

Article

SAR Image Recognition with Monogenic Scale Selection-Based Weighted Multi-task Joint Sparse Representation

Zhi Zhou , Ming Wang, Zongjie Cao * and Yiming Pi

School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; zzhuestc@163.com (Z.Z.); mingwangesperer@gmail.com (M.W.); ymp@uestc.edu.cn (Y.P.)

* Correspondence: zjcao@uestc.edu.cn; Tel.: +86-028-6183-0379

Received: 13 February 2018; Accepted: 21 March 2018; Published: 22 March 2018



Abstract: The monogenic signal, which is defined as a linear combination of a signal and its Riesz-transformed one, provides a great opportunity for synthetic aperture radar (SAR) image recognition. However, the incredibly large number of components at different scales may result in too much of a burden for onboard computation. There is great information redundancy in monogenic signals because components at some scales are less discriminative or even have negative impact on classification. In addition, the heterogeneity of the three types of components will lower the quality of decision-making. To solve the problems above, a scale selection method, based on a weighted multi-task joint sparse representation, is proposed. A scale selection model is designed and the Fisher score is presented to measure the discriminative ability of components at each scale. The components with high Fisher scores are concatenated to three component-specific features, and an overcomplete dictionary is built. Meanwhile, the scale selection model produces the weight vector. The three component-specific features are then fed into a multi-task joint sparse representation classification framework. The final decision is made in terms of accumulated weighted reconstruction error. Experiments on the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset have proved the effectiveness and superiority of our method.

Keywords: monogenic signal; scale selection; multi-task sparse representation

1. Introduction

Synthetic aperture radar (SAR) automatic target recognition (ATR) is becoming increasingly important as the development of radar technology continues [1]. Much research has been based on SAR images [2]. However, the traditional pixel intensity cannot be always treated as a reliable feature in the classification field. In recent years, the monogenic signal has been used for SAR ATR because of its ability to capture the characteristics of SAR images. The monogenic signal is a generalized extension of the analytic signal in high-dimensional space, and was first introduced by Felsberg and Sommer in 2001 [3]. Similarly to the Hilbert transform to a one-dimensional (1-D) analytic signal, the original signal can be orthogonally decomposed into three components with the Riesz transform. The three components are local energy, local phase, and local orientation. The decomposed operation is usually used for high-dimensional signal analysis and processing. This decoupling strategy makes it possible to deal with many problems in the field of image processing, especially when the traditional pixel intensity can not be treated as a reliable feature in the classification field. Furthermore, the monogenic signal has the ability to capture broad spectral information, which is useful for SAR image recognition. Then, the monogenic scale space is proposed to unify scale-space theory and the phase-based signal

processing technique [4,5]. The monogenic signal has provided a new viewpoint and method in the field of low-level image processing, due to its ability to provide local features of the phase-vector and attenuation in scale space. In [6], the monogenic approach allows for analyzing the local intrinsic two-dimensional (2-D) features of low image data through the information fusion of monogenic signal processing and differential geometry. In [7], a new complex wavelet basis is constructed from the complex Riesz operator to specify multiresolution monogenic signal analysis. A novel application of wavelets to coherent optical imaging is also shown. In [8], the monogenic signal is applied to color edge detection and color optical flow. In recent years, the application of the monogenic signal has been extended into the field of pattern recognition as research continues. In [9], the random monogenic signal is used to distinguish between unidirectional and nonunidirectional fields. In [10], the monogenic signal is exploited at multiple scales to extract features from subtle facial micro-expressions and a linear support vector machine (SVM) is used for classification. In [11], the texture and motion information which is extracted from the monogenic signal is used for facial expression recognition.

Since multiple features, including the monogenic signal and Gabor wavelets, can be extracted from original images, there is a growing tendency to combine the multiple features for image recognition. Multiple features extracted from original images can provide different information with respect to target characteristics when compared with a single-feature approach. Therefore, approaches based on the combination of feature modalities can fuse their benefits and have better classification capability and higher classification accuracy [12]. Much of the study has been focused on multiple kernel learning, and the core ideas on the similarity functions between images are linearly combined [13,14]. Since sparse representation classification (SRC) methods have powerful expressive ability and simplicity, other studies on feature combination aim at building a multi-task joint sparse representation classification (MTJSRC) model. The SRC concept was firstly introduced in [15]. Then, the application of SRC was rapidly extended to classification in other domains [16,17]. In [18], SRC with optimizing kernel principal component analysis (PCA) is used to perform pattern classification task based on a data set of SAR images. Recently, multiple views of SAR images or multiple feature extracted from SAR images were employed as inputs in the SRC field. Multi-view images carry more information than single-view images. The multi-view image-based approaches can fuse the benefits and improve the performance of target recognition. Similarly, feature combination based on an MTJSRC model can also achieve better recognition performance. According to the idea mentioned above, a multi-task joint sparse representation classification approach is proposed [19]. In [20], a joint sparse representation-based multi-view method is proposed to exploit the inter-correlations among the multi-view images and improve the performance of a recognition system. In [21], the component-specific feature descriptor of a monogenic signal is proposed and the multiple features are then fed into the MTJSRC to make the final decision.

A large number of components of the monogenic signal can be acquired from the original SAR images because of three components (local amplitude, local phase, and local orientation) and their multiple scales. This characteristic of the monogenic signal provides a valid option for MTJSRC. Meanwhile, MTJSRC can effectively fuse the benefits of the multiple features and improve the recognition accuracy. As a result, there are many advantages with respect to the joint use of the monogenic signal and MTJSRC. However, the large data size produced by the monogenic signal makes the computation load increase rapidly. Moreover, different scales of the monogenic signal carry different amounts of information for target classification. In particular, some scales of the monogenic signal have little or even a negative impact on the recognition accuracy. Hence, a measurement model is needed to calculate the informativeness of each component at different scales. On the other hand, the traditional sparse representation algorithms based on monogenic signal make the final decision by accumulating the total reconstruction error directly, which totally ignores the heterogeneity of the three components of the monogenic signal. Since a measurement model for the informativeness of each component at different scales is established, the weight vector associated with the reconstruction

error can be also determined adaptively. The process of making decisions based on the measurement model becomes more reasonable and the recognition accuracy is improved.

In this paper, a scale selection method, based on a weighted multi-task joint sparse representation (abbreviated as WTJSR), is proposed for SAR image recognition. Three components of the monogenic signal at different scale spaces are extracted from the original SAR images, which carry rich information for target classification. Meanwhile, the data set becomes enormous. Then, a scale selection model based on the Fisher discrimination criterion is designed. The higher the score, the more discriminative the components are at the corresponding scale. A global Fisher score is proposed to measure the discriminative ability of components at each scale. The less discriminative scales are abandoned and the rest of the components are concatenated to three component-specific features. Meanwhile, the adaptive weight vector is provided by the scale selection model. The three component-specific features are then fed into a tri-task joint sparse representation classification framework. The final decision is made by the cumulative weighted reconstruction error. Our contributions are shown below:

- (1) We introduce a novel joint sparse representation method (WTJSR) with the components of the monogenic signal
- (2) We propose a scale-selection model based on a Fisher discrimination criterion to effectively use the information contained in monogenic signal and establish the adaptive weight vector due to the heterogeneity of the three component-specific features.
- (3) We perform comparative experiments under different conditions.

The rest of this paper is organized as follows. In Section 2, the SRC and MTJSRC are introduced. In Section 3, the monogenic signal is introduced and Fisher's discrimination criterion-based monogenic scale selection is proposed. The WTJSR is analyzed. Afterwards, several experiments are presented in Section 4. Conclusions are provided in Section 5.

2. Related Work

This section briefly introduces two prior concepts, SRC and MJSRC. Some necessary terms are first described to facilitate the following description. Assume that the image size is w by h pixels. Every image is transformed to a column vector \mathbb{R}^m where $m = w \times h$.

2.1. SRC

The sparse signal representation technique has been extensively studied over the last decade. The resurgent development of many theoretical analysis frameworks [22,23] and effective algorithms [24] has been witnessed. The applications of the sparse signal representation technique mainly include radar imaging [25,26], image restoration [27], image classification [28,29], and pattern recognition [15,30]. The key of sparse signal model is based on the fact that a certain signal can be represented by an overcomplete basis set (dictionary). The convex strategy is generally adapted to finding an optimal solution [31,32].

The number of training samples from all the classes is n . Let $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}] \in \mathbb{R}^{m \times n_k}$ be the concentration of the training samples from the k th ($k = 1, 2, \dots, K$) class, where $n = n_1 + n_2 + \dots + n_K$. Sparse representation is based on a simple assumption that the new query sample of the same class lies in the same subspace of training samples with a much lower dimension [33–35]. Therefore, the test sample \mathbf{y} with the k th class can be represented as

$$\mathbf{y} = \mathbf{x}_{k,1}\alpha_{k,1} + \mathbf{x}_{k,2}\alpha_{k,2} + \dots + \mathbf{x}_{k,n_k}\alpha_{k,n_k}. \quad (1)$$

The vector of scalar coefficients is $\alpha_k = [\alpha_{k,1}, \alpha_{k,2}, \dots, \alpha_{k,n_k}]^T \in \mathbb{R}^{n_k}$.

As a matter of fact, the class of the test sample is almost unknown initially. Therefore, the test sample \mathbf{y} is represented by the whole training group. Let $n = \sum_{k=1}^K n_k$ denote the total number of training

samples and $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K] \in \mathbb{R}^{m \times n}$ represent the training set. Based on the whole training set, the test sample \mathbf{y} can be rewritten as

$$\mathbf{y} = \mathbf{X}_1\alpha_1 + \mathbf{X}_2\alpha_2 + \dots + \mathbf{X}_K\alpha_K = \mathbf{X}\alpha \quad (2)$$

where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K]^T \in \mathbb{R}^n$ is the coefficient vector. Most of the elements in the representative vector α have values of zero, except those related to the training samples of the k th class. Theoretically, the target recognition problem is converted into solving the linear system $\mathbf{y} = \mathbf{X}\alpha$.

Frequently, the solution to the equation $\mathbf{y} = \mathbf{X}\alpha$ is not unique because of the fact that $m < n$. A popularly used solver seeks the sparsest linear combination of the dictionary to represent the test sample [36]:

$$\min_{\alpha} \|\alpha\|_0, \text{ s.t. } \|\mathbf{y} - \mathbf{X}\alpha\| \leq \varepsilon \quad (3)$$

where ε is the allowed error tolerance. Traditionally, sparsity is measured using the l_0 -norm by counting the number of nonzero elements. The problem of finding the optimally sparse representation, i.e., minimum $\|\alpha\|_0$, is a combinatorial optimization problem. However, Equation (3) has been proven to be non-deterministic polynomial (NP) hard and it is difficult to find an optimal solution in theory [37]. Considering the difficulty in solving large combinatorial problems, some algorithms such as orthogonal matching pursuit and relaxed formulations have been proposed over the years [33,38,39]. Since the solution is sparse enough, Equation (3) can be relaxed as l_1 -norm minimization [40]

$$\min_{\alpha} \|\alpha\|_1, \text{ s.t. } \|\mathbf{y} - \mathbf{X}\alpha\| \leq \varepsilon. \quad (4)$$

In the ideal situation, most of the elements in the representative α are zero except those associated with the ground-truth class. In practice however, the recovered sparse representation vector has most of its nonzero elements with large absolute values associated with the ground-truth class, while small or zero values are distributed elsewhere. Hence, the minimum reconstruction error criterion is used to decide the label of the test sample \mathbf{y} with the optimal solution $\hat{\alpha}$ [41];

$$\text{class}(\mathbf{y}) = \arg \min_{k \in \{1, 2, \dots, K\}} \|\mathbf{y} - \mathbf{X}_k \hat{\alpha}_k\|. \quad (5)$$

2.2. MTJSRC

SRC is proposed to deal with the classification problem with a single test sample or feature. MTJSRC is developed to expand the SRC algorithm in multiple features. Suppose that the number of features extracted from the original images is P . For each modality (task) index, $p = 1, 2, \dots, P$, denote $A^p = [A_1^p, A_2^p, \dots, A_K^p]$ as the p th training feature modality matrix. $k = 1, 2, \dots, K$ refers the columns associated with the k th class. A multi-task linear representation problem with P modalities of features extracted from the original data can be described as:

$$\mathbf{y}^p = \mathbf{X}^p \alpha^p, p = 1, 2, \dots, P. \quad (6)$$

Inspired by SRC, the sparse representation of each modality $\mathbf{y}^p, p = 1, 2, \dots, P$ can be obtained from the optimization of the following l_1 -norm problem:

$$\min_{\alpha} \|\alpha\|_1, \text{ s.t. } \|\mathbf{y}^p - \mathbf{X}^p \alpha^p\| \leq \varepsilon \quad (7)$$

The optimal solution $\hat{\alpha}^p, p = 1, 2, \dots, P$ can be obtained by computing the optimization problem repeatedly P times. Given all of the sparse representation vectors, the minimum reconstruction error criterion accumulated over all the P tasks is used to decide the label of the original test sample \mathbf{y} :

$$class(\mathbf{y}) = \arg \min_{k \in \{1,2,\dots,K\}} \sum_{p=1}^P \left\| \mathbf{y}^p - \mathbf{X}_k^p \hat{\mathbf{a}}_k^p \right\|. \tag{8}$$

When $P = 1$, SRC is obtained.

3. Monogenic Scale Selection-Based Weighted Multitask Joint Sparse Representation Classification

3.1. Monogenic Signal

The monogenic signal is an extension from analytic signal theory and Hilbert transform theory [3]. The analytic signal is a complex-valued representation of a 1-D real valued signal $f^{(1)}(x)$:

$$f_a(x) = f^{(1)}(x) + if_H(x), x \in R. \tag{9}$$

where $f_H(x) = f^{(1)} * \frac{1}{\pi x}$ denotes the Hilbert-transformed signal. In the above definition, the local amplitude $A(x)$ and the local phase $\phi(x)$ are shown in another appearance of the analytic signal:

$$\begin{aligned} A(x) &= \sqrt{f^{(1)}(x)^2 + f_H(x)^2} \\ \phi(x) &= \text{atan2}\left(f_H(x), f^{(1)}(x)\right) \end{aligned} \tag{10}$$

where the local amplitude $A(x)$ represents the local energy, and the local phase $\phi(x)$ changes if the local structure varies.

In order to deal with high-dimensional signals like images and videos, the monogenic signal has been developed. The monogenic signal is built around the Riesz transform, which is an extension of the Hilbert transform from 1-D to N-D. In the 2-D case, the spatial representation of the Riesz kernel is

$$(R_x(\mathbf{z}), R_y(\mathbf{z})) = \left(\frac{x}{2\pi|\mathbf{z}|^3}, \frac{y}{2\pi|\mathbf{z}|^3} \right), \mathbf{z} = (x, y) \in \mathbb{R}^2 \tag{11}$$

and its frequency response is

$$(H_u(\mathbf{u}), H_v(\mathbf{u})) = \left(-i \frac{u}{|\mathbf{u}|}, -i \frac{v}{|\mathbf{u}|} \right), \mathbf{u} = (u, v) \in \mathbb{R}^2. \tag{12}$$

The expression of Riesz-transformed signal in the spatial domain is

$$f_R(\mathbf{z}) = -\frac{\mathbf{z}}{2\pi|\mathbf{z}|^3} * f(\mathbf{z}) \tag{13}$$

where $*$ denotes the convolution.

For an image $f(\mathbf{z})$, the monogenic signal is composed of the image signal $f(\mathbf{z})$ itself and its Riesz transformed one $f_R(\mathbf{z})$:

$$f_M(\mathbf{z}) = f(\mathbf{z}) - (i, j) f_R(\mathbf{z}) \tag{14}$$

where i and j are the imagery units, and $\{i, j, 1\}$ forms an orthonormal basis in \mathbb{R}^3 . Similarly to the analytic signal, the original image signal is decomposed orthogonally into three components: local amplitude A , local phase ϕ , and local orientation θ , which can be generated as

$$\begin{aligned} \mathbf{A} &= \sqrt{f(\mathbf{z})^2 + |f_R(\mathbf{z})|^2} \\ \phi &= \text{atan2}\left(|f_R(\mathbf{z})|, f(\mathbf{z})\right) \in (-\pi, \pi] \\ \theta &= \text{atan}\left(f_y(\mathbf{z})/f_x(\mathbf{z})\right) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right] \end{aligned} \tag{15}$$

where $f_x(\mathbf{z}) = f(\mathbf{z}) * h_x(\mathbf{z})$ denotes the i -imaginary component and $f_y(\mathbf{z}) = f(\mathbf{z}) * h_y(\mathbf{z})$ represents the j -imaginary component. Similarly to the analytic signal, the local amplitude A describes the local energetic information and the local phase ϕ denotes the local structural information. The major difference is that the monogenic signal has one more component, i.e., local orientation, which describes the geometric information.

The signals are usually of finite length in applications. Therefore, in order to extend the signal to be infinite, a bandpass filter is used before the Reisz transform. Then, the monogenic signal can be rewritten as

$$f_M(\mathbf{z}) = (h_b(\mathbf{z}) * f(\mathbf{z})) - (i, j) (h_b(\mathbf{z}) * f_R(\mathbf{z})) \quad (16)$$

where $h_b(\mathbf{z})$ denotes the bandpass filter. The log-Gabor filters have the ability to catch broad spectral information [42]. Hence, the log-Gabor filter bank is employed in this paper. The frequency response of the log-Gabor filter can be generated as

$$G(\omega) = \exp \left\{ - \left[\log \left(\frac{\omega}{\omega_0} \right) \right]^2 / \left[\log \left(\frac{\sigma}{\omega_0} \right) \right]^2 \right\} \quad (17)$$

where ω_0 is the center frequency and σ is the scaling factor of the bandwidth. Given that s is the scale index, the multiresolution monogenic signal representation can be acquired, which forms the monogenic scale space. The monogenic signal with scale S ($S = 10$) is shown in Figure 1.

Suppose the scale parameter is S . The monogenic scale space $\mathbf{f} = \{\mathbf{f}_A, \mathbf{f}_\phi, \mathbf{f}_\theta\}$ of local amplitude, local phase, and local orientation can be described as

$$\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_S, \phi_1, \phi_2, \dots, \phi_S, \theta_1, \theta_2, \dots, \theta_S\}. \quad (18)$$

3.2. Fisher Discrimination Criterion-Based Monogenic Scale Selection

As we can see from Equation (3), the data size of the feature set \mathbf{f} is increased by a factor of $3S$ compared with that of the original image data set. The feature set normally leads to considerable computational complexity because of redundancy and high dimensions, which makes it difficult to be applied in the recognition system directly. In order to deal with this problem, a multitask joint sparse representation strategy based on a scale selective model is proposed in this paper without concatenating all the features at S scales together in the learning system [21,43]. The features at some scales may be less discriminative or even have a negative effect on classification. In order to verify this statement, a typical example and some analyses will be given later.

A monogenic scale selection method based on Fisher's discrimination criterion is proposed to solve this problem. This method is aimed at finding the most discriminative features in monogenic scale space. As we discussed before, the three components of the monogenic signal, i.e., local amplitude, local phase, and local orientation are three different types of features. Fisher's discrimination criterion is employed separately to the scale space of each component.

In the scale space of the amplitude features, let the training samples of the k th class in the s th scale be A_k^s . The within-class and between-class distances (denoted as $Y_W(\mathbf{A}_k^s)$ and $Y_B(\mathbf{A}_k^s)$) in the scale space of the amplitude features with the k th class are, respectively, defined as

$$Y_W(\mathbf{A}_k^s) = \sum_{\mathbf{a}_i \in A_k^s} (\mathbf{a}_i - \mathbf{m}_k^s)^T (\mathbf{a}_i - \mathbf{m}_k^s) \quad (19)$$

$$Y_B(\mathbf{A}_k^s) = n_k (\mathbf{m}_k^s - \mathbf{m}^s)^T (\mathbf{m}_k^s - \mathbf{m}^s) \quad (20)$$

where \mathbf{m}_k^s and \mathbf{m}^s are the mean vectors of A_k^s and A^s , respectively, which are defined as

$$\mathbf{m}_k^s = \frac{1}{n_k} \sum_{\mathbf{a}_i \in A_k^s} \mathbf{a}_i \quad (21)$$

$$\mathbf{m}^s = \frac{1}{n} \sum_{\mathbf{a}_j \in \mathbf{A}^s} \mathbf{a}_j. \quad (22)$$

where n_k is the number of training samples of the k th class in the scale space of the amplitude features, and n is the number of training samples in the scale space of the amplitude features. $(\cdot)^T$ denotes the transpose of a matrix or vector.

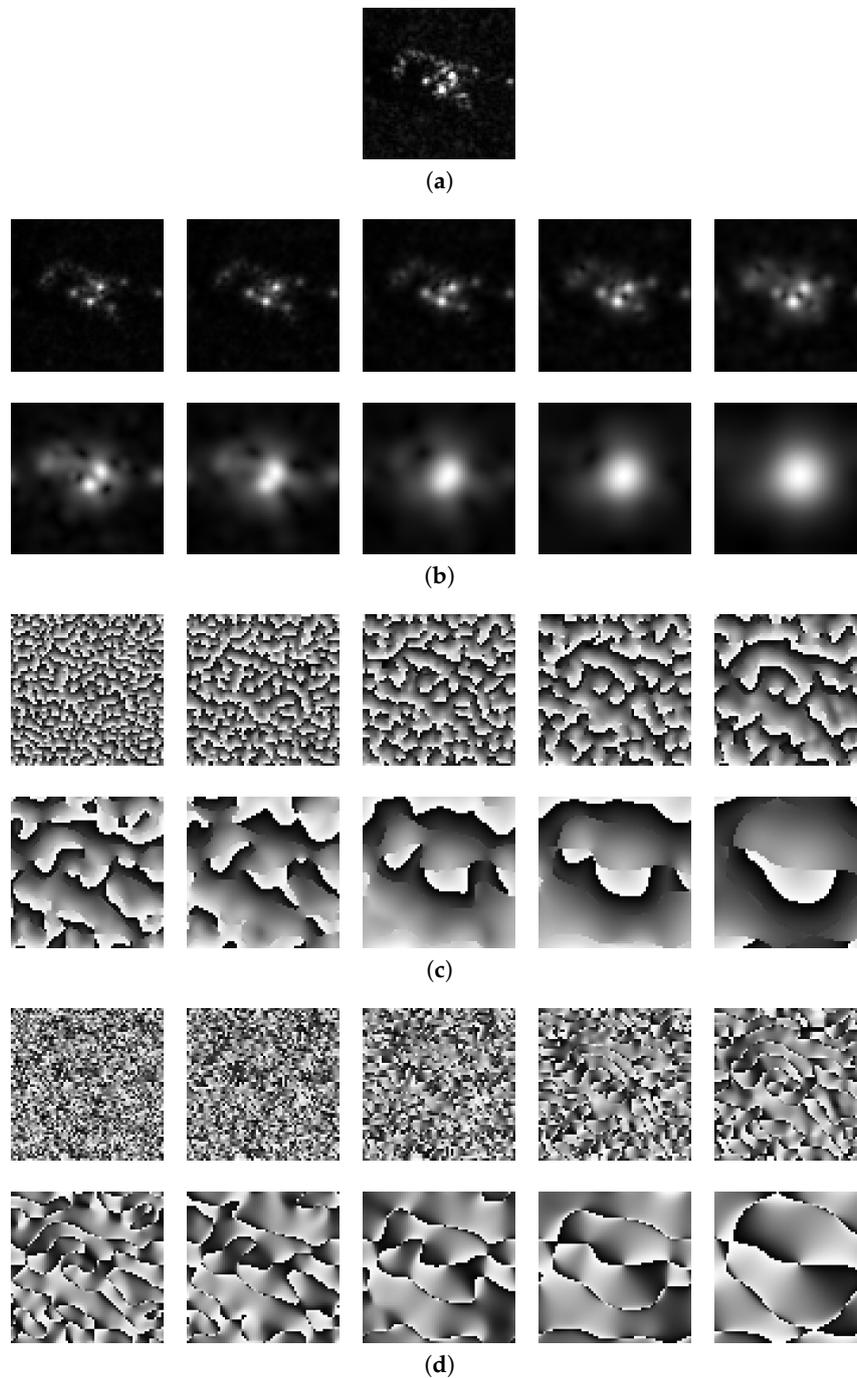


Figure 1. Monogenic signal with scale S ($S = 10$). (a) original SAR image; (b) amplitude scale space; (c) phase scale space; and (d) orientation scale space.

According to Fisher's linear discriminant analysis, the classification accuracy is associated with the within-class and between-class distance. Furthermore, the within-class distance should be minimized

and the between-class distance should be maximized in order to achieve high recognition accuracy. Inspired by previous work, the Fisher score of the k th class in the s th amplitude scale is defined as

$$AC_k^s = \frac{Y_B(\mathbf{A}_k^s)}{Y_W(\mathbf{A}_k^s)}. \quad (23)$$

The matrix $\mathbf{AC} \in \mathbb{R}_K^S$ of the Fisher score in the amplitude scale space can be described as

$$\begin{bmatrix} AC_1^1 & AC_1^2 & \cdots & AC_1^S \\ AC_2^1 & AC_2^2 & \cdots & AC_2^S \\ \vdots & \vdots & \ddots & \vdots \\ AC_K^1 & AC_K^2 & \cdots & AC_K^S \end{bmatrix}. \quad (24)$$

Each row vector of the local Fisher score matrix \mathbf{AC} is normalized to obtain the feature weight in each class and scale space. Obviously, the larger the value of AC_k^s , the more discriminative the amplitude feature of the k th class in s th scale. Therefore, the matrix \mathbf{AC} provides the indicator for choosing the most discriminative features of each class in the scale space. The global Fisher score in s scale can be generated as

$$AC_s = AC_1^s + AC_2^s + \cdots + AC_K^s. \quad (25)$$

\mathbf{AC} can be rewritten as

$$\mathbf{AC} = (AC_1, AC_2, \cdots, AC_S). \quad (26)$$

Clearly, \mathbf{AC} is the vector of global Fisher score, which has provided the information of how representative the features are in the scale space. Similarly, it is easy for us to acquire weight vector of the phase scale space and the orientation scale space, respectively

$$\mathbf{PC} = (PC_1, PC_2, \cdots, PC_S) \quad (27)$$

$$\mathbf{OC} = (OC_1, OC_2, \cdots, OC_S). \quad (28)$$

The row vectors, \mathbf{AC} , \mathbf{PC} , and \mathbf{OC} are sorted in descending order. Supposing V is the number of selected scales in each component scale space, the chosen features of $3V$ scales will be applied in the classification system rather than the whole monogenic scale space. Hence, the chosen scale matrix \mathbf{D} can be generated as

$$\begin{bmatrix} AD_1 & AD_2 & \cdots & AD_V \\ PD_1 & PD_2 & \cdots & PD_V \\ OD_1 & OD_2 & \cdots & OD_V \end{bmatrix} \quad (29)$$

where the elements AD_v , PD_v , and OD_v are the chosen scale numbers of the amplitude, phase and orientation scale space, respectively. The component with the corresponding scale can be described as

$$\begin{bmatrix} \mathbf{AE}_1 & \mathbf{AE}_2 & \cdots & \mathbf{AE}_V \\ \mathbf{PE}_1 & \mathbf{PE}_2 & \cdots & \mathbf{PE}_V \\ \mathbf{OE}_1 & \mathbf{OE}_2 & \cdots & \mathbf{OE}_V \end{bmatrix}. \quad (30)$$

3.3. Classification via Tri-Task Joint Sparse Representation of Monogenic Signal

The data size is still too large for classification after scale selection. Therefore, the independent and identically distributed (IID) Gaussian random project matrix is employed for the components of the monogenic signal (local amplitude, local phase, local orientation) at V scales. This projection is aimed at reducing the dimension and redundancy. After projection, each component is then transformed into vectors.

Finally, the obtained vectors are concatenated to generate a component-specific feature vector, which can be described as

$$\begin{aligned} \chi_A &= [\text{vec}(\mathbf{AE}_1); \text{vec}(\mathbf{AE}_2); \cdots; \text{vec}(\mathbf{AE}_V)] \\ \chi_P &= [\text{vec}(\mathbf{PE}_1); \text{vec}(\mathbf{PE}_2); \cdots; \text{vec}(\mathbf{PE}_V)] \\ \chi_O &= [\text{vec}(\mathbf{OE}_1); \text{vec}(\mathbf{OE}_2); \cdots; \text{vec}(\mathbf{OE}_V)] \end{aligned} \quad (31)$$

where $\text{vec}(\cdot)$ denotes the reshaping operation from a matrix to a vector. The generation of monogenic-specific features based on scale selection is shown in Figure 2.

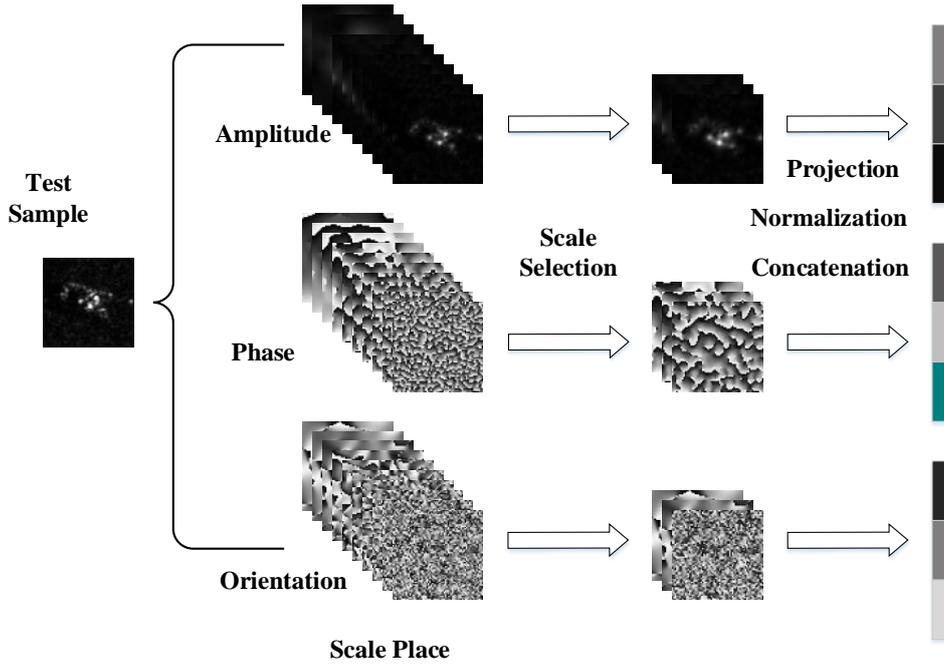


Figure 2. Generation of monogenic component-specific features with scale selection.

After scale selection and feature concatenation, the multi-task joint sparse representation turns out to be a tri-task joint sparse representation. Three overcomplete dictionaries can be built based on the training sample set $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}] \in \mathbb{R}^{m \times n_k}$. Let the monogenic component-specific feature corresponding to the test sample $\mathbf{x}_{k,j}$ be $(\chi_A(\mathbf{x}_{k,j}), \chi_P(\mathbf{x}_{k,j}), \chi_O(\mathbf{x}_{k,j}))$ by Equation (31). The overcomplete dictionary can be formulated as

$$\begin{aligned} \mathbf{X}^1 &= [\chi_A(\mathbf{x}_{1,1}), \chi_A(\mathbf{x}_{1,2}), \cdots, \chi_A(\mathbf{x}_{K,n_K})] \\ \mathbf{X}^2 &= [\chi_P(\mathbf{x}_{1,1}), \chi_P(\mathbf{x}_{1,2}), \cdots, \chi_P(\mathbf{x}_{K,n_K})] \\ \mathbf{X}^3 &= [\chi_O(\mathbf{x}_{1,1}), \chi_O(\mathbf{x}_{1,2}), \cdots, \chi_O(\mathbf{x}_{K,n_K})] \end{aligned} \quad (32)$$

The test sample \mathbf{y} can be also described as

$$\mathbf{y}^1 = \chi_A(\mathbf{y}), \mathbf{y}^2 = \chi_P(\mathbf{y}), \mathbf{y}^3 = \chi_O(\mathbf{y}). \quad (33)$$

Similarly to the multi-task joint sparse representation, the minimum reconstruction error criterion accumulated over all the three task is used to decide the label of the original test sample:

$$\text{class}(\mathbf{y}) = \arg \min_{k \in \{1,2,\dots,K\}} \sum_{p=1}^3 w_p \|\mathbf{y}^p - \mathbf{X}_k^p \hat{\mathbf{a}}_k^p\|. \quad (34)$$

where $\mathbf{W} = (w_1, w_2, w_3)$ denotes the weight vector. Since the three components of monogenic signal show different characteristics, the elements of the weight vector W should not be considered equally as usual. Moreover, the weight vector should be adaptive when the training data set changes. The value of w_p is larger when the corresponding component is more discriminative. The elements of weight vector are the global Fisher scores of each component-specific feature, which can be generated as

$$\begin{aligned} w_1 &= \sum_{v=1}^V AC_v \\ w_2 &= \sum_{v=1}^V PC_v. \\ w_3 &= \sum_{v=1}^V OC_v \end{aligned} \quad (35)$$

The proposed method in this paper, i.e., monogenic scale selection-based weighted tri-task joint sparse representation classification (WTJSRC), is outlined in Algorithm 1.

Algorithm 1 Monogenic scale selection-based WTJSRC for SAR image classification.

Input:

SAR image data \mathbf{R} , original training set $\mathbf{X} \in \mathbb{R}^{m \times n}$ (with n samples) and test set $\mathbf{Y} \in \mathbb{R}^{m \times l}$ (with l samples) from K classes;

The number of total scales S ;

The number of selected scales V ;

Output:

Identity ζ for all test samples;

1: **BEGIN**

2: Acquire the monogenic signal of all the original training samples by (8) and (9), from which the V discriminative scales can be selected by Fisher discrimination criterion;

3: Build the weight vector $W = (w_1, w_2, w_3)$ with global Fisher score of each component;

4: Generate the monogenic component-specific features with scale selection by (31);

5: Build three overcomplete dictionaries $\mathbf{X}^1, \mathbf{X}^2, \mathbf{X}^3$;

6: **for** $j = 1$ to l **do**

7: $\zeta_j = \arg \min_{k \in \{1, 2, \dots, K\}} \sum_{p=1}^3 w_p \left\| \mathbf{y}^j - \mathbf{X}_k^p \hat{\alpha}_k^p \right\|$.

8: **end for**

9: **return** ζ ;

10: **END**

4. Experimental Results

The Moving and Stationary Target Acquisition and Recognition (MSTAR) public database is used to evaluate the performance of the proposed method. SAR images in MSTAR dataset have a resolution of $0.3 \text{ m} \times 0.3 \text{ m}$ with HH polarization. The azimuth angles of SAR imagery are from 0° to 360° and adjacent angle intervals are from 1° to 2° . The SAR images are of 128×128 pixels in size and are cropped in pretreatment process to extract 64×64 patches from the center. After pretreatment, all the SAR images are of 64×64 pixels in size. The IID Gaussian random project matrix is used to reduce the dimension of each component from 64×64 to 12×12 . For the multiresolution monogenic signal representation, 10 is the scale index ($S = 10$) in this paper. To verify the effectiveness of our proposed method, several methods shown in Table 1 are studied.

In the rest of the section, several experiments are carried out to evaluate the performance of our method proposed in this paper.

Table 1. Approaches to be compared in this paper.

Abbreviation	Full Name	Reference
CondGauss	conditional Gaussian model	[44]
SVM	support vector machine	[45]
AdaBoost	feature fusion via boosting on radial basis function net classifiers	[46]
IGT	iterative graph thickening	[47]
SRC	sparse representation classification	[15]
TJSR	joint sparse representation classification of monogenic signal	[43]
WTJSR	weighted tri-task joint sparse representation	-

4.1. Scale Parameter Experiments

First of all, the estimation of the selected scale parameter (V) in a 10-scale space is considered. Since monogenic signal is acquired from the original SAR images, it is essential to determine the optimal value of V . The dataset in Table 2 is used to compare the performance of our method with each V (from 1 to 10). As shown in Table 2, three targets (BMP2, BTR70, T72) are employed. Among them, BMP2 and T72 have several variants with small structural modifications (denoted by series number). Training set is composed with the standards (sn_c9563 for BMP2, sn_c71 for BTR70 and sn_132 for T72) at the depression angle of 17° .

Table 2. Data set 1.

Type	Training Set (17°)	Testing Set (15°)
BMP2	233(Sn_9563)	195(Sn_9563)
		196(Sn_9566)
		196(Sn_c21)
BTR70	233(Sn_c71)	196(Sn_c71)
T72	232(Sn_132)	196(Sn_132)
		195(Sn_812)
		191(Sn_s7)

The determination of V is based on data set 1 by comparing the performance of our method with each V (from 1 to 10). The component-specific features vary with each V (from 1 to 10), which is shown in Figure 3. In order to get rid of the influence of the IID Gaussian matrix, the same Gaussian random matrix is employed to process both the training and testing samples. In addition, 10 Gaussian matrices are used separately, and the final decision is made by the residual value in each class.

The recognition rate of the value of V is shown in Table 3 and the computational time of each V is shown in Figure 4.

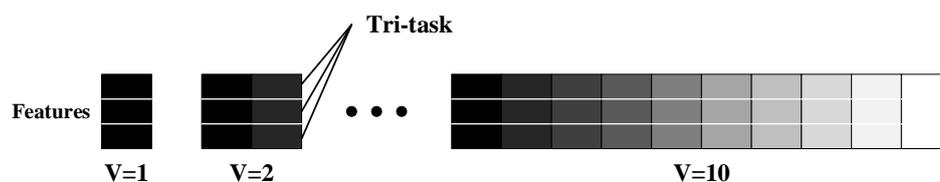
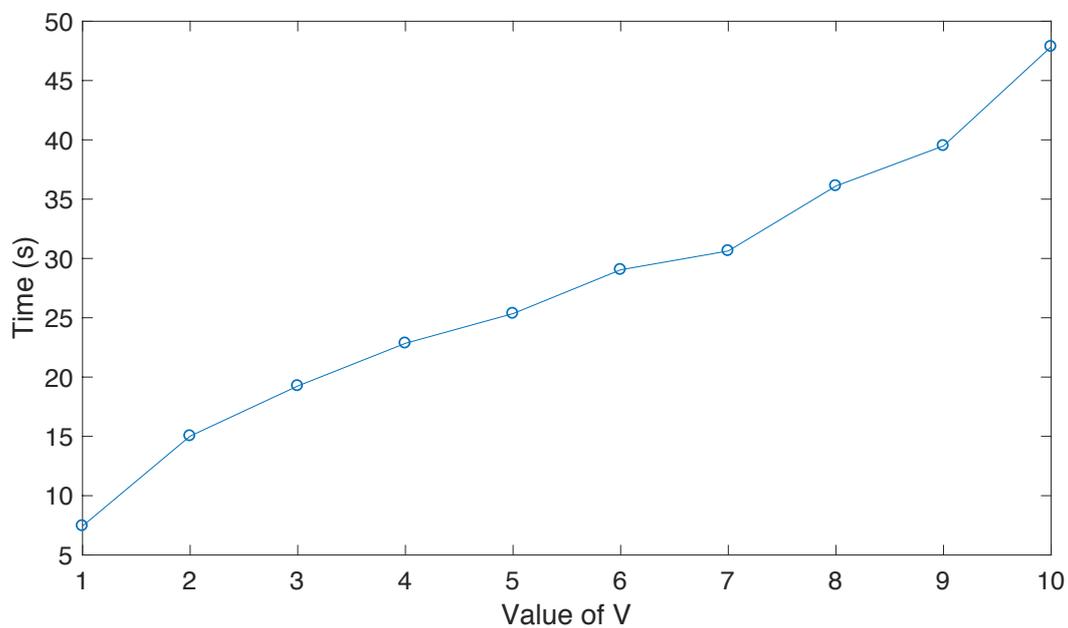


Figure 3. Component-specific features with each selected scale parameter (V). The darker color of components indicates a larger global Fisher score.

Table 3. Recognition rate with each V .

V	BMP2 (%)	BTR70 (%)	T72 (%)	Average (%)
1	88.08	98.47	79.73	86.10
2	91.99	99.49	86.60	90.77
3	94.21	99.49	90.21	93.26
4	95.74	100	91.92	94.72
5	96.59	100	94.50	96.19
6	96.25	100	95.19	96.34
7	96.25	100	94.67	96.11
8	96.25	100	93.70	95.70
9	94.88	100	93.70	95.11
10	95.25	100	93.35	95.12

**Figure 4.** Computational time of each V on data set 2.

4.2. SAR Image Classification under Standard Operating Conditions

We focus on performance evaluation of our method under standard operating conditions (SOCs). The testing data set is a collection of images with all ten classes acquired at the depression angle of 15° . The training data set is a collection of images with all ten class acquired at the depression angle of 17° . Similarly to the data set 1, only the variants of BMP2 and T72, Sn_9563, and Sn_132 (in bold in Table 4), are available for training. The confusion matrices under SOCs are shown in Figure 5. The computation costs of sparse representation-based approaches under SOCs are shown in Table 5.

Table 4. Data set 2.

Type	Training Set (17°)	Testing Set (15°)
BMP2	233(Sn_9563)	195(Sn_9563)
	232(Sn_9566)	196(Sn_9566)
	233(Sn_c21)	196(Sn_c21)
BTR70	233(Sn_c71)	196(Sn_c71)
T72	232(Sn_132)	196(Sn_132)
	231(Sn_812)	195(Sn_812)
	228(Sn_s7)	191(Sn_s7)
BTR60	256	195
2S1	299	274
BRDM2	298	274
D7	299	274
T62	299	273
ZIL131	299	274
ZSU234	299	274

The series numbers for BMP2 and T72 are in parentheses. The bold of BMP2 and T72 are used for training.

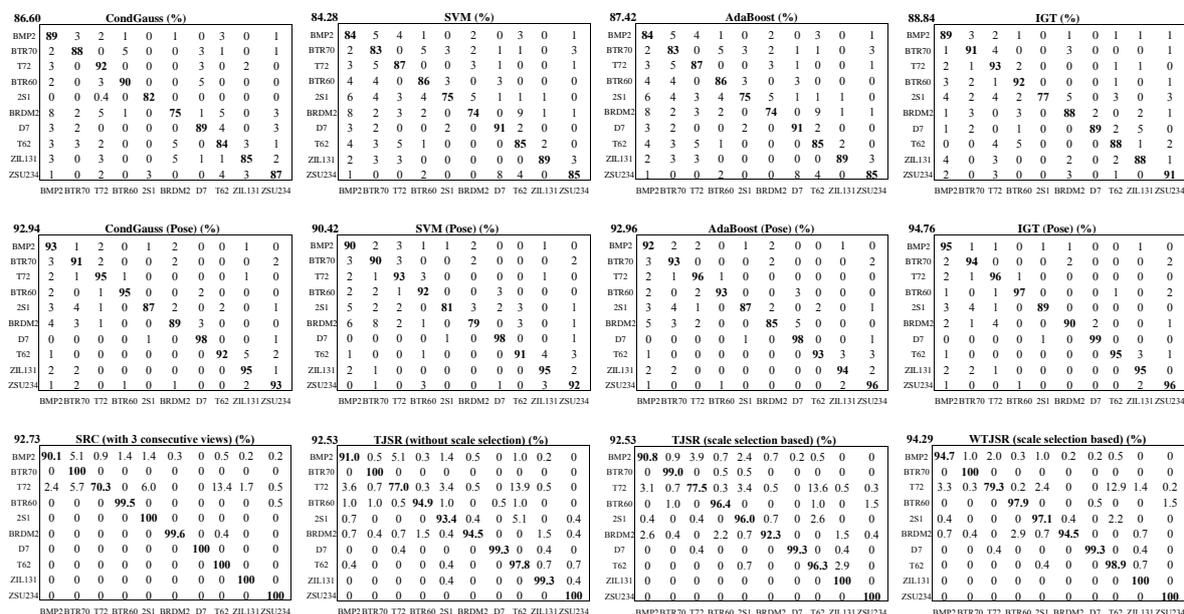


Figure 5. Confusion matrices under SOC. Each row in the subfigures denotes the ground-truth class of the test sample. Each column in the subfigures shows the class predicted by different methods. The element in bold in the upper left corner represents the overall recognition rate. The diagonal elements except for the one in the upper left corner describe the recognition rate of each class. The rest of the elements denote the misclassification rate of each class.

Table 5. Computational cost under SOCs.

Methods	Offline Training (s)	Testing (s)
SRC	5.6	7.4
TJSR	74.1	47.9
WTJSR	130.7	25.4

4.3. SAR Image Classification under Extended Operating Conditions

Two experiments are designed to evaluate the performance of our proposed method under extended operating conditions (EOCs). First all, three types of targets (2S1, BRDM2, and ZSU234) in the MSTAR data set at depression degree of 30° are used as test data set, which is shown in Table 6. The corresponding confusion matrices are shown in Table 7.

Table 6. Test data set under extended operating condition-1 (EOC-1; large depression variation).

Type	Serial Number	Depression Angle	Number of Images
2S1	b01	30°	288
BRDM2	E-71		287
T72	A64		288
ZSU234	d08		288

Table 7. Confusion matrices under EOC-1 (large depression variation).

Type	2S1 (%)	BRDM2 (%)	T72 (%)	ZSU234 (%)
2S1	97.2	1.4	4.2	1.4
BRDM2	4.2	88.9	1.4	5.6
T72	6.9	3.1	86.5	3.5
ZSU234	1.0	0	1.6	97.4
Total (%)	92.6			

In another EOC test scenario, the algorithm is evaluated with respect to target configuration and version variants. Considering several variants of BMP2 and T72, the SAR images of BMP2 (Sn_9563) and T72 (Sn_132) collected at the depression degree of 17° are used as the training data set, while the testing data set includes SAR images of BMP2 (Sn_9566, Sn_c12) and T72 (Sn_812 and Sn_s7) collected at the depression degree of 15° . The data set under EOC-2 is shown in Table 8. The corresponding confusion matrices are shown in Table 9.

The recognition rate of our proposed method is compared with three widely cited methods including k-nearest neighbor (KNN) and three sparse representation-based methods under EOC, as shown in Table 10. TJSR(1) denotes the TJSR method without scale selection, while TJSR(2) represents the scale selection-based TJSR method.

Table 8. Test data set under EOC-2 (version variants).

	BMP2	T72	BTR60	T62	Depression
Train	233(Sn_9563)	232(Sn_132)	256	299	17°
Test	196(Sn_9566) 196(Sn_c21)	195(Sn_812) 191(Sn_s7)	195	273	15°

Table 9. Confusion matrices under EOC-2 (version variants)

Type	BMP2 (%)	T72 (%)	BTR60 (%)	T62 (%)
BMP2	94.1	4.8	0.3	0.8
T72	4.4	74.1	0.8	20.8
BTR60	0	0.5	99.5	0
T62	0	0	0	100
Total (%)	90.4			

Table 10. Comparison of methods under EOCs

Method	KNN	SVM	IGT	TJSR(1)	TJSR(2)	WTJSR
EOC-1 (%)	79	81	85	88	90	93
EOC-2 (%)	86	85	88	87	89	90

5. Discussion

5.1. Scale Parameter Analysis

From Table 3, the recognition rate increases with the value of V up to $V = 7$. The recognition rate drops after the value of V increases to 8. The main reason is that the features at some scales are less discriminative and even have a negative effect on classification. The result also indicates that the features of the low Fisher score reduce the recognition rate of our method, which partly improves the effectiveness of Fisher's discrimination criterion.

The computational time of each V on data set 2 is shown in Figure 4. The computation time was recorded using a personal computer with a 4-core Intel processor of 4.0 GHz and 8 GB RAM. The results are the average of 30 experiments. We can see that the computation time increases with the increase of the value V . The reason is that the dimensions of component-specific features increase with the increase in the value V , as shown in Figure 3. Experimental results show that Fisher's discrimination criterion-based monogenic scale selection reduces the computation load and improves the accuracy of the identification as compared with the method adapting all the scale space into the classification.

Based on the above analysis, we may assess that there is an appropriate tradeoff between the recognition rate and the computation load when $V = 5$.

5.2. Analysis of the Recognition Rate under SOC

We first compare our proposed scale selection-based WTJSR approach with the other three methods based on sparse representation. As we can see from Figure 5, the total recognition rates for our proposed approach are 1.39% and 1.40% better than for TJSR with and without scale selection, respectively. This means our scale selection and weight vector based on Fisher's discrimination criterion contribute positively to the classifier. Although the recognition rates of BTR60, 2S1, BRDM2, D7 and T62 based on our approach are lower than with SRC, the overall recognition rate of our proposed approach is 1.56% higher than for SRC with three consecutive views. Since multi-view images carry more information than a single-view image, the recognition rates with multiple-view images are higher than those with a single-view image in sparse representation-based approaches. Our method based on a single-view image not only closes the gap but also achieves a higher overall recognition rate. The recognition rate of scale selection-based TJSR is a little bit higher than for TJSR without scale selection. As the weight vector can be acquired offline from the training data set, we may assess that the less informative features are abandoned and the computation load decreases to nearly 50% by scale selection (Figure 4). The scale selection-based TJSR has a smaller computation with no recognition rate impairments as compared with TJSR without scale selection.

We also compare scale selection-based WTJSR with four widely cited approaches in ATR literature. The methods, including the conditional Gaussian model (CondGauss), SVM, AdaBoost, and iterative

graph thickening (IGT) with and without a pose estimator as proposed in [46], are studied. As we can see, the pose estimator can significantly improve the performance of the four methods. In other words, the performance of the four methods depends strongly on the accuracy of pose estimation. The performance of our proposed method is much better than the four methods without pose estimation. Even in the case when the CondGauss, SVM, and AdaBoost algorithms use a pose estimator as a preprocessing step, the recognition rate of our method is still 1.35%, 3.87%, and 1.33% better than that of the three methods, respectively. The recognition rate of our method is slightly lower than that of IGT with pose estimation. All the results under SOCs prove the superiority of our proposed method.

In addition, we compare the computational costs of sparse representation-based approaches. The computational cost mainly involves two parts: the offline training and online testing. The computational cost of training is of little importance for SAR ATR because the process of training can be run offline. As we can see, the computational cost of TJSR and WTJSR is higher than for SRC. The reason is that the monogenic signal has large number of components at different scales and causes too much of a burden for computation onboard. The computational cost of WTJSR is lower than for TJSR due to the scale selection model of WTJSR. This result proves the effectiveness of the scale selection model of WTJSR.

5.3. Analysis of the Recognition Rate under EOCs

The experiment under EOCs aims to investigate the practicability of the method proposed in this paper. Under EOC-1, the overall recognition rate of the proposed WTJSR method is 14%, 12%, 8%, 5%, and 3% better than for the competing methods KNN, SVM, IGT, TJSR, scale selection-based TJSR, and WTJSR, respectively. Under EOC-2, the proposed WTJSR method still achieves the highest recognition rate at 90%. There are visible signs of the improvement in the recognition rate under EOCs. The experimental results indicate that our method is more robust with respect to large depression variation and version variants than the competitors.

6. Conclusions

This paper presents a scale selection-based tri-task joint sparse representation method for SAR image recognition. Our proposed approach can effectively process the huge data volume of the monogenic signal and reduce the negative effect of the less informative scale space. In addition, the adaptive weight vector is proposed based on the scale selection model due to the heterogeneity among the three component features of the monogenic signal.

We also illustrate the recognition rate of our method by experiments under SOCs and EOCs. The results of our method are compared with not only state-of-art algorithms such as SVM, AdaBoost, CondGauss, and IGT, but also sparse representation-based algorithms such as SRC and TJSR. The recognition rate of our method is 1.39% and 1.40% better than that of TJSR with and without scale selection, respectively. The scale selection-based TJSR has a smaller computational load but no recognition rate impairments as compared to TJSR without selection. Furthermore, the weight vector based on Fisher's discrimination criteria can effectively improve the recognition rate. The experimental results show the effectiveness of our method. We conclude that it is necessary to evaluate the reliability of the components of the monogenic signal at different scales, and the adaptive weight process is a very important step in classification algorithms based on the monogenic signal due to the heterogeneity among the three component features.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (NSFC) under Grant U1433113, Fundamental Research Funds for the Central Universities under Grant ZYGX2015J020, and in part by Project 2011 for Collaborative Innovation at the Center of Information Sensing and Understanding.

Author Contributions: Zhi Zhou designed the recognition system, performed the experiments and wrote the paper; Ming Wang designed the experiments; Zongjie Cao conceived the study; and Yiming Pi provided guidance during the whole research process.

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

References

1. Tait, P. *Introduction to Radar Target Recognition*; The Institution of Engineering and Technology (IET): Stevenage, UK, 2005; Volume 18.
2. Dang, S.; Cui, Z.; Cao, Z.; Liu, N. SAR Target Recognition via Incremental Nonnegative Matrix Factorization. *Remote Sens.* **2018**, *10*, 374.
3. Felsberg, M.; Sommer, G. The monogenic signal. *IEEE Trans. Signal Process.* **2001**, *49*, 3136–3144.
4. Felsberg, M.; Sommer, G. The Monogenic Scale-Space: A Unifying Approach to Phase-Based Image Processing in Scale-Space. *J. Math. Imag. Vis.* **2004**, *21*, 5–26.
5. Felsberg, M.; Duits, R.; Florack, L. The Monogenic Scale Space on a Rectangular Domain and its Features. *Int. J. Comput. Vis.* **2005**, *64*, 187–201.
6. Wietzke, B.L.; Sommer, G.; Schmaltz, C.; Weickert, J. Differential geometry of monogenic signal representations. In *International Workshop on Robot Vision*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 4931, pp. 454–465.
7. Unser, M.; Sage, D.; Ville, D.V.D. Multiresolution Monogenic Signal Analysis Using the Riesz–Laplace Wavelet Transform. *IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc.* **2009**, *18*, 2402–2418.
8. Demarcq, G.; Mascarilla, L.; Berthier, M.; Courtellemont, P. The Color Monogenic Signal: Application to Color Edge Detection and Color Optical Flow. *J. Math. Imag. Vis.* **2011**, *40*, 269–284.
9. Olhede, S.C.; Ramirez, D.; Schreier, P.J. Detecting Directionality in Random Fields Using the Monogenic Signal. *IEEE Trans. Inf. Theory* **2014**, *60*, 6491–6510.
10. Oh, Y.H.; Ngo, A.C.L.; See, J.; Liong, S.T.; Phan, C.W.; Ling, H.C. Monogenic Riesz wavelet representation for micro-expression recognition. In *Proceedings of the IEEE International Conference on Digital Signal Processing*, Singapore, 21–24 July 2015; pp. 1237–1241.
11. Huang, X.; Zhao, G.; Zheng, W.; Pietikainen, M. Spatiotemporal Local Monogenic Binary Patterns for Facial Expression Recognition. *IEEE Signal Process. Lett.* **2012**, *19*, 243–246.
12. Yuan, X.T.; Liu, X.; Yan, S. Visual classification with multitask joint sparse representation. *IEEE Trans. Image Process.* **2012**, *21*, 4349–4360.
13. Gu, Y.; Wang, C.; You, D.; Zhang, Y.; Wang, S.; Zhang, Y. Representative multiple kernel learning for classification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2852–2865.
14. Wang, Q.; Gu, Y.; Tuia, D. Discriminative multiple kernel learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3912–3927.
15. Wright, J.; Yang, A.Y.; Sastry, S.S.; Sastry, S.S.; Ma, Y. Robust Face Recognition via Sparse Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *31*, 210–227.
16. Huang, J.B.; Yang, M.H. Fast sparse representation with prototypes. In *Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, 13–18 June 2010; pp. 3618–3625.
17. Shin, Y.; Lee, S.; Woo, S.; Lee, H.N. Performance increase by using a EEG sparse representation based classification method. In *Proceedings of the IEEE International Conference on Consumer Electronics*, Las Vegas, NV, USA, 11–14 January 2013; pp. 201–203.
18. Lin, C.; Wang, B.; Zhao, X.; Pang, M. Optimizing kernel PCA using sparse representation-based classifier for mstar SAR image target recognition. *Math. Probl. Eng.* **2013**, *2013*, 847062.
19. Yuan, X.T.; Yan, S. Visual classification with multi-task joint sparse representation. In *Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, USA, 13–18 June 2010; pp. 3493–3500.
20. Zhang, H.; Nasrabadi, N.M.; Zhang, Y.; Huang, T.S. Multi-View Automatic Target Recognition using Joint Sparse Representation. *IEEE Trans. Aerosp. Electron. Syst.* **2012**, *48*, 2481–2497.
21. Dong, G.; Kuang, G.; Zhao, L.; Lu, J.; Lu, M. Joint sparse representation of monogenic components: With application to automatic target recognition in SAR imagery. In *Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Quebec City, QC, Canada, 13–18 July 2014; pp. 549–552.
22. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **1996**, *267*–288.
23. Candes, E.J.; Romberg, J.K.; Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **2006**, *59*, 1207–1223.

24. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202.
25. Samadi, S.; Çetin, M.; Masnadi-Shirazi, M.A. Sparse representation-based synthetic aperture radar imaging. *IET Radar Sonar Navig.* **2011**, *5*, 182–193.
26. Austin, C.D.; Ertin, E.; Moses, R.L. Sparse signal methods for 3-D radar imaging. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 408–423.
27. Qian, Y.; Ye, M. Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 499–515.
28. Yang, F.; Gao, W.; Xu, B.; Yang, J. Multi-frequency polarimetric SAR classification based on Riemannian manifold and simultaneous sparse representation. *Remote Sens.* **2015**, *7*, 8469–8488.
29. Song, S.; Xu, B.; Yang, J. SAR target recognition via supervised discriminative dictionary learning and sparse representation of the SAR-HOG feature. *Remote Sens.* **2016**, *8*, 683.
30. Cao, Z.; Xu, L.; Feng, J. Automatic target recognition with joint sparse representation of heterogeneous multi-view SAR images over a locally adaptive dictionary. *Signal Process.* **2016**, *126*, 27–34.
31. Wright, J.; Ma, Y.; Mairal, J.; Sapiro, G.; Huang, T.S.; Yan, S. Sparse representation for computer vision and pattern recognition. *Proc. IEEE* **2010**, *98*, 1031–1044.
32. Bruckstein, A.M.; Donoho, D.L.; Elad, M. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.* **2009**, *51*, 34–81.
33. Mallat, S.G.; Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.* **1993**, *41*, 3397–3415.
34. Tropp, J.A.; Gilbert, A.C.; Strauss, M.J. Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. *Signal Process.* **2006**, *86*, 572–588.
35. Tropp, J.A.; Gilbert, A.C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **2007**, *53*, 4655–4666.
36. Elad, M.; Figueiredo, M.A.; Ma, Y. On the role of sparse and redundant representations in image processing. *Proc. IEEE* **2010**, *98*, 972–982.
37. Natarajan, B.K. Sparse approximate solutions to linear systems. *SIAM J. Comput.* **1995**, *24*, 227–234.
38. Tropp, J.A. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inf. Theory* **2004**, *50*, 2231–2242.
39. Gao, X.; Wang, N.; Tao, D.; Li, X. Face sketch-photo synthesis and retrieval using sparse representation. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1213–1226.
40. Candes, E.J.; Tao, T. Decoding by linear programming. *IEEE Trans. Inf. Theory* **2005**, *51*, 4203–4215.
41. Du, Q.; Zhang, L.; Zhang, B.; Tong, X.; Du, P.; Chanussot, J. Foreword to the special issue on hyperspectral remote sensing: Theory, methods, and applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 459–465.
42. Yang, M.; Zhang, L.; Shiu, S.C.K.; Zhang, D. Monogenic binary coding: An efficient local feature extraction approach to face recognition. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1738–1751.
43. Dong, G.; Kuang, G.; Wang, N.; Zhao, L.; Lu, J. SAR target recognition via joint sparse representation of monogenic signal. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3316–3328.
44. O’Sullivan, J.A.; Devore, M.D.; Kedia, V.; Miller, M.I. SAR ATR performance using a conditionally Gaussian model. *IEEE Trans. Aerosp. Electron. Syst.* **2001**, *37*, 91–108.
45. Zhao, Q.; Principe, J.C. Support vector machines for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2001**, *37*, 643–654.
46. Sun, Y.; Liu, Z.; Todorovic, S.; Li, J. Adaptive boosting for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 112–125.
47. Srinivas, U.; Monga, V.; Raj, R.G. SAR Automatic Target Recognition Using Discriminative Graphical Models. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 591–606.

