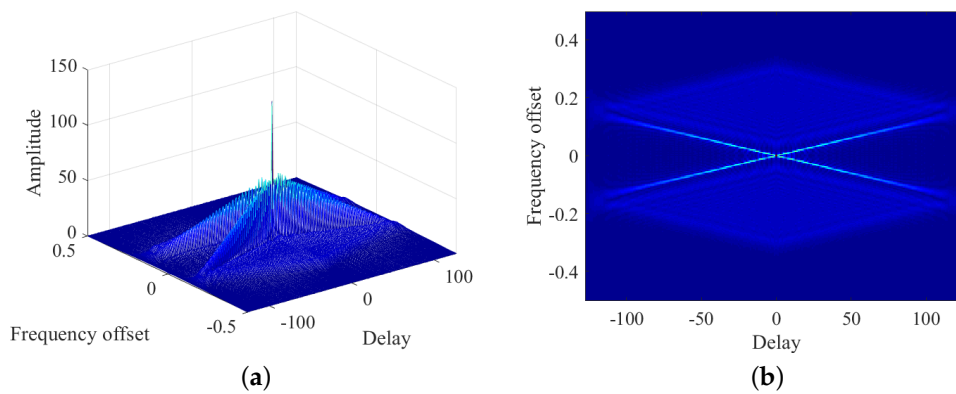


Figure 2. AF of LFM signal. (a) Three-dimension map of AF, (b) Two-dimension map of AF.

Figure 3 shows the slices of the LFM signal near the zero-frequency offset, which is from the frequency offset from 0 ($\nu = 0$) to 5 ($\nu = 5$):

A class-dependent method to classify radar emitter is proposed by Gillespie and Atlas [12,13], which is used to extract initial features of radar signals, followed by a kernel optimization scheme in ambiguity function plane. Paper [8] applied Direct Discriminant Ratio (DRR) as a criterion to rank the kernel, defined as:

$$DRR(\tau, \nu) = \sum_{i=1}^C \sum_{j=1}^C | \bar{A}^i - \bar{A}^j |^2, \tag{2}$$

here, \bar{A}^i means the "average" auto-ambiguity function of Class i . Frequency offset ν is usually set near zero in order to rank the points along the major direction of ambiguity function distribution. By this norm, AF-RS extracts the slices near the origin in delay-frequency offset domain as a feature set. The most representative slice is selected as the feature, resulting in a great reduction in data dimension and an obvious enhancement of identification accuracy. However, there are still some shortcomings, mainly in: (1) Only the slice at a fixed frequency is extracted as a feature, while other slices may contribute to identification in other offsets. (2) Compared with the zero-frequency-offset slice, the representative slice has more feature fusion optimization process, which probably makes it difficult to achieve real-time optimization. It must be based on recognition rate feedback, resulting in an increase in the computational complexity.

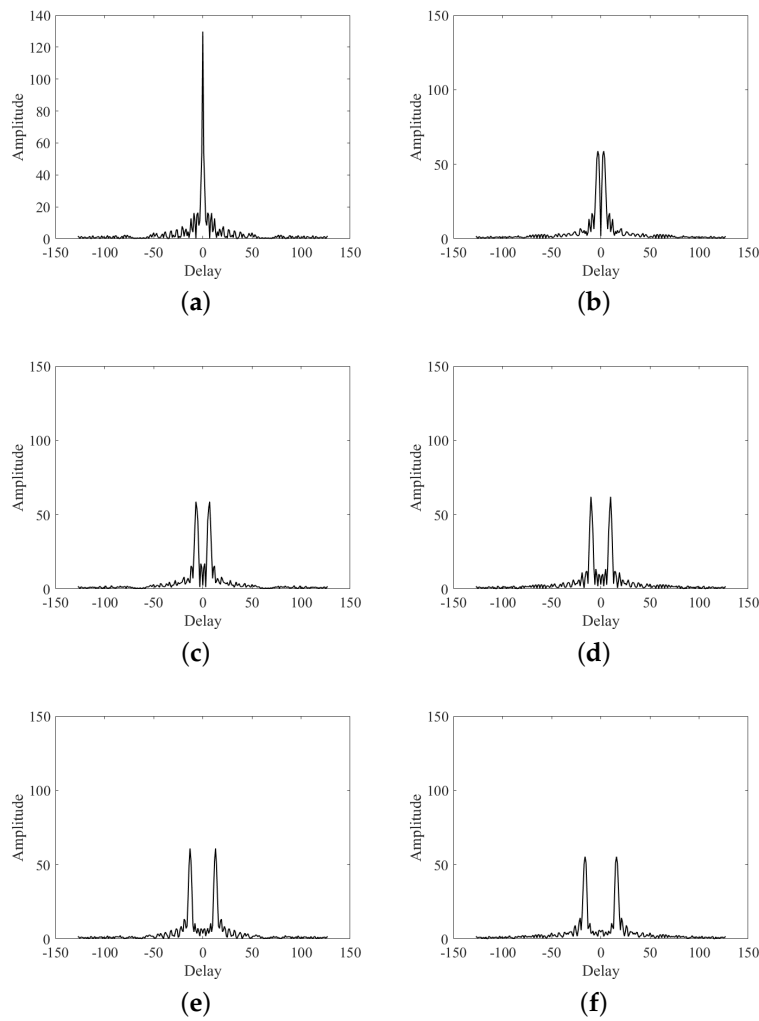


Figure 3. AF-RS in different offset. (a) $\nu = 0$, (b) $\nu = 1$, (c) $\nu = 2$, (d) $\nu = 3$, (e) $\nu = 4$, (f) $\nu = 5$.

2.2. CS-MASK

A compressed sensing mask is based on that ambiguity function and Wigner-Ville distribution is a two-dimensional Fourier transform pair relationship. After transforming the signal into the ambiguity domain, the compressed sensing theory is used to optimize the sparse ambiguity function mask feature.

According to the expression of ambiguity function, the two-dimensional inverse Fourier transform of the Wigner-Ville distribution is the ambiguity function, shown below:

$$A_f(\tau, \nu) = \int_{-\infty}^{\infty} f\left(t + \frac{\tau}{2}\right) f^*\left(t - \frac{\tau}{2}\right) e^{-j\nu\tau} dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} WVD_f(t, f) e^{-j(\nu\tau + \tau f)} dt df, \quad (3)$$

here $WVD_f(t, f)$ is Winger-Ville Distribution (WVD) of signal $f(t)$, a common quadratic time frequency representation tool, defined as:

$$WVD_f(t, f) = \int_{-\infty}^{\infty} f\left(t + \frac{\tau}{2}\right) f^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} dt. \quad (4)$$

Assuming WVD and ambiguity as length N points sequences, the matrix correlation of them as follows, which are related by two-dimensional Fourier transform:

$$A_{f(N \times 1)} = \Psi_{(N \times N)} \cdot W_{D(N \times 1)}, \quad (5)$$

where Ψ is 2D Fourier transform matrix in size of $N \times N$, and W_D is the matrix form of Winger-Ville Distribution of the processed signal $f(t)$. To reconstruct Wigner-Ville distribution from ambiguity function with high accuracy and low data dimension, a measurement matrix Φ in size of $M \times N$ ($M \ll N$) used in the compressed sensing is defined. Feature $A_{f(M \times 1)}^{Mask}$, an M sparse measurement of AF, is obtained as follows:

$$A_{f(M \times 1)}^{Mask} = \Psi_{(N \times N)} \cdot A_{f(N \times 1)} = \Phi_{(M \times N)} \cdot \Psi_{(N \times N)} \cdot W_{D(N \times 1)}. \tag{6}$$

An ideal feature should include as much information as possible with a low dimension. The sparse feature can be obtained according to compressed sensing theory, which can be optimized as:

$$A_{f(M \times 1)}^{Mask} = \underset{A}{\operatorname{arg\,min}} \|A_f\|_1, \tag{7}$$

$$\text{s.t. } \sum_{k=1}^N \sum_{n=1}^M \frac{1}{MN} [\mathcal{F}_{2d}^{-1}\{A_f\} - W_D] \leq \varepsilon, \tag{8}$$

where \mathcal{F}_{2d}^{-1} is 2D Fourier inverse transform, ε represents a user-specified bound to constrain the error in an acceptable range. Because the 2D Fourier transform of ambiguity function is WVD, the efficacy of the feature CS-MASK can be examined by the error between its 2D Fourier inverse transform and WVD according to Equation (8).

Figure 4 represents the ambiguity function and CS-MASK feature of an LFM signal. The signal length is 256, the size of the original ambiguity function is 256×256 , while the size of CS-MASK extracted according to Equations (7) and (8) is 17×17 , meaning the feature dimension has been greatly reduced.

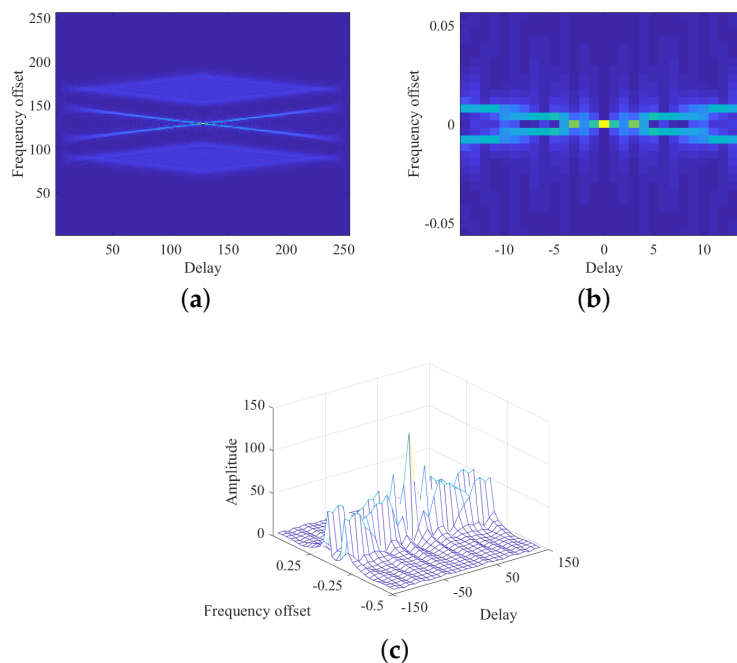


Figure 4. Comparison of AF-RS and CS-MASK of LMF signal. (a) AF of LMF signal, (b) Two dimensional CS-MASK, (c) Three dimensional CS-MASK.

It can be seen from Figure 4c that although CS-MASK only takes the information in a 17×17 region centered near the origin, it contains the most information about energy and frequency change.

The amount of information contained in CS-MASK can be measured with its recovered WVD. Figure 5 is a comparison of WVD of the original signal and WVD of the signal reconstructed by CS-MASK.

Nevertheless, there are still two points which can be improved in this method: (1) The selection of the mask size is based on the optimization criterion, which needs to be selected by the reconstructed signal WVD, so the calculation amount is large, although the rapid optimization method in two projects is proposed [7]; (2) The size of the optimized feature is not fixed but will change with certain factors such as signal modulation and noise interference.

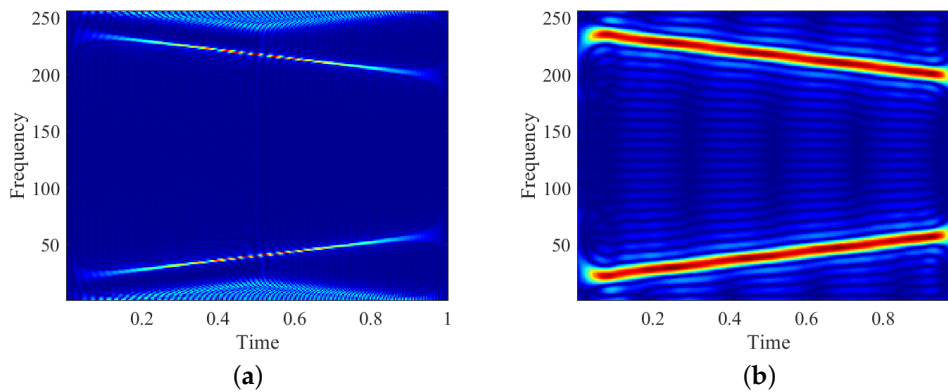


Figure 5. WVD of LFM signal (a) WVD of original signal. (b) WVD of signal reconstructed by CS-MASK.

3. Linear Relevance Propagation

The LRP method is generally used in the middle layer visualization and feature inversion in image processing with DNN. LRP is based on the layer-wise conservation principle which states that relevance R is conserved in process of back-propagating from one layer to the next, i.e., [11],

$$R^{(1)} = \dots = R^{(l)} = R^{(l+1)} = \dots = f(x). \tag{9}$$

The relevance of the last layer is equivalent to the classification function $f(x)$ of the input data. In each layer, relevance R is defined as below:

$$R^{(l)} = \sum_{n=1}^N r_n^{(l)}, \tag{10}$$

where $r_n^{(l)}$ is the relevance of the neurons in each layer. The heatmap $h_n = r_n^1$, the relevance in the first hidden layer, is obtained by the following rule:

$$r_n^{(l)} = \sum_m (\alpha \frac{z_{n,m}^-}{\sum_{n'} z_{n',m}^-} + \beta \frac{z_{n,m}^+}{\sum_{n'} z_{n',m}^+}) r_m^{(l+1)}, \tag{11}$$

where α and β are two fixed parameters during iterations, and $z_{n',m}^+$ and $z_{n',m}^-$ are the positive and negative part of $z_{n',m}$ severally. $z_{n',m}$ is the contribution of n th neuron at layer l to the activation of the m th neuron in the next layer $l + 1$,

$$z_{n',m} = a_n^{(l)} \omega_{n,m}^{(l,l+1)}, \tag{12}$$

here $a_n^{(l)}$ is the activation of neuron n in layer l , and $\omega_{n,m}^{(l,l+1)}$ is the weights from n th neuron in layer l to m th neuro in layer $l + 1$.

In this paper, LRP method is used to create a heatmap for the recognition result of each method that reflects which parts of the original data contribute positively to recognition of network and which contributes negatively. Combined with this heatmap, we can understand this network, perhaps even manage to interpret how it works. This research is conducted in the simulation part.

4. Synchrosqueezing Transform

SST, proposed by Daubechies et al., is based on wavelet transform and empirical mode decomposition (EMD) [9]. SST is a post TF processing tool which compresses the spectrum in TFD by squeezing the energy distributed in a frequency range into some certain instantaneous frequency points. SST has been widely applied in many fields of biomedical signals, geographic signals, sonar signals, etc. [14,15].

In this paper, Short Time Fourier Transform (STFT), a commonly used time-frequency analysis tool, is adopted as time frequency transform. STFT converts the signals into a 2D time-frequency domain by performing Fourier Transform in a fixed window traversing the time domain. Considering a signal $f(t)$, the STFT of $f(t)$ with a window $g(t)$ in Schwartz class, is defined as [16–18]:

$$V_f^g(t, f) = \int_{\mathbb{R}} f(\tau)g^*(\tau - t)e^{-i2\pi f\tau} d\tau, \tag{13}$$

where $g(t)^*$ is the complex conjugate of $g(t)$, and f is frequency.

Signal $f(t)$ to be considered in synchrosqueezing techniques in the STFT is defined by:

$$f(t) = \sum_{k=1}^K f_k(t), \text{ with } f_k(t) = a_k(t)e^{i\phi_k(t)}, \tag{14}$$

where K is a constant, $a_k > 0$ is a continuously differentiable function, ϕ_k is a two times continuously differentiable function satisfying $\phi'_k(t) > 0$ and $\phi'_{k+1}(t) > \phi'_k(t)$ for all t . In this context, ideal T-F (ITF) representations can be defined as:

$$IFT_q(t, f) = \sum_{k=1}^K a_k(t)^q \delta(2\pi f - \phi'_k(t)), \tag{15}$$

where q is a positive integer depending on the chosen T-F distribution (TFD), and δ refers to Dirac Distribution. The STFT being is associated with $q = 1$.

Base on this theory, the synchrosqueezing operator reassigns the STFT as follows:

$$T_f^g(t, f) = \frac{1}{g(0)^*} \int_{\mathbb{R}} V_f^g(t, f) e^{i2\pi f t} \delta(2\pi f - \hat{\omega}_f(t, f)) df, \quad g(0)^* \neq 0 \tag{16}$$

where $\hat{\omega}_f(t, f) = \phi'(\hat{\tau}_f(t, f))$. $\hat{\tau}_f(t, f)$ is a time reassignment factor, defined as [17]:

$$\hat{\tau}_f(t, f) = \frac{1}{T_f^g(t, f)} \int \int_{\mathbb{R}^2} \tau W_g(\tau - t, \nu - 2\pi f) W_f(\tau, \nu) \frac{d\tau d\nu}{2\pi}, \quad T_f^g(t, f) \neq 0, \tag{17}$$

here W_g and W_f are WVD of $g(t)$ and $f(t)$.

By squeezing the frequency components within a certain range of instantaneous frequency, an energy-concentrated TF representation can be obtained. Then, the reconstruction of the original signal is computed by

$$f(t) = \frac{1}{C_h} \int_{-\infty}^{+\infty} T_f^g(t, f) \frac{df}{f}, \quad C_h \neq 0, \tag{18}$$

where $C_h = Re(C_\varphi C_\psi)$, $C_\varphi = e^{-i2\pi f t f^2}$ and $C_\psi = \frac{1}{2} \int_0^{+\infty} \overline{\hat{\psi}(\xi)} \frac{1}{\xi} d\xi$. $\overline{\hat{\psi}(\xi)}$ is the complex conjugate of the Fourier transform of the mother wavelet $\psi(t) = \frac{1}{\sqrt{\omega\pi}} e^{-\frac{t^2}{2}} e^{i2\pi t}$ ($\omega = 2\pi f$). More details can be found in [19].

Because of the property that the energy of a signal is mainly concentrated near the instantaneous frequency points, it is only necessary to extract vital frequency, energy, and time-varying information. The effect of first-order SST is better only when the second derivative of the phase with respect to time is approximately zero or much smaller than its first derivative. However, for complex frequency modulation signals, such as secondary frequency hopping signals, the second derivative of phase is not negligible. Hence, vertical synchrosqueezing transform (VSST) is proposed [20]. In this paper, VSST is used as the TF analysis tool because the experimental datasets contain numerous complex modulation signals.

5. The Proposed SST Feature Extraction Algorithm

In this section, the SST feature extraction is proposed to achieve an accurate recognition rate. The property of VSST is elaborated and some experimental results of VSST are shown in Section 5.1. Additionally, the multipath effect, a common phenomenon in radar signal processing, usually leads to an adverse impact on feature representation. VSST can mitigate the multipath effect, which is discussed in Section 5.2 and the SST feature extraction method is presented in Section 5.3

5.1. Vertical Synchrosqueezing Transform

The latest research results [20] have achieved vertical synchrosqueezing transform (VSST) and even high-order form. For nonlinear FM signals, the concentration of the transform is further improved. D.H.Pham gives a VSST expression based on STFT:

$$T_f^g(t, f) = \frac{1}{g(0)^*} \int_{\mathbb{R}} V_f^g(t, f) e^{i2\pi f t} \delta(\omega - \hat{\omega}_f(t, f)) df, \quad g(0)^* \neq 0, \tag{19}$$

$$\tilde{\omega}_f(t, f) = \begin{cases} \hat{\omega}_f(t, f) + \hat{q}_f(t, f)(t - \hat{\tau}_f(t, f)) & , \partial_t \hat{\tau}_f(t, f) \neq 0 \\ \hat{\omega}_f(t, f) & , \text{otherwise} \end{cases}$$

Here, $\hat{\omega}_f(t, f)$ and $\hat{\tau}_f(t, f)$ are TF rearrangement factors, $\tilde{\omega}_f(t, f)$ is the second-order estimate of the instantaneous frequency of signal. Compared with the first-order SST, VSST obtains more accurate instantaneous frequency estimation by adding the adjustment factor $\hat{q}_f(t, f)$, meaning the TF concentration can be further improved. The original signal can be quickly recovered using the following formula [20]:

$$f(t) = \int_{\mathbb{R}} T_f^g(t, f) df. \tag{20}$$

Figure 6 shows the VSST and the inverse transform of synthetic linear frequency modulation pulse signal. The signal information is shown below:

$$sig(t) = \cos(2\pi 30t^2 + 2\pi 20t). \tag{21}$$

The normalized fs = 256, the total length of the signal is 512 and the signal to noise ratio (SNR) is 0 dB.

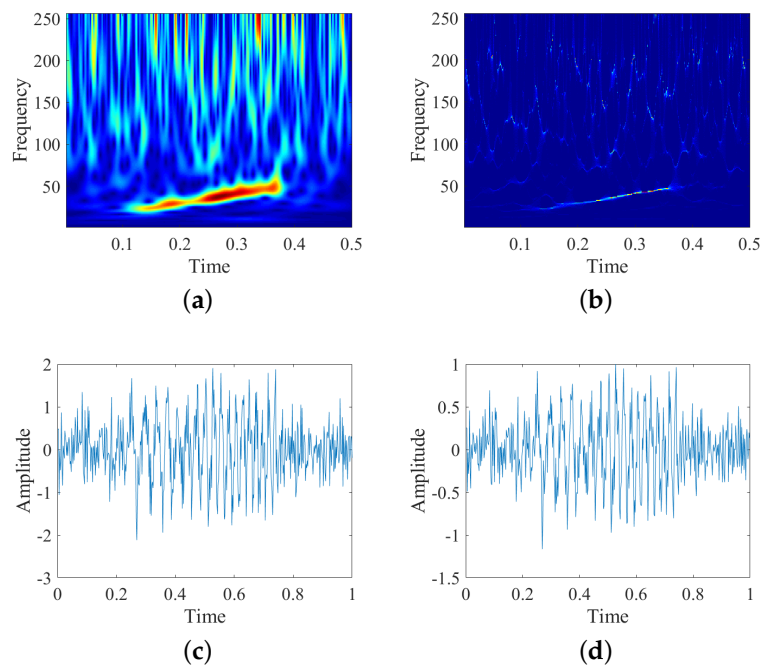


Figure 6. STFT and SST of the simulation signal $sig(t)$ and the comparison of original signal and reconstructed signal. (a) STFT, (b) SST, (c) original signal, (d) reconstructed signal.

From Figure 6, VSST retains the time-varying information of the original signal and has a lossless inverse transformation, which brings great convenience for feature extraction and identification. Figure 7 is an enlarged partial view of VSST. It can be seen that most energy is concentrated near instantaneous frequency points. Compared to STFT and WVD, SST is a more accurate TF representation because of its perfect TF concentration.

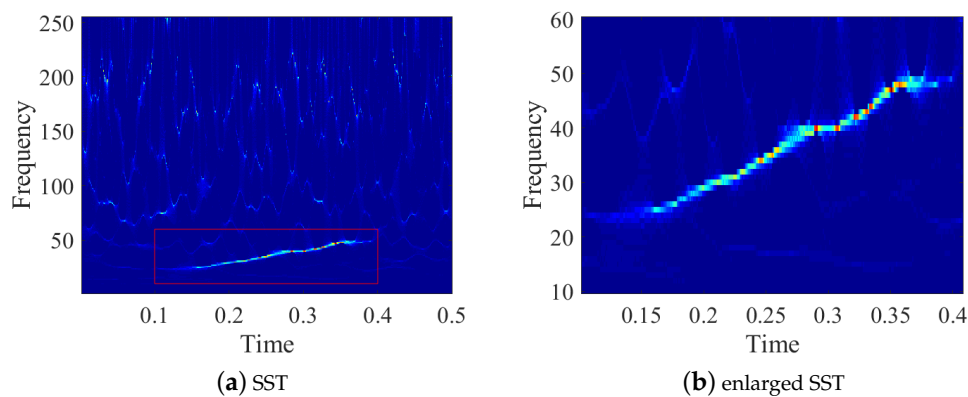


Figure 7. SST and enlarged part.

5.2. Suppression on Multipath Effect by VSST

The multipath effect refers to a phenomenon that after electromagnetic wave propagates through different paths, each component arrives at the receiver in different time, which causes great interference due to the superposition of each phase [20]. The multipath effect denotes as:

$$f_m(t) = f(t) + E_m * f(t - \tau), \tag{22}$$

where f_m is the signal with multipath effect, E_m is the ratio of amplitude of multipath effect to that of signal. Figure 8 shows the SST of the simulation signal $sig(t)$ in Equation (21) with multipath effect under different E_m . Here, time delay has two conditions: $\tau = 16$ and 128 sampling points. The SNR is 0 dB. From Figure 8a,d,g,j, it can be seen that multipath effect can cause a great adverse impact on SST feature. Figure 8c,f,i,l manifest that VSST can separate the multipath component from the signal and suppress it effectively.

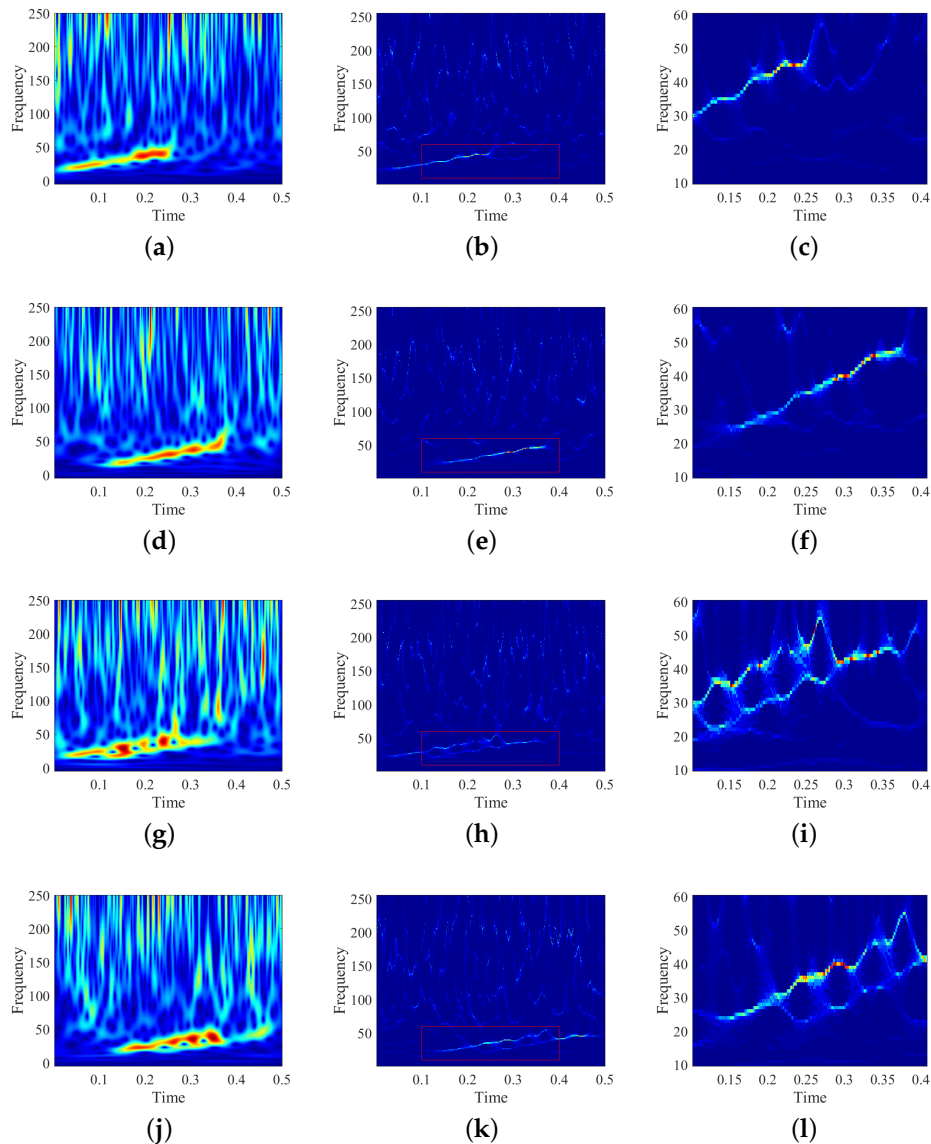


Figure 8. STFT and SST of the signal with multipath effect in different E_m and τ . (a–c) are STFT, SST, and enlarged part of SST with $E_m = 0.2$ and $\tau = 16$. (d–f) are STFT, SST, and enlarged part of SST with $E_m = 0.2$ and $\tau = 128$. (g–i) are STFT, SST, and enlarged part of SST with $E_m = 0.8$ and $\tau = 16$. (j–l) are STFT, SST, and enlarged part of SST with $E_m = 0.8$ and $\tau = 128$.

5.3. Feature Extraction

It is clear from the various experimental results of SST that energy from the signal in TFD is mainly concentrated within a limited range, while some redundant information is distributed around the rest area. Based on this property, SST feature will be derived as follows: First, STFT together with VSST are performed on one signal to obtain the TF spectrum. Second, the energy sequence $E(n)$ and the phase trajectory sequence $f(n)$ are extracted from the SST domain. It is worth noting that

in order to embody more detailed information, the compression bandwidth of the energy sequence can be appropriately extended. Finally, the two sequences are merged into one new feature matrix and then sent to the classifier for identification. The feature matrix is shown below:

$$\mathbf{M} = \begin{bmatrix} f(1) & f(2) & \dots & f(n) \\ E(1) & E(2) & \dots & E(n) \end{bmatrix} \quad (23)$$

Take the SST feature of the real measured signal from DatasetII as an example (The description of DatasetII is given in the next section), see in Figure 9, the length of the signal is $N = 500$, and the size of original spectrum matrix is 250×500 . In comparison, the size of the feature matrix extracted by SST is reduced to 2×500 .

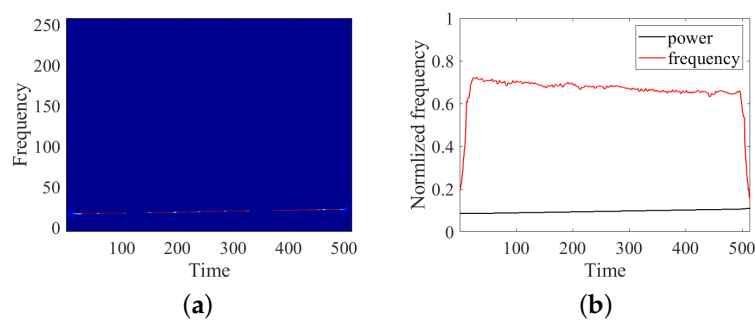


Figure 9. The SST feature of a sample in DatasetII. (a) VSST. (b) Feature matrix.

The process of the proposed algorithm (including feature extraction and LRP interpretation) is given as follows and Figure 10:

Step1: Signal Preprocessing.

Preprocess the signals to obtain invalid and neat signals including demean, aligning and normalization. (Denoise is not considered in this paper since it may remove important details by accident, while SST can keep the main part of the signal within the narrow band and remove the rest of the signal with noise.) The obtained data usually contain abnormal value, noise, and interference which influence the identification in this algorithm. In addition, normalization is required in order to avoid the adverse impact of different signal values for identification.

Step2: Feature Extraction.

In this step, firstly it is required to perform STFT on the processed signals. Then SST is used to suppress the energy of STFT. In SST spectrum, instantaneous frequency and suppressed energy are extracted as a new TF feature matrix.

Step3: Identification.

Finally, the extracted SST feature in Step2 is delivered into DNN for training and identification. It should be noted that the structure of DNN for each type of data is identical. A better recognition rate means that the feature may include more useful information for training the network.

Step4: Interpretation.

Combined with the LRP algorithm, heatmaps of various signals can be obtained. Based on these heatmaps, we can manage to understand how the networks work, and even try to give a reasonable interpretation of its identification mechanism.

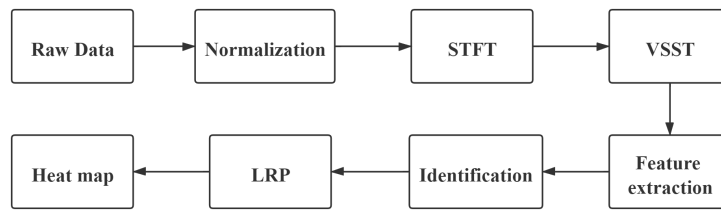


Figure 10. Flowchart of data processing.

6. Experimental Results

In this section, the supervised classifier is applied to identify the four sets of measured radar data, and compare SST feature with other three features to verify the effectiveness of the proposed method. DatasetI contains 10 types of signals from 10 identical types of civil aviation meteorological radars. It includes 500 approximate single-frequency signal samples; the next three data sets are all collected from radar signal generators. DatasetII contains 10 types of 500 single frequency signal samples; DatasetIII contains 10 types of 500 LFM signal samples; DatasetIV contains 10 types of 500 LFM signal samples. The reason for choosing these four data sets is that they are very typical and representative real measured signals of civil radars. We randomly select a signal from each data set to perform STFT and SST in order to show the TF characteristics of each data set, shown in Figure 11. Table 1 shows the processing time to obtain SST feature of each data set.

The extracted SST features of each data set are then sent to the classifier for training and identification. The training rate is set from 10% to 70%, and deep neural network (DNN) with fully connected layers is used as a classifier. The recognition rates under different features are shown in Tables 2–5 and Figure 12. It can be seen that for DatasetI, SST feature has only an inapparent advantage over the power spectrum (PSE), AF-RS and CS-MASK [7,8]. For the single-frequency signal in DatasetII, the effect of SST features is even slightly worse than PSE and CS-MASK. It is because the form of single-frequency signals is simple, and the energy is highly concentrated at the origin of AD domain, hence the information extracted by the mask is very complete. From the perspective of PSE, single-frequency signals are time-invariant signals, so one-dimensional energy information can represent signals enough, while SST feature may introduce redundant information which reduces the recognition rate.

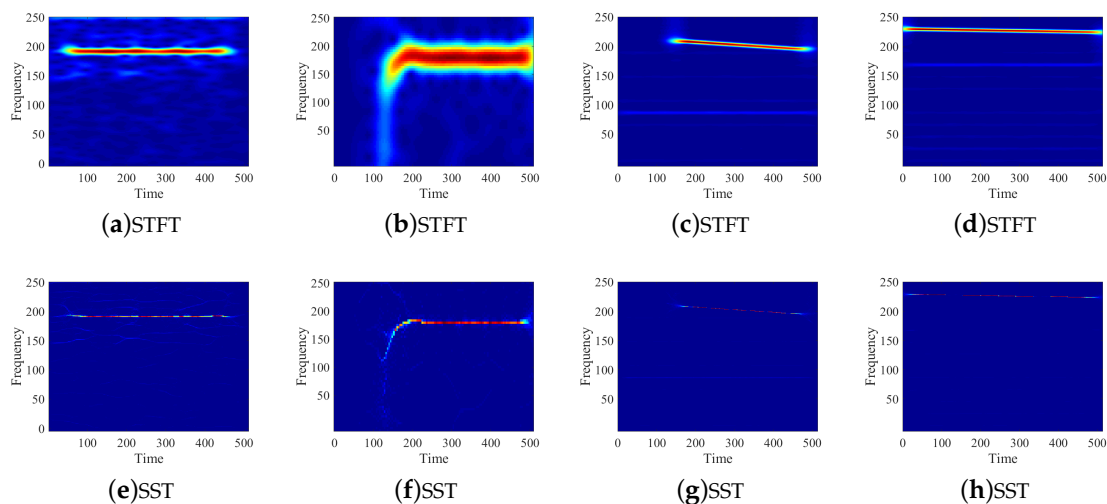


Figure 11. STFT and SST spectrum of Dataset, the signal in (a) and (e) is from DatasetI, the signal in (b) and (f) is from DatasetII, the signal in (c) and (g) is from DatasetIII, the signal in (d) and (h) is from DatasetIV.

Table 1. Processing time for each dataset.

Dataset	DatasetI	DatasetII	DatasetIII	DatasetIV
Processing time/ms	64.60	65.42	65.86	76.12

Table 2. Recognition rate for DatasetI.

Training Rate	10	20	30	40	50	60	70
PSE	93.10	92.18	90.61	96.43	91.78	92.80	94.01
AF-RS	88.69	91.27	89.96	91.06	88.59	89.58	87.70
CS-MASK	92.01	90.75	92.87	90.57	90.03	93.17	90.16
SST	85.68	91.45	91.56	94.03	92.71	94.45	96.02

Table 3. Recognition rate for DatasetII.

Training Rate	10	20	30	40	50	60	70
PSE	56.60	73.87	86.67	93.62	93.01	86.69	95.34
AF-RS	62.00	63.25	69.73	71.33	84.67	80.02	86.00
CS-MASK	93.54	91.26	95.98	93.54	92.51	94.19	91.92
SST	67.18	62.40	78.74	89.45	81.33	84.59	84.42

Table 4. Recognition rate for DatasetIII.

Training Rate	10	20	30	40	50	60	70
PSE	41.56	62.50	82.54	90.67	91.60	93.50	92.67
AF-RS	52.39	64.75	74.86	84.33	81.60	83.00	87.33
CS-MASK	52.56	72.75	82.57	83.33	83.20	89.00	86.00
SST	88.67	93.75	95.14	98.00	99.20	98.23	98.60

Table 5. Recognition rate for DatasetIV.

Training Rate	10	20	30	40	50	60	70
PSE	80.55	82.90	89.15	92.95	96.45	93.05	96.15
AF-RS	70.65	71.25	72.95	73.55	76.05	86.95	84.05
CS-MASK	69.10	72.75	73.65	75.95	75.30	81.00	75.30
SST	89.30	97.75	98.70	98.30	98.00	98.45	99.35

In contrast, for DatasetIII and IV, the SST feature obtains the best recognition rate among these four features. This is because the modulation form of LFM signal is more complicated than single-frequency signal, and theoretically it needs more accurate information to identify. SST feature contains more detailed time-varying information of signal than other three features, so it can obtain a better representation of LFM signal.

To further verify the effectiveness of the SST feature, the LRP algorithm [10,11] is used to obtain the TF heatmaps in reverse. The parameters α and β in Equation(11) are set as $\alpha = 3$, and $\beta = -1.5$. The number of the hidden layers in auto-coder is 3, and *relu* is selected as the activation function of each layer. Figure 13 shows heatmaps of deep network for LFM signals from DatasetIII and DatasetIV. The brightness of pixels represents the contribution to identification of deep network. It is apparent from Figure 13 that the brightest region is concentrated near the SST feature, meaning SST feature is akin to the part of spectrum which contributes most to the identification of deep network. In other words, the decisive information in deep network is not disarray of some pixels in spectrum, but some special parts of the spectrum make the most contribution. It can be observed that there exists a horizontal stripe and two vertical stripes in Figure 13a,b. The horizontal stripe may be caused by harmonic components in the signal, and the vertical stripes may be caused by phase hopping. In Figure 13c,d, a dim horizontal stripe appears within the first quarter of the entire time,

which is probably due to the noise at the beginning from signal generators. The specific causes of this phenomenon need to be further studied. These experimental results demonstrate the validity of the SST feature, and also reflect that recognition of the deep network is possibly interpretable.

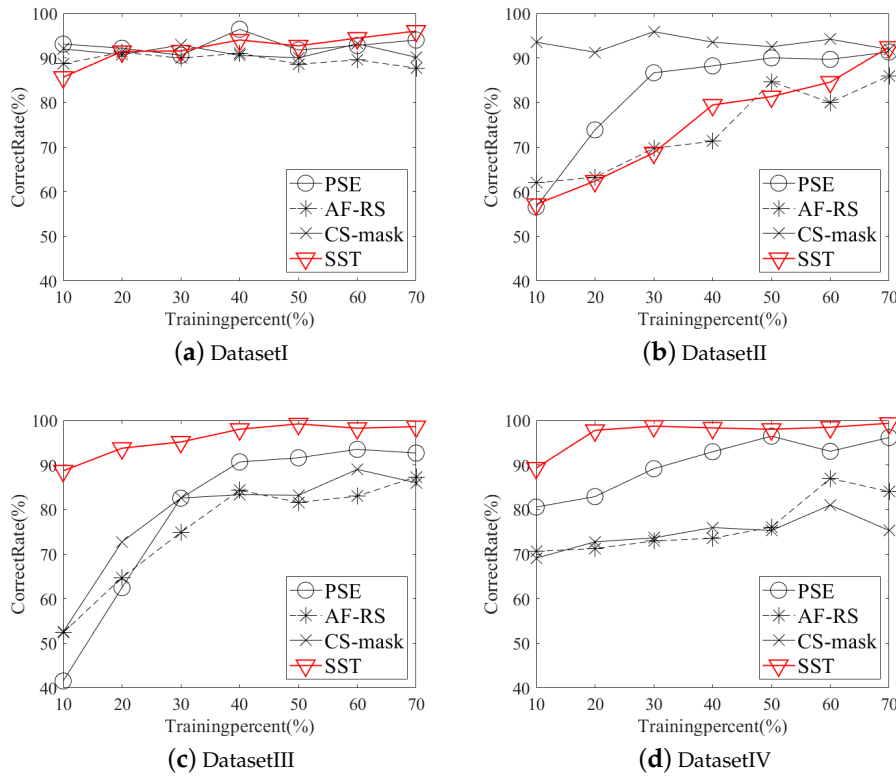


Figure 12. Recognition rate for each Dataset.

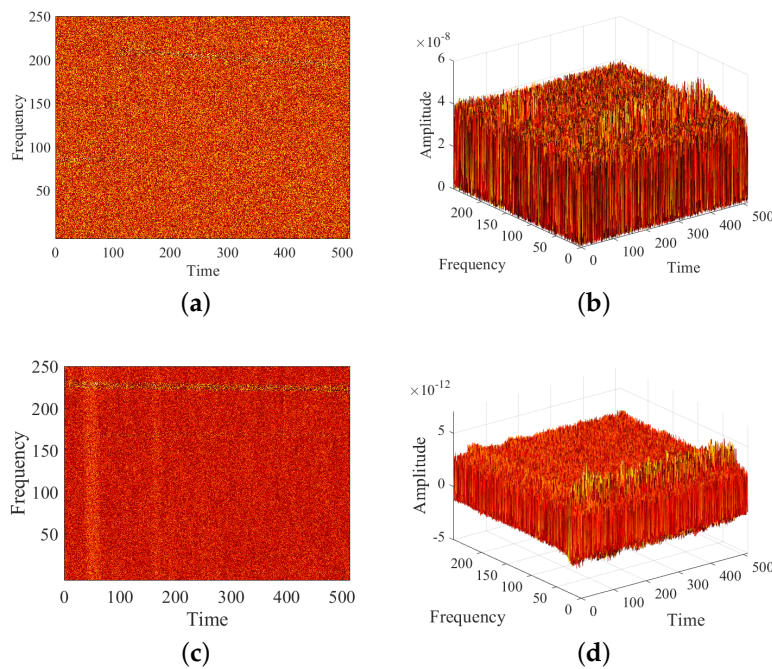


Figure 13. Heatmaps of the signals from DatasetIII and DatasetIV. (a) Two dimensional heatmap of signal from DatasetIII, (b) Three dimensional heatmap of signal from DatasetIII, (c) Two dimensional heatmap of signal from DatasetIV, (d) Three dimensional heatmap of signal from DatasetIV.

7. Conclusions

In this paper, we proposed a time-frequency feature extraction method based on SST. The energy sequence and phase trajectory sequence are merged into a feature matrix, which retains almost complete time-frequency information, and holds clear physical meaning with a small dimension. The proposed feature can obtain better recognition rates especially for complex modulation signals due to its high energy concentration and precise instantaneous frequency which can contain more beneficial information, meanwhile, suppress redundant or even negative information, such as noise and multipath effect, to classifier. The superiority of the method in comparison to PSE, AF-RS, and CS-MASK [7,8] is verified by the identification of the measured radar data. Furthermore, combined with the proposed feature and LRP algorithm [11], it is possible to understand and interpret the inner mechanism of the identification by deep networks.

Author Contributions: The contributions of the authors are as follows. Conceptualization, M.Z.; Methodology, Z.F.; Software, X.Z. (Xianda Zhou); Investigation R.X.; Data Curation Y.Q. and X.Z. (Xinliang Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Shaanxi Province (2019JM-412) and the National Natural Science Foundation of China (61701374).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ye, W.; Peng, C. Recognition Algorithm of Emitter Signals Based on PCA+CNN. In Proceedings of the 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12–14 October 2018; pp. 2410–2414.
- Mingqiu, R.; Jinyan, C.; Yuanqing, Z. Classification of Radar Signals Using Time-frequency Transforms and Fuzzy Clustering. In Proceedings of the 2010 International Conference on Microwave and Millimeter Wave Technology, Chengdu, China, 8–11 May 2010; pp. 2067–2070.
- Yang, L.B.; Zhang, S.S.; Xiao, B. Radar Emitter Signal Recognition Based on Time-frequency Analysis. In Proceedings of the IET International Radar Conference, Xi'an, China, 14–16 April 2013; pp. 1–4.
- Wang, X.; Huang, G.; Zhou, Z.; Gao, J. Radar Emitter Recognition Based on the Short Time Fourier Transform and Convolutional Neural Networks. In Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, China, 14–16 October 2017; pp. 1–5.
- Wang, C.; Wang, J.; Zhang, X. Automatic Radar Waveform Recognition Based on Time-frequency Analysis and Convolutional Neural Network. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2437–2441.
- Zhu, J.; Zhao, Y.; Tang, J. Automatic Recognition of Radar Signals Based on Time-frequency Image Character. *IET Int. Radar Conf.* **2013**, *63*, 1–6.
- Zhu, M.; Zhang, X.; Qi, Y.; Ji, H. Compressed Sensing Mask Feature in Time-Frequency Domain for Civil Flight Radar Emitter Recognition. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AL, Canada, 15–20 April 2018; pp. 2146–2150.
- Wang, L.; Ji, H.; Shi, Y. Feature Extraction and Optimization of Representative-slice in Ambiguity Function for Moving Radar Emitter Recognition. In Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Dallas, TX, USA, 14–19 March 2010; pp. 2246–2249.
- Daubechies, I.; Lu, J.; Wu, H.T. Synchrosqueezed Wavelet Transforms: An Empirical Mode Decomposition-like Tool. *Appl. Comput. Harm. Anal.* **2011**, *30*, 243–261. [[CrossRef](#)]
- Gurbuz, S.Z.; Amin, M.G. Radar-Based Human-Motion Recognition With Deep Learning: Promising Applications for Indoor Monitoring. *IEEE Signal Proc. Mag.* **2019**, *36*, 16–28. [[CrossRef](#)]
- Amin, M.G.; Erol, B. Understanding Deep Neural Networks Performance for Radar-based Human Motion Recognition. In Proceedings of the 2018 IEEE Radar Conference (Radar-Conf18), Oklahoma City, OK, USA, 23–27 April 2018; pp. 1461–1465.
- Gillespie, B.W.; Atlas, L.E. Optimizing Time-Frequency Kernels for Classification. In *IEEE Transactions on Signal Processing*; IEEE: Piscataway, NJ, USA, 2001; pp. 485–496.

13. Gillespie, B.W.; Atlas, L.E. Optimization of Time and Frequency Resolution for Radar Transmitter Identification. In Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Phoenix, AR, USA, 15 March 1999, Volume 3; pp. 1341–1344.
14. Herry, C.L.; Frasch, M.; Seely, A.J.; Wu, H.T. Heartbeat classification from single-lead ecg using the synchrosqueezing transform. *Physiol. Meas.* **2017**, *38*, 171. [[CrossRef](#)]
15. Huang, Z.; Zhang, J.; Zhao, T.; Sun, Y. Synchrosqueezing S-Transform and Its Application in Seismic Spectral Decomposition. *IEEE Trans. Geo Sci. Remote Sens.* **2016**, *54*, 817–825. [[CrossRef](#)]
16. Fourer, D.; Auger, F.; Czarnecki, K. Chirp Rate and Instantaneous Frequency Estimation: Application to Recursive Vertical Synchrosqueezing. *IEEE Signal Proc. Lett.* **2017**, *24*, 1724–1728. [[CrossRef](#)]
17. Auger, F.; Flandrin, P.; Lin, Y. Time-Frequency Reassignment and Synchrosqueezing: An Overview. *IEEE Signal Proc. Mag.* **2013**, *30*, 32–41. [[CrossRef](#)]
18. Fourer, D.; Auger, F. Second-order Time-Reassigned Synchrosqueezing Transform: Application to Draupner Wave Analysis. In Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO), Coruna, Spain, 2–6 September 2019; pp. 1–5.
19. Thakur, G.; Wu, H.T. Synchrosqueezing-based Recovery of Instantaneous Frequency from Nonuniform Samples. *SIAM J. Math. Anal.* **2010**, *43*, 2078–2095. [[CrossRef](#)]
20. Oberlin, T.; Meignen, S.; Perrier, V. Second-order synchrosqueezing transform or invertible reassignment? towards ideal time-frequency representations. *IEEE Trans. Signal Proc.* **2015**, *63*, 1335–1344. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).